SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF

Simon A Levin

Princeton University, Princeton, NJ, USA

VOLUME 1

ELSEVIER

AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO

Academic Press is an imprint of Elsevier

ACADEMIC PRESS

Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xxxvii

Contributors to Volume 1 xliii

Contributors xlvii

Foreword to the First Edition lxvii

Preface to the First Edition lxix

Preface to the Second Edition lxxi

How to use of the Encyclopedia lxxiii

VOLUME 1

A

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems 1
Charles T Driscoll

Adaptation 15
Michael R Rose and Molly K Burke

Adaptive Radiation 21
Rosemary G Gillespie, George K Roderick, and Francis G Howarth

Aesthetic Factors 37
Gordon H Orians

Africa, Ecosystems of 45
J Michael Lock

African and Asian Savannas 58
Mahesh Sankaran and Jayashree Ratnam

Agricultural Invasions 75
David Pimentel

Agriculture, Industrialized 85
Prabhu Pingali and Melinda Smale

Agriculture, Nutrient Management, and Water Quality 95
Andrew N Sharpley

Agriculture, Sustainable 111
G Philip Robertson and Richard R Harwood

Agriculture, Traditional 119
Miguel A Altieri

Agrobiodiversity 126
Louise E Jackson, L Brussaard, PC de Ruiter, Unai Pascual, Charles Perrings, and K Bawa

Air Pollution 136
Mike Ashmore

Alpine Ecosystems 148
Christian Körner

Amazon Ecosystems 158
Ghillean T Prance

Amphibians, Biodiversity of 169
Ross A Alford, Stephen J Richards, and Keith R McDonald

Antarctic Ecosystems 179
Peter Convey

Aquaculture 189
Max Troell, Nils Kautsky, Malcolm Beveridge, Patrik Henriksson, Jurgenne Primavera, Patrik Rönnbäck, and Carl Folke

Arachnids 202
Jonathan A Coddington and Robert K Colwell

Archaea, Origin of 218
Costantino Vetriani

Arctic Terrestrial Ecosystems 227
Terry V Callaghan, Nadya Matveyeva, Yuri Chernov, Niels M Schmidt, Rob Brooker, and Margareta Johansson

Arthropods (Terrestrial), Amazonian 245
Joachim Adis

Asia, Ecosystems of 254
Elgene O Box and Kazue Fujiwara

Atmospheric Gases 281
Donald J Wuebbles

Australia, Biodiversity of Ecosystems 291
Raymond Louis Specht and Alison Specht

B

Bacterial Biodiversity 307
Erko Stackebrandt

Bacterial Genetics 317
Stanley Maloy

Beetles 326
Henry A Hespenheide

Biodiversity and Cultural Ecosystem Services 332
Franz W Gatzweiler and Konrad Hagedorn

Biodiversity and Ecosystem Services 341
Sandra Quijas and Patricia Balvanera

Biodiversity and Human Health 357
Richard S Ostfeld and Felicia Keesing

Biodiversity and Pest Control Services 373
Stacy M Philpott

Biodiversity as a Commodity 386
Geoffrey Heal

Biodiversity, Definition of 399
Ian R Swingland

Biodiversity, Evolution and 411
Gregg Hartvigsen

Biodiversity-Friendly Farming 418
Joern Fischer, Claire Brittain, and Alexandra-Maria Klein

Biodiversity Generation, Overview 430
Paul H Harvey

Biodiversity, Human Well-Being, and Markets 435
Seth Binder and Stephen Polasky

Biodiversity Informatics 440
Robert A Morris

Biodiversity in Logged and Managed Forests 446
Reinmar Seidler and Kamaljit S Bawa

Biodiversity in Plant Breeding 459
Irwin L Goldman

Biodiversity of Harmful Marine Algae 470
Linda K Medlin and Allan D Cembella

Biodiversity of the Succulent Karoo 485
Philip W Rundel and Richard M Cowling

Biodiversity, Origin of 491
MEJ Newman and GJ Eble

Biodiversity-Rich Countries 497
José Sarukhán and Rodolfo Dirzo

Biodiversity State and Trends in Southeast Asia 509
Lian P Koh, Chris J Kettle, Douglas Sheil, Tien M Lee, Xingli Giam, Luke Gibson, and Gopalasamy R Clements

Biofuels and Biodiversity: The Implications of Energy Sprawl 528
Bruce A Robertson and Patrick J Doran

Biofuels and Biodiversity, Wildlife Habitat Restoration 540
Lisa A Schulte, Todd A Ontl, and GL Drake Larsen

Biogeochemical Cycles 552
Paul G Falkowski

Biogeographical Models 565
Thiago F Rangel and JAF Diniz-Filho

Biogeography, Overview 576
Mark V Lomolino

Bioprospecting 588
Paul Alan Cox and Steven King

Birds, Biodiversity of 600
Jeremy JD Greenwood

Boreal Forest Ecosystems 626
Jennie R McLaren and Roy Turkington

Breeding of Animals 636
David R Notter, Beate Scherf, and Irene Hoffmann

Butterflies 650
Philip J DeVries

C

Captive Breeding and Reintroduction 662
Katherine Ralls and Jonathan D Ballou

Captive Breeding and the Evolutionarily Significant Unit 668
Robin S Waples

Carbon Cycle 674
John Grace

Carnivores 685
Hans Kruuk

Carrying Capacity, Concept of 695
Gregg Hartvigsen

Cattle, Sheep, and Goats, Ecological Role of 702
Andreas Troumbis

CONTENTS OF ALL VOLUMES

VOLUME 1

A

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems 1
Charles T Driscoll

Adaptation 15
Michael R Rose and Molly K Burke

Adaptive Radiation 21
Rosemary G Gillespie, George K Roderick, and Francis G Howarth

Aesthetic Factors 37
Gordon H Orians

Africa, Ecosystems of 45
J Michael Lock

African and Asian Savannas 58
Mahesh Sankaran and Jayashree Ratnam

Agricultural Invasions 75
David Pimentel

Agriculture, Industrialized 85
Prabhu Pingali and Melinda Smale

Agriculture, Nutrient Management, and Water Quality 95
Andrew N Sharpley

Agriculture, Sustainable 111
G Philip Robertson and Richard R Harwood

Agriculture, Traditional 119
Miguel A Altieri

Agrobiodiversity 126
Louise E Jackson, L Brussaard, PC de Ruiter, Unai Pascual, Charles Perrings, and K Bawa

Air Pollution 136
Mike Ashmore

Alpine Ecosystems 148
Christian Körner

Amazon Ecosystems 158
Ghillean T Prance

Amphibians, Biodiversity of 169
Ross A Alford, Stephen J Richards, and Keith R McDonald

Antarctic Ecosystems 179
Peter Convey

Aquaculture 189
Max Troell, Nils Kautsky, Malcolm Beveridge, Patrik Henriksson, Jurgenne Primavera, Patrik Rönnbäck, and Carl Folke

Arachnids 202
Jonathan A Coddington and Robert K Colwell

Archaea, Origin of 218
Costantino Vetriani

Arctic Terrestrial Ecosystems 227
Terry V Callaghan, Nadya Matveyeva, Yuri Chernov, Niels M Schmidt, Rob Brooker, and Margareta Johansson

Arthropods (Terrestrial), Amazonian 245
Joachim Adis

Asia, Ecosystems of 254
Elgene O Box and Kazue Fujiwara

Atmospheric Gases 281
Donald J Wuebbles

Australia, Biodiversity of Ecosystems 291
Raymond Louis Specht and Alison Specht

B

Bacterial Biodiversity 307
Erko Stackebrandt

Bacterial Genetics 317
Stanley Maloy

Beetles 326
Henry A Hespenheide

Biodiversity and Cultural Ecosystem Services 332
Franz W Gatzweiler and Konrad Hagedorn

Biodiversity and Ecosystem Services 341
Sandra Quijas and Patricia Balvanera

Biodiversity and Human Health 357
Richard S Ostfeld and Felicia Keesing

Biodiversity and Pest Control Services 373
Stacy M Philpott

Biodiversity as a Commodity 386
Geoffrey Heal

Biodiversity, Definition of 399
Ian R Swingland

Biodiversity, Evolution and 411
Gregg Hartvigsen

Biodiversity-Friendly Farming 418
Joern Fischer, Claire Brittain, and Alexandra-Maria Klein

Biodiversity Generation, Overview 430
Paul H Harvey

Biodiversity, Human Well-Being, and Markets 435
Seth Binder and Stephen Polasky

Biodiversity Informatics 440
Robert A Morris

Biodiversity in Logged and Managed Forests 446
Reinmar Seidler and Kamaljit S Bawa

Biodiversity in Plant Breeding 459
Irwin L Goldman

Biodiversity of Harmful Marine Algae 470
Linda K Medlin and Allan D Cembella

Biodiversity of the Succulent Karoo 485
Philip W Rundel and Richard M Cowling

Biodiversity, Origin of 491
MEJ Newman and GJ Eble

Biodiversity-Rich Countries 497
José Sarukhán and Rodolfo Dirzo

Biodiversity State and Trends in Southeast Asia 509
Lian P Koh, Chris J Kettle, Douglas Sheil, Tien M Lee, Xingli Giam, Luke Gibson, and Gopalasamy R Clements

Biofuels and Biodiversity: The Implications of Energy Sprawl 528
Bruce A Robertson and Patrick J Doran

Biofuels and Biodiversity, Wildlife Habitat Restoration 540
Lisa A Schulte, Todd A Ontl, and GL Drake Larsen

Biogeochemical Cycles 552
Paul G Falkowski

Biogeographical Models 565
Thiago F Rangel and JAF Diniz-Filho

Biogeography, Overview 576
Mark V Lomolino

Bioprospecting 588
Paul Alan Cox and Steven King

Birds, Biodiversity of 600
Jeremy JD Greenwood

Boreal Forest Ecosystems 626
Jennie R McLaren and Roy Turkington

Breeding of Animals 636
David R Notter, Beate Scherf, and Irene Hoffmann

Butterflies 650
Philip J DeVries

C

Captive Breeding and Reintroduction 662
Katherine Ralls and Jonathan D Ballou

Captive Breeding and the Evolutionarily Significant Unit 668
Robin S Waples

Carbon Cycle 674
John Grace

Carnivores 685
Hans Kruuk

Carrying Capacity, Concept of 695
Gregg Hartvigsen

Cattle, Sheep, and Goats, Ecological Role of 702
Andreas Troumbis

VOLUME 2

Census of Marine Life 1
Darlene Trew Crist and Ronald O'Dor

Central America, Ecosystems of 23
Rodolfo Dirzo and Maria Argenis Bonilla

Cladistics 33
Ian J Kitching, Peter L Forey, and David M Williams

Cladogenesis 46
Christopher J Humphries

Climate Change and Ecology, Synergism of 58
Stephen H Schneider and Terry L Root

Climate Change and Extinctions 73
Joanne L Isaac and Stephen E Williams

Climate Change and Wild Species 79
Terry L Root, Stephen H Schneider, Rachel Warren, Jeff R Price, and Patricia R Mastrandrea

Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species 100
Joshua J Lawler, Carrie A Schloss, and Ailene K Ettinger

Climate, Effects of 115
GF Midgley and FI Woodward

Coastal Beach Ecosystems 128
Anton McLachlan and Omar Defeo

Coevolution 137
Douglas J Futuyma and André Levy

Commons, Concept and Theory of 149
Colin W Clark

Commons, Institutional Diversity of 155
Elinor Ostrom

Comparing Extinction Rates: Past, Present, and Future 167
Vânia Proença and Henrique Miguel Pereira

Competition, Interspecific 177
Bryan Shorrocks

Complementarity 192
Paul Williams and Lisa Manne

Complexity Versus Diversity 201
Shahid Naeem

Computer Systems and Models, Use of 213
Louis J Gross

Conservation and People 221
Michelle Marvier

Conservation and the World's Poorest of the Poor 230
Craig Leisher and M Sanjayan

Conservation Biology, Discipline of 238
Andrew P Dobson, Katarzyna Nowak, and Jon Paul Rodríguez

Conservation Efforts, Contemporary 249
Kristiina A Vogt, Jason J Scullion, Lloyd L Nackley, and Maura Shelton

Conservation Genetics 263
Katie Elizabeth Frith and A Rus Hoelzel

Conservation Movement, Historical 278
Curt Meine

Conserving Biodiversity Outside Protected Areas 289
Belinda Reyers

Convention on Biological Diversity 306
H David Cooper and Kieran Noonan-Mooney

Cooperation, Evolution of 320
Egbert Giles Leigh Jr.

Corals and Coral Reefs 330
Nancy Knowlton and Jeremy Jackson

Countryside Biogeography 347
Chase D Mendenhall, Carrie V Kappel, and Paul R Ehrlich

C_4 Plants 361
Rowan F Sage and Tammy L Sage

Crop Mixtures and the Mechanisms of Overyielding 382
Long Li, Lizhen Zhang, and Fusuo Zhang

Crustaceans 396
Marjorie L Reaka-Kudla

D

Darwin, Charles (Darwinism) 419
Michael T Ghiselin

Defenses, Ecology of 426
Phyllis D Coley and John A Barone

Defining, Measuring, and Partitioning Species Diversity 434
Hanna Tuomisto

Deforestation and Land Clearing 447
Jaboury Ghazoul

Desert Ecosystems 457
Exequiel Ezcurra and Eric Mellink

Desertification 479
James F Reynolds

Development by Design: Using a Revisionist History to Guide a Sustainable Future 495
Joseph M Kiesecker, Kei Sochi, Michael Heiner, Bruce McKenney, Jeffrey Evans, and Holly Copeland

Differentiation 508
Nicholas H Barton

Dinosaurs, Extinction Theories for 516
J David Archibald and Norman MacLeod

Diseases, Conservation and 523
Barbara A Han and Sonia Altizer

Dispersal Biogeography 539
Ran Nathan

Disturbance, Mechanisms of 562
Frank W Davis and Max A Moritz

Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems 568
Robert Keane

Diversity and Conservation of Neotropical Mammals 582
Ricardo A Ojeda

Diversity, Community/Regional Level 595
Howard V Cornell

Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea 608
Matthew Church and Daniela Böttjer

Diversity, Molecular Level 626
Carlos A Machado and Marcos A Antezana

Diversity, Organism Level 638
Daniel R Brooks and Deborah A McLennan

Diversity, Taxonomic versus Functional 648
John C Moore

Domestication of Crop Plants 657
Daniel Zohary

Dormancy and Diapause 665
Nelson G Hairston Jr. and Jennifer A Fox

Dynamic Global Vegetation Models 670
Iain Colin Prentice and Sharon A Cowling

E

Ecogenomic Sensors 690
Christopher A Scholin

Ecological Footprint, Concept of 701
William E Rees

Ecological Genetics 714
Beate Nürnberger

VOLUME 3

Ecology, Concept and Theories in 1
Peter Kareiva and Michelle Marvier

Ecology of Agriculture 9
Alison G Power

Economic Control of Invasive Species 16
David Finnoff, Jason F Shogren, Richard D Horan, Shana M McDermott, and Charles Sims

Economic Growth and the Environment 25
Karl-Göran Mäler

Economics of Agrobiodiversity 31
Unai Pascual, Louise E Jackson, and Adam G Drucker

Economics of the Regulating Services 45
Edward B Barbier

Economic Value of Biodiversity, Measurements of 55
Robert Mendelsohn and Seth Binder

Ecosystem, Concept of 59
Eugene P Odum

Ecosystem Function Measurement, Aquatic and Marine Communities 64
John T Lehman

Ecosystem Function Measurement, Terrestrial Communities 72
Sandra Díaz

Ecosystem Function, Principles of 90
Ross A Virginia and Diana H Wall

Ecosystem Services 96
Heather Tallis, Anne Guerry, and Gretchen C Daily

Ecosystems of South America 105
Luis Anibal Solorzano Cardenas

Ecotoxicology 118
JM Lynch, A Wiseman, and FAAM De Leij

Edible Plants 127
Eduardo H Rapoport and Barbara S Drausal

Education and Biodiversity 133
Shirley M Malcom

Effects of Ecotones on Biodiversity 142
Salit Kark

Elevational Trends in Biodiversity 149
John-Arvid Grytnes and Christy M McCain

El Niño and Biodiversity 155
Emily Meuser, Arne Ø Mooers, and Daniel FR Cleary

Endangered Amphibians 164
Tim Halliday

Endangered Ecosystems 169
Raymond C Nias

Endangered Freshwater Invertebrates 176
David L Strayer

Endangered Mammals 188
Peter Zahler and Tatjana Rosen

Endangered Marine Invertebrates 199
James T Carlton

Endangered Plants 205
Thomas J Stohlgren and Sunil Kumar

Endangered Reptiles 216
Tim Halliday

Endangered Terrestrial Invertebrates 219
Piotr Naskrecki

Endemism 228
RM Cowling

Energy Flow and Ecosystems 237
Alan P Covich

Energy Use, Human 250
Patrick Gonzalez

Environmental Ethics 267
Philip J Cafaro and Richard B Primack

Environmental Impact, Concept and Measurement of 278
Ellen W Chu and James R Karr

Estuarine Ecosystems 297
G Carleton Ray and Jerry McCormick-Ray

Ethical Issues in Biodiversity Protection 309
Philip J Cafaro and Richard B Primack

Ethnobiology and Ethnoecology 319
Gary J Martin

Eukaryotes, Origin of 329
B DeRennaux

Europe, Ecosystems of 333
Ladislav Mucina

Eutrophication and Oligotrophication 347
JoAnn M Burkholder and Patricia M Glibert

Evaluation of Ecosystem Service Policies from Biophysical and Social Perspectives: The Case of China 372
Jianguo Liu, Zhiyun Ouyang, Wu Yang, Weihua Xu, and Shuxin Li

Evolution in Response to Climate Change 385
Julie R Etterson and Ruth G Shaw

Evolution, Theory of 392
Gregory C Mayer and Catherine L Craig

Extinction, Causes of 401
Richard B Primack and Rachel A Morrison

Extinction in the Fossil Record 413
Jeffrey S Levinton

F

Feeding the World and Protecting Biodiversity 426
Paul C West, Reinette Biggs, Bruce A McKenney, and Chad Monfreda

Fires, Ecological Effects of 435
William John Bond

Fish Conservation 443
Carl Safina and Alan Duckworth

Fishes, Biodiversity of 456
Gene S Helfman

Fish Stocks 477
Rainer Froese and Daniel Pauly

Flies, Gnats, and Mosquitoes 488
Brian V Brown

Food Webs 497
Gary R Huxel and Gary A Polis

Forest Canopies, Animal Diversity 511
Terry L Erwin

Forest Canopies, Plant Diversity 516
Nalini M Nadkarni, Mark C Merwin, and Jurgen Nieder

Forest Ecology 528
Timothy J Fahey

Fossil Record 537
Sean R Connolly and Matthew A Kosnik

Framework for Assessment and Monitoring of Biodiversity 545
Francisco Dallmeier, Robert C Szaro, Alfonso Alonso, James Comiskey, and Ann Henderson

Freshwater Ecosystems 560
Robert G Wetzel

Freshwater Ecosystems, Human Impact on 570
Kaj Sand-Jensen

Functional Diversity 587
David Tilman

Functional Diversity Measures 597
Norman WH Mason and David Mouillot

Functional Groups 609
Robert S Steneck

Fungi 624
Thomas J Volk

G

Gene Banks 641
Simon H Linington and Hugh W Pritchard

Genes, Description of 654
Michael F Antolin and William C Black IV

Genetic Diversity 662
Eviatar Nevo

Geologic Time, History of Biodiversity in 678
James W Valentine

Global Declines of Amphibians 691
Kellie Whittaker, Michelle S Koo, David B Wake, and Vance T Vredenburg

Globalization Effects on Common Plant Species 700
Thomas J Stohlgren, Petr Pyšek, John Kartesz, Misako Nishino, Aní bal Pauchard, Marten Winter, Joan Pino, David M Richardson, John Wilson, Brad R Murray, Megan L Phillips, Laura Celesti-Grapow, and Jim Graham

Global Species Richness 707
Kevin J Gaston

Government Legislation and Regulations in the United States 712
Kathryn A Saterson

Grasshoppers and their Relatives 722
Piotr Naskrecki

VOLUME 4

Grassland Ecosystems 1
Osvaldo E Sala, Lucía Vivanco, and Pedro Flombaum

Grazing, Effects of 8
Mark E Hay and Cynthia E Kicklighter

Greenhouse Effect 18
Jennifer A Dunne, Stacy C Jackson, and John Harte

Guilds 33
Richard B Root

H

Habitat and Niche, Concept of 39
Kenneth Petren

Habitat Loss and Fragmentation 50
Heather Bird Jackson and Lenore Fahrig

Habitat Selection 59
Robert A Montgomery and Gary J Roloff

Hemiparasitism 70
David Smith, Todd J Barkman, and Claude W dePamphilis

Herbaceous Vegetation, Species Richness in 79
JP Grime and Jason D Fridley

Herbicides 87
Jodie S Holt

High-Elevation Andean Ecosystems 96
Mary TK Arroyo and Lohengrin A Cavieres

High-Temperature Ecosystems 111
Richard G Wiegert

Historical Awareness of Biodiversity 121
David Takacs

Hotspots 127
Ana SL Rodrigues

Human Impact on Biodiversity, Overview 137
Leslie E Sponsel

Human Impacts on Ecosystems: An Overview 153
Paul R Ehrlich, Claire Kremen, and Anne H Ehrlich

Hunter-Gatherer Societies, Ecological Impact of 162
Kathleen A Galvin and Tyler Beeton

Hybridization in Plants 166
Pamela S Soltis

Hymenoptera 177
Norman F Johnson

I

Identifying Conservation Priorities using a Return on Investment Analysis 185
Edward T Game

Impact of Ecological Restoration on Ecosystem Services 199
Holly P Jones

Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity 209
Thomas Kitzberger

Impact of Past Global Warming on Biodiversity 224
Gregory J Retallack

Implications of Urbanization for Conservation and Biodiversity Protection 231
Robert Ian McDonald

Inbreeding and Outbreeding 245
Katherine Ralls, Richard Frankham, and Jonathan D Ballou

Indicator Species 253
John H Lawton and Kevin J Gaston

Indicator Species: Computation 264
Pierre Legendre

Indigenous Peoples and Biodiversity 269
Víctor M Toledo

Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon 279
Giorgini A Venturieri

Indirect Land Use and Greenhouse Gas Impacts of Biofuels 293
Richard J Plevin and Daniel M Kammen

Insecticide Resistance 298
Ian Denholm and Greg Devine

Insects, Overview 308
Brian V Brown

In Situ, *Ex Situ* Conservation 313
Nigel Maxted

International Organizations and Biodiversity 324
Margaret Goud Collins

Intertidal Ecosystems 332
AJ Underwood and MG Chapman

Introduced Plants, Negative Effects of 345
William G Lee

Introduced Species, Impacts and Distribution of 357
Daniel Simberloff

Invertebrates, Freshwater, Overview 369
Margaret A Palmer and Holly L Menninger

Invertebrates, Marine, Overview 379
PJD Lambshead and PH Schalk

Invertebrates, Terrestrial, Overview 392
Olof Andrén

Island Biogeography 395
Dieter Mueller-Dombois

Isoptera 408
Takuya Abe and Masahiko Higashi

J

Justice, Equity and Biodiversity 434
Kai MA Chan and Terre Satterfield

K

Keystone Species 442
Bruce A Menge, Alison C Iles, and Tess L Freidenburg

L

Lake and Pond Ecosystems 458
Christian Lévêque

Landscape Corridors 467
Ellen I Damschen

Landscape Diversity 476
Debra PC Peters, Sarah C Goslee, Scott L Collins, and James R Gosz

Landscape Ecology and Population Dynamics 488
Scott M Pearson

Landscape Epidemiology 503
Felicia Keesing

Landscape Genetics 508
Andrew Storfer

Landscape Legacies 524
Jennifer M Fraterrigo

Landscape Modeling 531
Robert M Scheller

Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production 539
Wallace E Tyner and Farzad Taheripour

Land-Use Issues 555
Kees Klein Goldewijk

Land-Use Patterns, Historic 569
Oliver Rackham

Large-Scale Biodiversity Experiments 583
Andrew Hector and Michel Loreau

Latent Extinction—The Living Dead 590
Daniel H Janzen

Latitudinal and Elevational Range Shifts under Contemporary Climate Change 599
Jonathan Lenoir and Jens-Christian Svenning

Latitudinal Gradients of Biodiversity 612
Michael R Willig and Steven J Presley

Life Cycle Analysis of Biofuels 627
Jason Hill

Life History, Evolution of 631
Derek Roff

Limits to Biodiversity (Species Packing) 642
LB Slobodkin

Literary Perspectives on Biodiversity 649
William Howarth

Loss of Biodiversity, Overview 656
Robert Barbault

M

Macroscopic Patterns in Marine Plankton 667
Angel López-Urrutia

Mammals, Biodiversity of 681
Joshua R Ginsberg

Mammals, Conservation Efforts for 708
EJ Milner-Gulland, Sarah Durant, Rosie Woodroffe, and Richard Young

Mammals (Late Quaternary), Extinctions of 721
Paul S Martin

VOLUME 5

Mammals (Pre-Quaternary), Extinctions of 1
William A Clemens

Mangrove Ecosystems 10
Peter J Hogarth

Marine and Aquatic Communities, Stress from Eutrophication 23
Jonathan H Sharp

Marine Conservation in a Changing Climate 32
Malin L Pinsky, Kristy J Kroeker, Daniel J Barshis, and Cheryl A Logan

Marine Ecosystems 45
J Frederick Grassle

Marine Ecosystems, Human Impacts on 56
Juan Carlos Castilla

Marine Mammals, Extinctions of 64
Glenn R Van Blaricom, Leah R Gerber, and Robert L Brownell Jr.

Marine Protected Areas: Static Boundaries in a Changing World 94
Elizabeth Mcleod

Marine Sediments 105
Paul VR Snelgrove

Marine Symbioses: Metazoans and Microbes 116
Elisha M Wood-Charlson

Marine Viruses 127
Grieg F Steward, Alexander I Culley, and Elisha M Wood-Charlson

Market Economy and Biodiversity 145
Heather Klemick and R David Simpson

Mass Extinctions, Concept of 156
J John Sepkoski Jr.

Mass Extinctions, Notable Examples of 167
Richard J Twitchett

Measurement and Analysis of Biodiversity 178
Wade Leitner and Will R Turner

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data 195
Nicholas J Gotelli and Anne Chao

Mediterranean-Climate Ecosystems 212
Philip W Rundel and Richard M Cowling

Megaherbivores 223
Norman Owen-Smith

Metapopulations 240
Peter Chesson

Microbial Biodiversity 252
Kristin L Matulich, Devon J Mohamed, and Jennifer BH Martiny

Microbial Biodiversity, Measurement of 259
Kate M Scow, Egbert Schwartz, Mara J Johnson, and Jennifer L Macalady

Microbial Biogeography 271
China A Hanson, M Claire Horner-Devine, Jennifer BH Martiny, and Jed A Fuhrman

Microbial Diversity 280
Paul V Dunlap

Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program 292
Kenneth H Nealson and Shibu Yooseph

Microorganisms (Microbes), Role of 299
Tom Fenchel

Migration 309
Daniel I Rubenstein and Mace A Hack

Modeling Biodiversity Dynamics in Countryside and Native Habitats 321
Henrique M Pereira and Luís Borda-de-Água

Modeling Marine Ecosystem Services 329
Anne D Guerry, Mary H Ruckelshaus, Mark L Plummer, and Dan Holland

Modeling Terrestrial Ecosystem Services 347
Erik Nelson, Nirmal Bhagabati, Driss Ennaanay, Eric Lonsdorf, Derric Pennington, and Manu Sharma

Modern Examples of Extinctions 362
Gábor L Lövei

Molluscs 373
David R Lindberg

Moths 384
David L Wagner

Museums and Institutions, Role of 404
Michael J Novacek and Suzann L Goldberg

Myriapods 421
Alessandro Minelli and Sergei I Golovatch

N

Natural Extinctions (not Human Influenced) 433
Christopher N Johnson

Natural Reserves and Preserves 442
Alexander N Glazer

Nature-Based Coastal Defenses: Can Biodiversity Help? 451
Bregje K van Wesenbeeck, John N Griffin, Mark van Koningsveld, Keryn B Gedan, Michael W McCoy, and Brian R Silliman

Near East Ecosystems, Animal Diversity 459
Joseph Heller

Near East Ecosystems, Plant Diversity 478
Avinoam Danin

Neotropical Seasonally Dry Forests 488
Julio Calvo-Alvarado, Arturo Sánchez-Azofeifa, and Carlos Portillo-Quintero

Nest Parasitism 501
Scott K Robinson, Stephen I Rothstein, and Brian D Peer

Neutral Theory and Beyond 510
James O'Dwyer and Ryan Chisholm

Nitrogen Deposition and Terrestrial Biodiversity 519
Christopher M Clark, Yongfei Bai, William D Bowman, Jane M Cowles, Mark E Fenn, Frank S Gilliam, Gareth K Phoenix, Ilyas Siddique, Carly J Stevens, Harald U Sverdrup, and Heather L Throop

Nitrogen, Nitrogen Cycle 537
Sandy L Tartowski and Robert W Howarth

Nomenclature, Systems of 547
David L Hawksworth

North America, Patterns of Biodiversity in 560
Martin J Lechowicz

Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms 568
Laura F Landweber, Tamara L Horton, and Jonatha M Gott

O

Ocean Ecosystems 581
Richard T Barber

Oceanic Islands: Models of Diversity 590
Rosemary G Gillespie

Oil-Palm Plantations in the Context of Biodiversity Conservation 600
Erik Meijaard and Douglas Sheil

Origin and Diversification of Angiosperms 613
William L Crepet

Origin of Life, Theories of 628
Susanne Brakmann

Origins, Evolution, and Breakdown of Bacterial Symbiosis 637
Joel L Sachs

P

Paleoecology 645
Thompson Webb

Parasitism 656
Klaus Rohde

Parasitoids 674
HCJ Godfray

Pelagic Ecosystems 683
Andrea Belgrano, Sonia D Batten, and Philip C Reid

Pesticides, Uses and Effects of 692
Paul C Jepson

Pharmacology, Biodiversity and 703
Paul Alan Cox

Phenological Shifts in Animals Under Contemporary Climate Change 716
Marcel E Visser

VOLUME 6

Phenotype, a Historical Perspective 1
Robert J Berry

Photosynthesis, Mechanisms of 9
John A Raven

Phylogeny 16
Kevin C Nixon

Plankton, Status and Role of 24
Colin S Reynolds and Judit Padisák

Plant–Animal Interactions 39
Ellen L Simms

Plant Biodiversity, Overview 56
Jeannette Whitton

Plant Communities, Evolution of 65
Brian J Enquist, Bruce H Tiffney, John Haskell, and Karl J Niklas

Plant Conservation 76
Mike Maunder

Plant Invasions 90
David M Richardson and Petr Pyšek

Plant Phenology Changes and Climate Change 103
Tim Sparks and Annette Menzel

Plant–Soil Interactions 109
Joan G Ehrenfeld

Plant Sources of Drugs and Chemicals 129
William H Gerwick

Pollinators, Role of 140
David W Inouye

Pollution, Overview 147
William H Smith

Population Density 157
Brian H McArdle

Population Diversity, Overview 168
Jennifer BH Martiny

Population Dynamics 175
Alan Hastings

Population Genetics 182
Brian Charlesworth

Population Stabilization, Human 199
Samir KC

Population Viability Analysis 210
Hugh P Possingham, Michael A McCarthy, and David B Lindenmayer

Poverty and Biodiversity 220
Madhav Gadgil

Predators, Ecological Role of 229
James Estes, Kevin Crooks, and Robert D Holt

Primate Populations, Conservation of 250
Russell A Mittermeier and Matthew C Richardson

Priority Setting for Biodiversity and Ecosystem Services 261
Derric Pennington, Taylor Ricketts, and Robin Naidoo

Property Rights and Biodiversity 273
Susan S Hanna

Proteorhodopsins: Widespread Microbial Light-Driven Proton Pumps 280
Oded Béjà, Jarone Pinhassi, and John L Spudich

Protozoa 286
Bland J Finlay and Genoveva F Esteban

Psychrophiles 298
Richard Y Morita and Craig L Moyer

R

Rainforest Ecosystems, Animal Diversity 304
Gregory H Adler

Rainforest Ecosystems, Plant Diversity 313
Ian M Turner

Rainforest Loss and Change 322
Karndeo D Singh

Range Ecology, Global Livestock Influences 330
J Boone Kauffman and David A Pyke

Rare Biosphere 345
Carlos Pedrós-Alió

Recombination 353
Abraham B Korol

Reforestation 370
David Lamb

Religious Traditions and Biodiversity 380
Fikret Berkes

Remote Sensing and Image Processing 389
Ruth DeFries

Reptiles, Biodiversity of 400
F Harvey Pough

Residents and Transients in the Fossil Record 414
Albert J van der Meulen and Pablo Peláez-Campomanes

Resource Exploitation, Fisheries 420
John Beddington and Ray Hilborn

Resource Partitioning 429
FA Bazzaz and S Catovsky

Restoration of Animal, Plant, and Microbial Diversity 438
Edith B Allen, Joel S Brown, and Michael F Allen

Restoration of Biodiversity, Overview 453
Joy B Zedler and Roberto Lindig-Cisneros

Riparian Landscapes 461
Robert J Naiman, Henri Décamps, and Michael E McClain

River Ecosystems 469
KE Limburg, DP Swaney, and DL Strayer

Role and Trends of Protected Areas in Conservation 485
Timothy M Boucher, Mark Spalding, and Carmen Revenga

Role of Botanic Gardens 504
Peter Wyse Jackson and Lucy A Sutherland

S

Salmon 522
Michael H Schiewe

Scale, Concept and Effects of 532
David Clayton Schneider

Seagrasses 540
Carlos M Duarte

Slash-and-Burn Agriculture, Effects of 551
Stefan Hauser and Lindsey Norgrove

Social and Cultural Factors 563
Jeffrey A McNeely

Social Behavior 571
Daniel I Rubenstein and Dustin R Rubenstein

Soil Biota, Soil Systems, and Processes 580
David C Coleman

Soil Conservation 590
Dorota L Porazinska and Diana H Wall

South American Natural Ecosystems, Status of 599
Philip M Fearnside

Southern (Austral) Ecosystems 612
Peter H Weston and Robert S Hill

Speciation, Process of 620
Guy L Bush

Speciation, Theories of 629
Hope Hollocher

Species–Area Relationships 640
EF Conor and ED McCoy

Species Assemblages, Macroecology, and Global Change 651
Brian J McGill

Species Coexistence 667
Robert D Holt

Species, Concepts of 679
James Mallet

Species Distribution Modeling 692
Jane Elith and Janet Franklin

Species Diversity, Overview 706
A Ross Kiester

Species Interactions 715
Jessica J Hellmann

VOLUME 7

Stability, Concept of 1
Clarence L Lehman

Stewardship, Concept of 12
Peter Alpert

Storage, Ecology of 23
Caroline M Pond

Stress, Environmental 39
John Cairns Jr.

Subspecies, Semispecies, Superspecies 45
James Mallet

Subterranean Ecosystems 49
David C Culver and Tanja Pipan

Succession, Phenomenon of 63
HH Shugart

Sustainability and Biodiversity 71
Jeannine Cavender-Bares, James Heffernan, Elizabeth King, Stephen Polasky, Patricia Balvanera, and William C Clark

Systematics, Overview 85
Quentin D Wheeler

T

Taxonomy, Methods of 97
Richard Irwin Vane-Wright

Temperate Forests 112
John A Silander Jr.

Terrestrial Ecosystems 128
Stephen Roxburgh and Ian Noble

The Global Environment Facility: Financing the Stewardship of Global Biodiversity 136
Mark Zimsky, Gustavo Fonseca, Jaime Cavelier, Dirk Gaul, Jean-Marc Sinnassamy, Yoko Watanabe, and Ming Yang

The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity 144
Jeffrey J Opperman, Gerald E Galloway, and Stephanie Duvail

Thermophiles, Origin of 161
Anna-Louise Reysenbach and Margaret L Rising

The Value of Biodiversity 167
Partha Dasgupta, Ann P Kinzig, and Charles Perrings

Threatened Birds 180
S Jacob Socolar and David S Wilcove

Threatened Species: Classification Systems and Their Applications 191
Rebecca M Miller

Timber Industry 212
S Kellomäki, J Kouki, P Niemelä, and H Peltola

Tourism, Role of 222
Ralf Buckley

Traditional Conservation Practices 226
Carl Folke and Johan Colding

Translocation as a Conservation Strategy 236
Jessica J Hellmann

Trends in Nature Recreation: Causes and Consequences 241
Patricia Zaradic and Oliver RW Pergams

Trophic Cascades 258
Michael L Pace

Trophic Levels 264
Peter Yodzis

Tropical Forest Ecosystems 269
Gary S Hartshorn

Tropical Forest Regeneration 277
Robin L Chazdon

True Bugs and Their Relatives, Diversity of 287
Carl W Schaefer

U

Ultraviolet Radiation 296
Andrew R Blaustein and Catherine Searle

Urban–Suburban Biodiversity 304
Elizabeth M Cook, Rebecca L Hale, Ann P Kinzig, and J Morgan Grove

V

Valuing Ecosystem Services 314
Marc N Conte

Vents 327
Cindy Lee Van Dover

Vertebrates, Overview 333
Carl Gans and Christopher J Bell

Vicariance Biogeography 342
Christopher John Humphries

W

Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy 352
Rebecca L Goldman-Benner, Silvia Benitez, Alejandro Calvache, Aurelio Ramos, and Fernando Veiga

Wetland Creation and Restoration 367
William J Mitsch

Wetlands Ecosystems 384
Barbara L Bedford, Donald J Leopold, and James P Gibbs

Wildlife Management 403
David Saltz and Gary C White

Worms, Annelida 408
Kristian Fauchald

Worms, Nematoda 420
Scott Lyell Gardner

Worms, Platyhelminthes 437
Janine N Caira and D Timothy J Littlewood

Z

Zoos and Zoological Parks 470
Anna Marie Lyles and Dan Wharton

Glossary 481

Index 547

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 1

Joachim Adis
Max Planck Institute for Limnology, Germany, in cooperation with National Institute for Amazonian Research, Brasil

Ross A Alford
James Cook University, Townsville, QLD, Australia

Miguel A Altieri
University of California, Berkeley, CA, USA

Mike Ashmore
Stockholm Environment Institute, York, UK

Jonathan D Ballou
Smithsonian Conservation Biology Institute, Washington, DC, USA

Patricia Balvanera
Universidad Nacional Autónoma de México, Morelia, México

K Bawa
University of Massachusetts Boston, MA, USA

Kamaljit S Bawa
University of Massachusetts, Boston, MA, USA

Malcolm Beveridge
Worldfish Center, Lusaka, Zambia

Seth Binder
University of Minnesota, St Paul, MN, USA, and Yale University, New Haven, CT, USA

Elgene O Box
University of Georgia, GA, USA, and Yokohama National University, Japan

Claire Brittain
Leuphana University, Lueneburg, Germany

Rob Brooker
James Hutton Institute, Aberdeen, UK

L Brussaard
Wageningen University, Wageningen, The Netherlands

Molly K Burke
University of California, CA, USA

Terry V Callaghan
University of Sheffield, Sheffield, UK, and Royal Swedish Academy of Sciences, Stockholm, Sweden

Allan D Cembella
Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

Yuri Chernov
Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia

Gopalasamy R Clements
James Cook University, Cairns, QLD, Australia

Jonathan A Coddington
Smithsonian Institution, Washington, WA, USA

Robert K Colwell
University of Connecticut, CT, Storrs, USA

Peter Convey
British Antarctic Survey, Cambridge, UK

Richard M Cowling
Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

Paul Alan Cox
Institute for Ethnomedicine, Jackson, WY, USA

Philip J DeVries
Milwaukee Public Museum, Milwaukee, WI, USA

JAF Diniz-Filho
Federal University of Goiás, Goiás, Brazil

Rodolfo Dirzo
Stanford University, Stanford, CA, USA

Patrick J Doran
The Nature Conservancy, Lansing, MI, USA

GL Drake Larsen
Iowa State University, Ames, IA, USA

Charles T Driscoll
Syracuse University, Syracuse, NY, USA

GJ Eble
Santa Fe Institute, Santa Fe, NM, USA

Paul G Falkowski
Rutgers University, New Brunswick, NJ, USA

Joern Fischer
Leuphana University, Lueneburg, Germany

Carl Folke
Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

Kazue Fujiwara
University of Georgia, GA, USA, and Yokohama National University, Japan

Franz W Gatzweiler
University of Bonn, Bonn, Germany

Xingli Giam
Princeton University, Guyot Hall, Princeton, NJ, USA

Luke Gibson
National University of Singapore, Singapore, Republic of Singapore

Rosemary G Gillespie
University of California, Berkeley, CA, USA

Irwin L Goldman
University of Wisconsin-Madison, Madison, WI, USA

John Grace
University of Edinburgh, Edinburgh, UK

Jeremy JD Greenwood
British Trust for Ornithology, Thetford, UK

Konrad Hagedorn
Humboldt-University of Berlin, Berlin, Germany

Gregg Hartvigsen
State University of New York, Geneseo, NY, USA

Paul H Harvey
University of Oxford, Oxford, UK

Richard R Harwood
Michigan State University, East Lansing, MI, USA

Geoffrey Heal
Columbia University, New York, NY, and Stanford University, Stanford, CA, USA

Patrik Henriksson
Leiden University, Leiden, The Netherlands

Henry A Hespenheide
University of California, Los Angeles, CA, USA

Irene Hoffmann
Food and Agriculture Organization of the United Nations, Rome, Italy

Francis G Howarth
Bishop Museum, Honolulu, HI, USA

Louise E Jackson
University of California, Davis, CA, USA

Margareta Johansson
Lund University, Sweden, and Royal Swedish Academy of Sciences, Sweden

Nils Kautsky
Stockholm University, Stockholm, Sweden

Felicia Keesing
Bard College, Annandale-on-Hudson, NY, USA

Chris J Kettle
Universitatstrasse, Zurich, Switzerland

Steven King
Napo Pharmaceuticals, Inc., San Francisco, CA, USA

Alexandra-Maria Klein
Leuphana University, Lueneburg, Germany

Lian P Koh
Universitatstrasse, Zurich, Switzerland

Christian Körner
University of Basel, Basel, Switzerland

Hans Kruuk
Institute of Terrestrial Ecology, Huntingdon, UK

Tien M Lee
Columbia University, New York, NY, USA

Mark V Lomolino
Oklahoma Natural Heritage Inventory, and University of Oklahoma, Norman, OK, USA

Stanley Maloy
San Diego State University, San Diego, CA, USA

Nadya Matveyeva
Komarov Botanical Institute, Russian Academy of Sciences, Sankt-Petersburg, Russia

Keith R McDonald
Queensland Department of Environment and Heritage Protection, Atherton, QLD, Australia

Jennie R McLaren
University of Texas at Arlington, Arlington, TX, USA

Linda K Medlin
UPMC, University of Paris, Paris, France, and Observatoire Océanologique, Banyul-sur-Mer, France

J Michael Lock
Royal Botanic Gardens, London, UK

Robert A Morris
University of Massachusetts Boston, Boston, MA, USA

MEJ Newman
Santa Fe Institute, Santa Fe, NM, USA

David R Notter
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Todd A Ontl
Iowa State University, Ames, IA, USA

Gordon H Orians
University of Washington, Seattle, WA, USA

Richard S Ostfeld
Cary Institute of Ecosystem Studies, Millbrook, NY, USA

U Pascual
University of Cambridge, Cambridge, UK

Charles Perrings
Arizona State University, Tempe, AZ, USA

G Philip Robertson
Michigan State University, East Lansing, MI, USA

Stacy M Philpott
University of Toledo, Toledo, OH, USA

David Pimentel
Cornell University, Ithaca, NY, USA

Prabhu Pingali
International Maize and Wheat Improvement Center, Texcoco, Mexico

Stephen Polasky
University of Minnesota, St Paul, MN, USA

Ghillean T Prance
Royal Botanic Gardens, London, UK

Jurgenne Primavera
Southeast Asian Fisheries Development Centre, Tigbauan, Philippines

Sandra Quijas
Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Michoacán, Mexico

Katherine Ralls
Smithsonian Conservation Biology Institute, Washington, DC, USA

Thiago F Rangel
Federal University of Goiás, Goiás, Brazil

Jayashree Ratnam
National Center for Biological Sciences, Bangalore, India

Stephen J Richards
South Australian Museum, Adelaide, SA, Australia

Bruce A Robertson
Smithsonian Conservation Biology Institute, Washington, DC, USA, and Michigan State University, East Lansing, MI, USA

George K Roderick
University of California, Berkeley, CA, USA

Patrik Rönnbäck
Gotland University, Visby, Sweden

Michael R Rose
University of California, CA, USA

PC de Ruiter
Wageningen University, Wageningen, The Netherlands

Philip W Rundel
University of California, Los Angeles, CA, USA

Mahesh Sankaran
National Center for Biological Sciences, Bangalore, India, and University of Leeds, Leeds, UK

José Sarukhán
Universidad Nacional Autónoma de México, México

Beate Scherf
Food and Agriculture Organization of the United Nations, Rome, Italy

Niels M Schmidt
Aarhus University, Roskilde, Denmark

Lisa A Schulte
Iowa State University, Ames, IA, USA

Reinmar Seidler
Ashoka Trust for Research in Ecology and the Environment, India

Andrew N Sharpley
University of Arkansas, Fayetteville, AR, USA

Douglas Sheil
Mbarara University of Science and Technology, Kabale, Uganda; Center for International Forestry Research, Bogor, Indonesia, and Southern Cross University, Lismore, NSW, Australia

Melinda Smale
International Maize and Wheat Improvement Center, Texcoco, Mexico

Alison Specht
The University of Queensland, Brisbane, QLD, Australia

Raymond Louis Specht
The University of Queensland, Brisbane, QLD, Australia

Erko Stackebrandt
DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany

Ian R Swingland
University of Kent, Canterbury, UK

Max Troell
Beijer Institute, Sweden and Stockholm University, Stockholm, Sweden

Andreas Troumbis
University of the Aegean, Lesvos, Greece

Roy Turkington
University of British Columbia, Vancouver, BC, Canada

Costantino Vetriani
Rutgers University, New Brunswick, NJ, USA

Robin S Waples
NOAA Fisheries, Seattle, WA, USA

Donald J Wuebbles
University of Illinois at Urbana-Champaign, Urbana, IL, USA

CONTRIBUTORS

Takuya Abe[†]
Kyoto University, Kyoto, Japan

Joachim Adis
Max Planck Institute for Limnology, Germany, in cooperation with National Institute for Amazonian Research, Brasil

Gregory H Adler
University of Wisconsin at Oshkosh, Oshkosh, WI, USA

Ross A Alford
James Cook University, Townsville, QLD, Australia

Edith B Allen
University of California, CA, USA

Michael F Allen
Center for Conservation Biology, University of California, CA, USA

Alfonso Alonso
Smithsonian Conservation Biology Institute, Washington, DC, USA

Peter Alpert
University of Massachusetts, Amherst, MA, USA

Miguel A Altieri
University of California, Berkeley, CA, USA

Sonia Altizer
University of Georgia, Athens, GA, USA

Olof Andrén
Swedish University of Agricultural Sciences, Valsätra-Ultuna, Sweden

Marcos A Antezana
Universidade de Lisboa, Lisboa, Portugal

Michael F Antolin
Colorado State University, Fort Collins, CO, USA

[†]Deceased.

Mary TK Arroyo
Universidad de Chile, Santiago, Chile, and Instituto de Ecología y Biodiversidad, Santiago, Chile

Mike Ashmore
Stockholm Environment Institute, York, UK

Yongfei Bai
Chinese Academy of Sciences, Beijing, China

Jonathan D Ballou
Smithsonian Conservation Biology Institute, Washington, DC, USA

Patricia Balvanera
Universidad Nacional Autónoma de México, Morelia, México

Robert Barbault
Museum National d'Histoire Naturelle CNRS, Paris, France

Richard T Barber
Duke University, NC, USA

Edward B Barbier
University of Wyoming, Laramie, WY, USA

Todd J Barkman
Pennsylvania State University, University Park, PA, USA

John A Barone
Mississippi State University, Mississippi State, MS, USA

Daniel J Barshis
Hopkins Marine Station of Stanford University, Pacific Grove, CA, USA

Nicholas H Barton
IST Austria, Klosterneuburg, Austria

Sonia D Batten
Sir Alister Hardy Foundation for Ocean Science, Plymouth, UK

K Bawa
University of Massachusetts Boston, MA, USA

Kamaljit S Bawa
University of Massachusetts, Boston, MA, USA

FA Bazzaz
Harvard University, MA, USA

John Beddington
Imperial College of Science, Technology and Medicine, London, UK

Barbara L Bedford
Cornell University, Ithaca, NY, USA

Tyler Beeton
Colorado State University, Fort Collins, CO, USA

Oded Béjà
Technion – Israel Institute of Technology, Haifa, Israel

Andrea Belgrano
Institute of Marine Research, Lysekil, Sweden

Christopher J Bell
University of Texas at Austin, Austin, TX, USA

Silvia Benitez
The Nature Conservancy, Quito, Ecuador

Fikret Berkes
University of Manitoba, Winnipeg, Canada

Robert J Berry
University College London, London, UK

Malcolm Beveridge
Worldfish Center, Lusaka, Zambia

Nirmal Bhagabati
World Wildlife Fund – US, Washington, DC, USA

Reinette Biggs
Stockholm University, Stockholm, Sweden

Seth Binder
University of Minnesota, St Paul, MN, USA, and Yale University, New Haven, CT, USA

William C Black IV
Colorado State University, Fort Collins, CO, USA

Andrew R Blaustein
Oregon State University, Corvallis, OR, USA

J Boone Kauffman
Oregon State University, OR, USA

Maria Argenis Bonilla
Universidad Nacional de Colombia, Bogotá, Colombia

William John Bond
University of Cape Town, Rondebosch, South Africa

Luís Borda-de-Água
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Daniela Böttjer
University of Hawaii, Honolulu, HI, USA

Timothy M Boucher
The Nature Conservancy, VA, USA

William D Bowman
INSTAAR University of Colorado, Boulder, CO, USA

Elgene O Box
University of Georgia, GA, USA, and Yokohama National University, Japan

Susanne Brakmann
Evotec Biosystems AG, Germany

Claire Brittain
Leuphana University, Lueneburg, Germany

Rob Brooker
James Hutton Institute, Aberdeen, UK

Daniel R Brooks
University of Toronto, Toronto, Canada

Brian V Brown
Natural History Museum of Los Angeles County, Los Angeles, CA, USA

Joel S Brown
University of Illinois, IL, USA

Robert L Brownell Jr.
National Marine Fisheries Service, Pacific Grove, CA, USA

L Brussaard
Wageningen University, Wageningen, The Netherlands

Ralf Buckley
Griffith University, Gold Coast, QLD, Australia

Molly K Burke
University of California, CA, USA

JoAnn M Burkholder
North Carolina State University, Raleigh, NC, USA

Guy L Bush
Michigan State University, MI, USA

Philip J Cafaro
Colorado State University, Fort Collins, CO, USA

Janine N Caira
University of Connecticut, Storrs, CT, USA

John Cairns Jr.
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Terry V Callaghan
University of Sheffield, Sheffield, UK, and Royal Swedish Academy of Sciences, Stockholm, Sweden

Alejandro Calvache
The Nature Conservancy, Cartagena de Indias, Colombia

Julio Calvo-Alvarado
Instituto Tecnológico de Costa Rica, Cartago, Costa Rica

Luis Anibal Solorzano Cardenas
Gordon and Betty Moore Foundation, CA, USA

G Carleton Ray
University of Virginia, Charlottesville, VA, USA

James T Carlton
Williams College and Mystic Seaport, Mystic, CT, USA

Juan Carlos Castilla
Pontificia Universidad Católica de Chile, Santiago, Chile

S Catovsky
Harvard University, MA, USA

Jaime Cavelier
Global Environment Facility, Washington, DC, USA

Jeannine Cavender-Bares
University of Minnesota, St Paul, MN, USA

Lohengrin A Cavieres
Universidad de Concepción, Concepción, Chile, and Instituto de Ecología y Biodiversidad, Santiago, Chile

Laura Celesti-Grapow
Università La Sapienza, Rome, Italy

Allan D Cembella
Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

Kai MA Chan
University of British Columbia, Vancouver, BC, Canada

Anne Chao
National Tsing Hua University, Hsin-Chu, Taiwan

MG Chapman
University of Sydney, Sydney, NSW, Australia

Brian Charlesworth
University of Edinburgh, Edinburgh, UK

Robin L Chazdon
University of Connecticut, Storrs, CT, USA

Yuri Chernov
Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia

Peter Chesson
The University of Arizona, Tucson, AZ, USA

Ryan Chisholm
Smithsonian Tropical Research Institute, Panama, Republic of Panama

Ellen W Chu
University of Washington, Seattle, WA, USA

Matthew Church
University of Hawaii at Manoa, Honolulu, HI, USA

M Claire Horner-Devine
University of Washington, Seattle, WA, USA

Christopher M Clark
US Environmental Protection Agency, Washington, DC, USA

Colin W Clark
University of British Columbia, Vancouver, BC, Canada

William C Clark
Harvard University, Cambridge, MA, USA

Daniel FR Cleary
University of Aveiro, Aveiro, Portugal

William A Clemens
University of California, Berkeley, CA, USA

Gopalasamy R Clements
James Cook University, Cairns, QLD, Australia

Jonathan A Coddington
Smithsonian Institution, Washington, WA, USA

Johan Colding
Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

David C Coleman
University of Georgia, GA, USA

Phyllis D Coley
University of Utah, Salt Lake City, UT, USA

Margaret Goud Collins
International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

Scott L Collins
University of New Mexico, Albuquerque, NM, USA

Robert K Colwell
University of Connecticut, CT, Storrs, USA

James Comiskey
National Park Service, Fredericksburg, VA, USA

Sean R Connolly
James Cook University, Townsville, QLD, Australia

EF Conor
San Francisco State University, CA, USA

Marc N Conte
Stanford University, Stanford, CA, USA

Peter Convey
British Antarctic Survey, Cambridge, UK

Elizabeth M Cook
Arizona State University, Tempe, AZ, USA

Holly Copeland
The Nature Conservancy, Wyoming, WY, USA

Howard V Cornell
University of Delaware, Newark, DE, USA

Alan P Covich
University of Georgia, Athens, GA, USA

Jane M Cowles
University of Minnesota, Saint Paul, MN, USA

RM Cowling
University of Cape Town, Rondebosch, South Africa

Richard M Cowling
Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

Sharon A Cowling
University of Toronto, Ontario, Canada

Paul Alan Cox
Institute for Ethnomedicine, Jackson, WY, USA

Catherine L Craig
Harvard University, Cambridge, MA, USA

William L Crepet
Cornell University, NY, USA

Darlene Trew Crist
University of Rhode Island, North Kingstown, RI, USA

Kevin Crooks
Colorado State University, CO, USA

Alexander I Culley
University of Hawaii at Manoa, Honolulu, HI, USA

David C Culver
American University, Washington, DC, USA

Gretchen C Daily
Stanford University, Stanford, CA, USA

Francisco Dallmeier
Smithsonian Conservation Biology Institute, Washington, DC, USA

Ellen I Damschen
University of Wisconsin-Madison, Madison, WI, USA

Avinoam Danin
The Hebrew University of Jerusalem, Jerusalem, Israel

Partha Dasgupta
University of Cambridge, Cambridge and University of Manchester, Manchester, UK

J David Archibald
San Diego State University, San Diego, CA, USA

H David Cooper
Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada

R David Simpson
National Center for Environmental Economics, Washington, DC, USA

Frank W Davis
University of California, Santa Barbara, CA, USA

FAAM De Leij
University of Surrey, Guildford, UK

Henri Décamps
Université Paul Sabatier, Toulouse, France

Omar Defeo
Faculty of Sciences, UNDECIMAR, Montevideo, Uruguay

Ruth DeFries
Columbia University, NY, USA

Ian Denholm
Rothamsted Research, Harpenden, UK

B DeRennaux
University of Tennessee, Knoxville, TN, USA

Greg Devine
Queensland Health, Cairns, QLD, Australia

Philip J DeVries
Milwaukee Public Museum, Milwaukee, WI, USA

Sandra Díaz
Universidad Nacional de Córdoba – CONICET, Córdoba, Argentina

JAF Diniz-Filho
Federal University of Goiás, Goiás, Brazil

Rodolfo Dirzo
Stanford University, Stanford, CA, USA

Andrew P Dobson
Princeton University, Princeton, NJ, USA

Patrick J Doran
The Nature Conservancy, Lansing, MI, USA

GL Drake Larsen
Iowa State University, Ames, IA, USA

Barbara S Drausal
Universidad Nacional del Comahue, Bariloche, Argentina

Charles T Driscoll
Syracuse University, Syracuse, NY, USA

Adam G Drucker
Bioversity International, Rome, Italy

Carlos M Duarte
IMEDEA, CSIC-Universitat de les Illes Balears, Spain

Alan Duckworth
Blue Ocean Institute, Cold Spring Harbor, NY, USA

Paul V Dunlap
University of Maryland Biotechnology Institute, Baltimore, MD, USA

Jennifer A Dunne
Santa Fe Institute, Santa Fe, NM, USA

Sarah Durant
Zoological Society of London, London, UK, and Wildlife Conservation Society, Bronx, NY, USA

Stephanie Duvail
Institute of Research for Development, Nairobi, Kenya

GJ Eble
Santa Fe Institute, Santa Fe, NM, USA

Joan G Ehrenfeld[†]
Rutgers University, NJ, USA

Anne H Ehrlich
Stanford University, Stanford, CA, USA

Paul R Ehrlich
Stanford University, Stanford, CA, USA

Jane Elith
The University of Melbourne, Parkville, VIC, Australia

Driss Ennaanay
Riverside Technology, Inc., Fort Collins, CO, USA

Brian J Enquist
University of California at Santa Barbara, CA, USA

Terry L Erwin
Smithsonian Institution, Washington, DC, USA

Genoveva F Esteban
Bournemouth University, Poole, UK

James Estes
University of California, CA, USA

Julie R Etterson
University of Minnesota, Duluth, MN, USA

[†]Deceased.

Ailene K Ettinger
University of Washington, Seattle, WA, USA

Jeffrey Evans
The Nature Conservancy, Wyoming, WY, USA

Exequiel Ezcurra
University of California, Riverside, CA, USA

Timothy J Fahey
Cornell University, Ithaca, NY, USA

Lenore Fahrig
Carleton University, Ottawa, ON, Canada

Paul G Falkowski
Rutgers University, New Brunswick, NJ, USA

Kristian Fauchald
National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

Philip M Fearnside
National Institute for Research in the Amazon (INPA), Brazil

Tom Fenchel
University of Copenhagen, Copenhagen, Denmark

Mark E Fenn
Pacific Southwest Research Station, Riverside, CA, USA

Bland J Finlay
Queen Mary University of London, Wareham, UK

David Finnoff
University of Wyoming, Laramie, WY, USA

Joern Fischer
Leuphana University, Lueneburg, Germany

Pedro Flombaum
CONICET and Universidad de Buenos Aires, Buenos Aires, Argentina

Carl Folke
Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

Gustavo Fonseca
Global Environment Facility, Washington, DC, USA

Peter L Forey
The Natural History Museum, London, UK

Jennifer A Fox
Georgetown University, Washington, DC, USA

Richard Frankham
Macquarie University, North Ryde, NSW, Australia

Janet Franklin
Arizona State University, Tempe, AZ, USA

Jennifer M Fraterrigo
University of Illinois, Urbana, IL, USA

J Frederick Grassle
Rutgers University, New Brunswick, NJ, USA

Tess L Freidenburg
California Ocean Science Trust, Oakland, CA, USA

Jason D Fridley
Syracuse University, Syracuse, NY, USA

Katie Elizabeth Frith
Durham University, Durham, UK

Rainer Froese
IFM-GEOMAR, Kiel, Germany

Jed A Fuhrman
University of Southern California, Los Angeles, CA, USA

Kazue Fujiwara
University of Georgia, GA, USA, and Yokohama National University, Japan

Douglas J Futuyma
State University of New York at Stony Brook, New York, NY, USA

Madhav Gadgil
Indian Institute of Science, Bangalore, India

Gerald E Galloway
University of Maryland, College Park, MD, USA

Kathleen A Galvin
Colorado State University, Fort Collins, CO, USA

Edward T Game
The Nature Conservancy, Brisbane, QLD, Australia

Carl Gans
University of Texas at Austin, Austin, TX, USA

Scott Lyell Gardner
University of Nebraska-Lincoln, Lincoln, NE, USA

Kevin J Gaston
University of Exeter, Exeter, UK, and University of Sheffield, Sheffield, UK

Franz W Gatzweiler
University of Bonn, Bonn, Germany

Dirk Gaul
Global Environment Facility, Washington, DC, USA

Keryn B Gedan
Smithsonian Environmental Research Center, Edgewater, MD, USA

Leah R Gerber
Arizona State University, Tempe, AZ, USA

William H Gerwick
University of California, San Diego, CA, USA

Jaboury Ghazoul
Department of Environmental Sciences, ETH Zurich, Switzerland

Michael T Ghiselin
California Academy of Sciences, San Francisco, CA, USA

Xingli Giam
Princeton University, Guyot Hall, Princeton, NJ, USA

James P Gibbs
State University of New York, Syracuse, NY, USA

Luke Gibson
National University of Singapore, Singapore, Republic of Singapore

Rosemary G Gillespie
University of California, Berkeley, CA, USA

Frank S Gilliam
Marshall University, Huntington, WV, USA

Joshua R Ginsberg
Wildlife Conservation Society, Bronx, NY, USA

Alexander N Glazer
University of California, Berkeley, CA, USA

Patricia M Glibert
University of Maryland Center for Environmental Science, Cambridge, MD, USA

HCJ Godfray
University of Oxford, UK

Suzann L Goldberg
American Museum of Natural History, New York, NY, USA

Kees Klein Goldewijk
The Netherlands Utrecht University, Utrecht, The Netherlands

Irwin L Goldman
University of Wisconsin-Madison, Madison, WI, USA

Rebecca L Goldman-Benner
The Nature Conservancy, Durham, NC, USA

Sergei I Golovatch
Russian Academy of Sciences, Moscow, Russia

Patrick Gonzalez
University of California, Berkeley, CA, USA

Sarah C Goslee
USDA-ARS Pasture Systems and Watershed Management Research Unit, University Park, PA, USA

James R Gosz
University of Idaho, Moscow, ID, USA

Nicholas J Gotelli
University of Vermont, Burlington, VT, USA

Jonatha M Gott
Case Western Reserve University, Cleveland, OH, USA

John Grace
University of Edinburgh, Edinburgh, UK

Jim Graham
Colorado State University, Fort Collins, CO, USA

Jeremy JD Greenwood
British Trust for Ornithology, Thetford, UK

John N Griffin
University of Florida, Gainesville, Florida, USA

JP Grime
University of Sheffield, Sheffield, UK

Louis J Gross
University of Tennessee, Knoxville, TN, USA

John-Arvid Grytnes
University of Bergen, Bergen, Norway

Anne Guerry
Stanford University, Stanford, CA, USA

Anne D Guerry
Stanford University, Stanford, CA, USA

Mace A Hack
Nature Conservancy, NE, USA

Konrad Hagedorn
Humboldt-University of Berlin, Berlin, Germany

Nelson G Hairston Jr.
Cornell University, Ithaca, NY, USA

Rebecca L Hale
Arizona State University, Tempe, AZ, USA

Tim Halliday
Open University, Milton Keynes, UK

Barbara A Han
University of Georgia, Athens, GA, USA

Susan S Hanna
State University Corvallis, OR, USA

China A Hanson
University of California, Irvine, CA, USA

John Harte
University of California, Berkeley, CA, USA

Gary S Hartshorn
World Forestry Center, Portland, OR, USA

Gregg Hartvigsen
State University of New York, Geneseo, NY, USA

F Harvey Pough
School of Life Sciences, Rochester Institute of Technology, NY, USA

Paul H Harvey
University of Oxford, Oxford, UK

Richard R Harwood
Michigan State University, East Lansing, MI, USA

John Haskell
University of New Mexico, NM, USA

Alan Hastings
University of California, Davis, CA, USA

Stefan Hauser
IITA Democratic Republic of Congo, Kinshasa, DR Congo

David L Hawksworth
IUBS/IUMS International Committee on Bionomenclature, Ciudad Universitaria, Spain, and Natural History Museum, London, UK

Mark E Hay
Georgia Institute of Technology, Atlanta, GA, USA

Geoffrey Heal
Columbia University, New York, NY, and Stanford University, Stanford, CA, USA

Andrew Hector
University of Zürich, Zurich, Switzerland

James Heffernan
Florida International University, Miami, FL, USA, and Duke University, Durham, NC, USA

Michael Heiner
The Nature Conservancy, Colorado, CO, USA

Gene S Helfman
University of Georgia, Athens, GA, USA

Joseph Heller
The Hebrew University of Jerusalem, Jerusalem, Israel

Jessica J Hellmann
University of Notre Dame, Notre Dame, IN, USA

Ann Henderson
National Geographic Channel, Washington, DC, USA

Patrik Henriksson
Leiden University, Leiden, The Netherlands

Henry A Hespenheide
University of California, Los Angeles, CA, USA

Masahiko Higashi†
Kyoto University, Kyoto, Japan

Ray Hilborn
University of Washington, WA, USA

Jason Hill
University of Minnesota, St Paul, MN, USA

Robert S Hill
University of Adelaide, SA, Australia

Irene Hoffmann
Food and Agriculture Organization of the United Nations, Rome, Italy

Peter J Hogarth
University of York, York, UK

Dan Holland
National Oceanic and Atmospheric Administration, Seattle, WA, USA

Hope Hollocher
University of Notre Dame, Notre Dame, IN, USA

Jodie S Holt
University of California, Riverside, CA, USA

Robert D Holt
University of Florida, Gainesville, FL, USA

Richard D Horan
Michigan State University, East Lansing, MI, USA

Tamara L Horton
Princeton University, Princeton, NJ, USA

Francis G Howarth
Bishop Museum, Honolulu, HI, USA

Robert W Howarth
Cornell University, Ithaca, NY, USA

William Howarth
Princeton University, Princeton, NJ, USA

Christopher J Humphries
The Natural History Museum, London, UK

Christopher John Humphries
The Natural History Museum, London, UK

Gary R Huxel
University of Arkansas, Fayetteville, AR, USA

Alison C Iles
Oregon State University, Corvallis, OR, USA

David W Inouye
University of Maryland, MA, USA

†Deceased.

Joanne L Isaac
James Cook University, Townsville, QLD, Australia

Heather Bird Jackson
Carleton University, Ottawa, ON, Canada

Jeremy Jackson
Smithsonian Institution, Washington, DC, USA

Louise E Jackson
University of California, Davis, CA, USA

Stacy C Jackson
University of California, Berkeley, CA, USA

Peter Wyse Jackson
Missouri Botanical Garden, St Louis, Missouri, USA

S Jacob Socolar
Princeton University, Princeton, NJ, USA

Daniel H Janzen
University of Pennsylvania, Philadelphia, PA, USA

Paul C Jepson
Oregon State University, OR, USA

Margareta Johansson
Lund University, Sweden, and Royal Swedish Academy of Sciences, Sweden

Christopher N Johnson
James Cook University, Townsville, QLD, Australia

Mara J Johnson
University of California, Davis, CA, USA

Norman F Johnson
The Ohio State University, Columbus, OH, USA

Holly P Jones
University of California, Santa Cruz, CA, USA

Daniel M Kammen
University of California, Berkeley, CA, USA

Carrie V Kappel
University of California Santa Barbara, Santa Barbara, CA, USA

Peter Kareiva
National Marine Fisheries Service, Seattle, WA, USA

Salit Kark
The Hebrew University of Jerusalem, Jerusalem, Israel, and University of Queensland, Brisbane, QLD, Australia

James R Karr
University of Washington, Seattle, WA, USA

John Kartesz
Biota of North America Program, Chapel Hill, NC, USA

Nils Kautsky
Stockholm University, Stockholm, Sweden

Samir KC
International Institute for Applied Systems Analysis, Laxenburg, Austria

Robert Keane
Missoula Fire Sciences Laboratory, Missoula, MT, USA

Felicia Keesing
Bard College, Annandale-on-Hudson, NY, USA

S Kellomäki
University of Eastern Finland, Finland

Chris J Kettle
Universitatstrasse, Zurich, Switzerland

Cynthia E Kicklighter
Goucher College, Baltimore, MD, USA

Joseph M Kiesecker
The Nature Conservancy, Colorado, CO, USA

Elizabeth King
Princeton University, Princeton, NJ, USA, and University of Georgia, Athens, GA, USA

Steven King
Napo Pharmaceuticals, Inc., San Francisco, CA, USA

Ann P Kinzig
Arizona State University, Tempe, AZ, USA

Ian J Kitching
The Natural History Museum, London, UK

Thomas Kitzberger
INIBIOMA-CONICET and Universidad Nacional del Comahue, Bariloche, Argentina

Alexandra-Maria Klein
Leuphana University, Lueneburg, Germany

Heather Klemick
National Center for Environmental Economics, Washington, DC, USA

Nancy Knowlton
Smithsonian Institution, Washington, DC, USA

Lian P Koh
Universitatstrasse, Zurich, Switzerland

Michelle S Koo
University of California, Berkeley, CA, USA

Mark van Koningsveld
Van Oord Dredging and Marine Contractors, Rotterdam, The Netherlands, and Delft University of Technology, Delft, The Netherlands

Abraham B Korol
University of Haifa, Haifa, Israel

Christian Körner
University of Basel, Basel, Switzerland

Matthew A Kosnik
Macquarie University, North Ryde, NSW, Australia

J Kouki
University of Eastern Finland, Finland

Claire Kremen
University of California, Berkeley, CA, USA

Kristy J Kroeker
Hopkins Marine Station of Stanford University, Pacific Grove, CA, USA

Hans Kruuk
Institute of Terrestrial Ecology, Huntingdon, UK

Sunil Kumar
Colorado State University, Fort Collins, CO, USA

David Lamb
University of Queensland, Brisbane, Australia

PJD Lambshead
The Natural History Museum, London, UK

Laura F Landweber
Princeton University, Princeton, NJ, USA

Joshua J Lawler
University of Washington, Seattle, WA, USA

John H Lawton
Imperial College, London, UK

Martin J Lechowicz
McGill University, Montreal, QC, Canada

Tien M Lee
Columbia University, New York, NY, USA

William G Lee
Landcare Research, Hamilton, New Zealand

Pierre Legendre
Université de Montréal, Montréal, QC, Canada

Clarence L Lehman
University of Minnesota, St Paul, MN, USA

John T Lehman
University of Michigan, Ann Arbor, MI, USA

Egbert Giles Leigh Jr.
Smithsonian Tropical Research Institute, Panama, Republic of Panama

Craig Leisher
The Nature Conservancy, Millburn, NJ, USA

Wade Leitner
The University of Arizona, Tucson, AZ, USA

Jonathan Lenoir
Aarhus University, Aarhus, Denmark

Donald J Leopold
State University of New York, Syracuse, NY, USA

Christian Lévêque
Institut de Recherche pour le Développement (IRD), Paris, France

Jeffrey S Levinton
State University of New York, Stony Brook, NY, USA

André Levy
State University of New York at Stony Brook, New York, NY, USA

Long Li
China Agricultural University, Beijing, China, and Shihezi University, Shihezi City, China

Shuxin Li
Michigan State University, East Lansing, MI, USA

KE Limburg
State University of New York, NY, USA

David R Lindberg
University of California, Berkeley, CA, USA

David B Lindenmayer
The Australian National University, Canberra, Australia

Roberto Lindig-Cisneros
Universidad National Autónomo de México, México

Simon H Linington
Royal Botanic Gardens, London, UK

D Timothy J Littlewood
The Natural History Museum, London, UK

Jianguo Liu
Michigan State University, East Lansing, MI, USA

Cheryl A Logan
California State University Monterey Bay, Seaside, CA, USA

Mark V Lomolino
Oklahoma Natural Heritage Inventory, and University of Oklahoma, Norman, OK, USA

Eric Lonsdorf
Chicago Botanic Garden, Glencoe, IL, USA

Angel López-Urrutia
Instituto Español de Oceanografía, Gijón, Spain

Michel Loreau
McGill University, Montreal, QC, Canada

Gábor L Lövei
Aarhus University, Slagelse, Denmark

Anna Marie Lyles
Bio-Gist Ventures, Princeton, NJ, USA

JM Lynch
University of Surrey, Guildford, UK

Jennifer L Macalady
University of California, Davis, CA, USA

Carlos A Machado
University of Maryland, College Park, MD, USA

Norman MacLeod
Natural History Museum, London, UK

Shirley M Malcom
American Association for the Advancement of Science (AAAS), Washington, DC, USA

James Mallet
University College London, London, UK

Karl-Göran Mäler
Beijer International Institute of Ecological Economics, Stockholm, Sweden

Stanley Maloy
San Diego State University, San Diego, CA, USA

Lisa Manne
University of Toronto at Scarborough, Toronto, ON, Canada

Gary J Martin
The Global Diversity Foundation, Marrakech, Morocco

Paul S Martin
University of Arizona, Tucson, AZ, USA

Jennifer BH Martiny
University of California, Irvine, CA, USA

Michelle Marvier
Santa Clara University, Santa Clara, CA, USA

Norman WH Mason
Landcare Research, Hamilton, New Zealand

Patricia R Mastrandrea
Stanford University, Stanford, CA, USA

Kristin L Matulich
University of California, Irvine, CA, USA

Nadya Matveyeva
Komarov Botanical Institute, Russian Academy of Sciences, Sankt-Petersburg, Russia

Mike Maunder
Florida International University, FL, USA

Nigel Maxted
University of Birmingham, Birmingham, UK

Gregory C Mayer
University of Wisconsin-Parkside, Kenosha, WI, USA

Brian H McArdle
University of Auckland, Auckland, New Zealand

Christy M McCain
University of Colorado at Boulder, Boulder, CO, USA

Michael A McCarthy
University of Melbourne, Melbourne, Australia

Michael E McClain
UNESCO-Institute for Water Education (IHE), The Netherlands

Jerry McCormick-Ray
University of Virginia, Charlottesville, VA, USA

ED McCoy
University of South Florida, FL, USA

Michael W McCoy
University of Florida, Gainesville, FL, USA

Shana M McDermott
Dartmouth College, Hanover, NH, USA

Keith R McDonald
Queensland Department of Environment and Heritage Protection, Atherton, QLD, Australia

Robert Ian McDonald
The Nature Conservancy, Arlington, VA, USA

Brian J McGill
University of Maine, ME, USA

Bruce McKenney
The Nature Conservancy, Virginia, VA, USA

Bruce A McKenney
The Nature Conservancy, Charlottesville, VA, USA

Anton McLachlan
The University of Sydney, Sydney, NSW, Australia

Jennie R McLaren
University of Texas at Arlington, Arlington, TX, USA

Deborah A McLennan
University of Toronto, Toronto, Canada

Elizabeth Mcleod
The Nature Conservancy, Austin, TX, USA

Jeffrey A McNeely
The International Union for Conservation of Nature (IUCN), Gland, Switzerland

Linda K Medlin
UPMC, University of Paris, Paris, France, and Observatoire Océanologique, Banyul-sur-Mer, France

Erik Meijaard
People and Nature Consulting International, Jakarta, Indonesia; University of Queensland, Brisbane, QLD, Australia, and Center for International Forestry Research, Bogor, Indonesia

Curt Meine
Aldo Leopold Foundation, Baraboo, WI, USA

Eric Mellink
CICESE, Carretera Ensenada-Tijuana, Ensenada, BC, México

Robert Mendelsohn
Yale University, New Haven, CT, USA

Chase D Mendenhall
Stanford University, Stanford, CA, USA

Bruce A Menge
Oregon State University, Corvallis, OR, USA

Holly L Menninger
Cornell University, Ithaca, NY, USA

Annette Menzel
Technische Universität München, Germany

Mark C Merwin
The Evergreen State College, Olympia, WA, USA

Albert J van der Meulen
University of Utrecht, Netherlands

Emily Meuser
Simon Fraser University, Burnaby, BC, Canada

J Michael Lock
Royal Botanic Gardens, London, UK

GF Midgley
South African National Biodiversity Institute, Cape Town, South Africa

Rebecca M Miller
IUCN Red List Unit, Cambridge, UK

EJ Milner-Gulland
Imperial College London, London, UK

Alessandro Minelli
University of Padova, Padova, Italy

William J Mitsch
Florida Gulf Coast University, Naples, FL, USA

Russell A Mittermeier
Conservation International, VA, USA

Devon J Mohamed
Irvine Valley College, Irvine, CA, USA

Chad Monfreda
Arizona State University, Tempe, AZ, USA

Robert A Montgomery
Michigan State University, East Lansing, MI, USA

Arne Ø Mooers
Simon Fraser University, Burnaby, BC, Canada

John C Moore
Colorado State University, Fort Collins, CO, USA

J Morgan Grove
USDA Forest Service, Baltimore, MD, USA

Richard Y Morita
Oregon State University, OR, USA

Max A Moritz
University of California, Santa Barbara, CA, USA

Robert A Morris
University of Massachusetts Boston, Boston, MA, USA

Rachel A Morrison
Institution of Oceanography, La Jolla, CA, USA

David Mouillot
James Cook University, Townsville, Queensland, Australia

Craig L Moyer
Western Washington University, WA, USA

Ladislav Mucina
Curtin University, Perth, Australia

Dieter Mueller-Dombois
University of Hawai'i, Honolulu, HI, USA

Brad R Murray
University of Technology Sydney, Sydney, NSW, Australia

Lloyd L Nackley
University of Washington, Seattle, WA, USA

Nalini M Nadkarni
The Evergreen State College, Olympia, WA, USA

Shahid Naeem
Columbia University in the City of New York, New York, NY, USA

Robin Naidoo
World Wildlife Fund – US, Washington, DC, USA

Robert J Naiman
University of Washington, WA, USA

Piotr Naskrecki
Harvard University, Cambridge, MA, USA

Ran Nathan
The Hebrew University of Jerusalem, Jerusalem, Israel

Kenneth H Nealson
University of Southern California, Los Angeles, CA, USA

Erik Nelson
Bowdoin College, Brunswick, ME, USA

Eviatar Nevo
University of Haifa, Haifa, Israel

MEJ Newman
Santa Fe Institute, Santa Fe, NM, USA

Raymond C Nias
TierraMar Consulting, Sutherland, NSW, Australia

Jurgen Nieder
Botanische Institüt der Universität Bonn, Bonn, Germany

P Niemelä
University of Eastern Finland, Finland

Karl J Niklas
Cornell University, NY, USA

Misako Nishino
Biota of North America Program, Chapel Hill, NC, USA

Kevin C Nixon
Cornell University, NY, USA

Ian Noble
Australian National University, Canberra, ACT, Australia

Kieran Noonan-Mooney
Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada

Lindsey Norgrove
CABI, Delémont, Switzerland

David R Notter
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Michael J Novacek
American Museum of Natural History, New York, NY, USA

Katarzyna Nowak
Princeton University, Princeton, NJ, USA

Beate Nürnberger
The University of Edinburgh, Edinburgh, UK

Ronald O'Dor
Dalhousie University, Halifax, Canada, and Consortium for Ocean Leadership, Washington, DC, USA

James O'Dwyer
Santa Fe Institute, Santa Fe, NM, USA

Eugene P Odum
University of Georgia, Athens, GA, USA

Ricardo A Ojeda
Institute for Aridland Research (IADIZA-CONICET), Mendoza, Argentina

Todd A Ontl
Iowa State University, Ames, IA, USA

Jeffrey J Opperman
The Nature Conservancy, Chagrin Falls, OH, USA

Gordon H Orians
University of Washington, Seattle, WA, USA

Richard S Ostfeld
Cary Institute of Ecosystem Studies, Millbrook, NY, USA

Elinor Ostrom
Indiana University, Bloomington, IN, USA

Zhiyun Ouyang
Chinese Academy of Sciences, Beijing, China

Norman Owen-Smith
University of the Witwatersrand, Wits, South Africa

Michael L Pace
University of Virginia, Charlottesville, VA, USA

Judit Padisák
University of Pannonia, Hungary

Margaret A Palmer
University of Maryland, College Park, MD, USA

Claude W dePamphilis
Pennsylvania State University, University Park, PA, USA

U Pascual
University of Cambridge, Cambridge, UK

Unai Pascual
University of Cambridge, Cambridge, UK; Basque Centre for Climate Change (BC3), Bilbao, Spain, and IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Aníbal Pauchard
Universidad de Concepción and Institute for Ecology and Biodiversity (IEB), Concepción, Chile

Daniel Pauly
University of British Columbia, Vancouver, BC, Canada

Scott M Pearson
Mars Hill College, Mars Hill, NC, USA

Carlos Pedrós-Alió
Institut de Ciències del Mar, CSIC, Passeig Marítim de la Barceloneta, Barcelona, Spain

Brian D Peer
Western Illinois University, Macomb, IL, USA

Pablo Peláez-Campomanes
National Museum of Natural History, CSIC, Spain

H Peltola
University of Eastern Finland, Finland

Derric Pennington
University of Minnesota, St Paul, MN, USA, and World Wildlife Fund – US, Washington, DC, USA

Henrique M Pereira
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Henrique Miguel Pereira
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Oliver RW Pergams
Olive-Harvey College, Chicago, IL, USA

Charles Perrings
Arizona State University, Tempe, AZ, USA

Debra PC Peters
United States Department of Agriculture, Agricultural Research Service, Las Cruces, NM, USA

Kenneth Petren
University of Cincinnati, Cincinnati, OH, USA

G Philip Robertson
Michigan State University, East Lansing, MI, USA

Megan L Phillips
University of Technology Sydney, Sydney, NSW, Australia

Stacy M Philpott
University of Toledo, Toledo, OH, USA

Gareth K Phoenix
University of Sheffield, Sheffield, UK

David Pimentel
Cornell University, Ithaca, NY, USA

Prabhu Pingali
International Maize and Wheat Improvement Center, Texcoco, Mexico

Jarone Pinhassi
Linnaeus University, Kalmar, Sweden

Joan Pino
University of Barcelona, Barcelona, Spain

Malin L Pinsky
Princeton University, Princeton, NJ, USA

Tanja Pipan
Karst Research Institute at ZRC SAZU, Postojna, Slovenia

Richard J Plevin
University of California, Berkeley, CA, USA

Mark L Plummer
National Oceanic and Atmospheric Administration, Seattle, WA, USA

Stephen Polasky
University of Minnesota, St Paul, MN, USA

Gary A Polis†
University of California, Davis, CA, USA

Caroline M Pond
The Open University, Milton Keynes, UK

Dorota L Porazinska
University of Florida, FL, USA

Carlos Portillo-Quintero
Instituto Venezolano de Investigaciones Científicas (IVIC), Maracaibo, Venezuela

Hugh P Possingham
The University of Queensland, Queensland, Australia

Alison G Power
Cornell University, Ithaca, NY, USA

Ghillean T Prance
Royal Botanic Gardens, London, UK

Iain Colin Prentice
Macquarie University, North Ryde, Australia

Steven J Presley
University of Connecticut, Storrs, CT, USA

Jeff R Price
World Wildlife Fund, Washington, USA

Richard B Primack
Boston University, Boston, MA, USA

Jurgenne Primavera
Southeast Asian Fisheries Development Centre, Tigbauan, Philippines

Hugh W Pritchard
Royal Botanic Gardens, London, UK

Vânia Proença
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Petr Pyšek
Academy of Sciences and Charles University, Prague, Czech Republic

David A Pyke
USGS Forest and Rangeland Ecosystems Science Center, OR, USA

Sandra Quijas
Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Michoacán, Mexico

Oliver Rackham
University of Cambridge, Cambridge, UK

Katherine Ralls
Smithsonian Conservation Biology Institute, Washington, DC, USA

Aurelio Ramos
The Nature Conservancy, Cartagena de Indias, Colombia

Thiago F Rangel
Federal University of Goiás, Goiás, Brazil

Eduardo H Rapoport
Universidad Nacional del Comahue, Bariloche, Argentina

Jayashree Ratnam
National Center for Biological Sciences, Bangalore, India

John A Raven
University of Dundee, Dundee, UK

Marjorie L Reaka-Kudla
The University of Maryland, College Park, MD, USA

William E Rees
University of British Columbia, Vancouver, BC, Canada

Philip C Reid
Sir Alister Hardy Foundation for Ocean Science, Plymouth, UK

Gregory J Retallack
University of Oregon, Eugene, OR, USA

Carmen Revenga
The Nature Conservancy, VA, USA

†Deceased.

Belinda Reyers
Council for Scientific and Industrial Research, Stellenbosch, South Africa

Colin S Reynolds
Freshwater Biological Association and NERC Institute of Freshwater Ecology, Cumbria, UK

James F Reynolds
Duke University, Durham, NC, USA

Anna-Louise Reysenbach
Portland State University, Portland, OR, USA

Stephen J Richards
South Australian Museum, Adelaide, SA, Australia

David M Richardson
Stellenbosch University, Matieland, South Africa

Matthew C Richardson
Conservation International, VA, USA

Taylor Ricketts
University of Vermont, Burlington, VT, USA

Margaret L Rising
Portland State University, Portland, OR, USA

Bruce A Robertson
Smithsonian Conservation Biology Institute, Washington, DC, USA, and Michigan State University, East Lansing, MI, USA

Scott K Robinson
University of Florida, Gainesville, FL, USA

George K Roderick
University of California, Berkeley, CA, USA

Ana SL Rodrigues
Centre d'Ecologie Fonctionnelle et Evolutive, Montpellier, France

Jon Paul Rodríguez
Instituto Venezolano de Investigaciones Científicas, Maracaibo, Venezuela

Derek Roff
McGill University, Montreal, QC, Canada

Klaus Rohde
University of New England, NSW, Australia

Gary J Roloff
Michigan State University, East Lansing, MI, USA

Richard B Root
Cornell University, Ithaca, NY, USA

Terry L Root
Stanford University, Stanford, CA, USA

Patrik Rönnbäck
Gotland University, Visby, Sweden

Michael R Rose
University of California, CA, USA

Tatjana Rosen
Wildlife Conservation Society, Bronx, NY, USA

A Ross Kiester
Turtle Conservancy, NY, USA

Stephen I Rothstein
University of California, Santa Barbara, CA, USA

Stephen Roxburgh
Australian National University, Canberra, ACT, Australia

Daniel I Rubenstein
Princeton University, Princeton, NJ, USA

Dustin R Rubenstein
Columbia University, New York, NY, USA

Mary H Ruckelshaus
Stanford University, Stanford, CA, USA

PC de Ruiter
Wageningen University, Wageningen, The Netherlands

Philip W Rundel
University of California, Los Angeles, CA, USA

A Rus Hoelzel
Durham University, Durham, UK

Joel L Sachs
University of California, Riverside, CA, USA

Carl Safina
Blue Ocean Institute, Cold Spring Harbor, NY, USA

Rowan F Sage
The University of Toronto, Toronto, ON, Canada

Tammy L Sage
The University of Toronto, Toronto, ON, Canada

Osvaldo E Sala
Arizona State University, Tempe, AZ, USA

David Saltz
Ben Gurion University of the Negev, Israel

Arturo Sánchez-Azofeifa
University of Alberta, Edmonton, AB, Canada

Kaj Sand-Jensen
University of Copenhagen, Copenhagen, Denmark

M Sanjayan
The Nature Conservancy, Missoula, MT, USA

Mahesh Sankaran
National Center for Biological Sciences, Bangalore, India, and University of Leeds, Leeds, UK

José Sarukhán
Universidad Nacional Autónoma de México, México

Kathryn A Saterson
U.S. Environmental Protection Agency, NC, USA

Terre Satterfield
University of British Columbia, Vancouver, BC, Canada

Carl W Schaefer
University of Connecticut, Storrs, CT, USA

PH Schalk
ETI, Universiteit van Amsterdam, Amsterdam, Netherlands

Robert M Scheller
Portland State University, Portland, OR, USA

Beate Scherf
Food and Agriculture Organization of the United Nations, Rome, Italy

Michael H Schiewe
Anchor QEA, WA, USA

Carrie A Schloss
University of Washington, Seattle, WA, USA

Niels M Schmidt
Aarhus University, Roskilde, Denmark

David Clayton Schneider
Memorial University of Newfoundland, Newfoundland and Labrador, Canada

Stephen H Schneider†
Stanford University, Stanford, CA, USA

Christopher A Scholin
Monterey Bay Aquarium Research Institute, CA, USA

Lisa A Schulte
Iowa State University, Ames, IA, USA

Egbert Schwartz
University of California, Davis, CA, USA

Kate M Scow
University of California, Davis, CA, USA

Jason J Scullion
University of Washington, Seattle, WA, USA

Catherine Searle
Georgia Institute of Technology, Atlanta, GA, USA

Reinmar Seidler
Ashoka Trust for Research in Ecology and the Environment, India

J John Sepkoski Jr.†
University of Chicago, IL, USA

Manu Sharma
Stanford University, Stanford, CA, USA

Jonathan H Sharp
University of Delaware, Newark, DE, USA

Andrew N Sharpley
University of Arkansas, Fayetteville, AR, USA

Ruth G Shaw
University of Minnesota, St Paul, MN, USA

Douglas Sheil
Mbarara University of Science and Technology, Kabale, Uganda; Center for International Forestry Research, Bogor, Indonesia, and Southern Cross University, Lismore, NSW, Australia

Maura Shelton
University of Washington, Seattle, WA, USA

Jason F Shogren
University of Wyoming, Laramie, WY, USA

Bryan Shorrocks
University of Leeds, Leeds, UK

HH Shugart
University of Virginia, Charlottesville, VA, USA

Ilyas Siddique
Universidade Federal de Santa Catarina, Brazil

John A Silander Jr.
University of Connecticut, Storrs, CT, USA

Brian R Silliman
University of Florida, Gainesville, FL, USA

Daniel Simberloff
University of Tennessee, Knoxville, TN, USA

Ellen L Simms
University of California, Berkeley, CA, USA

Charles Sims
Utah State University, Logan, UT, USA

Karndeo D Singh
Academy of Forests and Environmental Sciences, New Delhi, India

Jean-Marc Sinnassamy
Global Environment Facility, Washington, DC, USA

†Deceased.

†Deceased.

LB Slobodkin
State University of New York-Stony Brook, Stony Brook, NY, USA

Melinda Smale
International Maize and Wheat Improvement Center, Texcoco, Mexico

David Smith
University of Maryland, College Park, MD, USA

William H Smith
Yale University, CT, USA

Paul VR Snelgrove
Memorial University of Newfoundland, St. John's, NL, Canada

Kei Sochi
The Nature Conservancy, Colorado, CO, USA

Pamela S Soltis
University of Florida, Gainesville, FL, USA

Mark Spalding
The Nature Conservancy, Newmarket, UK

Tim Sparks
Technische Universität München, Germany; Poznań University of Life Sciences, Poland, and University of Cambridge, UK

Alison Specht
The University of Queensland, Brisbane, QLD, Australia

Raymond Louis Specht
The University of Queensland, Brisbane, QLD, Australia

Leslie E Sponsel
University of Hawai'i, Honolulu, HI, USA

John L Spudich
The University of Texas Medical School, Houston, TX, USA

Erko Stackebrandt
DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany

Robert S Steneck
University of Maine, Orono, ME, USA

Carly J Stevens
The Open University, Milton Keynes, UK, and Lancaster University, Lancaster, UK

Grieg F Steward
University of Hawaii at Manoa, Honolulu, HI, USA

Thomas J Stohlgren
U.S. Geological Survey, Fort Collins, CO, USA

Andrew Storfer
Washington State University, Pullman, WA, USA

DL Strayer
Cary Institute of Ecosystem Studies, Millbrook, NY, USA

David L Strayer
Cary Institute of Ecosystem Studies, Millbrook, NY, USA

Lucy A Sutherland
Australian National Botanic Gardens, Canberra, ACT, Australia

Jens-Christian Svenning
Aarhus University, Aarhus, Denmark

Harald U Sverdrup
Lund University, Lund, Sweden

DP Swaney
Cornell University, NY, USA

Ian R Swingland
University of Kent, Canterbury, UK

Robert C Szaro
Smithsonian Conservation Biology Institute, Washington, DC, USA

Farzad Taheripour
Purdue University, West Lafayette, IN, USA

David Takacs
University of California, San Francisco, CA, USA

Heather Tallis
Stanford University, Stanford, CA, USA

Sandy L Tartowski
US Department of Agriculture, USA

Heather L Throop
New Mexico State University, Las Cruces, NM, USA

Bruce H Tiffney
University of California at Santa Barbara, CA, USA

David Tilman
University of Minnesota, St Paul, MN, USA

Víctor M Toledo
National University of México (UNAM), Morelia, México

Max Troell
Beijer Institute, Sweden and Stockholm University, Stockholm, Sweden

Andreas Troumbis
University of the Aegean, Lesvos, Greece

Hanna Tuomisto
University of Turku, Turku, Finland

Roy Turkington
University of British Columbia, Vancouver, BC, Canada

Ian M Turner
Singapore Botanic Gardens, Singapore

Will R Turner
The University of Arizona, Tucson, AZ, USA

Richard J Twitchett
Plymouth University, Plymouth, UK

Wallace E Tyner
Purdue University, West Lafayette, IN, USA

AJ Underwood
University of Sydney, Sydney, NSW, Australia

James W Valentine
University of California, Berkeley, CA, USA

Glenn R Van Blaricom
US Geological Survey and University of Washington, Seattle, WA, USA

Cindy Lee Van Dover
College of William & Mary, Williamsburg, VA, USA

Richard Irwin Vane-Wright
University of Kent, Cantenbury, UK; Natural History Museum, London, UK, and Canterbury Christ Church University, Cantenbury, UK

Fernando Veiga
The Nature Conservancy, Curitiba, Brazil

Giorgini A Venturieri
Universidade Federal de Santa Catarina, Florianópolis, Brazil

Costantino Vetriani
Rutgers University, New Brunswick, NJ, USA

Ross A Virginia
Dartmouth College, Hanover, NH, USA

Marcel E Visser
Netherlands Institute of Ecology, The Netherlands

Lucía Vivanco
Universidad de Buenos Aires, Buenos Aires, Argentina

Kristiina A Vogt
University of Washington, Seattle, WA, USA

Thomas J Volk
University of Wisconsin-La Crosse, La Crosse, WI, USA

Vance T Vredenburg
San Francisco State University, San Francisco, CA, USA

David L Wagner
University of Connecticut, Storrs, CT, USA

David B Wake
University of California, Berkeley, CA, USA

Diana H Wall
Colorado State University, Fort Collins, CO, USA

Robin S Waples
NOAA Fisheries, Seattle, WA, USA

Rachel Warren
University of East Anglia, Norwich, UK

Yoko Watanabe
Global Environment Facility, Washington, DC, USA

Thompson Webb
Brown University, RI, USA

Bregje K van Wesenbeeck
Unit for Marine and Coastal Systems, Deltares, MH, Delft, The Netherlands

Paul C West
University of Minnesota, St. Paul, MN, USA

Peter H Weston
University of Adelaide and National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, NSW, Australia

Robert G Wetzel
University of Alabama, Tuscaloosa, AL, USA

Dan Wharton
The Chicago Zoological Society, Brookfield, IL, USA

Quentin D Wheeler
Arizona State University, Tempe, AZ, USA

Gary C White
Colorado State University, Fort Collins, CO, USA

Kellie Whittaker
University of California, Berkeley, CA, USA

Jeannette Whitton
University of British Columbia, Vancouver, Canada

Richard G Wiegert
University of Georgia, Athens, GA, USA

David S Wilcove
Princeton University, Princeton, NJ, USA

David M Williams
The Natural History Museum, London, UK

Paul Williams
Natural History Museum, London, UK

Stephen E Williams
James Cook University, Townsville, QLD, Australia

Michael R Willig
University of Connecticut, Storrs, CT, USA

John Wilson
South African National Biodiversity Institute (SANBI), Claremont, South Africa

Marten Winter
Helmholtz Center for Environmental Research - UFZ, Leipzig, Germany

A Wiseman
University of Surrey, Guildford, UK

Elisha M Wood-Charlson
University of Hawaii at Manoa, Honolulu, HI, USA

Rosie Woodroffe
Zoological Society of London, UK

FI Woodward
University of Sheffield, Sheffield, UK

Donald J Wuebbles
University of Illinois at Urbana-Champaign, Urbana, IL, USA

Weihua Xu
Chinese Academy of Sciences, Beijing, China

Ming Yang
Global Environment Facility, Washington, DC, USA

Wu Yang
Michigan State University, East Lansing, MI, USA

Peter Yodzis
University of Guelph, Ontario, Canada

Shibu Yooseph
J. Craig Venter Institute, San Diego, CA, USA

Richard Young
Durrell Wildlife Conservation Trust, Trinity, Jersey

Peter Zahler
Wildlife Conservation Society, Bronx, NY, USA

Patricia Zaradic
Red Rock Institute, Bryn Mawr, PA, USA

Joy B Zedler
University of Wisconsin, WI, USA

Fusuo Zhang
China Agricultural University, Beijing, China

Lizhen Zhang
China Agricultural University, Beijing, China

Mark Zimsky
Global Environment Facility, Washington, DC, USA

Daniel Zohary
The Hebrew University, Jerusalem, Israel

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

A

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems

Charles T Driscoll, Syracuse University, Syracuse, NY, USA

Published by Elsevier Inc.

This article is a revision of the previous edition article by Hendrey, G., volume 1, pp 1–15, © 2001, Elsevier Inc.

Glossary

Acid neutralizing capacity (ANC) The amount of strong acid that is required to decrease the pH of water to a defined chemical endpoint. Acid neutralizing capacity is measured in units of $\mu eq\ l^{-1}$. It is a measure of the ability of water to neutralize inputs of strong acids.

Acid-sensitive area Characterized by low supply rates of acid neutralizing basic cations. These areas generally occur in high elevation forest regions and are characterized by shallow base poor soil.

Airshed or emission source area An area where emissions of air pollutants originate that are eventually deposited to a position on the Earth's surface.

Atmospheric deposition The process by which materials are deposited from the atmosphere to the Earth's surface.

Bioaccumulation A process by which contaminants, such as mercury, accumulate along food chains.

Methyl Mercury (MHg) The form of mercury largely produced by sulfate and iron reducing bacteria that strongly bioaccumulates along food chains.

Nitrogen saturation A condition where ecosystems, which were previously growth-limited by nitrogen inputs are shifting to limitation by some other resource.

pH The concentration of free hydrogen ion or proton, measured on a logarithmic scale. Thus a solution with a concentration of hydrogen ion of 1 $eq\ l^{-1}$ has a pH of 0 and is an extremely acidic solution. A hydrogen ion concentration of 0.000,01 $eq\ l^{-1}$, typical of a chronically acidic surface water, has a pH of 5. A hydrogen ion concentration of 0.000,000,1 $eq\ l^{-1}$ is pH 7, or neutral pH. Note that because pH is a logarithmic scale, as pH values decrease from neutral values, the amount of strong acid required to decrease the pH by a unit value increases greatly. Water with a pH of 5 has 100 times the hydrogen ion concentration as a pH 7; a pH of 4 has 10 times the hydrogen ion concentration of pH 5 or 1000 times the hydrogen ion concentration of a pH 7 solution.

Watershed An area of the landscape, which generally follows topographic features, where incoming precipitation collects and forms water drainage.

Introduction

Detailed studies by scientists for more than five decades have provided considerable insight on how atmospheric deposition alters the structure and function of ecosystems. When it was first identified, acidic deposition was viewed as a simple problem that was limited in scope. Scientists know now that acids, acidifying compounds, mercury, and other substances enter ecosystems from atmospheric deposition and are transported through soil, vegetation, and surface waters, resulting in a range of ecosystem effects (**Figure 1**).

In this article, information is synthesized on patterns of acidic and mercury deposition, and the effects of atmospheric sulfur, nitrogen, and mercury deposition on sensitive forest, and freshwater resources.

Emissions of Acidifying Substances and Acidic Deposition

Acidic deposition is largely comprised of sulfuric and nitric acid derived from sulfur dioxide and nitrogen oxides emissions, respectively, and ammonium resulting from emissions of ammonia. Sulfur dioxide and nitrogen oxides, originating from human activities, are largely emitted to the atmosphere by the burning of fossil fuels, while ammonia is largely the result of agricultural activities (Driscoll *et al.*, 2001a, 2003). While there are also natural sources of these materials, they tend to be small compared to human sources in developed regions of North America, Europe, and Asia. Once these compounds enter an ecosystem, they can acidify soil and surface waters, bringing about a series of ecological changes.

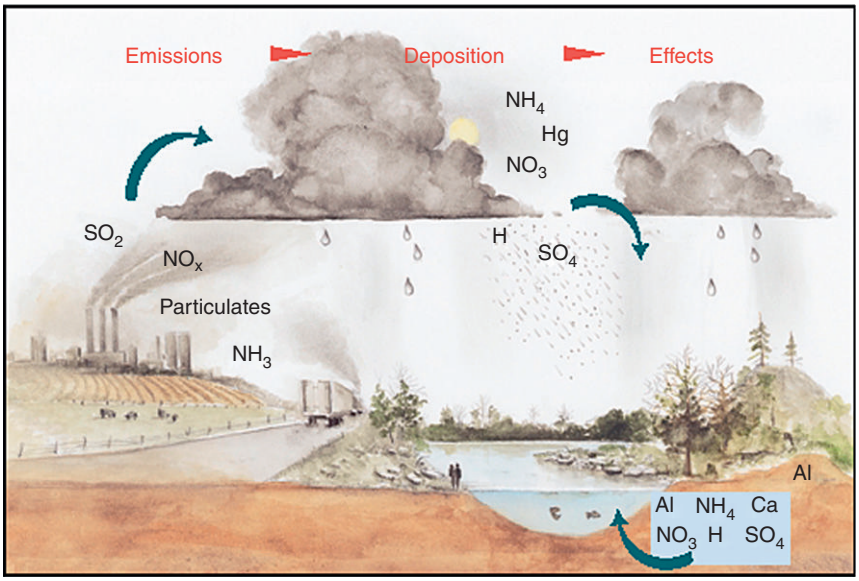


Figure 1 Conceptual diagram of acidic deposition. Acidic deposition is a complex environmental problem that originates with combustion of fossil fuels and leads to deposition of acids and results in a series of ecological effects (Driscoll *et al.*, 2001b). Courtesy of the Hubbard Brook Research Foundation.

The term acidic deposition or acidifying deposition encompasses all of the forms of these compounds that are transported from the atmosphere to the Earth, including gases, particles, rain, snow, clouds, and fog. Acidic deposition occurs as wet deposition: as rain, snow, sleet, or hail; as dry deposition, as particles or vapor; or as cloud or fog deposition, which is more common at high elevations or in coastal areas. Wet deposition is fairly well characterized by monitoring at about 250 National Atmospheric Deposition Programs (NADP; *http://nadp.sws.uiuc.edu/*) in the US. In contrast dry deposition is highly dependent on meteorological condition and vegetation characteristics, which can vary markedly over short distances in complex terrains. As a result dry deposition is generally poorly characterized and highly uncertain. Dry deposition is estimated through the Clean Air Status and Trends Network (CASTNet; *http://www.epa.gov/castnet/*), which includes about 100 sites in the US.

Sulfuric and nitric acids lower the pH of rain, snow, soil, lakes, and streams. Areas experiencing elevated acidic deposition include eastern North America (**Figure 2**), the western US, Europe, and Asia. In 2007–2009, wet deposition (i.e., deposition from forms of precipitation such as rain, snow, sleet, and hail) in acid-sensitive regions of the eastern US had average pH values of 4.4–5.0, which is about two to five times more acidic than background conditions.

Acidic deposition trends in the eastern US and Europe mirror emission trends in the atmospheric source area or airshed. Long-term data from across the eastern US and Europe show declining concentrations of sulfate in wet deposition since the mid-1970s, coincident with decreases in sulfur dioxide emissions. Based on these long-term data, a strong positive correlation exists between sulfur dioxide emissions in the source area and sulfate concentrations in wet deposition. It is expected that the sulfate concentration of wet deposition will decrease (or increase) in an approximately linear response to the decrease (or increase) of sulfur dioxide emissions in the atmospheric source area (Driscoll *et al.*, 2001a). These observations strongly suggest a cause and effect relationship between emissions of sulfur dioxide and deposition of sulfate. A similar relationship is starting to become evident between emissions of nitrogen oxides and wet deposition of nitrate. This relationship for nitrate is not as strong as the relationship for sulfur because emissions of nitrogen oxides were relatively constant through the 1980s and 1990s. However, in the US recent decreases in emissions of nitrogen oxides largely from electric utilities have resulted in decreases in atmospheric deposition of nitrate (Butler *et al.*, 2003).

Mercury Emissions and Deposition

Globally mercury is released to the environment in several ways, but the dominant pathway is airborne emissions and deposition. Fossil fuel combustion, and metal, gold and cement production rank among the highest human mercury emission sources worldwide (AMAP/UNEP, 2008). The largest single source of mercury emissions in the US is coal-fired power plants (Driscoll *et al.*, 2007a: http://www.epa.gov/ttn/chief/net/2005inventory. html).

The total amount of mercury released to the atmosphere worldwide is approximately 14,520,000 lbs (7,260 short tons), of which two-thirds originates from anthropogenic sources (Mason and Sheu, 2002). Mercury is a naturally occurring substance and about one-third of mercury emissions are from geologically generated sources. While overall mercury emissions in the US decreased from 490,000 lbs (245 tons) in 1990 to 228,000 lbs (107 tons) in 2005, emissions from coal-fired power plants did not change appreciably during that period and are now the largest source in the US.

One key to understanding where mercury deposits following release to the atmosphere is the nature of mercury emissions and deposition pathways. Mercury is emitted to the atmosphere in several different forms: elemental mercury,

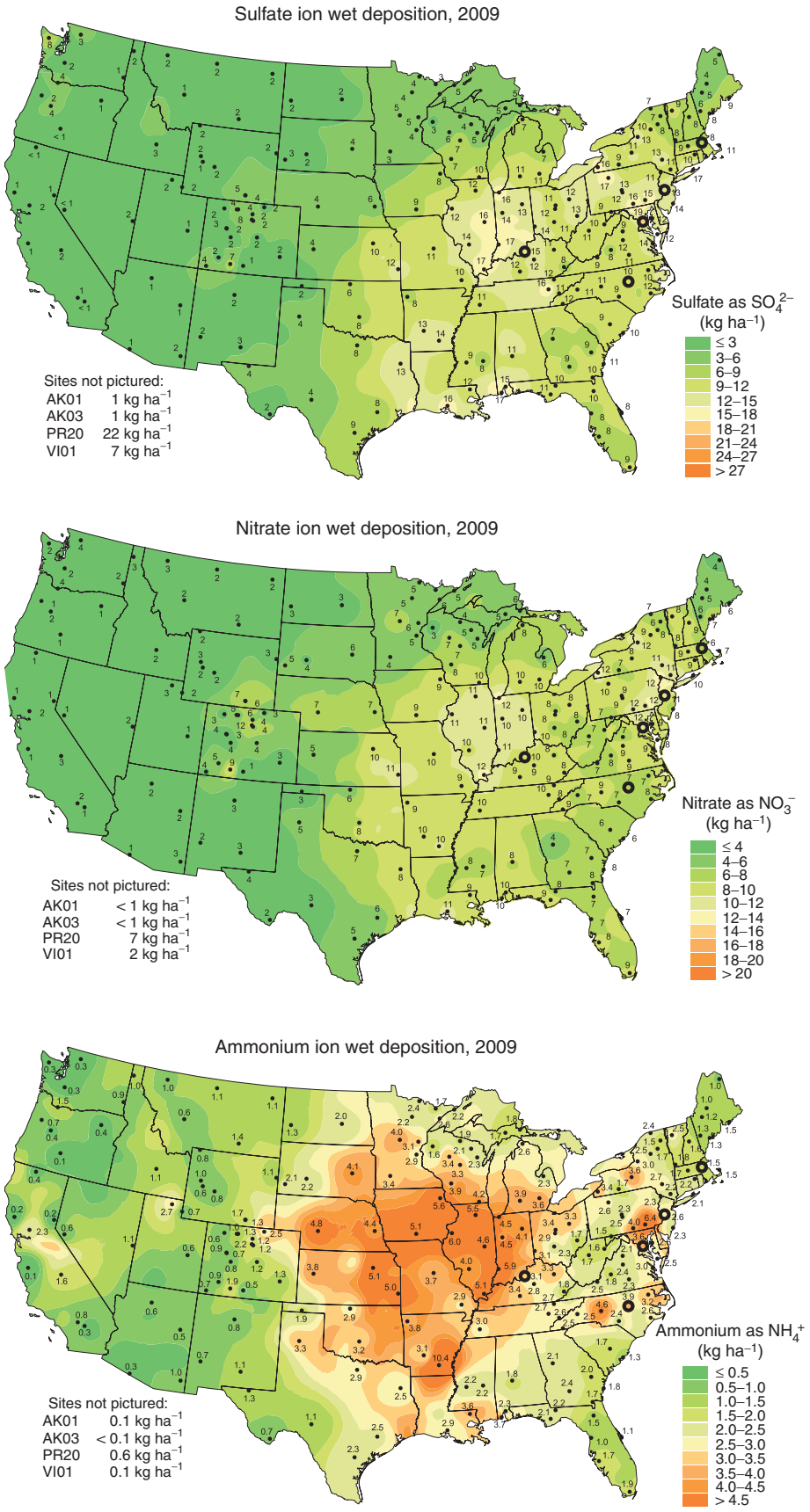


Figure 2 Map of wet sulfate, nitrate, and ammonium deposition for the US for 2009.

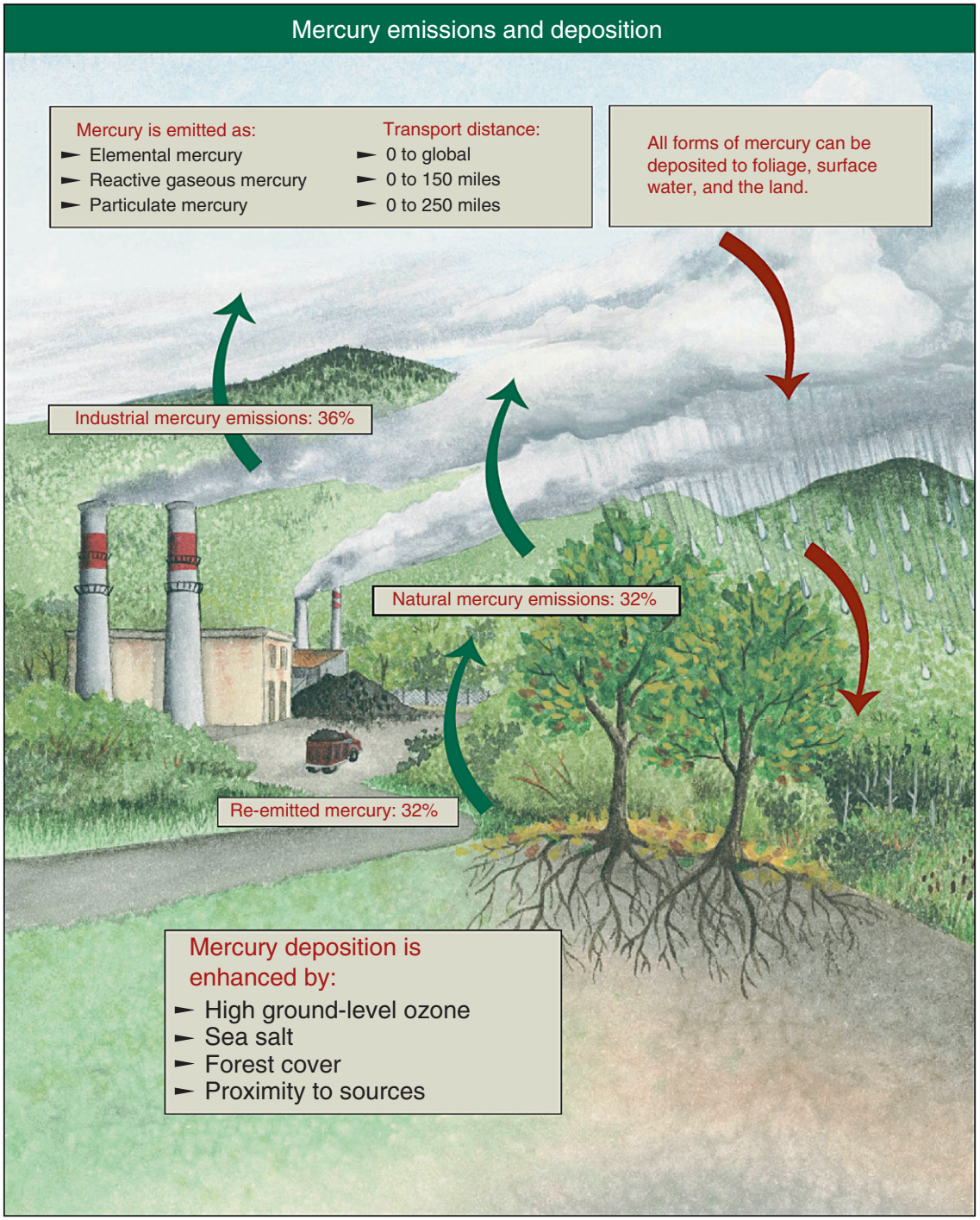


Figure 3 Different forms of mercury can travel varying distances before depositing to foliage, surface waters, and the land. Also, certain conditions enhance mercury deposition (Driscoll *et al.*, 2007b). Courtesy of the Hubbard Brook Research Foundation.

reactive gaseous mercury, and particulate mercury (**Figure 3**). The latter two forms together are known as oxidized (or ionic) mercury. Elemental mercury is a relatively inert gas that does not readily dissolve in water. Therefore, it generally travels farther in the atmosphere than oxidized mercury; it is truly a global contaminant with a residence time in the atmosphere of 0.5–2 years. Oxidized mercury is more chemically reactive and soluble in water than elemental mercury, and therefore generally deposits more rapidly, hours to days following emissions, and closer to the source (Driscoll *et al.*, 2007a). Elemental mercury can be converted to oxidized mercury in the atmosphere and deposited to the Earth's surface, and oxidized mercury can be converted to reduced mercury. Note this latter process largely occurs at the Earth's surface and is referred to as evasion or re-emission of mercury. Re-emission is an important process by which

previously deposited mercury can be emitted back to the atmosphere.

Scientists have identified pathways for direct deposition of elemental mercury. Trees and other vegetation take up elemental mercury by scavenging it out of the atmosphere with their rough leaf surfaces, or through gas exchange in pores on the leaf surface known as stomates (Driscoll *et al.*, 2007a). In forests of the Northeast, estimates suggest that up to 70% of total mercury deposition may occur through this "dry deposition" process (Driscoll *et al.*, 2007a). Based on what is known about how the various forms of mercury are transported or converted in the atmosphere, and the importance of dry deposition, it can be concluded that all forms of mercury (oxidized and elemental) may deposit locally or regionally.

Scientists use several methods to estimate mercury deposition from the atmosphere to the land surface. Like acidic deposition, mercury deposition occurs as both wet and dry deposition. Wet deposition in rain or snow is measured at a network of approximately 90 stations in North America known as the Mercury Deposition Network (MDN; http://nadp.sws.uiuc.edu/mdn/). MDN is part of the National Atmospheric Deposition Program (NADP). Dry deposition is not systematically measured in the US due to technical challenges, but it is estimated using a variety of approaches.

Effects of Acidic Deposition on Forest Ecosystems

In acid-sensitive regions, acidic deposition alters soils, stresses forest vegetation, acidifies lakes and streams, and harms fish and other aquatic life. These effects can interfere with important ecosystem functions and services such as forest and aquatic biodiversity and productivity, and soil and water quality. Years of acidic deposition have also made many ecosystems increasingly sensitive to continuing pollution. Moreover, the same pollutants that cause acidic deposition contribute to a wide array of other important environmental issues at local, regional, and global scales (**Table 1**).

Effects of Acidic Deposition on Forest Soils

Research has shown that acidic deposition has chemically altered soils with serious consequences for acid-sensitive ecosystems. Soils compromised by acidic deposition lose their ability to neutralize continuing inputs of strong acids, provide poorer growing conditions for plants, and extend the time needed for ecosystems to recover from acidic deposition.

Acidic deposition has altered and continues to alter soils in acid-sensitive regions in three important ways. Acidic deposition depletes available calcium and magnesium from exchange sites in soil; facilitates the mobilization of dissolved inorganic aluminum into soil water; and increases the accumulation of sulfur and nitrogen in soil.

Table 1 The links between sulfur dioxide and nitrogen oxide emissions, acidic deposition, and other important environmental issues

Problem	*Linkage to acid deposition*
Coastal, terrestrial, and freshwater eutrophication	Atmospheric deposition adds nitrogen to ecosystems, increasing productivity and altering species richness.
Mercury	Surface water acidification increases mercury accumulation in fish.
Visibility	Sulfate aerosols diminish visibility and views.
Tropospheric ozone	Emissions of nitrogen oxides contribute to the formation of ozone.

Loss of Calcium and Other Nutrient Cations

The cycling of calcium and other nutrient cations (magnesium, potassium) in forest ecosystems involves the inputs and losses of these materials (**Figure 4**). For most forest ecosystems, the supply of calcium and other nutrient cations largely occurs by weathering (i.e., the breakdown of rocks and minerals in soil). Calcium and other nutrient cations may also enter forests by atmospheric deposition, although this pathway is generally much smaller than weathering. Losses largely occur by vegetation uptake and drainage waters. An important pool of ecosystem calcium and nutrient cations is the soil available pool or the soil cation exchange complex. Plants are generally able to utilize this source of nutrients. Forest ecosystems that are naturally sensitive to acidic deposition are generally characterized by low rates of weathering and generally low quantities of available nutrient cations. Under conditions of elevated inputs of acidic deposition and subsequent transport of sulfate and nitrate in drainage waters, nutrient cations are displaced from available pools and leached from soil. This condition is not problematic for areas with high weathering rates and high pools of available nutrient cations. However, in acid-sensitive areas with shallow soil and which contain minerals that are resistant to weathering, the enhanced loss of calcium and other nutrient cations can result in a depletion of soil available pools. Acid sensitive conditions generally occur in high elevation often forested regions with acidic soils that are underlain by surficial materials or bedrock that are resistant to weathering; for the US acid sensitive regions are shown in **Figure 5**.

Over the last century, acidic deposition has accelerated the loss of large amounts of available calcium from acid-sensitive soil in acid-sensitive areas. Depletion occurs when nutrient cations are displaced from the soil by acidic deposition at a rate faster than they can be replenished by the slow breakdown of rocks and minerals or the deposition of nutrient cations from the atmosphere. This depletion of nutrient cations fundamentally alters soil processes, compromises the nutrition of some trees, and hinders the capacity for sensitive soils to ultimately recover. For example, more than half of the available calcium has been lost from soil at the Hubbard Brook Experimental Forest, New Hampshire over the past 60 years (Likens *et al.*, 1996). Note that while acidic deposition to acid-sensitive areas is decreasing in North America and Europe, and there is some associated recovery of the acid neutralizing capacity (ANC) of surface waters (*see* Surface Water Acidification), it appears that forest soils continue to exhibit net depletion of exchangeable nutrient cations.

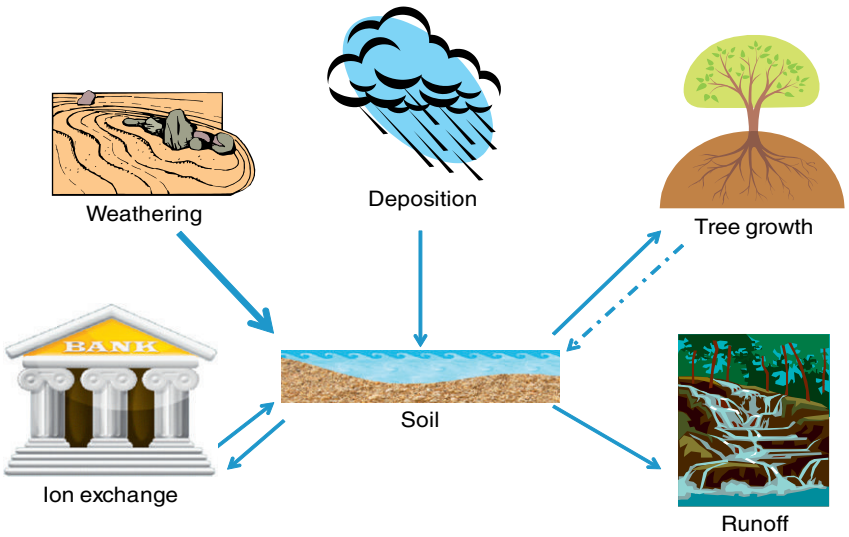


Figure 4 Conceptual diagram illustrating the inputs and losses of calcium and other nutrient cations for forest ecosystems. Ecosystems with low weathering rates have low supplies of calcium and other nutrient cations available to soil and trees. High inputs of acidic deposition increase the leaching of calcium and other nutrient cations from soil. In acid-sensitive regions with low weathering rates, inputs of acidic deposition can deplete soil pools of available calcium and other nutrient cations and limit quantities available for tree growth.

Mobilization of Aluminum

Aluminum is often released from soil to soil water, lakes, and streams in forested regions with high acidic deposition, low stores of available calcium, and high soil acidity. One of the most significant ecological effects of acidic deposition is the mobilization of aluminum from soil and a shift in the form of dissolved aluminum in water from nontoxic organic forms to highly toxic inorganic forms.

Concentrations of aluminum increase markedly with decreases in pH, particularly the toxic inorganic forms of aluminum. It is evident that concentrations of aluminum increase exponentially when surface water pH decreases below 6.0. Aluminum concentrations are thought to be ecologically significant when they increase to values above 2 μmol l^{-1} (Driscoll *et al.*, 2001a). This condition can occur when pH values decrease below 6.0.

High concentrations of dissolved inorganic aluminum can be toxic to plants, fish, and other organisms. Concentrations of dissolved inorganic aluminum in streams in acid sensitive areas of eastern North America and areas of Europe are often above levels considered toxic to fish, and much greater than concentrations observed in forest watersheds that receive low inputs of acidic deposition.

Accumulation of Sulfur and Nitrogen

Increases in atmospheric sulfur and nitrogen deposition will increase the quantities of these elements that accumulate in soil and ecosystems. This perturbation can have both positive and negative effects. The accumulation of sulfur in soil from atmospheric deposition occurs through adsorption of sulfate on mineral soil surfaces and increases inorganic sulfur through the accumulation of soil organic matter. This increase in soil sulfur will have limited direct adverse effects on soil quality, but following decreases in atmospheric sulfur deposition will release sulfate from soil to surface waters delaying recovery from acidic deposition. This process is particularly prominent in regions with unglaciated soil that have a large capacity to retain sulfate, such as the southeastern US (Galloway *et al.*, 1983).

Likewise, increased deposition of nitrogen increases ecosystem retention of nitrogen. Increases in nitrogen loading likely accelerate a process known as "nitrogen saturation" (Aber *et al.*, 1989). Nitrogen saturation is a condition where ecosystems, which were previously growth-limited by nitrogen inputs are shifting to limitation by some other resource. Nitrogen saturation results in changes in the structure and function of ecosystems and increases leaching losses of nitrate to surface waters contributing to acidification.

Elevated atmospheric nitrogen deposition can also have beneficial effects on ecosystems. As mentioned, nitrogen is generally the growth-limiting nutrient to terrestrial ecosystems. Increases in nitrogen loading should be accompanied by an increase in net ecosystem production and biomass. This perturbation enhances sequestration of carbon from the atmosphere.

Effects of Acidic Deposition on Trees

Acidic deposition has contributed to the decline of red spruce and sugar maple trees in the eastern US (**Figure 6**), and shifts in the distribution of tree species. Symptoms of tree decline include poor condition of the canopy, reduced growth, and unusually high levels of mortality. Declines of red spruce and sugar maple in the northeastern US have occurred during the past five decades. Factors associated with declines of both species have been studied and include important links to acidic deposition.

Red Spruce

Since the 1960s, more than half of large canopy red spruce in the Adirondack Mountains of New York and the Green Mountains of Vermont and approximately one quarter of large canopy red spruce in the White Mountains of New Hampshire have died (DeHayes *et al.*, 1999). Significant growth declines

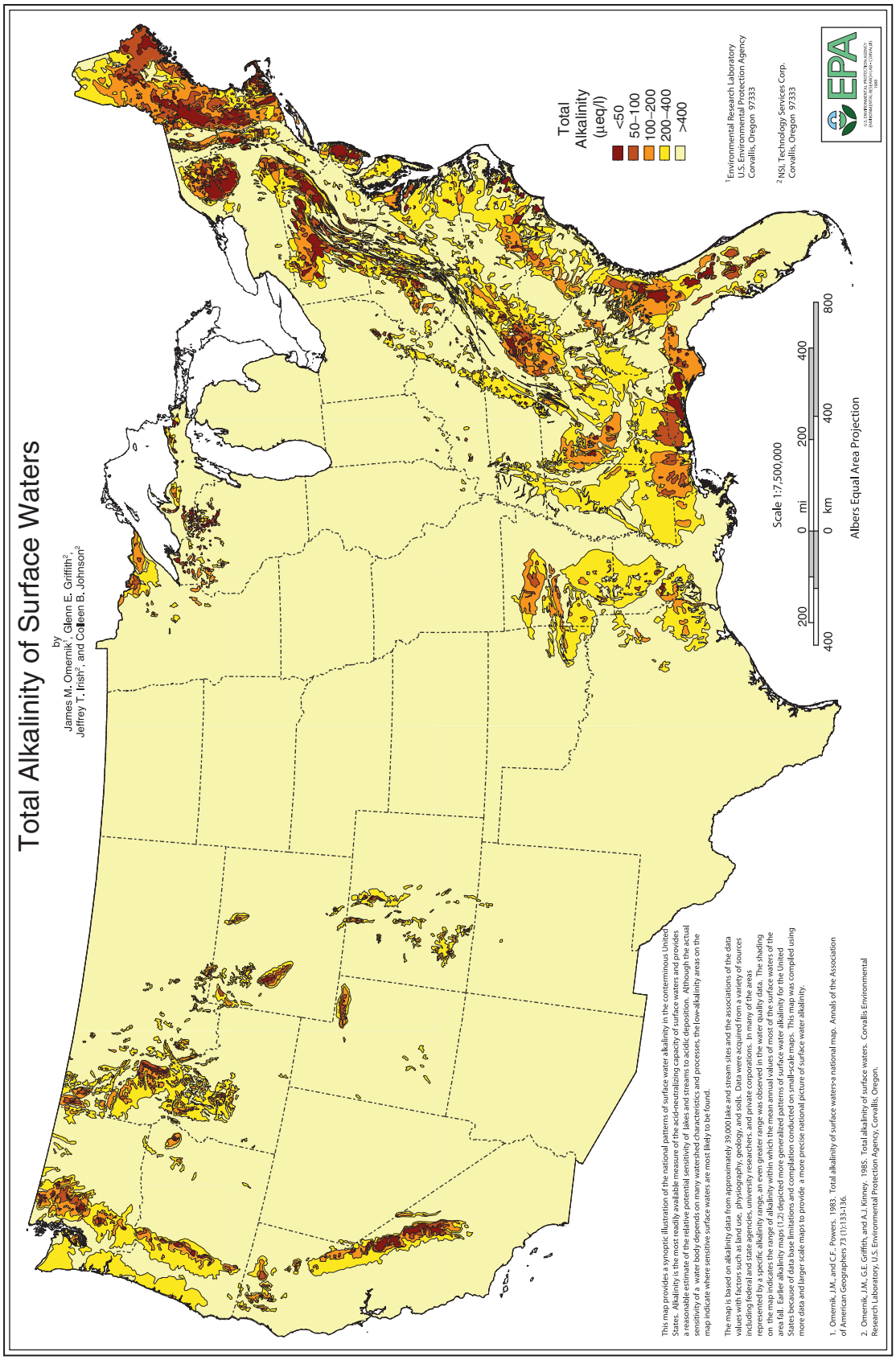


Figure 5 Map showing areas of the US with low values of acid neutralizing capacity in surface waters that are potentially sensitive to acidic deposition (Eilers and Selle, 1991).

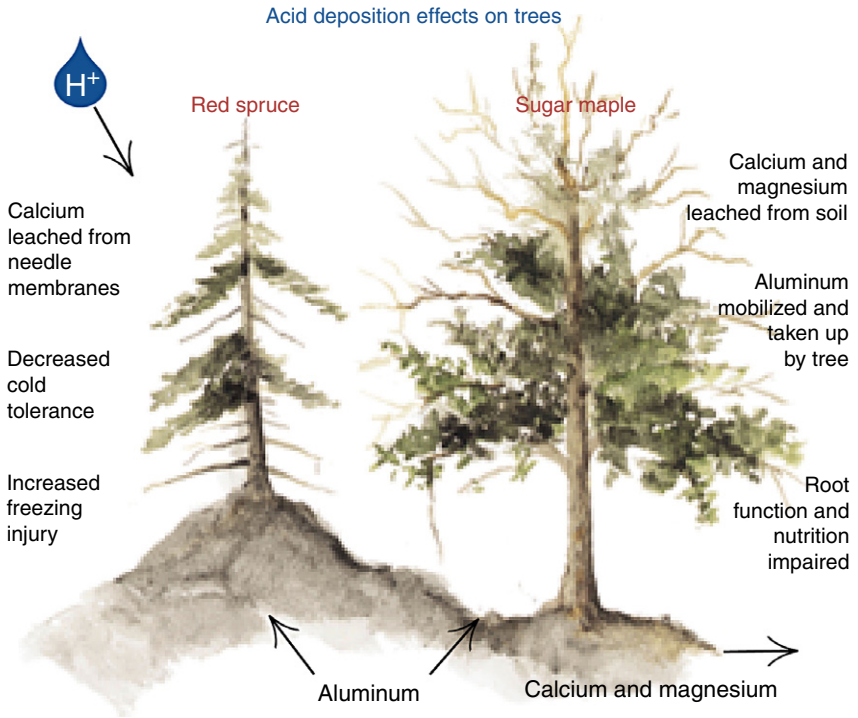


Figure 6 Conceptual diagram illustrating the mechanisms by which acidic deposition impacts red spruce and sugar maple (Driscoll *et al.*, 2001b). Acidic deposition impacts red spruce through loss of membrane calcium due to direct leaching from foliage or reduced uptake of calcium from soil. The loss of membrane calcium makes red spruce more susceptible to winter injury. Acidic deposition results in loss of soil available calcium and magnesium and less uptake by sugar maple. This condition may make sugar maple more susceptible to insect or drought stress. Courtesy of the Hubbard Brook Research Foundation.

and winter injury to red spruce have been observed throughout its range. Acidic deposition is a major cause of red spruce decline at high elevations in the Northeast.

Red spruce decline occurs by both direct and indirect effects of acidic deposition. Direct effects include the leaching of calcium from leaves and needles of trees (i.e., foliage), whereas indirect effects refer to acidification of the underlying soil chemistry.

The decline of red spruce is linked to the leaching of calcium from cell membranes in spruce needles by acid mist or fog. The loss of calcium renders the needles more susceptible to freezing damage, thereby decreasing the tolerance of trees to low temperatures and increasing the occurrence of winter injury and subsequent tree damage or death (DeHayes *et al.*, 1999). In addition, elevated aluminum concentrations in the soil may limit the ability of red spruce to take up water and nutrients through its roots. Water and nutrient deficiencies can lower the tolerance of trees to other environmental stresses and cause decline.

Sugar Maple

The decline of sugar maple has been studied in the eastern US since the 1950s. Extensive mortality among sugar maples appears to have resulted from deficiencies of nutrient cations (calcium, magnesium), coupled with other stresses such as insect defoliation or drought (Horsley *et al.*, 2000). The probability of decreases in the vigor of the sugar maple canopy or incidence of tree death increases on sites where the supply of calcium and magnesium to soil and foliage are lowest and stress from insect defoliation and drought is high. Low levels of nutrient cations can cause a nutrient imbalance and reduce the ability of a tree to respond to stresses such as insect infestation and drought.

Shifts in the Distribution of Tree Species with Increases in Atmospheric Nitrogen Deposition

Thomas *et al.* (2010) observed differential changes (both positive and negative responses) in annual aboveground carbon increment and 5-year survival rate as a function of total (wet + dry) inorganic nitrogen deposition for the 24 most common tree species in the northeastern US. Tree species such as red pine, red spruce, and white cedar show declining growth under elevated atmospheric nitrogen deposition, while 11 species show a positive growth response. This research suggests that different tree species are responding differently to increases in atmospheric nitrogen deposition and with ongoing elevated deposition there will be shifts in the species abundance.

Effects of Acidic and Mercury Deposition on Freshwater Aquatic Ecosystems

Surface Water Acidification

Acidic deposition degrades surface water quality by lowering pH (i.e., increasing acidity); decreasing ANC; and increasing

dissolved inorganic aluminum concentrations. While sulfate and to a lesser extent nitrate concentrations in lakes and streams have decreased in eastern North America and Europe over the last 20 years, they remain high compared to background conditions (e.g., approximately 20 $\mu eq\ l^{-1}$ for sulfate and 1 $\mu mol\ l^{-1}$ for nitrate). Moreover, improvement in other chemical conditions in many lakes and streams in acid-impacted regions has been limited.

Acidification of surface waters due to elevated inputs of acidic deposition have been reported in many acid-sensitive areas receiving elevated inputs of acidic deposition, including Great Britain, Nordic countries, Northern, Central and Eastern Europe, southwestern China, southeastern Canada, the northeastern US, the Upper Midwest and the Appalachian mountain region of the US. Large portions of the high elevation western US are also potentially sensitive to acidic deposition, however, atmospheric deposition to this region is relatively low. Concern over acidic deposition in the mountain West of the US has been overshadowed by potential effects of elevated nitrogen deposition, including eutrophication of naturally nitrogen-limited lakes (Fenn *et al.*, 2003).

To illustrate the regional impacts of acidic deposition, a comprehensive survey of lakes greater than 0.2 ha in surface area in the Adirondack region of New York was conducted in the mid-1980s to obtain detailed information on the acid-base status of waters in this region. Of the 1469 lakes surveyed, 24% had summer pH values below 5.0. Also 27% of the lakes surveyed were chronically acidic (i.e., ANC less than 0 $\mu eq\ l^{-1}$) and an additional 21% were susceptible to episodic, or short-term acidification (i.e., ANC between 0 and 50 $\mu eq\ l^{-1}$).

Seasonal acidification is the periodic increase in acidity and the corresponding decrease in pH and ANC in streams and lakes, which generally occurs during the higher flow fall, winter and spring periods. Episodic acidification is caused by the sudden pulse of acids and a dilution of base cations (i.e., calcium, magnesium, sodium, potassium) due to spring snowmelt and large rain events generally in the spring and fall. Acidic episodes are when the surface water ANC decreases to near or below 0 $\mu eq\ l^{-1}$ during events. Increases in nitrate are often important to the occurrence of acidic episodes. These conditions tend to occur when trees are dormant and therefore retain less nitrogen. At some sites short-term increases in sulfate and naturally occurring organic acids can also contribute to episodic acidification. Episodic acidification often coincides with pulsed increases in concentrations of dissolved inorganic aluminum. Short-term increases in acid inputs to surface waters can reach levels that are lethal to fish and other aquatic organisms. All of the acid-sensitive and acid-impacted regions discussed in this article have documented effects associated with episodic acidification.

Trends in surface water chemistry in Europe and eastern North America indicate that recovery of aquatic ecosystems impacted by acidic deposition is occurring over a large geographic scale since the early 1980s. Some regions are showing rather marked recovery, while others exhibit low or nonexistent increases in ANC. Based on long-term monitoring, virtually all surface waters impacted by acidic deposition in Europe and eastern North America exhibit decreases in sulfate concentrations. This pattern is consistent with decreases in emissions of sulfur dioxide and atmospheric sulfate deposition. The exception to this pattern is streams in the unglaciated southeastern US. Watersheds in the southeastern US exhibit strong adsorption of atmospheric sulfate deposition by highly weathered soils. In Europe, the most marked decreases in surface water sulfate have occurred in the Czech Republic and Slovakia, regions that experienced historically very high rates of atmospheric sulfate deposition (Evans *et al.*, 2001). More than half of the surface waters monitored in Europe are showing increases in ANC. The rate of ANC increase in Europe is relatively high. This pattern is, in part, not only due to the relatively high rates of sulfate decreases, but also to the fact that decreases in base cations are occurring at about half the rate of the decreases in sulfate plus nitrate, allowing for relatively large rates of ANC increases. In contrast, in the US only three regions are showing statistically significant increases in ANC; lakes in the Adirondacks and Upper Midwest and streams in the Northern Appalachian Plateau (Kahl *et al.*, 2004). In the US, decreases in the sum of base cations closely correspond to decreases in sulfate plus nitrate, limiting rates of ANC increase.

Three factors account for the relatively slow chemical recovery of the water quality of acid-impacted surface waters, despite the decreased deposition of sulfate. First, levels of acid-neutralizing base cations in streams have decreased markedly due to a loss of base cations from the soil and, to a lesser extent, a decrease in atmospheric inputs of base cations. Second, conditions of nitrogen saturation and nitrate leaching have contributed to the acidification of surface waters in many acid-impacted regions. Finally, sulfur has accumulated in the soil and is now being released to surface water as sulfate, even though sulfate deposition has decreased. It appears the only approach to accelerate the recovery of acid-impacted lakes is to make additional cuts in emissions of sulfur dioxide and nitrogen oxides.

The modest decreases in sulfate concentrations and increases in pH and ANC exhibited in some surface waters is an encouraging sign that impacted ecosystems are responding to emission controls and moving toward chemical recovery. Nevertheless, the magnitude of these changes is small compared to the magnitude of increases in sulfate and the decreases in acid neutralizing capacity that have occurred in acid-impacted areas following historical increases in acidic deposition. Moreover, as discussed above, in many acid-sensitive regions soils continue to acidify despite decreases in acidic deposition (*see* Loss of Calcium and Other Nutrient Cations).

Response of Aquatic Biota to Acidification of Surface Waters by Acidic Deposition

Decreases in pH and elevated concentrations of dissolved inorganic aluminum have resulted in physiological changes to organisms, direct mortality of sensitive life history stages, and diminished the species diversity and abundance of aquatic life in many streams and lakes in acid-impacted areas. Fish have received the most attention to date, but entire food webs are often adversely affected.

Decreases in pH and increases in aluminum concentrations have diminished the species diversity and abundance of plankton, invertebrates, and fish in acid-impacted surface

waters. A detailed summary of the response of aquatic biota to the acidification of surface waters is provided in **Table 2**.

In the Adirondacks, a significant positive relationship exists between the pH and ANC levels in lakes and the number of fish species present in those lakes (**Figure 7**). Surveys of 1469 Adirondack lakes conducted in 1984 and 1987 show that 24 percent of lakes (i.e., 346) in this region do not support fish. These lakes had consistently lower pH and ANC, and higher concentrations of aluminum than lakes that contained one or more species of fish. Experimental studies and field observations demonstrate that even acid-tolerant fish species such as brook trout have been eliminated from some waters.

Similar relationships are evident in surface waters in acid-impacted regions throughout the world. Studies demonstrate effects of acidic deposition on fish at three ecosystem levels:

- Effects on single organisms (condition factor – the relationship between the weight and the length of a fish). Fish condition factor is related to several chemical indicators of acid–base status, including minimum pH. This analysis suggests that fish in acidic streams use more energy to maintain internal chemistry that would otherwise be used for growth.
- Population-level effects (increased mortality). Bioassay experiments show greater mortality in chronically acidic streams than in high acid neutralizing capacity streams. Eggs and fry are sensitive life history stages for fish.
- Community-level effects (reduced species richness). The species richness of fish and other aquatic organisms decreases with decreasing acid neutralizing capacity and pH.

Although chronically high acid levels stress aquatic life, acidic episodes are particularly harmful because abrupt, large changes in water chemistry allow fish few areas of refuge. High concentrations of dissolved inorganic aluminum are directly toxic to fish and pulses of aluminum during acid episodes are a primary cause of fish mortality. High acidity and aluminum concentrations disrupt the salt and water balance of blood in a fish, causing red blood cells to rupture and blood viscosity to increase. Studies show that the viscous blood strains the heart of a fish, resulting in mortality.

Acid neutralizing capacity is commonly used to assess the acid–base status of surface waters. The US Environmental Protection Agency (USEPA, 2009) suggested five ANC classes coinciding with levels of concern for surface waters with respect to effects of acidic deposition that are linked to effects on aquatic biota: acute (ANC $< 0\ \mu eq\ l^{-1}$), representing chronically acidic conditions; severe (ANC 0–20 $\mu eq\ l^{-1}$), waters that experience acidic conditions and acidic episodes; elevated (ANC 20–50 $\mu eq\ l^{-1}$), waters that exhibit episodic acidification to low ANC values; moderate (ANC 50–100 $\mu eq\ l^{-1}$), waters that are moderately sensitive to episodic acidification; and low (ANC $> 100\ \mu eq\ l^{-1}$), waters with little sensitivity to acidic deposition (USEPA, 2009; Figure 7).

Mercury Contamination

The deposition of mercury onto the landscape is just one step in the complex suite of transformations of mercury from an inert element in the Earth's crust to a harmful contaminant in fish and wildlife (**Figure 8**). Most mercury entering the environment

Table 2 Biological effects of surface water acidification in North America

pH decrease	*General biological effects*
6.5–6.0	Small decrease in species richness of phytoplankton, zooplankton, and benthic invertebrate communities resulting from the loss of a few highly acid-sensitive species, but no measurable change in total community abundance or production. Some adverse effects (decreased reproductive success) may occur for highly acid-sensitive species (e.g., fathead minnow, striped bass).
6.0–5.5	Loss of sensitive species of minnow and dace, such as blacknose dace and fathead minnow; in some waters decreased reproductive success of lake trout and walleye, which are important sport fish species in some areas. Visual accumulations of filamentous green algae in the littoral zone of many lakes, in some streams. Distinct decrease in the species richness and change in species composition of the phytoplankton, zooplankton, and benthic invertebrate communities, although little if any change in total community biomass or production.
5.5–5.0	Loss of several important sport fish species, including lake trout, walleye, rainbow trout, and smallmouth bass; as well as additional nongame species such as creek chub. Further increase in the extent and abundance of filamentous green algae in lake littoral areas and streams. Continued shift in the species composition and decline in species richness of the phytoplankton, periphyton, zooplankton, and benthic invertebrate communities; decrease in the total abundance and biomass of benthic invertebrates and zooplankton may occur in some waters. Loss of several additional invertebrate species common in oligotrophic waters, including *Daphnia galeata mendotae, Diaphanosoma leuchtenbergianum, Asplanchna priodonta*; all snails, most species of clams, and many species of mayflies, stoneflies, and other benthic invertebrates. Inhibition of nitrification.
5.0–4.5	Loss of most fish species, including most important sport fish species such as brook trout and Atlantic salmon; few fish species able to survive and reproduce below pH 4.5 (e.g., central mudminnow, yellow perch, and in some waters, largemouth bass). Measurable decline in the whole-system rates of decomposition of some forms of organic matter, potentially resulting in decreased rates of nutrient cycling. Substantial decrease in the number of species of zooplankton and benthic invertebrates and further decline in the species richness of the phytoplankton and periphyton communities; measurable decrease in the total community biomass of zooplankton and benthic invertebrates in most waters. Loss of zooplankton species such as *Tropocyclops prasinus mexicanus, Leptodora kindtii,* and *Conochilis unicornis*; and benthic invertebrate species, including all clams and many insects and crustaceans. Reproductive failure of some acid-sensitive species of amphibians such as spotted salamanders, Jefferson salamanders, and the leopard frog.

Source: This table was previously published in Baker JP, Bernard DP, Christensen SW, and Sale MJ (1990) Biological effects of changes in surface water acid–base chemistry, Report SOS/T 13 for the National Acid Precipitation Assessment Program, Washington, DC.

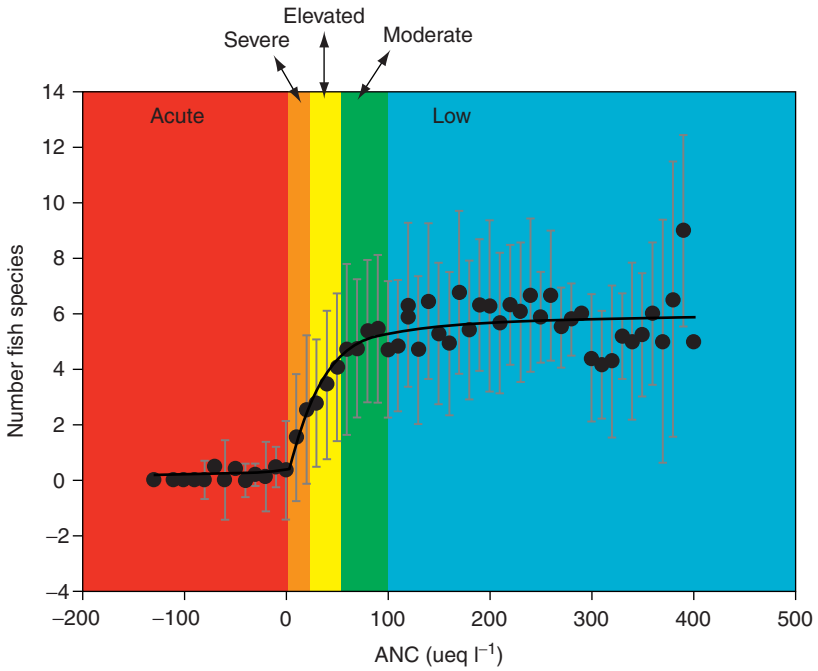


Figure 7 The mean number of fish species for lakes with different values of acid neutralizing capacity in the Adirondack region of New York. Also shown are the various sensitivity classes for aquatic biota.

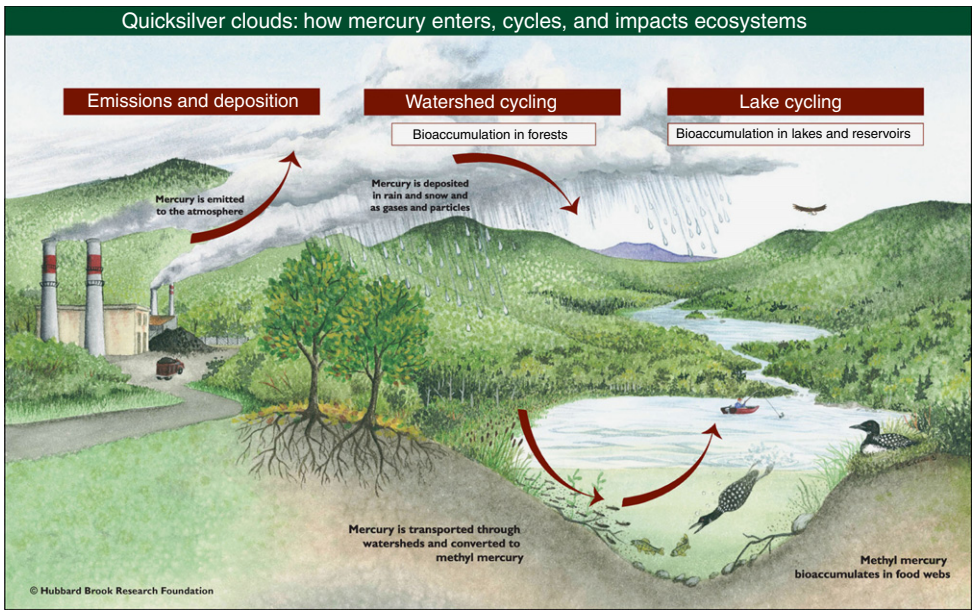


Figure 8 Conceptual diagram of atmospheric mercury deposition, transport, and transformations through a watershed-lake ecosystem, bioaccumulation along the food chain and exposure to humans and wildlife (Driscoll *et al.*, 2007b). Courtesy of the Hubbard Brook Research Foundation.

from atmospheric deposition occurs as inorganic mercury that poses no direct health risk. However, as this relatively benign form of mercury is transported with water through the watershed, it comes into contact with bacteria that process sulfate or iron in wetlands and lake or river sediments, and can be converted to the more bioavailable methyl mercury form. Relatively low concentrations of methyl mercury then bioaccumulate through the aquatic food chain and can reach very high concentrations in fish (**Figure 9**). Data show that methyl mercury concentrations increase nearly 10 million times as it is transferred from water to algae, invertebrates, small fish, larger fish, and to birds and other wildlife (**Figure 10**).

Much of the mercury entering the environment from atmospheric deposition is retained in soil and sediments. This

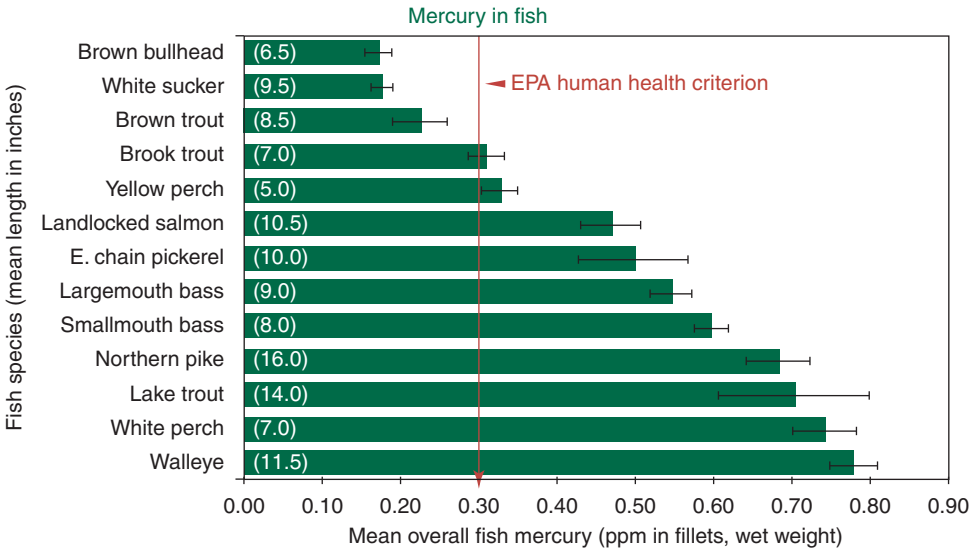


Figure 9 Mercury measurements in 13 of the most abundant freshwater fish species in eastern North America. Note 10 of 13 species exceed the EPA human health criterion of 0.3 μg g^{-1} (w/w; Driscoll *et al.*, 2007b). Courtesy of the Hubbard Brook Research Foundation.

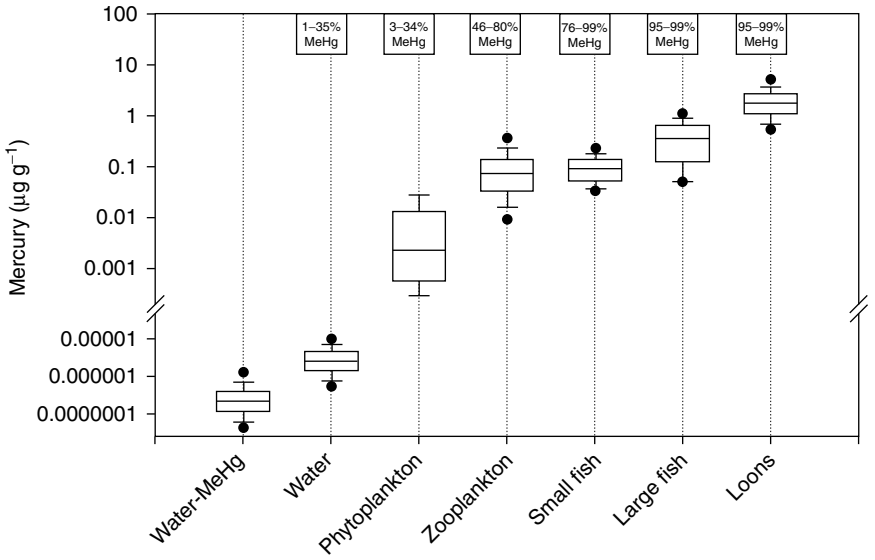


Figure 10 Concentrations and ranges of mercury in water and aquatic biota in eastern North America. Also shown are the percentages of total mercury that occurs as methyl mercury.

mercury represents a long-term legacy of atmospheric emissions and deposition. The extent to which this mercury is transported from soil to downstream surface waters is unclear, although a fraction is transported particularly during high flow events. Nevertheless, the mobilization of this legacy mercury would delay the recovery of contaminated ecosystems after the implementation of controls on mercury emissions.

Sensitivity to atmospheric mercury deposition refers to two processes that occur after mercury is deposited to the landscape: methyl mercury production and methyl mercury bioaccumulation (or transfer along the food chain). The rate at which ionic mercury is converted to methyl mercury and supplied to surface waters varies from watershed to watershed. The following are important watershed characteristics that influence methyl mercury production and transport: shallow hydrologic flowpaths, abundant shoreline wetlands, small watershed to lake area ratio, land cover (e.g., substantial wetlands and forestlands), and inputs of sulfate such as from acidic deposition (Driscoll *et al.*, 2007a). Where these conditions prevail, methyl mercury production and supply to surface waters are expected to be high.

The second component of sensitivity is "surface water sensitivity", or the ability of a lake or rivers to bioaccumulate methyl mercury to high levels in fish and wildlife. Several key factors controlling surface water sensitivity are: nutrient levels, food chain length, and surface water acidity.

Exposure to mercury largely occurs through consumption of fish or other food sources that contain elevated

concentrations of mercury (general methyl mercury; Driscoll *et al.*, 2007a). Mercury contamination is widespread. In the US, all 50 states have mercury advisories. These advisories cover 13 million lake acres and 764,000 river miles (Driscoll *et al.*, 2007a).

It has been estimated that between 300,000 and 400,000 children are born each year in the US at risk for impairment due to mercury poisoning as the result of fish consumption by the mother (Mahaffey *et al.*, 2004). Six percent of US women of childbearing age have blood mercury in excess of values deemed safe by the US Environmental Protection Agency (Schober *et al.*, 2003). Children exposed to high concentrations of methyl mercury can experience decreased ability to perform in school, declines in visual and spatial functions, delays in ability to recall and process information, and a general diminishment in intelligence. Several studies have also pointed to a connection between mercury exposure and cardiovascular disease in men. Elevated mercury in fish has been linked to decreased spawning success. In fish-eating wildlife, mercury exposure is associated with reduced reproductive success, impaired motor skills and death (Driscoll *et al.*, 2007a).

Recently, scientists have also shown that methyl mercury accumulates along the terrestrial food chain. Elevated concentrations have been observed in song birds in habitats at high elevations and in wetlands (Rimmer *et al.*, 2005)

Linkages among Acidic Deposition, Surface Water Acidification and Fish Mercury Concentrations

Several research studies indicate a linkage between acidic deposition and mercury concentrations in fish. Atmospheric deposition of sulfate associated with sulfur dioxide emissions provides a necessary substrate for methylating bacteria. Methylating bacteria convert ionic mercury into methyl mercury, the form of mercury, which bioaccumulates in fish and other wildlife.

Researchers have shown that experimental addition of sulfate to wetlands increase methylation and increase export of methyl mercury (Branfireun *et al.*, 1999). These studies infer that increasing sulfur dioxide emissions and sulfate deposition would result in increases in methyl mercury in the fish of receiving waters, or by analogy decreases in sulfate inputs would decrease fish mercury. In this regard, Drevnick *et al.* (2007) showed that decreases in fish mercury concentrations in Isle Royal in Lake Superior were coupled with historical decreases in atmospheric sulfate deposition.

Studies across eastern North America have shown that many lakes and reservoirs have elevated concentrations of mercury in fish tissue (Driscoll *et al.*, 2007a; Kamman *et al.*, 2005). Many of these surface waters have fish mercury content above the US Environmental Protection Agency health advisory action level of 0.3 $\mu g\, g^{-1}$ (**Figure 9**). In addition, numerous studies across eastern North America have reported increases in fish mercury concentrations with decreases in surface water pH (Driscoll *et al.*, 2007a).

Hrabik and Watras (2002) used observations from experimentally acidified Little Rock Lake, Wisconsin, and reference data to examine the relative contribution of atmospheric mercury deposition and acidic deposition in regulating changes in fish mercury concentrations. They observed that decreases in fish mercury in an experimentally deacidified basin exceeded those in the reference basin. Specifically, they found that approximately one-half of the changes in fish mercury concentrations over a six-year period could be attributed to deacidification (Hrabik and Watras, 2002). This study suggests that acidification of lakes by acidic deposition has enhanced fish mercury concentrations and that concentrations of mercury in fish may decrease in response to decreasing acidic deposition.

See also: Air Pollution. Biogeochemical Cycles. Freshwater Ecosystems, Human Impact on. Human Impacts on Ecosystems: An Overview. Nitrogen Deposition and Terrestrial Biodiversity. Nitrogen, Nitrogen Cycle

References

Aber JD, Nadelhoffer KJ, Steudler P, and Melillo JM (1989) Nitrogen saturation in northern forest ecosystems. *BioScience* 39: 378–386.

AMAP/UNEP (2008) Technical Background Report to the Global Atmospheric Mercury Assessment. AMAP/United Nations Environment Programme (UNEP) Chemical Branch.

Branfireun BA, Roulet NT, Kelly CA, and Rudd JWM (1999) In situ sulphate stimulation of mercury methylation in a boreal peatland: Toward a link between acid rain and methylmercury contamination in remote environments. *Global Biogeochemical Cycles* 13: 743–750.

Butler TJ, Likens GE, Vermeylen FM, and Stunder BJB (2003) The relation between NO_x emissions and precipitation NO_3^- in the eastern USA. *Atmospheric Environment* 37: 2093–2104.

DeHayes DH, Schaberg PG, Hawley GJ, and Strimbeck GR (1999) Acid rain impacts on calcium nutrition and forest health. *BioScience* 49: 789–800.

Drevnick PE, Canfield DE, Gorski PR, *et al.* (2007) Deposition and cycling of sulfur controls mercury accumulation in Isle Royale fish. *Environmental Science and Technology* 41: 7266–7272.

Driscoll CT, Evers D, Lambert KF, *et al.* (2007c) *Mercury Matters: Linking Mercury Science with Public Policy in the Northeastern United States*, vol. 1. Hubbard Brook Research Foundation. Science Links™ Publication, no. 3.

Driscoll CT, Han Y-J, Chen CY, *et al.* (2007a) Mercury contamination in forest and freshwater ecosystems in the Northeastern United States. *BioScience* 57: 17–28.

Driscoll CT, Lawrence GB, Bulger AJ, *et al.* (2001a) Acidic deposition in the northeastern United States: Sources and inputs, ecosystem effects, and management strategies. *BioScience* 51: 180–198.

Driscoll CT, Lawrence GB, Bulger AJ, et al. (2001b) *Acid Rain Revisited: Advances in Scientific Understanding Since the Passage of the 1970 and 1990 Clean Air Act Amendments*. Science Links™ Publication, vol. 1, no. 1. Hanover, NH: Hubbard Brook Research Foundation.

Driscoll CT, Whitall D, Aber J, *et al.* (2003) Nitrogen pollution in the northeastern United States: Sources, effects, and management options. *BioScience* 53: 357–374.

Eilers JM and Selle AR (1991) Geographic overview of the regional case study areas. In: Charles DF (ed.) *Acidic Deposition and Aquatic Ecosystems*. New York: Springer-Verlag.

Evans CD, Cullen JM, Alewell C, *et al.* (2001) Recovery from acidification in European surface waters. *Hydrology and Earth System Science* 5: 283–298.

Fenn ME, Baron JS, Allen EB, *et al.* (2003) Ecological effects of nitrogen deposition in the Western United States. *BioScience* 53: 404–420.

Galloway JN, Norton SA, and Church MR (1983) Freshwater acidification from atmospheric deposition of sulfuric acid: A conceptual model. *Environmental Science and Technology* 17: 541A–545A.

Horsley SB, Long RP, Bailey SW, and Hall TJ (2000) Factors associated with the decline disease of sugar maple on the Allegheny Plateau. *Canadian Journal of Forest Research* 30: 1365–1378.

Hrabik TR and Watras CJ (2002) Recent declines in mercury concentration in a freshwater fishery: Isolating the effects of de-acidification and decreased atmospheric mercury deposition in Little Rock Lake. *The Science of the Total Environment* 297: 229–237.

Kahl JS, Stoddard JL, Haeuber R, *et al.* (2004) Have U.S. surface waters responded to the 1990 Clean Air Act Amendments? *Environmental Science and Technology* 38: 485A–490A.

Kamman NC, Burgess NM, Driscoll CT, *et al.* (2005) Mercury in freshwater fish of northeast North America – a geographic perspective based on fish tissue monitoring databases. *Ecotoxicology* 14: 163–180.

Likens GE, Driscoll CT, and Buso DC (1996) Long-term effects of acid rain: Response and recovery of a forest ecosystem. *Science* 272: 244–246.

Mahaffey KR, Clickner RP, and Bodurow CC (2004) Blood organic mercury and dietary mercury intake: National health and nutrition examination survey, 1999 and 2000. *Environmental Health Perspectives* 112: 562–570.

Mason RP and Sheu G-R (2002) Role of the ocean in the global mercury cycle. *Global Biogeochemical Cycles* 16: 1093.

Rimmer CC, McFarland KP, Evers DC, *et al.* (2005) Mercury concentrations in Bicknell's Thrush and other insectivorous passerines in montane forest of northeastern North America. *Ecotoxicology* 14: 223–240.

Schober SE, Sinks TH, Jones RL, *et al.* (2003) Blood mercury levels in US children and women of childbearing age, 1999-2000. *Journal of the American Medical Association* 289: 1667–1674.

Thomas RQ, Canham CD, Weathers KC, and Goodale CL (2010) Increased tree carbon storage in response to nitrogen deposition in the US. *Nature Geosciences* 3: 13–17.

USEPA (2009) Risk and Exposure Assessment for Review of the Secondary National Ambient Air Quality Standards for Oxides of Nitrogen and Oxides of Sulfur. EPA-452/P-09-004a.

Adaptation

Michael R Rose and Molly K Burke, University of California, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Michael R. Rose, volume 1, pp 17–23, © 2001, Elsevier Inc.

Glossary

Adaptation *One or two of the following:* a beneficial construct produced by an omnipotent being; the process of change established by natural selection; and a biological character that gives increased Darwinian fitness.

Adaptationism The doctrine that all important evolutionary processes are dominated by natural selection, and that all significant biological characters increase an organism's fitness.

Epistasis Interactions between genes at different chromosomal locations in the determination of phenotypic character values.

Fitness Net reproductive output, discounted for any lack of viability.

Genetic drift Accidents of segregation and recombination causing evolutionary genetic change.

Genomics The study of the entire volume of an organism's hereditary information. Methods in population genomics tend to take one of two forms: (1) surveys of global gene expression in which a subset of the population's RNA is sequenced or otherwise quantified; or (2) surveys of allele frequency in which a sample of the population's DNA is sequenced or otherwise quantified.

Group selection Selection between different populations or sub-populations based on attributes of the entire group, where these attributes usually are either selected against or not favored at the level of individual selection.

Inbreeding The mating of close biological relatives.

Individual selection Selection driven by differences in the net reproduction of individual organisms.

Segregation Allocation of genetic variants ("alleles") to different gametes during sexual reproduction.

Teleology The imputation of goal-directed behavior or structures.

Historical Introduction

Classical Times

The concept of adaptation is older than any scientific concept of evolution, and certainly older than Darwin's theory of natural selection. The founder of academic biology, Aristotle, gave adaptive explanations for many of the features of the living, as well as nonliving, world. Thus the webbed feet of a frog can be said to be "for" efficient swimming, and thus can be explained as an illustration of the universe being well-made. This type of reasoning was commonplace in classical culture, which often assumed some type of benign natural order.

It is significant, however, that adaptive reasoning had its critics even in classical times. Lucretius, one of the most important of classical proto-scientists, was scathing about the wholesale imputation of function to body parts, even when such inferences were regarded as "common sense." Such criticisms of the concept of adaptation have waxed and waned ever since.

Pre-Darwinian Christendom

Biblical theology gave arguments about adaptation a new cast. The assumption that there was a single, benign, omnipotent Creator made the existence of well-constructed organisms a natural inference. From the beneficence of the Creator, each organism must have been given the specific characteristics best suited to its role in the Creation as a whole. Indeed, this concept of benign, and efficient, creation was extended to physics, especially by Isaac Newton, an avid believer. The orbits of the planets themselves were thereby interpreted as evidence of some type of adaptation, or suitedness, to a divine plan.

This universal adaptation gave rise to some interesting paradoxes for pre-Darwinian scientists. Did the Creator adapt organisms to the physical universe, or was the physical universe created to fit the organisms? Was life on earth based primarily on the chemistry of water and organic molecules because that was the biochemistry that could work on this particular planet? Or was the planet constructed by the Creator to fit the biochemistry that He already had in mind? How could such questions ever be resolved?

The Darwinian Theory of Adaptation

Darwin's theory of evolution by natural selection provided a natural solution to the two problems of what adaptations were and how they came about. On Darwin's theory, selection operating on heritable variation increased the frequency of individuals bearing attributes, which we can call "adaptations," which give them increased fitness. Fitness in turn was to be defined as the net reproductive rate of individual organisms, in the original version of Darwinism. Thus Darwin's theory proposed that the environment drove the evolution of adaptations, by determining the pattern of selection imposed on organisms.

Darwin's theory had a number of novel features for the biology of his time. First, it involved no omnipotent Creator,

beneficently organizing the arrangements of life. Second, there were no inner drives or teleologies shaping the process of organic change to an adaptive end, unlike the scheme of Lamarck and others, who were more influenced by Aristotle than Darwin was. Third, there was no overarching pattern to Darwin's process of evolution, and therefore adaptations might arise higgledy–piggledy, whenever selection made mortality or reproduction hinge on a particular attribute. All of these features made Darwin's new biology of adaptation distasteful to many of the older generation of Victorian biologists, who were highly teleological in their thinking, when not avidly creationist.

Two Common Definitions

The nature of Darwinian theory instilled a large degree of ambiguity in the term adaptation itself. For a creationist, there is no process of adaptation, only the fact of what the Creator made, and its beneficent nature. For the Darwinian, like the Lamarckian, there is necessarily a *process* of adaptation, a process by which adaptation is brought about. Then there is the *product* of adaptation as a process, which is called an adaptation, as well.

There has been some controversy as to which of these basic meanings of the term is really the true one, or the correct one. However, we can follow Ernst Mayr, or Karl Popper for that matter, and reject the need to find any essentially true definition of adaptation. Instead, the two alternative definitions can both be used, as appropriate.

Adaptation as a Process

The concept of adaptation as a process derives from the theory of natural selection. Therefore, a deeper consideration of this incarnation of adaptation requires study of natural selection. One of the basic intuitive expectations that most evolutionary biologists have is that natural selection should lead to the evolution of increased Darwinian fitness. However, this is not universally true. Mutation, segregation, recombination, meiotic drive, and frequency-dependent selection can force natural selection to produce a *decrease* in fitness under some conditions.

Usually this decrease in fitness is temporary. For example, if the heterozygote, at a locus with two alleles, is the most fit genotype, then, when the population is initially composed entirely of heterozygotes, segregation will cause an immediate decrease in fitness. But this effect is confined to the first generation. Subsequent generations will see increasing or stable mean fitness as natural selection brings the population to the stable gene frequency equilibrium, at which mean fitness will also be at a local maximum. Analogous processes can occur with other genetic processes, such as recombination. Again, in some of these cases, mean fitness does not continue to decline.

But there are also cases where fitness may continue to fall, and natural selection never produces a recovery in mean fitness. *Meiotic drive* is a well-known example of this. Meiotic drive occurs when some genes pervert segregation rates in their favor, so that, despite being deleterious, they spread through populations. Another situation in which mean fitness can fall is brought about by natural selection itself. The fitness of a mating with males and females can be a nonlinear function of the genotypes of the two organisms mating. If such nonlinearities are sufficiently severe, natural selection on fertility can actually drive fitness to lower and lower levels, at least in theory. No prominent empirical examples of this process are known, however, at this time.

The point that these examples serve to make is that, while it is conventional in evolutionary biology to expect improved adaptation from the action of natural selection, there is no absolute warrant, either in theory or in fact, for this assumption. Theory and experiment both indicate that a process of adaptation is usually brought about by natural selection. But it is not always brought about by natural selection.

When natural selection does act, however, to establish a process of adaptation, what can we say about that process? This is one of the three major research projects of evolutionary biology (the other two being the inference of phylogeny and the study of the genetic material used by evolution). Thus our understanding of adaptation as a process is undergoing continual upgrading as our understanding of natural selection improves.

At the most basic level, however, there are some essential features of adaptation by natural selection that can be taken as well-established. There is no general or consistent pattern to natural selection. Specific populations may undergo very intense selection for a short period of time. One of the best studied examples of this is the recent work on the evolution of Darwin's finches on the Galapagos, particularly the effects of drought on bill size (Grant, 1986). However, for most populations, it is usually very difficult to detect the action of natural selection. It is either too weak or too variable in direction (Abrahamson and Weis, 1997). Some of the cases where natural selection can be readily detected as working in each generation to produce adaptation involve human disturbances, ecological events that are unlikely to reflect the evolutionary situation of populations that have been spared artificial disruption. The classic example of this scenario is the evolution of wing camouflage in the moths of industrial Europe, in which natural selection was generated because soot blackened the tree trunks on which these moths rested, making the light-colored moths stand out against a black background. The very artificiality of this case, however, underscores the point that we do not normally find such cases of unequivocal selection when we study natural populations.

This leads to the next major point about adaptation as a process: it is hard to detect. Therefore, adaptation as a process tends to be assumed by evolutionary biologists more than it is actually demonstrated. And the teasing out of the mechanistic particulars of adaptation as a process is almost never accomplished. This central problem has led to a pervasive weakness in the scientific analysis of adaptation as a process, with unfortunate consequences that will be reviewed below.

Adaptation as a Product of Evolution

The view of adaptation as a product of evolution does not logically require that it be a product of natural selection.

An adaptation can arise evolutionarily from selection on some other character(s), or it might arise from some nonselective process, such as inbreeding or genetic drift. Thus, for example, a spider's web might have evolved because of selection for prey capture, yet it may also constitute an adaptation that enables spiders to obtain water from dew condensing on the web. This raises the question, if adaptation is divorced from the process by which it arose, then how is it to be distinguished from the other characteristics of an organism?

The conventional solution to this problem is to define adaptations as those products of evolution, however generated, that enhance the fitness of the organism. Nominally, this requires that fitness be measured with and without the character(s) that are presumed adaptations. This is a difficult enterprise for two reasons.

First, it is often difficult to perform the surgery, or other manipulation, required to make organisms without the adaptation in question. However, it is in precisely this area where significant progress has been made in studies of adaptation (e.g., Sinervo and Basolo, in Rose and Lauder, 1996). Evolutionary biologists are now successfully ablating tissues and grafting on additional body parts, in order to test the fitness consequences of the possession or loss of particular structures that are being evaluated for their status as adaptations. Manipulation of clutch size by removal or addition of eggs has been a traditional method in studies of vertebrate life-history adaptations. Research in this area now manipulates fertility and egg size by a variety of techniques, including microsurgery. The resources of modern molecular biology are likely to give evolutionary research even more power to manipulate phenotypes.

Second, the measurement of fitness is difficult in most organisms. In organisms that reproduce strictly by dividing in two, without sex, fitness can be measured fairly easily from estimates of viability between bouts of fission. In every other kind of organism, sex and variable numbers of offspring make the estimation of fitness extremely difficult. Perhaps the most difficult character of all for the estimation of fitness is male mating success. This difficulty arises because the attribution of maternity is usually fairly secure, while the attribution of paternity is often pure speculation. This is an area where the recent findings of behavioral ecology suggest considerable grounds for pessimism. Pairs of birds, for example, may indeed remain together for life, sharing the tasks of caring for young, foraging for food, and nest construction. But molecular genetic analysis of pedigrees frequently reveals that the "monogamous" female has had sex with another male of the species, while the male has himself dallied. Similar patterns are well-known from human paternity cases. Then there are the species that are either highly promiscuous, such as chimpanzees, or ejaculate gametes externally, such as most fish. In these species, there are no mated pairs to keep track of over the long-term. For all these reasons, estimating the Darwinian fitness of an individual having a particular phenotype is often extremely difficult, if not practically impossible.

The fallback position of many biologists, especially functional morphologists, comparative physiologists, and behavioral ecologists, has been to use a surrogate for fitness. Such surrogates might include mechanical efficiency, conservation of metabolic energy, number of copulations, and so on. The assumption is usually made that such surrogate measures will always be positively correlated with fitness. When they improve, fitness should increase. Unfortunately, it is likely to be precisely characters of this kind that will show diminishing returns, rather than a stable positive correlation with fitness. Mechanical efficiency is patently not the only impact of structure on fitness. Structures may be costly to develop, or they may impede movement. Evolution is unlikely to maximize each and every "design feature" of an organism, even if there were no genetic constraints preventing the realization of any particular phenotype. Therefore the expedient of using surrogates for fitness is not likely to be reliable in many cases.

If fitness itself cannot be accurately measured, and surrogates for fitness cannot be relied on either, it is difficult to see how the concept of adaptation as a product of evolution can be used in most cases. There are pleas to the effect that some characters are so intuitively beneficial that they cannot be reasonably denied the status of adaptations. Legs must be adaptations for terrestrial locomotion, large brains must be adaptations for life as a tool user, and so on. However, limbs may be used for many functions, not just locomotion. The human brain has also been explained as an adaptation for social behavior, not the use of tools. Supposedly obvious cases turn out to be far from obvious once a diversity of possible scientific interpretations is taken into account.

Evidence for Adaptation

If both basic definitions of adaptation are allowed, then there are two different lines of evidence for the existence of adaptation. The first is simply the action of natural selection. If adaptation is the process of natural selection, then any evidence for such selection is in turn evidence for adaptation. The second line of evidence is supplied whenever there are data showing an increase in fitness when a particular character is acquired. Together, the accumulated evidence bearing on both of these points helps establish the importance of adaptation as a feature and as an outcome of the evolutionary process.

Evidence for Natural Selection

If Darwin generally lacked evidence for natural selection in nature, modern evolutionary biology has supplied an abundance of such evidence (e.g., Endler, 1986). There are the classic studies of industrial melanism and there are the recent studies of drought selection in Darwin's finches (Grant, 1986). But there are many other examples of natural selection in the wild, dating back to Weldon's study of carapace width in estuarine crabs during the 1890s. Indeed, natural selection is such an obvious feature of the living world that it is now considered in discussions of such practical medical problems as the prescription of antibiotics and the treatment of the human immunodeficiency virus. Thus the general principle that there is a process of adaptation involving natural selection is not in any reasonable doubt.

The evidential problems instead concern the importance of the process in any particular instance. The idea of an adaptive process shaping the course of evolution is very attractive,

because it can be used to support the interpretation of evolutionary change in terms of natural selection. But as discussed already, the demonstration that such a process is occurring is usually very difficult. And the possibility that other evolutionary processes are involved, processes that do not involve adaptation by natural selection for the character of interest, cannot be dismissed out of hand. This renders most casual *post hoc* invocations of natural selection essentially dubious. Whatever the specific features of natural selection, casually invoking it as an explanation for all features of life is no longer reputable behavior in evolutionary biology.

This means that, while there are some specific studies that provide excellent evidence for adaptation by natural selection, in most cases scientists are not in a position to interpret an evolutionary process as being driven by natural selection. It may be allowed as a possibility, but further study is usually required before a particular evolutionary change can be taken as brought about by natural selection, even when such an interpretation seems intuitively natural.

Evidence for Increased Fitness

Even if it is difficult to establish the nature of the evolutionary process, surely the products of evolution are easier to categorize as adaptive? For the reasons listed above, however, it is often difficult to make an accurate determination concerning whether or not the possession of a particular character increases fitness. In particular, it is not enough to show that a particular function, say locomotion, has been improved, perhaps by a longer hind-limb, because such demonstrations do not define the effect on fitness as a whole. A particular function could be improved while fitness itself is reduced.

At present, some of the best demonstrations of adaptation come from the field of behavioral ecology. Of particular value have been manipulative experiments which change the behavior or morphology of study animals and plants. These studies have supplied many instances in which artificially created deviants have demonstrably reduced fitness (Sinervo and Basolo, in Rose and Lauder, 1996).

There is a great deal of potential for the study of molecular biology to extend the power of manipulation in the study of adaptation, particularly with genetic transformation and the insertion of genes with artificially inducible expression. Some of these studies have measured the effects on adult survival of gene insertions (Fleming and Rose, 1996). Fitness could also be measured in such experiments.

An alternative approach is to measure the relationship between the variation of a character and fitness in polymorphic populations. Much of modern evolutionary quantitative genetics collects data of this kind. One of the central concerns in these studies is the delimitation of optima for fitness as a function of quantitative characters.

Finally, natural selection can be imposed on populations in the laboratory, which results in perturbed values for selectable characters. If the selection is then relaxed, and the original character-state was adaptive, then the character should be driven back to its original state. This result has been observed in several studies (e.g., Service *et al.*, 1988; Teotonio and Rose, 2000 and Passananti *et al.*, 2004). Such patterns of reversion are expected to occur especially when there are tradeoffs between functional characters, such that high values of one character are associated with low values of other characters. Interestingly, these kinds of phenotypic reversions are not always accompanied by complete genotypic reversions (Teotonio *et al.*, 2009).

The Genomics of Adaptation

Genomics, a discipline still in its adolescence, has revolutionized the study of adaptation over the past decade. The tools of genomics, including genome-wide surveys of gene expression and allele frequencies, have developed into powerful means of describing differences, at the molecular level, between individuals with or without an inferred adaptive character. While these tools are sometimes used to describe natural populations, such as those situated along latitudinal gradients (e.g., Turner *et al.*, 2008), they are perhaps more useful when applied to populations evolved in the laboratory. Genomic comparisons of experimentally evolved populations to ancestral controls have usually employ microbial models with small genomes, such as bacteria (i.e., Barrick *et al.*, 2009) or yeast (i.e., Dunham *et al.*, 2002).

Recent work by Burke *et al.* (2010) offered the first genome-wide survey of allele frequencies in experimentally evolved *Drosophila* populations; in other words, they associated particular allelic variants with characters that underwent dramatic and reproducible change as a result of laboratory selection. The study used populations of fruit flies that had been selected for very rapid development for over 600 generations, which resulted in individuals with enhanced early-life fitness components, compared to individuals from control populations. Burke and colleagues identified many genomic regions that show strong allele-frequency differentiation in response to selection, and these regions included over 500 individual genes. But perhaps the real result of interest is that Burke and colleagues found no evidence for the complete fixation of any of the thousands of nucleotide polymorphisms that strongly responded to selection. In asexually reproducing systems, adaptation is thought to be driven by the sequential fixation of newly arising, beneficial mutations (e.g., Barrick *et al.*, 2009). Burke *et al.* (2010) failed to observe evidence of this phenomenon. Instead they found widespread heterozygosity across the genome, even at loci highly differentiated in terms of allele frequency. This raises important questions about the nature of adaptation, especially whether it operates by the same mechanisms in asexual versus sexual organisms. These questions will undoubtedly be addressed in the years to come, as the technology for exploring whole genomes becomes ever more informative and cost-effective.

Critique of Adaptationism

Adaptationism

The record of classical thinking, such as that of Aristotle, illustrates the extent to which the human mind is attracted to the idea of beneficent organization in the natural world. Many

classical scholars believed this even when they had no particular scientific theories to buttress the concept.

Darwinian evolution is thus an almost irresistible temptation for those who wish to infer function in the living world. Darwinism guarantees a role for natural selection in evolution, and it guarantees the existence of adaptation among the characters of organisms. But it does not guarantee that selection and adaptation must be everywhere prepotent, at all times, and in all respects.

Nonetheless, there is a variant form of Darwinism that flourished particularly in the 1950s and 1960s, a variant that assumed that all of the attributes of an organism are shaped by natural selection to the end of increased fitness. On this version of Darwinism, now called "adaptationism," all characters are adaptations and all nontrivial evolutionary processes are driven by natural selection. In effect, this school of thought made the study of evolution tantamount to the study of adaptation.

Among the effects of adaptationism on scientific practice was the notion that there must always be an adaptive explanation for each and every organ, structure, or behavior. Therefore, if an adaptive explanation for a particular structure has not been found, greater efforts must be made to discover its adaptive value. Alternative evolutionary processes (genetic drift, inbreeding, meiotic drive, etc.) must not be considered until all possible adaptive explanations have been tried and found wanting.

During its heyday, adaptationism put adaptation at the center of evolutionary biology, and to some extent at the center of all of biology. The many theoretical and experimental problems facing the study of adaptation were minimized or dismissed altogether.

The Rejection of Adaptationism

From the late 1960s until the early 1980s, adaptationism suffered a series of blows from which it has yet to recover. The first of these was the detection of a vast amount of molecular genetic variation, first by protein electrophoresis and later by DNA sequencing. The significance of this finding for adaptationism is that most species appear to have far more segregating genetic variation than is likely to be explicable in terms of natural selection. Therefore, natural selection probably isn't prepotent at the molecular level. The scientific consensus now is that many of the alleles that arise and eventually become fixed during evolution are merely neutral variants of already extant alleles. A great deal of genetic evolution has occurred, but much of it has not been driven by natural selection.

A second event was the publication of *Adaptation and Natural Selection* by Williams (1966). One of the common evasions of the adaptationists was to invoke group selection when they couldn't explain a particular character in terms of individual selection. Thus many of the social behaviors of colonially nesting birds were explained in terms of adaptations for group selection. Williams pointed out that many of these explanations were highly dubious. He argued that the inference or explanation of adaptations required greater restraint, particularly where social behavior was concerned. This undercut group selection, one of the ways in which adaptationists had been able to discover adaptations underlying seemingly maladaptive behavior, such as biological altruism. In so doing Williams also helped expose the extent to which adaptationism was based more on dogma than well-founded science.

The third, and culminating, event in the decline of adaptationism was the publication of the "Spandrels of San Marco" paper by Gould and Lewontin (1979). In this paper, Gould and Lewontin hold up for ridicule the adaptationist assumption that there has been a history of selection directly for each and every significant attribute of an organism. They follow Voltaire in his satirizing of such intellectual figures as Spinoza, particularly their boundless belief that "this is the best of all possible worlds," except that Gould and Lewontin satirize the adaptationist assumption of an all-powerful beneficent natural selection.

These events essentially undermined adaptationism as a dominant movement within evolutionary biology. Adaptationists remain scattered through biology, including such fields as molecular biology, comparative physiology, and systematics. But the powerful hold that they had on evolutionary biology in the 1950s was broken.

Adaptation after Adaptationism

Though adaptationism was clearly in error about the universality of adaptation in the living world, its demolition brought with it something of an overreaction. Many evolutionary biologists effectively rejected the concept of adaptation as a whole. They refused to work on the problem and they criticized those who did. Since 1980, the study of phylogeny has become the central concern of evolutionary biology. The study of adaptation, for some, is now a marginal, somewhat disgraceful, practice within biology as a whole.

Replacing the study of adaptation was the study of "constraints." Constraints in evolutionary biology are factors that prevent the achievement of an optimal adaptive outcome. Constraints have been discovered promiscuously in the evolutionary machinery: lack of genetic variation, linkage disequilibrium, too much environmental variation, too little environmental variation, epistasis, temporally variable selection, spatially variable selection, and so on. Evolutionary biology went from a doctrine where adaptation was everywhere to a doctrine where adaptation had disappeared, to be replaced by paralyzing constraints.

In the 1990s, there was some stabilization of views on the topic of adaptation. Less of a pariah among evolutionary topics, an edited volume entitled "Adaptation" was published in 1996. Evolutionary biologists were spending more time using experimental and other techniques that could test for adaptation, rather than simply assuming its presence or absence.

The comparative study of adaptation was greatly improved by an infusion of phylogenetic techniques. For example, if it was hypothesized that the gill structure of a fish species was an adaptation to a new way of life in salt water, but two species that had evolved in fresh water exclusively also have that gill structure, then the phylogenetic information indicates immediately that the basic adaptive hypothesis is not correct.

Laboratory selection is used more often now to study adaptation, with greater replication and greater attention to

designs that can be used to make inferences about selection. The great advantage of performing selection in laboratories is that selective process can be studied with greater statistical power and control, compared to the "experiments" of nature, all of which are unique and uncontrolled. For example, instead of studying the water physiology of two desert insect species compared to two forest insect species, evolutionary biologists instead select insects under conditions of desiccation, using replicated selection lines and controls maintained free of desiccation (Bradley *et al.*, 1999). Instead of dealing with possible historical accidents that might have differentiated species in the wild, selected laboratory populations provide good material for critically testing theories of adaptation to particular environmental conditions.

Another major development in the study of adaptation has been the use of natural populations, especially manipulated natural populations, in studies that approximate laboratory experiments. Reznick (Reznick and Travis, in Rose and Lauder, 1996) has studied guppy evolution in the streams of Trinidad. Multiple streams pass in parallel through highly uniform drainage systems, giving the streams isolated guppy populations, but populations that evolve under effectively identical conditions. This experimental system has provided tremendous opportunities for the study of adaptation in the wild, with both replication and controls.

The study of adaptation now proceeds with much more skepticism than in the past. At the same time, empirical methods for studying it have been greatly improved. The prospects have never been brighter for a genuine scientific analysis of adaptation, as opposed to the blithe speculations of the past.

References

Abrahamson WG and Weis AE (1997) *Evolutionary Ecology Across Three Trophic Levels: Goldenrods, Gallmakers and Natural Enemies*. Princeton, NJ: Princeton University Press.

Barrick JE, Yu DS, Yoon SH, *et al.* (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461: 1243–1247.

Bradley TJ, Williams AE, and Rose MR (1999) Phyiological responses to selection for desiccation resistance in *Drosophila melanogaster*. *American Zoologist* 39: 337–345.

Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, and Long AD (2010) Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* 467: 587–590.

Dunham MJ, Badrane H, Ferea T, *et al.* (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 99: 16144–16149.

Endler JA (1986) *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.

Fleming JG and Rose MR (1996) Genetics of aging in Drosophila. In: Schneider EL and Rowe JW (eds.) *Handbook of the Biology of Aging*. New York: Academic Press.

Gould SJ and Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society B* 205: 581–598.

Grant PR (1986) *The Ecology and Evolution of Darwin's Finches*. Princeton, NJ: Princeton University Press.

Passananti HB, Deckert-Cruz DJ, Chippindale AK, Le BH, and Rose MR (2004) Reverse evolution of aging in *Drosophila melanogaster*. In: Rose MR, Passananti HB, and Matos M (eds.) *Methuselah Flies: A Case Study in the Evolution of Aging*, pp. 296–322. Singapore: World Scientific.

Rose MR and Lauder GV (eds.) (1996) *Adaptation*. New York: Academic Press.

Service PM, Hutchinson EW, and Rose MR (1988) Multiple genetic mechanisms for the evolution of senescence. *Evolution* 42: 708–716.

Teotónio H, Chelo IM, Bradic M, Rose MR, and Long AD (2009) Experimental evolution reveals natural selection on standing genetic variation. *Nature Genetics* 41: 251–257.

Teotónio H and Rose MR (2000) Variation in the reversibility of evolution. *Nature* 408: 463–466.

Turner TL, Levine MT, Eckert ML, and Begun DJ (2008) Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. *Genetics* 179: 455–473.

Williams GC (1966) *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Adaptive Radiation

Rosemary G Gillespie and George K Roderick, University of California, Berkeley, CA, USA
Francis G Howarth, Bishop Museum, Honolulu, HI, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 25–44, © 2001, Elsevier Inc.

Glossary

Adaptive shift A change in the nature of a trait (morphology, ecology, or behavior) that enhances survival and/or reproduction in an ecological environment different from that originally occupied.

Allopatric speciation The process of genetic divergence between geographically separated populations leading to distinct species.

Character displacement Divergence in a morphological character between two species when their distributions coincide in the same ecological environment compared to overlap of the character in question in the two species when they are geographically separated.

Convergence The evolution of similar characters in genetically unrelated or distantly related species, often as the result of selection in response to similar environmental pressures.

Ecological release Expansion of habitat, or ecological environment, often resulting from release of species from competition.

Founder effect Random genetic sampling in which only a few "founders" derived from a large population initiate a new population. Since these founders carry only a small fraction of the parental population's genetic variability, radically different gene frequencies can become established in the new colony.

Key innovation A trait that increases the efficiency with which a resource is used and can thus allow entry into a new ecological zone.

Natural selection The differential survival and/or reproduction of classes of entities that differ in one or more hereditary characteristics.

Sexual selection Selection that acts directly on mating success through direct competition between members of one sex for mates or through choices made between the two sexes or through a combination of both modes.

Sympatric speciation The process of genetic divergence between populations occupying the same geographic range leading to distinct species.

Taxon cycle The repetitive pattern by which widespread dispersive stage I populations or species give rise to more restricted and specialized stage II populations or species; subsequent divergence leads to stage III local endemics.

History of the Concept

Beginning with the work of Darwin (1859) on the Galapagos fauna, the concept of adaptive radiation, in terms of diversification of ecological roles by means of natural selection, has been recognized. The term was first used by Osborn (1902) in describing parallel adaptations and convergence of species groups on different landmasses. Subsequently, it was developed as a major tenet for arguments presented in the modern synthesis by Huxley (1942). Simpson (1953), working on paleontological data, discussed the importance of key innovations in triggering adaptive radiation. For a detailed history of the concept of adaptive radiation, see Givnish (1997). Much recent information has been added, particularly during the past decade with the rise of molecular methods (Givnish and Sytsma, 1997).

Nonadaptive Radiations

The term "nonadaptive radiation" has been used to describe situations in which species proliferation has not been attended by diversification of ecological roles (Gittenberger, 1991). When proliferation is simply a consequence of isolation, with isolated sibling species maintaining similar ecological affinities, then the radiation cannot be considered "adaptive." As will be described later, isolation has been invoked to explain the initial divergence of taxa in some radiations (e.g., Galapagos finches and cichlid fish), with the adaptive phase not occurring until recently diverged sibling species become sympatric. However, there are some cases of nonadaptive radiation, with many allopatric and ecologically similar species. Most of these radiations are caused by changes in topography that, instead of opening up new habitats, have served simply to isolate a previously more widespread species. For example, isolated mountaintops and other continental refugia have allowed species long periods of evolution in isolation, without any ecological change. This may lead to patterns of considerable genetic distance between morphologically similar species from different isolates (Schneider and Moritz, 1999). Similarly, diversification of snails on islands has frequently been attributed to topographical isolation (e.g., Crete (Gittenberger 1991) and Medeira (Cameron *et al.*, 1996)). In general, it appears that (i) nonadaptive radiation occurs if there is isolation without any novel ecological opportunity and (ii) coexistence of species within a lineage will not occur in nonadaptive radiations but is a primary characteristic of adaptive radiations.

Factors Underlying Adaptive Radiation

The common requirement for triggering adaptive radiation is the opening up of ecological space. This may be allowed by intrinsic factors, i.e., something that changes in the organism to allow radiation to occur; for example, evolution of tolerance toward noxious plant chemicals (Farrell and Mitter, 1994; Mitter *et al.*, 1988). Alternatively, it may occur as a result of extrinsic factors; for example, it has been reported to occur in geological history after an influx of nutrients into the system (Vermeij, 1995), in recent evolutionary time when new islands are colonized (Wagner and Funk, 1995; Liebherr and Polhemus, 1997), and in ecological time when a new habitat opens (Rainey and Travisano, 1998).

For ancient radiations, it is often difficult to determine the relative importance of intrinsic and extrinsic factors in allowing adaptive radiation. Factors associated with such radiations include (i) coincidence (after a slight delay) with major extinction episodes (Sloan *et al.*, 1986) and (ii) radiation of a group frequently starting from a small, unimpressive set of species from an earlier period. For example, fossil ammonites (shelled cephalopod mollusks) reveal episodes of tremendous proliferation and extinction through the Devonian, Triassic, Jurassic, and Cretaceous (**Figure 1**; Lehmann, 1981). Echinoderms show a similar pattern, originating in the Ordovician and undergoing small radiations until all but one lineage went extinct by the end of the Permian (Smith, 1984). These then radiated extensively in the Triassic–early Jurassic, and the current diversity of forms remains similar to what arose at that time.

The great placental radiation (>4300 species) has been attributed partly to the extinction of many reptilian groups at the end of the Cretaceous (Simpson, 1953). The parallel adaptive radiation of marsupials in Gondwana has also been attributed to the Cretaceous extinctions and subsequent opening of ecological space (Springer *et al.*, 1997). However, within each lineage (placentals and marsupials) key innovations may have been involved: The radiation of ungulates and ruminants is associated with the opening up of the savannas (**Figure 2**) but would not have happened if the organisms did not develop the morphological and physiological features necessary to exploit the habitat. Similarly, the radiation of the diprotodontians appears to have commenced in the Eocene and may have been promoted by a key adaptation for herbivory (Springer *et al.*, 1997). The actual basis for radiations subsequent to extinction episodes is still a subject of debate, particularly because coincidence between extinction events and subsequent radiations is generally poor. Vermeij (1995) argued that there is a stronger coincidence of species diversification episodes with increases in nutrient input into the biosphere.

We consider factors underlying species proliferation under two headings: intrinsic factors, and the concept of "key innovations," and extrinsic factors, including environmental change and colonization of isolated landmasses.

Intrinsic Factors: Key Innovations

Simpson (1953) suggested that the evolution of a suite of traits, or key innovations, that increase the efficiency with which a resource is used might allow species to enter a "new" adaptive zone, and the ecological opportunity thus allowed might promote diversification. The concept of the key innovation is an essential element in hypotheses of the evolution of specialization and subsequent adaptive radiation in herbivorous insects. However, the nature of key innovations is not often clear. In an attempt to define more clearly the concept, Berenbaum *et al.* (1996) examined cytochrome P450S and its

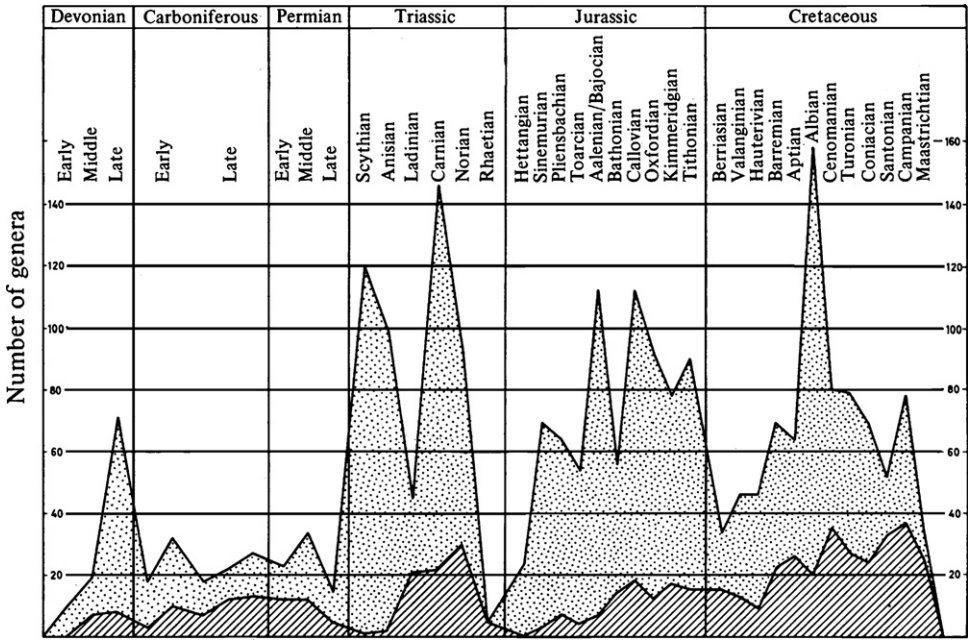


Figure 1 The differentiation of the ammonoids from the Devonian to the Cretaceous, based on the number of genera. Strippled area, new genera; Hatched area, continuous genera.

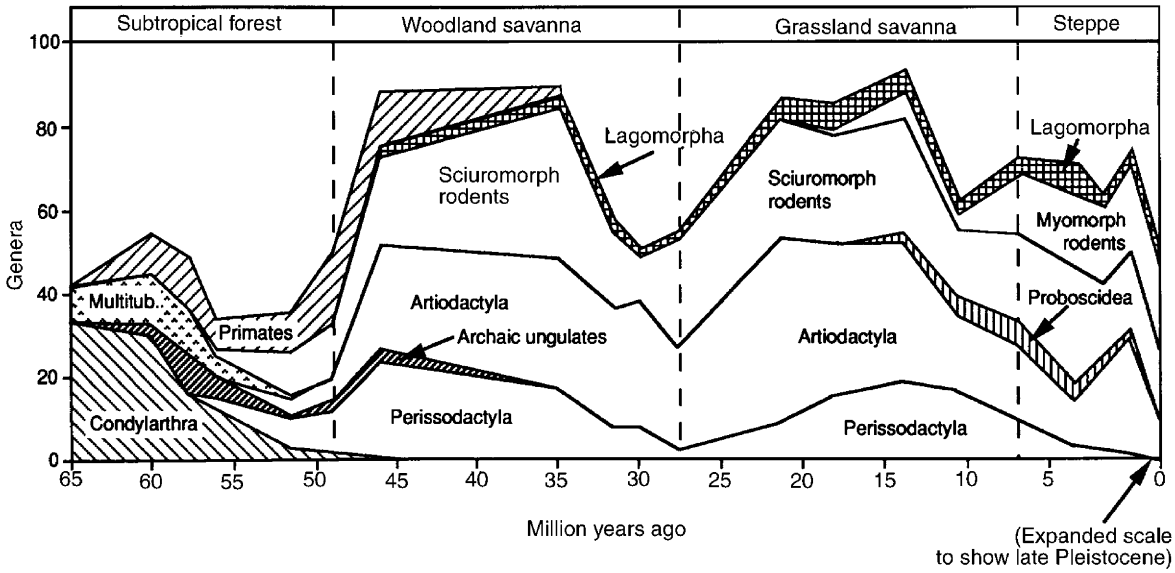


Figure 2 Plot of herbivore primary consumer diversity, taxonomic composition, and inferred dominant habitat type in North America for successive land-mammal ages during the Cenozoic. Vertical dashed lines indicate boundaries. Reprinted from MacFadden, 1992. Fossil horses: systematics, paleobiology, and evolution of the family Equidae.

relation to the adaptive radiation of butterflies. They found high levels of diversification in substrate recognition sites between species that do not share the same set of host plants; the reverse was true for those species that do share host plants. This result was taken to indicate that specialization may necessitate conservation of this region of the genome and could therefore be considered a key innovation.

Many attributes of species have been proposed as key innovations, or characteristics that have allowed diversification and proliferation. They generally involve the development of features that modify biotic interactions. Particular examples include the development of toxicity in plants (that allows them to "escape" predatory pressure of insects) and subsequent development of tolerance to the toxin in insects which allows them to radiate onto the plants. Symbioses are another possible "evolutionary innovation," allowing the abrupt appearance of evolutionary novelty (Margulis and Fester, 1991). They provide a possible avenue through which taxonomic partners can enter into a new set of habitats unavailable to one or both of the symbiotic partners alone. One of the most famous examples is the radiation of ruminants in the African savannas, which has been attributed partially to the development of gut endosymbionts and the concomitant ability to digest cellulose. Among the Foraminifera, Norris (1996) showed that photosymbiosis appeared in the fossil record in synchrony with the taxonomic differentiation of three of the dominant surface water foraminifera groups in the Paleocene and early Eocene. This radiation was not paralleled in the asymbiotic sister group. Symbiosis was suggested to provide a jump-start for diversification by providing the ecological opportunity.

In their classic paper, Ehrlich and Raven (1964) examined how interacting species in themselves may create ecological opportunity, and hence periodically enhance evolutionary rates, through a broad "coevolutionary" response. They hypothesized that, when plant lineages are temporarily freed from herbivore pressure via the origin of novel defenses, they enter a new adaptive zone in which they can undergo evolutionary radiation. However, if a mutation arose in a group of insects that allowed it to feed on one of these previously protected lineages of plants, it would also be free to diversify in the absence of competition. Ehrlich and Raven envisioned this as a step-like process in which the major radiations of herbivorous insects and plants have arisen as a consequence of repeated opening of novel adaptive zones that each has presented to the other over evolutionary history. This idea, termed the "escalation/diversification" hypothesis (Berenbaum and Feeny, 1981), has been supported by the work of Farrell and colleagues, who have studied insect diversification in the context of host plants (**Figure 3**). Repeated evolution of angiosperm feeding in phytophagous beetles is associated with an increased rate of diversification (Farrell, 1998). Similarly, there is consistently greater diversity among plants in which latex or resin canals have evolved as protection against insect attack (Farrell *et al.*, 1991).

Since Ehrlich and Raven, there has been a tremendous amount of research on the role of coevolution in dictating patterns of diversification. One of the major avenues that this research has taken is the study of the extent to which the phylogenetic order of divergence among herbivores or parasites corresponds to that among their hosts as a result of "parallel diversification" (Farrel and Mitter, 1994). A strongly corresponding evolutionary history might suggest a coevolutionary response between the host and the herbivore or parasite. However, there appear to be few cases, at least among insect–plant interactions, in which the phylogeny agreement is precise. In most cases there has been periodic transfer of species to more distantly related hosts.

How do the coevolution arguments invoke adaptive radiation? The situation that Ehrlich and Raven (1964) envisioned was one in which the host radiated prior to exploitation and subsequent radiation by the herbivore ("escape and radiation").

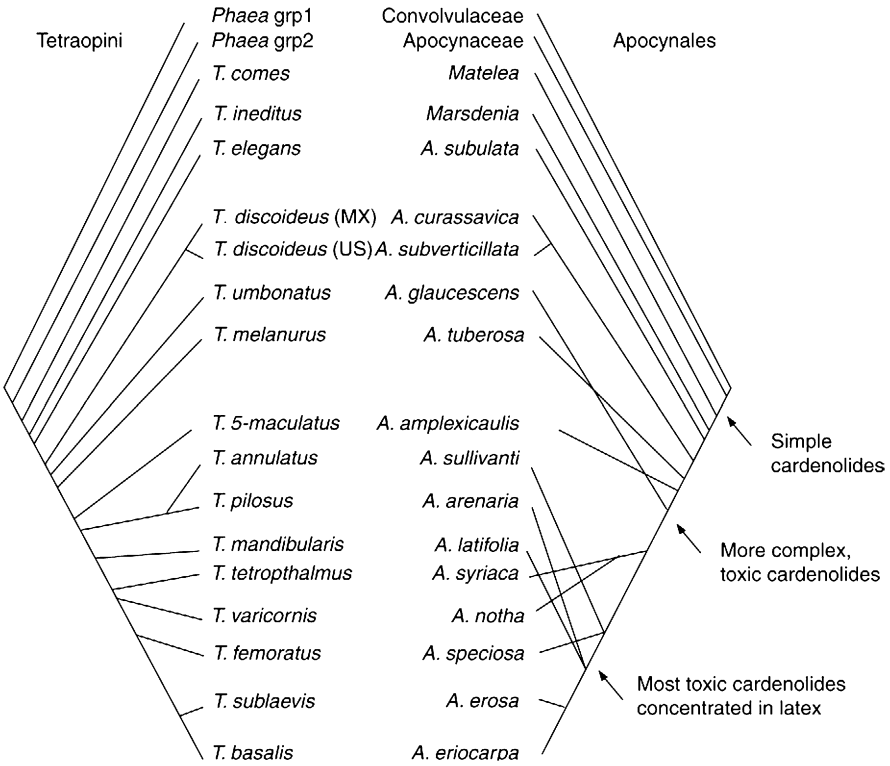


Figure 3 Phylogeny estimate of *Tetraopes* beetles based on morphology and allozymes, compared to literature based relationships of host plants. Host *Asclepias* shows an apparent progression toward increased complexity and toxicity of cardenolides, perhaps representing escape and radiation. Reproduced from Figure 6 in Farrell BD and Mitter C (1994) Adaptive radiation in insects and plants: Time and opportunity. *Am. Zool* 34: 57–69.

This might be considered analogous to the opening up of an array of ecological opportunities every time the innovation arose for either "escape" or "exploitation." The established diversity of hosts could provide the necessary diversity of ecological roles and associated adaptations. Where coevolution involves parallel diversification, adaptive radiation may not be involved. In particular, parallel diversification might be considered analogous to geographic separation, with divergence of the host causing isolation of the herbivore. This might then be considered a nonadaptive radiation. On the other hand, parallel diversification might cause an escalation in responses, with enhanced toxicity and reciprocal tolerance evolving in step-like progression (Berenbaum, 1983; Farrell and Mitter, 1994). In this latter scenario, adaptive radiation can be implicated for both the herbivore and the host.

Extrinsic Factors

Speciation rates are generally considerably higher in novel environments, whether a lake in the middle of a continent or an island in the middle of the ocean (Schluter, 1998; **Figure 4**).

Environmental Change

Environmental change has frequently been implicated in species radiations, with the opening up of new habitat. Diversification has frequently been suggested to occur under stressful conditions [e.g., for the origin of angiosperms (Shields, 1993) and the recent diversification of mole rats (Nevo *et al.*, 1984)]. However, any novel environment in which the organism is subjected to a new selective regime could be considered "stressful." In other words, an organism that successfully invades a novel environment will inevitably be subject to stressful conditions. However, no matter whether the novel environment is considered stressful or simply a situation in which the organism is subject to a novel set of selective forces, it does appear to be associated with acceleration in evolutionary rates (Nevo *et al.*, 1984; Shields, 1993).

The evolution and adaptive radiation of the African cichlids (**Figure 5**) appear to have been initiated by environmental change. Geological activity 20 million years ago (mya) caused the rivers in the area to become progressively meandric and swampy while still connected to the Zaire hydrological system. Over time, a mosaic of small, shallow, and isolated lakes developed, and finally the drainage system became closed and the lakes deepened (approximately 5 mya). The diversification of cichlid fish appears to have been initiated when river species moved into the swamp (Sturmbauer, 1998), and then successive radiations were associated with the development of protolakes and subsequently deep lakes.

In geological history, environmental changes appear to form the basis of the Phanerozoic revolutions (Vermeij, 1995): Increasing temperature and nutrient supplies as a result of submarine vulcanism may have triggered later Mesozoic and perhaps early Paleozoic diversification episodes. Similar

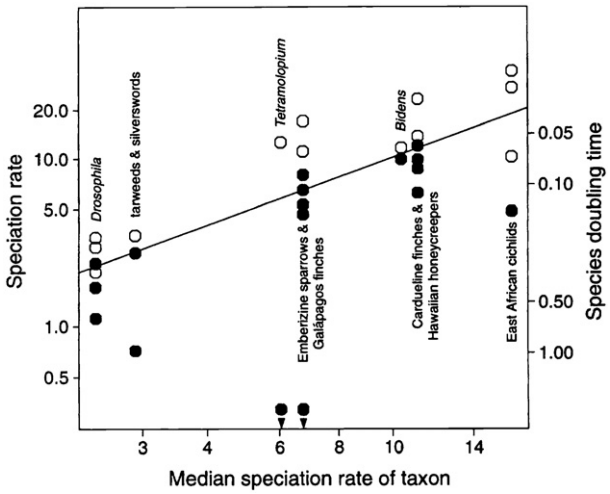


Figure 4 Per capita speciation rates of clades in novel environments (○) and in closely related "control" lineages inhabiting other environments (●). Rate estimates (y, left axis) are plotted against a dummy variable, the median of y-values. The solid line indicates $y = x$; points above the line therefore exhibit high rates. Rates were calculated from phylogenies based on allozyme frequencies. Time is measured in units of genetic distance (D). The calculation of rate y assumes exponential growth of species number: $y = \ln(N)/t$, where N is the number of extant species in a clade and t is its estimated time of origin. Corresponding times required for species number to double are indicated on the right; the number of species in a clade doubles after ln (2)/y time units. Reproduced from *Endless Forms: Species and Speciation*, ed. by D. J. Howard and S. H. Barlocher, © 1998 by Oxford University Press, Inc.

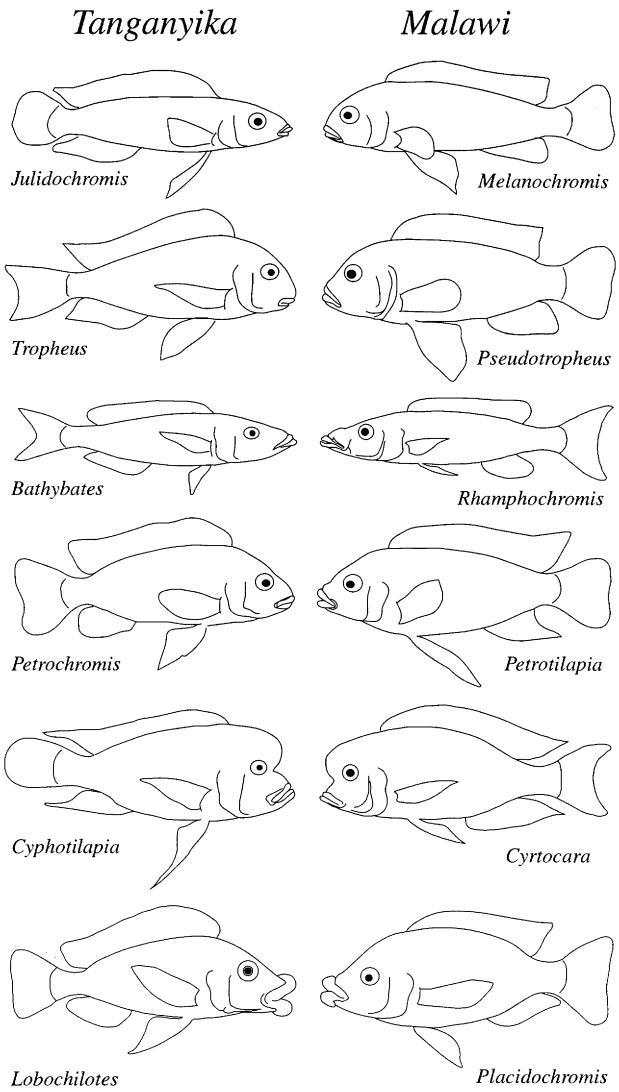


Figure 5 Adaptive radiation of cichlid fish, showing convergence in body shapes in Malawi and Tanganyika species. From TREE review. Reproduced from Meyer A (1993) Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends in Ecology & Evolution* 8: 279–284.

factors may underlie the iterative radiations of ammonoids throughout the geological record (Dommergues *et al.*, 1996). Each radiation appears to have originated from a few taxa, which went on to produce a wealth of morphological diversity. Although well documented for ammonoids, the pattern of iterative radiation is extremely rare in the fossil record. Within the total morphospace defined for the first three Jurassic stages, the radiation corresponds to a string of "events" separated by episodes of morphospace collapse. During each event, only a portion of morphospace (<45%) was filled; some morphs were reiterated, but each event had its own particular derived morphs.

Colonization of Isolated Landmasses (Particularly Oceanic Islands)

Isolated landmasses that have never been in contact with a source provide abundant opportunity in terms of newly available "niche space" for those taxa that manage to colonize them. Isolated archipelagoes are generally considered in a different category of adaptive radiations, although they are really just special cases of environmental change: the appearance of a new environment, which happens to be isolated in the ocean. This has much in common with the formation of, for example, a lake in the middle of a continent. In either case, newly created habitats that are isolated from a source of colonists provide an extraordinary opportunity for adaptive radiation. Both the novelty and the isolation are key features in allowing adaptive radiation in such areas. If a new habitat appears in close proximity to other such habitats, it will be colonized by taxa from those habitats. Species diversity patterns will then match closely the predictions of the MacArthur–Wilson model of island biogeography (MacArthur and Wilson, 1967); that is, species diversity patterns will be governed by ecological processes. As isolation from the source of colonists increases, fewer taxa will be able to colonize the new habitat, and the low rate of colonization may provide sufficient time for species diversification to occur. Adaptive radiations are most likely to occur at the extreme ends of the dispersal range of a given taxon (Whittaker, 1998).

Are Certain Taxa more Likely to Undergo Adaptive Radiation than Others?

Are species predisposed to undergo adaptive radiation because of a broad environmental tolerance, generalized feeding

patterns, or perhaps some proclivity to develop novel associations? This question has been developed by some authors. For example, Adler and Dudley (1994) compared patterns of adaptive radiation among birds and butterflies in the insular Pacific: Birds have undergone extensive adaptive radiation, whereas butterflies have not. They argued that speciation in butterflies may be constrained by the mechanics of insect-plant coevolution that prevents rapid diversification. However, this argument is not well supported because other insects with similar coevolutionary ties have undergone some of the most spectacular insular adaptive radiations known. It appears that almost any group of organisms is capable of undergoing adaptive radiation upon being provided ecological opportunity that it can exploit.

How does Adaptive Radiation get Started?

Initiation of Adaptive Radiation: Genetic Changes

Founder Events

The establishment of species in new environments inevitably involves sampling from the parent population. The size of the sample that can build a new population can be very small (cf. founder effects), although it need not necessarily be so. In particular, if, subsequent to colonization, a very small number of individuals were to proliferate rapidly, there would be little subsequent loss in genetic variability (Nei *et al.*, 1975). Consequently, the deleterious effects of inbreeding are largely mitigated. However, because the genes represented in the founding population are only a small sample of the original population, genetic drift may be pronounced.

The nature of genetic changes during shifts in population size, particularly those experienced during or after population bottlenecks, has been the subject of considerable controversy in recent years. Clearly, a crash in population size as a result of a genetic bottleneck or founder event will cause allele frequencies at some loci to differ from those of the parent population because of accidents of sampling (Templeton, 1980). The debate concerns the nature of genetic changes that occur subsequent to the bottleneck, during the period of population growth. Traditional arguments suggested that founder events may trigger rapid species formation (Carson and Templeton, 1984). However, recent arguments have largely refuted the contribution of founder events to reproductive isolation (Barton, 1996).

Other possible changes during founder events are due to genetic reorganization. Carson (1990) proposed that blocks of loci are destabilized when a newly founded colony undergoes a flush of exponential growth, during which time selection is relaxed and recombinants that ordinarily have low fitness survive. Release of additive genetic variance through change in epistatic interaction may allow formation of novel recombinants (Goodnight, 1988). Recent work has questioned the importance of epistatic interactions and the nature of "destabilization" among blocks of loci during periods of reduced population size. Slatkin (1996) discussed changes that might occur subsequent to reduction in population size in terms of conventional population genetics. Genetic drift will be the primary force causing genetic changes during the bottleneck. However, when the population starts to grow, the influence of genetic drift is expected to become weaker so that selection is most effective during and immediately after the period of rapid demographic expansion. On the other hand, Otto and Whitlock (1997) argued that, although it is true that the probability of fixation of beneficial alleles present during the bottleneck is increased during the subsequent period of population growth, a large proportion of alleles are lost during the bottleneck and few new mutations can occur while the population is at small size. The resultant effect of these opposing forces is that the number of beneficial mutations fixed per generation remains virtually unchanged by the bottleneck. Nevertheless, selection subsequent to a genetic bottleneck has an important effect on alleles that are initially rare and that would tend to be lost to stochastic events in populations of constant size (Slatkin, 1996). In addition, there is evidence that stress on a genetic system (e.g., as a result of a population crash) may activate transposable elements (Carson, 1990), which can exercise a mutagenic effect by interrupting structural or regulatory regions of the genes into which they become integrated. Therefore, mutation rates may actually be higher in founder populations, providing raw material on which selection can act, and which in turn could lead to a rapid recovery of genetic variability. There are thus two processes which can be associated with colonization events: (i) possible genetic changes/restructuring within the population and an expected loss of genetic variability and (ii) the subsequent recapturing of genetic variation during population expansion as a result of selection (Slatkin, 1996).

Rapid Proliferation and Hybridization

Differential mixing of characters during segregation of populations and species may occur as a result of hybridization of newly divergent taxa (Harrison, 1993). Behavioral changes during founder events may facilitate hybridization because it has been suggested that sexual interactions may lose specificity subsequent to a founder event (Kaneshiro, 1989). Closely related heterospecifics may therefore hybridize and/or introgress subsequent to colonizing a new landmass, but the extent of genetic exchange may differ between regions of the genome (DeSalle and Giddings, 1986). Differences may be particularly pronounced between character sets involving nuclear and extranuclear DNA, such as mitochondrial or chloroplast DNA. Extranuclear DNA differs from nuclear DNA because of its greater sensitivity to the effects of each founder event due to its smaller effective population size (one-fourth) relative to nuclear DNA, attributable to its transmission primarily through the female line, and its existence as a single copy (Avise, 1991). Analysis of extranuclear DNA information will result in a gene genealogy, but it is likely to provide an incomplete history of the organisms if much hybridization has taken place. The occurrence of hybridization may explain differences that have been found between nuclear and mitochondrial phylogenies, particularly for the Hawaiian *Drosophila* (DeSalle *et al.*, 1997). Indeed, natural hybridization with the formation of fertile hybrids has been documented between closely related species of *Drosophila* on the youngest island of the Hawaiian Islands (Carson, 1989). Among silverswords, hybridization has been

implicated as an important element in the adaptive radiation of the group (Baldwin, 1997).

Among the Galapagos finches, recent molecular data have failed to distinguish species limits, at least for the morphologically defined ground and tree finch species: Individuals representing different morphologically identified species are intermingled. This may be explained by interspecific hybridization and/or sorting of haplotypes (Freeland and Boag, 1999). In the case of ground and tree finches, both explanations may apply. The incomplete species differentiation within these taxa may be taken as an indication that adaptive radiation is currently ongoing.

Failure of molecular data to distinguish species limits has been found in several other adaptive radiations. For the cichlid fish of Lake Malawi, it has been suggested that speciation is occurring faster than alleles can become fixed within a species (Moran and Kornfield, 1993). Among Hawaiian crickets, Shaw (1996) comments on a discrepancy between phylogenies generated on the basis of song (Otte, 1994) with that generated from mtDNA variation. Although the discrepancy could be explained by problems with current taxonomic boundaries, and problems with mtDNA lineage sorting, she argues that hybridization and introgression appear to be the most likely explanation.

Although the number of species in which hybridization has been documented is currently small, it is likely to increase as researchers accumulate phylogeographic knowledge based on multilocus molecular data for different radiations.

Initiation of Adaptive Radiation: Ecological Changes

Populations frequently initiate a cycle of change in abundance and distribution upon colonization of a novel habitat or an unoccupied set of niches, during which they undergo ecological release, expand their range, and adopt a more generalized habit (Cox and Ricklefs, 1977). After their initial rise to dominance as widespread generalists, these taxa may subsequently be competitively displaced from much of their original range by younger relatives (Darlington, 1957). Wilson's (1961) "taxon cycle" was used to describe such changes in the distributional pattern of Melanesian ants. He suggested that widespread, dispersive populations give rise to many more restricted and specialized species. Regular changes in ecological and geographical distribution have been recognized in many island systems (Cox and Ricklefs, 1977).

Ecological release may be the precursor to adaptive radiation. This argument has been supported by recent experimental evidence: Rainey and Travisano (1998) studied adaptive radiation experimentally using the bacterium *Pseudomonas flourescens*. This interesting bacterium is known to evolve rapidly in novel environments, with evolutionary differences detectable in the morphology and ecological affinity of the phenotype. An isogenic population of one morph was propagated in (i) a spatially heterogeneous environment and (ii) a spatially homogeneous environment. In the spatially heterogeneous environment, extensive morphological diversification was found to occur within 3–10 days, resulting in three dominant morphs. This was taken as evidence for diversifying selection. In the spatially homogeneous environment, no morphological variation was found.

The Processes of Adaptive Radiation: Case Studies

What are the factors underlying diversification in an adaptive radiation? This question has fascinated biologists for many years. However, until recently, the characters involved in the adaptive radiation necessarily had to be used as the basis for phylogenetic inference. Given that adaptive radiations are characterized by tremendous levels of convergence in almost every morphological and ecological character, it has been very difficult to interpret evolutionary processes. As Givnish (1997) notes, "any rigorous, noncircular study of adaptive radiation must be based on a phylogeny that has been derived independently of the traits involved in that radiation." Recent advances in molecular techniques, although not without problems of their own, provide an opportunity to obtain such an independent assessment of evolutionary history (see review of arthropod radiations in Hawaii in Roderick and Gillespie, 1998). The following sections outline case studies from a range of taxa, most of which have used a combination of morphological and molecular techniques in an attempt to gain some understanding of processes underlying adaptive radiations.

Galapagos Finches

There are currently 14 recognized species of Darwin's finches in six genera, which have evolved from a common ancestor (**Figure 6**; Lack, 1947; Grant, 1986). Of these, 13 live in the Galapagos Islands. Based on morphological, behavioral, and ecological data, they have been divided into three lineages: First, the ground finches, *Geospiza* (6 species), which are found in more arid areas of the archipelago and feed on seeds on the ground. Three of the species which are considered the most "finch-like" differ primarily in body and beak size and are known as the large (*G. magnirostris*), medium (*G. fortis*), and small (*G. fuliginosa*) ground finches. The 3 other species of ground finches have longer beaks. Two feed on cactus flowers and pulp as well as seeds and are known as the large (*G. conirostris*) and small (*G. scandens*) cactus ground finches. Finally, the sharp-beaked ground finch (*G. difficilis*) supplements its diet with the eggs and blood of other birds and reptile ticks. Second, the tree finches, which are found mostly in trees and shrubs, are divided into 3 genera: *Cactospiza* [the woodpecker finch (*C. pallida*) and the mangrove finch (*C. heliobates*)], *Camarhynchus* [the large tree finch (*C. psittacula*), the medium tree finch (*C. pauper*), and the small tree finch (*C. parvulus*)], and *Platyspiza* (the vegetarian finch, *P. crassirostris*). All except *P. crassirostris* are insect eaters. Finally, the warbler-like finches, which are small with slender beaks, are in 2 genera: The warbler finch (*Certhidea olivacea*) catches insects like a warbler, and the Cocos finch (*Pinaroloxias inornata*) is the only Darwin finch that lives outside the Galápagos Archipelago. It appears to have colonized Cocos Island from the Galapagos.

Darwin's finches share common features of nest architecture, egg pattern, and courtship displays. They differ in song, morphology, and plumage. Based on morphology, allozyme, and DNA sequence data, the warbler finch *C. olivacea* appears to be closest to the ancestral form. However, recent molecular data indicate that the Cocos finch *P. inornata*

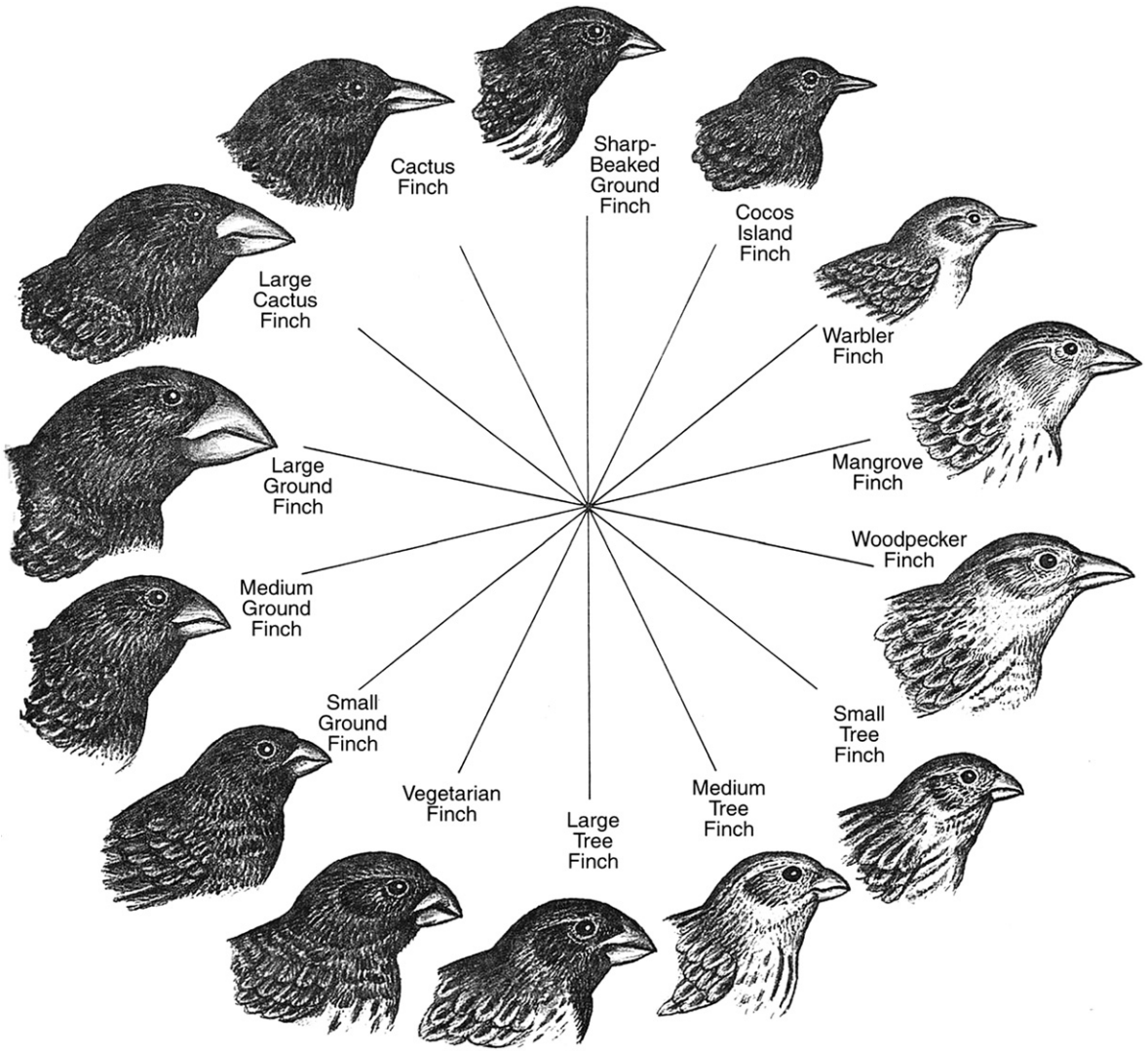


Figure 6 Adaptive radiation of Darwin's finches. Reproduced from Grant (1991).

is closest to the tree finches of the Galapagos (Sato *et al.*, 1999; Petren *et al.*, 1999). Also, the vegetarian finch appears to be ancestral to both the tree and ground finches rather than being a member of only the tree finch group.

Grant (1986) proposed a three-step process to speciation to explain the adaptive radiation of the group: (i) The ancestral species arrives in the archipelago; (ii) the species spreads to other islands and as a result of this, there will probably be some selection and hence differentiation in size because the islands are different; and (iii) members of the original and derived populations encounter each other, and as a result of competition for food and selection against intermediates character displacement causes rapid divergence in feeding structures between the species when they come together. Therefore, the radiation appears to be based on (i) the isolation of the archipelago, which provided ecological opportunity; (ii) considerable distances between the different islands, which has led to infrequent interisland exchange; and (iii) different environments on the different islands, which have selected for different feeding niches both within and between islands.

Hawaiian *Drosophila*

The Hawaiian Drosophilidae represent some of the most striking examples of adaptive radiation known for any group anywhere (Kaneshiro, 1988), with 337 species in the genus *Drosophila* (pomace flies), 19 in *Idiomyia* (picture-winged flies), 122 in *Scaptomyza*, and 11 in *Titanochaeta* (Hardy, 1965; Hardy and Kaneshiro, 1981). Based on the premise that founder events are the most important mechanism of speciation in the Hawaiian *Drosophila*, Kaneshiro (1983) proposed models suggesting the importance of sexual selection in driving species proliferation in these insects. The Hawaiian Drosophilidae, particularly the males of the picture-winged species, often have ornately patterned wings as well as unusual modifications of the mouthparts and legs. The extraordinary manifestations of these features in the male picture-winged species are frequently accompanied by elaborate courtship behavior. In classic studies of picture-winged *Drosophila*, Kaneshiro (1983) found evidence for asymmetrical sexual isolation: Females from a geologically older island were found

to be highly discriminating in terms of mate choice and would not mate with a closely related species from a younger island. In contrast, females from species on the younger island readily accepted males from the older island species. Based on these observations, Kaneshiro proposed that courtship requirements are relaxed during the early stages of colonization of a new island; sexual behavior may then become less constrained and simpler, and there may be more intraspecific variability. In such circumstances, there would be strong selection for less discriminating females because of the difficulties in finding mates and reproducing when the population size is small. Intrasexual selection may then operate to cause divergence of the sibling species during isolation because of a shift in the distribution of mating preferences during the founder/flush cycle (Carson, 1986). The result would be a shift in the mating system, which could then be fixed at a new "equilibrium."

Evolution of the Hawaiian *Drosophila* is commonly treated as an example of sexual selection influencing speciation (Carson, 1986). The nature of this influence is not entirely clear. *Drosophila heteroneura* and its close relative, *D. silvestris*, are partially sympatric in forests on the island of Hawaii, where they occasionally hybridize. The species are distinguished by the much broader head of *D. heteroneura*. Male aggression and male courtship are the major determinants of male mating success. Accordingly, it appears that male head width is subject to sexual selection through mate choice, although this, and perhaps related sexually dimorphic traits, may not be involved in behavioral isolation (Boake *et al.*, 1997).

A recent molecular phylogeny of picture-winged Hawaiian *Drosophila* showed that clades are characterized according to whether they breed on fungi, leaves, fruit, or bark (Kambysellis and Craddock, 1997). Suites of reproductive characters (in particular, ovarian egg and ovipositor traits) appear to have evolved together (Kambysellis, 1993), and adaptive shifts to new breeding sites appear to have been important at least in the early diversification of the group and in some recent speciation events (Kambysellis and Craddock, 1997). The interplay of such ecological shifts with sexual selection in allowing diversification of Hawaiian *Drosophila* remains unresolved.

Hawaiian Honeycreepers

The endemic Hawaiian subfamily Drepanidinae, or Hawaiian honeycreepers, includes 33 historically known species and at least 22 "fossil" species (**Figure 7**; James and Olson, 1991). The three tribes are (i) the Psittirostrini, Hawaiian finches (23 species), which includes 5 extant species, 4 historically extinct species, and 14 fossil species; (ii) the Hemignathini, Hawaiian creepers and allies (22 species), which includes 13 extant species, 3 species and 2 subspecies historically extinct, and 6 fossils; and (iii) the Drepanidini, Mamo's, Iiwi's, Apapane's, crested honeycreeper and allies (9 species), which includes 4 extant species, 3 species historically extinct, and 2 fossils. Based on the tremendous morphological and ecological differentiation known from historical and recent collections, these birds have been heralded as one of the best known examples of adaptive radiation. However, the high proportion of currently extinct species makes it difficult to develop hypotheses regarding processes underlying the adaptive radiation. Molecular studies based on extant species suggest that differentiation has occurred between islands only (Tarr and Fleischer, 1995). One might speculate that species have proliferated in a manner similar to that proposed for the finches in the Galapagos, as described previously.

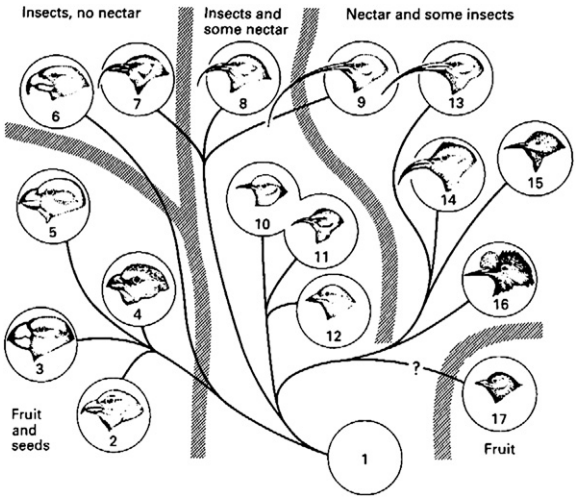


Figure 7 The inferred pattern of evolution of dietary adaptations as represented by 16 of the Hawaiian honeycreepers. 1. Unknown finch-like colonist from Asia; 2. *Psittirostra psittacea*, 'O' u; 3. *Chloridops kona*, Kona Grosbeak (extinct); 4. *Loxioides bailleui*, Palila; *Telespyza cantans*, Laysan Finch; 6. *Pseudonester xanthophyrs*, Maui Parrotbill; 7. *Hemignathus munroi*, 'Akiapola'au; 8. *H. lucidus*, Nuku Pu'u; 9. *H. obscurus 'Akialoa*; 10. *H. parvus*, 'Anianiau; 11. *H. virens, 'Amak-ihi*; 12. *Loxops coccineus*, Hawaii 'Akepa; 13. *Drepanis pacifica*, Mamo (extinct); 14. *Vestiaria coccinea*, Iiwi; 15. *Himatione sanguinea*, 'Apapane; 16. *Palmeria dolei*, 'Akohekohe; 17. *Ciridops anna*, 'Ula-'aihawane (extinct). Reproduced from Cox, Biogeography: An ecological and evolutionary approach (5th ed.), Oxford University Press.

African Cichlids

Lacustrine fish represent some of the most spectacular cases of adaptive radiation in vertebrates (McCune, 1997), with the best known being cichlids. Cichlids are spiny-rayed freshwater fishes that reach their most abundant diversity in Africa, particularly in the great east African lakes of Victoria (>400 species), Malawi (300–500 species), and Tanganyika (approximately 200 species) (**Figure 5**). In each of these lakes the fish exhibit spectacular diversity in trophic morphology, including specialist algal scrapers, planktivores, insectivores, piscivores, paedophages, snail crushers, and fin biters. In addition to their trophic diversity, they exhibit a striking array of color patterns. They also show complex mating behaviors, polyandrous mating systems, and a tendency to breed in leks. Several factors are considered to be involved in the diversification of these fish. One suggests that morphological adaptation is the primary event underlying speciation (Liem, 1973). The basis for this assertion is their trophic apparatus, which is unique in the possession of two sets of jaws—one in the mouth for sucking or scraping and the other in the throat for macerating, crushing, or piercing. These structures can be modified according to the diet of the fish. The second factor

considered to be involved in allowing reproductive isolation and hence diversification of the fish is sexual selection, which allows subsequent morphological differentiation.

Recent molecular work (Meyer, 1993; Meyer *et al*., 1996) has shown that cichlids in any one lake are far closer to each other than to morphologically similar species in other lakes. Evolution has generated similar ecomorphs in each of the lakes. There has been considerable debate regarding the nature of the mechanism underlying the proliferation of species within a lake, particularly whether it is considered "allopatric," "microallopatric," or "sympatric" (McCune and Love-joy, 1998). However, perhaps the most broadly accepted mechanism involves repeated isolation. For example, Lake Tanganyika has apparently undergone repeated decreases in water level (as much as 2000 ft). Populations may have been isolated in small pockets of water for long enough to become reproductively isolated. Sexual selection may play a role in allowing reproductive isolation of the fish subsequent to isolation because closely related species differ primarily in color pattern and not in jaw morphology (Albertson *et al*., 1999). It appears that morphological adaptation played a role in the early radiation of the group, resulting in distinct clades that differ in jaw morphology; only minor morphological differences are observed within genera (Stauffer *et al*., 1997). Recent diversification is largely associated with variation in adult male color pattern.

Hawaiian Silverswords

The Hawaiian silversword alliance has been considered "the best example of adaptive radiation in plants" (Raven *et al*., 1992). It consists of 30 species in three genera (*Wilkesia*, *Dubautia*, and *Argyroxiphium*) but with one common ancestor. Life form diversity includes trees, shrubs, mat plants, monocarpic and polycarpic rosette plants, cushion plants, and vines that occur across a broad environmental spectrum from rain forests to desert-like settings (Carr, 1985). Major ecological shifts have accompanied speciation in each of the major island-endemic lineages (**Figure 8**; Baldwin, 1997). The estimated minimum rate of diversification in the silversword alliance (0.56 ± 0.17 species per million years) is comparable to, or higher than, rates of several more ancient continental

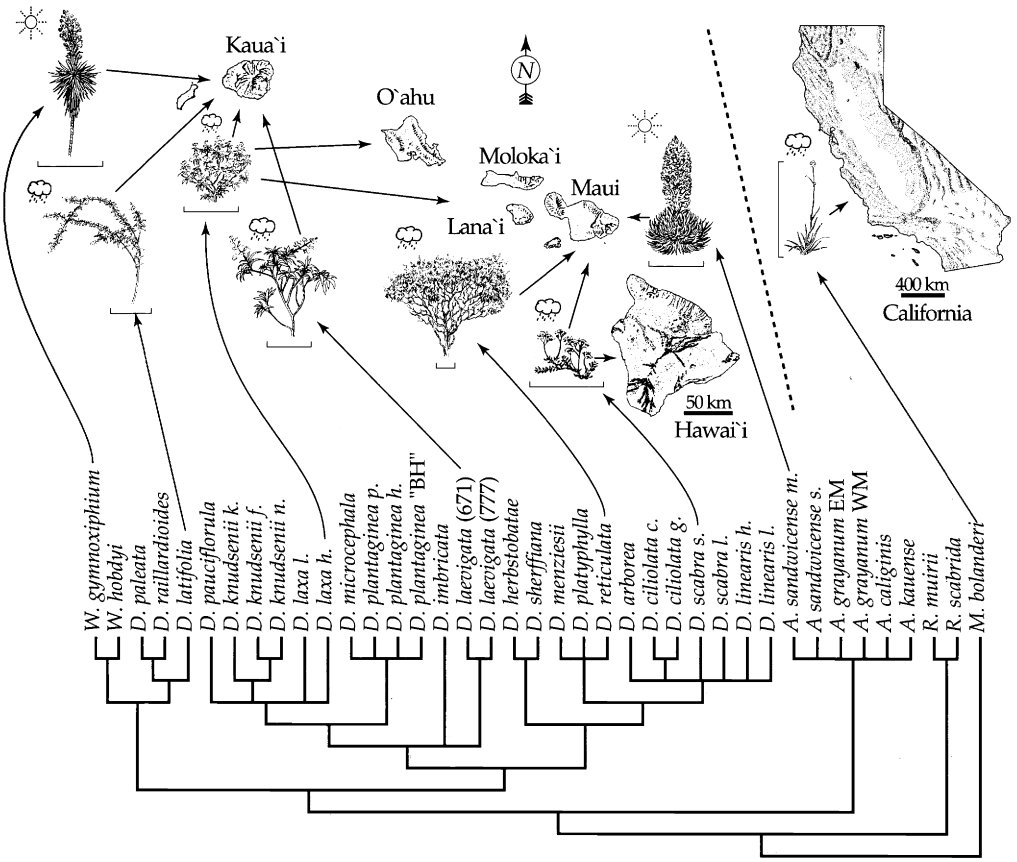


Figure 8 Phylogeny of the Hawaiian silversword alliance and the American tarweed genera *Madia* and *Raillardiopsis* based on DNA sequences. Generic abbreviations: A, *Argyroxphium*; D, *Dubautia*; M, *Madia*; R, *Raillardiopsis*; W, *Wilkesia*. EM, East Maui; WM, West Maui. Plant scale bars equal 1 meter. Arrows indicate island(s) or State (California) of occurrence, not necessarily the location within islands or California where the plants are found. Plant illustrations depict a variety of habits and habitat/island occurrences in the silversword alliance. Sun symbol, restriction to dry habitats; rain-cloud symbol, restriction to wet habitats (including bogs). Reproduced from Baldwin B (1997) Adaptive radiation of the Hawaiian silversword alliance: Congruence and conflict of phylogenetic evidence from molecular and non-molecular investigations. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 103–128.

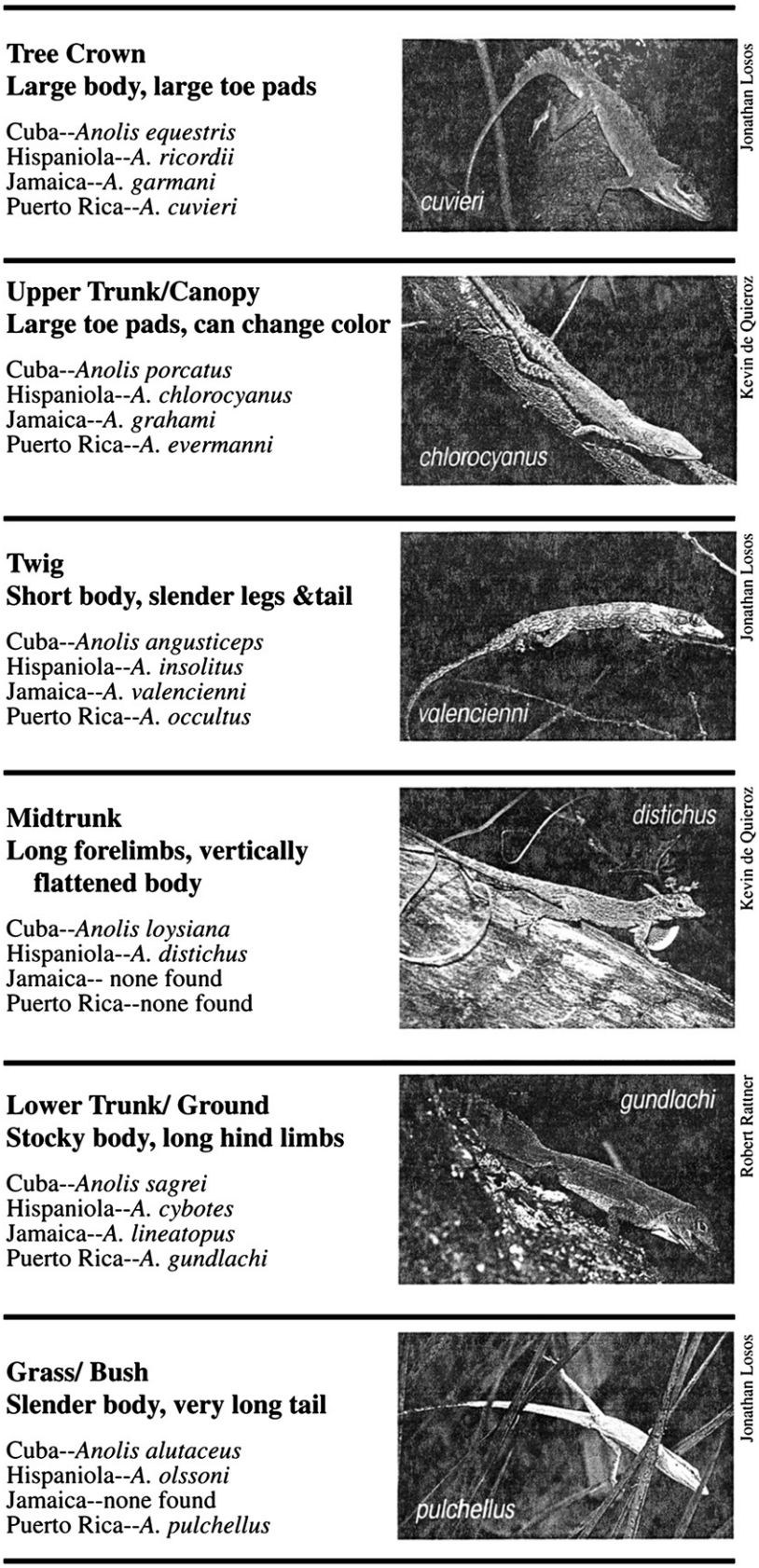


Figure 9 Adaptive radiation of Caribbean lizards. In the Greater Antilles, *Anolis* lizards that adapted to corresponding niches look alike, although they are not closely related. Shown is a sampler of niche holders listed by species name, with a photograph of one member of each category. Reproduced from Natural History, Dec-Jan 1997 v106 n11 p34(5), Darwin's Lizards: like Galapagos' finches, anoles of the Greater Antilles have proved to be eminently adaptable. Jonathan B. Losos; Kevin de Queiroz.

groups that have been regarded as examples of adaptive radiation (Baldwin and Sanderson, 1998).

Lizards in the Caribbean

Currently, 139 species of Caribbean lizards are recognized, 80% of which occur in the islands of the Greater Antilles (Cuba, Hispaniola, Jamaica, and Puerto Rico) (**Figure 9**). Losos (1992) and Losos *et al.* (1998) found that on each of the islands lineages have diversified in such a way as to occupy a range of ecological roles, with as many as 11 species occurring sympatrically. Different species live, for example, on twigs, in the grass, or on tree trunks near the ground. Twig anoles tend to be slender with short legs and tails, whereas the trunk-ground anoles are stocky with long legs and poorly developed toepads. Losos and colleagues recognized six habitat specialists and called these "ecomorphs," which are recognizable on the basis of morphological measurements. The same set of ecomorphs are found on each island (with a few exceptions): Four occur on all islands (trunk-ground, trunk-crown, crown-giant, and twig), one occurs on all islands except Jamaica (grass-bush), and one occurs only on Cuba and Hispaniola (trunk). Phylogenetic studies using mitochondrial DNA indicate that the same ecomorphs have evolved independently on each island, i.e., they appear to have arisen as a result of one-to-one convergence of the same set of ecomorph types on each island (Losos, 1992). However, the sequence by which they evolved differs on each island. Such tests of convergence are among the strongest available for evaluating the premise that the number and types of coexisting species are locally determined. The conclusion from this research was that "adaptive radiation in similar environments can overcome historical contingencies to produce strikingly similar evolutionary outcomes" (p. 2115, Losos *et al.*, 1998).

Sticklebacks in Deglaciated Lakes of Canada

Deglaciated lakes of coastal British Columbia, Canada, harbor a tremendous diversity in the form of undescribed sibling species of fish. There are numerous examples of species pairs in many fish families, and the repeated occurrence of such sympatric pairs has been attributed to novel ecological opportunity provided by deglaciation and recolonization from relatively depauperate faunas. In particular, the radiation of species of three-spined sticklebacks (a species complex currently classified under *Gasterosteus aculeatus*) includes some of the youngest species known—less than 13,000 years. No more than two species occur in any one lake, but pairs of species in different lakes appear to have evolved completely independently of other pairs. Results based on morphological, ecological, and molecular data suggest that species have diverged as a result of parallel bouts of selection for alternate trophic environments (**Figure 10**; Taylor *et al.*, 1997). Natural selection has been implicated as the major cause of evolution (Schluter, 1994; Schluter and McPhail, 1993). This study has been very important in highlighting the role of divergent natural selection as a mechanism underlying adaptive radiation—a view held widely by naturalists in the earlier part of the century although, until recently, without much support (Schluter, 1996). Ecological character displacement appears to underlie diversification in the sticklebacks, with the evolution of reproductive isolation a by-product of resource-based divergent natural selection.

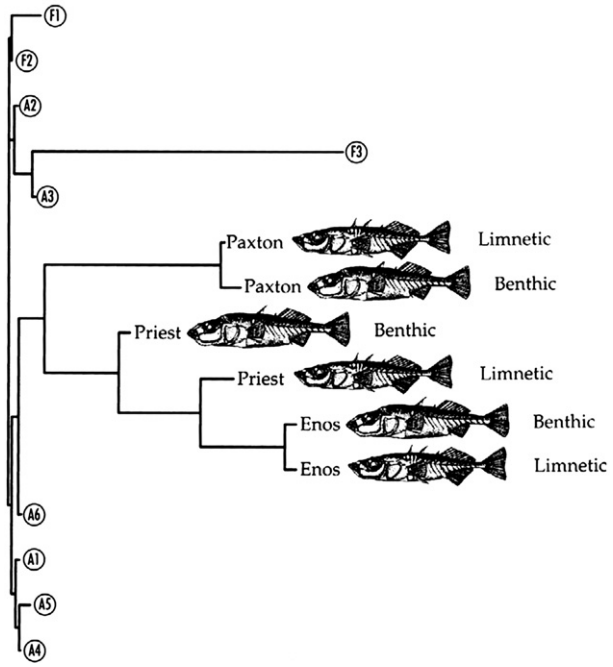


Figure 10 Ecological divergence in sticklebacks. A1–A6 are anadromous populations, F1 and F2 are populations resident in freshwater streams, and F3 is a solitary lake population. Benthic and limnetic species of fish are shown for Paxton, Priest, and Enos Lakes. Adapted from Taylor EB, McPhail JD, and Schluter D (1997) History of ecological selection in sticklebacks: uniting experimental and phylogenetic approaches. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 511–534. Cambridge: Cambridge University Press.

Hawaiian *Tetragnatha* Spiders

The long-jawed, orb-weaving spider genus *Tetragnatha* comprises 295 described species worldwide. The adaptive radiation of the genus in Hawaii has been uncovered only recently, with a total of 28 species described and many more undescribed (**Figure 11**; Gillespie *et al.*, 1998). This radiation spans a huge spectrum of colors, shapes, sizes, ecological affinities, and behaviors. Phylogenetic analyses have been performed on the "spiny-leg" clade of Hawaiian *Tetragnatha*, a lineage that has adopted a vagile, cursorial predatory strategy. On each island, species can be characterized as "green leaf-dwelling," "maroon," and "gray/black bark-dwelling." However, species on any one island are generally most closely related to each other, and each of the different ecomorphs appears to have evolved independently on the different islands (Gillespie *et al.*, 1997). There appears to have been a one-to-one convergence of the same set of ecomorph types on each island.

Hawaiian Swordtail Crickets

Thirty-five species of *Laupala* (Gryllidae) are found in rain forests across the Hawaiian archipelago, and up to 4 species

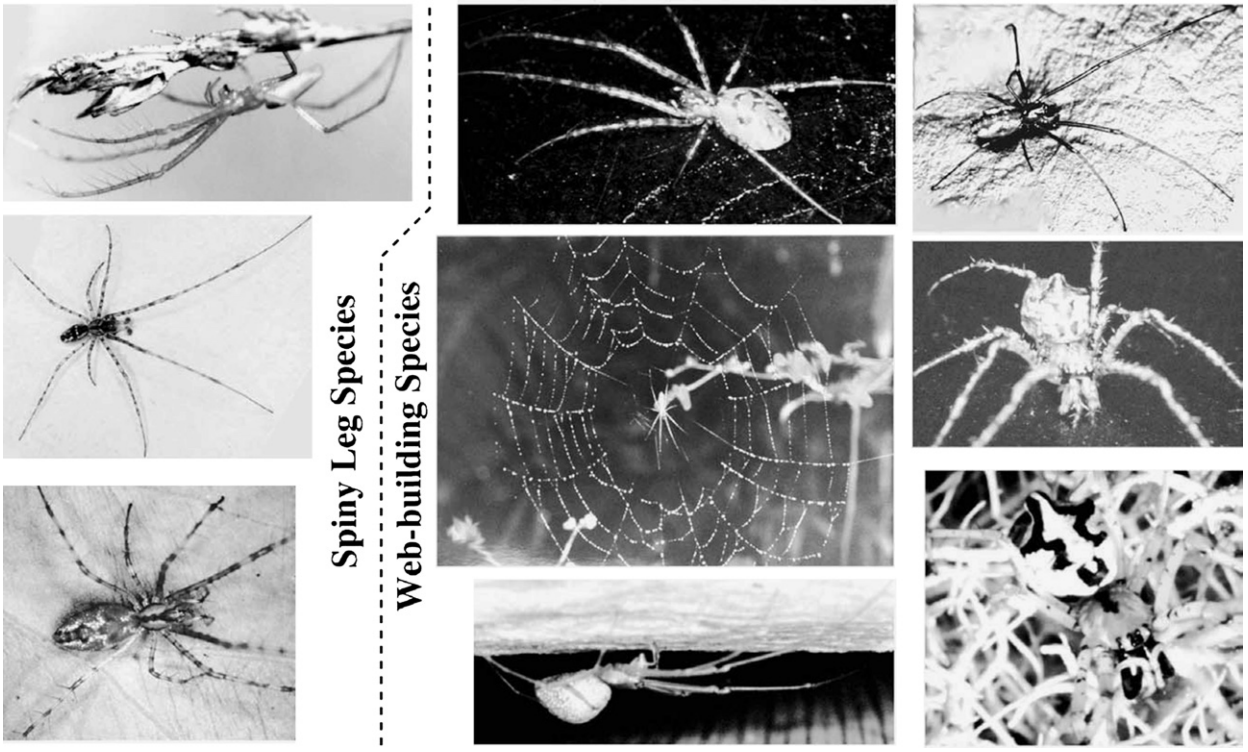


Figure 11 Adaptive radiation of Hawaiian *Tetragnatha* spiders. Shown are representative species from a lineage that has abandoned web building ("spiny-leg" species) and a large lineage of web-building species that show tremendous morphological and ecological diversity. Spiny-leg species: From top, *T. brevignaitha* and *T. kamakou* from wet forests on E. Maui; *T. pilosa* from wet forests on Kauai. Web building species: Five from wet forests on E. Maui (clockwise from bottom left): *T. filiciphilia*, *T. stelarobusta*, *T. paludicola*, *T. eurychasma*, and *T. trituberculata*; (bottom right) undescribed species from wet forests on Hawaii.

coexist on any given island (Otte, 1994). In common with other crickets, courting males "sing" to attract females by rubbing their forewings together. Each species has a unique song, and females respond preferentially and move toward the source of their species-specific song. It appears, therefore, that the song serves as an important mate recognition signal and variation in the acoustic system may serve as a basis for behavioral isolation between species. Phylogenetic analyses based on morphological (Otte, 1994) and molecular (Shaw, 1996) characters indicate that extensive intraisland species formation has occurred in this group. Differentiation between closely related species is associated with changes in song. These features of the *Laupala* system suggest that, unlike the Hawaiian *Drosophila*, founder events have not played a major role in the initial stages of species diversification (Shaw, 1996). Rather, differentiation appears to occur through the interaction of sexual selection on genetically well-structured populations. More work is needed to understand the relative importance of these factors in the initial divergence of species in this group.

Partula Land Snails

Land snails can exhibit tremendous phenotypic polymorphism with respect to the color, banding, and chirality of the shell. *Partula* is a group of species inhabiting the volcanic islands of the South Pacific. The evolution of this group parallels the geological history of the islands. They provide an example of classical adaptive radiation and also show that changes in ecologically important traits need not be accompanied by large genetic changes, and that the ecological changes can take place before reproductive isolation is complete (Johnson *et al.*, 1993; Murray *et al.*, 1993). Competition appears to be important in dictating the array of species at a site.

Bacteria in Culture

As described previously, experimental studies on a bacterium, *Pseudomonas flourescens*, have provided fascinating insights into possible processes underlying adaptive radiation. *Pseudomonas flourescens* was found to diversify rapidly from an isogenic population under conditions of environmental heterogeneity. Moreover, selection appeared to be the primary force maintaining diversity in the heterogeneous environment. The evolution of variant forms was found to follow a predictable sequence (cf. Caribbean lizards, stickleback fish, and *Tetragnatha* spiders), and competition is inferred to maintain the variation (Rainey and Travisano, 1998).

Conclusions

Three primary mechanisms have been inferred to underlie species proliferation in adaptive radiations: (i) geographical differentiation and subsequent character displacement (on the

basis of either ecologically or sexually important characters) in Galapagos finches and African cichlids; (ii) sexual selection, with or without genetic bottlenecks, in Hawaiian *Drosophila* and *Laupala* crickets; and (iii) divergent natural selection based on ecological shifts in Hawaiian *Tetragnatha* spiders, plants in the Hawaiian silversword alliance, lizards in the Caribbean, *Partula* land snails in the Pacific, stickleback fish in deglaciated lakes, and bacteria in culture.

The Future

Adaptive radiation, as a phenomenon, has tremendous research potential: The existence of a suite of closely related species adapted to exploit different habitats or lifestyles allows one to make comparative studies on the processes of speciation and selection in natural populations. Molecular systematics is providing a much better understanding of the evolutionary history of groups of closely related species and provides the opportunity for testing mechanisms underlying adaptive radiation (Givnish, 1997). The phylogenetic hypotheses provide a framework for examining the evolution of specific morphological, ecological, behavioral, and physiological adaptations and the circumstances in which they have arisen and have allowed adaptive radiation.

Adaptive radiations have recently become the focus of studies in conservation biology because they are frequently, particularly on islands, associated with high frequencies of endemism. As a corollary to the high endemism, many of the species that make up an adaptive radiation are often very rare and are characterized by very high extinction rates. Far more extinctions have been documented from islands than from continents: Of known extinctions, those from islands comprise 58% for mammals (most of which are absent from remote islands), 80% for mollusks, and 85% for birds (Whittaker, 1998). Similarly high extinction rates are known for radiations of lacustrine fish. The high rate of extinction of these narrowly endemic species has been greatly accelerated in recent years by the introduction of alien species into island environments. The plight of these extraordinary sets of species has largely been ignored because attention has focused on the devastation of forests in South America and Asia. However, setting aside even small reserves would serve to protect many species on islands from extinction. The microcosmal nature of species swarms makes such efforts at least feasible. However, without immediate action, few of the world's most spectacular radiations will survive far into the next millennium.

See also: Coevolution. Darwin, Charles (Darwinism). Dispersal Biogeography. Diversity, Molecular Level. Endemism. Habitat and Niche, Concept of. Population Genetics

References

Adler GH and Dudley R (1994) Butterfly biogeography and endemism on tropical Pacific Islands. *Biol. J. Linnean Soc.* 52: 151–162.

Albertson RC, Markert JA, Danley PD, and Kocher TD (1999) Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. *Proc. Natl. Acad. Sci. USA* 96: 5107–5110.

Avise JC (1991) Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annu. Rev. Genet.* 25: 45–69.

Baldwin B (1997) Adaptive radiation of the Hawaiian silversword alliance: Congruence and conflict of phylogenetic evidence from molecular and non-molecular investigations. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 103–128. Cambridge, UK: Cambridge Univ. Press.

Baldwin B and Robichaux R (1995) Historical biogeography and ecology of the Hawaiian silversword alliance (Compositae): New molecular phylogenetic perspectives. In: Wagner WL and Funk VA (eds.) *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*, pp. 259–287. Washington, DC: Smithsonian Institution Press.

Baldwin BG and Sanderson MJ (1998) Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proc. Natl. Acad. Sci. USA* 95: 9402–9406.

Barton NH (1996) Natural selection and random genetic drift as causes of evolution on islands. *Philos. Trans. R. Soc. London* 351: 785–795.

Berenbaum M and Feeny PP (1981) Toxicity of furanocoumarins to swallowtail butterflies: Escalation in a coevolutionary arms race? *Science* 212: 927–929.

Berenbaum MR, Favret C, and Schuler MA (1996) On defining "key innovations" in an adaptive radiation: Cytochrome P450s and Papilionidae. *Am. Nat.* 148(Suppl.): S139–S155.

Bernebaum M (1983) Coumarins and caterpillars: A case for coevolution. *Evolution* 37: 163–179.

Boake CRB, DeAngelis MP, and Andreadis DK (1997) Is sexual selection and species recognition a continuum? Mating behavior of the stalk-eyed fly *Drosophila heteroneura*. *Proc. Natl. Acad. Sci. USA* 94: 12442–12445.

Cameron RAD, Cook LM, and Hallows JD (1996) Land snails on Porto Santo: Adaptive and non-adaptive radiation. *Philos. Trans. R. Soc. London B.* 351: 309–327.

Carr GD (1985) Monograph of the Hawaiian Madiinae (Asteraceae): *Argyroxyphium, Dubautia, and Wilkesia*. *Allertonia* 4: 1–123.

Carr GD (1999) A reassessment of *Dubautia* (Asteraceae: Heliantheae-Madiinae) on Kauai. *Pacific Sci.* 53: 144–158.

Carson HL (1986) Sexual selection and speciation. In: Karlin S and Nevo E (eds.) *Evolutionary Processes and Theory*, pp. 93–107. New York: Academic Press.

Carson HL (1989) Natural hybridization between the sympatric Hawaiian Species *Drosophila silvestris* and *Drosophila Heteroneura*. *Evolution* 43: 190–203.

Carson HL (1990) Increased genetic variance after a population bottleneck. *Trends Ecol. Evol.* 5: 228–230.

Carson HL and Templeton AR (1984) Genetic revolutions in relation to speciation phenomena: The founding of new populations. *Annu. Rev. Ecol. Syst.* 15: 97–131.

Cox GW and Ricklefs RE (1977) Species diversity, ecological release, and community structuring in Caribbean land bird faunas. *Oikos* 29: 60–66.

Darlington Jr PJ (1957) *Zoogeography: The Geographical Distribution of Animals*. New York: Wiley.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection*. London: John Murray.

DeSalle R and Giddings LV (1986) Discordance of nuclear and mitochondrial DNA phylogenies in Hawaiian *Drosophila*. *Proc. Natl. Acad. Sci. USA* 83: 6902–6906.

DeSalle R, Brower AVZ, Baker R, and Remsen J (1997) A hierarchical view of the Hawaiian Drosophilidae. *Pacific Sci.* 51: 462–474.

Dommergues J-L, Laurin B, and Meister C (1996) Evolution of ammonoid morphospace during the early Jurassic radiation. *Paleobiology* 22: 219.

Ehrlich PR and Raven PH (1964) Butterflies and plants: A study in coevolution. *Evolution* 18: 586–608.

Farrell BD (1998) "Inordinate fondness" explained: Why are there so many beetles? *Science* 281: 555–559.

Farrell BD and Mitter C (1994) Adaptive radiation in insects and plants: Time and opportunity. *Am. Zool* 34: 57–69.

Farrell BD, Dussourd DE, and Mitter C (1991) Escalation of plant defense: Do latex and resin canals spur plant diversification? *Am. Nat.* 138: 881–900.
Freeland JR and Boag PT (1999) The mitochondrial and nuclear genetic homogeneity of the phenotypically diverse Darwin's ground finches. *Evolution* 53: 1553–1563.
Gillespie RG, Croom HB, and Hasty GL (1997) Phylogenetic relationships and adaptive shifts among major clades of tetragnathid spiders in Hawaii. *Pacific Sci.* 51: 380–394.
Gillespie RG, Rivera M, and Garb JE (1998) Sun, surf and spiders: Taxonomy and phylogeography of Hawaiian Araneae. In: Selden PA (ed.) *Proceedings of the 17th European Colloquium of Arachnology, Edinburgh 1997*, pp. 41–51. Burnham Beeches, Bucks, United Kingdom: British Arachnological Society.
Gittenberger E (1991) What about non-adaptive radiation? *Biol. J. Linnean Soc.* 43: 263–272.
Givnish T (1997) Adaptive radiation and molecular systematics: Issues and approaches. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 1–54. Cambridge, UK: Cambridge Univ. Press.
Givnish T and Sytsma KJ (1997) *Molecular Evolution and Adaptive Radiation*. Cambridge, UK: Cambridge Univ. Press.
Goodnight CJ (1988) Epistasis and the effect of founder events on the additive genetic variance. *Evolution* 42: 441–454.
Grant PR (1986) *Ecology and Evolution of Darwin's Finches*. Princeton, NJ: Princeton Univ. Press.
Hardy DE (1965) DipteraCyclorrhapha II, Series Schizophora, Section Acalypterae: I. Family Drosophilidae. *Insects Hawaii* 12: 1–814.
Hardy DE and Kaneshiro KY (1981) Drosophilidae of Pacific Oceania. In: Ashburner M, Carson HL, and Thompson JN (eds.) *The Genetics and Biology of Drosophila*, pp. 309–348. New York: Academic Press.
Harrison RG (1993) *Hybrid Zones and the Evolutionary Process*. Oxford: Oxford Univ. Press.
Huxley J (1942) *Evolution: The Modern Synthesis*. New York: Harper & Brother.
James HF and Olson SL (1991) Descriptions of thirty-two new species of birds from the Hawaiian Islands: Part II. Passeriformes. *Ornith, Monograph No. 46*. Washington, DC: American Ornithologists Union.
Johnson MS, Murray J, and Clarke B (1993) The ecological genetics and adaptive radiation of Partula on Moorea. *Oxford Surv. Evol. Biol.* 9: 167–238.
Kambysellis MP (1993) Ultrastructural diversity in the egg chorion of Hawaiian *Drosophila* and *Scaptomyza*: Ecological and phylogenetic considerations. *Int. J. Insect Morphol. Embryol:* 22: 417–446.
Kambysellis MP and Craddock EM (1997) Ecological and reproductive shifts in the diversification of the endemic Hawaiian *Drosophila*. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 475–509. Cambridge, UK: Cambridge Univ. Press.
Kaneshiro KY (1983) Sexual selection and direction of evolution in the biosystematics of the Hawaiian Drosophilidae. *Annu. Rev. Entomol.* 28: 161–178.
Kaneshiro KY (1988) Speciation in the Hawaiian *Drosophila*. *Bioscience* 38: 258–263.
Kaneshiro KY (1989) The dynamics of sexual selection and founder effects in species formation. In: Giddings LV, Kaneshiro KY, and Anderson WW (eds.) *Genetics Speciation and the Founder Principle*, pp. 279–296. Oxford: Oxford Univ. Press.
Klicka J and Zink RM (1997) The importance of recent ice ages in speciation: A failed paradigm. *Science* 277: 1666–1669.
Lack D (1947) *Darwin's Finches*. Cambridge, UK: Cambridge Univ. Press.
Lehmann U (1981) *The Ammonites*. Cambridge, UK: Cambridge Univ. Press.
Liebherr JK and Polhemus DA (1997) R. C. L. Perkins: 100 years of Hawaiian entomology. *Pacific Sci* 51: 343–355.
Liem KF (1973) Evolutionary strategies and morphological innovations: Cichlid pharyngeal jaws. *Syst. Zool* 22: 425–441.
Losos JB (1992) The evolution of convergent structure in Caribbean *Anolis* communities. *Syst. Biol.* 41: 403–420.
Losos JB, Jackman TR, Larson A, deQueiroz K, and Rodriquez-Scheltino L (1998) Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279: 2115–2118.
MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton, NJ: Princeton Univ. Press.
Margulis L and Fester R (1991) *Symbioses as a Source of Evolutionary Innovation*. Cambridge, MA: MIT Press.
McCune AR (1997) How fast is speciation? Molecular, geological, and phylogenetic evidence from adaptive radiation in fishes. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 585–610. Cambridge, UK: Cambridge Univ. Press.
McCune AR and Lovejoy NR (1998) The relative rates of sympatric and allopatric speciation in fishes. In: Howard DJ and Berlocher SH (eds.) *Endless Forms: Species and Speciation*, pp. 171–185. Oxford: Oxford Univ. Press.
Meyer A (1993) Phylogenetic relationships and evolutionary processes in East African cichlids. *Trends Ecol. Evol.* 8: 279–284.
Meyer A, Montero CM, and Spreinat A (1996) Molecular phylogenetic inferences about the evolutionary history of the East African cichlid fish radiations. In: Johnson T and Odada E (eds.) *International Decade of East African Lakes: The Limnology Climatology and Paleoclimatology of the East African Lakes*, pp. 303–323. New York: Gordon & Breach.
Mitter C, Farrell B, and Wiegmann B (1988) The phylogenetic study of adaptive zones: Has phytophagy promoted insect diversification. *Am. Nat.* 132: 107–128.
Moran P and Kornfield I (1993) Retention of an ancestral polymorphism in the Mbuna species flock (Teleostei: Cichlidae) of Lake Malawi. *Mol. Biol. Evol.* 10: 1015–1029.
Murray J, Clarke B, and Johnson MS (1993) Adaptive radiation and community structure of *Partula* on Moorea. *Proc. R. Soc. London B* 254: 205–211.
Nei M, Maruyama T, and Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution* 29: 1–10.
Nevo E, Filippucci MG, and Beiles A (1994) Chromosomal speciation and adaptive radiation of mole rats in Asia Minor correlated with increased ecological stress. *Proc. Natl. Acad. Sci. USA* 91: 8160–8164.
Norris RD (1996) Symbiosis as an evolutionary innovation in the radiation of Paleocene planktic foraminifera. *Paleobiology* 22: 461–480.
Osbora HF (1902) The law of adaptive radiation. *Am. Nat.* 36: 353–363.
Otte D (1994) *The Crickets of Hawaii: Origin, Systematics, and Evolution*. The Orthopterists' Society. Philadelphia: Academy of Natural Sciences.
Otto SP and Whitlock MC (1997) The probability of fixation in populations of changing size. *Genetics* 146: 723–733.
Petren K, Grant BR, and Grant PR (1999) A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proc. R. Soc. London B* 266: 321–329.
Rainey PB and Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature* 394: 69–72.
Raven PH, Evert RF, and Eichhorn SE (1992) *Biology of Plants*, 5th ed. New York: Worth.
Roderick GK and Gillespie RG (1998) Speciation and phylo-geography of Hawaiian terrestrial arthropods. *Mol. Ecol.* 7: 519–531.
Sato A, O'hUigin C, Figueroa F, Grant PR, Grant BR, Tichy H, and Klein J (1999) Phylogeny of Darwin's finches as revealed by mtDNA sequences. *Proc. Natl. Acad. Sci. USA* 96: 5101–5106.
Schluter D (1994) Experimental evidence that competition promotes divergence in adaptive radiation. *Science* 266: 798–801.
Schluter D (1996) Ecological causes of adaptive radiation. *Am. Nat.* 148: S40–S64.
Schluter D (1998) Ecological causes of speciation. In: Howard DJ and Berlocher SH (eds.) *Endless Forms: Species and Speciation*, pp. 114–129. Oxford: Oxford Univ. Press.
Schluter D and McPhail JD (1993) Character displacement and replicate adaptive radiation. *Trends Ecol. Evol.* 8: 197–200.
Schneider C and Moritz C (1999) Rainforest refugia and evolution in Australia's wet tropics. *Proc. R. Soc. London* 266: 191–196.
Shaw KL (1996) Sequential radiations and patterns of speciation in the Hawaiian cricket genus *Laupala* inferred from DNA sequences. *Evolution* 50: 237–255.
Shields O (1993) The origin and initial radiation of angiosperms in relation to Anthophytes. *Palaeobotanist* 42: 157–185.
Simpson GG (1953) *The Major Features of Evolution*. New York: Columbia Univ. Press.
Slatkin M (1996) In defense of founder-flush theories of speciation. *Am. Nat.* 147: 493–505.
Sloan RE, Rigby Jr JK, Van Valen LM, and Gabriel D (1986) Gradual dinosaur extinction and simultaneous ungulate radiation in the Hell Creek Formation. *Science* 232: 629–633.
Smith A (1984) *Palaeobiology*, London: Allen & Unwin.
Springer MS, Kirsch JAW, and Case JA (1997) The chronicle of marsupial evolution. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 129–161. Cambridge, UK: Cambridge Univ. Press.
Stauffer Jr JR, Bowers NJ, Kellogg KA, and McKaye R (1997) A revision of the blue-black *Pseudotropheus* zebra (Teleostei: Cichlidae) complex from Lake Malawi, Africa, with a description of a new genus and ten new species. *Proc. Acad. Nat. Sci. Philadelphia* 148: 189–230.

Sturmbauer C (1998) Explosive speciation in cichlid fishes of the African Great Lakes: A dynamic model of adaptive radiation. *J. Fish Biol.* 53(Suppl. A): 18–36.

Tarr CL and Fleischer RC (1995) Evolutionary relationships of the Hawaiian honeycreepers (Aves: Drepanidae). In: Wagner WL and Funk VA (eds.) *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*, pp. 147–159. Washington, DC: Smithsonian Institution Press.

Taylor EB, McPhail JD, and Schluter D (1997) History of ecological selection in sticklebacks: uniting experimental and phylogenetic approaches. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 511–534. Cambridge: Cambridge University Press.

Templeton AR (1980) The theory of speciation via the founder principle. *Genetics* 91: 1011–1038.

Vermeij GJ (1995) Economics, volcanoes, and Phaerozoic revolutions. *Paleobiology* 21: 125–152.

Wagner WL and Funk VA (1995) *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*. Washington, DC: Smithsonian Institution Press.

Whittaker RJ (1998) *Island Biogeography: Ecology, Evolution, and Conservation*. Oxford: Oxford Univ. Press.

Wilson EO (1961) The nature of the taxon cycle in the Melanesian ant fauna. *Am. Nat.* 95: 169–193.

Aesthetic Factors

Gordon H Orians, University of Washington, Seattle, WA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aesthetics The field of investigation, from philosophical and psychological perspectives, that attempts to discover the rules and principles that govern the sense of beauty and ugliness.

Beauty The qualities of a perceived or imagined object whereby it evokes feelings of admiring pleasure.

Cognition Act or process of knowing and understanding.

Habitat Type of environment in which an organism lives, grows, and reproduces.

Phobia Obsessive fear or dread of an object or situation.

Symmetry The correspondence in size, form, and arrangement of parts on opposite sides of a two- or three-dimensional object.

Aesthetic factors are those characteristics of a given object or situation that evoke a certain emotional response, either a sense of beauty, attractiveness, pleasure, symmetry, order, and so on, or, conversely, of ugliness, disorder, menace, disgust, or the like. Generally speaking, the aesthetic preferences that humans display in response to their environment, in such contexts as mate choice, foraging, and habitat selection, have been shaped by evolutionary experience and embody solutions for survival and reproductive success.

Introduction

Humans have strong emotional responses to living organisms and to natural and human-modified environments. Depending on the circumstances and needs of an individual, nature may evoke feelings of awe, respect, fear, loathing, longing, nostalgia, excitement, challenge, and belonging. These powerful emotions influence how we respond to nature, how we attempt to manipulate it, and why we care about it.

Strong emotional responses in all organisms, including humans, evolved because, on average, they increased fitness, that is, they improved the survival and reproductive success of the individuals that expressed them. For example, those of our ancestors who did not enjoy food and sex certainly were more poorly represented in subsequent generations than those who did enjoy and hence sought out food and sexual partners. Similarly, individuals who chose ecologically inferior environments in which to live should have been less well represented genetically in subsequent generations than individuals who made better habitat choices.

However, this truism does little more than establish general guidelines for investigating why human emotional responses to components of the environment are so varied. Nor does it take us very far in improving our understanding of how different emotional responses to varied environmental components resulted in fitness-improving behaviors. Specific theories and analyses are needed.

Evolutionary Approaches to Environmental Aesthetics

Why human emotional responses to nature would have influenced fitness is obvious. Our ancestors lived in environments devoid of modern conveniences. Their health, survival, and reproductive success depended on their ability to obtain and use environmental information wisely. They had to know how to interpret signals from the physical and biological components of the environment, and to adjust their behavior in response to them. They needed to evaluate habitats and their resources over space and time, and to adjust their use of habitats accordingly. Our ancestors also needed to be able to detect the presence of other humans, their effects on resources, and the direct dangers they posed or the help they might provide.

Making these evaluations and decisions was and still is difficult because the environment provides a far richer array of information than can possibly be assimilated and synthesized. Adaptive behavior requires selective attention to components of the environment that strongly influence fitness. In fact, the complex behavior of organisms would be impossible in the absence of neural filters that emphasize or deemphasize aspects of information emanating from the environment. These filters embody evolutionarily stored knowledge about the world that enables us to construct hypotheses capable of describing and understanding the external world. Extensive experiments on artificial intelligence have clearly shown that the rapid and efficient learning of a language performed by nearly every human child is quite impossible in the absence of preformed neural structures (Dennett, 1995).

These filters express themselves in all organisms as a sense of aesthetics and some form of logical analysis. In other words, aesthetic responses, rather than being a recently acquired capacity of little evolutionary significance, are fundamental to the ways in which organisms learn about the world and adapt to it. The English word *aesthetic* is derived from the adjectival form of the Greek *aisthanomai*, which means "to perceive." Thus, "aesthetic pleasure" means literally "pleasure associated with or deriving from perception." The central problem in aesthetics is to explain why pleasure and disgust have evolved to be so strongly associated with perception and recognition of certain kinds of objects.

The evolution of neural filters has led to what is called *biologically prepared learning* (Seligman, 1970). Biologically prepared learning theory asserts that evolution has predisposed humans, as well as individuals of other species, to learn easily and quickly, and to retain associations or responses that foster survival when certain objects or situations

are encountered. An important corollary is that even though modern societies have greatly reduced the real danger posed by many objects that evoke fears and phobias, fear and avoidance responses may nonetheless persist because selection against those responses is weak (Ulrich, 1993; Wilson, 1984). Biologically prepared learning in humans should be evident only for stimuli that have had significant influences on survival and reproductive success during our evolutionary history. Those responses should arise quickly when the appropriate conditioning stimuli appear, and they should be unusually resistant to extinction or forgetting. Adaptive responses to natural stimuli do not necessarily appear spontaneously or in the absence of learning, but they may.

Although aesthetic inquiries have been carried out since the middle of the eighteenth century, attention has been directed almost exclusively to responses to human creations rather than to nature. The neglect of nature is a by-product of the view that has dominated Western thinking for centuries – namely, that aesthetic experiences are molded primarily by cultural symbols and art forms. As a result, the study of aesthetics has been viewed as the domain of artists and philosophers. Attempts to explore the biophysical bases of aesthetic responses to environments were regarded, and still are by some people, as both futile and ideologically dangerous.

Culture and learning clearly exert strong influences on the ways humans perceive and respond to environmental information and they have important impacts on the symbolisms we attach to natural objects (Appleton, 1990; Dutton, 2009; Schama, 1995). But attempts to understand and interpret human aesthetic responses to environmental features without asking why they evolved have failed. People who studied the concept of beauty from a nonbiological perspective assumed that beauty was an intrinsic property of objects. They therefore looked for, and expected to find, correlations between the characteristics of objects and human aesthetic responses to them. This attempt largely failed because, as an evolutionary perspective immediately suggests, beauty is not an intrinsic property of the objects that we call beautiful. Rather, it is the product of interactions between traits of objects and the human nervous system that evolved so that we regard as beautiful those objects having properties that, if positively attended to, result in improved performance in some aspect of living (Appleton, 1975). Conversely, we regard as ugly those objects that should be avoided or destroyed. In essence, an evolutionary perspective suggests that concepts such as beauty and ugliness should be viewed from a functional rather than a structural perspective. In other words, emotional responses are best studied by asking "How did these responses help us solve problems?"

Environmental Information, Problem Solving, and Survival

Individuals of all species use information to make decisions that enhance their survival and reproduction. The array of information presented to an individual has components that are highly relevant to survival and reproductive success, and components that can be ignored with few or no adverse consequences. In addition, which components are important varies with time, location, and the needs of the individual.

Progress in the study of environmental aesthetics and problem solving has been aided by the development of systems for classifying environmental information into categories that roughly correspond to the kinds of decisions that organisms must make (Heerwagen and Orians, 1993). The basic problems that people (and other animals) must solve are: (1) protecting themselves from being injured or killed by other people or dangerous animals; (2) avoiding being attacked by parasites and disease-causing organisms that may be acquired from people or from other sources; (3) protecting themselves from adverse physical conditions; (4) acquiring enough of the right kinds of food; and (5) choosing good-quality associates for reproduction, foraging, protection, and achieving higher social status.

One useful classification divides information into categories based on the type of objects being identified. Some of these objects, such as food, water, and refuges, are resources. Gathering information about these resources should be pleasurable and exciting. Other objects, such as potentially dangerous animals, competitors, human enemies, and hazardous physical objects, pose dangers. Some of them are life-threatening, but others, although potentially dangerous, can be overcome. Gathering information about them should be accompanied by feelings of fear, anxiety, and apprehension rather than pleasure (**Table 1**).

Another classification of environmental information is based on the time frame over which the information is relevant. Although time is a continuous variable, humans long ago discovered that it is heuristically useful to divide time into categories. From the perspective of the study of environmental aesthetics, it is informative to divide time into categories that correspond to the time frames over which decisions about them matter (**Table 2**). Some environmental information

Table 1 Classification of information by type of object

Categogry	*Examples*	*Decision affected*
Inaminate objects		
Stationary	Landforms, glaciers	Bases of operations
Mobile	Rivers, ocean surface	Timings of foraging and exploratory activities
Animated objects		
Resources	Food items, fiber, medicine sources	Types, locations, and timing of foraging trips
Dangerous	Predators, large prey, venomous plants and animals, disease-causing, or disease-vectoring animals	Types, locations, and timing of foraging trips

Table 2 Classification of information in terms of the time frame of its relevance

Category	*Examples*	*Decision affected*
Short term (minutes to hours)	Weather changes (thunderclouds, wind)	Seeking shelter, initiating outdoor activities
	Appearance of dangerous animals, enemies	Immediate defensive actions
	Appearance of valuable prey illumination changes	Immediate hunting activities moves to appropriate locations for spending the night
Seasonal	Day-length changes, vegetative growth, flowering, precipitation changes	Shifts of hunting sites, planting and harvesting of crops
Multiyear changes	Vegetation succession, erosional changes (river meanders, lake sedimentation)	Shifts of hunting sites, movement of villages
Long term (decades to centuries)	Topography	Development of tradition

Source: Reproduced from Orians GH (1998) Human behavioral ecology: 140 years without Darwin is too long. *Bulletin of the Ecological Society of America* 79: 15–28.

signals events of immediate, temporary significance. Still other information signals seasonal changes that are associated with shifts in the types and locations of resources that will be available in the near future. Still other information signals features of the environment, such as status of vegetation and courses of rivers, that change slowly over periods of years or decades. Finally, other information arrives from objects that, measured in terms of human lifetimes, are permanent.

A third classification is based on the sensory mode in which the information arrives. What information can be perceived obviously depends on the sensory capacity of an organism and its neural processing system. The first, and still the most important sensory capacity in the living world as a whole, is chemical sensitivity. Molecules arrive by diffusion, augmented by movements of the medium. All chemicals are sensed when they come into contact with an organism's body, but, based on how fast they are carried by air currents, we distinguish verbally between the senses of taste and smell. The chemical knowledge possessed by a single bacterial cell exceeds what the most sophisticated organic chemist can measure. And because the sense of taste is based on chemical reactions, noses can be scaled down almost to the size of molecules.

Light sensitivity also evolved early in life's history on Earth. Bacteria evolved both photosensitivity – light-sensitive spots – and photosynthesis. The former vastly increased the distance over which objects could be detected; the latter generated nearly all the energy in today's biosphere. Because light travels in straight lines, at the speed of light, vision potentially provides accurate and time-specific information about shapes of objects and their location in space. Sound has similar properties to light, but location of objects is not given precisely. Animals have evolved complex ways of inferring locations of sources of sounds by comparing events at the two ears.

Nervous systems evolved in part because, by building up parallel computing units, they enabled organisms to speed up and amplify responses to sensory inputs. For motile organisms, speed of response is often extremely important. Parallel computing units also are the units in which neural filters are stored.

Responses to environmental information are complex because the significance of objects and events varies with the location and needs of observers. An approaching storm may be welcomed by a farmer whose crops are wilting from lack of rain, but despised by members of a family enjoying their first summer picnic. A vivid sunset enjoyed by people relaxing on the deck of their home may bring fear to a person walking on an African savanna too far from home to reach it before dangerous nocturnal predators become active. For these reasons, simple correlations between features of objects and aesthetic evaluations are unlikely ever to be discovered, although correlations between types of objects, contents, and aesthetic responses are likely.

Three functional concepts – prospect, refuge, and hazard – have guided recent approaches to the role of environmental aesthetics in problem solving (Appleton, 1975). *Prospect* refers to the ability of an individual to gather information about an environment with which to evaluate its characteristics and decide how to use it. Environments high in prospect offer rich opportunities for evaluation; environments low in prospect offer fewer opportunities. *Refuge* refers to the degree to which an environment provides security for an individual from negative agents while it is exploring and gathering information. *Hazard* refers to the dangers to which an individual would be exposed during information-gathering activities.

These concepts have been applied primarily to the initial evaluation and exploration of unfamiliar environments, but when combined with more recent developments in evaluating environmental information, they are readily applied and extended to a rich array of circumstances in which people must solve problems.

Emotional responses that assist in solving these problems necessarily include both positive and negative components. Because they are simpler and easier to study, the author discusses negative responses first.

Negative Emotional Responses

Many people have strong fearful reactions – phobias – to a wide variety of objects or situations (snakes, spiders, heights, closed spaces, open spaces, blood). The set of objects that evoke fearful responses appears to be the same in all industrialized societies for which data are available (Ulrich, 1993), but few data exist on the prevalence of phobias in developing nations and nontechnological societies.

The objects and situations that evoke fearful responses are generally ones that have been associated with threatening

situations during human evolutionary history. Precipitous cliffs are dangerous if approached too closely. Closed spaces offer few escape routes; people in wide open spaces are vulnerable to attacks by enemies and dangerous predators. Venomous and predatory animals have been significant sources of human injury and mortality for many millennia (Isbell, 2009). Interestingly, even though they account for trivial numbers of injuries and deaths in modern societies, spiders and snakes are the objects of the most prevalent phobias in Western societies today.

Research on human twins has provided convincing evidence that genetic factors play major roles in a wide range of human traits, including animal phobias and fear of open spaces. Imaginative research, especially by psychologists in Sweden and Norway, shows that biologically prepared learning plays a significant role in the acquisition and retention of phobias (see Ulrich, 1993, for an overview of experimental results). The most informative experiments employ a Pavlovian conditioning approach to compare the development of defensive and aversive conditioned responses (learned through repeated exposure or experience) to slides of fear-relevant and fear-irrelevant or neutral stimuli. Defensive responses are assessed by recording autonomic nervous system indicators such as skin conductance and heart rate.

Experimenters condition initial defensive responses by showing either fear-relevant (snakes or spiders) or neutral stimuli (geometric figures) paired with an aversive stimulus (the unconditioned stimulus), usually an electric shock intended to mimic a bite. This phase allows experimenters to compare the speed and strength of acquisition of defensive/aversive responses to fear-relevant and neutral stimuli.

Following an initial acquisition phase, the same stimuli are presented many times without electric shock reinforcement. This "extinction" phase allows comparison of the rate of extinction of the defensive/aversive response acquired earlier. The general result is that conditioned responses are usually, but not always, acquired more quickly, but that responses to snakes and spiders are reliably more resistant to extinction than responses to neutral stimuli. These responses cannot be the result of prior cultural reinforcement, because conditioned aversive responses to familiar modern dangerous stimuli, such as hand guns and frayed electrical wires, extinguish more rapidly than conditioned responses to snakes and spiders. In addition, aversive responses to fear-relevant natural stimuli can be elicited merely by telling a person that a shock will be administered. Aversive responses to fear-irrelevant natural stimuli cannot be elicited in this manner.

People also acquire much more persistent defensive reactions when watching an experimenter's reactions to fear-relevant than to fear-irrelevant stimuli. Similar results have been obtained by exposing rhesus monkeys to fear-relevant (toy snakes and crocodiles) and fear-irrelevant stimuli (toy rabbits).

Even more striking are the results of "backward masking" experiments, in which slides are displayed subliminally (for 15–30 ms) before being "masked" by a slide of another stimulus or setting. Even though the subjects are not consciously aware of having seen the stimulus slide, presentations of natural settings that contain snakes or spiders elicit strong aversive/defensive reactions in nonphobic persons. If a previous conditioning has already occurred, a masked subliminal presentation is sufficient to elicit defensive responses to the feared stimulus (Öhman and Soares, 1994).

Thus, a rich array of experimental results demonstrates that aversive responses develop more rapidly and persist longer to fear-relevant than to fear-irrelevant natural stimuli. In addition, defensive/aversive responses can develop to natural threat stimuli even though subjects are unaware that they have seen them. Such responses make adaptive sense. Threatening events need to be responded to quickly and automatically. Pleasures can be contemplated at leisure. It is better to respond quickly and sometimes be wrong than to be slow and mostly right! The lives or our ancestors, as they pursued food and sought to avoid being food for other animals, have left a strong, negative imprint in our brains (Rozin and Royzman, 2001).

Positive Emotional Responses

Research and theory are less well developed for positive emotional responses to environmental stimuli than for negative ones. This is in part due to the fact that positive conditioning experiments are typically more difficult to conduct than aversive conditioning experiments. In addition, positive emotional responses are very complex and difficult to quantify. Nevertheless, a significant body of literature is slowly accumulating.

Symmetry and Beauty

Much evidence has been gathered on emotional responses to the symmetry of objects. A rich body of literature demonstrates that humans find symmetrical objects, including abstract patterns, woody plants, and human bodies – especially faces – aesthetically pleasing. Moreover, sexual preferences for symmetrical facial features appear to be similar across human cultures. Other vertebrates also are sensitive to very small asymmetries (see Møller and Swaddle, 1997, for a review of the human and animal literature). Conversely, asymmetries typically evoke negative emotional responses, the most extreme of which is selective infanticide of asymmetrical newborn babies in many human cultures.

There are several nonmutually exclusive reasons why positive responses to symmetry may be adaptive. One is that symmetrical animals perform better physically than asymmetrical animals, just as symmetrical objects, such as bows, arrows, axes, boats, autos, and airplanes, function better than asymmetrical ones. Asymmetrical individuals may have been exposed to severely stressful environments that disrupted normal development, thereby rendering them less functionally adequate. Alcoholic mothers give birth to children with greater developmental asymmetries than do mothers who consumed less alcohol. Thus, it is not surprising that choices of objects to use, the design of objects, and choices of social partners have evolved to favor use of and association with symmetrical objects.

Second, asymmetrical individuals are likely to have genetic defects that could be passed on to offspring. For example,

humans suffering from trisomy-21 (Down syndrome) have noticeable asymmetries due to skeletal abnormalities. Greater asymmetry in dental traits is associated with genetic disorders. Therefore, avoiding as mates asymmetrical individuals would be likely to improve the quality of an individual's offspring.

Third, because diseases can cause asymmetrical development, asymmetrical individuals may be avoided because they may currently harbor communicable disease-causing microorganisms. Good symmetry signals a long history of good health!

Habitat Selection

Habitat selection is a vital decision in the lives of all organisms. When selecting a habitat an organism responds as *if* it understood the significance of objects, sounds, and odors for its future survival and reproductive success. Initial responses typically are emotional feelings that lead to rejection, exploration, or a certain use of the environment. Because the strength of these responses strongly influences immediate decisions about where to settle and what to do there, the emotional responses evoked by habitats should evolve to be positively correlated with the expected survival and reproductive success of an organism in those habitats. That is, "good" habitats should evoke strong positive responses; "poor" habitats should evoke weak or even negative responses.

Multiple spatial scales are important during habitat selection. At small spatial scales, the primary decisions are whether to accept or reject a specific food item or shelter. At medium spatial scales, the key decisions concern which activity to engage in and in which patches in the environment those activities will be carried out. At still larger spatial scales, decisions center on whether to initiate major changes in where to carry out activities, that is, to migrate or to shift the base of operations.

Theoretical explorations of these decisions typically include the development of "ideal" models, that is, models that assume perfect knowledge on the part of the decision maker. These models specify the best that an organism can do in solving a problem (e.g., obtaining the most food per unit hunting time). They are unrealistic because knowledge is always incomplete, but they are valuable for assessing the marginal value of additional information and for determining how well an animal using rather simple "rules of thumb" would perform compared to an "ideal" animal. Many of these rules of thumb include aesthetic responses that guide decisions. It is useful to begin with a consideration of medium-scale responses and then turn to microscale responses.

Habitat selection, viewed at medium spatial scales, has served as a perspective for a number of studies on human aesthetic responses to landscape features (Heerwagen and Orians, 1993). During most of human evolutionary history habitats rarely provided long-term reliable resources that enabled permanent occupation of sites. Frequent moves through the landscape were the rule even though traditional sites might be revisited on an annual basis. Because relatively few generations have passed since humans began to live in mechanized and urban environments, our evolutionarily based response patterns to landscapes are unlikely to have been substantially modified since the rise of industrialized, urban societies.

Human responses to environmental cues vary with a person's age, social status, and physiological condition. Nevertheless, positive responses to indicators of the presence of food, water, shelter, and protection from predators are widespread. So are negative responses to potential hazards, such as inclement weather, fire, dangerous predators, and barriers to movement. Although no direct evidence yet exists for genetic influences on these responses, a number of evolutionary hypotheses have generated predictions, some of which have been tested experimentally.

One approach is based on the fact that *Homo sapiens* evolved in African savannas and only recently has invaded other continents. Therefore, landscape features and tree shapes characteristic of high-quality African savannas are expected to be especially attractive to humans today. This hypothesis has been tested by determining the responses of people to tree shapes and by examining the features of "aesthetic environments," that is, those environments, such as parks and gardens, that are designed to make them attractive (Orians, 1986).

The shapes of trees that dominate savannas are good predictors of the resource-providing capacities of those environments. Therefore, people evolved to find more pleasing the shapes of trees that were prominent in environments that provided the highest-quality resources rather than shapes of trees that dominated poor-quality habitats (Coss, 2003). Trees that grow in the highest-quality African savannas have canopies that are broader than they are tall, trunks that bifurcate close to the ground, and layered canopies. People on three continents preferred Kenyan *Acacia tortilis* trees that had highly or moderately layered canopies, lower trunks, and higher canopy width/tree height ratios than trees with narrow canopies and trunks that bifurcated higher above the ground (Heerwagen and Orians, 1993). Similarly, college students in Australia, Brazil, Canada, Israel, Japan, and the United States preferred trees with broad spreading crowns over conical and columnar trees (Sommer and Summit, 1996).

Changes that landscape architects recommend to their prospective customers are another source of data on human responses to landscapes. Humphrey Repton, an eighteenth-century British landscape architect, presented his clients with "before" and "after" drawings of their estates (Repton, 1907). Because he presumably wished to encourage rather than discourage potential clients, Repton changed landscapes by creating more savanna-like scenes, by increasing visual access and penetrability of closed woods, by opening up distant views to the horizon (i.e., increased prospect), by adding refuges and cues signaling ease of movement, and by adding evidence of resource availability, particularly large mammals (Heerwagen and Orians, 1993).

Although a love of flowers is a pervasive human trait, it is not obvious why an omnivorous primate should take flowers to hospitals, bring them to dinner parties and house warmings, and annually spend billions of dollars on them. Nor is it obvious, given that flowers did not evolve their forms and colors to please us, why we should find them so aesthetically attractive. An evolutionary perspective suggests that flowers evoke strong positive feelings because they have long been associated with food resources. Flowers precede fruits, so

flowering plants provide excellent cues to timing and locations of future resources. In addition, flowers may attract animals, especially birds, that are potential human prey. In species-rich environments, paying attention to flowering plants may particularly enhance resource acquisition abilities in the future. Until the nineteenth century, honey was the only natural source of sugar; beekeeping is an ancient human enterprise.

No studies have investigated which traits of flowers evoke strong positive feelings, but the obvious changes produced in many species of flowers by artificial selection – increased size and duplication of floral parts – result in flowers similar to those that historically produced large nectar rewards. Anecdotal evidence suggests that strongly asymmetrical flowers, which usually produce large nectar rewards, are generally more attractive than symmetrical ones, suggesting a possible exception to the general human preference for symmetrical objects. Future studies of human aesthetic responses to flowers are likely to provide interesting results.

Foods and Food Selection

Eating is essential for survival. Foraging animals make two major types of decisions: what items are acceptable as food, and which acceptable items are actually eaten when encountered. Aesthetic responses are especially prominent in the first decision. The strong emotional responses associated with foods are not surprising because eating requires voluntary ingestion of foreign objects. To eat, the body's generally aversive responses designed to protect health must be overcome; vomiting and diarrhea, neither of them pleasant, are the primary postingestion defenses. These decisions are especially difficult for omnivores, who must ingest a variety of foods to achieve a balanced diet in varied environments where the array of available foods differs dramatically over space and time. Omnivores are expected to evolve only general food recognition mechanisms and should develop systems for classifying organisms with respect to their potential value as food.

People accept or reject foods for complex reasons (**Table 3**). Direct sensory responses – tastes good or tastes bad – are important, but they are only one component of acceptability. People also make decisions on the basis of anticipated consequences, both physiological and social, of eating a type of food. In addition, particular foods may acquire ideational features that limit their acceptability to particular situations or cause them to be categorically rejected. Because of the variety of reasons that influence which foods are in the set of what is acceptable, human cuisines differ more than would be expected simply from knowing what edible resources are available in the environments in which social groups live (Rozin, 1996, 2007).

Despite the idiosyncratic nature of human diets, a few generalizations have emerged. First, most items that evoke disgust are of animal origin. Plant parts rarely evoke disgust in any culture. However, "inappropriate" items are primarily vegetable in origin (Rozin and Fallon, 1981). Disgust may be an adaptation that deterred our ancestors from eating animal tissues, such as feces, rotting meat, and soft internal parts, all of which commonly harbor large numbers of potentially harmful microorganisms. Microorganisms have the ability to multiply rapidly, so there is no safe dose for ingesting them. Disgust may represent evolutionarily programmed intuitive microbiology.

Perhaps the most interesting feature of human cuisines is the stability of the major spices and sauces that characterize them (Nabhan, 2004). Humans are remarkably conservative in their food habits and are typically reluctant to try new foods or to abandon familiar ones. Traditional flavorings are high-priority culinary items; immigrant groups go to great lengths and expense to procure them in foreign settings. The deliberate manipulation of food by adding ingredients that reliably alter its taste is a uniquely human behavior. No other animals are known to do so. Which flavorings are used probably evolved in relation to what was available in the environment, but once established they are remarkably persistent. Certain flavorings probably signal that the food is safe. For example, the French refused to eat potatoes until Parmientier, an eighteenth-century French agriculturist, showed how to prepare them in familiar ways with familiar seasonings: butter, cheese, and herbs.

Cross-cultural similarities in which spices are used exist because spices inhibit or kill food-spoiling microorganisms. The most widely used spices all have strong antimicrobial properties (Billing and Sherman, 1998); mixes of them, which are common in many traditional recipes, are even more powerful. Not surprisingly, given that disease-causing organisms are more abundant in tropical than in temperate regions, the proportion of traditional recipes containing antimicrobial spices is inversely correlated with latitude.

Table 3 Reasons people accept or reject foods

Primary reason	*Examples*	*Rejection*
A. Sensory	Tastes good	
B. Anticipated consequences		
Short-term physiological effects	No adverse reaction	Allergic reaction, gastrointestinal upsets
Long-term physiological effects	Societal concepts of "healthy foods"	Social concept of "unhealthy foods"
Social effects	Results in acceptance by group (smoking, social drinking)	Causes rejection by group (foods not eaten by that group, e.g., pigs in some Mideast cultures)
C. Ideational		
Appropriateness	Appropriate ritual foods	Inappropriate nonedible contaminants (soil, weeds)

Source: Modified from Rozin P and Fallon A (1981) A perspective on disgust. *Psychological Review* 94: 23–41.

Alternative hypotheses – that spices provide micronutrients, disguise the taste and smell of spoiled foods, or increase perspiration and, thus, evaporative cooling – are not supported by the extensive data base (Billing and Sherman, 1998).

Restorative Responses

If aesthetic responses evolved because they enabled people to better solve life's problems, exposure to high-quality environments should be restorative, that is, it should reduce feelings of tension and stress. Stress reduction consistently emerges as one of the key benefits reported by recreationists in wilderness areas. Restoration from stress is also reported as a key benefit from time spent in urban parks with savanna-like vegetation and water (Schroeder, 1989). Patients recovering from surgery in hospitals with either views of natural vegetation or simulated views that depict natural scenes with water recover more rapidly and have less postoperative anxiety than patients with no access to natural views or who are presented with simulations of abstract designs. Many studies have shown that even a brief exposure to nature, real or *via* photographs, leads to positive emotional feelings, reductions in stress, and better performance on demanding tasks (see Ulrich, 1995, for a review). Clearly, the people's affiliative responses to nature have important implications for the design of work and living spaces and healthcare facilities that are just beginning to be implemented.

Aesthetics and Biodiversity

Although people are strongly attracted to living organisms, relationships between the attractiveness of an environment and the number of species in it are unclear. On the one hand, for example, the most highly evolved garden traditions – European formal gardens and Japanese gardens – are based on just a few species of woody plants. Landscape designers generally do not like the gardens of botanists because they are cluttered up with plants of too many species! Scenes of environments that contain a jumble of plants of many species receive low scores in psychological tests. People report that they are too difficult to interpret; it is difficult to determine how to enter and use them.

On the other hand, people take great pleasure in finding as many species of birds as they can on a given day and from assembling "life lists" of species they have seen. Generally, seeing more species is better than seeing fewer species. Journeys of hundreds or even thousands of miles to see a rare species not on one's life list are not uncommon in today's mobile society. People are also powerfully attracted to the unusual – rare species, individuals outside the normal range of the species, or individuals present at unusual times of the year or in unusual habitats.

Because these familiar patterns of human behavior have been subjected to remarkably little formal study, we can only speculate about why they have evolved. That environments with intermediate levels of biological complexity should be preferred over both simpler and more complex environments makes sense, because the range of resources present in an environment and the ability to find and use those resources probably peak at intermediate levels of complexity. Simple environments have too few resources; complex ones have so many that choosing among them becomes difficult. Developing a suitable classification system to guide responses to the components of complex environments may be especially difficult. People may have evolved to respond to rare and unusual events because they provide new information about the state of the environment. Not all novel events are associated with something important, but it may be best to pay attention to them to find out if they are rather than to ignore the signals. Novel events may indicate that current patterns of environmental use should be altered.

Aesthetics and Cognition

The human mind evolved into its current form long before the invention of agriculture and the dawn of the Industrial Revolution. It evolved its special characteristics in the service of our hunter-gatherer ancestors, who almost daily faced serious challenges from their physical, biological, and social environments. During the past century, many scientists found it difficult to explain why the demands of functioning in a preagricultural environment should have favored the evolution of the vast complexities of the human mind, which is capable of the many feats of which we are so proud. The apparent paradox arises because of a failure to appreciate that succeeding as a forager is probably more complicated than playing chess or doing calculus. To outwit nature, people need to use some form of intuitive scientific thinking, develop abstract conceptions, think about the future, compete in a rich social environment, and plan tactics. Aesthetic responses are basic components in all of these responses.

Mathematicians describe their search for important theorems as an aesthetic experience; they think of proven theorems as being beautiful. In fact, scientists in all fields regularly describe their models, experiments, and results in aesthetic terms. Thus, aesthetic sensibilities are imbedded deep in the human mind, having evolved in the service of choice of mates, habitats, and foods. An aesthetic sense functions as a holistic filter that helps the mind search efficiently for good solutions. It can do this because the filter has been molded by countless evolutionary experiences. Perhaps one of the most satisfying results of having a mind that was molded by natural selection is that we generally enjoy doing things that are good for us to do.

See also: Historical Awareness of Biodiversity. Human Impact on Biodiversity, Overview. Literary Perspectives on Biodiversity. Social and Cultural Factors. Tourism, Role of

References

Appleton J (1975) *The Experience of Landscape.* New York: John Wiley & Sons.
Appleton J (1990) *The Symbolism of Habitat: An Interpretation of Landscape in the Arts.* Seattle: University of Washington Press.

Billing J and Sherman PW (1998) Antimicrobial functions of spices: Why some like it hot. *Quarterly Review of Biology* 73: 3–49.
Coss RG (2003) The role of evolved perceptual biases in art and design. In: Voland E and Grammer K (eds.) *Evolutionary Aesthetics*, pp. 69–130. Heidelberg: Springer-Verlag.
Dennett DC (1995) *Darwin's Dangerous Idea. Evolution and the Meaning of Life.* New York: Simon & Schuster.
Dutton D (2009) *The Art Instinct. Beauty, Pleasure, and Human Evolution*. New York: Bloomsbury Press.
Heerwagen JH and Orians GH (1993) Humans, habitats, and aesthetics. In: Kellert SR and Wilson EO (eds.) *The Biophilia Hypothesis*, pp. 138–172. Washington, DC: Island Press.
Isbell Lynne A (2009) *The Fruit, The Tree, and The Serpent. Why We See So Well.* Cambridge, MA: Harvard University Press.
Møller AP and Swaddle JP (1997) *Asymmetry, Developmental Stability, and Evolution*. New York: Oxford University Press.
Nabhan Gary Paul (2004) *Why Some Like it Hot. Food, Genes, and Cultural Diversity.* Washington, DC: Island Press.
Öhman A and Soares JJF (1994) Unconscious anxiety: Phobic responses to masked stimuli. *Journal of Abnormal Psychology* 103: 231–240.
Orians GH (1986) An ecological and evolutionary approach to landscape aesthetics. In: Penning-Rowsell EC and Lowenthal D (eds.) *Landscape Meanings and Values*. London: Allen & Unwin.
Orians GH (1998) Human behavioral ecology: 140 years without Darwin is too long. *Bulletin of the Ecological Society of America* 79: 15–28.
Repton H (1907) *The Art of Landscape Gardening.* New York: Houghton Mifflin.
Rozin P (1996) Towards a psychology of food and eating: From motivation to module to model to marker, morality, meaning, and metaphor. *Current Directions in Psychological Science* 5: 18–24.
Rozin P (2007) Food and eating. In: Kitayama S (ed.) *Handbook of Cultural Psychology*, pp. 391–416. New York: The Guildford Press.
Rozin P and Fallon A (1981) A perspective on disgust. *Psychological Review* 94: 23–41.
Rozin P and Royzman EB (2001) Negativity bias, negativity dominance, and contagion. *Personality ad Social Psychology Review* 5: 296–320.
Schama S (1995) *Landscape and Memory.* New York: Viking Books.
Schroeder HW (1989) Environment, behavior, and design research on urban forests. In: Zube EH and Moore GT (eds.) *Advances in Environment Behavior and Design*, vol. 2, pp. 87–117. New York: Plenum Press.
Seligman MEP (1970) On the generality of the laws of learning. *Psychological Review* 77: 406–418.
Sommer R and Summit J (1996) Cross-national rankings of tree shape. *Ecological Psychology* 8: 327–341.
Ulrich RS (1993) Biophilia biophobia, and natural landscapes. In: Kellert SR and Wilson EO (eds.) *The Biophilia Hypothesis*, pp. 73–137. Washington, DC: Island Press.
Ulrich RS (1995) Effects of healthcare interior design on wellness: Theory and recent scientific research. In: Marberry SO (ed.) *Innovations in Healthcare Design*, pp. 88–104. New York: Van Nostrand–Reinhold.
Wilson EO (1984) *Biophilia*. Cambridge, MA: Harvard University Press.

Africa, Ecosystems of

J Michael Lock, Royal Botanic Gardens, London, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Basement complex The layer of highly metamorphosed sedimentary rocks, mainly of late Pre-Cambrian age, that underlies much of tropical Africa.
Catena A regularly recurring sequence of soil and vegetation types developed on an undulating landscape.
Fynbos Habitat type in southern Africa that is characterized by thickets and low shrubs, in which fire plays a dominant role in ecosystem maintenance. Plant endemism is particularly high in these areas.
Inselberg An isolated rock or rocky hill, often of great size (from the German: 'Island Mountain').
Miombo Woodland habitat type widespread in south-central Africa, characterized by numerous species of the tree genus *Brachystegia* and *Isoberlinia*, which form nearly closed canopies. Fire is an annual event in this habitat, which supports relatively low populations of large mammals.
Phytochorion (pl. phytochoria) Region within which a substantial proportion of the flora is endemic.
Savanna Tropical seasonal vegetation comprising an upper layer of trees of low stature and a lower layer made up largely of grasses. Fires are a regular feature.

The huge continent of Africa straddles the equator, extending to 37° N and 35° S. It has no marginal oceanic trenches and subduction zones and so lacks the extensive mountain ranges of the Americas, but much of the southern half of the continent is a high plateau close to 1000 meters above sea level, broken only by the southern extension of the Great Rift Valley and the somewhat lower basin of the Congo (Zaïre) River. Its great latitudinal range gives it an enormous variety of climates, and this variability is reflected in an extreme diversity of ecosystems. There are three major climatic zones: two at the extreme north and south of the continent, where the main rainfall season is the winter, and one in central Africa, where the rains fall mainly in the hot summer season. The two winter-rainfall areas are distinct from each other and from the tropical parts of Africa. These regions are briefly considered here; a fuller treatment appears in the entry for Mediterranean Ecosystems.

Introduction

White (1983) classified the vegetation of Africa into a number of phytochoria – regions within which a substantial proportion of the plants are endemic. These regions are also useful in helping to define zoogeographical regions, and comparisons can also be made with neighboring continents. Thus the Mediterranean phytochorion of North Africa is most closely related floristically to southern Europe and the Middle East. The Somali–Maasai phytochorion, which occupies the Horn of Africa and regions south to central Tanzania, shows floristic relationships with parts of Arabia and the western part of the Indian subcontinent. On the other hand, about 80% of the plant species that occur in the Guineo–Congolian phytochorion, which occupies the Congo Basin and extends westwards to Liberia, are endemic. The Cape Region, the second winter-rainfall area, is also extremely rich in species, almost all of which occur nowhere else.

The large island of Madagascar, off the southeast coast of Africa, has been isolated from the rest of the continent for over 100 million years and its flora and fauna are very different indeed from those of Africa. The ecology of Madagascar is treated, in brief outline (*see* Madagascar).

Major Environmental Factors of Continental Africa

Geology

Much of southern tropical Africa is a plateau close to 1000 m above sea level, most of which has not been submerged since the Tertiary period. This plateau is formed of Pre-Cambrian ("basement complex") rocks, mainly igneous and much metamorphosed, which have been deeply weathered. The topography of much of the plateau is gently rolling, often with isolated rocky inselbergs, and sometimes with flat-topped hills capped with secondary ironstone, a product of prolonged weathering in high temperatures.

In eastern Africa the plateau has been cut by the two branches of the Great Rift Valley, which extends from Israel to the Zambezi River. Rifting seems to have begun in the Miocene, and was (and still is) accompanied by volcanic activity. In Ethiopia great sheets of basalt were erupted and now form the Ethiopian Highlands. Scattered volcanoes line the branches of the Rift Valley; some, such as Elgon and Kilimanjaro, are now dormant, and others, such as Ol Doinyo Lengai (Tanzania) and Nyiragongo (Zaïre), are still active. The Rift Valley contains numerous lakes, some relatively shallow and saline (e.g., Lake Chala (Ethiopia), Lakes Turkana and Bogoria (Kenya), and Lakes Natron and Eyasi (Tanzania)) and others deep and freshwater (Lakes Albert and Edward (Uganda/Zaire), Lake Tanganyika, and Lake Malawi). The blocks between the branches of the Rift Valley have tilted as part of the same tectonic disturbances. This has impeded or even reversed

the flow of rivers, and led to the formation of extensive shallow lakes, such as Lake Victoria, and huge areas of swamp.

Climate

The climate of Africa (excluding the extreme north and south) is determined by the movement of the Inter-Tropical Convergence Zone (ITCZ). As the apparent position of the sun moves north and south with the seasons, the area where the sun is immediately overhead is heated more than the areas to the north and south. The heated air rises, and air is drawn in from north and south to replace it. The convergence of these two air masses, and the rising of the heated air, causes rain. The rain belt tends to lag behind the sun, so the rain belt reaches its farthest north in July–August, rather than at the summer solstice (June 21st), and its farthest south in January–February, not at the winter solstice (December 21st). Away from the equator, the rainfall produced by the ITCZ as it moves in one direction merges into that produced during its return, so that there is a single rainy season. Near the equator, however, there is a tendency for the two passages of the ITCZ to be separated by a dry season, giving two rainy seasons and two dry seasons in each year, but in the wettest regions these gaps hardly occur so that rainfall occurs throughout most of the year.

Superimposed on this basic seasonal pattern is variation in the total annual rainfall. In general, totals decline away from the equator, and also from west to east. In East Africa, air coming from the north-east has passed over the dry Arabian landmass, not the sea as in most of the rest of the continent, and this explains the lower rainfall. Areas close to the sea tend to be wetter than inland areas, except where, as in south-western Africa, cold sea currents offshore produce foggy conditions but low rainfall. Extremes include Sierra Leone, where the coastal capital, Freetown, has over 250 cm of rain each year, most of it falling in four months. Debunscha Point, on the seaward margin of Mt. Cameroon, receives over 10 m of rain in some years; this comes from the combination of a warm sea and rapidly rising ground close to the shore. Rainfall totals increase with altitude up to about 3000 m and then decline. At middle altitudes, the rain is supplemented by moisture condensing from clouds onto leaves and dripping to the ground; here humidity is extremely high for much of each day. Above the cloud layer, however, the climate is drier, and the upper regions of Mt. Kilimanjaro are a virtual desert with probably less than 25 cm of rain each year.

Soils

The soils of much of Africa are developed on ancient land surfaces and are the product of long weathering. Forest soils, developed under high rainfall, are generally very strongly leached, rich in clay, and yellowish brown in color. Drier forests have redder, somewhat less nutrient-poor soils. Savanna soils are often brownish and, again, rich in clay. On the older plateaux there are often regular sequences of soils ("catenas"), with leached sandy or gravelly soils with kaolinite clay minerals on the hill and ridge tops, finer-grained soils on the hill slopes, and deep, dark-colored clays in the valleys. These valley soils, which shrink and crack as they dry and swell again when they are wetted, become extremely sticky and difficult to work when they are seasonally flooded or waterlogged. These are the so-called "cotton-soils," rich in montmorillonite clay minerals, in which motor vehicles so easily become bogged.

As a result of prolonged weathering, most of the soils of the African plateau are nutrient-poor, with phosphorus in particular often being in short supply. Some of the more volatile elements, such as nitrogen and sulfur, can be lost in the smoke in the regular dry season fires.

In a few areas, such as the "copperbelt" of southern Zaïre, northern Zambia, and the Great Dyke in Zimbabwe, there are outcrops of soils rich in metals such as copper, cobalt, and chromium. These produce soils that are toxic to many plants. Often trees are scarce or absent, and the vegetation is mainly composed of grasses and a range of specialized herbs, often very local in their distribution.

Fire

The highly seasonal climates of most of tropical Africa provide ideal conditions for vegetation fires. During the wet season, grass growth is rapid, but during the dry season the aerial parts of the grasses dry out and burn very easily. Most of those parts of Africa with a seasonal climate are burned by vegetation fires every year. The exceptions are areas with a rainfall too low to produce the necessary volume of fuel, or so high that the tree cover is continuous and dense, preventing the growth of enough grass to provide fuel. The carbon dioxide produced by this biomass burning is highly significant in the annual atmospheric carbon dioxide budget of the world. Fire has certainly been important in African vegetation for a very long time. Natural fires, at long intervals, can be started by lightning, as well as by volcanic activity. Once humans began to use fire – and some have suggested that this may have been as much as a million years ago – fire frequency would have increased. The acquisition of the ability to make (not just use) fire would have increased this frequency further, and another increase surely came when safety matches became widely available.

The question of the influence of fire on the distribution of vegetation types in Africa has been extensively discussed (Lock, 1998). On the one hand there are the views of workers such as Aubréville, who considered fire to be a major destructive force responsible for the destruction of vast areas of the drier forests of Africa. The converse view, that fire has been a feature of African vegetation for so long that it is now a fundamental part of many ecosystems, is probably more widely held. Though fire may sharpen boundaries between vegetation types, it does not greatly affect their distribution. Assessment of the importance of fire also depends on viewpoint. Foresters generally disapprove of fire and expend much effort in its prevention, whereas pastoralists use fire extensively to remove old grass and to stimulate the development of young grass growth, as well as to check development of woody plants. Foresters, who want trees, generally recommend that fires should be started early in the dry season, when the fuel is still not completely dry ("early burning"). Pastoralists, who want grass, prefer a fire late in the dry season when the fuel is

completely dry and damage to woody growth is maximal ("late burning").

The presence of specific adaptations to fire in a number of plants points to fire having been an important factor for a long time. Several trees, such as the shea-butter tree (*Vitellaria paradoxa*) and species of *Combretum* and *Pterocarpus*, have seeds that, when they germinate, produce what appears to be a radicle that pushes into the ground. However, close examination shows that at least part of this is hollow and is actually formed by fusion of the stalks of the cotyledons (seed-leaves), which remain within the seed. The terminal bud, which lies between the bases of these stalks, is thus carried down below ground level so that it is at least partially protected from fire. Once well buried, the bud starts growth and the shoot breaks out through the side of the apparent root. It elongates and eventually emerges from the ground, but not before producing a number of reduced scale-like leaves, each with a bud in its axil. If the terminal bud is burned away, the lateral buds in the axils of the scale leaves start to develop, giving the seedling another chance of establishment.

Many of the grasses of the fire-swept regions have a long awn on their grains. This awn is hygroscopic, twisting and untwisting with daily changes in humidity. The tip of the grain is sharply pointed and bears stiff hairs. If the point enters a crack, the hairs prevent it from coming out again, and the regular movements of the awn will tend to drive it deeper into the soil, until it is protected from the main heat of the fire.

Past Climatic and Environmental Fluctuations

Tropical Africa has not escaped the climatic fluctuations of the last 2 million years. However, interpretation of changes in Africa is complicated by the effects of rifting and vulcanism, particularly in the east, as well as by the presence of many isolated mountains and mountain ranges. Evidence is also relatively scarce because the peat and other organic deposits that are widespread in cooler regions do not accumulate nearly so readily in tropical lowlands because high temperatures speed decomposition processes. The peat deposits that do exist are mainly on mountains, and so may not present a representative picture of the much more extensive lowlands.

However, there are other kinds of evidence available, such as the extent of active and inactive (fossil) sand dunes. In West Africa the fossil dunes extend 400–600 km south of the present limit of active dunes; the arid period during which these fossil dunes formed appears to have been 12,000 to 20,000 years ago. South of the equator, the Kalahari Sands extend far to the north and east of any presently active dunes, reaching the Zaïre River in the north. It appears that the most recent period of active dune growth coincided with that north of the equator. Lake levels, as marked by raised beaches, can also provide evidence, at least of times when lake levels stood higher than they do now. The study of plant distributions can also sometimes suggest former connections between similar but now isolated vegetation types. For instance, several tree genera with most of their species in West Africa have one or more species in the Eastern Arc montane forests of Tanzania. This suggests a connection between these forests areas in the distant past.

In general, periods marked by cold in high latitudes are correlated with periods of drier and cooler climate in tropical Africa, and warm periods in high latitudes with warmer, wetter, climate. During drier, cooler periods, forested regions contracted and grasslands spread. In the driest times, forests seem to have contracted into a number of refugia, in Sierra Leone and Liberia in West Africa, as well as Cameroon and Gabon, eastern Zaïre, and eastern Tanzania. Most of the evidence shows that there was an arid period between 12,000 and 20,000 years ago (Lind and Morrison, 1974). Prior to this there were certainly similar fluctuations, perhaps of greater magnitude, but their dating and extent are still in dispute. Since 12,000 years ago there have also been changes; the wettest period seems to have been between 8000 and 4000 years ago. At this time, Neolithic pastoralists inhabited large areas of the Sahara and left rock paintings showing elephants, giraffes, and antelopes in areas now far too dry to support them.

Major Phytogeographic and Ecoclimatic Zones of Continental Africa

Mediterranean North Africa

This region, with winter rainfall and hot dry summers, borders the Mediterranean Sea. It includes Morocco, Algeria, Tunisia, Libya, and Egypt. The wettest part is the west, and here also there are the high mountains of the Atlas range. Thousands of years of human settlement, agriculture, and grazing of domestic animals have greatly altered the ecosystems. In Roman times there were certainly lions (*Panthera leo*) (and therefore a substantial prey population) and probably elephants (*Loxodonta africana*) in this region, but all are now gone.

The wetter parts were probably originally covered with forest, but this is now represented only by tiny fragments; *Celtis australis* and *Pistacia atlantica* may have been important trees in the original forests. The drier forests were (and in places still are) dominated by evergreen oak (*Quercus ilex*), which casts a dense shade in which few other species can grow, or by cork oak (*Quercus suber*). There are also coniferous forests of species such as Aleppo pine (*Pinus halepensis*) and North African cedar (*Cedrus atlantica*). Other parts are covered by scrub, similar in physiognomy to the chaparral of California and the fynbos of the Cape Region of South Africa, made up of shrubs with small hard (sclerophyllous) leaves, such as the kermes oak (*Quercus coccifera*), wild olive (*Olea europaea*), and, in very degraded sites, the dwarf palm *Chamaerops humilis*. The shrubs are fire-resistant, sprouting from the base after fires. The gaps between the cushions and, lower down, between the sclerophyllous shrubs support a rich herb flora including many annuals (Fabaceae are abundant and diverse) and many plants springing from underground bulbs or corms (geophytes). These grow during the late winter and spring, flower, and then dry up in the baking heat and drought of the summer. At high altitudes, between 2800 and 3800 m, there is often a low scrub made up of spiny cushions, sometimes graphically referred to as hedgehog heath.

Cape Region of South Africa

The Cape Region also has a Mediterranean climate, although, being in the Southern Hemisphere, it enjoys a hot dry summer when North Africa is having a cool wet winter. Like North Africa, thicket and low scrub (known locally as fynbos) are the main physiognomic vegetation types, and fire is a regular influence on the vegetation. Many species appear to be fire-adapted. Some, such as the red-flowered, lily-like species of *Cyrtanthus*, flower only after fires, stimulated either by chemicals in the smoke or by the greater daily fluctuations in soil temperature that follow removal of the vegetation cover. Others, such as species of *Leucadendron* (a *Protea* relative), retain their seeds on the parent plant and only release them after fire. The range of fire-adapted species in fynbos, and the many ways in which they respond to fires, suggests that fire has been a feature of this vegetation type for a very long time.

The region is extremely diverse geologically and has numerous isolated mountain ranges. These factors, combined with long isolation, have given rise to an extraordinary diversity of plant species – estimates vary from about 7000 species in the 71,000 km^2 of the region (White, 1983) to 8600 species in an area of 91,000 km^2 (Cowling and Richardson, 1995, Cowling *et al.*, 1997). The genus *Erica* (heaths) has over 500 species in the Cape Region; other extremely diverse genera include *Aspalathus* (Fabaceae, 250 species) and *Muraltia* (Polygalaceae, 100 species). The family Proteaceae, including 85 species of *Protea*, is prominent among the larger woody plants, and is associated with an endemic pollinator, the Cape Sugar-bird (*Promerops cafer*). The lower-lying and more level areas have largely been converted to agricultural land, but the mountain ranges continue to provide refuges for the endemic flora, although invasion by woody species introduced from other regions with a similar climate (such as south and west Australia) is a major problem.

Originally there was a rich fauna of large mammals but these were heavily hunted by European settlers. Some still survive in reserves, but the quagga (a form of zebra, *Equus quagga*) is extinct, and the bontebok (*Damaliscus dorcas*) and white-tailed gnu (*Connochaetes gnou*) survive only on enclosed farms.

Tropical Africa

The region known as tropical Africa takes in most of the continent. The vegetation is determined by the climate, which is highly seasonal over much of the region. There are virtually no parts of Africa without some kind of a dry period; truly ever-wet climates like that of Singapore (where two weeks without rain is a drought) are virtually absent. The major exception lies along the equator in West-Central Africa; here the dry season lasts a month or less. The equatorial regions of eastern Africa, however, lie within the rain-shadow of the Arabian landmass; here, even on the equator, rainfall is low, and there tend to be two rainy seasons rather than one. Both of these vary in intensity, and the gaps between them also vary in length, so that this region tends to suffer more than most from periodic droughts.

Tropical Africa can be divided into two parts. South and east of a line from Ethiopia to the mouth of the Zaïre River, most of the land forms a dissected plateau lying at about 1000 m above the sea. This is split from north to south by the Great Rift Valley, which extends from Israel through the Red Sea, then across Ethiopia, Kenya, and into Tanzania. A western branch runs south through Uganda, along the western side of Tanzania, and ends in Malawi. Between the branches of the Rift Valley, the land surface has tilted in places, disrupting river flows and producing the huge but shallow Lake Victoria as well as the extensive swamps of Uganda. To the north and west of the high plateau, the general land surface is much lower. Here again rocks of the Basement Complex underlie most of the region, but younger rocks are found here and there. Most of tropical Africa is covered by woodland and various forms of wooded grassland or grassland, with forest occupying the basin of the Zaire River, and drier bushlands, thickets, and grasslands in the equatorial regions of eastern Africa.

Major Ecosystems of Tropical Africa

Forest

True tropical forest is confined in Africa to two main blocks: the basin of the Zaïre (Congo) River, extending north and west through Cameroon and Gabon into southern Nigeria and eastwards into western Uganda, and a region farther to the west stretching from western Ghana through Ivory Coast and Liberia into Sierra Leone. Elsewhere, isolated forest patches are found inland around the Ethiopian Highlands, along the eastern African coast, and in parts of upland Tanzania (the "Eastern Arc" forests). Recent work has shown that these isolated eastern African forests are rich in endemic plants and smaller animals (Burgess *et al.*, 2004a). Most forests are found where annual rainfall exceeds 120 cm and where the dry season is no more than four months long. Close to the climatic forest boundary, small differences in water availability can greatly affect the vegetation; forest may extend far into grassland along the banks of rivers as "gallery forest," and nonforest plants can be found well within the forest zone on rocky outcrops.

African forests are poor in species compared to those of Southeast Asia and the Amazon Basin, but they are still species-rich places. Surveys in Ghana found forests in which a 25-by-25-m patch contained more than 200 higher plant species. These surveys also showed that the wettest forests are the most species-rich. They do not, however, contain the tallest trees, which are found in drier, partially deciduous forest, perhaps because the soils there are not so leached of nutrients as those in the wetter areas. These drier forests are, at least in Ghana, richest in the number of timber trees (Hall and Swaine, 1981). The most prominent plant families in the forest canopy are the mahoganies (Meliaceae) and legumes (Fabaceae, particularly the subfamily Caesalpinioideae).

Following disturbance, the pioneers such as umbrella trees (*Musanga*, in Moraceae) and *Trema* (Ulmaceae) grow at 3–4 m per year and rapidly form a tree cover. These species shade the soil and improve conditions for the establishment of trees whose seeds can germinate in shade, and whose seedlings can tolerate low light intensities. These early-successional species include some of the mahoganies, such as *Khaya* and

Entandrophragma (sapele), whose seeds are wind dispersed. Studies in Uganda have suggested that eventually this "mixed forest" may give way to a more species-poor forest in which leguminous trees are common. Some of these leguminous trees form extensive areas of forest in which a single tree species is dominant – an unusual state in tropical forest, for which there is currently no satisfactory explanation. *Cynometra alexandri* in eastern Zaire and Uganda and *Gilbertiodendron dewevrei* in the Zaire River Basin form such monospecific forest stands.

Beneath the canopy trees grow smaller trees and shrubs, some of them young plants of canopy species, but others naturally small at maturity. It is usually quite easy to walk through undisturbed forest, as there are few herbs on the forest floor. Only where a tree has fallen, allowing more light to reach the ground, is there a dense mass of quick-growing herbs in the ginger (Zingiberaceae) and arrowroot (Marantaceae) families. Woody climbers (lianas) are also most frequent in disturbed forest, and can reach the tops of the tallest trees.

African forests are rich in animal species. The canopy is occupied by many species of monkey. Some, like the species of colobus (*Procolobus*, *Piliocolobus*, *Colobus*), are leaf-eaters. Others, such as the various small guenons (*Cercopithecus*), feed mainly on fruit (Kingdon, 1997). Different species of guenon specialize on different dietary mixes and several species can thus coexist in the same area of forest. Other tree dwellers include species of flying squirrel, and also true squirrels, most of them seed-eaters.

The great apes – gorilla (*Gorilla gorilla*), chimpanzee (*Pan troglodytes*), and bonobo (*Pan paniscus*) – travel mainly on the ground and feed both there and in the trees. All live in groups and exploit extensive home ranges that they know intimately, moving from one seasonal food resource to another. A high proportion of their food is fruit, and they are important dispersers of seeds. In the wettest and least seasonal forests there are also large ground-dwelling monkeys – the drill (*Mandrillus leucophaeus*) and the mandrill (*Mandrillus sphinx*). Like the apes, they are highly social, living in troops of a hundred or more individuals and traveling over a huge home range.

There are also ungulates in the rain forests of Africa, ranging in size from the tiny royal and dwarf antelopes (*Neotragus pygmaeus* and *N. batesii*, 25–30 cm at the shoulder) and the taxonomically very distinct water chevrotain (*Hyemoschus aquaticus*, only 30–40 cm), through several species of duiker (*Cephalophus*, small antelopes), to the large bongo (*Tragelaphus euryceros*). The Ituri forests of eastern Zaïre harbor the okapi (*Okapia johnstoni*), whose closest relative is the giraffe. All of these are browsers on the leaves of forest shrubs and herbs.

The largest forest animal is the elephant (*Loxodonta africana*). Forest elephants are often recognized as belonging to a separate subspecies from those of open country; they are smaller and tend to have smaller and straighter tusks. Elephants are voracious feeders; their inefficient digestive systems mean that they must feed for a large proportion of each day to obtain enough nourishment. They are particularly fond of clearings in forest, because here more of the foliage is accessible, and by concentrating their feeding in such areas they may prevent tree development and perpetuate the clearings. Elephants also eat large quantities of fruit. Trees such as *Panda oleosa*, *Desplatsia* spp., and *Balanites wilsoniana* have such large seeds that it is hard to know what animal other than an elephant could possibly disperse them. Attempts to germinate seeds of *Panda* that have not passed through an elephant have been unsuccessful (Hall and Swaine, 1981). The effects on such plants of the widespread decline in elephants may be severe.

Birds are abundant in African forests although the diversity is lower than in, for instance, South American forests. On the ground there are forest species of guineafowl and francolin, as well as the Congo peafowl (*Afropavo congensis*) – the latter confined to the Zaïre Basin. Sunbirds (Nectariniidae) occupy the flower-feeding niche, but also eat small insects. Canopy feeders include numerous species of bulbul and greenbul (Pycnonotidae) as well as small warblers, many in the genus *Apalis*. Some weaver birds (Ploceidae) and waxbills (Estrildidae) are confined to forests, although the majority of both are grassland birds. Parrots, of which the African gray parrot (*Psittacus erithacus*) is the best-known, fly screeching over the forest canopy, seeking fruiting trees. Parrots are, however, very much less diverse than in either South America or Australasia.

Seasonal Tropical Vegetation

Much of tropical Africa is occupied by vegetation that develops under a climate in which the year is divided into dry and wet seasons. The vegetation of these areas is of many kinds, ranging from woodland with an almost closed canopy at one extreme to dry open grasslands at the other. All of these vegetation types have been referred to as savanna, but recently there has been a tendency to attempt to discard this term. It has been argued that a word that can mean almost anything from closed forest to open grassland is too vague to have any utility in ecological discussions. White's (1983) classification avoids the term, and instead uses "woodland," "wooded grassland," "grassland," "bushland and thicket," and "shrubland" to define the various physiognomic vegetation types more precisely. These terms will be used here, but where the term "savanna" has wide currency, as in West Africa, it will be used.

In tropical seasonal vegetation, water is abundant and growth can be rapid during the wet season, but in the dry season the grass rapidly dries out and becomes flammable; fires are frequent and often annual. It is somewhat paradoxical that the fiercest fires occur where the rainfall is highest – and where, therefore, biomass production during the growing season is highest. Over much of the savanna regions of Africa there is one wet season and one dry season each year, but near the equator, particularly in East Africa, there can be two wet and two dry seasons each year (*see* Major Environmental Factors of Continental Africa: Climate), so that the seasonal cycle is six months rather than twelve.

The herds of wild ungulates and their associated predators for which Africa is so famous are all found in areas of seasonal tropical vegetation. Densities vary enormously, now largely because of human pressures, but there can be no doubt that densities varied greatly before human impacts became significant.

Woodland

In West Africa, where the isohyets (lines of equal rainfall) are more or less parallel to the coast, there are belts of vegetation that follow the isohyets. These woodlands and wooded grasslands were classified many years ago by the great French botanist Auguste Chevalier into three savanna types: Guinea, Sudan, and Sahel. In West Africa, use of the term "savanna" is long established, and Chevalier's zones are still recognized and used today. The southernmost zone, with the highest rainfall and the shortest dry season, is the Guinea Zone (Lawson, 1986).

Woodland has been called by many different names in the African ecological literature. Tree savanna is perhaps the most widely used, as well as the more descriptive "tall grass–low tree savanna." In West Africa, woodland is referred to as "Guinea Savanna," and in francophone countries it is often called "forêt claire." This is a confusing term, because it implies that woodland is essentially the same as true forest ("forêt dense"). This is not the case; surveys in Nigeria of adjacent forest and woodland have shown that they have very few species in common. Physiognomically, woodland consists of a single tree layer, sometimes with an almost closed canopy. Lianas are absent, as are epiphytes. Grass covers the ground beneath the trees.

The trees are often misshapen because the frequent fires kill the growing points and cause branching. Because grass productivity is high, fires are more intense than in other zones. The bark of many of the trees is thick and furrowed, which insulates the delicate growing tissue (cambium) from the heat of fires. Regeneration is difficult as young seedlings are in the hottest part of the fire. Some species, such as the shea-butter tree (*Vitellaria paradoxa*), have specialized germination mechanisms that may help to overcome this. Some grass seeds, mainly those from the subfamily Andropogoneae, also have a mechanism for evading fires (*see* Major Environmental Factors of Continental Africa: Fire).

The greatest obstacle to the development of the woodlands of Africa (many would say their most valuable conservator) is the tsetse fly (*Glossina morsitans, G. pallidipes*, and other species). These biting flies are the vectors of trypanosomiasis, a lethal protozoan disease of cattle, and of sleeping sickness in humans. Wild game animals are immune to trypanosomiasis, although they can carry the protozoan in their blood. Control of the fly has been attempted in many areas, using methods of varying destructiveness to the ecosystem. The fly requires shaded places in which to rest, so tree clearance has been used. Aerial spraying with DDT and other nonspecific insecticides has now been replaced by highly specific treatment with persistent insecticides of the undersides of large tree branches, where the flies rest. Other attempts to break the cycle of transmission have included the removal, by shooting, of wild ungulates. Researchers continue to look for effective and environmentally sound ways of tsetse control. An alternative that has been tried in some areas is to leave the flies alone and harvest the wild game instead of the cattle. This has an advantage in that each member of the diverse ungulate community occupies a different niche, so that the carrying capacity and yield are, at least in theory, higher than those from a monoculture of cattle.

The Guinea Savanna woodlands of West Africa have among their characteristic tree species *Lophira alata* (Ochnaceae) and *Daniellia oliveri* (Fabaceae). Slightly drier woodlands have species of *Isoberlinia* (Fabaceae). The main grasses are species of *Andropogon* and *Hyparrhenia*, both with awns on their grains. These large grasses stand 2–3 m high when mature. Many agronomists have been misled into thinking that this high standing crop also implies a high animal production potential, but this is not so. When the grasses are at their tallest, they are made up largely of stem, which is hard, woody, and of very low nutritional value. After a fire, and at the beginning of the rains, nutritious foliage is produced, but this does not last long.

This extreme seasonality in the quality and availability of grass goes a long way towards explaining the relative scarcity of large wild animals in this zone. Various antelopes, such as roan (*Hippotragus equinus*) and eland (*Taurotragus oryx*), occur here, as do smaller species such as bushbuck (*Tragelaphus scriptus*). Elephants occur at least seasonally. Predators are correspondingly also scarce, with leopard (*Panthera pardus*) and lion both occurring at very low densities.

The corresponding zone to the Guinea Savanna south of the equator is the miombo woodland (Frost *et al.*, 2002). "Miombo" is the local name for one of the commonest trees, *Isoberlinia tomentosa* (Fabaceae), but by far the most widespread and abundant trees belong to the genus *Brachystegia* (Fabaceae). There are about 25 species of this genus in the woodland regions of south-central Africa, all fairly similar to one another, mostly variable, and difficult to identify. Miombo covers vast areas of southern Tanzania, Zambia, northern Zimbabwe, southern Zaïre, and south-eastern Angola. Miombo is physiognomically rather different from Guinea Savanna; the trees are usually taller, often with an almost closed canopy, and the grass is somewhat shorter. However, fires are virtually an annual event. Although at first sight miombo woodland can appear rather monotonous, it shows considerable local variations related to topography. The ancient land surface on which it grows is weathered into a pattern of gentle hills and valleys.

The tops of the hills and ridges are dry and well drained, but there are often rock outcrops (inselbergs) around which water accumulates so that the vegetation around them is tall and dense. The rocks themselves, if not sloping too steeply, often bear a shallow turf that dries out completely during the dry season but, during the wet season, becomes saturated and supports many small ephemeral plants such as species of *Utricularia* (bladderworts), *Xyris* (yellow-eyed grasses), and various sedges (Cyperaceae). In West Africa there may be mats of the sedge *Afrotrilepis*, which can dry out completely and then rehydrate and resume growth when water is once again available. In southern Africa there can be clumps of the shrub *Myrothamnus flabellifolius*, the resurrection plant, which behaves in the same way (i.e., it is poikilohydric). These rocky hills are often home to a small and specialized antelope, the klipspringer (*Oreotragus oreotragus*), and to rock hyraxes (*Procavia johnstoni* and *Heterohyrax brucei*).

The typical miombo woodland occurs on the hill slopes, but even this is not uniform. Termite mounds, built by species of *Macrotermes*, can be up to 4 m high and 10 m in diameter. These termites are fungus gardeners; they collect dead plant material and carry it into the center of the nest. Here it is formed into a honeycomb-like structure that is colonized by

fungi that break down the cellulose and lignin in the wood. The termites feed on the fungal hyphae and are thus indirectly nourished by the wood. A by-product of this activity is the concentration of mineral nutrients and fine soil fractions in the mound so that, somewhat surprisingly, it becomes a favorable habitat for plant growth. Termite mounds bear much denser woody vegetation, often casting enough shade to exclude grasses and therefore fires.

The valley bottoms in the miombo ecosystem are seasonally waterlogged, preventing the growth of most trees. Tall grassland occupies this habitat, the grasses often forming well-spaced tussocks between which small plants can grow early in the wet season. The soils of these valley grasslands are often black clays that are extremely sticky when wet and very hard when dry.

A feature of miombo woodland, relevant to concerns about climate change and carbon sequestration, is that the root systems are very extensive so that almost half of the vegetation biomass is below ground. While there is much emphasis on the role of forest destruction on atmospheric carbon dioxide levels, very much less of forest biomass is below ground. The importance of woodlands like miombo to the world's carbon balance should not be underestimated and their value in this respect should be better known.

As in the Guinea Savanna, large mammal biomass is low in miombo. There are, however, many more species than in Guinea Savanna. Roan antelope occur, as well as the closely related sable antelope (*Hippotragus niger*) with its spectacular scimitar-shaped horns. Hartebeest (*Alcelaphus buselaphus*) occur as do elephant where there is permanent water within reasonable range. The spectacular greater kudu (*Tragelaphus strepsiceros*), with long, spirally twisted horns, is very much an animal of the miombo. Herds of Cape buffalo (*Syncerus caffer*) can be found, again where permanent water lies within a reasonable distance.

The annual cycle of the miombo woodland follows the rains. The trees often produce their new leaves before the first rains, presumably drawing on reserves of water from deep in the soil. This flush of new foliage is often reddish, perhaps because of secondary compounds that make the leaves unpalatable to predators. Flowering takes places fairly early in the wet season, giving the often bulky pods and seeds time to develop before the dry season begins. The pods ripen and burst to scatter their seeds during the dry season. The extent to which the leaves are shed varies from place to place and year to year according to the degree of drought and the intensity of the fire.

Drier woodlands are found both to the north and south of the equator. To the north, they fall into the zone called "Sudan Savanna" by Chevalier. This zone has a longer dry season and a lower annual rainfall than the Guinea Savanna. Typical tree species include *Parkia filicoidea* and *Piliostigma thonningii*, but this zone has a much higher human population and is more cultivated than the Guinea Savanna. There are a number of possible explanations for this; maybe all play a part. Because it is farther from the coast, it was less heavily raided and depopulated during the time of the slave trade. The relatively smaller growth of grass makes land clearance and maintenance easier, and the lower rainfall means that the soils are less intensively leached and therefore rather more fertile. Whatever the reason, this zone now supports high human populations who cultivate groundnuts, sorghum, and various kinds of millet. Many of the remaining trees survive because they are of some use; *Parkia* seeds and the pulp that surrounds them are a useful food, as are the leaves and fruit-pulp of the baobab tree (*Adansonia digitata*).

In southern Africa the corresponding zone is probably mopane woodland, which is not very intensively settled or cultivated. It is dominated by a single tree species, *Colophospermum mopane* (Fabaceae: Caesalpinioideae). Mopane grows in areas that are hotter and drier than those occupied by miombo. It is unusual in its family in being wind pollinated. Mopane woodland can support a wide range of large mammals, including elephant, black rhino (*Diceros bicornis*), eland, and impala (*Aepyceros melampus*).

Bushland and Thicket

The drier lowlands of eastern Africa and the Horn of Africa are occupied by woody vegetation that is low-growing (less than 10 m tall), often spiny, and leafless for a substantial part of the year. The rainfall in this region is always somewhat unpredictable, tending to fall in two wet seasons in each year, either or both of which may fail. The most obvious components of the vegetation are species of *Acacia* (Fabaceae: Mimosoideae). Almost all of these bear paired spines, sometimes hooked and very sharp. Another genus that is abundant and diverse here is *Commiphora* (Burseraceae); all the trees in the genus have a resinous scent and one is the main source of myrrh. (Another genus of the region, *Boswellia* (Burseraceae), yields the resin known as frankincense.)

Many other tree and shrubs occur so that this vegetation type is very rich in species. Few of these are common to the wetter woodland regions, and the Somali–Maasai area is recognized by White as a separate phytochorion; that is, a region with a distinctive flora. The floristic relationships of this area lie with southern Arabia and the north-western part of the Indian peninsula rather than with the rest of Africa. Somalia and northern Kenya in particular are rich in endemic plant species, perhaps partly because of long isolation, but also because of the existence of specialized habitats such as regions where the underlying rock is either limestone or gypsum (calcium sulfate) and the coastal strip near Hobyo (Somalia), where low-growing plants cover fossil sand dunes.

The animals of this area are also distinctive. Grevy's zebra (*Equus grevyi*) and the wild ass (*Equus africanus*) are virtually confined to this part of Africa, and the latter is very scarce. The very handsome and distinct reticulated giraffe (*Giraffa camelopardalis reticulata*) is confined to Somalia, southern Ethiopia, and northern Kenya. The beira (*Dorcatragus megalotis*), a specialized long-legged duiker, is confined to Somalia. The four species of dikdik (*Madoqua*), tiny antelopes that live alone or in pairs in bushland, in territories marked by dung piles, are virtually confined to this vegetation type. Several species of gazelle are also confined, or almost so, to this region. Finally, two remarkable long-necked gazelles, the dibatag (*Ammodorcas clarkei*) and the gerenuk (*Lithocranus walleri*) occur only here; both are exclusively browsers, stretching their long necks up into the bushes and often standing on their hind legs to reach taller growth. Their narrow muzzles allow them to pick out small leaves from between dense twigs or thorns.

The islands of Suquṭrā and Abd al Kūrī to the east of the tip of the Horn of Africa. Both have undoubtedly been isolated for a very long time and have no indigenous large mammals (although introduced livestock now abound). This has allowed the development of a remarkable flora that, although clearly part of the Somalia–Maasai Region, includes numerous endemics such as *Dendrosicyos socotranus* (a tree-forming representative of the cucumber family), *Dorstenia gigas* (Moraceae), and *Adenium socotranum* (Apocynaceae).

Grasslands

Many of the grasslands of tropical Africa are the product of special soils that prevent the growth of trees; seasonal waterlogging, shallow soils subject to extreme seasonal droughts, and high concentrations of metals such as copper and cobalt all lead to local grasslands. However, near the equator in eastern Africa there are extensive areas of grassland, sometimes with scattered flat-topped acacia trees (*Acacia tortilis*) or thicket clumps. Most of these areas lie between 1000 and 2000 m above sea level. The commonest and most prominent grass is the red oat grass (*Themeda triandra*). This species thrives under a regime of annual burning and light grazing, but it is vulnerable to overgrazing. Dense tussock-forming grasses such as *Sporobolus pyramidalis* tend to replace it if the grazing pressure is excessive.

These grasslands have long been the territory of pastoral people such as the Maasai, and it is possible that these people, who are well aware of the effects of fire on vegetation, have used it over the millennia to alter the balance from woodland or thicket to grassland on which to pasture their cattle. These people have also long lived in close proximity to huge populations of wild ungulates and their predators. These ungulate populations often make seasonal migrations to make best use of their range, such as that which straddles the Tanzania–Kenya border in the Serengeti and Mara region. The main species in these migratory populations are wildebeest (*Connochaetes taurinus*) and the plains zebra (*Equus quagga boehmii*); other more sedentary species include Thomson's gazelle (*Gazella rufifrons thomsoni*) and Grant's gazelle (*Gazella granti*). In wetter areas, with access to permanent water, other sedentary species such as Cape buffalo, elephant, and waterbuck (*Kobus ellipsiprymnus*) are found. The Serengeti migration is the best-known, but in southern Sudan there is also a huge migratory system involving tiang (*Damaliscus lunatus*), Mongalla gazelle (*Gazella rufifrons albonotata*), and white-eared kob (*Kobus kob leucotis*) (Cobb and Lock, 1988). These species spend the wet season in the tall grasslands towards the Uganda border. When the dry season begins, they move northwards and feed on the swamp grasses, such as *Oryza* (wild rice) and *Echinochloa*, exposed by the retreat of the Nile River floods. This migration crosses the line of the partially dug Jongei Canal – now abandoned due to civil war – which forms a significant obstruction.

The biomass of ungulates supported by some of these grasslands is very high, particularly where there are two rainy seasons each year and therefore at least some forage of good quality at all times. The Western Rift Valley in Uganda and eastern Zaïre is typical; here large herds of elephant, hippopotamuses from the lakes, and Cape buffalo form the main part of the biomass, but other species such as Uganda kob (*Kobus kob thomasii*), waterbuck, and warthog (*Phacochoerus africanus*) are also common.

In recent years, following the widespread destruction of the larger animals, the smaller ones have increased in numbers. The killing of many of the elephants has also led to widespread vegetation changes. *Acacia* trees, whose seedlings were formerly so regularly browsed and burned that they remained in a suppressed state in the grass, are now regenerating in many places, and thicket clumps dominated by the thorny scrambler *Capparis tomentosa* (Capparidaceae) are now spreading and coalescing to produce larger areas of thicket that are becoming home to giant forest hogs (*Hylochoerus meinertzhageni*), a species more commonly found in montane forests.

In all of these regions with numerous herbivores, there are high populations of predators. Lions are the most conspicuous, living in prides made up of one or more females and their offspring with at least one mature male in attendance. The females collaborate in hunting, and any kill is shared by all the pride. When a new male displaces another, his first action is to kill any cubs in the pride still dependent on their mothers. This action rapidly brings the lionesses into estrus and allows the new male to start passing on his genes with the minimum of delay.

Spotted hyenas (*Crocuta crocuta*) can also be very common and have been shown to function not only as scavengers but also as highly efficient predators in their own right. They live and hunt in matriarchal groups that hold and defend communal territories. The cheetah (*Acinonyx jubatus*) is a highly specialized and solitary cat, entirely dependent on its speed for running down prey. Its jaws are relatively weak, and many of its kills are lost to lions or hyenas against which it has virtually no defense. Finally, there is the wild dog (*Lycaon pictus*). This is another species that lives in groups; hunting is by a prolonged chase in which the members of the pack take turns in the lead. The leaders snap and tear at the hind end of the prey, eventually bringing it down. Wild dogs are short-lived, however, and rely on frequent large litters to maintain their numbers. They usually have a very wide home range and in many areas they are now very scarce or absent, and may well, with the cheetah, be Africa's most endangered carnivores.

The African grasslands are also home to large ground-living birds; some, like the ostrich (*Struthio camelus*), are flightless and others, like the kori bustard (*Ardeotis kori*), almost so. The large mammal populations provide food for scavengers; for instance, a carcass on the Serengeti Plains will attract four or five species of vulture that spend much of their day soaring on thermals; they watch both the ground and each other so that if one spots prey and starts to descend, others quickly follow.

Shrublands

Shrubland occupies the driest areas; under conditions of even lower rainfall the individual shrubs grow farther and farther apart until the land is best referred to as desert. The shrubs are generally between 10 cm and 2 m in height, and are of many different species and families. Members of the Acanthaceae, particularly the genera *Barleria* and *Blepharis*, are often prominent. Many are spiny, and all are highly facultative and irregular in their production of leaves and in their flowering and fruiting, often reproducing only when rainfall is exceptionally high or prolonged. Under these conditions, many annual grasses also appear between the shrubs. Perennial

drought-resistant grasses often occur among the shrubs, and some authorities believe that the natural state in these ecosystems is a drought-resistant grassland, which has been converted to a shrubland by overgrazing. However, at least some shrublands grow on soils developed from limestone or gypsum and so may be partly edaphically controlled.

Deserts

There are two main areas of real desert in Africa: the Sahara to the north of the equator and the much smaller Namib of southwestern Africa. Parts of the Horn of Africa are also extremely dry, but only small areas such as the Danakil Depression can be considered as true deserts, if true desert is defined as an area in which plants grow only where there is extra water either from springs or from runoff.

At least some rain falls over much of the Sahara each year. The exception is the eastern end, in the Nile Valley, where no rain may fall for many years in succession. Over much of the rest, enough rain falls in most years to produce a thin ephemeral vegetation in more favorable areas. Grasses are the main component of this and dry out quickly to form a standing hay that is a valuable food resource for both wild animals and the domestic animals of nomadic herders.

Formerly several ungulate species were not uncommon in the desert: the addax antelope (*Addax nasomaculatus*) and the scimitar-horned oryx (*Oryx dammah*) being the largest, with several smaller species of gazelle. Most of these have extremely efficient water-conserving strategies and are capable of surviving without drinking. Human population increase, with concomitant competition for space, grazing, and water, and the increased availability of firearms, mean that all of these species are now rare and endangered, and the scimitar-horned oryx is now officially extinct in the wild. Scattered over the desert are oases where springs from underground aquifers provide year-round water. Most of these oases support permanent human settlements, often dependent on sparse annual crops as well as dates produced by the date palms (*Phoenix dactylifera*) that thrive in these environments.

The Namib Desert in southwestern Africa is also very dry, at least near the coast, but the dryness is mitigated by frequent fogs arising from the cold Benguela Current immediately offshore. Many of the plants and animals in this region survive by collecting water from mist. Furthermore, a few rivers cross the coastal strip, forming linear oases in which water is almost always available. Animals such as gemsbok (*Oryx gazella*), and even a few herds of elephants, manage to survive in this harsh environment.

One plant family, Mesembryanthaceae, has speciated enormously in this region, and there are hundreds of species. Some of these are very simple in structure, producing clumps of short shoots each bearing just one pair of large thick leaves at any one time. The lower parts of these leaves are buried in the soil, thus avoiding the extreme heat of the surface, and the exposed surface is often translucent, allowing light to reach the buried part of the leaves. The exposed parts of the leaves are also often colored and textured just like the surrounding stones, which camouflages the plants; the members of one genus, *Lithops*, are known as "living stones." One truly remarkable plant grows in this area and nowhere else in the World – *Welwitschia mirabilis*. This is not a flowering plant, and neither is it a true conifer. Each plant consists of a thick woody stem, mostly buried in the ground, bearing two huge strap-shaped leaves that grow continuously from the base, lie on the ground, and wear away at the tips.

Montane and Afroalpine Ecosystems

Mountains in Africa are of two kinds: relatively recent volcanics such as Mts. Cameroon, Kilimanjaro, Elgon, and Kenya and the Ethiopian Highlands, and old up-thrust portions of basement complex such as the Ruwenzori Mountains and numerous lower ranges in Tanzania such as the Usambara, Uluguru, and Udzungwa Mountains.

Mountain slopes were originally forested, but because of their fertile soils the forests have been cleared for cultivation in many places. However, the higher forests tend to be cool and misty and much less attractive for agriculture, and some of these survive. Trees such as podocarps (*Podocarpus*, *Afropodocarpus*) and the mountain bamboo *Synarundinaria alpina* are common. Epiphytic bryophytes, ferns, and orchids are abundant. Above the montane forest is a belt of "giant heath" formed from large species of *Erica* that attain 10 m in height. As in the montane forest, epiphytic mosses and liverworts are very abundant, often forming dense mats covering the trunks and branches of the giant heaths.

Above the heath zone lies the afroalpine zone proper. The climate here has been described as "summer every day and winter every night." Hot sunny days (this region is often above the cloud zone) are followed by nights during which the temperature plunges to well below freezing, and frost and ice forms on the ground. Here giant groundsels (*Dendrosenecio*) and giant *Lobelia* grow scattered in a low shrubland of everlastings (*Helichrysum*) and *Alchemilla*. The giant groundsels and giant lobelias consist of rosettes of huge leaves borne at the ends of sparsely branched stems. (A very similar growth form is found in the genus *Espeletia* of the Andean paramos.) During the night, the leaves fold upwards to form a dense mass around the delicate terminal buds, thus protecting them from the cold. Water and mucilage accumulate in the rosette and also help to prevent the buds from freezing. The old leaves of some species do not fall, but accumulate as an insulating blanket around the stems; other species have thick corky bark that functions in a similar way.

Animals are generally sparse in montane forests and in the afroalpine zone. Various antelopes move up here from the surrounding savannas, as do elephants and the occasional buffalo. With these come predators, particularly leopards. Specifically montane animals include the mountain nyala (*Tragelaphus buxtoni*) of the heath zone in Ethiopia, and the giant forest hog, a huge, hairy, black pig that occurs in the montane forests of Kenya and elsewhere, but also at lower altitudes in western Uganda. Montane forests are rich in bird species, some of them occurring nowhere else. The giant lobelias are visited and pollinated by sunbirds, which are specialists on them.

Wetlands

The wetlands of Africa may be seasonal or permanent. Warping and subsidence of the earth's crust has produced several

extensive wetland areas, such as the inland delta of the Niger River in Mali, the Sudd of the Nile River in the Sudan, the inland delta of the Okavango in Botswana, and the Bangweolo swamps of northern Zambia. Many permanent swamps are dominated by the giant sedge *Cyperus papyrus*. Papyrus requires at least some water movement and a reasonable supply of nutrients to thrive; when it does, it attains heights of 3 m or more. Papyrus swamps are very species poor. There are few other plants, most of them climbers like the purple morning-glory, *Ipomoea rubens*, and the yellow-flowered pea, *Vigna luteola*. A swamp antelope, the sitatunga (*Tragelaphus spekei*), shelters in papyrus but finds little food there. Some birds, such as the golden-crowned gonolek (*Laniarius mufumbiri*) (a shrike) and Carruthers' cisticola (*Cisticola carruthersi*) (a small brown warbler) spend their whole lives in papyrus swamps, but most of the herons and other waterbirds are found along channels and at the swamp margins. A most striking bird of African wetlands is the shoebill (*Balaeniceps rex*), a large-billed stork that lives in the depths of swamps and feeds on fish.

Temporary pools are a common feature of the seasonal region of Africa. A pool will fill early in the rainy season, sometimes from rain and sometimes from overspill from rivers. Fish usually reach all but the most isolated pools; the air-breathing African lungfish (*Protopterus aethiopicus*) can survive in a mucous cocoon in the mud for many months, emerging when water returns. Catfish (*Clarias*) can move over land through wet grass, and species of killifish survive as resistant eggs in the mud, hatching when wetted.

Floating plants like the Nile cabbage (*Pistia*) and water lilies (*Nymphaea*), as well as submerged plants such as bladderworts (*Utricularia*) and hornwort (*Ceratophyllum*), appear from seed in the mud. In game reserves and parks hippopotamuses (*Hippopotamus amphibius*) often move into the pools and churn them into mud, as well as adding nutrients in their droppings. At the end of the wet season the pools start to dry up. Then flocks of herons, egrets, and storks arrive to feed on the fish trapped in the shrinking patches of water, and the hippos move back to permanent water or die.

Seasonally flooded grasslands are found around the edges of many of the swamp regions of Africa. Some of them, like those around the Sudd of southern Sudan, and the Kafue River flats of Zambia, are very large. Among the grasses, species of *Echinochloa* are often common; at least some of these are very tolerant of flooding and their stems can elongate and form a floating mat when water levels are high. Perennial wild rice (*Oryza longistaminata*) is also common. Many of these areas are occupied by pastoral groups such as the Dinka of southern Sudan. They are cattle herders and move seasonally to follow the rise and fall of the flood. There are also wild ungulates; the Nile lechwe (*Kobus megaceros*) is confined to the grasslands at the edges of the Sudd swamps, and the red lechwe (*Kobus leche*) is found in various seasonally flooded grassland areas of Zambia. These animals also follow the rise and fall of the flood.

Lakes

Africa has numerous freshwater lakes, some of them very large. They are either shallow or deep lakes. First, there are the deep lakes of the Rift Valleys: Lakes Tanganyika (33,000 km^2, maximum depth of 1460 m), Malawi (704 m), Edward, and Albert. These lakes are permanently stratified. In the temperate zones, low winter temperatures cool the surface waters to 4° C, at which temperature water is densest and sinks, carrying dissolved oxygen to the bottom of the lake. In the tropics there are no significant seasonal variations in temperature, the surface water is always warmer than the depths, and there is little circulation, so that the deeper layers are anaerobic and lifeless. In spite of this, the oxygenated upper layers have a rich fauna and are highly productive. These lakes are ancient, and speciation within them has been rapid. The diversity of molluscs in Lake Tanganyika, for instance, parallels that of the sea, with thick-shelled species occurring on exposed coastal rocks and species with long spines on the shells living on soft mud bottoms. The fish faunas also parallel those of the sea, with some species living among rocks near the shore, while others exploit the zooplankton and other food in the open water.

Then there are the shallower lakes, such as Lakes Victoria (69,000 km^2, maximum depth of 90 m), Kyoga, Chad, and Bangweolu. In these lakes the stirring action of the wind is enough to circulate the water and carry oxygen to the lake bottom. Most of the shallower lakes are relatively young in geological terms, but some of the groups of fishes in them have speciated explosively to produce "species flocks," each of whose members exploits a different, often extremely specialized, niche.

In Lake Victoria, which is less than a million years old, the cichlid mouth-brooding genus *Haplochromis* is represented by more than 150 species, including plant-eaters, snail-eaters, and fish predators. More bizarrely, there are also species that live by biting off scales and portions of the fins of others, by pulling mayfly larvae from their burrows in dead wood, and by (probably) sucking the eggs and young from the mouths of brooding females of other species. Sadly this astounding diversity is under extreme threat following the introduction of the Nile perch (*Lates niloticus*) to the lake. This is a voracious predator, formerly found only in the Nile system below the barrier of the Murchison Falls, but now introduced to Lake Victoria with the intention of increasing fishing yields. Water hyacinth (*Eichhornia crassipes*) has also been introduced to Lake Victoria and is altering the ecosystem enormously by forming extensive mats at the edge of the lake, which make access for fishing difficult. (These mats may also, however, be providing nursery areas for young fish.) Both Lake Malawi and Lake Tanganyika (deep lakes) have a similarly remarkable diversity of cichlid fishes and, so far, neither Nile perch nor water hyacinth has been introduced to them.

Finally, within the Rift Valley there are a number of lakes that are highly alkaline or saline. Soda (sodium carbonate) is obtained commercially from Lake Magadi in Kenya, and salt (sodium chloride) from Lake Katwe in western Uganda. The soda lakes of the western Rift Valley in Ethiopia, Kenya, and Tanzania, as is often the case with extreme environments, support relatively simple ecosystems in which a few species are present in great abundance. A planktonic blue-green alga, *Spirulina*, often forms almost a soup in the alkaline water. Planktonic copepods (small crustaceans) and chironomid (midge) larvae feed on the algae. Lesser flamingos (*Phoeniconaias minor*) also feed on the algae by filtering water through their beaks. Greater flamingos (*Phoenicopterus ruber*) have more widely spaced filters in their beaks, and feed mainly on

the larger copepods and midge larvae, often stirring them out of the mud with their feet. They are not as numerous as the lesser flamingos. Both species migrate up and down the Rift Valley in response to changes in water levels and algal concentration. They breed in huge colonies when water levels are right and food availability is maximal.

Coastal Ecosystems

Most of Africa's seashores are sandy or muddy; rock is rare and generally confined to isolated headlands. Coral reefs occur along the east coast from the Red Sea to Mozambique, but not in the west, where cold upwelling water and turbidity virtually exclude them.

The strand lines of sandy shores support open communities of creeping plants such as *Ipomoea pes-caprae* (Convolvulaceae), *Canavalia rosea* (Fabaceae) and *Remirea maritima* (Cyperaceae). The first two have large seeds that float in seawater and are probably thus dispersed; both species occur widely in the Old World tropics. On the landward side of this grow salt-resistant bushes of species such as *Sophora inhambanensis* and *Pemphis maritima*, often forming dense thickets. Such thickets may grade into forest, or, in parts of West Africa, into a wind-swept grassland with bush clumps sculpted by the wind. Some of the grassland plants exist here as prostrate ecotypes that maintain their prostrate growth form in cultivation.

More sheltered shores often support mangrove swamps, particularly in the neighborhood of the mouths of large rivers such as the Niger and the Rovuma (Tanzania). West African mangrove swamps are relatively poor in species; *Avicennia* and species of *Rhizophora* are frequent here. The east coast is richer in species. The mangrove ecosystem is extremely productive; the breathing roots of the mangroves support communities of seaweeds and oysters, while the mud surface teems with crabs and small fishes. The whole zone is a rich nursery for fish.

Some of the more sheltered sandy flats behind reefs, particularly in northern Kenya and southern Somalia, have extensive beds of seagrasses (*Cymodocea*, *Halodule*, and *Thalassodendron*) (Richmond, 1997). These are not true grasses, although their flat, parallel-sided leaves are similar, but are closer to the pondweeds (Potamogetonaceae). Some grow intertidally, but many extend well below low-tide level and may form extensive lawns on sandy substrates. Here they are grazed by dugongs (*Dugong dugon*). These marine vegetarian mammals are regularly hunted and are becoming scarce and endangered almost throughout their range.

Madagascar

Madagascar has been isolated from Africa for at least 140 million years, and from India for around 88 million years. It has a rich flora in which about 84% of the species are endemic, and in at least some cases their closest relatives occur in southern Asia rather than Africa. There are five endemic plant familes and 321 endemic genera. The isolation of Madagascar predated the major adaptive radiation of mammals that has occurred in Africa, and it lacks large grazing mammals – although a dwarf hippopotamus appears to have become extinct less than 1000 years ago – and large carnivores.

At present the largest wild mammal is the bush pig (*Potamochoerus larvatus*), which is believed to be a recent arrival. A group of primitive primates, the lemurs, has radiated into all the habitats on the island. Fossils show that they were formerly even more diverse than they are now; some extinct forms were much larger than any modern species. The largest carnivore is the fossa (*Cryptoprocta ferox*), a large mongoose-like animal that climbs trees well and is a specialist predator on lemurs. There are also fossil and subfossil remains of giant flightless birds (*Aepyornis*), the last of which seem to have become extinct only a few hundred years ago. Humans reached the island from the east perhaps 1500 years ago. At first their settlements were confined to the coast, but later they spread inland and colonized the central plateau. Humans have had a dramatic effect on the island's natural vegetation and habitats.

Flowering plants were in their earliest stages of evolution when the island became isolated, and a high percentage of species (80% in the legumes, Fabaceae), many genera, and some families are endemic. The vegetation of the island is very varied. On the eastern side, rainfall is high and there is little or no dry season. Here there are tropical rain forests, very diverse in composition, with no single species being dominant. They differ from the forests of mainland Africa in their lower stature (25–30 m), lack of large emergent trees, the abundance of small palms in the understory, and the frequent occurrence of climbing bamboos. These forests have been considerably reduced by clearance for agriculture, and only scattered fragments remain. Secondary forests are widespread, often characterized by the distinctive traveler's tree (*Ravenala madagascariensis*), with a single stem crowned by huge leaves arranged like a fan in two opposite rows. Higher up the forest takes on a more montane aspect; the trees are shorter and more branched, and epiphytic ferns and mosses are abundant. The highest mountains support a montane thicket of small-leaved ericoid shrubs such as *Erica* (Ericaceae), *Stoebe* (Asteraceae), and everlastings (*Helichrysum*, Asteraceae) (Koechlin *et al.*, 1974).

A drier form of forest or woodland seems also to have occupied much of the central plateau, but only tiny fragments remain, and these are under intense pressure from fire, agriculture, and wood cutting for charcoal. The commonest tree is tapia (*Uapaca bojeri*), which may owe its survival to its fire resistance. One legume genus with two species, *Peltiera*, has recently been described from forest fragments in this zone; only three specimens and no living plants are known and it seems likely that the genus was extinct before it was described. The forests of the central plateau have largely been replaced by a species-poor grassland that provides little protection to the soil from erosion so that gullies are widespread and deep.

In the western half of the island, dry deciduous or semi-deciduous forest survives here and there, particularly in limestone areas, which have often weathered to produce an inhospitable landscape of sharp ridges and pinnacles ("tsingy") that is very difficult to access and unsuited to any kind of agriculture.

The southern end of the island, particularly in the west, is very dry, and here a peculiar thorn forest is found in which the endemic cactus-like family Didieriaceae is common. Specialized lemurs (sifakas - *Propithecus verreauxi*) live in this thorn forest. This remarkable vegetation type is threatened by agriculture, particularly sisal cultivation, by grazing, and by cutting

for charcoal production. Perhaps because of the absence of large grazing animals, members of several plant families have developed a growth form in which leaves are absent and photosynthesis is carried out in the flattened stems. Several members of the family Fabaceae show this feature.

Another plant growth form perhaps more widely developed in the dry parts of Madagascar than in any other region is the "bottle-tree," in which a thick and swollen trunk supports a rather small crown. The genus *Adansonia*, with one species in Africa (the baobab) and one or two in Australia, has seven species in Madagascar. The flame-tree (*Delonix regia*), now an extremely widespread ornamental tree in the tropics, is one of ten *Delonix* species in Madagascar, with just one other in tropical Africa. Several of the Madagascar species are bottle-trees.

Some of the richest habitats in Madagascar are the rocky outcrops, perhaps because they are sheltered from fires and grazing animals. Numerous endemic species of *Aloe* and succulent spurges (*Euphorbia*), as well as strange single-stemmed spiny succulents (*Pachypodium*), are common on these rocky outcrops and make them striking refuges for the remarkable flora of this isolated island.

The Future for African Ecosystems

What of the future? The human species has been present in Africa for longer than in any other continent. For much of the time, however, humans would have lived as part of the normal ecosystem – essentially part of the wild fauna. Exactly when they began to rise to dominance and to influence the composition and distribution of ecosystems is hard to say. There is some evidence that humans have been using fire (although probably not making it) in Africa for as long as a million years. Fire is one of the most potent forces for change in tropical vegetation and an increase in fire frequency caused by humans may well have shifted the balance between woody and herbaceous vegetation towards the latter. The southward spread of pastoral peoples, with their knowledge of the power of fire to produce new grass and control trees, seems to have begun by at least 2000 years BC, perhaps in response to increasing Saharan aridity, and may have been an important force for vegetation change.

Domestication of crops may well have begun at the same time. African plants that have been domesticated include yams (*Dioscorea*), sorghum, upland rice (*Oryza*), and cowpea (*Vigna*), with oil palm (*Elaeis*) in the forest zone. Bananas (*Musa*) probably arrived from Asia in the first millennium AD, and New World crops such as maize (*Zea*), cassava (*Manihot*), tomato (*Lycopersicum*), and peppers (*Capsicum*) did not arrive until the fifteenth or sixteenth centuries AD.

Most of the people of Africa survive by subsistence agriculture, or by growing crops that are sold to the rapidly increasing town populations. Shifting cultivation, in which a piece of ground is cleared, cultivated for a few years, and then abandoned for a fallow period of varying length, is the traditional way of exploiting the nutrient-poor soils characteristic of much of Africa. Although this practice is often attacked as wasteful and destructive, it is a very satisfactory mode of land use, so long as population densities remain low and a long fallow period is possible. However, once populations increase, the length of the fallow period falls, as do yields. This increases pressure to find new agricultural land, such as may be opened up in forest by logging activities. The extraction roads that allow people into the forest and the clearings made during timber cutting provide sites for settlement. It is often said that most of the forest loss in Africa is not caused by timber extraction, but rather by the subsequent settlement.

Large-scale intensive agriculture, such as has led to the destruction of huge areas of tropical woodland in South America, has barely touched Africa. There are, however, signs that some countries outside Africa are looking to its relatively lightly populated regions as a potential 'bread basket' to address their own problems of food security. Large-scale commercial agriculture would not only impose heavy loads on fragile ecosystems, but also displace traditional farmers whose practices are in tune with their environment.

Human populations are increasing throughout Africa, in some nations at alarming rates. It is easy to overlook that an annual increase of 3.5% implies a population doubling every 15 years, and such rates are found in several African countries. Probably a majority of Africans are aware of the problem, but the absence of state care for the old is a considerable incentive to produce numerous children; even if child mortality is high, at least some will survive to provide care for the parents in their old age. Improvements in health care tend to come before reductions in birth rates, leading to lower death rates and longer life expectancies.

A rise in population increases the pressure on land, and therefore on natural ecosystems. Cultivators spread farther into areas moist enough for agriculture, and irrigation schemes push out the cultivable boundaries. Pastoralists increase their flocks and herds, which are viewed as cash on the hoof, whose numbers tend to increase until drought or disease cuts them back. Some of the more natural ecosystems, such as those of seasonally dry regions with high wild animal populations, are enclosed in national parks or game reserves, but unless these are very large they are invariably inadequate to accommodate migratory species or those, like the elephant, that normally range over a very wide area. Within the parks and reserves, tourist pressure can also be a problem.

Political instability and the wider availability of firearms are other threats. All of the larger antelopes of the Sahara and its fringes are now endangered because of hunting with guns and motor vehicles, as well as competition with domestic stock for forage and water. The migratory animal populations of southern Sudan have been greatly reduced by hunting with the many firearms now in the region. The state of the endemic large mammals in war-torn Somalia is largely unknown. In the 1970s and 1980s, elephants were hunted throughout Africa for their ivory.

Undisturbed forest is now a rare commodity in the whole of West and East Africa; only in the Congo River Basin are there tracts still in their original state, protected by their size and inaccessibility (Burgess *et al.*, 2004b). Exploitation of forests for timber will no doubt continue and probably intensify, not only for timber for export and internal use but also for charcoal to fuel the cookers of the cities. Additionally, more areas will be cleared for plantation agriculture and for cash crops such as oil palm and cocoa. The woodlands will continue to be heavily exploited for firewood and charcoal, and more favorable sites will be cleared and converted to large-scale agriculture, if soil

fertility can be maintained. Fire frequency in the grasslands will continue to increase as the human population increases, although the spread of agriculture tends to reduce fire by fragmenting the area that can be burned.

Yet here and there are signs that change may be taking place. In West Africa there is a tradition of preserving a patch of forest near every village to provide a burial ground, a home for the ancestral spirits, and a source of plants for medicine and poles for building. These sacred groves provide refugia for some forest species. In parts of the Guinea Savanna zone, schemes for sustainable exploitation of the woodlands for firewood and charcoal have been developed and may be taking hold. Overall, however, the future for many plants and animals in Africa, particularly those of the forests, is still bleak.

See also: Desert Ecosystems. Fires, Ecological Effects of. Mediterranean-Climate Ecosystems. Rainforest Ecosystems, Animal Diversity. Rainforest Ecosystems, Plant Diversity. Tropical Forest Ecosystems

References

Burgess N, Salehe J, Doggart N, *et al.* (2004a) Coastal forests of eastern Africa. In: Mittermeier RA, Robles Gil P, Hoffmann M, *et al.* (eds.) *Hotspots Revisited*, pp. 231–239. Washington/Mexico: Conservation International/CEMEX.

Burgess N, D'Amico Hales J, Underwood E, *et al.* (2004b) *Terrestrial Ecoregions of Africa and Madagascar: A Conservation Assessment.* Washington, DC: WWF & Island Press.

Cobb S and Lock M (1988) Wildlife. In: Howell P, Lock M, and Cobb S (eds.) *The Jonglei Canal. Impact and Opportunity*, pp. 350–374. Cambridge: Cambridge University Press.

Cowling RM and Richardson DM (1995) *Fynbos. South Africa's Unique Floral Kingdom.* Vlaeberg, South Africa: Fernwood Press.

Cowling RM, Richardson DM, and Pierce SM (1997) *Vegetation of Southern Africa.* Cambridge, UK: Cambridge University Press.

Frost PGH, Timberlake JR, and Chidumayo E (2002) Miombo-Mopane woodlands and grasslands. In: Robles Gil P, Mittermeier RA, Goettsch Mittermeier C, *et al.* (eds.) *Wilderness: Earth's Last Wild Places*, pp. 183–204. Washington/Mexico: Conservation International/CEMEX.

Hall JB and Swaine MD (1981) *Distribution and Ecology of Vascular Plants in a Tropical Rain Forest: Forest Vegetation in Ghana.* The Hague/Boston/London: Dr. W. Junk.

Kingdon J (1997) *The Kingdon Field Guide to African Mammals.* San Diego and London: Academic Press.

Koechlin J, Guillaumet J-L, and Morat P (1974, reprinted 1997) *Flore et Végétation de Madagascar.* Vaduz, Liechtenstein: ARG Gautner Verlag.

Lawson GW (ed.) (1986) *Plant Ecology in West Africa: Systems and Processes.* Chichester, UK: John Wiley & Sons.

Lind EM and Morrison MES (1974) *East African Vegetation.* London: Longman.

Lock JM (1998) Aspects of fire in tropical African vegetation. In: Huxley CR, Lock JM, and Cutler DF (eds.) *Chorology, Taxonomy and Ecology of the Floras of Africa and Madagascar.* Kew, London: Royal Botanic Gardens.

Richmond MD (1997) *A Guide to the Seashores of Eastern Africa and the Western Indian Ocean Islands.* Stockholm: SIDA.

White F (1983) *The Vegetation of Africa. A Descriptive Memoir to Accompany the UNESCO/AETFAT/UNSC Vegetation Map of Africa.* Paris: UNESCO.

African and Asian Savannas

Mahesh Sankaran, National Center for Biological Sciences, Bangalore, India, and University of Leeds, Leeds, UK
Jayashree Ratnam, National Center for Biological Sciences, Bangalore, India

© 2013 Elsevier Inc. All rights reserved.

Glossary

Arid A climate lacking in moisture, especially having insufficient rainfall to support many trees or woody plants.
C_3 plant A plant in which CO_2 is first fixed into a compound containing three carbon atoms before entering the Calvin cycle of photosynthesis. Most trees and temperate plants are C_3 plants.
C_4 plant A plant in which CO_2 is first fixed into a compound containing four carbon atoms before entering the Calvin cycle of photosynthesis. Most tropical grasses are C_4 plants.
Mesic A climate with medium moisture supply, neither too dry nor too wet, with sufficient rainfall to support trees and woody plants.
Monsoon A wind system that influences large climatic regions and reverses direction seasonally, bringing with it heavy rainfall to the region. Most of continental south Asia gets monsoon rains.
Savanna A mixed tree and grass system, with trees ranging from sparse to medium density, with an understory of grasses. Savannas are open, well lit systems. Rainfall usually occurs in the warmer, summer months with a dry period of 2–8 months. Fires are typical across savannas during drier months, and occur at intervals from 1 to 50 years.
Topkill Stem death as a result of fire, although the plant as a whole still survives and can resprout from root-stocks. Seedlings and saplings are particularly vulnerable to topkill.
Tropical dry forest A type of forest found in tropical to sub-tropical regions that has distinct rainy and dry seasons. Many tropical dry forest plants are adapted to withstand high temperatures and seasonal droughts.

Introduction

To most people, the term 'savanna' conjures up images of wide expanses of open grasslands, dotted with a few flat crowned trees, and grazed by large herds of roving ungulates (Huntley, 1982). Whilst these images, characteristic of nutrient-rich East African savannas, remain amongst the most enduring in the public consciousness, they fail to capture the diversity of physical, structural, and functional attributes that characterize the different savannas of the world.

The term *savanna* is thought to be derived from a Carib Amerindian word meaning 'treeless plain' (Scholes and Walker, 1993; Shorrocks, 2007). It appears to have first been recorded in print in Spanish by G.F. de Oviedo y Valdes (1535) who noted "This name 'sabana' is applied to land which is without trees, but with much grass either tall or short" (Bourlière and Hadley, 1970, 1983). Since then, the term has been widely used and variously defined, and extended to include trees (Bourlière and Hadley, 1983). In its currently accepted usage, it refers to systems with a continuous grass layer and a discontinuous stratum of trees (Scholes and Archer, 1997).

Tropical savannas can be more explicitly defined as mixed tree-C_4 grass systems, with an understory predominantly comprised of warm season grasses that utilize the C_4 photosynthetic pathway, while the trees utilize the C_3 pathway (Huntley, 1982; Scholes and Archer, 1997; Sage, 2004; Beerling and Osborne, 2006; Ratnam *et al.*, 2011). C_3 photosynthesis, in which CO_2 is first fixed to produce a three-carbon compound phosphoglyceric acid (PGA), is the more ancestral form, having dominated the process of photosynthetic CO_2 fixation for over 99% of the time since it evolved ~2.7 billion years ago (Osborne and Beerling, 2006). Plants with the C_4 photosynthetic pathway, predominantly grasses, evolved more recently (25–32 Mya (Million years ago)), and have a CO_2-concentrating mechanism based on four-carbon acids, which confers a high photosynthetic efficiency in warm climates and low CO_2 environments (Sage and Kubien, 2003; Sage, 2004; Osborne and Beerling, 2006; Osborne, 2008; Osborne and Freckleton, 2009).

As a biome, tropical savannas are relatively young when compared to other forested biomes of the tropics (Bourlière and Hadley, 1983). Their origin dates back to the late Miocene, ~8–10 Mya, corresponding to the near-synchronous expansion of C_4 grasses around the world, and represents one of the most dramatic examples of biome assembly in the geological record (Beerling and Osborne, 2006; Edwards *et al.*, 2010). The open habitats and newly available food resources associated with expansion of tropical savannas during the Miocene is also linked to the dramatic radiation and assembly of new communities of large herbivores (including horses, rhinos, antelope, and elephants), which evolved to exploit these newly available niches (see Beerling and Osborne, 2006). The spread of savanna ecosystems in Africa is also believed to have played a key role in the evolution of the early ancestors of humans and in the evolution of traits such as bipedalism and dietary adaptations to novel foods (Cerling *et al.*, 2011).

Today, savannas constitute the world's second largest biome, covering ~33 million km^2 or nearly 20% of the earth's land surface (Scholes and Walker, 1993; Ramankutty and Foley, 1999; Beerling and Osborne, 2006). They are widespread across Africa, Asia, South America, and Australia, and cover more than half the area of the southern continents. They support a fifth of the world's human population and a majority of its rangeland and livestock biomass (Scholes and

Archer, 1997). They are also home to some of the greatest diversity and densities of wild herbivores and carnivores of any ecosystem on earth.

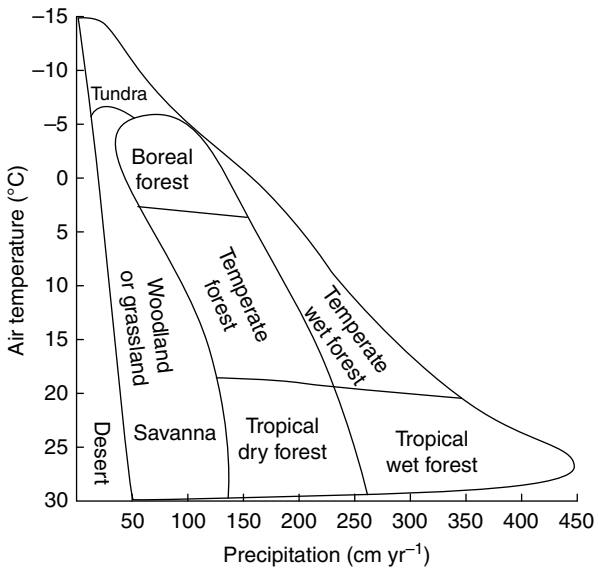


Figure 1 Distribution of major terrestrial biomes of earth with respect to mean annual precipitation and temperature (Redrawn with permission from Whittaker RH (1975) *Communities and Ecosystems*, 2nd edn. New York: McMillan).

Distribution of Savannas

Biogeographers have traditionally mapped the occurrence and distribution of biomes across the globe as a function of temperature and precipitation (e.g., Whittaker, 1975; **Figure 1**). In these frameworks, savannas occur between deserts and tropical dry forests, but the transitions between grassland, savanna, and woodland have not been easy to classify (**Figure 1**). This reflects a long standing unresolved issue in the understanding of the savanna biome. The limits to its distribution and the nature of the biome boundaries at both the arid and the mesic end are highly variable across the globe, and the dynamics underlying these patterns remain poorly understood. Many landscapes in the mesic end of the savanna biome, for example southern peninsular India, mainland Southeast Asia from Myanmar through northern Thailand, and the Guinean and Congolian savanna-forest regions in Africa are characterized by a dynamic mosaic of vegetation with open savannas, woodland savannas, and tropical dry forests occurring in a patchwork within the same climatic envelope (White, 1983; Stott, 1991; Ratnam *et al.*, 2011). Climate alone thus appears insufficient to explain the distribution of the savanna biome (Bond, 2008; Hirota *et al.*, 2011; Staver *et al.*, 2011). Further, the physiognomy of anthropogenic savannas is very often similar to that of natural ones, further complicating efforts to map the 'natural' occurrence of savannas (Bourlière and Hadley, 1970).

Tropical savannas dominated by C_4 grasses are extensive and largely occur in a latitudinal band $\sim 30^\circ$ N and S of the equator (**Figure 2**, Bond, 2008; Edwards *et al.*, 2010).

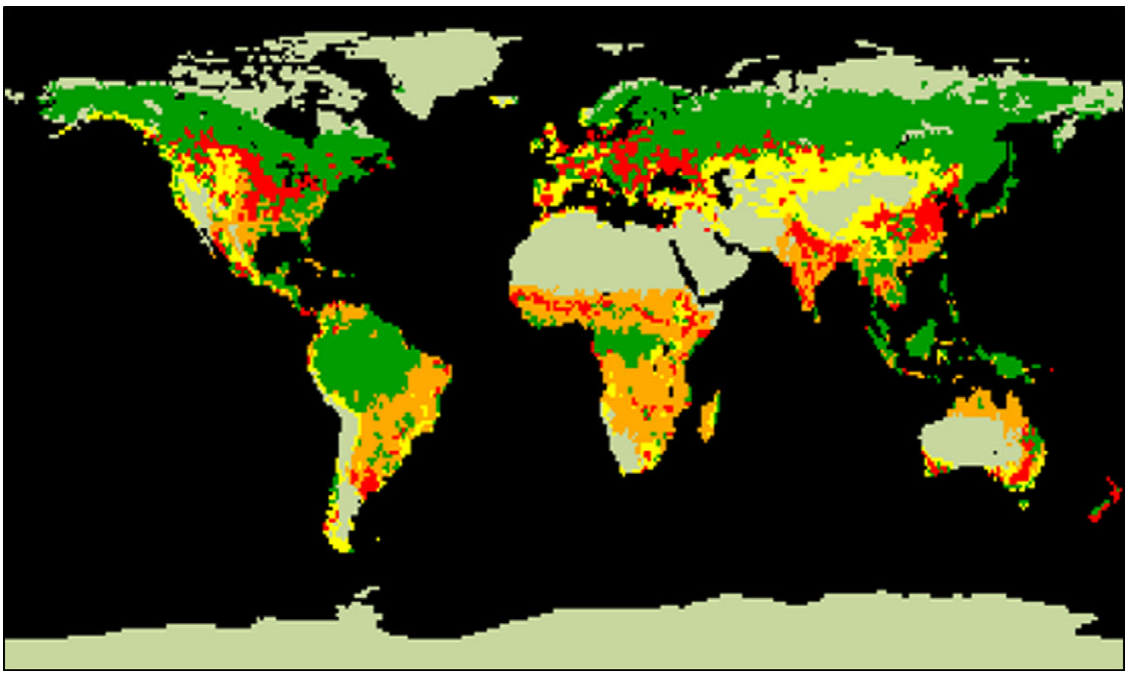


Figure 2 Global distribution of savannas and grasslands. Orange regions are systems where C_4 grasses predominate, while systems where C_3 grasses dominate are shown in yellow. Forests (green), croplands (red), and desert and ice-covered regions (beige-green) are also shown. The dots indicate regions where the geological history of C_4 grasses has been well described. Reproduced with permission from Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, and C_4 Grasses Consortium (2010) The origin of C_4 grasslands: Integrating evolutionary and ecosystem science. *Science* 328: 587–591; Map credit: Vernon Visser.



Figure 3 Variation in tree cover across different savanna sites: (a) open savanna with a short-grass understory in Samburu National Park, Kenya; (b) open savanna with a mid- to tall-grass understory in Kruger National Park, South Africa; (c) wooded savanna in Kruger National Park and (d) wooded savanna with tall-grass understory in Mudumalai Tiger Reserve, India. ((d): Courtesy of Edward February).

They occupy sites across a range of soil types and rainfall regimes (Huntley, 1982; Shorrocks, 2007). All savannas are characterized by a distinct dry season that ranges anywhere from 2 to 9 months of the year, and during which fires are a typical phenomenon (Schimper, 1903; Sarmiento, 1984; Stott, 1988; Scholes and Walker, 1993; Lloyd *et al.*, 2008). With this caveat, depending on the continent, rainfall in the wet season can range from 200 to 2500 mm (Nix, 1983; Stott, 1988; Lehmann *et al.*, 2011; Staver *et al.*, 2011).

Across their range of occurrence, savannas are extremely variable in their physical and structural attributes. They encompass a gradient from nearly pure grasslands to closed woodlands (**Figure 3**), exhibit differences in the characteristics of dominant trees (fine-leaved vs. broad-leaved), herbaceous vegetation (tall vs. short grasses, vegetated vs. bare patches), plant life-history characteristics (deciduous vs. evergreen trees, annual vs. perennial grasses), tree spatial patterns (random, regular, or clumped) and plant and soil nutrient status (nutrient-poor or dystrophic savannas vs. nutrient-rich or eutrophic savannas; House *et al.*, 2003; Sankaran and Anderson, 2009).

This inherent variability in savannas across their distributional range can pose problems when it comes to characterizing and mapping savannas, and developing a synthetic understanding of their functioning. As Lamotte and Bourlière (1983) noted: "… there is no such thing as a 'typical' savanna ecosystem. Rather, there is a gradient of related ecosystems ranging from open woodlands to almost treeless 'steppes'. However, these systems all share a number of common features which make them similar to each other rather than different". In the following sections, we focus on the commonalities of these systems in terms of the different driver variables, the interactions amongst which lead to the observed variability in their structure, composition, and functioning.

Drivers of Savanna Structure and Function

Savannas are ecologically unique in featuring the coexistence of two contrasting plant life-forms – trees and grasses – that share resources to a meaningful degree. How these two contrasting life-forms coexist without one displacing the other has been the subject of much research, and has been termed the 'savanna question' (Sarmiento, 1984). Traditionally, four key variables have been identified as being critical to determining savanna structure: water, nutrients, fire, and herbivory (Frost *et al.*, 1986). However, ecologists have been polarized in their views on the relative roles of 'bottom-up' forces (water, soil nutrients) versus 'top-down' forces (fire and herbivory) in fostering coexistence of both life-forms and determining the relative balance of trees and grasses in savannas.

Bottom-Up Control of Savannas

In the 'bottom-up' view, tree–grass coexistence and relative abundance in savannas is largely determined by the differential abilities of these two life-forms to partition and acquire limiting resources (Walter, 1971; Walker *et al.*, 1981; Walker and Noy-Meir, 1982; Scholes and Archer, 1997; van Langevelde *et al.*, 2003). Coexistence arises from spatial and/or temporal niche differences between trees and grasses in the acquisition of limiting resources, primarily water. The classic 'bottom-up' model of savannas is the root niche separation model (Walter, 1971). Here, spatial differences in rooting profiles allow trees and grasses to coexist by differentially exploiting soil water. Shallow-rooted grasses exploit only the upper layers of the soil profile, whereas deep-rooted trees are able to additionally access water from lower down in the soil profile. In this view, a characteristic tree–grass ratio is expected for a given set of rainfall and soil conditions, with tree cover increasing as one moves from arid to mesic sites. As rainfall increases across this gradient, more water percolates to depths, increasingly favoring trees over grasses. However, factors that alter the ratio of subsoil to topsoil water (e.g., variable rainfall patterns, grazing) can cause realized tree–grass ratios to deviate from this predicted ratio (Walker and Noy-Meir, 1982; van Langevelde *et al.*, 2003).

Temporal niche separation by phenology has also been proposed as a mechanism potentially contributing to tree–grass coexistence in savannas (Scholes and Archer, 1997; House *et al.*, 2003). The ability to store water and nutrients allows savanna trees to flush leaves and achieve full leaf expansion prior to, or shortly after the onset of the rains, while grasses only achieve their peak leaf area later in the growing season (Scholes and Archer, 1997). Savanna trees also typically retain their leaves for weeks following grass senescence (Scholes and Archer, 1997). Even though grasses are believed to be the better competitors for soil water during periods of growth overlap, exclusive access to soil water early and late in the growing season allows trees to persist in the system (Scholes and Archer, 1997; Sankaran *et al.*, 2004).

Most models of 'bottom-up' controls in savannas have focused on niche partitioning with respect to water rather than nutrients (Walter, 1971; Walker *et al.*, 1981; Walker and Noy-Meir, 1982; van Langevelde *et al.*, 2003). Since nutrient uptake is tightly coupled to water uptake, the models previously described are also applicable to the case of a single nutrient limiting both tree and grass growth. However, coexistence can also theoretically arise if trees and grasses are limited by, and have differential abilities to acquire, different nutrients.

In the 'bottom-up' perspective, savannas are viewed as 'stable' systems to the extent that disturbances such as fire and grazing, although capable of shifting the balance between trees and grasses, are not prerequisites for the persistence of both life-forms in the system.

Top Down Control of Savannas

The 'top-down' view represents a departure from traditional equilibrium models of savannas in that the emphasis is on demographic, rather than physiological mechanisms (Higgins *et al.*, 2000; van Wijk and Rodriguez-Iturbe, 2002; Sankaran *et al.*, 2004). Here, primacy is given to the roles of disturbances such as herbivory and fire in regulating savanna structure. Both fire and grazing act to regulate tree cover in savannas by imposing demographic bottlenecks, or in some cases eliminating bottlenecks, to tree recruitment and establishment (Higgins *et al.*, 2000; Sankaran *et al.*, 2004; Bond, 2008). The nature of the bottleneck can vary depending on the environment (Higgins *et al.*, 2000; Bond, 2008). In arid savannas, the primary bottleneck for trees is at the germination and seedling establishment stage (Jeltsch *et al.*, 1998; Higgins *et al.*, 2000; van Wijk and Rodriguez-Iturbe, 2002). Localized deposition of tree seeds in herbivore dung can provide more suitable conditions for tree seedling germination and establishment, thus eliminating the bottleneck and allowing trees to persist in the system (Jeltsch *et al.*, 1998). Similarly, grazing can reduce grass competition and favor tree establishment in areas where seedlings might otherwise have been outcompeted by grasses (Sankaran *et al.*, 2004). In contrast, in more mesic sites, disturbances such as fire and browsing serve to impose bottlenecks on seedling establishment and transition to adulthood. Frequent fires 'topkill' saplings, and result in direct mortality or force saplings to resprout from rootstocks (Higgins *et al.*, 2000). Adult savanna trees, however, are fairly immune to the effects of grass-fueled fires (Higgins *et al.*, 2000; Bond, 2008). Saplings only escape the 'fire trap' and transition into adult size classes when the interval between successive fires is long enough to allow them to grow above the flame zone (Higgins *et al.*, 2000). Browsers can similarly impose bottlenecks on tree establishment either directly as a result of browsing-induced mortality of seedlings, or indirectly, by slowing down seedling growth rates and thus maintaining them within the 'fire trap' (Sankaran *et al.*, 2004).

In the top-down view, savannas are essentially considered to be 'unstable' systems. Pure forests and grasslands are presumed to be the only stable states, and disturbances such as fire and grazing permit savannas to exist by 'buffering' the system against transitions to either extreme (Jeltsch *et al.*, 1996, 1998, 2000; Higgins *et al.*, 2000). Savannas exist because of factors that favor tree establishment at the arid end of the rainfall gradient, and factors that prevent canopy closure at the mesic end.

Limits to Woody Vegetation in Savannas

The importance of bottom-up (water, nutrients) and top-down forces (fire, herbivory) for savanna structure has been long recognized (Sarmiento, 1984; Frost *et al.*, 1986; Scholes and Archer, 1997; Stott, 1991), but their relative importance has also long been debated (Scholes and Archer, 1997; Bond *et al.*, 2003; Sankaran *et al.*, 2004; Bond, 2008). In an effort to resolve these issues, ecologists have recently begun to synthesize data across large environmental gradients in an attempt to gain a more comprehensive understanding of savanna woody community dynamics (Bond *et al.*, 2003; Sankaran *et al.*, 2005; Bucini and Hanan, 2007; Good and Caylor, 2011). Much of this synthesis has centered on African savannas, with very little from Asia, and as a result our knowledge of the dynamics of Asian savannas is rather limited.

For Africa, water availability has been identified as a critical determinant of savanna structure, setting limits to the maximum tree cover that can be supported in sites (**Figure 4(a)**, Sankaran *et al.*, 2005). Analysis of data from over 850 savanna sites distributed across the African continent reveals an upper bound on maximum potential tree cover in arid and semiarid savannas that receive between ~150 and 650 mm rainfall annually (**Figure 4(a)**, Sankaran *et al.*, 2005). Within this rainfall range, maximum potential tree cover that can be supported in sites increases linearly with annual rainfall (**Figure 4(a)**), but appears unrelated to fire frequency, herbivory, or soil properties (Sankaran *et al.*, 2005). Above the 650 mm threshold, mean annual rainfall appears sufficient to allow for a closed woody canopy, such that disturbances like fire and herbivory become critical for the persistence of both trees and grasses in the more mesic savannas (Bond *et al.*, 2003; Sankaran *et al.*, 2005). The importance of mean annual rainfall in setting limits to tree cover that can be supported in arid and semiarid savannas has also been documented in other continental scale analyses based on remotely sensed estimates of tree cover in African savannas (Bucini and Hanan, 2007; Good and Caylor, 2011; Greve *et al.*, 2011).

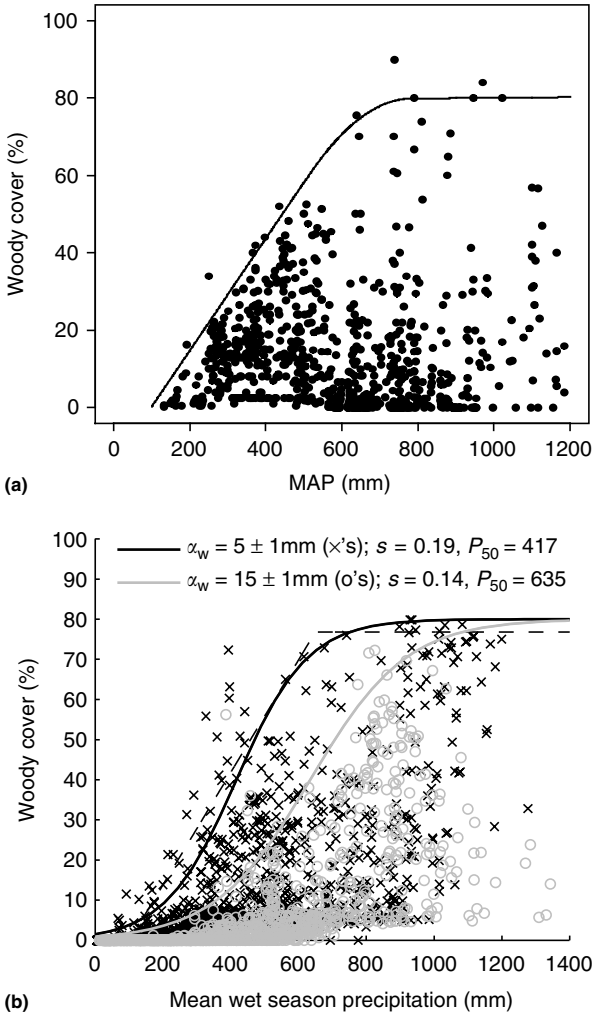


Figure 4 (a) Maximum tree cover in Africa's arid savannas (<650 mm MAP) is limited by water availability, while water availability in mesic savannas (>650 mm MAP) appears sufficient to support a closed canopy. The upper bound on tree cover was fitted using a 99th quantile piecewise linear regression model. The breakpoint, that is, rainfall at which maximum tree cover is attained lies in the interval 650±134 mm MAP. Data are from 854 sites across Africa (Reproduced from Sankaran M, Hanan NP, Scholes RJ, *et al.* (2005) Determinants of woody cover in African Savannas. *Nature* 438: 846–849). (b) Effects of wet season rainfall depth (α_w) on woody cover in African savannas. The dark line represents the maximum potential woody cover that can be supported in sites with less intense rainfall events (mean rainfall depth α_w=5±1 mm; x symbols), and the gray line for sites with more intense rainfall events (α_w=15±0.5 mm; o symbols). Sites with less frequent but more intense rainfall events support lower tree cover than sites with more frequent but less intense events. Tree cover data are derived from MODIS and precipitation data from TRMM (Reproduced from Good SP and Caylor KK (2011) Climatological determinants of woody cover in Africa. *Proceedings of the National Academy of Sciences* 108: 4902–4907).

Besides total rainfall, rainfall seasonality has also been shown to play an important role in determining tree cover in sites (**Figure 4(b)**, Good and Caylor, 2011). Good and Caylor (2011) demonstrated that for areas with similar seasonal rainfall amounts, sites characterized by frequent, less intense precipitation events tended to have higher tree cover than sites with less frequent, but more intense precipitation events. When precipitation events are intense, more water runs off as surface flow and less percolates into soils, reducing water availability and lowering tree cover compared to sites where rainfall events are less intense.

In other recent work examining the limits of savanna occurrence across the southern continents, Lehmann *et al.* (2011) also highlight the importance of rainfall seasonality as a predictor of savanna occurrence in any given landscape. The authors suggest a mechanistic explanation for the same: rainfall seasonality is important because it simultaneously decreases rates of woody canopy closure (Sarmiento, 1984; Good and Caylor, 2011) and increases the probabilities of disturbances such as fire that open up woody canopies (Archibald *et al.*, 2009). The counter-balance between these two processes is critical to maintain the open state that characterizes savannas.

Although the existence of an upper bound on tree cover in arid and semiarid savannas is consistent with the predictions of 'bottom-up' or niche-based models of tree–grass coexistence (Walker *et al.*, 1981; Walker and Noy-Meir, 1982), the actual mechanisms underlying the observed pattern remain unclear. For one, Walter's root niche separation hypothesis (Walter, 1971), which is the mainstay of 'bottom-up' models, is not empirically supported in many savannas (Scholes and Archer, 1997; Higgins *et al.*, 2000; Sankaran *et al.*, 2004; Bond, 2008; Kulmatiski *et al.*, 2010). Root-niche separation models also do not consider the seedling stages of trees when there is no spatial separation of roots and seedlings compete directly with grasses for soil resources (Sankaran *et al.*, 2004; Bond, 2008). This, combined with the lack of empirical support, has led many authors to conclude that the root-niche separation model is not 'tenable' as a general explanation for the observed tree cover patterns in savannas (Bond, 2008). The role of phenological differences between trees and grasses in generating the observed bound on maximum tree cover in arid

and semiarid savannas is similarly unclear. Thus, while it is clear that resource availability, particularly water, plays a critical role in defining the maximum potential tree cover that can be supported at any site, the specific mechanisms generating the observed pattern remain unclear.

Factors Regulating Savanna Structure

The observed variability in tree cover across rainfall gradients in African savannas (**Figure 4**) suggests that maximum woody cover is rarely realized, with tree cover in most savannas well below the 'climatic' potential (Bond *et al.*, 2003; Bond, 2008; Sankaran *et al.*, 2008). Clearly, other processes besides water availability play a role in defining patterns of realized or observed tree cover in savannas. Analysis of observed tree cover data from a range of sites across Africa indicates that after water availability, fires accounted for much of the observed variation in tree cover, followed by soil properties and herbivory (Sankaran *et al.*, 2008). Similar analyses have not been conducted for Asian savannas; in fact, Asian savannas have been singularly omitted in many texts and large scale syntheses of savanna dynamics (Huntley and Walker, 1982; Lehmann *et al.*, 2011; Hirota *et al.*, 2011; Staver *et al.*, 2011). However, it is likely that all of these factors are just as important, even if to varying degrees, in regulating savanna structure in Asia.

Fire

Fires are well recognized as a major driver of savanna structure (Higgins *et al.*, 2000; Bond *et al.*, 2003, 2005; Bond, 2008). Savanna fires are typically surface fires, fueled by grasses, and rarely spread to tree canopies or kill adult trees (Higgins *et al.*, 2000; Bond, 2008). Their effects on savanna structure largely arise through the bottlenecks they impose on seedling and sapling survival and growth, with the extent and severity of the bottleneck varying depending on fire season, frequency, and intensity (Higgins *et al.*, 2000; Bond, 2008; Midgley *et al.*, 2010). Fires topkill saplings and force them to resprout to survive, thus 'trapping' them within the grass flame zone, or 'fire trap' (Higgins *et al.*, 2000). Fire impacts on regrowth and survival are highly dependent on size and species (Gignoux *et al.*, 1997; Mistry, 2000; Hoffmann and Solbrig, 2003; Midgley *et al.*, 2010). Within a species, larger saplings typically tend to survive better (Hoffmann and Solbrig, 2003; Bond, 2008), while across species growth and survival in the face of frequent fires can vary depending on factors such as bark thickness and bark moisture content (Gignoux *et al.*, 1997; Midgley *et al.*, 2010). For example, tree species characteristic of forests tend to have thinner barks than savanna species and are less fire tolerant than savanna species (Hegde *et al.*, 1998; Hoffmann *et al.*, 2009). Further, the ability of savanna tree saplings to 'juggle carbon' between growth, maintenance, and storage, such that root reserves are replenished before the next fire event, is probably a key factor allowing many savanna species to persist indefinitely within the flame zone (Schutz *et al.*, 2009; Wigley *et al.*, 2009). In *Acacia karroo*, root starch reserves can be replenished, and match those of unburned saplings, within a year following topkill; thus allowing it to survive in the face of very frequent fires (Schutz *et al.*, 2009). Species that lack the ability to replenish root reserves quickly are unlikely to persist in fire-prone environments. Fires thus act as a filter selecting for species with fire-adapted traits and are not only regulators of savanna structure, but also of its composition.

Besides modifying woody vegetation structure where they occur, fires also serve to extend the distribution of savannas beyond their climatically determined limits (Bond *et al.*, 2005). Bond *et al.* (2005) used dynamic global vegetation models (DGVMs) to compare potential and realized vegetation in a world with and without fire. Models that included fire better predicted current global vegetation distributions when compared to models in which fire had been 'switched off', particularly for systems dominated by C_4 grasses (**Figure 5**, Bond *et al.*, 2005; Bond, 2008). The models suggest that in the absence of fire, forests would nearly double their current extent, occupying regions that presently support savanna vegetation, both in Africa and Asia (**Figure 5**, Bond *et al.*, 2005).

Long-term fire exclusion experiments have similarly shown that fire is a critical factor for savanna trees and provides unequivocal evidence that tree biomass and species composition in many savannas, particularly the more mesic ones, are not at equilibrium with climate and resources, but rather depend on fire regimes (Bond *et al.*, 2003; Bond, 2008). Large scale analyses indicate that in many areas of the globe which are characterized by intermediate rainfall (1000–2500 mm annually) and mild seasonality (dry season <7 months), tree cover is bimodal, with both forest and savanna co-occurring within the same climatic envelope (Wilson and Agnew, 1992; Hirota *et al.*, 2011; Staver *et al.*, 2011). Forests and savannas here may represent alternate states differentiated by, and maintained through, fire feedbacks (Wilson and Agnew, 1992; Staver *et al.*, 2011). Frequent fire maintains open canopies with fire-tolerant C_4 grasses in the understory, which in turn promote fires and the savanna state in a positive feedback loop. Above a certain threshold of tree cover, shade-intolerant C_4 grasses are excluded from the system, reducing both fuel loads and fuel contiguity, inhibiting fires and maintaining the system in a forested state (Staver *et al.*, 2011).

Nutrients

Soil properties, including texture and the availability of mineral nutrients such as nitrogen (N) and phosphorus (P), also play an important role in regulating savanna structure (Scholes and Archer, 1997; Sankaran *et al.*, 2008). Tree cover tends to be higher in savannas on coarse-textured sandy soils compared to those on fine-textured clay soils (Williams *et al.*, 1996). Sandy soils allow for greater water percolation to depths, thus favoring deeper rooted woody vegetation (Walker and Noy-Meir, 1982). Tree cover across a range of African savanna sites has also been shown to have a strong negative dependence on soil N availability, with tree cover typically lower in N-rich sites (Sankaran *et al.*, 2008). Tree seedling survival and growth have been shown to decrease with increases in soil N, both directly due to the preemption of N by herbaceous vegetation, as well as indirectly as a result of faster depletion of soil water following the stimulation of herbaceous growth (Davis *et al.*, 1999; Kraaij and Ward, 2006; van der Waal *et al.*, 2009, 2011; Cramer *et al.*, 2010).

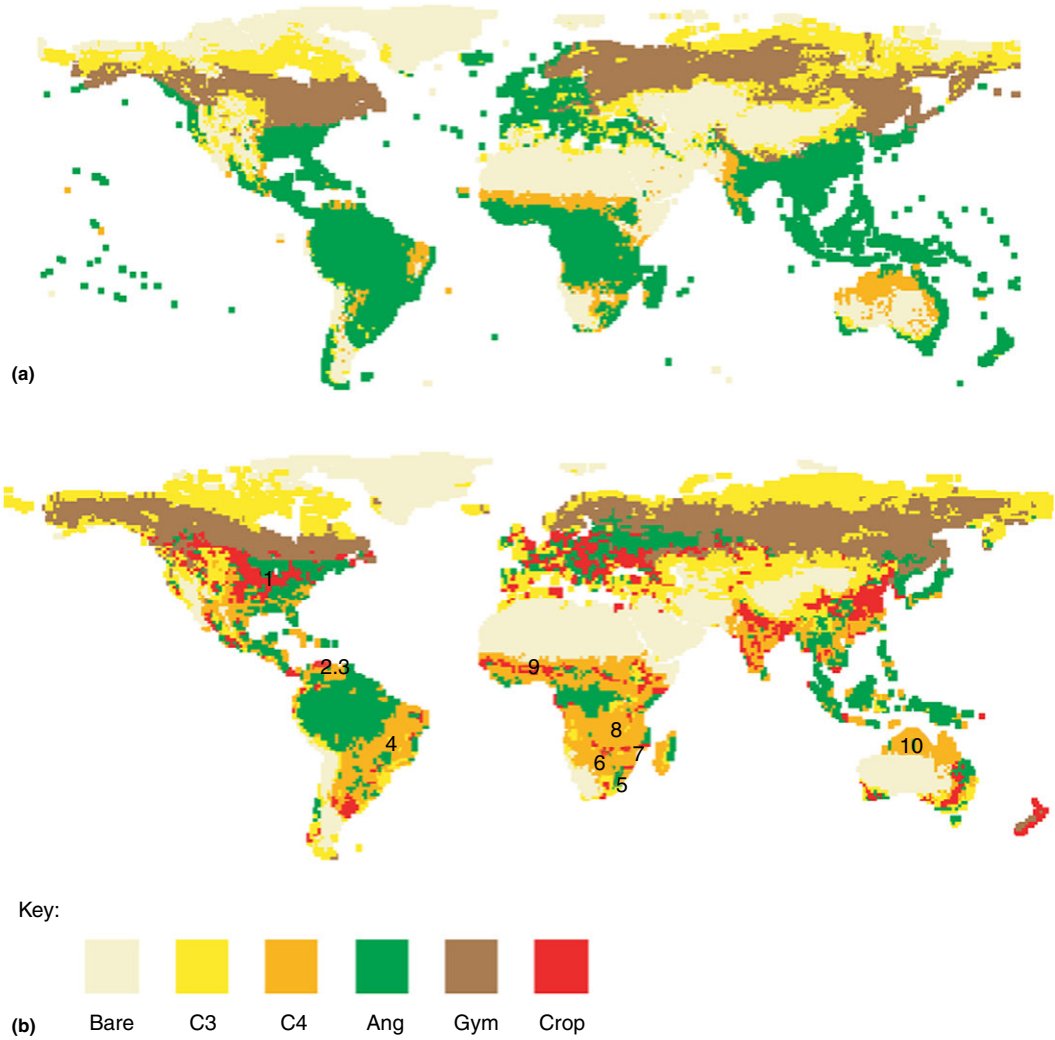


Figure 5 A comparison of (a) biome distributions simulated by a DGVM with fire switched off, and (b) actual vegetation distribution. The map of actual vegetation was sourced from ISLSCP (fttp:/daac.gsfc.nasa.gov/data/inter_disc/biosphere/land_cover/). Biomes are characterized by the dominant plant functional types: C_3 grasses or shrubs, C_4 grasses, Angiosperm trees (Ang), Gymnosperm trees mainly conifers (Gym). Croplands and bare ground (deserts, ice) are also shown. The numbers indicate locations of long-term field experiments of fire suppression (see Bond *et al.*, 2005; Bond and Keeley, 2005). The simulations indicate that high rainfall sites show a successional tendency toward forests in the absence of fire, suggesting that large areas of the globe are not at their climatic potential as a result of fires. Reproduced from Bond WJ, Woodward FI, and Midgley GF (2005) The global distribution of ecosystems in a world without fire. *New Phytologist* 165: 525–538, with permission from Wiley.

van der Waal *et al.* (2009) have shown that the negative effects of increased N availability on tree seedling growth and survival can be exacerbated during droughts, suggesting that the effects of soil N availability on tree–grass competition is likely to vary across rainfall gradients, with suppression of seedling growth by grasses more pronounced in arid, fertile areas compared to mesic, fertile areas.

Soil P availability has also been shown to be an important predictor of tree cover in a pan-African dataset, more so than soil N (Sankaran *et al.*, 2008). However, although total soil P explained a significant fraction of the observed variability in tree cover in this study, the nature of the relationship between total soil P and tree cover is complex and nonlinear (Sankaran *et al.*, 2008), potentially reflecting the fact that total soil P is not necessarily a good indicator of plant available P (Cramer *et al.*, 2010). Few studies have explicitly considered the effects of P availability on tree–grass competition, but available evidence suggests that, contrary to N availability, grass suppression of seedling growth may be greater in P deficient soils compared to P-rich soils (Cramer *et al.*, 2010). The reasons underlying these differential responses are unclear, and our understanding of the role of P availability in regulating savanna vegetation dynamics remains far from complete.

Although soil properties can have substantial impacts on tree cover and tree–grass interactions at local scales (White, 1983; Williams *et al.*, 1996; Staver *et al.*, 2011), the extent to which nutrient availability limits savanna distribution is unclear. Low nutrient availability has often been cited as a reason for the occurrence of savannas in areas with sufficient rainfall to support forests. For example, vegetation ecologists working

in the mesic savanna regions of mainland southeast Asia that receive between 800 and 2000 mm rainfall annually, have noted that open savanna formations tend to be associated with low nutrient edaphic conditions such as coarse sandy soils or rocky hill slopes (Puri, 1960; Champion and Seth, 1968; Blasco, 1983; Stott, 1984, 1990). However, Bond (2010), based on a nutrient-stock analysis of forests and savannas, concluded that low nutrient availability seldom constrains forest development. Rather than nutrient availability alone, its interactions with other factors such as fire and herbivory may be critical for explaining vegetation structure (Bond, 2010). In agreement with this, Lehmann *et al.* (2011) noted that soil fertility has contrasting associations with the occurrence of savannas at the arid and mesic ends of the biome. In arid regions, sites with high soil fertility support savannas, while the opposite is true for mesic regions where sites of low soil fertility are more likely to support savannas. The mechanisms underlying these effects are probably different. It is likely that fertile sites in arid regions attract herbivores which then play an important role in reducing grass cover and allowing trees to establish (Fritz and Duncan, 1994; Asner *et al.*, 2009; Staver *et al.*, 2009). In contrast, in sites with low soil fertility in mesic regions, slow rates of woody growth and canopy closure between successive fires might prevent the system from reaching a forested state that rainfall might otherwise support.

Herbivores

The importance of large mammalian herbivores for African savannas is well recognized (Scholes and Archer, 1997). Unfortunately, little is known about the impacts of native herbivores on savanna structure in Asia (Mistry, 2000). This lacuna is somewhat surprising given that Asia supports the highest richness and biomass of mammalian herbivores of any continent outside of Africa, including mega-herbivores like elephant and rhinoceros. Nevertheless, given the similarities in the herbivore assemblages between the two continents, lessons learnt from Africa are likely to be broadly applicable to Asian savannas.

The effect of mammalian herbivores on savanna vegetation is contingent on population densities and body size, and varies depending on whether the species is a browser, mixed feeder, or grazer (Augustine and McNaughton, 1998, 2004; Sankaran *et al.*, 2008). Mega-herbivores – those with a body mass >1000 kg – such as elephants, often tend to have a disproportionate effect on savanna woody vegetation and are capable of radically transforming savanna landscapes (Guldemond and van Aarde, 2008). In general, elephants tend to have a negative effect on woody vegetation, reducing tree cover and 'opening up' savannas (Sankaran *et al.*, 2008), with the magnitude of such effects varying with elephant densities and rainfall (Guldemond and van Aarde, 2008). Browsers and mixed-feeders, similarly, have negative effects on woody vegetation structure that arise from both their direct and indirect effects on tree recruitment (van Langevelde *et al.*, 2003; Augustine and McNaughton, 2004). Selective browsers can significantly depress growth rates and increase mortality of shrub seedlings and saplings, thus imposing strong bottlenecks to tree and shrub recruitment (Augustine and McNaughton, 2004). Browsers and mixed-feeders can also limit tree establishment indirectly in more mesic and fire-prone environments by reducing growth rates and maintaining seedlings and saplings within the flame zone; thus rendering them more prone to topkill from fires (Scholes and Archer, 1997; Sankaran *et al.*, 2008). Browsing on seedlings late in the growing season can also limit their ability to replenish carbohydrate reserves in root-stocks and reduce their resprouting vigor, making them more susceptible to mortality from subsequent fires (Schutz *et al.*, 2009; Wigley *et al.*, 2009). Large scale documented increases in tree cover accompanying the loss or reduction of browser populations following disease outbreaks like rinderpest are testimony to the impact of browsers on woody vegetation dynamics in savannas (Prins and van der Juegd, 1993).

Grazer effects on woody vegetation are more complex, with grazers having both positive and negative effects across gradients of grazing intensity (Sankaran *et al.*, 2008). For example, relief from grazing has been associated with increases in woody cover (Lenzi-grillini *et al.*, 1996). Similarly, overgrazing has also been implicated as a causal agent for shrub encroachment in grazing lands worldwide (Archer, 1995; Scholes and Archer, 1997; Roques *et al.*, 2001; Eldridge *et al.*, 2011). Enhanced woody cover under sustained grazing can result from multiple factors including (1) reduced competition from grazed grasses, (2) increased water availability for deeper rooted shrubs as a result of lowered uptake by grazed grasses, (3) reduced fire frequencies and intensities due to lower grass fuel loads, and (4) dispersal of woody seeds by grazers (Brown and Archer, 1987; Archer, 1995; Scholes and Archer, 1997; Roques *et al.*, 2001; Eldridge *et al.*, 2011). Further, grazers, particularly cattle, have also been shown to indirectly facilitate shrub recruitment by suppressing populations of seed and seedling predators, primarily rodents (Goheen *et al.*, 2010).

Classification of Savannas

Reflecting the above variability in the drivers and structural properties of savannas, scholars have classified savannas using a number of different criteria, depending upon the question of interest. In the following two sections, we describe the most widely used classifications of African and Asian savannas in some detail.

African Savannas

African savannas are probably the most well-researched savanna ecosystems in the globe, and consequently we have a very good idea of their distribution and vegetation. African savannas have either been distinguished based on their climatic and edaphic ranges, their physiognomy and structure, or the functional attributes of the dominant life-forms. These different systems are not always mutually exclusive and the descriptors are often used interchangeably by different sources in the literature.

Arid, Eutrophic Versus Mesic, Dystrophic Savannas

This classification highlights important differences in the water and nutrient status of savannas, which tend to covary at

the continental level in Africa (Huntley, 1982; White, 1983, **Figure 6(a)**). Savannas that occur on the more ancient surfaces of the African shield tend to be wetter and relatively nutrient poor ('dystrophic') while those that occur on the more 'recent' surfaces formed since the breakup of Gondwanaland tend to be drier and more fertile or 'eutrophic' (Scholes and Walker, 1993). This covariance at the continental scale results from the fact that nutrient leaching on the older, wetter sites has been more pronounced and prolonged than on the drier, younger surfaces on the continent, which also happen to occur on more nutrient-rich parent material (Scholes and Walker, 1993).

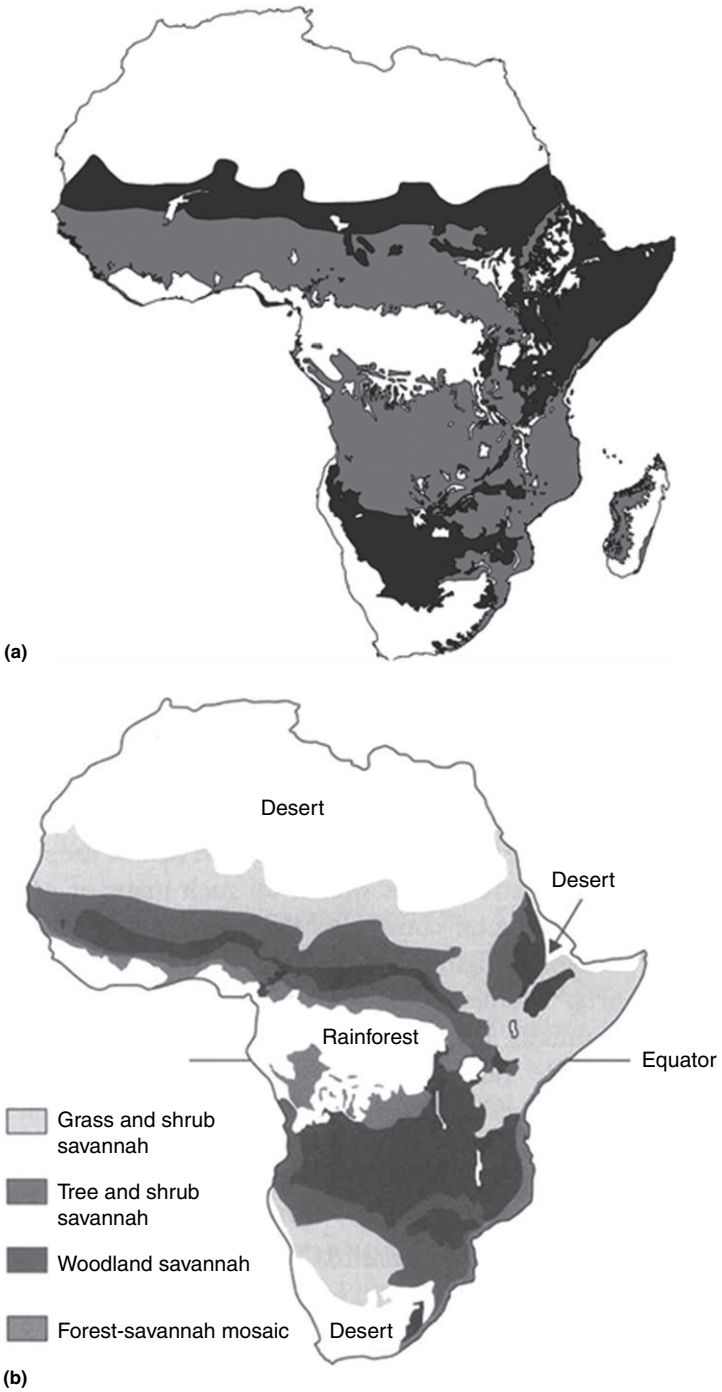


Figure 6 Distribution of (a) arid, eutrophic (black) and mesic, dystrophic (dark gray) savannas (map derived from White F (1983) The vegetation of Africa. Paris: UNESCO by combining different vegetation classes); and (b) savannas with different life-form associations across Africa. Reproduced from Shorrocks B (2007) *The Biology of African Savannas*. Oxford: Oxford University Press.

Broadly, arid, eutrophic savannas typically receive <650 mm Mean Annual Precipitation (MAP) and occur on calcareous and eutrophic noncalcareous soils (exchangeable bases >15 me/100 g clay, Huntley (1982)). Most of these savannas experience seven or more months of mid-summer droughts. Where they extend into subtropical areas, frost tends to be a regular phenomenon in arid savannas. Moist or humid savannas occur in areas that receive >650 mm MAP annually, and on dystrophic and mesotrophic noncalcareous soils (exchangeable bases <15 me/100 g clay; Huntley (1982)). Mid-summer drought is also a feature of these savanna types at the lower limit of their moisture range. Frost is generally absent over most of the range of moist savannas.

Despite the rather coarse-grained scale of such a classification, it has nevertheless proved quite useful for continental scale analyses of African savannas. In particular, the distribution and abundance of large mammalian herbivores in Africa is predictable to a large degree based on this functional classification, highlighting the importance of water and nutrients for savanna structure and function (Scholes and Walker, 1993).

Savannas With Different Life-Form Associations

Several authors have classified savanna types in Africa based on the dominant life-forms and structural attributes of these systems. Even while these classes are based on the appearance of the vegetation, they nevertheless convey a sense of the climatic zone in which a system occurs because vegetation structure is often associated in predictable ways with temperature, soil fertility, and rainfall. Here, we follow the classification scheme and descriptions developed by Shorrocks (2007) and simplified from White (1983), which combines elements of vegetation structure with information on species associations (**Figure 6(b)**). It is important to recognize that while such schemes differentiate savanna types at a continental scale, they fail to capture the inherent heterogeneity that characterizes savannas at regional or local scales.

Grass and Shrub Savannas

Familiar to most as the 'Sahelian' savanna in northern Africa, these arid savannas occupy the transition zone between the Sahara Desert in the north and the more humid savannas in the south, stretching from Senegal and Mauritania in the west to Sudan in the east. In the drier west (100–600 mm mean annual rainfall), the woody component is dominated by shrubby forms of *Acacia laeta*, *Acacia tortilis*, and *Balanites aegyptiaca* while the grass cover is dominated by annual species such *Cenchrus biflorus*, *Schoenefeldia gracilis* and *Aristida stipoides*. Where these savannas reach the horn of Africa, they merge into *Acacia–Commiphora* associations, which stretch east of the Ethiopian highlands and down into the Maasai drylands of Tanzania and southern Kenya. Here, where rainfall is bimodal and higher (600–800 mm mean annual rainfall), perennial grasses such as *Themeda triandra*, *Panicum coloratum*, *Aristida adscencionis*, and *Andropogon* and *Eragrostis* species dominate the grass community.

In southern Africa, these 'grass and shrub' formations form the transition between the southern deserts and northern mopane savanna. Mean annual rainfall ranges from 200 to 500 mm. In the northern part of this region, dense short savannas are characterized by the presence of species with succulent stems such as *Aloe dichotoma*, *Euphorbia guerichiana*, and *Moringa ovalifolia* predominately. Toward the south, the vegetation becomes more open and is dominated by karoo shrubs such as *Parkinsonia africana*, *Acacia newbrownii*, and *Boscia* species. Tufted grasses, mainly of the genus *Stipagrostis*, are found scattered in the understory.

Tree and Shrub Savanna

These savannas, which are more mesic than the 'grass and shrub' savannas above, occur in the regions immediately north and south of the rainforests and miombo woodlands of central Africa. In the northern section, mean annual rainfall ranges from 600 to 1000 mm and the climate is strongly tropical. The woody component is dominated by species of *Combretum* and *Terminalia*, with other common species being *Anogeissus leiocarpus*, *Boswellia papyrifera*, *Balanites aegyptica*, *Stereospermum kunthianum*, *Commiphora africana*, *Prosopis africana*, and *Ziziphus mucronata*. Tall grasses such as *Pennisetum purpureum* (elephant grass) and species of *Hyparrhenia*, *Cymbopogon*, and *Echinochloa* dominate the grassy layer.

South of the miombo woodlands of Central Africa, 'tree and shrub' savannas are characterized by the dominance of the mopane tree *Colophospermum mopane*, with its distinctive butterfly-shaped leaves and dense clonal stands. In the southeastern 'Zambezian' region, mean annual rainfall ranges from 450 to 700 mm, but some regions may receive as much as 1000 mm of rain. Here, the mopane tree shows great variation in its structural form, ranging from stunted 1–3 m bushes to 25 m tall trees depending on local soil conditions, with the two structural forms often occurring in a mosaic across the landscape (White, 1983). While mopane is dominant, other trees that occur in these formations include *Kirkia acuminata*, *Dalbergia melanoxylon*, *Adansonia digitata*, and species of *Combretum*, *Acacia*, and *Commiphora*. Typical grasses include species of *Aristida*, *Eragrostis* and *Urochloa*, *Digitaria eriantha*, *Brachiaria deflexa*, *Echinochloa colona*, *Cenchrus ciliaris*, *Enneapogon cenchroides*, *Pogonathria squarrosa*, and *Stipagrostis uniplumis*.

The southwestern 'Angolan' mopane covers vast regions of Namibia and Angola. Here, annual rainfall ranges from 400 to 600 mm, but is often highly unpredictable. In Namibian regions, mopane is often mono-dominant, forming dense single-species stands beneath which grass is virtually absent. In the Angolan region, mopane tends to occur as low, thorny bushland. Here it is associated with *Acacia kirkii*, *Acacia nilotica*, *Acacia hebeclada*, *Acacia erubescens*, *Balanites angolensis*, *Combretum apiculatum*, *Commiphora* species, *Dichrostachys cineria*, *Grewia villosa*, *Jatropha campestris*, *Melanthera marlothiana*, *Peltophorum africanum*, *Rhigosum brevispinosum*, *Rhigosum virgatum*, *Securinega virosa*, *Spirostachys africana*, *Terminalia*, and *Ximonia* species.

Woodland Savanna

'Woodland savanna' refers to the more mesic savannas where stands of trees occur in open formations, with some shrubs in a ground layer dominated by grasses. There are two major regions of woodland savanna in Africa – the large region called the *miombo* in central/south Africa and a smaller area called the *doka* in the north.

The miombo woodland covers an estimated 3 million square kilometers in Zimbabwe, Zambia, Mozambique, Angola, Malawi, Zaire, and Tanzania. It takes its name from the miombo tree, a species of *Brachystegia*, which is the dominant genus across this vast area. Mean annual rainfall in the center of the miombo region ranges from 800 to 1200 mm, but can decrease towards the south to 600–800 mm.

Typically, mature miombo trees are 15–20 m tall but heights can decrease in drier regions. Dominant tree species include several species of *Brachystegia*, and others such as *Marquesia macroura, Julbernardia globiflora, Julbernardia paniculata, Isoberlinia angolensis, Uapaca kirkiana, Monotes glaber, Faurea saligna, Faurea speciosa, Combretum molle, Albizia antunesiana, Strychnos spinosa, Strychnos cocculoides, Flacourtia indica*, and *Vangueria infausta*. Most of the miombo trees shed their leaves in the pronounced dry season and are bare for 2– 3 months. With high productivity from the mesic rainfall regime and a long dry season that leaves the vegetation dry, fire becomes an important ecological factor in the miombo woodland.

The northern doka savanna takes its name from the dominant tree in this region, *Isoberlinia doka*. This savanna is characterized by patches of dense trees, dominated by doka along with *Afzelia africana, Burkea africana, A. leiocarpus, Terminalia* species, and *Borassus aethiopium*.

Forest-Savanna Mosaic

The forest-savanna mosaics, as their name suggests, occur at the mesic limits of the savanna biome where savannas transition into closed-canopy forests. The savannas at these regions are densely wooded and structurally begin to approach the appearance of slightly open forests, but grasses are still dominant in the understorey (Bond and Parr, 2010; Ratnam *et al.*, 2011).

In Africa, these mosaic regions encircle the Congo basin, and are a dynamic and shifting matrix of savanna, forest, and grassland. To the north–west is the Guinean savanna forest mosaic which extends from Guinea and Ivory Coast through to Nigeria. To the north lies the Congolian forest-savanna mosaic, which extends through the Central African Republic, the Democratic Republic of Congo and into southern Sudan, while the Zambezian forest-savanna mosaic lies to the Southwest. With mean annual rainfall in these regions ranging from 1200 to 1600 mm, the vegetation has both forest and savanna elements. Trees common to wooded grasslands are widespread in the region and include such species as *Annona senegalensis, A. africana, Burkea africana, Butyrospermum paradoxum, Stereospermum kunthianum, Strychnos, Terminalia*, and *Vitex* species. Likewise, species common to dry forest such as *A. africana, Aningeria altissima, Chrysophyllum perpulchrum, Cola gigantea, Combretum collinum, Morus misozhygia*, and *Khaya grandifolia* also occur. Common grasses, many growing taller than two meters, include species of the genus *Andropogon, Hyparrhenia*, and *Loudetia*.

Asian Savannas

In contrast to the savannas of Africa, Asian savannas have been relatively less well studied, and consequently detailed knowledge about the distribution and dynamics of the different savanna types that occur in the region is lacking. Several factors have contributed to this lacuna.

Unlike Africa, where almost 50% of the landmass is arid to semiarid and dominated by savanna landscapes, much of continental south and Southeast Asia receives over 800 mm of rainfall annually, making forests and savanna woodlands much more dominant in this landmass. Further, historically, much of the early interest in the region's vegetation came from the perspective of timber and forestry operations of the early colonial period. Together, biogeography and history have resulted in a legacy of a markedly forest-centric approach to the understanding and study of vegetation in this region.

Most early regional vegetation classification schemes (e.g., Puri, 1960; Champion and Seth, 1968), which are still the major references used today, refer only to the 'forest types' of south and Southeast Asia with the word savanna rarely finding mention. Many vegetation formations in the Indian subcontinent that are functionally and structurally savannas, with woody tree or shrub components in a C_4 grass dominated understory, are nevertheless referred to as forests. These include the relatively open 'thorn and scrub forests' in the arid and semiarid regions of western India, and the more wooded 'mixed and dry deciduous forests' in the mesic regions of peninsular India (*sensu* Champion and Seth, 1968). Likewise, in Southeast Asian classification systems, the characteristic vegetation type of the lowland regions from northern Burma through Thailand, Laos, Cambodia, and Vietnam, an open dry deciduous dipterocarp forest with the ground covered by grasses, is variously referred to as 'forêts claire' or 'open forest' or 'savanna forest' (Boulbet, 1982; Stott, 1990). As Stott (1991) noted of these dry dipterocarp forests which bear striking similarities to the miombo woodlands of Africa: "To a large extent, these associations have been singularly neglected by scholars of savanna vegetation, and they remain very much the 'Cinderellas' of the subject, even in modern texts."

It is likely that the above nomenclature has contributed to the widespread perception that all of the open, savanna-like formations in the region today are derived, that is, they were originally forests that have been converted to savannas as a result of human activities and disturbances such as fire and grazing (Mishra, 1983; Gadgil and Meher-Homji, 1985; Yadava, 1990), but this is likely to be a misconception (Stott, 1991; Mistry, 2000; Shorrocks, 2007; Ratnam *et al.*, 2011). While it is undoubtedly true that long-term pervasive human presence has resulted in large tracts of 'derived' savannas in the region, it is also likely that several areas supported 'natural' savanna vegetation in the past, and continue to do so today (Stott, 1991; Pemadasa, 1991; Shorrocks, 2007; Ratnam *et al.*, 2011). In particular, several authors consider the dry dipterocarp forests of Southeast Asia to be natural 'climax' formations with edaphic origins on coarse sandy soils (Boulbet, 1982; Blasco, 1983; Stott, 1988, 1990). This idea is further supported by early observations in northern Thailand by Barrington (1931) who noted that "30 or 40 years of fire protection had no appreciable effect on the vegetation." Thus, even while these formations are referred to as 'forests', it is clear that their savanna-like characteristics have been long recognized.

As a consequence of the historical forest bias of vegetation scholars in this region, there is no detailed classification or

maps of the different types of Asian mixed tree–grass systems; **Figure 7** is extracted from a recent mapping of the vegetation of the region using satellite imagery (Blasco *et al.*, 1996), and while its scale of resolution does not detail different types of savanna, it provides the best available estimates of current day extents of mixed tree–grass systems in this region. Below, the authors draw on existing vegetation texts for the region (Champion and Seth, 1968; Blasco, 1983; Stott, 1984; Mistry and Stott, 1993) to reconsider the nature and extent of mixed tree–grass systems in south and Southeast Asia. In the interest of continuity, the authors retain the commonly used forest nomenclature (Champion and Seth, 1968) but focus on the tree–grass aspects and draw attention to their analogies with African savanna systems.

Dry Thorn Forests

These are open tree–grass systems of short stature with trees being 6–9 m in height. There is an assorted lower story of small trees and shrubs. There is usually thin grass growth, which appears abundant in the wet season, but much of the ground is bare in the dry season. The southern forms of this vegetation type extend across the dry tracts of peninsular India and along the lee of the Western Ghats, where mean annual rainfall typically ranges from 500 to 800 mm. The genus *Acacia* with its many species is characteristic above all others, often associated with species of *Ziziphus, Capparis,* and fleshy *Euphorbia* species. Perennial grasses like *Themeda quadrivivalis, T. triandra, Thelepogon elegans,* and *Sehima nervosum* are common in the understory. The northern forms spread over the semiarid regions of central India, east of the arid Thar Desert. These are shorter than the southern thorn forests, with members of *Acacia, Mimosa,* and *Prosopis* being dominant. The grasses tend to be annuals rather than perennials with *Lasiurus hirsutus, Panicum turgidum, Sporobolus marginatus,* and *Cenchrus* species being common. These systems find their closest structural and compositional analogs in the 'grass and shrub' savannas of arid Africa.

Dry Deciduous Forests

These woodlands, many of which have grassy understories, occur throughout the Indian peninsula with typical mean annual rainfall ranging between 1000 and 1400 mm, with

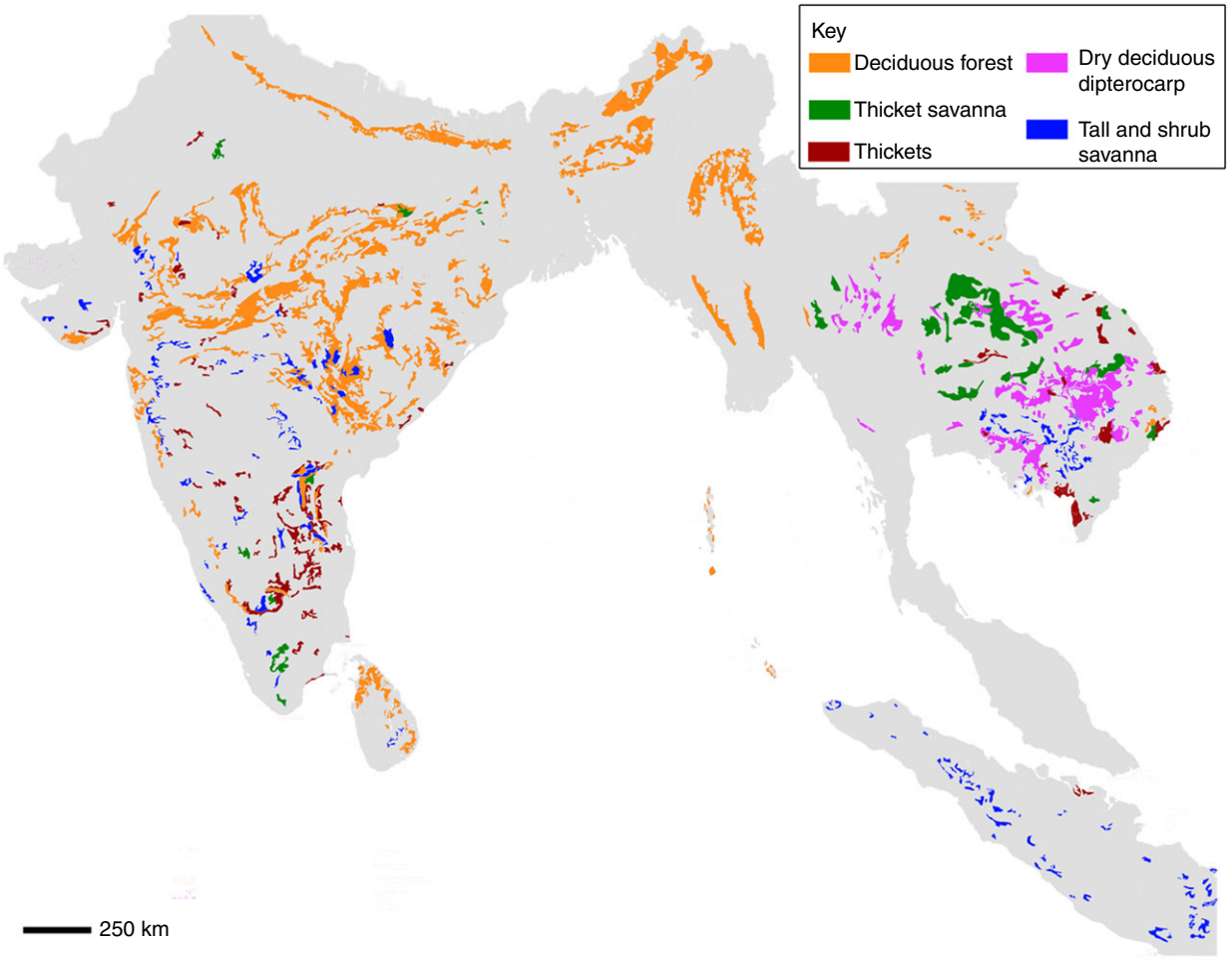


Figure 7 Present day extent of mixed tree–grass vegetation in continental south and Southeast Asia at a scale of 1 : 5 million. The areas in gray represent non-savanna habitats including agricultural lands (Redrawn from Blasco F, Bellan MF, and Aizpuru M (1996) A vegetation map of tropical continental Asia at 1 : 5 million. *Journal of Vegetation Science* 7: 623–634, with permission from John Wiley & Sons; Map credit: Varun Varma).

long dry seasons ranging from 5 to 7 months. While they occur on all well drained soils, they are particularly characteristic on rocky or shallow soils. The dominant tree species are *Tectona grandis*, *Anogeissus latifolia*, and species of *Terminalia*, *Boswellia*, *Diospyros*, and *Sterculea*. Grasses tend to be of medium height with *Heteropogon*, *Themeda*, and *Saccharum* dominating. The bamboo *Dendrocalamus strictus* is also widely present. Both fire and grazing are important in these landscapes, which have historically overlapped across their range with some of the densest human populations of the peninsula. Structurally, they are most similar to the 'tree and shrub' savannas of Africa. However, the factors that differentiate the 'savanna' formation from the 'forest' formation (without grassy understories) of this habitat type are unclear.

Moist Sal Savanna

These are open deciduous formations with heavy grass that occur patchily across the central plains of the Indian peninsula. Mean annual rainfall in these regions may range from 1400 to 1900 mm. The dominant tree species is the Sal or *Shorea robusta*, in association with *Careya arborea*, *Emblica officinalis* and *Wrightea tomentosa*. Predominant grasses in the understorey include *Imperata cylindrica*, *T. triandra*, *Cymbopogon*, and *Apluda* species. Given that these formations occur within a moist forest zone, Champion and Seth (1968) consider that both frequent fires and winter frosts are significant factors driving these formations.

Dipterocarp Savanna Forests

These dry open 'forests' stretch through seasonal mainland Southeast Asia from Vietnam in the east through Cambodia, Laos, Thailand, and Burma, up to the Manipur state in Northeast India. These regions experience mean annual rainfall in the range of 800–2000 mm, with dry seasons ranging from 5 to 7 months. Dominant tree species are *Dipterocarpus intricatus*, *Dipterocarpus tuberculatus*, *Dipterocarpus obtusifolius*, *S. robusta*, *Shorea roxburghi*, and *Shorea siamensis*. Other important tree species include, amongst others, *Dillenia* species, *Pinus merkusii*, *Pterocarpus macrocarpus*, *Terminalia* species, and *Xylia kerrii*. The dominant grasses include genera such as *Arundinella*, *Capellipedium*, *Heteropogon*, *Polytoca*, *Imperata and Themeda*, and pygmy bamboos of the genus *Arundinera*. Also characteristic of these formations are dwarf palms of the genus *Phoenix* and cycad members of the genus *Cycas* (Stott, 1990; Mistry and Stott, 1993). Structurally, these formations find their closest analogs in the miombo woodlands of Africa and the eucalypt woodland savannas of Northern Australia.

Biodiversity in Savannas

While savannas are not generally considered high biodiversity ecosystems like rainforests and coral reefs, these seemingly simple systems are distinctive formations and harbor many unique species and assemblages.

For African savannas, where data are extensive (**Figure 8**), it is apparent that these are floristically diverse systems (Menaut, 1983). The average plant richness of African savannas (calculated for landscape units of 10,000 km^2) is about 1750 species, with a lot of variation across savanna types, and a pattern of increasing plant richness from arid to mesic systems (Menaut, 1983). Of particular note, the miombo savannas of East Africa harbor >3000 species, rivaling the diversity of tropical rainforests. African savannas also support distinctive mammalian herbivore and carnivore assemblages, many of which may be more species-rich than mammal communities of rainforests (Scholes and Walker, 1993; Shorrocks, 2007). The rodent fauna of African savannas are also particularly diverse (Scholes and Walker, 1993), as are the birds (Fry, 1983; Shorrocks, 2007). Over 700 bird species from 57 families, second only to tropical rainforests, have been estimated to occur in just the lowland savannas of tropical Africa, of which eight families are confined to savanna habitats (Fry, 1983).

While biome-scale diversity estimates are not yet readily available for Asian mixed tree–grass systems, an examination of smaller scale studies reveals some trends in floristic diversity. In a pattern similar to that for total floristic diversity in African savannas, woody plant diversity in the Indian subcontinent appears to increase from arid to mesic systems. Species richness of trees and shrubs ranges from 20 to 35 in dry thorn forests, increasing to about 60–80 species in dry deciduous forests (Sukumar *et al.*, 1992; Pandey and Barik, 2006; Joseph *et al.*, 2008). In comparison to woody species, species richness in the herb–grass layers in these systems appears to be much higher. In a study of 40 savanna-grassland communities in southern India, Sankaran (2009) reports 278 species of herbs and grasses in the understory, with most species highly restricted in their distribution. Likewise, Singh *et al.* (2010) report 1011 species for the tropical dry scrub forests–grassland continuum in central India, with the grasses of the *Poaceae* family (112 species) being the most diverse group. For the dipterocarp savanna forests of Thailand, Stott (1990) reports a high diversity of understory species, including many geophytes. In terms of mammalian fauna, while the ungulate and carnivore communities of Asian savanna habitats are not as diverse as those of Africa, many of these (e.g., Greater one-horned rhinoceros, tiger) are unique to the Asian region so that the shrinking extents of these habitats is a serious threat to their conservation. Asian savannas also host a characteristic avifauna, albeit poorer in species richness when compared to Africa (Fry, 1983).

Savannas and the Global Carbon Cycle

Because of their large spatial extent – they cover an area greater than that occupied by boreal or temperate forests (Mouillot and Field, 2005) – savannas play a significant role in influencing local, regional, and global climate and biogeochemical cycles. Tropical savannas and grasslands store ~326 Gt of C, accounting for ~15% of global terrestrial carbon stocks (Grace *et al.*, 2006). Savannas are also remarkably productive ecosystems, with net primary production values ranging from 1 to 12 t C ha^{-1} depending on rainfall and woody cover (Grace *et al.*, 2006). At a global level, net primary production in savannas averages ~7.2 t C ha^{-1} year^{-1}, paralleling that of temperate forests (7.7 t C ha^{-1} year^{-1}) and third only to tropical rainforests (12.5 t C ha^{-1} year^{-1}). They sequester ~0.39 Gt C year^{-1}, contributing ~15% to the annual global carbon sink.

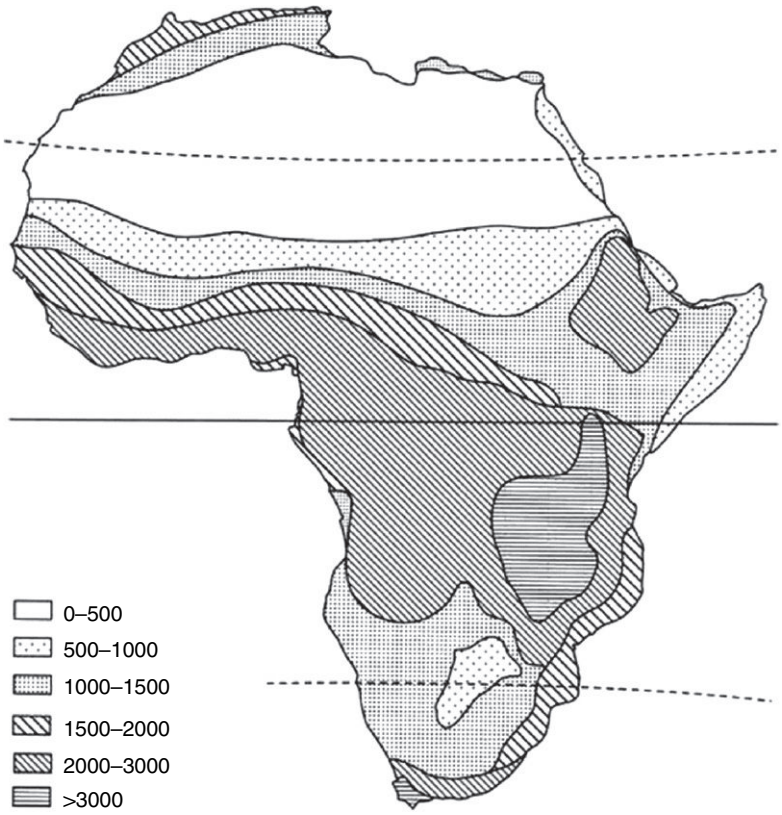


Figure 8 Floristic richness across the African continent, collated as number of species per 10,000 km^2. The solid line is the equator and the dotted lines represent the tropic of Cancer and tropic of Capricorn (Reproduced from Menaut J-C (1983) The vegetation of African savannas. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 109–149. Amsterdam: Elsevier Scientific Publishing Company).

Savannas are also the most fire prone biome in the world, accounting for nearly 85% of the global land area burnt annually (Mouillot and Field, 2005). It is estimated that nearly 311 and 35 million hectares of savanna vegetation burn annually in Africa and South Asia ($\sim$55% and 6% of global burnt area, respectively), with an additional 38 million hectares of Southeast Asian savannas burnt annually (Mouillot and Field, 2005). Collectively, savanna fires account for $\sim$1.67 Gt C year^{-1}, or half of the global carbon emissions from biomass burning (Mouillot *et al.*, 2006). Savanna fires are also important sources of NO_x ($NO + NO_2$) to the atmosphere, accounting for over half of that released by wildfires globally (Beerling and Osborne, 2006).

Clearly, savannas constitute an important component of the global carbon cycle. At the same time, savanna ecosystems worldwide are under intense anthropogenic pressure. Rate of loss of savanna habitats, though not well established, has been estimated to be as much as twice that of rainforests, potentially contributing a flux of carbon to the atmosphere matching that resulting from the deforestation of rainforests (Grace *et al.*, 2006). Protection of savannas from burning and grazing has been proposed as a mechanism to enhance terrestrial carbon sequestration in these biomes and worldwide (Grace *et al.*, 2006). While such measures are undoubtedly required in some regions, it is also important to recognize that these cannot be adopted as blanket measures in savannas worldwide. Fire and grazing are both integral components of savanna ecosystems, defining not only the structure of these ecosystems but also serving to maintain the unique diversity that characterizes them (Ratnam *et al.* 2011).

The Future of the Savanna Biome: Savannas and Climate Change

Savanna structure and dynamics are the consequence of a suite of interacting factors, and changing environmental drivers are likely to have significant impacts on savanna vegetation both directly, by influencing resource availability, and indirectly through feedback effects arising from changes in top-down controls. In fact, savanna ecosystems are believed to be amongst the most sensitive biomes to future climate change (Sala *et al.*, 2000; Bond *et al.*, 2003).

The strong dependence of savanna vegetation structure on water availability suggests that future changes in rainfall, both in the total amount and seasonality, are likely to have significant impacts on woody vegetation dynamics in savannas. Climate models predict both increases and decreases in future total rainfall in many savanna regions of the globe (IPCC, 2007). In addition, many regions are also likely to witness changes in rainfall distribution across the season, with the rain falling in fewer but more intense events with longer

intervening dry periods (IPCC, 2007). Different savanna regions are therefore likely to show divergent responses depending on the exact nature of the change in local precipitation regimes. Savannas are likely to shift to more wooded states in areas where precipitation increases, and toward more open states where precipitation declines, or where the rain arrives in fewer but more intense events (Sankaran *et al.*, 2005, 2008; Good and Caylor, 2011).

Increased atmospheric CO_2, however, is likely to favor woody proliferation (Polley *et al.*, 1997; Bond and Midgley, 2000). CO_2 fertilization stimulates C_3 photosynthesis more than C_4 photosynthesis (Sage and Kubien, 2003), and is thus likely to favor C_3 trees and shrubs and give them a competitive advantage over C_4 grasses. Besides such direct effects on growth, decreased transpiration rates under elevated CO_2 can increase subsoil moisture, additionally favoring the woody component (Polley *et al.*, 1997). Further, increased growth of tree and shrub seedlings and saplings under elevated CO_2 can also hasten their escape from the grass flame zone, and increase the resprouting ability of savanna species by enhancing carbohydrate reserves (Hoffmann *et al.*, 2000), thus reducing the strength of fire-imposed bottlenecks (Bond and Midgley, 2000). In fact, large scale increases in woody cover related to regional rather than local drivers, potentially increased CO_2, have already been widely reported (Wigley *et al.*, 2010 and references therein).

In addition to changes in atmospheric CO_2 concentrations, anthropogenic activity has also resulted in an estimated 13-fold increase in global N deposition (Galloway *et al.*, 2008) and a doubling of P deposition (Filippelli, 2008). Such widespread fertilization is likely to have significant impacts on savanna vegetation structure. While increased N-deposition can potentially cause savannas to shift to more open, grassy states (Sankaran *et al.*, 2008; van der Waal *et al.*, 2009, 2011), the impact of enhanced P-deposition on savannas remains unclear.

A substantial body of literature has developed over the years on the individual effects of different climate change drivers. These different drivers are likely to elicit divergent responses in savanna vegetation. Increased rainfall and elevated CO_2 can cause shifts toward more wooded states, while lowered rainfall amounts, more intense rainfall events, longer intraseasonal droughts and N-deposition should cause shifts toward more open states. However, it is important to recognize that different drivers do not act individually, but in concert, and are changing in ways that are spatially variable across the globe (IPCC, 2007). Predicting savanna responses to future global change will therefore require a more comprehensive understanding of the interactive effects of different global change drivers. Besides the direct effects on primary producers arising from altered resource availability, different global change drivers can also have effects that cascade through to higher trophic levels and feedback to influence vegetation dynamics. For example, changes in forage quality resulting from N-deposition or elevated CO_2, or changes in surface water availability as a result of altered rainfall regimes, can influence both the amounts and spatial pattern of herbivory and fires, with implications for savanna vegetation structure and dynamics at large scales. Even while our understanding of savanna ecosystems has substantially increased in the last two decades, more research that will enable us to better predict how the diverse savannas of the world will respond to global change is urgently needed.

See also: Africa, Ecosystems of. Asia, Ecosystems of. Australia, Biodiversity of Ecosystems. C_4 Plants. Climate, Effects of. Ecosystems of South America. Fires, Ecological Effects of. Grazing, Effects of. Megaherbivores

References

Archer S (1995) Harry Stobbs Memorial Lecture, 1993: Herbivore mediation of grass–woody plant interactions. *Tropical grasslands* 29: 218–235.

Archibald S, Roy DP, Van Wilgen BW, and Scholes RJ (2009) What limits fire? An examination of drivers of burnt area in Southern Africa. *Global Change Biology* 15: 613–630.

Asner GP, Levick SR, Kennedy-Bowdoin T, *et al.* (2009) Large-scale impacts of herbivores on the structural diversity of African savannas. *Proceedings of the National Academy of Sciences USA* 106: 4947–4952.

Augustine DJ and McNaughton SJ (1998) Ungulate effects on the functional species composition of plant communities: Herbivore selectivity and plant tolerance. *Journal of Wildlife Management* 62: 1165–1183.

Augustine DJ and McNaughton SJ (2004) Regulation of shrub dynamics by native browsing ungulates on East African rangeland. *Journal of Applied Ecology* 41: 45–58.

Barrington AHM (1931) *Forest Soil and Vegetation in the Hlaing Forest Circle, Burma.* Burma Forest Bulletin, No 25, Ecology Series No. 1. Rangoon: Government Printing and Stationery.

Beerling DJ and Osborne CP (2006) The origin of the savanna biome. *Global Change Biology* 12: 2023–2031.

Blasco F (1983) The transition from open forest to savanna in continental south-east Asia. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 167–181. Amsterdam: Elsevier Scientific Publishing Company.

Blasco F, Bellan MF, and Aizpuru M (1996) A vegetation map of tropical continental Asia at 1 : 5 million. *Journal of Vegetation Science* 7: 623–634.

Bond WJ (2008) What limits trees in C_4 grasslands and savannas? *Annual Review of Ecology and Systematics* 39: 641–659.

Bond WJ (2010) Do nutrient-poor soils inhibit development of forests? A nutrient stock analysis. *Plant and Soil* 334: 47–60.

Bond WJ and Keeley JE (2005) Fire as a global 'herbivore': The ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution* 20: 387–394.

Bond WJ, Midgley GF, and Woodward FI (2003) What controls South African vegetation – climate or fire? *South African Journal of Botany* 69: 79–91.

Bond WJ and Midgley JJ (2000) A proposed CO_2-controlled mechanism of woody plant invasion in grasslands and savannas. *Global Change Biology* 6: 865–869.

Bond WJ and Parr CL (2010) Beyond the forest edge: Ecology, diversity and conservation of the grassy biomes. *Biological Conservation* 143: 2395–2404.

Bond WJ, Woodward FI, and Midgley GF (2005) The global distribution of ecosystems in a world without fire. *New Phytologist* 165: 525–538.

Boulbet J (1982) Evolution des paysages vegetaux en Thailande du nord-est, *Publication de l' Ecole Francais d'Extreme Orient CXXXVL*. Paris: Ecole Francais d'Extreme Orient.

Bourlière F and Hadley M (1970) The ecology of tropical savannas. *Annual Review of Ecology and Systematics* 1: 125–152.

Bourlière F and Hadley M (1983) Present-day savannas: An overview. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 1–17. Amsterdam: Elsevier Scientific Publishing Company.

Brown JR and Archer SR (1987) Woody plant seed dispersal and gap formation in a North American subtropical savanna fauna woodland: The role of domestic herbivores. *Vegetation* 73: 73–80.

Bucini G and Hanan NP (2007) A continental-scale analysis of tree cover in African savannas. *Global Ecology and Biogeography* 16: 593–605.

Cerling TE, Wynn JG, Andanje SA, *et al.* (2011) Woody cover and hominin environments in the past 6 million years. *Nature* 476: 51–56.

Champion HG and Seth SK (1968) *A Revised Survey of the Forest Types of India.* Nasik: Government of India Press.

Cramer MD, van Cauter A, and Bond WJ (2010) Growth of N_2-fixing African savanna *Acacia* species is constrained by below-ground competition with grass. *Journal of Ecology* 98: 156–167.
Davis MA, Wrage KJ, Reich PB, *et al.* (1999) Survival, growth, and photosynthesis of tree seedlings competing with herbaceous vegetation along a water–light–nitrogen gradient. *Plant Ecology* 145: 341–350.
Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, and C_4 Grasses Consortium (2010) The origin of C_4 grasslands: Integrating evolutionary and ecosystem science. *Science* 328: 587–591.
Eldridge DJ, Bowker MA, Maestre FT, *et al.* (2011) Impacts of shrub encroachment on ecosystem structure and functioning: Towards a global synthesis. *Ecology Letters* 14: 709–722.
Filippelli GM (2008) The global phosphorous cycle: Past, present and future. *Elements* 4: 89–95.
Fritz H and Duncan P (1994) On the carrying capacity for large ungulates of African savanna ecosystems. *Proceedings of the Royal Society of London B* 256: 77–82.
Frost PG, Medina E, Menaut JC, Solbrig O, Swift M, and Walker BH (eds.) (1986) *Response of Savannas to Stress and Disturbance. Biology International Special.* Paris: IUBS.
Fry CH (1983) Birds in savanna ecosystems. In: Bourlière F (ed.) *Tropical savannas. Ecosystems of the World 13*, pp. 337–357. Amsterdam: Elsevier Scientific Publishing Company.
Gadgil M and Meher-Homji VM (1985) Land use and productive potential of Indian savannas. In: Tothill JC and Mott JJ (eds.) *Ecology and Management of the World's Savannas*, pp. 107–113. Canberra: The Australian Academy of Sciences.
Galloway JN, Townsend AR, Erisman JW, *et al.* (2008) Transformation of the nitrogen cycle: Recent trends, questions and potential solutions. *Science* 320: 889–892.
Gignoux J, Clobert J, and Menaut JC (1997) Alternate fire resistance strategies in savanna trees. *Oecologia* 110: 576–583.
Goheen JR, Palmer TM, Keesing F, Riginos C, and Young TP (2010) Large herbivores facilitate savanna tree establishment via diverse and indirect pathways. *Journal of Animal Ecology* 79: 372–382.
Good SP and Caylor KK (2011) Climatological determinants of woody cover in Africa. *Proceedings of the National Academy of Sciences* 108: 4902–4907.
Grace J, Jose SJ, Meir P, Miranda HS, and Montes RA (2006) Productivity and carbon fluxes of tropical savannas. *Journal of Biogeography* 33: 387–400.
Greve M, Lykke AM, Blach-Overgaard A, and Svenning J (2011) Environmental and anthropogenic determinants of vegetation distribution across Africa. *Global Ecology and Biogeography* 20: 661–674.
Guldemond R and van Aarde R (2008) A meta-analysis of the impact of African elephants on savanna vegetation. *The Journal of Wildlife Management* 72: 892–899.
Hegde V, Chandran MDS, and Gadgil M (1998) Variation in bark thickness in a tropical forest community of Western Ghats in India. *Functional Ecology* 12: 313–318.
Higgins SI, Bond WJ, and Trollope WSW (2000) Fire, resprouting and variability: A recipe for grass–tree coexistence in savanna. *Journal of Ecology* 88: 213–229.
Hirota M, Holmgren M, Van Nes EH, and Scheffer M (2011) Global resilience of tropical forest and savanna to critical transitions. *Science* 334: 232–235.
Hoffmann WA, Adasme R, Haridasan M, *et al.* (2009) Tree topkill, not mortality, governs the dynamics of savanna-forest boundaries under frequent fire in central Brazil. *Ecology* 90: 1326–1337.
Hoffmann WA, Bazzaz GA, Chatterton NJ, Harrison PA, and Jackson RB (2000) Elevated CO_2 enhances resprouting of a tropical savanna tree. *Oecologia* 123: 312–317.
Hoffmann WA and Solbrig OT (2003) The role of topkill in the differential response of savanna woody species to fire. *Forest Ecology and Management* 180: 273–286.
House J, Archer S, Breshears DD, Scholes RJ and NCEAS Tree–Grass Interaction Participants (2003) Conundrums in mixed woody–herbaceous plant systems. *Journal of Biogeography* 30: 1763–1777.
Huntley BJ (1982) Southern African savannas. In: Huntley BJ and Walker BH (eds.) *Ecology of Tropical Savannas*, pp. 111–119. Berlin: Springer-Verlag.
Huntley BJ and Walker BH (eds.) (1982) *Ecology of Tropical Savannas.* Berlin: Springer-Verlag.
Intergovernmental Panel on Climate Change (IPCC) (2007) *Climate Change 2007: The Physical Science Basis. Summary for Policymakers.* New York: Cambridge University Press.
Jeltsch F, Milton SJ, Dean WRJ, and van Rooyen N (1996) Tree spacing and coexistence in semiarid savannas. *Journal of Ecology* 84: 583–595.
Jeltsch F, Milton SJ, Dean WRJ, van Rooyen N, and Moloney KA (1998) Modelling the impact of small-scale heterogeneities on tree–grass coexistence in semi-arid savannas. *Journal of Ecology* 86: 780–793.
Jeltsch F, Weber GE, and Grimm V (2000) Ecological buffering mechanisms in savannas: A unifying theory of long-term tree–grass coexistence. *Plant Ecology* 161: 161–171.
Joseph S, Sudhakar Reddy C, Pattanaik C, and Sudhakar S (2008) Distribution of plant communities along climatic and topographic gradients in Mudumalai Wildlife Sanctuary (southern India). *Biological Letters* 45: 29–41.
Kraaij T and Ward D (2006) Effects of rain, nitrogen, fire and grazing on tree recruitment and early survival in bush-encroached savanna, South Africa. *Plant Ecology* 186: 235–246.
Kulmatiski A, Beard KH, Verweij RJT, and February EC (2010) A depth-controlled tracer technique measures vertical, horizontal and temporal patterns of water use by trees and grasses in a subtropical savanna. *New Phytologist* 188: 199–209.
Lamotte M and Bourlière F (1983) Energy flow and nutrient cycling in tropical savannas. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 583–603. Amsterdam: Elsevier Scientific Publishing Company.
van Langevelde F, van de Vijver CADM, Kumar L, *et al.* (2003) Effects of fire and herbivory on the stability of savanna ecosystems. *Ecology* 84: 337–350.
Lehmann CER, Archibald SA, Hoffmann WA, and Bond WJ (2011) Deciphering the distribution of the savanna biome. *New Phytologist* 191: 197–209.
Lenzi-grillini CR, Viskanic P, and Mapesa M (1996) Effects of 20 years of grazing exclusion in an area of the Queen Elizabeth National Park, Uganda. *African Journal of Ecology* 34: 333–341.
Lloyd J, Bird MI, Vellen L, *et al.* (2008) Contributions of woody and herbaceous vegetation to tropical savanna ecosystem productivity: A quasi-global estimate. *Tree Physiology* 28: 451–468.
Menaut J-C (1983) The vegetation of African savannas. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 109–149. Amsterdam: Elsevier Scientific Publishing Company.
Midgley JJ, Lawes MJ, and Chamaillé-Jammes S (2010) Savanna woody plant dynamics: The role of fire and herbivory, separately and synergistically. *Australian Journal of Botany* 58: 1–11.
Mishra R (1983) Indian savannas. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 151–166. Amsterdam: Elsevier Scientific Publishing Company.
Mistry J and Stott P (1993) The savanna forests of Manipur State, India: An historical overview. *Global Ecology and Biogeography Letters* 3: 10–17.
Mistry J (2000) *World Savannas: Ecology and Human Use.* Harlow: Prentice Hall.
Mouillot F and Field CB (2005) Fire history and the global carbon budget: A 1° x 1° fire history reconstruction for the 20th century. *Global Change Biology* 11: 398–420.
Mouillot F, Narasimha A, Balkanski Y, Lamarque J, and Field CB (2006) Global carbon emissions from biomass burning in the 20th century. *Geophysical Research Letters* 33: L01801.http://dx.doi.org/10.1029/2005GL024707.
Nix HA (1983) Climate of tropical savannas. In: Bourlière F (ed.) *Tropical Savannas. Ecosystems of the World 13*, pp. 37–61. Amsterdam: Elsevier Scientific Publishing Company.
Osborne CP and Beerling DJ (2006) Nature's green revolution: The remarkable evolutionary rise of C_4 plants. *Philosophical Transactions: Biological Sciences* 361: 173–194.
Osborne CP (2008) Atmosphere, ecology and evolution: What drove the Miocene expansion of C_4 grasslands? *Journal of Ecology* 96: 35–45.
Osborne CP and Freckleton RP (2009) Ecological selection pressures for C_4 photosynthesis in the grasses. *Proceedings of The Royal Society, Series B* 276: 1753–1760.
Pandey HN, and Barik SK (eds.)(2006) *Ecology, Diversity, and Conservation of Plants and Ecosystems in India.* New Delhi: Regency Publications.
Pemadasa MA (1991) Tropical grasslands of Sri Lanka and India. In: Werner PA (ed.) *Savanna Ecology and Management*, pp. 51–56. London: Blackwell Scientific Publications.
Polley HW, Mayeux HS, Johnson HB, and Tischler CR (1997) Viewpoint: Atmospheric CO_2, soil water, and shrub/grass ratios on rangelands. *Journal of Range Management* 50: 278–284.
Prins HHT and van der Juegd HP (1993) Herbivore population crashes and woodland structure in East Africa. *Journal of Ecology* 81: 305–314.
Puri GS (1960) *Indian Forest Ecology: A comprehensive survey of vegetation and its environment in the Indian subcontinent* Vol. 1. New Delhi: Oxford Book and Stationary Company.
Ramankutty N and Foley JA (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13: 997–1027.
Ratnam J, Bond WJ, Fensham RJ, *et al.* (2011) When is a 'forest' a savanna, and why does it matter? *Global Ecology & Biogeography* 20: 653–660.

Roques KG, O'Connor TG, and Watkinson AR (2001) Dynamics of shrub encroachment in an African savanna: Relative influences of fire, herbivory, rainfall and density dependence. *Journal of Applied Ecology* 38: 268–280.
Sage RF and Kubien DS (2003) *Quo vadis* C_4? An ecophysiological perspective on global change and the future of C_4 plants. *Photosynthesis Research* 77: 209–225.
Sage RF (2004) The evolution of C_4 photosynthesis. *New Phytologist* 161: 341–370.
Sala OE, Chapin III FS, Armesto JJ, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.
Sankaran M, Ratnam J, and Hanan NP (2004) Tree–grass coexistence in savannas revisited – insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology Letters* 7: 480–490.
Sankaran M, Hanan NP, Scholes RJ, *et al.* (2005) Determinants of woody cover in African Savannas. *Nature* 438: 846–849.
Sankaran M, Ratnam J, and Hanan N (2008) Woody cover in African savannas: The role of resources, fire and herbivory. *Global Ecology and Biogeography* 17: 236–245.
Sankaran M (2009) Diversity patterns in savanna grassland communities: Implications for conservation strategies in a biodiversity hotspot. *Biodiversity and Conservation* 18: 1099–1115.
Sankaran M and Anderson MT (2009) Management and restoration in African savannas: Interactions and feedbacks. In: Hobbs RJ and Suding KN (eds.) *New Models for Ecosystem Dynamics and Restoration*, pp. 136–155. Washington: Island Press.
Sarmiento G (1984) *The Ecology of Neotropical Savannas*. Cambridge: Harvard University Press.
Schimper AFW (1903) *Plant Geography Upon A Physiological Basis*. Oxford: Clarendon Press.
Scholes RJ and Archer SR (1997) Tree–grass interactions in savannas. *Annual Review of Ecology and Systematics* 28: 517–544.
Scholes RJ and Walker BH (1993) *An African Savanna: Synthesis of the Nylsvley Study*. Cambridge, UK: Cambridge University Press.
Schutz AEN, Bond WJ, and Cramer MD (2009) Juggling carbon: Allocation patterns of a dominant tree in a fire-prone savanna. *Oecologia* 160: 235–246.
Shorrocks B (2007) *The Biology of African Savannas*. Oxford: Oxford University Press.
Singh KP, Shukla A, and Singh JS (2010) Floristic diversity and taxonomic profile of Achanakmar–Amarkantak Biosphere Reserve, Central India. *Journal of the Bombay Natural History Society* 107: 135–145.
Staver AC, Archibald S, and Levin SA (2011) The global extent and determinants of savanna and forest as alternate biome states. *Science* 334: 230–232.
Staver AC, Bond WJ, Stock WD, *et al.* (2009) Browsing and fire interact to suppress tree density in an African savanna. *Ecological Applications* 19: 1909–1919.
Stott P (1984) The savanna forests of mainland southeast Asia: An ecological survey. *Progress in Physical Geography* 8: 315–335.
Stott P (1988) The forest as Phoenix: Towards a biogeography of fire in mainland south-east Asia. *Geographical Journal* 154: 337–350.
Stott P (1990) Stability and stress in the savanna forests of mainland south-east Asia. *Journal of Biogeography* 17: 373–383.
Stott P (1991) Recent trends in the ecology and management of the world's savanna formations. *Progress in Physical Geography* 15: 18–28.
Sukumar R, Dattaraja HS, Suresh HS, *et al.* (1992) Long term monitoring of vegetation in a tropical deciduous forest in Mudumalai, southern India. *Current Science* 62: 608–616.
van der Waal C, de Kroon H, de Boer WF, *et al.* (2009) Water and nutrients alter herbaceous competitive effects on tree seedlings in a semi-arid savanna. *Journal of Ecology* 97: 430–439.
van der Waal C, Kool A, Meijer SS, *et al.* (2011) Large herbivores may alter vegetation structure of semi-arid savannas through soil nutrient mediation. *Oecologia* 165: 1095–1107.
van Wijk MT and Rodriguez-Iturbe I (2002) Tree–grass competition in space and time: Insights from a simple cellular automata model based on ecohydrological dynamics. *Water Resources Research* 38: 18.11–18.15.
Walker BH, Ludwig D, Holling CS, and Peterman RM (1981) Stability of semi-arid savanna grazing systems. *Journal of Ecology* 69: 473–498.
Walker BH and Noy-Meir I (1982) Aspects of stability and resilience of savanna ecosystems. In: Walker BJ and Huntley BH (eds.) *Ecology of Tropical Savannas*, pp. 556–590. Berlin: Springer-Verlag.
Walter H (1971) *Ecology of Tropical and Subtropical Vegetation*. Edinburgh: Oliver and Boyd.
White F (1983) *The Vegetation of Africa*. Paris: UNESCO.
Whittaker RH (1975) *Communities and Ecosystems*, 2nd edn. New York: McMillan.
Wigley BJ, Cramer MD, and Bond WJ (2009) Sapling survival in a frequently burnt savanna: Mobilization of carbon reserves in *Acacia karroo*. *Plant Ecology* 203: 1–11.
Wigley BJ, Bond WJ, and Hoffman MT (2010) Thicket expansion in a South African savanna under divergent land use: Local vs. global drivers? *Global Change Biology* 16: 964–976.
Williams RJ, Duff GA, Bowman DMJS, and Cook GD (1996) Variation in the composition and structure of tropical savannas as a function of rainfall and soil texture along a large-scale climatic gradient in the Northern Territory, Australia. *Journal of Biogeography* 23: 747–756.
Wilson JB and Agnew ADQ (1992) Positive-feedback switches in plant communities. *Advances in Ecological Research* 23: 263–335.
Yadava PS (1990) Savannas of north-east India. *Journal of Biogeography* 17: 385–394.

Agricultural Invasions

David Pimentel, Cornell University, Ithaca, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Competition Use or defense of a resource, such as a desirable growing site for a plant and a food source or shelter for an animal, by an individual that reduces the availability of the resource for another individual.

Competitive exclusion Extirpation or extinction of one species by another in a given area through competition for resources.

Exotic (nonindigenous) species A species that occurs in an area outside of its historically known range and that has been introduced into a new habitat or ecosystem, either intentionally or accidentally.

Native (indigenous) species A species that occurs naturally in a given area.

Nonindigenous crops Crop plants that did not originate in the range where they are cultivated. Most or 99% of all crop species grown in the US are introduced or nonindigenous species and have been introduced intentionally.

Introduction

Thousands of species have been intentionally introduced to new regions for explicit agricultural purposes, such as the growing of crops and pest control. Yet many species are also accidentally introduced, and these biological "invaders" end up doing harm to the areas they invade. They can cause major economic losses in agriculture, forestry, and other segments of the world economy if crops or other products are destroyed. One recent study reported that more than 50,000 species that have been introduced into the US since Columbus' time have resulted in approximately $238 billion in damages annually (Pimentel *et al.*, 2000; Pimentel, 2011). In addition, introduced species often have substantial negative impacts on global ecological integrity; for example, the presence of an alien invader can result in the extinction of many native species.

It is difficult and complicated to estimate the full extent of the ecological damages to world agriculture caused by exotic species and the number of species extinctions they have caused, because only approximately 20% of the estimated 10–15 million species in the world have been described and cataloged. Nonetheless, about 400 of the 958 species on the Threatened or Endangered Species List in the US are considered at risk primarily because of competition and predation by nonindigenous species. Other species are continually threatened by hybridization with alien species or ecosystem changes caused by biological invaders. Estimation of the economic impacts of nonindigenous pests affecting world agriculture is also difficult; however, sufficient data are available to quantify some aspects of the economic impacts on agriculture. This article assesses the magnitude of environmental impacts and some of the economic costs associated with the diverse nonindigenous species that have invaded world agriculture.

Exotic Species Introductions in the United States

Most plant and vertebrate animal introductions in the US agriculture have been intentional, whereas most invertebrate animal and microbe introductions have been accidental. In the recent past, the rate of and risk associated with exotic species introductions have increased enormously because the human population is growing so rapidly; this ever-increasing number of humans has substantially altered, and continues to alter, the global environment (Pimentel *et al.*, 2001; Pimentel, 2011). In addition, large numbers of people are traveling faster and farther and more goods and materials are being traded among nations all over the world (Bryan, 1996), so the opportunity for even more species introductions is on the rise.

Nonindigenous Plants

Introduced Crop Plants

Most alien plants now established in the US were introduced for food, fiber, or ornamental purposes. An estimated 5000 plant species have escaped these realms and now exist in US natural ecosystems. This number is substantial when compared with a total of about 18,000 species of native US plants.

Ornamental plants are the largest group of nonindigenous plants that have become established in nature. These plants include those found in cultivated gardens, such as purple loosestrife (*Lythrum salicaria*). Of the approximately 25,000 cultivated plant species, mostly ornamentals brought into Florida, more than 900 have become established in surrounding natural ecosystems (Simberloff *et al.*, 1997). More than 3000 plant species have been introduced into California. New York state also has a large percentage of introduced plant species: 1082 exotics versus 1940 native species.

In crop systems, including forage crops, an estimated 500 introduced plant species have become weed pests. Most of these were accidentally introduced with crop seeds, from ship ballast, or from various imported dead and live plant materials. Some of these accidentally introduced weeds, like yellow rocket (*Barbarea vulgaris*) and Canada thistle (*Cirsium arvense*), have become major weed pests in US gardens and commercial agriculture.

Most of the 5000 nonindigenous plants established in US natural ecosystems have displaced some native plant species. Alien grasses introduced as forage species have diminished the

health of natural ecosystem functions as a result of nutrient losses, changes in microclimates, altered vegetation succession, and increased incidence of fires. Nonindigenous weeds are spreading and invading approximately 700,000 ha of US wildlife habitat each year (Babbitt, 1998).

One of these weed pests is the European purple loosestrife, which was introduced in the early nineteenth century as an ornamental garden plant. It has been spreading at a rate of 115,000 ha per year and is changing the basic structure of most of the wetlands that it has invaded. The dominant, monotypic stands of purple loosestrife have reduced the biomass of 44 native plants and endangered the wildlife that depends on these native plants.

Economic Costs of Introduced Plants

In US agriculture, weeds cause a reduction of 12% in crop yields; in economic terms, that represents an approximate $33 billion loss in crop production annually, based on the potential crop value of all US crops of more than $267 billion per year. Using the estimate that about 73% of the weeds are nonindigenous, it follows that around $24 billion of these crop losses are due to introduced weeds. In addition, approximately $4 billion in herbicides are applied to US crops, of which about $3 billion can be attributed to nonindigenous weed control.

US pastures provide about $10 billion in forage crops annually, with total losses due to weeds estimated at nearly $2 billion. Since about 45% of these weed species are nonindigenous, the author estimates that forage losses due to these weeds are close to $1 billion each year.

Some introduced weeds are actually toxic to cattle and wild ungulates. For example, leafy spurge (*Euphorbia esula*) contains a toxic substance that, when eaten by cattle, acts as an irritant, emetic, and purgative and can cause dysentery and sometimes death. Some animals, after unpleasant experiences, learn to avoid eating the plant. If the vegetation on grazing land is composed of just 10–20% leafy spurge, the cattle will not graze there and the value of the land as forage is lost. Total losses resulting from the prevalence of this weed on western rangelands account for nearly $110 million per year due to toxic effects and reduced grazing by the cattle (this cost is included in the $1 billion in losses mentioned earlier) (OTA, 1993).

Several nonindigenous thistles also replace desirable native plant species in pastures, rangelands, and forests and similarly reduce cattle grazing. According to the US Secretary of the Interior Babbitt (1998), ranchers spend about $5 billion each year to control invasive nonindigenous weeds in pastures and rangelands; despite those expensive efforts, though, these weeds continue to spread.

Lawn, garden, and golf course management costs about $36 billion per year. A significant proportion of these management costs is related to pest control. In addition to the high cost of this commercial pest control, it is estimated that about $1.3 billion (included in the $36 billion) is spent on outdoor residential weed, insect, and disease pest control each year. Given the high percentages of nonindigenous species, the author estimates that about $500 million is spent on residential nonindigenous weed control, and an additional $1 billion is invested in nonindigenous weed control on golf courses. Therefore, the total economic cost of introduced weeds on US cropland, pastureland, and other uses is about $35 billion annually (**Table 1**).

Nonindigenous Vertebrates

Mammalian Introductions

Introduced mammal species include dogs, cats, horses, burros, cattle, sheep, pigs, goats, monkeys, and deer (Layne, 1997). Several of these animals escaped or were released into the wild, and many have become pests by either preying on native animals, grazing on vegetation, or intensifying soil erosion.

For example, goats (*Copra hircus*) introduced on San Clemente Island, California, are responsible for the extinction of eight endemic plant species and the endangerment of eight other native plant species on the island (Kurdila, 1995). Horses (*Equus caballus*) and burros (*Equus asinus*) released in the western US have reached populations of up to 50,000 animals. These animals graze heavily on native vegetation, allowing nonindigenous annuals to displace native perennials. Furthermore, burros inhabiting the northwestern US compete for the primary food sources of native bighorn sheep and

Table 1 Economic losses to biological pest invaders in agriculture in the US, UK, Australia, South Africa, India, and Brazil (in billions of dollars)

Pest invader	*US*[a]	*UK*	*Australia*	*South Africa*	*India*	*Brazil*	*Total*
Plant weeds	$35.000	$1.200	$2.400	$2.500	$23.724	$12.300	$77.124
Mammals							
Rats	19.000	4.400	1.200	2.700	25.000	3.200	$55.500
Other	1.310	1.200	5.580	—	—	—	$8.090
Birds	2.100	0.270	—	—	—	—	$2.370
Arthropods	15.000	0.840	1.000	1.700	10.200	4.600	$33.340
Microbes							
Plant pathogens	23.500	1.700	2.700	2.900	21.400	18.300	$70.500
Livestock diseases	14.000	—	0.021	1.000	—	—	$15.021
Total	$109.910	$9.610	$12.901	$10.800	$80.324	$38.400	$261.945

[a]Includes pest control costs for lawns, gardens, and golf courses. See text for details.

See Table 1 of Chapter 17 of *Invasive Species* book for livestock disease figure – The author has added the US column and bottom row of total – also *see* Conclusions. For the final costs for US biological pest invaders – *see* Microbe Introduction in Agriculture.

seed-eating birds, thereby reducing the abundance of these native animals (Kurdila, 1995).

Overall, large populations of feral horses and burros are costly in terms of lost forage for livestock and wildlife. Based on a minimum of $10 per animal-unit-month (AUM) per hectare, and assuming 10 ha of pasture per animal per year, these large animals exact damage of at least $5 million per year. This cost does not include the extensive environmental damage to ecosystems, which can be up to 10 times this amount when increased soil erosion and the reduction of native plants and associated native animals are considered.

Pigs (*Sus scrofa*), which are native to Eurasia and North Africa, have been introduced into some US parks (like the California coastal prairie) for hunting, and have substantially changed the vegetation in those places. Feral pigs are also a serious problem for US agriculture. In Hawaii, more than 80% of the soil is bare in regions inhabited by pigs (Kurdila, 1995). This disturbance allows annual plants to invade the overturned soil and intensifies soil erosion. Pig control per park in Hawaii, with about 1500 pigs per park, costs about $150,000 each year (based on $100 spent per pig for control). Assuming that the three parks in Hawaii have similar pig control problems, the total is $450,000 per year just in Hawaii (Zuniga, personal communication).

Pigs also have become a serious problem in Florida, where the feral pig population has risen to more than 500,000. In a few cases these invaders are beneficial – for example, pigs serve as a food source for the endangered Florida panther (Layne, 1997). For the most part, though, in Florida and elsewhere, pigs damage hay, grain, peanut, soybean, cotton, and various vegetable crops. Pigs also damage fences, livestock watering devices, ponds, and young livestock. In addition, they transmit and are reservoirs for serious diseases that affect both humans and livestock, like brucellosis, pseudobrucellosis, and trichinosis.

Nationwide, there are an estimated 4 million feral pigs (Pimentel *et al.*, 2000). Control costs for feral pigs range from $36 per pig for poison control to $140 per pig for hunting with dogs. Estimated damages are at least $300 per pig per year. Assuming that 4 million feral pigs inhabit the US, the yearly cost is about $800 million per year. This is a conservative estimate because pigs cause significant environmental damage that is not easily translated into dollar values.

In addition to large animals like horses and pigs, many small mammals have been introduced both accidentally and intentionally into the US. Some examples of these species are the European (black or tree) rat (*Rattus rattus*), Asiatic (Norway or brown) rat (*Rattus norvegicus*), house mouse (*Mus musculus*), nutria (*Myocaster coypus*), European rabbit (*Oryctolagus cuniculus*), cat (*Felis cattus*), and dog (*Cards familiaris*) (Layne, 1997).

Some rodents have become serious pests on farms, in industries, and in homes (Layne, 1997). On farms, rats and mice are particularly abundant and destructive. On poultry farms there is approximately one rat for every five chickens (Pimentel, 2011). Assuming this ratio, the total rat population on US poultry farms may easily number more than 1.4 billion. Even with a decline in the rat population since these observations were made, the author estimates that the number of rats on poultry and other farms is still approximately 1 billion. Given an estimated 1 rat per person, there are an estimated 250 million rats in US urban and suburban areas. All told, there is a total of around 1.25 billion rats in the US.

Various studies report that an individual rat consumes or destroys from 210 to 740 kg of grain each year. If it is conservatively assumed that each rat consumes or destroys grain and other materials valued at $15 per year, then the total cost of destruction by rats in the US is more than $19 billion per year. In addition, rats cause fires by gnawing electric wires, pollute foodstuffs, and act as major vectors in the transmission of several diseases, including salmonellosis and leptospirosis, and to a lesser degree plague and murine typhus.

The Indian mongoose (*Herpestes auropunctatus*) was first introduced into Jamaica in 1872 for the biological control of rats in sugarcane plantations. It was soon introduced to Puerto Rico, other West Indian Islands, and Hawaii. The mongoose was effective in reducing the Asiatic rat population in some areas, but with reduced competition the European rat returned to being a major pest in sugarcane fields. Moreover, the mongoose preyed heavily on ground-nesting birds and subsequently reduced their numbers. It also preyed on amphibians and reptiles that were beneficial to biological pest control, resulting in a minimum of 7–12 extinctions in Puerto Rico and other islands of the West Indies. In addition, the mongoose emerged as the major vector and reservoir of rabies and leptospirosis in Puerto Rico and other islands. Based on public health effects, poultry losses, extinctions of amphibians and reptiles, and the elimination of many native birds, we estimate that the mongoose causes approximately $50 million in damages each year in Puerto Rico and the Hawaiian Islands (Pimentel, 2011; Zuniga, personal communication).

Finally, most dogs introduced into the US have been domesticated, although some have escaped into the wild. Some of these feral dogs run in packs and kill deer, rabbits, and domestic cattle, sheep, and goats. It is estimated that feral dog packs in Texas cause more than $5 million in livestock losses each year. Dog packs have also become a serious problem in Florida (Layne, 1997). Assuming $5 million in damages for the other 47 continental states, total losses in livestock kills per year would be approximately $10 million per year.

Exotic Birds

Of the 97 bird species introduced into the US, only 5% are considered beneficial; a majority (56%) are pests (Temple, 1992). Several nonindigenous bird species, including chickens and pigeons, were introduced into the US for agricultural purposes.

The common myna (*Acridotheres tristis*), introduced into Hawaii for biological control, helped control pest cutworms and armyworms in sugarcane. However, it became the major disperser of seeds of the introduced pest-weed *Lantana camara*. To cope with the weed problem, Hawaii was forced to introduce insect biocontrol agents for the weed (Kurdila, 1995).

The English or house sparrow (*Passer domesticus*) was introduced into the US intentionally in 1853 to control the canker worm. By 1900, the birds themselves were considered pests because they damage plants around homes and public buildings and consume wheat, green corn, and the buds of fruit trees. About 70% of the house sparrow's diet consists of grains (the introduced alfalfa weevil accounts for the other

30%). Furthermore, English sparrows harass robins, Baltimore orioles, yellow-billed cuckoos, and black-billed cuckoos. They also displace bluebirds, wrens, purple martins, and cliff swallows from their nest sites, and are associated with the spread of about 29 diseases that affect both humans and domestic livestock.

Similarly, European starlings (*Sturnus vulgaris*) are serious pests and are estimated to occur at densities of more than 1 per hectare in agricultural regions. They are capable of destroying approximately $2000 worth of cherries per hectare. In grain fields, starlings consume about $6 worth of grain per hectare. Therefore, assuming approximately $5 per hectare for all damages to agriculture crop production in the US, the total economic loss due to starlings would total approximately $800 million per year. In addition to these economic damages, starlings also have some negative environmental impacts. For example, starlings are aggressive and have displaced numerous native bird species, which can disrupt precariously balanced ecosystems. Starlings also have been implicated in the transmission of 25 diseases, including parrot fever and other diseases in humans and livestock.

The domestic pigeon (*Columba livia*) has been introduced for agricultural production and it has invaded most cities of the world, including US cities. Pigeons are considered a nuisance because they foul buildings, statues, cars, and sometimes pedestrians with their droppings, and because they feed on agricultural grain. Each pigeon consumes an average of 15 kg of grain per year (Smith, 1992). The economic impacts caused by fouling are estimated to be at least $9 per pigeon per year (based on $9 in control costs per pigeon). Assuming there is 1 pigeon per hectare in urban areas or approximately 0.5 pigeon per person in urban areas, common pigeons cause at least $2.2 billion in damages each year. This estimate does not take into consideration pigeons' role as reservoirs for greater than 50 serious human and livestock diseases, including parrot fever, ornithosis, histoplasmosis, and encephalitis.

Thus, if we assume $800 million per year in economic losses from starlings, $2 billion per year from pigeons, and $200 million for house sparrows and other birds, the damages from nonindigenous pest birds are estimated to be $3 billion per year.

Nonindigenous Amphibians and Reptiles

The cane toad (*Bufo marinus*) and bull frog (*Rana catesbeiana*) were introduced into Florida, Puerto Rico, Hawaii, and other warm regions, in some cases for the biological control of pest insects. However, the cane toad, with its toxic skin glands, has proved lethal to birds, dogs, cats, and other mammals.

Nonindigenous Invertebrates and Microbes

Approximately 4600 arthropod species (2600 species in Hawaii and more than 2000 in the continental US) have been introduced into the US. Eleven earthworm species and nearly 100 aquatic invertebrate species also have been introduced (OTA, 1993). More than 95% of these introductions were accidental – many invertebrate species gained entrance *via* plant introductions, in soil, or in water ballast from ships.

Insects and Mites

Approximately 1000 nonindigenous insect and mite species are pests in crops, stored-food products, and structures. Hawaii has about 5200 identified native insect species, and an additional 2600 introduced insect species (Howarth, 1990). Introduced insects account for 98% of the pest insects in that state. In addition to Florida's 11,500 indigenous insect species, 949 immigrant species have invaded the state (42 species were introduced for biological control). In California, the 600 introduced species are responsible for 67% of all crop losses. Some of the California pests include the cottony cushion scale (*Icerya purchasi*) and alfalfa weevil (*Hypera postica*).

Each year pest insects destroy about 13% of potential crop production at a cost of about $33 billion in US crops. Considering that about 40% of the pests were introduced, we estimate that these alien pests cause about $13 billion in crop losses each year. In addition, about $1.2 billion in pesticides are applied for all insect control each year. The portion applied against introduced pest insects costs approximately $0.5 billion. Therefore, the total cost for introduced nonindigenous insect pests is approximately $13.9 billion annually.

Lawn, garden, and golf course management activities cost about $36 billion annually. More than $3 billion per year is related to the expenses of pest control. Assuming the presence of 40% nonindigenous insect pests, the author estimates the control costs in lawns, gardens, and golf courses to be about $1.5 billion each year.

Other introduced insect species have become pests of livestock and wildlife. For example, the imported red fire ant (*Solenopsis invicta*) kills poultry chicks, lizards, snakes, and ground-nesting birds. In some areas fire ants are extremely abundant, with as many as 367 nests per hectare. The estimated damage to livestock, wildlife, and public health caused by fire ants in Texas is estimated to be $300 million. An additional $200 million per year is invested in controlling these ants. Assuming equal damages in other such infested southern states, the fire ant damages total approximately $2 billion per year.

Nonindigenous Earthworms

There are approximately 70 native US earthworm species. In disturbed habitats, 11 species of introduced earthworms have reduced the numbers of some of the native species. Although detailed data are not available concerning impacts of nonindigenous earthworms on US earthworm species and soil quality, earthworms are generally beneficial to soil productivity and formation in agriculture.

Microbe Introductions in Agriculture

Although some microbes were intentionally introduced into the US for wine and cheese making, most were accidentally introduced and some have become serious pests. The number of microbe species introduced into the US cannot be estimated because several thousand species can exist in a single gram of soil. Only a small fraction of all of these species have been identified.

More than 100 species of microbes have been intentionally introduced for processing wine, beer, cheese, and other foods. In addition, about 50 microbes have been introduced

for the biological control of pest insects. A strain of *Bacillus thuringiensis* (BT) was introduced and has been used extensively to control pest caterpillars. Other strains of BT have been developed to control beetles and mosquitoes.

An estimated 121 species of microbes, mostly introduced inadvertently in seeds and other parts of host plants, have become major crop pests in the US. US crop losses to all plant pathogens total approximately 12% of crops planted, a loss of $33 billion per year. Approximately 65% of these crop losses – an estimated $21 billion per year – are attributable to nonindigenous plant pathogens. In addition, $0.72 billion is spent annually on fungicides to control plant pathogens; approximately $0.5 billion of this goes toward the control of nonindigenous plant pathogens.

As mentioned earlier, lawn, garden, and golf course management activities have an annual cost of about $36 billion; most of this is spent on pest control. Assuming that about $3 billion of management costs are related to plant pathogen control and 65% of the pests are nonindigenous pathogens, we estimate the costs caused by introduced plant pathogens in lawns, gardens, and golf courses to be $2 billion each year. Therefore, the damages caused by nonindigenous pathogens and the attempts to control them total about $22.5 billion annually.

When all of the foregoing economic costs in crop and pastureland losses and for pest control for nonindigenous weeds, mammals, birds, insects, livestock pathogens, and plant pathogens are combined, the annual total is approximately $109.91 billion (**Table 1**) to the US economy.

Exotic Species Introductions in the United Kingdom

Nonindigenous Plants

There are a total of 26,000 introduced plant species in the UK. An estimated 25,000 species are established in UK botanical gardens alone, 14,000 species are cultured as commercial horticultural crops, and there are 6000 species of noncultivated aliens. There are 1169 species of exotic plants known to be established in natural ecosystems. The native flora of the UK numbers only 1515 species (Crawley *et al.*, 1996).

Most of the alien plants occur in relatively few types of habitats. More than 80% of the alien species are present in waste ground areas, urban sites, roadway sides, and agricultural habitats. An estimated 63% of the alien plants occur in hedges and scrub areas; croplands and gardens harbor about 43% of the alien species (Crawley *et al.*, 1996). Finally, about 40% of alien species occur in rock walls and woodlands. It is interesting that plant communities like grazed, mesic grasslands and native *Pinus sylvestris* woodlands contain no alien plant species.

In UK agriculture, weeds cause an average reduction of about 10% in crop yields; however, this loss can be as high as 32% in some crops. In economic terms, about $2.8 billion in potential crop production is annually lost due to weed infestations. Given the estimate that about 43% of the weeds are alien, it follows that $1.2 billion of the crop losses are due to introduced weeds (**Table 1**).

Nonindigenous Vertebrates

Introduced Mammals

The total number of mammalian species in the UK is 54; 17 of these are alien species. The mammal introductions include domestic animals like dogs, cats, cattle, horses, sheep, and pigs, as well as other nondomesticated mammals. Some of the species that were intentionally or accidentally introduced into the UK include the gray squirrel (*Sciurus carolinensis*), European rabbit, North American mink (*Mustela vison*), brown rat, and black rat (Lever, 1994). All of these animals, except for cattle and horses, have escaped into the wild and are now well established.

In the UK, rodents are serious pests on farms, in industries, and in homes. The estimated number of rats are estimated at five per person (based on the rat-to-person ratio in the US), including rats on farms (Pimentel *et al.*, 2000). Thus the total number of rats is 295 million. Using the same value of $15 for food and other goods damaged per rat, the cost of rats in the UK is about $4.4 billion annually.

The European rabbit is also abundant in the UK, with densities in some areas of up to 30 rabbits per hectare. Assuming approximately 10 rabbits per hectare on the 7 million ha of cropland in the UK with an estimated $11 damage per rabbit, the total annual economic damages from European rabbits are $800 million per year. They are reported to reduce wheat production by about 5%–8%, and to reduce forage production for livestock by about 20%. With 11 million ha of pasture for livestock, we estimate these losses to be $400 million per year. Thus, the total damage from the European rabbit in the UK is $1.2 billion per year.

Introduced Birds

Out of the 542 bird species in the UK, 47 are alien species. Some of the introduced species include the Canada goose (*Branta canadensis*) and the little owl (*Athene noctua*). Only one of the 47 introduced bird species is currently causing major ecological or economic problems – the common pigeon. This bird does invade agriculture, but is most common in cities and towns (Lever, 1994). Pigeons are a particularly serious problem for reasons discussed earlier, namely, the pollution of city surfaces with their droppings and their consumption of grain in agriculture.

Assuming there is 1 pigeon per hectare, or 0.5 pigeon per person in urban areas (as in the US), then there are approximately 30 million pigeons in the UK. The estimated damage that a pigeon causes is a minimum of $9 per year to crops and structures in cities and towns (Pimentel *et al.*, 2000). Therefore, pigeon damages are estimated to be at least $270 million per year. This does not include the role of pigeons as reservoirs for greater than 50 human and livestock diseases, including parrot fever, ornithosis, histoplasmosis, and encephalitis. Pigeons are also responsible for transmission of at least three diseases to UK poultry, including Newcastle disease.

Nonindigenous Invertebrates and Microbes

Introduced Arthropods

There are approximately 23,000 native species and 1700 introduced species of arthropods in the UK. An estimated

1500 species are of economic importance; about 169 alien species are considered pests. An estimated 30% of the crop losses in the UK are associated with the introduced arthropod pests, as compared with about 40% in the US (Pimentel *et al.*, 2000).

Arthropods damage or destroy approximately $2.8 billion in crops in the UK each year, based on the average of 10% crop losses per year. With about 30% of these losses due to introduced arthropods, they cause an economic loss of about $840 million per year.

Introduced Crop Plant Pathogens

An estimated 74% of the plant pathogens in the UK were introduced when seeds and other crop parts were brought into the country for agriculture. Approximately $8 billion is lost to all pests in crop production; about 8% of total potential production is lost to plant pathogens at a cost of about $2.3 billion per year. If 74% of these crop losses are due to introduced plant pathogens, then about $1.7 billion per year is associated with introduced microbes in crops.

Exotic Species Introductions in Australia

Nonindigenous Plants

There are approximately 20,000 vascular plant species in Australia (Pimentel, in press), including an estimated 1952 alien species. The rate of alien species introductions into Victoria, Australia, alone has been five to six species per year during the past century. Many of these species have become weeds and have invaded a wide range of environments. The invasive plants are a serious problem in both agricultural and wild ecosystems, where they disrupt key natural ecosystems, alter fire regimes, and reduce the resources for native animals. An estimated 60% of the weed species in crops in Australia are alien (based on a survey of major weeds in cereal crops).

The introduced blackberry (*Rubus proceus*) alone causes $77 million worth of damages to crop production each year. The indirect and direct losses due to all weeds in pastures are estimated to be $970 million per year. Weeds are estimated to cause about $4 billion in total damages in crops and pastures. Therefore, invasive plants cause approximately $2.4 billion per year in losses to agriculture (**Table 1**).

Nonindigenous Vertebrates

Introduced Mammals

There are presently 20 introduced mammals in Australia, including the European rabbit, hares (*Lepes europaeus*), water buffaloes (*Bubalis bubalis arnee*), cats, dogs, foxes (*Vulpes vulpes*), sheep, goats, cattle, horses, camels (*Camelus dromedarius*), pigs, and donkeys. In comparison, the number of native mammals is 227. The populations of some of these introduced mammals in the Northern Territory are quite high: buffaloes, 340,000; horses, up to 300,000; donkeys, up to 140,000; and camels, up to 30,000. The populations of these herbivores are much too high, but the costs of implementing a control program are also very high. For example, the cost of controlling buffalo is nearly $30 per animal.

The herbivorous animals significantly reduce the vegetative cover by overgrazing and this intensifies soil erosion and encourages annual plants and inedible shrubs in pastures (Lever, 1994). Surprisingly, the most serious pest herbivorous mammal in Australia is not one of these large-hoofed herbivores, but rather the European rabbit. In Tasmania, the rabbit population in 1952 reached a density of up to 250 rabbits per hectare. The total number of rabbits in Australia ranges from 200 to 300 million. Approximately 15 rabbits consume the equivalent pasture forage of one sheep (Emmerson and McCulloch, 1994). The impact of rabbits on sheep production per year is estimated to be $110 million, including reduced sheep production and rabbit control costs. If we assume a conservative estimate of only 0.5 rabbit per hectare for cropland and pastureland and also assume that each rabbit causes a minimum of $5 damage, then on the 465 million ha of Australian cropland and pastureland, rabbits are causing at least $1.2 billion per year in damages.

Feral pigs are also a serious problem in Australia, as they damage fences and spread animal diseases, including tuberculosis, brucellosis, rabies, and foot-and-mouth disease (Lever, 1994). The number of feral pigs ranges from 4 to 20 million. These pigs damage crops, kill lambs, and damage the natural environment; they are estimated to cause at least $80 million per year in damages (Emmerson and McCulloch, 1994).

The house mouse damages crops, houses, farm machinery, and livestock production. Estimates for these annual losses range from $50 to $100 million per year. Far more serious than the house mouse, though, are the invading brown and black rats. In the US, there are 4.6 rats per person (Pimentel *et al.*, 2000). Assuming the same number of rats per person in Australia and that each rat causes $15 of damage, then for a human population of 18 million, rats are causing approximately $1.2 billion per year in damages in Australia.

Introduced pet cats and feral cats are also a serious problem, especially for native bird, mammal, and amphibian populations. There are an estimated 18 million feral cats in Australia and each cat is estimated to kill 8 birds per year. Assuming that 144 million birds per year are killed, and that each bird has a value of $30, then damages from cats alone reach about $4.3 billion per year.

Introduced Birds

Australia has 850 bird species, of which about 70 are alien species. Most of the introductions have been intentional, including the English starling, English house sparrow, and common pigeon. These birds are often restricted to cities and towns. Pigeons cause similar problems in Australia as they do elsewhere.

Introduced Reptiles and Amphibians

There are about 700 species of reptiles and amphibians in Australia, but only 2 of them are nonindigenous. One of these introduced amphibians, the cane toad (*B. marinus*), was introduced as a biological control agent for insect pests in sugarcane fields. Unfortunately, the toad has become a pest itself because it is poisonous to dogs, cats, and other mammals that attack it.

Nonindigenous Invertebrates

Introduced Arthropods

There are an estimated 108,000 arthropod species in Australia, with 54,000 species of native and nonnative insects and mites and 10,000 mollusk species. Crop losses due to insects and mites are estimated at 10.7% of Australia's gross potential production of $22 billion per year. An estimated 36% of the pest arthropods in Australia are alien species. Based on arthropod-caused crop losses of about $2.4 billion per year, the exotic pests account for losses of $860 million per year. In addition, three exotic insects and mites cause $228 million per year in damages to the wool industry alone. Thus, we estimate that exotic insect and mite species in Australia cause losses of at least $1 billion per year.

Introduced Crop Plant Pathogens

If the total potential crop production in Australia is $22 billion per year and about 15.2% of crop losses are due to plant pathogens, the economic costs of these pathogens total about $3.3 billion per year. Because a large number of plant pathogens are introduced with crop seeds and other plant parts, an estimated 82% of the plant pathogens in crops are believed to be alien species (based on plant pathogens in field crops). Therefore, Australia loses $2.7 billion per year in crops from exotic plant pathogens.

Microbe Introductions Affecting Livestock

There are an estimated 44 exotic animal diseases that could infect livestock in Australia. One exotic livestock disease, sheep pox, costs the wool industry an estimated $21 million per year (Pimentel, unpublished data).

Exotic Species Introductions in South Africa

Nonindigenous Plants

There are 24,000 plant species in South Africa, including an estimated 8750 alien species. Most of the alien species were introduced from South America and Australia and have in turn invaded a wide range of environments. A total of 273 species of introduced plants in South Africa are serious weeds in crops (Bromilow, 1995).

Reduced crop production due to all weeds is 16.6% of potential crop production and total $3.7 billion per year in losses (total potential agricultural production is $22 billion per year). Assuming that 67% of the weeds in crops are alien (Bromilow, 1995), then the total loss in crop production due to alien weeds is $2.5 billion per year (**Table 1**).

Two of the most serious plants invading pasturelands include the shrub *L. camara* and the cactus plant *Opuntia ficus-indica*. In addition to these two weeds, there are approximately 800 alien weed species out of a total of 1604.

Nonindigenous Vertebrates

Introduced Mammals

There are 16 species of introduced mammals in South Africa, including the European rabbit, hares, water buffaloes, cats, dogs, sheep, goats, cattle, horses, camels, pigs, and donkeys. The total number of mammal species in South Africa, including alien species, is 247. As in Australia, the herbivore populations in South Africa are much too high for the resources that are available, and they significantly reduce the vegetative cover by overgrazing. This intensifies soil erosion and encourages some annual plants, weeds, and inedible shrubs to take over the pasturelands (Pimentel *et al.*, 2001).

Feral pigs are also a serious problem; they damage fences and spread disease in South Africa much as they do in Australia. The estimated control cost per feral pig in the US is about $100 (Pimentel *et al.*, 2000), about the same as in South Africa.

Rats are a serious problem in South African agriculture as well as in urban areas. There are an estimated 4.6 rats per person in South Africa (based on data from the US; Pimentel *et al.*, 2000), which has a human population of 39 million people. Thus, there are about 179 million rats in South Africa; each rat is assumed to cause $15 in damages per year. From these estimates, losses from rat damage total $2.7 billion per year.

Introduced Birds

Of the 725 bird species in South Africa, only eight are alien species. Some of the introduced species include the English starling, common pigeon, Indian myna, and English house sparrow. Only the starling and pigeon are currently causing economic problems, and their damage is restricted to cities and towns. The problems with starlings and sparrows are very similar to the negative effects that these birds have elsewhere; namely, they pollute structures with their droppings and cause agricultural losses by consuming grains and fruits. In addition, pigeons, starlings, and sparrows are known reservoirs and vectors of up to 50 different diseases of humans and livestock.

Nonindigenous Invertebrates and Microbes

Introduced Arthropods

An estimated 80,000 species of insects, 6000 species of spiders, and numerous other arthropod species exist in South Africa. About 20% of these are believed to be exotic. One of the most serious invaders is the Argentine ant (*Linepithema humile*), which causes major problems because it destroys native vegetation, including endangered plants. The same ant also negatively affects native ants and other species of arthropods by competing with them for limited resources.

Insect and mite pests in agriculture cause 16.7% or $3.7 billion in losses of potential crop production each year. Because approximately 45% of the insect and mite pests are exotic, the economic losses to exotic pests are estimated to be $1.7 billion per year.

Introduced Crop Plant Pathogens

Approximately 85% of the plant pathogens that attack crops in South Africa are nonnative species (based on an assessment of diseases of fruits and vegetables). Most of these pathogens were introduced with the introduction of crops into South Africa. In forests, around 69% of the pathogens are exotic. Plant pathogens in South Africa cause an estimated 15.6% or

$3.4 billion per year loss of the potential crop production. Since 85% of the pathogens are exotic, crop losses to exotic species total $2.9 billion per year.

Microbe Introductions Affecting Livestock

Several serious livestock diseases, including tuberculosis, brucellosis, East Coast fever, anthrax, and rinderpest, infect livestock and other animals in South Africa (Pimentel, unpublished data). Estimates suggest that these livestock diseases are causing losses of around $1 billion per year (**Table 1**).

Exotic Species Introductions in India

Nonindigenous Plants

There are 45,000 plant species in India, and an estimated 18,000 of these are alien species (Saxena, 1991). Many of the alien species have become weeds and have invaded a wide range of environments. Several weed species have been introduced along with the introductions of new crops. Weeds are estimated to cause a 30% loss in potential crop production, which totals about $54 billion per year. Assuming that 42% of the weeds in crop production are alien, the total cost associated with invading weeds is $22.8 billion per year (**Table 1**).

L. camara, a shrub introduced from South America as an ornamental plant, is a major weed in India, and it has invaded most pasturelands (13.2 million ha). Lantana is toxic to cattle and the cost of controlling it is $70 per hectare. Only about 4% of India's land area is in pasture, yet total damage per year from Lantana is $924 million.

Nonindigenous Vertebrates

Introduced Mammals

There are approximately 30 species of nonnative mammals in India, including cats, dogs, sheep, goats, pigs, axis deer (*Axis axis*), house mice, and rats. The total number of mammal species in India, including alien species, totals 320 (Pimentel *et al.*, 2001). As in other countries, the introduced herbivorous animals significantly reduce the vegetative cover by overgrazing, resulting in intensified soil erosion and increased invasion of pastureland by inedible weeds and shrubs.

Rats number at least 2.5 billion in India, or about 2.7 rats per person (Pimentel *et al.*, 2001). They attack crops in the field and are estimated to reduce potential crop yields by about 2%. In addition, rats are especially serious pests of stored grain supplies. Various studies report that an individual rat consumes or destroys about 210 kg of grain per year in India, and up to 740 kg of grain per year in Pakistan. Rats are estimated to destroy about 12 million tons of grain per year in India (Pimentel, unpublished data), in addition to damaging other foods, goods, and structures. In India the author estimates that each rat causes at least $10 of damage per year, and thus they are responsible for at least $25 billion annually (Pimentel *et al.*, 2001). In addition, rats are major vectors of and carriers of 38 human and livestock diseases. An average of 250,000 people die each year from the plague in India.

Introduced Birds

The number of bird species in India is 1221, including migrants and vagrants. Introduced species number only four, and include the English house sparrow, common pigeon, black francolin, and Alexandrine parakeet. In neighboring Pakistan, sparrows reduce potential wheat yields by 170,000 tons at a cost of $26 million per year (Pimentel, 2011). Based on these data, losses in India are calculated to be $50 million per year. Also, the common pigeon is a problem in agriculture, as well as in cities and towns, where it consumes grains and fouls buildings, statues, cars, and the occasional unlucky pedestrian. Pigeons are also involved in the spread of about 50 diseases that affect humans and livestock.

Nonindigenous Invertebrates and Microbes

Introduced Mollusks

An estimated 1500 species of land mollusks exist in India; several of these are exotic. The giant African snail (*Achatina fulica*) destroys from 0.29 to 4.3 g of grain per day per snail from eight major crops.

Introduced Arthropods

There are more than 30,000 species of arthropods in India. An estimated 560 mite species and about 600 species of insects are crop pests. About 30% of the pest species are introduced arthropods, and as a group they reduce potential crop production by 18.7%. Based on total potential crop production in India of $181 billion per year, crop losses to alien arthropods total $10.2 billion per year.

Introduced Crop Plant Pathogens

Plant pathogens reduce potential crop production in India by approximately 16%, at a total cost of $29 billion per year. There are approximately 30,000 species of plant pathogens in India, and about 74% of the major pathogens are exotic species (based on the major plant pathogens in vegetable crops). Thus, the total cost of invading plant pathogens to crops in India is about $21.4 billion per year.

Microbe Introductions Affecting Livestock

Several major diseases of livestock cause significant losses in India, most significantly foot-and-mouth disease. During 8 months in 1996, nearly 50,000 cases were reported at an estimated cost of $17,000 per year.

The combined cost of crop losses that can be attributed to nonindigenous plants, mammals, arthropods, and plant pathogens is $80 billion. Thus the negative effect of exotic species on Indian agriculture rivals that found in the US.

Exotic Species Introductions in Brazil

Nonindigenous Plants

Of the 55,000 plant species in Brazil, an estimated 21.1%, or 11,605, are alien species. Many of the alien species have become weeds and invaded a range of environments. In crop production, alien species make up 75% of the weed species. Weeds are estimated to destroy about 13.4% of Brazil's

potential crop and pasture production, or about $12.3 billion per year (**Table 1**).

Nonindigenous Vertebrates

Introduced Mammals

The number of introduced mammals in Brazil is estimated to be 30, including cats, dogs, sheep, goats, cattle, horses, pigs, and donkeys. The total number of mammal species in Brazil, including alien species, is 428.

On the outskirts of São Paulo in the community of Taboao de Sera, estimates suggest that there are 12,500 feral dogs and 4600 feral cats (Pimentel *et al.*, 2001). The feral dogs attack livestock and other animals. Feral cats are also a serious problem, as they destroy birds and other native animals in Brazil.

It is estimated that there are about 320 million rats in Brazil. Assuming that each rat causes $10 in damages, the total cost of rats is estimated to be $3.2 billion per year.

Introduced Birds

The number of native bird species in Brazil is 1635. Only three, the English house sparrow, common wax-bill (*Estrilda astrild*), and common pigeon, are introduced species. The sparrow and pigeon populations cause major ecological and/or economic problems. The common waxbill is not considered a serious problem because it feeds primarily on the introduced guinea grass (*Panicum maximum*) and does little damage to this fast-growing grass (Pimentel, unpublished data).

Nonindigenous Invertebrates and Microbes

Introduced Arthropods

Invertebrate species number more than 100,000 in Brazil; about 70,000 are arthropod species. About 14.4% of potential crop production is destroyed by insects and mites, and approximately 35% of these pests are exotic. The calculated loss of crops to exotic insects and mites is estimated to be $4.6 billion per year.

Introduced Crop Plant Pathogens

There are an estimated 100,000 species of microbes in Brazil; around 75% of the microbes that attack crops are exotic. Like most countries, most of these plant pathogens were introduced via alien crop species. Plant pathogens are estimated to cause 13.5% in crop losses each year. If 75% of the plant pathogen species are exotic, estimated losses from alien species total $18.3 billion per year.

Conclusions

With approximately 400,000 nonindigenous species in various nations worldwide, if even a small fraction of these invaders are harmful, significant agricultural problems can result. While nearly all of our crop and livestock species are nonindigenous and have proven essential to the viability of the world's agriculture and economy, exotic species invasions do result in many negative financial effects. This article shows that the economic damages of these alien species to the agricultural economies of just six nations (including, in some cases, control costs) amount to approximately $262 billion each year (see **Table 1**). If this figure is extrapolated to all nations based on costs per person, the total cost of exotic species to world agriculture would be about $943 billion per year.

Precise economic data for some of the most ecologically damaging biological invaders are not available. Mammals, including horses, water buffaloes, and goats, have been responsible for reducing the productivity of pastures and rangelands, as well as the extinction of many plant species in various ecosystems. In other areas, feral pigs and feral dogs have had serious impacts on agriculture, but as of yet few data are available concerning these animals.

Most of the exotic species in the world have been introduced within the past 70 years. Accelerated international trade and travel, plus rapidly increasing human numbers, guarantee that the threat from nonindigenous species invasions is still growing. The true challenge lies not in assigning precise dollar values to exotic species losses but in preventing further biological introductions and the resultant damage to managed and natural ecosystems. Although policies and practices to prevent the accidental and intentional introduction of exotic species are improving, the world's nations are woefully short of allocating sufficient resources to the problem in proportion to the risks. Both natural and managed ecosystems need to be protected from additional costly damages resulting from introduced species.

See also: Endangered Ecosystems. Extinction, Causes of. Introduced Species, Impacts and Distribution of. Pesticides, Uses and Effects of. Plant Invasions. Range Ecology, Global Livestock Influences

References

Babbitt B (1998) Statement by Secretary of the Interior on invasive alien species, *Proceedings of the National Weed Symposium*. Bureau of Land Management Weed Page. April 8–10.

Bromilow C (1995) *Problem Plants of South Africa*. Arcadia, South Africa: Briza.

Bryan RT (1996) Alien species and emerging infectious diseases: Past lessons and future applications. In: Sandlund GT and Schel PJ (eds.) *Proceedings of the Norway/UN Conference on Alien Species*, 1–5 July 1996, pp. 74–80. Trondheim, Norway: Norwegian Institute for Nature Research.

Crawley MJ, Harvey PH, and Purvis A (1996) Comparative ecology of the native and alien floras of the British Isles. *Philosophical Transactions of the Royal Society of London, Series B Biological Science* 351: 1251–1259.

Emmerson G and McCulloch J (1994) *Feral Peril: Queensland's Introduced Plants and Animals*. Brisbane, Australia: Queensland Parliamentary Library.

Howarth FG (1990) Hawaiian terrestrial arthropods: An overview. *Bishop Museum Occasional Papers* 30, pp. 4–26.

Kurdila J (1995) The introduction of exotic species into the United States: There goes the neighborhood. *Environmental Affairs* 16: 95–118.

Layne JN (1997) Nonindigenous mammals. In: Simberloff D, Schmitz DC, and Brown TC (eds.) *Strangers in Paradise*, pp. 157–186. Washington, DC: Island Press.

Lever C (1994) *Naturalized Animals: The Ecology of Successfully Introduced Species*. London: T. & A. D. Poyser Natural History.

Office of Technology Assessment (OTA) (1993) *Harmful Nonindigenous Species in the United States*. Washington, DC: U.S. Government Printing Office.

Pimentel D (2011) Environmental and economic costs associated with alien invasive species in the United States. In: Pimentel D (ed.) *Biological Invasions: Economic and Environmental Costs of Alien Plant, Animal, and Microbe Species*, 2nd edn., pp 411–430. Boca Raton, FL: CRC Press.

Pimentel D, Lach L, Zuniga R, and Morrison D (2000) Environmental and economic costs associated with nonindigenous species in the United States. *BioScience* 50(1): 53–65.

Pimentel D, McNair S, Janecka J, *et al.* (2001) Ecological and economic threat of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems and Environment* 84: 1–20.

Saxena KG (1991) Biological invasions in the Indian subcontinent: Review of invasion plants. In: Ramakrishnan PS (ed.) *Ecology of Biological Invasions in the Tropics*, pp. 53–73. New Delhi: International Scientific Publications.

Simberloff D, Schmitz DC, and Brown TC (1997) *Strangers in Paradise*. Washington, DC: Island Press.

Smith RH (1992) Rodents and birds as invaders of stored-grain ecosystems. In: Jayas DS, White NDG, and Muir WE (eds.) *Books in Soils, Plants, and the Environment: Stored-Grain Ecosystems*, pp. 289–323. New York: Marcel Dekker.

Temple SA (1992) Exotic birds, a growing problem with no easy solution. *The Auk* 109: 395–397.

Agriculture, Industrialized

Prabhu Pingali and Melinda Smale, International Maize and Wheat Improvement Center, Texcoco, Mexico

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 85–97, © 2001, Elsevier Inc.

Glossary

Alleles Variants of nucleotides within the genome that cause changes in a given protein.

Crop genetic diversity: The diversity of the sets of genes carried by different individuals within a crop species. It occurs in the form of nucleotide variation within the genome.

Genetic diversity The diversity of the sets of genes carried by different organisms. It occurs in the form of nucleotide variation within the genome. When this variation causes a change in a given protein, the variants are called alleles. Allelic variation occurs at various genetic loci or gene positions within a chromosome.

Genetic erosion The loss of genes from a gene pool attributed to the elimination of populations caused by factors such as the adoption of high-yielding varieties, farmers' increased integration into the market, land clearing, urbanization, and cultural change.

Genetic variation The allelic variation that occurs at various genetic loci or gene positions within a chromosome. Genetically variable loci are termed polymorphic or are said to show polymorphism.

Landraces These have a complex nature; therefore, it is not possible to give an all-embracing definition. Landraces are the originally adapted but variable crop populations on which farmers based their selections. They are geographically or ecologically distinctive populations which are conspicuously diverse in their genetic composition both between populations and within them. They have certain genetic integrity and are recognizable morphologically; farmers have names for them and different landraces are understood to differ in adaptation to soil type, time of seeding, date of maturity, height, nutritive value, use, and other properties.

Loci Gene positions within a chromosome. Allelic variation occurs at various genetic loci.

Spatial and/or temporal patterns of genetic Variation The systematic changes in the alleles occurring at specific loci along spatial and/or temporal dimensions.

Introduction

During the past 200 years, and in a more accelerated way since 1960, the pattern of crop genetic diversity in the fields of the developing world has changed fundamentally. The germplasm that dominates the area planted with major cereals has shifted from the locally adapted populations that farmers historically selected from seed they saved—often called "landraces"—to the more widely adapted seed types produced by scientific plant breeding programs and purchased by farmers—often called "modern varieties." These yield-enhancing seed types enabled the intensification of agriculture in areas of the world with high population densities. Initially, they diffused through the environments best suited for their production, spreading later—and unevenly—into less favored areas.

Economic growth, increasing incomes, and migration of labor from rural to urban areas lead to the commercialization of agriculture. Commercial crop production is characterized by the controlled application of water and other purchased inputs, such as fertilizer, fungicides, or pesticides to modern seed types. Sustained yield increases in commercial agriculture have depended not only on continued modifications in crop management practices but also on successive genetic improvements that are accomplished through recombining and exchanging diverse sources of germplasm. The spatial and temporal diversity among modern varieties in farmers' fields is determined in large part by the economic factors affecting their profitability and the performance of agricultural research institutions and seed industries.

In contrast, the genetic content and the geographical distribution of landrace populations are influenced more by natural selection pressures and the seed and crop management practices of traditional farming communities. Typically, it is not profitable to grow them with large amounts of purchased inputs since they are often grown principally for home consumption of food or feed and the size of their commercial market is limited. Seed markets for these varieties are generally local and rules for the exchange for seed, grain, and other products are more personal.

Today, the modernization of agriculture in the developing world is incomplete. In some areas with rain-fed, upland, or heterogeneous growing environments, landraces still dominate crop area or both modern varieties and landraces coexist. As the process of modernization continues, more farmers will be integrated into specialized product markets. Whereas others will work outside the farm and eventually leave it. Even then, local or export demand for certain quality attributes may provide market-based incentives for continued cultivation of landraces. In some places, there will be no alternative but to grow landraces for food since there may be no urban target for migration.

This article outlines some of the major implications of the modernization of agriculture for crop genetic diversity, focusing on the world's three principal cereals—rice, wheat, and maize. Wheat is a major tradable crop grown principally in temperate zones over a wide geographic area and often on relatively large, mechanized farms of industrialized countries.

The world's main wheat-producing regions are southern Russia and the Ukraine, the central plains of the United States and adjacent areas in Canada, northwest Europe, the Mediterranean basin, southwestern Australia, and, in the developing world, India, China, and Argentina. Rice is the most important food grain in the developing world. Almost all the world's rice is both produced and consumed on numerous small farms within Asia, with less than 5% of production entering trade. India, China, and Indonesia are the major producers. Maize is the world's most widely grown cereal in terms of growing environments and, after wheat, the most extensively traded cereal. Although virtually all the wheat and rice traded internationally are destined for human consumption, most of the maize is used to feed livestock. Maize is a very important part of the human diet in Africa, parts of Latin America, and Asia. Major maize-producing regions include the United States and parts of Europe, Mexico, Brazil, Argentina, South Africa, India, and China.

The next section summarizes the factors driving the intensification of agriculture. The section on *The Implications of Intensification for Crop Genetic Diversity* explains the implications of intensification for the patterns of genetic variation in the fields of the developing world. The section on *The Commercialization of Crop Production* discusses how commercialization affects these patterns and presents hypotheses concerning the future for crop genetic diversity.

The Intensification of Crop Production

What is the Intensification of Crop Production?

The intensification of agriculture refers to the increase in output per unit of land used in production, or land productivity. Population densities, expressed by the ratio of labor to land, explain much about where and under which conditions this process has occurred. The transition from low-yield, land-extensive cultivation systems to land-intensive, double- and triple-crop systems is only profitable in societies in which the supply of uncultivated land has been exhausted. The process of agricultural intensification has been observed in traditional and modern agricultural societies. The movement from forest and bush fallow systems of cultivation to annual and multicrop cultivation systems, whereby plots of land are cultivated one or more times per year, has generally been influenced by increasing population densities in traditional societies. It is no accident that the modern seed-fertilizer revolution has been most successful in densely populated areas of the world, where traditional mechanisms for enhancing yields per unit area have been exhausted.

Intensive cultivation will also be observed in areas with lower population densities provided that soil conditions are suitable and markets are accessible. Intensification occurs in the less densely populated areas for two reasons: (i) Higher prices and elastic demand for output imply that the marginal utility of effort increases—hence, farmers in the region will begin cultivating larger areas, and (ii) higher returns to labor encourage migration into well-connected areas from neighboring regions with higher transport costs. Examples of regions with low population density but intensive, market-oriented production are the central plains of Thailand and parts of South America's southern cone. If the conditions described are not present, labor and other costs associated with intensive agriculture are substantially higher than its incremental economic returns.

Intensification of land use and the adoption of yield-enhancing technologies have occurred in traditional and modern agricultural systems. In the case of Africa, Pingali *et al.* (1997) documented the movement from shifting cultivation to permanent agriculture with increases in population densities and improvements in market infrastructure. As land became scarce, traditional farming communities across sub-Saharan Africa began to extract increasingly higher levels of output from their land through investments in land improvements and soil fertility management. The intensification of traditional farming systems is a process that the more densely populated regions of Asia had been through several decades and in some cases centuries earlier. The application of modern scientific innovations in the form of high-yielding seeds and fertilizer allowed the extremely land-scarce regions of Asia to achieve levels of land productivity that were not possible through the exclusive reliance on traditional systems of farmer innovation.

Seed Technical Change: Traditional and Modern

The genetic evolution of cultivated crops is closely interwoven with the evolution of human civilization and crop husbandry. The recognition of wild species of cereal crops was first accomplished by primitive societies of hunters and gatherers, who harvested them for food. The domestication of wild species occurred when societies of shifting cultivators first began to cultivate food crops. Sedentary or permanent agricultural systems led to the emergence of ancestors of modern-day landraces of cereal crops. Although landraces have evolved over thousands of years under farmer management of natural selection, varieties have been bred by plant breeding programs for about 100 years. The first high-yielding maize hybrids were developed about 50 years ago. The high-yielding varieties of rice and wheat with semidwarf stature were developed less than 50 years ago and have been successfully adopted only in intensive agricultural production systems. Modern cereal cultivars have developed through three main phases of selection: (i) subconscious selection by the earlier food growers in the process of harvesting and planting, (ii) deliberate selection among variable material by farmers living in settlements and communities, and (iii) purposeful selection by professional breeders using scientific methods.

The main attainment of the first phase was to make the crop more suitable for planting and harvest by humans, threshing or shelling, and consumption. Higher germination rates, more uniform growing periods, resistance to shattering, and palatability were some of the achievements of this effort. In the second phase, many farmers exerted pressures continuously in numerous directions, resulting in variable populations that were adapted to local growing conditions and consumption preferences. These are broadly known as landraces. During the third phase, fields of cereals have become more uniform in plant types with less spontaneous gene

exchange. Planned gene migration has increased, however, with the worldwide exchange of germplasm. The products of the third phase are loosely referred to as "modern varieties."

The genetic changes embodied in seed can lead to change in the productivity of land (yield) both directly and indirectly, in combination with other inputs. A primary example of such technical change is the "green revolution." A term which has been used in various contexts to symbolize various types of social and technical change, it is first said to have been used by a USAID administrator to describe the rapid movement through Asia of new wheat and rice varieties coupled with increased use of fertilizer and irrigation. Here, we use it to refer specifically to the widespread adoption of semidwarf rice and wheat varieties in the developing world during the late 1960s and early 1970s. When grown with increased levels of fertilizer and a controlled water supply, these varieties performed significantly better than the varieties they replaced. Initially, they spread rapidly throughout many of the irrigated zones of the developing world where rice and wheat cultivation was concentrated and where population densities were high. Later, more widely adapted descendants of these varieties spread gradually into less favorable environments, including rain-fed areas with relatively modest production potential. Their diffusion was faster in the plains and valleys, diminishing up the hillsides and in more heterogeneous environments. The term modern varieties is also used to refer more exclusively to semidwarf varieties of rice and wheat.

Conservationists who observed the popularity of the green revolution varieties expressed concern for the possible loss of valuable genetic resources and narrowing in the genetic stock that serves as the basis for crop improvement (Frankel, 1970; Harlan, 1992; Hawkes, 1983). In the next section, the implications of the intensification of agriculture for crop genetic diversity are outlined. First, crop genetic diversity is defined. Some historical perspective is then provided on the spatial and temporal distributions of genetic variation in rice, wheat, and maize. Current estimates of areas planted with modern and landrace varieties of these cereals are reported. Evidence on the genetic diversity of modern varieties and landraces follows.

The Implications of Intensification for Crop Genetic Diversity

Crop genetic diversity broadly defined refers to the genetic variation embodied in seed and expressed when challenged by the natural and human selection pressures that shape the environment in which it grows. In applied genetics, diversity refers to the variance among alternative forms of a gene (alleles) at individual gene positions on a chromosome (loci), among several loci, among individual plants in a population, or among populations (Brown *et al.*, 1990). Diversity can be measured by accessions of seed held in gene banks, lines or populations utilized in crop-breeding programs, or varieties cultivated by farmers (cultivars).

The relationship between precise quantitative measurements and what can be casually observed among the plants growing in a field, and between these measurements and what can be observed in other environments, is indirect. Crop genetic diversity cannot be literally or entirely observed at any point in time; it can only be indicated with reference to a specific crop population and analytical perspective. To understand the implications of agricultural intensification to crop genetic diversity, it is first necessary to gain a spatial and temporal perspective of the variation in crop plants.

Current Spatial Patterns of Genetic Variation

The adoption of modern cereal varieties has been most widespread in land-scarce environments and/or in areas well connected to domestic and international markets. Even in these areas, the profitability of modern variety adoption has been conditioned by the potential productivity of the land under cultivation. For instance, while modern rice and wheat varieties spread rapidly through the irrigated environments, their adoption has been less spectacular in the less favorable environments—the drought-prone and high-temperature environments for wheat and the drought- and flood-prone environments for rice. Maize, as discussed later, has an even spottier record in terms of farmer adoption of modern varieties and hybrids. For all three cereals, traditional landraces continue to be cultivated in the less favorable production environments throughout the developing world.

Recent estimates of the areas planted with modern varieties and landraces are shown in **Table 1**. Based on data from a global wheat survey conducted by the Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT) in 1997, approximately 80% of the wheat area in the developing world was sown to semidwarf varieties, with the remainder split almost equally between improved tall varieties and landraces or varieties with unknown ancestry. The relative importance of tall wheat varieties remains greater in the industrialized than in the developing countries, probably for reasons related to wheat-growing environment and management practices.

Most of the extensive area sown to wheat landraces is found in parts of Turkey, Iran, Afghanistan, and Ethiopia. Pockets of landrace diversity for special traits may also be found throughout the Mediterranean region (Morocco, Tunisia, Syria, Egypt, Cyprus, Portugal, Spain, and Italy) as well as Nepal, Outer Mongolia, and Tibet, although field research and data analysis would be needed to confirm these hypotheses. Work in Mexico, Guatemala, Honduras, Peru, and Bolivia indicates the presence of unique landrace varieties that were probably introduced by Spanish immigrants in the sixteenth century. No significant areas are believed to be sown to wheat landraces in the former Soviet Union or in Eastern Europe, despite the historical importance of these areas for major progenitors such as the so-called "Turkey" and "fife" wheat. The loss of landrace populations in the former Soviet Union was no doubt a consequence of collective agriculture that was practiced in the region for several decades.

Approximately three-fourths of the rice area in Asia, which produces most of the world's rice, is thought to be sown to semidwarf varieties. Semidwarf varieties dominate the irrigated rice ecosystems of Asia and cover substantial areas in the favorable rain-fed lowlands (M. Jackson and G. Khush, personal communication). In the more marginal rain-fed lowland environments, the deep-water environments, and the

Table 1 Percentage distribution of rice, wheat, and maize area by type of germplasm in the 1990s

Region	*Wheat*			*Maize*			*Rice*		
	Semidwarf improved	*Tall improved*	*Landraces/ unknown*	*Hybrid*	*Improved open-pollinated*	*Landraces/ unknown*	*Semidwarf improved*	*Other improved*	*Landraces/ unknown*
Sub-Saharan Africa[a]	66	14	20	38	8	54	25	15	60
West Asia/North Africa	66	10	24	22	7	71	11		
Asia	86	7	6	70	7	23	73	13	14
Latin America	90	9	1	43	5	52	59	36	5
All developing countries	81	9	11	53	7	40	71		
Industrialized countries	55	45	Trace	99	1	Trace	78		

[a]Data for rice in sub Saharan Africa are West Africa only.

Sources: IRRI World Rice Statistics (1995, 1999), CIMMYT Global Maize and Wheat Impacts Surveys (1992, 1997), Heisey, *et al.* (1999), Morris and López-Pereira (1999), Luis Roberto Sanint (CIAT), and Timothy Dalton (WARDA).

upland rice environments, farmers continue to grow landraces adapted to those particular environmental niches. Although more accurate data concerning rice landraces in Asia are now being compiled at the International Rice Research Institute (IRRI), the most comprehensive data from a national-level sample survey conducted by the Department of Agricultural Extension in Bangladesh (1996–1997) suggests that although farmers in that country still grow large numbers of named varieties including landraces (more than 300), more than 20 modern varieties cover nearly 65% of the total rice area. In West Africa, semidwarf varieties cover essentially all the irrigated lowlands and one-third of the rain-fed lowlands, but landraces and other improved varieties appear to occupy most of the area in uplands, mangrove, and deep-water flooded areas. Particularly in the deep-water/floating rice environment near Mopti in Mali, Birnin Kebbi in Nigeria, and in northern Guinea are found *Oryza glaberrima* (African rice) and tall traditional *O. sativa* (Asian rice) cultivars. The West African Rice Development Association (WARDA) has recently developed hybrid crosses of *O. sativa* and *O. glaberrima* species. In East Africa, the area in Madagascar seems to be approximately evenly split between traditional and modern varieties. In Latin America, tall and landrace varieties cover most of the upland area, whereas few are found in the irrigated and rain-fed lowlands.

In the uplands of Asia, traditional varieties still dominate, with "pockets" of important diversity in the Bastar Plateau (Madhya Pradesh, eastern India), parts of northern Bangladesh and the Chittagong Hill Tracts in that country, northern states of Myanmar, almost all of Laos, and parts of Cambodia (M. Jackson and G. Khush, personal communication). Oka (1988) reported that the genetic diversity in landraces of Asian rice was most prevalent in the area extending over Assam, Bangladesh, Burma (now Myanmar), Thailand, Laos, and Yunnan, China. Although Oka also reported that diversity was relatively high in Indonesia, which lies outside the center of domestication, much of Indonesian ricelands are now planted with modern varieties.

Data from global maize surveys conducted by CIMMYT in 1992 and 1997 indicate that, relative to wheat and rice, far less of the maize area in the developing world is planted with maize types released by plant breeding programs. Hybrids appear to occupy an increasing proportion of the area planted with modern maize types, but in zones such as Latin America, most of the maize area is still planted with landraces.

At least some of the area listed under landraces is planted with populations that result from the genetic integration of modern varieties with landraces when farmers save seed or plant seed of different types in adjacent fields. Similarly, a substantial proportion of the maize area in sub-Saharan Africa is planted with advanced generations of improved varieties whose seed farmers could not afford to replace on a regular basis introgressed with landrace populations brought to the continent with the slave trade several centuries ago. Although there appears to be a high proportion of maize landraces grown in West Asia and North Africa, the total area in maize is limited in this region and many of these are likely of unknown origin. In the industrialized world, a negligible percentage of maize area is planted with either improved open-pollinated varieties or landraces. These are specialty maizes or "heirloom" varieties grown for fresh consumption, popcorns, or ornamental corns such as those marketed on holidays in the United States.

The data in **Table 1** confirm that although much of the genetic variation in the world's wheat and rice area is shaped today by the efforts of modern plant breeders, a substantial part of the genetic variation in the maize grown in the developing world remains in the hands of farmers, some of which are among the poorest in the world. In some zones of production, such as Turkey and Iran for wheat or parts of Latin America for maize, fairly large contiguous areas may be planted to landraces. Often, in regions of crop domestication and diversity, landraces persist "as patches and islands of farming systems" (Brush, 1995, p. 246). Harlan (1992, p. 147) invoked the term "microcenters of diversity" to describe "relatively small regions, 100–500 km across, in which may be

packed an astonishing variation" within extensive areas of apparent uniformity in parts of Turkey, the Caucasus, Iran, and Afghanistan. Oka (1988) called the Jeypore Tract in India a microcenter of diversity in Asian rice. Jackson (personal communication) refers to some of the upland rice areas of Asia as "pockets" of diversity.

The change in the crop genetic landscape from predominantly traditional to largely modern patterns of genetic variation occurred during the past 200 years and at an accelerated rate since the 1960s. Whether the change to modern varieties has resulted in a narrowing of genetic diversity remains largely unresolved for many reasons, which are discussed in the following sections and further in Smale (1998) and Wood and Lenné (1997).

Comparing Genetic Diversity in Landraces and Modern Varieties

A major problem in assessing whether genetic narrowing has occurred with the replacement of landraces by modern varieties is the magnitude of the sampling and measurement effort that would be required to test the hypothesis in a meaningful way. In most cases it would be impossible to locate reliable samples of the landraces originally grown in an area now planted with modern varieties since this process occurred over time and unevenly across environments.

In some sense, the genetic diversity of landraces and modern varieties is incomparable by definition since its structure is distinct for each. Hawkes (1983) wrote that landraces, which are mixtures of genotypes, "could not even be called varieties." He called the range of genetically different varieties available to breeders "the other kind of diversity" (pp. 100–101). Harlan (1992) described a landrace as "an integrated unit" of "component genotypes" that have adjusted to one another over the generations as well as to the local environment, both natural and man-made (p. 148). A field planted with a rice or wheat landrace may be viewed by a plant breeder as a mixture of several lines but viewed by the farmer as one single variety because of its recognizably distinct agronomic, processing, or consumption characteristics. Vaughan and Chang (1992) described traditional rice types that are mixtures or composites of morphologically distinguishable types grown together deliberately.

The breeding system of the crop plant also affects the structure of genetic variation. Landraces generally contain some heterozygous material, though the extent of segregation is considerably greater for open-pollinated than for self-pollinated species. Maize is a cross-pollinating species, and maize in Mexico is often cited as an example of deliberate manipulation of the composition of landraces by farmers. Landrace varieties of maize evolve continuously through the purposeful mixing of seed lots of the same varieties or introgression by farmers, as well as inadvertently when fields of maize planted with different varieties flower simultaneously. A single "race" or race complex, as understood by maize geneticists, contains numerous genetically distinct farmers' varieties.

The distinction between modern varieties and landraces can also be blurred for predominantly outcrossing crops such as maize, making it difficult to determine which population is under study. Many small-scale, subsistence-oriented maize farmers promote hybridization between improved varieties and landraces by growing them together or in neighboring fields and producing what farmers in Mexico call "creolized" varieties. Also termed "rustication" or simply "adaptation," this process may enable improved varieties to fit better the need of local farmers.

These points imply that comparing counts of landraces and modern varieties as an index of genetic narrowing may not make sense. They also imply that even if reliable samples of the landraces originally cultivated in an area could be obtained, analyses comparing their genetic diversity might provide only part of the answer regarding genetic narrowing. Although the landrace in the farmers' field is a heterogeneous population of plants, it is derived from generations of selection by local farmers and is therefore likely to be local in adaptation. The plants of a modern variety are uniform but the diverse germplasm in the genetic background may enable them to adapt more widely. The diversity in a modern variety may not be expressed until challenged by the environment. On the other hand, the landrace may carry an allele that occurs rarely among modern varieties and is a potentially valuable source of genetic material not only for the farmer that grows it today but also for future generations of producers and consumers.

When did Genetic Narrowing Occur?

Another problem in assessing the relationship of modern varieties to genetic narrowing is the temporal point of reference. Porceddu *et al.* (1988) described two major stages of genetic narrowing in wheat during modern times. The first occurred in the nineteenth century when scientific plant breeding responded to the demand for new plant types. Farming systems emerged that were based on the intensive use of land and labor, livestock production, and the use of organic manure. Changes in cultivation methods favored genotypes that diverted large amounts of photosynthates into the ear and grain. Bell (1987) reports that the engineering innovations of the late nineteenth century led to the establishment of extensive wheat-growing areas in North America, Australia, and parts of South America. Mechanization of agriculture dictated uniformity in plant type.

According to Porceddu *et al.* (1988), a second stage of narrowing occurred in the twentieth century, when genes were introduced to produce major changes in plant type. Use of the dwarfing genes *Rht1* and *Rht2*, for example, conferred a positive genotype-by-environment interaction in which yield increases proved greater given a certain combination of soil moisture, soil fertility, and weed control. Varieties carrying these dwarfing genes were developed by N. Borlaug with the national breeding program in Mexico and later by the International Maize and Wheat Improvement Center (CIMMYT). They became known as the green revolution wheats.

Evenson and Gollin's (1997) summary of the history of rice breeding suggests a process of continual expansion and narrowing of the genetic pool. Organized breeding efforts probably date earlier than 1000 AD in China. Modern efforts

can be traced to the late nineteenth century in several parts of Asia. In temperate east Asia, the first significant advances were made by Japanese farmers and scientists when they developed relatively short-statured and fertilizer-responsive cultivars. Known as the *rono* varieties, these belonged to the *japonica* class of rice and were widely cultivated in Japan as early as the 1890s. During the Japanese occupation of Taiwan in the early part of the twentieth century, Japanese scientists sought to adapt these varieties to the more tropical conditions of Taiwan. At the same time, researchers in tropical Asia were seeking more productive varieties of rice from the *indica* and *javanica* classes of rice. After World War II, the United Nations Food and Agricultural Organization initiated a program to cross *indica* rice with *japonicas* as a means of increasing rice yields, culminating in the formation of the IRRI and the green revolution varieties of rice.

To Vaughan and Chang (1992), genetic narrowing in modern rice began early in this century. Development projects, population increases, and forest clearing in Asia were the primary causes of the loss of wild and cultivated rice landraces. In the Mekong Delta, the replacement of traditional deep-water rice by irrigated rice occurred with drainage and irrigation schemes that were introduced during the French colonial period.

Goodman (1995) reports that the major portion of the variability now found in maize developed before European contact (circa 1500), and several of the most widely grown races, including the commercially important Corn Belt dents, developed later. During the "corn show era" in the nineteenth century, U.S. farmers exhibited their open-pollinated varieties locally and emphasis was placed on uniformity and conformity to an "ideal type." By the early 1950s, essentially all of the maize grown in the Corn Belt was double-cross hybrid. After the late 1950s, increasingly more farmers in the U.S. Corn Belt grew single-cross rather than double-cross hybrids. Because single-cross seed must be produced on an inbred line, this type of selection contributed to a marked loss of variability in U.S. breeding materials. According to Goodman, a countervailing influence during the past 25 years has been the emphasis by public researchers on development of improved maize populations.

Not all scientists agree about what constitutes genetic narrowing or precisely when such narrowing has occurred. For instance, in contradiction with Porceddu *et al.* (1988), Hawkes (1983) cites the introduction of *Rht1* and *Rht2* genes into Western wheat breeding lines as an example of how diversity has been broadened by scientific plant breeders. The Japanese line Norin 10 carried the dwarfing genes from the landrace Daruma, believed to be of Korean origin. Similarly, the efforts to increase rice yields by crossing *japonica* and *indica* classes of rice extended the gene pool accessible to rice breeders. As these examples suggest, in modern agriculture, today's broadening of the genetic pool in a plant breeding program may lead to a narrowing of the breadth of materials grown by farmers precisely because such innovations often produce varieties that are popular.

Trends in Genetic Diversity of Modern Varieties of Rice, Wheat, and Maize

Part of the concern for genetic narrowing is based on the perception that, with time, conventional plant breeding practices inevitably restrict the genetic base of modern varieties. The evidence from studies on the parentage of modern varieties lends little support to the view (Witcombe, 1999). In an analysis of genealogies of 1709 modern rice varieties, Evenson and Gollin (1997) found that although a variety released in the 1960s had 3 landraces in its pedigree, recent releases have 25 or more. The complexity of rice pedigrees, in terms of parental combinations, geographical origin, and number of ancestors, has expanded over time. A similar pattern has been shown for about 800 wheat varieties released in the developing world since the 1960s (Smale, 1997). The average number of distinct landraces found in bread wheat pedigrees increased from approximately 20 in the mid-1960s to about 50 in 1990.

Skovmand and de Lacy (1999) analyzed the distance among coefficients of parentage for a historical set of CIMMYT wheat varieties during the past four decades. Their results show a rate of increase in genealogical diversity that is positive but decreases over time, with marked expansion in genealogies from 1950 to 1967 and gradual flattening through the 1990s. If progenitors were recycled and reused, the distance among them would decrease over time and the slope of the line would be negative.

Less evidence is available worldwide on trends in the pedigrees or ancestry of maize varieties than for rice and wheat, in part because this information is confidential in an increasingly privatized industry. Following the epidemic of corn blight in the U.S. crop in 1970, the National Research Council (1972) concluded that the genetic base of maize in the United States was sufficiently narrow to justify concern. Duvick (1984) found that during the 10 years following the 1970 epidemic, breeders had broadened their germplasm pools.

Molecular markers, like genealogies, can be used to construct indicators of the latent diversity in a set of crop populations. Using molecular markers, Donini *et al.* (2000) concluded that there is no objective evidence to support the assertion that modern plant breeding has reduced the genetic diversity of U.K. wheats since 1930. Recent molecular evidence for a set of CIMMYT wheats indicates that genetic distance has been maintained among major parents and popular varieties during the past 30 years (unpublished data). Since many of the varieties of spring bread wheat grown in the developing world have a combination of CIMMYT and locally bred materials in their ancestry (Heisey *et al.*, 1999), these data represent a lower bound on actual genetic diversity. Furthermore, the genetic diversity that is accessible to conventional plant breeders today includes not only spring bread wheat, of course, but also wheat types with different growing habits, close relatives, and wild grasses. Techniques of biotechnology may traverse the species barriers faced by conventional breeders.

The Commercializaton of Crop Production

What is Commercialization of Crop Production?

Economic growth, urbanization, and the withdrawal of labor from the agricultural sector have led to the increasing

commercialization of agricultural systems. Subsistence-oriented monoculture food production systems give way to a diversified market-oriented production system. Agricultural commercialization means more than the marketing of agricultural output: It means that product choice and input use decisions are based on the principles of profit maximization. Commercial reorientation of agricultural production occurs for the primary staple cereals and for the so-called high-value cash crops. Commercialization of agricultural systems leads to greater market orientation of farm production; progressive substitution of nontraded inputs in favor of purchased inputs; and the gradual decline of integrated farming systems and their replacement by specialized enterprises for crop, livestock, poultry, and aquaculture products (Pingali, 1997).

On the demand side the process of agricultural commercialization is triggered by rapid income growth and the consequent diversification in food demand patterns. A slowdown in income-induced demand for rice and for coarse grains is accompanied by a shift of diets to bread and higher valued foods such as meat, fruit, and vegetables. These dietary transitions are induced by the growth in per capita income and by the rapid migration of population to urban areas. The need to provision the rapidly growing cities of the world also acts as an impetus for the transformation of food production systems.

On the supply side, growing factor scarcities contribute to the demise of subsistence agricultural systems. Although growing land and water scarcity can be compensated for with increasing scientific knowledge and farmer management, farmer time required for sustaining productivity and profitability of intensive food production systems will become increasingly scarce. The collapse of subsistence systems will come about because of the competing demands for farmers' time. Although the speed of the structural transformation differs substantially across countries they are all moving in the same direction.

Seed Industries

As the orientation of crop production shifts from subsistence toward commercial objectives, the locus of crop improvement and seed distribution moves from individual farmers toward an organized seed industry composed of specialized private and public organizations. In terms of an increased reliance on commercially produced seed, this has occurred substantially faster for maize than for rice and wheat. In a stylized depiction of the maize seed industries in developing countries, subsistence production is characterized by open-pollinated varieties improved through farmer selection and on-farm seed production with local seed markets governed by custom. In a fully commercial system, the predominant seed type is a hybrid that is purchased annually. Seed is traded globally and is a product of specialized research that is both privately and publicly funded. The exchange of seed and the genetic resources used to improve it are enabled and protected by strict forms of intellectual property rights.

For rice and wheat, which are self-pollinating crops, the incentives for privatization of research have not always been as strong as those for maize, although this depends on the institutional and economic context. In industrialized countries, profound changes in science and in intellectual property protection during the past 20–30 years have been associated with a higher rate of investment in agriculture by the private sector than the public sector and a shift in the composition of private investment from agricultural machinery and processing into chemical research and plant breeding. Although privatization is greatest in the maize seed industry, it is also occurring in wheat to a lesser extent and particularly in Europe. There is very little private sector rice breeding anywhere in the world. Almost all the research on rice has been conducted by the public sector, and most of this has taken place in Asia. In the developing world, there is increasing privatization of the maize seed industry but rice and wheat remain primarily public.

In commercial systems of rice, wheat, and maize, recent changes in the structure of the seed industry are likely to have implications for the utilization of modern patterns of genetic variation. The global seed industry has integrated both vertically (within production processes) and horizontally (among production processes) into "life science" firms that combine seed, chemical, and pharmaceutical businesses. As part of this structural change, firms are engaging in strategies to ensure more exclusive proprietary rights, including, for the first time in history, patents on genetically modified organisms.

The implications of these changes for the exchange and utilization of the genetic resources that are used in breeding modern varieties are unknown. Efforts are under way to harmonize intellectual property regimes globally through international trade agreements, but differences between developed and developing countries, as well as among developing and developed countries, pose challenges. Small public plant breeding programs in developing countries are not on the same footing with respect to investments and legal clout as the life science conglomerates. The maize, wheat, and rice industries are likely to be affected in different ways given the nature of economic incentives associated with seed reproduction. Patents are only one type of intellectual property right; in addition to intellectual property rights, national seed regulatory systems will have a strong impact on farmers' access to seed and the conservation of plant genetic diversity.

The Implications of Commercialization for Crop Genetic Diversity

Agricultural commercialization influences the extent of crop genetic diversity in two ways: (i) through changes in land use patterns and (ii) through crop choice changes in the irrigated as well as the rain-fed environments. The organization and management of food production systems in both the irrigated and the rain-fed environments are affected by economic growth. The opportunity cost of family labor can be expected to increase equally in the high- and low-potential areas since the populations in both environments are responding to nonagricultural employment opportunities. The declining viability of subsistence production systems can also be expected to be similar. The movement from subsistence to market-oriented, rain-fed production could follow a general pattern: (i) the abandonment of highly drought-prone

environments, especially in areas where the opportunities for groundwater exploitation are limited; (ii) the shift from small subsistence farms to mechanized cultivation of large farms; and (iii) where dry season water supplies are available, increased areas under vegetables, feed grain and fodder crops, and other high-valued crops. Cereal crop production would continue to have a comparative advantage in the rain-fed environment, primarily because of the high cost of modifying the environment in order to make it suitable for noncereal crops. Low-input, low-yield cereal production systems, rice and maize, grown on consolidated holdings may emerge as the most viable option for the rain-fed environments. The irrigated environments, on the other hand, would shift from being predominantly under cereal monoculture to a highly diversified production system. The implications of these changes in crop choice on genetic diversity in the irrigated and rain-fed environments are discussed in the following sections.

Implications When Intensification has Occurred

In the irrigated rice and wheat production zones of the developing world, commercialization has had little impact on crop genetic diversity beyond that of agricultural intensification. Since the modern varieties grown there are varieties rather than hybrids, their seed is saved and spreads now, as in the past, from farmer to farmer. This will change if hybrids, or certain types of transgenic varieties that require annual seed purchase, are developed. It may also change if the demand for labor-saving technologies such as herbicides leads to the use of herbicide-tolerant varieties. These must be purchased annually to prevent the carryover of weed seed.

The effects of commercialization on maize crop genetic diversity, independent of those associated with agricultural intensification, are much more pronounced because of the historical importance of hybrids relative to improved, open-pollinated varieties. The distinctive biological properties of maize plants (in particular, their propensity for open pollination and their tendency to segregate) make it difficult for farmers to maintain the genetic purity of maize seed saved from their own harvest. Commercial maize growers are therefore dependent on reliable external sources of affordable seed in a way that growers of self-pollinated rice and wheat are not. The reliance on the seed industry will continue to grow for maize if farmer use of genetically engineered seed increases in importance in the future.

The survival of landrace diversity for cereal crops in the high-potential, irrigated environments would depend on farmer incentives for maintaining that diversity. To a large extent, farmer incentives to do so would depend on the market demand for the unique quality characteristics that are present in some of the landraces. The importance of Basmati rice in the irrigated production zones of India and Pakistan is an example of how market demand for quality can influence the survival of traditional varieties and landraces. Even where modern varieties are used exclusively, diversity within the plant has increased over time, as discussed previously, by the introduction of new gene pools through breeding.

Implications when Intensification has not Occurred

The areas of the developing world where modern varieties are not widely grown are typically marginal for production of the crop or are inadequately served by markets and infrastructure. In some areas, agricultural research has been unable to produce varieties demonstrating an obvious yield advantage or commercial seed systems have not had the incentive to do so because of the small size of the market or fluctuating effective demand.

As suggested previously, this may remain the case for sizable portions of the developing world's maize area. The proportion of cultivated area that is irrigated is far less for maize than for rice and wheat, whereas the use of purchased inputs remains modest. Most of the farmers who grow maize in developing countries face difficult and variable maize production environments and cultivate it with the primary objective of meeting subsistence requirements. These farmers have little incentive to make investments in fertilizer, pesticides, and other modern means of coping with disease and weather since their traditional varieties do not respond as well to these as modern varieties. In many of these production zones, it is not easy to breed well-adapted materials and there are few profits to be earned for seed companies.

In the most difficult environments, commercialization is likely to lead to the complete abandonment of crop production, as has already occurred in parts of Asia. When there are limited opportunities for migration but environments are too marginal for specialized agricultural production to be profitable, farmers may remain on small landholdings and grow landraces for subsistence.

Even when a zone may be suitable for the production of modern varieties, the development of commercial seed systems is not sufficient to ensure that they will replace landraces in the near future because markets are imperfect. In some local communities, the specific varietal traits demanded by farmers (grain quality, fodder, and suitability for a certain soil type) cannot be obtained through the production of modern varieties or procured through impersonal market transactions so that farmers must rely on their own production or that of nearby farmers for their supply of a valued attribute. The specialized uses of certain landrace varieties for medicinal purposes, rituals and festivals, and culinary practices have been extensively documented.

Small-scale farmers' choice to grow more than one variety simultaneously is likely to reflect their need to address numerous concerns that no single variety can satisfy (Bellon, 1996). Farmers often choose to grow both landraces and modern varieties. Zimmerer (1999) found that the capacity of farmers to grow diverse food plants (including maize) in Peru and Bolivia depends on whether they can cultivate them in combination with commercially developed, high-yielding varieties. Vaughan and Chang (1992) noted that the rapid changes that occur with natural calamities are more likely to have a greater impact on the loss of rice genetic diversity than the farmer-driven, incremental changes that are regularly occurring, many of which enhance diversity (Dennis, 1987). Meng *et al*. (1998) concluded that multiple factors, including missing markets, yield risk, grain quality, and agroclimatic constraints, influence the probability that a Turkish household

will grow a wheat landrace; a change in any single economic factor is unlikely to cause farmers to cease growing it.

Viewed in the conventional microeconomic literature as partial adoption, this observed pattern has been explained theoretically through attitudes toward risk and uncertainty, nonexisting markets, and differential soil quality or nutrient response combined with fixity or rationing (Meng, 1997; Smale *et al.*, 1997). Although treated as a transitional period to full adoption (or replacement), the coexistence of modern varieties and landraces may represent an equilibrium if one or several of these aspects persist despite economic change. Then there are locally based economic incentives for farmers to continue to grow landraces.

Even when the pressures for market integration are strong, the coexistence of modern varieties and landraces may also persist with certain types of market-based incentives, as discussed previously. In the early phases of economic growth when rural populations move to urban areas, market integration exerts pressures for uniformity in the attributes of coarse grains. Localized preferences diminish in favor of cheaper, bulk-marketed grains. The elasticity of demand for staple grains declines as income increases, and it is sometimes negative. Generally, rice substitutes for maize, and wheat substitutes for rice.

The income elasticity of demand for attributes of the grains may be higher, however, than the income elasticity of demand for the cereal (Pingali *et al.*, 1997). For example, a notable pattern of rice consumption is that, with growing incomes, people express preferences for higher quality rice once their calorie needs have been met. High-income consumers spend more on rice by paying higher prices for varieties with preferred eating quality which they substitute for the lower quality variety consumed when their income levels were lower. In Asia, traditional varieties are generally of higher quality and fetch premium prices in the market. Thailand still grows low-yielding traditional rain-fed varieties extensively for the export market. When the income level was low, South Korea used to grow the modern ''tongil'' variety, but this was replaced by relatively low-yielding traditional *japonica* rice as consumers expressed preference for *japonicas* by offering higher prices. In response to these market signals farmers are eager to grow even low-yielding, high-quality rices because the higher prices more than compensate for their lower yields. Because rice scientists have had limited success in developing high-yielding cultivars with better eating quality, the price difference between the standard- and high-quality varieties has been increasing in Asian markets.

The post-industrial agricultural economy is characterized by growth in demand for an array of increasingly specialized goods and services. Some product quality attributes are associated with features of the production process (organic/inorganic). Some are extrinsic (origin or effects on animal welfare), whereas others cannot be discerned without laboratory tests (genetically modified organisms). The elasticity of demand for such attributes is likely to increase with very high levels of income. Under these conditions, global market integration may provide market-based incentives for continued cultivation of patches of diverse landraces. In addition to their demonstrated private economic value to the farmers who grow them, some of these hold potential for niche markets and exports.

Conclusions

In areas of the world that are more favorable for agricultural intensification, the conversion from landrace varieties to modern varieties has been almost complete for rice, wheat, and maize. In some of the remaining crop-growing environments, and especially in relatively small pockets of crop diversity called microcenters, farmers who are linked to commercial agriculture through labor markets or specialized product markets continue to grow landraces—often in combination with modern varieties. Some cannot obtain or afford to purchase seed on a routine basis, so they both purchase and save the seed of modern types. Others purposefully adapt modern varieties to their own conditions by saving and selecting the seed, genetically integrating modern and landrace types.

The structure of genetic diversity is distinct for modern varieties and landraces. Both are essential to the future food supply. Conservationists propose that landrace diversity must be maintained not only in preserved stocks called *ex situ* collections but also *in situ*, or in the fields of farmers. The future of landrace cultivation appears to be uncertain. Although some view the replacement of landraces by modern varieties as an inevitable product of agricultural commercialization, idiosyncratic growing environments and consumer preferences may provide economic incentives for their continued cultivation by farmers—although on a limited scale. It appears unlikely that modern varieties of rice, wheat, and maize will entirely replace landraces in the near future, although it is difficult to postulate about equilibrium areas planted with each type of germplasm since the equilibrium itself shifts with technical and economic change.

Continued genetic improvement does not necessarily lead to loss of genetic diversity in areas where modern varieties dominate, especially when access to germplasm is relatively unrestricted and innovative plant breeding strategies may be employed. Access to diverse sources of germplasm is therefore of great importance to the success of public and private breeding programs for the supply of varieties in modern agriculture. The continued advances in yield potential that are a necessary (although not a sufficient) condition for alleviating hunger are thought to depend on increasingly complex combinations of genes and novel alleles. Landraces and wild relatives have served as repositories for resistance to biotic and abiotic stress when these are absent in advanced breeding materials. Even in the parts of the world where the ''ancient patterns of diversity'' may still be found, access to the products of modern plant breeding is often integrated economically or genetically to generate more resilient and sustainable systems.

See also: Agriculture, Sustainable. Agriculture, Traditional. Genetic Diversity. Herbicides. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Pesticides, Uses and Effects of

References

Bell GDH (1987) The history of wheat cultivation. In: Lupton FGH (ed.) *Wheat Breeding: Its Scientific Basis*. London: Chapman & Hall.

Bellon MR (1996) The dynamics of crop infraspecific diversity: A conceptual framework at the farmer level. *Econ. Bot.* 50: 26–39.

Brown AHD, Clegg MT, Kahler AL, and Weir BS (eds.) (1990) *Plant Population Genetics: Breeding and Genetic Resources*. Sunderland, MA: Sinauer.

Brush SB (1995) *In situ* conservation of landraces in centers of crop diversity. *Crop Sci.* 35: 346–354.

CIMMYT Economics Program (1999) Global Maize and Wheat Impacts Surveys, 1992, 1997. Mexico, D.F.: CIMMYT

Dennis JV (1987) *Farmer management of rice variety diversity in Northern Thailand Ph.D. thesis*. Ithaca, NY: Cornell University.

Donini P, Law JR, Koebner RMD, Reeves JC, and Cooke RJ (2000) Temporal trends in the diversity of UK wheat. *Theor. Appl. Genet.* 100: 912–917.

Duvick DN (1984) Genetic diversity in major farm crops on the in press - must have been published by now/farm and in reserve. *Econ. Bot.* 38(2): 161–178.

Evenson RE and Gollin D (1997) Genetic resources, international organizations, and improvement in rice varieties. *Econ. Dev. Cultural Change* 45(3): 471–500.

Frankel OH (1970) Genetic dangers of the Green Revolution. *World Agric.* 19: 9–14.

Goodman MM (1995) Maize. In: Smartt J and Simmonds NW (eds.) *The Evolution of Crop Plants*, pp. 193–202. New York: Wiley.

Harlan JR (1992) *Crops and Man*. American Society of Agronomy/Crop Science Society of America. WI: Madison.

Hawkes JG (1983) *The Diversity of Crop Plants*. Cambridge, MA: Harvard Univ. Press.

Heisey PW, Lantican M, and Dubin HJ (1999) Assessing the benefits of international wheat breeding research: An overview of the global wheat impacts study. In: Pingal PL (ed.) *CIMMYT 1998–99 World, Wheat Facts and Trends. Global Wheat Research in a Changing World: Challenges and Achievements*. Mexico City, Mexico: CIMMYT.

IRRI (1995) World Rice Statistics, 1993–1994 Manila, Philippines: IRRI.

IRRI (1999) World Rice Statistics, 1998–1999 Manila, Philippines: IRRI.

Meng ECH (1997) *Land Allocation Decisions and In Situ Conservation of Crop Genetic Resources: The Case of Wheat Landraces in Turkey*. Ph.D. thesis. University of California: Davis.

Meng ECH, Taylor JE, and Brush SB (1998) Implications for the conservation of wheat landraces in Turkey from a household model of varietal choice. In: Smale M (ed.) *Farmers, Gene Banks and Crop Breeding: Economic Analyses of Diversity in Wheat, Maize, and Rice*, pp. 127–143. Boston: Kluwer/CIMMYT.

Morris ML and López-Pereira MA (1999) *Impacts of Maize Breeding Research in Latin America, 1966–67*. Mexico City, Mexico: CIMMYT.

National Research Council (1972) *Genetic Vulnerability of Major Crops*. Washington, D.C.: National Academy of Sciences.

Oka HI (1988) *Origin of Cultivated Rice*. Tokyo: Japan Scientific Societies Press/ Elsevier.

Pingali PL (1997) From subsistence to commercial production systems: The transformation of Asian agriculture. *Am. J. Agric. Econ.* 79: 628–634.

Pingali PL, Hossain M, and Gerpacio RV (1997) *Asian Rice Bowls: The Returning Crisis?* Wallingford, UK: CAB International.

Porceddu E, Ceoloni C, Lafiandra D, Tanzarella OA and Scarascia Mugnozza GT (1988) Genetic resources and plant breeding: Problems and prospects, In *The Plant Breeding International, Cambridge Special Lecture*, pp. 7–21. Cambridge, UK: Institute of Plant Science Research.

Skovmand B and de Lacy IH (1999) Parentage of a historical set of CIMMYT wheats. 1999 Annual Meeting Abstracts. Madison, WI: American Society of Agronomy.

Smale M (ed.) (1998) *Farmers, Gene Banks and Crop Breeding Economic Analyses of Diversity in Wheat, Maize, and Rice*. Mexico/Norwell, MA: CIMMYT/Kluwer.

Smale M, Just RE, and Leathers HD (1994) Land Allocation in HYV Adoption Models: An Investigation of Alternative Explanations. *Am. J. Ag. Econ.* 76: 535–546.

Vaughan D and Chang TT (1992) *In situ* conservation of rice genetic resources. *Econ. Bot.* 46: 369–383.

Witcombe JR (1999) Does plant breeding lead to a loss of genetic diversity? In: Wood D and Lenné J (eds.) *Agrobiodiversity: Characterization, Utilization, and Measurement*. Wallingford, UK: CAB International.

Wood D and Lenné J (1997) The conservation of agrobiodiversity on-farm: questioning the emerging paradigm. *Biodiversity and Cons.* 6: 109–129.

Zimmerer KS (1999) Overlapping patchworks of mountain agriculture in Peru and Bolivia: Toward a regional–global landscape model. *Human Ecol.* 27: 135–165.

Agriculture, Nutrient Management, and Water Quality

Andrew N Sharpley, University of Arkansas, Fayetteville, AR, USA

Published by Elsevier Inc.

Glossary

Beneficial management practices (BMPs) BMPs are effective, practical, structural, or nonstructural methods that are designed to prevent or reduce the movement of sediment, nutrients, pesticides, and other chemical contaminants from the land to surface or ground water, or that otherwise protect water quality from the potential adverse effects of agricultural activities. These practices are developed to achieve a cost-effective balance between water quality protection and the agricultural production (e.g., crop, forage, animal, and forest).

Concentrated animal feeding operations (CAFOs) A CAFO is an animal feeding facility that confines animals for more than 45 day in an area that does not produce vegetation during the growing season.

Conservation practices Any action taken to produce environmental improvements, particularly with respect to agricultural nonpoint-source emissions. The term is used broadly to refer to structural practices, such as buffers, as well as nonstructural practices, such as in-field nutrient management planning and application. Conservation Practice standards have been developed by NRCS and are available at: http://www.nrcs.usda.gov/Technical/Standards/nhcp.html

Edge-of-field nutrient loss A term that refers to the nitrogen and phosphorus that is lost or exported from fields in agricultural production.

Eutrophication An increase in the rate of supply of organic matter to an ecosystem.

Nonpoint source A diffuse source of chemical and/or nutrient inputs not attributable to any single discharge (e.g., agricultural runoff, urban runoff, and atmospheric deposition).

Nutrients Inorganic chemicals (particularly nitrogen, phosphorus, and silicon) required for the growth of animals and plants, including crops and phytoplankton.

Point sources Readily identifiable inputs where treated wastes are discharged from municipal, industrial, and agricultural facilities to the receiving waters through a pipe or a drain.

Water quality standards Risk-based requirements, which set site-specific allowable pollutant levels for individual water bodies, such as rivers, lakes, streams, and wetlands. States set water quality standards by designating uses for the water body (e.g., recreation, water supply, aquatic life, and agriculture) and applying water quality criteria (numeric pollutant concentrations and narrative requirements) to protect the designated uses.

Water-use designation The water quality standards regulation requires that States and authorized Indian Tribes specify appropriate water uses to be achieved and protected. Appropriate uses are identified by taking into consideration the use and value of the water body for public water supply, for protection of fish, shellfish, and wildlife, and for recreational, agricultural, industrial, and navigational purposes. In designating uses for a water body, States and Tribes examine the suitability of a water body for the uses based on the physical, chemical, and biological characteristics of the water body, its geographical setting and scenic qualities, and economic considerations.

Watershed The drainage basin contributing water, organic matter, dissolved nutrients, and sediments to a stream or a lake.

Introduction

Since the late 1960s, point sources of water pollution have been reduced in the US due to their ease of identification and the passage of the Clean Water Act in 1972. However, water quality problems remain, and as further point-source control becomes less cost-effective, attention is being directed toward the role of agricultural nonpoint sources in water quality degradation. In a 1996 U.S. Environmental Protection Agency (EPA) report, nonpoint-source nutrients were the primary source of concern in 40% of rivers, 50% of lakes, and 60% of estuaries surveyed and listed as impaired (U.S. Environmental Protection Agency, 1996). As in the US, the European Union Water Framework Directive (WFD) (Council of European Communities, 2000) now requires widespread control of N and P inputs to rivers specifically to improve riverine ecology (Hilton *et al.*, 2006).

Besides soil and pesticide loss from agriculture, most environmental concerns center on nonpoint transport of the nutrients phosphorus (P) and nitrogen (N), which are essential inputs for optimum crop and animal production. However, dramatic improvements in the economic efficiency of grain and animal production over the past 50 years have been coupled with increasing surface and ground water quality problems (Boesch *et al.*, 2001; Rabalais *et al.*, 2001). Owing to its much lower mobility in soil, P concerns focus mainly on its transport in surface runoff, whereas N concerns revolve around nitrate movement through the soil profile to ground water (**Figure 1**).

Phosphorus inputs to fresh waters can accelerate eutrophication (Carpenter *et al.*, 1998; Schindler, 1977). Although nitrogen and carbon (C) are also essential to the growth of aquatic biota, most attention has focused on P because of the difficulty in controlling the exchange of N and C between the atmosphere and water, and fixation of atmospheric N by some blue-green algae. Thus, control of P inputs is critical to reducing freshwater eutrophication. For water bodies with a naturally higher salt content, as in estuaries, there

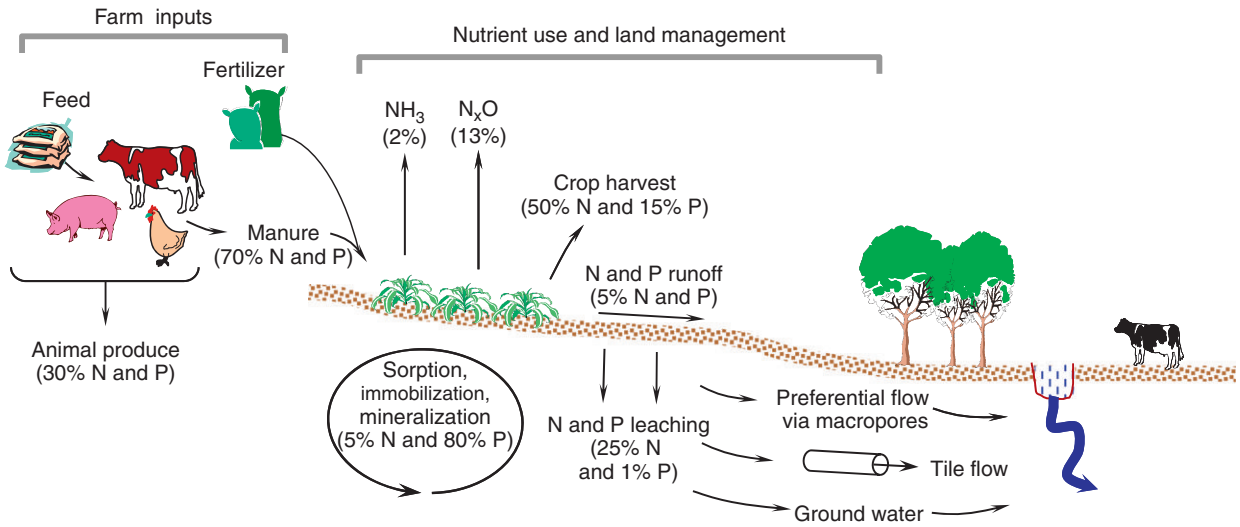


Figure 1 Factors affecting the fate of N and P in agriculture, representing their potential impact on soil, water, and air quality. Numbers in parentheses are based on an approximate farm nutrient balance and the relative fate of N and P as a percentage of load (Farm inputs) or percentage of fertilizer and manure (Nutrient use and land management). Adapted from Bouwman AF and Booij H (1998) Global use and trade of feedstuffs and consequences for the nitrogen cycle. *Nutrient Cycling in Agroecosystems* 52: 261–267; Duxbury JM, Harper LA, and Mosier AR (1993) Contributions of agroecosystems to global climate change. In: Harper LA, Mosier AR, Dubury JM, and Rolston DE (eds.) *Agroecosystem Effects in Radiatively Important Trace Gases and Global Climate Change*. Special Publication 55, pp. 1–18. Madison: American Society of Agronom; Howarth RW, Anderson DA, Church TM, *et al.* (2000) *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. Washington, DC: National Research Council. National Academy Press, 405 pp, and Sims JT and Sharpley AN (eds.) (2005) *Phosphorus: Agriculture and the Environment*. American Society of Agronomy Monograph. Madison, WI: American Society of Agronomy.

are likely unique site-specific critical concentrations of both N and P that generally limit aquatic productivity (Howarth *et al.*, 2000).

Eutrophication has been identified as the main problem in surface waters with impaired water quality (U.S. Environmental Protection Agency, 2002). Eutrophication restricts water use for fisheries, recreation, industry, and drinking, due to the increased growth of undesirable algae and aquatic weeds and oxygen shortages caused by their death and decomposition. Also, many drinking water supplies throughout the world experience periodic massive surface blooms of cyanobacteria (Kotak *et al.*, 1993). These blooms contribute to a wide range of water-related problems including summer fish kills, unpalatability of drinking water, and formation of trihalomethane during water chlorination (Kotak *et al.*, 1994; Palmstrom *et al.*, 1988). Consumption of cyanbacterial blooms or water-soluble neuro- and hepatoxins released when these blooms die can kill livestock and may pose a serious health hazard to humans (Lawton and Codd, 1991; Martin and Cooke, 1994). Recent outbreaks of the dinoflagellate *Pfiesteria piscicida* in the eastern US coastal waters have been linked to excess nutrients in affected waters (Burkholder and Glasgow, 1997). Neurological damage in people exposed to the highly toxic volatile chemical produced by this dinoflagellate has dramatically increased public awareness of eutrophication and the need for solutions.

Nitrate is a water quality concern because it has been linked to methemoglobinemia in infants, to toxicities in animals, and to increased eutrophication in both fresh and saline (e.g., estuaries) waters (Amdur *et al.*, 1991; Sandstedt, 1990). The EPA has established a maximum contaminant level for nitrate-N in drinking water of 10 mg l^{-1} (45 mg nitrate per liter) to protect babies under 3–6 months of age (U.S. Environmental Protection Agency, 2002). This segment of the population is most sensitive because bacteria that live in an infant's digestive tract can reduce nitrate to nitrite, thereby causing hemoglobin to transform into methemoglobin, which interferes with the oxygen-carrying ability of blood (Amdur *et al.*, 1991). Under anaerobic soil conditions, denitrifying bacteria readily convert excess nitrate into N gases (primarily N_2), reducing the quantity of nitrate that can potentially leach to ground water supplies. However, the production of N_xO gases contributes to the accumulation of greenhouse gases in the atmosphere (**Figure 1**).

To address these problems with drinking water treatment, some areas, such as New York State, now treat the cause of eutrophication rather than its effects. In the early 1990s, New York City decided that remediation of the sources nutrients in its water supply watersheds was more economical than building a new $8-billion water treatment facility. For agricultural land, progress has been made in developing and implementing farm-specific management plans and best management practices. Other nutrient management programs mandated include storm water management, septic system rehabilitation and replacement, and forestry management programs.

Research has identified agricultural land-use practices that are of highest risk for nutrient loss at field and farm scales, and the forms in which N and P are exported (Sharpley *et al.*, 2005; Withers and Lord, 2002). Their impact on stream water quality at a watershed scale is less well defined because field and farm inputs are modified by watershed hydrology, in-stream processing, and contributions from a diverse range of other rural nutrient sources, including farmyard and septic tanks runoff

and rural sewage treatment works (Jarvie *et al.*, 2006 Gburek and Sharpley, 1998). In response, new requirements to improve the ecological status of rivers focus attention on identifying nutrient sources and their ecological impacts in rivers.

Phosphorus and N are typically treated separately by scientists and environmental managers, but such a separation is artificial as P and N occur simultaneously in watersheds and farmers manage them together.

Water Quality Criteria

Guidelines for management within agricultural catchments are based on definitions of impairment and acceptable conditions. U.S. EPA established the National Regional Nutrient Criteria Program (**Table 1**; U.S. Environmental Protection Agency, 2001). Similarly, in Europe, the EU WFD requires the setting of "reference conditions" for different water body types to define how far they have become impaired (Council of European Communities, 2000). Reference condition is defined as "no or only very minor alterations" resulting from human activities to the water body's physiochemical, biological, and hydrological properties.

In the US, baseline or reference conditions of pristine waters are defined by monitoring total P, total N, chlorophyll-*a*, and clarity in waters draining areas with minimal human impact (Gibson *et al.*, 2000). Because it can be argued that most, if not all, lakes have been impacted by human activity to some degree, reference conditions represent the least impacted conditions or what is considered to be most attainable. The following guidelines are used to define minimally impacted lakes: catchments with $<1\%$ urban land use, $>65\%$ of lakeshore that has at least 10 m of natural bank-side vegetation, no point source discharge inputs, and there is no control of water-level fluctuations (Gibson *et al.*, 2000). Baseline concentrations of total N and P for freshwater systems in the continental US are shown in **Table 1**. The difference between baseline and current N and P levels is used to determine and target reductions required in a catchment to alleviate water quality impairment. In the EU, countries have governmental powers to implement the WFD and water quality goals are achieved through river basin management planning.

Even so, water quality concerns associated with agriculture have arisen at a catchment rather than farm scale, reflecting the cumulative impacts of individual nutrient and land management activities on farms. In many areas, agriculture's contribution to water quality impairment has been, or is, masked by nonagricultural nonpoint and point sources to varying degrees. This raises issues of scale (both spatial and temporal), difficulties attributing cause and effect to ensure adaptive management, and the need to consider interactions with other nutrient sources, which must all be taken into account when developing catchment strategies for agricultural management that sustain productive viability and maintain water quality.

The Evolution of Agriculture

After World War II, fertilizer production and distribution became cheaper. Between 1950 and 1990, N use in the US increased fourfold (8.5 million tons) and P use two fold (0.9 million tons) and agricultural production more than doubled (Evans *et al.*, 1996). US farm land decreased from 500 to 400 million ha (20%) and the number of farms from 5.6 to 2.1 million (63%), whereas the average farm size has increased from 90 to 200 ha (120%). There was also a switch from crop- to livestock-based systems in states such as Alabama, Arkansas, Delaware, North Carolina, and Texas (Kellogg *et al.*, 2000;

Table 1 Baseline N and P concentrations for each of the aggregated nutrient Ecoregions in the US for freshwater systems

Aggregated ecoregion		*Total N (mg l^{-1})*		*Total P ($\mu g\ l^{-1}$)*	
Region	*Description*	*Rivers and streams*	*Lakes and reservoirs*	*Rivers and streams*	*Lakes and reservoirs*
I	Willamette and central valleys	0.31	–	47	–
II	Western forested mountains	0.12	0.10	10	9
III	Xeric west	0.38	0.40	22	17
IV	Great Plain Grass and Shrub Lands	0.56	0.44	23	20
V	South Central cultivated Great Plains	0.88	0.56	67	33
VI	Corn Belt and northern Great Plains	2.18	0.78	76	38
VII	Mostly glaciated dairy region	0.54	0.66	33	15
VIII	Nutrient poor largely glaciated upper Midwest and Northeast	0.38	0.24	10	8
IX	Southeastern temperate forested plains and hills	0.69	0.36	37	20
X	Texas-Louisiana coastal and Mississippi alluvial plain	0.76	–	128[a]	–
XI	Central and eastern forested uplands	0.31	0.46	10	8
XII	Southern coastal plains	0.91	0.52	40	10
XIII	Southern Florida coastal plains	–	1.27	–	18
XIV	Eastern coastal plains	0.71	0.32	31	8

[a]This high value may either be a statistical anomaly or reflects a unique condition.

Source: Adapted from U.S. Environmental Protection Agency (2001) *Ecoregional Nutrient Criteria*. EPA-822-F-01-010. USEPA, Office of Water (4304). Washington, DC: U.S. Government Printing Office, http://www.epa.gov/waterscience/criteria/nutrient/9docfs.pdf.

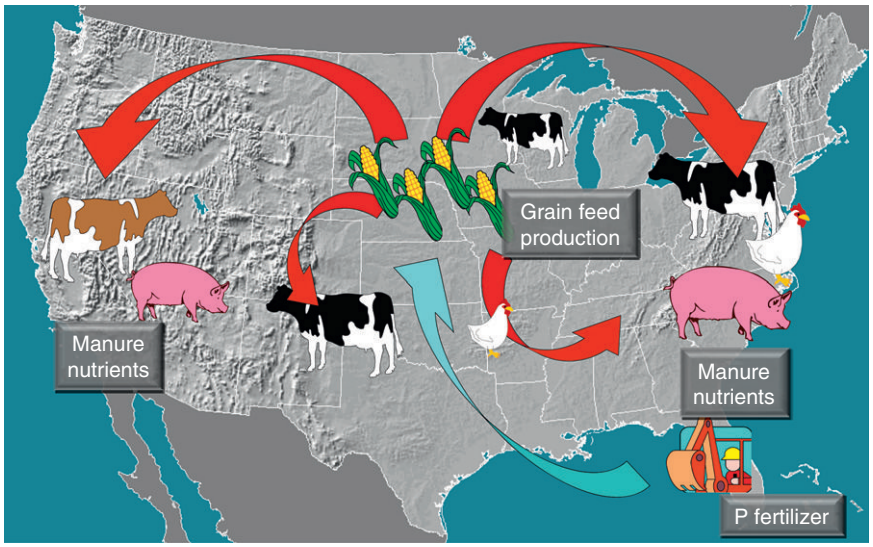


Figure 2 Grain and animal production systems have evolved in spatially separate areas leading to localized accumulations of nutrients in manure.

Lanyon, 2000). Over the past 10 years, the number of dairies has decreased by 40%, but herd size has increased by 50%, whereas the number of hogs has increased 18% on 72% fewer farms. Similarly, 97% of poultry production comes from farms with more than 100,000 birds and over a third of beef production from <2% of the feedlots. As a result of this regional intensification, modern farming systems have become fragmented, with increasing separation of crop and livestock production, even across regional boundaries (Lanyon, 2005). Similar trends occurred in Europe (Withers *et al.*, 2002).

As a consequence of the spatial separation of crop and animal production systems, fertilizers are produced (i.e., N) or imported (i.e., P) to areas of grain production (**Figure 2**). The grain and harvested N and P are then transported to areas of animal production, when inefficient animal utilization of nutrients in feed (<30% utilized) is excreted as manure. This has led to a large-scale, one-way transfer of nutrients from grain- to animal-producing areas and dramatically broadened the emphasis of nutrient management strategies from field and watershed levels to national scales (**Figure 2**; Lander *et al.*, 1998; Lanyon, 2005; Sharpley *et al.*, 2005). In the western US, irrigated lands, animal production, and grain production areas are still contiguous; however, large-scale consolidation has created much larger concentrated animal feeding operations (CAFOs), which makes economical redistribution of manure to croplands difficult (Kellogg *et al.*, 2000).

Managing Nutrients in Agriculture

The Fate of N and P in Agriculture

The fate of N and P in agriculture is depicted in **Figure 1**. Approximate percentage inputs of N and P, as feed or fertilizer, and their respective fate when fed to animals or applied to land, were obtained from several studies, as summarized below.

Nitrogen

About 30% of feed N is accounted for in animal products sold, with the remainder excreted in manure and available for land application. In rain-fed systems, an average of 40–60% of the N applied in fertilizer and manure is removed in crop harvest (**Figure 1**; Howarth *et al.*, 2000). Of the remainder, some is stored as organic N in the soil, some is volatilized to the atmosphere, and some leaches to ground and surface waters. A variety of factors, including soil properties, climate, fertilizer, manure, and land management, influence the fate of applied N (Howarth *et al.*, 1996). For typical rain-fed farming in the US, roughly 15% of field-applied N is volatilized to the atmosphere as ammonia and N_xO (**Figure 1**; Bouwman and Booij, 1998). This N is deposited back onto the landscape, often near the source of volatilization, although sometimes traveling long distances within an airshed (Holland *et al.*, 1999). Again, for typical farming systems in the US, the percentage of applied N that leaches to ground and surface waters varies between 10% and 40% for loam and clay soils, and 25% and 80% for sandy soils (Howarth *et al.*, 1996).

As a consequence of irrigation in the western US, the nitrate concentration has increased in shallow aquifers underlying major grain and livestock production areas (CAST, 1985; Powers and Schepers, 1989). Corn is the primary crop that receives most of the N fertilizer and manure. Most of the leaching losses occur during the nongrowing season because rainfall and irrigation during the peak growing season often just meet crop water requirements. If excess irrigation results in too high a soil moisture content at the end of the season, any nitrate remaining in the soil may be leached during the following spring and fall when rainfall often exceeds evapotranspiration demands (Hergert, 1986; Klocke *et al.*, 1999).

Phosphorus

Less than a third of feed P is utilized by animals, with the remainder excreted in manure and applied to land for crop use. Several soil and crop factors, including soil P sorption

capacity, crop type, P application type, method and rate, and land management, influence plant P uptake (Pierzynski and Logan, 1993; Sims and Sharpley, 2005). Phosphorus loss by leaching is generally greater in sandy, organic, or peaty soils – those with low-P adsorption capacities, and in soils with substantial preferential flow pathways (Bengtson *et al.*, 1988; Sharpley and Syers, 1979; Sims *et al.*, 1998) (**Figure 1**). Phosphorus losses from tile-drained fine-textured soils receiving manure are generally low (Hergert *et al.*, 1981). Soils that allow substantial subsurface export of dissolved P are common in parts of the northeast Coastal Plains region of the Delmarva Peninsula, and are important to consider in minimizing coastal water eutrophication in these regions. However, P loss in surface runoff is generally greater than in subsurface flow, depending on the rate, time, and method of P application; form of fertilizer or manure applied; amount and time of rainfall after application; and land cover (Sharpley and Rekolainen, 1997).

Phosphorus losses are often agronomically small (generally <2 kg P per hectare), representing a minor proportion of applied fertilizer or manure P (generally $<5\%$). Phosphorus uptake and harvest removal by crops ranges from 10% to 40% of applied P (**Figure 1**), due to the rapid and only slowly reversible sorption of P to Al, Fe, and Ca compounds in soil.

Farm Inputs and Nutrient Balances

The potential for P and N surplus at the farm scale can be much greater in CAFO than in cropping systems when nutrient inputs become dominated by feed rather than fertilizer (**Table 2**). A CAFO is defined as an animal feeding facility that confines animals for more than 45 days in an area that does not produce vegetation during the growing season (see http://www.epa.gov/region07/water/cafo/). With a greater reliance on imported feeds, only 30% of P and 55% of N in purchased feed for a 74,000 layer operation on a 12-ha farm in Pennsylvania could be accounted for in farm outputs (**Table 2**). These nutrient budgets clearly show that the largest input of nutrients to CAFOs, and thus the primary source of any on-farm nutrient excess, is in animal feed. Similar annual surpluses of N and P were reported for grain-poultry farms in Delaware (Sims, 1997) and grain-dairy farms in Wisconsin (Bundy and Kelling, 2001; Powell *et al.*, 2002) and New York (Klausner *et al.*, 1998). Results of US census surveys over the past 15 years show a steady increase in the amount of N and P that exceeds crop requirements at a farm level for all livestock operations (Kellogg *et al.*, 2000). Thus, many of the water quality concerns we now face are a result of this imbalance in system inputs and outputs of P and N, caused by the intensification of CAFOs.

Regional Nutrient Imbalances

The already difficult nutrient management implications of this spatial separation and intensification of grain and animal production in farming systems are exacerbated by the localization of the problem to specific areas (**Figure 2**). The increasing livestock density in certain regions of the US is due, in part, to strategic planning by the animal industry and market integrators, economic pressures, and consumer demands, all of which are external to individual farmer decisions. This has led to an accumulation of manure N and P in certain regions. Kellogg *et al.* (2000) estimated the amount of manure produced from animal distribution numbers from the 1997 US Census of Agriculture (**Figure 3**), using standard values of manure production and nutrient concentration for each animal type. Estimates of excess N and P were calculated based on crop N and P removal (Lander *et al.*, 1998) and the

Table 2 Farming system and nutrient budget

Farming system	*Nutrient input in*		*Output in produce (kg ha^{-1} year^{-1})*	*Surplus (kg ha^{-1} year^{-1})*
	Feed (kg ha^{-1} year^{-1})	*Fertilizer (kg ha^{-1} year^{-1})*		
Nitrogen budget				
Cash crop[a]	–	246	238	8
Dairy[b]	155	40	75	120
Hog[c]	390	10	120	280
Poultry[d]	5800	–	1990	3810
Phosphorus budget				
Cash crop[a]	–	59	38	21
Dairy[b]	30	11	15	26
Hog[c]	105	–	30	75
Poultry[d]	1560	–	440	1120

[a]30 ha cash crop farm growing corn, alfalfa, and hay; "output in produce" includes nitrogen fixed by alfalfa (79 kg nitrogen per hectare per year) not otherwise tallied in the nutrient inputs.

[b]40 ha farm with 65 dairy Holsteins averaging 6600 kg milk per cow per year, 5 dry cows, and 35 heifers. Crops were corn for silage and grain, alfalfa, and rye for forage.

[c]30 ha farm with 1280 hogs; surplus includes 36 kg phosphorus and 140 kg nitrogen per hectare per year manure exported from the farm.

[d]12 ha farm with 74,000 poultry layers; surplus includes 180 kg phosphorus and 720 kg nitrogen per hectare per year manure exported from the farm.

Source: Data adapted from Lanyon LE (2000) Nutrient management: Regional issues affecting the Chesapeake Bay. In: Sharpley AN (ed.) *Agriculture and Phosphorus Management: The Chesapeake Bay*, pp. 145–158. Boca Raton, FL: CRC Press, and Bacon SC, Lanyon LE, and Schlauder Jr. RM (1990) Plant nutrient flow in the managed pathways of an intensive dairy farm. *Agronomy Journal* 82: 755–761.

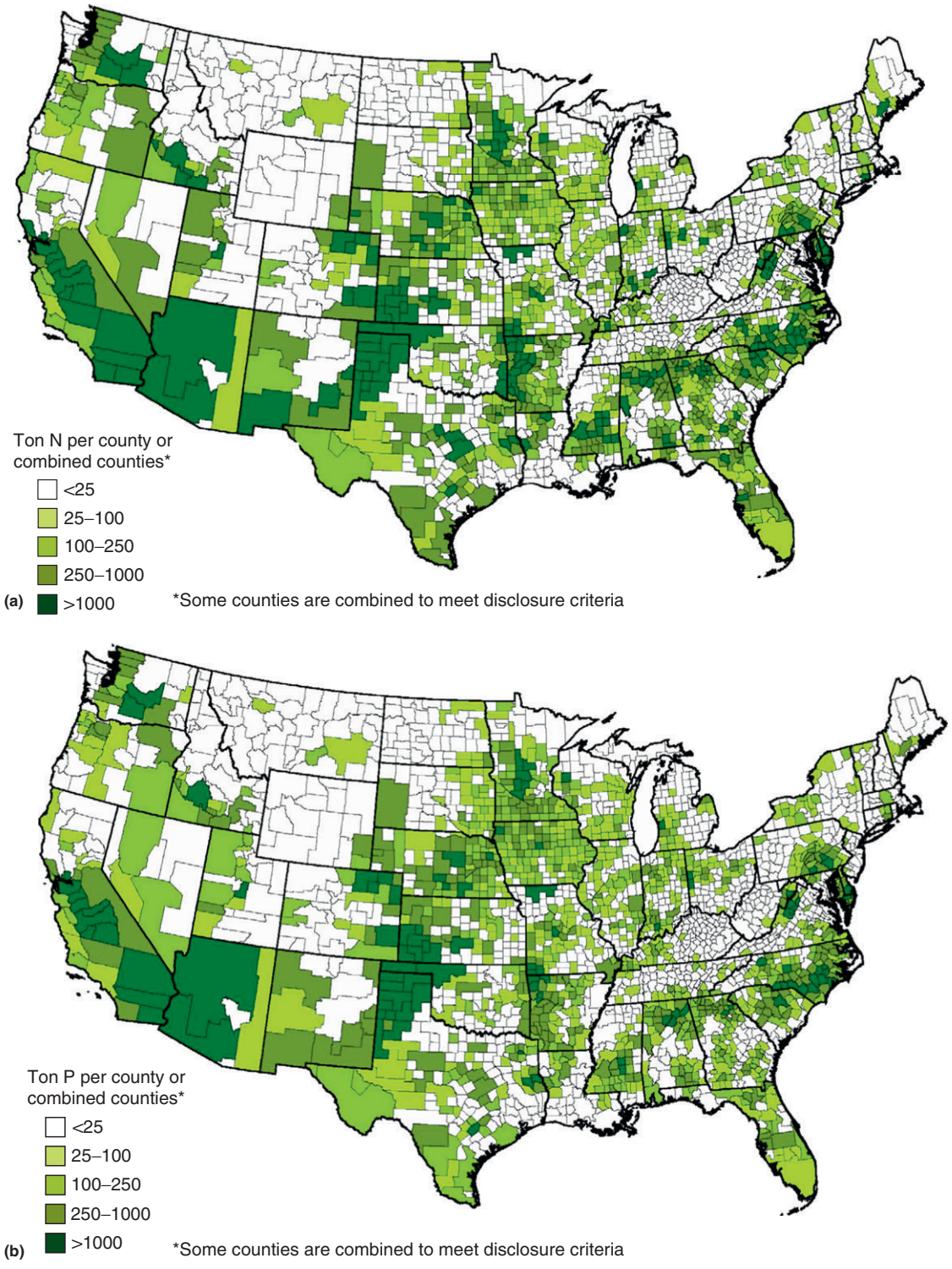


Figure 3 (a) Excess manure N assuming no export of manure from the farm using 1997 census data. (b) Excess manure P assuming no export of manure from the farm, using 1997 census data.

assumption that all suitable crop and pasture land was available for manure application (**Figure 3**; Kellogg *et al.*, 2000). Most areas with CAFOs have some excess N (**Figure 3(a)**) and P (**Figure 3(b)**). The regions of greatest excess are generally in animal production areas of the southern tier states. With the exception of the Midwest and central Great Plains, where P excesses were substantially greater than N, most regions had similar amounts of excess N and P in manure (**Figure 3**).

The significance of these regional N and P accumulations is highlighted by their close proximity to nutrient-sensitive waters in these regions (e.g., Neuse River, North Carolina; Chesapeake Bay; Great Lakes; Lake Champlain; Gulf of Mexico; Illinois River water supply, Arkansas and Oklahoma; and North Bosque River, Texas). Kellogg (2001) combined an assessment of the excess amounts of N and P in manure relative to potential crop assimilation (**Figure 4**), with environmental factors governing their potential for loss to surface and ground water; these factors included rainfall and the leaching, runoff, and erosion potential of area soils. The resulting nitrate leaching and P vulnerability indices are presented in **Figure 4(a)** and **4(b)**, respectively. This approach allowed Kellogg (2001) to identify potential priority watersheds for water quality protection, where both high manure nutrients can be applied to the land and environmental indicator scores are higher than other areas of the country. These would be the watersheds where government programs could be targeted first to quickly meet the goals of protecting watersheds from contamination by manure nutrients. The "environmental setting" is not addressed by these vulnerability indexes. For example, some areas of the country, such as the Chesapeake Bay and Florida, have a greater potential for use impairment than other areas of the country because of the high demand for use of the water resource for recreation, commercial activities, shoreline housing, drinking water, and general esthetics. Even so, while these maps were developed a decade ago, priority areas are still the same and the same challenges remain.

In the Great Plains, desirable topography and climate are the primary reasons contributing to a major increase in cattle feeding during the past 20 years (Kellogg *et al.*, 2000). A large portion of the grain production in this region occurs on irrigated lands that rely on ground water sources. Much of the intermountain west and the far west rely on surface water from reservoirs, but there is also substantial irrigation from ground water (Spalding and Exner, 1993). A major impact from concentrated livestock operations in these regions is nitrate enrichment of ground water. Because these regions have typically low rainfall amounts, surface runoff is limited, and denitrification, which decreases the impact of nitrate available for enrichment of ground water, is less of a factor than it is in the humid east. Instances of high nitrate in ground water, of which a part of the source can be traced to manure, include Nebraska, Minnesota, Kansas, Colorado, Arizona, New Mexico, Texas, and California.

Remedial Beneficial Management Practices (BMPs)

There are many BMPs that can be implemented over a wide range of scales to minimize the loss of N and P from agriculture to surface and ground waters (**Table 3**). These are commonly grouped into measures that seek to reduce the input of N and P onto farms and ring them into closer balance with outputs in produce; manage on-farm nutrient sources through appropriate rate, timing, and method of N and P application; and measure the potential for N and P transport to surface and ground waters (**Table 3**). These measures are also depicted in **Figure 5**.

Farm Input Decisions

Approximately 70% of the N and P fed to animals is excreted (**Figure 1**; Lanyon, 2005); thus, managing animal nutritional requirements is critical to on-farm N and P management. However, several surveys conducted in the Northeast (Tylutki and Fox, 2000), Midsouth States (Sansinena *et al.*, 1999), and Wisconsin (Powell *et al.*, 2001) indicate that P was fed 30–40% above the National Research Council (2001) dietary guidelines of about 0.36% P. There are several reasons for this overfeeding of P, which include rations formulated with wide "safety margins" or "insurance" levels of P, a lack of ration formulation based on animal groups, and relatively inexpensive P supplements (CAST, 2002; Dou *et al.*, 2001).

Carefully matching dietary P inputs to animal requirements can reduce the amount of P excreted by animals (Poulsen, 2000; Valk *et al.*, 2000). For instance, in a survey of Wisconsin dairy farms, Powell *et al.* (2001) found that decreasing dietary P level from an excessive 0.55% P to the National Research Council (NRC) level of 0.36% P would reduce the number of farms and acreage with an excess P balance by approximately two-thirds (Powell *et al.*, 2002). Recently, Ebeling *et al.* (2002) showed that increasing the P concentration in a dairy cow diet increased the potential for P loss in runoff from land-applied manures, even with similar P application rates. This difference is likely due to a greater proportion of manure P being water soluble in high-P (40%) than low-P diet manure (29%). Similar trends of increased P loss in runoff as a function of greater water-soluble manure P are also evident for bee, pigs, and poultry (Kleinman *et al.*, 2002, 2005b).

Implementing a carefully planned diet tailored to meet the specific N and P requirements of animals in each phase of their growth will minimize nutrient loss to the environment in feces, urine, and gases. Reducing farm inputs of N and P in animal feed is a very effective BMP that can contribute to bringing about a lasting decrease in N and P loss to the environment. This is a direct result of the fact that farm N and P surpluses are governed by feed imports (**Table 2**), as well as the fact that farm-gate inputs of fertilizer N and P can be carefully matched to crop needs by plant tissue and soil testing. Other nutrient management measures (**Table 3**) are generally aimed at decreasing the potential for N and P loss and seen as short-term "band aids" and not solutions to environmental concerns.

Increase Nutrient-Use Efficiency in Animals

A reduction in the quantity of N and P excreted by livestock can be accomplished by supplementing livestock diets with enzymes to enhance the digestion of feedstuffs. Research indicates that an enzyme-treated diet decreased total N excretion 6% in late-lactating cows but increased N excretion by 3% in early-lactating cows (Knowlton *et al.*, 2002). A concurrent decrease (9.3 g day^{-1} or 10.1%) in total P excretion was observed in the early-lactating cows, but a change was not observed in the late-lactating cows. The results of adding amino acids and phytase to the diet of laying hens were much more dramatic. Daily total N and P excretion was reduced 45% and 48%, respectively (Keshavarz and Austic, 2003). Although a marked decrease in total N and P excretion was observed in

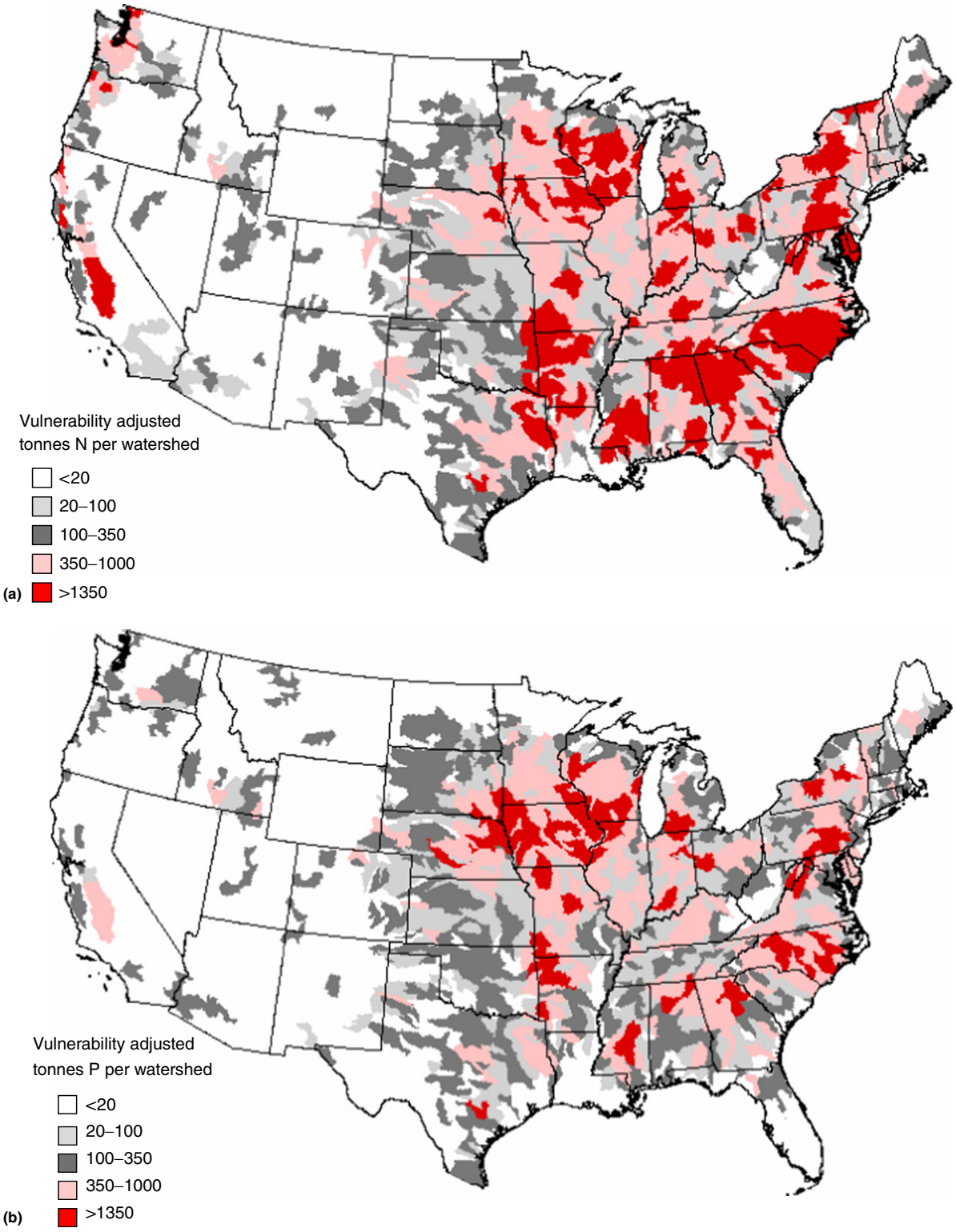


Figure 4 (a) Leaching vulnerability index for manure N available for land application, using 1997 census data. (b) Runoff vulnerability index for manure P available for land application, using 1997 census data.

Table 3 Beneficial management practices to minimize the loss of N and P from agriculture

Practice	*Description*	*Impact on loss*[a]	
		N	*P*
Farm inputs			
Crop hybrids	Low phytic acid corn reduces P in manure	Neutral	Decrease
Feed additives	Enzymes increase nutrient utilization by animals	Decrease	Decrease
Feed supplements	Match animals nutritional requirements	Decrease	Decrease
Source management			
Crop requirements	Nutrient applications based on crop N and/or P needs	Decrease	Decrease
Presidedress N Test	Can aid accurate split N applications	Decrease	Neutral
Soil P testing	Nutrient applications based on soil P availability	Neutral	Decrease
Tissue testing	N applications can be tailored to crop needs	Decrease	Neutral
Cover crops/residues	If harvested can reduce residual soil nutrients	Decrease	Decrease TP increase DP
Site-specific management	Use of Geographical Information System (GIS) and Global Positioning Systems (GPS) to apply and manage nutrient sources	Decrease	Decrease
Method of application	Incorporated, banded, or injected in soil	Decrease	Decrease
Rate of application	Match crop needs	Decrease	Decrease
Source application	Sources can differ in their P and N availability	Decrease	Decrease
Timing of application	Avoid application to frozen ground Apply during season with low runoff probability	Decrease	Decrease
Transport management			
Conservation tillage	Reduced and no-till increases infiltration and reduces soil erosion	Decrease TN increase NO_3	Decrease TP increase DP
Strip cropping, contour tillage, and terraces	Reduces transport of sediment-bound nutrients	Decrease TN neutral NO_3	Decrease TP neutral DP
Conservation cover	Permanent vegetative cover increases soil infiltration and water holding capacity	Decrease	Decrease
Invert stratified soils	Redistribution of surface P through profile by plowing	Neutral	Decrease
Buffer, riparian, wetland areas, and grassed waterways	Removes sediment-bound nutrients, enhances denitrification	Decrease	Decrease TP neutral DP
Critical source area treatment	Target sources of nutrients in a watershed for remediation	Decrease	Decrease

[a]TN is total N, NO_3 is nitrate, TP is total P, and DP is dissolved P.

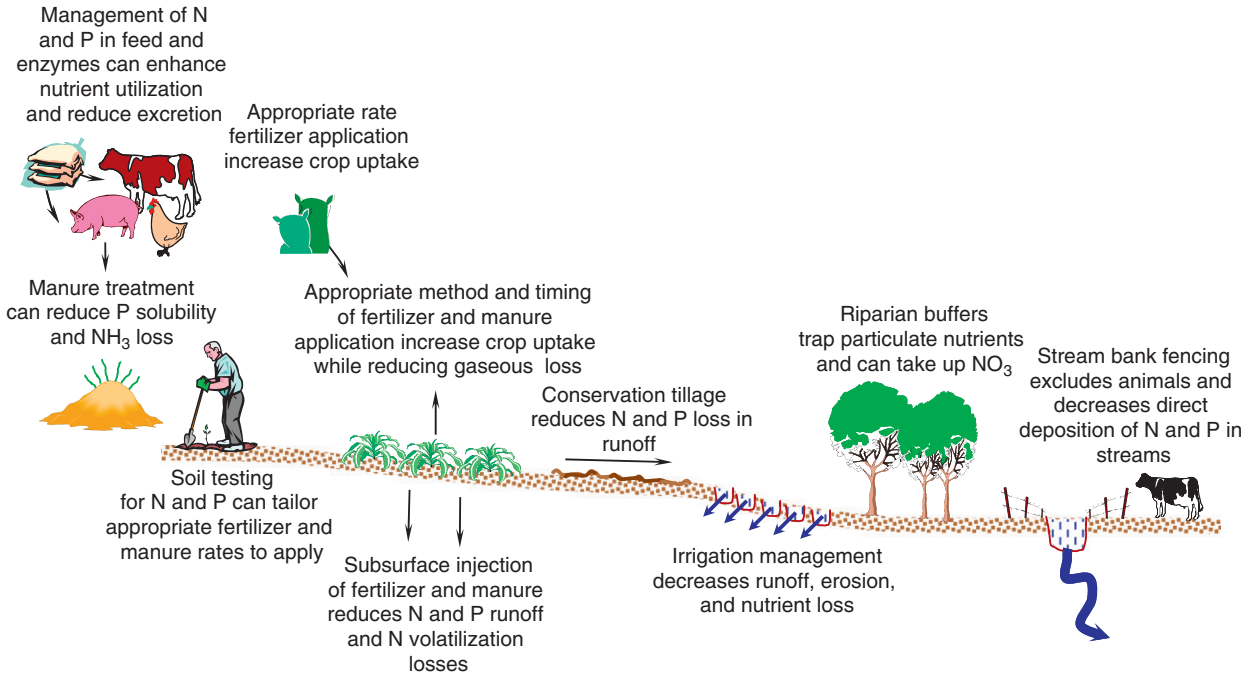


Figure 5 The management of N and P in agricultural systems.

this poultry study, the cost of the successful amino acid-supplemented diet was \$311 compared with \$126 t^{-1} for the control diet. Despite these dramatic results, the economic incentive to implement these new diets may not exist.

Corn hybrids are also available that contain low amounts of indigestible phytate P. Pigs and chickens fed "low-phytic acid" corn grain excreted less P in manure than those fed conventional corn varieties (Ertl *et al.*, 1998). This study also showed that P availability to nonruminants from low-phytate, high-available phosphate corn is about two to three times higher than from normal corn. Currently, the challenge to plant breeders is to incorporate the low-phytate trait into commercial grain hybrids with other agronomically desirable traits.

Nutrient Management

Soil and Plant Testing

Considerable attention has been paid to developing soil N tests for corn, with the preplant nitrate test successfully used in semiarid regions of the Western and Great Plains (Hergert, 1987); however, the preplant nitrate test does not work well in more humid regions (Bundy *et al.*, 1999). The preplant tests' lower success rate in more humid climates results from the increased probability of leaching and denitrification that occurs between soil sampling and plant N demand, as well as not representing the amount of N mineralization that might occur during the growing season. A modified approach to the preplant was proposed by Magdoff *et al.* (1984) to be used in the more humid corn production areas, known as the presidedress soil nitrate test to determine if and how much N fertilizer is needed to supplement any early season N mineralization. These soil nitrate tests are effective methods for reducing the amount of N fertilizer applied to corn, which reduces nitrate leaching potential by avoiding N applications in excess of plant requirements, although the window of opportunity between soil sampling and N fertilizer application is very small – an obstacle for many producers.

Because N availability is a function of multiple processes within the soil, varying spatially and temporally, more recent research efforts have focused on methods that evaluate the N status of the growing crop. Varvel *et al.* (1997) demonstrated that chlorophyll meters could be used as an index for N status in corn, and N fertilizer-use efficiency could be improved with this technique. Remotely sensing crop reflectance has been successfully used in a similar capacity, detecting N deficiency (Osborn *et al.*, 2002) and using an index for developing N recommendations.

Fertilizer P rates are usually established by crop need and modified by the amount already in the soil, as determined by established soil test P methods (Cox, 1994). In the case of commercial fertilizer P, applications can easily be tailored to match crop needs and minimize excessive soil P accumulation, because an economic disincentive exists to avoid applying excess fertilizer P. For instance, Dodd and Mallarino (2005) showed that annual fertilizer P rates of about 15 kg P per year maintained near-optimum soil test P levels (16–20 mg kg^{-1} as Bray-1P) and corn–soybean (*Glycine max* (L.) Merr.) yields for Mollisols in Iowa. However, applications of manure have, until recently, been made to meet the N needs of the crop, which has resulted in the build-up in soil test P above levels needed for optimum crop yields, and increases the risk of P loss in runoff (Daverede *et al.*, 2003; Pote *et al.*, 1996; Torbert *et al.*, 2002).

Application Method

Nitrogen fertilizers are applied by various methods, including injection, broadcasting, banding, and with irrigation water. The method selected usually depends on the machinery, fertilizer type, and compatibility with a farmer's overall crop management practices and usually has little impact on crop response to or environmental impact of the fertilizer. There are three exceptions: (1) when urea is applied to the soil surface, especially in no-till, (2) when N fertilizer is band-applied or injected into a microfeature that improves localized water infiltration, and (3) when N fertilizer is sprinkler-applied on sandy soils.

Urea applied to the soil surface is susceptible to considerable loss as volatilization of ammonia gas (as much as 30% of applied N; Fox *et al.*, 1986). Urea volatilized as ammonia poses an environmental risk as atmospheric ammonium fallout. Ammonium deposition poses an environmental threat to pristine ecosystems that are sensitive to slight changes in N inputs (e.g., estuaries or native forests); however, the impact of ammonia volatilization from urea fertilizers on such ecosystems is not fully understood.

Because of the relative immobility of P in the soil profile, placement of fertilizer P generally is more critical for plant availability than in the case of fertilizer N. Long-term studies in the northern Great Plains show that high rates of broadcast P (90 kg P per hectare) can have long-term effects (17 years) on soil test P, wheat yields (Halvorson and Black, 1985) and profitability (Jose, 1981; Halvorson *et al.*, 1986). Several studies show a greater yield response to surface or subsurface band application of fertilizer P at low rates, compared with broadcasting or mixing (Bailey and Grant, 1989; Lamond, 1987). In fact, Welch *et al.* (1966) observed greater P uptake and yield for corn with a combined banded (50%) and broadcast (50%) application (40 kg ha^{-1}).

Manure application methods and timing relative to rainfall influence the concentration of N and P in runoff (Dampney *et al.*, 2000; Sims and Kleinman, 2005). Several studies show, for example, reductions in N and P losses with an increase in the length of time between manure application and surface runoff (Djodjic *et al.*, 2000; Edwards and Daniel, 1993; Sharpley, 1997). These reductions can be attributed to the reaction of added P with soil and dilution of applied P by infiltrating water from rainfall that did not cause surface runoff. The incorporation of manure into the soil profile, either by tillage or by subsurface placement, reduces the potential for P loss in surface runoff. Rapid incorporation of manure also reduces ammonia volatilization and potential loss in runoff, as well as improving the N : P ratio for crop growth. For example, Mueller *et al.* (1984) showed that incorporation of dairy manure by chisel plowing reduced total P loss in runoff from corn 20-fold, compared with no-till areas receiving surface applications. In fact, P loss in runoff declined because of a lower concentration of P at the soil surface and a reduction in runoff with incorporation of manure (Mueller *et al.*, 1984; Pote *et al.*, 1996).

Rate and Frequency of Application

Nitrogen fertilizer should be applied nearest the time when needed by a crop. This minimizes opportunities for leaching below the root zone or gaseous losses via denitrification and, consequently, should result in the greatest proportion of applied N used by the crop. Although farmers might recognize this situation, deciding when to apply fertilizer depends on logistical and time constraints unique to every farm. Often, N fertilizer is applied much earlier than required by the crop. In the upper Midwest and Northeast, farmers sometimes are faced with the opportunity to apply N fertilizer on frozen soils. As a general rule, fertilizer should be applied when runoff potential is low. Fertilizer application to frozen soils should be avoided.

Many studies show that the loss of P in runoff relates directly to the rate and frequency of applied fertilizer P (Sharpley *et al.*, 2007; Sims and Kleinman, 2005). Avoiding fertilizer applications within a few days of expected rainfall can minimize runoff losses (Sims and Kleinman, 2005). Even so, fertilizer rates can easily be tailored to crop needs on the basis of soil test P, minimizing the potential for soil P build up and P loss in runoff. In fact, the use of fertilizer P has declined steadily since 1995 (Howarth *et al.*, 2000). This decline is most likely due to greater awareness and advice on the large savings in fertilizer costs afforded by nutrient management planning on farms.

Rainfall intensity and duration, as well as when rainfall occurs relative to when manure is applied, are all factors that influence the concentration and overall loss of manure N and P in runoff. The relationship between potential loss and application rate, however, is critical to establishing BMP strategies. Although soil P clearly is important in determining P loss in surface runoff, the rate and frequency of applying P to soil can override soil P in determining P loss (Sharpley *et al.*, 2007). Also evident are the long-lasting effects of applied manure on increased concentrations of P in surface runoff. For instance, Pierson *et al.* (2001) found that a poultry litter application tailored to meet pasture N demands elevated surface runoff P for up to 19 months after application. Although few studies have evaluated the loss of P in surface runoff as a function of application frequency, more frequent manure applications can be expected to increase soil P rapidly (Haygarth *et al.*, 1998; Sharpley *et al.*, 1993; 2004; Sims *et al.*, 1998), with a concomitant increase in runoff P loss.

Transport Management

Runoff Potential

The potential for runoff from a given site is important in determining the loss of P and, to a lesser extent, N and is thus a critical component of nutrient management strategies. For instance, Pionke *et al.* (1999 and 2000) generalized that about 90% of annual P export (0.63 kg ha^{-1} $year^{-1}$) from an agricultural watershed in the Chesapeake Bay Basin occurs during storm flow as a result of runoff from saturated areas extending no more than 60 m from the stream channel. In contrast, only 20% of N (6 kg ha^{-1} $year^{-1}$) is exported annually in storm runoff. Distance from where runoff is generated to a stream channel influences N and P loss and thus must be a nutrient management consideration (Gburek and Sharpley, 1998). Runoff, and nutrients carried by it, can be reduced or even intercepted by infiltration and deposition, respectively, before reaching a stream channel. Generally, the further a field is from a stream channel, the lower the potential for runoff to contribute nutrients to the stream. Many states, therefore, have adopted the premise of implementing more restrictive nutrient management strategies, particularly for P, on fields close to streams (Sharpley *et al.*, 2003).

Erosion Potential

A sequence of changing tillage practices in several watersheds in Oklahoma enabled comparison of surface and ground water impacts associated with native grass, conventionally tilled wheat, and no-till wheat (Sharpley and Smith, 1994). In the late 1970s, conversion of native grass to conventionally tilled wheat increased soil loss dramatically (**Table 4**). From 1979 to 1990, fertilizer N and P applications associated with conventionally tilled wheat also contributed to increased losses (3.74 kg N and 0.55 kg P per hectare per year) in surface runoff compared with native grass (**Table 4**). Over the same period, dissolved P and total P concentrations in surface runoff averaged 0.20 and 4.50 mg l^{-1}, respectively, whereas nitrate-N in ground water under two conventionally tilled wheat watersheds averaged 4.5 mg l^{-1} (**Figure 6**). Following conversion of two watersheds to no-till wheat in 1984, the total P concentration in runoff declined 25-fold by 1995, to 0.18 mg l^{-1} (**Figure 6**). This related to 18.9 kg ha^{-1} $year^{-1}$ less total N and 3.44 kg ha^{-1} $year^{-1}$ less total P in surface runoff from no-till wheat than conventionally tilled wheat (**Table 4**). However, both the dissolved P concentration (0.20–0.75 mg l^{-1}) and loss in runoff (0.26–0.74 kg ha^{-1} $year^{-1}$), as well as nitrate-N in ground water, increased (4.5–29.0 mg l^{-1}) following the conversion to no-till wheat (**Figure 6** and **Table 4**).

Similar amounts of fertilizer P and N were applied to the conventionally tilled and no-till wheat watersheds (about 12 kg N and 60 kg P per hectare per year; Sharpley and Smith, 1994). The increase in nitrate leaching, therefore, might be

Table 4 Annual sediment, N, and P loss as a function of tillage management of watersheds at El Reno and Woodward, Oklahoma

Study period	*Management*	*Sediment (kg ha^{-1} $year^{-1}$)*	*Dissolved P (kg ha^{-1} $year^{-1}$)*	*Total P (kg ha^{-1} $year^{-1}$)*	*Total N*
1977–1978	Native grass	103	0.03	0.08	0.50
1979–1983	Native grass	38	0.05	0.09	1.10
	Conventional till wheat	575	0.12	0.64	4.81
1984–1990	Native grass	34	0.07	0.15	1.31
	Conventional till wheat	11,300	0.26	4.87	26.12
	No-till wheat	625	0.74	1.43	7.19

Source: Adapted from Sharpley AN and Smith SJ (1994) Wheat tillage and water quality in the Southern Plains. *Soil and Tillage Research* 30: 3 3–38.

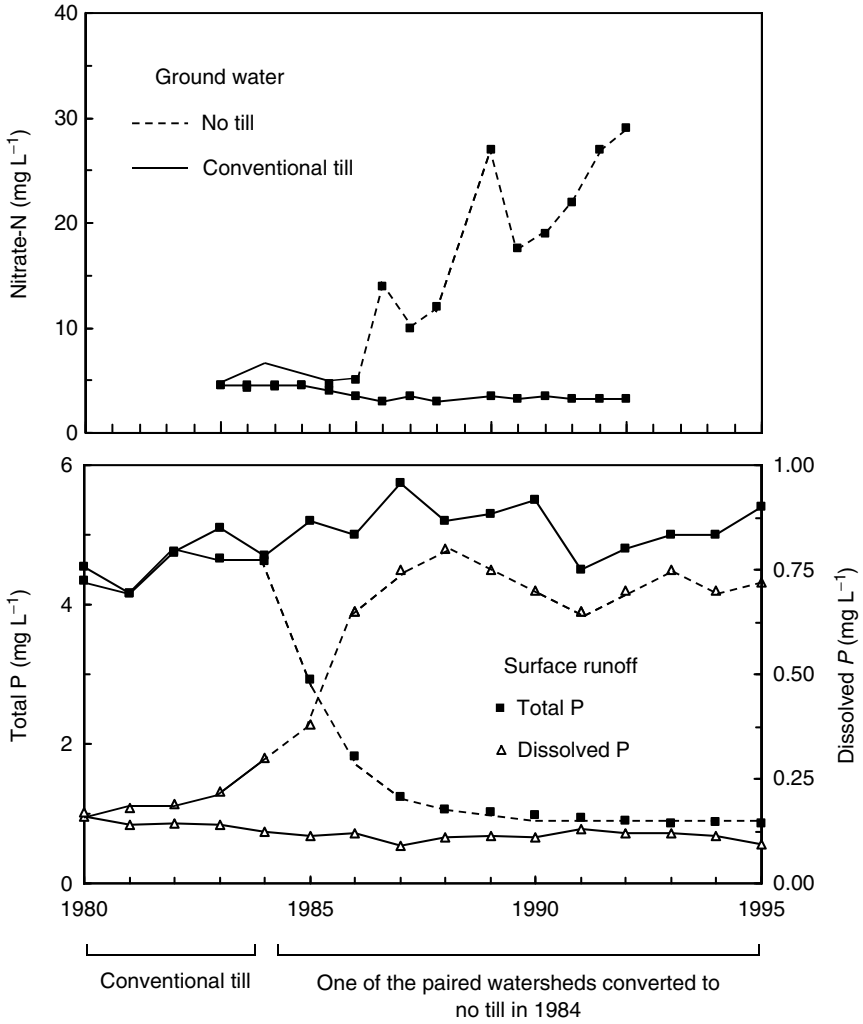


Figure 6 Mean annual nitrate-N concentration of ground water and dissolved and total P concentration runoff as a function of tillage management of watersheds in Oklahoma, USA. Data adapted from Sharpley AN and Smith SJ (1994) Wheat tillage and water quality in the Southern Plains. *Soil and Tillage Research* 30: 33–38.

attributed, in part, to greater soil water content under no-till compared with conventionally tilled wheat – a result of the increased residue cover enhancing percolation and soil water storage and reducing evaporation losses. The increase in dissolved P runoff can be attributed to leaching of P from crop residue and preferential transport of clay-sized particles under no-till compared with conventional tillage. In addition, such practices as stripcropping and contour planting reduce soil erosion potential and enhance infiltration.

These positive and negative impacts of conservation tillage practices on N and P loss potential should be considered in the development of sound nutrient management strategies. Clearly, a technically sound framework must be developed that recognizes critical sources of N and P export from agricultural watersheds so that optimal strategies at farm and watersheds scales can be implemented to best manage both N and P.

Permanent Vegetation

Keeping land in permanent cover, such as grass or cover crops, reduces the risk of runoff and erosion, increases infiltration, and thereby minimizes the loss of N and P. Cover crops protect the soil surface from raindrop impact, improve infiltration relative to bare soil, and trap eroded soil particles (Sharpley and Smith, 1991). Where dissolved P transport is the primary concern, cover crops may reduce runoff and, consequently, runoff P load. Cover crops are unlikely to affect dissolved P concentrations in runoff, however. Kleinman *et al.* (2005a) found that cover crops reduced the total P concentration in springtime runoff to 36% of the dissolved P in runoff from conventional corn. But dissolved P concentration was not significantly different between cover crops and conventional corn because dissolved P was controlled by soil P content rather than by soil erosion.

Grassed waterways are designed to trap sediment and reduce channel erosion. In some cases, waterways are installed as cross-slope diversions to intercept runoff and reduce effective slope length. Chow *et al.* (1999) estimated that installation of grassed waterways and terraces in combination reduced annual soil erosion 20-fold in a New Brunswick, Canada potato field.

Riparian/Buffer Areas

Healthy riparian areas can reduce N and P export, increase wildlife numbers and diversity, and improve aquatic habitat. In addition to acting as physical buffers to sediment-bound nutrients, plant uptake captures N and P, resulting in short-term and long-term accumulations of nutrients in biomass (Groffman *et al.*, 1992; Peterjohn and Correll, 1984; Uusi-Kämppä *et al.*, 2000). Enhanced denitrification in riparian areas can reduce the loss of N from agricultural fields to stream corridors (Howarth *et al.*, 2000; Lowrance *et al.*, 1985).

The effectiveness of conservation buffers as a nutrient management practice can vary significantly. For instance, the route and depth of subsurface water flow paths through riparian areas can influence nutrient retention. Conservation buffers are most efficient when sheet flow occurs, rather than channelized flow, which often bypasses some of the retention mechanisms. Those areas must be managed carefully to realize their full retention and filtration capabilities.

Critical Source Area Management

Transport of N and P from agricultural watersheds depends to a large extent on the coincidence of source (soil, crop, and management) and transport factors (runoff, erosion, and proximity to water course or body). Source factors relate to watershed areas with high potential to contribute to N and P export. For N, amounts applied in excess of crop requirements can be leached from the soil profile by percolating water. For P, source areas are often spatially confined and limited in extent, generally reflecting soil P status and fertilizer and manure P inputs (Gburek and Sharpley, 1998; Pionke *et al.*, 2000).

Transport factors determine whether nutrient sources translate into nutrient losses. Nitrogen transport from agricultural land generally occurs on a watershed scale. Source factors tend, therefore, to govern the magnitude of N loss, whereas transport factors govern the time lag or delay caused by percolation through various soil layers. Nitrogen loss on cropland depends on the balance between N added (amount, timing, and form of N) and crop removal. The rate of N loss through leaching (transport factor) depends on soil properties (primarily texture and permeability) and the amount of water available to drain through the soil profile (rainfall and irrigation).

Phosphorus transport generally occurs from hydrologically active areas in a watershed where surface runoff contributing to streamflow is coincident with areas of high soil P (Gburek and Sharpley, 1998; Gburek *et al.*, 1996). For example, Pionke *et al.* (1996) reported that 90% of annual P exported ($0.63\ kg^{-1}\ ha^{-1}\ year^{-1}$) from a south central Pennsylvania watershed occurred during storm flow as surface runoff on saturated areas extending no more than 60 m from the stream channel and that itself accounted for only 10% of annual runoff. In contrast, only 20% of nitrate export was associated with storm events; the remaining 80% was exported in baseflow ($6\ kg^{-1}\ ha^{-1}\ year^{-1}$), which came from a larger recharge area of the watershed. Even in regions where subsurface flow pathways dominate, areas contributing P to drainage water appear to be restricted to soils with high soil P saturation and hydrologic connectivity to the drainage network. For example, Schoumans and Breeuwsma (1997) found that soils with high-P saturation contributed only 40% of the total P load, whereas another 40% came from areas where the soils had only moderate P saturation but some degree of hydrological connectivity with the drainage network.

Water Quality Response to Land-Use Change

Critical sources areas of P on agricultural lands have increased due to the several decades of N-based manure applications that have resulted in the localized accumulation of P in soil. This stored P in soil and fluvial sediments is referred to as legacy P. Legacy P is particularly problematic because soil and sediment can serve as a long-term source of P to the overlying water column for several decades. Within freshwater ecosystems, legacy P can contribute to the external loading of rivers, lakes, and bays. Considering both soil and sediment sources, legacy P can mask reductions in P loss from agricultural lands with the implementation of BMPs.

In fluvial systems, the release of legacy P from deposited sediments can be influenced by the oxygen status of overlying waters, where reducing conditions favor the release of sediment-bound P. Also, the beneficial effects of BMPs on lower P inputs to stream and rivers will decrease P concentrations that trigger the release of P from sediments, masking the downstream benefits of BMPs on P fluxes. Finally, P-enriched sediments are resuspended during high flows.

Currently, a basin-level response to practices cannot be expected to be observed for some time. A better understanding of the spatial and temporal aspects of basin-level responses is needed, as well as on other scales at which response can occur in a more timely fashion. This would likely be smaller subwatershed scales, where local water quality and quantity benefits may become evident more quickly (Bishop *et al.*, 2005) and that can enhance BMP adoption. Research on pre- and post-BMP implementation has indicated the value of small-scale monitoring and/or modeling of BMP effects at a local level (Chu *et al.*, 2004; Mostaghimi *et al.*, 2001). The modeling and monitoring study by Chu *et al.* (2004) indicated that climate variability plays a major role in compounding the short-term BMP effect in water quality improvements, even in small watersheds.

The importance of long-term monitoring in terms of assessing watershed response to conservation management is demonstrated by the Lake Erie Agricultural Systems for Environmental Quality (LEASEQ) Project (Richards *et al.*, 2002b, 2009). Phosphorus loads in two Ohio watersheds (Maumee and Sandusky River Watersheds) with major tributaries to Lake Erie, monitored since 1975 to determine the effect of adopting BMPs such as conservation tillage and nutrient management planning in predominantly row-crop agriculture (corn, soybean, and wheat), averaged an 8% increase in flow whereas the mean annual flow-weighted concentrations of suspended sediment, total P, and dissolved P decreased 23%, 44%, and 86%, respectively (Richards *et al.*, 2002a). Since 1995, annual flow-weighted concentrations of dissolved P have increased, whereas particulate (and total P) have continued to decline (Richards and Baker, 2010). The trend of increasing dissolved P and decreasing total P may be attributed to a combination of several factors: a change in the rainfall distribution pattern; a buildup at the soil surface with

conversion to no-till cropping; and increased application of fertilizer and manure, without incorporation in the fall and winter. Although this project did show that water quality can change in response to management change at a watershed scale, lessons learnt are broadly applicable to other watersheds.

Conclusions

In the past, separate strategies for P and N have been developed and implemented at farm or watershed scales. Because of differing biology, chemistry, and flow pathways of P and N in soil, these narrowly targeted strategies may lead to conflicting or suboptimal advice. As a result, the prevention of P and N loss from agricultural systems needs to focus on defining, targeting, and remediating source areas of P that combine high soil P levels with high erosion and surface runoff potentials and source areas of N that coincide with soils of high permeability. Thus, differing levels of management may be appropriate for different areas of a watershed. Overall, inputs in fertilizer and manures should be carefully matched with crop needs over the whole watershed. Short-term remediation of P loss should focus on critical source areas of P export where high soil P, P application, and zones of surface runoff and erosion coincide. However, lasting improvements in water quality can only be achieved by balancing system inputs and outputs of both P and N. Clearly, the management of agricultural nutrients involves a complex suite of options that must be customized to meet site-specific needs that are depicted in **Figure 5**. Even so, the long-term impacts have been and remain difficult to quantify.

A range of voluntary and regulatory measures can be used to encourage implementation of nutrient management strategies as part of conservation programs to protect soil and water resources. In general, the success of these measures relates to how well farmers can afford to implement new management strategies and the concomitant level of support or incentives for their adoption. In some instances, local or regional governmental controls may be necessary to enhance prompt adoption of practices that will have a positive influence on environmental outcomes.

In some areas, we cannot and should not expect that pristine waters are achievable with ever-increasing population densities and more intensive agricultural production systems to meet demand. The bottom line is that this may require either a reassessment of water-use designations and/or far-reaching societal commitment and support of agricultural system changes.

See also: Agriculture, Industrialized. Agriculture, Sustainable. Agriculture, Traditional. Biodiversity of Harmful Marine Algae. Conservation Efforts, Contemporary. Environmental Impact, Concept and Measurement of. Eutrophication and Oligotrophication. Freshwater Ecosystems. Freshwater Ecosystems, Human Impact on. Indirect Land Use and Greenhouse Gas Impacts of Biofuels. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production. River Ecosystems. Soil Conservation

References

Amdur MO, Dull J, and Klassen ED (eds.) (1991) *Casarett and Doull's Toxicology*, 4th edn. New York, NY: Pergamon Press.

Bailey LD and Grant CA (1989) Fertilizer phosphorus placement studies on calcareous and noncalcareous chernozemic soils: Growth P-uptake and yield of flax. *Communications in Soil Science and Plant Analysis* 20: 635–654.

Bengtson RL, Carter CE, Morris HF, and Bartkiewics SA (1988) The influence of subsurface drainage practices on nitrogen and phosphorus losses in a warm, humid climate. *Transactions of the ASAE* 31: 729–733.

Bishop PL, Hively WD, Stedinger JR, Rafferty MR, Lojpersberger JL, and Bloomfield JA (2005) Multivariate analysis of paired watershed data to evaluate agricultural best management practice effects on stream water phosphorus. *Journal of Environmental Quality* 34: 1087–1101.

Boesch DF, Brinsfield RB, and Magnien RE (2001) Chesapeake Bay eutrophication: Scientific understanding, ecosystem restoration, and challenges for agriculture. *Journal of Environmental Quality* 30: 303–320.

Bouwman AF and Booij H (1998) Global use and trade of feedstuffs and consequences for the nitrogen cycle. *Nutrient Cycling in Agroecosystems* 52: 261–267.

Bundy LG and Kelling KL (2001) Reacting to high N costs and limited supplies. *Wisconsin Crop Manager* 8: 10–13.

Bundy LG, Walters DT, and Olness AE (1999) *Evaluation of Soil Nitrate Tests for Predicting Corn Nitrogen Response in the North Central Region.* Madison, WI: Wisconsin Agricultural Experiment Station and College of Agricultural and Life Sciences. University of Wisconsin. North Central Regional Research Publication no. 342. Available on-line at http://www.soils.wisc.edu/extension/Word_Documents_PDF2/soilNO3.pdf (last verified 1 September 2005).

Burkholder JA and Glasgow Jr. HB (1997) *Pfiesteria piscicidia* and other Pfiesteria-dinoflagellates behaviors, impacts, and environmental controls. *Limnology and Oceanography* 42: 1052–1075.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, and Smith VH (1998) Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecolological Applications* 8: 559–568.

CAST – Council for Agricultural Science and Technology. 1985. Agricultural and groundwater quality. *Rep. 103.* Ames, IA: CAST.

CAST – Council for Agricultural Science and Technology. 2002. Animal diet modification to decrease the potential for nitrogen and phosphorus pollution. *Issue Paper* 21, 16 pp. Ames, IA: Council for Agricultural Science and Technology.

Chow TL, Rees HW, and Daigle JL (1999) Effectiveness of terraces/grassed waterway systems for soil and water conservation: a field evaluation. *Journal of Soil and Water Conservation* 54: 577–583.

Chu TW, Shirmohammadi A, Montas H, and Sadeghi AM (2004) Evaluation of SWAT Model's Sediment and Nutrient Components in Piedmont Physiographic Region of Maryland. *Transactions of American Society of Agricultural Engineers* 47(5): 1523–1538. http://europa.eu.int/comm/environment/water/water-framework/index_en.html.

Council of European Communities (2000) Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the Community action in the field of water policy. *Official Journal L* 327: 0001–0073. http://europa.eu.int/comm/environment/water/water-framework/indexen.html

Cox FR (1994) Current phosphorus availability indices: characteristics and shortcomings. In: Havlin JL, Jacobsen JS, Lekiam DF, Fixen PE, and Hergert GW (eds.) *Soil Testing: Prospects for Improving Nutrient Recommendations.* Soil Science Society of America Special Publication No. 40, pp. 101–114. Madison, WI: Soil Science Society of America.

Dampney PMR, Lord EI, and Chambers BJ (2000) Development of advice for farmers and advisors. *Soil Use and Management* 16: 162–166.

Daverede IC, Kravchenko AN, Hoeft RG, *et al.* (2003) Phosphorus runoff: Effect of tillage and soil phosphorus levels. *Journal of Environmental Quality* 32: 1436–1444.

Djodjic F, Ulen B, and Bergstrom L (2000) Temporal and spatial variations of phosphorus losses and drainage in a structured clay soil. *Water Research* 34: 1687–1695.

Dodd JR and Mallarino AP (2005) Soil-test phosphorus and crop grain yield responses to long-term phosphorus fertilization for corn-soybean rotations. *Soil Science Society of America Journal* 69: 1118–1128.

Dou A, Galligan TD, Ramberg Jr. CF, Meadows C, and Ferguson JD (2001) A survey of dairy farming in Pennsylvania: Nutrient management practices and implications. *Journal of Dairy Science* 84: 966–973.

Ebeling AM, Bundy LG, Powell MJ, and Andraski TW (2002) Dairy diet phosphorus effects in phosphorus losses in runoff from land-applied manure. *Soil Science Society of America Journal* 66: 284–291.

Edwards DR and Daniel TC (1993) Drying interval effects on runoff from fescue plots receiving swine manure. *Transactions of the American Society of Agricultural Engineering* 36: 1673–1678.

Ertl DS, Young KA, and Raboy V (1998) Plant genetic approaches to phosphorus management in agricultural production. *Journal of Environmental Quality* 27: 299–304.

Evans R, Cuffman-Neff LC, and Nehring R (1996). Increases in agricultural productivity, 1948–1993. *Updates on Agricultural Resources and Environmental Indicators* No. 6. U.S. Department of Agriculture-Economic Research Service. Washington, DC: U.S. Government Printing Office, 85 pp.

Fox RH, Kern JM, and Piekielek WP (1986) Nitrogen fertilizer source, and method and time of application effects on no-till corn yields and nitrogen uptakes. *Agronomy Journal* 78: 741–746.

Gburek WJ and Sharpley AN (1998) Hydrologic controls on phosphorus loss from upland agricultural watersheds. *Journal of Environmental Quality* 27: 267–277.

Gburek WJ, Sharpley AN, and Pionke HB (1996) Identification of critical source areas for phosphorus export from agricultural catchments. In: Anderson MG and Brookes S (eds.) *Advances in Hillslope Processes*, pp. 263–282. Chichester, UK: John Wiley & Sons.

Gibson GR, Carlson R, Simpson J, *et al.* (2000) *Nutrient criteria technical guidance manual: lakes and reservoirs* (EPA-822-B00-001). U.S. Environmental Protection Agency, Washington, DC. Washington, DC: U.S. Government Printing Office.

Groffman PM, Gold AJ, and Simmons RC (1992) Nitrate dynamics in riparian forests: microbial studies. *Journal of Environmental Quality* 21: 666–671.

Halvorson AD and Black AL (1985) Long-term dryland crop responses to residual phosphorus fertilizer. *Soil Science Society of America Journal* 49: 928–933.

Halvorson AD, Black AL, Watt DL, and Leholm AG (1986) Economics of a one-time phosphorus application in the northern Great Plains. *Applied Agricultural Research* 1: 137–144.

Haygarth PM, Chapman PJ, Jarvis SC, and Smith RV (1998) Phosphorus budgets for two contrasting grassland farming systems in the UK. *Soil Use and Management* 14: 160–167.

Hergert GW (1986) Nitrate leaching through sandy soil as affected by sprinkler irrigation management. *Journal of Environmental Quality* 15: 272–278.

Hergert GW (1987) Status of residual nitrate-nitrogen soil tests in the United States of America. In: Brown JR (ed.) *Soil Testing: Sampling Correlation, Calibration, and Interpretation*. Special Publication no. 21. Madison, WI: Soil Science Society of America.

Hergert GW, Klausner SD, Bouldin DR, and Zwerman PJ (1981) Effects of dairy manure on phosphorus concentrations and losses in tile effluent. *Journal of Environmental Quality* 10: 345–349.

Hilton J, O'Hare M, Bowes MJ, and Jones I (2006) How green is my river? A new paradigm of eutrophication in rivers. *Science of the Total Environment* 365: 66–83.

Holland EA, Dentener FJ, Braswell BH, and Sulzman JM (1999) Contemporary and pre-industrial global reactive nitrogen budgets. *Biogeochemistry* 46(1/3): 7–43.

Howarth RW, Anderson DA, Church TM, *et al.* (2000) *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. Washington, DC: National Research Council. National Academy Press. 405 pp.

Howarth RW, Billen G, Swaeny D, *et al.* (1996) Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. *Biogeochemistry* 35: 75–139.

Jarvie HP, Neal C, and Withers PJA (2006) Sewage-effluent phosphorus: A greater risk to river eutrophication than agricultural phosphorus? *Science of The Total Environment* 360: 246–253.

Jose HD (1981) An economic comparison of batch and annual phosphorus fertilizer application in wheat production in western Canada. *Canadian Journal of Soil Science* 61: 47–54.

Kellogg RL (2001) Potential priority watersheds for protection of water quality from contamination by manure nutrients. *Animal Residuals Management Working Paper*. USDA, Natural Resources Conservation Service and Economic Research Service, General Services Administration, National Forms and Publication Center, Fort Worth, TX, 20 pp. http://www.nrcs.usda.gov/technical/NRI/pubs/wshedpap_w.pdf

Kellogg RL, Lander CH, Moffitt DC and Gollehon N (2000) Manure nutrients relative to the capacity of cropland and pastureland to assimilate nutrients: Spatial and temporal trends for the United States. *Resource Assessment and Strategic Planning Working Paper* 98-1, USDA, Natural Resources Conservation Service and Economic Research Service, General Services Administration, National Forms and Publication Center, Fort Worth, TX, 140 pp. http://www.nrcs.usda.gov/technical/NRI/pubs/manntr.pdf

Keshavarz K and Austic RE (2003) The use of low-protein, low-phosphorus, amino acid- and phytase-supplemented diets on laying hen performance and nitrogen and phosphorus excretion. *Poultry Science* 83: 75–83.

Klausner SD, Fox DG, Rasmussen CN, *et al.* (1998) Improving dairy farm sustainability. 1. An approach to animal and crop nutrient management planning. *Journal of Production Agriculture* 11: 225–232.

Kleinman PJA, Salon P, Sharpley AN, and Saporito LS (2005a) Effect of cover crops established at time of corn planting on phosphorus runoff from soils before and after dairy manure application. *Journal of Soil and Water Conservation* 60: 311–322.

Kleinman PJA, Sharpley AN, Moyer BG, and Elwinger G (2002) Effect of mineral and manure phosphorus sources on runoff phosphorus losses. *Journal of Environmental Quality* 2026–2033.

Kleinman PJA, Wolf AM, Sharpley AN, Beegle DB, and Saporito LS (2005b) Survey of water extractable phosphorus in livestock manures. *Soil Science Society of America Journal* 69: 701–708.

Klocke NL, Watts DG, Schneekloth JP, Davison DR, Todd RW, and Parkhurst AM (1999) Nitrate leaching in irrigated corn and soybean in a semi-arid climate. *Transactions of the ASAE* 42: 1621–1630.

Knowlton KF, McKinney JM, and Cobb C (2002) Effect of direct-fed fibrolytic enzyme formulation on nutrient intake, partitioning, and excretion in early and late lactation Holstein cows. *Journal of Dairy Science* 85: 3328–3335.

Kotak BG, Kenefick SL, Fritz DL, Rousseaux CG, Prepas EE, and Hrudey SE (1993) Occurrence and toxicological evaluation of cyanobacterial toxins in Alberta lakes and farm dugouts. *Water Research* 27: 495–506.

Kotak BG, Prepas EE, and Hrudey SE (1994) Blue green algal toxins in drinking water supplies: Research in Alberta. *Lake Line* 14: 37–40.

Lamond RE (1987) Comparison of fertilizer solution placement methods for grain sorghum under two tillage systems. *Journal of Fertilizers* Issues 4: 43–47.

Lanyon LE (2000) Nutrient management: Regional issues affecting the Chesapeake Bay. In: Sharpley AN (ed.) *Agriculture and Phosphorus Management: The Chesapeake Bay*, pp. 145–158. Boca Raton, FL: CRC Press.

Lanyon LE (2005) Phosphorus, animal nutrition and feeding: Overview. In: Sims JT and Sharpley AN (eds.) *Phosphorus; Agriculture and the Environment*, pp. 561–586. Madison, WI: American Society of Agronomy. American Society of Agronomy Monograph.

Lander CH, Moffitt D, and Alt K (1998) Nutrients available from livestock manure relative to crop growth requirements. Resource Assessment and Strategic Planning Working Paper 98-1, USDA-Natural Resources Conservation Service, Washington, DC: http://www.nrcs.usda.gov/technical/land/pubs/nlweb.html. Last verified 1 September 2005.

Lawton LA and Codd GA (1991) Cyanobacterial (blue-green algae) toxins and their significance in UK and European waters. *Journal of the Institution of Water & Environmental Management* 5: 460–465.

Lowrance RR, Leonard RA, and Sheridan JM (1985) Managing riparian ecosystems to control non-point pollution. *Journal of Soil and Water Conservation* 40: 87–91.

Magdoff FR, Ross D, and Amadon J (1984) A soil test for nitrogen availability to corn. *Soil Science Society of America Journal* 48: 1301–1304.

Martin A and Cooke GD (1994) Health risks in eutrophic water supplies. *Lake Line* 14: 24–26.

Mostaghimi S, Brannan KM, Dillaha TA, and Bruggeman AC (2001) Best management practices for nonpoint source pollution control. In: Ritter WL (ed.) *Agricultural Nonpoint Source Pollution: Watershed Management and Hydrology*, pp. 59–89. Boca Raton, FL: Lewis Publishers.

Mueller DH, Wendt RC, and Daniel TC (1984) Phosphorus losses as affected by tillage and manure application. *Soil Science Society of America Journal* 48: 901–905.

National Research Council (2001) *Nutrient Requirements of Dairy Cattle*, 7th rev. edn. Washington, DC: National Academy of Sciences.

Osborn SL, Schepers JS, Francis DD, and Schlemmer MR (2002) Detection of phosphorus and nitrogen deficiencies in corn using spectral radiance measurements. *Agronomy Journal* 94: 1215–1221.

Palmstrom NS, Carlson RE, and Cooke GD (1988) Potential links between eutrophication and formation of carcinogens in drinking water. *Lake and Reservoir Management* 4: 1–15.

Peterjohn WT and Correll DL (1984) Nutrient dynamics in an agricultural watershed: Observations on the role of a riparian forest. *Ecology* 65: 1466–1475.

Pierson ST, Cabrera ML, Evanylo GK, *et al.* (2001) Phosphorus and ammonium concentrations in surface runoff from grasslands fertilized with broiler litter. *Journal of Environmental Quality* 30: 1784–1789.

Pierzynski GM and Logan TJ (1993) Crop, soil, and management effects on phosphorus soil test levels. *Journal of Production Agriculture* 6: 513–520.

Pionke HB, Gburek WJ, and Sharpley AN (2000) Critical source area controls on water quality in an agricultural watershed located in the Chesapeake Basin. *Ecological Engineering* 14: 325–335.

Pionke HB, Gburek WJ, Sharpley AN, and Schnabel RR (1996) Flow and nutrient export patterns for an agricultural hill-land watershed. *Water Resources Research* 32: 1795–1804.

Pionke HB, Gburek WJ, Schnabel RR, Sharpley AN, and Elwinger GF (1999) Seasonal flow, nutrient concentrations and loading patterns in stream flow draining an agricultural hill-land watershed. *Journal of Hydrology* 220: 62–73.

Pote DH, Daniel TC, Sharpley AN, Moore Jr. PA, Edwards DR, and Nichols DJ (1996) Relating extractable soil phosphorus to phosphorus losses in runoff. *Soil Science Society of America Journal* 60: 855–859.

Powell JM, Jackson-Smith DB, and Satter LD (2002) Phosphorus feeding and manure nutrient recycling on Wisconsin dairy farms. *Nutrient Cycling in Agroecosystems* 62: 277–286.

Powell JM, Wu Z, and Satter LD (2001) Dairy diet effects on phosphorus cycles of cropland. *Journal of Soil and Water Conservation* 56: 22–26.

Power JF and Schepers JS (1989) Nitrate contamination in North America. *Agriculture, Ecosystems and Environment* 26: 165–187.

Poulsen HD (2000) Phosphorus utilization and excretion in pig production. *Journal of Environmental Quality* 29: 24–27.

Rabalais NR, Turner RE, and Wiseman Jr. WJ (2001) Hypoxia in the Gulf of Mexico. *Journal of Environmental Quality* 30: 320–329.

Richards RP, Baker DB, and Crumrine JP (2009) Improved water quality in Ohio tributaries to Lake Erie: A consequence of conservation practices. *Journal of Soil and Water Conservation* 64(3): 200–211.

Richards RP, Baker DB, Crumrine JP, and Sterns AM (2010) Unusually large loads in 2007 from the Maumee and Sandusky Rivers, tributaries to Lake Erie. *Journal of Soil and Water Conservation* 65: 450–462.

Richards RP, Baker DB, and Eckert DJ (2002a) Trends in agriculture in the LEASEQ watersheds, 1975–1995. *Journal of Environmental Quality* 31: 17–24.

Richards RP, Calhoun FG, and Matisoff G (2002b) The Lake Erie agricultural systems for environmental quality project: An introduction. *Journal of Environmental Quality* 31: 6–16.

Sandstedt CA (1990) *Nitrates: Sources and their Effects upon Humans and Livestock*. Washington, DC: The American University.

Sansinena M, O'Connor LD, Fox DG, Van Socst PJ, and Sniffen CJ (1999) A survey of trends and rationales for dietary phosphorus recommendations among mid-south dairy nutritionists. *Proceedings of the 1999 Mid-South Ruminant Nutrition Conference* 6–7 May, pp. 51–54. Dallas, TX: Texas AgriLife Extension Service.

Schindler DW (1977) Evolution of phosphorus limitation in lakes. *Science* 195: 260–262.

Schoumans OF and Breeuwsma A (1997) The relation between accumulation and leaching of phosphorus: Laboratory, field and modelling results. In: Tunney H, Carton OT, Brookes PC, and Johnston AE (eds.) *Phosphorus Loss from Soil to Water*, pp. 361–363. Cambridge, UK: CAB International Press.

Sharpley AN (1997) Rainfall frequency and nitrogen and phosphorus in runoff from soil amended with poultry litter. *Journal of Environmental Quality* 26: 1127–1132.

Sharpley AN and Rekolainen S (1997) Phosphorus in agriculture and its environmental implications. In: Tunney H, Carton OT, Brookes PC, and Johnston AE (eds.) *Phosphorus Loss from Soil to Water*, pp. 1–54. Cambridge, UK: CAB International Press.

Sharpley AN and Smith SJ (1991) Effect of cover crops on surface water quality. In: Hargrove WL (ed.) *Cover Crops for Clean Water*, pp. 41–50. Ankeny, IA: Soil and Water Conserv. Soc.

Sharpley AN and Smith SJ (1994) Wheat tillage and water quality in the Southern Plains. *Soil and Tillage Research* 30: 33–38.

Sharpley AN and Syers JK (1979) Loss of nitrogen and phosphorus in tile drainage as influenced by urea application and grazing animals. *New Zealand Journal of Agriculture Research* 22: 127–131.

Sharpley AN, Herron S, and Daniel TC (2007) Overcoming the challenges of phosphorus-based management in poultry farming. *Journal of Soil and Water Conservation* 62: 375–389.

Sharpley AN, McDowell RW, and Kleinman PJA (2004) Amounts, forms and solubility of phosphorus in soils receiving manure. *Soil Science Society of America Journal* 68: 2048–2057.

Sharpley AN, Smith SJ, and Bain R (1993) Effect of poultry litter application on the nitrogen and phosphorus content of Oklahoma soils. *Soil Science Society of America Journal* 57: 1131–1137.

Sharpley AN, Withers PJA, Abdalla C, and Dodd A (2005) Strategies for the sustainable management of phosphorus. In: Sims JT and Sharpley AN (eds.) *Phosphorus; Agriculture and the Environment*, pp. 1069–1101. Madison, WI: American Society of Agronomy. American Society of Agronomy Monograph.

Sharpley AN, Weld JL, Beegle DB, *et al.* (2003) Development of phosphorus indices for nutrient management planning strategies in the U.S. *Journal of Soil and Water Conservation* 58: 137–152.

Sims JT (1997) Agricultural and environmental issues in the management of poultry wastes: Recent innovations and long-term challenges. In: Rechcigl J and MacKinnon HC (eds.) *Uses of By-Products and Wastes in Agriculture*, pp. 72–90. Washington, DC: American Chemical Society.

Sims JT and Kleinman PJA (2005) Managing agricultural phosphorus for environmental protection. In: Sims JT and Sharpley AN (eds.) *Phosphorus; Agriculture and the Environment*, pp. 1021–1068. Madison WI: American Society of Agronomy. American Society of Agronomy Monograph.

Sims JT and Sharpley AN (eds) (2005) *Phosphorus: Agriculture and the Environment*. American Society of Agronomy Monograph. Madison, WI: American Society of Agronomy.

Sims JT, Joern BC, and Simard RR (1998) Phosphorus losses in agricultural drainage: Historical perspective and current research. *Journal of Environmental Quality* 27: 277–293.

Spalding RF and Exner ME (1993) Occurrence of NO_3 in Groundwater – A Review. *Journal of Environmental Quality* 22: 392–402.

Torbert HA, Daniel TC, Lemunyon JL, and Jones RM (2002) Relationship of soil test phosphorus and sampling depth to runoff phosphorus in calcareous and noncalcareous soils. *Journal of Environmental Quality* 31: 1380–1387.

Tylutki TP and Fox DG (2000) Quality control in herd nutrient management. *Proceedings of the 2000 Cornell Nutrition Conference for Feed Manufacturers*, Ithaca, NY, pp. 130–143.

U.S. Environmental Protection Agency (1996) *Environmental indicators of water quality in the United States*. EPA 841-R-96-002. U.S. EPA, Office of Water (4503 F). Washington, DC: U.S. Government Printing Office.

U.S. Environmental Protection Agency (2001) *Ecoregional nutrient criteria*. EPA-822-F-01-010. USEPA, Office of Water (4304). Washington, DC: U.S. Government Printing Office. http://www.epa.gov/waterscience/criteria/nutrient/9docfs.pdf

U.S. Environmental Protection Agency (2002) Nutrient water quality criteria, summary table for the nutrient criteria documents. http://www.epa.gov/waterscience/criteria/nutrient/ecoregions/

Uusi-Kämppä J, Braskerud B, Jansson H, Syverson N, and Uusitalo R (2000) Buffer zones and constructed wetlands as filters for agricultural phosphorus. *Journal of Environmental Quality* 29: 151–158.

Valk H, Metcalf JA, and Withers PJA (2000) Prospects for minimizing phosphorus-excretion in ruminants by dietary manipulation. *Journal of Environmental Quality* 29: 28–36.

Varvel GE, Schepers JS, and Francis DD (1997) Ability for in-season correction of nitrogen deficiency in corn using chlorophyll meters. *Soil Science Society of America Journal* 61: 1233–1239.

Welch LF, Mulvaney DOL, Boone LV, McKibben CE, and Pendleton JW (1966) Relative efficiency of broadcast versus banded phosphorus for corn. *Agronomy Journal* 58: 283–287.

Withers PJA and Lord EI (2002) Agricultural nutrient inputs to rivers and groundwaters in the UK: policy, environmental management and research needs. *Science of The Total Environment* 282: 9–24.

Withers PJA, Edwards AC, and Foy RH (2002) Phosphorus cycling in UK agriculture and implications for phosphorus loss from soil. *Soil Use Manage* 17: 139–149.

Wu Z, Satter LD, Blohowiak AJ, Stauffacher RH, and Wilson JH (2001) Milk production, phosphorus excretion, and bone characteristics of dairy cows fed different amounts of phosphorus for two or three years. *Journal of Dairy Science* 84: 1738–1748.

Agriculture, Sustainable

G Philip Robertson and Richard R Harwood, Michigan State University, East Lansing, MI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bush-fallow cropping system An agronomic system in which soil fertility is maintained by allowing native vegetation to regrow following several years of cropping.
Conservation tillage Tillage that reduces soil disruption in order to conserve organic matter and water and reduce erosion.
Ecosystem services The benefits people obtain from ecosystems.
Legumes Plants belonging to the family Leguminosae, many of whose members can form symbiotic associations with nitrogen-fixing bacteria.
Multifunctional cropping system An agronomic system that incorporates different combinations of herbaceous crops, trees, and animals within a single farm unit.
Nitrogen fixation The transformation of atmospheric nitrogen into a form usable by plants.
Resource conservation The protection and enhancement of resources on which sustainability depends.
Rotation The sequence of crops grown in a single field.
Soil organic matter Material in soil containing organic carbon derived from the decomposition of mainly plant residue.
Trophic complexity The number and types of organisms that feed at different trophic levels within a community; also known as food web complexity.

Elements of Sustainability

Strictly defined, sustainability denotes any system capable of persisting. Because persistence depends on scale, *sustainable agriculture* is entirely scale dependent. An agricultural practice that is sustainable at the scale of an individual field may lack sustainability at the larger farm scale if the inputs required to maintain stable production eventually exceed the capacity of the farm to provide them. Likewise, farm-scale sustainability must be evaluated in the context of local and global regions, and global sustainability requires consideration of global output versus the long-term costs of that output and the ability of global resources to accept those costs.

Of course costs are not simply economic. They are also social and environmental: If a cropping system requires large inputs of nitrogen fertilizer that leak from the system to pollute groundwater drinking supplies and distant coastal fisheries, a system that may be sustainable at the field scale becomes nonsustainable at the farm and regional scale – even though the long-term supply of fertilizer is stable and the economic cost of fertilizer to the farmer is easily borne by higher grain production. Landscape models now being developed for farmer decision making on alternatives and for development extension planning are becoming increasingly available (Nassauer and Opdam, 2008).

Sustainable agriculture must therefore be defined not just in terms of its long-term economic productivity but also in terms of its environmental and social benefits and costs. Many of these costs are value-laden. For a society that values family farms, for example, sustainability must be evaluated in terms of the impact of farming practices on the social structure of rural communities: Do markets and policies favor large-scale producers to the exclusion of those who operate family farms and the rural communities in which they live? For a society that values biodiversity, sustainability must be evaluated in terms of the impact of farming practices on wildlife health and habitat. These considerations take sustainability arguments into sometimes contentious territory because people have different social and cultural values. What is socially acceptable in one nation or to one segment of society may be socially unacceptable to another. These differences must be clearly defined when evaluating sustainability.

More specifically defined, sustainable agriculture is any suite of agronomic practices that are

- economically viable,
- environmentally safe, and
- socially acceptable.

These elements provide an operational definition of agricultural sustainability. There are many ways to blend them into an overarching definition, and many authors – and policymakers – have done so. In 1990, for example, the US Congress defined sustainable agriculture as "an integrated system of plant and animal production practices having a site-specific application that will, over the long term:

- satisfy human food and fiber needs,
- enhance environmental quality and the natural resource base on which the agricultural economy depends,
- make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls,
- sustain the economic viability of farm operations, and
- enhance the quality of life for farmers and society as a whole."

One should always keep in mind the subjective nature of these constructs, however, and that sustainability is a relative term. Although it may be difficult to foresee what ultimately will be sustainable, one can usually judge whether one set of practices is likely to be more sustainable than another at a societal scale.

Development of the Sustainability Concept

Sustainable agriculture as a descriptive term evolved toward common usage in the United States in the early 1980s as a mixture of concepts, ideas, values, and development direction that many believe is a vision for what agriculture should be. As with most visions, it has had strong impetus from a small group of critics of the conventional paradigm, notably Wes Jackson in his 1980 book *New Roots for Agriculture* (Jackson, 1980), and by 1984 the term was in widespread use. By 1991, the term had been fully "legitimized" as evidenced in the National Research Council's (NRC) documentation of the many U.S. Department of Agriculture, university, and other programs in sustainable agriculture throughout the United States and abroad (NRC, 1991).

Many of the philosophical roots for sustainable agriculture that coalesced in the 1980s were traced from as far back as the Greek and Roman philosophers by Harwood (1990). The public agenda debate over sustainable agriculture of the 1980s embodied many of the older concepts but certainly not any one in its entirety. The notion of land ownership, cultivation, and personal connectedness to the land as a grounding for personal responsibility, morality, and sense of purpose can be found in the writings of English philosopher John Locke in the 1600s, those of Thomas Jefferson in the late 1700s, and recently in the poetry and books of Wendell Berry. Wes Jackson applied the concept to the 1990s in his book *Becoming Native to This Place* (Jackson, 1994), which focused on homecoming and on being native, i.e., developing an ecological literacy in a holistic sense, through being, living, and having personal experience with one's ecosystem. The values in having a sense of place, a connectedness, have changed from an earlier focus on moral and political values; these values have nevertheless become a part of the vision.

The concepts of conservation and natural resource preservation are woven throughout sustainability, derived from the earlier work of Aldo Leopold and others. Louis Bromfield in *Malabar Farm* (Balantine, 1947) and the prolific writings of Edward Faulkner, beginning with *Plowman's Folly* (Faulkner, 1943), had major impact. Faulkner's work had a significant influence on both JI Rodale and his son, Robert, who carried the notion further to one of regenerative agriculture (Rodale, 1983), defined as an agriculture which not only maintains the natural resource base but also restores and increases its productive potential.

Organic, biodynamic, and the many schools of thought and broad literature surrounding these terms made major contributions to the sustainability debate. The "humus farming" school, with its focus on management of soil organic matter and the interconnectedness of soil health, plant health, and that of animals and humans, was a foundation for this philosophy. Several schools of organic practice originated in Europe and England, but the movement was popularized as organic agriculture in the United States by JI Rodale in his widely read book *Pay Dirt* (Rodale, 1945).

Each of these philosophical roots contained the notion of a holistic structure of agriculture – an interconnectedness between people, other living organisms, and the soil. Organic agriculture today continues a focus on "the living soil," on optimizing the use of biological processes and on avoiding the use of synthetic chemicals and fertilizers. Most sustainable advocates agree with a biological focus and hope to reduce but not necessarily eliminate chemical use.

Major impetus for the new vision also evolved from the 1960s era of intense agricultural development. The magnitude of the chemical revolution with its newly available pesticides, herbicides, and fungicides plus the availability of concentrated, more easily handled fertilizers very much narrowed the global agricultural development focus to the high productivity of major cereal crops. The very real specter of massive starvation in China and hunger and food deficits in India, Bangladesh, and many other countries drove the agricultural development of the 1960s. The low cost of oil and gas and prosperous northern economies all combined for infrastructure development, high agricultural inputs, and the well-known green revolution of this period.

A backlash of environmental impact, embodied in Rachel Carson's *Silent Spring* (Carson, 1962), the oil crisis of the early 1970s, and concern over a growing gap between rich and poor people and nations all set the stage for the new vision. The senescence and decline of rural communities in the United States during the 1970s and 1980s added urgency to the social dimensions of agriculture. The narrowness and focus of the development debate of the 1960s set the stage for a new, broader vision that included revisiting the old and much introspection amid assurances of decreasing real food prices and near-term global food security.

The 1980s and early 1990s produced a plethora of sustainable agriculture development writings. *Future Horizons: Recent Literature in Sustainable Agriculture* (Hegyes and Francis, 1997) is an excellent review of this broad literature. The terms "low-input sustainable" and "alternative agriculture" appeared frequently and are still used by many non-American authors, particularly with reference to organic production systems. The *American Journal of Alternative Agriculture* was founded in the mid-1980s, and in 1989 the NRC's report, *Alternative Agriculture* (NRC, 1989) brought the name "alternative agriculture" to prominence. Alternative agriculture was defined, in these sources, as any food or fiber production that has:

- a more thorough incorporation of natural processes;
- reduced use of off-farm inputs, with less harm to the environment and consumers;
- a more productive use of the biological and genetic potential of plants and animals;
- a better match between cropping patterns and the physical capacity of lands; and
- an improved emphasis on conservation of soil, water, energy, and biological resources.

Alternative agriculture is not synonymous with organic agriculture (which completely avoids synthetic chemical inputs), but they share many of the same farm management practices and approaches.

The 1980s and 1990s led us from a primary focus on engineering and chemistry in agriculture toward a greater emphasis on biology – from an age of "alchemy" to the age of "algeny" (Rifkin, 1983). The adoption of the term "agroecology" signaled a reemphasized trend in holistic thinking and analysis. A shift in emphasis within the field of ecological

science toward managed ecosystems in the late 1980s added significant perspective. Entomologists, agronomists, and other scientists began to use an ecological, process-oriented approach. Stephen Gliessman's *Agroecology: Ecological Processes in Sustainable Agriculture* (Gliessman, 1998) is an example of one book on this topic. The infusion of ecological thinking has added clarity to our understanding of overlays of subsystems, of spatial hierarchies, and of how complexity makes up an agricultural ecosystem. It has added to the understanding of multiple functions of an agricultural system and to the notion of ecosystem services. Most importantly, agroecology has taken the analysis of sustainability to a process level of understanding that allows us to understand gradients of change in production systems across time and space.

Giving greater voice in development direction to farmers has been another significant dimension in sustainable agriculture. It was very apparent that much, if not most, of alternative agriculture had originated with farmers. Philosophical direction had been heavily influenced by farmer-writers. The resurgence of on-farm research in the developing world of the 1970s (Harwood, 1979) was followed by similar emphasis with farmer collaborators in the United States in the 1980s and 1990s. Farmers have been increasingly invited to serve on steering committees and research grant award committees for sustainable agriculture projects.

Farm family well-being and that of their rural communities has been another major area of merger and inclusion in the sustainability vision. The many links and interdependencies between human and community development and sustainable agriculture are being considered in the structure and design of food systems. In 1986, Dahlberg, presented a comprehensive and thought-provoking collection of writings on social, economic, and structural issues (Dahlberg, 1986). The importance of farm size, community interaction, and the global structure of the food system is seen to be critical to both social and economic well-being. Heffernan (1997) forcefully makes the point that with globalization of capital markets and the resulting centralization of control and ownership (of both input supply and product handling and processing) have come a reduction in market competition, a shift in the balance of economic power away from the producer, and a replacement of farm-level production instability with greater macroeconomic instability in the marketplace. Much of the current literature on the social dimensions of sustainable agriculture is found in *Agriculture and Human Values*, the Journal of the Agriculture, Food, and Human Values Society.

Economic dimensions of sustainable agriculture have been typically associated with whole-farm studies comparing organic with "conventional" farms in the late 1970s, exemplified by paired comparison studies in the Midwest (e.g., Lockeretz *et al.*, 1981). Other, single-farm studies have been reported, such as in the case studies of the NRC's report *Alternative Agriculture* (NRC, 1989). Most of these studies have shown organic farms, in most years, to be as profitable as conventional farms. Most of the sustainable agricultural production research of the past two decades has focused on comparisons of crop rotations and use of cover crops and other systems component practices.

There has been a growing crescendo of voices critical of the failure to account for the side effects, the "external costs" of conventional agriculture (e.g., Roberston and Swinton, 2005). These external factors impact communities, the environment, and human health. There is increasing criticism of the emerging structure, which includes farm scale, the patterns of movement of food, and the corporate concentration of input supply and processing on a global scale. A significant component of the sustainable agriculture debate concerns the desirability of some portion of foods being of local origin, with the size of that portion related to location and community development status (Shuman, 1998; Harwood, 1998).

Scientific projections for the required incremental and transformative changes needed in US agriculture in the coming decades are based on the integration of biological, social and economic interactions across a range of scales to provide the necessary increase in resource use efficiency to supply the multitude of products and services expected by society. Robustness at varying scales, based on productivity, resilience and resistance to shock must be increasingly incorporated into systems design (NRC, 2010).

Long-standing schools of thought and practice in agriculture, influenced by changes in science and technology and in food system structure, have thus provided much of the content of the present-day sustainable agriculture agenda or vision. The amalgam of ideas and component factors has provided an extremely rich background from which development direction can be modified. The breadth of the agenda and the level of dissatisfaction with the current system point to a very major underlying problem of the global, monetarily driven process that is directing and fueling current change.

Most of the sustainable agricultural debate concerns differences in goals and in ethical and value dimensions between farmers, agriculture as a sector, and national, international, and global interests as pointed out so clearly by Dahlberg (1985). Many of the resources used and managed by agriculture and many of the services and outputs from agriculture that are critical to ecological and human well-being lie outside the monetary process that is currently driving global agriculture. If the marketplace does not put value on those dimensions and the political process either cannot or will not value them, there is a high level of disarray. Much of today's sustainable agriculture debate revolves around these value differences. Many of these sustainability issues are deeply imbedded in the current public debate over genetically modified organisms.

Indicators of Sustainability

What agricultural practices are sustainable? This is an area of intensive research. Sustainable practices must meet the three criteria defined *in* Elements of Sustainability: They must be economically viable, environmentally safe, and socially acceptable. There is no single prescription for sustainability; sustainable practices will vary by cropping system, local environment, and socioeconomic system. Nonetheless, emerging research results suggest that locally sustainable systems tend to be more resource conservative than less sustainable systems and tend to rely less on external subsidies and more on internal ecosystem services.

Resource Conservation

Resource conservation means that those resources on which sustainability depends are conserved and even enhanced by agronomic management. Soil organic matter is a good example of an ecosystem resource that is easily reduced without effective management. Soil organic matter declines rapidly in almost all cropping systems following initial cultivation – typically to 40–60% of original values within a few decades. However, soil organic matter is a valuable resource, providing habitat and energy for soil organisms, a soil structure favorable for plant growth and water retention, and a chemical structure favorable for nutrient retention (Robertson and Grandy, 2006).

The loss of soil organic matter is often associated with a need for greater external inputs. Cropping practices that conserve or enhance soil organic matter buildup will invariably enhance the environmental and often the economic sustainability of cropping systems. Crops grown in high-organic matter soils have a better water and nutrient environment than similar crops grown in soils that are depleted in organic matter, and thus they may require fewer external inputs for the same productivity. Additionally, less soil erosion and lower runoff from high-organic matter soils better protects downstream environments from agronomic impact. Therefore, cropping practices that conserve soil organic matter can be considered more sustainable than those that do not.

Often, however, there are trade-offs that require any specific conservation effort to be evaluated in the overall context of sustainability. For example, conservation tillage typically slows or stops soil organic matter loss and thus can be considered a resource-conserving, sustainable cropping practice. However, tillage controls weeds in cropping systems, and in the absence of tillage weed control is typically achieved with herbicides, which have environmental and economic costs different from those of tillage. Is the maintenance of soil organic matter as sustainable in light of a more intensive reliance on herbicides? Ideally, such trade-offs can be minimized. For example, winter cover cropping can also reduce soil organic matter loss and additionally can reduce nitrate leaching and suppress weeds, without the need for additional herbicide (Snapp *et al.*, 2005). Each cropping practice must be evaluated in a whole-system context to adequately evaluate its contribution to a system's sustainability.

Ecosystem Services

Ecosystem services are the benefits people obtain from ecosystems, often categorized as provisioning, regulating, cultural, and supporting (Millennium Ecosystem Assessment, 2003). Unmanaged systems provide such services as a matter of course. Farms are historically valued for their provisioning services – food, fuel, and fiber production – but in fact provide a whole host of other services ranging from pollination and disease regulation (regulating services) to soil and nutrient conservation (supporting services) to aesthetic amenities and a sense of place (cultural services) (Swinton *et al.*, 2007). Ecosystem services are integral to the functioning of healthy ecosystems.

In modern cropping systems many services provided by the original ecosystem before its conversion to agriculture have been suppressed or ignored in favor of services provided by external inputs. In a nitrogen-poor native ecosystem, for example, biological nitrogen fixation by native legumes such as clover (*Trifolium* spp.) might be a principal source of fixed nitrogen; modern cropping systems instead rely almost exclusively on industrially fixed nitrogen provided as inorganic fertilizer (Robertson and Vitousek, 2009). In an unmanaged system, insect herbivory is suppressed largely by trophic and structural complexity that enables insect and vertebrate predators to keep plant pests at bay. In most modern systems insect pests are controlled with insecticides, which also kill insect predators. Managing a cropping system with legumes or with greater plant diversity (either within fields or in the landscape) would allow the ecosystem to provide more of the services now provided via external inputs. Legume cover crops can reduce the need for external nitrogen, and greater plant diversity can provide the structural complexity and refugia needed to support predator populations in otherwise monospecific landscapes.

Just as for practices intended to enhance resource conservation, practices established to reintroduce or enhance existing ecosystem services need to be evaluated on the basis of their total net contribution to sustainability. Although nitrogen fixation by legumes can lower the need for fertilizer inputs and benefit soil organic matter buildup as well as provide winter habitat for predaceous insects, there is little evidence that legume-fixed nitrogen is conserved more tightly than fertilizer-derived nitrogen. Thus, there may be no downstream environmental benefit associated with this ecosystem service. Likewise, animal manure produced on-farm and recycled back to the field may be less conserved than fertilizer nitrogen if the manure is added at a time of low plant nutrient demand. Ongoing research is helping to identify ways in which management can add ecosystem services that both enhance resource availability and reduce the environmental costs of agriculture. At the societal scale there is ongoing debate on how to value services provided by farms to their neighboring communities (Robertson and Swinton, 2005). Should farmers be compensated for managing their land in ways that provide services to local, regional, and national communities? Such payments occur on a small scale in some parts of Canada, the United States, and Europe today.

Examples of Sustainable Cropping Systems

Bush-Fallow Rotation Systems

Perhaps the best documented example of a locally sustainable cropping system is the bush-fallow rotation, also known as swidden and slash and burn agriculture, indigenous to many cultures before the advent of continuous cropping systems several hundred years ago, and still evident in the humid tropics today. In the absence of population change, the bush-fallow system allows a tract of forest or savanna to provide food with few subsidies other than human labor.

In these systems, a small section of native vegetation is cut and cropped. Crop nutrient needs are met by the

decomposition of soil organic matter and perhaps by leguminous crops. Sufficient pest protection is provided by crop rotation, complex crop mixtures, and the proximity of fields to native vegetation. Weed suppression is performed by hand.

Once soil nutrients are depleted to levels that significantly compromise crop productivity, the plot is abandoned to "bush fallow" and another plot is cleared from native vegetation and cropped. Meanwhile, the now-abandoned but newly fallowed plot is undergoing secondary succession with attendant recovery of soil nutrient stores. By the time several more plots have been sequentially cut, cropped, and fallowed, the original plot will have recovered much of its original fertility and be ready to be cleared and cropped again.

Such agronomic systems are sustainable indefinitely as long as each cropped area is allowed sufficient time to recover its original fertility. However, when land becomes scarce because of development or population growth or both, the system can quickly fail. Native vegetation brought out of fallow too quickly will provide soil fertility for only a portion of the former cropping periods, so the crop portion of the rotation will either be shorter or yield less, forcing more of the native vegetation to be brought out of fallow earlier than planned in order to feed a growing population. Eventually, little native fallow will remain and crops will be grown continuously on soils that now lack much of their former fertility. One of today's greatest agronomic challenges is providing nitrogen and phosphorus to cropping systems that until recently have been in bush-fallow rotations, especially in sub-Saharan Africa where fertilizer is largely unavailable and most food is grown on small holdings of a few hectares. The maintenance of soil quality and adequate levels of soil organic matter to provide it are major concerns of tropical agronomists. Promising recent advances include the incorporation of more biodiverse rotations with modest fertilizer inputs to provide services elsewhere provided by intensive external inputs (Snapp *et al.*, 2010).

Multifunctional Agriculture

The successor to simple bush-fallow systems is mixed farming systems that have several production enterprises of different herbaceous crops, trees, animals, or combinations of crops and animals. In less developed or unstable economies requiring a high level of local, community, and farm family self-reliance, the production of a wide array of goods was primarily to meet family and local market needs and to ensure a year-round supply of food, fuel, and building materials. Farm and landscape-level diversity optimized stability within local environments and increased the resiliency of the system to a wide variety of disturbances. The diversity of land use provided a wide range of ecosystem services, including precipitation management, groundwater recharge, wildlife habitat, an environment usually conducive to adequate pest-predator balance, and some mitigation of harsh climatic conditions. The mixed plant community provided shade, wind protection, privacy, and many other, often seasonal, assorted products and services. This range of outputs has recently been termed the multifunctional character of agricultural land (Food and Agriculture Organization, 1999; Boody *et al.*, 2005) shown in **Figure 1.**

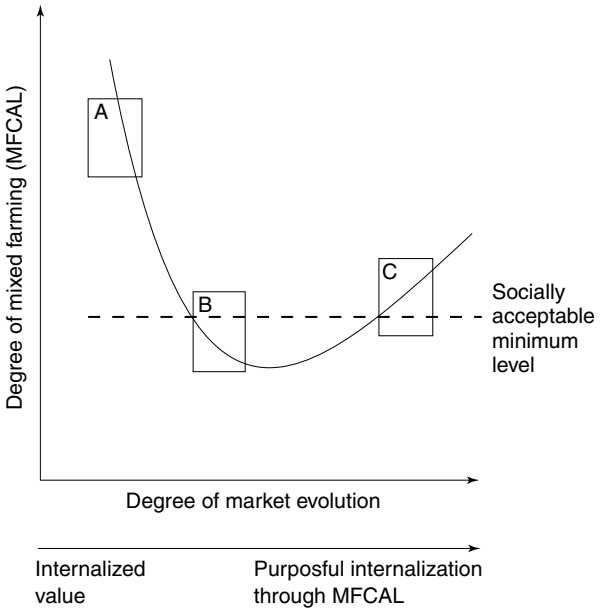


Figure 1 Changes in the purposeful development of the multifunctional character of agricultural land (MFCAL) as markets evolve (adapted from Fresco L, personal communication).

As infrastructure and markets develop, the need for a broad range of products and services decreases. When the costs of adverse environmental impacts such as groundwater contamination are not internalized, or when farmers are not rewarded for ecosystem services that their farms provide to the community or region, they do not include such values in their farm enterprise unless they are motivated and willing to make an altruistic contribution. Many farmers, in fact, do this now, but ultimately the more narrowly focused economic marketplace rules. Today's farms in highly developed economies frequently have a level of product and land use specialization that is well below an acceptable standard for long-term environmental and resource sustainability (**Figure 1, Box B**). In other words, the production base, the environment, and its ecosystem have not been stabilized and are being degraded. With continued market evolution, farmers may be increasingly compensated for the full range of ecosystem services as well as actual product output that they provide (Shuman, 1998; Soule and Piper, 1992; Robertson and Swinton, 2005) (**Figure 1, Box C**).

A more immediate incentive is to add crops and/or livestock to provide integration efficiencies. These efficiencies include the increase in yield of one crop following another, the savings in nutrient inputs, or the reduction in pest control costs (**Table 1**).

An increasing amount of information on the efficiencies of specific technologies for integration is becoming available in the scientific literature. The reduction in input requirements is often a key part. There is less direct research information on the relationship of many of these practices for mitigating environmental harm. An exception is the wealth of data on reduced soil erosion as a result of reduced or zero tillage. Currently, the predictive models of loss of pesticides, nutrients, or crop or animal residues are rudimentary. Direct

Table 1 Michigan maize, soybean, and wheat rotation efficiencies

System, advantages
Maize after beans
30 kg ha^{-1} nitrogen credit
No rootworm scouting or control costs
6–10% yield advantage
Maize after soybeans dry beans wheat (Michigan, second, third year of rotation)
No nitrogen credit (since maize follows wheat)
No rootworm control costs
Window for perennial weed control (either mechanical or chemical)
Greater than 10% yield advantage (because of the preceding bean/wheat sequence)
Maize after wheat plus frost-seeded clover
40 kg ha^{-1} nitrogen credit (60–70 kg N ha^{-1} with presidedress nitrogen test)
No rootworm control costs
At least 15% yield advantage
30–50% yield advantage if the farm is organic, where maize-after-maize is not advisable

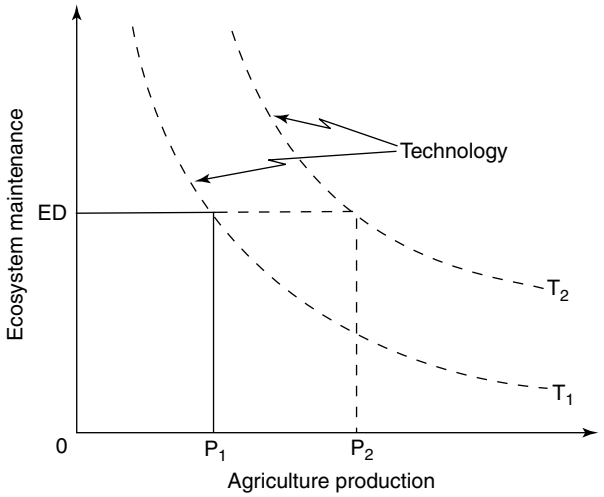


Figure 2 Effect of technology selection on the relationship between agricultural production and ecosystem maintenance. ED, environmental disturbances; T_1, inappropriate technology; T_2, more appropriate technology; P_1, agricultural production with technology T_1; P_2, agricultural production with technology T_2 (adapted from NRC (1993)).

measurements of loss from alternative rotations and use of cover crops are very difficult, expensive, and location specific. These rotation and cover crop practices are widely acknowledged as fundamental to sustainability. Their efficiencies are being quantified with respect to yield, input reduction, and soil quality and the prevention of soil loss. Michigan data show, for instance, that wheat in rotation loses less than 20 kg N ha^{-1} year^{-1} via groundwater leaching. Well-fertilized continuous corn averages 50 kg N ha^{-1} year^{-1}. Most U.S. farmers use at least a two-crop rotation.

Animal integration in crop systems is declining in the United States. Poultry and turkeys are increasingly produced in specialized production facilities not located on the farms where their feed is produced. They are usually located in areas where agricultural land is available for manure application, often on a contract "disposal" basis. The level of crop or animal diversity that is appropriate on a farm to balance the market forces for specialization with the need for biological efficiency and ecosystem maintenance is very situation specific. As enterprise integration increases with an effective level of appropriate technologies and their effective management (**Figure 2**, technology T_2), agricultural output can be maintained at a much higher level for a given amount of ecosystem disturbance. In other words, sustainable agriculture can maintain productivity at a much lower level of ecosystem disturbance. Very large-scale operations tend to have less diversity, in part because of the greater difficulty of managing diverse enterprises. Crop and animal management requires numerous and often frequent decisions to be made as conditions change that are often stimulated by visual, difficult to measure changes. The frequent presence and sensitivity of the manager, the experience in production management, and the ability to make decisions place limits on the scale of highly diversified operations. Every farm owner experiences this tension.

On a global scale, under conditions of high population density, small farms, and the need for producing a wide array of products in often marginal production environments, a very

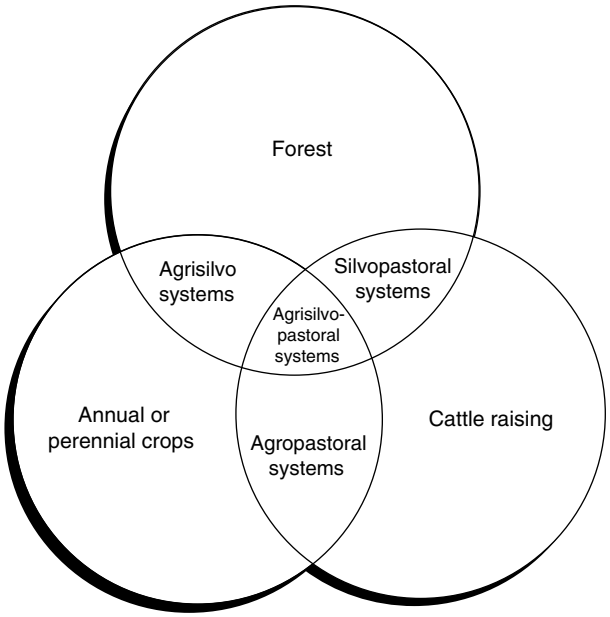


Figure 3 Combinations of enterprises in mixed farming systems (redrawn from Serrão EAS and Homma AKA (1993) Brazil. In: National Research Council. (eds.) *Sustainable Agriculture and the Environment in the Humid Tropics*, pp. 265– 351. Washington, DC: National Academy Press, with permission from Dr. Scott).

diverse type of farm enterprise mix is common. Trees become a very important part of farm productivity in the higher rainfall areas where they are a part of the native vegetation. Animals are more often than not a part of the enterprise because they consume crop residues and add significantly to overall productivity (**Figure 3**). In most developing countries there is very little area of undisturbed forest, and in only a few in which large tracts of land remain is there cattle only on farms. In the humid tropics the mixture of crops, trees, and animals

(agrisilvo-pastoral systems) represent the great majority of farms. Where human population is relatively high (>300–500 persons per square kilometer) in rural areas, if there is poverty combined with modest levels of rainfall (less than 1500 mm $year^{-1}$) and/or cool temperatures for part of the year, fuel for cooking and heating becomes a problem. Resource degradation and loss of production potential often occur as the standing stocks of carbon (particularly in trees) and eventually the soil carbon stocks are reduced as crop and animal residues are burned. The system rapidly loses crop nutrient holding and recycling capacity, and its ability to intercept and retain rainfall decreases. Most developing country mixed farms have a larger portion of cash grain crops.

Many countries in eastern Africa (Kenya, Ethiopia, Rwanda, Tanzania, and Uganda) have these exacting conditions. Rural well-being depends on small farms (between 0.5 and 1.5 ha) supporting families by providing food, fuel, building materials and cash income (**Figure 4**). Ruminant animals (often dairy cows or dairy goats) are a key part of that production. The productivity of the animals and their welfare through the dry season depends on fodder trees and other perennials, which are nearly always mixed into the landscape in intricate patterns. There is always a high diversity of food and fodder crop species, which protect the soil from erosion in sometimes heavy rainstorms and give the systems overall stability and resilience. Improved varieties of both trees and crop plants and improved animal breeds are critical to ongoing improvements (Pye-Smith, 2010).



Figure 4 In Kenya's densely populated Central Province, multifunctional agriculture includes grain and tree crops as well as ruminant animals such as dairy cows that depend on trees for fodder (credit: ICRAF/C. Pye-Smith).

Conclusions

Our global food and fiber production systems are undergoing an enormous transformation, driven by rapid advances in the sciences of biotechnology, engineering, and food processing and chemistry. The increasing centralization in the manufacture of agricultural inputs and in the collection and processing of food is causing huge economic and social change. Many social and political values are being challenged. The global marketplace forces product uniformity and the geographical concentration of its production are driving farmers toward a level of specialization that results in farms with much less diverse crop and animal enterprises on the landscape than that desirable for the maintenance of many ecosystem services. Markets have not yet matured to adequately value these services, especially those that affect environmental quality, nor in most places have governments established disincentives in order to protect them. The greatest challenge to sustainability, with its many economic, environmental, and social dimensions, is the lack of public awareness, vision, and will to implement necessary changes. In some cases, research is needed to clarify the value of alternative strategies and to provide additional options for sustainable management.

In many if not most places food is not being produced sustainably, i.e., in a manner that is economically viable, environmentally benign, and socially acceptable to many who are affected by its production. On the other hand, new research is showing that sustainable cropping systems can be designed to operate effectively, using ecological knowledge to substitute for some of the management options now provided by external inputs, and in a way that has a less adverse environmental and social impact than conventional management. A better understanding of agricultural ecosystems and the services they provide, coupled with new technology, may provide the advances needed to feed – sustainably – a burgeoning global population in the decades to come.

See also: Agriculture, Industrialized. Agriculture, Traditional. Biodiversity as a Commodity. Ecosystem Services. Herbicides. Nitrogen, Nitrogen Cycle. Pesticides, Uses and Effects of. Soil Conservation

References

Altieri M (1995) *Agroecology: The Science of Sustainable Agriculture*. Boulder, CO: Westview Press.

Boody G, Vondracek B, Andwo DA, *et al.* (2005) Multifunctional agriculture in the United States. *BioScience* 55: 27–38.

Carroll CR, Vandermeer JH, and Rosset P (1990) *Agroecology*. New York: McGraw-Hill.

Carson R (1962) *Silent Spring*. Boston: Houghton Mifflin.

Dahlberg KA (1985) Ethics, values and goals in agricultural systems and agricultural research. In: Edens TC, Fridgen C, and Battenfield SL (eds.) *Sustainable Agriculture and Integrated Farming Systems: 1984 Conference Proceedings*, pp. 202–218. East Lansing: Michigan State University Press.

Dahlberg KA (ed.) (1986) *New Directions for Agriculture and Agricultural Research: Neglected Dimensions and Emerging Alternatives*. Totowa, NJ: Rowman and Allanheld.

Faulkner E (1943) *Plowman's Folly and a Second Look*. Washington, DC: Island Press.

Food and Agricultural Organization (FAO) (1999) *Conference on the Multifunctional Character of Agricultural Land*, Rome. Maastricht, The Netherlands, 12–17 September.

Francis CA, Flora CB, and King LD (eds.) (1990) *Sustainable Agriculture in Temperate Zones*. New York: Wiley.

Gliessman SR (1998) *Agroecology: Ecological Processes in Sustainable Agriculture*. Chelsea, MI: Ann Arbor Press.

Harwood RR (1979) *Small Farm Development: Understanding and Improving Farming Systems in the Humid Tropics*. Boulder, CO: Westview Press.

Harwood RR (1990) A history of sustainable agriculture. In: Edwards CA, Lal R, Madden R, Miller R, and House G (eds.) *Sustainable Agricultural Systems*, pp. 3–19. Ankeny, IA: Soil and Water Conservation Society.

Harwood RR (1998) Directions for farm change: Bringing it all together. In: Cavigelli MA, Deming SR, Probyn LK, and Harwood RR (eds.) *Michigan Field Crop Ecology: Managing Biological Processes for Productivity and Environmental Quality*, pp. 80–86. East Lansing: Michigan State University Press. Extension Bulletin No. E2646.

Heffernan WD (1997) Domination of world agriculture by transnational corporations. In: Madden JP and Chaplow SG (eds.) *For All Generations Making World Agriculture More Sustainable*, pp. 173–181. Glendale, CA: World Sustainable Agriculture Association.

Hegyes G and Francis CA (eds.) (1997) *Future Horizons: Recent Literature in Sustainable Agriculture*. University of Nebraska, Lincoln: Center for Sustainable Agricultural Systems.

Jackson W (1980) *New Roots for Agriculture*. Lincoln: University of Nebraska Press.

Jackson W (1994) *Becoming Native to This Place*. Lexington: University Press of Kentucky.

Lockeretz W, Shearer G, and Kohl DH (1981) Organic farming in the Corn Belt. *Science* 211: 540–547.

Millennium Ecosystem Assessment (2003) Ecosystems and their services. In: *Ecosystems and Human Well-Being: A Framework for Assessment*, pp. 49–70. Washington, DC: Island Press.

Nassauer JI and Opdam P (2008) Design in science: Extending the landscape ecology paradigm. *Landscape Ecology* 23: 633–644.

National Research Council (NRC) (1989) *Alternative Agriculture*. Washington, DC: National Academy Press.

National Research Council (NRC) (1991) *Sustainable Agricultural Research and Education in the Field: A Proceedings*. Washington, DC: National Academy Press.

National Research Council (NRC) (2010) *Toward Sustainable Agricultural Systems in the 21st Century*. Washington, DC: National Academy Press.

Pye-Smith C (2010) Fodder for a better future: How agroforestry is helping to transform the lives of smallholder dairy farmers in East Africa. *ICRAF Trees for Change no. 6*. Nairobi: World Agroforestry Centre.

Rifkin J (1983) *Algeny*. New York: Viking.

Robertson GP and Grandy AS (2006) Soil system management in temperate regions. In: Uphoff N, Ball AS, Fernandes E, *et al.* (eds.) *Biological Approaches to Sustainable Soil Systems*, pp. 27–39. Boca Raton, FL: CRC Press.

Robertson GP and Swinton SM (2005) Reconciling agricultural productivity and environmental integrity: A grand challenge for agriculture. *Frontiers in Ecology and the Environment* 3: 38–46.

Robertson GP and Vitousek PM (2009) Nitrogen in agriculture: Balancing the cost of an essential resource. *Annual Review of Environment and Resources* 34: 97–125.

Rodale JI (1945) *Pay Dirt*. Emmaus, PA: Rodale Press.

Rodale R (1983) Breaking new ground: The search for sustainable agriculture. *The Futurist* 1: 15–20.

Serrão EAS and Homma AKA (1993) Brazil. In: National Research Council. (eds.) *Sustainable Agriculture and the Environment in the Humid Tropics*, pp. 265–351. Washington, DC: National Academy Press.

Shuman MH (1998) *Going Local: Creating Self-Reliant Communities in a Global Age*. New York: Free Press.

Snapp SS, Swinton SM, Labarta R, *et al.* (2005) Evaluating benefits and costs of cover crops for cropping system niches. *Agronomy Journal* 97: 322–332.

Snapp SS, Blackie MJ, Gilbert RA, Bezner-Kerr B, and Kanyama-Phiri GY (2010) Biodiversity can support a greener revolution in Africa. *Proceedings of the National Academy of Sciences of the USA* 107: 20 840–20 845.

Soule JD and Piper JK (1992) *Farming in Nature's Image: An Ecological Approach to Agriculture*. Washington, DC: Island Press.

Swinton SM, Lupi F, Robertson GP, and Hamilton SK (2007) Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. *Ecological Economics* 64: 245–252.

Agriculture, Traditional

Miguel A Altieri, University of California, Berkeley, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 109–118, © 2001, Elsevier Inc.

Glossary

Agroecosystem A simplified natural ecosystem subjected to exploitation for purposes of food and fiber production.
Biodiversity Diversity of microbial, animal, and plant species in an ecosystem that performs distinct ecological functions and services.
Ethnoecology Study of the various forms of traditional environmental knowledge that are characteristic of specific ethnic groups and that translate into natural resource management.
Polyculture Intensive growing of two or more crops either simultaneously or in sequence on the same piece of land.
Sustainable agriculture Form of agriculture that is environmentally sound, culturally sensitive, socially acceptable, and economically viable.
Traditional agriculture Indigenous form of ecologically based agriculture resulting from the coevolution of local cultural and environmental systems.

Introduction

One of the salient features of traditional farming systems throughout the developing world is their high degree of biodiversity. These traditional systems have emerged over centuries of cultural and biological evolution and represent the accumulated experiences of indigenous farmers interacting with the environment without access to external inputs, capital, or modern scientific knowledge (Chang, 1977; Grigg, 1974). Using inventive self-reliance, experiential knowledge, and locally available resources, traditional farmers have often developed farming systems that generate sustained yields (Harwood, 1979). In Latin America alone, more than two and a half million hectares are under traditional agriculture in the form of raised fields, polycultures, and agro-forestry systems, documenting the successful adaptation of these farming practices to difficult environments (Altieri, 1991).

Many of these traditional agroecosystems, still found throughout the Andes, Meso-America, and the lowland tropics, constitute major *in situ* repositories of both crop and wild plant germplasm. From an agroecological perspective, these agroecosystems can be seen as a continuum of integrated farming units and natural or semi-natural ecosystems where plant gathering and crop production are actively pursued. Plant resources are directly dependent on management by human groups; thus, both species and genetic diversity have evolved in part under the influence of farming practices shaped by particular cultures and the forms of sophisticated knowledge they represent (Nabhan, 1983).

Perhaps the key to understanding how traditional farmers maintain, preserve, and manage biodiversity is to recognize the complexity of their production systems. Today, it is widely accepted that indigenous knowledge is a powerful resource in its own right and complementary to knowledge available from Western scientific sources. Therefore, in studying such systems, it is not possible to separate the study of agricultural biodiversity from the study of the culture that nurtures it.

This article explains the features of the biodiversity inherent in traditional agroecosystems, and the ways in which farmers apply local knowledge to manage such biodiversity to satisfy subsistence needs and to obtain ecological services. Traditional agriculture is rapidly disappearing in the face of major social, political, and economic changes. The conservation and management of this agrobiodiversity will be possible only if they are linked to the preservation of the cultural diversity and economic viability of the local farming populations.

Biodiversity Features of Traditional Agroecosystems

Traditional farming systems commonly support a high degree of plant diversity in the form of polycultures and/or agroforestry patterns (Chang, 1977; Clawson, 1985). This strategy of minimizing risk by planting several species and varieties of crops stabilizes yields over the long term, promotes diet diversity, and maximizes returns even with low levels of technology and limited resources (Harwood, 1979). Such biodiverse farms are endowed with nutrient-enriching plants, insect predators, pollinators, nitrogen-fixing and nitrogen-decomposing bacteria, and a variety of other organisms that perform various beneficial ecological functions.

Traditional multiple-cropping systems provide as much as 15–20 percent of the world food supply (Francis, 1986). Polycultures constitute at least 80 percent of the cultivated area in West Africa and predominate in other parts of Africa as well (Norman, 1979). At the same time, much of the production of staple crops in the Latin American tropics occurs in polycultures. More than 40 percent of the cassava, 60 percent of the maize, and 80 percent of the beans in the region grow in mixtures with each other or other crops (Francis, 1986). Polycultures are also very common in parts of Asia where upland rice, sorghum, millet, maize, and irrigated wheat are the staple crops. Lowland (flooded) rice is generally grown as a monoculture, but in some areas of Southeast Asia farmers

build raised beds to produce dryland crops amid strips of rice (Beets, 1982).

Tropical agroecosystems composed of agricultural and fallow fields, complex home gardens, and agroforestry plots commonly contain well over 100 plant species per field, and these are used as construction materials, firewood, tools, medicines, livestock feed, and human food. Examples include multiple-use agroforestry systems managed by the Huastecs and Lacondones in Mexico, the Bora and Kayapó Indians in the Amazon River basin, and many other ethnic groups who incorporate trees into their production systems (Wilken, 1977).

In the Latin American tropics, home gardens are a highly efficient form of land use, incorporating a variety of crops with different growth habits. The result is a structure similar to that of tropical forests, with diverse species and a layered physical configuration (Denevan *et al.*, 1984). In Mexico, for example, Huastec Indians manage a number of fields, gardens, and forest plots that may harbor a total of about 300 species. Small areas around their houses commonly average 80–125 useful plant species, mostly native and medicinal plants. Huastec management of the noncrop vegetation in these complex farm systems has influenced the evolution of individual plants and the distribution and composition of the crop and noncrop communities.

In these "forestlike" agricultural systems, nutrient cycles are tight and closed. In traditional shaded coffee plantations (where *Inga* and *Erythrina* are common tree species), the total nitrogen input from the decomposition of shade tree leaves, as well as from litter and symbiotic fixation, can be well over ten times higher than the net nitrogen output in the coffee harvest, which usually averages 20 kg/ha/yr. Clearly, the system amply compensates for the nitrogen loss by harvest with a subsidy from the shade trees. In Mexico, farmers encourage the growth of native leguminous trees in cultivated fields (Wilken, 1977). From Puebla and Tehuacán south through Oaxaca, farms with light to moderately dense stands of mesquite (*Prosopis* spp.), guaje (*Leucaena esculenta*), and guamuchil (*Pithecellobium* spp.) are a familiar sight. Stand density varies from fields with only a few trees to virtual forests with crops planted beneath them. A slightly different practice is found near Ostuncalco, Guatemala, where rigorously pruned sauco (*Sambucus mexicana*) stumps dot maize and potato fields. Sauco leaves and small branches are removed annually, scattered around individual crop plants, and then chopped and interred with broad hoes. Local farmers claim that crop quality and yields in the sandy volcanic soils of this region depend on the annual application of this method (Wilken, 1977).

Many traditional agroecosystems are located in centers of crop diversity, and thus contain populations of variable and adapted landraces as well as wild and weedy relatives of crops (Harlan, 1976). Clawson (1985) described several systems in which tropical farmers plant multiple varieties of each crop; this practice supports both intraspecific and interspecific diversity, and also enhances harvest security. For example, in the Andes, farmers cultivate as many as 50 potato varieties in their fields (Brush, 1982). Similarly, in Thailand and Indonesia, farmers maintain a diversity of rice varieties adapted to a wide range of environmental conditions, and they regularly exchange seeds with each other (Grigg, 1974). The resulting genetic diversity heightens resistance to diseases that attack particular strains of the crop and enables farmers to exploit different microclimates and to derive multiple nutritional and other uses from the genetic variation among the species.

Many plants within or around traditional cropping systems are wild or weedy relatives of crop plants. In fact, many farmers "sponsor" certain weeds in or around their fields that may have positive effects on soil and crops, or that serve as food, medicines, ceremonial items, teas, soil improvers, or pest repellents. In the Mexican Sierras, the Tarahumara Indians depend on edible weed seedlings or "quelites" (e.g., *Amaranthus, Chenopodium, Brassica*) in the early season from April through July, a critical period before crops mature from August through October. Weeds also serve as alternative food supplies in seasons when maize or other crops are destroyed by frequent hail storms (Bye, 1981). In barley fields, it is common for Tlaxcalan farmers to maintain *Solanum mozinianum* at levels up to 4500 plants/ha; this yields about 1300 kg of fruit, a significant contribution to agricultural subsistence (Altieri and Trujillo, 1987).

Farmers also derive other benefits from weeds, such as increased gene flow between crops and their relatives. In Mexico, when the wind pollinates maize, natural crosses occur with wild teosinte growing in the field borders, resulting in hybrid plants. Certain weeds are used directly to enhance the biological control of insect pests, as many flowering weeds attract predators and parasites of pests to their pollen and nectar. Other farmers allow weeds such as goosegrass (*Eleusine indica*) in bean fields to repel *Empoasca* leafhoppers, or wild *Lupinus* as a trap plant for the pestiferous scarab beetle (*Macrodactylus* sp.), which otherwise would attack corn (Altieri, 1993).

However, diversity is maintained not only within a cultivated area. Many farmers maintain natural vegetation adjacent to their fields, and thus obtain a significant portion of their subsistence requirements through gathering, fishing, and hunting in habitats that surround their agricultural plots. For the P'urhepecha Indians who live around Lake Pátzcuaro in Mexico, gathering is part of a complex subsistence pattern that is based on multiple uses of their natural resources. These people use at least 224 species of native and naturalized vascular plants for dietary, medicinal, household, and fuel needs (Caballero and Mapes, 1985).

Depending on the level of biodiversity of closely adjacent ecosystems, farmers accrue a variety of ecological services from surrounding natural vegetation. For example, in western Guatemala, the indigenous flora of the higher-elevation forests provide valuable native plants that serve as a source of organic matter to fertilize marginal soils, for each year farmers collect leaf litter from nearby forests and spread it over intensively cropped vegetable plots to improve tilth and water retention. Some farmers may apply as much as 40 metric tons of litter per hectare each year; rough calculations indicate that a hectare of cropped land requires the litter production of 10 ha of regularly harvested forest (Wilken, 1977).

Clearly, traditional agricultural production commonly encompasses the multiple uses of both natural and artificial ecosystems, where crop production plots and adjacent habitats are often integrated into a single agroecosystem.

The Complex Nature of Traditional Farmers' Knowledge

Ethnoecology is the study of the natural world knowledge systems of indigenous ethnic people. This knowledge has many dimensions, including linguistics, botany, zoology, craft skills, and agriculture, and is derived from the direct interaction between humans and their environment. In such a system, cognition and perception select the most adaptive and useful environmental information, and this "successful" knowledge is preserved from generation to generation through oral or experimental means. Indigenous peoples' knowledge about soils, climates, vegetation, animals, and ecosystems usually results in multidimensional productive strategies (i.e., the use of multiple ecosystems with multiple species), and these strategies generate (within certain ecological and technical limits) the food self-sufficiency of farmers in a region (Netting, 1993).

Captivated by the ecological intricacies of these traditional agricultural systems, many scientists are now beginning to show interest in them. As scientists search for ways to remedy the deficiencies of modern agriculture, they recognize that indigenous farmers' knowledge may hold vital information for the future of world agriculture. After centuries of cultural and biological evolution, these farmers have developed locally adapted, complex farming systems that have helped them to sustainably manage a variety of environments and to meet their subsistence needs, without depending on modern agricultural technologies.

For many agricultural scientists, four aspects of these traditional knowledge systems are relevant: knowledge of the environment, folk taxonomies, knowledge of farming practices, and the experimental nature of traditional knowledge (Altieri, 1987).

Knowledge of the Environment

Indigenous knowledge about the physical environment is often very detailed. Many farmers have developed traditional calendars to control the scheduling of agricultural activities, and many sow according to the phase of the moon, believing that there are lunar phases of rainfall. They also cope with climatic seasonality by utilizing weather indicators based on the phenologies of local vegetation.

Soil types, degrees of soil fertility, and land-use categories are also discriminated in detail. Soil types are commonly distinguished by color, texture, and sometimes taste. Shifting cultivators usually classify their soils based on vegetation cover. In general, peasants identify soil types based on the nature of the peasant's relationship to the land (Williams and Ortiz-Solario, 1981). Aztec soil classification systems were very complex, recognizing more than two dozen soil types identified by origin, color, texture, smell, consistency, and organic content. These soils were also ranked according to agricultural potential, which was used in both land-value evaluations and rural census. Today, Andean peasants in Coporaque, Peru, recognize four main soil classes, where each class has specific characteristics matching the most adequate cropping system (Brush, 1982).

Biological Folk Taxonomies

Many complex knowledge systems that are used by indigenous people to group together plants and animals have been well documented (Berlin *et al.*, 1973). The traditional name of a plant or animal usually reveals that organism's taxonomic status, and researchers have found that, in general, there is a good correlation between folk taxa and scientific taxa.

The classification of animals, especially insects and birds, is widespread among indigenous farmers. Insects and related arthropods have major roles as crop pests, as causes of disease, as food, and as medicinal products, in addition to their importance in local myth and folklore. In many regions of the world, agricultural pests are tolerated because they also constitute agricultural products; that is, traditional agriculturalists may consume plants and animals that would otherwise be considered pests (Brokensha *et al.*, 1980).

Ethnobotanies are the most commonly documented folk taxonomies (Alcorn, 1984). The ethnobotanical knowledge of certain campesinos in Mexico is so elaborate that the Tzeltal, P'urepecha, and Yucatan Mayans can recognize more than 1200, 900, and 500 plant species, respectively (Toledo *et al.*, 1985). Similarly, !ko bushwomen in Botswana were able to identify 206 out of 211 plants collected by researchers (Chambers, 1983), while Hanunoo swidden cultivators in the Philippines could distinguish over 1600 plant species (Grigg, 1974).

Knowledge of Farming Practices

As more scientific research is conducted, many of the traditional farming practices once regarded as primitive or misguided are being recognized as sophisticated and appropriate. For example, when confronted with specific problems of slope, flooding, drought, pests, diseases, or low soil fertility, small farmers throughout the world have developed unique management systems aimed at overcoming these constraints (Klee, 1980). In general, traditional agriculturalists have adjusted to environmental constraints by concentrating on a few characteristics and processes that incorporate the following structural and functional elements (Gliessman, 1998; Altieri and Anderson, 1986):

a. They combine high species numbers and structural diversity in time and space (through both vertical and horizontal organization of crops).
b. They exploit the full range of microenvironments (which differ in soil, water, temperature, altitude, slope, fertility, etc.) within a field or region.
c. They maintain closed cycles of materials and wastes through effective recycling practices.
d. They rely on the complexity of biological interdependencies, resulting in some degree of biological pest suppression.
e. They rely on local resources plus human and animal energy, thereby using low levels of technology input.
f. They rely on local varieties of crops and incorporate the use of wild plants and animals. Production is usually for local consumption. The level of income is low; thus the influence of noneconomic factors on decision making is substantial.

The Experimental Nature of Traditional Knowledge

The strength of traditional people's knowledge is that it is based not only on acute observation but also on trial and error and experimental learning. The experimental approach is very apparent in the selection of seed varieties for specific environments, but it is also implicit in the testing of new cultivation methods to overcome particular biological or socioeconomic constraints. In fact, Chambers (1983) argued that farmers often achieve a richness of observation and a fineness of discrimination that would be accessible to Western scientists only through long and detailed measurement and computation.

Yet only recently has some of this traditional knowledge been described and documented by researchers. The evidence suggests that the finest discrimination develops in communities where the environments have great physical and biological diversity and/or in communities living near the margins of survival (Chambers, 1983). Also, older community members possess greater, more detailed knowledge than younger members (Klee, 1980).

The Ecological Services of Biodiversity in Traditional Agroecosystems

In traditional agroecosystems, complex and diversified cropping systems are vital because the interactions among crops, animals, insects, and trees result in beneficial synergisms that optimize soil fertility, pest control, and productivity (Altieri, 1995; Harwood, 1979; Richards, 1985). Among the ecological services are the following:

1. By interplanting, farmers take advantage of the capacity of cropping systems to reuse their own stored nutrients. The tendency of some crops to deplete the soil is counteracted by interplanting other crops that enrich the soil with organic matter. Soil nitrogen, for example, can be increased by incorporating legumes in the crop mixture, and phosphorus assimilation can be enhanced by growing crops with mycorrhizal associations.
2. The complex structure of traditional agroecosystems minimizes crop loss to insect pests through a variety of biological mechanisms. The intercropping of diverse plant species provides habitats for the natural enemies of insect pests as well as alternative host plants for pests. For example, a crop may be planted as a diversionary host to protect other more susceptible or more economically valuable crops from serious damage. The diversity of crops grown simultaneously in polycultures helps prevent the buildup of pests on the comparatively isolated plants of each species. Where shifting cultivation is practiced, the clearing of small plots from secondary forest vegetation also permits the easy migration of natural pest predators from the surrounding forest.
3. Increasing the species and/or genetic diversity of cropping systems is a key strategy to minimize losses from plant diseases and nematodes (types of round-worms that are among the most widespread and damaging of agricultural pests). The mixing of different crop species or varieties can delay the onset of diseases, reduce the spread of disease-carrying spores, and modify environmental conditions such as humidity, light, temperature, and air movement so that they are less favorable to the spread of certain diseases.
4. Many intercropping systems prevent competition from weeds, chiefly because the large leaf areas of their complex canopies prevent sufficient sunlight from reaching sensitive weed species. In general, the extent to which weeds present a problem depends on the type of crops and the proportion of the different species grown, their density, where they are planted, the fertility of the soil, and management practices. Weed suppression can be enhanced in intercrop systems by adding crop species that inhibit weed germination or growth. Crops such as rye, barley, wheat, tobacco, and oats release toxic substances into the environment, either through their roots or from decaying plant material. Such toxins inhibit the germination and growth of some weed species such as wild mustard (*Brassica* spp.) and poppy.
5. The integration of animals (e.g., cattle, swine, poultry) into farming systems, in addition to using them for milk, meat, and draft needs, adds another trophic level to the system, making it even more complex. Animals are fed crop residues and weeds with little negative impact on crop productivity, and this serves to turn otherwise unusable biomass into animal protein. Animals also recycle the nutrient content of plants by transforming them into manure. Furthermore, the need for animal feed broadens the crop base to include plant species that are useful for conserving soil and water (Reijntjes *et al.*, 1982). Legumes are often planted to provide quality forage, but they also improve the nitrogen content of soils. Integrated crop–livestock systems usually take the form of a crop–pasture rotation in which the pasture phase "charges" the system with nutrients and organic matter and the cropping phase "extracts" the accumulated nutrients. This balances biomass and nutrient inputs and outputs.

Preserving the Biodiversity of Traditional Agroecosystems

As many rural societies undergo the conversion from a subsistence economy to a cash agricultural economy, the loss of biodiversity in their ecosystems is mounting at an alarming rate. Because many peasants are directly linked to the market economy, external economic forces are increasingly influencing production by favoring genetically uniform crops and mechanized and/or agrochemical practices. Many landraces and wild plant relatives are being abandoned, which may cause them to become relic populations or even extinct. In some areas, land scarcity (mostly a result of uneven land distribution) has forced changes in land use and agricultural practices, which in turn have caused the disappearance of habitats that formerly maintained useful noncrop vegetation, including wild progenitors and weedy forms of crops (Altieri *et al.*, 1987).

In many parts of the world, genetic erosion is occurring at a fast pace because farmers are having to quickly change their

farming systems because of economic, technical, and social pressures. As farmers adopt high-yield modern varieties (HYVs), they often subdivide their farming systems into commercial (mostly devoted to HYVs) and subsistence sectors, growing native varieties in the latter. The greatest loss of traditional plant varieties is occurring in lowland valleys close to urban centers and markets (Brush, 1986).

Given these destructive trends, many scientists and development workers have emphasized the need for *in situ* conservation of native crop genetic resources and the environments in which they occur (Prescott-Allen and Prescott-Allen, 1981). However, most researchers believe that *in situ* preservation of landraces would require a return to or the preservation of microcosms of traditional agricultural systems, which some regard as an unacceptable and impracticable proposition (Frankel and Soulé, 1981). Nevertheless, the maintenance of traditional agroecosystems may be the only sensible strategy to preserve *in situ* repositories of crop germplasm. Although most traditional agroecosystems are undergoing some process of modernization or drastic modification, the conservation of crop genetic resources can still be integrated into agricultural development, especially in regions where rural development projects preserve the vegetation diversity of traditional agroecosystems and are anchored in the peasant rationale to utilize local resources and their intimate knowledge of the environment (Alcorn, 1984; Nabhan, 1983).

Previous recommendations for *in situ* conservation of crop germplasm emphasized the development of a system of village-level landrace custodians (a farmer curator system) whose purpose would be to continue growing a limited sample of endangered landraces native to the region (Mooney, 1983). One suggestion for preserving crop-plant diversity was for governments to set aside carefully chosen 5-by-20-km strips of land at as few as 100 sites around the world where native agriculture is still practiced (Wilkes and Wilkes, 1972). But given the increasing impoverishment and lack of income-generating alternatives for many rural populations in less developed countries, a proposition of this kind is clearly unrealistic since it fails to address the subsistence needs of these populations. In many areas where the urgent short-term goal of the local people is survival, diverting the limited land available for conservation purposes per se might prove totally inappropriate. A more feasible approach would be to support sustainable farming systems that incorporate native crops and wild/weedy relatives within and around production fields, as well as appropriate technologies aimed at upgrading food production for self-sufficiency (Altieri and Merrick, 1987). Such efforts would ensure that germplasm preservation remains linked to the economic and agricultural viability of local populations.

If biodiversity conservation is to succeed among small farmers, conservation goals and rural development efforts must be integrated to give equal importance to local resource conservation, food self-sufficiency, and equitable market participation. Any attempt at *in situ* crop genetic conservation must struggle to preserve the agroecosystem in which these resources occur (Nabhan, 1983). In the same vein, preservation of traditional agroecosystems cannot be achieved unless the sociocultural stability of the local community is also assured (Altieri, 1995).

An examination of effective grassroots rural development programs in less developed countries suggests that the process of agricultural improvement must (a) utilize and promote autochthonous knowledge and resource-efficient technologies, (b) emphasize use of local and indigenous resources, including valuable crop germplasm as well as essentials like firewood resources and medicinal plants, and (c) remain a self-contained, village-based effort with the active participation of the local people (Altieri, 1987). The subsidizing of a peasant agricultural system with external resources (e.g., pesticides, fertilizers, and irrigation water) can bring high levels of productivity, but such a system would then be sustainable only at high external cost and would depend on the uninterrupted availability of commercial inputs. In contrast, an agricultural strategy based on a diversity of plants and cropping systems can bring moderate to high levels of productivity through the manipulation and exploitation of the resources internal to the farm and can be sustainable at a much lower cost and for a longer period of time (Gliessman, 1998).

Using Biodiversity-Based Strategies to Support Traditional Agriculture

By understanding the common features of traditional agriculture, such as the capacity to bear risk, the use of biological folk taxonomies, and the production efficiencies derived from multiple and symbiotic crop mixtures, agricultural scientists have been able to develop technologies that support the needs and circumstances of specific groups. While subsistence farming generally lacks the potential for producing a meaningful marketable surplus, it does ensure food security. Many scientists wrongly believe that traditional systems do not produce more because hand tools and draft animals put a ceiling on productivity. However, where productivity is low, the cause appears to be social, not technical. When the subsistence farmer succeeds in providing food, there is no pressure to innovate or to enhance yields. Yet research shows that increased productivity is possible when traditional crop and animal combinations are adjusted and when labor and local resources are used more efficiently (Pretty, 1995).

As the inability of the Green Revolution to improve production and farm incomes for the very poor became apparent, growing enthusiasm for established, traditional agricultural practices generated a renewed quest in the developing world for affordable, productive, and ecologically sound technologies that could enhance small farm productivity while conserving resources. In the Andean altiplano, development workers and farmers have reconstructed a 3000-year-old indigenous farming system at an altitude of almost 4000 m. These indigenous farmers were able to produce food in the face of floods, droughts, and severe frosts by growing crops such as potatoes, quinoa, oca, and amaranthus in raised fields or "waru-warus," which consisted of platforms of soil surrounded by ditches filled with water (Browder, 1989).

Technicians have now assisted local farmers in reconstructing 10 ha of these ancient farms, with encouraging results, which later led to a substantial expansion of the area under warus. For instance, yields of potatoes from waru-warus

can surpass yields from chemically fertilized fields. Recent measurements indicate that waru-warus produce 10 tons of potatoes per hectare compared to the regional average of 1–4 tons/ha.

This combination of raised beds and canals has proven to have remarkably sophisticated environmental effects. During droughts, moisture from the canals slowly ascends the crop roots by capillary action, and during floods, furrows drain away excess runoff. Waruwarus also reduce the impact of temperature extremes. Water in the canal absorbs the sun's heat by day and radiates it back by night, thereby helping protect crops from frost. On the raised beds, nighttime temperatures may be several degrees higher than in the surrounding area. The system also maintains its own soil fertility. In the canals, silt, sediment, algae, and organic residues decay into a nutrient-rich muck that can be dug out seasonally and added to the raised beds. There is no need for modern tools or fertilizers, and the main expense is manual labor to dig canals and build up the platforms. This ancient technology is proving so productive and inexpensive that it is now being actively promoted throughout the Andean altiplano.

One of the early projects advocating the reconstruction of traditional farming systems occurred in Mexico in the mid-1970s when the then existing Instituto Nacional de Investigaciones sobre los Recursos Bioticos (INIREB) unveiled a plan to build "chinampas" in the swampy region of Veracruz and Tabasco. Chinampa agriculture was perfected by the Aztec inhabitants of the Valley of Mexico prior to the Spanish Conquest. It involves the construction of raised farming beds in shallow lakes or marshes, and represents a self-sustaining system that has operated for centuries as one of the most intensive and productive ever devised by humans. Until the last several decades, chinampas demanded no significant capital inputs yet maintained extraordinarily high yields year after year. A wide variety of staple crops, vegetables, and flowers are often mixed with an array of fruit trees and bushes. Abundant aquatic life in the canals provides valuable sources of protein for local diets (Gliessman, 1998).

Now threatened by the sprawling growth of Mexico City and its suburbs, chinampas have nearly vanished except in a few isolated areas. Regardless, this system still offers a promising model as it promotes biological diversity, thrives without chemical inputs, and sustains year-round yields. When INIREB first began to establish the chinampa system in the lowland tropics of Tabasco, implementation and adoption met with mixed success. Some critics felt that no market outlets were explored or developed for the new outputs produced by the community. Nevertheless, the "raised beds" of Tabasco (or camellones chontales) are still in full operation in the swamps of this region, and apparently the local Chontal Indians have full control of them. The Chontal practice traditional agriculture, and these raised beds produce a great variety of products, which in turn have enhanced the income and food security of these "swamp farmers."

In a completely different ecoregion in the Andes, several institutions have engaged in programs to restore abandoned farming terraces and build new ones. In the Colca Valley of southern Peru, PRAVTIR (Programa de Acondicionamiento Territorial y Vivienda Rural) sponsors terrace reconstruction by offering peasant communities low-interest loans, seeds, and other inputs to restore large areas of abandoned terraces. The main advantages of using terraces are that they minimize risks in times of frost or drought, reduce soil loss, amplify the cropping options because of microclimate and hydraulic differences, and thus improve crop yields. Yield data from new bench terraces showed a 43–65 percent yield increase in potatoes, maize, and barley compared to yields of these crops grown on sloping fields. One of the main constraints of this technology is its high labor intensity, requiring about 350–500 worker-days per hectare for the initial building of the terraces. Such demands, however, can be buffered when communities organize and share tasks (Browder, 1989).

Another example of how a biodiversity-based approach can support or even resurrect traditional agriculture is occurring on Chiloé Island in southern Chile. This is a secondary center of origin of potatoes, and development workers are currently tapping the ethnobotanical knowledge of elderly female Huilliche Indians in an effort to slow genetic erosion and to recover some of the original native potato germplasm. They intend to provide impoverished farmers with locally adapted varieties that can produce without the use of agrochemical fertilizers. After surveying several agroecosystems on Chiloé, technicians collected hundreds of samples of native potatoes still grown by local farmers, and with this material, and in collaboration with farmers, they established community seed banks where more than 120 traditional varieties are grown year after year and are subjected to selection and seed enhancement. In this way, an *in situ* conservation program has been initiated involving farmers from various rural communities, thus ensuring the active exchange of varieties among participating farmers. As more farmers become involved, this strategy will provide a continuous supply of seeds to resource-poor farmers and will also create a repository of vital genetic diversity for future regional crop improvement programs (Altieri, 1995).

Conclusions

A key conclusion that emerges from the relevant anthropological and ecological literature is that, when not disrupted by economic or political forces, indigenous modes of food production generally preserve rather than destroy biodiversity and natural resources. In fact, in any particular region, capitalist development through the promotion of large-scale, energy-intensive, commercial agriculture is bound to deplete natural resources more than some of the existing traditional systems. A number of studies have proven that many traditional agricultural systems are highly sustainable and productive, offering an alternative to the capital-intensive agriculture currently promoted by many development and governmental agencies. Besides employing crop diversity, traditional farmers use a set of practices that often cause minimal land degradation. These include the use of terraces and hedgerows in sloping areas, minimal tillage, mulching, small field sizes, and long fallow cycles (Grigg, 1974; Brush, 1982; Richards, 1985; Netting, 1993). It is clear that this more traditional strategy is both ecologically informed and environmentally sound, as the agricultural practices that are most likely to endure are those that deviate least from the native diversity of the natural

plant communities within which they exist (Altieri, 1995; Gliessman, 1998).

This assessment of traditional subsistence agriculture does not romanticize its origins or practitioners, nor does it consider development per se to be detrimental. The intention is rather to stress the demonstrated value of traditional agriculture in the preservation of biodiversity, native crop diversity, and the adjacent vegetation communities (Toledo, 1980). Basing a rural development strategy on traditional farming and ethnobotanical knowledge not only assures the continual use and maintenance of valuable genetic resources, but also allows for the diversification of peasant or other indigenous subsistence strategies (Alcorn, 1984; Caballero and Mapes, 1985), which is a crucial issue in times of economic uncertainty.

The study of traditional agroecosystems and the ways in which indigenous peoples maintain and use biodiversity can facilitate the discovery of valuable agroecological principles, which in turn can contribute to the development of more sustainable agroecosystems and biodiversity conservation strategies in both developed and less developed countries.

See also: Agriculture, Industrialized. Agriculture, Sustainable. Ethnobiology and Ethnoecology. Grazing, Effects of. Indigenous Peoples and Biodiversity. Land-Use Patterns, Historic. Traditional Conservation Practices

References

Alcorn JB (1984) *Huastec Mayan Ethnobotany*. Austin: University of Texas Press.

Altieri MA (1987) *Agroecology: The Scientific Basis of Alternative Agriculture*. Boulder, Colorado: Westview Press.

Altieri MA and Anderson MK (1986) An ecological basis for the development of alternative agricultural systems for small farmers in the Third World. *Amer. J. Alternative Agriculture* 1: 30–38.

Altieri MA and Merrick LC (1987) *In situ* conservation of crop genetic resources through maintenance of traditional farming systems. *Econ. Botany* 4: 86–96.

Altieri MA, Anderson MK, and Merrick LC (1987) Peasant agriculture and the conservation of crop and wild plant resources. *J. Soc. Conservation Biol.* 1: 49–58.

Beets WC (1982) *Multiple Cropping and Tropical Farming Systems*. Boulder, Colorado: Westview Press.

Berlin B, Breedlove DE, and Raven PH (1973) General principles of classification and nomenclature in folk biology. *Amer. Anthropologist* 75: 214–242.

Brokensha DW, Warren DM, and Werner O (1980) *Indigenous Knowledge Systems and Development*. Lanham, Maryland: University Press of America.

Brush SB (1982) The natural and human environment of the central Andes. *Mountain Res. Develop.* 2: 14–38.

Brush SB (1986) Genetic diversity and conservation in traditional farming systems. *J. Ethnobiol.* 6: 151–167.

Bye RA (1981) Quelites—Ethnoecology of edible greens—Past, present and future. *J. Ethnobiol.* 1: 109–123.

Caballero JN and Mapes C (1985) Gathering and subsistence patterns among the P'urhepecha Indians of Mexico. *J. Ethnobiol.* 5: 31–47.

Chacon JC and Gliessman SR (1982) Use of the "non-weed" concept in traditional agroecosystems of southeastern Mexico. *Agro-Ecosystem* 8: 1–11.

Chambers R (1983) *Rural Development: Putting the Last First*. Essex, United Kingdom: Longman Group Limited.

Chang JH (1977) Tropical agriculture: Crop diversity and crop yields. *Econ. Geogr.* 53: 241–254.

Clawson DL (1985) Harvest security and intraspecific diversity in traditional tropical agriculture. *Econ. Botany* 39: 56–67.

Denevan WM, Treace JM, Alcorn JB, Padoch C, Denslow J, and Paitan ST (1984) Indigenous agroforestry in the Peruvian Amazon: Bora Indian management of swidden fallows. *Interciencia* 9: 346–357.

Francis CA (1986) Variety development for multiple cropping systems. *CRC Crit. Rev. Plant Sci.* 3: 133–168.

Frankel OH and Soulé ME (1981) *Conservation and Evolution*. Cambridge, United Kingdom: Cambridge University Press.

Gliessman SA, Garcia E, and Amador A (1981) The ecological basis for the application of traditional agricultural technology in the management of tropical agro-ecosystems. *Agro-Ecosystems* 7: 173–185.

Grigg DB (1974) *The Agricultural Systems of the World: An Evolutionary Approach*. Cambridge, United Kingdom: Cambridge University Press.

Harlan JR (1976) The possible role of weed races in the evolution of cultivated plants. *Euphytica* 14: 173–176.

Harwood RR (1979) *Small Farm Development—Understanding and Improving Farming Systems in the Humid Tropics*. Boulder, Colorado: Westview Press.

Klee GA (1980) *World Systems of Traditional Resource Management*. New York: J. Wiley & Sons.

Mooney PR (1983) The law of the seed. *Develop. Dialogue* 1: 1–172.

Nabhan GP (1983) *Papago Indian Fields: Arid Lands Ethnobotany and Agricultural Ecology*. Unpubl. Ph.D. diss. Tucson: University of Arizona.

Norman MJT (1979) *Annual Cropping Systems in the Tropics*. Gainesville: University Presses of Florida.

Prescott-Allen R and Prescott-Allen C (1981) *In Situ Conservation of Crop Genetic Resources: A Report to the International Board for Plant Genetic Resources*. Rome: IBPGR.

Richards P (1985) *Indigenous Agricultural Revolution*. Boulder, Colorado: Westview Press.

Toledo VM (1980) La ecologia del modo campesino de producion. *Antropologia y Marxismo* 3: 35–55.

Toledo VM, Carabias J, Mapes C, and Toledo C (1985) *Ecologia y Autosuficiencia Alimentaria*. Mexico City: Siglo Vientiumo Editors.

Wilken GC (1970) The ecology of gathering in a Mexican farming region. *Econ. Botany* 24: 206–245.

Wilken GC (1977) Integrating forest and small-scale farm systems in middle America. *Agro-Ecosystems* 3: 291–302.

Wilkes HG and Wilkes KK (1972) The green revolution. *Environment* 14: 32–39.

Williams BJ and Ortiz-Solario C (1981) Middle American folk soil taxonomy. *Ann. Assoc. Amer. Geographers* 71: 335–358.

Agrobiodiversity

Louise E Jackson, University of California, Davis, CA, USA
L Brussaard and PC de Ruiter, Wageningen University, Wageningen, The Netherlands
Unai Pascual, University of Cambridge, Cambridge, UK
Charles Perrings, Arizona State University, Tempe, AZ, USA
K Bawa, University of Massachusetts Boston, MA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–13,
© 2007, Elsevier Inc.

Glossary

Agricultural intensification Process by which farming systems increasingly use high levels of nonrenewable, purchased inputs, e.g., inorganic fertilizers and other agrochemicals for pest suppression, substitute mechanization, and fossil fuels for human labor and increase capital invested per unit of land.

Agrobiodiversity The variety and variability of living organisms that contribute to food and agriculture in the broadest sense, and that are associated with cultivating crops and rearing animals within ecological complexes.

Conservation agriculture Farming practices that maintain a permanent or semipermanent organic soil cover, as a growing crop or a dead mulch to protect the soil physically and to feed soil biota, which improves soil quality.

Ecoagriculture Land use approaches that achieve agricultural production and sustain rural livelihoods in ways that also protect wild plant and animal species and the natural ecosystem services upon which both humans and wildlife depend.

Ecosystem approach Strategy for the integrated management of land, water, and living resources that promotes conservation and sustainable use in an equitable way.

Ecosystem services Ecological functions and processes by which the environment provides for human needs, e.g., clean water, timber, habitat for fisheries, pollination, and esthetic values.

Insurance hypothesis Biodiversity insures that ecosystems maintain functions because more species guarantee that some will continue to function even if others fail, thereby enhancing the capacity to recover from, and mitigate the risks caused by disturbance.

Land sparing On-farm management practices to increase yields per unit land area with the goal of minimizing the pressure to convert land in natural ecosystems for agriculture, so that wildlife in natural ecosystems persists, but at the potential cost of decreasing wildlife populations in and around intensified farmland.

Market-based instrument Policy mechanism in which a change in technology, behavior, or the mix of inputs and outputs is encouraged through financial incentives, such as subsidies, taxes, or market creation; its use is intended to align the social and private values of agrobiodiversity.

Market failure The lack of appropriate markets or the incapacity of existing markets to reflect the social value of agrobiodiversity through prices.

Niche complementarity Differential acquisition and use of resources by different species; niches of different species differ which results in more complete utilization of resources in spatially or temporally complex environments.

Organic agriculture Farming practices that rely on biological inputs and diversity for soil fertility and pest and disease suppression; organic farms that are certified by government policies or private organizations are not allowed to use synthetic pesticides or fertilizers.

Planned versus unplanned biodiversity Crops and livestock that are purposefully introduced and maintained in an agroecosystem (planned biodiversity) versus associated biodiversity that includes all soil biota, herbivores, carnivores, decomposers, and any other species that exist in or colonize the agroecosystem (unplanned biodiversity).

Resilience Capability of a system to adapt to external changes by either returning to its original state or evolving into a state preferable to the initial one; persistence, adaptiveness, variability, and unpredictability are all attributes that contribute to resilience.

Social benefits of conservation Array of benefits enjoyed by society from the conservation of agrobiodiversity that need not coincide with the values reflected by food and fiber markets which often accrue to the individual users of agrobiodiversity such as farmers; the wedge between private and social value reflects the external value of agrobiodiversity.

Use-value of agrobiodiversity Benefits accrued to users of agrobiodiversity directly by enhancing agricultural productivity, including increases in yields and savings in artificial inputs, and indirectly by enhancing stability and resilience of the agricultural system; nonuse values include the "option" value that arises from the future potential benefits of agrobiodiversity not necessarily known in the present, and the "existence value" stemming from human ethical considerations relating to matters such as the extinction of species and ecosystem.

Wildlife-friendly farming On-farm management practices as benign to wildlife as possible at the potential cost of decreasing yields; reduction in pesticides and fertilizers and their effects on nontarget organisms, and retention of patches of nonfarmed habitats and less-intensively farmed seminatural habitats within the agricultural landscape.

Introduction

Agrobiodiversity has always formed the basis for human food production systems, and has provided cultural, spiritual, religious, and esthetic value for human societies (Brush, 2004). Planned agrobiodiversity is the biodiversity of the crops and livestock chosen by the farmer (**Figure 1**), while unplanned biodiversity refers to the associated biota, for example, soil microbes and fauna, weeds, herbivores, and carnivores, colonizing the agroecosystem and surviving according to the local management and environment (Vandermeer and Perfecto, 1995). Rapid loss of agrobiodiversity is now occurring worldwide, as an indirect result of population growth in the developing world and the high per capita consumption of natural resources in the industrialized world, affecting the production of food, fuel, and fiber and also many ecological services such as those supporting water supplies, health, and wildlife habitat. Most (75%) of the world's poor people live in rural landscapes, and will be especially vulnerable to these changes (WRI, 2005).

A wide range of disciplinary sciences has contributed to understanding the functions of agrobiodiversity, as illustrated by the following examples. The agricultural sciences have shown that crop and animal breeding can select for genes to increase agricultural productivity (e.g., Cooper *et al.*, 2001). Crop diversity, in both space and time, can enhance nutrient use efficiency, and a diverse soil community may regulate tighter nutrient cycling, reduce pests and diseases, and improve soil structure (e.g., Brussaard *et al.*, 2007). Ecological research has shown that biodiversity can increase the stability of grassland ecosystems (Loreau *et al.*, 2002), which has implications for the management of pastures for livestock production. Additionally, the structure of agricultural landscapes, that is, mosaics of agricultural and nonagricultural ecosystems, is increasingly recognized as an important factor for the utilization and conservation of both on- and off-farm biodiversity. Economics has shown that by investing in biodiversity, environmental and market risks faced by producers can potentially be reduced, especially over the long term (Perrings, 1998). Interdisciplinary approaches in which the agroecological value of biodiversity is combined with the socioeconomic value, especially those focused on how managing biodiversity can be enhanced in agricultural practices, are now viewed as an important new direction of agrobiodiversity research (Jackson *et al.*, 2005). Although the agrobiodiversity literature focuses mainly on farming systems, many of the issues and approaches are also relevant to food, fiber, and timber production in natural ecosystems such as grazed grasslands, managed forests, and fisheries.

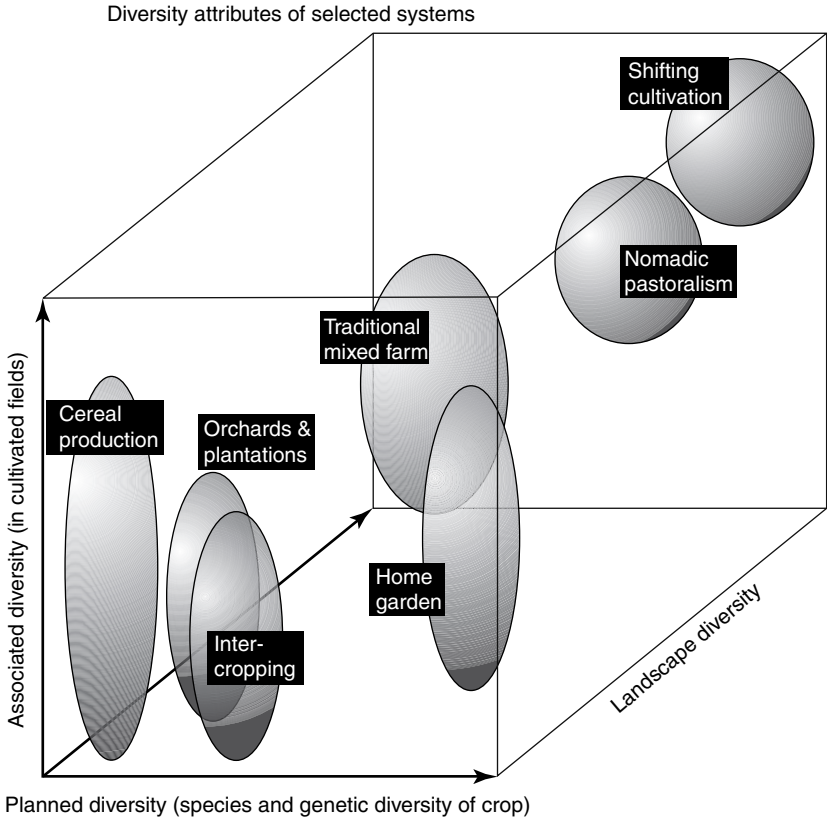


Figure 1 Planned and unplanned (or associated) biodiversity in selected farming systems. Both types of biodiversity support ecosystem services such as nutrient cycling, pest control, and pollination. Low values of associated biodiversity within cultivated fields are indicative of intensive farming systems, and thus are present in several of the types of farming systems. Cereal production, as an example, occurs in many permutations worldwide, and a wide range of associated biodiversity is possible. The same pattern probably characterizes livestock husbandry. Reproduced with permission from Figure 26.4, p. 758, in Cassman KG and Wood S (2005) Cultivated systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystems and Human Well-Being Current State and Trends*, Ch 26. Washington, DC: Island Press.

This is because the adoption of biodiversity-based practices for agriculture is not solely based on functions and services that society as a whole obtains from such functions. It is individual farmers who are ultimately the agents who decide how much natural capital to conserve and utilize, and will do that based on their own objectives and needs, and the social, economic (e.g., markets and policies), and environmental conditions in which they operate. The multiple interactions between farmers and the natural resources they use, and between farmers and the institutional environment, create different farming practices and styles, which result in local specificity and regional heterogeneity of the natural and social environments which may either enhance or reduce agrobiodiversity. The private and social values of biodiversity often differ widely (**Figure 2**). Individual farmers tend to react to the private use-value of biodiversity when reflected in the marketplace and thus typically ignore the "external" benefits of conservation that accrue to the wider society. This currently implies an overexploitation of biodiversity, thus imposing excessive costs on society at large. Institutional intervention through incentives is therefore needed to ensure that agrobiodiversity will be used and conserved to its full potential of functional benefits.

Fortunately, awareness of the need for sustainable agricultural production has increased, in response to the unprecedented population growth, food demand, and regionally high per capita use of natural resources and global environmental change that is now occurring. The focus is shifting to a greater reliance on ecological goods and services, which is especially important for modern intensive agriculture, since there will be less-damaging effects on environmental quality and on the biodiversity of wildlands. Conservation of existing biodiversity in agricultural landscapes and the adoption of biodiversity-based practices have been proposed as ways to increase the sustainability of agricultural production (Collins and Qualset, 1999; McNeely and Scherr, 2003). For example, in the Millennium Ecosystem Assessment (MEA, 2005), biodiversity is viewed as an important coping strategy against agricultural risks in an uncertain future. A successful transition to a biodiversity-based paradigm for agriculture requires fundamental knowledge and understanding of biodiversity as natural capital for providing ecosystem goods and services for agriculture, the direct and indirect use-value in economic terms that are derived from these goods and services, and the social forces, including poverty, that will promote or impede its sustained adoption (Swift *et al.*, 2004; Wood and Lenné, 2005).

Agrobiodiversity and Modern Agricultural Intensification

Rapid Land-Use Change to Meet World Food Demand

Modern intensive cropping systems have greatly increased the global food supply by utilizing high-yielding genotypes that demand high levels of synthetic inputs (e.g., inorganic fertilizers and other agrochemicals for pest suppression, irrigation, and fossil fuels) (Matson *et al.*, 1997; Evenson and Gollen, 2003). However, there have been significant environmental costs in terms of increased pollution, loss of biodiversity in both agroecosystems and in wildlands, and loss of traditional knowledge associated with biodiversity. Furthermore, economic subsidies paid to modern intensive agriculture have also led to the oversupply, from ecological and economic points of view, of crops and associated agrichemicals (MEA, 2005).

The world's population of 6.3 billion people is projected to grow to 9 billion by 2050. At present, 10% of the global land area is under modern, intensive agricultural use, 17% is under

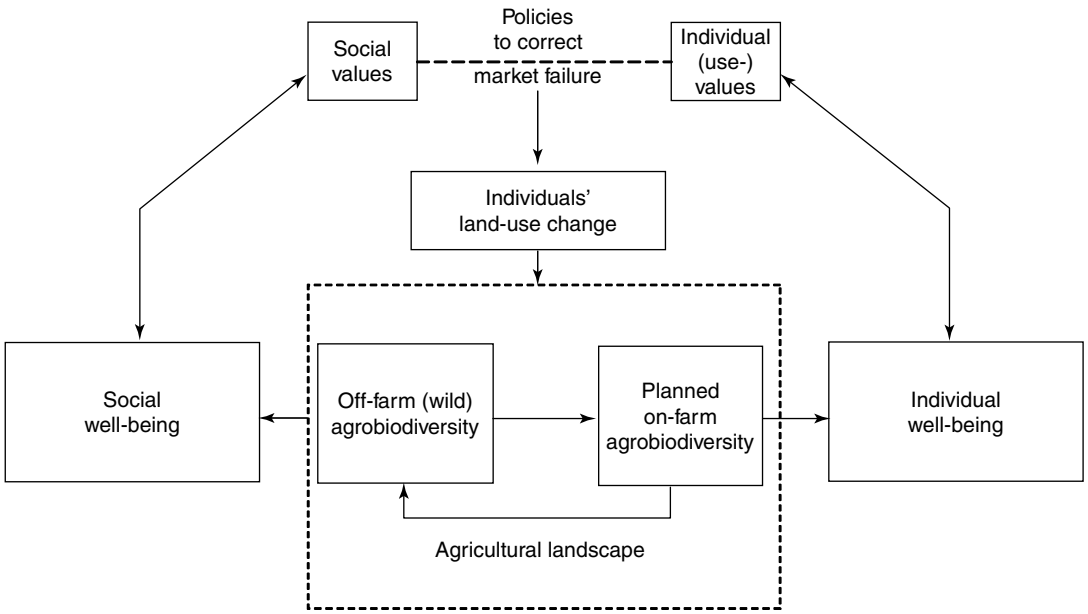


Figure 2 Linkages between social and individual or private values of biodiversity conservation. When the social value outweighs the individual value, then biodiversity is underprovided by the direct land users. Policies are needed to correct such market failure.

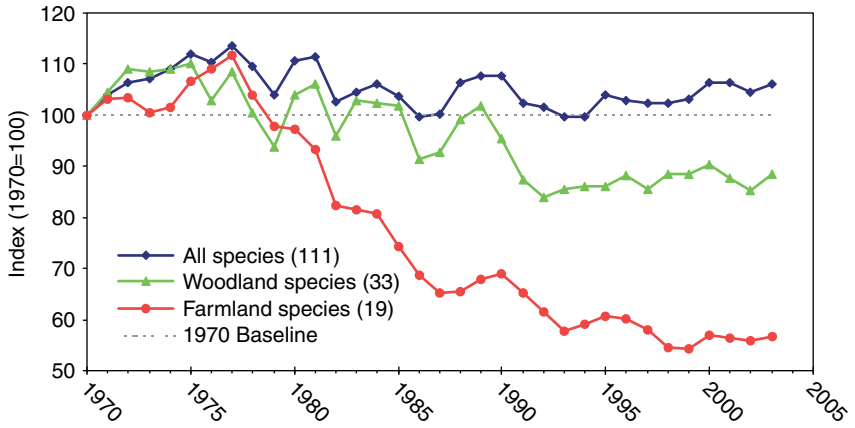


Figure 3 Marked declines in bird species populations across the United Kingdom are due to agricultural intensification (Donald *et al.*, 2001), especially many upland grassland bird species (Henderson *et al.*, 2004). The index is the percentage change in bird species in various categories relevant to a 100% baseline value for 1970. Reproduced from DEFRA (2004) Indicators of Sustainable Development. Sustainable Development Unit, Department of Environment, Food, and Rural Affairs, UK, http://www.defra.gov.uk/corporate/sdstrategy/chapter3.htm (accessed November 2006).

extensive use associated with a lower application of artificial inputs, and 40% is grazed by domestic livestock (Wood *et al.*, 2000; Mooney *et al.*, 2005). It is projected that by 2050, food production must double to meet human needs and hence it is expected that more land will be converted into agriculture. Further intensification on already converted land is also expected to increase through greater reliance on synthetic, off-farm inputs unless more sustainable paradigms are developed (Clay, 2004).

Irrigated and pasture lands are both expected to double in area by 2050, with a net loss of 10^9 ha of wildlands worldwide, thereby increasing the global pressure on biodiversity in natural ecosystems. Agricultural intensification not only leads to a loss of diversity of breeds and associated agroecosystem biodiversity, such as farm birds (**Figure 3**), beneficial insects, and soil biota, but it can also put wild biodiversity at risk through gene flow from domesticated varieties to wild species, exposure to potentially virulent pathogens *via* cross-species transmission, and adverse effects of agrochemicals on nontarget species in adjacent wildland ecosystems. It is possible that a wide range of ecosystem services are thereby affected, including but not limited to provisioning services that support production of food, fuel, and fiber; regulating services such as pollination and pest control; and supporting services such as nutrient cycling and water purification.

Several of the projected long-term human population growth trends are likely to impact agrobiodiversity through land-use changes in agricultural landscapes (Kates and Parris, 2003). Cities will expand to accommodate the growing urban population, and rural populations will proportionately decrease. In turn, agroindustrialization will increase in many regions, resulting in increased processing of food, use of off-farm sources of inputs, distribution of globalized farm technologies, and vertical coordination, in which management units at different stages of production and processing are owned by the same firm, and product flows are coordinated through administrative means (Reardon and Timmer, 2005). Large agribusiness firms may increasingly dominate food production systems, with concomitant increases in large-scale farming and marketing operations, which may jeopardize the livelihoods of small farmers. In contrast, stabilization of agricultural livelihoods of rural populations in many low-income countries may occur if higher energy prices for transport and lower communication costs provide incentives for less-concentrated patterns of settlement, and greater reliance on local agricultural products.

Besides population dynamics, global impacts of climate change are also expected to alter agricultural landscapes. Developing regions, and Africa in particular, are most vulnerable to the impacts of projected changes because existing widespread poverty limits adaptation capabilities. **Figure 4** shows the predicted transition between pastoral and mixed agro-pastoral systems in Africa. The land-use change projections indicate that for Africa, there will be a decrease in the total area in which cropping will be possible by 2050 (Thornton *et al.*, 2002). Coupled with population growth, and the decline of large tracks of arable land, further agricultural intensification is very likely.

Agrobiodiversity and Sustainable Agriculture

The broad concept of "sustainable agriculture" is characterized by a set of complementary goals: producing more food and fiber, protecting the resource base upon which agriculture depends, and promoting social well-being. A widely accepted definition of sustainable agriculture in the United States is one that "will over the long term[1]

- satisfy human food and fiber needs
- enhance environmental quality and the natural resource base upon which the agricultural economy depends
- make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls

[1]1990 Farm Bill (US Government 1990).

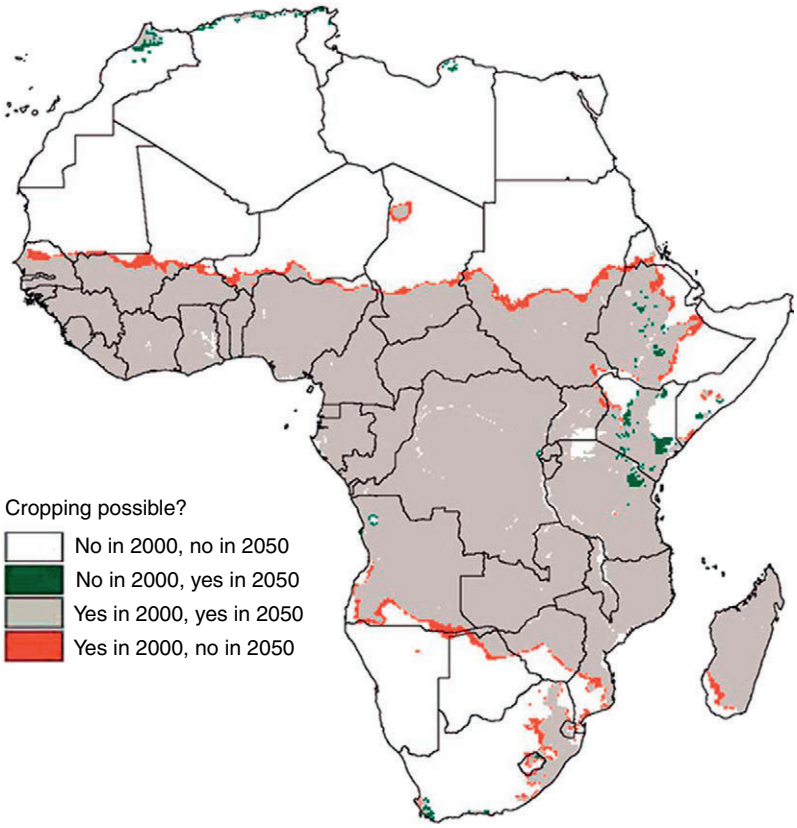


Figure 4 Predicted transitions of cropping systems in Africa due to global climate change between 2000 and 2050. Reproduced from Thornton PK, Kruska RL, Henninger N, *et al.* (2002) *Mapping Poverty and Livestock in the Developing World*, 124pp. Nairobi, Kenya: ILRI (International Livestock Research Institute).

- sustain the economic viability of farm operations
- enhance the quality of life for farmers and society as a whole."

The utilization of agrobiodiversity provides a set of tools and approaches to intensify agricultural production and still support sustainable agriculture (Srivastava *et al.*, 1996).

Agriculture is both a threat to biodiversity and a key to its survival. Unless agriculture is sustainably intensified, many remaining wild areas will succumb to the plow, ax, or herds of grazing livestock. With global demand for food and other agricultural products expected to at least double – and perhaps even triple – over the next 50 years, the productivity of existing farms and rangelands will have to increase dramatically. How this intensification process plays out in the decades ahead will largely determine how many species – and their habitats – will survive into the coming century.

Traditional agriculture is regarded as more sustainable than intensified agriculture, due to strategies that minimize risk and stabilize yields, promote diversity in diets, and utilize practices that require low levels of technology and efficient use of limited inputs (Altieri, 2004). At present, most of the world's population lives in rural areas and depends to some extent on traditional agricultural systems (Brookfield *et al.*, 2003). But as the pressure for more food increases, innovative biodiversity-based solutions are needed to support sustainable agriculture relevant to specific locations and farming systems. Some will build upon traditional practices, such as replanting of clonal rubber trees within agroforests to avoid low-income regrowth phases in Sumatra (Joshi *et al.*, 2003), or utilizing species-rich Mediterranean grasslands that have higher water-use efficiency than monocultures to increase water-use efficiency compared with monoculture (Caldeira *et al.*, 2001). Others may be based on production systems that mimic natural systems, for example, domestication of stress-tolerant native perennial grasses such as grain crops, as promoted by the Land Institute (Salina, Kansas). But the main source of biodiversity-based innovation for sustainable agriculture is likely to come from ecosystem approaches that blend ecology, agronomic and horticultural technologies, and socioeconomic analysis to develop high-yielding and sustainable farming systems. Current examples that have been shown in some cases to achieve these objectives are conservation agriculture, which uses farming practices that maintain a permanent or semipermanent organic soil cover to increase soil quality and reduce fuel use; and organic agriculture, which relies on biological inputs and diversity for pest management and soil fertility.

Agrobiodiversity and Resilience to Environmental Change

The capability of a system to adapt to external changes by either returning to its original state or by evolving into a state preferable to the initial one is referred to as resilience. Resilience also involves a quantitative/dynamic aspect in terms of the rate with which the system returns to equilibrium, for

example, highly diverse soil communities were more resistant to an imposed stress than low-diversity communities (Griffiths *et al.*, 2000). A basic hypothesis, shared by ecologists and economists, is that various functions of biodiversity contribute to the resilience of agroecosystems. For example, increasing the diversity of annual crops and trees in agroecosystems can improve the spatial and temporal efficiency of resource use as well as the economic viability for farmers by diversifying their revenues under uncertain market conditions with respect to volatile market price changes of inputs, crops, or tree products.

Biodiversity is thought to enhance the capacity to recover from disruption of functions, and the mitigation of risks caused by disturbance. This may be most important at larger spatial and temporal scales, by providing insurance value, especially when dispersal abilities of organisms allow for immigration within the landscape (Loreau *et al.*, 2003, **Figure 5**). The insurance hypothesis proposes that species or phenotypes that appear to be functionally redundant for a specific ecosystem process at a given time and space may actually diverge in response to a new set of environmental conditions. Redundancy is strongly dependent on the type of agroecological functions. For example, "narrow" processes such as soil nitrification are associated with low redundancy, while broad processes such as soil organic matter decomposition or N-mineralization are carried out by a large variety of species, hence redundancy is high.

At the larger landscape scale, ecosystem functions would be expected to be more stable and less subject to fluctuation in species-rich landscapes. Heterogeneous composition of ecosystems in agricultural landscapes may thus provide insurance value that is not detected by the local-scale experiments that are typical of most agricultural research. More diversity in crops, management strategies, and mosaics of ecosystems in the landscapes may ultimately be socially desirable. On the short term, less agrobiodiversity would make the system more stable in the sense that there is less variation in producer or consumer behavior following minor perturbations. But on the long term, it may also reduce the capacity of that system to absorb greater environmental or economic risks. The uncertainties associated with rising CO_2 concentrations and temperature suggest that agricultural responses to global climate change may benefit from the resilience that results from increasing agrobiodiversity at the landscape level.

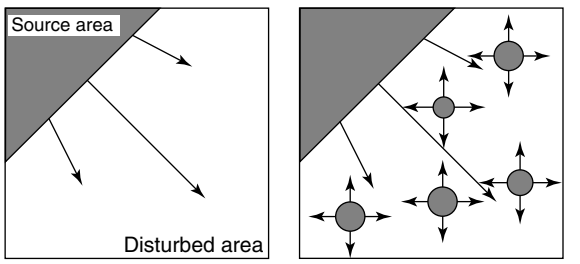


Figure 5 The biological legacy or ecological memory present in or near a disturbed area affects the reestablishment and organization of biota. Future community composition depends on spatial configuration of habitats and distance to source areas. Severe uniform disturbances result in loss of many species (left). Colonization and reorganization require that disturbed areas are in close proximity to undisturbed patches (right). Reproduced from Bengtsson J (2002) Disturbance and resilience in soil animal communities. *European Journal of Soil Biololgy* 38: 119–125.

Strategies to Protect Wildlife in Agricultural Landscapes

There has been much debate over cropland expansion versus intensification in agricultural landscapes to increase food supply and enhance environmental services and biodiversity of wild species (Donald, 2004; Green *et al.*, 2005; Mooney *et al.*, 2005). Usually, two scenarios are considered: (1) wildlife-friendly farming would use lower agrichemical inputs to reduce impact on nontarget biota, but would expand production of less-intensive, often low-yielding cultivation systems into wildlands to meet food demands or (2) land sparing by intensifying agriculture using agrichemical inputs and monocultures to produce more food per unit land area would result in less increase in the area of cultivated lands at the potential cost of decreasing wildlife populations in and around farmland, but would save wildland biodiversity. These extremes are too simplistic. For one reason, high-input, modern agriculture can quickly expand to dominate an agricultural landscape, as the efficiency of production often increases with the volume of goods, that is, economy of scale, so that land sparing does not materialize. Another reason is that biodiversity-friendly agriculture can offer sustainable alternatives that maintain or increase agriculture productivity with positive benefits for wildlife, for example, ecoagriculture (McNeely and Scherr, 2003), so that both wildlife-friendly farming and land sparing are achieved.

Forest–agricultural ecotones are the areas that adjoin forests and other natural habitats and that lie between forests and zones of intensive agriculture. They often support both high biodiversity and human populations (Rosenzweig, 2003). Conservation can be enhanced if there is less-direct dependence on forests in the form of collection of nontimber forest products, and instead, greater choices for livelihoods derived from diversification of farming systems and use of nonforest products. Nongovernmental organizations, such as ATREE and Ecoagriculture Partners, have effectively demonstrated that imparting ideas for simple technologies can increase yields, profitability, and diversification of livelihoods, thereby reducing the impact on the biodiversity within the forest zone (McNeely and Scherr, 2003; Bawa *et al.*, 2007).

When native vegetation is cleared for agriculture, habitats which were once continuous become divided into separate fragments, which become islands isolated from each other by cropland, pasture, and urban areas. These islands are not favorable for the set of species which require interior habitat, and cannot survive in the habitat along edges of fragments, which experience a different microenvironment. One solution is to plant corridors of native vegetation to link fragments; another is to increase the amount of interior habitat. Landscape planning to reduce habitat loss and fragmentation has focused on population persistence, mainly of birds and

mammals, with respect to species' sensitivity to isolation and area effects. The effect of the characteristics of the matrix of ecosystems in the landscape is gradually gaining more attention (Swihart and Moore, 2004).

Ecological Functions of Agrobiodiversity

Agrobiodiversity is most likely to enhance ecosystem functioning when a unique or complementary effect is added to an ecosystem, for example, by planting genotypes with specific genes for higher yield or pest resistance, mixing specific genotypes of crops (Zhu *et al.*, 2000), using more crops in rotation, supporting more parasitoids or insect enemies with specific roles in controlling pests (Tscharntke *et al.*, 2005), or including a plant functional group, such as a legume, that increases nitrogen inputs and cycling (Drinkwater *et al.*, 1998). Current evidence suggests that merely adding more species to most agroecosystems has little effect on function, given the redundancy in many groups, especially for the soil biota (Swift *et al.*, 2004).

Agrobiodiversity: Genetic and Population Diversity

Genetic and population diversity has been successfully used in many ways for direct-use value for agricultural production. In modern production systems, the focus has been on identifying traits that increase yield, keeping up the treadmill of rapidly changing virulence of specific diseases, meeting the demand for resistance to an increased range of biotic and abiotic stresses, or increasing the temporal and spatial production of commodities that are destined for different uses or planted at different times of the year (Cassman *et al.*, 2003). Some of the gene pool available for crop improvement in the future has been safeguarded in gene banks for the major crops (**Table 1**).

In traditional production systems, farmers and communities opt to maintain a number of traditional varieties because they (1) meet their multiple needs more closely than modern ones, for example, grains that provide food for humans and forage for animals; (2) perform better under low levels of external inputs such as pesticides and herbicides; and (3) are less sensitive to abiotic stress, due to ecophysiological traits or due to the use of genotypic mixtures that minimize risk in seasonal differences (Ceccarelli, 1996).

Historically, agricultural breeders used landraces and crop wild relatives to introduce improved pest resistance, yield, and quality genes into crops and livestock (Cooper *et al.*, 2001). In most cases, crop breeders have selected for morphological and developmental traits, rather than increased rates of photosynthesis or other physiological processes (Sinclair *et al.*, 2004). Breeding has been facilitated by understanding the genetic aspects of adaptation, and developing innovative breeding methods that are more environmentally and culturally acceptable than genetic engineering, such as marker-assisted selection (Gepts, 2002). Readily available genetic sequence information and genetic and physical maps open new possibilities for introgressing traits for stress tolerance and disease resistance (Engels *et al.*, 2002) and will further increase utilization of agrobiodiversity at the genetic and population levels.

Agrobiodiversity: Communities and Ecosystems

Increased plant species richness can have significant effects on ecosystem processes, based on mixtures of approximately 1–20 species in grasslands (Loreau *et al.*, 2002). This can be attributed to niche complementarity (i.e., differential resource used by different species), positive interactions between organisms, and greater resource capture than in species-poor communities (Tilman *et al.*, 2001). By inference, small increases in biodiversity may have a large effect on ecosystem functions such as productivity in agricultural communities composed of only a few species (de Wit, 1960), yet trophic interactions may make outcomes more complicated (Lewis *et al.*, 1997).

A good example of the benefits of using an ecosystem approach, rather than a reductionist approach that focuses on specific organismal interactions is the role of agrobiodiversity in pest-control strategies such as planting crop mixtures, integrating plants with natural defences against herbivore attack, for example, toxin biosynthesis, releasing natural enemies, and enhancing the complexity of the soil food web (van Bruggen and Termorshuizen, 2003). However, these must be integrated with management practices of the specific cropping system

Table 1 Percentage of gene pool remaining to be collected for major world crops, based on surveys of crop experts. The numbers lack precision due to the possibility of duplication and to incomplete knowledge of the gene pool of each crop

Crop	*Estimated, mid-1980s percentage*	*Estimated 1986 percentage*	*Estimated, mid-1990s percentage*
Wheat	10	5	5
Rice	8–15	30	5
Maize	2	10	5
Sorghum	25	20	20
Pearl millet	<30	20	
Potato	10–20	5	5
Sweet potato	>50	40	50
Cassava	25–33	65	
Soybean	30	40	
Common bean	<50	50	35

Source: Reproduced from Fowler C and Hodgkin T (2004) Plant genetic resources for food and agriculture: Assessing global availability. *Annual Review of Environmental Research* 29: 143–179.

(Lewis *et al.*, 1997). Ideally, species assemblages will provide a set of multifunctions, but these raise issues about trade-offs (Gurr *et al.*, 2003). For example, weeds can favor natural enemies of insect pests by providing nonhost foods (e.g., pollen, nectar, alternative hosts, and prey) and shelter. Thus, weed-management practices, for example, herbicide application, must be evaluated in terms of a range of different biota and functions, which are unique to specific agroecosystems.

Changes in community structure often occur simultaneously with changes in inputs, making it hard to differentiate the effects of biodiversity. For example, long-term organic management can lead to higher biodiversity of plants, arthropods, and soil microbial functional groups and greater soil C sequestration, than long-term conventional management. However, the subsequent changes in ecosystem function cannot be simply attributed to higher biodiversity, because other inputs involved with organic production also vary substantially, for example, fertilizers, manures, and legume-based crop rotations (Mäder *et al.*, 2002).

Species interactions can be complicated, and are not always repeatable. As an example, a more complex rotation can alter crop nutritional properties but also increase attacks by insect pests. For soil biota, it is especially difficult to attribute ecosystem functions to biodiversity, since numerous species, many of which are unidentifiable, contribute to soil activity and aboveground responses (Swift *et al.*, 2004; Wardle *et al.*, 2004). Thus, idiosyncratic effects on soil biota can occur, even with the same plant species, or same plant community, or a similar soil type.

Based on these considerations, assessing species diversity and community assemblages for multifunctionality, along with gauging inputs to maximize economic benefit and environmental quality, is important for sustainable agriculture. This requires interdisciplinary research, an ecosystem approach, and often site-specific analyses across different types of gradients.

Agrobiodiversity: Landscapes

The insurance hypothesis of biodiversity, that is, that higher numbers of species increase resilience and reorganization after disturbance, may be most relevant at the landscape level (Loreau *et al.*, 2003; Swift *et al.*, 2004). Agricultural landscapes that are composed of a mosaic of well-connected early and late successional habitats may also be more likely to harbor biota that contribute to regulating and supporting services for agriculture, compared with simple landscapes (Bengtsson *et al.*, 2003). This idea, however, is not easy to demonstrate. As an example, European agrienvironmental schemes support direct payments to farmers to provide environmental benefits, but the observed species richness of birds and vascular plants increased slightly or not at all, whereas there were more consistent increases for arthropods (Kleijn and Sutherland, 2003). This may have been due to the low frequency of natural and seminatural areas in these agricultural landscapes, and thus, there were few opportunities for dispersal of species and functional groups of insects from relatively undisturbed habitats into agricultural production areas (Duelli and Obrist, 2003). But a more general problem is finding truly paired replicates that constitute a robust research design.

Research on landscape-level biodiversity requires close association with farmers and other land managers, and involvement with stakeholders that care about the boundaries between agricultural and nonagricultural ecosystems. These human interactions are necessary for both the conservation of undisturbed natural ecosystems and traditional human land-use systems. Landscape-level biodiversity research poses some major issues related to scale. Agroforestry research in Southeast Asia has shown that a useful approach is to emphasize both local and regional data collection, and assess trade-offs between different types of services provided under landscape scenarios that vary in the fallowing frequency and intensification both with direct data collection and with models (Murdiyarso *et al.*, 2002; van Noordwijk, 2002).

The economic assessment of services provided by agrobiodiversity at the landscape level is contingent on the juxtaposition of different types of agricultural and nonagricultural ecosystems. Simple calculations have shown that pollination services contributed by wild bees increase yields and thus profitability on a farm field (Kremen *et al.*, 2002; Ricketts *et al.*, 2004). Even so, there are costs associated with leaving land out of agricultural production to protect the bees' nearby undisturbed habitats. Modeling approaches can show the trade-offs between various aspects of environmental quality, economics, and social welfare that change when biodiversity-based practices are adopted, especially when trade-offs are complex and depend strongly on scale (Hietala-Kiovu *et al.*, 2004; van Noordwijk, 2002).

Agrobiodiversity Utilization and Conservation: Socioeconomic Considerations

The economic value of the changes in agrobiodiversity and its associated services is rarely reflected in current market transactions. Therefore farmers, as the direct users of agrobiodiversity and usually decision-makers in general, do not pay sufficient attention to these services. This implies that there is too little incentive to maintain sufficient diversity to meet the needs of society in the context of multifunctional agricultural systems. The main problem is that markets for agricultural inputs and outputs do not reflect the full social opportunity costs. Economists refer to this as the problem of market failure, and to the resultant loss of biodiversity as an external effect of agricultural markets that creates excessive social costs. Many policies, particularly in agriculture, aggravate the problem by modifying agricultural prices in ways that increase the gap between the value of agrobiodiversity faced by private users (e.g., farmers) and the social value that better reflects the perceived scarcity of agrobiodiversity by society, for example, price supports for intensively managed crops. An alternative is rewards or payments for ecosystem services that result from adoption of biodiversity-friendly practices (**Figure 6**).

To resolve these problems, a different outlook on biodiversity in agricultural landscapes is needed (Perrings *et al.*, 2006). Biodiversity should be considered part of the natural capital that generates a flow of ecosystem services, upon which society depends, that is, interest on that capital. Just as private

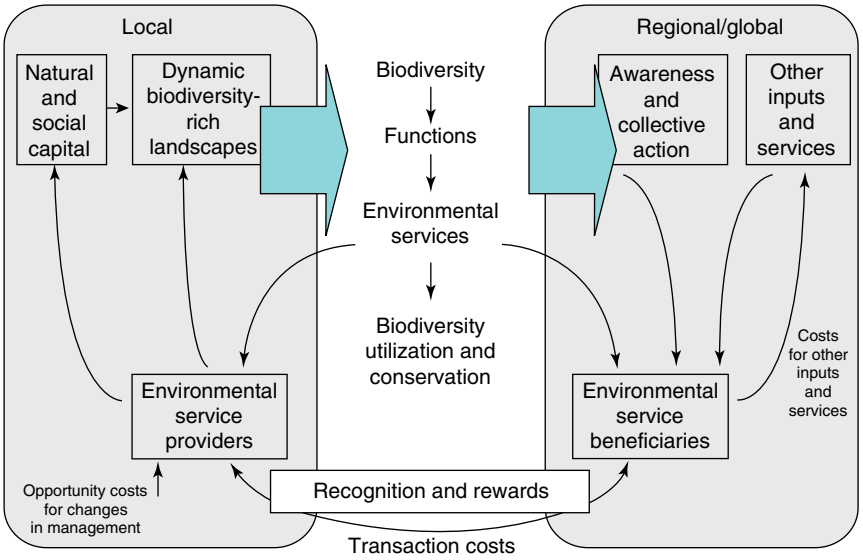


Figure 6 Linkages between providers and beneficiaries of environmental services derived from dynamic biodiversity-rich landscapes, each supporting biodiversity utilization and conservation by different mechanisms. Reproduced from van Noordwijk M, Chandler FJ, and Tomich TP (2004) *An Introduction to the Conceptual Basis of RUPES: Rewarding Upland Poor for the Environmental Services They Provide*, 46pp. Bogor: ICRAF-Southeast Asia.

investors select a portfolio of produced capital to maintain the return on capital over a range of market risks, so must society choose the mix of genes, species, and ecosystems to maintain the flow of ecosystem services over a range of risks for environmental quality and human well-being, including poverty alleviation. Risks may arise from changes in that mix, and these risks must be understood and used to inform conservation strategy and policy.

New types of communication and cooperation are necessary for evaluating the potential for the utilization and conservation of biodiversity in agricultural landscape, for example, among agriculturalists, ecologists, and economists to identify and establish adequate assessment strategies (Robertson and Swinton, 2005), between anthropologists and ecologists to preserve ethnobiological species and functions (Brush, 2004), and between conservation biologists and agriculturalists to resolve priorities for managing genetic, species, and ecosystem diversity in agricultural landscapes (Banks, 2004). In addition, bridging between the natural and the social sciences will improve the success of engaging farmers and other stakeholders to use biodiversity-based solutions for increasing agricultural production in a sustainable manner (Pretty and Smith, 2004). These are themes promoted by the international program of biodiversity science, DIVERSITAS, which has identified a science agenda for agrobiodiversity in support of sustainable agriculture that promotes interdisciplinary communication, with the goal of inspiring and facilitating new research approaches for understanding the role of biodiversity in agricultural landscapes (Jackson *et al.*, 2005).

See also: Agriculture, Sustainable. Ecology of Agriculture. Ecosystem Function, Principles of. Ecosystem Services. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Restoration of Animal, Plant, and Microbial Diversity. Scale, Concept and Effects of. Stewardship, Concept of. The Value of Biodiversity

References

Altieri MA (2004) Linking ecologists and traditional farmers in the search for sustainable agriculture. *Frontiers in Ecology and the Environment* 2: 35–42.

Banks JE (2004) Divided culture: Integrating agriculture and conservation biology. *Frontiers in Ecology and the Environment* 2: 537–545.

Bawa KS, Joseph G, and Setty S (2007) Poverty, biodiversity and institutions in forest-agriculture ecotones in the Western Ghats and Eastern Himalayas of India. *Agriculture, Ecosystems and Environment* 121: 287–295.

Bengtsson J (2002) Disturbance and resilience in soil animal communities. *European Journal of Soil Biololgy* 38: 119–125.

Bengtsson J, Angelstam P, and Elmqvist T (2003) Reserves, resilience and dynamic landscapes. *Ambio*. 32: 389–396.

Brookfield H, Parsons H, and Brookfield M (2003) *Agrobiodiversity: Learning from Farmers Across the World*, 343pp. United Nations: University Press.

Brush SB (2004) *Farmers Bounty: Locating Crop Diversity in the Contemporary World*, 352pp. New Haven, CN: Yale University Press.

Brussaard L, de Ruiter PC, and Brown GG (2007) Soil biodiversity for agricultural sustainability. *Agriculture, Ecosystems and Environment* 121(3): 233–244.

Caldeira MC, Ryel RR, Lawton JH, and Pereira JS (2001) Mechanisms of positive biodiversity-production relationships: Insights provided by δ13C analysis in experimental Mediterranean grassland plots. *Ecology Letters* 4: 439–443.

Cassman KG, Doberman A, Walters DT, and Yang H (2003) Meeting cereal demand while protecting natural resources and improving environmental quality. *Annual Review of Environmental Research* 28: 1–44.

Cassman KG and Wood S (2005) Cultivated systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystems and Human Well-Being Current State and Trends*, Ch 26. Washington, DC: Island Press.

Ceccarelli S (1996) Adaptation to low/high input cultivation. *Euphytica* 92: 203–214.

Clay JW (2004) *World Agriculture and the Environment: A Commodity-by-Commodity Guide to Impacts and Practices*, 282pp. Washington, DC: Island Press.

Collins WW and Qualset CO (1999) *Biodiversity in Agroecosystems*, 334pp. Boca Raton, FL: CRC Press.

Cooper HD, Spillane C, and Hodgkin T (eds.) (2001) *Broadening the Genetic Base of Crop Production*, 452pp. New York, USA: CABI Publishing.
DEFRA (2004). Indicators of Sustainable Development. Sustainable Development Unit, Department of Environment, Food, and Rural Affairs, UK, http://www.defra.gov.uk/corporate/sdstrategy/chapter3.htm (accessed November 2006).
de Wit CT (1960) On competition. *Verslagen van Landbouwkundige Onderzoekingen* 66(8): 1–82.
Donald PF (2004) Biodiversity impacts of some agricultural commodity production systems. *Conservation Biology* 18: 17–37.
Donald PF, Green RE, and Heath MF (2001) Agricultural intensification and the collapse of Europe's farmland bird populations. *Proceedings of the Royal Society of London B* 268: 25–29.
Drinkwater LE, Wagoner P, and Sarrantonio M (1998) Legume-based cropping systems have reduced carbon and nitrogen losses. *Nature* 396: 262–265.
Duelli P and Obrist MK (2003) Regional biodiversity in an agricultural landscape: The contribution of seminatural habitat islands. *Basic and Applied Ecology* 4: 129–138.
Engels JMM, Rao VR, Brown AHD, and Jackson MT (2002) *Managing Plant Genetic Diversity*, 384pp. Wallingford, UK: CABI Publishing.
Evenson RE and Gollen D (2003) Assessing the impact of the Green Revolution, 1960 to 2000. *Science* 300: 758–762.
Fowler C and Hodgkin T (2004) Plant genetic resources for food and agriculture: Assessing global availability. *Annual Review of Environmental Research* 29: 143–179.
Gepts P (2002) A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Science* 42: 1780–1790.
Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–555.
Griffiths BS, Ritz K, and Bardgett RD (2000) Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions: An examination of the biodiversity-ecosystem function relationship. *Oikos* 90: 279–294.
Gurr GM, Wrattan SD, and Luna JM (2003) Multi-function agricultural biodiversity: Pest management and other benefits. *Basic and Applied Ecology* 4: 107–116.
Henderson IG, Fuller RJ, Conway GJ, and Gough SJ (2004) Evidence of declines in populations of grassland-associated birds in marginal upland areas of Britain. *Bird Study* 51(1): 12–19.
Hietala-Kiovu R, Lankoski J, and Tarmi S (2004) Loss of biodiversity and its social cost in an agricultural landscape. *Agriculture, Ecosystems and Environment* 103: 75–83.
Jackson L, Bawa K, Pascual U, and Perrings C (2005) *Agrobiodiversity: A New Science Agenda for Biodiversity in Support of Sustainable Agroecosystems*, 40pp. Paris, France: DIVERSITAS Report No. 4.
Joshi L, Wibawa G, Beukema H, Williams S, and van Noordwijk M (2003) Technological change and biodiversity in the rubber agroecosystem of Sumatra. In: Vandermeer J (ed.) *Tropical Agroecosystems*, pp. 133–157. FL, USA: CRC Press.
Kates RW and Parris TM (2003) Long-term trends and a sustainability transition. *Proceedings of the National Academy of Sciences* 100: 8062–8067.
Kleijn D and Sutherland WJ (2003) How effective are European agri-environment schemes in conserving and promoting biodiversity? *Journal of Applied Ecology* 40: 947–969.
Kremen C, Williams NM, and Thorp RW (2002) Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences* 99: 16812–16816.
Lewis WJ, van Lenteren JC, Phatak SC, and Tumlinson JH III (1997) A total system approach to sustainable pest management. *Proceedings of the National Academy of Sciences* 94: 12243–12248.
Loreau M, Mouquet N, and Gonzalez A (2003) Biodiversity as spatial insurance in heterogenous landscapes. *Proceedings of the National Academy of Sciences* 100: 12765–12770.
Loreau M, Naeem S, and Inchausti P (eds.) (2002) *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*, 295pp. Oxford, UK: Oxford University Press.
Mäder P, Bach AF, Dubois D, Gunst L, Fried P, and Niggli U (2002) Soil fertility and biodiversity in organic farming. *Science* 296: 1694–1697.
Matson PA, Parton WJ, Power AG, and Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 277: 504–509.
McNeely JA and Scherr SJ (2003) *Ecoagriculture: Strategies to Feed the World and Save Wild Biodiversity*, 323pp. Washington, DC: Island Press.
MEA (Millennium Ecosystem Assessment) (2005) *Ecosystems and Human Well-Being: Biodiversity Synthesis*, 100pp. Washington, DC: World Resources Institute.
Mooney HA, Cooper A, and Reid W (2005) Confronting the human dilemma. *Science* 434: 561–562.
Murdiyarso D, van Noordwijk M, Wasrin UR, Tomich TP, and Gillison AN (2002) Environmental benefits and sustainable land-use options in the Jambi transect, Sumatra. *Journal of Vegetation Science* 12: 429–438.
Perrings C (1998) Resilience in the dynamics of economy–environment systems. *Environmental and Resource Economics* 11: 503–520.
Perrings C, Jackson L, and Bawa K (2006) Biodiversity in agricultural landscapes: Saving natural capital without losing interest. *Conservation Biology* 20: 263–264.
Pretty J and Smith D (2004) Social capital in biodiversity conservation and management. *Conservation Biology* 18: 631–638.
Reardon T and Timmer CP (2005) Transformation of markets for agricultural output in developing countries since 1950: How has thinking changed? In Handbook of Agricultural Economics: Agricultural Development: Farmers. In: Evenson R, Pingali P, and Schultz TP, (eds.) *Farm Production, and Farm Markets*, vol. 3A, pp. 195–205. Holland: Elsevier.
Ricketts TH, Daily GC, Ehrlich PR, and Michener CD (2004) Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences* 101: 12579–12582.
Robertson GP and Swinton SM (2005) Reconciling agricultural productivity and environmental integrity: A grand challenge for agriculture. *Frontiers in Ecology and the Environment* 3: 38–46.
Rosenzweig ML (2003) Reconciliation ecology and the future of species diversity. *Oryx*. 37: 194–205.
Sinclair TR, Purcell LC, and Sneller CH (2004) Crop transformation and the challenge to increase yield potential. *Trends in Plant Science* 9: 70–75.
Srivastava JP, Smith NJH, and Forno DA (eds.) (1996) *Biodiversity and Agricultural Intensification: Partners for Development and Conservation*. Washington, DC, USA: The World Bank.
Swift MJ, Izac AMN, and van Noordwijk M (2004) Biodiversity and ecosystem services in agricultural landscapes—are we asking the right questions? *Agriculture, Ecosystems and Environment* 104: 113–124.
Swihart RK and Moore JE (2004) *Conserving Biodiversity in Agricultural Landscapes: Model Based Planning Tools*. USA: Purdue University Press.
Thornton PK, Kruska RL, Henninger N, *et al.* (2002) *Mapping Poverty and Livestock in the Developing World*, 124pp. Nairobi, Kenya: ILRI (International Livestock Research Institute).
Tilman D, Reich P, Knops J, Wedin D, Mielke T, and Lehman C (2001) Diversity and productivity in a long-term grassland experiment. *Science* 294: 843–845.
Tscharntke T, Klein AM, Steffan-Dewenter I, and Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity—ecosystem service management. *Ecology Letters* 8: 857–874.
van Bruggen AHC and Termorshuizen AJ (2003) Integrated approaches to root disease management in organic farming systems. *Australasian Plant Pathology* 32: 141–156.
van Noordwijk M (2002) Scaling trade-offs between crop productivity, carbon stocks and biodiversity in shifting cultivation landscape mosaics: The FALLOW model. *Ecological Modelling* 149: 113–126.
van Noordwijk M, Chandler FJ, and Tomich TP (2004) *An Introduction to the Conceptual Basis of RUPES: Rewarding Upland Poor for the Environmental Services They Provide*, 46pp. Bogor: ICRAF-Southeast Asia.
Vandermeer J and Perfecto I (1995) *Breakfast of Biodiversity: The Truth about Rainforest Destruction*, 192pp. Oakland, CA: Food First Books.
Wardle DA, Bardgett RD, Klironomos JN, Setala H, van der Putten WH, and Wall DH (2004) Ecological linkages between aboveground and belowground biota. *Science* 304: 1629–1633.
Wood S and Lenné JM (2005) 'Received Wisdom' in agricultural land use policy: 10 Years on from Rio. *Land Use Policy* 22: 75–93.
Wood S, Sebastian K, and Scherr SJ (2000) *Analysis of Global Ecosystems: Agroecosystems*, 94pp. Washington, DC: International Food Policy Research Institute and World Resources Institute.
WRI (World Resources Institute) (2005) *The Wealth of the Poor: Managing Ecosystems to Fight Poverty*. Washington, DC: World Resources Institute.; http://pdf.wri.org/wrr05_full_hires.pdf (accessed on January 2007).
Zhu YY, Chen HR, and Fan JH (2000) Genetic diversity and disease control in rice. *Nature* 406: 718–722.

Air Pollution

Mike Ashmore, Stockholm Environment Institute, York, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acidification The process of soil or water becoming more acidic, with a loss of base cations (calcium, magnesium, potassium) and an increase in hydrogen and aluminum ion concentrations.

Acidophytic Species of plant adapted to grow on an acidic substrate.

Acute toxicity A damaging effect caused by a single short period of exposure to high concentrations of a pollutant.

Bioaccumulation The accumulation of toxic substances within an organism, primarily due to the rate of uptake by the organism being significantly greater than the loss rate.

Chronic toxicity A damaging effect caused by a long period of exposure to low or moderate concentrations of a pollutant.

Deposition The rate of influx of a substance from the atmosphere, usually expressed as mass per unit area of ground.

Detoxification Reduction in the toxic effect of a pollutant by its chemical or biological transformation to another less toxic chemical.

Eutrophication The enhancement of nutrient levels within an ecosystem, due to direct discharge into waters or aerial deposition to terrestrial systems.

Nitrophytic Species of plants that either tolerate or require high levels of nutrients, and in particular nitrogen.

Ombrotrophic Ecosystems or plant species only receiving nutrients in precipitation.

Phytotoxic Harmful to plants.

Air pollution results from emissions from a range of sources, including industry, power generation, transport, and domestic sources. Its local impacts on biodiversity can be dramatic close to large point sources of emissions, but significant impacts on biodiversity over much wider areas can also result from the long-range transport of pollutants. This article provides a brief description of air pollution sources and pathways, and then considers the major mechanisms by which air pollution can affect biodiversity. It then describes the evidence of such impacts in the field, focusing on a limited number of case studies. It finally considers the issues connected with establishing credible air quality guidelines to protect biodiversity, and the implications of global changes in air pollution concentrations for biodiversity.

Sources and Environmental Pathways of Air Pollutants

Historical Overview

The impact of air pollutants derived from human activity on vegetation and animals has been recognized for many centuries. The English diarist John Evelyn wrote a pamphlet on air pollution in London, published in 1661, in which he noted, "It is the horrid smoke which kills our Bees and Flowers abroad, suffering nothing in our gardens to bud, display themselves or ripe." There is also evidence from historical records of changes in biodiversity as a result of pollution. For example, until the early 1800s the dominant vegetation of the upland bogs of the Peak District in the southern Pennines of England was *Sphagnum* moss species. However, with the onset of the industrial revolution in the surrounding valleys, air pollutant levels in the region increased rapidly. The *Sphagnum* species began to disappear, to be replaced by vegetation dominated by cotton grass (*Eriophorum vaginatum*). Overall, there has been a reduction in species diversity, as species associated with the ombrotrophic *Sphagnum* species, such as sundews (*Drosera* species) and marsh andromeda (*Andromeda polifolia*), have also been lost. There is little doubt in this case that the change in community composition was a result of the direct elimination of the sensitive *Sphagnum* species by increased sulfur deposition.

During the past century, the range and quantities of chemicals discharged into the atmosphere from industry, transport, agriculture, energy production, domestic heating, and many other human activities, have increased dramatically. Once discharged into the atmosphere, these compounds are physically dispersed in the atmosphere and may undergo chemical transformations that alter their potential environmental impact. The environmental impact of air pollutants will depend on their concentration in the environment, or on the dose received or accumulated by the target organism.

Sources, Distribution, and Effects of Major Air Pollutants

Air pollutant problems vary greatly in their spatial scales. Some are very local in character, with the environmental impact of the pollution restricted to the immediate vicinity of, for example, a road or a factory. Other problems are regional in character as a result of the long-range transport of pollutants such as acid deposition and tropospheric ozone. Similarly, pollutant impacts may vary on different temporal scales. Some impacts, for example, are the result of an accidental release of large pollutant concentrations, which may cause an immediate impact on biodiversity, and from which there may be a slow and gradual recovery, whereas others are the result of an accumulation of pollutant deposition over years or even decades.

This article provides an overview of the ways in which pollutants can affect ecological processes and biodiversity.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00283-5

Table 1 Summary of major sources and impacts of air pollutants of relevance to biodiversity

Pollutant	*Major sources*	*Major impacts*	*Scale of effects*
Sulfur dioxide (SO_2)	Power generation: industry, domestic, and commercial heating	Forest decline; elimination of lichens and bryophytes	Local
Nitrogen oxides (NO_x) and ammonia (NH_3)	Power generation and transport (NO_x); intensive agriculture (NH_3)	Altered plant growth and enhanced stress sensitivity; soil acidification and eutrophication	Local, regional
Acid deposition	Secondary pollutant formed from SO_2 and NO_x	Soil and freshwater acidification: forest decline	Regional
Ozone (O_3)	Secondary pollutant formed from hydrocarbons and NO_x	Reduced plant growth: forest decline	Regional
Toxic metals (e.g., lead and cadmium)	Smelting industry; transport (lead)	Reduced soil microbial activity; reduced soil invertebrate populations	Local, regional
Persistent organic pollutants (POPs)	Industry; fuel combustion; pesticide use	Bioaccumulation in food chain	Local, global

Inevitably, it is not possible to provide a comprehensive account of the effects of the vast range of contaminants emitted into the atmosphere by human activity; therefore, this article will concentrate on those air pollutants for which the greatest evidence exists of impacts on biodiversity. **Table 1** summarizes the major sources of these pollutants, their major ecological impacts, and the spatial scale of these impacts.

Deposition and Environmental Pathways

The pathways by which pollutants enter ecosystems may have direct relevance to the nature of their impacts. When pollutants enter as gases, they may be taken up directly into foliage, through stomata, or directly inhaled by animals. However, in other cases pollutants may enter ecosystems in rainfall, in cloud or mist, or as particles.

The pathways of aerially deposited pollutants through ecosystems, and the mechanisms by which they affect biodiversity, are varied. There are three major groups of air pollutants considered in this article which are of significant concern in terms of biodiversity and which differ generically in the nature of their pathways from the atmosphere to the site of biological impact.

1. Acidifying and eutrophifying pollutants, such as sulfur dioxide, nitrogen oxides, and ammonia, which are also deposited as sulfate, nitrate, and ammonium in wet deposition. These gases may have direct effects on organisms, but over longer periods they can also lead to acidification and eutrophication of soils and freshwaters, with implications for biodiversity.
2. Photochemical oxidants, primarily ozone, which are not emitted directly into the atmosphere but are secondary pollutants formed as a result of reactions involving nitrogen oxides and volatile organic compounds, which require sunlight and high temperatures. Hence photochemical oxidants are characteristic of warm climates. The main impacts of ozone on biodiversity are through direct uptake into, and effects on, leaf tissue.
3. Metals are deposited primarily in rainfall or in particulate matter, but atmospheric deposition is only one source of these chemicals; they can also result from soil contamination or from discharges into the marine or freshwater systems. The impacts of metals result primarily from their accumulation in soils at concentrations which are toxic to soil organisms or plant roots, or through leaching into freshwaters.

There are a much wider range of specific pollutants which are not considered in this article but which may have impacts around specific industrial sources – for example, hydrogen fluoride and hydrogen chloride. Another group of chemicals, persistent organic pollutants (POPs), are of concern because of the potential for significant bioaccumulation in the food chain. There is evidence that the atmosphere can act as a pathway to redistribute these compounds through "global distillation," in which compounds volatilize at ambient temperatures in warmer parts of the planet and are then redeposited at cooler latitudes. Thus, it is the bioaccumulation of these compounds in the polar regions which is of greatest concern; for example, Antarctic fish remote from any immediate pollution sources have been reported to contain concentrations of certain POPs which are as high as those in North Sea fish.

Mechanisms of Impacts on Biodiversity

Long-Term and Short-Term Effects on Individual Organisms

Before considering the impacts of air pollution on biodiversity at a community level, it is essential to assess the ways in which individual organisms respond to air pollutants. The impact of a pollutant on any individual organism is complex and may involve many factors. The most important of these factors is the dose of the pollutant received. This will depend on the concentration of the pollutant in the atmosphere, or other relevant medium, and the duration of exposure to it. Short-term effects of air pollution exposure at high concentrations result in acute toxicity which is usually characterized by direct

damage to exposed tissue and visible leaf injury on vegetation. In contrast, long-term effects of air pollution exposure, which may result from much lower concentrations, result in chronic toxicity, which is usually characterized by alterations in physiology, growth, and reproduction. In general, the relationship between pollutant concentration, or dose, and the response of an organism is classified as follows.

1. When the pollutant is present in sufficiently high concentrations, the organism may be killed outright.
2. At lower concentrations, the organism is able to survive, but its performance is adversely affected. For example, rates of growth or reproduction may be reduced, or changes may occur in the patterns of development or in resource allocation.
3. At even lower concentrations, the physiology and growth of the organism may be unaffected under optimal environmental conditions, but the pollutant may cause subtle morphological, physiological, or behavioral changes which lead to an altered tolerance of other environmental stresses.
4. Finally, in the case of certain pollutants, there may be positive effects at low concentrations. For example, sulfur dioxide can stimulate plant growth at low concentrations, especially in soils which are sulfur deficient, whereas nitrogen (N) deposition can also increase the growth rate of many fast-growing competitive plant species

This range of different biological responses, depending on the severity of the air pollution stress, means that different types of mechanisms exist for impacts on biodiversity. Thus, for 1, direct elimination of species may lead directly to a reduction in biodiversity. For 2, reductions in growth rates may affect the longer term viability of populations and lead to local extinctions; however a more likely outcome is a shift in competitive balance between species, which may lead to a changed community composition. In the case of 3, the pollutant has no direct effect on biodiversity but may exacerbate, or mitigate, the impacts of other stress factors. Finally, for 4, although the effect of the air pollutant on a species may be positive, increases in growth of one species may alter the competitive balance and lead to reduced populations or local extinctions of other species. Air pollution may also have effects on some species indirectly through its detrimental impacts on other components of the ecosystem. For example, elimination of forest cover through direct effects of air pollution will cause large changes in the microclimate of the forest floor, with consequences for ground vegetation and soil biodiversity. Several of these mechanisms may operate simultaneously, as explained for the case of N deposition in the section Effects of Nitrogen Deposition on Plant Species Richness and Diversity.

There is a large degree of variation between species in their sensitivity to particular pollutants; there is also often substantial genetic variation in response within species. This variation broadly relates to the ability of the organism to restrict pollutant uptake or, once it has been taken up, to detoxify, metabolize, or sequester the pollutant. Within the same genotype, other factors such as age and growth stage may also influence sensitivity to air pollutants. Finally, the dynamics of these responses are important, with pollutant exposure frequently inducing adaptive biochemical, physiological, or morphological responses which lead to a reduction in its adverse effects.

Interactions between Air Pollutants and Other Environmental Factors

The effect of a given exposure to an air pollutant may also depend on other environmental conditions which can modify responses to the pollutant. First, they may modify pollutant dose; for example, soil water stress may lead to stomatal closure and thus to a reduced uptake of air pollutants by plants and hence reduced pollutant damage. Second, environmental factors may reduce the capacity of an organism to detoxify and assimilate pollutants; for example, it is well established, from both controlled experiments and field observations, that sulfur dioxide is more phytotoxic when plants are growing at low temperatures or under low light conditions. Third, exposure to pollutants may lead to changes in the morphology or physiology of the organism which make it more sensitive to environmental stresses; for example, increased deposition of N has been shown to increase the likelihood of defoliating pest outbreaks in heathland ecosystems.

Air pollutants rarely occur alone, and the responses to pollutant mixtures may be very different from those to the individual constituent pollutants. These interactions can operate in several different ways. First, two pollutants taken up together may have a greater effect than would be expected from knowledge of the effects of each pollutant – a so-called synergistic response. For example, uptake of sulfur dioxide and nitrogen dioxide together often has synergistic effects on vegetation. Although synergistic interactions between air pollutants have received considerable attention, there are many instances in which the effect of a pollutant mixture is not different from that of the individual pollutants or indeed there is a reduced effect – an antagonistic interaction. Second, over the longer term, the deposition of one pollutant can affect the uptake and impacts of another contaminant; for example, there is evidence that freshwater acidification resulting from deposition of sulfate and nitrate can cause increased bioaccumulation of metals such as lead, cadmium, and mercury in fish and birds. Finally, problems such as forest decline may result from complex interactions between a whole range of air pollutants, biotic, and abiotic stresses, and management factors.

Evolutionary Responses to Air Pollution Exposure

In competitive situations, it is likely that selection pressures will act in favor of individuals that are more resistant to a particular stress factor, and there is no reason that air pollution should be an exception to this, given the established intraspecific variation in response to pollution. A well established example is the evolution of industrial melanism in moths and ladybirds. Darker forms of these species have a selective advantage because of reduced predation where trees have a darkened bark, and a reduced lichen cover, caused by high concentrations of sulfur dioxide (SO_2) and smoke.

Improvements in air quality have led to a reversal of this evolutionary trend, with a decreased proportion of darker forms in the population.

Evolution of tolerance to air pollution has also been demonstrated in vegetation. Studies of emission sources in both the USA and the UK have shown that tolerance to SO_2 in annual and perennial species increases with increasing concentrations closer to the source. There is also evidence that decreasing urban SO_2 concentrations are associated with a loss of this tolerance in grass species, suggesting that the SO_2-tolerant genotypes are at a selective disadvantage in the absence of the pollutant. Where the pollutant is more widely dispersed, it is more difficult to demonstrate the phenomenon, although it would be expected to occur. For example, temporal and spatial associations between ozone exposures and the ozone tolerance of local populations of the annual species *Plantago major* strongly suggest an evolution of tolerance.

There is little evidence of the impact of evolution of air pollution tolerance on the overall genetic variation, or fitness, of a population. Where pollution stress is high, the frequency of resistant individuals is low, and a large proportion of the population is eliminated, the effect on overall genetic variation in the population is likely to be substantial. However, with less severe pollution levels, the effects may be relatively small.

Effects on Interactions between Organisms

When an air pollutant is present at a concentration which affects the physiology, growth, or reproduction of individual organisms, it is clear that the potential exists to influence the outcome of competition between species. In the case of competition between plant species, simple experiments with air pollutants such SO_2 and ozone (O_3) have shown, as expected, that when a pollution-sensitive and pollution-resistant species (often a clover and grass species, respectively) are grown together, the presence of the pollutant shifts the balance in favor of the latter. However, although such simple experiments clearly demonstrate that pollutants can modify the outcome of plant competition, it is doubtful whether they provide much guidance about how air pollution modifies interactions between plant species in real communities. For example, the vertical stratification of forest communities, which cannot be readily reproduced in simple competition experiments, may have a major effect on community responses. Field studies of forest stands which have been damaged by industrial pollution typically report a decline in density of the dominant overstory tree species. However, in some cases this can be accompanied by an increase in density in the shrub layer as a result of the release of competitive suppression. Competitive exclusion is recognized as a major mechanism for N deposition causing reductions in species richness and loss of rare species, as fast-growing species with a high N demand increase their growth rates in response to the additional N.

Air pollution is also known to alter the relationships between plants and insect herbivores. For example, high concentrations of nitrogen oxides are probably a major factor in the high numbers of aphids, and other phytophagous insects, found alongside major roads. **Figure 1** illustrates some of the potential interactions between air pollution, insect pests, and host plants. At high concentrations, there may be direct effects on the insects, but at lower concentrations the effect is mediated primarily through chemical changes in the host plant. In the case of SO_2 and NO_2, the effects on amino acid composition may be particularly important, whereas increased leaf N concentrations as a result of elevated N deposition can increase their palatability to most insect herbivores. Effects on the natural enemies of insect herbivores are another possible mechanism of response. These primary interactions can, as **Figure 1** demonstrates, lead to a range of secondary interactions affecting both plant performance and populations of other species at higher trophic levels.

Air pollutants can similarly modify the interactions between plants and fungal pathogens. Many pathogens are very sensitive to pollutants such as SO_2, and the absence of certain diseases producing clear visible symptoms, such as tar spot on sycamore (*Rhytisma acerinum*), has been used as a bioindicator of elevated concentrations of SO_2. Chemical and biological changes on the leaf surface induced by pollutants may critically affect the performance of fungal pathogens and other leaf microflora. Increased levels of pathogenic fungi have been reported when N deposition rates increase, possibly as a result of the decreased levels of phenolic compounds in the leaf. Pollutant exposure can also have an adverse effect on fungal mycorrhizal associations, which are crucial to the stress tolerance (including heavy metal tolerance) and competitive ability of many plant species. This may occur because of changes in soil chemistry, for example, due to mobilization of aluminum by acid deposition or accumulation of heavy metals, or because of effects above-ground leading to reduced carbon partitioning to the roots.

Effects on Bird and Mammal Biodiversity

There is very limited evidence of effects of air pollution on bird or mammal populations. Almost certainly, any effects are not due to direct effects on these organisms but due to changes in food quality or prey species composition caused by air pollution. Although there have been suggestions that declines in certain bird species in north-west Europe are associated with effects of increased atmospheric N deposition, the supporting evidence of changes induced by air pollution across the entire food webs that support these bird populations is lacking. More direct evidence is available where aquatic food webs have been disrupted by water acidification. For example, decline of the common loon, a large fish-eating predator with iconic status in Canada, have been linked to reduced breeding success on acidified lakes because of the reduced availability of fish.

Evidence of Impacts on Biodiversity

The previous section summarized the key factors which may influence the impact of air pollution on biodiversity. In

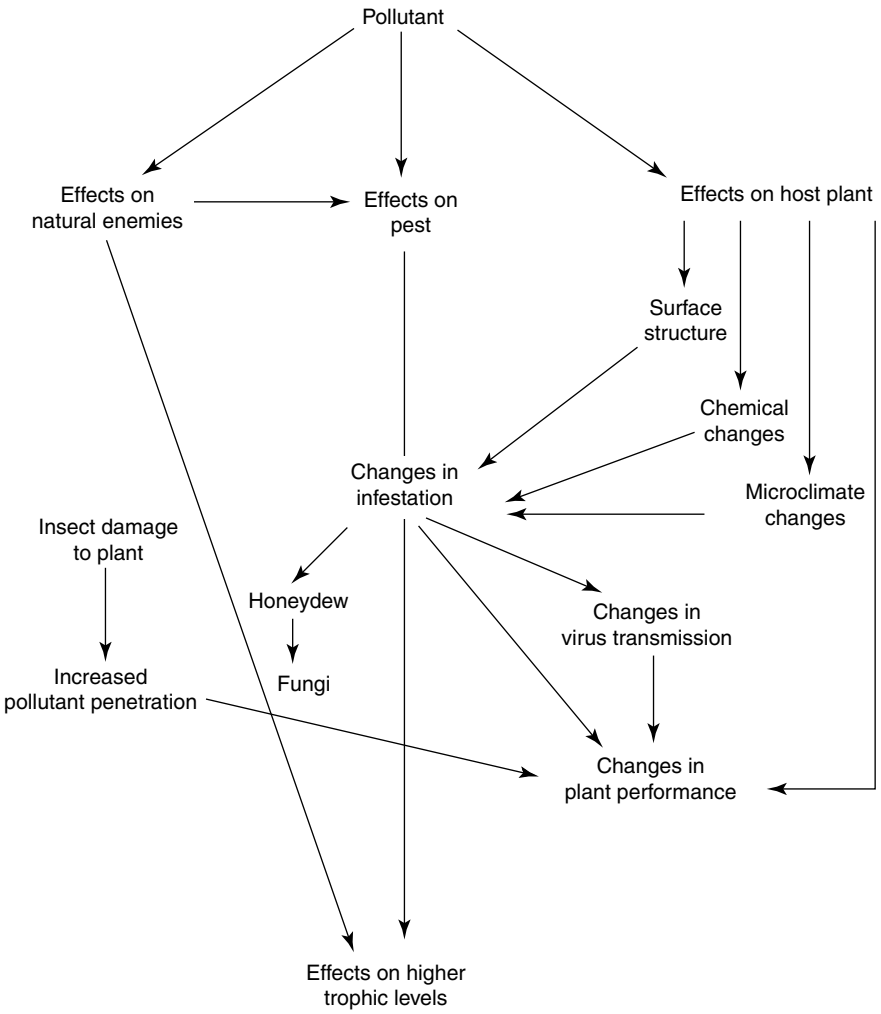


Figure 1 Model of potential interactions between air pollution, insect herbivores, and host plants. Reproduced from Bell JNB, McNeill S, Houlden G, Brown VC, and Mansfield PJ (1993) Atmospheric change: Effect on plant pests and diseases. *Parasitology* 106: S11–S24.

this section, some examples of actual impacts of air pollution on biodiversity in the field are discussed in more detail to illustrate the types of effects which have been found. In some cases, there is evidence of both the declines in biodiversity at historic high levels of air pollution and of recovery as air pollution levels have decreased. It should be noted that the evidence proving the link between exposure to air pollution and loss of biodiversity in the field is variable in quality. Where the effects are severe and localized, spatial associations between pollutant deposition and biodiversity may be readily established and, especially when the species present before operation of a local source are known, a causal link may readily be inferred. However, where the concern is related to a regional-scale deposition of air pollutants at more moderate levels and the effects on biodiversity are more subtle, possibly involving interactions between several factors, the causal link is much more difficult to establish from field observations alone, and experimental evidence allied to an understanding of the underlying mechanisms becomes more important. The examples discussed in the following sections illustrate a range of cases, from local to regional.

Effects of Sulfur Dioxide and Metals near Large Smelters

It is clear that, when present in high enough concentrations, air pollutants can cause the complete elimination of all plant species. For example, near the large Sudbury smelter in Ontario in the early 1970s, areas devoid of vegetation occurred up to 8 km from the source, and species numbers and productivity were reduced up to 20–30 km from the smelter. **Figure 2** illustrates the relationships between the stem basal area, species number, and the diversity index of overstory vegetation, and the distance from the smelter; similar relationships were found for ground vegetation, although the trend with distance was not as strong, and for insect communities. These effects were primarily attributed to the combined effects of SO_2 and metal emissions from the smelter. There was also evidence of acidification and metal accumulation in lakes near the smelter, with adverse effects on species numbers for plankton, macrophytes, and fish.

It is interesting to compare these results with the effects of large point sources of metals, which, unlike Sudbury, have

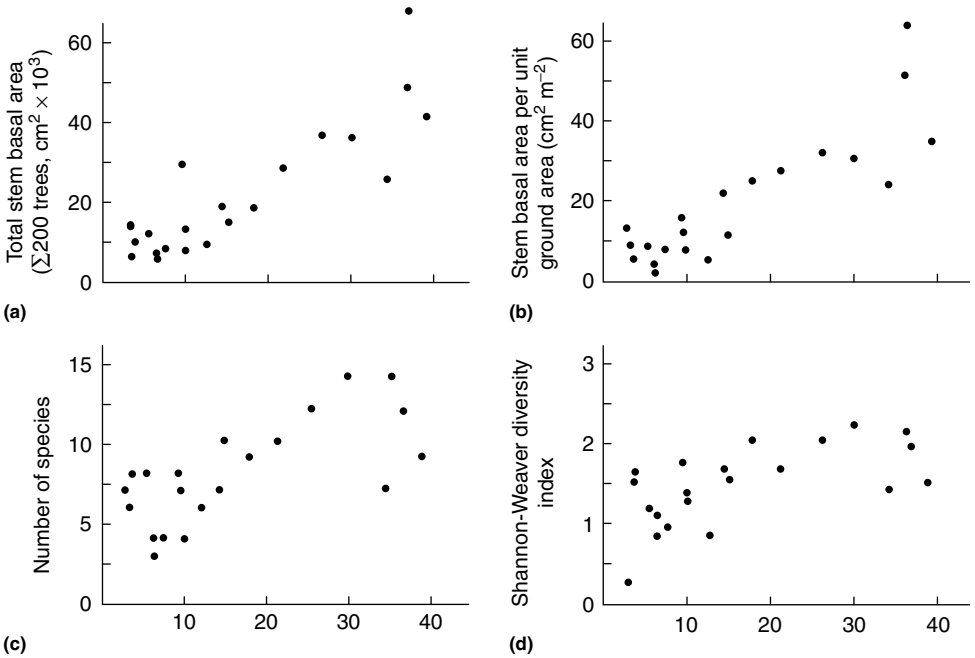


Figure 2 Forest overstory characteristics, expressed as (a) total stem basal area, (b) per unit ground area, (c) number of species, and (d) Shannon–Weaver diversity index, with distance south of the Sudbury smelter complex. Reproduced from Freedman B (1995) *Environmental Ecology*, 2nd edn. London: Academic Press, and redrawn from Freedman B and Hutchinson TC (1980) *Canadian Journal of Botany* 58: 2123–2140.

minimal SO_2 emissions. One of most intensively studied areas is near the brass works at Gusum, Sweden, where the metals have accumulated in the surrounding soil over a period of three centuries. **Figure 3** shows the relationship between soil copper concentrations, which increase with proximity to the smelter, and the numbers of taxa of different groups of organisms. Lead and zinc are associated with copper, and may be contributing to these effects on biodiversity. Vascular plant species are relatively unaffected in terms of numbers of taxa, although cover and growth are reduced near the smelter; in contrast, the number of taxa of mosses, earthworms, ground lichens, and macrofungi are all significantly reduced at higher metal concentration.

The local impact of smelters also depends on soil and climate. For example, the impact of two large nickel smelters in the Kola peninsula, which produce large emissions of metals and SO_2, may be exacerbated by their location in the Arctic, where nutrient turnover rates are low and the low temperatures are likely to increase the sensitivity of vegetation to SO_2.

How fast can biodiversity recover once such large point sources of air pollution are closed? Evidence suggests that chemical recovery of soils from acidification, and the accumulation of metals, may take several decades or more. Rates of chemical recovery vary between pollutants. For example, a series of studies in woodlands near a major metal smelter in south-west England revealed that cadmium and zinc concentrations in the surface horizons have decreased since a reduction in emissions in the mid-1970s but that concentrations of lead and copper have continued to increase in the surface layers. The biological implications of these slow rates of chemical recovery are illustrated by studies since the closure of the Sudbury smelter complex, where the gradient of disturbance and impoverished biodiversity is still visible 30 years later up to a distance of 50 km. Active restoration of vegetation, for example, through liming, fertilization, and seeding and planting, has accelerated the recovery of vascular plant biodiversity; however, this has not necessarily led to the recovery of species richness and diversity in other groups, such as insects and nonvascular plants.

Effects of Air Pollution on Lichen Biodiversity

The impacts of air pollution, and SO_2 in particular, on lichen species provide an illustration of larger scale changes in biodiversity. The disappearance of many lichen species from European and North American cities during the nineteenth and twentieth centuries has been well documented; in the UK, between 30% and 90% of species were lost from areas in which air pollution was a dominant factor. Many of these species were highly sensitive to the direct effects of relatively low concentrations of SO_2. In contrast, other lichen species are relatively tolerant, and several different scales have been developed to map SO_2 concentrations based on the occurrence of particular groups of lichen species.

In many European and North American cities, SO_2 concentrations have decreased dramatically in recent decades, but the reinvasion of lichens in response to this decline has been patchy and variable. In London, for example, it has been found that some relatively pollution-sensitive lichens have reinvaded more quickly than species which are more tolerant. There are several possible explanations for these different rates of recovery. One factor is the speed of dispersal of vegetative propagules into areas of decreasing SO_2; certain

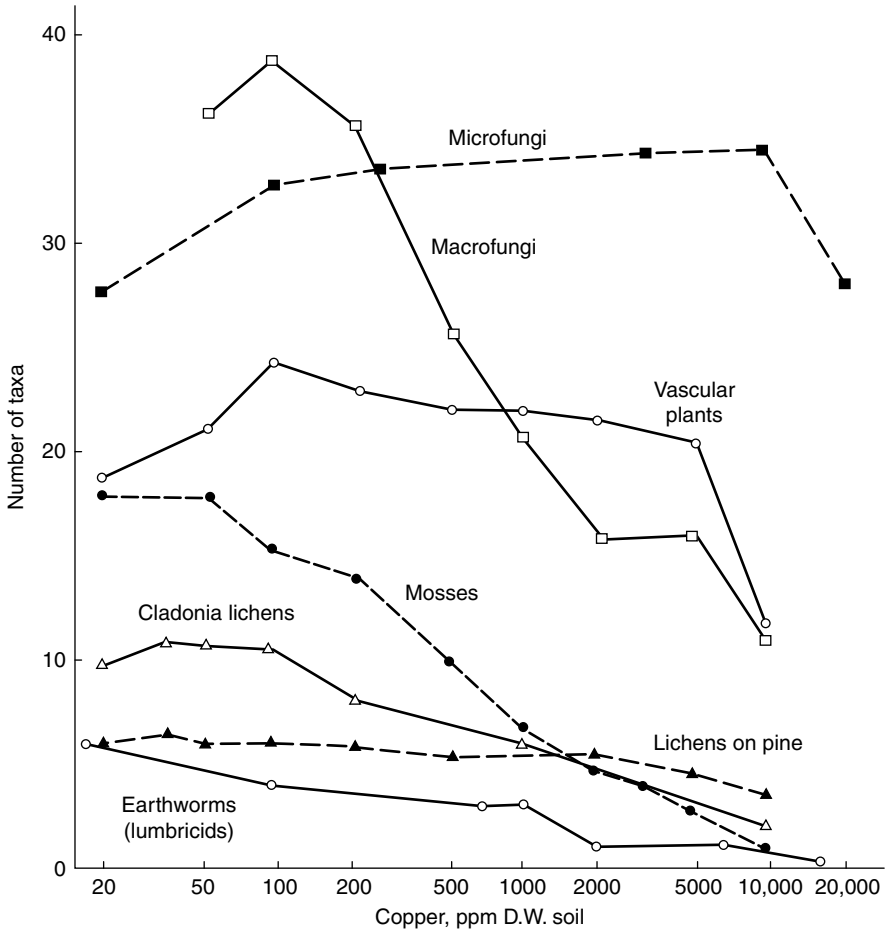


Figure 3 Species richness of different groups of organisms related to copper concentrations in the surface soil of forests near the Gusum brass works. Reproduced from Freedman B (1995) *Environmental Ecology*, 2nd edn. London: Academic Press, and redrawn from Tyler G (1994) *Ambio* 13: 18–24.

species (so-called "zone-skippers") are able to disperse much more rapidly than others. Another important factor is changes in bark chemistry. Recovery of epiphytic lichen species seems to be more rapid on tree species such as ash and willow, which have a relatively high pH bark, than on species with acidic bark, such as oaks. A third factor is the importance of other air pollutants, especially ammonia and nitrogen oxides, in influencing lichen biodiversity. For example, increased levels of ammonium deposition have also been associated with an increase in the cover and distribution of nitrophytic species and a loss of more acidophytic species in The Netherlands, whereas in London, there is evidence that the number of lichen species declines once nitrogen oxide concentrations exceed a critical threshold.

Effects of Nitrogen Deposition on Plant Species Richness and Diversity

The impacts of N deposition on terrestrial ecosystem biodiversity arise through a variety of mechanisms shown in **Figure 4**, the relative importance of which will depend on the ecosystems under consideration. These mechanisms include direct toxicity, eutrophication, acidification, and exacerbation of other stresses.

An example of significant changes in biodiversity which are caused by N deposition, and which involve a number of the mechanisms shown in **Figure 3**, is the loss in recent decades of heathlands dominated by ericaceous shrubs, such as *Calluna vulgaris*, which have been replaced by acid grassland communities dominated by *Molinia caerulea* and *Deschampsia flexuosa*. This has occurred in areas such as those in the Netherlands, where there are high rates of ammonia emission from intensive agriculture. Here, the primary mechanism is a shift in the competitive balance between *Calluna* and the grass species; experimental studies clearly demonstrate that young plants of the grass species are better able to respond to increased levels of N and thus can out-compete *Calluna*.

However, the field situation is more complex because of the effect of competition for light; once a *Calluna* canopy is established, invasion by grasses is unlikely, even when there are high levels of N availability. Increased N deposition also acts by increasing the sensitivity of *Calluna* to biotic and abiotic stress factors, which will lead to canopy breakdown, and hence an opportunity for grass invasion. In particular, outbreaks of heather beetle infestations, which are favored by higher leaf

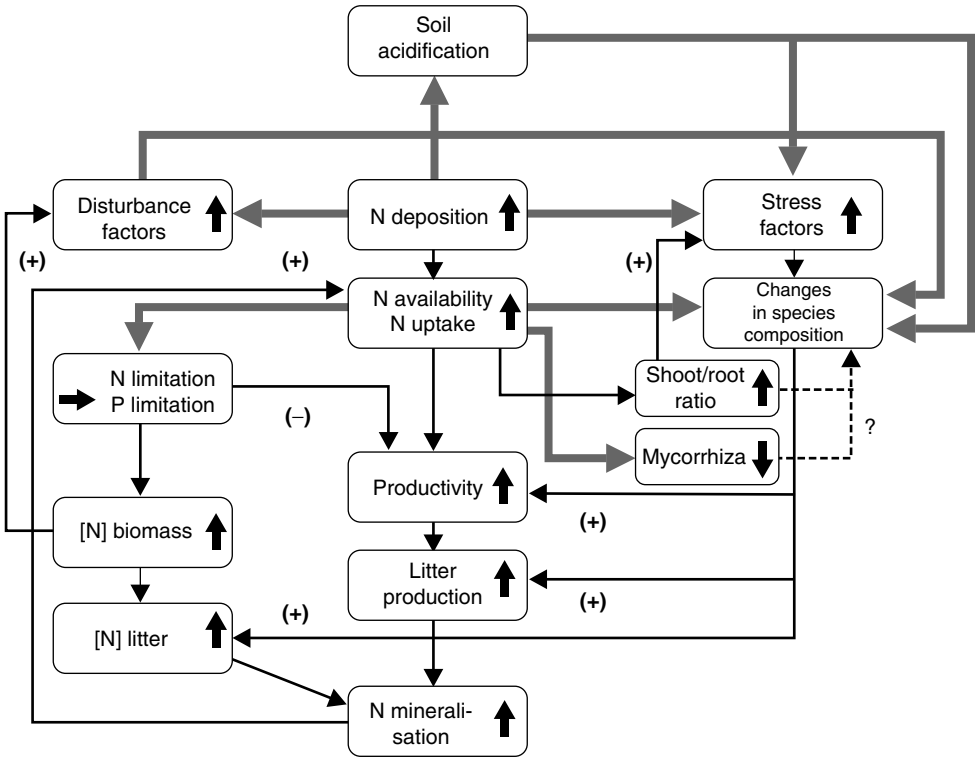


Figure 4 Summary of the main mechanisms for effects of increased nitrogen deposition on terrestrial ecosystems. The arrows inside the boxes indicate the direction of any effect, whereas feedbacks are indicated as positive (+) or negative (−). Black thin arrows indicate short-term effects, whereas gray thick arrows indicate long-term effects. Published from Bobbink and Lamers (2002).

N concentrations, can cause widespread defoliation of the *Calluna* canopy. Finally, it is likely that management practices are also significant; there have been reduced rates of grazing and sod-cutting in recent years, and these traditional management practices have been important in maintaining the low nutrient status of the heathland communities. Thus, in this case, changes in species composition may have resulted from a complex mixture of factors, including pollutant deposition, management practices, competition for light and nutrients, and the effects of climatic stress and insect herbivores.

A synthesis of the available evidence across Europe has identified a range of effects of elevated N deposition on terrestrial biodiversity including, *inter alia*: reduced species richness in acid grasslands; loss of characteristic of bryophyte species and carnivorous species, such as sundews, from peatlands; reduced species diversity and invasion of grasses in coastal dunes; and loss of epiphytic lichen species in forests. Close to large agricultural point sources, much of the N deposition is in the form of ammonia. There is a strong body of evidence of adverse effects, especially on lichen and bryophyte species, close to these point sources of ammonia. Nitrophytic species, which are sensitive to direct effects of high ammonia concentrations, decline, as do acidophytic epiphyte species, because ammonia, as a basic gas, increases tree bark pH. On a regional scale, focused field surveys have shown a close relationship between elevated N deposition and some components of biodiversity. For example, a survey of an acid grassland community in the UK showed a reduction in species richness as N deposition increased (**Figure 5**). In particular,

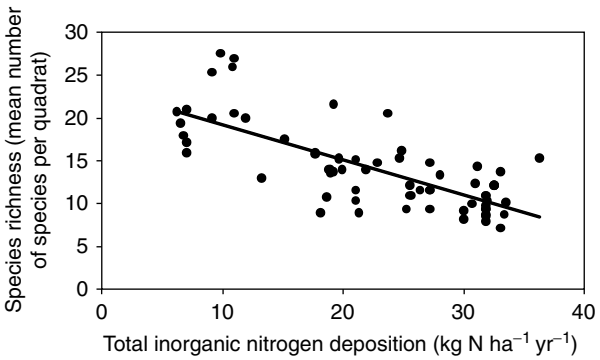


Figure 5 Species richness in 2 m × 2 m quadrats in acid grassland communities across the UK expressed as a function of nitrogen deposition. Reproduced from Stevens CJ (2004) Impact of nitrogen deposition on the species richness of grasslands. *Science* 303: 1876–1879.

forb species richness was reduced by about 75% across this gradient of N deposition. Detailed analysis of the survey results suggests that acidification was a stronger driver of the observed effects than eutrophication in these poorly buffered soils.

A number of recent studies in Europe have also analyzed data from historic and more recent large-scale ecological surveys. These studies have consistently shown either (1) a decline in species richness and, more specifically, a decline in species associated with nutrient-poor habitats or (2) little

change or even an increase in species richness, but a decline in species adapted to a low nutrient availability, including species of high conservation importance. Taken together, there is now a large body of experimental and field evidence to suggest that N deposition has had a significant adverse effect on biodiversity in many habitats of conservation importance in western and central Europe, and there is increasing evidence of such effects in North America.

As for effects of acidification and metal accumulation around smelters, the timescale for ecosystem recovery from long-term impacts of N deposition may be substantial. Most, but not all, experimental studies suggest that recovery of species number and diversity may take a minimum of 1–2 decades and, where the original cumulative deposition of N was large, recovery may take significantly longer. For example, in areas of the Netherlands where ammonium deposition has decreased greatly in the last 20–30 years, soil conditions still prevent the establishment of sensitive heathland and grassland species of high conservation value, and recovery of these species is likely to require active restoration of the necessary soil chemical conditions through nutrient removal and liming.

Effects of Ozone on Species Composition and Diversity

In the case of ozone, current trends in Europe and North America suggest that peak concentrations during photochemical episodes have declined over the past 20–30 years in response to national and regional measures to control emissions. However, at the same, northern hemispheric annual mean concentrations have increased significantly, primarily driven by global increases in emissions of precursors, and are projected to increase further over the current century. The implications of these changes in ozone exposure patterns for biodiversity are quite uncertain.

In the past, there was strong evidence from field observations linking cause and effect when ozone occurred in very high concentrations. For example, the area where the impacts of ozone on forests have been most intensively studied is in the San Bernadino mountains, which surround the city of Los Angeles. Effects of ozone began to be observed on the forest community in the 1960s. The most dominant species of these mixed-conifer forests prior to European settlement were ponderosa pine (*Pinus ponderosa*) and Jeffrey pine (*Pinus jeffreyi*) because of their tolerance of the frequent wildfires; however, these also proved to be the most sensitive species to ozone. In many areas, both species showed severe visible injury and reduced needle longevity. These were associated with reduced radial growth or even years with missing growth rings. Trees affected by ozone were more susceptible to attack by bark beetles, which was often the direct cause of mortality. Outbreaks of bark beetles were associated with drought years and high ozone concentrations. Regeneration in these forests was greater for tree species such as white fir or cedar species, which are more resistant to ozone. The patterns of change in community composition are confounded by the role of fire since current fire exclusion policies favor replacement of ponderosa pine and Jeffrey pine by more fire-sensitive species which also happen to be more ozone tolerant. Similar effects of long-term exposure to high concentrations of ozone have been reported in other forest ecosystems, such as the declines in fir vitality which have been reported in the mountains near Mexico City.

Recent experimental studies in chamber systems have shown significant effects of small increases in mean background ozone concentration on a range of grassland species. This suggests that the increases in background ozone concentrations that are currently occurring across the northern hemisphere have the potential to have large-scale adverse effects on biodiversity. However, experimental studies are typically for periods of 1–2 years, and very few experimental studies have examined the long-term effects of ozone under field conditions. Furthermore, those species identified as ozone-sensitive based on field surveys of visible symptoms are not necessarily those on which ozone has the greatest impacts on cover or biomass. For example, an experimental study of an early successional forest community showed that ozone decreased species richness, diversity, and evenness; however, the species which dominated the community in the highest ozone treatment was blackberry, which is considered to be highly ozone sensitive on the basis of visible injury responses.

There is an increasing recognition that long-term effects of ozone on biodiversity have to be considered as just one element of wider scale environmental change. Even in the San Bernadino mountains, where ozone concentrations were historically very high, analysis of changes in species composition over more than a 25-year period suggests that a number of factors are important, including management, N deposition, and water stress as well as ozone exposure.

Effects of, and Recovery from, Water Acidification

The previous sections have focused on terrestrial, rather than aquatic, biodiversity. However, one of the most substantial large-scale impacts of air pollution over recent decades has been the effects of surface water acidification, primarily due to deposition of sulfate, in sensitive regions of northern and western Europe and north-eastern North America. This water acidification has been associated with effects on a range of groups, including invertebrates, fish (especially salmonids), algae, macrophytes,and piscivorous birds as well as impacts on a range of other ecosystem services. These effects occurred once the limited buffering potential of base cations in catchment soils were exhausted, resulting in increased concentrations of hydrogen and aluminum ions that were toxic to a range of aquatic fauna, and reduced pH and alkalinity that affected a range of aquatic plants. Loss of acid-sensitive species then had wider effects on aquatic food webs.

Since the 1980s, there have been substantial reductions in sulfur emissions across these regions, although the size and timing of these reductions varies. The evidence of biological recovery is limited and variable, but illustrates the complexity of recovery in biodiversity once ecosystems have been significantly affected by the air pollution. In Canada, for example, the evidence of chemical and biological recovery is very limited, except in regions, such as that around Sudbury (*see* Effects of Sulfur Dioxide and Metals near Large Smelters),

where there have been dramatic reductions in sulfur deposition. In the UK, in contrast, evidence of recovery of invertebrate and fish populations is widespread following a process of chemical recovery; however, biological recovery is far from complete and the invertebrate assemblages that are developing are not identical to those found in similar streams that have not been affected by acidification. Indeed, pre-acidifcation species assemblages may never be completely replaced.

A number of factors have been identified that affect both the degree of chemical recovery and the nature of biological recovery. First, the degree of emission reduction needs to be sufficient to reverse acidification. Second, there is a potential for nitrate leaching to increase in regions with high rates of N deposition, reversing or delaying the recovery from acidification. Third, management is important; active restoration, for example through liming, has clearly accelerated chemical recovery in many locations, but areas with extensive planting of conifers, for instance in the UK, have tended to show greater degrees of acidification, and slower rates of chemical recovery, than those with no afforestation.

In the UK, dispersal is not a significant barrier to reestablishment of acid-sensitive invertebrate species, since adults arrive but fail to establish viable populations. One important factor restricting species re-establishment, even when mean pH has reached appropriate values is the effect of acid episodes. Such episodes, which may be linked, for example, to droughts or to sea-salt events, can lead to short periods when hydrogen and aluminum concentrations are still toxic to acid-sensitive groups such as mayflies. A trend to warmer wetter winters is also a factor, as this also increases the frequency of acid episodes. Such factors may explain why, although some acid-sensitive invertebrate species have reappeared, these are only a small proportion of the acid-sensitive species that were originally identified. Competition for ecological niches from acid-tolerant populations that have now become established may be another significant factor.

Air Quality Guidelines to Protect Biodiversity

The long-range transport of air pollutants, and their regional-scale impacts, means that local measures to protect biodiversity in specific areas, such as national parks or nature reserves, may have little value. For example, many of the national parks in the USA are affected by ozone pollution, and many sites of high conservation value in upland areas of the UK have been affected by acidification. It is therefore clear that the prevention of adverse effects on biodiversity requires measures to reduce emissions and to manage air quality in the regions of concern.

However, complete control of atmospheric emissions is impossible, and the costs of decreasing air pollution emissions typically increase greatly as the degree of removal increases. Thus, it is important to assess how far pollutant emissions should be reduced in order to ensure an acceptable degree of protection for biodiversity. In Europe, the concept of a critical load has been used to assess the need for further measures to protect biodiversity; a critical load is defined as "a quantitative estimate of an exposure to one or more pollutants below which significant harmful effects on specified sensitive elements of the environment do not occur according to present knowledge." The critical load refers to the pollutant input in all forms of deposition; critical levels, defined in terms of atmospheric concentrations over a given averaging time, have also been defined for pollutant gases. The critical load concept is based on a precautionary approach and relates to the long-term cumulative effects of deposition on ecosystem function and biodiversity. Thus, wherever current deposition rates exceed critical loads, but field studies demonstrate no adverse effects on biodiversity, action to reduce emission rates may still be justified to prevent future damage. Dynamic modeling can be a useful complement to the application of critical load values, by indicating the timescale over which changes in air pollutant emissions will lead to a change in soil or water chemistry that would be critical for particular groups of species.

Although the critical load concept has been valuable in providing an effects-based approach to international negotiations on trans-boundary air pollution in Europe, it is important to appreciate the difficulties of defining thresholds, in terms of pollutant deposition or concentration, for adverse effects on biodiversity. These difficulties arise from two major sources.

1. There are conceptual problems in defining what is meant by a significant harmful effect, both scientifically and in terms of societal judgment. For example, are small shifts in species composition of urban lichen communities, or changes in species composition of the invertebrate fauna of a woodland floor which do not affect nutrient cycling, actually significant harmful effects? This issue often needs to be addressed in areas where other human impacts, such as changes in land use, have already changed substantially the structure and composition of plant and animal communities.
2. Although the adverse effects of air pollution on biodiversity may be clear in situations in which they are present in high concentrations, as a threshold concentration is approached the effects become gradually more subtle and are difficult to detect under field conditions. Although experimental dose–response relationships may be able to define thresholds for short-term effects on single organisms under controlled conditions, the long-term cumulative effects of air pollution on complex communities cannot be reliably assessed using current experimental approaches.

Implications of Global Changes in Air Pollution for Biodiversity

In the past few decades, measures have been taken in Europe and North America which have reduced atmospheric emissions of many pollutants. Although recovery of biodiversity in ecosystems affected by these pollutants may take time, and may not result in the same community as that originally present in a particular location, major new impacts on biodiversity due to increases in atmospheric emissions are unlikely in these regions. In contrast, the past two decades have

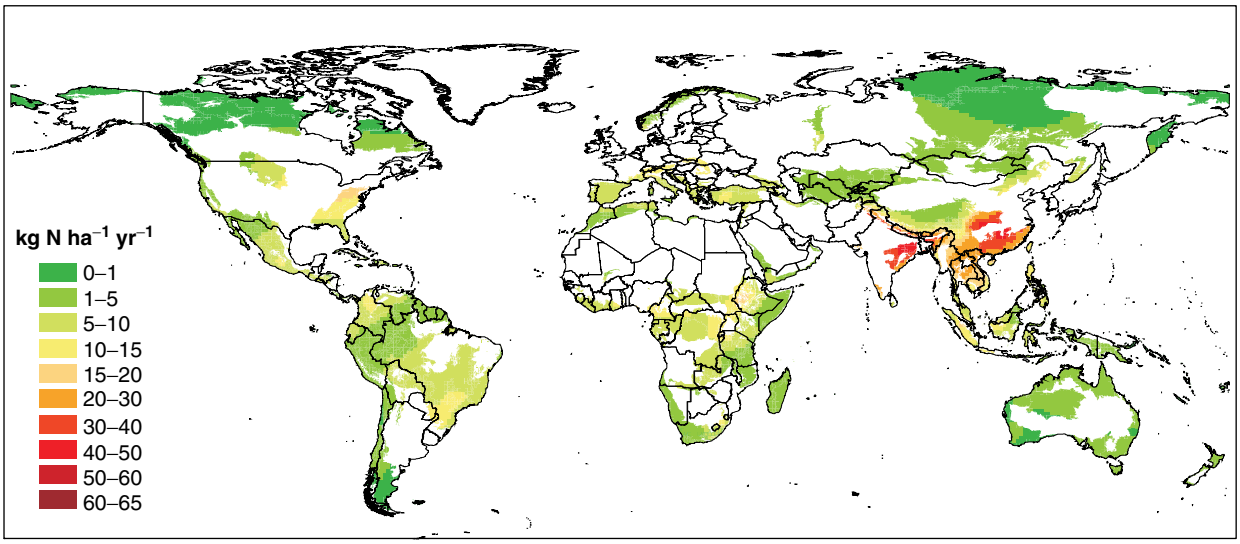


Figure 6 Distribution of nitrogen deposition in the G200 ecoregions of the world. Reproduced from Bobbink R, *et al.* (2010) *Ecological Applications* 20: 30–59, using modeled nitrogen deposition from Dentener, *et al.* (2006) and G200 data from Olson and Dinerstein (2002).

seen large increases in atmospheric emissions in parts of Asia, Africa, and Latin America.

To illustrate the threat to biodiversity from these global changes in patterns of air pollution emissions, **Figure 6** shows modeled rates of N deposition in 2000 in the G200 ecoregions, that is, global priority conservation areas that contain distinct and unique assemblages of natural communities and species. It is clear that high rates of N deposition were not only found in Europe and North America, but also in China and parts of south- and south-east Asia. In all these regions, critical loads that have been recommended to protect sensitive ecosystems in Europe and North America were exceeded in 2010. The consequences of excess N deposition for biodiversity were described in the section Effects of Nitrogen Deposition on Plant Species Richness and Diversity in general terms, but their specific impacts in sensitive ecosystems in east, south-east, and south Asia are unknown.

Model projections suggest that N deposition will increase not only in these rapidly developing regions of Asia, but also in many parts of Africa and Latin America, by 2030, although the validity of any projections depends critically on economic growth rates and the extent to which pollutant emissions are controlled. What is certain is that the authors lack experimental and field data to assess the consequences for local biodiversity of these increases in N deposition, as well as in sulfur deposition, ozone concentrations, and exposure to other pollutants. These large gaps in knowledge are critical for any assessment of the consequences of both current and future air pollution emissions for global biodiversity.

Concluding Remarks

This article has summarized some key aspects of air pollutant impacts on biodiversity. Although the text has, for the sake of clarity, focused on a limited range of pollutants, many of the same principles apply in the case of other air pollutants and other ecological impacts. Although in the past, very high concentrations of air pollutants were the dominant factor causing local changes in biodiversity, most of the situations of concern today in Europe and North America tend to involve more widely dispersed pollutants at lower concentrations, the effects of which may only become apparent over many years or even decades. In such cases, air pollution is no longer the single dominant factor but one of a range of biological, climatic, and soil factors which may influence biodiversity. If we want to fully understand the role of air pollution in such situations, it is essential that we gain further understanding of the ways in which pollution can interact with these other factors and base our analysis of pollutant impacts more clearly within an ecological framework. It is also vital that we acquire more knowledge in the many areas of the planet where recent and future increases in air pollutant exposure are a significant potential threat to the characteristic species and ecosystems.

Appendix

List of Courses

1. Environment Science
2. Environment Studies
3. Biological Sciences

See also: Adaptation. Atmospheric Gases. Greenhouse Effect. Nitrogen, Nitrogen Cycle

References

Ashmore MR (2005) Assessing the global impacts of ozone on vegetation. *Plant Cell and Environment* 28: 949–964.

Bell JNB, McNeill S, Houlden G, Brown VC, and Mansfield PJ (1993) Atmospheric change: Effect on plant pests and diseases. *Parasitology* 106: S11–S24.

Bell JNB and Treshow M (2002) *Air Pollution and Plant Life*, 2nd edn. Chichester: John Wiley.

Bobbink R, Hicks K, Galloway J, *et al*. (2010) Global assessment of nitrogen deposition effects on terrestrial plant diversity: A synthesis. *Ecological Applications* 20: 30–59.

Dise NB, Asshmore MR, Belzazid S, *et al*. (2011) Nitrogen as a threat to European terrestrial biodiversity. In: Sutton MA, Howard CM, Erisman JW, *et al*. (eds.) *European Nitrogen Assessment*, ch. 20. Cambridge: Cambridge University Press.

Emberson L, Ashmore M, and Murray F (eds.) (2003) *Air Pollution Impacts on Crops and Forests: A Global Assessment*. London: Imperial College, Press.

Sokhi R (ed.) (2008) *World Atlas of Atmospheric Pollution*. London: Anthem Press.

Sutton MA, Reis S, and Baker SMH (eds.) (2009) *Atmospheric Ammonia: Detecting Emission Changes and Environmental Impacts*. New York: Springer.

Taylor GE Jr, Pitelka LF, and Clegg MT (eds.) (1991) *Ecological Genetics and Air Pollution*. New York: Springer-Verlag.

Wellburn A (1994) *Air Pollution and Climate Change: The Biological Impact*, 2nd edn. Harlow, UK: Longman.

Alpine Ecosystems

Christian Körner, University of Basel, Basel, Switzerland

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alpine Refers to the life zone above the climatic high-elevation treeline, irrespective of latitude. Though originating in the Alps (a pre-Roman word for high mountains), the term is applied globally. The reader should be aware that this term is often used in a much wider sense in common language, and is also applied to regions with mountains in general, including settlements and resorts, which is not the scientific meaning here.

Apomixis A very common but "hidden" mode of clonal propagation by seeds, the embryos of which are 100% genetic copies of the source plant. Seeds are produced without fertilization, but often pollination is required to induce apomictic seed production. Apomictic plants also reproduce sexually, but these are very rare events.

Clonal growth A vegetative mode of propagation and expansion of plants by runners, tillers, or plant fragment dispersal, which is very important in the alpine life zone. (See also apomixis.) Clonal plants also produce sexual offspring by seed, but their clonal propagules often show higher survival. All clonal offspring of one source plant have the same genome and hence belong to the same "genet."

Ecotype or ectotypic Refers to genetic (evolutionary) differentiation within a given species (a specific "race") that reflects an obvious advantage in a given environment. Ecotypic traits are retained when individuals are transplanted into a different environment where these traits have no advantage.

Life-form The size and stature of a plant under natural life conditions. Environmental constraints can cause the life-form to differ substantially from that of a plant that develops under more favorable conditions, where genotype morphology, the "growth form," finds full expression. For instance, the growth form "tree" may be modified to the life-form "shrub" in the alpine zone.

Microclimate The climate that plants, small animals, and soil microbes experience, and that differs substantially from the "macroclimate" reported by weather stations. This difference is related to surface warming by the sun or cooling at night, as well as wind shelter effects, and is largely driven by relief, exposure, ground cover, and plant stature.

Treeline Also known as the forest line, this describes the high-elevation limit of (usually fragmented) forest. Most often there is no "line" visible in the landscape, so treeline position represents a convention. "Outpost" tree individuals may occur at higher elevations (the tree species line), and the boundary of closed, tall forest with timber-size stems (the timber-line) commonly occurs at an elevation that is 50–150 m lower. The whole transition from timber-line to the tree species line is called the "treeline-ecotone."

The Alpine Life Zone

The lower boundary of the alpine life zone is, by definition, the natural climatic high-elevation treeline. Where a treeline is missing, as is the case in some dry continental areas or because of deforestation, the treeline-isotherm is still applied, irrespective of the presence or absence of trees. This is nothing more than a practical convention. The actual natural treeline and hence the lower end of the alpine zone do not form a sharp boundary, for patches of stunted trees and alpine plants often intermingle. This transition zone is termed the 'treeline-ecotone'. The common climatological denominator of this boundary is a mean temperature during the growing season of 6.5 °C, with a minimum length of the growing season of 94 days (Körner, 2012); this definition applies worldwide (see the discussion in Körner, 2003). The length of the growing season within the alpine life zone varies from 12 months in the tropical alpine zone to merely 4–6 weeks in alpine snowbed vegetation at higher latitudes.

About 4 million km^2 fall into the alpine life zone (*ca.* 2.6% of the global land area outside Antarctica). Because not all of this alpine area is covered by vegetation – some consists of bare rock, rock fields, scree, and glaciers – the vegetated alpine land area is estimated to be *ca.* 3 million km^2 (see Körner *et al.*, 2011).

Two-thirds of the global alpine area is situated in the temperate and subtropical zones and only 10% occurs in the tropics (**Table 1**). Partly this is because mountains in the tropics need to be very high, at least 3600–4000 m, to permit tropical alpine vegetation to occur, whereas at the polar circles mountain heights of only 600–800 m are required to support alpine habitat. These are the approximate treeline elevations of the respective climatic zones. At about 4000 m, treeline reaches its average highest elevations in the subtropics (local extremes up to 4900 m, **Figure 1**). At high latitudes (>65 to 70° N), alpine vegetation merges with arctic tundra. Despite a number of common taxa and the overwhelming influence of low mean temperature, the arctic tundra life zone is very different from the alpine zone in terms of climate, land surface structure, and vegetation, hence several authors (e.g., Löve and Billings) have recommended that alpine vegetation not be referred to as "alpine tundra." The upper limits of vascular plant occurrence are commonly 1000 to 1500 m above the lower limits of the alpine zone (the treeline), but some extreme high-elevation outposts of higher plants are found up to 4500 m in the temperate zone and up to 6400 m in the

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00217-3

Table 1 The global area (million km^2) of bioclimatic mountain belts (rugged terrain only), as defined for the Global Mountain Biodiversity Assessment (GMBA) Mountain Portal (www.mountainbiodiversity.org). GS = growing season. Temperatures refer to season mean air temperatures: M (%) = percent of total mountain area (100% = 16.5 million km^2), T (%) = percent of total terrestrial area outside Antarctica (100% = 134.6 million km^2). Total mountain area (16.51 million km^2) is defined by ruggedness, i.e., all terrain with at least 200 m of elevational contrast depicted among 9 cells of 30'' size in a 2′30'' pixel (4.6 × 4.6 km at the equator; Körner, *et al.*, in press)

	Area million km^2%	*M%*	*T%*
Thermal belts			
1. Nival (<3.5 °C, GS<10 days)	0.53	3.24	0.40
2. Upper alpine (<3.5 °C, GS>10 days<54 days)	0.75	4.53	0.56
3. Lower alpine <6.4 °C, GS<94 days)	2.27	13.74	1.68
The treeline			
4. Upper montane (>6.4<=10 °C)	3.39	20.53	2.51
5. Lower montane (>10<=15 °C)	3.74	22.64	2.78
6. Remaining mountain area with frost (>15 °C)	1.34	8.11	0.99
7. Remaining mountain area without frost (>15 °C)	4.49	27.22	3.34
Total	16.51	100.00	12.26

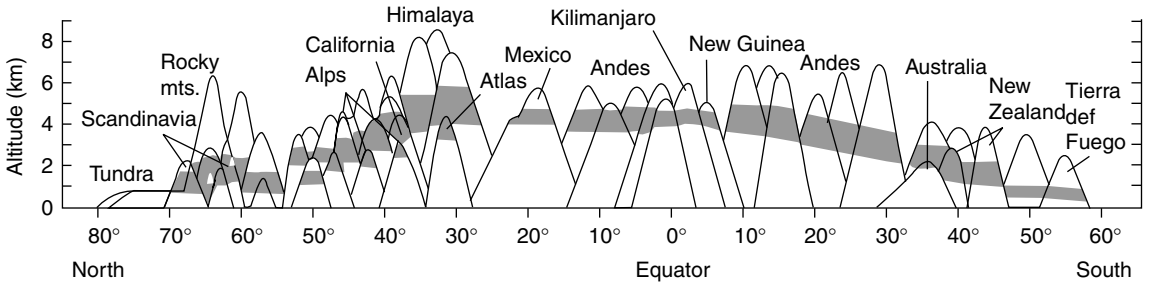


Figure 1 The alpine life zone occurs at all latitudes, though at contrasting elevations. Reproduced from Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

Table 2 The global area (million km^2) of alpine belts (upper alpine belt <3.5 °C, GS>10d <54d ; lower alpine belt (<6.4 °C, GS<94d) of rugged terrain only, as defined for the GMBA Mountain Portal (www.mountainbiodiversity.org). Temperatures refer to season mean air temperatures. M (%) = percent of total mountain area (100% = 16.5 million km^2), T (%) = percent of total terrestrial area outside Antarctica (100% = 134.6 million km^2). Rugged terrain as defined in **Table 1**

Region (million km^2)	*Continent*									
	As	*Eu*	*Af*	*N-A*	*Gld*	*S-A*	*Aus*	*Oce*	*Total*	*(%)*
All land area	44.57	9.84	30.01	22.08	2.15	17.82	7.7	0.45	134.62	100.0
All rugged area	8.9	0.92	1.19	2.92	0.1	2.16	0.13	0.18	16.51	12.26
Nival belt	0.2	0.06	<0.01	0.15	0.08	0.03	<0.01	<0.01	0.53	0.40
Upper alpine belt	0.45	0.04	<0.01	0.18	0.01	0.07	0	<0.01	0.75	0.56
Lower alpine belt	1.43	0.07	<0.01	0.45	<0.01	0.31	<0.01	0.01	2.27	1.68
Total alpine and nival	2.08	0.17	<0.01	0.78	0.1	0.41	<0.01	0.01	3.55	2.64

As, Asia; Eu, Europe; Af, Africa; N-A, North America; S-A, South America; GLD, Greenland; Aus, Australia and New Zealand; OCE, Oceania (including the large islands of SE Asia).

subtropics, where the uppermost individual of a higher plant was found in the Himalayas (Miehe, 1997). An important feature of alpine life is its isolation. Mountaintops with their high biological diversity represent islands or archipelagos surrounded by lowlands, where most alpine organisms cannot survive (**Table 2**).

Various aspects of plant diversity in high mountain systems have been reviewed previously. A selection of such synthetic papers or volumes is included in the References. Full bibliographic references to original studies (mentioned in the text by authors' names only) can be found in Körner (2003).

Habitat Diversity, a Key to Alpine Plant Diversity

Climatic Microhabitats

There is a common belief that the alpine life zone is very hostile to plants and animals and that low temperatures restrict life activities, including productivity. This generalization is wrong for two reasons: (1) as will be shown, temperatures are not always that low, and (2) this belief reflects our human perspective of what is "cold." For organisms adapted to alpine life conditions the temperatures they experience are not

necessarily "cold." In fact, if temperatures were higher, most of these organisms would suffer or even die, many of them because they would be outcompeted by other species that do better at such higher temperatures.

The most important characteristic of the alpine life zone is its fragmentation into a multitude of microhabitats created by relief, exposure, and slope, which interact with solar radiation and wind to cause soil moisture, temperature, and substrate quality to vary enormously over very short distances (Scherrer and Körner, 2010). It is this diversity of microhabitats, the steepness of the terrain in particular, that makes the alpine life zone so different from arctic tundra, and that is responsible for its much greater organismic richness. Across a few meters one can easily find snowbed plants, just emerging from melting snow, and succulent plants on outcrops, which perform crassulacean acid metabolism (CAM) as do many hot desert plants. In fact, these succulents may regularly experience temperatures close to 50 °C on steep south slopes under full midday sun, even at high latitudes.

In addition to these microclimatic determinants imposed by land surface structure, plants themselves influence their micro-environment (**Figure 2**). Depending on life-form, plants may decouple themselves from ambient air conditions. Under direct insolation, compact cushion plants or prostrate dwarf shrubs have been shown to warm up to tropical temperatures. These life-form-dependent microclimates largely disappear under thick clouds or at night, but storage of warmth in the topsoil is also strongly influenced by the type of plant cover, and so is radiative cooling during clear nights.

Hence, the climate that alpine plants experience can be very different from what might be expected based on elevation alone or data measured at weather stations. Such weather data monitor the temperature, humidity, and windspeed that occur outside the calm boundary layer into which alpine plants (leaves in particular) and their animal and microbe partners are commonly nested. A widespread assumption is that alpine plants are small and prostrate because their growth is temperature limited. The reality is that, because alpine plants are small (and mostly stay small when grown at higher temperatures), they may periodically escape the cold and, at least during sunny hours, experience radiative warming, and thus are not always colder than lowland plants. This is well reflected in their thermal optima for photosynthesis, which (in the temperate zone) were shown to be similar to those of lowland plants. In contrast, night-time temperatures may be prohibitive for plant growth at alpine elevations, and thus limit structural investments of assimilates acquired during the day. Because of the strong link between plant life-form and microclimate, alpine biodiversity can be understood only if one accounts for plant structural diversity.

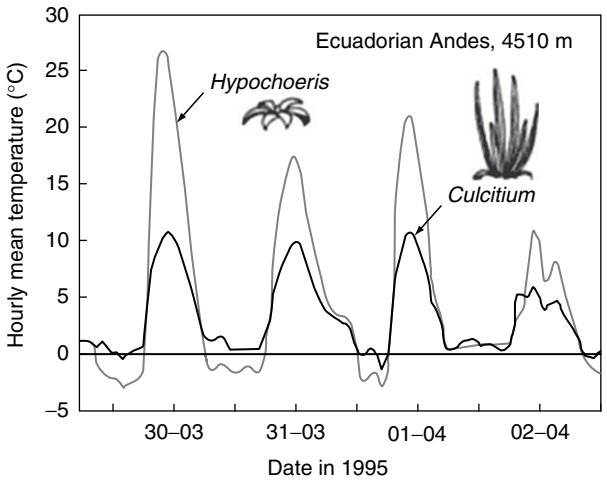


Figure 2 The diversity of plant morphologies creates diverse microclimates. Note the different temperatures in upright and prostrate leaves of two species in the Ecuadorian Andes. Reproduced from Diemer MD, in Körner Ch (2003) *Alpine Plant Life.* Heidelberg, Germany: Springer.

Diversity of Substrates

Alpine microhabitats may belong to a suite of different land surface structures and soils, the ten most important of which are common to all mountains:

- exposed rock terraces and rock crevices
- block fields scree and mixed scree/rock slopes or flats
- drained ridges and plateaus
- periodically wet depressions, gullies, or snowbeds
- gentle slopes with relatively stable soil
- steep slopes with creeping soil
- flats with cryogenic structure (e.g., hummocks, polygons)
- mires or other wet ground
- springs, water flows, or lakes.

These ten habitat types are not exhaustive. Special habitats not found everywhere include sand drifts or dunes and salt flats (in some semiarid subtropical mountains), disturbed surfaces due to animal trampling, rest places, or burrowing, avalanche tracks, and man-made landscapes such as pastures.

Each of these types of habitats may be found at different exposures to sun and wind (so the number of substrate types multiplies with the number of microclimatic conditions, despite some redundancy), and each of these combinations in turn may be found on very different parent rock material such as calcareous or siliceous, mixed metamorphic, or volcanic material. Depending on wetness and elevation, plants may have converted the top layers of the substrate to an extent that it becomes chemically independent from the parent rock (e.g., humic turf of pH 3.7 overlaying calcareous bedrock). Varying degrees of erosion and soil formation will then enhance the spatial heterogeneity of the substrate. The availability of soil nutrients and the extent of soil development are strongly relief driven, but plant life-forms also determine local nutrient retention. By forming dense cushions or tough tussocks, plant litter and thus the precious nutrients contained in it are prevented from being blown or washed away. When trapped beneath or strongly attached to the plant, litter can decompose and nutrients can recycle locally. Plant life-form, root and rhizome structure in particular, also influences substrate stability and soil formation (*see* Potential Functions of Alpine Biodiversity).

To understand alpine biodiversity, it is important to be aware of this extensive microenvironmental patchiness of the alpine life zone, and the capacity of plants to influence their

life conditions substantially. Taken together, the foregoing combinations of microenvironmental conditions yield hundreds of very specific niches, each preferred by a different combination of species.

Plant Diversity in the Alpine Life Zone

Diversity of Morpho-types

Though the alpine life zone is not as rich in life-forms as a humid tropical forest, the diversity of morpho-types found here is surprisingly high. There are ten principal groups of life-forms, eight of higher plants and two of cryptogams, irrespective of whether individuals perform clonal growth. The first four groups are most important:

- low stature or prostrate woody shrubs
- graminoids such as grasses and sedges, many forming tussocks
- herbaceous perennials, often forming rosettes, and
- cushion plants of various types.

Less common or of more regional importance are:

- giant rosettes of tropical mountains
- geophytes, mainly confined to mountains with a pronounced seasonality
- succulents, with both stem and leaf succulence, and
- annuals (sometimes biannuals), which become quite rare at high elevations.

The remaining two life-forms are cryptogams, that is, desiccation-tolerant, nonflowering plants:

- bryophytes ("mosses"), in some areas also ferns and lycopods, and
- lichens (including fruticose, foliose, and crustaceous).

These life-forms, in mixtures of varying abundance of each, compose the "alpine vegetation." In addition, algae and fungi play an important role. The diversity of plant structures is further enhanced by various modes of clonal propagation, which become increasingly important as elevation increases. The following is a short list of the diversity of clonal structures (**Figure 3**):

- tussock graminoids
- stoloniferous graminoids
- mat- or cushion-forming forbs
- stoloniferous or rhizomatous forbs
- creeping dwarf shrubs
- prostrate dwarf shrubs
- viviparous plants
- accidential clonal plants (fragmentation by external forces)

These typologies do not account for plant height or degree of horizontal spreading, both of which vary considerably among species and microclimates. Overall, the diversity of plant stature in the alpine life zone is perhaps nearly as large as the taxonomic diversity (see Halloy and Mark, 1996). It is obvious that these structural features are significant functional attributes, strongly associated with microhabitat preference and microhabitat tolerance.

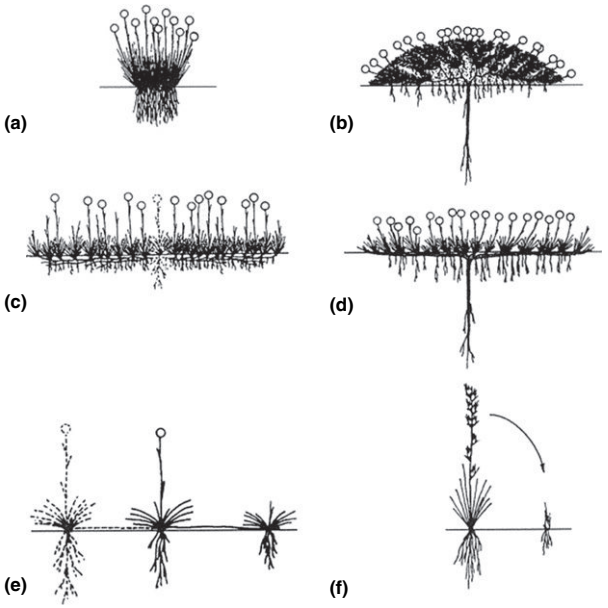


Figure 3 Diversity of clonal growth in alpine plants. The examples shown here include (a) graminoid tussocks, (b) cushions with adventitious roots and potential later fragmentation, (c) mats of rhizomatous forbs or graminoids (d) mats of proliferating forbs that retain a primary root (e) stoloniferous plants, and (f) "viviparous" plants producing bulbils or floral plantlets. Reproduced from Hartmann H, from Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

Diversity of Physio-types

The term physio-type is used here to circumscribe the physiological attributes of alpine plants, which may be as diverse as the variety of life-forms. Certain morphological attributes are functionally linked to physio-types. The following is only a brief summary. Readers with an interest in this field will find an extended treatment of the subject in Körner (2003).

Rates of plant photosynthesis, how photo-assimilates are invested in the plant body, nutrient use, water relations, stress resistance, and secondary metabolites all vary substantially among alpine species. The predominance of physical limitations to growth, particularly at the uppermost limits where plants can grow and survive, may be expected to narrow the spectrum of possibilities, in the ultimate case by permitting only one way to survive. Surprisingly this is not so. Even in habitats that by all standards can be rated as "extreme," one can find a suite of physio-types (often associated with specific morpho-types). This is a most important point for the study of alpine plant adaptation. The selection of a single species for study inevitably will produce data with no generalization potential. "The alpine physio-type" does not exist. One randomly selected species or small group of species may represent a very special case, and be all but "typical."

A good example is the way plants invest in biomass. Although this finds expression in morphology, the quantitative aspects of it are directly related to plant metabolism. Plants may favor roots, stems, storage organs (all three are net sinks for carbon), or leaves (the net source of carbon). How plants invest is key to the understanding of whole-plant carbon balance and

to growth and reproduction. Co-occurring species, equally successful in terms of abundance and often found together in the same habitat, may represent the left and right tails of a frequency distribution of these traits at a common attitude (**Figure 4**). There seems to be a multitude of ways to cope with the demands of life conditions even at extremely high elevations.

Another example is the way alpine plants construct their leaves in terms of leaf dry matter investment per unit leaf area. "Expensive" (thick) leaves contain a lot of dry matter per unit leaf area, whereas "cheap" (thin) leaves contain little. Again the diversity of this trait expressed as "leaf mass per area" (LMA, or its reciprocal specific leaf area, SLA) is very large. There is nothing like a typical alpine SLA (**Figure 5**; for a full account of such leaf traits in a global comparison see Körner *et al.*, 1989). The same applies to leaf nitrogen concentration and mobile carbohydrates, although there are slight overall trends with increasing elevation for SLA to decrease and nitrogen and mobile carbohydrate concentrations to increase (leaves becoming more "expensive"). However, such means across many species need to be treated with care, given the great diversity. It cannot be concluded that alpine plants have lower SLA when a substantial fraction of all studied alpine species does not fit this pattern. The species from different elevations that are included or excluded from such community subsamples will always affect the mean.

The same applies to any other physiological trait that has been studied in more than one species. For example, freezing resistance varies enormously in alpine plant species. According to Larcher, some species can survive any low winter temperature that could occur on Earth (less than $-70\,^{\circ}C$), and others are killed by only $-12\,^{\circ}C$. Freezing resistance is one of the few traits where this diversity of responses could be explained by habitat characteristics, in this case the predictability of snow cover in winter. But for most other traits such clear-cut causal links have not been found. It rather seems there are many different ways to cope with similar problems. These different traits may be "nonfunctional" in the sense of not being critical for survival and reproduction under "normal" situations, but some may become decisive under very specific and rare

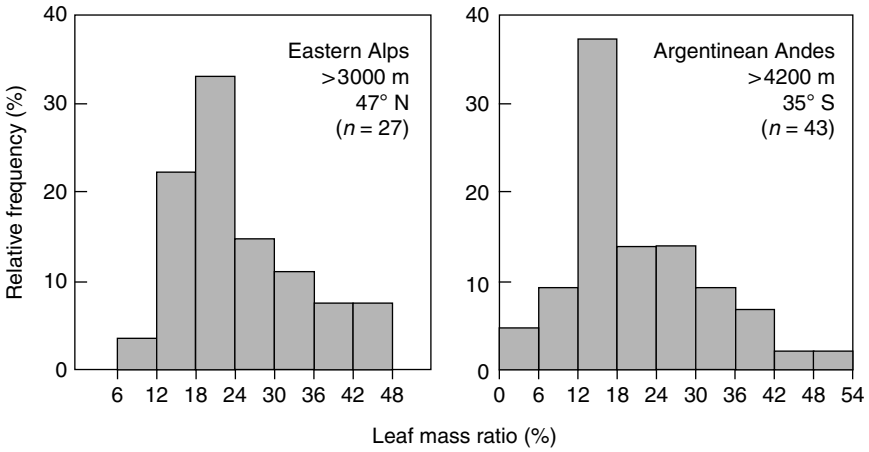


Figure 4 Frequency distribution of leaf mass ratio (LMR; the fraction of leaf dry matter versus total plant dry matter) in the Alps and the Andes. In the Alps, for instance, co-occurring species are found in the left (*Ranunculus glacialis*) and right (*Cerastium uniflorum*) tails, reflecting the great diversity of plant dry matter investment even at extremely high elevations. The same is true for any other mountain region, including the example for Argentina shown in the right diagram. Modified from Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

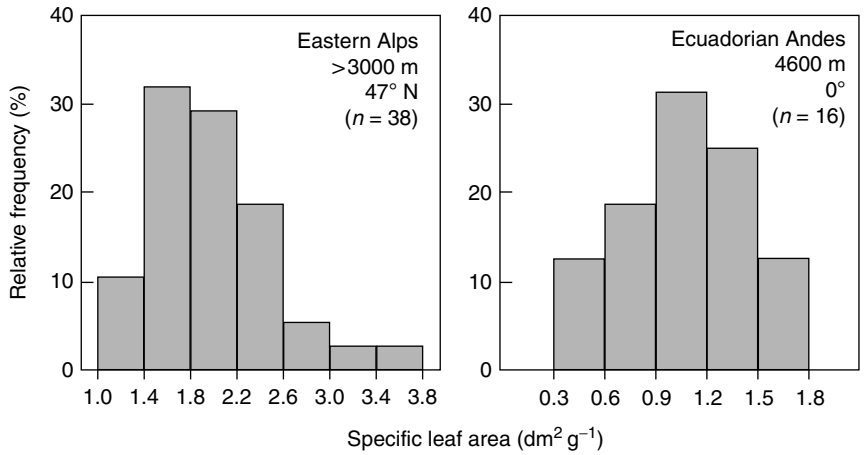


Figure 5 Frequency distribution of specific leaf area (SLA) in alpine species in the Alps and the Andes. Reproduced from Körner, *et al.*, 1989.

conditions through which they might have been selected for as advantageous (see the discussion of the ecological function of diversity *in* Potential Functions of Alpine Biodiversity).

Diversity of Reproduction

Recruitment *via* seed is rather risky in alpine environments because the seedling needs to establish in an often short season and on a potentially hostile substrate. Hence the most common strategies for alpine plants to survive and persist are long life and clonal propagation. The production of seed itself does not seem to be a problem and alpine plants produce lots of viable seed in most years, as has been known from the beginning of alpine plant research in the nineteenth century (see review by Körner, 2003). For a functional understanding of alpine biodiversity, seed production is very important, as rare as successful seedling establishment might be. It is through sexual reproduction that the genetic diversity is retained and by which the ecotypic differentiation of traits, essential in the fragmented and ever-changing alpine environment, develops. The most important pollinators are flies; at lower alpine elevations, bumblebees, solitary bees, and butterflies become more prominent (with decreasing importance by order; a detailed analysis for the southern Andes was done by Arroyo). Roughly one-fourth of alpine plant species are wind pollinated (such as Poaceae, Cyperaceae, and Polygonaceae).

Current evidence suggests that outbreeding is the dominant form of sexual reproduction, but most alpine plants can also self (which secures some seed production if pollen transfer fails because of bad weather), and some are even obligatory selfers. The most extreme forms of retaining the parental genome is apomixis, where embryos are genetic copies of the source plant. Surprisingly, both clones and apomicts exhibit high intraspecific genetic diversity, indicating that sexual reproduction also occurs in such plants. Through genetic fingerprinting (a recent review by Witte and Stöcklin, 2010), it was shown that some of these obligatory clonal alpine plants can produce very large genets and be thousands of years old, but clones of different genets were often found to be intermingled. Yet clonal propagation must not be seen as a substitute to sexual propagation. Both the sexual and clonal modes occur simultaneously for most of the time. There is no indication of genetic depauperation within species with increasing elevation, but this is a field that needs more research. Current knowledge does not suggest that alpine plant diversity is limited by reproductive constraints.

It is well established through classic transplantation experiments in California, the Rocky Mountains, and the Alps that ecotypic differentiation does occur among populations from different elevations, so that genotypes of the same species from high elevation differ in stature, phenology, and physiology from lower-elevation genotypes when grown in a common garden (see review by Clements *et al.*, 1950), Hence, in addition to species diversity, discussed next, there is also a high degree of within-species diversity.

Taxonomic Diversity

The total alpine flora of the globe consists of approximately 8000 to 10,000 species of flowering plants (Körner, 2003). These belong to about 100 (± 10) families and 2000 genera. Hence, one-fourth to one-third of all plant families of higher plants have representatives in the alpine life zone. Assuming a total known global flora of 250,000 species, alpine species contribute about 4% of the global plant species diversity. Given that only 2.6% of the global land area outside Antarctica falls in the alpine zone, global alpine plant diversity per unit land area is in fact higher than the global mean of all other biomes. Although such gross means need to be treated with great caution, one can at least conclude that plant species diversity in the alpine life zone is comparatively high, in view of the generally less luxurious life conditions found there.

Whether high mountains are biodiversity hot spots (Barthlott *et al.*, 1996) is a question of scale. The preceding numbers strictly refer to the alpine life zone as defined in the section The Alpine Life Zone. If one selects a census area that includes the full spectrum of elevations (e.g., a cross section through a whole mountain range), it may encompass almost all life zones on Earth from humid tropical forest at the bottom to glacier forefields at the top. High mountains, those in tropical latitudes in particular, indeed represent an incredible compression of biomes. Climatic and vegetation zones separated by several thousands of kilometers at sea level may be found within 50 km or less horizontal distance on the flanks of the major subtropical or tropical mountain systems. However, the topic of life zone compression in mountains exceeds the frame-work of this article, which is restricted to the discrete alpine part of mountain vegetation.

A single mountain system such as the Rocky Mountains, the Alps, the Caucasus, the Venezuelan part of the Andes, or the mountains of New Zealand commonly has an alpine flora consisting of 600 to 1500 species. This needs to be considered in view of a total flora of the arctic tundra of about 1000 to 1500 species (depending on how sub-species are ranked). In this respect, biodiversity of the alpine life zone is outstanding. A distinct mountain region such as the Teton Range in Wyoming, the Snowy Mountains of Australia, or the central part of the Swiss Alps commonly contains roughly 200–400 alpine species, a number that is surprisingly constant across the globe (most ranges harbor around 250 species). On a single sample plot (e.g., 100×100 m) one may find one-third of the total regional flora. An analysis of the Swiss Alps by Wohlgemut revealed that alpine plant species diversity increases with the size of the observation area up to 20 km^2, but beyond that species numbers level off.

Plant species diversity generally decreases with increasing elevation (**Figure 6a**), although there may be intermittent peaks where two altitudinal life zones merge. Again this is a question of scale. The decline of species diversity is particularly impressive in the uppermost part of the alpine life zone (**Figure 7**).

The diversity of cryptogams, mosses and lichens in particular, is relatively high in the alpine life zone. Species numbers may be of the same order of magnitude as for vascular plants, depending on humidity. With greater humidity one generally finds higher numbers of cryptogams. In some parts of temperate zone and subarctic mountains the biomass of fruticose lichens in alpine grassland is double that of higher plants. Since mosses and lichens are desiccation tolerant, they can occupy bare rock and scree and are among the first organisms to initiate humus accumulation on raw substrates.

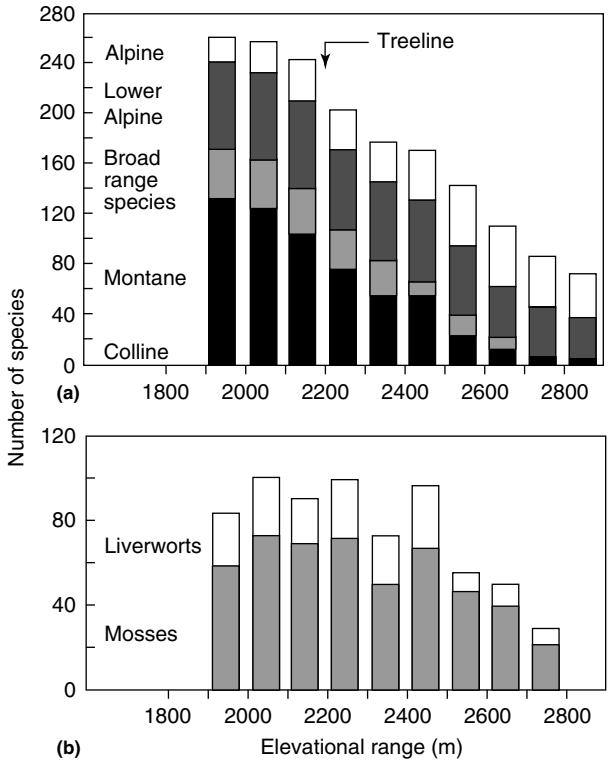


Figure 6 Altitudinal trends in species diversity of higher plants in the Alps ((a) after Theurillant JP and Schlüssel A) and bryophytes ((b) after Geissler P and Velluti C). Adapted from Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

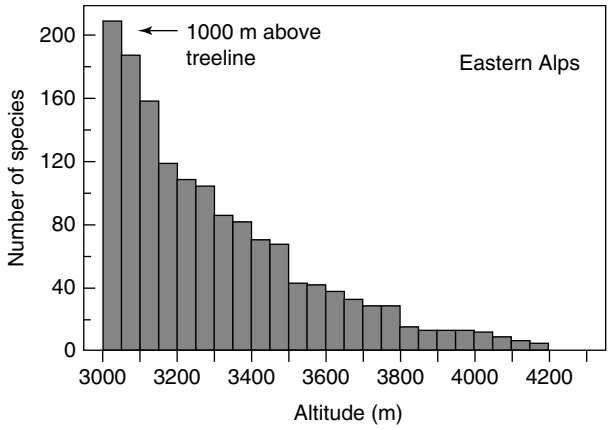


Figure 7 Altitudinal trends in plant species diversity in the uppermost range of higher plant life in the Alps. Reproduced from Grabherr G, from Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

Although cryptogams have almost no elevation limits as long as there is some bare substrate (rock lichens are reported at 7200 m elevation in the Himalayas), the number of cryptogam species also declines with greater elevation, as is shown in **Figure 6b** for bryophytes.

Besides the major role of microhabitat diversity discussed earlier, there are also historical reasons for alpine plant diversity. While the lowland flora changed owing to global climatic changes, such as after the ice ages, a fraction of the then cold-adapted lowland flora migrated to high elevations and enriched the existing older stock of alpine species. This is how the edelweiss (*Leontopodium alpinum*) became a recent addition to the alpine flora of the Alps when the ice retreated. Agakhanjanz and Breckle (1995) suggested that in some mountains (e.g., the Pamirs) tectonic uplift matched time constants of speciation and contributed to the autochthonous species richness.

In summary, the alpine life zone is rich in plant species that belong to a multitude of morphological, physiological, and reproductive types and that inhabit a broad spectrum of microhabitats, reflecting both historical and even geological events. Because alpine plants are small, species diversity may exceed 50 species per 1 m^2 of land area, which is among the highest in the world. Alpine plants also provide a varied diet of food for alpine animals and microbes, the diversity of which will be briefly touched on in the following section.

Diversity of Alpine Animals and Microorganisms

Animal Diversity

The great variety of animals from mites to birds, and the greater fragmentation of expertise in animal sciences, may explain why there has been no attempt at a global synthesis of animal diversity in the alpine life zone. Franz (1979), in his German book on high mountain ecology, referred to a great number of otherwise scattered observations, and Meyer and Thaler (1995) attempted a summary for invertebrates, with a focus on the Alps. The following brief account leans heavily on the latter publication.

The only resident vertebrates in the uppermost part of the alpine zone (often called the nival zone) are voles, with the record finding by Halloy at *ca.* 6000 m elevation at a geothermic vent on Sucompa volcano in the Argentinean Andes. Snow voles are very abundant and active even above 3200 m in the Alps. At these elevations birds are visitors, but only a few hundred meters lower they are resident as the plant cover becomes more regular. In the Argentinean Andes, ducks breed at 4250 m elevation (personal observation at an alpine lake in the Cumbres Calchaquies). Snakes, lizards, and frogs are reported to occur up to 1000 m above treeline. Large mammal grazers (e.g., ibex in the Alps and guanacos in the southern Andes) occur at almost any alpine elevation, except for peaks surrounded by glaciers or inaccessible cliffs. These grazers profoundly influence the development of alpine plant diversity. In the lower and middle alpine belts, wild herbivores have been replaced by domestic herbivores in many areas (sheep, goats, yaks), which continue to exert selective pressure on vegetation. Some of these animals graze the highest fragments of vegetation, as was reported for yaks, which were found grazing above 5000 m in the dry part of the Himalayas (Miehe, 1997).

Quantitative data for some important groups of invertebrates have been compiled by Meyer and Thaler (1995). The major losses of taxonomic groups as elevation approaches the upper limits of closed vegetation are those of earthworms, gastropods, grasshoppers, Hymenoptera, and beetles (**Figure 8**), with the latter showing the greatest decline

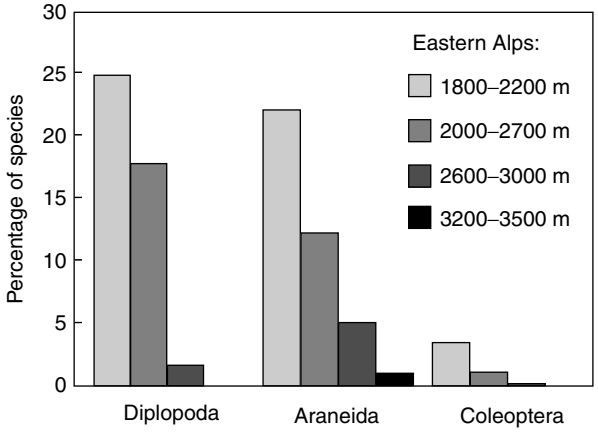


Figure 8 Examples of elevational trends in invertebrate species diversity in the Alps. Reproduced from Meyer E and Thaler K (1995) Animal diversity at high altitudes in the Austrian Central Alps. In: Chapin FS III and Körner Ch (eds.) *Arctic and Alpine Biodiversity: Patterns, Causes and Ecosystem Consequences*, pp. 97–108. Heidelberg, Germany: Springer.

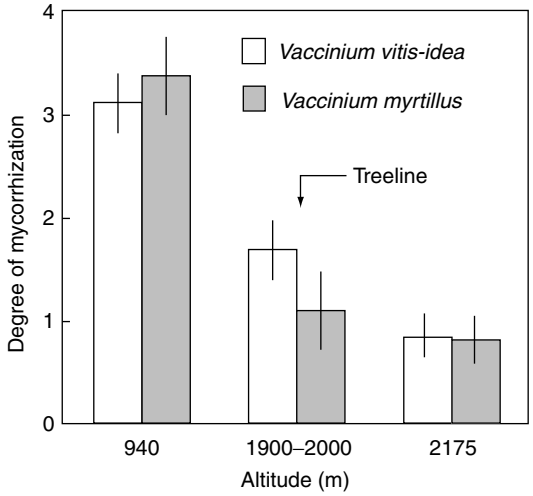


Figure 9 Elevational trends in mycorrhizae occurrence in the Alps. Reproduced from Haselwandter K and Read DL, from Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

already below treeline. By contrast, flies, spiders, and springtails remain quite abundant (at 300 m above treeline, one-fifth of the total regional number of 200 springtail species were found in the Tirolian Alps). According to Swan, a salticid spider species was collected in the Himalayas at an altitude above any plant growth, possibly living on the aerial import of small arthropods from lower elevations. Only 2% of the 400 species of spiders of the central Alps of Tirol (Austria) are regularly found in the uppermost alpine zone. The more open vegetation becomes, the more invertebrate life is linked to the occurrence of compact plant forms, such as cushion plants.

In their overview, Meyer and Thaler noted that among invertebrates the herbivore species numbers declined more rapidly than species numbers in higher trophic levels. The reduced species number at "extreme" habitats is often balanced by greater individual densities per species. Overall, animal species diversity in the alpine life zone follows similar elevational trends as in plant species diversity, but in terms of the total numbers of alpine species, animals may exceed plants by factors of five to 10 (this is a personal guess). Establishing such diversity ratios would be an important contribution to the understanding of alpine biodiversity in general.

Microbial Diversity

As the altitude above treeline increases, animals make a smaller contribution to plant litter decomposition and soil formation and microbes become more important, even as their species numbers decline steadily. Diversity of soil fungi decreases with altitude and mycorrhization, the important plant root–fungus symbiosis, also decreases (**Figure 9**). All known types of mycorrhizae occur in alpine soils: ectomycorrhizae (e.g., on Salix, Dry as, Polygonum, and Kobresia spp.), ericoid mycorrhizae (Ericaceae), vesicular–arbuscular (VA) mycorrhizae (most forbs, grasses, and some sedges), and even orchid mycorrhizae. Nonmycorrhizal plant species can also be found (Gardes and Dahlberg, 1996). So-called arbuscular mycorrhizae (Glomeromycota) and dark-septate hyphae type species (belonging to Ascomycota) are found even in the highest rock and scree habitats, though at greatly reduced abundance (for references see Körner, 2003). Yet, even completely isolated plants above 3000 m in the Alps were found with intense mycorrhization (reported by Oehl; see Körner, 2011).

The most robust of all organisms, the bacteria plus some unicellular fungi, retain a high diversity and abundance in the alpine life zone. Schinner and coworkers isolated 130 different strains of microorganisms in alpine environments of the western and eastern Alps that can survive and multiply at 0 °C. Of these, 77% were bacteria, 20% yeasts, and only 3% hyphal fungi. A very detailed analysis of bacterial diversity on Niwot Ridge, Colorado, by Mancinelli showed that Pseudomonas and Bacillus were the most abundant genera. Dinitrogen-fixing as well as nitrifying and denitrifying bacteria are abundant in alpine soils. Bacteria have no altitudinal limits as long as some organic dust and short spells with liquid water occur, and Swan reported a number of taxa isolated from substrate collected at 8400 m altitude on Mt. Everest, the environment on Earth that he thinks is most comparable with that of Mars.

Taken together, these observations indicate that mycorrhizae are a common element of alpine plant life, being more prominent in lower alpine elevations and in infertile soils and becoming rare only with isolated plants on high mountain peaks, where soils have little carbon. Rich bacterial life is found even at the highest elevations, indicating a capacity for metabolism under the most extreme conditions.

Potential Functions of Alpine Biodiversity

Among the theories that attempt to explain the functional significance of organismic diversity, the insurance theory seems to be most relevant for the alpine life zone. In simple terms it says that a species-rich, functionally partly redundant

organismic "work force" ensures the functional integrity of ecosystems even if some of the organisms die out. A complete loss of species by extreme stress or through pathogens would be nearly irreversible in steep alpine terrain, because it is the presence of at least some species that prevents the soil from being washed away. The sustained functional integrity of alpine ecosystems is inevitably tied to the presence of soil, and this presence is inevitably dependent on the roots and rhizomes that hold it. Strong-rooted plants are the keystone elements for the preservation of the alpine ecosystem in steep terrain. A high diversity of species is commonly associated with a high diversity of rooting patterns, which in combination create the mechanical strength required to hold the soil (Pohl *et al.*, 2009).

Messerli and Ives estimated that 10% of the global human population depends directly, and 40% of the population indirectly, on mountain ecosystems, and thus the stability of their upper part, the alpine zone, is of critical significance to society (Körner and Ohsawa, 2006). Supplies of drinking and irrigation water, as well as the safety of hydroelectric schemes and transport routes, depend on intact upslope soil conditions, and alpine biodiversity helps to ensure ecosystem health. Of course alpine biodiversity also provides other things. Alpine meadows and fellfields are very attractive landscapes for human recreation, and with their high biodiversity they are of prime conservation value; in many regions they represent the last undisturbed natural areas. Because the alpine life zone is represented across the globe, it is also ideally suited for global monitoring of biological responses to atmospheric change.

Alpine Biodiversity and Global Change

Global change has many facets, all related to the ever-increasing use of resources and land area because of human population growth and the increasing consumption of goods. The most severe impacts on alpine ecosystems are (1) land use practices (2) potential global warming with associated changes of snow cover and permafrost, and (3) increasing wet nitrogen deposition. Other aspects of global change, such as atmospheric CO_2 enrichment by itself, increasing ultraviolet radiation due to ozone layer depletion, and air pollutants other than nitrogen loading, seem to be of minor significance on a global scale (see review by Körner, 2003).

Land use, in particular the intensification of pasturing or the reverse, the abrupt abandonment of former, traditionally pastured alpine terrain, exerts the greatest influence. The steeper the terrain, the more critical these effects become. Overgrazing generally diminishes plant diversity and ruins the protective plant cover within a few seasons, with self-repair of the ecosystem commonly occurring at a much slower rate than soil erosion. In contrast, moderate and well-managed grazing can increase biodiversity and create short and dense swards of vegetation that are extremely robust against erosion and increase catchment water yield. Some of the biodiversity hot spots of the Alps, the Caucasus, and the Himalayas are traditional pastures in the lower alpine belt. There is no way of maintaining these systems by "let alone" grazing strategies, because herds of domestic animals tend to crowd certain areas and ignore others. Maintaining these intact alpine grasslands

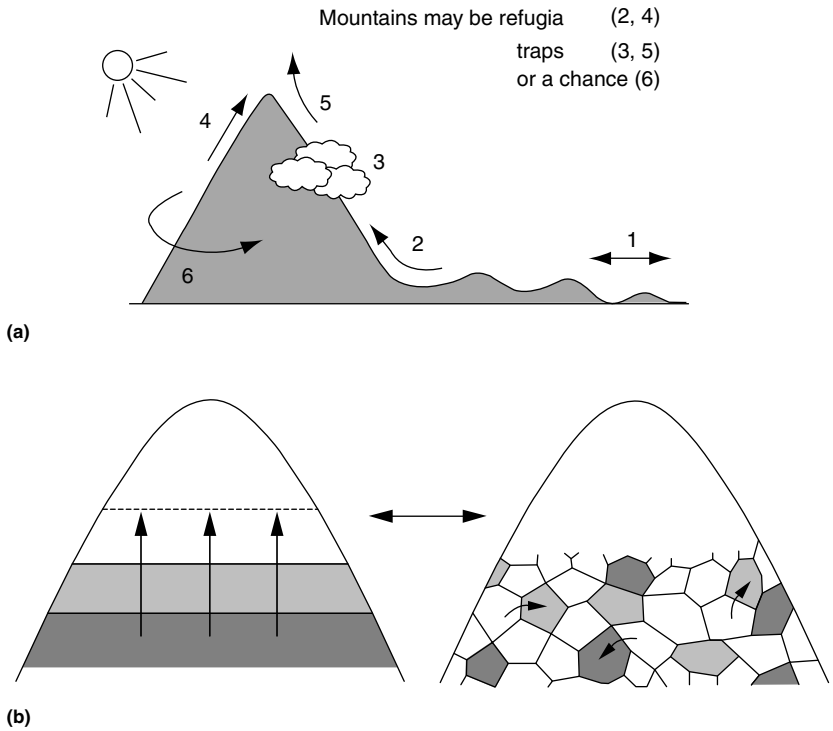


Figure 10 (a) Species responses to climatic warming. Mountains may be refugia (2, 4), traps leading to local exinction (3, 5), or a chance to escape climate warming by topography effects (6). Lowland species often have to move greater distances (1). (b) Upslope migration of species or niche filling? Both responses are likely with climatic warming and will alter local species abundance. Niche filling, that is, "horizontal reallocation," will precede longer-term changes of the position of whole vegetation belts.

is very important for at least three reasons: (1) their often very impressive biological richness, (2) their continued potential as a "clean" food source, and (3) their positive influence on water yield. The abrupt abandonment of pastures leads to an unstable transition phase that may last for a century before new, well-adapted communities of species are able to return. A later reversal back to the previous biodiverse pasture-land is often impossible within a reasonable time and with affordable effort, because the biological structure and the soils of these pastures took many hundreds or thousands of years to develop, but are rapidly converted.

Climate warming will affect alpine biodiversity in subtle ways because of the great variety of intermingled microhabitats (*see* Habitat Diversity, a Key to Alpine Plant Diversity). Although a number of authors documented some upslope migration of species, the more important changes possibly happen among microhabitats, with new niches filled by species and other niches abandoned (**Figure 10**; Scherrer and Körner, 2010). These mosaics of life conditions represent a certain margin of safety against the loss of alpine species in a slightly warmer overall climate. However, the abundance of species will change as the abundance of their microhabitats changes. Effects on snow cover and snow duration will be more critical than temperature *per se*. Exceptions to these scenarios are mountains that are not high enough, in which case the current alpine biota will find no upslope escape if it gets warmer.

Since alpine vegetation is well adjusted to cope with low soil nutrient levels, the regional increase of soluble nitrogen deposition will influence biodiversity. More vigorous and nitrogen-demanding species are likely to gain space over slower-growing, smaller species. Since these more vigorous species are commonly also less resistant to stress, nitrogen deposition can increase the sensitivity of certain ecosystems, while others (e.g., pioneer vegetation) may profit. It was shown that even minute additions of nitrogen fertilizer – less than is contained in many places in lowland rain-water – can create drastic changes in the alpine flora (Körner, 2003).

In summary, the greatest risk of loss of biological diversity in the alpine life zone is human land use. However, land use can also contribute to the maintenance of highly diverse and stable ecosystems in the lower alpine belt if sustainable management practices are applied.

See also: Antarctic Ecosystems. Europe, Ecosystems of. Grazing, Effects of

References

Agakhanjanz O and Breckle SW (1995) Origin and evolution of the mountain flora in middle Asia and neighboring mountain regions. In: Chapin FS III and Körner Ch (eds.) *Arctic and Alpine Biodiversity: Patterns, Causes and Ecosystem Consequences*, pp. 63–80. Heidelberg, Germany: Springer.

Barthlott W, Lauer W, and Placke A (1996) Global distribution of species diversity in vascular plants: Towards a world map of phytodiversity. *Erdkunde* 50: 317–327.

Clements FE, Martin EV, and Long FL (1950) *Adaptation and Origin in the Plant World. The Role of Environment in Evolution*. Waltham, MA: The Chronica Botanica Company.

Franz H (1979) *Ökologie der Hochgebirge*. Stuttgart: Verlag Eugen Ulmer.

Gardes M and Dahlberg A (1996) Mycorrhizal diversity in arctic and alpine tundra: An open question. *New Phytologist* 133: 147–157.

Halloy SRP and Mark AF (1996) Comparative leaf morphology spectra of plant communities in New Zealand, the Andes and the European Alps. *Journal of the Royal Society of New Zealand* 26: 41–78.

Körner C and Ohsawa M (2006) Mountain systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystem and Human Well-being: Current State and Trends*, vol. 1, Millennium Ecosystem Assessment, pp. 681–716. Washington: Island Press.

Körner Ch (2003) *Alpine Plant Life*. Heidelberg, Germany: Springer.

Körner Ch (2011) The coldest places on earth for higher plant life. *Alpine Botany* 121: 11–22.

Körner Ch (2012) *Alpine Treelines*. Basel: Springer.

Körner Ch, Neumayer M, Pelaez Menendez-Riedl S, and Smeets-Scheel A (1989) Functional morphology of mountain plants. *Flora* 182: 353–383.

Körner Ch, Paulsen J, and Spehn EM (2011) A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data. *Alpine Botany* 121: 73–78.

Meyer E and Thaler K (1995) Animal diversity at high altitudes in the Austrian Central Alps. In: Chapin FS III and Körner Ch (eds.) *Arctic and Alpine Biodiversity: Patterns, Causes and Ecosystem Consequences*, pp. 97–108. Heidelberg, Germany: Springer.

Miehe G (1997) Alpine vegetation types of the central Himalaya. In: Wielgolaski FE (ed.) *Polar and Alpine Tundra: Ecosystems of the World 3*, pp. 161–184. Amsterdam: The Netherlands: Elsevier.

Pohl M, Alig D, Körner Ch, and Rixen Ch (2009) Higher plant diversity enhances soil stability in disturbed alpine ecosystems. *Plant and Soil* 324: 91–102.

Scherrer D and Körner Ch (2010) Topography-controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. *Journal of Biogeography* 38: 406–416.

de Witte LC and Stöcklin J (2010) Longevity of clonal plants: Why it matters and how to measure it. *Annals of Botany* 106: 859–870.

Amazon Ecosystems

Ghillean T Prance, Royal Botanic Gardens, London, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Caatinga Within the Amazon region this is applied to open forest on white sand, which occurs in the basin of the Rio Negro. The term is also applied to the semidesert region of northeast Brazil; it is an indigenous word from the Tupi language meaning an open place.

Capoeira Brazilian term for secondary forest on cleared ground.

Cerrado The Brazilian term for the large area of savanna and savanna forests that are the dominant vegetation of the Planalto of Central Brazil.

Igapó Vegetation periodically flooded by acidic black water or clear water rivers.

llanos The term used in Colombia and Venezuela for the extensive grass-dominated, open savannas of the orinoco river region.

Inselberg Granitic domes that rise above the forest in the older geological formations to the north and south of the alluvial plains of Amazonia and elsewhere.

Rivers in Amazonia The three main types of rivers in the Amazon region are known as white, black, and clear water. White water rivers are muddy with much suspended sediment and are neutral or only slightly acidic. Black water rivers are dark because of dissolved tannic matter and are acidic (pH about 4). Clear water has neither mud nor humic matter and is usually slightly acidic.

Tepui Venezuelan term for the sandstone table mountains of the Guayana Highland or Lost World region.

Terra firme Brazilian term for areas that are above the level of periodic inundation by the rivers.

Várzea Brazilian term for vegetation periodically flooded by white water rivers.

Most people think of the Amazon region as being covered by a uniform green carpet of rain forest. This is far from reality, because Amazonia is a complex mosaic of ecosystems varying from typical tall rain forest to open grassland savannas to scrubby vegetation on white sand that resembles the chaparral of western California. Even when one type of vegetation is studied, many local variations occur depending on local edaphic conditions, the flooding regime, or rainfall. This great variety of ecosystems has been defined in different ways by different authors depending on their approach to the subject. For example, some authors have concentrated on the physiognomy and structure of the vegetation (Hueck, 1966; Ducke and Black, 1953; Prance, 1987) and others, such as Holdridge (1972), on the life zones that define vegetation based on climatic factors such as rainfall, temperature, and altitude. In this account, the primary separation of ecosystems is based on floristics, that is, the individual species that make up an ecosystem, and on their physiognomy and structure.

In addition to the previously mentioned factors that account for the variety of ecosystems, the history of the region must also be taken into consideration. Over time the Amazon ecosystems have by no means been stable. The distribution of the different types of vegetation has fluctuated with changes in worldwide climate. The distribution and species composition of forest and nonforest biomes changed continuously during the Pleistocene and earlier epochs. What today is continuous forest was once broken up into isolated blocks interspersed with deciduous forest and savanna ecosystems that later coalesced when the climate again became favorable for forest. These changes have added to the overall complexity of the current Amazon ecosystems. The principal vegetation formations are described here.

Nonflooded Terra Firme Ecosystems

The primary division of Amazonian ecosystems is between those that are subject to periodic flooding by the seasonal rise and fall of river water levels and those that are above the flood level. In Brazil the nonflooded land is known as terra firme, and since this term is widely applied in the literature about Amazonian vegetation it is also used here. Within the Amazon region there are forests and savannas on both terra firme and floodplain.

Rain Forest on Terra Firme

This is the single most widespread vegetation type of Amazonia and covers approximately 50% of the region. Rain forest occurs where there is heavy rainfall, usually greater than 2000 mm y^{-1} that is not markedly seasonal. Some of the other types of vegetation described here occur in areas with a marked dry season or with less rainfall, for example, on the southern fringes of Amazonia where transition forest and semideciduous forest occur.

The rain forest on terra firme is characterized by its dense closed canopy at about 25–35 m above the forest floor. Above the canopy a number of tall emergent tree species rise up to about 50 m. The dense canopy means that little light (about 3%) penetrates to the forest floor and the herbs and shrubs growing there are adapted to living in low light. The periodic sunflecks are of considerable importance to those species and to the seedlings of trees. Tree species range from light demanders that grow quickly where light is abundant, but that survive poorly in shade, to shade-tolerant species whose seedlings can survive and grow in low light (Pires and Prance, 1984).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00243-4

Rain forest on terra firme is most notable for its amazing diversity of tree species. There have now been numerous inventories of trees in different parts of Amazonia. The highest number of trees of 10 cm diameter or more was recorded in western Amazonia at Cuyabeno in Ecuador, where Valencia *et al.* (1994) found 307 tree species on a single hectare. Inventories have listed from 81 to 307 species depending on the soil, rainfall, and the seasonality of the climate (**Table 1**). When the total number of woody species in a hectare is counted, there can be as many as 550 species.

A summary of the α-diversity of the terra firme forest is given in ter Steege *et al.* (2003) based on 275 one-hectare plots from all over the Amazon region. These authors found that dry season length, while only weakly correlated with average tree α-diversity, is a strong predictor of tree density and of maximum tree α-diversity.

The floristic composition is not uniform and some species have very local distributions. As a result, the lowland forest of Amazonia has been divided into a number of distinct phytogeographic regions based on the distribution of species. The first important subdivision of the region was offered by Ducke and Black (1953): this was later modified slightly by Prance (1977). The seven major phytogeographic regions can be explained by the history of the region and by the present-day climate. For example, the Atlantic coastal region has a wet almost aseasonal climate, but farther inland around Santarém, Brazil, there is a region with a much drier climate and a strong dry season. Farther to the west near the Andes in Peru, the climate is much wetter and less seasonal. Each of these regions falls into different phytogeographic zones (see Daly and Prance, 1989, for a review).

Figure 1 shows the phytogeographic regions with the distributions of some plant species superimposed. While some species have extremely restricted local ranges and help to define the different regions, others are widespread and occur throughout the entire region, for example, *Caryocar glabrum*,

Table 1 Results of some 1-ha inventory plots in the Amazonian rain forest where trees of 10 cm diameter or more were counted

Site	*No. of species*	*No. of trees*	*Habitat*
Cuyabeno, Ecuador	307	693	Aseasonal rain forest
Yanomono, Peru	283 (+17 lianas)	583	Aseasonal rain forest
Cocha Cashu, Peru	189	650	Seasonal rain forest
Xingu River, Brazil	162	567	Seasonal rain forest
San Carlos de Rio Negro	83	744	Aseasonal forest on poor oxisol
Beni, Bolivia	94	649	Seasonal forest

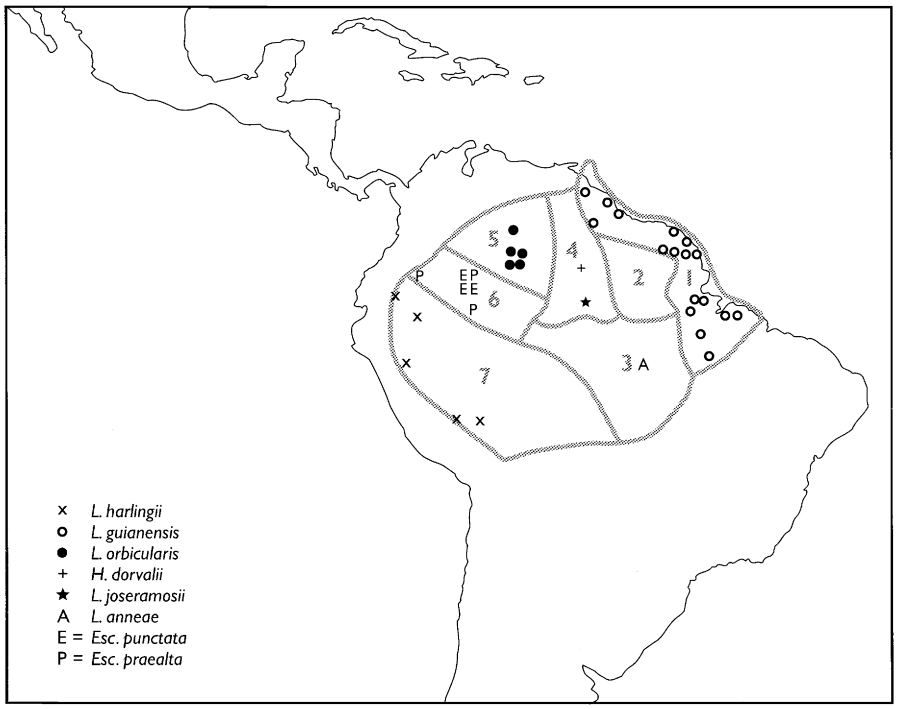


Figure 1 Map showing the phytogeographic divisions of the Amazon rain forest as defined by Prance (1977). The symbols represent the distributions of some species of trees typical to each region. L = *Licania* (Chrysobalanaceae), H. = *Hirtella* (Chrysobalanaceae), Esc = *Eschweilera* (*Lecythidaceae*).

Licania heteromorpha, and *Parinari excelsa*. The distribution of individual species is also dependent on altitude, soil, drainage, and topography, as well as climatic factors. There is not much structural or physiognomic difference in terra firme forest between the different phytogeographic regions, but the canopy height can vary considerably from one place to another.

Between 1972 and 1983, the RADAMBRASIL project surveyed the entire Brazilian Amazon using side-scanning radar. It distinguished two main types of forest on terra firme based on topography: forest in the lowlands on relatively flat ground and hill forest on undulating land between 250 and 700 m. The dense lowland forest occurs mainly on the area of Tertiary and Quaternary sediments between the major rivers on land below 250 m. This forest usually has a canopy at about 30 m and frequent emergent trees up to 50 m, for example, the Brazil nut (*Bertholletia excelsa*) or the angelim pedra (*Dinizia excelsa*), the largest tree in the forest. The understory has relatively few shrubs and consists mainly of young individuals of canopy tree species. Above 250 m the lowland forest tends to be replaced by hill forest, which has a lower canopy and its emergents reach to about 35 m tall. The canopy is more open and consequently the understory is often much thicker and has a greater diversity of shrubs and herbs. The physiognomy of hill forest is far more varied than that in lowland forest, and there are often marked differences between slopes and valleys in both physiognomy and species composition, adding to the total species diversity of the region. Much of the variation in composition of terra firme forests is related to edaphic conditions. A study across a 750 km transect of terra firme forest in western Amazonia indicated two sharp discontinuities in community structure at the genus level (Pitman *et al.*, 2008). This variation is congruent with a change in soil texture due to a subterranean paleoarch. A second discontinuity is noted between clay and white sand soils.

Most trees of the terra firme forest have large heavy fruit or seeds that drop to the ground. After this initial dispersal by gravity they are often carried farther by secondary dispersal agents such as rodents and other terrestrial animals. In contrast to savanna and riverine vegetation, most terra firme forest species are not adapted to long-distance dispersal. This must affect the rate of expansion of the forest into potentially colonizable habitats following climate change or human-caused forest destruction. The various species of primates in the canopy are important dispersers of seeds as well as seed predators.

White Sand Oligotrophic Formations

The edaphic factor that has the most striking effect on vegetation on terra firme is white sand soil. This habitat is scattered throughout the region, but especially between the Rio Branco and the Rio Negro in Brazil (**Figure 2**). These areas are characterized as podzols, or leached pure white sand, known as regosols in Brazilian soil terminology. The origin of these sands is either from the erosion of sandstone mountains in the northern part of Amazonia or from uplifted former river or sea beaches throughout the region. This soil is nutrient deficient and, as a consequence, has a distinctive lower vegetation that varies from an open savanna-like formation to low forest. In the Old World tropics this is called heath forest, but a variety of local terms have been applied in the Neotropics, such as Amazonian caatinga, campina, campinarana, chavascal, and charravascal. The commonest term for campina forest in Amazonia is caatinga, but this leads to confusion because the same term is used for the semidesert thorn-scrub of northeastern Brazil.

The largest area of campina forest in the upper Rio Negro region has a canopy at about 18 m. The trees tend to be of a much smaller diameter than in terra firme forest and the branching is tortuous and does not form a closed canopy. As a result, the ground vegetation is abundant and rich and it contains many endemic species, especially in the families Bromeliaceae, Marantaceae, and Rapateaceae. There are also abundant epiphytic Orchidaceae, Araceae, and pteridophytes on the tortuous branches of the trees.

In addition to the almost continuous area of campina forest in the upper Rio Negro, there are many small isolated patches of campina especially in the lower Rio Negro around Manaus. These form small islands of campina surrounded by forest on terra firme. Some of the species that dominate these campina islands are *Aldina heterophylla*, *Humiria balsamifera*, and *Pradosia schomburgkiana*. These campina islands have trees with particularly tortuous branches that are heavily loaded with epiphytes. Besides the epiphyte families mentioned earlier, Gesneriaceae and *Peperomia* in the Piperaceae are common and they support ant gardens where ants nest among the roots and form a rich layer of detritus and disperse the seeds. In some campinas there are open areas where bare sand occurs that is often covered by the blue-green alga *Stigonema tomentosum*, as well as small islands of shrubs and low trees with abundant lichens around them. There is good evidence from the presence of pottery shards that these open areas were cleared by indigenous peoples about 800 years ago. The recolonization by vegetation has been extremely slow because of the lack of nutrients in the soil, and so many open patches still remain today. Often associated with areas of white sand are areas of black soils (*terra preta*) of anthropogenic origin (see Glaser and Woods, 2004). The vegetation on these soils tends to be much richer (Paz-Rivera and Putz, 2009). The black soils are being overexploited for agriculture and use in horticulture.

A study of a campina island near Manaus showed that three-fourths of the species of trees and shrubs are adapted to long-range dispersal by birds, bats, or wind. Therefore, campina species tend to have a much greater capacity for long-distance dispersal than species of the terra firme forest.

White sand areas occur in many other places throughout Amazonia. One large area is that of Serra do Cachimbo on the border between the Brazilian states of Pará and Mato Grosso. This area has a similar physiognomy to the campinas of the Rio Negro, but the species composition differs considerably and there are a number of species that are endemic to the sands of Cachimbo. To the west of this region there are white sand areas in Amazonian Peru and Colombia and in the Brazilian state of Acre. In the Guianas, white sand areas are abundant and are covered by forest or by open savanna. Dense forests on white sand in Guyana are dominated by the legume species *Eperua falcata* and two other species of *Eperua* and are termed wallaba forest. To the south of Amazonia there are

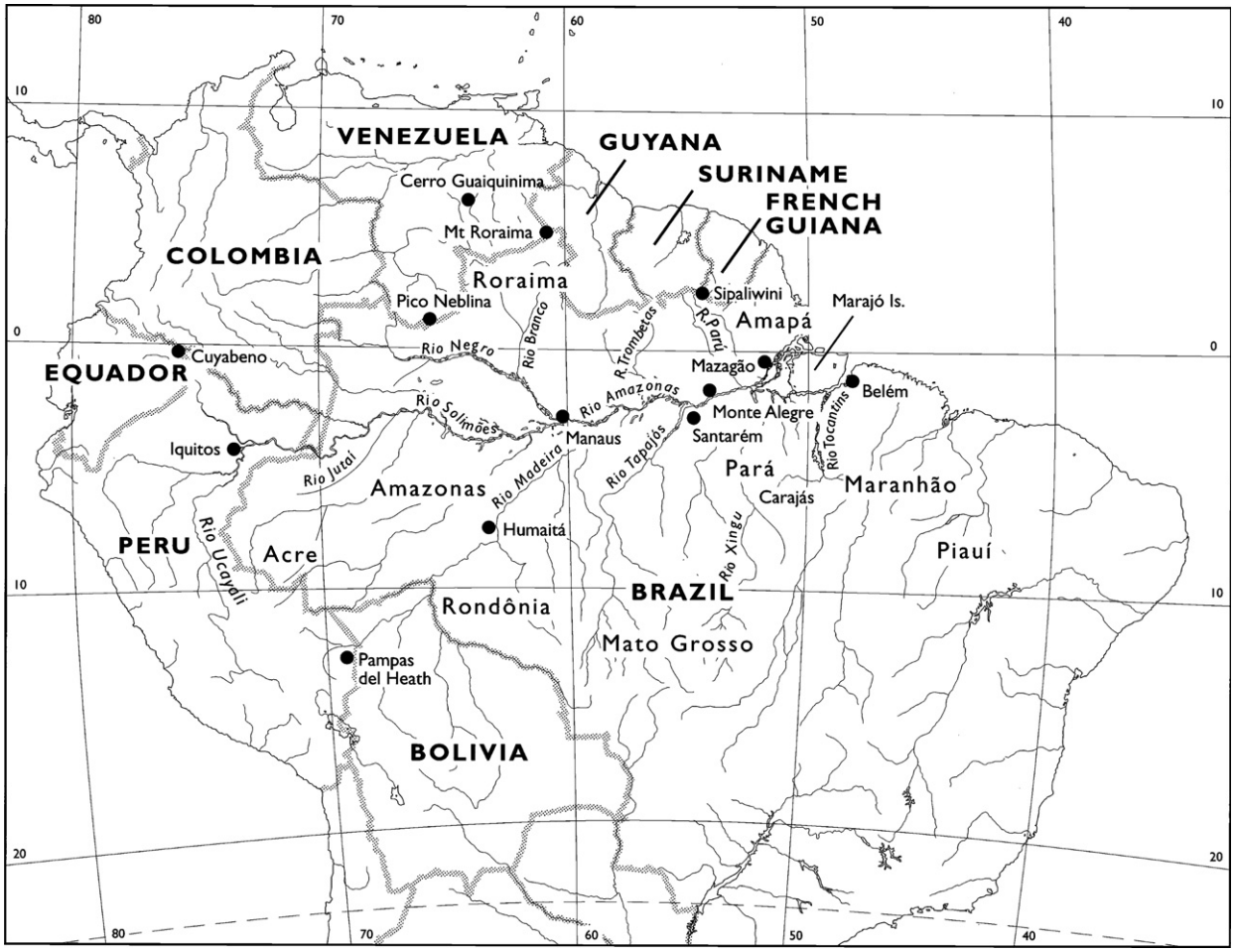


Figure 2 Map of northern South America, showing principal cities and rivers and localities mentioned in text.

extensive areas of white sand vegetation in the Chapada de Parecís in Rondônia, where many of the same genera occur as in the Rio Negro region, for example, *Abolboda*, *Clusia*, *Humiria*, *Paepalanthus*, *Retyniphyllum*, *Ternstroemia*, and *Xyris*.

The other white sand formation that occurs in Amazonia is called restinga. This low scrub habitat along the seacoast of Brazil is found on sand dunes beyond the influence of salt water and inland from beaches or mangrove forest. Within Amazonia, restinga is not extensive but it extends southward around the coast to southern Brazil. Restinga is a rather uniform scrub characterized by the presence of *Chrysobalanus icaco*, a species that also occurs in similar habitats in West Africa. Common herbs include the ubiquitous *Ipomoea pes-caprae* and the sedge *Bulbostylis capillaris*; shrubs include *Hibiscus tiliaceus* and *Manilkara triflora*.

Dry Semideciduous Southern Fringe Forests and Cerradão

This formation occurs predominantly in southeastern Amazonia on the border with the cerrado (the savanna) of the Central Brazilian Plateau. This type of dry forest also extends into the cerrado region, where it is termed cerradão. In this region the climate is much more seasonal, drier, and less humid. As a result, the trees lose some of their leaves in the dry season to make the forest semideciduous in marked contrast to the evergreen terra firme forest. Dry forests occur in small patches and do not occupy an extensive area. A small amount of dry forest also occurs in northern Amazonia in Roraima state and in southern Guyana. The dry forests really form part of the transition zone between Amazonian high forests and regions of savanna and are interspersed with the various types of transition ecosystems described in the following section. Ivanuskas *et al.* (2008) suggested that the transition forests of the upper Xingu river region be given a distinct phytogeographic recognition as "Evergreen Seasonal Forest." This was markedly distinct in their analysis based on physiognomy, floristic composition as well as edaphic factors.

The semideciduous forest is not rich in endemic species. Some of the common species that are widespread and characteristic of this formation include *Bowdichia virgiloides*, *Centrostigma macrophyllum*, *Combretum leprosum*, *Erythrina ulei*, *Lafoensia pacari*, *Magonia glabrescens*, *Attalea speciosa*, *Physocalymma scaberrimum*, and *Vochysia haenkiana*.

A study of an area of this type of forest in southern Amazonia at Sarare in Mato Grosso found that it contained many species typical of Amazonian terra firme, such as *Enterolobium schomburgkii*, *Hernandia guianensis*, *Parkia* spp.,

Schizolobium amazonicum, *Simaruba amara*, and *Spondias lutea*, mixed with others more characteristic of the forests of São Paulo, such as *Aspidosperma*, *Myroxylon*, and *Poeppigia procera*. There are also a number of species in this belt of transition forest that do not occur in central Amazonia within the dense forest region. The best known of these species is mahogany, *Swietenia macrophylla*, which occurs around the western and southern fringes of Amazonia. The genus *Myroxylon* is another example of this type of distribution. The other region that contributes to the semideciduous ecosystem is the cerrado region of central Brazil. Species such as *Apuleia leiocarpa*, *Er. ulei*, *Plathymenia foliosa*, and *Phys. scaberrimum* are typical cerrado species that extend their ranges into the semideciduous forest.

Transition Ecosystems

The transition forests form a belt of vegetation between the terra firme forests of Amazonia and the savannas of Roraima and Rupununi to the north and the cerrado of central Brazil to the south.

Transition forests mostly occur in regions with a strongly seasonal climate, but the boundary between them and the tall rain forest is not a clear-cut line. Rain forests extend farther into areas with seasonal climates that are on soils where the periodic water stress is compensated for by good water retention. This can occur because of either the physical properties of the soil or the proximity of rivers. Conversely, seasonal forests can extend into areas with little dry season in places that easily become dry, such as ridge crests and coarsely textured soils, and also on oligotrophic soils. Thus the superimposition of a climate map on that of vegetation will not necessarily correspond well and the forest boundaries are reticulate. The rain forest to seasonal forest ecotone still remains one of the least-studied ecosystems of Amazonia, yet it is in this boundary where most deforestation is occurring because of its accessibility to the developed south of Brazil and because the soil there tends to be more fertile than under evergreen forests. It is also easier to farm in a seasonal region when there is a time to burn the cut forest and to dry crops. In addition to the semideciduous forest, there are various, more specialized formations that tend to be dominated by one or a few species only. The three principal types of transition forest are described in the following.

Babassu Palm Forest

Babassu is a species of palm, *At. speciosa* that dominates the transition forests of the southeastern fringe of Amazonia in the Brazilian states of Maranhão, Tocantins, and Pará. In a babassu palm forest, the trees are widely spaced and do not form a closed canopy. The tree height is from 10 to 25 m and many pure stands of *Attalea* occur. As a result, species diversity is extremely low and there are few endemics. *At. speciosa* is a fire-resistant species and this is probably the reason for its abundance in a transition region to savanna. The fires that occur naturally in the cerrado probably help to reduce the competition from other species and allow babassu to dominate. The extent of babassu forest has probably been increased by deforestation both historically by indigenous peoples and more recently by clearance for farming. Babassu occurs as a component of semideciduous forest, and when the latter is cleared, it is often one of the few species that survives the burning, hence the recent increase in area covered by babassu forest. Babassu is also an extremely useful species and oil is extracted commercially from the kernel of the fruit. The woody endocarp is used to make charcoal or as a fuel for factories, and the outer mesocarp is used as an edible flour.

Liana Forest

As the name implies, in a liana forest woody vines or lianas predominate. It is open forest with well-spaced trees that are often completely entwined by lianas. In Brazil this formation is known as cipoal or mata de cipó (cipó being the Brazilian word for liana). The tree species in liana forest are some of the commoner species of terra firme forest, but they occur at a much lower diversity. Babassu palm also occurs frequently in liana forest, but there is a far greater quantity of lianas here, particularly in the plant families Bignoniaceae, Dilleniaceae, Leguminosae, Malpighiaceae, and Menispermaceae. To walk through this forest, one has to cut one's way through a thicket of tangled liana stems.

Liana forest is especially abundant in the area between the Rios Tapajós Xingu, and Tocantins in southern Pará, Brazil, but it also occurs in small patches in Roraima state in the north. The exact reason for the occurrence of this type of forest has not yet been established. One study showed that the soil under liana forest did not differ from that of nearby terra firme forest. It is thought to occur in regions where the forest has been disturbed by either human intervention or natural climate cycles.

Some of the typical Amazonian tree species that occur in liana forest include *Astronium gracile*, *Astronium lecointei*, the Brazil nut (*Be. excelsa*), *Elizabetha paraensis*, and *Sapium marmieri*.

Bamboo Forests of Western Amazonia

Bamboo forest is an open type of transition forest that occurs mainly in southwestern Amazonia in Amazonian Peru, Bolivia, and the Brazilian state of Acre. In Acre, bamboo forest is abundant and covers a large area. In this forest various species of bamboo (*Bambusa* subgenus *Guadua* and the genus *Merostachys*) are extremely frequent in both the understory and the canopy, where they reach up to 30 m. Bamboo forest contains relatively few tree species and those that do occur are mostly common species of the terra firme forest. The bamboos are interspersed with trees, which lend support to their arching branches. By branching the bamboos spread far over the forest canopy and can easily be picked out in aerial photographs. Botanist Thomas Soderstom measured a bamboo culm of 29.77 m in length with a side branch of 9.69 m and a secondary branch that was a further 4.96 m. The belt of transition forest that extends across the southern fringe of Amazonia is dominated by bamboos to the west, lianas in the central part, and babassu palms to the east.

Floodplain Ecosystems

Almost 10% of the Amazon region is subject to periodic flooding due to the rise and fall of river levels in most of the region and backup of the rivers from high tides in the delta region. River levels are very seasonal and the magnitude of change is considerable. In central Amazonia around Manaus, the Rio Negro and Rio Solimões undergo an average annual change in water level of 10 m, and the difference between extremely high flood years and the extreme low level is 15 m. The record high level at Manaus was recorded in 2009, followed by a record low in 2010. The forests beside the rivers are flooded to a depth of several meters for 5–9 months each year. The two main flooded ecosystems are those of the white water rivers called várzea and those of the black and clear water rivers called igapó (Prance, 1979).

Várzea Forest

Várzea forest is formed on soil that is flooded by the alluvial-rich white water rivers. The soils are therefore much more fertile than those of the terra firme because of the annual deposit of alluvial matter. For this reason, much of the várzea forest has been replaced by agricultural systems that are more sustainable than those on terra firme. Seasonal várzea is forest that is periodically flooded by the seasonal rise and fall of river level, and tidal várzea is flooded twice a day by fresh water as a result of backup caused by the tide in the Atlantic Ocean. There are few differences between seasonal and tidal varzea and most plant species can occur in both types.

Várzea forest is tall and is physiognomically quite similar to terra firme forest, but it is much less diverse in species. Buttressed trunks are common in várzea forest trees, and lianas are also common. The várzea often has a rich understory with a large number of species of Zingiberaceae, Marantaceae, and Heliconiaceae. In upper Amazonia the várzea forests are much richer in species composition than those of lower Amazonia. Várzea forests tend to have high riverbanks that form natural levées. Some of the typical species are *Buchenavia macrophylla*, *Ceiba pentandra* (the kapok tree), *Crataeva benthamii*, *Gustavia hexapetala*, *Hevea brasiliensis* (the rubber tree), *Hura crepitans*, *Piranhea trifoliata*, *Rheedia brasiliensis*, and *Virola surinamensis*, as well as numerous species of palms such as *Astrocaryum jauari*.

In the lower Amazon, where the river is very wide, especially between Itacoatiara in Amazonas and Monte Alegre in Pará, the várzea forest is often accompanied by large meadows of grassland known as canarana (*Echinochloa spectabile* being the dominant species). This region has narrow bands of forest on higher ground beside the rivers and extensive grass meadows on the lower, more flooded ground farther away from the river margins; the lowest areas form lakes. When the waters recede and the lakes dry up, the open ground quickly becomes covered with grass, increasing the area of the meadows. The arboreal species that are most common in this type of várzea are *Bombax munguba*, *Calycophyllum spruceanum*, *Cr. benthamii*, *Cordia tetrandra*, *Pithecelobium multiflorum*, *Muntingia calabura*, and *Salix martiana* var. *humboltiana*. The latter species of willow occurs mainly on muddy beaches that are in the process of formation.

Igapó

The black water and clear water rivers form a very different ecosystem than do the white water rivers. In this case there is little or no sediment and the water is acid and nutrient poor. In black water rivers the water contains much humic matter. The rivers flow from either sandy or rocky terrain. The largest black water river is the Rio Negro and the three major clear water rivers are the Rios Tapajós, Tocantins, and Xingu.

In these rivers in the dry season the banks are usually gradually sloping sandy beaches with trees on them, but in the rainy season they are transformed into an inundated igapó forest where only the canopy is above water. Characteristic species include *Alchornea castaniifolia*, *Couepia paraensis*, *Exellodendron barbatum*, *Eschweilera tenuifolia*, *Neoxythece elegans*, and *Virola elongata*. Three common palms of the igapó of the Rio Negro are the piassaba, *Leopoldinia piassaba*, the closely related *Leopoldinia pulchra* and the caranaí (*Mauritiella aculeata*).

The physiognomy of the igapó forest is usually rather different from that of the terra firme and várzea forests. This forest is of a lower stature and the trees are more scattered. Many species are confined to the igapó forest but local endemism is rare. One interesting local endemic, *Polygonanthus amazonicus*, occurs on the sandy beaches near Maués in Pará, and it is still of uncertain plant classification.

Near Manaus, where the large black water Rio Negro and the white water Amazon (called the Rio Solimões above Manaus) merge, there are extensive areas of inundation forest that are influenced by both black water and white water. A study of this area showed that the floodplain forest contains a mixture of species that are characteristic of varzea and igapó as well as many of the species that are common to both ecosystems, such as *Astrocaryum jauari*, *Calophyllum brasiliense*, *Eschweilera albiflora*, *Gustavia augusta*, *Hevea spruceana*, *Pari. excelsa*, and *Vatairea guianensis*.

The seeds of both várzea and igapó species are often light and have many different flotation mechanisms, such as spongy tissue, oil bodies, hollow areas, or a light mesocarp. In some cases it is the seed that floats, and in others the entire fruit floats to ensure dispersal by the water currents, for example, in the macucu (*Ald. heterophylla*). In the rubber tree the seeds float, and in the case of *Montrichardia arborescens*, an aquatic aroid, the whole fruit floats.

Other riverside trees have a fleshy fruit that is often colored green. These are the ichthyochorous or fish-dispersed species. When the fruit of these trees falls into the river in the flood season, fish can be seen grabbing them rapidly. The local people know that they can catch these fruit-eating fish by using tree fruit as bait on their hooks. For example, the fruit of *Alc. castaniifolia*, *Carapa guianensis*, and *Astrocaryum jauari* are frequently used as bait. The seeds of most species of the igapó and várzea forest are dispersed either by water or by fish. Some of the shrubs that grow on riverbanks, such as *Myrciaria dubia* or the camu camu, *Alc. castaniifolia*, and *Couepia uiti*, produce fruit as the river level rises. The fruit become submerged and

are then readily available to their fish dispersal agents, who pluck them from the trees.

Lakes, Swamps, and Other Aquatic Systems

Lakes

Not all of the flooded ground is covered by the tall, species diverse, várzea and igapó forests. There are many lakes scattered throughout Amazonia and some of them are extensive, such as Lago Manacupuru in Amazonas and Lago Grande in Pará. In the upper Amazon in Peru and in the upper reaches of some of the main Amazon tributaries, oxbow lakes are also common.

The lakes of Amazonia have an extensive macrophyte flora. White water lakes beside the main Amazon River that are cut off from river flow in the dry season are the habitat of the world's largest water lily, *Victoria amazonica*. This spectacular plant is pollinated by scarab beetles that the lily traps inside its flowers for 24 h. There are many species of Pontederiaceae, including the water hyacinth *Eichhornia crassipes*, which has become a noxious weed in many other parts of the tropics. Another common floating aquatic in the same family is *Reussia rotundifolia*.

In most lakes the aquatic flora can be divided into the free-floating species, those that are rooted to the bottom of the lake, and those growing around the margin of the lake. In the free-floating group there are several aquatic ferns such as *Azolla*, *Salvinia*, *Ceratopteris pterioides*, and *Marsilia polycarpa*. Others include a floating spurge, *Phyllanthus fluitans*, a legume, *Neptunia oleracea*, and *Ludwigia helminthorriza* in the willow-herb family or Onagraceae. Some of the rooted aquatics are *Limnocharis flava*, *Pacourina edulis*, *Caperonia castaneifolia*, *Aeschynomene rudis*, *Ludwigia decurrens*, and *Sphenoclea zeylanica*. Growing around the lake margin are common species such as *Cassia occidentalis*, *Cyperus mutisii*, *Paspalum fasciculatum*, *Hyptis parkeri*, and *Diodea kuntzei*. When the lake level rises, large mats of floating vegetation, mainly the grasses *Panicum* and *Echinochloa*, are washed out into the river and can be seen floating downstream.

Permanent Swamps and Buritizal

In a few places in Amazonia the water never drains and permanent swamp occurs. The permanent swamp forest has a low species diversity and covers only a small area of the region. The soil under the forest is a eutrophic humic gley and in some areas of dystrophic humic gley palm swamp occurs. This is usually dominated by species of *Mauritia* or *Euterpe*. There are several extensive areas of *Mauritia flexuosa* palm swamp that are locally called buritizal. These occur in eastern Amazonia in the state of Maranhão, in central Amazonia in the basin of the Rio Jutaí, and in the west near Iquitos, Peru. Since almost all parts of both *Mauritia* and *Euterpe* are useful, these oligarchic or one-species forests have great potential for sustainable use. *Euterpe oleracea* occurs in large quantities in the swamp forests of Marajó Island in the Amazon delta, where it is harvested extensively for both its fruit and as a source of palm hearts or palmito. *Mauritia flexuosa* also occurs in the gallery forests of the Roraima savanna and in várzea forest and is not confined to the permanent swamp forest.

Pirizal

Pirizal refers to a type of vegetation that occurs in small restricted areas of eastern Amazonia. A pirizal is a shallow lake or pond with stagnant water and a large number of rooted plants that emerge above the water. The commonest species are the large sedge, *Cyperus giganteus*, and *Thalia geniculata*, an aquatic member of the Marantaceae. There are also a number of typical floating aquatics such as species of *Cabomba*, *Eichhornia*, *Limnathemum*, *Nymphaea*, *Sagittaria*, and *Salvinia*. The *Mauritia* palm often grows around the margins. The pirizals occur as enclaves in the midst of dense forest, mainly in the state of Amapá in an area along the coast, especially in the region of Mazagão. There are also sedge-dominated communities in Amapá called cariazal. These are dominated by *Diplasia karataefolia*, the local name of which is cariá.

Mangrove Forest

The mangrove forest of Amazonia occurs as a narrow littoral belt in the coastal area that is subject to salt water inundation by tidal movements. Amazonian mangrove forests are species poor and have a rather uniform formation. The commonest species is the red mangrove, *Rhizophora mangle*, which occurs in the saltiest water. *Rhizophora racemosa* occurs in less salty brackish water, and the intermediate hybrid species *Rhizophora harrisonii* occurs in the area in between. All three species are found in the mangroves of the Amazon coast.

The other common species of this formation are *Avicennia nitida*, the white mangrove, *Laguncularia racemosa*, and *Conocarpus erectus*. *Avicennia* grows inland for a considerable distance upriver into the zone of fresh water, whereas all three species of *Rhizophora* do not extend beyond the influence of salt water. In areas of clay salt beaches the small grass *Spartina brasiliensis* plays an important role in stabilizing the soil as it occurs in thick dense carpets. Some of the species that are common around the fringes of Amazonian mangrove are *Annona palustris*, *Hi. tiliaceus*, *Pithecellobium cochleatum*, *Pterocarpus officinalis* and the mangrove fern *Acrostichum aureum*. The Atlantic coastal mangrove forests of South America are species poor in comparison to those along the Pacific coast. A good review of studies of the Amazonian mangrove vegetation is given in Menezes *et al.* (2008).

Amazonian Savannas

Nonflooded Savannas and Cerrado

Scattered throughout the rain forest region of Amazonia are a number of patches of open grassland or savanna. Some of these cover small areas and others are quite extensive. The largest ones are the Roraima–Rupununi savanna in Guyana and northern Brazil, which covers 54,000 km^2, the Sipaliwini savanna on the border between Brazil and Suriname, the Pampas del Heath in Peru, and the Humaitá savannas in the south of Amazonas state. Amazonia is bordered to the south by the largest savanna area of South America, the cerrado of the Central Brazilian Plateau. There are also banks of cerrado that extend into the rain forest region of southern Amazonia, but the isolated Amazonia savannas are a rather different

formation from the cerrado. The Amazonian savannas have less local endemism and species diversity than the cerrado, few suffrutices (woody plants with underground branches and short aerial shoots which die back after fires and resprout), and less tortuously branched trees, indicating that fire is not as strong a factor as it is in the drier cerrado.

The cerrado region of central Brazil has an extremely well defined dry season, the air humidity is often low, and the soils are very deep. As a consequence, roots are also deep and are specialized to reach water at considerable depth when the upper layers of soil are dry. When the topsoil is scraped off by a tractor, immediately there is a large amount of sprouting from vegetative organisms such as xylopodia. In contrast, in Amazonian savannas the roots are more superficial and there are not many vegetative parts in the soil to allow vegetative reproduction. Also, the climate in Amazonia is more humid and the relative humidity of the air is higher. Furthermore, in Amazonian savannas there is never such a dense arboreal cover as in the more closed types of savanna, whereas in the cerrados, dense closed vegetation (cerradão) is common and is similar to the dry semideciduous forest. Ratter *et al.* (2003) analyzed the species composition of 376 areas of cerrado and Amazonian savannas and found that in both systems of classification that they used the Amazonian disjunct sites formed a distinct group separated from the cerrado samples.

The two ecosystems have many species in common, for example, *Antonia ovata*, *Byrsonima verbascifolia*, *Curatella americana*, *Hancornia speciosa*, *Palicourea rigida*, *Qualea grandiflora*, *Salvertia convalarriodora*, and *Tabebuia caraiba*. The wide distribution of these and other species is evidence of a historically more continuous distribution of savannas during previous periods of drier and cooler climate. However, it has been shown that many savanna species have a much greater capacity for long-distance dispersal than rain forest trees. Bird- and wind-dispersed diaspores are common in savanna ecosystems, enabling distribution into isolated islands of savanna.

Most of the larger areas of savanna, such as those of Roraima, the savannas of the Rio Paru region of Pará state, and the extensive savannas around Santarém, also in Pará, occur within the drier and more seasonal regions of Amazonia. A belt of drier seasonal climate extends from Roraima to Santarém with under 2000 mm of rainfall and a strong dry season. Other savannas, especially some of the smaller patches, occur for edaphic reasons, such as an underlying hardpan.

There are a large number of savannas in Amazonian Venezuela and these were subdivided into three types by Otto Huber (1982): (1) grassy llanos and llanos-type savannas confined to the north of the region, (2) grassy inundated savannas of the Manapiare–Parucito basin, and (3) the true Amazonian savannas of the central and western part of Amazonas Territory. Unlike many of the Brazilian savannas, the Amazonian savannas of Venezuela have a high rate of endemism. They are mainly of edaphic, pre-Quaternary origin, whereas the llanos and inundated savannas are relict areas from the Pleistocene arid periods. The savannas of Amazonia are therefore not just one single ecosystem.

Other savannas also vary greatly in their botanical diversity and degree of endemism. The Roraima, Sipaliwini, and Humaitá savannas have few endemic species, whereas a few savannas have a rich and diverse flora. Most notable are those of the Rio Cururu region of Pará, which occur over sandstone. The sandstone savannas of the Guayana region of Venezuela are also richer in endemism than those of central Amazonia. Savannas that occur over special edaphic conditions are generally much richer than savannas occurring in terrain with similar conditions to those of the surrounding forest. These edaphic savannas are probably older and more stable over time.

Along the northern fringes of Amazonia, other extensive savannas are the llanos of Colombia and the contiguous Gran Sabana of Venezuela. There are also extensive upland savannas in the Departments of San Martín and Madre de Dios in Peru. Overall, savanna covers approximately 15% of the Amazon region and so it adds considerable species diversity. Many of the savanna ecosystems are under threat by overuse for farming and agriculture. For example, much of the cerrado region is being destroyed to plant soy and the Roraima–Rupununi savannas are being intensively burned by cattle ranchers and more recently also for soy. In the dry El Niño years of 1998 and 2005, the fires spread into the rain forest, causing extensive destruction. Nepstad *et al.* (2004) showed that there is much reduced tree growth and how susceptible the terra firme forest is to fire in these years of drought.

Flooded Savannas of Eastern Amazonia

In the delta region of Amazonia in the state of Amapá and in the eastern half of Marajó Island, there are large areas of inundated savanna or campos de várzea as they are called locally. Other inundated savannas occur farther inland in Pará, especially between the Rios Xingu and Tapajós, and in the basin of the Rio Madeira in Amazonas, but these are much smaller than those of Marajó and Amapá. There are many species of grass in the flooded savannas, such as *Eragrostis hypnoides*, *Eragrostis glomerata*, *Parathera prostrata*, *Paspalum orbiculatum*, and *Phenakospermum guianense*. But the main difference between flooded and nonflooded savannas is the greater frequency of sedges (Cyperaceae) such as *Cy. giganteus* in the wettest areas and *Cyperus luzulae*, *Cyperus ferax*, and *Scleria geniculata* in the less flooded areas. Shrubs are less abundant in flooded savanna, but in areas that are grazed by cattle *Artemisia artemisiifolia* becomes so abundant that it can completely dominate an area because it is unpalatable to cattle. *Ipomoea fistulosa*, a shrub that is toxic to cattle, also increases on grazed land for the same reason. Other species characteristic of flooded savannas include *Aeschynomene sensitiva*, *Alternanthera philoxeroides*, *Capironia fistula*, *Cassia reticulata* (known locally as mata pasto or pasture killer because of its tendency to invade pasture), *Clitoria triquetrum*, *Justicia obtusifolia*, *Mimosa pigra*, *Montrichardia linifolia*, *Phaseolus lineatus*, *Rhabdadenia macrostoma*, *Sesbania exasperata*, and *Th. geniculata*.

In the flood season, various floating aquatic species, such as *Ceratopteris pteridioides*, *Eichhornia azurea*, *Nep. oleracea*, *Pistia stratiotes*, and *Salvinia radula*, multiply with great rapidity.

Montane Fringing Ecosystems

Around the northern and southern edges of Amazonia, on the older pre-Cambrian formations of the Guiana and Brazilian

crystalline shields, there are a number of hills and mountains, and to the west Amazonia is bordered by the Andes Mountains. These upland areas, which enclose the lowland Amazon basin on three sides, add considerably to the biological diversity of the Amazon ecosystem. Small changes in altitude or topography can cause significant changes in the vegetation type, physiognomy, species composition, and climate.

The Sandstone Tepuis of the Guayana Highland

An area with a complex mixture of vegetation types is the Guayana Highland, which is dominated by dramatic sandstone table mountains or tepuis. The mountain slopes are covered by lower and upper montane forest and often there is an abrupt change to open savanna and swamp formations on the summits of the mountains because of the sheer sandstone cliff faces. The highest mountains have less forest on their summits and more open vegetation except for gallery forest along the streams. The lower mountains have considerable areas of forest on their summits. For example, on Cerro Guaiquinima in Venezuela, one of the largest sandstone tepuis, 40% of the summit is covered by forest. Both tall forest and relatively low dense forest, as well as intermediate types, occur on Guaiquinima. There is a strong relationship in species content between the forest on the summit of this and other tepuis and the forest of the neighboring lowlands. Species endemism is high on the sandstone tops of the Guayana Highlands, and many genera have a series of closely related species with one occurring on each of the larger mountains, for example, the pitcher plant genus *Heliamphora* or the shrub *Tepuianthus*.

Two outlying tepuis, Tepequém and Araca, occur well within the Brazilian Amazon region, and Mounts Roraima and Neblina both lie on the Brazilian border with Venezuela (see Prance and Johnson, 1992). The latter is the highest tepui, reaching 3045 m in altitude. The Guayana Highland region is a complex mosaic of tepuis interspersed with rain forest or savanna depending on the locality.

Granite Inselbergs

Scattered throughout the Guianas and northern Amazonian Brazil are a series of granitic outcrops or domes that rise above the rain forest to a height of 300–800 m. In some cases these hills are covered by dense forest (e.g., Palunlouiméempeu and Mitraka), but most of them are characteristic inselbergs with low scrub forest and open areas of exposed rocks. These are obviously islands of a special type of vegetation surrounded by rain forest. The tops of the inselbergs are well drained and become very dry in the dry season, and so the vegetation is often dominated by sclerophyllous plants or species with other adaptations to drought, such as the orchid *Cyrtopodium andersonii* with large pseudobulbs that store water or various cacti (*Epiphyllum* and *Melocactus*). The dominant shrub is usually a species of *Clusia* and other terrestrial orchids include *Epidendrum nocturnum* and *Encyclia ionosma*, and the bromeliad *Pitcairnia geyskesii* is also common. Many of the plant species on the inselbergs of the Guianas are confined to that habitat. There has been considerable speciation in adaptation to the summit of inselbergs with their arid dry season conditions and very humid rainy season conditions. However, in contrast to tepuis, there has been much less speciation between the different inselbergs, and local endemics are rare.

Low Hills within Amazonia

In addition to the tepuis and inselbergs scattered throughout the Amazon region are several lower hills that can have distinct vegetation. These are more frequent to the north of the Amazon River and near the Brazilian Shield region and are mostly outliers of the ancient shields.

Small changes in altitude can have extremely important effects on vegetation type, physiognomy, species composition, and climate. For example, outside Amazonia, small low hills in the arid northeast of Brazil tend to accumulate cloud and therefore are covered by tall moist forest called brejo. The brejo forests contain many Amazonian species. The different effects between larger and smaller mountains (known as the Massenerhebung effect), as well as latitude, on the vegetation type are quite striking in the Neotropics. Consequently, there are many small patches of montane or cloud forest on small outcrops throughout the lowland region. Also the altitudinal limits of the different forest types vary considerably depending on the effects of local climate, soil, and latitude. Some of the species and genera that occur on the higher mountains also occur on these areas, for example, *Retinophyllum* and *Pagamea* in the Rubiaceae.

The Andes Foothills

The Amazon region is bordered on the west by the Andes, which rise abruptly out of the lowland rain forest. Unlike the tepuis and inselbergs, the Andes are a relatively recently formed mountain chain and have a distinct zonation of the vegetation according to altitude. The forests of the eastern slopes fall into the ecosystems of Amazonia and add considerable variety of both vegetation types and species.

The lowland forest extends up the mountains to between 700 and 1000 m. Above that altitude it is replaced first by lower montane, and then by montane vegetation types. The montane forests can generally be divided into three zones: lower montane forest, upper montane forest, and subalpine forest.

Lower montane forest begins between 700 and 1200 m. It is quite similar to lowland rain forest, but a large number of species drop out and are replaced by more upland species. This forest tends to be lower in stature than that of the surrounding lowland area, and it has fewer woody vines and fewer buttressed trees. Vascular epiphytes are common and mesophyll leaf types are predominant. Many of the lowland taxa persist into this zone, such as species of *Licania* (Chrysobalanaceae) and *Eschweilera* (Lecythidaceae), but a number of distinctly highland elements also enter the lower montane forest, for example, in the Colombian Andes, *Brunellia comocladifolia*, *Alchornea bogotensis*, and *Cinchona calycina*. Palms such as *Euterpe purpurea* and *Wettinia cladospadyx* are quite abundant, and woody vines include *Anomospermum occidentale* and species of *Passiflora* and *Paullinia*. This zone typically

extends upward to between 1800 and 2400 m depending on latitude and local conditions.

The upper montane forest begins at 1800–2400 m and may extend in places up to 3400 m. It is usually of lower stature than the lower montane forest, with a predominance of microphyllous trees. Vascular epiphytes are still common and woody lianas are rarer. An increasing number of species characteristic of higher altitude enter the flora, for example, *Brunellia occidentalis*, *Weinmannia balbisiana*, *Symplocos pichindensis*, and the vines *Hydrangea peruviana* and *Liabum megacephalum*.

The subalpine forest, including the dwarf elfin forests, is of frequent occurrence in the Andes and extends locally into Central America. It is microphyllous or, at its altitudinal extreme, nanophyllous or dwarf-leaved. Few vascular epiphytes and climbers occur, but there are abundant bryophytes and lichens. This formation may extend up to 3800 m in some places, above which it is replaced by the herbaceous alpine formations such as páramo, puna and *Polylepis* forest which are not discussed here. The subalpine forest has few predominantly lowland genera but many that are characteristic of the highlands, for example, species of *Weinmannia* (Cunoniaceae) and *Gynoxys* (Asteraceae), *Brunellia* (Brunelliaceae), *Clusia* (Clusiaceae), *Befaria* (Ericaceae), *Miconia* (Melastomataceae), and *Rhamnus* (Rhamnaceae).

A most interesting type of dry forest occurs in small isolated dry valleys that are scattered along the Andes due to local weather conditions They are significant as a link between the larger arid areas, and they were probably more nearly connected during the dry periods of the Pleistocene and early Holocene (see Pennington *et al.*, 2000). Important genera in these valleys include *Acacia*, *Bursera*, *Caesalpinia*, *Cercidium*, *Prosopis* (*Prosopis chilensis* and *Prosopis limensis*), and many Cactaceae. Linares-Palomino *et al.* (2003) conducted a multivariate analysis of the tropical dry forests some of which are on the fringes of Amazonia.

Campo Rupestre

The campos rupestres, or open formations on rocks, are generally confused with open savanna and the orchard savanna on terra firme, but they are quite different physiognomically and floristically. They develop on rocks and in rocky terrain, because they suffer drought in the dry season despite the equatorial humid climate, and because there is no possibility for water retention and all rainfall runs off immediately. Typical species of the cerrado do not occur here, such as *Cu. americana*, *Ha. speciosa*, *Salvertia convallariodora*, and *Q. grandiflora*. Instead there are various species of *Byrsonima*, *Clusia*, *Norantea*, and *Vellozia*, certain Bromeliaceae (*Bromelia*, *Dyckea*, and others), and many Eriocaulaceae (*Paepalanthus* and *Syngonanthus*). Lichens are also frequent and cover many of the rocks.

Campos rupestres are also quite common in central Brazil, where they are often confused with cerrados. For example, Serra do Cipó in Minas Gerais, much cited in the literature because of its interesting landscape and large number of Velloziaceae, is mostly composed of campo rupestre. In Amazonia, Serra do Cachimbo, in the southwest of Pará, has an area of campo rupestre that covers about 16,000 km^2. Another large patch is the campo of the Rio Cururú, in the Tapajós River basin. At Ariramba, on the Rio Trombetas, both savanna and campo rupestre occur.

The canga or ironstone vegetation, which occurs on Serra dos Carajás over the iron deposit there, is also a special form of campo rupestre. On this formation, various extra-Amazonian species are common, such as *Pilocarpus microphyllus*, which occurs also in the northeastern states of Piauí and Maranhão; the extra-Amazonian genus *Callisthene* is represented by *C. microphyllus* of central Brazil, which occurs only within Amazonia on Serra dos Carajás. Also found on Carajás is the curious and beautiful treelet *Norantea goyasensis*. The branches of this thick-barked treelet sometimes lengthen and turn it into a vine. A review of vegetation on canga and the threats to it is given in Jacobi *et al.* (2007).

The Amazonian montane forests also show a strong affinity to campo rupestre and some could even be classified as floresta rupestre (i.e., forest on rock). In the mountain areas there has been a great accumulation of organic matter because, although the forests sometimes endure short dry periods, they are maintained by the high air humidity and mist.

Secondary Forest

Unfortunately, secondary forest is of increasing importance in Amazonia owing to the amount of felling of primary forest that is taking place. Obviously the species that now dominate secondary areas had a natural distribution prior to the occurrence of man-made secondary forest. A few secondary forest species are found throughout the tropics, and others grow in open spaces beside rivers, on old landslides, in gaps caused by tree falls or dead trees, or where forest has been felled by severe storms. Because the seeds and stumps have not been destroyed by fire, regenerated vegetation of naturally cleared areas is different from that in areas burned by man. Secondary forest species do not generally occur in the transition regions between forest and savanna. Modern vegetation maps show an increasingly large area of a new man-made vegetation and also of agricultural areas, but these are not discussed here. Some of the characteristic genera of secondary forests are *Cecropia*, *Byrsonima*, and *Vismia*. In Brazil, secondary forest is known by the term capoeira. More recently attention has been drawn to the importance of secondary forests for the biodiversity that they hold, for example, Chazdon *et al.* (2009). A comparison of the diversity of primary and secondary forests is given by Barlow *et al.* (2007) and they concluded the primary forest is of irreplaceable value. Almeida *et al.* (2010) provided an estimate of the area of secondary forest in Amazonian Brazil.

Conclusions

This summary of the principal vegetational formations that occur in Amazonia shows the considerable diversity that accounts for the approximately 30,000 species of vascular plants within the region. Many of the ecosystems described here could be subdivided further into other different local

communities. A knowledge of this ecosystem diversity and the distribution of species is an essential tool for the conservation and sustainable use of the region. To preserve the diversity of Amazonia, it is vital to conserve areas of each different type of vegetation because of the number of unique species of plants and animals that each ecosystem contains.

See also: Ecosystems of South America. Rainforest Ecosystems, Animal Diversity. South American Natural Ecosystems, Status of. Southern (Austral) Ecosystems

References

Almeida CA, Valeriano DM, Escada MIS, and Renno CD (2010) Estimativa de área de vegetação secundária na Amazônia Legal Brasileira. *Acta Amazonica* 40: 289–302.

Barlow J, Gardner TA, Araujo IS, *et al.* (2007) Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. *Proceedings of the National Academy of Sciences of the USA* 104: 18555–18560.

Boom BM (1986) A forest inventory in Amazonian Bolivia. *Biotropica* 18: 287–294.

Campbell DG, Daly DC, Prance GT, and Maciel UN (1986) Quantitative ecological inventory of terra firme and várzea tropical forest on the Rio Xingu, Brazilian Amazon. *Brittonia* 38: 369–393.

Chazdon RL, Harvey CA, Komar O, *et al.* (2009) Beyond reserves: A research agenda for conserving biodiversity in human-modified tropical landscapes. *Biotropica* 41: 142–153.

Daly D and Prance GT (1989) Brazilian Amazon. In: Campbell DG and Hammond HD (eds.) *Floristic Inventory of Tropical Countries*, pp. 401–426. New York: The New York Botanical Garden

Ducke A and Black GA (1953) Phytogeographical notes on the Brazilian Amazon. *Anais da Academia Brasileira de Ciencias* 25(1): 1–46.

Gentry AH (1982) Patterns of neotropical species diversity. *Evolutionary Biology* 15: 1–84.

Gentry AH (1988) Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the USA* 85: 156–159.

Glaser BL and Woods WI (eds.) (2004) *Amazonian Dark Earths: Explorations in Space and Time*. Berlin: Springer-Verlag.

Holdridge LR (1972) *Forest Environments in Tropical Life Zones: A Pilot Study*. Oxford/New York: Pergamon Press.

Huber O (1982) Significance of savanna vegetation in the Amazon Territory of Venezuela. In: Prance GT (ed.) *Biological Diversification in the Tropics*, pp. 221–244. New York: Columbia University Press.

Hueck K (1966) *Die Walder Südamerikas*. Stuttgart, Germany: Gustav Fischer Verlag.

Ivanuskas NM, Monteiro R, and Rodrigues RR (2008) Classificação fitogeográfica das florestas do Alto Rio Xingu. *Acta Amazonica* 38: 387–402.

Jacobi CM, do Carmo FF, Vincent RC, and Stehman JR (2007) Plant communities on ironstone outcrops: A diverse and endangered Brazilian ecosystem. *Biodiversity and Conservation* 16: 2185–2200.

Linares-Palomino R, Pennington RT, and Bridgewater S (2003) The phylogeography of the seasonally dry forests in Equatorial Pacific South America. *Candollea* 58: 473–490.

Menezes MPM, de U, Berger, and Mehlig U (2008) Mangrove vegetation in Amazonia: A review of studies from the coast of Pará and Maranhão States, north Brazil. *Acta Amazonica* 38: 403–420.

Nepstad D, Lefebvre P, Da Silva UL, *et al.* (2004) Amazon drought and its implication for forest flammability and tree growth: A basin wide analysis. *Global Change Biology* 10: 704–717.

Paz-Rivera C and Putz FE (2009) Anthropogenic soils and tree distributions in a lowland forest in Bolivia. *Biotropica* 41: 665–675.

Pennington RT, Prado DE, and Pendry CA (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography* 27: 261–273.

Pires JM and Prance GT (1985) The vegetation types of the Brazilian Amazon. In: Prance GT and Lovejoy TE (eds.) *Key Environments: Amazonia*, pp. 109–145. Oxford, UK: Pergamon Press.

Pitman NCA, Mogollón H, Dávila N, *et al.* (2008) Tree community change across 700 m of lowland Amazonian forest from the Andean foothills to Brazil. *Biotropica* 40: 525–535.

Prance GT (1977) The phytogeographic subdivisions of Amazonia and their influence on the selection of biological reserves. In: Prance GT and Elias TS (eds.) *Extinction is Forever*, pp. 195–212. New York: The New York Botanical Garden.

Prance GT (1979) Notes on the vegetation of Amazonia III: The terminology of Amazon forest types subject to inundation. *Brittonia* 31: 26–38.

Prance GT (1987) Vegetation. In: Whitmore TC and Prance GT (eds.) *Biogeography and Quaternary History in Tropical America*. Oxford Monographs on Biogeography 3, pp. 28–44. Oxford, UK: Clarendon Press.

Prance GT and Johnson DM (1992) Plant collections from the plateau of Serra do Aracá (Amazonas, Brazil) and their phytogeographic affinities. *Kew Bulletin* 47: 1–224.

RADAMBRASIL *Levantamento de recursos naturais Vegetação*. vol. 8, pp. 307–403. Rio de Janeiro: Min. das Minas e Energia, Dept. Nacional da Produção Mineral Projeto Radambrasil.

Ratter JA, Bridgewater S, and Ribeiro JF (2003) Analysis of the floristic composition of the Brazilian cerrado vegetation III: Comparison of the woody vegetation of 376 areas. *Edinburgh Journal of Botany* 60: 57–109.

ter Steege H, Pitman N, Sabatier D, *et al.* (2003) A spatial model of tree α-diversity and tree density for the Amazon. *Biodiversity and Conservation* 12: 2255–2277.

Valencia R, Balslev H, and Paz Y Miño G (1994) High tree alpha diversity in Amazonian Ecuador. *Biodiversity and Conservation* 3: 21–28.

Amphibians, Biodiversity of

Ross A Alford, James Cook University, Townsville, QLD, Australia
Stephen J Richards, South Australian Museum, Adelaide, SA, Australia
Keith R McDonald, Queensland Department of Environment and Heritage Protection, Atherton, QLD, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Amphibians are tetrapod vertebrates They differ from the other tetrapods (the reptiles, birds, and mammals) in that their eggs are anamniotic; they are relatively simple and are enclosed in a jelly capsule.

Anamniotic Eggs are not surrounded by the complex membranes that distinguish the amniotic eggs of reptiles, birds, and mammals.

Cloaca The common chamber in which the reproductive, excretory, and digestive tracts of amphibians unite before exiting the body.

Metapopulation A group of local populations among which individuals migrate relatively frequently; however, the rate of migration is slow enough that the populations fluctuate independently.

Paedomorphosis Reproduction while retaining at least some larval characteristics.

Tetrapods The terrestrial vertebrate classes amphibians, reptiles, birds, and mammals, so named because they primitively possess four legs.

The Evolutionary History of Amphibians

Amphibians first appeared during the Late Devonian, about 360 million years ago (Ma). There is a general consensus that all amphibians shared a common ancestor, a sarcopterygian (fleshy-finned) bony fish in the class Osteichthyes. All sarcopterygians have paired fins with limb-like bones, and many exhibit other anatomical features, such as lungs, that ally them with the tetrapods. It is not likely that either of the extant groups of sarcopterygians, the dipnoans (lungfishes), and crossopterygians (lobe-finned fishes), contained the ancestor of the amphibians. The earliest major radiation of terrestrial vertebrates occurred during the Carboniferous Period (*ca.* 335 Ma). By the end of the Triassic, about 200 Ma, nearly all of the large ancestral amphibians were extinct. The subclass Lissamphibia, the modern amphibians, appeared during the Triassic and is the only group that has survived to the present.

The ancestry of modern amphibians is poorly understood because there is a sparse fossil record linking primitive amphibians to the three modern orders. The earliest fossil Lissamphibian is *Triadobatrachus massinoti* from the early Triassic, about 230 Ma. *Triadobatrachus* is similar to the modern frogs but is not considered to belong to the order Anura. The Anura (frogs) and Gymnophiona (caecilians) appeared during the early Jurassic (about 190 Ma), whereas the Caudata (salamanders) appeared in the Middle Jurassic, 170–150 Ma. Although the fossil evidence is sparse, phylogenetic analyses of shared derived morphological characters and of molecular characters strongly suggest that the Lissamphibia share a common early amphibian ancestor and that the Lissamphibia and the amniote tetrapods (reptiles, birds, and mammals) originated from different early amphibians. It is possible; however, that the Lissamphibia is a polyphyletic group, with one or more of the modern orders having an independent origin from the subclasses Temnospondylii or Microsauria.

Historical Biogeography and Current Diversity of Modern Amphibians

Modern amphibians are found in all continents except Antarctica. However, the three modern orders show disparate patterns of fossil and recent diversity that are associated with their different histories on the landmass of Pangaea and subsequently on landmasses associated with Gondwana and Laurasia. The evolution of salamanders is linked closely with landmasses derived from Laurasia. All fossil and living species occur in the Northern Hemisphere, with the exception of the lungless salamanders (family Plethodontidae) which invaded South America when the isthmus of Panama completed the connection between North and South America. In contrast, the caecilians and most frogs are associated predominantly with the southern Gondwanan landmasses. The caecilians occur throughout the tropics, except in Madagascar and to the east of Wallace's Line in Australasia. They are the least diverse group of living amphibians, with five families and 188 species, although their highly cryptic nature suggests that many species may remain to be discovered and described. The greatest diversity of species occurs in northern South America and Central America. Frogs are the most widespread of the three orders. Although the breakup of Pangaea probably isolated salamanders on the Laurasian continents and many groups of frogs on the Gondwanan continents, several clades of frogs have dispersed to all continents that support amphibians.

Frost *et al.* (2006) proposed a sweeping revision of the taxonomy of amphibians at all levels. Many aspects of their proposals are still being actively debated. To make it easier to find more information on taxa that may be of interest and relate the basic facts presented here to the broader literature, we have largely retained the traditional scheme. There are now more described species of amphibians than there are of mammals (Stuart *et al.*, 2004; **Table 1**). More than 600 salamander species are currently known in 10 families, with the highest diversity found in North America. One family, the

Table 1 Richness of modern amphibian taxa

Taxon	*Families*	*Genera*	*Species*
Anura	33	354	6012
Caudata	10	59	608
Gymnophiona	5	33	188
Total	48	446	6808

Plethodontidae of North and South America, contains more than half the known salamander species.

The frogs are placed in 25 families, and more than 6000 species are known, approximately 80% of which are in tropical regions. Roughly 1000 species have been described since the first edition of this encyclopedia, and estimates suggest that another one to two thousand species remain to be described. The greatest diversity of frogs occurs in South America, but other tropical areas such as Southeast Asia, New Guinea, Madagascar, and central Africa also have highly diverse frog faunas.

Four of the five most diverse frog families, the Hylidae, Ranidae, Bufonidae, and Microhylidae, are each found across several continents – distributions that reflect successful dispersal and long evolutionary histories. The most speciose family is the Leptodactylidae, which is restricted to South and Central America and contains 22% of living frogs. It contains the genus *Eleutherodactylus*, which was the most diverse vertebrate genus with more than 550 species until recent revisions created many new genera from many of its species.

Basic Morphology and Functional Anatomy of Modern Amphibians

All modern amphibians have complex glandular skins and all but caecilians lack scales. Their skins are kept moist by the secretions of mucous glands, whereas granular glands produce a variety of toxins that serve to deter predation and also produce a variety of antimicrobial peptides that deter microbial infections. They shed their skins periodically and usually consume them as they are shed. Their color patterns are produced by xanthophores, iridophores, and melanophores. Many species are able to alter the shape of pigment cells and the distribution of pigments within them and can rapidly change color.

The eyes of most species contain photoreceptors of several types, probably giving at least limited color vision, can be covered by moveable eyelids, and focus by moving the lens using ciliary muscles. Many frogs and salamanders have a wide binocular field of vision and have a good depth perception. Many species use olfaction in the detection and capture of prey and have well-developed olfactory systems and Jacobson's organs. Larval amphibians and the adults of many aquatic species have lateral line systems. Amphibians have no external ears; their tympanum (when present) is at the surface of the body. The middle ears have a columella bone (also found in reptiles and birds and modified in mammals), which transmits vibrations from the tympanum to the inner ear. In anurans, there is a second bony element in the middle ear, the operculum, which can be engaged by muscular contractions, and modifies hearing to increase the reception of low-frequency sounds.

Most terrestrial amphibians respire (in part) with lungs, and have largely separate pulmonary and systemic blood circulations. Their hearts typically have three chambers – two atria that receive blood from the body and the lungs and one ventricle that serves to pump blood to the lungs and the body. Although the ventricle is usually single, a combination of partial septa and the high viscosity of blood reduce mixing of the blood streams flowing to the lungs and the remainder of the body. Many species carry out a considerable fraction of gas exchange across the moist skin; for the lungless salamanders (family Plethodontidae) this is the primary mode of respiration.

The kidneys of amphibians function in maintaining water and ionic balance. In species with aquatic larval stages, most water and ion balance needs are reversed at metamorphosis; the freshwater larvae must cope with a hypo-osmotic environment, to which they lose ions and from which they gain water, whereas the terrestrial stages must cope with rates of evaporative water loss that are typically high because of their moist skins. Aquatic individuals tend to excrete dilute solutions of ammonia, whereas terrestrial animals excrete more concentrated solutions of urea, although no amphibian excretes fluid urine that is more concentrated than the blood plasma. Water uptake in most amphibians does not involve drinking but is accomplished by active or passive transport of water across the skin.

Amphibians are ectotherms, relying on the external environment as a source of body heat. All aquatic forms and many terrestrial forms are also poikilotherms or thermoconformers, with body temperatures that do not differ from their environments. Some species control their body temperatures to some extent by selecting microhabitats that provide appropriate temperatures, and some species of frogs periodically increase their body temperatures by basking in sunlight. The ability of many terrestrial amphibians to thermoregulate by basking is limited by their need to conserve water while maintaining a moist skin. However, many species of frogs have developed mechanisms that allow them to increase their resistance to evaporative water loss. Some species of frogs and salamanders can be active at temperatures approaching 0 °C, whereas others can tolerate body temperatures well above 40 °C. Some temperate species can survive exposure to temperatures below 0 °C by supercooling. A few species can survive partial freezing by a combination of ice-nucleating proteins that encourage freezing of extracellular fluids and high intracellular glucose concentrations that prevent intracellular water from freezing and keep cells from dehydrating.

Special Features of Caecilians

Caecilians are aquatic and burrowing animals that superficially resemble large earthworms. Adults range from approximately 10 to 150 cm in length. They have elongate bodies with distinct annuli, which are grooves delineating their body segments. They are limbless, and their tails are reduced or absent. Their eyes are reduced and are covered by skin. They are unique among the Lissamphibia in possessing

dermal scales, which occur in the annuli of some species. Their skulls are heavily ossified and completely roofed. Caecilians possess a unique chemosensory organ, the tentacle, which extends a short distance from the surface of the head, emerging from a skull opening between the eyes and the nostrils.

Special Features of Salamanders

Salamanders are typically four-limbed animals with relatively long tails. They superficially resemble lizards but lack epidermal scales and claws. Salamanders range from about 30 mm to 2 m in total length. Their limbs are relatively small and are reduced or lost in some terrestrial and aquatic species. Their skulls typically show the loss of many bony elements. Salamanders lack external ears and, with the exception of weak distress calls in some species, do not vocalize. The most diverse group of salamanders, the Plethodontidae of the Americas, is lungless. Salamanders of many groups exhibit various degrees of paedomorphosis, which in some species is facultative, allowing them to exploit high-quality aquatic environments while retaining the option of a terrestrial stage. Salamander larvae are carnivorous and usually have well-developed external gills.

Special Features of Anurans

Anurans are invariably four-limbed, and terrestrial juveniles and adults completely lack true tails. Adults range from about 1 to 30 cm in length. Their hind-limbs and feet are greatly elongated. The radius and ulna of the forelimb and the tibia and fibula of the hindlimb are usually fused. There are no more than nine trunk vertebrae, and most modern frogs lack free ribs. The caudal vertebrae are fused into a rod-shaped urostyle that is associated with the elongate pelvis. Most of these features are probably the result of adaptation for light weight and strength for jumping and show interesting convergences with similar adaptations in birds. The lightweight skulls of frogs are large relative to their body size and lack many bones. All but one species lack teeth on the dentary bone of the lower jaw. The tongue is attached at the front of the mouth in most species and is flipped forward rapidly to capture prey. The ears of frogs often have external tympanic membranes.

The skins of frogs depart from the usual amphibian pattern in several ways. Some species have additional types of glands: lipid glands which secrete lipids that reduce rates of evaporative water loss or breeding glands that produce sticky secretions which adhere the male to the female during amplexus. Skin lipids and other as yet poorly understood modifications of the skin allow some frogs to achieve rates of evaporative water loss as low as those of some lizards. Frogs also have distinct differences between their dorsal and ventral skins. The ventral skin usually has fewer granular and mucous glands, and many terrestrial species have an area called the pelvic patch in which thin and highly vascularized skin can be exposed. Water uptake from substrates via the pelvic patch is the primary means by which almost all species of frogs obtain water. Burrowing desert frogs of several lineages form cocoons by repeatedly shedding their skins and retaining the shed layers to reduce evaporative water loss while they are buried.

Frogs have a variety of adaptations to structures other than the skin for maintaining water balance. Most species store copious quantities of dilute urine in the bladder when water is available and withdraw water from this pool to replace evaporative losses. Desert species and species that inhabit brackish water can allow high concentrations of urea to accumulate in the body fluids. A few species can excrete nitrogenous wastes as uric acid, which minimizes the water lost when it is excreted.

Reproductive Biology and Life Histories of Modern Amphibians

The typical life history of the Lissamphibia includes a complex life cycle in which eggs are deposited in freshwater habitats, where larvae grow and develop, eventually metamorphosing into terrestrial juveniles and leaving the water. Although this is a common pattern found in all three modern orders, it is far from universal. One of the major features that set amphibians apart from the remainder of terrestrial vertebrates is their extremely wide range of life histories and modes of reproduction, many of which occur in all three extant orders (Duellman and Trueb, 1986). The presence of a wide range of life histories and reproductive modes suggests that the typical amphibian life cycle does not reflect a failure to adapt to the terrestrial environment but rather serves as an adaptation that allows female amphibians to produce very large numbers of eggs with a small investment in each and to exploit freshwater habitats for larval growth and development.

Caecilians

Limited observations have been made on the reproductive biology of caecilians and very little information on their courtship behavior is available. It appears that fertilization is internal via protrusion of the male cloacal wall. Although vocalizations have been reported for some species, it is not known if they are linked to reproduction.

Caecilians are oviparous or viviparous. In viviparous species gestation may take up to a year with reproduction occurring only every 2 years, and nutrition can be supplied to the young within the oviduct. Oviparous species lay terrestrial eggs but the larvae can be aquatic or can complete development within the egg. Oviparous caecilians produce more offspring than viviparous species. In oviparous species, parental care of eggs is common.

Salamanders

Many salamanders conform to the typical amphibian complex life cycle. Most produce aquatic eggs and larvae, which metamorphose into terrestrial juveniles. As in the frogs, there is a great deal of variation beyond this basic mode. Larval salamanders are relatively similar to adults; this has allowed species belonging to several groups to evolve the ability to

reproduce while retaining larval characters, either facultatively or obligately.

An example of the complexity and variability of salamander life histories is the life history of red-spotted newts (*Notophthalmus viridescens*, family Salamandridae). Their life history begins with aquatic eggs which hatch into aquatic larvae. The larvae typically develop within one summer and metamorphose into a morphologically distinct terrestrial juvenile stage, the eft. The eft stage last from 1 to 8 years, depending on temperature and food availability. Efts undergo a second transformation into adults, which then return annually to water to breed. Larvae that encounter very favorable conditions and grow rapidly can retain some larval morphological characteristics, remain permanently aquatic, and reproduce as paedomorphic adults. Larvae that encounter slightly less favorable conditions can bypass the eft stage and metamorphose directly into the adult morphology.

Courtship in salamanders can be quite elaborate, incorporating chemical, visual, and tactile cues. During the breeding season males of aquatic species can develop enlarged fins and become more brightly colored. There may be other morphological changes to the head glands, tooth structure, musculature, and skin during the breeding season. Many salamanders have specialized glands that secrete compounds used as olfactory signals during courtship.

Salamanders have internal or external fertilization. External fertilization involves deposition of the sperm on the egg mass. Salamanders with internal fertilization either transfer sperm directly from male to female or exhibit a unique form of sperm transfer in which a bundle of sperm (the spermatophore) is deposited by males. The spermatophore is picked up in the female cloaca and stored in a special structure in the cloaca, the spermatheca. The stored sperm can remain viable until ovulation, which may occur from a few days to 2.5 years later. In live-bearing salamanders the sperm enters the oviduct, in which the eggs are internally fertilized.

Both male and female salamanders of some species exhibit parental care (**Figure 1(a)**), although it is practiced mostly by terrestrial breeders and may serve to protect the eggs from predation, fungus infection, and desiccation. Adults attend clutches for up to 9 months. In a few species communal clutches are attended by aggregations of females, and eggs are fertilized by many different males.

Anurans

The most common reproductive strategy in frogs involves a complex life cycle with externally fertilized aquatic eggs which produce highly specialized larvae – tadpoles. The tadpoles grow and develop for some period in the water and then undergo a radical metamorphosis and emerge as terrestrial juvenile frogs. However, frogs have evolved a remarkable variety of reproductive strategies, most involving a trend toward protection of eggs and larvae from the hazards of the aquatic environment. Pough *et al*. (2004) described 29 combinations of egg deposition sites, including still water and fast-flowing water, terrestrial nests, and on or in the body of the male or the female, and they also described tadpole development, ranging from the typical free-swimming feeding larva to direct development inside the oviduct. Because amphibian eggs lack a water-resistant shell, the greatest diversity of reproductive modes occurs in the humid tropics where eggs can survive for long periods in the permanently moist terrestrial environment. Complex behaviors including egg guarding and embryo transport (on the dorsum, in specialized pockets or pouches, in the mouth, and in *Rheobatrachus* species in the stomach), and unusual morphological structures such as skin pockets to provide protection for developing embryos, are associated with many reproductive modes. Many separate lineages of anurans have evolved direct development, bypassing the aquatic larval stage. Most direct developers have some form of parental care of the eggs. Several species bear live young, and at least one of these, *Nectophrynoides occidentalis*, actually provides oviductal nutrition to its developing embryos.

In tropical regions where conditions for reproduction are favorable throughout the year, breeding can be unseasonal, and females may lay multiple clutches in 1 year. In more temperate or high-altitude regions breeding is typically strongly seasonal, occurring only during short periods of the year when temperature and rainfall reach critical levels. Under these conditions females generally lay only one clutch of eggs each year.

Male frogs vocalize mainly to attract mates (**Figure 1(b)**) and to advertise their presence and sometimes their status to other males. Calls are species-specific and females typically respond only to males of their own species. Males of most species possess a single or double vocal sac, which serves as a resonator and in at least some species as a sound radiator. Vocal sacs also conserve energy by allowing passive reinflation of the lungs as the vocal sac contracts after a call. Some male frogs, such as the Australian torrent frog (*Litoria nannotis*), lack a vocal sac but can still produce a surprisingly loud call. Environmental conditions such as temperature affect vocalizations. At colder temperatures, notes and pulses are produced at slower rates, but the length of the call increases. The dominant frequencies in the calls of most frogs are lower than 5000 Hz, although those of some small species are higher. Within species, variation in the dominant frequencies, pulse rates, and durations of calls often reflects male body size; therefore, the call may indicate male quality as well as male location. Females of many species use these characteristics to choose their mates from among competing males. Males vocalize from species-specific locations, which can be in water, on or beneath the ground, in vegetation from near ground level to high in trees, and even under water (several species including African clawed frogs, *Xenopus*).

Female frogs do not have a vocal sac and very few vocalize. Some female frogs produce a scream when distressed, and reproductively active females of some species call in response to male advertisement calls. Females of a few species call in other contexts, such as territoriality and during mating.

The posture of frogs during the fertilization of eggs is called amplexus; in most species this involves the male grasping the female from above (**Figure 1(c)**). The exact posture adopted depends on the morphology and relative size of the male and female. The two most common positions involve the male grasping the female in front of the back legs (inguinal amplexus) or the front legs (axillary amplexus). Amplexus is aided in many species by specialized patches of skin called

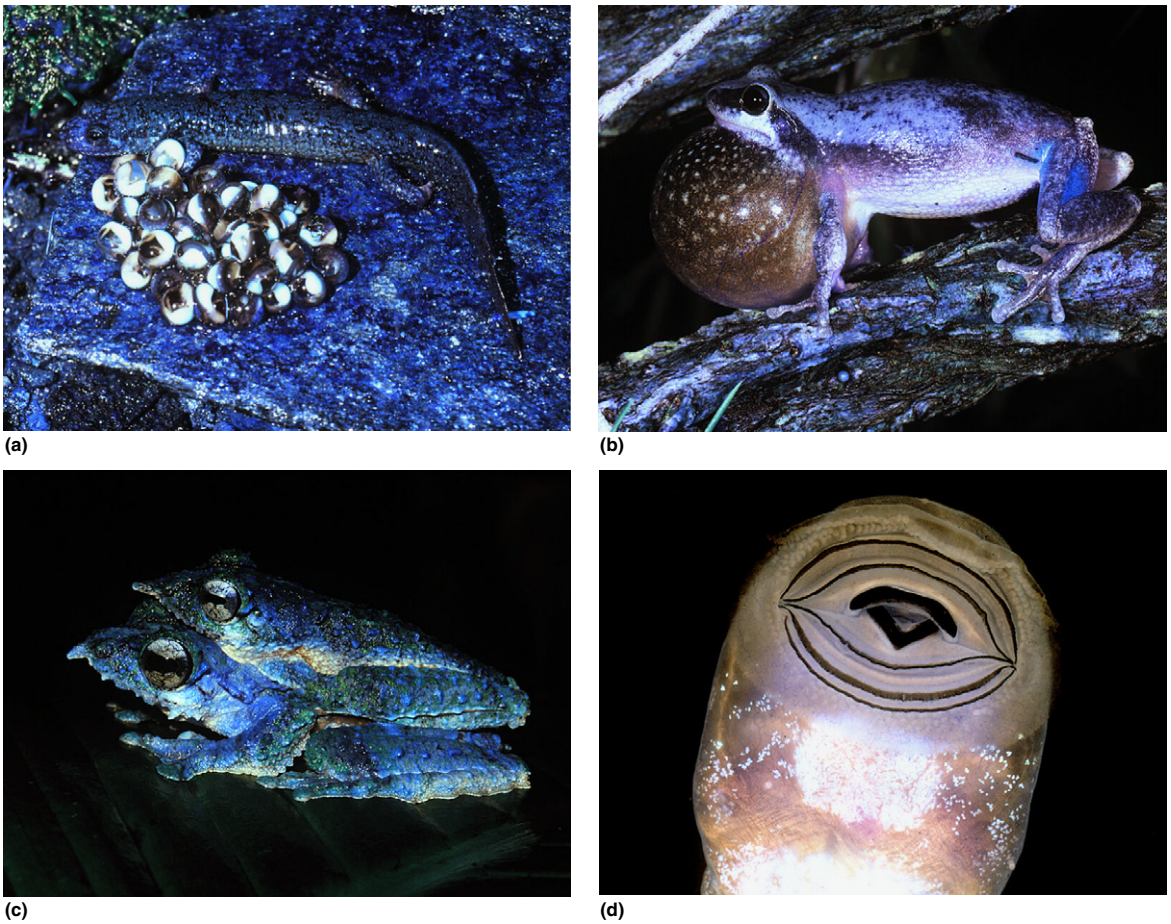


Figure 1 (a) A black-bellied salamander *Desmognathus quadromaculatus* guards its eggs (removed from (and replaced in) a small stream, Appalachian Mountains). (b) A male Australian desert treefrog, *Litoria rubella*, produces an ear-splitting call with the aid of its balloon-like vocal sac. (c) Spike-nosed frogs, *Litoria prora*, from New Guinea, in amplexus. The function of the rostral (nose) spine is unknown. (d) A suctorial tadpole of the Australian treefrog *Litoria nyakalensis*. The upper and lower jaw sheaths, surrounded by keratinized teeth and oral papillae set in an oral disc, are clearly visible. These tadpoles use their large ventral oral discs as suckers to attach to rocks in fast-flowing rain forest streams.

nuptial pads on the forelimbs of males. Pairs remain in amplexus whereas the male sheds sperm onto the eggs as they are released by the female. Fertilization in nearly all frogs is external but several species accomplish internal fertilization by cloacal apposition. The frog *Ascaphus truei*, commonly called the "tailed frog", breeds in fast-flowing streams of the Northwest of North America and carries out internal fertilization using the tail, which is actually an intromittent organ formed from an extension of the male's cloaca.

Ecology and Functional Morphology of Larval Amphibians

Caecilians

Relatively little is known about the larvae of caecilians. They are more similar to adults than are those of frogs or salamanders. Externally, they closely resemble adults but have gill slits and fins. Free-living caecilian larvae have external gills and a lateral line system. Their mouth and dentition resemble those of adults. They lack the tentacle organ that appears on the head of adults; this appears at metamorphosis.

Salamanders

Salamander larvae are much more similar to adults than are the tadpoles of frogs. Larval salamanders have external gills that are not completely covered by an operculum. Some embryonic salamanders have paired lateral projections from the head called balancers; in some species these persist for a short period following hatching. Most species possess well-developed fore- and hindlimbs through most of the larval period. Their bodies are laterally compressed compared to those of adults, and their tails are also relatively thinner and deeper. Their skins contain lateral line organs (neuromasts) and are thinner and less glandular than those of adults. Their dentition is different from adults, and their tongues are rudimentary. Their eyes lack lids. Larval salamanders are almost all carnivorous, usually feeding on zooplankton and larval insects. The larvae of larger species can also feed on small vertebrates. The larvae of some species have alternative morphologies; the

typical morphology is usually a planktivorous carnivore, but when conditions are favorable some individuals develop relatively larger heads and more powerful jaws, adopting a "cannibal" morphology that allows them to prey more effectively on small vertebrates, including conspecifics.

As in larval frogs, there is considerable variation among and within species in rates of larval growth and development. Both the minimum and the maximum development rates for salamanders are slower than those for frogs. Salamanders can take from 6 weeks to 5 years to complete larval development. Within species, rates respond to both temperature and food availability, and salamanders have the additional option, apparently not available to frogs, of changing the relative rates of development of somatic and reproductive structures so that they mature sexually without losing all larval characters.

Anurans

The tadpole larvae of frogs are highly specialized for growth and development in the aquatic environment (McDiarmid and Altig, 1999). They have an oval head–body region and a long tail, which is laterally compressed and includes a central area of musculature and dorsal and ventral fins of thin, lightly vascularized tissue. The tail is supported only by a notochord. Despite their very different body form, they swim and turn as rapidly and efficiently as fishes of similar body sizes. They have external gills, however, late in embryogenesis a fold of epithelium, the operculum, covers them, enclosing them and forming the opercular chamber. They filter feed using a mechanism that is very different from the oral morphology of adults. They pump water though the mouth and opercular chamber, where food particles are removed by branchial filters and entrapment in strands of mucus. Some tadpoles can remove particles as small as 0.126 μm from the water. Water is ejected through the nostrils of most species and through the spiracle, which is usually a single, tubular structure leading out of the opercular chamber that can be located midventrally or on the left side of the body. The unique mouthparts of tadpoles typically include an oral disc with transverse rows of keratinized teeth that are used to scrape particles into suspension. Keratinized sheaths on the jaws serve as cutting and biting surfaces. The oral apparatus is variously modified and sometimes allows attachment to the substrate via suction (**Figure 1(d)**). The relatively long, coiled intestine fills most of the body cavity. Tadpoles are typically thought of as microphagous herbivores that feed on algae and small parts of higher plants, but most species will feed on animal material when it is available. Tadpoles often scavenge on dead animals in the water and frequently prey on amphibian eggs. They typically feed at high rates, retain food in their guts for relatively short periods, and extract only a small fraction of the energy from ingested food. This makes them an important source of the fine particulate organic matter that many aquatic invertebrates depend on for food.

Tadpoles often hatch before the opercular chamber forms, and many species initially do not swim but attach to a substrate using adhesive organs located posterior to the mouth. Swimming usually commences with the closing of the operculum. Tadpoles lack limbs at hatching. The rear legs usually develop slowly, starting as limb buds at the posterior end of the body and developing over a long period. The forelimbs develop within the opercular chamber and are visible only after they erupt fully formed through the opercular wall at the onset of metamorphosis.

A typical pattern of tadpole growth and development would include one to a few days as a nonswimming hatchling with exposed gills followed by several weeks to months as a swimming and feeding larva. During this period, the tadpole grows dramatically; many species increase their body mass by hundreds of times and some by thousands. During this period, the hindlimbs begin to grow and slowly develop through the remainder of larval life.

Metamorphosis begins with the eruption of the forelimbs and involves drastic changes to all elements of the structure and function of the body. Reorganization of the mouth and digestive tract allows a switch from larval filter feeding to adult carnivory. The tail fins and musculature are broken down and reabsorbed. Most elements of the chondrocranium are reshaped and realigned, and the branchial apparatus ceases to be a support for gills and takes on a role as support for the adult tongue. The external and middle ears form, and in all but a few aquatic species the lateral line system disappears. The eye changes in size and structure, the photopigments of the rods change from porphyropsin to rhodopsin, and eyelids appear. The axial skeleton is reorganized and many elements are ossified. The lungs, which begin to develop and function during the larval period of many species, enlarge and take on their role as a major respiratory structure. The complexity of the skin increases, with the epidermis increasing from two layers of cells to five or six, many with specialized functions. The kidney, which in tadpoles is relatively simple and excretes excess water and ammonia, becomes more complex to serve its new function of conserving water and excreting urea. The gonads differentiate at about the time of metamorphosis.

Rates of growth and development of tadpoles are typically highly variable within species, responding to environmental temperature, food availability, and the density of tadpoles of their own and other species. Many species that inhabit unpredictable environments, such as temporary ponds, can have larval periods from 2 weeks or less up to months. Some species regularly spend 1 year or more (maximum 3 years) as larvae. The interaction between rates of growth and development in tadpoles has produced a rich literature that examines how and why this interaction is controlled. In general, it appears that rates of growth control rates of development during the earlier part of the free-living tadpole stage (Wilbur and Collins, 1973). Larvae that are growing slowly develop more slowly than fast growers, and changes in growth rate caused by changes in environmental conditions are mirrored by alterations in developmental rate. Fast-growing tadpoles tend to metamorphose both earlier and at larger sizes than slower growing individuals. Very slow growers appear to regulate their developmental rate so that they metamorphose near a species-specific minimum size after a larval period that may vary greatly in length. This flexibility in relative rates of growth and development is lost late in the larval period at a point that may vary among taxa. Theory suggests that the regulation of these rates ultimately responds to natural selection acting in a way that depends on the relative rates of growth and survival in aquatic and terrestrial environments. These ideas have been

applied to life-history transitions in many other taxa, including fishes, insects, and plants.

Because frogs typically deposit large numbers of eggs during relatively short breeding seasons, densities of tadpoles are often high. The success or failure of a cohort of tadpoles in a typical temporary freshwater habitat depends on a highly complex and unpredictable set of factors, including the density of tadpoles of their own species and other species, which control the degree of intraspecific and interspecific competition for food and space; the number and species of predators present; and the duration of the aquatic phase of the habitat. Competition and predation are both controlled by the choice of time and place of breeding by adult frogs and to some extent by microhabitat selection within habitats by tadpoles. The outcomes of species interactions involving tadpoles can be altered by changes in the timing of breeding, and microhabitat selection by tadpoles can depend on the species and sizes of other tadpoles present.

Tadpoles are preyed on by a wide variety of vertebrate and invertebrate predators, for whom they constitute a valuable resource. Major predators include fishes, salamanders and salamander larvae, and the aquatic larvae of insects such as dragonflies, damselflies, and beetles. Some species alter their color patterns or tail morphology in response to the presence of certain predators. Vulnerability of tadpoles to predators typically decreases as the tadpoles grow and develop, and many tadpoles exhibit short-term behavioral responses to predators, such as decreasing their levels of activity or switching from midwater feeding to substrate feeding, that appear to decrease their vulnerability. Most of these responses also decrease the tadpoles' rates of growth and development, leading to trade-offs that have been explored by behavioral ecologists.

Behavior and Ecology of Postlarval Amphibians

The ecological breadths of the three classes of modern amphibians are reflected in their geographical distributions (Savage, 1973). The caecilians are restricted by both thermal and water requirements to relatively low latitudes and elevations, and they do not occur where mean annual temperatures are less than about 12 °C or total annual precipitation is less than about 1000–2000 mm. The salamanders have less restrictive ecological requirements and are distributed across a broader range of habitats, occurring from low to high latitudes and elevations. Their major limitation is clearly moisture; they do not occur in areas that have prolonged dry seasons and only rarely in areas with total rainfall less than 1000 mm per annum. The anurans as a group can tolerate wide ranges of both temperature and water availability, and they occur at all but the highest latitudes and elevations and in all but the driest deserts.

Almost all adult amphibians are relatively nonselective carnivores that ingest invertebrate and vertebrate prey small enough to be swallowed whole. Amphibians are either sit-and-wait predators or active foragers, but do not engage in cursorial pursuit of prey. Sit-and-wait predators locate their prey primarily using vision, whereas active foragers often use olfaction.

Salamanders use a variety of methods for prey capture. Many larvae, and adults of aquatic taxa, are suction feeders, using rapid depression of the floor of the mouth and throat to pull in water and prey. Terrestrial salamanders usually feed by extending their large, fleshy tongues, which adhere to prey and pull it into their mouths. The tongue of salamanders is attached at the base and is protruded and elongated by muscles and fluid tension. The degree of attachment of the tongue and the length to which it is protruded vary among taxa. Captured prey are ingested whole. Terrestrial anurans also capture prey by protruding their tongues. These are usually attached near the front of the lower jaw and are protruded by literally flipping them forward and downward. Prey that adhere to the tongue are drawn into the mouth and swallowed whole.

Amphibians generally appear to grow throughout their lives, but rates of growth decline drastically after reproductive maturity is attained. In captivity, many species can live for decades, and even in nature some species live for extended periods. In general, amphibian life spans appear to be limited more by environmental hazards than by aging and senescence.

Many species of frogs and salamanders occupy relatively small home ranges during the nonbreeding season and migrate to breeding habitats for reproduction. The nonbreeding home range is aggressively defended by some species. Some species return to their natal ponds to breed, whereas others may simply migrate to a suitable body of water. In temperate regions, some species may also regularly migrate to overwintering sites. Amphibians use celestial navigation, light polarization, and the Earth's magnetic field for orientation during migrations. There is considerable movement among local populations in many species, particularly by anurans. Some of this movement is dispersal by juveniles, but some is due to longer range movements by terrestrial adults. At least one species of frog, *Bufo marinus*, can move great distances over relatively short periods. The range of the introduced Australian populations of this species has expanded by approximately 30 km per year for an extended period, and the animals that arrive first in new habitat are adults. Detailed studies of the local movements of adult *B. marinus* suggest that they are nomadic. Many mark-recapture studies of frogs have found high rates of disappearance of marked frogs and of the appearance of unmarked animals, suggesting that adults of many species may at least occasionally disperse to new areas.

The relatively high rates of migration between local populations found in many amphibian species suggest that groups of local populations often form metapopulations. It is important to recognize this because the dynamics of metapopulations are governed by different factors than those of local populations, and the conservation of metapopulations requires a different approach than the conservation of local populations.

Because amphibians typically produce relatively large numbers of relatively small eggs, their populations can increase rapidly in size when reproduction is successful. It is likely that populations of most amphibians normally fluctuate fairly widely over time (Alford and Richards, 1999). It may be common for local populations of many species to fluctuate to extinction, followed by relatively rapid recolonization by immigrants from adjacent local populations belonging to the same metapopulation. It is likely that for many species the persistence of most local populations has little bearing on whether the regional metapopulation will persist. However, there may be a few critical local populations that either serve as reliable sources of

colonists or serve as stepping stones for migration between more widely distributed local populations.

It has recently become recognized that in some, and possibly many, species, rates of movement among breeding sites are so high that what had initially been thought of as isolated populations, one at each breeding site, and subsequently as metapopulations, with some breeding sites acting as sources and some as sinks, are in fact single populations, with unified dynamics, and that the appearance and disappearance of species at breeding sites may reflect short-term changes in habitat selection, rather than colonization and extinction. Understanding the nature of populations of threatened species is a very important step in conserving them, and much more needs to be done before the population biology of most amphibians is well understood.

Amphibian Conservation

Human Uses of Amphibians

Amphibians have featured prominently in many human cultures through stories, song, and poetry. In urban areas they are frequently found coexisting successfully with humans in parks and garden ponds. By consuming large numbers of insects, they are important in controlling agricultural pests and reducing disease transmission. Frogs are an important source of protein in some subsistence cultures. In affluent countries, frogs are imported for consumption in gourmet restaurants. Hundreds of millions of frogs have been exported from Southeast Asia and the Indian subcontinent, resulting in increased insect pest populations. Frogs have also become model organisms in ecological, embryological, physiological, and genetic research.

Amphibian skin contains a wide variety of chemicals, including complex amines, alkaloids, and polypeptides, some of which have pharmacological properties. Many skin secretions are effective against amphibian bacterial and fungal infections. Secretions of the South American frog *Epipedobates tricolor* contains a constituent, epibatidine, that blocks pain 200 times more effectively than morphine. Poison dart frogs of the family Dendrobatidae harbor many exceptionally toxic skin compounds and one species, *Phyllobates terribilis*, contains sufficient toxin in a single frog to kill several adult humans. This toxin is smeared on darts used by the Choco Indians of western Colombia for hunting monkeys and other large game. There are connections between frog diets and the presence of skin toxins; some dendrobatid frogs with high toxin levels in the wild gradually lose their toxicity when held in captivity. Although indigenous cultures have recognized the toxic and medicinal properties of frog skin for centuries, their potential for development as medicines using scientific methods has only recently emerged as a significant field of research, which has already yielded many compounds that may be of great use in combating a variety of illnesses, including the HIV virus.

Amphibians as Components of Ecosystems

Amphibians form vital links in many food webs. They consititute the highest vertebrate biomass in some ecosystems and occupy an intermediate position in many food chains. Aquatic larval amphibians are herbivorous to omnivorous (most Anura) or carnivorous (Caudata and Apoda) and are significant prey items for a wide variety of vertebrate and invertebrate predators. Larval anurans often reach extremely high densities and can have a significant impact on nutrient cycling within freshwater habitats. Because tadpoles feed on algae and other aquatic material, they play an important role in the transfer of plant energy to predators of tadpoles. They also facilitate the breakdown of plant and animal biomass into plant nutrients; the extra nutrients regenerated by tadpoles can increase rates of primary production in freshwater systems, providing food at a higher rate to the tadpoles themselves. Adult amphibians feed on a wide variety of live food. Most are generalists that consume any living creature smaller than their gape size. However, some have specialized to feed exclusively on narrow ranges of food items, such as worms, ants, and even snails. Their effects on populations of their prey are probably important in many natural communities. Predators of amphibian eggs, tadpoles, and adults include other amphibians, spiders, insects, mammals, birds, and reptiles (especially snakes). Because of their important roles in ecosystems, population declines or extinctions of amphibians may have significant impacts beyond the amphibian species affected.

Amphibian Diversity and Threats to It

There are almost 7000 recognized amphibian species (**Table 1**), and many new species are being discovered and described every year. It has been estimated that the eventual number of species will be between 8000 and 10,000. Amphibians occupy all continents except Antarctica, and they are found in habitats ranging from arid deserts and saline mangrove swamps to tropical rain forests and mountain peaks more than 4000 m high. More than 80% of amphibians are found in the tropics, with an estimated 44% of the world's amphibians in the tropical Americas. Many tropical regions, including New Guinea, Southeast Asia, and parts of South America, have been inadequately surveyed, and there is no doubt that many new species will be discovered in these areas.

The conservation status of amphibians has deteriorated rapidly in recent years. The 1996 International Union for the Conservation of Nature (IUCN) *Red Data Book* listed five amphibian species as extinct and 124 as threatened, which represented 25% of the species for which assessment of conservation status had been undertaken at that time. Concern over the status of amphibians led the IUCN to conduct the Global Amphibian Assessment (GAA; Stuart *et al.*, 2004), which enlisted the aid of scientists throughout the world in determining the status of amphibians in their country or region. The results of this are frightening. In 2004, 1856 species, 32% of all known amphibians, were listed as threatened by the IUCN, with 427 listed as critically endangered. Thirty-four species were identified as being extinct, with nine having become extinct since 1980, and 122 as being possibly extinct, with 113 possible extinctions since 1980. In all, the GAA estimated that the conservation status of 435 species had become worse since 1980. All of these are probably

underestimates; although the GAA used the knowledge of more than 500 amphibian specialists, insufficient information was available on 1294 species (22.5%) to determine their conservation status, and many of the thousands of undescribed species that probably exist may also be under threat.

The Problem of Amphibian Population Declines

All amphibian populations fluctuate, and assessing the significance of downward trends in amphibian populations has been difficult. However, well-documented declines and disappearances of amphibian species and entire suites of species occurred over wide areas in the 1980s and 1990s, and continued in some places in the 2000s. These widespread losses of species and populations in a relatively short time frame, including dramatic extinctions of previously abundant species in protected areas such as national parks, are evidence that amphibians are experiencing a severe crisis.

Many causes have been postulated for amphibian declines, including habitat loss and modification, predation, environmental toxicity, disease, immunosuppression, ultraviolet radiation, changes in climate or weather patterns, and combinations of these. It is likely that many of these have caused some of the declines that have occurred (Alford and Richards, 1999; Alford, 2010). The GAA identified three major likely causes of declines in the 435 species whose conservation status had worsened since 1980. Overexploitation by humans appeared to be affecting 12% of these species, habitat loss 42%, and "enigmatic decline," which was defined as declines occurring where there is suitable habitat and no other identifiable cause, was suggested as affecting the remaining 46%. Many of the enigmatic declines have involved rapid crashes of populations in relatively pristine habitats, such as montane rain forests. Lips *et al.* (2006) monitored a species-rich community in Panama before and during such a crash, and determined that it was caused by an outbreak of chytridiomycosis, caused by the amphibian chytrid skin fungus, *Batrachochytrium dendrobatidis*. This recently discovered disease is the proximate cause of many enigmatic amphibian declines. It infects the outer layer of amphibian skin via motile zoospores, which can be transmitted through contact or shared water. The host–pathogen ecology of this disease is complex; its prevalence and virulence are strongly affected by environmental temperature and humidity, and the timing of outbreaks of chytridiomycosis is probably affected by weather, and perhaps by changes in weather patterns related to global climate change. It is a very unusual disease in that it infects a very wide range of host species and can drive many of them to extinction, because resistant alternative hosts are common and it may also persist in the environment, independent of amphibian hosts.

In every area where there have been outbreaks of chytridiomycosis, some species have been driven to local extinction or very low numbers whereas others have persisted with endemic infections of the pathogen. The resistance of species to the development of disease is likely to be controlled by many factors, including the behavior and microhabitat selection of hosts, host population density, the types and quantities of antimicrobial skin peptides secreted by hosts, and the composition of the natural assemblage of microbiota inhabiting the skin, some of which can compete with and inhibit the growth of *B. dendrobatidis*. In some places, including tropical Australia and some locations in Central America, some species have reappeared at some sites after initially disappearing. Research is underway to determine whether these reappearances result from increased disease resistance in amphibians, decreased virulence of the pathogen, effects of short-term variation in weather patterns, combinations of these, or alternative mechanisms.

It is likely that the pathogen has recently increased its geographic range, and it may even have spread globally from a single place of origin, possibly aided by human activities such as the extensive trade in amphibians for food, medical and scientific purposes, and as pets. Although genetic studies favor the hypothesis of a relatively recent spread and origin for most strains of the pathogen, the evidence is not conclusive. For example, in the California Sierras, there appears to be fine-scale genetic structure, suggesting that the pathogen may have been present for some time, despite the fact that declines linked to Bd have only recently been noted in the area (Morgan *et al.*, 2007). As more, and more thorough, surveys for its presence have been conducted, it has become clear that it is present in all continents occupied by amphibians, but appears to be absent from some areas, such as parts of far northern Australia, some parts of continental Asia, and the island of Madagascar. Some quarantine measures have been instituted, and more are needed. Knowledge of the status and distribution of the disease should increase since chytridiomycosis has been listed as a notifiable disease by the world organization for Animal Health (OIE).

At present, the only practical approaches to managing species that are severely threatened by chytridiomycosis are quarantine and the establishment of chytridiomycosis-free colonies of captive animals. Neither of these are likely to prove practical in the very long term, since the pathogen is clearly highly dispersive and captive management is expensive in resources and vulnerable to disruption. Approaches that have been proposed and are in various stages of development include probiotic manipulations of amphibian skin microbiota, artificial selection for increased resistance in captive populations, with resistant animals eventually reintroduced to the wild, and translocations of animals from populations that have developed naturally increased resistance.

A major hindrance to our understanding of amphibian population declines is a general lack of information about amphibian autecology and host–pathogen interactions, about how amphibian populations normally fluctuate, and about the extent to which they operate as metapopulations (Heyer *et al.*, 1994). In the short term, a great deal of research aimed at increasing our knowledge in these areas and testing hypotheses about population declines will be vital for identifying causal factors and determining how to eliminate or mitigate them.

Conclusion: The Amphibian Success Story and Its Future

Contrary to popular opinion, modern amphibians are not a relictual remnant of the ancestors of other tetrapods but are a highly successful group in their own right. There are more

species of amphibians than there are of mammals. Modern amphibians occur in nearly all of Earth's terrestrial habitats, from within the Arctic Circle to tropical deserts. Groups of modern amphibians that need to conserve water have evolved impermeable skins, cocoons, and the ability to excrete uric acid. Groups that need to breed outside water have evolved a startling array of reproductive adaptations; amphibians have the widest range of reproductive modes of any tetrapods. These include aquatic eggs and larvae, many species with terrestrial eggs, and truly viviparous species in which the mother provides nutrition in addition to the yolk during development. This diversity indicates that the typical reproductive pattern, with aquatic eggs and larvae, must not represent a constraint that has limited their success. It probably represents a successful adaptation that allows the exploitation of aquatic habitats by terrestrial species and allows a much higher fecundity than is available to species that must provision their eggs with enough yolk for complete development.

The ability to respond to environmental challenges has allowed the modern amphibians to persist and flourish during and through periods of dominance of terrestrial habitats by other tetrapod groups. They have outlived the early dominant amphibians, several waves of dominance by reptiles including the dinosaurs, and many radiations and extinctions of mammals. The current declines and disappearances of many amphibians may be an early manifestation of a general crisis in biodiversity, and unless they are halted it seems likely that future generations will be denied a full appreciation of the remarkable extent of amphibian diversity. It also seems likely, though, that as long as terrestrial habitats continue to accommodate vertebrate life, amphibians will persist.

List of Courses

1. Herpetology
2. Vertebrate Zoology
3. Conservation Biology

See also: Birds, Biodiversity of. Fishes, Biodiversity of. Mammals, Biodiversity of. Reptiles, Biodiversity of

References

Alford RA (2010) Declines and the global status of amphibians. In: Sparling DW, Lindner G, Bishop CA, and Krest SK (eds.) *Ecotoxicology of Amphibians and Reptiles*, 2nd edn., pp. 13–46. USA: SETAC press.

Alford RA and Richards SJ (1999) Global amphibian declines: A problem in applied ecology. *Annual Review of Ecology and Systematics* 30: 133–165.

AmphibiaWeb: Information on amphibian biology and conservation. (web application) (2011) Berkeley, California: AmphibiaWeb. Available from: http://amphibiaweb.org/ (accessed 19 March 2011).

Duellman WE and Trueb L (1986) *Biology of Amphibians*. Baltimore: Johns Hopkins University Press.

Frost DR, Grant T, Faivovich J, *et al.* (2006) The amphibian tree of life. *Bulletin of the American Museum of Natural History* 370.

Heyer WR, Donnelly MA, McDiarmid RW, Hayek LC, and Foster MS (1994) *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*. Washington, DC: Smithsonian Institution Press.

Lips KR, Brem F, Brenes R, *et al.* (2006) Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. *Proceedings of the National Academy of Sciences, USA* 103: 3165–3170.

McDiarmid RW and Altig R (1999) *Tadpoles: The Biology of Anuran Larvae*. Chicago: University of Chicago Press.

Morgan JAT, Vredenburg VT, Rachowicz LJ, *et al.* (2007) Population genetics of the frog-killing fungus *Batrachochytrium dendrobatidis*. *Proceedings of the National Academy of Sciences, USA* 104: 13845–13850.

Pough FH, Andrews RM, Cadle JE, Crump ML, Savitsky AH, and Wells KD (2004) *Herpetology*, 3rd edn. Engle-wood Cliffs, NJ: Prentice Hall.

Savage JM (1973) The Geographic Distribution of Frogs Patterns and Predictions. In: Vial JL (ed.) *Evolutionary Biology of the Anurans Contemporary Research on Major Problems*, pp. 315–445. Columbia: University of Missouri Press.

Stuart SN, Chanson JS, Cox NA, *et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 178–1786.

Wilbur HM and Collins JP (1973) Ecological aspects of amphibian metamorphosis. *Science* 182: 1305–1314.

Antarctic Ecosystems

Peter Convey, British Antarctic Survey, Cambridge, UK

Published by Elsevier Inc.

Glossary

Ablation Direct transfer of water molecules from ice crystals to the vapor phase, without transition through the liquid phase.
Alien Nonindigenous species whose introduction and transient or established presence at an Antarctic location is known or considered most likely to be linked with human activity.
Anhydrobiosis Ability of certain organisms to lose virtually all water from their bodies under certain conditions, yet remain viable when subsequently rehydrated.
Cryoturbation Process by which soil particles and stones are moved and mixed by the frequent formation and subsequent melting of ice crystals in the soil column.
Cryptogam Plants without apparent reproductive organs; reproduction by spores.
Fellfield Dry, windswept habitat of cold regions comprising mineral soil, gravels, stones, and rock, dominated by cryptogams (especially mosses and lichens) and, outside the Antarctic continent, by compact cushion-forming phanerogams and short grasses.
Microclimate Climate experienced over a scale of mm to cm, within which the majority of Antarctic terrestrial biota (invertebrates, plants, and microbes) exist.
Microhabitat Small-scale habitat within which an organism or community exists.
Nunatak Isolated ice-free summit of a mountain whose bulk is permanently below the surface level of an ice sheet.
Phanerogam Plants with conspicuous reproductive organs (flowers); reproduction by seeds.
Phylogeography Studies concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially within and among closely related species.
Propagule Any part of an organism capable of developing into a new individual (e.g., seed, spore, gemma, lichen soredium, egg, shoot fragment, etc.).
Sub-nivean Environment beneath a temporary, seasonal, or permanent covering of snow.
Synanthropic Referring to a species whose presence is intimately linked with human activity (i.e., only occurring in occupied buildings etc.).

Antarctica

Introduction to the Continent

Antarctica is distinct amongst the Earth's continents, not least because it is the only one not to have a long-term history of human occupation and influence. Antarctica is isolated from other continents by the 1000 km Drake Passage, and by 4000–5000 km from Australia and South Africa (**Figure 1**). More than 99.6% of its area is covered permanently by snow or ice, and ice-free ecosystems are limited to exposed nunataks, cliffs, and seasonally snow and ice-free areas. Terrestrial habitats are best represented in coastal regions, where most of the biodiversity is found, particularly along the Antarctic Peninsula, and in the major mountain ranges inland. Most ice-free areas are of small extent and isolated from others by distances of tens to hundreds of kilometers. However, in the McMurdo Dry Valleys of southern Victoria Land, exposed areas of several hundred square kilometers exist. Large areas of the continent are classified as cold deserts, with extremely low precipitation rates (much of which is then lost by ablation).

In contrast with the Arctic, which includes the northern fringes of continents surrounding a generally shallow ocean, Antarctica is an ice-bound continental mass surrounded and isolated by a large extent of cold ocean. This simple geographical difference is sufficient to create the large climatic differences found between the two polar regions at any specific latitude, and leads to a fundamental difference in colonization patterns. With continuous southward land connections, recolonization of Arctic regions following post-Pleistocene deglaciation was relatively simple. In contrast, the extreme isolation of terrestrial and freshwater habitats across Antarctica from plausible refugia or other continental sources of colonists, combined with the obliteration of most but not all habitats during glaciation, have led to much lower colonization rates and levels of diversity for many faunal, floral, and microbial groups.

Tectonic, Glacial, and Biological History

Antarctica was an integral element of Gondwana. As these elements separated, its last continental links were with Australia and South America. The connection with South America, via what was to become the Antarctic Peninsula, was lost approximately 30–35 million years ago (Livermore *et al.*, 2005). At this time, fossil evidence indicates that Antarctica possessed a cool temperate fauna and flora. The separation of Antarctica and South America allowed development of circumpolar ocean currents, eventually isolating Antarctica from lower latitude sources of heat energy, and accelerating the continent's gradual cooling and formation of continental-scale ice sheets. The extent of glaciation has varied widely both geographically and over time during the last *c.* 40 my (Tripati *et al.*, 2005). Tundra-like ecosystems persisted in Antarctica through this period until at least 12–13 Ma (Lewis *et al.*, 2008) and possibly more recently (see discussion in Convey *et al.*, 2008).

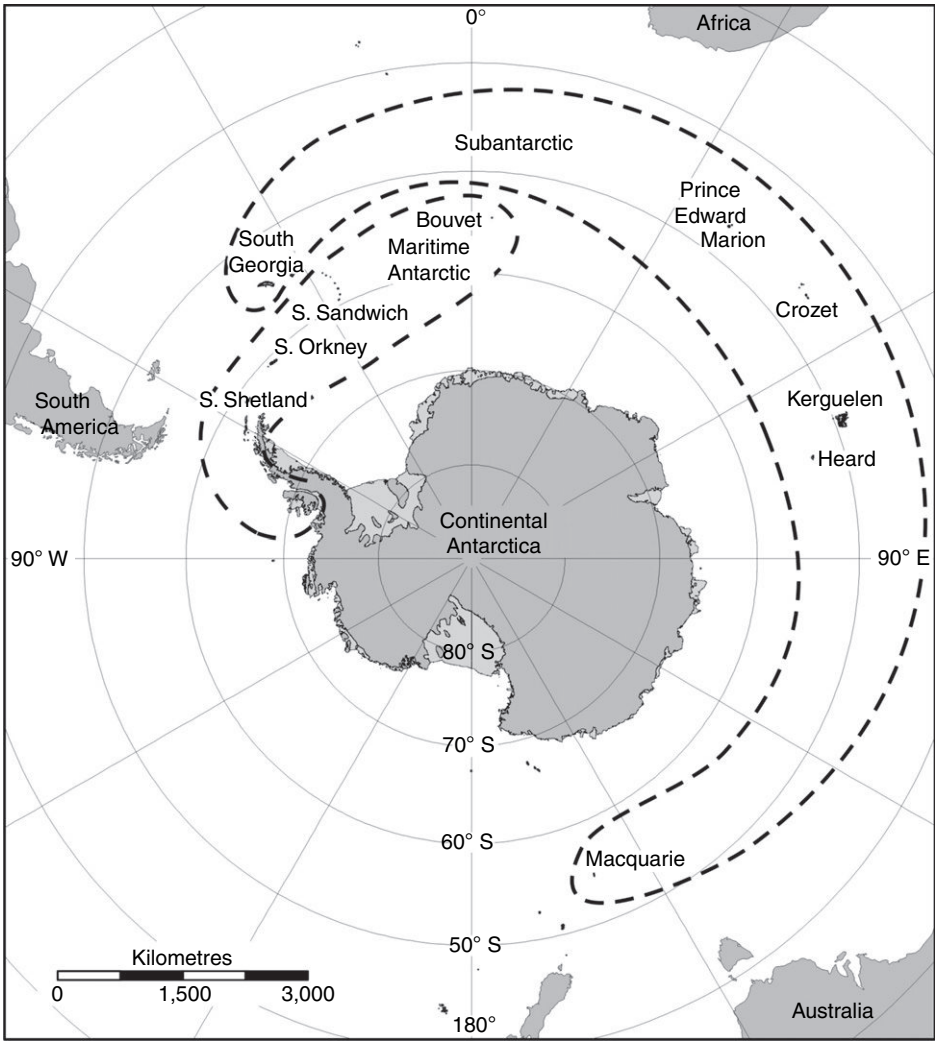


Figure 1 Antarctica's geographical relationship with the other Southern continents, and the three commonly described terrestrial biogographic zones within the Antarctic region.

Pleistocene glaciation mirrored that of the Northern Hemisphere, with the continent spending the majority of time over at least the last 800,000 years in the glacial phase of glacial–interglacial cycles (EPICA, 2004). At maximum, the continental icecap was considerably deeper than now, and the continent was surrounded by extensive ice shelves. The island groups of the South Shetlands, South Orkneys, and South Georgia are similarly thought to have been the centers of large ice caps, with current glaciological models suggesting that terrestrial habitats would have been limited to isolated and higher altitude inland nunataks. However, there is increasing recognition that reconciling the next generation of finer resolution glaciological models with biogeographical and phylogeographical evidence supporting ancient and relictual distributions is a pressing challenge (Convey and Stevens, 2007; Convey *et al.*, 2008, 2009; Vyverman *et al.*, 2010).

Antarctic Terrestrial Biogeographic Zones

Most authors have long recognized three biogeographic zones: the sub-, maritime, and continental (or frigid) Antarctic (**Figure 1**), whose ecosystems are distinctively different. This formulation is retained in the descriptions below, for ease of comparison with the existing literature. However, recent plant biogeographic analyses (Peat *et al.*, 2007) refined the boundaries and components of the maritime Antarctic, whereas Chown and Convey (2007) highlighted the biogeographic strength of the division between the southern Antarctic Peninsula and the remainder of the continent, defining the "Gressitt Line," which is analogous to the Wallace Line of southeast Asia.

Sub-Antarctic

The core sub-Antarctic islands include a ring of isolated Southern Ocean islands and Archipelagos. Except South Georgia, Heard, and McDonald Islands, the majority of these (Marion and Macquarie Islands, and the Crozet and Kerguelen archipelagos) lie either close to or north of the Polar Frontal Zone, a circumpolar oceanic feature where cold Antarctic surface waters sink below warmer sub-Antarctic waters (Selkirk, 2007). Some authors include further groups (Falkland

Table 1 Typical air and microhabitat temperature ranges experienced in summer and winter in each of the Antarctic biogeographical zones

Antarctic zone	*Months with positive mean air temperatures*	*Air temperature range (°C)*		*Microhabitat temperature range (°C) (exposed habitat minima in parentheses)*	
		Summer	*Winter*	*Summer*	*Winter*
Sub-Maritime	6–12	−2–+25	−10–+15	−5–+30	> −2–(−10)
	1–4	−10–+15	−45–+5	−5–+30	−20–+3 (−45)
Continental coastal	0–1	−20–+10	−40–−5	< −10–+30	(−40)
Inland (deserts and nunataks)	0	−31–+9	−48–−12	−29–+17	−47–−16 (< −60)

Note that few data are available from microhabitats in the continental Antarctic, particularly during winter; considerably more negative minima would be expected in sites without snow cover.

Islands, Gough Island, Îles Amsterdam and St Paul, New Zealand's outlying groups of Snares, Campbell, Bounty, and Chatham) either within the sub-Antarctic or in a separate "mild Antarctic" classification. More correctly these are southern cool temperate or ocean temperate, containing faunal and floral groups not otherwise represented in the sub-Antarctic. Sub-Antarctic island climates are strongly oceanic and, with the occasional exception of South Georgia, they are not influenced by seasonal pack or fast ice. Mean air temperatures of most islands are low but positive year-round, and precipitation is high (**Table 1**). With the exception of South Georgia and Heard Island, most islands are now or have been only partially glaciated.

Maritime Antarctic

The maritime Antarctic also experiences a considerable but more seasonal oceanic influence on climate. The zone includes the western coastal regions of the Antarctic Peninsula to Alexander Island (*ca.* 72° S), along with the Scotia arc island archipelagos (South Shetland, South Orkney, and South Sandwich Islands), and the isolated Bouvetøya and Peter I Øya. Mean air temperatures are marginally positive for 1–4 months of the year; however, both summer maxima and winter minima (**Table 1**) are buffered by the surrounding ocean, the surface temperature of which varies annually between *ca.* −2 °C and +1 °C. Short thaws are possible in all winter months. Precipitation is generally high but varies with location and time. The South Sandwich archipelago and Bouvetøya are unusual in being geologically recent (1–3 my old) and active volcanic islands, with unique biological communities associated with geothermal activity (Convey *et al.*, 2000; Convey and Smith, 2006). Together with analogous areas on Deception Island (South Shetland Islands) (Smith, 2005) and in Victoria Land (continental Antarctic) (Skotnicki *et al.*, 2001), these are exceptional in Antarctica.

Continental Antarctic

This zone conventionally comprises most of the continental area of Antarctica, including all of East (or Greater) Antarctica, the Balleny Islands, and the eastern side of the Antarctic Peninsula. However the inclusion of the latter element, other than on climatic grounds, is put into doubt by recent analyses (Peat *et al.*, 2007; Chown and Convey, 2007). With the exception of the Dry Valley region of Victoria Land, terrestrial habitats are of limited extent and great isolation. In addition to the cold deserts, they include exposed coastal regions similar to those of the maritime Antarctic and inland nunataks. Positive mean air temperatures are achieved for <1 month in coastal locations and never inland (**Table 1**). Air temperatures rarely if ever become positive even for short periods in summer. As elsewhere, microhabitat temperatures do not track air temperatures closely, due to warming from solar radiation in summer, and protection from winter extremes by snow cover. Beyond the Antarctic Circle, habitats inevitably experience winter darkness with no thermal input, during which those that are unprotected (e.g., exposed mountain ridges, cold deserts) are expected to track air temperature patterns and minima closely.

Antarctic Terrestrial Biota

Habitats

In the maritime and continental zones, with the exceptions of steep cliffs and exposed mountain ridges, most terrestrial habitats are covered seasonally by snow or ice, protecting from extreme temperature variation and wind abrasion. Although habitats in the maritime and continental zones may be free of seasonal snow cover, if at all, for periods of only days or weeks up to *c.* 5 months, many sub-Antarctic islands experience only intermittent snow cover, often restricted to higher altitudes. Even on the coldest islands sub-nivean microhabitat temperatures are often sufficient to allow year-round biological activity. This is not the case in the maritime and continental zones, where biological processes are arrested by low winter temperatures.

Antarctic soils are typically poorly developed, with low organic content (Beyer and Bölter, 2002). Formation, development, and stability of soils are heavily influenced by cryoturbation. There is a clear dichotomy between those of the sub-Antarctic and the other two zones, with only the latter possessing a widespread permafrost layer. Brown soils are associated only with the larger stands of higher plants (phanerogams) in the maritime Antarctic, are not present in the continental zone, but are more widespread in the sub-Antarctic. Deep peat deposits have developed under extensive valley bog communities in the sub-Antarctic, and are radiocarbon dated as forming soon after the end of the Pleistocene glaciation (10–20,000 years b.p.). Peat deposits are much

Table 2 Dominant biotic components of typical ecosystems of each of the Antarctic biogeographical zones

Zone	*Flora*	*Fauna*	*Microbiota*
Sub	Phanerogams, cryptogams; open fellfield (*cf.* maritime zone) at higher altitude; alien species	Arthropods (esp. Diptera, Coleoptera); micro-arthropods (Acari, Collembola); micro-invertebrates (Nematoda, Tardigrada, Rotifera); alien vertebrates and invertebrates	Various groups
Maritime	Closed and open cryptogamic communities; phanerogams very limited	Micro-arthropods; micro-invertebrates; Diptera very limited	algal and cyanobacterial mats; foliose algae; protozoa
Continental	Cryptogams, very limited extent	Micro-invertebrates; micro-arthropods more limited	Algal and cyanobacterial mats; protozoa; endolithic microbiota

more restricted in the maritime zone (maximum age 5–6000 years b.p.) and not found in the continental Antarctic.

Terrestrial Fauna

Native terrestrial vertebrates, limited to a single endemic insectivorous passerine and freshwater ducks on South Georgia and Kerguelen, and two scavenging sheathbills, form a very small element of the fauna of Antarctica. There are many records of vagrant birds, particularly from sub-Antarctic islands, the majority of which succumb to the extreme conditions or predators. There are no native mammals, reptiles or amphibians. Terrestrial habitats are influenced (guano deposition, aerosol dispersal, trampling) by the abundant Antarctic marine vertebrates (penguins, seabirds, seals), which come ashore to breed, rest or molt. Most impacts are coastal, but birds such as Antarctic and snow petrels and south polar skuas also breed on nunataks several hundred km inland.

The terrestrial faunas of each Antarctic zone, therefore, consist of invertebrates (**Table 2**). The sub-Antarctic islands include "higher" insects and other arthropods, although many familiar groups are not present (e.g., Odonata, Trichoptera) or are represented by very few species (e.g., Aranaea, Isopoda, Lepidoptera, Hymenoptera and Hemiptera). The most diverse insect groups are Diptera and Coleoptera. Molluscs and annelid worms are also present along with communities of microarthropods (Acari, Collembola) and microinvertebrates (Nematoda, Tardigrada and Rotifera). These have received less critical attention than the larger and more obvious groups, as is also the case with the (polyphyletic) Protozoa.

Maritime Antarctic faunas include fewer higher taxa, with only two chironomid midges (Diptera) present. The most studied faunal elements are a minority of the species of Acari and Collembola, although microinvertebrates are also well represented. Species diversity in all groups is low, but population densities are often high, in the range 10^4–10^7 individuals m^{-2}, comparable with or greater than many temperate ecosystems. Food webs have a simple structure, with very low representation of predators. Microinvertebrates include genera with characteristic trophic preferences (e.g., algivory, bacterivory, fungivory, and nematophagy); however, little detailed autecological work has been attempted and the diets of Antarctic taxa are virtually unknown (Hogg *et al.*, 2006).

The fauna of the continental Antarctic gives a further simplification, including no insects, and with microarthropods only reaching levels seen in the maritime zone in the much more limited areas of vegetation and nutrient enrichment. This zone includes the simplest faunal ecosystems on the planet, where even nematodes are absent (Convey and McInnes, 2005).

Terrestrial Flora

Large-scale patterns in the Antarctic flora mirror those of the fauna (**Table 2**). Smith (in Laws, 1984) provides the most authoritative ecological description and comparative treatment of Antarctic floras, which remains valid despite many advances in detailed taxonomic knowledge. Monographs are available of the Antarctic lichen (Øvstedal and Smith, 2001) and bryophyte (Ochyra *et al.*, 2008) floras.

The flora of the sub-Antarctic is the richest and most diverse, and shows affinities with either South America (Fuegia) or Australasia (New Zealand). Trees and shrubs are absent (excepting the dwarf shrub *Coprosina* on Macquarie Island). Sub-Antarctic vegetation is superficially similar to, and climate, land form and pedology comparable with those of Arctic tundra, but the two regions differ in several important respects. Sub-Antarctic soils do not contain a permafrost layer, whereas the only woody plants are dwarf-shrub-like rhizomatous herbs (*Acaena*). The vegetation includes exceptional development of large rosette-forming "megaherbs" and tall grasses, encouraged by the lack of native grazers. Phanerogams are best developed in low altitude coastal regions. With increasing altitude, they are replaced by a cryptogamic fellfield flora, closely similar to that of the maritime Antarctic.

The most favorable habitats of the maritime Antarctic are dominated by cryptogamic communities of carpet- and turf-forming mosses. Extensive vegetated areas are limited to a narrow altitudinal range (to *c.* 150 m.a.s.l.) in coastal regions. Inland, at higher altitudes, and at more exposed coastal sites, open fellfield communities occur, including cushion and turf-forming mosses, liverworts and crustose, foliose, and fruticose lichens. Phanerogams are represented by the hairgrass, *Deschampsia antarctica*, and pearlwort, *Colobanthus quitensis*, which also occur in the sub-Antarctic and the South American Andes.

Cryptogamic vegetation is fragile and, lacking roots, sensitive to physical disturbance, for example in the vicinity of penguin and seal colonies. Such areas tend to be dominated by the foliose alga, *Prasiola crispa*. One consequence of the rapid recovery of Antarctic fur seal (*Arctocephalus gazella*) populations, following near extinction through hunting, has been the occupation of new summer resting and molting sites within terrestrial habitats of the sub and maritime Antarctic. The plant communities of these areas are unable to withstand excessive seal trampling and nutrient enrichment and have largely been destroyed.

Continental Antarctic vegetation is categorized within the same system as that of the maritime zone, although lichen sub-formations predominate. No phanerogams are present, the extent of closed cryptogamic communities is much more limited and peat formations are absent.

Microbial Systems

Microbial autotrophs form the basis of polar terrestrial ecosystem processes (Vincent, 1988; Friedmann, 1993; Wynn-Williams, 1996), playing pivotal roles in the processes of primary colonization and stabilization of mineral soils, which allow secondary colonization and succession by other microbiota, plants and Metazoa. Autotrophic cyanobacteria and algae are the primary colonists, followed secondarily by bacteria, fungi, and protozoans (whose significant Antarctic members are heterotrophic flagellates, gymnamoebae, testate amoebae and ciliates).

In addition to the largely edaphic (on or within soils), epiphytic (on surfaces of living plants and lichens) and epilithic/hypothithic (on exposed or undersurfaces of rocks) habitats typically also occupied by faunal and floral communities, microbial ecosystems may also utilize cryophilic (between ice crystals) and endolithic (within surface few mm of rock matrix) habitats. The latter habitat is further divided into chasmoendolithic (within fissures and cracks open to the rock surface) and cryptoendolithic (within tiny cavities of the rock matrix). These cryptic habitats represent one limit to life on Earth and have been proposed as models assisting development of exobiological techniques (Wynn-Williams, 1996).

Microbial ecosystems have received most attention in more extreme terrestrial habitats, although the same groups are present and important in all three Antarctic zones. Groups such as algae and cyanobacteria form filaments and mats within water bodies and on/in the surface layers of damp soils. They are well represented within the maritime zone, and are often a climax community of the continental zone (**Table 2**). In large parts of the continental Antarctic Dry Valleys no detectable life survives on the surface of soils or rock.

Freshwater Systems

Freshwater ecosystems of the Antarctic possess a very simple structure when compared with those of the Arctic and lower latitudes. Perhaps most striking is the absence of fish. Their faunas include a single predatory diving beetle (only South Georgia), several Crustacea (cladocerans, copepods, ostracods, and anostracans), adventitious micro arthropods, nematodes, tardigrades, rotifers, and protozoans. Ecosystem structure in the maritime Antarctic is even more simplified, the plankton containing single copepod (*Boeckella poppei*) and anostracan (*Branchinecta gaini*) herbivores, and the predatory copepod *Parabroteas sarsi* in addition to benthic cladocerans and ostracods and the microscopic groups. *B. poppei* provides an intriguing link between the maritime and continental Antarctic, which otherwise do not share any crustacean species.

The trophic impact of metazoans in lakes of all three zones is thought to be minimal, although data are lacking and there is disagreement over their importance in the sub-Antarctic lakes of South Georgia. With this exception, top–down grazing control is reduced, and the "microbial loop" - a microbial food web consisting of phytoplankton, bacteria and protozoans – assumes importance (Laybourn-Parry, in Lyons *et al.*, 1997).

Lakes in the maritime and continental zones are seasonally or permanently covered by ice. Some continental lakes are meromictic or hypersaline. Some are thought to be at least hundreds of thousands of years old. Lake Vostok, a 200 km long, 500 m deep lake beneath at least 3 km thickness of the continental icecap, is the largest of many subglacial lakes across Antarctica, with the possibility of harboring microorganisms effectively isolated since the formation of a permanent ice barrier.

Present Day Biodiversity

Within Antarctica

Attempts to catalog and compare biodiversity of the three Antarctic biogeographical zones continue to be hindered by two fundamental problems – a lack of adequate or comparable sampling coverage, and taxonomic uncertainty (particularly the likelihood of extensive synonymy), both of which apply in varying extent to all the groups encountered (Chown and Convey, 2007). Currently, sufficient data do not exist to allow rigorous comparisons for any microbial groups. Worldwide as well as in the Antarctic, a tiny proportion of prokaryotes have been identified, with virtually no mention of species distribution. However, increasingly, the application of molecular biological techniques is indicating both a higher than previously realised level of microbial diversity in Antarctica (Cowan *et al.*, 2002), and that much of this diversity shows large levels of divergence (implying long-term isolation) from sequences available in current molecular databases (De Wever *et al.*, 2009; Vyverman *et al.*, 2010). This suggests that the "global ubiquity hypothesis," which predicts that microbial groups do not face the same obstacles in dispersal as larger organisms, does not apply to Antarctic systems to the same extent as it does elsewhere.

The biodiversity of groups for which there are reasonable data in each of the Antarctic zones is summarized in **Table 3** and **Table 4**. These data illustrate a large-scale trend of decreasing diversity with increasing environmental extremes from the sub to the continental Antarctic. However, although there have been several proposals of similar trends occurring within zones (particularly along the Antarctic Peninsula and in Victoria Land), the true picture appears to be much more

Table 3 Biodiversity of plant taxa in the three Antarctic biogeographical zones

Zone	*Flowering plants*	*Ferns and club-mosses*	*Mosses*	*Liverworts*	*Lichens*	*Macro-fungi*
sub-Antarctic	60	16	250	85	250	70
maritime Antarctic	2	0	100	25	250	30
continental Antarctic	0	0	25	1	150	0

Note that figures presented are approximate, as it is likely that (i) new species records will be obtained through more directed sampling, (ii) a significant number of unrecognized synonymies are likely to exist and (iii) taxonomic knowledge of some Antarctic groups is incomplete. The Scientific Committee on Antarctic Research, through its biological program 'Evolution and Biodiversity in Antarctica', maintains up-to-date and open access databases of known biodiversity in most terrestrial faunal and floral groups (see www.eba.aq).

complex (Adams *et al.*, 2006; Maslen and Convey, 2006; Peat *et al.*, 2007), in part related to the presence of suitable microhabitats, and in part to glacial history and the hypothesized existence of refuge sites and biodiversity hotspots.

Origin and Antiquity of Biota

The biota of Antarctica is most closely related to that of other southern continents. There are also examples of plant and microbial groups, with known ability to disperse *via* aerial propagules, with bipolar or alpine distributions (Longton, 1988; Broady, in Wynn-Williams, 1996; Øvstedal and Smith, 2001; Pearce *et al.*, 2009). Gross similarities between soil faunal communities of both polar regions (the dominance of Acari, Collembola and other microinvertebrates) indicate the generalist function of these groups in soil ecosystems worldwide.

Except in the sub-Antarctic, where some islands experienced incomplete glaciation, many coastal terrestrial habitats are likely to have been obliterated during glaciation. As a large proportion of extant biota in all three zones is found in low-lying coastal areas, it is unlikely that they survived glacial maxima in higher altitude nunatak sites where, even now, they are not represented. However, at least among the Acari, Collembola, Tardigrada, Nematoda, and the lichens, many species and some genera are endemic to one or more of the Antarctic zones, supporting a prolonged regional presence (Øvstedal and Smith, 2001; Pugh and Convey, 2008). The wide distributions of some representatives of these groups within zones must be as a result of postglacial expansion from as yet unidentified refugia. In contrast, mosses show much lower levels of endemism (Ochyra *et al.*, 2008), suggesting that the contemporary flora may have developed mainly through postglacial recolonization. The use of molecular biological approaches to estimate divergence dates clearly identifies examples of regional presence over millions of years in the continental, maritime and sub-Antarctic (Stevens *et al.*, 2006; Mortimer *et al.*, 2010), as well as separation events estimated to have taken place at least 30–40 Ma (Allegrucci *et al.*, 2006), some of which may be linked to the geological split of the Antarctic Peninsula from South America (see Convey and Stevens, 2007; Convey *et al.*, 2008, 2009a; Vyverman *et al.*, 2010).

Table 4 Biodiversity of native terrestrial invertebrates in the three Antarctic biogeographical zones

Group	*Sub-Antarctic*	*Maritime Antarctic*	*Continental Antarctic*
Protozoa[a]	83		33
Rotifera[a]	>59	>50	13
Tardigrada	>34	26	19
Nematoda[a]	>22	28	14
Platyhelminthes	4	2	0
Gastrotricha	5	2	0
Annelida (Oligochaeta)	23	3	0
Mollusca	3/4	0	0
Crustacea (terrestrial)	4	0	0
Crustacea (nonmarine)	44	10	14
Insecta (total)	210	35	49
(Mallophaga	61	25	34)
(Diptera	44	2	0)
(Coleoptera	40	0	0)
Collembola	>30	10	10
Arachnida (total)	167	36	59
(Araneida	20	0	0)
(Acarina[a]	140	36	29)
Myriapoda	3	0	0

[a]Large changes likely with future research due to current lack of sampling coverage, expertise and/or synonymy (see also www.eba.aq).
ND – number of representatives of group unknown.

Natural Colonization Processes

Contemporary Colonization

Although there are clear oceanic and atmospheric features that form barriers isolating Antarctica from lower southern latitudes, over long timescales these barriers are porous to the movement of biota (Barnes *et al.*, 2006). Aerobiological sampling demonstrates the presence of viable colonizing propagules in the airspora, both from "local" (Antarctic) and distant sources (Pearce *et al.*, 2009), although propagule densities are several orders of magnitude lower than obtained in comparable temperate and tropical studies. Most records refer to microbial or lower plant groups (bacteria, cyanobacteria, fungi, algae,and bryophytes) and lichens. Direct evidence of colonization, on both short and long distance scales, is given by records of new populations on previously barren ground, and of previously unrecorded species in known sites. Compelling evidence is provided by the

colonization of heated ground at the few, very isolated, volcanically active sites spread around Antarctica. These sites are recent, short-lived and of limited area. Their bryophyte and arthropod (and, presumably, microbial) communities contain species not found elsewhere in the Antarctic, but with wider distributions in the sub-Antarctic, South America, or New Zealand (Convey *et al.*, 2000). Other than these exceptional instances, confirmation of natural faunal colonization events is problematical, as many Antarctic distributional records stem from specific studies at single locations by individual taxonomic experts.

Mechanisms

Natural colonization mechanisms may be classified into (i) directed active movement, (ii) assisted transport by other species or debris and (iii) passive transport by air or water currents (Hughes *et al.*, in Bergstrom *et al.*, 2006). No Antarctic terrestrial biota show migratory behavior and option (i) is discounted in explaining current patterns of Antarctic terrestrial biodiversity. However, several vertebrate and invertebrate species with well-known active dispersal characteristics are recorded as vagrants at sub and even maritime Antarctic locations, and some have become established on sub-Antarctic islands in recent years (Frenot *et al.*, 2005).

Assisted transport of nonparasitic invertebrates, plant seeds, and other propagules by vertebrates (birds), other invertebrates, and debris (e.g. driftwood) has been proposed as an important mechanism of Antarctic colonization by several authors (reviewed by Barnes *et al.*, 2006). Transport of terrestrial biota is most likely to be rapid, and *via* vectors that spend time associated with land masses during the austral winter, such as skuas, gulls, and sheathbills (thus coming into contact with terrestrial biota) rather than marine mammals and birds that spend the winter period wholly at sea.

Passive transport in either air or water currents is also plausible, but there is little explicit evidence of its occurrence. Some arthropods, such as oribatid mites and springtails, show rafting behavior or high tolerance of submersion in seawater, as well as including many littoral or marine representatives. Others show no such ability. Transport in the air column presents the twin challenges of low temperature and desiccation. Although many Antarctic biota possess well developed resistance adaptations to these stresses, such transport is most likely to be survived by organisms with specific dispersal (lichen soredia, moss spores) or resistant stages (tardigrade "tuns," anhydrobiotic nematodes, rotifers, protozoan cysts). However, no attempt has yet been made to model or experimentally determine potential or actual survival of such processes.

Dispersing Propagules

Investment in dispersal features is not a characteristic of the "adversity-selected" life histories typical of Antarctic plants and animals (Convey, 1996). Nevertheless, all Antarctic organisms require the ability to disperse over a range of scales, from the mm/cm required to move between areas of favorable microhabitat, through the m/km required for local colonization of icefree ground within the Antarctic, to the hundreds or thousands of km necessary to allow intra or intercontinental dispersal. The wide distribution of most Antarctic species in suitable habitats, particularly within the maritime and continental zones, combined with the recent age of most of these islands of habitat, argues for their success in dispersal.

Various Antarctic microinvertebrates (protozoans, rotifers, tardigrades, and nematodes) can potentially disperse in a resistant desiccated (anhydrobiotic) state. Likewise, algal and cyanobacterial mats and mosses survive long periods of unfavorable conditions in a desiccated state and are subjected to fragmentation and wind dispersal. Lichens, bryophytes, and fungi possess sexually and asexually produced dispersing propagules.

Future Trends

Anthropogenic Influences

Both the Antarctic continent and Southern Ocean were immune from human impact until recent times. Since the early nineteenth century, marine ecosystems have been devastated, at first due to the uncontrolled exploitation of marine mammal (seal, whale) populations and, more recently, the continuing though more controlled impact of various fisheries either directly on their target species or on nontarget bycatch. Human impact on the terrestrial environment commenced with visits of sealing and whaling vessels to sub and maritime Antarctic islands (Convey and Lebouvier, 2009), followed by the first landings on the Antarctic Peninsula and continent approximately 100 years ago. Initiated by the International Geophysical Year (1957/58), many scientific research stations have been erected around the continent, importing personnel and materials into the biome. This has been followed since about 1970 by rapidly increasing tourist operations. Antarctic terrestrial ecosystems are fragile and sensitive to disturbance, often existing on the same areas of icefree ground where research stations are established, and also where large concentrations of wildlife attract tourists. Therefore, human presence has inevitably disturbed and, locally, destroyed areas of terrestrial habitat (Tin *et al.*, 2009). Even before the expansion of these direct impacts, the global consequences of human activity were detectable throughout Antarctica in the form of chemical pollutants (Bargagli, 2005).

Introduction of Alien Organisms

The most obvious direct biological impact of human activities is in the introduction of alien organisms. Accidental introductions and deliberate transplant experiments have shown that a wide range of flora, fauna, and microbes are capable of surviving and establishing viable populations, whereas an even greater number have been recorded on a transient or synanthropic basis (Frenot *et al.*, 2005; Greenslade, 2006; **Table 5**). The greatest impacts have been on sub-Antarctic ecosystems, with the longest record of human influence and least extreme climate. Here, various vertebrates have been introduced both accidentally (rats, mice) and deliberately (fish, chickens, rabbits, cats, pigs, sheep, moufflon, cattle, and

Table 5 The occurrence of alien non-indigenous species across Antarctic biogeographical zones

	Entire sub-Antarctic	*Maritime Antarctic*	*South Georgia*	*Marion*	*Prince Edward*	*Crozet*	*Kerguelen*	*Heard*	*Mac Donald*	*Macquarie*
Dicotyledons	62	0	17	6	2	40	34	0	0	2
Monocotyledons	45	2	15	7	1	18	34	1	0	1
Pteridophytes	1	0	1	0	0	1	1	0	0	0
Total nonindigenous plants	108	2	33	13	3	59	69	1	0	3
Invertebrates	72	2–5	12	18	1	14	30	3	0	28
Vertebrates	16	0	3	1	0	6	12	0	0	6

Extracted from Frenot *et al.* (2005); see also Greenslade (2006) for a detailed description of established and transient alien species, and species recorded only synanthropically, from sub-Antarctic Macquarie Island.

reindeer) to most major islands (Convey and Lebouvier, 2009). In some cases, the consequences of these introductions are irreversible, whereas in others effective control measures appear to be impracticable. The ecological impact of some introductions (e.g., cats, grazing mammals) has been drastic but, generally, appropriate studies have not been made. Introduced species may have both direct (e.g., predation of bird eggs and terrestrial invertebrates, trampling and grazing of plants) and indirect impacts (e.g., alteration of habitat structure leading to changes in species dominance or behavior) on native species.

Introductions of invertebrates are less well documented at most locations. All alien invertebrates (with the exception of rabbit fleas used to disperse the myxoma virus amongst some introduced rabbit populations) have been introduced accidentally, with stores, food, equipment, domestic animals, etc, associated with human settlement (Frenot *et al.*, 2005). Although many have persisted only as long as human occupation continued, some have become established (particularly in the sub-, but also in the maritime Antarctic), with evidence of competitive displacement and local extinction of native species, and of new trophic links being introduced to terrestrial ecosystems.

Introductions of flowering plants, bryophytes, and microbes have also accompanied human occupation. Again, the sub-Antarctic has been most affected (**Table 5**), although many Arctic, sub-Antarctic, and some temperate vascular plants are also capable of long-term survival of maritime Antarctic conditions. Data on microbial introductions are sparse, although the process is inevitable via human transport (see Broady, in Wynn-Williams, 1996) and complete control measures impracticable. Given the possible evolutionary isolation of Antarctic microbes, the potential for damage to this unique biological resource should not be underestimated.

Regional and Global Environmental Change

Description and understanding of global climate change currently receives much attention from scientists, politicians, media and the general public. Most global climate models agree in predicting that any regional temperature increases will be greatest and most rapid at high latitudes. There is clear evidence of rapid regional warming trends at several maritime and sub-Antarctic research stations, along with changes in other environmental variables affecting various parts of the continent (Turner *et al.*, 2009; Convey *et al.*, 2009b). Warming rates along the Antarctic Peninsula are among the most rapid seen worldwide.

Another important prediction of climate change models is of changing patterns of precipitation, altering water input to terrestrial habitats. Spatially detailed predictions are not yet available for Antarctica, although the role of water as possibly the single most important factor limiting distribution of Antarctic terrestrial biota is well recognized. At a local scale, water availability to terrestrial ecosystems can undergo drastic change as a result of factors as simple as the complete exhaustion of a snow bank or, conversely, increased release of water from melting icefields. These effects may be caused by changes in either the timing or amount of snow accumulation, or the duration of positive summer temperatures (i.e., in combination with temperature amelioration).

A separate anthropogenic influence on Antarctic and high latitude Southern Hemisphere ecosystems – the ozone hole – has existed since the early 1980s. Its formation leads to a temporary but drastic reduction in stratospheric ozone concentration, allowing a disproportionate increase in the exposure of terrestrial (and shallow marine) ecosystems to potentially damaging shorter wavelength UV–B radiation. Levels experienced at any site vary depending on the pattern of movement of the ozone hole, and modulation by factors such as clouds and snow cover but, at worst, exposure levels (integrated over time) comparable to those of the tropics are experienced.

Biological Consequences of Climate Change

Prediction and identification of the effects of climate change processes in the Antarctic terrestrial environment has received critical attention over the last decade (see Kennedy, 1995; Convey, 2003; Bergstrom *et al.*, 2006 for review). Many indigenous Antarctic terrestrial biota exhibit sufficient ecophysiological flexibility to absorb and even benefit from both the direct and indirect (e.g., changed nutrient availability) effects of predicted levels of change. However, how individual species responses may be integrated to community or ecosystem level is unknown in Antarctic systems, although recognized as a key research challenge (Day, 2001; Convey, 2003). Any amelioration is likely to influence Antarctic ecosystems and biodiversity

further by easing the constraints that limit colonization and establishment of alien species (Frenot *et al.*, 2005).

The effects of warming are expected to be greater at high latitude sites that already have a very restricted thermal energy budget, where the relative importance of a small temperature increment will be of greater significance. Thus the main consequences of temperature amelioration (assuming continued water availability) are to increase the effective length of season and thermal energy budget, in turn leading to increased growth, reproductive rates, and population sizes both of photosynthetic autotrophs and heterotrophic microbial and invertebrate species. The evidence available from observational and manipulation studies in the Antarctic supports this contention (Convey, 2010). As also reported from the Arctic, there is considerable variability in the responses seen, with some biota responding very rapidly or significantly, whereas others, or their habitats, appear to be robust or buffered to change (Hodkinson *et al.*, 1998).

Some climate change processes are likely to be deleterious for Antarctic biota, at least at a local scale. Any decrease or total loss of water input to ecosystems could lead to local extinctions and drastic changes in ecosystem structure. Likewise, should maximum exposure to ozone hole related UV–B radiation surpass the tolerances of exposed biota, particularly colonizing microbiota with crucial roles in soil stabilization and permitting secondary colonization by other organisms, some terrestrial habitats may become barren. Such drastic eventualities are yet to be reported from the Antarctic. Much more likely and ecologically relevant, but harder to quantify, are changes in patterns of organism investment in the ecophysiological and biochemical responses (e.g., cryoprotectant, UV screening pigment production, and repair processes) required to permit survival of changing patterns and intensities of environmental stress (see Snell *et al.*, 2009 for an example relating to UV screening pigments in plants).

Climate amelioration will lead both to an increased area of available terrestrial habitat and a longer "window of opportunity" within which to complete the dual processes of colonization and establishment. With or without human assistance, this will lead to increases in biodiversity and changed community composition, and, in turn, to greater trophic complexity and the inclusion of higher trophic levels. The resistance of indigenous Antarctic biota to these processes is unknown, although they are generally thought to be poor competitors.

Acknowledgments

I thank the British Antarctic Survey (BAS) for resources and the many BAS colleagues, and members of the international SCAR biological community, who have been involved in many discussions over time that have contributed to the development of the ideas in this paper.

See also: Alpine Ecosystems. Arctic Terrestrial Ecosystems. Climate Change and Ecology, Synergism of. Desert Ecosystems. Dispersal Biogeography. Diversity, Community/Regional Level. Endemism. Energy Flow and Ecosystems. Geologic Time, History of Biodiversity in. Greenhouse Effect. Human Impacts on Ecosystems: An Overview. Human Impact on Biodiversity, Overview. Introduced Species, Impacts and Distribution of. Invertebrates, Terrestrial, Overview. Latitudinal and Elevational Range Shifts Under Contemporary Climate Change. Phenological Shifts in Animals Under Contemporary Climate Change. Psychrophiles. Soil Biota, Soil Systems, and Processes. Stress, Environmental. Terrestrial Ecosystems. Ultraviolet Radiation. Vicariance Biogeography

References

Adams B, Bardgett RD, Ayres E, *et al.* (2006) Diversity and distribution of Victoria Land biota. *Soil Biology Biochemistry* 38: 3003–3018.

Allegrucci G, Carchini G, Todisco V, *et al.* (2006) A molecular phylogeny of Antarctic Chironomidae and its implications for biogeographical history. *Polar Biology* 29: 320–326.

Bargagli R (2005) *Antarctic Ecosystems: Environmental Contamination, Climate Change, and Human Impact*. Berlin: Springer.

Barnes DKA, Hodgson DA, Convey P, *et al.* (2006) Incursion and excursion of Antarctic biota: Past, present, and future. *Global Ecology and Biogeography* 15: 121–142.

Bergstrom D, Convey P, and Huiskes AHL (eds.) (2006) *Trends in Antarctic Terrestrial and Limnetic Ecosystems*. Dordrecht: Springer.

Beyer L and Bölter M (2002) *Geoecology of Antarctic Ice Free Coastal Landscapes*. Berlin: Springer.

Chown SL and Convey P (2007) Spatial and temporal variability across life's hierarchies in the terrestrial Antarctic. *Philosophical Transactions of the Royal Society, Series B* 362: 2307–2331.

Convey P (1996) The influence of environmental characteristics on life history attributes of Antarctic terrestrial biota. *Biological Reviews* 71: 191–225.

Convey P (2003) Maritime Antarctic climate change: signals from terrestrial biology. In: Domack E, Burnett A, Leventer A, *et al.* (eds.) *Antarctic Peninsula Climate Variability: Historical and Palaeoenvironmental Perspectives*. Washington, DC: American Geophysical Union.

Convey P (2010) Terrestrial biodiversity in Antarctica – recent advances and future challenges. *Polar Science* 4: 135–147.

Convey P, Bindschadler RA, di Prisco G, *et al.* (2009b) Antarctic climate change and the environment. *Antarctic Science* 21: 541–563.

Convey P, Gibson J, Hillenbrand C-D, *et al.* (2008) Antarctic terrestrial life – challenging the history of the frozen continent? *Biological Reviews* 83: 103–117.

Convey P and Lebouvier M (2009) Environmental change and human impacts on terrestrial ecosystems of the sub-Antarctic islands between their discovery and the mid-Twentieth Century. *Papers and Proceedings of the Royal Society of Tasmania* 143: 33–44.

Convey P and McInnes SJ (2005) Exceptional, tardigrade dominated, ecosystems from Ellsworth Land. *Antarctica Ecology* 86: 519–527.

Convey P and Smith RIL (2006) Thermal relationships of bryophytes from geothermal habitats in the South Sandwich Islands, maritime Antarctic. *Journal of Vegetation Science* 17: 529–538.

Convey P, Smith RIL, Hodgson DA, *et al.* (2000) The flora of the South Sandwich Islands, with particular reference to the influence of geothermal heating. *Journal of Biogeography* 27: 1279–1295.

Convey P and Stevens MI (2007) Antarctic Biodiversity. *Science* 317: 1877–1878.

Convey P, Stevens MI, Hodgson DA, *et al.* (2009a) Exploring biological constraints on the glacial history of Antarctica. *Quaternery Science Reviews* 28: 3035–3048.

Cowan DA, Russell NJ, Mamais A, *et al.* (2002) Antarctic dry valley mineral soils contain unexpectedly high levels of microbial biomass. *Extremophiles* 6: 431–436.

Day TA (2001) Multiple trophic levels in UV–B assessments - completing the ecosystem. *New Phytologist* 152: 183–186.

De Wever A, Leliaert F, Verleyen E, *et al.* (2009) Hidden levels of phylodiversity in Antarctic green algae: further evidence for the existence of glacial refugia. *Proceedings of the Royal Society London B*. http://dx.doi.org/10.1098/rspb.2009.0994.

EPICA (2004) Eight glacial cycles from an Antarctic ice core. *Nature* 429: 623–628.

Frenot Y, Chown SL, Whinam J, *et al.* (2005) Biological invasions in the Antarctic: extent, impacts and implications. *Biological Reviews* 80: 45–72.

Friedmann EI (1993) *Antarctic Microbiology*. New York: Wiley.
Greenslade P (2006) *Macquarie Island*. Hobart: Australian Antarctic Division.
Hodkinson ID, Webb NR, Bale JS, *et al*. (1998) Global change and Arctic ecosystems: conclusions and predictions from experiments with terrestrial invertebrates on Spitsbergen. *Arctic and Alpine Research* 30: 306–313.
Hogg ID, Cary SC, ik P, *et al*. (2006) Biotic interactions in Antarctic terrestrial ecosystems: Are they a factor? *Soil Biology and Biochemistry* 38: 3035–3040.
Kennedy AD (1995) Antarctic terrestrial ecosystem responses to global environmental change. *Annual Review of Ecology, Evolution, and Systematics* 26: 683–704.
Laws RM (ed.) (1984) *Antarctic Ecology,* vol. 1. London: Academic Press
Lewis AR, Marchant DR, Ashworth AC, *et al*. (2008) Mid-Miocene cooling and the extinction of tundra in continental Antarctica. *PNAS* 105: 10676–10680.
Livermore RA, Nankivell AP, Eagles G, *et al*. (2005) Paleogene opening of Drake passage. *Earth and Planetary Science Letters* 236: 459–470.
Longton RE (1988) *The Biology of Polar Bryophytes and Lichens*. Cambridge: Cambridge University Press.
Lyons WB, Howard-Williams C, and Hawes I (eds.) (1997) *Ecosystem Processes in Antarctic Icefree Landscapes*. Rotterdam: Balkema.
Maslen NR and Convey P (2006) Nematode diversity and distribution in the southern maritime Antarctic – clues to history? *Soil Biology and Biochemistry* 38: 3141–3151.
Mortimer E, Jansen van Vuuren B, Lee JE, *et al*. (2010) Mite dispersal among the Southern Ocean Islands and Antarctica before the last glacial maximum. *Proceedings of the Royal Society Series B*. http://dx.doi.org/10.1098/rspb.2010.1779.
Ochyra R, Smith RIL, and Bednarek-Ochyra H (2008) *The illustrated moss flora of Antarctica*. Cambridge: Cambridge University Press.
Øvstedal DO and Smith RIL (2001) Lichens of Antarctica and South Georgia. *A guide to their Identification and Ecology*. Cambridge: Cambridge University Press.
Pearce DA, Bridge PD, Hughes KA, *et al*. (2009) Microorganisms in the atmosphere over Antarctica. *FEMS Microbiology Ecology* 69: 143–157.
Peat HJ, Clarke A, and Convey P (2007) Diversity and biogeography of the Antarctic flora. *Journal of Biogeography* 34: 132–146.
Pugh PJA and Convey P (2008) Surviving out in the cold: Antarctic endemic invertebrates and their refugia. *Journal of Biogeography* 35: 2176–2186.
Selkirk PM (2007) The nature and importance of the sub-Antarctic. *Paleontol Stratigraf Litolog Papers and Proceedings of the Royal Society of Tasmania* 141: 1–6.
Skotnicki ML, Selkirk PM, Broady PA, *et al*. (2001) Dispersal of the moss *Campylopus pyriformis* on geothermal ground near the summit of Mount Erebus and Mount Melbourne, Victoria Land, Antarctica. *Antarctic Science* 13: 280–285.
Smith RIL (2005) The thermophilic bryoflora of Deception Island: unique plant communities as a criterion for designating an Antarctic Specially Protected Area. *Antarctic Science* 17: 17–27.
Snell KRS, Kokubun T, Griffiths H, *et al*. (2009) Quantifying the metabolic cost to an Antarctic liverwort of responding to UV-B radiation exposure. *Global Change Biology* 15: 2563–2573.
Stevens MI, Greenslade P, Hogg ID, *et al*. (2006) Examining Southern Hemisphere springtails: Could any have survived glaciation of Antarctica? *Molecular Biology and Evolution* 23: 874–882.
Tin T, Fleming Z, Hughes KA, *et al*. (2009) Impacts of local human activities on the Antarctic environment: a review. *Antarctic Science* 21: 3–33.
Tripati A, Backman J, Elderfield H, *et al*. (2005) Eocene bipolar glaciation associated with global carbon cycle changes. *Nature* 436: 341–346.
Turner J, Bindschadler R, and Convey P (eds.) (2009) *Antarctic Climate Change and the Environment*. Cambridge: Scientific Committee on Antarctic Research.
Vincent WF (1988) *Microbial Ecosystems of Antarctica*. Cambridge: Cambridge University Press.
Vyverman W, Verleyen E, Wilmotte A, *et al*. (2010) Evidence for widespread endemism among Antarctic micro-organisms. *Polar Science* 4: 103–113.
Wynn-Williams DD (ed.) (1996) Antarctic microbial diversity. *Biodiversity and Conservation* 5(11) (Special issue).

Aquaculture

Max Troell, Beijer Institute, Sweden and Stockholm University, Stockholm, Sweden
Nils Kautsky, Stockholm University, Stockholm, Sweden
Malcolm Beveridge, Worldfish Center, Lusaka, Zambia
Patrik Henriksson, Leiden University, Leiden, The Netherlands
Jurgenne Primavera, Southeast Asian Fisheries Development Centre, Tigbauan, Philippines
Patrik Rönnbäck, Gotland University, Visby, Sweden
Carl Folke, Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Kautsky, N., Folke, C., Roennbaeck, P., Troell, M., Beveridge, M., Primavera, J., volume 1, pp 185–198, © 2001, Elsevier Inc.

Glossary

Aquaculture The farming of aquatic organisms, including fish, mollusks, crustaceans, and aquatic plants. Farming implies some sort of intervention in the rearing process to enhance production, such as regular stocking, feeding, or protection from predators. Farming also implies individual or corporate ownership of the stock being cultivated (Definition by FAO).

Broodstock Fish or shellfish from which a first or subsequent generation may be produced in captivity, whether for growing as aquaculture or for release to the wild for stock enhancement.

Ecosystem service Ecosystem services are the benefits people obtain from ecosystems. These include provisioning services such as food and water; regulating services such as flood and disease control; cultural services such as esthetic and recreational, values; and supporting services, such as nutrient cycling, which maintain the conditions for life on Earth.

Farming intensity In a broad continuum, extensive systems are those which are closest to natural fisheries, requiring minimal inputs and offering relatively low yields, whereas intensive systems require a large amount of inputs to maintain an artificial culture environment, with high yields. Between these extremes are the varying degrees of semi-intensive aquaculture, where definitions are less distinct: (i) extensive aquaculture does not involve feeding of the organism, (ii) semi-intensive aquaculture involves supplementation of natural food by fertilization and/or the use of feeds, and (iii) intensive aquaculture is when the culture species is maintained entirely by feeding with nutritionally complete diets.

Feed conversion The efficiency of farmed animals to incorporate given feed into biomass. Feed conversion is usually expressed in terms of the feed conversion ratio of weight of feed provided to fish/shellfish flesh biomass harvested. The ratio is affected by the relative moisture content of both feed and aquaculture product as well as by the metabolic characteristics of the farmed species, farming techniques, and husbandry.

Life cycle assessment (LCA) Environmental framework incorporating the whole production chain with the intention of (1) producing an inventory of the economic and environmental inputs and outputs to each stage of a product/service life cycle and (2) quantifying a subset of the environmental impacts potentially associated with those flows using standardized impact assessment methods.

Seed A term used to describe eggs, larvae, postlarvae, or juveniles (fry and fingerlings) stocked into aquaculture production systems.

Spawner Mature individual of a stock responsible for reproduction.

Introduction

Aquaculture, the aquatic counterpart of agriculture, has grown rapidly in recent decades, and today it produces almost as much fish and shellfish as fisheries. Aquaculture is the main means for obtaining more food from our aquatic environments in the future. Impacts of aquaculture on biodiversity arise from the consumption of resources, such as land (or space), water, seed, and feed, their transformation into products valued by society, and the subsequent release into the environment of greenhouse gases and wastes from uneaten food, fecal, and urinary products, chemotherapeutants as well as microorganisms, parasites, and feral animals. Negative effects may be direct, through release of eutrophicating substances, toxic chemicals, the transfer of pathogens diseases and parasites to wild populations stock, and the introduction of exotic and genetic material into the environment, or indirect through loss of habitat and niche space and changes in food webs. Today, large quantities of fish are caught to produce fishmeal and fish oil – an important source of ingredients for protein and fatty acids in feeds for many fish and shrimp species – which besides being a questionable usage of a valuable food resource also may contribute to the high fishing pressure on some stocks with potential negative consequences for marine food webs. Despite advances being made within the feed industry, resulting in decreased feed conversion ratios and development of suitable alternatives to fish resources, the aquaculture industry's use of global fishmeal and fish oil increased three-fold between 1992 and 2006 (Hasan and Halwart, 2009).

The life cycles for most aquaculture species have been successfully closed but the culture of some, especially marine,

farmed fish, and shrimp, is still partially dependent on the capture of larvae, postlarvae, or gravid females from the wild. This can result in both over-fishing and bycatch, representing losses to capture fisheries and biodiversity. Large areas of critical habitats such as wetlands and mangroves have historically been lost due to the siting of aquaculture developments and pollution, resulting in reduced biodiversity and recruitment to capture fisheries. Today the values of coastal wetlands are better recognized and regulations make it difficult for large-scale aquaculture development in sensitive coastal areas (e.g., mangroves, seagrass beds) in most countries.

The magnitude of biodiversity loss from aquaculture development generally increases with scale, intensity of resource use, and net production of wastes. It is also very much dependent on the location of production site, species being cultured, the aquaculture system being used, method of cultivation and management, and the aquaculture site. In some cases aquaculture may also increase local biodiversity, for example, when ponds are constructed in dry areas, through restocking activities and from integrated aquaculture. Its role in maintaining cultural diversity must also be acknowledged.

Aquaculture Development and Practices

The farming of aquatic plants and animals is several thousands of years old. Nevertheless, it must be regarded as a largely post-World War II phenomenon. In 1950, global farmed fish and shellfish production was approximately 2 mmt (million metric tons) and largely confined to areas of Asia. During the last three decades, aquaculture production has increased by approximately 7–11% per year. Production for 2008 was 33.9 mmt of finfish, 5 mmt of crustaceans, 13.1 mmt of mollusks, and 15.7 mmt of aquatic plants (FAO, 2010). Fish produced by aquaculture now accounts for half of all fish directly consumed by humans. More than 400 species of fish and shellfish are farmed; the range includes giant clams that obtain most of their nutrients from symbiotic algae, various species of carps that are largely herbivorous, and Atlantic salmon and marine fish species that are carnivorous (**Table 1**).

Aquaculture typically involves the enclosure of a species in a secure system under conditions in which it can thrive. Interventions in the life cycle range from exclusion of predators and control of competitors (extensive aquaculture) to enhancement of food supply (semi-intensive) or even the provision of all nutritional requirements (intensive). Intensification of production also implies increasing the number of individuals per unit area, which decreases the local demand for land/sea space but instead requires greater use and management of inputs and a greater reliance on technology and fossil energy.

Aquaculture is an economic activity that uses and transforms natural aquatic resources into commodities valued by society and in so doing it may impact on biodiversity, essentially due to the consumption of resources, the transformation process (aquaculture), and the production of wastes (Naylor *et al.*, 2000; Boyd *et al.*, 2005; Diana, 2009) (Figure 1). Looking at the diversity of farming systems it is easy to appreciate that the biophysical impacts of aquaculture activities, that is, magnitude and spatial scale, vary enormously. Technical and economic inputs, such as construction materials and energy, in many traditional aquaculture systems form only a small part of the inputs needed. The main and critical inputs are instead natural resources and to some extent also labor. Together with ecosystem services these ultimately determine the limits for the local and global expansion of aquaculture. The magnitude and type of resource use and impacts of aquaculture are, however, very much dependent on species cultured, farming system, intensity of farming methods, and management.

Production practices and their impacts on aquatic ecosystems vary widely across species. Mollusks such as scallops and mussels are generally farmed along subtidal or intertidal coastlines where wild-caught or hatchery-reared seed are grown in bags set on the sea bottom or on stakes and suspended ropes. The animals rely entirely on prevailing supplies of plankton and organic particles for food. Even though no feed is added, large densities of filter-feeding organisms may cause reversible local organic accumulation, resulting in

Table 1 Summary of the most important aquaculture species groups, farming systems and methods in terms of production

Group	*System*	*Method of culture*
Plants		
Eucheuma, Kappaphycus, Gracilaria, Gelidium, Caulerpa	Stakes, rafts, long-lines, beds	Extensive
Bivalve molluscs		
Oysters (*Crassostrea*, spp.), mussels (*Mytilus* spp.), Cockles, Abalone	Rafts, long-lines, stakes, beds, tanks	Extensive
Shrimps and prawns		
Shrimps (*Penaeus* spp., Litopenaeus vannamei)	Ponds	Extensive, semi-intensive, intensive
Marine, brackishwater fish		
Milkfish (*Chanos chanos*), yellowtail (*Seriola* spp.), groupers (*Epinephelus* spp.), mullets (*Mugil* spp., *Liza* spp.), cobia (*Rachycentron canadum*)	Ponds, cages,	Semi-intensive, intensive
Freshwater/diadromous fish		
Chinese carps, Indian carps, tilapia, Atlantic salmon (*Salmo salar*), trout, catfish	Ponds, cages, tanks	Extensive, Semi-intensive, intensive

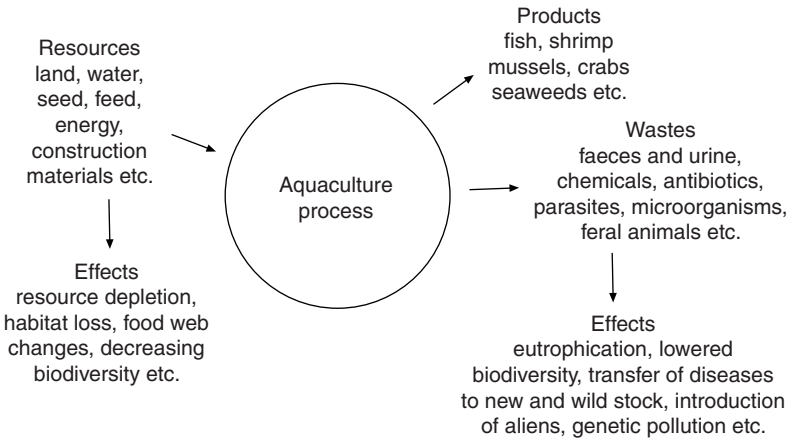


Figure 1 Diagram illustrating the principal direct and indirect effects from aquaculture on biodiversity, through the use of resources and the generation of wastes. It should be noted that additional impacts associated with specific culture systems exist, and that as impacts resulting from management are not included. Note that other impacts being specific to system. Details are given in the text.

negative effects on benthic diversity, and possibly also, if farmed in large quantities, impact on pelagic plankton communities. A range of systems – ponds, tanks, or cages – are used to farm finfish. The majority of carp and other freshwater species farmed in the tropics and subtropics are herbivores/omnivores and are grown in ponds fertilized by supplemental feeds. In contrast, most diadromous and marine finfish, including both tropical and temperate species, are farmed intensively in floating net cages and are to a large extent reliant on nutritionally complete fishmeal and fish oil-based diets. For some marine fish species farmed mainly in Southeast Asia, including also freshwater species like *Pangasius* catfish (e.g., in Vietnam) large amounts of fish, so-called trash fish, are being used directly as feed (Huntington and Hasan, 2009).

Penaeid shrimps, dominating crustacean farming, are reared in semi-intensive or intensive coastal pond systems. The shrimps depend mainly on formulated pellet feeds, aeration to replenish dissolved oxygen, and pumped seawater to dilute pollutants and flush out harmful metabolites. Shrimp postlarvae are either derived from captured wild spawners/broodstock or directly collected from the sea – something that is true also for the freshwater shrimp *Macrobrachium* (Ahmed *et al.*, 2010).

The aquaculture process in itself may affect biodiversity as a result of disturbance through increased road and boat traffic. High densities of farmed fish and food often attract predators and scavengers such as wild fish, gulls, and seals. These can come into conflict with farmers and may be killed, either accidentally (entanglement in nets) or deliberately (shooting and trapping), or if they become established they may displace sensitive local species (Boyd *et al.*, 2005).

Feed Resources

Extensive or semi-intensive aquaculture, for example, pond farmed carps and filter-feeding bivalves, depends either on natural production or agricultural wastes and some generally locally made feed. Filter feeders like mussels do not depend on addition of feed; however, very dense farming of mussels and bivalves in semi-enclosed coastal areas may in exceptional cases reduce fisheries stocks by shortening the linkages comprising the food web. In Rio Arosa, Spain, for example, overgrazing of the phytoplankton population by filter-feeding mussels resulted in zooplankton starvation and the subsequent collapse of the sardine fishery (GESAMP, 1991).

Nutrient-rich materials added to stimulate the growth of algae and other food items, and on-farm feeds, based largely on cheap, locally available agricultural by-products, augmented by household scraps and perhaps small amounts of fishmeal, are used to supplement the food in extensive ponds. However, in the intensive production systems that predominate in temperate aquaculture, and are increasing in the tropics, the farmed animals are to varying degrees reliant on nutritionally complete commercial feeds containing fishmeal and fish oil. Marine fish species and shrimps have high demand for fish resources in the feed but also freshwater fish species like carp, tilapia, and catfish, which being herbivores or omnivores, are also increasingly being farmed using formulated feeds containing various percentage of fishmeal and fish oil (Tacon and Metian, 2008).

Diets for salmonids, seabass, sea bream, and other carnivores are largely composed of fishmeal and fish oil. Although it may be possible to replace much of the fishmeal used in intensive fish diets with proteins of plant origin (e.g., oilseeds) (Stickney *et al.*, 1996; Hasan and Halwart, 2009), requirements for essential amino acids, especially cystine and methionine, is still to a large extent being met by fishmeal. It remains to be seen whether commercial plant protein-based diets can be developed in an industry in which the product is competing with many others for customer attention and in which profit margins are increasingly being squeezed. Depending on the source and inclusion rate, oilseed meals can compromise palatability, growth (Stickney *et al.*, 1996), and profitability. Any decrease in palatability or diet digestibility may aggravate waste loadings to the environment. The issue of fish oils is even more pressing than that of fishmeal (Naylor *et al.*, 2009). Aquatic carnivores are poor at using carbohydrate to supply energy requirements, relying instead on protein and lipid (Cowey and Sargent, 1977). The substitution of fish oils with vegetable oils in freshwater carnivorous or omnivorous

fish diets is more straightforward than for marine and diadromous carnivorous species, such as Atlantic salmon, which require n-3 highly unsaturated fatty acids. Progress has been made with respect to alternatives sources like rapeseed and linseed as well as synthesis of n-3 fatty acids, which can supplement various feeds while still resulting in acceptable high growth and quality (Naylor *et al.*, 2009).

Currently, approximately one-third of the total harvest from capture fisheries is destined for nonfood use, of which most is used to produce fishmeal and fish oil for aquaculture (Tacon and Metian, 2009) (Figure 2). Fish species being used for reduction to fishmeal includes species such as Peruvian and Japanese anchovy, Blue whiting, Atlantic herring, chub mackerel, Chilean jack mackerel, capelin, European pilchard (Kaushik and Troell, 2010). Many fishes exploited for feed are fast growing opportunistic species and therefore can sustain a heavy fishing pressure. However, production of some of these species is also constrained by climatic variability associated with El Nino–Southern Oscillation events (FAO, 1997; NRC, 1999) and impacts on marine biodiversity from intensive fishing on pelagic forage fish species, including reduction of available food supplies for marine predators and valuable species consumed by humans, have not been studied in a systematic way (Smith *et al.*, 2011). In Europe, the crash of North Sea capelin and herring stocks has been attributed to over-fishing and this may have caused depletion of other wild fish stocks (e.g., cod) and the starvation of seals and seabird chicks (Vader *et al.*, 1990; Naylor *et al.*, 2000). Declining capelin populations in the western Gulf of Alaska are implicated in the decrease of harbor seal and sea lions in the early 1980s (Hansen, 1997). A strong interaction between anchoveta and seabird and mammal populations has also been well documented for the Peruvian upwelling system (Pauly and Tsukayama, 1987). Recent focus on fishing for krill near Antarctic for aquaculture feeds may result in negative ecosystem consequences for a system that is not yet fully understood.

Low value fish, that is, so-called "trash fish" (from rivers, lakes, and the sea) has increasingly been used directly as feed in aquaculture, especially in Asian countries, with implication for both biodiversity and food security (Huntington and Hasan, 2009). Despite significant progress being made on reducing inclusion of fishmeal and fish oil in feeds, and finding alternative feed ingredients (e.g., plants and microorganisms)

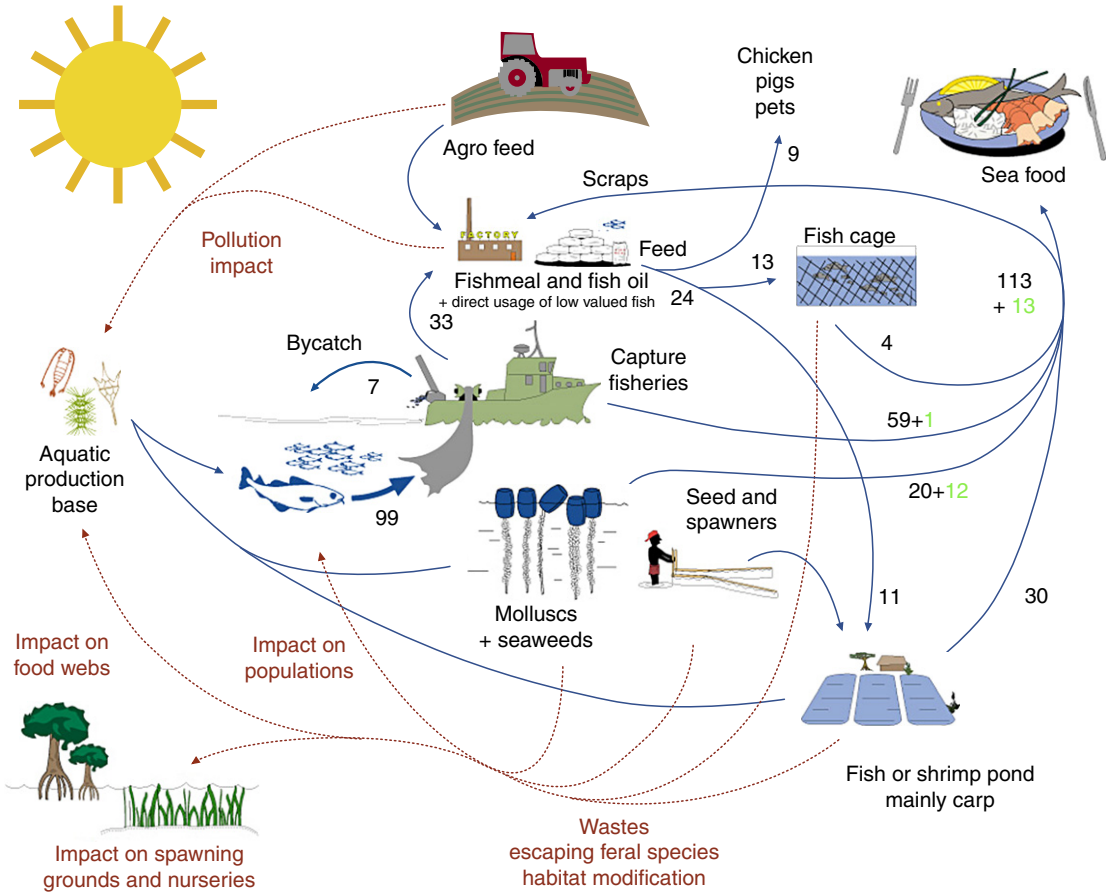


Figure 2 Ecological links between aquaculture and capture fisheries. Thick blue lines refer to main flows from aquatic production base through fisheries and aquaculture to human consumption of seafood. Numbers refer to 2006 data and are in units of megatons (million metric tons) of fish, shellfish and seaweeds (green). Thin blue lines refer to other inputs needed for production. Hatched red lines indicate negative feedbacks. Flows to, for example, pond aquaculture from cage aquaculture also exist as used feed fish generates surplus of fishmeal (not in Figure) (Graphic outline from Naylor *et al.*, 2000. Data from Tacon AGJ and Metian M (2008) Global overview on the use of fish meal and fish oil industrially compounded aquafeeds: Trends and future prospects. *Aquaculture* 285: 146–158).

(Tacon and Metian, 2008; Naylor *et al.*, 2009), some intensive and semi-intensive aquaculture systems use more fish protein to feed the farmed species than is ultimately harvested. Also, the increase in global aquaculture production has resulted in a net increase in total demand for fish resources in aquaculture and it is estimated that 68% and 98% of global fishmeal and fish oil, respectively, were utilized by the aquaculture sector in 2006 (Tacon and Metian, 2008; FAO, 2010). Alternatives to fishmeal and fish oil include soy meal, soy oil, livestock co-products, vegetable oil, single cell, and other locally available resources (e.g., snails in Bangladesh). Other commonly used feed ingredients for many freshwater fish species include rice and wheat bran, maize gluten meal, and cassava meal (Naylor *et al.*, 2009).

In 2006, 74.2% of total aquaculture production was of species feeding low in the aquatic food chain, including aquatic plants, filter-feeding molluscs, and herbivorous and omnivorous finfish species. Farming of fish species such as carps, tilapia, and catfish dominates production and even though this production is increasingly based on fish resources, it results in a net production of fish. This will, however, change if farming methods continue to intensify and increasingly utilize higher quality fishmeal-based feeds or feeding with low-valued fish (Tacon and Metian, 2009). For other fish species, for example, marine and some diadromous species (e.g., salmon), production requires more fish as feed than is ultimately produced. For example, approximately 3.4 kg of wild fish is used to produce 1 kg of farmed salmon (estimated IFFO values for 2010 in Tacon and Metian, 2008). The culturing of such species thus leads to a net loss in fish protein and fish oil.

Human consumption of seafood was 112 mmt in 2006, of which 59 mmt of fish, crustaceans, and mollusks come from capture fisheries, whereas 52 mmt are from aquaculture (Tacon and Metian, 2008; FAO, 2010) (Figure 2). Total capture fisheries was 99 mmt but 7 mmt of this is discarded as bycatch and 33 mmt being used for fishmeal production or as direct feed, often for aquaculture (Tacon and Metian, 2009; FAO, 2010). An additional few mmt of processing scraps from aquaculture and fisheries are also converted into fishmeal. Thus, 64% of total fish resources are being transformed to fishmeal and fish oil or used directly as animal feeds. Twenty-four mmt are currently used in aquaculture, the rest being used for chickens, pigs, and other animals, as well as health food supplements (approximately 9 mmt) (Tacon and Metian, 2009). However, even though farming efficiency improves and substitutes for fish in feeds are being developed, an increasing proportion of fish resources will probably be used for aquaculture feeds as supplies are unlikely to expand and as aquaculture production continues to grow (and demand for higher trophic level species) and production methods of pond fishes in major producer countries such as China intensify (Rana *et al.*, 2009).

Land, Water, and Energy

Land is needed to build fish or shrimp ponds or establish tank-based operations, whereas fish cages, pens, and mussel and seaweed farm operations occupy areas of lakes, rivers, and the sea. Fish ponds are usually sited in agricultural land and this arguably contributes positively to the floral and faunal diversity of agriculture landscapes. However, unproductive, boggy areas of agricultural land have often been used, and since such boundary ecosystems, or ecotones, may serve as reserves for species in areas otherwise surrounded by monocultures of crops, this may in fact reduce biodiversity. Land is becoming scarcer and the increasing competition with other users, for example, agriculture and urban development, puts pressure on aquaculture to minimize appropriation of land. This has promoted the intensification of farming systems with high stocking densities, resulting in higher dependence on external concentrated inputs , for example, feed, energy, and chemicals. On the positive side, less land is needed per metric ton of fish production and the resultant more concentrated wastes are also more amenable to treatment. However, if treatment is not done the environmental impacts may be severe.

Large areas of tropical coastal wetlands and mangroves have historically been converted to fish and shrimp ponds, resulting in impoverished biodiversity and recruitment to fisheries, with consequences for local communities and regional economies (Rönnbäck *et al.*, 2002; Walters *et al.*, 2008). Thus, since the 1400s, hundreds of thousands of hectares of mangroves have been transformed into milkfish ponds in Indonesia and the Philippines. In recent decades, shrimp farming has been responsible for a significant share of the conversion of coastal and supratidal areas, for example, 102,000 ha of mangrove forests in Vietnam during 1983–1987 (Tuan, 1997) and 65,000 ha in Thailand during 1961–1993 (Menasveta, 1997). When the full range of ecological effects associated with mangrove habitat loss are accounted for, the net production in fish and shrimp aquaculture may be negative (Rönnbäck, 1999; Barbier *et al.*, 2008). Coastal ecosystems, such as mangroves, seagrass beds, and coral reefs, provide habitats and nursery areas for many fish and invertebrate species caught in coastal and offshore fisheries (Robertson and Duke, 1987). A positive relationship between commercial fish/shrimp landings and mangrove area has also been documented throughout the tropics (Pauly and Ingles, 1986; Rönnbäck, 1999). To identify and value total commercial and subsistence fisheries catch supported by mangroves, economic analyses must take into account: (1) the large number of resident and transient species that utilize mangroves as habitat; (2) the direct and indirect subsidies of shrimp trawlers and mangroves, respectively, to total fisheries catch; and (3) the aquaculture industry's dependence on inputs like seed, broodstock, and feed (Rönnbäck, 1999; Walters *et al.*, 2008). By acknowledging these support functions, the potential life-support value of mangroves to fisheries is in the order of 1–10 mmt of fish and shellfish per hectare per year (first sale value ≈$1000–10,000 US in developing countries) (Rönnbäck, 1999). In addition, mangroves also harbor a wide array of non-marketed fish, crustacean, and mollusk species, whose subsistence harvest constitutes an important protein source for coastal communities. Moreover, mangroves are closely linked to habitat conditions and associated biodiversity of coral reefs and seagrass beds through the biophysical interactions in the coastal seascape (Moberg and Rönnbäck, 2003; Nagelkerken, 2009). Almost one-third of the

world's marine fish species are associated with coral reefs, and fish catches from reefs contribute to approximately 10% of human fish consumption at a global level and much higher in developing countries (Weber, 1993; Moberg and Folke, 1999).

Aquaculture requires large amounts of water to physically support the farmed animals, replenish oxygen, and remove wastes. The impacts of aquaculture on the quantity and quality of water resources have direct impacts on associated aquatic biodiversity. Large amounts of water pass through cages and pens (in the sea or lakes) but there is no net removal from the system. It is important to distinguish between water consumption and water withdrawal, the former implying that water diverted from streams, rivers or aquifers lost through evaporation or seepage, while in the latter water is returned to the environment to be reused or restored (Verdegem and Bosma, 2009). Consumptive water use in aquaculture has mainly impacted on the reduction in downstream freshwater flows and groundwater resources (Boyd *et al.*, 2005). In land-based systems, aquaculture not only borrows water, returning it in a more degraded form, but also consumes it or accelerates its loss from surface to groundwater or the atmosphere (Boyd, 2005; Boyd *et al.*, 2005). Thus, by creating ponds, especially in areas of poor (sandy/loam) soils or high temperatures, evaporation and seepage are increased and 1–3% of the fish pond volume may be lost in this way each day (up to 45 $m^3\ kg^{-1}$ produced in ponds) (Verdegem *et al.*, 2006; Dugan *et al.*, 2007). Such losses may be particularly significant in arid or semi-arid areas of the world, such as Israel, where fish pond design and management practices have had to be changed in order to reduce surface water losses. Conversely, the incorporation of a fish pond into small rural farms has been shown to improve water conservation (i.e., creating a water reservoir) (Dey *et al.*, 2007). In addition, recent analysis shows that indirect freshwater consumption in aquaculture can be significantly higher than direct consumption (Verdegem *et al.*, 2006). This is because formulated feeds indirectly consume large quantities of water through crop and livestock production. About 1.2 m^3 of water is needed to produce 1 kg of grain. Fish and crustaceans require less grain compared to other terrestrial animal production, and even if water consumption through feed is large total freshwater consumption by aquaculture is small compared to , for example, agriculture. Total freshwater use depends on system and practice but it has been estimated that the 8,750,000 ha freshwater and 2,333,000 ha brackish water ponds for aquaculture consume approximately 429 $km^3\ yr^{-1}$ (16.9 $m^3\ kg^{-1}$), which represents only 3.6% of flowing water globally (Verdegem and Bosma, 2009).

Intensification of the aquaculture sector has led to an increasing dependence on external energy inputs throughout the production chain. Feed is commonly the main energy demanding source for fed aquaculture systems, while pumping, water purification and aeration may contribute significantly to 'closed' systems (Troell *et al.*, 2004; Henriksson *et al.*, 2012). Transportation, in turn, often accounts for a surprisingly small fraction of energy inputs (less than 1% for Norwegian salmon) (Pelletier *et al.*, 2009) while holding aquaculture products at warehouses can be an energy consumption hotspot for live organisms such as mussels and abalone (Iribarren *et al.*, 2010). Cumulative energy demand has been shown to provide a good indicator of both carbon footprint and ecological footprint while the true environmental consequences related to energy production depend largely on the energy carrier and country of production (Huijbregts *et al.*, 2010).

Wild Capture of Larvae and Spawners

The farming of shrimp and fish depends on larvae collected from the wild or reared in hatcheries from eggs of wild or farmed broodstock or spawners (Figure 2). Although most of the aquaculture industry today relies on hatchery-produced fry and fingerlings derived from parents bred in captivity, some tropical marine fish and shrimp culture still depends on capture of wild broodstock or juveniles (i.e., India, Bangladesh, The Philippines, Indonesia, etc.). Besides adverse effects on wild stocks of the target species, large bycatch of other larvae are killed in the process, representing losses to capture fisheries and biodiversity (Naylor *et al.*, 2000; Ahmed *et al.*, 2010). The quantities of bycatch associated with such wild catches are directly proportional to the natural abundance of the target species for culture. For example, milkfish *Chanos chanos* constitute only 15% of total fry in inshore collections by seine net (Bagarinao and Taki, 1986). The annual utilization of approximately 1.7 billion wild fry for stocking in Philippine milkfish ponds in the end of the 1990s (Bagarinao, 1997; 1998) corresponded to a loss of nearly 10 billion fry of other fish species. Many shrimp and prawn ponds in India and Bangladesh are still to a large extent dependent on stocking with wild-caught seed of *Penaeus monodon* and *Macrobrachium rosenbergeii* (Hoq *et al.*, 2001; Ahmed *et al.*, 2010). The peak period for this fishery reached one billion *P. monodon* seed collected annually in southeast Bangladesh (Dev *et al.*, 1994). Up to 900 fish and other shrimp fry species are discarded for every targeted shrimp being collected from estuarine and coastal waters in Bangladesh (Ahmed *et al.*, 2010). This is sufficiently large to potentially cause major impacts on biodiversity and capture fisheries production, but to date remains unstudied.

The development of hatcheries for cultured shrimp and marine fish species (e.g., milkfish) has reduced dependence on wild seed. However, it has also increased the demand for wild-caught mature (spawners) and immature (broodstock) shrimp adults (Rönnbäck *et al.*, 2003). For example, because the species is rare, wild collection of *P. monodon* broodstock and spawners can lead to large amounts of bycatch, with consequences for fisheries and biodiversity.

Impacts of Wastes, Chemicals, Diseases, and Feral Animals

The term 'wastes' in the current context is used to mean not only food, fecal and urinary products, and chemicals, but also microorganisms, parasites, and feral animals that may be introduced and subsequently thrive in aquaculture environments which may subsequently be released to the wider environment (Figure 1) (Beveridge, 2004; Hargrave, 2005). The release of uneaten food and fecal and urinary wastes may lead to eutrophication and oxygen depletion, the magnitude of the

impact depending on the type and size of operation and the nature of the site, ecosystem characteristics and assimilative capacity. Local effects from nutrient release into marine ecosystems, from coastal and near-shore intensive aquaculture operations, especially shrimp pond and fish cage farming, have been well studied (Islam, 2005). For example, eutrophic conditions from uncontrolled fish pen and cage operations in Pangasinan, Philippines, included increased ammonia, nitrate, nitrite, and phosphate concentrations, and low dissolved oxygen levels, leading to a major fish kill in 2002 valued at US$ 9 million (San Diego-McGlone *et al.*, 2008).

Sediment impacts associated with cage farming can be severe, resulting in anoxic sediments where fauna been lost. The effect on sediments is, however, mainly local and biodiversity will return a few years after production stops. However, studies documenting more large scale and long-term ecosystem changes from excessive nutrient emissions are few. Avoiding release of wastes, thereby minimizing environmental impacts, has been one of the drivers toward land-based 'closed' production systems. This has, however, come at the price of higher energy dependency (Henriksson *et al.*, 2012). However, semi-intensive, extensive, traditional, polyculture, and integrated systems generally assimilate much wastes internally (Beveridge, 2004; Edwards, 2009; Troell, 2009) resulting in fewer wastes being discharged to surrounding ecosystem.

Chemotherapeutants, including antimicrobial compounds and pesticides, are mainly used in intensive fish and shrimp cultures to control bacterial, fungal, and parasitic diseases. In shrimp farming, prohibited chemicals may be used due to lack of legislation or poor implementation of regulations. Farmers also use a range of vitamins, immunostimulants, disinfectants, and chemotherapeutants and employ chemicals for pond soil and water treatment (Gräslund and Bengtsson, 2001). The impacts of many of these chemicals are largely unknown. The release of antibiotics into the environment can affect microbial biodiversity and antimicrobial drug resistance but has reduced hugely in the farmed salmon industry due to the introduction of vaccines (Adams, 2009). Controls on use are increasingly strict, especially for aquaculture products aimed at the European and North American market (Beveridge *et al.*, 2010), but the controls target mainly the finalized products and not the aquaculture production process.

Impacts of aquaculture operations on biodiversity are strongly linked to introductions of exotic species and strains (Naylor *et al.*, 2001). Worldwide transfers and introductions of the few preferred shrimp culture species, including *P. monodon*, *Litopenaeus vannamei*, and *Marsupenaeus japonicus*, were numerous in the early decades of commercialized farming (Kautsky *et al.*, 2000). Such introductions and transfers may lead to competition with endemic fauna, genetic introgression with local fauna, and introduction of pathogens and parasites (Naylor *et al.*, 2001). Tilapia have had a long history of deliberate (for aquaculture and fisheries) and accidental introductions to some 90 countries and territories. However, recent studies suggest that impacts of tilapia introductions in Asia on aquatic ecosystem structure and function have been relatively small and often outweighed by the socioeconomic benefits (Gross, 1998; Arthur *et al.*, 2010). Atlantic salmon have escaped from culture facilities both within the geographic range of wild Atlantic salmon as well as in Pacific waters, and are now found as far north as the Bering Sea and as far south as Chile. Increasing evidence suggests that escapes may have direct genetic impacts on wild populations through hybridization (Walker *et al.*, 2006). Larger numbers of escapes also increase the likelihood of hybridization between feral farmed Atlantic salmon and wild fish in populations that are locally endangered or close to extinction (Slaney *et al.*, 1996; Gross, 1998; Hindar *et al.* 2006). In addition to consequences for the population gene pool and fitness, there are many potential ecological impacts associated with feral fish. Feral Atlantic salmon may compete extensively with wild salmon species for food and space, disturb native spawning sites, and introduce new diseases and parasites into wild populations (Jonsson and Jonsson, 2006).

The consequence of the large-scale introduction of the shrimp *P. vannamei* into Asia and the Pacific on biodiversity of indigenous cultured or wild shrimp populations is still uncertain as insufficient time and research have been conducted on this issue and there is therefore a need for caution (Briggs *et al.*, 2004). Recent field surveys in the Bangpakong Estuary in Thailand (where this species contributed >99% of total 2007 marine shrimp production) show that the species is persistent in the wild, and the presence of Taura Syndrome Virus (TSV) has been detected in seven local species (Senanan *et al.*, 2010). Moreover, laboratory studies show that *P. vannamei* can tolerate environmental conditions in the estuary and seeks food better than some of the indigenous species. The development of genetically modified species for aquaculture is currently being heavily debated (Kapuscinski *et al.*, 2007), and of particular interest is the 'super salmon' that has been developed by Aquabounty and is awaiting approval by the US Federal Drug Administration to be farmed commercially. The animal has a single copy of a DNA sequence that includes code for a Chinook growth gene as well as regulatory sequences derived from Chinook salmon and ocean pout (Marris, 2010).

Numerous wild salmon stocks in Canada and Norway have been infested by sea lice, either through the release of juvenile farmed salmon or large densities of fish cages in coastal salmon migration routes (Krkosek, 2010). The mechanisms and linkages to farmed salmon in Canada, however, is still being debated (Marty *et al.*, 2010).

Diseases are prevalent in many types of aquaculture, especially in highly stocked, intensive production systems, for example, Atlantic salmon, marine fish species, marine shrimp, etc. For example, the introduction of shrimp postlarvae and broodstock from areas affected by the White Spot Syndrome Virus (WSSV) and TSV resulted in the rapid spread of these pathogens throughout most of the shrimp-growing regions in Asia and Latin America, respectively (Kautsky *et al.*, 2000). A native of Asia, where it has caused multimillion-dollar shrimp crop losses, the WSSV has also been detected in wild and cultured shrimp in Texas and South Carolina. The virus was probably introduced by release of untreated wastes from plants processing imported Asian shrimp into coastal waters and by use of imported shrimp as bait in sport fishing or as fresh food for rearing other aquatic species (Lightner *et al.*, 1997). Another major shrimp virus, the Infectious Hypodermal and Hematopoietic Necrosis Virus (IHHNV), was probably introduced to the Americas from Asia through the import of live *P. monodon* in the early 1970s (Lightner *et al.*,

1999). In the Philippines, IHHNV prevalence in various wild populations of *P. monodon* has been correlated with intensification of shrimp culture methods and with mangrove status (Belak *et al.*, 1999). Higher incidence of viral infection in wild shrimp has been observed in areas with intensive shrimp farms and severely degraded mangroves.

The multibillion dollar Chilean salmon industry was during 2008 hit by Infectious Salmon Anemia (ISA) virus, resulting in millions of dead salmon and the layoff of thousands of fish farm workers. The main salmon farming areas in Chile are around Chiloe Island and this area is now being contaminated. Salmon producers are therefore reducing farm and fish densities as well as seeking to transfer their businesses further south. Whether the move will be successful or not is uncertain because this area has previously been almost free of farming operations due mainly to its remoteness and concern for the integrity of the seascape. Another factor that may compromise the successful translocation of salmon farming to the south is that the ISA virus already appears to be present in these waters.

In contrast to shrimp and salmon, comparatively few diseases have been reported for carps, tilapia, and milkfish, particularly from extensive and other low-density culture systems. The current trend toward intensification in rearing ponds and cages, however, may create stressful conditions through deterioration of water quality, excessive stocking, and polluted water inflow that predispose the fish to disease. The farming of *Pangasius* catfish in Vietnam has increased rapidly, reaching more than 1 mmt in 2008, and farming is characterized by holding high densities of fish in the ponds, resulting in increased risks for diseases and heavy usage of antibiotics (Phan *et al.*, 2009).

Resource Use and Carrying Capacity – Energy Analysis, Ecological Footprint and Life-Cycle Assessment

To reduce the risk of resource constraints and impacts on biodiversity, a shift to aquaculture production systems that use less valuable resources and emit wastes that do not exceed the assimilative capacity of the environment must occur. There is a need to recognize and manage nature's life support and its provision of ecosystem services, on which economic development and human welfare depends. Besides traditional EIA one way of identifying the broader demands for natural resources and ecosystem services of aquaculture has been to estimate the ecosystem area – the ecological footprint – functionally required to support the activity. When problems beset aquaculture operations, solutions focus on the pond or cage unit, and the fact that the farm is part of a much larger ecosystem with which it interacts is generally not considered. Surrounding ecosystems provide the feed, seed, clean water, and other essential resources and services, including waste assimilation. This unvalued work of nature sets the limits to culture levels without compromising biodiversity or causing pollution or disease problems. The footprint concept has proven to be very useful in illuminating the nonpriced and often unperceived work of nature that forms the basis for economic activities such as aquaculture. It is, however, not a detailed tool for measuring environmental capacity or sustainability in aquaculture, but an excellent tool for communicating human dependence on life-support systems.

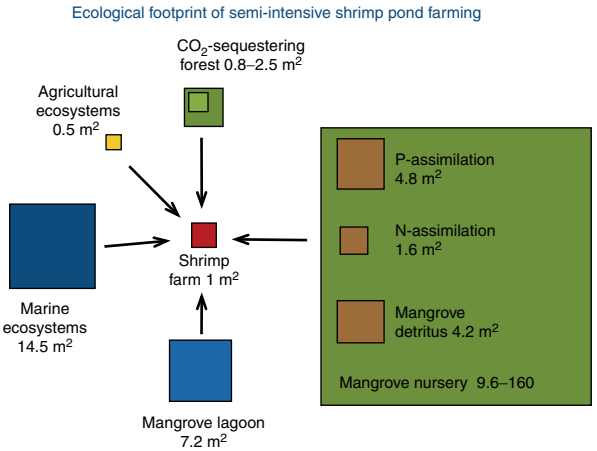


Figure 3 Ecosystem support areas required to sustain a semi-intensive shrimp farm in a coastal mangrove area of Columbia/ square meters of support area needed per square meter of pond area.

An illustration of how the footprint concept can be used for aquaculture is provided by Larsson *et al.* (1994). They estimated the spatial ecosystem support, or footprint, for a semi-intensive shrimp farm in a coastal mangrove area in Colombia. Support included food inputs, nursery areas, clean water, as well as waste processing and the support area was 35–190 times the surface area of the farm (Figure 3). The details of the footprint include: mangrove nursery area required to produce the shrimp seed for stocking – up to 160 times the pond area; if located close to the farm, the same mangrove area could also supply natural food inputs (4.2 m^2 m^{-2} shrimp pond area) and absorb polluting nutrients (2–22 m^2 m^{-2} pond area) in the farm effluents; feed pellets, a major input to a shrimp farm, needed a marine area of 14.5 m^2 to produce the fish, and an additional agricultural area of 0.5 m^2 for the agriculture ingredients used in feed pellets; 7.2 m^2 was needed to provide clean lagoon water to the ponds; and 0.8–2.5 m^2 of forest area per square meter shrimp pond area was needed to sequester the CO_2 associated with fossil fuel consumption at the farm.

Footprint size changes with farming intensity, that, is, a higher stocking density requires more food inputs and produces more wastes. However, less intensive production operations stock lower densities of species and thus require larger areas to produce the same biomass of produce. Pressure on local ecosystems can be reduced to some extent by importing some inputs (e.g., feeds) from other areas and by investing in shrimp hatcheries. Although producing seed in hatcheries considerably reduces shrimp and fish larvae bycatch that would otherwise be recruited populations exploited by fisheries, it also increases demand for wild-caught spawners and broodstock (Rönnbäck *et al.*, 2003). Other ecosystem services, however, such as clean water supply and waste assimilation, must be located close to the farming area. Up to a certain level of farming intensity this may pose little problem, but the

whole operation may collapse when the dynamic carrying/assimilating capacity of the local environment is exceeded unless extensive and costly water treatment facilities are built. The footprint concept provides an early warning device when the level of carrying capacity is being approached. Integrated farming technologies that re-circulate resources and wastes within the farm may be one way of reducing the footprint (Troell, 2009).

Aquaculture has directly contributed to the loss of important ecosystem functions (and biodiversity) through land and seascape transformation, and also more indirectly through, for example, pollution. However, aquaculture has also enhanced provisioning services, both in the agriculture landscapes and in the seascape, thus leading to improved welfare through livelihood diversification. Aquaculture can be a viable substitute for some sectors of today's terrestrial animal production (i.e., as an important source of micronutrients and animal-based protein and lipids), proven to be highly resource consuming (e.g., beef feedlot systems). The question is how to balance the negative and positive consequences from aquaculture development. The landscape and seascape are today increasingly being managed for multiple functions and services in addition to provision of food and fiber, and this requires the integration of ecological and socioeconomic research, policy innovation, and public education. The recently developed 'Ecosystem approach to aquaculture' (FAO,

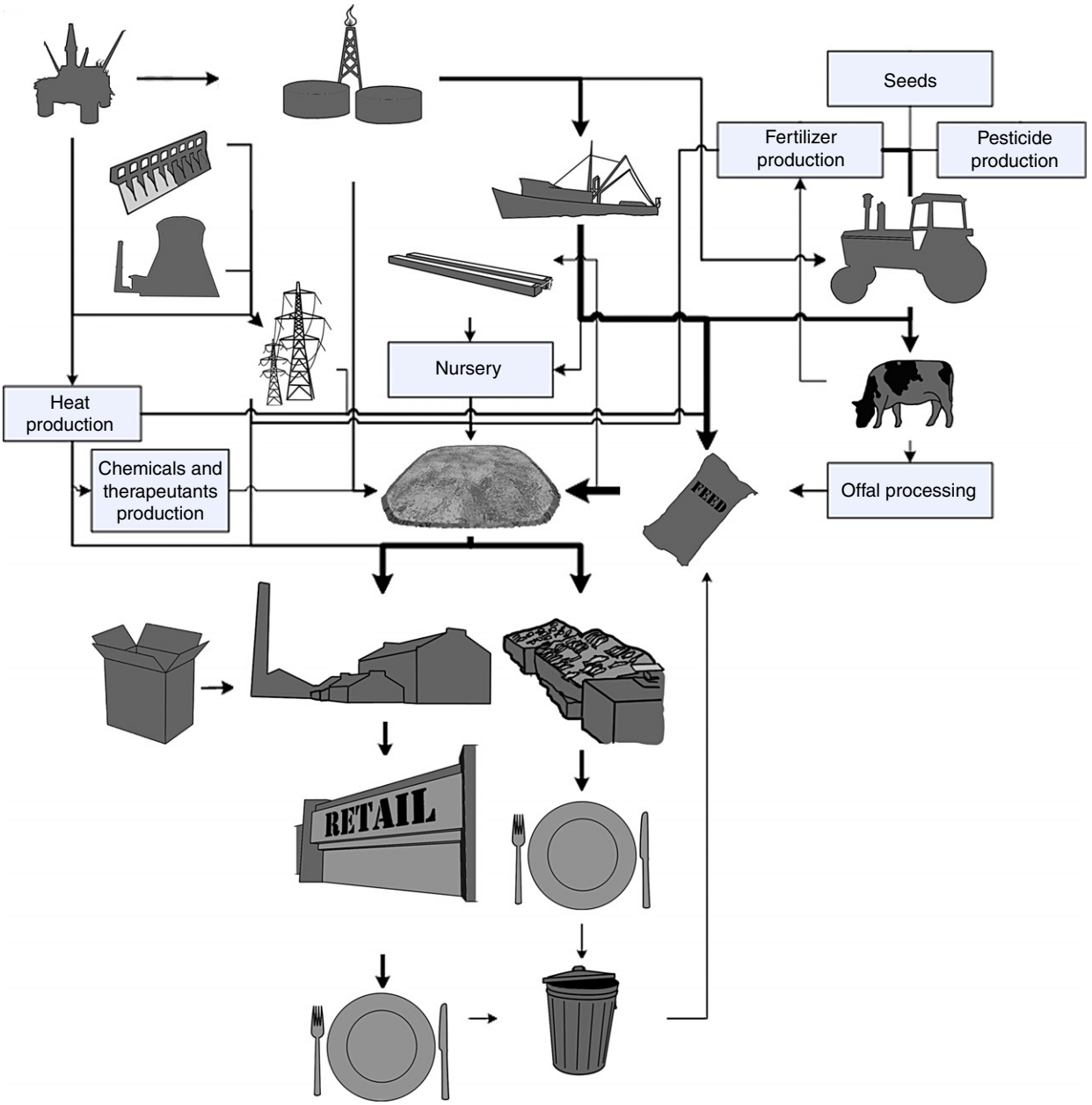


Figure 4 Illustration of system boundaries being identified in Life Cycle Analysis of an aquaculture production system.

2010) is an example of a broader systematic perspective on aquaculture development. This is a 'strategy for the integration of the activity within the wider ecosystem in such a way that it promotes sustainable development, equity, and resilience of inter-linked social and ecological systems', and could potentially force changes in human behavior with respect to understanding ecosystems functioning and the need for developing institutions capable of integrating different sectors at multiple scales.

Life cycle analysis (LCA) is a standardized eco-efficiency measurement framework that is increasingly being used to build inventories and estimate one or several environmental impacts in aquaculture (Pelletier *et al.*, 2007; Henriksson *et al.*, 2012). LCA identifies system boundaries at the broadest scale, both temporal and spatial, and is a much more quantitative tool than Ecological Footprints (Figure 4). Thus, the width of environmental impacts that can be measured allows for more detailed estimates of the environmental interactions. Climate change, habitat change, pollution and overexploitation are impact categories commonly entailed within the LCA framework (Guinèe, 2002), addressing all major drivers, except invasive species, for biodiversity loss, identified by the Millennium Ecosystem Assessment (2005).

Integrated Aquaculture

Integrated aquaculture may offer opportunities for the efficient usage of water and utilization of nutrients, and increased productivity and profits, providing in a single package practical and creative solutions to some problems of waste management and pollution (Neori *et al.*, 2004). Thus, the resulting environmental impacts from aquaculture and various resource limitations (water, feed, energy, etc.) may find their solutions in integrated cultivation techniques (Troell *et al.*, 2004). Traditional inland aquaculture, comprising diverse systems that use on-farm or locally available resources, still dominates in Asia, but they are increasingly being replaced by industrial-based aquaculture technologies (Edwards, 2009). Traditional polyculture/integrated systems make efficient use of inputs and generate less waste, thus adding to net food supplies locally and regionally at relatively minor environmental and social expense. Traditional pond cultures of herbivorous fish species (e.g., carps in China) have been viable for centuries and their existence is the proof of sustainable integrated farming systems. Here, raising poultry and livestock is integrated with fish farming, and the principal linkages between the systems are animal manure and other agriculture waste products. Compost is used to fertilize the pond water for proliferation of natural organisms as natural feeds for fish from juvenile to adult. However, despite its environmental advantages the increased global demand for food cannot be met by traditional extensive production systems (due to demand on space and to low productivity). Even though technological development and improved management has resulted in increased efficiency and environmental performance of some intensive single species aquaculture systems, we need to ask what information embedded in traditional integrated systems might be useful for aquaculture development. This information could then be added to recent research on integrated aquaculture systems. Integrated aquaculture is certainly not a panacea for aquaculture development, but should be looked upon as one potential approach among many others facilitating sustainable development.

The concept of ecological engineering has gained interest in an aquaculture context, with negative environmental effects being remedied by species integration for nutrient trapping or recirculation. The recent development and promotion of integrated aquaculture in coastal areas has focused on modern integrated approaches, mainly from temperate and subtropical regions (China, Canada, South Africa, the Mediterranean Sea, i.e., IMTA systems (Integrated Multi-trophic Aquaculture)) involving integration of fish, seaweeds, crustaceans, and mollusks (Chopin *et al.*, 2002, 2008; Neori *et al.*, 2004;). Such systems try to tackle the throughput characteristics of many farming systems, in which large amounts of wastes are released to the environment. Thus, in addition to output of particulate wastes, aquaculture also releases dissolved nutrients, and generally less than one-third of the nutrients added through feed will be removed through harvest in intensive fish and shrimp farming (Hargrave, 2005; Islam, 2005). Treatment of effluents usually involves a high degree of technology, and therefore high costs, which implies that release of untreated water is the rule rather than the exception (especially in many developing countries). Besides the improved environmental performance, the benefits from integration include economic gains resulting from co-production of different products.

Tropical integrated mariculture systems involve various types of polyculture, sequential integration, temporal integration, and different integration with mangroves (Troell, 2009). An example of the latter is the use of mangroves as biofilters to process effluents from aquaculture ponds. Results from the central Philippines indicate that 1.8–5.4 ha of mangroves are needed to assimilate nitrate wastes from 1 ha of shrimp pond (Primavera *et al.*, 2007). Mangroves and aquaculture are not necessarily incompatible and guides for responsible use of mangrove areas for aquaculture have been developed (Bagarinao and Primavera, 2005; Primavera *et al.*, 2009). However, any development of aquaculture within or nearby mangroves needs to be carefully evaluated to identify net benefits and impacts on environment, biodiversity, ecosystem services, and implications for a variety of stakeholders.

Conclusions

Since the principle of aquaculture is to reroute flows of energy and matter from the ecosystem into those species that we culture, aquaculture, like agriculture, will always affect the environment to some extent. This interaction and alteration of supporting environment is unavoidable, but it should not be done in a fashion that results in the capacity of ecosystems to sustain social and economic development to be diminished. Recent research has revealed that aquaculture systems are strongly coupled to nature's subsidies and services to sustain production, and for many species this subsidy expands beyond the local to the global scale.

Conversion of ecosystems, the extraction of feed resources, collection of larvae and broodstock, spread of exotic species, diseases and effluents, may all add to a spiral whereby total

fish supplies are reduced over time. Thus, the real challenge for aquaculture is to develop farming practices that are in tune with ecosystem processes and functions in a fashion that enhances aquatic food production. There is great potential to develop techniques that work with nature's dynamics. There is also potential to redirect unsustainable modes of production into practices that contribute to and enhance nature's support capacity not only for aquaculture but also for other human activities dependent on aquatic ecosystems. There is no doubt that the aquaculture sector will move in this direction. Governmental policies and institutional frameworks are required that can make such a transition possible. The internalization of the costs of deterioration of supporting areas caused by farming will create incentives for the industry to take a more sustainable path (Folke *et al.*, 1994). This should of course be something that all food production systems (i.e., fisheries, crop, and animal agriculture) strive towards.

The role of the consumers may be critical in shaping farm management practices. Aquaculture products are increasingly traded through multinational supermarkets that are highly responsive to customer opinion and demands. If farmed aquatic foods become associated in the public's mind with poor environmental management and its direct effects on biodiversity, then supermarkets may well refuse to stock the produce. This is today a reality and many labeling schemes are under development (e.g., Aquaculture Dialogue, WWF).

See also: Carrying Capacity, Concept of. Census of Marine Life. Crustaceans. Ecological Footprint, Concept of. Ecosystem, Concept of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Function, Principles of. Introduced Species, Impacts and Distribution of. Mangrove Ecosystems. Marine and Aquatic Communities, Stress from Eutrophication. Resource Exploitation, Fisheries. Salmon

References

Adams A (2009) Advances in disease diagnosis, vaccine development and other emerging methods to control pathogens in aquaculture. In: Burnell G and Allan G (eds.) *New Technologies in Aquaculture*, pp. 197–214. Boca Raton, FL: CRC.

Ahmed N, Troell M, Allison EH, and Muir JF (2010) Prawn postlarvae fishing in coastal Bangladesh: Challenges for sustainable livelihoods. *Marine Policy* 34: 218–227.

Arthur RI, Lorenzen K, Homekingkeo P, *et al.* (2010) Assessing impacts of introduced aquaculture species on native fish communities: Nile tilapia and major carps in SE Asian freshwaters. *Aquaculture* 299: 81–88.

Bagarinao T (1997) The milkfish fry shortage in the Philippines and the supply from fisheries and hatcheries. *UPV Journal of Natural Science* 2: 146–166.

Bagarinao T (1998) Historical and current trends in milkfish farming in the Philippines. In: de Silva S (ed.) *Tropical Mariculture*, pp. 381–448. London: Academic Press.

Bagarinao T and Taki Y (1986) The larval and juvenile fish community in Pandan Bay, Panay Island, Philippines. In: Uyeno T, Arai R, Taniuchi T, and Matsuura K (eds.) *Indo-Pacific Fish Biology: Proceedings of the Second International Conference on Indo-Pacific Fishes*, pp. 728–739. Tokyo: Ichthyological Society of Japan.

Bagarinao TU and Primavera JH (2005) *Code of Practice for Sustainable Use of Mangrove Ecosystems for Aquaculture in Southeast Asia*. Iloilo, Philippines: SEAFDEC Aquaculture Department.

Barbier EB, Koch EW, Silliman BR, *et al.* (2008) Coastal ecosystem-based management with nonlinear ecological functions and values. *Science* 319: 321–323.

Belak J, Dhar AK, Primavera JH, *et al.* (1999) Prevalence of viral diseases (IHHNV and WSSV) in Penaeus monodon from the Philippines andits association with mangrove status and shrimp culture systems. In: Alcivar-Warren A (ed.) *Proceedings of the Aquaculture and Conservation of Marine Shrimp Biodiversity Symposium*. North Grafton, MA: Tufts University School of Veterinary Medicine.

Beveridge MCM (2004) *Cage Aquaculture*, 3rd edn. Oxford: Blackwells.

Beveridge MCM, Phillips MJ, Dugan P, and Brummett R (2010) Barriers to aquaculture development as a pathway to poverty alleviation and food security. In: OECD (ed.) *Advancing the Aquaculture Agenda: Workshop Proceedings*, pp. 345–359. Paris: OECD Publishing.

Boyd CE, McNevin AA, Clay J, and Johnson HM (2005) Certification issues for some common aquaculture species. *Reviews in Fisheries Science* 13: 231–279.

Briggs M, Funge-Smith S, Subasinghe R, and Phillips M (2004) *Introductions and movement of* Penaeus vannamei *and* Penaeus stylirostris *in Asia and the Pacific*. Bangkok, Thailand: FAO Regional Office for Asia and the Pacific.

Chopin T, Buschmann A, Halling C, *et al.* (2002) Integrating seaweeds into aquaculture systems: A key towards sustainability. *Journal of Phycology* 37: 975–986.

Chopin T, Robinson SMC, Troell M, Neori A, Buschmann A, and Fang JG (2008) Ecological engineering: Multi-trophic integration for sustainable marine aquaculture. In: Jorgensen SE and Fath B (eds.) *Encyclopedia of Ecology*, pp. 2463–2475. Amsterdam: Elsevier.

Cowey CB and Sargent JR (1977) Lipid nutrition in fish. *Comparative Biochemistry and Physiology* 57: 269–273.

Deutsch L, Troell M, Limburg K, *et al.* (2011) Trade of Fisheries Products–implications for marine ecosystems and their services. In: Köllner T (ed.) *Ecosystem Services and Global Trade of Natural Resources Ecology, Economics and Policies*, 304 pp. Routledge.

Dev AK, Das NG, and Alam MM (1994) Colossal loss of shell-fish and fin-fish postlarvae for indiscriminate catch of Penaeus monodon fry along the Cox's Bazaar – Teknaf coast of Bangladesh. In: Wells PG and Ricketts J (eds.) *Coastal Zone Canada 94, 'Cooperation in the Coastal Zone' Conference Proceedings*, pp. 1530–1546. Dartmouth, Nova Scotia, Canada: Bedford Institute of Oceanography.

Dey MM, Kambewa P, Prein M, *et al.* (2007) World Fish Centre. Impact of the Development and Dissemination of Integrated Aquaculture Technologies in Malawi. In: Waibel H and Zilberman D (eds.) *International Research on Natural Resource Management*, pp. 118–140. Rome: FAO, and Cambridge: CAB International.

Diana JS (2009) Aquaculture production and biodiversity conservation. *Bioscience* 59: 27–38.

Dugan P, Sugunan VV, Welcomme RL, *et al.* (2007) Inland fisheries and aquaculture. In: Molden, *et al.* (eds.) *Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture*, pp. 459–483. London: Earthscan and Colombo: International Water Management Institute.

Edwards P (2009) Traditional Asian aquaculture: Definition, status and trends. In: Burnell G and Allan G (eds.) *New Technologies in Aquaculture, Improving Production Efficiency, Quality and Environmental Management*, pp. 1029–1063. Cambridge, UK: Woodhead Publishing Ltd.

FAO (1997) Empirical Investigation on the Relationship Between Climate and Small Pelagic Global Regimes and El Niño-Southern Oscillation (ENSO), Food and Agriculture Organization of the United Nations, Rome: FAO Fisheries Circular No. 934.

FAO (2010) *The State of World Fisheries and Aquaculture 2010*. Rome: FAO.

Folke C, Kautsky N, and Troell M (1994) Cost of eutrophication in Salmon farming. *Environmental Management* 40: 173–182.

GESAMP (IMO/FAO/UNESCO-IOC/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection) (1991) Reducing environmental impacts of coastal aquaculture. Reports and Studies. GESAMP 47. Rome: FAO.

Gräslund S and Bengtsson BE (2001) Chemicals and biological products used in South-East Asian shrimp farming, and their potential impact on the environment – a review. *The Science of the Total Environment* 280: 93–131.

Gross MR (1998) One species with two biologies: Atlantic salmon (*Salmo salar*) in the wild and in aquaculture. *Canadian Journal of Fisheries and Aquatic Science* 55: 131–144.

Guinèe JB (ed.) (2002) *Handbook on Life Cycle Assessment: Operational Guide to the ISO Standards*. Dordrecht: Kluwer Academic Publishers.

Hansen DJ (1997) Shrimp fishery and capelin decline may influence decline of harbor seal (*Phoca vitulina*) and northern sea lion (*Eumatopias jubatus*) in

western Gulf of Alaska. In: *Proceedings of the International Symposium on the Role of Forage Fishes in Marine Ecosystems*, pp. 197–207. Sea Grant College Program Report No. 97-01. Fairbanks, AK: University of Alaska

Hargrave B (ed.) (2005) *Environmental Effects Of Marine Finfish Aquaculture: The Handbook Of Environmental Chemistry,* vol. 5. Springer-Verlag

Hasan MR and Halwart M (2009) *Fish as Feed Inputs for Aquaculture: Practices, Sustainability and Implications*. Rome: FAO; FAO Fisheries and Aquaculture Technical Paper. No. 518.

Henriksson P, Pelletier N, Troell M, Tyedmers P (2012) Life cycle assessment and its application to aquaculture production systems. In: Meyers RA (ed.) *Encyclopedia of Sustainability Science and Technology.* 1st Edition. New York: Springer-Verlag.

Hindar K, Fleming IA, McGinnity P, and Diserud O (2006) Genetic and ecological effects of salmon farming on wild salmon: Modeling and experimental results. *ICES Journal of Marine Science* 63: 1234–1247.

Hoq ME, Islam MN, Kamal M, and Wahab MA (2001) Abundance and seasonal distribution of Penaeus monodon postlarvae in the Sundarbans mangrove, Bangladesh. *Hydrobiologia* 457: 97–104.

Huijbregts M, Hellweg S, Frischknecht R, *et al.* (2010) Cumulative energy demand as predictor for the environmental burden of commodity production. *Environmental Science Technology* 44: 2189–2196.

Huntington TC and Hasan MR (2009) Fish as feed inputs for aquaculture – practices, sustainability and implications: A global synthesis. In: Hasan MR and Halwart M (eds.) *Fish as Feed Inputs for Aquaculture: Practices, Sustainability and Implications*, FAO Fisheries and Aquaculture Technical Paper. No. 518, pp. 1–61. Rome: FAO.

Iribarren D, Moreira M, and Gumersindo F (2010) Revisiting the life cycle assessment of mussels from a sectorial perspective. *Journal of Cleaner Production* 18: 101–111.

Islam MS (2005) Nitrogen and phosphorus budget in coastal and marine cage aquaculture and impacts of effluent loading on ecosystem: Review and analysis towards model development. *Marine Pollution Bulletin* 50: 48–61.

Jonsson B and Jonsson N (2006) Cultured Atlantic salmon in nature: A review of their ecology and interactions with wild fish. *ICES Journal of Marine Science* 63: 1162–1181.

Kapuscinski AR, Hayes KR, Li S, and Dana G (eds.) (2007) *Environmental Risk Assessment of Genetically Modified Organisms*, vol. 3. Methodologies for Transgenic Fish. Oxford: CAB International

Kaushik S and Troell M (2010) Taking the fish-in fish-out ratio a step further. *Aquacul*ture *Europe* 35: 15–17.

Kautsky N, Rönnbäck P, Tedengren M, and Troell M (2000) Ecosystem perspectives on management of disease in shrimp pond farming. *Aquaculture* 191: 145–161.

Krkosek M (2010) Sea lice and salmon in Pacific Canada: Ecology and policy. *Frontiers in Ecology and the Environment* 8: 201–209.

Larsson J, Folke C, and Kautsky N (1994) Ecological limitations and appropriation of ecosystem support by shrimp farming in Colombia. *Environmental Management* 18: 663–676.

Lightner DV, Redman RM, Poulos BT, *et al.* (1997) Risk of spread of penaeid shrimp viruses in the Americas by the international movement of live and frozen shrimp. *Review Science and Technology* 16: 146–160.

Lightner DV, Redman RM, Poulos BT, *et al.* (1999) The penaeid shrimp viruses TSV, IHHNV, WSSV and YHV: Current status in the Americas, available diagnostic methods and management strategies. *Journal of Applied Aquaculure.* 9: 27–52.

Marris E (2010) Transgenic fish go large. *Nature* 467: 259.

Marty GD, Saksida SM, and Terrance JQII (2010) Relationship of farm salmon, sea lice, and wild salmon populations. *Proceedings of the National Academy of Sciences* 2010.http://dx.doi.org/10.1073/pnas.1009573108.

Menasveta P (1997) Mangrove destruction and shrimp culture systems. *World Aquaculture* 28: 36–42.

Millennium Ecosystem Assessment (2005) Current state and trends 2005. http://www.maweb.org

Moberg F and Folke C (1999) Ecological goods and services of coral reef ecosystems. *Ecological Economics* 29: 215–233.

Moberg F and Rönnbäck P (2003) Ecosystem services in the tropical seascape: Ecosystem interactions, substituting technologies, and ecosystem restoration. *Ocean and Coastal Management* 46: 27–46.

Nagelkerken I (ed.) (2009) *Ecological Connectivity Among Tropical Coastal Ecosystems*, 615 p. London: Springer.

Naylor R, Goldberg R, Primavera J, *et al.* (2000) Effects of aquaculture on world fish supplies. *Nature* 405: 1017–1024.

Naylor RL, Hardy RW, Bureau DP, *et al.* (2009) Feeding aquaculture in an era of finite resources. *Proceedings of the National Academy of Sciences* 106: 15103–15110.

Naylor RL, Williams SL, and Strong DR (2001) Aquaculture – a gateway for exotic species. *Science* 294: 1655–1656.

Neori A, Chopin T, Troell M, *et al.* (2004) Integrated aquaculture: Rationale, evolution and state of the art emphasizing seaweed biofiltration in modern mariculture. *Aquaculture* 231: 361–391.

NRC (National Research Council) (1999) *Sustaining Marine Fisheries*. Washington, DC: National Academy Press.

Pauly D and Ingles J (1986) The relationship between shrimp yields and intertidal vegetation (mangrove) areas: A reassessment. In: Yáñez-Arancibia A and Lara-Domínguez AL (eds.) *Ecosistemas de Manglar en América Tropical. Instituto de Ecología A.C. México, UICN/ORMA, Costa Rica* MD, NMFS Silver Spring, USA: NOAA.

Pauly D and Tsukayama I (eds.) (1987) *The Peruvian Anchoveta and its Upwelling Ecosystem: Three Decades of Change*. Manila, The Philippines: ICLARM.

Pelletier N, Ayer N, Tyedmers PN, *et al.* (2007) Impact categories for Life Cycle Assessment research of seafood production: Review and prospectus. *International Journal of Life Cycle Assessment* 12: 414–421.

Pelletier N, Tyedmers P, Sonesson U, *et al.* (2009) Not all salmon are created equal: Life cycle assessment (LCA) of global salmon farming systems. *Environmental Science and Technology* 43: 8730–8736.

Phan LT, Bui TM, Nguyen TTT, *et al.* (2009) Current status of farming practices of striped catfish, Pangasianodon hypophthalmus in the Mekong Delta, Vietnam. *Aquaculture* 296: 227–236.

Primavera JH, Altamirano JP, Lebata MJHL, delos Reyes Jr. AA, and Pitogo CL (2007) Mangroves and shrimp pond culture effluents in Aklan, Panay Island central Philippines. *Bulletin of Marine Science* 80: 795–804.

Primavera JH, Binas JB, Samonte-Tan GPB, *et al.* (2009) Mud crab penculture – replacement of fish feed requirement and impacts on mangrove community structure. *Aquaculture Research* 41: 1211–1220.

Rana KJ, Siriwardena S, and Hasan MR (2009) *Impact of Rising Feed Ingredient Prices on Aquafeeds and Aquaculture Production*, FAO Fisheries and Aquaculture Technical Paper 541. Rome: FAO.

Robertson AI and Duke NC (1987) Mangroves as nursery sites: Comparisons of the abundance and species composition of fish and crustaceans in mangroves and other nearshore habitats in tropical Australia. *Marine Biology* 96: 193–205.

Rönnbäck P (1999) The ecological basis for the economic value of Mangrove forests in seafood production. *Ecological Economics* 29: 235–252.

Rönnbäck P, Bryceson I, and Kautsky N (2002) Coastal aquaculture development in Eastern Africa and the Western Indian Ocean: Prospects and problems for food security and local economies. *Ambio* 31: 537–542.

Rönnbäck P, Troell M, Zetterström T, and Babu DE (2003) Mangrove dependence and socio-economic concerns in shrimp hatcheries of Andhra Pradesh, India. *Environmental Conservation* 30: 344–352.

San Diego-McGlone ML, Azanza RV, Villanoy CL, and Jacinto GS (2008) Eutrophic waters, algal bloom and fish kill in fish farming areas in Bolinao, Pangasinan, Philippines. *Marine Pollution Bulletin* 57: 295–301.

Senanan W, Panutrakul S, Barnette P, *et al.* (2010) Ecological risk assessment of an alien aquatic species: A case study of *Litopenaeus vannamei* (Pacific Whiteleg Shrimp) Aquaculture in the Bangpakong River, Thailand. In: Hoanh CT, Szuster BW, Suan-Pheng K, Ismail AM, and Noble AD (eds.) *Tropical Deltas and Coastal Zones. International Water Management Institute*, pp. 64–79. CABI, UK and Cambridge, MA: WorldFish Center, International Rice Research Institute, Food and Agriculture Organization of the United Nations – Regional Office for Asia and the Pacific, and CGIAR Challenge Program on Water and Food.

Slaney TL, Hyatt KD, Northcote TG, and Fielden RJ (1996) Status of anadromous salmon and trout in British Columbia and Yukon. *Fisheries* 21: 20–35.

Smith AD, Brown CJ, Bulman CM, *et al.* (2011) Impacts of fishing low-trophic level species on marine ecosystems. *Science* 21 July.

Stickney RR, Hardy RW, Koch K, *et al.* (1996) The effect of substituting selected oil seed protein concentrates for fishmeal in rainbow trout *Oncorhynchus mykissdiets. Journal of World Aquaculture Society* 27: 57–63.

Tacon AGJ and Metian M (2008) Global overview on the use of fish meal and fish oil industrially compounded aquafeeds: Trends and future prospects. *Aquaculture* 285: 146–158.

Tacon AGJ and Metian M (2009) Fishing for aquaculture: Non-food use of small pelagic forage fish – A global perspective. *Reviews in Fisheries Science* 17: 305–317.

Troell M (2009) Integrated marine and brackishwater aquaculture in tropical regions; research, implementation and prospects. In: Soto D (ed.) *Integrated Mariculture: A Global Review*, FAO Fisheries and Aquaculture Technical Paper No. 529. Rome: FAO.

Troell M, Tydemer P, Rönnbäck P, and Kautsky N (2004) Aquaculture – energy use. In: Cleveland C (ed.) *Encyclopedia of Energy*, pp. 97–108. Elsevier Inc.

Tuan MS (1997) Building up the strategy for mangrove management in Vietnam. In: Hong PN, Ishwaran N, San HT, Tri NH, and Tuan MS (eds.) *Proceedings of Ecotone V, Community Participation in Conservation, Sustainable Use and Rehabilitation of Mangroves in Southeast Asia*, pp. 244–255. Vietnam: UNESCO, Japanese Man and the Biosphere National Committee and Mangrove Ecosystem Research Centre.

Vader W, Barret RT, Erikstad KE, and Strann K-B (1990) Differential responses of common and thick-billed murres to a crash in the capelin stock in the southern Barents Sea. *Scientia Marina* 67(Suppl. 2): 33–45. Studies in Avian Biology 14, 175-180.

Verdegem MCJ and Bosma RH (2009) Water withdrawal for brackish and inland aquaculture, and options to produce more fish in ponds with present water use. *Water Policy* 11: 52–68.

Verdegem MCJ, Bosma RH, and Verreth JAV (2006) Reducing water use for animal production through aquaculture. *Water Resources Development* 22: 101–113.

Walker AM, Beveridge MCM, Crozier W, O'Maoleidigh N, and Milner N (2006) Monitoring the incidence of escaped farmed Atlantic salmon, *Salmo salar* L, in rivers and fisheries of the United Kingdom and Ireland: Current progress and recommendations for future programmes. *ICES Journal of Marine Science* 63: 1201–1210.

Walters B, Rönnbäck P, Kovacs J, *et al.* (2008) Ethnobiology, socio-economics and management of Mangroves: A Review. *Aquatic Botany* 89: 220–236.

Weber P (1993) Reviving the coral reefs. In: Brown LR (ed.) *State of the World*. New York: W.W. Norton.

Arachnids

Jonathan A Coddington, Smithsonian Institution, Washington, WA, USA
Robert K Colwell, University of Connecticut, CT, Storrs, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 199–218, © 2001, Elsevier Inc.

Glossary

Booklungs One to four pairs of abdominal respiratory organs consisting of a thin, multifolded membrane (the book's "pages") over which blood circulates and that is open to an air-filled cavity on the outside, itself open to the exterior via a spiracle. Gases passively diffuse back and forth across the membrane.

Chelicerae (chelate) The first pair of preoral appendages. They are at most three segmented, usually two, and usually the distal segment acts against the penultimate to grab or hold prey or objects. If the basal segment has a finger-shaped outgrowth against which the distal segment operates, the chelicerae are chelate (as in scorpions or harvestmen). If not, the chelicerae are subchelate, as in spiders and tailless whip scorpions. In parasitic mites the chelicerae are modified into piercing stylets.

Monophyly A true, historical, evolutionary lineage consisting of an ancestor and all of its descendants; defined by shared, derived characters.

Ovoviviparous Young are born alive, but the mother simply retains eggs within her body until they hatch.

Paraphyly A group consisting of an ancestor and only some of its descendants. Defined by primitive characters.

Pedipalps The second pair of preoral appendages. They are multisegmented and primitively leg-like. They may be raptorial or sensory (like antennae) or used as walking legs.

Phoresy A method of long-range dispersal in which the dispersing animal attaches itself to another animal (e.g., beetle, wasp, or bird) that carries the disperser along with it until the disperser drops off or disembarks.

Polyphyly A group in which the most recent common ancestor of the included taxa is excluded from the group. Defined by convergent, nonhomologous characters.

Spermatophore A chitinous container produced by the male to hold sperm. It may be attached to the substrate for the female to find or passed to the female from the male during mating.

Spinnerets Usually three, rarely four pairs of modified terminal abdominal appendages in spiders bearing one to hundreds of hollow spigots from which silk is drawn.

Tracheae A system of hollow, branched or unbranched air-conducting tubes used for respiration, opening via abdominal spiracles. They may or may not extend into the cephalothorax or legs.

Trichobothria Long, delicate, slender setae set in broad, shallow innervated sockets in the cuticle. Trichobothria are sensitive to vibration or near-field air movement and are a major sense organ of arachnids.

Overview of Arachnida

The known diversity of arachnids is approximately 640 families, 9000 genera, and 93,000 species (Table 1), but there are many thousands of new mite and spider species still undescribed and hundreds to thousands of undescribed species in the remaining orders. Together with the marine horseshoe crabs (Xiphosura) and sea spiders (Pycnogonida), arachnids

Table 1

Latin name[a]	*Common name*	*No. of Families*	*No. of Genera*	*No. of Described species*
Arachnida	Arachnids	648	9241	±92,680
Araneae	Spiders	108	3200	±37,000
Palpigradi	Micro-whip scorpions	2	6(5)	80
Uropygi	Whip scorpions	1	16	101
Amblypygi	Whip spiders, tailless whip scorpions	5	20	126
Schizomida	Schizomids	2	31	195
Solifugae	Wind spiders	12	153	1,065
Pseudoscorpiones	Pseudoscorpions	24	430	3,100
Ricinulei	Ricinuleids	1	3	53
Opiliones	Harvestman	44	1554	±4,500
Scorpiones	Scorpions	18	156	1,260
Acari	Mites, ticks	431	3672	45,200

[a]Modified from Adis J and Harvey MS (2000) How many Arachnida and Myriapoda are there worldwide and in Amazonia? *Stud. Neotropical Fauna Environ.* 35: 139–141, with data for the Acari from Walter DE and Proctor HC (1999) *Mites: Ecology, Evolution and Behaviour.* Wallingford: Univ. of New South Wales Press and CABI.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00150-7

comprise the arthropod subphylum Chelicerata, named for the characteristic first pair of preoral appendages, the chelate, or pinching, mouthparts. In some arachnid groups the chelicerae are further modified into venomous fangs or piercing stylets to suck body or plant fluids. Arachnids are the only terrestrial chelicerates. Along with the insects, arachnids are by far the most species-rich, abundant, and widely distributed terrestrial arthropods. Acarologists (scientists who study mites and ticks) estimate that there may be as many mite species as beetles, implying that total extant arachnid diversity may exceed 1 million species. Arachnids are an important component of every terrestrial ecosystem, but, apart from several specialized mite lineages, none are aquatic or marine. Although most arachnids share many ancestral similarities in body plan and lifestyle, many extremely specialized groups exist, especially among the mites (Acari).

Although arachnids are commonly mistaken for some sort of peculiar insect, the groups are quite distinct and only distantly related. Arachnids have four pairs of walking legs rather than three (except the young stages of mites and the related Ricinulei), only two (not three) major body parts, and simpler chelate mouth-parts rather than the more complex feeding apparatus of insects. The anterior body part is specialized for locomotion and the posterior for digestion and reproduction. Arachnids lack the wings, antennae, and compound eyes usual in insects. In many groups the first pair of walking legs (e.g., amblypygids, uropygids, schizomids, solifuges, palpigrades, and many mites) are elongate and function in much the same way as insect antennae. **Figure 1** depicts a consensus view of the position of arachnids in the arthropod evolutionary tree.

Arachnids and hexapods (which insects dominate) differ in fundamental ways, possibly because their marine ancestors were already distinct lineages in the Silurian and both colonized land independently. They therefore solved the fundamental challenges of terrestrial existence (support, breathing, water balance, reproduction in dry environments, and nitrogenous waste management) in different ways. The arachnid skeleton is hydraulic; arachnids, except scorpions and pseudoscorpions, lack extensor muscles at key joints. Instead, the animal pumps blood into the limb to extend it. The basic arachnid uses two (or four) pairs of book lungs (gas-permeable, gill-like membranes with blood on one side and open to the air on the other) to exchange carbon dioxide for oxygen rather than tracheae. Many arachnids possess rudimentary tracheal systems that supplement or replace book lungs, but ventilation is passive, not active. Unlike insect tracheae, arachnid tracheae generally do not ramify throughout the entire body or penetrate inside body cells. For all these reasons, insect tracheae more efficiently deliver oxygen directly to tissue, and insects in general can lead more energy-intensive lifestyles than arachnids. Arachnids have significantly lower metabolic rates than other terrestrial arthropods, especially insects. Whereas male insects transfer sperm directly to the female using an intromittent organ, in most arachnids males either ejaculate onto a special structure and carry the sperm mass with a specialized appendage until the female is encountered or they deposit the sperm mass in a specially built receptacle (spermatophore) fixed to the substrate, which the female picks up. Major exceptions include the harvestmen, astigmatic mites, and spider mites, which transfer sperm to the female by means of an intromittent organ, and the water mites (Prostigmata: Hydracarina), which transfer sperm directly by opposing the male gonopore to the female gonopore.

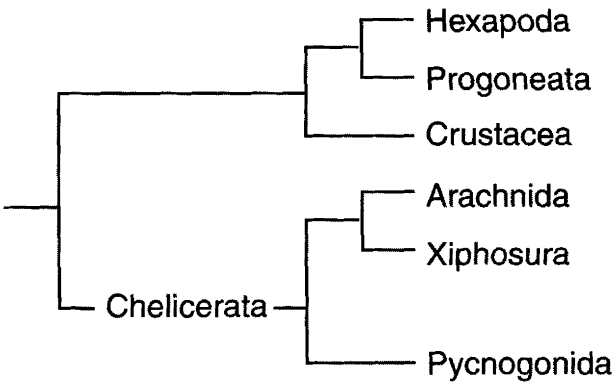


Figure 1 Phylogenetic position of Arachnida among arthropods.

Arachnids are peculiar among animals in using guanine (three nitrogen atoms per molecule) as well as the much more common uric acid (two nitrogen atoms per molecule) to eliminate nitrogenous wastes. Insects have compound eyes, which provide relatively excellent vision. Arachnids lost compound eyes early in their evolutionary history but retain usually one to five (commonly four) pairs of simple eyes, much inferior in acuity to compound eyes. Schizomids, palpigrades, ricinuleids, mites, and other mainly litter-dwelling arachnid groups are nearly always blind. Vision is much less important to arachnids than vibration. Many structures (slit sense organs, trichobothria, and lyriform organs) are specialized to detect minute vibrations and slight air currents.

Arachnids are also peculiar in that species of most arachnid groups digest food externally. They have a strong pumping stomach that rhythmically vomits and sucks digestive juice through a preoral cavity formed by the basal articles of the pedipalps back and forth over their prey. The process continues until only the hard, indigestible parts of the prey remain. Only liquids or very small particles are actually ingested. Some major groups of arachnids have internal digestion, however. Opilionids and several groups of mites are particulate feeders, and parasitism of plants, vertebrates, and invertebrates has arisen repeatedly in the mites. External digestion is a major obstacle to life in fresh water or the sea. Arachnids are also notable for their ability to withstand starvation. Fasts of weeks or even months are routine for larger arachnids. Some scorpions and mygalomorph spiders live for more than 1year without food, and adult soft ticks (Argasidae) can survive for years without feeding.

Reproduction and Growth

Like all arthropods, arachnids grow by molting their exoskeletons and expanding the larger skin beneath with blood pressure before it hardens into the usual tough covering. The number of molts to maturity varies widely between 3 and 10–12; five is perhaps the most common. Life spans also vary greatly. The majority of mites and spiders live less than 1 year,

but several years is common among the larger forms, and mygalomorph spiders can live 20–30 years in captivity. Some arachnids cease molting at adulthood but others continue to molt periodically until death.

The ancestral reproductive pattern of sperm transfer is via spermatophore, modified in spiders, harvestmen, mites, and ricinuleids. Except among permanently social species and many mites, the sex ratio is equal, and parthenogenesis is rare. In the vast majority of species, males and females meet only to mate; cohabitation and parental care are uncommon. Nevertheless, various spiders, scorpions, schizomids, uropygids, and amblypygids may carry and feed their young, and in some harvestmen the male cares for the eggs in a specially built nest. Pseudoscorpion females nourish eggs with secretions from their bodies, and scorpions bear only live young. Females commonly guard their eggs until they hatch, but the young are usually abandoned soon thereafter. Mites are more diverse in reproductive strategies than nonacarine arachnids.

Ecology

Despite huge numbers of species (**Table 1**), arachnid biology is coherent in many ways (even though exceptions to nearly every generalization exist). Most arachnid orders consist of fluid-feeding predators, and predation still dominates these groups today. However, opilionids and two of the three orders of mites are particulate-feeders on detritus, fungi, and small invertebrates. Additionally, parasitism of vertebrates, invertebrates, and plants has arisen numerous times within the Acari and radiations within these lineages account for most of the 45,000 described species of mites.

Non-acarine arachnids tend to be at the top of the terrestrial invertebrate food chain wherever they occur. At one site in Israel, mites comprised 35% of the total soil arthropod population; in the Amazon ranges from 35 to 55% have been reported. Some harvestmen eat dead or decaying animal or plant material. Arachnids are generally nocturnal, despite numbers of diurnal harvestmen, spiders, and mites. Nocturnal forms hide in dark crevices and burrows during the day; several orders are morphologically specialized to inhabit small spaces. With the exception of the wind spiders (Solifugae), arachnids tend to be torpid and sedentary—none fly, for example, nor do any move constantly like ants or other active insects. The basic arachnid forages with a "sit and wait," solitary strategy. They move rarely and wait for prey to encounter them. Prey is then seized with a quick strike and immobilized. Highly organized social systems are known only in a few spiders, but loose aggregations (spiders, harvestmen, and pseudoscorpions) are not uncommon, usually in response to high prey density or limited refugia. Most arachnids are well adapted to last for long periods (weeks or months) without food; in the laboratory some have survived for more than 1 year. Only jumping spiders among arachnids have notably good vision; otherwise, they orient primarily via vibrations and touch.

Arachnids occupy all terrestrial habitats—deserts, forests, tundra, grasslands, mountaintops, soils, litter, caves, etc. Hydrachtinid mites (approximately 5000 species) are important components of most freshwater ecosystems; other mites are parasites of marine organisms, whereas others inhabit marine sediments, including the deepest oceanic trenches. Otherwise, arachnids are exclusively terrestrial. A few groups, such as scorpions and wind spiders, conserve body water as well as any arthropod and thus tend to dominate in deserts. The majority need moist conditions to survive. Schizomids, palpigrades, and ricinuleids are apparently restricted to the interstices of moist tropical and subtropical leaf litter or equally constant and moist habitats.

Phytogeny and Taxonomy

The most commonly encountered arachnids are spiders, scorpions, harvestmen, and mites, but the class contains seven smaller groups of terrestrial arthropods less familiar to the general public (**Figure 1**). The largest and heaviest arachnid is the African scorpion *Pandinus imperator*, which may reach a length of 18–20 cm. The smallest are perhaps the gall mites at 80 μm. No group of arachnids is well-known by vertebrate, butterfly, or vascular plant standards. Popular manuals are available for only a smattering of the most common species of spiders in Europe, North America, and Japan; all others require technical literature and specialist knowledge to identify. Myriad species are undescribed and undiscovered; at best, the approximately 93,000 known arachnid species are but one-third of the probable total, and probably much less. Most undescribed arachnid species are mites.

The study of arachnids is called arachnology. The principal international scientific society for nonacarines has approximately 600 members but many more belong only to regional societies. The taxonomy of arachnids is still a monumental task and an obstacle to better ecological and biotechnological understanding of arachnids, but the number of arachnid taxonomists is small and decreasing, and new students are not being trained. There are no comprehensive arachnology texts appropriate for university teaching (three exist for Acari), although modern "biologies" are available for spiders, scorpions, solifuges, pseudoscorpions, and various aspects of mite biology.

Paleontology

Arachnida were among the earliest terrestrial animals. The marked similarity between fossil and recent forms in overall body plan and morphology suggests few changes over hundreds of millions of years. The earliest sites for terrestrial arachnids are Early Devonian (400 Ma) and Late Silurian (414 Ma); the extinct arachnid order Trigonotarbida figures prominently, but mites are also present. The fossil record of all arachnids is comparatively poorly known. The 13 living and extinct orders are still known from less than 50 major time horizons since the Silurian; gaps remain more common than fossils. Arachnids seem to have invaded land in the Silurian and reached a pinnacle of ordinal diversity by the Carboniferous. The latest dates for the extinct orders Trigonotarbida and Phalangiotarbi are Permian and Carboniferous, respectively. No order seems to have succumbed to the end Cretaceous event that eliminated the dinosaurs. The earliest arachnid fossils are aquatic scorpions from the Late Silurian.

Spiders, pseudoscorpions, terrestrial scorpions, and mites are known from the Devonian. The scorpion-like Eurypterida, which are sister to true Arachnida, also may have become extinct at the Permo–Triassic boundary; the youngest fossils are also Carboniferous. The phylogeny in **Figure 2** implies that many arachnid clades must predate all scorpion fossils; an alternate opinion, based on less evidence, is that scorpions instead are sister to eurypterids, and that the phylogeny of **Figure 2** will be shown to be incorrect. Small size may explain the rarity of schizomid and palpigrade fossils, but mites in the Devonian show that small arachnid fossils can persist. Given the sparse record, it may well be that all arachnid orders were distinct and essentially modern in appearance by end Silurian–Early Devonian times. The most detailed arachnid fossils are "inclusions" in amber, which is fossilized gum or tree sap. The organisms became caught in the wet sap and were fossilized along with the sap. Cenozoic amber arachnid fossils 40–80 million years old can mostly be placed in modern genera; in some cases, the fossils are difficult or impossible to distinguish from modern species. Ecological relations are also preserved: Parasitic mites occur on their insect hosts, for example, thus documenting the antiquity and diversity of mite parasitism.

Araneae

Araneae is the second largest arachnid order with 108 families, 3200 genera, and approximately 37,000 species described (second to mites). Spiders are distinguished from other arachnids by their silk-producing spinnerets at the end of the abdomen and the prosomal poison glands exiting through their chelicerae modified as fangs (**Figure 3**). Spiders are among the very few animals that use silk throughout their entire lives. A narrow stalk joins the abdomen to the prosoma, allowing great flexibility and precise orientation of the abdominal spinnerets. Pedipalps are leg-like and short. Colors are predominantly dull tans, browns, and blacks, but spiders are occasionally very colorful, even iridescent. The abdomen of most spiders shows no trace of ancestral segmentation, unlike that of other arachnids. The Neotropical tarantula *Theraphosa leblondi* (Theraphosidae) is the largest spider at about 10 cm in body length. The smallest spiders are the tiny orb-weaving Symphytognathidae; adults are less than 1mm long.

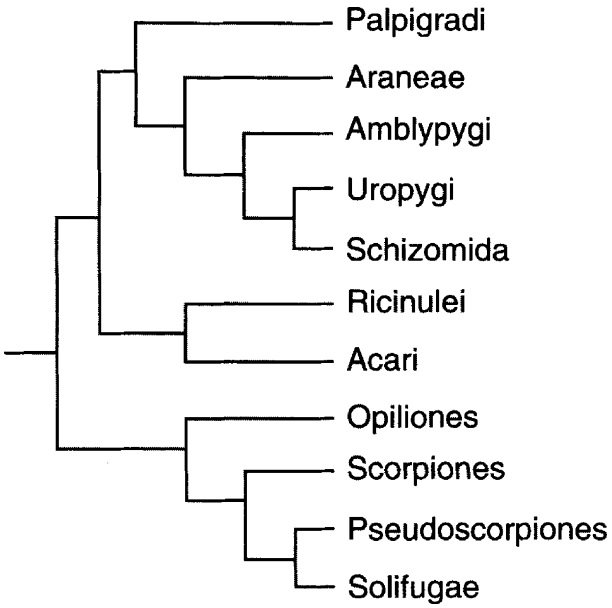


Figure 2 Phylogeny of arachnid orders.

Ecology

Spiders are very abundant. One calculation estimated 5 million animals per hectare in an English meadow. Another found 29,000 per cubic meter in an English sewage treatment plant. These are extreme values, but "average" nondesert habitat probably supports at least 1 and as many as 800 spiders per square meter. Many spider species disperse by ballooning. Ballooning spiders spin a silk line until it is caught by the wind and lifted aloft—potentially for hundreds or thousands of kilometers. One study estimated that 216,000 spiders per hectare may balloon into tilled fields during the growing season. Point diversities per hectare may vary from 100 species in moderate temperate zones to 600 or more in the wet tropics. Boreal diversities are less, perhaps 20–50 species per hectare.

Spiders are the only animals to use silk throughout their lives. Silk is one of the strongest and toughest natural fibers known and compares favorably with the best man-made filaments. Eight different kinds of spider silk have been discovered, but the maximum made by any single species is seven. Orb web spinners, for example, make two varieties of stiff, tough silk for weight-bearing structural fibers and safety lines, cement to fix silk to itself and substrate, sticky silk or glue to capture prey, rubbery silk to carry the sticky silk in the web, specialized silk for eggsacs, and thin, weak silk spun as multiple fibers to wrap prey, cradle eggs, and for other general purposes. All spiders are capable of spinning silk as soon as they leave the eggsac, and all make at least safety lines ("draglines") and the cement to attach them to substrate. Approximately half of spider species spin webs to capture prey. Web architectures are taxonomically specific and provide



Figure 3 Spider (Araneae: Ctenidae, *Cupiennius* sp.).

many clues for reconstructing spider evolution. The remaining spiders are ambush predators such as the crab spiders, which lurk inside flowers to attack pollinators, tube or retreat dwellers that forage in the very limited area at the burrow mouth, or vagabond predators such as wolf, ground, or jumping spiders.

Reproduction and Growth

Reproduction in spiders requires the male to ejaculate sperm onto a specially constructed sperm web. He then sucks up the sperm into specialized sperm transfer organs at the pedipalpal tips. Only adult males have such structures; their form is usually species-specific. With his palps filled, the male then searches for females. Females are usually more sedentary than males. Life is short for adult males, both because of predation and because they eat little as adults. The duration of courtship and copulation varies from seconds to days, but it ends with the insertion of the male pedipalp and transfer of sperm into the female gonopore. Females store the sperm and sometime later construct a silken eggsac into which 1–2500 eggs are placed. Eggs are fertilized only as they exit the female's body. Parental care is rare and highly variable, ranging from simple guarding of eggsac to actively feeding the babies. In at least one crab spider the female dies as the juveniles emerge from the eggsac, which then eat their mother's body. About 20 species of spiders are extremely social. Males are rare (1:40–100), generations overlap, food is shared, and prey capture and brood care are cooperative. Colony sizes range from a dozen to several thousand individuals.

Phylogeny and Taxonomy

Spiders are the seventh largest zoological order on Earth (after Coleoptera, Hymenoptera, Lepidoptera, Diptera, Hemiptera, and Acari), and of these they are the only one for which all taxonomic literature is fully cataloged. Catalogs greatly facilitate all kinds of research because scientists can easily determine the current taxonomic status and history of any described species and thus decide whether a given specimen belongs to a described species or not. At higher taxonomic levels (approximately, family) a basic, first-draft phylogeny is nearly complete. At the species level the easiest species-specific characters are found in the male and female genitalia. Species boundaries in spiders are generally clear-cut. However, the species taxonomy of spiders is based overwhelmingly on morphology. If more costly and sophisticated molecular methods were routinely applied, the number of distinguishable spider "taxa" would certainly increase.

The fundamental phylogenetic division in spiders is between the primitive mesotheles (spinnerets towards the middle of the abdomen) and the opisthotheles (spinnerets terminal). Within opisthotheles there are again two basic groups, the mygalomorphs (tarantulas and their allies) and araneomorphs (so-called "true" spiders). Mesotheles and mygalomorphs are not particularly diverse at the species level; araneomorphs currently include 94% of all known spider species and this disparity will certainly increase. Within araneomorphs the basal taxa are a few relictually distributed families in north and south temperate regions. Araneomorph haplogynes are diverse, but again comprise relatively few species. The araneomorph Entelegynae includes the bulk of modern spiders. Seven spider families currently contain more than 1000 species—all are entelegyne.

During a recent 39-year period, an average of 314 new species were described per year (12,200 total), but an annual average of 104 old names were synonymized, for a net gain of approximately 200 species per year (8800 total). Estimates of total spider diversity range from 76 to 170,000. Lower estimates mostly extrapolate from the proportions of new versus known species in taxonomic publications or are based on comparisons to well-known groups. The higher estimates take into account that many regions, particularly those richest in spider species, are disproportionately undercollected. In any case, the real diversity of spiders will never be known because a potentially great fraction will certainly go extinct before being discovered, much less described.

Major Lineages

Mesothelae

Liphistiidae is the only extant family and is limited to areas of Southeast Asia and Japan. Only a few dozen species are known, but some are common where they occur. Liphistiids retain many primitive morphological features, such as eight (rather than six or fewer) segmented spinnerets that insert anteriorly rather than terminally on the semi-segmented abdomen. Their biology may likewise represent the ecological "ground plan" for spiders. Liphistiids live in silk-lined tubes equipped with rudimentary trap doors in banks and cave entrances. Sometimes, silk "trip lines" lead away from the burrow entrance to extend the sensory radius of the animal. They are nocturnal, ambush predators. They live for 5–8 years, are remarkably sedentary, and consume a catholic diet of mainly walking prey.

Mygalomorphae

Mygalomorphs include the tarantulas or baboon spiders (Theraphosidae), trap-door spiders (Ctenizidae, Actinopodidae, Migidae, etc.), purse web spiders (Atypidae), funnel web spiders (Hexathelidae), and several other families with no common name. Mygalomorphs number 15 families and approximately 2200 species, but several of the families are para- or polyphyletic. A more realistic estimate is 20–30 "family-level" groups. Like mesotheles, mygalomorphs tend to live in burrows and forage at the burrow entrance or for a very limited distance around it. Some theraphosids are arboreal and spin elaborate silken retreats. Diplurids make extensive webs but are virtually unique among mygalomorphs. The venomous Australian funnel web spiders (*Atrax and Hadronyche:* Hexathelidae) were responsible for many deaths until an antivenin was developed in the 1980s. The large theraphosid baboon spiders are not seriously venomous to humans, despite their popular reputation.

Araneomorphae

Araneomorphs include approximately 94% of known spider species. Even the most primitive araneomorphs are very different from mygalomorphs and mesotheles. Basal araneomorphs

tend to be much smaller, and most are obligate web spinners with elaborate spinnerets capable of making adhesive "cribellate" silk. Cribellate silk is adhesive due to the extremely fine threads drawn from the cribellum, the very modified and fused anterior pair of spinnerets. Adhesive silk makes feasible a greater variety of web architectures, and these basal araneomorphs spin elaborate catching webs. The sister group of all remaining araneomorphs is the family Hypochilidae; its dozen or so species are limited to the Appalachian Mountains, a few places in western North America, and equally restricted sites in China.

Haplogynae Haplogynae comprises 17 families of spiders of diverse habits and worldwide distribution. Filistatidae are cribellate web-spinning spiders. Their web architecture is not much different from that of the mygalomorph diplurids—a sheet that narrows to a silk-lined retreat in a tube or crevice. Diguetids make dry silk webs. The majority of haplogynes are leaf-litter specialists and are vagabond, webless predators. The venomous brown recluse spiders (Loxosceles spp.) are haplogyne sicariids. The sister group of *Loxosceles* in southern Africa (Sicarius) has also been implicated in medically serious bites.

Entelegynae The remaining 70 families of spiders are the Entelegynae. Entelegynes share many evolutionary novelties. Females have a convoluted abdominal plate protecting their gonopore, which male genitalia must navigate successfully to achieve insemination. Sperm are stored in a unique "flow-through" system so that the female reproductive tract has two apertures to the outside. Female entelegynes also make special silk used exclusively in eggsacs, although its exact role is unknown. The lateral eyes possess a canoe-shaped tapetum that in some spiders enables orientation via polarized light. Although the higher phylogeny has been worked out for many entelegyne lineages, some very large ones remain unstudied, and the relationships between entelegyne lineages are also controversial.

Lycosoidea includes 11 families of mainly hunting spiders, some with common names: lynx, wolf, fishing, or tropical wolf spiders. Lycosidae (wolf spiders) and Pisauridae (fishing spiders) are common, cosmopolitan lycosoid spiders. Lycosoids occur in all terrestrial habitats, and some are semiaquatic in their ability to run across the surface of the water or dive beneath the surface. Web spinning is rare among lycosoids; some may have regained it after evolutionary loss. Most species are vagabond predators or, occasionally, tube dwellers. Active at night, they move sporadically or wait until prey approaches and then attack with powerful front legs and chelicerae. They are built strongly and run and jump with agility. The South American ctenid *Phoneutria* is venomous to humans.

Dionycha includes 17 families of spiders with two tarsal claws and a tuft of hair rather than the more common three claws and no claw tufts. The monophyly of Dionycha is by no means certain. Dionycha are also hunting spiders and have habits similar to those of lycosoids. Crab spiders (Thomisidae) wait for insect prey in flowers. Jumping spiders (Salticidae), the largest spider family, can be very brightly colored and often prefer to jump rather than walk. Their vision is superior to any other arachnid; salticids are the only sizable spider lineage that is strictly diurnal.

Orbiculariae includes 14 families of spiders and approximately 12,000 species. Most orbicularians spin prey-catching webs, but a few groups have secondarily lost the web-spinning habit. The primitive web architecture seems to be the orb—the classic spider web of radially symmetric, stiff, dry spokes supporting a spiral of sticky silk—but more orbicularian species have lost or modified the orb architecture than retained it. Web spiders rely exclusively on webs for prey capture. Araneidae (common orb weavers), Linyphiidae (sheet weavers), and Theridiidae (cobweb weavers) are the largest families. The venomous widow spiders (*Latrodectus*) are theridiids and are distributed worldwide; several species are spread by humans and are now cosmopolitan.

Scorpiones

About 1256 species of scorpions are currently known in 156 genera and 18 families (**Figure 4**). Scorpions are one of the better collected arachnid groups so that huge increases in diversity are not as likely as in mites or spiders. Estimates of total diversity run as high as 7000 species. All scorpions have large, obvious pedipalps modified as pincers, both body regions are broadly joined, the distal abdomen is narrowed into a flexible tail bearing a venomous stinger at the end, and ventrally. The abdomen bears a pair of comb-like sensory appendages



Figure 4 Scorpion (Scorpiones: Vaejovidae, *Uroctonus* sp.) (photograph by James C. Cokendolpher).

known as pectines. Colors vary from translucent to brown or black. Curiously, they fluoresce under ultraviolet light, a discovery that has galvanized recent field research on these animals. The longest is *Hadogenes troglodytes* at 21 cm, but *Pandinus imperator* is nearly as long and much heavier. The external morphology of recent scorpions is impressively similar to Silurian fossils. Formerly scorpions were thought to be the sister group of all other arachnids because they closely resemble the extinct marine eurypterids that are the sister group of all arachnids. Better analysis of morphological data (**Figure 2**), weakly corroborated by molecular evidence, suggests that scorpions are more deeply imbedded in the arachnid clade and merely retain many primitive features. The issue is controversial.

Scorpions are the only arachnids with a narrow postabdomen (''tail'') terminating in a venomous sting. The sting is most often used for defense, although scorpions will sting large or strong prey. The sting of most scorpions is painful—like wasp or hornet stings—but not dangerous. Characteristically, scorpions with slender pedipalps are more prone to sting their prey, whereas those with robust pedipalps tend to crush prey. The Central American genus *Centruroides*, Brazilian *Tityus*, and Old World *Androctonus*, *Leiurus*, *Mesobuthus*, and *Parabuthus* are very venomous and medically important. In Mexico, *Centruroides* spp. sting 300,000 and kill 1000 people annually; *Androctonus*, *Leiurus*, and *Mesobuthus* kill thousands annually in Egypt and Pakistan alone. Excepting ticks that spread disease, scorpions are by far the most dangerous arachnids to humans. Scorpions are correspondingly prominent in mythology and folklore (e.g., the zodiacal constellation ''Scorpio''). Scorpion venoms typically contain multiple low-molecular-weight proteinaceous neurotoxins. Scorpion blood inactivates scorpion venom, but if the venom is injected directly in a nerve, the animal rapidly dies. *Parabuthus transvaalicus* and *P. villosus* squirt venom to damage corneas, like spitting cobras.

Ecology

Scorpions are most diverse in deserts or similar dry areas, although they are reliably present in moist ecosystems if the temperature is not too cold. They now occur on all major landmasses except Antarctica. Favored habitats are burrows, under bark, stones, or logs, or inside small crevices. Burrows may be as deep as 40–80 cm, serving to escape the hot daytime temperatures in deserts. Because they like hard substrates and dry conditions, scorpions adapt well to human structures. Most scorpions are ground dwelling, but many species are arboreal, especially in the neotropics. In canopy fogging at four Amazonian sites in Peru, all of approximately 100 specimens were Buthidae (J. Ochoa, personal communication). A few are limited to lightless caves.

Scorpions are almost invariably nocturnal, although the African *P. villosus* is predominantly diurnal. The eyes seem to detect luminosity at best. Prey movements are detected by tarsal sense organs at distances up to 15 cm, and prey are attacked in a single motion. At distances up to 30 cm prey are located through orientation responses. The large, pincher-like pedipalps immobilize prey; thereafter, pieces are torn off by the chelicerae and digested in the pre-oral cavity before being sucked into the gut. Scorpions can be important consumers in some communities. In Israel, *Scorpio maurus* annually ate an average of 11% of the isopod population, which was not the only item in their diet. At moderate densities of 1.5 kg/ha, *Urodacus yaschenkoi* ate an annual average of 7.9 kg/ha of prey. Cannibalism and predation by other scorpion species are thought to be the most important sources of mortality, but other top invertebrate predators (e.g., spiders) and vertebrates are also important scorpion predators. Generally, mortality is highest immediately after birth, lower for intermediate-aged animals, and high for adults (e.g., 65, 30, and 60%, respectively, per year for the Australian *Urodacus manicatus*). Scorpion mortality is particularly high among males because of their high activity levels and mobility during the breeding season. Cannibalism by females is a significant cause of male death. Biased adult sex ratios of 1.2–1.4 : 1 are typical. Communal behavior, however, does occur. For example, family groups up to 15 individuals of the Brazilian *Opisthacanthus cayaporum* cooperate to construct and occupy communal chambers in the center of termite mounds. The African *Heterometrus* spp. also construct and share a communal burrow, inhabited by individuals of various ages. ''Piles'' of 20–30 individuals of *Centruroides exilicauda* are found in the winter months. Groups of 5 individuals of *Mesobuthus martensi*, all of the same age and all with their heads oriented toward a central spot, have been found under wet rocks in the intertidal zone.

The vast majority of scorpion species are subtropical or tropical. Point diversity (the number of species sympatric at one site) peaks in subtropical deserts and is particularly high (6 or 7, with a maximum of 12) in Baja California. Two to three species per site is more usual. *Vaejovis littoralis* reaches unusually high densities (8–12/m^2) in the drift line along the Gulf of California. The North American *Paruroctonus boreus* occurs as far north as British Columbia and Alberta, and the European *Euscorpius germanus* reaches the southern Alps. Even tropical scorpions sometimes inhabit extreme conditions; *Orobothriurus crassimanus* was collected at 5560 m in the Peruvian Andes.

Reproduction

Reproduction in scorpions is via a spermatophore attached to the substrate. The male completes production of the spermatophore inside his body, deposits the sperm inside, and attaches the spermatophore to the substrate, all the while holding on to the female during preliminary courtship. The spermatophore is ''spring-loaded'' and catapults the sperm mass into the female gonopore when a lever is touched. Scorpions are exclusively ovoviviparous or viviparous. The 1–105 young are born live and cling to the mother for the first few molts. A few species are parthenogenetic. Scorpions live 4 or 5 years (rarely 8); they do not molt as adults.

Phylogeny and Taxonomy

The higher classification of scorpions has changed dramatically as classical data have been reinterpreted phylogenetically. The old system proposed a few, huge, polyphyletic families about which nothing much in general could be said; now 16 or 18 families with increasingly coherent biologies are recognized. Species limits in scorpions are often difficult because scorpion genitalia are usually not species-specific. By tradition, scorpion taxonomists use the subspecies category more than most arachnologists. About 150 subspecies are recognized in addition to the 1260 species, but because these are easily distinguished they are probably distinct species. A classic example is the 25 non-overlapping, fully distinct subspecies of *Scorpio maurus*. Species-level taxonomic characters include the surface sculpturing of the exoskeleton, morphometric data, the number and position of pedipalpal trichobothria, and the hemispermatophores—internal male structures that produce the spermatophore.

Major Lineages

The basal division in Scorpiones is between the buthoids and remaining scorpions. New and Old World buthids are also distinct lineages. Scorpionoids and the vaejovoid–chactoid lineage are the remaining major scorpion lineages. Chaerilidae and Pseudochactidae (*Chaerilus*, 21 species; *Pseudochactus ovchinnikovi*, from Kazakhstan) are monogeneric and enigmatic; they are like none of the other scorpion families and their relation to other major lineages is obscure. They may be basal buthoid groups.

Buthidae is the largest and most widely distributed scorpion family with approximately 74 genera and 531 species. Buthidae is most diverse in the African tropics and Palearctic regions. Buthids tend to have slender, elongate pincers, a robust tail, and usually a tubercle under the sting. All scorpions considered dangerous to humans are buthids. Buthids are also the most diverse ecologically and occupy humid, mesic, and dry habitats. The small family Microcharmidae is an Afrotropical buthid segregate with two genera and six species.

The Scorpionoidea (36 genera and 355 species) is a large, monophyletic lineage that includes Bothriuridae, Diplocentridae, Heteroscorpionidae, Hemiscorpiidae, Ischnuridae, Scorpionidae, and Urodacidae. Scorpionidae lack a tubercle under the sting and the sides of the sternum are parallel. The family contains the genus *Scorpio* from the Mediterranean and Near East, much mentioned in classical Greek, Egyptian, and Christian myths, and *Pandinus*, the giant African black scorpion. The longest and heaviest scorpions are scorpionids, which are exclusively Old World.

The relationships and monophyly of the vaejovoid–chactoid lineages (43 genera and 495 species: Chactidae, Euscorpiidae, Iuridae, Scorpiopidae, Superstitionidae, Troglotayoscidae, and Vaejovidae) are the most problematic areas of scorpion higher taxonomy and phylogeny. The large families Iuridae and Chactidae, in particular, are doubtfully monophyletic, although each includes many clearly valid groups. Together, the vaejovoid–chactoid lineages comprise about one-fourth of known scorpion species, including the most common species in North America.

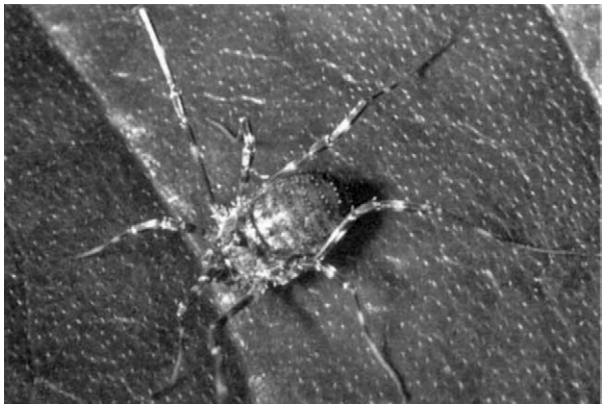


Figure 5 Harvestman (Opiliones: Phalangiidae, *Hadrobunus* sp.).

Opiliones

The described world fauna of Opiliones (harvestmen or daddy long legs) comprises approximately 44 families, 1554 genera, and about 5000 species (**Figure 5**). The largest harvestman is *Trogulus torosus* at 2.2 cm long. The anterior and posterior body regions are broadly joined and the abdomen is rather short, giving the body a wider and rounder appearance than those of other arachnids. The second pair of walking legs is usually the longest. Opilions have just two eyes (cave or litter groups are sometimes blind). All harvestmen have a pair of glands that open via large pores on the anterolateral margins of the body; the function of their secretions is apparently diverse. Harvestmen are the only arachnids in which males have a penis. Females have a long, flexible, and extensible ovipositor (as do many mites). The most common group in north temperate regions, the Phalangioidea (daddy long legs), have soft and flexible bodies, weak mouthparts, and extremely long legs (commonly 15 times the body length), but just as many harvestmen are fantastically armored with bizarre, huge chelicerae, raptorial pedipalps, and short, stiff legs. Others are mite-like, and still others are dorsoventrally flattened. Colors are usually subtle patterns of brown, gray, or black, but tropical forms can rival anything seen in spiders or mites. In the long-legged forms, the distal leg tip is divided into numerous false segments that form a prehensile tip. It can be wrapped around objects to achieve a very firm and adaptable grip. The long second legs may double as feelers.

The respiratory system is exclusively tracheate. The very long-legged Phalangioidea have accessory spiracles on distal leg articles. Touch and vibration perception, as in most other arachnids, seems to be the dominant sense. The eyes at best distinguish light and dark and direction of light. Opilions consume a broader diet than any other arachnids other than mites. The basic pattern is predation, but some, for example, specialize on snails, which are otherwise rarely consumed by nonacarine arachnids. Opilions also are known to eat dead insects, fruit, and decaying vegetable matter. Unlike other arachnids, harvestmen can ingest solid particles, as shown by sclerotized bits in their excrement.

Many harvestmen defend themselves against attack by shedding legs. One study of two species in Louisiana found

that about half the animals had lost one leg, but less than 10% lacked three. Seven-legged harvestmen seem to survive and function as well as intact animals. A shed leg continues to jerk and twitch attractively for minutes, permitting the harvestman to escape. All harvestmen have paired "repugnatorial" glands on the front margin of the body. When legs are pinched a droplet of fluid appears at the orifice which may be dabbed on an attacker with a leg or allowed to evaporate. The secretions can also help aggregating species to find each other, and some show broad antibiotic and antifungal activity, presumably useful to litter and soil dwelling forms. Quinones are a major ingredient. Soil dwelling harvestmen and the short-legged Laniatores are slow compared to long-legged harvestmen.

Ecology

In general, harvestmen prefer moist, or at least not xeric, environments. The mite-like Cyphophthalmi live in dark leaf litter, caves, or under stones. The largely tropical and usually short-legged Laniatores move slowly over vegetation or the forest floor. The usually long-legged Eupnoi can be anywhere, but their very long legs with prehensile tips are specialized for crossing the large gaps between the leaves of trees, shrubs, and herbs. The northern European opilionid fauna comprises approximately 24 species and that of North America approximately 235 species. In temperate regions, diversities of more than a dozen species per hectare are uncommon.

Reproduction

Uniquely among non-acarine arachnids, sperm transfer occurs via the male penis. Mating occurs quickly and seemingly without courtship. The male faces the female and pushes the penis underneath her body or between her chelicerae and into the gonopore. After insemination, the sexes separate and continue their solitary wanderings. Females use their long, flexible ovipositors to deposit eggs in suitably protected areas. Trogulids deposit their eggs only in empty snail shells, and other groups oviposit beneath stones, deep into soil, underneath bark, or in bore holes in plant stems left by insects, usually abandoning the eggs once laid. Newly hatched animals are active and resemble adults. Five to eight molts to maturity are common. An unusual reversal of sex role occurs in the Panamanian harvestman *Zygopachylus albomarginis*. Males fight to occupy existing mud nests or construct their own. Females wander between nests, courting the males, mating, and ovipositing in a series of nests. They have nothing more to do with the offspring. Males accumulate eggs of different ages and from different females and defend the eggs against conspecifics and ants.

Phylogeny and Taxonomy

The phylogeny of Opiliones has recently been clarified at the superfamily level, but many additional changes are expected in familial arrangements. Some families seem to be based only on primitive features (e.g., Travuniidae, Phalangodidae, and Triaenonychidae). Eupnoi and Dyspnoi classically formed the suborder Palpatores, but increasing evidence indicates that this taxon is paraphyletic. The number of recognized families has approximately doubled in the past 20 years. At the species level the morphology of male genitalia is especially diagnostic.

Major Lineages

Cyphophthalmi

This suborder contains five (or six) families and about 100 species. It is sister to all remaining Opiliones (the "Phalangida"). Cyphophthalmi are eyeless, live in deep moist leaf litter or caves, and range in size from 1 to 7 mm. The animals have a hardened plate covering the entire dorsal surface, and they resemble mites. *Siro* is North American and European, but the family also occurs in Southeast Asia, Turkey, Japan, Mexico, and South Africa. Life spans of up to 7 years have been reported.

Eupnoi

This group contains two superfamilies, including the classic daddy long legs (Phalangioidea: Phalangiidae, Sclerosomatidae, Megalopsalididae, and Neopilionidae) of soft-bodied, long-legged harvestmen. *Phalangium opilio* is common around buildings and introduced throughout the world. The sclerosomatid *Leiobunum* spp. are common in North American and European forests, in which they move easily across the upper vegetation. They are predators and scavengers. Caddoidea includes only one or two families of harvestmen with enlarged eyes and short legs, sometimes common on tree trunks.

Dyspnoi

This group also includes two superfamilies, but has many fewer species than Eupnoi. Ischyropsalidoidea contains three families and just seven genera. *Ischyropsalis* is European and feeds on snails. Troguloidea contains four families, two monogeneric and several genera in the remaining families. *Nemostoma* is common in caves. Trogulidae are peculiar harvestmen that look like giant, flattened mites. Legs are very short. Trogulids live under stones and in leaf litter; snails are an important part of their diet.

Laniatores

This group is morphologically more diverse than the previous groups and tends to be more diverse in the tropics, including many colorful species. Three super-families are recognized: Traviunoidea (4 families), Oncopodoidea (1 family), and Gonyleptoidea (18–20 families). Colloquially, laniatorids are known as "short-legged" harvestmen because the most common laniatorids do have short legs, but some agoristenids, cosmetids, gonyleptids, caelopygines, progonyleptoidellines, and mitobatines have legs comparable in length to those of Palpatores. The Gonyleptoidea contains 18–20 families with raptorial pedipalps for prey capture and enlarged fourth coxae with spreading appendages—perhaps a defense against being swallowed whole by predators or dragged down the burrow of a parasitoid. Gonyleptidae are typical heavily armored and spiny, often colorful, and usually larger than 4 mm. They are exclusively Neotropical and, with more than 100 genera, are one of the largest families. The closely related Cosmetidae superficially resemble gonyleptids—large, heavily armored, often colorful, and slow-moving New World harvestmen.

Phalangodidae is a large cosmopolitan family of more than 150 genera defined mainly by the features they lack. For years it has been used to file taxa with no obvious relatives, and therefore its biology makes little sense. Although most diverse in the northern neotropics, their distribution includes southern North America. Their pedipalps are flattened with spiny margins.

Smaller Arachnid Orders

Amblypygi (whip spiders or tailless whip scorpions), Uropygi (whip scorpions or vinegaroons), Schizomida (no common name), and Palpigradi (microwhip scorpions) are all small, unfamiliar arachnid groups that are closely related to each other and to spiders (**Figure 2**). Like primitive spiders, all have two pairs of abdominal book lungs, although the second pair is missing in the tiny Schizomida and palpigrades lack both. The first pair of walking legs is elongate and feeler-like, with false articulations in the terminal articles to promote flexibility. Pseudoscorpiones (pseudoscorpions), Solifugae (wind spiders), and Ricinulei (no common name) are a miscellany of remaining arachnid orders related to various orders already discussed, as illustrated in **Figure 2**.

Amblypygi

Approximately 125 species of amblypygids in 20 genera and 5 families are known (**Figure 6**). The American *Acanthophrynus coronatus* is the biggest, at about 4.5 cm in body length. The usual adult body size is about 4–6 cm. The first walking legs of amblypygids are enormously long, and fully stretched a large animal can span 50 cm. Whip spiders have no tail, and their pedipalps are modified into fierce-looking, spiny raptorial appendages. Amblypygids are easily recognized by their extremely long front legs and raptorial pedipalps. Colors are dull brown or black. The body is flat and leg insertion twisted so that limbs fold in the same plane as the body (like a crab), permitting the animals to edge through thin openings such as cracks in hollow trees. They move sideways more easily than forward or back. Amblypygids hunt by drifting their long front legs gently over the surface around them to locate prey and using their raptorial pedipalps to pounce. Like spiders, all the abdominal ganglia have migrated into the prosoma in which, fused, they form a brain. Reproduction is via a spermatophore. Like uropygids, amblypygid females carry their egg clutches inside a membrane of dried mucus glued to their ventral abdomens. The 15–50 young hatch and remain inside this membrane until they have undergone their first molt. The young cling to the mother for a short time.

Whip spiders live in subtropical and tropical areas, in forests and often in caves. They are exclusively nocturnal and fairly common. During the day they hide in hollow trees or logs, under loose bark, or under large logs. Only one species lives in leaf litter and is not known to burrow. Their diet seems to be a broad range of smaller arthropods in their environment.

Uropygi

About 100 species of whip scorpions are known in 16 genera and one family, Thelyphonidae (**Figure 7**). At 7.5 cm, the largest is *Mastigoproctus giganteus* of North America, but 3–5 cm is usual. Whip scorpions are easy to recognize by the long posterior whip or flagellum (highly modified terminal abdominal segments). Colors are brown to black. Uropygids have defensive anal glands that accurately spray an acid-smelling fluid at attackers. The smell explains the common name "vinegaroon." The fluid of *M. giganteus* is 85% acetic acid but it also contains substances to reduce the surface tension of the epicuticle so that the acetic acid can spread widely and penetrate. Vinegaroons are supposedly not sensitive to their own spray. Reproductive habits are known only for a few species, but presumably sperm transfer occurs via a spermatophore glued to the substrate in all species. Some species have a lengthy courtship—10 hr to several days. Females keep their 12–40 eggs attached to their ventral abdomen. Female *Thelyphonus* build a deep burrow and do not feed



Figure 6 Whip spider or tailless whip scorpion (Amblypygi: Phrynidae, *Phrynus sp.*).



Figure 7 Whip scorpion (Uropygi: Thelyphonidae, *Mastigoproctus* sp.).

while guarding the eggs for 4 or 5 weeks. Uropygids may live 6–8 years or even longer.

Uropygids are tropical to subtropical animals. They hide in leaf litter, burrows, under logs and rocks, and inside dark crevices or holes during the day, emerging at night to hunt. Not much is known about their prey, but presumably it consists of other small ground-dwelling insects, arachnids, and crustaceans, which they crush between their pedipalpal segments prior to ingestion.

Schizomida

About 200 species in 30 genera and two families (Protoschizomidae and Hubbardiidae) are known (**Figure 8**). Schizomids (there is no common name) are most like tiny uropygids, but the abdominal flagellum is short (three segments but rarely five). Colors are usually light yellow to tan to dull green. Assiduous searching in moist tropical litter usually turns up a schizomid. The largest schizomid known is *Agastoschizomus lucifer* at about 12.7 mm long, but 3 mm is typical. Most schizomids lack eyes entirely and live in moist tropical or subtropical leaf litter, under stones, in logs, moist crevices, caves, and so on. Reproduction is via a spermatophore attached to the substrate, after which females lay 6–30 eggs that are glued to the ventral female abdomen until the young emerge. At two Amazon sites schizomid abundances ranged from 5 to 110 animals per square meter per month.

Like the uropygids they have been reported to produce a defensive chemical smell. They can move backwards rapidly and have enlarged femora apparently used to hop backwards.

Palpigradi

There is only one family (Eukoeneniidae) with five or six genera and about 80 species. Palpigrades (microwhip scorpions) are tiny, light yellow to white, delicate soil and leaf litter specialists with a cosmopolitan distribution. The largest is *Eukoenenia draco* at 2.8 mm long, but 1 or 2 mm is typical. They resemble the young of Uropygi. Like whip scorpions, they have a wide preabdomen and a multi-segmented whip-like postabdomen. Like many soil organisms, palpigrades lack most of the organ systems required by large animals that live in drier and less constant environments. They lack eyes, respiratory organs, and a circulatory system but have innervated setae that detect vibrations. Most species are known from the tropics; several palpigrades apparently live in intertidal or shallow marine habitats. *Eukoenenia janetscheki* was the only species found at two Amazon sites, but it was fairly abundant (30–120 animals per square meter per month).

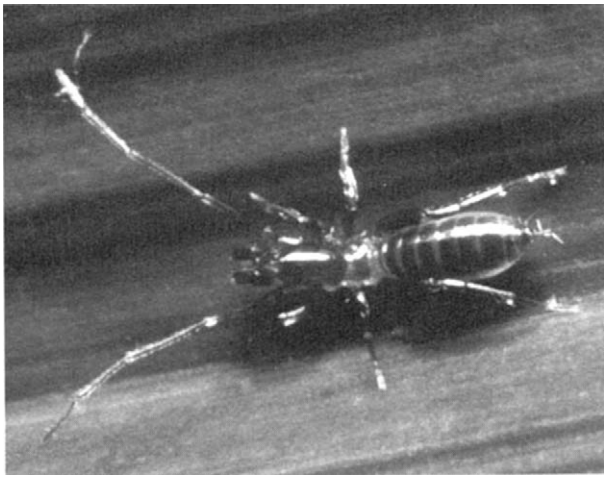


Figure 8 Schizomid (Schizomida: Hubbardiidae) (photograph by Jeremy A. Miller).

Ricinulei

One family (Ricinoididae), three genera, and 53 species of ricinuleids are known (**Figure 9**). Ricinuleids, like schizomids and palpigrades, are soil-dwelling specialists. The largest is *Ricinoides afzelli* at 10mm. long, but most are 3–5 mm. Unlike the preceding forms, ricinuleids are heavily armored with a thick exoskeleton and a kind of visor or flap (the cucullus) that folds over the mouthparts. Colors are reddish to brown or black. Males have third legs modified for sperm transfer. Some have "eye-spots" that are able to sense light. Females lay one or two eggs at a time, which they may carry about with them. Adults may live 10 years. Formerly thought to be extremely rare, ricinuleids are not uncommon in Neotropical leaf litter. They are also known from Africa, but not from the Asian tropics.

Pseudoscorpiones

About 3100 species of Pseudoscorpions (or "false scorpions") have been described, classified in approximately 430 genera, 24 families, and two suborders (Epiocheirata and Iocheirata) (**Figure 10**). The North American fauna comprises about 350 species. One local fauna in the Amazon comprised nine species and Mediterranean ecosystems around Perth, Australia,



Figure 9 Ricinuleid (Ricinoididae).

and other Old World tropical moist ecosystems are comparable (M. Harvey, personal communication). Pseudoscorpions look like tiny scorpions without the tail and pectines. Colors are tan to dark brown or black. Their chelate pedipalps can have venom glands and the chelicerae have silk glands. The largest is *Garypus titaneus* at 12 mm long, but most are 2 or 3 mm. They are flattened for life in crevices, usually under bark or in leaf litter, but a few forms live on ocean shores where they are truly intertidal. Numerous species live among the hairs of mammals where they feed upon other arthropod parasites. At least one species has the body, legs, and palps modified for living on a Celebese rat. *Chelifer cancroides* is cosmopolitan and lives in houses where it hunts book lice and other small insects. Pseudoscorpions are sometimes found in aggregations of dozens of individuals. Reproduction, as in most other arachnids, is via a spermatophore. The female builds a silken nest and secretes a "brood sac" attached to her body and in which she nourishes the 3–40 embryos with maternal secretions. Juveniles leave the mother soon after hatching. Pseudoscorpions are famous for their tendency to disperse by phoresy. Mature females (rarely males or juveniles) use their pedipalps to hold onto larger insects, which carry them from place to place as they fly about. Phoretic specimens in fossil amber show that this behavior is at least 10 million years old.

Solifugae

About 1065 species of solifuges are known, grouped in about 153 genera and 12 families (**Figure 11**). The greatest known diversity occurs in Namibia. Wind spiders (sun spiders or camel spiders) look somewhat like hairy, stout, fast-running spiders, but their chelicerae are fearsomely enlarged and both the leg-like pedipalps and the long and slender first legs are used as feelers. The fourth femurs bear ventral sensory structures shaped like inverted mushrooms. Colors are tan to sandy yellow. The largest is *Galeodes caspius* at 7 cm, but 2 or 3 cm is usual. Wind spiders can run extremely fast (53 cm/sec) for short bursts, but like most arachnids they cannot sustain rapid locomotion for long periods. The solifuge respiratory system is only tracheate and is comparable in complexity to that of insects. They are generally nocturnal and exclusively carnivores and, like any of the larger arachnids, will take small mice, lizards, amphibians, and so on if the opportunity arises. A few species feed only on termites. Unlike most other arachnids, solifuges are active, cursorial predators. Wind spiders are most common in tropical and subtropical dry or desert regions, but they are absent from Australia. Many inhabit burrows, where they may stay up to 9 months of the year, depending on rainfall. Reproduction is via spermatophore, which the male produces during courtship and transfers into the female gonopore with his chelicerae. The female digs a deep burrow to deposit egg masses that may contain 5–164 eggs. Females may lay 1–5 egg masses. Some females remain with the eggs until hatching. Newly hatched wind spiders are gregarious and remain in the burrow for the first two molts.

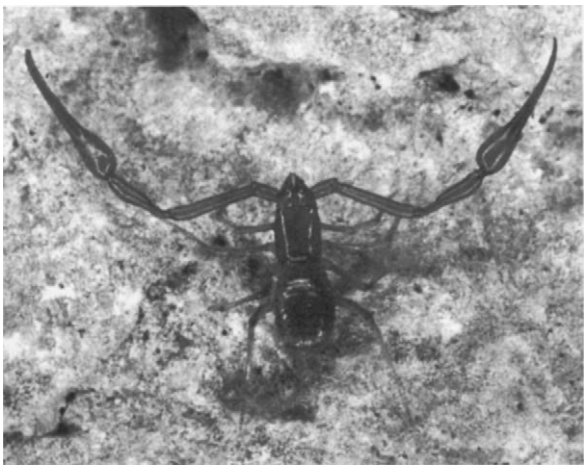


Figure 10 Pseudoscorpion (Pseudoscorpiones: Neobisiidae, *Tartarocreagis infernalis* (Muchmore) (photograph by James C. Coken-dolpher).

Acari

Mites and ticks are classified as Acari, from a Greek word meaning a thing too tiny to be divided (**Figure 12**). At more than 45,000 species, Acari is already the most diverse arachnid order, but acarologists estimate that between half a million and more than 1 million species actually exist. Despite being far more abundant and diverse than any other order of arachnids, mites and ticks are nonetheless much less familiar to most people than spiders, harvestmen, or scorpions. Likewise, scientific knowledge of mites lags far behind our understanding of most other arachnid groups. The primary reason for both public and scientific ignorance of mites is their small size.

Since we sometimes find them on ourselves and our pets, because they are relatively large, and because of their role as serious vectors of human disease, the ticks are probably the most familiar group of mites. In fact, ticks (which have their own name only because of their importance to humans) represent a small offshoot on the evolutionary tree of mites. Other kinds of mites are vastly more common and diverse



Figure 11 Wind spider or solifuge (Solifuga: Eremobatidae).

than the ticks and include not only other parasites but also followers of many other ways of life. In this article, unless otherwise distinguished, the term *mites* is intended to include ticks and will be used interchangeably with Acari.

The Acari are distinguished from other arachnids principally by the lack of conspicuous body segmentation, the broad union of the leg-bearing part of the body (the podosoma) with the posterior part (the opisthosoma) to form a single unit (idiosoma), the aggregation of the mouthparts into a distinct anterior body region (the gnathosoma), and the first active immature stage, the larva, having only three pairs of legs. The gnathosoma is not a true head, like that of an insect, in that it possesses neither eyes nor antennae and does not contain the brain. (Most mites have no eyes at all, but those that do bear them on the idiosoma. The brain is situated anterior to the stomach in the idiosoma.) Although usually compact in form, like that of a spider mite or tick, the overall shape of the mite body varies greatly. Species that live in polypore fungi or hair follicles are elongated to fit their habitat, and mites found deep in soil tend to be worm-like. Ticks and some other mites can tolerate enormous distention of the body wall to accommodate food.

Like other chelicerate arthropods, mites have no antennae, but in many groups the first pair of legs carry chemosensory and tactile organs and are not used, or little used, for locomotion. The pedipalps, integrated laterally into the gnathosoma, are also primarily sensory in function. The chelicerae, the primary food-getting organs, vary greatly in form, in accord with the varied feeding habits of mites. In many species, the chelicerae of males are so modified for sperm transfer that adult males scarcely feed.

Ecology

Mites are found in virtually every habitat on Earth. On land, legions of free-living mites populate soils in all habitats, often numbering in the hundreds of thousands per square meter in the soil (50,000–250,000 per square meter, down to 5 cm depth, in forests; 20,000–100,000 in grasslands; and 500–1000 in deserts). In many soils, including deserts, mites can be found at depths up to 10 m, where they follow plant root systems. Although the habits and behavior of many soil mites remain a mystery, it is known that soil mites fill many ecological roles. Many eat decaying organic matter (and the microorganisms it contains); others consume fungi, algae, and bacteria; and still others are predators, feeding on nematodes, other mites, or various stages of insects. Oribatid mites in the soil facilitate plant growth by dispersing the spores of mycorrhizal fungi, which form mutually beneficial and often essential associations with higher plants.

Mites are common and often abundant on plants, on which many species feed on leaf, stem, or root tissue by piercing cell walls. Others specialize on plant reproductive parts, including nectar, pollen, and fruit. Still other plant-dwelling mites are fungivores (eating fungi on leaves or stems) or predators, mostly of plant-feeding mites. Unlike life in the soil, life on plant surfaces constantly threatens mites with desiccation. Hiding from direct sun on the underside of leaves offers some protection because the "boundary" layer of still air is kept relatively moist by plant transpiration.

In addition, many plants produce domatia—hollow, mite-sized cavities or pockets in the leaf. The adaptive function of domatia is controversial, but experiments have shown that removal of domatia from some plants that normally have them reduces the density of beneficial predatory mites and increases the density of herbivorous mites.

Some fungus-feeding mites have specialized cavities or pouches in the body wall (sporothecae) in which they carry spores or propagules of fungi that they depend on for nourishment.

Parasitic and phoretic mites exploit most terrestrial vertebrate species, most other arthropod groups, mollusks, and even marine invertebrates (**Figure 13**). Most species of birds, for example, generally support many "feather" mite species found nowhere else. Four or five different, host-specific species inhabit different feather types (flight feathers, tail feathers, or body feathers), whereas different species live on the surfaces of feathers, others in the interior lumen of the quills, others on or in the skin, and yet others in the nasal passages. One particular species of parrot hosts more than 20 species of feather mites. Bird nests support additional, phoretic species that feed on debris or on other nest associates. Both parasitic and phoretic species generally colonize young birds while in the natal nest and travel with the host until it builds its own nest. Similar sets of species inhabit the bodies of mammals and reptiles and, to a lesser degree, amphibians.

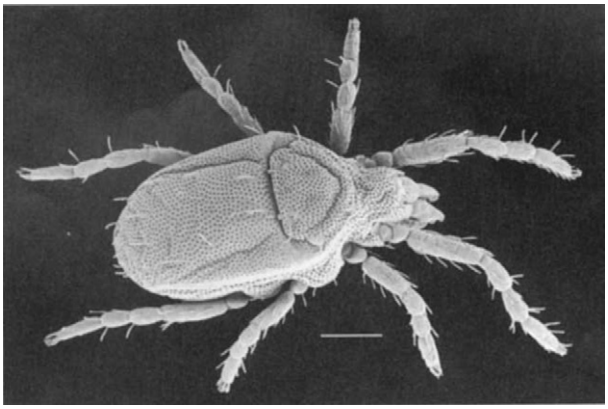


Figure 12 Free-living yellow moss mite (Acari: Penthalodidae, *Stereotydeus* sp.) (photomicrograph by David Evans Walter; scale bar = 100 μm).

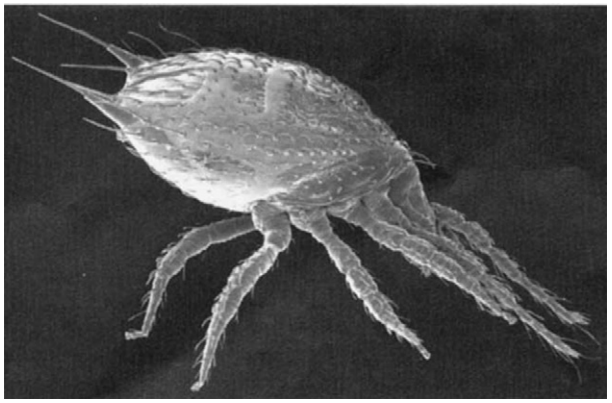


Figure 13 Predatory sejine mite (Acari: Sejiidae, *Epicroseius* sp.) (photomicrograph by David Evans Walter).

Mites are also common associates of insects and other invertebrates. Some are parasitic and deleterious to their hosts, such as the tracheal mite and *Varroa* mite parasites of honeybees or moth ear mites. Some are commensal (e.g., feeding on insect waste products). Others are believed to benefit the host, perhaps by consuming other harmful mites, parasitic nematodes, or decaying matter that would otherwise nourish fungi or bacteria that might threaten the host or its young. Some insects have acarinaria, special pockets in the outer body wall, that are routinely occupied by phoretic mites. The very existence of acarinaria, which have no other known function other than to house and transport mites, is evidence for a net beneficial relationship between these mites and their insect hosts. Parasitic mites have been shown to be capable of mediating horizontal gene transfer between insect host species.

All ticks require blood from a vertebrate host for development and reproduction but live off the host when not feeding or mating. Chiggers and many other mite species must live a portion of their lives, and others all their lives, as parasites of animals. Other mites exploit animal hosts not for food but strictly for transport—a phenomenon called phoresy (see the following section). Parasitism is thought to have evolved, in some lineages, from phoretic associations.

Mites are common and diverse inhabitants of freshwater and marine habitats. In fresh water, adults and nymphs prey on other mites, crustaceans, and aquatic insects. The larvae of aquatic mites are commonly parasitic on insects and crustaceans, and nymphs and adults of unionicolid mites parasitize bivalve mollusks. Other aquatic mites graze algae and fungi growing on aquatic plants. Mites are common on seashores, including the intertidal zone. The sea is home to predatory and algae-feeding mites. Some mites have been found at depths up to 7000 m at hydrothermal vents, and one group lives in the gut of sea cucumbers. Mites are the only arachnids found in the ocean depths.

Closer to home, even the cleanest office, classroom, or home normally supports dust mites (e.g., *Dermatophagoides* spp.), which live on organic matter in house dust. Some mites are stored-product pests, feeding on grains, cheese, and other dry foods. Even more intimately, as you read this encyclopedia there is a very good chance that the follicles and sebaceous glands of your forehead or eyelids house normally benign, microscopic mites (*Demodex* spp).

Reproduction and Dispersal

Reproductive patterns in mites are remarkably diverse. Many mites are ordinary diploids (having an equal number of chromosomes from each parent) with two sexes, requiring fertilization by the male for females to reproduce. In other groups, males are absent, and females bear female young from unfertilized eggs. In still other mites (e.g., spider mites), females require sperm from a male to produce (diploid) daughters, but sons are haploid (they have only the maternal chromosomes) since they are produced from unfertilized eggs. In other species (e.g., some phytoseiid mites), both sexes are diploid and all eggs require fertilization to develop, but only the mother's chromosomes are active in males (called paternal genome loss).

Modes of sperm transfer are diverse. Some mites transfer sperm directly through copulation, using legs, palps, or other specialized body parts to accomplish the transfer. Other mites rely on spermatophores placed on the substrate, with or without active enticement of the female to take up the sperm.

Development of the egg takes place in most mites after oviposition. In some species (e.g., moth ear mites and hummingbird flower mites) the eggs are laid *communally*. In other mites, the young leave the mother's body as active larvae. In some mites (e.g., *Siteroptes graminum*) sisters and brothers both hatch and mate within the mother's body, and both mother and sons die when the fertilized daughters emerge—truly material for a Greek tragedy.

The development of most mites includes one or more active stages between egg and adult. (Some Heterostigmata have no active immature stages.) In nearly all groups, the first active immature stage, called the larva, has three pairs of legs, and later stages, including the adult, have four pairs. (The plant-parasitic Eriophyoidea are exceptional in having two pairs of legs in all active stages, and females of many species of insect parasitic Podapolipidae have three, two, or even only one—the anterior—pair of legs.) The number of nymphal stages between larva and adult varies from none to three, depending on the group. Several groups of mites have metabolically inactive stages that can endure stressful periods. Life expectancy among mites also varies widely. Some mites live only about 1 week or so (e.g., hummingbird flower mites), whereas others require several years to complete the life cycle (e.g., certain ticks, oribatid mites, and water mites).

Because mites are so small and (unlike most insects) are wingless, selection for effective dispersal from one food patch to another has produced many special adaptations. Parasitic mites, for example, must have a way to find a new host—or ensure that their offspring do—before the current host dies or becomes unsuitable. In the case of plant-feeding mites, some species disperse aerially, either by simple passive transport in the wind or, in the case of some spider mites (Family Tetranychidae), by "ballooning" on strands of silk. (The silk-producing glands in these mites are not homologous to those of spiders.) Other plant-feeding mites disperse phoretically on plant-visiting animals. For example, nectar- and/or pollen-feeding insects and birds on all continents carry nectar- and/or pollen-feeding mites as hitchhikers—a way of life that has many independent evolutionary origins.

Some groups of mites have special, non-feeding stages adapted for phoretic dispersal. In others, adults and sometimes later nymphal stages mount phoretic hosts to reach new hosts or habitats when local mite population density becomes high, when the host is ready to disperse (e.g., emerging adult insects), or when a food plant becomes unsuitable.

Phylogeny, Taxonomy, and Current Knowledge

Only 90 species of Acari were known to Linnaeus, all of which he placed in the genus *Acarus*. Currently, perhaps 45,000 species in 431 families and 3672 genera have been described compared to 30,000 species estimated 30 years ago. The rate of description of new species is still very high—several hundred species per year—and every systematic acarologist has dozens or hundreds of undescribed species sitting on shelves in vials

of alcohol, awaiting description and classification. In fact, the taxonomic impediment to full knowledge of the Acari is monumental:— The number of practicing, modern mite systematists is no more than 50, and the rate of training of new systematists is alarmingly low. Moreover, exploration and species discovery in tropical forests (especially the forest canopy), lakes, and streams is still in its earliest stages. Most estimates of the total world fauna of Acari range from 500,000 to more than 1 million species, but some acarologists even suggest that mites rival insects in worldwide number of species. In any case, mites represent by far the most species-rich and ecologically diverse arachnid group. **Table 1** gives summary data on mite diversity.

Major Lineages

Given the highly incomplete knowledge of the Acari, it is not surprising that their phylogeny, classification, and nomenclature are far from settled. It is generally recognized, however, that the Acari comprise three lineages that diverged from one another in very ancient times: the Opilioacariformes and Parastiformes (which together constitute the Anactinotrichida) and the Acariformes (which constitute the Actinotrichida). Their relationships to one another and to non-acarine arachnid groups have been debated for decades, but the current opinion appears to be that the Acari, so defined, is a monophyletic group.

The taxonomic rank of acarine lineages is also controversial. The tradition among zoologists in general and arachnologists is to treat Acari as an order, but acarologists now treat it as a subclass or even a class, with subsidiary mite taxa then being orders or superorders. Clearly, this is not the place to attempt a resolution of this controversy. The taxonomy in the following sections reflects the prevailing acarological point of view that major groups should rank as orders, despite the incongruity with the rest of this article.

Order Opilioacariformes

The small order Opilioacariformes (with 17 species described of an estimated 85–170 total) with a single family (Opilioacaridae) are large mites (1.5–2.3 mm in length) that somewhat resemble small harvestmen. Some genera occur in dry climates, often under litter or stones, and others in tropical forest litter, where they feed as omnivores or predators.

Order Parasitiformes

The Order Parasitiformes includes three Suborders: Holothyrida, Mesostigmata, and Ixodida.

Suborder Holothyrida The Holothyrida is a small group (with 32 species described, of an estimated 160–320 total) of large (2–7 mm) saprophagous and predaceous mites of temperate and tropical forest litter, known only from Pacific-Indian Ocean islands and the Australasian and Neotropical regions.

Suborder Mesostigmata The cosmopolitan suborder Mesostigmata, in contrast, is extremely rich in species (with 11,632 species described of an estimated 100,000–200,000 total) and diverse in habits, ranging in size from 0.2 to 4.0 mm. Mesostigmatic mites range from saprovores (eating dead or decaying organic matter), fungus feeders, predators living in soil, litter, beach wrack, dung, or rotting wood, to pollen and nectar feeders. Repeatedly, mesostigmatic lineages have evolved close relations with other arachnids (spiders, amblypygids, and scorpions), myriapods, insects, and vertebrates and their nests. Some are endoparasitic (living inside vertebrate respiratory tracts), and others are ectoparasitic. Many others are phoretic, with a variety of feeding habits. Some prey on mite and insect pests of orchards and stored food products. Some of these beneficial predators have been genetically engineered to be more acaricide resistant, enhancing integration of biological and chemical control of pest mite species.

Suborder Ixodida With 880 species described, of an estimated 1000–1200 total, the ticks (suborder Ixodida) are taxonomically the best-known major group of Acari primarily because of their medical and economic importance. Ticks are considerably larger than most other Acari; adults range in size from 1.7 to 12.7 mm in the unfed state. All ticks feed on the blood of terrestrial vertebrates as ectoparasites. (Sea snakes also have ticks.) Ticks are capable of carrying and transmitting to humans and their domesticated animals more kinds of disease-causing organisms than any other group of blood-feeding arthropods. These agents include viruses, spirochaetes, rickettsiae, anasplasmas, bacteria, piroplasma, and filariae. Important tick-borne human diseases include Russian spring–summer encephalitis, Colorado tick fever, African relapsing fever, Lyme disease, Rocky Mountain spotted fever, Siberian tick typhus, Q-fever, monocytic and granulocytic ehrlichiosis, Kyasanur forest disease, and tularemia.

Order Acariformes

The order Acariformes comprises three suborders—Prostigmata, Astigmata, and Oribatida—with a fourth group, the Endeostigmata (120 described species, perhaps 1200–2400 total) of uncertain rank and affinities but often treated as Prostigmata.

Suborder Prostigmata Prostigmatic mites (with 17,000 described species of an estimated 320,000–640,000 total) range in size from 0.1 to 2 mm, with a few of the giant red velvet mites as large as 16 mm, and vary widely in body form, color, habitat, and feeding adaptations. They include predatory and omnivorous species living in organic strata of soils and on algae, lichens, mosses, trees, and shrubs. The Prostigmata also include obligately plant-feeding groups, including the spider mites (Tetranychoidea), false spider mites (Tenuipalpidae), and eriophyoid mites, many of which are serious economic pests of field crops, orchards, and greenhouse plants, in some cases acting at vectors of damaging plant viruses. Some groups inhabit the soil; species of the family Nematalycidae have been found at depths of up to 3 m in sand dunes. Species of at least five families of prostigmatic mites are known from Antarctica, and others live in caves. Some prostigmatic groups live in the sea, where they are predaceous or algivorous. Many families of this order are specialized for life in springs, streams, rivers, waterfalls, lakes, bogs, or aquatic interstitial habitats of all descriptions. Some groups live in thermal waters; mites of the genus *Thermacarus* inhabit hot springs, thriving in water up to 50 °C, in which

the larvae parasitize amphibians. As in the Mesostigmata, parasitism has arisen repeatedly in independent lineages within the Prostigmata. Some are ectoparasites of slugs, scorpions, insects, or vertebrates; some are parasitoids of eggs or larvae of insects; and some are endoparasites of insect respiratory and reproductive organs, vertebrate respiratory tracts, the quills of birds, the skin of turtles, the mantle cavity of mollusks, the guts of echinoderms, or the gills of decapod crustaceans. A few species in the order are vectors of human disease (scrub typhus), and some cause mange and other skin diseases in domesticated animals. *Demodex* spp, which live (generally benignly) in hair follicles and sebaceous glands of the human face, are prostigmatic mites, as is the straw itch mite *Pyemotestritici* which can cause severe skin lesions in humans.

Suborder Oribatida The Oribatida is remarkable in being the only major group of mites in which a great diversity of species (11,000 described species of an estimated 33,000–110,000 total) has evolved without the evolution of parasitism. Although a few species are phoretic as adults on insects, the rest of this vast radiation are free-living inhabitants of the soil, forest litter, tree holes, bark, twigs, leaves, mosses, lichens, algae, freshwater vegetation, bogs, and the intertidal zone. Most oribatid mites feed on fungi, but some consume dead woody material or algae, and a few are predators on nematodes, rotifers, and other small invertebrates. Oribatida (and many species classified as Astigmatata, which are actually derived oribatids) feed on paniculate matter. Through their feeding, they help to maintain soil structure. Unknown for other mites, many oribatid species sequester calcium and other minerals in their thickened cuticle. As with Prostigmata, oribatid species are also found on continental Antarctic. Some oribatids are intermediate hosts of cestode tapeworms whose final hosts are herbivorous mammals. Oribatid mites range in size from 0.2 to 1.5 mm.

Suborder Astigmata The Astigmata, closely allied to the Oribatida, are primarily associates of arthropods, vertebrates, and other animal groups, although a few species are free-living in all life stages (with 4500 described species of an estimated 90,000–180,000 total). This lineage is thought to be derived from within the Oribatida, but a revised classification, with names of major groupings to reflect this relationship, has yet to be promulgated. A key evolutionary innovation of the Astigmata is a nonfeeding, immature stage especially adapted for phoretic dispersal or tolerance of adverse conditions. Astigmatic mites include fungus feeders, a few plant feeders, a few predators, and mites with mouthparts specialized for filter feeding in water. Some feed on algae in the intertidal zone, in water-filled tree holes, or in sap exudates. Several groups inhabit the nests of insects, mammals, and birds as ectoparasites or commensals. The specialized deutonymphs in several families are parasitic in the hair follicles of mammals. The insect hosts of some species produce acarinaria. One group lives as commensals on the gills of hermit crabs. A large radiation of astigmatic mites lives on feathers and within quills of birds as well as in the avian respiratory tract and air sacs. Another group has radiated among mammals as parasites of the skin, hair, follicles, respiratory passages, ears, and even (rarely) the digestive tract. Some astigmatic mites feed in dung or carrion, and others are important pests of stored food products, including not only grains but also stored meat and dried fish. This order also includes house dust mites (*Dermatophagoides* spp.), the itch or scabies mite (*Sarcoptes scabiei*), and mange mites.

Acknowledgments

We gratefully thank Valerie Behan-Pelletier, Clifford Desch, Mark Harvey, Gustavo Hormiga, Adriano Kury, Evert Lindquist, Roy Norton, Barry O'Connor, Lorenzo Prendini, Heather Proctor, Jeffrey Shultz, William Shear, and David Walter for comments on the manuscript, whole or in part. Estimates of mite species numbers are based on Walter and Proctor (1999).

See also: Arthropods (Terrestrial), Amazonian. Crustaceans. Insects, Overview. Invertebrates, Freshwater, Overview. Invertebrates, Marine, Overview. Invertebrates, Terrestrial, Overview. Myriapods

References

Adis J and Harvey MS (2000) How many Arachnida and Myriapoda are there worldwide and in Amazonia? *Stud. Neotropical Fauna Environ.* 35: 139–141.

Bruin J, van der Geest LPS, and Sabelis MW (eds.) (1999) *Ecology and Evolution of the Acari*. New York: Kluwer.

Coddington JA and Levi HW (1991) Systematics and evolution of spiders (Araneae). *Annu. Rev. Ecol. Syst.* 22: 565–592.

Evans GO (1992) *Principles of Acarology*. Wallingford, UK: CAB International.

Foelix RF (1996) *Biology of Spiders*, 2nd ed. Oxford: Oxford University Press.

Harvey MS (1992) The phylogeny and systematics of the Pseudoscorpionida (Chelicerata: Arachnida). *Invertebrate Taxonomy* 6: 1373–1435.

Houck MA (1994) *Mites: Ecological and Evolutionary Analyses of Life History Patterns*. New York: Chapman & Hall.

Johnston DE, Kethley JB, and O'Connor BM (1982) Acari. In: Parker SP (ed.) *Synopsis and Classification of Living Organisms*, pp. 111–169. New York: McGraw-Hill.

Kaestner A, Levi HW, and Levi LR (1980) *Invertebrate Zoology*, Vol. 2. New York: Krieger.

Krantz GW (1978) *A Manual of Acarology*. Corvallis: Oregon Univ. Book Stores.

Parker SP (ed.) (1992) *Synopsis and Classification of Living Organisms*, Vol. 2. New York: McGraw-Hill.

Polis G (ed.) (1990) *The Biology of Scorpions*. Stanford, CA: Stanford Univ. Press.

Punzo F (1998) *The Biology of Camel Spiders (Arachnida: Solifugae)*. New York: Kluwer.

Shear WA (1999) *Introduction to Arthropoda and Cheliceriformes. Microscopic Anatomy of Invertebrates: Chelicerate Arthropoda*, Vol. 8A, pp. 1–19. New York: Wiley-Liss.

Shultz JW (1990) Evolutionary morphology and phylogeny of Arachnida. *Cladistics* 6: 1–38.

Stockwell SA (1989) Revision of the phylogeny and higher classification of scorpions (Chelicerata). PhD dissertation. Berkeley: University of California.

Treat AE (1975) *Mites of Moths and Butterflies*. Ithaca, NY: Comstock.

Walter DE and Proctor HC (1999) *Mites: Ecology, Evolution and Behaviour*. Wallingford: Univ. of New South Wales Press and CABI.

Weygoldt P (1996) The evolutionary morphology of whip spiders: Towards a phylogenetic system (Chelicerata: Arachnida: Amblypygi). *J. Zool. Syst. Evol. Res.* 34: 185–202.

Wheeler WC and Hayashi CY (1998) The phylogeny of the extant chelicerate orders. *Cladistics* 14: 173–192.

Wrensch DL and Ebbert MA (eds.) (1993) *Evolution and Diversity of Sex Ratios in Insects and Mites*. New York: Chapman & Hall.

Archaea, Origin of

Costantino Vetriani, Rutgers University, New Brunswick, NJ, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 219–230, © 2001, Elsevier Inc.

Glossary

Acidophiles From the Latin *acidus* (sour) and the Greek *philos* (loving). Includes organisms that grow optimally at low pH.

Archaea One of three domains of life. From the Greek *archaios* (ancient, primitive). Prokaryotic cells; formerly called Archaeobacteria.

Bacteria One of three domains of life. From the Greek *bacterion* (staff, rod). Prokaryotic cells; formerly called Eubacteria.

Crenarchaeota One of two kingdoms of organisms of the domain Archaea. From the Greek *crene* (spring, fountain), for the resemblance of these organisms to the ancestor of the Archaea, and *archaios* (ancient). Include sulfur-metabolizing, extreme thermophiles.

Eukarya One of three domains of life. From the Greek *eu* (good, true) and *karion* (nut; refers to the nucleus). Eukaryotic cells; formerly called Eucaryotes.

Euryarchaeota One of two kingdoms within the domain Archaea. From the Greek *eurys* (broad, wide), for the relatively broad patterns of metabolism of these organisms, and *archaios* (ancient). Includes halophiles, methanogens, and some anaerobic, sulfur-metabolizing, extreme thermophiles.

Halophiles From the Greek *halos* (salt) and *philos* (loving). Includes organisms that grow optimally at high salt concentrations.

Hyperthermophiles From the Greek *hyper* (over), *therme* (heat), and *philos* (loving). Includes organisms that grow optimally at temperatures higher than 80 °C.

Korarchaeota Proposed third kingdom within the domain Archaea. From the Greek *koros* (young man), for the early divergence of this group during the evolution of the Archaea, and *archaios* (ancient). Includes a small group of ribosomal RNA sequences retrieved from geothermally heated sediments.

Mesophiles From the Greek *mesos* (middle) and *philos* (loving). Includes organisms that grow optimally at temperatures between 20 and 50 °C.

Methanogens Strictly anaerobic Archaea that produce (Greek *gen*: to produce) methane.

Phylogeny The study of the evolutionary relationships among organisms or genes.

Ribosomal rRNA Universally distributed molecule among cellular life forms. Widely used to infer the evolutionary relationships among organisms.

Thermophiles From the Greek *therme* (heat) and philos (loving). Includes organisms that grow optimally at temperatures between 50 and 80 °C.

Introduction

Historical Background

For more than 50 years, the notion that there were two basic kinds of living organisms, prokaryotes and eukaryotes, was generally accepted. The placement of a specific organism within one or the other group was based on overall morphological and phenotypic similarities. In 1965, Zuckerkandl and Pauling for the first time suggested that sequences of molecules could be used to reconstruct evolutionary history. Thus, they opened the way to molecular phylogeny, a discipline that studies the evolutionary relationships among organisms or genes. Using molecular phylogeny, in the late 1970s Carl Woese and co-workers found evidence that life consisted not of two but of three distinct groups of organisms: eukaryotes and two kinds of prokaryotes, the Eubacteria and the Archaeobacteria. Their evidence was based on the phylogenetic analysis of a single molecule, the small-subunit ribosomal RNA (SSU rRNA), which is now generally considered an excellent molecule for studying the evolutionary relationships among organisms. The tripartite view of life was formally proposed by Woese *et al.* in 1990 in the form of three domains: the Eucarya (formerly Eukaryotes), Bacteria (formerly Eubacteria), and Archaea (formerly Archaeobacteria).

Although Archaea resemble Bacteria morphologically, they differ in a variety of cellular and genetic features. From a cellular and biochemical standpoint, Archaea have some unique characteristics, such as the composition of their cell wall and membranes. From a genetic standpoint, Archaea have a unique combination of characteristics once thought to be exclusive to either the Bacteria or Eucarya. Recent analysis of genomic sequences revealed that there is a tendency for archaeal genomes to be "grab-bags" or chimeras of both bacterial and eukaryotic sequences.

Since many Archaea thrive under unusual environmental conditions that are lethal to most organisms, such as high temperature, high salinity, or extreme pH, they provide experimental models to study adaptations to extreme environments. From an evolutionary standpoint, the study of modern Archaea may offer valuable insights into the nature of the evolution of biological processes and the origin of life.

Taxonomy and Phylogeny

Molecular phylogenetic studies involving the comparison of SSU rRNA sequences revealed that the domain Archaea

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00177-5

consists of two kingdoms: the Crenarchaeota and the Euryarchaeota (Figure 1). Physiologically characterized members of the Crenarchaeota include extremely thermophilic microorganisms belonging to the order Thermoproteales (e.g., *Thermoproteus* and *Pyrobaculum*) and to the proposed order Igneococcales (e.g., *Pyrodictium* and *Desulfurococcus*), respectively, and thermoacidophilic microorganisms belonging to the order Sulfolobales (e.g., *Sulfolobus* and *Acidianus*). The Euryarchaeota comprise a metabolically versatile group, which includes all the methanogenic Archaea (orders Methanopyrales, Methanococcales, Methanobacteriales, Methanomicrobiales, and Methanosarcinales), the extreme halophiles (order Halobacteriales), the extreme thermophiles (orders Thermococcales and Archaeoglobales), and the extremely acidophilic microorganisms belonging to the order Thermoplasmales.

The recent application of molecular approaches to survey microbial diversity in natural environments has provided a new perspective on the distribution and evolutionary relationships of the Archaea. A small group of SSU rRNA sequences retrieved from geothermally heated sediments were found to be different enough from all other archaeal rRNA sequences to warrant the proposal of a new kingdom within the Archaea, namely, the Korarchaeota (see Figure 3). Furthermore, the Crenarchaeota, once thought to be restricted to high-temperature environments, have been found to be ubiquitously distributed, challenging earlier evidence that these microorganisms were confined to high-temperature ecological niches.

Diversity, Ecology, and Physiology of the Archaea

In general, the Archaea fall within three general physiological types: (i) thermophiles that live at high temperatures (some thermophiles are adapted to live in extremely low pH conditions and are known as thermoacidophiles); (ii) methanogens that inhabit strictly anaerobic environments and convert carbon dioxide and simple organic molecules to methane; and (iii) halophiles that inhabit highly saline environments and are characterized by a chemoorganotrophic, aerobic metabolism. These three major groups of Archaea are adapted to live in unusual environments and are commonly known as extremophiles. However, molecular surveys have recently revealed that Archaea are much more diverse than previously known, and that they are also widespread in more common biotopes. Non-extremophilic Archaea have been found in marine plankton, terrestrial soils, and freshwater and marine sediments. Because these newly discovered Archaea have resisted cultivation to date, their physiology and metabolic potential are unknown.

Thermophilic Archaea

Extremely thermophilic Archaea, or hyperthermophiles, comprise a group of microorganisms which are adapted to grow at temperatures higher than 80 °C. Hyperthermophilic Archaea are generally restricted to environments in which geothermal energy is available, such as hot springs, solfataras, geothermally heated marine sediments, and submarine hydrothermal vents. These environments are rich in sulfur and sulfides, and consequently many of the thermophiles have a sulfur-dependent metabolism. Extremely thermophilic Archaea can carry out a variety of respiratory processes and in most cases elemental sulfur is used as either an electron donor or an electron acceptor. Elemental sulfur is formed from geothermal H_2S either by spontaneous oxidation of H_2S with O_2 or through the reaction between H_2S and SO_2. Submarine volcanic environments include shallow (2–50 m) and deep-sea hydrothermal vents (to depths exceeding 3700 m), where the pressure, even at shallow depths, can raise the boiling point of water sufficiently to select for organisms capable of growth at temperatures higher than 100 °C.

Among the Crenarchaeota, *Pyrodictium* and *Pyrolobus* (order Igneococcales) are chemolithotrophic sulfur-dependent hyperthermophiles whose maximum growth temperatures of 110 and 113 °C, respectively, represent the upper temperature limits for life known so far. *Pyrodictium* is a strict anaerobe and grows on H_2 and S^0. *Pyrolobus* is unusual in that it is capable of reducing both NO_3^- and $S_2O_3^{2-}$ to NH_4^+ and H_2S, respectively, with H_2 as the electron donor. *Desulfurococcus* and *Staphylothermus* (order Igneococcales) are phylogenetically clearly separate from the *Pyrodictium* group (Figure 1). These coccoid or disc-shaped organisms have an optimal growth temperature higher than 85 °C and, in contrast with the *Pyrodictium* group, a maximum temperature not higher than 100 °C. They can grow chemolithoautotrophically by sulfur reduction to H_2S or heterotrophically by sulfur respiration of various organic substrates. *Thermoproteus* and *Thermofilum* (order Thermoproteales) are rod-shaped hyperthermophiles that grow in mildly acidic conditions at temperatures up to 95 °C. They are both strict anaerobes that can grow chemolithotrophically on H_2 or chemoorganotrophically on complex carbon substrates with S^0 as an electron acceptor. *Pyrobaculum aerophilum* (order Thermoproteales) is a rod-shaped hyperthermophile capable of aerobic respiration in the presence of very low oxygen concentrations ($\sim$0.3%) and nitrate reduction under strictly anaerobic conditions.

Solfataras are terrestrial volcanic environments which abound in sulfur and are gassed by emanating steam carrying CO_2, H_2, and H_2S, with temperatures up to 100 °C. Thermoacidophilic members of the order Sulfolobales inhabit acidic solfataras and oxidize H_2S or S^0 to H_2SO_4 using organic carbon or fixing CO_2 as a carbon source. *Sulfolobus* grows at, or close to, the surface of acidic solfataras at temperatures between 75 and 85 °C and low pH. No thermoacidophilic, aerobic sulfur oxidizers of the *Sulfolobus* type have been isolated from deep-sea hydrothermal vents. This is related to the steep gradient that occurs between the rising hot, anoxic hydrothermal fluid and the cold, oxygenated ambient sea water, which does not provide an ideal niche for these type of organisms. *Acidianus* (order Sulfolobales) is a close relative of *Sulfolobus* that grows at an optimum temperature of 85 °C and pH 2. *Acidianus* is a more versatile organism than *Sulfolobus* in that it is capable of growth under both aerobic and anaerobic conditions. During aerobic growth, S^0 is oxidized to H_2SO_4, whereas under anaerobic conditions H_2 is oxidized and S^0 is concomitantly reduced to H_2S.

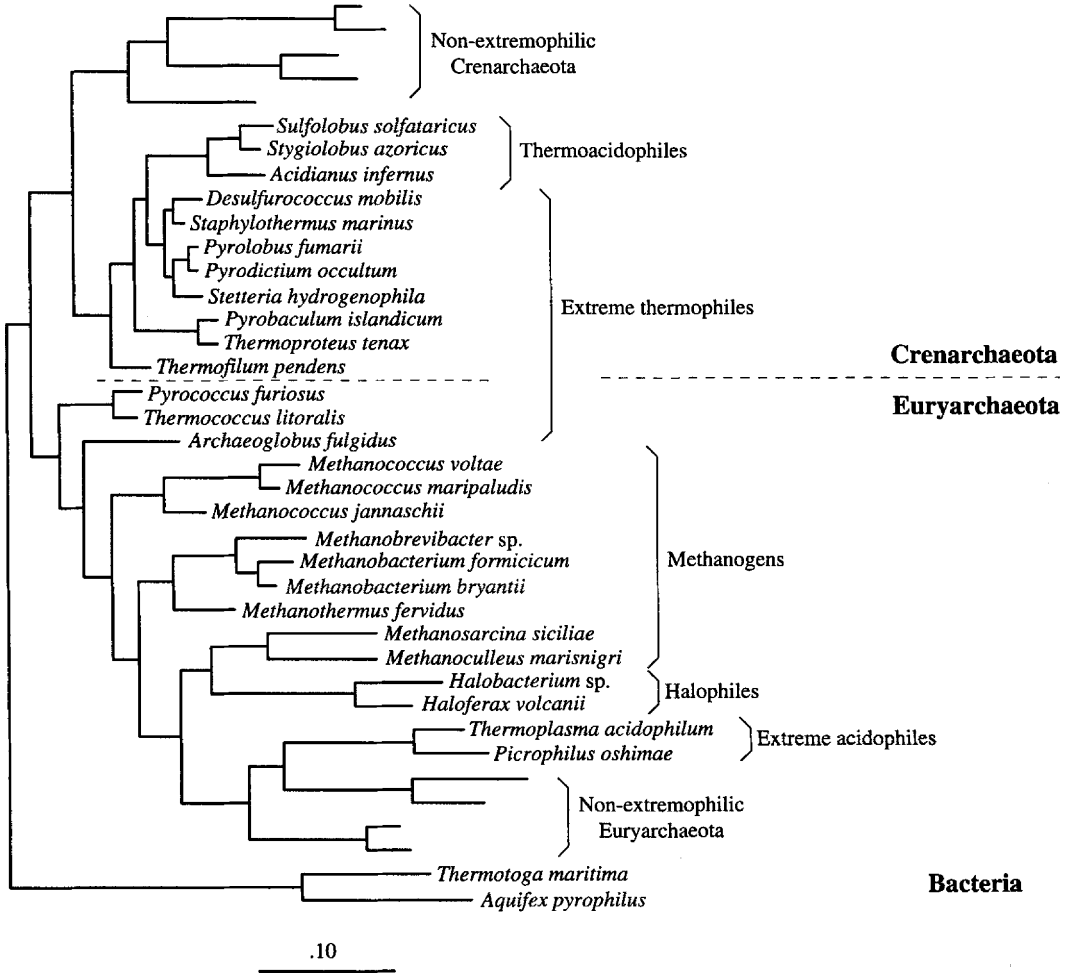


Figure 1 Unrooted phylogenetic tree showing the relationships among representative species of Archaea. Branch lengths (indicated by horizontal distances from nodes) are proportional to the evolutionary distances determined by comparative analyses of SSU rRNA sequences. The scale represents the expected number of changes per sequence position.

Among the Euryarchaeota, *Pyrococcus* and *Thermococcus* (order Thermococcales) are two closely related, coccoid, sulfur-reducing hyperthermophiles that differ primarily in their optimal growth temperatures (100 and 88 °C, respectively) (**Figure 2**). These organisms are obligate anaerobic chemoorganotrophs that utilize proteins and other complex organic mixtures and reduce S^0 to H_2S. *Archaeoglobus* (order Archaeoglobales) is an extremely thermophilic, sulfate-respiring organism which has been isolated from shallow and deep-sea hydrothermal environments. *Archaeoglobus* grows at temperatures up to 95 °C using H_2, lactate, or complex organic mixtures as electron donors and reducing SO_4^{2-} and $S_2O_3^{2-}$ to H_2S. *Ferroglobus* (order Archaeoglobales) is a coccoid, hyperthermophilic organism capable of oxidizing Fe^{2+} at neutral pH under anoxic conditions.

The order Thermoplasmales includes two moderately thermophilic organisms (optimum temperature growth approximately 60 °C), *Thermoplasma* and *Picrophilus*, isolated from coal refuse piles and acidic solfataras. *Thermoplasma* is an acidophilic, cell wall-less Archaeon that resembles bacterial *Mycoplasma* species. It can grow aerobically and anaerobically. Under anaerobic conditions there is a requirement for S^0, which is reduced to H_2S. *Picrophilus* is an extremely acidophilic, heterotrophic organism adapted to grow optimally at pH 0.7. **Table 1** summarizes some characteristics of the thermophilic Archaea.

Methanogenic Archaea

Methanogenic *Archaea* are characterized by their metabolic capability of producing methane. Methanogenesis is a strictly anaerobic respiratory means of metabolism that produces cellular energy in the form of ATP through the reduction of carbon dioxide, formate, or CO (CO_2-reducing methanogens), methanol or methylamines (methylotrophic methanogens), or acetate (aceticlastic methanogens), respectively, to methane. Typical methanogenic reactions are

$$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O \quad (CO_2\text{-reducing methanogens})$$
$$4CH_3OH \rightarrow 3CH_4 + CO_2 + 2H_2O \quad (\text{Methylotrophic methanogens})$$
$$CH_3COOH \rightarrow CH_4 + CO_2 \quad (\text{Acetoclastic methanogens})$$

Most methanogens are capable of autotrophic growth using molecular hydrogen to reduce CO_2 in a multistep reaction that requires coenzymes unique to this group of

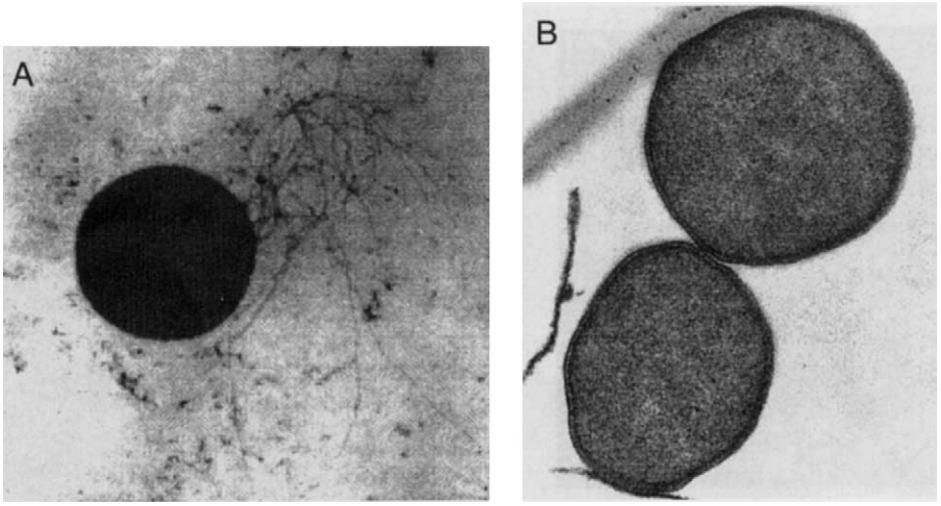


Figure 2 Electron micrographs of extremely thermophilic Archaea. (A) Platinum-shadowed *Thermococcus peptonophilus* cell showing the flagella. (B) Pyrococcus horikoshii cells.

Table 1 Habitat and general characteristics of thermophilic *Archaea*

Kingdom Order Genus	*Habitat*	*Morphology*	*Temperature (°C)*		
			Optimum	*Maximum*	*Optimum pH*
Crenarchaeota					
Igneococcales					
Pyrodictium	Coastal and deep-sea hydrothermal vents	Disc-shaped	105	110	6
Pyrolobus	Deep-sea hydrothermal vent, Mid-Atlantic Ridge	Lobed cocci	105	113	5.5
Desulfurococcus	Hot springs and solfataras	Sphere	85	95	6
Staphylothermus	Submarine hydrothermal vent, Italy	Spheres in clumps	92	98	6–7
Thermoproteales					
Thermoproteus	Hot springs and solfataras	Rod	88	96	6
Thermofilum	Solfataras	Rod	88	95	5.5
Pyrobaculum	Submarine hydrothermal vents and solfataras	Rod	95	100	6
Sulfolobales					
Sulfolobus	Acidic solfataras	Lobed sphere	75–85	87	1–5
Acidianus	Acidic solfataras	Sphere	85–90	95	2
Desulfurolobus	Acidic solfataras	Lobed sphere	80	87	2.5
Stygiolobus	Hot spring, Azores	Lobed sphere	80	89	3
Euryarchaeota					
Thermococcales					
Pyrococcus	Coastal and deep-sea hydrothermal vents	Sphere	100	106	6–8
Thermococcus	Coastal and deep-sea hydrothermal vents	Sphere	88	98	6–7
Archaeoglobales					
Archaeoglobus	Coastal and deep-sea hydrothermal vents	Cocci	83	95	7
Ferroglobus	Submarine hydrothermal vent, Italy	Cocci	85	95	7
Thermoplasmales					
Thermoplasma	Coal refuse pile, solfataras	Sphere, filaments	59	63	2
Picrophilus	Acidic solfatara, Japan	Cocci	60	65	0.7

organisms, such as methanofuran, methanopterin, F_{420}, F_{430}, and coenzyme M. Methanofuran is a low-molecular-weight phenol derivative involved in the first step of methanogenesis. Methanopterin is a pterin-containing coenzyme which resembles the vitamin folic acid and serves as a Cl carrier during the reduction of CO_2 to CH_4. Coenzyme F_{420} is a flavin derivative involved in redox reactions. The oxidized form of F_{420} has a characteristic blue-green fluorescence at 420 nm and is very useful in recognizing methanogens microscopically. Coenzymes M and F_{430} play a crucial role in the final step of the

Table 2 General characteristics of methanogenic *Archaea*

Order Genus	*Morphology*	*Substrates*	*Temperature (°C)*
Methanobacteriales			
Methanobacterium	Long rods	$H_2 + CO_2$, formate	35–40
Methanobrevibacter	Short rods	$H_2 + CO_2$, formate	30–38
Methanosphaera	Cocci	Methanol + H_2 (both needed)	36–40
Methanothermus	Rods	$H_2 + CO_2$	83–88
Methanococcales			
Methanococcus (mesophilic sp.)	Irregular cocci	$H_2 + CO_2$, pyruvate + CO_2, formate	35–40
Methanococcus (thermophilic sp.)	Irregular cocci	$H_2 + CO_2$	88
Methanomicrobiales			
Methanomicrobium	Short rods	$H_2 + CO_2$, formate	40
Methanogenium	Irregular cocci	$H_2 + CO_2$, formate	30–57
Methanospirillum	Spirilla	$H_2 + CO_2$, formate	30–40
Methanoplanus	Plate-shaped cells	$H_2 + CO_2$, formate	32–40
Methanoculleus	Coccus	$H_2 + CO_2$, formate	37–60
Methanosarcinales			
Methanosarcina	Large irregular cocci in packets	$H_2 + CO_2$, methanol, methylamines, acetate	35–50
Methanolobus	Irregular cocci in aggregates	Methanol, methylamines	30–40
Methanohalobium	Irregular cocci	Methanol, methylamines	50
Methanococcoides	Irregular cocci	Methanol, methylamines	23–35
Methanohalophilus	Irregular cocci	Methanol, methylamines, methyl sulfides	26–36
Methanothrix	Long rods to filaments	Acetate	35–60
Methanopyrales			
Methanopyrus	Rods in chains	$H_2 + CO_2$	100

methanogenesis as part of the methyl reductase system. F_{430} is a yellow, soluble, nickel-containing tetrapyrrole that, unlike F_{420}, does not fluoresce.

All methanogenic Archaea are strictly anaerobic and occupy anoxic habitats, such as sediments, marshes, waterlogged soils, submarine hydrothermal vents, and the digestive tracts of animals. Methanogens are usually abundant in environments depleted in Fe^{3+}, NO^{3-}, and SO_4^{2-} because these substrates stimulate growth of anaerobic bacteria that can outcompete them. Therefore, methanogenesis is usually limited in sulfate-rich marine environments because the methanogens have to compete with sulfate-reducing bacteria for H_2. Since oxidation of H_2 with sulfate as the electron acceptor is thermodynamically more favorable than when CO_2 is the electron acceptor (as in methanogenesis), sulfate reducers are usually favored. However, since methylotrophic methanogens can use noncompetitive substrates (such as methylamines, which are inaccessible to sulfate-reducers), they can be detected in some sulfate-rich environments. Most methanogens have been found to contain the nitrogenase protein complex; therefore, they are also capable of anaerobic nitrogen fixation under certain growth conditions.

The methanogenic Archaea is a diverse group with many species; it is the largest taxonomic group of Archaea. The order Methanobacteriales includes mainly rod-shaped methanogens that grow by CO_2 reduction, although a few species of the genus *Methanosphaera* are cocci that grow only by using H_2 to reduce methanol to methane. *Methanobrevibacter* species are very short rods which use H_2 or formate to reduce CO_2 to CH_4. The genus *Methanothermus* includes extremely thermophilic methanogens isolated from terrestrial geothermal springs. They can grow at 88 °C on CO_2 and H_2.

The order Methanomicrobiales encompass a genetically diverse group of CO_2-reducing methanogens that grow at temperatures ranging from 30 to 60 °C.

The order Methanosarcinales includes morphologically diverse organisms such as *Methanosarcina*, which form irregular spheroid bodies occurring alone or in aggregates of cells, and *Methanothrix*, which forms sheathed rods. Both *Methanothrix* and *Methanosarcina* are acetoclastic methanogens, although the latter can also use methanol and methylamines as substrates for methanogenesis.

The order Methanococcales includes mesophilic and extremely thermophilic methanogens characterized by irregular coccoid morphology. *Methanococcus jannaschii* was isolated from a deep-sea hydrothermal vent system where it grows at 88 °C on CO_2 and H_2 released in volcanic gases.

Methanopyrus (order Methanopyrales) is an extremely thermophilic methanogen isolated from hydrothermally influenced sediments, and it is capable of autotrophic growth at temperatures up to 110 °C. **Table 2** summarizes some characteristics of the methanogenic Archaea.

Halophilic Archaea

The halophilic Archaea occur in environments characterized by high salinity. Salt-neutral lakes, saline soils, solar evaporation marine salterns (areas where sea salt is produced), and subsurface haline deposits are among the most common habitats for halophiles. Most of the halophilic Archaea are

Table 3 Habitat and general characteristics of halophilic *Archaea*

Order Genus Species	*Habitat*	*Morphology*	*Physiological characteristics*
Halobacteriales			
Halobacterium		Rod-shaped	
H. Salinarum	Salted fish, hypersaline lakes		Amino acids, phototrophic
Haloferax		Rod/disk-shaped	
H. mediterranei	Salterns		Carbohydrates
H. denitrificans	Salterns		Nitrate respiration
Halorubrum		Rod-shaped	
H. lacusprofundi	Deep lake, Antarctica		Carbohydrates
H. sodomense	Dead Sea		Phototrophic
Haloarcula		Pleiomorphic	
H. vallismortis	Death Valley		Carbohydrates
H. marismortui	Dead Sea		Nitrate respiration
Halococcus		Cocci	
H. morrhuae	Dead sea		Amino acids, Nitrate respiration
H. saccharolyticus	Salterns		Carbohydrates
Natronobacterium		Rod-shaped	
N. gregoryi	Soda lakes		Organic acids, ph 9.5–10
Natronococcus		Cocci	
N. occultus	Soda lakes		Amino acids, ph 9.5–10
Natrialba		Rod-shaped	
N. Asiatica	Sea sand		Nonpigmented, extreme halophile

red or orange because of the presence of carotenoid pigments in the cell envelope. Frequently, due to their abundance, they impart a red color to the brine. Alkaliphilic halophiles are found in soda lakes, which are highly alkaline environments whose high pH (8 to >12) is due to high levels of carbonate. Halophiles require at least 1.5 *M* NaCl for growth, and optimal salt concentrations are usually in the range of 2–4 *M* NaCl.

Neutrophilic and alkaliphilic halophiles both belong to the order Halobacteriales. The members of the neutrophilic group, represented by the genera *Halobacterium*, *Halococcus*, *Haloarcula*, and *Haloferax*, grow optimally under conditions of high magnesium and sodium concentrations (0.5 and 4.0 *M*, respectively) at pH ranging from 5 to 8. Alkaliphilc halophiles (genera *Natronobacterium* and *Natronococcus*) grow optimally at low magnesium concentrations and pH of approximately 10.

Halophilic Archaea have evolved several physiological adaptations that permit their growth in habitats with salt concentrations that cause cellular dehydration and protein denaturation in other organisms. Halophilic Archaea such as *Halobacterium* resist high salt concentrations by pumping large amounts of K^+ from the environment into the cell such that the concentration of K^+ inside the cell is greater than the concentration of Na^+ outside the cell. This mechanism allows *Halobacterium* to remain in positive water balance and avoid dehydration.

The halophilic Archaea are aerobic and grow heterotrophically using carbohydrates, alcohols, organic acids, and amino acids. *Halobacterium* species are normally aerobic but can grow anaerobically in the presence of light. Under oxygen-limiting conditions and in the presence of light, *Halobacterium* inserts large amounts of a protein called bacteriorhodopsin into the cytoplasmic membrane. This purple pigment, which adsorbs light strongly at approximately 570 nm, acts as a light-driven proton pump and leads to the establishment of an electrochemical membrane potential. Its equilibration can be accompanied by the generation of ATP. Thus, bacteriorhodopsin facilitates a special kind of photophosphorylation. The energy obtained by this mechanism complements that obtained from aerobic substrate oxidation. **Table 3** summarizes some characteristics of the halophilic Archaea.

Nonextremophilic Archaea

The application of SSU rRNA-based molecular techniques to survey natural microbial populations has recently revealed that Archaea are virtually ubiquitous. Therefore, our view of the Archaea as highly specialized microorganisms adapted to survive in extreme environments is gradually changing to that of a very diverse group of organisms, including more moderate representatives.

Non-extremophilic Archaea have been identified in a variety of environments, such as temperate and cold marine planktonic habitats, freshwater sediments, soils, terrestrial subsurface environments, and permanently cold deep-sea sediments, and in association with several marine metazoans. To date, none of the non-extremophilic Archaea have been isolated in pure culture; therefore, their phenotypic and physiological characteristics remain unknown. However, their ubiquitous distribution and high phylogenetic diversity suggests a wide range of ecological and physiological adaptations.

The distribution and phylogenetic affiliation of the non-extremophilic Archaea are summarized in **Table 4**.

Origin and Evolution of the Archaea

Archaea occupy a pivotal phylogenetic position between the two other domains of life, Bacteria and Eukarya (**Figure 3**).

With the recent sequencing of the entire genomes of several Archaea, it has become evident that archaeal genomes share bacterial and eukaryotic features. Furthermore, the availability of an increasing number of gene sequences from organisms belonging to the three domains of life allows the identification of a "universal" set of protein families present in all three domains. This universal set of proteins can then be used as an estimate of the genome content and the cellular processes that were present in the last universal ancestor of living organisms. However, factors such as horizontal gene transfers between distantly related groups and gene loss make it difficult to draw conclusions about the biology of the last universal ancestor, and different theories have been proposed.

The following sections describe different traits of the Archaea which can be used to infer some properties of the last universal ancestor as well as to provide insights of the origin and evolution of the Archaea.

Table 4 Distribution and phylogenetic affiliation of nonextremophilic *Archaea*

Distribution	*Phylogenetic affiliation*
Marine Habitats	
Surface and deep waters (up to 3000 m)	*Crenarchaeota*, *Euryarchaeota*
Temperate coastal sediments (12 m)	*Crenarchaeota*, *Euryarchaeota*
Low-temperature deep-sea sediments (1500 to 4500 m)	*Crenarchaeota*, *Euryarchaeota*
Temperate microbial mats at deep-sea hydrothermal vent	*Crenarchaeota*, *Euryarchaeota*
Antarctic low-temperature surface waters	*Crenarchaeota*, *Euryarchaeota*
Salt marsh	*Euryarchaeota*
Associated with Marine Metazoans	
Gut of abyssal holothurian *Oneirophanta mutabilis* (4870 m)	*Crenarchaeota*
Digestive tract of fish	*Crenarchaeota*, *Euryarchaeota*
Tissues of sponge *Axinella mexicana* (10–20 m)	Crenarchaeon, *Crenarchaeota symbiosum*
Freshwater Habitats	
Lake sediments	*Crenarchaeota*, *Euryarchaeota*
Terrestrial Habitats	
Soils	*Crenarchaeota*, *Euryarchaeota*
Subsurface paleosol (188 m)	*Crenarchaeota*
Contaminated aquifer	*Crenarchaeota*, *Euryarchaeota*
Rice roots	*Crenarchaeota*, *Euryarchaeota*

Early Respiratory Processes

The presence of genes encoding for homologous proteins in organisms belonging to both domains Archaea and Bacteria allows for the reconstruction of cellular processes that were present early in the history of life on Earth. In particular, terminal oxidases belonging to oxygen, nitrate, sulfate, and sulfur respiratory pathways are present in both domains.

Cytochrome oxidase, which catalyzes the reduction of oxygen to water and acts as a redox-linked proton pump, is the key enzyme of aerobic metabolism. This enzyme is present in members of both the Crenarchaeota (*Sulfolobus*, *Acidianus*, and *Pyrobaculum*) and the Euryarchaeota (*Halobacterium* and *Natronobacterium*) as well as in many thermophilic Bacteria (e.g., *Thermus* and *Aquifex*). It is plausible that in the early atmosphere, prior to the advent of oxygenic photosynthesis, microaerophilic organisms similar to the present-day archaeon *Pyrobaculum* may have been able to thrive under extremely low oxygen concentrations, known to have existed as a result of the photolysis of water.

The ability to use sulfate as a terminal electron acceptor is a characteristic common to many mesophilic and thermophilic

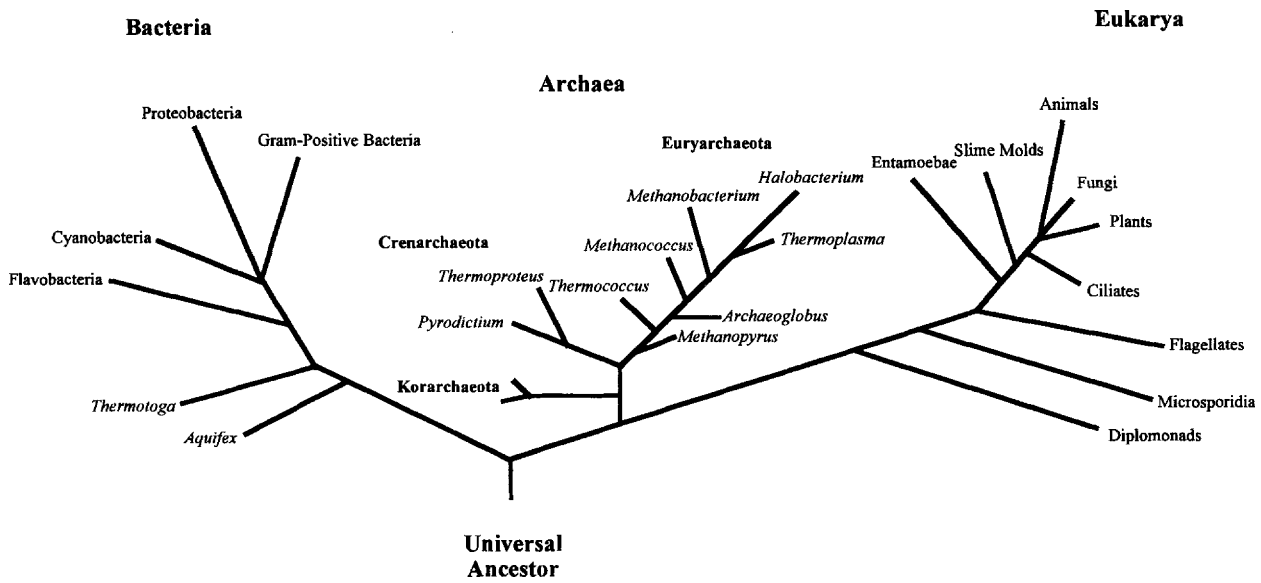


Figure 3 Schematic representation of the rooted "universal" SSU rRNA phylogenetic tree showing the position of the Archaea relative to the Bacteria and the Eukarya.

bacterial species and to the thermophilic archaeon *Archaeoglobus*. In addition, a sulfite reductase-type protein has been identified in *Pyrobaculum*. The enzyme dissimilatory sulfite reductase catalyzes the six-electron reduction of (bi)sulfite to sulfide, which is the central energy-conserving step of sulfate respiration. Phylogenetic evidence suggests an ancestral origin of sulfate respiration, a finding which is consistent with the notion that sulfates of magmatic origin were common in the Archaean time. In addition to sulfate reduction, *Archaeoglobus* produces very small quantities of methane. Factor F_{420} and methanopterin, two coenzymes involved in methanogenesis, are also present in *Archaeoglobus*, although this organism appears to lack other cofactors normally present in methanogens. This apparently intermediate type of metabolism between methanogenesis and sulfur reduction suggested that *Archaeoglobus* may represent a transitional form in the diversification of Archaea from sulfur-respiring to methane-producing and halophilic Archaea. The branching position of *Archaeoglobus* in the SSU rRNA tree is consistent with this interpretation (**Figure 1**).

Sulfur respiration has also been proposed to be an ancestral mode of energetic metabolism on the base of two considerations: (i) Volcanic-derived S^0 was probably one of the most abundant electron acceptors in the early atmosphere, and (ii) the capacity to reduce S^0 to H_2S is common among anaerobic, hyperthermophilic members of both Archaea and Bacteria.

Adaptive Features to High Temperature

Archaea exhibit a strong adaptive capacity to extremely high temperatures. Since microorganisms are isothermal, they have to develop strategies to avoid thermal stress. Thus, in order to survive at temperatures in excess of 100 °C, hyperthermophiles have to adapt their cell inventory to function optimally at the temperatures in the niches they occupy. Many different mechanisms have been hypothesized by which hyperthermophiles can simultaneously retain stability and plasticity at high temperatures.

One important mechanism of thermoadaptation is the biosynthesis of thermostable proteins and enzymes that maintain sufficient structural integrity to allow optimal catalytic efficiency at high temperatures. Thermostable proteins are often characterized by a very efficient packing density in the hydrophobic core of the molecule as well as by shorter connecting elements (loops) between regions of secondary structure. An elevated number of ionic interactions has frequently been found in highly thermostable proteins. Ionic interactions among monomers are thought to stabilize multimeric proteins at high temperatures. Overall, these characteristics lead to decreased flexibility in the polypeptide chain, a required feature that compensates for increased thermal fluctuations at high temperatures.

Most living beings need the strand-opening potential of negative DNA supercoiling to allow transcription and other DNA-dependent processes. In bacterial nucleo-proteins and in eukaryal chromosomes, DNA has a negative superhelicity as a result of topoisomerase activity and wrapping of DNA around histone cores, respectively. Histone-like proteins may play an important role in certain thermophilic Archaea, protecting DNA against thermal denaturation and degradation. The DNA-binding proteins isolated from both *Thermoplasma acidophilum* and *Methanothermus fervidus* stabilize double-stranded DNA molecules *in vitro* by increasing their melting temperatures. A novel class of topoisomerases has been identified in both hyperthermophilic Archaea (e.g., *Sulfolobus*) and Bacteria (e.g., *Thermotoga*). Since these enzymes catalyze positive supercoiling into the DNA, they have been referred to as reverse gyrases. Reverse-gyrase activity is widely distributed in hyperthermophiles and appears to be, at least at critical regions, a requirement to prevent DNA denaturation at high temperatures. Thus, the reverse girase may be considered as an ancestral trait retained by present-day hyperthermophilic Archaea.

Evolution

Hyperthermophilic Archaea have been proposed as analogs for the early life on Earth. If life arose and evolved under high-temperature, reduced, sulfur-rich conditions, then the requirement of thermophilic Archaea for high temperature environments as well as their predominantly anaerobic, sulfur-metabolizing phenotype suggest they are good analogs to test such a theory. Two independent lines of evidence suggest that the phenotype of the extant hyperthermophiles may represent remnants of the ancestral Archaea. First, the hyperthermophilic trait is the only one found in both archaeal kingdoms, the Crenarchaeota and the Euryarchaeota. Second, most hyperthermophilic lineages are more slowly evolving than any other archaeal lineage. This is particularly evident when the rate of evolution of the SSU rRNA genes (indicated as the branch length in the phylogenetic tree) is considered (**Figure 1**). Mesophilic methanogens, halophiles, and the recently discovered non-extremophilic Archaea appear to evolve much faster than hyperthermophiles (**Figure 1**). Furthermore, mesophilic methanogens, such as members of the Methanococcales, evolve more rapidly than their extremely thermophilic relatives. Therefore, the assumption that more slowly evolving lineages (thermophilic) tend to retain more of the ancestral traits than do rapidly evolving lineages (non-thermophilic) suggests a high-temperature ancestry of the Archaea, which later adapted to lower temperatures.

A different theory supporting a non-thermophilic ancestor to extant life forms has been recently formulated on the base of the strong correlation between the Guanine + Cytosine (G + C) nucleotide content of rRNA genes and the optimal growth temperature of microorganisms. The analysis of rRNA genes from hyperthermophiles suggests that survival in high-temperature conditions requires a high G + C content in rRNA genes. However, the estimated G + C contents of ancestral organisms appeared to be incompatible with life at high temperature and suggested a moderate environmental growth temperature for the universal ancestor. This theory leads to the conclusion that present-day thermophily evolved from mesophilic organisms via adaptation to high temperature.

In principle, the nature of an ancestral organism may be inferred from the distribution of homologous traits among its descendants. However, the reconstruction of early evolutionary histories is complicated by phenomena such as horizontal

gene transfer (gene swapping) among organisms. As new microbial genomes have become available, it has become increasingly evident that extensive gene swapping, even among distantly related organisms, may have occurred. This tends to blur the history of the early stages of life on Earth; therefore, the nature of the last universal ancestor remains elusive. At the same time, comparative functional genomics is providing a new perspective of the characteristics of the universal ancestor. Because of their evolutionary history, Archaea are likely to play a critical role in our understanding of the early events in the evolution of cells and the origin of life.

See also: Eukaryotes, Origin of. Microbial Biodiversity. Origin of Life, Theories of. Psychrophiles. Thermophiles, Origin of

References

Aravalli RN, She QX, and Garrett RA (1998) Archaea and the new age of microorganisms. *TREE* 13: 190–194.

Barns SM, Delwiche CF, Palmer JD, and Pace NR (1996) Perspectives on archaeal diversity, thermophily and monophily from environmental rRNA sequences. *Proc. Natl. Acad. Sci. USA* 93: 9188–9193.

Blöch E, Burggraf S, Fiala G, *et al.* (1995) Isolation, taxonomy and phylogeny of hyperthermophilic microorganisms. *World J. Microbiol. Biotechnol.* 11: 9–16.

Brown JR and Doolittle WF (1997) Archaea and the Prokaryote-to-Eukaryote transition. *Microbiol. Mol. Biol. Rev.* 61: 456–502.

Bult CJ, White O, Olsen GJ, *et al.* (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273: 1058–1073.

Burggraf S, Huber H, and Stetter KO (1997) Reclassification of the crenarchaeal orders and families in accordance with 16S rRNA sequence data. *Int. J. Syst. Bacteriol.* 47: 657–660.

Castresana J and Moreira D (1999) Respiratory chains in the last common ancestor of living organisms. *J. Mol. Evol.* 49: 453–460.

Ciaramella M, Cannio R, Moracci M, Pisani FM, and Rossi M (1995) Molecular biology of extremophiles. *World J. Microbiol. Biotechnol.* 11: 71–84.

DeLong EF (1992) Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* 89: 5685–5689.

Fuhrman JA, McCallum K, and Davis AA (1992) Novel major archaebacterial group from marine plankton. *Nature (London)* 356: 148–149.

Galtier N, Tourasse T, and Gouy M (1999) A nonhyperthermophilic common ancestor to extant life forms. *Science* 283: 220–221.

Gonzalez JM, Kato C, and Horikoshi K (1995) *Thermococcus peptonophilus* sp. npv., a fast-growing, extremely thermophilic archaebacterium isolated from deep-sea hydrothermal vents. *Arch. Microbiol.* 164: 159–164.

Konig H and Stetter KO (1989) Archaeobacteria. In: Staley JT, Bryant MP, Pfennig N, and Holt JG (eds.) *Bergey's Manual of Determinative Bacteriology,* vol. 3, pp. 2171–2173. Baltimore: Williams & Wilkins.

Lopez-Garcia P (1999) DNA supercoiling and temperature adaptation: A clue to early diversification of life? *J. Mol. Evol.* 49: 439–452.

Vetriani C, Jannasch HW, MacGregor BJ, Stahl DA, and Reysenbach A-L (1999) Population structure and phylogenetic characterization of marine benthic Archaea in deep-sea sediments. *Appl. Environ. Microbiol.* 65: 4375–4384.

Vetriani C, Maeder DL, Tolliday N, *et al.* (1998) Protein thermostability above 100 °C: A key role for ionic interactions. *Proc. Natl. Acad. Sci. USA* 95: 12300–12305.

Wagner M, Roger AJ, Flax JL, Brusseau GA, and Stahl DA (1998) Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J. Bacteriol.* 180: 2975–2982.

Woese CR, Kandler O, and Wheelis ML (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* 87: 4576–4579.

Zuckerkandl E and Pauling L (1965) Molecules as documents of evolutionary history. *J. Theor. Biol.* 8: 357–366.

Arctic Terrestrial Ecosystems

Terry V Callaghan, University of Sheffield, Sheffield, UK, and Royal Swedish Academy of Sciences, Stockholm, Sweden
Nadya Matveyeva, Komarov Botanical Institute, Russian Academy of Sciences, Sankt-Petersburg, Russia
Yuri Chernov, Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia
Niels M Schmidt, Aarhus University, Roskilde, Denmark
Rob Brooker, James Hutton Institute, Aberdeen, UK
Margareta Johansson, Lund University, Sweden, and Royal Swedish Academy of Sciences, Sweden

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Terry V. Callaghan, Nadya Matveyeva, Yuri Chernov, Rob Brooker, volume 1, pp 231–247, © 2001, Elsevier, Inc.

Glossary

Arctic Geographically, the region lying to the north of latitude 66.7° N, but environmentally and in the context of this article, the region to the north of the climatically controlled northern latitudinal treeline that corresponds approximately to the mean July isotherm of 10 °C.

Permafrost The phenomenon of ground materials (e.g., soil and bedrock) that are at or below 0 °C for 2 or more years. Usually, but not always, any water present is in the solid state. The Arctic is characterized by the presence of large, continuous areas of permafrost that have the form of lower soil layers permanently frozen with a shallow (usually <1 m) 'active layer' that freezes and thaws each year and accommodates belowground biological activity.

Refugia In the context of the Arctic, these are land areas that were not covered by ice sheets or glaciers during the last glaciation. Consequently, some biota could survive there and recolonize adjacent areas when the ice retreated. Therefore, these areas are associated with high biodiversity and endemism. Refugia usually occurred in coastal areas (now continental shelves) and ice-free mountain tops or "nunataks."

Sub-Arctic The ecotone (ecological boundary zone) connecting the treeless tundras in the north with the taiga or coniferous boreal forests in the south. The area is characterized by the presence of scattered, deciduous or coniferous trees of low stature and is sometimes termed "forest tundra." (Other Arctic vegetation zones are described in the text.)

Treeline Here, denotes the northern latitudinal and altitudinal distribution limit of vegetation in which trees are a conspicuous but not necessarily a dominant element. Individual, isolated trees or 'oases' of trees in environmentally benign areas are considered to be beyond the treeline.

Tundra This is a type of vegetation characteristically occupying the Arctic. However, the term is used in many ways, from characterizing individual plant communities of the Arctic, which consist of dwarf shrubs and sedges, to characterizing all vegetation above the altitudinal treeline and between the latitudinal treelines and the poles in both hemispheres. In this article, the term is used in the Russian sense to characterize Arctic vegetation lying between the taiga and the region of the polar deserts.

Introduction

Although Arctic environments have undergone dramatic changes over millions of years, the twentieth century was associated with particularly rapid changes in many aspects of the physical environment and particularly with rapid cultural and sociological changes in the north, leading to the increasing exploitation and fragmentation of wilderness areas. Because future environmental changes were predicted to be even greater during this century (ACIA, 2005; IPCC, 2007), it became increasingly necessary to document, monitor, and understand the biological resources of the Arctic. Understanding the patterns, causes, and consequences of biodiversity throughout the Arctic's lands, as highlighted by Chapin and Körner (1995), is an important aspect of this challenge. However, the Arctic should not be seen as a region remote from the more populated regions of the world at lower latitudes: Several hundreds of millions of birds overwintering in temperate regions migrate to summer nesting grounds in the Arctic, and the functions of many Arctic ecosystems in sequestering the greenhouse gas carbon dioxide in soils have contributed to cooling the earth's surface since at least the end of the last ice age.

This article introduces the reader to current patterns of biodiversity in terrestrial Arctic ecosystems and presents further details of the physical, historical, and biotic environments that have shaped these patterns. The authors then show the importance of the Arctic's biodiversity in terms of ecosystem function and its provision of resources for human life support and welfare (Chapin *et al.*, 2005). Finally, they show how the Arctic's biota is threatened by numerous factors and discuss the major challenges to further understanding the biota.

Definitions and General Environmental Characteristics of the Arctic

The Arctic is difficult to define, basically because of two confounding issues – the earth's inclination to the sun and environmental, including climatic, conditions (Nutall and Callaghan, 2000). Strictly, the Arctic is the area north of latitude 66.7° N, where the sun does not set below the horizon at midnight on midsummer's night and does not rise above the horizon at midday on midwinter's day. This astronomical feature sets a definitive photoperiod for the Arctic and reduces the amount of solar heat absorbed by the earth's surface. Hence, the Arctic is characterized by having long, cold, and dark winters and short, cool summers with the midnight sun visible.

The energy balance of the earth's surface in the north is complicated, however, by the connectivity of the Arctic's oceans and climate with those further south. The transport of heat from southern latitudes into the Arctic in warm ocean currents creates environments which are much warmer than those that can be considered as typically Arctic, for example, those in northern Fennoscandia (**Figure 1**), which are warmed by the Gulf Stream. In contrast, where returning cold water flows southward, unusually cold climates exist for particular latitudes: Tundra environments exist in eastern Canada (**Figure 1**) at latitudes at which British agriculture and forestry thrive. Proximity to the coast affects terrestrial temperatures by the formation of mists and increasingly by loss of sea ice (Bhatt *et al.*, 2010). On land, topography also affects temperatures locally, particularly in summer, whereas proximity to large lakes affects temperatures in winter (Zhenlin *et al.*, 2011). As one moves closer to the North Pole, progressively smaller differences in topography become more important in determining temperature.

Often, therefore, the definition of the Arctic depends on the subject and the scientific discipline. Because all biological processes depend on chemicophysical reactions, temperature is an important determinant of biological activity. Also, because nearly all food webs ultimately depend on primary producers that usually require light for photosynthesis, the strong seasonal variations in photoperiod are also important for biological activity. For biologists, therefore, definitions that incorporate the sensitivity of flora and fauna to low temperatures and specific photoperiods are particularly appropriate. Thus, the 10° summer isotherm (where the mean monthly temperature is at or below 10 °C) and the northern limit of forests (the latitudinal treeline), both of which represent the southern boundary of the "low Arctic" in Western terminology and the "tundra zone" in Russian terminology, are commonly used to define the Arctic (**Figure 1**).

The extent of the Arctic is totally dependent on its definition. Using botanical definitions for the Arctic, Bliss and Matveyeva (1992) calculate that it comprises approximately 7.5 million km^2, which is approximately 5.5% of the earth's land surface. The Arctic stretches over 33° of latitude from 84° N in Greenland to 51° N in Hudson Bay, Canada (see Jonasson *et al.* in Nutall and Callaghan, 2000). Almost 2 million km^2 is permanent snow and ice. Distances around the Arctic decrease dramatically toward the North Pole. For example, the circumference of the equator is 40,076 km, and only 13,752 km at the 70° parallel. This has great implications for connectivity between animal and plant populations and species.

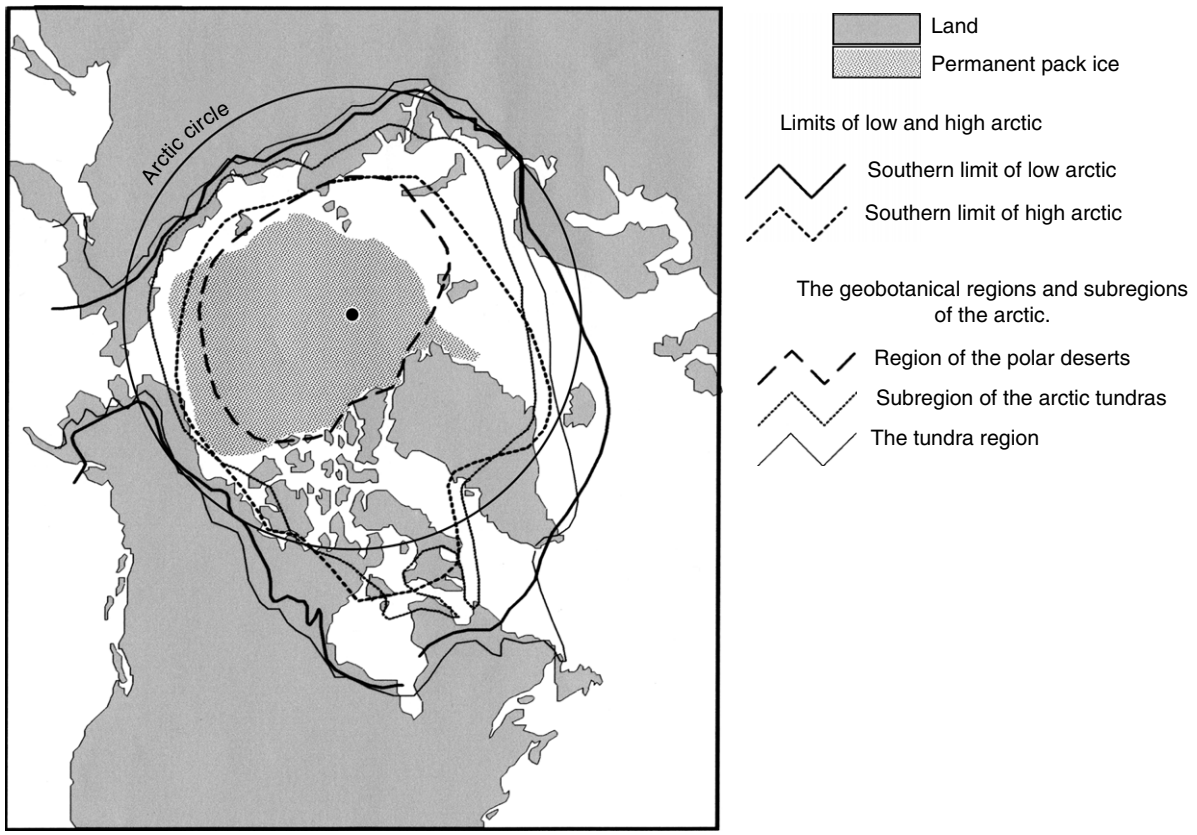
The complexity of defining the Arctic is amplified when the region is subdivided into different zones. Although it is clear that the Arctic is not a homogeneous environment, definitions of subzones vary between Eastern and Western scientific traditions (Bliss and Matveyeva, 1992; Walker *et al.*, 2005; **Table 1**).

Despite these classifications, the 'sub-Arctic' region is often referred to and represents the ecotone between the tundra and taiga and the forest tundra. In fact, there is a continuous gradient of environmental severity within the Arctic from the boreal forest zone at its southern boundary to the polar deserts of the far north, even if this is interrupted in some places by mountain chains and water bodies. Along this gradient, the nature of the species changes from erect woody plants to low-growing herbs, mosses, and lichens, and the complexity of the canopy in vertical and horizontal dimensions decreases from south to north.

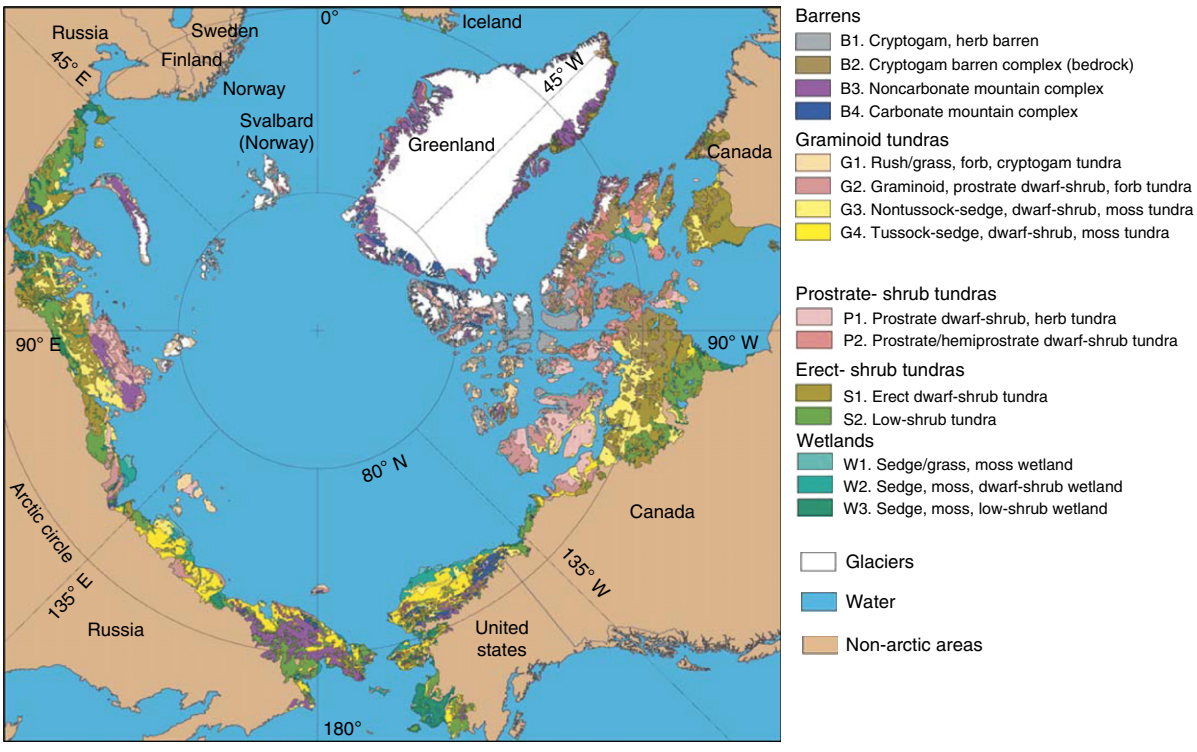
Great temperature variations exist throughout the Arctic. From the southern boundary, where the mean July temperature is 10 °C and there are more than 1000° days >0 °C, there is a decrease to the high Arctic, where the mean July temperature can be less than 2 °C with just greater than 110 day year^{-1}. Precipitation also decreases toward the north, from more than 1000 mm year^{-1} to approximately 50 mm year^{-1}. Low temperatures and low precipitation exert critical limitations on the productivity of Arctic ecosystems (Jonasson *et al.* in Nutall and Callaghan, 2000). The net primary production varies from approximately 1 g m^{-2} year^{-1} in polar deserts to 150–800 g m^{-2} year^{-1} in the sub-Arctic. In the low Arctic, values can range between 100 and 1200. Although there is a good correlation between temperature and productivity, the mechanisms are complex. In many areas, it has been suggested that the most important aspect of temperature is indirect in that it constrains the rate of nutrient cycling in cold Arctic soils, thereby resulting in nutrient limitation throughout most of the Arctic (Jonasson *et al.* in Nutall and Callaghan, 2000).

The great spatial differences in temperature regimes throughout the current Arctic are accompanied by even greater variations over time. Arctic landmasses and oceans are relatively young and have been formed by the migrations of landmasses from southerly latitudes. Also, most of the Arctic's land and oceans have been covered even in summer by ice, but only during the past 1.8 million years. The last glacial period, which ended approximately 10,000 years ago displaced terrestrial, freshwater, and marine species southward from much of the Arctic. A similar process occurred over large lowland areas of the Russian Arctic when they were entirely inundated by the marine transgression (Aleksandrova, 1980; Yurtsev, 1997).

The Arctic's current biota can be seen, therefore, as relatively young and, on land, it can be interpreted as depauperate remnants of previous floras and faunas that existed in preglacial eras. For example, forests covered areas of Greenland, Svalbard, and the Canadian high Arctic Archipelago during the Tertiary Period, ending approximately 3 million years ago (this signified a warmer Arctic but also an Arctic landmass that was a few degrees latitude further south). The megafauna, such as



(a)



(b)

Figure 1 (a) Contrasting concepts of vegetation zones of the Arctic according to Bliss LC and Matveyeva NV (1992) Circumpolar arctic vegetation. In: Chapin III FS, Jefferies RL, Reynolds JF, Shaver GR, Svoboda J, and Chu EW (eds.) *Arctic Ecosystems in a Changing Climate*, pp. 59–89. San Diego: Academic Press; Aleksandrova VD (1980) *The Arctic and Antarctic: Their Division in to Geobotanical Areas*. Cambridge, UK: Cambridge University Press, and Walker DA, Raynolds MK, Daniëls FJA, *et al.* (2005) The Circumpolar Arctic vegetation map. *Journal of Vegetation Science* 16: 267–282. (b) **Figure 1b** is a generalized version of the CAVM (Reproduced from CAVM Team (2003) Circumpolar Arctic Vegetation Map, scale 1:7,500,000. Conservation of Arctic Flora and Fauna (CAFF) Map No. 1. U.S. Fish and Wildlife Service, Anchorage, Alaska.).

Table 1 Various classifications of Arctic vegetation taken from Walker *et al.*, 2005. Note that the references within the table to each classification are included in Walker *et al.* (2005)

CAVM (Walker et al.*, 2005)*			*Russia*			*North America*					*Fennoscandia*	
Subzone	*Vertical structure of plant cover*	*Horizontal structure of plant cover*	*Alexandrova (1980)*	*Yurtsev (1994)*	*Matveyeva (1998)*	*Polunin (1951)*	*Edlund (1990) Elund and Alt (1990)*	*Bliss (1997)*	*Daniels* et al. *(2000)*	*Walker* et al. *(2002)*	*Tuhkanen (1986)*	*Elvebakk (1999)*
A	Mostly barren. In favorable microsites, 1 lichen or moss layer <2 cm tall, very scattered vascular plants hardly exceeding the moss layer.	<5% cover of vascular plants, up to 40% cover by mosses and lichens.	Northern polar desert Southern polar desert	High Arctic tundra	Polar desert	High Arctic	Herbaceous and cryptogam	High Arctic	Arctic herb	Cushion forb	Inner polar Outer polar	Arctic polar desert
B	2 layers, moss layer 1–3 cm thick and herbaceous layer, 5–10 cm tall, prostrate dwarf shrubs <5 cm tall	5–25% cover of vascular plants, up to 60% cover of cryptogams.	Northern Arctic tundra	Arctic tundra: northern variant	Arctic tundra	Middle Arctic	Herb prostate shrub transition Prostrate shrub		Northern arctic dwarf shrub	Prostrate dwarf shrub	Northern Arctic	Northern arctic tundra
C	2 layers, moss layer 3–5 cm thick and herbaceous layer 5–10 cm tall, prostrate and hemiprostrate dwarf shrubs <15 cm	5–50% cover of vascular plants, open patchy vegetation.	Middle arctic tundra Southern arctic tundra	Arctic tundra: southern variant	Typical tundra		Dwarf and prostrate shrub		Middle arctic dwarf shrub	Hemiprostrate dwarf shrub	Middle Arctic	Middle Arctic tundra
D	2 layers, moss layer 5–10 cm thick and herbaceous and dwarf-shrub layer 10–40 cm tall.	50–80% cover of vascular plants, interrupted closed vegetation.	Northern sub-Arctic tundra Middle sub-arctic tundra	Northern hypoarctic tundra		Low Arctic	Low erect shrub	Low Arctic	Southern Arctic dwarf shrub	Erect dwarf shrub	Southern Arctic	Southern Arctic tundra
E	2–3 layers, moss layer 5–10 cm thick, herbaceous/ dwarf shrub layer 20–50 cm tall, sometimes with low-shrub layer to 80 cm	80–100% cover of vascular plants, closed canopy.	Southern sub-Arctic tundra	Southern hypoarctic tundra	Southern tundra				Arctic shrub	Low shrub		Arctic shrub – tundra

Source: Reproduced from Walker DA, Raynolds MK, Daniëls FJA, *et al.* (2005) The Circumpolar Arctic vegetation map. *Journal of Vegetation Science* 16: 267–282.

the mammoth, of the Russian Arctic and extensive tundra steppe communities disappeared during the early Holocene.

Current vegetation patterns in the Arctic are to a large extent controlled by the glacial retreat during the Holocene (Raynolds and Walker, 2009), and much of the current Arctic floras immigrated into the Arctic during this period. For instance, on Svalbard, long-distance plant colonization has occurred repeatedly and from several source regions (Alsos *et al.*, 2007). However, some Arctic areas, "refugia" and "nunataks," remained ice free and supported biota for periods long enough to facilitate the establishment of endemic taxa. The eastern Russian Arctic was largely free of ice during the last ice age and is now associated with many endemic species (CAFF, 1999). Some species, such as *Pedicularis dasyantha* from Svalbard, are thought to have survived in a refugium and then spread outward (Odasz, as quoted by McGraw in Chapin and Körner, 1995).

During the recolonization of the Arctic, the continuity of landmasses, as well as the decreasing longitudinal distances toward the pole, allowed the movement of biota into, out of, and around the Arctic, for example, across the Bering land bridge. This connectivity allowed, and continues to allow, the mobility of many organisms such as marine mammals, migrating caribou, and particularly migratory birds. Such mobility adds considerable complexity to understanding the biodiversity in the Arctic because some of the causal factors are operative when migratory organisms are in other latitudes.

Despite this general connectivity, barriers to dispersal of species occur throughout the Arctic. In Alaska, the Brooks mountain range runs approximately east–west and separates the taiga forest zone to the south from the tundra zone in the north. In the Canadian Arctic, much of the Northwest Territories is an archipelago, and in Greenland, particularly along the east coast, glaciers flowing to the sea interrupt the continuity of ice-free coastal lands, whereas the Greenland ice sheet in general is an effective barrier covering the majority of the country. In Fennoscandia, the mainland is represented mainly by sub-Arctic areas in the mountains, and high Arctic environments are found only on the Svalbard archipelago. In the Russian Arctic, the Ural Mountains running north–south separate the vast landmass longitudinally. However, on the Taymyr Peninsula, the greatest latitudinal continuity of land reaches into the Arctic.

Patterns of Biodiversity in the Arctic

It is not possible to state the exact number of species in the Arctic because of a lack of taxonomic knowledge (particularly for the Protozoa, Nematoda, terrestrial Oligochaeta, many Acari taxa, and insect families in the orders Diptera, Hymenoptera, and Lepidoptera) and incompatible synonyms used in different countries. Estimating microbial diversity is even more problematic although these molecular techniques are improving the situation. However, some estimates of general species number can be made (Matveyeva and Chernov in Nutall and Callaghan, 2000; updated in Callaghan *et al.*, 2005, in ACIA, 2005; **Table 2**).

The number of species in the majority of plant and animal classes that are of importance in the higher Arctic environments amounts, on average, to approximately 2.5% of the worldwide species number for a given taxon (**Table 2**). The majority of the higher terrestrial faunal taxa are represented in the Arctic, although on a species number basis the Arctic contains less than 1% of the world's fauna. In comparison, the angiosperms, the most advanced plant taxon, have fewer species in the Arctic than the cryptogams (**Table 2**). Thus, the Arctic's biota is not just an extremely impoverished replication of the world's biota but also has a very distinctive structure as exemplified by the proportions of the highest taxa of plants and animals.

Classes that are well represented in, and are therefore particularly well adapted to, Arctic environments constitute approximately 3–11% or more of global diversity. These include birds, springtails, mosses, and lichens. The largest classes

Table 2 Biodiversity estimates for the Arctic compared with world biota[a]

Taxon								
Animals			*Plants*			*Fungi*		
Group	*Number*	*% of world biota*	*Group*	*Number*	*% of world biota*	*Group*	*Number*	*% of world biota*
Mammals	75	1.7	Angiosperms	2130	0.9	Fungi	2500	2.3
Birds	240	2.9	Monocotyledons	490	0.7			
Insects	3700	0.4	Dycotyledons	1640	0.9			
Diptera	1600	0.9	Gymnosperms	16	2.1			
Beetles	700	0.1 > 0.2	Pteridophytes	72	0.7			
			Bryophytes	900	3.7			
Springtails	400	6.0	Mosses	633	4.3			
Spiders	300	1.7	Liverworts	267	2.7			
Mites	700	1.9	Lichens	1750	9.6			
Other groups[b]	600	–	Algae	1200	3.3			
Total estimate	6000	–		6068	–			

[a]Modified from Matveyeva and Chernov in Nutall M and Callaghan TV (2000) *The Arctic: Environment, People, Policy*. Reading, UK: Harwood, and updated in Callaghan TV, Björn LO, Chernov Y, *et al.* (2005) Tundra and polar desert ecosystems. In *ACIA, Arctic Climate Impacts Assessment*, pp. 243–352. Cambridge, UK: Cambridge University Press, and with information kindly provided by Prof. FJA Daniels, University of Muenster, assuming estimates of relative proportions of Dycotyledons to Monocotyledons on one hand, and Mosses to Liverworts on the other, have remained unchanged since the 2005 publication.

[b]Amphibians (7), Centipedes (10), Mollusks (3), Annelids (70), and Nematodes (~500).

of global biota (e.g., insects and flowering plants) have the lowest level of representation in the Arctic at 0.4% and 0.7%, respectively. At lower taxonomic levels – orders, families, and genera – some taxa are endemic to the Arctic. For instance, all species of the order of the diving birds Gaviiformes are found in the Arctic.

Although the proportion of various taxa that are found in the Arctic may be very small, there is a common misunderstanding that this equates to low species numbers in Arctic communities (Chapin and Körner, 1995). This is not necessarily the case. For example, the wide ecological distribution of many Arctic plants, together with their small stature, results in high numbers of species per unit area in certain communities. Thus, on Taymyr, Matveyeva (Matveyeva and Chernov, in Nutall and Callaghan, 2000) recorded from 110 to 182 species per sample plot (100 m^2). There were 40–50 species within 1 m^2 and up to 25 species within 0.01 m^2. Even in the polar desert, approximately 50–60 species were recorded per 25 m^2 in both the Eurasian and the Canadian Arctic.

As with diversity at the community level, low absolute species numbers do not necessarily equate to low functional diversity within communities. Genetic heterogeneity of ecotypes is a reservoir of genotypic variation on which selection acts, fitness is adjusted, and adaptation becomes fixed (Crawford and Chapman in Callaghan *et al.*, 1995). Subspecific genetic variation occurs in all aspects of Arctic plants such as physiology, morphology, phenology, growth form, growth rate, and reproductive development (McGraw in Chapin and Körner, 1995). However, only some of these differences can be observed in the field (Murray in Chapin and Körner, 1995), and information on population differentiation based on isoenzyme or DNA analyses is rare compared with, for example, that of temperate latitudes. In addition, the longevity of Arctic plants (Callaghan and Emanuelsson, 1985) and the long periods required for many Arctic plant species to reach a mature growth form from seed imposes severe methodological constraints on partitioning variations in growth forms within a species between phenotypic plasticity and genetic differentiation. However, there are many examples of plant species with high levels of ecotypic differentiation, for example, *Oxyria digyna* (photosynthetic and respiration rates), *Dryas octopetala* and *Saxifraga oppositifolia* (respiration rates and growth form), *Eriophorum vaginatum* and *Carex aquatilis* (phenology), *Phleum alpinum* (growth rates), and *Hylocomium splendens* (growth form).

Possibly as a result of high genetic and ecotypic variability, in combination with low species diversity and a lack of species to occupy available ecological niches in the high Arctic, a phenomenon exists when one and the same species dominates in different communities. Matveyeva and Chernov (in Nutall and Callaghan, 2000) call this "superdominance." Examples are the lemming *Lemmus sibiricus*, widely distributed in practically all terrestrial habitats from the very wet depressions up to summits in mountains; the crane fly *Tipula carinitrons*; and the circumpolar plants *Arctophila fulva*, *Betula nana sensu lato*, *Carex stans*, *Cassiope tetragona*, *Dryas punctata*, *Dupontia fisheri*, *Eriophorum angustifolium*, *E. vaginatum*, *H. splendens sensu lato*, *Aulacomnium turgidum*, *Drepanocladus uncinatus*, and *Ptilidium ciliare*.

Despite the wide ecological amplitudes of many Arctic species, still strong geographical trends are associated with their distributions. At the species level, two basic types of distributions within the Arctic can be classified: polyzonal and zonal (**Table 3**). Polyzonal species have wide distributions in both the tundra and the taiga and possibly elsewhere. Zonal Arctic species, omitting topographical distinctions, can be subdivided into hyperarctic, euarctic, hemiarctic, and hypoarctic groups (**Table 3**; Chernov and Matveyeva in Nutall and Callaghan, 2000). These groups represent a distributional progression from the polar deserts to the taiga boundary described later and are described with examples in **Table 3**. Furthermore, they provide a useful basis for projecting the sensitivity of a species to climate change in that the polyzonal species are likely to be relatively resilient to change whereas the zonal Arctic hyperarctic species are likely to suffer reductions in appropriate habitats.

The Arctic contains extreme environments – polar deserts, snow, and ice – and these host specially adapted organisms called "extremophiles" (Bell, in press). Ice masses can be considered as ecosystem habitats with two key glacial ecosystems, one inhabiting the glacier surface (the supraglacial ecosystem) and the other at the ice-bed interface (the subglacial ecosystem) (Hodson *et al.*, 2008). The supraglacial ecosystem is characterized by microbes such as bacteria, algae, phytoflagellates, fungi, viruses and occasional rotifers, tardigrades, and diatoms within the snowpack, supraglacial streams, and melt pools (cryoconite holes). Perhaps the most noticeable are the snow algae, which can turn snow packs into a pink color. The subglacial system is dominated by aerobic/anaerobic bacteria and most probably viruses in basal ice/till mixtures and subglacial lakes (Hodson *et al.*, 2008). All these glacial ecosystems are sensitive to climate change. However, whereas rapid, melt-driven habitat changes affect microorganisms in many supraglacial ecosystems, much slower processes affect the subglacial ecosystem.

Even more extreme habitats are presented by permafrost. Large numbers (more than in the soil layers above) of viable ancient microbes are present within permafrost. They have been extracted from depths of 400 m in temperatures of $-17\,^{\circ}C$ (Gilichinsky, 2002). Spore-forming and spore-less bacteria, green algae and cyanobacteria, yeast, and mycelium fungi have been found. Evidence of extracellular enzymes and pigments suggests that many of these organisms are not just viable but also active even though the age of the isolates can be millions of years. These microorganisms are the only organisms that have remained viable over geological time and conceptually, and they could help to resolve fundamental scientific issues such as the periods for which organisms can remain viable on the earth or even on Mars. Climate change could result in the release of some organisms in the surface layers of thawing permafrost with unknown health consequences, whereas permafrost drilling could potentially release the oldest organisms.

Latitudinal Patterns of Diversity

Great changes in biodiversity within the Arctic often reflect sharp thermal gradients that have no analogies in other

Table 3 Diversity changes with latitude in the Arctic region

Category	*Distribution*	*Examples*		
		Plants	*Birds*	*Mammals and invertebrates*
Polyzonal	Taiga and far to the north in tundra but usually in local habitats and wet depressions	Soil algae; the mosses *Hylocomium splendens sensu lato, Aulacomnium turgidum*, and *Racomitrium lanuginosum*; the liverwort *Ptilidium ciliare*; the lichens *Cetraria islandica, Psora decipiens*, and *Cladina rangiferina*; the vascular species *Cardamine pratensis, Chrysosplenium alternifolium*, and *Eriophorum angustifolium*; the sedge *Carex duriuscula* the herb *Helictotrichon krylovii*; the moss *Tortula ruralis*	The common raven *Corvus corax*, the peregrine falcon *Falco peregrinus*, the white wagtail *Motacilla alba*, and the northern wheatear *Oenanthe oenanthe*	The wolf *Canis lupus,* the ermine *Mustella ermine*, the weasel *M. nivalis*, the vole *Microtus gregalis*, and the mite *Chiloxanthus pilosus*
Zonal boreal	Not abundant and constrained to the south of the Arctic in benign habitats such as river valleys, south-facing slopes, and wet areas	Tree species of *Larix*; the orchid *Corallorhiza*; the shrub *Salix myrtylloides*; the sedges *Carex chordorrhiza*; the herbs *Allium schoenoprasum, Cortusa matthiolii, Galium densiflorum, Sanguisorba officinalis*; and forest mosses *Climacium dendroides, Pleurozium shreberi*, and *Rhytidiadelphus triquetrus*	The forest birds *Turdus miscus* and *T. polaris* (thrushes); chiff-chaffes *Phylloscopus trochilis* and *P. collybita*; and river ducks *Anas acuta, A. penelope*, and *A. crecca*	Reindeer and the wolverine *Gulo gulo*
Zonal arctic hyperarctic	Polar desert and in the northern-most part of the tundra zone	Almost no plants are restricted to these zones: The following have their highest frequencies there– the grasses *Phippsia algida* and *Poa abbreviata*; the herbs *Cerastium regelii, Draba oblongata, D. subcapitata, Saxifraga hyperborea*, and *S. oppositifolia*; the mosses *Dicranoweisia crispula, Bryum tortifolium, Orthothecium chryseum* and *Seligeria polaris*; and the lichens *Cetraria delisei, C. elenkenii, Dactylina ramulosa, D. madreporiformis*, and *Thamnolia subuliformis*	The dovekie *Alle alle*, the ivory gull *Pagophila eburnea*, and snipe *Calidris alba* and *C. maritime*	Polar bear and the collembolan *Vertagopus brevicaudus*
Euarctic	Northern part of the tundra zone, rare in the southern part	The dwarf shrubs *Salix polaris* and *S. arctica* (this group is relatively small, but it has an important value in the subdivision of the tundra zone into subzones)	The black-bellied plover *Squatorola squatorola*, the curlew sandpiper *Calidris ferruginea*, the snowy owl *Nyctea scandiaca*, and the snow-bunting *Plectrophenax nivalis*	The lemming *Dicrostonyx torquatus* and the crane fly *Tipula carinifrons*

(*Continued*)

Table 3 Continued

Category	*Distribution*	*Examples*		
		Plants	*Birds*	*Mammals and invertebrates*
Hemiarctic	Throughout the tundra zone but most frequent in the middle	Most of the dominant plant species: the grasses *Arctophila fulva, Dupontia fisheri*; the sedges *Carex bigelowii/arctisbirica and Carex stans*; the shrub willow *Salix reptans*, the dwarf shrubs *Dryas punctata/octopetala* and *Cassiope tetragona*; the mosses *Tomenthypnum nitens, Drepanocladus intermedius*, and *Cinclidium arcticum*; the herbs *Lagotis minor*, and *Pedicularis hirsuta*, the moss *Polytrichum juniperinum*	The lapland longspur *Calcarius lapponicus*, the lesser golden plover *Pluvialis dominica*, and the dunlin *Calidris alpina* and *C. minuta*	The lemming *Lemmus sibiricus*, the bumblebees *Bombus hyperboreus* and *B. balteatus*, and the ground beetles *Amara alpina, Pterostichus costatus*, and *Syrphus tarsatus*
Hypoarctic	Optima in the southern tundra subzone	This group characterizes the southern tundra subzone; the shrubs *Betula nana/exilis* and sedge *Eriophorum vaginatum*	The ptarmigan *Lagopus lagopus*, the spotted redshank *Tringa erythropus*, the little bunting *Emberiza pusila*, and the bartailed godwit *Limosa lapponica*	The vole *Microtus middendorffi*, the ground beetle *Carabus truncaticolus*, the bumblebee *Bombus cingulatus*, and the spiders *Alopecosa hirtipes* and *Lycosa hirta*

Source: Compiled from information in Matveyeva and Chernov in Nutall M and Callaghan TV (2000) *The Arctic: Environment, People, Policy*. Reading, UK: Harwood.

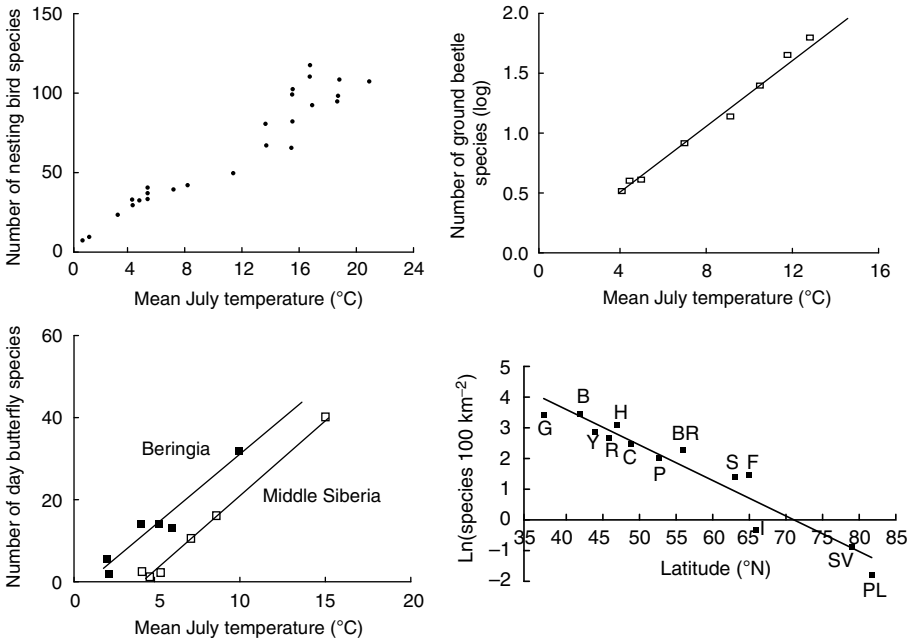


Figure 2 Relationships between biodiversity and latitude for various groups of animals and plants (reproduced from Matveyeva and Chernov, in Nutall M and Callaghan TV (2000) *The Arctic: Environment, People, Policy*. Reading, UK: Harwood). Note the logarithmic scales on the right-hand graphs. G, Greece; B, Bulgaria; Y, Yugoslavia; H, Hungary; R, Rumania; C, Czechoslovakia; P, Poland; Br, Britain; S, Sweden; F, Finland; I, Iceland; SV, Svalbard; PL, Peary Land, northeast Greenland.

biomes. In the Siberian sector of the tundra zone over a distance of 900 km, the mean July temperature decreases from 12 to 2 °C, whereas in the boreal forest belt, a comparable 10 °C change in mean July temperature occurs over almost 2000 km, a range across which there are three natural life zones. The gradient of summer temperature of 12 °C at the treeline of the Taymyr Peninsula to 2 °C in the polar desert is associated with a decrease in the number of vascular plants in local floras from approximately 250 in the south to approximately 50 in the north. Analogous decreases in plant diversity in Canada also exist (Rannie, 1986). A similar pattern occurs in the animal world, for instance, in day butterflies, beetles and birds (**Figure 2**), and spiders. At a larger scale, there is a general decrease in biodiversity with an increase in latitude from temperate to Arctic ecosystems (**Figure 2**).

In the plant kingdom, patterns of decrease in biological diversity toward the North Pole differ among taxa. Some diminish their significance in the biota or even disappear (Ericoids), others change proportionally to the level of general diversity (Fabaceae and Rosaceae), and a third type retain a high level of biological diversity in the Arctic where their proportion becomes higher (Saxifragaceae and Brassicaceae). Not only is there a decrease in plant species and family diversity with increasing latitude, but also the paucity of life forms within the plant kingdom of the Arctic results in a simpler vertical vegetation structure than that of the forested regions further south. In the Arctic, two- or three-layered vegetation, with the height of the tallest shrub layer up to 2 m, changes into the nearly two-dimensional plant cover of polar deserts where most of the biota are concentrated into a thin film of less than 5 cm above the ground and no more than 5 cm below the ground.

Because any species belongs to a particular life (or growth) form (**Figure 3**), the changes in species composition inevitably lead to changes in the proportion of life forms in a wide sense. The proportion of rooted species (flowering plants) that control their water content to merely adnate species (cryptogams) decreases from approximately 1:2 at the treeline to approximately 1:5 in the polar desert. Tall shrubs are common in the south of the tundra, but absent in the north. Dwarf shrubs are widely spread throughout the tundra zone, but absent in polar desert ecosystems. The proportion of long rhizomatous herbs decreases whereas that of tap-rooted plants increases in the northern direction. Loose caespitose species are replaced by dense tufted cushions or mat-forming species (**Figure 3**). Within a species, there is also a tendency to form more compact growth forms in the north compared to in the south.

Latitudinal patterns of biodiversity within the Arctic may reflect global trends. Passerines and other bird taxa that are important in the tropics (e.g., kingfishers, woodpeckers, parrots, hummingbirds, and pigeons) decrease successively as a proportion of the avifauna from the equator to the poles. Passerine birds comprise approximately 60% of the avifauna in temperate and tropical forests, approximately 50% up to the northern taiga landscapes, 40% in forest-tundra, and 20% in the northern part of the tundra zone. In contrast, the proportion of Charadriiformes (40%) and geese-like birds (20%) in the Arctic avifauna is high and decreases toward the south. Charadriiformes are reduced to 20% in the north taiga and 10% in the south, and they contribute only 5% in the tropics.

Circumpolar Diversity

The Arctic's biota have a relatively large number and proportion of circumpolar species compared with circumboreal

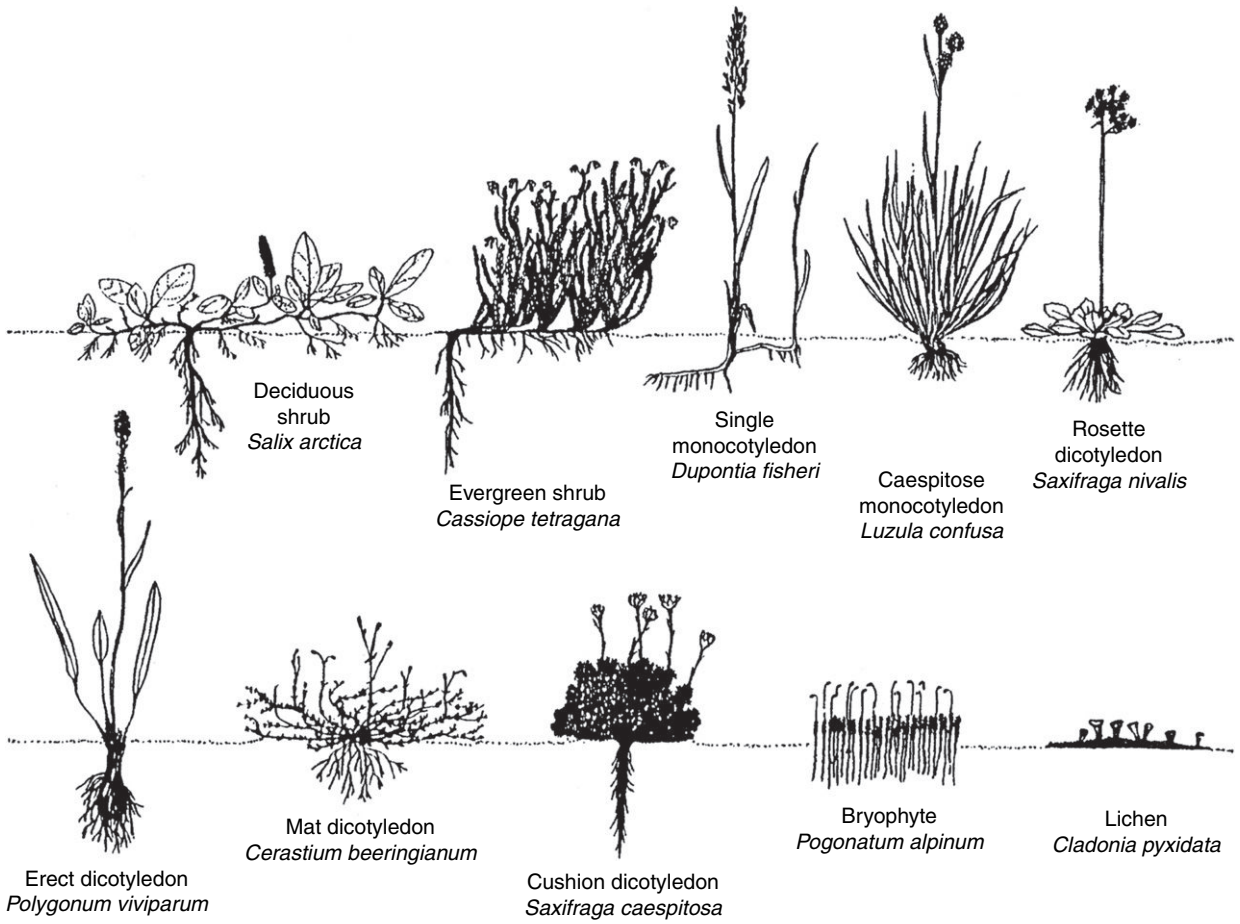


Figure 3 Growth forms of tundra plants, with example species. Plant functional type classifications basically follow the growth forms depicted and usually do not specify any particular function even though different classifications might be appropriate for individual functions. Reproduced from Webber PJ (1978) Spatial and temporal variation of the vegetation and its productivity. In: Tieszen LL (ed.) *Vegetation and Production Ecology of an Alaskan Arctic Tundra*, pp. 37–112. New York: Springer.

species. One of the obvious reasons for this is the decreasing circumference of latitudes toward the poles. Circumpolar vascular plants comprise approximately 45–50% of the extant flora in the south, 60–65% in the north of the tundra zone, and more than 70% in the polar desert.

Less than half of the Arctic birds may be considered circumpolar because many species formally considered as circumpolar have large gaps in their distribution. Thus, approximately 20% of the Arctic avifauna are true circumpolar species, and only 10% of the tundra terrestrial fauna are circumpolar species. Only 2 of 50 tundra *Tipulidae* species have true circumpolar distributions. Only 1 species is conditionally circumpolar among approximately 40 leaf-eater beetles (Chrysomelidae) known in the Arctic. In two very common plant genera, *Saxifraga* and *Draba*, each of which includes approximately 40 species in the Arctic, only 25% (10 in each) have true circumpolar distributions. The rich genus *Pedicularis* has 6 circumpolar species out of 24, but 5 of these have large gaps in Europe, Greenland, or eastern North America. In the *Ranunculus* genus, 9 of 25 species are circumpolar. Only 1 of the 10 species of *Dryas* in the Arctic, *D. punctata sensu stricto*, has a true circumpolar distribution.

Species that form monotypic genera have the most pronounced circumpolar distribution. Examples from the animal world are arctic fox (*Vulpes lagopus*), snowy owl (*Nyctea scandiaca*), old squaw duck (*Clangula hyemalis*), and calliphorid (*Boreellus atriceps*) and from the plant world the grasses *Arctophilia fulva*, *Dupontia fisheri*, and *Pleuropogon sabinei*; the herbs *Koenigia islandica* and *Oxyria digyna*. The two latter species have enormous geographic ranges, being circumpolar and with extensions far south: *K. islandica* is bipolar.

Most species have distributions narrower than circumpolar. In general, there are five main types of longitudinal distribution: Species can have (1) true circumpolar distribution or, if represented on both the North American and the Eurasian continent, they can have (2) amphi-Atlantic or (3) amphi-Pacific (Beringian) distributions or they can occur only on (4) Eurasian or on (5) North American continents. Within each continent, there are an additional five subregions, giving 10 zonations in total (Porsild, 1959): They can have a wide geographic range (Eurasian species or North American species) or can occur in a more restricted area (e.g., Siberian or Alaskan species).

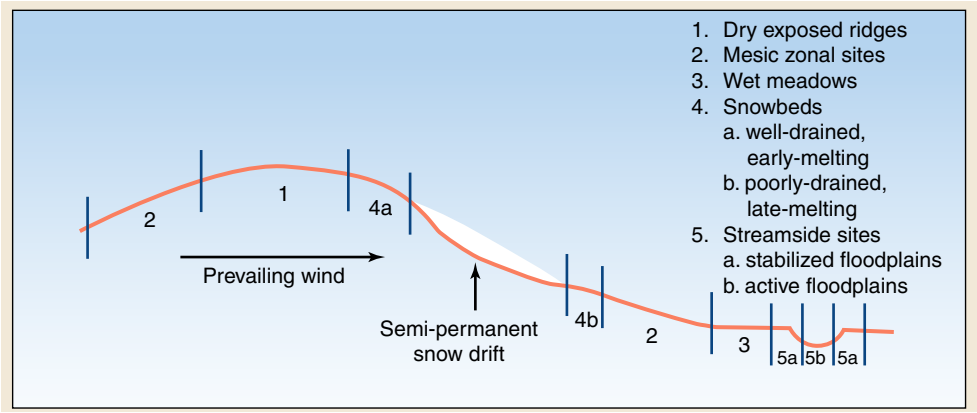


Figure 4 Local topography. Small-scale variations in topography determine the patterns of plant communities. This idealized mesotopographic gradient shows five microsites commonly found in Arctic landscapes (Walker *et al.*, 2002). Reproduced from Walker DA, Raynolds MK, Daniëls FJA, *et al.* (2005) The Circumpolar Arctic vegetation map. *Journal of Vegetation Science* 16: 267–282.

Effect of Topography on Biodiversity

The macrorelief represented by Arctic lowlands and mountain ranges is the main cause of interlandscape diversity and imposes altitudinal biodiversity patterns on longitudinal and latitudinal ones. The diversity of mesorelief forms is responsible for the diversity of the vegetation at the intralandscape level through their impacts on local temperature and snow accumulation. Throughout a range of tundra environments, there are recurring patterns in the distribution of diversity relative to topography (**Figure 4**). Using Russian terminology, zonal (i.e., mesic) tundra communities occupy the central environmental position within the landscape, namely, the "plakor" or watershed. These communities consist of continuous lichen–moss cover (pleurocarpous green mosses and fruticose lichens) with either a more or less sparse layer of shrubs, dwarf shrubs, sedges, and cotton grasses in the south or grasses and dicotyledonous herbs in the north. These communities have the richest species composition and are considered the true tundra communities for a given site. In addition, there is a rich diversity of intrazonal communities with fewer species and life forms developing along environmental gradients that deviate from zonal communities, including mires (without fruticose lichens), grasslands on south-facing slopes (with only few mosses), snowbed vegetation on the north-facing slopes, dwarf shrub heaths on fellfields (with few mosses), dense shrub thickets in deep valleys, and salt marshes.

Although different types of vegetation can be distinguished within a landscape, even greater diversity in vegetation pattern can occur over shorter distances due to nanorelief (e.g., patterned ground). A host of forms of patterned ground are found (Washburn, 1956). Within a given patterned ground feature, substrate stability can vary widely, influencing disturbance, soil nutrient availability, soil moisture, microclimate, and thickness of the active layer above permafrost. The relief of the patterned ground feature may influence the duration and height of the snow within a few square meters, for example, between hummocks and hollows. The extremely localized ability of Arctic plants to influence microenvironments leads to a great complexity of plant diversity, biomass, net primary production, and soil flora and fauna characteristics with small (>10 cm) differences in relief.

An additional resource increasing landscape diversity is the presence of diversity 'focal points' (Walker, in Chapin and Körner, 1995) or 'oases' (Edlund and Alt, 1989; Svoboda and Freedman, 1994). Moist areas below snowbeds in the Canadian polar deserts have a plant cover many orders of magnitude higher than that of the surrounding barrens and contain most of the local plant species diversity. Pingos, with their dome shape, steep slopes of diverse exposures, and gravel and sand substrate, are also 'hot spots of diversity' on the flat Alaskan coastal plain dominated by tussock tundra (Walker, in Chapin and Körner, 1995).

Localized topographical extremes can be associated with extreme species distribution outliers. Deep valleys with thick snow cover in winter and warm weather in summer at latitude 75° N on Taymyr have dense willow thickets of *Salix lanata* up to 2 m tall. Warm-water springs in Alaska, Chukotka, and middle-west Greenland create localized warm microclimates, and many stands in the surroundings of these hot springs support extreme range disjunctions of southern, even boreal, species. Examples include stands of *Populus balsamifera* in the northern foothills of the Brooks Range and the Chukotka Peninsula, well to the north of the treeline; ferns, orchids, and umbellifers near warm springs on Disko Island; and dwarf shrubs in some inner fjord areas of Svalbard. Steep south-facing river banks combine warm and well-drained habitats with the possibility of migration along the river corridor and are another focal point of community diversity, often supporting trees within the tundra. Animals also create diversity hot spots, when they are dead, by providing nutrients and moisture in polar deserts and fellfields and, when they are alive, by bringing back nutrients to their nesting places or dens. Examples are bird cliff communities, fox and wolf dens, and lemming gardens. Such outliers represent potential "inocula" for recolonization (*see* Threats to Biodiversity in the Arctic) and plant community change in a future warmer climate and support animal species with more southerly

distributions. This process may have occurred approximately 12,000 or 13,000 years ago in Beringea, when birch rapidly spread and created a 'birch zone.'

Controls of Biodiversity

A series of filters has selected, and is selecting, species and genotypes that occur in the Arctic (Walker, in Chapin and Körner, 1995). The first filter is the presence of a species in the region, the second set of filters is the biogeography of species within the Arctic, and the third set of filters is internal filters within communities and the environment. Körner (in Chapin and Körner, 1995) adds a time dimension to the geographical filters. For example, the current range of species and communities in the Arctic has been determined by survivors from the thinning of past populations and floras during glacial periods together with the species that immigrated into the Arctic during the Holocene. In his Alpine examples, Körner describes a second filter, grazing and freezing temperatures, and a third filter that acts at the microscale and specifically selects plant species that can tolerate certain microclimates, soil disturbance levels, and moisture regimes.

In the Arctic, during the Pleistocene, it could be argued that the first filter was environmental in that only freezing-tolerant plant species could survive in a region that became 'Arctic' in character. Later, during the Holocene, the next set of filters could be seen as relating to migration and dispersal abilities in that many species followed the retreating ice margins northward, whereas some glacial relicts spread outward from their refugia. It is possible that this filter could at least partially explain the high abundance of cryptogams, which have easily dispersed spores, relative to vascular plants in the Arctic. In general, however, it can be assumed that the main set of filters on current Arctic biodiversity is environmental rather than constraints on migration and the 'available species pool' (see, e.g., Alsos *et al.*, 2007). This assumption is also supported because, for example, latitudinal treelines are often situated more to the south now than during the earlier Holocene (although trees never reached the high Arctic, even during the climatic optimum of the Holocene). Also, herbivore diversity and population sizes are limited by low primary production resulting primarily from low nutrient availability. Any local increase in nutrient availability (e.g., through animal activity) dramatically stimulates biodiversity and productivity.

The next set of filters, as suggested by Körner, act at the microscale and consist of numerous interactions among plants in communities, between plants and herbivores, and between all organisms and the microenvironment. These fine-scale filters select for organisms that are preadapted to Arctic environments or that have developed specific adaptations. Among the animals, preadaptations can be seen in traits common in boreal forest animals. In plants, preadaptations are particularly evident in those with Arctic–Alpine distributions (i.e., with populations in Alpine areas to the south of the Arctic) and in those from nutrient-poor bogs of the temperate region. Specific adaptations of plants to the Arctic environment have been considered to be few. A constraint on adaptation rate has been the generally young nature of the Arctic flora, the longevity of many plants, and the sporadic successful completion of reproduction and seedling recruitment in areas of closed vegetation (Callaghan and Emanuelsson, 1985).

The fine-scale filters on biodiversity arising from biotic interactions were mentioned earlier. Plant competition is not thought to be a major force displacing species in the Arctic. Indeed, species removal experiments often result in little compensatory growth of remaining species, though the shrub species dwarf birch *B. nana* seems capable of displacing other shrubs in a warmer Arctic (Bret-Harte *et al.*, 2001). Instead, facilitation and succession are important and might explain to some extent the similarity of plant aggregations throughout large areas of the Arctic. If there were any competitive 'tensions' within these aggregations, they would be expected to have broken down over time or over geographical ranges.

Herbivory is only sometimes a filter on biodiversity. Perhaps in general, herbivores do not eat particular plant species to extinction and, indeed, can stimulate plant diversity. Also, in some cases, herbivores may act as buffers against other, for example, climate-induced, changes of the vegetation (Post and Pedersen, 2008; Olofsson *et al.*, 2009). In other cases, such as grazing by snow geese in the Hudson Bay area, high population pressure can lead to ecological cascades ending with loss of plant populations and ultimately soil erosion. Similar dramatic impacts on vegetation can occur when population peaks of two different herbivores, both feeding on the same plant, occur within a short period of time. Analogs of tundra vegetation in Fennoscandia can be produced when sub-Arctic birch forests are defoliated by insects and regenerating trees are browsed by reindeer. Also, browsers such as the megafauna of the Pleistocene and early Holocene had the ability to maintain open vegetation (Zhimov *et al.*, in Chapin and Körner, 1995), whereas grazers can control dominance in plant communities, for example, by suppressing moss growth. Although such impacts can be dramatic at the local scale, the generally wide distributions of many Arctic species ensure survival of the species. Survival of rare species and ecotypes, however, has a greater risk.

Consequences of Biodiversity

Environmental Consequences of Biodiversity

One way in which the importance of biodiversity within processes can be readily considered is by understanding the potential impact of biodiversity decline. Primary successional processes are indicative of the potential impact of biodiversity loss. One of the most thoroughly studied successional series within the Arctic is that present at Glacier Bay, Alaska (Chapin *et al.*, 1994). The retreat of the glaciers during the past 200 years has led to the development of a series of communities whose spatial arrangement can be considered analogous to a sequence through time. At the foot of retreating glaciers is a harsh environment of exposed glacial till that lacks soil organic matter and soil nutrients, and is a potential zone of drought. Mats of nitrogen-fixing cyanobacteria, lichens, and liverworts, with scattered forbs (e.g., the nitrogen-fixing *Dryas drummondii*), colonize these environments. The process of pedogenesis, involving the accumulation of organic matter

and plant-available soil nutrients, then proceeds and the succession sequence passes through many stages from alder woodland (*Alnus sinuata*) to sitka spruce (*Picea sitchensis*) forest. The seedlings of key species from each successional stage at Glacier Bay were found to germinate more readily within the preceding successional stage than within that of their own (Chapin *et al.*, 1994). Every step of the successional process enhances the development of species dominant within the next stage. Given the loss of any individual step within this series, the composition of the community may develop along an entirely different trajectory, or the environment may never develop beyond a given point.

The nature of Arctic environments is such that events leading to the initiation of primary succession, for example, ice-based processes such as frost heave, avalanche, and debris flows and high runoff due to permafrost-impeded drainage, are common. For any site undergoing primary succession, the availability of propagules of primary colonizers within that area is essential. An underlying level of diversity is necessary throughout the environment for appropriate colonizer species to be within the range of disturbance events. For example, in the Alaskan tundra, gravel pads and borrow pits are often initially colonized by species from nearby riparian systems. Without habitat diversity at the landscape level, these colonizing species would be absent. Therefore, both the distribution and the absolute number of organisms are important for ecosystem function. Organisms may maintain the habitat diversity of Arctic systems. The activity of burrowing animals such as ground squirrels, lemmings, or Arctic foxes may provide exposed areas of disturbed soil that act as potential refugia for plant primary colonizers and faunal diversity, thus promoting floral diversity. Another example is animals acting as seed dispersers, thereby aiding not only dispersion but also biodiversity in general (Bruun *et al.*, 2008).

As succession proceeds, the potential for interaction between organisms increases because their zones of influence are more likely to overlap with increasing organism density (this is especially true of plants). Because of an increase in the number of species, the potential complexity of interactions within a community also increases. Therefore, there may be a parallel increase in the level of biodiversity and the degree and complexity of interactions between individuals. This type of pattern might be found along the latitudinal gradient from high Arctic polar semidesert toward more productive tundra environments.

Types of interaction include direct resource-based interactions (e.g., competition) and indirect interactions, possibly involving higher trophic levels. For example, there is evidence that herbivores may encourage the development of vascular plants by preventing the formation of dense moss mats in certain tundra environments (Zhimov *et al.*, in Chapin and Körner, 1995; Van der Wal and Brooker, 2004). Because of the mobile nature of larger herbivores (e.g., reindeer and musk ox), their importance in the functioning of a community at any given time may not be clear. As in the case of plant primary colonizers, the key role that they play, and their importance in terms of maintaining diversity at both the species and the community level, may be apparent only at particular points in time or space.

Facilitative interactions may be of particular importance within Arctic communities. In the case of plant interactions,

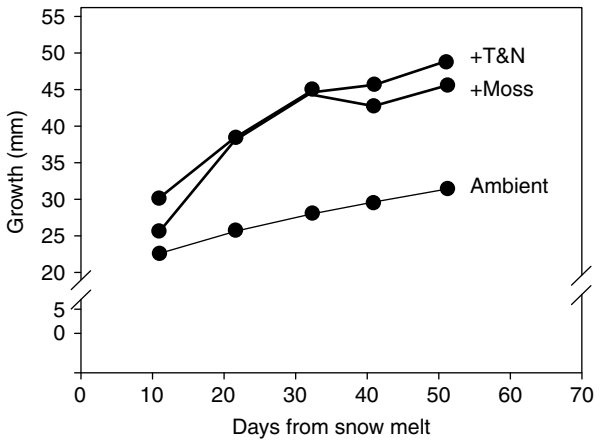


Figure 5 Impact of mosses in determining the growth of vascular plants in the high Arctic (an example is the grass *Poa arctica* growing in cushions of the moss *Drepanocladus uncinatus*). The impact of the moss on the growth of the grass, when compared with the grass growing outside moss cushions, is equivalent to growing the grass under plastic greenhouses with nutrient addition (+T and N) (reproduced from Press MC, Callaghan TV, and Lee JA (1998) How will European arctic ecosystems respond to projected global environmental change? *Ambio* 27: 306–311).

facilitation may act through the amelioration of environmental severity and may occur within all successional stages, including mature communities. Facilitative plant species may be keystone species within an environment; for example, in areas of frost heave-disturbed scree, the growth of spreading plant species with persistent nets of rhizomes or roots has been shown to stabilize the scree and to allow the development of other plant species within the stabilized zone. A critical factor in this case is the morphology of the colonizing plant. Other types of plant growth form may have different facilitative effects. Although all plants may provide shelter to their neighbors, certain species and growth forms, especially cushion plants, can develop in the absence of this initial shelter and can raise temperature differentials between their meristems and the air by more than 25 °C. They subsequently provide shelter for others. In one Alpine example, five species of cushion plant, including *Silene acaulis*, together 'hosted' 93 other species of plants. In another example, grasses growing in moss cushions in the high Arctic grew dramatically faster than those growing outside moss cushions (**Figure 5**). Diversity of form (**Figure 3**) is therefore perhaps one aspect of Arctic biodiversity that may be critical for ecosystem function.

Differences in plant growth form also strongly affect the movement of resources within an environment. Spreading rhizomatous species that maintain physiological connections between individuals may move nutrient resources away from localized resource-rich areas. Species with spreading root systems but a localized canopy may in turn concentrate resources within a given area. Similarly, differences in the behavior of animals may also lead to variability in resource distribution and resource movement throughout the environment. During the winter period when their grazing is confined to roots, rhizomes, and the stem bases, lemmings can accumulate within their nesting areas small reservoirs of nutrients in plant tissue and droppings. Other grazing species such as geese

and muskoxen may redistribute nutrients throughout the environment.

Not only are nutrient distribution processes dependent on the variability in life form, and hence diversity, but also nutrient cycling and soil conditions (e.g., temperature, moisture, and pH) may be strongly related to the diversity of organisms within an ecosystem. Hobbie (in Chapin and Körner, 1995) describes how differences in plant traits may directly and indirectly influence biogeochemical processes within tundra plant communities. For example, species with high productivity rates (generally deciduous and graminoid species) may increase the rate of nutrient turnover, woody species may increase the sequestration of carbon, and graminoid species may attract herbivores and increase the degree of herbivore nutrient cycling. Interspecific variability in rooting depth and leaf area influences soil moisture conditions, which in turn may influence soil nutrient availability, soil biota diversity, ion uptake, and soil pH. Leaf litter tissue quality differences may also influence soil microbial processes and decomposition rates. Mixtures of litter from more than one Arctic plant growth form, particularly hemiparasitic herbs and dwarf shrubs, have been shown to decompose at a faster rate than single-species litters (Quested *et al.*, 2002). The diversity of a plant community will therefore have important consequences for soil environmental conditions and terrestrial nutrient cycling, which is a major constraint on productivity of Arctic ecosystems.

Species diversity may also affect atmospheric environmental conditions by influencing soil–atmosphere exchange processes. Verville *et al.* (1998) found that sedges contribute to methane emission from Arctic wet meadow communities by acting as a conduit for methane release from anaerobic soils (rather than by producing methane directly). In contrast, mosses tend to limit the evolution of methane (**Figure 6**).

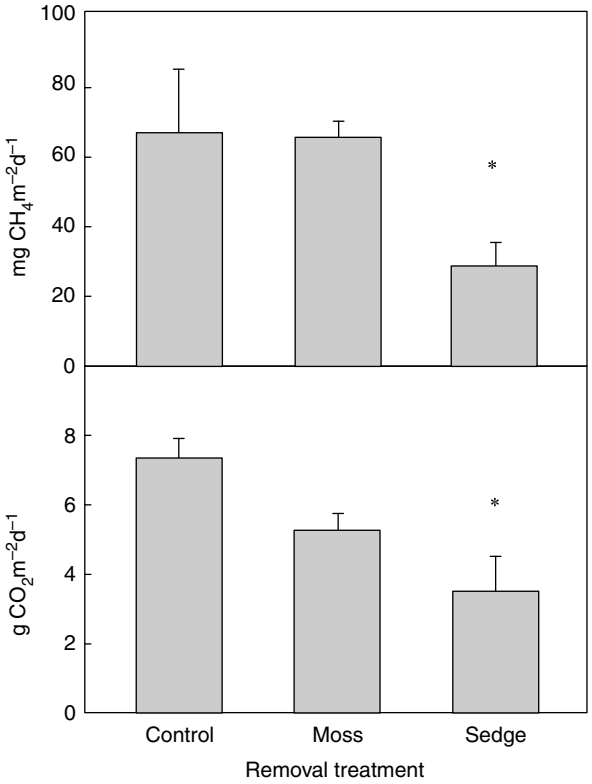


Figure 6 Methane and carbon dioxide fluxes in short-term species removal plots in wet meadow tundra. Sedges acted as a conduit for methane emission and carbon dioxide uptake: Mosses did not. *Statistically significant differences among species-removal treatments ($p=0.05$). Reproduced from Verville JH, Hobbie SE, Chapin III FS, and Hooper DU (1998) Response of tundra CH_4 and CO_2 flux to manipulation of temperature and vegetation. *Biogeochemistry* 43(3): 215–235.

Natural Resources and Environmental Exploitation

An analysis of ecosystem services from the Arctic was presented by Chapin *et al.*, 2005 during the Millennium Assessment of Ecosystems. The Arctic's ecosystems are important for both their provisioning services (e.g., food) and their regulatory services (e.g., drawdown and storage of atmospheric carbon). At the most simplistic level, biodiversity can be considered a natural resource. One key future of Arctic biodiversity as a resource is its unique components, that is, species (and processes) that are confined to Arctic ecosystems. However, existence per se is not commonly considered to be a resource, and the tendency is to regard resources as items for exploitation by humans, either currently or in the future.

Certain types of human exploitation of Arctic natural resources are limited by biodiversity and by productivity, namely, those in which an organic crop is removed from the environment. Biodiversity determines the number of resources available, and productivity limits the rate at which they can be exploited. For humans to live within the Arctic using extant natural resources, they must travel over a wide area to obtain and concentrate sufficient energy, nutrients, and raw materials. The Saami people of northern Fennoscandia follow their reindeer because productivity in general is very low and it is necessary for the reindeer herds to move over a vast area in order to obtain sufficient resources. This is true both for the semidomesticated reindeer herds of the Saami and for the wild caribou herds of North America. One impact of low biodiversity in this context may be the necessity of exploiting one species intensively, such as the Saami exploitation of reindeer, rather than being able to switch exploitation between species on a seasonal basis.

Nonindigenous human exploitation of wildlife has involved hunting to accumulate raw material for manufacturing processes (e.g., furs and blubber). The replacement of many animal-derived products with synthetic materials derived from oil has reduced the pressure on Arctic animal populations. However, hunting of Arctic wildlife still occurs, either by indigenous peoples, such as the Saami or Inuit, or for sport. This latter type of hunting is just one particular branch of the development of tourism within the Arctic. Ecotourism depends heavily on the wildlife of the Arctic. The biodiversity of the Arctic, including unique Arctic species, and the paucity of species and barren nature of high-latitude Arctic systems in particular (i.e., their inherently low biodiversity) may be features of Arctic environments that are attractive to visitors. In this case, biodiversity can be an exploitable resource, and as ecotourism develops, perhaps the monetary value attached to the biodiversity of the Arctic will increase.

A more intangible natural resource dependent on Arctic biodiversity is the potential for future acquisition of knowledge. The low biodiversity and process simplicity within Arctic ecosystems makes them a valuable resource for testing ecological theories. Arctic ecosystems may provide a second type of scientific resource, that is, they might contain undiscovered or unrecognized useful natural substances. Although the low biodiversity of Arctic ecosystems suggests that the number of possibly useful substances is less than that in diverse ecosystems (e.g., rainforest), unique substances may be found that have evolved in response to the Arctic environment. For example, it may be that the Arctic provides humans with a source of natural antifreeze chemicals. Similarly, the mechanism whereby some Arctic species can tolerate extreme anoxia might be important in understanding the problems associated with the susceptibility of some trees and crops to waterlogging.

Mineral resources are another natural product extracted from Arctic environments. However, in this case the problem of exploitation is not one of low biodiversity but rather of exploiting the resources without damage to biodiversity, as discussed in the following section.

Threats to Biodiversity in the Arctic

The Arctic is currently undergoing rapid changes: cultural and sociological changes resulting in exploitation of natural and nonrenewable resources, changes in climate, and changes in pollution – both local and transboundary (Nutall and Callaghan, 2000; Callaghan *et al.*, 2005; Post *et al.*, 2009; Callaghan and Tweedie, 2011). Many of these changes are predicted to accelerate in the future, and these have implications for changes in biodiversity within the Arctic.

Extractive industries (oil, nickel, etc.) have resulted in large local impacts on species, including reductions in the diversity of soil microorganisms (Evdokimova, in Callaghan *et al.*, 1995). Examples of impacts of mineral extraction are found in towns of the Kola Peninsula, but there has been surprising resilience of the vegetation in that pollution-tolerant forms are being actively selected for (M. Kozlov, personal communication). Oil spills in Alaska and Russia have had dramatic effects on local populations of both plants and animals. However, the large and widespread distributions of Arctic biota provide a system in which changes in biodiversity from this source over wide areas are unlikely. More serious is the slow recovery potential of Arctic vegetation that exacerbates disturbances to ecosystems.

Harvesting of natural resources on land in the Arctic consists mainly of hunting, collecting (e.g., berries), and reindeer husbandry. Effects of hunting have probably had serious impacts on a small proportion of Arctic biota over the Holocene, and it is widely debated whether northern megafauna, such as the mammoth, were hunted to extinction at the beginning of the Holocene. Within the past 300 years, hunting has probably hastened the demise of the great auk, and following the advent of firearms in the north, wolf populations have been reduced to just a few in the sub-Arctic of Fennoscandia. Currently, international agreements protect some species such as polar bear and some subpopulations have increased. However, the health of many species of birds that migrate to the Arctic is determined more by hunting practices in the overwintering grounds than in the Arctic: This is more difficult to legislate.

Changes in reindeer husbandry practices in Fennoscandia during the past 300–400 years from hunting to herding and, recently, from herding to almost farming in some areas have had impacts on predators (hunted by husbanders) over large areas and on grazing pasture productivity and biodiversity. Restrictions on the nomadic lifestyles of reindeer husbanders both in Fennoscandia (due, for example, to constraints of national boundaries) and in areas of the Russian Arctic (due to imposition of a sedentary lifestyle for women and children) have resulted in large-scale impacts on vegetation with an often dramatic reduction in lichen cover. Changes in lichen biodiversity remain to be documented but are likely to be pronounced.

Pollution in the Arctic is still generally less than in more southerly regions, and the pathways of pollutants into the Arctic and their concentrations within Arctic biota have been well documented recently (Reiersen, in Nutuall and Callaghan, 2000). However, little is known about the sensitivity of Arctic biota to the wide range of pollutants that are found there.

During the 1.8 million years of the Pleistocene, the Arctic has undergone severe climatic fluctuations associated with glacial and interglacial periods. Even within the last interglacial period of the Holocene, climate has changed greatly, resulting in a current cool period associated with lower and more southerly treelines in many places in the Arctic than found in the early Holocene. Since the Little Ice Age, however, which ended in the north Atlantic region approximately 150 years ago, mean annual temperatures have been increasing. Arctic climate has now entered a unique period relative to the instrumental record and, in the case of summer temperatures, relative to 2000-year reconstructions of past variations: This trend is expected to continue (Walsh *et al.*, 2011). Analysis of *in situ* and satellite data shows evidence of different regional snow cover responses to this warming and increasing winter precipitation over the past 40–50 years. In the past decades, sea ice has been diminishing, which adds to the warming of adjacent terrestrial environments (Bhatt *et al.*, 2010). The warming has been followed by an increase in the 'greenness index' of vegetation especially in the low Arctic, as seen by satellite images, which suggests an increase in the growing season of approximately 11 days over the past 20 years. The recent increases in temperature in the Arctic have been associated with increased pest outbreaks in the taiga, increased frequency of forest fires, and a warming of permafrost with an increase in disturbance events (ACIA, 2005; AMAP, 2011).

General circulation models that predict future climates for various concentrations of greenhouse gases all agree that future warming will be greatest in the Arctic, and particularly during wintertime. Because of the strong correlations between biodiversity and temperature discussed previously (*see* Latitudinal Patterns of Diversity), it is expected that future warming will have a large impact on biodiversity in the Arctic. Experiments in a range of Arctic habitats that simulate warmer conditions show that one of the first effects of warming on plant communities is a change in the dominance of vascular plant species existing when the manipulations started and an

increase in biomass (Chapin *et al.*, 1995; Press *et al.*, 1998; Arft *et al.*, 1999). However, as the vascular plants increase in biomass, lichen biomass and cover decrease, and some moss species (e.g., *Hylocomum splendens*) also decrease in abundance. In the case of lichens, the decreases seen in the experiments parallel decreases found along natural geographical gradients of temperature. The major implications of the experiments are therefore that, over the medium term, warming will change the dominance among vascular plant species, the abundance of lichens and mosses will decrease, and the immigration of more southerly species will happen. Over the longer term, clearly, biodiversity of vascular plants will increase but that of cryptograms will decrease. In addition to this decrease in lichens is the impact of changing reindeer husbandry discussed previously (*see* Natural Resources and Environmental Exploitation). Overall, the Arctic's important role as a reserve of primitive plants is threatened, and the great cover of lichens in the Arctic will have an effect on their biodiversity at a global level.

In contrast to predicted decreases in biodiversity of the mosses and lichens, warming may increase diversity of all plant groups in some places. Glacier forefield expansion as a result of glacier retreat during warming and the open ground of polar deserts and semideserts offer open niches for the establishment of new plants. Because colonization of these habitats is usually by seeds and spores, it can expand the local species and subspecific genetic diversity. This is in marked contrast to closed areas of vegetation in the mid- and sub-Arctic, where plant establishment is mainly via clonal growth. However, even here, increases in disturbance due to thawing permafrost are likely to open up niches and allow new genets to establish.

Current Status of Arctic Biodiversity

The Conservation of Arctic Flora and Fauna working group of the Arctic Council has initiated a circumpolar assessment of the Arctic biodiversity, and some important findings have already been presented in the overview report *Arctic Biodiversity Trends 2010: Selected Indicators of Change* (Conservation of Arctic Flora and Fauna, 2010). The assessment reports that the unique Arctic habitats are changing rapidly and that this will have severe implications for the entire Arctic. Many populations of the selected Arctic species examined are seemingly stable or increasing, including some of the more abundant seabirds, such as common eiders (*Somateria mollissima*). However, other species central to the Arctic are declining, including reindeer (caribou) and lemmings. Also, some important geographical patterns emerged from the assessment, and whereas the high Arctic vertebrates are declining markedly, vertebrates living in the low Arctic and sub-Arctic have relatively stable or increasing populations. Similar differences were observed when comparing the terrestrial, limnic, and marine biomes, and whereas vertebrate species in the terrestrial biome generally declined, vertebrates in the limnic and marine biomes increased. The Arctic Biodiversity Assessment has identified multiple external stressors impacting the Arctic's diversity, including human exploitation and contamination, but emphasises that the changing climate is emerging as the single most important stressor of Arctic biodiversity.

Another recent assessment of changing Arctic ecosystems and biodiversity is the International Polar Year Project "Back to the Future" in which multidecadal changes were recorded by site revisits and data retrieval (Callaghan and Tweedie, 2011). General patterns were hard to identify: Biodiversity in some areas had remained surprisingly stable, whereas in others, dramatic increases in cover and immigration of a thermophilic species were seen (Callaghan *et al.*, 2011, in Callaghan and Tweedie, 2011). The greatest responses were usually in the woody species, in line with the "shrubification" of the tundra (Sturm *et al.*, 2001; Tape *et al.*, 2006). Although the sample size was low, a drying pattern was seen in which, for example, ponds had diminished in size. This is characteristic of many areas of the Arctic, while waterlogging occurs in others. Although the net balance of these two processes is critically important in determining future biodiversity, ecosystem services, and biogeochemical cycling, the likely balance is not yet known. An important conclusion of the study is that local factors can override the general climatic controls on diversity, range extension, cover, and productivity of plants. For example, Van Bogaert *et al.* (2011) demonstrated that treeline could relocate upward and downward and remain stable all within the same area experiencing the same climatic changes. In this case, the local controls were release from reindeer herding pressure leading to an upward movement of the treeline, insect pest outbreaks that lowered the treeline, and steep rocky slopes that resulted in a stable treeline. Though the Arctic Biodiversity Assessment and the Back to the Future Projects have made substantial progress in evaluating changing biodiversity of the Arctic, the thorough assessment of the Arctic's biodiversity is hampered by the lack of sufficient data, even for charismatic species such as the polar bear. The Arctic Council initiative Circumpolar Biodiversity Monitoring Program (CBMP; http://cbmp.arcticportal.org/) is therefore a cornerstone in the effort to conserve Arctic biodiversity through the development of long-term programs for monitoring the biodiversity in the Arctic.

Current Relevant International Organizations and Activities

For major recent assessments of biodiversity, changes in biodiversity, and research and monitoring activities based on Arctic Terrestrial biodiversity, follow the following links:

Arctic Climate Impact Assessment (ACIA) www.acia.uaf.edu
Conservation of Arctic Flora and Fauna (CAFF) www.caff.is
Circumpolar Arctic Biodiversity Monitoring Programme (CBMP) http://caff.is/monitoring
The International Arctic Science Committee (IASC) http://iasc.arcticportal.org/
International Network for Terrestrial Research and Monitoring in the Arctic (INTERACT) www.eu-interact.org/
Intergovernmental Panel on Climate Change (Polar and Ecosystems Chapters) (IPCC) www.ipcc.ch
International Study of Arctic Change (ISAC) http://www.arcticchange.org/
Snow, Water, Ice and Permafrost in the Arctic (SWIPA) www.amap.no/swipa

World Wildlife Fund (Arctic Office) (WWF) http://wwf.panda.org/what_we_do/where_we_work/arctic/

Conclusion

Although biodiversity in the Arctic is low compared with that in other regions, a surprisingly varied array of landscapes, ecosystems, species, ecotypes, and topographic micropatterns exist. Relationships between diversity and environment are superficially simple because land-use impacts are relatively few in the Arctic. However, beneath the surface, interactions between the snow patterns and low temperatures of the Arctic and its biota are complex: Temperature and snow have many direct effects on the biota but also control nutrient cycling in cold soils, which severely constrains the abundance of vascular plant species and the organisms that depend on them. Importantly, interactions between plants and climate are bidirectional: Plant growth forms facilitate dramatic differentials to occur between plant and air temperatures, whereas tundra soils have sequestered the greenhouse gas carbon dioxide throughout much of the Arctic. Currently, rapid sociological and environmental changes are occurring in the Arctic. Although low biodiversity is compensated for to some extent by large and widespread populations, the vast wilderness areas are nevertheless under numerous threats. Our major challenge is to further document (Conservation of Arctic Flora and Fauna, 2001, 2010) and monitor the biota in the remote Arctic wildernesses and to understand both how biodiversity patterns are formed and how they contribute to the functioning of ecosystems. Only then can the conservation strategies be established that can cope with rapid environmental change. The Arctic Biodiversity Assessment and the CBMP are important steps toward a common strategy for the conservation of Arctic biodiversity.

See also: Antarctic Ecosystems. Boreal Forest Ecosystems. Latitudinal Gradients of Biodiversity

References

ACIA (2005) *Arctic Climate Impact Assessment – Scientific Report*, 1st edn. Cambridge, UK: Cambridge University Press.

Aleksandrova VD (1980) *The Arctic and Antarctic: Their Division into Geobotanical Areas*. Cambridge, UK: Cambridge University Press.

Alsos IG, Eidesen PB, Ehrich D, *et al.* (2007) Frequent long-distance plant colonization in the changing Arctic. *Science* 316: 1606–1609.

AMAP (2011) *Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2011*. Oslo: Arctic Monitoring and Assessment Programme (AMAP).

Arft AM, *et al.* (1999) Responses of tundra plants to experimental warming: Meta-analysis of the international tundra experiment. *Ecological Monographs* 69: 491–511.

Bell EM (ed.) (2012) *Life at Extremes: Environments, Organisms and Strategies for Survival*. Wallingford, UK: CABI.

Bret-Harte MS, Shaver GR, Zoerner JP, *et al.* (2001) Developmental plasticity allows Betula nana to dominate tundra subjected to an altered environment. *Ecology* 82: 18–32.

Bruun HH, Lundgren R, and Philip M (2008) Enhancement of local species richness in tundra by seed dispersal through guts of muskox and barnacle goose. *Oecologia* 155: 101–110.

Bhatt US, Walker DA, Raynolds MK, *et al.* (2010) Circumpolar arctic tundra vegetation change is linked to sea ice decline. *Earth Interactions* 14: 1–20.

Bliss LC and Matveyeva NV (1992) Circumpolar arctic vegetation. In: Chapin III FS, Jefferies RL, Reynolds JF, Shaver GR, Svoboda J, and Chu EW (eds.) *Arctic Ecosystems in a Changing Climate*, pp. 59–89. San Diego: Academic Press.

Callaghan TV, Björn LO, Chernov Y, *et al.* (2005) Tundra and polar desert ecosystems. In *ACIA, Arctic Climate Impacts Assessment*, pp. 243–352. Cambridge, UK: Cambridge University Press.

Callaghan TV and Emanuelsson U (1985) Population structure and processes of tundra plants and vegetation. In: White J (ed.) *The Population Structure of Vegetation*, pp. 399–439. Dordrecht: Junk.

Callaghan TV, Oechel WC, Gilmanov T, *et al.* (eds.) (1995) Global Change and Arctic Terrestrial Ecosystems. *Proceedings of Papers Contributed to the International Conference*, Oppdal, Norway, August 21–26. European Commission Ecosystems Research Report No. 10. European Commission Luxembourg.

Callaghan TV and Tweedie CE (eds.) (2011) Multi-decadal changes in tundra environments and ecosystems: The international polar year back to the future project. *Ambio* 40(6): 555–716.

Chapin III FS and Körner Ch (eds.) (1995) *Arctic and alpine biodiversity: Patterns, causes, and ecosystem consequences. Ecological Studies* 113–332 pp.

Chapin III FS, Walker LR, Fastie CL, and Sharman LC (1994) Mechanisms of primary succession following deglaciation at Glacier Bay, Alaska. *Ecological Monographs* 64(2): 149–175.

Chapin III FS, Shaver GR, Giblin AE, Nadelhoffer KJ, and Laundre JA (1995) Responses of arctic tundra to experimental and observed changes in climate. *Ecology* 76: 694–711.

Chapin FS, *et al.* (2005) Polar systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystems and Human Well-being: Vol. 1: Current State and Trends*, pp. 717–746. Washington, District of Columbia: Island Press.

Conservation of Arctic Flora and Fauna Working Group (1999) *CAFF Technical Report No. 3 – Atlas of rare endemic vascular plants of the Arctic*. Technical Report. Conservation of Arctic Flora and Fauna Working Group (CAFF).

Conservation of Arctic Flora and Fauna (2001) *The Arctic Flora & Fauna: Status and Conservation Assessment*. Akyreuri, Iceland: CAFF International Secretariat.

Conservation of Arctic Flora and Fauna (2010) *Arctic Biodiversity Trends 2010: Selected Indicators of Change*. Akyreuri, Iceland: CAFF International Secretariat.

Edlund SA and Alt BT (1989) Regional congruence of vegetation and summer climate patterns in the Queen Elisabeth Islands, Northwest territories, Canada. *Arctic* 42: 3–23.

Gilichinsky D (2002) In: Britton G (ed.) *Encyclopedia of Environmental Microbiology*, pp. 932–956. Chichester, UK: Wiley.

Hodson A, *et al.* (2008) *Ecological Monographs* 78(1): 41–67.

IPCC (2007) In: Solomon S, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis*. Cambridge, UK and New York: Cambridge University Press. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Nutall M and Callaghan TV (2000) *The Arctic: Environment, People, Policy*. Reading, UK: Harwood.

Olofsson J, Oksanen L, Callaghan T, Hulme PE, Oksanen T, and Suominen O (2009) Herbivores inhibit climate-driven shrub expansion on the tundra. *Global Change Biololgy* 15: 2681–2693.

Porsild N (1959) *Circumpolar Arctic Flora*. Oxford: Clarendon.

Post E, Forchhammer MC, Bret-Harte MS, *et al.* (2009) Ecological dynamics across the Arctic associated with recent climate change. *Science* 325: 1355–1358.

Post E and Pedersen C (2008) Opposing plant community responses to warming with and without herbivores. *PNAS* 105: 12353–12358.

Press MC, Callaghan TV, and Lee JA (1998) How will European arctic ecosystems respond to projected global environmental change? *Ambio* 27: 306–311.

Quested H, *et al.* (2002) The hemiparasitic angiosperm *Bartsia alpina* has the potential to accelerate decomposition in sub-arctic communities. *Oecologia* 130: 88–95.

Rannie WF (1986) Summer air temperature and number of vascular species in arctic Canada. *Arctic* 39: 133–137.

Raynolds MK and Walker DA (2009) Effects of deglaciation on circumpolar distribution of arctic vegetation. *Canadian Journal of Remote Sensing* 35: 118–129.

Sturm M, Racine CH, and Tape KD (2001) Increasing shrub abundance in the Arctic. *Nature* 411: 546–547.

Svoboda J and Freedman B (eds.) (1994) *Ecology of a Polar Oasis. Alexandra Fjord, Ellesmer Island, Canada*. Toronto: Captus University Publications.

Tape KD, Sturm M, and Racine CH (2006) The evidence for shrub expansion in Northern Alaska and the Pan-Arctic. *Global Change Biology* 12: 686–702.

Van Bogaert R, *et al*. (2011) A century of tree line changes in sub-Arctic Sweden show local and regional variability and only a minor role of 20th Century climate warming. *Journal of Biogeography* 38: 907–921.

van der Wal R and Brooker RW (2004) Mosses mediate grazer impacts on grass abundance in arctic ecosystems. *Functional Ecology* 18: 77–86.

Verville JH, Hobbie SE, Chapin III FS, and Hooper DU (1998) Response of tundra CH4 and CO2 flux to manipulation of temperature and vegetation. *Biogeochemistry* 43(3): 215–235.

Walker DA, Gould WA, Maier HA, and Reynolds MK (2002) The Circumpolar Arctic Vegetation Map: AVHRR-derived base maps, environmental controls, and integrated mapping procedures. *International Journal of Remote Sensing* 23: 4551–4570.

Walker DA, Raynolds MK, Daniëls FJA, *et al*. (2005) The Circumpolar Arctic vegetation map. *Journal of Vegetation Science* 16: 267–282.

Walsh JE, Overland JE, Groisman PY, and Rudolf B (2011) Chapter 2: Arctic climate: Recent variations. *Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2011*, pp. 13. Oslo: Arctic Monitoring and Assessment Programme (AMAP).

Washburn AL (1956) Classification of patterned ground and review of suggested origins. *Geological Society of America* 67: 823–866.

Yurtsev B (1997) Effect of climate change on biodiversity of arctic plants. *Ecological Studies* 124: 229–244.

Zhenlin Y, Hanna E, Callaghan TV, and Jonasson C (2011) How can meteorological observations and microclimate simulations improve understanding of 1913–2010 climate change around Abisko, Swedish Lapland? *Meteorological Applications* 18. DOI: 10.1002/met.276.

Arthropods (Terrestrial), Amazonian

Joachim Adis, Max Planck Institute for Limnology, Germany, in cooperation with National Institute for Amazonian Research, Brasil

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–12,
© 2001, Elsevier Inc.

Glossary

Arboricolous Tree inhabitants ("arboreal").
Assemblage Collective occurrence of several individuals representing the same species or several species without interspecific relationships.
Bionomics Mode of life of a species.
Biotope Characteristic living space, of distinguishable nature compared with its environment, of a (bio) community (also "(bio)coenosis").
Bivoltine Species with two generations per year.
Community Collective occurrence of individuals representing several species with, at least in part, interspecific relationships (also "coenosis").
Flood resistance Submersion ability of weeks or months.
Flood tolerance Submersion ability of few hours up to several days.
Habitat Characteristic living space or site of a species.
Plastron Thin layer of air, held by specific body structures, into which oxygen from the surrounding water is added by means of Diffusion to the same extent as oxygen is withdrawn by breathing.
Plurivoltine Species with several generations per year (also "multivoltine").
Quiescence Dormancy in which inhibition of development depends directly on environmental factors.
Terricolous Soil inhabitants.
Univoltine (Bivoltine) Species with one (two) generations per year.

Taxa and Number of Species

Approximately 2% of the world's Arachnida and 3% of the Myriapoda live in Amazonia. However, Ricinulei represent 28%, Schizomida and Scolopendromorpha 9%, and Pauropoda more than 7% of species known worldwide in the respective taxa (**Table 1**). To estimate the extant species in Amazonia is difficult, especially in those cases in which numbers differ by two or three different orders of magnitude from the species described worldwide. For mites which, like spiders, are considered as a mega- or hyperdiverse taxon in the Arachnida, opinions of taxonomists differ substantially. This is also due to different methods used to derive such estimates (Platnick, 1999). In millipedes, ~250 species have been described from the Amazon drainage basin but 5000 are estimated to exist (**Table 1**). This is based on the assumption that most genera originated in the Andes, where most of the known species now occur but there have been few sampling efforts.

Megadiversity and high densities are attributed to Neotropical insects, in particular those inhabiting forest canopies. The dominant taxa in the canopy of nonflooded primary forests in Central Amazonia are Formicidae and Diptera. In studies by Guerrero in 1995 and 1996 near Manaus (Reserva Ducke), these two taxa represented 52% and 10%, respectively, of the 325 arthropods per square meter obtained on average by fogging the canopy of 40 trees (nine species of the families Sapotaceae and Lecythidaceae) once each during the dry and rainy seasons. In another study in this reserve, two canopies of a widely distributed Amazonian tree, *Goupia glabra* Aubl. (Celastraceae, height 38 and 45 m) were fogged in intervals of 6 or 24 months (1991–94). A total of 95 ant species were found on a single tree. In comparison, there are 105 ant species recorded from all over Germany. Of the 124 ant species obtained from both trees, more than one-third were represented by singletons, that is, one specimen per species. Data also indicated a biotic interaction between predatory ants (probably *Crematogaster* spp.) and the gall-building Cecidomyiidae and between Cecidomyiidae and the parasitic Hymenoptera. This was not determined from previous studies in the tropical canopy (Adis *et al.*, 1998a).

Comparative with the high density and biodiversity of Formicidae and Diptera in tree canopies is that of Acari and Collembola on the forest floor. These two taxa represent between 50% and 80% of the 26,000–74,000 terricolous arthropods extracted per square meter from 14 cm soil depth in nonflooded "terra firme" forests of Central Amazonia. However, owing to taxonomical difficulties in both groups, data on species diversity are scarce and based on adult animals, although just as many immatures have been caught. The highest number of oribatid mites, 97 morphospecies, were reported by Wunderle from a primary forest near Pucallpa, Peru. In the Manaus area, Franklin found 71 and 74 morphospecies of oribatid mites in a primary forest on yellow latosol and on white sand, respectively, and 57 morphospecies in a 3-year-old secondary forest on yellow latosol (previous primary forest that was cut and burned). The ratio between the total number of specimens and morphospecies decreased accordingly, being 53.3, 36.3, and 22.8, respectively. In the same forests, Oliveira obtained 74, 65, and 65 morphospecies, respectively, of Collembola (53–64% represented Isotomidae and 7–16% Entomobryidae), but specimens: morphospecies ratio was lowest in the primary forest on white sand (49.5) compared with the primary and secondary forests on yellow latosol (67.6 and 61.8, respectively). Statistical analysis of data indicated that morphospecies similarity is dependent on both the nature of the soil (in springtails) and human disturbance (in oribatid mites). Only 32 oribatid morphospecies were obtained from pastures near Manaus.

Table 1 Families, genera, and species of taxa representing the Arachnida and Myriapoda in the world and in Amazonia (up to 1999) as well as estimates of existing species (fossil taxa not included)

		Families	*Genera*	*Species*	
				Described	*Estimated*
Arachnida	World	>566	>8869	>92,529	
	Amazonia	>136	>480	>1592	
Araneae	World	106	±3200	±37,000	76,000–170,000
	Amazonia	>65	>300	>1000	4000–8000
Palpigradi	World	2	6	80	100
	Amazonia	1	1	1	
Uropygi (Thelyphonida)	World	2	16	101	
	Amazonia	2	2	3	10
Amblypygi	World	5	20	126	
	Amazonia	1	2	11	
Schizomida	World	2	31	195 (219)	
	Amazonia	1	3	10	
Solifugae	World	12	153	1065	(1115)
	Amazonia	1	1	1(2)	
Pseudoscorpiones	World	24	430	3100	3500–5000
	Amazonia	12	31	75	>150
Ricinulei	World	1	3	53	85
	Amazonia	1	1	15	25
Opiliones	World	44	>1554	>4559	
	Amazonia	13	75	160	
Cyphophthalmi	World	6	28	109	
	Amazonia	3	4	4	
Palpatores	World	15	>250	1000–2000	
	Amazonia	1	5	20	
Laniatores	World	23	1276	3460	
	Amazonia	9	66	136	
Scorpiones	World	18–20	156	1250–1500	6000–7500
	Amazonia	4	12–14	68–111	200
Acari	World	350–422	3300–4000	45,000	0.5–1 million
	Amazonia	35	50	150–300	20,000–250,000
Myriapoda	World	158 (159)	2167	>15096	59,400–92,500
	Amazonia	28 (29)	>94	>423	5600–7600
Chilopoda	World	20 (21)	325	>3196	6850–6950
	Amazonia	8 (9)	26	>115	400
Scutigeromorpha	World	1	16	>80	100–150
	Amazonia	1	1	2	5
Lithobiomorpha	World	2	95	1500	2000
	Amazonia	1	1	2	5
Craterostigmomorpha	World	1	1	1	?5
	Amazonia	0	0	0	0
Scolopendromorpha	World	3	33 (32)	515–616	700–800
	Amazonia	2	11	52	90
Geophilomorpha	World	13 (14)	180	1100	4000
	Amazonia	4 (5)	13	60	300
Diplopoda	World	131	1800	11,000	50,000–80,000
	Amazonia	16	55–60	250	5000–7000
Pauropoda	World	5	29	700	2000–5000
	Amazonia	2	8	52	>200
Symphyla	World	2	13	200	?500
	Amazonia	2	4	5	10–20

Note: Numbers in parentheses include forthcoming descriptions and revisions.

Source: Modified from Adis J and Harvey MS (2000) How many Arachnida and Myriapoda are there worldwide and in Amazonia. *Stud. Neotrop. Fauna Environ.* 35(2): 139–141.

Origin and Geographical Distribution of Species

Little is known about the centers of evolution and dispersal for arthropod taxa that occur in Amazonia. Several pathways are currently being discussed.

Origin in the Neotropics with a Generally Wide Range of Geographical Distribution (Chilopoda: Scolopendromorpha)

No scolopendromorph taxon of generic level is known as Amazonian endemic as it is at least present in other areas of

the Neotropical realm. Studies of Schileyko showed that only 25% of the total 52 species currently known in Amazonia, representing 6 of the total 11 genera, seem to be endemic. The wide geographical distribution might be related to the relatively large size, high mobility, swimming ability, and predaceous mode of life of an ancient group.

Origin in the Andes and/or the Guyanian Shield and Subsequent Dispersal into the Amazon Basin (Archaeognatha: Meinertellidae, Coleoptera: Carabidae, Diplopoda: Polydesmida)

Representatives of the Meinertellidae probably advanced from the Andes into the lower situated forests of the Amazon Basin. During this process, the formerly petrophilous animals, inhabiting mountain floors, adapted to an arboricolous mode of life in the trunks and canopies of trees. They even acquired a vertical jumping ability on the tree trunk, which was found for the first time in the arboricolous species *Neomachilellus adisi* in a black-water inundation forest of Central Amazonia. In these periodically flooded forests, the Meinertellidae succeeded to colonize the forest floor as well. Flood resistance of eggs (observed in *N. scandens*) and the determination of egg development by the flood pulse (= quiescence duration correlated with the period of inundation) are prerequisites for a potential distribution on the waterway (on logs or floating meadows). *N. adisi* was also found in an inundation forest above Leticia (Columbia) and might have reached Central Amazonia on the Rio Solimões-Amazonas. Similarly, the *Neomachilellus* species which currently inhabit the Caribbean and the coastal regions of the United States might originally have been transported by ocean currents from the mouth of the Rio Amazonas and the Rio Orinoco to these areas. An actually waterproof, resistant blastoderm cuticula and a presumably long duration of egg development during transportation (1 year or more, as in European species) certainly favor a geographical distribution on waterways. An interruption of the quiescence owing to drying of the means of transport used, a rapid eclosion of juvenile animals, and a fast achievement of maturity (all observed in *N. scandens*) favor the new colonization of a biotope as well. Useful for this colonization is the swimming ability and a temporal survival on the water surface (flood tolerance), which was reported for several species of the Machiloidea (Adis, 1992; Sturm, 1984).

Erwin and Pogue presented biogeographical maps with centers of endemism, based on taxa distribution and cladograms, for sister groups of the arboricolous carabid beetles genus *Agra* from Neotropical forests. Patterns of species evolution are directed, in part, from the North Atlantic coast (including the northern Guyanian shield) to upper Amazonia (*palmata* lineage) and from the North Atlantic coast to either upper and middle Amazonia or lower Amazonia (*erythropus* lineage). According to Erwin (1991), these hypothetical "centers and corridors of radiation" can be discovered and targeted for long-term protection through analyses of diverse arthropod groups (beetles in particular) and detection of congruent patterns among radiation lineages.

Origin in Nonflooded and/or Floodplain Forests of the Amazon Basin (Scorpiones, Chilopoda: Geophilomorpha)

Many scorpion species apparently require strict environmental conditions and thus are found only in narrow zones of distributions. Very few species currently known from eastern Amazonia are present in western Amazonia. One of the four principal "corridors of distribution" (or radiations) in South America is located in Central Amazonia. Rupture zones and ecobarriers between these corridors are attributed to geohistorical events. For instance, the rupture zone of distribution found in the region of the Orinoco delta might be due to the proposed existence of a large inland lake in Central Amazonia during the late Pleistocene and early Holocene (Lourenço, 1994).

In the Geophilomorpha, half of the ~100 Neotropical species representing the genera *Schendylops*, *Pectiniunguis*, and *Ribautia* seem to be endemic to areas identified as Pleistocene forest refuges (*see* Species Richness: Causes and Maintenance) based on evidence from woody angiosperms (Prance, 1982). Valleys of the major rivers (Amazon, Orinoco, Paraná, and Paraguay) are suggested to be the main routes of subsequent distribution, even to high-altitude sites (Pereira *et al.*, 1997).

Origin in Treeless Floodplains of the Amazon Basin (Coleoptera: Carabidae)

According to Erwin, the main centers of evolution in carabid beetles are open floodplains along rivers or lakeshores in the equatorial tropics with somewhat stable conditions. Taxon pulses which effect these flying and highly vagile waterside generalists caused, by means of adaptations, their dispersal in three directions: (i) colonization of treetops (canopy specialists), (ii) colonization of mountains (altitude specialists), and (iii) colonization of temperate regions (climate specialists). For the colonization of temperate regions, seasonality is considered as a basic requirement. Seasonal changes, primarily owing to periodical precipitation (= primary or ultimate ecofactor), cause seasonal inundations in riverine habitats (e.g., Central Amazonia) or dryness of the soil owing to a decreasing river level (e.g., East Zaire). Both situations cause an interruption of the reproduction in carabid beetles during these unfavorable conditions. They induce a gonad dormancy which is regulated by changing temperatures in the habitat (= secondary or proximate ecofactor). For this reason, Central Amazonian inundation forests have been postulated to represent one of the locations at which seasonality has been generated in carabids and other arthropods (Adis, 1992; Erwin, 1998).

Species Richness: Causes and Maintenance

Biological diversification in the Neotropics is the subject of ongoing and frequently antagonistic discussions. Several theories have been proposed to explain the high species richness of Amazonian plants and animals, mostly in a deductive sense, based on selected groups. Studies on distribution patterns aim to detect centers of endemism. It is hoped that biogeographical maps, now also based on cladistic methods, will elucidate evolutionary patterns among these centers.

However, it is accepted that databases are still insufficient (with the possible exception of those for birds) for the type of detailed study needed to understand distribution patterns for most Neotropical groups. Full knowledge of distribution will not be obtained until we understand what sets the limits, for example, the edges of species distribution (Vanzolini and Heyer, 1988).

Past assessments of this type have been based mostly on the presence of Pleistocene refugia (*see* Origin and Geographical Distribution of Species), also called the "Haffer effect." This is the theory that Neotropical forest fragmentation occurred during northern ice advances due to Milankovitch cycles and the temporary existence of forests and grasslands (savanna), called "island" refugia, that caused temporal centers of endemism (Haffer as cited in Prance, 1982; Haffer, 1996).

Studies on 162 species of Neotropical butterflies (Nymphalidae) by Brown indicated the existence of 50 principal forest centers of evolution and endemism. Smaller areas, recognizable as 38 forest refuges, presumably acted during the last major dry, cold spell in the Quartenary (20,000–13,000 years BP). Studies by Erwin and Pogue on arboricolous carabid beetles, adapted to life in Neotropical forest canopies (genus *Agra*), showed that species group distribution is much broader. In this case, areas considered to be centers of endemism are geographically larger than that in butterflies and vertebrates (Prance, 1982; Vanzolini and Heyer, 1988).

Haffer's forest refuge model as stated deals only with nonflooded terra firme forests, their fragmentation, and/or replacement by savannas (*sensu lato*). There has been no indication that riverine forests disappeared along the lengths of rivers, but rather it appears that savannas were behind such forests as they are today. For this reason, Erwin and Adis consider Central Amazonian floodplain forests, especially along black-water rivers in the Negro delta (called igapó), to be short-term refuges and long-term evolutionary centers (Prance, 1982; Adis, 1984). Flood cycles greatly influence speciation rate and dispersion of arthropods. Evidence from Rio Tarumã Mirím, an affluent of the Rio Negro above Manaus, suggests alternation of two types of inundation: annual flooding and continental water table rise during the past million years. The area examined originally was covered by a nonflooded terra firme forest. Periods of high sea level caused a backup of main riverine courses in Amazonia. During these periods, the affluent Tarumã Mirím was high enough that annual water fluctuations of the Rio Negro led to an inundation lasting several months in the forest under study. The first inundation was several million years ago. Since that time, various periods of high sea level (according to Fairbridge) have occurred (e.g., 250,000, 170,000, and 85,000 years ago; **Figure 1**), with annual flooding of 10,000–30,000 years duration. Approximately 6000 years ago, the sea level reached a height of about 5 m below mean sea level. Since then, the examined forest has been inundated again. Flood cycles were responsible for extant vegetation formation in the igapó. During high sea level, the (lower) bank of the Rio Negro suffered longer periods of inundation and, according to Sioli, possible formations of large "ria lakes." Igapó forests were drowned (at least partially) and backed up minor tributaries which subsequently had annual flooding (e.g., Rio Tarumã Mirím; **Figure 1a**). They replaced nonflooded terra firme forests that previously occupied the area. Intermittent dry spells during high sea level periods may have caused

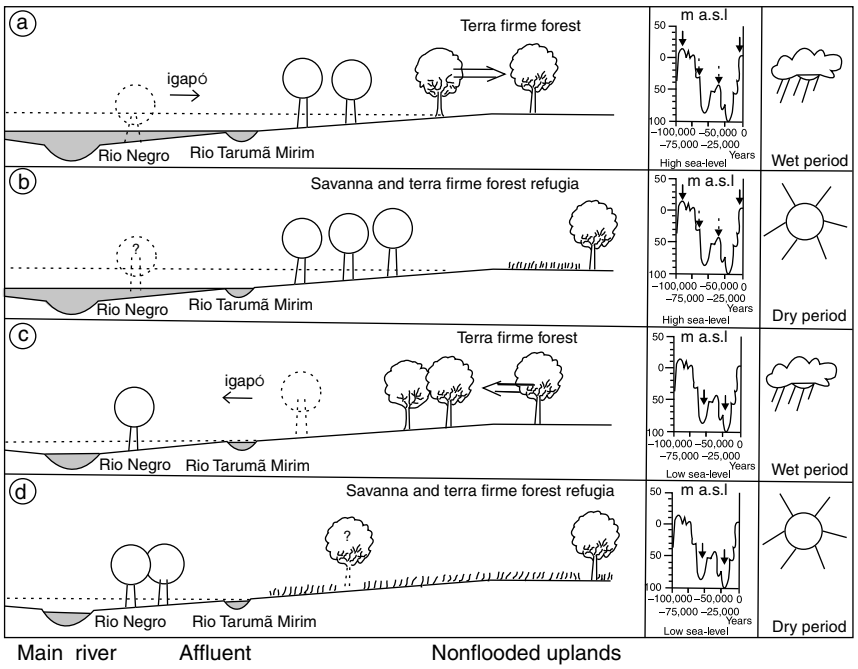


Figure 1 Postulated formations and shifts of black-water inundation forests along main rivers and affluents of Central Amazoni within the past 100,000 years due to changes in sea level and climate. (See text for further explanation.) Modified from Adis J (1984) "Seasonal igapó"-forests of Central Amazonian black-water rivers and their terrestrial arthropod fauna. In: Sioli H (ed.) *The Amazon. Limnology and Landscape Ecology of a Mighty Tropical River and Its Basin. Monographiae Biologicae*, pp. 245–268. Dordrecht: Junk.

formation of savanna and nonflooded terra firme forest refugia (**Figure 1b**), as postulated by Haffer. During low sea level periods, annual inundation occurred only along main rivers (e.g., the Rio Negro); this has been the case for at least 1 My, probably including all glacial advances of the Pleistocene and perhaps even since the Andes took their present configuration. Nonflooded vegetation on terra firme ("uplands") spread extensively, forcing igapó forests along affluents to retreat to main rivers (**Figure 1c**). Intermittent dry periods, according to Haffer, again caused savanna formation with possible nonflooded forest refugia on terra firme and "igapó forest refuges" (**Figure 1d**).

Flood cycles have a dramatic impact on inundation forest faunas. This is especially important when two kinds of flooding patterns manifest themselves through time. In the greater region of Amazonia, there could have been periods favoring lakeshore species, then inundation forest species, and then riparian species (for discussion of carabid beetles, see Erwin and Adis as cited in Prance, 1982). Large water bodies and drying affluents, with possibly additional regression of nonflooded forests on terra firme in both cases, cause interruption and isolation of igapó forests. Species vagility then seems to become a crucial factor in sorting out and isolating gene pools. Thus, small barriers may be highly disruptive to populations of individuals with low vagility, particularly nonflying soil arthropods. Supposing that intense competition and predation promote rapid speciation (especially in igapó forests, in which flooding mixes terrestrial and canopy biotas during half the year) and that the isolation of igapó forests is sufficiently long for rapidly evolving species, this model represents another key to Amazonian species richness. In fact, recent data indicate that (i) igapó forests retain many endemic species; (ii) the fauna of igapó forests is different in composition and activity patterns when compared with that of adjacent nonflooded terra firme forests; and (iii) in igapó forests across the Rio Negro and its tributaries, taxonomically different subspecies and species are found.

A third theory considers the high plant and animal diversity of Amazonian nonflooded terra firme to be a response to extremely low nutrient concentrations in geochemically impoverished ecosystems (Fittkau, 1973; Klinge, 1973). Species richness and simultaneously occurring high diversity of moist tropical ecosystems do not reflect high nutrient supply but rather a mode of adaptation to continuous restriction of nutrients or food substances under otherwise permanently favorable living conditions. Plants and animals are assumed to act as highly efficient "nutrient traps." During evolution of multiple life forms, shortages must have been of significant importance at all times.

A model group supporting this theory are the carabid communities on fig fruit falls. Paarmann and Adis sampled 8962 beetles on 65 fruit falls from 10 fig species between 1991 and 1996 in a nonflooded terra firme forest (Reserva Ducke) near Manaus. Eight of the 36 species collected represented the spermatophagous genus *Notiobia*, which accounted for 92% of all carabid beetles obtained. Only one *Notiobia* species was dominant. The abundance distribution of the eight *Notiobia* species was very similar on fruit falls of the two most common fig species as well as on fruit falls of the remaining fig species. The dominance structure varied considerably between individual fruit falls. Also, the abundance distribution changed during the course of a single fruit fall. Only two of the eight *Notiobia* species were found to be specialized fig seed feeders, able to reproduce only on fig fruit falls (Vanicek *et al.*, 1994). The remaining six species of this genus use fig fruit falls as stepping stones between fruit falls of their host trees, which are separated by time and space.

Another theory for the high within-community diversity of tropical arthropods is that continuous stochastic local disturbances in nature are assumed to prevent the achievement of any long-term equilibrium (climax) state. These so-called stochastic nonequilibrium models assume that the presence of a species at a vacant site is important. This may represent an advantage against all species that arrive later. Niche overlaps are assumed to be very common within species-rich communities. As a result, neither successional nor a climax community can emerge. This is in contrast to the so-called deterministic equilibrium models, which are based on the ecological niche. Each organism maintains a defined position in its environment and, driven by competition, the system goes through defined successional stages and a structurally predictable climax equilibrium results (Linsenmair, 1990).

Examples that clearly support one or the other mechanism are not available for Amazonian arthropods. For example, concerning the diversity of the seed-feeding ground beetle community of the genus *Notiobia*, they are specialized on certain seeds as food for a successful larval development (deterministic process) and yet the nonfig seed specialists consume fig seeds to survive periods of food shortage. The necessity to do so limits the tendency to specialize. If stochastic processes do play a role in maintaining the diversity of the seed-feeding *Notiobia* species, this role is probably less important. Thus, they are "chance specialists" adapted to the unpredictability of fruit falls. The fig seed specialists belong to the "touring group," moving from one fig fruit fall to the next, depending on a sufficient density and an annually equal distribution. The aseasonal fruiting of figs is caused by their pollination biology.

Biotope and Habitat Specificity of Species

Conservation or sustainable use of Amazonian ecosystems demand information on biodiversity. Owing to the lack of basic comparative data on arthropods, recent recommendations mostly relate to the high diversity and rarity of species found in certain regions, which are called areas of "high biological value" (Anonymous, 1991; Dinerstein *et al.*, 1995; Kress *et al.*, 1998), rather than to specific indicator species of biotopes and/or habitats revealing endemism and possible centers of radiation. Two long-term studies on pseudoscorpions and tiger beetles demonstrate the validity for such information.

Pseudoscorpions were studied intensively in Central Amazonian biotopes between 1975 and 1990. The 35,000 specimens collected represented 26 genera and 60 species (Adis and Mahnert, 1990). Of these, 29 species (48%) occurred exclusively in nonflooded terra firme forests and 25 species (42%) were restricted to floodplain forests. Species from terra firme forests were more terricolous, with 69%

Table 2 Species of tiger beetles (Carabidae: Cicindelinae) sampled in white-water floodplains (Várzea) of the Rio Solimões-Amazonas and on nonflooded terra firme near Manaus, Central Amazonia

	Várzea floodplains		*Terra firme uplands*
Species	9		15
Forest	3		7
Joint		1	
Open areas	6	0	8
Joint			
	Cylindera (Plectographa) suturalis (Fabricius) (O, d)		*Aniara sepulcralis* (Fabricius) (O, d)
	Megacephala (Tetracha) sobrina punctata Castelnau (O, n)		*Brasiella (Cicindela) argentata* (Fabricius) (O, d)
	Megacephala (Tetracha) spinosa (Brullé (O, n))		*Brasiella (Cicindela) pretiosa* (Dokhtouroff) (O, d)
	Megacephala (Phaeoxantha) aequinoctialis Dejean (O, n)		*Cenothyla varians* (Gory) (F, d)
	Megacephala (Phaeoxantha) klugi Chaudoir (O, n)		*Ctenostoma (Stenoctenostoma) aspendum* Bates (F, ?)
	Odontocheila confusa (Dejean) (F, d)		*Ctenostoma formicarium* (Fabricius) (F, ?)
	Pentacomia (Mesacanthina) cribrata (Brullé) (O, d)		*Cylindera (Cylindera) morio* (Klug) (O, d)
	Pentacomia (Pentacomia) egregia (Chaudoir) (F, d)		*Megacephala (Tetracha) bilunata* Klug (O, n)
	Pentacomia (Poecilochila) lacordairei Gory (F, d)		*Odontocheila cayennensis* (Fabricius) (F, d)
			Odontocheila chrysis (Fabricius) (O, d)
			Odontocheila luridipes (Dejean) (F, d)
			Odontocheila margineguttata (Dejean) (O, d)
			Odontocheila nigrotarsalis Horn (F, d)
			Pentacomia (Poecilochila) lacordairei Gory (F, d)
			Pentacomia (Poecilochila) ventralis (Dejean) (O, d)

F, forests; O, open areas; d, diurnal; n, nocturnal.

Source: Modified from Adis J, Paarmann W, Amorim MA, Arndt E, and da Fonseca CRV (1998b) On occurrence, habitat specificity and natural history of adult tiger beetles (Coleoptera: Carabidae: Cicindelinae) near Manaus, Central Amazonia, and key to the larvae of tiger beetle genera. *Acta Amazon* 28(3): 331–344.

living in litter and soil. Species from floodplain forests were predominantly arboricolous, with 88% living in the trunk and/or canopy region. Differences in habitat selection are attributed to the impact of the flood pulse in Amazonian floodplains, that is, long-term inundation of 5–7 months' duration. Only 11 species were found in forests inundated by white-water (Rio Solimões-Amazonas) compared with 20 species found in black-water forests (Rio Negro). This is attributed to the differences observed between biotopes (flora and soil structure) as well as between the two river systems (current and sediment load; Adis as cited in Junk, 1997).

Databases such as these help to answer general questions about specific biotopes. For example, the terricolous pseudoscorpions of terra firme forests on white sand soil in Central Amazonia (called campinarana) represent the species spectrum of primary terra firme forests on yellow latosol and show no endemic species (Adis and Mahnert, 1993). This reinforces the geological results that indicate that campinarana does not represent forests along former riverbeds which have dried up, as previously suggested, but rather the final stage of podzolization (i.e., the transformation of clayey latosols to white sand podzols by long-term weathering and leaching processes).

Of the 24 species of tiger beetles (Carabidae: Cicindelinae) studied between 1991 and 1998 in the Manaus area (**Table 2**), 15 were found in nonflooded terra firme areas. Of these, seven species lived in forests. Five of them, being diurnal, inhabited the floor (**Figure 2**: CVA, OCA, OLU, ONI, and PLA), and two species were canopy dwellers (CAS and CFO). Eight species, one nocturnal (MBI) and seven diurnal (ASE, BAR, BPR, CMO, OCH, CAS, OMA, and PVE), lived in open areas. Three of these were found on mostly bare latosol (BAR, OCH, and PVE). The other five species occurred on mostly bare white sand (ASE, BPR, CAR, CMO, and MBI). The species assemblage in open areas indicated an impact by human intervention: *Pentacomia ventralis* (PVE) is known as a species that inhabits extensive cleared forest sites and *Odontocheila margineguttata* (OMA) occurs in secondary forests. Another nine tiger beetle species were recorded from white-water floodplains (várzea; **Table 2**). Of these, three diurnal species inhabited inundation forests (**Figure 3**: OC, PE, and PL). Six species, two diurnal (CS and PC) and four nocturnal (MA, MK, MP, and MS), lived in open areas. Four were found on bare beaches (CS, PC, MP, and MS). The other two species (MA and MK) inhabited grass-grown areas behind the beaches which were mainly composed of annual and perannual aquatic grasses. Only one species, *P. lacordairei*, occurred in the forests of both nonflooded terra firme and floodplains (**Table 2**).

Data clearly show the biotope and habitat specificity of these tiger beetle assemblages. Moreover, the life cycle of species on nonflooded terra firme differed greatly from those inhabiting floodplains, which indicates a high adaptation to the respective biotope and habitat. Data such as these might also help to elucidate questions of general interest. For example, the dichotomy of forest or open-habitat assemblages might be important for understanding tiger beetle distribution in Amazonia. Similar patterns were reported by Heyer for frogs east of the Andes (Vanzolini and Heyer, 1988). Data also support the potential use of tiger beetles as bioindicators for monitoring the degradation and regeneration of Amazonian forests. In Venezuela, forest-floor species assemblages changed significantly with the degree of forest disturbance, and each stage of disturbance was characterized by a particular subset of species (Rodríguez *et al.*, 1998).

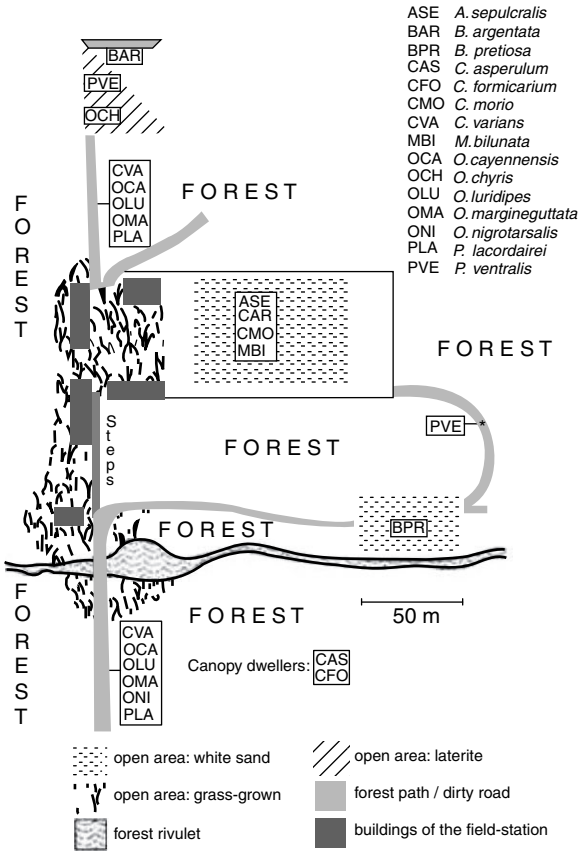


Figure 2 Species of tiger beetles (Carabidae: Cicindeline) occurring in the vicinity of the field station at the Adolpho Ducke forest reserve (Reserva Ducke) near Manaus, Central Amazonia. Reproduced from Adis J, Paarmann W, Amorim MA, Arndt E, and da Fonseca CRV (1998b) On occurrence, habitat specificity and natural history of adult tiger beetles (Coleoptera: Carabidae: Cicindelinae) near Manaus, Central Amazonia, and key to the larvae of tiger beetle genera. *Acta Amazon* 28(3): 331–344.

Morphological Species

In millipedes, taxonomy is mainly based on the structure of genitalia in males, the gonopods. However, in the process of adaptation to new environments, "pioneers" apparently change their external structure more quickly than their gonopods, which results in difficulties in defining species. According to Hoffman (1990), the diversification of the external body form with little modification of gonopod structures in the order Polydesmida seems to be associated with those taxa (usually genera) that seem to have recently occupied a new area or biotope independently of close relatives and are indulging in a burst of adaptive radiation unimpeded by sibling competition. It is presumed that after an initial period of great diversification some selection of successful lineages would occur. During this time, body form types, stabilized by adaptational factors, would tend to remain constant. The generic variability, however, might be expressed much more rapidly in structures such as genitalia, which are less directly influenced by environmental constraints. Thus, a period would ensue in which these species groups (or genera) would consist of species readily distinguishable by highly distinctive gonopods. Such a stasis would endure for a long time unless it were again possible for a fragment of a group (probably migrant populations of one or two species) to escape into a new area or new ecological niche and begin the cycle anew. This situation seems to be different from the punctual evolution of new clades in that two basically different character systems alternate in phase with the chronological status of the organisms in "new" and "old" territory and thus underscore the principle that phylogeny cannot be adequately interpreted outside the context of biogeography.

New insights derived from studies on the Neotropical genus *Pycnotropis* support these postulations. Among the 26 species known, only three are restricted to Amazonian white- and mixed-water inundation forests, whereas the remainder are found in nonflooded terra firme forests. Biogeographically, the history of *Pycnotropis* can be viewed as one implying the origins in the western Andes, the postulated center of origin (and radiation) of the entire subfamily Amplininae, with subsequent waves of downstream dispersal throughout the Amazon Basin. The present-day distribution of *Pycnotropis* is the result of numerous, apparently ongoing vicariance events involving repeated gene flows downstream of the Amazon and its tributaries (radiation corridors) from source areas, with colonization and recolonization of inundation forests and/or nonflooded terra firme habitats. The genus is still in a stage of very active speciation because at least some of its species display pronounced variation in both external and gonopod characters. This is supported by preliminary tests concerning genetic variation in *Pycnotropis tida*, a species that has been described twice due to variation in morphological characters. The pattern of a somewhat "unsettled" speciation process seems to contrast the one observed in the chelodesmid genus *Camptomorpha*, which likewise is species-rich in the Andes and abundantly represented in the Amazon Basin. It is, however, a good example of a taxon in the "stable body-variable gonopod" stage, with each species being distinct morphologically (Golovatch *et al.*, 1998).

Genetical Species

Comparative phenological studies on arthropod species that inhabit both floodplain and nonflooded terra firme forests in Central Amazonia suggested an uni- or bivoltine life cycle of populations in annually inundated forests and a plurivoltine mode of life of populations in terra firme forests. Representatives of both populations were morphologically alike; hence, they represented the same species (e.g., in Pseudoscorpiones, Diplopoda, Symphyla, and Archaeognatha).

In the case of *N. scandens*, a member of the Meinertellidae (Archaeognatha, Insecta), the two populations in terra firme and floodplain forests were first postulated to be "biotope-specific races." Morphological characters that would justify separation into two species could not be found by the taxonomist, at least with traditional research methods. However, protein analyses by means of electrophoresis (testing 15 enzymes) revealed that there is no gene flow between populations from floodplain and terra firme forest types, even when they are only 50 m distant from each other. Thus, this "species" was proposed to represent two different species based on ecophenological and genetic characteristics. Without this

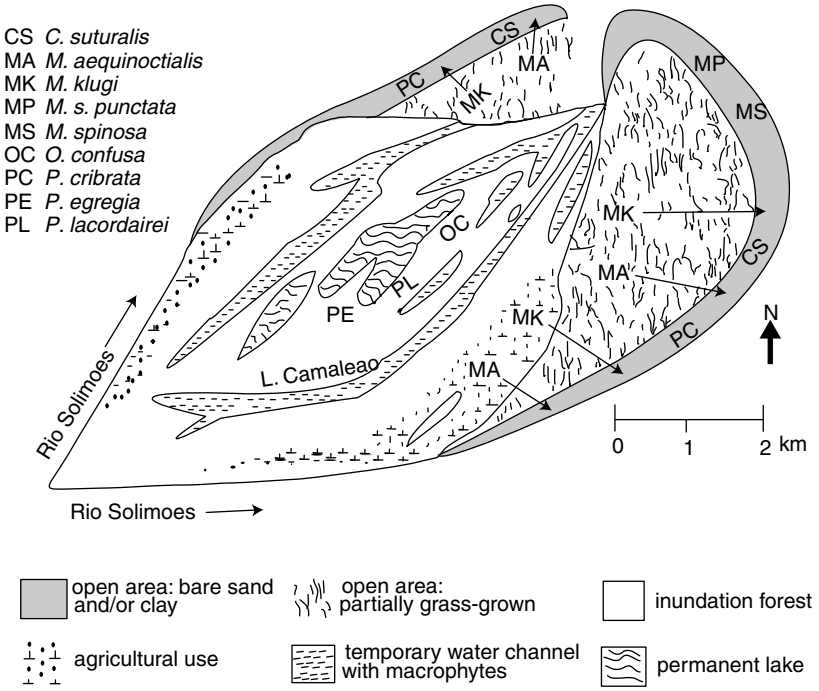


Figure 3 Species of tiger beetles (Carabidae: Cicindelinae) found on Ilha de Marchantaria, the first island in the Rio Solimões near Manaus, Central Amazonia. Reproduced from Adis J, Paarmann W, Amorim MA, Arndt E, and da Fonseca CRV (1998b) On occurrence, habitat specificity and natural history of adult tiger beetles (Coleoptera: Carabidae: Cicindelinae) near Manaus, Central Amazonia, and key to the larvae of tiger beetle genera. *Acta Amazon* 28(3): 331–344.

genetical background, problems arise if the two genetical species occur sympatrically because, due to the intermixture of data, the ecophenological differences are hardly or not at all recognizable. This was the case in a mixed-water inundation forest near Manaus (Wolf and Adis, 1992). Until now, culturing and breeding of Amazonian Meinertellidae was unsuccessful; thus, interbreeding experiments could not be performed.

The polydesmidan millipede *P. tida* (*see* Morphological Species) represents a second example. In the Manaus area, populations that inhabit secondary terra firme forests breed throughout the year, whereas populations of white- and mixed-water inundation forests show a defined, univoltine reproduction during the nonaquatic phase (Vohland and Adis, 1999). The genetic data were obtained from allozyme analyses (testing 14 enzymes by electrophoresis) and from a specific satellite DNA. In this case, results suggested that individuals from the two biotopes represented populations of the same species, although genotypic structures among and within local populations indicated processes of ongoing genetic differentiation (Bachmann *et al.*, 1998). A successful interbreeding thereafter reinforced the genetical results.

For taxonomists, the use of genetic techniques and the consideration of ecological as well as ethological data represent the final, although mandatory, stage of analysis in systematic biology. In the case of Amazonian Pseudoscorpiones and Symphyla, both genetic and breeding exercises have not been successful.

Prospects

Amazonian terrestrial arthropods are special: they have adapted in some way to their environment. Only detailed studies at the species level will reveal the "strategies" which have evolved. For example, there are (i) grasshoppers on aquatic macrophytes in the Rio Amazonas with spiny front legs, like mantids, that follow their prey into tree canopies during high water (*Phlugis teres*, Tettigoniidae); (ii) a small, flood-resistant "terrestrial" millipede which lives up to 11 months submerged in inundation forests along the Rio Negro, breathing under water by means of a plastron and feeding on algae (*Gonographis adisi*, Pyrgodesmidae); and (iii) larvae of a carabid beetle (normally a predaceous taxon) which accomplish their development feeding on fig seeds (*N. flavicintus*) but apparently depend on larvae of their sister species (*N. pseudolimbipennis*) which are able to open the seed shells with their morphologically larger mandibles. Some insights into this fascinating world have been achieved since Bates voyages on the "River Amazons." However, it took about 150 years to adopt his suggestion to study 1 acre of rain forest in more detail. Two approaches are currently being taken. One is to develop an All-Taxa-Biodiversity-Inventory and the other to develop an All-Biota-Taxon-Inventory (Platnick, 1999). Both enterprises require an intensified training of tropical taxonomists and ecoentomologists.

Acknowledgments

I gratefully acknowledge the valuable comments received on a draft of this article by Terry L. Erwin, Sergei I. Golovatch, Richard L. Hoffman, Wilfried Paarmann, Norman I. Platnick, and Harald Sioli.

See also: Amazon Ecosystems. Arachnids. Beetles. Forest Canopies, Animal Diversity. Hymenoptera. Myriapods

References

Adis J (1984) "Seasonal igapó"-forests of Central Amazonian black-water rivers and their terrestrial arthropod fauna. In: Sioli H (ed.) *The Amazon. Limnology and Landscape Ecology of a Mighty Tropical River and Its Basin. Monographiae Biologicae*, pp. 245–268. Dordrecht: Junk.

Adis J (1992) Überlebensstrategien terrestrischer Invertebraten in Überschwemmungswäldern Zentralamazoniens. *Verh. Naturwiss. Ver. Hamburg (NF)* 33: 21–114.

Adis J, Harada AY, da Fonseca CRV, Paarmann W, and Rafael JA (1998a) Arthropods obtained from the Amazonian tree species "Cupiuba" (*Goupia glabra*) by repeated canopy fogging with natural pyrethrum. *Acta Amazon* 28(3): 273–283.

Adis J and Harvey MS (2000) How many Arachnida and Myriapoda are there worldwide and in Amazonia. *Stud. Neotrop. Fauna Environ.* 35(2): 139–141.

Adis J and Mahnert V (1990) On the species composition of Pseudoscorpiones (Arachnida) from Amazonian dryland and inundation forests in Brazil. *Rev. Suisse Zool.* 97: 49–53.

Adis J and Mahnert V (1993) Vertical distribution and abundance of pseudoscorpions (Archachnida) in the soil of two different Neotropical primary forests during the dry and rainy seasons. *Mem. Qd Mus.* 33(2): 431–440.

Adis J, Paarmann W, Amorim MA, Arndt E, and da Fonseca CRV (1998b) On occurrence, habitat specificity and natural history of adult tiger beetles (Coleoptera: Carabidae: Cicindelinae) near Manaus, Central Amazonia, and key to the larvae of tiger beetle genera. *Acta Amazon* 28(3): 331–344.

Anonymous (1991). Workshop 90: Biological Priorities for Conservation in Amazonia (map and explanatory text). Washington, DC: Conservation International.

Bachmann L, Tomiuk J, Adis J, and Vohland K (1998) Genetic differentiation of the millipede *Pycnotropis epiclysmus* inhabiting seasonally inundated and non-flooded Amazonian forests. *J. Zool. Syst. Evol. Res.* 36: 65–75.

Dinerstein E, Olson DM, Graham DJ, *et al.* (1995) *A Conservation Assessment of the Terrestrial Ecoregions of Latin America and the Caribbean.* Washington, DC: The World Bank.

Erwin TL (1991) An evolutionary basis for conservation strategies. *Science* 253: 750–752.

Erwin TL (1998) Evolution at the equator: Arboreal and alticolous beetles and their taxon pulses with descriptions of a new *Agra* subclade and its species (Coleoptera: Carabidae: Lebiini). *Mus. Reg. Sci. Nat. Torino* 491–510.

Fittkau EJ (1973) Artenmannigfaltigkeit amazonischer Lebensräume aus okologischer Sicht. *Amazoniana* 4(3): 321–340.

Golovatch SI, Vohland K, Hoffman RL, *et al.* (1998) Review of the Neotropical millipede genus *Pycnotropis* Carl, 1914 (Diplopoda, Polydesmida, Aphelidesmidae). *Amazoniana* 15(1/2): 67–102.

Haffer J (1996) Time's cycle and time's arrow in the history of Amazonia. In: Pavan C and de Araujo MC (eds.) *Uma Estratégia Latino-Americana para a Amazônia. Vol. I*, pp. 21–49. Brasilia: Min. do Meio Amb., dos Rec. Hídricos e da Amaz, Legal.

Hoffman RL (1990) Myriapoda 4. Polydesmida: Oxydesmidae. In: von Wermuth H and Fischer M, (eds.) *Das Tierreich,* Vol. 107. Berlin: de Gruyter.

Junk WJ (ed.) (1997) The Central Amazon flood plain. *Ecology of a pulsing system.* In *Ecological Studies,* Vol. 126: Berlin: Springer.

Klinge H (1973) Struktur und Artenreichtum des zentralamazonischen Regenwaldes. *Amazoniana* 4(3): 283–292.

Kress WJ, Heyer R, Acevedo P, *et al.* (1998) Amazonian biodiversity: Assessing conservation priorities with taxonomic data. *Biodiv. Conserv.* 7: 1577–1587.

Linsenmair KE (1990) Tropische Biodiversität: Befunde und offene Probleme. *Verh. Disch. Zool. Ges.* 83: 245–261.

Lourenço W (1994) Biogeographic patterns of tropical South American scorpions. *Stud. Neotrop. Fauna Environ.* 29(4): 219–231.

Penny ND and Arias JR (1982) *Insects of an Amazon Forest.* New York: Columbia University Press.

Pereira LA, Foddai D, and Minelli A (1997) Zoogeographical aspects of Neotropical Geophilomorpha (Chilopoda). *Ent. Scand. Suppl.* 51: 77–86.

Platnick NI (1999) Dimensions of biodiversity: Targeting mega-diverse groups. In: Cracraft J and Grifo F (eds.) *The Living Planet in Crisis: Biodiversity Science and Policy*, pp. 33–52. New York: Columbia University Press.

Prance GT (ed.) (1982) Biological diversification in the tropics. In *Proceedings of the Fifth International Symposium of the Association for Tropical Biology.* New York: Columbia University Press.

Rodríguez JP, Pearson DL, and Barrera R (1998) A test for the adequacy of bioindicator taxa: Are tiger beetles (Coleoptera: Cicindelidae) appropriate indicators for monitoring the degradation of tropical forests in Venezuela? *Biol. Conserv.* 83: 69–76.

Sturm H (1984) Zur Systematik, Biogeographie und Evolution der südamerikamschen Meinertellidae (Machiloidea, Archaeognatha, Insecta). *Z. Zool. Syst. Evol.* 22: 27–44.

Vanicek M, Adis J, and Paarmann W (1994) Untersuchungen zur Biologie dreier Laufkaferarten (Coleoptera, Carabidae, Harpalini) aus amazonischen Regenwaldem. *Andrias* 13: 161–168.

Vanzolini PE and Heyer WR (eds.) (1988) In *Proceedings of a Workshop on Neotropical Distribution Patterns.* Rio de Janeiro: Academia Brasileira de Ciências.

Vohland K and Adis J (1999) Life history of *Pycnotropis tida* (Diplopoda: Polydesmida: Aphelidesmidae) from seasonally inundated forests in Amazonia (Brazil and Peru). *Pedobiology* 43: 231–244.

Wolf HG and Adis J (1992) Genetic differentiation between populations of *Neomochilellus scandens* (Meinertellidae, Archaeognatha, Insecta) inhabiting neighbouring forests in Central Amazonia. *Verh. Naturwiss. Ver. Hamburg (NF)* 33: 5–13.

Asia, Ecosystems of

Elgene O Box and Kazue Fujiwara, University of Georgia, GA, USA, and Yokohama National University, Japan

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition, volume 1, pp. 261–291, © 2001, Elsevier Inc.

Glossary

Angiosperms Flowering vascular plants in which the seeds, which are produced from the ovules of the flowers, are enclosed in fruits developed from the matured ovaries; the largest class of vascular plants, as opposed to cryptogams (mosses, etc., which do not produce seeds) and conifers (in which the seeds are not enclosed in fruits).

Biodiversity Biological diversity, i.e., the variety of living organisms in an area, including the variety in genes, species, functional types of organisms, and ecosystems.

Biome A major terrestrial ecological community and landscape type, characterized by more or less uniform physiognomy of its potential natural vegetation and with characteristic fauna and flora, such as the tropical rainforest, the warm deserts, or the temperate grasslands.

Boreal Pertaining to the northern high latitudes (but not polar), regions which contain large continental landmasses and thus continental climates with moderate summers but long, severely cold winters (from Greek *boreas*, the north wind).

Ecosystem A community of organisms (plants, animals, and microbes) and their physical environment interacting as an ecological unit, such as a lake, a wetland, a forest, or an agricultural landscape.

Endemic Occurring only within a comparatively restricted geographic range within a specific region.

Global circulation The regular global circulation pattern of the earth's atmosphere, which generates the world's basic climate types; the system involves an Intertropical Convergence zone of low pressure near the equator, subtropical high-pressure belts near the Tropics of Cancer and Capricorn, trade winds flowing from these high-pressure belts toward the equatorial low, and westerly winds in the midlatitudes.

Monsoon system A wind system covering the eastern half of Asia in which winter cooling of the large landmass produces strong, stable high pressure, with clear skies and outward flow of cold, dry air, and summer warming produces low pressure, drawing wet air masses, with clouds and rain, inward from the adjacent oceans.

Pleistocene glaciation The expansion of ice sheets over the large northern continents, especially northern Europe and North America (excluding the northwest) during the "Ice Ages."

Potential natural vegetation The vegetation cover which would develop naturally in an area and become stable (not replaced by a subsequent stage) if all outside disturbances were eliminated.

Species richness The total number of species of an area or ecosystem.

Temperate Pertaining to the climates and landscapes of the midlatitude regions, which are seasonally warmer and cooler, with winter frost even in most coastal areas and at least partial dormancy or collapse of the vegetation.

Tropical Astronomically, the region lying between the Tropic of Cancer (23 1/2° N) and Tropic of Capricorn (23 1/2° S); more generally, involving the climates and landscapes characteristic of this region, which are essentially frost free in the lowlands and permit biological activity as long as water is available.

Vegetation The total plant cover of an area.

Zonation The tendency of climate, soil, and natural landscape types to occur in distinct latitudinal zones (e.g., tropical, subtropical, and temperate), as generated by the global atmospheric circulation system.

Introduction

Asia the world's largest continent, extending from the Mediterranean Sea in the west to the Japanese Archipelago in the east, and from polar landscapes in northern Siberia to hot deserts in the southwest and tropical regions south of the Himalaya and China. Tropical Asia includes perhumid equatorial rainforest climates as well as seasonally wet and dry tropical savannas, raingreen woodlands, and monsoon forests, all extending to some degree from India to the Philippines and East Indies. Dry regions of Asia include an eastward extension of the hot, subtropical Sahara desert into southwestern Asia, large expanses of higher latitude interior deserts with cold winters, and also a small 'mediterranean' region and an east–west strip of temperate grassland (steppe) which spans almost the entire continent. The main area of temperate forests is in East Asia, with deciduous and mixed forests in the north, rich evergreen "laurel" forests in the south, and extensive areas of mostly secondary pine forests throughout. Mountain areas of Asia have many endemic species, especially relict conifers, and the mountains of East Asia represent the only truly large mountainous area in the warm-temperate and humid subtropical zones of either hemisphere. Northern Asia was less glaciated than Europe or North America and contains enormous expanses of relatively diverse Siberian conifer forests, polar and upland tundra, and Lake Baikal (the world's deepest). Wetlands in Asia include mangroves, coastal strand forests, salt marshes and estuaries, and terrestrial swamp forests, marshes, and bogs, most of which are highly productive and represent critical habitat for terrestrial as well as aquatic

animals. Despite long histories of human habitation and landscape alteration, Asia retains the highest biodiversity of any continent, due not only to its size but also to its climatic and topographic complexity and its complex geological and evolutionary history. This biodiversity is threatened in Asia as elsewhere by human overpopulation and overdevelopment.

Asia the Region

Asia is traditionally separated from Europe, with which it shares the world's single largest land-mass, Eurasia (**Figure 1**). Even by this cultural definition, however, Asia remains the world's largest continent and extends almost halfway around the world, from a small western coastline on the Mediterranean Sea (about 25° E in Turkey) eastward to the Bering Strait (170° W) separating it from Alaska. Asia also spans almost the entire latitudinal range of the Northern Hemisphere, from a northern coastline in the Arctic zone to islands lying slightly south of the equator in Indonesia. Asian superlatives include the world's

- highest mountain (Mt. Everest, 8848 m) and mountain range (Himalaya);
- largest highland plateau (Tibet);
- lowest terrestrial elevation (Dead Sea, − 400 m);
- largest (Caspian Sea, 143,200 km^2) and deepest (Lake Baikal, 1620 m) lakes; and
- coldest, most continental region except for Antarctica (northeastern Siberia), with mean monthly temperature ranging from 20 to − 60 °C and extremes to − 80 °C.



Figure 1 Location and main physiographic features of Asia. Asia extends almost halfway around the world, from Turkey and Arabia in the west to the Bering Strait in the east. It also extends from the Arctic Ocean in the north to slightly south of the equator in Indonesia. Middle Asia extends across the Turanian Basin from the Caspian Sea to the Tien-Shan and Altai mountains and is largely steppe and desert. Central Asia includes the high, cold Tibetan Plateau and the desert basins to its north, on into the steppes of Mongolia. Siberia (Asiatic Russia) is the mostly forested region north of the east–west mountains and Middle Asian deserts. Other regions and the geological development of Asia are described in the text.

Also found in Asia are

- the world's third (Borneo) and fifth (Sumatra) largest islands;
- four of the world's seven rivers over 5000 km long (Yang-tze, Ob'-Irtysh, Yenisey-Angara, and Huang-He);
- one of the world's three regions with essentially no rainfall (Tarim Basin); and
- two of the world's three stations with more than 10 m of average annual rainfall (windward slopes of easternmost India at Cherrapunji and Yakushima Island of Japan).

In addition, Asia has been home to some of the oldest human cultures and to areas of very long, continuous human habitation, especially in the Middle East, India, and Southeast and East Asia.

Asia represents the largest part of Laurasia, the northern half of the supercontinent Pangaea, which existed in the late Paleozoic and broke up during the Mesozoic (**Table 1**; Raven and Axelrod, 1974). Eurasia began to separate from North America around 180 million years ago, but contact was reestablished in the Cretaceous when Siberia approached Alaska, forming Beringia. At about this time, southwestern Asia also began to be influenced by northward-moving Africa, a portion of Gondwana, the southern part of Pangaea.

The present-day geologic and physiographic structure of Asia, however, is dominated by the Himalayan Mountains and Tibetan Plateau, which arose when India, another piece of former Gondwana, collided with Asia in the Eocene. In addition to the Himalaya, a mountain node was created at its western end, from which other major mountain ranges radiate in various directions. The most important of these is the Tien Shan-Altai system, which extends to the north–northeast into Mongolia, dividing interior temperate Asia into

- a western part, called Middle Asia in Russian literature (essentially Turkestan and the Turanian Basin), influenced far inland by westerly winds from the Mediterranean; and

Table 1 Some major events in asia's geological and evolutionary history[a]

		Geological and climatic events	*Biotic events*
Paleozoic Era (>225 million BP)			
Silurian	425–400 m	Climate similar to present?	First vascular and land plants
Carboniferous	345–280 m	Warm and humid	Forests of ferns, horsetails, etc.
Permian	from 280 m	Pangaea; glaciation at S pole	First ginkgos and mosses
Mesozoic Era (225–65 million BP)			
Triassic	225–185 m BP	Breakup of Pangaea begins; climate cool and humid	Ginkgos and cycads dominant
Jurassic	185–135 m	Breakup of Pangaea: W. Gondwana	Conifers and cycads dominant
Cretaceous	135–65 m	"Greenhouse" climate, high sea level	Angiosperms dominant
		Major separations	
		Gondwana (130–100 m BP)	
		Eurasia and N. America (180 m, initially)	Modern plant orders exist; some modern families and genera exist
		Europe and N. America (80 m; not north)	
		Major contacts	
		SW Asia and Africa (100? m)	
		Siberia and Alaska (70 m)	
Tertiary Period (of Cenozoic Era)			
Paleocene	65 m BP	Africa and Europe separate temporarily	
Eocene	60–38 m	Final separation of Europe and N America; Beringia becomes warmer	Most modern families, many genera exist
		Himalaya and central Asia plateaus form as India collides with Asia	Madro-Tertiary geoflora First primates
Oligocene	38–26 m	Alpine orogeny stresses Middle Asian mountains	
Miocene	26–7 m	Australia approaches SE Asia; islands form Global climates become cooler and drier	
Pliocene	7–2 m	Arabia splits from Africa and contacts Asia, forcing Turkey westward	Most modern plant taxa; herbaceous plants abundant
Quaternary Period (of Cenozoic Era)			
Pleistocene	2–0.01 m BP	Glacial and interglacial periods;	N–S migration of vegetation zones, with loss
	18 k	N Europe glaciated, Siberia less	of taxa unable to cross barriers
	14–12 k	Last glacial maximum: sea 150 m lower, land bridges and wider coastal plains	Tropical areas drier
		Warming and glacial retreat	
Holocene	from 10 k	Continued warming	Temperate and high-latitide soils and biomes redevelop
	8–5 k	Hypsithermal period (warmest)	
	5 k	Sea reaches present level; cooling	Paludiflcation in Siberia

[a]Times are in millions (m) or thousands (k) of years before present (BP).

Sources: Pearson (1995); Bridges (1990), and Raven PH and Axelrod DI (1974) Angiosperm biogeography and past continental movements. *Ann. Missouri Bot. Garden* 61: 539–673.

- an eastern part, called Central Asia (essentially the Tibetan Plateau and northwestern China, plus Mongolia) dominated by the Asian monsoon system.

Other radial ranges from the central node include the Hindu Kush mountains to the southwest, which separate Afghanistan from Pakistan; the Paropamisus and Kopet Dagh further west, which separate Afghanistan and Iran to the south from Turkestan to the north; and the Kunlun and Karakoram ranges, which run east–west across the Tibetan Plateau. South of Tibet, of course, lies the Deccan Plateau of peninsular India, with its Eastern and higher Western Ghats ranges. Southeast of Tibet, the Yunnan-Guizhou Plateau (southwestern China) represents an upland transition to still mountainous Southeast Asia, with its Dawna and Arakan ranges (Burma) and Annamese cordillera.

In southwestern Asia, the Iranian Plateau (extending into Afghanistan), the Anatolian Plateau of central Turkey, and the Arabian Plateau (peninsula) have all been affected to some extent by the interaction between Africa and Asia. More linear mountain ranges in Asia include the Urals, the Hinggan-Ling and Changbai-Shan in northeastern China, and the Sikhote Alin, Kamchatkan, and other ranges of eastern Siberia. Major interior basins include the Turanian Basin east of the Caspian Sea, the West Siberian Lowland (the world's largest marshland), and the Tarim Basin of Central Asia. Other lowlands include the densely populated East China Plain and many river valleys of Asia: Ganga, Indus, Tigris–Euphrates, Brahmaputra, Irrawaddy, Yangtze, and Huang-He. Less populated lowlands extend along the Siberian coast (with the Taimyr peninsula) and the valleys of the Amur and north-flowing Lena, Yenisey-Angara, and Ob'-Irtysh rivers. The remaining structure of Asia is provided by the volcanic arcs forming the Kuril, Japanese, Okinawan, Philippine, and Indonesian archipelagos, along with continental islands such as Sri Lanka, Hainan, Taiwan, and the smaller Siberian Arctic islands and Andaman-Nicobar chain. All of these areas represent regions of localized environmental conditions and ecosystems, with more or less individual characteristics and diversity.

Climatic zonation in general and major regions in Asia are summarized in **Table 2**. The single most important climatic feature of Asia is the monsoon system, which affects an area from Japan and northeastern China in the east, through China and Southeast Asia, to India and Sri Lanka in the south. The cooling of the vast mid-latitude land-mass in the autumn and winter produces a very large, stable center of high atmospheric pressure over Mongolia and northern China, which feeds cold, dry air outward over the rest of East, Southeast, and South Asia throughout the winter. As a result, and aided perhaps by the many east–west mountain ranges, mean winter temperatures are generally low but extremes are not far below the means. In summer, continental warming causes rising air, which draws wet air masses in off the adjacent Pacific and Indian Oceans, greatly accentuating the rainier season which, due to global atmospheric circulation, would occur in summer over most of this area anyway. The "summer monsoon" arrives as two almost stationary rain fronts which, over the course of about one month, penetrate from Kerala across the rest of India and from Japan across Korea into eastern China. Although the monsoon quickly brings warm rains to south Asia, it brings a cooler, mistier rainy season as it crosses Japan into Korea and China. The cold, dry winter is much more reliable than the summer rains. Heavy winter snow can delay the warming of the

Table 2 Global climatic zonation in lowlands, with regions in asia

Zone	*Temperature extremes*	*Significance*	*Asian examples*
Tropical	No frost or other "cold" temperatures ever	Sensitivity of many tropical plants to non-freezing cold	South Asia and East Indies
Subtropical	Occasional frost or near frost, not every year and not below about $-1\,^{\circ}C$	Frost sensitivity of tropical evergreen and most other tropical plants	Southern China, Taiwan, south Asian foothills, northern Arabian Peninsula
Warm-temperate	Light to moderate frost, every year or nearly so; absolute minima not $< -15\,^{\circ}C$	Annually leaf-changing broad-leaved evergreens tolerate, (sub)tropical evergreens may not	Eastern, inland SE and SW China, southern Japan; Mediterranean region
Temperate	Significant frost every year; occasional temperatures below about $-15\,^{\circ}C$	Coldness tolerance limit for evergreen broad-leaved plants (e.g., internal ice formation)	Northern China and Japan, Korea, interior Asia from Mongolia to Turkey
Cool-temperate	Moderate to significant frost every year, plus cool summer; minima can be $> -15\,^{\circ}C$ if oceanic	Growing-season warmth marginal for many typical temperate plants, including deciduous	Hokkaido (minima $< -15\,^{\circ}C$)
Boreal	Cool, short summer and long severe winter; absolute minima $< -15\,^{\circ}C$, perhaps $< -40\,^{\circ}C$	Growing season insufficient for most deciduous trees (exceptions: larch, birch, etc.)	Most of Siberia, northern Manchuria
Polar	Short summers below $10\,^{\circ}C$ and long severe winters below $0\,^{\circ}C$; extremes $< -40\,^{\circ}C$ unless oceanic	Growing season too cold for wood-producing enzymes; no trees or large shrubs	North-coastal Siberia, islands of Arctic Ocean

landmass to such an extent that the summer monsoon "fails" in places such as India, with disastrous effects on crops.

The Philippines and East Indies are less affected by the monsoon and more by the seasonal shift in global circulation. Near the equator, warm perhumid conditions continue throughout most of the year, with only a short dry season if any at all. The length and degree of the dry season increase away from the equator, reaching several months in the northern Philippines but also in some eastern parts of Indonesia (Lesser Sunda Islands as well as eastern Java and southernmost Borneo and Sulawesi).

Southwestern Asia is not affected by the monsoon but rather by the pervasive effect of the subtropical high-pressure belt. This belt of descending dry air migrates north–south seasonally, creating "Mediterranean" climates with dry summers in Turkey and throughout the Mediterranean borderlands. To the south, however; subtropical high pressure remains the dominant influence throughout the year, causing the desert belt across North Africa, the Arabian Peninsula, and as far east as the Thar desert in Pakistan and western India.

Northern Asia (Siberia) is dominated by its high latitude and large land area, resulting in ultracontinental climates with short summers and long severe winters rivaling the northern polar region in degree of cold.

Biodiversity

Angiosperms are thought to have originated in South–east Asia, and indeed Asia is thought to contain many other foci of more recent biotic radiation as well. During the Pleistocene, when many species were lost from heavily glaciated Europe and North America, glaciation in Asia was less extensive, especially in the east. The Scandinavian ice sheet extended east to the Urals, but Siberia remained largely ice-free except in the northwest and in the mountainous northeast. A large Pleistocene refuge for temperate Asian plants was also readily available in southern China. These climatic (and related sea-level) changes, plus the physiographic complexity of Asia, all contributed to the development and maintenance of high biotic diversity in Asia, both in and within genera. Among vascular plants, for example, both Siberia and temperate east Asia generally show higher diversity than comparable latitudes in Europe and North America.

The most complete compilation of recent estimates of both species richness and the protection status of taxa and ecosystems is titled *Global Biodiversity* (Groombridge, 1992), from which summaries will appear throughout this article. Species richness and endemism of plants in the different countries and regions of Asia are summarized in **Table 3**. China and Malesia are thought to have about 30,000 plant species each (most not in common), India may have 15,000 species, and tropical and subtropical Asia may have a total of at least 50,000 species. By comparison, Latin America may have as many as 85,000 species, but this includes the region from Mexico to temperate southern South America. Within Asia, China has the largest number of plant species at 30,000, followed by Indonesia (20,000), India (15,000), Malaysia (12,000) and Thailand (12,000). A special feature of East Asia is its large number of endemic plant families, many of which are mono-specific, as shown in **Table 4**. Species richness is lowest in the dry areas of the Arabian Peninsula and mainland southwestern Asia, but rates of endemism are high in some of these drier areas. Inventories are incomplete in all areas, with only small Brunei claiming more than 20% completion. No inventory statistics at all were available for Southeast Asia, except for Thailand and Myanmar.

Estimates of species richness and endemism in four major groups of animals, but without indication of inventory status, are provided in **Table 5** (from Groombridge, 1992). Species richness is generally lowest in the drier areas, but the number of reptiles is still high. Numbers of birds are also fairly high in mainland southwestern Asia. Other sources of biodiversity data include the United Nations Environment Programme *Global Biodiversity Assessment* (Heywood, 1995) and the various publications of the World Resources Institute.

Regionalization and Biomes of Asia

Global regionalization of climatic types, as related to causal mechanisms (latitude and global circulation) and natural ecosystems, has been portrayed perhaps best by the widely used global system of Heinrich Walter (1985). This and the major large-area landscape types, called biomes (e.g., tropical rain forest and temperate grassland), provide a geographic framework for regionalizing the ecosystems of Asia. An attempt is made in **Table 6** to juxtapose the main biome types, the climate type in which they occur, the corresponding Bailey ecoregion class, and the corresponding International Union for the Conservation of Nature biogeographic provinces in Asia. Note that the "Savanna" ecoregion also includes the region of tropical moist and dry deciduous forests ("monsoon forests") that stretches across South and Southeast Asia and which corresponds to the larger tropical wet–dry climatic regions of Africa and South America, with their areas of tropical deciduous forest, woodland, and savanna. Tropical deciduous forests are not recognized in some newer classifications because they have been so completely converted into savanna. Areas of tropical deciduous forest still remain in Asia, however, especially in the Ghats Mountains of India and from the Deccan Plateau to interior Southeast Asia. Note also that one must be careful with Russian and Chinese terminologies which use terms such as "Hot/Warm Continental" and "subtropical" for temperate forest regions which have cold to severely cold winters.

Using the synthesis presented in **Table 6**, Asia can be divided into four main regions of bioclimates and associated ecosystems:

1. Tropical Asia, including the tropical rain forest and monsoon forest regions, the tropical islands, and the various elevational belts of tropical mountains
2. the dry region of subtropical southwestern and interior Middle and Central Asia, including the subtropical and interior deserts, the temperate grasslands, the Mediterranean-type ecosystems in the southwest, and the drier mountain ranges
3. East Asia, the monsoonal but extra-tropical region of subtropical and warm-temperate evergreen broad-leaved forests ("laurel forests") and of temperate deciduous forests,

Table 3 Plant species richness and endemism in Asia[a]

Location	*Angiosperms*	*Gymnosperms*	*Ferns*	*No. of Endemics*	*Percentage endemism*[b]	*Date*[c]	*Status*[d]
Arabian Peninsula							
Bahrain	195	1	1	0	0	1991	1 c
Kuwait	234	1	1	0	0	1991	1 c
Oman	1018	3	14	74	7.1	1991	1 c
Qatar	220	1	0	0	0	1991	1 c
Saudi Arabia	1729	8	22	34	1.9	1991	2 c
United Arab Emirates	340	2	5	0	—	1991	l c
Yemen (P.D.R.)	1373	3	41	58	4.1	1991	2 c
Yemen (Arab Rep.)	959	1	14	77	7.9	1991	1 c
Southwest Asia (mainland)							
Afghanistan	3500	—	—	—	[30–35]	1989–1991	2 e2
Cyprus	1650	12	20	88	5.2	1977–1985	l c
Iran	6500	33	—	—	[30–35]	1989–1991	1 e2
Iraq	2914	7	16	190	6.5	1966–1986	1 c
Israel	2294	8	15	155	6.7	1982–1984	l c
Jordan	2200	6	6	—	—	1982–1985	2 c
Lebanon	2000	12	40	—	[10]	1984–1991	2 e3
Syria	2000	12	40	—	[10]	1984–1991	2 e3
Turkey	8472	22	85	2651	30.9	1988	3 e2
South Asia							
Bangladesh	5000	—	—	—	—	1972	2 e2
Bhutan	5446	22	—	50–100	1.4	1991	3 e1
British Indian Ocean Territory	100	1	—	0	0	1971	1 e1
India	15,000	—	1000	5000	31.3	1983–1984	2 e2
Maldives	260	2	15	5	1.8	1983	1 c
Nepal	6500	23	450	315	4.5	1978–1982	2 c
Pakistan	4917	21	—	372	7.5	1986	2 e1
Sri Lanka	2900	—	314	900	28.0	1982–1983	2 c
Southeast Asia (mainland)							
Cambodia	—	—	—	—	—	—	—
Laos	—	—	—	—	—	—	—
Myanmar	7000	—	—	1071	15.3	1961	4 e2
Thailand	12,000	25	600	—	—	1979–1985	3 e2
Vietnam	—	—	—	—	—	—	—
Insular and Peninsular Southeast Asia							
Brunei	3000	28	—	7	0.2	1990	5 e2
Indonesia	20,000	—	2500	15,000	66.7	1991	4 e3
Malaysia	12,000	—	500	—	—	1991	3 e3
Philippines	8000	31	900	3500	39.3	1982–1991	3 e2
Singapore	2000	2	166	1	0.1	1989–1991	1 e1
East Asia							
China	30,000	200	2000	18,000	55.9	1991	3 e2
Hong Kong	1800	4	180	25	1.3	1978–1991	2 e2
Japan	4700	42	630	2000	37.2	1987	1 c
Korea (North)	2898[e]	—	—	107	14.0[e]	1976–1983	— c
Korea (South)	2898[e]	—	—	224	14.0[e]	1976–1983	— c
Mongolia	2272	—	—	229	10.1[a]	1984	1 c
Taiwan	2983	20	565	—	[25]	1982–1991	1 c

[a]Compiled by the World Conservation Monitoring Centre (WCMC). The last two columns refer to the date and status of country inventories, using codes from the original source.
[b]Percentage endemism is calculated from the data, unless in square brackets.
[c]Date is the date of the information provided to the WCMC.
[d]Status includes percentage completion (first number) and method (counted/estimated): 1, <5%; 2, 5–10%; 3, 10–15%; 4, 15–20%; 5, >20%; c, counted; el, approximate count; e2, extrapolation, e3, estimated based on any available information and comparable floras.
[e]Data for North and South Korea are combined; no data are available for combined Yemen.
Source: The data are regrouped by region from those presented in the Global Biodiversity report (Groombridge, 1992, p. 80).

Table 4 Endemic families of east Asia[a]

Polygeneric families	
Helwingiaceae	(4–5/4–6)
Carlemanniaceae	(2/4, E. Himalaya, Assam, N. Myanmar, N. Indo-China to Yunnan; 1 disjunct in Sumatra)
Podoaceae	(2/3, central Yunnan to E. Himalaya and N. Thailand)
Monogeneric families	
Hostaceae	(10–40)
Stachyuraceae	(10)
Cephalotaxaceae	(9)
Aucubaceae	(7–11, to E. Himalaya and N. Myanmar, 1 to Russian Far East)
Torricelliaceae	(2, E. Himalaya, N. Myanmar to central China)
Triplostegiaceae	(2 1 disjunct in Sulawesi and New Guinea)
Euptelaceae	(2)
Trapellaceae	(1–2, extending to Russian Far East)
Monospecific families	
Nandinaceae	
Circaeasteraceae	(not Japan, excluding *Kingdonia)*
Tetracentraceae	(not Japan)
Cercidiphyllaceae	(not Himalaya)
Euryalaceae	(not Himalaya but to India and Bangladesh)
Plagiopteraceae	(disjunct in India, Bangladesh, lower Myanmar, Thailand, Guangxi)
Pottingeriaceae	(Assam, N. Myanmar, NW Thailand to Yunnan)
Trochodendraceae	(not mainland China)
Bretschneideraceae	(China to N. Laos and N. Thailand)
Dipentodontaceae	(central China to NE India, Himalaya)
Sargentodoxaceae	(only China to N. Vietnam and N. Laos)
Rhoiptelaceae	(only SW China to N. Vietnam)
Sladentiaceae	(only SW China and adjacent Myanmar, Thailand)
Acanthochlamydaceae	(only in Hengduan mtn. to E. Tibet)
Davidiaceae	(only in China)
Ginkgoaceae	(only in China)
Eucommiaceae	(only mainland China)
Glaucidiaceae	(only in Japan)
Pteridophyllaceae	(only in Japan)
Sciadopityaceae	(only in Japan)

[a]Numbers in parentheses represent number of genera followed by number of species or number of species only (if monogeneric).
Source: Reproduced from Wu and Wu, 1996.

with the more humid uplands and mountain ranges in these areas

4. Boreal and polar Asia, including the boreal forests of Siberia, northeasternmost China, and northern Mongolia and the tundra landscapes across the northern Siberian coastal plain and in the mountain areas of eastern Siberia.

Treatment of the various ecosystems of Asia will follow this framework, with an emphasis on vegetation since it provides the basic physical and trophic structure of ecosystems. Wetlands, coastal ecosystems, and some artificial ecosystems will be treated separately.

Estimates of vegetation cover type, percentage cover, and carbon storage, by country and region, are summarized in **Table 7** (Groombridge, 1992). The most extensive forest cover is in eastern to southern Asia and in Siberia, with non-forest landscapes dominating southwestern Asia. These data probably overestimate the extent of remaining forest.

Allotted space does not permit a detailed description of regional vegetation composition or even a more complete bibliography. Some general treatments, from which one can find references to more regional descriptions, include the following:

Global: Archibold (1995) and Walter (1985)
Tropical: Champion and Seth (1968), Mani (1974), and Whitmore (1984)
Dry region: Walter and Box (1983) and Zohary (1973)
East Asia: Missouri Botanical Garden (1983), Numata (1974), and Wang (1961)
Boreal/polar: Walter (1974)
Mountains: Chen *et al.* (1986), Troll (1972), and Walter (1974).

For animals, one might start with the extraordinary syntheses of Schaller (1998).

Tropical Asia

Tropical Asia is the region of mainly low-lying topography and warm, at least seasonally humid climates which extends from India to the Philippines and East Indies (i.e., excluding all of dry southwestern Asia). Much of the region is coastal lowland, and the highest point is on an island, at Mount Kinabalu (4110 m) in northern Borneo. The region can be divided climatically into two main zones:

1. The perhumid equatorial zone, covering most of the East Indies and southern Philippines, the Malay Peninsula, and extending along the windward southwest-facing coastlines of mainland southern Asia to southern Kerala and Sri Lanka.
2. The seasonally wet–dry zone to the north, mainly across interior southern Asia from India to Vietnam and the northern Philippines, but also including the eastern Sunda Islands (6–11° S) and some other island areas, including rain shadows such as eastern and northern Sri Lanka.

The natural biome of the equatorial zone is tropical rain forest and other humid tropical evergreen forests, whereas the wet–dry zone contains natural landscapes ranging from moist and dry deciduous forests to savanna and small areas of dry evergreen forest mainly in Thailand. Other ecosystems, such as swamp forests, kerang (sclerophyll scrub), and viny limestone forests, occur on particular substrates. These main characteristic vegetation types of tropical Asia and their topographic or substrate affinities are summarized in **Table 8**. Characteristic vegetation belts in mountains are summarized in **Table 9** (Whitmore, 1984).

Humid Tropical Forest

Tropical rain forest is tall, multi-layered, evergreen broad-leaved forest characterized by many epiphytes, climbing plants, and a tremendous richness of tree species with dark green, thin-coriaceous leaves, many with attenuated "drip tips" for better drainage but otherwise all looking very similar. The air is constantly humid, but the canopy is exposed each day to high solar radiation, resulting in a drier canopy microclimate and a more perhumid microclimate in the forest understory. Many tree

Table 5 Species richness and endemism of higher animals in Asia

Region	Mammals		Birds		Reptiles		Amphibians	
	Species known	Endemic species	Species known	Endemic species	Species known	Endemic species	Species known	Endemic species
Mediterranean–Middle Asia								
Afghanistan	123	0	456	0	103	—	6	1
Cyprus	21	0	80	2	23	1	4	0
Iran	140	4	—	1	164	3	11	5
Iraq	81	1	145	1	81	—	6	0
Israel	—	2	169	0	—	—	—	0
Jordan	—	0	132	0	—	—	—	0
Lebanon	52	0	124	0	—	—	—	0
Syria	—	0	165	0	—	—	—	0
Turkey	116	0	284	0	102	5	18	2
Arabian Peninsula								
Bahrain	—	0	—	0	25	0	—	0
Kuwait	—	0	27	0	29	0	2	0
Oman	46	3	—	0	64	11	—	0
Qatar	—	0	—	0	17	0	—	0
Saudi Arabia	—	1	59	0	84	5	—	0
United Arab Emirates	0	0	—	0	37	1	—	0
Yemen	—	1	—	8	77	25	—	1
Tropical Asia								
Bangladesh	109	0	354	0	119	—	19	0
Bhutan	109	0	448	0	19	—	24	0
British Indian Ocean Territory	—	0	—	0	—	—	—	0
Brunei	155	0	359	0	44	—	76	0
Cambodia	117	0	305	0	82	—	28	0
India	317	38	969	69	389	156	206	110
Indonesia	515	165	1519	258	511	150	270	100
Laos	173	0	481	1	66	—	37	1
Malaysia	264	14	501	4	268	—	158	39
Maldives	—	0	24	0	—	—	—	0
Myanmar	300	8	867	4	203	29	75	9
Nepal	167	1	629	1	80	—	36	7
Pakistan	151	3	476	0	143	22	17	2
Philippines	166	90	395	172	193	131	63	44
Singapore	57	1	118	0	—	—	—	0
Sri Lanka	86	12	221	20	144	75	39	19
Thailand	251	5	616	2	298	39	107	13
Vietnam	273	5	638	12	180	—	80	26
East Asia								
China	394	62	1100	63	282	—	190	131
Hong Kong	38	0	107	0	61	0	23	2
Japan	90	29	>250	20	63	28	52	35
Korea (North)	—	0	—	0	19	1	13	0
Korea (South)	49	0	—	0	18	—	13	1
Mongolia	—	6	—	0	—	—	—	0
Taiwan	62	13	160	15	67	20	26	6

Source: As in Table 3, the data are regrouped by region from those presented in the *Global Biodiversity* report (Groombridge, 1992, p. 139).

species respond to this difference with somewhat smaller, thicker canopy leaves and larger, thinner (and often darker) understory leaves, all on the same tree. In true tropical rain forest, the only seasonality is the difference between day and night. As a result, the main phenological functions of the trees tend to occur simultaneously. A given tree may be seen to be flowering, flushing new leaves, dropping individual or small bunches of old leaves, and carrying ripening or ripe fruit simultaneously on different branches.

As one proceeds away from the equator, subtle climatic cues begin to appear: seasonal periods of less precipitation and slight seasonal temperature variations. The result is a "tropical seasonal evergreen forest" which is still tall and very diverse but in which phenological functions become synchronized and more sensitive) plant types become somewhat less abundant, such as epiphytes with only aerial roots. Further from the equator, as the tropical dry season begins to become more pronounced, the forest becomes a "tropical

semi-evergreen forest" containing some deciduous trees and still fewer epiphytes (Walter, 1985). In most tropical trees, deciduousness appears to be a facultative characteristic. This is demonstrated by two facts:

1. Many tropical tree species are evergreen in one part of their range but lose their leaves in a drier part.
2. Many individual tropical trees have been observed to keep almost all their leaves in wetter years and lose most or all of their leaves in drier years.

Temperate tree species, on the other hand, often die within a few years if planted in the tropics, apparently due to lack of the seasonal cues necessary for vernalization.

Table 6 Biomes, ecoregions, and biogeographic provinces of Asia

Biome region[a]	*Climate*[b]	*Ecoregions*[c]	*Biogeographic province*[d]
Tropical rain forest	I	Tropical rain forest	Malabar rain forest Ceylonese rain forest Bengalian rain forest Burman rain forest Indochinese rain forest South Chinese rain forest Malayan rain forest
Tropical deciduous forest (moist, dry types)	II	Savanna	Indus–Ganges monsoon forest Burman monsoon forest Thailand monsoon forest Mahanadian monsoon forest Coromandel monsoon forest Ceylonese monsoon forest Deccan thorn forest
Tropical islands		—	Laccadives Islands Maldives and Chagos Islands Cocos–Keeling and Christmas Islands Andaman and Nicobar Islands Sumatra Java Lesser Sunda Islands Celebes Borneo Philippines
Subtropical desert	III	Tropical/subtropical desert	Arabian desert Thar desert
Mediterranean scrub	IV	Mediterranean	Mediterranean sclerophyll
Laurel forest (evergreen broad-leaved)	Ve	Humid-subtropical	Chinese subtropical forest Japanese evergreen forest
Subtropical islands			Taiwan Ryukyu Islands
Summergreen forest (temperate deciduous)	VI	Hot/warm continental	Manchu–Japanese mixed forest Oriental deciduous forest West Anatolian temperate forest Kamchatkan
Temperate grassland	VII	Temperate steppe	Pontic steppe Mongolian–Manchurian steppe
Temperate desert	VIIa	Temperate desert	Anatolian–Iranian desert Iranian (subalpine) desert Turanian desert Takla Makan–Gobi desert Tibetan desert
Boreal forest	VIII	Subarctic	West Eurasian taiga East Siberian taiga
Tundra	IX	Tundra	Low Arctic tundra High Arctic tundra Arctic cold desert

(*Continued*)

Table 6 Continued

Biome region[a]	*Climate*[b]	*Ecoregions*[c]	*Biogeographic province*[d]
Multi-zonal mountains	X	Mountains	Caucasian–Iranian highlands Altai highlands Pamir–Tien Shan highlands Hindu Kush highlands Himalayan highlands Sichuan–Yunnan highlands
Lakes systems	—	—	Aral Sea Lake Baikal

[a]Biome regions are denoted by names commonly used in regional and global treatments of terrestrial vegetation and biomes (Reproduced from Archibold OW (1995) *Ecology of World Vegetation*. London: Chapman & Hall, and Walter (1985) *Vegetation of the Earth and Ecological Systems of the Geobiosphere*, 3rd ed. New York: Spinger-Verlag, and the "Ecosystems of the World" book series).

[b]Climate types are those of Walter (1985): I, equatorial; II, tropical summer-rain; III, subtropical arid; IV, mediterranean; Ve, warm-temperate (humid east sides); Vm, marine west-coast (perhumid); VI, typical temperate (humid) VII, temperate continental (subhumid); VIIa, temperate arid; VIII, boreal; IX, polar; X, highlands (unclassified).

[c]Ecoregions are from Bailey RG (1996) *Ecosystem Geography*. New York: Springer. See also the world map from 1989.

[d]Biogeographic provinces are from the classification by Udvardy MDF (1975) IUCN Occasional Paper No.18. Morges, Switzerland: International Union for the Conservation of Nature and Natural Resources.

Tropical rain forest has its largest extent in Asia from peninsular Malaysia through the East Indies and southern Philippines. It can also be found, however, along the mainland coast from peninsular Thailand and Myanmar to Bangladesh, in Kerala and southwestern Sri Lanka, and northward through the mountains of Southeast Asia, Myanmar, and Assam to southernmost China (Yunnan) and the foothills of the eastern Himalaya (Whitmore, 1984). In these more northerly areas, extending beyond the Tropic of Cancer, the rain forest takes on a more subtropical character but maintains high diversity and its multi-layered structure, with abundant epiphytes and lianas.

Tropical rain forest in general is composed of many tree species from families such as the Lauraceae, Rubiaceae, Moraceae, Meliaceae, and Leguminosae, the latter two usually with compound leaves and somewhat smaller leaflets than the typical rain forest leaf. Tropical rain forest reaches its greatest species richness in Southeast Asia, but here another family, the Dipterocarpaceae, provides more of the rain forest species than does any other family elsewhere. Some of the most important tree genera in Malaysia, Indonesia, and adjacent mainland areas include *Dipterocarpus*, *Shorea*, *Hopea*, *Dryobalanops*, and *Vatica* (Dipterocarpaceae); *Mitragyna* and other Rubiaceae; trees with large buttress roots such as *Heritiera* (Sterculiaceae); *Aglaia* and *Dysoxylum* (Meliaceae); *Mangifera* and other anacards; *Ficus* and *Artocarpus* (Moraceae); other trees with large edible fruits such as *Durio* (Bombacaceae); *Koompassia* and other legumes; and other families such as Dilleniaceae, Flacourtiaceae, and Thymeliaceae. Most genera are mainly tropical. The higher understories contain trees of the same physiognomy but also various palms (e.g., *Licuala*), even climbing palms (rattans), and other climbers such as stranglers (especially *Ficus* spp.). The ground layer is largely composed of frutescent but entirely herbaceous monocots such as *Heliconia* and *Alpinia* (both Zingiberidae). Tall bamboos, of genera such as *Calamus* and *Phyllostachys*, may form colonial stands where the forest has been disturbed frequently and opened, especially in mountains. Secondary vegetation may also involve larger-leaved, deciduous, fast-growing arborescents from the characteristic pioneer genera *Macaranga* and *Mallotus* (analogous to *Cecropia* in the tropical Americas). Some other characteristic secondary genera include *Melastoma*, *Glochidion* (Euphorbiaceae), and *Trema* (Ulmaceae).

Birds, amphibians such as tree frogs, reptiles such as tree-dwelling lizards, and countless insects are the main groups of animals in tropical rain forests. Mammals are less numerous but include both larger ground-dwellers such as rodents and flying mammals (i.e., bats). Most animals are probably insectivorous, although many are also mixed feeders, including various mammals which range between the canopy and the ground. Some larger ground animals are herbivores, eating mainly plant roots and fallen fruit. Ants are an especially interesting and important group for their role in dispersing plant diaspores and in decomposition and nutrient cycling.

Monsoon Forest (Moist Deciduous Forest)

Monsoon forest is not always a well-defined term but is used most commonly for the tall tropical "moist deciduous forest" with fairly large, soft (malacophyllous) deciduous leaves which occurs in areas with essentially perhumid rain forest conditions during a wet season which is longer than the dry season. Since the "rain-green" leaves wilt when soil water is depleted (about a month after the end of the rainy season), there are generally no bright colors in the tropical "autumn" (with the exception of *Excoecaria agallocha* in mangroves). In the tropical "spring," however, many trees bloom brightly in the month before the rainy season, cued by increasing temperatures, and the new leaves of many species are also bright reddish before enlarging and receiving their full dose of chlorophyll. Moist deciduous forest is not as species-rich as the rain forest but may still reach canopy heights of 30 m and have more species than temperate forests.

Moist deciduous forest occurs in the Western Ghats (Kerala) and in those parts of lowland and mountainous eastern India, Myanmar, and Southeast Asia which have relatively short dry seasons (Champion and Seth, 1968). In the Western

Table 7 Estimates of vegetation type and percentage cover in Asia

Region	*Coastal aquatic (%)*	*Major wetlands (%)*	*Desert/ semidesert (%)*	*Polar/alpine (%)*	*Grass/shrub (%)*	*Crop and settlements (%)*	*Interrupted woods (%)*	*Major forests (%)*	*Carbon (kg/m^2)*
World	4	2	13	12	20	11	17	22	3.1
Asia	4	1	16	9	24	17	10	18	2.6
Middle East									
Cyprus						100			0.8
Iran	1		30		41	7	17	4	1.4
Iraq			30		34	33	2		1.0
Israel					40	30	30		1.7
Jordan			44		49	7			0.6
Lebanon						25	75		2.5
Syria			18		41	36	3	1	1.0
Turkey	3		9		37	18	25	7	1.9
Arabian Peninsula									
Kuwait	9		91						0.3
Oman	13		46		34	7			0.5
Qatar					100				0.9
Saudi Arabia	1		62		33	2	1	1	0.5
United Arab Emirates	3		76		21				0.2
Yemen	7		11		74	3	6		1.6
USSR (former)	3	2	5	26	10	8	21	26	3.2

South Asia									
Afghanistan			11	16	62		7	4	1.2
Bangladesh	7				5	42	25	20	4.6
Bhutan				25		25	19	31	3.7
India	3	0	2	2	12	44	23	14	2.7
Nepal				23			35	42	4.4
Pakistan	1	2	21	7	46	13	9	0	1.1
Sri Lanka	32				42	3	13	10	1.7
Southeast Asia									
Brunei		67				33			2.3
Cambodia	7	4				19	5	65	6.5
Indonesia	24	9			4	9	14	40	5.4
Laos						7	7	86	9.1
Malaysia	8	3			2	10	33	45	6.7
Myanmar	5			1	2	16	23	52	6.7
Philippines	32				2	22	13	31	4.7
Thailand	7					41	7	45	5.5
Vietnam	17	1			4	30		49	6.0
East Asia									
China	1	1	14	22	21	17	5	18	2.4
Japan	21			1	5	18	27	30	4.2
Korea (North)	12				18	14	16	39	3.8
Korea (South)	10				24	20	24	22	3.4
Mongolia			32	5	52	1	3	7	1.4
Taiwan	29					12	6	53	5.5

Source: Estimates are from the *Global Biodiversity* report (Groombridge, 1992, p. 251), based mainly on Olson *et al.* (1983).

Table 8 Tropical humid forest types of southeastern Asia

Climate	*Soil water*	*Zone*	*Elevation*	*Soil type*	*Forest type*
Everwet	Dry land	Inland	Lowland (<1200 m)	Zonal	Lowland rain forest
			Mountains		
			1200–1500 m		Lower montane rain forest
			1500–3000 m		Upper montane rain forest
			3000 m to treeline		Subalpine forest
			Lowland (mostly)	Podzolized sands	Heath forest
				Limestone	Limestone forest
				Ultrabasic rock	Forest on ultrabasic rock
		Coastal		Sand	Beach vegetation
	Near surface	Saltwater			Mangrove forest
		Brackish water			Brackish water forest
		Fresh water		Oligotrophic peat	Peat swamp forest
			Wet permanently	Eutrophic (muck and mineral)	Freshwater swamp forest
			Wet periodically		Seasonal swamp forest
Wet–dry	Moderate dry season	Inland			Semievergreen forest
	Increasing seasonal deficit				Moist deciduous forest
					Dry deciduous forest

Source: Reproduced from Whitmore (1975), and Steenis (1950).

Table 9 Generalized vegetation belts in mountains of tropical southeastern Asia

Belt	*Elevation (m)*	*Floristic zone*	*Main taxa*
Upper montane	1500–2100	Montane ericaceous	Ericaceae, Myrtaceae, conifers
Lower montane	1200–1500	Oak-laurel	Fagaceae, Lauraceae
	750–1200	Upper dipterocarp	*Shorea*, *Dipterocarpus*
Lowland	300–750	Hill dipterocarp	Lowland + *Shorea curtisii*
	0–300	Lowland dipterocarp	Many dipterocarps, especially *Dipterocarpus*, *Shorea*, *Dryobalanops aromatica*

Source: Modified from Symington (1943). The oak-laurel zone corresponds closely to subtropical/warm-temperate laurel forests further north (Ohsawa, 1995).

Ghats, one of the most important trees is teak (*Tectona grandis*), which grows very straight but has unusually large, soft leaves very subject to insect herbivory. Other important trees include *Shorea* spp., *Terminalia tomentosa*, *Lagerstroemia lanceolata*, *Grewia tiliaefolia*, and *Pterocarpus marsupium*. To the east one encounters sal (*Shorea robusta*), a dipterocarp and another of the most important forest trees of tropical Asia, economically and ecologically. Moist deciduous forests with *Shorea* continue eastward from the Eastern Ghats through Assam and Myanmar into parts of Southeast Asia. The composition and biogeography of the moist as well as dry deciduous forests of India are well described by Mani (1974), who provides photographs of various forest areas which are now probably gone. Other moist deciduous forests occur in scattered mountain areas of Sri Lanka and the East Indies but have been largely destroyed over most of their area.

Tropical Dry Forest, Scrub, and Savanna

"Dry deciduous forest" occurs where the dry season is at least as long as the wet season and involves shorter, somewhat more open forests dominated by trees with usually smaller leaves, including many Leguminosae. Species richness and leaf area are lower than in monsoon forest or rain forest, and canopy height is usually not more than about 15 m. In still drier areas, dry forest may grade into dense but shorter thorn-scrub, dominated by sprawling, thorny arborescents with sparse foliage of quite small leaves (or leaflets), many of which may be Leguminosae. In other drier areas, especially on relatively flat terrain, trees may be only widely scattered in savannas dominated by tall to short seasonal grasses. In most areas, all three of these vegetation structures may be interspersed in mosaic landscapes, with forest now usually replaced by more extensive areas of degradation stages: thorn-scrub, sclerophyll scrub, and derived savanna. This kind of rain-green mosaic landscape occurs over much of the Deccan Plateau and Ganga Valley, in the "dry belt" of interior Myanmar, and in much of interior Southeast Asia. It also occurs on the drier northern to eastern side of Sri Lanka and in drier parts of the northern Philippines and the East Indies, especially in the Lesser Sunda Islands of eastern Indonesia. An interesting variant of this rain-green response to the tropical dry season is represented by the somewhat open "dry evergreen forest" of interior Thailand.

The gradual transition between moist and dry deciduous forest occurs in particular across the Deccan Plateau, from moist deciduous forest in the Western Ghats to thorn-scrub and savanna in the rain shadow immediately eastward and to dry deciduous forest, and again to moist deciduous forest further eastward. Teak and sal still occur in the dry forest, but trees such as *Diospyros melanoxylon* (ebony), *Anogeissus latifolia*

(Combretaceae), leguminous taxa such as *Acacia* (especially A. *catechu*) and *Cassia*, and low arborescents such as *Zizyphus* become more important. This is also the home of the evergreen but semi-parasitic sandalwood tree (*Santalum album*). Further east, in interior Southeast Asia, the main type of dry deciduous forest is "dry dipterocarp forest," reaching over 20 m in height and with canopy species such as *Dipterocarpus tuberculatus*, D. *intricatus*, *Shorea obtusa*, and Sh. *siamensis*. In addition, dry evergreen forest dominated by evergreen dipterocarps such as *Hopea ferrea* and *Shorea henryana* covers significant areas, especially in Thailand.

Humid Tropical and Subtropical Mountains

The main truly tropical or subtropical humid mountain areas in Asia are the lower slopes of the Himalaya (south side), the Western and Eastern Ghats in peninsular India, the mountains of Southeast Asia, and the smaller ranges and isolated volcanic peaks of the East Indies, the Philippines, and Sri Lanka. A general zonation scheme for the vegetation of mountains in tropical Asia was shown in **Table 9**. As one ascends any mountain range, two climatic changes occur which together quickly reduce the effect of any dry season which may exist in the lowlands:

1. Temperature decreases (approximately 6 °C per kilometer), reducing potential water loss to evapo-transpiration.
2. Humidity and precipitation increase, at least up to the quasi-permanent cloud belt, which is usually 1000–1500 m above the base of the mountains (lower in more humid areas).

Thus, a tropical dry season in the surrounding lowland may disappear completely with a climb of 1000 m or less. At some elevation, generally between 800 and 1200 m, the lowland forest is replaced by a belt of tropical montane forest reminiscent of the evergreen laurel forest of the warm-temperate zone and dominated by temperate taxa such as Ericaceae and Ilex, plus subtropical taxa such as Myrsinaceae, Symplocaceae, Rubiaceae, and understory Lauraceae and Theaceae. From the lowland forest to this montane forest, leaf size typically changes from normal mesophyll to a mix of smaller notophyll and microphyll (Ohsawa, 1995). Above this there may be a low, dense, distinctly microphyll cloud forest of laurophyll treelets covered by mosses and other epiphytes. Of the truly tropical Asian mountains (i.e., excluding the upper parts of the Himalaya), however, only Mt. Kinabalu is high enough for an alpine belt and it is only rock (old volcanic plug) at its top.

Mount Kinabalu (4110 m, northern Borneo) is the highest mountain between the Himalaya and New Guinea (geologically part of Australia). Due to its location near the equator, Mt. Kinabalu contains taxa from both hemispheres, including some distinctly Australian elements such as *Leptospermum* and *Platycladus*. Due to the discontinuous nature of the Southeast Asian land-mass and its fluctuating biogeographic history, Mt. Kinabalu also represents a true biogeographic "island in the sky" and one of the most interesting goals of any knowledgeable geobotanist. The lower slopes are covered by lowland and montane (subtropical) rain forest, followed by laurel forest rich in *Syzygium* and other amphi-tropical taxa. Near the probably edaphic treeline (around 3500 m), on shallow soil, is a belt of low, open forest dominated by *Leptospermum*, with extremely small (leptophyll) leaves reminiscent of Australian forms and the *Kunzea* woodlands of New Zealand. Other interesting features include 30-cm-wide, brown, foul-smelling, fly-pollinated flowers of *Rafflesia* and the diversity of *Nepenthes* species, with their insectivorous pitchers hanging on vines.

The mountains of mainland Southeast Asia are less described but generally show the same vertical zonation as described elsewhere (Whitmore, 1984). Evergreen *Castanopsis*, *Lithocarpus*, and *Cyclobalanopsis* (evergreen *Quercus*), dominants in the laurel forests of East Asia, extend far into the tropics in Southeast Asia to form extensive montane forests. Satellite data have suggested that some relatively large areas of montane rain forest and laurel forest may remain.

In the Western Ghats of Kerala, moist deciduous forest forms a lower montane forest belt, followed upward by semi-evergreen transition and then evergreen laurophyll cloud forest. The same zonation probably occurred on other mountains of southern India and Sri Lanka but has been largely destroyed, converted especially for tea plantations. A similar zonation occurs on the lower southern slopes of the eastern Himalaya but involves more subtropical species.

Dry Regions of West and Interior Asia

The dry regions of Asia include the subtropical deserts of southwestern Asia, the adjacent region of Mediterranean-type ecosystems from Turkey to Israel, the Asian portion of the temperate grassland corridor from the Ukraine to Mongolia and northern China, and the semi-desert and desert regions of both Middle Asia (Turkestan) and Central Asia (Gobi and northwest China as well as the high, dry Tibetan Plateau). Major mountain ranges separate some of these and adjacent regions, especially the Himalaya, which blocks moist monsoonal air from the south, and the Tien Shan system, which blocks moist air from the Mediterranean and thus separates summer-dry Middle Asia from summer-rain Central Asia. Most of these dry climates are also continental, reaching coastal areas only in the Middle East. The Mediterranean region contains landscapes ranging from small areas of sclerophyll forest and montane conifer forest to sclerophyll shrublands and dwarf-scrub. The temperate grasslands range from tall-grass prairies and meadow-steppes in the north to drier, open steppes in the south, following the north–south moisture gradient which obtains across most of the continent. These steppes eventually grade southward into the temperate semi-deserts and deserts.

Subtropical Deserts

Deserts (*sensu strictu*) have essentially no vegetation, whereas semi-deserts have scattered plants. The subtropical arid zone of Asia contains large areas of extreme desert: the Rub al-Khali (empty quarter) of the southeastern Arabian Peninsula, some areas of the Syrian desert (in Iraq and northern Arabia, not Syria), and the Dasht-e-Kavir and smaller areas of interior Iran. The rest of the subtropical desert area, from the Middle East to westernmost India, is mostly semi-desert, with widely scattered xeromorphic dwarf-shrubs, desert grasses, and

ephemeral herbs on coarse, often gravelly or even rubbly substrates.

Desert conditions may have existed since the earlier Paleozoic. Whether the desert flora arose this early, or in the Miocene (expanding during the colder, drier Pliocene and Pleistocene), the desert flora of Asia has probably been closely related to that of northern Africa for a long time. Although total species richness is low and perhaps because there are essentially no true native stem-succulents, the vegetation of the Middle Eastern deserts shows a wide variety of adaptations of form and seasonality to avoid dehydration (Walter, 1985; Zohary, 1973). Some dwarf-shrubs are evergreen but greatly reduce their leaf area at the beginning of the dry season (e.g., *Artemisia monosperma*, *Reaumuria palestina*, *Salsola villosa*, *Suaeda palestina*, and *Zygophyllum dumosum*). Some are wintergreen shrubs, losing their leaves in summer (e.g., *Lycium arabicum* and *Anagyris foetida*), whereas some (generally dwarf-shrubs) shed their leaves but retain evergreen stems, as in *Retama* and *Calligonum* species. Some have grayish, soft, deciduous leaves but keep smaller leaves during the summer (e.g., *Artemisia herba-alba*). Some are totally leafless, such as *Ephedra*. Perhaps most interesting are the leafless shrubs with modular stems and green, photosynthetic but disposable bark segments (e.g., *Haloxylon persicum* and *Anabasis articulata*). Many of these desert dwarf-shrubs extend well into the region of cold-winter continental deserts, in Middle and even Central Asia.

Where more water is available, there may be scattered trees such as acacias, deciduous thorn-shrubs such as *Noaea*, or desert grasses such as *Aristida*. Date palms (*Phoenix dactylifera*) occur around oases, where groundwater is reliable and near the surface. If seasonal pools form, small marshes of sedges and reeds occur. Halophytes include species such as *Halocnemum strobilaceum* on very saline sites and *Nitraria retusa* in brackish water. Deciduous but otherwise juniper-like tamarisk trees (*Jamarix* spp.) also occur where groundwater can be reached, generally at greater depth.

These dry areas are the landscapes of Judaism, Christianity, and Islam. These are also the landscapes of animals such as camels, which sweat but tolerate enormous loss of body water, and the oryx and eland, which pant to moisten nasal sinuses, thus cooling the blood supplied to their brains. There are also many rodents, living in underground burrows and foraging only at night, as well as ectothermic lizards and other reptiles. Most birds spend less time in deserts but may still be dependent on desert ecosystems during the cooler season, migrating to northern Eurasia in summer.

The two main mountain ranges in the subtropical desert zone are the Hejaz (Yemen and southwestern Saudi Arabia) and the Zagros and other mountains of southern and central Iran. The montane belt of the Hejaz contains dry juniper woodlands (*Juniperus*).

Mediterranean Ecosystems

Mediterranean-type climates have cool winters with some precipitation and long, dry summers with no significant precipitation perhaps for several consecutive months. As a result, the primary growing season is restricted to springtime, when temperatures are increasing and soil water is readily available. Plant activity may continue into the summer as long as root systems can still reach deeper soil water. A shorter, secondary growing season often occurs in the autumn, when westerlies again bring some precipitation but before temperatures drop too low. The most characteristic plant adaptation to this climate is woody vegetation with deep root systems and hard (sclerophyll) evergreen leaves, often with waxy cuticles and/or oily secondary compounds which retard water loss from exposed surfaces. This type of "mediterranean sclerophyll" vegetation occurs in all five of the world's mediterranean-climate regions and represents a striking example of convergent evolution. The mediterranean region of Eurasia lies mainly between southern Europe and North Africa but extends eastward along the coastal rim of western and southern Turkey, to Cyprus, and eastward from Syria, Lebanon and Israel to more continental Iraq, the southern Caucasus region, and Iran.

Three general types of mediterranean landscapes can be recognized, differentiated primarily by total water availability. Where sufficient soil water can be reached throughout the year, forests composed of sclerophyll oaks such as *Quercus ilex* covered the Mediterranean borderlands (more in the more humid west) before they were cut by the ancient Romans to build their navies. Where less but still some water is available through the summer, perhaps at greater depth, the landscape is dominated by the characteristic dense sclerophyll woodlands and scrub denoted by their French name *maquis*. Where soil water is not available through the summer, more open stands of smaller, summer-deciduous dwarf shrubs dominate landscapes generically called *garrigues* (which may be patchy and may include some sclerophylls). Other plant types, however, also occur in mediterranean ecosystems on permanently more moist sites, including more water-demanding, evergreen laurel, myrtle, oleander, and even soft-leaved, winter-deciduous shrubs.

The Mediterranean region of Eurasia has been settled and modified by humans for several millennia, with consequent massive soil erosion and other degradation. As soil was lost, both soil water-holding capacity and local precipitation were reduced, resulting in replacement of forest and woodland by various types of garrigues in typical degradation stages.

In Mediterranean mountains, where winter temperatures may become too cold for sclerophylls, dry forests of pines are common, especially in Turkey. Cypresses (*Cupressus* spp.) and true cedars (*Cedrus*) may also occur, the latter occurring only in the mediterranean mountains of Lebanon (*C. libani*) and Cyprus (*C. brevifolia*), in the Atlas Mountains of northwestern Africa (*C. atlantica*), and in the drier western part of the Himalaya (*C. deodara*). Summergreen trees may also appear in mediterranean mountains and may form deciduous forests in sub-mediterranean borderlands (mainly in Europe), in continental areas with colder winters (such as the southern Caucasus and northern Iran), and on particular biotopes such as continuously moist flood-plains.

Many of the familiar mediterranean plants of the more humid western Mediterranean do not extend into the Asian region to the east. In particular, the stem-evergreen "broom" form of the western Mediterranean and Atlantic Europe (genera *Genista*, *Cytisus*, etc.) appears not to be important in southwestern Asia. Some species, however, extend throughout

the western and eastern Mediterranean areas, such as *Quercus cerris*, *Ceratonia siliqua*, *Juniperus oxycedrus*, *Myrtus communis*, and *Pistacia lentiscus*. Others are replaced to the east by similar species, such as *Quercus calliprinos*, *Qu. ithaburensis*, *Pistacia palestina*, and *Pinus brutia*. On the other hand, some species widespread throughout the Mediterranean region may have come originally from the east, such as Aleppo pine (*Pinus halepensis*). The vegetation of mediterranean Asia is well described by Zohary (1973, pp. 130–165).

Temperate Grasslands

In temperate Asia, especially in the west but also in the east, precipitation decreases southward while temperatures increase. As a result, climatic dryness increases southward and the temperate grasslands of Asia occur primarily in a rather narrow corridor that extends east–west across the continent, from the Black Sea region to northeastern China, between the boreal forest to the north and drier semi-desert and desert areas to the south. The Russian-derived term "steppe" normally means a short, discontinuous grassland in English-language ecology but is traditionally used in Asia for all grasslands. The Asian grassland zone can be divided as follows:

1. A discontinuous strip of "forest steppe" in the north (mainly *Populus* and *Betula*), analogous to the aspen "grove belt" of North America
2. A zone of tall-grass "meadow steppe" with many colorful forb species, analogous to the tall-grass prairies of North America
3. A zone of shorter, more open steppe, less developed north of the Black Sea but widening east of the Caspian and extending southward toward the desert areas of Turkestan and the Gobi

This zonation, its soils, and steppe ecosystems have been well described in the Russian ecological literature, much of which is essentially unknown in the West. In China also, scientific study of the grasslands (and other vegetation regions) goes back well before the Communist period but is unknown in the West.

The Eurasian steppe corridor is broken, largely by the Altai Mountains, into the Russian–Siberian steppes and the grasslands of Mongolia and northern China (Lavrenko as cited in Walter, 1974, pp. 163–165). In the western section, the steppe is composed especially of *Stipa* species, such as *Stipa capillata*, *S. lessingiana*, *S. stenophylla*, and other important grasses such as *Festuca sulcata*, *Koeleria gracilis*, *Dactylis glomerata*, *Poa botryoides* and other *Poa* spp., various *Agropyron* spp., *Bromus inermis*, and species of *Carex*. Especially characteristic of the western steppe are the many small geophytes ("ephemeroids"; Walter, 1974) from genera such as *Tulipa*, *Crocus*, *Ornithigalum*, *Bulbocodium*, and *Hyancinthella* as well as *Poa bulbosa*. Tall forbs in the meadow-steppe include *Anemone patens*, *Adonis vernalis*, *Salvia nutans*, *S. superba*, and *Delphinium confusum*. These are accompanied by some shorter shrubs and semishrubs, including the genera *Spiraea*, *Caragana*, and *Artemisia*. The main tree species in the forest-steppe, a wooded tall-grass prairie, are *Quercus robur*, *Betula verrucosa*, and *Populus tremula*.

In the mountains of Middle Asia, montane steppes occur below 2000 m and taller meadow-like grasslands above, including a park-like forest-steppe transition. These may involve grasses such as *Alopecurus pratensis* and *D. glomerata* and tall forbs, such as *Delphinium confusum*, *Ligularia altaica*, *Aquilegia sibirica*, and *Scabiosa alpestris*.

The grasslands of Central Asia, east of the Altai Mountains, involve mainly the Mongolian steppe and the grasslands of Manchuria. In the Mongolian steppe the most important grasses are *Stipa* species such as *S. krylovii*, and *Aneurolepidium chinense*, plus *K. cristata*, *F. sulcata*, and *P. botryoides* from the west. These are joined by xeric shrubs such as *Caragana* species and *Artemisia frigida*. The Manchurian grassland is largely a meadow-steppe dominated by *Aneurolepidium chinense*, along with *A. pseudo-agropyron* and *Stipa baicalensis*, plus forbs such as *Tanacetum sibiricum* (Compositae) and xeric shrubs such as *Artemisia sibirica* (Walter, 1974). On the lower slopes of the southern Hinggan Mountains is a forest-steppe transition to montane forests of *Larix dahurica*, *Betula platyphylla*, and more temperate forests of *Quercus mongolica*, *Tilia mongolica*, and *Pinus tabulaeformis*.

The development of the world's temperate grasslands is closely linked with populations of large ungulates, such as the well-known bison of North America. In Eurasia some of the larger animals, most highly endangered, include Przewalski's horse (*Equus przewalskii*), the Saiga antelope (*Saiga tatarica*), the Zheiran gazelle (*Gazella subgutturosa*), and the wild camel (*Camelus ferus*). The ranges of these and many smaller grassland animals also extended into the adjoining semi-desert and desert zones. Smaller animals include many rodents, such as jereboas, gerbils, and small hamsters.

Temperate Semi-deserts and Deserts

The temperate deserts and semi-deserts of Asia, with cold winters, occur south of the temperate grasslands in three general areas (Walter and Box, 1983):

1. In Middle Asia (Turkestan), from the Caspian Sea (below sea level) to the foothills of the Altai and Tien Shan mountains
2. In Central Asia, including Dzungaria (northern Xinjiang), southern Mongolia (Gobi), part of Chinese Inner Mongolia (e.g., Ordos desert), the Tarim Basin (southern Xinjiang), and the smaller Tsaidam Basin (Qinghai), all approximately 500–1500 m above sea level
3. On the high, cold Tibetan Plateau, with an average elevation of approximately 4000 m

Separating Middle and Central Asia is the Tien Shan mountain system, running generally from southwest to northeast. This complex mountain system, with the associated Pamir to the south and Altai to the north, rises from the dry continental interior and constitutes to some extent a multi-zonal orobiome of its own (Walter and Box, 1983), with well developed coniferous forests related more to those of humid eastern Asia.

The climates of the temperate deserts are warm to hot in the summer (except for Tibet) and severely cold in the winter, except in the southern parts of Middle Asia (Turkmenistan), near the Iranian border. Precipitation in Middle Asia occurs

mainly in the spring and autumn, with very scant snow cover in winter and essentially no precipitation in summer. In Central Asia precipitation is confined almost completely to summer, brought by monsoon storms which manage to penetrate into the continental interior. Soils in all regions are coarse and skeletal, salinized in areas of internal drainage. The Central Asian deserts are totally continental and separated from the subtropical deserts of southwestern Asia. Although the Middle Asian deserts share some species with the subtropical deserts, they also are distinct from the subtropical deserts, both climatically and physiographically, separated by the Kopet Dagh and other mountains of Iran and Afghanistan.

Middle Asian Deserts

A characteristic of the sandy deserts of Middle Asia is the dominance of both small, typical and much larger xeromorphic shrubs, some leafless and some at least partly evergreen. Reaching 2 m in height, leafless *Haloxylon persicum* extends over the whole region and even into the Gobi desert. It is joined in the Kara-Kum desert by *H. ammodendron*, a leafless tree form reaching 8 m in some locations. This is the only cold-winter desert region in the world in which a tree form gains even local dominance. Other important desert shrubs include *Artemisia pauciflora*, *A. terrae-albae*, *Halostachys caspica*, *Anabasis salsa*, *Halocnemum strobilaceum*, and *Kochia prostrata*, along with steppe and desert grasses such as *Festuca sulcata*, *Stipa capillata*, *Koeleria cristata*, and *Agropyron repens*. In areas of flat terrain and finer soil, mini-geophytes such as *Poa bulbosa* form vernal carpets of miniature flowers (Walter, 1974, p. 254).

Dzungaria, Gobi, and Tarim Basin

Dzungaria is the region east of the Dzungarian Gate, a gap in the Tien Shan–Altai mountain system, and extends across northern Xinjiang. This corridor between Middle Asia and the Gobi, as well as the Gobi itself, represents a region of rockier substrates and deserts dominated by typical desert shrubs such as *Artemisia*, *Anabasis*, and *Calligonum*. The most important grasses are *Stipa* bunch-grasses. Isolated trees of *Populus diversifolia*, *Tamarix ramosissima*, and occasionally *H. ammodendron* may occur along streams.

The Takla Makan (Tarim Basin), on the other hand, is a large sand sea within the horseshoe (open to the east) formed by the Tien Shan on the north, the Pamirs in the center, and the Karakoram and Kunlun ranges on the south. Together they form one of the most effective rain shadows in the world, and precipitation has very rarely ever been recorded in the Tarim Basin. There is little vegetation on the shifting sand areas, but runoff from the mountains does generate intermittent and permanent streams which extend into the basin. Along these streams one can find riparian strips of the phreatophytic *Tamarix ramosissima* as well as the widespread *Populus diversifolia* and *Ulmus pumila*. In some areas there are even extensive marshes of *Phragmites*, *Typha*, and *Scirpus* species, which are especially important for the large numbers of birds which overwinter in India and migrate to Siberia for the summer.

Tibetan Plateau

The high, dry Tibetan Plateau (average elevation approximately 4000 m) is a region of dry steppes, alpine mats (especially of *Kobresia* species), dwarf conifer scrub of *Juniperus* and *Sabinia*, and other sparse, low-growing, often cushion-shrub vegetation adapted to the harsh conditions. Many familiar genera from Central Asia (and Middle Asia) are represented, including *Caragana*, *Reaumuria*, *Acanitholimon*, *Astragalus*, *Tanacetum*, *Artemisia*, *Ptilagrostis*, and *Festuca*. Some familiar species also extend into the Tibetan highlands, including *Eurotia ceratoides* and *Kochia prostrata*. East Asian elements are important only in the east but include *Gentiana*, *Primula*, *Saxifraga*, *Saussurea*, and *Rhododendron*. Holarctic elements are few, but some are very abundant, especially in the alpine mats and marshes: *Kobresia*, *Carex*, *Eleocharis*, *Eriophorum*, and *Juncus*. Where moisture permits, subalpine tall-forb stands include such genera as *Aconitum*, *Delphinium*, *Ligularia*, *Polygonum*, and *Rheum*. These different geoelements all represent taxa which migrated into Tibet after the last glaciations.

Humid Monsoon Asia (Extratropical)

Monsoon Asia is the region dominated by the seasonally alternating monsoon wind system, which brings wet oceanic air masses onshore in the summer and blows dry and cold outward from the continental interior in winter (mainly eastward and southward). The region lies east of the Altai–Tien Shan mountains and extends from southern Siberia, through China, to Korea and Japan in the east and to tropical Asia in the south. In most of this region the only major mountains run east–west, so mean winter temperatures are low but outbreaks of even colder Siberian air are largely excluded. Except toward the interior, summer brings adequate rainfall throughout for forest growth. In fact, in Japan and some parts of eastern China, there is rarely any climatic drought and forests can be very mesomorphic. Humid monsoon Asia (i.e., forested East Asia) can be divided into two regions: (i) temperate deciduous forest in the typical-temperate and cool-temperate areas to the north and (ii) evergreen broad-leaved "laurel" forest in the warm-temperate and humid-subtropical zones to the south. Pine forests are important secondary landscapes in both areas. Mountains in East Asia are especially interesting for the rich variety of endemic conifer taxa which they contain.

Laurel Forest (Evergreen Broad-leaved Forest)

Laurel forests are evergreen broad-leaved forests of warm-temperate and humid-subtropical climates, dominated by trees with intermediate-sized, dark green (shade-tolerant), thin-coriaceous but mesomorphic leaves (laurophylls) such as are especially characteristic of the laurel family (Lauraceae). Forests dominated by laurophyll trees are rather dark and somber, with low light levels below the canopy, and are evergreen from top to bottom. In some respects laurel forests can be thought of as an extension of the tropical rain forests into the warm-temperate zone. East Asia contains the world's largest area of such forests, due to its abundant rainfall and winters without severely low temperatures, at least in coastal areas or on islands. Laurel forests occur generally south of about 35° N latitude but extend to 38° N on both sides of Japan. Counterparts occur in southern Brazil and small areas

of eastern Australia, as well as moist depressions ("bay forests") in the southeastern United States. Cool-temperate analogs occur in New Zealand and Tasmania.

In East Asia, laurel forests occur in southeastern China, in drier southwestern China (Sichuan and Yunnan, with different but largely vicariant species), in southern Japan, and in a small strip across southernmost Korea (Wang, 1961). In Southeast Asia, laurel forests also ascend into the mountains to form montane forests just above the tropical seasonal evergreen forests of the lowlands (Whitmore, 1984; Ohsawa, 1995). The main canopy tree genera of laurel forest are *Persea* (=*Machilus*), *Cinnamomum*, *Beilschmiedia*, etc. (Lauraceae); *Castanopsis*, evergreen *Quercus* (=*Cyclobalanopsis*), and *Lithocarpus* (Fagaceae); *Schima* (Theaceae), *Ilex* (Aquifoliaceae), *Michelia* (Magnoliaceae), and others, all with very similar laurophyll physiognomy. Understory trees and arborescents are largely from the same families. Lauraceae tend to be especially important in more perhumid (e.g., coastal) areas, whereas evergreen Fagaceae tend to become more important inland, especially in southwestern China.

In humid East Asia, the largest species turnover between the polar region and the tropics occurs within the bioclimatic zone of evergreen broad-leaved forests—without major change in the forest physiognomy. This occurs in the Okinawa Islands of Japan and in southeastern China, as temperate species abruptly disappear, including many laurophyll tree species, and are replaced by an essentially tropical flora which includes many new tree genera with essentially the same evergreen broad-leaved structure.

Temperate Deciduous Forest

The temperate deciduous forests of East Asia cover the northern half of Japan (including Hokkaido), most of Korea, and the northeastern part of China, especially Manchuria. This region is smaller than the region of evergreen broad-leaved forest and smaller than the deciduous forest region in eastern North America. Field estimates of primary productivity in Japan suggested that the deciduous forests were less productive than in North America and less productive than some montane conifer forests of Japan. Nevertheless, the deciduous forests of East Asia appear to have more species and genera, including almost all genera found in eastern North America (except *Carya*) and in Europe, plus some additional ones (e.g., *Cercidiphyllum*).

Japan and Korea

The deciduous forests of Japan are especially rich, due perhaps to the perhumid maritime climate and total lack of water stress during the growing season. Deciduous forests cover most of northern Honshu, dominated mainly by *Fagus crenata* but with many other taxa, such as *Fraxinus lanuginosa*, *Tilia japonica*, and *Magnolia obovata*. Understories are rich with maples (e.g., *Acer mono*, *A. japonicum*, and *A. palmatum*) and other mesic understory elements such as *Styrax obassia*, *Sorbus alnifolia*, *Cercidiphyllum japonicum*, *Aesculus turbinata*, *Callicarpa japonica*, *Hamamelis obtusata*, *Viburnum*, and *Euonymus*. One interesting transitional forest type occurs in parts of northeastern Honshu and involves co-dominance by *F. crenata* and *Abies firma*, a temperate-zone fir. Forests of *Fagus japonica* generally occur only in the areas of transition to warm-temperate forest. An especially characteristic feature of the deciduous *Fagus* forests of Japan is the prevalence of short but broad-leaved bamboos of the genus *Sasa* (and recently also *Sasamorpha*), which form dense understories 0.5–1 m high where there is continuous snow cover. Other common herb-layer taxa include *Disporum*, *Viola*, *Chimaphila*, *Mitchella*, *Carex*, and ferns. Floodplain and other especially moist forests are dominated by ashes (*Fraxinus sieboldiana* and *F. spaethiana*), elms (e.g., *Ulmus davidiana*), *Pterocarya rhoifolia*, and other typical floodplain taxa (e.g., *Alnus hirsuta*). Also belonging to the temperate zone are the hemlock (*Tsuga sieboldii*) forests which occur most commonly on rocky substrates at the base of lower mountain slopes. Drier areas (still without water deficit) are generally covered by deciduous oak forests, including interior areas, lower mountains (e.g., *Qu. serrata*), and most of Hokkaido (*Quercus mongolica* var. *grosseserrata*).

The deciduous forests of Korea occur in a distinctly more continental climate, in which oaks (e.g., *Quercus mongolica*), and also *Carpinus*, become more important. Much of the Korean landscape is still recovering from the Korean War, after which the southern part was seeded with North American pines. In deciduous forest areas of both Japan and Korea, as in the evergreen forest region, native pines such as *Pinus densiflora* play a very important role in successional landscapes.

Manchuria and Eastern China

The forests of northeastern China and the Amur valley have a shorter growing season than in Japan but are still very rich. Canopy trees may include such species as *Quercus mongolica*, *Fraxinus mandschurica*, *Tilia amurensis* and *T. mandschurica*, *Ulmus macrocarpa* and *U. davidiana*, *Phellodendron amurense*, *Maackia amurensis*, *Juglans mandschurica*, *Acer mono*, *Betula dahurica* and *B. costata*, and *Populus koreana* and *P. ussuriensis*. A common temperate conifer in these forests is *Pinus koraiensis*. Understories may include trees of mainly the same species plus shrubs such as *Corylus heterophylla*, *Deutzia amurensis*, *Euonymus pauciflora*, *Philadelphus schrenkii*, *Syringa amurensis*, and *Viburnum* spp.

Further south, in northern, eastern, and into southeastern China, the deciduous and evergreen broad-leaved forests are almost totally destroyed by millennia of human habitation. As a result, it is difficult even for local scientists to identify the boundaries of the two forest regions. In hill and mountain areas of northern China (Liaoning, Shandong, Hebei, and Beijing) remaining mature forest types are dominated by *Quercus mongolica*, *Qu. liaotungensis*, and other *Quercus* species along with *Celtis*, *Fraxinus*, *Tilia*, *Betula*, and *Ulmus* as subordinate canopy elements. Understories may have elements such as *Lindera obtusiloba*, *Corylus heterophylla*, *Lespedeza bicolor*, and *Spiraea trilobata*. *Betula* and *Populus* forests occur as secondary types. *Quercus variabilis* forests occur as transitional forests south of the main deciduous forest region.

Among the most interesting deciduous forests are those involving various species of beech (*Fagus*), generally covering small areas at elevations from 1300 to 2200 m in central eastern China (Zhejiang to Guizhou provinces). Deciduous trees in interior southern China may have only short leafless periods. This is also one of very few areas in the world in

which deciduous and evergreen broad-leaved trees coexist in semi-evergreen broad-leaved forests.

Pine Forests

Pine forests represent extensive landscapes in East Asia due to long histories of continuing forest cutting. Until after World War II, many of the evergreen broad-leaved forest areas were managed as coppice forests. Other areas, however, were burned and cleared much earlier. In hilly and mountainous terrain, soil erosion followed clearing and broad-leaved forests could no longer recover. These areas usually became pine forest and were maintained by intermittent natural and induced burning of the understory. Much of southern Japan, the Korean peninsula, and upland warm-temperate and subtropical China are now covered by such pine forests.

In Japan, Korea, and some parts of northeastern China, these secondary pine forests are mainly dominated by *Pinus densiflora*. In eastern and southeastern China, pines with more warm-temperate to subtropical distributions become dominant, especially *P. massoniana* and *P. tabulaeformis*. In southwestern China, forests of P. *armandii* occur in Sichuan and on the Yunnan plateau, whereas subtropical pine forests of *P. kesiya* (cf. *P. khasya*) become important in the mountainous transition southward into lowland Xishuangbanna and the mountains of northern Thailand. In addition to pine, the warm-temperate to subtropical "Chinese fir" (*Cunninghamia lanceolata*) is also important in secondary landscapes and, like pine, is widely planted.

Mountains of Humid East Asia

The mountains of East Asia represent the only large mountainous area in the warm-temperate and humid subtropical zones of either hemisphere (although smaller areas do occur in Mexico, southern Brazil, eastern Australia, and northern New Zealand). The various mountain systems of humid East Asia, including at least the more humid eastern Himalaya, thus represent areas of unusual diversity and endemism, especially among what appear to be relict conifers. Endemic conifer genera include *Amentotaxus*, *Cathaya*, *Cephalotaxus*, *Keteleeria*, *Cryptomeria*, *Cunninghamia*, *Fokienia*, *Glyptostrobus*, *Sciadopitys*, *Taiwania*, and *Thujopsis*, at least some of which once occurred more widely. In addition, more widespread conifer genera such as *Abies*, *Chamae-cyparis*, *Picea*, *Taxus*, *Thuja*, and *Tsuga* have relatively high species diversity in these areas. In most of the area, the lowland biome is evergreen broad-leaved forest, which may extend well into the montane belt in most places, with well-developed *Rhododendron* belts and a more warm-temperate character. As winter becomes colder upward, summergreen deciduous forest may occur, but the forest may also grade directly into evergreen high-mountain conifer forest. None of these warmer mountain systems is high enough to have an alpine belt except the Himalaya.

In the western Himalaya, montane conifer forests are dominated by *Pinus roxburghiana* and the deodar cedar (*Cedrus deodara*), a preferred ornamental tree throughout the world. To the northwest, in the Tien Shan and Altai systems, the montane conifer forests are simpler, dominated often by *Picea schrenkiana* but including forests of shade-tolerant *Pinus sibirica* as well as boreal *Picea* and *Abies*, larches, and *Pinus sylvestris* (Walter, 1974; Walter and Box, 1983, Chap. 7). In western Tibet, the main conifers are *Abies delavayi* and *A. webbia*, along with *Picea likiangensis*. This is also the area of the dawn redwood (*Metasequoia glyptostroboides*), first identified from fossils, only discovered in the wild in the early 1900s but now widely planted as a street tree throughout China.

Extensive evergreen oak forests, involving *Quercus semicarpifolia* and many other species, still occur on the southern slopes of the Himalaya and may represent a somewhat unique warm-temperate forest biome. At higher elevations, *Betula albosinensis*, *Acer tetramerum*, and many other deciduous species form belts of summergreen forest in some areas. Subalpine conifer forests may contain species such as *Picea likiangensis* (better known from Manchuria), *Abies squamata* and *A. georgei*, *Tsuga dumosa*, and even the deciduous *Larix potaninii*. The best developed *Rhododendron* belts are in the eastern Himalaya.

On the lower parts of the eastern Tibetan Plateau, *P. likiangensis*, *A. squamata*, and *A. georgei* form tall coniferous forests, followed upward by *Abies faxoniana*, *Picea asperata*, and *P. noveitchii* (Archibold, 1995). *Larix* may be admixed, and *Sabina* and *Cupressus* form shorter woodlands in drier areas. Toward the Yunnan Plateau to the southeast, *Tsuga yunnanensis* becomes important, with *Picea brachytyla* and *Abies fabri* at higher elevation. At lower elevation, the eastern Tibetan Plateau (mainly in Sichuan) also contains two of China's biological treasures: the range of the Giant Panda at the Wolong Preserve and E-mei Shan (3099 m), one of China's four sacred mountains, with a well-developed zonation (and footpath) from laurel forest to subalpine conifer forest of *A. fabri*. To the south, at elevations not higher than about 1500 m on the Yunnan Plateau, smaller areas of laurel forest remain in the Ailao Shan and other ranges southward into Xishuangbanna.

In the transition region of eastern China (summergreen to evergreen laurel forest), two other accessible mountains also have well-developed zonations, which are botanically well described: Huang-Shan in southeastern Anhui province and Tianmu-Shan in northern Zhejiang province. The mountains of Taiwan are especially steep, and the montane laurel forests are consequently relatively well preserved.

In northeastern China, along the Korean border, the vertical zonation of the Changbai-Shan can be described as follows:

0–500 m: summergreen deciduous forest
500–1100 m: mixed deciduous and conifer forest (*P. koraiensis* plus much secondary forest of *Betula platyphylla*)
1100–1700 m: coniferous forest (*Picea jezoensis* var. *komarovii*, *P. koraiensis*, with *Abies nephrolepis*, *Acer*, *Betula*, *Sorbus*)
1700–2000 m: Birch woodland and krummholz (*Betula ermanii*)
2000–2600 m: Alpine tundra, mostly heath of Vaccinium uliginosum and Dryas octopetala, plus Rhododendron aureum, Rh. confertissimum, etc.
2600–2700 m: Alpine cold-desert (snow still just melting in early September)

Further to the north, larch forests dominated by *Larix gmelini* occur in the northern Da Hinggan Mountains near the Siberian border (Missouri Botanical Garden, 1983; Wang, 1961).

Montane forests and zonation in Korea can perhaps best be seen on Mt. Seolag (1308 m). The zonation is similar to that described for Changbai-Shan but with a greater importance of deciduous *Quercus* and *Carpinus* forests at low elevation and the lack of an alpine belt.

The mountain vegetation of Japan has been very extensively studied and described in the evergreen as well as the deciduous forest zones (Numata, 1974). Among the most impressive native coniferous forests of Japan are the tall forests of *Cryptomeria japonica* occurring on Yakushima island and in some areas with more than 3000 mm of precipitation on Shikoku and Honshu. These forests could be called temperate rain forests. Forests of *Chamaecyparis obtusifolia* are also tall and impressive. Other montane conifers in Japan include *Abies homolepis*, *A. mariesii*, and *A. veitchii* mainly on Honshu; *A. sachalinensis* mainly on Hokkaido; *Picea jezoensis* (Hokkaido) and *P. glehnii*; *Larix leptolepis* (on old lava); *Tsuga sieboldii* (lower mountains) and subalpine *Ts. diversifolia*', *Thuja standishii*; and *Thujopsis dolobrata* (endemic genus).

In most of northeastern Asia, the subalpine belt is characterized by dense stands of the krummholz species *Pinus pumila*, which also occurs in Siberia. In more maritime areas approaching the boreal zone, such as Hokkaido, conifers may be replaced by forests of birch, especially *B. ermanii*.

Boreal and Polar Asia

Boreal and polar Asia extends from the Ural Mountains to the Pacific Ocean and from islands in the Arctic Ocean to the transition to temperate climates in the south, approximately along the southern boundary of Siberia. In Europe, this transition takes the form of mixed boreo-nemoral forests of deciduous broad-leaved trees and conifers, narrowing eastward from Poland through Moscow and on to the Urals around Sverdlovsk. In Asia, this transition is a more discontinuous zone of mixed forests and forest-steppe, running from the southern Urals past the southern end of Lake Baikal and on to northeastern China, where it widens again between the Siberian larch taiga to the north and the deciduous forests of Manchuria. In the middle section, the transition is mainly a narrow belt of aspen (*Populus tremula*) and birch parklands between subboreal or boreal forests to the north and the grasslands of Mongolia to the south.

During the Pleistocene, much of Siberia remained unglaciated but covered by tundra and cold steppe vegetation, which extended southward to the grasslands with no intervening coniferous forest zone. The boreal forest which now covers most of Siberia is thus a relatively recent phenomenon. This boreal forest is mainly an evergreen coniferous forest in the west but opens to lighter larch (*Larix*) forest and woodland in the east, where winters are most severe. The west-Siberian lowland is a vast complex of boreal forest and wetlands, whereas northeastern Siberia is quite mountainous. The larch region of interior eastern Siberia contains the most continental climates in the world, with mean July temperatures usually exceeding 20 °C but with mean January temperatures generally near or below −30 °C, reaching −51 °C at Oimyakon and Verkhoyansk in the northeast. The transition to polar tundra involves a zone of open forest–tundra woodland, which reaches its farthest point north (in the world) in the Taimyr Peninsula of northwestern Siberia. Polar tundra forms a narrow belt along the northern coastal plain completely across Eurasia.

Boreal Forest

If the subboreal zone is included, four main types of boreal forest can be recognized (Walter, 1974):

1. Mixed forests of summergreen broad-leaved trees and conifers (mainly *Picea* and *Pinus*), along with localized pine forests
2. Dense forests of spruce (*Picea*), fir (*Abies*), and five-needle pine (*Pinus sibirica*)
3. More open forests and woodlands of Scots pine (*Pinus sylvestris*)
4. Generally open forests and woodlands of larch (*Larix*), a genus of deciduous conifers

Of these, the dense forests of spruce, fir, and Siberian pine are generally called "dark taiga," whereas the more open forests and woodlands of pine and larch are called "light taiga."

Boreo-nemoral mixed forests, with *Acer*, *Tilia*, *Ulmus*, *Fraxinus*, *Quercus*, *Picea*, *Pinus*, etc., are better developed in eastern Europe than in Asia but reappear in Manchuria and the Russian Far East and some smaller areas such as the subboreal region on the west side of Lake Baikal. In more continental areas, *Betula* and *Populus* may be the only important broad-leaved trees, and pines may be the main conifers, especially if the soil is coarse and less humic. At the southern end of Lake Baikal (52–54° N, 450 m), boreo-nemoral mixed forests can be found in moist alluvial lowlands, whereas dark taiga, with larch and five-needle pine, occurs on uplands higher than about 700 m.

Dark taiga is composed of shade-tolerant conifers, mainly Siberian spruce (*Picea obovata*), five-needle Siberian pine (*P. sibirica*), and Siberian fir (*Abies sibirica*). All three of these extend from the Urals into southeastern Siberia (if not completely to the Pacific). Dark taiga is limited geographically by the extreme winter cold, yielding to deciduous larch taiga generally where mean January temperature falls below about −30 °C (extremes to below −60 °C). The range of dark taiga is thus mainly European Russia through western Siberia and in a narrowing zone into southeastern Siberia. Within this range, dark taiga occurs over most of the area, with more open forests of *P. sylvestris* confined mainly to sandy, upland rock, or other drier substrates. In some smaller areas, mainly in mountains, other species of *Picea* and *Abies* appear as well as other five-needle pines such as *P. cembra* and *P. pumila*, often occurring as subalpine krummholz. Larches, often large, also occur widely as companion species within the dark taiga.

Scots pine (*P. sylvestris*) has the widest distribution of any tree species in Eurasia and covers large areas of light taiga throughout the European and Asian boreal zone and occurs in mixed forests in Europe, the Caucasus, and northern Manchuria. Boreal pine forests are usually open and have understories of two main types: *Cladonia* and similar lichens (plus mosses) on drier sites such as sand, and ericads such as *Vaccinium* on more moist sites.

Most of the boreal zone of central and eastern Siberia is covered by extensive open forests and woodlands of larch, especially *Larix dahurica* but also other species with smaller ranges. *Larix sibirica* is more important as a companion in the dark taiga of western Siberia. As in many higher latitude mountainous regions, the larch woods may occur mainly on slopes while flat valley bottoms remain largely treeless, a pattern attributed to cold-air drainage. The ground vegetation in these areas and in the larch stands as well is composed largely of boreal and tundra dwarf-shrubs such as *Vaccinium uliginosum*, *Ledum*, *Dryas*, and *Arctous*, plus abundant mosses and lichens. In moist areas shrubs from the genera *Juniperus*, *Salix*, *Alnus*, and *Betula* often become important understory elements.

The larch woods of interior northeastern Siberia in particular occur where annual precipitation is generally under 300 mm and sometimes under 200 mm. Although the growing season is thus a period of climatic water deficit, the trees survive from soil water which is effectively rationed to them over the course of the summer by the slow thawing of the soil (Walter, 1974, p. 84). Larches are outcompeted by evergreen conifers in some boreal areas but have no tree competitors in the coldest parts of eastern Siberia. Larch-dominated landscapes thus extend from the polar treeline across much of Siberia southward to Mongolia and into inland areas of the Russian Far East (Primorye and Khabarovsk provinces) and mountains of northernmost Manchuria.

Mountains within the boreal zone may quickly reach treeline. On lower slopes, however, the boreal forest may be somewhat different from the surrounding flatter lowlands. In the Ural Mountains, forests on the wetter western slopes generally involve *P.* sylvestris, *Abies sibirica*, *Picea abies*, *Betula pubescens*, and krummholz of *P. pumila*, whereas forests on the drier eastern slopes are mainly *P. sylvestris*. In less extreme, maritime areas of eastern Siberia, the deciduous larch taiga may be replaced again by evergreen conifers, such as *Abies nephrolepis* around the Sea of Okhotsk and A. *sachalinensis* on Sakhalin island (also northern Japan). Stunted forests of B. *ermanii* appear in the most maritime areas (e.g., Kamchatka, Sakhalin, and northern Japan).

Polar and Mountain Tundra

Polar ecosystems, generally called tundra, occur across the northern coastal plain of Eurasia but also in both upland and high-mountain situations, such as the lower mountains of boreal northeastern Siberia and the high Tibetan Plateau. Upland tundra in the boreal zone represents the first upward vegetation belt and is thus commonly called montane tundra. In high mountains, tundra-like ecosystems occur above the alpine treeline and are generally called alpine tundra. In all three situations, these are treeless landscapes dominated by mosaics of the following basic associations:

1. Largely evergreen, microphyllous dwarf-scrub on locally somewhat higher, drier areas
2. Wet graminoid vegetation in the lowest areas, generally flat, marshy depressions
3. Mixes of evergreen and summergreen plants, herbaceous and dwarf-shrub, including minigeophytes, on the broad, slight slopes between the higher and lower areas
4. Mainly mosses and lichens on the most extreme areas, including the coldest but also the most exposed upland areas where little snow remains

In mountains, two additional types can be identified. One is the "snow valleys" (Schneetälchen), depressions with deep snow accumulations which have shorter snow-free seasons and specialized floras composed of plants which can complete their life cycles in periods as short as one month in summer. The other is the extensive alpine mats, often dominated by *Kobresia*, found in drier high-mountain areas such as the Tien Shan, Hindu Kush, and Tibetan highlands, generally above 3000 m.

The polar tundra extends from extreme northern Fennoscandia across northern Russia to coastal northeastern Siberia in a narrow strip which is widest on the Taimyr peninsula. Due to continentality effects, however, the Taimyr is also the area where boreal forest extends farthest north, reaching about 72.5° N (as larch forest) north of Khatanga. On the flat coastal topography, herbaceous tundra is perhaps most extensive. The polar tundra can be represented zonally, however, in four subzones from south to north:

1. The transitional forest-tundra of largely dwarf-shrub tundra with scattered but still tall individual conifers (mainly larches in the east and spruces in the west)
2. Dwarf-shrub tundra, dominated largely by shrub forms from summergreen tree genera, mainly *Betula*, *Salix*, and *Alnus*
3. Largely herbaceous tundra with few dwarf-shrubs but still a continuous vegetation cover including many mosses and lichens
4. The discontinuous High Arctic cold-desert with vegetation restricted to more scattered individuals of a few vascular plants (especially *Dryas* spp.) and various lichens and mosses

The forest-tundra and dwarf–shrub tundra zones of the lowlands are better represented in western Siberia, but montane analogs are fairly widespread in the lower mountain areas east of the Lena River in eastern Siberia. Cold-desert, on the other hand, is a bit more widespread in the east and on the Arctic islands.

Two other prominent features of the polar zone are the prevalence of permafrost and substrates divided into stone polygons by frost action. Permafrost occurs where mean annual temperatures are below 0 °C, i.e., north of about 62° N continuously and discontinuously in uplands further south. Polygon substrates (and other features such as pingos) occur where seasonal freezing and thawing of "soil" which is largely water forces the larger solids (i.e., rocks) to the surface in a slow, bubble-like fashion that distributes them into polygons which become connected into polygon networks. The interiors of the polygons are lower, flat, and wetter, whereas the rocky borders, only 10–20 cm higher, represent drier biotopes.

Polar tundra in particular (including boreal montane tundra) involves many taxa with circumpolar distributions, including genera such as *Carex*, cotton grass (*Eriophorum*), and other sedges; many ericads (e.g., *Vaccinium*, *Arctostaphylos*, and *Cassiope*); other forbs, such as *Dryas*; and both evergreen dwarf-shrubs (e.g., *Empetrum*) and sometimes taller deciduous dwarf-shrubs (e.g., *Betula*, *Salix*, and *Alnus*). For nonvascular

plants, an even larger percentage of the taxa are circumpolar, including mosses such as *Sphagnum*, *Polytrichum*, and *Hylocomnium*, and lichens such as *Cladonia* and *Cladina*. As in the boreal zone, the flora of eastern polar Siberia is richer than that of western Siberia due to its lack of Pleistocene glaciation.

Most of the same genera, of vascular and nonvascular plants, also occur in the high-mountain tundra. Alpine treeline in mountains of northern Asia is commonly formed by *P. pumila* krummholz, whereas the alpine tundra typically involves familiar heath taxa such as *Empetrum*, *Phyllodoce*, *Cassiope*, and *Dryas* as well as dwarf *Salix* and sedges. In general, however, alpine vegetation is more species rich than is corresponding polar tundra, perhaps due to the lack of permafrost and the greater number of microhabitats in the more heterogeneous terrain of high mountains. Many more localized species usually occur. Some important genera of Asian high mountains which are not important (if present at all) in polar tundra include *Geum*, *Gentiana*, and *Leontopodium*.

Despite its relatively low productivity, the polar tundra supports many animal species and in surprisingly large numbers, at least in summer. Caribou (reindeer) winter in the boreal zone and migrate to the tundra during the summer. Birds migrate from much greater distances away, some from as far as tropical Asia. Other birds, such as ptarmigan, overwinter near the polar zone. Mammals such as polar bears and some small rodents (e.g., lemmings) also live year-round in the polar zone.

Wetlands and Coastal Ecosystems

Wetlands include mangroves, other coastal ecosystems such as strand forests, beach vegetation, salt marshes, deltas, and estuaries and terrestrial wetlands such as swamp forests, marshes, bogs, and (*sensu lato*) even lakes and streams. Most wetlands are highly productive systems when not badly degraded, have important economic and social benefits for local human populations, and perform useful ecological services such as cleansing wastewater. Wetlands also support a wide range of animal life, often far beyond their own borders, as in the case of migratory birds, fish, and shellfish.

Mangroves are salt-tolerant forests growing in the intertidal zone, thus flushed by saltwater tides once or twice daily but also influenced by the inflow regime of freshwater runoff from the land. Behind many mangroves are so-called "back mangroves" which are less influenced by wave energy, generally less salty, and contain different and usually more species. Mangroves have developed in part by trapping fine sediment, so the mangrove substrate is usually deep mud (but also shallower mud and sand). As a result, mangroves are home to abundant burrowing shrimp, crabs, and other crustaceans as well as spiders, insects, and some birds. The total number of plant "mangrove species" in the world has generally been estimated at fewer than 100, involving especially the genera *Rhizophora* and *Avicennia* as well as particularly Asian genera such as *Bruguiera*, *Ceriops* and *Kandelia* (Rhizophoraceae),

Table 10 Mangrove areas in Asia

Country	*Total mangrove area (ha)*	*Number of protected mangrove areas*
East Asia		
China	67,000	24
Hong Kong	0	9
Japan (Ryukyu Islands)	400	4
Taiwan	174	4
Southeast Asia		
Brunei	7000	3
Cambodia	10,000	0
Indonesia (including Irian Jaya)	4,251,011	152
Malaysia	630,000	99
Myanmar	517,000	6
Philippines	400,000	59
Singapore	1800	2
Thailand	268,693	17
Vietnam (South)	370,000	2
South and Southwest Asia		
Andaman Islands	(50,000)	25
Bangladesh	410,000	5
India	306,000	9
Iran	23,717	3
Maldives	—	0
Pakistan	249,500	2
Sri Lanka	120,000	9
Arabian Peninsula		
Bahrain	40	1
Saudi Arabia, Oman, Qatar	—	0
United Arab Emirates	—	6
Yemen	—	5

Source: Data are from Groombridge B (1992) *Global Biodiversity Status of the Earth's Living Resources*, 325 pp. London: World Conservation Monitoring Center/Chapman & Hall.

Sonneratia, and *Aegiceras*. Back mangroves contain additional taxa such as *Lumnitzera*, *Xylocarpus*, the mangrove palm (*Nypa fruticans*), *Acanthus*, the widespread fern *Acrostichum aureum*, and even the buttressed tree *Heritiera littoralis* (Sterculiaceae) in some cases. In areas protected from direct hits by typhoons, mangrove forests can grow to about 35 m in height. In more exposed areas, however, and in areas with shallower or hard mud, mangroves are reduced to only a few meters in height.

Table 11 Most seriously threatened wetlands in Asia

- Bangladesh
  - Chalan Beel*
  - Haor Basin of Sylhet and Eastern Mymensingh
  - Dubriar Haor*
  - The Sundarbans
  - Wetlands in Pablakhali Wildlife Sanctuary
  - Chokoria Sundarbans*
- Bhutan
  - Boomthang Valley
- Burma (Myanmar)
  - Irrawaddy Delta
- China
  - Yancheng Marshes
  - Shijiu Hu
  - Shengjin Hu and the lower Yangtze Lakes
  - Shengjin Hu
  - Xi Jiang (Pearl River) Delta*
  - Tuosu Hu (Kurlyk Nor) and Kuerhleiko Hu
- India
  - Dal Lake
  - Wular Lake
  - Harike Lake
  - Jheels in the vicinity of Haidergarh*
  - Dahar and Sauj (Soj) Jheels
  - Southern Gulf of Kutch
  - Gulf of Khambhat
  - Khabartal
  - Dipor (Deepar) Bheel
  - Logtak Lake
  - Salt Lakes Swamp*
  - The Sunderbans
  - Chilka Lake
  - Kolleru Lake
  - Estuaries of the Karnataka coast
  - Kaliveli Tank and Yedayanthittu estuary
  - The Cochin backwaters
  - Wetlands in the Andaman and Nicobar Islands
- Indonesia
  - Banyuasin Musi River Delta
  - Muara Cimanuk*
  - Sukolilo
  - Cilacap and Sagara Anakan
  - Danau Bankau, other swamps of Barito Basin*
  - Banau Sentarum
  - Wetlands in Manusela proposed national park
  - Wasur and Rawa Biru
  - Japan
  - Shonai-Fujimae Tidal Rats and Inner Ise Bay
  - Lake Shinji and Lake Nakaumi
- Korea
  - South Kanghwa and North Yongjong mudflats
  - Mudflats of South Yongjong and adjacent islands
  - Namyang Bay
  - Asan Bay
  - Kum, Mankyung and Tangjin estuaries
- Malaysia
  - Sedili Kecil swamp forest
  - Klang Islands: Pulau Ketam*
  - Kapar Forest Reserve
  - North Selangor swamp forest
  - Marintaman Mengalong*
  - Tempasuk Plain
  - Lawas mangroves
  - Trusan-Sundar mangroves
  - Limbang mangroves
  - Maludam swamp forest
  - Sarawak Mangrove Forest Reserve
- Nepal
  - Begnas Tal*
- Pakistan
  - Khabbaki Lake*
  - Siranda Lake*
  - Hawkes Bay/Sandspit Beaches and adjacent creeks
  - Clifton Beach
  - Korangi and Gharo Creeks
  - The Outer Indus Delta
- Philippines
  - Pangasinan Wetlands*
  - Manila Bay*
  - Laguna de Bay
  - Tayabas Bay including Pagbilao Bay
  - Inabanga Coast
  - Mactan, Kalawisan and Cansaga Bays*
  - Agusan marsh
  - Lake Leonard*
  - Davao Gulf
  - Liguasan marsh
- Singapore
  - Serangoon estuary*
- Sri Lanka
  - Thandamannar Lagoon*
  - Chundikkulam Lagoon
  - Chalai Lagoon*
  - Periyakarachchi and Sinnakarachchi Lagoons*
  - Mahaweli Ganga floodplain system
  - Maha Lewaya and Karagan Lewaya
  - Lunama Kalapuwa and Kalametiya Kalapuwa*
  - Bellanwilla-Attidiya marshes
- Taiwan
  - Tatu estuary
  - Tungshih (Ton-Shou) mangroves*
- Thailand
  - Gulf of Thailand
  - Pak Phanang estuary
  - Pa Phru
- Vietnam
  - Red River Delta
  - Red River estuary
  - Mekong Delta
  - Nam Can mangrove forest

*Sites are considered to be already too degraded to merit any special conservation effort.

Source: Reproduced from Scott DA and Poole CM (1989) *A Status Overview of Asian Wetlands*, No. 53 AWB. Malaysia: Kuala Lampur, via Groombridge (1992, p. 303).

In areas with higher salinity (drier climates and less freshwater inflow), mangroves are also shorter and may have saline lagoons behind them instead of more productive back mangroves.

Mangroves in Asia are considered to be the most diverse in the world and covered about 7.7 million ha at the time of the *Global Biodiversity* census (Groombridge, 1992). As shown in **Table 10**, most of this area was in Indonesia (about half in Sumatra), Malaysia (about half in Sabah), Bangladesh, the Philippines, Vietnam, India, Thailand, and Pakistan (see complete statistics in Groombridge, 1992, p. 325). Another 3 million ha belonged to Indonesia but are in Irian Jaya (New Guinea), which geologically is part of Australia, not Asia. These statistics are old, and mangroves have been declining steadily and often rapidly, being converted for development, urban sprawl, or mining (tin); cut for wood; or simply degraded by overuse.

On higher-energy tropical and subtropical coastlines, where the substrate is usually sand, the first beach pioneer inland from the water line is usually the creeping vine *Ipomoea pes-caprae* (Convolvulaceae). Behind this the first stable vegetation is often a dense thicket of large-sclerophyll shrubs, *Scaevola* (Goodeniaceae) and sometimes *Messerschmidtia* (Boraginaceae), plus some pandans and various smaller species which elsewhere might pass as weeds. On many sandy tropical shorelines, coconut palms (*Cocos nucifera*) have established naturally or been planted and cultivated. Strand forests may develop behind beaches and on raised coral terraces, dominated by tree species such as *Pongamia pinnata*, *Ficus microcarpa*, and *Bischoffia javanica*.

At higher latitudes, where winter cold precludes tropical taxa and the previously mentioned systems, salt marshes are the most common shoreline ecosystem. These are also highly productive ecosystems, rivaling tropical rain forests in some cases, and can cover quite wide areas along more concave shorelines, where the intertidal range is higher and thus reaches farther inland. In east Asia, salt marshes extend as far north as the Russian Far East north of Vladivostok. Of course, estuaries also extend to higher latitudes and have been the basis for human settlement in coastal areas for millennia because of their abundance of fish and shellfish. All of these coastline ecosystems, though narrow, are important as buffers between the ocean and the more widespread, typical forests behind them.

Inland, the western Siberian lowland (between the Urals and the Yenisey River, about one-third of all of Siberia) is mostly a vast boreal swamp and may rank as the world's largest single wetland. Boreal landscapes in general, however, are often slowly shifting mosaics of forested uplands and lower lying fens and bogs. In the cold climate most bogs accumulate peat, which raises the land surface such that former depressions become convex surfaces fed only by nutrient-poor precipitation rather than groundwater. These are the highly acidic "raised bogs" which contain mainly ericaceous dwarf-shrubs (e.g., *Oxycoccus* and *Andromeda*), sedges (e.g., *Eriophorum*), and mosses, especially *Sphagnum* species. Around the edges of these (and other) bogs may be stunted birches, larches, and other trees tolerant of saturated soils and high acidity. Although individually small and not stable over long time scales, these bogs cover a very large area and store much carbon, which may be released to the atmosphere as global warming warms and dries these areas.

Bogs and (less acidic) fens also occur in the temperate zone to the south but especially in moist, cooler areas such as northern Japan. More important in drier climates are probably

Table 12 Original and remaining areas of tropical moist forest in Asia (km^2)[a]

Location	*Original extent*	*Remaining extent*			*% Moist forest remaining*	
		Atlas maps[b]	*Map date*	*FAO (1988)*[c]	*Map data*	*FAO (1988)*
Southeast Asia						
Brunei	5000	4692	1988	3230	94	65
Cambodia	160,000	113,250	1971	71,680	71	45
China/Taiwan	340,000	25,860	1979	125,860	8	—
Indonesia	1,700,000	1,179,140	1985–89	1,138,950	69	67
Laos	225,000	124,600	1987	78,100	55	35
Malaysia	320,000	200,450	—	209,960	63	66
Peninsular	(130,000)	(69,780)	1986	—	54	—
Sabah	(70,000)	(36,000)	1984	—	51	—
Sarawak	(120,000)	(94,670)	1979	—	79	—
Myanmar	600,000	311,850	1987	313,090	52	52
Philippines	295,000	66,020	1988	95,100	22	32
Singapore	500	20	(1980s)	—	4	—
Thailand	250,000	106,900	1985	83,350	43	33
Vietnam	280,000	56,680	1987	75,700	20	27
South Asia						
Bangladesh	130,000	9730	1981–86	9270	7	7
India	910,000	228,330	1986	504,010	25	55
Sri Lanka	26,000	12,260	1988	16,590	47	64

[a]Subtotals for subunits are in parentheses.
[b]Rain forests and monsoon forests.
[c]Closed broad-leaved plus coniferous forests.
Source: Data are from the Global Biodiversity report (Groombridge, 1992, p. 262).

the freshwater marshes, which are of great importance to birds migrating between wintering areas in tropical Asia (e.g., India) and summer breeding grounds in Siberia. Such marshes may occur in depressions as well as along streams, especially in the deserts of western China, the grasslands of northern China and Mongolia, and in Middle Asia and the highlands of Tibet. These marshes are composed primarily of reeds (*Phragmites*), cattails (Typha), sedges (*Cyperus, Scirpus*, and *Carex*), and similar herbaceous plants.

Among lakes and streams, the important ecological distinction is between more aerated, nutrient-rich flowing waters and standing waters which can become quite oligotrophic, especially in cooler climates. Lake Baikal, in southern Siberia, is the world's deepest lake (1620 m) and contains about 20% of all the world's (unfrozen) fresh water. The lake contains about 600 plant species (35% endemic) and 1200 animal species (60% endemic). It is also large enough to produce a distinct limnoclimate around its edges, in which species from boreal and polar areas occur in small disjunct populations. Another well-known lake in Asia is Dal Lake in Kashmir, which has a flat bottom, is only a few meters deep, and has been filling rapidly over the past decades with sediment from

Table 13 Total numbers of threatened plant and vertebrate species by country

Country	*Plants*	*Mammals*	*Birds*	*Reptiles*	*Amphibians*	*Fish*
Mediterranean–Southwest Asia						
Afghanistan	4	13	13	1	1	0
Cyprus	43	1	17	1	0	0
Iran	301	15	20	4	0	2
Iraq	1	9	17	0	0	2
Israel	3	8	15	1	1	0
Jordan	752	5	11	0	0	0
Lebanon	5	4	15	1	0	5
Syria	11	4	15	1	0	0
Turkey	1944	5	18	5	1	5
Arabian Peninsula						
Bahrain	0	1	4	0	0	1
Kuwait	1	5	7	0	0	0
Oman	2	6	8	0	0	2
Qatar	0	0	3	0	0	0
Saudi Arabia	2	9	12	0	0	0
United Arab Emirates	0	4	7	0	0	0
Yemen	134	6	9	0	0	0
Tropical Asia						
Bangladesh	33	15	27	14	0	0
Bhutan	15	15	10	1	0	0
British Indian Ocean Territory	0	0	0	0	0	0
Brunei	40	9	10	3	0	2
Cambodia	11	21	13	6	0	5
India	1336	39	72	17	3	2
Indonesia	70	49	135	13	0	29
Laos	3	23	18	5	0	5
Malaysia	522	23	35	12	0	6
Maldives	0	1	1	0	0	0
Myanmar	—	23	42	10	0	2
Nepal	33	22	20	9	0	0
Pakistan	14	15	25	6	0	0
Philippines	159	12	39	6	0	21
Singapore	19	4	5	1	0	1
Sri Lanka	220	7	8	3	0	12
Thailand	68	26	34	9	0	13
Vietnam	338	28	34	8	1	4
East Asia						
China	350	40	83	7	1	7
Hong Kong	5	1	9	2	0	0
Japan	41	5	31	0	1	3
Korea (North)	0	5	25	0	0	0
Korea (South)	33	6	22	0	0	0
Mongolia	0	9	13	0	0	0
Taiwan	95	4	16	0	0	0

Source: Data are from the *Global Biodiversity* report (Groombridge, 1992, p. 239).

the surrounding mountains, which have been denuded largely since the 1950s. This and many other lakes are also extremely important for migrating birds. In temperate and tropical Asia, the large rivers are especially important, even though their ecosystems are often badly degraded or at least greatly altered by introductions of exotic species.

A listing of the most seriously threatened wetlands in Asia is shown in **Table 11**. Some sites are denoted as already too degraded to merit any special conservation effort. Many other wetlands are not on this priority list but were or have since become threatened to the point that their biodiversity as well as their benefits to local human societies are being rapidly lost.

Artificial Ecosystems

One other class of ecosystems, especially important in long-settled Asia, is artificial ecosystems such as those involving tree plantations, urban landscapes, rice paddies, cropland, managed ponds, etc. Plantations of native conifers are especially extensive in the mountains of Japan, but much more widespread, especially in tropical Asia, are plantations of exotic eucalyptus, Australian pine (*Casuarina*), and other trees not from any part of Asia. Although these plantations, which mainly provide fuelwood, may take the pressure off more natural, native forests, they also damage the soil such that it no longer supports the normal diversity of soil and understory plants, animals, and microbes, especially decomposer organisms which break down and cycle minerals.

Rice paddies and cropland do support some degree of diversity among species which are commonly associated with certain types of management regimes. These largely weedy species may not be species of stable native ecosystems, but they do form somewhat consistent communities. Ponds and other water bodies managed for fish raising, etc. are also common in warmer parts of Asia. The species diversity is low and involves mainly exotic, weedy species such as *Myriophyllum* or water hyacinth (*Eichhornia crassipes*).

Urban landscapes are extensive and ever-growing in Asia as elsewhere. Urban ecosystems may involve allées and other plantings for landscape amelioration but may also involve seminatural wooded, grassy, and aquatic communities. Some of the methods for building semi-natural forests in densely built-up areas are quite well developed, especially in Japan (Miyawaki et al., 1987).

Conservation Status

The conservation status of natural and even of modified but functioning, quasinatural ecosystems in Asia varies widely. Some vast tracts of boreal forest appear untouched, but one cannot fly over these areas without seeing smoke (often industrial) rising from what appear to be very isolated places. Other forest areas have almost completely disappeared. It has been estimated that less than 1% of the natural evergreen broad-leaved forest area of Japan remains, and natural broad-leaved as well as conifer forests in China are confined entirely to mountain areas, where they now are also facing encroachment. Tropical forests have also largely disappeared, remaining over large areas mainly in Southeast Asia (see **Table 7**). The especially well-developed tropical rain forests of Borneo, with canopy heights commonly near 50 m, have largely been cut within the past 20 years, mostly for export to Japan. Some estimates of original and remaining area of tropical rain forest are shown in **Table 12**, but the data for remaining area (all pre-1990) must be overestimates.

Country totals of threatened plant and vertebrate species, and also fish, in Asia are summarized in **Table 13**. Although these statistics reflect different levels of coverage and different methodologies, they give an indication of the magnitude of the problem of threats to biodiversity by changing conditions—mostly changes in land use by humans. Large numbers of threatened plants are recognized only in Turkey in the west, in Yemen on the Arabian Peninsula, in China in east Asia, and in India, Malaysia, Vietnam, and Sri Lanka in tropical Asia—although many others must also be threatened in other areas. The pattern for animals is more even across countries and regions, but with Indonesia or China recognizing the largest numbers of threatened species in all five categories.

Species and ecosystems are endangered throughout Asia, as elsewhere. For further reading, see in particular the various recent summaries of the status of conservation efforts and strategies in all parts of the world (Hoyt, 1994; Heywood, 1995) as well as more standard treatments of the general problems of wildlife conservation.

See also: Africa, Ecosystems of. Ecosystems of South America. Europe, Ecosystems of. Near East Ecosystems, Animal Diversity. North America, Patterns of Biodiversity in

References

Archibold OW (1995) *Ecology of World Vegetation*. London: Chapman & Hall.
Bailey RG (1996) *Ecosystem Geography*. New York: Springer.
Box EO (ed.) with Peet RK, Masuzawa T, Yamada I, Fujiwara K, and Maycock PF (Co-eds.) (1995) *Vegetation Science in Forestry: Global Perspective Based on Forest Ecosystems of East and Southeast Asia*, Handbook of Vegetation Science, Vol. 12. Dordrecht: Kluwer.
Champion HG and Seth SK (1968) *A Revised Survey of the Forest Types of India*. Delhi: Government Printer.
Groombridge B (1992) *Global Biodiversity Status of the Earth's Living Resources*. London: World Conservation Monitoring Center/Chapman & Hall.
Heywood VH (1995) *Global Biodiversity Assessment*. Cambridge, UK: UN Environment Programme/Cambridge University Press.
Hoyt JA (1994) *Animals in Peril: How "Sustainable Use" Is Wiping Out the World's Wildlife*. Garden City, NY: Avery.
Mani MS (ed.) (1974) *Ecology and Biogeography in India*. Monographiae Biologicae, Vol. 23. The Hague: Junk.
Missouri Botanical Garden (1983) Biogeographical relationships between temperate eastern Asia and temperate eastern North America. *Ann. Missouri Bot. Garden* 70: 421–749.
Miyawaki A, Bogenrieder A, Okuda Sh, and White J (1987) Vegetation Ecology and Creation of New Environments. Proceedings of 1984 IAVS Symposium. Tokyo: Tokai University Press.
Numata M (ed.) (1974) *The Flora and Vegetation of Japan*. Tokyo/Amsterdam: Kodansha/Elsevier.

Ohsawa M (1995) Latitudinal comparison of altitudinal changes in forest structure, leaf type, and species richness in humid monsoon Asia. *Vegetatio* 121: 310.
Raven PH and Axelrod DI (1974) Angiosperm biogeography and past continental movements. *Ann. Missouri Bot. Garden* 61: 539–673.
Schaller GB (1998) *Wildlife of the Tibetan Steppe*. Chicago: University of Chicago Press.
Scott DA and Poole CM (1989) *A Status Overview of Asian Wetlands*, No. 53 AWB. Malaysia: Kuala Lampur.
Troll C (1972) *Geoecology of the High Mountain Regions of Eurasia*, Erdwissenschaftl. Forschung, Vol. 4. Wiesbaden: Franz-Steiner-Verlag.
Udvardy MDF (1975) IUCN Occasional Paper No. 18. Morges, Switzerland: International Union for the Conservation of Nature and Natural Resources.
Walter H (1974) *Die Vegetation Osteuropas, Nord- und Zentralasiens*, "Vegetation der einzelnen Groβräume" series, Vol. VII. Stuttgart: Gustav-Fischer-Verlag.
Walter H (1985) *Vegetation of the Earth and Ecological Systems of the Geobiosphere*, 3rd ed. New York: Spinger-Verlag.
Walter H and Box EO (1983) Deserts of Eurasia. In: West NE (ed.) *Temperate Deserts and Semi-Deserts*. Ecosystems of the World, Vol. 5. Amsterdam: Elsevier.
Wang CW (1961) *The Forests of China, with a Survey of Grassland and Desert Vegetation*. Maria Moors Cabot Foundation Publ. No. 5. Cambridge, MA: Harvard University.
Whitmore TC (1984) *Tropical Rainforests of the Far East*, 2nd ed. Oxford, UK: Clarendon. (1st ed. 1975).
Zhengyi Wu and Sugong Wu (1998) A proposal for a new floristic kingdom (realm)—the East Asiatic Kingdom, its delineation and characteristics. In: Zhong A and Wu S (eds.) *Floristic Characteristics and Diversity of East Asian Plants*, pp. 3–42. Beijing: China Higher Education Press; and Berlin: Springer-Verlag.
Zhengyi Wu, Sugong Wu, Zhang A-L, and Wu S-G (eds.) (1996) *Floristic Characteristics and Diversity of East Asian Plants*. Berlin: Springer Verlag.
Zohary M (1973) *Geobotanical Foundations of the Middle East*, "Geobotanica Selecta" series, Vol. 3. Stuttgart: Gustav-Fischer-Verlag.

Atmospheric Gases

Donald J Wuebbles, University of Illinois at Urbana-Champaign, Urbana, IL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Atmospheric lifetime Timescale characterizing the rate of removal of a gas from the atmosphere, generally defined as the time to remove 63.2%, the e-folding time.

Greenhouse gases Gases like water vapor and carbon dioxide that can affect climate through their absorption and re-emission of terrestrially emitted infrared radiation.

Mixing ratio Ratio of the concentration of a gas to the concentration of air. It can be defined in terms of volume (e.g., molecules per m^3) or mass (e.g., g m^{-3}).

Oxidizing capacity Self-cleansing ability of the atmosphere through oxidative reactions.

Ozone Molecule composed of three oxygen atoms that is extremely important to life on Earth because of its absorption of solar ultraviolet radiation; ozone is also a greenhouse gas.

Photochemical source or sink Production or loss of an atmospheric gas involving photodissociation or chemical reactions.

Photodissociation Destruction of a molecule through absorption of solar radiation and subsequent breaking of one or more chemical bonds.

Introduction

As seen in Table 1, the bulk of the Earth's dry atmosphere is composed of molecular nitrogen (N_2, 78.08%), molecular oxygen (O_2, 20.95%), and argon (Ar, 0.93%). While the remaining gases comprise only about 0.03% of the Earth's atmosphere, many of them are extremely important to the climate and habitability of our planet. These gases are the primary focus of this article. The concentrations of many of these gases are also being affected by human activities.

Some of the gases listed in Table 1, like ozone (O_3), affect the transmission of solar radiation in the atmosphere. Others, like carbon dioxide (CO_2), affect the absorption of terrestrially produced infrared radiation and are called greenhouse gases. These types of gases are therefore linked to the physical climate system of the Earth. Increasing concentrations of greenhouse gases have resulted in concerns about changes in climate, or global warming as it is often referred to in the public media. Some trace gases, such as sulfur dioxide (SO_2) and hydrogen sulfide (H_2S), can affect climate through their role as precursors in the production of atmospheric particles. The resulting sulfate particles can scatter solar radiation, thus preventing some of it from reaching the Earth's surface. These particles can also serve as condensation nuclei in the formation of clouds.

The ozone layer is a term that refers to the distribution of ozone that is naturally formed in the stratosphere. This layer protects life on Earth from harmful levels of solar ultraviolet radiation (UV-B). Chlorofluorocarbons (CFCs) and other chlorinated and brominated halocarbons are emitted by a variety of human activities. Atmospheric measurements have clearly corroborated theoretical studies showing that the chlorine and bromine released from the destruction of these halocarbons in the stratosphere have been reacting to destroy significant amounts of ozone over the last few decades.

Gases like hydroxyl (OH), ozone, and hydrogen peroxide (H_2O_2) control the oxidizing capacity of the atmosphere, and destroy many pollutants emitted into the atmosphere through chemical reactions. Hydroxyl has been referred to as the "atmospheric vacuum cleaner." It is generated primarily by the interactions of water vapor, ozone, and UV-B, and is destroyed primarily through reaction with carbon monoxide (CO) and methane (CH_4).

Emissions of trace gases in urban regions can result in atmospheric pollution referred to as photochemical smog. For example, ozone, a potential cause of health problems in polluted regions, can be formed in the air by chemical reactions following emissions of nitrogen oxides (NO_x) and volatile organic gases (VOCs), including hydrocarbons. NO_x is the sum of the concentrations of nitric oxide (NO) and nitrogen dioxide (NO_2), while NO_y also includes other reactive nitrogen gases like nitric acid (HNO_3).

Missing from Table 1 is water vapor, a gas with highly variable concentration, whose amount in the atmosphere is dependent on evaporation and hydrological cycle processes associated with weather. The concentration of atmospheric water vapor can be affected at the local scale by human activities but is not thought to be directly affected at the global scale by human activities. However, climate change can affect the amount of water vapor; the atmosphere can hold more water vapor under a warmer climate. Water vapor is also the most important of the greenhouse gases; its changes in concentration are generally considered in climate studies as being part of the climate response to other forcings, either natural or human-driven.

Biodiversity affects the concentrations of atmospheric trace gases and can also be affected by changing concentrations of atmospheric gases. The biosphere is an important component in the biogeochemical cycles of carbon, nitrogen, and sulfur. As a result, the sources and sinks of many atmospheric gases are strongly affected by changes in the biosphere. For example, deforestation is contributing to the increase in concentration of carbon dioxide. Also, biospheric growth from increasing amounts of atmospheric carbon dioxide can increase the rate of removal of CO_2 from the atmosphere. Local air pollution, changes in climate, and changes in stratospheric ozone can

Table 1 Typical concentrations and primary sources of some of the important gases in the atmosphere (based on WMO, 2010, and Blasing, 2010; wherever possible, values correspond to the globally averaged concentration in 2009)

Constituent	*Chemical formula*	*Volume mixing ratio (dry air)*[a]	*Major sources and remarks*
Nitrogen	N_2	78.084%	Biological
Oxygen	O_2	20.948%	Biological
Argon	Ar	0.934%	Inert
Carbon dioxide	CO_2	386 ppm	Combustion, ocean, biosphere
Methane	CH_4	1.8 ppm	Biogenic, human-related
Hydrogen	H_2	0.55 ppm	Biogenic, human-related, photochemical
Nitrous oxide	N_2O	0.32 ppm	Biogenic, human-related
Carbon monoxide	CO	0.05–0.2 ppm	Photochemical, human-related
Ozone (troposphere)	O_3	0.01–0.5 ppm	Photochemical
Ozone (stratosphere)	O_3	0.5–10 ppm	Photochemical
Nonmethane hydrocarbons	C_2H_2, etc.	5–20 ppb	Biogenic, human-related
CFC-11	$CFCl_3$	240 ppt	Human-related
CFC-12	CF_2Cl_2	536 ppt	Human-related
CFC-113	CCl_2FCClF_2	76 ppt	Human-related
Halon-1211	$CBrClF_2$	4.3 ppt	Human-related
Halon-1301	$CBrF_3$	3.3 ppt	Human-related
Methyl bromide	CH_3Br	10 ppt	Ocean, human-related
Methyl chloride	CH_3Cl	550 ppt	Ocean, biogenic, human-related
Nitrogen species	NO_y	10 pptv–1 ppm	Soils, lightning, human-related
Ammonia	NH_3	10 pptv–1 ppb	Biogenic
Hydroxyl	OH	0.1 pptv–10 ppt	Photochemical
Peroxyl	HO_2	0.1 pptv–10 ppt	Photochemical
Hydrogen peroxide	H_2O_2	0.1–10 ppb	Photochemical
Formaldehyde	CH_2O	0.1–1 ppb	Photochemical
Sulfur dioxide	SO_2	10 pptv–1 ppb	Photochemical, volcanic, human-related
Di-methyl sulfide	CH_3SCH_3	10–100 ppt	Biogenic
Carbon disulfide	CS_2	1–300 ppt	Biogenic, human-related
Carbonyl sulfide	OCS	500 ppt	Biogenic, volcanic, human-related
Hydrogen sulfide	H_2S	5–500 ppt	Biogenic, volcanic

[a]ppm = parts per million by volume = number of molecules in 10^6 molecules of air; ppb = parts per billion by volume = number of molecules in 10^9 molecules of air; ppt = parts per trillion by volume = number of molecules in 10^{12} molecules of air.

affect the biosphere and have an impact on biodiversity. For example, acid rain or acidification resulting from emissions of nitrogen- and sulfur-containing gases can affect landscapes in many ways.

The Changing Composition

Without human intervention, concentrations of many atmospheric gases would be expected to change slowly. Ice core measurements of the gases trapped in ancient ice bubbles indicate that this was the case before the last century. We also know that major natural events, like a volcanic eruption, can enhance concentrations of some gases, such as sulfur dioxide, for as long as a few years. However, since the beginning of the Industrial Age, emissions associated with human activities have risen rapidly. Agriculture, industry, waste disposal, deforestation, and especially fossil fuel use have been producing increasing amounts of carbon dioxide, methane, nitrous oxide (N_2O), CFCs, and other important gases. Because of these increasing emissions, atmospheric levels of these gases have been building at an unprecedented rate, resulting in effects on ozone, another radiatively important gas, and raising concerns regarding the impact of these gases on climate. Although all of these gases are having some influence on climate, changes in the amount of atmospheric carbon dioxide is of particular interest in the major concerns about a human-induced effect on temperature, precipitation, and other aspects of our climate system. In addition, changes in the human-related emissions and atmospheric concentrations of small particles or aerosols are also influencing climate. The following discussion focuses on the changing concentrations and budgets of a number of these gases.

Carbon Dioxide

Of all the greenhouse gases, carbon dioxide has been undergoing the largest changes in concentration. It is also the gas of most concern for analyses of the potential human effects on climate. Accurate measurements of atmospheric CO_2 concentration began in 1958 at the Mauna Loa Observatory in Hawaii. **Figure 1** shows that the annually averaged concentration of CO_2 in the atmosphere has risen from 316 parts per million (ppm) in 1959 to 386 ppm in 2009. The CO_2 measurements exhibit a seasonal cycle, which is mainly caused by the seasonal uptake and release of atmospheric CO_2 by terrestrial ecosystems. The average annual rate of increase over the whole time period is about 1.2 ppm or 0.4% per year, with

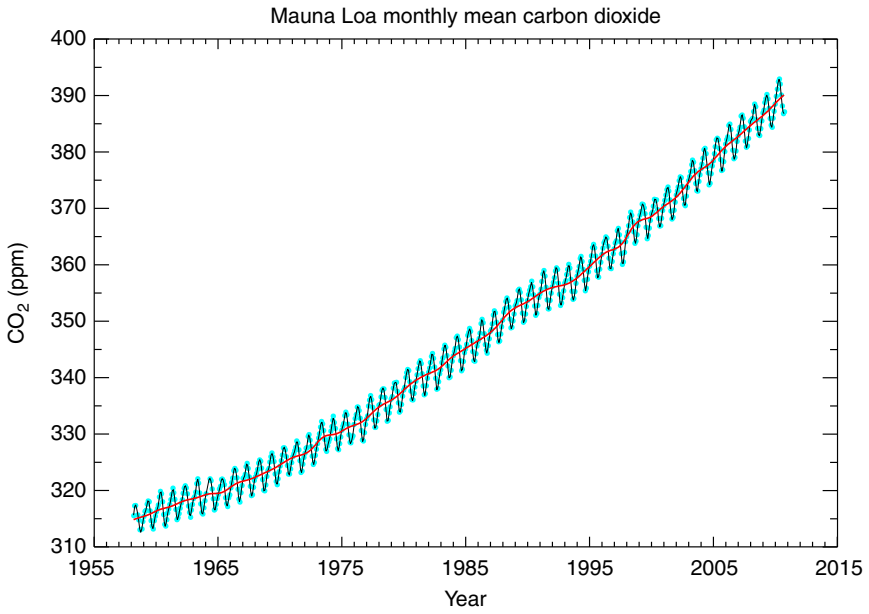


Figure 1 Observed monthly average of carbon dioxide (CO_2) concentration (ppmv, parts per million by volume) from Mauna Loa, Hawaii (measurements began by CD Keeling and continue under the auspices of the US National Oceanic and Atmospheric Administration (NOAA)). Seasonal variations are primarily due to the uptake and production of CO_2 by the terrestrial biosphere.

the rate of increase over the last decade being about 1.6 ppm yr^{-1}. A significant slowing in the growth rate occurred in 1992 (when it dropped to 0.6 ppm), but by 1994 the rate had risen above the average over the decade. This anomaly in the carbon cycle system appears to be related to atmospheric aerosols resulting from the Mount Pinatubo volcanic eruption in June 1991 and the corresponding effects on solar irradiances and natural emissions/uptake of CO_2. A rate increase in 1998 was related to biospheric response to a major El Nino event.

Measurements of CO_2 concentration in air trapped in ice cores indicate that the preindustrial concentration of CO_2 was approximately 280 ppm. These data show that carbon dioxide concentrations fluctuated by ± 10 ppm around 280 ppm for well over a 1000 years until the recent increase to the current 380+ ppm, an increase is greater than 30%.

Why has the atmospheric concentration of CO_2 increased so dramatically? Analyses with models of the atmosphere–ocean–biosphere system of the carbon cycle indicate that human activities are primarily responsible for the increase in CO_2. Two types of human activities are primarily responsible for emissions of CO_2: fossil fuel use, which released about 6.0 gigatons of carbon (ppb; GtC a gigaton is 10^9 tons) into the atmosphere in 1990, and land use, including deforestation and biomass burning, which may have contributed about 1.6 ± 1.0 GtC in addition to that from fossil fuels. Since 1751 approximately 337 billion tons of carbon have been released to the atmosphere from the consumption of fossil fuels and cement production. Half of these emissions have occurred since the mid-1970s. By 2007, the CO_2 contribution from fossil fuel burning was 8.3 GtC yr^{-1} (Boden *et al.*, 2010), an increase of 1.7% relative to 2006.

This added atmospheric carbon resulting from human activities is redistributed within the atmospheric, oceanic, and biospheric parts of the global carbon cycle, with the dynamics of this redistribution determining the corresponding rise in atmospheric CO_2 concentration. About 40% of the added CO_2 from human activities remains in the atmosphere while about 60% is transferred to the oceans and terrestrial biosphere.

Analyses of the carbon budget previously implied a mismatch between observed levels of CO_2 and known loss processes. This sink has been well explained in terms of increased net carbon storage by the terrestrial biomass stimulated by the CO_2 fertilization effect (increased plant growth in an atmosphere of higher CO_2 concentration) and other biospheric processes.

Emissions from Fossil Fuel Consumption

Carbon dioxide is emitted when carbon-containing fossil fuels are oxidized by combustion. Carbon dioxide emissions depend on energy and carbon content, which ranges from 13.6 to 14.0 MtC EJ^{-1} for natural gas, 19.0–20.3 for oil, and 23.9–24.5 for coal. Other energy sources, such as hydro, nuclear, wind, and solar, have no direct carbon emissions. Biomass energy, however, is a special case. When biomass is used as a fuel, it releases carbon with a carbon-to-energy ratio similar to that of coal. However, the biomass has already absorbed an equal amount of carbon from the atmosphere prior to its emission, so that net emissions of carbon from biomass fuels are zero over its life cycle.

Human-related emissions from fossil fuel use have been estimated as far back as 1751. Before 1863, emissions did not exceed 0.1 GtC yr^{-1}. However, by 2007 they had reached 8.3 GtC yr^{-1}, giving an average emission growth rate slightly greater than 3% per year over the last two and a half centuries. Recent growth rates have been significantly lower, at 1.8% per year between 1970 and 1995. Emissions were initially dominated by coal. Since 1985, petroleum products have been the

main source of emissions despite their lower carbon content. The regional pattern of emissions has also changed. Global emissions were once dominated by Europe and North America, but now developing nations, especially China and India, are providing an increasing share of emissions. By 2007, developing countries (including China and India as part of the Annex I nations being used in international policy discussions) account for more than half of global emissions.

Emissions from Land Use Changes

The biosphere holds approximately 560 GtC in the form of aboveground biomass, and an additional 1200 GtC in soils and detritus. These pools form the principal reservoirs from which terrestrial systems can exhaust or sequester carbon. Evaluations of carbon releases from vegetation and soils based on changes in land use indicate that the global net flux during the period 1850–2000 was 148.6 Pg C, about 55% of which was from the tropics (Houghton, 2008). During the period 1990–2005, the greatest regional flux was from South and Central America (11.3 Pg C). The global total flux averaged 1.5 Pg C yr^{-1} during the 1980s and 1.56 Pg C yr^{-1} during the 1990s, dominated by fluxes from tropical deforestation. The global total flux averaged 1.47 Pg C yr^{-1} during the period from 2000 to 2005.

Methane

Although methane's atmospheric abundance is less than 0.5% that of CO_2 on a molecule-by-molecule basis, a molecule of CH_4 is approximately 50 times more effective as a greenhouse gas in the current atmosphere than CO_2. When this is combined with the large increase in its atmospheric concentration, methane becomes the second most important greenhouse gas of concern to climate change. Based on analyses of ice cores, the concentration of methane has more than doubled since preindustrial times. The current globally averaged atmospheric concentration of methane is about 1.8 ppm (**Figure 2**). Continuous monitoring of methane trends in ambient air indicated that from 1979 to 1989, the concentration increased at an average of about 16 parts per billion (ppb by volume) or 1% per year, but has declined since then, reaching roughly zero-change in the early 2000s. Since 2006, the concentration of CH_4 has been increasing at a rate of roughly 5 ppb yr^{-1}.

There are a number of ideas to explain these rapid changes in the growth rate, ranging from suggested reductions in emissions from anthropogenic or natural sources to a slowing in the rate of CH_4 removal. However, the best indications are that the sharp increase and the subsequent dip in the early 1990s were connected to changes in atmospheric chemistry and temperature induced by the 1991 Pinatubo eruption. Yet the cause of the longer term global decline in methane growth is still not well understood, although it may be that much of the earlier rapid increase in methane emissions from agricultural sources has been slowing down. The recent increase in emissions is also not well understood.

The Methane Budget

Methane emissions come from a number of different sources, both natural and anthropogenic. One type of human-related emissions arises from biogenic sources from agriculture and waste disposal, including enteric fermentation, animal and human wastes, rice paddies, biomass burning, and landfills. Emissions also result from fossil fuel-related methane sources such as natural gas loss, coal mining, and the petroleum industry. Methane is emitted naturally by wetlands, termites, other wild ruminants, oceans, and hydrates. Based on recent estimates, current human-related biogenic, and fossil fuel-related sources for methane are approximately 275 and 100 tetragrams, or 10^{12} g) CH_4 yr^{-1}, respectively, while total natural sources are around 170 $TgCH_4$ yr^{-1}.

Because of the variety of methane sources, emissions are affected by numerous factors, including energy use, human population distributions, agricultural practices, and climate. These factors complicate the resolution of past emissions and make predictions of future CH_4 emissions difficult. Estimates of historical methyne emissions can be obtained by analyses of past measurements of atmospheric CH_4 concentrations and by analyses of emissions from the different sources. Such studies suggest that, although methyne emissions have climbed rapidly over the past four decades, the relative importance of agricultural sources may be declining, whereas the importance of nonagricultural sources related to fossil fuel use and waste disposal is on the increase.

In contrast to the numerous sources of methane, there are only one major and two minor sinks for tropospheric methane. Reaction with the hydroxyl radical is responsible for the removal of approximately 510 $TgCH_4$ yr^{-1} (88% of the total sink). The remainder of the CH_4 is removed through reactions with soil (30 $TgCH_4$ yr^{-1}, or ~5%) or transport to the stratosphere (40 $TgCH_4$ yr^{-1}, or ~7%).

Nitrous Oxide

Nitrous oxide is a greenhouse gas that on a molecule-to-molecule basis is 200 times more efficient than CO_2 in absorbing infrared radiation. Also, through reactions with excited oxygen atoms, N_2O is the primary source of the nitrogen oxides that account for a significant fraction of the natural destruction of ozone in the stratosphere. Atmospheric measurements and ice core data indicate a continuous increase in N_2O from a preindustrial concentration of about 275 ppb. In 2007, the mean atmospheric concentration of N_2O was about 320 ppb, with a current growth rate of 0.2–0.3% per year.

Nitrous oxide is produced by various natural and human-related sources, most of which are not well quantified. Current estimates of mean emissions from natural sources include: oceans, 3.8 TgN yr^{-1}; and soils, 6.6 TgN yr^{-1}. Anthropogenic sources include agriculture, 2.8 TgN yr^{-1}; biomass burning, 0.7 TgN yr^{-1}; industrial sources, 0.7 TgN yr^{-1}; and river and estuaries, 1.7 TgN yr^{-1}. Large uncertainties associated with the numerous small sources that make up the N_2O budget make it difficult to fully explain its increase in concentration.

The major sink for N_2O is photodissociation by sunlight in the stratosphere. The current best estimate for stratospheric removal, based on stratospheric chemistry modeling, is 12.5 TgN yr^{-1}. There is some evidence that N_2O is consumed by certain soils, but there are not enough data to make a reasonable global estimate of this sink.

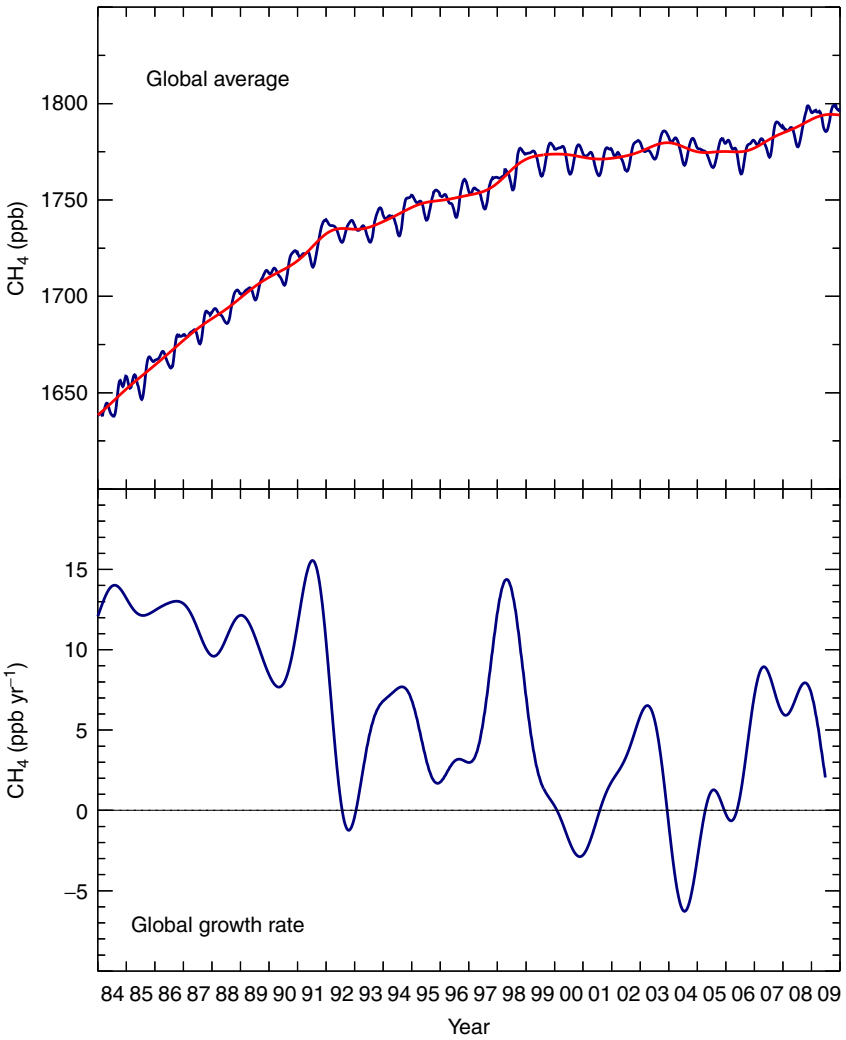


Figure 2 (Top) Globally averaged atmospheric methane (CH_4) concentrations (parts per billion (ppb) by volume) derived from the NOAA Climate Monitoring and Diagnostics Laboratory air-sampling sites by Dlugokencky EJ, Bruhwiler L, and White JWC (2009) Observational constraints on recent increases in the atmospheric CH_4 burden. *Geophysical Research Letter* 36: L18803. doi:10.1029/2009GL039780. The solid line is a deseasonalized trend curve fitted to the data. The dashed-line is a model-estimated (that accounts for CH_4 emissions and loss in the atmosphere) calculated trend that fit to the globally average values. (Bottom) Atmospheric CH_4 instantaneous growth rate (ppb^{-1} yr), which is the derivative with respect to the trend curve shown in the bottom panel.

Chlorofluorocarbons and Other Halocarbons

Halocarbons are greenhouse gases that can contribute to climate change, and they are also largely responsible for stratospheric ozone loss over recent decades. Because of their dual impact, they are of particular concern to environmental health. Among the most potent halocarbons in the current atmosphere are the CFC's CFC-11 ($CFCl_3$) and CFC-12 (CF_2Cl_2). One molecule of CFC-11 or CFC-12 in the atmosphere is, respectively, 12,400 and 15,800 times more effective as a greenhouse gas than one molecule of CO_2. With the exception of the naturally occurring portions of CH_3Cl and CH_3Br emissions, all the halocarbons in the atmosphere are man-made. Their inertness and long lifetimes have made them attractive chemicals for use as propellants, refrigerants, fire retardants, and other industrial applications. Natural sources of CH_3Cl from ocean surface waters and wood-rotting fungi may account for as much as half of atmospheric CH_3Cl, with biomass burning (some of which is natural) being the other primary source. Roughly half of the CH_3Br comes from the oceans, while biomass burning and its use as a soil fumigant account for the majority of its remaining emissions into the atmosphere.

Halocarbons containing chlorine or bromine are of particular concern with regard to destruction of stratospheric ozone. Bromine and chlorine effectively catalyze ozone destruction cycles. These halocarbons, plus others such as the perfluorocarbons (PFCs) and the hydrofluorocarbons (HFCs) that contain fluorine instead of chlorine, also have the potential to affect climate change since these chemical species characteristically have strong infrared absorption features.

Measurements of CFCs and other compounds by the NOAA Climate Monitoring and Diagnostics Laboratory at sites throughout the world are shown in **Figure 3**. CFC-11 and CFC-12 have the largest atmospheric concentrations, at 0.24 and

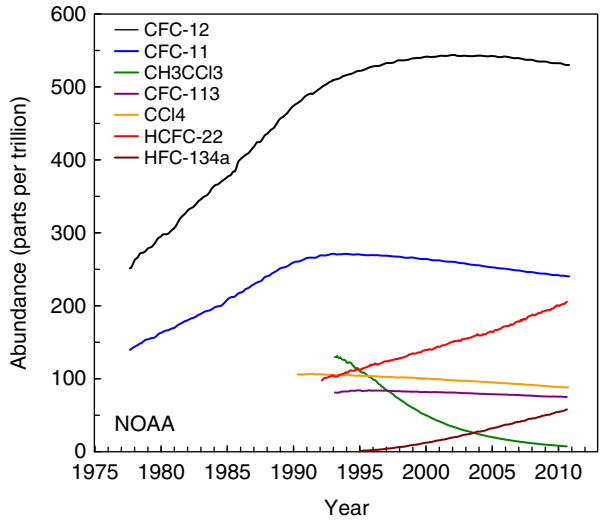


Figure 3 Changes in measured concentrations of various important halogenated compounds over time, showing the initial rise until the last few decades and the beginning of a recent decline in several of the important compounds. Data are based on measurements from NOAA (these NOAA data courtesy of J Elkins, S Montzka, and G Dutton).

0.54 ppb, respectively. The tropospheric concentrations of both of these gases were increasing at about 4% per year in the early 1990s, but are now declining. The use of these compounds has diminished greatly, with all but a few minor essential applications being banned as a result of the international agreement of the Montreal Protocol to protect the stratospheric ozone layer. The atmospheric concentrations of CFC-113 ($C_2F_3Cl_3$) were increasing at about 10% per year in the early 1990s but has also slowed greatly, with a current concentration of about 0.08 ppbv. Abundances of the shorter-lived controlled compound methyl chloroform (CH_3CCl_3), have declined appreciably. Hydrochlorofluorocarbon (HCFC) 22 (CHF_2Cl), a refrigerant often found in home air conditioners, has gained increased use as a replacement for CFCs and its concentration had been increasing at about 5–6% per year since 1995 but has slowed down recently. The measured concentrations of several other replacements are also increasing.

All of the fully halogenated CFCs have long atmospheric lifetimes. The lifetime of CFC-11 is about 50 years, while the lifetime of CFC-12 is about 100 years. The atmospheric lifetimes of HCFCs, HFCs, and other halocarbons containing hydrogen tend to be much shorter than those of the CFCs. Because of these shorter lifetimes, less of the replacement compounds containing chlorine and bromine reach the stratosphere, and they have less effect on ozone than the CFCs.

Bromine is more effective at destroying ozone than chlorine. Therefore, early rapid increases in atmospheric concentrations of bromine-containing halons, most notably Halon-1301 (CF_3Br) and Halon-1211 (CF_2ClBr), have caused concern. Despite their control under the 1997 Montreal Protocol on Substances That Deplete the Ozone Layer, the concentrations of several halons, including H-1211 and H-1301, continue to increase. Primary destruction of these compounds occurs through photolysis, resulting in long atmospheric lifetimes (65 years for H-1301 and 20 years for H-1211). However, these compounds currently have small atmospheric concentrations, about 4 ppt or less, and hence their contribution to the absorption of infrared radiation is considered minimal.

Other Gases

Carbon Monoxide

In addition to its potential effects on health in polluted areas, carbon monoxide affects the oxidizing capacity of the atmosphere through its reactivity. The reaction of carbon monoxide with OH is the primary sink for atmospheric OH. This reaction also yields an additional source of the greenhouse gas CO_2.

The global emissions of CO are still poorly understood. Sources of carbon monoxide include incomplete combustion processes (complete combustion yields CO_2 rather than CO). Emissions of CO from fossil fuel combustion peak between 30° and 60° N latitude. More than 70% of the CO from the biomass burning source is emitted in tropical regions. Chemical decomposition of methane and other hydrocarbons is also an important source. It is estimated that about two-thirds of current atmospheric CO results from human activities. Total sources are about 1100 TgC yr^{-1} with a large uncertainty range. As a result of its short lifetime (2–3 months) and the high spatial variability of its sources, the atmospheric concentration of CO varies greatly in time and space. Annually averaged concentrations of CO peak at about 0.2 ppm at high northern latitudes; minimum concentrations of about 0.06 ppm occur throughout most of the Southern Hemisphere. Long-term trends up to 1990 suggest that CO concentrations had been increasing in the Northern Hemisphere at about 1% per year, with little evidence of an increase in the Southern Hemisphere. After 1990, concentrations show little change.

Nonmethane Hydrocarbons

Nonmethane hydrocarbons (NMHCs) in the presence of nitrogen oxides (NO_x) contribute to the formation of tropospheric ozone and stratospheric H_2. They can also react with nitrogen species to produce the atmospheric gas peroxyacetyl nitrate (PAN), which is a long-lived reservoir of reactive nitrogen. Oxidation of NMHCs is an additional source of CO and ultimately CO_2.

There are many anthropogenic and natural sources emitting NMHCs into the atmosphere. Anthropogenic sources result from all aspects of human activity, including chemical manufacturing, vehicle exhaust, food processing, refuse disposal, biomass burning, and energy production. Global emissions from anthropogenic activities are estimated at roughly 140 Tg yr^{-1}. Emissions from the natural sources are less well known.

Nitrogen Oxides

Emissions of nitrogen oxides have been a cause of concern because of their role as primary pollutants in photochemical smog and their contribution to acid wet and dry deposition. Nitrogen oxides are also important because of their indirect effect on climate through their role in affecting global ozone

concentrations. Although NO_x species are relatively short-lived, they can react chemically with hydrocarbons to produce PAN. The constituent PAN provides a reservoir for nitrogen oxides that can be transported long distances to affect ozone chemistry well downstream from the sources. The lifetime of PAN depends strongly on temperature.

There is significant uncertainty in the sources of nitrogen oxides, with the total source being estimated to be roughly 42–47 TgN yr^{-1}. The largest sources of reactive nitrogen in the troposphere are fossil fuel combustion, biomass burning, lightning discharges, microbial activity in soils, aircraft emissions, and transport from the stratosphere. Emissions of nitrogen oxides from combustion of fossil fuel have increased globally at 1–2% per decade during this century, resulting in increased tropospheric concentrations particularly over continents in the lower atmosphere and in the flight corridors used by commercial aircraft. Nitrogen dioxide is an important absorber of visible solar radiation, and it could affect climate directly if tropospheric and stratospheric concentrations continue to increase.

Sulfur Gases

Emissions of sulfur dioxide and other gases can result in the formation of aerosols that can affect climate. Aerosols affect climate directly by absorption and scattering of solar radiation and indirectly by acting as cloud condensation nuclei (CCN). A variety of analyses indicate that human-related emissions of sulfur, and the resulting increased sulfuric acid concentrations in the troposphere, may be cooling the Northern Hemisphere sufficiently to compensate for much of the warming expected from greenhouse gases. Volcanic emissions can influence climate for short periods (1–3 years) through emissions of sulfur dioxide into the lower stratosphere.

More than half of the SO_2 emitted into the atmosphere comes from human-related sources, mainly from the combustion of coal and other fossil fuels. Most of these emissions occur in the Northern Hemisphere. Analyses indicate that anthropogenic emissions have grown dramatically during this century. Other SO_2 sources come from biomass burning, from volcanic eruptions, and from the oxidation of di-methylsulfide (DMS) and hydrogen sulfide (H_2S) in the atmosphere. DMS and H_2S are primarily produced in the oceans. Atmospheric SO_2 has a lifetime of less than a week, leading to formation of sulfuric acid and eventually to sulfate aerosol particles. Gas-to-particle conversion can also occur in cloud droplets; when precipitation does not occur soon, the evaporation of such droplets can then leave sulfate aerosols in the atmosphere.

During periods of low volcanic activity, carbonyl sulfide (COS) is thought to be responsible for the maintenance of the sulfuric aerosol layer found in the lower stratosphere. Natural emissions explain most of the COS in the present atmosphere, while the relatively long atmospheric lifetime (about 2 years) of COS explains why much of it reaches the stratosphere before its conversion to sulfuric acid aerosol. However, if sources of COS (or its precursor, CS_2) were to increase dramatically, the background aerosol layer concentration would increase, with significant implications for climate.

The Importance of Stratospheric Ozone

Ozone (O_3) is composed of three oxygen atoms and is a gas at atmospheric pressures and temperatures. Most of the ozone (about 90%) exists in the stratosphere, the layer of the atmosphere about 10–50 km above the Earth's surface. The remaining ozone is in the troposphere, the lower region of the atmosphere extending from the Earth's surface up to roughly 10 km at mid-latitudes and 16 km in the tropics. **Figure 4** shows a typical integrated column of ozone, referred to as the total ozone column, for various seasons based on satellite observations. Despite the fact that the primary production of ozone occurs in the tropics and the mid-latitudes, the largest amounts of ozone are found at high latitudes as a result of the pole-ward transport of ozone by atmospheric dynamical processes. The large decrease in ozone over springtime Antarctica, termed the "Antarctic ozone hole," can also be seen in **Figure 4**.

Ozone in the troposphere and stratosphere is chemically identical, but it has very different effects on life on the Earth depending on its location. Stratospheric ozone plays a beneficial role by absorbing UV-B, thus preventing biologically harmful levels of UV radiation from reaching the Earth's surface. It is the absorption of solar radiation by ozone that explains the increase in temperature with altitude in the stratosphere. Concerns about increased UV-B from the decreasing levels of ozone have been the driver for policy actions to protect the ozone layer.

Ozone is also a greenhouse gas, with a large infrared absorption band in the atmospheric window, at 9.6 μm. It is the balance between the solar and infrared radiative processes that determines the net effect of ozone on climate. Increases in ozone in the stratosphere above about 30 km tend to decrease the surface temperature as a result of the increased absorption of solar radiation, effectively decreasing the solar energy that would otherwise warm the Earth's surface. Below 30 km, increases in ozone tend to increase the surface temperature, and the infrared greenhouse effect dominates in this region.

Closer to the Earth's surface, ozone displays its destructive side. Ozone is a strong oxidizer. Hence, direct exposure to high levels of ozone has toxic effects on human health and plants. Although ozone is a major component of photochemical smog in urban areas, this ozone is generally not thought to be a significant contributor to the global ozone budget. Balloon measurements suggest that tropospheric ozone at the global scale has been increasing.

Stratospheric Ozone Trends

Concentrations of ozone in the stratosphere result from chemical production and destruction processes in combination with transport processes. Production of ozone in the stratosphere results primarily from photodissociation of oxygen molecules. The destruction of ozone occurs mainly through catalytic reactions with other gases, such as chlorine and bromine. The total amount of ozone in the stratosphere will remain fairly constant (relative to the well-recognized seasonal variations) as long as there is no change in the destruction rate and the transport of ozone out of the

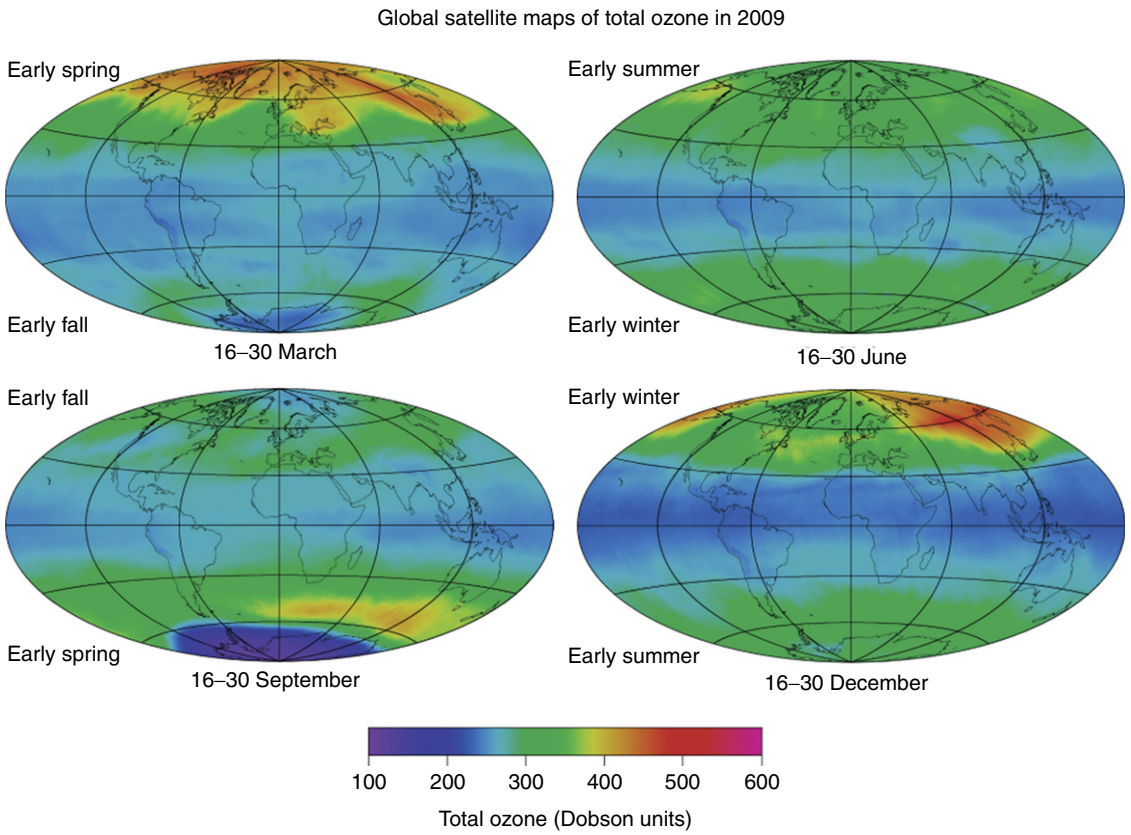


Figure 4 Integrated column of ozone, referred to as the total ozone column, for 2009 seasonal averages based on satellite observations. Total ozone values are often reported in Dobson units (DU). Typical values vary between 200 and 500 DU over the globe. The September measurements show the Antarctica ozone "hole." Reproduced from 20 Questions and Answers section of World Meteorological Organization (2010) *Scientific Assessment of Ozone Depletion: 2009. Global Ozone Research and Monitoring Project, Report 52.* Geneva: World Meteorological Organization.

stratosphere. However, increasing inputs of chlorine and bromine into the stratosphere over the past few decades have changed this balance.

Measurements of ozone by satellite and ground-based instruments over the last several decades indicate that stratospheric ozone levels have decreased significantly. Amounts of ozone in the global atmosphere have decreased globally by more than 5% since about 1970 until the mid-1990s (the maximum decrease in ozone in 1992–1993 is in part related to the effects of sulfur emissions on heterogeneous chemistry in the stratosphere from the Mount Pinatubo eruption in 1991). **Figure 5** shows global averaged changes in ozone based on satellite and ground-based measurements. Satellite measurements began in late 1978, while a reasonably representative global network of ground-based stations has been in operation since the 1960s. In the data shown here, the seasonal variations and the effects of the variations in the solar flux due to the 11-year solar cycle have been removed (which would produce about a 1.2–1.5% variation from solar minimum to solar maximum).

Significant decreases in the total (integrated column of) ozone are found in both the Northern and Southern Hemispheres at the middle and high latitudes, with no significant change in the tropics. Much larger decreases in total ozone are found at latitudes greater than 60°, particularly in the Southern Hemisphere, as discussed later. Measurements indicate that the total ozone column at mid-latitudes in the Northern Hemisphere has decreased by 1.3% per decade in the summer months and 2.7% per decade in the winter months since 1970. Satellite and ozonesonde (balloon) data sets indicate that ozone is particularly decreasing in the lower stratosphere, accounting for a major fraction of the trend in total ozone, although there is also significant ozone destruction occurring in the upper stratosphere. At the end of the ozone record in **Figure 5**, there is an upturn in ozone associated in part with the recovery from the effects of the Mount Pinatubo volcanic eruption in 1991. There are also indications that part of the increase in ozone is related to the effects of the Montreal Protocol to reduce levels of stratospheric chlorine.

Beginning in the late 1970s, a special phenomenon began to occur in the springtime over Antarctica, referred to as the Antarctic ozone "hole." A large decrease in the total ozone, now greater than a 60% decrease relative to prehole levels, has been observed in the springtime (September to November) above Antarctica. Joe Farman of the British Antarctic Survey and coauthors documented this rapid springtime decrease in Antarctic ozone over the ozone measurement station at Halley Bay, Antarctica, attracting the attention of the scientific community. Decreases in the total ozone column of more than 50% were found, compared with historical values observed by both ground-based and satellite techniques. Measurements made in 1987 indicated that more than 95% of the ozone over

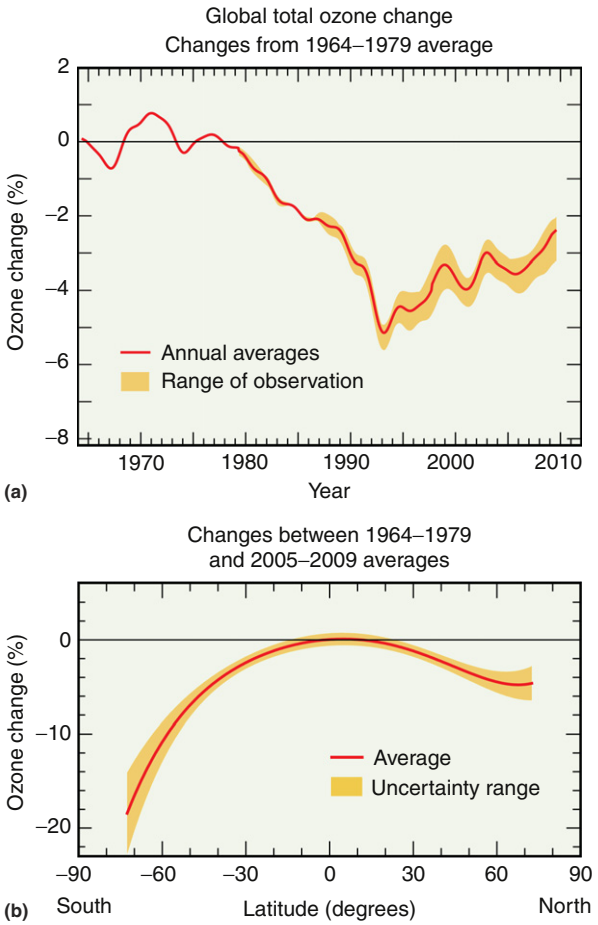


Figure 5 Deviations in total ozone with time relative to 1964–1979 from ground-based and satellite measurements. (a) shows the globally averaged trends and (b) shows the overall trend relative to latitude, showing that the largest changes in ozone have occurred at high latitudes. Reproduced from 20 Questions and Answers section of World Meteorological Organization (2010) *Scientific Assessment of Ozone Depletion: 2009. Global Ozone Research and Monitoring Project, Report 52.* Geneva: World Meteorological Organization.

Antarctica at altitudes from 13 to 22 km had disappeared during September and October. More recently, the Antarctic ozone holes since 1992 have been quite comparable, being the biggest (areal extent) and the deepest (minimum amounts of ozone overhead), with ozone being locally depleted by more than 99% between about 14 and 19 km in October. This continues through 2010, but the extent of the Antarctic ozone decrease should lessen over the coming decades.

Causes and Consequences of Stratospheric Ozone Depletion

The connection between potential environmental effects and man-made CFCs was first raised by Dr. Mario Molina and Dr. Sherwood Rowland in 1974, when they suggested that chlorine from these compounds could destroy stratospheric ozone. Research findings since then have continued to support the significant role that these compounds play in affecting the global distribution of ozone. In addition, it is recognized that other gases containing chlorine and bromine, which are even more reactive with O_3 than chlorine, are also affecting ozone. To a lesser extent, the increasing concentrations of other gases like CO_2, CH_4, and N_2O are also affecting stratospheric ozone.

The inverse relationship between changes in ozone and UV-B is well established by both theoretical analyses and observations. A number of studies have shown that the corresponding increase in UV-B at the ground resulting from ozone depletion can lead to increased incidences of skin cancers, cataracts, and other effects on humans and animals.

The recognition of the deleterious effect of chlorine and bromine on ozone spawned international action to restrict the production and use of CFCs and halons to protect stratospheric ozone. These included the 1987 Montreal Protocol on Substances That Deplete the Ozone Layer and the subsequent 1990 London Amendment, 1992 Copenhagen Amendment, and the 1997 Montreal Amendment. The agreements initially called for reduction of CFC consumption in developed countries. A November 1992 meeting of the United Nations Environment Programme held in Copenhagen resulted in substantial modifications to the protocol because of the large observed decrease in ozone, and called for the phase-out of CFCs, carbon tetrachloride (CCl_4), and methyl chloroform (CH_3CCl_3) by 1996 in developed countries. As part of this, the USA, through the Clean Air Act, has also eliminated the production and import of these chemicals. Production of these compounds is to be totally phased out in developing countries by 2010, while production of halons in developed countries was stopped in 1994. Human-related production and emissions of methyl bromide are not to increase after 1994 in developed countries, with total elimination by 2005.

Projected Trends in Ozone

Projected changes in globally averaged total ozone, using several assumptions about future emissions of CFCs, halons, and their replacements, have been evaluated using models of atmospheric processes. With the original Montreal Protocol provisions, there would still have been a significant reduction in total ozone, as much as 15% by 2050 relative to 1980 levels, according to some analyses. It is only under the London and Copenhagen Amendments and more recent provisions, which call for the complete phase-out of CFCs and halons and other halocarbons, that the ozone reduction trend is reversed. The largest ozone reductions are analyzed to have been reached in the late 1990s. After this, ozone begins to recover, although it is not until about 2050 or later that the 1980 level of global total ozone would again be expected (and more likely, much later over Antarctica). Thus, according to these models, it will roughly be the middle of the century before the chlorine and bromine in the stratosphere are reduced to levels corresponding to those when the Antarctic ozone "hole" first began. Also, without any control measures and assuming unfettered growth, the global mean total column ozone would decrease about 30% by 2050, decreasing further with time. Because of the increasing levels of CO_2, CH_4, and N_2O, which can have either radiative or chemistry effects on ozone, the distribution of ozone in the future stratosphere may not look like that in 1980 at any time this century.

Air Pollutants in the Lower Atmosphere

The elevated concentration of oxidants in urban regions is a difficult problem in many parts of the world. In the USA, progressively tighter emissions controls have been implemented since the late 1970s to control production of ozone. Most of these controls have been aimed at emissions of NO_x and VOCs, particularly the former. Nonetheless, many urban areas still fail to meet the federal standards. In addition, there is growing concern that enhanced levels of ozone concentrations in rural areas downwind of urban areas are resulting in harmful effects on agricultural and forest ecosystems. Part of the difficulty is understanding how emissions of hydrocarbons from the biosphere affect ozone levels in urban areas.

The Future and Atmospheric Gases

Changes in climate associated with increasing concentrations of greenhouse gases could have significant effects on the biosphere and on biodiversity. Corresponding changes in the biosphere could affect biogeochemical cycles and the sources and sinks of atmospheric gases, leading to further changes in the climate. Such interactions with the biosphere and with biodiversity are inadequately considered in current climate projections. The effects of such interactions require much more study.

Stratospheric ozone levels should largely recover at mid-latitudes by mid-century. However, the timing of this recovery will largely depend on the emissions of halocarbons and on the emissions of other gases that influence stratospheric chemistry. Changes in the oxidative capacity could also change the amounts of gases reaching the stratosphere and thus the rate of and extent of recovery. Climate change would also likely influence stratospheric temperature and winds, thus further affecting the rate of ozone recovery. The long-term impacts of the reduced levels of ozone on the biosphere and on biodiversity remain poorly understood.

Attempts to control pollution in urban regions continue to increase. However, there is still much to learn before such controls can be fully effective. The US National Academy of Sciences has stated that too little is known about the transport and deposition of atmospheric gases, including toxics and nutrients, to the biosphere and their interaction with biota. Further study of the rates of chemical exchange between the atmosphere and ecosystems will provide better quantitative estimates of atmospheric chemical impacts on the biosphere and biospheric emissions to the atmosphere.

See also: Air Pollution. Carbon Cycle. Climate Change and Ecology, Synergism of. Climate, Effects of. Nitrogen, Nitrogen Cycle. Ultraviolet Radiation

References

Blasing TJ (2010) Recent greenhouse gas concentrations. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory,* http://cdiac.ornl.gov/pns/current_ghg.html.http://dx.doi.org/10.3334/CDIAC/atg.032.

Boden TA, Marland G, and Andres RJ (2010) Global, regional, and national fossil-fuel CO_2 emissions. *Carbon Dioxide Information Analysis Center.* Oak Ridge, TN, USA: Oak Ridge National Laboratory, US Department of Energy.http://dx.doi.org/10.3334/CDIAC/00001_V2010.

Dlugokencky EJ, Bruhwiler L, and White JWC (2009) Observational constraints on recent increases in the atmospheric CH_4 burden. *Geophysical Research Letter* 36: L18803.http://dx.doi.org/10.1029/2009GL039780.

Houghton RA (2008) Carbon flux to the atmosphere from land-use changes: 1850–2005. *In TRENDS: A Compendium of Data on Global Change.* Oak Ridge, TN, USA: Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy.

World Meteorological Organization (2010) *Scientific Assessment of Ozone Depletion: 2009. Global Ozone Research and Monitoring Project, Report 52.* Geneva: World Meteorological Organization.

Australia, Biodiversity of Ecosystems

Raymond Louis Specht, The University of Queensland, Brisbane, QLD, Australia
Alison Specht, The University of Queensland, Brisbane, QLD, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biogeographic region Region that contains a "homogeneous" suite of floristic groups sorted by some classificatory program.
Community diversity Number of ecosystems and associated plant and animal species per 1° latitude by 1° longitude grid-cell.
Diagnostic species Species that emerge from semiquantitative analysis as being the key species associated with a floristic group.
Evaporative aerodynamics Study of the influence of aerodynamics on the development of both leaf structure and foliage distribution throughout a plant community.
Floristic group Plant community that contains a "homogeneous" suite of species sorted by some classificatory program.
Foliage projective cover (FPC) The horizontal cover of leaves (in crowns of varying foliage density) in overstory and understory strata of plant community, measured using vertical cross-wire sighting tubes.
Shoot growth Biomass production (per hectare) of foliage shoots produced annually in the overstory of a plant community.
Species richness Number of species of plant or animal per hectare (per 100 hectares for rainforest remnants).

Outline

Species richness (number of species per hectare) of overstory and understory plants in floristic groups from arid to perhumid climatic zones in the tropical north to the temperate south of Australia, together with species richness of resident vertebrates and epigaeic invertebrates, is correlated with annual energy-fixation (per hectare) in each ecosystem.

Australian Ecosystems

Australia is a continental mass lying south of the equator between latitudes 11° and 44°. The climate varies markedly with geographical position: summer rainfall predominates in the north, winter rainfall is a characteristic of the south, and the eastern seaboard shows a more general distribution of rainfall throughout the year. Marked evaporative gradients from the humid coastal fringe to the arid inland are found in the south, east, and north of the continent. The structure of the vegetation varies along these evaporative gradients: from the dense rainforests in eastern Australia, to the tall eucalypt forests of the south-west and south-east (extending northwards at the fringe of subtropical and tropical rainforests), to the eucalypt forests/woodlands, with either grassy or heathy understory, distributed throughout the continent. Heathlands and grasslands are found in tropical to temperate climes, the mallee eucalypt vegetation is found on the calcareous soils of the south, and a variety of other vegetation structures are found in the vast arid zone and in the alpine, coastal, and wetland landscapes of the continent.

Most families and many genera now found in the Australian flora developed over 100 million years ago, before the Gondwanan supercontinent began to break up. Vast expanses of infertile lateritic soils developed in the subtropical and tropical climates of that time (Prescott and Pendleton, 1952; Specht and Specht, 1999). Subtropical and tropical rainforests were widespread, with cool temperate *Nothofagus* forests at higher altitudes (Specht, 1981b; Specht *et al.*, 1992). Gondwanan and Old World Tropical C4 grasses (67 genera, including two species), were widespread at the lateritic earths of tropical Australia and extended southwards on medium-nutrient soils far to the south (Clifford and Simon, 1981). Many woody plant species on the sandstone outcrops of northern Australia are identical with those found on the sandstone hills of western Bengal and central India (Hooker, 1860; Specht, 1958; Specht, 1981a; Specht, 1988a). In the south of the continent, sclerophyllous (heathy) vegetation (± an overstory of *Eucalyptus* spp.) flourished on sandstone and quartzitic outcrops and formed wet-heathlands with species of Restionaceae and Cyperaceae on sandy out-wash plains (Johnson and Briggs, 1981; Specht *et al.*, 1992). Today, most of the Australian *Acacia* species, widespread across the continent, belong to the subgenus *Racosperma*; five *Acacia* species, related to the Gondwanan parent genus of *Acacia* in South America, Africa, and India, still survive in northern Australia (Specht 1981a).

The genus *Eucalyptus* (Myrtaceae) appears to have evolved in the Early Tertiary across the southern part of the continent, the earliest pollen appearing in the Oligocene about 34 million years ago, extending from New Zealand in the east to the Ninety-East Ridge in the west (Martin, 1981; Pole, 1989; Pole, 1993). The fossil fruits of *Eucalyptus patagonica*, found in Eocene deposits in South America (Hill, 1994 and Hill, 2004), however, suggest that Australasian eucalypts may have evolved long before, in Gondwanaland. Some of the fossil pollen grains of the Early Tertiary resemble those of *E. spathulata* (in the subgenus *Symphyomyrtus*), now a stunted mallee species

growing in south-west Western Australia (Martin and Gadek, 1988). Fossil leaf impressions indicate that tropical *Corymbia* (bloodwoods and spotted gums), until recently considered a subgenus of *Eucalyptus,* evolved at the periphery of the rainforests, even as far south as Tasmania (Ladiges *et al.*, 2003).

The onset of aridity during the Cainozoic (Martin, 2006), especially when global temperatures rose by about 5 °C during the Mid-Miocene (Feary *et al.*, 1991; Shackleton and Kennett, 1975), caused fragmentation of the humid/perhumid Gondwanan vegetation. Less humid elements in the flora began to expand and evolve, eventually to cover much of the continent. In the drier center of the continent, extensive stands of a tall shrubland of mulga (*Acacia aneura*) were mixed with hummock grasslands dominated by the spiny spinifex grass, *Triodia basedowii*, on low-nutrient soils (Specht and Specht, 1999). Chenopod low shrublands (with stands of *Casuarina pauper*) occupied the Early Miocene calcareous soils in the south of the continent (Specht and Specht, 1999).

The vegetation of Australia today comprises those taxa that were able to survive as the evaporative climate became drier. A distinctive heathy understory has survived on the nutrient-poor lateritic podsols, also on quartzitic/sandstone and acid granitic rocks, characteristic of a large part of the original Gondwanan super-continent. A savanna understory of Gondwanan C4 grasses survived under eucalypt open-forests and woodlands on medium-nutrient lateritic earth. The attributes that enabled most of the Gondwanan flora to survive over the last 50 million years also facilitates their rapid regeneration from underground organs and epicormic buds whenever a disturbance such as a fire occurs.

Evaporative Aerodynamics

Water – its availability through annual and supra-annual cycles, and the aridity of the atmosphere prevalent over most of the continent – is the primary limiting factor in ecosystem development in the Australian continent. The distribution of foliage in the overstory and understory canopies, and the eco-physiological attributes of the foliage (such as the "sclerophylly" characteristic of much Australian vegetation) is primarily governed by water availability in relation to atmospheric demand. The growth of new shoot apices (in both overstory and understory strata) is in equilibrium with the evaporative aerodynamics of the atmosphere flowing over and through the plant community (Specht, 1972; Specht and Specht, 1999). The amount of available water affects the vertical growth of each foliage shoot (and the number of essentially identical leaves produced annually), but not the horizontal cover of the foliage throughout the plant community, (Specht and Specht, 1989a) which is determined by evaporative demand, a relatively constant factor from year to year.

After any perturbation, the horizontal cover of the foliage and its eco-physiological characteristics are rapidly restored in both the overstory and understory of an evergreen plant community. The ratio of actual to potential evapotranspiration (per hectare) from the plant community is linearly related to the amount of water available during each month; the slope of this linear regression is called the Evaporative Coefficient (Specht, 1972). Values of this community–physiological constant have been computed for meteorological stations throughout Australia and the resulting isolines are mapped in **Figure 1**.

Major Plant Communities in Australia

Plant communities are fundamental to ecosystem definition. The great number of tree species (20–140) that coexist in subtropical and tropical plant communities in Australia, from the arid to the perhumid evaporative zones, however, makes the definition of statistically "homogeneous" suites within the vegetation a challenge. This is not faced by temperate plant ecologists who deal with simple systems with relatively low numbers of overstory species per hectare (1–10 in number). The polythetic-divisive classificatory program TWINSPAN (Hill, 1973) was used to assist the objective identification of plant communities, henceforth called TWINSPAN floristic groups.

TWINSPAN floristic groups were defined in the 29 Australian plant formations (Specht *et al.*, 1995). To accomplish this, complete species lists (several floristic lists could not be included) reported for plant communities throughout the continent since the 1890s were collated and arranged according to structural formation (**Table 1**). These were collated into 16 large data banks and were analyzed using TWINSPAN by the CSIRO Division of Computing Research based on the presence or absence of a species in each list.

In each case, the communities were first sorted by the program according to climatic zone, emphasizing the importance of climate to species occurrence. First, the tropical/subtropical plant communities in northern Australia were separated from the temperate/montane plant communities in southern Australia. Then the tropical communities were separated from the subtropical, and the temperate from the montane. Further subdivision separated the humid plant communities from the subhumid, and so on for plant communities that had developed in different microenvironments. The number of TWINSPAN floristic groups and their subdivisions that were recorded in each 30′ latitude by 30′ longitude grid-cell are shown in the *Conservation Atlas of Plant Communities in Australia* (Specht *et al.*, 1995). The diversity of floristic groups is greatest in the perhumid zone, and least in the arid evaporative zone.

TWINSPAN analysis resulted in the definition of 338 overstory floristic groups. A further 60 understory floristic groups were recognized in open-structured plant communities (Specht *et al.*, 1995).

Diagnostic Species in Australian Floristic Groups

The plant species, which define the floristic groups (Specht *et al.*, 1995) determined by TWINSPAN in the 29 Australian plant formations (**Table 1**) are listed below. Each floristic group has been given a code: capital letters for the formation (e.g., RF = closed-forest (rainforest), T = open-forest and woodland), followed by numerals. The numerals refer to the dichotomous

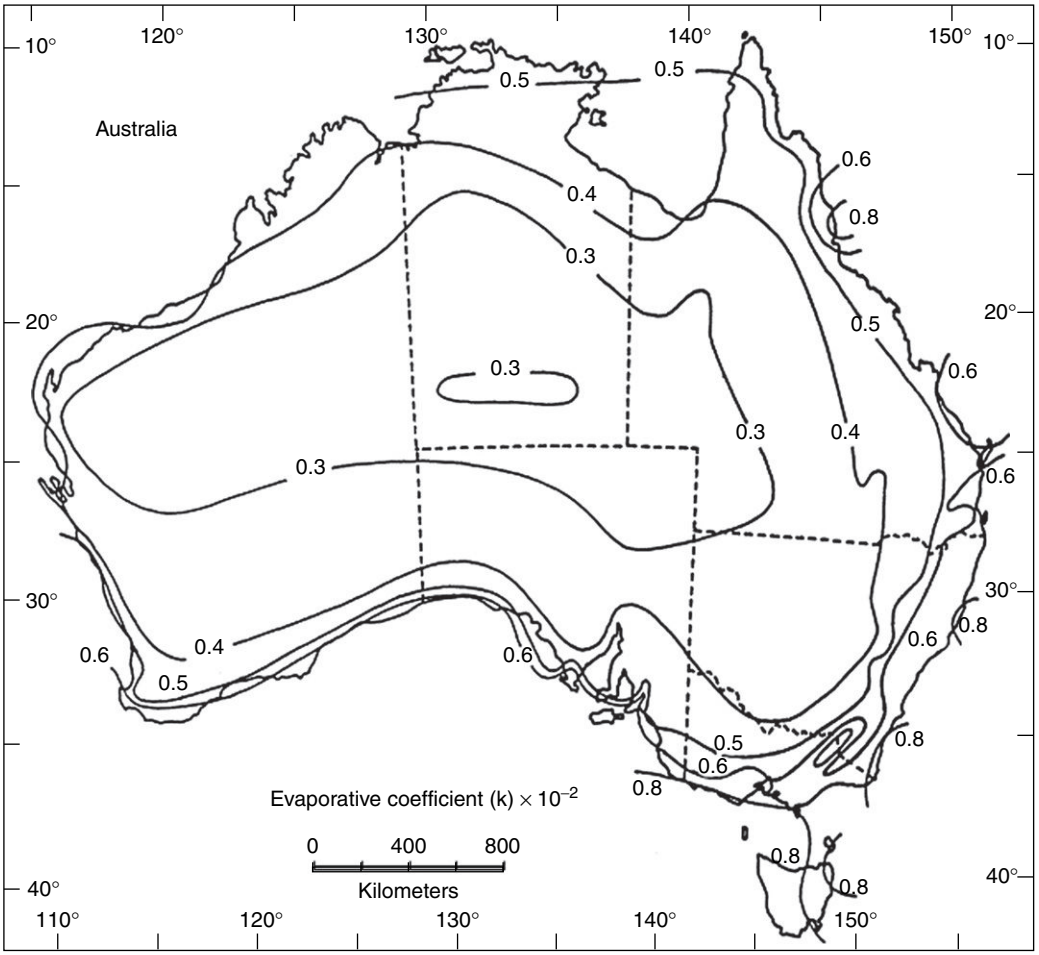


Figure 1 Map of Australia showing isolines of the Evaporative coefficient (k). Climatic zones may be defined as follows: Arid zone, $k<0.35$; semi-arid zone, $k \equiv 0.35$–0.45; subhumid zone, $k \equiv 0.45$–0.55; humid zone, $k \equiv 0.55$–0.75; perhumid zone, $k>0.75$. (All values of k are $\times 10^{-2}$ per mm of soil water available per month.) (Specht, 1972; Specht and Specht, 2002).

subdivisions in the TWINSPAN classification; within a formation, groups having consecutive odd–even pairs of numerals are most closely related (e.g., RF35 and RF36). In species-rich formations, such as closed-forests, subdivisions of the primary TWINSPAN groups are shown by the numbers in parentheses.

Closed-forests (rainforests) – tropical north-eastern Australia (Webb *et al.*, 1984)

RF35(140) *Drypetes deplanchei*
RF35(141) *Endiandra glauca – Polyscias australiana*
RF35(142) *Calophyllum australianum – Castanospermum australe*
RF35(143) *Polyalthia nitidissima*
RF36(144) *Agathis robusta – Diospyros pentamera*
RF36(145) *Myristica insipida*
RF36(146) *Polyscias australiana – Podocarpus neriifolius*
RF36(147) *Archontophoenix alexandrae – Ficus racemosa*
RF37(75) *Acronychia acidula – Dendrocnide photinophylla*
RF37(148) *Archontophoenix alexandrae – Castanospermum australe*
RF37(149) *Musgravea heterophylla – Normanbya normanbyi*
RF38(76) *Beilschmiedia tooram – Endiandra sankeyana*
RF38(155) *Cryptocarya angulata – C. corrugata*
RF38(308) *Cryptocarya mackinnoniana – Flindersia bourjotiana*
RF38(309) *Balanops australiana – Lomatia fraxinifolia*
RF39 *Laccospadix australasica – Steganthera macooraia*

Closed-forests (rainforests) – subtropical eastern Australia (Webb *et al.*, 1984)

RF40(160) *Austromyrtus acmenoides – Litsea leefeana – Sloanea woollsii*
RF40(163) *Helicia ferruginea – Litsea leefeana*
RF40(322) *Rhodamnia argentea – Syzygium francisii*
RF40(323) *Cryptocarya obovata – Sloanea woollsii*
RF40(324) *Acmena ingens – Zanthoxylum brachyacanthum*
RF40(325) *Schizomeria ovata – Argyrodendron actinophyllum*
RF41(165) *Cupaniopsis anacardioides*
RF41(167) *Celtis paniculata – Cryptocarya triplinervis*
RF41(328) *Dendrocnide* spp. *– Ficus macrophylla – Toona ciliata*
RF41(329) *Acmena smithii – Ficus coronata – Toona ciliata*
RF41(332) *Cryptocarya glaucescens – Mischocarpus pyriformis*
RF42(168) *Orites excelsa – Sloanea woollsii*
RF42(169) *Nothofagus moorei – Cryptocarya foveolata – Orites excelsa*
RF42(170) *Nothofagus moorei – Tristaniopsis laurina*
RF42(171) *Doryphora sassafras*
RF43(86) *Acmena smithii – Ficus coronata – Eupomatia laurina*

Table 1 Structural characteristics of major plant formations in Australia[a]

Formation	*Climatic zone*	*Overstory*			*Understory*
		Stand height (m)	*Foliage projective cover (%)*	*Leaf-specific-weight*[b] *(mg cm^{-2})*	
Overstory with trees					
Closed-forest	Perhumid	20[c]–10	100–70	8–16	Regeneration saplings, shade-tolerant plants
Open-forest	Humid	>30–10	70–30	10–28	Grassy, heathy
Woodland	Subhumid	20–10	30–10	13–40	Grassy, heathy
Open-scrub	Semiarid	5–2	70–30	18–50	Grassy, heathy
Tall shrubland	Arid	5–2	30–10	20–>50	Grassy, hummock grass, chenopods
Treeless vegetation					
Heathland	Humid–subhumid	2–0.25	70–30	Sclerophyllous	–
Dwarf heathland (fellfield)	Montane	<0.25	30–<10	Sclerophyllous	–
Tussock grassland	Humid–subhumid	1–0.25	70–30	Herbaceous	–
Open tussock grassland	Semiarid	<0.25	30–10	Herbaceous	–
Hummock grassland	Semiarid–arid	1–0.5	70–10	Sclerophyllous	–

[a]Based on Specht *et al.* (1995) and Specht and Specht (2002).
[b]Leaf/leaflet area decreases and leaf specific weight increases in the 10 °C temperature gradient from tropical to temperate regions.
[c]In the tropics and subtropics, there are emergent trees and strangler figs, lianas and epiphytes.

RF43(87) *Acacia melanoxylon – Acmena smithii*
RF44(88) *Cryptocarya floydii – Pittosporum undulatum*
RF44(178) *Podocarpus elatus – Syzygium oleosum*
RF44(359) *Araucaria cunninghamii – Flindersia australis*
RF44(358) *Araucaria cunninghamii – Argyrodendron trifoliolatum – Drypetes* sp.
RF45(180) *Araucaria cunninghamii – Argyrodendron trifoliolatum – Cryptocarya bidwillii*
RF45(181) *Murraya ovatifoliolata*
RF45(182) *Araucaria cunninghamii – Austromyrtus bidwillii*
RF45(183) *Strychnos axillaris*
RF46(93) *Apophyllum anomalum – Maytenus cunninghamii*
RF46(184) *Cupaniopsis anacardioides – Pleiogynium timorense*
RF46(185) *Casearia multinervosa – Croton insularis*

Closed-forests (rainforests) – temperate south-eastern Australia (Webb *et al.*, 1984)

RF7 *Nothofagus cunninghamii – Athrotaxis cupressoides – Diselma archeri*
RF13 *Nothofagus cunninghamii – Atherosperma moschatum*
RF24 *Nothofagus cunninghamii – Pittosporum bicolor*
RF25 *Atherosperma moschatum – Pittosporum bicolor*

Semi-deciduous closed-forests – monsoonal northern Australia (Russell-Smith, 1991)

MRF1 *Aglaia sapidina – Hydriastele wendlandiana*
MRF2 *Acacia auriculiformis – Carpentaria acuminata*
MRF3 *Acmenosperma claviflorum – Macaranga involucrata*
MRF4 *Calophyllum sil – Ficus hispida – Ilex arnhemensis*
RF34(69) *Carallia brachiata – Ilex arnhemensis*
MRF5 *Acmenosperma claviflorum – Litsea breviumbellata*
MRF6 *Omalanthus novo-guineensis – Xanthostemon eucalyptoides*
MRF7 *Allosyncarpia ternata – Calophyllum sil*
MRF8 *Allosyncarpia ternata – Alyxia ruscifolia*
MRF9 *Aglaia rufa – Diospyros maritima – Secamone elliptica*
MRF10 *Barringtonia acutangula – Livistona benthamii*
MRF11 *Ficus racemosa – Melaleuca argentea*
MRF12a *Canarium australianum – Celtis philippensis*
MRF12b *Alstonia spectabilis – Denhamia obscura*
MRF13 *Ficus coronulata – F. virens – Glochidion apodogynum*
MRF14 *Brachychiton collinus – Callicarpa candicans*
MRF15 *Brachychiton spectabilis – Ziziphus quadrilocularis*
RF47 *Brachychiton chillagoensis – Terminalia aridicola*
RF16(132) *Gyrocarpus americanus – Lysiphyllum cunninghamii*
RF16(32) *Ficus platypoda* var. *minor*

Semi-deciduous closed-forests – subtropical eastern Australia (Forster *et al.*, 1991)

DS8(16) *Cupaniopsis parvifolia – Diospyros humilis – Owenia venosa*
DS8(17) *Cupaniopsis parvifolia – Maclura cochinchinensis*
DS9(18) *Archidendropsis thozetiana – Cupaniopsis wadsworthii*
DS9(19) *Archidendropsis thozetiana – Brachychiton rupestris*
DS10(20) *Pleiogynium timorense – Pouteria sericea*
DS10(21) *Alchornia ilicifolia – Cryptocarya triplinervis*
DS11(22) *Cupaniopsis parvifolia*
DS11(23) *Cupaniopsis parvifolia – Araucaria cunninghamii*
DS12(24) *Mischocarpus pyriformis – Alyxia ruscifolia*
DS12(25) *Mischocarpus pyriformis – Araucaria cunninghamii*
DS13(26) *Araucaria cunninghamii – Alectryon tomentosus – Cleistanthus cunninghamii*
DS13(27) *Araucaria cunninghamii – Alectryon tomentosus – A. subcinereus*
DS14(28) *Cryptocarya macdonaldii – C. laevigata*
DS14(29) *Cryptocarya macdonaldii*
DS15(30) *Araucaria cunninghamii – Calamus muelleri -Wilkiea macrophylla*
DS15(31) *Araucaria cunninghamii – Maclura cochinchinensis – Smilax australis*

DS40 *Alectryon forsythii – A. subdentatus – Notelaea microcarpa*

Eucalypt open-forests and woodlands – monsoonal northern Australia (R. L. Specht and D. W. Drake, unpublished data)

T187a *Eucalyptus leucophloia*
T189 *Eucalyptus microneura*
T359 *Eucalyptus brownii – E. crebra*
T368 *Eucalyptus cullenii – E. shirleyi – Corymbia erythrophloia*[1]
T369 *Corymbia papuana – C. erythrophloia*
T371 *Corymbia papuana – C. polycarpa – Eucalyptus platyphylla*
T372 *Eucalyptus tectifica – Corymbia grandifolia – C. foelscheana*
T373a *Eucalyptus tetrodonta – E. miniata – Corymbia ferruginea*
T738 *Eucalyptus whitei – Corymbia papuana*
T739 *Coymbia papuana – C. erythrophloia*

Eucalypt open-forests and woodlands – subtropical eastern Australia (R. L. Specht and D. W. Drake, unpublished data)

T168 *Eucalyptus andrewsii – E. youmanii – E. cameronii*
T168/343 *Eucalyptus deanei – E. campanulata*
T346 *Euc. campanulata – E. laevopinea – E. cypellocarpa – E. obliqua*
T346b *Eucalyptus saligna – E. microcorys*
T351 *Eucalyptus campanulata – E. saligna*
T352 *Eucalyptus amplifolia – Angophora subvelutina*
T360 *Eucalyptus chloroclada – Angophora costata* subsp. *leiocarpa*
T361 *Eucalyptus chloroclada – E. sideroxylon*
T362 *Eucalyptus microcarpa*
T363a *Eucalyptus microcarpa – E. populnea*
T364 *Eucalyptus populnea – E. orgadophila*
T365 *Eucalyptus melanophloia*
Calcolu *Callitris glaucophylla*
Serp-R *Eucalyptus fibrosa* subsp. nov. *Corymbia xanthope – Macrozamia* sp. nov.
Serp-W *Eucalyptus crebra – Corymbia erythrophloia – Macrozamia* sp. nov.
Serp-B *Eucalyptus ophitica – Xanthorrhoea glauca*
T706 *Eucalyptus pilularis – Corymbia gummifera – Angophora* spp.
T707 *Eucalyptus paniculata – E. propinqua*
T710 *Euc. acmenoides – E. pilularis – E. microcorys – Syncarpia glomulifera*
T710b *Eucalyptus grandis*
Eucbiturb *Eucalyptus biturbinata*
Euccloe *Eucalyptus cloeziana – E. tenuipes*
Shill *Syncarpia hillii*
T711 *Eucalyptus signata – Corymbia intermedia*
T712 *Eucalyptus crebra – Corymbia citriodora – C. trachyphloia*
T713 *Corymbia intermedia – Lophostemon* spp.
T714 *Eucalyptus tereticornis – Lophostemon suaveolens*
T715 *Eucalyptus tereticornis – Corymbia tessellaris – C. clarksoniana*
T716a *Eucalyptus crebra – E. tereticornis – Corymbia maculata*
T716b *Eucalyptus moluccana – E. tereticornis*
T717 *Eucalyptus crebra – Corymbia maculata*

Eucalypt open-forests and woodlands – temperate south-eastern Australia (R. L. Specht and D. W. Drake, unpublished data)

T172 *Eucalyptus regnans*
T328a *Eucalyptus odorata – E. porosa*
T329 *Eucalyptus leucoxylon*
T330 *Eucalyptus fasciculosa – E. leucoxylon*
T331 *Eucalyptus baxteri – E. fasciculosa*
Eucpryor *Eucalyptus pryoriana – E. viminalis* subsp. *cygnetensis*
Eucnit *Eucalyptus nitida*
T338 *Eucalyptus albens – E. blakelyi – E. melliodora*
T339 *Eucalyptus dealbata – E. blakelyi – E. sideroxylon*
T340 *Eucalyptus blaxlandii – E. goniocalyx*
T341 *Eucalyptus macrorhyncha – E. polyanthemos*
T342 *Eucalyptus dives – E. macrorhyncha – E. radiata*
T347 *Eucalyptus cypellocarpa – E. obliqua – E. radiata*
T348 *Eucalyptus globoidea – E. sieberi*
T349 *Eucalyptus globoidea – Corymbia gummifera*
T350 *Eucalyptus piperita – E. sieberi-Corymbia gummifera*
T40 *Eucalyptus cordata*
T664 *Eucalyptus nitida*
T665 *Eucalyptus cephalocarpa – E. globoidea*
T666 *Eucalyptus viminalis*
T667 *Eucalyptus obliqua*

Eucalypt open-forests and woodlands – montane south-eastern Australia (R. L. Specht and D. W. Drake, unpublished data)

T8 *Eucalyptus coccifera*
T36 *Eucalyptus obliqua – E. amygdalina – E. globulus*
T38 *Eucalyptus delegatensis – E. dalrympleana – E. viminalis*
T75 *Eucalyptus urnigera – E. delegatensis*
T79 *Eucalyptus brookeriana – E. delegatensis*
T148 *Eucalyptus obliqua*
T149 *Eucalyptus delegatensis – E. regnans – E. obliqua*
T156 *Eucalyptus delegatensis – E. pauciflora – E. rubida*
T157 *Eucalyptus delegatensis – E. pauciflora – E. rodwayi*
T334 *Eucalyptus pauciflora – E. rubida*
T335 *Eucalyptus pauciflora – E. stellulata*

Eucalypt open-forests and woodlands – temperate south-western Australia (R. L. Specht and A. J. M. Hopkins, unpublished data)

SW20 *Eucalyptus gomphocephala – Agonis flexuosa*
SW26 *Eucalyptus loxophleba*
SW43a *Eucalyptus marginata – Corymbia calophylla*
SW47 *Eucalyptus marginata – Banksia* spp. *– Allocasuarina humilis*
SW85 *Eucalyptus diversicolor*
SW96a *Eucalyptus wandoo*
SW97a *Eucalyptus salmonophloia*
SW100 *Eucalyptus loxophleba – E. salmonophloia*

Eucalypt open-forests and woodlands – Australian wetland forests (Specht, 1990)

T188 *Eucalyptus coolabah*, *E. microtheca*, etc.
T191 *Eucalyptus ochrophloia – E. coolabah*
T190 *Eucalyptus camaldulensis*
T352b *Casuarina cunninghamiana*
T354 *Melaleuca quinquenervia – Eucalyptus robusta*
T366a *Eucalyptus largiflorens*
T366b *Allocasuarina luehmannii*
T370 *Eucalyptus leptophleba*
SW4a *Eucalyptus rudis – Melaleuca rhaphiophylla*
SW4b *Eucalyptus rudis – Corymbia calophylla*
SW84 *Agonis hypericifolia – Banksia littoralis*
T801 *Melaleuca leucadendra – M. viridiflora*

Mallee eucalypt open-scrubs – monsoonal northern Australia (R. L. Specht, unpublished data)

[1]*Corymbia* was formerly regarded as a subgenus of *Eucalyptus*.

Euccupu *Eucalyptus cupularis*
Eucmann *Eucalyptus mannensis*
Eucnorm *Eucalyptus normantonensis*
Eucpers *Eucalyptus persistens*
Eucoxym *Eucalyptus oxymitra*
Eucpach *Eucalyptus pachyphylla*
Eucset *Corymbia setosa*
Eucthoz *Eucalyptus thozetiana – Acacia catenulata – A. aneura*

Mallee eucalypt open-scrubs – subtropical eastern Australia (R. L. Specht, unpublished data)
Eucappr *Eucalyptus approximans*
Eucbake *Eucalyptus bakeri*
Euccurt *Eucalyptus curtisii*
Serp-GW *Eucalyptus serpenticola – E. nortonii – Allocasuarina ophiolitica*

Mallee eucalypt open-scrubs – temperate south-eastern Australia (R. L. Specht, unpublished data)
M7 *Eucalyptus diversifolia – E. cosmophylla – E. rugosa -Melaleuca uncinata*
M10 *Eucalyptus behriana*
M13 *Eucalyptus diversifolia – E. socialis*
M23 *Eucalyptus socialis – E. dumosa – E. gracilis – E. polybractea*
M24 *Eucalyptus incrassata – E. viridis – E. polybractea*
M25 *Eucalyptus incrassata – E. foecunda – Melaleuca uncinata*
M4 *Eucalyptus socialis – E. dumosa*
M44 *Eucalyptus socialis – E. dumosa – E. gracilis – Acacia rigens*
M45 *Eucalyptus socialis – E. dumosa – E. gracilis – Enchylaena tomentosa*
Eucinte *Eucalyptus intertexta – Callitris glaucophylla*
Euckits *Eucalyptus kitsoniana*
Eucobtus *Eucalyptus obstans – E. luehmanniana*

Mallee eucalypt open-scrubs – temperate south-western Australia (R. L. Specht and A. J. M. Hopkins, unpublished data)
SW27 *Eucalyptus diversifolia – E. oleosa – Melaleuca lanceolata*
SW51 *Eucalyptus leptopoda – Callitris preissii – Melaleuca cordata*
SW101b *Eucalyptus eremophila – E. forrestiana – E. oleosa*
SW196 *Allocasuarina campestris*
SW197 *Allocasuarina campestris – Acacia acuminata -Eucalyptus obtusiflora*

Heathlands and related shrublands – monsoonal northern Australia (R. L. Specht and W. E. Drake, unpublished data)
H7 *Corymbia ferruginea – Jacksonia odontoclada – Plectrachne pungens*
H7a *Calytrix exstipulata – Grevillea pteridifolia – Persoonia falcata*
H12 *Asteromyrtus lysicephala -Jacksonia thesioides*
H13 *Neofabricea myrtifolia – Neoroepera banksii*
Baecfrut *Baeckea frutescens*
Leptpurp *Leptospermum purpurascens*

Heathlands and related shrublands – subtropical eastern Australia (R. L. Specht and W. E. Drake, unpublished data)
H40 *Allocasuarina littoralis – Alphitonia excelsa*
H82 *Melaleuca nodosa – Gahnia sieberana – Blechnum indicum*
H166 *Leptospermum polygalifolium – Phyllota phylicoides*
H167 *Leptospermum polygalifolium – Platysace linearifolia*
Xanth *Xanthorrhoea* spp.

Heathlands and related shrublands – temperate south-eastern Australia (R. L. Specht and W. E. Drake, unpublished data)
H44a *Melaleuca uncinata*
H44b *Banksia ornata – Xanthorrhoea caespitosa – Allocasuarina pusilla*
H45 *Melaleuca uncinata – Xanthorrhoea semiplana* var. *tateana*
H84 *Banksia ericifolia – Hakea dactyloides*
H170 *Banksia marginata – Epacris impressa*
H171 *Banksia marginata – Epacris impressa – Pteridium esculentum*

Heathlands and related shrublands – montane south-eastern Australia (Kirkpatrick, 1989)
AV1 *Carex gaudichaudiana – Myriophyllum pedunculatum*
AV2 *Carex gaudichaudiana*
AV3 *Baeckea gunniana – Epacris breviflora – Carex gaudichaudiana*
AV4 *Baeckea gunniana – Epacris paludosa – E. lanuginosa*
AV5 *Oxylobium alpestre – Olearia phlogopappa – Phebalium ovatifolium*
AV6 *Grevillea australis – Epacris paludosa – Poa* spp.
AV7 *Leucopogon montanus – Poa* spp.
AV8 *Podocarpus lawrencei*
AV9 *Microtrobos niphophilus – Epacris serpyllifolia*
AV10 *Abrotanella forsterioides – Richea scoparia*
AV11 *Carpha rodwayi – Oreobolus pumilio – Phyllachne colensoi*
AV12 *Diselma archeri – Microcachrys tetragona*
AV13 *Gaultheria depressa – Senecio leptocarpus*
AV14 *Epacris serpyllifolia – Orites revoluta – Richea sprengelioides*
AV15 *Epacris serpyllifolia – Richea scorparia – Senecio leptocarpus*
AV16 *Eucalyptus vernicosa – Nothofagus cunnunghamii*
AV-Ngum *Nothofagus gunnii*

Heathlands and related shrublands – temperate south-western Australia (George *et al.*, 1979; R. L. Specht and A. J. M. Hopkins, unpublished data)
SW22 *Allocasuarina humilis – Calothamnus quadrifidus – Dryandra nivea*
SW46 *Allocasuarina humilis – Dryandra nivea – Hakea lissocarpha*
SW47b *Banksia attenuata – B. menziesii*
SW99a *Allocasuarina acutivalvis*

Tussock grasslands (R. L. Specht and M. P. Bolton, unpublished data)
G8 *Astrebla elymoides – Iseilema membranaceum*
G11 *Dichanthium sericeum*
G18 *Astrebla* spp. – *Dichanthium fecundum – Iseilema* spp.
G19 *Astrebla elymoides – A. lappacea*
G56 *Xerochloa* spp. – *Eleocharis* spp. – *Oryza sativa*
G69 *Themeda triandra* (Basalt Plains, western Victoria)
G114 *Themeda triandra* (coastal headlands, Queensland)
GDanth *Danthonia caespitosa*

***Acacia* vegetation – subhumid, subtropical eastern Australia** (R. L. Specht, M. P. Bolton, and A. Specht, unpublished data)
HW11 *Acacia harpophylla – Casuarina cristata*
DA8a *Acacia catenulata*
DA8b *Acacia shirleyi*
DA18 *Acacia cambagei*
DA40 *Acacia pendula – Atriplex nummularia*
DAargy *Acacia argyrodendron*
DAspar *Acacia sparsiflora – A. burrowii*
DAsuth *Acacia sutherlandii – Astrebla* spp.

Acacia vegetation – Australian Arid Zone (R. L. Specht, M. P. Bolton, and A. Specht, unpublished data)
DA14a *Acacia ligulata*

DA19 *Acacia aneura – Eucalyptus populnea – Eremophila mitchellii*
DA21 *Eucalyptus terminalis – Acacia ligulata – A dictyophleba*
DA22a *Acacia aneura – Senna* spp. *– Eremophila* spp.
DA22d *Acacia sclerosperma – A. aneura*, etc.
DA24 *Eucalyptus terminalis – Acacia dictyophleba – Senna artemisioides*
DAgeor *Acacia georginae*
DA41a *Acacia aneura – Alectryon oleifolius*
C24 *Acacia papyrocarpa – Maireana sedifolia*
DA46a *Acacia kempeana – A. aneura*
DA47 *Acacia acradenia – A. tenuissima – Senna artemisioides*
DA51 *Eucalyptus gamophylla – Acacia* spp. *– Dodonaea viscosa*
DA62 *Eucalyptus kingsmillii – E. youngiana – E. gamophylla – Triodia* sp.
DA62a *Eucalyptus gongylocarpa – Triodia basedowii*
DA63a *Eucalyptus gamophylla – Acacia* spp. *– Thryptomene maisonneuvei*
DA100g *Eucalyptus dichromopholoia – Acacia lysiphloia – Acacia* spp.
DA101c *Acacia pachycarpa – Eucalyptus dichromophloia*
Casdec *Allocasuarina decaisneana – Triodia basedowii*
DAstow *Acacia stowardii – Senna* spp. *– Eremophila* spp.
Erem *Eremophila* spp. *– Dodonaea spp. – Senna* spp. *– Acacia* spp.

Hummock grasslands – Australian Arid Zone (R. L. Specht and E. E. Hegarty, unpublished data)
HG1 *Triodia pungens*
HG2 *Triodia basedowii*
HG3 *Triodia wiseana*
HG4 *Triodia intermedia – Eriachne obtusa*
HG5 *Triodia inutilis – Triodia* spp.
HG6 *Plectrachne schinzii*
HG7 *Triodia clelandii – T. irritans*
HG8 *Triodia basedowii – Zygochloa paradoxa*
HG9 *Triodia irritans*
HG10 *Triodia hubbardii*

Chenopod low shrublands – southern Australian Arid Zone (R. L. Specht and A. Specht, unpublished data)
C4 *Atriplex vesicaria – Maireana sedifolia*
C7 *Myoporum platycarpum – Atriplex vesicaria* ± *Cratystylis* sp.
C13 *Alectryon oleifolius* ± *Atriplex vesicaria* ± *Maireana* spp.
C20 *Halosarcia* spp. (inland salt marsh)
C21 *Atriplex vesicaria – Maireana aphylla – M. planifolia*
C22 *Casuarina pauper – Maireana pyramidata*
C23a *Atriplex vesicaria – A. nummularia – Acacia pendula*
C24 *Acacia papyrocarpa – Maireana sedifolia*

Aquatic vegetation – tropical and subtropical northern Australia (R. L. Specht and M. P. Bolton, unpublished data)
A12 *Monochoria cyanea – Potamogeton tricarinatus*
A15/26 *Caldesia oligococca – Eleocharis dulcis – Nymphaea gigantea*
A27 *Caldesia oligococca – Nymphaea gigantea*
A28 *Typha domingensis – Caldesia oligococca*
A29 *Typha domingensis*
Restio *Leptocarpus spathaceus – Restio tetraphyllus*

Aquatic vegetation – temperate southern Australia (Aston, 1973; R. L. Specht and M. P. Bolton, unpublished. data)
A16 *Eleocharis acuta – Isolepis fluitans – Lepilaena* spp.
A17 *Lemna minor – Isoetes muelleri*
A18/19 *Eleocharis acuta – E. sphacelata*
A22 *Eleocharis acuta – E. sphacelata – Typha* spp. *– Maundia* sp.
A40/43 *Eleocharis acuta – E. sphacelata – Typha* spp. *– Marsilea* sp.
A41 *Eleocharis acuta – E. sphacelata – Typha* spp. *– Myriophyllum* sp.
A42 *Eleocharis acuta – E. sphacelata – Typha* spp.
A46 *Eleocharis acuta – Typha spp. – Potamogeton crispus*
A47 *Eleocharis acuta – Typha* spp.
Baumea *Baumea juncea – B. rubiginosa*
Gahnia *Gahnia filum – G. trifida*

Coastal dune vegetation (Specht, 1993)
CD8 *Olearia axillaris – Capparis spinosa – Spinifex longifolius*
CD12 *Banksia integrifolia – Casuarina equisetifolia – Spinifex sericeus*
CD13 *Calophyllum inophyllum – Spinifex sericeus*
CD14 *Calophyllum inophyllum – Casuarina equisetifolia – Spinifex longifolius*
CD15 *Casuarina equisetifolia – Argusia argentea – Spinifex longifolius*
CD18 *Olearia axillaris – Acacia cyclops – Spinifex hirsutus*
CD19 *Allocasuarina verticillata – Olearia axillaris – Spinifex sericeus*
CD20 *Banksia integrifolia – Leptospermum laevigatum – Leucopogon* sp.
CD21 *Banksia integrifolia – Leptospermum laevigatum – Myoporum* sp.
CD22 *Banksia integrifolia – Leptospermum laevigatum – Cupaniopsis* sp.
CD23 *Banksia integrifolia – Casuarina equisetifolia – Isolepis nodosa*
Pisonia *Pisonia grandis*

Coastal wetland vegetation (mangroves, salt marshes, and brackish wetlands) (Saenger *et al.*, 1977; Specht, 1981b)
CW5 *Avicennia marina – Aegiceras corniculatum – Halosarcia indica*
CW5d/19d *Casuarina glauca*
CW6 *Rhizophora stylosa – Avicennia marina – Aegialitis annulata*
Nypa *Nypa fruticans*
Melacac *Melaleuca acacioides*
CW7 *Aegialitis annulata – Avicennia marina – Rhizophora* spp.
CW19 *Avicennia marina – Aegiceras corniculatum – Halosarcia pergranulata*
CW36 *Avicennia marina – Halosarcia halocnemoides*
Melhal *Melaleuca halmaturorum*
CW37 *Avicennia marina – Halosarcia pergranulata*
Meleric *Melaleuca ericifolia – M. squarrosa*
CW8 *Sclerostegia arbuscula* (salt marsh)

Biogeographic Regions in Australia

The presence or absence of TWINSPAN floristic groups in each 1° latitude by 1° longitude grid-cell of the Australian continent was analyzed using the agglomerative classification program, PATN (Belbin, 1989). The symmetric form of the Kulczynski measure with $\beta \equiv -0.1$ was used to determine the association between community groups in each grid cell. This has been found to produce the most useful classification for such ecological applications (Faith *et al.*, 1987) and in this case produced the most "sensible" results. The results are plotted as a dendrogram in **Figure 2**. The resulting distribution of the

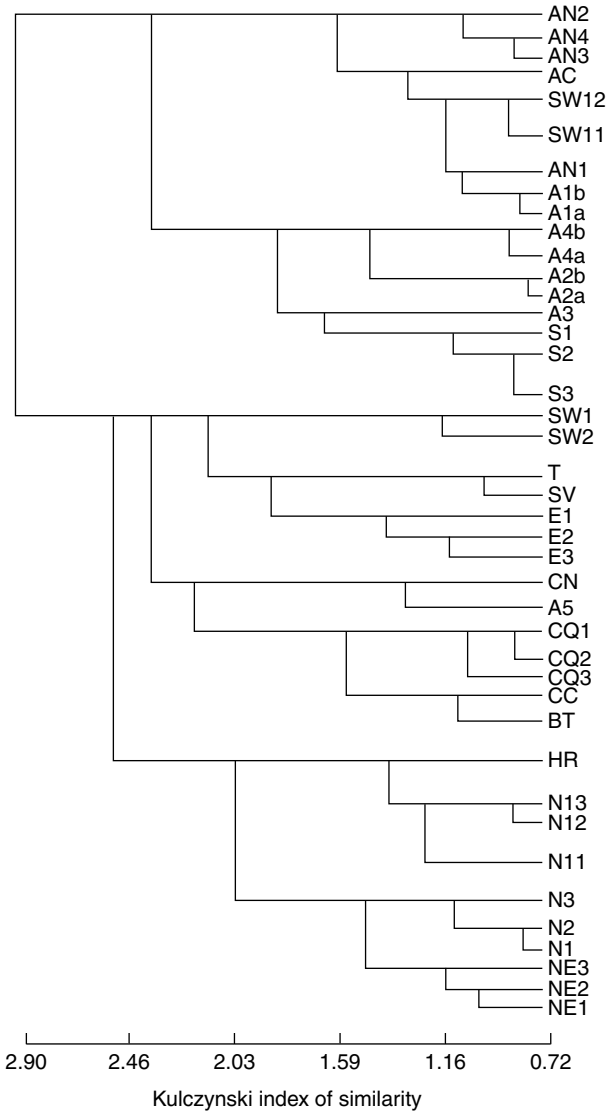


Figure 2 Dendrogram showing the relationships of the biogeographic regions of Australia derived from a Kulczynski symmetric pattern analysis of the major TWINSPAN floristic groups recorded in 1° latitude by 1° longitude grid-cells throughout Australia (Specht and Specht, 2002).

biogeographic regions identified by the foregoing analysis is plotted in **Figure 3**.

The first subdivision of the Kulczynski Symmetric dendrogram (**Figure 2**) separated the arid and semi-arid biogeographic regions of Western Australia, Northern Territory, South Australia, and north-western Victoria. Three biogeographic regions (AN2–4) were defined in the northern section of the arid zone, an area that receives mostly summer rainfall. The biogeographic regions of Central Australia (AC) and the Ashburton (ANI) of Western Australia receive both summer and winter rainfall. Although summer rains occur erratically, winter rainfall is predominant in the southern arid zone (A1a–b, A2a–b, A3, A4a–b) and in the semi-arid climatic zone (SWI1–2 and S1–3). The TWINSPAN floristic groups on the mid-tertiary limestone sediments from the Nullarbor Plain (A3) across South Australia (S1 and S2) into north-western Victoria (S3) are closely related.

The complex of TWINSPAN floristic groups (A1a–b) in the southern section of the arid zone of Western Australia are closely interrelated (and probably should not be separated); they are affiliated with the vegetation of the Ashburton region (AN1), where summer rainfall predominates. The northern section (A2a–b) of the arid zone in South Australia is distinctly separated from the southern arid vegetation (A4a–b) in which outliers of the mallee open-scrub (dominant in the semi-arid climate of S1–3) are to be found.

The second subdivision of the Kulczynski Symmetric dendrogram differentiated both the vegetation of the humid/perhumid section of south-western Australia (SW1–2, with a Mediterranean-type climate) and all the biogeographic regions in eastern Australia (where rainfall is evenly distributed throughout the year). Biogeographic regions were defined in the humid/perhumid climates of Tasmania to south-eastern Queensland (T, Tasmania; SV, Southern Victoria; and E1–3, Eastern Australia). Subhumid biogeographic regions were defined in central Queensland (CQ1–3) and in central New South Wales (CN), with related biogeographic regions (BT, the Barkly Tableland; CC, the Channel Country; and A5, Far Western New South Wales) in the semi-arid to arid zones of both states.

The third subdivision of the Kulczynski Symmetric dendrogram defined the biogeographic regions in northern Australia, where summer rainfall alternates with winter drought. The biogeographic regions of Cape York Peninsula (NE1) and the Gulf of Carpentaria (NE2) in Queensland are closely related to the wet tropical region (NE3) that extends southwards along the coast toward the tropic of Capricorn.

The TWINSPAN floristic groups of the humid/subhumid section of the Kimberley (N3) and the Top End of the Northern Territory (N1–2) are interrelated and were differentiated from the inland, semi-arid/sub-humid biogeographic regions (NI1–3) across northern Australia and the Hamersley Ranges (HR) in Western Australia.

Diagnostic Floristic Groups in Australian Biogeographic Regions

The TWINSPAN groups associated with each of the biogeographic regions are shown below.

N1 (humid north-western W.A.) – RF34(69), MRF12a, T371, T372.
N2 (humid N.T.) – MRF12a, T187a, T371, T372, T373a, G56, H7.
N3 (humid Cape York Peninsula and Gulf, Qld.) – RF35(140), RF35(141), T370, T372, T373a, G18, H12, H13.
N4 (Broome-Derby, W.A.) – RF16(132), T372.
N5 (southern Kimberley region, W.A.) – RF16(132), T187a, T371, T372.
N6 (subhumid N.T.) – T187a, T371, T372, H7.
N7 (semi-arid N.T.) – T187a, T372, DA8b.
N8 (Pilbara region, W.A.) – T372, DA100g.
NE1 (subhumid Cape York Peninsula, Qld.) – T188, T189, T368, T369, T371, T714, T801, G18.
NE2 (perhumid north-eastern Qld.) – RF35(142), RF35(143), RF36(144), RF36(145), T368, T369, T371, T710, T712, T713, T714, T715, T801.

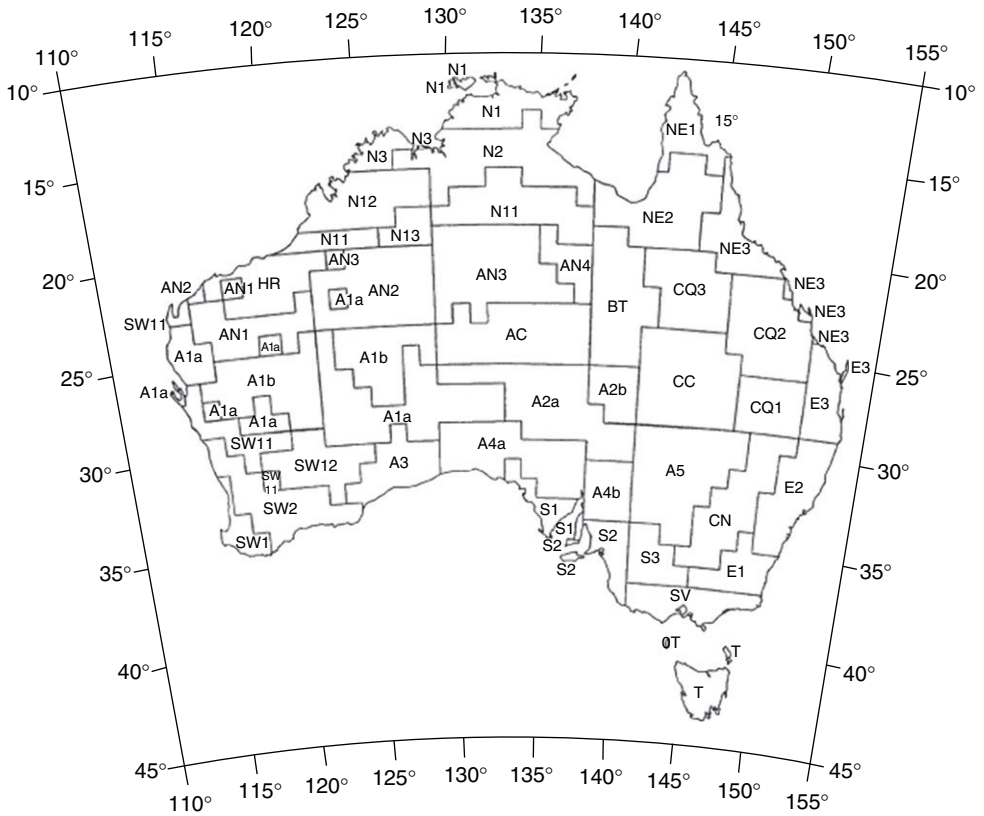


Figure 3 Biogeographic regions of Australia derived from a Kulczynski Symmetric Pattern Analysis of the major TWINSPAN floristic groups recorded in 1° latitude by 1° longitude grid-cells throughout Australia (Specht *et al.*, 1995).

NE3 (subhumid north-central highlands, Qld.) – T188, T189, T368, T369, T714, DA8b.
NE4 (subhumid central Qld.) – RF46(184), T188, T364, T352b, T365, T369, T714, T716a, T717, HW11, DA8a.
NE5 (subhumid south-eastern Qld.) – T188, T360, T363a, T364, T365, T369, T716a, T717, Calcolu, HW11.
BT (Barkly Tableland, Qld.) – T188, G8, G19.
CC (Channel Country, Qld.) – T188, DA18, DA19, C20.
SE1 (humid south-eastern Qld.) – RF45(180), RF41(332), T354, T707, T710, T711, T713, T714, T715, T716a, T717, G11, H82, H166, H167.
SE2 (perhumid Border Ranges, N.S.W.–Qld.) – RF42(169), RF40(322), RF40(325), T346, T707, T710, H82, H84, H166, H167.
SE3 (central to northern coasts and highlands, N.S.W.) – T334, T347, T350, T354, T707, T710, T716b, H82, H84, H166, H167.
SE4 (south-eastern and montane N.S.W.–Vic.) – RF25, T38, T79, T156, T157, T334, T335, T347, T348, T349, T350, AV4, AV6, H84.
SE5 (south-western to south-eastern Vic.) – RF24, T331, Eucpryor, T341, T342, T666, T667, G69, H170, H171.
SE6 (Tasmania) — RF7, RF13, T8, T36, T40, T148, T149, Eucnit, T666, T667, AV14, AV15, H170, H171.
SE7 (northern tablelands, N.S.W.) – T338, T339, T362, T363a, T717, Calcolu.
SE8 (southern and central tablelands, N.S.W.) – T190, T338, T339, T341, T362, T363a, T716b, Calcolu.
SE9 (humid Mt. Lofty Ranges and South East District, S.A.) – T190, T328a, T329, T330, T331, T667, H170.
S1 (Nullarbor Plain, S.A.) – M13, M45, C4, C20.
S2 (northern Eyre and Yorke Peninsulas, S.A.) – M25, M45, C24, DA14a, H44b.
S3 (subhumid S.A.) – T190, T328a, T329, T330, T331, T366a, T366b, T667, M13, M25, H44b, H170.
S4 (north-western Vic.) – T190, T329, T362, T366a, T366b, M10, M25, M45, H44a, H44b.
SW1 (humid to perhumid south-western W.A.) – SW4a, SW20, SW43a, SW47, SW84, SW85.
SW2 (subhumid to semi-arid south-western W.A.) – SW22, SW46, SW96a, SW97a, SW100, SW196.
A1a (south-eastern Pilbara District, W.A.) – C20, DA14a, DA22a, DA51, DA62, DA63a.
A1b (southern part of Murchison District to Great Victoria Desert, W.A.) – C20, DA22a, DA62, SW197.
A1c (northern part of Murchison District to Great Victoria Desert, W.A.) – C20, DA22a, DA62, DA63a.
A1d (southern N.T.) – C20, DA21, DA46a.
A1e (Great Victoria Desert, S.A.) – C20, C22, DA62.
A2a (Great Sandy Desert, W.A.) – C20, DA100g, DA101c, HG6.
A2b (Tanami Desert, N.T.) – C20, DA47, HG6.
A3a (northern S.A.) – C20, DA21, DA41a, DA46a.
A3b (Simpson Desert, S.A.) – C20, HG2, HG8.
A4a1 (Exmouth Gulf, W.A.) – C20, DA14a.

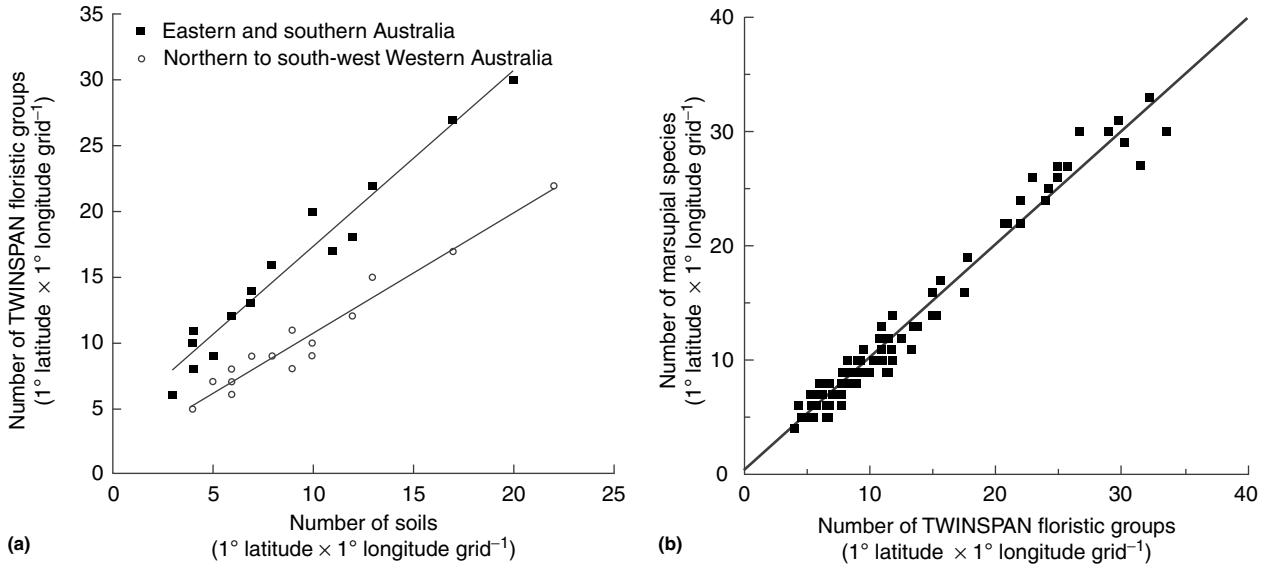


Figure 4 (a) The number of TWINSPAN floristic groups plotted against the number of principal soil types (Northcote *et al.*, 1960–1968) found in the same grid-cell of 1° latitude by 1° longitude dimensions in eastern and southern Australia (South-eastern Australian floristc groups = 1.34 soil groups + 3.83, $n = 15$, $r^2 = 0.96$) and on the nutrient-poor soils of northern, central, and western Australia (Western Australian floristic groups = 0.92, soil groups + 1.51, $n = 15$, $r^2 = 0.95$). (Specht and Specht, 2002; Specht and Tyler, 2010). (b) Number of marsupial species plotted against the number of TWINSPAN floristic groups for the same area (number of species of marsupials = 0.37 + 0.99 and number of floristic groups, $n = 90$, $r^2 = 0.97$).

A4a2 (north-eastern Wheat Belt, W.A.) – C20, DA14a, SW100, SW196.
A4b (Goldfields, W.A.) – C20, DA14a, DA62, SW26, SW51, SW99a, SW101b.
A5a (Nullarbor Plain to Flinders Ranges, S.A.) – C4, C7, C13, C20, C21, C24, DA41a.
A5b (North East District, S.A.) – C4, C7, C13, C20, C22, DA41a, T366a.
A6 (western N.S.W.) – C4, C20, C22, C23a, DA41a, T366a.

Community Diversity of Australian Ecosystems

Over large areas, such as the 1° latitude by 1° longitude grid-cell, the diversity of habitats is high, resulting in a high diversity of TWINSPAN floristic groups. This is illustrated by a positive correlation between the number of TWINSPAN floristic groups in each grid cell and the number of soil types occurring within them (**Figure 4(a)**). This has influence on other biotic groups, illustrated by the positive relationship between the species richness of marsupials per unit area reported by Pianka and Schall (1981), and the gamma diversity of plant communities **(Figure 4(b))**.

Species Richness of Australian Ecosystems

Species richness of overstory and understory strata

The potential for annual foliage growth increases in the 10 °C temperature gradient from temperate to tropical Australia (**Table 2**). The numbers of both stems and species recorded in the overstory (per hectare) increases exponentially with temperature zone and with moisture status (**Table 2** and **Figure 5(a)**) (after Specht and Specht, 1993, 1994).

Photosynthesis in overstory leaves, and hence potential annual foliage growth, depends primarily on the amount of solar radiation incident on the plant community, this declining from tropical to temperate Australia. In the open-structured plant communities of the wet–dry tropics and subtropics, a high percentage of direct-beam solar radiation reaches the understory through the gaps in the overstory canopy; species richness (the number of species per hectare) can become higher than that of the exposed overstory and follows the exponential curve shown for the overstory (**Figure 5(b)**). In more temperate latitudes, solar radiation reaches the Earth at a lower angle, and consequently the amount of radiation that reaches the understory is low (Specht *et al.*, 1992; Specht and Specht, 1999), resulting in only a slight increase from the semi-arid to the perhumid zone, with the proportion of cryptogamic species substantial in the most humid climates.

On nutrient-poor soils, (residual lateritic podsols, arenaceous, acid granitic, and serpentinitic soils, Quaternary sands, etc.) common in many parts of Australia, the species richness of the overstory is low, below that on higher-fertility soils in the same evaporative regime (Specht and Specht, 1989b). The species richness of the understory is positively related to solar radiation penetration as foliage cover in the overstory decreases (**Figure 6**).

Species Richness of Vertebrates and Invertebrates

The species richness of vertebrates recorded in a plant community is positively related to the species richness of vascular

Table 2 Structural characteristics of mature Australian plant communities[a]

Climate	*Stand height (m)*	*Stand density (stems ha^{-1})*	*Foliage Projective Cover (%)*	*Annual shoot growth (kg ha^{-1})*	*Overstory species (ha^{-1})*	*Understory species (ha^{-1})*
Tropical						
Perhumid	30–(45[b])–24	>5,000–2,500	100–65	12,500	140	150[c]
Humid	24–17	2,500–800	65–45	7,600	35	50
Subhumid	17–14	800–450	45–35	5,700	20	30
Semiarid	14–10	450–200	35–25	4,800	12	25
Arid	<10	>200	>25	3,600	8	15
Subtropical						
Perhumid	30–(50[b])–22	4,500–1,200	100–65	10,500	90	120[c]
Humid	22–15	1,200–500	65–45	6,000	20	60
Subhumid	15–12	500–300	45–35	4,400	10	40
Semiarid	12–8	300–200	35–25	3,500	6	35
Arid	<8	<200	<25	2,800	4	25
Temperate						
Perhumid	30–(60[b])–20	550–400	100–65	4,800	10	45[c]
Humid	20–14	400–300	65–45	2,700	5	45[c]
Subhumid	14–11	300–250	45–35	1,600	4	44
Semiarid	11–7	250–200	35–25	1,100	3	43
Arid	<7	<200	<25	600	2	40

[a]Based on Specht and Specht (1989a).
[b]Stand heights of emergent eucalypts at the edge of rainforest stands.
[c]The species richness of understories in perhumid to humid climates contains a mixture of regenerating overstory species as well as shadeloving vascular plants and cryptogams.

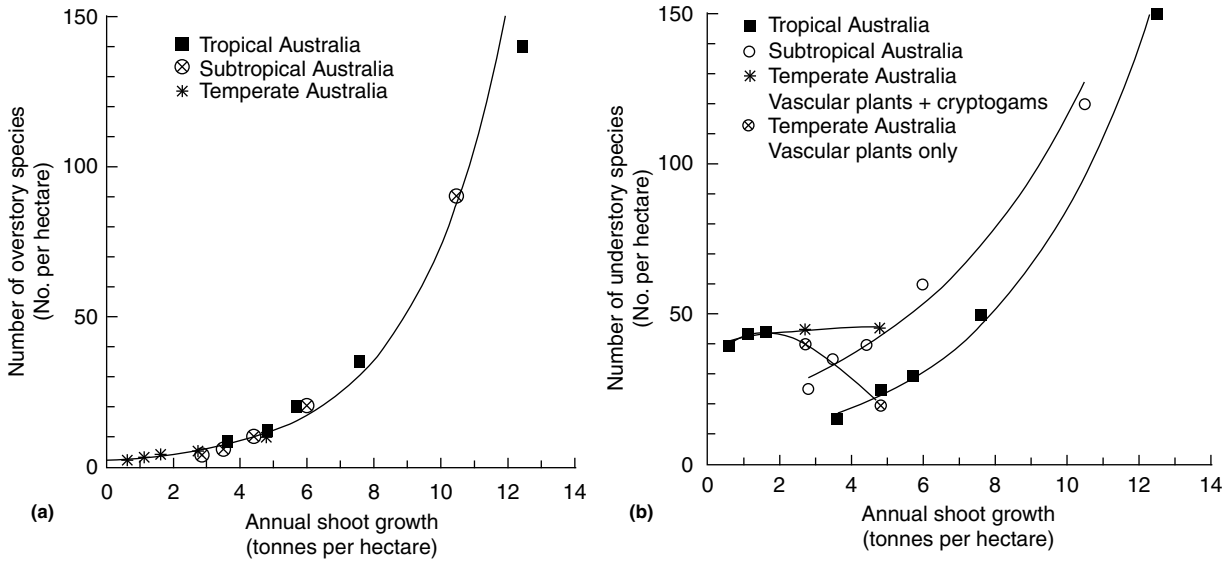


Figure 5 (a) Species richness (number of species per hectare) of the overstory increases exponentially as annual shoot growth of this stratum increases: $ln\text{Ns} \equiv 0.70 + 0.36$ annual shoot growth ($n \equiv 12$, $r^2 \equiv 0.98$). Species richness (N) increases exponentially (1) in the evaporative gradient from the arid zone (Evaporative coefficient $k < 0.35 \times 10^{-2}$) to the perhumid zone (Evaporative coefficient $k \equiv 1.0 \times 10^{-2}$), and (2) in the 10 °C temperature gradient from temperate to subtropical to tropical Australia (Specht and Specht, 1993). (b) The low amount of solar radiation that penetrates the overstory to reach the understory in the temperate latitudes of Australia has little effect on species richness (number of species per hectare) of this stratum. In the tropics and subtropics, however, the relationship of species richness of the understory to annual shoot growth (of the upper stratum) is similar to that shown by the overstory (Specht and Specht, 1993, 1994, 2002).

plants found in the understory stratum (**Figure 7(a)** and 7(b)), as indeed is the number of frog and snake species (Braithwaite *et al.*, 1985; P. C. Catling, in Specht, 1988b; Cody, 1994a and Cody, 1994b; Specht, 1994; Specht and Specht, 1999; Specht and Tyler, 2010). These may also be related to potential energy-fixation. An inverse correlation has been observed, however, for lizards, presumably reflecting their food and habitat preference.

It would also appear from the limited data available that the species richness of epigaeic invertebrates collected in pit-fall traps decreases as the annual input of leaf litter (Specht and Brouwer, 1975) from the overstory decreases from the humid to the semi-arid climatic zone in the Mediterranean-type climate of southern Australia (**Table 3**; after Greenslade and Majer, 1985, and Specht, 1988).

Dryland Vegetation Continua

As the depth of nutrient-poor soils becomes shallower, less soil water is available for evapotranspiration, the transport of phosphate and other nutrient ions in the transpiration stream is reduced, and the growth of the overstory is limited (**Figure 8(a)**; Specht and Specht, 2010). The low level of phosphate and nitrogen ions in the soil influences mycorrhizal association in roots of seedling eucalypts (Burrell, 1981). In extreme situations, such as on serpentine substrates, the foliage projective cover of the overstory, the annual shoot growth (per hectare), and the stand height are all reduced (**Figure 8(b)**). As more solar radiation penetrates the gaps between the trees, both the foliage projective cover and the species richness of the understory increases. The sum of the foliage covers of overstory and understory strata remains a constant (**Figure 8(c)**) in equilibrium with the evaporative aerodynamics (**Figure 1**) of the area (Specht and Morgan, 1981; Specht, 1983; Specht *et al.*, 1990).

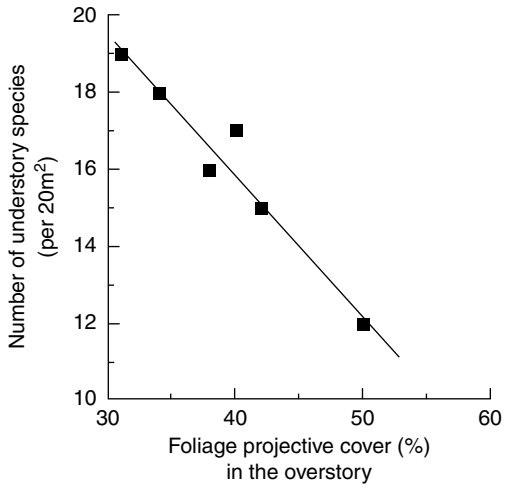


Figure 6 The relationship between species richness of the understory in the open-forest on the serpentinite soil catena in Central Queensland (Batianoff *et al.*, 1997; Specht *et al.* 2006).

Wetland Vegetation Continua

The density of the overstory, and hence its foliage projective cover, decreases as the soils become more waterlogged (anaerobic) in wetland continua (Specht, 2009). The foliage projective cover of the wetland understory increases to maintain the constant sum of the foliage projective covers of overstory and understory strata. However, unlike the increasing species richness in the vegetation continuum on dryland soils described in the previous section, species richness of the wetland understory decreases in proportion to the number of months the roots are subjected to waterlogging.

Postfire Succession

The shading effect of the overstory on the species richness in the understory has been demonstrated when long-lived shrubby species, such as *Banksia ornata* (Proteaceae),

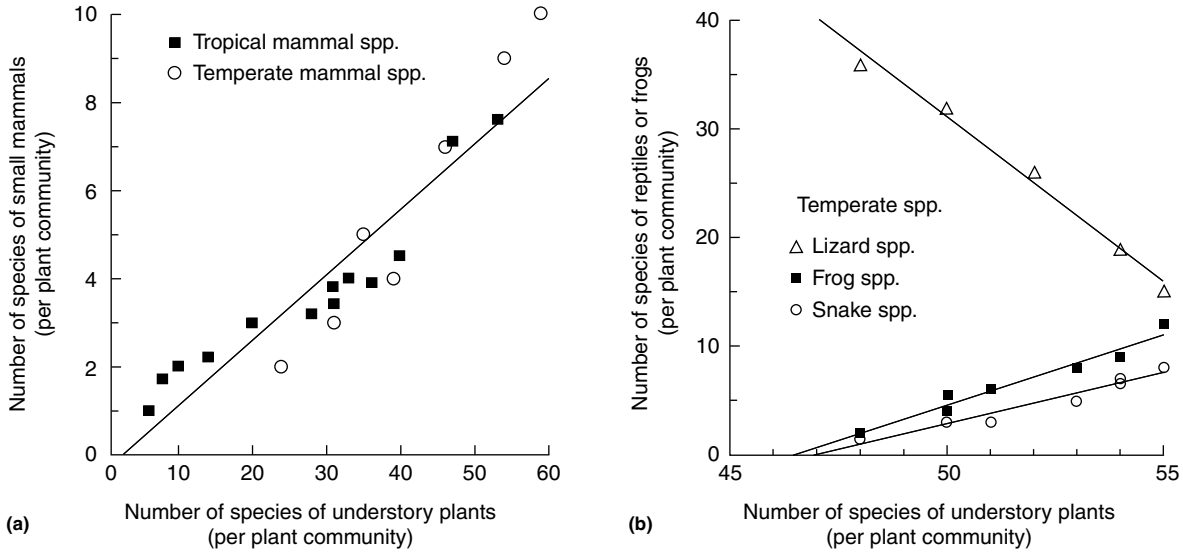


Figure 7 (a) The number of species of small mammals recorded in ecosystems in tropical and temperate Australia plotted against the number of species of vascular plants recorded in the understory of each plant community. (b) The number of species of lizards, snakes, and frogs recorded in ecosystems in temperate Australia plotted against the number of species of vascular plants recorded in the understory of each plant community (based on Specht and Specht, 2002).

Table 3 Number of groups[a] of epigaeic invertebrates collected (throughout the year) in pitfall traps established within four vegetation formations typical of humid to semiarid mediterranean-type climates of southern Australia[b]

Locality	*Vegetation formation*	*Evaporative coefficient*	*Annual leaf litter (kg ha^{-1})*	*Epigaeic invertebrates (number of groups)*[c]
Dwellingup, W.A. (32°43′ S, 116°04′ E)	Eucalypt open-forest	0.53×10^{-2}	1,300	19
Kuitpo, S.A. (33°15′ S, 138°43′ E)	Eucalypt woodland	0.45×10^{-2}	1,050	13
Wyperfield, Vic. (35°35′ S, 142°00′ E)	Eucalypt open-scrub	0.42×10^{-2}	950	12
Wyperfield, Vic. (35°35′ S, 142°00′ E)	Heathland	0.40×10^{-2}	850	10

[a]Invertebrate groups – Acarina, Araneae, Isopoda, Amphipoda, Diplopoda, Chilopoda, Collembola, Insecta (Thysanura, Blattodea, Isoptera, Dermaptera, Orthoptera, Embioptera, Psocoptera, Hemiptera, Homoptera, Thysanoptera, Coleoptera, Diptera, Lepidoptera, and Hymenoptera, including Formicidae).
[b]After Greenslade and Majer (1985) and Greenslade and Majer, in Specht (1988b).
[c]Monthly actual/potential evapotranspiration per millimeter of available water.

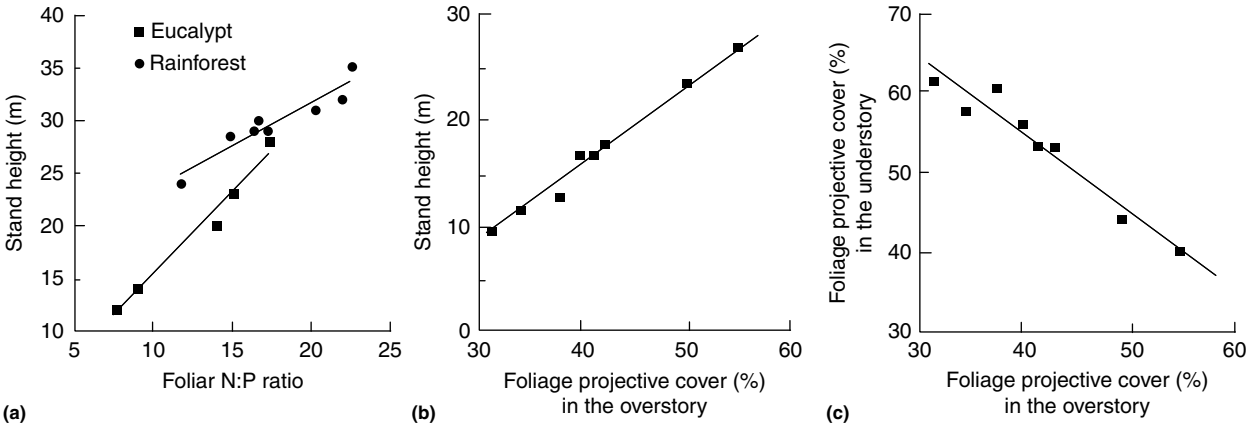


Figure 8 The relationship between (a) stand height and nitrogen and phosphorus supply, and for serpentine substrates, (b) stand height of the overstory, and (c) foliage cover of the understory and overstory foliage cover in the open-forest on the serpentine soil catena in Central Queensland.

regenerate from seed in a secondary succession following the fire that regularly razes Australian heathlands (**Figure 9**). Application of phosphate fertilizer to these nutrient-poor ecosystems results in more rapid growth of overstory shrubs to the detriment of the understory species (Specht and Specht, 1989c).

Long-lived overstory species, regenerating from root-stocks after fire, rapidly develop maximum foliage projective cover in the overstory which shades out understory plants. However, the sum of the foliage projective covers of overstory and understory strata remains a constant throughout the post-fire succession (Specht, 1969).

Tall open-forests of mountain ash (*Eucalyptus regnans*) in south-eastern Australia are killed by fire. Dense stands of saplings soon regenerate and over-shadow developing understory plants. Gradually over many decades the foliage projective cover of the overstory thins to the norm for the aerodynamic climate of the region and enables the understory to develop (Ashton, 1976).

Invasive Species in the Savanna Understory

In the Mediterranean-type climate of southern Australia, the grassy understory of savanna woodlands has been invaded by short-lived herbaceous species. These introduced plants deplete soil water during their spring growing-season, before the late spring into summer growth rhythm of the native perennial grasses. Species richness of the understory is increased as the growth of the perennial grasses slowly declines (Specht, 2000). The species richness of the understory is further increased by destruction of overstory plants (Specht *et al.*, 1990).

Species Richness of Disjunct Stands of Rainforest

Foliage at the edges of isolated rainforest stands, in habitats well-watered throughout the year is abraded by aerodynamic forces, thus reducing the total amount of solar energy fixed by the stand. As the area of the rainforest stand increases, the

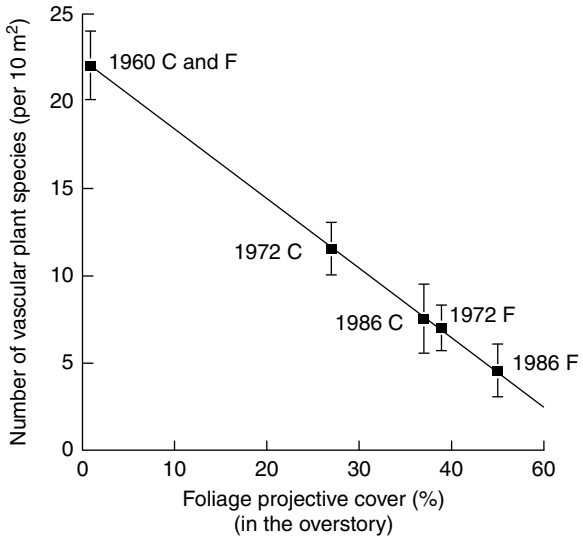


Figure 9 The species richness of Dark Island heathland, South Australia (Specht and Specht, 1989c) decreased as foliage cover of the overstory (*Banksia ornata*, *B. marginata*, *Allocasuarina pusilla*, and *Xanthorrhoea caespitosa*) increased with time after fire control (C) and with phosphate fertilizer (F).

number of species of vascular plants, also resident birds, increases asymptotically until it reaches a maximum when the area of the stand reaches 100–200 hectares (Specht, 1988; Specht, 2007).

Conclusions

The vegetation of Australia in the Cretaceous–Early Tertiary, fifty- to a hundred-million years ago, had developed on infertile lateritic soils, which extended across the Gondwanan super-continent of which Australia was part. The tropical and temperate rainforests, savanna and heathy floras of Australia therefore show strong family and generic (sometimes specific) affinities with the floras of India, Africa and, to a lesser extent, with South America (Specht, 1981).

Although Australia was at this time contiguous with Antarctica at latitude 60–65° S, palaeo-oxygen analyses of sea-bed sediments indicate that the climate was subtropical with a mean annual temperature of 19.5 °C (Specht *et al.*, 1992). Increasing aridity during the Tertiary (Martin, 2006), as well as the movement of the Australian Tectonic Plate northward towards the Equator at 66 mm per year (Wellman and McDougall, 1974), imposed great stresses on the Australian vegetation. Remnants of the Gondwanan flora persisted as fragments in more humid habitats. In addition, a great diversity of taxa evolved during the Early Tertiary to occupy the greater part of the continent, which today experiences varying degrees of seasonal drought (Specht, 1981; Specht and Specht, 1999). The convergence of the Australian Plate with the Sunda Plate about 15 million years ago allowed some migration of the Australian flora northward onto Sundaland, but little, if any, migration southward from southeast Asia.

The evaporative and frictional forces of the atmosphere, flowing over and through the vegetation across the Australian continent, molded foliage distribution and structure within all plant communities; distinctive plant formations resulted from the perhumid to the arid evaporative zone, from the tropical to the temperate region. Species richness (per hectare) of vascular plants in the overstory of Australian plant communities is correlated with the annual shoot growth (per hectare) of that stratum. Similarly, the species richness of vascular plants in the understory is related to annual shoot growth in that stratum – an attribute influenced by the amount of solar radiation transmitted through the overstory stratum.

The species richness (per hectare) of nonarboreal vertebrates is correlated with species richness of the understory stratum. Species richness of epigaeic invertebrates appears to depend on the annual input of leaf litter from the overstory foliage. The diversity of TWINSPAN floristic groups and associated vertebrates is also correlated with the optimal annual shoot growth (per hectare) of the overstory stratum of Australian plant communities.

The community-physiological processes that determine the species richness of vascular plants, vertebrates, and invertebrates and the diversity of floristic groups need urgent scientific investigation before they will be better understood.

See also: Africa, Ecosystems of. Antarctic Ecosystems. Asia, Ecosystems of. Ecosystems of South America

References

Ashton DH (1976) The development of even-aged stands of Eucalyptus regnans F. Muell. in Central Victoria. *Australian Journal of Botany* 24: 397–414.

Aston HI (1973) *Aquatic Plants of Australia*, 368 pp. Melbourne: Melbourne University Press.

Batianoff GN, Reeves RD, and Specht RL (1997) The effect of serpentine on vegetation structure, species diversity and endemism in Central Queensland. In: Jaffré T, Reeves RD, and Becquer T (eds.) *The Ecology of Ultramafic and Metalliferous Areas*, pp. 147–153. New Caledonia: ORSTOM Center de Nouméa.

Belbin L (1989) *PATN Technical Reference*. Canberra: CSIRO Division of Wildlife and Ecology.

Braithwaite RW, Winter JW, Taylor JA, *et al.* (1985) Patterns of diversity and structure of mammalian assemblages in the Australian tropics. *Australian Mammalologist* 8: 171–186.

Burrell JP (1981) Invasion of coastal heaths of Victoria by Leptospermum laevigatum (J. Gaertn.) F. Muell. *Australian Journal of Botany* 29: 747–764.

Clifford HT and Simon BK (1981) The biogeography of Australian grasses. In: Keast A (ed.) *Ecological Biogeography of Australia*, pp. 537–554. The Hague, Netherlands: Dr W. Junk.

Cody ML (1994a) Bird diversity within and among Australian heathlands. In: Arianoutsou M and Groves RH (eds.) *Plant–Animal Interactions in Mediterranean-Type Ecosystems*, pp. 47–61. Dordrecht, Netherlands: Kluwer Academic Publishers.

Cody ML (1994b) Mulga bird communities. I. Species composition and predictability across Australia. *Australian Journal of Ecology* 19: 206–219.

Faith DP, Minchin PR, and Belbin L (1987) Compositional dissimilarity as a robust measure of ecological distance: A theoretical model and computer simulation. *Vegetation* 69: 57–68.

Feary DA, Davies PJ, Pigram CJ, *et al.* (1991) Climatic evolution and control of carbonate deposition in north-east Australia. *Palaeogeography,*

Palaeoclimatology, Palaeoecology (Global and Planetry Change Section) 89: 341–361.
Forster PI, Bostock PD, Bird LH, *et al.* (1991) Vineforest Plant Atlas for South-East Queensland. *Queensland Herbarium, Brisbane* 496.
George AS, Hopkins AJM, and Marchant NG (1979) The heathlands of Western Australia. In: Specht RL (ed.) *Ecosystems of the World. vol. 9A. Heathlands and Related Shrublands. Descriptive Studies*, pp. 211–230. Amsterdam: Elsevier.
Greenslade P and Majer, JD (eds.) (1985) *Soil and Litter Invertebrates of Australian Mediterranean-Type Ecosystems*. Western Australian Institute of Technology, School of Biology, Perth, Bulletin No. 12.
Hill MO (1973) Reciprocal averaging: an eigen vector method of ordination. *Journal of Ecology* 61: 237–249.
Hill RS (1994) The history of selected Australian taxa. In: Hill RS (ed.) *History of Australian Vegetation: Cretaceous to Recent*, pp. 390–419. Cambridge: Cambridge University Press.
Hill RS (2004) Origins of the south-eastern Australian vegetation. *Philosophical Transactions of the Royal Society of London B* 359: 1537–1549.
Hooker JD (1860) On the flora of Australia, being part of an Introductory Essay to Flora Tasmaniae. In: *The Botany of the Antarctic Voyage of H. M. Discovery Ships Erebus and Terror, in the Years 1839–1843*, Part 3, vol. 1 , pp. xxvii–cxxviii. London: Lords Commisioners of the Admiralty.
Johnson LAS and Briggs BG (1981) Three old southern families – Myrtaceae, Proteaceae and Restionaceae. In: Keast A (ed.) *Ecological Biogeography of Australia*, pp. 427–469. The Hague, Netherlands: Dr W. Junk.
Kirkpatrick JB (1989) The comparative ecology of mainland Australian and Tasmanian alpine vegetation. In: Good R (ed.) *The Scientific Significance of the Australian Alps*, pp. 127–142. Canberra: Australian Academy of Science.
Ladiges PY, Udovicic F, and Nelson G (2003) Australian biogeographical connections and the phylogeny of a large genera in the plant family Myrtaceae. *Journal of Biogeography* 30: 989–998.
Martin HA (1981) The Tertiary flora. In: Keast A (ed.) *Ecological Biogeography of Australia*, pp. 391–406. The Hague, Netherlands: Dr W. Junk.
Martin HA (2006) Cainozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments* 66: 533–563.
Martin HA and Gadek PA (1988) Identification of Eucalyptus spathulata pollen and its presence in the fossil record. *Memoirs of the Association of Australian Palaeontologists* 5: 311–327.
Northcote KH, Hubble GD, Isbell RF, *et al.* (1960–1968) *Atlas of Australian Soils*, pp. 1–10. Melbourne: CSIRO Australia and Melbourne University Press.
Pianka ER and Schall JJ (1981) Species densities of Australian vertebrates. In: Keast A (ed.) *Ecological Biogeography of Australia*, pp. 1675–1694. The Hague, Netherlands: Dr. W. Junk.
Pole M (1989) Early Miocene floras from central Otago, New Zealand. *Journal of the Royal Society of New Zealand* 19: 121–125.
Pole M (1993) Early Miocene flora of the Manuherikia Group, New Zealand. 7. Myrtaceae, including Eucalyptus. *Journal of the Royal Society of New Zealand* 23: 313–328.
Prescott JA and Pendleton RL (1952) Laterite and lateritic soils. *Commonwealth Bureau of Soil Science and Technology, Communication No* 47: 51.
Russell-Smith J (1991) Classification, species richness and environmental relations of monsoonal rainforest in northern Australia. *Journal of Vegetation Science* 2: 259–278.
Saenger P, Specht MM, Specht RL, *et al.* (1977) Mangal and coastal salt marsh communities in Australasia. In: Chapman VJ, (ed.) *Ecosystems of the World. Wet Coastal Formations,* vol. 1, pp. 293–345. Amsterdam: Elsevier.
Shackleton NJ and Kennett JP (1975) Palaeotemperature history of the Cenozoic and the initiation of Antarctic glaciation: Oxygen and carbon isotope analyses in DSDP sites 277, 279, 281. *In: Initial Reports of the Deep Sea Drilling Project, 2.* pp. 743–755. Washington: United States Government Printing Office.
Specht A (1988). In: Planners North Pty Ltd. (ed.) *Big Scrub Conservation Strategy. Resource Material*, vol. 2, pp. 2.1–2.72. Australia: New South Wales National Parks and Wildlife Service Sydney.
Specht A and Specht RL (1993) Species richness and canopy productivity of Australian plant communities. *Biodiversity and Conservation* 2: 152–167.
Specht A and Specht RL (1994) Biodiversity of overstory trees in relation to canopy productivity and stand density in the climatic gradient from warm temperate to tropical Australia. *Biodiversity Letters* 2: 39–45.
Specht RL (1958) Geographical relationships of the flora of Arnhem Land. In: Specht RL and Mountford CP, (eds.) *Records of the American-Australian Scientific Expedition to Arnhem Land. Botany and Plant Ecology,* vol. 3, pp. 415–478. Melbourne: Melbourne University Press.
Specht RL (1969) A comparison of the sclerophyllous vegetation characteristic of Mediterranean-type climates in France, California and southern Australia. 1. Structure, morphology and succession. 2. Dry matter, energy and nutrient accumulation. *Australian Journal of Botany* 17(277–292): 293–308.
Specht RL (1972) Water use by perennial, evergreen plant communities in Australia and Papua New Guinea. *Australian Journal of Botany* 20: 273–299.
Specht RL (1981a) Evolution of the Australian flora: Some generalizations. In: Keast A (ed.) *Ecological Biogeography of Australia*, pp. 783–806. The Hague, Netherlands: Dr. W. Junk.
Specht RL (1981b) Biogeography of halophytic angiosperms (salt-marsh, mangrove, sea-grass). In: Keast A (ed.) *Ecological Biogeography of Australia*, pp. 575–590. The Hague: Junk.
Specht RL (1983) Foliage projective covers of overstory and understory strata of mature vegetation in Australia. *Australian Journal of Ecology* 8: 433–439.
Specht RL (1988a) Origin and evolution of terrestrial plant communities in the wet-dry tropics of Australia. *Proceedings of the Ecological Society of Australia* 15: 19–30.
Specht RL (ed.) (1988b) *Mediterranean Type Ecosystems. A Data Source Book.* Dordrecht, Netherlands: Kluwer Academic Publishers.
Specht RL (1990) Forested wetlands in Australia. In: Lugo AE, Brinson MM, and Brown Sandra (eds.) *Ecosystems of the World. Wetland Forests,* vol. 16, pp. 387–406. Amsterdam: Elsevier.
Specht RL (1993) Dry coastal ecosystems of Australia — An overview of the dune vegetation. In: van der Maarel E, (ed.) *Ecosystems of the World. Dry Coastal Ecosystems. Africa, America, Asia, and Oceania,* vol. 2B, pp. 223–237. Amsterdam: Elsevier.
Specht RL (1994) Species richness of vascular plants and vertebrates in relation to canopy productivity. In: Arianoutsou M and Groves RH (eds.) *Plant-Animal Interactions in Mediterranean-Type Ecosystems*, pp. 15–24. Dordrecht, Netherlands: Kluwer Scientific Publications.
Specht RL (2000) Savanna woodland vegetation in the South-East District of South Australia: the influence of evaporative aerodynamics on the foliage structure of the understory invaded by introduced annuals. *Australian Ecology* 25: 588–599.
Specht RL (2007) Species richness of rainforest stands on nonserpentinite and serpentinite substrates in the Rockhampton-Marlborough area of Central Queensland. *Proceedings of the Royal Society of Queensland* 113: 17–35.
Specht RL (2009) Structure and species richness in wetland continua on sandy soils in subtropical and tropical Australia. *Australian Ecology* 34: 761–772.
Specht RL, Batianoff GN, and Reeves RD (2006) Vegetation structure and biodiversity along the eucalypt forest to rainforest continuum on the serpentinite soil catena in a sub-humid area of Central Queensland, Australia. *Australian Ecology* 31: 394–407.
Specht RL and Brouwer YM (1975) Seasonal shoot growth of Eucalyptus spp. in the Brisbane area of Queensland (with notes on shoot growth and litter fall in other areas of Australia). *Australian Journal of Botany* 23: 459–474.
Specht RL, Dettmann ME, and Jarzen DM (1992) Community associations and structure in the Late Cretaceous vegetation of southeast Australasia and Antarctica. *Palaeogeography, Palaeoclimatology, Palaeoecology* 94: 283–309.
Specht RL, Grundy RI, Specht A, *et al.* (1990) Species richness of plant communities: Relationship with community growth and structure. *Israel Journal of Botany* 39: 465–480.
Specht RL and Morgan DG (1981) The balance between the foliage projective covers of overstory and understory strata in Australian vegetation. *Australian Journal of Ecology* 6: 193–202.
Specht RL and Specht A (1989a) Canopy structure in Eucalyptus-dominated communities in Australia along climatic gradients. *Oecologia Plantarum* 10: 191–202.
Specht RL and Specht A (1989b) Species richness of overstory strata in Australian plant communities. The influence of overstory growth rates. *Australian Journal of Botany* 37: 321–336.
Specht RL and Specht A (1989c) Species richness of sclerophyll (heathy) plant communities in Australia. The influence of overstory cover. *Australian Journal of Botany* 37: 337–350.
Specht RL and Specht A (2002) *Australian Plant Communities: Dynamics of Structure, Growth and Biodiversity.* Melbourne: Oxford University Press.
Specht RL and Specht A (2010) The ratio of foliar nitrogen to foliar phosphorus: A determinant of leaf attributes in life-forms of subtropical and tropical rainforests. *Australian Journal of Botany* 58(7): 527–538.
Specht RL, Specht A, Whelan MB, *et al.* (1995) *Conservation Atlas of Plant Communities in Australia.* New South Wales: Southern Cross University, Lismore.

Specht RL and Tyler MJ (2010) The species richness of vascular plants and amphibia in major plant communities in temperate to tropical Australia: Relationship with annual biomass production. *International Journal of Ecology*, Article ID 635852, 17 pp. doi: 10.1155/2010/635852.

Specht RL and Tyler MJ (2010) The species richness of vascular plants and amphibia in major plant communities in temperate to tropical Australia: Relationship with annual biomass production. *International Journal of Ecology*, 17 pp. doi: 10.1155/2010/635852.

Webb LJ, Tracey JG, and Williams WT (1984) A floristic framework of Australian rainforests. *Australian Journal of Ecology* 9: 169–198.

Wellman P and McDougall I (1974) Cainozoic igneous activity in eastern Australia. *Tectonophysics* 23: 49–65.

B

Bacterial Biodiversity

Erko Stackebrandt, DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 325–337, © 2001, Elsevier Inc.

Glossary

Domains The highest taxonomic rank defined to classify organisms into Archaea, Bacteria, and Eucarya, which differ from each other in fundamental genomic and phenetic properties.

Endosymbiont Specialized form of symbiosis in which one partner thrives within cells, lumen, or tissues of the host organism; most obligate endosymbionts belong to the group of uncultured organisms.

Epigenetic level Study of molecules which are the product of gene expression.

Homology Denoting common ancestry. Structures, processes, sequences, behaviors, etc. are said to be homologous if there is evidence that they are derivations from a common ancestral structure. In molecular biology the term indicates a significant degree of similarity between DNA or proteins.

Phylogeny Natural relatedness among life forms; the science of ordering the genealogy of organisms into a family tree.

Prokaryotes Life forms that are members of the domains Archaea and Bacteria as opposed to Eucarya, comprising organisms with a cell nucleus.

rDNA Genes coding for rRNA that play a fundamental role in the translation process; the most thoroughly studied molecule in prokaryotic cells; used in comparative phylogenetic studies.

Uncultured prokaryotes Organisms, the presence of which has been detected by molecular methods in the environment but they have not been cultured under artificial laboratory conditions.

Recognizing Gaps in Knowledge

Due to progress in methodologies and concepts, facets of biodiversity have not been covered equally in different areas in bacteriology, but progress must be viewed as a logical consequence of available technologies. Introduction of groundbreaking methods is usually followed by a period of increased knowledge in those fields for which the methods were developed. Numerous examples exist in microbiology, such as the elucidation of (i) the ultrastructure of the prokaryotic cell, following the development of the electron microscope; (ii) metabolic and biochemical pathways, following the introduction of the isotope label technique and enzymology; (iii) anaerobic organisms, following the development of appropriate anaerobic cultivation technology; and (iv) cell constituents, following the introduction of the amino acid analyzer, gas chromatography, high-performance liquid chromatography, and thin layer chromatography. Recently, the introduction of high-resolution gel electrophoresis led to one- and two-dimensional fingerprinting methods for proteins, ribonucleic acids (RNAs), and deoxyribonucleic acids (DNA); application of restriction enzymes, cloning strategies, and polymerase chain reaction (PCR) technology led to improved sequence analysis of genes and genomes; and cloning of PCR fragments of environmental DNA, *in situ* hybridization, and the development of gradient gel electrophoresis revealed a larger spectrum of prokaryotic diversity than previously known. The latter insights have been used to develop strategies in which DNA, isolated from environmental samples, is expressed to yield a spectrum of novel enzymes not detected in strains available from biological resource centers. As a result, taxonomists learned that the vast majority of prokaryotic species are still undetected, and physiologists, biochemists, and geneticists can expect to find novel lines of descent containing organisms that express fundamental deviations from currently known biochemical pathways. New models will help to better understand the structure and function of living matter and the role of prokaryotes in maintaining the biosphere.

The number of validly described species of animals, plants, and lower eukaryotes is approximately 400 times larger than

the number of prokaryotic species (1,600,000 : 4,000). This fraction of bacterial and archaeal species is surprisingly low considering that prokaryotic species evolved eons ago, exploring and occupying any niche that has been investigated for the presence of prokaryotic organisms. Insects, on the other hand, which comprise more than 1 million species, evolved late in evolution—less than 600 million years ago during the Cambium. Certainly, differences in the species definition that exist between the biological species of higher evolved eukaryotes and the pragmatically defined prokaryotic species contribute to the tremendous discrepancy, but it has been shown convincingly that the vast majority of prokaryotic strains which are part of the free-living microbial community species have not been cultured. Endosymbiotic prokaryotes which are not free living but firmly associated with eukaryotic cells are another source of uncultured organisms. The limitation of recognizing the richness of microbial organisms is most likely due to the use of a restricted spectrum of enrichment media which selects for a very narrow spectrum of organisms that compete best under artificial laboratory conditions.

The discussion of diversity issues and the increasing awareness of the importance of microorganisms in maintaining the biosphere is embedded in discussion of the implementation of the articles of the Convention on Biological Diversity (CBD). The driving force behind the CBD has been the recognition of the possibility of a significant reduction or loss of biological diversity at source by human activities, and the preamble admits a general lack of information and knowledge regarding biological diversity. Nowhere is the lack of knowledge more acute than for microbial diversity (prokaryotes, fungi, yeasts, and heterotrophic and autotrophic protists) and there is a widespread appreciation among microbiologists that cultured microorganisms represent a very small, not necessarily ecologically important, fraction of natural microbial diversity.

Considering the total number of nucleotides per genome, far more recognizable diversity can be seen at the molecular level than at the epigenetic level. Compared to the total number of described species, the number of organisms analyzed in genome sequencing projects is still small (about 2.5%), but this portion will increase rapidly and genomic screening and data mining will develop into dominating biological disciplines in the future. However, taxonomists will not lose interest in the more traditional properties of a cell because of the demanding process of describing species and genera. The polyphasic approach to classification requires information about morphology, ultrastructure, metabolism, chemical structure of cell constituents, and genomic features. Thus, from more than 100 years of bacterial classification, an enormous wealth of phenotypic and epigenetic data have accumulated and are of benefit to bacterial taxonomy, which many scientists regard as the mother of biological sciences. However, taxonomists are concerned that this spectrum of diversity will have to be sacrificed for an approach that considers molecular/genomic data more appropriate for delineating taxonomic ranks. Indeed, this cut will happen sooner or later, but it does not mean that the avalanche of genomic data will bury all activities directed toward the elucidation of the phenotype.

Unraveling Phylogenetic Diversity

Whereas the past century unraveled the diversity of epigenetic properties, i.e., those characters that are the result of gene expression, the next century will prove a tremendous wealth of information on genomic properties. The terms genomics and reverse genetics have been coined for the strategy of obtaining value-added information from sequences of genomes and genes. Data mining and meganetworking programs will replace the simple search for similarities between a few homologous genes, and it can be expected that the importance of phenotypic traits will be reintroduced as a consequence of the desire to understand the horizontal and vertical flow of genes and their regulation and expression.

The combination of information of a (small) portion of the genome with phenotypic properties is nowhere exemplified better than in modern classification strategy of prokaryotes. The first component, the backbone of the system, is provided by the primary structure of homologous molecules which have accompanied the organisms since early evolution and they can be ranked according to their evolutionary history. The number and nature of sequence differences among proteins and genes coding for rRNA and proteins allow the recognition of pairs or groups of organisms which evolved from a common ancestor, and the order in which lineages have emerged in time facilitates decisions about the grouping of organisms.

The main problem that has emerged in phylogenetic studies during the past 10 years is the question of whether to place emphasis on the genealogical relationships derived from comparative analyses of a single homologous gene, a gene cluster, or even large parts of the genome. It has been demonstrated convincingly that the historic fate of ribosomal (r)DNA genes does not necessarily represent the fate of other genes because the phylogenetic branching patterns of different genes may show significant deviations. Analysis of the presence and position of genes in completely sequenced prokaryotic and eukaryotic genomes and horizontal gene transfer among genomes with a high plasticity has played an enormous role in designing and shaping early evolved organisms, and the chimerical genomic structure of descendants of early evolving organisms has been proven beyond doubt. The 16S rDNA is the most widely analyzed molecule in phylogenetic studies because of its alternating change of degrees of sequence conservatism that allows the recognition of most distant relationships and moderate and close relationships. The availability of a database comprising more than 10,000 sequences of prokaryotic strains makes it easy to either unambiguously affiliate a new isolate to one of the 4000 species or to postulate the finding of a new taxon in case no highly similar match is obtained with the isolate and a recognized species. Analysis of rDNA is the gold standard for analyzing phylogenetic relatedness; Schleifer and Ludwig (1989) provided excellent evidence that the rDNA data are closely matched by results of comparative analyses of other conservative molecules responsible for central cell functions. Second, the phylogenetic branching patterns provide a scenario in which clusters of related organisms also share a high portion of epigenetic properties. The resolution power of the 16S rDNA is restricted, however, because neither very early

nor very recent evolutionary events are well resolved. Branching patterns are influenced not only by the treeing algorithms applied for inferring phylogenetic relationships but also by the size of the database. A phylogenetic tree is a dynamic construct which will to some extend change its topology with any new sequence added. Consequently, it is difficult for the scientist to judge whether the phylogenetic tree generated is a true reflection of the evolutionary history of the analyzed molecule. If the tree is generated for the purpose of making conclusions on the fate of genomic and epigenetic characteristics of the organisms concerned, conclusions should be underpinned by the analysis of additional evolutionary markers.

Increasingly, genes with a resolution higher than 16S rDNA are sequenced for determining intrageneric and intraspecies relatedness (e.g., genes *hsp65* and *hsp70* coding for heat shock proteins and *gyraseB*). Traditionally, and still required in the polyphasic approach to classification, DNA–DNA hybridization studies are performed at the level of highly related species and at the interspecies level, but the laborious experimental burden and lack of a cumulative database exclude wide application.

Diversity Extends to the Strain Level

In the past, selected physiological reactions and computer-assisted techniques were applied to elucidate intraspecific relationships. This goal can also be achieved by generating patterns from DNA, RNA, and proteins that represent a strain-specific fingerprint or a bar code-type profile. Spontaneous mutations occur at an average rate of about 10^{-7} per gene and generation in prokaryotes but vary to a great extent in genetic loci. The extent to which strains of a species differ from each other also depends on the relative evolutionary time during which strains evolved from each other. Because prokaryotic species are man-made constructs established to facilitate taxonomy, strains included in a species may vary significantly from each other in phylogenetic depth and hence in the extent to which macromolecules diverged. The complexity of the patterns obtained depends on the size, length, and degree of conservatism of the macromolecule of choice and the tools used to cleave, amplify, hybridize, and separate these markers. Separation of DNA fragments by pulsed field electrophoresis, probing of defined genes with labeled probes, visualization of bands via computerized, laser-analyzed densitometer scanning, and the use of fully automated, reproducible techniques have improved the resolution and the monitoring part of the analyses significantly. Pattern identification is rapid, discriminating, and applicable to any species for which DNA, RNA, and whole cell proteins can be isolated. Providing the potential for discrimination, isolates and reference strains that exhibit a high degree of pattern similarity can be considered related. These techniques complement traditional typing methods used mainly in the clinical environment, such as serotyping, biotyping, and phage typing. The decision of whether a strain with a unique pattern actually belongs to a described species or should be described as a new species requires more quantitative methods at the genomic level that allow one to measure the degree of relatedness.

Phylogenetic Diversity and Distribution of Phenotypic Traits

The phylogenetic and phenotypic separation of the domain Archaea from the domains Bacteria and Eucarya (**Figure 1**) is the most exciting result since the introduction of comparative sequence analysis by Woese and Fox in the mid-1970s. The presence of two prokaryotic domains in which members are defined by clearly different genomic and phenotypic properties has changed fundamentally the hypothesis on the dichotomy of life forms and revolutionized ideas about the evolution of the eukaryotic cell. Domains were introduced by Woese, Kandler, and Wheelis to denote that these primary lines of descent constitute higher entities than the traditional eukaryotic kingdoms. The most significant epigenetic differences among members of Archaea and Bacteria are the compositions of the cell wall and fatty acid and the modification pattern of tRNA.

Figure 1 schematically depicts the tripartition of the domain Archaea that guided the description of three kingdoms, the Euryarchaeota, the Crenarchaeota, and the Korarchaeota, for some uncultured organisms. No kingdoms have been described for the rich phylogenetic structure of the domain Bacteria, outlined in **Figure 2**, because of the significantly large number of organisms and lineages involved, which are not always well separated from each other. The order at which these lineages evolved is of low statistical significance but the phylogenetic composition of organisms within these lineages can be recovered by analyses of other genes, such as those coding for 23 rRNA, 5S rRNA, ribosomal proteins, ATPase, elongation factors, and heat shock protein HSP70. In some lineages certain characteristics are indeed of phylogenetic significance, such as morphology and/or ultrastructural features (*Thermotogales, Planctomycetales, Verrucomicrobiales, Spirochaetales, and Myxobacteriales*), chemotaxonomic properties such as cell wall composition (*Thermotogales/Deinococcus*, clostridia, and *Actinobacteria*) or lack thereof (*Planctomycetales*), and physiology, i.e., the composition of the photosynthetic apparatus (*Chloroflexus Chlorobiales*, and cyanobacteria). Most lineages, however, have a wide variation of morphological, chemical, ultrastructural, and biochemical diversity, some traits of which may have been acquired in the course of their evolution by horizontal gene transfer, whereas others may have evolved as a response to occupying new environmental niches (e.g., autotrophic and chemolithotrophic forms). Traits formerly believed to be of monophyletic origin—and hence of taxonomic value—lost their significance in classification when their polyphyletic origin was demonstrated or when these properties were found to be of little genomic stability. In general, this is true for morphology, spore formation, the relationship to oxygen, the presence of a photosynthetic apparatus, gliding motility, and many other characters. In modern classification, these properties are no longer used as the sole basis for the description of higher taxa, such as families, orders, and classes; today, the rational for a higher taxon is primarily the distinct phylogenetic grouping of its members, whereas their phenotypic description may be rather broad.

A good example of the inability of phenotypic properties to serve as phylogenetic markers is provided by members of the class *Proteobacteria*, which includes the majority of Gram-negative bacteria. This class is highly diverse with respect to

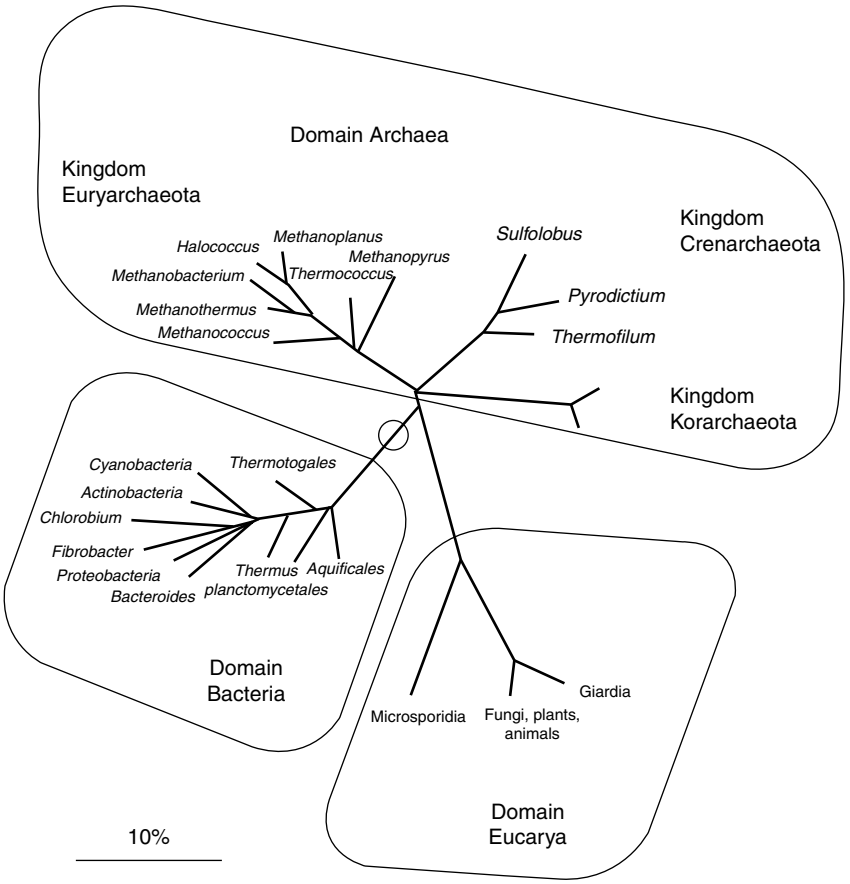


Figure 1 Schematic illustration of the main lines of 16S rDNA sequences showing the tripartition of the domains Bacteria, Archaea, and Eucarya. The circle indicates the approximate position of the root of the tree, making the domains Archaea and Eucarya phylogenetic neighbors. The scale bar indicates 10 estimated changes per nucleotide position.

physiological and morphological properties. Morphological diversity ranges from simple spherical forms to the highly complex fruiting bodies of myxobacteria. Physiologies include chemolithoautotrophy, photosynthesis, fermentation, anaerobic respiration, and nitrogen fixation. To reliably affiliate a new isolate to a described genus by these properties is improbable; the chance of doing so is increased by the presence of unique biochemical properties, such as nitrate or ammonium oxidation (*Nitrobacter* in the α subclass and *Nitrosomonas* and relatives predominantly in the β subclass, respectively) and sulfate reduction (*Desulfovibrio* and relatives; δ subclass). Sulfur and sulfate oxidation as well as anaerobic photosynthesis are poor phylogenetic markers because members of *Thiobacillus* are found in the α, β, and γ subclasses and the sulfurless photosynthetic organisms occupy different sublines of descent within the α and β subclasses.

As a consequence of the recognized discrepancies between the phylogeny of prokaryotic taxa and their previous taxonomic treatment, the recent phylogeny-oriented classification system of genera and higher taxa was developed to primarily match the 16S rDNA data. Newer textbooks have adopted the modern approach, in which, by providing the phylogenetic framework, the origin and evolution of certain phenotypic traits may be better understood than expressed by traditional superficial lumping.

The Diversity of Symbiotic Prokaryotes

Endosymbiotic associations, recognized more than a century ago, initially concentrated mainly on the elucidation of the origin of chloroplasts. Later, the importance of eukaryote–prokaryote relationships was recognized for associations between plants and nitrogen-fixing bacteria, e.g., legumes and *Rhizobium* species, monocots and *Azospirillum* species, *Casuarina* sp. and *Frankia* sp., ferns and *Anabena* sp., and the prokaryotic origin of plant mitochondria. However, most endosymbiotic relationships between microorganisms and their eukaryotic hosts were mainly descriptive and the taxonomic affiliation of the vast majority of uncultured microorganisms remained virtually unknown. In contrast, microbial partners participating in nonobligate symbiotic relationships were identified long before the molecular era; there are numerous associations for the ectosymbiotic microbiota of the rumen, intestine, gut, and skin.

Identification of Symbionts

Most endosymbiotic bacteria are defined as "as yet uncultured" organisms, many of which do not exist as a pure culture within the host tissue. Hence, genes coding for rRNA

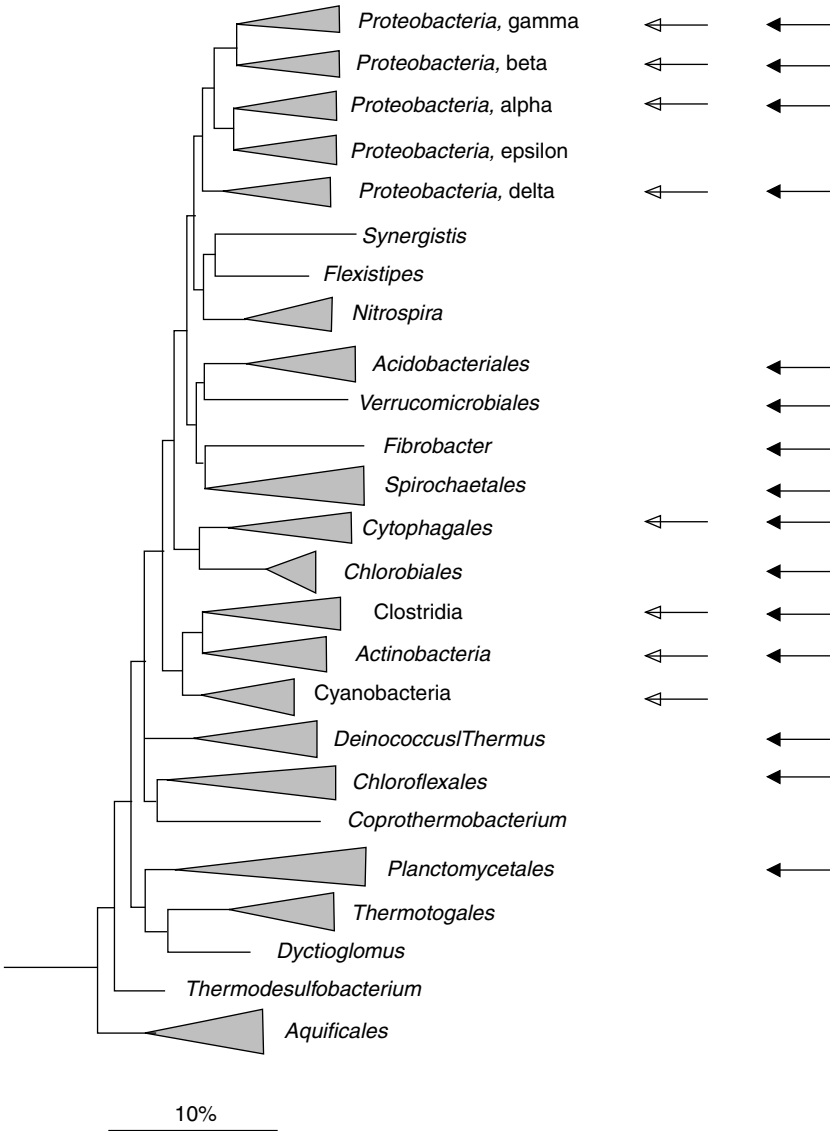


Figure 2 Relative order of the main lines of descent within the domain Bacteria based on 16S rDNA sequences. The horizontal portion of triangles indicates the phylogenetic depth of that lineage, whereas the size of triangles is not indicative of species numbers. Lineages from which symbionts evolved are indicated by open arrowheads. Lineages which encompass sequences of as yet uncultured free-living bacteria are indicated by solid arrowheads. The scale bar indicates 10 estimated changes per nucleotide position.

cannot be isolated selectively but must be identified within clone libraries consisting of PCR-amplified rDNA using prokaryote-specific PCR primers and total DNA extracted from the plant, animal, or even the prokaryotic cell. If the host contains a single symbiotic partner only, the clone library will consist exclusively of the one unique rDNA insert; if the association is more complex, the clone library will contain phylogenetically different inserts. Authentification of the symbiont and verification of the location of the putative endosymbiont within the host tissue is performed by fluorescence *in situ* hybridization techniques as elegantly developed by Stahl, Amann, and coworkers. In most cases the 16S rDNA sequence will provide sufficient unique nucleotide stretches to allow generation of symbiont-specific oligonucleotide probes.

Not until 1982 was *Prochloron didemni*, the endosymbiont of the ascidian *Lissoclinum patellum*, identified by molecular techniques. Ten years later, a wealth of information was available on the molecular phylogeny of symbionts and endosymbionts from a broad spectrum of eukaryotic hosts, ranging from protozoa to vertebrates. The availability of a large database, consisting of thousands of 16S rDNA sequences of free-living prokaryotic species, facilitates the search for the phylogenetic affiliation of the more than 500 16S rDNA sequences available for host-associated bacteria. Fundamental questions about the identity of the prokaryotic partner, the evolution of symbiotic relationships, and the mechanisms of symbiont transmission can now be addressed.

The Evolutionary Origin of Symbionts

Most endosymbionts investigated to date originate from ancestors within the domain Bacteria, in which they are found in a few main lines of descent (open arrowheads in **Figure 2**). The ability to thrive in certain anaerobic protozoa of the genera Metopus, Plagiopyla, and Trimyema appears to be widespread among methanogenic Archaea (kingdom Euryarchaeota). No endosymbionts have been described to share a common ancestry with those prokaryotes which define the most deeply branching lineages, such as the Crenarchaeota, domain Archaea, and *Aquificales*, *Thermotogales*, and other branches comprising thermophilic and phototrophic organisms of the domain Bacteria. Because symbionts and nonsymbionts share more than 80% 16S rDNA sequence similarity it can be concluded that the invasion of the eukaryotic host by prokaryotic cells must have occurred less than 2 Gy ago. **Figure 3** shows recent evolutionary events by plotting geological time against 16S rDNA similarity values determined for the origin of organisms defined by key physiological types (i.e., oxygen-generating photosynthesis by cyanobacteria, fermentative metabolism in facultative anaerobic bacteria, and origin of respiration chain in aerobic bacteria). Thus, as derived from 16S rDNA similarity values, the origin of endosymbionts correlates with the origin of the eukaryotic cell and endosymbiosis has occurred repeatedly in (perhaps all) eukaryotic lineage.

In order to study the history of symbiotic associations, the phylogenetic trees of hosts and their symbionts should be compared. Few data are available, with the most convincing study being that on the endosymbionts of aphids. The topology of the symbiont *Buchnera aphidicola* tree is completely concordant with host phylogeny based on morphology. The fossil and biogeographic time points for the aphid phylogeny have been used by Wilson and Bauman to calibrate the 16S rDNA of the closely related endosymbionts ($>8\%$ similarity). The value of 1% fixed substitutions per 25–50 million years determined for the symbionts of aphids is similar to the value of 1% per 50 million years determined on the basis of a broader range of nonobligate symbiotic relationships (e.g., Rhizobium/legumes, Photobacterium/fish, and enterobacteria/mammals) and to the value of 1% per 60 million years for the past 500 million years (**Figure 3**).

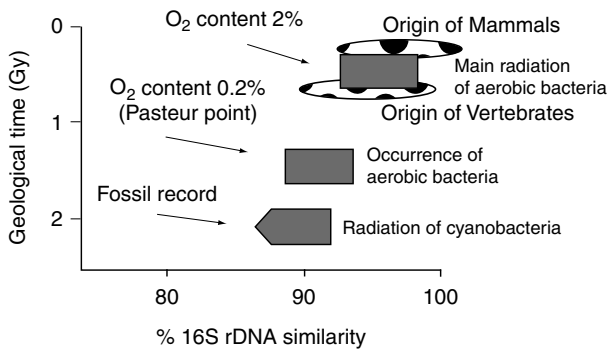


Figure 3 Correlation plot of geological time (as deduced from the fossil record and oxygen content) and 16S rDNA similarity values. Only the past 2 billion years are shown. The origin of aerobic bacteria and symbiotic bacteria correlates with the increased content of oxygen in the atmosphere and the eukaryotic cell, respectively.

The host advantage of the association has been unraveled in a few cases, such as the removal of hydrogen produced from hydrogenosomes of ciliates by archaeal methanogenic endosymbionts, provision of nutritional carbon to the host bivalves by sulfur-oxidizing gill symbionts, or essential amino acids to aphids by their endosymbionts. Application of the PCR techniques has allowed the elucidation of the transmission route of symbionts in ovaries, testis, and gill tissue of tropical lunicid bivalves and deep sea bivalves.

Phylogenetic Affiliation of Endosymbionts

The majority of endosymbionts cluster phylogenetically most closely with Gram-negative free-living bacteria. It can therefore be deduced that Gram-negative bacteria are the most successful candidates for forming symbiosis, including obligate endosymbiotic associations. A few nonproteobacterial symbioses have been described, such as those found between wood-eating cockroaches and termites and spirochetes and Gram-positive host-associated bacteria, such as the microparasite *Pasteuria penetrans* (in root-knot nematodes), *Epulopiscium* (a fish symbiont), and *Frankia* (nitrogen fixing on *Casuarina* and relatives). The association between *Sphingobacterium comitans*, a member of the *Bacteroides/Flavobacterium* phylum, and the myxobacterium *Chondromyces crocatus* is an example of a prokaryote–prokaryote symbiosis. Endosymbionts of Archaea are members of the kingdom Euryarchaeota and they are, in contrast to those of Bacteria, organisms containing a Gram-positive cell wall. These anaerobic symbionts, originating from ancestors of the families *Methanosarcinaceae*, *Methanomicrobiaceae*, and *Methanocorpusculaceae*, have been identified in several genera of termites and protozoa. The basis of the symbiotic interaction appears to be hydrogen transfer from host to endosymbiont, which can use the gas for methanogenesis.

Symbionts of the Proteobacteria

Proteobacteria embrace organisms known to have close associations with eukaryotic hosts as pathogens, parthenogenesis bacteria, incompatibility bacteria, symbionts, and organelles (such as the mitochondria of plants which evolved from α proteobacterial ancestors). Endosymbionts and nonobligatory associates of the same host may belong to different phylogenetic groups, indicating that the same host is susceptible to more than a single invasion process and that not all symbiotic relationships result in obligate endosymbiosis.

Symbionts of the α Proteobacteria

The α subclass of Proteobacteria contains a wide spectrum of organisms that are closely associated with eukaryotic cells. Prime examples are members of the nitrogen-fixing genera *Rhizobium*, *Sinorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, and *Azorhizobium*. Pathogens include *Afipia*, *Brucella*, *Bartonella*, *Rickettsia*, *Ehrlichia*, *Orienta*, and *Anaplasma*. A highly related cluster of host-associated organisms, including symbionts of insects, cytoplasmatically inherited bacteria such as the parthenogenesis bacteria (PB), and the cytoplasmatic

incompatibility bacteria (CIB) is related to *Wolbachia pipientis*. Other members of this group are the PB and CIB of *Culex* and *Drosophila*. Because of the degree of 16S rDNA similarity among cultured strains, these symbionts must be considered members of the same species. These relationships demonstrate that the common ancestor of *Wolbachia* and relatives invaded a broad spectrum of insect hosts which, as shown by the high values of up to 99% 16S rDNA similarity, must have occurred recently in evolution.

Symbionts of the β Proteobacteria

Members of the β subclass of Proteobacteria encompass a wide range of mainly pathogenic plant- and animal-associated bacteria, such as members of *Burkholderia, Azoarcus, Ralstonia, Bordetella, Kingella, Eikenella*, and *Neisseria*. Also included in this subclass is the kinetoplast of *Crithidia*. The diversity of endosymbionts, however, is rare and restricted to the endosymbionts of the mealy bugs, which are moderately related members of the genera *Ralstonia* and *Burkholderia*.

Symbionts of the γ Proteobacteria

By far the higher number of endosymbionts of insects and vertebrates are members of the γ subclass of Proteobacteria. In addition, this taxon contains a wide spectrum of animal and human pathogens and nonobligate symbionts such as members of *Enterobacteriaceae, Legionellaceae, Pasteurellaceae, Vibrionaceae, Pseudomonas* (sensu stricto), and *Acinetobacter*. Many endosymbionts cluster according to the phylogenetic rank of their hosts, which may be indicative of coevolution events: (i) The primary endosymbionts of giant ants, aphids, tse-tse, and the sweet potato white fly form a phylogenetically coherent cluster within the radiation of enterobacteria; (ii) the symbionts of fish with light organs, such as the deep-sea anglerfish and the flashlight fish, cluster with different members of *Vibrio*; (iii) two moderately related subclusters consist of the gill symbionts of bivalves. These organisms are remotely related to methylotrophic bacteria. The latter symbionts are sulfuroxidizing organisms which provide their hosts with nutritional carbon. Nitrogen-fixing symbiosis is not restricted to plants but the gland of an invertebrate shipworm contains large numbers of γ proteobacteria which possess the ability to digest cellulose and fix nitrogen.

The Diversity of the Uncultured Free-Living Organisms

For more than a century, assessment of prokaryotic species has been evaluated by the culturing approach. The number of different growth media is unknown, but all aimed to recover the largest possible diversity of organisms. However, it is not the enrichment and isolation but rather the lack of cheap and reliable molecular identification methods that still slow microbiologists in their attempts to classify strains to the species level and to describe new prokaryotic species. The number of novel strains that has been eliminated during the isolation process, or which were only included in a biased search for specific properties before they disappeared in nonpublic resource collections, cannot be counted. It is unknown how many strains have been investigated in parallel. Although this problem may one day be overcome by more facilitated species definitions and the availability of a global network of biological information, microbiologists are currently confronted with the problem that there is a remarkable difference in the number and morphology of organisms in natural samples with enrichment cultures and isolated colonies. Staley and Konopka introduced the phrase "great plate anomaly" to indicate that only a small fraction of prokaryotic species observed under the microscope will grow under artificial laboratory conditions.

The Vast Majority of Prokaryotic Species Have Not Been Cultured

Pace and colleagues first suggested that rRNA sequences could be used to characterize natural communities without the need to culture. The delay in publishing the first studies by the research groups of Giovannoni and Ward was largely due to the need to develop robust and simple technologies whereby 16S rRNA sequences could be recovered from complex mixtures of environmental nucleic acids and then individually sequenced. Since then, there have been numerous 16S rRNA- and 16S rDNA-based studies in which sequences were analyzed to explore microbial diversity in different environments. Most of these studies differ in details of methodologies, such as isolation of nucleic acids, PCR amplification conditions, and source of cloning vectors and ligation enzymes, but in one aspect the outcome of all of these studies was similar: The vast majority of the more than 1200 environmental partial 16S rDNA sequences, deposited in public databases, were not identical with and often not even similar to the homologous sequences of described species accessible in the extensive databases of cultured bacteria (**Figure 2**, solid arrowheads). Also, the sequences were rarely identical to sequences obtained from strains that were isolated in parallel to the molecular work from the same environment (**Figure 4**). As judged from the low degree of sequence similarity, one could even conclude that many sequences are indicative of the presence of higher taxa. This finding reinforced the previously mentioned idea that the vast majority of species have not yet been isolated.

Additional molecular techniques have been developed for understanding the composition of microbial communities that reach beyond the mere assessment of phylogenetic relationships of clones and strains. Among others, methods comprise (i) the application of gradient electrophoresis of PCR-amplified 16S rDNA sequences to facilitate recognition of changes of populations in time and space, (ii) the development of biological probes to detect the presence of genes encoding metabolic enzymes or to identify bacterial species directly in an environmental sample, (iii) the development of biosensors to determined microprofiles of inorganic compounds, (iv) flow cytometry and cell sorting to enumerate and separate groups of organisms according to size and taxon specificity, (v) subtractive hybridization to facilitate comparative analysis of environmental samples, and (vi) extension of the database to include genes other than 16S rDNA (e.g., *nif* genes).

The environments discussed in the following sections were selected because they provide the largest database of

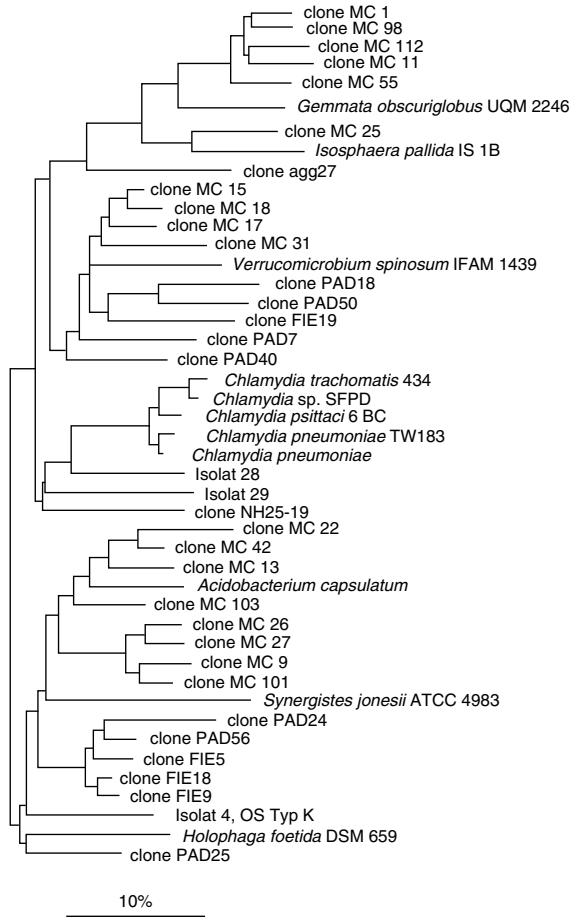


Figure 4 Distribution of 16S rDNA clone sequences, representing uncultured bacteria, within the radiation of 16S rDNA sequences of some cultured bacteria. Origin of clone sequences: MC, Australian soil; PAD, Japanese paddy field; FIE, Japanese soy bean field; NH, north Pacific Ocean. The scale bar indicates 10 estimated changes per nucleotide position.

phylogenetic information on uncultured organisms. Many other environments have been investigated, but the overall picture of prokaryotic diversity is less focused; these include, fresh water, paddy fields, marsh oil, marine plants, bioremediation sites, bioleach reactors, or the multistructured associations between prokaryotes and eukaryotic cells.

The Marine Environment

The Uncultured Archaea

The use of archaeal-specific 16S rDNA PCR primers and subsequent analysis of clone libraries from DNA of oxygenated coastal surface waters and oligotrophic open-ocean samples by the groups of DeLong and Fuhrman revealed the widespread occurrence of two types of archaeal diversity. Analyses of different marine sites, including surface and subsurface waters of the Pacific, Atlantic, Antarctic coastal waters, offshore slope regions, and a deep-sea marine holothurian, have subsequently confirmed the affiliation of archaeal sequences to these two groups. Clone sequences of the first group contained 16S rDNA genes of Crenarchaeota organisms found at depths >100 to 4800 m, whereas others were found at different marine sites at more shallow depths. Organisms from which sequences originated constitute a significant component of approximately 5–14% of the marine picoplancton assemblages and were novel and closely related to the archaean "Crenarchaeum symbiosum," a symbiont of the sponge *Axinella mexicana*. At a lower level of relatedness these sequences form a separate branch at the basis of the kingdom *Crenarchaeota*, where they show some distant relatedness to a group of sequences isolated from sediment organisms of the Obsidian Pool, Yellowstone.

A few sequences have been described to belong to the second sequence group originating from Euryarchaeota and are distantly related to the terrestrial species *Thermoplasma acidophilum*. Except for a holothurian archaeal sequence, they have been retrieved from the same environment from which the crenarchaeotal sequences were obtained, but mostly at lower depth (0–100 m). The high sequence similarity of members of this group from geographically separate sites contrasts the finding by Munson and colleagues, who demonstrated the high degree of phylogenetic diversity between archaeal sequences retrieved from 16S rDNA clone libraries generated from material sampled in marsh sediment samples and adjacent vegetative marshland in the United Kingdom. In this study, clone sequences formed about 15 different phylogenetic groups, each of which was highly to distantly related to cultivated Euryarchaeota species.

The origin of marine archaea is unclear because no representative has been cultured. One may argue that these prokaryotes originate from dormant stages of released commensales or symbionts of marine invertebrates. This view, however, is contradictory to the high cell number of living organisms of this group—up to 14% of the total community. Disturbed deep-sea sediments, which have not been investigated, may be another possible source for these organisms. In contrast, euryarchaeal clone sequences retrieved from surface water material may be of coastal and even terrestrial origin, considering that there is a specific relationship of these sequences to some of the DNA retrieved from coastal salt marsh.

The Uncultured Bacteria

Phylogenetic analysis of rDNA from phytoplankton of the Sargasso Sea by Giovannoni, Britschgi, and coworkers revealed that the majority of the mostly eubacterial sequences were novel to systematists. Some clones represented oxygenic phototrophs and could be assigned to the *Synechococcus* group of the *Cyanobacteriales*. Groups of related sequences of mainly the α and γ Proteobacteria were identified and defined as SAR (Sargasso) groups, showing distant relatedness to cultured bacteria of *Shewanella, Vibrio*, and *Oceanospirillum*. Analysis of the phytoplankton from the north Pacific near Hawaii by Schmidt and colleagues led to the unexpected result that the population was very similar to that of the Atlantic Ocean. Some sequences were similar to those of common cultured marine organisms (e.g., *Vibrio, Pseudoalteromonas, and Chromatium*), whereas others represented novel lineages which were distantly related to the *Fibrobacter* group and *Chlorobium*.

Following studies extended the range of sampling sites in the Atlantic and Pacific Oceans, the Antarctic Sea, and the Mediterranean Sea. Basically, the findings of the first studies were confirmed in that many of the new sequences were highly

related, although not identical, to those defined earlier, irrespective of the location. The ecological role of the hitherto undescribed organisms remains unresolved, although their general physiological capacity is probably similar to that of described species because α Proteobacteria encompass mainly lithotrophic and oligocarbophilic organisms, whereas γ Proteobacteria and cytophages exhibit strong hydrolyzing and degrading capacities. The few sequences of Gram-positive bacteria are probably of terrestrial origin because they are mainly found in coastal regions and in sediments which must be regarded as a deposit mainly for endospore-forming organisms.

Hot Spring Environments

Research in hot spring environments has been stimulated by the discovery of *Thermus aquaticus* as a species of high biotechnological value. The community of mats especially provided an excellent comparison of microbial composition as assessed by selective culture techniques introduced by Brock and Castenholz. Results of the sequence investigations by Ward, Weller, and colleagues clearly demonstrated that even such a rather closed system inhabits a phylogenetically diverse community. Consistent with the findings in the marine environment, none of the recovered sequences closely resembled sequences from cultured taxa isolated from similar environments. Sequences of cultured organisms believed to constitute a major component of the mat community, such as *Synechoccous lividus* or *Chloroflexus aurantiacus*, were not recovered. Analysis of prokaryotic DNA from hot springs located in Yellowstone National Park by Pace and coworkers revealed an unexpectedly large number of distinct bacterial sequences, indicative of novel main lines of descent. Archaeal sequences were determined to be specifically related to sequences from the Crenarchaeota. A few sequences showed high similarities with 16S rRNA sequences of cultivated Archaea, (e.g., *Desulfurococcus mobilis*, *Pyrobaculum islandicum*, and *Thermofilum pendens*) but were not identical to any. The archaeal sequences from the hot spring environment were not closely related to the novel archaeal sequences retrieved from the marine environment, which appear to possess a similar position intermediate to the Crenarchaeota. Results of these studies indicated that the domain Archaea possesses a third line of descent, the kingdom Korarchaeota, suggesting that not only the phylogenetic but also the physiological diversity of the Archaea are significantly larger than reflected by the few cultured representatives.

Soil Environment

In contrast to the early extensive work on different marine sites and hot springs, determination of microbial biodiversity in soil was delayed until appropriate methods were developed that circumvented methodological difficulties such as the isolation of PCR-able DNA from humic acid-containing soil and semiquantitative cell recovery. Despite these shortcomings studies by Liesack and Triplett clearly demonstrated that soil samples from the Southern and Northern Hemispheres contain a rich and varied bacterial flora, including nonthermophilic archaeal members from the kingdom Crenarchaeota.

The first soil sample investigated by sequencing and probing 16S rRNA genes was located in a subtropical, moderately acidophilic, and forested environment in Queensland, Australia. Unexpectedly, only a few sequences were obtained from commonly isolated soil organisms, such as *Streptomyces*, although members of this genus were cultured in large numbers from the same soil sample. The reasons for the low representation of Gram-positive bacteria in libraries generated with universal 16S rDNA primers are not known, but this may be explained by the cell wall structure of resting and dormant cells which fail to disintegrate under mild enzymatic lysis. By using actinomycete-specific primers, Embley and colleagues were indeed able to demonstrate the presence of a rich diversity of these taxa in soil. Alternatively, Gram-positive bacteria may be a minor (numerically) component of the soil flora, with selective isolation exaggerating their numbers. Some Australian soil sequences were closely related to those of nitrogen-fixing species of the α Proteobacteria, but the majority of clone sequences represented novel groups that were only remotely related to known taxa, e.g., *Planctomycetales*, *Actinobacteria*, *Verrucomicrobium*, *Acidiphilium*, and *Thiobacillus acidiphilius*.

Several actinobacterial clone sequences, retrieved from rDNA of different soil types and different geographical locations, formed two clusters grouping remotely with members of *Rubrobacter*, *Acidicrobium*, and *Atopobium*. In addition to their presence in Australian forested soil, their occurrence was verified in a hot spring (Australia), geothermally heated soil (New Zealand), paddy fields and soybean fields (Japan), cultivated soil (Mexico), peat bog and garden soil (Germany), grassland soil (The Netherlands), and forest soil (Finland) (**Figure 4**). These sequences formed fractions of 1–23% of the respective clone libraries and one of these organisms constituted about 6% of the metabolically active part of a Dutch grassland soil community as shown by ribosomal RNA analysis. It can be deduced that these uncultured organisms are distributed worldwide and play a physiologically important role in the soil ecology.

Conclusions

Despite tremendous progress in the elucidation of prokaryotic diversity, many pitfalls have been identified which influence the composition of sequences in a clone library and hence these data can be used neither to quantify nor to qualify the composition of communities. Any estimation of the relation of cultured and as yet noncultured organisms is nothing more than a guess and not supported by scientific data. Most environmental analyses have revealed a heterogeneous mixture of deep and shallow branching lineages, very few of which have shown close relationships to cultured taxa. Many lineages are very closely related to each other, some of this diversity is due to microheterogeneity at the level of rrn operons within a single cell, whereas others may represent true strain diversity.

Although the contribution of rDNA and rRNA to microbial ecology must be considered significant, one molecule alone cannot nearly cover all facets of microbial ecology. Not only must the function of an ecosystem be deduced from analyses of genes expressed through rRNA, mRNA, and proteins, and

not only should the network of broad physiological interactions be verified by *in vitro* reconstitution of isolates but also all these strategies must include data on thorough physical and chemical analysis of the natural sample. Because ecological interactions are performed by strains and not by species, the 16S rDNA is not the appropriate tool to unravel the diversity of this high phylogenetic level. Strain diversity may thus be one or more magnitudes higher than mirrored by the analysis of such an evolutionary conserved gene. Ecological niches must be defined and the difficulty in doing so is increased with the complexity of the sample. Soil samples, for example, are homogenized in that during the isolation of DNA many individual microniches with their individual populations are destroyed. In order to understand ecological interactions, population sizes must be known, strain richness and strain abundance must be assessed, the biochemical diversity of strains must be recognized by applying functional probes for the detection of the expression of specific genes, and physical probes must be applied for the assessment of the chemical and physical conditions of the environment. Modern environmental studies must maintain the isolation component because biotechnological exploitation of cosmid libraries is as desirable as the increase in knowledge of the evolution, phylogeny, and ecology of pure cultures.

See also: Archaea, Origin of. Bacterial Genetics. Eukaryotes, Origin of. Microbial Biodiversity, Measurement of. Thermophiles, Origin of

References

Embley TM and Finlay BJ (1994) The use of small subunit rRNA sequences to unravel the relationships between anaerobic ciliates and their methanogen endosymbionts. *Microbiology (Reading)* 140: 225–235.

Ochman H and Wilson AC (1987) Evolution in bacteria: Evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol* 26: 74–86.

Pace NA (1997) A molecular view of microbial diversity and the biosphere. *Science* 276: 734–740.

Schleifer KH and Ludwig W (1989) Phylogenetic relationships among bacteria. In: Fernholm B, Bremer K, and Jörnwall H (eds.) *The Hierarchy of Life*, pp. 103–117. Amsterdam: Elsevier.

Service RF (1997) Microbiologists explore life's rich, hidden kingdoms. *Science* 275: 1740–1742.

Von Wintzingerode F, Göbel U, and Stackebrandt E (1997) Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. *FEMS Microbiol, Rev* 21: 213–229.

Ward DM, Ferris MJ, Nold SC, and Bateson MM (1998) A natural view of microbial diversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev.* 62: 1353–1370.

Woese CR (1987) Bacterial evolution. *Microbiol. Rev.* 51: 221–271.

Zuckerkandl E and Pauling L (1965) Molecules as documents of evolutionary history. *J. Theor. Biol* 8: 357–366.

Bacterial Genetics

Stanley Maloy, San Diego State University, San Diego, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Michael Travisano, volume 1, pp 339–350, © 2001, Elsevier Inc.

Glossary

Alleles Alternative forms of a gene. For example, the mutants *putA601* and *putA736* each have a different mutation in the *putA* gene.

Ames test A genetic test for the identification of carcinogens based on their mutagenic activity initially developed by Bruce Ames. The test relies on the ability of a chemical or physical agent to promote the reversion of different classes of mutations that cause histidine auxotrophy.

Auxotroph A mutant that will only grow when a particular nutritional requirement (e.g., amino acid, nucleotide, or vitamin) is provided.

Genotype A specific description of the genetic constitution of an organism. The genotype is defined by the allelic form of each gene in an organism, but for simplicity usually only differences from the wild type are described.

Mutagen A chemical or physical agent that increases the frequency of mutation, usually by directly damaging the DNA.

Mutation Any heritable alteration in the base sequence of the genetic material.

Phenotype The appearance or other observable characteristics of an organism. The phenotype expressed by an organism depends on the particular forms of its genes (e.g., its wild-type or mutant alleles) and the environmental conditions.

Suppression The restoration (or partial restoration) of a wild-type phenotype by a second mutation. There are many different potential mechanisms of suppression.

Transposon A genetic element which, in addition to encoding the proteins required for its own transposition, confers one or more new observable phenotypes (often resistance to one or more specific drugs) on the host cell.

Mutations and Mutagenesis

A powerful feature of bacterial genetics is the ability to examine very large populations of cells (typically $>10^{10}$) for extremely rare types of mutants. Mutations can arise in a variety of ways, including rare errors in DNA synthesis or DNA repair. The probability that a spontaneous mutation will affect a particular base pair varies from about 10^{-7} to 10^{-11} per generation. Thus, in a population of bacteria, about 1 in 10^9 cells may have a mutation in a particular gene. The frequency of mutations can be increased by certain chemical or physical agents called mutagens. Mutagens may act by increasing the frequency of errors during DNA synthesis or repair or by directly altering the DNA. Direct exposure of bacteria to a mutagen may increase the frequency of mutations in a population of cells from 10^3- to 10^6-fold. Thorough genetic analysis requires many types of mutations, but each particular method of mutagenesis yields different subsets of mutations.

Effects of Mutations

Bacterial mutants are typically described by comparison to a standard, well-characterized, reference strain called the 'wild-type' strain. Bacterial mutants have often lost some growth property (e.g., failure to utilize a particular carbon or nitrogen source or failure to grow without a particular nutrient), or acquisition of some new growth property (e.g., ability to grow in the presence of some toxic substance). Genes can be divided into two categories based on the phenotypes of the corresponding mutants: nonessential gene products are only required under specific growth conditions, whereas essential gene products are required under all conditions. The genes of lactose catabolism are nonessential because they are only required for growth on medium with lactose as the sole carbon source. In contrast, the genes encoding RNA polymerase are essential because they are required for growth on all media. Null mutations in a nonessential gene will prevent growth on a medium that requires that gene product but such mutants will still grow on other media. In contrast, null mutations in an essential gene are lethal. Consequently, such mutants cannot be recovered from haploid bacteria. Nevertheless, it is possible to isolate more subtle mutations in essential genes. For example, it is possible to isolate mutations that alter a subunit of RNA polymerase that make the organism resistant to the antibiotic rifampicin. It is also possible to isolate mutations where some phenotype is observed under certain 'nonpermissive' conditions but not under other 'permissive' conditions (**Table 1**).

Because not all mutations have an observable effect, it is important to distinguish the genotype from the resulting phenotype. In bacterial genetic nomenclature a three-letter mnemonic refers to a pathway or discrete cluster of physiologically connected systems. A fourth, capitalized letter represents a particular gene of that set. The genotype is written in lower case letters and italicized (e.g., *purB*), with a plus superscript indicating the wild-type genotype. (The *purB* gene encodes one of the enzymes required for purine biosynthesis.)

The phenotype is indicated by the same mnemonic but the first letter is upper case and it not italicized (e.g., PurB), with a plus superscript indicating the functional phenotype and a minus indicating a mutant phenotype. The genotype of a cell is usually inferred from its phenotype but may also be determined indirectly by recombination experiments or directly by DNA sequencing

Isolation and Characterization of Bacterial Mutants

Genetic analysis begins with the isolation of mutants that affect some property of the bacteria. Since mutations are very rare, mutants must be isolated from large populations of wild-type cells. Thus, some tactic is needed to find the rare mutants within a vast excess of parental bacteria. It is possible to identify mutations by physical methods such as DNA sequencing, but because mutations are so rare this approach is not usually practical. Instead, mutants are usually identified on the basis of an observable effect on the physiology of the cell. Detection of mutants requires a genetic selection or screen (**Table 2**).

A selection is an experimental arrangement that allows specific mutants, but not the parental cells, to grow. Genetic selections are very powerful because they allow the direct isolation of rare mutations from a very large population of cells. Some useful selections include resistance to antibiotics, resistance to phage, or the ability to grow on a medium where the parental cells cannot grow. For example, selection for streptomycin-resistant (Str^R) mutants simply requires exposing a large number of bacteria to medium containing streptomycin-sensitive bacteria are killed by the antibiotic and resistant mutants grow.

The ability to grow bacteria on solid media (i.e., agar plates) under conditions where each cell forms a single colony allows one to carry out screens that distinguish a particular mutant from other bacteria in the population. If the mutation is relatively common and no direct selection for the mutant phenotype is available, it is possible to screen for mutants on media where both the mutant and parental cells grow, but where the mutant has a readily scorable phenotype. For example, mutants of *Escherichia coli* unable to degrade lactose can be identified on indicator media, such as MacConkey-lactose color indicator plates. MacConkey-lactose plates contain a carbon source that can be used by Lac^+ cells, and peptides that can be used as a carbon source by both Lac^+ and Lac^- cells. Both the Lac^- mutant and the Lac^+ parental cells can therefore grow on MacConkey-lactose medium. However, MacConkey-lactose medium also contains a pH indicator that is colorless at high pH but red when the pH decreases owing to fermentation of the lactose, so Lac^+ bacteria form red colonies and Lac^- bacteria form white colonies. Hence, it is possible visually to screen for Lac^- mutants by looking for rare white colonies on MacConkey-lactose plates.

Screens for other types of mutants are not this simple. For example, when mutations disrupt a biosynthetic pathway, the bacteria cannot grow unless the endproduct of the pathway is provided in the medium (permissive medium). It is not possible to isolate such 'auxotrophic mutants' by directly screening for growth on media lacking the endproduct because the desired mutants will not grow (nonpermissive medium). However, auxotrophs are readily identified by replica plating. The bacteria are grown on permissive medium at a density of several hundred colonies per plate. Each of these 'master plates' is then replicated onto two other plates, containing either the permissive medium or the nonpermissive medium. The bacteria are transferred to an identical position on each of the replica plates. Once colonies are identified that grow only on the permissive medium, the mutants can be isolated from the master plate (**Figure 1**).

The difference between a selection and a screen has important practical consequences. Compare the selection for a Str^R mutant with a screen for a His^- mutant. Isolation of a Str^R colony is a direct selection because as many as 10^{10} cells can be

Table 1 Some types of conditional mutations used in bacterial genetics

Conditional mutation	*Permissive conditions*	*Nonpermissive conditions*
Temperature sensitive (Ts)	30 °C	42 °C
Cold sensitive (Cs)	42 °C	30 °C
Osmoremedial	High osmotic strength	Low osmotic strength
Suppressor sensitive	Host with suppressor mutation	Host lacking suppressor mutation

Table 2 Isolation of bacterial mutants

Approach	*Features*	*Sensitivity*	*Examples*
Selection	Condition where only mutant cells can grow	$>10^{-9}$	Resistance to antibiotics or toxic substrate analogs; resistance to phage
Screen	Condition where both mutants and parental cells can grow, but the mutants have a phenotype that is distinguishable from the parent	10^{-2}–10^{-3}	Indicator medium to identify mutants unable to ferment a carbon source; replica plating to identify auxotrophs
Enrichment	Condition where survival of the mutant is favored over the parental cells; usually employs an antibiotic that selectively kills growing cells or a condition that kills cells able to incorporate a particular substrate; a genetic screen is needed to identify the resulting mutants	10^{-2} per cycle	Penicillin enrichment; D-cycloserine enrichment; radioisotope suicide

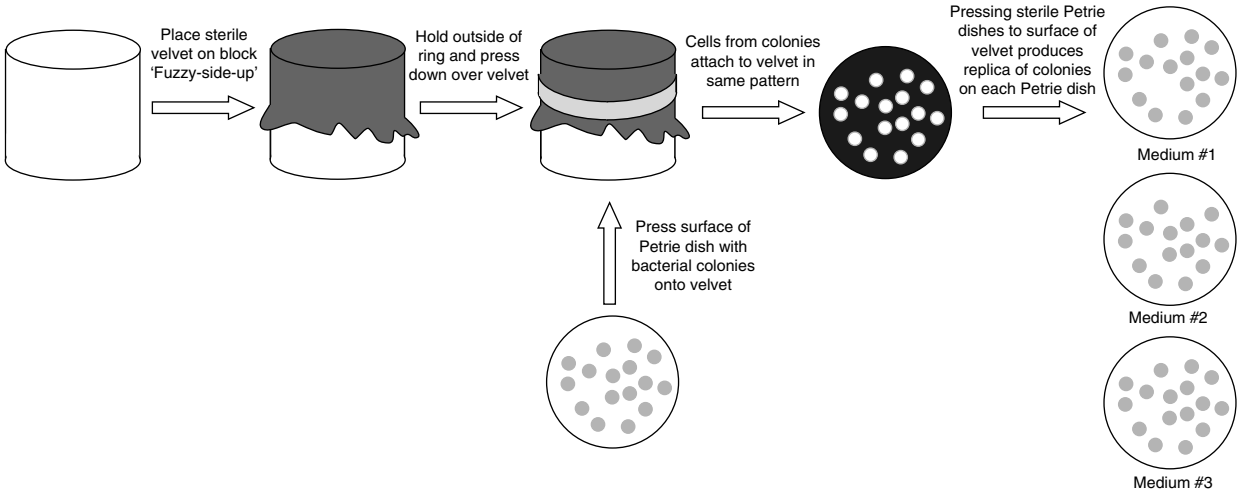


Figure 1 Replica plating. This simple technique makes it possible to test many individual mutants for their ability to grow on a variety of different types of media. After pressing the sterile plates to the velvet replica of an original ('master') plate, the newly inoculated plates are incubated to allow the colonies to grow.

spread on a single plate containing streptomycin and only Str^R mutants will form colonies. Thus, Str^R mutants as rare as 1 in 10^{10} can be isolated easily. In contrast, finding an auxotrophic mutant involves screening through a large number of colonies for one that fails to grow in the absence of the required nutrient. To score the growth behavior of individual colonies, only a few hundred colonies can be examined on any given plate. Thus, if 1 in 10^5 cells in the population were a particular auxotrophic mutant, thousands of plates would be needed to find a single mutant. Treating populations of cells with a mutagen may increase the fraction of mutants to 1 in 10^3, allowing screening for the desired mutant on a reasonable number of replica plates. Nevertheless, random mutagenesis may not increase the mutant fraction sufficiently to screen easily for the desired mutants, and mutagenesis may lead to other, undesirable mutations as well.

Isolating rare mutants is generally achieved by using some sort of enrichment, a method that favors the growth of the desired mutants relative to nonmutant bacteria. Penicillin enrichment is a classical example of this approach. This antibiotic disrupts the synthesis of the bacterial cell wall. Nongrowing bacteria do not engage in cell wall synthesis, so penicillin only kills actively growing bacteria. The differential survival of growing versus nongrowing bacteria can be used to enrich for a desired mutant. For example, penicillin enrichment can be used to isolate a rare auxotrophic mutant from a population of bacteria. If a mixture of wild-type and auxotrophic bacteria are suspended in the nonpermissive medium containing penicillin, the auxotrophic mutants do not grow and thus will survive, whereas 99% of the wild-type bacteria will grow and be killed by the penicillin. When the surviving bacteria are washed free of the penicillin, and resuspended in permissive medium, both the mutant and wild-type bacteria will grow, but the ratio of mutant to wild-type bacteria will be enriched 100-fold. Repeating this procedure multiple times eventually increases the proportion of mutants in the population (**Figure 2**).

Genetic Exchange

Exchange of DNA between bacteria plays an important role in evolution. Gene transfer in nature can result in the acquisition of antibiotic resistance and new virulence traits. Gene transfer is also a useful tool in the laboratory, allowing genetic mapping and complementation tests, and the construction of bacterial strains with multiple mutations. The three most common methods of gene transfer between bacteria are transformation, conjugation, and transduction (**Table 3**). A suitable selectable marker is required to identify recipients that have inherited the desired region of donor DNA.

Transformation

The uptake of naked DNA is called transformation. Some species of bacteria are naturally transformable. At some stage during growth they express gene products that facilitate the uptake of exogenous DNA, a condition called 'competence.' The physiological conditions required to induce competence differs for different species of bacteria. However, most natural transformation seems to involve the degradation of one strand of the exogenous DNA during the transfer of the other strand of DNA into the cell. Stable inheritance of the donor DNA requires either that it be replicated (e.g., some plasmids and phage) or that it integrate into the recipient chromosome via homologous recombination.

Many types of bacteria are not naturally transformable but can be induced to take up DNA by treatment with specific chemicals or by electric shock, processes that are mechanistically different from natural transformation. A common method of chemical transformation in *E. coli* uses hypotonic Ca^{2+} shock. To prepare competent bacteria by this method, an early-exponential phase culture of cells is suspended in a cold hypotonic $CaCl_2$ solution. When DNA is added to these bacteria it forms a calcium–DNA complex that adsorbs to the

cell surface. The bacteria are then briefly warmed (heat shocked), which allows the DNA to enter the cell. In electroporation, cells are exposed to an electric field, so that a voltage potential develops across the membrane, transiently forming small pores that allow entry of exogenous DNA. In contrast to natural transformation and chemical transformation, which only seem to work in certain bacteria, a wide range of bacteria can be transformed by electroporation.

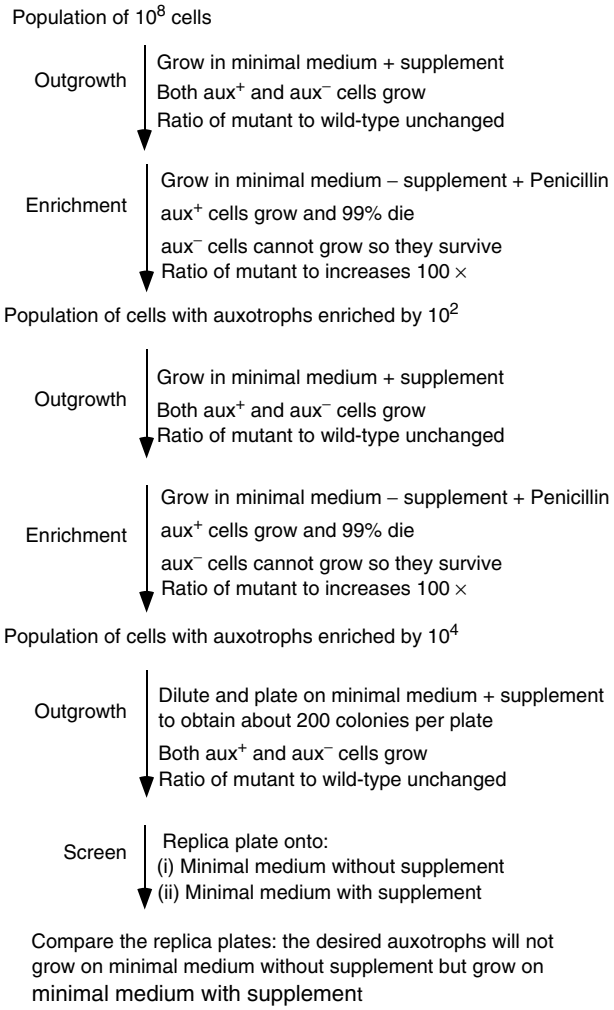


Figure 2 An example of the use of penicillin enrichment to isolate auxotrophic (aux^-) mutants from a population of prototrophic (aux^+) bacterial cells. Each cycle results in about 100 × enrichment of the desired mutants.

Conjugation

DNA can also be transferred between bacteria by a process called conjugation. Conjugation occurs via cell–cell contact and the formation of a mating channel. This structure is formed by specific proteins that form a pore between the juxtaposed membranes through which single-stranded DNA and some associated proteins are transferred into the recipient cell. Conjugation requires four events: (1) contact between donor and recipient cells and formation of a mating appendage; (2) nicking of the donor DNA at specific sites; (3) translocation of the nicked strand of donor DNA to the mating bridge and into the recipient cell; and (4) replication of the transferred DNA in the recipient cell. The proteins required for conjugation are usually encoded on specific plasmids or transposons. These conjugal plasmids or transposons can transfer themselves or, if integrated into another region of the genome, any adjacent genomic sequence. Thus, conjugation can result in the transfer of very large DNA fragments – even chromosome length DNA fragments. Furthermore, some conjugal plasmids are quite promiscuous. In addition to transferring DNA into a wide variety of bacteria, some conjugal plasmids can transfer DNA into Archaea and eukaryotes as well (**Figure 3**).

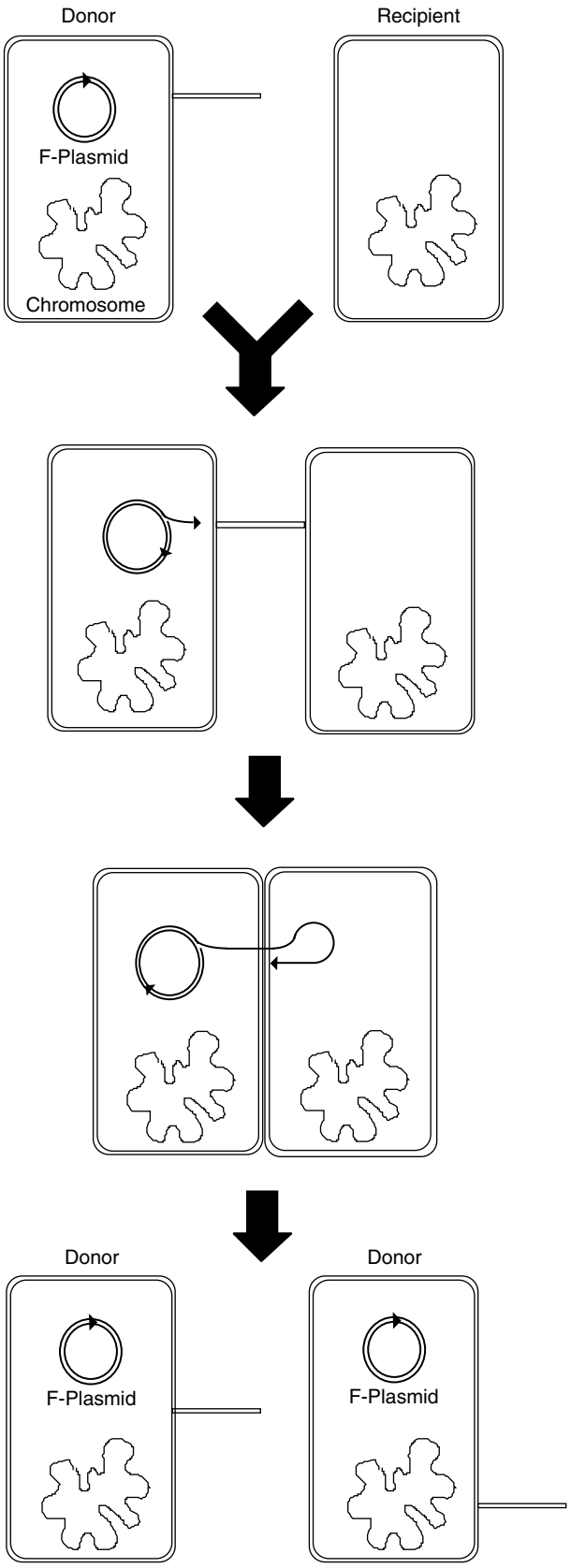
Transduction

Transduction is the transfer of bacterial DNA from one cell to another by a phage particle. Phage particles that contain bacterial DNA are called transducing particles. There are two types of transduction: generalized and specialized.

Generalized transduction requires an error during the packaging phase of phage maturation, such that random, phage-size fragments of chromosomal DNA get inserted into the phage head in place of phage DNA. Thus, in a lysate of a generalized transducing phagea some particles contain DNA obtained from the host bacterium rather than phage DNA. The bacterial DNA fragment can be derived from any part of the bacterial chromosome. When these phage particles adsorb to a recipient, the double-stranded chromosomal DNA is injected. Stable inheritance of the donor DNA requires either

Table 3 Common mechanisms of natural gene transfer in bacteria

Transfer mechanism	*Vehicle*	*Properties*
Transformation	–	Uptake of naked DNA from environment; recipient cell must be competent
Conjugation	Conjugal plasmid or transposon	Requires cell-to-cell contact; very long fragments of DNA can be transferred to recipient cell
Generalized transduction	Phage	Phage-length fragments of chromosomal DNA transferred to recipient cell; phage head only carries host DNA
Specialized transduction	Lysogenic phage	Short fragments of chromosomal DNA transferred to recipient cell; phage head carries host DNA covalently attached to phage DNA



that it be replicated (e.g., following transfer of plasmids) or that it integrate into the recipient chromosome via homologous recombination. Generalized transducing particles arise during lytic growth of both virulent and lysogenic phages (**Figure 4**).

Specialized transduction results from the aberrant excision of an integrated lysogenic phage. In contrast to generalized transduction, the phage head packages contain a contiguous molecule having both host and phage DNA. Only regions of the host DNA that flank the prophage are packaged.

By using genetic tricks to force lysogenic phages to integrate at many sites in the genome, it is possible to isolate specialized transducing phage from many different regions of a bacterial genome. Stable inheritance of specialized transducing fragments can occur either by phage-specific processes or by homologous recombination mediated by bacterial recombinases.

Genetic Analysis of Bacterial Mutants

Three of the most important tools for the analysis of bacterial mutations are suppressor analysis, genetic recombination, and complementation. Suppressor analysis facilitates the dissection of the roles of gene products and how they interact with other gene products. Recombination allows the construction of new combinations of genes and the elucidation of the positions of genes on a chromosome relative to other genes. Complementation can reveal how many genes are responsible for a particular phenotype and can distinguish regulatory genes from regulatory sites.

Reversion and Suppressor Analysis

Mutant organisms can regain their wild-type characteristics by means of a mutation that restores function, by a process called reversion. Reversion can either be owing to 'true reversion,' a back mutation that restores the original genotype, or 'suppression,' a second mutation that produced a change that partially or fully compensated for the first mutation. The ability to analyze very large populations of bacteria facilitates the isolation of rare classes of such revertants.

The reversion frequency is a useful criterion for classifying mutants. The reversion frequency can distinguish a single point mutant from a double mutant or a deletion mutant: single point mutants (which can be repaired by a single back mutation) revert at a much higher frequency than double mutants (which require two changes and thus revert very rarely) or deletion mutants (which cannot directly revert). The effect of

Figure 3 Conjugation is a process of DNA transfer via direct contact between a donor cell that carries a conjugal plasmid such as the F-plasmid, and a recipient cell that lacks the conjugal plasmid. A pilus protruding from the surface of the donor cell recognizes the recipient cell and promotes close contact between the two cells, forming a bridge through which DNA is transferred. On cell–cell contact the conjugal plasmid is nicked at a specific site called oriT (indicated by the arrowhead), and the DNA is transferred from this site into the recipient cell. After conjugation is complete, both cells have a complete copy of the conjugal plasmid.

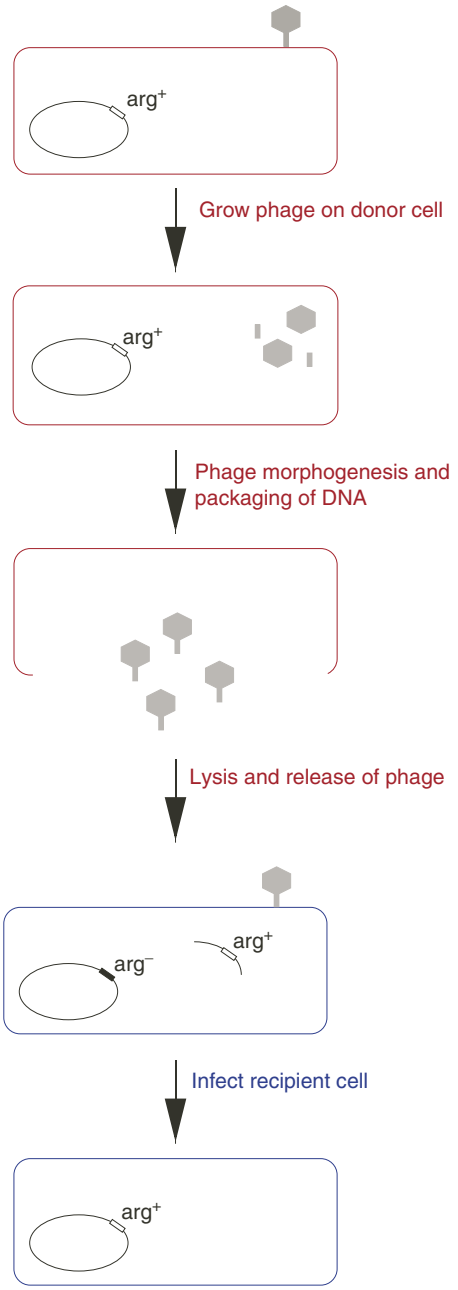


Figure 4 Transduction is the process of packaging of a portion of the genome from a donor cell into a virus particle (e.g., a bacteriophage) and the subsequent delivery of the genetic material to a recipient cell. Once in the recipient cell the fragment of donor DNA can recombine with the genome of the recipient yielding a new phenotype (in the case shown, via conversion of the recipient from an arginine auxotroph into an arginine prototroph). The altered recipient cell is called a transductant.

mutagens on the reversion frequency can also be used to distinguish different types of mutations – reversion of frameshift mutations is stimulated by frameshift mutagens, and so on. For most purposes this approach has been superceded by direct DNA sequence analysis; however, the reversion of known bacterial mutations remains an important assay to detect potential human carcinogens (for example, using the Ames test).

Second site mutations that result in suppression can occur within the same gene as the original mutation (intragenic suppression), or within a different gene (intergenic suppression). Intragenic suppressors can provide information about the role of particular amino acids in a protein, protein folding, and interactions between different amino acids within a protein. Intergenic suppressors can occur in a variety of ways, and each provides different insights into the structure or function of the gene products involved. Suppressor mutations often result from 'gain-of-function' mutations that are dominant over the wild-type gene. Thus, characterization of suppressors requires genetic recombination to construct strains with different combinations of mutant alleles, and complementation analysis to determine dominance.

Homologous Recombination

Genetic recombination is the physical breakage, exchange, and rejoining of two DNA molecules. Homologous or general recombination can be mediated by several different pathways in bacteria. Each of these pathways requires the RecA protein to align the DNA molecules between regions of substantial DNA sequence identity. The DNA molecules are broken between random but matching nucleotides, and then the DNA fragments are exchanged and rejoined to form two new combinations of genes. For example, recombination between two DNA molecules with the genotypes a^+b and ab^+ can yield two recombinant DNA molecules with the genotypes a^+b^+ and ab. Owing to the efficiency of gene transfer and the ability to work with large populations of bacteria, recombination analysis can be sufficiently sensitive to detect recombination events between adjacent base pairs.

Recombination can be used not only to construct strains with different genotypes, but can also reveal the relative map location of genes. The probability of recombination between any two adjacent base pairs is very low, but the probability of recombination between base pairs within a homologous DNA sequence is essentially random. Hence, the physical distance between two genes located on the same DNA molecule determines the frequency of recombination between the genes: the probability of recombination is less if the genes are close to each other than if the genes are farther apart (**Figure 5**).

Genetic mapping exploits the recombination frequency between genes to measure the relative distance between genes. In bacterial genetics, the probability that recombination did not occur between genes is usually determined. If recombination does not occur between two genes, the genes will be coinherited. For two genetic markers on the same DNA molecule, the closer two genetic markers are to each other, the more often they will be coinherited. The frequency that two genes are coinherited is defined as their linkage. Determining the linkage of two genetic markers is called a two-factor cross. It is also possible to determine the relative location of genetic mutations by using three genetic markers (three-factor crosses) or by genetic crosses involving a set of defined deletion mutations (deletion mapping). Although it is possible to determine the relative location of genes by hybridization or DNA sequencing, genetic mapping often provides a simple and

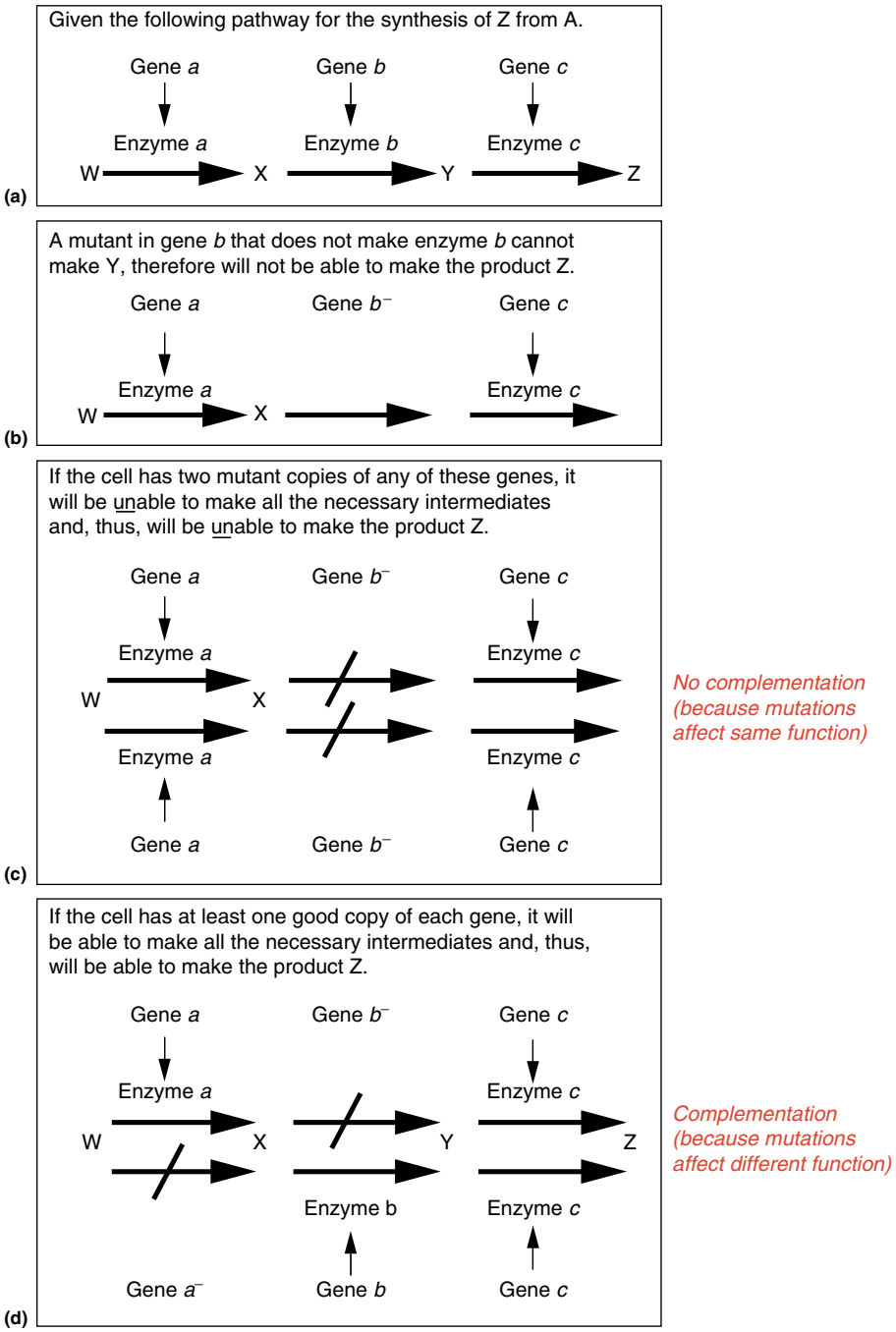


Figure 5 Complementation analysis provides a genetic test for particular functions encoded by different genes. (a) Assume that in a cell substrate W is converted in three enzymatic steps to the endproduct Z, and that each of these enzymatic steps is encoded by a different genes (*a, b,* and *c*). (b) If any one of these genes is defective in a haploid bacterial cell, then that cell will be unable to make the endproduct. (c) If two copies of these genes are placed in a single bacterial cell, and if both copies are defective for the same gene, then the cell will be unable to make the endproduct. In this case, we say that the second copy did not complement the original mutation. (d) If the two copies of the genes are defective for different genes, then at least one functional gene is present for each step in the pathway, and the endproduct will be made. If the second copy compensates for the defective function, then we say that the mutation was complemented.

inexpensive way to determine rapidly the location of mutations in bacteria.

Recombination also provides an invaluable tool for constructing strains with multiple mutations. If a mutation involves a directly selectable phenotype, it can be transferred with ease from a donor to an appropriate recipient strain. If the mutation cannot be directly selected, linkage to an adjacent, selectable marker can be used to move the mutation into a recipient strain. For example, an auxotrophic mutation may be transferred by selecting for coinheritance of a nearby gene.

To determine the genetic or biochemical effects of a mutation, it is necessary to compare a mutant with a strain that only differs by a single mutation. If several mutations are present, it is not obvious which of the mutations is responsible for an observed phenotypic change. Two organisms that differ by only a single mutation are said to be isogenic or to have the same genetic background. The most common way to ensure that two strains are isogenic is to transfer a small region of DNA carrying the mutation into the parental strain by recombination.

Complementation Analysis

A particular phenotype frequently reflects the activity of many genes. To understand any genetic system it is essential to know the number of genes and regulatory elements involved. Since multiple genes that affect the same function may map very close to each other, it is not possible to determine if two mutations are in the same or different genes simply from the recombination frequency. For a variety of reasons, it is also often difficult to prove that two mutations affect the same gene product by DNA sequence analysis. However, this question can be addressed by genetic complementation.

Bacteria are normally haploid, but complementation requires the maintenance of two copies of a particular gene in the same cell. In bacteria this can be done by constructing a partial diploid or 'merodiploid' that carries two copies of the relevant genes. Partial diploids may provide one copy of the gene on the chromosome and a second copy of the gene on a plasmid. For example, the partial diploid *lacZ*/*lacZ*$^+$ has both a copy of the *lacZ* gene and second copy with the *lacZ*$^+$ gene. If the functional copy of these genes encodes a protein that can diffuse through the cell to perform its function, and the second copy has a loss-of-function mutation, the functional copy of the gene will be dominant over the mutant copy. Thus, even though the *lacZ* gene on the chromosome cannot produce beta-galactosidase, the plasmid-bone *lacZ*$^+$ gene makes beta-galactosidase so the cell is phenotypically Lac$^+$. The mutant is complemented by the wild-type gene, indicating that the mutation is recessive. In contrast, the partial diploid *lacZ*/*lacZ* cannot make beta-galactosidase so the cell is phenotypically Lac$^-$. If two recessive mutants fail to complement, they affect the same gene (**Figure 6**).

Complementation analysis in bacteria usually follows the grouping of mutations by genetic mapping experiments. This reduces the number of partial diploids that must be constructed because genes that map far from each other are clearly different. In addition, complementation analysis should involve partial diploids with an equal number of copies of each gene. If one of the genes is present in excess (for example, with genes cloned on multicopy plasmids) artifacts can occur which may be very misleading.

Portable Regions of Homology

Although most genes have a specific, defined location in a bacterial genome, the genome is by no means static. Recombination between repeats of homologous DNA sequences can result in the duplication, deletion, or inversion of the intervening DNA. It is possible to select for insertion of such homologous sequences at strategic positions in the genome. Such 'portable regions of homology' can be generated by transposable elements, which are mobile segments of DNA that move to new locations at low frequency. Alternatively, having specific DNA fragments on a circular DNA molecule (e.g., a plasmid or phage) can permit the targeted insertion of DNA into a particular site on the chromosome. These two types of insertions have a wide variety of genetic uses. A few examples include: construction of insertion mutants with complete loss of function; construction of deletion mutations with defined endpoints; construction of duplication mutations for complementation studies; integration of other genetic elements at particular sites in the genome by homologous recombination; transfer of mutations by selection for an associated genetic marker (e.g., antibiotic resistance); isolation of linked genetic markers with a selectable phenotype; and construction of operon or gene fusions.

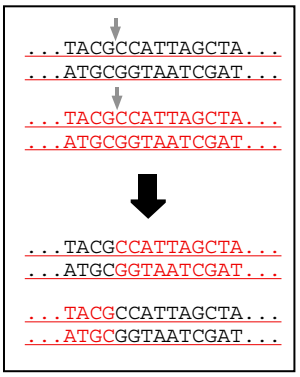


Figure 6 Recombination is the breaking and joining of two homologous DNA molecules. The exchange occurs between two homologous base pairs, generating a join point between the two parental DNA molecules. Because the recombination can occur between any of the base pairs in homologous sequences, the further apart two genes are, the more likely there will be a recombination event between them.

Transposable Elements

Transposable elements in bacteria include insertion sequences and transposons. Insertion sequences are short elements (typically less than 5 kb) that only encode functions required for their own transposition. Transposons are typically longer (>5 kb) and encode other gene products (e.g., antibiotic resistance) in addition to the functions required for transposition. The frequency of transposition of these elements is typically low, although the frequency varies over a wide range (10^{-7}–10^{-2} per generation). Transposition requires both cutting the DNA at each end of the transposon and the target DNA site, joining the ends of the transposon and target DNA molecules, and DNA replication. Transposition can occur by a replicative mechanism (requiring replication of the entire transposon) or nonreplicative mechanism (requiring only replication of short fragments at the end of the insertion site). Both mechanisms are independent of the homologous recombination machinery of the host.

Transposable elements play an important role in bacterial evolution, including the transfer of antibiotic resistance genes between bacteria and promoting chromosome rearrangements. In addition, transposable elements are useful tools in bacterial genetics because they provide selectable markers and portable regions of homology that can be used to facilitate genetic recombination.

Integration of Circular DNA Molecules

In addition to transposons, it is possible to construct small repeats of homologous DNA by integration of a circular DNA molecule with a cloned fragment of chromosomal DNA. This approach is useful for construction of defined duplications for complementation analysis and for construction of insertion mutations on the chromosome. Recombination between the homologous sequences of the resulting duplication can result in allele exchange, moving a mutation from a cloned sequence onto the chromosome DNA.

Summary

A hallmark of bacterial genetics is the ability to analyze very large populations of cells to identify rare genetic events. Although a wide variety of genetic tricks have been developed for specific purposes in particular bacteria, bacterial genetics relies on a relatively small core of tools for dissection of the structure and function of genes. The essential tools include the isolation of mutations, the ability to transfer genes between bacterial strains, the ability to isolate recombinants, and the ability to do complementation tests. These tools have been finely honed for a select group of model bacteria, including *E. coli*, *Salmonella enterica*, and *Bacillus subtilis*. The concepts developed for these model bacteria are readily applicable to other bacteria as well, although the experimental details typically require adaptation for each particular species of bacteria.

References

Hughes K and Maloy S (eds.) (2007) *Advanced Bacterial Genetics. Methods Enzymology*, vol. 421. San Diego, CA: Academic Press.

Maloy S, Cronan J, and Freifelder D (1995) Microbial Genetics. 2nd edn. Boston, MA: Jones and Bartlett Publishers.

Maloy S, Hughes K, and Casadesus J (2011) *The Lure of Bacterial Genetics*. Washington, DC: ASM Press.

Maloy S, Stewart V, and Taylor R (1996) *Genetic Analysis of Pathogenic Bacteria*. Plainview, New York: Cold Spring Harbor Laboratory Press.

Miller JH (1992) *A Short Course in Bacterial Genetics*. Plainview, New York: Cold Spring Harbor Laboratory Press.

Beetles

Henry A Hespenheide, University of California, Los Angeles, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 351–358, © 2001, Elsevier Inc.

Glossary

Fauna The animal species living in a defined area.
Leaf beetles Members of the family Chrysomelidae.
Phytophagous Feeding on plants; also, herbivorous.
Weevils Members of the beetle family Curculionidae.

Introduction

In terms of numbers of species, arthropods are the largest phylum of living organisms, insects are the largest group within the arthropods, and beetles are the largest order among the insects. Beetles presently comprise about a quarter of all described, extant organisms. This numerical dominance led the British biologist J. B. Haldane to answer the question of what one might learn about the mind of the Creator from a study of his creation with the famous reply: "An inordinate fondness for beetles."

In 1982, T. L. Erwin used samples of beetles to estimate that the number of species of living organisms may be 30 million or more. The ensuing controversy (especially Stork, 1988) rekindled interest in the more general question of the magnitude of global biodiversity (May, 1988) and how to measure it. These discussions did not question the usual assumption that beetles represent the largest part of global biodiversity, although mites and nematodes may in fact be more diverse than beetles. More recently and narrowly, a proposal to "explain" the diversity of beetles has itself engendered controversy. Both questions are reviewed here.

What Beetles Are?

The order Coleoptera is usually divided into four suborders and 160–170 families depending on the author. The suborders Archostemata and Myxophaga each consist of four families and fewer than 100 extant species. The Adephaga includes about a dozen families, with most of the species in the large family of ground beetles Carabidae. All other beetles—150 families and 90% of the 340,000 species—are classified as Polyphaga. Somewhat more than two-thirds of the species belong to the eight largest families discussed here, each with more than 10,000 species, but there are 30 other families that have between 1000 and 10,000 species **Table 1**). The most thorough review of the biology of members of the order is found in Crowson (1981).

Because they are well sclerotized, there is a good fossil record for beetles. Beetle-like insects first appeared in the Lower Permian, modern families are known from the Cretaceous, and extant genera date from the Oligocene and Miocene. Ecologically, beetles are the most diverse order of insects, which is a primary factor in their great species richness. There are large families of beetles that live in fresh-water habitats, that are predaceous and parasitic, and that feed on fungi and detritus, but the largest proportion of described species feed in some way on plants. Among these phytophagous beetles, the largest number of species are wood-borers; a much smaller number feed directly on leaves, and some feed on seeds.

Diversity—How Many Beetles Are There?

Gaston (1991) has reviewed current estimates of the number of insect species described, favoring those of J. F. Lawrence for beetles. Lawrence (1982) put the numbers of beetles at 340,500 species divided into about 25,000 genera and placed in 169 families. His estimate of the number of species falls in the middle of estimates (290,000–370,000) cited by Gaston, and roughly equal to the sum of the number of species in the three next largest insect orders (Diptera, 98,000–120,000 species; Lepidoptera, 112,000–165,000 species; Hymenoptera, 100,000–130,000 species). Estimates of the total number of living species, described and undescribed, are difficult to relate to the number of described species for two major reasons. First, most taxonomic research has been carried out in the north temperate regions of the world on organisms that live there, whereas most insect species are tropical. Second, different groups of insects have received differential amounts of study from scientists, both within the beetles and between beetles and the other insect orders. This article briefly reviews what is known about the major beetle families, considers explanations for the success of the order, reviews a few ecological and biogeographic patterns in beetle diversity, and then discusses the problems in estimating what we don't yet know about beetle diversity.

As stated earlier, eight families of beetles currently account for more than half of all beetle species. However, unequal efforts in the collection and study of different groups very likely mean that the relative sizes of these families will change when the less well studied taxa become better known. Some of the better-studied groups have already begun to be used to test hypotheses of mechanisms that might promote or maintain beetle diversity; for example, whether speciation in plant-feeding beetles parallels speciation in their hosts. The eight major families are considered in order of their traditional phylogenetic placement.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00007-1

Table 1 Diets and numbers of species and genera of the larger families of beetles (more than 1000 species)[a]

Taxon	*Genera*	*Species*	*Diet*[b]							
			A	*P*	*D*	*F*	*Ph*	*W*	*S*	*L*
Adephaga										
Caraboidea										
Carabidae	**1500**	**30,000**		P						
Dytiscidae	120	3000	A							
Polyphaga										
Hydrophiloidea										
Hydrophilidae	125	2000	A							
Histeroidea										
Histeridae	200	3000		P						
Staphylinoidea										
Liodidae	250	2000				F				
Scydmaenidae	75	2000			D					
Staphylinidae	**1500**	**30,000**		P	D	F				
Pselaphidae	650	5000				F				
Scarabaeoidea										
Lucanidae	100	1200					Ph			
Scarabaeidae	**2000**	**25,000**					Ph			
Buprestoidea										
Buprestidae	**400**	**15,000**						W		
Elateroidea										
Elateridae	400	9000					Ph	W		
Eucnemidae	190	1200						W		
Cantharoidea										
Lycidae	150	3500		P						
Lampyridae	100	2000		P						
Cantharidae	135	5000		P						
Bostrychoidea										
Anobiidae	180	2150						W		
Cleroidea										
Cleridae	150	4000		P						
Melyridae	200	5000		P						
Cucujoidea										
Nitidulidae	160	3000			D					
Cucujidae	75	1200		P						
Erotylidae	30	2500				F				
Endomychidae	120	1300				F				
Coccinellidae	500	4500		P						L
Colydiidae	180	1300		P						
Mordellidae	100	1200					Ph	W		
Oedimeridae	100	1000						W		
Anthicidae	100	3000					Ph			
Meloidae	120	3000		P						
Tenebronidae	**1700**	**18,000**					Ph			
Chrysomeloidea										
Cerambycidae	**4000**	**35,000**						W		
Bruchidae	60	1500							S	
Chrysomelidae	**2500**	**35,000**								L
Curculionoidea										
Anthribidae	325	2600				F				
Attelabidae	100	2100								L
Apionidae	26	2200					Ph			
Brenthidae	325	2300						W		
Curculionidae	**4500**	**50,000**					Ph	W		

[a]Data from Lawrence JF (1982) Coleoptera. In: *Synopsis and Classification of Living Organisms*, 482–553. New York: McGraw-Hill. Families listed in boldface account for more than two-thirds of all beetle species.

[b]A = aquatic, D = detritus, F = fungi, L = leaves, P = phytophagous, S = seeds, W = wood-boring.

Carabidae

The ground beetles (30,000 described species, including the tiger beetles, sometimes placed in the separate family Cicindelidae) are primarily predatory. Although some carabid groups are very rich in tropical regions, others do not show as great an increase in diversity there as do other beetles. Despite the vernacular family name, tropical forms are commonly arboreal. Tiger beetles are well studied taxonomically, have been used to test ecological and evolutionary hypotheses, and have been suggested as indicator species for biodiversity studies (Pearson and Cassola, 1992).

Staphylinidae

The rove beetles (30,000 species) are both numerous in species and diverse ecologically, including predatory, fungus, and detritus feeders. The staphylinids are the least well known family of beetles, with perhaps only 10% of extant species described, and may eventually prove to be the largest family. They are abundant in canopy fogging samples in tropical regions.

Scarabaeidae

The scarabs (25,000 species) are primarily phytophagous, but with one group well known as dung beetles. Because they are popular with collectors, they are among the best known of the beetle groups. The actual number of extant species is unlikely to be more than twice as many as those already described.

Buprestidae

The jewel beetles (15,000 species) include a majority of wood-boring species and several groups of leaf-miners. Although larger species are popular with collectors, the smaller species are primarily subtropical and tropical and very poorly known. The genus *Agrilus* may be the largest genus of living organisms with 2000 described species and many more undescribed—there are nearly 350 undescribed species among more than 600 known from México alone. At La Selva Biological Station in Costa Rica, 24 of 29 *Agrilus* (83%) are undescribed; of 206 Buprestidae, 147 (71%) are leaf-miners and 62% are undescribed. The family should more than double in size when completely known.

Tenebrionidae

The darkling beetles (18,000 species) are phytophagous and have been relatively popular and well studied in temperate and subtropical regions. I have seen no discussion of its current taxonomic status or estimates of undescribed species.

Cerambycidae

The long-horned beetles (35,000 species) are wood-borers, very popular among collectors, and well studied. Of 274 species collected at La Selva, only 23 (8%) are undescribed, although another 28 (10%) are still undetermined. It is possible that the number of species will double when the rich South American fauna is more fully known. This and the following two families constitute the Phytophaga and are the largest beetle families in described species.

Chrysomelidae

The leaf beetles (35,000 species) are by far the largest group of beetles that primarily feed on the leaves of plants. Some groups within the family are relatively well studied, even within tropical areas (Cassidinae), whereas others are not (Alticinae). Members of the family are common in canopy samples and have been shown to speciate in parallel with their plant hosts. The success of beetles overall has been attributed to the success of this and the other two families in the Phytophaga. When the total chrysomelid fauna is known, it will probably double the number of currently known species.

Curculionidae

The weevils (50,000 species) are the largest family of living organisms. Although they are primarily phytophagous and associated with angiosperm plants in a variety of ways, many (and perhaps most) are wood-borers (many Cryptorhynchini, Molytini, and Zygopini) rather than leaf-eaters, in some cases cultivating fungi in their tunnels (Platypodinae and some Scolytinae). The great majority of weevils are endophytic; the few that feed directly on leaves often do so as leaf-miners. Sampling at tropical localities has yielded large numbers of weevils. At Barro Colorado Island in Panamá, three years of light trapping (Wolda *et al.*, 1998) yielded 1240 species of Curculionoidea (excluding Platypodinae and Scolytinae, but including about 200 species from four other families). Sampling at La Selva has collected well over a thousand species of weevils, including 504 in the largest tribe, Zygopini, of which 425 (84%) are undescribed. The number of species is probably two to three times those presently known.

Explanations for the Species Richness of Beetles

Farrell (1998) recently claimed to have verified an earlier proposal that the great species richness of Coleoptera is due to the association of the largest lineages within the Phytophaga (Cerambycidae, Chrysomelidae, Curculionidae) with angiosperm plants. Indeed, beetles feed on angiosperms in a number of ways. The narrowest form of phytophagy is to feed on leaves (foliivory), but very few beetles do that, primarily Chrysomelidae and leaf-miners in several families. As foliivores, the Lepidoptera are much more successful (most of 140,000 species). A few phytophagous beetles feed on seeds (Bruchidae, some Curculionidae), but most are wood-borers (Buprestidae, Cerambycidae, many Curculionidae). Perhaps more than qualities of the host itself, it is the ability of beetles to feed endophagously in wood and other plant tissues that makes them so successful. Anderson (1995) has attributed the success of weevils, the most species-rich lineage, to the evolution of the rostrum as an aid to oviposition in plant tissues. It is more difficult for Diptera and Lepidoptera to feed

endophagously because adults lack chewing mouth-parts with which to exit plant tissues.

Another problem with the Farrell hypothesis is that even if the Phytophaga were removed from the beetles, the remaining taxa are still twice as numerous as any other insect order. If the Staphylinidae are as species rich as has been suggested (300,000 species), that family alone would be more than twice as large as the currently known Phytophaga and about the same size as estimates of total Phytophaga. The success of beetles is almost certainly due to their use of so many major resources, not just their intensive use of one of them. Of the 8 largest beetle families, the Carabidae are primarily predaceous and the Staphylinidae are partly or largely so; 10 of the other 30 families with more than 1000 species are also primarily predaceous. Five of the larger families and some Staphylinidae feed on fungi, which is another major adaptive zone for beetles. Aquatic habitats are used by two large families of beetles and only by Diptera among the other three major orders.

Ecological samples of canopy insects, although dominated by phytophages, include significant portions of species of other feeding types. If only half of all beetle species feed on angiosperm plants, it could be argued that beetles are in fact *less* successful than expected, for the reason that angiosperms constitute a much larger resource in terms of productivity and biomass than the other trophic levels in which beetles feed with equal success. In sum, there is no single simple explanation for the great diversity of beetles. Although association with angiosperms has obviously been a major factor in the diversification of beetles, it has been so in a very complex way that may be due less to the plants as a resource and more to the features of the beetles ("preadaptations") that allowed them to use the plants endophagously.

Taxonomic and Ecological Patterns

Most large taxa have greater numbers of species in tropical regions than in temperate ones, and that pattern is likely to hold for families of beetles. Increase in species numbers at lower latitudes is not usually a matter of uniform increases in all genera within the taxon, but rather of changes in dominance among taxa. For example, in the Buprestidae, wood-boring genera dominate in temperate and subtropical regions both north and south of the equator, but leaf-mining genera dominate near the equator. Among the leaf-miners, the genera *Brachys* and *Taphrocerus* dominate in temperate and subtropical regions, but the genera *Hylaeogena*, *Leiopleura*, and *Pachyschelus* dominate near the equator. In the Chrysomelidae, the subfamily Clytrinae has a peak in species richness in the subtropics, as does the zygopine weevil genus *Cylindrocopturus*.

More complex relationships between species diversity and ecological diversity are suggested by some data. The wood-boring Cerambycidae increase in absolute species richness as one moves towards the tropics, but F. T. Hovore (unpublished data) has found that species richness relative to potential woody plant host species diversity actually decreases at lower latitudes because of a greater increase in the number of woody plant taxa (**Figure 1**). This is not unexpected in that many tropical plant species are rare, and Southwood and others have

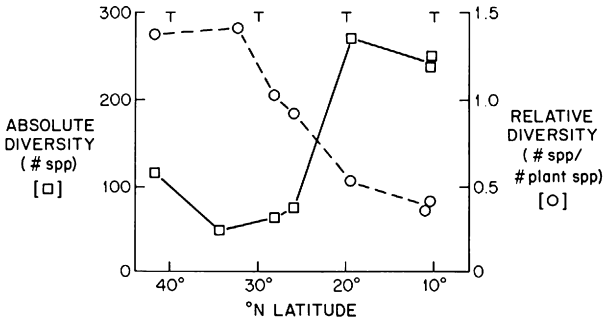


Figure 1 Diversity of the beetle family Cerambycidae at several sites as a function of latitude. Number of species is given on the left axis (squares)—the number increases from the temperate zone to the tropics. Relative diversity is shown on the right axis, expressed as the number of species of beetles compared to the number of species of potential woody plant hosts (circles)—the number of beetle species increases less rapidly than the number of potential hosts. (Data from F. T. Hovore, unpublished.)

shown that rarer plants have fewer host-specific herbivores than do more common plants. In the tropical Cerambycidae, many of the species are associated with a few larger, more abundant host plant genera such as *Inga* and *Ficus*, and host generalists may be common.

Mimicry is more common in tropical beetle species than in temperate members of the same taxon. The proportion of species participating in mimicry complexes has been shown to increase with decreasing latitude in *Agrilus* (Buprestidae) mimicking the ant genus *Zacryptocerus* and in zygopine weevils mimicking a variety of models. On the other hand, mimicry of and by clytrine chrysomelids is most frequent in the subtropical regions of México, where the clytrines are most species rich.

How Much Don't We Know?—Local Faunas

Estimates of how many beetle species exist in various families and the contribution of beetles to global biodiversity should be placed in the context of our extensive, basic ignorance of the ecology and distribution of beetles at local and regional scales. Except perhaps for England, there is no place on earth of any size and ecological diversity where one can make a complete list of all beetle species. In areas where checklists have been attempted, the larval biology and population ecology of many or most species are completely unknown. Lists of described species are incomplete, and a significant number of species are undescribed in all but the most popularly collected taxa. A few small taxa have been surveyed both intensively at a few sites and more or less extensively at scattered sites (Pearson and Cassola, 1992). In general, recent extrapolations to global diversity estimates have been made from samples of local faunas.

Beetles occupy such a wide variety of ecological roles in communities that no single method of sampling is adequate, even when the taxonomic scope of the beetles studied is relatively narrow. In recent years, Malaise trapping and canopy fogging have become popular, but flight intercept traps, sweep

netting, blacklighting, Berlese sampling, pitfall traps, baiting, beating, systematic rearing, and other methods have been used as well. Malaise traps sample active, flying insects (e.g., Mordellidae), and beetles in Malaise samples have been used as indicators of specific habitats. Light trapping (Wolda *et al.*, 1998) is successful in sampling cryptorhynchine and molytine weevils, but is poor for diurnally active zygopines and baridines. Canopy fogging has been widely used in recent years to sample beetle faunas.

Several major projects have attempted to inventory the arthropod faunas of single, diverse localities. The Arthropods of La Selva (ALAS) project in Costa Rica, for example, uses a combination of four standardized sampling techniques (Malaise trapping, canopy fogging, litter sampling, and light trapping), as well as group-specific collecting methods, to survey a variety of arthropods including several beetle groups. A similar variety of sampling techniques were employed in Sulawesi (Hammond, 1993). These and other such projects in tropical regions have been slow in producing more than partial faunal lists because of the massive taxonomic problems associated with rich tropical faunas that have been largely unstudied (see the following sections). The current ecological concept of metapopulations raises questions about the meaning of local faunas by suggesting that local diversity is a dynamic function of regional diversity. The idea that some "sink" populations are maintained by regular immigration from "source" populations elsewhere may explain anecdotes of the sudden and irregular appearance and disappearance of species of leaf-mining Buprestidae at La Selva in Costa Rica.

Estimating Diversity Globally

The problem of rare species has both ecological and taxonomic–evolutionary implications. Rare species are difficult to sample and difficult to interpret when they are sampled. Apparent rarity can be due to several factors, including methodological artifacts, disturbance by "tourists," seasonal phenology, cyclic populations, or true biological rarity. Furthermore, small populations may be characteristic of tropical species. Three years of light trapping by Wolda at Barro Colorado Island, Panamá, yielded over 95,000 weevil specimens, but 28% of the 1239 species were represented by a single specimen. In smaller samples at six other localities in the same study, unique specimens represented 39–51% of the species. Canopy fogging samples show comparably high proportions of rare species. Reliable taxonomic decisions about what is or is not a valid biological species often require series of specimens to interpret variation; adequate series for studies are even more necessary when specimens from several localities are involved. Many described tropical species are known only from unique type specimens. The question of what is a species has been justifiably raised in discussions of biodiversity at regional and global geographic scales.

Erwin (1982) startled the scientific world by calculating that 1200 beetle species (955 counted, 206 weevils estimated, then "rounded up" to 1200) sampled from 19 individual trees of *Luehea seemannii* implied that there were 30 million species of insects worldwide. His extrapolation depended on three important quantities: his estimate of the proportion of host-specific beetles, the fraction found exclusively in the canopy, and the assumption that all tree species have the same number of herbivores. The last of these is certainly not true—rare tree species have fewer host-specific insects (see earlier discussion).

Empirical studies that sample both the canopy and lower levels in the forest have not shown a large and/or distinct canopy fauna. Gaston (1991) interviewed taxonomic specialists of many insect taxa, including for beetles, and concluded that the proportions of undescribed species in tropical samples were high but not high enough to warrant an estimate of 30 million species (if 1 million species are known and 30 million exist, undescribed species would comprise 97% of samples).

Other means of estimating global species richness have been used. Rates of description of new species of beetles ("trend lines") have been extrapolated to an asymptote, but this method incorrectly assumes constant levels of scientific study and publication. May (1988) compared observed with theoretical distributions of body size and abundance to estimate the numbers of undescribed species, primarily those in the smaller size classes. It is certainly true that larger organisms are more likely to be collected and described than are smaller organisms. In each of the other major orders of insects, it is the smaller species that are least well known: for Lepidoptera, it is the microlepidoptera; for Hymenoptera, the smaller parasitic wasps in the Chalcidoidea; for Diptera, the small Nematocera and Phoridae. In beetles, Gaston showed that larger beetles were described earlier and that the average size of the species being described decreases as one nears the present. Of the zygopine weevil species identified at La Selva, 57% of weevils larger than 4 mm in length are undescribed (88 measured), but 81% of those less than 4 mm are undescribed (146 measured). Methods of estimating unsampled species from theoretical abundance frequency distributions depend on the mode of the distributions being defined, which is not usually possible for tropical samples. It seems to be true for all estimates of global beetle or insect species numbers, regardless of the sampling method used, that what is not known is so large compared to what is known that a precise estimate is impossible.

What do We Know and How Likely are We to Ever Know?

Beetles are a very large group of organisms and are diverse both taxonomically and ecologically. The 169 families and 340,000 described species currently comprise somewhat less than half of all known insects and about a quarter of all living organisms. About 40 families of beetles have more than 1000 described species, and 8 families have more than 10,000 species. About half of all beetle species feed on plants, usually as endophages on woody or other support tissue, but there are a large number of predatory species, as well as lesser but significant numbers of fungivores, detritivores, and aquatic forms. This ecological diversity precludes simple explanations for the diversity of beetles. However, many beetle families show general patterns of distribution, for example, greater diversity in tropical latitudes and the greater importance there of antipredator defenses such as mimicry.

These crude generalizations aside, we actually know rather little about most beetles. Because of the popularity of certain beetle groups, the various families have been unequally studied. Parts or all of the families Carabidae, Scarabaeidae, Buprestidae, Tenebrionidae, and Cerambycidae have been relatively better studied, whereas the three largest families (Staphylinidae, Chrysomelidae, and Curculionidae) are the most poorly studied. Ironically, as we have become aware of the existence of large numbers of undescribed species by using new sampling methods in previously little-collected tropical areas, the number of professional taxonomists able to name and describe species has been declining. Many positions for taxonomists have been lost at most major museums, funding for taxonomic research has declined, and the trend in research has been toward genetic analyses for the purpose of elucidating phylogenetic relationships rather than describing species. Money has become available for sampling and conducting species inventories, especially in threatened tropical areas, but not for processing and studying the material collected beyond simple sorting of a few target or focal taxa and counting of their putative "morphospecies." May (1988) showed that, although beetles are the largest group of organisms, they have the lowest rate of publications per described species of all organisms except nematodes. He expressed amazement that there is, in fact, no complete worldwide catalog of insects that would allow one to know how many are actually described (hence the varying "estimates" of even the numbers of described species).

At the level of local or regional faunas, there are few sites outside of Western Europe where beetle faunas are more than partially known. The richest areas—the tropics, and especially the Neotropics—are virtually unknown faunistically outside of samples taken at a few widely scattered sites. Many large regions of the tropics have never been significantly collected for any beetle groups. Costa Rica's National Biodiversity Institute (INBio), including the ALAS project at La Selva, has made the most extensive collections toward a regional fauna of any tropical country, but their collections favor larger species, are only partially mounted and sorted, and are only beginning to be studied. Overall, then, there is no good estimate of global species richness for beetles and few data on which to base such an estimate.

As the magnitude of how much we still don't know becomes clearer, there are serious questions about whether we will ever understand the extent of beetle diversity in light of the declining number of taxonomically trained scientists and the possibility that many species will go extinct because of human activities before they are ever collected or studied. The decline in the number of scientists who are able or willing to describe species is unlikely to be halted. Furthermore, widespread beetle extinction over the next century is especially likely in the face of the accelerating environmental destruction of the richest sites in the tropics to accommodate growing human populations and, even more significantly, to maintain Western European and North American standards of living. Unfortunately, it is a safe guess that we will never completely know how many beetles there were.

Acknowledgments

I thank F. T. Hovore for reviewing this manuscript and allowing me to publish some of his data. I have been supported by the Research Committee of the UCLA Academic Senate and by grants in support of the ALAS Project (National Science Foundation Grants BSR 9025024, DEB 9401069, and DEB 9706976).

See also: Arthropods (Terrestrial), Amazonian. Insects, Overview. Invertebrates, Terrestrial, Overview

References

Anderson RS (1995) An evolutionary perspective on diversity in Curculionoidea. *Mem. Entomol. Soc. Washington* 14: 103–114.

Crowson RA (1981) *The Biology of the Coleoptera*. London: Academic Press.

Erwin TL (1982) Tropical forests: Their richness in Coleoptera and other arthropod species. *Coleopts. Bull* 36: 74–75.

Farrell BD (1998) "Inordinate fondness" explained: Why are there so many beetles? *Science* 281: 555–559.

Gaston KJ (1991) The magnitude of global insect species richness. *Conservation Biol* 5: 283–296.

Hammond PM (1993) Insect abundance and diversity in the Du-moga-Bone National Park, N. Sulawesi, with special reference to the beetle fauna of lowland forest in the Toraut region. In: Knight WJ and Holloway JD (eds.) *Insects and the Rain Forests of South East Asia (Wallacea). Symposium of the Royal Entomology Society of London*, pp. 197–254. London: Royal Entomology Society of London.

Hespenheide HA (1994) An overview of faunal studies. In: McDade L, Bawa KS, Hespenheide HA, and Hartshorn GS (eds.) *La Selva: Ecology and History of a Neotropical Rainforest*, pp. 238–243. Chicago: University of Chicago Press.

Lawrence JF (1982) Coleoptera. In: *Synopsis and Classification of Living Organisms*, 482–553. New York: McGraw-Hill.

May RM (1988) How many species are there on earth? *Science* 241: 1441–1449.

Pearson DL and Cassola F (1992) World-wide species richness patterns of tiger beetles (Coleoptera: Cicindelidae): Indicator taxon for biodiversity and conservation studies. *Conservation Biol* 6: 376–391.

Wolda H, O'Brien CW, and Stockwell HP (1998) Weevil diversity and seasonality in tropical Panama as deduced from light-trap catches. *Smithsonian Contr. Zool* 590: 1–79.

Biodiversity and Cultural Ecosystem Services

Franz W Gatzweiler, University of Bonn, Bonn, Germany
Konrad Hagedorn, Humboldt-University of Berlin, Berlin, Germany

© 2013 Elsevier Inc. All rights reserved.

Glossary

Culture Culture can be understood as the mental models, norms, values, codes of conduct, habits and rules which have evolved within different groups of humans to facilitate and constrain human perceptions, feelings, and actions among each other and between humans and their natural environment. Hofstede (2001, p. 10) defines culture as the "interactive aggregate of common characteristics that influence a human's group response to its environment" and explains, with reference to Guilford (1959), that "culture is to a human collectivity what personality is to an individual." Culture can be expressed at different levels of depth as visible and invisible manifestations. Although values are at deeper levels, tend to change slowly (Williamson, 2000, p. 597; North, 1991, p. 111) (Williamson places values at the embeddedness level of societies at which he also places norms, customs, and traditions. At that level, Williamson explains (in Williamson, 2000, p. 596), institutions change in the order of centuries or millennia. North states that institutions at that level have a "…pervasive influence upon the long-run character of economies." Both assume change at evolutionary time scales.) and are invisible, rituals, heroes, and symbols are outer layers of cultural manifestations around this center and are visibly expressed in practices. The UNESCO Declaration on Cultural Diversity (2002) defines culture "as the set of distinctive spiritual, material, intellectual, and emotional features of society or a social group, that encompasses, in addition to art and literature, lifestyles, ways of living together, value systems, traditions, and beliefs"(UNEP, 2003).

Institutions The definition of what an institution is and what it does differs according to different models of human behavior. The cognitivist or social constructivist model and the new institutional economists' model differ mainly in the role which the individual has in the construction of institutions. According to the social constructivist model the way humans perceive the external world (including society itself) is socially constructed and social behavior of an individual is shaped in the context and by the institutions of society. According to Scott (1995, p. 33, cited from Vatn, 2005, p. 10) institutions "consist of cognitive, normative, and regulative structures and activities that provide stability and meaning to social behavior. Institutions are transported by various carriers – cultures, structures, and routines – and they operate at multiple levels of jurisdiction." In contrast, the new institutionalist model sees institutions as the "rules of the game in a society, or more formally, are the humanly devised constraints that shape human interaction" (North, 1990, p. 3, cited from Vatn, 2005, p. 10).

Memes Memes, in analogy to genes, are a basic unit of cultural transmission which are passed on from one collective memory to another, also over generations, by a process of imitation (Dawkins, 2006). Examples of memes are virtues, ideas, beliefs, fashions, values, and norms.

Structural coupling Structural coupling is here referred to as the perception, recognition, and conceptualization of the natural or social environment, representation of it as cognitive or mental maps, and the adjustment and readjustment of behavior according to the valuation. This representation is not merely visual or static, it also serves as a guide for decision making and action according to a certain rationality. The cognitive guide, or compass, is part of what constitutes "culture" and valuation serves as a feedback and calibration mechanism within the process of structural coupling. Valuation informs the conscious mind about the consequences of actions guided by particular sets of values.

Valuation language Juan Martinez-Alier (2003) provides examples of languages of valuation: the economic language, the languages of territorial rights, environmental justice, livelihood, and sacredness. Valuation languages articulate particular values, concerns, beliefs, or expectations. Duygu Avci and Özkaynak (2010), for example, review valuation languages of resistance in environmental conflicts over resources in Turkey.

Value In many economic models value is something people are willing to give up something for in order to get or enjoy it (Kumar, 2010, p. 19). The Millenium Assessment (2005) has defined value as "the contribution of an action or object to user specific goals, objectives, or conditions." Values can be viewed as manifestations of culture. They can take the form as social moral norms or as preferences and measures of importance; both strongly influencing behavior. Values guide decision making. In neoclassical economics "value" refers to the ability of a good or service to satisfy desires or wants. Values as ethical and moral norms refer to principles of what people consider to be good and bad or right and wrong. Different types of values are not always commensurable (e.g., the value of bread and the value of friendship). In neurobiology values are associated to states of the neurobiological system which produces emotions and feelings. Values are different states of affective systems of response, or emotional significance given to experiences. This understanding of values from the neuroniological perspective is relevant because values are not to be confused with objective pre-existing facts which only need to be revealed by means of specific valuation methods. Deontological positions and viewpoints emphasis norm-type of values and consequentialist positions emphasis preference-type of values. Whether norm-type or preference-type of values influence behavior depends on different rationalities of human behavior, the (social and biophysical) context of the decision making situation,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00423-8

including the options the decision maker has and the individual brain, especially with regards to its level of consciousness.

Value articulating institutions According to Vatn (2005, p. 301) and Jacobs (1997) value articulating institutions are the institutional structures or the rules according to which values are articulated and according to which the valuation process is carried out. For example, contingent valuation, multicriteria analysis and different deliberative valuation methods. These institutions define who participates, in which role, how (e.g., in meetings or individual interviews), what counts as data and what form the data should take, and the data handling process (e.g., how data are weighed or aggregated). The choice of a value articulating institution may strongly influence which preferences or values become articulated.

Introduction

There is a dual relationship between biodiversity and cultural ecosystem services. On the one hand biodiversity enables the provision of "cultural ecosystem services" which are beneficial for human use. On the other hand culture and its institutions provide a social lens through which ecosystem services are not only perceived and conceptualized but also valued. Values are culturally specific and a compass for human action, also on biodiversity. This dual and mutual relationship is interlinked and subject to change. Properties of ecosystems that people regard as useful may change, even if the ecosystem itself remains rather constant, and changes in social values may lead to changes in ecosystems and the services they provide.

In the following the authors will explore the linkages between biodiversity and cultural ecosystem services by describing what types of cultural ecosystem services are provided by biodiversity, explaining how culture shapes the way nature is perceived and valued and how these values are expressed. They examine the properties of this mutual relationship and the particular role that languages of valuation and value articulating institutions play within cultures as regulation and feedback mechanisms and for achieving sustainable outcomes from the interactions between biodiversity and culture.

Ecosystems and Their Cultural Services

Any analysis is explicitly or implicitly guided by a conceptualization of the subject matter that is to be investigated. For our purpose the concepts of biodiversity, ecosystem functions, services, benefits, and values are refered as suggested by the group of authors involved in the study on The Economics of Ecosystems and Biodiversity (Kumar, 2010, pp. 17–19). According to The Economics of Ecosystems and Biodiversity (TEEB), ecosystem functions "are defined as a subset of the interactions between ecosystem structure and process that underpin the capacity of an ecosystem to provide goods and services." The relationship between ecological processes and structure, functions, and services, and the benefits derived and values attached to them is depicted in **Figure 1**. Biomass production, for example, is required to maintain a population of buffalos, which can be hunted to provide the service of food from which humans benefit. Another example related to cultural ecosystem services: the structure and processes of ecosystems and biodiversity provide information services for science, leisure or spiritual experience and cultural expression.

Ecosystem services (provisioning, regulation, habitat, and cultural) are therefore both scientific concepts and social constructs of biophysical functions that are useful for human

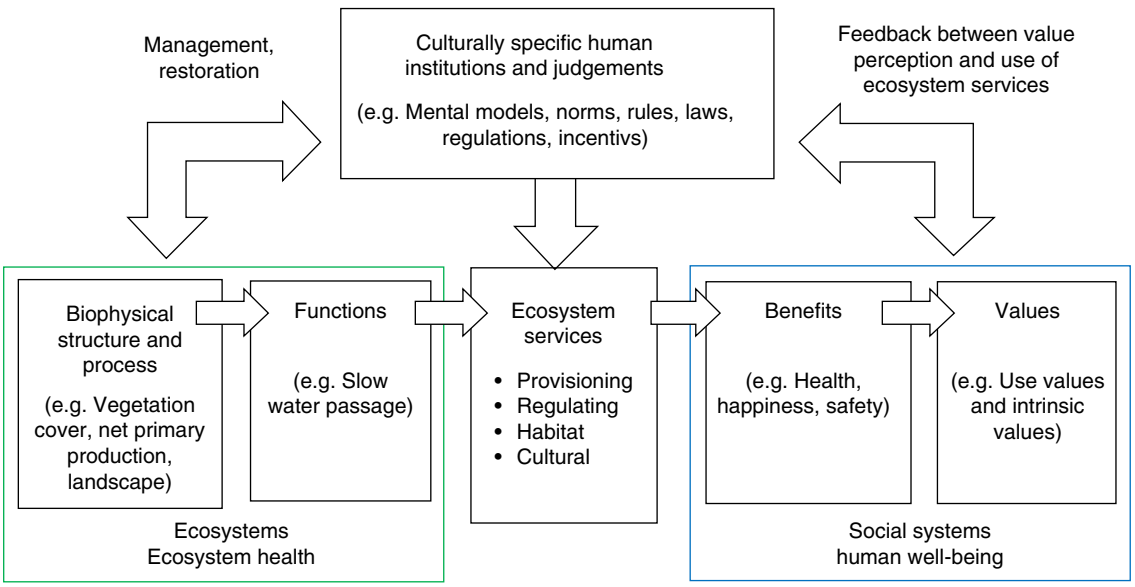


Figure 1 Ecosystem functions provide services for the benefit and well-being of people.

Table 1 Comparison of maximum monetary values (in $ ha^{-1} year^{-1}) of ecosystems services types in different biomes

	Provisioning services	*Regulating services*	*Habitat services*	*Cultural services*
Open oceans	22	62	0	0
Coral reefs	20,892	33,640	56,137	1,084,809
Coastal systems	7549	30,451	164	41,416
Coastal wetlands	8289	135,361	68,795	2904
Inland wetlands	9709	23,018	3471	8399
Rivers and lakes	5776	4978	0	2733
Tropical forests	9384	7135	5277	1426
Temperate forests	1736	456	2575	96
Woodlands	862	1088	0	0
Grasslands	715	2067	298	11

Note: Values are taken from min–max ranges of a number of individual case studies for each biome and show only the maximum values. The valuation studies which have been referred to for generating the values do not value and compare all services for each biome. That means, for example, the same case study in which the provisioning services of grasslands have been valued, does not also value the cultural services of the same biome, so that the values are not comparable. They stem from an attempt to provide an overview of values found in the literature for each biome but are useful for getting an idea of the importance of each ecosystem service.
Source: Reproduced from de Groot RS and Kumar P (2010) Estimates of monetary values of ecosystem services. In: Kumar P (ed.) *The Economic of Ecosystems and Biodivesity – Ecological and Economic Foundations*, pp. 368–401. London, Washington: Earthscan.

well-being (**Figure 1**). Provisioning ecosystem services includes, for example, the provision of food, water, raw materials, genetic resources, medicinal, and ornamental resources. Regulation services refer to services such as climate regulation, the regulation of water flows, waste treatment, erosion prevention, and the like. Habitat services maintain genetic diversity and the life cycles of species.

Examples of specific cultural ecosystem services are provided by de Groot (1992, pp. 120–128) and de Groot and Ramakrishnan (2005) and the value of cultural ecosystem services as compared to provisioning, regulating, and habitat services is shown in **Table 1**:

- Esthetic information and recreation, for example,
  - The value of housing in esthetically pleasing environments is often higher when houses are located in scenic natural environments.
  - For recreational purposes people look for landscapes which vary from those they are used to from daily experience. The Millenium Assessment (2005, p. 14) estimates that nature- and culture-based tourism employs approximately 60 million people and generates approximately 3% of global gross domestic product (GDP).
- Information for cognitive development, spiritual, and religious experience, for example,
  - The notion that ancestors observed the same landscape as those people grew up in and spent their childhood, can lead to a feeling of unity with nature.
  - Spiritual or religious values are often attached to nature areas, water or single species, like trees, in some societies (e.g., a holy tree).
  - Very often, in indigenous societies, spiritual and religious ties to nature are reflected in social norms and customs to ensure adequate behavior toward nature. There are examples of the values of indigenous forest gardens in West Kalimantan which had been manifested in local legal systems and sanctioning mechanisms (Gatzweiler, 2003).
  - Sacred natural sites and cultural landscapes institutionalized all over the world usually have a substantial meaning for people in the regions (UNESCO, 2003, 2008a).
- Inspiration for culture, art and design, for example,
  - Nature is used as a motive in many books, magazines, films, and other media. It is an important motive in paintings, sculptures, and popular art. Mountains, trees, leaves, or animals are used as symbols in national flags.
  - Architecture and fashion is inspired by naturally occurring shapes and designs.
- Information for education, history and science, for example,
  - Science uses functions of nature as a service of information provision in medicine or bio-technology. The estimated size of the market for bio-prospecting in 2020 is 100 million USD year^{-1} (Bishop, 2010, p. 11).
  - The field of bionics mimics structures, processes, and systems in nature for engineering and modern technology.

The linkages between biodiversity and the diversity of cultural services it provides are manifold. The authors of the report of an international workshop on the linkages between biological and cultural diversity (UNESCO, 2008b) identified the following areas of interdependence:

- Language and linguistic diversity (e.g., terms, concepts and categories related to nature).
- Material culture (e.g., objects produced from or created by biodiversity).
- Knowledge and technology (e.g., traditional and local knowledge about places, ecological relations; early warning systems, risk management; traditional medicine).
- Modes of subsistence (e.g., resource-based livelihoods, plant and animal domestication, and breeding).
- Economic relations (e.g., management of common pool resources, partnerships based on trading natural resources).
- Social relations (e.g., attachment to places, such as sacred sites; social roles related to differential resource use; gender specific environmental knowledge; political relations defined by control over resources; legal institutions, such as customary law and contemporary conventions).

The particularity of cultural ecosystem services, however, is that they are not only perceived as such but also influence and shape the process of perception itself. The values attached by people to different cultural services is strongly dependent on the cultures of those people. In other words, observing cultural ecosystem services not only mobilizes cognitive schemata internalized by humans but also contributes to how nature-related cognitive schemata form. Most obvious is this relationship in the example of values derived from cultural ecosystem services in indigenous societies (Gatzweiler, 2004, p. 149). In this example, the values attached to individual species of local ecosystems are translated into sanctioning mechanisms and part of the local indigenous legal system. Sanctions for cutting valuable trees do not only reflect the value of the tree in terms of its use values but also the dependence of the group on an intact forest environment in which a single tree can play an essential role.

Conceptualizing biodiversity as "goods and services" provided by the ecosystem for human use is one example of a cognitive construction used in economics to guide human behavior in interaction with its natural environment. From the perspective of economics the value of the environment to people depends on the goods and services it yields. These goods and services are valued for their usefulness to human beings, and different cultures articulate their value by means of different types of 'valuation languages' (Martinez-Alier, 2003). The grammar of those valuation languages, i.e., the rules according to which value is articulated, is referred to as 'value articulating institutions' (Vatn, 2005).

Valuation languages, value articulating institutions and other systems of rules and organization including the incentives and rewards that result from such institutions and governance structures are socially constructed in human societies, and they have their own specific dynamics. A particular feature of cultural ecosystem services is that they depend on both functioning of the external environment and the internal processes of human perception, cognition, and of what valuing such "services" means.

This awareness is part of the cognitive process in which valuation plays an important feedback mechanism. Knowledge about the process of cognition and social constructions plays an important role for how people choose to value biodiversity and ecosystem services and how they are managed in accordance to those values. Knowledge regarding the importance of ecosystem functions for human well-being is fed by the information provision service of cultural ecosystem services.

Culture and Ecosystem Services

Cultural differences are fundamental in the way people conceive and act on the natural environment and also how they perceive nature and understand it as, for example, a provider of ecosystem services. Descola (in Descola and Pálsson, 1996, p. 82) suggests an analytical model to describe how different societies objectify their relationship to nature: (1) modes of identification (e.g., the relation between humans and nonhumans becomes social and through this social relation the distance between both is reduced and socially defined, like in animism, totemism, and naturalism), (2) modes of interaction (reciprocity, predation, protection between species), and (3) modes of categorization (e.g., analogy and metaphoric similarities based on morphological resemblances). Descola's analytical approach is useful in that it enables us to place the scientific and naturalist perspective within the Judeo-Christian tradition of control and utilitarianism, and the historical separation of social and biological sciences and the consequent epistemological distinction between culture and nature.

Other epistemological strands see biological diversity as being profoundly interrelated with human's cultural diversity. The term 'bio-cultural diversity' reflects the belief that humans and nature are connected and both constitute the diversity of life on earth. As such, bio-cultural diversity includes biological, cultural, and linguistic diversity (Maffi, 2001; Maffi and Woodley, 2010) and emphasizes that the structure of ecosystems and human systems are mutually linked (Berkes *et al.*, 2000). This linkage encompasses three components (**Figure 2**):

1. The human brain as neural network, enabling cognition and (in combination with other brains) the social construction of what constitutes value and the production of culture.
2. Culture as collective construction within societies.
3. The natural environment, perceived as nature, biodiversity, or environmental goods and services and its information provision functions.

Culture here refers to the mental models, norms, values, codes of conduct, habits, and rules which have evolved within different groups of humans to facilitate and constrain human perceptions, feelings, and actions among each other and between humans and their natural environment.

Hofstede (2001, p. 10) defines culture as the "interactive aggregate of common characteristics that influence a human's group response to its environment" and explains, with reference to Guilford (1959), that "culture is to a human collectivity what personality is to an individual." Culture can be expressed at different levels of depth as visible and invisible manifestations. Although values are at deeper levels, they tend to change slowly (Williamson, 2000, p. 597; North, 1991, p. 111). (Williamson places values at the embeddedness level

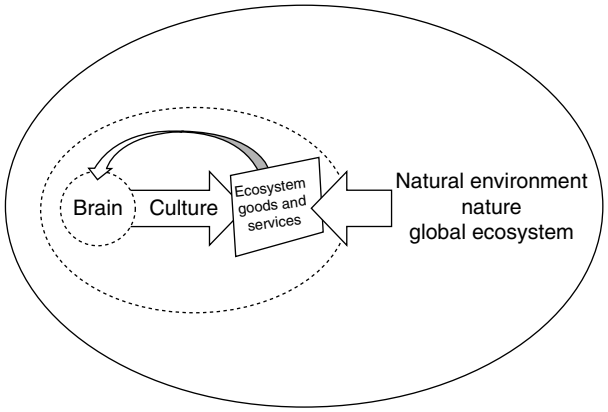


Figure 2 The coevolution of culture and perception of ecosystem services.

of societies at which he also places norms, customs, and traditions. At that level, Williamson explains (Williamson, 2000, p. 596), institutions change in the order of centuries or millennia. North states that institutions at that level have a "…pervasive influence upon the long-run character of economies." Both assume change at evolutionary time scales.) and are invisible, rituals, heroes and symbols are outer layers of cultural manifestations around this center and are visibly expressed in practices.

Culture therefore has an enormous influence on how individual brains perceive their environment and how humans act on it. Through culture our natural environment is perceived, recognized, and conceptualized and thereby built into the cognitive maps that our brain helps us to construct. Such cognitive maps or internal "neuropsychological structures" (Wexler, 2008, p. 5), guide human behavior, like a compass. Especially during childhood and preadulthood the human brain is strongly being altered and shaped by culture, whereas the individual is capable of altering culture in later stages of its ontogenesis (Wexler, 2008). Culture is thereby formed through a constant adjustment process (e.g., learning) between internal (brain) and external (ecosystem) structures.

However, internal neuropsychological maps or mental maps are not "objective", one-to-one, congruent representations of the external world. Also, a map or a compass alone, do not help in deciding which direction to go and which action to take. What is useful in addition to a neuropsychological representation of the external world (cognition) is the ability to make those representations felt; that is consciousness.

Damasio (1994, 1999, 2009, 2010) explains that every representation is connected to an emotion and to feelings. Although emotions are bodily states rapidly generated in the brain stem and allow humans and animals to react without thinking, feelings are conscious perceptions (awareness) of emotional states. Feelings are between the calculative mind, which can trade off costs and benefits as cold facts without ever coming to a decision, with emotions, which weigh or value the factual analysis and enable decisions to be made. Values are different states of affective systems of response which give emotional significance to experiences.

The process of culture being formed by a constant adjustment process between internal and external structures has also been referred to as "structural coupling", a process that is structure determined and structure determining and that leads to structural congruence between two or more systems. Rizzello (1999, p. 79) has described the process as cognitivism. This process of matching internal representations with external (changing) environments (Wexler, 2008, p. 15), or structural coupling, is as important for survival as diversity in social and ecological systems. Diamond provides examples and explains the collapse of societies in history as a result of a failure to adjust established social or cultural values to the requirements of new environments (Diamond, 2011).

Applied to socio-economic and ecological systems, Norgaard (1994, p. 77) has explained structural coupling in the development process as a coevolutionary process in which he defines subsystems such as values, knowledge, technology and social organization, which interact with each other and the environment in certain ways, depending on their specific traits or culturgenes (Norgaard, 2010, p. 3).

Culturgenes, according to Lumsden and Wilson (1981), or "memes" according to Richard Dawkins (2006, p. 189) are a basic unit of cultural transmission, similar to genes, which "propagate themselves in the meme pool by leaping from brain to brain via a process, which in the broad sense can be called imitation" (Dawkins, 2006, p. 192). Examples of memes are virtues, ideas, beliefs, fashions, values, and norms passed on from one generation to another through social interaction, language and imitation. In that process values are transmitted through the internalization of norms (Gintis, 2004).

According to Hofstede (2001) the way people value, think about, and behave on nature is rooted in the different layers of cultural manifestations and transferred from one person to another and one generation to another, by means of language and how value is articulated. Two ways through which the values of a culture are transmitted are through languages of valuation and value articulating institutions.

Languages of Valuation and Value Articulating Institutions

When culture is transferred from one generation to another, language and discourse play the role of important transmission mechanisms. The way people value biodiversity in different cultures is expressed in different valuation languages (Martinez-Alier, 2003, p. 79). "Languages of indigenous territorial rights, human rights, livelihood, sacredness, environmental values, esthetic values, and cultural values…" (Martinez-Alier *et al.*, 2010, p. 154) or religious languages of social justice (Martinez-Alier, 2003, p. 207) are examples of languages through which value is expressed.

Which language dominates and influences decision-making depends, amongst others, on who has the power of imposing a particular language over others (Martinez-Alier, 2001). Today one of the most dominant and influential languages of valuation is that of economics and chrematistics (Pack, 2010, p. 16; Marx, 1994, p. 260), which has evolved from an economic culture or economism (Economism is the "changing mix of knowledge beliefs and values related to, reinforced by, and essential to the short run operation of the industrial order" (Norgaard, 2010). Freudenberg *et al.* (2009) conceptualize economism around the assumption that (people) know their needs, wants, are self-interested and motivated mainly by economic incentives. In economism "markets are the arbiters of social exchange" and thereby economism "divorces material context from culture... (and)…privileges material well-being over other contributors to human health and wholeness."). It is a language used and understood by most people on earth, across all cultures. The conception of nature as a service provider stems from this economic culture. Economics as an epistemological tradition has also evolved from a specific socio-cultural context and economic values are part of culturally constructed realities and belief systems of particular societies (Wilk and Cliggett, 2007).

Economic values are expressed and communicated by means of opportunity costs and measured in monetary units

Table 2 Value articulating institutions and respective normative and epistemological stances

Value articulating institutions	*Normative and epistemological stance*
Contingent valuation method	*Cartesianism*: Value is pre-existing and needs to be discovered/revealed by, for example, willingness to pay questions. Separation between values and facts, human and nature. Substitutability between money values and ecosystem goods and services
Deliberative or social process methods	*Democracy stance*: Value is constructed in social processes. Previously unknown values evolve from deliberation and debate
Multicriteria methods	*Complexity*: Value is understood in terms of ranked importance. Irreducible plurality of analytical perspectives

Source: Modified from O'Connor M, Funtowicz S, Aguiler-Klink F, Spash CL, and Holland A (1998) *Valuation for Sustainable Environments. The VALSE Project, Full Final Report.* Ispra, Italy: European Commission Joint Research Centre, and Brondizio ES and Gatzweiler FW (2010) The socio-cultural context of ecosystem and biodiversity valuation. In: Kumar P (ed.) *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations*, ch. 4, p. 163. London, Washington: Earthscan.

that Martinez-Alier (2003, p. 228) refers to as "language of chrematistics". Despite this dominance of economics (Costanza, 1991; Common and Stagl, 2005) multiple valuation languages can be applied in decision-making, "without being previously translated into a common monetized bottom line" (Martinez-Alier *et al.*, 2010, p. 154; Zografos and Martínez-Alier, 2009). This diversification of languages of valuation is part of the diversity of values and cultures.

Along the same lines (Zografos and Martínez-Alier, 2009, p. 1726) analyzed a case study of wind farm conflict in rural Catalonia and come to the conclusion that "the absence of (diverse) opportunities for meaningful deliberation in decision making and the predominance of decisional bottom lines curtail claims to fairer distribution of costs and benefits from locally hosted energy developments, as well as alternative landscape value claims, and that this fuels conflict."

The choice of a particular valuation language is not a "natural" process in the sense that people do not make conscious choices. Languages of valuation by cultural or interest groups (e.g., the industry, church, political parties, academics, or indigenous people) also carry political and social goals, i.e., choosing how to perceive nature, for example, as ecosystem services, is purpose driven. Ellen (in Ellen and Fukui, 1996, p. 28) explains that particular ways of conceiving the environment can serve the interest of particular groups.

"Whether reinventing an image of the noble ecological savage, or defining nature with a price tag, or constructing a shared image of the global environment, these models carry political and social goals and have far-reaching consequences" (Brondizio and Gatzweiler in Kumar, 2010, p. 153). The field of political ecology is much about the conflicts and politics of how people interact, solve conflicts and make decisions on how to shape nature (Robbins, 2004).

The political ecology aspect of valuation also explains why the linkage between cultural and biological diversity is not direct and clear, although popularly, high cultural diversity is frequently propagated as a precondition of high biological diversity. High cultural diversity can mean "many different but small culturally distinct groups" who may each have close traditional ties to their natural environment but weak decision making power. In contrast, another group of people which may be culturally less diverse but therefore dominant in their decision making influence can transform biodiversity according to the values it holds. This can lead to conservation or reduction of biodiversity by the group with more influential decision making powers and not by the group which is culturally more diverse. Examples are plenty among the indigenous people's movements around the world.

Although choosing a particular valuation language is driven by purpose, the way according to which values are articulated within a specific institutional and cultural setting (once the choice has been made) is determined by rules – the so-called value articulating institutions. Vatn (2005) refers to valuation methods as value articulating institutions. These institutions define certain models of human behavior and different norms of rationality and as such define the assumptions and rules under which people articulate values.

Although the institutions of "economism" can have a dominating influence on human behavior and the choice of values, the institutionalism as understood by Vatn (2005, p. 105) rejects the sole norm of rationality in welfare theory and the idea that people are merely individualistic calculators striving to maximize their own utility. Rather, as Sagoff (1988) puts it, people "act within institutions of public decision making" and these institutions, which are culturally specific, define what is right and wrong, good or bad, valuable or worthless (Vatn, 2005, p. 105; Wilk and Cliggett, 2007). Brondizio and Gatzweiler (in Kumar, 2010, p. 163) present an overview of how different normative and epistemological stances of value are related to different value articulating institutions (**Table 2**) and conclude that for the purpose of biodiversity valuation, choosing a value articulating institution can be more important than calculating a correct value.

Biodiversity, Culture, and Sustainability

Cultural diversity has been recognized together with biological diversity as important ingredients for sustainable development (UNEP, 2003). Although there is no common interdisciplinary conceptual and methodological framework that clearly defines the dynamic interface between biological and cultural diversity the authors have pointed to linkages which are important to consider for achieving the goal of sustainable development. On that path the "TEEB" report (Kumar, 2010) has been a milestone.

One of the core messages of the TEEB report is that while aiming at a full account of all economic values is necessary for sustainable use of biodiversity, it is not sufficient. Gowdy (1997) (in Kumar, 2010, p. 272) explains that "Biodiversity value can be seen as layers of a hierarchy moving from market value, to nonmarket value to humans, to ecosystem value.

These various levels of biodiversity value point to the need for a pluralistic and flexible methodology to determine appropriate policies for its use and preservation….ultimately economic value represents only a small portion of the total value of biodiversity and ecosystems". The backward link to culture again is that such a pluralistic methodology involves the recognition of a diversity of valuation languages and value articulating institutions, which are part of cultural diversity.

Culture as defined by the UNESCO Declaration on Cultural Diversity (the UNESCO Declaration on Cultural Diversity (2002) defines culture "as the set of distinctive spiritual, material, intellectual and emotional features of society or a social group, that encompasses, in addition to art and literature, lifestyles, ways of living together, value systems, traditions, and beliefs.") is (also) understood as a characteristic of people with a common mode of thinking and acting (UNEP, 2003). Although it is difficult to evaluate the impact a specific ethnic or cultural group has on biodiversity (not least because of the political ecology and political economy aspects mentioned above) it is possible to describe the quality of this relationship when considering 'a culture of acting' or an 'epistemological culture'. Different cultures of thinking and acting thereby take into account and recognize more or less different types of values. If confined to market values only, the likelihood is high that noneconomic aspects of biodiversity are left aside and the sustainability of actions remains uncertain.

A culture which attaches and communicates only a narrow set of values for biodiversity will shape biodiversity according to those values it recognizes and appreciates most. Attaching economic values to biodiversity will strengthen its resource character – those aspects of biodiversity which are known to humans, which they perceive as useful and for which they are willing to pay. The result of such behavior is the defect "economic compass": economic incentives and institutions are blind for the noneconomic values of biodiversity which are lost.

Attaching also noneconomic and intrinsic values to biodiversity (Gatzweiler, 2008), values which are difficult or impossible to express monetarily, which follow not only the consequentialist view in the tradition of economics (meaning that the value of biodiversity is a means to a higher end which is the well-being of humans.), but also the deontological position (meaning that biodiversity has a value independent from its usefulness to humans, according to moral standards independent from its usefulness or contributions to higher ends.), will shape our environment and produce cultural landscapes which can not immediately be explained by their economic usefulness to humans. Just those values however are transmitted by and produce the cultural ecosystem services.

Consequently, if the diversity of cultures per se and the diversity of values and institutions are not a sufficient explanation for the extent to which biodiversity is conserved or lost, the power and opportunity to translate values into action and implement institutions in a way that it influences behavior through effective institutional environment, could be a better explanation of the linkage between biodiversity and culture and how it needs to be shaped for sustainability. This notion on how systems of rules and governance structures emerge in concrete action arenas and action situations in order to co-ordinate the coevolution between ecological and social systems has been operationalized in an analytical frame work developed by Hagedorn *et al.*, called "IoS" (Hagedorn *et al.*, 2002; Hagedorn, 2008) (the IoS framework provides the function of a research heuristic, which relates four key elements of institutional innovation and institutional performance in social–ecological systems (SESs) to each other: transactions, actors, institutions, and governance structures. According to the IoS-Framework, the properties of the respective transactions and the characteristics of the involved actors determine which institutions (sets of rules, norms, and values) develop and through which governance structures (organizational forms) these institutions will in practice be implemented. Ecological and social systems find themselves in a permanent process of coevolution. Each system is dependent on the other; their existence, thus, depends on their reciprocal ability to adapt. To assure such a coadaptation, "IoS" are required. These are sets of rules that, through linkages to the appropriate organizational forms, make felt their coordinating effects. They either emerge from processes of institutional change or are explicitly created. IoS, thus, take over a coordinating function between social and ecological systems and order their reciprocal influences. The relationship between humans and nature is, thus, seen as a socially constructed culture, in which the emergence or creation of new institutions or institutional change is always a question of reasons or drivers, which may or may not be of a practical nature. They result from, for example, the desire to economize on transaction costs or resolve conflict between actors, but may also have reasons that are far beyond such economic or interest-based motives, for example, preserving nature for its own right. This approach has two main advantages: Firstly by including nature-related transactions the impact of Ecological, biological, geological, and other physical systems on how cultures evolve and how they shape our habits and behavior is explicitly made a part of analyzing SESs. Secondly, it enables such analyses of those institutions and governance structures which represent nature-related and socially obstructed cultures such as biodiversity cultures to become concrete and operational. This contributes to a better and also more detailed understanding of processes of institutional and organizational change with an emphasis on the rules, norms, values and organization of the use, and protection of natural systems such as soil, water, air, energy, climate, biodiversity, genetic resources, vegetation and animal populations, ecosystem, and natural common-property goods such as fish, forest, game, and rangeland resources.)

To conclude:

a. The cultural ecosystem services comprise a set of largely nonconsumptive benefits valued for their contribution to human well-being.
b. These services are frequently among the most valuable nonmarketed ecosystem services.
c. A particular feature of cultural ecosystem services is that they not only depend on the functioning of the external environment but also on the internal processes of human perception, cognition and of what valuing such services means.

d. Consciousness is being aware of the self and which role it plays in the cognitive process in which valuation plays an important feedback mechanism.
e. The values of cultural ecosystem services frequently lie outside the market, are incommensurable, and are expressed in valuation languages that are culturally specific. Different valuation languages play an important role in conflicts over access to resources and recognition of rights.
f. For the purpose of biodiversity valuation and depending on the socio-cultural context of valuation, the choice of a particular value articulating institution can be more important than calculating a correct value.

Achieving society's equity, sustainability and efficiency goals requires that they all be taken into account in social decision-making. This pluralism of valuation systems can be taken into account by maintaining diversity of valuation languages and not restricting the choice of diverse value articulating institutions.

References

Berkes F, Folke C, and Colding J (2000) *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience*. Cambridge: Cambridge University Press.

Bishop J (2010) TEEB – The Economics of Ecosystems and Biodiversity. *Report for Business – Executive Summary*.

Brondizio ES and Gatzweiler FW (2010) The socio-cultural context of ecosystem and biodiversity valuation. In: Kumar P (ed.) *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations*, ch. 4. London, Washington: Earthscan.

Common MS and Stagl S (2005) *Ecological Economics: An Introduction*. Cambridge: Cambridge University Press.

Costanza R (1991) *Ecological Economics: The Science and Management of Sustainability*. New York: Columbia University Press.

Damasio AR (1994) *Descartes' Error. Emotion, Reason and the Human Brain*. New York: Quill/Harper Collins.

Damasio AR (1999) *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*. New York, London: Harvest Book, Harcourt Inc.

Damasio AR (2009) Neuroscience and the emergence of neuroeconomics. In: Glimcher PW, Camerer, Fehr, and Poldrack (eds.) *Neuroeconomics. Decision Making and the Brain*, pp. 209–213. London, UK: Academic Press.

Damasio AR (2010) *Self Comes to Mind. Constructing the Conscious Brain*. New York: Pantheon Books.

Dawkins R (2006) *The Selfish Gene*. New York: Oxford University Press.

de Groot RS (1992) *Functions of Nature: Evaluation of Nature in Environmental Planning, Management and Decision Making*. Groningen, The Netherlands: Wolters-Noordhoff.

de Groot RS and Kumar P (2010) Estimates of monetary values of ecosystem services. In: Kumar P (ed.) *The Economic of Ecosystems and Biodivesity – Ecological and Economic Foundations*, pp. 368–401. London, Washington: Earthscan.

de Groot RS and Ramakrishnan PS (2005) Cultural and amenity services. In: Hassan Rashid, Scholes Robert, and Ash Neville (eds.) *Millenium Ecosystem Assessment. Ecosystems and Human Well-Being: Current State and Trends, Vol. 1, Findings of the Condition and Trends Working Group*, ch. 17. Washington: Island Press.

Descola P and Pálsson G (1996) *Nature and Society: Anthropological Perspectives*. London: Routledge.

Diamond J (2011) *Collapse: How Societies Choose to Fail Or Succeed*. USA: Penguin Group.

Duygu Avci FA and Özkaynak B (2010) Valuation languages in environmental conflicts: How stakeholders oppose or support gold mining at Mount Ida, Turkey. *Ecological Economics* 70: 228–238.

Ellen RF and Fukui K (1996) *Redefining Nature: Ecology, Culture and Domestication*. London, UK: Berg.

Freudenberg N, Klitzman S, and Saegert S (2009) *Urban Health and Society: Interdisciplinary Approaches to Research and Practice*. San Francisco, CA: Jossey-Bass.

Gatzweiler F (2003) *The Changing Nature of Economic Value: Indigenous Forest Garden Values in Kalimanatan, Indonesia*, 1st edn. Vol. 16. Germany: Shaker Verlag GmbH.

Gatzweiler F (2008) Beyond economic efficiency in biodiversity conservation. *Journal of Interdisciplinary Economics* 20: 217–240.

Gintis H (2004) The genetic side of gene-culture coevolution: Internalization of norms and prosocial emotions. *Journal of Economic Behavior & Organization* 53(1): 57–67.

Gowdy J (1997) The value of biodiversity. *Land Economics* 73: 25–41.

Guilford JP (1959) *Personality*. New York, NY: McGraw-Hill.

Hagedorn K (2008) Particular requirement for institutional analysis in nature-related sectors. *European Review of Agricultural Economics* 35(3): 357–384.

Hagedorn K, Arzt K, and Peters U (2002) Institutional arrangements for environmental co-operatives: A conceptual framework. In: Hagedorn K (ed.) *Environmental Cooperation and Institutional Change: Theories and Policies for European Agriculture. New Horizons in Environmental Economics*, pp. 3–25. Cheltenham, UK, and Northampton, MA, USA: Edward Elgar.

Hofstede DG (2001) *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, 2nd edn. London: Sage Publications, Inc..

Jacobs M (1997) Environmental valuation, deliberative democracy and public decision-making. In: Foster J (ed.) *Valuing Nature? Economics, Ethics and Environment*. London: Routledge.

Kumar P (2010) *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations*. London: Earthscan.

Lumsden CJ and Wilson EO (1981) *Genes, Mind, and Culture: The Coevolutionary Process*. Cambridge, MA: Harvard University Press.

Maffi L (2001) *On Biocultural Diversity: Linking Language, Knowledge, and the Environment*. Washington, DC: Smithsonian Institution Press.

Maffi L and Woodley E (2010) *Biocultural Diversity Conservation – Google Books*. London, Washington, DC: Earthscan. Available at: http://books.google.com/booksid=IrCoSHBNPMQC&dq=biocultural%20diversity&source=gbs_similarbooks [Accessed April 4, 2011].

Martinez-Alier J (2001) Mining conflicts, environmental justice, and valuation. *Journal of Hazardous Materials* 86(1–3): 153–170.

Martinez-Alier J (2003) *The Environmentalism of the Poor: A Study of Ecological Conflicts and Valuation New edition*. Cheltenham and Massachusetts: Edward Elgar Publishing Ltd.

Martinez-Alier J, Kallis G, Veuthey S, Walter M, and Temper L (2010) Social metabolism, ecological distribution conflicts, and valuation languages. *Ecological Economics* 70(2): 153–158.

Marx K (1994) *Selected Writings*. Indianapolis, IN: Hackett Publishing.

Millenium Ecosystem Assessment (MA) (2005) In: Hassan Rashid, Scholes Robert, and Ash Neville (eds.) *Ecosystems and Human Well-Being: Current State and Trends, Vol. 1, Findings of the Condition and Trends Working Group*. Washington: Island Press.

Norgaard RB (1994) *Development Betrayed: The End of Progress and a Coevolutionary Revisioning of the Future*. London, UK: Routledge.

Norgaard RB (2010) A coevolutionary interpretation of ecological civilization. In: *Proceedings of the Fourth International Symposium on Ecological Civilization*. Claremont, California.

North DC (1991) Institutions. *The Journal of Economic Perspectives* 5(1): 97–112.

O'Connor M, Funtowicz S, Aguiler-Klink F, Spash CL, and Holland A (1998) *Valuation for Sustainable Environments. The VALSE Project, Full Final Report*. Ispra, Italy: European Commission Joint Research Centre.

Pack SJ (2010) *Aristotle, Adam Smith and Karl Marx: On Some Fundamental Issues in 21st Century Political Economy*. Cheltenham, UK: Edward Elgar Publishing.

Rizzello S (1999) *The Economics of the Mind*. Cheltenham, UK: Edward Elgar.

Robbins P (2004) *Political Ecology: A Critical Introduction*. Oxford, UK: Wiley-Blackwell.

Sagoff M (1988) *The Economy of the Earth: Philosophy, Law, and Environment*. Cambridge: Cambridge University Press.

Scott WR (1995) *Institutions and Organisations*. Newbury Park, CA: Sage.

UNEP (2003) *Cultural Diversity and Biodiversity for Sustainable Development*. Nairobi, Kenya: UNEP.

UNESCO (2002) Declaration on Cultural Diversity http://unesdoc.unesco.org/images/0012/001271/127160m.pdf.

UNESCO (2003) International workshop on the importance of sacred natural sites for biodiversity conservation. In: *Proceedings of the Kunming and Xishuangbanna Biosphere Reserve*. People's Republic of China, 17–20 February. Paris: UNESCO.

UNESCO (2008a) *Proceedings of the Joint Regional Seminar of the Ecotone-SeaBRnet 2007, and the Ninth Conference of the China Biosphere Reserves Network (CBRN): Cultural Diversity – A Foundation for Biodiversity Conservation and Sustainable Development*. Jakarta: UNESCO.

UNESCO (2008b) Links between biological and cultural diversity – concepts, methods and experiences. *Report of an International Workshop*. Paris: UNESCO.

Vatn A (2005) *Institutions and the Environment*. Cheltenham, UK: Edward Elgar Publishing.

Wexler BE (2008) *Brain and Culture: Neurobiology, Ideology, and Social Change*. Cambridge, MA, London, UK: MIT Press.

Wilk RR and Cliggett L (2007) *Economies and Cultures: Foundations of Economic Anthropology*. Boulder, CO: Westview Press.

Williamson OE (2000) The new institutional economics: Taking stock, looking ahead. *Journal of Economic Literature* 38(3): 595–613.

Zografos C and Martínez-Alier J (2009) The politics of landscape value: A case study of wind farm conflict in rural Catalonia. *Environment and Planning A* 41(7): 1726–1744.

Biodiversity and Ecosystem Services

Sandra Quijas, Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Michoacán, Mexico
Patricia Balvanera, Universidad Nacional Autónoma de México, Morelia, México

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity Variety of living genotypes, populations, species, communities, and landscapes; for the case of species it includes the number of different species, their relative abundance, differences in their composition or functional attributes, their spatial patterns, and the complexity of their trophic interactions.

Ecosystem process Transfers of energy and materials given by interactions among abiotic (nonliving) and biotic (living) components of ecosystems.

Ecosystem resilience Ability of an ecosystem to maintain its functioning in the face of changing environmental conditions and disturbances.

Ecosystem service Ecosystem component or process that directly contributes to human well-being as it is consumed, experienced, or it regulates the environmental conditions in which humans live.

Ecosystem service provider Components of populations, communities, trophic groups, landscapes, and habitats that are necessary for the delivery of ecosystem services.

Human well-being People's condition and capacities allowing secure access to food, water, energy, and shelter of adequate quality to meet their needs, and ensure good health as well as social connections.

Meta-analysis Statistical synthesis of the results of separate studies (quantitative research synthesis).

Spatial and temporal scale Spatial or temporal extent within which biodiversity interacts with ecosystem processes, ecosystem processes operate, and ecosystem services are delivered.

Why Worry About the Links Between Biodiversity and Ecosystem Services?

We are increasingly concerned about the impacts of human enterprise on biodiversity and ecosystem functioning, and their ability to continue to deliver the ecosystem services that are essential for human life. To meet humanity's rapidly increasing demand for fuel, water and food, human impacts on ecosystem have also been increasing, accumulating and interacting, leading to irreversible changes in the way ecosystems function. Anthropogenic climate change caused by an increase in the atmospheric concentration of greenhouse gases, large additions of nitrogen and phosphorus activated by humans to soil, water and atmosphere resulting from modern agriculture, and rates of biodiversity loss have all crossed boundaries for sustainable functioning of the Earth system (Rockstrom *et al.*, 2009).

Species are lost today at rates 10–100 times higher than those found in the fossil record. Predictions for extinction rates in the twenty-first century are up to 100 times higher than those found today (Millennium Ecosystem Assessment, 2005; Rockstrom *et al.*, 2009).

Given that ecosystem processes depend on living organisms, and that ecosystem functioning is given by the combination of such ecosystem processes, a key question to be answered is how will biodiversity loss impact the ability of ecosystems to continue meeting human needs.

The concept of ecosystem services was developed to show explicitly the tight connections between human well-being and the adequate provision of ecosystem services. The drive to satisfy human needs is threatening the maintenance of ecosystem functioning, while at the same time human impacts on ecosystems threaten their continued ability to satisfy human needs.

Introduction to the Concepts

New concepts have been coined in the past few years to advance key frontiers of knowledge, and old ones have been used in different contexts to address the multifaceted nature of the relationship between biodiversity and ecosystem services.

Biodiversity

Biodiversity is a term used to encompass all the variability found in living organisms in terrestrial and aquatic systems. This variability can be observed at different levels of organization. The variability among genomes of individuals, among populations of individuals within the same species, among species within a community, or among communities within a landscape can all be analyzed and may be linked to ecosystem services (Diaz *et al.*, 2006). Various mechanisms can generate biodiversity effects on ecosystem functioning. Species can be complementary in the way they contribute to the function and thus higher species richness leads to higher magnitude of the function. Sampling (or selection) effect is given by the increasing probability of including a highly efficient species for a particular function with increasing species richness increasing the probability of choosing this species as a result of random sampling. These two mechanisms are the most frequently called for; in fact, the most common is the combination of both. Facilitation occurs when a given species has a positive effect on the functional contribution of another one. To date, despite the accepted importance of these mechanisms, actual tests of their operation are much less frequent (Cardinale *et al.*, 2011).

The simplest and more direct way to analyze the variability is through the counting of the number of different types of

units, otherwise called richness. The most common use of this term is in reference to species richness, the number of species found in a community. The relative abundance of species is also a part of this variability, given that an ecosystem largely dominated by a single species will function in a different way than one with the same richness but with even abundance of species. Differences in species composition, not only in terms of their taxonomic identity, but also in terms of the differences in the way they respond to environmental conditions and contribute to ecosystem functions, will lead to changes in ecosystem services. The variability in morphological or functional attributes of species is called functional diversity. The distribution patterns of species in space are also relevant to ecosystem functioning, and species can either be evenly distributed across space or rather clumped into patches of different sizes. The structure of a food web, in terms of the number of trophic levels and the number of linkages between species is also relevant to ecosystem functioning and has been referred to as vertical diversity.

A rich literature has developed exploring the links between biodiversity and ecosystem services in many ways. Such a wealth of data and theory has led to syntheses on the role of species richness in ecosystem functioning and provision of ecosystem services that will be discussed below. A large fraction of the explorations of the links between biodiversity and ecosystem services focus on species richness (**Figure 1**). Yet, the roles of species composition or identity, functional diversity, species' relative abundance, species' spatial distribution patterns, and trophic diversity are increasingly being recognized and investigated (**Table 1**); the same is true for the roles of genetic diversity in ecosystem functioning and ecosystem services. To date there is a very large amount of evidence available that indicates that increased producer species richness increases the efficiency by which plants and algae convert these into standing biomass; this is true for temperate grasslands, agricultural systems, marine coastal systems and lakes (Cardinale *et al.*, 2011).

Ecosystem Processes and Ecosystem Functioning

Ecosystem processes are the interactions among abiotic or nonliving and biotic or living components of terrestrial and aquatic ecosystems. Such interactions encompass transfers of energy and materials (Chapin *et al.*, 2000).

One of the central processes in terrestrial and aquatic ecosystems is the transformation of solar energy into biomass through photosynthesis. Primary productivity, the change in primary producer biomass per unit area and time, is the result of photosynthesis, respiration, and decomposition leading to net plant growth. Numerous studies have explored the links between biodiversity and some components of this process, largely as total biomass of primary producers or the rates at which this accumulates. To date there is a very large amount of evidence available that indicates that increased producer species richness increases the efficiency by which plants and algae convert these into standing biomass; this is true for temperate grasslands, agricultural systems, marine coastal systems and lakes (Cardinale *et al.*, 2011).

The biomass stored in primary producers is either consumed or decomposed. The effects of biodiversity have been frequently studied for secondary production (consumption) or decomposition. Such studies have measured biomass consumed or decomposed, amounts of carbon liberated to the atmosphere, as well as respiration rates to assess the activity of decomposers.

The cycling of several key elements – phosphorus, nitrogen, sulfur and carbon – can also be linked to biodiversity. The loss of species of primary producers, particularly plants in

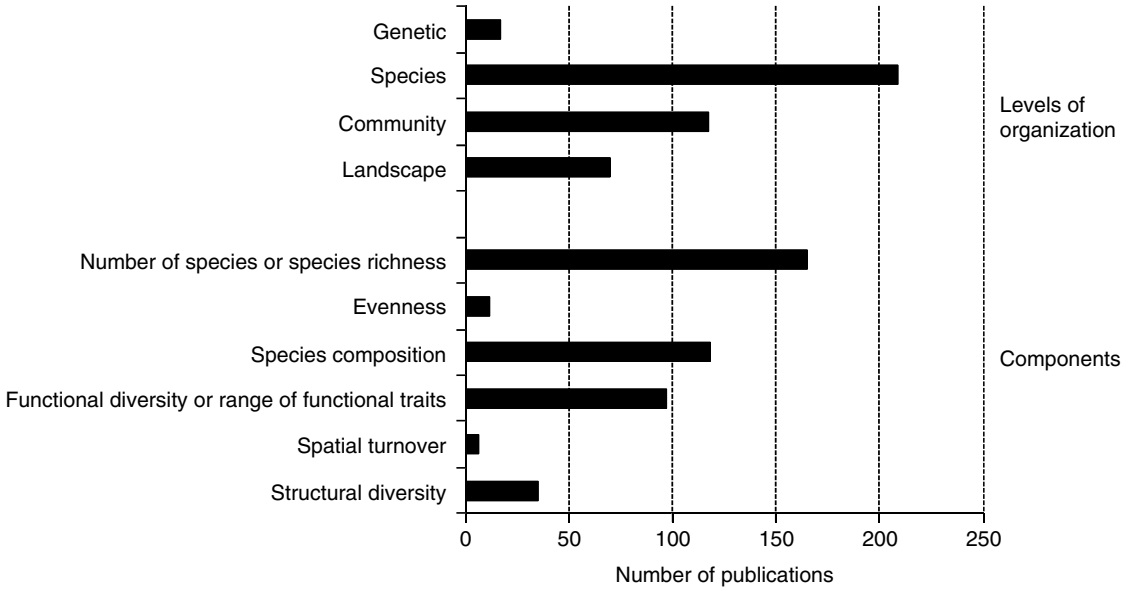


Figure 1 Components and levels of organization of biodiversity that have been related to ecosystem services. Results are derived from a search in ISI Web of Science and Biological Abstracts in May 2011 for all past references using the term in the graph and "(plant or invertebrates or nonvascular plants or microbes or aquatic plants or vertebrates) and ecosystem services".

Table 1 Selected examples of ecosystem services and their ecosystem service providers. Characteristics of the providers refer to the level of organization or the components of diversity involved in service delivery

Ecosystem service provider	*Characteristics of the providers*	*Processes*	*Service*	*Type of ecosystem service*	*Spatial scale*	*References*
Soil invertebrates, soil microorganisms, plant and animal	Populations, species, functional groups	Elemental transformation, Soil structure modification	Soil generation	Supporting	Local	van der Heijden *et al.* (1998) and Barrios (2007)
Vegetation	Populations, species, communities, ecosystems	Primary production	Agricultural food production	Provisioning	Local to landscape	Jackson *et al.* (2009)
Leaf litter and soil invertebrates, soil microorganisms, nitrogen-fixing plants, plant and animal production of waste products	Populations, species, functional groups	Decomposition, Nutrient cycling	Regulation of soil productive potential	Regulating	Local	van der Heijden *et al.* (1998) and Barrios (2007)
Insects, birds, mammals	Populations, species, functional groups	Trophic interactions, Movement of pollen and seeds by animals	Pollination	Regulating	Local to landscape	Kremen *et al.* (2007) and Klein *et al.* (2009)
Microorganisms, invertebrate parasitoids, vertebrate predators	Populations, species, functional groups	Trophic interaction	Regulation of human disease vectors	Regulating	Local to global	Jones *et al.* (2008) and Keesing *et al.* (2010)
Many types of organisms	Populations, species, communities, ecosystems	Ecological–social dynamics	Aesthetic and spiritual inspiration, sense of identity	Cultural	Local to global	Chan *et al.* (2011)

temperate grasslands, has been shown to reduce the efficiency by which they assimilate inorganic resources such as soil nitrate (Cardinale *et al.*, 2011). Yet, the information available to date does not allow generalizing these findings to other elements or to other ecosystems. This is particularly worrying given that the concentrations and cycles of these elements have been substantially altered by human activities over the past two centuries, strongly modifying the functioning of ecosystems and their ability to deliver ecosystem services (Millennium Ecosystem Assessment, 2005; Rockstrom *et al.*, 2009).

We know little about the role of biodiversity in the different components of the water cycle. Plant diversity, in the form of functional and structural diversity may play a key role in evapotranspiration, surface and underground runoff (Chapin *et al.*, 2002; Diaz *et al.*, 2006). The diversity of soil organisms (such as earthworms, micro-arthropods, and microbes) is related to soil physical properties such as the presence of macropores to aid infiltration (Barrios, 2007). Yet, not enough experimental nor observational data is consistently available to confirm the role of biodiversity on these processes.

Ecosystem functioning encompasses many of the abovementioned processes and the interactions among them. Functioning includes the ability of the ecosystem to cope with natural and anthropogenic disturbance by converging around a particular state rather than changing dramatically into a very different condition for the components and processes that make them up. The latter is called ecosystem resilience.

The potential links between biodiversity and multiple ecosystem processes or ecosystem functioning are often referred to in scientific literature. Yet, only a fraction of such references correspond to actual data on such links.

Ecosystem Services

Ecosystem services are the links between ecosystems and human societies. In the broadest sense they are the benefits societies obtain from ecosystems (Millennium Ecosystem Assessment, 2005). More specifically, services are produced by living or nonliving components of ecosystems, through the conditions and processes in ecosystems. They contribute to human well-being in different ways: Services can be directly consumed, as in the case of water or food. Services can be experienced, as in the case of the scenic beauty or the sense of awe that a waterfall, a mountain covered by vegetation or a monarch butterfly can instill in us. Services also contribute to the fundamental environmental conditions for human life, such as the regulation of relatively stable climatic conditions and protection against extreme events.

Ecosystem services have also been defined as those components and processes of ecosystems that contribute directly to human well-being (Luck *et al.*, 2009). By emphasizing the direct nature of such connections an effort is being made to distinguish ecosystem processes from ecosystem services.

There is a fine line between ecosystem processes and ecosystem services. Some authors have considered ecosystem processes as those ecosystem services that support the delivery of all other types of services. The advantage of this position is that it is possible to encompass both direct benefits to human societies as well as ecosystem functions and the origin and maintenance of biodiversity within the term ecosystem services. This approach can be very useful when communicating with decision makers.

Yet, in order to quantify the amount of services, a clear distinction between supply, service and value can be useful. Supply is the contribution of ecosystem processes and components to a service that can potentially benefit societies. Service is the actual benefit societies obtain from the service, which can include water consumption or the number of people that benefit from regulation of flooding. Value reflects how societies consider the relative importance of a service; it can be monetary, as an expression of markets or preferences or embedded into cultural perspectives (Chan *et al.*, 2011). In this section we will dissect ecosystem processes from ecosystem services in this way.

The key message is that services or benefits to human societies depend on ecosystems and their functioning. In turn, such functioning is dependent to some degree on biodiversity.

The Millennium Ecosystem Assessment (2005) classifies services into four types: supporting, provisioning, regulating, and cultural services. The authors do not include supporting services, as they consider them to be ecosystem processes that indirectly benefit societies by supporting one of the other three types of services.

Provisioning Services

Provisioning services are the tangible resources or goods that people obtain from ecosystems. These are finite, can be renewable, and can be directly consumed, appropriated, and traded.

Food provision is among the most important services. Food can be obtained from agricultural activities, aquaculture, hunting, gathering, and fisheries. Other key provisioning services include fodder, wood, fibers, biofuels, and a diversity of abiotic and biotic resources derived from terrestrial and aquatic systems.

Regulating Services

Regulating services result from the contribution of multiple ecosystem processes to ecosystem functioning, specifically to the regulation of the conditions where humans live and make a living. Such regulation determines both the average and the variance in such conditions.

Regulating services include the regulation of local or global climate, of disease vectors and incidence, of pests, crop pollination, soil fertility, and soil erosion. They also regulate the amount, temporality, and quality of water provision, as well as the impact of severe weather conditions on ecosystems and people.

Cultural Services

Cultural services are ecosystems' contribution to the nonmaterial benefits that arise from the interaction between people and ecosystems. These benefits include a range of capabilities and experiences (Chan *et al.*, 2011).

Cultural ecosystem services include the sense of place or identity linked to a particular ecosystem and its management. They also include activities such as recreation and work. The existence and bequest value of a site or a species for future

generations, or the sense of awe and spiritual or esthetic inspiration are also considered cultural services.

Ecosystem Service Provider

An ecosystem service provider is a biotic component of an ecosystem that is necessary for the delivery of an ecosystem service. Ecosystem providers can include populations, communities, trophic groups, landscapes, and habitats (Kremen, 2005; Luck *et al.*, 2009).

The concept of the ecosystem service provider was developed to identify the component(s) of the ecosystem that should be managed in order to ensure the sustained delivery of each service. Vascular and nonvascular plants, terrestrial and aquatic invertebrates, microbes, and vertebrates have been explicitly recognized as ecosystem service providers (see **Table 1**). Genetic diversity, species diversity, species richness, functional diversity, and vertical diversity have also been identified as such.

Human Well-Being

Human well-being is a complex concept that is not universal but rather context dependent. The different components of such well-being as well as their relative importance depend on the cultural, geographic, and historic development of societies.

The Millennium Ecosystem Assessment (Millennium Ecosystem Assessment, 2005) identified food, water and other materials, security, good health, and good social relations as the key components of well-being. A recent report by another interdisciplinary international team (Stiglitz *et al.*, 2010) emphasizes the need to simultaneously consider the following dimensions: material living standards (income, consumption, and wealth), health, education, personal activities including work, political voice and governance, social connections and relationships, environment (present and future conditions), and insecurity (of an economic as well as physical nature). To assess them both people's objective conditions and capabilities should be taken into account.

Human well-being is directly linked to the biodiversity that surrounds the many rural communities of the world, from which they obtain food, medicine, materials, and inspiration. The role that biodiversity plays in ecosystem processes, the services they provide and thus human well-being will be discussed below.

The Shape of the Relationship Between Species Diversity and Ecosystem Functions and Services

The shape of the relationship between the species diversity and the magnitude of ecosystem function or service is quite relevant for management. If the shape is asymptotic this means that loosing the few first species may have little effect on the function, but the magnitude of the impacts of losing further species can increase in non-linear ways leading to a dramatic decline in the function or service once enough species have been lost (**Figure 2**). Experimental studies much more frequently find such a pattern, rather than a linear

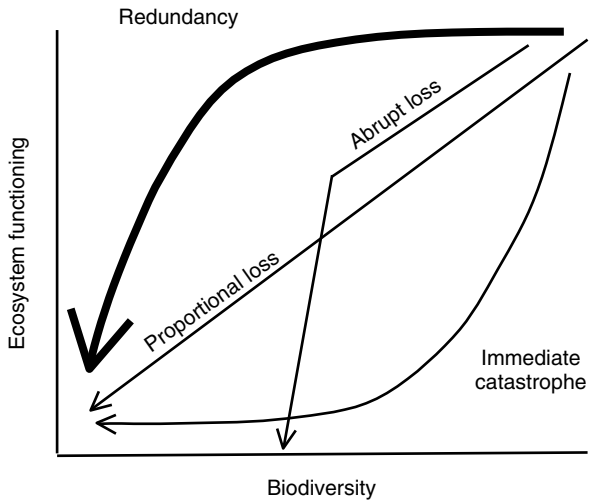


Figure 2 Possible ecological consequences of diversity loss. (1) The redundancy response predicts that initial losses of diversity will be accompanied by minimal change in the ecosystem process because some fraction of species are redundant for that process; however, at some point, loss of species leads to rapid declines in ecological function. (2) Ecosystem function declines proportionally to species loss. (3) An abrupt loss in ecosystem functioning is given by the loss of a keystone species or last member of a key functional group, or the addition of a new species trait. (4) Even minimal species loss leads to an immediate catastrophe and large declines in the functioning of ecosystems. A thicker line for the redundancy model indicates that it is the most frequently found pattern (modified from Letourneau DK, Armbrecht I, Rivera BS, *et al.* (2011) Does plant diversity benefit agroecosystems? A synthetic review. *Ecological Applications* 21: 9–21).

relationship, or one where losing the first few species has the most dramatic effects (Cardinale *et al.*, 2011). Similarly, observational studies have also found such an asymptotic pattern (Balvanera *et al.*, 2005). This means that losing species in already depauperate systems can have dramatic cons'equences for ecosystem functioning and ecosystem services.

Spatial and Temporal Scales

Spatial or temporal scale refers to the extent of the area or the duration of time. A related but different issue is the resolution, or grain, that refers to the smallest detectable event or property at a particular scale. Given that distribution patterns of biodiversity vary at different spatial and temporal scales, given that ecosystem processes operate at different scales, and given that services are delivered at different scales it is essential to clearly define the scale of analysis.

Spatial and temporal scales are continuous and dissected in a variety of ways, but for analytical purposes the authors have arbitrarily defined three contrasting scales here. The local spatial scale was defined here for areas up to 1 ha. This scale encompasses the area used for many of the local biodiversity assessments in terrestrial habitats (e.g., 0.1 ha assessment, or 1 ha assessment). Also, the experimental manipulations of biodiversity have used experimental units within this area range; for the case of those involving primary producers

(the most frequent ones) the median size of the experimental unit is 0.003 ha, and the largest ones have an area of 0.5 ha.

The landscape scale was defined as that ranging from 1 to 1000 ha. This range encompasses the area included into an individual piece of land to the whole area one can see from an elevated point under low to a very high visibility. This area encompasses well-known gradients in environmental conditions, a combination of individual lots of lands with natural, seminatural and anthropogenic conditions (e.g., agriculture and agroforestry), and can include one or various rural settlements. The regional scale to global scale was defined as that which includes 1000 ha and more. This range of areas can include a few rural settlements, one or many municipalities or urban settlements, all the way to whole states and countries. At this spatial scale use of remotely sensed data or spatially explicit models are needed to assess biodiversity distribution patterns as well as the magnitude of ecosystem processes and services.

The Complex Linkages between Biodiversity and Ecosystem Services: Selected Examples

What would happen if we lost most of the biodiversity? Would the delivery of ecosystem services be dramatically impaired? Do all ecosystem services show the same sensitivity to the loss of biodiversity? How universal is the asymptotic relationship between biodiversity and service provision?

All these are very complex and profound questions that have not been addressed thoroughly and systematically for all ecosystem services and at all spatial scales. The authors focus below on four very important and intensively studied ecosystem services. The authors show that they all depend on biodiversity, though in very different ways. The authors also show how they have been analyzed in very different ways.

Agricultural Food Production

Food production derived from agriculture depends on a wide diversity of crop species found across the planet. As many as 100 different crops used for human food are registered in global agricultural databases, and many more are locally grown and consumed. Yet, 63% of food production (as measured on a dry biomass basis) relies on only five species: sugar cane, maize, wheat, rice and potatoes (FAO, 2011).

Understanding the complex links between food production derived from agriculture and biodiversity is as relevant as ever, as feeding an expected nine billion people poses a paramount challenge. Most food production today is derived from highly intensive agricultural systems with low diversity. In fact, the sites with the highest agricultural yields (e.g., sites in China with 1.3 metric tons per hectare by year (MT); (FAO, 2011)) are those where most plants correspond to only one species and to one or a few genetic combinations (e.g., wheat and rice with 0.32 MT in China or sugar cane with 0.78 MT in Brazil; (FAO, 2011)). Also, the expansion and intensification of agricultural systems are the major driver of biodiversity loss of terrestrial habitats. On the other hand, the ability of diverse cropping systems, diverse agricultural landscapes, and biodiversity friendly agricultural systems (e.g., organic) to meet global food needs is being debated.

The diversity of plant crops, as well as that of soil microorganisms and vertebrates, weeds, herbivores, and carnivores that interact with them is all relevant to the provision of this service. Also, diversity at different trophic levels, ranging from genetic diversity within populations to the diversity of communities within landscapes is relevant.

Agrobiodiversity, that is the variety and variability of living organisms that contribute to food and agriculture as well as the knowledge associated with them, can enhance agricultural food production. A recent meta-analysis of 552 experiments (Letourneau *et al.*, 2011) showed a general trend for positive effects of biodiversity. Diverse cropping schemes (polycultures) have been shown to have a significantly lower abundance of pests (herbivores), higher abundance of natural enemies of such pests (predators and parasites), decreased damage by pests, though at the cost of a small but significant reduction in crop yield (**Figure 3**).

Genetic diversity within populations, population diversity within crop species and community diversity within agricultural landscapes contribute to food production derived from agriculture. Farmers, particularly in developing countries, use a wide variety of traditional crop local varieties, or landraces, that still dominate production in many rural landscapes. Genetic and population diversity confers resilience in the face of abiotic (e.g., drought) and biotic (e.g., pests) stress, leading to more stable yields and higher adaptation to stressful conditions; various study cases support this well, though a systematic broader assessment is still needed to test the generality of such statements. Also, diversity of communities within agricultural landscapes, including low to high diversity agroecosystems and a mosaic of well-connected successional (recovering after agricultural use) and conserved patches of habitat provide increased resilience (Jackson *et al.*, 2009).

Regulation of Soil Productive Potential

Soils provide physical support, as well as water and nutrients needed for plants to grow, all of which can be encompassed into the regulation of soil productive potential. Services such as food production from agricultural systems, firewood, timber and biofuel production, and indirectly the regulation of soil erosion depend on the adequate regulation of soil conditions.

A tight but much more complex and less understood link exists between biodiversity and soil productive potential. The maintenance of the function of a large array of different types of organisms within soils is critical for maintaining agricultural, pastoral and forestry yield over the long term. Managing for increasing short term yield, such as the addition of inorganic fertilizers, may jeopardize the long term maintenance of biodiversity and this service. Yet, management is complex as it can hardly be targeted at specific groups or functions that we understand little.

Many different types of organisms (**Table 2**) participate in the regulation of soil productive potential (Barrios, 2007). Plants establish symbiotic associations with microorganisms that enhance the availability of otherwise limiting nutrients;

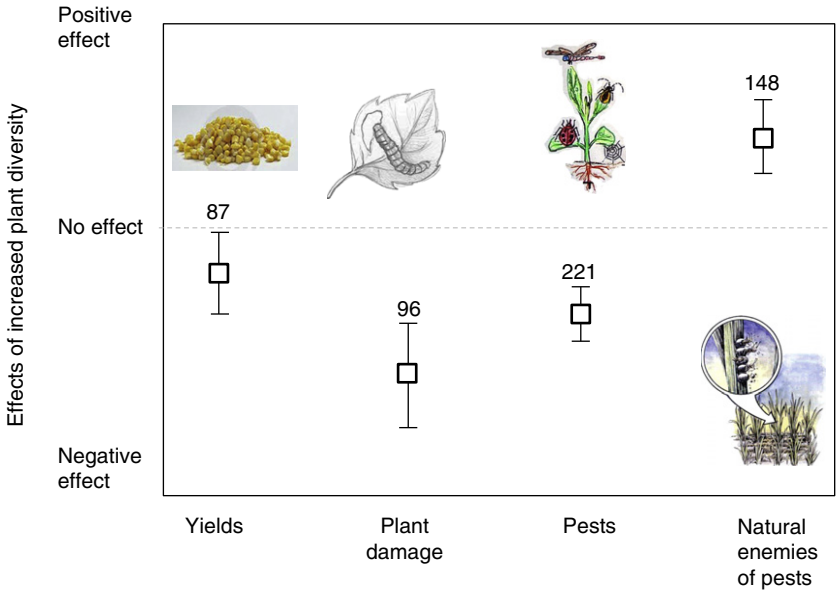


Figure 3 Effects of increased diversity associated to more diverse cropping schemes (polycultures) relative to less diverse ones or crop monocultures on various ecosystem services. The mean of the proportional effects of diversity and the 95% confidence intervals values derived from the meta-analysis are shown for each response variable. Significant effects correspond to average and confidence interval values that do not overlap with the no-effect line. Numbers above each result indicate the number of independent measurements found in the literature used for the analyses (modified from Winfree R, Aguilar R, Vazquez DP, LeBuhn G, and Aizen MA (2009) A metaanalysis of bees' responses to anthropogenic disturbance. *Ecology* 90: 2068–2076).

Table 2 Soil organisms and key functional groups of soil biota, soil processes they influence and ecosystem services they provide

Soil organisms	*Key functional groups*	*Selected soil processes*	*Selected ecosystem services*	*Service classification*
N-fixing organisms, mycorrhiza	Microsymbionts	Microsymbiosis	Nutrient uptake	Supporting
Cellulose and lignin degraders	Decomposers	Decomposition	Nutrient cycling	Supporting
Nitrifiers, denitrifiers	Elemental transformers	Elemental transformation	Nutrient cycling	Supporting
Earthworms, termites	Ecosystem engineers	Soil structure modification	Regulation of soil erosion, C sequestration, water flow and storage	Regulating
White grubs, plant–parasitic nematodes, root–rots	Soil-borne pest and disease	Herbivory, parasitism	Biological control of pests and disease	Regulating
Grazers	Microregulators	Grazing of decomposers	Nutrient cycling	Supporting

Source: Modified from Barrios E (2007) Soil biota, ecosystem services and land productivity. *Ecological Economics* 64: 269–285.

soil bacteria (e.g., Rhizobium) in symbiosis with leguminous plants fix atmospheric nitrogen into the soil, whereas soil-borne fungi in symbiosis with roots of many vascular plants, particularly the so-called arbuscular mycorrhizal fungi enhance phosphorus acquisition through the branching filamentous structure of fungi (hyphae). Decomposition of organic materials into simpler molecules is given by small invertebrates that break down dead tissues; by bacteria, fungi, protozoa, and invertebrates that chemically degrade dead organic matter further with enzymes; and by microorganisms that generate soluble and volatile organic and inorganic compounds. Earthworms, ants, and termites modify soil structure by removing soil and creating structures by their activity; root–rots, plant–parasitic nematodes, and scarab beetle larvae are responsible for soil-borne diseases, whereas soil grazer, predators, or parasites drive the local food web (**Table 2**).

Richness of various types of organisms is linked to some of the processes related to the regulation of soil productive potential. Recent meta-analyses (Balvanera *et al.*, 2006; Hoeksema *et al.*, 2010; Quijas *et al.*, 2010) have shown that higher plant species richness contributes to significantly higher plant root biomass and decomposer activity, higher mycorrhizal richness leads to higher plant nutrient

concentration, and higher decomposer richness to higher decomposer activity. Instead, higher plant species richness contributes to lower decomposer biomass and soil nutrient supply.

Species composition and particularly functional composition and diversity are expected to have a paramount importance for the regulation of soil productive potential given the wide range of organism involved in the different processes. Growth form composition (e.g., legume and hemiparasitic plants) and growth form diversity (e.g., life form, root depth, photosynthetic capacity, and nitrogen-fixing capacity) contribute to higher nutrient uptake through increased functional diversity (de Bello *et al.*, 2010).

The links between biodiversity and the regulation of soil productive potential occur at local scales (Barrios, 2007). Soils are highly heterogeneous in terms of their physical, chemical, and microclimatic characteristics, as well as in terms of the spatial (horizontal and vertical) distribution of the soil biota, at scales ranging from microns to meters. Soil biota is concentrated in hot-spots of activity that are mostly associated with the availability of carbon substrates and water availability.

A better understanding of the spatial distribution patterns of soil biota at a wide range of spatial scales and the consequent impacts on soil processes is needed to allow for managing soils and maintaining the range of ecosystem services they provide.

Pollination

Mobile organisms pollinate many important crops and other plant species of commercial and social importance. In the absence of the native organisms that have contributed to this service, domestic European bees (*Apis mellifera*) have been introduced; in a few cases, pollination has been performed by hand by humans.

Pollination, a regulation service, provides benefits to human populations by increasing the production of those commercial and subsistence crops, fiber, fodder, wood, and non-timber forest products that require a mobile organism to transport male pollen grains into feminine floral structures for seed production. Pollination is also a supporting service, or ecosystem process, that enables the reproduction of many wild plants.

Pollination is important for 35% of the global crop production. Between 60% and 90% of wild plants (depending on the type of biome) require pollination; a larger proportion of the species benefit from flower visitation by mobile organisms.

Mobile organisms that visit flowers provide pollination services at the local scale at which such interaction occurs. Yet, these organisms move freely within and among habitats, and thus their individual, population, and community dynamics depend on the spatial distribution of resources at landscape or larger spatial scales (Kremen *et al.*, 2007).

Pollination is provided by wild animals. Bees are the most important pollinators. Yet, butterflies, moths, flies, beetles, wasps, birds, bats, and other invertebrates, and mammals are important for this service. Commercially managed bees are very important for the provision of the service in agricultural contexts; *Apis mellifera* is the most important pollination service provider globally to farmers.

Rich pollinator assemblages enhance pollination in different ways. Richer species are often the most abundant and show the highest contribution to pollination. Different species pollinate different species within a diverse agroecosystem. Also, different species visit preferentially different parts of the plant and thus contribute to a more thorough pollination. Different species respond differently to changes in climatic conditions and resource availability. They are also differently sensitive to the impacts of agricultural intensifications. Thus richer pollinator communities show higher resilience (Klein *et al.*, 2009).

Landscape structure and diversity determine the spatial and temporal availability of food, and sites most suitable for nesting, hibernating, and mating. Thus landscape quality is determinant for the maintenance of pollinators and the services they provide.

Wild and managed pollinators have suffered dramatic declines in recent years. Severe declines in *A. mellifera* populations, those of the major groups of pollinators and of the plants that they pollinate have been documented, despite the pervading lack of long term, globally distributed data (though the data records for *A. mellifera* begin in 1908). Land use change and changes in landscape structure are the major drivers of this change. Also, agricultural intensification, particularly the use of pesticides, that directly affect pollinator populations, and that of herbicides, fertilizers, irrigation, tillage, or fire, that affects the availability and spatial distribution of resources as well as environmental conditions, change the total abundance, composition, and richness of pollinator assemblages (Winfree *et al.*, 2009). Yet, the reduced amount of independent experiments or observations has hindered testing for the generality of the impacts of the above factors. To date, only habitat change has shown consistent negative significant effects on bee richness and abundance (**Figure 4**).

The above results show that pollinator decline is undeniable and that decline in the abundance and richness of pollinators has negative consequences for pollination. Yet, the consequent threats to global agricultural food production are still under debate.

Regulation of Human Infectious Disease

It is not possible to assert that the more diversity the better the regulation of human infectious diseases. In fact, evidence to date shows that the maintenance of some elements of biodiversity can cause disease, while the maintenance of other elements is key to regulating both the incidence and prevalence of disease.

Regulation of the appearance and transmission of infectious diseases to human populations is a regulating service. The service is provided by several processes such as the regulation of population sizes of vectors, hosts and pathogens, the regulation of interactions between them and human populations, and the regulation of their natural enemies.

Infectious diseases result from the interaction of pathogens or parasites, vectors, hosts, and the environment. Frequently many species are involved, including many hosts, many vectors, and a wide range of species with which they interact. Pathogens include helminthes, fungi, protozoa, viruses or prions, and bacteria or rickettsiae; bacteria and rickettsiae

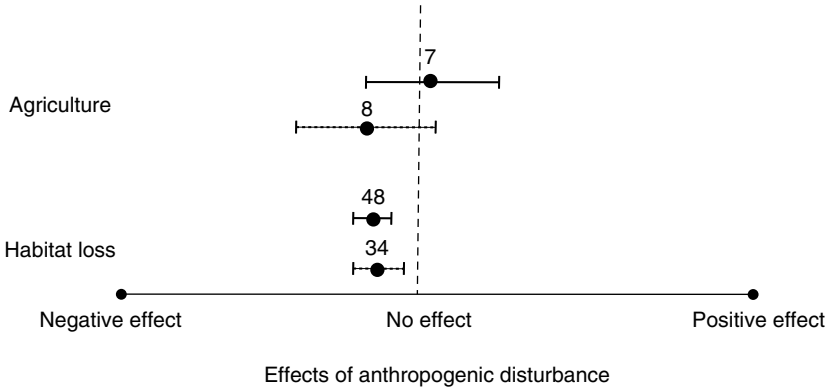


Figure 4 Effects of different anthropogenic disturbance factors on (a) bee abundance (solid line) and (b) bee richness (dashed line). Mean effect and 95% confidence intervals are shown for each type of disturbance. Significant effects correspond to average and confidence interval values that do not overlap with the no-effect line. Numbers above each result indicate the number of independent measurements found in the literature used for the analyses (modified from Balvanera P, Pfisterer AB, Buchmann N, *et al.* (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9: 1146–1156, clear open symbol, and Quijas S, Schmid B, and Balvanera P (2010) Plant diversity enhances provision of ecosystem services: A new synthesis. *Basic Applied and Ecology* 11: 582–593, solid symbol).

Table 3 Diseases whose transmission is affected by habitat or biodiversity loss and the associated mechanism. S indicates a suggested mechanism, while D denotes that mechanisms have been demonstrated in diverse studies

Disease	*Mechanism*	*Main organisms that affects the disease*
Coral diseases	Changes in host/vector abundance [D]	Coral reefs
Malaria	Changes in host/vector abundance [D]	Humans (mammals, vertebrates)
Helminthic parasite of fish	Changes in host/vector abundance [S]	Aquatic vertebrates (fishes and sharks)
Puccinia rust infection of ryegrass	Changes in host/vector abundance [S]	Plants
Amphibian limb malformation	Changes in host/vector/parasite behavior [D]	Amphibian (salamanders, frogs and toads)
Bacteriophage of Pseudomonas syringae	Changes in host/vector/parasite behavior [D]	Plants
Fungal disease of Daphnia	Changes in host/vector/parasite behavior [D]	Plank tonic crustaceans
Schistosomiasis	Changes in host/vector/parasite behavior [D]	Humans (invertebrates, vertebrates)
Trematode diseases of snails and birds	Changes in host/vector/parasite behavior [D]	Invertebrates, vertebrates
Hantavirus disease	Changes in host/vector abundance and host/vector/parasite behavior [D]	Humans (mammals, vertebrates)
Lyme disease	Changes in host/vector abundance and host/vector/parasite behavior [D]	Humans (mammals, vertebrates)
West Nile fever	Changes in host/vector abundances [S], host/vector/parasite behavior [S]	Humans (birds, vertebrates)

Source: Modified from Keesing F, Belden LK, Daszak P, *et al.* (2010) Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468: 647–652.

caused nearly 55% of infectious diseases. Zoonotic pathogens are those that complete their whole life cycle within animal species but can also infect humans, through links with other vertebrates. More than 70% of human diseases are caused by such pathogens associated to wild fauna, and their incidence is increasing with time (Jones *et al.*, 2008).

The nature of the links between biodiversity and the regulation of infectious diseases is quite complex. On one hand, higher biodiversity can be associated with a larger diversity of new pathogens; on the other hand, higher biodiversity has been associated with reduced transmission probability of already established pathogens and with the reduced establishment of new ones. Yet most evidence points to the latter. Mathematical models, individual studies, and recent syntheses have shown that a higher diversity of hosts is associated with less frequent and slower disease transmission, and thus reduced disease incidence (Keesing *et al.*, 2010).

At local to regional scales (Keesing *et al.*, 2010), biodiversity loss affects the emergence and transmission of infectious disease through changes in the abundance of hosts or vectors, changes in the behavior of hosts, vectors, or parasites, and changes in the condition of hosts or parasites (**Table 3**). Anthropogenic impacts on ecosystems have led to impacts on the regulation of infectious diseases through losses of natural enemies of vectors, increases in population densities of host species, changes in genetic diversity of vectors or pathogens (such as antibiotic resistant bacteria or pesticide resistant mosquitoes), as well as increased frequency of soil, water, or airborne transmission of pathogens.

At global scales, many social and ecological factors are related to the increasing incidence of infectious diseases (that peaked in the 1980s). The highest incidence occurs at high latitudes (30–60° N and 30–40° S), in areas with the highest population density. Yet, host species richness is a significant

predictor of the appearance of infectious diseases associated to zoonotic transmissions from wild species (Jones *et al.*, 2008).

Overall, reduced species richness has been related to increased disease incidence. Yet, further understanding of the role played by biodiversity in the regulation of infectious diseases is needed to better assess the risks of biodiversity loss on human health.

Analyzing the Links Between Biodiversity and Ecosystem Services at Different Spatial Scales

The links between biodiversity and ecosystem services have been studied at a wide range of spatial scales. The approaches used, hypotheses tested, and the types of results change clearly across these scales.

Local Scale

A large amount of work has been done to explore the links between biodiversity and ecosystem functioning at local scales, some of which has been used to assess the effects of biodiversity on ecosystem services. Experimental studies, observational studies, and modeling efforts have led to recent syntheses of this vast literature.

Qualitative syntheses have addressed the role played by various types of diversity. The first qualitative synthesis (Hooper *et al.*, 2005) emphasized the strong influence that species' functional characteristics have on ecosystem processes, independently of their relative abundance. Another one (Diaz *et al.*, 2006) suggested multiple potential links between genetic diversity, species richness, functional composition, and landscape structure on a range of ecosystem services. Yet the generality of the patterns was debated.

Quantitative meta-analyses have assessed the generality of the effects of species richness on ecosystem functioning derived from experimental studies. Higher plant species richness leads consistently and significantly to increased standing biomass of primary producers and increased efficiency in the assimilation of inorganic resources by these primary producers in terrestrial and aquatic systems. Data for other functions and groups is still insufficient to draw general conclusions (Cardinale *et al.*, 2011).

Experimental studies have also been used to assess the role of biodiversity in the provision of ecosystem services and synthesized using meta-analysis (**Figure 5**). Higher plant species richness leads to higher provision of plant products, greater pest regulation (lower primary consumer biomass), greater erosion control, and higher regulation of invasive species and pathogens. Also, increased micorrhizae richness led to higher provision of plant products and higher erosion control. Yet, higher species richness led to lower low or nonsignificantly different soil fertility regulation (Balvanera *et al.*, 2006; Quijas *et al.*, 2010).

The effects of species richness on ecosystem stability, measured as the inverse of the variance of an ecosystem process through time, have been assessed in a few long-term experiments. Higher species richness led to higher stability in some cases and to nonsignificant changes in stability in others (Balvanera *et al.*, 2006). Higher plant species richness consistently contributed to higher security (less variance) in the provision of plant products (Quijas *et al.*, 2010).

The above results are derived from extrapolations from links between ecosystem processes actually measured in experiments and ecosystem services that are potentially associated with them. Focusing on individual ecosystem service providers and making explicit these links has helped clarify these connections (Quijas *et al.*, 2010). Yet, direct measures of services from biodiversity experiments are needed.

Observational studies have in many cases measured ecosystem services directly. Recent syntheses for the case of agricultural food production, pollination, and regulation of infectious diseases have been summarized above. A meta-analysis of the effect of richness in tree plantations has shown that mixing tree species generally increases plantation growth rate.

Experimental and observational studies seldom address cultural ecosystem services. Yet, one experiment in temperate grasslands showed that people perceive changes in species richness, composition, and evenness. People were most attracted to the most diverse arrays (Lindemann-Matthies *et al.*, 2010).

Functional attributes of species and communities have strong influence on ecosystem service provision through their effect on ecosystem processes. A recent synthesis has shown that quantitative associations between species' traits and ecosystem services are available to date for a range of organisms, with a large proportion of data for plants and soil invertebrates. Key traits such as body size, canopy size, and specific leaf area or root architecture are related to a suite of ecosystem services (de Bello *et al.*, 2010).

Individual species have differential contributions to a particular function. The functional structure approach can be used to assess the relative contribution of each individual species to an additive ecosystem function, in the same way that rank-abundance curves assess the relative contribution of each species to the total abundance. Pollination in watermelon organic farms in California found within conserved matrices is provided by 11 bee species; eight of these species contribute with quite similar amounts of pollen grains. Instead, in organic farms found within agricultural matrices only four species pollinate watermelon; only one species contributes to more than 90% of the pollen grains deposited. Also, the relationship between species' relative abundance and relative functionality can be used to identify abundant species that are functionally important, or rare species that contribute little to the function, or species whose relative contribution to the function is significantly different than that expected by their relative abundance. Carbon stocks in the tropical rain forest of Chiapas, Mexico, are provided by 169 tree species; one abundant species with dense wood and large trunks, *Dialium guianense*, contributes to more than 10% of the abundance and of the carbon stocks; a fast growing species with thin stems, the pioneer species *Cecropia obtusifolia*, contributes less to carbon than expected from its relative abundance; various species of the genus *Ficus* contribute to less than 1% of carbon stocks and to less than 0.1% of the abundance, and yet

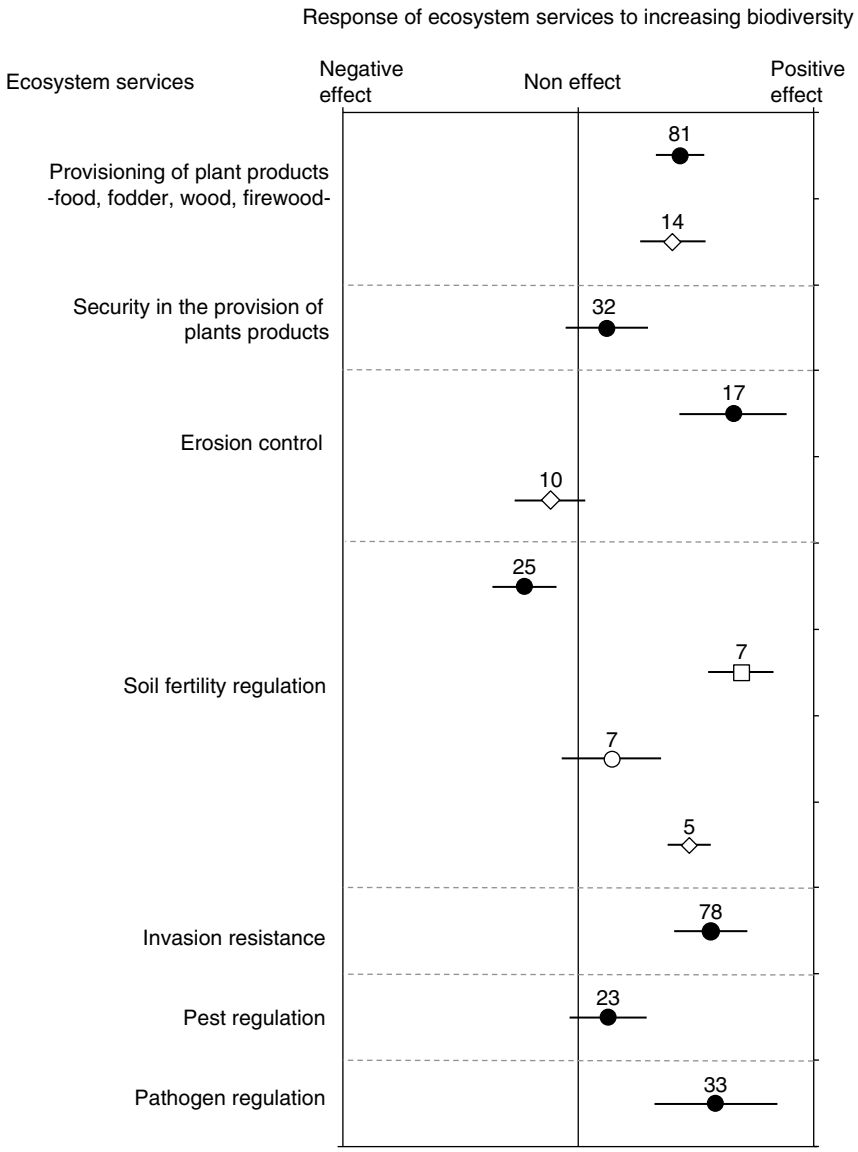


Figure 5 Effects of experimental manipulations of biodiversity on ecosystem services. Symbols show differential effects of trophic level manipulated: circle, primary producers; diamond, mycorrhiza; square, decomposer. Numbers above each result indicate the number of independent measurements found in the literature used for the analyses (modified from Cardinale BJ, Matulich KL, Hooper DU, *et al.* (2011) The functional role of producer diversity in ecosystems. *American Journal of Botany* 98: 572–592, and Chapin FS, Zavaleta ES, Eviner VT, *et al.* (2000) Consequences of changing biodiversity. *Nature* 405: 234–242).

have been considered keystone species due to their critical interactions with many bird and mammal species (Balvanera *et al.*, 2005).

Analyses of biodiversity effects on functions or services have mostly focused on a single function. Yet, a key question is how will species loss affect overall ecosystem functioning. Conceptual models and experimental studies have shown that increasing species richness increases the maintenance of multiple ecosystem processes; there are many species that contribute significantly to various functions (Gamfeldt *et al.*, 2008).

Evidence from local-scale experimental and observational studies points to positive effects of biodiversity on service provision, and to the paramount role played by species composition or functional attributes. Key functional attributes such as leaf phenology, leaf distribution over stems and canopy size and architecture have been shown to be important in determining plant species' contribution to climate regulation (de Bello *et al.*, 2010).

Overall, while the local-scale experimental approaches may provide key insights on the way biodiversity relates to functioning, their potential applications to services is somehow limited. To date, only a small range of services have been assessed (e. g. soil fertility, fodder provision, biocontrol). Also, the realism of local scale biodiversity experimental manipulations and their potential applications to management relevant conditions, including larger spatial scales and human interventions, has been debated.

Landscape Scale

Analyzing the links between biodiversity and ecosystem services at landscape scales offers very different perspectives than those at the local scale. Although food production, or soil fertility regulation are provided at local spatial scales, many ecosystem services such as pollination, pest regulation, regulation of landslides or coastal protection depend on ecosystem processes and ecosystem service providers that occur at landscape scales. Mobile organisms that move freely through the landscape provide some of the services; the attributes of the landscape in terms of the types of land-use change found and their spatial arrangement are relevant at this scale; complex trophic cascades that encompass organisms with home ranges at a variety of spatial scales are associated with service provision. Also, societies and ecosystems interact at landscape scales through land-use decisions that often lead to land-cover changes or to agricultural intensification.

Natural experiments, those in which different diversity levels are found due to natural or anthropic causes, replace experimental manipulations of species richness at this scale. As a consequence, correlations between diversity and services can be observed, but can be the result of direct effects of changes in abiotic conditions on diversity, direct effects on the processes or services, or mediate the links between them. Testing for the underlying mechanisms is more complex. Yet, manipulations of some groups of species are possible at this scale.

Experimental manipulations of individual groups of species have shown positive correlations between diversity and ecosystem services. A meta-analysis of experimental bird exclosures in diverse agroecosystems in Central America showed that bird functional species richness, richness of functional groups and functional diversity were positively correlated with pest regulation services, measured as removal of arthropods (Philpott *et al.*, 2009). The contribution of all species and functional groups were not equal, and one species and two of the functional groups contributed disproportionately to the service. Yet, direct links to a service would require focusing on crop consumer arthropods.

Species loss at one particular trophic level can have cascading effects on many other levels, and thus have profound consequences on ecosystem services. Mathematical models have shown that many species of decomposers can be lost before observing a dramatic reduction in service provision; yet, the fraction of species lost that leads to reduction in ecosystem services increases as upper trophic levels are considered (plants, herbivores, and predators). Moreover, rapid declines in service provision are predicted as predators are lost (**Figure 6**). Given that the upper trophic levels are more vulnerable to land use change, as they have less functional redundancy, fewer individuals per species and are characterized by larger home ranges, rapid declines in ecosystem services are to be expected (Raffaelli, 2004; Dobson *et al.*, 2006).

Yet, the responses of ecosystem processes and services associated with species loss vary between different trophic levels. A meta-analysis of removal experiments in agroecosystems (Mooney *et al.*, 2010) showed that excluding vertebrate insectivores such as birds, bats, and lizards increased the abundance of predaceous and herbivorous arthropods, that is both those species that are considered pests in agroecosystems (herbivores) and their natural enemies (predators); in fact, the strength of the effect of vertebrate insectivores on both groups was the same. Exclusions also showed reduced plant damage and increased plant biomass. Thus the role played by birds, bats, and lizards in pest regulation in agroecosystems can be seen both as positive (by decreasing herbivores, decreasing plant damage, and increasing plant biomass) and negative (by decreasing natural enemies of herbivores).

One of the main drivers behind species loss in terrestrial environments is habitat loss, or habitat degradation. Thus biodiversity effects on ecosystem services at landscape scales can be assessed indirectly through changes in land cover and in the spatial distribution of habitat patches. The mechanisms underpinning the maintenance of biodiversity in complex landscapes such as agricultural landscapes are increasingly being understood. Individual species' dynamics is given by that of individual populations that are found locally as well by the metapopulation, given by the group of populations found across the landscape. In similar ways, interactions among species within and among trophic levels occur at different spatial scale. Thus, structurally complex agricultural landscapes allow for the maintenance of biodiversity and are expected to contribute to the maintenance of ecosystem services such as pest regulation while ensuring food production (Tscharntke *et al.*, 2005).

One well-studied aspect of the role played by landscape structure on ecosystem provision is isolation from natural habitat. A meta-analysis of landscape effects on crop pollination services (Ricketts *et al.*, 2008) showed a consistent decline in pollinator richness with increasing isolation; richness drops to half of its maximum value at distances of 1500 m. A faster decay rate was found in visitation rate, which drops to half of its maximum value at a distance of 700 m, with steeper decay rates in tropical regions respective to temperate ones.

Ecosystem services can also be provided by the physical structure that results from the establishment and growth of species. That is the case of coastal protection which is provided by mangroves, salt marshes, coral reefs, and sea grass beds. Wave height decreases as it moves inland from a shoreline with mangroves or salt marshes. Decreases in wave height have been associated with increasing habitat distance inland from the shoreline for mangroves and salt marshes. Wave attenuation decreases with increasing water depth above sea grass beds and near-shore coral reefs, and with increasing grass cover in coastal dunes (Barbier *et al.*, 2008).

Given that changes in species richness, species' functional attributes and land cover characteristics are all relevant to determining change in ecosystem service provision, the integration of many of those aspects is needed to characterize the role played by biodiversity in service provision at landscape scales. Ecosystem processes and services are strongly determined by species' functional characteristics, or functional traits, and largely by those of dominant species. However, abiotic conditions can have direct effects on the services; also, nondominant species can have disproportionately large effects on the function. Thus the role played by biodiversity on ecosystem services can be explored using a step-wise analysis by dissecting the individual contribution of the above factors on ecosystem services, and then combining those into the

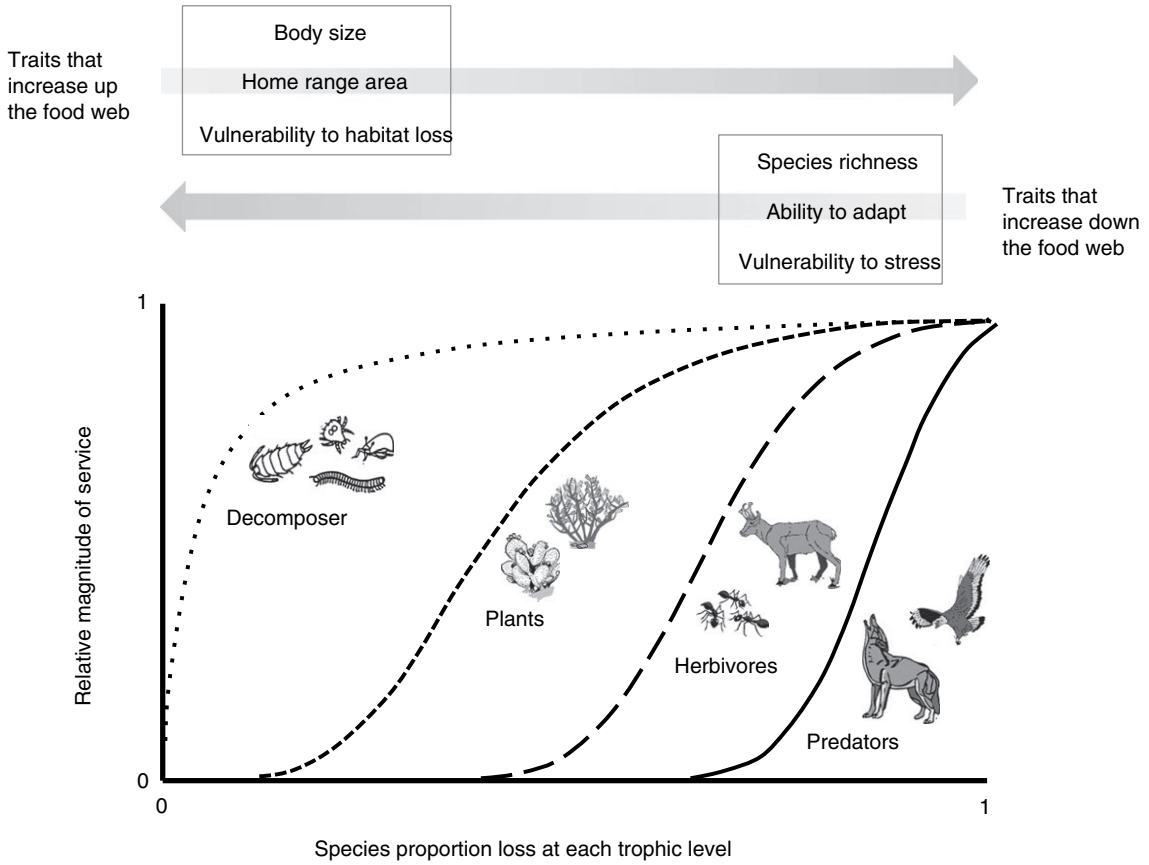


Figure 6 Declines in species diversity and corresponding impacts on ecosystem services for different trophic levels. A mathematical model suggests that declines in ecosystem services will initially be slow but will then accelerate as species from higher trophic levels are lost at faster rates and the consequences of their loss is most dramatic (modified from Dobson A, Lodge D, Alder J, *et al.* (2006) Habitat loss, trophic collapse, and the decline of ecosystem services. *Ecology* 87: 1915–1924).

best possible predictive model as has been suggested recently (Diaz *et al.*, 2007).

Spatially explicit models of the links between biodiversity and ecosystem services can be built based on the understanding of the tight links between land cover and ecosystem service provision, as well as those between species' functional attributes (Lavorel *et al.*, 2011). Thus land use can have both direct effects on ecosystem services, and those that affect species' functional attributes and in turn affect ecosystem service provision. Given that individual species' traits are relevant to various ecosystem services, and that land use change has impacts on multiple ecosystem services, models have been developed to assess the direct and indirect impacts of land use change on multiple ecosystem services. The recently developed tool to map and model ES called InVEST (Integrated Valuation of Ecosystem Services and Tradeoffs) uses LUC maps as a basis for further assessing quantitatively ES delivery in terrestrial and marine ecosystems.

At landscape scale the needs and perceptions of different social actors converge, each leading to different land use change decisions and fostering the provision of particular ecosystem services. Links between biodiversity and ecosystem services would then need to incorporate such complexity. Relevant ecosystem services can be related to species' functional traits and functional diversity based on scientific evidence and local quantifications. Yet this can also be done from the perspective of the local social actors based on their perceptions of such links and their preferences for different ecosystem services. A conceptual framework that integrates all the above has been developed (Diaz *et al.*, 2011), that results from a combination of approaches into an interdisciplinary (ecological and social sciences) and intersectorial (scientists and relevant local actors) perspective of the links between biodiversity and ecosystem services (**Figure 7**).

Regional to Global Scale

From regional to global scales exploration of the links between biodiversity and ecosystem services is much more related to their potential spatial congruence than their functional relationships. The analyses rely on spatially explicit models that predict both biodiversity and service provision that are based on topographic, climatic, and land use information. The key question is whether most diverse areas are also those that are most critical for the provision of individual services or that of combinations of services. To date, less than

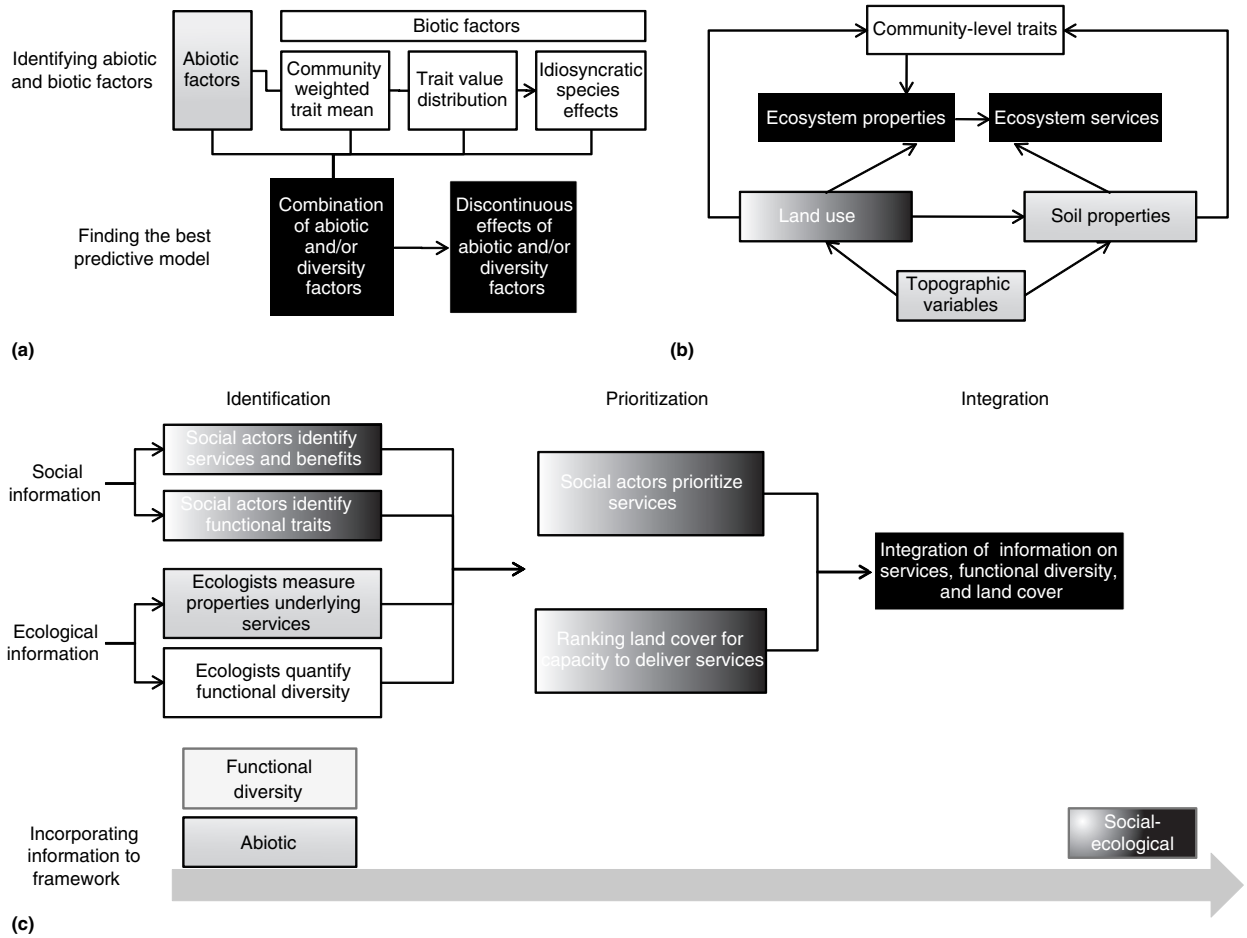


Figure 7 Diagrammatic conceptual frameworks for analyzing the links between abiotic factors (light gray), functional diversity (white), social–ecological interactions (black to white gradient) and ecosystem properties, and services (black). (a) Ecosystem services predicted from plant functional diversity (modified from Diaz S, Lavorel S, de Bello F, *et al.* (2007) Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences of the United States of America* 104: 20684–20689). (b) Predicting ecosystem services from ecosystem properties, ecosystem management and species' functional traits (modified from Lavorel S, Grigulis K, Lamarque P, *et al.* (2011) Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology* 99: 135–147). (c) Integrating ecological and social information about the links between biodiversity, ecosystem services and land use (modified from Diaz S, Quetier F, Caceres D, *et al.* (2011) Linking functional diversity and social actor strategies in a framework for interdisciplinary analysis of nature's benefits to society. *Proceedings of the National Academy of Sciences of the United States of America* 108: 895–902).

a dozen such studies are available at scales ranging from regional to global.

Spatially explicit models to predict the distribution of biodiversity have been developed for the past two decades. Instead, the development of analogous models for the case of ecosystem service delivery is very recent, and started to be developed only a decade ago. To date close to 50 different publications have addressed this issue from different perspectives. Yet, the broadest initiative to date on the subject, both in terms of the types of services analyzed, the range of data sources needed, and the number of places in the globe where it has been applied is the Natural Capital Project initiative (www.naturalcapitalproject.org). The initiative developed spatially explicit models for many ecosystem services as well as analytical tools to assess the management implications of the resulting maps (Kareiva *et al.*, 2011).

No consistent evidence of spatial concordance between biodiversity and ecosystem services has been found. This is true both for regions within a state (Chan *et al.*, 2006) and for the whole globe (Naidoo *et al.*, 2008). Also, the spatial patterns of provision of individual services are not correlated and thus critical areas for the maintenance of one service is not similar to that of another service (Chan *et al.*, 2006). When combining multiple ecosystem services with a total ecosystem service value the concordance between biodiversity and the services was not found either (Turner *et al.*, 2007).

The lack of congruence in the spatial patterns of biodiversity distribution and ecosystem service delivery is expected from a conceptual perspective. The lack of congruence in the spatial patterns of different groups of species was found a decade ago, and is consistent with the fact that each group is

differentially affected by changes in environmental conditions and that the scale at which such interaction is given changes with the size of the organism. In similar ways, the authors have seen here that different ecosystem service providers are relevant for the delivery of different services; ecosystem service providers and services respond differently to changes in environmental conditions and to anthropogenic disturbance.

Despite the lack of congruence in the distribution patterns of biodiversity and ecosystem services, management can be aimed toward maximizing the possibility of protecting both. Reserve selection algorithms have been used to identify the combination of areas that would maximize the conservation of biodiversity and the conservation of ecosystem services (Chan *et al.*, 2006). Spatially explicit models of biodiversity and ecosystem services have been used to predict the outcomes of different land use change scenarios and explore management alternatives that would maximize conservation of both biodiversity and ecosystem services. Stakeholder preferences for scenarios that provide multiple services while ensuring the conservation of biodiversity in a case study in the state of Oregon suggest opportunities for win–win policies that will ensure both (Nelson *et al.*, 2009).

To date, the identification of key areas of biodiversity or service provision heavily relies on the limitations inherent in the spatial modeling tools. In general, the analysis of different scales (total area studied) encompasses changes in the resolution (pixel size) of the study. Changes in resolution have been shown to lead contrastingly different patterns of concordance between biodiversity and service provision (Anderson *et al.*, 2009).

Understanding the links between biodiversity and ecosystem services at regional to global scales is still developing. Systematic comparisons with the equivalent methods are needed among regions and among scales. Also tools for this assessment might need to be refined.

Conclusions

Research available to date has shown that biodiversity plays a prominent role in the delivery of many key ecosystem services. A wide range of types of organisms, their genetic, population, and community diversity, as well as the landscape diversity and the trophic networks among them are the key ecosystem service providers. Biodiversity is linked to agricultural food provision, soil fertility regulation, pollination, and human disease regulation in a variety of ways; ongoing studies have contributed significantly to understand the ways in which biodiversity is important for ecosystem services. Most of the data available was derived from experimental manipulations of biodiversity and its effects on ecosystem services or on a few key ecosystem services. Direct measurements of a wide variety of ecosystem services are needed as well as more data at landscape scales, at which the most relevant connections between biodiversity and services are expected. Systematic comparisons of the role played by biodiversity at a set of spatial scales for many ecosystem services are needed. Also, the final effects on biodiversity and human well-being through ecosystem services are still to be fully understood.

Acknowledgments

We are very thankful to the many colleagues who have shaped our visions of the interactions between biodiversity and ecosystem services over the past few years. This paper constitutes a partial fulfilment of the Graduate Program in Biological Sciences of the National Autonomous University of México (UNAM). S Quijas acknowledges the scholarship and financial support provided by the National Council of Science and Technology (CONACyT), and UNAM. P. Balvanera was supported by DGAPA-UNAM for sabbatical at the Department of Biology, Stanford University.

See also: Agrobiodiversity. Diseases, Conservation and. Ecosystem Services. Human Impact on Biodiversity, Overview. Scale, Concept and Effects of. Soil Conservation

References

Anderson BJ, Armsworth PR, Eigenbrod F, *et al.* (2009) Spatial covariance between biodiversity and other ecosystem service priorities. *Journal of Applied of Ecology* 46: 888–896.

Balvanera P, Kremen C, and Martinez-Ramos M (2005) Applying community structure analysis to ecosystem function: Examples from pollination and carbon storage. *Ecological Applications* 15: 360–375.

Balvanera P, Pfisterer AB, Buchmann N, *et al.* (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9: 1146–1156.

Barbier EB, Koch EW, Silliman BR, *et al.* (2008) Coastal ecosystem-based management with nonlinear ecological functions and values. *Science* 319: 321–323.

Barrios E (2007) Soil biota, ecosystem services and land productivity. *Ecological Economics* 64: 269–285.

de Bello F, Lavorel S, Diaz S, *et al.* (2010) Towards an assessment of multiple ecosystem processes and services via functional traits. *Biodiversity and Conservation* 19: 2873–2893.

Cardinale BJ, Matulich KL, Hooper DU, *et al.* (2011) The functional role of producer diversity in ecosystems. *American Journal of Botany* 98: 572–592.

Chan KMA, Goldstein J, Satterfield T, *et al.* (2011) Cultural services and non-use values. In: Kareiva P, Tallis H, Ricketts T, Daily GC, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 206–228. Oxford: Oxford University Press.

Chan KMA, Shaw MR, Cameron DR, Underwood EC, and Daily GC (2006) Conservation planning for ecosystem services. *Plos Biology* 4: 2138–2152.

Chapin FS, Matson PA, and Mooney HA (2002) *Principles of Terrestrial Ecosystem Ecology*. New York: Springer-Verlag.

Chapin FS, Zavaleta ES, Eviner VT, *et al.* (2000) Consequences of changing biodiversity. *Nature* 405: 234–242.

Diaz S, Fargione J, Chapin FS, and Tilman D (2006) Biodiversity loss threatens human well-being. *Plos Biology* 4: 1300–1305.

Diaz S, Lavorel S, de Bello F, *et al.* (2007) Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences of the United States of America* 104: 20684–20689.

Diaz S, Quetier F, Caceres DM, *et al.* (2011) Linking functional diversity and social actor strategies in a framework for interdisciplinary analysis of nature's benefits to society. *Proceedings of the National Academy of Sciences of the United States of America* 108: 895–902.

Dobson A, Lodge D, Alder J, *et al.* (2006) Habitat loss, trophic collapse, and the decline of ecosystem services. *Ecology* 87: 1915–1924.

FAO (2011) *The State of Food and Agriculture*. Rome: Food and Agriculture Organization of the United Nations.

Gamfeldt L, Hillebrand H, and Jonsson PR (2008) Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology* 89: 1223–1231.

Hoeksema JD, Chaudhary VB, Gehring CA, *et al.* (2010) A meta-analysis of context-dependency in plant response to inoculation with mycorrhizal fungi. *Ecology Letters* 13: 394–407.

Hooper DU, Chapin FS, Ewell JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75: 3–35.

Hooper DU and Vitousek PM (1997) The effects of plant composition and diversity on ecosystem processes. *Science* 277: 1302–1305.

Jackson L, Rosenstock T, Thomas M, Wright J, and Symstad A (2009) Managed ecosystems: Biodiversity and ecosystem functions in landscape modified by human use. In: Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*, pp. 178–194. Oxford: Oxford University Press.

Jones KE, Patel NG, Levy MA, *et al.* (2008) Global trends in emerging infectious diseases. *Nature* 451: 990-U4.

Kareiva P, Tallis H, Ricketts TH, Daily GC, and Polasky S (eds.) (2011) Natural capital. *Theory and Practice of Mapping Ecosystem Services*. Oxford: Oxford University Press

Keesing F, Belden LK, Daszak P, *et al.* (2010) Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468: 647–652.

Klein AM, Müller C, Hoehn P, and Kremen C (2009) Understanding the role of species richness for crop pollination services. In: Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*, pp. 195–208. Oxford: Oxford University Press.

Kremen C (2005) Managing ecosystem services: What do we need to know about their ecology? *Ecology Letters* 8: 468–479.

Kremen C, Williams NM, Aizen MA, *et al.* (2007) Pollination and other ecosystem services produced by mobile organisms: A conceptual framework for the effects of land-use change. *Ecology Letters* 10: 299–314.

Lavorel S, Grigulis K, Lamarque P, *et al.* (2011) Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology* 99: 135–147.

Letourneau DK, Armbrecht I, Rivera BS, *et al.* (2011) Does plant diversity benefit agroecosystems? A synthetic review. *Ecological Applications* 21: 9–21.

Lindemann-Matthies P, Junge X, and Matthies D (2010) The influence of plant diversity on people's perception and aesthetic appreciation of grassland vegetation. *Biological Conservation* 143: 195–202.

Luck GW, Harrington R, Harrison PA, *et al.* (2009) Quantifying the contribution of organisms to the provision of ecosystem services. *Bioscience* 59: 223–235.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-being: Synthesis*. Washington, DC: Island Press.

Mooney KA, Gruner DS, Barber NA, Van Bael SA, Philpott SM, and Greenberg R (2010) Interactions among predators and the cascading effects of vertebrate insectivores on arthropod communities and plants. *Proceedings of the National Academy of Sciences of the United States of America* 107: 7335–7340.

Naidoo R, Balmford A, Costanza R, *et al.* (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9495–9500.

Nelson E, Mendoza G, Regetz J, *et al.* (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment* 7: 4–11.

Philpott SM, Soong O, Lowenstein JH, *et al.* (2009) Functional richness and ecosystem services: Bird predation on arthropods in tropical agroecosystems. *Ecological Applications* 19: 1858–1867.

Quijas S, Schmid B, and Balvanera P (2010) Plant diversity enhances provision of ecosystem services: A new synthesis. *Basic Applied and Ecology* 11: 582–593.

Raffaelli D (2004) How extinction patterns affect ecosystems. *Science* 306: 1141–1142.

Ricketts TH, Regetz J, Steffan-Dewenter I, *et al.* (2008) Landscape effects on crop pollination services: Are there general patterns? *Ecology Letters* 11: 499–515.

Rockstrom J, Steffen W, Noone K, *et al.* (2009) A safe operating space for humanity. *Nature* 461: 472–475.

Stiglitz JE, Sen A, and Fitoussi JP (2010) Mismeasuring our lives: Why GDP doesn't add up. *The Report of the Comission on the Measurement of Economic Performance and Social Progress*. New York: The New Press.

Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* 8: 857–874.

Turner WR, Brandon K, Brooks TM, Costanza R, da Fonseca GAB, and Portela R (2007) Global conservation of biodiversity and ecosystem services. *Bioscience* 57: 868–873.

van der Heijden MGA, Klironomos JN, Ursic M, *et al.* (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 396: 69–72.

Winfree R, Aguilar R, Vazquez DP, LeBuhn G, and Aizen MA (2009) A meta-analysis of bees' responses to anthropogenic disturbance. *Ecology* 90: 2068–2076.

Biodiversity and Human Health

Richard S Ostfeld, Cary Institute of Ecosystem Studies, Millbrook, NY, USA
Felicia Keesing, Bard College, Annandale-on-Hudson, NY, USA

Published by Elsevier Inc.

Glossary

Dilution effect The phenomenon whereby increases in biodiversity are associated with decreases in the transmission, risk, or incidence of infectious disease.
Disease incidence The number of new cases of a disease per population per unit time, often expressed as a rate.
Disease prevalence The number of disease cases per population.
Horizontal transmission The transmission of pathogens within single generations of hosts.
Infection prevalence The number of infections (or infected individuals) per population.
Infectious disease State of compromised health in a host caused by a pathogen infection.
Reservoir competence The degree to which a host is capable of supporting pathogen proliferation and transmission to other hosts or to vectors.
Vector (of pathogen) An animal, typically an arthropod, which transmits pathogens from one host to another.
Vertical transmission The transmission of pathogens across generations of hosts.
Zoonosis Disease of humans caused by pathogens that proliferate within and are transmitted from nonhuman vertebrate animals.

Introduction

Biodiversity – the variety of biological entities, from genes to species to ecosystems – is a very broad concept. With such a broad definition, it seems obvious that human health depends on biodiversity. From the plant and animal fibers we wear, to the plants, animals, and fungi we eat, to the natural compounds we use to treat infections, there is no aspect of human health and well-being that does not rely directly or indirectly on the variety of life. Even many synthetic materials are modeled after natural products. But we also sense – indeed, we have direct evidence that not every change in the earth's biodiversity results in a detectable change in human health. The loss of particular genes or species from natural ecosystems might be irrelevant to human health, whereas the loss of others could be cataclysmic. The loss of harmful organisms, such as smallpox virus, certainly is an example of biodiversity loss, yet it would be likely to increase human health. To analyze the effects of biodiversity loss or biodiversity change on human health, it is necessary to subdivide the breadth of biodiversity into component parts. It is also useful to subdivide human health into components in order to understand how specific aspects of biodiversity affect specific aspects of human health.

Subdividing Biodiversity

Scientific study of the linkages between biodiversity and human health has focused predominantly on species diversity, with some research pursuing the role of genetic diversity, and very little focusing on diversity among ecological communities or within entire ecosystems. This review will focus on how changes in species diversity affect human health. Species diversity is typically defined in one of two different ways. Species richness is the number of different species that occurs in a defined area. In practice, estimates of species richness (or any other metric of species diversity) are typically restricted to a particular taxonomic group (e.g., trees and mammals), or a particular trophic level (e.g., primary producers and herbivores). Species evenness is a measure of the relative abundances of species within a community. Quantitative metrics of species diversity, such as the Shannon Index or Simpson Index, combine both species richness and species evenness to derive a value that characterizes a community. So for instance, a three-species community in which each species has 33% of the individuals would be considered more diverse (higher evenness) than another community with the same three species but in which one species constitutes 95% of the community. In addition to the number of species and their relative abundances, the identity of species can be critical in understanding the importance of biodiversity. As natural communities change in biodiversity (species richness or evenness), it matters which species are added or lost. This is because species tend to vary strongly in their functional roles. Moreover, species tend to have vastly different probabilities of disappearing from communities as biodiversity is lost or of invading communities as biodiversity increases. Therefore, a critical aspect of biodiversity is the species composition of communities.

When asking how species diversity affects human health, then, it is necessary to be clear about which measure of biodiversity we are considering. An additional complexity is that the three aspects mentioned above – richness, evenness, and composition – are often interrelated in natural systems. For example, we can return to our three-species hypothetical community, with each species comprising 33% of the individuals. If one species is lost and the remaining two do not change in relative abundance, then richness will decline without a concomitant decline in evenness. However, because species in communities typically interact, the loss of one species is likely to affect the remaining species differently, such that one increases in abundance at the expense of the other. In this example, species richness and evenness would be interdependent. This effect is magnified by the fact that species within ecological communities have vastly different

probabilities of declining or disappearing as biodiversity is lost. For example, large-bodied animals, dietary specialists, predators, and rare species are much more likely to disappear from communities than are smaller animals, generalists, herbivores or omnivores, and abundant or widespread species. This example illustrates that species richness, species evenness, and species composition are interdependent.

Until recently, the predominant focus in the literature about diversity and human health has been on the effects of the disappearance of individual species rather than on biodiversity, which is an aggregate measure of multiple biological entities. Below the literature has been included on how individual species, or groups of species, affect human health only when those species are relevant to biodiversity by virtue of being under threat of serious decline or extinction.

Subdividing Human Health

Different components of biodiversity can affect different components of human health. The impact of changes in biodiversity on human health can be organized into four categories. First, changes in biodiversity can affect scientists' abilities to conduct basic research in the biomedical sciences. This occurs when species that are or might be useful as animal models of human health are affected by drivers of biodiversity loss. Second, the development of future treatments for human disease can be strongly affected by biodiversity loss. A large percentage of current treatments (e.g., pharmaceutical drugs) was derived from biological species, some of which are threatened with decline or extinction. Much effort is also spent to discover biological species or compounds that can serve as sources of preventative or curative therapies, and biodiversity loss threatens such discovery. Third, humans rely on many different plant, animal, and microbial species for food, and the loss of biodiversity can affect health through the direct and indirect impacts of nutrition. The diversity of current and potential food species takes on additional importance as climate change alters the conditions under which many of our currently used crops and livestock were selected. Degrading environmental conditions will increase the necessity for incorporating diverse species and genotypes into human nutrition. Fourth, changes in biodiversity can affect the abundance, distribution, and transmission dynamics of infectious agents that attack humans, livestock, or crop plants, with strong consequences for human health.

To reflect the major research foci developing in the past 10–15 years, the authors will describe how changes in biodiversity affect: (1) basic biomedical research; (2) identification and development of preventative and therapeutic agents; (3) nutrition; and (4) risk and incidence of infectious diseases.

Biodiversity and Basic Biomedical Research

Nonhuman animals have served as models of human biomedical research for centuries. In some respects, each such use of a nonhuman model could be considered relevant to a discussion of biodiversity and human health because all these model organisms contribute to the earth's biodiversity. In other respects, however, most of the traditional model organisms are not highly relevant to discussions of the interactions between biodiversity and human health because they are not imperiled and because they are constructs of human genetic modification and do not represent natural biodiversity. Examples of the latter include the many laboratory genotypes of the laboratory mouse *Mus musculus*, the lab rat *Rattus norvegicus*, the fruit fly *Drosophila melanogaster*, and the roundworm *Caenorhabditis elegans*. But in the case of nontraditional models of human biology and medicine, some species are themselves imperiled or belong to taxonomic groups of which some members are of conservation concern. Below the authors will discuss two examples of how these species have contributed or might contribute to a basic understanding of human disease. The underlying theme is that some wildlife species are able to perform functions that solve physiological challenges similar to those humans also face. Further study of these adaptations is likely to lead to a better understanding of, and solutions to, human diseases of various kinds. Continued declines in biodiversity could foreclose these opportunities to improve public health.

Bears and Human Physiology

Most species of bear (Family Ursidae) undergo long periods of seasonal quiescence called "denning" or hibernation. Hibernation by bears differs from that of "true hibernators" (e.g., some bats, ground squirrels, and dormice) in that bear body temperatures decline by only a few degrees but not to ambient and bears are easily awakened. Bear hibernation can continue for many months, during which the bears do not eat, drink, urinate, or defecate, and they do not even move at all. To contend with the accumulation of urine, which continues during hibernation, bears transport urine through the lining of the bladder into the bloodstream, where the urea is degraded into amino acids that can be used to synthesize proteins (Chivian and Bernstein, 2008). Buildups of urea characterize human patients with end-stage renal disease, which is currently treatable only with dialysis or kidney transplantation. Considerable interest exists, therefore, in understanding the physiological mechanisms by which bears recycle urea into proteins so that treatments for human renal disease can be developed.

Nutritional deficiencies and immobility by humans can cause dramatic loss of bone mass leading to osteoporosis, which is a major public health problem worldwide. In bears, however, the months of fasting and immobility during hibernation do not cause a loss of bone mass, because bears recycle the calcium lost from bone back into forming new bone (Chivian and Bernstein, 2008). Apparently, hibernating bears are able to down-regulate the activity of osteoclasts, cells that help regulate bone resorption, and up-regulate osteoblasts and fibroblasts, cells that form new bone and cartilage, respectively (Chivian and Bernstein, 2008). The biochemical extract responsible for these regulatory activities in bears has reversed bone loss in ovariectimized rats, which simulates treatment for osteoporosis in postmenopausal women (Chivian and Bernstein, 2008).

As large-bodied predators, bears are more vulnerable to extinction than are smaller-bodied animals and those at lower trophic levels. Consequently, the ability of scientists to make use of the unique adaptations of bears for the betterment of human health is vulnerable.

Cone Snails and Human Physiology

The genus *Conus* is a speciose group of mostly tropical marine snails with many members that are probably rare and perhaps in danger of extinction (Chivian and Bernstein, 2008). Too little is known about their abundance and population trends to accurately characterize their conservation status, but many occupy sites (coral reefs and mangrove habitats) that are being destroyed and degraded at a rapid pace, and many are subject to unregulated collecting by amateur malacologists. Cone snails are predators that capture mobile prey (e.g., fishes) by harpooning them with a specialized "tooth" that contains a cocktail of toxic peptides. Each species of *Conus* apparently has a unique mixture of >100 such peptides and the entire genus might produce as many as 100,000 different toxic peptides (Chivian and Bernstein, 2008). Although these peptides, called conopeptides, have a common impact on their victims – paralysis – each appears to have a different mechanism of action by virtue of an affinity for different cell surface receptors. Many act by blocking ion channels that control concentrations of ions such as sodium, potassium, calcium, and chloride on each side of the cell membrane. Because controlling the flux of ions to and from cells is critical for many physiological functions, the nature and mechanism of action of conopeptides is a fertile ground for basic biomedical researchers. For example, recently a conopeptide was discovered that has a highly specific, strong affinity for a particular type of acetylcholine receptor in vertebrate nervous systems and muscle cells. Acetylcholine receptors control ion channels responsible for regulating nerve cell impulses, but many such receptors exist, each with a different regulatory function. The newly discovered conopeptide's extreme specificity for a particular acetylcholine receptor suggests that other conopeptides will be similarly highly specific. Having a diverse group of highly specific peptides should pave the way for basic research into the proximate causes of diseases in which abnormal acetylcholine regulation plays a key role, including Parkinson disease, Alzheimer disease, epilepsy, and alcoholism (Chivian and Bernstein, 2008).

Biodiversity and Preventative/Therapeutic Agents

Humans have used natural compounds as medicines since antiquity. In modern times, active compounds from plants, animals, and microbes are important worldwide in both scientifically evaluated and "traditional" medicines. At least half of the most prescribed medicines in the USA contain active compounds directly or indirectly derived from natural sources (Newman *et al.*, 2008). The enormous diversity of chemical compounds produced by the earth's biota constitutes a rich source of molecules of potential use in protecting or improving human health. Discovery of natural compounds that could be used as medicines proceeds either by targeted screening or by untargeted screening, the latter of which requires enormous numbers of species to be tested because the odds of each individual species containing useful compounds are so low, (Newman *et al.*, 2008). Screening can be targeted by using prior information about traditional human use of particular organisms to solve specific health problems (e.g., (Rosenthal and Preszler, 2002). Screening can also be targeted by focusing on species whose natural histories suggest that they share pathogens or other health challenges with humans. For example, saprophytic bacteria and fungi often compete for resources, and both types of microbes might be expected to produce chemical compounds that inhibit their competitors. Consequently, antibacterial agents are often sought in certain fungi, and antifungal and antibacterial agents are often sought in certain types of bacteria.

Newman *et al.* (2008) provide many examples of plants, animals, and microbes that produce compounds that scientists have developed into medicines. Examples of plants include the analgesic morphine from opium poppies (*Papaver somniferum*); vincristine and vinblastine as highly potent anticancer compounds from rosy periwinkle (*Catharanthus roseus*); the precursor of aspirin in willows (*Salix* spp.); and the anti-HIV drugs called calanolides from tropical trees of the genus *Calophyllum*. Examples of animals include potent anticoagulants from medicinal leeches (*Hirudo medicinalis*) and canine hookworms (*Ancylostoma caninum*), and angiotensin-converting-enzyme (ACE)-inhibiting compounds instrumental in controlling hypertension that were initially discovered in the venom of a pit viper *Bothrops jararaca*. Examples from microbes include penicillin and its many derivatives, initially found in *Penicillium notatum* molds; the aminoglycoside antibiotics (including streptomycin) from an actinomycete bacterium *Streptomyces griseus*; and the cholesterol-controlling statins, initially isolated from several *Penicillium* species.

Although the organisms mentioned above have provided compounds of enormous benefit to human health and well-being, none are imperiled or of conservation concern, a fact that complicates the argument that these are examples of strong connections between biodiversity and human health. However, the sheer number of medically useful natural compounds already discovered (Newman *et al.*, 2008; Rosenthal and Preszler, 2002) combined with the tiny fraction of species that have been thoroughly screened for such compounds, and the precipitous loss of biodiversity worldwide, together suggest that we are rapidly losing potential sources of medicines as biodiversity declines.

Biodiversity and Nutrition

Biodiversity is essential for the production of both plant-based and animal-based foods. Throughout history, wild plants and animals have been chosen for domestication and selectively bred to improve desirable traits. Biodiversity in agroecosystems can affect human health through several mechanisms, including increased resistance to pests and pathogens, insurance against current and future environmental changes, improved nutrition through dietary diversity, and better immune defenses against infectious diseases.

Monocultures of genetically similar or identical crops are notoriously vulnerable to massive disease-caused failures, because pathogens can encounter huge numbers of susceptible hosts leading to rapid spread (Elton, 1958; Keesing *et al.*, 2006). Examples of monoculture crop disasters that have been caused by pathogens include wheat failures during ancient Roman times caused by stem rust (*Puccinia graminis*), destruction of rye crops due to ergot poisoning during Europe's Middle Ages, the decimation of French grape vineyards caused by downy mildew disease, and the Irish potato famine of the mid-nineteenth century, caused by *Phytophthora infestans* (Hillel and Rosenzweig, 2008). Experimental plantings of diverse polycultures of different strains of rice have vastly reduced the burden of rice blast disease in China, avoiding the need for massive fungicide application and improving productivity and profits (Zhu *et al.*, 2000). Although in recent years genetic engineering has emerged as a means of adding genetic variants to produce desirable traits such as pest resistance, drought resistance, or nutritional value, the creation of genetic variants through cross-breeding with wild relatives remains an essential means of improving crops.

Current agricultural productivity and resistance to pests and pathogens can be increased by both genetic diversity within populations of plants and animals already used for agriculture and by species diversity of related wild plants and animals. This diversity is also likely to be critical in confronting future environmental challenges arising from climate change, introduced pests, emerging pathogens, and nutrient-depleted soils. Traits that support high productivity under current environmental conditions are not likely to support equally high productivity as those conditions change, because of trade-offs among specific adaptations. These observations should serve as a caution to agronomists. Because strong selection reduces genetic diversity by promoting some alleles at the expense of others, and because only relatively few plant and animal species are currently profitable, modern agricultural practices tend to erode crop biodiversity (Hillel and Rosenzweig, 2008).

A recent study in western Kenya showed that farms with higher diversity of edible plant species provided a greater diversity of essential nutrients, including protein, carbohydrates, vitamins A and C, iron, zinc, and folate (DeClerck *et al.*, 2011). Low nutritional diversity of crop plants within these agricultural fields was associated with high rates of anemia among associated community members, whereas high nutritional diversity of crop plants within these agricultural fields was strongly correlated with the absence of anemia (DeClerck *et al.*, 2011).

Crops that require animal pollination comprise about one third of agricultural food production globally (Klein *et al.*, 2007). Biodiversity loss in guilds of both avian and mammalian pollinators can compromise crop production, especially from tropical fruit trees (Kunz *et al.*, 2011; Whelan *et al.*, 2008). But more importantly, the loss of insect pollinators, especially native bees, from many ecosystems is causing extensive declines in pollination success and hence yields of vegetable and fruit crops globally (Kremen and Ostfeld, 2005; Potts *et al.*, 2010; Winfree, 2010). Insufficient levels of pollination caused by the loss of insect pollinators have caused declines of up to 70% in crop productivity (Hillel and Rosenzweig, 2008).

Biodiversity and Infectious Disease

Four recent reviews of the scientific literature have revealed many cases in which high biodiversity is associated with low risk or incidence of infectious disease, and conversely, many cases in which biodiversity loss is associated with increased disease risk or incidence (Johnson and Thieltges, 2010; Keesing *et al.*, 2006, 2010; Pongsiri *et al.*, 2009). Basic epidemiological models of a simplified system consisting of a single host species and a single pathogen specialized on that host show at least five mechanisms by which adding one or more species to the system (increasing diversity) can reduce the spread of infection (Keesing *et al.*, 2006). Increased diversity can reduce transmission by: (1) reducing the rate of encounter between susceptible and infectious individuals, a mechanism termed encounter reduction; (2) reducing the probability of pathogen transmission when encounters do take place (transmission reduction); (3) decreasing the abundance of susceptible individuals (susceptible host regulation); (4) increasing the recovery rate of infected individuals (recovery augmentation); and (5) increasing the death rate of infected individuals (infected host mortality) (**Figure 1**). In disease systems in which a vector (usually an arthropod such as a mosquito or tick) is important in transmitting pathogens, an additional mechanism can occur. In these systems increased diversity can: (6) decrease the abundance of vectors (vector regulation). In addition, in these vector-borne disease systems, increased diversity can reduce encounter rates between vectors and either infected or susceptible hosts, leading to a second form of encounter reduction (*see* Conceptual Antecedents). Widespread support exists for three of these mechanisms: encounter reduction; susceptible host regulation, and vector regulation; whereas studies of the other mechanisms are largely lacking. Although in theory increased diversity could increase risk or incidence of infectious diseases, empirical studies demonstrating such an effect are rare (Keesing *et al.*, 2010).

The phenomenon whereby increased diversity decreases disease transmission has been termed the "dilution effect" (Norman *et al.*, 1999; Ostfeld and Keesing, 2000a, b; Vanbuskirk and Ostfeld, 1995), and the reverse phenomenon has been termed the "amplification effect" (Keesing *et al.*, 2006). In the following sections the author will describe some historical antecedents to the notion that high biodiversity reduces disease transmission, followed by discussions of specific diseases in which biodiversity has been shown to reduce pathogen transmission, in each case referring to the underlying mechanisms whenever these are known.

Conceptual Antecedents

An important precursor to the modern study of the relationship between biodiversity and infectious disease transmission is the notion of zooprophylaxis, which the World Health Organization (Organization, 1982) defined as "the use of wild or domestic animals, which are not the reservoir hosts of a given disease, to divert the blood-seeking mosquito vectors from the human hosts of that disease." This definition, which follows an earlier treatment by (Brumpt, 1944–1945),

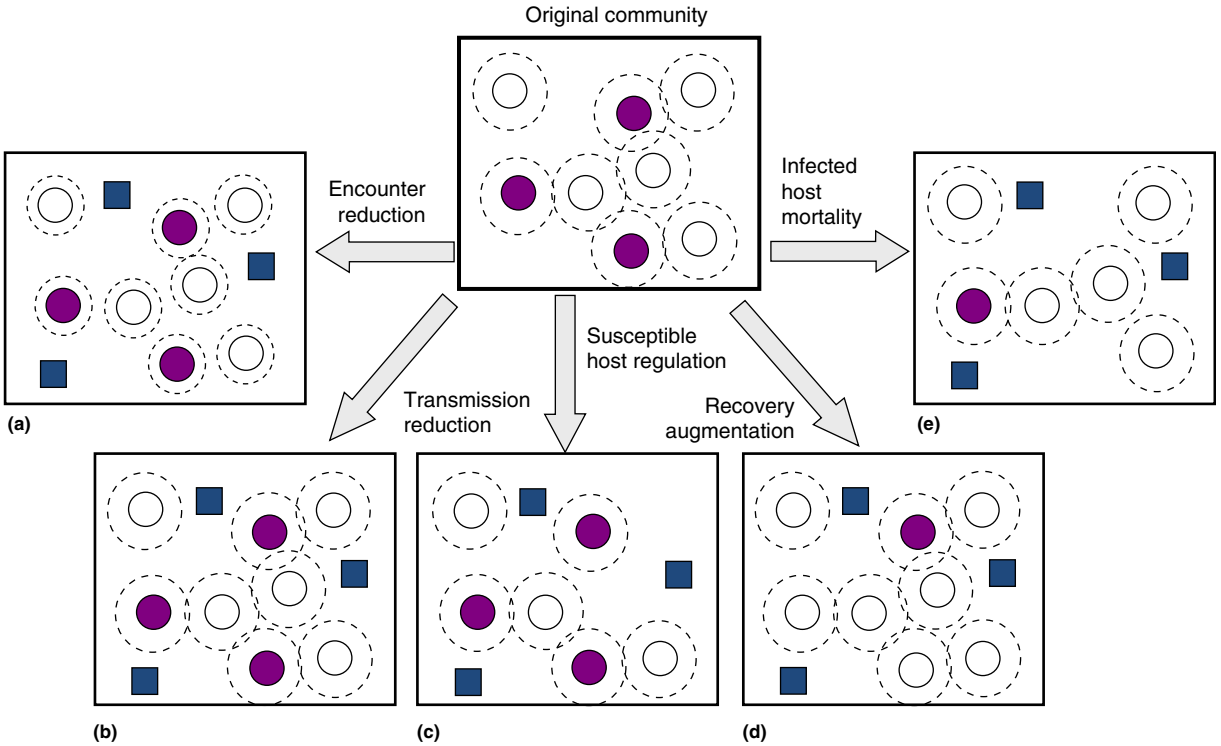


Figure 1 Conceptual model of the mechanisms by which diversity could reduce disease risk in a specialist host-pathogen system for a non-vector-borne disease. The original community consists of a single species with some individuals infected (filled circles) and some individuals uninfected and therefore susceptible (open circles). Each individual uses a particular home range (dashed lines). The addition of a second species (black squares) could (a) result in a reduction in space use by the host species, thus reducing encounters between susceptible and infected individuals (encounter reduction); (b) reduce the probability of transmission given encounters, as indicated here by no increase in the number of infected individuals despite contacts that could lead to transmission (transmission reduction); (c) reduce the number of susceptible hosts (susceptible host regulation); (d) increase the rate of recovery of infected individuals, as indicated by some infected individuals becoming uninfected (recovery augmentation); or (e) increase the mortality rate of infected individuals (infected host mortality). Redrawn from Keesing F, Holt RD, and Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecology Letters* 9: 485–498. doi: DOI 10.1111/j.1461-0248.2006.00885.x, with permission from Nature.

implies that diversity has increased via the addition of non-human animals and that the mechanism of reduced disease transmission is the diversion of vector blood meals away from humans. Note that according to this definition, zooprophylaxis is a manifestation of encounter reduction pertaining specifically to mosquitoes, although the concept has sometimes been extended to other arthropod vectors of human pathogens (Service, 1991). Therefore, zooprophylaxis is not equivalent to the dilution effect, but rather describes one manifestation of one of the five specific mechanisms by which the dilution effect can operate. Another precursor to modern study of the dilution effect is Elton's (Elton, 1958) recognition that plant diseases are reduced in intensity or distribution in more diverse plant communities, especially when that higher diversity reduces the population density of the host plant. Note that this concept of a protective effect of diversity on disease transmission is a manifestation of susceptible host regulation (Keesing *et al.*, 2006). For decades or centuries before Elton, agronomists have used various forms of intercropping (the practice of growing mixtures of crop plants, either simultaneously or sequentially) to increase yields. One of the main mechanisms by which crop diversity can increase yields is by reducing abundance of or attack rates by plant pathogens (Altieri, 1999).

Lyme Disease

Lyme disease is caused by the spirochete bacterium *Borrelia burgdorferi*, which is transmitted among wildlife hosts and from wildlife to humans via the bite of an ixodid tick. In Europe and Asia the main vectors are *Ixodes ricinus* and *Ixodes persulcatus*, respectively, and in North America, the main vector is the blacklegged tick *Ixodes scapularis*. Because *Ixodes* ticks do not transfer the bacterium from mother to offspring, each new generation of larval ticks hatches free of infection. After hatching from eggs, larval ticks take a single 3–4-day-long blood meal from a vertebrate host, during which they can acquire infection with *Bo. burgdorferi* if the host itself is infected. This larval blood meal is also critical in determining whether the tick survives to the next stage – the nymph – because larvae attempting a blood meal are susceptible to host grooming, which is an important cause of mortality (Keesing *et al.*, 2009). Nymphal ticks are responsible for transmitting infection to both reservoir hosts and human victims (Ostfeld,

2011). The risk to people of contracting Lyme disease is strongly correlated with the abundance of infected nymphal ticks in the environment (Ostfeld, 2011).

Vertebrate biodiversity is important in determining both the abundance of nymphal ticks and the proportion of nymphs that are infected with *Bo. burgdoferi*. Blacklegged ticks will readily attempt to feed on dozens of different species of mammals, birds, and lizards. Indeed, they have been reported infesting >112 different host species (Keirans *et al.*, 1996). Larval ticks appear to ''prefer'' some hosts more than others when offered a choice in a laboratory setting (Shaw *et al.*, 2003; Slowik and Lane, 2009), but these preferences tend to be weak, and the ticks' extraordinarily wide distribution among hosts underscores their lack of specificity in host associations. The abundance (intensity) of larval ticks on any particular host appears to be quite sensitive to the availability of alternative hosts in the community. For example, Brunner and Ostfeld (2008) showed that, as the population density of one important host, the eastern chipmunk (*Tamias striatus*), decreased, the abundance of larval ticks on another key host, the white-footed mouse (*Peromyscus leucopus*), increased up to 20-fold. Apparently, because these ticks are only weakly motile and require a host to draw near in order for them to attach, the relative abundances of different hosts strongly affect encounter rates between ticks and those hosts and thus the distribution of larvae across the host community.

Different species of hosts have vastly different probabilities of supporting blood feeding by ticks and tick survival, and ticks feeding on different hosts have different probabilities of acquiring infection with the Lyme disease agent. Among mammal and bird hosts that are commonly parasitized by blacklegged ticks in the eastern USA, the white-footed mouse is the host most likely to support tick survival during blood feeding (Keesing *et al.*, 2009), overwinter survival after the blood meal (Brunner *et al.*, 2011), and acquisition of *Bo. burgdorferi* spirochetes (LoGiudice *et al.*, 2003, 2008). Chipmunks and shrews are intermediate in most respects, and most other members of the host community tend to be poor hosts and poor reservoirs for the Lyme disease bacterium. Therefore, white-footed mice play a critical role in the risk of human exposure to Lyme disease. The abundance of infected nymphal ticks is strongly correlated with the population density of mice in space and time (Ostfeld *et al.*, 2006). As nonmouse hosts are added to, or increase in abundance in, host communities, the risk of human exposure to Lyme disease declines (Ostfeld and Keesing, 2000a; Schmidt and Ostfeld, 2001).

White-footed mice are highly resilient to anthropogenic and natural disturbance and are often most abundant in disturbed, degraded, or fragmented habitats, including suburban and exurban homes (Ostfeld, 2011). This high ecological resilience seems to be facilitated by extreme habitat and dietary generalization, small body size, strong dispersal ability, and high reproductive potential. Other vertebrates that are more specialized, larger bodied, weaker dispersers, and slower in reproductive rates tend to decline or disappear in highly disturbed, degraded, or fragmented habitats. A critical consequence of these differences in ecological resilience is a tendency for mice to persist, and even thrive, when vertebrate biodiversity is lost (LoGiudice *et al.*, 2008; Ostfeld, 2011). Thus, low-diversity vertebrate communities typically contain abundant mice but few other species, whereas high-diversity communities contain mice and many other species in high relative abundances. These selective declines or local extinctions of host species that tend to buffer Lyme disease transmission appear to underlie the observation that highly fragmented landscapes and small forest fragments maintain the highest abundances of infected blacklegged ticks (Allan *et al.*, 2003; Brownstein *et al.*, 2005; Jackson *et al.*, 2006) (**Figure 2**).

Observations from the Lyme disease system allowed Ostfeld and Keesing (Ostfeld and Keesing, 2000a) to characterize the general features that are necessary and sufficient to produce a dilution effect (i.e., a negative correlation between biodiversity and disease risk) in any vector-borne disease system. The conditions are: (1) a vector must be at least somewhat generalized in its choice of hosts (i.e., a nonspecialist). The more specialized the vector is on one or a few hosts, the less important overall diversity of the host community will be. (2) The pathogen must rely largely on host-to-vector transmission rather than on vertical transmission across generations of vectors. For pathogens that are efficiently transmitted transovarially (from mother vector to offspring), host-to-vector transmission is relatively less important in determining infection prevalence and therefore transmission risk. (3) The various host species from which the vector feeds must differ in their ''reservoir competence,'' or their probability of infecting a feeding vector in nature. If host species were interchangeable in their reservoir competence, then the identity of the blood meal host, and hence host diversity, would not matter to transmission dynamics. More recently, Keesing *et al.* (2009) added a correlate of condition (3), namely that the various host species differ in their ''permissiveness'' to vector feeding, or their probability of allowing the vector to feed successfully versus killing it during grooming. (4) The host species most likely to successfully provide the vector with a blood meal and to infect it must be those that persist in species-poor communities. If the most competent reservoir hosts tended to occur only in the most diverse communities, then high-diversity ecosystems would produce the highest risk – the opposite of the dilution effect. When these four conditions are met, vectors co-occurring with low diversity host communities will tend to bite the most competent reservoirs and the most permissive hosts, leading to higher abundance and/or infection prevalence in the vector population. The consequence will be heightened risk of transmission to people. Ample evidence now exists to support the widespread occurrence of these four conditions, which indicates that the dilution effect is a general phenomenon (Keesing *et al.*, 2010).

West Nile Virus (WNV)

WNV is a flavivirus (relative of the yellow fever virus) that is amplified in certain species of birds and transmitted among wildlife hosts by mosquitoes, predominantly members of the genera *Culex* and *Aedes*. When mosquitoes infected with WNV bite humans, serious illness or death can result. The virus was restricted to parts of Africa, Asia, and Europe until 1999 when it was inadvertently introduced into the New York City area.

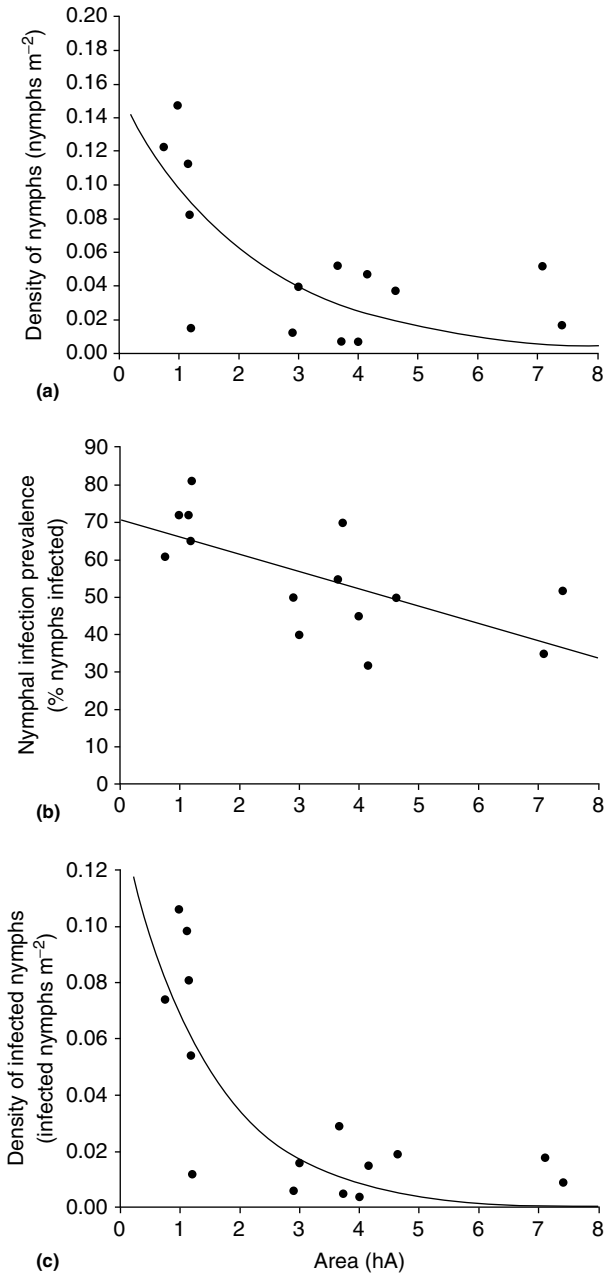


Figure 2 The total density of nymphal blacklegged ticks *Ixodes scapularis* (a) percentage of nymphs infected with *B. burgdorferi* (b) and density of infected nymphs (c) in 14 forest fragments of Dutchess County, New York. Reprinted from Allan BF, Keesing F, and Ostfeld RS (2003) Effect of forest fragmentation on Lyme disease risk. *Conservation Biology* 17: 267–272, with permission from Wiley.

A mere 3–4 years after its appearance in New York the virus had dispersed all the way across North America to the west coast, causing illness or death in tens of thousands of people and countless wild birds as it spread. Throughout much of North America, it remains a serious health threat to both people and native avifauna (Allan *et al.*, 2009; Kilpatrick *et al.*, 2006a; LaDeau *et al.*, 2007, 2008).

The WNV system shares some important features with the Lyme disease system. The virus relies predominantly on horizontal (host to vector to host) transmission rather than on vertical (transovarial) transmission across generations of mosquitoes. Therefore, similar to Lyme disease, the prevalence of infection in the vector depends on which hosts are bitten and how likely those hosts are to transmit the pathogen to the feeding mosquito. Unlike the Lyme disease system, in which tick abundance is determined at least partly by the abundance of some hosts, mosquito abundance appears to be less dependent on host abundance and more dependent on the availability of aquatic breeding sites (Dobson, 2004). This independence of vector abundance from host abundance simplifies the system in that the major impact of hosts is on vector infection prevalence rather than on vector abundance. Similar to the tick vectors of Lyme disease, the mosquitoes that transmit WNV are host generalists – they readily take blood meals from many different species of birds, mammals, and reptiles within ecological communities (Kilpatrick *et al.*, 2007; Komar *et al.*, 2003; LaDeau *et al.*, 2008; Marra *et al.*, 2004). *Culex* and *Aedes* mosquitoes do show preferences for some hosts more than others (Kilpatrick *et al.*, 2006b), and their high mobility relative to ticks means that they can more readily reject potential hosts that are less preferred. Nevertheless, at the population level, it is clear that vector mosquitoes bite most or all available members of the vertebrate community. Within West Nile-endemic zones, ample evidence indicates that infection is widespread among birds, mammals, and some reptiles (LaDeau *et al.*, 2007, 2008; Marra *et al.*, 2004), confirming that the vectors are generalists despite having host preferences.

Similar to the Lyme disease system, the various hosts for mosquitoes differ dramatically in their ability to support replication by the pathogen, the length of time they remain infected, and their probability of transmitting virus to feeding mosquitoes (Komar *et al.*, 2003). Although studies at different localities sometimes implicate different bird species as reservoirs, the general pattern is that American robins, house sparrows, common grackles, American crows, blue jays, and house finches are the most competent reservoirs for WNV within bird communities. Many other passerines (songbirds), nonpasserines, mammals, and reptiles, either fail to permit the virus to amplify sufficiently, producing very low concentrations of virus in their blood (low viremia) or they permit a high viremia but for only a very brief period. Either of these two processes will make these bird species inefficient at transmitting the virus to mosquitoes and therefore cause them to be incompetent reservoirs.

Ezenwa *et al.* (2006) were the first to test for an association between avian diversity and the prevalence of WNV in mosquitoes and humans. They reasoned that, because nonpasserine birds were generally incapable of producing high viremias, the presence of a high diversity of these bird species should cause mosquitoes to feed predominantly on poor reservoir hosts, reducing transmission rates of WNV from host to vector. To test this hypothesis, they captured hundreds of batches of *Culex* mosquitoes from West Nile-endemic sites in eastern Louisiana where they also estimated species richness in the local bird communities. They found that the proportion of mosquitoes infected with WNV was strongly negatively correlated with species richness of nonpasserines and uncorrelated with species richness of passerines. An interesting

additional result was that the population density of infected *Culex* mosquitoes was also negatively correlated with nonpasserine diversity, suggesting that nonpasserines might be killing mosquitoes that attempt to feed (similar to what Keesing *et al.* (2009) observed for ticks and Lyme disease). In addition, examining county-specific data throughout Louisiana, Ezenwa and colleagues found that the incidence of West Nile disease in humans was negatively correlated with species richness of nonpasserines in that county (**Figure 3**).

Three subsequent studies also found strong support for a dilution effect in WNV, although in all these cases it was species diversity of passerines that reduced virus prevalence. Swaddle and Calos (2008) selected 65 adjacent pairs of counties in the eastern USA such that each pair consisted of one county with human cases of West Nile encephalitis and an adjacent county without. They tested the hypothesis that the West Nile-negative county in each pair would contain higher bird species diversity, as measured by the Breeding Bird Survey data from the United States Geological Survey. Indeed, Swaddle and Calos found that the pairs of counties with the largest difference in bird diversity also had the largest difference in prevalence of West Nile encephalitis, with higher bird diversity being correlated with lower disease in humans.

Allan *et al.* (2009) examined human incidence of West Nile disease and bird diversity from 742 US counties within 38 states during years 2002–2004 of rapid, cross-country spread. They calculated bird diversity from the Breeding Bird Survey data using the Shannon Diversity metric, which emphasizes both species evenness and richness. First, they calculated the total "community competence" of birds by multiplying the abundance of each species by its reservoir competence and summing the species-specific totals. Community competence provides an estimate of the likelihood that a particular bird community will produce infected mosquitoes. They found that species diversity of bird communities was strongly negatively correlated with community competence, a result that parallels that in the Lyme disease system. In other words, host communities with low diversity tend to be biased toward highly competent reservoirs, whereas high diversity communities contain poorer reservoirs. By comparing different models for explaining the variation among counties in West Nile disease, Allan and colleagues found that bird diversity was the strongest and most consistent factor – counties with low avian diversity had, on average, much higher rates of West Nile disease than did those with high diversity (**Figure 4**).

Koenig *et al.* (2010) examined the role of bird diversity in population declines of American crows during the initial

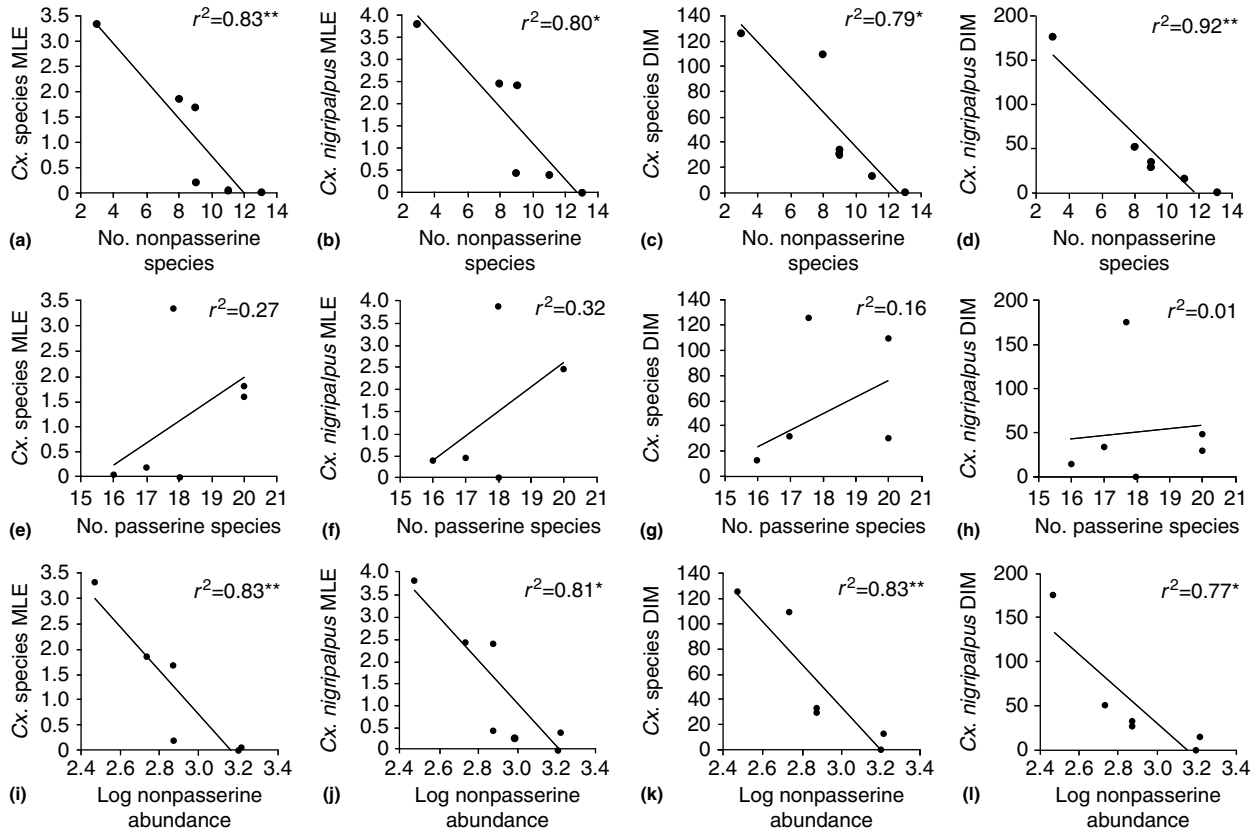


Figure 3 Various measures of risk of human exposure to mosquitoes infected with West Nile virus as a function of the species richness of songbirds (passerines) and nonpasserine birds. Both the infection prevalence of mosquitoes (maximum likelihood estimate (MLE)) and the density of infected mosquitoes (DIM) strongly, negatively correlated with species diversity of nonpasserines, but were unrelated to species diversity of passerines. Reprinted from Ezenwa VO, Godsey MS, King RJ, and Guptill SC (2006) Avian diversity and West Nile virus: Testing associations between biodiversity and infectious disease risk. *Proceedings of the Royal Society B-Biological Sciences* 273: 109–117. doi: DOI 10.1098/rspb.2005.3284, with permission from Royal Society of Publishing.

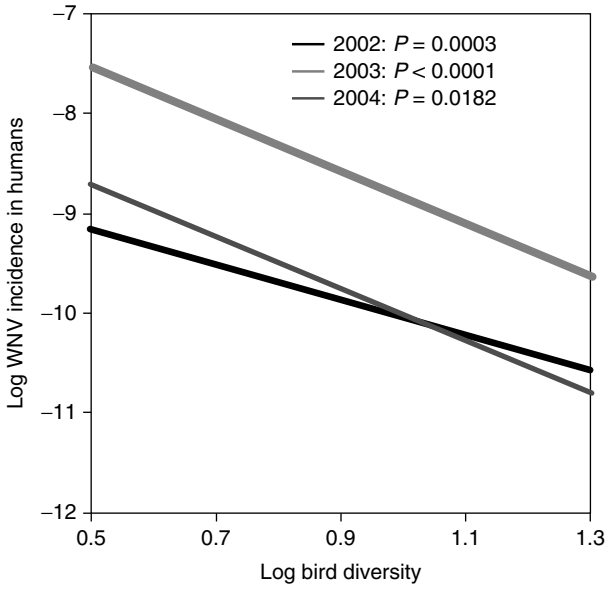


Figure 4 Negative correlations between the number of human cases of disease caused by West Nile virus in an American county and the average species diversity of birds in that county as determined by the breeding bird survey. See Allan *et al.* (2009) for details. Reprinted from Ostfeld RS (2009) Biodiversity loss and the rise of zoonotic pathogens. *Clinical Microbiology and Infection* 15: 40–43, with permission from Springer.

5-year sweep of WNV across USA. Crows can experience rapid mortality after infection resulting in severe population declines as WNV invades new areas (LaDeau *et al.*, 2007); therefore crow population declines are a good proxy for transmission rates within wildlife communities. Koenig and colleagues quantified the severity of crow declines using the Breeding Bird Survey data for individual survey routes conducted before and after invasion by the virus. They then tested the ability of several factors to explain the severity of crow declines, including: the year WNV was first detected in the state; population density of crows; bird diversity (Shannon index); urbanization (estimated by human population density); total number of individual birds; mean rainfall; mean maximum temperature; and a measure of spatial self-correlation in the data (to represent the likelihood that closer sites are more similar). The causal variables best supported by the data were bird diversity (crows were more likely to decline where bird diversity was low), urbanization (crow declines were more likely in more urbanized areas), and crow density (declines were more likely where initial crow density was high). They found some evidence that areas hit earlier by WNV tended to show more pronounced crow declines, but there was little support for effects of total bird abundance, rainfall, or temperature. The patterns were spatially self-correlated.

Hamer *et al.* (2011) examined the relationship between bird diversity and WNV infection in *Culex pipiens* mosquitoes in Chicago, IL. They assessed mosquito abundance and infection prevalence, total bird diversity (richness and Shannon index), and the diversity of the birds from which the mosquitoes fed based on molecular diagnostic tests of blood in captured mosquitoes. These data allowed them to test the hypothesis that selective feeding by the mosquitoes was strong enough to weaken the effect of the diversity of available bird hosts. If feeding preferences are sufficiently strong, host availability will be less important. Hamer and colleagues estimated each bird species' contribution to the pool of infected mosquitoes by calculating host-specific "force of infection" (the squared fraction of total blood meals taken from a host species times that species' reservoir competence), which also allowed them to calculate "community force of infection" by summing across all host species present at a site. Of the 23 sites in which all data were collected, 11 had sufficient samples of bird-fed mosquitoes for statistical analysis of all factors. Although they found strong apparent host preferences overall, these preferences differed among sites, suggesting that preferences are not fixed but might vary with local host availability. In addition, it is important to note that true feeding preferences (i.e., the degree of bias in the distribution of attempted feedings compared to the availability of hosts) might not be revealed by analyzing only those mosquitoes that successfully fed. Successful feedings undoubtedly represent a subset of feeding attempts. If some hosts for mosquitoes are better able to defend themselves against mosquitoes, potentially causing mosquito mortality (as has been shown for ticks and their hosts: (Keesing *et al.*, 2009), then the distribution of successful feedings across host availability could distort true preferences considerably.

In Hamer and colleagues' study, diversity of birds available and diversity of birds represented in the blood-fed mosquitoes were not correlated. Hamer and colleagues used a maximum likelihood statistical approach to compare explanatory models of mosquito infection prevalence with WNV. Interestingly, despite apparently strong host preferences by mosquitoes, the best supported model invoked an interaction between bird diversity and the community force of infection. The majority of infected mosquitoes had fed on American robins, house sparrows, and house finches, species that tend to reach high abundance in low-diversity communities (Allan *et al.*, 2009). A plausible interpretation of these results is that selective feeding by mosquito vectors causes a multiplicative effect with diversity, whereby bird species with high reservoir competence are both preferred by mosquitoes and more likely to be abundant in low-diversity communities. In other words, the effect of feeding preferences compounds the effect of bird diversity. This interpretation reinforces the notion (Ostfeld and Keesing, 2000a; Ostfeld and LoGiudice, 2003; Schmidt and Ostfeld, 2001) that the nonrandom sequence by which species are lost as biodiversity declines is critical in understanding the relationship between biodiversity and disease risk.

Hantavirus Pulmonary Syndrome

The hantaviruses (family *Bunyaviridae*) appear to be common rodent-associated (or occasionally insectivore-associated) viruses that are widespread in Europe, Asia, North America, and South America. They appear to be nonpathogenic in their small-mammal hosts (Yates *et al.*, 2002), although more research is required to substantiate this. Virus particles can be shed at a high rate in the first few weeks after infection in the host's feces, urine, and saliva, after which the host's immune system often substantially reduces or eliminates viremia. Virus

is transmitted among individual small mammals during contacts such as biting, scratching, licking, and grooming. Cohabiting a burrow or nest with an infectious animal might also result in transmission to uninfected individuals. Humans can be exposed to hantaviruses if they inhale airborne particles of small-mammal excreta that contain virus particles. Such events appear to be uncommon and possible only when people dwell in heavily rodent-infested closed spaces such as cabins, sheds, or homes (Yates *et al.*, 2002). Although human diseases caused by hantaviruses are rare worldwide, they can be quite devastating. For instance, hantavirus pulmonary syndrome, which occurs in the Americas, has a case fatality rate of about 30% (Yates *et al.*, 2002), which is among the highest of any infectious disease.

Different types of hantavirus, usually named after the geographic location where they were first described, tend to attack different tissues in the human host. In Asia and Europe, hantavirus disease usually involves the kidneys or peripheral circulatory system, whereas in the Americas, the hantaviruses tend to cause pulmonary (lung) disease. Each type ("species") of hantavirus tends to be associated with only one species, or a group of closely related species, of reservoir host, although spillover to other species is possible. The list of rodent species that act as key reservoir hosts worldwide includes: the Norway rat (*R. norvegicus*), the black rat (*R. rattus*), the meadow vole (*Microtus pennsylvanicus*), the bank vole (*Myodes glareolus*), the deer mouse (*Peromyscus maniculatus*), the white-footed mouse (*Pe. leucopus*), the hispid cotton rat (*Sigmodon hispidus*), the rice rat (*Oryzomys palustris*), and the small vesper mouse (*Calomys laucha*). It is noteworthy that each of these species is a habitat generalist, geographically widespread, locally highly abundant, and resilient to anthropogenic disturbance. Consequently, these reservoir species tend to occur and even thrive in low-diversity animal communities, potentially leading to higher transmission risk when biodiversity is low.

This observation led Ostfeld and Keesing (2000a) to ask whether the dilution effect (higher disease transmission when diversity is low) operates in the hantavirus system. Ostfeld and Keesing analyzed the data published by Mills *et al.* (1998), who surveyed Sin Nombre hantavirus infection prevalence in deer mouse populations in nine national parks scattered throughout the USA. Mills and colleagues also surveyed small-mammal biodiversity in those parks. A negative correlation between small-mammal diversity and deer mouse infection prevalence would support a dilution effect. Although the relationship between these two variables appeared to be negative, the correlation was not statistically significant (Ostfeld and Keesing, 2000a). Later, Calisher *et al.* (2002) compared infection prevalence in deer mice and small-mammal biodiversity in three sites in Colorado. In one site, where hantavirus prevalence was 13%, only two species of small mammal were trapped. In a second site, where hantavirus prevalence was 9%, four small-mammal species occurred. In the third site, hantavirus prevalence was only 2% and 19 species of small mammals were present. More recently, Dizney and Ruedas (2009) asked whether mammalian diversity and hantavirus infection prevalence in deer mice were correlated in a set of five field sites around Portland, OR. Captures of about 5000 individuals of 21 species of small and medium-sized mammals demonstrated that the sites varied dramatically in their biodiversity. Dizney and Ruedas found strong support for the dilution effect. All sites that maintained moderate to high mammalian diversity were characterized by low (approximately 2%) hantavirus infection prevalence in deer mice, whereas the one site with low mammalian diversity had 14% infection prevalence in deer mice, a result that was highly statistically significant (**Figure 5**). In Montana, Carver *et al.* (2011) examined the relationship between prevalence of Sin Nombre hantavirus infection in deer mice and the presence and abundance of other rodents (voles) that are not hosts for this virus. They analyzed 15 years of live-trapping data on three field plots on which deer mice were always present but voles were only occasionally trapped. They found that the presence of voles was associated, on average, with an

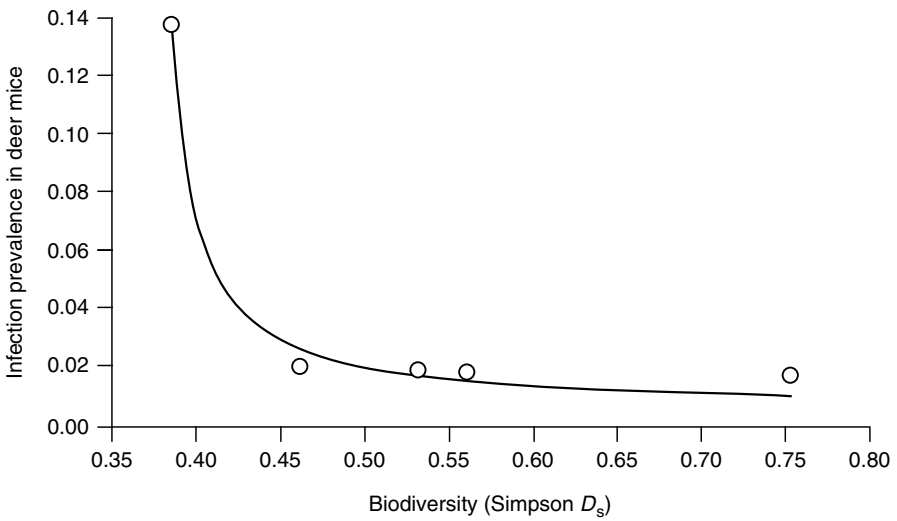


Figure 5 Prevalence of Sin Nombre hantavirus infection in populations of deer mice (*Peromyscus maniculatus*) as a function of mammal species diversity at sites in Oregon. Reprinted from Dizney LJ and Ruedas LA (2009) Increased host species diversity and decreased prevalence of Sin Nombre virus. *Emerging Infectious Diseases* 15: 1012–1018. doi: DOI 10.3201/eid1507.081083, with permission from CDC.

extraordinary 54–64% reduction in infection prevalence of deer mice. Deer mouse infection prevalence did not vary with vole population density; the mere presence of voles, even in low numbers, apparently was sufficient to suppress mouse infection rates. On the Channel Islands of southern California, Orrock *et al.* (2011) found that the species diversity of predators on small mammals (carnivores and raptors) was a highly significant predictor of hantavirus infection prevalence in deer mice.

Studying a European hantavirus (Puumala virus) in northern Belgium, Tersago *et al.* (2008) asked what factors might contribute to variation in infection prevalence in the reservoir host, the bank vole. Studying 14 sites over 2 years, Tersago and colleagues found that hantavirus infection prevalence varied from 0% to 16%, with vole populations at <20 individual per hectare having no detectable hantavirus. Where hantavirus was present, infection prevalence in bank voles was strongly, negatively correlated with the relative abundance of a nonreservoir host, the wood mouse (*Apodemus sylvaticus*), which was the only other common rodent at these sites. Once again, these results strongly support a dilution effect. Piudo *et al.* (2011) assessed the causes of variation in infection prevalence of an Argentine hantavirus, Andes virus, in its reservoir host, *Oligoryzomys longicaudatus* in both peridomestic and forest habitats of Patagonia. They found that *Ol. longicaudatus* was the numerically dominant small-mammal species in peridomestic habitats with low diversity but not in "sylvan" ones, supporting a major tenet of the dilution effect (that principal reservoirs tend to predominate in disturbed, low diversity habitats). Piudo and colleagues found that the number of infected *Ol. longicaudatus* was negatively correlated with small-mammal diversity during the same month and in the prior 2 months; however, they did not find a significant correlation between mammal diversity and infection prevalence in the reservoir host.

In all of the above-described studies of the relationship between biodiversity and hantavirus infection prevalence, the study design was correlational and no experimental manipulations were imposed. Cause and effect relationships can be postulated but rarely rigorously tested with such study designs, and mechanisms that underlie the putative causes of reduced infection prevalence are difficult to elucidate. But substantial progress has been made in understanding the mechanisms underlying the dilution effect in hantaviruses. The Dizney and Ruedas (2009) study did not find that higher mammalian diversity was associated with lower abundance of deer mouse reservoirs, nor that hantavirus infection prevalence was positively correlated with deer mouse abundance. They therefore conclude that the "susceptible host regulation" mechanism was not supported but that the "encounter reduction" mechanism was supported (see **Figure 1**). Studying hantavirus dynamics in Utah, Clay *et al.* (2009b, c) and Lehmer *et al.* (2008) found a somewhat more complex set of relationships to be influencing infection prevalence in deer mice. Among 16 large field sites, Clay and colleagues found a negative correlation between nocturnal mammal diversity and infection prevalence of deer mice with Sin Nombre virus, supporting the dilution effect. They found that both the population density of deer mice and the average survival (persistence) of individual mice were lower where diversity was higher. But infection prevalence was negatively correlated only with deer mouse survival and not with deer mouse density (Clay *et al.*, 2009a, d; Lehmer *et al.*, 2008). When they monitored encounter rates and durations between mice in foraging plots, Clay *et al.* (2009d) found that high mammalian diversity was significantly correlated with reduced frequency of encounters between mice, supporting the encounter reduction mechanism of Keesing *et al.* (2006). The weight of evidence for the hantavirus pulmonary syndrome system in western North America suggests that high diversity of mammals causes deer mice to increase their frequency of encounter with heterospecifics at the cost of encounters with conspecifics. Heterospecific (interspecific) encounters typically do not result in hantavirus transmission, whereas hantavirus transmission rates correlate positively with the frequency of intraspecific encounters.

The role of biodiversity in the transmission dynamics of hantaviruses was recently studied experimentally by Suzan *et al.* (2009) in the Azuero Peninsula of southwestern Panama, where an outbreak of hantavirus pulmonary syndrome killed $>20\%$ of the ca 50 human victims (Ruedas *et al.*, 2004). Two hantaviruses had been described in this region, Choclo virus and Calabazo virus, although only the former is known to be pathogenic in humans. The natural reservoirs of Choclo and Calabazo virus are, respectively, the pigmy rice rat (*Oligoryzomys fulvescens*) and the cane mouse (*Zygodontomys brevicauda*) (Ruedas *et al.*, 2004). Suzan and colleagues established 24 study sites in patches of tropical forest in the vicinity of the outbreak of human cases. In all sites they established live-trapping grids to assess the abundance and diversity of all small mammals present, using standard mark-recapture techniques. They designated 16 of the sites as "removal" plots, from which they permanently removed all species except pigmy rice rats and cane mice; rice rats and cane mice were marked and released after drawing a blood sample, but were otherwise unmanipulated. On the remaining eight sites – the control sites – the investigators marked and released all animals captures. This design was intended to experimentally reduce small-mammal diversity according to patterns observed in nature, in which diverse communities contain rice rats, cane mice, and many other species of rodents and small marsupials, and species-poor communities contain rice rats, cane mice, and few other species (Ruedas *et al.*, 2004; Suzan *et al.*, 2008a, b). In the 16 experimental plots from which nonreservoir species were removed, two dramatic effects were observed. First, population densities of the two reservoir species were significantly increased relative to their densities on the control plots (where no removals took place). Second, the removal of nonreservoir species was associated with a dramatic increase in the proportion of the reservoir populations that was infected with hantavirus. In the control plots, about 8–12% of rice rats and cane mice were infected, whereas on the plots from which nonreservoirs were removed, infection prevalences of rice rats and cane mice were about three times higher, about 35%. Thus, this experimental manipulation of diversity supports two mechanisms underlying the dilution effect. The increased population density of reservoir species where other species were removed suggests that high mammalian diversity regulates the abundance of reservoir hosts. And the increased infection prevalence in reservoir populations suggests that high diversity reduces encounter rates

between susceptible and infected members of the reservoir populations that can result in transmission.

Schistosomiasis

Schistosomiasis is a parasitic disease caused by a trematode (fluke) that uses freshwater snails as an intermediate host and a vertebrate as a definitive host (defined as the host in which sexual reproduction takes place). The disease is endemic to 74 countries and affects more than 200 million people worldwide, almost exclusively in tropical and subtropical countries (Committee, 2002). Infected humans release schistosome eggs when they urinate or defecate near freshwater sources, and these eggs hatch into a free-swimming form called miracidia. Miracidia orient toward freshwater snails, burrowing in through the snail's foot, and once inside this intermediate host they replicate rapidly, producing thousands of new parasitic forms called cercariae. Infected snails release dozens of cercariae each day, and these cercariae swim about seeking a suitable vertebrate host. The trematode that causes most schistosomiasis cases in humans, *Schistosoma mansoni*, uses snails in the genus *Biomphalaria* as the intermediate host. *Biomphalaria* snails can coexist with other species of freshwater snails and these snails can attract schistosome miracidia, but non-*Biomphalaria* snails tend not to become infected and in fact can regulate miracidia by consuming them.

Johnson *et al.* (2009) conducted an experimental laboratory study to test whether increased diversity of snail species reduced human risk of exposure to schistosomiasis. Using 200 experimental containers, Johnson and colleagues housed *Biomphalaria glabrata* snails either alone or together with other snail species that do not support replication by miracidia. The numbers of *Bi. glabrata* snails was held constant among containers, some of which also contained either *Helisoma trivolvis* or *Lymnaea stagnalis* snails, or both. After adding known concentrations of *Sc. mansoni* miracidia to the containers, Johnson and colleagues monitored the proportion of snails infected and the numbers of cercariae released by the snails, both in total and per snail. Cercarial release rates are strong predictors of human risk. In containers with two or three species of snail (*Biomphalaria* plus one or both of the other snails), both the proportion of snails infected and the cercarial release rates were strongly reduced as compared to low diversity containers containing *Biomphalaria* alone. *Biomphalaria* snails raised alone released, on average, between two and five times as many schistosome cercariae as did the snails raised together with one or two other species (**Figure 6**). The mechanism by which *Helisoma* and *Lymnaea* snails reduced disease risk was the deflection of miracidia away from *Biomphalaria*, in which they replicate rapidly, to these other snail hosts, which act as dead ends. Thus, high snail diversity reduces disease risk via the encounter reduction mechanism.

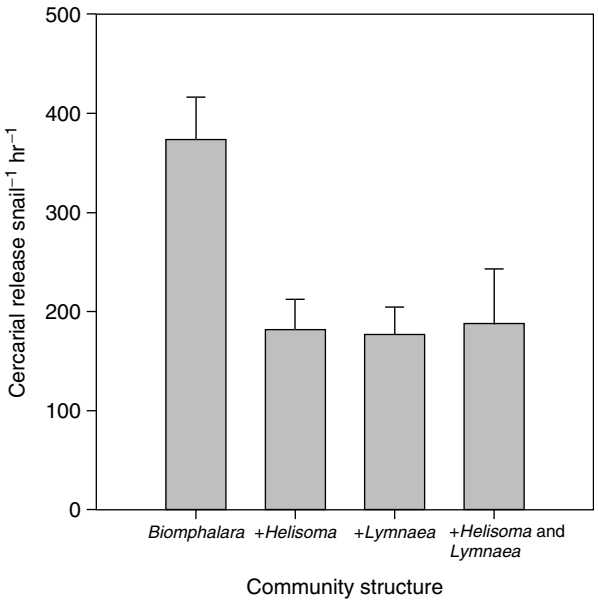


Figure 6 The rate of release of infective stages of *Schistosoma mansoni* parasites by their main intermediate hosts, *Biomphalaria* snails. *Biomphalaria* snails were raised either alone or together with one or two other snail species. Increased diversity of snails strongly reduced the release of infective cercariae. Reprinted from Johnson PTJ, Lund PJ, Hartson RB, and Yoshino TP (2009) Community diversity reduces *Schistosoma mansoni* transmission, host pathology and human infection risk. *Proceedings of the Royal Society B-Biological Sciences* 276: 1657–1663. doi: DOI 10.1098/rspb.2008.1718, with permission from Royal Society of Publishing.

Global Scale Patterns of Diversity and Human Infectious Diseases

The previous section, Biodiversity and Infectious Disease, describes theory and data that demonstrate a negative effect of high biodiversity on the transmission of infectious agents and therefore risk or incidence of infectious disease. An alternative perspective has recently been explored by researchers asking whether global patterns of human infectious diseases are positively correlated with global patterns of biodiversity. One of the best-established patterns in the field of ecology is the latitudinal gradient in species diversity, such that for most taxa, more species occur in low-latitude areas than in higher-latitude areas of similar size (Pianka, 1966). Recently, researchers have asked whether this same pattern pertains to parasites and pathogens of humans. It is plausible that the greater the diversity of human pathogens, the greater the burden of disease. Moreover, because many infectious agents of humans are or were zoonotic, it is possible that disease risk will be higher where vertebrate biodiversity (representing the sources of zoonotic pathogens) is higher. We might therefore expect a higher burden of infectious disease in the tropics, where pathogen and zoonotic host diversity are higher, than at higher latitudes. Such hypotheses provide an apparent contrast to the dilution effect, because they emphasize the possibility of a positive correlation between diversity and disease risk.

Recently, Guernier *et al.* (2004) argued that, like their free-living counterparts, parasitic and infectious agents of humans also exhibit their highest levels of diversity at low latitudes. Many of the species present at low latitudes also occur at higher latitudes, with relatively few species occurring only at higher latitudes, leading to a geographically nested pattern of these infectious agents. The same general pattern pertained to

viral, bacterial, fungal, protozoal, and helminthic pathogens. Guernier and colleagues postulated that the latitudinal gradient in pathogen species richness was caused by a gradient in the maximum range in precipitation, rather than by total precipitation or average temperature. The question remains, however, whether human disease risk is higher where the diversity of potential pathogens is higher.

Dunn *et al.* (2010) addressed this question by regressing the prevalence of human diseases caused by helminths, protozoal parasites, and two bacterial pathogens (22 diseases in total, with prevalence categorized as endemic, sporadic, or nonendemic) in various political units against the species richness of viral, bacterial, protozoan, and helminthic pathogens (347 pathogens in total) in similar political units. The disparate taxonomies and sample sizes used in the two variables were a consequence of data availability. They found a positive relationship between disease prevalence and pathogen species richness. Dunn and colleagues further found a positive association between total pathogen species richness and the species richness of birds and mammals, which act as zoonotic hosts. Using structural equation modeling, Dunn and colleagues concluded that bird and mammal species richness drives richness of human pathogens, which in turn drives the prevalence of the 22 diseases in their database.

In contrast, Jones *et al.* (2008) compiled a dataset consisting of 335 infectious diseases of humans having emerged after 1940, and found that, "The highest concentration of emerging infectious disease (EID) events per million square kilometres of land was found between 30 and 60 degrees north and between 30 and 40 degrees south, with the main hotspots in the northeastern United States, western Europe, Japan and southeastern Australia." (Jones *et al.*, 2008, p. 991). In other words, emergence of infectious human diseases is concentrated in the temperate zone and not in the tropics. Jones and colleagues were concerned that the geographic pattern of EIDs could be biased by geographic variation in the abundance or activity of scientists likely to detect and report these diseases. They attempted to account statistically for this suspected "reporting bias" by including as an independent variable in their spatial analyses the frequency of each country listed as the address for each author in all papers published in the *Journal of Infectious Diseases* since 1973. When they accounted for reporting bias in this fashion, Jones and colleagues found that the emergence of drug-resistant pathogens and all zoonotic pathogens from nonwildlife remained more likely at higher than lower latitudes, whereas emergence of zoonotic pathogens from wildlife and vector-borne pathogens were not significantly associated with latitude. Jones and colleagues found that wildlife species richness was significantly, positively correlated with emergence events of zoonotic pathogens from wildlife hosts, but their analysis indicated that the effect was very weak (odds ratio of 1.01, or a 1% increase in emergence probability over the range of species richness) (**Figure 7**).

From these studies, it is difficult to draw strong conclusions about geographic associations between wildlife species richness, species richness of human pathogens, and either the probability of disease emergence or prevalence of human disease. A positive relationship between wildlife species richness and species richness of zoonotic pathogens seems well supported, but whether higher pathogen richness leads to greater incidence of human diseases – emerging or established – is not yet resolved. Poor resolution might be partly due to the disease data used in these analyses. The variables used range from crude categories of prevalence (endemic, sporadic, or nonendemic) to counting instances of emergence. Owing to data limitations, more relevant quantitative data on incidence rates (cases per capita per unit time) or risk (probability of exposure as represented by abundance of infected hosts or

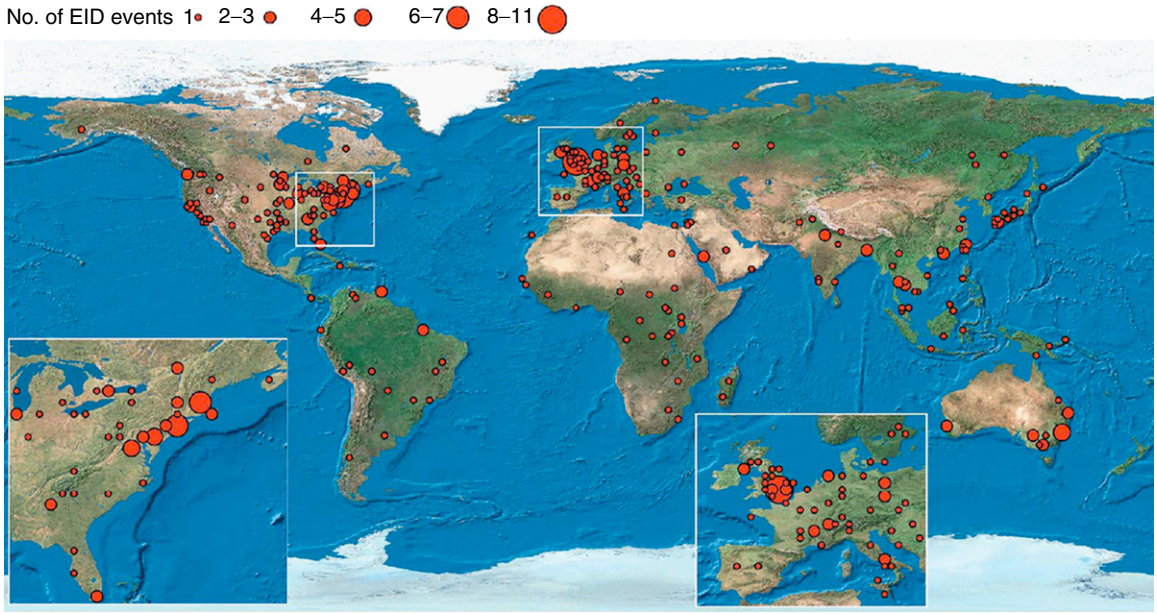


Figure 7 Number of emerging infectious disease events compiled by Jones *et al.* (2008), with size of circle proportional to the number of events. See Jones and colleagues for details on how EID events are defined and located.

vectors) are not available at these broad geographic scales. Similarly, biodiversity data are often crude and perhaps inappropriate for some analyses. Diversity data typically consist of counts of species (species richness), rather than containing information on evenness or species composition. Moreover, they are generally obtained by superimposing various species' range maps, which account for species presence or absence at extremely coarse scales, over political units in which disease data occur over quite different scales.

A critical source of confusion in the literature on diversity-disease relationships is the existence of variable metrics of both diversity and disease. As argued above from both theoretical and empirical bases, species richness alone is expected to contain relatively little information of use in predicting pathogen transmission (Ostfeld and Keesing, 2000a; Schmidt and Ostfeld, 2001). Metrics that include species richness and evenness provide more information, but the changes in species composition that accompany changes in richness and evenness are required in many cases to understand impacts on pathogen transmission (Keesing *et al.*, 2006, 2010). Similarly, counts of emergence events provide relatively little information on the true burden of infectious disease. More information is contained in crude prevalence metrics such as endemic or sporadic disease. But ecological risk (the abundance of infected hosts or vectors) is a more precise metric of the probability of exposure, holding human behavior constant, and incidence rates are the foundation of understanding true exposure. In addition, coarse-scale geographic analyses of correlations between the size of the potential species pool representing pathogen diversity and disease emergence or prevalence do not by themselves inform biodiversity policy or management to reduce disease risk. At best, geographic gradients in pathogen diversity provide the upper bound of how many infectious diseases human populations might be expected to encounter in a given location. But these gradients do not tell us how we might expect disease risk to change when human activities change biodiversity.

Concluding Comments

Biodiversity affects human health in myriad ways and in some cases the relationship is complex. In a broad sense, human health relies fundamentally on biodiversity, because virtually all elements necessary for health, including food, medicine, shelter, and energy, require diverse life forms. To focus the discussion of biodiversity-human health interactions, it is useful to ask how *changes* in biodiversity affect *changes* in human health. Considerable information now exists to describe how various human activities are both reducing and changing the composition of biodiversity in virtually all ecosystems on earth. This essay has reviewed our knowledge of how those changes affect basic biomedical research, preventative and therapeutic agents, nutrition, and infectious disease. This knowledge is required for sound policy and management decisions that will affect future changes in biodiversity.

Strong evidence indicates that species of animals, plants, and various microbes that either currently do or potentially will contribute importantly and perhaps uniquely to basic biomedical research are threatened with extinction. Similarly, many species with the potential to provide important medicines and foods are currently threatened, and many have probably already disappeared without documentation. To help advance policy that will maximally protect human health benefits, we need to pursue the relationship between the likelihood that species will be valuable as foods or medicines and their vulnerability to decline and extinction.

Detailed studies of several important zoonotic diseases strongly support a protective role of high biodiversity. Reduced rates of pathogen transmission and disease risk that accompany high diversity have been documented in infectious diseases of humans, wildlife, and plants (Keesing *et al.*, 2010). The generality of this pattern appears to result from the following observations: (1) different species within host communities play dramatically different roles in supporting transmission of pathogens, with some amplifying and others suppressing pathogens. Similarly, hosts can differ dramatically in how they affect survival and population dynamics of arthropod vectors of those pathogens; and (2) the host species most likely to support pathogen replication and transmission tend to be ecologically resilient species that persist or even thrive in disturbed habitats that lose biodiversity. The host species most likely to suppress pathogen transmission tend to be less resilient and therefore to disappear preferentially as human activities reduce biodiversity. The mechanisms that underlie the correlation between ecological resilience and host quality are only beginning to be assessed.

In some cases, it is plausible that biogeographic areas with high native biodiversity contain more potential sources of zoonotic pathogens, such that risk of infectious diseases emerging is higher in more diverse than in less diverse places. However, once a pathogen has emerged, the body of scientific studies suggests that transmission will be inhibited by high diversity. Moreover, if the potential sources of zoonotic pathogens tend to be the species that are resilient to anthropogenic disturbances, then the loss of biodiversity will be likely to increase the probability of emergence anywhere along a biogeographic gradient in native biodiversity.

See also: Biodiversity and Ecosystem Services. Biodiversity, Human Well-Being, and Markets. Bioprospecting. Deforestation and Land Clearing. Diseases, Conservation and. Global Declines of Amphibians. Global Species Richness. Habitat Loss and Fragmentation. Human Impact on Biodiversity, Overview. Landscape Epidemiology. Loss of Biodiversity, Overview. Parasitism. Pharmacology, Biodiversity and. Species–Area Relationships. Valuing Ecosystem Services

Appendix

List of Courses

1. Conservation Biology
2. Ecology
3. Epidemiology
4. Global Change

References

Allan BF, Keesing F, and Ostfeld RS (2003) Effect of forest fragmentation on Lyme disease risk. *Conservation Biology* 17: 267–272.

Allan BF, Langerhans RB, Ryberg WA, *et al.* (2009) Ecological correlates of risk and incidence of West Nile virus in the United States. *Oecologia* 158: 699–708. http://dx.doi.org/10.1007/s00442-008-1169-9.

Altieri MA (1999) The ecological role of biodiversity in agroecosystems. *Agriculture Ecosystems & Environment* 74: 19–31.

Brownstein JS, Skelly DK, Holford TR, and Fish D (2005) Forest fragmentation predicts local scale heterogeneity of Lyme disease risk. *Oecologia* 146: 469–475. http://dx.doi.org/10.1007/s00442-005-0251-9.

Brumpt E (1944–1945) Anophelisme sans paludisme et regression spontanee du paludisme. *Annales de parasitologie* 20: 67–91.

Brunner JL, Cheney L, Keesing F, *et al.* (2011) Molting Success of Ixodes scapularis Varies Among Individual Blood Meal Hosts and Species. *Journal of Medical Entomology* 48: 860–866. http://dx.doi.org/10.1603/Me10256.

Brunner JL and Ostfeld RS (2008) Multiple causes of variable tick burdens on small-mammal hosts. *Ecology* 89: 2259–2272.

Calisher CH, Root JJ, Mills JN, and Beaty BJ (2002) Assessment of ecologic and biologic factors leading to hantavirus pulmonary syndrome, Colorado, USA. *Croatian Medical Journal* 43: 330–337.

Carver S, Kuenzi A, Bagamian KH, *et al.* (2011) A temporal dilution effect: Hantavirus infection in deer mice and the intermittent presence of voles in Montana. *Oecologia* 166: 713–721. http://dx.doi.org/10.1007/s00442-010-1882-z.

Chivian E and Bernstein A (2008) Threatened groups of organisms valuable to medicine. In: Chivian E and Bernstein A (eds.) *Sustaining Life: How Human Health Depends on Biodiversity*, pp. 203–286. New York: Oxford University Press.

Clay CA, Lehmer EM, Jeor SS, and Dearing MD (2009a) Sin Nombre virus and rodent species diversity: A test of the dilution and amplification hypotheses. *PLoS ONE* 4: e6467.

Clay CA, Lehmer EM, Jeor SS, and Dearing MD (2009b) Sin Nombre virus and rodent species diversity: A test of the dilution and amplification hypotheses. *PLoS ONE* 4. doi: Artn E6467 Doi 10.1371/Journal.Pone.0006467.

Clay CA, Lehmer St EM, Jeor S, and Dearing MD (2009c) Testing mechanisms of the dilution effect: Deer mice encounter rates, Sin Nombre virus prevalence and species diversity. *Ecohealth* 6: 250–259. http://dx.doi.org/10.1007/s10393-009-0240-2.

Clay CA, Lehmer EM, Previtali A, Jeor SS, and Dearing MD (2009d) Contact heterogeneity in deer mice: Implications for Sin Nombre virus transmission. *Proceedings of the Royal Society B-Biological Sciences* 276: 1305–1312. http://dx.doi.org/10.1098/rspb.2008.1693.

Committee WHOE (2002) Prevention and control of schistosomiasis and soil-transmitted helminthiasis. In: W. H. Organization (ed.) *WHO Technical Report Series* 912, pp. 57. Geneva: World Health Organization.

DeClerck FAJ, Fanzo J, Palm C, and Remans R (2011) Ecological approaches to human nutrition. *Food and Nutrition Bulletin* 32: S41–S50.

Dizney LJ and Ruedas LA (2009) Increased host species diversity and decreased prevalence of Sin Nombre virus. *Emerging Infectious Diseases* 15: 1012–1018. http://dx.doi.org/10.3201/eid1507.081083.

Dobson A (2004) Population dynamics of pathogens with multiple host species. *American Naturalist* 164: S64–S78.

Dunn RR, Davies TJ, Harris NC, and Gavin MC (2010) Global drivers of human pathogen richness and prevalence. *Proceedings of the Royal Society B-Biological Sciences* 277: 2587–2595. http://dx.doi.org/10.1098/rspb.2010.0340.

Elton CS (1958) *The Ecology of Invasions by Animals and Plants*. London: Methuen & Co.

Ezenwa VO, Godsey MS, King RJ, and Guptill SC (2006) Avian diversity and West Nile virus: Testing associations between biodiversity and infectious disease risk. *Proceedings of the Royal Society B-Biological Sciences* 273: 109–117. http://dx.doi.org/10.1098/rspb.2005.3284.

Guernier V, Hochberg ME, and Guegan JFO (2004) Ecology drives the worldwide distribution of human diseases. *PloS Biology* 2: 740–746. doi: ARTN e141 DOI 10.1371/journal/pbio.0020141.

Hamer GL, Chaves LF, Anderson TK, *et al.* (2011) Fine-scale variation in vector host use and force of infection drive localized patterns of West Nile virus transmission. *PLoS ONE* 6. doi: ARTN e23767 DOI 10.1371/journal.pone.0023767.

Hillel D and Rosenzweig C (2008) Biodiversity and food production. In: Chivian E and Bernstein A (eds.) *Sustaining Life: How Human Health Depends on Biodiversity*, pp. 325–382. New York: Oxford University Press.

Jackson LE, Hilborn ED, and Thomas JC (2006) Towards landscape design guidelines for reducing Lyme disease risk. *International Journal of Epidemiology* 35: 315–322. http://dx.doi.org/10.1093/ije/dyi284.

Johnson PTJ, Lund PJ, Hartson RB, and Yoshino TP (2009) Community diversity reduces *Schistosoma mansoni* transmission, host pathology and human infection risk. *Proceedings of the Royal Society B-Biological Sciences* 276: 1657–1663. http://dx.doi.org/10.1098/rspb.2008.1718.

Johnson PTJ and Thieltges DW (2010) Diversity, decoys and the dilution effect: How ecological communities affect disease risk. *Journal of Experimental Biology* 213: 961–970. http://dx.doi.org/10.1242/Jeb.037721.

Jones KE, Patel NG, Levy MA, *et al.* (2008) Global trends in emerging infectious diseases. *Nature* 451: 990–993. http://dx.doi.org/10.1038/Nature06536.

Keesing F, Belden LK, Daszak P, *et al.* (2010) Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468: 647–652. http://dx.doi.org/10.1038/Nature09575.

Keesing F, Brunner J, Duerr S, *et al.* (2009) Hosts as ecological traps for the vector of Lyme disease. *Proceedings of the Royal Society B-Biological Sciences* 276: 3911–3919. http://dx.doi.org/10.1098/rspb.2009.1159.

Keesing F, Holt RD, and Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecology Letters* 9: 485–498. http://dx.doi.org/10.1111/j.1461-0248.2006.00885.x.

Keirans JE, Hutcheson HJ, Durden LA, and Klompen JSH (1996) *Ixodes scapularis* (Acari: Ixodidae): Redescription of all active stages, distribution, hosts, geographical variation, and medical and veterinary importance. *Journal of Medical Entomology* 33: 297–318.

Kilpatrick AM, Daszak P, Jones MJ, Marra PP, and Kramer LD (2006a) Host heterogeneity dominates West Nile virus transmission. *Proceedings of the Royal Society B-Biological Sciences* 273: 2327–2333. http://dx.doi.org/10.1098/rspb.2006.3575.

Kilpatrick AM, Kramer LD, Jones MJ, Marra PP, and Daszak P (2006b) West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. *PloS Biology* 4: 606–610. doi: Artn E82 Doi 10.1371/Journal.Pbio.0040082.

Kilpatrick AM, LaDeau SL, and Marra PP (2007) Ecology of West Nile virus transmission and its impact on birds in the western hemisphere. *Auk* 124: 1121–1136.

Klein AM, Vaissiere BE, Cane JH, *et al.* (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B-Biological Sciences* 274: 303–313. http://dx.doi.org/10.1098/rspb.2006.3721.

Koenig WD, Hochachka WM, Zuckerberg B, and Dickinson JL (2010) Ecological determinants of American crow mortality due to West Nile virus during its North American sweep. *Oecologia* 163: 903–909. http://dx.doi.org/10.1007/s00442-010-1627-z.

Komar N, Langevin S, Hinten S, *et al.* (2003) Experimental infection of north American birds with the New York 1999 strain of West Nile virus. *Emerging Infectious Diseases* 9: 311–322.

Kremen C and Ostfeld RS (2005) A call to ecologists: Measuring, analyzing, and managing ecosystem services. *Frontiers in Ecology and the Environment* 3: 540–548.

Kunz T, Braun de Torrez E, Bauer D, Lobova T, and Fleming. TH (2011) Ecosystem services provided by bats. *Annals of the New York Academy of Sciences* 1223: 1–38.

LaDeau SL, Kilpatrick AM, and Marra PP (2007) West Nile virus emergence and large-scale declines of North American bird populations. *Nature* 447: 710–713. http://dx.doi.org/10.1038/Nature05829.

LaDeau SL, Marra PP, Kilpatrick AM, and Calder CA (2008) West Nile virus revisited: Consequences for North American ecology. *Bioscience* 58: 937–946. http://dx.doi.org/10.1641/B581007.

Lehmer EM, Clay CA, Pearce-Duvet J, Jeor SS, and Dearing MD (2008) Differential regulation of pathogens: The role of habitat disturbance in predicting prevalence of Sin Nombre virus. *Oecologia* 155: 429–439. http://dx.doi.org/10.1007/s00442-007-0922-9.

LoGiudice K, Duerr STK, Newhouse MJ, Schmidt KA, Killilea ME, and Ostfeld RS (2008) Impact of host community composition on Lyme disease risk. *Ecology* 89: 2841–2849.

LoGiudice K, Ostfeld RS, Schmidt KA, and Keesing F (2003) The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. *Proceedings of the National Academy of Sciences of the United States of America* 100: 567–571. http://dx.doi.org/10.1073/pnas.0233733100.

Marra PP, Griffing S, Caffrey C, *et al.* (2004) West Nile virus and wildlife. *Bioscience* 54: 393–402.

Mills JN, Johnson JM, Ksiazek TG, *et al.* (1998) A survey of hantavirus antibody in small-mammal populations in selected United States National Parks. *American Journal of Tropical Medicine and Hygiene* 58: 525–532.

Newman DJ, Kilama J, Bernstein A, and Chivian E (2008) Medicines from nature. In: Chivian E and Bernstein A (eds.) *Sustaining Life: How Human Health Depends on Biodiversity*, pp. 117–162. New York: Oxford University Press.

Norman R, Bowers RG, Begon M, and Hudson PJ (1999) Persistence of tick-horne virus in the presence of multiple host species: Tick reservoirs and parasite mediated competition. *Journal of Theoretical Biology* 200: 111–118.

Organization W.H. (1982) *Manual on Environmental Management for Mosquito Control with Special Emphasis on Mosquito Vectors*. Geneva, Switzerland: WHO Offset Publication.

Orrock JL, Allan BF, and Drost CA (2011) Biogeographic and ecological regulation of disease: Prevalence of Sin Nombre virus in island mice is related to island area, precipitation, and predator richness. *American Naturalist* 177: 691–697. http://dx.doi.org/10.1086/659632.

Ostfeld R and Keesing F (2000a) The function of biodiversity in the ecology of vector-borne zoonotic diseases. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* 78: 2061–2078.

Ostfeld RS (2009) Biodiversity loss and the rise of zoonotic pathogens. *Clinical Microbiology and Infection* 15: 40–43.

Ostfeld RS (2011) *Lyme Disease: The Ecology of a Complex System*. New York: Oxford University Press.

Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, and Keesing F (2006) Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. *Plos Biology* 4: 1058–1068. doi: Artn E145 Doi 10.1371/Journal.Pbio.0040145.

Ostfeld RS and Keesing F (2000b) Biodiversity and disease risk: The case of Lyme disease. *Conservation Biology* 14: 722–728.

Ostfeld RS and LoGiudice K (2003) Community disassembly, biodiversity loss, and the erosion of an ecosystem service. *Ecology* 84: 1421–1427.

Pianka ER (1966) Latitudinal gradients in species diversity: A review of concepts. *The American Naturalist* 100: 33–46.

Piudo L, Monteverde MJ, Walker RS, and Douglass RJ (2011) Rodent community structure and Andes virus infection in Sylvan and Peridomestic Habitats in Northwestern Patagonia, Argentina. *Vector-Borne and Zoonotic Diseases* 11: 315–324. http://dx.doi.org/10.1089/vbz.2009.0242.

Pongsiri MJ, Roman J, Ezenwa VO, *et al.* (2009) Biodiversity loss affects global disease ecology. *Bioscience* 59: 945–954. http://dx.doi.org/10.1525/bio.2009.59.11.6.

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, and Kunin WE (2010) Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology & Evolution* 25: 345–353. http://dx.doi.org/10.1016/j.tree.2010.01.007.

Rosenthal JP and Preszler T (2002) Biodiversity in biomedical research. In: Aguirre AA, Ostfeld RS, Tabor GM, House C, and Pearl MC (eds.) *Conservation Medicine: Ecological Health in Practice*, pp. 327–344. New York: Oxford University Press.

Ruedas LA, Salazar-Bravo J, Tinnin DS, *et al.* (2004) Community ecology of small mammal populations in Panama following an outbreak of Hantavirus pulmonary syndrome. *Journal of Vector Ecology* 29: 177–191.

Schmidt KA and Ostfeld RS (2001) Biodiversity and the dilution effect in disease ecology. *Ecology* 82: 609–619.

Service MW (1991) Agricultural-Development and Arthropod-Borne Diseases - a Review. *Revista De Saude Publica* 25: 165–178.

Shaw MT, Keesing F, McGrail R, and Ostfeld RS (2003) Factors influencing the distribution of larval blacklegged ticks on rodent hosts. *American Journal of Tropical Medicine and Hygiene* 68: 447–452.

Slowik TJ and Lane RS (2009) Feeding preferences of the immature stages of three western North American Ixodid ticks (Acari) for avian, reptilian, or rodent hosts. *Journal of Medical Entomology* 46: 115–122.

Suzan G, Armien A, Mills JN, *et al.* (2008a) Epidemiological considerations of rodent community composition in fragmented landscapes in Panama. *Journal of Mammalogy* 89: 684–690.

Suzan G, Marce E, Giermakowski JT, *et al.* (2008b) The effect of habitat fragmentation and species diversity loss on Hantavirus prevalence in Panama. *Animal Biodiversity and Emerging Diseases: Prediction and Prevention* 1149: 80–83. http://dx.doi.org/10.1196/annals.1428.063.

Suzan G, Marce E, Giermakowski JT, *et al.* (2009) Experimental evidence for reduced rodent diversity causing increased hantavirus prevalence. *PLoS ONE* 4: e5461.

Swaddle JP and Calos SE (2008) Increased Avian diversity is associated with lower incidence of human West Nile infection: Observation of the dilution effect. *PLoS ONE* 3: e2488.

Tersago K, Schreurs A, Linard C, Verhagen R, Van Dongen S, and Leirs H (2008) Population, environmental, and community effects on local bank vole (*Myodes glareolus*) Puumala virus infection in an area with low human incidence. *Vector-Borne and Zoonotic Diseases* 8: 235–244. http://dx.doi.org/10.1089/vbz.2007.0160.

Vanbuskirk J and Ostfeld RS (1995) Controlling Lyme-disease by modifying the density and species composition of tick hosts. *Ecological Applications* 5: 1133–1140.

Whelan CJ, Wenny DG, and Marquis RJ (2008) Ecosystem services provided by birds. *Year in Ecology and Conservation Biology* 1134: 25–60. http://dx.doi.org/10.1196/anndis.1439.003.

Winfree R (2010) The conservation and restoration of wild bees. *Year in Ecology and Conservation Biology* 1195: 169–197. http://dx.doi.org/10.1111/j.1749-6632.2010.05449.x.

Yates TL, Mills JN, Parmenter CA, *et al.* (2002) The ecology and evolutionary history of an emergent disease: Hantavirus pulmonary syndrome. *Bioscience* 52: 989–998.

Zhu YY, Chen HR, Fan JH, *et al.* (2000) Genetic diversity and disease control in rice. *Nature* 406: 718–722.

Biodiversity and Pest Control Services

Stacy M Philpott, University of Toledo, Toledo, OH, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Agricultural intensification A process of change in crop and noncrop vegetation and management practices on farms. Changes in vegetation include reducing the numbers of crop species and varieties, and limiting trees, trap crops, and weeds. Other local changes in management include increasing the application of chemical pesticides and fertilizers, increased tillage and irrigation, and heavier mechanization. At the landscape level, intensification includes converting natural habitat to crop fields, destroying edge habitats, simplifying landscapes, avoiding fallows, and fragmenting natural habitat.

Agrobiodiversity Typically refers to the level of species diversity (with some research in this context on genetic diversity or diversity among communities and ecosystems). At the species level, agrobiodiversity refers to species richness, species evenness, and community composition (or varietal richness, evenness, and composition) of both planned and unplanned vegetation, and all other organisms on a farm and in farming landscapes.

Biological control The use of naturally occurring or introduced natural enemies (e.g., predators and parasitoids) to suppress the populations of, and damage caused by, pests.

Complementarity Resource partitioning among consumer species that results in utilization of a greater extent of available resources.

Intraguild predation Species interaction where predators feed on herbivores as well as other predators, and where trophic levels are not well defined.

Introduction

A major focus of agricultural research is investigating the relationship between biodiversity and pest control. As natural systems are converted to agriculture, provisioning of some vital ecosystem services can decline (Kremen *et al.*, 2004). Ecosystem services, such as pest control, are ecosystem processes that improve and sustain human life (Daily, 1997). In agricultural systems, pest control problems may be exacerbated by biodiversity loss. Pests, defined as noxious weeds, insects, mites, and fungi, cause economic damage to crop plants (Hill, 1987). Many different forms of pest control are used, including cultural control, mechanical control, chemical control, and biological control. Biological pest control involves attempts to use natural enemies (e.g., predators and parasitoids) to suppress the populations of, and damage caused by, pests. Biological pest control has been used for centuries. Early records from ancient China (near 300 AD) indicate that farmers used ants as natural enemies in orange groves to control mite populations (Huang and Yang, 1987). Classical biological control has focused on introduction of species into agricultural areas in order to reduce pests to below economic damage levels. By the 1980s, 160 species of predatory arthropods and 16 insectivorous birds had been released for pest control in the USA (Letourneau *et al.*, 2009). To date, more than 2000 species have been released worldwide (van Lenteren *et al.*, 2006).

With agricultural intensification, use of biological methods of pest control has declined, being replaced with chemical control and the use of genetically modified organisms to combat pests (Benbrook, 2001). But such changes incur environmental costs and exacerbate biodiversity losses in agricultural landscapes (Bengtsson *et al.*, 2005; Relyea, 2005). Pesticides can cause reproductive problems for agricultural and off-farm biodiversity, and contaminate waterways significantly. Secondary pest outbreaks, and nontarget mortality of natural enemies, often lead to increasing pesticide use, and continuing dependence on pesticides to combat ever-increasing pest loads (e.g., pesticide treadmill) (Van den Bosch, 1989).

Although ecological researchers have long maintained an interest in understanding relationships between diversity and trophic interactions, there is renewed interest in understanding natural processes of pest control in agroecosystems, due to growth of organic and ecological agriculture. Using biological control of pests necessitates understanding relationships between agricultural management, pest and natural enemy communities, and agricultural landscapes. Thus biodiversity impacts on pest control can be evaluated at a number of levels of biological organization (e.g., genetic diversity within a species, species diversity within a taxon, functional diversity within trophic level, and interactions within complex food webs) and at distinct spatial scales (e.g., farm plot, habitat, landscape, and region) (Letourneau and Bothwell, 2008) (**Figure 1**).

At the level of biological communities, agroecological research examines how two components of agrobiodiversity affect pest control. Planned biodiversity includes crop plants, other vegetation, and livestock chosen by the farmer to be in the agroecosystem. Associated biodiversity includes all other organisms that occur and survive in an agroecosystem depending on the management style chosen by the farmer and the surrounding landscape (Vandermeer and Perfecto, 1995). Thus researchers have focused on understanding how changes in agroecosystem management and agricultural landscapes directly affect natural enemies and pests, and specifically have examined the mechanisms underlying relationships between biodiversity of natural enemies, pests, and plants, and pest control services.

The study explores how biodiversity, both within farms and across landscapes, affects pest control. It focuses on how

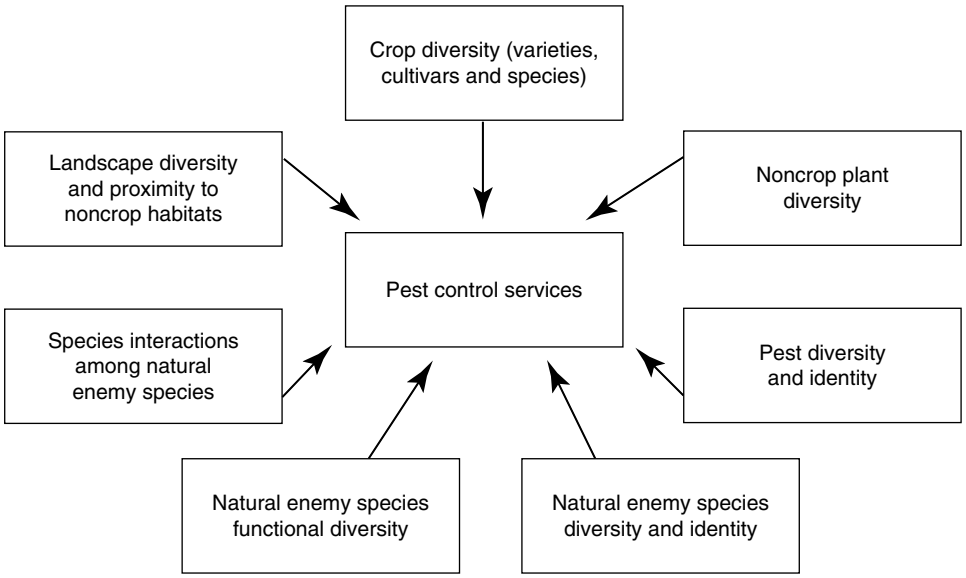


Figure 1 Major influences of biodiversity (at several levels of biological organization and distinct spatial scales) on pest control. Each of the levels of diversity influence pest control in several ways and interact resulting in complex multitrophic interactions.

changes in crop and noncrop plant diversity, natural enemy diversity, and changes in diversity at multiple trophic levels result in complex effects on pest control. It then shifts to patterns beyond the farm by examining how landscape diversity including maintaining noncrop vegetation nearby to farm fields, proximity to natural habitats, and habitat heterogeneity affect pest control. Finally, the study discusses conservation biological control and provides some conclusions. Although pests occur in a variety of ecosystems, from a wide array of taxonomic groups, it focuses on insect pests in terrestrial agricultural systems, and includes examples from other systems and pests where relevant to inform the discussion.

Agricultural Intensification and Impacts on Biodiversity

Creation and intensification of agriculture result in biodiversity loss with important implications for pest control. Destruction of habitat, conversion to agriculture, and simplification of landscape structure are the principal causes for biodiversity declines (Harrison and Bruna, 1999; Bianchi *et al.*, 2006). Furthermore, agricultural habitats are created at an ever-accelerating rate (Bawa *et al.*, 2004), and agricultural intensification results in dramatic biodiversity losses (Letourneau and Bothwell, 2008). Although natural ecosystems have largely intact food webs, with naturally occurring processes that maintain pest populations at certain levels, natural control is lost with agricultural intensification (Swift and Anderson, 1993). Agricultural intensification refers to two general processes: (1) changes in the vegetation diversity in an agroecosystem (including crop species and varieties and other vegetation components such as trees, trap crops, and weeds) and (2) changes in management practices and intensity of production including soil amending, chemical use, tillage, and irrigation, among others (Altieri, 1999). Agricultural intensification includes modifications at the local (e.g., shortening crop rotation cycles, decreasing crop diversity, increasing inputs, implementation of genetically modified crops, increased tillage, increasing field size, and increased mechanization) and landscape scales (e.g., converting natural habitat to fields, destroying edge habitats, simplifying landscapes, avoiding fallows, and fragmenting natural habitat) (Tscharntke *et al.*, 2005). Intensified, or modern, agriculture is highly simplified and includes low diversity of plant species and crop varieties (e.g., about 70 species are planted in 1440 million ha) (Altieri, 1999). The impacts of intensive agricultural production on biodiversity and ecosystem services are a major research focus.

Intensification affects biodiversity generally, but natural enemies, especially predator species, tend to be more strongly affected by habitat disturbance and loss leading to shifts in the ratios of prey to predator species, and subsequent effects for ecosystem processes (Bruno and Cardinale, 2008). A recent meta-analysis, covering 66 studies, found that both the species richness and abundance of predatory insects (as well as birds and spiders) was significantly higher in organic compared with conventional (or modern, intensive) farms, whereas the abundance of pest species and nonpredatory insects was higher in conventional farms (Bengtsson *et al.*, 2005). Specific groups of predators, such as carabids, are more abundant, diverse, and more evenly distributed in organic compared with conventional farms (Kromp, 1999). Frequently, organic agriculture benefits species richness and abundance of plants, predatory arthropods, and nonpredatory arthropods (Hole *et al.*, 2005; Letourneau and Bothwell, 2008). Other low-intensity agroecosystems, such as shaded tropical agroforests, harbor a high richness of predator species (e.g., ants and birds) (Perfecto *et al.*, 1996; Philpott *et al.*, 2008). Similarly, shade tree diversity within cacao agroforests positively impacts parasitoid wasp richness (Sperber *et al.*, 2004). The characteristics of organic farms likely to increase biodiversity of

natural enemies include agrochemical elimination, crop rotation, maintaining areas of seminatural vegetation, and preservation of mixed, or species-rich, farming (Hole *et al.*, 2005; Macfadyen *et al.*, 2009). In agroforests, increases in canopy diversity, density, and structural complexity correlate with increased predator diversity (Philpott *et al.*, 2008). Thus, less-intensive agricultural systems, such as organic and ecological farms, that include a high plant and natural enemy diversity, noncrop habitats, and landscape heterogeneity may support higher biodiversity, promoting pest control services. Despite the large number of studies that demonstrate higher abundance of natural enemies in organic farms, relatively few have examined differences in pest control, or cascading effects on crop yields (Letourneau and Bothwell, 2008). These and studies comparing pest control in diverse versus species poor farms, farms in simple versus complex landscapes, and the effects of releases of single versus multiple biological control agents are more common, and essential for understanding how pest control is affected by biodiversity.

Vegetation Diversity and Pest Control

The number of crop plants and varieties, weeds, planned weed strips, shade trees, and the temporal and spatial organization of plant species all affect pest populations and natural enemy assemblages. Andow (1991) reviewed the impacts of vegetation diversity on arthropod populations to test the resource concentration and natural enemies hypotheses proposed by Root (1973). The natural enemies hypothesis states that in polycultures (fields with multiple crop species), natural enemy abundance and diversity are higher due to more continuous food availability compared with monocultures (fields with a single crop species) (Root, 1973). The resource concentration hypothesis states that in monocultures, specialized herbivores have a more concentrated and unlimited food supply, thus supporting higher populations (Root, 1973). Thus, herbivore, or pest loads are affected by vegetation diversity in two major ways, with a number of possible mechanisms explaining each. Andow (1991) found that across 209 studies (and 287 herbivore species) reviewed, 51.9% of herbivore populations were denser in polycultures and 15.3% had higher population densities in polycultures. However, more herbivore species were found in polycultures than in monocultures. Across all studies, evidence indicated that both resource concentration and natural enemy increases suppress herbivores.

Crop Diversity

Increasing crop diversity can be accomplished across both space and time and can affect pest control via several mechanisms. Crop diversity can be altered by increasing the number of cultivars or varieties of a single species (e.g., increasing genetic diversity), increasing the species diversity of crops, adding crop rotations, and by increasing the architectural diversity of the crops. For several decades, scientists predicted that pest outbreaks should be more frequent in monocultures than in polycultures because crops associated with taxonomically diverse plantings will be less frequently attacked than those associated with simple mixtures (Elton, 1958; Root, 1973). Andow (1991) reviewed several hypotheses of how crop diversity may reduce herbivore loads and increase natural enemies. Herbivores, especially specialists, he wrote, should be less common in polycultures because they have problems locating host plants, because host-location cues may be interrupted by the mix of plant species, and plant quality is more variable. Parasitoids and predators should better deter pests in polycultures because they can switch prey when certain species become rare, natural enemy reproduction takes place more often in polycultures, and there are alternative hosts in polycultures. Natural enemy populations are further benefited in polycultures because therein exist prey refuges that allow prey populations to survive, stabilize the population fluctuations between prey and predators, and thereby make continual prey suppression possible. Several studies show empirical support for these hypotheses.

Increasing crop diversity has variable impacts on herbivore populations. Increasing genetic diversity can protect crops from pests, including crop diseases. Increasing the crop species can limit the dispersal of disease spores, and at the same time increase crop yields (Altieri, 2004). For example, in a study in China, increasing rice varieties planted across a large area decreased attack by rice blast (by 94%) and increased yields (by 89%) compared with single-variety monocultures (Zhu *et al.*, 2000). Although examples with diseases are somewhat more common, increased genetic diversity can also hinder insect pests. Increasing the diversity of willows in a field decreases the population density, oviposition rate, and plant damage caused by herbivorous beetles (Peacock and Herrick, 2000). Large-scale increases in corn acreage for biofuel production (19% increase between 2006 and 2007 in the USA) have negatively impacted pest control, yet increasing varieties of corn planted may promote pest control (Landis *et al.*, 2008).

Intercropping, or the cultivation of two or more crops such that they interact biologically, can reduce pest populations by decreasing resources available, specifically to specialist herbivores. Intercrops can be mixed or in neighboring rows or strips, where crops are grown in the same field, or relay crops where crops are grown one after the other (Vandermeer, 1992). Myriad examples empirically demonstrate that pests are less abundant in intercrops due to mechanisms associated with the resource concentration hypothesis (Vandermeer, 1992); however, intercropping may increase resources available to generalist herbivores (Schellhorn and Sork, 1997; Zehnder *et al.*, 2007). Timing of planting and fallow lands (a temporary lack of crop diversity) can be very important in regulating pests as is discontinuity in monocultures (Altieri, 1999). Providing some permanent vegetation helps maintain populations of natural enemies (Altieri, 1999). Temporal increases in crop diversity from using crop rotations can diminish pest problems. Rotations, or successional changes in crop species planted in a single field, can minimize weed and disease problems and can lower insect pest populations (McLaughlin and Mineau, 1995, and references therein).

Increasing crop diversity can also lead to increased architectural diversity, or from a pest's-eye view, habitat complexity. Habitat complexity can strongly affect both the abundance and diversity of natural enemies and their ability to capture

prey items (pests). Langellotto and Denno (2004) examined the impacts within plant and within habitat complexity on natural enemy communities. They found that in seven of nine natural enemy guilds, and for natural enemies overall, increasing habitat complexity increased abundance, and decreasing habitat complexity lowered abundance. They also reported possible mechanisms for these observations and did not find that increased prey abundance mediated observed increases in natural enemy abundance. Instead, refuge from other predators, additional provided resources (e.g., alternative prey, nectar, and pollen), and more effective prey capture appeared to play important roles. In contrast, habitat complexity may hinder pest control if parasitoid host-location cues are interrupted, or if parasitoid search efficiency is reduced (Andow, 1991).



Figure 2 A strawberry field in Watsonville, California, with sweet alyssum (*Brassicaceae*) featured prominently in the foreground. This plant species attracts beneficial insects (i.e., lacewings, syrphid flies) to the strawberry fields where they feed on pests and benefit the crop plant.

Other Vegetation in Crop Fields

Allowing noncrop plants to grow within crop fields can also increase natural enemy activity and decrease pest pressure. Diversity and abundance of weeds, and trap crops, in agricultural systems may also decrease pest densities because weeds may provide alternative resources for or otherwise harbor populations of natural enemies (Altieri, 1999). Especially in orchards and agroforests, natural enemy diversity may be higher, and pest populations lower, where weeds or other understory plants are maintained (Altieri, 1999). Adding strips of weedy plants into crops (including vineyards) can increase pest control services (Berndt *et al.*, 2002; Jacometti *et al.*, 2007). Weed strips, especially of perennial plants, can also provide habitat for natural enemies, and early sources of predators in to crop fields. The so-called "beetle banks" or strips of tussock grasses, for instance, are used throughout Europe to provide overwintering spots for predators of crop pests (Gurr *et al.*, 2003, and references therein). The location of such strips (often in the center of fields) can further enhance the degree to which pests are suppressed in the crops themselves.

Trap crops, additional plant species planted usually as decoys for pests, increase plant diversity in fields, can lower pest populations in crops, and often harbor natural enemy populations (**Figure 2**). Trap crops are used in a number of crops. In cotton fields in Australia, alfalfa strips are used to both attract cotton pests (e.g., *Creontiades dilutus*) and to provide habitat for natural enemies. When the alfalfa is mowed, natural enemies will move into cotton crops, thus this technique can be used to enhance biological control when needed (Gurr *et al.*, 2003). In citrus groves in China, an aster (*Ageratum conyzoides*) is commonly planted to encourage and stabilize populations of natural enemies of herbivorous mites (*Panonychus citri*) (Liang and Huang, 1994). However, increasing floral resources in agricultural systems may also benefit herbivores by increasing herbivore fitness or masking necessary host-location odors for the parasitoids (Lavandero *et al.*, 2006). In fact, some plant species used to increase parasitoid populations may simultaneously benefit herbivores, thus selection of the plant species used to enhance diversity and resources for natural enemies must be done with care (Lavandero *et al.*, 2006). Finally, alley cropping, or planting rows of woody plants in crop fields, can also reduce pest pressure. In a number of countries, legume trees are planted alongside maize crops which depending on the context can increase soil quality, reduce pest pressure, and increase natural enemy populations, or can increase seed predation on the crop, and increase root competition (Schroth *et al.*, 2000).

Natural Enemy Diversity and Pest Control

Ample research has examined the relationships between predator diversity and pest control (e.g., Pimentel, 1961; Root, 1973; Andow, 1991; Gliessman, 1989) and the complex interactions that arise between predator diversity and prey (Ives *et al.*, 2005, and references therein). Further, there is a long history of debate among experts in biological control, and conflicting evidence as to whether single or multiple species introductions of pest control agents are better for controlling pests (e.g., Cardinale *et al.*, 2003; Bianchi *et al.*, 2006). Increases in predator diversity do not always result in increased pest control over single-species treatments (Finke and Denno, 2004) and many biological control strategies prove successful with the introduction of just a few natural enemy species (Myers *et al.*, 1989; Denoth *et al.*, 2002). In fact, biological control strategies employed with multiple species can actually hinder biological control efforts (Denoth *et al.*, 2002). Yet, several empirical studies show that presence of multiple predators can enhance prey risk for important crop pests (Losey and Denno, 1998; Cardinale *et al.*, 2003; Schmidt *et al.*, 2003; Snyder and Ives, 2003). The variation in results, however, begs the question as to which mechanisms result in different outcomes in multipredator experiments.

There are several different mechanisms that could potentially contribute to both risk enhancement for prey (better pest control) and risk reduction (worse pest control). Both risk reduction and risk enhancement for pests can result in nonadditive effects – effects that do not add up to the sum of their

parts (Sih *et al.*, 1998; Denoth *et al.*, 2002; Bruno and Cardinale, 2008). Risk enhancement can result from the sampling effect, facilitation, species complementarity, and increased abundance of predators. Risk reduction can result from antagonistic effects such as aggression, cannibalism, and intraguild predation (IGP). Results from simulation models show that the sampling effect and complementarity among species tend to more strongly affect prey suppression than other mechanisms (Ives *et al.*, 2005). While different mechanisms may show prominence in certain studies or certain agroecosystems, it is likely that a variety of mechanisms operate simultaneously (Cardinale *et al.*, 2003). Each of these mechanisms and resulting effects is discussed in the following sections.

Mechanisms of Predator Diversity Effects

Sampling and Selection Effects

The sampling effect, also called the selection effect, selection probability, or lottery model, indicates that as richness increases, the chances of finding a species with strong or unique effects on lower trophic levels also increases (Huston, 1997; Ives *et al.*, 2005). Thus, in a community with higher biodiversity, it is more probable that one or two species responsible for large effects will be present. Sampling effects may occur where certain species have disproportionately large effects in a community, or where a single species has relatively greater abundance, prey capture ability, longevity, reproductive capacity, or competitive ability (Letourneau *et al.*, 2009). In biological control efforts, the sampling effect may be evident with releases of specialist species. In a review of natural enemy introductions to control agricultural pests, Denoth *et al.* (2002) found that in more than 50% of the successful introductions of multiple enemy species a single species was actually responsible for the successful control. They attribute this to the sampling effect whereby adding additional species increased the chance of having a useful one. Of course, sampling effects that result in introduction of a damaging or disruptive natural enemy could diminish pest control (Letourneau *et al.*, 2009).

Predator abundance usually increases as predator diversity increases, making it difficult to distinguish between the effects of abundance and richness, especially in field-based studies. Where correlations between natural enemy richness and pest control are encountered, increased pest control could result from simple increases in natural enemy abundance (Van Bael *et al.*, 2008). In the lab, experimental treatments can be set up to distinguish the two. With replacement designs, diversity is manipulated, but density or biomass of natural enemies is held constant. In additive designs, diversity is manipulated while initial density or biomass of natural enemies is held constant. Replacement designs are generally more useful for disentangling mechanisms of biodiversity and ecosystem function relationships, and specifically species–specific effects, whereas additive designs are more appropriate tests of complementarity or nonadditive effects (Ives *et al.*, 2005). In addition, additive designs more likely represent what is happening in agricultural fields with enhanced vegetation, predator refuges, and other complex structural components.

Facilitation

Facilitation occurs when two or more different species of natural enemies enhance the effect of another. There are many examples of predator facilitation in the literature that result in increased pest control. For example, predators that forage on vegetation (e.g., coccinellid beetles) will often scare pests (e.g., aphids) who then fall to the ground and are preyed on by ground-foraging carabid beetles (Losey and Denno, 1998). Thus the coccinellid assists the carabids to locate prey by chasing away aphids. Clearly, however, for such facilitation to occur and result in synergistic effects, in particular, the two predator species must forage in the same part of the season, at the same time of day, and not interfere with the prey capture rates of the other species (Losey and Denno, 1999).

Complementarity

Complementarity is based on the principle that as consumer species utilize different resources, thereby partitioning them, a greater extent of the available resources will be consumed (Loreau *et al.*, 2001). If a diverse suite of natural enemies partition resources, or are complementary, this may thus increase pest control. The effects of natural enemies that feed on different prey species, different life stages of a single prey species, or that forage or feed in different microhabitats of agroecosystems, or at the different times of day or seasons may combine in a complementary fashion (e.g., Bruno and Cardinale, 2008; Letourneau *et al.*, 2009). Complementarity often leads to increases in prey risk enhancement but may depend on the degree to which different natural enemies actually partition resources (Bruno and Cardinale, 2008). Although theory predicts that organisms partition resources, little empirical data support that complementarity increases pest suppression. This may be due to difficulty of assessing natural enemy diets and host preferences or because many studies are conducted in homogeneous agricultural fields without much option for partitioning (Ives *et al.*, 2005; Bruno and Cardinale, 2008).

Nonetheless, some studies show that species complementarity in diverse natural enemy assemblages increases pest control. Bogran *et al.* (2002) found that three species of parasitoids attacking a whitefly pest in cotton preferentially attacked larvae of different sizes and in spatially distinct areas of the cotton plants, thus leading to higher parasitism rates where all three species co-occurred. Finke and Snyder (2008) used an experimental test to separate the effects of species richness and resource partitioning and found evidence to support the latter. They released different mixes of specialist parasitoids (reared on one species of aphid hosts to which they show host "loyalty") and generalist parasitoids (reared on three aphid species) into large field cages with mixes of the aphids. They found that increasing the diversity of specialist parasitoids resulted in higher parasitism rates and lower aphid populations; however, increasing the richness of generalist species did not. Thus they empirically demonstrated the importance of resource partitioning as a mechanism driving biodiversity effects. Neumann and Shields (2008) found that releasing a combination of nematodes with complimentary foraging strategies (e.g., an ambush and a cruiser nematode) significantly reduced alfalfa insect damage compared with controls and one single–species treatment; however, not all

combinations of nematodes provided effective control. Finally, Williams-Guillén *et al.* (2008) compared the single- and multitaxon impacts of birds and bats on arthropod removal in coffee agroforests and found that each predator taxon provided higher pest control services during different seasons (birds during winter when migratory birds form a large fraction of the community). Further, likely because of the temporal separation in foraging times, birds and bats acting together had the greatest negative effect on arthropod loads.

Functional Diversity

Although species richness has been most often used as a metric of diversity, functional group richness has been invoked as a better predictor of ecosystem service as traits of organisms more strongly relate to functions than do taxonomic classifications (Tilman *et al.*, 1997; Diaz and Cabido, 2001). A functional group is a grouping of species based on similarity in behavioral, morphological, physiological, or resource use traits (Petchy and Gaston, 2006; Philpott *et al.*, 2009). Functional diversity of natural enemy characteristics, then, might be more important to consider than the taxonomic richness of predators itself, as this relates more strongly to ecosystem function (Hooper *et al.*, 2005).

Functional diversity, especially of important predators, may be more strongly affected by agricultural intensification than species richness. For example, Flynn *et al.* (2009) found that bird functional diversity declined with agricultural intensification more quickly than did species richness. Further, Schweiger *et al.* (2007) found that declines in specialist parasitoids were stronger than generalist parasitoids with habitat degradation. This may be especially important if species complementarity is an important mechanism maintaining positive effects of predator diversity on pest control. Philpott *et al.* (2009) examined patterns behind significant positive relationships between richness of insectivorous birds and arthropod removal in tropical agroforestry systems. They divided birds into functional groups based on characteristics related to predatory function (e.g., body size, diet, foraging strategy, and strata) and then correlated changes in functional richness with both species richness and pest control function. Species richness and functional richness were highly correlated across the nine study sites examined, and functional richness correlated significantly with arthropod removal. However, simple species richness remained a better predictor of ecosystem function than functional richness either because the traits included were not sufficient to explain all variation, or because presence of important predator species played a more important role.

Functional Redundancy and the Insurance Hypothesis

Functional redundancy may result in increased pest control under the insurance hypothesis. Functional redundancy, or a lack of complementarity among co-occurring species, indicates that natural enemies share traits such as similar foraging modes, diets, and strategies, and thus can be placed in the same functional group. Removing functionally redundant species from a community in theory has no effect on the ecosystem, whereas adding functionally redundant species can increase interspecific competition but does not affect pest control function (Straub *et al.*, 2008, and references therein). However, agricultural systems are constantly disturbed, and even functionally equivalent species may respond differently to environmental changes (Perfecto *et al.*, 2004). Thus, the importance of maintaining functionally redundant species can be supported under the insurance hypothesis.

The insurance hypothesis suggests generally that functional redundancy is important for maintaining ecosystem services should conditions change (Yachi and Loreau, 1999). In the context of pest control, this means that as crops are harvested, fields are tilled, or as weed strips are mowed, the resulting changes to natural enemy effectiveness will be buffered by the presence of functionally redundant species within the agroecosystem. Functional redundancy has been inferred from studies finding no effect of increased predator diversity on pest control. In some cases, spider species richness does not affect predation (Sokol-Hessner and Schmitz, 2002) and the diversity of aphid-feeding natural enemies does not increase predation (Chang, 1996). But, functional redundancy is difficult to show unequivocally because neutral effects of predator species may result from a combination of other effects (Straub *et al.*, 2008).

Intraguild Interference

Intraguild interference refers to negative interactions between natural enemy species including predation, cannibalism, predator avoidance behavior, and predator–predator competition (**Figure 3**) (Lang, 2003). Among phytophagous insects released as biological control agents there is a very high degree of competitive interactions. Denno *et al.* (1995) found that in 91% of releases (45 studies), biological control agents competed with each other. If such interactions are sufficiently antagonistic, this may result in herbivore release and potential pest outbreaks, but depends on the nature of the interactions



Figure 3 *Azteca instabilis* ants attacking a lepidopteran larva in a coffee agroecosystem. Ants are often effective predators; however, they can interfere with the activities of other natural enemies such as parasitoids, spiders, and coccinellid beetles resulting, at times, in risk reduction for prey when in diverse assemblages. Photo by D. Gonthier.

involved (Rosenheim *et al.*, 1999; Snyder and Wise, 2001; Perez-Lachaud *et al.*, 2004; Ives *et al.*, 2005). IGP occurs where trophic levels are not clearly defined, and predator species feed on herbivores as well as other predators (Polis *et al.*, 1989; Rosenheim *et al.*, 1995, Ives *et al.*, 2005).

Interspecific interactions result in both enhanced and reduced prey risk. Herbivore suppression may decline due to IGP (Rosenheim *et al.*, 1995; Snyder and Wise, 2001; Finke and Denno, 2005). Finke and Denno (2005) examined the effects of several spider species, a coccinellid, and a mirid bug on prey in salt marshes and found that increasing spider richness reduced prey suppression due to both IGP and nonlethal effects of some spider species on others. However, the effects of IGP by spiders changed with habitat complexity; in complex habitats where mirids find refuges to escape spider predation, effects on herbivores were greater than in simple habitats where the combination of predator species resulted in risk reduction for prey (Finke and Denno, 2002). Ant foraging interrupts spiders, resulting in lower predation rates (probably due to both IGP and direct interference) (Halaj *et al.*, 1997). Pell *et al.* (2008) report that the exotic coccinellid predator *Harmonia axyridis* interferes with other predators via cannibalism, interfering with oviposition of parasitoids, and by feeding on parasitized pest eggs (coincidental IGP). IGP by the ladybeetle does not diminish with increased prey and contributes to its success as an exotic invasive species in agricultural landscapes (Snyder *et al.*, 2004). In contrast, IGP sometimes results in risk enhancement for prey. Lang (2003) investigated the interactions between carabid beetles and spiders (lycosids and lyniphiids) in winter wheat. Carabid beetles negatively affected lycosid abundance, likely due to IGP, or by altering emigration rate of spiders out of carabid-free cages. In contrast, linyphiid abundance was not affected by the presence of carabids. Overall, predators did not negatively affect parasitism rates on aphids, and the overall impacts of predators, despite evidence of IGP, were synergistic resulting in higher predation rates where predator diversity was higher.

The strength of intraguild interactions in determining pest control outcomes is highly context dependent; effects vary with habitat complexity, prey density, type of IGP, and size and mobility of the intraguild predator and prey species (Müller and Brodeur, 2002; Pell *et al.*, 2008). Intraguild interactions most likely occur when natural enemies have similar hunting modes and foraging locations (Schmitz, 2007). Strength of intraguild effects also depends on plant architecture. In recently cut alfalfa fields, carabids are effective and quick predators on aphids, and also reduce predation rates leaving a longer term positive impact on pests (Snyder and Ives, 2001). Where plants are taller, carabids were no longer effective predators and still reduced parasitism thus having only negative impacts on biological control.

Finally, effects of IGP may be stronger for omnivorous rather than coincidental IGP. Coincidental IGP occurs where predators eat parasitized hosts, whereas omnivorous IGP occurs where predators consume other predators. Because coincidental IGP is coupled with direct predation on pests (e.g., intraguild predators eat prey and parasitoids simultaneously), omnivorous IGP most strongly interferes with pest suppression (Straub *et al.*, 2008). In biological control strategies, coincidental IGP actually can increase pest suppression (Rosenheim and Harmon, 2006). Thus, in sum, even where IGP occurs, a diverse assemblage of predators may still promote better pest control than a species-poor assemblage without interspecific effects (Letourneau *et al.*, 2009).

Evidence that Predator Diversity Enhances Pest Control

Several compelling examples demonstrate that natural enemy diversity can enhance pest control, many from controlled lab and cage studies in temperate agricultural studies that are discussed above (*see* Facilitation and Complementarity). In addition, Cardinale *et al.*, 2003 investigated the effects of a coccinellid, a parasitic wasp, and the damselbug on pea aphids on alfalfa in large field exclosures and found increased pest control (and increased yields) when all three species were together. However, establishment rate of natural enemy species (for insect pests but not for weed pests) can be significantly lowered when multiple species of agents are released (Denoth *et al.*, 2002). This could be due to competitive exclusion, to bias in the data set, or due to the fact that managers often continue releasing agents in sequence until one is successfully established (Denoth *et al.*, 2002). Lower establishment could also be due to intraspecific aggression, or to IGP among agents released, or those already residing within the target agricultural systems, as discussed above (*see* Intraguild Interference).

Relatively little manipulation of predator diversity has been conducted in tropical agricultural systems, but there are increasing numbers of compelling examples from coffee and cacao agroforests that both invertebrate and vertebrate natural enemy diversity relates to increased pest control. Tylianakis *et al.* (2008) examined a number of agricultural systems (pasture, rice coffee, abandoned coffee, forest) and found that parasitism of nectar and pollen-feeding wasps was higher where parasitoid diversity was higher. Perfecto *et al.* (2004) excluded birds from coffee plants in Mexico and found that suppression of an artificial outbreak was greater in farms with higher diversity and abundance of birds. They attributed the increase in pest removal to increased abundance of a particular insectivore species in the more complex coffee habitat, thereby providing field evidence for the sampling effect. Borkhataria *et al.* (2006) found in shade coffee farms that birds alone and the combination of birds and lizards significantly reduced arthropod abundance (but not the abundance of parasitoids or arthropod predators), and that coffee leafminers responded weakly to predator removal. Further, the combination of birds and lizards was an additive effect. They concluded that vertebrate predator diversity is important in controlling pests, and that in their system, vertebrate predators do not disrupt arthropod natural enemies through IGP.

Several authors have conducted meta-analyses to review how biodiversity affects herbivore or pest densities. The results widely vary, perhaps due to the studies included, the criteria for inclusion of studies, or the experimental method used. Cardinale *et al.* (2006) reviewed studies to examine the effects of consumer diversity on resource depletion. Of 111 experiments, eight covered the effects of terrestrial predators on prey. Overall, a diverse mix of predators (>3 species) removed prey better than a nondiverse mix, but prey suppression was not greater for a diverse mix than for the single best predator

species thus supporting the sampling effect. Schmitz (2007) examined studies specifically to evaluate support for different mechanisms driving multipredator effects. He found that in just under half of the studies examined (45.6%) that predator diversity resulted in risk enhancement or additive effects of predators, and in almost as many cases (40.3%), predator diversity resulted in risk reduction. He found widespread evidence for substitutable effects and for interspecific interference (including IGP) among natural enemy species; further, risk enhancement with multiple predator species was more likely in lab compared with field studies.

Finally, Letourneau *et al.* (2009) reviewed 62 studies, yielding 266 comparisons of diverse versus nondiverse mixtures of natural enemies. They found that in 69.5% of comparisons increased natural enemy diversity resulted in increases in pest suppression; in 30% of cases, increased diversity resulted in decreases in pest suppression. They also compared several study characteristics to determine how effects differed across system and location. In temperate areas, pest suppression as a result of natural enemy diversity was significant in agricultural systems, but not in natural systems, and pest suppression occurred in both temperate and tropical agricultural systems. Additionally, mean effect sizes (magnitude of pest suppression) were greater in cage compared with field studies. Thus a great deal of empirical evidence indicates that natural enemy diversity can enhance prey risk, but diversity effects are far from consistent.

One major limitation of studies conducted to date is the relatively low number of natural enemy species included in high-diversity treatments. In most meta-analyses, reporting the results of multipredator impacts on single prey species, the mean diversity of predators and parasitoids included is between three and four species (Borer *et al.*, 2005; Letourneau *et al.*, 2009). However, this is a far cry from the actual diversity of predators recorded, even in species-poor temperate agroecosystems. Natural enemy species richness for single herbivore species ranges from 13 to 86 species in several systems in the USA and northern Mexico (Letourneau *et al.*, 2009, and references therein). Further, when full prey–natural enemy communities are reported, numbers increase even more dramatically. More than 220 species of birds reportedly feed on agricultural pests in the USA (Letourneau *et al.*, 2009), and hundreds of species of predators feed on insect pests in tropical agricultural systems. Furthermore, the magnitude of diversity effects on pest suppression increases significantly with predator richness (Letourneau *et al.*, 2009), thus more studies are needed that manipulate a greater number of species in field or lab experiments, or new methods and models need be developed to examine the impacts of multiple predator species in highly species-rich communities.

Changes in Diversity at Multiple Trophic Levels

As agroecosystems are complex, so are the interactions therein, and interactions between diversity at multiple trophic levels may affect pest control. Differences in both plant diversity and predator diversity have distinct effects on pest populations. Yet, there may be complex interactions between diversity at different trophic levels that influences pests. Although many focus on diversity at the natural enemy trophic level, it is clear from other work presented that plant diversity has strong bottom-up effects in agricultural systems (Root, 1973; Andow, 1991). Moreover, multipredator effects are rarely placed within the context of other interactions and most studies aimed at testing multipredator effects examine only predator, prey, and plant trophic levels. However, adding vertical diversity within food webs (e.g., a fourth trophic level) may alter biodiversity effects at lower trophic levels (Duffy *et al.*, 2007). In the few studies conducted to date, for example, addition of herbivore species can both weaken and strengthen the relationship between biodiversity and ecosystem function (Duffy *et al.*, 2007).

A few studies have examined whether the effects multiple predator species have on herbivores and plants may be altered by the presence of a larger predator species, or a parasite of one or more predator species, and those that manipulate diversity at multiple levels are rare. Aquilino *et al.* (2005) manipulated predator diversity (from one to three species) and plant diversity (from one to two species). They used three plant species (alfalfa, clover, and fava beans) and three predator species (two species of coccinellid and a damselbug). They found that increasing predator diversity resulted in an increase in predation on pea aphids (on both monocultures and polycultures), but that increasing plant diversity decreased predation rate by the same magnitude (in both single- and multipredator treatments). Negative effects of IGP on a single pest species may disappear in studies including alternate prey species (Schmitz, 2007) indicating that even simple changes to prey richness may affect the effects of predator diversity. Finally, Macfayden *et al.* (2009) used a food web approach to examine whether diversity changes at all trophic levels influenced whether organic farms are more resistant to establishment of novel pests. In comparing food web networks in 10 pairs of organic and conventional farms in England, they found that organic farms had higher richness of plants, herbivores, and parasitoid species. However, in farms with more species, food web connectivity was lower, and parasitism rates and numbers of parasitoid species attacking herbivores did not differ in the diverse organic farms and the conventional farms. Thus, in systems where complexity of an entire food web was examined, increased predator diversity did not necessarily result in increased pest control. Despite the difficulty in manipulating the vast array of species in multiple trophic levels, more complete investigation will be necessary for fully understanding how differences in diversity at multiple trophic levels affect pest control (Ives *et al.*, 2005).

Agricultural Landscapes and Biodiversity

Agricultural landscapes represent a wide variety of habitats including crops, noncrop vegetation patches, woodlands, wetlands, grasslands, and forests. In such landscapes, the presence of noncrop vegetation, the distance to natural areas, and the complexity of the landscape can all affect pest control.

Noncrop Habitats

Hedgerows, live fences, and other linear habitats within agricultural systems provide habitat for birds, bats, dung beetles,

and butterflies (Harvey *et al.*, 2005) and specifically for several groups of invertebrate natural enemies (Bianchi *et al.*, 2006). As agricultural habitats are constantly disturbed, hedgerows and other crop margins provide stable resource bases for natural enemies (Bianchi *et al.*, 2006). The characteristics of noncrop habitats that benefit natural enemies include providing alternate prey, nectar and pollen, nesting sites, and host plants necessary for reproduction and life-cycle completion (Landis *et al.*, 2000; Bianchi *et al.*, 2006), and are similar to benefits provided by vegetation diversity within crop fields. Hedgerows and field margins increase the movement of predators across agricultural landscapes similarly to how a high-quality matrix may increase movement of organisms between forest fragments (Vandermeer and Carvajal, 2001; Tscharntke *et al.*, 2005). In some cases, increases in predator diversity in hedgerows can increase pest control. For example, linear vegetation strips in vineyards in California facilitate movement of natural enemies in the grapes, and thereby increase pest control (Benton *et al.*, 2003). Field margins and hedgerows can slow movement of fungal pathogens and can serve as barriers to the movement of pests, thereby improving pest control (Altieri, 1999). However, if the field margins provide alternative habitat for beneficial insects and other predators such that they forage more in the margins than in the crop, or prefer the crop-margin habitats more than the natural habitats, then this may harm pest control (Benton *et al.*, 2003; Bianchi *et al.*, 2006).

Proximity to Natural Areas

Distance to natural areas such as woodlands and grasslands can also affect pest control because natural enemies may find refuge in nearby natural habitats. Species of higher trophic levels, such as natural enemies, tend to be more strongly negatively affected by isolation from natural habitats (Klein *et al.*, 2006, and references therein). These declines in richness are likely due to more unstable populations of predator and parasitoid species, increased energy requirements, and a lack of food and nectar sources when far from natural habitats (Klein *et al.*, 2006). In agroforest landscapes in Indonesia, both diversity of natural enemies and parasitism rates of trap-nesting hymenopteran brood declined with distance from forest, and for natural enemies, this landscape factor was a more important predictor than light intensity or the number of plant species in crop areas (Klein *et al.*, 2006). Furthermore, in coffee agroforests in Mexico, ant richness declines markedly with distance from forest fragments, especially in less diverse agroecosystems leading to less diversity of predatory species further from forest fragments (Perfecto and Vandermeer, 2002; Armbrecht and Perfecto, 2003). Predation rates may also decrease with increased distance from noncrop habitats at field margins. Parasitism rate declines, likely because parasitoids (and predators) are more susceptible to habitat fragmentation than herbivores (Kruess and Tscharntke, 1994; Bianchi *et al.*, 2006). Thus landscapes with a greater diversity of habitats, and especially with smaller habitat patches may be preferable for increasing natural enemy function.

Habitat Heterogeneity

At the landscape level, habitat heterogeneity can strongly benefit pest control services. Including a high degree of habitat heterogeneity in agricultural landscapes (including many different types of crop fields, natural habitat areas, hedgerows, fencerows, wetlands, etc.) can increase the diversity of natural enemies in crop fields, and also can provide stability of resources for maintaining natural enemy populations (**Figure 4**) (Altieri, 1999). Habitat heterogeneity increases



(a) (b)

Figure 4 Two contrasting agricultural landscapes differing in habitat heterogeneity. The high heterogeneity landscape from Sumatra, Indonesia (a) contains a high diversity of both crops (banana, mango, coffee, and rice) and includes forest trees and weed patches. The low heterogeneity landscape (b) near Toledo, Ohio, shows large corn and soybean monocultures and a very small forest fragment. Most evidence indicates pest control is more effective in high diversity landscapes.

with increased number of habitats, and generally with smaller patch size; spatial arrangement of patches may also be important. In landscapes with small patches, natural enemies may be better able to reach all areas of a crop field, and colonize early in the growing season. Heterogeneous landscapes may support higher abundance and diversity of natural enemies simply due to the different preferences of different species for different habitats (Bianchi *et al.*, 2006). Boatman (1994) demonstrated that carabid beetles move from field margins approximately 15–30 m into field margins, and thus the locations of field margins could be maximized to increase the abundance of polyphagous predators. Furthermore, some natural enemies can move only short distances into crop fields from margins but the dispersal distances vary with species examined (Nicholls *et al.*, 2001; Bianchi *et al.*, 2006). Thies and Tscharntke (1999) studied the effects of habitat complexity of noncrop habitats and their role on oilseed rape crops. They found that presence of structurally complex noncrop habitats related to lower parasitism of the pest than presence of simple noncrop habitats nearby.

Even in intensive, extensive agricultural systems, landscape diversity can be important in promoting biological control. Marino and Landis (1996) explored parasitoid diversity and attacks in cornfields in Michigan embedded in complex (small plots with abundant hedgerows) and simple (large plots with rare hedgerows) landscapes near and far from hedgerows. They found that parasitoid richness was similar in both habitats, and parasitism rates were five times higher in complex than in simple landscapes, but that distance from hedgerows did not affect parasitism in either landscape. In the Midwest US, relative removal rates of aphid pests in soybean fields were increased by landscape diversity (e.g., number and evenness of different habitats) (Gardiner *et al.*, 2009). In addition, abundance of coccinellid beetles (the main predator encountered in soy fields) was positively affected by the amount of natural habitat in the surrounding area. Conversely, in these same landscapes, increases in area planted with corn (largely for biofuel production) resulted in strong decreases in biological control of soybean aphids due to declines in habitat heterogeneity and specifically losses of area formerly in fallow or conservation lands (Landis *et al.*, 2008).

A few recent reviews have specifically targeted effects of habitat heterogeneity at the landscape scale on pest control. Bianchi *et al.* (2006) conducted a literature review to examine the impacts of landscape complexity on natural enemy activity in relation to pest pressure. They defined complex landscapes as those with a high proportion of noncrop habitats (e.g., forests, hedgerows, tree lines, grasslands, wetlands, and fallows) and with small patches (large perimeter to area ratios). In 74% of the studies examined, pest control was enhanced in complex landscapes. Further, pest pressure (defined as population densities, crop injury, and survival and population growth rate of aphids) was reduced in complex landscapes in 45% of observations. They highlighted that landscapes with herbaceous vegetation (80% of studies with enhanced natural enemy activity), woody vegetation (71%), and landscape patchiness (70%) were those that most related to increased natural enemy activity.

Studies since this time have documented that parasitism rate increases with forest area (measured at multiple scales), proximity to forest, and proximity to road edges (Bianchi *et al.*, 2008). But just as with relationships between natural enemy diversity and pest control, the mechanisms underlying relationships between landscape diversity and pest control services need to be more specifically examined. Benton *et al.* (2003) reported that a mosaic of farm fields connected by noncrop habitat benefits birds, predatory ground beetles, and spiders. Thus in sum, several aspects of agricultural landscapes, including noncrop habitat, and spatial structure of the landscape affect pest control services provided by natural enemies, and movement of natural enemies across agricultural landscapes (Tscharntke *et al.*, 2005).

Conservation Biological Control

Conservation biological control is a process by which managers manipulate plants and other aspects of agricultural landscapes in order to increase abundance and diversity of natural enemies (Barbosa, 1998; Fiedler *et al.*, 2008). Normally, managers increase enemy populations by planting nectar sources, floral resources, seed production, and plants that support alternative prey, or shelter (Landis *et al.*, 2000). Out of 34 studies that have evaluated the impact of habitat management (e.g., purposeful plant additions) to increase natural enemies, most have focused on just four plant species, all annual, and most not native to the study area (Fiedler *et al.*, 2008). One study did evaluate the effectiveness of several native perennial plants in attracting natural enemies compared with plants commonly examined in habitat management trials (Fiedler *et al.*, 2008, and references therein). Many of the native plants screened attracted high numbers of natural enemies, more so than the commonly used nonnative plants, and in addition, these other plants may provide additional ecosystem services (e.g., increasing pollinator abundance and rural beauty) (Fiedler *et al.*, 2008).

Because impacts of natural enemy richness on prey suppression are variable and context dependent, making clear recommendations to farmers is difficult. Understanding the traits that increase pest suppression is key (Straub *et al.*, 2008). If, for example, the selection effect is operative in a particular agroecosystem, one could recommend the important predator, rather than planning to conserve natural enemy diversity more broadly (Straub *et al.*, 2008). Likewise, promoting the conservation of functionally redundant species within a community should not negatively impact function, and to the contrary may improve pest control services if conditions change (Yachi and Loreau, 1999; Straub *et al.*, 2008). Further, there are often nontarget impacts of introduced biological control agents, and if many species need to be introduced, there is even greater chance of impacts on nontarget species (Denoth *et al.*, 2002). Thus, considering pest control using naturally occurring predators and parasitoids is important.

Conclusions

Several levels of biological and habitat diversity affect pest control in complex manners. Vegetation diversity including crop genetic diversity, crop species richness, and noncrop plants

in farm fields can increase the number and function of natural enemies in crop fields. Predator diversity is a strongly context-dependent predictor of pest control, sometimes resulting in increased risk for pests where functional richness or complementarity is high, or where single species of effective predators are found. Yet IGP and other interspecific interactions between natural enemy species may result in risk reduction. At the habitat level, vertical structure of crop plants, weeds, and shade trees may promote population stability and diversity of natural enemies. Landscape complexity, including incorporating hedgerows and other types of noncrop habitat, as well as maintaining highly complex landscapes with a high amount of natural habitat seem to best promote pest control services.

Pest control, however, as an important ecosystem service in human-managed systems, should be examined in a larger context. First, pest control has a long history, and has in agroecological farms moved away from a pest elimination mentality toward understanding complex ecological interactions. For example, in traditional Guatemalan agroecosystems, farmers have intricate knowledge of their agricultural systems, and insects that feed on crops; however, they do not consider them pests (Morales and Perfecto, 2001). Instead, they consider the multitude of techniques used (e.g., intercropping, natural composting, and allowing for survival of beneficial insects) as a way to keep the insects from becoming pests. Many traditional agricultural systems, which incorporate traditional knowledge, and high levels of planned and associated biodiversity have complex ecological webs that result in high levels of natural pest suppression (Gliessman, 1989); Vandermeer *et al.*, 2010). Thus considering not only strategies for eliminating pests, but also employing strategies to maintain complex food webs may be warranted.

Whatever pest control strategies are promoted should also be considered in the context of farmers whose livelihoods strongly depend on crop production. For example, Steffan-Dewenter *et al.* (2007) examined a number of ecosystem services provided in shaded and unshaded cacao agroforests in Indonesia. They discovered that conversion of forest to cacao systems negatively affected plant biomass and a range of ecosystem services, but had little effect on species richness of associated biodiversity, overall. Reducing shade cover in the cacao farms, however, resulted in nonlinear changes in biodiversity, ecosystem services, cacao production, and income. Reducing shade cover in cacao farms from 80% to 40% resulted in higher levels of ecosystem services provided, and marginal increases in income. Eliminating shade cover altogether (to 0%) drastically lowered ecosystem function, but boosted incomes by 40%, and thus provided obvious benefits to farmers, at the cost of other ecosystem benefits. Thus any strategy aimed at increasing pest control services, or conservation biological control need consider not only the ecological principles discussed here, but also the farmers' livelihood.

Appendix

List of Courses

1. Agroecology
2. Biodiversity and Ecosystem Function
3. Ecosystem Services
4. Integrated Pest Management
5. Sustainable Agriculture

See also: Agrobiodiversity. Biodiversity and Ecosystem Services. Ecology of Agriculture

References

Altieri MA (1999) The ecological role of biodiversity in agroecosystems. *Agriculture, Ecosystems & Environment* 74: 19–31.

Altieri MA (2004) Linking ecologists and traditional farmers in the search for sustainable agriculture. *Frontiers in Ecology and the Environment* 2: 35–42.

Andow D (1991) Vegetational diversity and arthropod population response. *Annual Review of Entomology* 36: 561–586.

Aquilino KM, Cardinale BJ, and Ives AR (2005) Reciprocal effects of host plant and natural enemy diversity on herbivore suppression: An empirical study of a model tritrophic system. *Oikos* 108: 275–282.

Armbrecht I and Perfecto I (2003) Litter-twig dwelling ant species richness and predation potential within a forest fragment and neighboring coffee plantations of contrasting habitat quality in Mexico. *Agriculture, Ecosystems & Environment* 97: 107–115.

Barbosa P (1998) *Conservation Biological Control*. San Diego, CA: Academic Press.

Bawa KS, Kress WJ, Nadkami NM, *et al.* (2004) Tropical ecosystems into the 21st century. *Science* 306: 227–228.

Benbrook C (2001) Do GM crops mean less pesticide use? *Pesticide Outlook* 12: 204–207.

Bengtsson J, Ahnström J, and Weibull AC (2005) The effects of organic agriculture on biodiversity and abundance: A meta-analysis. *Journal of Applied Ecology* 42: 261–269.

Benton TG, Vickery JA, and Wilson JD (2003) Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology & Evolution* 18: 182–188.

Berndt LA, Wratten SD, and Hassan P (2002) Effects of buckwheat flowers on leafroller (*Lepidoptera: Tortricidae*) parasitoids in a New Zealand vineyard. *Agricultural and Forest Entomology* 4: 39–45.

Bianchi F, Booij C, and Tscharntke T (2006) Sustainable pest regulation in agricultural landscapes: A review on landscape composition, biodiversity and natural pest control. *Proceedings of the Royal Society B: Biological Sciences* 273: 1715.

Bianchi FJJA, Goedhart P, and Baveco J (2008) Enhanced pest control in cabbage crops near forest in The Netherlands. *Landscape Ecology* 23: 595–602.

Boatman N (1994) *Field Margins: Integrating Agriculture and Conservation*. Surrey, UK: British Crop Protection Council.

Bogran CE, Heinz KM, and Ciomperlik MA (2002) Inter-specific competition among insect parasitoids: Field experiments with whitefiles as hosts in cotton. *Ecology* 83: 653–668.

Borer ET, Seabloom EW, Shurin JB, *et al.* (2005) What determines the strength of a trophic cascade? *Ecology* 86: 528–537.

Borkhataria RR, Collazo JA, and Groom MJ (2006) Additive effects of vertebrate predators on insects in a Puerto Rican coffee plantation. *Ecological Applications* 16: 696–703.

Bruno JF and Cardinale BJ (2008) Cascading effects of predator richness. *Frontiers in Ecology and the Environment* 6: 539–546.

Cardinale BJ, Harvey CT, Gross K, and Ives AR (2003) Biodiversity and biocontrol: Emergent impacts of a multi-enemy assemblage on pest suppression and crop yield in an agroecosystem. *Ecology Letters* 6: 857–865.

Cardinale BJ, Srivastava DS, Duffy JE, *et al.* (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443: 989–992.

Chang GC (1996) Comparison of single versus multiple species of generalist predators for biological control. *Environmental Entomology* 25: 207–212.

Daily G (1997) *Nature's Services: Societal Dependence on Natural Ecosystems*. Washington: Island Press.

Denno RF, McClure MS, and Ott JR (1995) Interspecific interactions in phytophagous insects: Competition reexamined and resurrected. *Annual Review of Entomology* 40: 297–331.

Denoth M, Frid L, and Myers JH (2002) Multiple agents in biological control: Improving the odds? *Biological Control* 24: 20–30.

Diaz S and Cabido M (2001) Vive la difference: Plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution* 16: 646–655.

Duffy JE, Cardinale BJ, France KE, *et al.* (2007) The functional role of biodiversity in ecosystems: Incorporating trophic complexity. *Ecology Letters* 10: 522–538.

Elton CS (1958) *The Ecology of Invasions by Animals and Plants*. London, UK: Methuen.

Fiedler AK, Landis DA, and Wratten SD (2008) Maximizing ecosystem services from conservation biological control: The role of habitat management. *Biological Control* 45: 254–271.

Finke DL and Denno RF (2002) Intraguild predation diminished in complex-structured vegetation: Implications for prey suppression. *Ecology* 83: 643–652.

Finke DL and Denno RF (2004) Predator diversity dampens trophic cascades. *Nature* 429: 407–410.

Finke DL and Denno RF (2005) Predator diversity and the functioning of ecosystems: The role of intraguild predation in dampening trophic cascades. *Ecology Letters* 8: 1299–1306.

Finke DL and Snyder WE (2008) Niche partitioning increases resource exploitation by diverse communities. *Science* 321: 1488.

Flynn D, Gogol-Prokurat M, Nogeire T, *et al.* (2009) Loss of functional diversity under land use intensification across multiple taxa. *Ecology Letters* 12: 22–33.

Gardiner M, Landis D, Gratton C, *et al.* (2009) Landscape diversity enhances biological control of an introduced crop pest in the north-central USA. *Ecological Applications* 19: 143–154.

Gliessman SR (1989) *Ecological Basis for Sustainable Agriculture*. New York, NY: Springer-Verlag.

Gurr GM, Wratten SD, and Luna JM (2003) Multi-function agricultural biodiversity: Pest management and other benefits. *Basic and Applied Ecology* 4: 107–116.

Halaj J, Ross DW, and Moldenke AR (1997) Negative effects of ant foraging on spiders in Douglas-fir canopies. *Oecologia* 109: 313–322.

Harrison S and Bruna E (1999) Habitat fragmentation and large-scale conservation: What do we know for sure? *Ecography* 22: 225–232.

Harvey CA, Villanueva C, Villacís J, *et al.* (2005) Contribution of live fences to the ecological integrity of agricultural landscapes. *Agriculture, Ecosystems & Environment* 111: 200–230.

Hill D (1987) *Agricultural Insect Pests of Temperate Regions and their Control*. New York: Cambridge University Press.

Hole DG, Perkins AJ, Wilson JD, *et al.* (2005) Does organic farming benefit biodiversity? *Biological Conservation* 122: 113–130.

Hooper DU, Chapin FS, Ewel JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75: 3–35.

Huang HT and Yang P (1987) The ancient cultured citrus ant. *BioScience* 37: 665–671.

Huston MA (1997) Hidden treatments in ecological experiments: Re-evaluating the ecosystem function of biodiversity. *Oecologia* 110: 449–460.

Ives AR, Cardinale BJ, and Snyder WE (2005) A synthesis of subdisciplines: Predator–prey interactions, and biodiversity and ecosystem functioning. *Ecology Letters* 8: 102–116.

Jacometti MA, Wratten SD, and Walter M (2007) Management of understorey to reduce the primary inoculum of *Botrytis cinerea*: Enhancing ecosystem services in vineyards. *Biological Control* 40: 57–64.

Klein A, Steffan-Dewenter I, and Tscharntke T (2006) Rain forest promotes trophic interactions and diversity of trap-nesting Hymenoptera in adjacent agroforestry. *Journal of Animal Ecology* 75: 315–323.

Kremen C, Williams NM, Bugg RL, Fay JP, and Thorp RW (2004) The area requirement of an ecosystem service: Crop pollination by native bee communities in California. *Ecology Letters* 7: 1109–1119.

Kromp B (1999) Carabid beetles in sustainable agriculture: A review on pest control efficacy, cultivation impacts and enhancement. *Agriculture, Ecosystems & Environment* 74: 187–228.

Kruess A and Tscharntke T (1994) Habitat fragmentation, species loss, and biological control. *Science* 264.

Landis DA, Gardiner MM, Van Der Werf W, and Swinton SM (2008) Increasing corn for biofuel production reduces biocontrol services in agricultural landscapes. *Proceedings of the National Academy of Sciences* 105: 20552.

Landis DA, Wratten SD, and Gurr GM (2000) Natural enemies of arthropod pests in agriculture. *Annual Review of Entomology* 45: 175–201.

Lang A (2003) Intraguild interference and biocontrol effects of generalist predators in a winter wheat field. *Oecologia* 134: 144–153.

Langellotto GA and Denno RF (2004) Responses of invertebrate natural enemies to complex-structured habitats: A meta-analytical synthesis. *Oecologia* 139: 1–10.

Lavandero B, Wratten SD, Didham RK, and Gurr G (2006) Increasing floral diversity for selective enhancement of biological control agents: A double-edged sward? *Basic and Applied Ecology* 7: 236–243.

Letourneau DK and Bothwell SG (2008) Comparison of organic and conventional farms: Challenging ecologists to make biodiversity functional. *Frontiers in Ecology and the Environment* 6: 430–438.

Letourneau DK, Jedlicka JA, Bothwell SG, and Moreno CR (2009) Effects of natural enemy biodiversity on the suppression of arthropod herbivores in terrestrial ecosystems. *Annual Review of Ecology, Evolution, and Systematics* 40: 573–592.

Liang W and Huang M (1994) Influence of citrus orchard ground cover plants on arthropod communities in China: A review. *Agriculture, Ecosystems & Environment* 50: 29–37.

Loreau M, Naeem S, Inchausti P, *et al.* (2001) Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science* 294: 804–808.

Losey JE and Denno RF (1998) Positive predator-predator interactions: Enhanced predation rates and synergistic suppression of aphid populations. *Ecology* 79: 2143–2152.

Losey JE and Denno RF (1999) Factors facilitating synergistic predation: The central role of synchrony. *Ecological Applications* 9: 378–386.

Macfadyen S, Gibson R, Polaszek A, *et al.* (2009) Do differences in food web structure between organic and conventional farms affect the ecosystem service of pest control? *Ecology Letters* 12: 229–238.

Marino PC and Landis DA (1996) Effect of landscape structure on parasitoid diversity and parasitism in agroecosystems. *Ecological Applications* 6: 276–284.

McLaughlin A and Mineau P (1995) The impact of agricultural practices on biodiversity. *Agriculture, Ecosystems & Environment* 55: 201–212.

Morales H and Perfecto I (2001) Traditional knowledge and pest management in the Guatemalan highlands. *Agriculture and Human Values* 17: 49–63.

Müller CB and Brodeur J (2002) Intraguild predation in biological control and conservation biology. *Biological Control* 25: 216–223.

Myers JH, Higgins C, and Kovacs E (1989) How many insect species are necessary for the biological control of insects? *Environmental Entomology* 18: 541–547.

Neumann G and Shields EJ (2008) Multiple-species natural enemy approach for biological control of alfalfa snout beetle (Coleoptera: Curculionidae) using entomopathogenic nematodes. *Journal of Economic Entomology* 101: 1533–1539.

Nicholls CI, Parrella M, and Altieri MA (2001) The effects of a vegetational corridor on the abundance and dispersal of insect biodiversity within a northern California organic vineyard. *Landscape Ecology* 16: 133–146.

Peacock L and Herrick S (2000) Responses of the willow beetle *Phratora vulgatissima* to genetically and spatially diverse *Salix* spp. plantations. *Journal of Applied Ecology* 37: 821–831.

Pell JK, Baverstock J, Roy HE, Ware RL, and Majerus MEN (2008) Intraguild predation involving *Harmonia axyridis*: A review of current knowledge and future perspectives. *BioControl* 53: 147–168.

Perez-Lachaud G, Batchelor TP, and Hardy ICW (2004) Wasp eat wasp: Facultative hyperparasitism and intraguild predation by bethylid wasps. *Biological Control* 30: 149–155.

Perfecto I, Rice RA, Greenberg R, and VanderVoort ME (1996) Shade coffee: A disappearing refuge for biodiversity. *Bioscience* 46: 598–608.

Perfecto I and Vandermeer J (2002) The quality of agroecological matrix in a tropical montane landscape: Ants in coffee plantations in southern Mexico. *Conservation Biology* 16: 174–182.

Perfecto I, Vandermeer JH, Bautista GL, *et al.* (2004) Greater predation in shaded coffee farms: The role of resident neotropical birds. *Ecology* 85: 2677–2681.

Petchy OL and Gaston KJ (2006) Functional diversity: Back to basics and looking forward. *Ecology Letters* 9: 741–758.

Philpott SM, Arendt WJ, Armbrecht I, *et al.* (2008) Biodiversity loss in Latin American coffee landscapes: Review of the evidence on ants, birds, and trees. *Conservation Biology* 22: 1093–1105.

Philpott SM, Soong O, Lowenstein JH, *et al.* (2009) Functional richness and ecosystem services: Bird predation on arthropods in tropical agroecosystems. *Ecological Applications* 19: 1858–1867.

Pimentel D (1961) Species diversity and insect population outbreaks. *Annals of the Entomological Society of America* 54: 76–86.

Polis GA, Hurd SD, Jackson CT, and Sanchez-Pinero F (1989) Multi-factor population limitation: Variable spatial and temporal control of spiders on Gulf of California islands. *Ecology.*

Relyea RA (2005) The impact of insecticides and herbicides on the biodiversity and productivity of aquatic communities. *Ecological Applications* 15: 618–627.

Root RB (1973) Organization of a plant-arthropod association in simple and diverse habitats: The fauna of collards (*Brassica oleracea*). *Ecological Monographs* 43: 95–124.

Rosenheim JA and Harmon JP (2006) The influence of intraguild predation on the suppression of a shared prey population: An empirical assessment. In: Brodeur J and Boivin G (eds.) *Trophic and Guild Interactions in Biological Control*, pp. 1–20. Dordrecht, The Netherlands: Springer.

Rosenheim JA, Kaya HK, Ehler LE, Marois JJ, and Jaffee B (1995) Intraguild predation among biological-control agents: Theory and evidence. *Biological Control* 5: 303–335.

Rosenheim JA, Limburg DD, and Colfer RG (1999) Impact of generalist predators on a biological control agent, *Chrysoperla carnea:* Direct observations. *Ecological Applications* 9: 409–417.

Schellhorn NA and Sork VL (1997) The impact of weed diversity on insect population dynamics and crop yields in collards, *Brassica oleracea* (Brassicaceae). *Oecologia* 111: 233–240.

Schmidt MH, Lauer A, Purtauf T, *et al.* (2003) Relative importance of predators and parasitoids for cereal aphid control. *Proceedings of the Royal Society B: Biological Sciences* 270: 1905–1909.

Schmitz OJ (2007) Predator diversity and trophic interactions. *Ecology* 88: 2415–2426.

Schroth G, Krauss U, Gasparotto L, Duarte Aguilar JA, and Vohland K (2000) Pests and diseases in agroforestry systems of the humid tropics. *Agroforestry Systems* 50: 199–241.

Schweiger O, Musche M, Baily D, *et al.* (2007) Functional richness of local hoverfly communities (Diptera, Syrphidae) in response to land use across temperate Europe. *Oikos* 116: 461–472.

Sih A, Englund G, and Wooster D (1998) Emergent impacts of multiple predators on prey. *Trends in Ecology & Evolution* 13: 350–355.

Snyder WE, Clevenger GM, and Eigenbrode SD (2004) Intraguild predation and successful invasion by introduced ladybird beetles. *Oecologia* 140: 559–565.

Snyder WE and Ives AR (2001) Generalist predators disrupt biological control by a specialist parasitoid. *Ecology* 82: 705–716.

Snyder WE and Ives AR (2003) Interactions between specialist and generalist natural enemies: Parasitoids, predators, and pea aphid biocontrol. *Ecology* 84: 91–107.

Snyder WE and Wise DH (2001) Contrasting trophic cascades generated by a community of generalist predators. *Ecology* 82: 1571–1583.

Sokol-Hessner L and Schmitz OJ (2002) Aggregate effects of multiple predator species on a shared prey. *Ecology* 83.

Sperber CF, Nakayama K, Valverde MJ, and Neves FS (2004) Tree species richness and density affect parasitoid diversity in cacao agroforestry. *Basic and Applied Ecology* 5: 241–251.

Steffan-Dewenter I, Kessler M, Barkmann J, *et al.* (2007) Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification. *Proceedings of the National Academy of Sciences* 104: 4973.

Straub CS, Finke DL, and Snyder WE (2008) Are the conservation of natural enemy biodiversity and biological control compatible goals? *Biological Control* 45: 225–237.

Swift and Anderson (1993) Biodiversity and ecosystem function in agroecosystems. In: Schultze E and Mooney HA (eds.) *Biodiversity and Ecosystem Function*, pp. 57–83. New York: Springer.

Thies C and Tscharntke T (1999) Landscape structure and biological control in agroecosystems. *Science* 285: 893.

Tilman D, Knops J, Wedin D, *et al.* (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277: 1300–1302.

Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity–ecosystem service management. *Ecology Letters* 8: 857–874.

Tylianakis JM, Rand TA, Kahmen A, *et al.* (2008) Resource heterogeneity moderates the biodiversity-function relationship in real world ecosystems. *PLoS Biology* 6: 947–956.

Van Bael S, Philpott SM, Greenberg R, *et al.* (2008) Birds as predators in tropical agroforestry systems. *Ecology* 89: 928–934.

Van den Bosch R (1989) *The Pesticide Conspiracy.* Berkeley: University of California Press.

van Lenteren JC, Bale J, Bigler E, Hokkanen HMT, and Loomans AM (2006) Assessing risks of releasing exotic biological control agents of arthropod pests. *Annual Review of Entomology* 51: 609–634.

Vandermeer J and Carvajal R (2001) Metapopulation dynamics and the quality of the matrix. *American Naturalist* 158: 211–220.

Vandermeer J and Perfecto I (1995) *Breakfast of Biodiversity: The Truth About Rain Forest Destruction.* Oakland, CA: Institute for Food and Development Policy.

Vandermeer J, Perfecto I, and Philpott S (2010) Ecological complexity and pest control in organic coffee production: Uncovering an autonomous ecosystem service. *Bioscience* 60: 527–537.

Vandermeer JH (1992) *The Ecology of Intercropping.* New York, NY: Cambridge University Press.

Williams-Guillén K, Perfecto I, and Vandermeer J (2008) Bats limit insects in a neotropical agroforestry system. *Science* 320: 70.

Yachi S and Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences* 96: 1463–1468.

Zehnder G, Gurr GM, Kühne S, *et al.* (2007) Arthropod pest management in organic crops. *Annual Review of Entomology* 52: 57–80.

Zhu Y, Chen H, Fan J, *et al.* (2000) Genetic diversity and disease control in rice. *Nature* 406: 718–722.

Biodiversity as a Commodity

Geoffrey Heal, Columbia University, New York, NY, and Stanford University, Stanford, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 359–376, © 2001, Elsevier Inc.

Glossary

First theorem of welfare economics If all goods are private and all private and social costs are equal, then an economy with a complete set of competitive markets operates in a way that is Pareto efficient.

Insurance A contract under which a person or institution buys the right to be compensated in the event of a specified harmful occurrence. The payment, referred to as the insurance premium, is paid before it is known whether the harmful event will occur.

Intellectual property rights The rights of developers of ideas and techniques to require payment for their use by others and to prevent their use by others unless such payment has been made.

Invisible hand The term used by Adam Smith to describe the capacity of a decentralized market system to attain efficiency.

Pareto efficient A pattern of operation of an economy is said to be Pareto efficient if it is impossible to change its operation so that everyone gains or at least someone gains and no one loses.

Public good A good whose consumption is nonrival and nonexcludable. Nonrival means that one person's consuming it does not preclude another from doing likewise. Nonexcludable means that the provider of the good cannot ensure that only those who have paid can benefit from its provision. Knowledge is a public good: My knowing something does not conflict with your knowing the same fact, and those who develop knowledge cannot ensure that only people and institutions that have contributed to the costs can benefit. An apple, in contrast, is a private good: If I eat it, you cannot. Also, apple producers can ensure that only those who pay for them can eat them.

Social costs The total costs to society of an action. These may exceed (or in some cases be less than) the private costs, which are the costs of that action to the individual or institution executing it.

Many distinguished scientists have argued strenuously for the conservation of biodiversity as a significant human priority (Ehrlich, 1988; Raven and Williams, 1999; Wilson, 1988). Because economic activity is the main driver of biodiversity loss, such conservation would undoubtedly have important economic implications. It would require economic changes and would certainly be associated with economic costs, although these could be more than offset by the gains. The proposal implicit in these arguments is in effect that we—society—should "buy" biodiversity by changing our economic modus operandi and incurring conservation costs.

What are we buying if we make such a deal? What is the nature of biodiversity as an economic commodity and why does it matter? How would its conservation contribute economically to our well being? These are the themes I address in this article. I consider three issues:

What are the economic functions of biodiversity?
What kind of commodity is biodiversity?
Does our usual economic mechanism, the market system, have the capacity to appreciate the economic value of biodiversity?

In other words, I first characterize biodiversity from an economic perspective, and then consider the capacity of our main economic institutions to realize the value of biodiversity and ensure that it is treated in a way commensurate with its importance. With regard to the first of these tasks, assessing why biodiversity is important economically, I draw extensively on recent literature in ecology, which I do with some trepidation because this is not my field of expertise.[1] However, this is unavoidable: A serious analysis of the economic contributions of biodiversity has to draw on scientific understanding of how diversity contributes to the functioning of the natural environments that provide crucial infrastructure to human societies.

I begin with an attempt to explain what biodiversity contributes, in economic terms, to human societies. Why is biodiversity important? The reasons can be classified under the following headings. Biodiversity provides or enhances:

Ecosystem productivity
Insurance
Knowledge
Ecosystem services

There is some overlap between these concepts, but nevertheless they are helpful as a guide to thinking through the issues. All of them are economically important categories.

Biodiversity and Productivity

How does biodiversity contribute to productivity? There is experimental evidence that plant systems with more biodiversity are on average more productive than those with less biodiversity. A good illustration of this is work done by David

[1]My understanding of the relevant issues has benefited from conversations with Gretchen Daily, Paul Ehrlich, David Tilman, and Peter Vitousek.

Tilman at the University of Minnesota (Tilman and Downing, 1994; Tilman *et al.*, 1996, 1997). He planted many similar plots of land with a variety of grassland plants, some with many species and some with a much smaller number. Each plot was planted with the same mix year after year, and several indicators of plot performance were recorded, including the amount of biomass grown and the proportion of the nutrients available that were taken up by the plants. Biomass refers to the total dry body weight of the plants: It is a measure of the amount of carbon from the atmosphere that is photosynthesized into carbohydrate. Tilman and others performing similar studies found that, on average, during a period of about 20 years, plots with a more diverse collection of species performed better than those with a less diverse collection.

What does more or less diverse mean here? There are two dimensions to diversity: diversity of functional groups, or of plant types, and diversity of plant species within a functional group. Plants are classified into functional groups on the basis of their intrinsic physiological and morphological characteristics, such as whether they fix nitrogen, have three carbon or four carbon photosynthetic pathways, or are woody (Tilman *et al.*, 1997). These characteristics influence the plant's resource requirements, seasonality, and life history. A key aspect of diversity is measured by the number of different functional groups represented by the plants on the plot. This is called, not surprisingly, functional diversity. Species diversity refers to the number of different species within each functional group, or to the total number of species present. This latter measure is sometimes called diversity per se. Clearly, there is a correlation between diversity per se, the total number of different species present, and functional diversity. One cannot add more species without eventually adding more functional groups as well. Another related determinant of productivity is the composition of the functional groups present: Productivity may depend not only on the number of such groups represented but also on their identities because some groups may be more important than others in contributing to productivity or resilience.

The average amount of biomass grown per year on a plot of a given size increased with the diversity of functional groups represented, as shown in **Figure 1**. The increase leveled off after a certain point. Tilman and co-workers also found more nutrient uptake and better soil quality on plots with a more diverse collection of plant species. Furthermore, the plots that were more diverse in this sense were also more robust in the face of climatic fluctuations. It appears from this work that both functional diversity and species diversity are important in maintaining productivity and resilience. Having functionally similar plants that respond differently to environmental fluctuations contributes to resilience. It ensures that whatever the environmental conditions, there will be plants of a given functional type that thrive under those conditions. The functional composition of the community will therefore not be changed by environmental fluctuations. In contrast, if the members of different functional groups respond differently to environmental fluctuations, then these fluctuations will alter the functional composition of the community and therefore its ecological characteristics, and it will be less resilient in the face of such fluctuations (Chapin *et al.*, 1997). Chapin *et al.* (p. 503) argue that

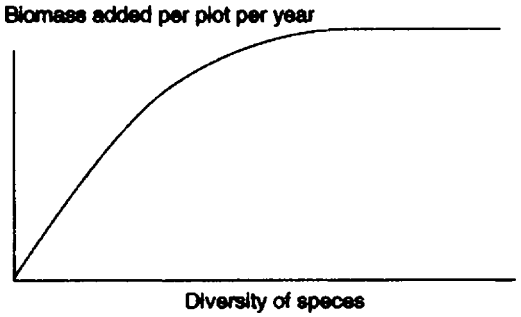


Figure 1 Average amount of biomass grown per year on a plot of a given size.

> Genetic and species diversity per se are important to long-term maintenance of community and ecosystem structure and processes. This argues that no two species are ecologically redundant, even if they appear similar in their ecosystem effects under one particular set of environmental conditions.

What is the mechanism behind these results? There is still dispute about this. One possibility is that each plant type has a range of climatic conditions to which it is best suited. Climate varies from year to year in terms of temperature, rainfall, and in many other ways, and if a plot contains only one plant type then in many years it will not have any plants well adapted to the climates of those years. If, however, a plot carries many types, then in most years there will be some that are well adapted to the climate of that year, and on average its productivity will be greater. This is another illustration of the old proverb, "Don't put all your eggs in one basket." Analytically, it illustrates the same point as the economic argument for holding a diversified portfolio of stocks. The more diversified the portfolio, the less vulnerable one is to conditions that are bad for particular stocks or stocks in a particular sector of the economy. A robust portfolio should have stocks that do well in times of growth, stocks that do well in times of high interest rates, and stocks that do relatively well in times of recession. It should have different types of stocks. Diverse plots will also be more resilient to climatic variation, as a part of the same phenomenon. This is related to the insurance role of biological diversity, here manifested as higher average productivity. An alternative explanation for the greater productivity of mixed plant communities has been suggested based on the interactions between the fungal communities associated with the roots of the plants and the diversity of the fungi in the soil, which affects the effectiveness of nutrient uptake (van der Heijden *et al.*, 1998; Read, 1998).

Similar studies have been conducted for microbial communities and have found similar results. Again, they show that more diverse communities are on average more stable and robust in the face of environmental fluctuations (McGrady-Steed *et al.*, 1997; Grime, 1977; Hooper and Vitousek, 1997; Naeem and Li, 1997).

There are other arguments about why diversity raises productivity, mostly specific to particular biological communities. Trees in forests provide a good example. Some trees are tolerant of shade and others of bright light. A forest of a uniform tree species will consist entirely of trees of one of

these types. In contrast, consider a forest with tall shade-intolerant trees forming the upper canopy. These need bright light and in the upper canopy leaves receive this light. Below them are shorter trees that are more shade tolerant, and perhaps below these is another layer of even more shade-tolerant trees or shrubs. With such an arrangement of diverse species, bright light falls on those that most need it, and the light that passes through their canopy and is of lesser intensity then falls on plants well suited to it. Total photosynthesis—that is, conversion of carbon dioxide in the air to carbohydrates with the aid of light energy—will be greater under such a regime than in a single-species forest (Aber and Melillo, 1991).

These arguments show that diversity is important in ensuring the productivity and robustness of natural plant communities, and therefore of the ecosystems that are based on them. Diversity also helps natural ecosystems to make the best adjustments to conditions that vary over time or over space. Without the appropriate level of diversity, natural ecosystems cannot adjust to natural variations in the environment.

Through its role as the raw material in plant breeding, biodiversity also contributes substantially to the productivity of agricultural systems. New and higher yielding plant and animal varieties are generated from the natural variation in plants and animals. The large increases in grain yields of the "green revolution" of the 1960s and 1970s, which were responsible for keeping food output increasing in parallel with populations in developing countries, were largely achieved by use of genetic diversity in the plant populations. Estimates suggest that as much as $1 billion has been added to the value of the U.S. agricultural output each year for the past half century as a result of plant breeders' use of genetic diversity. Specifically, in the past half century there has been a doubling in yields of rice, barley, soybeans, wheat, cotton, and sugarcane; a threefold increase in tomato yields; and a quadrupling in yields of maize, sorghum, and potato (U.S. Congress Office of Technology Assessment, 1987). All this has been based on and derived from genetic variability in the underlying plant populations. In economic terms, this variability is an asset—and one that has yielded a high return at little cost. Diversity also contributes to productivity in agriculture through the practice of crop rotation. Changing from one crop to another from one year to the next can enhance soil productivity by increasing the nutrients in the soil. For example, rotating a crop such as wheat with a legume that fixes nitrogen can prevent long-term nitrogen loss and reduce the need for nitrogen fertilizers. This type of crop rotation was characteristic of traditional agricultural practices in medieval Europe.

It is important to note that although biodiversity contributes to the productivity of both natural ecosystems and agricultural systems, it does so through different mechanisms. Natural systems benefit directly from a diverse mix of species: Agricultural systems benefit from the existence of a pool of genetic variability on which breeders can draw. Agricultural systems are usually monocultures, consisting of a single species grown intensively over a large area. Its growth is supported by applications of water, fertilizers, pesticides, and weed killers. Farmers manage cropland so as to ensure that crop growth is not limited by lack of water or nutrients, and that the main food crop does not have to compete with other species or with pests during its growth. Farmers create and maintain an artificial environment and then plant a crop that is optimally adjusted to this environment—an approach that is radically different from the natural growth process.

Biodiversity and Insurance

A dramatic illustration of the insurance role of biodiversity comes from the recent history of rice production. The prosperity and comfort of literally billions of people depend on the rice harvest. In the 1970s, a new virus, the grassy stunt virus, carried by the brown plant hopper, threatened the Asian rice crop. This appeared to be capable of destroying a large fraction of the crop and in some years destroyed as much as one-fourth of the crop. Developing a form of rice resistant to this virus became of critical importance. Rice breeders succeeded in this task with the help of the International Rice Research Institute (IRRI) in the Philippines. The IRRI conducts research on rice production and holds a large seed bank of seeds of different varieties of rice and the near relatives of rice. In this case, the IRRI located a variety of wild rice that was not used commercially but which was resistant to the grassy stunt virus. The gene conveying resistance was transferred to commercial rice varieties, yielding commercial rice resistant to the threatening virus. This would not have been possible without genes from a variety of rice that was apparently of no commercial value. Without this variety, the world's rice crop, one of its most important food crops, would have been seriously damaged. Interestingly, the variety of wild rice that was resistant to the virus was found in only one location—a valley that was flooded by a hydroelectric dam shortly after the IRRI found and took into its collection the critical rice variety. The same situation was repeated later in the 1970s, and similar situations have occurred with other food crops, particularly corn in the United States (Myers, 1997). There is every reason to expect that events such as these will recur regularly: Planting large areas with genetically identical plants greatly increases the chances that once a disease starts it will spread with dangerous speed through the entire area and crop. A report by the Committee for Agriculture, Science and Technology (1999, p. 13) emphasized this point:

> Because of the increasingly high densities and large areas over which they are now grown, both livestock and crop plants are continually acquiring new diseases and pests, and existing diseases and pests are continually evolving new strains that overcome the defenses of particular breeds or strains. This is exacerbated by the accidental transport of diseases around the world. These diseases and pests destabilize agricultural systems. For instance, areas of western Minnesota and eastern North and South Dakota no longer can produce viable wheat and barley crops because of new strains of scab and vomit toxin for which no crop varieties have sufficient genetic resistance. Indeed, catastrophic attacks of disease, invasions of insects, and climatic extremes have caused wholesale crop destruction and ensuing famines whenever crops had insufficient diversity to provide at least some plants with the ability to withstand the assaults. Outbreaks of avian flu in the Chesapeake Bay area regularly result in rigorous quarantines of poultry houses due to the extremely high density of poultry farms in this region.
>
> Disease problems, as old as agriculture, are recorded in myth and in written history, and still exist. Red rust on wheat in Roman times, mass poisoning from ergot-tainted rye during the middle ages, the Irish potato famine of the 19th century, and the Southern

corn leaf blight in 1970 all were due to insufficient biodiversity in the affected crops. The severity of the 1998 Hong Kong chicken epidemic was likely exacerbated by the lack of diversity in disease resistance as well as by the high chicken densities in the production facilities.

The continual accrual of new diseases can be countered only if breeders can find sufficient genetic diversity within a crop or its relatives. Even the full complement of natural genetic variation, though, may not be sufficient to stop some diseases. Consider, for instance, the impacts of chestnut blight, an introduced disease that devastated what was once the dominant tree of the eastern United States, but which now occurs only as rare stump sprouts. Despite the vast geographic expanse and genetic diversity of the native North American chestnut, there is no known genetic resistance to its pathogen. Similarly, in vast areas of west and central Africa, livestock genetic resistance to the debilitating effects of trypanosomiasis is found only in a few unproductive local breeds. Despite massive efforts, the genetic mechanisms governing this resistance are not yet well understood.

A lethal disease of corn, or wheat or rice, were it to appear, would devastate agriculture and human society. The only insurance that society has against such a catastrophe is biodiversity. Genetic diversity within a crop plant or animal species and its relatives might allow resistant strains to be discovered and used. Similarly, a diversity of potential food plants might allow another species to become an effective substitute for a major crop species that was lost to disease.

These cases illustrate clearly the insurance role of biodiversity. It is an important defense against disaster in the form of new diseases. The pathogens that cause disease are evolving continually, in an attempt to outwit our defenses against them. A clear example of this phenomenon is the evolution of antibiotic resistance among bacteria. The bacteria that cause several once common diseases in humans are now showing resistance to their principal controls, to the great concern of public health authorities. The same is happening with the pathogens that cause disease in crops and in commercial animals. Without reserves of genetic variability we may not be able to develop varieties of our agricultural crops and animals that can resist these new disease varieties. Indeed, it is precisely genetic variability in the pathogens that allows them to develop resistance. Genetic variability means that some of the disease-causing pathogens are naturally relatively unaffected by our defenses against them, which may be in the form of weed killers, insecticides, or vaccinations for livestock. These more resistant specimens are the ones that survive and from which new subsequent generations are bred. Therefore, pathogens use against us the mechanisms that we will use against them if we preserve and use genetic diversity. Without this diversity, we have disarmed unilaterally in the war against our most threatening enemies.

There is another important role for the insurance provided by biodiversity—to provide variability that could be critical in responding to the environmental changes wrought by humans. Human activity is changing the climate and the sea level, and it is making many more local changes in the environment. A hotter climate may require different crop varieties. An increase in sea level may lead to increased salinity in groundwater, and therefore to a need for crop varieties that are salt tolerant. A good example of the value of diversity in the context of a changed environment is the evolution of plants that have grown on mine wastes in the United Kingdom. These wastes are rich in heavy metals and are poisonous to most plant species (Antonovics *et al.*, 1971). The existence of a pool of genetic diversity allows us to find plants that could tolerate these poisons and even help to remove them from the soil. We are making changes to the global environment on an unprecedented scale, and biodiversity might be critical in allowing us to respond to the consequences of these changes. Population growth and environmental change mean that we now need the insurance provided by biodiversity more than ever before.

Biodiversity and Genetic Knowledge

The third reason I previously gave for the importance of biodiversity is that it is a source of knowledge. We can learn from natural organisms how to make chemicals that have important and valuable properties. A good example is provided by the polymerase chain reaction (PCR). This reaction is central to culturing DNA specimens for analysis—as in forensic tests used in trials such as the O. J. Simpson trial, and in many processes central to the biotechnology industry. Culturing requires an enzyme that is resistant to high temperatures. Enzymes with the right degree of temperature resistance were found in hot springs in Yellowstone National Park, and the heat resistance of these was then used to create an enzyme that could be used to culture DNA specimens. This enzyme is now central to the rapidly growing biotechnology industry. There are many more less complex examples. In fact, 37% by value of the pharmaceuticals sold in the United States are or were originally derived from plants or other living organisms (Carte, 1996). Aspirin is derived from the bark of willow trees. The bark of yew trees has been used to derive a drug that is effective against ovarian cancer (Stierle *et al.*, 1993). A derivative of the rosy periwinkle flower is being used to cure childhood leukemia. The key point is that certain plants and animals are known to produce substances that are highly active pharmacologically. Plants that live in insect-infested areas produce substances that are poisonous to insects, and these have been used as the basis for insecticides. Some snakes produce venom that paralyzes parts of the nervous system, and others produce venom that reduces blood pressure. Some insects produce anticoagulants. All of these have been adapted for medical use. There is little that is new in these observations: They form the basis for many traditional medicines, which rely heavily on plants. Shakespeare refers to this in *Romeo and Juliet* (II: iii):

O! mickle is the powerful grace that lies
In herbs, plants, stones and their true qualities:
For nought so vile that on earth doth live
But to the earth some special good doth give,
Within the infant rind of this weak flower
Poison hath its residence and medicine power.

Biodiversity and Ecosystem Services

I previously mentioned the role of genetic diversity in providing raw material for selective breeding, the traditional way of developing new crop or animal varieties that are more productive, more disease resistant, hardier, or more desirable in some other way. I have also mentioned its role in ensuring

the productivity of ecosystems, and in ensuring their robustness against diseases and pathogens. There are other more complex ways in which biodiversity is essential to the proper functioning of ecosystems and to the delivery of the ecosystem services on which human beings are so dependent.

There are cases in which the full diversity of organisms in an ecosystem is required for that system to function and to provide services to human societies, and the removal or addition of even a single type of organism can have extraordinarily far-reaching consequences. "Keystone species" provide a convincing illustration. Ecologists use the term keystone species to describe a species whose removal will cause an entire ecosystem to change substantially. A widely cited example is that of sea otters on the California coast. The removal of sea otters as a result of hunting them for their pelts led to far-reaching and undesirable changes in the California coastal ecosystems. Sea otters eat sea urchins, which in turn graze on kelp plants. Without control of the urchin population by otters, the urchins will destroy the kelp beds, completely changing the marine coastal environment. Removal of otters led to a greatly impoverished coastal environment, which was restored in part to its original state with a ban on otter hunting. Another example of the role and impact of a keystone species is provided by the removal of kangaroo rats from an area of the Chihuahuan desert, which led to a threefold increase in the yields of grasses and to far-reaching changes in the desert ecosystem. In this case, the rats had played a key role by eating seeds and disturbing the soil, and their removal consequently changed the plant balance (Power *et al.*, 1996).

Not only can the removal of a species lead to major changes in an ecosystem but also the introduction of a new species (a so-called exotic species) can lead to a profound transformation of the system. A dramatic example is the introduction of the rinderpest virus into East Africa in 1890. This initially attacked domestic and wild cattle and then spread. By 1892, 95% of the wildebeest in the Serengeti region had died, in addition to most of the domestic cattle. Wildebeest are one of the main grazers and also the main food sources for carnivorous predators (lions, leopards, and hyenas) in the Serengeti; therefore, their virtual elimination led to profound changes in the system. In the 1930s, the introduction of a vaccination against rinderpest reestablished the original system (Aber and Melillo, 1991). The point of these examples is that we cannot easily tell a priori which species are essential and which are not. There is often a risk that an apparently small change in a set of species will have effects far beyond those initially anticipated. The degree of interdependence between different species is great; therefore, human beings may depend on many more species than we would expect from a first analysis of the situation. Abelard (an eleventh-century French theologian) suggested that any organism has a role to play and a reason for existing: "Whatever is generated is generated by some necessary cause, for nothing comes into being except there be some due cause for it" (as quoted in National Research Council, 1999). Also, John Donne (as quoted in Hemingway, 1940), an English metaphysical poet of the seventeenth century, wrote that

> No man is an island, entire of itself; every man is a piece of the continent, a part of the maine; if a clod be washed away by the sea, Europe is the less, as well as if a promontory were, as well as if a manor of thy friends or thine own were; Any man's death diminishes me, because I am involved in mankind; And therefore never send to know for whom the bell tolls: It tolls for thee.

There is an ecological equivalent to this: No species is an island, entire of itself, not even *Homo sapiens*. Any species' extinction may diminish us because we depend on many species. To repeat: The loss of even apparently unimportant species can have immensely costly consequences because of the complex patterns of interdependence between species. In the end, the loss of an apparently small and unimportant group of species could threaten the provision of ecosystem services that are essential to humanity. The distinguished biologist E. O. Wilson once said of microbes that "We need them but they don't need us." This is why many scientists see a serious risk in the current rate of species extinction: They cannot be precise about the dangers involved but nonetheless believe that there is a real risk of costly consequences. To give this point some substance, I mention a possible relationship between the extinction of passenger pigeons and the introduction of Lyme disease into American society. When Europeans first arrived in the United States, the passenger pigeon was probably the most abundant bird in the country. Its population was estimated in the billions. It traveled around in flocks of hundreds of thousands, flocks so large that their passing darkened the sky for many minutes at a time. By 1914 they were extinct, annihilated by a combination of hunting and destruction of the habitat that they needed for survival. It seemed unbelievable that an animal so abundant could be reduced to extinction so fast. A possible connection between this extinction and the emergence of Lyme disease events has recently been proposed. A letter to *Science* in 1998 made the following suggestion (Blockstein, 1988, p. 1831):[2]

> There is another possible twist to the complicated ecological chain of events presented by Clive G. Jones *et al.* (*Reports*, 13 Feb., p. 1023) whereby the incidence of Lyme disease might increase following population increases of mice allowed by a big mast year of acorns. A major competitor of deer and mice for these bumper crops has been absent from the eastern deciduous forests for a century. The extinct passenger pigeon (*Ectopistes migratorius*) was a nomadic wanderer that specialized on a diet of the superabundant, but unpredictable, crops of mast. With a population estimated at 2 to 5 billion, concentrated in enormous flocks, passenger pigeons congregated wherever there were huge crops of mast. The birds were so efficient at denuding the woods of nuts that many observers noted that native wildlife and feral hogs could not find sufficient food after a pigeon flock had passed through. Is it possible that, in the presence of passenger pigeons, the population explosions of mice in mast years, reported by Jones *et al.* would have been less likely. Could the outbreaks of Lyme disease in the late 20th century have been a delayed consequence of the extinction of the passenger pigeon?

The point here is that passenger pigeons ate acorns and beechnuts, both of which were abundant in the forests of the northeastern United States. The demise of the pigeons led to an increase in the food available for other animals that ate these, including mice. Mice are the main breeding ground and hosts of the parasites that cause Lyme disease, and it is reasonable that the explosion of food for mice led to an increase

[2] I am grateful to Paul Ehrlich for this reference.

in their population and thus in the population of Lyme disease vectors. (Abundant acorn crops always lead to increases in the population of mice.) The disease vectors transfer from mice to deer, which browse in the same forests and on the same foods and then move across territory likely to be used by humans, grazing on grass on lawns and fields. Therefore, the extinction of passenger pigeons could have been instrumental in causing the spread of Lyme disease to humans. This illustrates well the extraordinary complexity of the web of life and of the connections between different species and between species and human welfare. No one could reasonably have anticipated this connection between passenger pigeons and Lyme disease. No analysis of the consequences of the loss of this bird could have anticipated such an outcome. Indeed, the bird was so abundant that it must have been difficult to anticipate that human activity could drive it extinct.

The message of these examples is that it is difficult to foresee the consequences of a change in the biodiversity of an ecosystem. Even an apparently small change can lead to dramatic alterations in the system's ability to function and to provide the services on which human beings are dependent. There is another aspect of this phenomenon. A particular role in an ecosystem may be played at different times or in different circumstances by quite different plants or animals. The type of tree that stabilizes soil on a north-facing slope at a certain latitude may not grow on a south-facing slope at that latitude so that a different species is needed there to maintain the physical stability of the system. As a consequence, the set of species required for a certain type of ecosystem to function may vary greatly from region to region. In fact, we know of no single subset of species that on their own would serve to operate all ecosystems and provide all ecosystem services in all regions of the planet. Therefore, diversity in a given location may increase productivity and ecosystem functions in that location, whereas diversity at the regional or global level is actually necessary for the operation of important ecosystems in all geographic regions. Although individual species may possibly be redundant in some locations, it is possible that on the global scale few if any are really redundant. A clear statement on this topic is given by Chapin *et al.* (1997, p. 505):

> The abundance of species with similar ecological effects should give stability (resistance and resilience) to ecosystems in the face of increasingly rapid human-induced environmental change. Loss of a keystone species or of all species in a major functional group will, by definition, have large ecosystem effects. Efforts to identify and protect such species and groups often yield demonstrable near-term benefits. Of increasing concern is the loss of species that have similar ecosystem effects but differ in their environmental responses. Loss of such species may reduce ecosystem resilience and the capacity to adjust to ever-increasing rates of environmental change. This latter role of diversity is not adequately represented in current international conventions, but it may be one of the most important mechanisms by which we sustain the long-term functioning of ecosystems and the services they provide to society.

Biodiversity and Markets

Returning to the economic questions, to what extent can we hope to commercialize these contributions of biodiversity? Obviously, they are economically important, but can the market capture them? Can the economic contributions of biodiversity be used to generate incomes to the owners of biodiversity that will provide them with incentives to conserve it? Can they make conserving tropical forests more attractive than clearing them? This is a critical question: Forest owners do not conserve forest because they are important to humanity. Rather, if they conserve them, they do so because they can profit from so doing. Therefore, we need to know whether the important services provided by biodiversity can be the basis for profits from forest conservation or for conservation of sources of biodiversity more generally. In this context, I again examine the categories under which I classified the contributions of biodiversity:

Increasing productivity
Providing insurance
Providing knowledge
Maintaining ecosystem services

One of the most fundamental insights into the operation of a competitive market economy is that under certain conditions it will align individual and social interests and provide incentives that lead to an efficient outcome. In the famous words of Adam Smith in his *An Enquiry into the Nature and Causes of the Wealth of Nations* (1977),

> Every individual … neither intends to promote the public interest, nor knows how much he is promoting it. He intends only his own security, his own gain. And he is in this led by an invisible hand to promote an end which was no part of his intention. By pursuing his own interest he frequently promotes that of society more effectively than when he really intends to promote it.

This is a beautiful metaphor: Market forces are an invisible hand, steering us to act in the interests of society as a whole when in fact we only seek to promote our own interests. Recently, this insight was formalized and made more precise via some important propositions from economic theory, including the First Theorem of Welfare Economics. These propositions state that if all goods are private goods (i.e., there are no public goods) and the private and social costs of all activities are equal, then a competitive market economy is Pareto efficient (i.e., operates so as to leave unexploited no possibilities for mutual gain). This is a remarkable result and provides the basis for economists' belief in the efficacy of market systems and the desirability of market-based approaches to economic organization. However, in the context of biodiversity the restriction that all goods be private is critical.

What is the public–private good distinction? A public good is one that has two properties: My consumption does not interfere with yours, and the provider cannot prevent nonpayers from benefiting from the good. Such goods are said to be nonrival and nonexcludable, in contrast with private goods, which are both rival and excludable. A seat at the opera is a private good: If I sit in it you cannot, and the management can certainly exclude nonpayers from seeing the performance. Law and order, in contrast, is a public good: It benefits everyone in the region in which it is enforced and the benefits cannot be restricted to those who have contributed to its costs.

Markets are not good at providing public goods: Their nonexcludability makes it difficult for the provider to earn a good return on the costs of providing them. Knowledge is an interesting and relevant example: It is naturally a public good because it can be passed costlessly from one person to another and enjoyed by all of them, even though none of them, or perhaps only the first, paid for it. Hence the existence of intellectual property rights, instituted as a means of ensuring some return on the generation of knowledge. Because of the difficulty in appropriating the returns on their provision, markets tend to underprovide public goods relative to an economist's concept of what is efficient for society. Consequently, they have traditionally been provided by the public sector. Some of the services provided by biodiversity are public goods, although biodiversity does not fit the traditional mold for public goods completely. The extent of biodiversity is not something that can be determined by the public sector because it is the result of literally billions of land-use choices throughout the world. It is also strongly influenced by issues such as climate change, which are again driven by billions of heating and transportation choices. In fact, biodiversity has been called a privately produced public good. For the remainder of this argument, the key point is that the First Theorem fails in the context of public goods, and indeed it has long been recognized that markets will underprovide public goods relative to the level that would be required for efficiency.

One further basic economic point is needed—a familiarity with the diamond–water paradox and the limitations of market prices as indicators of "importance" to society. Discussion of this paradox helps us to be clear that the price of a good does not reflect its importance in any overall social or philosophical sense. Very unimportant goods can be valued more highly by the market than (have higher prices than) very important goods. The classic illustration of this is the diamonds and water paradox, which perplexed economists through the eighteenth and nineteenth centuries until its resolution by Alfred Marshall. The point here is that water is clearly more important to human society than diamonds, but diamonds trade in the market at prices far in excess of those fetched by water. Why? Marshall's answer was simple and is now part of common knowledge: Price is set by supply and demand. The market price is the price at which the amount supplied is also the amount demanded. In the case of water, the supply (at least in Marshall's time) was so large as to exceed the amount that could possibly be demanded at any price. Consequently, the price was zero: Water was free. Now, of course, the demand for water has increased greatly as a result of population growth and increasing prosperity, whereas the supply has remained approximately constant so that water is no longer free. For diamonds, being naturally scarce, the desire for ownership always exceeded that which could be accommodated naturally. The market price was high as a result of competition between rich people for the few diamonds available.

What are the implications for biodiversity? Simply that even if it is of great importance to society, and is not a public good, it will not necessarily be possible to convert this importance into value in the market place. The balance of supply and demand will be critical here, as shall be shown in the context of the commercialization of the genetic knowledge inherent in biodiversity.

Productivity

To the extent that diversity increases productivity in agricultural systems, we would expect that farmers would be willing to pay for it. Arable farmers achieve some of the benefits of diversity by crop rotation, i.e., rotating between a series of different crops in successive years. The different crop types make different demands on the soil and contribute different nutrients to it. However, the range of crops used for this purpose is quite limited and does not contribute in any substantial way to the conservation of biodiversity. The other mechanism through which diversity contributes to productivity in agriculture is via its contribution to the breeding of new plant species that are better adapted to emerging conditions or more resistant to new diseases. At this point, the productivity and knowledge roles of biodiversity merge, and I comment on the possibility of commercializing the knowledge role later. These comments on biodiversity's contribution to knowledge will also apply to the contribution that biodiversity makes to human societies via the breeding of new varieties.

There appears to be some appreciation of the benefits of diversity in tropical agriculture, in which there is a tendency to grow several crops together or to grow crops in a way that conserves the original forest. Traditionally, coffee was grown as an understory plant beneath high tropical forest trees: This benefits the coffee, which is a shade-tolerant plant, takes full advantage of the light available in the region, and avoids the need to destroy the main forest trees to make land available for coffee growing. This practice also allowed other commercial crops to be grown with the coffee, such as citrus fruits and avocados, thus allowing farmers to diversify their risks. In cases such as this there is a contribution to the preservation of diversity because of the conservation of the forest. Studies have shown that forests converted for production of shade-grown coffee retain a very large proportion of their original biodiversity and that growing in this way is less expensive per pound produced than plantation growing. This cost difference reflects in part the greater productivity of diverse ecosystems and the more effective cycling of nutrients in these plant communities (the total yield of coffee per hectare, however, is less; Perfecto *et al.*, 1996). In this case, it seems that there can be some conditions under which the productivity enhancements of diverse systems can be realized commercially, with attendant benefits for biodiversity conservation.

Insurance

Insurance is clearly something for which there is a demand. Most of us insure our homes and our cars and have health insurance. Therefore, perhaps the insurance role of biodiversity is one for which people will pay. The difficulty here is that until recently this insurance has been provided as a public good—indeed, as a global public good. Consider in the light of the public–private good distinction the conservation of rice varieties by the IRRI mentioned previously and in particular

the use of one of these to provide a defense against the grassy stunt virus. In cases such as this, the insurance was provided as a public good. It was available to all for the cost of buying the new variety of rice incorporating resistance to the grassy stunt virus. You did not have to pay an insurance premium to benefit from the insurance. The developers of the new variety could not exclude from using it those who did not contribute to its development by paying insurance premiums.

At this point, I again digress briefly into economics. Risk, such as the risk of the destruction of a part of a crop, is managed in a market system by contingent contracts. A contingent contract is a contract that pays a specified amount if and only if a specified contingency occurs. Buying the contract reduces the risk to you that is associated with the specified event. If it occurs and has negative consequences, then you are to some degree compensated by the payment made to you under the contract. An insurance contract is a classic example of a contingent contract: You buy it by paying a premium and it entitles you to a payment if and only if a specified event—the insured peril—occurs. It is crucial for the efficient management of risks that the payment through which one purchases a contingent contract should not be contingent. You pay the insurance premium whether you use the insurance or not: The only uncertainty or contingency is whether you can make a claim under the policy.

It follows that the problem of the public good nature of the insurance provided by the existence of new varieties could not be overcome by charging premium prices for them. This price would only be paid after the product had been developed and needed. If the new variety were never needed, then this price premium would never be paid. No contingent contract is being sold in this framework.

There is a possibility that the public good nature of the insurance will change. Public goods can become private goods through either institutional or technical change. A good example is television broadcasting. Until about 10 years ago, this was a public good par excellence: A broadcaster could not exclude from viewing a program anyone in the reception area so that the non-excludability property of public goods held, and of course there is no rivalry in consumption. My viewing a TV broadcast in no way interferes with your viewing. The development of scramblers changed this. A TV broadcast can now be scrambled so that it can only be viewed by those who have purchased a descrambler. Broadcasters can now exclude those who have not paid from viewing their programs, which are therefore no longer public goods. A public good has been privatized.

Developments in the area of intellectual property rights for agricultural biotechnology could change the situation in a similar way for the insurance value of biodiversity. These may lead to the privatization of a hitherto public good. Crop developers are increasingly patenting genes developed to enhance the properties of crops, including such properties as their taste, productivity, and insect resistance. Most agricultural biotechnology companies have a large and growing portfolio of patents on genes and on genetically modified plant varieties. They are also aggressively defending these patents, to the extent of developing and introducing "terminator genes" that will ensure that the properties conveyed by patented genes will not be transferred to offspring of the plant. Consequently, users cannot breed from seed that they have bought but must purchase more from the supplier. It is possible that plant breeders with such a tight hold on their intellectual property will be able to extract from users a great enough return to justify substantial investment in biodiversity conservation. The aggressive enforcement of intellectual property rights might act here like scramblers did with TV broadcasts, effectively privatizing a previously public good. Were it to happen, this would increase substantially the incentives for biodiversity conservation, but possibly at the cost of restricting access to the latest agricultural technology to those with the ability to pay a premium.

There is already a good illustration of the possible side effects of privatization of genetic knowledge. The bacteria *Baccilus thuringiensis* (*Bt*) produces a toxin that kills many crop pests and is itself biodegradable. For this reason, organic farmers use it as a pesticide: Because it is biodegradable it leaves no dangerous residues on the crops on which it is used. Monsanto and Novartis recently incorporated genes from *Bt* into transgenic crops. The presence of these genes in the plants means that the plants benefit from the defense provided by the *Bt* bacteria, which is the production of proteins that are toxic to the main pests of cotton and corn. Because the *Bt* genetic defense against pesticides is now widely used, it is possible that pests will develop resistance to it. Resistance develops faster the more widely a defense is used. Therefore, within a few years there may be generations of crop pests that are immune to this way of defending crops, i.e., immune to the *Bt* toxin. Monsanto and Novartis will seek to develop variants on the *Bt* genes, and with their scientific and financial resources they may succeed. If so, they will sell a new generation of transgenic plants with defenses against the new generation of pests. However, these will be their proprietary products, covered by their patents. Because of the development of resistance to the *Bt* toxin, organic farmers will no longer be able to use *Bt* as a harmless pesticide. They will be forced to buy proprietary defenses against pests. This is a clear illustration of the two factors mentioned previously: development of intellectual property rights in genetic knowledge leading to both stronger incentives to conserve and to develop further and also leading to a restriction of access. In the case of transgenic crops including the *Bt* defenses, Monsanto recently announced that all growers of *Bt* corn will be required to grow plots of nonengineered corn that are at least 20% of the size of the engineered crops. The aim here is to provide sufficient nonengineered corn that the development of resistant pests will be delayed or possibly even prevented. ("Monsanto concession", 1999).

Knowledge

Here, I discuss the role of biodiversity as a source of knowledge. There are two different areas of application here: the development of medical products by the pharmaceutical industry and the development of new or better crops by the agricultural biotechnology industry.

In applications of biodiversity to the development of pharmaceuticals there has already been some progress toward commercialization. Recognition of the likelihood that tropical plants contain chemicals that could be forerunners of

pharmaceuticals has led most major drug companies to pursue bioprospecting as a way of finding new pharmacologically active substances to serve as a basis for drug development. Typically, they have sought these compounds in the tropics, in areas where there is extensive interspecies competition, or in other extreme areas. They have been willing to pay quite substantial sums for access to these regions, and they have made deals with host countries that involve giving them a royalty on the products that might eventually be based on their prospecting. Such royalties could be large relative to the incomes of the countries concerned. Merck, Inc., one of the largest pharmaceutical companies in the United States, has an agreement with a Costa Rican agency called InBio (Institute National de La Biodiversidad) for bioprospecting rights in Costa Rica. The terms of the agreement are that Merck paid InBio a fixed sum, $1.35 million, to be used for forest conservation in exchange for the right to receive samples collected by InBio and to use these as the basis for new product development. Should any of them prove commercially successful, Merck will pay InBio a royalty on the revenues generated. Similar agreements are in place between other U.S. pharmaceutical companies and regions of Central and South America.

The discovery of the PCR enzyme and the agreement between Merck and InBio and several other drug discoveries based on plants from developing countries led to a wave of optimism, perhaps excessive, about the potential commercial value of *in situ* biodiversity in developing countries. What in fact is the commercial potential here? There is no question that pharmaceutical and agricultural products of great human and commercial value have been and will in the future be developed from the biodiversity in tropical countries. The key question is how much of this will be returned as a reward for the conservation of the originating biodiversity. In answering this question, one must take note of several points. The first point is related to the discussion of the diamond–water paradox: Not everything that is important will have value in the marketplace because there is nothing associated with being important that rules out the possibility that the supply may exceed the demand. Some commentators have seen this as a critical issue in the market for biodiversity-based genetic information. There are literally millions of organisms in the world that might provide genetic information, and if we do not know which ones will provide valuable information and which will not, then the supply of potential genetic leads is huge, possibly greater than pharmaceutical companies can process. In such a situation, the market price for such leads would be near zero. However, recent calculations suggest that in a small number of the world's biodiversity hot spots, bioprospecting rights may be worth as much as $9000 per hectare. This is small in relation to the amount that might ultimately be derived from drug sales but still large relative to other uses of the land. In fact, it is approximately a century's worth of ranching income (Rausser and Small, 1998). The key insight in these calculations is that prior knowledge of the nature of the ecosystems in a location can improve estimates of the probability of finding commercially interesting compounds there and can suggest where search would be profitable and where success is unlikely. In this case, not all organisms or all leads are equivalent: Some have much greater chances of success than others. Prior scientific information can change the odds of success from one 1 in 10,000 to approximately an order of magnitude better. In practical terms, this means that developing countries can clarify the commercial attractions of their biodiversity by performing research on the ecosystems of which it is a part. This is similar to a country with potential oil reserves engaging in basic geological prospecting before seeking to negotiate leases for oil development. The results may be positive or negative, but either way they will give the country a better view of its prospects. In those cases in which the research is positive, the impact on the value of prospecting rights could be large.

Another important point is that an immense amount of human skill and expertise are needed to develop a plant specimen into a commercial drug. Typically, there will be a minimum of 10 or more years of work by hundreds of skilled people working with millions of dollars of sophisticated equipment. Unfortunately, very few plant extracts actually produce drug leads (i.e., contain pharmacologically active compounds with no obvious ill effects)—probably less than 1 in 10,000. Of these few, very few become commercial drugs—less than 1 in 100. On average, perhaps 1 in 250,000 samples collected leads to a commercial drug. ("When Rhetoric Hits Reality," 1998).

Therefore, the chance of any individual bioprospecting operation leading to a commercially valuable drug is very small indeed. Also, even if it does produce a drug, tens or even hundreds of millions of dollars will have to be invested, with significant chances of failure. Developments in biotechnology are currently altering this picture. They are reducing the time needed for testing and development and giving greater insights into the kinds of chemicals likely to be successful. By reducing the costs of drug development based on bioprospecting, they are making bioprospecting more attractive. For example, the cost of screening 10,000 samples for pharmaceutical potential 10 years ago would have been $6 million: Today it is $150,000 (Reid *et al.*, 1993). Simultaneously, advances in knowledge are also making more effective alternative methods of drug development, based on understanding of the cellular and genetic mechanisms of disease. In total, the picture that emerges is one of heavily guarded optimism. Bioprospecting does have economic value, and technological developments may be increasing that value. However, in the short term we cannot expect great sums of money to flow to the conservation of biological diversity because of bioprospecting possibilities.

There is another problem to be overcome in establishing an income flow from bioprospecting—a problem with intellectual property rights. The same plant may occur in several different regions, and the same or similar chemicals may occur in different plants. Therefore, the same or similar drugs may be derived by different routes from different plants or different geographic regions. Research toward a commercial product has to be well under way before it is patentable; therefore, there is always a risk of being blocked by a prior patent.

There is also a risk, regarding the conservation of biodiversity, that biodiversity is valuable but leads to no direct

commercial application. In a recent article on the value of marine bioprospecting, Carte (1996, p. 284) stated,

> Although many of these products are not likely to become therapeutics, the information gained from studying them is likely to lead to the development and understanding of novel molecular targets, which may in turn lead to the development of new therapeutic agents.

This is a classic statement of the importance of basic knowledge. Basic knowledge is a public good: Its importance is great but it is not patentable and not something that can be appropriated by a typical bioprospecting contract with an element of royalty payment or revenue sharing. The development of basic knowledge is typically publicly funded precisely because its economic benefits, although potentially immense, are difficult to appropriate. More constructive thinking about how to realize the undoubted importance of bioprospecting in terms of income for conservation is needed. A further illustration of this point is provided by the example of taxol mentioned earlier. Taxol is a promising anti-tumor agent in breast and ovarian cancers that can be extracted from a fungus that lives in the phloem (inner bark) of the Pacific yew tree. Taxol was first isolated from the tree itself, but the tree is relatively rare and slow growing, and it produces little taxol; therefore, a search for other sources was initiated (Stierle *et al*., 1993). Ultimately, little in the way of economic returns may flow to the regions in which taxol was discovered.

Biodiversity can also be applied to the development of new or better crops: To date, this has probably been the most important commercial application of biodiversity. As noted previously, the existence of a pool of genetic variation provides plant and animal breeders with the raw materials for developing new varieties and more productive or resilient variants of existing varieties. The existing varieties of a commercially important crop are usually the property of a commercial firm or of a research facility and are protected by patents. For example, the University of California owns the patents of many varieties of strawberries. These patents cover varieties that are best suited to different soil types and different weather conditions, that are least prone to spoiling during transport and storage, and so on. For most other commercial crops, the varieties are owned by seed companies, whose main asset is often the intellectual property represented by their ownership of patents to widely used varieties. In this context, the market can certainly recognize the value of biodiversity, provided that it is of the type that seems likely to contribute to the development and refinement of commercially important crops. Unusual variants of commercial crops, such as early variants of wheat, corn, or soybeans, would qualify, and possibly so too would their near genetic relatives. However, biodiversity more broadly would probably not derive a value through this process, even though genes from unrelated plants might enhance the commercial potential of existing crops.

Ecosystem Services

Perhaps the most promising approach is to consider selling the services of natural ecosystems and using the revenue to provide incentives for conserving the biodiversity that supports them. Selling services provided by natural ecosystems can potentially provide incentives for the conservation of these systems and thus indirectly of biodiversity. For example, watersheds provide economically valuable services for which there is a market, and indeed recognition of this has already led to the conservation of significant forest areas (Chichilnisky and Heal, 1998). Some commentators (Reid, 1998) have suggested that as much as 15% of the earth's land area serves as watersheds for large cities and so could legitimately be conserved on the basis of the watershed services that it supplies. Many of these watersheds are areas of considerable biological uniqueness and their conservation would be a major advance for the conservation of biodiversity. Their functioning as watersheds probably depends substantially on the continuation of their current levels of biodiversity.

Ecotourism is based on the preservation of intact ecosystems and the more appealing elements of biodiversity to be found in some of these systems, and revenues from this are providing powerful incentives for the conservation of several important ecosystems (Chichilnisky and Heal, 1998; Heal, 2000a). Insofar as ecotourism conserves tropical ecosystems, it contributes to the conservation of biodiversity. In certain regions, particularly southern Africa, ecotourism is making a major contribution to the conservation of biodiversity. In Angola, Botswana, Kenya, Malawi, Mozambique, Namibia, South Africa, Tanzania, Zambia, and Zimbabwe, approximately 18% of the total land area is now devoted to the support of wildlife. This is both a significant amount in total and a massive increase relative to 15 or 20 years ago (Bond, 1993; Cumming, 1990a, and b; Cumming and Bond, 1991). This has led to an increase in the populations of several previously endangered species and to more robust breeding populations of many important birds and mammals. All of this has been driven by the fact that in many parts of this region land devoted to ecotourism and sport hunting can earn a higher return than land devoted to more conventional agriculture. **Figure 2** illustrates these phenomena: It shows for Zimbabwe both the growth increase in land whose primary use is the support of wildlife and the favorable economic returns to wildlife conservation relative to agriculture in parts of the region. (Regions III–V are either arid regions or wetlands.) Although the growth of ecotourism as an economically important activity has increased most in southern Africa, this region is not unique in being able to generate conservation incentives by tourism. Central and South America, and parts of Asia, are now benefiting from ecotourism, which is providing significant returns to the conservation of certain endangered species in these regions (Heal, 2000b; Freese, 1999; Honey, 1999).

Finally, there is a real prospect of commercializing some of the carbon sequestration services of forests under the terms of the Kyoto Protocol, which provides for compensation for some carbon sequestration activities via its provisions for joint implementation and through its clean development mechanism. I have estimated elsewhere (Heal, 2000b) that this could generate an income as high as $50–100 per hectare per year, which is high enough to radically change the incentives for forest conservation and hence biodiversity conservation.

There are also other economic mechanisms through which the goods and services provided by tropical forests and their

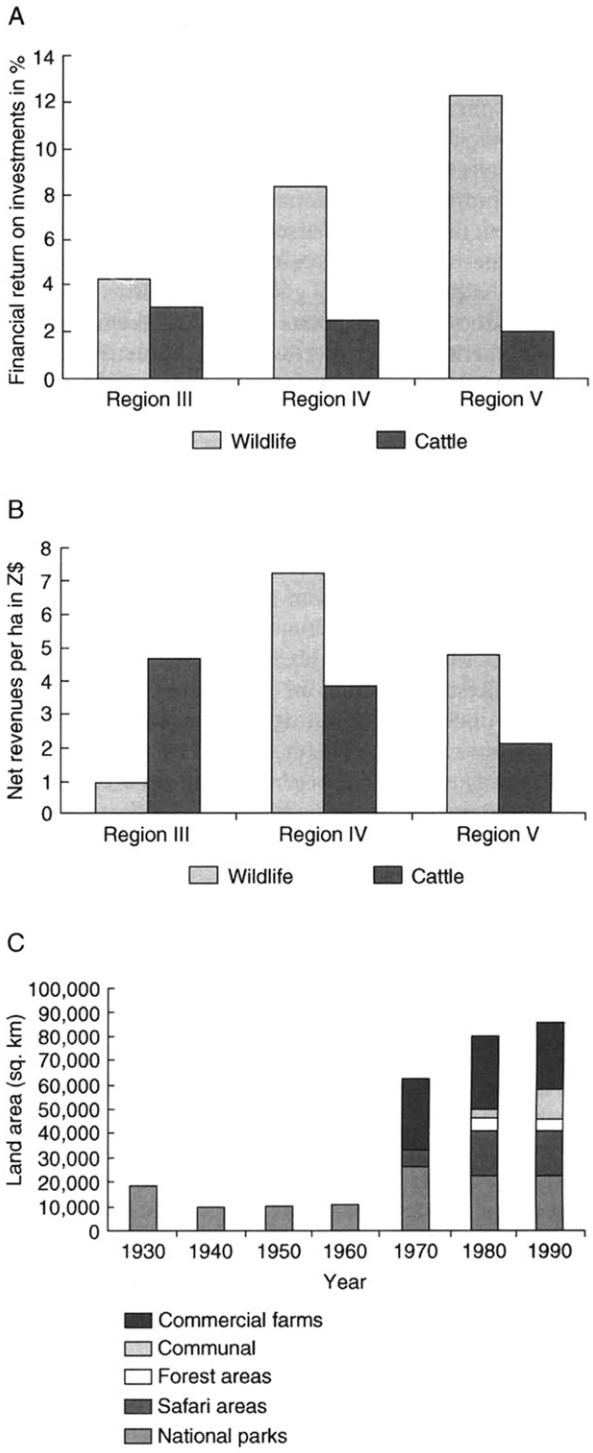


Figure 2 (A) Financial returns to wildlife and cattle ranching in Zimbabwe. (B) Returns per hectare from cattle and wildlife in Zimbabwe. (C) Land area (square kilometers) for wildlife conservation in Zimbabwe.

associated diversity might be marketed, although it is not clear whether these mechanisms can be applied generally. One is the sale of "nontimber forest products," which are commercially valuable forest products that are not produced by tree cutting. They are therefore not timber, and their harvest and sale are compatible with conservation of the forest. They include various tropical fruits, vines that can be used as ropes, rattan (which grows among the trees), resins such as latex, and plants used as medicine by local populations. This last use is important. Approximately 4 billion people have no access to Western-style medicine and in times of sickness depend on plant extracts; this is known as ayurvedic medicine. This market is important in human terms, and although the amounts of money involved are small by comparison with those in Western medicine, they could still be enough to provide a significant incentive. Recent estimates suggest that in some tropical forests the collection and sale of nontimber forest products could yield as much as $60–140 per hectare per year, although many other studies have suggested significantly lower numbers (Grimes *et al.*, 1994; Houghton and Mendelsohn, 1996; Peters *et al.*, 1989). These are significant amounts. Given that they are sustainable, i.e., can be earned year after year, they are probably sufficient to justify conservation of forests even in the face of pressure for logging. However, currently there are a limited number of studies on this form of forest use so that the generalizability of these numbers is not clear, and whether the implied levels of harvesting are truly sustainable is also unclear.

The different mechanisms for generating income from natural ecosystems and the biodiversity that powers them are not mutually exclusive: The same land area could earn income by all mechanisms. A forest could obtain returns from carbon sequestration, bio-prospecting, nontimber products, managing a watershed, and ecotourism. In fact, the region of the Mata Atlantica (Brazilian coastal rain forest) inland from Rio de Janeiro is in a position to do exactly this. It manages the watershed for Rio in much the way that the Catskills region does for New York. It also manages the stream flow of the river Rio Paraibo do Sul, which provides most of the electric power for Rio via hydropower. These two services make it truly a major utility for Rio, with great economic value. Additionally, it supports a wide range of endemic species, sequesters carbon, and acts as a magnet for tourists. Currently, the region obtains a financial return only on one of these activities—ecotourism. In a case such as this, it is clear that the economic incentives for conservation could be immense, if we were only to do effectively what we already know how to do.[3]

Conclusions

What can we conclude about biodiversity, its economic importance, and the prospects for generating income based on this? Clearly, its importance is great, in the sense that it has been and remains a key contributor to human well-being. We understand enough of this contribution for its magnitude to

[3]All of these measures are consistent with the United Nations Convention on Biological Diversity, to which many nations (although not the United States) have subscribed. This states in article 11 that "Each Contracting Party shall, as far as possible and appropriate, adopt economically sound measures that act as incentives for the conservation and sustainable use of biological diversity".

be clear, although there are probably many aspects that are still unknown to us. Economically, the following is a key question: Can this importance be reflected in a commensurate income yielded by the conservation and use of biodiversity and of goods and services based on it?

In essence, much of what biodiversity contributes is, or has traditionally been, a public good. Its contribution to the functioning of natural ecosystems is nonexcludable and nonrival, as is its insurance value. Also, knowledge is the quintessential public good. Despite the overwhelming publicness of much that biodiversity offers, there is some prospect of commercializing a limited part. Some of the services provided by natural ecosystems can be privatized and sold, generating a return to the conservation of the biodiversity that supports them. In the case of knowledge, the prospect of commercialization rests on the ability to establish intellectual property rights that will effectively privatize some of the public good aspects of its insurance and knowledge functions. Probably the most fundamental of these functions, like basic scientific research and development, will never be privatized and will thrive only with financial support via other mechanisms.

In summary, biodiversity is important economically. There are some market-based approaches to obtaining returns from biodiversity that can provide incentives for its conservation. If fully implemented, they would have a profound and positive impact. However, they would not solve the problem of biodiversity conservation. They would need to be supplemented by non-market measures such as the Endangered Species Act (ESA), the Convention on International Trade in Endangered Species (CITES), the Marine Mammals Protection Act, the Fishery Conservation and Management Act, and others. I am not suggesting that these are good examples of nonmarket policies: Indeed, in several cases they are not. The point is that they illustrate the type of option available as an alternative to the market. Of course, except for CITES, the previously listed measures are all domestic U.S. measures. To be really effective, we would need global equivalents—a Global Endangered Species Act, for example. Given the political complexity of the domestic ESA, this prospect should illustrate clearly the desirability of market-based policies whenever possible.

It is important to note that even regulatory approaches such as the ESA can give rise to market-based incentives if coupled with provisions for measures such as mitigation banking. In the case of the ESA, several states have modified this in a way that has proven highly effective and has introduced economic incentives where none existed in the initial formulation of the act. A good illustration is an agreement reached between the International Paper Corporation and the U.S. Fish and Wildlife Service concerning the red cockaded woodpecker. This bird is endangered and nests in forests owned by International Paper. The U.S. Fish and Wildlife Service and International Paper reached the following arrangement: A target number of breeding woodpecker pairs was agreed on, and provided that this number is attained or exceeded, then International Paper will be regarded as complying with the ESA, whatever modifications it might make in the habitat under its control. Furthermore, the agreement also provided that any surplus of breeding pairs over this number could be "banked." This means that it could be used by the company to offset ESA requirements with respect to red cockaded woodpeckers elsewhere, or title to the surplus could be sold to other landowners and used by them to gain some measure of exemption (Jorling, 2000; Noss *et al.*, 1997). The important point is that the costs of compliance with the ESA have been reduced by this agreement, without reducing its effectiveness. Indeed, there are additional benefits: Because the production of nesting pairs over a target level is saleable, International Paper now actually has an economic incentive to encourage the endangered species, something it never had with a strict interpretation of the ESA. Recent reports suggest that International Paper may be able to sell banked breeding pairs for as much as $100,000 per pair (W. Coleman, personal communication). Similar mitigation banking systems have been put in place for wetlands, which have to be conserved under the Clean Water Act and which also provide important habitat for endangered or threatened species.

In the cases of the red cockaded woodpecker and wetlands, regulation has produced a market. The market occurs as a result of efforts to meet the regulations at minimum costs and has the positive effect of providing stronger incentives than the original regulation for the conservation of the endangered species or habitat. In principle, we do not need to go through the process of regulation to set up a market. The most obvious move from an economic perspective is instead to create a market in situations in which this seems needed by using the state as a buyer. In other words, in the case of the red cockaded woodpecker, the government could just pay landowners on whose land the woodpeckers breed. This immediately establishes an incentive for conservation without the political and other costs of regulation. Systems of this type are currently being used on a trial basis in parts of the United Kingdom. In regions in which hill farming is economically marginal, and has been subsidized by the European Community's Common Agricultural Policy, the farming subsidies have been replaced by subsidies for preserving the ecological integrity of the countryside.[4] Specifically, farmers are paid to grow and maintain hedgerows, ponds, wetlands, wildflowers, and coppices that provide habitat for birds and mammals. The *Financial Times*, the European equivalent of the *Wall Street Journal*, recently commented in an editorial that in many cases it would make more economic sense to pay farmers to provide "ecological sustenance" rather than to grow food that is clearly surplus to the region's needs and cannot compete on world markets ("Too much food," 1998). This solution is less attractive than establishing a market in which consumers buy directly, as in the cases of ecotourism or watersheds, because the money spent by the government has to be raised as tax revenue from individuals or corporations. This can produce a loss of economic efficiency elsewhere in the economy. However, if the services provided by biodiversity are public goods, then there may be little alternative to having the public sector act as a buyer on behalf of society as a whole.

See also: Deforestation and Land Clearing. Ecosystem Services. Keystone Species. Land-Use Issues. Market Economy and

[4]This system, known as Tyr Cwmen, is in operation in parts of South Wales.

Biodiversity. Plant Sources of Drugs and Chemicals. The Value of Biodiversity

References

Aber JJD and Melillo JM (1991) *Terrestrial Ecosystems*. Philadelphia: Saunders.

Antonovics J, Bradshaw AD, and Turner RG (1971) Heavy metal tolerance in plants. *Advances in Ecological Research* 7: 1–85.

Blockstein D (1988) Lyme disease and the passenger pigeon. *Science* 279: 1831.

Bond I (1993, April) The economics of wildlife and land use in Zimbabwe: An examination of current knowledge and issues. World Wildlife Fund (WWF) Multispecies Animal Production Systems Project, Project Paper No. 36. WWF Program Office, P.O. Box CY 1409, Causeway, Harare, Zimbabwe: WWF.

Carte BK (1996) The biomedical potential of marine natural products. *BioScience* 46(4): 271–286.

Chapin FS III, Walker BH, Hobbs RJ, *et al.* (1997) Biotic control over the functioning of ecosystems. *Science* 277: 500–504.

Chichilnisky G and Heal GM (1998) Economic returns from the biosphere. *Nature* 391: 629–630.

Committee on Agriculture Science and Technology (CAST) (1999, February) Benefits of biodiversity, Task Force Report No. 133. CAST, 4420 West Way, Ames, IA 50015-3447.

Cumming DHM (1990a, June) Developments in game ranching and wildlife utilization in East and Southern Africa. World Wildlife Fund (WWF) Multispecies Animal Production Systems Project, Project Paper No. 13. Zimbabwe: WWF Program Office.

Cumming DHM (1990b, June) Wildlife and the market place: A view from Southern Africa. World Wildlife Fund (WWF) Multispecies Animal Production Systems Project, Project Paper No. 12. Zimbabwe: WWF Program Office.

Cumming DHM and Bond I (1991, July) Animal production in Southern Africa: Present practices and opportunities for peasant farmers in arid lands. World Wildlife Fund (WWF) Multispecies Animal Production Systems Project, Project Paper No. 22. Zimbabwe: WWF Program Office.

Daily GC (1997) *Nature's Services: Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.

Ehrlich PR (1988) The loss of diversity: Causes and consequences. In: Wilson EO (ed.) *Biodiversity*. Washington, DC: National Academy of Sciences/Smithsonian Institution.

Freese CH (1999) *Wild Species as Commodities: Managing Markets and Ecosystems for Sustainability*. Washington, DC: Island Press.

Grime JP (1977) Biodiversity and ecosystem function: The debate deepens. *Science* 277: 1260.

Grimes A, Loomis S, Jahnige P, *et al.* (1994) Valuing the rain forest: The economic value of nontimber forest products in Ecuador. *Ambio.* 23(7): 405–410.

Heal GM (2000a) Sustainability and markets. In: Stewart R (ed.) *Topics in Environmental Law*. Cambridge, UK: Cambridge Univ. Press.

Heal GM (2000b) *Earthkeeping*. Washington, DC: Island Press.

Hemingway E (1940) *For Whom the Bell Tolls*. New York: Scribner.

Honey M (1999) *Ecotourism and Sustainable Development: Who Owns Paradise?* Washington, DC: Island Press.

Hooper DU and Vitousek PM (1997) The effect of plant composition and diversity on ecosystem processes. *Science* 277: 1302–1305.

Houghton KT and Mendelsohn RO (1996) An economic analysis of multiple-use forestry in Nepal. *Ambio.* 25(3): 156–159.

Jorling TC (2000) Incentive-based management strategies: A "common-sense" approach to conservation of biological diversity. *Sci. Total Environ.* 240: 1999.

McGrady-Steed J, Harns PM, and Morin PJ (1997) Biodiversity regulates ecosystem predictability. *Nature* 390: 162–165.

Monsanto concession on engineered corn (1999) *Nature* 397: 98.

Myers N (1997) Biodiversity's genetic library. In: Daily GC (ed.) *Nature's Services: Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.

Naeem S and Li S (1997) Biodiversity enhances ecosystem reliability. *Nature* 390: 505–509.

National Research Council (NRC) (1999) A framework for managing biodiversity. Report of the Committee on the Economic and Non-Economic Value of Biodiversity Washington, DC: NRC.

Noss RF, O'Connell MA, and Murphy DD (1997) *The Science of Conservation Planning: Habitat Conservation under the Endangered Species Act*. Washington, DC: Island Press.

Perfecto I, Rice RA, Greenberg R, and van der Voort ME (1996) Shade coffee: A disappearing refuge for biodiversity. *BioScience* 46(8): 598–608.

Peters CM, Gentry AH, and Mendelsohn RO (1989) Valuation of and Amazonian rainforest. *Nature* 339: 655–656.

Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *BioScience* 46(8): 609–620.

Rausser GC and Small AA (1998, March) Valuing research leads: Bioprospecting and the conservation of genetic resources Paper presented at the Conference on Managing Human-Dominated Ecosystems at the Missouri Botanical Gardens, St. Louis, MO.

Raven PH and Williams T (1999) *Nature and Human Society: The Quest for a Sustainable World*. Washington, DC: National Academy of Sciences.

Read D (1998) Biodiversity: Plants on the web. *Nature* 396. (News and Views).

Reid WV (1998, March) A business plan for ecosystem services: Extending the New York City watershed model to other geographic regions and other ecosystem services. Paper presented by the World Resources Institute at the Conference on Managing Human-Dominated Ecosystems at the Missouri Botanical Gardens, St. Louis, MO.

Reid WV, Laird SA, Meyer CA, *et al.* (1993) *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development*. Washington, DC: World Resources Institute.

Smith A (1977) *An Enquiry into the Nature and Causes of the Wealth of Nations*. Chicago: Univ. of Chicago Press.

Stierle, *et al.* (1993) Taxol and taxane production by *Taxomyces andreanae*, and endophytic fungus of Pacific Yew. *Science* 260: 214–216.

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, and Siemann E (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277: 1300–1305.

Tilman GD and Downing JA (1994) Biodiversity and stability in grasslands. *Nature* 367: 363–365.

Tilman GD, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379.

Too much food (1998, November 17). *Financial Times*.

U.S. Congress Office of Technology Assessment (1987) Washington, DC: U.S. Congress.

van der Heijden MGA, Klironomos JN, Ursic M, *et al.* (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 396: 69–71.

When rhetoric hits reality in debate on bioprospecting (1998) *Nature* 392: 535–540.

Wilson EO (ed.) (1988) *Biodiversity*. Washington, DC: National Academy of Sciences/Smithsonian Institution.

Biodiversity, Definition of

Ian R Swingland, University of Kent, Canterbury, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity/biological diversity Species, genetic, and ecosystem diversity in an area, sometimes including associated abiotic components such as landscape features, drainage systems, and climate.

Diversity indices Measures that describe the different components of biodiversity, such as species richness (alpha diversity), beta and gamma diversity, endemicity, and higher taxon richness.

Ecosystem diversity Diversity of habitats, ecosystems, and the accompanying ecological processes that maintain them.

Endemicity State of a species or other taxon being restricted to a given area, such as a specific habitat, region, or continent.

Flagship species Charismatic or well-known species that is associated with a given habitat or ecosystem and that may increase awareness of the need for conservation action.

Genetic diversity Genetic variety found within or among species; this diversity allows the population or species to adapt and evolve in response to changing environments and natural selection pressures.

Keystone species Species that has a disproportionately greater effect on the ecological processes of an ecosystem, and whose loss would result in significantly greater consequences for other species and biotic interactions.

Organismal (species) diversity Number and relative abundance of all species living in a given area.

Species richness Absolute number of species living in a given area (also called alpha diversity), giving equal weight to all resident species.

Use values Values that are obtained by using a natural resource, such as timber, fuelwood, water, and landscapes. These include direct, indirect, option, and nonuse values.

The word biodiversity is a modern contraction of the term biological diversity. Diversity refers to the range of variation or variety or differences among some set of attributes; biological diversity thus refers to variety within the living world or among and between living organisms.

What is Biodiversity?

The term "biodiversity" was first used in its long version (biological diversity) by Lovejoy (1980) and is most commonly used to describe the number of species. Recognizing that conventional methods of determining, and separating, species were inadequate, others elaborated the definition by including the variety and variability of living organisms.

These reduced and simple definitions, which embrace many different parameters, have been much elaborated and debated in the last three decades (*see* Section on Definition of Biodiversity); on this definition hangs the outcome of important scientific considerations, international agreements, conventions, conservation initiatives, political debates, and socioeconomic issues. Indeed, while the word "biodiversity" has become synonymous with life on earth, the term is commonly used in the fields of politics and environmental technology in addition to various scientific disciplines (Ghikrov, 1996). The US Strategy Conference on Biological Diversity (1981) and the National Forum on Biodiversity (1986) in Washington, DC, were the critical debates in crafting a definition, and it was the proceedings from the latter, edited by E.O. Wilson, that "launched the word 'biodiversity' into general use" (Harper and Hawksworth, 1994).

In measuring biodiversity, it is necessary to deconstruct some of the separate elements of which biodiversity is composed. It has become widespread practice to define biodiversity in terms of genes, species, and ecosystems, for example, "the abundance, variety, and genetic constitution of native animals and plants" (Dodson *et al.*, 1998). Biodiversity also encompasses all five living kingdoms, including fungi. However, biodiversity does not have a universally agreed on definition and it is often redefined on each occasion according to the context and purpose of the author.

Definition of Biodiversity

"Biodiversity" is a relatively new compound word, but biological diversity (when referring to the number of species) is not. Over the last decade its definition has taken a more reductionist turn. Possibly the simplest definition for biodiversity, lacking in specificity or context, is merely the number of species. Yet many have argued that biodiversity does not equate to the number of species in an area. The term for this measure is species richness (Fiedler and Jain, 1992), which is only one component of biodiversity. Biodiversity is also more than species diversity (simply called diversity by some authors), which has been defined as the number of species in an area and their relative abundance (Pielou, 1977).

DeLong (1996) offered a more comprehensive definition:

Biodiversity is an attribute of an area and specifically refers to the variety within and among living organisms, assemblages of living organisms, biotic communities, and biotic processes, whether naturally occurring or modified by humans. Biodiversity can be measured in terms of genetic diversity and the identity and number of different types of species, assemblages of species, biotic communities, and biotic processes, and the amount (e.g., abundance, biomass, cover, rate) and structure

of each. It can be observed and measured at any spatial scale ranging from microsites and habitat patches to the entire biosphere.

This definition allows for modification according to the context in which it is used.

Various authors have proposed specific and detailed elaborations of this definition. Gaston and Spicer (1998) proposed a three-fold definition of "biodiversity" – ecological diversity, genetic diversity, and organismal diversity – while others conjoined the genetic and organismal components, leaving genetic diversity and ecological diversity as the principal components. These latter two elements can be linked to the two major "practical" value systems of direct use/genetics and indirect use/ecological described by Gaston and Spicer (1998). Other workers have emphasized a hierarchical approach or hierarchies of life systems.

In contrast, some argue that biodiversity, according to the definition of biological, does not include the diversity of abiotic components and processes, and that it is inaccurate to identify ecological processes, ecosystems, ecological complexes, and landscapes as components of biodiversity. The term ecological, as used in the sense of ecological system (ecosystem), encompasses both biotic and abiotic components and processes. Therefore, ecological diversity is a more appropriate term for definitions that include the diversity of ecological processes and ecosystems. However, ecological processes, it has been argued, should be included in the definition of biodiversity, the reasoning being that "although ecological processes are as much abiotic as biotic, they are crucial to maintaining biodiversity." Similarly, a US Bureau of Land Management advisory group included ecological processes in their definition of biodiversity in response to criticism that the Office of Technology Assessment's (1987) definition did not consider ecosystem form and function. Other writers point out that even though ecological processes are often cited as being crucial to maintaining biodiversity (Reid and Miller, 1989; Noss and Cooperrider, 1994; Samson and Knopf, 1994), this does not warrant the inclusion of ecological processes into the meaning of biodiversity. For example, Reid and Miller (1989) and Agarwal (1992) distinguished between biodiversity and the processes and ecological diversity that maintain it.

Nevertheless, the jargon word "biodiversity" is, by its very origin, fundamentally indefinable, being a populist word invented for convenience. Its invention has had beneficial effects by fueling research projects, mainly in ecology and systematics, and scientists have been drawn into contributing to the debate by the need to show that biodiversity is useful to humans and necessary for the proper functioning of ecosystems. Conservation (i.e., management) of biodiversity is axiomatic to these two concerns and lies behind the scientific need to define the term within whatever context is appropriate, since no general definition will be suitable when applied across a range of situations.

Biodiversity conservation requires the management of natural resources, and this in turn requires the measurement of these resources. Biodiversity measurement implies the need for some quantitative value that can be ascribed to the various measurements so these values can be compared. Among the first scientists to measure diversity were Fisher, Corbet, and Williams (1943), who approximated the frequency distribution of the species represented by 1,2,3,4… (and so on) individuals by the logarithmic series αx, $\alpha x2/2$, $\alpha x3/3$, $\alpha x4/4$…, where the constant α has been found to be a measure of species diversity. Species diversity is low when the number of species is growing slowly with respect to the increase in number of individuals, and it is high when the number of species is growing quickly.

If the need to quantify biodiversity drives the fundamental meaning of biodiversity, the definition may be limited to that which can be readily measured given current understanding and technologies. Such a definition of biodiversity could change over time as ideas, technology, and resources for measuring diversity change. DeLong (1996) suggested that an operational "clause" should be added to the definition of biodiversity, namely, that "biodiversity is…as measured in terms of…." This approach provides a link to management while distinguishing between what biodiversity is (a state or attribute) and how it is measured. It also allows the operational clause to be adjusted over time without changing the fundamental meaning of the term. A definition of biodiversity should portray the full scope of what the term means, not just what can be measured and managed. In contrast, monitoring or management objectives must be attainable to be effective. Recognizing the distinction between a definition and management objectives should reduce the confusion between the meaning of biodiversity and the objectives for achieving biodiversity goals.

Biodiversity is a broad totality and often embraces elements beyond species diversity or numbers. For example, a major debate in biological sciences over many decades has been that of pattern versus process, especially in systematics and evolutionary studies. Molecular biology and systematics have enabled ecologists to see that inferred history is important in framing appropriate questions, and this understanding has precipitated a real integration of these twin hierarchies – pattern (e.g., diversity) and process (e.g., evolution). Fundamental divisions remain, such as "straight" parsimony (i.e., pattern) versus maximum likelihood (i.e., process) in the phylogenetic interpretation of sequence data.

It is apparent that the term biodiversity still lacks consistent meaning within the field of natural resource management. Michael Soulé found it shocking that "we are still trying to define biological diversity after all of the efforts of the Office of Technology Assessment and E.O. Wilson's book, Biodiversity" (Hudson, 1991). It is still defined in different ways by different people; some characterize biodiversity as being a widely used term "having no unified definition" and others emphasize or limit the meaning of biodiversity to that of native biodiversity. Some writers have included human alterations of biological communities in the scope of biodiversity (Bryant and Barber, 1994). Angermeier (1994) argued that "the absence of a 'native' criterion within the definition (of biodiversity) severely compromises biodiversity's utility as a meaningful biological concept," reasoning that native biodiversity is more valuable than artificial diversity and should be the primary focus of conservation efforts. The conservation of native biodiversity appears to be the theme of biodiversity conservation texts (Wilson and Peter, 1988; Hunter, 1996). Conversely, others argued that an important component of biodiversity is

maintained by traditional farming techniques. In the context of conserving biodiversity, Reid and Miller (1989) and Bryant and Barber (1994) discussed the importance of genetic diversity within species of cultivated plants. Biodiversity within agricultural plants is important for pest management in agroecosystems and sustainable agriculture.

An accepted fundamental definition of biodiversity is needed for conservation planning, as are effective communication and cooperation within and among different countries, governments, agencies, disciplines, organizations, and private landowners. Cooperation among these entities has been identified as being necessary for the conservation of biodiversity (Babbitt, 1994). Knopf (1992) asserted that the definitions of biodiversity are "as diverse as the biological resource." Definitions of biodiversity range in scope from "the number of different species occurring in some location" to "all of the diversity and variability in nature" and "the variety of life and its processes." A more comprehensive definition is "the variety of living organisms, the genetic differences among them, the communities and ecosystems in which they occur, and the ecological and evolutionary processes that keep them functioning, yet ever changing and adapting" (Noss and Cooperrider, 1994).

This plethora of terms and definitions is one of the major stumbling blocks to reaching agreement in problem solving and decision making. If entities in a planning process view biodiversity in fundamentally different ways, agreement on management objectives and strategies for biodiversity conservation will be impaired. (Swingland, 1999).

The differences between these conceptual perspectives on the meaning of biodiversity, and the associated semantic problems, are not trivial. Management intended to maintain one facet of biodiversity will not necessarily maintain another. For example, a timber extraction program that is designed to conserve biodiversity in the sense of site species richness may well reduce biodiversity measured as genetic variation within the tree species harvested. Clearly, the maintenance of different facets of biodiversity will require different management strategies and resources, and will meet different human needs.

Even if complete knowledge of particular areas could be assumed, and standard definitions of diversity are derived, the ranking of such areas in terms of their importance with respect to biological diversity remains problematic. Much depends on the scale that is being used. Thus, the question of what contribution a given area makes to global biological diversity is very different from the question of what contribution it makes to local, national, or regional biological diversity. This is because, even using a relatively simplified measure, any given area contributes to biological diversity in at least three different ways – through its richness in numbers of species, through the endemism (or geographical uniqueness) of these species (e.g., Mittermeier *et al.*, 1992), and on the basis of degree of threat. The relative importance of these three factors will inevitably change at different geographical scales, and sites of high regional importance may have little significance at a global level. None of these factors includes any explicit assessment of genetic diversity.

Although the word biodiversity has already gained wide currency in the absence of a clear and unique meaning, greater precision will be required of its users if policy and programs are to be more effectively defined in the future.

Genetic Diversity

Genetic diversity is reliant on the heritable variation within and between populations of organisms. New genetic variation arises in individuals by gene and chromosome mutations, and in organisms with sexual reproduction it can be spread through the population by recombination. It has been estimated that in humans and fruit flies alike, the number of possible combinations of different forms of each gene sequence exceeds the number of atoms in the universe. Other kinds of genetic diversity can be identified at all levels of organization, including the amount of DNA per cell and chromosome structure and number. Selection acts on this pool of genetic variation present within an interbreeding population. Differential survival results in changes of the frequency of genes within this pool, and this is equivalent to population evolution. Genetic variation enables both natural evolutionary change and artificial selective breeding to occur (Thomas, 1992).

Only a small fraction ($<1\%$) of the genetic material of higher organisms is outwardly expressed in the form and function of the organism; the purpose of the remaining DNA and the significance of any variation within it are unclear (Thomas, 1992). Each of the estimated 109 different genes distributed across the world's biota does not make an identical contribution to overall genetic diversity. In particular, those genes that control fundamental biochemical processes are strongly conserved across different taxa and generally show little variation, although such variation that does exist may exert a strong effect on the viability of the organism; the converse is true of other genes. A large amount of molecular variation in the mammalian immune system, for example, is possible on the basis of a small number of inherited genes (Thomas, 1992).

Species Diversity

Historically, species are the fundamental descriptive units of the living world and this is why biodiversity is very commonly, and incorrectly, used as a synonym of species diversity, in particular of "species richness," which is the number of species in a site or habitat. Discussion of global biodiversity is typically presented in terms of global numbers of species in different taxonomic groups. An estimated 1.7 million species have been described to date; estimates for the total number of species existing on earth at present vary from 5 million to nearly 100 million. A conservative working estimate suggests there might be around 12.5 million.

When considering species numbers alone, life on earth appears to consist mostly of insects and microorganisms. The species level is generally regarded as the most natural one at which to consider whole-organism diversity. While species are also the primary focus of evolutionary mechanisms, and the origination and extinction of species are the principal agents in governing biological diversity, species cannot be recognized

and enumerated by systematists with total precision. The concept of what a species is differs considerably among groups of organisms. It is for this reason, among others, that species diversity alone is not a satisfactory basis on which to define biodiversity.

Another reason why a straightforward count of the number of species provides only a partial indication of biological diversity concerns the concept of degree or extent of variation that is implicit within the term biodiversity. By definition, organisms that differ widely from each other in some respect contribute more to overall diversity than those that are very similar. The greater the interspecific differences (e.g., by an isolated position within the taxonomic hierarchy), then the greater contribution to any overall measure of global biological diversity. Thus, the two species of Tuatara (genus Sphenodon) in New Zealand, which are the only extant members of the reptile order Rhynchocephalia, are more important in this sense than members of some highly species-rich family of lizards. A site with many different higher taxa present can be said to possess more taxonomic diversity than another site with fewer higher taxa but many more species. Marine habitats frequently have more different phyla but fewer species than terrestrial habitats, that is, higher taxonomic diversity but lower species diversity. By this measure, the Bunaken reef off the north coast of Sulawesi has the highest biodiversity on earth. Current work is attempting to incorporate quantification of the evolutionary uniqueness of species into species-based measures of biodiversity.

The ecological importance of a species can have a direct effect on community structure, and thus on overall biological diversity. For example, a species of tropical rain forest tree that supports an endemic invertebrate fauna of a hundred species makes a greater contribution to the maintenance of global biological diversity than does a European alpine plant that may have no other species wholly dependent on it.

Ecosystem Diversity

While it is possible to define what is in principle meant by genetic and species diversity, it is difficult to make a quantitative assessment of diversity at the ecosystem, habitat, or community level. There is no unique definition or classification of ecosystems at the global level, and it is difficult in practice to assess ecosystem diversity other than on a local or regional basis, and then only largely in terms of vegetation. Ecosystems are further divorced from genes and species in that they explicitly include abiotic components, being partly determined by soil/parent material and climate.

To get around this difficulty, ecosystem diversity is often evaluated through measures of the diversity of the component species. This may involve assessment of the relative abundance of different species as well as consideration of the types of species. The more that species are equally abundant, then the more diverse that area or habitat. Weight is given to the numbers of species in different size classes, at different trophic levels, or in different taxonomic groups. Thus a hypothetical ecosystem consisting only of several plant species would be less diverse than one with the same number of species but that included animal herbivores and predators. Because different weightings can be given to these different factors when estimating the diversity of particular areas, there is no one authoritative index for measuring ecosystem diversity. This obviously has important implications for the conservation ranking of different areas. In examining beta diversity (i.e., the change in species composition between areas), the only reliable predictor of community similarity is to compare the species composition of the site immediately adjacent.

Biodiversity: Meaning and Measurement

Species Diversity

A.S. Corbet, on analyzing a large collection of butterflies from Malaya, remarked on the decrease in number of new species with an increasing number of individuals. He thought that the resulting distribution could be described by a hyperbola, but R.A. Fisher, to whom Corbet sent his results, suggested that a negative binomial distribution would be much more appropriate (Williams, 1964). As mentioned earlier, Fisher, Corbet, and Williams (1943) approximated the frequency distribution of the species represented by 1,2,3,4… (and so on) individuals by the logarithmic series αx, $\alpha x2/2$, $\alpha x3/3$, $\alpha x4/4$…, where the constant α is a measure of species diversity. Species diversity is low when the number of species rises slowly with an increase in the number of individuals, and diversity is high when the number of species rises quickly.

Species diversity measurement was thus clearly formulated more than 50 years ago and a particular index was proposed. Fisher *et al*. attempted to find some general "rule" or "law" according to which the numerical abundances of different species were related to each other. In many communities, the number of species with given abundance could be approximated by the log-normal distribution. If species are classified in accordance with their abundance in logarithmically increasing classes – so-called "octaves" (i.e., the first octave contains 1–2 individuals, the second contains 2–4 individuals, the third has 4–8, the fourth has 8–16, and so on) – then the number of species per "octave" shows a truncated normal distribution. If a sample contains a high number of species and individuals, we can usually obtain a log-normal distribution, and it is obviously more tractable than the logarithmic series.

MacArthur (1957) went further by proposing an interesting model that assumed that boundaries between niches in resource–niche hypervolume are set at random, whereas the relative abundances of species are proportional to these sections of hypervolume. This model became widely known as the "broken-stick" or MacArthur's model. The distribution of abundance prescribed by MacArthur's model is much "flatter" (i.e., the contrast between given species and the next in the sequence is less) than in the case of a logarithmic series (Ghikrov, 1996).

It has become clear that there is no universal type of distribution of relative abundance that corresponds to all real communities, though such distributions change in the course of succession according to a particular pattern. The dominance of a few of the most abundant species is more pronounced at the early stages of succession, while later the species of

intermediate abundance become more significant (Whittaker, 1972). A comprehensive understanding of the underlying mechanisms that result in a given pattern of species abundance still eludes scientists.

Another line of species diversity studies was connected with the use of special indices proposed to measure diversity without reference to some hypothetical distribution of relative abundance. A great variety of indices were proposed that assess the number of species and the proportions in abundance of different species. Among others, there was the very popular index that is based on Shannon's formula derived from information theory: $H = \sum pi \log pi$ where pi is the proportion of the total number of individuals that belong to the ith species.

In a seminal work on the measurement of diversity, Whittaker (1972) introduced the concepts of alpha, beta, and gamma diversity. The measurements just described, giving diversity values for single sites, are examples of alpha diversity. The beta and gamma diversity concepts relate to changes in diversity between sites at local (beta) and geographical (gamma) scales. An essential part of these relational concepts is the idea of species turnover – the degree to which species replace other species at different sites. For use in assessing the relative value of multiple sites for the conservation of biodiversity, the idea of species turnover is translated into the principle of complementarity (*see* Section on Loss of Biodiversity and Causes), which can be implemented in combination with a taxonomic diversity index.

Taxonomic Diversity

Biodiversity measurements that measure genetic difference directly, or indirectly through use of the taxonomic (cladistic) hierarchy (Williams *et al.*, 1991), are currently being used. The indirect taxonomic approach is more practical because we already have a "rule of thumb" taxonomic hierarchy (which is being steadily improved through the application of cladistic analysis, notably to molecular data), whereas reliable estimates of overall genetic differences between taxa are virtually nonexistent (abridged from Vane-Wright, 1992).

Based on the shared and unshared nodes between taxa (equivalent to position in the taxonomic hierarchy), a number of taxonomic diversity indices have now been developed. Of these, the most distinct are root weight, higher taxon richness, and taxonomic dispersion. The first places highest individual value on taxa that separate closest to the root of the cladogram and comprise only one or relatively few species; in effect this gives high weighting to relict groups (Vane-Wright, 1996). Higher taxon richness favors taxa according to their rank and number of included species. Dispersion, the most complex of the measures proposed so far (Williams *et al.*, 1991), endeavors to select an even spread of taxa across the hierarchy, sampling a mixture of high, low, and intermediate ranking groups.

For a given group these measures, together with simple species richness if desired, can be used to compare the biotic diversity of any number of sites. The measures can also be expressed as percentages. Thus a site with viable populations of all species in a group would have a diversity score of 100%, whereas a site without any species of the group in question would score zero. In reality, of course, most sites have only a selection of species, and so receive various intermediate scores. Such assessments allow us to compare all sites with each other, and rank them individually from highest to lowest diversity (Vane-Wright, 1996). However, if we then take some conservation action (such as conserving a particular site), the same measures are unlikely to be directly comparable for making a second decision (such as choosing a second conservation site). This is because, in most real situations at least, there will be considerable overlap in the presence of species at particular sites.

Community Diversity

Early ecologists did not confine themselves to measuring species diversity. They also tried to understand the relationship of diversity with other features of the community (e.g., Williams, 1964; Whittaker, 1972). The dependence of species diversity on the structural complexity of the environment was demonstrated (MacArthur and MacArthur, 1961), as was the role of predation (Addicott, 1974) and periodical disturbance (Sousa, 1979) in determining a given level of diversity. The relationship between the species diversity and standing crop of a community was also shown (Ghilarov and Timonin, 1972).

Margalef (1957) was the first to use the Shannon index (though expressed in a different form). He proposed to evaluate the level of community organization in terms of information theory. Margalef stimulated many ecologists to quantitatively measure the species diversity of different communities or of the same community in different stages of its development. At that time, there was a widespread belief that with a single numerical value, an assessment could be made of some very significant feature of community structure. Many ecologists believed that in measuring species diversity at the community level they were using an approach that was fundamental to an understanding of diversity (Ghikrov, 1996).

Ecologists have measured diversity either by estimating species richness (number of species) in an area, or by using one or more indices combining species richness and relative abundance within an area. Some attempts have also been made to measure change in species richness (species turnover) between areas. These solutions to the problem of measuring biodiversity are limited because species richness takes no account of the differences between species in relation to their place in the natural hierarchy. Moreover, relative abundance is not a fixed property of a species, for it varies widely from time to time and place to place. In many environments most taxa are virtually or even completely unknown.

Conservation biologists, or applied ecologists, have called for a measurement of diversity that is more clearly related to overall genetic difference. An example concerns the problem of differential extinction. In World Conservation Strategy (IUCN/UNEP/WWF, 1980), it is noted that "the size of the potential genetic loss is related to the taxonomic hierarchy because…different positions in this hierarchy reflect greater or lesser degrees of genetic difference…. The current taxonomic hierarchy provides the only convenient rule of thumb for determining the relative size of a potential loss of genetic material."

Synthesis

A model incorporating island biogeographic theory, species abundance, and speciation, and that produces a fundamental biodiversity number (θ) that is closely associated with species richness and abundance in an equilibrium metapopulation, has been proposed in Hubbell's unified theory (1997). This model assumes zerosum community dynamics or a saturated, totally stochastic local community, which limits its application, but it advances the study of species richness and relative abundance if others can extend its usefulness to the nonequilibrium systems that characterize the real world.

Biodiversity: Changes in Time and Space

Changes Over Time

The fossil record is very incomplete, which emphasizes the marked variation between higher taxa and between species in different ecosystems in the extent to which individuals are susceptible to preservation and subsequent discovery. Chance discovery has played a large part in compiling the known fossil record, and interpretation by paleontologists of the available material is beset by differences of opinion. Thus, the record is relatively good for shallow-water, hard-bodied marine invertebrates, but poor for most other groups, such as plants in moist tropical uplands.

Two relevant points appear to be well substantiated. First, taxonomic diversity, as measured by the number of recognized phyla of organisms, was greater in Cambrian times than in any later period. Second, it appears that species diversity and the number of families have undergone a net increase between the Cambrian and Pleistocene epochs, although interrupted by isolated phases of mass extinction (few of which are reflected in the fossil record of plants).

Changes in Space

Species diversity in natural habitats is high in warm areas and decreases with increasing latitude and altitude; additionally, terrestrial diversity is usually higher in areas of high rainfall and lower in drier areas. The richest areas are tropical moist forest and, if current estimates of the number of microfaunal species (mainly insects) of tropical moist forests are credible, then these areas, which cover perhaps 7% of the world's surface area, may well contain over 90% of all species. If the diversity of larger organisms only is considered, then coral reefs such as Bunaken (see earlier) and, for plants at least, areas with a Mediterranean climate in South Africa and Western Australia may be as diverse. Gross genetic diversity and ecosystem diversity will tend to be positively correlated with species diversity.

What are not fully understood are the reasons for the large-scale geographic variation in species diversity, and in particular for the very high species diversity of tropical moist forests. The origin of diversity through the evolution of species and the maintenance of this diversity both need more study before they are better understood. This will require consideration of the present and historic (in a geological or evolutionary sense) conditions prevailing in particular areas, principally climatic but also edaphic and topographic. Climatically benign conditions (warmth, moisture, and relative seasonality) over long periods of time appear to be particularly important.

Climax ecosystems will be more diverse than areas at earlier successional stages, but an area with a mosaic of systems at different successional stages will probably be more diverse than the same area at climax provided that each system occupies a sufficiently large area of its own. In many instances, human activities artificially maintain ecosystems at lower successional stages. In areas that have been under human influence for extended periods, notably in temperate regions, maintenance of existing levels of diversity may involve the maintenance of at least partially man-made landscapes and ecosystems, mixed with adequately sized areas of natural climax ecosystems.

Loss of Biodiversity and Causes

Species extinction is a natural process that occurs without the intervention of humans since, over geological time, all species have a finite span of existence. Extinctions caused directly or indirectly by humans are occurring at a rate that far exceeds any reasonable estimates of background extinction rates, and to the extent that these extinctions are correlated with habitat perturbation, they must be increasing.

Quantifying rates of species extinction is difficult and predicting future rates with precision is impossible. The documentation of definite species extinctions is only realistic under a relatively limited set of circumstances, for example, where a described species is readily visible and has a well-defined range that can be surveyed repeatedly. Unsurprisingly, most documented extinctions are of species that are easy to record and that inhabit sites that can be relatively easily inventoried. The large number of extinct species on oceanic islands is not solely an artifact of recording, because island species are generally more prone to extinction as a result of human actions.

Most global extinction rates are derived from extrapolations of measured and predicted rates of habitat loss, and estimates of species richness in different habitats. These two estimates are interpreted in the light of a principle derived from island biogeography, which states that the size of an area and of its species complement tend to have a predictable relationship. Fewer species are able to persist in a number of small habitat fragments than in the original unfragmented habitat, and this can result in the extinction of species (MacArthur and Wilson, 1967). These estimates involve large degrees of uncertainty, and predictions of current and future extinction rates should be interpreted with considerable caution. The pursuit of increased accuracy in the estimation of global extinction rates is not crucial. It is more important to recognize in general terms the extent to which populations and species that are not monitored are likely to be subject to fragmentation and extinction (Temple, 1986).

Loss of biodiversity in the form of domesticated animal breeds and plant varieties is of little significance in terms of overall global diversity, but genetic erosion in these populations is of particular human concern in so far as it has

implications for food supply and the sustainability of locally adapted agricultural practices. For domesticated populations, the loss of wild relatives of crop or timber plants is of special concern for the same reason. These genetic resources may not only underlie the productivity of local agricultural systems but may also, when incorporated into breeding programs, provide the foundation of traits (disease resistance, nutritional value, hardiness, etc.) that are of global importance in intensive systems and that will assume even greater importance in the context of future climate change. Erosion of diversity in crop gene pools is difficult to demonstrate quantitatively, but can be indirectly assessed in terms of the increasing proportion of world cropland planted to high-yielding, but genetically uniform, varieties. Genetic modification of organisms, varieties, or cultivars for food production, pharmaceuticals, and other products, which has caused concern in some countries but not others, may also contribute to the loss of biodiversity.

Humans exterminate species either directly by hunting, collection, and persecution or indirectly through habitat destruction and modification. Overhunting is perhaps the most obvious direct cause of extinction in animals, but it is undoubtedly far less important than the indirect causes of habitat modification in terms of overall loss of biodiversity. Hunting selectively affects the targeted species, as well as plant and animal species whose populations are subsequently affected either negatively or positively, and so it has important implications for the management of natural resources. Genetic diversity in a hunted population is liable to decrease as a result of the same factors. The genetic diversity represented by populations of crop plants or livestock is also likely to decline as a result of mass production, for the desired economics of scale demand high levels of uniformity.

Sustained human activity will affect the relative abundance of species and in extreme cases may lead to extinction. This may result from the habitat being made unsuitable for the species (e.g., clear-felling of forests or severe pollution of rivers) or through the habitat becoming fragmented (discussed earlier). Fragmentation divides previously contiguous populations of species into small subpopulations. If these are sufficiently small, then chance processes lead to higher probabilities of extinction within a relatively short time. Major changes in natural environments are likely to occur within the next century as a result of changes in global climate and weather patterns. These will cause greatly elevated extinction rates.

Maintaining Biodiversity

In Situ Conservation

The maintenance of biological diversity is the sustainable management of viable populations of species or populations *in situ* or *ex situ*. The maintenance of a significant proportion of the world's biological diversity only appears feasible by maintaining organisms in their wild state and within their existing range. This allows for continuing adaptation of wild populations by natural evolutionary processes and, in principle, for current utilization practices to continue. For such maintenance to succeed, it almost invariably requires enhanced management through the integrated, community-based conservation of protected areas.

Over the last 30 years, conservation biologists have struggled with the concept of the maintenance of biodiversity in highly diverse environments like rain forests. Analytical techniques (neural-net models) that allow us to reconstruct past distributions of forest types present an opportunity to predict past contractions and expansions of forest forms, and the likelihood of refugia surviving climate change. Such extrapolations must be treated with caution, as pollen samples from Brazil (for example) disproved modeling predictions that savanna grasslands should have been extant, when in fact tropical and temperate forests were present. Various authors also opposed the Pleistocene refugia hypothesis (Haffer, 1969) for the Amazon region because some evidence demonstrated the lack of rain forest fragmentation during that era. In the biogeographical zones of the Australian wet tropics, there is a strong correlation between diversity patterns and reputed rain forest refugia in both species and genetic diversity. However, this appears to have been caused by differential extinction rates in differently sized refugia rather than by allopatric speciation in the Pleistocene. Others have emphasized that a greater concentration on the Pliocene or before would be useful, since most tropical species radiations occurred before the Pleistocene.

The local-determination hypothesis of species diversity (Rosenzweig, 1995), which predicts similar species diversity in similar habitats, has also been challenged. In sister taxa of plants, the net diversification was significantly higher in Asia than in North America for genera shared between the two continents. Greater insights into the effects of current ecology on the local diversity of an area may be assisted by considering the relative ages of clades, which could establish species proliferation rates between regions, thus advancing the local versus regional diversity debate (Ricklefs and Schluter, 1993). They also tested the taxon cycle theory (Wilson, 1961) using phylogenies of bird species and showed that older species' lineages had more restricted ranges, smaller habitat breadth, and more fragmented distributions, and were closer to extinction than younger species.

In efforts to conserve biodiversity, preserving genetic dissimilarity is often a higher priority than maintaining genes of considerable similarity. Recent work shows that genetic divergence in mammals increases from the headwaters to the mouth as a river gets broader and thus becomes a greater barrier to populations on opposite banks; this effect promotes species diversity through allopatric speciation. Headwater species are basal in the phylogeny, and shared haplotypes occur only at the headwaters; this research is a contribution to Wallace's riverine diversification hypothesis in the Amazon basin.

A central question in the design of effective conservation programs is what geographical regions to protect in order to maintain the most biological diversity. The term biodiversity hotspot was coined by Myers (Myers, 1990) and most commonly refers to regions of high species richness. GAP analysis is used to identify gaps in existing protected area networks (Scott *et al.*, 1993); it uses algorithms to select the minimum set of grid cells that encompass the unprotected species. Rarity and endemicity have been used to define hotspots in bird conservation (Balmford and Long, 1994), and species richness and endemism have been used to rank countries (McNeely *et al.*, 1990).

Hotspots are also defined as those areas with the greatest number of threatened species.

In setting conservation priorities, assumptions are made that indicator groups (e.g., macroorganisms such as birds, mammals, and plants) are good predictors of biological diversity in general. Another question that arises is how best to analyze biodiversity information to generate accurate and useful analyses that will inform conservation decisions. On a large scale, some concordance is found between bird diversity across continents with insect diversity (Pearson and Cassola, 1992), and in endemism patterns across taxa (Lawton, 1994); but at a finer spatial scale this correlation begins to break down. Richness in genera and families are good predictors of species richness at a finer level (Balmford *et al.* (1996a) and Balmford *et al.* (1996b)). However, species richness is not a good measure with which to identify hotspots for conservation because it overlooks rare species, although as the sample area for hotspots is increased, more rare species are included as a simple function of arithmetic progression. Rarity and endemicity are efficient indices for selecting the most parsimonious number of sites, but compared to complementarity measures they are less useful in defining conservation priorities.

A good conservation measure is complementarity, where the species complement of a reserve or area is identified and then further sites are found that add the greatest number of new species; this is akin to the portfolio approach (Swingland, 1997). Another method using integer linear programming to choose the optimal set of sites (maximal-covering-location; Church *et al.*, 1996) is limited to small datasets and does not achieve the greatest conservation gain for the fewest additional sites. Clearly, combining an ecosystem portfolio approach with a richness or endemism assessment would be effective, but differing approaches are needed according to the conservation goal and data availability.

Ex Situ Conservation

Viable populations of many organisms can be maintained in cultivation or in captivity. Plants may also be maintained in seed banks and germplasm collections; similar techniques are under development for animals (storage of embryos, eggs, and sperm, i.e., "frozen zoos") but are more problematic. *Ex situ* conservation is extremely costly in the case of most animals, and while it would in principle be possible to conserve a very large proportion of higher plants *ex situ*, this would be feasible for only a small percentage of the world's organisms. Furthermore, it often involves a loss of genetic diversity through founder effects and the high probability of inbreeding (Milner-Gulland and Mace, 1998).

Contextual Variations of the Definition

Derivation of "Biodiversity"

The definition of biodiversity put forth by the Office of Technology Assessment (1987) appears to be the most widely cited basis for other published definitions (Scott *et al.*, 1995). However, the OTA did not explain why they defined the term as they did, nor did they cite any supportive documentation. One problem with relying solely on authoritative sources for definitions of biodiversity is that different authorities have defined the term in fundamentally different ways.

"Bio" is derived from the Greek word bios, meaning life. "Biological" and "biotic" are terms that refer to life, living organisms, assemblages of living organisms, and the activities and interactions of living organisms. The scope of the term "biological" can be further understood in the context of components and processes that are considered biological. Defining biodiversity (i.e., diversity) is more difficult because it continues to be defined in several fundamentally different ways. In definitions of biodiversity, diversity has been characterized as (1) the number of different types of items, (2) the number of different types of items and their relative abundance, and (3) variety. Characterization of diversity in discussions of biodiversity has also included the structural complexity of landscapes (Huston, 1994).

Classifying Biodiversity

The classification of biodiversity can be divided into those authors who consider biodiversity to be a state and those who believe that it is a measure of the state.

Most authors have defined biodiversity as a state or attribute, for example, "biodiversity is the variety of…" or "variety and variability of…" (Noss and Cooperrider, 1994). Standard dictionaries have classified diversity as a state, condition, or quality (Soukhanov *et al.*, 1988).

Other definitions of biodiversity limited the scope of the attribute to explicit, quantifiable dimensions or measures, for example, "biodiversity is the number of…" or "the number and relative abundance of…" (Office of Technology Assessment, 1987). This emphasis on quantitative, operational definitions of biodiversity, and criticisms of nonquantitative definitions (Angermeier, 1994; Hunter, 1996) may signal a potential shift in the classification of the term from an attribute to a measure of an attribute. In the ecological and natural resource management literature, Pielou (1977) and others have treated diversity as a one- or two-dimensional attribute of a community (e.g., diversity is "the number of" or "the number and relative abundance of"). More recently, it has been defined as a measure or index of those attributes; for example, diversity is a "measure of…"(Noss and Cooperrider, 1994). Operational definitions of biodiversity (Angermeier, 1994) provide impetus to define biodiversity in quantitative terms as Hunter (1996) recommended.

Attributes of Biodiversity

Another way of delineating the meaning of a term is to list its characteristics, properties, qualities, and parts. Noss (1990) recognized three main attributes of biodiversity: composition, structure, and function.

Composition addresses the identity and richness of biotic components, and the relative amount (e.g., abundance, cover, biomass) of each (Noss, 1990). Biotic components of ecosystems include genes, organisms, family units, populations, age classes, species and other taxonomic categories, trophic levels of animals (e.g., herbivores, predators), animal guilds and assemblages, plant communities, and interacting

assemblages of plants, animals, and microorganisms (i.e., biotic communities).

Structural attributes of biodiversity refer to the various vertical and horizontal components of a community or landscape (Noss, 1990) and the organizational levels of plant and animal populations and assemblages (Gaston and Spicer, 1998; Hunter, 1996). Considering only biotic, vegetative components of a landscape, horizontal structure consists of the size, shape, and spatial arrangement and juxtaposition of different plant communities; vertical structure consists of the foliage density and height of different vegetation layers (Noss, 1990). Structure can also refer to population, age and trophic structure, and other levels of community organization (Hunter, 1996).

The inclusion of structure in the meaning of biodiversity provides linkages with other concepts, such as habitat diversity and the plant community concept, for both of which vegetation structure is an important differentiating attribute. Structure may have been left out of most definitions of biodiversity because the concept of biodiversity evolved from the concept of ecological diversity, which primarily focused on species diversity (Fisher *et al.*, 1943). Interestingly, 20 years ago it was asserted that measurements of diversity should not preclude structural diversity even though the term is most often used in reference to species diversity. Diversity can also be used in reference to niche width and the structural complexity of habitats.

Biotic functions represent the third component of biodiversity, and these include processes such as herbivory, predation, parasitism, mortality, production, vegetative succession, nutrient cycling and energy flow through biotic communities, colonization and extinction, genetic drift, and mutation (Noss, 1990). Biotic processes can be addressed in terms of the identity and number of different types of processes, as well as the rate (e.g., predation rate) at which each process operates.

Diversity of biotic components and processes can be observed at many biogeographic scales, from microsites and larger-scale landscape elements (e.g., vegetation types, habitat types, range sites) to regional landscapes, biomes, continents, hemispheres, and the entire biosphere (Noss, 1990; Huston, 1994; Hunter, 1996). Although these are scales at which biodiversity can be observed, they are not necessarily scales of biodiversity because most include abiotic (e.g., geological) features. Biodiversity can also be observed at several organism-based scales, including individual organisms, populations, species, and assemblages (e.g., guilds and plant communities), which themselves can be observed at various biogeographical scales.

Biological Resource Asset and Management Objectives

The contextual variations in the definition of biodiversity depend on what use is being made of the biological resource asset (or bioasset), and thus the asset management objective. Biological resource values consist of direct use, indirect use, and option and nonuse values. For the purposes of assessing potential use, they can be further classified as follows:

- Direct use values of major extractive products. Principally, this would include forestry for timber and commercial fisheries in the case of terrestrial and marine systems. Extraction of these products often involves substantial investment in capital equipment by large nonlocal firms, and the products are transported and sold in well-developed markets far from their original source.
- Direct use values of "minor" extractive products. These are naturally or seminaturally occurring products that require labor-intensive gathering or harvesting activities, often carried out by local people. Examples include rattan, fuelwood, seaweed, wild foods, artisanal fisheries, aquarium fish, and medicinal herbs. These may be collected for sale, barter, or home consumption.
- Direct use values that require the extraction of only a small amount of biological material for *ex situ* research or storage. This includes extraction of material for biological inventories, germplasm banks, and industrial research. Extraction is often accomplished during short or long expeditions that traverse large areas to collect representative samples of biological material.
- Direct use values that are nonextractive, but often require considerable on-site interaction of the user with the resource. This includes ecotourism, recreation, on-site research, and other major "nonconsumptive" activities occurring principally in protected areas. These activities are characterized by the need to provide food, lodging, and transport to the participants.
- Indirect use values that accrue on site. The primary feature of these values is that they support or protect the basic functioning of the protected area. Examples include nutrient cycling, stabilization of soils in erosion-prone areas, coastal zone stabilization, and biological support to local ecosystems. As a result of their nature, the value of these on-site functions is likely to be a component of all the other direct and nonuse values generated by the area.
- Indirect use values that accrue off site. The value of these functions – such as watershed protection, natural ecosystems protected as national parks in generating income from wildlife tourism, protection of fisheries' nurseries and subsistence fisheries, and climate regulation – may be very large or very small depending on their relative importance to the support or protection of off-site economic activity.
- Option values. Because option values may be associated with each and every use value, they are considered only where they may be of potential significance in conjunction with the particular type of product or service.
- Nonuse values. By their very nature these values occur at a distance from the resource and require no extraction or physical interaction with the resource, for example, stewardship, ethics, cultural belief, and esthetics.

These foregoing values are only indirectly related to biological diversity. That is, a certain level of species richness is required for these functions but there is not necessarily a direct correlation between the value of the ecosystem and its diversity. Thus, mangrove ecosystems are generally of far lower diversity than adjacent lowland terrestrial forests, but in resource terms they are likely to be of comparable value. The savannas of eastern and southern Africa, which are of great importance in generating revenues from tourism, are less diverse than the

moist forests in these countries, which have far less potential for tourism.

Cave Canem or the Precautionary Principle

At present, humans actively exploit a relatively small proportion of the world's biological diversity. Many other potential, yet undiscovered, optional and nonuse values of biodiversity exist. These factors support a precautionary approach to maintaining biological diversity. In this case, the precautionary principle argues that actions should be taken to prevent further loss of biodiversity and potentially irreversible consequences before all biological uncertainties are resolved. Yet in conserving biodiversity, there must come a point at which the projected costs required to protect and maintain it will outweigh any probable benefits.

If species are to be viewed as a resource, and their maintenance is to be cost-effective, conservation should concentrate on systems and areas rich in species, and on those species known to be useful, or regarded as having a high probability of being useful. Thus biodiversity and its conservation would be defined purely along operational or cost–benefit lines. This bioasset perspective on biodiversity would therefore rest on economic arguments more than biological ones.

Biodiversity has been identified as important for ecosystem health, medicinal values, agricultural purposes, and esthetic and recreational values (Noss and Cooperrider, 1994). Noss (1990) characterized an operational definition as one that is responsive to real-life management and regulatory questions, adding that such a definition is unlikely to be found for biodiversity. Angermeier (1994) referred to an operational definition in a similar way, and Hunter (1996) suggested that a quantitative definition is needed for monitoring biodiversity and developing management plans. However, some writers assert that the confounding of definition and application is partly to blame for the confusion over how biodiversity concepts can be practically implemented.

Implications of Variations in the Definition

The need for an unequivocal and precise meaning of biodiversity that is scientifically sensible and universally applicable is imperative to help guide the design of policy and programs for the future, as well as to make critical decisions in the present. Currently, such a definition does not exist. As a concept, biodiversity is both ubiquitous and useful, particular and confusing; and for this reason it is constantly redefined on nearly every occasion.

One of the many reasons for this state of affairs is that the definition of biodiversity affects objectives in national and regional research and conservation management, and in international funding priorities. One could easily promote a timber extraction or nontimber forest product program that conserves species richness (i.e., numbers of species) at the expense of genetic diversity. Indeed, a current research program to stimulate or increase the range of tropical tree species not currently in trade, as a way to take the pressure off overexploited species, may be misguided. It may lead to increased genetic as well as species impoverishment when foresters expand the number of species they take and select only the best and most mature specimens, thus removing the most productive and healthiest genetic stock.

Apart from the principal definitions of biodiversity discussed earlier, such as the highest number of species (i.e., species richness) and the highest level of species endemicity (Myers, 1990) or taxal endemicity (called critical faunas analysis), interpretations of pure or applied definitions are becoming more common within the vocabulary of conservation and biodiversity utilization when determining biodiversity management priorities. Some examples are national biodiversity programs that maintain "biodiversity portfolios"; biodiversity defined as flagship or keystone species diversity; viability modeling (population viability analysis) defining the species' populations to be prioritized; population analysis defining sustainability and thus defining a species' status; projects that focus on the feasibility of integrating the targeted species, assemblages, or ecosystems with the needs of local human populations and sustainable use; and (lastly) political exigency (Swingland, 1997). Although the conservation policy of a country may be driven by more pressing needs – family planning, education, politics, internal conflict, financial planning and investment, individual vested interests – current policy and decisions are also being made on the foregoing biodiversity bases rather than along strict academic lines.

Endemicity and species richness are useful starting points in defining priorities on the global level, but without information on the possibility of extinction using viability modeling or population analysis, the urgency of a given conservation action cannot be assessed. Moreover, with the increasing emphasis on the integration of local people into conservation programs to minimize long-term costs and to provide a more stable basis for the people and their natural environment, the potential for community-based conservation, coupled with sustainable use, cannot be ignored. Since national or external funding will generally provide essential support for most projects, the ecological importance of an area relative to others using an ecosystem diversity (or portfolio) approach will be a major selection criterion. The presence of a flagship or keystone species will also be significant in raising such funds. Clearly political exigencies or pure chance can enter the situation, and scientists have yet to articulate whether genetic diversity should be used as the key measure. In the absence of realistic methods of quantifying these biodiversity characteristics, they must remain imponderable objectives for the moment.

The differing approaches being advocated for biodiversity conservation are not just guided by the available methodologies but are also symptomatic of the underlying philosophies. The evolution-based approach is predominantly the preserve of biologists, and it is concerned with the maintenance of diversity as an unqualified objective unaffected by economics. The need for conservation and the uses of biodiversity – the resource-based argument – are what are used to "sell" the proposition to decision makers and policy-makers. Where these factors come together, the ideal of ecological sustainability and the conservation methods of achieving it will be possible. Because so much is now formally invested in using the word "biodiversity", its definition will continue to

play a crucial role in both conservation planning and public policy.

See also: Biodiversity, Origin of. Ecology, Concept and Theories in. Ecosystem, Concept of. Genetic Diversity. Habitat and Niche, Concept of. Loss of Biodiversity, Overview. Measurement and Analysis of Biodiversity. Taxonomy, Methods of

References

Addicott JF (1974) Predation and prey community structure: An experimental study of mosquito larvae on the protozoan communities of pitcher plants. *Ecology* 55: 475–492.

Agarwal A (1992) Sociological and political constraints to biodiversity conservation: A case study from India. In: Sandlund O, Hindar K, and Brown AHD (eds.) *Conservation of Biodiversity for Sustainable Development*, pp. 293–302. Oslo, Norway: Scandinavian University Press.

Angermeier PL (1994) Does biodiversity include artificial diversity? *Conservation Biology* 8: 600–602.

Babbitt B (1994) Protecting biodiversity. *Nature Conservancy* 44: 1621.

Balmford A and Long A (1994) Avian endemism and forest loss. *Nature* 372: 623–624.

Balmford A, Green MJB, and Murray MG (1996a) Using higher-taxon richness as a surrogate for species richness: I. Regional tests. *Proceedings of the Royal Society of London, Series B* 263: 1267–1274.

Balmford A, Jayasuriya AHM, and Green MJB (1996b) Using higher-taxon richness as a surrogate for species richness: II. Local applications. *Proceedings of the Royal Society of London, Series B* 263: 1571–1575.

Bryant D and Barber C (1994) *World Resources: 1994–95*. New York: World Resources Institute/Oxford University Press, pp. 147–164.

Church RL, Stoms DM, and Davis FW (1996) Reserve selection as a maximal covering location problem. *Biological Conservation* 76: 105–112.

DeLong DC (1996) Defining biodiversity. *Wildlife Society Bulletin* 24: 738–749.

Dodson SI, Allen TFH, Carpenter SR, *et al.* (1998) *Ecology*. New York: Oxford University Press.

Fiedler PL and Jain SK (eds.) (1992) Conservation biology: The theory and practice of nature conservation. *Preservation and Management*. New York: Chapman and Hall.

Fisher RA, Corbet AS, and Williams CB (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12: 42–58.

Gaston and Spicer (1998) *Biodiversity: An Introduction*. Oxford, UK: Blackwell Science.

Ghikrov A (1996) What does "biodiversity" mean—Scientific problem or convenient myth? *Trends in Ecology and Evolution* 11: 304–306.

Ghilarov A and Timonin AG (1972) Relations between biomass and species diversity in marine and freshwater zooplankton communities. *Oikos* 23: 190–196. F.

Haffer J (1969) Speciation in Amazon forest birds. *Science* 165: 131–137.

Harper JL and Hawksworth DL (1994) Biodiversity: Measurement and estimation. *Philosophical Transactions of the Royal Society of London, Series B* 345: 5–12.

Hubbell SP (1997) A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs* 16: S9–S21.

Hudson WE (ed.) (1991) *Landscape Linkages and Biodiversity*. Washington, DC: Island Press.

Hunter Jr. ML (1996) *Fundamentals of Conservation Biology*. Cambridge, MA: Black-well Science.

Huston MA (1994) *Biological Diversity: The Coexistence of Species on Changing Landscapes*. New York: Cambridge University Press.

Knopf FL (1992) Focusing conservation of a diverse wildlife resource. *Transactions of the North American Wildlife and Natural Resources Conference* 57: 241–242.

Lawton JH (1994) Population dynamic principles. *Philosophical Transactions of the Royal Society of London, Series B* 334: 61–68.

Lovejoy TE (1980) The Global 2000 Report to the President. In: Barney GO, (ed.) *The Technical Report*, vol. 2, pp. 327–332. New York: Penguin.

MacArthur RH (1957) On the relative abundance of bird species. *Proceedings of the National Academy of Sciences USA* 45: 293–295.

MacArthur RH and MacArthur J (1961) On bird species diversity. *Ecology* 4: 594–598.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Margalef R (1957) La teoria de la informacíon en ecologia. *Memories de la Real Academia de Ciencias y Artes (Barcelona), 3rd series* 32: 373–449.

McNeely JA, Miller KR, Reid WV, Mittermeier RA, and Werner TB (1990) *Conserving the World's Biological Diversity*. Gland, Switzerland/Washington, DC: International Union for Conservation of Nature and Natural Resources/World Resources Institute/Conservation International/World Wildlife Fund–U.S./World Bank.

Milner-Gulland EJ and Mace R (1998) *Conservation of Biological Resources*. Oxford, UK: Blackwell Science.

Mittermeier RA, Carr JL, Swingland IR, Werner TB, and Mast RB (1992) Conservation of amphibians and reptiles. In: Adler K (ed.) *Herpetotlogy: Current Research on Amphibians and Reptiles*, pp. 59–80. Society for the Study of Amphibians and Reptiles. Contributions to Herpetology.

Myers N (1990) Threatened biotas: "Hotspots" in tropical forests. *Environmentalist* 8: 1–20.

Noss RF (1990) Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355–364.

Noss RF and Cooperrider AY (1994) *Saving Nature's Legacy: Protecting and Restoring Biodiversity*. Washington, DC: Island Press.

Office of Technology Assessment (1987) *Technologies to Maintain Biological Diversity*. Washington, DC: U. S. Govt. Printing Office.

Pearson DL and Cassola F (1992) World-wide species richness patterns of tiger beetles (Coleoptera: Cicindelidae): Indicator taxon for conservation and biodiversity studies. *Conservation Biology* 6: 376–391.

Pielou EC (1977) *Mathematical Ecology*. New York: John Wiley & Sons.

Reid WV and Miller KR (1989) *Keeping Options Alive: The Scientific Basis for Conserving Biodiversity*. Washington, DC: World Resources Institute.

Ricklefs RE and Schluter D (1993) *Species Diversity in Ecological Communities*. Chicago: University of Chicago Press.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.

Samson FB and Knopf FL (1994) A framework to conserve biological diversity through sustainable land management. *Transactions of the North American Wildlife and Natural Resources Conference* 59: 367–377.

Scott JM, Abies ED, Edwards Jr. TC, *et al.* (1995) Conservation of biological diversity: Perspectives and the future for the wildlife profession. TWS committee report. *Wildlife Society Bulletin* 23: 646–657.

Scott JM, Davis F, Csuti B, *et al.* (1993) GAP analysis: A geographic approach to protection of biological diversity. *Wildlife Monographs* 123: 1–41.

Soukhanov AH, Ellis KE, Harris DR, DeVenne PB, and Severynse M (eds.) (1988) *Webster's* II *New Riverside University Dictionary*, p. 1563. New York: The Riverside Publ. Co.

Sousa WP (1979) Disturbance in marine intertidal boulder fields: The nonequilibrium maintenance of species diversity. *Ecology* 60: 1225–1239.

Swingland IR (1997) Global conservation and the sciences: People, policy and pennies. In: Van Abbema J (ed.) *Conservation, Restoration and Management of Tortoises and Turtles*, pp. 453–464. New York: New York Turtle and Tortoise Society.

Swingland IR (1999) Commercialisation, structure and sustainability of biodiversity conservation. In: Walkey M, Swingland IR, and Russell S (eds.) *New Practices in Integrated Protected Area Management*. London: Chapman and Hall.

Temple SA (1986) Predicting impacts of habitat fragmentation on forest birds: A comparison of two models. In: Verner J, Morrison ML, and Ralph CJ (eds.) *Wildlife 2000: Modelling Habitat Relationships of Terrestrial Vertebrates*, pp. 301–304. Madison: University of Wisconsin Press.

Thomas R (1992) Genetic diversity. In: World Conservation Monitoring Centre (ed.) *Global Biodiversity: Status of the Earth's Living Resources*, pp. 1–6. London: Chapman and Hall.

Vane-Wright RI (1992) Systematics and diversity. In: World Conservation Monitoring Centre (ed.) *Global Biodiversity: Status of the Earth's Living Resources*, pp. 7–12. London: Chapman and Hall.

Vane-Wright RI (1996) Systematics and the conservation of biological diversity. *Annals of the Missouri Botanic Gardens*.

Whittaker RH (1972) Evolution and measurement of species diversity. *Taxon* 1: 213–251.

Williams CR (1964) *Patterns in the Balance of Nature*. New York/London: Academic Press.

Williams PH, Humphries CJ, and Vane-Wright RI (1991) Measuring biodiversity: Taxonomic relatedness for conservation priorities. *Australian Systematic Botany* 4: 665–679.

Wilson EO (1961) The nature of the taxon cycle in the Melanesian ant fauna. *The American Naturalist* 95: 169–193.

Wilson EO and Peter FM (eds.) (1988) *Biodiversity*. Washington, DC: National Academy Press.

World Conservation Strategy: Living Resource Conservation for Sustainable Development. Gland, Switzerland: International Union for Conservation of Nature and Natural Resources/U.N. Environment Programme/World Wildlife Fund.

Biodiversity, Evolution and

Gregg Hartvigsen, State University of New York, Geneseo, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity A measure of the relative abundance of species or higher taxonomic units found in a certain area at a particular time. In ancient communities, this is often simply referred to as the number of species or taxa found at a site, usually referred to as "species richness" in the ecological literature.

Community A group of interacting organisms, representing multiple species in a given location.

Complex adaptive systems A group of individuals or types that exhibit variability in which large-scale patterns emerge from small-scale interactions. The system is adaptive in the sense that it is subject to extrinsic selection that leads to a change in structure and dynamics over time.

Evolution The morphological or genetic change in species over time. Small changes that do not lead to reproductive isolation among members of a group are referred to as "microevolution." Speciation, or the generation of new species, is generally referred to as "macroevolution."

Fossil Anything found in a strata of rocks or sediments that is recognized as the remains of an organism from a former geological time.

Species A group composed of individuals that interbreed or potentially interbreed and are separated reproductively from other such groups.

Taxon (taxa) A group of individuals representing a classification of organisms, such as genus, family, or order. Higher taxa are those above the species level.

Introduction

The earliest record of life on Earth is from approximately 3.6 billion years ago. Aside from the various hypotheses as to how life arose, a topic beyond the scope of this article, the fossil record clearly reveals that early life forms were single-celled organisms 00of the group Archaea with structures similar to modern bacteria. The Earth at that time was far different than today and would have been uninhabitable to the vast majority of the contemporary species. Our understanding of these life forms that have arisen over this time to the present relies mainly on the fossilized remains of organisms representing an array of groups that far exceeds those living today. Estimates for the diversity of life on Earth, based on the small sample of discovered fossils, suggest that greater than 99.9% of species have evolved and disappeared through extinction. Although the fossil record contains discontinuities and absences of these taxa, evidence is more than sufficient to reveal the evolution of the diversity of life and even provide details in the changing of communities, or groups of species, over time. This level of detail is astounding, but it is only suggestive of far greater details likely to be uncovered in the coming decades. In addition, recent development and use of computer models is enabling us to both understand these paleobiological data and make predictions of how evolution influences the dynamics of communities over time. The author suggests that the coupling of our rapidly increasing knowledge of the fossil record and our use of computer modeling will help us to both understand and predict the influence of humans on the patterns of biodiversity.

The Fossil Record

What Does the Fossil Record Indicate?

In general, the fossil record is an outstanding record of the changes in biodiversity, revealing both an increase and turnover in taxa over time. Data for marine organisms suggest that the diversity of families has generally increased during the Phanerozoic, which includes the past 540 million years (My), the time for which fossilized remains of multicellular organisms exist (**Figure 1**; Sepkoski, 1991). A broader survey suggests that diversity has increased almost exponentially for all known families, genera, and species during this period

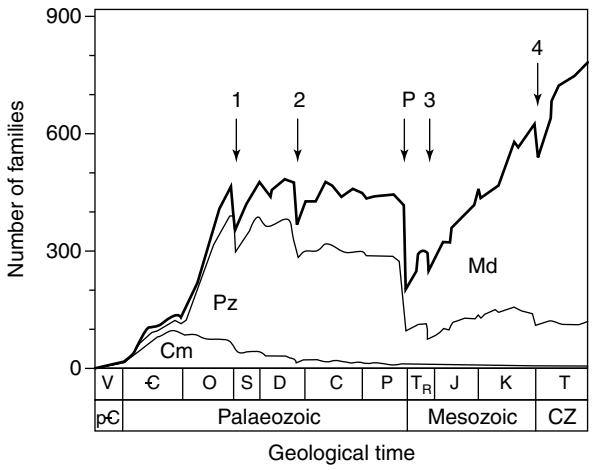


Figure 1 The diversity of skeletonized marine animal families during the Phanerozoic. The bold line represents total numbers of families known from the fossil record. Five major extinction events are indicated: 1, end Ordovician; 2, Late Devonian; P, end Permian (largest); 3, end Triassic; 4, end Cretaceous. V, Vendian; C, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; TR, Triassic; J, Jurassic; K, Cretaceous; T, Tertiary. Faunal assemblages are categorized as Cm, Cambrian; Pz, Paleozoic; and Md, Modern. Reproduced from Sepkoski Jr. JJ (1991) Diversity in the Phanerozoic oceans: A partisan review. In: Dudley EC (ed.) *The Unity of Evolutionary Biology*, pp. 210–236. Portland, OR: Dioscorides Press.

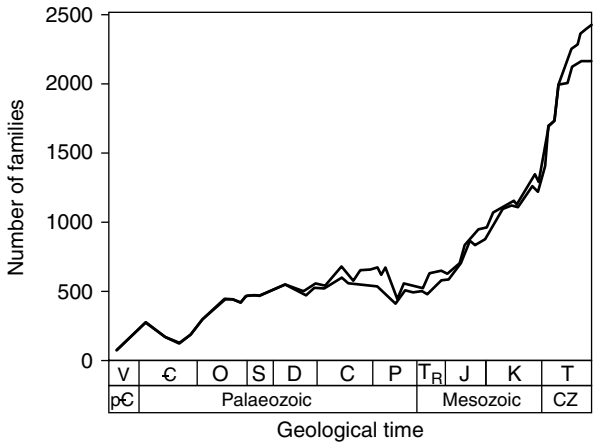


Figure 2 Diversity has increased exponentially for all known families, genera, and species during the Phanerozoic. The top and bottom lines represent maximum and minimum numbers of families, respectively. See the legend to **Figure 1** for abbreviations. Reproduced from Benton MJ (1995) Diversification and extinction in the history of life. *Science* 268: 52–58.

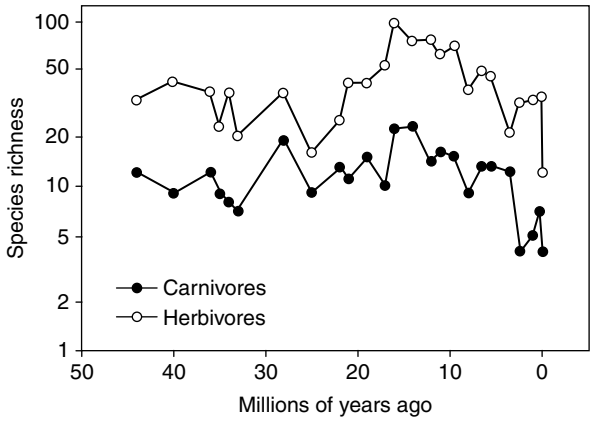


Figure 3 The number of large mammal species from North America during the past 44 million years. Note the correlation between the diversity of predators and their potential prey ($r^2 = 0.43$; $p < .001$). Data from Van Valkenburgh B and Janis CM (1993) Historical diversity patterns in North American large herbivores and carnivores. In: Ricklefs R and Schluter D (eds.) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*, pp. 330–340. Chicago: University of Chicago Press.

(**Figure 2**; Benton, 1995). More recently, works by Courtillot and Gaudemer (1996) and Pave *et al.* (2002) have argued that biodiversity has increased logistically, with a relatively rapid increase followed by a leveling of diversity over time (see **Figure 1**).

Increases in taxa, however, have been discontinuous, with major catastrophic losses of taxa in the form of mass extinctions, such as that which occurred at the end of the Cretaceous period that led to the loss of the dinosaurs. Thus, the accumulation of biodiversity in the form of numbers of taxa has been discontinuous and has exhibited a skewed distribution in the magnitude of the size of extinction events, ranging from relatively low-level background extinction rates to large-scale mass extinction events.

In general, we know that species are generally shortlived, averaging a duration on the order of 5 My, with few taxa being long-lived. Most species are, and are inferred to have been, composed of relatively few individuals which are locally distributed and few species being relatively abundant and widely distributed. We also know that most genera have few species, whereas most families have few genera. These patterns have led and continue to lead to high probabilities that any particular taxon is likely to be relatively short-lived as opposed to long-lived, where the mean duration of taxa exceeds the median duration.

The Number of Taxa has Increased over Time

A clear and dramatic trend exists in the fossil record of an increase in diversity. This pattern exists for taxa ranging from the level of species to phyla, and it has been well documented for plants, invertebrates, and vertebrates. This pattern is seen, for example, in the number of marine families over time during the Phanerozoic eon, since about 540 million years ago (**Figure 1**). The major additions to diversity in the fossil record include skeletonized marine invertebrates during the Cambrian, especially trilobites, brachiopods, and archaeocyathans. Corals, cephalopods, ostracods, crinoids, and starfish arose through the remainder of the Paleozoic, and bivalves, gastropods, echinoids, teleost fish, and marine reptiles arose during the Mesozoic. Diversity increased on land and included the evolution of vascular plants (Silurian and Devonian), gymnosperms (Carboniferous), and angiosperms (Jurassic). Insect diversity exploded in the late Paleozoic, followed by social insects in the Jurassic. Major clades of terrestrial vertebrates arose by filling vacant niches, including amphibians (Devonian), reptiles (Carboniferous), pterosaurs and dinosaurs (Triassic and Jurassic), birds (Jurassic), and mammals in the early Tertiary.

The pattern, however, is one of fluctuations over time in the shorter term. The number of species of relatively large herbivores and carnivores in North America during the past 44 My, for example, appears to exhibit little long-term trend toward increased or decreased diversity (**Figure 3**). Interestingly, however, these data suggest that the number of predators depends on the number of potential prey species over time. There is a significant correlation between the number of predators and preys which is consistent with the hypothesis that interspecific associations are interdependent within communities over time.

Extinction Events are Followed by Evolutionary Radiations

Extinction and the origin of species define the dynamics of the fossil record. Species, for instance, typically last on the order of 1–10 My and appear to go extinct in a nonrandom fashion. Therefore, species arise and ultimately disappear throughout time, producing a general pattern of loss referred to as "background" extinction (**Figure 4**).

Biodiversity at various levels of organization has increased during the Phanerozoic. There exist, however, at least five catastrophic extinction events that have led to a decrease in the number of taxa, and therefore biodiversity, that greatly exceed the background extinction rate (**Figure 4**). In particular, the extinction event at the end of the Permian (**Figure 1**, "P")

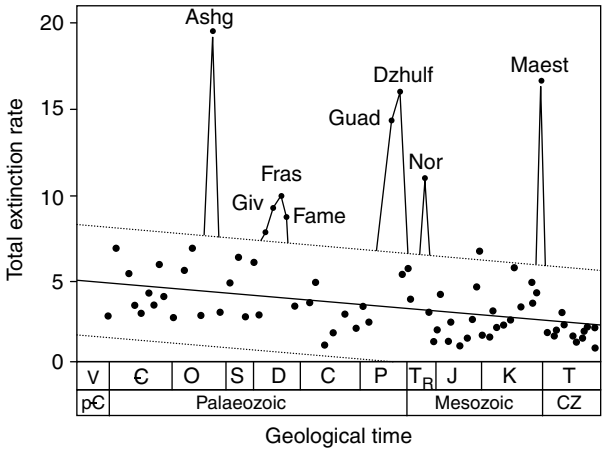


Figure 4 Background and mass extinction of marine animal families during the Phanerozoic. Rates are reported as the number of families lost per million years. Extinctions within dotted lines represent background rates, whereas solid lines above the upper dotted line represent mass extinctions. Mass extinctions: ASHG, Ashgillian (end Ordovician); GIV, Givetian; FRAS, Frasnian; FAME, Famennian (Late Devonian); GUAD, Guadalupian; DZHULF, Dzhulfian (end Permian); NOR, Norian (end Triassic); MAEST, Maastrichtian (end Cretaceous). For other abbreviations, see the legend to **Figure 1**. Reproduced from Raup DM and Sepkoski Jr. JJ (1982) Mass extinctions in the marine fossil record. *Science* 215: 1501–1503.

resulted in a loss of a maximum estimated 96% of species on Earth (although this is likely an overestimate; Raup, 1994).

The overall pattern of extinction appears to follow a well-defined "kill curve," proposed by Raup, that indicates a skewed distribution of species loss ranging over the gradient where most time periods exhibit low rates of extinction (background) and few periods of very high rates of extinction (mass extinctions). These data, therefore, suggest a predictable pattern, based on probability, that describes species extinctions.

Following extinctions there is general radiation, or rapid evolutionary diversification, of taxa. This is seen in Sepkoski's database of marine families (**Figure 1**) and even more clearly in Benton's larger database of families from both terrestrial and marine systems following the five major extinction events (**Figure 2**).

The rates of species origination have varied over time and are highest following periods of mass extinction and in times when new ecospace became available, such as during the initial rise of diversity during the late Vendian and Cambrian (**Figures 1** and **2**). Speciation was highest for terrestrial families during the Priabonian (late Eocene), most likely due to the increase in insect diversity.

How Good Is the Fossil Record?

The fossil record is far from a complete representation of all of the Earth's biotic history. There are only bits and pieces of evidence from organisms and we are able to date strata from which organisms derive only with varying levels of accuracy. The fossil record is better, in general, for recent strata and decreases in completeness with age. The fossils we do have, however, enable us to construct a logical progression in both complexity of form and diversity over time. The idea that evolution is progressive, however, has long been abandoned in light of the many examples of lineages that do not follow straight progressions of simple to more complicated forms over time. Moreover, extant species that appear ancient, such as the coelacanth, horseshoe crab, and cockroach, are neither unchanged from their ancestors nor are their longevities unexpected. By chance we can expect variance in the longevity among lineages, with many short-lived species and few long-lived species.

Understandably, only certain morphological characteristics have been susceptible to fossilization. Of those that are fossilized we often find only disarticulated parts of organisms, such as a tooth, part of a shell, or perhaps a bone fragment, separated over varying distances. Soft-bodied organisms and soft tissues are rarely observed. To complicate matters, these fragments are susceptible to movement both spatially (e.g., down-stream after death) and vertically in the stratigraphy, which makes fossils appear to move temporally. This occurs when sedimentation rates are very low, allowing noncontemporaneous organisms to become fossilized together, or when layers are mixed or folded, a process referred to as "time averaging" of the fossil record. However, taphonomy, or the study of postmortem behavior of organic remains, indicates that a vast majority of organisms are not transported out of their original habitats and therefore can be used to construct probable past communities. Evidence for this derives primarily from the observation of contemporary local death assemblages in which individuals rarely move across community boundaries.

Microfossils such as pollen are abundant in lake sediment and can be used for quantitative reconstruction of local plant communities. Pollen from sediments on the order of a few hundred thousand years has been used to reasonably describe assemblages. Little is known about the role that evolution played in plant communities from sediments more than 1 My old. The pollen record during the Holocene (e.g., since the past glacial period ended about 16,000 years ago) has been widely and more confidently used to reconstruct past plant communities. Therefore, we remain perplexed with regard to the various assemblages of taxa that likely resided concurrently in more ancient communities, but for more recent times we have increasing confidence in our effort to reconstruct communities from fossil remains.

What Does the Fossil Record Indicate About Communities?

The interpretation of community structure in the fossil record is more problematic than documenting the change in taxa over time. The primary difficulty lies in determining whether fossilized taxa occurred contemporaneously. Postmortem movement of organisms and their parts appears to take place over relatively short distances and within time-averaged ranges of species, which is a pattern found for both marine and terrestrial taxa. Therefore, optimistically, it appears that cooccurring taxa through the fossil record will likely represent evidence of past communities.

Evidence from pollen records, however, can produce a paradox in our interpretation of past plant communities. Over long periods of time, pollen deposition in lake sediments

generally reflects nearby plant communities accurately. The distance-dependent deposition of pollen is generally correlated with the sum basal area of trees (cross-sectional area of tree diameters) within a few square kilometers of the surrounding area, although many factors influence the pollen collected from lakes and the surrounding vegetation, such as relative decomposition rates and the relative efficiency of a lake to capture pollen. Pollen records from lake sediments, however, suggest that the composition of plant species changes over time and that the species change independently of each other in response to rapid climate change following the Wisconsin glacial period in North America. In addition, there is a consistent pattern that, among late Quaternary pollen records for basins of different sizes and from different regions (temperate and tropical), communities appear to be the product of independently acting species.

The rain of leaf litter and other debris, likewise, is well correlated in contemporary communities with the abundance of trees within local communities. Unfortunately, leaf litter, in general, rapidly decomposes and rarely provides evidence of past communities.

The fossil record remains difficult to interpret with regards to community composition and dynamics over time. Considerable research is needed and is under way on the dynamics of communities over long periods of time. In particular, we need to resolve conditions that lead to the following paradox: Taphonomy of current assemblages suggests fossil taxa may represent contemporaneous communities, whereas pollen studies suggest that at least plant species have acted relatively independently over time. Additional data are needed to help reconstruct the composition of past communities accurately. In addition, the expanding use of detailed DNA sequence data will undoubtedly help to resolve phylogenetic relationships over evolutionary time.

Current Patterns of Biodiversity

Biodiversity is heterogeneously distributed on Earth, with the general pattern that species diversity is greatest near the equator and decreases with increasing latitude. Hypotheses and evidence for these patterns abound. In the following sections, the author briefly introduces how we define biodiversity, discuss very general trends in communities, and apply these to understanding biodiversity of communities over time.

How do we Define Biodiversity?

Before discussing patterns of biodiversity change over time, the author needs to define diversity. In general, paleontologists consider biodiversity to be equivalent to the number of unique taxa. In the ecological literature, diversity generally incorporates both the number of species and equitability or the relative abundance of individuals among species. A simple hypothetical example is the relative diversities of two communities. Assume that both have 100 individuals within two species. One community has 99 individuals of one species and the second community has 50 individuals of both species. The first community is essentially a monoculture of species 1. The second community is considered a more diverse community, despite the fact that both communities have the same number of species present.

Accounting for equitability in the fossil record, however, is quite problematic because of the spotty nature of the preserved specimens. It must be recognized that fossils in an assemblage will not likely constitute a random sample of individuals from a community. Rather, the remains of fossil communities are likely to be composed of a skewed distribution of individuals representing only a subset of those species that contained structures susceptible to fossilization. Other species possibly escaped fossilization completely or were missed in the sample from which the fossils were found. In addition, the subset that is found is more than likely to be small and therefore highly susceptible to the variance incurred from the chance occurrence of individuals of rare species. This is an example of the problem referred to as the ''central limit theorem'' in statistics in which sampling of a skewed distribution (the reality of the fossil communities) yields a normally distributed population (not an accurate representation of the fossil communities). In other words, we are most likely to find fossils of the most abundant species, yielding an underestimate of the true number of species that occurred at that time in that community. Some fossil assemblages, however, will contain samples of relatively rare taxa, making these taxa appear more abundant than they really were. In addition, it is difficult to assess whether forms represent truly different taxa. We are more likely than not to lump taxa together because of our inability to know whether similar fossils represent members of an interbreeding population.

Patterns of Species Change Over Time

Simple change in the number of species over time can be expressed as a rate, similar to the change in a population over time. We can represent this change in the number of taxa over time as $dN/dt = RN$, where N represents the number of species in a genus or genera in a family, and $R = S - E$, where S is the speciation rate and E is the extinction rate. We may therefore assume $dN/dt = 0$ to represent a null hypothesis against which we test for significant changes in the rates of species additions or subtractions. There are several problems with this approach, however, since our null hypothesis is that all species are expected to eventually become extinct. Therefore, we must define a time constraint of interest, concluding that if $R < 0$, then the taxa of interest appear to be on the decline. Comparisons of the fossil record have been found to differ from this random null hypothesis. The database assembled by Sepkoski (**Figure 1**) suggests that diversity increases logistically, whereas the data set assembled by Benton (**Figure 2**) suggests diversity has increased exponentially during the past 600 My.

The work previously mentioned by Courtillot and Gaudemer (1996) and Pave *et al.* (2002) suggests that exponential growth for species number over time might not be supported by the data. These analyses are challenging, however, in the light of imagining the mechanisms that would bring about regulation in the number of species over time.

Are Communities Saturated with Species?

In general, it appears that many communities are at or near a saturation level for the number of species they contain. The evidence for this derives indirectly from several sources, including the relationship of families depicted previously in the fossil record and from the ecological literature in which is found a general inability of introduced species to successfully invade communities. Species introductions usually fail, and when they succeed, the species usually have little consequence on natural systems, despite the numerous celebrated examples of introduced species altering native communities. Occasionally, even with the loss of native dominant species such as the American chestnut tree to a fungal blight, the effect on natural processes may be equivocal.

The result of this contemporary observation suggests that communities may reach taxa saturation. If this pattern holds for prior paleocommunities, then this suggests that ecological and evolutionary processes operate to fill available niche space with either colonizing species or through the process of speciation. Inspection of either Sepkoski's or Benton's databases for the number of taxa over time suggests that the Earth is not at a carrying capacity. The following is the obvious question stemming from this: Where should we expect to find new species occurring, especially in light of the estimation that humans are currently consuming approximately 40% of the Earth's productivity?

Community Stability

Researchers in the field of ecology, and particularly conservation biology, have long been interested in the dynamics of communities and the rate of species change over time, also referred to as "species turnover." Hutchinson and his student MacArthur were early advocates of the notion that stability of communities is positively related to the number of species residing in the community. This relationship, quantified by MacArthur and incorporated into the influential theory of island biogeography developed in collaboration with Wilson, portrayed ecological systems as having stable equilibria, resulting from the trade-off between colonization and extinction rates on islands. Recently, this theory has been extended to systems that, in general, function like islands.

Although intuitively attractive, the positive relation between diversity and stability may not be universal. In addition, researchers have to assess the role diversity plays in community dynamics differently. For instance, in 1973, May suggested that there may be an inverse relationship between species diversity and community stability in model food webs, as measured by the long-term maintenance of species composition. Alternatively, research by Tilman on grasslands in Minnesota suggests that community resistance to drought, as measured by the maintenance of productivity, is positively related to the number of species in the community. These results suggest a complex interplay between diversity and various measures of stability. Understanding the role of diversity on the functioning of ecosystems is of particular concern as we face rapid changes in our global climate.

Communities as Complex Adaptive Systems

Recent developments in our understanding of community assembly, organization, and development over ecological and evolutionary time suggest that communities are more complex than previously believed. Classic approaches to understanding the dynamics of communities portrayed them as being composed of individual species that interact ecologically with little regard to spatial location or temporal dynamics. Recently, we have come to appreciate the intricacies of small-scale spatial and temporal partitioning of resources and interactions among co-occurring organisms. In addition, in the past 30 years we have come to appreciate the immense genetic diversity within species that enables species to adapt to biotic and abiotic factors that operate spatially within communities (Thompson, 1994).

Complex adaptive systems can be defined as groups of individuals that exhibit variability, interact with a subset of individuals within the global population (usually within some local neighborhood), and are susceptible to an autonomous, selective process that usually leads to individuals experiencing differential reproduction. These rules apply quite generally to species, communities, and other systems such as economies, allowing them to adjust to changing environmental conditions. Natural communities fit this general class of systems very well, and recent theoretical work has described the similar dynamics among diverse systems. Recent work, for instance, suggests that population and community dynamics may be influenced by the process of evolution acting on genetic diversity among individuals (Hartvigsen and Levin, 1997). It is likely that this new theoretical approach will lead to a greater understanding of how communities respond evolutionarily to disturbances over time scales greater than years.

The Effect of Disturbance on Community Biodiversity: Evolutionary versus Ecological Time

The causes of various mass extinction events are not known, but hypotheses do exist. The most recent mass extinction at the end of the Cretaceous was most likely caused by an asteroid impacting the Earth, forming the Chicxulub crater on the north side of the Yucatan Peninsula. The causes of the other mass extinctions are unresolved but may be the result of similar massive disturbances. The probability distribution of taxa loss appears to follow what we expect to occur by chance and matches the loss of taxa due to random factors. Occasionally, we should expect to lose many species, whereas over most time periods, only a few species should be expected to be lost. Therefore, the greater the disturbance, the greater the effect (reduction) on biodiversity.

The probability of species loss depends on many factors, including the size of the populations of species within the community. Smaller populations are more likely to go extinct than larger populations, a process referred to as demographic stochasticity. The Gambler's Ruin is a metaphor that is commonly associated with this process. The idea is that a gambler in a casino who has a limited number of betting units to lose (e.g., quarters for a slot machine or equivalents to the minimum bet at a black jack table) is more likely to go bankrupt simply by chance compared to a gambler with many betting

units. Likewise, stochastic, or random, changes in population size lead to small populations being more likely to reach the absorbing state of zero individuals than a large population. Therefore, and not surprisingly, small populations are at the greatest risk of extinction simply by chance.

Communities composed of species with small populations are therefore likely to suffer higher levels of species loss than communities composed of species with large populations. Raup (1994) argues that there appears to be little predictability in which species will be eliminated and that extinction does not appear to be selective. Species with narrow geographic ranges may be more susceptible to extinction, whereas species with large geographic ranges are not likely to experience disturbances so great as to influence their entire population.

Small populations typically exhibit lower genetic variability than large populations and are more likely to become locally extinct. Small populations, however, may experience more rapid evolution due to reduced genetic inertia. This may be seen in the founder effect, in which a few individuals begin a new population leading to allopatric speciation.

Without knowing the probability distribution of extrinsic factors causing large-scale mass extinctions, it is difficult to extrapolate from the fossil record if the species that went extinct in fact had small populations. In addition, it does not appear that evolution during the past several hundred years has reduced the likelihood of extinction events. There is no clear reduction in the apparent risk of such mass extinctions with time in the fossil record. Using the "all things being equal" assumption, the fossil record (**Figures 1** and **2**) suggests that, with a general trend toward increased diversity over time, the condition of having more taxa should lead to increases in the number of taxa lost due to disturbances. That is, if the probability of extinction of any particular species is constant, then we should see increased extinction rates as diversity has increased. However, as shown in **Figure 4**, the background extinction rate appears to have decreased with time. There remains the need to evaluate the risk of extinction by determining factors such as the absolute abundance of individuals within species and how this influences the extinction risk. For example, if increased diversity over time has led to smaller population sizes due to the ecological partitioning of resources, we may anticipate a trend toward increased extinction rates.

Experimental work on laboratory and field communities suggests that decreasing biodiversity also leads to a reduction in ecological time of ecosystem function, such as total productivity and carbon dioxide sequestration (Tilman and Downing, 1994; Naeem *et al.*, 1994). More recently Oswald and Steadman (2011) published data that suggested that the extinction of a cowbird (*Pandanaris convexa*) during the Pleistocene was associated with the extinction of large herbivores. These examples support the idea that positive feedback mechanisms may operate so that disturbed communities are more likely to lose species in addition to those initially lost due to decreased population size.

Models of the Influence of Evolution on Ecological Systems

The effect of evolution on the likelihood of species extinction and community stability represents a fundamental problem in the field of evolutionary ecology and is of great importance to our ability to conserve species over long time periods. Theoretical work suggests that the process of evolution influences population and community dynamics (Hartvigsen and Levin, 1997; Doebeli, 1996). Hartvigsen and Levin, using a spatially explicit individual-based model that accounts for genetic variability at the individual level, leading to evolution over time, show that the process of evolution in a model system can slow population dynamics, suggesting that the risk of species extinction declines due to reduced fluctuations resulting from demographic stochasticity. If this occurs in real systems, then we might expect a positive feedback mechanism to operate such that a population reduced in size is likely to experience a loss in genetic diversity. In addition to other known factors that increase the extinction risk of small populations, including inbreeding depression and social dysfunction (Allee effect), a loss in genetic diversity may increase population fluctuations, thus increasing the likelihood that such a species will go extinct.

The Future Record

Documenting Earth's Current Mass Extinction Event

The extent of the current extinction event will only be known in retrospect. However, there are patterns that are recognizable today. For example, there exists a good record of species that have gone extinct during the past 400 years (Smith *et al.*, 1993). A minimum of 485 animal species and 584 plant species became extinct during this period, with approximately half of these occurring during the twentieth century. This includes the loss of nearly 1% of all known bird species during the past 100 years, mostly due to human activity. This loss suggests an average species duration of about 10,000 years, which is two or three orders of magnitude shorter than has occurred on average in the fossil record (in other words, this represents an extinction rate two or three orders of magnitude greater than the background extinction rate in the fossil record). It is this rate of known extinction, compared to average extinctions in the fossil record, that suggests humans are the impetus for a mass extinction event.

More troubling is a recent analysis of extinction rates in amphibians which suggests that the current rate of loss dwarfs the background extinction rate observed in the fossil record (over the past several major extinction events). This rate may even be, since 1980, several orders of magnitude greater than seen previously (McCallum, 2007).

The Big Question: How Will the Process of Evolution Influence Biodiversity Dynamics?

Evolution generates new species in the aftermath of mass extinction. This is well supported by the fossil record and appears, in geological time, to be rapid. The current extinction event taking place on Earth, however, appears to be more rapid than has occurred previously (Smith *et al.*, 1993). Speciation will undoubtedly lead to the rise of new taxa but over time scales that are likely to be too long to have a noticeable effect in our lifetime. In addition, and more troubling, is the suggestion that

relatively young taxa are likely to be more susceptible to extinction due to the smaller population sizes and the relatively limited geographic ranges (Boyanjian, 1991).

The other role that evolution plays in biodiversity is its ability to influence the stability of communities. Little is known about this role of evolution. Empirical data for plant communities suggest that stability is correlated with biodiversity, such that more diverse communities are more resilient to disturbance. Natural communities are often composed of genetically variable individuals, which provide the foundation that enables species to evolve in response to biotic and abiotic factors over time. Also, communities with more species are in general more diverse genetically. If the time frame of environmental change can be matched by selection operating on genetic diversity, communities may resist disturbances by adjusting to changes over time. The analogy of species wandering over an "adaptive landscape" (sensu Sewall Wright) produces an image of communities functioning as complex adaptive systems. Ecosystems that function as complex adaptive systems may resist environmental fluctuations. If, however, environmental disturbances are abrupt, communities may either not be able to react quickly enough or suffer overloads that essentially shake species out of the community. There is evidence that communities harbor redundancy at the species level, but we have not yet determined patterns that identify which species are necessarily important or unimportant for community function. Therefore, the loss of species from communities is likely to decrease, if only slowly, the stability of communities.

Theoretical work suggests that the process of evolution may increase the resilience of communities to disturbance. This area of inquiry is in its infancy and will likely become an important and rapidly developed discipline in the following decades.

Conclusion

Biodiversity, measured as the number of taxa, has increased on Earth over time. This pattern appears to follow either an exponential or a logistic growth pattern of taxa during the past 540 My. The increase in taxa, however, has been discontinuous due to extinction events. There appears to be a probability distribution of extinctions with small background extinctions occurring frequently and large catastrophic mass extinctions occurring rarely. Biodiversity has repeatedly and rapidly rebounded following such extinction events.

Biodiversity currently appears to be near a saturation level, based on the observation that most species that are introduced to areas fail to successfully become established. However, the pattern of increase in the number of taxa (e.g., families) over time does not appear to be at or near an asymptote (**Figure 2**). The extreme alteration of the earth's biosphere by humans is likely to lead to a mass extinction event, currently on the order of 10–100 times the background extinction rate observed in the fossil record. In areas where habitats are disturbed rather than destroyed, an explosive radiation of new species may occur, manifested, however, only over geological time.

The process of evolution may help stabilize communities. However, current human-induced worldwide disturbances are likely to lead to reduced population sizes and a resulting loss in genetic diversity, ultimately increasing the likelihood that more species will go extinct.

Appendix

List of Courses

1. Principles of Ecology
2. Biodiversity
3. Evolution
4. General Biology

See also: Archaea, Origin of. Carrying Capacity, Concept of. Diversity, Community/Regional Level. Fossil Record. Limits to Biodiversity (Species Packing). Mass Extinctions, Notable Examples of. North America, Patterns of Biodiversity in

References

Benton MJ (1995) Diversification and extinction in the history of life. *Science* 268: 52–58.

Boyanjian GE (1991) Taxon age and selectivity of extinction. *Paleobiology* 17: 49–57.

Courtillot V and Gaudemer Y (1996) Effects of mass extinctions on biodiversity. *Nature* 381: 146–148.

Doebeli M (1996) Quantitative genetics and population dynamics. *Evolution* 50: 532–546.

Hartvigsen G and Levin S (1997) Evolution and spatial structure interact to influence plant–herbivore population and community dynamics. *Proceedings of the Royal Society of London B Biological Sciences* 264: 1677–1685.

McCallum ML (2007) Amphibian decline or extinction? Current declines dwarf background extinction rate. *Journal of Herpetology* 41: 483–491.

Naeem S, Thompson LJ, Lawler SP, Lawton JH, and Woodfin RM (1994) Declining biodiversity can alter the performance of ecosystems. *Nature* 368: 734–737.

Oswald JA and Steadman DW (2011) Late Pleistocene passerine birds from Sonora, Mexico. *Palaeogeography, Palaeoclimatology, Palaeoecology* 301: 56–63.

Pavé A, Hervé J-C, and Schmidt-Lainé C (2002) Mass extinctions, biodiversity explosions and ecological niches. *Comptes Rendus Biologies* 325: 755–765.

Raup DM (1994) The role of extinction in evolution. *Proceedings of the National Academy of Sciences of the United States of America* 91: 6758–6763.

Raup DM and Sepkoski Jr. JJ (1982) Mass extinctions in the marine fossil record. *Science* 215: 1501–1503.

Sepkoski Jr. JJ (1991) Diversity in the Phanerozoic oceans: A partisan review. In: Dudley EC (ed.) *The Unity of Evolutionary Biology*, pp. 210–236. Portland, OR: Dioscorides Press.

Smith FDM, May RM, Pellew R, Johnson TH, and Walker KS (1993) Estimating extinction rates. *Nature* 364: 494–496.

Thompson JN (1994) *The Coevolutionary Process*. Chicago: University of Chicago Press.

Tilman D and Downing JA (1994) Biodiversity and stability in grasslands. *Nature* 367: 363–365.

Biodiversity-Friendly Farming

Joern Fischer, Claire Brittain, and Alexandra-Maria Klein, Leuphana University, Lueneburg, Germany

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity The diversity of genes, species, communities, and ecosystems.
Connectivity The connectedness of habitat, land cover, or ecological processes from one location to another, or throughout an entire landscape.
Countryside biogeography A field aimed at understanding the diversity, distribution, and conservation of species in agricultural and other human-modified landscapes.
Ecosystem disservices Functions or processes provided by ecosystems with a negative impact on people.
Ecosystem functions Functions or processes provided by ecosystems.
Ecosystem services Functions, processes, and resources provided by ecosystems with benefits for people.
Holistic resource management A philosophy and practice of livestock grazing management, used by a minority of ranchers in Africa, North America, and Australia, attempting to integrate economic, social, and ecological goals.
Land sparing The spatial separation of biodiversity conservation and agriculture.
Landscape A human-defined area, typically ranging in size from approximately 1 to 1000 km^2.
Landscape heterogeneity The variability in land cover types within a given landscape.
Matrix The dominant and most extensive patch type in a landscape, which exerts a major influence on ecosystem processes.
Native vegetation Plant species growing naturally in a certain region.
Patch A relatively homogenous area within a landscape that differs markedly from its surroundings.
Seminatural vegetation Near-natural vegetation associations that provide suitable habitat for a large number of native species.
Structural complexity The complexity of the vertical architecture of vegetation, including a ground layer, shrub layer, and tree layer(s).

Article Outline

As the global population continues to increase, ever more land is used for agricultural production. What does this mean for biodiversity conservation? Agricultural land use constitutes one of the greatest threats to biodiversity. A key challenge is to find ways in which agricultural land use and biodiversity conservation can be integrated. A rapidly growing body of research has demonstrated a variety of generic measures that typically help to conserve biodiversity in agricultural landscapes. Such measures include local-scale activities that farmers can undertake and landscape-scale measures which can be encouraged through regional coordination and appropriate incentives. Drawing on examples from around the world, this article provides an up-to-date overview of current knowledge on biodiversity-friendly farming.

Introduction

In 1700, nearly half of the global land surface was free of human settlements or agricultural land use, with the remainder covered by seminatural vegetation, including small-scale agriculture and human settlements (Ellis *et al.*, 2010). Today, the vast majority of Earth's surface has been significantly influenced by human activities. Estimates are that only 25% of land remains in largely natural condition, and 40% is used for the production of crops and livestock grazing (Ellis *et al.*, 2010; Asner *et al.*, 2004; Foley *et al.*, 2005). Together, cropping and grazing thus account for a vast amount of "farmland." Globally, farmland use is expected to further increase in extent and intensity, with demand for food predicted to double from 2000 levels by 2050 (Tilman *et al.*, 2002).

Traditionally, biodiversity conservation has focused on protected areas (Diamond, 1975; Margules and Pressey, 2000). Although a system of dedicated nature reserves is an important backbone to biodiversity conservation, it has become increasingly clear that reserves alone will be inadequate for biodiversity conservation (Lindenmayer and Franklin, 2002; Tscharntke *et al.*, 2005). Presently, approximately 12% of global land is located in protected areas (Hoekstra *et al.*, 2005), or in other words, well over 80% of land is not formally protected. What happens to this land will shape the future of terrestrial biodiversity to a large extent. Reserves alone are simply too small, not sufficiently connected, and not safe enough from anthropogenic threats such as climate change to provide a viable safeguard for biodiversity conservation on their own (Daily, 1999, 2001; Lindenmayer and Fischer, 2006).

Given the need to look beyond protected areas, there has been a surge of interest in the conservation of farmland biodiversity. The subdiscipline of countryside biogeography seeks to understand patterns of species distribution in human-dominated landscapes, including generic relationships between the characteristics of land use, species composition, and species traits (Daily, 1999; Daily *et al.*, 2003; Horner-Devine *et al.*, 2003; Mayfield, *et al.*, 2005). A key motivation for this subdiscipline is the realization that, especially in the tropics,

there is great potential to integrate biodiversity conservation and agricultural production (Steffan-Dewenter *et al.*, 2007). Not tapping into this potential would mean losing some of the most important remaining opportunities for biodiversity conservation.

A second factor sparking increasing interest in the conservation of farmland biodiversity is that a myriad of species declines has been documented in farmland in a wide range of nations (Attwood, *et al.*, 2009; Brown *et al.*, 2008; Conrad *et al.*, 2006; Kleijn *et al.*, 2009; Murphy, 2003). As a result, government initiatives in many countries now try to foster biodiversity conservation on farmland, including through agri-environment schemes in Europe (Kleijn and Sutherland, 2003; Vickery, *et al.*, 2004; Whittingham, 2007), stewardship payments in Australia (Fischer *et al.*, 2010), or the Conservation Reserve Program in the US (Dunn *et al.*, 1993; Ribaudo *et al.*, 1990). Partly in response to biodiversity decline, there is a growing interest in farming practices trying to cater for both biodiversity conservation and commodity production. Examples of such practices include organic agriculture (Batáry *et al.*, in press; Badgley *et al.*, 2007; Hole *et al.*, 2005; Rundlöf *et al.* 2010), integrated pest management (Fitt *et al.*, 2009; Kogan, 1998), or livestock grazing according to the principles of "holistic resource management" (Stinner, *et al.*, 1997; Savory and Butterfield, 1999).

In this article, the authors provide an overview of generic measures that can be used to make farming more biodiversity-friendly. Although the authors provide a range of specific examples, their goal is to give an overview, rather than offer prescriptive advice for any particular farming system. The article is divided into three sections. The first section outlines a range of local-scale measures that typically benefit farmland biodiversity conservation. The second section documents some existing criticisms of measures to integrate biodiversity conservation and agricultural commodity production. As part of this section, the authors critically appraise a putative alternative solution – namely "land sparing," which is the spatial separation of biodiversity conservation and agriculture. The third and final section discusses why it is important to move from a local-scale focus to a landscape focus that considers whole farmland mosaics.

Local Measures for Biodiversity-Friendly Farming

Several management actions at the local scale can improve the friendliness of farming operations for biodiversity. The local scale is a useful starting point in many cases, because it is the scale at which individual farmers make management decisions. Most of the management actions discussed in this section target a combination of patterns and processes – for example, by changing a certain landscape pattern (e.g., the retention of native vegetation), certain ecosystem functions are improved (e.g., pollination), which in turn benefits biodiversity as a whole.

Retaining Native Vegetation Adjacent to Farmed Areas

Although some species can persist in intensively farmed areas, many more are able to persist in native or seminatural vegetation. It follows that retaining such vegetation in areas within or adjacent to farmland benefits farmland biodiversity as a whole. Examples of the benefits of native vegetation for farmland biodiversity abound, and come from all over the world. For example, native bee populations in a range of different ecosystems typically decrease with distance from native or seminatural vegetation (Ricketts *et al.*, 2008) (**Figure 1**) (**Box 1**). Similarly, many bird species depend on patches of

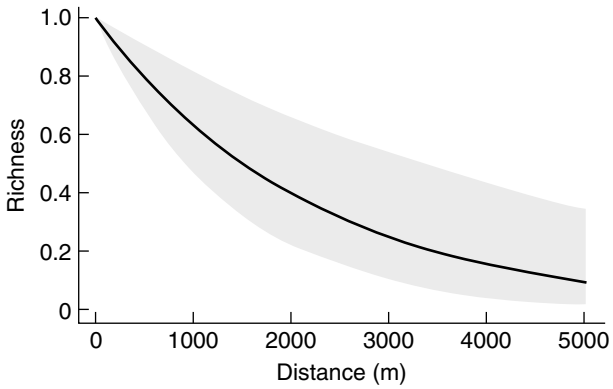


Figure 1 Decay in the species richness of pollinators with increasing distance from native or seminatural vegetation, based on a meta-analysis including 23 studies with 16 crops on five continents. Reproduced from Ricketts TH, Regetz J, Steffan-Dewenter I, *et al.* (2008) Landscape effects on crop pollination services: Are there general patterns? *Ecology Letters* 11: 499–515, with permission from Wiley.

Box 1 The value of native and seminatural vegetation for pollinators

Pollinator richness and the rate of flower visitation in a crop generally are affected positively by the amount of native vegetation in the surrounding landscape (Morandin and Winston, 2006), and affected negatively by distance to the nearest patch of such vegetation (**Figure 1**; for a review see Ricketts *et al.*, 2008). Bees require pollen and nectar, and native vegetation can provide foraging resources for bees at times when there is no crop in flower. Bees have a variety of nesting strategies, for example, some nest in the ground and require undisturbed bare soil, whereas others nest in hollow or pithy stems (Michener, 2000). A range of materials may be used in nest construction including mud, pebbles, resin and plant material. Areas of native vegetation can provide nesting sites and building materials for bees, which are essential for their reproduction.

In intensively used agricultural landscapes there may be little remaining natural vegetation, and this has implications for how to best implement biodiversity-friendly farming. In Californian almond orchards, for example, flower visitation by non-*Apis* insects was greater in orchards with more than 30% natural vegetation within a 1-km radius than in isolated orchards with less than 5% natural vegetation (**Figure 2**). However, while native vegetation clearly benefits bees in this region, many almond orchards occur in locations with little remaining natural vegetation nearby. In such cases, retaining or creating a seminatural strip of vegetation near the orchard's edge can be beneficial for pollinators. Orchards with such strips had non-*Apis* insects visiting flowers at levels intermediate between those in isolated orchards and those with high surrounding natural vegetation (**Figure 2**). Hence, while large areas of native vegetation are ideal for pollinator diversity, retaining or enhancing seminatural vegetation can also be beneficial.

native vegetation. In recently cleared landscapes, such as the wheat-sheep belt of eastern Australia, native vegetation is critically important for a wide range of bird species, and land clearing has caused the decline of many woodland birds (Lindenmayer *et al.*, 2010; Ford *et al.*, 2009; Recher, 1999).

In landscapes with a long history of human use, such as agricultural landscapes in Europe, it is often seminatural vegetation (rather than strictly native vegetation) that provides biodiversity benefits. Hedgerows and other vegetated field margins, for example, were originally established to delineate field boundaries (Sklenicka *et al.*, 2009), but now fulfill an important function in providing biodiversity benefits (**Figure 2**), including for small mammals (Wegner and Merriam, 1979), butterflies (Schmitt and Rakosy, 2007; Harvey *et al.*, 2005), and birds (Wegner and Merriam, 1979; Harvey *et al.*, 2005).

Because a substantial amount of literature on farmland biodiversity and the umbrella concept of "countryside biogeography" has come from recently cleared landscapes, one concern might be that in fact, even with the retention of patches of native vegetation, these landscapes will still continue to lose species for many decades into the future (due to an "extinction debt"; see for example, Robinson (1999), Kuussaari *et al.* (2009)). While this is likely in some landscapes (e.g., southeastern Australia; see Fischer *et al.* (2010)), there are also examples of agricultural landscapes that have long maintained a rich diversity of species, despite ongoing agricultural land use. Examples include the highly diversified agricultural landscapes in the Western Ghats of India (Ranganathan *et al.*, 2008), or the foothills of the Carpathians in Romania (Schmitt and Rakosy, 2007). The landscape-scale

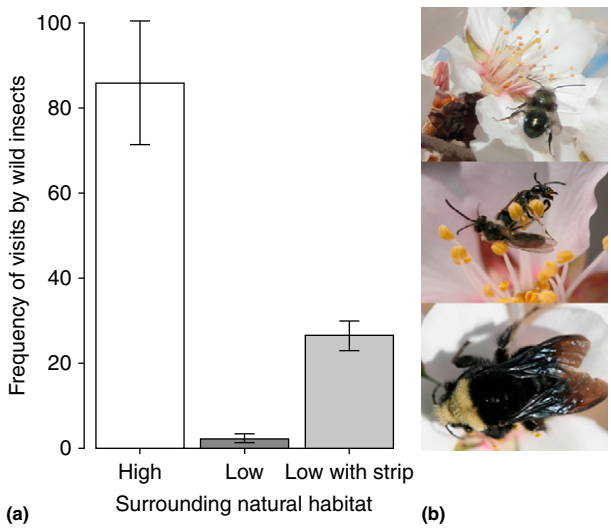


Figure 2 (a) The relationship between the frequency of visits by wild insects to almond flowers in California orchards (mean per orchard ± SE) and the amount of surrounding natural habitat. The amount of surrounding natural habitat at a 1-km radius was classed as: low (≤5%, $n = 5$), low with a strip of seminatural vegetation ($n = 5$), and high (>30%, $n = 5$). Observations in each orchard were undertaken on three separate days in 2008. (b) Non-*Apis* visitors to almond flowers included (from top to bottom): the blue orchard bee *Osmia lignaria*, sweat bees (family Halictidae), and bumble bees such as *Bombus vonesenskii*. Unpublished data and photos by Claire Brittain and Alexandra-Maria Klein.

diversity (or landscape heterogeneity) of these areas is likely to be an important safeguard for their rich diversity of species – the authors discuss the importance of landscape heterogeneity for biodiversity below (*see* Biodiversity-Friendly Farming at the Landscape Scale).

A simple rule of thumb is that more natural or seminatural vegetation leads to more biodiversity. The species-area relationship is one of the few established principles (or laws) in ecology, and applies broadly across many different ecosystems (MacArthur and Wilson, 1967; McGuinness, 1984; Rosenzweig, 1995). However, there will also be species that persist in farmland itself for all or part of their life cycle. Farmland provides critically important resources for such species, such as pollen and nectar for pollinators from mass flowering crops (Winfree *et al.*, 2007). This means that retaining native vegetation should not be the only strategy – what happens on the farmland itself is equally important (Pereira and Daily, 2006).

Maintaining Structural Complexity in Farmed Areas

While natural areas adjacent to farmland provide an important source pool for many species, farmland itself also can provide habitat for some species. Specifically, farmed areas are often more hospitable if their structure mimics that of native vegetation (Fischer *et al.*, 2006). In ecosystems originally dominated by woody vegetation, structural similarity to native vegetation typically equates to structural complexity – structurally complex agricultural areas, generally speaking, thus are more beneficial for farmland biodiversity than structurally simplified areas.

There are many examples of this relationship. One of the best known ones is that of "shade coffee" plantations, especially when compared with "sun coffee" plantations (Rappole *et al.*, 2003). Shade coffee plantations that incorporate old native trees are structurally much more complex than sun coffee plantations, and a wide range of species can find suitable habitat in shade coffee (**Box 2**) (**Figures 3** and **4**). Arecanut plantings in India (and possibly other parts of Asia), too, can harbor a surprisingly rich bird fauna, partly because they are structurally similar to native forest patches (Ranganathan *et al.*, 2008). A similar situation also exists for reptiles in Australian farming landscapes. Provided the presence of native tussock grasses, fallen branches, and patches of leaf litter, a wide range of reptile species can use areas that are grazed by livestock (Fischer *et al.*, 2005).

Notably, not all elements of structural complexity are equally important. Some elements are disproportionately important relative to their size or footprint in the landscape. Examples of such "keystone structures" are ephemeral water bodies in north-eastern German fields (Tews *et al.*, 2004), or tree hollows in Australian forests (Gibbons and Lindenmayer, 2002). Both are used by a large number of species, which fundamentally depend on the existence of these structures.

One increasingly well-known example of keystone structures in farming landscapes is that of scattered trees (Manning *et al.*, 2006). Such trees are common in agricultural landscapes around the world, including commercially grazed oak savannas in Oregon (DeMars *et al.*, 2010), wood pastures in Europe (Bergmeier *et al.*, 2010), and agricultural landscapes in

Box 2 The value of structural complexity for biodiversity: The base of coffee plantations

Coffee is produced in tropical areas around the globe. It is grown either in monocultures exposed to full sun, or under a layer of shade trees, which can exhibit different levels of structural complexity. Moguel and Toledo (1999) described the different forms of structural complexity in coffee systems of Mexico (**Figure 3**). Similar structures can be found in other parts of the world such as South-East Asia (Klein *et al.*, 2002, **Figure 4**).

Many studies have highlighted the potential of complex shade coffee to provide habitat for a diversity of plants (Hylander and Nemomissa, 2008), arthropods (Nestel *et al.*, 1993; Perfecto and Snelling, 1995), amphibians (Pineda *et al.*, 2005), birds (Tejeda-Cruz and Sutherland, 2004), and mammals (Williams-Guillén *et al.*, 2006). In addition, shade coffee may also "soften" the landscape matrix as a whole (Perfecto and Vandermeer, 2002; Pineda *et al.*, 2005) and may play a buffering role for natural rainforests (Moguel and Toledo, 1999; Tejeda-Cruz and Sutherland, 2004). Producing coffee under a complex array of shade trees has become a widely applied conservation measure, and has led various organizations to develop "biodiversity-friendly," "wildlife-friendly," or "bird-friendly" production systems, some of which are independently certified (Rappole *et al.*, 2003; Rainforest Alliance, 2009).

Although complex shade coffee systems provide habitat to a relatively large number of species, rare forest specialists often cannot be conserved in such systems (Rappole, *et al.*, 2003; Guevara, 2005). Saving patches of native vegetation ("land sparing"; see main text) therefore is an important complementary strategy to enhance biodiversity in landscapes used for coffee production. Certification programs for biodiversity-friendly coffee recognize the unique value of natural tropical rainforests for rare species. For this reason, certification is only being granted for areas that were previously used more intensively (Greenberg, 2001; Sustainable Agriculture Network, 2009) – this avoids the conversion of natural rainforest to (supposedly biodiversity-friendly) coffee production. Such conversion is a real threat to rainforests in some poor countries (Tejeda-Cruz *et al.*, 2010). To counteract this threat, conservation initiatives need to understand the socioeconomic drivers of land use change, and need to monitor the actual land use changes and conservation outcomes. Biodiversity conservation in landscapes used for coffee production thus requires a mix of strategies, including measures to ensure biodiversity-friendly production methods, but also measures to protect remnant patches of natural rainforest.

the wheat-sheep belt of eastern Australia (Ozolins, *et al.*, 2001; Maron and Fitzsimons, 2007). Scattered trees are disproportionately important because the marginal contribution of a single, semi-isolated tree to local biodiversity is many times larger than that of a single tree which is part of a large patch. In Australia, Fischer *et al.* (2010) documented that in a given 2 ha area used for livestock grazing, bat activity was ten times higher near a single tree than in an area without trees; and ten times higher again when there was a small clump of three to five trees (**Figure 5**). After that, bat activity barely increased with the presence of additional trees – the first few trees thus were most important.

Minimizing Chemical Use

The potentially disastrous consequences for biodiversity of chemicals in the environment have long been known. Most famously, in the 1960s, Rachel Carson warned against the then widespread use of the pesticide dichlordiphenyltrichlorethan because of its far-reaching biological consequences throughout the food web (Carson, 1962). Since then, and partly triggered by Carson's efforts, a range of restrictions have been placed on the use of chemicals in many jurisdictions. But still, both the use of inorganic fertilizers and pesticides has increased steadily since the 1960s. While this has contributed to increasing agricultural productivity, it poses a range of threats to biodiversity (Tilman *et al.*, 2002), especially in poor tropical countries where many agro-chemicals are used that are banned in other wealthier countries (Wanger *et al.*, 2010).

Many ecosystems around the world have been negatively affected by fertilizer application. A useful conceptual model of the effects of inorganic fertilizers on grassland systems was developed by McIntyre and Lavorel (2007). They recognized several different states in grassland systems, from species-rich native systems to species-poor fertilized (or even sown) pastures. For transitions between the various states, the two most important drivers recognized were livestock grazing pressure and fertilizer use. Especially in systems that are naturally nutrient-poor, fertilizer application can be disastrous. For example, fertilizer application would permanently destroy traditional hay meadows in Romania (Akeroyd and Page, 2007), and fertilizer application is one of the main reasons why trees are no longer naturally regenerating over vast areas of Australia's wheat-sheep belt (Fischer *et al.*, 2009).

The use of pesticides is similarly problematic, with global pesticide use predicted to further grow (Tilman *et al.*, 2002; Wanger *et al.*, 2010). The impact of pesticides on biodiversity can be difficult to tease apart from other factors associated with agricultural intensification. However, there is evidence of widespread negative impacts on biodiversity (Geiger *et al.*, 2010; Relyea, 2005; Pimentel, *et al.*, 1992). Nontarget impacts are of particular concern. Such impacts occur when species not targeted by a particular chemical are negatively affected. Nontarget impacts recently have been implied as part of the reason for widely documented bee declines (Potts *et al.*, 2010; Oldroyd, 2007). Organisms that live in agricultural landscapes, such as honey bees, can be exposed to many different pesticides over long periods (Frazier *et al.*, 2008; Chauzat and Faucon, 2007), with unknown cumulative and synergistic effects. To reduce the pressure on wildlife from pesticides, several management measures can be effective. Such measures range from conservation headlands where pesticides are reduced along the edge of the crop (Longley *et al.*, 1997), to field margins that are sown with seed mixes to provide resources for wildlife (such as nectar for bees; see **Box 1**). In addition, patches of native or seminatural vegetation including hedgerows can protect wildlife from pesticide exposure (Otto *et al.*, 2009).

Finally, not just fertilizer and pesticides, but also other chemicals associated with industrial farming can cause unforeseen environmental problems. A high-profile example involved several species of vultures (*Gyps* spp.) on the Indian subcontinent (Oaks *et al.*, 2004). Rapid population declines of vultures had been documented starting in the 1990s. In 2004, a link was reported between these population declines and a veterinary drug, namely the anti-inflammatory drug diclofena.

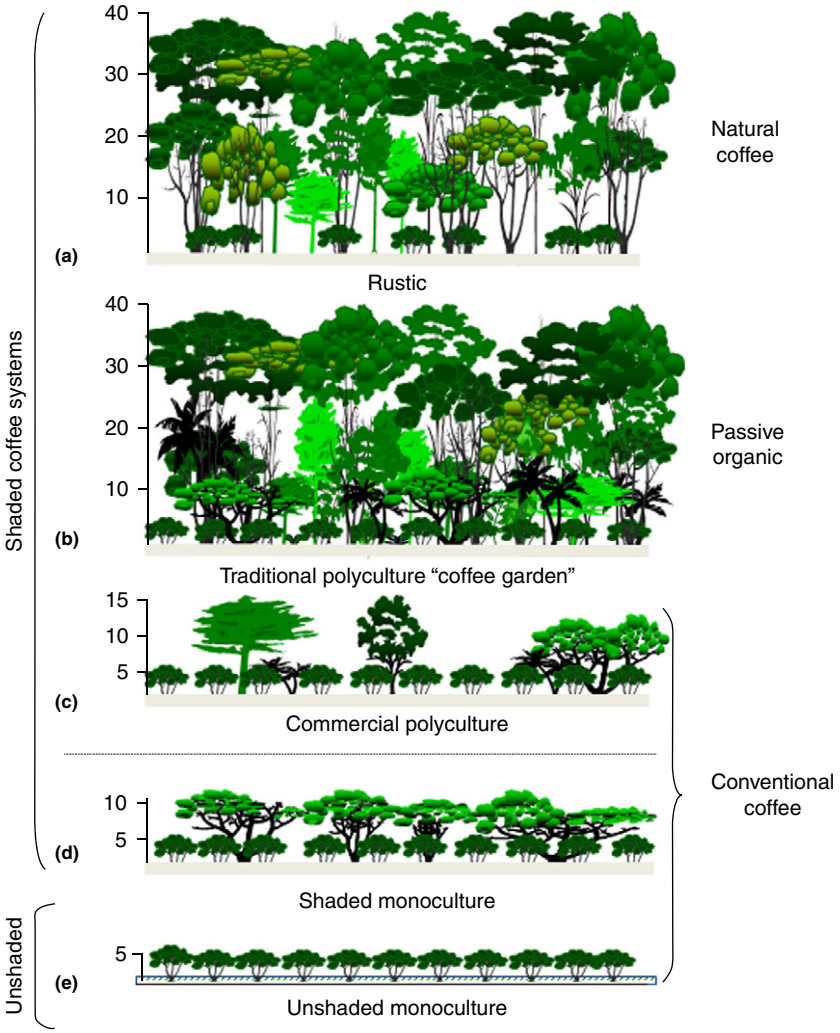


Figure 3 Schematic representation of increasing structural complexity in coffee-growing systems. From top to bottom (a–e) structural complexity decreases with the transition from shaded polyculture to unshaded monoculture. Reproduced from Moguel P and Toledo VM (1999) Biodiversity conservation in traditional coffee systems of Mexico. *Conservation Biology* 13: 11–21, with permission from Wiley.

This readily available drug was widely used to treat domestic livestock. However, the primary food source of vultures in this area was dead domestic livestock. Residues of the drug in dead livestock were highly significantly correlated with renal failure in the vultures, suggesting that the drug was the likely cause of unintended ecological consequences throughout the food chain (Oaks *et al.*, 2004).

Given the varied negative and difficult to predict consequences of chemical use for biodiversity, it follows that (generally speaking) farming operations are more biodiversity-friendly if chemical use is minimized. This is strongly supported through a range of papers on organic farming, where chemical use is absent or severely restricted. Recent reviews consistently show that on-average biodiversity increases in organic farming systems when compared with conventional systems (Batáry *et al.*, 2011; Hole *et al.*, 2005; Bengtsson, *et al.*, 2005); with the benefits often depending on the surrounding landscape context (*see* Biodiversity-Friendly Farming at the Landscape Scale).

The Timing of Disturbance Events

In addition to the general principles of maintaining an appropriate structure and minimizing chemical use, some location-specific farming practices can also be beneficial for biodiversity. Such practices often revolve around the balance and timing between periods of disturbance and periods of rest for a particular parcel of land. Examples exist both for livestock grazing and crop production.

For livestock grazing, two variables are important: the average stocking rate and the degree to which livestock stay in one place or move around the landscape. The stocking rate is of obvious importance because a higher number of livestock per unit area means greater pressure on the ecosystem. Frequently observed ecological consequences of overly high stocking rates include an impoverished flora and insect fauna, soil compaction, patches of bare soil, and subsequent risk of soil erosion (Oliver *et al.*, 2005; Yates *et al.*, 2000; Greenwood and McKenzie, 2001). By contrast, the relationship between



Figure 4 (a) Structural complexity in a traditional forest polyculture with coffee in Central Sulawesi, Indonesia; (b) coffee harvest in a traditional coffee polyculture in Manabi, Ecuador; and (c) ripe coffee berries of the polyculture in Ecuador.

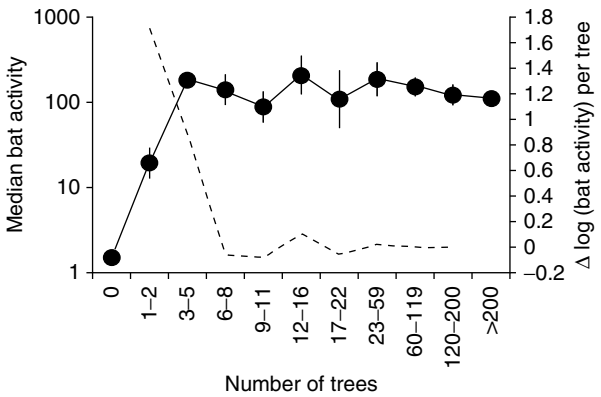


Figure 5 The calling activity of bats, as measured at 63 individual 2-ha sites in a livestock grazing landscape in eastern Australia. Solid lines show means (+/− standard error; left vertical axis); dotted lines show mean marginal change per additional tree (right vertical axis). The figure shows that the first few trees were particularly important in attracting bats to an area, suggesting that scattered trees are "keystone structures" in this example. Reproduced from Fischer J, Stott J, and Law BS (2010) The disproportionate value of scattered trees. *Biological Conservation* 143: 1564–1567.

livestock movement and biodiversity is less well understood. Many traditional livestock grazing systems that have also maintained high biodiversity are characterized by substantial livestock movement. One example is traditional sheep grazing in Transylvania, Romania (Huband, 2007). A single shepherd typically is responsible for many dozen sheep, which are owned by many different villagers. The shepherd takes these sheep out into pastures surrounding the village and returns to the village only approximately once a week, before taking the sheep out again to a different set of pastures (Akeroyd, 2006). During their time in the pastures, the sheep are in constant movement. Grazing intensity is thus very high during the time a flock of sheep is present at a given location – but the same location is unlikely to be grazed again for many days or weeks.

A combination of high intensity grazing with prolonged rest periods between grazing events is also at the heart of rotational grazing. Rotational grazing is practiced, for example, on some ranches (sometimes also called "farms") in the US, Australia, and parts of southern Africa. Advocates of rotational grazing believe it fundamentally benefits ecosystems because of the long rest periods between grazing events. The rationale is that short bouts of intense disturbance leave an ecosystem less damaged in the long term than constant low-grade pressure (Savory and Butterfield, 1999). While this rationale is appealing, empirical evidence on demonstrable benefits of rotational grazing is more difficult to come by (Briske *et al.*, 2008). In southeastern Australia, Fischer *et al.* (2009) demonstrated that native tree regeneration was more likely under rotational grazing than under continuous grazing. However, many examples exist where rotational grazing has not lived up to its promise. For example, rotational grazing systems in Arizona (USA) failed to maintain ground cover in drought conditions, with negative consequences for grassland birds (Bock and Bock, 1999). The usefulness of this practice therefore appears to depend on the location where it is practiced, and on the specific details of its implementation.

For cropping systems, too, the timing of management actions is important. In some parts of the tropics, agricultural systems are shifting toward shorter fallow periods (Dalle and De Bois, 2006), and European countries such as England have moved away from a mixture of spring and winter-sown crops in lowland agricultural land, toward a more simplified landscape of winter-sown cereals (Henderson *et al.*, 2009). These changes pose a potential threat to biodiversity because recovery time between disturbance events is reduced. Fallow periods associated with crop rotation can be beneficial for soil

microbial diversity (Lupwayi *et al.*, 1998), soil fertility, and natural pest management, thereby reducing the need for pesticides and fertilizers (Tilman *et al.*, 2002). For example, in North Carolina, Maize fields under 2- and 4-year rotations had significantly more pest predators and fewer corn root infestations than fields not in rotation (Brust and King, 1994). There is less information available on the impacts of different crop rotations on above-ground biodiversity. In Iowa, the richness of carabid beetles was higher in a 4-year (corn/soybean/triticale-alfalfa/alfalfa) crop rotation, than a 2-year (corn/soybean) crop rotation, most likely due to the inclusion of alfalfa (O'Rourke *et al.*, 2008). Similarly, in the UK, the populations of several species of birds increased over 5 years, following the introduction of a mixed rotation into a winter wheat system (Henderson *et al.*, 2009). While the general notion of crop rotation is beneficial for biodiversity in all of these examples, there are likely to be differential responses by various taxa, depending on the extent to which different species rely on particular resources and the timing of their life stages relative to crop management activities.

Maintaining Key Ecosystem Functions: Pollination and Pest Control

The processes of pollination and pest control are fundamentally important for biodiversity-friendly agricultural production. Conserving natural habitat, establishing conservation headlands, and sowing native flowers in field margins are all actions that can benefit pollinators on farmland (**Box 1**), along with reducing pesticide use or farming organically (*see* Minimizing Chemical Use). Maintaining and promoting pollination services in agricultural landscapes will require the conservation or establishment of resources necessary for populations of pollinators (Ricketts *et al.*, 2008). Such resources include food and nesting habitat (**Box 1**). Conserving and encouraging native pollinators can bring benefits to farmers by reducing dependence on a single species for pollination, namely the honey bee (*Apis mellifera*). In the watermelon, a crop whose production is largely dependent on pollination by bees, organic fields close to natural habitat received sufficient pollination services for fruit set just from wild bees (Kremen *et al.*, 2002). Moving away from reliance on a single species for pollination to a more diverse pollinator community is desirable as a safeguard against unforeseen environmental changes, especially in the context of widespread reductions in honey bee colonies (Potts *et al.*, 2010; Ellis *et al.*, 2010).

Biodiversity-friendly farming can also bring benefits through natural pest control. Organic farming (Letourneau and Goldstein, 2001) and noncrop habitat such as field margins (Bianchi *et al.*, 2006) can increase the number of predators for crop pests, thereby reducing the need for pesticide applications. Greater species evenness and natural pest control have been found in organic compared to conventional farms and in field enclosures where plant biomass was greater (Crowder *et al.*, 2010). Rotational and mixed cropping can also increase natural pest control on farms (Bianchi, *et al.*, 2006; Brust and King, 1994). Natural pest control may reduce the cost of pesticide applications for the individual farmer in the short term and may help to reduce the long-term costs of widespread pesticide use (Pimentel, 2009). It also represents a more sustainable option in the long term, given the evolution of resistance of pests to commonly used pesticides (Brattsten *et al.*, 1986).

Criticisms of Biodiversity-Friendly Farming

Although many conservation biologists have passionately advocated biodiversity-friendly farming as a partial solution to the global biodiversity crisis, biodiversity-friendly farming has also been met with a range of criticisms. First, many governments have taken to the notion of biodiversity-friendly farming, and have established schemes that financially reward farmers for management actions that are deemed beneficial for biodiversity. However, reviews of the benefits of such schemes have found mixed results (Kleijn and Sutherland, 2003; Kleijn *et al.*, 2006; Knop *et al.*, 2006). Possible reasons for a lack of demonstrable benefits of some agri-environment schemes that support biodiversity-friendly practices include a lack of consideration of landscape context (Fahrig *et al.*, 2011; Whittingham, 2007) (*also see* Biodiversity-Friendly Farming at the Landscape Scale) and inadequate monitoring (Kleijn and Sutherland, 2003).

Second, some elements of biodiversity may be worth having on a farm from a conservation perspective, but in fact may cause substantial harm to crops. While the rhetoric of biodiversity-friendly farming is rife, in some ecosystems and for some crops, there are very real limits to the reality of its implementation. Monitoring programs typically consider only the positive aspects of biodiversity-friendly farming schemes, but do not consider disservices related to reduced productivity or higher production costs, resulting, for example, from wildlife destroying part of the harvest or acting as vectors for pathogen infections (Zhang *et al.*, 2007; Dunn, 2010).

Third, some conservation biologists have claimed that an exclusive focus on biodiversity-friendly farming is overly simplistic, and quite possibly detrimental to biodiversity conservation in some situations. Probably the most influential paper on this subject was published by Green *et al.* (2005). The main argument presented by these authors was based on the premise that biodiversity-friendly farming, because it is less intensive, typically results in lower yields per unit area. To meet the same demand for food, it then follows that a greater area is needed for biodiversity-friendly agriculture as opposed to intensive agriculture (see also Balmford *et al.*, 2005). Green *et al.* (2005) argued that in a scenario of intensive agriculture, the land not used for agriculture could be "spared" for nature. Depending on the response of biodiversity to increasingly intensive land uses, it is then possible that a series of relatively small areas "spared" from agriculture could support a greater diversity of species than vast areas of biodiversity-friendly farmland which incorporate no protected areas (Balmford, *et al.*, 2005; Green *et al.*, 2005).

The fundamental point of this argument is important – biodiversity-friendly farming must not be seen as a panacea for the problem of biodiversity decline in all ecosystems (Mattison and Norris, 2005). But there are many subtleties that suggest reality is in fact more complicated than a simple choice between "biodiversity-friendly farming" and "land sparing." Perhaps the most important criticism of the "land sparing" model is that some farming systems can produce

relatively high yields, while also being more biodiversity-friendly than industrial alternatives (Badgley *et al.*, 2007). Moreover, from the perspective of an individual farmer, high yields are not necessarily important to make a living. Some Australian cattle farmers, for example, have reported that biodiversity-friendly low-input systems have become increasingly attractive because input costs have risen – while these farmers now stock fewer cattle than in the past, their profits have in fact remained stable (J. Fischer, pers. obs.).

Intensive agriculture, typically assumed to be an unavoidable part of the notion of "sparing land" for nature, also involves a range of additional problems. Most importantly, the negative environmental effects of intensive agriculture are not constrained to the actual location that is being farmed. Rather, there is a range of "externalities," that is, negative and often diffuse environmental consequences that occur elsewhere and are not easily accounted for. Such externalities include nutrient runoff into water bodies (Olsson *et al.*, 2008) and chemical drift into nearby areas (Duncan *et al.*, 2008), but also less tangible contributions to global environmental degradation, for example, through carbon emissions resulting from intensive land use (e.g., Garnett, 2009).

In addition, the notion that some land can be used intensively but other land "spared" within the same region may be unrealistic. Typically, intensive land use in a given area attracts even more intensive land-use to that area (Fearnside, 2002; Matson and Vitousek, 2006). Intensification thus often begets more intensification, rather than encouraging the preservation of land for nature.

Many variables influence the extent to which an area is suitable for intensive agriculture. Less intensive, biodiversity-friendly farming is most easily implemented in areas that have low productivity or steep topography, are heterogeneous and divided into small parcels of land, and occur near markets for environmentally benign (e.g., organic) food (Fischer *et al.*, 2008). In reality, biodiversity-friendly farming and intensive farming with areas "spared" for nature are therefore not truly opposed options. Rather, they occur on a continuum of levels of intensity and spatial use of the environment, and their adoption is influenced by a range of variables (**Figure 6**, Fischer *et al.*, 2008). Because the world is a heterogeneous place, it is likely that there are some places where indeed, intensive agriculture might be preferable (Fischer *et al.*, 2009). However, in many areas, more biodiversity-friendly farming is possible – sometimes even without a loss in productivity (Stinner *et al.*, 1997). Notably, discussions about land sparing versus biodiversity-friendly farming are ongoing (Phalan *et al.*, 2011), and as yet, no broad consensus has been reached.

Biodiversity-Friendly Farming at the Landscape-Scale

Local decisions about how to farm have regional effects, and vice versa. This is true not only in a socioeconomic sense (e.g., intensification often begets more intensification; (*see* Criticisms of Biodiversity-Friendly Farming), but also in an ecological sense (e.g., through either negative or positive neighborhood effects). On this basis, ecologists are increasingly interested not only in the local effects of different types of farmland on biodiversity, but also in the effects of entire farmland mosaics. Two landscape properties are particularly interesting in this context: heterogeneity and connectivity.

Landscape Heterogeneity

An increasingly consistent theme emerging from studies on biodiversity-friendly farming is that, other things being equal, heterogeneous farming landscapes support more biodiversity than homogenous farming landscapes (Daily *et al.*, 2003; Benton *et al.*, 2003; Haslem and Bennett, 2008). Policies aiming to promote landscape heterogeneity therefore can benefit the biodiversity of farming landscapes.

In addition, heterogeneity at the landscape scale can influence species richness at the local scale (Weibull *et al.*, 2003; Clough *et al.*, 2005). Local variables and landscape variables can interact, and it is their combination that shapes biodiversity at any given point in a farming landscape. In an influential paper, Tscharntke *et al.* (2005) proposed a general hypothesis of how the local effect of farming practices might depend on landscape context. Based on preliminary evidence,

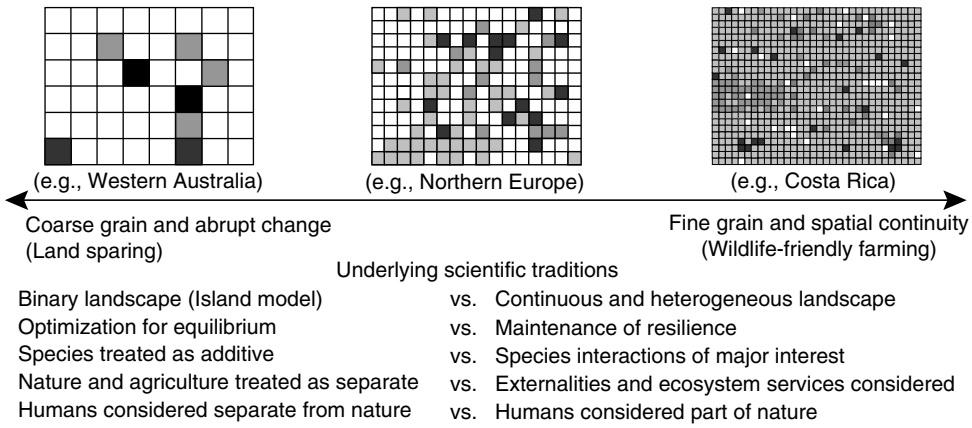


Figure 6 Wildlife-friendly farming (or biodiversity-friendly farming) and land sparing occur on a continuum. A finer spatial grain in land use tends to be correlated with more biodiversity-friendly practices. Many underlying scientific traditions influence whether in a given situation, land sparing or wildlife-friendly farming are considered preferable. Reproduced from Fischer J, Brosi B, Daily GC, *et al.* (2008) Should agricultural policies encourage land sparing or wildlife-friendly farming? *Frontiers in Ecology and the Environment* 6: 380–385, with permission from ESA.

they proposed that complex or heterogeneous landscapes supported high biodiversity across the landscape. But further, they suggested that although such landscapes are important for biodiversity conservation, local measures to further enhance biodiversity in specific locations were likely to be relatively ineffective in heterogeneous landscapes. By contrast, in simplified or homogenous landscapes, total biodiversity is likely to be a lot lower, but partly for this reason, local measures to enhance biodiversity at specific locations might be more effective in such landscapes.

Evidence in support of this conceptual model is increasing. For example, in Sweden, local organic management only increased species richness and abundance in field margins in homogeneous, rather than heterogeneous landscapes (Rundlöf and Smith, 2006; Rundlöf, *et al.*, 2008). Similarly, in Germany, bee diversity was consistently greater in organic than conventional wheat fields, but the difference in diversity between paired organic and conventional sites was greater in a homogeneous landscape context than in a heterogeneous context (Holzschuh *et al.*, 2007). Most recently, Batáry et al. (2011) undertook a meta-analysis of the local effects of organic farming on a range of taxonomic groups in both homogenous and heterogeneous landscapes. While they found partial evidence in support of the original proposition, the model did not universally hold across all taxonomic groups or types of farming systems. On this basis, evidence is strong that landscape heterogeneity increases biodiversity at both local and landscape scales; and it appears that often (but not always) the effectiveness of local management actions is contingent on the landscape context.

Landscape Connectivity

Connectivity is likely to enhance biodiversity in farming landscapes. Structural links between similar patches of vegetation (i.e., "structural connectivity") can help to facilitate the movements of individuals throughout the landscape (thereby providing "functional connectivity") (Uezu *et al.*, 2005). Such connectivity is important to enable the movements of individuals on a daily basis (e.g., foraging movements; Saunders, 1980), to facilitate the exchange of individuals between subpopulations on an occasional basis (Hanski and Simberloff, 1997), and to facilitate gradual, large-scale movements of species (e.g., poleward or upward migration of species in response to climate change; Hickling, *et al.*, 2006; Parmesan, 2006).

Connectivity in farming landscapes can be maintained via three complementary mechanisms. The first is managing the actual farmland matrix in a way that enhances its use by species (*see* Maintaining Structural Complexity in Farmed Areas). For example, Gascon and Lovejoy (1998) assessed the connectivity of cleared areas in the Amazon. They likened these areas to a filter – some species could move through cleared areas, whereas others could not. In this case, the "pore size" of the filter was partly determined by its vegetation structure, with more forest-like structure more likely to be used by forest birds. More generally, compared to a given species' core habitat, a structurally similar matrix is more likely to facilitate the movement of this species than a structurally dissimilar matrix (Donald and Evans, 2006; Fischer, *et al.*, 2006).

The second mechanism that can enhance connectivity is that of stepping stones. At the landscape scale, the best known example of such stepping stones is that of scattered trees and the connectivity they provide for birds (*see* Maintaining Structural Complexity in Farmed Areas). For example, in Mexico, Guevara and Laborde (1993) demonstrated that frugivorous birds preferentially moved through the landscape via areas where trees were scattered relatively densely. A similar effect was documented for woodland birds in Australia (Fischer and Lindenmayer, 2002), and radio-tracking studies in Costa Rica also have reported positive effects of scattered trees on connectivity (Sekercioglu *et al.*, 2007; Gillies and St. Clair, 2010).

Finally, the most widely known method to enhance connectivity is that of corridors. In agricultural landscapes, such corridors can be traditional hedgerows (Wegner and Merriam, 1979), live fences (Harvey *et al.*, 2005), or linear restoration plantings (Lindenmayer *et al.*, 2010). As with all other kinds of structural connectivity, the mere presence of a structural link does not guarantee that the target species will actually move through a given corridor. Many factors influence whether a corridor is actually used, including a given species' biology, the presence of predators, and the physical characteristics of the corridor (Lindenmayer and Fischer, 2006).

Conclusion

Biodiversity-friendly farming can be enhanced through a series of local management measures, which relate to the retention of native vegetation, the maintenance of structural complexity, the avoidance of chemicals, and the use of regionally-proven beneficial farming methods (many of which are related to the balance between disturbance events and periods of rest). In addition, local farmland biodiversity is influenced by its landscape context. Other things being equal, heterogeneous, or complex farmland mosaics support greater biodiversity than homogenous or simplified farming landscapes. Biodiversity-friendly farming thus is enhanced by maintaining native vegetation, encouraging heterogeneity and structural complexity, and avoiding the use of chemicals. While industrial farming may be appropriate in some situations, biodiversity-friendly farming offers opportunities for biodiversity conservation in many parts of the world.

Appendix

List of Courses

1. Conservation Biology

See also: Agriculture, Industrialized. Agriculture, Sustainable. Agriculture, Traditional. Agrobiodiversity. Biodiversity and Pest Control Services. Biofuels and Biodiversity: The Implications of Energy Sprawl. Biofuels and Biodiversity, Wildlife Habitat Restoration. Conservation Biology, Discipline of. Countryside Biogeography. Crop Mixtures and the Mechanisms of Overyielding. Economics of

Agrobiodiversity. Feeding the World and Protecting Biodiversity. Grassland Ecosystems. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Land-Use Issues. Landscape Corridors. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production. Loss of Biodiversity, Overview. Oil-Palm Plantations in the Context of Biodiversity Conservation. Pollinators, Role of. Soil Conservation

References

Akeroyd J (2006) *The Historic Countryside of the Saxon Villages of Southern Transylvania*. Saschiz: Fundatia Adept.

Akeroyd JR and Page N (2007) The Saxon villages of southern Transylvania: Conserving biodiversity in a historic landscape. In: Gafta D and Akeroyd J (eds.) *Nature Conservation: Concepts and Practice*. New York: Springer.

Asner GP, Elmore AJ, Olander LP, Martin RE, and Harris AT (2004) Grazing systems, ecosystem responses, and global change. *Annual Review of Environment and Resources* 29: 261–299.

Attwood SJ, Park SE, Maron M, *et al.* (2009) Declining birds in Australian agricultural landscapes may benefit from aspects of the European agri-environment model. *Biological Conservation* 142: 1981–1991.

Badgley C, Moghtader J, Quintero E, *et al.* (2007) Organic agriculture and the global food supply. *Renewable Agriculture and Food Systems* 22: 86–108.

Balmford A, Green RE, and Scharlemann JPW (2005) Sparing land for nature: Exploring the potential impact of changes in agricultural yield on the area needed for crop production. *Global Change Biology* 11: 1594–1605.

Batáry P, Baldi A, Kleijn D, and Tscharntke T (2011) Landscape-moderated biodiversity effects of agri-environmental management: A meta-analysis. *Proceedings of the Royal Society B: Biological Sciences Series B* 278: 1894–1902.

Bengtsson J, Ahnstrom J, and Weibull AC (2005) The effects of organic agriculture on biodiversity and abundance: A meta-analysis. *Journal of Applied Ecology* 42: 261–269.

Benton TG, Vickery JA, and Wilson JD (2003) Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology & Evolution* 18: 182–188.

Bergmeier E, Petermann J, and Schroeder E (2010) Geobotanical survey of wood-pasture habitats in Europe: Diversity, threats and conservation. *Biodiversity and Conservation* 19: 2995–3014.

Bianchi FJJA, Booij CJH, and Tscharntke T (2006) Sustainable pest regulation in agricultural landscapes: A review on landscape composition, biodiversity and natural pest control. *Proceedings of the Royal Society B: Biological Sciences* 273: 1715–1727.

Bock CE and Bock JH (1999) Response of winter birds to drought and short-duration grazing in southeastern Arizona. *Conservation Biology* 13: 1117–1123.

Brattsten LB, Holyoke CW, Leeper JR, and Raffa KF (1986) Insecticide resistance: Challenge to pest management and basic research. *Science* 231: 1255–1260.

Briske DD, Derner JD, Brown JR, *et al.* (2008) Rotational grazing on rangelands: Reconciliation of perception and experimental evidence. *Rangeland Ecology & Management* 61: 3–17.

Brown GW, Bennett AF, and Potts JM (2008) Regional faunal decline – Reptile occurrence in fragmented rural landscapes of south-eastern Australia. *Wildlife Research* 35: 8–18.

Brust GE and King LR (1994) Effects of crop rotation and reduced chemical inputs on pests and predators in maize agroecosystems. *Agriculture, Ecosystems & Environment* 48: 77–89.

Carson R (1962) *Silent Spring*. Boston: Houghton Mifflin.

Chauzat M-P and Faucon J-P (2007) Pesticide residues in beeswax samples collected from honey bee colonies (*Apis mellifera* L.) in France. *Pest Management Science* 63: 1100–1106.

Clough Y, Kruess A, Kleijn D, and Tscharntke T (2005) Spider diversity in cereal fields: Comparing factors at local, landscape and regional scales. *Journal of Biogeography* 32: 2007–2014.

Conrad KF, Warren MS, Fox R, Parsons MS, and Woiwod IP (2006) Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biological Conservation* 132: 279–291.

Crowder DW, Northfield TD, Strand MR, and Snyder WE (2010) Organic agriculture promotes evenness and natural pest control. *Nature* 466: 109–112.

Daily GC (1999) Developing a scientific basis for managing Earth's life support systems. *Conservation Ecology* 3: 14. http://www.consecol.org/vol3/iss2/art14.

Daily GC (2001) Ecological forecasts. *Nature* 411: 245.

Daily GC, Ceballos G, Pacheco J, Suzan G, and Sanchez-Azofeifa A (2003) Countryside biogeography of neotropical mammals: Conservation opportunities in agricultural landscapes of Costa Rica. *Conservation Biology* 17: 1814–1826.

Dalle SP and De Bois S (2006) Shorter fallow cycles affect the availability of noncrop plant resources in a shifting cultivation system. *Ecology and Society* 11: e2.

Demars CA, Rosenberg DK, and Fontaine JB (2010) Multi-scale factors affecting bird use of isolated remnant oak trees in agro-ecosystems. *Biological Conservation* 143: 1485–1492.

Diamond JM (1975) The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation* 7: 129–145.

Donald PF and Evans AD (2006) Habitat connectivity and matrix restoration: The wider implications of agri-environment schemes. *Journal of Applied Ecology* 43: 209–218.

Duncan DH, Dorrough J, White M, and Moxham C (2008) Blowing in the wind? Nutrient enrichment of remnant woodlands in an agricultural landscape. *Landscape Ecology* 23: 107–119.

Dunn CP, Stearns F, Guntenspergen GR, and Sharpe DM (1993) Ecological benefits of the conservation reserve program. *Conservation Biology* 7: 132–139.

Dunn RR (2010) Global mapping of ecosystem disservices: The unspoken reality that nature sometimes kills us. *Biotropica* 42: 555–557.

Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, and Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography* 19: 589–606.

Ellis JD, Evans JD, and Pettis J (2010) Colony losses, managed colony population decline, and Colony Collapse Disorder in the United States. *Journal of Apicultural Research* 49: 134–136.

Fahrig L, Baudry J, Brotons L, *et al.* (2011) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. *Ecology Letters* 14: 101–112.

Fearnside PM (2002) Can pasture intensification discourage deforestation in the Amazon and Pantanal regions of Brazil? In: Wood CH and Pozzo R (eds.) *Deforestation and Land Use in the Amazon and Pantanal Regions of Brazil*. Gainesville: University of Florida Press.

Fischer J, Brosi B, Daily GC, *et al.* (2008) Should agricultural policies encourage land sparing or wildlife-friendly farming? *Frontiers in Ecology and the Environment* 6: 380–385.

Fischer J, Brosi B, Daily GC, *et al.* (2009) Fostering constructive debate: A reply to Chappell *et al. Frontiers in Ecology and the Environment* 7: 84.

Fischer J, Fazey I, Briese R, and Lindenmayer DB (2005) Making the matrix matter: Challenges in Australian grazing landscapes. *Biodiversity and Conservation* 14: 561–578.

Fischer J and Lindenmayer DB (2002) The conservation value of paddock trees for birds in a variegated landscape in southern New South Wales. 2. Paddock trees as stepping stones. *Biodiversity and Conservation* 11: 833–849.

Fischer J, Lindenmayer DB, and Manning AD (2006) Biodiversity, ecosystem function and resilience: Ten guiding principles for off-reserve conservation. *Frontiers in Ecology and the Environment* 4: 80–86.

Fischer J, Sherren K, Stott J, Zerger A, Warren G, and Stein J (2010) Towards landscape-wide conservation outcomes in Australia's temperate grazing region. *Frontiers in Ecology and the Environment* 8: 69–74.

Fischer J, Stott J, and Law BS (2010) The disproportionate value of scattered trees. *Biological Conservation* 143: 1564–1567.

Fischer J, Stott J, Zerger A, Warren G, Sherren K, and Forrester R (2009) Reversing a tree regeneration crisis in an endangered ecoregion. *Proceedings of the National Academy of Sciences of the United States of America* 105: 10386–10391.

Fischer J, Zerger A, Gibbons P, Stott J, and Law BS (2010) Tree decline and the future of Australian farmland biodiversity. *Proceedings of the National Academy of Sciences of the United States of America* 107: 19597–19602.

Fitt G, Wilson L, Kelly D, and Mensah R (2009) Advances with integrated pest management as a component of sustainable agriculture: The case of the Australian cotton industry. In: Peshin R and Dhawan AK (eds.) *Integrated Pest Management: Dissemination and Impact*. Netherlands: Springer.

Foley JA, Defries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.

Ford HA, Walters JR, Cooper CB, Debus SJS, and Doerr VAJ (2009) Extinction debt or habitat change? – Ongoing losses of woodland birds in north-eastern New South Wales, Australia. *Biological Conservation* 142: 3182–3190.

Frazier M, Mullin C, Frazier J, and Ashcraft S (2008) What have pesticides got to do with it? *American Bee Journal* 148: 521–523.

Garnett T (2009) Livestock-related greenhouse gas emissions: Impacts and options for policy makers. *Environmental Science & Policy* 12: 491–503.

Gascon C and Lovejoy TE (1998) Ecological impacts of forest fragmentation in central Amazonia. *Ecology* 101: 273–280.

Geiger F, Bengtsson J, Berendse F, *et al.* (2010) Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. *Basic and Applied Ecology* 11: 97–105.

Gibbons P and Lindenmayer DB (2002) *Tree Hollows and Wildlife Conservation in Australia*. Collingwood: CSIRO Publishing.

Gillies CS and St. Clair CC (2010) Functional responses in habitat selection by tropical birds moving through fragmented forest. *Journal of Applied Ecology* 47: 182–190.

Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–555.

Greenberg R (2001) Criteria working group thought paper. Smithsonian Migratory Bird Center, Washington, DC, USA. Available from http://nationalzoo.si.edu/ConservationAndScience/MigratoryBirds/Coffee/thoughtpaper.pdf.].

Greenwood KL and Mckenzie BM (2001) Grazing effects on soil physical properties and the consequences for pastures: A review. *Australian Journal of Experimental Agriculture* 41: 1231–1250.

Guevara R (2005) Saprotrophic mycelial cord abundance, length and survivorship are reduced in the conversion of tropical cloud forest to shaded coffee plantation. *Biological Conservation* 125: 261–268.

Guevara S and Laborde J (1993) Monitoring seed dispersal at isolated standing trees in tropical pastures: Consequences for local species availability. *Vegetatio* 108: 319–338.

Hanski I and Simberloff D (1997) The metapopulation approach, its history, conceptual domain and application to conservation. In: Hanski I and Gilpin ME (eds.) *Metapopulation Biology*. San Diego: Academic Press.

Harvey CA, Villanueva C, Villacis J, *et al.* (2005) Contribution of live fences to the ecological integrity of agricultural landscapes. *Agriculture Ecosystems & Environment* 111: 200–230.

Haslem A and Bennett AF (2008) Birds in agricultural mosaics: The influence of landscape pattern and countryside heterogeneity. *Ecological Applications* 18: 185–196.

Henderson IG, Ravenscroft N, Smith G, and Holloway S (2009) Effects of crop diversification and low pesticide inputs on bird populations on arable land. *Agriculture, Ecosystems & Environment* 129: 149–156.

Hickling R, Roy DB, Hill JK, Fox R, and Thomas CD (2006) The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology* 12: 450–455.

Hoekstra JM, Boucher TM, Ricketts TH, and Roberts C (2005) Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters* 8: 23–29.

Hole DG, Perkins AJ, Wilson JD, Alexander IH, Grice F, and Evans AD (2005) Does organic farming benefit biodiversity? *Biological Conservation* 122: 113–130.

Holzschuh A, Steffan-Dewenter I, Kleijn D, and Tscharntke T (2007) Diversity of flower-visiting bees in cereal fields: Effects of farming system, landscape composition and regional context. *Journal of Applied Ecology* 44: 41–49.

Horner-Devine MC, Daily GC, Ehrlich PR, and Boggs CL (2003) Countryside biogeography of tropical butterflies. *Conservation Biology* 17: 168–177.

Huband S (2007). *The role of Romanian pastoralists in conserving agricultural biodiversity, Edinburgh*. PhD Thesis, The University of Edinburgh.

Hylander K and Nemomissa S (2008) Home garden coffee as a repository of epiphyte biodiversity in Ethiopia. *Frontiers in Ecology and the Environment* 6: 524–528.

Kleijn D, Baquero RA, Clough Y, *et al.* (2006) Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology Letters* 9: 243–254.

Kleijn D, Kohler F, Baldi A, *et al.* (2009) On the relationship between farmland biodiversity and land-use intensity in Europe. *Proceedings of the Royal Society B-Biological Sciences* 276: 903–909.

Kleijn D and Sutherland WJ (2003) How effective are European agri-environment schemes in conserving and promoting biodiversity? *Journal of Applied Ecology* 40: 947–969.

Klein A-M, Steffan-Dewenter I, Buchori D, and Tscharntke T (2002) Effects of land-use intensity in tropical agroforestry systems on coffee flower-visiting and trap-nesting bees and wasps. *Conservation Biology* 16: 1003–1014.

Knop EVA, Kleijn D, Herzog F, and Schmid B (2006) Effectiveness of the Swiss agri-environment scheme in promoting biodiversity. *Journal of Applied Ecology* 43: 120–127.

Kogan M (1998) Integrated pest management: Historical perspectives and contemporary developments. *Annual Review of Entomology* 43: 243–270.

Kremen C, Williams NM, and Thorp RW (2002) Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences of the United States of America* 99: 16812–16816.

Kuussaari M, Bommarco R, Heikkinen RK, *et al.* (2009) Extinction debt: A challenge for biodiversity conservation. *Trends in ecology & evolution (Personal edition)* 24: 564–571.

Letourneau DK and Goldstein B (2001) Pest damage and arthropod community structure in organic vs. conventional tomato production in California. *Journal of Applied Ecology* 38: 557–570.

Lindenmayer DB and Fischer J (2006) *Habitat Fragmentation and Landscape Change*. Washington, DC: Island Press.

Lindenmayer DB and Franklin J (2002) *Conserving Forest Biodiversity*. Covelo, California: Island Press.

Lindenmayer DB, Knight EJ, Crane MJ, Montague-Drake R, Michael DR, and Macgregor CI (2010) What makes an effective restoration planting for woodland birds? *Biological Conservation* 143: 289–301.

Longley M, Čilgi T, Jepson PC, and Sotherton NW (1997) Measurements of pesticide spray drift deposition into field boundaries and hedgerows: 1. Summer applications. *Environmental Toxicology and Chemistry* 16: 165–172.

Lupwayi NZ, Rice WA, and Clayton GW (1998) Soil microbial diversity and community structure under wheat as influenced by tillage and crop rotation. *Soil Biology and Biochemistry* 30: 1733–1741.

Macarthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Manning AD, Fischer J, and Lindenmayer DB (2006) Scattered trees are keystone structures – Implications for conservation. *Biological Conservation* 132: 311–321.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.

Maron M and Fitzsimons JA (2007) Agricultural intensification and loss of matrix habitat over 23 years in the West Wimmera, south-eastern Australia. *Biological Conservation* 135: 587–593.

Matson PA and Vitousek PM (2006) Agricultural intensification: Will land spared from farming be land spared for nature? *Conservation Biology* 20: 709–710.

Mattison EHA and Norris K (2005) Bridging the gaps between agricultural policy, land-use and biodiversity. *Trends in Ecology & Evolution* 20: 610–616.

Mayfield MM, Boni ME, Daily GC, and Ackerly D (2005) Species and functional diversity of native and human-dominated plant communities. *Ecology* 86: 2365–2372.

Mcguinness KA (1984) Equations and explanations in the study of species-area curves. *Biological Reviews* 59: 423–440.

Mcintyre S and Lavorel S (2007) A conceptual model of land use effects on the structure and function of herbaceous vegetation. *Agriculture, Ecosystems & Environment* 119: 11–21.

Michener CD (2000) *The Bees of the World*. The Johns Hopkins University Press

Moguel P and Toledo VM (1999) Biodiversity conservation in traditional coffee systems of Mexico. *Conservation Biology* 13: 11–21.

Morandin LA and Winston ML (2006) Pollinators provide economic incentive to preserve natural land in agroecosystems. *Agriculture, Ecosystems & Environment* 116: 289–292.

Murphy MT (2003) Avian population trends within the evolving agricultural landscape of eastern and central United States. *Auk* 120: 20–34.

Nestel D, Dickschen F, and Altieri MA (1993) Diversity patterns of soil macro-Coleoptera in Mexican shaded and unshaded coffee agroecosystems: An indication of habitat perturbation. *Biodiversity and Conservation* 2: 70–78.

Oaks JL, Gilbert M, Virani MZ, *et al.* (2004) Diclofenac residues as the cause of vulture population decline in Pakistan. *Nature* 427: 630–633.

Oldroyd BP (2007) What's killing American honey bees? *PLoS Biol* 5: e168.

Oliver I, Garden D, Greenslade PJ, *et al.* (2005) Effects of fertiliser and grazing on the arthropod communities of a native grassland in south-eastern Australia. *Agriculture Ecosystems & Environment* 109: 323–334.

O'Rourke ME, Liebman M, and Rice ME (2008) Ground beetle (Coleoptera: Carabidae) assemblages in conventional and diversified crop rotation systems. *Environmental Entomology* 37: 121–130.

Olsson P, Folke C, and Hughes TP (2008) Navigating the transition to ecosystem-based management of the Great Barrier Reef Australia. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9489–9494.

Otto S, Lazzaro L, Finizio A, and Zanin G (2009) Estimating ecotoxicological effects of pesticide drift on nontarget arthropods in field hedgerows. *Environmental Toxicology and Chemistry* 28: 853–863.

Ozolins A, Brack C, and Freudenberger D (2001) Abundance and decline of isolated trees in the agricultural landscapes of central New South Wales, Australia. *Pacific Conservation Biology* 7: 195–203.
Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics* 37: 637–669.
Pereira HM and Daily GC (2006) Modeling biodiversity dynamics in countryside landscapes. *Ecology* 87: 1877–1885.
Perfecto I and Snelling R (1995) Biodiversity and the transformation of a tropical agroecosystem: Ants in coffee plantations. *Ecological Applications* 5: 1084–1097.
Perfecto I and Vandermeer J (2002) Quality of agroecological matrix in a tropical montane landscape: Ants in coffee plantations in southern Mexico. *Conservation Biology* 16: 174–182.
Phalan B, Balmford A, Green RE, and Scharlemann JPW (2011) Minimising the harm to biodiversity of producing more food globally. *Food Policy* 36.
Pimentel D (2009) Environmental and economic costs of the application of pesticides primarily in the United States. In: Peshin R and Dhawan AK (eds.) *Integrated Pest Management: Innovation-Development Process*. Netherlands: Springer.
Pimentel D, Acquay H, Biltonen M, *et al.* (1992) Environmental and economic costs of pesticide use. *BioScience* 42: 750–760.
Pineda E, Moreno C, Escobar F, and Halffter G (2005) Frog, bat, and dung beetle diversity in the cloud forest and coffee agroecosystems of Veracruz, Mexico. *Conservation Biology* 19: 400–410.
Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, and Kunin WE (2010) Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology & Evolution* 25: 345–353.
Potts SG, Roberts SPM, Dean R, *et al.* (2010) Declines of managed honey bees and beekeepers in Europe. *Journal of Apicultural Research* 49: 15–22.
Rainforest Alliance (2009) *Sustainable Agriculture – Coffee*. San José, Costa Rica. http://rainforest-alliance.org/agriculture/crops/coffee.
Ranganathan J, Daniels RJR, Chandran MDS, Ehrlich PR, and Daily GC (2008) Sustaining biodiversity in ancient tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America* 105: 17852–17854.
Rappole JH, King DI, and Rivera JHV (2003) Coffee and conservation. *Conservation Biology* 17: 334–336.
Recher HF (1999) The state of Australia's avifauna: A personal opinion and prediction for the new millennium. *Australian Zoologist* 31: 11–29.
Relyea RA (2005) The impact of insecticides and herbicides on the biodiversity and productivity of aquatic communities. *Ecological Applications* 15: 618–627.
Ribaudo MO, Colacicco D, Langner LL, Piper S, and Schaible GD (1990) Natural resources and users benefit from the Conservation Reserve Program. *Agricultural Economic Report 627*. Washington, DC: US Department of Agriculture, Economic Research Service.
Ricketts TH, Regetz J, Steffan-Dewenter I, *et al.* (2008) Landscape effects on crop pollination services: Are there general patterns? *Ecology Letters* 11: 499–515.
Robinson WD (1999) Long-term changes in the avifauna of Barro Colorado Island, Panama, a tropical forest isolate. *Conservation Biology* 13: 85–97.
Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.
Rundlöf M, Edlund M, and Smith HG (2010) Organic farming at local and landscape scales benefits plant diversity. *Ecography* 33: 514–522.
Rundlöf M, Nilsson H, and Smith HG (2008) Interacting effects of farming practice and landscape context on bumble bees. *Biological Conservation* 141: 417–426.
Rundlöf M and Smith HG (2006) The effect of organic farming on butterfly diversity depends on landscape context. *Journal of Applied Ecology* 43: 1121–1127.
Saunders DA (1980) Food and movements of the short-billed form of the White-tailed Black Cockatoo. *Australian Wildlife Research* 7: 257–269.
Savory A and Butterfield J (1999) *Holistic Management*. Washington, DC: Island Press.
Schmitt T and Rakosy L (2007) Changes of traditional agrarian landscapes and their conservation implications: A case study of butterflies in Romania. *Diversity and Distributions* 13: 855–862.
Sekercioglu CH, Loarie SR, Brenes FO, Ehrlich PR, and Daily GC (2007) Persistence of forest birds in the Costa Rican agricultural countryside. *Conservation Biology* 21: 482–494.
Sklenicka P, Molnarova K, Brabec E, *et al.* (2009) Remnants of medieval field patterns in the Czech Republic: Analysis of driving forces behind their disappearance with special attention to the role of hedgerows. *Agriculture Ecosystems & Environment* 129: 465–473.
Steffan-Dewenter I, Kessler M, Barkmann J, *et al.* (2007) Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification. *Proceedings of the National Academy of Sciences of the United States of America* 104: 4973–4978.
Stinner DH, Stinner BR, and Martsolf E (1997) Biodiversity as an organizing principle in agroecosystem management: Case studies of holistic resource management practitioners in the USA. *Agriculture Ecosystems & Environment* 62: 199–213.
Sustainable Agriculture Network (2009) http://www.rainforest-alliance.org/agriculture/documents/sust_ag_standard.pdf.
Tejeda-Cruz C, Silva-Rivera E, Barton JR, and Sutherland WJ (2010) Why shade coffee does not guarantee biodiversity conservation. *Ecology and Society* 15: 13.
Tejeda-Cruz C and Sutherland WJ (2004) Bird responses to shade coffee production. *Animal Conservation* 7: 169–179.
Tews J, Brose U, Grimm V, *et al.* (2004) Animal species diversity driven by habitat heterogeneity/diversity: The importance of keystone structures. *Journal of Biogeography* 31: 79–92.
Tilman D, Cassman KG, Matson PA, Naylor R, and Polasky S (2002) Agricultural sustainability and intensive production practices. *Nature* 418: 671–677.
Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity – Ecosystem service management. *Ecology Letters* 8: 857–874.
Uezu A, Metzger JP, and Vielliard JME (2005) Effects of structural and functional connectivity and patch size on the abundance of seven Atlantic Forest bird species. *Biological Conservation* 123: 507–519.
Vickery JA, Bradbury RB, Henderson IG, Eaton MA, and Grice PV (2004) The role of agri-environment schemes and farm management practices in reversing the decline of farmland birds in England. *Biological Conservation* 119: 19–39.
Wanger TC, Rauf A, and Schwarze S (2010) Pesticides and tropical biodiversity. *Frontiers in Ecology and the Environment* 8: 178–179.
Wegner JF and Merriam G (1979) Movements by birds and small mammals between a wood and adjoining farmland habitats. *Journal of Applied Ecology* 16: 349–358.
Weibull A-C, Östman Ö, and Granqvist Å (2003) Species richness in agroecosystems: The effect of landscape, habitat and farm management. *Biodiversity and Conservation* 12: 1335–1355.
Whittingham MJ (2007) Will agri-environment schemes deliver substantial biodiversity gain, and if not why not? *Journal of Applied Ecology* 44: 1–5.
Williams-Guillén K, Mccann C, Martínez Sánchez JC, and Koontz F (2006) Resource availability and habitat use by mantled howling monkeys in a Nicaraguan coffee plantation: Can agroforests serve as core habitat for a forest mammal? *Animal Conservation* 9: 331–338.
Winfree R, Griswold T, and Kremen C (2007) Effect of human disturbance on bee communities in a forested ecosystem. *Conservation Biology* 21: 213–223.
Yates CJ, Norton DA, and Hobbs RJ (2000) Grazing effects on plant cover, soil and microclimate in fragmented woodlands in south-western Australia: Implications for restoration. *Austral Ecology* 25: 36–47.
Zhang W, Ricketts TH, Kremen C, Carney K, and Swinton SM (2007) Ecosystem services and dis-services to agriculture. *Ecological Economics* 64: 253–260.

Biodiversity Generation, Overview

Paul H Harvey, University of Oxford, Oxford, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 403–409, © 2001, Elsevier Inc.

Glossary

Key innovation A property of a species that promotes the process of speciation.

Molecular phylogenies Phylogenies constructed from comparisons of homologous genetic material.

Radiation All the species descended from a single common ancestor; tends to be a speciose group in comparison with others of a similar age.

Reconstructed evolutionary trees Phylogenies linking contemporary species, therefore containing no information about lineages that have gone extinct.

Sister taxa The expression has two uses, according to context. First, from a species perspective, the sister taxon of a particular species is the one with which it had the most recent common ancestor. Second, from the perspective of a node in a bifurcating phylogenetic tree, each daughter lineage gives rise to a set of contemporary species. The two sets are sister taxa with reference to the node.

The Fossil Record and Contemporary Biodiversity

The generation of biodiversity involves, in large part, the origins of species. The fossil record, though notoriously incomplete, provides the most direct evidence for tracing temporal changes in biodiversity, including both the origin and extinction of lineages. Biases in preservation, often confounded with differences in taxonomic interest among paleontologists, mean that much more is known about the history of some taxa than others. For example, May (1996) points out that while around 95% of fossil databases consist of shallow-water marine invertebrates, most of today's known species are terrestrial, with 56% being insects. At the same time, Benton and Storrs' (1996) analyses suggest that more is known about the fossil history of vertebrates, most of which have been terrestrial, than of echinoderms, which are marine.

Benton's (1993) extensive compilations of stratigraphic incidence records for different lineages reveal the extent of well-known mass extinctions, followed by the subsequent regeneration of biodiversity. The rises in biodiversity that follow mass extinctions are presumably a consequence of the exploitation of (and speciation into) vacant niches. Also it is clear that the demise of some radiations is accompanied or followed by the rise of others; dinosaurs and mammals are cases in point. The causes of such complementarities have been hotly debated for many years. Similarly, there are times when one radiation is accompanied by, or shortly follows, another. The angiosperm plants are a massive radiation, as are the beetles. Cause and effect have been firmly suggested: the radiation of angiosperm plants provided a new set of niches into which beetles could speciate (Farrell, 1998). Arguments such as this are frequently anecdotal in that they involve post hoc explanations that require further testing, which often proves very difficult. For example, while the angiosperm plant radiation can explain about one-half of the species number in the beetle radiation (there are about 300,000 known beetle species), the other half remains unexplained. As Barraclough *et al.* (1998) point out, the angiosperm radiation does not explain the large radiation of 50,000 mainly carnivorous and fungivorous beetle species in the Cucujoidea or the equally large radiation of predatory beetle species in the Staphylinoidea.

Reconstructed Versus Real Evolutionary Trees

The reasons for the origins of contemporary biodiversity are becoming better understood, largely as a consequence of advances in molecular genetics. Gene sequence analysis has become routine in many laboratories, and sequence data are deposited in openly accessible data banks. When equivalent sequences are available from different species, phylogenetic trees can be reconstructed. If a molecular clock has been operating, these trees show when in relative time any pair of species last shared a common ancestor. When the molecular clock can be dated, the trees can be calibrated in real time. Such phylogenies show the relationships among contemporary species, and contain no explicit information about extinct lineages. However, the structure of such trees can be used to demonstrate which evolutionary models are more likely than others to explain the origins of biodiversity in particular taxa.

When a model-based approach is used to analyze the origins of biodiversity, it soon becomes apparent how easily intuition can lead us astray. For example, the tree reproduced in **Figure 1** shows the relationships among contemporary salamander species from the genus *Plethodon*. The branching rate in the tree seems to have increased recently. This, however, does not necessarily mean that the net rate of speciation has increased. The simplest model to describe the generation of biodiversity is a constant-rates model in which the rate of speciation (lineage splitting) and the rate of extinction have been the same in all lineages at all times since the root of the tree. Obviously the rate of speciation will have been greater

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00010-1

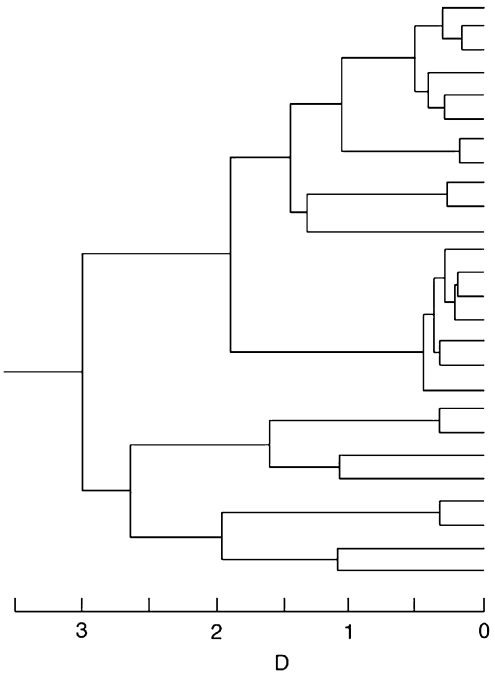


Figure 1 A phylogenetic tree reconstructed from genetic data showing the relationships among salamanders of the genus *Plethodon*. *D* is Rogers' genetic distance. The phylogeny reconstruction assumes a molecular clock, meaning that nodes are placed approximately correctly in relative time. No absolute timescale is available because the rate of ticking of the molecular clock is not known with precision.

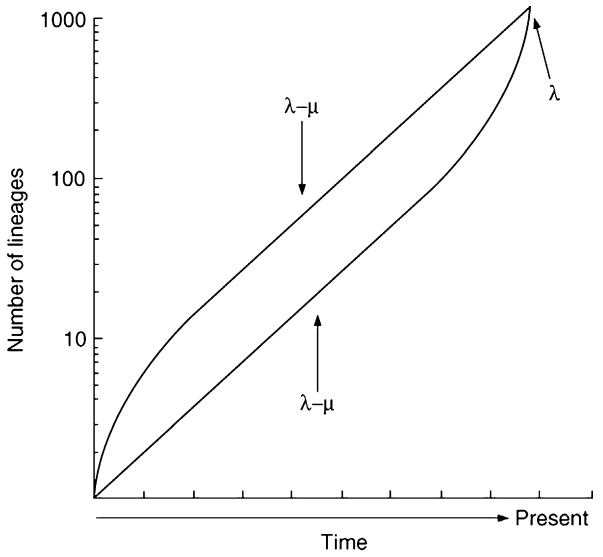


Figure 2 The expected forms of an actual phylogeny (top line) and a reconstructed phylogeny (bottom line) produced by a birth-death process, in which the constant per-lineage speciation rate (λ) is greater than the constant per-lineage extinction rate (μ). Note that the two curves would be superimposed if the extinction rate was zero. The initial steep slope of the line for the actual phylogeny is a consequence of only a sample of phylogenies surviving to the present and those are likely to be the ones that got off to a flying start. Note also that as μ increases from zero toward λ, so does the change in slope of the reconstructed line, so that a steep upturn describes a phylogeny where μ is high in comparison with λ. Reproduced from Harvey PH, May RM, and Nee S (1994) Phylogenies without fossils. *Evolution* 48: 523–529.

than the rate of extinction, or the tree would not have grown. Under such a model, if the rate of extinction is zero, a lineages-through-time plot based on the relationships among contemporary species would be a straight line with a slope, λ, the speciation rate. As the rate of extinction (μ) increases, so does curvature of the line upward toward the present (**Figure 2**). This means that the plethodontid salamander tree, with its increased branching rate toward the present, is representative of a phylogeny generated under a high rate of extinction relative to speciation. Of course, speciation and extinction rates could have been changing through time, but the simple constant-rates model is sufficient to explain the data. If we accept this model, then it is also possible to estimate speciation and extinction rates or, more usefully for conservation biologists, the net rate of speciation (λ–μ) and the danger of a clade going extinct (μ/λ) (**Figure 3**). Similar analyses have been carried out for diverse taxa, including primates, carnivores, *Drosophila*, and birds.

Biodiversity cannot increase indefinitely. Eventually, any adaptive radiation will run out of niches to occupy and net speciation rates will decrease. Such an effect is apparent when Sibley and Ahlquist's (1990) molecular phylogeny of birds is analyzed as a lineages-through-time plot (**Figure 4**). Instead of the line steepening toward the present, there is a general leveling off. In the absence of fossil material, which is relatively sparse for birds, it is not possible to say whether, in fact, speciation rates have decreased through time, extinction rates have increased, or both. Sibley and Ahlquist's phylogeny is one of the earlier molecular phylogenies and is based on

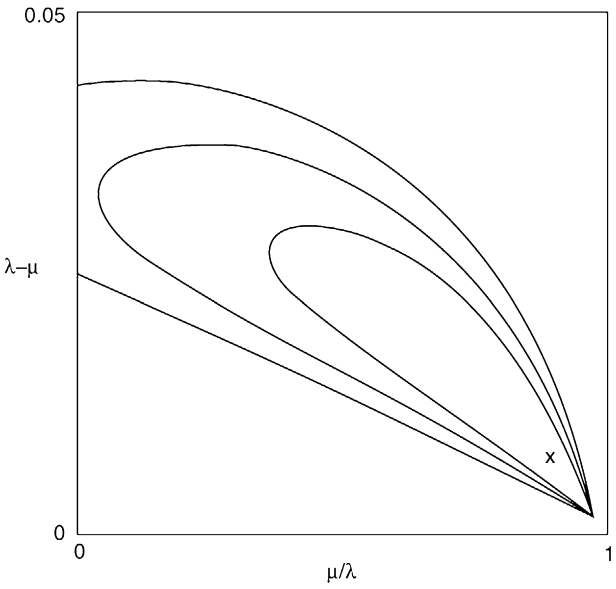


Figure 3 Contour plot of the likelihood surface for the *Plethodon* data (**Figure 1**). The maximum likelihood estimate, the peak of the surface, is marked by an X. The lines surrounding it approximate to the 90, 95, and 99% confidence limits, respectively. Reproduced from Nee S, Holmes EC, May RM, and Harvey PH (1995) Estimating extinction from molecular phylogenies. In: Lawton JH and May RM (eds.) *Estimating Extinction Rates*, pp. 164–182. Oxford, United Kingdom: Oxford University Press.

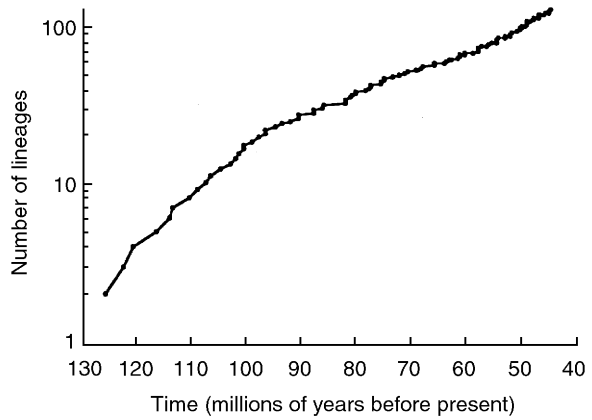


Figure 4 A lineages-through-time plot for the molecular phylogeny of birds. All families are represented, that is, all lineages that were present about 45 million years ago and have left descendants in the present day. Accordingly, we should not expect the leveling off of the curve under a constant-rate process: either speciation rate has been decreasing over time or extinction rate has been increasing. There is no evidence for sharp kinks in the curve, which would provide evidence for mass extinction events. For example, the mass extinction at the end of the Cretaceous about 65 million years ago is thought to have resulted in the loss of about 75% of species in many taxa. Reproduced from Nee S, Mooers AØ, and Harvey PH (1992) The tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl. Acad. Sci. USA* 89: 8322–8326.

DNA–DNA hybridization data: DNA from different species is annealed, and then heated to dissociate the strands. The more similar that strands are in base composition, the higher the temperature at which they dissociate. More recent studies based on direct comparison of sequences largely support Sibley and Ahlquist's original phylogeny, which, for example, indicated for the first time that the Australian passerines are an independent radiation from other passerines. In addition to many sequence-based studies that lend support for the general phylogeny, one reports a slowing down in the net rate of speciation in eight out of nine genera studied.

If we are to better understand the reasons for the generation of biodiversity, it is critical that molecular phylogenies are properly related to the known fossil record and to our understanding of biogeographic events. For example, the molecular phylogeny of birds reveals two aberrant clades that were speciating at a higher rate when that for others was slowing down. The two radiations resulted in the Passeriformes (seed-eating birds) and the Cicconiformes (mainly shore-birds). The fossil record of both groups is poor, but it is clear that the radiations occurred at a time when landmasses were warmer and shorelines longer than previously. Though cause and effect may be inferred, the unique events described leave the way open for other possible explanations. One way to strengthen interpretations is to show similar increases in biodiversity associated with the similar environmental changes on repeated occasions. If it could be shown, for example, that bursts in speciation among shorebirds were repeatedly associated with increases in the length of shoreline, it would be less likely that a third unknown environmental factor was responsible for changes in shorebird speciation rates.

Key Innovations, Sister Taxon Analysis, and the Origins of Adaptive Radiations

It is not necessary that increases in biodiversity occur in response to environmental changes: so-called key innovations may evolve that allow the invasion of currently unoccupied niches. Liem (1973, 1980) suggested that one such innovation is responsible for the much higher rates of speciation among cichlid fishes than among their sister taxa. Liem pointed out that a small shift in position of a single muscle attachment ultimately allowed the pharyngeal bones of cichlids to manipulate their prey items while still holding them. As a consequence, the premaxillary and mandibular jaws were freed to evolve along new routes that did not involve manipulating prey. This, the argument runs, allowed cichlids to evolve a whole new diversity of feeding mechanisms and there seems little doubt that the adaptive radiation of cichlids in African lakes, in the face of competition from other fish families, resulted from their evolved diversity of feeding mechanisms. However, the problem of nonreplication occurs: Lauder (1981) correctly pointed out that any other characteristic common to cichlid fishes but that differs from a sister group could, in principle, be the key innovation that resulted in high speciation rates. Conclusions would be strengthened if the same key innovation had been associated with high species diversity in several taxa.

The first example of a statistically supported successful sister group comparison showing a repeated correlate of high species diversity comes from the Mitter *et al.* (1988) analysis of phytophagy, the habit of feeding on vascular plants, among insects. There are considerable barriers to the evolution of phytophagy. Behavioral and morphological adaptations are required to reduce the risk of desiccation, to remain attached to host, and to deal with low-nutrient food. It had been argued that, once those barriers were overcome, the diversity of phytophages would be promoted by both the great diversity of plant species and of plant parts that could not be exploited by a single species, and by the absence of competitors. Mitter *et al*. performed 13 sister taxon comparisons between the numbers of species in a phytophagous clade versus its non-phytophagous sister clade. Since sister clades originated at the same time, they have had identical times for diversification. In 11 of the 13 comparisons, there were more species in the phytophagous clade than in the sister group, and in each of those cases the difference was greater than twofold.

Sister taxon comparisons using newly established molecular phylogenies are beginning to allow the first real statistical tests of many long-standing hypotheses. For example, Darwin (1871) suggested that sexual selection by female choice might increase the rate of reproductive divergence between populations, and thereby increase the speciation rate of a clade. More than a century later, Barraclough *et al.* (1995) tested the idea using Sibley and Ahlquist's phylogeny of passerine birds. Since it is generally accepted that mate choice is responsible for the evolution of sexual dichromatism in birds, Barraclough *et al*. were able to use sexual dichromatism as a measure of the importance of mate choice for a species. In significantly more than 50% of sister taxon comparisons, the clade with the higher proportion of sexually dichromatic species was more speciose than its sister clade, thereby supporting Darwin's suggestion.

Evolution within Lineages: The Comparative Method

It is clear that some characteristics of species can promote speciation itself. But species differ from each other in many ways: morphologically, anatomically, physiologically, behaviorally, and by life history. Making sense of the reasons for that diversity is the job of comparative biologists.

Fewer Causes than Characters

Traits have not evolved independently from each other, and particular characteristics of the environment select for whole suites of traits. For example, among many mammalian orders such as the primates, females tend to be spatially distributed in relation to food resources. If food is in large clumps, females live in large groups, but if food is evenly distributed, females are often territorial. Males distribute themselves in response to female grouping patterns, having been selected to monopolize mating access to as many females as possible. When females live alone in territories, males follow suit and share territorial defense (monogamous, as in gibbons). If females live in small groups, then a single male may be able to defend a group of females from mating access by other males (single-male societies). But if females live in groups that are sufficiently large that they cannot be defended by a single male, then more than one male may join the group (multimale societies). Among monogamous species, the males and females have similar weaponry, such as canines, having been selected equally to defend their territory. In single-male and multimale species, males have enlarged canines relative to females of the species, so that they can defend the group of females, or at least any that are in estrus, against mating access from other males. However, in multimale species, there is always the potential for females to be mated by more than one male; males in such species have markedly enlarged testes as a consequence of selection to produce more sperm per ejaculate, thereby increasing the chances of success in sperm competition. Males of monogamous and single-male species are not subject to selection for sperm competition. Here, then, the distribution of resources influences differences in social behavior, weaponry, and testes size. Indeed, if we know a male primate's testes and relative canine sizes, it is possible to say whether he belongs to a monogamous species (both small), a single-male species (large canines, small testes), or a multimale species (both large).

Identifying Independent Evolutionary Events

As mentioned in the previous section, there are many fewer explanatory factors than there are variable traits across species. How do we identify the factors that are responsible for the patterns of biodiversity that we see today? Species may share traits in common for two evolutionary reasons: they inherited them from a common ancestor or their ancestors independently evolved the traits. Closely related species share more traits through common ancestry, whereas more distantly related species share more through convergent evolution. This does not necessarily mean that traits cannot evolve rapidly in response to selection. For example, when a new niche appears, the species to successfully invade that niche is most likely to be one that came from a similar niche elsewhere. Very minor changes in phenotype may be sufficient to adapt the new species to its new home. However, the fact that closely related species are subject to similar selective pressures does mean that cross-species correlations can be misleading. For example, in a sample of vertebrates, we might find that those with feathers laid eggs while those with fur gave birth to live young. If the sample contains just birds and mammals, we should be wary of saying that the two sets of characters had responded to the same selective pressure. The way forward is to identify evolutionary independent origins of traits and niche occupancy. For example, we know that on many separate occasions bright coloration among insects has evolved together with distastefulness to predators. Because the two traits seem to have evolved in concert, and there is good reason to expect distasteful prey to advertise the fact to predators, we are more confident that they are part of the same adaptive complex.

For continuously varying characters, such as body size, a particularly elegant technique was developed by Felsenstein (1985) that allows comparative biologists to partition out independent evolutionary change and seek correlated evolution among traits. The key realizations are (1) that differences between each pair of sister taxa in a phylogeny evolved independently, and (2) likely ancestral character states could be estimated if characters had evolved according to a Brownian motion model of evolution. To determine whether characters had evolved independently of each other, it is simply necessary to plot sister taxa differences against each other. If one

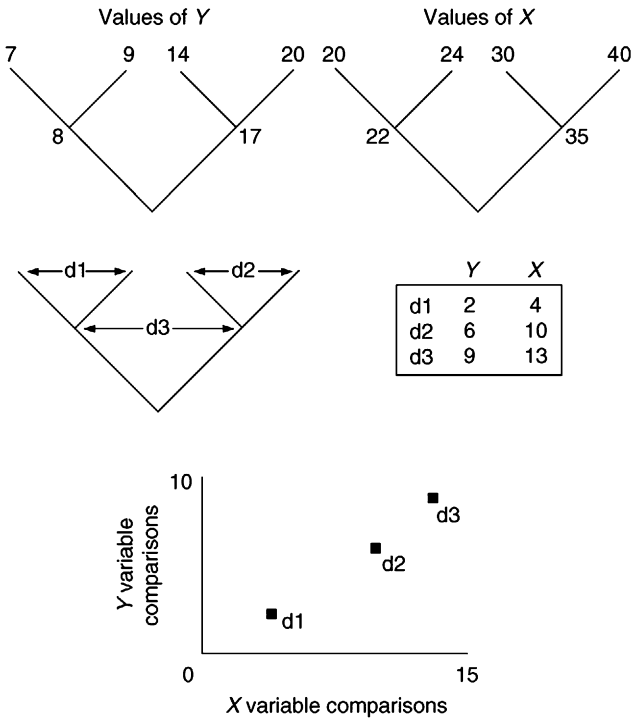


Figure 5 The comparative method of independent contrasts. Comparisons d1, d2, and d3 are independent of each other under a Brownian motion model of character evolution. Reproduced from Harvey PH and Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford, United Kingdom: Oxford University Press.

character evolved independently of another, differences between sister taxa would not be correlated (**Figure 5**).

An example of the use of independent contrast analyses is Kelly and Woodward's (1995) analysis of the correlates of variation in carbon isotope composition ($\partial^{13}C$) among plant species. Three cross-species correlates of $\partial^{13}C$ had previously been identified: altitude, latitude, and growth form. When the data were analyzed using phylogenetically independent contrasts, the correlation with latitude dropped out, and when altitude was controlled for there was no correlation with growth form. Alternative interpretations of the reasons for species differences in $\partial^{13}C$ were thereby reduced in number. Indeed, it was concluded that differences in atmospheric composition (CO_2 and O_2 partial pressures) may be sufficient to explain observed differences in carbon isotope composition.

See also: Adaptive Radiation. Biodiversity, Origin of. Cladogenesis. Evolution, Theory of. Extinction in the Fossil Record. Fossil Record. Phylogeny. Speciation, Process of

References

Barraclough TG, Barclay MVL, and Vogler AP (1998) Species richness: Does flower power explain beetle-mania? *Current Biol* 8: 843–845.

Barraclough TG, Harvey PH, and Nee S (1995) Sexual selection and taxonomic diversity in passerine birds. *Proc. Roy. Soc. London Ser. B* 259: 211–215.

Benton MJ (1993) *The Fossil Record 2*. London: Harper Collins/Academic.

Benton MJ and Storrs GW (1996) Diversity in the Past: Comparing cladistic phylogenies and stratigraphy. In: Hochberg ME, Clobert J, and Barbault R (eds.) *Aspects of the Genesis and Maintenance of Biological Diversity*. Oxford, United Kingdom: Oxford University Press.

Darwin C (1871) *The Descent of Man and Selection in Relation to Sex*. London: Murray.

Farrell BD (1998) "Inordinate fondness" explained: Why are there so many beetles? *Science* 281: 555–559.

Felsenstein J (1985) Phylogenies and the comparative method. *Amer. Naturalist* 125: 1–15.

Harvey PH, May RM, and Nee S (1994) Phylogenies without fossils. *Evolution* 48: 523–529.

Harvey PH and Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford, United Kingdom: Oxford University Press.

Kelly CK and Woodward FI (1995) Ecological correlates of carbon isotope composition of leaves: A comparative analysis testing for the effects of temperature CO_2, and O_2 partial pressures and taxonomic relatedness on $\partial^{13}C$. *J. Ecol* 83: 509–515.

Lauder GV (1981) Form and function: Structural analysis in evolutionary morphology. *Paleobiology* 7: 430–442.

Liem KF (1973) Evolutionary strategies and morphological innovations: Cichlid pharyngeal jaws. *Systematic Zool* 22: 425–441.

Liem KF (1980) Adaptive significance of intra- and interspecific differences in the feeding repertoires of cichlid fishes. *Amer. Zool* 20: 295–314.

May RM (1996) Introduction. In: Hochberg ME, Clobert J, and Barbault R (eds.) Aspects of the Genesis and Maintenance of Biological Diversity. Oxford, United Kingdom: Oxford University Press.

Mitter C, Farrell B, and Wiegmann B (1988) The phylogenetic study of adaptive zones: Has phytophagy promoted diversification? *Amer. Naturalist* 132: 107–128.

Nee S, Holmes EC, May RM, and Harvey PH (1995) Estimating extinction from molecular phylogenies. In: Lawton JH and May RM (eds.) *Estimating Extinction Rates*, pp. 164–182. Oxford, United Kingdom: Oxford University Press.

Nee S, Mooers AØ, and Harvey PH (1992) The tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl. Acad. Sci. USA* 89: 8322–8326.

Sibley CG and Ahlquist JE (1990) *Phylogeny and Classification of Birds*. New Haven, Connecticut: Yale University Press.

Biodiversity, Human Well-Being, and Markets

Seth Binder, University of Minnesota, St Paul, MN, USA, and Yale University, New Haven, CT, USA
Stephen Polasky, University of Minnesota, St Paul, MN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ecosystem services Goods and services provided by nature that contribute to human well-being such as clean water and charismatic wildlife.

Efficiency In economics, efficiency may refer to Pareto efficiency, which occurs when making one person better off necessarily makes someone else worse off. Efficiency may also refer to Kaldor–Hicks (or potential Pareto) efficiency, which occurs when those who are made better off by a change could not fully compensate those who are made worse off by that change. Efficiency implies that society cannot improve unambiguously on its current use of resources; it says nothing about distributional equity.

Externalities Costs or benefits of an action taken by one party that directly impact the welfare of other parties that are not reflected by market prices.

First fundamental theorem of welfare economics Mathematical proof that if competitive markets exist for all goods and there are no externalities or other market imperfections, then the market system leads to a Pareto-efficient outcome.

Public goods Goods whose consumption is nonrival and nonexcludable. Nonrival means that one person's consuming it does not preclude another from doing likewise. Nonexcludable means that the provider of the goods cannot ensure that only those who have paid can benefit from its provision. A fireworks display is an example of a public good.

Risk Situation in which outcomes are stochastic (not deterministic) but each alternative has a known probability of occurrence. Risk may be idiosyncratic (specific to an individual) or systemic (shared).

Introduction

Human actions have become important drivers of global change, including land-use change, changes in nutrient cycles and pollution, species declines and extinction from overharvesting, and climate change, all of which can lead to significant biodiversity loss. Although biodiversity loss may be viewed as a loss in its own right (intrinsic value), it can also have serious negative consequences on human well-being. Humans depend on biodiversity to produce ecosystem goods and services, ranging from foods, fuels, fibers, and medicines to ecosystem functions that provide water purification, soil enrichment, nutrient retention, and climate regulation.

This entry discusses two important and related topics. First, what are the various types of benefits that biodiversity provides to humans? These benefits can either directly affect human well-being through harvesting or other activities or indirectly benefit people through the role biodiversity plays in supporting ecosystem functions that lead to the provision of ecosystem services. The sources of benefits might come from components of biodiversity (i.e., individual species) or from the complex interactions among biotic and abiotic elements of ecosystems. Second, what is the potential for markets to provide incentives for the conservation of biodiversity? Market forces often lead to the loss of biodiversity. However, when biodiversity conservation is important for human well-being, there is the potential for appropriately designed market mechanisms to maintain or enhance biodiversity through payments that reflect value to ensure the continued supply of benefits.

Direct and Indirect Value

Some components of biodiversity are valuable because they contribute directly to human well-being. People value particular plants and animals because they provide nutritious food or medicines. People may also value particular plants and animals because they have spiritual or symbolic value. Changes in other components of biodiversity, such as the abundance of soil microorganisms, do not have a direct impact on human well-being.

Many things that people do not value directly, however, still have value. Intermediate goods and services are inputs into the provision of final goods and services that people directly value. Soil microorganisms have value because they are essential for plant growth and without which there would be no agricultural crops. The interactions of organisms in a system ultimately determine the provision of ecosystem services that humans directly value. Thus, nearly all components of biodiversity are likely to have at least indirect value because of the role they play in determining the functioning of ecosystems that support the provision of final ecosystem services that have direct value to people.

An analogy can be made between biodiversity and economic infrastructure. Most people do not care directly about infrastructure, but deterioration of infrastructure will cause declines in the provision of final goods and services. Just as ball bearings indirectly contribute to well-being because they are important intermediate goods, soil microorganisms and other components of biodiversity are important components in the production of ecosystem services on which society depends. Ideally, the indirect and direct value of biodiversity would be reflected in markets so that society could be assured that economic activity took proper account of the costs and benefits of degrading or enhancing biodiversity. As will be discussed, however, this is not always the case.

Economic valuation methods are well suited for capturing the values of final goods and services. However, much of the value of biodiversity comes indirectly through the complex interactions among sets of organisms or functional groups

that collectively determine ecosystem functions that lead to provision of ecosystem services. It is more difficult to use economic valuation methods to uncover the indirect value of biodiversity. Valuing the contributions of biodiversity to the provision of ecosystem services requires understanding not only the value of final goods and services but also the ecological role that biodiversity plays in the provision of these goods and services. Only integrated ecological and economic analysis can uncover the indirect value of biodiversity.

The Contributions of Biodiversity to Human Well-Being

Commodity Production and Ecosystem Services

Biodiversity plays an important direct role in the production of commodities, such as agricultural crops, timber, and fish. Here it is the particular species (e.g., maize and bluefin tuna) that have value. This type of value is often readily observable because harvesting these species yields marketed commodities with explicit prices. These prices also reflect the indirect value of biodiversity. In addition, biodiversity plays an important indirect role in determining ecosystem functions on which valuable ecosystem services, such as the bluefin tuna fishery, depend.

If species diversity promotes ecosystem function and this enhances the production of ecosystem goods and services, then species diversity provides a significant indirect benefit to society. A number of experimental studies, primarily but not exclusively in grassland ecosystems, have found a causal relationship between species richness and ecosystem function. Although some early results led to questions regarding the extent to which the apparent relationship could be explained instead as a statistical inevitability of having more species (see Grime, 1997), more recent work demonstrates functional effects of species diversity beyond those expected from selection effect or portfolio effects (Tilman *et al.*, 2001). Although a debate persists regarding the relative importance of these effects, evidence suggests that biodiversity improves human well-being through contributions to ecosystem function that underpin ecosystem services.

On the face of things, biodiversity and the production of agricultural commodities appear to be at odds. Monoculture agriculture is the antithesis of diversity. All else equal, greater diversity appears to mean less of the most valuable species, and it may also preclude economies of scale in production. However, evidence from experimental plots has demonstrated "transgressive overyielding," the production of biomass at higher levels of biodiversity in excess of yields from the best-performing monoculture (Nyfeler *et al.*, 2009). These effects appear to be the results of important ecological interactions among species and to increase over time (Allen *et al.*, 2011; Cardinale *et al.*, 2007; Fornara and Tilman, 2008; Nyfeler *et al.*, 2009; Tilman *et al.*, 2001, 2006).

Studies have yet to demonstrate transgressive economic overyielding (i.e., superior profitability). For increases in biomass to translate into increases in value of production, the percentage increase in quantity must exceed the percentage decrease in average value per unit of biomass, net of harvest costs. Research to date leaves open the possibility that more diverse systems could generate larger returns for agricultural commodities, timber, or fish. Much less is known about the diversity–function relationship in forest and marine systems due to the logistical difficulties of establishing rigorous controlled experiments in such systems.

In the same way that biodiversity has been shown to increase the production of biomass, it may also enhance other ecosystem functions with implications for human well-being. The more "functional contexts" considered, the greater the number of species that appear to play important roles in ecosystem service provision (Isbell *et al.*, 2011), which itself suggests an indirect value for biodiversity. However, increased ecosystem function does not necessarily mean increased quantity or value of ecosystem services. Only a subset of the components or characteristics of a system will be affected by changes in function, and only a subset of these provide direct benefits to society.

Such caveats notwithstanding, biodiversity does appear to be important to ecosystem services. Enhanced biodiversity has been shown to lead to greater soil carbon storage (Fornara and Tilman, 2008; Tilman *et al.*, 2006), which helps to mitigate the release of carbon dioxide to the atmosphere. Greater species diversity has also been shown to increase nitrogen uptake (Fornara and Tilman, 2008; Reich *et al.*, 2001; Tilman *et al.*, 1996), reducing the flow of nitrogen pollutants to water bodies.

Risk and Insurance Benefits

Beyond the possibility of direct and indirect contributions to mean levels of commodity production and other ecosystem services, biodiversity may also reduce the dispersion of returns over time. In experimental plots, diversity has been shown to increase ecosystem stability and decrease the variability of yields more so than an equivalent portfolio of monocultures (Tilman *et al.*, 2006). Biodiversity can also be an important factor over the longer term by maintaining varieties or species better suited to stochastic future conditions (Brock and Xepadadeas, 2003). By reducing risk – for instance, risk associated with weather fluctuations, pests, and pathogens in an agricultural setting – biodiversity contributes an insurance value to society (Perrings *et al.*, 2009). For individual farmers, whether the value of mitigating risks specific to them (i.e., idiosyncratic risk) via enhancing biodiversity is worthwhile depends on costs in terms of reduction in expected returns and the availability of traditional insurance. Even in cases of systemic risk, for which markets do not provide insurance, the stability-enhancing benefits of biodiversity may or may not outweigh the costs of increasing diversity.

Uncertainty, Irreversibility, and the Benefit of Options

Biodiversity provides benefits to society in the face of uncertainty, where the probabilities of certain outcomes are unknown, as well as in the face of risk, where probabilities can be assessed. We have uncertainty, for example, over ecosystem conditions in 2100 or beyond given global change and over the preference of future generations living in a changed

environment. Where the options available to future generations depend on continued existence of particular species, and where agricultural and land-use decisions can have irreversible effects on the continued existence of many species, biodiversity has an "option" value (Fisher and Hanemann, 1986). We would be willing to pay a certain amount of money today to retain the option of having the species in the future, even if it provides no instrumental value today and no certain instrumental value in the future.

Knowledge Benefits

We also have uncertainty over the benefits that undiscovered or unstudied species might bring. This uncertainty is perhaps most salient in the case of the knowledge benefits that such species might provide. Although knowledge benefits ultimately derive from particular species, uncertainty over which species will provide those benefits creates value for diversity *per se* (Polasky and Solow, 1995). Of particular interest has been the potential pharmaceutical value of compounds found in certain organisms, although the pharmaceutical value of the marginal species is likely to be quite small (Simpson *et al.*, 1996). In a similar vein, agronomists use knowledge of many different species to breed crop varieties with superior yield and/or resistance characteristics.

Biodiversity can also generate knowledge benefits in a variety of contexts in which engineers seek efficiencies through biomimicry, the emulation of the forms or processes that other organisms have evolved for "solving" problems. Biomimetic products under development include drag-reducing coating for marine vessels based on water fern, adhesive tape that mimics the hairy feet of insects, and artificial "leaves" capable of splitting hydrogen and oxygen from water molecules.

Esthetic, Spiritual, Cultural, and Recreational Benefits

Biodiversity directly contributes to human well-being in the form of esthetic, spiritual, cultural, and recreational benefits. People value "charismatic megafauna" (e.g., lions, tigers, and whales), and they value flora and fauna that defines regions in which they live or travel. A number of studies have shown that people are willing to pay more for property in more biodiverse areas. For example, Bark *et al.* (2009) find that the diversity of riparian vegetation increases nearby home values in Tucson, AZ, USA. This suggests that diversity, along with other characteristics of the riparian ecosystem, contribute indirectly to one or more amenities. That biodiversity contributes to human well-being is also evident in many recreational contexts. Ecotourism in places like the Great Barrier Reef and the Galapagos Islands are testament to the value that people place on biodiversity. To the extent that people value the chance to see particular rare or exotic species, biodiversity contributes indirectly to human well-being through the support of those individual species. However, there is also evidence that ecotourists value variety (diversity) *per se*. For instance, Naidoo and Adamowicz (2005) find ecotourists are willing to pay more to visit African rainforest reserves with higher avian species diversity, all else equal. There is also considerable evidence that people value knowing that species exist even when they do not expect to observe them or directly benefit from them in any way ("existence value").

Markets, Market Failure, and Policy for Biodiversity

Markets and Efficiency

Markets work by bringing together consumers who wish to buy goods and services with producers who wish to sell goods and services. Markets can be an effective way to provide information across a vast decentralized interconnected web of producers and consumers. Market prices provide signals to producers about how valuable the good or service is to consumers. Similarly, market prices provide signals to consumers about what it costs producers to provide the good or service. If the value of a good or service to consumers is greater than its cost to produce, there are potential gains from trade and a market price will provide an incentive for producers to supply the item to consumers. The first fundamental theorem of welfare economics states that when producers and consumers interact in competitive markets with full information and without externalities, the market outcome will be Pareto efficient, i.e., the only way to make one person better off comes at the cost of making someone else worse off.

When markets work well they provide incentives to produce goods and services at the lowest possible cost and to allocate them to those consumers who value them the most. It is important to note, however, that value depends not only on consumer preferences but also on ability to pay. Those with greater income or wealth can express greater value for goods and services and markets will tend to cater to their needs. Market outcomes, even when efficient, are not necessarily equitable. Market outcomes can result in unequal distribution of income, with the poor not getting much say in what is produced and no guarantee that even their basic consumption needs will be met.

Although market forces have often led to the loss of biodiversity, it is possible to design markets to provide incentives to conserve or enhance biodiversity. Just as with other goods and services, markets can provide incentives to provide valuable ecosystem services. In cases where biodiversity is an input into the provision of ecosystem services, having consumers pay for services provides an incentive to landowners (or others whose decisions affect biodiversity) to conserve biodiversity in order to maintain the continued supply of ecosystem services. This logic has fueled interest in "payments for ecosystem services (PES)" and related market-based approaches to conservation (Engel *et al.*, 2008).

For most goods and services traded in markets, buyers pay sellers the full cost of production and enjoy all of the benefits that the good or service provides. This result occurs when goods and services are rival: the consumption of a good or service by one individual means there is less available for others. It also occurs when the good or service is excludable: the producer can control access to the good or service so that only those who pay can consume. Under these conditions, markets can provide and allocate goods efficiently.

Public Goods and Externalities

Unlike most market goods, the existence of biodiversity and the ecosystem services that biodiversity generates are public goods whose consumption is in many respects nonrival and nonexcludable. The pleasures of observing wildlife, the benefits of climate regulation through carbon sequestration or provision of clean water through nutrient retention are all pure public goods. If these benefits are available to anyone, then they are typically available free of charge to everyone. Markets fail to provide adequate amounts of public goods. Because public goods are nonexcludable, consumers can free ride; they can enjoy the good or service for free as long as it is available. Producers have no incentive to undertake the cost of providing a good or service knowing that they cannot collect revenues to cover costs.

Even when goods and services are not pure public goods, markets may fail to supply efficient levels of provision. Consider the case of fisheries or forests where numerous harvesters can extract resources. Such common-pool resources can suffer from overharvesting because individual harvesters ignore the costs that resource depletion imposes on other harvesters (the "tragedy of the commons"). This case can be thought of as a consequence of externalities, where costs or benefits are not internalized by an individual decision maker.

What Values Are Internalized through Markets?

The degree to which markets will realize the benefits of ecosystem services generated by biodiversity depends on the degree to which those benefits can be internalized. To the extent that biodiversity contributes to rival and excludable private goods already traded in markets, its indirect value will be internalized. For example, the value of soil microorganisms that contribute to soil fertility and enhance agricultural yields is realized by farm owners through increased revenues. Farm owners have an incentive to maintain soil fertility because it enhances their bottom line. However, farm owners do not have an incentive to maintain other forms of biodiversity that enhance ecosystem services that contribute to public goods, such as nutrient retention that improves water quality downstream. The fact that we observe low levels of biodiversity in agro-ecosystems suggests that many of the benefits from maintaining biodiversity are public goods or, alternatively, that land managers are ill-informed about the potential benefits of biodiversity.

Markets can internalize some benefits even when ecosystem services are not fully rival and excludable, as the proliferation of privately owned ecotourism destinations attests. Large benefits create an incentive to make goods excludable so that providers can charge for access. Game ranches can do this on private land by fencing and guarding property boundaries and limiting access to only those who pay for it. Markets for pharmaceuticals and agricultural biotechnology also can capture benefits through patents and contracts. However, because knowledge has aspects of a public good in that it is nonrival, and to the extent that others can invent around patents or pirate patented goods, nonexcludable, the market will tend to under-supply these knowledge-driven benefits.

When markets fail to capture the value of biodiversity, the perceived opportunity cost of biodiversity loss will be too low, and the decline of biodiversity too high. Correcting this imbalance requires intervention by governments to "get the prices right," either by using taxes and subsidies or by establishing regulations and property rights that help to create more efficient markets.

Internalizing Benefits through Policy or Other Means

Ironically, one answer to market failure resulting from public goods or externalities is the creation of new markets with the help of public policy. Market failures can be the result of the failure of markets to exist. For instance, the creation of a market for carbon would force producers and consumers to internalize the costs of carbon emission and the benefits of carbon storage. A governmentally imposed market for carbon, in addition to affecting decisions about energy production and use, would affect land markets by making lands with greater carbon sequestration and storage potential more valuable, and thus internalize the indirect value that biodiversity plays in carbon storage. Similarly, PES program focused on water quality or habitat provision would enhance the value of biodiversity conservation to the extent that doing so enhances the provision of these services. Programs such as the Conservation Reserve Program in the USA and PES (Pago de Servicios Ambientales) in Costa Rica compensate landowners for management regimes meant to enhance provision of ecosystem services and help to internalize some of the societal benefits of biodiversity conservation that landowners otherwise do not capture.

Although markets may offer economic efficiency, it may not be practical or feasible to create markets to provide incentives to maintain biodiversity. Markets require functioning institutions that can enforce property rights and, in the case of ecosystem services like water purification or climate regulation, the ability to effectively monitor the contribution of producers to the provision of ecosystem services. The ability to monitor and enforce may be lacking in countries with weak governance or simply too costly relative to the benefits of establishing a market.

Other means to try to internalize benefits include community-based programs of self-regulation and certification programs. Some communities have managed to sustainably manage resources such as fisheries, communal grazing lands, and forests through self-regulation, social customs, and norms (Ostrom, 1990). These work best in close-knit, small, and stable communities where people can oversee the actions of their neighbors. Finally, certification efforts aim to provide consumers with information about how the products they buy are produced in a sustainable manner. If consumers are sufficiently environmentally aware and back up this awareness with their pocketbooks, then certification can provide incentives for biodiversity conservation through the sustainable production of goods and services. However, there is no guarantee that consumers will not free ride and simply choose goods and services that are the least expensive.

Finally, society may opt for direct government regulation ("command and control") to preserve biodiversity for the sake of human well-being. Government regulations can outlaw destructive practices or actions detrimental to biodiversity.

Bans on trading listed species across international borders under the Convention on Trade in Endangered Species is one example. However, failure to provide adequate resources for enforcement will drive activity into the black market rather than eliminate it. Moreover, the costs associated with such regulations may exceed their benefits, in which case human well-being will decline overall even if biodiversity is enhanced.

Appendix

List of Courses

1. Environmental Economics
2. Ecological Economics
3. Conservation Biology
4. Natural Resource Management

See also: Economic Value of Biodiversity, Measurements of. Ecosystem, Concept of. Property Rights and Biodiversity

References

Allen E, Weisser W, Weigelt A, Roscher C, Fischer M, and Hillebrand H (2011) More diverse plant communities have higher functioning over time due to turnover in complementary dominant species. *Proceedings of the National Academy of Sciences of the United States of America* 108: 17034–17039.

Bark RH, Osgood DE, Colby BG, Katz G, and Stromberg J (2009) Habitat preservation and restoration: Do homebuyers have preferences for quality habitat? *Ecological Economics* 68: 1465–1475.

Brock WA and Xepadadeas A (2003) Valuing biodiversity from an economic perspective: A unified economic, ecological and genetic approach. *American Economic Review* 93: 1597–1614.

Cardinale BJ, Wright JP, Cadotte MW, *et al.* (2007) Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences of the United States of America* 104: 18123–18128.

Engel S, Pagiola S, and Wunder S (2008) Designing payments for environmental services in theory and practice: An overview of the issues. *Ecological Economics* 65: 663–674.

Fisher AC and Hanemann WM (1986) Option value and the extinction of species. In: Smith VK (ed.) *Advances in Applied Micro-Economics*. Greenwich, CT: JAI Press.

Fornara DA and Tilman D (2008) Plant functional composition influences rates of soil carbon and nitrogen accumulation. *Journal of Ecology* 96: 314–322.

Grime JP (1997) Biodiversity and ecosystem function: The debate deepens. *Science* 277: 1260–1261.

Isbell F, Calcagno V, Hector A, *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature* 477: 199–203.

Naidoo R and Adamowicz V (2005) Economic benefits of biodiversity exceed costs of conservation at an African rainforest reserve. *Proceedings of the National Academy of Sciences of the United States of America* 102: 16712–16716.

Nyfeler D, Huguenin-Elie O, Suter M, Frossard E, Connolly J, and Lüscher A (2009) Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. *The Journal of Applied Ecology* 46: 683–691.

Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.

Perrings C, Baumgärtner S, Brock WA, *et al.* (2009) The economics of biodiversity and ecosystem services. In: Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*. New York: Oxford University Press.

Polasky S and Solow A (1995) On the value of a collection of species. *Journal of Environmental Economics and Management* 29: 298–303.

Reich PB, Knops J, Tilman D, *et al.* (2001) Plant diversity enhances ecosystem responses to elevated CO_2 and nitrogen deposition. *Nature* 410: 809–812.

Simpson RD, Sedjo RA, and Reid JW (1996) Valuing biodiversity for use in pharmaceutical research. *Journal of Political Economy* 104: 163–185.

Tilman D, Reich PB, Knops J, Wedin D, Mielke T, and Lehman C (2001) Diversity and productivity in a long-term grassland experiment. *Science* 294: 843–845.

Tilman D, Reich PB, and Knops JMH (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441: 629–632.

Tilman D, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379: 718–720.

Biodiversity Informatics

Robert A Morris, University of Massachusetts Boston, Boston, MA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Circumscription In the context of a name, the collection of things designated by that name.
Controlled vocabulary A set of terms, typically with an assigned meaning, used in datasets or for data exchange, to promote ease of data comparison.
Data-centric Particularly concerned with data.
Electronic resource Data or information in electronic form, especially accessible on the Internet. Also applied to Internet services, q.v.
EXIF Acronym for Exchangeable Image File Format. The full name is rarely used. It describes data often captured by digital cameras in addition to the image.
Geospatial Concerned with locations on the Earth.
GO Acronym for the Gene Ontology. This is a formal ontology used as a controlled vocabulary for gene annotation.
IANA Acronym for the Internet Assigned Numbers Authority. It is responsible for administering the syntax and use of named and numbered objects on the internet, such as the addresses that identify particular computers connected to the Internet.
IETF Acronym for the Internet Engineering Task Force, a body that promulgates standards for technical aspects of the Internet.
Informaticist One who practices informatics.
Informatics The practice of designing, implementing, deploying, and managing computerized information systems.
Internet service A remote software program on the Internet accessed by a user or by another program and which provides for some action to take place, or some information to be provided.
JPEG Acronym for Joint Photographic Experts Group. It describes a particular file format for digital images. The full name is rarely used.
Ontology A controlled vocabulary, especially machine readable, that provides for standard terminology, including relationships between the terms, for use in data description.
Optical character recognition The process by which an electronically scanned image of text is separated into individual characters and made available for electronic text processing such as by word processors. Often written only in its acronym, OCR.
Provenance A record, often electronic, of the origin of, and changes to data. Data provenance may include a history of the changes to the data, an authority for the data and its changes, and a way to recover the original.
Time stamp An electronic attribute of a data record or dataset that indicates when it was created or the last time it was changed.

History

Data is older than writing itself and arguably arose as early as the eighth millennium BC in the form of clay tokens that probably signified quantities of agricultural goods. Possibly as early as the sixth, but unambiguously by the third millennium BC, curation of that data had become more refined by sealing it inside clay envelopes with descriptions of the contents stamped on the outside. We can identify the earliest form of metadata in those stampings – the description of data or as it is now often characterized, data about data. The ubiquity of computerized acquisition, storage, and exchange of data creates a critical need for high-quality metadata, which is itself computerized. While many scientists recognize the need to be able to exchange electronic data in standardized formats, only recently is the awareness rising about the need for standards for the representation of metadata.

Many data-centric communities have metadata standards on which metadata standards for biodiversity information have been modeled. Historically, standardized electronic metadata is unambiguously recognized in the electronic cataloging of library collections promulgated in the Dublin Core Metadata Initiative (popularly known as the Dublin Core), organized originally by the Online Consortium for Library Cataloging (OCLC). Dublin Core is a standard for describing library collection items (whether electronic or not). Natural history museums, herbaria, and botanical gardens have collections whose curation shares much structure with library collection management. Thus it is not surprising that the electronic description of specimens represents early standardization efforts by taxonomists, who, in at least one currently important case, emulated the OCLC efforts with a now widely used standard called the Darwin Core. This standard and a related, more comprehensive specimen description standard with the acronym ABCD (Access to Biological Collection Data) may be considered as modeling specimens as data, and therefore these standards merit the title metadata. Life scientists might argue this point, but computer scientists rarely distinguish data from metadata other than that metadata is data about some other data. Other early metadata standards of importance to biodiversity arose from governmental efforts to provide for the interchange of geospatial data as this became more easily accessible in electronic form throughout the last decade of the twentieth century. Recently the International Standards Organization (ISO) has established international standards for representing geospatial metadata – including the so-called country profiles, which recognize special governmental geospatial metadata needs of member countries.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00262-8

Because so much geospatial data is provided by governments, scientists can expect that systems worldwide will begin to be compliant with the ISO standards.

Kinds of Metadata

Although there is no widespread agreement on any classification, metadata can be roughly divided into several kinds: acquisition, format, curatorial, normalization, and descriptive metadata. Thus, the Ecological Metadata Language (EML) recognizes information about where and when a certain set of data was gathered (acquisition metadata), who is responsible for its care (curatorial metadata), what are the units of measurement and permitted values for each field (normalization), whether the data is numeric or text (format), and what observations or measurements the various data represent (descriptive). Scientists have long recognized the need for standards in the expression of this sort of information, but all too often do not provide standardized metadata for their computerized data. For example, early in scientific education one is taught about the International System of Units (SI, from the French *Système international d'unités*) and a set of standard prefixes for scaling the fundamental units. This is a prime example of a standard for the expression of metadata, and without appeal to it or something equivalent, a table of measurements comprising only numbers is useless. Yet in computerized data, such tables are sometimes all that are recorded. Another kind of metadata, which allows the user to trace the history of changes to a database, is sometimes called "provenance" metadata. Although studied by computer scientists, its importance has received little attention among biodiversity informaticists, most of whose metadata standards presently express change in history only weakly.

Purposes of Metadata

Suitability of Data for Scientific Purpose

In the absence of adequate metadata, it may be quite difficult to decide whether a particular dataset is suitable for a particular scientific inquiry, or to know how it might be compared with data from another source. For example, the US Geological Survey offers vegetation coverage data for the Arctic National Wildlife Refuge, with a carefully constructed metadata file containing, among other things, the descriptive names of 23 classes of land cover that are represented in the satellite-derived data, along with a reference to a technical report describing how that data was deduced from the raw remote-sensing data. With this metadata, a climate change investigator might better support a study investigating change in vegetation correlated to climate change data than if only the raw remote-sensing data was considered. In such cases, the metadata plays a role much like the familiar methods section of traditional scientific publications; it allows those considering the data analysis to decide whether the data was properly used.

Data Discovery

On the Internet it is often quite difficult to locate and exploit electronic resources, especially those hidden behind databases (sometimes called the deep web). This problem is even more severe when the search is not controlled by a human but rather by a semiautonomous software application, for example, an ecological modeling program using data distributed around the Internet. Although a human usually controls data access, when the data is in electronic form it is always accessed in the first instance by software, such as a web browser or a database application package. For control by software, two important pieces of metadata are always needed: where to locate the data and how to access it.

Data access software can be made substantially more powerful if it can automatically discover the location, access methods, and standard metadata of new data sources that have become available subsequent to the deployment of the software. Several solutions to this are common. Among them are data portals, and dereferenceable globally unique identifiers. The latter have become particularly useful for "Linked Open Data," a Semantic Web technology described below.

Data Portals

A data portal is a single Internet site at which the data discovery and delivery tasks are delegated to software. Superficially, the data portal eliminates the data discovery problem. In reality, the portal software has merely taken the discovery burden off the user (whether human or software) onto itself, often by some form of metadata registry access. Expertly run, this can be very useful. For example, The Global Biodiversity Information Facility (GBIF) operates a portal that aggregates several hundred million records that document the occurrence of a species, based either on a specimen in a natural history collection, or on a field observation. These records come from several hundred different data providers that software or humans can query for more detailed information.

Globally Unique Identifiers

Substantial attention is now being given to a special kind of metadata to assist in the discovery of electronic resources. This is the Globally Unique Identifier (GUID). One GUID scheme widely used by publishers is the Digital Object Identifier (DOI). Like all GUID schemes, DOIs must be issued according to some process that guarantees no two electronic objects are ever assigned the same DOI, and a nonprofit organization is established to do that. The Biodiversity Information Standards (TDWG) organization (named from its historical name, the Taxonomic Databases Working Group) and GBIF have explored a different GUID scheme for their data-exchange standards. This is the Life Sciences Identifier (LSID), which is also in some use for genomics data. GUID is a rather generic term, with no standards for its use. The Internet Engineering Task Force (IETF) has promulgated a standard for a kind of GUID called a Uniform Resource Identifier (URI), along with a procedure for registering the syntax of a URI with the Internet Assigned Numbers Authority (IANA). Both DOI and LSID are registered URI schemes. An increasingly important registered URI scheme long in use on the Internet is that based on the familiar syntax for the address of web sites, the

so-called "http URI scheme." All of the examples are "dereferencable URIs," meaning that from the URI alone, software can reliably discover and retrieve datasets or even individual records from, or at least further information about, the online resource identified by the URI.

Data Provenance and Quality Control

Because taxonomic synonymy may, for example, make a binomial species name ambiguous, data can for some purposes have its utility impaired unless some authority for the name is given. When the data involve the identification of a taxon, an ecosystem type, or identification appealing to some other classification of the objects represented by the data, the metadata often contains the authority for the classification as well as for the particular identification. Particularly important, but often overlooked, it can be critical in such a case to have metadata indicating the date at which the classification was made. This is because the names of things – for example, taxa and places – can have their circumscription change over time by such activity as taxonomic revision or geopolitical events. In the absence of such time stamp metadata, it may be impossible to tell exactly what is described by a name found in a dataset. Time stamps are an example of what is called record-level metadata, because they often must be applied to each data record rather than to the dataset as a whole unless changes to individual data entries are prohibited. Other important provenance metadata might include information about the authority for a taxonomic determination and the original source of the record in the case that a third party, for example, a portal, is providing the data.

Data provenance can help provide evidence for the quality of individual data records by providing a trail to the original source of the data, and to changes that may have been made. However, the use of provenance is not the only way to automatically find some kinds of data errors, such as those arising from human data entry or from programming errors. Systems that apply rules about the semantic nature of the data are also often employed to find logical inconsistencies in the data. For example, if a dataset of species occurrence contains an attribute describing whether the species is terrestrial or marine, then software with access to the boundaries of marine bodies can detect such errors as the exchange of latitude and longitude values in a dataset that might result in data showing the occurrence of a terrestrial organism in an ocean.

Annotation

In genome science it is commonplace to annotate data with additional information about a gene's function, its relation to other genes, or other attributes beyond just its nucleotide sequence. Data searches based on such annotation can help find patterns in genomic data that represent a higher level of knowledge than the sequence alone. Recent informatics research efforts have begun to explore biodiversity data record annotation using approaches and technologies similar to those of genomics, with the expectation of similar benefits.

Beyond Spreadsheet Headings: Structured Metadata

Obstructions to Machine Integration of Heterogeneous Biodiversity Data

Even when describing the same scientific concerns, for example, the same group of taxa and the same ecological questions, scientists often have experiment, modeling, or collection protocols that suit their precise investigation. Variation in these gives rise to heterogeneous data, which can be very difficult to share and combine. In the past, this problem was often addressed with tedious cooperation between database administrators in ways that might require special-purpose manipulation that often did not survive evolution of the database, such as adding new data tables. With dramatic decreases in the costs of gathering data and making it available immediately on the Internet, these *ad hoc* data integration methods are no longer satisfactory and new support for data integration is emerging.

There are several requirements for addressing this problem technologically. The most important is controlled vocabularies within disciplines to provide common terminology for representations of data and metadata. The second is a common way to express that terminology – a sort of lingua franca – suitable for computer processing. A widely accepted solution to the latter problem has arisen to meet the needs of non-scientific users of the Internet – principally those engaged in electronic commerce – that has been embraced by many scientific communities. It is a computer language called the eXtensible Markup Language (XML). Correspondingly, many scientific groups have been organized to produce common terminologies expressed in XML.

XML and Data on the Internet: Metadata for Metadata

XML is now a ubiquitous way of expressing data in an extremely portable fashion, making it particularly easy to share data on the Internet. XML data is often characterized as self-describing, which in one sense can be considered as carrying its metadata with it. It does so by surrounding data with tags carrying descriptive metadata. XML is expressed in human-readable form, but in reality most XML is too complex for humans to read it without special software. However, because so much Internet technology has been devoted to human-to-human communication, human-readable forms are more reliably exchanged than many specially encoded forms of computerized data. This fragment of a type specimen record for the oak *Quercus obtusata* illustrates the point. It is expressed in the Darwin Core XML form. The tags are the expressions in brackets.

```
<response xmlns="http://digir.net/schema/protocol/2003/1.0">
<header>
<version>1.0.0</version>
<sendTime>2006-08-17T05:25:19</sendTime>
<source>
resource="rsr25d058f8a625b6c7b2573e2ed947f9e9">
http://192.38.112.87:8079/botanik/DigirProvider/index_html
</source>
```

</header>
<content xmlns:darwin = "http://digir.net/schema/conceptual/darwin/2003/1.0"
xmlns:xsd = "http://www.w3.org/2001/XMLSchema"
xmlns:xsi = "http://www.w3.org/2001/XMLSchema-instance" >
<record>
<DateLastModified>2005-02-04T18:24:32</DateLastModified>
<BasisOfRecord>Herbarium specimen</BasisOfRecord>
<ScientificName>Quercus obtusata</ScientificName>
<Subspecies>var. brevipes</Subspecies>
<CatalogNumber>13367</CatalogNumber>
<Kingdom>Plantae</Kingdom>
<ScientificNameAuthor>Oerst.</ScientificNameAuthor>
<Genus>Quercus</Genus>
<Phylum>Vascular Plants</Phylum>
<CollectionCode>V</CollectionCode>
<Notes>On sheet: Liebm. Pl. Mex. nr. 3526.
Quercus obtusata Thumb.
et Bonpl. gamma Hartwegi (Benth.)
Alph. DC. Syn. Quercus Hartwegii Benth. S. Felipe; 5/42.
Leg. Liebmann. Label1: 3526 Very?? Liebmannii
Label2: California Academy of Sciences Isotype</Notes>
<InstitutionCode>C</InstitutionCode>
<TypeStatus>Isotypus</TypeStatus>
<Species>obtusata</Species>
</record>
</content>
</response>

Source: The Darwin Core data record for *Quercus obtusata* was obtained from the Type database of the Danish Biodiversity Information Facility.

If one takes the liberty of considering the specimen itself as data, then the text and numbers between the tags are a kind of metadata for that data, and the tags are metadata for that metadata. In practice, the specimen record would normally reside in a standard database, in which case the XML tags can then more clearly be seen as metadata, for example, the names of attributes in a relational database. The Darwin Core metadata is not arbitrarily organized. Instead it corresponds to a published set of structural rules playing much the same role as the taxonomic nomenclatural codes play for taxonomic names. In this case, these rules are themselves expressed in XML, in a form called XML-Schema, and accepted as an international standard by TDWG. The Darwin Core is employed in the GBIF portal, as is a related but more complex XML-Schema, the ABCD, both mentioned earlier. Tightly controlled by the XML-Schema rules, these standards make it possible to combine and compare data across species, data providers, locations, or any of the other kinds of data they represent.

Ecological Metadata Language

EML is the metadata standard most directly addressing ecological data. It has many components of each type of metadata mentioned earlier in this article (*see* Kinds of Metadata). Of special note, the organization promulgating EML – the Knowledge Network for Biocomplexity – also provides software that supports metadata repositories and catalogs (and even collaborative data repositories) in ways that make it easy for scientists to maintain their data in their own database systems or spreadsheets. Although EML is represented in XML, these software products make it unnecessary for scientists to know about the inner workings of either XML or EML.

Below is a fragment of a metadata item represented with EML tags. It expresses where, how, with what kind of numeric measurement, and with what units and precision a certain temperature is regularly recorded.

<attribute scope = "document" >
<attributeName>NoonTempAtDock</attributeName>
<attributeLabel>Local noon surface temperature</attributeLabel>
<attributeDefinition>Surface temperature at dock at local noon, measured with DataTemp Model 73 Temp Logger
</attributeDefinition>
<storageType typeSystem = "http://www.w3.org/2001/XMLSchema-datatypes" >float<storageType>
<measurementScale>
<interval>
<unit>
<standardUnit>Celsius</standardUnit>
</unit>
<precision>0.001</precision>
<numericDomain>
<numberType>real</numberType>
</numericDomain>
</interval>
</attribute>

Darwin Core and BioCase Access to Biological Collections Data Metadata Standards for Specimens

The Darwin Core and the ABCD metadata standards both aim to describe natural history specimens and both have been extended to descriptions of field observations of organisms. These two standards exhibit a common tension in the design of metadata standards: whether to be economical or to be inclusive. For example, the ABCD standard admits seven different kinds of specimens or observations (microbial, mycological, herbarium, botanical garden, plant genetic resources, zoological, and paleontological), each with a different structure. By contrast, Darwin Core makes no such distinctions. While it is easier to conform to a simpler standard – perhaps leading to wider acceptance – it is correspondingly more difficult to use metadata thus expressed to locate data with specific properties. For example, using Darwin Core could require special knowledge about the nature of institutions holding the records to ascertain whether a particular plant specimen is alive or not, whereas in ABCD such specimens would typically be tagged as being in a botanical garden.

Metadata Standards for Geospatial Data

Locations and geographic extent of biota, populations, and ecosystems are so central to biodiversity information that access to geospatial data plays a key role in many biodiversity

information systems. All the standards discussed above have some way to describe geospatial data, but often their commonality lies only in the simplest expressions, such as that for the latitude and longitude of a point and an error radius. The Open Geospatial Consortium (OGC) has developed not only standards for geospatial data and metadata, but also guidelines for how to exploit those standards in the development of other standards. A somewhat simpler standard was proposed in 2010 by the IETF and is finding wide use for web-accessible data, particularly where only point data is available.

Emerging Metadata Technologies and New Applications of Metadata

The Semantic Web: Machine Reasoning about Comparing Databases

An increasingly important role of biodiversity metadata arises from the ease with which data is now shared on the World Wide Web coupled with the difficulty of locating and evaluating the utility of such data for any particular scientific inquiry.

In an attempt to mitigate issues surrounding the aggregation of datasets on the Internet with widely differing descriptions of their data, the World Wide Web Consortium (W3C) has promulgated a metadata language with the acronym RDF (Resource Description Framework). This specialized language (most often expressed in XML) builds on a history of computer science research dealing with diverse descriptions of the same thing. A further specialization of RDF (together with some other W3C XML standards) forms the Web Ontology Language (OWL). Computer ontologies support the expression – in this case in XML – of the objects of a discipline, their properties and the relations between those properties, between objects, and between classes of objects. Other ontology expression languages important for biological sciences include those of the Unified Medical Language System (UMLS) and of the Open Biomedical Ontology (OBO) format. The well-known Gene Ontology (GO) has expressions in OBO and OWL.

An important focus of ontology languages is their support for machine reasoning. By use of relations expressed in OWL or OBO, and to a lesser extent in RDF in general, it is possible for computer software to draw conclusions about the relations between data records. This holds the promise of successful data discovery and integration when different data sources have different controlled vocabulary in their metadata or when the seeker has a different way of conceptualizing the data than might be represented in available databases. In some simple cases, machine reasoning can even discover inconsistencies within or between datasets that might otherwise lie hidden. Somewhat more advanced is the capability for such software to test whether a given dataset is logically consistent with one or another model expressed as an ontology. This promise is in its infancy though, in part because it is quite easy when using these technologies to build ontologies that are computationally intractable. Ongoing biodiversity informatics efforts in ontology development and application hold promise in describing such diverse studies as phenotypic variation, ecological relations, and the application of environmental sensor data to ecological investigations.

RDF has found use in a simple set of Internet best practices collectively denoted "Linked Open Data (LOD)." These practices make it easy for online data holders to link their datasets to other related ones. The resulting "web of data" comprising all these links make it easy to provide services that an ordinary web browser can navigate. More sophisticated browser applications allow the user to choose the data relations that they wish to guide their browsing. Such facilities reduce the burden and increase the utility of scientific data sharing.

Extracting Knowledge from Scientific Literature

XML is derived from a computer tagging language called the Standard Generalized Markup Language (SGML), which dates to research in the 1960s at the IBM Corporation. SGML was designed primarily to tag (or "mark up") electronic documents. In this context, the tags are often called document markup. Such tags help identify the structure and meaning of the document, and – using XML – are now being rapidly applied to digitized forms of existing scientific and other scholarly and cultural literature. In this process, images of the pages are made by the use of scanners, and optical character recognition turns these images into plain text. The markup is applied by knowledgeable humans, or increasingly often, by software programs with human intervention only for proofreading and correction. Nevertheless, without full automation the markup task is immense. By the end of 2010, the Biodiversity Heritage Library had scanned nearly 33 million pages of published literature, but the best present markup correction tools require 2–3 min per page of effort by biologists. As with data, such markup forms a kind of descriptive metadata for the form and content of the documents and enhances the ability of the software to discover and extract information from literature not previously available in digital form. Recent advances in electronic publication make it easier to do this as part of the production of born-digital publications, and the resulting markup can provide such things as intelligent searches of newly published biosystematics literature.

Amplifying the Utility of Digital Media

An increasingly important form of data consists of digital media such as images, sounds, and movies. Many of the digital formats for the exchange of these media support limited metadata standards, which, moreover, can be embedded in the digital representations along with the data, making the metadata curation more robust. For example, most support an early metadata standard with the acronym EXIF that consists mainly of acquisition metadata that is often automatically added to the media by the acquisition system itself (including, e.g., acquisition date, image size, camera, and lens details). For web-accessible data, the ability of the data to carry its own metadata provides some advantages: once the location and electronic format of the desired data is known, the precise description of the data can, in principle, be extracted from the data itself, much as happens with XML data representations.

EXIF metadata can be manipulated and extracted by many image-processing programs.

A 10-year-old standard for image data files, JPEG2000, is an improved format over the classical Joint Photographic Experts Group (JPEG) format widely used by digital cameras. JPEG2000 can represent still pictures, video, and audio. Arbitrary XML data can be embedded in a JPEG2000 file, which can carry multiple images or other media within the file. In particular, JPEG2000 files can carry extensive metadata, conveying much more information than is specified by EXIF, which has very little provision for descriptive metadata. For example, remote-sensing vegetation images can carry with them the vegetation types corresponding to each color in the image and suitable software could automatically construct a color key when displaying the image, or images of plant specimens could include Darwin Core specimen records with them.

A recently adopted TDWG standard, the Audubon Core Multimedia Resource Metadata Schema makes it easy for providers of biodiversity images, video, and sound to better to describe the content of their media resource. In many cases, this metadata alone can be combined with other species occurrence data for studies such as niche modeling, seasonal variation, climate change, migration, and other investigations that depend on geocoded occurrence data.

Scientific Workflow Systems

Workflow systems are software systems that conveniently allow data analysis to be done automatically with analysis provided by packages chosen and organized by the user. Typically, these provide for inclusion of popular statistical and modeling packages that are already familiar to scientists. They often make use of the metadata for their inputs to ensure that the data are meaningfully combined by the packages, and some more advanced workflow systems are capable of locating and exploiting data on the internet in combination with the user's own data.

See also: Economic Value of Biodiversity, Measurements of. Framework for Assessment and Monitoring of Biodiversity

References

Edwards JL, Lane MA, and Nielsen ES (2000) Interoperability of biodiversity databases: Biodiversity information on every desktop. *Science* 29(September): 2312–2314.

Guralnick R and Neufeld D (2005) Challenges building online GIS services to support global biodiversity mapping and analysis: Lessons from the mountain and plains database and informatics project. *Biodiversity Informatics* 2: 56–69.

Hey T, Tansley S, and Tolle K (2009) *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Redmond, WA: Microsoft Research.

Jones MB, Berkley C, Bojilova J, and Schildhauer M (2001) Managing scientific metadata. *IEEE Internet Computing* 5(5): 59–68.

Leong LKW, Coddington P, and Wendelborn A (2005) Data grid services for biodiversity informatics. *International Journal of Information Technology* 11(3): 29–39.

Michener W and Brunt J (eds.) (2000) *Ecological Data – Design, Management and Processing*. Malden, MA: Blackwell Science.

Page RDM (2006) Taxonomic names, metadata, and the Semantic Web. *Biodiversity Informatics* 3: 1–15. ISSN: 1546-9735.

Schmandt-Besserat D (1992) *Before Writing*. Austin: University of Texas Press.

Zhao Y, Raicu I, and Foster I (2008) Scientific workflow systems for 21st century, New Bottle or New Wine? *IEEE Congress on Services - Part I, 6-11 July 2008*. pp. 467–471. doi:10.1109/SERVICES-1.2008.79

Biodiversity in Logged and Managed Forests

Reinmar Seidler, Ashoka Trust for Research in Ecology and the Environment, India
Kamaljit S Bawa, University of Massachusetts, Boston, MA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Dipterocarp A member of the Dipterocarpaceae family (comprising 22 genera) of South Asian and African timber trees.
Mast fruiting A period or year in which a heavy crop of fruits/seeds is synchronously produced by trees and shrubs: This phenomenon, uniquely characterized by synchronicity, high variability, and periodicity of heavy fruit production, distinguishes it from nonmasting plants.
Natural regeneration The process of replacement by natural seedlings in gaps created by the selective cutting of marketable timber.
Pioneer guild The first groups of species to colonize a newly formed or denuded habitat.
Primary forest Forests that appear to be undisturbed by human influence.
Secondary forest Successional forests growing in areas where in the past the forest cover had been completely removed.
Shifting agriculture The practice of clearing a plot of land for cultivation for a short period of time, then abandoning it and allowing it to revert to its natural vegetation when the cultivation moves to another plot.
Silviculture An applied science and branch of forestry concerned with the theory and practice of controlling forest establishment, composition, and growth.
Succession Replacement of one kind of population/community by another in a habitat through a regular progression over time.
Taungya The Burmese term for an agricultural system in which crops are interplanted with plantations of trees. As the trees grow and shade the areas, cultivation of crops is abandoned.

Introduction

In the nearly 10 years since the original version of this review was written, hundreds of studies of forest management on biodiversity have been performed and many dozens have been published in peer-reviewed journals. The basic conclusions of the earlier review, presented rather tentatively at the time, have generally been strengthened by this increase in evidence; more importantly, significant new generalizations can now be made with some confidence.

Clearly, forests are being managed for a variety of goals. The precise impact of different management systems on biological diversity is still difficult to predict in particular cases, although there have been literally hundreds of studies of forest ecosystems managed for the production of timber over the last decade. Here the authors provide a perspective on changes in biodiversity in forests managed principally for timber. The concepts and principles of large-scale forest management for a more diverse range of goods and services are still under development, although traditional societies have managed forests for such purposes for centuries. Here the authors first outline various systems of management and then describe the extent of knowledge about their effects on forest biodiversity. The effects of these systems vary according to a number of factors, including the type of forest, the landscape matrix in which the forest is embedded, the nature and intensity of the management system, and the scale and method of analysis chosen by the observer. Forests are complex ecosystems with subtle and complex interrelationships, many of them unknown in detail and changing over time. The potential for variation in each of these factors is great, making generalization across specific cases tenuous.

The final section looks at efforts – increasing in number over the last 10 years – to extract general principles from the great multitude of individual studies of the effects of forest management. Individual studies have examined a vast range of forests types, harvesting and management systems, and target taxa. They have been carried out at a range of scales, from microhabitat to landscape, and have measured a broad panoply of physical parameters. Extracting the essential lessons from this mass of material is difficult but urgent, since the major barriers to more consistent and effective forest management remain political and institutional. Overcoming such barriers requires the clear and persistent communication of major trends.

Forest diversity may be examined at various levels, including ecosystem, landscape, population, and genetic levels. It is advantageous to keep these multiple levels in mind as one considers forest management systems. Since the changes effected by management on forest biodiversity do not happen all at once, but are mediated through a chain or sequence of influences both spatially and over time, this discussion is organized sequentially. The authors begin by describing the structural changes imposed on forests by management operations or activities. These structural changes determine the ensuing abiotic or environmental conditions within the forest at the small and medium (local and regional) scales. Structural changes and altered abiotic conditions combine to produce a variety of biotic changes, including changes in species composition and abundance, productivity, population density and distribution, and so forth. It is thought that these changes may also affect long-term ecological interactions and processes. Thus, the original structural alterations to forests result in changed environmental conditions, which in turn trigger

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00248-3

changes at the species, population, and community levels. Large reductions in effective population sizes may result in genetic alterations and impoverishment. In addition, it is important to keep in mind that virtually every managed forest experiences a variety of indirect synergistic anthropogenic effects, namely human activities not part of the management itself but facilitated by it. These may include increased incidence of wildfire, hunting, shifting agriculture, and others. Taken together, these activities may multiply the overall impact of the management system.

Management Systems

Systems of forest management vary widely throughout the world in response to ecological, social–institutional, and political factors. Use of the phrase management system need not imply that management is in each case applied in a planned, thoughtful, or consistent manner, but only that it is systematic in the sense of being a pattern of human manipulative activities governed by a specific common goal. These activities may be carefully planned and controlled interventions or they may be haphazard patterns of one-time individual extractive activities, without much overall guidance. Thus, we use the word management rather broadly.

Timber Management

Clearing Systems

At one extreme of a continuum of management systems for the extraction of timber are systems of forest clearing, with or without replanting to stimulate regeneration. Clearing systems have been used most often in temperate forests. These forests have relatively low numbers of tree species and therefore a relatively high density of commercially useful species per hectare. However, some tropical forests, such as some of the dipterocarp-dominated forests of Southeast Asia, may also have sufficiently dense stands of commercial trees to make clear-cutting an economic option. If there are residual noncommercial trees, they may be girdled or poisoned to produce even-aged stands in which commercial species predominate. There have been experiments with various geometries of clear-cuts, intended to aid in natural regeneration and the preservation of wildlife habitat. Examples are the checkerboard patterns in the Pacific Northwest of the United States and Canada and the strip shelter-belt system of Palcazu in Peru. Patterns of this type have as their goal the reduction in area of the individual gaps, and thus the reduction in distance between the edge of the remaining forest and the center of clear-cut areas. The underlying idea is to increase the likelihood of propagules from the forest finding their way into the clear-cut area and regenerating the stand. There have also been attempts to develop timber management systems that use natural gaps as models, but these are considered selective logging systems (*see* Selective Logging Systems).

Other forest management systems involving clearing are replacement systems, in which plantations of indigenous or exotic species are established after clearing. These have been widely used in temperate forests and less so in the tropics, partly because they often involve a larger initial investment in silviculture. The even-aged stands which result are seen as advantageous in terms of accounting and the efficiency of silvicultural operations, but they have been vulnerable to outbreaks of pests and disease and, as highly simplified ecosystems, have been criticized as deleterious to biodiversity.

Selective Logging Systems

In many species-diverse tropical forests, clear-cutting is neither an economically nor an ecologically appropriate option, since only a few tree species are commercially accepted – although the number of accepted species is rising fast in response to a number of technical and industrial drivers. Moreover, conspecifics are widely scattered, and infrastructure including roads and rural mills may be sparse. In such managed forests, trees are selectively logged. Selective logging is a general term encompassing a wide array of management systems that vary widely with respect to spatial and temporal scale, harvesting intensity, planning, and oversight. Selectivity of species to be removed does not necessarily translate into selectivity of overall impact, so the residual forest may be affected indiscriminately even by the selective extraction of only a few trees per hectare.

Tropical forest management systems were adapted from the Central European forestry tradition, which was exported to India and Burma during the nineteenth century. Natural regeneration systems, or polycyclic systems, minimize silvicultural interventions by relying on natural regeneration after the harvest of a relatively low number of trees per hectare. These systems result in uneven-aged and multispecies stands, which are thought to provide the best opportunities for biodiversity conservation as part of the management plan. However, a variety of assumptions about regeneration patterns must be made, many of which have been questioned by foresters, and conclusive evidence of long-term sustainability is lacking.

Natural regeneration systems with adaptations to regional conditions have been developed in Malaysia (selection management system), Ghana (modified selection system), Suriname (Celos silvicultural system), Trinidad (periodic block system), and Queensland, Australia (Queensland selective logging system). Each of these has made important strides toward solving technical silvicultural problems, but in many cases sociopolitical obstacles have been more severe. Nevertheless, research is showing that where careful reduced-impact logging (RIL) is carried out, the impacts can be much smaller than in conventional logging (e.g., Smith *et al.*, 2005; Meijaard *et al.*, 2006; Lagan *et al.*, 2007; Putz *et al.*, 2008), and guidelines for tropical and temperate zone forests are being published (e.g., Thorpe and Thomas, 2007).

Salvage Logging

Salvage logging is the practice of extracting boles from burned-over or otherwise disturbed forest areas in order to minimize the loss of commercial timber. Boles tend to degrade rapidly after death due to wood-boring insects, pathogens, and fungi, so post-fire salvage logging is often performed rapidly in the months following a fire. The ecological consequences of such rapid operations have been studied only recently, in particular since a controversial study was published in *Science* in 2005. It

is now clear that salvage logging can inhibit natural regeneration, help spread invasive species, distribute sediments into streams, alter plant species composition, and damage a suite of fire-associated animal species (Kurulok and Macdonald, 2004; Lindenmayer and Noss, 2006; Herrando *et al.*, 2009). These effects may be even stronger than in conventional logging because after major disturbance, forest stands and soils are particularly vulnerable to cumulative effects. The removal of 'biological legacies' including above-ground and underground decomposing biomass, CWD, ash, standing snags, and other natural structures that create special microhabitats, can have very long-lasting impacts (hundreds of years).

Nontimber Forest Products (NTFPs)

Timber is not the only product to be taken out of forests and marketed, nor is it the only product for which forests are managed. NTFPs include fruits, nuts, fungi, fibers, medicinal and ornamental plants, mosses, dyes, resins, gums, fuel-wood, charcoal, leaves as fodder, poles for local construction, honey, syrup, fish, and game, as well as other animal products. In some forests, these may constitute important and large-scale commercial resources; in others, they may have great local importance but fail to appear in commercial markets. Management systems for NTFPs, too, run the gamut from traditional, socially sanctioned systems to those that are legally organized and monitored. In some cases, extraction of NTFPs may affect forest biodiversity even more than timber extraction, since NTFP extraction is often done over a long period of time – over many human generations in parts of southern Asia, Southeast Asia, and Africa – and may constitute an intensive use of some species or areas. Most of the economically significant NTFP extraction systems are today found in tropical forests and include extraction of Brazil nuts, rattans, ornamental plants, and animals, fruits, and medicinal plants. Maple syrup is an economically significant and commercially developed NTFP of temperate forests.

Shifting Agriculture and Traditional Long-Term Intensive Forest Management Practices

As primarily agricultural systems, the effects of shifting agriculture on biodiversity are treated extensively in another article. Shifting agriculture is by no means the only traditional forest management practice, however. Some traditional practices have been very intensive, especially in high population density areas of Asia, Europe, and Central America, and some have been sustained over the very long term. In Japan and Europe, coppicing and pollarding systems, in which trees are repeatedly pruned back for their leaves or poles, have been practiced for centuries. These traditional systems have preserved habitat for a diversity of plants, birds, arthropods, fungi, reptiles, and amphibians that are absent from younger successional forests in the same areas. Traditions of pasturing domestic animals in commons woods, and the retaining of hedgerows, copses, and windbreaks within intensively managed agricultural landscapes, have all contributed to the survival of woodland diversity in Europe. In the tropics, traditional systems of *taungya* and other intensive agroforestry systems have for centuries combined high forest productivity with high species diversity. *Taungya* is a Burmese word, but similar traditions of agroforestry and the combination of tree crops with ground crops and kitchen gardens may be found in Brazil, Indonesia, Malaysia, and the Indian subcontinent. There are currently attempts to reintroduce and adapt some of these traditional techniques to modern conditions, since they are thought to hold promise both ecologically and economically.

Sacred groves found throughout the world that have been "managed" for centuries for spiritual, religious, and utilitarian reasons are beyond the scope of the present discussion.

Managed Forests as Distinct from Secondary Forests

The phrase secondary forest has been used in a variety of ways in the scientific literature of recent decades. The authors reserve the phrase secondary forest for successional forests growing in areas where the forest cover has at some time in the past been completely removed, causing a break in the continuity of the vegetative cover over time, and where this break in continuity can be detected structurally or floristically. Forest historians, geographers, and archeologists have been pushing back the date for earliest detectable human impacts on forests throughout the world, and it is becoming clear that few if any forests are primeval in the literal sense of the word. Nevertheless, the authors call forests that show no obvious structural or floristic traces of human influence undisturbed primary forests.

From the perspective of forest biodiversity, the distinction between secondary and managed primary forests is important. Secondary forests consist of earlier successional stages; they are dependent on seed dispersal from outside for their regeneration and the continued process of succession. The implications of this for biodiversity will be discussed later.

Structural Alterations

Forest structure is the three-dimensional arrangement of trees and other plants, in combination with nonliving spatial elements such as soils, slopes, and hydrology. In short, structure is the physical geography of the forest, considered at a range of spatial scales. Structure includes such characteristics as canopy and understory geometry, continuity or fragmentation of canopy cover, homogeneity or patchiness of species distribution through the landscape, soil structure, and the species composition and age structure of stands. Most of these elements may be considered at micro, local, or landscape scales. Forest structure influences forest biodiversity directly through the formation of microhabitats as well as the determination of larger-scale habitat characteristics, but the relationship between structure and species diversity is complex and not well understood.

The most immediate effects of selective timber extraction systems on forest structure, of course, are the removal of the individual target trees, together with the associated incidental damage. Incidental damage generally includes removal of

surrounding vegetation to construct access roads and skid trails, damage to vegetation during felling and skidding, compaction, and scraping of soils. Harvest of as little as 3% of the trees in an area, as is not atypical in the diverse neotropical forests, may reduce canopy cover by half through incidental damage. Canopy reduction by 75% is not uncommon in Asian dipterocarp forests, where a total of nearly half the basal area of the forest may be removed. However, damage to basal area of as little as 4 or 5% has been documented in Amazonia, when less than a single mahogany tree per hectare was removed. And noncommercial harvesting projects have obtained even better harvest-to-damage ratios in Costa Rica, Suriname, Trinidad & Tobago, Ghana, Madagascar, Queensland, Australia, and other sites.

In conventional operations, the impact on undergrowth can also be substantial. In a logged area in neotropical French Guyana, nearly half of the undergrowth was destroyed during the removal of an average of only three trees per hectare. Removal and damage of vegetation opens the canopy, creating gaps and artificial edges, and often lowers the average height of the canopy, thus altering the internal vertical habitat structure of the remaining forest and, in some cases, removing most or all individuals of the larger and older size classes.

Little precise information exists on the sorts of structural damage ensuing from NTFP extraction. It is clear that damage from intensive harvesting of leaves for fodder, poles for construction, fruits, and firewood can all be substantial.

In the temperate zone forests, the structural changes in managed forests often result from changes in species composition and age structure of overstorey trees, as the multispecies stands are replaced by even-aged populations of commercially valuable native or exotic species. Such changes set in motion a number of changes in ecosystem structure (biodiversity) and function.

Vertical Structure

The vertical stratification of forest vegetation has been the subject of some debate among plant ecologists. It is clear that many forest organisms (birds, insects, and other arthropods, certain herps) partition habitat in the vertical dimension. It is less clear what effect timber management may have on this dimension of forest structure. Selective timber extraction may simplify structure, for instance, by eliminating or preferentially damaging one or several strata. However, it has been suggested that one reason for the increase in local abundance of certain species groups in secondary forests is the greater complexity of the understory structure in younger, less well developed forests. The enhanced presence of short-lived treelets, shrubs, and herbaceous plants in disturbed forests is a matter of record (Richards, 1996).

Canopy

The forest canopy, the highest layer of the vertical structure, includes most of the interface between leaf and light. As such, it is the area of greatest energy input into the forest ecosystem. In tropical or temperate evergreen rain forest, this interface may be virtually unbroken over large areas, and there is often a very distinct division, in terms of biodiversity, between the canopy layer and the understory. In seasonal, dry, or gallery-type forests, the canopy may be much more permeable to incoming light and correspondingly less distinct in its flora and fauna.

Gaps and Edges

The creation of forest gaps through disturbances such as selective logging both enhances and diminishes species diversity and richness in different ways. Gaps increase the local heterogeneity of habitats and create microhabitats that are rare in mature forest, but they simultaneously limit the extent of undisturbed forest, thereby reducing regional habitat heterogeneity. The question of which tendency will predominate is at the heart of discussions about management for biodiversity. The answer depends partly on the characteristics of the focal species or group, partly on the extractive methods employed, and partly on the spatial scale of analysis. Alpha diversity, which is a measurement of species diversity at the community level, does not necessarily co-vary with beta diversity, namely the overall turnover of species associations in contiguous habitat patches across a landscape. In other words, local diversity may not mirror or represent regional diversity. There have been a number of suggestions that the creation of structural gaps through management systems can or should mimic the dynamics of natural gap creation through treefall, with the idea that community dynamics would thus be minimally affected. Other authors, however, have been less sanguine about the practicality of this (for reviews see, e.g., Bengtsson *et al.*, 2000 for European forests).

Gaps and forest edges may often constitute habitats not unlike the forest canopy in respect of abiotic factors such as insolation, wind, temperature extremes, and rainfall, and may therefore exhibit an analogous flora and fauna.

Forest Fragmentation

The term fragmentation is often used to summarize the landscape-level structural changes to the forest exerted by a range of human activities. Forest fragmentation is a process that may take place over centuries, decades, or years. Often, this process starts with selective logging within a matrix of natural forest and may extend over time through the development of progressively more intensive agricultural landscapes (*agroscapes*). These agroscapes may eventually themselves become the matrix, finally leaving only isolated "islands" of residual forest. In developed countries, urbanization, suburbanization, "exurbanization," and the establishment of extensive transport infrastructure have all contributed to forest fragmentation, especially over the past 50 years. The process of forest fragmentation, and its effects on various aspects of biodiversity, has proven to be a rich mine for ecological study. Several studies show that fragmented forests have less biodiversity than contiguous forests and that ecological and evolutionary processes that maintain biodiversity are compromised in forest fragments.

Spatial Mosaic of Forest Types

Both short- and long-term forest management actions have been shown to make radical changes in the spatial distribution

of forest types and communities at the landscape scale. Natural forests are mosaics of species associations responsive in their distribution to aspect and degree of slope, altitude, exposure to winds and storms, fire history, and soil characteristics. A forest's spatial pattern and landscape context are thought to influence the regeneration of logged areas through seed and pollen dispersal dynamics and recruitment rates. Management can alter the natural forest mosaic by superimposing on it separate patterns reflecting land-use history, regeneration dynamics and history, or simply the pattern of access for logging (Cannon *et al.*, 1998; Foster, 1992). Altered distribution of forest types, even within a continuous forest cover, can have potentially harmful (or beneficial) effects on populations of some wildlife species.

Soil Structure and Coarse Woody Debris

Effects of management systems on soils include alteration of the microstructure of large areas of forest soil through compaction and scraping, alteration of the forest floor profile through the creation of artificial pits and mounds (or the elimination of natural ones) by bulldozers, changes to the hydrology of the affected region, and losses of organic matter and nutrients that potentially threaten fertility of the site. The latter problem is perhaps most serious in dryland areas and some areas of the humid tropics, where soil structure is often less robust, lacking the highly developed humus layer and deep mineral soil profile of many forests in temperate climates. In some tropical and temperate rain forests, heavy rainfall tends to exacerbate soil problems through increased runoff from the exposed surfaces and compacted soils of roads and trails, the associated erosion of soil surfaces, and increased clogging and silting of lakes, streams, and wetlands. Old tractor tracks and landings may remain sources of direct runoff for decades, due to their low infiltration capacities. Amounts of sediment washing into streams have regularly been shown to be higher and more variable after logging disturbances. Increases in sediment off-flow by factors of between two and 20 have been measured. These sediments often contain substantial amounts of organic material. Leaching of nutrients, as large volumes of rainwater fall directly onto exposed earth and percolate through the soil, is another problem, the severity of which depends both on the characteristics of the soil substrate and on the logging techniques employed. These soil changes may over time alter species composition and structure of forest vegetation at all levels.

The absence of downed logs and decomposing coarse woody debris (CWD) in temperate zone production forests has been blamed for reductions in soil fertility. A number of plant, animal, and fungi groups depend on complex forests floor structure for habitat and nutrient availability. Several studies have also confirmed that a variety of forest stream fauna, both temperate and tropical, need CWD lying in and across the stream channel. Clear-cuts, elimination of riparian buffer strips and channel cleaning are all highly destructive practices associated with some management approaches.

All of these problems with soils are amenable to improvement through careful logging procedures. The US state of New Hampshire, the Canadian province of British Columbia, the Forest Guild of New Mexico and the University of Maine are just some of the groups that have published guidelines for the maintenance of CWD in forestry operations (for further references, see e.g., Evans and Kelty, 2010).

Abiotic Changes

Insolation, Temperature, and Wind Patterns

Direct structural changes may cause changes in a series of abiotic factors, again on several scales. Opening of the canopy alters the insolation regime in dense-canopied forests, though it may have less effect in forests with naturally sparse canopies. Sunlight reaching the forest floor has been measured at as little as 2% in closed-canopy tropical forests; selective logging raises this percentage manifold. Greater insolation has two kinds of effects: it increases the daily energy flux, especially to the lower strata of vegetation, and it raises the average temperature of both air and soil in the forest undestory. Partial or complete removal of the forest canopy also increases the movement of the air column throughout the understory in response to winds above the canopy, affecting water retention, relative humidity, and transpiration rates in the understory. These changes affect the microclimates within the forest, making it warmer and drier on average. This reduces the suitable habitat for some species of plants and animals, while increasing it for others.

Hydrology

Changes in forest floor hydrology may contribute to an increase in soil disturbances such as landslides and erosion, especially in the wet tropics, where rainfall is extreme. Such disturbances may increase the area of early successional vegetation; in areas where the organic component of the topsoil layer has been scoured away, the process of succession may be halted in its early stages for a long time. Even lesser levels of disturbance to the soil due to logging have been shown to reduce concentrations of CO_2, organic matter, and nutrients through leaching. Such changes will alter competitive balances among plant populations in favor of species adapted to disturbed habitats, and to the detriment of those specialized for the relatively consistent conditions of closed-canopy forest. Some forest ecosystems are especially sensitive to changes in hydrology. Among these are mangrove forests, which depend on the regular flushing action of the tides and may be strangled by influxes of sediment from eroding soils. Lowland swamp forests, including Brazilian *varzea*, likewise may be gradually replaced by *terra firma* forests under conditions of sedimentation. Riparian forests of various kinds are important refuges for specialized plant communities, especially in arid regions. Scrub forest and other drylands forests are vulnerable to regional changes in rainfall patterns, since their plant communities may be at the edge of their ranges or environmental tolerances. Some of the most interesting and endangered forest types, in short, may be compromised by the hydrological changes that often accompany logging and other forest use.

Biotic Changes

Up till now the authors have been describing those measurable structural, abiotic effects of logging that have been documented in forests of various types worldwide. What effects do these changes have on the biodiversity, especially the species, which make up the forest?

Plants form the first, or *autotrophic*, layer of the forest ecosystem. The autotrophs are able to build organic compounds from inorganic building blocks, using energy from the sun. The other (*heterotrophic*) organisms in a forest are directly or indirectly dependent on the transformative power of photosynthetic plants for their livelihood. Therefore the authors will consider plants first. Fungi have been implicated in crucial symbiotic relationships with many forest plants and will be considered next. The most numerous primary consumers of plants in forests are the invertebrates. Primary consumers transform and recycle the organic compounds they derive from ingesting plants. This group will be discussed third. Some vertebrates, the group the authors review fourth, are also primary consumers (*herbivores*), others are secondary or tertiary consumers (*carnivores* and *omnivores*). *Detritivores*, which make up the final stage in the recycling of organic materials, are a very important component in forest ecosystems; the authors will discuss them in the context of effects on soils.

Plants

Since plants are stationary, they are most immediately influenced by changes in their abiotic surroundings. Such changes may influence plant species composition, density, evenness, community structure and associations, and plant interactions including mutualisms and parasitisms.

Early Successional Phases

A *pioneer guild* of plants is characterized by lack of shade tolerance and therefore tends to be favored in the light-rich microhabitats created by gaps. Pioneer species tend to be widespread, with airborne pollen and small airborne seeds, fast growth rates, and tolerance of a wide variety of environmental conditions, allowing them to disperse quickly into disturbed habitats. Hence, they are characteristic of managed forests, and the pioneer guild is relatively homogeneous in species composition over large regions. This can be confusing to researchers, since localized increases of plant species diversity are common following logging. Logging expands the area of forest occupied by early successional phases, boosting representation of pioneer species in small samples. These local increases in diversity are not, however, matched on the regional level, since the local increases themselves come from the regional pool. Thus, logged forests tend to become increasingly homogeneous on a larger scale.

Exotic Invasive Species

Loss of tree cover often permits the recruitment of "naturalized exotics" (species that are not natural to the area), since the special conditions under which native species are competitively superior no longer exist. The presence of exotic plants tends to homogenize habitats and hence biodiversity. Examples of the many widespread and aggressive exotic plant problems are *Impatiens* (an Asian ornamental) in the neotropics, and the reed *Phragmites australis* in the United States.

Fungi and Mycorrhizae

A number of studies have shown that the presence in the soil of inoculate of certain species of symbiotic fungi is necessary for the regeneration and successful growth of a number of tree species, and the same may be true for other rooted plants. Soil mycorrhizae invade growing root tips and make available to them, in the course of their own metabolic processes, nitrogen and other nutrients that are otherwise locked up in the soil. Mycorrhizal fungi require certain soil conditions to flourish. When soil environmental conditions change radically – as a result of clear-cutting, for instance – these fungi may go locally extinct. Without healthy populations of the appropriate fungal symbionts, many tree species may be unable to regenerate successfully.

Invertebrates

We still know the least about the effects of logging on the first tier of forest-inhabiting animals, the invertebrates. One of the central unknowns is the number of invertebrate species inhabiting tropical forests, which is not known even to the nearest order of magnitude. Nevertheless, data on some taxa are available, and it is possible to distinguish some trends. Canopy denizens, especially flying insects such as butterflies and moths as well as flies, wasps, and bees, may be more easily seen or trapped in lower, disturbed forest, and hence may show up more easily in surveys. Like nectar-feeding and frugivorous bird species, some invertebrates may also benefit from the increased flowering and fruiting among early successional trees and vines in logged tropical forests. And many groups are adapted to the microclimates of forest edges and gaps, and so may benefit on the population level from structural changes and forest fragmentation. Nevertheless, lepidopterans and some dipterans and bee species are clearly reduced on a regional scale in disturbed tropical forests, because there are groups with very specific adaptations to the environmental conditions of closed-canopy forest. For example, the hives of the large Asian honeybee (*Apis dorsata*) are usually found on very large old trees. Certain groups of tropical flying insects depend on vertebrate dung for part or all of their nutritional requirements. These may be strongly affected by reductions in mammal and bird populations that may accompany disturbance, although some may equally benefit from the presence of horses, cattle, or other domestic animals in the landscape. Other tropical winged insects are well adapted to the conditions of closed forest and have coevolved with particular resources that may be absent or rare in disturbed forest. Examples are euglossine bees and orchid flowers, parasitic wasps and fig fruits, or lepidopteran larvae. These groups may become rare when their hosts become rare.

In the temperate zone, lepidopterans, as well as other flying invertebrates, such as dragonflies, mayflies, and damselflies, are rarely found under closed forest canopy due to both their

temperature requirements and their dependence on bodies of water during part of their life cycle. Nevertheless, these too may be affected by uncontrolled logging activities that disturb their habitats within forests, such as bogs, swampy areas, ponds, and streams. Since boreal forests are relatively species poor, approaching the conditions of a monoculture, it is unclear to what extent cyclic events – such as insect infestations or fire – may be a part of their natural disturbance regime.

The invertebrate fauna of leaf litter and soil surface shows divergent responses to logging activities. Some groups (beetles, cockroaches, and millipedes) increase significantly in both species richness and abundance in logged tropical forests. Others (spiders, mites, scorpions, springtails, and termites) decrease. Still others, such as ants, seem to be reduced in species richness while remaining abundant in absolute numbers. Many detritivores (wood-boring beetles, certain termites, and soil mites as well as fungi) are habitat-dependent on standing or downed rotting wood. Such groups appear to have been negatively affected by long-term timber management in Europe and the American Southeast, where old-growth forest structure including snags and large rotting logs has been effectively eliminated over large areas.

The effects of forest management on another important group of invertebrates, the parasites, are little understood. It is known that parasites are some of the strongest regulators of agricultural pests, and there are data showing that tropical agriculture, chronically vulnerable to pests, may actually be more successful where patches of natural forest are left between fields. Parasites and predators on insect pests may breed in forests and disperse into nearby fields, holding down pest populations.

It has often been observed that a dramatic increase in mosquito populations is a common phenomenon in logged areas, where drainage patterns are often affected. In the tropics this may constitute an important human health concern, and malaria is increasing in some areas.

In general, the invertebrates are still the least well-known of the macroscopic fauna, yet their aggregate biomass is very large and their ecological influence is difficult to overestimate. E.O. Wilson has written about the essential role played by the "little things" in maintaining a variety of ecological functions (Wilson, 1987). The reduction or absence of groups of these forest denizens may have as yet unforeseen long-term consequences for forest ecology and ecosystem functioning.

Vertebrates

John Terborgh observed that, in addition to the critical role played by Wilson's "little things" as the foundation of an ecological pyramid, there is an important top–down regulatory role played by the "big things," namely vertebrates and especially the large predators (Terborgh, 1988). In sites where top predators are absent, their prey populations, most of which are herbivores, tend to proliferate. This dynamic has been documented in tropical sites such as at the Smithsonian Institution Tropical Research Institute station on Barro Colorado Island, Panama, which lacks jaguar, in temperate parks devoid of wolf and puma populations, and in expanding suburbia all across North America, where white-tail deer and Canada geese are increasingly being viewed as pests by the human population. Many of the proliferating tropical herbivores, such as the large rodents agouti and capybara, are seed predators and may be affecting the regeneration of large-seeded tree species. In temperate parks, large deer and elk populations are browsing deciduous seedlings to the ground, effectively preventing their regeneration. The lack of top predators in managed forests may significantly alter forest structure and composition, both in the short and the long term.

Vertebrate populations in forests coming under management may be affected by losses or gains in (1) habitat area or (2) food resources. The direction of change is determined both by the particular needs of each species and by the nature of the management system. The many permutations of these variables explain why we see divergent responses to management among vertebrates.

Habitat Loss or Increase

The loss of important habitat elements tends to affect the more specialized and the less mobile species most strongly, since they may not be able to disperse to new habitats. Loss of the appropriate microclimate, sufficient kinds and amounts of cover, nesting sites, or even perch sites may force species to local extinction. We know that some vertebrates depend on habitat elements found only in old, undisturbed forests. A well-known example is the northern spotted owl in the Pacific Northwest of the United States, which nests in the large hollow trees that are usually eliminated by timber management. A range of large temperate and tropical birds, including large woodpeckers, hornbills, and quetzals, likewise nest exclusively in hollows and may use and reuse the same holes year after year. Some mammals, too, are critically dependent on old-growth habitat elements for survival. The American marten, for instance, which has very little body fat yet remains active throughout the winter in cold and snowy environments, needs resting places below the surface of the snow, often under rotting stumps or large logs. Thus, although animals like the omnivorous marten may benefit from the increased seasonal availability of berries or other resources in logged forest gaps, their territories must also include areas of older forests with woody debris.

It has been shown that a number of smaller forest birds, both temperate and tropical residents, avoid gaps in forest cover such as those created by logging. Some are reluctant to cross gaps of as little as a few tens of meters. This may result from fear of predation or from physiological adaptation to a narrow range of understory temperatures and humidity levels. Under the increasingly predominant conditions of forest fragmentation, these species may find it difficult to locate appropriate habitat, even when it is available nearby.

Terrestrial amphibians, including frogs and salamanders, may often find the relatively desiccating conditions of open-canopy disturbed forest too warm and dry. Because they "breathe" partly by the exchange of gases directly across the skin, amphibians are particularly vulnerable to alterations in their microclimates. One element in the present worldwide crisis among amphibian populations may be reductions in suitable cool, moist, and shaded old-growth habitat.

However, some vertebrate populations may do well in disturbed habitats in secondary succession. Examples of this

group that have been studied are mice in temperate forest fragments, tenrecs in Madagascar, and other rodents in tropical and temperate forests, all of which may benefit from the increase in undergrowth and herbaceous ground cover accompanying management.

Loss or Increase of Food Resources

Opening the forest canopy may often stimulate a flush of new leaves or fruits as the trees and understory respond to the more generous light allowance. Many browsing and fruit-eating vertebrates benefit from this change and may seek out areas of secondary succession such as logged areas. Populations of large folivorous mammals, including elephant, tapir, duiker, gaur, bearded pig and peccary in the tropics, or deer and elk in the temperate zone, often do well in secondary forest. In the absence of predation, including human hunting, they may themselves become agents of serious ecological change by grazing and trampling large areas of forest understory. However, many of these large browsing mammals are important game animals, especially in tropical forests where hunting is particularly difficult to control. Many have evolved shyness responses to humans, and most have greatly reduced populations in areas where hunters have access.

Birds

The effects of logging on the species richness of birds are highly variable. In logged tropical forest, certain feeding guilds experience a reduction in population ranging from 25% up to 100%. Hard-hit guilds include terrestrial foragers (insectivores and frugivores) and the small insectivorous species associated with understory habitats, particularly those of the forest interior. These decreases are likely the result of reductions in the insect resource base or microhabitat changes. These findings are not completely consistent among studies, however. Some studies, for instance, have found small increases in abundance for terrestrial, understory, and foliage-gleaning species after logging in Malayasia. Substantial local increases in abundance have also been noted for generalist feeders that can supplement nectar and fruit resources with insects. The opening of the canopy often, though not always, increases the availability of nectar and fruits in early successional patches, in turn supporting more abundant populations of generalist feeders.

Primates

Primates have particularly complex responses to logging and to the presence of humans in managed forests. In the Neotropics, generalist feeders (*Pithecia* spp., *Saguinus* spp.) have been found to do well in disturbed areas, whereas frugivores (e.g., *Chiropotes* spp., *Ateles* spp.) tend to leave logged-over areas, especially when their food trees have been extracted. However, there are exceptions, such as *Callicebus torquatus*, a generalist feeder that was found to be absent from logged forest. Old World frugivores such as *Pongo pygmaeus* and *Pan troglodytes* decline sharply in population density after logging, whereas folivores (e.g., *Colobus* spp., *Gorilla gorilla*) may seek out disturbed forest for the leaf flush.

Behavioral Changes

Behavioral changes among vertebrate populations have received little attention until recently. There are substantial difficulties inherent in studying and quantifying them. Nevertheless, they should be considered a significant group of changes because their effects may be delayed until one or several generations after the disturbance itself. Examples, which have been noted, are changes in social organization and breeding behavior among primates in Malaysia, and in nesting behavior and dispersal among birds in fragmented habitats. Among animals with complex response patterns, there are doubtless many other instances yet to be studied of behavioral changes stemming from human interventions. These may affect population dynamics years or decades after an event such as logging.

Domestic Animals and Nonforest Species

Domestic animals and introduced species have been important influences on forests throughout history. In European oak and beech forests, for instance, pigs and other grazing animals were grazed in commons forests on a seasonal basis to take advantage of fruit masting. Sheep and goats have virtually eliminated forest regeneration on the once lushly tree-clad slopes of the Mediterranean and throughout the Middle East. Today, browsing of cattle is a major influence on remaining forest fragments in many areas of the tropics.

The movement of nonforest species into logged and disturbed forests is the origin of some of the apparent increases in species diversity in logged forest plots. More species of coleoptera, for instance, have been counted in logged tropical forests than in neighboring primary forest. The species lists in logged areas, however, included the names of field species that had entered the forest after the canopy had been degraded.

Aquatic Fauna

The loss of aquatic biodiversity is understudied in the tropics. Concern regarding changes in hydrology resulting from forest disturbance overlap with local community concerns, since the need to protect sources of clean water is becoming acute in many places as human populations grow. There is more information about the effects of logging on streams and aquatic life in the temperate zone. The removal or degradation of tree cover on slopes, including the construction of logging roads, often results in topsoil loss and erosion. Soil is washed downhill and ends up in streams and rivers. Silt deposition causes declines among anadromous fishes, bottom-feeding river fishes, bivalves, and other invertebrates. Heavy silt loads can have major impacts far downstream – even in some cases degrading coastal habitats such as tropical coral reefs, mangroves, and offshore fisheries. Changes in the temperature of stream water as a result of streamside vegetation loss can also have far-reaching effects on faunal diversity.

Ecological Interactions

Structural, abiotic, and biotic changes are likely to alter a range of ecological interactions and evolutionary processes. Changes

in hydrology, soils, and soil fauna, for example, should influence nutrient dynamics, which may in turn have further impacts on biodiversity. Biotic changes alone can alter species interactions. Again little is known about the impact of extractive regimes on ecosystem function and ecological interactions. The impacts are likely to be most dramatic where extracted species constitute dominant elements of forest ecosystems as, for example, the dipterocarp tree species in Southeast Asian dipterocarp forests. Indeed, much of the recent evidence that logging may influence vital interactions and ecological processes comes from such forests.

Pollination, seed dispersal, and seed predation are basic processes that profoundly influence reproductive output and regeneration of plants. By lowering the density of harvested species, logging can decrease the resources available to pollinators, animal seed dispersers, and seed predators. The vast majority of tree species are outcrossed, and in tropical forests, where most tree species are pollinated by animals, substantial declines in the abundance of food plants can result in changes in pollinators' foraging behavior. One study in Southeast Asian *dipterocarp* forests confirmed that logging in forests containing *Shorea siamensis* increased the average distances among the remaining trees of that species. The large distances made it difficult for the bees that pollinate *S. siamensis* to move from tree to tree, concomitantly limiting outcrossing potential. Opening up the canopy also led to colonization of the forest floor by flowering plants that attracted bees to the understory, further depriving the canopy trees of their pollinators (Ghazoul, 2001).

Reduction in the density of reproductive individuals can also result in a substantial decline in the number of seedlings. Studies have compared seedling production in logged and unlogged dipterocarp forests of Indonesia (Curran *et al.*, 1999). Logging reduced the number of reproductive individuals per hectare to 3% of the original. Seedling production in the logged forest, following a mast fruiting year, was a mere 15% of that in the nearby unlogged forests.

The same study also documented indirect effects of logging. Dipterocarps undergo mast fruiting every few years associated with local El Nino-Southern Oscillations (ENSO). Logging around the Gunung Palung National Park, the site of the study, has fragmented the once contiguous forest. Changes in land-use have apparently affected local ENSO climate conditions. As a result, spatial synchrony in mast fruiting and the level of seed production have been reduced. One result of the natural mast fruiting pattern is that copious amounts of seed are produced simultaneously. It is assumed that this "satiates" seed predators, leaving a substantial surplus of intact seeds for regeneration. However, with the reduction in seed production, seed predators are not getting "satiated," and there is little seed surplus. Even in the national park, where there is no logging, there has been inadequate regeneration because of the decrease in seed production and the movement of seed predators into the park from surrounding areas lacking mature reproductive trees.

A more recent comparative study of dipterocarp regeneration in Borneo (Bischoff *et al.*, 2005) found that 13 years after logging, succession was not proceeding as expected – valuable light hardwood dipterocarps were being suppressed. Bischoff *et al.* suggested that this may have been the result of logging forest stands in a late but not mature stage of succession. A mast fruiting event had occurred shortly before the logging, and had produced a large cohort of largely noncommercial dipterocarp seedlings, which then benefited from the canopy opening and prevented commercial dipterocarp regeneration. Bischoff *et al.* traced this situation back to the previous major (natural) disturbance, from which the forest was still in recovery. This event was presumed to have taken place over a hundred years before, about 1878 – a salutary reminder that forest history is a central, but often overlooked, factor that may contribute strongly to the formation of present forest conditions.

These studies from Southeast Asian dipterocarp forests illustrate several of the pathways by which logging can have far-reaching and perhaps unexpected effects on forest dynamics:

a. Removal of mature dipterocarps and the opening of the canopy (both structural changes) produce new microclimatic conditions and an altered array of understory species (a biotic change);
b. these changes combine to produce altered foraging patterns among bees and other pollinators, compromising their efficiency in an essential ecological interaction;
c. meanwhile, fragmentation and degradation of the forest cover alters local climatic patterns (an abiotic change), which again affect an ecological process, the mast fruiting pattern;
d. this alters seed predation dynamics and affects regeneration, potentially influencing forest structure and composition into the long term.

Considerably more work is needed to fully understand the consequences of the altered ecological relationships brought about by logging and other extractive activities. Studies on the effects of logging have largely been restricted to the impacts on structural and biotic components. Removal of nontimber forest products may also be expected to have an impact on a diverse range of ecological interactions and processes but has received even less systematic study.

Genetic Effects and Evolutionary Processes

Alterations in density of tree species may also have a profound effect on evolutionary processes. Most trees are strongly outcrossed; outcrossing rates are highly dependent on population density. Low population density can restrict the movement of pollen among trees and result in a high level of inbreeding. Although it has been difficult to document the deleterious consequences of a decrease in outcrossing rate, it has been confirmed in several species that inbreeding does increase with a decrease in density.

In the tropics where many tree species typically occur in low densities, reduction in density due to extraction may further decrease population sizes. Small population sizes may also contribute to inbreeding. Thus, inbreeding may be increased both by the effects of decreases in density on the mating system and by decrease in overall population size.

Extraction can also remove the better-adapted genotypes from the population, leading to *dysgenic* selection, or selection for less well-adapted genotypes. Populations subject to harvest

have lower levels of heterozygosity and lower overall levels of genetic variation. Inbreeding in logged populations resulting from a decrease in density may further decrease overall levels of heterozygosity.

Genetic and evolutionary effects can be cumulative. Fragmentation, decrease in density, and reduction in effective population size all contribute to inbreeding. In trees that are highly outcrossed, increasing levels of inbreeding may be particularly deleterious in the face of a changing environment. The genetic consequences of management interventions in forest ecosystems remain poorly explored.

Synergistic Anthropogenic Effects

Still today, only a minority of extractive forest operations is performed under a management plan, and most of these are in the developed, less biodiversity-rich countries. In the tropics, by contrast, only about 96.3 million hectares of forest was under any kind of active management plan in 2005. According to the International Tropical Timber Organization's (ITTO) comprehensive review of the status of tropical forest management (ITTO, 2006), that represents 27% of the total natural 'production forests' of the tropics (352.5 million hectares). The ITTO survey, based primarily on data submitted by the then 33 producer-member countries of the ITTO, also asserted that 25.2 million hectares of tropical 'production forest' (7% of total tropical production forest), plus 11.2 million hectares of tropical 'protection forest' (2.4% of this designation), are being managed 'sustainably'. Together with 14.3 million hectares of planted forests under certification for sustainable management, this would sum to a grand total of 858.84 million hectares of tropical 'Permanent Forest Estate', of which 36.4 million hectares were being managed sustainably in 2005. If that is true, it would represent a 25-fold increase over the less than 1 million hectares of sustainably managed tropical forests estimated by the Poore *et al.* study of 1989. Nevertheless, it would still mean that only about 4.2% of all tropical forests were being sustainably managed in 2005.

The absence of planning and long-term control most often results in a series of secondary effects, including increased hunting, fire, invasion by exotic species, mining activities, shifting agriculture, and illegal logging, which may far exceed the effects on biodiversity of the original management operations.

It has been noted that large amounts of secondary vegetation and brush tend to alter the understory structure of cutover areas. The role of such vegetation and brush in fueling forest fires, both in temperate forests and even in the relatively fire-resistant moist tropical forests, has received increasing attention in recent years (e.g., Nepstad *et al.*, 1999; Blate, 2005; Laurance and Useche, 2009; Lindenmayer *et al.*, 2009). It is clear that large-scale fires are altering microclimatic conditions over wide areas of tropical Asia and South America. In some of the drier temperate forests, for instance parts of the US West and the pine barrens of the Atlantic seaboard, vegetation communities have evolved under the influence of regular low-intensity fires. The suppression of such fires during the past century has altered the makeup of these communities and in some cases lowered their resistance to large destructive conflagrations. Reintroduction of controlled fire regimes is beginning to allow the native species to return and compete successfully with exotic species. A somewhat contrasting case is found in certain areas of the dry tropics, which are, as a group, some of the most highly altered and degraded forest ecosystem types. Here, degradation of the native plant assemblages and hydrology has virtually eliminated the forest's ability to regenerate itself in the face of continual incursions of fire. In such areas, putting a stop to forest fires over at least the next several decades is seen as a key step to restoring the natural plant communities.

Sharp increases in hunting after logging, too, remain the object of attention (Struhsaker, 1997; Oates, 1999; Robinson *et al.*, 1999; Laurance *et al.*, 2006; Poulsen *et al.*, 2009). Logging requires roads. Particularly in the tropics, where hunting is often unregulated, logging roads provide easy access for hunters and their vehicles into previously remote areas. Logging crews themselves are often the first hunters in a newly accessed area, both for their own consumption and for commercial markets. Others quickly follow them, and pressure on the game animal species can become intense.

In many parts of the tropics, one of the most intractable effects of logging is settlement by colonists eager for new agricultural land. Colonists follow the logging crews, entering the forest along logging roads, carving out homesteads, and planting crops. Since such movements tend to fragment and finally convert forest to agricultural land, they are clearly some of the most deleterious influences on biodiversity. Further extractive activities associated with wresting a living from the forest, such as mining, fishing, and trapping animals and birds, tend to accompany colonization and increase pressure on forest resources in a multitude of ways.

Demonstrating the multiple indirect effects that logging and its attendant activities such as road-building can have in remote tropical forests show that destructive fires too are closely associated with roads in the Brazilian Amazon. Interestingly, indigenous reserves, strictly protected reserves and restricted use reserves all had about the same effect in reducing the incidence of destructive fires, at the rather coarse scale of the satellite data used in this study.

Anthropogenic impacts compounding and altering the effects of logging are by no means limited to the tropics. The spread of invasive species like the Hemlock Wooly Adelgid in North America has interacted in some surprising ways with timber harvest, including encouraging the use of 'salvage' logging and biomass removal (e.g., Kizlinski *et al.*, 2002). The proliferation of nonnative tree species in intensively harvested southern European forests affects forest elements such as overwintering bird species, especially bark foragers (Laiolo *et al.*, 2003). Asner *et al.* (2010) produced an important attempt to integrate information about the expected impacts of climate change and land-use change, finding that combined impacts could drive significant changes in vegetation over more than 80% of Amazonian humid forests.

Essential Trends and Concluding Remarks

Clearly the extraction of forest products in managed ecosystems has an impact on biodiversity. However, natural

ecosystems are dynamic and subject to all sorts of perturbations. The critical questions in the case of managed ecosystems are the extent to which biodiversity is lost and the degree to which the losses are irreversible. A landscape perspective is important for addressing these questions, because biodiversity may increase locally while decreasing at the regional level. Moreover, some local changes may be manifestations of regional level changes, and vice versa.

In the last 10 years, there have been a number of significant attempts to synthesize what is known about the biodiversity effects of forest management. Among these is Hill and Hamer's (2004) analysis of the role of observational scale in producing results that sometimes appear contradictory – in particular as regards the conditions under which logging can appear to increase species richness. Different taxa show nonlinear effects of sampling scale: the bird studies sampling at relatively *larger* scales and the Lepidopteran studies sampling at relatively *smaller* scales tended to report increased diversity following disturbance.

From this and other synthetic efforts, the following broad principles and predictions emerge:

1. Disturbances (such as logging) affect biodiversity at the scale of the *forest stand* (scale of hectares or less than one hectare) by mixing edge species and nonforest species into the pool of forest species, raising species richness at that scale. This effect is the result of the new canopy openings produced by all logging systems. Since diversity of plants, invertebrates and small mammals is often measured at that scale, these taxa often show increases in species richness in logged stands. Mobile taxa such as birds, however, tend to show reductions in species richness at this scale.
2. At the *whole-forest* scale (scale of tens of hectares, over 25 ha), many birds, less-sedentary butterflies and moths, primates and other large mammals may be censused, as well as trees. Species that prefer or depend on new-leaf flush, flowering and fruiting plants (especially pioneer species) are drawn to disturbed forest areas, sometimes producing inflated species numbers after selective logging events. Species that have specialized requirements for the habitat conditions provided by old-growth forests – including narrow temperature ranges in the understory, prevailing humidity, restricted air movement, large standing snags and CWD – tend to disappear from managed forests, so their numbers go down at the whole-forest scale unless sufficient refuges are available in the form of undisturbed areas. The presence of such refuges produces a mosaic of areas of different disturbance intensities at the whole-forest scale, raising species richness overall.
3. At the *landscape* scale (scale of tens or hundreds of ha), β diversity or species turnover typically goes down after logging, as managed forests are relatively homogeneous. All logging systems increase the size, number, and distribution of canopy gaps across a forested landscape. This inevitably affects the relative abundance of tree species, favouring pioneer and early successional species while rendering shade-tolerant late-successional species – many of which are rare in nature – prone to regional extinction. This effect may be minimized by careful selective logging systems with set-asides of sensitive areas (swamp-land, riparian areas, steep slopes, etc.) to maintain habitat heterogeneity, and limitation of secondary effects like hunting and burning. It is these uncontrolled secondary effects that are responsible for most local and regional extinctions and extirpations.
4. Future studies should account for scale effects explicitly by sampling across a wider range of spatial scales. Crucially, logging alters the relationship between α and β diversity (point diversity vs. species turn-over between sampling points). The direction and extent of the change, however, depends on the taxon being observed. This effect is difficult to quantify using available data, because taxonomic groups and sampling scale co-vary quite strongly.
5. "Conservation value" is a separate measure that must be treated as distinct from diversity, or species richness. Even in those studies that find overall increased species richness after logging, species of restricted ranges and specialized requirements often decrease or disappear, while the more generalist taxa flourish.
6. Attempts to define management procedures or outcome measurements based on indicator species, focal species or threshold levels of vegetation cover have never become widely influential. There are a number of ecological and technical reasons why such management 'short-cuts' are difficult to apply when making real-life decisions (see Lindenmayer *et al.*, 2006 for a summary). In short, ecological complexity and the complexity of environmental change mean that extracting individual phenomena as 'representatives' of broader processes is always an uncertain task.

Here we mention just a few examples of syntheses of the effects of forest management systems on particular taxa or functional groups:

- Gray *et al.* (2005) analyzed data from 57 published studies (covering 1214 species) on avian guilds and trophic organization to show how and why different groups respond differently to forest disturbance.
- Reviewing work in boreal forests, Kreuzweiser *et al.* (2008) found soils responses to logging to be highly variable and site-specific, dependent on endogenous factors including soil types, site conditions, and hydrological connectivity, as well as external factors such as the timing of harvest activities, postlogging weather patterns, atmospheric pollution inputs, and advancing climate change.
- Sodhi *et al.* (2009) analyzed the results of 120 published studies to produce a generalized statistical measure of 'ecological health' in SE Asia. The results, largely confirming intuitive understandings, are encapsulated in concluding statements such as 'SE Asia's biotas are extremely sensitive to human-induced disturbances'; 'large-scale regional deforestation may depress reproductive success'; 'deforestation and disturbance may negatively affect the demographics of SE Asian biotas' and 'the negative effects of deforestation may be partially mitigated if forests are allowed to regenerate'.
- Meijaard *et al.* (2006) observed that timber management guidelines to date mostly focus on trees and structural issues and say relatively little about managing for wildlife populations. They produce a sensible summary of

recommendations, specifically for Borneo but applicable throughout the tropics (and indeed globally), including maintaining understory vegetation and large trees for fruit, seed, dead wood, and tree hollow production, limiting canopy gaps, and reducing hunting and wildlife trade in timber concessions.

- Newton *et al.* (2009) focus on the kinds of information and analysis needed, in principle, for the further development of integrated forest modeling frameworks and participatory research approaches where there is considerable uncertainty about many parameters – as there often is still today, especially in tropical forest contexts. These approaches include Bayesian Belief Networks, spatial multi-criteria analysis, and scenario planning. Only some of these are likely to be amenable to simple, nonmathematical techniques such as could be used by nonacademic or community-level managers, especially in rural parts of developing countries. Management approaches requiring the extensive use of modeling will always limit the ability of communities to participate actively, and the ability of managers to involve and gain the trust of communities. Fortunately, it is possible to extrapolate a few relatively simple predictions about general trends, though these will always need to be adjusted to local conditions.

What is now urgently needed are sustained efforts to synthesize and communicate the knowledge encapsulated in the hundreds of publications available on the detailed effects of human commercial and management activities on forest biodiversity. Two examples of such synthetic programs are:

1. The program in Ecological Synthesis at the Fenner School of the Australian National University, led by Fischer, Lindenmayer and Cunningham (http://fennerschool-research.anu.edu.au/cle/metaanalysis/; see e.g., Lindenmayer & Hunter 2010), and
2. The Centre for Evidence-Based Conservation (CEBC) at Bangor University, Wales, (www.cebc.bangor.ac.uk) and its associated websites and programs: www.conservationevidence.com, run from the Department of Zoology, University of Cambridge, and the Collaboration for Environmental Evidence (www.environmentalevidence.org). CEE has set up an Environmental Evidence Library on line, containing systematic reviews of evidence on the effectiveness of human interventions in environmental management and the environmental impacts of human activities.

A landscape perspective is crucial to unraveling the confounding but synergistic effects of habitat fragmentation and climate change on biodiversity. A wide outlook is critical to understanding the impacts on the ecological and evolutionary processes and the ecosystem services that have only recently begun to be drawn into studies concerned with biodiversity in managed ecosystems. Finally, the authors do not imply that biodiversity in managed forests will always substantially decline. The low-impact harvesting regimes that are being explored all over the world may allow significant and critical goods and services to be conserved and sustainably used, particularly if such management systems are well integrated with their social and cultural milieus and are responsive to the social needs of specific regions.

See also: Agriculture, Sustainable. Deforestation and Land Clearing. Fires, Ecological Effects of. Forest Canopies, Animal Diversity. Forest Ecology. Predators, Ecological Role of. Timber Industry

References

Asner GP, Loarie SR, and Heyder U (2010) Combined effects of climate and land-use change on the future of humid tropical forests. *Conservation Letters* 3: 395–403.

Bawa KS and Seidler R (1998) Natural forest management and conservation of biodiversity in tropical forests. *Conservation Biology* 12: 46–55.

Bischoff W, Newbery DM, Lingenfelder M, *et al.* (2005) Secondary succession and dipterocarp recruitment in Bornean rain forest after logging. *Forest Ecology and Management* 218(1-3): 174–192.

Blate GM (2005) Modest trade-offs between timber management and fire: Susceptibility of a Bolivian semi-deciduous forest. *Ecological Applications* 15(5): 1649–1663.

Bruenig EF (1996) *Conservation and Management of Tropical Rainforests: An Integrated Approach to Sustainability.* Wallingford: CAB International.

Cannon CH, Peart DR, and Leighton M (1998) Tree species diversity in commercially logged Bornean rainforest. *Science* 281: 1366–1368.

Curran LM, Caniago I, Paoli GD, *et al.* (1999) Impact of El Nino and logging on canopy tree recruitment in Borneo. *Science* 286: 2184–2188.

Evans AM and Kelty MJ (2010) *Ecology of Dead Wood in the Northeast.* Santa Fe, NM: The Forest Guild.

Forman RTH (1995) *Land Mosaics: The Ecology of Landscapes and Regions.* Cambridge: Cambridge University Press.

Foster DR (1992) Land-use history (1730–1990) and vegetation dynamics in central New England, USA. *Journal of Ecology* 80: 753–772.

Frumhoff P (1995) Conserving wildlife in tropical forests managed for timber. *Bioscience* 45: 456–464.

Ghazoul J and McLeish M (2001) Reproductive ecology of tropical forest trees in logged and fragmented habitats in Thailand and Costa Rica. *Plant Ecology* 153: 335–345.

Gray MA, Baldauf S, Mayhew PJ, and Hill JK (2005) The response of avian feeding guilds to tropical forest disturbance. *Conservation Biology* 21(1): 133–141.

Grieser Johns A (1997) *Timber Production and Biodiversity Conservation in Tropical Rain Forests.* Cambridge: Cambridge University Press.

Herrando S, Brotons L, Guallar S, Sales S, and Pons P (2009) Post-fire forest management and Mediterranean birds: The importance of the logging remnants. *Biodiversity Conservation* 18: 2153–2164.

Hill JK and Hamer KC (2004) Determining impacts of habitat modification on diversity of tropical forest fauna: The importance of spatial scale. *Journal of Applied Ecology* 41: 744–754.

ITTO (2006/1) Status of tropical forest management 2005: Summary report. *A Special Edition of the ITTO Tropical Forest Update.* Yokohama: ITTO.

Kizlinski ML, Orwig DA, Cobb RC, and Foster DR (2002) Direct and indirect ecosystem consequences of an invasive pest on forests dominated by Eastern Hemlock. *Journal of Biogeography* 29: 1489–1503.

Kreutzweiser DP, Hazlett PW, and Gunn JM (2008) Logging impacts on the biogeochemistry of boreal forest soils and nutrient export to aquatic systems: A review. *Environmental Reviews* 16: 157–179.

Kurulok S and Macdonald E (2003/2004) Impacts of post-fire salvage logging on tree regeneration and plant communities in the mixed-wood boreal forest of Alberta. *SFM Network Project Report.* Edmonton: University of Alberta.

Lagan P, Mannan S, and Hisashi M (2007) Sustainable use of tropical forests by reduced-impact logging in Deramakot Forest Reserve, Sabah, Malaysia. *Ecological Research* 22(3): 414–421.

Laiolo P, Caprio E, and Rolando A (2003) Effects of logging and non-native tree proliferation on the birds overwintering in the upland forests of northwestern Italy. *Forest Ecology and Management* 179(1-3): 441–454.

Laurance WF and Bierregaard RO (1997) *Tropical Forest Remnants.* Chicago: University of Chicago Press.

Laurance WF, Croes BM, Tchignoumba L, *et al.* (2006) Impacts of roads and hunting on Central African rainforest mammals. *Conservation Biology* 20(4): 1251–1261.

Laurance WF and Useche DC (2009) Environmental synergisms and extinctions of tropical species. *Conservation Biology* 23(6): 1427–1437.

Lindenmayer DB, Franklin JF, and Fischer J (2006) General management principles and a checklist of strategies to guide forest biodiversity conservation. *Biological Conservation* 131: 433–445.

Lindenmayer DB and Hunter ML (2010) Some guiding concepts for conservation biology. *Conservation Biology* 24(6): 1459–1468.

Lindenmayer DB, Hunter ML, Burton PJ, and Gibbons P (2009) Effects of logging on fire regimes in moist forests. *Conservation Letters* 2(6): 271–277.

Lindenmayer DB and Noss RF (2006) Salvage logging, ecosystem processes, and biodiversity conservation. *Conservation Biology* 20(4): 949–958.

Meijaard E, Sheil D, Nasi R, and Stanley SA (2006) Wildlife conservation in Bornean timber concessions. *Ecology and Society* 11(1): Article No. 47. http://www.ecologyandsociety.org/vol11/iss1/art47/

Nepstad DC, Verissimo A, Alencar A, *et al.* (1999) Large-scale impoverishment of Amazonian forests by logging and fire. *Nature* 398: 505–508.

Newton AC, Cayuela L, Echeverría C, *et al.* (2009) Toward integrated analysis of human impacts on forest biodiversity: lessons from Latin America. *Ecology and Society* 14(2) http://www.ecologyandsociety.org/vol14/iss2/art2/

Oates JF (1999) *Myth and Reality in the Rain Forest: How Conservation Strategies are Failing in West Africa*. Berkeley: University of California Press.

Poore D, Burgess P, Palmer J, Rietbergen S, and Synnot T (1989) *No Timber Without Trees: A study for ITTO.* London: Earthscan Publications.

Poulsen JR, Clark CJ, Mavah G, and Elkan PW (2009) Bushmeat supply and consumption in a tropical logging concession in Northern Congo. *Conservation Biology* 23(6): 1597–1608.

Putz FE, Sist P, Fredericksen T, and Dykstra D (2008) Reduced-impact logging: Challenges and opportunities. *Forest Ecology and Management* 256: 1427–1433.

Richards PW (1996) *The Tropical Rainforest: An Ecological Study*, 2nd edn. Cambridge: Cambridge University Press.

Robinson JG, Redford KH, and Bennett EL (1999) Wildlife harvest in logged tropical forests. *Science* 284: 595–596.

Smith RGB, Nichols JD, and Vanclay JK (2005) Dynamics of tree diversity in undisturbed and logged subtropical rainforest in Australia. *Biodiversity and Conservation* 14: 2447–2463.

Sodhi NS, Lee TM, Koh LP, and Brook BW (2009) A meta-analysis of the impact of anthropogenic forest disturbance on Southeast Asia's biotas. *Biotropica* 41(1): 103–109.

Struhsaker TT (1997) *Ecology of an African Rain Forest: Logging in Kibale and the Conflict between Conservation and Exploitation.* Gainesville: University Press of Florida.

Sturtevant BR, Fall A, Kneeshaw DD, *et al.* (2007) A toolkit modeling approach for sustainable forest management planning: Achieving balance between science and local needs. *Ecology and Society* 12(2): Article No. 7. http://www.ecologyandsociety.org/vol12/iss2/art7/

Terborgh J (1988) The big things that run the world—A sequel to E.O. Wilson. *Conservation Biology* 2: 402–403.

Thorpe HC and Thomas SC (2007) Partial harvesting in the Canadian boreal: Success will depend on stand dynamic responses. *The Forestry Chronicle* 83(3): 319–325.

Wilson EO (1987) The little things that run the world. *Conservation Biology* 1(4): 344–346.

Biodiversity in Plant Breeding

Irwin L Goldman, University of Wisconsin-Madison, Madison, WI, USA

Published by Elsevier Inc.

Glossary

Commodification The act of making something a commodity; or a traded item in the marketplace that is not differentiated from other similar types. An example of an agricultural commodity is Number 2 Yellow Corn, which has certain market characteristics and can be produced by many different cultivars of corn.

Cultivar Literally "cultivated variety," or named entity developed by plant breeders and used by farmers to plant a crop.

Genetic variation Differences within and/or among organisms that are due to genetic causes.

Genetically modified organism An organism produced through introduction of a transgene or transgenes, most often by means of biotechnology.

Genotype The genetic makeup of an organism.

Landrace Broadly adapted population used by farmers in a specific geographic region. Often developed by farmers and often possesses substantial variation compared to modern cultivars.

Plant breeding The systematic process of genetic improvement of plants for food and agricultural uses.

Trait A quality or characteristic that an organism may possess through inheritance.

Transgene/transgenic Possessing a gene or genes that were transferred from another organism, most often by means of biotechnology.

Variety Synonymous with cultivar.

Don Duvick stated that "the goal of crop agriculture is to reduce biodiversity (biodiversity refers to the range of organisms present in the environment, and genetic diversity refers to the variation within and among those organisms that is due to genes) in favor of increased and easily accessible food supplies" (Duvick, 2007). Indeed, modern agriculture itself is characterized by ongoing attempts to minimize the number of species cultivated or raised, along with their attendant pathogens and pests, while maximizing the yield of food and fiber. For example, a modern Midwestern US farm may possess only a few major species such as corn, soybean, wheat, or alfalfa, in contrast to a much larger species diversity on the same farm that would have included dairy cattle, swine, sheep, poultry, vegetables, herbs, and other grains prior to the middle of the twentieth century.

Industrialization of crops such as corn during the twentieth century has provided opportunities to produce many more products from the same raw material, including feeds, oils, starches, food additives, fabrics, packaging materials, sweeteners, and fuels. Plant breeding (plant breeding is the systematic process of genetic improvement of plants for food and agricultural uses) has played a critical role in this process by producing improved cultivars capable of multiple end uses. Commodification of the raw product reduces prices for consumers and processors and makes more goods and services available with less cost to the public. Government price support programs have fostered cropping systems that focus on fewer, major crop species and their supporting technologies, such as herbicides, insecticides, and fertilizers. This in turn has kept commodity prices low and made food prices a much smaller share of the US family budget than in many other countries worldwide. While the efficiency of this approach is one of the hallmarks of US agriculture, the concomitant reduction in biodiversity on the farm is a similarly defining aspect of our modern farming system.

But has this reduction in biodiversity been important for US agriculture, and is it important for our future? Has this reduction in biodiversity also caused important changes in genetic diversity? By focusing our goals on fewer refined species throughout the process of domestication and breeding, humans have substantially reduced the available biodiversity within these species. This in turn has lead to cases of genetic bottlenecks, genetic vulnerability, and crop failures. But it has also led to huge increases in agricultural productivity; and interestingly, some breeding efforts have also reversed genetic bottlenecks, resulting in increases in phenotypic diversity of useful traits. Modern biotechnology has both increased genetic diversity through identification of useful allelic variants and reduced genetic diversity through the widespread deployment of single transgenes in major commodity crops across many countries. During all of agriculture's history, introgression of genes from wild and feral species has continued to modify domesticated plants and bring both wanted and unwanted genetic diversity into the mix.

The US public and private sectors have each contributed in many ways to both reductions and expansions of genetic diversity in crop breeding, fueling the growth of modern scientific breeding through the development of theory and practice, and fostering the growth of agricultural production and food processing technologies. Today, the public sector focuses more on the preservation and enhancement of genetic diversity, through gene banks and prebreeding activities, while the private sector focuses primarily on the development of elite cultivars for crop production.

Since the 1980s, many countries have adopted a legal framework that supports the protection of intellectual property rights for crop germplasm. This further fueled the growth of the agricultural private sector, which today dominates crop breeding activities on a global scale. In addition, the global focus on germplasm rights has resulted in treaties that

emphasize bilateral agreements between countries for germplasm exchange. While these treaties may create new remunerative relationships between countries, they may also reduce the trade in genetic resources and reduce biodiversity.

The plant breeding landscape may thus be characterized by many fluxes of biodiversity in different directions, some of which are in tandem, some of which are competing, and some of which are entirely independent of each other. The paradoxical nature of these fluxes encompasses (1) breeders both reducing and enhancing diversity for various traits, (2) biotechnologists and breeders incorporating novel allelic variants in the form of transgenes or other unique variants into crop cultivars and deploying these cultivars over vast acreages of agricultural land, (3) consumer demand driving reductions in diversity and increases in uniform commodities, (4) government programs sustaining diversity levels by supporting current major crops and their cultivation systems, (5) a legal framework that supports intellectual property rights for crop germplasm and mandates bilateral agreements between countries for future germplasm exchange, (6) preservation, maintenance, and enhancement of genetic diversity through public gene and seed banks, and (7) natural and artificially induced gene transfer from wild relatives. All of these elements contribute to the dynamic patterns found in crop genomes and on farmer's fields. The goal of this article is to explain and highlight the nature of these fluxes in crop biodiversity and to characterize their importance and future in plant breeding terms.

Modern Scientific Plant Breeding

Although plant breeding dates to the beginnings of agriculture, modern scientific breeding dates back only a little more than 100 years. Mendel's work describing the fundamentals of heredity was published in 1865 but lay dormant until 1900, when it was more widely publicized in Europe. After 1900, the genetic revolution that ensued had a strong impact on plant breeding. Although breeders had unknowingly been using many procedures that exploited genetic variation via crossing and selection in the nineteenth century, the merging of quantitative genetic theory with Mendelian heredity provided a clear foundation for the modern science to begin (Janick and Goldman, 2003).

For cross-pollinated crops, the first half of this period was characterized by refinement of landraces and open-pollinated populations and substantial efforts to produce inbred lines from these populations. An important element of this early work was the recognition that crops have a distinct center or centers of origin, which correspond to their centers of diversity. These ideas were first advanced by Nicolai Vavilov, a Russian botanist who was also one of the world's leading geneticists (Crow, 1993). Through the work of Vavilov, plant breeders became aware of the need to collect, preserve, and study landraces and other germplasm from the various centers of origin of their crop species. This in turn opened up opportunities for exploration and trade of germplasm resources specifically for the purpose of breeding.

Once plant breeders had evaluated landraces and open pollinated populations for useful traits, and sufficient inbreeding had been accomplished, inbred lines were crossed into hybrid combinations to form productive F1 hybrids. By the 1970s, F1 hybrids could be found not only in agronomic crops like corn, but also in horticultural crops such as onions and carrots, as well as dozens of ornamental species. However, the drive toward productive F1 hybrids was part of the reason that diversity issues aroused great concern among plant scientists and the general public.

Perhaps the most defining twentieth century moment in crop genetic diversity occurred in 1970 when Southern corn leaf blight, caused by *Helminthosporium maydis* (Nisik and Miyake) (now called *Cochliobolus heterostrophus* (Drechs.) Drechs), caused widespread destruction of the US corn crop. A majority of the F1 hybrid corn in the US contained "Texas" sterile cytoplasm for the purpose of making F1 hybrids, and the Texas cytoplasm was coincidentally susceptible to this pathogen. The genetic uniformity of approximately 80% of the US corn crop stimulated government and private sector scientists to look for alternatives, and spawned numerous commissions to study the genetic variability of US crops. One of the most important landmarks to result from that work was a report from the National Research Council (1972) that highlighted the need for genetic diversity in our crops (National Academy of Sciences, 1972). This report described that the key lesson of the 1970 epidemic was that genetic uniformity was the basis of crop vulnerability. The Committee on Genetic Vulnerability of Major Crops who authored the report found that most major crops were "impressively uniform genetically and impressively vulnerable." Importantly, the Committee also found that the uniformity derived from "powerful economic and legislative forces," and that the situation posed national challenges that must be addressed. One of their recommendations was a national facility for continuously maintaining gene pools at home and abroad, which today is carried out by the National Plant Germplasm System. The fact that the Committee identified the importance of "economic and legislative forces" 40 years ago underscores one of the key messages of this article: fluxes affecting biodiversity in plant breeding are many, varied, and complex.

In the intervening 40 years since the National Academy of Sciences report, genetic resources and genetic variation have been important foundation concepts in plant breeding. Scientific societies such as the Crop Science Society of America created space for study of these issues, and in 1990 included a section called Plant Genetic Resources in their chapter. Scientific journals such as *Crop Science* began to devote separate sections to genetic resources, highlighting the importance of this area to plant improvement and production agriculture. Numerous studies and publications have highlighted the importance of identifying, preserving, and maintaining crop genetic diversity for the future of food and agriculture, and public funding has been directed toward increasing global efforts in these areas. In contrast to the earlier part of the twentieth century, plant breeders today possess a high level of awareness and readiness with respect to the maintenance of genetic diversity.

It is, however, still important to note that the replacement of older cultivars with newer, higher-performing ones is a hallmark of successful plant breeding. Historical descriptions of crop breeding regularly mention the disappearance of older cultivars and the replacement of such favorites by newer, high performing varieties. Such disappearance can result in a loss of



Figure 1 Homogeneity in crop fields in the US Wheat harvest in the Palouse range of Washington. Courtesy USDA.

biodiversity. In some ways, the rapid progress of modern plant breeding is responsible for the parallel growth in trade of heirloom seeds and for the activities of organizations such as Seed Savers Exchange (Decorah, Iowa). These efforts include elements of gene conservation and preservation, a great appreciation for classic and familiar cultivars, and social commentary on modern agriculture. Kloppenburg (2004) eloquently described how commercialization of technologies such as F1 hybrid seed have contributed to reduced diversity and further commodification of our crops (**Figure 1**). He also points out how current efforts to combat these trends have flowered in recent decades, providing avenues for resurrecting the public aspects of crop breeding. Despite this, examples of the homogenization of crop germplasm may be found among certain industrial crops where strict standards are required by new cultivars. The American Malting Barley Association provides a list of criteria that must be satisfied for new barley cultivars (**Figure 2**), suggesting the possibility of narrowing the germplasm pool if those criteria are met. However, there is contradictory evidence (see Does Variation Become Reduced Through Breeding?) on whether such focused breeding programs actually reduce genetic diversity (Rasmusson and Phillips, 1997).

Barley factors
Plump kernels
Thin kernels
Germination
Protein
Skinned and broken kernels
Malt factors
Total protein
On 7/64 screen
Measures of malt modification
Beta glucan
F/C difference
Soluble/total protein
Turbidity
Viscosity
Congress wort
Soluble protein
Extract
Color
FAN
Malt enzymes
Diastatic power
Alpha amylase

Figure 2 Quality guidelines for commercial malting barley that are made available to barley breeders, as published by the American Malting Barley Association, Inc. There are 18 criteria in total, and each criterion has a certain range of acceptable values, for example, plump kernels must be $>90\%$ and thin kernels must be $<3\%$.

Does Variation Become Reduced Through Breeding?

A common refrain about modern plant breeding is the concern about its potential to reduce genetic diversity by increasing selection pressure on the uniform conditions typical of modern agriculture. Plant breeding is an evolutionary process that results in changes in the genetic makeup of plant populations. Some changes in diversity are to be expected, but patterns of change have not always been consistent, and it is not accurate to say that plant breeding is unidirectional with respect to diversity. Breeders incorporate or make use of many of the same elements of natural selection in the core aspects of breeding programs, such as fecundity, heritable variation, and differential reproduction and survival. Variation is both created and destroyed in breeding populations, and one of the key goals of breeders is to manage this variation to the betterment of crop cultivars. An example of the creation of variation is the recombination that takes place in thousands of meiotic events as a breeder makes crosses between strains in a breeding program. This recombination generates thousands of new genetic combinations, some or all of which may never have previously been produced in any breeding program (**Figures 3** and **4**). It is also possible that some of these recombination events will result in transgressive segregation, resulting in individuals that fall outside of the mean values of parental traits. These new genetic combinations may be planted and evaluated by the breeder. A breeder evaluating these recombinant populations may choose an individual or



Figure 3 Phenotypic variation in maize kernels in a collection from the GEM (Germplasm Enhancement of Maize) project. Courtesy USDA.



Figure 4 Phenotypic variation in bean (*Phaseolus vulgaris* L.) seed coat patterning and coloring. Courtesy USDA.

individuals at the tail end of the population and discard the rest. Variation is therefore both created and discarded in each breeding cycle.

Many examples suggest a loss of genetic variation and allelic diversity through concerted breeding efforts, but a substantial body of work demonstrates that the opposite is also possible. One of the most famous selection experiments of the twentieth century is the Illinois Long Term Selection experiment conducted with maize at the University of Illinois in Urbana (**Figure 5**). The origins of the experiment were influenced primarily by work conducted in France by early breeders of the fodder beet (Hopkins, 1899; Troyer, 1996; Dudley and Lambert, 1992), and the early focus of this experiment was modification of the corn kernel's chemical makeup. The experiment was conceived by Cyril Hopkins, a chemist. The goal of his project was based on "Darwin's principle of selection," which was employed to modify the chemical composition of the corn kernel (Goldman, 2004). Perhaps the most unusual result of this experiment is the fact that genetic variation remains even after more than 100 generations of selection in a closed population. The observation of continuing gain from selection despite what would appear to be large reductions in genetic variance and large increases in inbreeding that have



Figure 5 Queen Anne's Lace, also known as wild carrot, is a ubiquitous weed that may have unique genes for crop improvement.

accompanied more than 100 generations of selection is startling. The paradox of this long term selection experiment is the cause for celebration among plant breeders, who often see it in terms of a limitless source of variation for selection and a reason they can continue to make gain from selection in breeding populations (Goldman, 2004). While it is possible that maize is unique among crop plants for the flexibility of its genome, these results suggest that genetic diversity may not be fully understood, even in the context of closed breeding populations studied for centuries.

Rasmusson and Phillips (1997) described a paradox involving continued plant breeding progress for barley (*Hordeum vulgare* L.) malting traits despite what appeared to be a continually narrowing genetic base for the crop. The authors suggested *de novo* variation in the form of transposable elements, transcription factors, intragenic recombination, epigenetic phenomena, and the like, may be arising during the selection process, thereby creating variation on which plant breeders can select. While our understanding of such phenomena in the context of breeding populations is currently limited, it is likely that research investments in this area will yield important insights into the flux of genetic diversity during the breeding process. If, like in the case of the Illinois Long Term Selection experiment, breeding populations are capable of producing substantial and useful *de novo* variation while under selection, the apparent loss of biodiversity through narrow breeding objectives may be mitigated. However, much more remains to be learned on this topic before plant breeders can claim *de novo* variation as a cornerstone of their breeding programs.

Huang *et al.* (2007) investigated whether modern winter wheat (*Triticum aestivum* L.) breeding in Europe has resulted in a loss of genetic diversity within and among wheat varieties. They evaluated 511 widely grown varieties from Central and Northern Europe using microsatellite markers. By far, the largest source of variance in their study was the variance among

varieties within decadal groups between the period 1945 and 2000. Only a small share of the variance was due to differences between the decadal groups. Their analysis suggested that while allelic changes have occurred in modern plant breeding, no apparent loss of genetic diversity or allele number was detected in this collection of varieties. Fu *et al.* (2003) studied a similar question in oats (*Avena sativa* L.) and found certain specific marker alleles that were reduced in diversity from 1886 to 2001, but that average genetic diversity across 96 Canadian oat cultivars as described by these markers did not decrease in a systematic fashion. Fu and Somers (2009) examined genetic diversity in wheat cultivars released from 1845 to 2004 using 370 simple sequence repeat markers. These workers found allelic reductions in every part of the wheat genome, with significant reduction beginning in the 1930s. While they concluded that the study provided clear evidence for reductions in genetic diversity through breeding, it is not clear to what extent these simple sequence repeat markers represent a loss of useful genetic variation for crop improvement. Indeed, many molecular markers are neutral to phenotype and thus it is difficult to extrapolate from such markers to phenotype. It is, however, important to note that the allelic reduction noted was substantial and genome-wide, suggesting broad patterns in the wheat genome that are worthy of further study.

An explosion of interest in certain crops has meant a corresponding increase in the available genetic resources for plant breeders. This *et al.* (2006) indicated that the genomic resources available to the grape (*Vitis vinefera* L.) research community have increased substantially since 2001, partly because of the renewed interest in this crop. Sexual recombination and asexual reproduction have expanded and diversified grape germplasm collections (This *et al.*, 2006), and more germplasm is currently cataloged and available to breeders than ever before. In this way, consumer demand for grape products such as wine, table grapes, and raisins has resulted in increased investment in grape germplasm collections and breeding efforts, which in turn means more available grape diversity. Public and commercial interests have the potential to drive diversification and expansion of breeding programs. By the same token, crops for which consumer demand has weakened may find correspondingly less breeding activity. Table beet (*Beta vulgaris* L.), a minor crop in the US, is no longer bred by most US vegetable seed companies and is represented in only one public US breeding program (Goldman and Navazio, 2002). Genetic resources for this crop may not be as diversified and well-characterized as a result, and the total amount of diversity found decades ago may not be able to be preserved or maintained.

Genetic Bottlenecks

A feature of most domesticated plants is a reduction in genetic diversity following domestication (Gepts, 2004). Many studies using molecular markers or gene sequence information have demonstrated reductions in diversity during and after domestication. However, some studies based on phenotypic data have shown increases in phenotypic diversity following domestication. One of the central theses of Darwin's *Variation in Plants and Animals Under Domestication* (Darwin, 1868) was that variation is created and preserved in those plant organs on which humans depend. Seeds of *Phaseolus vulgaris* L., for example, have highly varied seed coat patterns, likely because variants have been observed, collected, and preserved by humans. The same is true for shape and color of potato tubers. Traits with very strong selective advantage, such as those of interest to humans for food production, are likely to be preserved under domestication. Molecular markers are often neutral and may not be selectively advantageous. It is therefore difficult to predict the degree to which they would survive under domestication. In general, domestication events seem to be associated with a loss of diversity, but intense selection following domestication appears to be responsible for increasing phenotypic diversity. The bottleneck through which many of our crop species has come has itself suggested breeding strategies that seem to show some promise.

Genes from related wild species have been promising for traits like disease and pest resistance, but breeders have historically shied away from them for productivity traits. This is primarily because wild species chromosome segments also may carry unfavorable alleles that can dramatically alter crop productivity. Tanksley and McCouch (1997) examined wild rice species accessions and crossed them to elite inbred lines. They found increases in yield from early-generation progeny of these crosses, suggesting wild species had much more to contribute in terms of productivity traits than previously thought. Gur and Zamir (2004) described useful and productive variation from wild species of tomato that was capable of increasing yield over elite commercial cultivars while using less water. Interestingly, they describe the variation in these wild species as "unused," which suggests that modern plant breeding has left useful variation in its germplasm banks, waiting to be incorporated into cultivars. Zamir and colleagues also created an introgression population of *Solanum pennelli* chromosome segments in an elite *Lycopersicon esculentum* background (Eshed and Zamir, 1995). Some of these introgression lines exhibited higher yield than elite genotypes, suggesting great promise for wild species chromosome segments. Johal *et al.* (2008) have recently described a technique known as mutant-assisted gene identification and characterization (MAGIC) that makes use of mutants or variants in a trait of interest to identify novel variants for the trait. This technique, and others like it, has the potential to identify useful natural variation that can be incorporated into breeding populations (**Figure 6**). Increasingly, breeders have become interested again in variation that is natural and the result of millennia of selection in agricultural conditions (Zamir, 2008). Studies such as these have fueled renewed interest in gene banks and wild species, not only as a source of important disease and pest resistant characters, but also as a source for productivity traits. In addition, the incorporation of wild species genes into elite inbred lines and cultivars has resulted in greater diversity in breeding populations. This is yet another example of how breeding activities can contribute to increasing diversity in crop germplasm.

Gene Pool Concepts

Crop genetic resources are the genes and gene combinations available for crop improvement. For the past 40 years, since

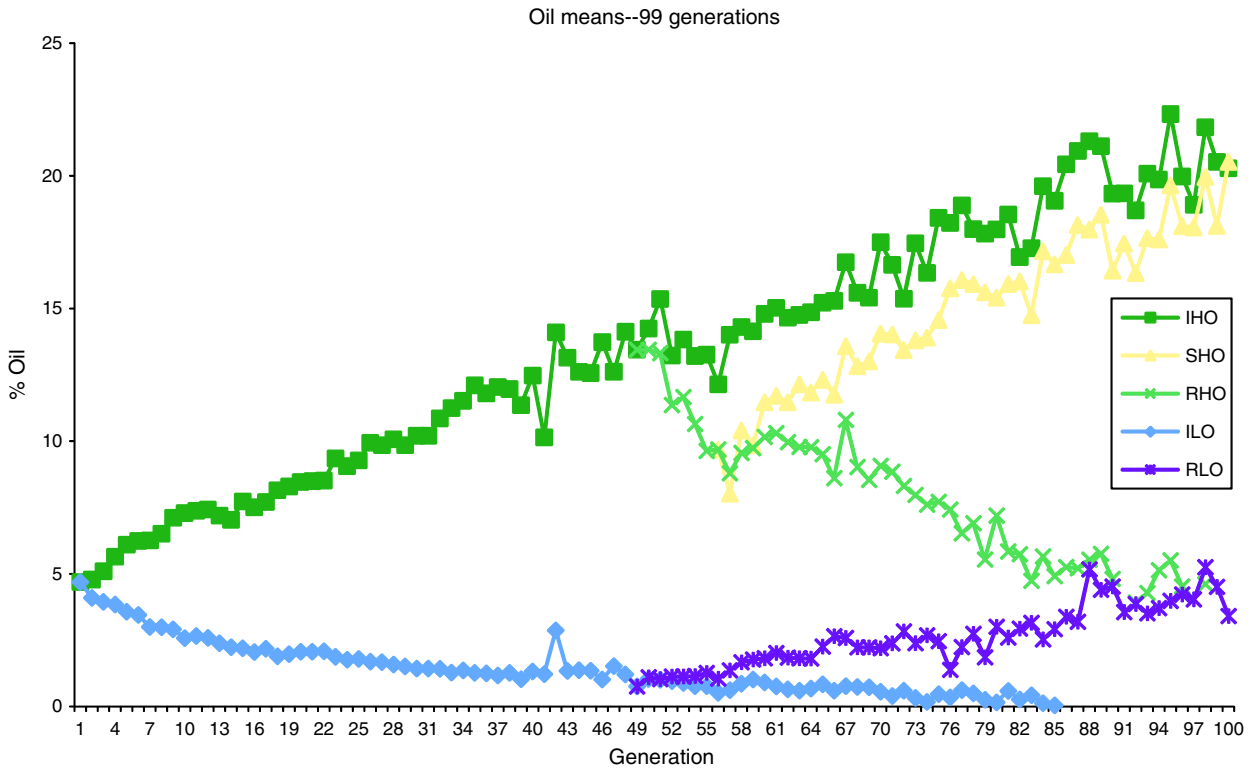


Figure 6 Response to selection for oil and protein in maize in the Illinois Long Term Selection Experiment (Reproduced from Dudley JW and Lambert RJ (1992) Ninety generations of selection for oil and protein in maize. *Maydica* 37: 1–7.).

the work of Harlan (1971) and Harlan and de Wet (1971), crop genetic resources have been categorized in three gene pools. Gene Pool I included the crop and its wild progenitors, and is typically consistent with the biological species concept (Gepts, 2006). Crossing should be successful among species in this gene pool. Gene Pools II and III reflect increasing genetic distance away from the crop species, respectively, with crosses much more difficult, if not impossible, to make. Occasionally, such crosses require the aid of techniques such as embryo rescue and tissue culture. Gepts has also suggested (2002, 2006) the need for enumeration of a Gene Pool IV that reflects the genetic diversity that can now be accessed through biotechnology. Some biotechnological approaches make it possible to clone nucleic acids from any organism and insert them into a vector that can be introduced into a plant. This makes genetic resources available for crop improvement that were not previously possible to access. Interestingly, Gene Pool IV may be accessed by any of the other three Gene pools, though it will likely require molecular techniques in each case. Gepts (2006) pointed out that while one of the contributions of biotechnology to plant breeding should be a broadening of the gene pool and resulting biodiversity increases, it has thus far had a limited role because of the limited number of genes thus far available and our lack of understanding about their interaction with the plant genome.

It should also be mentioned that farmers themselves are critical curators of genetic diversity for current and future plant breeding efforts. While many US farmers may purchase seed from companies each year, farmers in the developing world collect, curate, and maintain genetic diversity for themselves and their neighbors. This is true for many crops, including major industrial crops such as corn, where farmers in the developing world still grow the majority of the world's acreage and manage the seed resources for the crop. Pearl millet (*Pennisetum glaucum* L.) landraces are grown and maintained by farmers throughout the world. A study by vom Brocke *et al.* (2003) examined the management practices of farmers in Rajasthan, India with respect to pearl millet germplasm. These workers found that crop management practices and seed systems practiced by farmers in the region were the primary determinants of genetic diversity patterns of the landraces and cultivars grown. Some regions exhibited landraces that possessed more intravillage variation while other regions showed landraces with more intervillage variation, depending on seed practices. These practices are fundamental plant breeding practices, and as such are important regulators of genetic diversity in crop germplasm.

Duvick (2007) described genetic diversity in place and genetic diversity in time as ways that farmers and agriculturists have preserved biodiversity. Genetic diversity in place is a type of spatial diversity, where farmers plant several kinds of crops or several varieties of a crop to insure some of them would survive poor weather or pests. In general, farmers still practice the use of genetic diversity in place by planting different cultivars of the same crop in the same field with the hope that some will do well regardless of the environmental conditions. Other farmers practice a more dramatic form of diversity in place by growing many species in the same field, or by intercropping two or more different species in the same field. The use of multiple species of crop in the same field is sometimes

Regional stations	Clonal repositories	Crop collections	Genetic stocks
NC-7, Iowa	Brownwood, Texas	Arctic plants, Alaska	Maize, Illinois
NE-9, New York	Corvallis, Oregon	Arid land plants, California	Rice, Stuttgart
S-9, Georgia	Davis, California	Clover, Kentucky	Peas, Washington
W-6, Washington	Geneva, New York	Cotton, Texas	Tomato, California
Special collections	Hilo, Hawaii	Desert Legumes, Arizona	
National seed storage, Colorado	Mayaguez, Puerto Rico	National Arboretum, DC	
Plant Quarantine, Maryland	Miami, Florida	Potato, Wisconsin	
	Riverside, California	Soybean, Illinois	
		Small grains, Idaho	
		Tobacco, North Carolina	
		Ornamentals	

Figure 7 Key stations, repositories, collections, and stock centers of the National Plant Germplasm System of the US Department of Agriculture.

referred to as a polyculture. The classic Central and South American example of this type of intercropping is maize and beans grown in the same hill, where the beans twine around the maize plant for support. Native Americans also combined these plantings with squash, increasing native diversity levels. Another form of increased genetic diversity in a cultivar is the multiline, which is essentially a mechanical mixture of individual lines, each bred to have a specific trait. An example of such a cultivar would be one bred to combine resistance to individual races of a pathogen by combining inbred lines, each carrying a resistance allele.

Genetic diversity in time is based on the fact that for most crops, seed planted in any given year is not genetically identical to the seed produced by the crop in a subsequent year (Duvick, 2007). This constant shifting of genotypes, characteristic of the process of evolution via natural selection, means that cultivars exhibit changes in their genetic makeup over time simply through the process of cultivation. Duvick wisely observed that "the sequential genetic changes represented continuing adaptation to prevailing constraints to yield."

Many crop species now have a number of resources attached to them with respect to germplasm preservation. For example, key crop species are often preserved in *ex situ* and *in situ* collections worldwide. Examples of the former include gene banks and seed banks, while the latter includes biodiversity preserves that attempt to conserve germplasm as a living organism in its habitat. Care is taken to include germplasm from various Gene Pools in these collections, and many collections have curators whose specialty involves characterization and preservation of the germplasm. Some of the collections include "core collections," which are meant to capture a full range of genetic diversity in a limited number of germplasm accessions. These accessions then can be used by researchers and breeders to evaluate particular traits, such as disease or pest resistance, and conclusions can be drawn from this sample that may be extrapolated to the entire collection. Breeders maintain their own collections of germplasm for use in current and future situations, along with segregating populations derived from crosses between Gene Pools. Some crops may also have genetic stock centers maintained by public funds that allow for the creation, maintenance, and study of unique genetic materials for research and breeding. All of these structures help preserve biodiversity in plant breeding and ensure a future where genetic resources will be available (**Figure 7**). In some cases, these structures increase biodiversity and foster a climate where new genetic combinations can be created and studied.

Exploration and Trade

It is worth noting that the voyage of Columbus was to find a route to Asia in order to acquire spices for Europe. His trips led to what first became known in 1972 as the "Columbian

Exchange'' (Crosby, 2003), where plant species and other organisms moved between Europe and the Americas. This single event, or series of events, led to explosive growth in crop genetic resources for both Europe and the Americas, and diversified the genetic resources in both continents. It also had the effect of homogenizing agricultural landscapes on the two continents, for example, by increasing the growth of corn in Europe and the growth of wheat in the Americas. An individual's diet in each country may have become more diversified, but agricultural land in both continents may have come under more similar patterns of use. The scale is important. Locally, regions were diversified and biodiversity increased. Globally, regions became more similar. Some would argue that the crops were successful in getting humans to spread them even farther and prepare landscapes for them to thrive (Pollan, 2006). Regardless, exploration and trade shuffled genetic resources between localities that had never seen them before, while also making some aspects of crop agriculture more uniform. This example highlights the push and pull of biodiversity, reinforcing the notion of contrasting fluxes operating in a complex landscape.

The US government became actively involved in plant and seed acquisition in the nineteenth century and began building a national plant germplasm system, which today is managed by the US Department of Agriculture and holds many thousands of accessions of hundreds of species. The US Department of Agriculture still employs scientists whose primary role is to collect and characterize wild crop relatives. While substantial genetic resources are held in this repository and made freely available upon request, humans rely on only approximately 100–150 crop species for their survival, and of these, only approximately 15 make up the bulk of calories in the US diet. One could argue that biodiversity is preserved by the existence of seed banks such as the National Seed Storage Laboratory in Fort Collins, Colorado. But one could also argue that biodiversity is threatened because the existence of some of these species is only as a static entry in a germplasm bank and not as a growing and evolving species in either agriculture or nature. Regardless, plant breeding does and will continue to depend on these genetic resources for its future. These accessions make up the raw material for many of the day-to-day activities of plant breeders. The simple fact that they are preserved secures a measure of biodiversity for the planet and provides hope for future breeding efforts.

The Legal Framework

Two treaties cover the legal framework for crop biodiversity: the UN convention on Biological Diversity, and the International Treaty on Plant Genetic Resources for Food and Agriculture. The former covers genetic diversity in animals, plants, and microbes, while the latter focuses on global food security. The Convention on Biological Diversity moves away from the concept of germplasm and genetic diversity as the common heritage of humankind, and focuses instead on the development of bilateral agreements that include benefit sharing among countries that hold or desire genetic resources. The US has not ratified the Convention on Biological Diversity, but has on certain occasions participated in bilateral agreements with other countries for germplasm exchange. In the main, however, the US National Plant Germplasm System has a fundamental goal to make germplasm freely available to any and all requestors; a situation that is almost unique among world germplasm banks. The International Treaty on Plant Genetic Resources for Food and Agriculture focuses solely on plant genetic resources used to produce food, and covers 64 crops. This treaty contains a pool of genetic resources that are available freely to those countries that sign on and agree to use the material only for research and plant breeding activities for the benefit of food and agriculture.

Patents for plant germplasm have also played an important role in the diversity landscape for the past 25 years. Although prior to 1980, the only patents available for plant germplasm were for asexually propagated species, changes in US law allowed for cultivars, genes, plant breeding processes, and associated technologies to be patented. In some cases, utility patents were issued that covered broad-based breeding procedures and the cultivars they produced. In the 1970s, the Plant Variety Protection act was adopted, which led to increased opportunity for protection of seed-propagated crop species. Importantly, the Plant Variety Protection Act did not restrict breeding with cultivars protected in this way, whereas cultivars that were patented could not be used before obtaining a license from the owner.

The implications of these two treaties and the trend toward plant patents with respect to biodiversity are important for plant breeding. In a previous era, the concept that germplasm was the common heritage of humankind prevailed, and germplasm was exchanged freely among plant breeders. Public sector plant breeders often adopted ''codes of ethics'' or protocols that provided for attribution, when appropriate, for the use of germplasm from other breeding programs. However, free exchange was the rule rather than the exception. Since the 1980s, public institutions with breeding programs have sought to secure revenues for germplasm developed by their scientists. Initially, this was accomplished by individual breeders who sought to bring royalties for germplasm directly back to their breeding programs. But with cutbacks in public funding and a national trend toward identifying new revenue sources for public institutions in the 1990s, many sought to formalize contractual agreements for germplasm whereby licensees or purchasers would pay royalties to the technology transfer arm of the institution. Many institutions sought legal protection for plant germplasm, in the same way a private company would normally do so. This aspect of technology transfer is only approximately 20 years old in the US and it may be too early to tell if it has been a successful model. Anecdotally, however, there is some evidence that it has reduced germplasm exchange among breeding programs because of protective clauses in germplasm agreements. In contrast, there is also some evidence that royalties generated have helped support public breeding programs, which otherwise may have folded due to lack of public funding. The paradox inherent in this institutional solution is obvious: public breeding programs in danger of becoming extinct rely more and more on royalty generation, which has the potential to further reduce germplasm exchange, limit breeding activities, and thereby impinge on genetic gain. Recently, some have suggested open-source models of germplasm release similar to approaches used by the computer

software industry (Open Plant Breeding Foundation, website). Such approaches have the potential to reaffirm the public nature of public breeding programs, but the issue of financial support for such programs is still a concern.

Commodity Agriculture and Biodiversity

One of the most remarkable and successful features of US agriculture is the way animals and crops have become commodified, increasing the efficiency of the market, reducing costs for processors and consumers, and hedging natural and human-caused risks inherent in agriculture. Commodity agriculture has also spawned markets for speculators, who link farmers with end-users in an attempt to profit from changes in commodity prices. One could argue that consumers ultimately benefit from this arrangement through lower prices and more reliable supply of the commodities they need. However, commodity agriculture favors uniformity of product and therefore may impinge on crop diversity. For example, the very nature of commodity corn means that any particular corn kernel is not differentiated from any other corn kernel, and their equal treatment leads the crop itself to be considered monolithic with respect to the market. Because many types of commodity corn are not identity preserved (there are certainly types that are, including corn for ethanol and certain types of animal feed), there is no corresponding incentive for breeders, farmers, or processors to demand diversity. In fact, this system provides incentive for field homogeneity. While one might argue that breeders must still develop ever higher-performing cultivars and thereby increase new genetic diversity into the mix, it is also true that the drive toward sameness could have an impact on genetic diversity. As our commodity markets have become global, this situation promises more homogeneity with respect to crop diversity in the future.

Green Revolution Genes and Biotechnology

Worldwide food production on a per capita basis is higher today than it was 50 years ago (Trewavas, 2001). Plant breeding and its associated technologies have played a significant role in these increases. One example of this was the introduction of dwarfing genes in both wheat and rice that led to the Green Revolution. The Green Revolution refers to technological change in agriculture that occurred in the developing world during the 1960s. The powerful combination of a shift in the harvest index by partitioning more energy into grain production and less to overall biomass, along with significant increases in fertilizers and pesticides, allowed for massive increases in food production in the developing world. The father of the Green Revolution was Normal Borlaug, the only plant breeder to have received a Nobel Prize for Peace for his work on improving world agriculture. The genes used by Borlaug that were responsible for producing dwarf strains of wheat and rice were later discovered to interfere with the signal transduction pathway of gibberellic acid, an important growth hormone (Hedden, 2003). The dwarfing effect produced the dramatic change in harvest index and biomass partitioning that drove increased grain yields.

The wheat dwarfing genes responsible for these changes, *Rht-B1b* and *Rht-D1b*, originated in Japan and were bred into American cultivars in the middle of the twentieth century. Today, these genes are present in approximately 70% of commercial wheat cultivars worldwide. Various alleles of a single recessive gene, *sd1*, are responsible for the semidwarf phenotypes of rice, which have also been used to make dramatic improvements in rice yields around the world. The *sd1* alleles also interfere with gibberellic acid biosynthesis, making modification of this growth hormone a key step toward increasing crop yields during the twentieth century.

From the point of view of biodiversity, one could argue that the widespread deployment of single genes such as *sd1* or *Rht-Blb* in the majority of rice or wheat cultivars grown across the world means that biodiversity is reduced. The homogeneity associated with cultivars carrying these genes may seem to be limited to single genes, but one may also imagine that other genes in the wheat and rice genomes have been similarly selected to perform well in concert with the dwarfing genes. This in turn may unwittingly lead to more crop uniformity. However, one could also argue that the use of these genes has fueled tremendous expansion of these crops throughout the world, increasing the number and scope of breeding programs worldwide, and increasing the number of environments in which these crops can be successfully grown. As breeders incorporate these genes into new genetic backgrounds, additional diversity will be generated. Green Revolution genes may therefore serve as an example of a reduction in biodiversity of crop cultivars as well as a stimulus for generated new diversity and adaptation via breeding.

Biotechnological innovations have revolutionized plant breeding in the past 30 years. Among the most controversial and powerful of these innovations has been the development of transgenic, or genetically modified, crops. The advent of this technology followed the successful development of many workhorse tools such as gene cloning, plant transformation, regeneration, and suitable vectors for hosting genes of interest. By the late 1980s, crop cultivars carrying transgenes were being tested, and the first commercially available transgenic crop was released in 1993. By 1995, substantial acreage in the US was devoted to transgenic corn and soybean for herbicide and insect resistance. The extremely rapid adoption of these technologies demonstrates their utility to farmers, processors, and consumers, though they are not entirely without controversy in the US Still, the US marketplace shows no sign of slowing down with respect to the use of transgenic crops; however the use has been confined primarily to corn, soybean, cotton, and canola. Some transgenic squash is produced, as is some transgenic papaya, and both sugarbeet and alfalfa have recently been deregulated for use as transgenic crops in the US Many transgenes have been inserted into a broad range of crop species in the context of research programs, but there is currently no sign of new transgenic crops entering the commercial marketplace other than the primary crops mentioned above (corn, soybean, cotton, and canola). Some consumers have expressed concerns about certain aspects of genetic modification, including the possibility of gene flow to wild or weedy relatives, the possibility of creating novel, allergenic proteins in transgenic species, and food safety aspects of transgenics, and a lengthy and at times rancorous debate has

ensued in the US and in Europe, where transgenic crops have only recently begun to appear.

Transgenic technology means that nucleic acids from any source may be added to the plant genome. In theory, this greatly expands the possibility of increasing crop biodiversity through the addition of genes from any organism with the potential to improve crops. Likewise, it is also the source of concern for many consumers, who see transgression of the species and kingdom boundaries as potentially problematic. It is important to note that the species and kingdom boundaries are occasionally transgressed in nature, and it is in exactly one such example of this transgression that transgenic crops were originally produced: the soil bacterium *Agrobacterium tumefaciens*, which normally infects plant roots and transfers its nucleic acids to the host. From a biodiversity point of view, transgenic manipulation has the potential to expand biodiversity through the addition of genes from plants that would not have been able to cross in nature or under controlled conditions. Earlier, the author cited Gepts' assertion that a Gene Pool IV should be identified, which would represent those genes found anywhere in nature that have the potential to improve crop genomes. Indeed, one of the most remarkable aspects of transgenic technology is the limitless nature of the possible gene combinations that could be created. So, from a biodiversity point of view, there is much potential for enhancement through this particular tool of plant breeding.

There is a high cost associated with the development and deployment of transgenes in modern agriculture, and that cost is frequently prohibitive for smaller-scale seed companies to develop transgenic cultivars. Larger companies may be able to afford licensing fees, which reach into the millions of dollars, in order to license and obtain transgenes that can then be incorporated into their cultivars. Over time, this situation has the potential to limit the number of entities able to make the investment to produce transgenic cultivars, and may therefore reduce genetic diversity in crop germplasm as a result. In addition, as described above (*see* Green Revolution Genes and Biotechnology) for genes associated with the Green Revolution wheats and rices, the widespread deployment of single genes over millions of agricultural acres both expands the reach of these cultivars and simultaneously increases the genetic uniformity of elite crop cultivars. As transformation technology decreases in price, it may be possible to reverse this trend. However, it remains to be seen how much variation will be limited and how much will be expanded as a result of transgenic technologies.

An interesting recent development in the deployment of transgenes sheds some light on possible biodiversity scenarios for the future. Hutchison *et al.* (2010) evaluated the impact of transgenic maize carrying the *Bt* gene for resistance to European corn borer. Transgenic maize carrying the *Bt* gene had a suppressive effect on European corn borer in both *Bt* corn crops and in non*Bt* crops in neighboring fields. The suspected reason for the benefit to non*Bt* crops is that the female corn borers lay eggs indiscriminately in both *Bt* and non*Bt* crops. Thus, farmers may choose to grow nontransgenic maize, or a mixture of nontransgenic and transgenic maize, and be able to reap the benefits of the technology simply by being in close proximity to others who use the technology. The implications of these findings are of interest for the evolution of pest resistance in particular and for biodiversity in general. If this "halo effect" phenomenon is generally true for transgenic crops or pest resistant crops, the evolution of pests may be slowed and the deployment of transgenes may be less pronounced than one might expect. Interestingly, Ronald (2011) describes studies where the planting of *Bt* crops has had the effect of increasing biodiversity. She cites the work of Marvier *et al.* (2007), who showed that nontarget invertebrates such as spiders, mites, and insects not targeted by *Bt* crops are more abundant in *Bt* corn and *Bt* cotton fields than in conventional fields managed with insecticides.

Summary

Despite the obvious reduction in genetic diversity that can be caused through focused plant breeding efforts, the author has endeavored to show that fluxes in biodiversity in plant breeding are not unidirectional. Rather, plant breeding has the potential to both decrease and increase biodiversity and genetic diversity depending on its objectives. Domestication and concerted breeding efforts have certainly reduced genetic diversity in some cases, but breeding has also expanded genetic diversity in other instances, and detailed analysis should be conducted on a crop-by-crop basis in order to assess, which fluxes in biodiversity have been most important. Today's plant breeding landscape also includes a complex legal framework that has the potential to inhibit germplasm exchange and reduce biodiversity. At the same time, public investments in gene banks and other preservation efforts have resulted in dramatic improvements in collecting biodiversity and hold great promise for future plant breeding. Biotechnology has the potential to result in enhanced genetic diversity of crop species, but deployment of transgenes on a landscape level is fraught with complex scientific and societal questions. At this point, biotechnological approaches have not made as much of an impact on biodiversity as they might in the future. The process of modern plant breeding has impacted crop biodiversity in many ways, not only through the genetic manipulation that is at the core of the breeding activity, but also by the surrounding economic, political, and legal forces that are a part of today's society. Plant breeders are fortunate to be aware of diversity issues in their daily work and have clearly made great strides to protect and enhance biodiversity. Our future food security will depend on their success in this area.

Appendix

Relevant Courses at UW-Madison

1. Agronomy 338, Plant Breeding and Biotechnology
2. Horticulture-Agronomy 501, Principles of plant breeding
3. Horticulture-Agronomy 502, Techniques of plant breeding

See also: Agricultural Invasions. Agriculture, Sustainable. Agrobiodiversity. Biodiversity-Friendly Farming. Breeding of Animals.

Captive Breeding and the Evolutionarily Significant Unit. Conservation Genetics. Darwin, Charles (Darwinism). Diversity, Molecular Level. Diversity, Organism Level. Ecology of Agriculture. Economics of Agrobiodiversity. Edible Plants. Endangered Plants. Evolution, Theory of. *In Situ*, *Ex Situ* Conservation. Feeding the World and Protecting Biodiversity. Human Impact on Biodiversity, Overview. Hybridization in Plants. Inbreeding and Outbreeding. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Loss of Biodiversity, Overview. Plant Biodiversity, Overview. Plant Conservation. Property Rights and Biodiversity

References

Crosby Jr Alfred W (2003) *The Columbian Exchange: Biological and Cultural Consequences of 1492; 30th Anniversary Edition*. Greenwood Publishing Group

Crow JF (1993) Nicolai Vavilov: Martyr to Genetic Truth. *Genetics* 134: 1–4.

Darwin C (1868) *The Variation of Animals and Plants Under Domestication*. London: John Murray.

Dudley JW and Lambert RJ (1992) Ninety generations of selection for oil and protein in maize. *Maydica* 37: 1–7.

Duvick DN (2007) Breeding of plants. *Encyclopedia of Biodiversity* pp. 1–12. Oxford, UK: Elsevier.

Eshed Y and Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTL. *Genetics* 141: 1147–1162.

Fu Y-B and Somers DJ (2009) Genome-wide reduction of genetic diversity in wheat breeding. *Crop Science* 49: 161–168.

Fu Y-B, Peterson GW, Scoles G, Rossnagel B, Schoen DJ, and Richards KW (2003) Allelic diversity changes in 96 Canadian oat cultivars released from 1886 to 2001. *Crop Science* 43: 1989–1995.

Gepts P (2004) Crop domestication as a long term selection experiment. *Plant Breeding Reviews* 24: 1–44.

Gepts P (2002) A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Science* 42: 1780–1790.

Gepts (2006) Plant genetic resources conservation and utilization: The accomplishments and future of a societal insurance policy. *Crop Science* 46: 2278–2292.

Goldman IL (2004) Intellectual legacy of the Illinois long term selection experiment. *Plant Breeding Reviews* 24: 61–78.

Goldman IL and Navazio JP (2002) History and breeding of table beet in the United States. *Plant Breeding Reviews* 22: 357–388.

Gur A and Zamir D (2004) Unused natural variation can lift yield barriers in plant breeding. *PLoS Biology* 2: 1610–1615.

Harlan JR (1971) Agricultural origins: Centers and non-centers. *Science* 174: 468–474.

Harlan JR and de Wet JMJ (1971) Towards a rational classification of cultivated plants. *Taxon* 20: 509–517.

Hedden P (2003) The genes of the green revolution. *Trends in Genetics* 19: 5–9.

Hopkins C (1899) Improvement in the chemical composition of the corn kernel. *Illinois Agricultural Experiment Station Bulletin* 55: 205–240.

Huang X-Q, Wolf M, Ganal MW, Orford S, Koebner RMD, and Roder JS (2007) Did modern plant breeding lead to genetic erosion in European winter wheat varieties? *Crop Science* 47: 343–349.

Hutchison WD, Burkness EC, Mitchell PD, *et al.* (2010) Area wide suppression of European corn borer with Bt maize reaps savings to non-Bt maize growers. *Science* 330: 222–225.

Janick J and Goldman IL (2003) Horticulture, horticultural science, and 100 years of ASHS. *HortScience* 38: 487–504.

Johal GS, Balint-Kurti P, and Weil C (2008) Mining and harnessing natural variation: A little MAGIC. *Crop Science* 48: 2066–2073.

Kloppenburg JR (2004) *First the Seed: The Political Economy of Plant Biotechnology*. Madison, WI: University of Wisconsin Press.

Marvier M, McCreedy C, Regetz J, and Kareiva P (2007) A meta-analysis of effects of Bt cotton and maize on nontarget invertebrates. *Science* 316: 1475–1477.

National Research Council (1972) *Genetic Vulnerability of Crops*. pp. 5–22, 286–304. Washington, DC: National Academy of Sciences.

Open Plant Breeding Foundation. http://opbf.org/

Pollan M (2006) *The Omnivore's Dilemma*. New York: Penguin Press.

Rasmusson DC and Phillips RL (1997) Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. *Crop Science* 37(2): 303–310.

Ronald PC (2011) Plant genetics, sustainable agriculture and global food security. *Genetics* 188: 11–20.

Tanksley SD and McCouch SR (1997) Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* 277: 1063–1066.

This P, Lacombe T, and Thomas MR (2006) Historical origins and genetic diversity of wine grapes. *Trends in Genetics* 22: 511–519.

Trewavas AJ (2001) The population/biodiversity paradox. Agricultural efficiency to save wilderness. *Plant Physiology* 125: 174–179.

Troyer AF (1996) Illini corn breeders: Their quest for quality and quantity. *American Seed Trade Association Hybrid Corn-Sorghum Research Conference* 50: 56–67.

vom Brocke K, Christinck A, Weltzien E, Presterl T, and Geiger HH (2003) Farmers' seed systems and management practices determine pearl millet genetic diversity patterns in semiarid regions of India. *Crop Science* 43: 1680–1689.

Zamir D (2008) Plant breeders go back to nature. *Nature Genetics* 40: 269–270.

Biodiversity of Harmful Marine Algae

Linda K Medlin, UPMC, University of Paris, Paris, France, and Observatoire Océanologique, Banyul-sur-Mer, France
Allan D Cembella, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

Published by Elsevier Inc.

Glossary

Allelochemical A bioactive secondary metabolite with no known major function in primary metabolism but with the capacity to affect the growth, survival, or behavior of target organisms. Many allelochemicals act as chemical defense agents, although they may also mediate beneficial interactions with the targets.

Bootstrap support A method of attempting to estimate confidence levels of inferred relationships in phylogenetic trees. The bootstrap analysis resamples the original data matrix with replacement of the characters and reconstructs the tree. The bootstrap support for any internal branch is the number of times it was recovered during the bootstrapping procedure.

Clade A group of species that shared a common ancestor (i.e., monophyletic).

Heterokonts A monophyletic group (also known as *stramenopiles*) that includes all cells with heterodynamic flagella: the mature flagellum bearing small tripartite tubular hairs termed *mastigonemes* and the second being smooth.

Maximum likelihood (ML) A method of inferring phylogenetic relationships using a prespecified (often user-specified) model of sequence evolution. Given a tree (a particular topology, with branch lengths), the ML process asks the question, "What is the likelihood that this tree would have given rise to the observed data matrix, given the prespecified model of sequence evolution?"

Microsatellite Short tandem repeats (e.g., ACn, where $n > 8$) of nucleotide sequences; the tandem units can be dinucleotides, trinucleotides or tetranucleotides. Each individual should have a unique number of repeat units for each microsatellite. When two individuals share the same number of repeat units, then they are more closely related than those whose repeat units differ in number.

Molecular clock A hypothesis that mutation rates and substitution rates do not vary among lineages in a tree, thus the timing of divergences can be dated from known times of divergence in the fossil record. However, it has been shown that rates of evolution along a lineage can vary. Two different kinds of clocks can be used: a fixed or a relaxed clock, depending on whether the rate of evolution is allowed to vary.

Outgroup A species or other taxonomic group chosen to establish the root of any phylogeny

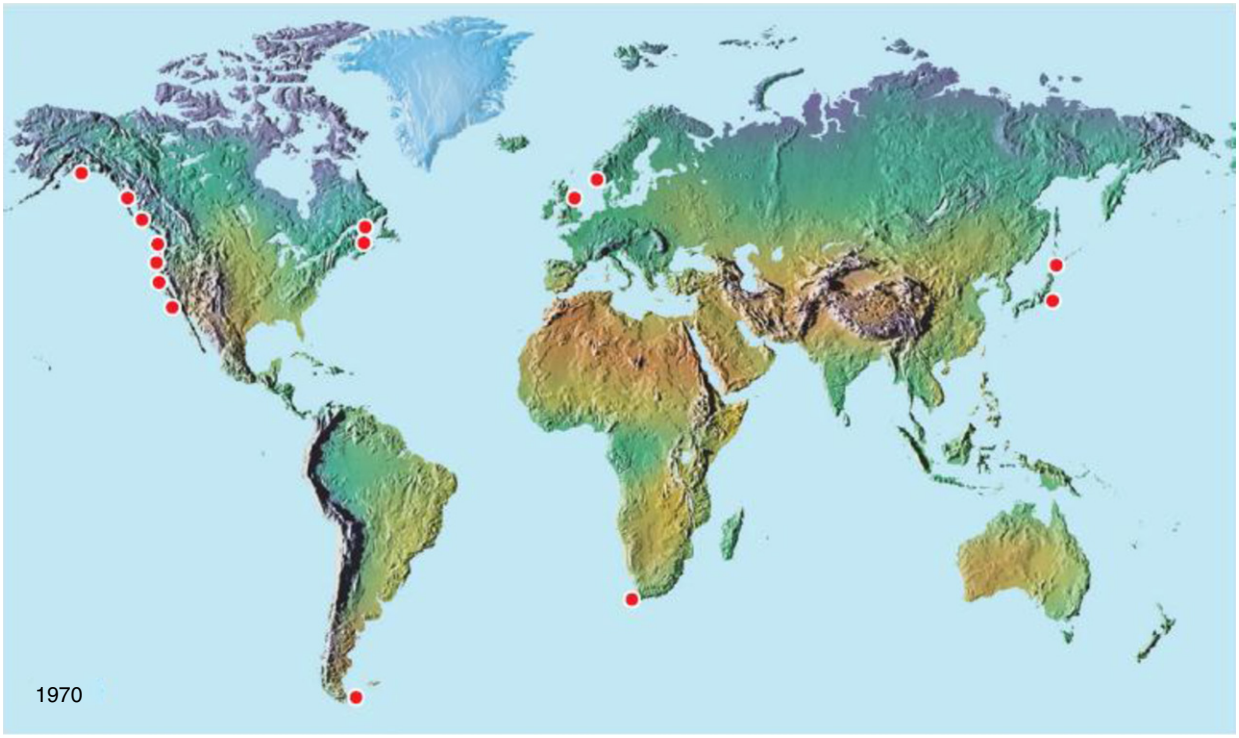
Ribotyping Fingerprinting of genomic DNA restriction fragments that contain all or part of the ribosomal RNA gene to identify and distinguish among strains or species, which are thereby defined as *ribotypes*.

Distribution and Biogeography

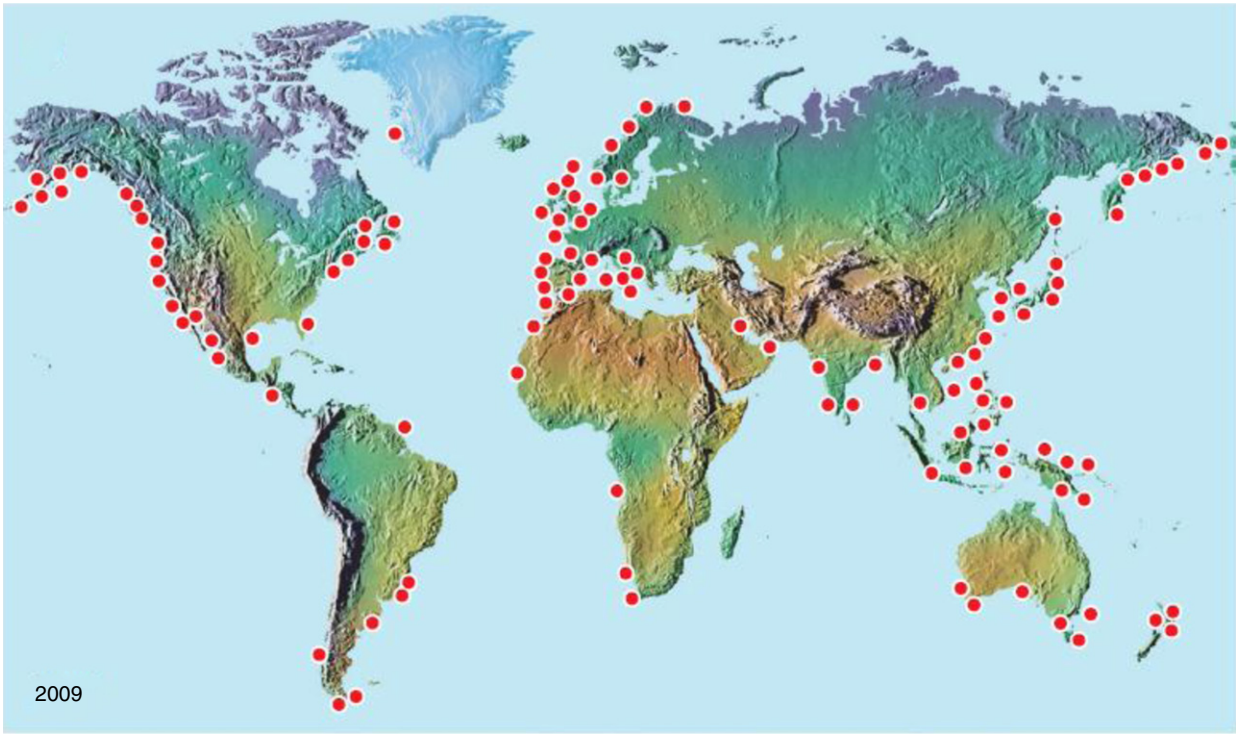
Harmful marine algal taxa are globally distributed from tropical to polar latitudes and occupy ecological niches ranging from brackish water, such as the Baltic Sea, to oceanic environments. Nevertheless, high-biomass harmful blooms tend to be predominant in coastal and shelf seas and high-magnitude effects are most often expressed in these areas. Whereas an apparent global increase in the frequency, magnitude, and biogeographical range in harmful algal blooms (HABs) over the last several decades has often been cited and appears to be legitimate in a general sense (Hallegraeff, 1993) (**Figure 1**), the effects on local and regional biodiversity remain unresolved. Mechanisms invoked for the "global spreading hypothesis" (Smayda, 1990) include transport of exogenous species via ship ballast-water transport or transfer of aquaculture stock, changes in ocean currents, and human intervention in coastal zones (e.g., increased eutrophication, coastal engineering, dredging, and sediment disruption). The socioeconomic impact of HABs has clearly increased in recent years, perhaps partly caused by accelerated and modified use of coastal resources, but it is premature to conclude that this accurately represents verifiable changes in biodiversity. Enhanced scientific awareness and refined monitoring programs for HAB taxa and associated toxins have also contributed to the apparent global increase scenario.

Whether or not discovery of previously undetected or undescribed HAB taxa in a new environment represents a true shift in biodiversity or merely reflects inadequate observations of cryptic species is often difficult to determine. There are, however, well-documented cases of apparent range extension of HAB species and associated syndromes, often rather suddenly, into new environments. For example, until recently, the dinoflagellate *Karenia brevis* responsible for neurotoxic shellfish poisoning (NSP) was apparently confined to the Gulf of Mexico and the east coast of Florida. However, in 1987, an unusual northward flow pattern of the Gulf Stream apparently distributed a major bloom of *K. brevis* into North Carolina waters, and since then such blooms are recurrent along the coast of the Carolinas (Tester *et al.*, 1991). Extreme weather events, such as hurricanes, are known to expand the existing distribution of cyst-producing toxic dinoflagellates. As a classic example, a toxigenic population of the dinoflagellate *Alexandrium tamarense* was apparently introduced from coastal Nova Scotia to New England, where it established resident populations as far south as Cape Cod in the aftermath of a hurricane in 1972 (Anderson, 1994).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00404-4



(a)



(b)

Figure 1 Distribution of events up to (a) 1970 and (b) 2009, respectively, where paralytic shellfish poisoning toxins were detected in shellfish or fish, thereby providing strong circumstantial evidence for a range expansion in the distribution of the causative toxigenic dinoflagellate blooms. Reproduced with permission from US National Office for Harmful Algal Blooms.

Phylogenetic and Taxonomic Relationships

The HAB phenomenon is characterized based on a societal definition ("harmfulness") and thus not surprisingly the associated taxa do not fall neatly into tight taxonomic or monophyletic groups. Marine HAB events have been attributed to members of various eukaryotic algal groups, including dinoflagellates (Dinophyta), prymnesiophytes or haptophytes (Haptophyta), diatoms (Bacillariophyta), raphidophytes (Raphidophyta), and pelagophytes (Pelagophyta); certain cyanobacteria from marine and brackish habitats are also included because of their consequences to ecosystems and human health. In the following sections, the authors discuss three major marine groups of harmful algae.

Dinoflagellates

Origin of the Group

The dinoflagellates are monophyletic, with *Oxyrrhis* (predinoflagellate) lying outside the core dinoflagellates, and sister to the perkinsid flagellates that cause malaria (Apicomplexa) with high bootstrap support to form a clade that is sister to the ciliates (e.g., *Paramecium*), again with high bootstrap support (Leander and Keeling, 2004). Collectively, these groups comprise the superphylum or kingdom Alveolata (Adl *et al.*, 2005), defined morphologically by the series of cortical membranes or membrane bags beneath the plasmalemma and tubular cristae in the mitochondria. A less robust sister relationship (<50%) is recovered in most phylogenetic trees with the stramenopiles or heterokont organisms (Leander and Keeling, 2004), which together with the cryptomonads and haptophytes form the chromalveolates.

The evolution and morphological adaptation of the dinoflagellates, leading to a final radiation of the free-living autotrophic forms, has been reviewed (Saunders *et al.*, 1997). In this scheme, the Noctilucales, which undergo an unusual gametic meiosis and lack chloroplasts, are basal to the core dinoflagellates. Myzotosis or peduncle feeding, whereby a tube from a heterotrophic dinoflagellate attaches to a prey cell and the content of the prey cell is sucked into the predator, is proposed to have evolved early in the dinoflagellates. Many dinoflagellates possess a membrane-bound cell wall, known as a *theca*, that is divided into plates of cellulose (*armor*) within membrane vesicles. The tabulation of these plates (number, structure, and orientation) provides a distinctive morphological feature for taxonomic and phylogenetic analysis. Subsequent to the radiation of the core dinoflagellates, there are three alternative evolutionary hypotheses to explain the separation of thecate and nonthecate genera (Bujak and Williams, 1981), *viz.*, the plate increase, the plate reduction, and the plate fragmentation models, with the first and third models placing the naked or nonthecate Gymnodiniales in a derived position. The peridinioid (thecate) taxa are in a different evolutionary position in each of the three models.

New phylogenetic trees derived from molecular data must be evaluated in light of these evolutionary schemes based on morphology, but there are few such trees covering the entire range of dinoflagellates, with most molecular trees focusing only on a group of species or closely related genera. Saldarriaga *et al.* (2004) have produced one of the earliest trees that sampled across most major core dinoflagellates. Using ciliates as an outgroup, the parasitic and atypical taxa diverge in exactly the sequence predicted by their morphological features until the divergence of the core dinoflagellates. After these divergences, clades of nonthecate gymnodinioid dinoflagellates diverge several times, alternating with clades of thecate peridinioid dinoflagellates. Each clade consists of monophyletic well-supported genera, with the exception of *Gymnodinium*, which is paraphyletic. However, relationships between the clades are not supported. There is a final divergence of the Gonyaulacales, but the Prorocentrales are paraphyletic; thus no real support is provided for any of the morphological models. Adding more taxa to the tree has not really improved the situation but has yielded a few new surprises. **Figure 2** shows a tree of dinoflagellate phylogeny from a maximum likelihood analysis of the 18S rRNA gene. Divergences from the ciliates to the core dinoflagellates follow a similar pattern as seen in the other trees. The core dinoflagellates diverge simultaneously into four major clades. The first major clade contains a mixture of gymnodinioid and peridinioid taxa with nonthecate *Amphidinium* spp. often occurring as a basal divergence in a peridinioid clade, which suggests that a naked form gave rise many times to a thecate genus. Dinophysiales are a basal divergence, and the Prorocentrales are split into benthic and planktonic clades not too distantly related. Hoppenrath and Leander (2010) used the heat shock protein (HSP) 90 gene and Zhang *et al.* (2007) used three genes to find that the Prorocentrales were a monophyletic lineage. The second major group in the large rRNA tree (**Figure 2**) is a smaller gymnodinioid clade. The third major clade contains a mixture of gymnodinioid and peridinioid taxa. However, Noctilucales are embedded in this clade, a position also recovered in the HSP gene tree by Hoppenrath and Leander (2010). Gonyaulacales are a final divergence. The fourth major clade consists primarily of naked dinoflagellates with the Suessiales (found as coral endosymbionts) as a final divergence. With a three-gene concatenated dataset, Zhang *et al.* (2007) found *Amphidinium* at the base of the core dinoflagellate lineage. Their major well-supported divergence was between endosymbiotic taxa and free-living taxa, with the Gonyaulacales being monophyletic with one exception.

Origin of the Toxic Species

The toxigenic species are not a monophyletic group and although the various genera containing toxic members are more or less monophyletic, they are distributed among the four different major clades shown in **Figure 2**. The dinoflagellate toxins are structurally diverse, ranging from linear and macrocyclic polyketides to tetrahydropurine alkaloids, and are derived via divergent pathways. Thus, the various major types of toxins likely evolved more than once among the dinoflagellates, and perhaps even within clades. Alternatively, if toxins have evolved only once, then there have been many multiple losses of toxigenic capability. Clade 1 includes the genera *Dinophysis* and *Prorocentrum* associated with diarrhetic shellfish poisoning, the karlotoxin-producing genus *Karlodinium* sister to *Azadindium*, which contains an azaspiracid toxin–producing



Figure 2 RaxML phylogenetic tree from 1490 small subunit rRNA sequences from dinoflagellates and closest sister groups. The core dinoflagellates have a simultaneous divergence of four major clades. All major clades are collapsed, showing the dominant cell morphology in the clade (either peridinioid (thecate) or gymnodinioid (naked)). Toxic genera are mentioned where they occur throughout the tree.

species and the ichthyotoxic *Pfeisteria*; each of these genera is in a different monophyletic clade within clade 1 and have nontoxic species as their closest sister group. The NSP toxin producers of the genus *Karenia* are found in clade 2. In clade 3, the monophyletic genus *Alexandrium*, containing approximately a dozen species that produce paralytic shellfish poisoning (PSP) toxins, is sister to the ciguatera toxin–producing genus *Gambierdiscus*, which in turn is sister to the thecate genus *Pyrodinium*, including *P. bahamense* var. *compressum* that also synthesizes PSP toxins. *Gymnodinium catenatum*, a nonthecate PSP toxin producer, is the only toxic species found in clade 4.

Haptophytes

Origin of the Group

Haptophytes are unicellular algae that are important members of the marine phytoplankton involved in many important biochemical cycles. Cells possess two smooth flagella and another organelle, called a *haptonema*, inserted between the flagella. The cells are covered by organic scales, which are calcified in one group, the coccolithophorids. The true sister group of the haptophytes is unknown but is weakly allied with the Cryptophyta (see review in Lane and Durnford, 2010). Haptophytes comprise two clades that correspond to the two classes. In the class Coccolithophyceae are several clades that correspond to order level taxonomy in the group.

Origin of the Toxic Species

Among the haptophytes, only a few planktonic species are responsible for harmful and nuisance events through the production of high biomass, which may be accompanied by apparent toxicity effects on marine fauna – for example, mass fish mortalities. Members of the genera *Chrysochromulina* and *Prymnesium* are associated with ichthyotoxicity in coastal marine, brackish, and even freshwater systems, whereas *Phaeocystis* blooms can produce copious foam that forms unpleasant aggregations on beaches but are not known to be toxigenic. In **Figure 3**, only the order Prymnesiales is shown because all toxic haptophyte species fall into one clade, namely, the family Prymnesiaceae within this order. This is in contrast to the dinoflagellates for which toxigenic taxa are distributed throughout the phylogenetic tree.

The genus *Chrysochromulina* is paraphyletic and falls into two clades, one of which contains the toxic species as a separate and well-supported subclade. The high bootstrap and morphological support for major clades and subclades in the Prymnesiales justified the recent substantial taxonomic revision (Edvardsen *et al.*, 2011), whereby all species in one of

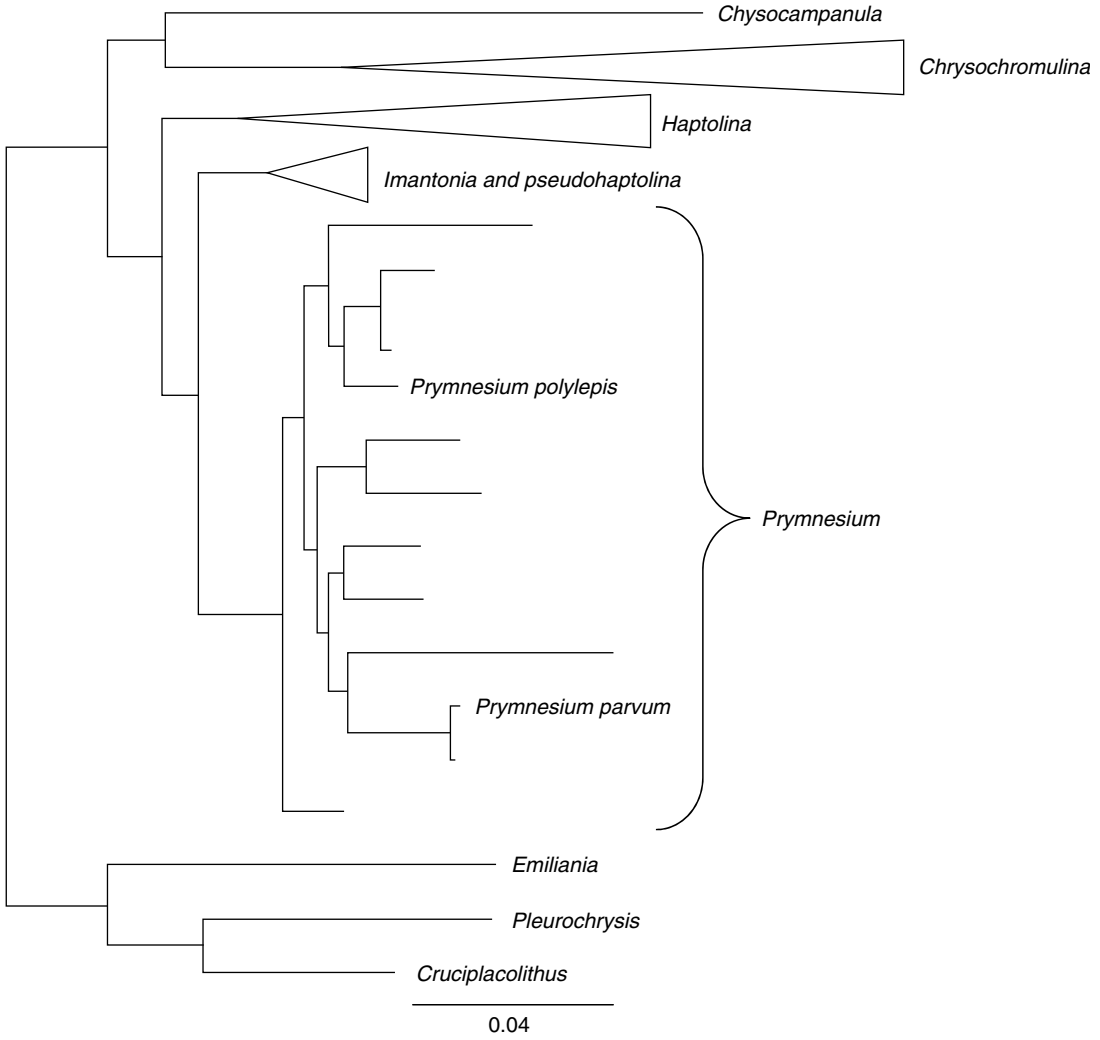


Figure 3 Consensus Bayesian tree based on concatenated nuclear 18S and partial 28S and plastid 16S ribosomal encoding DNA sequences of members of the Prymnesiales. *Cruciplacolithus neohelis*, *Emiliania huxleyi*, and *Pleurochrysis carterae* were used as outgroups. Redrawn from Edvardsen B, Eikrem W, Throndsen J, Saez A, Probert I, and Medlin LK (2011) Ribosomal DNA phylogenies and a morphological revision set the basis for a new taxonomy of the Prymnesiales (Haptophyta). *European Journal of Phycology* 46: 202–228.

the subclades have been transferred into *Prymnesium*. This genus includes the toxic species *Chrysochromulina polylepis* (now as *Prymnesium polylepis*); with *Prymnesium parvum*, these are the two main fish-killing species in the haptophytes.

Heterokonts

Origin of the Group

The Heterokonts are another major eukaryotic radiation, often linked to the alveolate lineage with moderate bootstrap support (Leander and Keeling, 2004). The basal lineages are heterotrophic and photosynthetic lineages are the last major divergence in the group (Cavalier-Smith and Chao, 2006). All toxigenic and otherwise harmful groups of heterokont algae occur within the photosynthetic lineage.

Origin of the Toxic Species

Within the pigmented heterokonts, two major classes or phyla include many toxic species. Among the class Raphidophyceae, some strains of the genera *Chattonella* and *Heterosigma* at high cell concentrations can kill fish, but the exact mode of action and the nature of the "toxins" remains unresolved (Rensel and Whyte, 2003). Other raphidophytes, such as members of the freshwater genus *Gonyostomum*, may secrete copious amounts of mucilage that are detrimental to fish gills but show no evidence for a defined toxin.

Among the Bacillariophyta, harmful effects may be produced by alternative mechanisms, primarily by toxin production or by physical-mechanical damage. In the latter case, some centric diatoms, such as *Chaetoceros* spp., have a life form (cells in a chain with long spiny protuberances called *setae*) that can clog fish gills and thereby kill fish, but they do not produce any toxins. Toxin producers are only found so far among the marine pennate diatoms (**Figure 4**), largely or exclusively belonging to the genus *Pseudo-nitzschia*. This genus contains more than a dozen species known to produce the neurotoxin domoic acid, the causative agent of amnesiac shellfish poisoning. Most of the older-named species were originally placed in the genus *Nitzschia* because the genus *Pseudo-nitzschia* was originally its own genus before it was reduced to a section of the genus *Nitzschia* and then elevated back to genus status by Hasle (1993). From a diversity perspective, the toxigenic and taxonomic status remains confusing because extensive molecular analyses continue to recover cryptic species. As a result, many toxin-producing *Pseudo-nitzschia* have been transferred or promoted from *forma* or varieties of *Pseudo-nitzschia* to species level, and new species have been described, based on morphological characteristics normally at the limit of resolution of the light microscope. Electron microscopy must be used to confirm their identity in field samples as well as to describe them as new taxa. Thus, name changes among the taxa abound. For example, the first pennate diatom known to produce domoic acid was originally reported as *Navia pungens* forma *multiseries*, then as *Pseudo-nitzschia pungens* f. *multiseries*, and finally as *P. multiseries* (Bates *et al.*, 1998). Early reports of domoic acid in marine species of *Nitzschia sensu stricto* and the brackish water species, *Amphora coffaeformis*, require confirmation. Further complications arise from the fact that toxigenicity may be lost in cultures without sexual production and domoic acid production within strains is highly inducible or repressible. Finally, even among clearly defined strains belonging to the same morphospecies or genetic species of *Pseudo-nitzschia*, the capacity for domoic acid production may be highly inconsistent and is often related to the stage in the growth cycle with many strains only becoming toxic as they enter stationary growth.

Population Genetics and Bloom Dynamics

Protein variants known as *isozymes*, with the same enzymatic functional role but differing slightly in 1°, 2°, 3° or 4° structure, were the first molecular markers for species- and population-level studies on marine species, including those of HAB taxa. Small differences in their molecular size or isoelectric point enable separation of isozymes by electrophoresis. Isozymes were the markers of choice for early investigations because they were quick and easy to resolve and detect, and they reflect primary gene products. But the requirement that isozymes must still be functional in the biochemical pathways strongly limits the number of possible mutations and, therefore, the number of alleles and the heterozygosity of this marker. Another disadvantage of this marker is that protein content and enzyme activity and thus the detection of isozymes is strongly influenced by the environment and physiological status of the cells.

The goal of most early molecular studies on microalgae based on isozyme analysis was to resolve species-level issues among species with conflicting or little morphological resolution rather than to study genetic structure within bloom populations. Thus, the early isozyme results were applied to the recognition of cryptic species or the recognition of previously discounted morphological markers for separation of members of a species complex. In the first inter- and intraspecific study of members of the *A. tamarense* and *catenella* species complex, isozyme analyses showed a high degree of enzymatic heterogeneity among isolates from the west coasts of the USA and Canada but indicated that isolates from the same locality were most closely related (Cembella and Taylor, 1986). In contrast, a relative lack of enzymatic heterogeneity was revealed by a similar analysis of *Alexandrium* populations from the east coast of the USA (Hayhome *et al.*, 1989). Isozyme data suggested a common origin for the East Coast populations and supported the dispersal hypothesis along the East Coast of the USA from Canada down to Massachusetts, as related to hydrographic events dissipating a massive red tide that occurred in 1972. Isolates of the dinoflagellate *Gambierdiscus toxicus* from similar geographical regions within the tropics were shown not to be closely related by isozyme comparison, which suggested a multiclonal origin for these populations (Chinain *et al.*, 1997). Among HAB taxa, isozyme studies are largely confined to the dinoflagellates, as there are no published studies for the toxic haptophytes and only very preliminary investigations have been conducted on toxic *Pseudo-nitzschia* spp. (Skov *et al.*, 1997).

As a consequence of the limitations of isozyme analysis to reflect genomic differences at high resolution, alternative marker types were later developed that directly addressed the nucleic acid level and hence are relatively refractory to

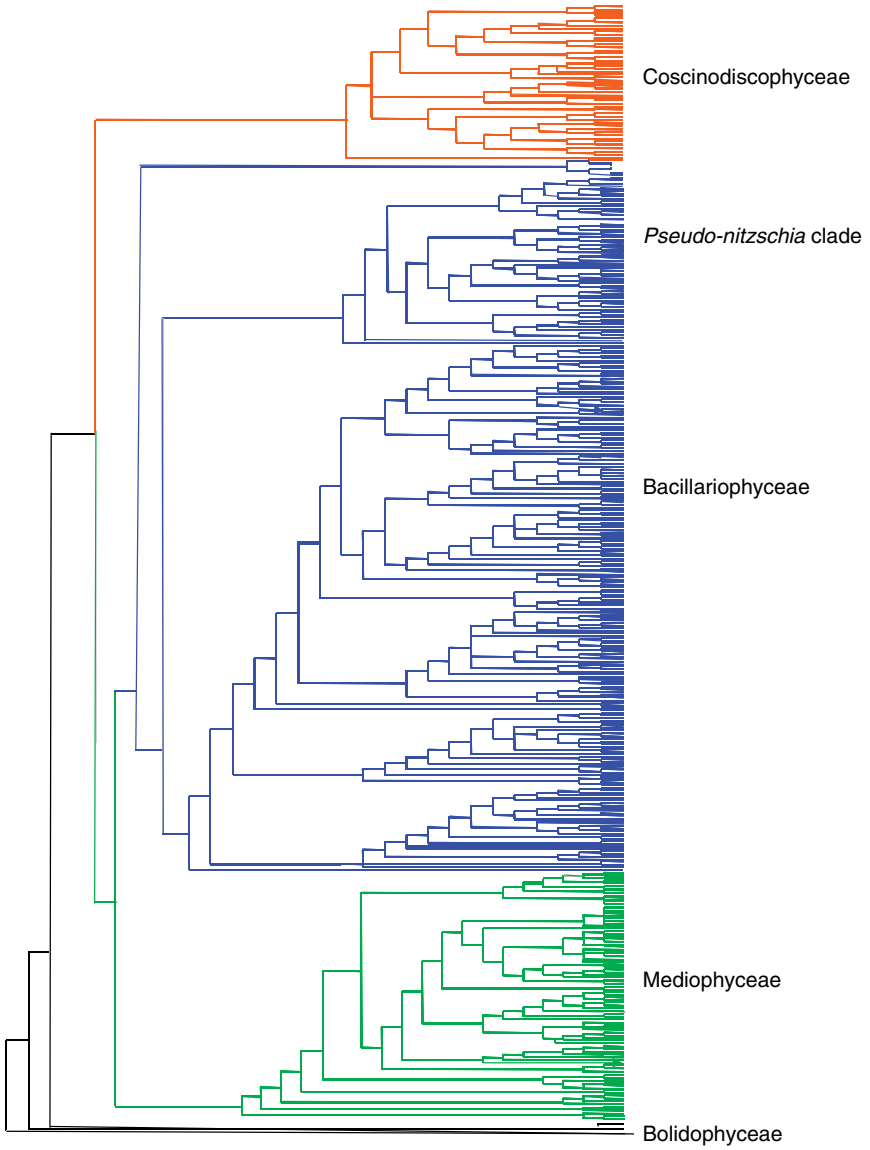


Figure 4 Baysian analysis of 920 diatom taxa rooted with the Bolidomonads. Three monophyletic classes occur, each shown in a different color. The toxic diatom genus *Pseudo-nitzschia* occurs as one of the most derived or advanced pennate diatom genera. Redrawn from Bowler C, Allen AE, Badger JH, *et al.* (2008) The *Phaeodactylum* genome reveals the dynamic nature and multilineage evolutionary history of diatom genomes. *Nature* 456: 239–244.

environmental variation and shifts in physiological status. Diversity below the species level is now most robustly measured by DNA- or RNA-fingerprinting methods of which microsatellites are the most computationally intensive. Recently, there has been a rapid increase in studies using DNA microsatellite markers to estimate gene flow and resolve dispersal mechanisms in natural populations as evidenced by the development of microsatellite markers for many species, with particular focus on harmful species (see references in **Table 1**). The primary limiting factor to developing microsatellites and performing ecological analyses on large spatial and temporal scales is the need to make single-cell isolations from planktonic populations and to maintain them in culture for analysis. Amplification of microsatellites from single dinoflagellate cells (Henrichs *et al.*, 2008) represents a way to overcome culture and sampling bias for planktonic organisms.

Cryptic speciation in several harmful algal species has been studied in more detail using sequence analysis of rapidly evolving genomic regions, such as the internal transcribed spacer (ITS) and the hypervariable D1/D2 region of the large subunit (LSU) rRNA gene. Within the *Alexandrium tamarense–fundyense–catenella* species complex, perhaps the most intensively studied HAB group, isolates were shown to be related by geographic origin rather than by morphological affinities (Scholin *et al.*, 1995), with the rRNA gene data essentially confirming the early isozyme analyses. *Alexandrium* isolates from within the species complex will interbreed more successfully if they have similar isozyme patterns from two

Table 1 Summary of microsatellite studies on harmful dinoflagellates

Species	*Source*
Alexandrium tamarense/fundyense/catenella NA	Nagai, *et al.* (2004, 2007a), Alpermann, *et al.* (2009)
Alexandrium tamarense/fundyense/catenella TA	Nagai, *et al.* (2006b), Nishitani, *et al.* (2007)
Alexandrium minutum	Nagai, *et al.* (2006a)
Cochlodinium polykrikoides	Nagai, *et al.* (2009)
Heterocapsa circularisquama	Nagai, *et al.* (2007b)
Heterosigma akashiwo	Nagai, *et al.* (2006c)
Karenia brevis	Renshaw, *et al.* (2006)
Lingulodinium polyedrum	Frommlet and Iglesias-Rodríguez (2008)
Akashiwo sanguinea	Cho, *et al.* (2009)

All species listed are known to produce an identifiable toxin, except as indicated in bold.

different locations than will isolates from the same locations but with different isozyme patterns (Sako *et al.*, 1990). Currently, the geographic clade names have been replaced with numbers because their present-day known distribution exceeds the original areas (Lilly *et al.*, 2007). Thus, the North American clade is now termed Group I with primarily a Pacific distribution but also found in northern Europe, the Mediterranean clade of Group II is still endemic to the Mediterranean; the Western European clade equals Group III with one Japanese relative, the temperate Asian clade equals Group IV with relict populations in the Mediterranean, and the tropical Asian clade equals Group V known only from Australia.

The worldwide biogeographic dispersal of ancestral populations of the *A. tamarense* species complex from the Pacific into the Atlantic has also been hypothesized from rRNA gene data. Using a molecular clock, John *et al.* (2003) proposed a scenario for the biogeographic history of this complex from a single global ancestor. The later divergences of the nontoxic Western European clade from the toxic temperate Asian (TA) could be dated to the closing of the Tethys Sea, which likely explains the presence of relict cells of this clade in the Mediterranean. The separation of the nontoxic Mediterranean clade from the toxic North American clade was dated to the rising of the Isthmus of Panama. With time, the Mediterranean clade went extinct in the western tropical Atlantic waters, but it thrives today in the Mediterranean, where it is now endemic.

Microsatellites have clarified the origin of the TA clade of the *Alexandrium tamarense–fundyense–catenella* species complex. Lilly *et al.* (2002), comparing Mediterranean and Asian strains with LSU rRNA sequence data, concluded that the presence of TA strains in the Mediterranean was a clear case of ballast-water introduction. Penna *et al.* (2005), using ITS sequence data, also concluded that it was likely the result of ballast-water introduction. Masseret *et al.* (2009), using microsatellites, showed the Mediterranean populations to be clearly distinct and distant from the Asian strains but could not offer any reasonable explanation for their occurrence in the Mediterranean. However, referring back to the historical biogeography of the species complex by John *et al.* (2003), the closing of the Tethys Sea, which is the vicariant event separating the West European from the TA strains, the presence of the TA genotypes in the Mediterranean is likely a relict population from when the two groups were once joined. Their notable presence in Mediterranean French lagoons is likely caused by a change in environmental conditions that yield dramatic blooms. The first hierarchical study of genetic diversity has been conducted on populations of *A. tamarense* (NA clade) off the Orkney Islands, UK (Alpermann *et al.*, 2009, 2010). On a global scale, the Scottish east coast populations found near the Orkney Islands are more closely related to Japanese Pacific isolates than to populations on the east coast of North America, which suggests that the Orkney populations were introduced to that region by cells transported directly across the Arctic Ocean. This interpretation is in contrast to the hypothesis put forward by Medlin *et al.* (1998) that these populations entered from the Pacific and moved along coastal pathways of the eastern side of North America until they reached the Gulf Stream, which carried them across to the Scottish east coast. Moving to the diversity of local populations, at one site off the Orkney Islands, four populations were discovered that could interbreed, and these populations were proposed to represent different year classes that had hatched from local cyst beds (Alpermann *et al.*, 2010).

Karenia brevis has been shown to be a single population in the Gulf of Mexico (Renshaw *et al.*, 2006 and unpubl.) based on microsatellite analysis, although molecular probes have also supported the recognition that the Florida red tides previously assumed to be monospecific *K. brevis* blooms are often multispecific (Steidinger *et al.*, 2008). In *Alexandrium minutum*, several microsatellites have been found to occur on the same chromosome (Nagai *et al.*, 2006b). *Alexandrium minutum* from the Atlantic Ocean and the Mediterranean Sea can be segregated into four genetic populations that more or less correspond to four regional seas in the Mediterranean and the relatedness of these four populations can be explained by circulation patterns in the Mediterranean (Casabianca *et al.*, 2012).

Among HAB taxa, most microsatellites have been developed for dinoflagellates but also for a few raphidophytes, such as *Chattonella* spp. (Demura *et al.*, 2007) and *Heterosigma akashiwo* (Nagai *et al.*, 2006b). Microsatellites for the raphidophytes were found to be very heterozygous but have not yet been applied for ecological interpretations at the population level.

Almost all dinoflagellates are nominally haploid in the vegetative stage, with exceptions found among the Noctilucales. A number of HAB species, particularly among the thecate dinoflagellates, produce a resistant benthic resting stage (or cyst) within their life history as an overwintering or survival strategy (**Figure 5**). Differentiation of haploid gametes

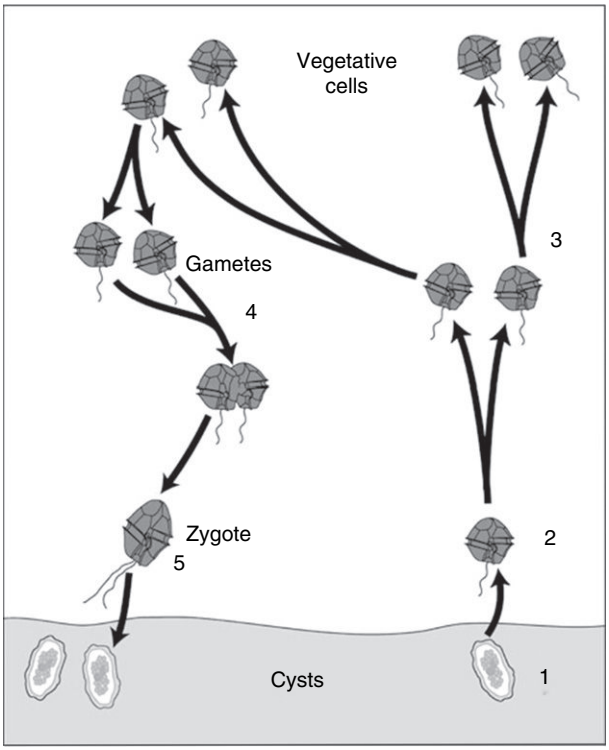


Figure 5 Schematic representation of the life history of *Alexandrium* species, typical of dinoflagellates for which the motile vegetative cells are haploid – that is, they have one set of chromosomes. Excystment from the benthic resting cyst (1) leads to formation of a motile planomeiocyte stage (2) that subsequently divides to form vegetative cells (3). Some vegetative cells become gametes, initiating the sexual phase whereby the gametes fuse (4) to yield a motile planozygote with double the number of chromosomes (5). The planozygote can transform into a resting cyst (hypnozygote) to complete the cycle or for some species can undergo a reduction division to yield a vegetative cell. Modified from Sellner KG, Doucette GJ, and Kirkpatrick GJ (2003) Harmful algal blooms: Causes, impacts and detection. *Journal of Industrial Microbiology and Biotechnology* 30: 383–406, with permission from Springer.

followed by fusion to form diploid planozygotes leads ultimately to nonmotile sexual resting cysts (or hypnozygotes) and thereby provides the opportunity for both genetic recombination and the maintenance of functional diversity within and among populations. The timing of excystment and survival success of hatched cysts as a contribution to the vegetative population or "bloom" is therefore a key determinant of the population genetic structure. This was demonstrated in a conceptual population model for the dinoflagellate genus *Alexandrium* (Alpermann *et al.*, 2009) based on the analysis of multiple isolates from a geographical population using microsatellites and amplified-fragment length polymorphism (AFLP). Clonal selection and shifts in genotype frequencies caused by variations in selective constraints of the environmental regimes is the most likely explanation for observed population genetic substructures revealed by these genetic markers. The interplay between genetic structure and population diversity, as modulated by mating compatibility within contemporaneous populations and environmental factors affecting survival, is also supported by observation of two genetically distinct subpopulations of *A. fundyense* (Group I) in the northeastern USA, comprising either early bloom or late-bloom genotypes (Erdner *et al.*, 2011).

Apparent biogeographical distinctiveness among clades may arise via reproductive barriers based on limited sexual compatibility. Brosnahan *et al.* (2010) mated strains of Group I and Group III ribotype of the dinoflagellate *A. tamarense*, which yielded true sexual resting cysts, but the subsequently germinated cells did not survive. Such reproductive barriers are indicative that these ribotypes constitute different biological species and also suggests that such barriers may limit successful invasions of foreign ribotypes and range extension via lethal hybridizations. This interpretation is consistent with the mating compatibility model developed earlier for *A. tamarense* (as *Alexandrium excavatum*) by Destombe and Cembella (1990) based on crossing multiple clonal strains and monitoring the presence of fusing gametes, cyst formation, and subsequent germination success. This model postulates a complex spectrum of mating compatibility and affinities rather than two defined parental mating types, even within a geographical population.

Chemical Ecology in Species Interactions Affecting Diversity

As members of pelagic or benthic communities, HAB taxa are subject to species interactions that may alternatively promote or suppress the population growth rate and biomass yield. Population growth may be limited not only by abiotic factors (light, temperature, nutrients, advection) but also by species interactions, such as grazing, parasitism, viral and bacterial attack, and resource competition. Although it is often assumed that HAB taxa achieve high biomass and monospecific dominance via high intrinsic growth rates, this is not always the case. For example, many dinoflagellates, including most harmful taxa, are comparatively slow growing and poor nutrient competitors with respect to contemporaneous diatoms (Smayda, 1997).

In addition to the selective advantage of swimming behavior and vertical migration capabilities of marine flagellates, many HAB flagellates possess the ability to produce toxins or other allelochemicals that may affect species interactions (Cembella, 2003). Among the eukaryotic flagellates, these toxins include linear and polycyclic polyethers and tetrahydropurines, such as saxitoxin and analogues. Within the diatom genus *Pseudo-nitzschia*, some strains of a few species also produce the neurotoxic secondary amino acid known as domoic acid and its derivatives.

The widespread although not ubiquitous distribution of potent toxins and other allelochemicals among phylogenetically diverse groups of HAB taxa suggests that these secondary metabolites have a defined functional role and provide a selective advantage, perhaps related to competition, predator–prey interactions, or other forms of chemical communication (Legrand *et al.*, 2003). The high potency of these secondary metabolites against mammals and cultured cells led to the assumption that the compounds are

components of a "watery arms race" (Smetacek, 2001) against competitors and predators in marine ecosystems. This molecular target idea is further supported by evidence that the toxic profiles are typically stable and highly polymorphic within a clone but subject to high variation within and among species and populations (Cembella, 2003). The chemical defense hypothesis has been intensively investigated over the past three decades in studies with a wide variety of toxigenic taxa producing various toxin spectra, and in combinations with putative predators and competitors. Indeed, in a review of interactions between toxic phytoplankton and metazoan grazers, Turner *et al.* (1998) reported that in some cases exposure of grazers to toxigenic prey caused predator avoidance behavior, incapacitation, or mortality, as would be predicted from the chemical defense hypothesis. In other cases, the grazers appeared not to be negatively affected by the presence of toxins in the prey, but the responses were always highly species- and strain-specific. With respect to microzooplankton grazer interactions, such as dinoflagellates, ciliates, and other protists against toxic HAB taxa, the evidence does not typically support a chemical defense function for the known toxins (Ianora *et al.*, 2011).

This lack of general confirmation of expected toxicity responses should not be interpreted to imply that the toxins play no role in species interactions. In fact, recent studies provide evidence of co-evolutionary mechanisms in chemically mediated species interactions among predators and competitors that could lead to shifts in diversity and dominant genotypes in natural bloom populations. Waterborne cues of predatory copepods can induce a shift up in cell toxin content and changes in gene expression profiles in potential prey, such as *Alexandrium* spp. (Wohlrab *et al.*, 2010; Yang *et al.*, 2011). This functional genomic approach indicated that regulation of serine–threonine kinase signaling pathways has a major influence in directing the copepod cues into different intracellular cascades and networks in toxic prey dinoflagellates (Wohlrab *et al.*, 2010), and may serve as a survival mechanism for dinoflagellates with the capacity for toxin production.

In addition to the known toxins of high potency against mammals, many marine HAB flagellates can synthesize and release allelochemical substances (often poorly chemically characterized) that are capable of immobilizing or lysing cells of potential predators and competitors (Ma *et al.*, 2011). Although the mode of action is unknown or poorly characterized, these allelochemicals may act as "toxins" of high adaptive significance in an ecological sense because of their high lytic activity in natural populations.

The taxonomic and phylogenetic distribution of the known mammalian toxins and apparently unrelated allelochemicals is particularly intriguing from a diversity perspective. For example, the saxitoxin analogues associated with PSP are produced not only by certain species and strains of the related gonyaulacoid dinoflagellates *Alexandrium* spp. and *P. bahamense* but also by the distantly related gymnodinioid dinoflagellate *G. catenatum*. The same toxin group then reappears in members of several cyanobacterial genera, primarily from fresh and brackish water; in the complete operon fully sequenced from *Alexandrium*, however, only three of the eight genes are of cyanobacterial origin (Hackett *et al.*, in press). Clearly, most of the genes are of a eukaryotic and independent origin.

Most of the polyether and tetrahydropurine toxin derivatives produced among HAB dinoflagellates are constitutive metabolites; if present within a strain, the toxin composition is a rather stable phenotypic trait. In contrast, production of the neurotoxin domoic acid by some strains of the diatom *Pseudo-nitzschia* is highly inducible or repressible, depending on nutritional status and other extrinsic factors (Bates, 1998). Among HAB dinoflagellates, the toxin composition is typically represented by a suite of closely related derivatives and is fixed genetically within a clonal strain (Anderson *et al.*, 1990). This led to attempts to apply toxin profiling as a molecular fingerprint or phenotypic marker at both the species and population levels (Cembella, 1998 and references therein). The composition of saxitoxin analogues varies widely within and among *Alexandrium* species, but toxin profiles are generally distinguishable from those of other toxic dinoflagellate genera (*Pyrodinium* and *Gymnodinium*) and cyanobacteria. Within the *A. tamarense* species, complex toxin profiles are generally too diverse and variable for species discrimination, but this chemotaxonomic marker has been successfully applied to differentiate among geographical populations within a species based on cultured isolates (Cembella *et al.*, 1987; Anderson *et al.*, 1994). Comparison of PSP toxin composition of field populations of *A. tamaren*se from different sampling sites in eastern Canada (Cembella and Destombe, 1996) showed that geographical populations from the Bay of Fundy and the St Lawrence estuary were more homogeneous in toxin composition than those from eastern Nova Scotia. Such population-level differences in toxin composition can be interpreted as an indication that the St. Lawrence populations are well mixed genetically and perhaps originate from the same cyst beds, whereas populations from eastern Nova Scotia are more reproductively isolated.

Even if the toxin profiles are genetically determined, the interpretation of biodiversity patterns and variation at the species and population levels based on toxin composition of HAB taxa is nevertheless subject to critical limitations. If based on one or a few cultured isolates, then the toxin phenotypic range is not determined and the profiles represent only those of survivor genotypes. Alternatively, if based on directly sampling of mixed field populations, the toxin profile will reflect the average of the multiple genotypes, reflecting their relative abundance and cellular toxin composition and concentration. Indeed, in an integrated study of clonal variation in toxin composition, allelochemical activity, and genetic markers (microsatellites and AFLP) among a high number of isolates (88 clones) of *A. tamarense* from a geographical population on the east coast of Scotland, Alpermann *et al.* (2010) found high heterogeneity in all phenotypic and genotypic characteristics within the population. Despite hierarchical grouping according to toxin profiles, there was no clear linkage to the molecular markers AFLP and microsatellites or to the allelochemical activity measured as lytic effects. Although similar comparative data are not yet available for other toxigenic HAB taxa, there are preliminary indications from biogeographical studies of toxin polymorphisms (Gribble *et al.*, 2005) that such phenotypic heterogeneity and cryptic diversity within and among populations are common features.

Diversity and Global Change Processes

With respect to general biodiversity of plankton, climate-mediated shifts in ocean regimes are expected to lead to (1) range expansion of warm-water species toward higher latitudes and niche displacement of cold-water species; (2) temporal shifts in the frequency and abundance of particular taxa, including the timing of seasonal succession processes; and (3) alteration of food-web interactions, including grazing and prey selectivity. The extent to which these changes will affect the diversity and biogeographical distribution of HAB taxa (reviewed by Dale *et al.*, 2006; Moore *et al.*, 2008; Hallegraeff, 2010) will be a function of their respective ability to adapt to and colonize such emerging niches provided via changes in nutrient concentration and ratios, degree of stratification, and abiotic factors, such as temperature, pH and salinity regime, etc.

A few recent studies have attempted to establish causal links between exceptional or unprecedented algal blooms and climate shifts and anomalies (e.g., Belgrano *et al.*, 1999; Skjoldal and Dundas, 1991; Cloern *et al.*, 2005; Moore *et al.*, 2008), but conclusive evidence remains lacking. Although scenarios for predicted biodiversity shifts are often directly coupled to climate change and variability (e.g., temperature rise, oceanic carbon flux), these issues must be considered in a broader context to include synergistic or compensating interactions at the local and regional scale. For example, nutrient enrichment of coastal waters may interact with increasing temperatures and CO_2 levels to affect the frequency and abundance of sympatrically distributed HAB species, such as the raphidophyte *Heterosigma akashiwo* and the dinoflagellate *Prorocentrum minimum* (Fu *et al.*, 2008) but specific outcomes are difficult to predict in nature. As an unintended consequence on diversity, eutrophication abatement strategies may lead to a decrease in the frequency and magnitude of ichthyotoxic blooms of rapidly growing high-biomass producing species to be replaced by blooms of highly toxigenic species causing shellfish toxicity even at relatively low cell concentrations. This is apparently the case for the biodiversity regime shift in the Seto Inland Sea of Japan, with decreasing fish kills linked to blooms of raphidophytes over the last several decades but increasing incidences of shellfish toxicity associated with toxic dinoflagellates (Imai *et al.*, 2006). Such complex interactions may also account for the dramatic increase in ciguatera fish poisoning over the past century in the South Pacific, in particular in French Polynesia (Hallegraeff, 2003). The causative factors are unclear but may include reef damage by military and recreational development, increases in typhoon frequency and intensity, and climate-related effects on coral erosion and destruction via photobleaching and acidification. In any case, the primary causative organism *Gambierdiscus toxicus* has now apparently spread from the Great Barrier Reef into southeast Australian sea-grass beds in association with a strengthening of the East Australian Current (Hallegraeff, 2010). Within the Mediterranean region, *Gambierdiscus* populations are also expanding their distributional range (Aligizaki *et al.*, 2008), although this cannot be unequivocally linked to climate-mediated processes.

Conceptual models and hypothetical scenarios of the effects of global change on the diversity and biogeographical distribution of HAB species are based primarily on extrapolations of variation observed over the last several decades (short timescale) or derived from the fossil record. Global change processes affecting HAB diversity are often interpreted within the context of climate variation and oscillations (atmospheric CO_2 levels, sea surface temperature, etc.), but direct anthropogenically mediated modification of coastal ecosystems via coastal engineering, introduction of invasive species and eutrophication are also expected to induce biodiversity shifts on a local and regional scale. Recent decadal scale observations, however, do tend to indicate shifts in abundance and distribution of plankton species, such as those documented from the Hardy Continuous Plankton Recorder (CPR) data from the North Sea and adjacent waters since the 1940s (Reid *et al.*, 2003). Most multidecadal sets are inadequate with respect to revealing frequency and abundance shifts of HAB plankton. The CPR transects were originally designed to focus primarily on zooplankton, and although they have proven useful for studying coccolithophorid distributions in recent years, many of the cryptic HAB taxa, particularly small-celled species, are not well represented or have been previously overlooked in the surveys.

Nevertheless, there are a few documented diversity changes in the range and abundance of particular HAB taxa, related either to interannual variation in oceanic regime or to catastrophic events such as hurricanes. Based on decadal scale analysis, there is circumstantial evidence of a coincidence between the magnitude of blooms of the toxigenic dinoflagellate *P. bahamense* in the southern Pacific and the quasi-periodic climate phenomenon known as El Niño Southern Oscillation (ENSO) (Azanza and Taylor, 2001). Exceptionally high-magnitude PSP episodes caused by blooms of the *A. tamarense/catenella* along the Pacific Northwest coast of the USA were also temporally correlated with ENSO events in seven of nine cases between 1941 and 1984 (Erickson and Nishitani, 1985). A putative link between ENSO events and a massive offshore bloom of the toxigenic diatom *Pseudo-nitzschia* in summer 1998 along the coast of California was associated with higher than normal water temperatures in the eastern Pacific during the summer and fall of 1997 followed by a transition from El Niño to cooler La Niña conditions (Bargu *et al.*, 2010). Although the evidence is not conclusive in any of these cases, climate-mediated changes on oceanic regimes, such as ENSO and North Atlantic oscillation (Belgrano *et al.*, 1999), may prove to be a major factor in determining biodiversity shifts and periodicity in the appearance and abundance of HAB taxa, particularly in coastal zones.

As previously mentioned, rising sea surface temperatures associated with global warming may be expected to favor a range expansion toward higher latitudes, but this will also bring a disproportionate increase in the abundance of marine flagellates, which tend to thrive in warmer stratified water regimes. Rising sea surface temperatures have already been associated with increases in dinoflagellates in the North Atlantic, the North Sea, and the Baltic Sea, as well as to seasonal shifts in the timing of dinoflagellates in the plankton successional cycle (Dale *et al.*, 2006). Global warming may also tend to favor differentially blooms of cyanobacteria in marine and freshwater ecosystems, including those of harmful and toxigenic species (reviewed by Paerl and Huisman, 2008 and 2009). Nevertheless, these apparent shifts in group diversity and

dominance cannot yet be unequivocally linked to climate-mediated differential effects on particular HAB taxa. Dale *et al.* (2006) argue convincingly that because most HAB taxa are coastal or estuarine and broadly eurythermal, modest water temperature increases of a few degrees associated with global warming are unlikely to be the major determinant of their dominance and persistence.

Climate change is expected to lead to disruptive changes in wind velocity and seasonal shifts in ocean stratification parameters, such as upwelling and downwelling of coastal and shelf seawaters. Many HABs, particularly of dinoflagellates, are associated with postupwelling relaxation events in areas, such as the California coast and within the Benguela Current system off southwest Africa (Pitcher and Boyd, 1996). Wind-driven effects on HAB dynamics and diversity are linked not only to stratification but also to upwelling of macronutrients to surface waters (Kudela *et al.*, 2010). These combined wind effects tend to favor the development of blooms of toxigenic chain-forming dinoflagellates, such as those of *G. catenatum* off the Iberian coast (Fraga and Bakun, 1990) and the *catenella* morphotype of the *A. tamarense* species complex along the west coast of North America (Langlois, 2001). In the case of the fish-killing raphidophyte *Chattonella marina* in Hiroshima Bay, Japan, blooms were apparently enhanced subsequent to typhoon-induced runoff of terrestrial nutrients, but such blooms are also likely associated with excystment triggers and the capacity for vertical migration as adaptive strategies (Hallegraeff, 2010).

With few exceptions, HAB events are typically confined to coastal and shelf seas, therefore a rise in sea level and decreased salinity associated with shifts in precipitation patterns and melting of polar ice sheets may increase the potential for blooms of species favored by low salinity and highly stratified water columns. Examples of such species would include the ichthyotoxic haptophytes, such as *Pr. parvum* and *Chrysochromulina* (now *Pr.*) *polylepis* and perhaps coccolithophorids in fjord systems. Ocean acidification linked to increasing atmospheric carbon dioxide has been predicted to have a deleterious effect on calcifying organisms, including coccolithophorids (Riebesell *et al.*, 2000). However, there is wide variation in strain response and also evidence that in complex assemblages as found in nature and not in a culture vessel increased CO_2 will actually increase calcification (Beaufort *et al.*, 2011). In principle, any change in climate could open new niches for other groups – for example, noncalcareous dinoflagellates – but it is difficult to predict whether or not harmful taxa would be particularly advantaged.

Among the thecate dinoflagellates, the benthic cyst stages (following sexual reproduction) of some species are subject to fossilization. The fossil morphotypes can be related to the vegetative stages of extant species, thereby yielding a potential paleo-oceanographic time series of the distributional shifts in the taxon. The toxigenic dinoflagellate *P. bahamense* provides a particularly instructive example of climate-mediated shifts. This species is presently restricted to tropical and subtropical coastal waters of the Atlantic, the Caribbean, and the Indo-West Pacific. Distribution of *P. bahamense* is strongly associated with mangrove-fringed areas (MacLean, 1977; Azanza and Taylor, 2001), therefore it would be expected that occurrence of this species may be reduced in distribution and even disappear from areas subject to mangrove destruction. The fossil cyst record of this species (as *Polysphaeridium zoharyii*), extending to the substantially warmer Eocene period 50 million years ago, indicates a much wider range of distribution, including the Portuguese coast, than indicated in recent records. Although today absent from the nontropical Australian coast, approximately 100,000 years ago this species ranged as far south as Sydney Harbour (McMinn, 1989).

Given the diverse and polyphyletic evolutionary history of HAB species and the consequent lack of an apparent common ecophysiological adaptive strategy, it is difficult to construct general predictive scenarios of "winners and losers" in response to global change. Autecology of individual species and even intrapopulation responses to ecophysiological shifts must be more thoroughly investigated. Group-specific predictions based on known responses to increased temperature, enhanced surface stratification, intensification or weakening of nutrient upwelling, stimulation of photosynthesis by elevated CO_2, changes in land runoff and nutrient availability, and decreased ocean pH are clearly inadequate. Genetic diversity has been investigated in only a handful of species, so it is also impossible to predict whether or not these species have sufficient genetic diversity to withstand major climate shifts and not becoming globally or locally extinct. Certainly climate change will fragment populations and affect gene flow among them.

The Way Forward

The lack of a scientifically based definition of what constitutes a HAB taxon coupled with the multiplicity of strategies and mechanisms whereby the harmful effects are expressed suggest that it is unrealistic to expect a "unified field theory" to encompass all HAB dynamics and evolution in the context of biodiversity. Until recently, ecophysiological and genetic studies on HAB taxa have followed the autecological model in which one or a few cultured isolates are taken to be representative of the response and diversity in natural populations. The fallacy of this approach was shown by Wood and Leatham (1992), who documented in their review significant variation in response among all strains tested within a given microalgal species for a wide variety of responses to environmental stressors. Thus, this acceptance of only a few isolates or strains as representative of a species has seriously hampered our understanding of the relationship between genotypic and phenotypic variation in the microalgae, including HAB taxa, and how this interaction provides the basis for adaptation and ecosystem resilience.

Advanced studies of the genetic composition and gene expression potential of various HAB taxa – the functional ecogenomic approach – can now provide insights into the capacity for survival and adaptation at both the species and population level. High-throughput sequencing (e.g., next-generation 454-pyrosequencing) coupled with bioinformatics platforms offer the capacity for sequencing and annotation of entire genomes of a wide spectrum of eukaryotic microalgae. For HAB species, several such initiatives are underway or essentially completed, for example, for the occasionally toxigenic diatom *Pseudo-nitzschia pungens* (Armbrust, 2009), and more are in the pipeline. Although the enormous genome size

(from 3000 to 215,000 MB) and complexity of the dinoflagellates remains a challenge for whole genome sequencing, *Heterocapsa circularisquama* is presently being sequenced. This dinoflagellate was chosen because the genome of the virus that kills it has been sequenced. Alternatively, limited genomic investigations based on expressed sequence tags have already contributed critical insights into the genome of several HAB species (Yang *et al.*, 2010; Jaeckisch *et al.*, 2011), including the diversity of genes associated putatively with bloom dynamics and toxin biosynthesis. Attempts to model HAB dynamics must take into account the inherent cryptic diversity within and among populations, particularly with respect to toxigenesis and growth regulatory genes. New molecular probe technologies offer the ability to detect and identify genotypic variation within HAB species and populations, in some cases at the individual cell level. When implemented on deployable sensor platforms, this opens the possibility for near-real-time assessment of shifts in biodiversity and gene frequency and expression in natural HAB populations, even over seasonal and annual bloom cycles.

Acknowledgments

This work was supported in part by a grant from the EU FP7 201724 MIDTAL to LKM. The contribution of ADC from Alfred Wegener Institute was provided via the Helmholtz Society Programme on Earth and Environment as part of the Coast Topic (WP2) on Integrating Evolutionary Ecology into Coastal and Shelf Seas Processes.

Appendix

List of Courses

1. Ecology and Evolution of Marine Protists
2. Marine Phytoplankton
3. Marine Chemical Ecology
4. Taxonomy and Phylogeny of Microalgae
5. Evolution of Eukaryotes

See also: Biodiversity, Evolution and. Census of Marine Life. Defining, Measuring, and Partitioning Species Diversity. Diversity, Molecular Level. Diversity, Organism Level. El Niño and Biodiversity. Evolution in Response to Climate Change. Global Species Richness. Landscape Ecology and Population Dynamics. Loss of Biodiversity, Overview. Marine Ecosystems. Microbial Biodiversity. Pelagic Ecosystems. Population Genetics. Species Assemblages, Macroecology, and Global Change

References

Alpermann T, Beszteri B, John U, Tillmann U, and Cembella A (2009) Implications of life-history transitions on the population genetic structure of the toxigenic marine dinoflagellate *Alexandrium tamarense*. *Molecular Ecology* 18: 2122–2133.

Alpermann TJ, Tillmann U, Beszteri B, Cembella AD, and John U (2010) Phenotypic variation and genotypic diversity in a planktonic population of the toxigenic marine dinoflagellate *Alexandrium tamarense* (Dinophyceae). *Journal of Phycology* 46: 18–32.

Adl SM, Simpson AGB, Farmer MA, *et al.* (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* 52: 399–451.

Aligizaki K, Nikolaidis G, and Fraga S (2008) Is *Gambierdiscus* expanding to new areas? *Harmful Algae News* 36: 6–7.

Anderson DM (1994) Red tides. *Scientific American* 271: 52–58.

Anderson DM, Kulis DM, Doucette GJ, Gallagher JC, and Balech E (1994) Biogeography of toxic dinoflagellates in the genus *Alexandrium* from the northeastern United States and Canada. *Marine Biology* 120: 467–478.

Anderson DM, Kulis DM, Sullivan JJ, Hall S, and Lee C (1990) Dynamics and physiology of saxitoxin production by the dinoflagellates *Alexandrium* spp. *Marine Biology* 104: 511–524.

Armbrust EV (2009) The life of diatoms in the world's oceans. *Nature* 459: 185–192.

Azanza RV and Taylor FJR (2001) Are *Pyrodinium* blooms in the Southeast Asian region recurring and spreading? A view at the end of the millennium. *AMBIO* 30: 356–364.

Bargu S, Silver MW, Goldstein T, Roberts K, and Gulland F (2010) Climate events, foraging patterns, and toxic blooms: Complexity of domoic acid-related sea lion strandings in Monterey Bay, California. *Marine Ecology Progress Series* 418: 213–222.

Bates SS (1998) Ecophysiology and metabolism of ASP toxin production. In: Anderson DM, Cembella AD, and Hallegraeff GM (eds.) *Physiological Ecology of Harmful Algal Blooms*, pp. 405–426. Heidelberg, Berlin/New York: Springer.

Bates SS, Garrison DL, and Horner RA (1998) Bloom dynamics and physiology of domoic acid-producing *Pseudo-nitzschia* species. In: Anderson DM, Cembella AD, and Hallegraeff GM (eds.) *Physiological Ecology of Harmful Algal Blooms*, pp. 267–292. Heidelberg, Berlin/New York: Springer.

Beaufort L, Probert I, de Garidel-Thoron T, *et al.* (2011) Sensitivity of coccolithophores to carbonate chemistry and ocean acidification. *Nature* 476: 80–83.

Belgrano A, Lindahl M, and Hernroth B (1999) North Atlantic oscillation, primary productivity, and toxic phytoplankton in the Gullmar Fjord, Sweden (1985–1996). *Proceedings of the Royal Society B* 266: 425–430.

Brosnahan ML, Kulis DM, Solow AR, *et al.* (2010) Outbreeding lethality between toxic Group I and non toxic Group III *Alexandrium tamarense* spp. isolates: Predominance of heterotypic encystment and implications formating interactions and biogeography. *Deep-Sea Research Part II* 57: 175–189.

Bujak JP and Williams GL (1981) The evolution of dinoflagellates. *Canadian Journal of Botany* 59: 2077–2087.

Casabianca S, Penna A, Pecchioli, Jordi A, Basterretexa G, and Vernesi C (2012) Population structure and connectivity of the harmful dinoflagellate *Alexandrium minutum* in the Mediterranean Sea. *Proceedings of the Royal Society B* 279: 129–138.

Cavalier-Smith T and Chao EE-Y (2006) Phylogeny and megasystematics of phagatrophic heterokonts (Kingdom Cromista). *Journal of Molecular Evolution* 62: 388–420.

Cembella AD (1998) Ecophysiology and metabolism of paralytic shellfish toxins in marine microalgae. In: Anderson DM, Cembella AD, and Hallegraeff GM (eds.) *Physiological Ecology of Harmful Algal Blooms*, pp. 381–403. Heidelberg, Berlin/New York: Springer.

Cembella AD (2003) Chemical ecology of eukaryotic microalgae in marine ecosystems. *Phycologia* 42: 420–447.

Cembella AD and Destombe C (1996) Genetic differentiation among *Alexandrium* populations from eastern Canada. In: Yasumoto T, Oshima Y, and Fukuyo Y (eds.) *Harmful and Toxic Algal Blooms*, pp. 447–450. Paris: Intergovernmental Oceanographic Commission of UNESCO.

Cembella AD, Sullivan JJ, Boyer GL, Taylor FJR, and Andersen RJ (1987) Variation in paralytic shellfish toxin composition within the *Protogonyaulax tamarensis/catenella* species complex; red tide dinoflagellates. *Biochemical Systematics and Ecology* 15: 171–186.

Cembella AD and Taylor FJR (1986) Electrophoretic variability within the *Protogonyaulax tamarensis/catenella* species complex, Pyridine linked dehydrogenases. *Biochemical Systematics and Ecology* 143: 311–323.

Chinain M, Germain M, Sako Y, Pauillac S, and Legrand A-M (1997) Intraspecific variation in the dinoflagellate *Gambierdiscus toxicus* Dinophyceae. I. Isozyme analysis. *Journal of Phycology* 33: 36–43.

Cho S-Y, Nagai S, Nishitani G, and Han M-S (2009) Development of compound microsatellite markers in red-tide-causing dinoflagellate *Akashiwo sanguinea* (Dinophyceae). *Molecular Ecology Resources* 9: 915–917.

Cloern JE, Schraga TS, Lopez CB, Knowles N, Labiosa RG, and Dugdale R (2005) Climate anomalies generate an exceptional dinoflagellate bloom in San Francisco Bay. *Geophysical Research Letters* 32: 1–5. L14608. doi: 10.1029/2005GL023321.

Dale B, Edwards M, and Reid PC (2006) Climate change and harmful algal blooms. In: Granéli E and Turner J (eds.) *Ecology of Harmful Algae*, Ecological Studies Vol. 189, pp. 367–378. Dordrecht, The Netherlands: Springer-Verlag.

Demura M, Kawachi M, Kunugi M, Nishizawa T, Kasai F, and Watanabe MM (2007) Development of microsatellite markers for the red tideforming harmful species *Chattonella antiqua*, *C. marina*, and *C. ovata* (Raphidophyceae). *Molecular Ecology Notes* 7: 315–317.

Destombe C and Cembella A (1990) Mating-type determination, gametic recognition and reproductive success in *Alexandrium excavatum* (Gonyaulacales, Dinophyta), a toxic red-tide dinoflagellate. *Phycologia* 29: 316–325.

Edvardsen B, Eikrem W, Throndsen J, Saez A, Probert I, and Medlin LK (2011) Ribosomal DNA phylogenies and a morphological revision set the basis for a new taxonomy of the Prymnesiales (Haptophyta). *European Journal of Phycology* 46: 202–228.

Erdner DL, Richlen M, McCauley LAR, and Anderson DM (2011) Intrapopulation diversity and dynamics of a widespread bloom of the toxic dinoflagellate *Alexandrium fundyense*. *PLoS One* 6: e22965.

Erickson GM and Nishitani L (1985) The possible relationship of El Nino/Southern oscillation events to interannual variation in *Gonyaulax* populations as shown by records of shellfish toxicity. In: Wooster WS and Fluharty DL (eds.) *El Nino North: Nino Effects in the Eastern Subarctic Pacific Ocean*, pp. 283–290. Seattle, WA: Washington Sea Grant Program.

Fraga S and Bakun A (1990) Global climate change and harmful algal blooms: The example of *Gymnodinium catenatum* on the Galician Coast. In: Smayda TJ and Shimizu Y (eds.) *Toxic Phytoplankton Blooms in the Sea. Developments in Marine Biology,* vol. 3. pp. 59–65. New York: Elsevier.

Frommlet JC and Iglesias-Rodríguez MD (2008) Microsatellite genotyping of single cells of the dinoflagellate species *Lingulodinium polyedrum* (Dinophyceae): A novel approach for marine microbial population genetic studies. *Journal of Phycology* 44: 1116–1125.

Fu FX, Zhang Y, Warner ME, Feng Y, Sun J, and Hutchins DA (2008) A comparison of future increased CO_2 and temperature effects on sympatric *Heterosigma akashiwo* and *Prorocentrum minimum*. *Harmful Algae* 7: 76–90.

Gribble KE, Keafer BA, Quilliam MA, *et al.* (2005) Distribution and toxicity of *Alexandrium ostenfeldii* (Dinophyceae) in the Gulf of Maine, USA. *Deep-Sea Research II* 52: 2745–2763.

Hackett JD, Wisecaver JH, Brosnahan ML, Kulis DM, Anderson DM, and Plumley FG (in press) Independent evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. *Molecular Biology and Evolution*.

Hallegraeff GM (1993) A review of harmful algal blooms and their apparent global increase. *Phycologia* 32: 79–99.

Hallegraeff GM (2010) Ocean climate change, phytoplankton community responses, and harmful algal blooms: A formidable predictive challenge. *Journal of Phycology* 46: 220–235.

Hasle GR (1993) Nomenclatural notes on marine planktonic diatoms. The family Bacillariaceae. *Nova Hedwigia Beih* 106: 315–321.

Hayhome BA, Anderson DM, Kulis DM, and Whitten DJ (1989) Variation among congeneric dinoflagellates from the northeastern United States and Canada. I. Enzyme electrophoresis. *Marine Biology* 101: 427–435.

Henrichs DW, Renshaw MA, Santamaria CA, Richardson B, Gold JR, and Campbell L (2008) PCR amplification of microsatellites from single cells of *Karenia brevis* preserved in Lugol's iodine solution. *Marine Biotechnology* 10: 122–127.

Hoppenrath M and Leander BS (2010) Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences. *PLOS One* 5: e13220.

Ianora A, Bentley MG, Caldwell GS, *et al.* (2011) The relevance of marine chemical ecology to plankton and ecosystem function: An emerging field. *Marine Drugs* 9: 1625–1648.

Imai I, Yamaguchi M, and Hori Y (2006) Eutrophication and occurrences of harmful algal blooms in the Seto Inland Sea, Japan. *Plankton and Benthos Research* 1: 71–84.

Jaeckisch N, Yang I, Wohlrab S, *et al.* (2011) Comparative genomic and transcriptomic characterization of the toxigenic marine dinoflagellate *Alexandrium ostenfeldii*. *PLoS One* 6(12): e28012.

John U, Fensome RA, and Medlin LK (2003) The application of a molecular clock based on molecular sequences and the fossil record to explain the biogeographic distribution within the *Alexandrium tamarense* "species complex". *Molecular biology and Evolution* 20: 1015–1027.

Kudela RM, Seeyave S, and Cochlan WP (2010) The role of nutrients in regulation and promotion of harmful algal blooms in upwelling systems. *Progress in Oceanography* 85: 122–135.

Lane CE and Durnford DG (2010) Endosymbiosis and the evolution of plastids. In: Oren, A and Thane Papke R (eds.) *Molecular Phylogeny of Microorganisms*, pp. 187–221. Norfolk, UK: Caister Academic Press.

Langlois G (2001) Marine biotoxin monitoring in California, 1927–1999. In: RaLonde R (ed.) *Harmful Algal Blooms on the North American West Coast*, pp. 31–34. Fairbanks, AK: University of Alaska Sea Grant College Program.

Leander BS and Keeling P (2004) Early evolutionary history of Dinoflagellates and Apicomplexa (Alveolata) as inferred from HSP90 and actin phylogenies. *Journal of Phycology* 40: 341–350.

Legrand C, Rengefors K, Fistarol GO, and Granéli E (2003) Allelopathy in phytoplankton: Biochemical, ecological and evolutionary aspects. *Phycologia* 42: 406–419.

Lilly EL, Halanych KM, and Anderson DM (2007) Species boundaries and global biogeography of the *Alexandrium tamarense* complex (Dinophyceae). *Journal of Phycology* 43: 1329–1338.

Lilly EL, Kulis DM, Gentian P, and Anderson DM (2002) Paralytic shellfish poisoning toxins in France linked to a human-introduced strain of *Alexandrium catenella* from the western Pacific: Evidence from DNA and toxin analysis. *Journal of Plankton Research* 24: 443–452.

Ma H, Krock B, Tillmann U, Bickmeyer U, Graeve M, and Cembella A (2011) Mode of action of membrane-disruptive lytic compounds from the marine dinoflagellate *Alexandrium tamarense*. *Toxicon* 58: 247–258.

MacLean JL (1977) Observations on *Pyrodinium bahamense* Plate, a toxic dinoflagellate in Papua New Guinea. *Agricultural Journal* 24: 131–138.

McMinn A (1989) Late Pleistocene dinoflagellate cysts from Botany Bay, New South Wales, Australia. *Micropaleontology* 35: 1–9.

Masseret E, Grzebyk D, Nagai S, *et al.* (2009) Unexpected genetic diversity among and within populations of the toxic dinoflagellate *Alexandrium catenella* as revealed by nuclear microsatellite markers. *Applied and Environmental Microbiology* 75: 2037–2045.

Medlin L, Lange M, Wellbrock U, *et al.* (1998) Sequence comparison links toxic European isolates of *Alexandrium tamarense* from the Orkney Islands to toxic North American stocks. *European Journal of Protistology* 34: 329–335.

Moore SK, Trainer VL, Mantua NJ, *et al.* (2008) Impacts of climate variability and future climate change on harmful algal blooms and human health. *Environmental Health* 7: S4.

Nagai S, Lian C, Hamaguchi M, Matsuyama Y, Itakaru S, and Hogetsu T (2004) Development of microsatellite markers in the toxic dinoflagellate *Alexandrium tamarense* (Dinophyceae). *Molecular Ecology Notes* 4: 83–85.

Nagai S, Lian C, Yamaguchi S, *et al.* (2007a) Microsatellite markers reveal population genetic structure of the toxic dinoflagellate *Alexandrium tamarense* (Dinophyceae) in Japanese coastal waters. *Journal of Phycology* 43: 43–54.

Nagai S, McCauley L, Yasuda N, *et al.* (2006a) Development of microsatellite markers in the toxic dinoflagellate *Alexandrium minutum* (Dinophyceae). *Molecular Ecology Notes* 6: 756–758.

Nagai S, Nishitani G, Sakamoto S, *et al.* (2009) Genetic structuring and transfer of marine dinoflagellate *Cochlodinium polykrikoides* in Japanese and Korean coastal waters revealed by microsatellites. *Molecular Ecology* 18: 2337–2352.

Nagai S, Nishitani G, Yamaguchi M, *et al.* (2007b) Development of microsatellite markers in the noxious red tide-causing dinoflagellate *Heterocapsa circularisquama* (Dinophyceae). *Molecular Ecology Notes* 7: 993–995.

Nagai S, Sekino M, Matsuyama Y, and Itakura S (2006b) Development of microsatellite markers in the toxic dinoflagellate *Alexandrium catenella* (Dinophyceae). *Molecular Ecology Notes* 6: 120–122.

Nagai S, Yamaguchi S, Lian LC, Matsuyama Y, and Itakura S (2006c) Development of microsatellite markers in the noxious red tide causing algae *Heterosigma akashiwo* (Raphidophyceae). *Molecular Ecology Notes* 6: 477–479.

Nishitani G, Nagai S, Masseret E, *et al.* (2007) Development of compound microsatellite markers in the toxic dinoflagellate *Alexandrium catenella* (Dinophyceae). *Plankton and Benthos Research* 2: 128–133.

Paerl HW and Huisman J (2008) Blooms like it hot. *Science* 320: 57–58.

Paerl HW and Huisman J (2009) Climate change: A catalyst for global expansion of harmful cyanobacterial blooms. *Environmental Microbiology Reports* 1: 27–37.

Penna A, Garcés E, Vila M, *et al.* (2005) *Alexandrium catenella* (Dinophyceae), a toxic ribotype expanding in the NW Mediterranean Sea. *Marine Biology* 148: 13–23.

Pitcher GC and Boyd AJ (1996) Across-shelf and alongshore dinoflagellate distributions and the mechanisms of red tide formation within the southern

Benguela upwelling system. In: Yasumoto T, Oshima Y, and Fukuyo Y (eds.) *Harmful and Toxic Algal Blooms*, pp. 243–246. Paris: Intergovernmental Oceanographic Commission of UNESCO.

Reid PC, Colebrook JM, Matthews JBL, Aiken J and Continuous Plankton Recorder Team (2003) The continuous plankton recorder: Concepts and history, from plankton indicator to undulating recorders. *Progress in Oceanography* 58: 117–173.

Rensel JE and Whyte JNC (2003) Finfish mariculture and harmful algal blooms. In: Hallegraeff GM, Anderson DM, and Cembella AD, (eds.) *Manual on Harmful Marine Microalgae* vol. 11. pp. 693–722. Paris: UNESCO. ch. 25, Monographs on Oceanographic Methodology.

Renshaw MA, Soltysiak K, Arreola D, *et al.* (2006) Microsatellite DNA markers for population genetic studies in the dinoflagellate *Karenia brevis*. *Molecular Ecology Notes* 6: 1157–1159.

Riebesell U, Zondervan I, Rost B, Tortell PD, Zeebe RE, and Morel FMM (2000) Reduced calcification of marine plankton in response to increased atmospheric CO_2. *Nature* 407: 364–367.

Saunders GW, Hill DRA, Sexton JP, and Anderson RA (1997) Small subunit ribosomal RNA sequences from selected dinoflagellates: Testing classical evolutionary hypotheses with molecular systematic methods. *Plant Systematics and Evolution* 11(supplement): 237–259.

Saldarriaga JF, Taylor FJR, Cavalier-Smith T, Medden-Duer S, and Keeling PJ (2004) Molecular data and the evolutionary history of dinoflagellates. *European Journal of Protistology* 40: 85–111.

Sako Y, Kim CH, Ninomiya H, Adachi M, and Ishida Y (1990) Isozyme and cross analysis of mating populations in the *Alexandrium catenella/tamarense* species complex. In: Granéli E, Sundström B, Edler L, and Anderson DM (eds.) *Toxic Marine Phytoplankton*, pp. 320–329. New York: Elsevier.

Scholin CA, Hallegraeff GM, and Anderson DM (1995) Molecular evolution of the *Alexandrium tamarense* species complex (Dinophyceae), dispersal in the North American and West Pacific regions. *Phycologia* 34: 472–485.

Skjoldal HR and Dundas I (1991) The *Chrysochromulina polylepis* bloom in the Skagerrak and Kattegat in May–June 1988: Environmental conditions, possible causes, and effects. *ICES Cooperative Research Report* 175: 1–59.

Skov J, Lundholm N, Pocklington R, Rosendahl S, and Moestrup Ø (1997) Studies on the marine planktonic diatom *Pseudo-nitzschia*. 1. Isozyme variation among isolates of *P. pseudodelicatissima* during bloom in Danish coastal waters. *Phycologia* 36: 374–380.

Smayda TJ (1990) Novel and nuisance phytoplankton blooms in the sea: Evidence for a global epidemic. In: Granéli E, Sundström B, Edler L, and Anderson DM (eds.) *Toxic Marine Phytoplankton*, pp. 29–40. New York: Elsevier.

Smayda TJ (1997) Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography* 42: 1137–1153.

Smetacek V (2001) A watery arms race. *Nature* 411: 745.

Steidinger KA, Wolny JL, and Haywood AJ (2008) Identification of Kareniaceae (Dinophyceae) in the Gulf of Mexico. *Nova Hedwigia Beiheft* 133: 269–284.

Tester PA, Stumpf RP, Vukovich FM, Folwer PK, and Turner JT (1991) An expatriate red tide bloom: Transport, distribution, and persistence. *Limnology and Oceanography* 36: 1053–1061.

Turner JT, Tester PA, and Hansen PJ (1998) Interactions between toxic marine phytoplankton and metazoan and protistan grazers. In: Anderson DM, Cembella AD, and Hallegraeff GM (eds.) *Physiological Ecology of Harmful Algal Blooms*, pp. 453–474. Heidelberg, Berlin/New York: Springer.

Wohlrab S, Iversen MH, and John U (2010) A molecular and co-evolutionary context for grazer induced toxin production in *Alexandrium tamarense*. *PLoS One* 5: e15039.

Wood AM and Leatham T (1992) The species concept in phytoplankton ecology. *Journal of Phycology* 28: 723–729.

Yang I, John U, Beszteri S, *et al.* (2010) Comparative gene expression in toxic versus non-toxic strains of the marine dinoflagellate *Alexandrium minutum*. *BMC Genomics* 11: 248.

Yang I, Selander E, Pavia H, and John U (2011) Grazer-induced toxin formation in dinoflagellates: A transcriptomic model study. *European Journal of Phycology* 46: 66–73.

Zhang H, Bhattacharya D, and Lin S (2007) A three-gene dinoflagellate phylogeny suggests monophyly of Prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*. *Journal of Molecular Evolution* 65: 463–474.

Biodiversity of the Succulent Karoo

Philip W Rundel, University of California, Los Angeles, CA, USA
Richard M Cowling, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aeolianites Wind-deposited sediments.
Aphyllous Leafless.
CAM metabolism Crassulacean acid metabolism is a carbon uptake system utilized by many succulent plants; these plants fix carbon dioxide during the night when evapotranspiration is relatively low, storing it as organic acids. The stored CO_2 is released during the day and refixed in the typical light reactions of photosynthesis, thereby increasing water use efficiency.
Drought-deciduous Descriptive of plant species that lose their leaves during the dry season as soil moisture becomes limited.
Edaphic Related to or caused by particular soil conditions, as of texture or drainage, rather than by physiographic or climatic factors.
Fynbos Evergreen sclerophyll shrubland dominating the Cape Region of South Africa.
Geophytes Plant with succulent storage tissues such as bulbs, corms, or rhizomes.
Inselberg Isolated mountain that stands above lowland plains below.
Orthophyllous Pollinated by birds.
Petaloid Describing flowers with petals present, and typically showy.
Succulent Having fleshy tissues that store moisture.

Introduction

The Succulent Karoo comprises an area of approximately 103,000 km^2 stretching down the Atlantic coast of southwestern Namibia and South Africa, as well as in dry intermontane valleys and basins within the Cape fynbos biome (**Figure 1**). This desert area is characterized by a dominance of winter rainfall and shows strong floristic affiliation with the Cape Floristic Region such that most of the area has been included in a Greater Cape Floristic Region (Born *et al.*, 2007). Of special note is the fact that the Succulent Karoo is one of only two biodiversity hotspots that are entirely arid, with the other being the newly recognized Horn of Africa hotspot (Mittermeier *et al.*, 2005).

The Succulent Karoo is commonly divided into two broad components (Jürgens, 1991). The first, Namaqualand-Namib, extends along the west coast of South Africa and southern Namibia. It is a winter-rainfall desert (30–300 mm $year^{-1}$) with a mild climate moderated by cold Atlantic Ocean currents. The second zone, the Southern Karoo, has a more interior position in elevated intermontane valleys and basins, and thus experiences more extremes of temperature variation than the Namaqualand region (Desmet and Cowling, 1999b). It has a major peak of rainfall in spring and a second peak in autumn, although summers are invariably dry and appreciable winter rain does fall. Annual rainfall ranges from 180 to 350 mm.

The Succulent Karoo is exceptional among the world's desert biomes because of its spectacular diversity of plant species (Cowling *et al.*, 1998; Cowling and Pierce, 1999). The most remarkable component of this diversity lies with succulent plant species, which occur here with richness unparalleled to any other region (van Jaarsveld, 1987). It is this dominance of succulent-leaved plant species that gives rise to the name, Succulent Karoo, in contrast to the larger Nama Karoo biome of South Africa (Cowling *et al.*, 1999). Although the Succulent Karoo and summer-rainfall Nama Karoo biomes were once lumped together as a warm desert unit of southern Africa with subtropical affinities, the winter-rainfall region has been shown to have strong affinities to the Cape Floristic Region, whereas the Nama Karoo retains subtropical African affinities (Born *et al.*, 2007; Verboom *et al.*, 2008).

The rich biodiversity of the Succulent Karoo is due in large part to an extensive and complex array of subtle biotic changes, derived from topographic, edaphic, and climatic diversity in the biome's landscape (Cowling *et al.*, 2009; Linder, 2008). This landscape lacks high mountains but varies from young coastal dunes across older dune surfaces to upland areas of granitic ranges in Namaqualand, and interior valleys. Many species are extreme habitat specialists with highly limited ranges. Local endemism (i.e., the restriction of species to extremely small ranges of <50 km^2) is most pronounced among small Mesembryanthemaceae and other succulents, as well as in geophytes (Desmet and Cowling, 1999a). Similar patterns of compositional change along gradients have been observed for certain groups of invertebrates (Colville *et al.*, 2002).

Dwarf shrublands dominated by low shrubs and subshrubs with evergreen succulent leaves are found throughout the Succulent Karoo and represent the characteristic plant cover (Cowling *et al.*, 1994, 1999). These drought-adapted plants have fleshy leaves or stems for water storage. In the Succulent Karoo, there are approximately 1700 species of leaf succulents, and this dominance is unique among the world's deserts (Cowling *et al.*, 1998, 1999). The recent (Plio-Pleistocene) and explosive diversification of the Mesembryanthemaceae, the largest group, has been described as an event unrivaled among flowering plants (Klak *et al.*, 2004). There are 108 genera and 1132 species reported for this family in the Succulent Karoo (**Table 1**). Stem succulents are also found here (around 140 species), as are seasonal bulbs and annuals that display magnificent spring blooms in the open spaces between the shrubs, particularly during the spring in Namaqualand (Cowling and Pierce, 1999).

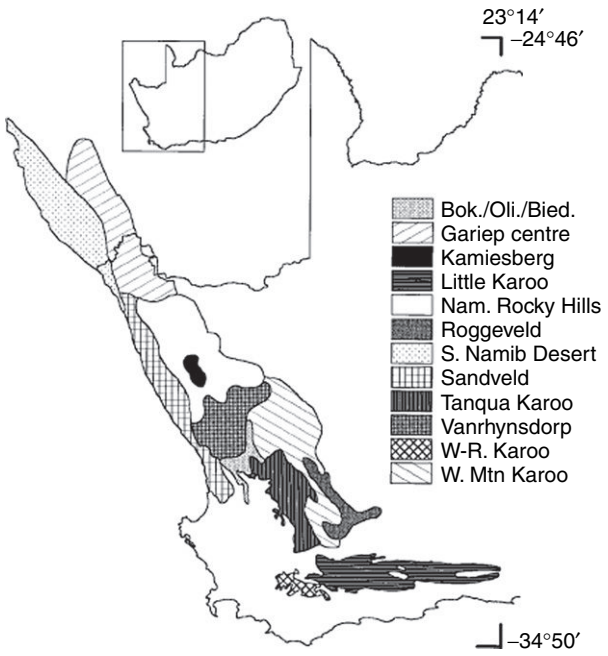


Figure 1 Bioregions of the Succulent Karoo (modified after Hilton-Taylor C (1996) Patterns and characteristics of the flora of the Succulent Karoo Biome, southern Africa. In: van der Maesen UE, van der Burgt XM, and van Medenbach de Rooy JM (eds.) *The Biodiversity of African Plants*, pp. 58–72. Dordrecht: Kluwer). (Bok./Oli./Bied. = Bokkeveld/Olifants River/Biedou bioregion; Nam. Rocky Hills = Namaqualand Rocky Hills; S. Namib Desert = Southern Namib Desert; Vanrhynsdorp = Vanrhynsdorp Centre, W-R. Karoo = Worcester-Robertson Karoo; W. Mtn Karoo = Western Mountain Karoo.)

Table 1 The 10 largest plant families in the Succulent Karoo. The Mesembryanthemaceae is here considered separate for the Aizoaceae

Family	*Genera*	*Species*
Mesembryanthemaceae	108	1132
Asteraceae	108	777
Iridaceae	29	408
Scrophulariaceae	39	349
Fabaceae	57	315
Poaceae	84	297
Crassulaceae	5	202
Hyacinthaceae	22	172
Asphodelaceae	8	149
Euphorbiaceae	13	140
Total		3941

Bioregions

Twelve geographic bioregions have been identified within the Succulent Karoo (**Figure 1**), reflecting geomorphological variation, floristic affinities, and patterns of endemism (Hilton-Taylor, 1996). Ten of these lie in the Namaqualand-Namibian area of the Succulent Karoo, whereas two others form the Southern Karoo area. The Southern Namib Desert encompasses the northern extent of the Succulent Karoo, largely comprised by the Sperrgebeit in southwestern Namibia and the northwestern South African coast from Port Nolloth to Alexander Bay. The Orange River, the largest river in southern Africa, reaches the coast in this region and forms the boundary between South Africa and Namibia. Diamond mining concessions for this area date back a century, and have restricted public and scientific access. Although mining activities have sharply transformed a narrow belt along the coast (Desmet and Cowling, 1999a), much of the remainder of the area is in pristine condition. A notable site within this area is the fog flora of lichens that completely dominates an area of the coast between Port Nolloth and Alexander Bay. There are 776 plant species reported for the Sperrgebeit, with a third of these endemic (Burke, 2004).

The Gariep Center, which includes upland areas of southwestern Namibia and the greater Richtersveld area, has the richest flora of succulent plants in the world. This region contains an estimated 2700 plant species, with leaf and stem succulents as the overwhelming majority of the flora (van Wyck and Smith, 2001). The approximately 20% of these species that are endemic include many charismatic succulents with highly restricted ranges. The Richtersveld National Park, leased to the South African National Parks by the indigenous local population, protects much of the rugged mountainous portions of this bioregion, and is home to diverse wildlife.

The Namaqua Rocky Hills comprise the largely granitic uplands of central Namaqualand, south of the Greater Richtersveld bioregion. The Namaqua Rocky Hills are famous for their spectacular displays of spring wildflowers, and their remarkable diversity and endemism of geophytes (van Rooyen *et al.*, 1990). Agriculture and heavy grazing, the latter particularly on communal lands, have transformed much of the level terrain within this bioregion (Hoffman and Rohde, 2007; Thompson *et al.*, 2009). The recently established Namaqua National Park is protecting a large part of this region. The Kamiesberg Mountains rise to 1700 m elevation within the Namaqua Rocky Hills, and this elevation provides sufficient rainfall to support patches of fynbos vegetation. The Kamiesberg supports a rich flora of approximately 1000 species, including 80 endemics, most of which are geophytes (van Wyck and Smith, 2001).

The Sandveld comprises a long and narrow belt of the Namaqualand coast dominated by geologically young sandy soils, mostly alkaline but also acid aeolianites. The vegetation is distinct structurally and floristically and gradually grades southward into the coastal thickets and fynbos of the Cape Floristic Region.

The Vanrhynsdorp Center is largely composed of the Knersvlakte, a broad dry plain of approximately 5200 km^2 bounded on the east by the Bokkeveld Mountains. Extensive plains covered by quartz pebbles occur across this bioregion, a consequence of Pliocene uplift and subsequent erosion of quartz veins. The Knersvlakte is rich in plant diversity, despite its barren appearance, with dwarf succulents scattered in huge densities among the quartz pebbles of its plains (Schmiedel and Jürgens, 1999). A flora of 1324 species is reported for this bioregion; some 15% of these are endemic. South of the Vanrhynsdorp Center lies the Bokkeveld/Oliphants River/Biedou bioregion, which forms a transition to the Cape Region communities to the south.

The Western Mountain Karoo occupies high country along the Great Escarpment, a feature that separates the coastal and upland plateaus. Frost is common during the winter months and succulent cover is much sparser than elsewhere in the Succulent Karoo. This bioregion harbors approximately 2500 species, 10% of which are endemic (van Wyck and Smith, 2001). Most of the endemics are geophytes. The Tanqua Karoo occupies a very arid basin that lies between the rain-trapping Cape Fold Mountains to the west, and the Great Escarpment (Western Mountain Karoo) to the east. Owing to the low rainfall ($>20\,\text{mm year}^{-1}$ falls in the center of the basin), vegetation cover is sparse. However, the area does support a surprisingly rich but poorly explored flora.

The final two bioregions of the Succulent Karoo lie within the Southern Karoo bioregion. These regions experience a very different climate regime to the Namaqualand region because of their location away from the influence of the cold Atlantic Ocean. Summer temperatures regularly reach 40 °C, and summer storms penetrate to this region from the warm Indian Ocean to the south, allowing for a biseasonal pattern of rainfall. Nonetheless, mean annual rainfall is relatively reliable – at least by global standards – but less so than for Namaqualand.

The Worcester/Robertson Karoo bioregion lies along the Central Breede River Valley in a rainshadow created by the surrounding Cape Fold Mountains. The bioregion has a flora of 1500 species of which about 8% are endemic; some three quarters of the endemics are succulents (van Wyck and Smith, 2001). Extensive irrigated agriculture and vineyards have transformed much of this area, although pristine patches of landscape remain. The final bioregion, the Little Karoo, comprises a dissected basin lying between the Langeberg and Swartberg Cape Fold Mountains. The flora comprises more than 2500 species of which about 10% are endemic; succulents comprise about one quarter of the endemics (van Wyck and Smith, 2001; Vlok and Schutte-Vlok, 2010). Many of the succulent endemics grow on quartz fields similar to those of the Knersvlakte, described above (Schmiedel and Jürgens, 1999). Extensive areas of this bioregion have been heavily impacted by ostrich farming.

Biodiversity

Plants

For an arid biome, the Succulent Karoo has extraordinarily high plant diversity, including the richest succulent flora in the world (Hilton-Taylor, 1996; Cowling *et al.*, 1998). The size of the flora is estimated at 6356 species, including floristically related communities present on the Bushmanland inselbergs lying in the Nama Karoo biome to the east of the Gariep bioregion, with about 40% of these endemic. Many of these are point endemics, confined to a single habitat that covers a very small area. Approximately 100 genera are endemic to the Succulent Karoo, most of which are succulents (van Wyck and Smith, 2001). The Mesembryanthemaceae, considered here as a segregate family from the Aizoaceae, exhibits a remarkable degree of biodiversity in the Succulent Karoo, with 108 genera and 1132 species. This total for the Mesembryanthemaceae comprises about 18% of the flora in this single family. The largest 10 families make up more than 60% of the total vascular plant flora. The Asteraceae is the second largest family, making the Succulent Karoo of global significance for the evolution of this family. Three families of geophytes are included in the 10 largest groups present in the Succulent Karoo. These are the Iridaceae, the Hyacinthaceae, and the Asphodelaceae.

Notable plant species found in this hotspot include the botterboom (*Tylecodon paniculatus*), a stem succulent that has glossy leaves in winter and red flowers in summer, and the halfmens (*Pachypodium namaquanum*), a strange stem succulent endemic to the Richtersveld that can grow up to 4 m tall. Clusters of halfmens stems tend to face toward the north, giving the appearance of groups of people gazing northward. The terminal crowns of leaves, which resemble hairy human heads, enhance the impression. The scientific explanation for this unusual orientation is that the plants, which grow on shaded slopes, lean northward in order to ensure that their leaves and developing flowerheads, produced during the cool, foggy winter months, are maximally exposed to solar radiation (Rundel *et al.*, 1995).

Vertebrates

Vertebrate diversity is moderately high in the Succulent Karoo (**Table 2**), particularly for reptiles (Vernon, 1999; Mittermeier *et al.*, 2005). There are 75 mammal species reported in the Succulent Karoo, of which four are endemic or near endemic. These are De Winton's golden mole (*Cryptochloris wintoni*, vulnerable), Van Zyl's golden mole (*Cryptochloris zyli*), Namaqua dune mole rat (*Bathyergus janetta*), and a bat (*Mysotis seabrai*). Major concentrations of large mammals, including elephant (*Loxodonta africana*, vulnerable status), black rhinoceros (*Diceros bicornis*, critical status), mountain zebra (*Equus zebra*), quagga (*Equus quagga*), giraffe (*Giraffa camelopardalis*), and African buffalo (*Syncerus caffer*), once roamed the plains and gallery forests along the Orange River in the Succulent Karoo. These populations have now disappeared. Small herds of gemsbok (*Oryx gazella*) and springbok (*Antidorcas marsupialis*) are present today.

The avifauna of the Succulent Karoo includes approximately 226 species, one of which is endemic, Barlow's lark

Table 2 Biodiversity and endemism of vascular plants and vertebrates in the Succulent Karoo

Taxonomic group	*Species*	*Endemic species*	*Percent endemism*
Plants	6356	2439	38.4
Mammals	75	2	2.7
Birds	226	1	0.4
Reptiles	94	15	16.0
Amphibians	21	1	4.8
Freshwater Fishes	28	0	0.0

Source: Reproduced from Driver A, Desmet PG, Rouget M, Cowling RM, and Maze KE (2003) *Succulent Karoo Ecosystem Plan: Biodiversity Component Technical Report.* Cape Town: Cape Conservation Unit, Botanical Society of South Africa, and Mittermeier RA, Gil PR, and Hoffman M, *et al.* (2005) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions.* Mexico City: CEMEX.

(*Certhilauda barlowi*). Other species in the biome include the black harrier (*Circus maurus*, vulnerable status) which has the most restricted range of the world's 13 harrier species, Karoo bustard (*Eupodotis vigorsii*), Ludwig's bustard (*Neotis ludwigii*), Karoo chat (*Cercomela schlegelii*), dune lark (*Certhilauda erythrochlamys*), and dusky sunbird (*Nectarinia fusca*).

Reptile diversity is relatively high in the Succulent Karoo, with approximately 94 species, of with 15 are endemic. All of the endemics are geckos and lizards, representing about a quarter of the nearly 60 geckos and lizard species in the biome. These endemics include seven species of girdled lizards of the genus *Cordylus*, including the armadillo girdled lizard (*C. cataphractus*, vulnerable status). This lizard has a heavily armored body and spiny tail, and is known for rolling into a tight ball when threatened. As in the other parts of South Africa, tortoise diversity is high in the Succulent Karoo with seven taxa, two of which are endemic. These are the Namaqualand tent tortoise (*Psammobates tentorius trimeni*) and the Namaqualand speckled padloper (*Homopus signatus signatus*).

Amphibians are poorly represented in the Succulent Karoo with 21 species. All of these species are frogs, and only one is endemic. This is the desert rain frog (*Breviceps macrops*, vulnerable status), which occurs along the Namaqualand coast of South Africa north to Luderitz in southwestern Namibia.

There are approximately 26 species of freshwater fish in the Succulent Karoo, none of which are endemic.

Invertebrates

Invertebrate diversity is relatively high in the Succulent Karoo, and evidence suggests that more than half of the species in many insect groups are endemic (Vernon, 1999). Of the 50 species of scorpions found in the hotspot, nearly 22 are endemic. Monkey beetles, which are almost exclusively found in southern Africa, are concentrated in the area. Along with many types of wasps and bees, these beetles pollinate a diverse assemblage of endemic plant species (Colville *et al.*, 2002). Perhaps the most unusual invertebrates found here are the long-tongued flies (Nemestrinidae), which can have mouthparts up to 50 mm long. The flies are the exclusive pollinators of 28 different plant species in Namaqualand (Manning and Goldblatt, 1996).

Plant Growth Forms and Vegetation Structure

The growth form mix of Namaqualand communities is unique among the world's desert ecosystems. Dwarf (<0.5 m) succulent shrubs (chamaephytes), mainly members of the Mesembryanthemaceae, Crassulaceae, and Asteraceae, are the dominant growth form over much of the Succulent Karoo, particularly in Namaqualand (Cowling *et al.*, 1994, 1999). Such a dominance of leaf succulents contrasts sharply with other deserts, including winter-rainfall ones, where this growth form is rare (Esler *et al.*, 1999). Many but not all of these succulents utilize Crassulacean acid metabolism (CAM) metabolism, with both obligate CAM and CAM flexibility present (Rundel *et al.*, 1999).

Many Succulent Karoo leaf succulents have shallow (0.1–0.2 m) rooting systems (Eccles *et al.*, 1999; Esler *et al.*, 1999; Carrick, 2003), even when growing in deep soils. A consequence of this unusual below- and aboveground morphology is that members of this growth form are vulnerable to drought-induced mortality after rare episodes of lower-than-average rainfall. Indeed, in contrast to many other desert ecosystems where the dominant perennial component is very long-lived, many of Namaqualand's leaf succulents have relatively short lifespans (5–10 years) (Jürgens, 1991).

A unique feature of the succulent flora of the Succulent Karoo is the presence of numerous minute (<25 mm tall) succulents belonging to the Mesembryanthemaceae (e.g., *Argyroderma*, *Conophytum*, and *Fenestraria*), Crassulaceae (*Crassula* and *Tylecodon*), Portulacaceae (*Anacampseros*), and Asphodelaceae (*Bulbine*, *Haworthia*) (van Jaarsveld, 1987). Many of these grow embedded in the soil with only their upper surface exposed. The occurrence of such minute perennial growth forms is paradoxical in the context of a desert environment since their small storage organs offer little capacity to endure prolonged moisture stress. Many species grow in shaded rock crevices where they are sheltered from excessive radiation and can exploit run off from rainfall, fog, and dew, whereas others are associated with exposed sites, especially quartz fields (Schmiedel and Jürgens, 1999). Ultimately, the evolution of this intriguing growth form would only be possible in an environment characterized by low but highly predictable inputs of moisture, thereby enabling plants to replenish moisture reserves on a regular basis. Indeed, there is good evidence that in southern Africa, succulence, especially the incidence of dwarf leaf succulent forms, is correlated with rainfall reliability and seasonality (winter rain) rather than annual totals (Cowling *et al.*, 1999).

Taller succulent shrubs and succulent trees are relatively uncommon in the flora of the Succulent Karoo. Many of the larger shrubs are stem succulent *Euphorbia* species, which are particularly common in the warmer and drier northern reaches, especially the greater Richtersveld. This area is also home to most of the biome's small but charismatic arborescent succulent flora, namely *Aloe dichotoma*, *A. pillansii*, *A. ramosissima* (Asphodelaceae), and *P. namaquanum* (Apocynaceae). Shrubs with drought-deciduous, orthophyllous leaves, and aphyllous forms, which comprise the dominant perennial cover in other warm, winter-rainfall deserts (Esler *et al.*, 1999), are relatively rare in the Succulent Karoo where the great majority of nonsucculent woody species are evergreen (Cowling *et al.*, 1994). Evergreen shrubs and trees include lineages endemic to the Karoo-Namib Biome (e.g., *Pteronia* and other Asteraceae) as well as species of tropical affinity, related to southern Africa's extensive subtropical thicket ecosystems. The latter, a relatively small flora comprising species of *Euclea*, *Diospyros* (Ebenaceae), *Rhus*, *Ozoroa* (Anacardiaceae), *Ficus* (Moraceae), and *Schotia* (Fabaceae), are mainly restricted to water courses and run-on sites such as the base of large, granite domes. These tall shrubs and low trees are likely relicts of the late Tertiary, summer-rainfall scrub forests or thickets that once dominated local landscapes of the Succulent Karoo. They are extremely long-lived, have vertebrate-dispersed propagules, and probably establish seedlings

on rare occasions. This life history contrasts strongly with that exhibited by the short-lived and recruitment-dependent lifestyles of most other Namaqualand perennials (Cowling *et al.*, 1999).

Another unique feature of the Succulent Karoo growth form mix is the high diversity and local abundance of geophytes. Namaqualand is home to approximately 480 species (17.5% of the flora) comprising approximately 100 genera in 19 families. At a smaller spatial scale, petaloid monocot geophytes make up approximately 16% of the flora (550 spp. in 150 km^2) at Goegap Nature Reserve in the Namaqua Rocky Hills (van Rooyen *et al.*, 1990), and nearly one third of the flora (280 spp. in 1.2 km^2) at Nieuwoudtville on the western fringe of the Hantam-Roggeveld bioregion. The proportional representation of geophytes in other semiarid, winter-rainfall biomes is much lower. Namaqualand geophytes encompass a wide variety of leaf types, including species with large, broad leaves fully pressed against the soil surface as well as species with tightly coiled and variously pleated leaves. Storage organ size varies from very large in many amaryllids to very small in *Oxalis*. However, the majority of species have small storage capacity; bulbs and corms of most geophytes are 1–3 cm in diameter. If geophytes are functionally interpreted as underground succulents, then their success in Namaqualand, like that of small-bodied aboveground succulents, may partly result from the biome's reliable moisture regime (Procheş *et al.*, 2005).

Annuals, which are not especially conspicuous in undisturbed vegetation, are nevertheless an important component of the Namaqualand flora, comprising 28% of the flora at Goegap Nature Reserve (van Rooyen *et al.*, 1990). Generally, proportional representation of annuals is much lower than in semiarid, winter-rainfall regions of the Mediterranean Basin and the New World. Spectacular displays of spring-flowering annuals, especially on old fields, are a popular tourist feature. The reliability of this display, which is marketed as an annual ecotourism event, is a function of the biome's predictable rainfall regime (Cowling and Pierce, 1999).

Threats to the Succulent Karoo

Two major threats to the ecology and biodiversity of the Succulent Karoo have been identified. These are threats from global climate change and threats from land degradation. The potential impacts of global change are large. Biome-level approaches likely underestimate the risk of species diversity loss from climate change impacts in the Succulent Karoo biome because of the large number of narrow range endemics that would likely suffer significant range dislocation with changing conditions (see Midgley *et al.*, 2002).

Irreversible land transformation associated with cultivation and urban areas in the Succulent Karoo is not extensive because of the low population density of the area. Extensive livestock dominates 90% of the biome, and it is this land use that has the greatest impact on biodiversity. For example, in the Little Karoo, 28% of the area of Succulent Karoo vegetation has been irreversibly transformed by overgrazing (Thompson *et al.*, 2009). Although overgrazing in communal lands represents a threat to the biota of the Succulent Karoo, the recent purchase of 330,000 ha by the South African government for communal landowners should go some way to alleviating population pressures that were a consequence of past apartheid policies. However, stocking rates on newly acquired lands will need to be conservative to maintain conservation values (Hoffman *et al.*, 2007; O'Farrell *et al.*, 2010).

Diverse groups of grazing animals were once widely present in the Succulent Karoo, with seasonal migrations from lowlands (winter) to uplands (summer). As a result of the demise of these herds, grazing as a land use is not incompatible with biodiversity conservation if managed properly. Well-managed grazing may help maintain available niches for plant diversity. There are sufficiently extensive areas of intact habitat for many options to design and implement large conservation corridors (Cowling *et al.*, 1999; Desmet *et al.*, 2002). Implementation will, however, require some restoration in mined and severely overgrazed areas and mechanisms to reduce grazing pressure.

Nevertheless, the Succulent Karoo is vulnerable to land use pressures from overgrazing on communal lands, ostrich farming in the southeast, mining activities, and the illegal collection of plants and animals for trade. Of these, the greatest threats to biodiversity are associated with overgrazing (Alsopp, 1999, Todd and Hoffman, 1999) and mining. Overgrazing is most problematic in communal areas, which support at least double the prescribed stocking rate. Agriculture is restricted to a limited number of irrigable habitats that collectively occupy a small area. Similarly, off-road impacts are also restricted spatially. In the Southern Karoo ostrich farming has seriously transformed large areas. Diamond mining along the northwestern coast of South Africa and southwestern coast of Namibia has had major impacts on local areas but protected other pristine areas (Desmet and Cowling, 1999a).

Succulent Karoo habitats and species are inadequately represented in protected areas; only 3.5%, of the biome is formally protected in seven statutory reserves (Driver *et al.*, 2003), the largest of which is the 162,445 ha Richtersveld National Park, where livestock grazing is permitted. Owing to the high turnover of species along environmental and geographic gradients, almost the entire Succulent Karoo will be required to achieve representation targets for its entire species (Lombard *et al.*, 1999). Therefore, the current conservation strategy – supported by numerous local and international conservation organizations that participate in the Succulent Karoo Ecosystem Planning (SKEP) project – is to identify and implement incentives to encourage good stewardship on private and communal land, and to purchase land for protected areas in priority sites that can sustain tourism ventures. Good progress is being made on both fronts. For example, 3.2 million ha will shortly be included in protected areas associated with the Sperrgebiet-|Ai-|Ais-Richtersveld Transfrontier Park; approximately 50,000 ha have been recently added to the protected area system in the Kamiesberg, Sandveld, and Vanrhynsdorp bioregions; and approximately 20,000 ha to those in the Little Karoo. SKEP has initiated numerous community-based and private conservation initiatives, all of which bode well for the long-term conservation of the Succulent Karoo.

See also: Africa, Ecosystems of. Desert Ecosystems. Desertification. Hotspots. Mediterranean-Climate Ecosystems. Plant Conservation

References

Allsopp N (1999) Effects of grazing and cultivation on soil patterns and processes in Namaqualand. *Plant Ecology* 142: 179–187.

Born J, Linder HP, and Desmet PG (2007) The Greater Cape Floristic Region. *Journal of Biogeography* 34: 147–162.

Burke A (2004) A preliminary account of patterns of endemism in Namibia's Sperrgebiet – the Succulent Karoo. *Journal of Biogeography* 31: 1613–1622.

Carrick PJ (2003) Competitive and facilitative relationships among three shrubs, and the role of browsing intensity and rooting depth in the Succulent Karoo, South Africa. *Journal of Vegetation Science* 14: 761–772.

Colville J, Picker MD, and Cowling RM (2002) Species turnover of monkey beetles (Scarabaeidae: Hopliini) along environmental and disturbance gradients in the Namaqualand region of the Succulent Karoo, South Africa. *Biodiversity and Conservation* 11: 243–264.

Cowling RM, Esler KJ, Midgley GF, and Honig MA (1994) Plant functional diversity, species diversity and climate in arid and semi-arid southern Africa. *Journal of Arid Environments* 27: 141–158.

Cowling RM and Pierce SM (1999) *Namaqualand, A Succulent Desert*. Cape Town: Fernwood Press.

Cowling RM, Pressey RL, Lombard AT, Desmet PG, and Ellis AG (1999) From representation to persistence: Requirements for a sustainable system of conservation areas in the species-rich mediterranean-climate desert of southern Africa. *Diversity and Distributions* 5: 51–71.

Cowling RM, Rundel PW, Desmet PG, and Esler KJ (1998) Extraordinarily high regional-scale plant diversity in southern African arid lands: Subcontinental and global comparisons. *Diversity and Distributions* 4: 27–36.

Cowling RM, Rundel PW, and Esler KJ (1999) Namaqualand, South Africa – An overview of a unique winter-rainfall desert ecosystem. *Plant Ecology* 142: 3–21.

Desmet PG and Cowling RM (1999a) Biodiversity, habitat and range-size aspects of a flora from a winter rainfall desert in northwestern Namaqualand, South Africa. *Plant Ecology* 142: 23–33.

Desmet PG and Cowling RM (1999b) The climate of the Karoo – a functional approach. In: Dean WRJ and Milton SJ (eds.) *The Karoo: Ecological Patterns and Processes*. Cambridge: Cambridge University Press.

Desmet PG, Cowling RM, Ellis AG, and Pressey RL (2002) Integrating biosystematic data into conservation planning: Perspectives from the Succulent Karoo. *Systematic Biology* 51: 317–330.

Driver A, Desmet PG, Rouget M, Cowling RM, and Maze KE (2003) *Succulent Karoo Ecosystem Plan: Biodiversity Component Technical Report*. Cape Town: Cape Conservation Unit, Botanical Society of South Africa.

Eccles NS, Esler KJ, and Cowling RM (1999) Spatial pattern analysis in Namaqualand desert plant communities: Evidence for general positive interactions. *Plant Ecology* 142: 71–85.

Esler KJ, Rundel PW, and Cowling RM (1999) The Succulent Karoo in a global context: Plant structural and functional comparison with North American winter-rainfall deserts. In: Dean WRJ and Milton SJ (eds.) *The Karoo: Ecological Patterns and Processes*, pp. 303–313. Cambridge: Cambridge University Press.

Hilton-Taylor C (1996) Patterns and characteristics of the flora of the Succulent Karoo Biome, southern Africa. In: van der Maesen LJE, van der Burgt XM, and van Medenbach de Rooy JM (eds.) *The Biodiversity of African Plants*, pp. 58–72. Dordrecht: Kluwer.

Hoffman MT, Allsopp N, and Rohde RF (2007) Sustainable land use in Namaqualand, South Africa: Key issues in an interdisciplinary debate. *Journal of Arid Environments* 70: 561–569.

Hoffman MT and Rohde RF (2007) From pastoralism to tourism: The historical impact of changing land use practices in Namaqualand. *Journal of Arid Environments* 70: 641–658.

van Jaarsveld E (1987) The succulent riches of South Africa and Namibia. *Aloe* 24: 45–92.

Jürgens N (1991) A new approach to the Namib Region. I: Phytogeographic subdivision. *Vegetatio* 97: 21–38.

Klak C, Reeves G, and Hedderson T (2004) Unmatched tempo of evolution in southern African semi-desert ice plants. *Nature* 427: 63–65.

Linder HP (2008) Plant species radiations: Where, when, why? *Philosophical Transactions of the Royal Society B* 363: 3097–3105.

Lombard AT, Hilton-Taylor C, Rebelo AG, Pressey RL, and Cowling RM (1999) Reserve selection in the Succulent Karoo, South Africa: Coping with high compositional turnover. *Plant Ecology* 142: 35–55.

Manning JC and Goldblatt P (1996) The *Proseca peringuyei* (Diptera: Nemenstrinidae) pollination guild in southern Africa: Long-tongued flies and their tubular flowers. *Annals of the Missouri Botanical Garden* 83: 67–86.

Midgley GF, Hannah L, Millar D, Rutherford MC, and Powrie LW (2002) Assessing the vulnerability of species richness to anthropogenic change in a biodiversity hotspot. *Global Ecology and Biogeography* 11: 445–451.

Mittermeier RA, Gil PR, Hoffman M, *et al.* (2005) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Mexico City: CEMEX.

O'Farrell PJ, Reyers B, Le Maitre DC, *et al.* (2010) Multi-functional landscapes in semi arid environments: Implications for biodiversity and ecosystem services. *Landscape Ecology* 25: 1231–1246.

Procheş Ş, Cowling RM, and du Preez DR (2005) Patterns of geophyte diversity and storage organ size in the winter-rainfall region of southern Africa. *Diversity and Distributions* 11: 101–109.

van Rooyen MW, Theron GK, and Grobbelaar N (1990) Life form and dispersal spectra of the flora of Namaqualand. *Journal of Arid Environments* 19: 133–145.

Rundel PW, Cowling RM, Esler KJ, *et al.* (1995) Winter growth phenology and leaf orientation in the stem succulent Pachypodium namaquanum (Apocynaceae) in the Succulent Karoo of the Richtersveld, South Africa. *Oecologia* 101: 472–477.

Rundel PW, Esler KJ, and Cowling RM (1999) Ecological and phylogenetic patterns of carbon isotope discrimination in the winter rainfall flora of the Richtersveld, South Africa. *Plant Ecology* 142: 133–148.

Schmiedel U and Jürgens N (1999) Community structure on unusual habitat islands: Quartz fields in the Succulent Karoo, South Africa. *Plant Ecology* 142: 57–69.

Thompson M, Vlok J, Rouget M, Hoffman MT, Balmford A, and Cowling RM (2009) Mapping grazing-induced degradation in a semi-arid environment: A rapid and cost effective approach for assessment and monitoring. *Environmental Management* 43: 585–596.

Todd SW and Hoffman MT (1999) A fence-line contrast reveals effects of heavy grazing on plant diversity and community composition in Namaqualand, South Africa. *Plant Ecology* 142: 169–178.

Verboom GA, Archibald JK, Bakker FT, *et al.* (2008) Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both? *Molecular Phylogenetics and Evolution* 51: 44–53.

Vernon CJ (1999) Biogeography, endemism, and diversity of animals in the Karoo. In: Dean WRJ and Milton SJ (eds.) *The Karoo: Ecological Patterns and Processes*, pp. 57–85. Cambridge: Cambridge University Press.

Vlok J and Schutte-Vlok AL (2010) *Plants of the Klein Karoo*. Hatfield, South Africa: Umdaus Press.

van Wyck B and Smith G (2001) *Regions of Floristic Endemism in Southern Africa: A Review with Emphasis on Succulents*. Hatfield, South Africa: Umdaus Press.

Biodiversity, Origin of

MEJ Newman and GJ Eble, Santa Fe Institute, Santa Fe, NM, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 411–417, © 2001, Elsevier Inc.

Glossary

Cambrian explosion A period of about 25 million years during the Cambrian, when the diversity of multicellular life greatly increased in a relatively short time.

K-T event A notable large-scale extinction event occurring about 65 million years ago, at the boundary of the Cretaceous and the Tertiary, involving the extinction of about 70% of extant species, including the dinosaurs and many other land-dwelling vertebrates.

Monograph effect A term for the fact that the apparent existence of intensive evolutionary activity in a given interval actually reflects the fact that researchers have thoroughly studied a particular section of the fossil record which reveals this activity, while other such investigations elsewhere might just as well have produced the same number of previously unknown species.

Pull of the recent A term reflecting the fact that recently formed rock is more accessible and offers better preserved fossils; this may produce a bias toward the conclusion that diversity has increased toward the present time.

Taxon sorting The argument that in certain taxa, species undergo relatively rapid speciation and extinction, while in others the species speciate and die out relatively slowly; over time this difference will mean that the highly volatile families become extinct more quickly than the less volatile ones.

Life began on Earth about 3.5 billion years ago in the Archean Eon, probably with the appearance first of self-reproducing RNAs and later of prokaryotic single-celled organisms. No actual fossils date from this far back, and so our knowledge of Archean life is sketchy at best. Approximately 2.5 billion years ago the last stocks of elemental iron in the earth's crust were converted to oxides as a result of photosynthesis by cyanobacteria, and the planet's atmosphere changed from reducing to oxydizing, making oxygen-breathing life possible. The Proterozoic Eon which followed saw the first appearance of eukaryotes (organisms with nucleated cells and mitochondria), sexual reproduction, and multicellular organisms, although the precise timing and even the order of these innovations are disputed. The earliest firm evidence of metazoan multicellularity dates back to about 575 million years ago.

The end of the Proterozoic at the Vendian–Cambrian boundary about 545 million years ago marks the start of the Cambrian explosion, a period of approximately 25 million years during which, for unknown reasons, multicellular life underwent a period of extraordinary diversification, producing a multitude of new evolutionary lineages in a comparatively short time, including almost all of the major metazoan body plans seen today as well as many which have become extinct. The fossil record from the Cambrian onward—an interval known as the Phanerozoic Eon—is good by comparison with that of the Precambrian largely because multicellular organisms, whose fossils are easier to find and study than those of single-celled organisms, became numerous.

The Phanerozoic fossil record consists of about a quarter of a million known species. The vast majority of these are marine species, mostly invertebrates such as mollusks, brachiopods, corals, or foraminifera as well as sea-dwelling plants. Terrestrial (i.e., land-dwelling) organisms are far less well preserved because the depositional regimes under which fossils are formed are less reliable and uniform on land. Even for marine species, there are many biases in the fossil record, among which the following are some of the most important:

1. Deposition can vary greatly from one time to another. Even in the oceans, in which deposition is relatively reliable, there are periods during which preservation is excellent and others in which it is poor. A sudden period of poor preservation can even give the appearance of a large extinction event because the number of preserved organisms decreases substantially.
2. The "Pull of the Recent" is the name given to the greater accessibility of recently formed rock and the better preservation of fossils in these rocks. The Pull of the Recent implies that we have a more complete record of recent times than we do of times long past, and this could produce the misleading impression of an increase in diversity toward the present where none exists. Despite this bias, it is still believed (as discussed later) that there has been a real increase in biodiversity over the course of the Phanerozoic.
3. The "monograph effect" refers to the sudden apparent burst of new species resulting from the attention paid by a particularly zealous researcher or group of researchers to a particular section of the fossil record: The thorough investigation of a short interval can turn up many previously unknown species, making that interval appear to be one of especial evolutionary activity when in fact a similarly thorough investigation of another interval would have turned up just as many species. A similar effect is produced by the discovery of a particularly rich and well-preserved fossil bed dating from one particular time.
4. Although it is now possible to date fossils within a few million years with high confidence, more accurate dating in the fossil record is still problematic. Absolute dates are derived from slowly decaying radioactive isotopes, whereas relative dates of different fossils are derived from

stratigraphic evidence, both geological and from fiducial lineages (so-called "index fossils"). Some geographic areas and intervals of time can be dated with high accuracy—a million years or better—but others are not nearly as good.

Despite these biases, the overall pattern of life during the Phanerozoic survives the vagaries of the record and is now reasonably well understood.

Figure 1 is a graph of (marine) biodiversity as a function of time during the Phanerozoic. The horizontal axis measures time before the present, whereas the vertical one measures the number of known families of marine animals at a succession of times. Each point corresponds to one stratigraphic stage; stages are irregular intervals of time of average duration of about 7 million years which are based on widely accepted stratigraphic features. The data are taken from the compilation of marine families by Sepkoski (1992).

One of the most notable features of **Figure 1** is a substantial increase in the number of known families over the course of the Phanerozoic, particularly in the 250 million years since the Late Permian mass extinction. Overall, the number of known families increases by more than a factor of 10 from the earliest Cambrian to the present. In considering this increase, one must first ask whether it is genuine or an artifact of biases in the fossil record. As mentioned previously, there is certainly a bias occurring because of the better preservation and greater availability of fossils in more recent rocks. There is no doubt that some of the increase apparent in **Figure 1** is a result of this bias. On the other hand, the use of numbers of families as a measure of diversity, rather than species, tends at least partly to remove this bias since we need find only one species belonging to a certain family in order to establish the existence of that family at any given time. Therefore, even during periods of relatively poor species sampling, a family with a sufficiently large number of species can still contribute to our estimate of diversity. Furthermore, there are plausible biological reasons to believe in a real diversity increase: The increase seems to some extent to be the result of the evolution of organisms into new ways of life. The spread of life onto the land in the Silurian and Devonian is a particularly prominent example of this, but among marine organisms the trend is also clear. It seems reasonable to suppose that as life spreads into new environments, the total number of species (or families) which the planet can support will also increase.

Many models have been put forward for diversity increases in the fossil record. Benton (1995) suggested that the increase is an exponential one, indicating perhaps a stochastically constant rate of creation of new taxa from existing ones, with no detectable limit to ultimate levels of diversity. A more detailed analysis has been given by Sepkoski (1981), who suggested that the diversification of life occurred in three distinct phases, each typified by vigorous turnover within a particular subset of species or "fauna." The first phase, during the Cambrian, was dominated by groups in which the rate of origination of new species was relatively high, such as trilobites, inarticulate brachiopods, and monoplacophorans. The second phase, lasting from the Ordovician to the end of the Permian, was dominated by groups with intermediate origination rates, such as crinoids, articulate brachiopods, and cephalopods.

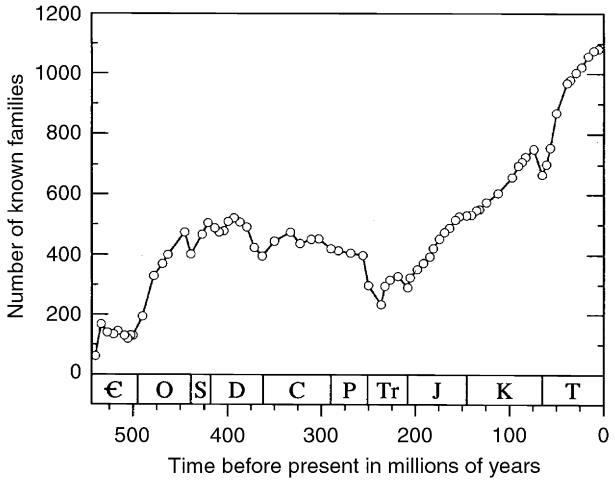


Figure 1 The number of known marine families alive over the time interval from the Cambrian to the present (data from acompilation by Sepkoski JJ Jr. (1992) A compendium of fossil marine animal families, 2nd edition. *Milwaukee Public Mus. Contrib. Biol. Geol.* 83).

Finally, there was a third phase, lasting from the beginning of the Triassic up until the Recent, which was dominated by bivalves, gastropods, echinoids, and marine vertebrates—groups which have relatively low rates of origination. Sepkoski (1984) suggested that within each of the three faunas diversification proceeds initially exponentially but eventually saturates, implying the existence of equilibrium "carrying capacities" for the global ecosystem. The superimposition of this "logistic" growth within each fauna gives rise to the observed pattern of near-monotonic diversity increase in **Figure 1**. A logistic diversification model was also explored by Courtillot and Gaudemer (1996) but without allusion to distinct faunas.

An important feature of the pattern of biodiversity shown in **Figure 1** which can be explained in terms of the logistic growth model is the apparent plateau in diversity throughout the greater part of the Paleozoic, from the Late Ordovician to the end of the Permian. (The same pattern is also present in genus-level data (Sepkoski, 1997), although there is in general more fluctuation in the numbers of genera so the plateau is not as clear.) One explanation for this plateau is that it represents the saturation phase of the logistic diversification of Sepkoski's second fauna, which continues mostly uninterrupted for approximately 200 million year until the Late Permian mass extinction. This is not the only possibility, however; it has also been suggested that the plateau is more the result of high extinction rates in the later Paleozoic than it is the result of low origination rates (Stanley, 1999). In both cases, however, the ultimate limit on diversity is presumed to be some effective carrying capacity of the ecosystem.

Another notable feature of **Figure 1** is a dip in diversity at several points, particularly at the ends of the Devonian (D), Permian (P), and Cretaceous (K) periods. These dips correspond to major mass extinctions that can be seen more clearly in **Figure 2**, which shows the percentage of extant families of marine animals becoming extinct in each stratigraphic stage of the Phanerozoic. The data are again taken from the compilation by Sepkoski (1992). We normally distinguish five major mass extinctions during the Phanerozoic. These "big five" are

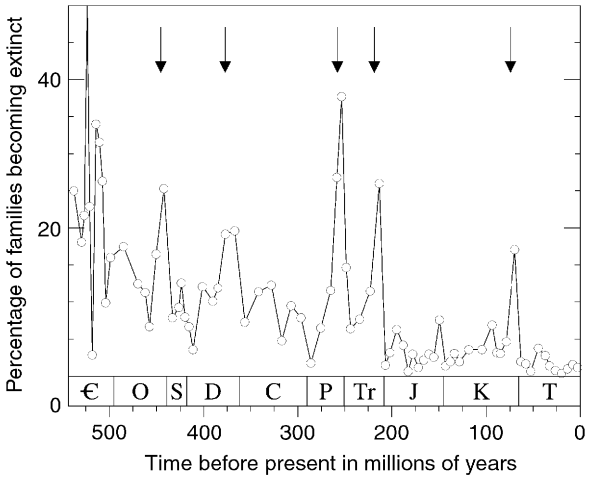


Figure 2 Estimated extinction of marine animals in families per stratigraphic stage since the Cambrian as a percentage of the total number of families in existence. The arrows indicate the positions of the "big five" mass extinction events discussed in the text.

marked with arrows in Figure 2. A sixth period of heightened extinction is visible in the Cambrian but is thought to be primarily an artifact of sampling biases rather than a real extinction event. Many smaller extinction peaks are also apparent in Figure 2 but are of secondary importance. The basic features of the big five events are as follows.

The end-Ordovician event about 440 million years ago appears to have occurred in two bursts, separated by about 1 million years, which between them wiped out about 85% of then living species. The event was confined to marine species since multicellular life had not yet colonized the land. Particularly affected were brachiopods, bivalves, echinoderms, bryozoans, and corals. The immediate cause of extinction appears to have been the continental drift of a significant land mass into the south polar region, causing a global temperature decrease, glaciation, and consequent lowering of the sea level, which destroyed species habitats around the continental shelves. The sea level rose again with the end of the glacial interval about 1 million years later and caused a second burst of extinction.

The Late Devonian extinction approximately 360 million years ago is complex and poorly understood. It is probably in fact composed of many separate events (as many as seven) spread over about 25 million years, including particularly notable extinctions at the ends of the Givetian, Frasnian, and Famennian stages. Overall, about 80% of living species died out in the Late Devonian. Particularly hard hit were corals, brachiopods, bryozoans, ammonoids, and fish. The causes of these extinctions are unclear. The leading theories suggest that changes in sea level and ocean anoxia, possibly triggered by global cooling or oceanic volcanism, were most likely responsible, although the impact of an extraterrestrial body such as a comet has also been considered.

The Late Permian extinction approximately 250 million years ago was the largest extinction event of all time, killing approximately 95% of marine species and about 70% of land-dwelling ones. Like the end-Ordovician event, it seems to have been composed of two bursts, separated in this case by an interval of about 10 million years, with the second being the larger of the two. Notable extinction happened again among brachiopods, ammonoids, and corals as well as gastropods, echinoderms, and, unusually, insects. Despite an enormous amount of research of the subject, the causes of the Late Permian event are still a subject of debate. It is clear, however, that the sea level rose during this period, levels of oxygen in the oceans were low, and carbon dioxide levels were high. There is some suggestion that a cometary impact may have been involved, or a shift in ocean circulation driven by climate change, or carbon dioxide and sulfur release following large-scale volcanic activity. The Late Permian event had a profound effect on the terrestrial ecosystem which is still being felt today, a quarter of a billion years later. A particularly notable example among marine faunas is that of the bivalves, a relatively minor group during the Paleozoic that took advantage of the ecological vacuum left by the extinction to establish a solid grip on shallow-water environments, leading to their dominance over the previously very successful brachiopods.

The end-Triassic extinction approximately 210 million years ago is probably the most poorly understood of the big five extinction events. It appears to have killed about 80% of species then living, either in one burst or possibly in two separated by about 20 million years. Major extinction is observed particularly among ammonoids, bivalves, gastropods, and brachiopods. Leading theories of the causes of the end-Triassic event are ocean anoxia, massive volcanism, or possibly a bolide impact.

The end-Cretaceous event, usually called the Cretaceous–Tertiary or K–T event, has attracted the most popular interest of any extinction because it saw the end of the dinosaurs, but it was in fact the smallest by quite a wide margin of all the big five. The K–T event appears to have been a single pulse of extinction approximately 65 million years ago which wiped out about 70% of all species then living. In addition to the dinosaurs, it extinguished many other land-dwelling vertebrates, especially large-bodied ones, along with large numbers of (marine) bivalves, gastropods, and foraminifera. The proximal cause of the K–T event was almost certainly the impact of a large comet or meteor near the present site of the town of Puerto Chicxulub on the Yucatán peninsula in eastern Mexico, with an associated drop in sea level and possibly short-term cooling or heating or acid rain. It has also been noted that average extinction rates were higher than the historical average for some time in advance of the K–T boundary, indicating that long-term environmental change may also have played a part.

Table 1 summarizes the fraction of families and species killed in each of the big five mass extinctions. In fact, the fraction of species killed cannot be measured directly from the fossil record since the preservation of individual species is too unreliable; the majority of species appear in only one stratigraphic stage so that no statistically significant estimate of their origination or extinction time can be made. The fraction of families killed can be calculated with reasonable certainty: Many families have reasonable coverage in the fossil record so that good estimates can be made of origination and extinction times. Estimates of species kill are then extrapolated from the family-level estimates by a technique known as reverse

Table 1 Extinction intensities at the family and species level for the big five mass extinctions of the Phanerozoic[a]

Extinction	*Family loss (observed) (%)*	*Species loss (estimated) (%)*
End Ordovician	26	84
Late Devonian	22	79
Late Permian	51	95
End Triassic	22	79
End Cretaceous	16	70

[a]Estimates of family extinction are obtained from directed analysis of the fossil record, whereas species loss is inferred using a statistical technique called "reverse rarefaction."

Source: Reproduced from Jablonski D (1995) Extinctions in the fossil record. In: Lawton JH and May RM (eds.) *Extinction Rates*. Oxford: Oxford Univ. Press.

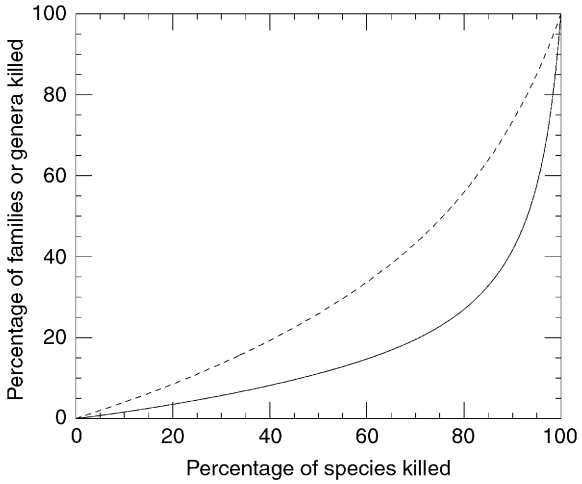


Figure 3 Rarefaction curves for the extinction of families (solid line) and genera (dashed line) calculated for echinoderms by Raup (1979). The curves indicate what percentage of families or genera are expected to become extinct during an event which kills a given percentage of species. Used in reverse, the curves also allow us to estimate what percentage of species became extinct in a particular event, given an observed percentage of family or genus kill.

rarefaction which works as follows. First, one denotes the probability of extinction of a particular species within a family as p. If the family contains only one species, then p is also the probability of extinction of the family. If it contains two species, however, then the probability of extinction of the family is p^2. In general, if the family contains n species, then the probability of its extinction is p^n. Then one averages this over many families to find the average probability of extinction of a family. In order to perform the average, we need to know the distribution of the sizes of families. Following Raup (1979), data from living species are usually used for this purpose. Calculating the average probability of family kill gives us a "rarefaction curve" (**Figure 3**) for the fraction of families or genera killed as a function of the fraction of species killed. "Reverse rarefaction" is the name given to the use of this curve "backwards" to estimate what fraction of species must have been killed in order to produce an observed level of family kill. This is the method which is used to calculate the data in **Table 1**.

Another overall feature of the extinction record shown in **Figure 2** is a decline in the level of background extinction (i.e., not mass extinction) during the Phanerozoic. Like the increase in standing diversity discussed earlier, it is important to establish whether this decline is a real effect or the result of some bias in the record. It is difficult to imagine that a decrease in the number of families becoming extinct could arise from better preservation toward the Recent. Rather, one would expect the reverse since, if extinction were actually approximately constant but more fossil families were preserved, one would expect an increase in apparent extinction. What is the explanation for the decrease? Many hypotheses have been put forward.

Flessa and Jablonski (1985) suggested that the decrease could be a result of increases in the average number of species per family toward the Recent. If families increase in size (as seems to have been the case), then the probability of extinction, which was denoted as p^n, will decrease as n, the number of component species, becomes larger. A related suggestion has been made by Boyajian (1986), who proposed that families "age" over time, meaning that they accumulate more species (on average) the longer they live and therefore become resistant to extinction. If the number of such aged families increased over time, this could explain the decrease in extinction rates. This explanation receives some support from the fossil record, which indicates that families have indeed tended to live longer in recent times than they did in the early Phanerozoic. The average lifetime of a family in the Paleozoic is estimated to be about 30 million years but increases to about 80 million years in the post-Paleozoic.

Gilinsky and Bambach (1987) and Gilinsky (1994) made a slightly different argument that the decrease is a "taxon sorting" effect. Suppose that taxa (e.g., families) vary in volatility. That is, the species within some of them undergo comparatively rapid speciation and extinction and turnover is high, whereas the species within others speciate and die out relatively slowly. For the highly volatile families, the time to extinction of the entire family, as a result of extinction of all its members, will be shorter than that for less volatile ones by precisely the ratio of the respective volatilities, meaning that the highly volatile families become extinct more quickly. Over the course of a long period of time, therefore, the families that remain will tend to be the less volatile, longer lived ones. Taxon sorting in effect produces a selection pressure at the family or other higher taxonomic level. Another possible explanation for the increase in the lifetimes of taxa and decrease in extinction rates is that species have really become more highly evolved toward the Recent and are better adapted to survive (Raup and Sepkoski, 1982). This explanation, however, is not a widely accepted one.

Although extinction has received the lion's share of paleontological attention in recent years, origination is equally important in the study of Phanerozoic biodiversity. We have already discussed origination in the context of the "three faunas" view of the diversification of life. **Figure 4** shows a detailed plot of the measured pattern of origination of families during the Phanerozoic. Many prominent features are visible. The major peaks in the Cambrian period are a reflection of the burst in metazoan diversity of the Cambrian explosion. Since the plot is one of origination rates as a fraction of standing diversity, however, the size of these peaks is partly just a result

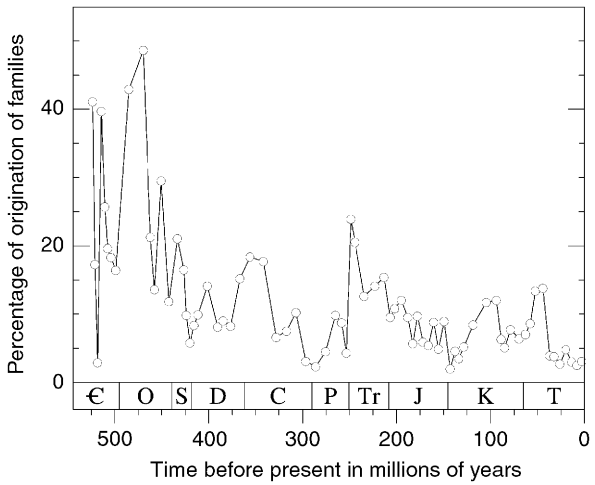


Figure 4 Estimated origination of marine animals in families per stratigraphic stage since the Cambrian as a percentage of the total number of families in existence.

of low Cambrian diversity (**Figure 1**). Another burst of origination is visible in the Ordovician. This burst was responsible for the corresponding major increase in diversity shown in **Figure 1**. After the Ordovician, origination is more subdued, although substantial surges did take place in the Triassic, Jurassic, Cretaceous, and during the Cenozoic.

On the whole, origination tends to vary less in magnitude and to be less episodic than extinction (Foote, 1994). Nonetheless, it appears to be bursts of origination, rather than declines in extinction rates, which account for most of the major evolutionary radiations. The basic features of the Ordovician, Jurassic–Cretaceous, and Cenozoic radiations are as follows.

Origination during the Ordovician occurred particularly among groups such as brachiopods, bryozoans, and mollusks, increasing the dominance of these groups while groups such as trilobites shrank in importance. Although phyla had basically ceased to be produced by the Ordovician, the period led to substantial production of body plans at the level of classes and orders.

The Jurassic–Cretaceous radiation produced substantial increases in the diversity of gastropods, bivalves, echinoids, and fishes—groups that were relatively minor during the Paleozoic. Ecological expansion in the form of the appearance of fundamentally new modes of life (Bambach, 1985) and the increasing occurrence of predation (Vermeij, 1977) have been advanced as important contributors to this pattern.

The Cenozoic radiation is certainly in part a product of the Pull of the Recent, but there are also genuine bursts of origination in the Paleocene and Eocene. In the marine realm, the diversity of veneroid bivalves, neogastropods, irregular echinoids, gymnolaemate bryozoans, and fishes increased substantially, whereas in the terrestrial realm the mammal and angiosperm radiations are particularly prominent.

Although mass extinctions are usually explained as the results of various kinds of exogenous stress on the ecosystem, such as climate change or bolide impact, originations seem to be more often the result of biological phenomena such as evolutionary innovations or changes in the structure of ecological assemblages. However, as a close inspection of **Figures 1**, **2**, and **4** reveals, origination is often intensified in the aftermath of mass extinction events, suggesting that physical disturbances do have an indirect effect on origination, probably through removal of incumbent species, through selectivity in extinction (Jablonski, 1989), and through particular patterns of rediversification of groups and repopulation of ecospaces during the recovery period following extinction. Recoveries typically take a few million years to run their course and reestablish previous levels of diversity and ecological heterogeneity (Erwin, 1998). Some recoveries do not begin immediately after an extinction event; there is a lag, perhaps as long as a few million years, probably the result of lingering environmental stresses and the proliferation of opportunistic species. A paradigmatic example of recovery is the Triassic recovery which followed the end Permian (**Figure 4**). A conspicuous spike in origination is observed beginning about 5 million years after the end of the extinction and continuing for about another 5 million years.

A long-term decline in origination over the Phanerozoic is as clear a feature of the fossil record as is the decline in extinction. Taxon sorting of the kind described earlier in the context of extinction is a plausible explanation (Gilinsky, 1994). Alternatively, the decline in origination may reflect a genuine decrease in the rate of evolutionary innovation as a result of either ecological or developmental restriction (Gilinsky and Bambach, 1987; Eble, 1999).

It is also instructive to examine origination rates at the level of higher taxa such as classes and orders. Origination at these levels reflects real innovations (to the extent that they are visible in the fossil record) rather than just sheer numbers of species. Indeed, there is no reason why the two need necessarily be correlated. During the Cambrian, for example, changes in total diversity (e.g., measured by the number of families) are modest, but it is in the Cambrian that one finds the most significant generation of new body plans—the "Cambrian explosion." This aspect of biodiversity having to do not with numbers of evolving entities but with their distinction is often referred to as "morphological diversity," "biodisparity," or simply "disparity" (Foote, 1997). Significant discrepancies between biodiversity and biodisparity can exist, such as faster morphological diversification early on in the evolution of clades, despite low levels of taxonomic diversity, and later deceleration despite continued increases in diversity. Further understanding of biodiversity in the fossil record may hinge on a better understanding of such patterns.

See also: Diversity, Organism Level. Eukaryotes, Origin of. Fossil Record. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of. Origin of Life, Theories of. Paleoecology. Species Diversity, Overview

References

Bambach R (1985) Classes and adaptive variety: The ecology of diversification in marine faunas through the Phanerozoic. In: Valentine JW (ed.) *Phanerozoic Diversity Patterns: Profiles in Macroevolution*. Princeton, NJ: Princeton Univ. Press.

Benton MJ (1995) Diversification and extinction in the history of life. *Science* 268: 52–58.
Boyajian GF (1986) Phanerozoic trends in background extinction: Consequence of an aging fauna. *Geology* 14: 955–958.
Courtillot V and Gaudemer Y (1996) Effects of mass extinctions of biodiversity. *Nature* 381: 146–148.
Eble GJ (1999) Originations: Land and sea compared. *Geobios* 32: 223–234.
Erwin DH (1998) The end and the beginning: Recoveries from mass extinctions. *Trends Ecol. Evol.* 13: 344–349.
Flessa KW and Jablonski D (1985) Declining Phanerozoic background extinction rates: Effect of taxonomic structure? *Nature* 313: 216–218.
Foote M (1994) Temporal variation in extinction risk and temporal scaling of extinction metrics. *Paleobiology* 20: 424–444.
Foote M (1997) The evolution of morphological diversity. *Annu. Rev. Ecol. Syst.* 28: 129–152.
Gilinsky NL (1994) Volatility and the Phanerozoic decline of background extinction intensity. *Paleobiology* 20: 445–458.
Gilinsky NL and Bambach RK (1987) Asymmetrical patterns of origination and extinction in higher taxa. *Paleobiology* 13: 427–445.
Jablonski D (1989) The biology of mass extinction: A palaeontological view. *Philos. Trans. R. Soc. B* 325: 357–368.
Jablonski D (1995) Extinctions in the fossil record. In: Lawton JH and May RM (eds.) *Extinction Rates*. Oxford: Oxford Univ. Press.
Raup DM (1979) Size of the Permo-Triassic bottleneck and its evolutionary implications. *Science* 206: 217–218.
Raup DM and Sepkoski JJ Jr. (1982) Mass extinctions in the marine fossil record. *Science* 215: 1501–1503.
Sepkoski JJ Jr. (1981) A factor analytic description of the Phanerozoic marine fossil record. *Paleobiology* 7: 246–267.
Sepkoski JJ Jr. (1984) A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* 10: 246–267.
Sepkoski JJ Jr. (1992) A compendium of fossil marine animal families, 2nd edition. *Milwaukee Public Mus. Contrib. Biol. Geol.* 83.
Sepkoski JJ Jr. (1997) Biodiversity: Past, present, and future. *J. Paleontol.* 71: 533–539.
Stanley SM (1999) The myth of the diversity plateau: High extinction rates, not ecological saturation, stifled the diversification of Paleozoic marine life. *Geol. Soc. Am. Meeting Abstr.* 31: A337.
Vermeij GJ (1977) The Mesozoic marine revolution: Evidence from snails, predators and grazers. *Paleobiology* 3: 245–258.

Biodiversity-Rich Countries

José Sarukhán, Universidad Nacional Autónoma de México, México
Rodolfo Dirzo, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Beta diversity Turnover of species within a small spatial scale resulting from highly differing ecological conditions.
Biogeographic regions Major areas of the world where large assemblages of species originated and that differ considerably from other regions.
Convention on Biodiversity (CBD) An international convention signed by 175 countries, which originated at the Earth Summit of Rio de Janeiro in 1992.
Cultural diversity Variation of races, languages, and habits including art that define a given culture.
Database Assemblage of information organized in tables with a logical structure. It is a fundamental tool used to gather, organize, and analyze biodiversity information.
Endemism Confinement of a species or other taxonomic groups to a given region, for example, an island or a country.
Megadiverse An area or a country possessing a much larger proportion of species than would be expected by its extent, latitudinal position, and other factors.
Parataxonomist Laymen and peasants trained to collect and carry out preliminary identification of botanical and zoological specimens. They are widely used by INBio in Costa Rica.
Phylum A major category, between kingdom and class, of taxonomic classification.
Primary production Amount of energy produced by photosynthetic organisms in a community.

Introduction

The fact that different countries harbor contrasting degrees of biological diversity has been recognized since the time of the great naturalists–travelers (including Darwin, Humboldt, Wallace, and others). More recently, the occurrence of biodiversity-rich countries has been described in more quantitative terms, under the notion of "megadiversity countries." Conservationist Russell Mittermeier developed it with an initial emphasis on tropical primates. Later, it was extended to all types of ecosystems and several groups of organisms (Mittermeier *et al.*, 1997). The concept of megadiversity countries is also related to that of centers of diversity – species-rich areas with a high concentration of endemic organisms. However, the centers of diversity are naturally defined spatial units, whereas megadiversity countries are geopolitical entities. Not surprisingly, therefore, there is no precise numerical coincidence between centers of diversity and megadiverse countries. For instance, at the global scale, the Plant Conservation Office of the International Union for the Conservation of Nature (http://www.iucn.org) recognizes 234 centers of plant diversity, whereas the number of megadiverse countries, from the botanical point of view, is only approximately 10. However, some spatial overlap is observed: Of the total number of plant diversity centers, close to 50% are located within these 10 megadiversity countries. That such concentrations of the planet's biodiversity are located in a handful of countries constitutes a matter of national, regional, and global importance.

Facets of Biodiversity in Biodiversity-Rich Countries

To examine the underlying causes for the concentration of biodiversity in specific countries, the authors first discuss the criteria on which the definition of megadiversity is based. Consideration of species richness is the predominant approach in ranking areas according to their biodiversity. However, in this article, and following Mittermeier *et al.* (1997) and Dirzo and Raven (2003), megadiversity is defined on the basis of six facets of biological diversity.

Species Richness

The authors consider species richness of vascular plants and of the four major groups of animals for which knowledge is better: mammals, birds, amphibians, and reptiles. Species richness in these groups of organisms, and plant species in particular, can be taken as correlates of other lesser-known organisms, and as a surrogate of total species richness. When available, species richness of other groups of organisms is also used.

Concentration of Endemisms

The concentration of endemisms refers to the percentage of endemic species present among a country's total resident species. This criterion is also based on the best-known groups of organisms: plants, mammals, birds, amphibians, and reptiles.

Ecological Heterogeneity and Diversity of Habitats

The diversity of habitats refers to the diversity of distinct ecological settings, usually expressed by the diversity of vegetation types. Closely related to this facet is the concept of beta diversity, the spatial turnover of species between nearby or adjacent habitats. Indeed, high beta diversity typically results from the heterogeneity of ecological conditions.

Presence of Tropical Forest Ecosystems

Given that tropical forests, particularly rain forests, are widely recognized as the most species-rich terrestrial ecosystem (Dirzo and Raven, 2003), their presence is used as another indicator of megadiversity.

Presence of Marine Ecosystems

The presence of marine ecosystems is included because they hold the greatest diversity of taxa of higher rank ("deep diversity"). In particular, marine ecosystems are rich in phyla of animals: 35 animal phyla are known from the sea, 14 of them being exclusively marine, whereas only 11 are terrestrial and only one exclusively so.

Cultural Diversity

The cultural diversity criterion addresses the diversity of ethnic groups present in a given country. Its relevance in diversity-rich countries stems from the strong association between biological diversity and cultural diversity. Different cultural dimensions in different countries have played a significant role in the modes in which the conjunction of biological diversity, including natural and domesticated biodiversity, is perceived, maintained, used, and valued by different ethnic groups.

Underlying Causes of Diversity in Biodiversity-Rich Countries

The predominant underlying causes that determine the existence of biodiversity-rich countries are latitudinal position, physical factors that increase primary production, historical (e.g., geological and tectonic) factors, environmental heterogeneity, and the presence of tropical rain forests. Most of the information regarding the relationship between these factors and biodiversity, however, considers species richness.

Latitudinal Location

Latitudinal location is not a determining factor of biodiversity per se. Rather, it is a proximal correlate of biodiversity: For almost all groups of organisms, species diversity increases toward equatorial latitudes. For example, on land, species richness and latitude are negatively correlated among mammals, birds, and trees (**Figure 1**).

In consequence, countries located at lower latitudes tend to be rich in species of many groups. **Table 1** shows the contrast between tropical and temperate countries in terms of the number of plant and mammal species. Regardless of size, the sample of countries located close to tropical latitudes has plant and mammal diversities that are about twice as large as those of the temperate countries sample. Although some temperate (e.g., US) or partly temperate (e.g., Australia) countries of large area reach values of species richness that are comparable to those of tropical countries, the tropical nations of this sample constitute a list of potential biodiversity-rich countries.

The pattern of decline in species richness with latitude in terrestrial ecosystems is paralleled in marine systems, although available information relates to other groups of organisms. For example, Arctic waters contain approximately 100 species of tunicates (sea squirts), temperate seas contain some 400 species, whereas waters at tropical latitudes have more than 600 species. Although data are not readily available for individual countries, the similarity of latitudinal patterns in terrestrial and marine systems suggests that a comparable trend exists on a per country basis, which again indicates that low-latitude countries are likely biodiversity-rich countries on the basis of marine species richness too. In sum, using the species richness criterion, biodiversity-rich countries are expected to be located in low-latitude regions of the world.

Physical Factors of the Environment Leading to Increased Primary Production

Consistently, primary production increases with precipitation, available solar energy, and nutrient availability (e.g., soil fertility). The highest terrestrial primary production is found in areas with high rainfall and year-round warm temperatures, whereas production decreases with lowered temperatures and the occurrence of frost. In marine settings, primary production increases with nutrient concentration, as is typical in some coastal zones, and it is lowest in open ocean, where nutrient

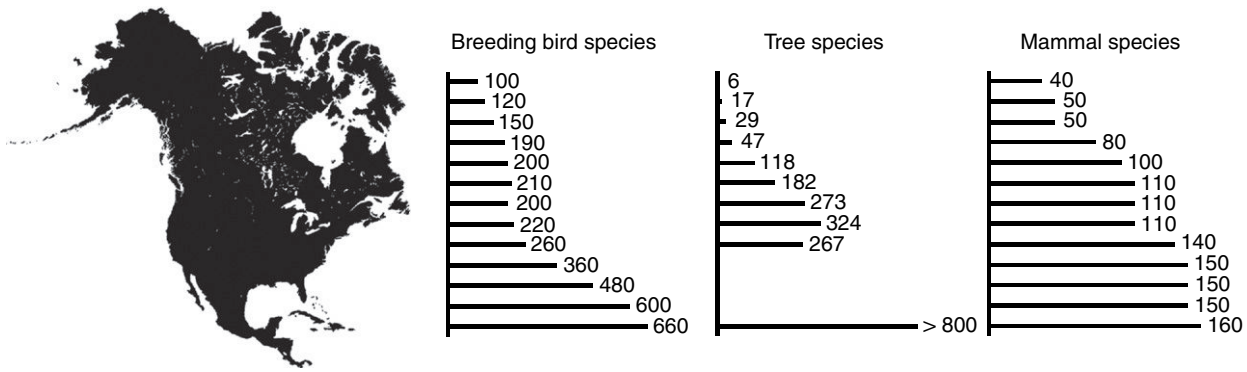


Figure 1 Latitudinal gradients of species richness for birds, trees, and mammals in North and Central America. Bars and numbers correspond to species richness according to latitudinal position on the map to the left. Reproduced with permission from Figure 4.5 of Meffe GK and Carroll CR (1997) *Principles of Conservation Biology*, 2nd edn. Sunderland, MA: Sinauer Associates Inc.

Table 1 Numbers of plant and mammal species in a selected sample of tropical and nontropical (mainly temperate) countries

Tropical	*Plants*	*Mammals*	*Temperate*	*Plants*	*Mammals*
Brazil	49,520–56,000	541–648	Argentina	10,000	338–374
Colombia	45,000–51,000	456	Australia	16,000–24,716	349–386
Congo Democratic Republic	11,000	430	Canada	2920	202
Ecuador	17,600–21,100	372–382	Egypt	2066	105
India	45,000	390–412	France	6000	123
Indonesia	37,000	670	Japan	4700	144
Malaysia	15,000	298	Morocco	7000	336
Mexico	30,000	450–535	South Africa	23,456	247
Peru	>20,000	467–508	UK	1550	77
Venezuela	15,000–21,000	363	USA	18,956	440

Source: Data derived from several sources, mainly National Reports for the Convention on Biological Diversity (CBD, 2009–2010), and IUCN Red List of Threatened Species (2010 Report, http://www.iucnredlist.org).

concentration becomes limiting. The relationship between productivity and biodiversity, however, is not always lineal. For example, in many ecosystems, species richness increases with the increment in primary production at low productivity levels, reaches a plateau at moderate levels of production, and then decreases at high levels of primary production. Thus, areas with environmental factors leading to increased primary production have, up to a certain level, a greater resource base that can support a wider range of species and typically have high species richness. As a result, countries with latitudinal and elevation ranges that include tropical and lowland regions, as well as coastal areas, where nutrient-rich waters upwell, tend to be megadiverse. Brazil, Peru, Colombia, and Mexico are clear examples of countries where such environmental conditions coincide.

Historical Factors

Historical factors are also critical in determining species richness and concentration of endemisms. Over a geological timescale, the continued rise in species richness is explained by the "provincialization" that accompanied the breakup of Pangaea and, later, Gondwanaland. Following their separation, evolution on each of the newly produced "provinces" (continents) produced many new taxa that were continental endemics and, indeed, much of the increase in species numbers of the Cenozoic era appears to be due to such continental isolation. As a consequence, several countries of South America (Brazil and Colombia), Africa (Democratic Republic of Congo), and Southeast Asia (Indonesia), are extremely rich in species, but they share few species in common because of their many millions of years of separation.

In general, areas that are geologically older have higher species richness than similar but younger areas, older areas having had more time to receive species dispersing from other regions of the world, and more time for existing taxa to undergo adaptive speciation in response to local conditions. The effect of geological age on species richness is well exemplified among marine ecosystems. The species richness of reef-forming corals is several times greater in the Indian Ocean and western Pacific than in the Caribbean Sea and adjacent Atlantic Ocean, which are considerably younger than the former. Similarly, mangrove ecosystems typically contain a much higher number of arboreal mangrove species in the Old World than in the Neotropics. Marine species richness of countries associated with these older seas is therefore remarkably high, thus contributing to the existence of megadiverse countries in some of these areas. Australia is a clear example: Its Great Barrier Reef contains approximately 8% of the world's fish species in just 0.1% of the global ocean surface area.

Regions rich in endemism tend to have experienced major events that caused fragmentation of the ranges of many species. Such fragmentation events leading to geographic isolation include continental drift, mountain building, and sea level changes. Subsequent to their isolation, taxa may undergo evolutionary radiation in a given locality. For example, isolation and subsequent evolutionary radiation due to continental drift have been fundamental to the generation of a high degree of endemism in the biota of Australia, Papua New Guinea, and Madagascar. Mountain building that led to isolation, explosive radiation, and a high accumulation of endemism is well represented in biodiversity-rich countries such as Colombia, Peru, and Mexico. Likewise, the remarkable diversity of fishes and other groups of organisms in large tropical lakes and rivers is due to rapid evolutionary radiation in such isolated and usually productive ecosystems, as exemplified by those aquatic ecosystems in Brazil, Colombia, and Peru.

As expected, islands often have high proportions of endemic taxa. Australia, a continent-island of remarkable size, is one of the world's leaders in the endemism of mammals, birds, reptiles, and amphibians. Nevertheless, because islands frequently have a relatively depauperate biota, such high levels of endemism do not necessarily coincide with high richness of species, except on those islands where other driving factors are of exceptional importance, such as the large tropical archipelago of Indonesia.

Another aspect related to history is the distribution of floristic and faunistic assemblages of different geographic affinity, which define the so-called biogeographic regions or provinces, and reflect shared evolutionary histories and ecologies of species. The relevance of this aspect of biogeography lies in the fact that a few countries of the world are located at the meeting ground of adjacent biogeographic provinces. Three remarkable examples of this are the Nearctic and Neotropical regions that meet in Mexico, the Oriental and Palearctic regions in China, and the Indomalaysian and

Australoasiatic regions in Indonesia. The meeting points of biotas of different biogeographic affinity are so striking that, for example, on a single Mexican cloud forest site one can find gum trees and oaks, together with tropical trees, epiphytic cacti, and tree ferns among many other biogeographically different plant lineages.

A final aspect of historical nature is the magnitude of the coastal/marine contour in relation to land area. Overall, when this ratio is high, some components of biodiversity, including ecological diversity and species richness, tend to be high. Though coastlines and the presence of seas are common to many countries, the ratio is particularly high in islands and some countries that are rich in biodiversity (**Table 2**). Notable examples include Indonesia (an archipelago of several hundred islands), Australia, Madagascar, and Mexico. **Table 2** shows that most of the megadiversity countries (cf. Mittermeier *et al.*, 1997) have coastline-to-area ratios that are relatively similar, with the exception of the Philippines and Indonesia, two countries constituted by archipelagoes, in which the ratios are exceptionally high. Mexico, in addition, includes extensive coastlines on two major oceans, the Pacific and Atlantic. Colombia is in a similar situation, although the ratio is lower. Mexico also has a sea of its own, the Sea of Cortez, all of which has resulted in significant levels of marine species richness, reflected by the fact that 51% of the marine mammal species of the world and close to 80% of all sea turtle species are present in Mexican coastal waters. It is worth noting that no land-locked country, regardless of its size, is a significant megadiversity country.

Ecological Heterogeneity

Species richness and the generation of endemism tend to be greater where there is significant ecological heterogeneity that allows for genetic isolation, local adaptation, and speciation to take place. Such situations can occur because of the existence of a series of isolated mountain peaks, large valleys, or drainage systems that become isolated and separated into smaller systems or areas that, being geologically heterogeneous and complex, produce a variety of soil conditions with well-defined boundaries between them. In addition to the promotion of speciation, these situations may provide the necessary ecological heterogeneity on which a variety of communities can develop. This is reflected in two important – and related – aspects of the definition of biodiversity-rich countries: The occurrence of an ample diversity of vegetation types or ecosystems and the occurrence of a high turnover of species between adjacent localities, leading to a spatially structured diversity of species known as beta diversity.

Regarding the diversity of ecosystems and vegetation types, Colombia, Mexico, Indonesia, Peru, and India are remarkable in that their territories range from savannas and arid and semiarid ecosystems, to evergreen tropical ecosystems, and from coastal vegetation types (e.g., mangroves) to alpine and subalpine ecosystems. Brazil, in addition to its great predominance of tropical rain forests, includes other tropical vegetation types such as the semiarid cerrado and caatinga – two types of tropical dry forest.

Regarding species turnover, available data on individual countries are extremely limited. However, it can be inferred from the diversity of vegetation types and the magnitude of environmental heterogeneity that the countries that are prolific in these attributes are likely to have high levels of beta diversity. One illustrative case is Mexico, where tropical dry forests expand along a geographic gradient from approximately 15° to 26° N latitude in a variety of ecological conditions (Trejo and Dirzo, 2002). At a given locality within this range, the number of arboreal species (with 1.0 cm or greater diameter at breast height) per 0.1 ha is remarkably high, with an average of 75 ($N=20$ sites). However, the average proportion of species shared among sites is only 9%, implying that the turnover of tree species in biodiversity-rich countries is very high and that even a given vegetation type is composed of a series of distinct assemblages of species.

Presence of Tropical Rain Forests

Several of the determining factors of species richness and endemism discussed above are coincident in tropical rain forests and coral reefs (see **Table 2**). In particular, the presence of

Table 2 Presence of coral reefs and the relationship between a country's surface area and coastline for the twelve most important megadiversity countries defined by Mittermeier *et al.* (1997)

Country	*Presence of coral reefs*	*Area (km^2)*	*Coastline (km)*	*Ratio coastline/area*
India	Yes	3,287,590	7000	0.002
Philippines	No?	300,000	36,289	0.120
Madagascar	Yes	587,041	4828	0.008
Australia	Yes	7,741,220	25,760	0.003
Colombia	Yes	1,138,914	3208	0.002
Brazil	Yes	8,511,965	7491	0.0008
Mexico	Yes	1,958,201	11,592.77	0.005
Peru	No	1,285,215.6	2414	0.001
Venezuela	Yes	916,445	2800	0.003
Ecuador	Very small	275,800	2237	0.008
China	No	9,596,961	14,500	0.001
Indonesia	Yes	1,919,442	54,716	0.028

Source: Defined by Mittermeier RA, Goettsch-Mittermeier C, and Robles-Gil P (1997) *Megadiversidad: Los países Biológicamente más ricos del Mundo*. Mexico City, Mexico: Agrupación Sierra Madre, SC.

tropical rain forests is a significant indicator of the potential for a country to be in the biodiversity-rich group, given the compelling evidence that these forests are the most species rich compared to any other ecosystems (Dirzo and Raven, 2003).

Top Biodiversity-Rich Countries

The previous description gives some indication of the most likely nations to be considered biodiversity-rich countries. In their analysis of megadiversity countries, Mittermeier *et al.* (1997) examined 17 countries that collectively possess 66–75% of the planetary biodiversity (terrestrial, marine, and freshwater). However, not all of these examined countries comply with all the facets of biodiversity discussed here: Although some countries are prominent in one or more of those facets, they are not prominent in, or lack, some of them. A more quantitative analysis by the same authors led to a shorter list of countries based on a hierarchical scoring of the five most prominent values (scored from 1 to 5) for species richness and magnitude of endemism of plants, mammals, birds, reptiles, and amphibians. Additionally, even though information is much more limited, consideration was also given to freshwater fishes, butterflies, and tiger beetles (family Cicindelidae). Thus, a country that had the highest values of any of these components was assigned a score of 5, the second highest was given the score of 4, and so on. The total sum of these scores gave a value that could then be used to compare the different countries. For example, Brazil was highest in four components: species richness of plants, mammals, and freshwater fishes, and plant endemism (i.e., $5+5+5+5=20$). It also had three scores of 4 (species richness of amphibians and butterflies, and endemism of amphibians $=12$), four scores of 3 (species richness of birds and tiger beetles, and bird and tiger beetle endemism $=12$), one score of 2 (bird endemism), and two scores of 1 (species richness and endemism of reptiles $=2$). The total sum for Brazil is therefore 48. The corresponding values for the twelve most biodiversity-rich countries are shown in **Figure 2**.

This synthetic analysis confirms the suggestions of the previous sections and can be summarized as follows. Brazil is the most prominent biodiversity-rich country, followed by Indonesia and Colombia. At a considerable distance, Australia and Mexico follow in the fourth and fifth positions, respectively. The insular character of Madagascar highlights its outstanding endemism and places this country in the overall sixth place, close to Mexico. The other countries in **Figure 2** are notable for their species richness (Peru, seventh position; China, eighth position; Ecuador, eleventh position; and Venezuela, twelfth position) or their degree of endemism (Philippines, ninth position). India is remarkable in both species richness and endemism, although it ranks in the tenth position. All seven of the most prominent diversity-rich countries share the following common biological attributes: (1) They have tropical rain forest ecosystems within their territories. (2) They also have marine ecosystems and, to a varying degree, a high coast-to-land ratio (see **Table 2**). (3) Their ecosystem diversity is considerable and therefore their beta diversity is expected to be high. (4) They have a very rich diversity of cultures. (5) They have considerable levels of threat to their biodiversity and cultural diversity, and (with the exception of Australia) all of them are developing countries. In particular, recent information on species of plants, birds, mammals, amphibians, and reptiles under several categories of extinction risk suggest that no less than 50% of the seriously threatened biological diversity at the global level is concentrated in the top megadiversity countries, as the authors discuss in the next section.

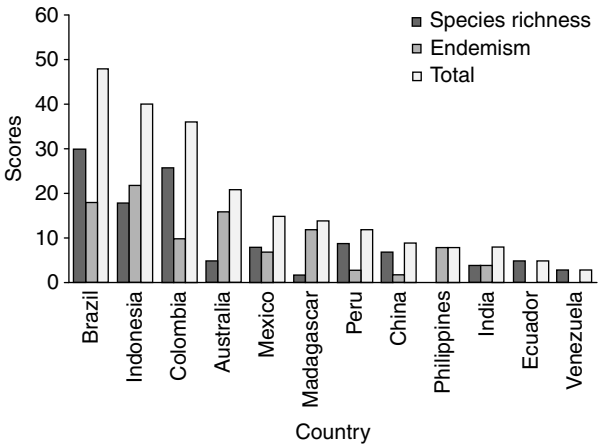


Figure 2 The ranking of the twelve countries with highest biodiversity based on scores of species richness and endemism. Derived from Mittermeier RA, Goettsch-Mittermeier C, and Robles-Gil P (1997) *Megadiversidad: Los países Biológicamente más ricos del Mundo*. Mexico City, Mexico: Agrupación Sierra Madre, SC, with permission from Cemex. See text for details.

Threats to Biodiversity in Biodiversity-Rich Countries

These countries are experiencing significant alterations to their natural ecosystems, and such degradation is a serious threat to their biological resources. The main current threats to biodiversity in these countries are the current levels and patterns of land use that lead to severe deforestation, habitat fragmentation, and species overexploitation. Most of these countries are expected to be impacted, as well, by climatic change and the synergistic interactions of climatic change and other anthropogenic threats. An analysis of climate change and impacts on biodiversity-rich countries is out of the scope of this article, but given their immediacy, the authors focus on the effects of land use change and overexploitation.

Deforestation and Habitat Fragmentation

A commonly used metric of habitat destruction is the rate of deforestation. **Table 3** presents FAO's compiled statistics on deforestation rates for the periods 1980–89 and 2000–10 (Food and Agriculture Organization (FAO), 2011). Rates of annual deforestation for the 12 biodiversity-rich countries in the previous decade ranged from 0.4% (Peru) to 2.3% (India and Ecuador). Only four countries had deforestation rates below 1% per year. Interestingly, in the period 2000–10, all biodiversity-rich countries showed a decline, in some cases

Table 3 Demographic, social, physical, and resource utilization parameters[a], including deforestation in megadiverse countries[b]

Country	*Population in 2010 ($\cdot 10^6$)*	*Growth Rate (1993–2000)*	*Rural population (% of total in 1993)*	*Area ($ha \cdot 10^6$)*	*Forested area (% of total)*	*Cultivated area (% of total)*	*Area under irrigation (% of total)*	*Firewood use ($m^3 \cdot 10^6$)*	*Yearly deforest. rate (%, 1980–1989; 2000–2010)*
Brazil	195.4	0.7	23	851.2	57.3	4.9	6.7	194.27	0.7; 0.5
Indonesia	232.5	1.0	67	190.5	58.7	9.9	24.3	149.06	0.8; 0.5
Colombia	46.3	1.3	28	113.9	43.9	3.4	13.5	17.22	1.7; 0.2
Australia	21.5	1.0	10.9	771.3	18.8	6	4.6	–	–; 0.4
Mexico	110.6	0.9	22.2	195.8	24.9	11.8	26.3	1.3	1.3; 0.3
Ecuador	13.8	1.2	43	28.3	55	5.7	34.1	2.3	2.3; 1.8
Madagascar	17.3	3.2	74	58.7	39.5	4.4	42.1	1.2	1.2; 0.4
Peru	29.5	1.9	29	128.5	66	2.6	37.6	0.4	0.4; 0.2
Venezuela	29.0	1.5	8	91.2	32.9	3.5	5.9	0.7	0.7; 0.6
China	1354.1	0.6	71	956,100	13.5	9.6	53.6	220.6	–; 1.6
India	1214.5	1.3	74	328,759	20.8	50.5	28.9	262.8	2.3; 0.5
Philippines	93.6	1.7	48	30	45.3	18.4	28.6	35.98	1.5; 0.7

[a]*Source*: UNDP Human Development Reports, 1996, 2010. New York.
[b]FAO (2011) Annual report.

considerable, of the yearly deforestation rates. Even though these recent proportional figures seem low, for most countries these values represent a significant annual loss of absolute area and, potentially, a disturbance of long-term consequences, especially in light of the synergies between land use change and climatic change. In addition, rates of deforestation do not reflect the fact that the remaining areas of habitat frequently are left in the form of an archipelago of forest islands immersed in a matrix of transformed habitat. Forest fragmentation has been recognized as a significant factor leading to the local extinction of species, reduction of genetic variability, and disruption of several ecosystem processes and services.

Consistent with this decline in deforestation in the biodiversity-rich countries, worldwide deforestation has declined from *ca*. 10 million ha $year^{-1}$ in the period 1980–99 to 7.13 million ha $year^{-1}$ in the period 2000–10. Yet, we emphasize that habitat loss is still very large, particularly in some biodiversity-rich countries (e.g., Brazil). Habitat loss is recognized as the main driver of biological extinction, and botanist Peter Raven estimates that an average of 50,000 species may be lost per year during the next several decades, of which only some 7000 will have been recognized and named.

Overexploitation of Species

The overexploitation of species is much more difficult to document and no reliable, compiled statistics are available. Overexploitation of the majority of targeted marine fishes is very well documented, but this generally occurs in the "commons" of international waters. However, some isolated reports indicate that overexploitation is a significant threat to biodiversity in biodiversity-rich countries. For example, in the Brazilian Amazon, studies by zoologist Carlos Peres indicate that hunting of vertebrates (mammals and birds) can kill approximately 23 million animals per year. In summary, the great biological richness present in the biodiversity-rich countries is seriously threatened and urgent measures are needed to deal with this problem.

Socio-Demographic Characteristics and their Relation to Biodiversity

Of all the issues related to biological diversity, there are very few whose study makes any sense outside of the context of its relation with human populations. Indeed, human populations – their very existence, cultural evolution, size, and distributions – have been and still are intricately dependent and interactive with the biological diversity available to them in the regions where they live. Nowhere are these relations between population, cultures, and biological diversity more evident than in the majority of the biodiversity-rich countries, where the richest biological diversity is present, and where many of the ancient and most diverse human cultures and civilizations emerged. In many biodiversity-rich countries, indigenous groups comprise a high percentage of the population, commonly ranging from 10% to nearly 80% in the case of Papua New Guinea.

A measure of the cultural diversity of a country may be the number of languages spoken by its people. A total of 6700 languages are spoken in the world, and more than half are found in the 15 most biologically diverse countries (e.g., Papua New Guinea alone has more than 800 languages). This relationship between biological and cultural diversity – biocultural diversity – is explored by Víctor M. Toledo elsewhere in this Encyclopedia (Indigenous Peoples, Biodiversity and). Usually such a relationship has resulted in cultural traits and land use practices that have increased biological diversity, mostly through the development of new, cultivated species and their wild relatives. As illustrated in **Figure 3**, countries of both high biological and cultural diversity have generally been defined as centers of origin of crops, of management of wild biodiversity, and areas of development of agricultural technologies.

Rural Populations and Territorial Distribution

As **Table 3** shows, in stark contrast with economically developed regions of the world, most biodiversity-rich countries have increased their populations, and have growth rates that

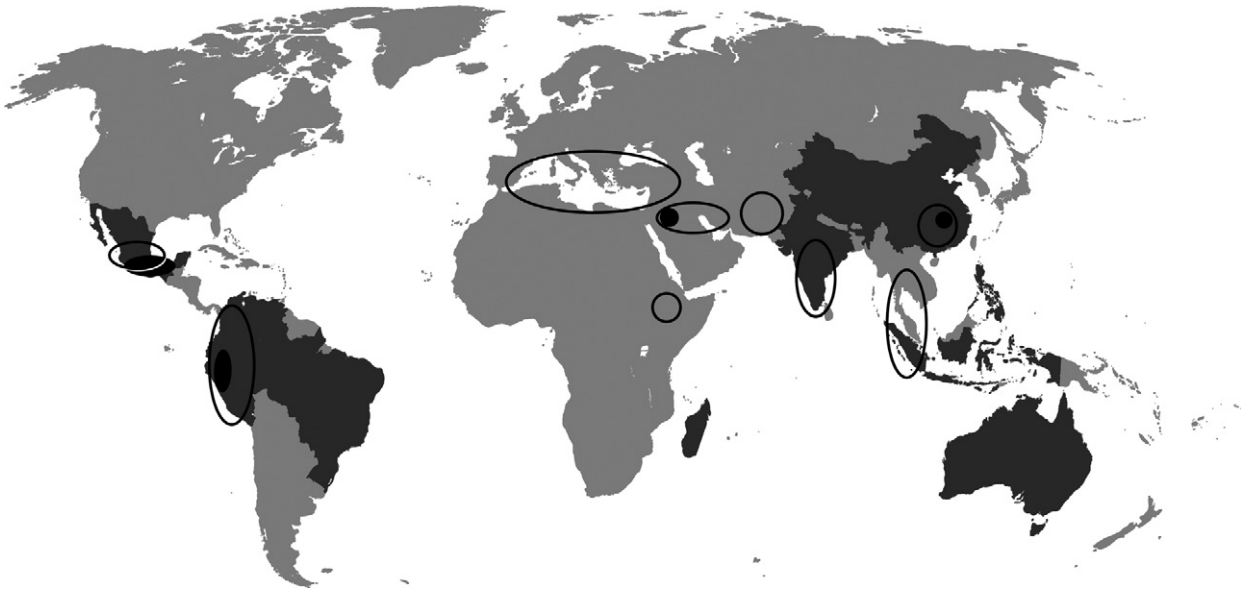


Figure 3 Distribution of the twelve biodiversity-rich countries (shown in black), the centers of origin of crops (open circles), and centers of agricultural development (filled circles).

range from 0.6 (China) to 3.2 (Madagascar). They also have a considerable proportion of their total population living in rural areas (**Table 3**). Consistent with their relatively high coverage of forested area (with the notable exception of China; **Table 3**), these groups are highly dependent on biodiversity (e.g., firewood use; **Table 3**), and are usually deeply knowledgeable of its uses and ecological traits. Additionally, in many cases a large proportion of these peoples live in small villages, most of which are distributed within the high biodiversity regions in each country. Because of their close relationship with the surrounding environment and resources, most of these communities are intimately dependent on the conditions of the ecosystems in which they live for their quality of life and the maintenance of their well-being. They have been termed as "ecosystem people."

Land Ownership by Indigenous Groups

The patterns of resource use and land ownership practiced by indigenous groups are normally the result of well-established traditions, forms of social organization, and consensus building. Communal land ownership is common, and it can take a variety of forms in different countries. Frequently, indigenous groups own or control extensive regions of natural forests and other ecosystems in these countries. For example, the indigenous groups in the Amazon basin possess more than 100 million ha of forested land. In Papua New Guinea, the many cultural groups that live therein own close to 97% of the national territory. In Mexico, nearly 75% of all the areas listed as priorities for conservation are under communal ownership ("ejidos"), mostly by indigenous or rural groups.

There is an abundant literature describing the benefits and disadvantages of communal land ownership in relation to the rational management and conservation of natural resources, particularly in biologically rich countries. After reviewing much of this literature and visiting the sites and talking to people, it is apparent to the authors that when a strong and healthy system of internal governance exists, allowing participation of all interested parties, the community uses and conserves its resources in an admirable fashion (see an account of examples in Bray *et al.*, 2005). When those internal structures of governance break down for whatever reason, the communal system of land ownership may become as destructive of its resources as any other system.

Because of the often large indigenous population in an area, their widespread territorial distribution, and the patterns of land ownership, any program of biodiversity use and conservation in most of these countries must take these stakeholders into account. Such inclusion should not be undertaken as an unavoidable necessity, but rather as a wise approach to any conservation program. In addition to being the owners of the areas whose resources and services conservationists and governments wish to conserve, these indigenous groups typically possess an enormous wealth of valuable and ecologically informed knowledge, which is often the result of a long tradition of an orally transmitted culture from one generation to the next, and that makes them truly valuable partners in conservation, management, or restoration programs.

Although a great deal of traditional ecological knowledge exists among the indigenous groups of biodiversity-rich countries, the long-term sustainable use of many of those natural resources remains uncertain and difficult. This problem is caused by a number of factors, including rising pressure from growing populations to acquire land, increased demands for food production, and the actions of economic models and interests, often of foreign origin, that are likely to degrade or destroy biologically rich ecosystems.

Levels of Scientific Development

Another characteristic that is (deplorably) common to the majority of biodiversity-rich countries (with the sole exception

of Australia) is that they have a relatively incipient scientific infrastructure, which in many instances is reflected in the very small number of scientists (taxonomists, botanists, ecologists, etc.) that study their natural resources.

Although it is difficult to ascertain to what extent this is the result of the history of colonization – another important characteristic shared by all biodiversity-rich countries – they have in common that they have been dominated by a European country for diverse lengths of time, from a few decades to several centuries. In these cases, an appallingly poor effort was made by the European countries to promote the technical and scientific development of the colonies. Additionally, the subsequent development of these countries was characterized by a neglect of the traditional knowledge on natural resource management of their ethnic communities.

Very often the most important collections of flora and fauna of these countries are found in institutions, herbaria, and museums located in developed countries. This can play a significant role in limiting scientific training and consequently the national capacity to assess and manage natural resources. As described earlier, a great proportion of the planet's terrestrial biodiversity is located in tropical forests, and in general within countries that often lack the necessary scientific and institutional infrastructure. The authors believe that this is a major issue that must be addressed when considering the future of the conservation of the world's biological diversity. A discussion of this point is presented in the remainder of this article.

Human and Institutional Capacity for Biodiversity Knowledge and Conservation

With a few exceptions, a shared characteristic of the biodiversity-rich countries is that a very large proportion of the scientific collections that represent their biodiversity wealth are deposited in institutions located in diversity-poor but economically rich nations. Most of the vast collections of the Earth's biota (some 3 billion specimens, excluding microorganisms) are located in the largest natural history museums and herbaria of the developed world, resulting from expeditions supported by developed countries during many decades, or even centuries. Many of these collections represent a heritage of significant historical value, since they contain biological information of regions that have been seriously altered and that presently have lost much of their original floras and faunas.

Much of the reason why these collections are not located in their "resident" countries is that only approximately 6% of the world's scientists live in the countries that house close to 75% of the planet's biodiversity. Different estimations coincide in calculating the number of taxonomists in the world to be approximately 7000. If a similar 6% ratio holds for taxonomists, this means that the biodiversity-rich countries have fewer than 500 taxonomists. We consider this to be an underestimation, and that there must be at least twice that number in the biodiversity-rich countries. However, even with this higher figure, this represents an inadequately small number of specialists who are able to adequately study an enormously rich biota.

Comprehensive and scientifically sound biological collections, together with the necessary taxonomic expertise that must go hand-in-hand with the collections, are a fundamental need to advance the ability of a country to know, understand, manage, and preserve its biological heritage. This was recognized at a high official level by the 175 signatories of the Convention on Biodiversity (CBD) during their fourth Conference of the Parties (COP) in Bratislava, Czech Republic, in May of 1998, who concluded that there exists what they called the "taxonomic impediment." This "impediment" results from the severe limitations of institutional and human infrastructure that affect the majority of biodiversity-rich countries. The very small number of taxonomists, systematists, and ecologists, as well as the absence, or very limited number, of zoological museums and herbaria that have reliable and long-term institutional support in most biodiversity-rich countries illustrates these limitations.

Table 4 is a compilation of the herbaria (with the number of specimens they contain), botanical gardens (with the number of taxa reported under cultivation), and zoological collections in the 12 most biologically diverse countries. It is clear that the majority have a serious shortage of scientific collections of their floras and faunas, with the notable exception of Australia and to a lesser extent India, China, and Indonesia in Asia and Brazil, Mexico, and Colombia in the Americas.

It is worth noting that for developed countries of the temperate regions the situation is dramatically different. The UK, France, and the US, for example, have between 10 and 100 times more herbaria and specimens housed in them than any of the 12 biodiversity-rich countries. The comparison is almost the same for zoological collections and botanical gardens. Obviously, only a fraction of those specimens belongs to their national floras and faunas, which makes them, together with a few other European countries like the Netherlands and Germany, the most important repositories of the biological information of the majority of biodiversity-rich countries.

At the Bratislava COP meeting, an agreement was reached on the imperative need to remove this "impediment" with the aim of fulfilling the objectives set by the CBD within a shorter period of time. However, the training of personnel at a high level, as is required for specialist taxonomists and systematists, and the extensive development of scientific collections (which in turn requires long-term institutional support) are complex, lengthy, and expensive processes. Most of the largest of such institutions in the world are the result of a tradition of many decades, and even centuries, of sustained institutional effort and economic support. This does not mean that those countries that do not have, or that need to develop more, such human and physical infrastructures should abandon any attempts to embark themselves in this process. Yet it seems to be far more effective to simultaneously dedicate efforts to develop a strong capacity to acquire, systematize, interpret, and utilize the biological information already existing in the many local and foreign museums and herbaria. Examples of such efforts, although not abundant, demonstrate different strategies in how to bridge the gap of knowledge about the biological diversity of a country. Some of these examples are presented next.

Table 4 Institutional infrastructure related to biodiversity in megadiverse countries

Country	*Number of herbaria*	*Accessions in herbaria (10^6)*	*Number of botanical gardens*	*Number of taxa in botanical gardens (10^3)*	*Number of zoological collections*
Brazil	88	3.2	23	17	18
Indonesia	6	1.67	5	70	3
Colombia	23	0.66	13	3	8
Australia	38	5.32	63	100	4
Mexico	52	2.3	35	7	25
Madagascar	2	0.08	2	5	2
Peru	11	0.45	6	–	3
China	306	15.5	68	27	7
Philippines	9	0.29	9	17	6
India	52	3.7	68	85	27
Ecuador	10	0.26	2	0.5	3
Venezuela	15	0.63	7	1	15

Source: Reproduced with permission from Heywood VH and Watson RT (eds.) (1995) *Global Biodiversity Assessment*. Cambridge, UK: Cambridge University Press.

Case Studies of Managing Biological Information in Biodiversity-Rich Countries

As the authors have mentioned, by far the majority of the scientific information about the flora, fauna, and microorganisms that live in a region or country is housed in the scientific collections of plants and animals. Such information is complemented to a certain degree by works such as written floras and faunas, updated checklists of organisms, and detailed ecological accounts of groups of plants or animals. However, only the scientific collections contain the wealth of information that covers the historical dimension of which organisms used to live in what now are profoundly disturbed areas. A second dimension, the geographical one, is also characteristic of scientific collections, found in the representation of specimens that records the spatial and ecological distribution of a species and its morphological and genetic variance within its range.

Despite the considerable wealth of information contained in most scientific collections, the access to such information has been essentially restricted to the scientists (taxonomists and systematists) who work on and take care of those collections. However, there has been an intense effort in assembling different catalogs and directories for biodiversity information in several developed countries. These mostly take the form of accounts of collections for diverse organisms, catalogs of described species, lists of endangered species, bibliographic sources, or environmental studies. Paradoxically, several of the best examples of development at the national level of the capacity to acquire, systematize, and utilize information about the biota of a country occur in some biodiversity-rich countries, as described next.

The Comisión Nacional Para el Conocimiento y Uso de la Biodiversidad (CONABIO)

CONABIO (http://www.conabio.gob.mx.), of Mexico, was created in March 1992 as an interministerial organization with the mission of coordinating the actions and studies related to the knowledge and preservation of biological diversity, as well as promoting and stimulating scientific research activities for the exploration, study, protection, and utilization of biological resources. Its aims are to preserve Mexico's ecosystems and to develop criteria for their sustainable management. Although CONABIO is placed within the Ministry of the Environment, and receives funding from federal sources, it operates with a private trust fund which allows CONABIO to obtain, in addition to federal funds, economic support from contracts, donations, and services, which help it to run a very efficient operation. The main functions of CONABIO are to: (1) Establish a national program for biological inventories. (2) Synthesize the information relating to biological resources in a permanently updated database. (3) Design and implement a National Biodiversity Information System. (4) Promote projects directed to the actual and potential use of biological resources that are both conventional and nonconventional. (5) Provide advice to decision makers in the government, civil society, and industry on technical and scientific aspects related to conservation and management of natural resources. (6) Produce educational materials for the general public, especially for younger people, and provide information about biological diversity and natural resources to all components of Mexican society. A recent effort of CONABIO has been a three-volume compilation on the ecosystems of Mexico, including a thorough account of known biodiversity of the country; a detailed assessment of the status and trajectories of change of Mexico's ecosystems, and a description of the national efforts to address the conservation and intelligent use of Mexico's "natural capital"(Sarukhán *et al.*, 2009).

In a period of only twenty years, CONABIO has become the central Mexican institution that compiles data and generates information on biodiversity and natural resources, and in so doing it has been able to bridge the gap between scientific research efforts and the production of information relevant to policy-making in the conservation of natural resources. One of its fundamental activities is the construction of a large, geo-referenced biodiversity database, developed with the participation of most of the taxonomic community of Mexico and also using the repatriation of information on Mexican

specimens housed in foreign herbaria and museums that hold large collections of specimens of Mexican organisms. Currently, these data are housed in various Mexican institutions, and are accessible to any interested person.

CONABIO does not conduct biological inventories or fieldwork on its own. Rather, it uses the biodiversity information already stored in national and foreign scientific institutions and funds a large part of the research carried out by Mexican scientists in the areas that have been identified as priorities in its mandate. The extensive program for the repatriation of data on Mexican species will continue over the coming years. The biodiversity database is constructed on the basis of the information contained in the label of each specimen included, which is taxonomically validated by experts in the different groups of organisms and then precisely georeferenced. At present the database includes information on approximately 5.0 million specimens, and is constructed into a geographic information system (GIS) containing a wide array of data on climate, geology, topography, soils, vegetation types, land use patterns, and other factors.

The Environmental Resources Information Network (ERIN), Australia

ERIN (http://www.erin.gov.au) is Australia's oldest of all national entities devoted to compiling the biodiversity information of a country and building a database with that information. It has been a model and inspiration for other such systems around the world. The Australian government established it in 1989 as part of the Federal Department of the Environment. As is the case with CONABIO, ERIN's aim has been to provide biodiversity and environmental information to assist planning and decision making in the fields of conservation and natural resource management. It also draws information from many sources and institutions and its database is constructed on the basis of specimen information and compiled into a very rich GIS. ERIN is the main source for an array of fairly sophisticated data analyses about the distribution, conservation, and management of the Australian biota and serves to generate information for decision-making. It has become an increasingly powerful analysis tool for many studies of the biogeography and ecology of the country's plant and animal species and the assemblages and ecosystems that they constitute. It now covers subjects that range from endangered species to the effects of drought and pollution.

The Instituto Nacional de Biodiversidad (INBio), Costa Rica

Started in October 1989, INBio (http://www.inbio.ac.cr) has pioneered a substantial effort to survey the biodiversity of Costa Rica. It is a private, nonprofit organization, entitled to receive both governmental and private funds for its operation. Its Board of Directors and the General Assembly include representatives from different sectors of Costa Rican society who are stakeholders in the efforts to conserve the natural resources in the country.

One of INBio's central aims is to carry out a basic national biodiversity inventory. To this end, it has developed an active program of biological surveys of the major groups of organisms and is building scientific collections of the different groups. This effort has involved the deployment of a substantial number of field assistants (mostly laymen and peasants) who are qualified as "parataxonomists" and who do most of the collecting of specimens and the initial taxonomic identification. INBio currently maintains a biodiversity database that provides geo-referenced information (by latitude–longitude), particularly of the specimens that have been collected since it started operating. Another of its stated aims is to support biodiversity conservation and promote new opportunities for sustainable development within the social and economic contexts of Costa Rica. It has also developed an important public education program on the subject.

A third area of INBio activity has been the prospecting for natural products in plants and animals of Costa Rica, which has been an especially active component within its various programs. The organization has a close collaboration with different governmental agencies in charge of the conservation and management of natural resources, and it has served as a constant source of advice and information to them in their programs.

The South African Botanical Diversity Network (SANBI)

SANBI (http://www.ckw@nbipre.nbiac.za) originated in the early 1990s as a multinational program closely linked to the National Botanical Institute (NBI) of South Africa; it was then called SABONET. It has now evolved into a very active institution with a mission to strengthen the level of biological expertise in the country, expand and improve scientific collections, and stimulate closer collaborative links among botanists and zoologists of the 10 countries that form the South African region.

The main objective of SANBI is "exploring, revealing, celebrating, and championing biodiversity for the benefit and enjoyment of South African people." SANVI is also responsible for ensuring that biodiversity knowledge influences policy, management, and decision-making. The three main areas of research of SANBI are biosystematics research and biodiversity collections, applied biodiversity research, and climate change and bio-adaptation. SANBI offers training courses and workshops to enable better environmental management practices and human capital development. It also organizes botanical expeditions to poorly known areas of the region, and publishes a newsletter. SANBI's homepage contains an ample description of its research and education programs.

Biota-Fapesp

Established in 1999, and conceived as a virtual Biodiversity Institute, this Brazilian organization (www.biota.org.br) deals with the biological diversity only of the State of Sâo Paulo. It aims at characterizing and inventorying the biodiversity of the State, defining mechanisms for its conservation, assessing its economic potential and its sustainable use. Like other similar institutions, it aims at attaining information that can be useful for decision making in relation to the use and management of biodiversity, as well as in sharing freely and amply all the information gathered.

Sistema Nacional de Datos Biológicos (SNDB), Argentina

The SNDB (www.datosbiologicos.mincyt.gov.ar) is the national Argentinian organization dedicated to the compilation of biodiversity information of the country. Established in 2009, it is part of the Ministry of Science, Technology, and Productive Innovation. Its mission is "to assemble a unified database on biological information, based on taxonomic, ecological, cartographic, bibliographic, and ethnobiological data, also comprising catalogs on the use of natural resources and related subjects." The system has a central node and several institutional nodes; it is built by 99 scientific collections housed in 24 institutions. Several of the collections also contribute to a variety of entities such as the Global Biodiversity Information Facility and the Inter American Biodiversity Information Network. The SNDB supports research proposals of the institutions that form part of it.

National Biodiversity Authority (NBA), India

Established by the Government of India, in 2003, the NBA (www.nbindia.org) has as its functions to advise the government on matters related to the conservation of biodiversity, sustainable use of its components and equitable sharing of benefits arising from the utilization of biological resources. It also regulates activities and issues guidelines for access to biological resources. In stark contrast with similar organizations in other countries, the NBA in India possesses regulatory capacities.

Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Colombia

Instituted in 1995, in Colombia, the Humboldt Institute (www.humboldt.org.co/avh) has the mission of "promoting, coordinating, and conducting research that contributes to the knowledge, conservation, and sustainable use of biodiversity as a factor of development and well-being of Colombian society." A civil corporation linked to the Ministry of the Environment, its mandate is to carry out scientific research on the biodiversity of the country. It is also responsible for conducting the National Biodiversity System, in coordination with close to 100 Colombian institutions, and for developing the country's national biodiversity inventory. A section of the Institute studies the territorial and historic processes linked to the way in which use, transformation, and maintenance of biodiversity takes place. The Institute also has undertaken, in collaboration with IUCN, the task of producing red book lists of the threatened species of Colombia.

Conclusions

The world's biological riches are unevenly distributed across the continents, and this distribution coincides with the same areas where many ancient and diverse cultures developed. These areas of high biological diversity now face threats not only from the activities and growing demand for natural resources of today's much larger native populations, but also from demands that originate in distant, developed countries. Presently, these resources and ecosystems are under direct and imminent threat. Yet we now have a more complete understanding of these ecologically diverse systems and the impacts of human activities on them on local, regional, and global scales.

Biodiversity-rich countries find themselves pressed between the responsibility, and even the fundamental desire, to preserve their biological diversity and the inescapable need to attend to the living requirements and well-being of their growing populations. The difficulty lies in meeting both the justified demands for development and increased well-being, and the responsibility for the conservation of those countries' natural capital. We believe that in order for these countries to increase the possibilities of having a successful outcome in this apparent and not justified contention, serious consideration should be given to the following two points.

1. It is impossible to make informed decisions about the conservation of a country's biological diversity if a scientifically sound platform of information does not exist. Such a platform is built with the training of specialists (taxonomists, systematists, ecologists, social scientists, etc.), the funding of extensive fieldwork, and the buildup of scientific collections or of information on the biological diversity of a country. These are long-term (many decades long) enterprises that should be undertaken in a carefully planned manner. Although these efforts are under way, a much faster and efficient way to collect and analyze information is to generate biodiversity databases using new field data and the vast amounts of existing information in herbaria and museums – all incorporated into an accessible GIS format. The authors believe that the examples discussed here for some countries offer an array of models that can be followed by other biodiversity-rich countries to enable them to better understand the extent and characteristics of their biological wealth.
2. In biodiversity-rich countries with large indigenous or rural populations, the conservation of biological diversity will require an honest and equitable collaboration with the owners of the resources (indigenous people or rural communities) that are the object of conservation or management plans. If this is not done beginning with the initial stages of planning, there will be limited success, if any, in such efforts. However, by carrying out successful collaborative programs using this approach, it will be possible to conserve a nation's biodiversity and the cultures, languages, and traditions of the indigenous peoples who own and inhabit the land. Sustained efforts toward achieving this type of partnership for conservation of biocultural diversity in biodiversity-rich countries will pay off in the long run. Very simply, this "experiment" has seldom been done correctly and, once proven practicable and successful, it could then be adapted and applied to a variety of social, economic, and biological situations.

Appendix

List of Courses

1. Anthropology
2. Biodiversity Science
3. Biogeography

4. Conservation Biology
5. Ecology
6. Tropical Biology
7. Political Ecology

See also: Biodiversity as a Commodity. Central America, Ecosystems of. Indigenous Peoples and Biodiversity. Social and Cultural Factors. Tropical Forest Ecosystems

References

Bray DB, Merino L, and Barry D (2005) *The Community Forests of Mexico: Managing for Sustainable Landscapes*. Austin, TX: University of Texas Press.

Dirzo R and Raven PH (2003) Global state of biodiversity and loss. *Annual Review of the Environment and Resources* 28: 137–167.

Food and Agriculture Organization (FAO) (2011) *State of the World's Forests*. Rome: Food and Agriculture Organization of the United Nations.

Heywood VH and Watson RT (eds.) (1995) *Global Biodiversity Assessment*. Cambridge, UK: Cambridge University Press.

Meffe GK and Carroll CR (1997) *Principles of Conservation Biology*, 2nd edn. Sunderland, MA: Sinauer Associates Inc.

Mittermeier RA, Goettsch-Mittermeier C, and Robles-Gil P (1997) *Megadiversidad: Los países Biológicamente más ricos del Mundo*. Mexico City, Mexico: Agrupación Sierra Madre, S.C.

Sarukhán J, Carabias J, Koleff P, *et al.* (2009) *Capital Natural de Mexico. Síntesis: Conocimiento Actual, Evaluación y Perspectivas de Sustentabilidad*. Mexico City, Mexico: Comisón Nacional para el Conocimiento y Uso de la Biodiversidad.

Trejo I and Dirzo R (2002) Floristic diversity of Mexican seasonally dry tropical forests. *Biodiversity and Conservation* 11: 2048–2063.

Biodiversity State and Trends in Southeast Asia

Lian P Koh and Chris J Kettle, Universitatstrasse, Zurich, Switzerland
Douglas Sheil, Mbarara University of Science and Technology, Kabale, Uganda; Center for International Forestry Research, Bogor, Indonesia, and Southern Cross University, Lismore, NSW, Australia
Tien M Lee, Columbia University, New York, NY, USA
Xingli Giam, Princeton University, Guyot Hall, Princeton, NJ, USA
Luke Gibson, National University of Singapore, Singapore, Republic of Singapore
Gopalasamy R Clements, James Cook University, Cairns, QLD, Australia

Published by Elsevier Inc.

Glossary

ENSO El Nino/La Nina-Southern Oscillation is a climate pattern that happens across the tropical Pacific Ocean in approximately 5-year intervals. It is characterized by variations in the ocean surface temperature and air surface pressure in the eastern and western tropical Pacific Oceans, respectively.
Indo-China Mainland Southeast Asia.
Malesia The Malay Peninsula, Malay archipelago, and New Guinea. Biogeographically, this combines Sundaland, Wallacea (including the Phillipines), and New Guinea.
Recalcitrant Seeds that have high moisture content and lose viability if moisture drops below a critical amount limiting their storage for long periods.
Sahul New Guinea, Australia, and their islands.
Sundaland The Malay Peninsula along with Sumatra, Java, Borneo, and associated islands.
Wallacea Lands east and north of Borneo, west of New Guinea and East of Wallace's Line, including the Phillipines, Sulawesi, Lombok, Sumbawa, Sumba, Timor, the Moluccas (e.g., Halmahera and Seram), Flores, and the other islands of Nusa Tenggara.

Introduction

Southeast Asia has been variously defined. Herein, "Malay Archipelago" states of Brunei, East Malaysia (Sabah and Sarawak), East Timor, Indonesia, the Philippines, and Singapore and the mainland states of Cambodia, Laos, Burma (Myanmar), Thailand, Vietnam, and Peninsular Malaysia are included. Australia, Bangladesh, China, India, and their island territories are excluded. Papua New Guinea and the Solomon Islands are also excluded. Indonesian Papua and neighboring islands of eastern Indonesia are included in national statistics but are judged biologically distinct from Southeast Asia proper.

The region's biota has complex origins (e.g., Morley, 2000; Briggs, 2003). Climate, sea level fluctuations, and tectonic motion have repeatedly joined and severed biotas (Morley, 2000; Wurstera *et al.*, 2010), whereas geology, terrain and climate provide diverse local environments (Zeuner, 1941).

The genus *Homo* has lived in Southeast Asia for over a million years (Corvinus, 2004). The species distributions of flora and fauna, including many fruit, bamboo, and palms, reflect millennia of human interaction. In this article, the state and trends of terrestrial, freshwater, and mangrove ecosystems in Southeast Asia are considered.

Habitat Loss, Degradation, and Recovery

Lowland and Montane Forests

The lowland rainforests of Southeast Asia are among the most diverse biomes, (Brooks *et al.*, 2006; Normile, 2010) and globally significant terrestrial carbon sinks (Saatchi *et al.*, 2011). The forests are vulnerable to clearance for agriculture as well as fire, mining, and unsustainable logging. The degradation of these forests also has major implications for the substantial numbers of forest-dependent rural poor (Poffenberger, 2009). Southeast Asia's forests face serious threats. The region has the highest rates of deforestation of any tropical region (Laurance, 1999, 2007), with rates between two and three times higher than those of other tropical regions (Achard *et al.*, 2002; Sodhi and Brook, 2006). Based on current deforestation rates of 0.71% per year (Achard *et al.*, 2002), 74% of Southeast Asia's forests could be cleared by 2100, which is predicted to cause the extinction of between 13% and 42% of all species in the region by the end of this century (Brook *et al.*, 2003).

Based on the latest Global Forest Resources Assessment (FAO, 2010), Southeast Asia has lost 33.2 million ha of forests in the past two decades at an annual deforestation rate of 0.72% and now sustains 214.1 million ha of forests. Forests in Indonesia span 94.4 million ha, the third largest forested area of any tropical country (FAO, 2010). The great extent of Indonesia's forests, representing 44.1% of all forests in Southeast Asia, highlights the importance of this country for regional conservation efforts. However, forest cover in Indonesia declined by 24.1 million ha between 1990 and 2010 at an annual deforestation rate of 1.14% (FAO, 2010). Although the absolute area of forest loss was greatest in Indonesia, Myanmar, Cambodia, and East Timor also had annual deforestation rates exceeding 1%. In contrast, Vietnam's forest cover actually expanded during this time period, from 9.4 million ha in 1990 to 13.8 million ha in 2010, an increase of 4.4 million ha, of which most (2.5 million ha) came from planted forests. Overall, the extent of planted forests in Southeast Asia increased from 11.7 million ha in 2000 to 14.5 million ha in 2010, at an annual growth rate of 2.14% (FAO, 2010).

In insular Southeast Asia, rates of deforestation have declined in the past two decades. In the late 1990s, fires during a severe El Niño-Southern Oscillation (ENSO) event destroyed very large areas of forest in Indonesia, which could explain the higher deforestation rates in the 1990s (Miettinen *et al.*, 2011a). Within this area, the eastern lowlands of Sumatra and the peatlands of Sarawak, Borneo had exceptionally high levels of forest loss with annual rates exceeding 5%. Overall, however, Indonesia's deforestation rates appear to be slowing down; the rate between 1990 and 2000 was 1.75% per year, 0.31% per year between 2000 and 2005, and 0.71% per year between 2005 and 2010 (FAO, 2010). In mainland Southeast Asia, Thailand's deforestation rates, although already relatively low, also appear to be slowing; deforestation rates of 0.28% per year from 1990 to 2000 and 0.11% per year from 2000 to 2005 changed to a net forest increase of 0.08% per year from 2005 to 2010 (however, this may simply reflect the fact that most of Thailand's forests have already been cleared). Although changes in forest cover have stabilized in some Southeast Asian nations, much of these changes represent the replacement of primary forests with plantation or secondary forests, which offer less conservation value (Gibson *et al.*, 2011).

The area of primary forest in Southeast Asia decreased from 66.3 million ha in 2000 to 64.0 million ha in 2010, a loss of 2.3 million ha (FAO, 2010). Indonesia suffered the greatest loss of primary forest among any nation, losing 2.0 million ha between 2000 and 2010 at an annual rate of 0.42%. Cambodia lost the next largest amount of primary forest and had the highest annual rate of primary forest loss of 3.48%, and Brunei (0.91%) and Vietnam (0.85%) also experienced high rates of primary forest loss. However, several nations in Southeast Asia reported no change in the extent of primary forests between 1990 and 2010, including Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, and East Timor.

Secondary forests may represent a key habitat for biodiversity conservation in Southeast Asia because of their great extent and their potential to regenerate into old-growth forests (Koh, 2007; Chazdon *et al.*, 2009). Although secondary forests were not included in the most recent Forest Resources Assessment, based on the previous assessment, secondary forests comprised 67.5% of the total native forest area (FAO, 2005; Koh, 2007). These potentially critical forests continue to be lost at rapid rates in almost all Southeast Asian nations, with the exception of Vietnam, the only country in Southeast Asia where the extent of secondary forests actually increased between 1990 and 2005.

Hosting half the world's cloud forests, with a total area of 32 million ha, Southeast Asia's montane forests are also under threat (Sodhi and Brook, 2006). Although these montane forests receive considerably less attention than their lowland counterparts, representing less than 10% of studies on biodiversity conducted across Southeast Asia's major terrestrial and freshwater ecosystems, they are important ecosystems and support exceptional levels of biodiversity (Peh *et al.*, 2011). Rates for lowland evergreen forests (1.2% per year) were higher than for montane forests (0.1% per year) in insular Southeast Asia (Miettinen *et al.*, 2011a). These forests are also highly threatened and are disappearing at much faster rates compared with lowland tropical forests in the region (Waggener and Lane, 1997). Indonesia, which has the greatest cover of montane forests at 19.5 million ha, is losing these forests at a rate of 3.05% per year from 2000 to 2005, a full order of magnitude higher than its rate of total forest loss for the same time period (FAO, 2005). Other countries in Southeast Asia that have more than a million hectare of montane forests include Malaysia, Myanmar, Thailand, and Vietnam. Although total forest cover is increasing in Vietnam, montane forests are disappearing from this country at a rate of 1.14% per year (FAO, 2005).

Singapore is a small island country located in the heart of Southeast Asia. In the past two centuries, more than 99% of Singapore's original forests were destroyed, leading to major losses of biodiversity (Corlett, 1992; Brook *et al.*, 2003). Singapore represents a worst-case environmental scenario that other nations in the region might be trending toward. For example, Java, the smallest of Indonesia's five large islands (the others being Borneo, Sumatra, Sulawesi, and Papua) and one of the most densely populated islands in the world with a population of more than 130 million people, now retains just 7% forest cover (Miettinen *et al.*, 2011a). Similarly, almost 90% of Thailand's original forests have been cleared (FAO, 2010), and continued agricultural and urban developments will only further reduce the extent of forests elsewhere in Southeast Asia. These examples suggest that similar losses in biodiversity will result throughout Southeast Asia unless ongoing clearance of forest is greatly restricted.

Rivers, Peat Swamp Forests, and Mangroves

To meet the increasing anthropogenic demand for food, water, and energy, freshwater ecosystems across the world are being extensively altered through the building of dams (Dudgeon, 2000b; Nilsson *et al.*, 2005), pollution (Driscoll *et al.*, 2007), introduction of exotic species (Leprieur *et al.*, 2008; Sanderson *et al.*, 2009), conversion of natural habitats (Allan and Flecker, 1993; Sanderson *et al.*, 2009), and the direct overharvesting of biological resources (Humphries and Winemiller, 2009). This section serves to highlight the causes, state, and consequences of habitat loss and degradation, as well as discuss the prospects of recovery, in two main inland aquatic and semiaquatic systems in Southeast Asia: rivers and streams, and peat swamp forests (PSFs), as well as a coastal aquatic system: mangroves (**Figure 1**).

Rivers and Streams

The biodiversity in rivers and streams in Southeast Asia are threatened by a myriad of anthropogenic activities. Deforestation/land-use change within a drainage basin (**Figure 2**) is one of the major threats to biodiversity in tropical freshwaters (Dudgeon, 2000a; Winemiller *et al.*, 2008) and especially so for Southeast Asia where deforestation rates are highest across the tropics in the period 1990–1997 (Achard *et al.*, 2002). Deforestation in the drainage basin and riparian zone leads to increased runoff, sedimentation, and incidence of flash floods (Dudgeon, 1999), thereby changing the natural seasonal flow regime of river systems. Sedimentation negatively impacts fish communities by reducing habitat complexity, altering available food types (Berkman and Rabeni, 1987), reducing



Figure 1 Main freshwater habitats across an altitudinal gradient. (a) A hill stream of the upper Katingan basin in Central Kalimantan, Indonesia. (b) A lowland swamp forest stream in Singapore. (c) A large river channel (Kahayan River) in Central Kalimantan, Indonesia (Photos: Heok Hui Tan (a) and (c).



Figure 2 (a) Deforested hill stream and (b) lowland stream in Kalimantan, Indonesia.

foraging efficiency (Bergstedt and Bergersen, 1997), and reducing spawning success (Burkhead and Jelks, 2001; Sutherland, 2007). Sedimentation may also lead to increased water turbidity and a reduction in primary production owing to decreased light penetration (Kottelat and Whitten, 1996), thus affecting river food webs. The removal of riparian vegetation also reduces the food availability for fishes that depend on allochthonous food sources, for example, terrestrial insects and terrestrial plant material (Inger and Chin, 1962; Knöppel, 1970; Dudgeon, 1983). In Sarawak, Malaysian Borneo Iwata *et al.* (2003) demonstrated that riparian deforestation (through slash-and-burn agriculture) increased sedimentation and lowered the abundance and/or diversity of benthic taxa (periphyton, aquatic insects, shrimps, crabs, and benthic

fishes) in hill streams. In Singapore, deforestation and canalization has led to the extinction of 11 (out of 46) native freshwater fishes (Giam *et al.*, 2011b).

Impoundments and dams are a major threat to the biodiversity of rivers in Southeast Asia. The impact of dams is greatest in the large rivers of mainland Southeast Asia, for example, Chao Phraya, Mekong, and Salween. In particular, proposed impoundments along the Mekong River have attracted considerable attention from conservationists and fishery scientists (Dudgeon, 2000b; Dugan *et al.*, 2010; Baran *et al.*, 2011; Ferguson *et al.*, 2011). As of 2009, 124 dams are already operating, under construction, or planned along the tributaries of the Lower Mekong, with a projected capacity of 29,760 M when completed (Mekong River Commission, 2010). Along the main stem (primary downstream segment) of the Upper Mekong in China, four dams (Manwan, Dachaoshan, Jinghong, and Xiaowan) have already been completed, with a further three in the construction or planning stage (Mekong River Commission, 2010). No dam currently exists along the main stem of the Lower Mekong, but 11 projects are being proposed, most of which are in Lao PDR (Mekong River Commission, 2010; Baran *et al.*, 2011; Ferguson *et al.*, 2011).

Dams impact both upstream and downstream freshwater biota in a variety of ways. First, the flooding that occurs upstream of the dam creates a new still water (lentic) habitat, which is oxygen poor (anoxic) and affected by sedimentation (Nilsson *et al.*, 2005). Fish and other freshwater organisms adapted to flowing waters may become locally extinct. At the same time, by trapping coarser grained and greater quantities of sediments upstream, dams alter the sediment transport regime in downstream regions, which in turn leads to changes in habitat quality for fish and other organisms (Poff *et al.*, 1997; Roberts, 2001; Poff and Hart, 2002). Second, dams alter the natural flood pulse of the river system, which contributes to the drying out of wetlands (Lemly *et al.*, 2000), and a reduction in the size and productivity of the floodplain system, and reduces the amount of food available for predatory fish (Wootton *et al.*, 1996). Third, dams fragment what were previously continuous, free-flowing habitats, preventing obligate migratory fish from completing their migrations to reproduce or feed (Roberts, 2001; Dugan, 2008). At the same time, by isolating small fish populations and preventing gene flow between adjacent populations, dams reduce the overall genetic diversity in the isolated populations (Williams, 2006). The impacts of dams on Southeast Asian biodiversity have been clearly documented. For example, the impoundment of Pak Mun Dam on the Mun River, a tributary of the Mekong in Thailand, resulted in declines in fish species richness both downstream and upstream of the dam (Roberts, 2001) and declines in species richness and ichthyomass in the reservoir area (Jutagate *et al.*, 2001).

Pollution affects most freshwater systems in the world (Dudgeon *et al.*, 2006) and Southeast Asia is no different (Dudgeon, 1992, 2000a). Rapid population growth, and intensive agriculture and industrial development have contributed to the deterioration of water resources. Untreated municipal and industrial wastes released into rivers are major problems in most Southeast Asian countries (Ministry of Natural Resources and Environment, 2006; Department of Environment Malaysia, 2007; Mekong River Commission, 2007; Thailand State of Pollution Report Working Group, 2011) (**Figure 3**). Urban sewage increases biochemical oxygen demand and subsequently decreases the amount of dissolved oxygen in water, resulting in a reduction of aquatic biodiversity (Ramírez *et al.*, 2008). Runoff in agricultural areas, which is rich in fertilizers, may also lead to eutrophication thus threatening river and stream biodiversity. Effluents discharged by industries along the Chao Phraya River in Thailand have resulted in heavy metal pollution and their accumulation in several fish species at the river mouth (Polprasert, 1982). In the Kelian River in Kalimantan, Indonesian sediment pollution due to open-cut gold mining resulted in a decline in macroinvertebrate richness and abundance as well as food web complexity (Yule *et al.*, 2010).



Figure 3 Floating toilets constructed in a lowland stream in Kalimantan Tengah, Indonesia. Municipal waste is disposed streamside, whereas human waste is released directly into the stream.

The threats outlined above often function synergistically to further threaten fish biodiversity and other organisms. For example, plantation conversion in Southeast Asia is likely to impact fish not only through increased sedimentation and water temperature, but also pesticide and fertilizer pollution. The building of hydroelectric dams may also speed up industrial development that may compound pollution problems if unmatched by the development of sewage treatment facilities (Dudgeon, 2000b). Sometimes the conservation impacts may go beyond aquatic environments. For example, the inundation of human settlements upstream of dams may lead to the migration of people into nearby forests resulting in forest degradation and further biodiversity impacts.

PSFs and Blackwater Streams

PSFs – a lowland forest ecosystem formed on peat soils that is derived from woody plant debris in areas with high rainfall, high temperature, and poor drainage (Posa *et al.*, 2011) – are mainly found on the islands of Sumatra, Borneo, and Peninsular Malaysia. Most of these PSFs are ombrogenous: they obtain their nutrients directly from rain and plant remains only (Whitmore, 1984). It can be classified as a semiaquatic system as the rain-fed water table fluctuates with changes in rainfall intensity but always remains close to or above the peat surface throughout the year (Page *et al.*, 1999). The tea-colored

water draining from PSFs is very acidic (pH 3–4.5), owing to the high concentration of humic acid and tannins (Whitten *et al.*, 2000) and is commonly termed "blackwater."

Despite being poorly drained and relatively infertile, Sundaland's PSFs are being converted to industrial-scale forestry and monoculture plantations at alarming rates (Koh *et al.*, 2011b; Miettinen *et al.*, 2011b). Among the different subregions, the rate of loss is highest in Sarawak (− 8.1% per annum) and Sumatra (− 5.2% per annum). Across the region, only 47 833 km^2 of PSFs remain, approximately 36% of their historical extent.

The high rates of deforestation in the blackwaters of PSFs are likely to threaten its unique assemblage of fish species via a number of processes. Clearance of PSFs stops the supply of senesced leaves that provides the likely basal resource for the aquatic foodweb – dissolved organic carbon (Yule, 2010). Where PSFs are converted to plantations, the physicochemical properties of the blackwaters are likely to change owing to burning, drainage, and the application of lime (to increase pH of peat) and fertilizers (Yule, 2010). In addition to biodiversity impacts, the degradation and conversion of PSFs also lead to the loss of substantial amounts of biomass and soil carbon (Koh *et al.*, 2011b).

Mangroves

Mangroves are forests that occur in marine or brackish environments along sheltered estuaries and river banks (Giesen *et al.*, 2006; FAO, 2007). They provide important ecosystem services for both humans (e.g., fuel wood and construction material) and wildlife (e.g., breeding nursery for fish and crustaceans) (Giesen *et al.*, 2006). Southeast Asia originally harbored more than 6.3 million ha of mangroves, the largest mangrove area of any region in the world (Giesen *et al.*, 2006). However, a sizeable proportion of these mangrove forests have since been converted. According to FAO (2007), 1.7 million ha of mangroves have been cleared between 1980 and 2005 in Southeast Asia. In terms of area, mangrove loss was highest in Indonesia; 1.3 million ha of mangroves were cleared at a rate of 1.5% annually. In terms of percentage loss, Timor-Leste and Vietnam were the highest in the region, experiencing an annual loss of 3.4% and 2.1%, respectively (**Table 1**). The loss of mangroves can be attributed to commercial logging that includes charcoal and fuelwood, aquaculture, agriculture, and conversion for human settlements (Giesen *et al.*, 2006; FAO, 2007; Spalding *et al.*, 2010). The main drivers of conversion are different for each country (Giesen *et al.*, 2006) and the reader is referred to the excellent country-level reviews by Giesen *et al.* (2006), FAO (2007), and Spalding *et al.* (2010).

Table 1 Decrease in mangrove area in Southeast Asia from 1980 to 2005

Countries	*Area (ha)*		*Change 2005–1980*		
	1980	*2005*	*Area (ha)*	*%*	*% yr^{-1}*
Brunei	18,400	18,400	0	0.0	0.0
Cambodia	91,200	69,200	− 22,000	− 24.1	− 1.1
Indonesia	4,200,000	2,900,000	− 1,300,000	− 31.0	− 1.5
Malaysia	674,000	565,000	− 1,09,000	− 16.2	− 0.7
Myanmar	555,500	507,000	− 48,500	− 8.7	− 0.4
Philippines	295,000	240,000	− 55,000	− 18.6	− 0.8
Singapore	500	500	0	0.0	0.0
Thailand	280,000	240,000	− 40,000	− 14.3	− 0.6
Timor-Leste	4250	1800	− 2450	− 57.6	− 3.4
Vietnam	269,150	157,000	− 112,150	− 41.7	− 2.1
Total	6,388,000	4,698,900	− 1,689,100	− 26.4	− 1.2

Source: Mangrove area data collated from FAO (2007) The world's mangroves 1980–2005: A thematic study prepared in the framework of the global forest resources assessment 2005. *FAO Forestry Paper 153*. Rome, Italy: Food and Agricultural Organization of the United Nations.

Note that Lao PDR is excluded from the table as it is a landlocked country with no mangrove forest

Prospects of Recovery

For these habitats to recover, habitat degradation must first be abated. However, conservation strategies must incorporate human development goals for them to be practicable and effective (Dudgeon, 1992; Giam *et al.*, 2011a). High rates of deforestation in the last 20 years (Achard *et al.*, 2002; Miettinen *et al.*, 2011a, b) suggest that forest loss and conversion is unlikely to cease in the near future. To slow down forest loss and conserve intact forested watersheds, one possible strategy is to shift agricultural expansion into degraded lands. This must be coupled with the development of wastewater treatment facilities and tighter regulation of pesticide and fertilizer use (Dudgeon, 1992) in these areas. The carbon payment scheme known as Reduced Emissions from Deforestation and Forest Degradation *plus* conservation, sustainable management of forests, and enhancement of forest carbon (REDD +) (UNFCCC, 2010) might help fund the conservation of intact areas, especially carbon-rich PSFs. In watersheds where the complete protection of forests is not possible, riparian buffers should be established to mitigate the effects of sediment (Ziegler *et al.*, 2005; Lorion and Kennedy, 2009) and nutrient pollution (Peterjohn and Correll, 1984).

The development of hydroelectric dams is likely to continue given the rising energy demands of the region (Dudgeon, 1992). Future development of dams should take into account the trade-offs between economic benefits and impacts to biodiversity and human livelihoods. The opening of water channel (sluice) gates during fish migration periods may help to maintain the biological connectedness of rivers. For example, when all the sluice gates of the Pak Mun Dam were opened between July 2001 and June 2002, some recovery of fish species richness, including migratory fish, was observed (Jutagate *et al.*, 2005).

The rate of mangrove conversion in Southeast Asia has not slowed down substantially for the last 20–25 years (Giesen *et al.*, 2006; FAO, 2007; Spalding *et al.*, 2010). Despite the bleak outlook for mangroves, there are causes for optimism. As biological and economic impacts of mangrove conversion (e.g., coastal fisheries productivity declines (Barbier, 1993)) become better understood, mangroves are increasingly being protected (Spalding *et al.*, 2010) and rehabilitated or restored (Macintosh *et al.*, 2002; Giesen *et al.*, 2006). Strategies for restoration are reviewed in Giesen *et al.* (2006). In addition to the rehabilitation and restoration of degraded mangroves, governments must muster the political will to stop degradation of the remaining intact mangroves of the region.

Species Status and Trends

Plants

Southeast Asia possesses an estimated vascular flora of around 60,000 species (20,000–25,000 species in Indo-China and 42,000 species in Malesia; Frodin, 2001). A generic level count for the regional flora is provided in **Table 2**. Species level summaries are less certain but Indonesia alone has an estimated 38,000 species of which around 18,700 are endemic (BAPPENAS, 2003). For example, Mt Kinabalu in Sabah, Borneo has around 5000 species in 1200 km^2 (Beaman, 2005). Plant diversity peaks in the wet equatorial lowland forests of Borneo, Sumatra, and Peninsular Malaysia. The 52-ha plot at Lambir, Sarawak, averages 247 species of tree over 10-cm diameter per hectare and possess 1003 such tree species in total (Losos and Leigh, 2004). A nondestructive survey of all vascular plants in just 200 m^2 of lowland forest in Tesso Nilo, Sumatra cataloged 217 species (Gillison, 2001).

The flora remains incompletely known. As recently as the year 2000, Laos's main commercial rattan was undescribed (Prance *et al.*, 2000). In Borneo 15–35% of the flora likely remains uncollected (Beaman and Burley, 2003). Surprisingly, some estimates suggest that, although three-quarters of Southeast Asia's mosses (76%) are already known, knowledge of at least some higher plants is considerably less (only 28% complete for the Fabaceae; see Giam *et al.*, 2010b).

Of 1371 red-listed plant species in Southeast Asia, 292 are critically endangered, 196 endangered, 737 vulnerable, 31 near-threatened, and 110 data-deficient. Five plant extinctions are documented: two mangos (*Mangifera casturi* Kosterm and *Mangifera rubropetala* Kosterm), two dipterocarp trees (*Dipterocarpus cinereus* Sloot and *Shorea cuspidata* Ashton), and one herb – the "woolly stalked Begonia" (*Begonia eiromischa* Ridl.). Only the mangos survive in *ex situ* collections (IUCNRedlist.org, accessed 23 August 2011). A recent detailed assessment of the 155 Dipterocarp trees of Peninsular Malaysia finds that five of them are critically endangered, 35 are endangered, 42 are vulnerable, and 46 are near-threatened (Chua *et al.*, 2010). One endemic species, *Shorea kuantanensis* P.S. Ashton, is judged extinct, whereas nine of 11 other considered possible extinct in a previous summary (Symington, 2004) have been rediscovered (Chua *et al.*, 2010).

Although ecological relationships are poorly explored, the region's pollination systems are considered relatively robust (Corlett, 2004). Exceptions include figs (*Ficus* spp.) where loss of Agaonidae wasp pollinators and thus of viable seed production has been observed following droughts (Harrison, 2000). There are concerns with bat-pollinated species such as durian and pete/petai (*Durio* spp. and *Parkia* spp.) but data are lacking. Comparison suggests that loss of vertebrate seed dispersers is generally a greater and more immediate concern than loss of pollinators – although both remain concerns in the longer term (Corlett, 1998, 2004).

Insects

The effects of land-use change and deforestation on the megafauna of Southeast Asia have been well reported (Sodhi *et al.*, 2009). However, the responses of less "charismatic" groups, such as insects, to habitat disturbance remain relatively poorly understood (Dunn, 2005). This is alarming, considering the global dominance of insects among animal communities in terms of species richness, abundance, and biomass (Wilson, 1987), as well as the overarching importance of insects in providing ecosystem services for human societies (e.g., pollination of crops) (Nabhan and Buchmann, 1997).

Data on the conservation status of insects are scarce because the International Union for Conservation of Nature (IUCN) Red List is highly biased toward larger and better studied taxa (e.g., vertebrates; Rodrigues *et al.*, 2006). In Southeast Asia, 87 terrestrial insect species are listed as threatened on the IUCN Red List, which represents 12% of total threatened terrestrial insect species worldwide. These threatened species are classified as "critically endangered" (CR), "endangered" (EN), or "vulnerable" (VU; IUCN, 2011). Of the 87 threatened terrestrial insect species, 43 are

Table 2 Generic richness by regions in Southeast Asia

Region	*Genera*	*Endemic genera*	*Genera standardized by area*[a]	*Genera standardized by collecting intensity*[a]
Cambodia	686	?	598	922
Laos	767	4	622	1028
Burma/Myanmar	1304	10	812	1987
Thailand	1605	17	1097	1959
Vietnam	1401	21	1061	1642
Borneo	1396	31	841	1704
Java	1403	3	1307	1206
Nusa Tengarra	812	1	802	964
Malaya	1457	19	1354	1288
Moluccas	898	3	1026	1009
Philippines	1406	20	1062	1467
Sulawesi	1000	5	850	1285
Sumatra	1300	7	877	1696

[a]Regression-based estimates correcting for area and collection effort (for full details see Marsh *et al.*, 2009)

Source: Reproduced from Marsh ST, Brummitt NA, de Kok RPJ, and Utteridge TMA (2009) Large-scale patterns of plant diversity and conservation priorities in South East Asia. *Blumea* 54: 103–108, with permission from Nationaal Herbarium Nederland.

Lepidoptera (butterflies or moths – CR: 1, EN: 14, and VU: 28) and 42 are Odonata (dragonflies or damselflies – CR:6, EN: 9, and VU: 27). There is also one vulnerable ant species (*Monomorium hospitum*; VU) and one extinct stick insect (*Pseudobactricia ridleyi*). Because of the high proportion of endemic species in Southeast Asia, the loss of many of these threatened species likely will result in global extinctions.

Butterflies are among the best-studied insect groups in Southeast Asia in terms of taxonomy and biogeography (D'Abrera, 1982, 1985, 1986; Corbet and Pendlebury, 1992; Igarashi and Fukuda, 1997, 2000). They are highly sensitive to habitat disturbance and have commonly been used as an indicator taxon for ecological research (Koh and Sodhi, 2004). Most field studies reveal that deforestation has significant impacts on butterfly diversity. For example, Willott *et al.* (2000) found that primary forests in the Danum Valley in Borneo had significantly lower species richness (121 vs. 151), Fisher's α-diversity (31 ± 3 (95% confidence interval) vs. 40 ± 4) and rarified species richness (120 ± 3 vs. 143 ± 4) than secondary logged forests.

Apart from the Lepidoptera, bees are the other insect group that has been relatively well studied in this region. Liow *et al.* (2001) found that below-canopy bee abundance was significantly higher in primary forests than in other land uses. In Sumatra and Indonesia, primary forests were also shown to contain higher bee richness, diversity, and abundance than disturbed lands, such as plantations (Sakagami *et al.*, 1990). Dung beetles is another group that have been the focus of ecological studies often because of the ecosystem services they provide, including nutrient cycling and seed dispersal (Davies *et al.*, 2001). Not surprisingly, a study on the impacts of forest conversion to *Acacia* and *Albizia* plantations in Sabah demonstrated that beetle species that are closely associated with primary forests were most negatively affected by forest conversion (Davies *et al.*, 2001). Termites and ants also provide important ecosystem services. In Sumatra, the species richness and relative abundance of termites were highest in primary forests compared with secondary forests and agricultural lands (Jones *et al.*, 2003). Furthermore, of the 34 termite species recorded in primary forests, only one could be found in cultivated areas. Similarly, Bruhl *et al.* (2003) found that less than half of the leaf litter ant species in contiguous forests were recorded in forest fragments.

Reptiles and Amphibians

Amphibians

Southeast Asia is a hotspot for amphibian species richness and endemism (Bickford *et al.*, 2010). More than 730 species of amphibians have been described for the region (IUCN, 2011) and many more are likely undescribed (Giam *et al.*, 2010a, 2011c). The conservation status of most amphibians has been assessed as a result of the 2004 Global Amphibian Assessment (Stuart *et al.*, 2004) and subsequent yearly assessments by IUCN (IUCN, 2011). Of the 738 native amphibians assessed on the 2011 IUCN Red List, 135 are threatened with extinction (i.e., Red List categories, critically endangered, endangered, or vulnerable), while a further 76 are near-threatened. Many species are poorly known (Rowley *et al.*, 2010); 268 species were categorized as data-deficient, that is, their conservation status could not be accurately assigned due to lack of data. In particular, all but one of the 24 caecilians (amphibians that superficially resemble earthworms or snakes) native to the region were classified as data-deficient. Population trends (in terms of population and range size) of 268 species are deteriorating, whereas those of 314 species are currently unknown.

Habitat destruction associated with logging and agricultural conversion is a major threat to amphibians in Southeast Asia (Bickford *et al.*, 2010) as most of the region's amphibians are forest specialists (Rowley *et al.*, 2010). In Sulawesi, Indonesia, amphibian species richness and abundance were higher in primary and secondary forests compared with agroforests and open areas (Wanger *et al.*, 2010). Illegal logging, in addition to mercury and sediment pollution resulting from illegal gold mining, is threatening the persistence of the only known lungless frog in the world (*Barbourula kalimantanensis*; Bickford *et al.*, 2008). Overexploitation of amphibians for food, traditional medicine, and the pet trade are major threats to amphibians in Southeast Asia (Rowley *et al.*, 2010). Predicted climate warming is likely to exacerbate the impacts of habitat degradation and overexploitation (Bickford *et al.*, 2010).

Reptiles

In 2008, Southeast Asia (excluding Philippines and islands east of Western Nusa Tenggara) harbored 975 native reptile species (Das, 2010). However, the conservation status of most of these species remains unknown – less than a quarter have been assessed by the IUCN as of October 2011 (IUCN, 2011). The proportion of species that were assessed differed across major groups of reptiles. The conservation statuses of crocodiles (83% of species assessed; 5 of 6 species) and turtles (75%, 40 of 53 species) were better elucidated compared to snakes (27%, 132 of 484 species) and lizards (15%, 65 of 432 species). Of the 242 species that were assessed, 50 species were found to be threatened (i.e., categorized as critically endangered, endangered, or vulnerable). A further 80 species were classed as data-deficient. Based on available species assessments, the crocodiles and turtles appear to be most threatened. Thirty-five of 40 assessed turtle species and four of five assessed crocodile species were deemed to be threatened. The main threats to turtles appear to be from the wildlife trade, bushmeat hunting, degradation of habitat for nesting (for sea turtles), pollution, and bycatch mortality (i.e., mortality from entanglement in fishing nets; Shanker and Pilcher, 2003; Pacini and Harper, 2008; Nijman, 2010; IUCN, 2011). Crocodiles are threatened mainly by habitat degradation (Pacini and Harper, 2008); persecution and exploitation for bushmeat and wildlife trade are also important threats (IUCN, 2011).

The low proportion of threatened lizards and snakes (relative to turtles and crocodiles) may have arisen partly due to a lack of information on population and range trends of many species. Many of the data-deficient and unassessed species may actually be threatened. For example, Thirakhupt's Bent-Toed Gecko (*Cyrtodactylus thirakhupti*) is restricted to a single cave in a forested limestone hill in Surat Thani province, Southern Thailand (Pauwels *et al.*, 2004; Das, 2010) and is

likely to be classified as threatened owing to its highly restricted range. To inform conservation priorities, the conservation statuses of unassessed species must be elucidated as soon as possible. Habitat degradation is listed as the main threat to the two red-listed lizards (*Bronchocela smaragdina* and *Cyrtodactylus cavernicolus*) and many of the threatened snakes (IUCN, 2011). As the majority of range-restricted reptiles are confined to the rapidly diminishing lowland rainforests (Das, 2010; Miettinen *et al.*, 2011a), species declines are imminent.

Birds

Birds represent one of the best-studied groups in Southeast Asia. Among the main tropical regions in the world, Southeast Asia has the highest mean proportion of country-endemic (9%) and the highest mean proportion of threatened bird species (6%; Sodhi *et al.*, 2010a). The relatively high degree of endangerment may be related to the high rate of recent deforestation observed in this tropical region. Up to 67% of forest bird species have been lost as a result of more than 90% forest loss in the highly urbanized Singapore representing one of the world's highest estimates of long-term loss of forest species (Castelletta *et al.*, 2000). In addition, local extinctions have also been reported for small, yet largely forested Indonesian and Malaysian islands (Trainor, 2007; Sodhi *et al.*, 2010b). Using estimates of forest loss and the species area relationship, Brook *et al.* (2003) projected that between 16% and 32% of Southeast Asian forest bird species will become extinct by 2100.

Endemic forest species of insular Southeast Asia have been shown to be negatively affected by deforestation and various forms of agricultural land use (Indonesia, Lee *et al.*, 2007; Maas *et al.*, 2009; the Philippines, Posa and Sodhi, 2006). For example, more sensitive species such as woodpeckers may take a longer time to reach predisturbance levels (Styring and Ickes, 2001). However, some bird communities did recover after several decades (Wong, 1986; Yap *et al.*, 2007). In fact, some logged forests in Southeast Asia can support similar or even higher species richness of birds than undisturbed forests, which may be due to improved conditions for widespread forest edge species (Johns, 1996; Edwards *et al.*, 2011). Similarly, human-modified habitats such as rubber and oil palm plantations or cacao and coffee agroforests may have higher bird species richness than open and more disturbed habitats, but the species similarity with forests is often extremely low (Aratrakorn *et al.*, 2006; Danielsen and Heegaard, 1995; Peh *et al.*, 2005; Sodhi *et al.*, 2005a; Waltert *et al.*, 2004). For instance, only 27% of observed forest species were also found in coffee plantations in Sumatra (Philpott *et al.*, 2008).

High densities of tall shade trees and proximity to natural forest may improve overall conservation value of agroforests (e.g., in cacao; Abrahamczyk *et al.*, 2008; Clough *et al.*, 2009). Having tall forest trees in otherwise altered agricultural landscapes (oil palm, cacao, and rubber) and the protection of forest fragments could be an alternative to protected primary forest fragments and are particularly critical for the remaining widespread, disturbance-tolerant forest species (Abrahamczyk *et al.*, 2008; Castelletta *et al.*, 2005; Clough *et al.*, 2009; Peh *et al.*, 2006; Sodhi *et al.*, 2005b). However, small fragments (< 100 ha) have been shown to continue to lose species even decades after isolation and are colonized by generalist nest predators such as house crows (*Corvus splendens*; Sodhi *et al.*, 2005c, 2006a). Overall, with increasing land-conversion pressure for industrial-scale agriculture productions (e.g., oil palm), the long-term persistence of native Southeast Asian forest bird species appears to be particularly in jeopardy.

Mammals

Southeast Asia hosts exceptional levels of mammal diversity, with the highest degrees of endemism of any tropical region (Sodhi *et al.*, 2010b). However, mammals in Southeast Asia also have the highest proportion of threatened species among any tropical region (Morrison *et al.*, 2007; Schipper *et al.*, 2008; Sodhi *et al.*, 2010b), and an estimated 85% of species are at risk of extinction during this century due to deforestation (Sodhi and Brook, 2006). A recent meta-analysis found that mammals were the most sensitive taxonomic group to forest disturbance in Southeast Asia (Sodhi *et al.*, 2009), suggesting that they are most adversely affected by increasing human pressures in the region.

Forest loss and disturbance affect mammals in different ways. Despite general trends of decreased diversity and abundance (Sodhi *et al.*, 2009), many species are able to persist or recover following forest degradation (for a comprehensive review, see Sodhi *et al.*, 2010a). In Borneo, selectively logged forests can support large numbers of mammal species (e.g., the tarsier *Tarsius bancanus*), but the response varies by species based on ecological requirements (Meijaard *et al.*, 2008). In northern Borneo, small mammal diversity was lower in selectively logged forest sites compared to undisturbed forests; however, logged forests retained many species found in undisturbed forests, highlighting the conservation value of these forests (Wells *et al.*, 2007). Secondary forests in northern Borneo were also found to support diverse assemblages of up to 27 different mammal species, but Acacia and oil palm plantations offered little to no support (Bernard *et al.*, 2009; McShea *et al.*, 2009). The Philippines hosts an impressive mammal assemblage, with 90% of its more than 100 nonvolant species found nowhere else (Rickart *et al.*, 2011). In northern Luzon Island, species richness of small mammal was greatest in undisturbed forest habitats, but moderately disturbed forests also supported high levels of diversity and abundance (Rickart *et al.*, 2011).

Bats comprise a third of all Southeast Asian mammals, and of Southeast Asia's 330 species, 60% are endemic to the region (Kingston, 2010). Despite this, bats remain some of the least known vertebrates of Southeast Asia and 40% of species are predicted to be extinct by 2100 due to deforestation (Kingston, 2010). Although forest loss in the region continues, large forest fragments in Peninsular Malaysia can retain high levels of bat diversity (Struebig *et al.*, 2010).

Other mammals in the region are at high risk of extinction due to illegal poaching. Southeast Asia's pangolin species are threatened by trade for use in food, medicine, and clothing, particularly in Vietnam and China (Newton *et al.*, 2008).

Vietnam's primates are also highly threatened by poaching, with 21 of 24 species listed as threatened (Nadler *et al.*, 2007). Although hunting has greatly reduced populations of large ungulates throughout Southeast Asia, some populations have recovered following regulation of illegal poaching (e.g., the muntjac *Muntiacus muntjak*; Steinmetz *et al.*, 2010).

Large mammals may face the greatest threats because they are particularly vulnerable to habitat loss and are often targeted for hunting. Three of the region's flagship species, rhinos, elephants, and tigers, persist in isolated pockets of forest throughout their historic ranges and are increasingly rare throughout Southeast Asia (Clements *et al.*, 2010a; Lynam, 2010). In Sumatra, tigers were found in both primary and selectively logged forests, but tiger densities were higher in primary forests, probably because selectively logged forests offer greater access to hunters (Linkie *et al.*, 2008). Although secondary and selectively logged forests both sustain populations of elephants and tigers in Malaysia, the conversion of forests into agriculture has destroyed large areas of critical habitat for these species (Clements *et al.*, 2010a). Agricultural expansion has also led to human–wildlife conflicts that often result in the death of these endangered animals (Nyhus and Tilson, 2004; Clements *et al.*, 2010a).

Threats to Biodiversity

Industrial Agriculture

Industrial agriculture has become an increasing threat of forests and biodiversity across the tropics (Rudel *et al.*, 2009a, b). In Southeast Asia, oil palm agriculture has been expanding at an astonishing rate of 470,000 ha per annum between 1999 and 2009 (FAO, 2011). Currently, Indonesia and Malaysia produce 170 million metric tons of oil palm fruits each year, which accounts for 80% of the global production (FAO, 2011).

Conversion of natural habitats into a monoculture such as oil palm, almost by definition, implies a drastic loss in biodiversity. In fact, oil palm plantations contain less than half as many vertebrate species as primary forests (Fitzherbert *et al.*, 2008). Forest bird species declined by 73–77% (Koh and Wilcove, 2008), and only 10% of mammal species were detected in oil palm plantations (Maddox *et al.*, 2007). Endangered species such as the Sumatran tiger (*Panthera tigris sumatrae*), tapir (*Tapirus indicus*) ,and clouded leopard (*Neofelis nebulosa*) have not been recorded in oil palm plantations; and most mammals prefer marginal and heavily degraded lands (e.g., shrublands) to oil palm (Maddox *et al.*, 2007). Mammals that do occur in oil palm plantations tend to be of low conservation concern and are dominated by a few generalist species such as the wild pig (*Sus scrofa*), bearded pig (*Sus barbatus*), leopard cat (*Prionailurus bengalensis*), and common palm civet (*Paradoxurus hermaphroditus*).

The impacts of industrial agriculture on invertebrates are more variable compared with vertebrates (Fitzherbert *et al.*, 2008). Conversion of forests to oil palm resulted in 79–83% declines in the species richness of forest butterflies (Koh and Wilcove, 2008). However, ants, moths, and bees demonstrated a higher tolerance for the oil palm habitat (Danielsen *et al.*, 2008). Nevertheless, studies consistently showed a dominance of nonforest invertebrate species in oil palm plantations (Danielsen *et al.*, 2008). On average, only 15% of vertebrate and invertebrate species recorded in primary forest could be found in oil palm plantations (Fitzherbert *et al.*, 2008). Not surprisingly, plant diversity within oil palm plantations was also impoverished compared with forests due to regular maintenance (e.g., weeding) of oil palm fields (Danielsen *et al.*, 2008; Fitzherbert *et al.*, 2008).

Climate Change and Forest Fires

Climate change impacts on biodiversity have been relatively mild compared with the other main threats in the last century. However, it is very likely that climate change will be one of the most dominant biodiversity threats this century (MEA, 2005). In particular, strong field evidence of climate change effects has been reported in the montane areas in Southeast Asia, where moth ranges have shifted to higher elevations in the last 42 years (e.g., Sabah, Malaysia; Chen *et al.*, 2009). In addition, recent reviews of the literature and guide books revealed that certain amphibians and reptiles, and birds, respectively, appear to be responding to global warming by shifting their elevational ranges upwards (Peh, 2007; Bickford *et al.*, 2010). Because Southeast Asia consists largely of islands with few mountain ranges, many affected island species may not be able to migrate or disperse under climate change (Bickford *et al.*, 2010). As a consequence, high elevational species may be adversely impacted by rapid climate change, indirectly affecting species assemblages and communities, and ecosystem functioning and services.

Forest fires in Southeast Asia can negatively impact species and alter communities. For example, Sumatran bird diversity declined with open field species replacing forest specialists (e.g., understorey insectivores) over time (Adeney *et al.*, 2006). ENSO-induced wildfires in Borneo also reduced butterfly species richness significantly (Cleary *et al.*, 2003). Both studies suggest that wildfires have strong impacts on species communities in Southeast Asia. The two recent ENSO-induced wildfire episodes (i.e., 1982–83 and 1997–98) revealed the potential for destruction of high-intensity fires on forest biodiversity (Baker *et al.*, 2009; Page *et al.*, 2009). However, more critically, the changing fire regimes (increasing occurrence of large fires) due to both growing human pressures (through forest clearance) and projected climate change may further result in a huge loss of forest with severe impacts on biodiversity. They might also pose serious public health risks to the human populations in this region (Lohman *et al.*, 2007; Taylor, 2010). In essence, forest loss and fragmentation, and climate change, would lead to the edges of forest patches drying up, a regional drop in evapotranspiration rate, and large-scale shifts in rainfall due to a disruption of the Indian monsoon and ENSO cycle, all potentially enhancing fire risk and frequency, water loss, and thermal stress (Brook *et al.*, 2008). Together, the synergisms of these threats would place the biodiversity and people of Southeast Asia under grave danger. As such, large-scale and regionally coordinated fire management plans are urgently needed to mitigate the projected high-intensity and high-risk wildfires (Lohman *et al.*, 2007; Johnson and Dearden, 2009).

Overexploitation and Illegal Trade of Wildlife

Overexploitation of wildlife can cause irreversible loss of species and populations, ecosystem perturbations, economic loss, food shortages, and spread of pathogens. In the last 50 years, the need to hunt for subsistence in tropical Asia has been increasingly supplanted by the need to hunt for the market (Corlett, 2007) – this has led to population declines and local extirpation of threatened species in certain areas (Lee *et al.*, 2005; Phoonjampa and Brockelman, 2008). More charismatic threatened species (e.g., rhinos, elephants, and tigers; Clements *et al.*, 2010a), however, are increasingly being hunted to fuel the trade of illegal wildlife and their products (e.g., ingredients in traditional Chinese medicine; **Figure 4**). The illegal wildlife trade refers to the unlawful trade, smuggling, poaching, capture, or collection of endangered species or their parts (including animals or plants that are subject to harvest quotas and permit regulation; South and Wyatt, 2011).

Southeast Asia is now regarded as one of the world's largest illegal wildlife trading hubs, with almost every country serving as source and/or transit point to markets beyond the region (i.e., China being the largest market followed by the USA; Felbab-Brown, 2011). The value of this illicit market in this region was recently estimated to be worth around USD 10–30 million annually (Then, 2010; Felbab-Brown, 2011).

The Convention on International Trade in Endangered Species (CITES; www.cites.org), which is an international agreement among governments to ensure that the international trade of wildlife specimens does not threaten their survival, affords varying degrees of protection to more than 30,000 species of plants and animals. A nongovernmental organization (NGO) known as TRAFFIC (www.traffic.org) is involved in monitoring the legal and illegal trade of CITES species in Southeast Asia. Despite regulatory mechanisms such as CITES, large volumes of CITES-listed species permitted for trade continue to be exported illegally above set quotas (Nijman, 2010). Furthermore, species that are prohibited for international trade continue to be sought after by poachers due to the relatively low risk and high profitability of this business (**Figure 5**). Examples of wildlife being persecuted for the illegal trade in Southeast Asia include mammals (e.g., bears, Shepherd and Nijman, 2007; elephants, Shepherd and Nijman, 2008; felids, Ng and Nemora, 2007; pangolins, Pantel and Chin, 2010; primates, Nijman, 2005; rhinos, Ahmad Zafir *et al.*, 2011); reptiles (e.g., snakes, Auliya 2010; tortoises and turtles, Nijman and Shepherd, 2007); birds (Shepherd *et al.*, 2004); and plants (The World Bank, 2005).

Although legal and illegal hunting of species in this region is regulated to some extent by national wildlife laws, hunting quotas continue to be exceeded and hunting prohibitions may be ignored partly due to the lack of mechanisms in place to monitor population trends of species being hunted (Clements *et al.*, 2010a). The harsh reality is that enforcement systems in this region have largely failed to curb wildlife overexploitation, which is showing no sign of abating (Bennett, 2011; **Figures 6** and **7**).

Others: Invasive Species and Disease

Invasive species can be defined as exotic species that adversely affect environments they invade. In Southeast Asia, invasive species can cause the displacement of native biodiversity due to synergistic factors such as an invasive species' advantageous life-history traits (in certain plants and molluscs; Osunkoya *et al.*, 2005; Clements *et al.*, 2006), competition for similar resources (among birds; Huong and Sodhi, 1997), ecosystem modification (by plants and butterflies; Peters, 2001; Ghazoul, 2004), hybridization, environmental disturbance, and economic loss (Brook *et al.*, 2003). The Global Invasion Species Database (www.issg.org/database) contains the most comprehensive list of invasive species present in Southeast Asia. Indeed, every country in the region contains invasive species, with Indonesia harboring the highest number (78 species), and plants and fish being the most common invasive animal groups (Peh *et al.*, 2010). Continental islands such as Singapore appear to be the worst-case scenario for invasibility with alien species dominating persistently disturbed areas, but very few introduced species have successfully invaded closed-canopy forests (Corlett, 2010). However, in natural aquatic systems, the negative impacts of invasive species appear to be more pervasive (e.g., among fish; Nguyen and de Silva, 2006). Ecosystems that are subjected to increasing levels



Figure 4 One of Southeast Asia's endangered mammals, the Malayan tiger (*Panthera tigris*), photocaptured in a selectively logged forest in Malaysia. This charismatic species is being persecuted by illegal wildlife traders for products such as ingredients in traditional Chinese medicine.



Figure 5 A group of poachers caught on a camera trap trespassing into selectively logged forests from a highway in Malaysia. Their possession of a modified shotgun indicates that they are targeting endangered large mammals, possibly for the lucrative wildlife trade.



Figure 6 Agricultural expansion in the montane forests of Peninsular Malaysia has increased levels of siltation and chemicals in downstream water systems with negative effects on freshwater biodiversity.

of disturbance can also become more susceptible to the establishment of invasive species. For example, more disturbed peat swamps in Malaysia appear to be more prone to invasion by noxious weeds (Yule, 2010).

The transmission of diseases and pathogens that threaten Southeast Asia's biodiversity has undoubtedly been facilitated by globalization. One particular emerging disease of great concern to herpetologists is chytridiomycosis, a fatal

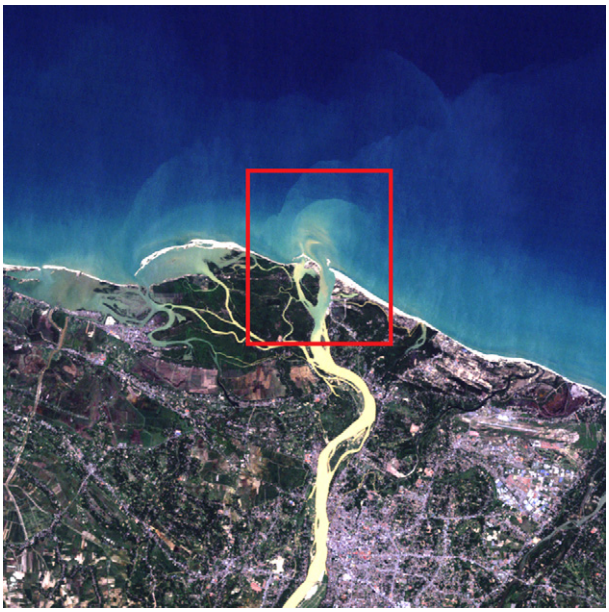


Figure 7 A LANDSAT 5 satellite image taken in August 2004 over Peninsular Malaysia showing an influx of sediments (red box) from upstream developments along the Kelantan River into the marine environment. Data available from the US Geological Survey.

amphibian-specific skin disease caused by a fungal pathogen *Batrachochytrium dendrobatidis*. Although this chytrid fungus has mainly been recorded in tropical regions outside Southeast Asia, it was recently recorded for the first time among amphibian species in Indonesia (Kusrini *et al.*, 2008). Commercial trade and invasive species represent the more likely vectors of pathogens, but large-scale mortality of native biota as a result of pathogens introduced from alien vectors are yet to manifest in this region (Peh, 2010). However, approximately 75% of human diseases having links with wildlife is of concern (Patz *et al.*, 2004). In Indonesia, for example, the risk of transmitting zoonotic diseases between fruit bats and humans appears to be high given the former is a carrier of Henipaviruses (Harrison *et al.*, 2011), which have caused numerous fatalities among animals and humans in Malaysia and Singapore (e.g., Nipah virus; Breed *et al.*, 2006). The deadly H5N1 influenza virus also remains a major threat to both humans and wild birds in this region. For example, this virus has been found in bar-headed geese that died *en masse* in certain areas of China after being infected; this species may pose a threat to native avifauna in the region as it migrates to aggregation and breeding grounds in Myanmar (Chen *et al.*, 2005).

Challenges and Opportunities

Enhance Protected Area Networks

Protected area networks represent one of our core strategies in biodiversity conservation and protection of ecosystem services (Chape *et al.*, 2005; Luck *et al.*, 2009). Reserves in Southeast Asia have been shown to be safe havens for butterflies, forest endemic birds and bats, for instance (Koh and Sodhi, 2004; Lee *et al.*, 2007; Struebig *et al.*, 2010). In addition, protected areas can be effective in buffering against human encroachment such as deforestation (e.g., Myanmar, Songer *et al.*, 2009). However, these reserves are also under increasing threats from isolation due to extensive forest loss and degradation in areas surrounding them (DeFries *et al.*, 2005). Furthermore, reserves in Southeast Asia may only be "paper parks" that are poorly managed and suffer illegal logging (e.g., in Kalimantan, Indonesia; Curran *et al.*, 2004). The proportion of natural forest under protection across the countries is also relatively low ($<10\%$), with nations like Myanmar and Vietnam having $<5\%$ of land area currently protected (Rao *et al.*, 2002; Sodhi *et al.*, 2010a). This is despite the fact that the region has the highest deforestation rate as well as the lowest remaining forest cover (Sodhi *et al.*, 2010a). Poverty and corruption may also be tied to higher incidence of fires in Southeast Asia parks (Wright *et al.*, 2007). Nevertheless, parks, if effective, have the potential to alleviate human poverty rates as recently demonstrated for Thailand (Andam *et al.*, 2010).

The future conservation battlegrounds are expected to shift under future climate and land-use changes (Lee and Jetz, 2008). In particular, Southeast Asia will continue to suffer heavy conversion pressure as global human populations and demand for agricultural products soar this century (Lee and Jetz, 2008). Much of the unprotected land is projected to be transformed into human-modified landscapes thus increasingly reducing the connectivity between reserves. This potentially hampers species dispersal, which may be the most serious in high-altitude sites where species may be pushed out of the reserve boundaries due to increasingly unsuitable climatic niches (Hannah *et al.*, 2007). In Thailand, for instance, the national protected area network has been assessed for the conservation of tropical forest butterflies using high resolution regional land-use and climate change projections (Klorvuttimontara *et al.*, 2011). Studies evaluating the effectiveness of reserves in protecting biodiversity under global change will be extremely valuable for future conservation planning and should be a priority in Southeast Asia. As such, it becomes particularly important and urgent to rapidly expand the reserve network of the region, considering projected change impacts; increasing connectivity among reserves; and improving park enforcement, governance, and management in order to combat impending global change.

Safeguard Biodiversity in Human-Dominated Landscapes

Well-enforced protected areas have demonstrated their effectiveness at preserving biodiversity in some countries (e.g., in the Philippines; Samoilys *et al.*, 2007). Improving the management of protected areas is therefore one of the most important strategies to safeguard biodiversity in human-dominated landscapes across the region; even a modest increase in funds can help improve their effectiveness (Bruner *et al.*, 2001). Community-based conservation within protected areas should also be explored as local collaborative programs can aid wildlife population recovery (e.g., in Thailand; Steinmetz *et al.*, 2010). Given that primary forests are irreplaceable in sustaining biodiversity (Gibson *et al.*, 2011), remaining tracts of primary forests with adequate habitat connectivity should be prioritized for protection (e.g., by

prohibiting road development; Laurance *et al.*, 2009) and gazetted as protected areas if possible.

Degraded habitats such as selectively logged forests must still be afforded greater protection as they are biologically important (Edwards *et al.*, 2011), supporting some of the world's most endangered species (e.g., the tiger *Panthera tigris*; Rayan and Shariff, 2009). Selectively logged forests of high conservation value should therefore be prioritized for improved sustainable management (Giam *et al.*, 2011a, b) and protected from conversion to commercial plantations (e.g., Aziz *et al.*, 2010), which should only be allowed to expand on preexisting cropland and degraded habitats (Koh and Wilcove, 2008). One mechanism through which selectively logged forests in Southeast Asia can be better protected is through the United Nations Framework Convention on Climate Change (UNFCCC) framework "REDD+" (UNFCCC, 2010). Although REDD schemes show great promise in protecting the region's imperiled biodiversity (e.g., for mammals in Indonesia; Venter *et al.*, 2009), and web-based decision-support tools are now available (Koh *et al.*, 2011a; available at www.REDDcalculator.com), many political, economic, and social challenges have to be overcome to ensure success (Butler *et al.*, 2009; Clements *et al.*, 2010b; Sanderbrook *et al.*, 2010).

To mitigate specific threats such as illegal hunting of threatened species in and around forests, NGOs have established antipoaching groups (**Figure 8**) in association with enforcement authorities to combat the threat of poaching (e.g., in Malaysia and Indonesia; Clements *et al.*, 2010a; Ahmad Zafir *et al.*, 2011). At a larger scale, regional governments are taking encouraging steps by improving outdated national wildlife laws (e.g., the Wildlife Conservation Act 2010 in Malaysia) and forming an intergovernmental wildlife enforcement network (asean-wen.org), which has had some success in clamping down on illegal wildlife traders (Lynam, 2010).

To better protect native fauna from invasive species, recommendations (in ascending order of gravity) would be: (1) prevention of their introduction entirely, (2) swift eradication of an incipient population, and (3) harnessing a combination of technologies to mitigate the threats from an established population (Corlett, 2010; Simberloff, 2010). Scientists that discover the presence of potentially invasive species should not only publish their findings as soon as possible, but also alert authorities to urgently implement appropriate control measures (e.g., **Figure 9**; Tan and Clements, 2011). Although zoos and gardens should be well equipped with strategies to minimize the risk of escape from invasive species (Dawson *et al.*, 2008), custom departments should adopt stricter quarantine measures for species arriving via the horticultural and pet trade, which are considered the main avenues for the arrival of invasive species (Corlett, 2010).

To deal with the threat of emerging infectious diseases (EIDs) to native biodiversity, control measures include: (1) minimizing contact with the direct host, (2) monitoring of intermediate hosts, (3) improving biosecurity on farms, and (4) better disease recognition and diagnosis (Field and Mackenzie, 2004). Studies aimed at identifying traits that predict the vulnerability of taxa to EIDs should be carried out well in advance of their declines as recovery of population crashes can be very difficult, if not, impossible (e.g., Bielby *et al.*, 2008). Regarding the threat of pollution, several



Figure 8 A ranger from the Rhino Protection Units (RPUs) of the Rhino Conservation Program in Indonesia, a good example of an antipoaching initiative formed by an NGO in collaboration with government authorities. RPUs not only conduct intelligence-driven patrols in protected areas trying to save the Sumatran Rhino from extinction, but they also help remove poaching threats to other endangered species. (Photo credit: Ahmad Zafir Abdul Wahab).

conventions have been signed to address the negative effects of marine pollution to biodiversity (see Todd *et al.*, 2010). A reduction in pollution from wildfires may be achieved through several priority actions (see Lohman *et al.*, 2007), especially through the enactment and enforcement of laws that prohibit further conversion of peatlands. For example, a 2-year moratorium on peatland conversion in Indonesia was passed in 2010 in return for monetary payments from the Norwegian government on successful implementation.

Despite some conservation successes at varying scales (Sodhi *et al.*, 2011), much more needs to be done (beyond the limits of this article) to safeguard biodiversity in human-dominated landscapes within Southeast Asia. Ultimately, biodiversity conservation requires greater political will and support from civil society, including industry, environmental groups, and the wider public (Clements *et al.*, 2010a; Bennett, 2011), without which there will be no halt to the current slide of the region's biodiversity toward extinction (Sodhi *et al.*, 2004).

Forest Restoration

Reversing these trends through forest restoration is attracting attention across the tropics (Braxton Little, 2008; Chazdon,



Figure 9 A potentially invasive snail, *Limicolaria flammea*, recently discovered in Singapore, possibly introduced from Africa via the horticultural trade. Scientists who discovered it alerted government authorities through letters and a publication, but political will to eradicate this species has been weak and reports now indicate that it has expanded its distribution.

2008; Normile, 2009; Lamb, 2011), motivated by the need to mitigate greenhouse gas emissions (Drummond *et al.*, 2010), restore key habitats for species of conservation concern, and enhance the biodiversity value of human-modified landscapes (Gardner *et al.*, 2009). For example, the restoration of logged-over forests in Sabah has resulted in more rapid recovery of insectivorous bird populations than in naturally regenerating logged forest (Edwards *et al.*, 2009).

Ecological Restoration programs will require, among other things, tree nurseries that have adequate stocks of seedlings of native tree species. Many of the larger tree species (particularly among the Dipterocarpaceae) reproduce during infrequent community-wide "general flowering" events (masting). The dipterocarps, which comprise 10% of tree species and 80% of canopy species on the island of Borneo, have recalcitrant seeds. These biological constraints present a major challenge to collection of seeds and seedlings that form the basis of forest restoration projects across dipterocarp-dominated lowland forest (Kettle *et al.*, 2011a, b).

The knowledge base for implementing successful forest restoration has grown over several decades of silvicultural and ecological research in tropical Asia (Adjers *et al.*, 1995; Kettle, 2010). Dissemination of the existing knowledge via Internet-based virtual networks of restoration projects would enhance knowledge transfer and technical capacity. Development of transboundary phenology monitoring programs will ensure that flowering and mast fruiting events are capitalized to their full advantage. The existing networks of botanical gardens and Forest Research Institutes should also play a substantial role in capacity building at the local and regional scales.

Appendix

List of Courses

1. Conservation Biology
2. Managing Ecosystems for Resistance and Resilience
3. MA in Conservation Biology

See also: Agriculture, Sustainable. Biodiversity in Logged and Managed Forests. Biodiversity-Rich Countries. Climate Change and Extinctions. Deforestation and Land Clearing. Endangered Amphibians. Endangered Ecosystems. Endangered Mammals. Endangered Plants. Endangered Reptiles. Endangered Terrestrial Invertebrates. Human Impact on Biodiversity, Overview. Land-Use Issues. Mangrove Ecosystems. River Ecosystems. Threatened Birds. Tropical Forest Regeneration

References

Abrahamczyk S, Kessler M, Putra D, Waltert M, and Tscharntke T (2008) The value of differently managed cacao plantations for forest bird conservation in Sulawesi, Indonesia. *Bird Conservation International* 18: 349–362.

Achard F, Eva HD, Stibig H-J, *et al.* (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297: 999–1002.

Adeney JM, Ginsberg JR, Russell GJ, and Kinnaird MF (2006) Effects of an ENSO-related fire on birds of a lowland tropical forest in Sumatra. *Animal Conservation* 9: 292–301.

Adjers G, Hadengganan S, Kuusipalo J, Nuryanto K, and Vesa L (1995) Enrichment planting of dipterocarps in logged-over secondary forests: Effect of width, direction and maintenance method of planting line on selected Shorea species. *Forest Ecology and Management* 73: 259–270.

Ahmad Zafir AW, Azlan M, Lau CF, *et al.* (2011) Now or never: What will it take to save the Sumatran rhinoceros (*Dicerorhinus sumatrensis*) from extinction? *Oryx* 45: 225–233.

Allan JD and Flecker AS (1993) Biodiversity conservation in running waters. *BioScience* 43: 32–43.

Andam KS, Ferraro PJ, Sims KRE, Healy A, and Holland MB (2010) Protected areas reduced poverty in Costa Rica and Thailand. *Proceedings of the National Academy of Sciences of the United States of America* 107: 9996–10001.

Aratrakorn S, Thunhikorn S, and Donald PF (2006) Changes in bird communities following conversion of lowland forest to oil palm and rubber plantations in southern Thailand. *Bird Conservation International* 16: 71–82.

Auliya M (2010) Conservation status and impact of trade on the Oriental rat snake *Ptyas mucosa* in Java, Indonesia. *TRAFFIC Southeast Asia Report*. Petaling Jaya, Malaysia: TRAFFIC Southeast Asia.

Aziz SA, Laurance WF, and Clements R (2010) Forests reserved for rubber. *Frontiers in Ecology and Environment* 8: 178.

Baker PJ and Bunyavejchewin S (2009) Fire behavior and fire effects across the forest landscape of continental Southeast Asia. In: Cochrane MA (ed.) *Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamics*, pp. 311–334. Chichester, UK: Praxis Publishing.

BAPPENAS (2003) Indonesian Biodiversity Strategy and Action Plan (IBSAP). *Report of Badan Perencanaan Pembangunan Nasional*. Indonesia: The National Development Planning Agency Jakarta.

Baran E, Larinier M, Ziv G, and Marmulla G (2011) Review of the fish and fisheries aspects in the feasibility study and the environmental impact assessment of the proposed Xayaburi Dam on the Mekong mainstream. *Report prepared for the WWF Greater Mekong*. Gland, Switzerland: WWF International.

Barbier EB (1993) Habitat-fishery linkages and mangrove loss in Thailand. *Contemporary Economic Policy* 21: 59–77.

Beaman JH (2005) Mount Kinabalu: Hotspot of plant diversity in Borneo. (Biologiske Skrifter 55). In: Friis I and Balslev H (eds.) *Plant Diversity and Complexity Patterns: Local, Regional and Global Dimensions*, pp. 103–127. Copenhagen: Royal Danish Academy of Sciences and Letters.

Beaman JH and Burley JS (2003) Progress in the floristic inventory of Borneo. In: Padoch C and Peluso NL (eds.) *Borneo in Transition: People, Forests, Conservation, and Development*, pp. 93–113. Selangor, Malaysia: Oxford University Press.

Bennett EL (2011) Another inconvenient truth: The failure of enforcement systems to save charismatic species. *Oryx* 45: 476–479.
Bergstedt LC and Bergersen EP (1997) Health and movements of fish in response to sediment sluicing in the Wind River, Wyoming. *Canadian Journal of Fisheries and Aquatic Sciences* 54: 312–319.
Berkman HE and Rabeni CF (1987) Effect of siltation of stream fish communities. *Environmental Biology of Fishes* 18: 285–294.
Bernard H, Fjeldsa J, and Mohamed M (2009) A case study on the effects of disturbance and conversion of tropical lowland rain forest on the non-volant small mammals in north Borneo: Management implications. *Mammal study* 34: 85–96.
Bickford D, Howard SD, Ng DJJ, and Sheridan JA (2010) Impacts of climate change on the amphibians and reptiles of Southeast Asia. *Biodiversity and Conservation* 19: 1043–1062.
Bickford D, Iskandar D, and Barlian A (2008) A lungless frog discovered on Borneo. *Current Biology* 18: R374–R375.
Bielby J, Cooper N, Cunningham AA, Garner TW, and Purvis A (2008) Predicting susceptibility to future declines in the world's frogs. *Conservation Letters* 1: 82–90.
Braxton Little J (2008) *Regrowing Borneo's Rainforest – Tree by tree. Scientific American*. New York: Nature America, Inc.
Breed AC, Field HE, Epstein JH, Plowright RK, and Daszak P (2006) Emerging henipaviruses and flying foxes – Conservation and management perspectives. *Biological Conservation* 131: 211–220.
Briggs JC (2003) The biogeographic and tectonic history of India. *Journal of Biogeography* 30: 381–388.
Brook BW, Sodhi NS, and Bradshaw CJA (2008) Synergies among extinction drivers under global change. *Trends in Ecology and Evolution* 23: 453–460.
Brook BW, Sodhi NS, and Ng PKL (2003) Catastrophic extinctions follow deforestation in Singapore. *Nature* 424: 420–423.
Brook BW, Sodhi NS, Soh MCK, and Lim HC (2003) Abundance and projected control of invasive house crows in Singapore. *Journal of Wildlife Management* 67: 808–817.
Brooks TM, Mittermeier RA, da Fonseca GAB, *et al.* (2006) Global biodiversity conservation priorities. *Science* 313: 58–61.
Bruhl C, Eltz T, and Linsenmair K (2003) Size does matter: Effects of tropical rainforest fragmentation on the leaf litter ant community in Sabah. *Malaysia Biodiversity Conservation* 12: 1371–1389.
Bruner A, Gullison R, Rice R, and Da Fonseca G (2001) Effectiveness of parks in protecting tropical biodiversity. *Science* 291: 125–128.
Burkhead NM and Jelks H (2001) Effects of suspended sediment on the reproductive success of the tricolor shiner, a crevice-spawning minnow. *Transactions of the American Fisheries Society* 130: 959–968.
Butler RA, Koh LP, and Ghazoul J (2009) REDD in the red: Palm oil could undermine carbon payment schemes. *Conservation Letters* 2: 67–73.
Castelletta M, Sodhi NS, and Subaraj R (2000) Heavy extinctions of forest-dependent avifauna in Singapore: Lessons for biodiversity conservation in Southeast Asia. *Conservation Biology* 14: 1870–1880.
Castelletta M, Thiollay JM, and Sodhi NS (2005) The effects of extreme forest fragmentation on the bird community of Singapore Island. *Biological Conservation* 121: 135–155.
Chape S, Harrison J, Spalding M, and Lysenko I (2005) Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society of London: Biological Sciences* 360: 443–455.
Chazdon RL (2008) Beyond deforestation: Restoring forests and ecosystem services on degraded lands. *Science* 320: 1458–1460.
Chazdon RL, Peres CA, Dent D, *et al.* (2009) The potential for species conservation in tropical secondary forests. *Conservation Biology* 23: 1406–1417.
Chen H, Smith GJD, Zhang SY, *et al.* (2005) H5N1 virus outbreak in migratory waterfowl. *Science* 436: 191–192.
Chen IC, Shiu H-J, Benedick S, *et al.* (2009) Elevation increases in moth assemblages over 42 years on a tropical mountain. *Proceedings of the National Academy of Sciences of the United States of America* 106: 1479–1483.
Chua LSL, Suhaida M, Hamidah M, and Saw LG (2010) *Malaysia Plant Red List – Peninsular Malaysian Dipterocarpaceae*. Kuala Lumpur, Malaysia: FRIM.
Cleary DFR (2003) An examination of scale of assessment, logging and ENSO-induced fires on butterfly diversity in Borneo. *Oecologia* 135: 313–321.
Clements R, Koh LP, Lee TM, Meier R, and Li D (2006) Importance of reservoirs for the conservation of freshwater molluscs in a tropical urban landscape. *Biological Conservation* 128: 136–146.
Clements R, Mark Rayan D, Ahmad Zafir AW, *et al.* (2010a) Trio under threat: Can we secure the future of rhinos, elephants and tigers in Malaysia? *Biodiversity and Conservation* 19: 1115–1136.
Clements RG, Sayer J, Boedhihartono AK, *et al.* (2010b) Cautious optimism over Norway–Indonesia REDD pact. *Conservation Biology* 24: 1437–1438.
Clough Y, Dwi Putra D, Pitopang R, and Tscharntke T (2009) Local and landscape factors determine functional bird diversity in Indonesian cacao agroforestry. *Biological Conservation* 142: 1032–1041.
Corbet A and Pendlebury H (1992) *The Butterflies of the Malay Peninsula*, 4th edn. Kuala Lumpur, Malaysia: Malayan Nature Society.
Corlett R (2010) Invasive aliens of tropical East Asian islands. *Biodiversity and Conservation* 19: 411–423.
Corlett RT (1992) The ecological transformation of Singapore, 1819–1990. *Journal of Biogeography* 19: 411–420.
Corlett RT (1998) Frugivory and seed dispersal by vertebrates in the Oriental (Indomalayan) region. *Biological Reviews* 73: 413–448.
Corlett RT (2004) Flower visitors and pollination in the Oriental (Indomalayan) Region. *Biological Reviews* 79: 497–532.
Corlett RT (2007) The impact of hunting on the mammalian fauna of tropical Asian forests. *Biotropica* 39: 292–303.
Corvinus G (2004) *Homo erectus* in East and Southeast Asia, and the questions of the age of the species and its association with stone artifacts, with special attention to handaxe-like tolos. *Quaternary International* 117: 141–151.
Curran LM, Trigg SN, McDonald AK, *et al.* (2004) Lowland forest loss in protected areas of Indonesian Borneo. *Science* 303: 1000–1003.
D'Abrera B (1982) *Butterflies of the Oriental Region. Part I*. Melbourne, Australia: Hill House.
D'Abrera B (1985) *Butterflies of the Oriental Region. Part II*. Melbourne, Australia: Hill House.
D'Abrera B (1986) *Butterflies of the Oriental Region. Part III*. Melbourne, Australia: Hill House.
Danielsen F, Beukema H, Burgess ND, *et al.* (2008) Biofuel plantations on forested lands: Double jeopardy for biodiversity and climate. *Conservation Biology* 23: 348–358.
Danielsen F and Heegaard M (1995) Impact of logging and plantation development on species diversity: A case study from Sumatra. In: Sandbukt Ø (ed.) *Management of Tropical Forests: Towards an Integrated Perspective*, pp. 73–92. Oslo: University of Oslo.
Das I (2010) *A Field Guide to the Reptiles of Southeast Asia: Myanmar, Thailand, Laos, Cambodia, Vietnam, Peninsular Malaysia, Singapore, Sumatra, Borneo, Java, Bali*. London: New Holland Publishers.
Davies A, Holloway J, Huijbregts H, Krikken J, Kirk-Spriggs A, and Sutton S (2001) Dung beetles as indicators of change in the forests of northern Borneo. *Journal of Applied Ecology* 38: 593–616.
Dawson W, Mndolwa AS, Burslem D, and Hume PE (2008) Assessing the risks of plant invasions arising from collections in tropical botanical gardens. *Biodiversity and Conservation* 17: 1979–1995.
DeFries R, Hansen A, Newton AC, and Hansen MC (2005) Increasing isolation of protected areas in tropical forests over the past twenty years. *Ecological Applications* 15: 19–26.
Department of Environment Malaysia (2007) *Malaysia Environmental Quality Report 2007*. Petaling Jaya, Malaysia: Sasyaz Holdings.
Driscoll CT, Han Y-J, Chen CY, *et al.* (2007) Mercury contamination in forest and freshwater ecosystems in the northeastern United States. *BioScience* 57: 17–28.
Drummond SP, Wilson KA, Meijaard E, *et al.* (2010) Influence of a threatened-species focus on conservation planning. *Conservation Biology* 24: 441–449.
Dudgeon D (1983) The utilization of terrestrial plants as a food source by the fish stock of a gently sloping marginal zone in Plover Cove Reservoir, Hong Kong. *Environmental Biology of Fishes* 8: 73–77.
Dudgeon D (1992) Endangered ecosystems: A review of the conservation status of tropical Asian rivers. *Hydrobiologia* 248: 167–191.
Dudgeon D (1999) *Tropical Asian streams: Zoobenthos, Ecology, And Conservation*. Hong Kong: Hong Kong University Press.
Dudgeon D (2000a) The ecology of tropical Asian rivers and streams in relation to biodiversity conservation. *Annual Review of Ecology and Systematics* 31: 239–263.
Dudgeon D (2000b) Large-scale hydrological alterations in tropical Asia: Prospects for riverine biodiversity. *BioScience* 50: 793–806.
Dudgeon D, Arthington AH, Gessner MO, *et al.* (2006) Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews* 81: 163–182.
Dugan P (2008) Mainstream dams as barriers to fish migration: International learning and implications for the Mekong. *Catch and Culture* 14: 9–15.

Dugan PJ, Barlow C, Agostinho AA, *et al.* (2010) Fish migration, dams, and loss of ecosystem services in the Mekong basin. *Ambio* 2010: 344–348.

Dunn RR (2005) Modern insect extinction, the neglected majority. *Conservation Biology* 19: 1030–1036.

Edwards DP, Ansell FA, Ahmad AH, Nilus R, and Hamer KC (2009) The value of rehabilitating logged rainforest for birds. *Conservation Biology* 23: 1628–1633.

Edwards DP, Larsen TH, Docherty TDS, *et al.* (2011) Degraded lands worth protecting: The biological importance of Southeast Asia's repeatedly logged forests. *Proceedings of the Royal Society B: Biological Sciences* 278: 82–90.

FAO (2005) Global forest resources assessment 2005: Progress towards sustainable forest management. *FAO Forestry Paper 147*. Rome, Italy: Food and Agricultural Organization of the United Nations.

FAO (2007) The world's mangroves 1980–2005: A thematic study prepared in the framework of the global forest resources assessment 2005. *FAO Forestry Paper 153*. Rome, Italy: Food and Agricultural Organization of the United Nations.

FAO (2010) *Global Forest Resources Assessment 2010: Main Report*. Rome, Italy: Food and Agriculture Organization of the United Nations.

FAO (2011) *FAOSTAT: Statistical Databases and Data-Sets*. URL http://faostat.fao.org

Felbab-Brown N (2011) The disappearing act: The illicit trade in wildlife in Asia. *Reuters June* 2011.

Ferguson JW, Heaney M, Dugan P, and Barlow C (2011) Potential effects of dams on migratory fish in the Mekong River: Lessons from Salmon in the Fraser and Columbia Rivers. *Environmental Management* 47: 141–159.

Field HE and Mackenzie J (2004) Novel viral encephalitides associated with bats (Chiroptera) – Host management strategies. *Archives of Virology* 18: 113–121.

Fitzherbert EB, Struebig MJ, Morel A, *et al.* (2008) How will oil palm expansion affect biodiversity? *Trends in Ecology and Evolution* 23: 538–545.

Frodin D (2001) *Guide to the Standard Floras of the World*, 2nd edn. Cambridge, UK: Cambridge University Press.

Gardner TA, Barlow J, Chazdon R, *et al.* (2009) Prospects for tropical forest biodiversity in a human-modified world. *Ecology Letters* 12: 561–582.

Ghazoul J (2004) Alien abduction: Disruption of native plant-pollinator interactions by invasive species. *Biotropica* 36: 156–164.

Giam X, Clements RC, Aziz SAA, Chong KY, and Miettinen J (2011a) Rethinking the 'back to wilderness' concept for Sundaland's forests. *Biological Conservation*

Giam X, Ng TH, Lok AFSL, and Ng HH (2011b) Local geographic range predicts freshwater fish extinctions in Singapore. *Journal of Applied Ecology* 48: 356–363.

Giam X, Ng TH, Yap VB, and Tan HTW (2010a) The extent of undiscovered species in Southeast Asia. *Biodiversity and Conservation* 19: 943–954.

Giam X, Scheffers BR, Sodhi NS, Wilcove DS, Ceballos G, and Ehrlich PR (2011c) Reservoirs of richness: Least disturbed tropical forests are centres of undescribed species diversity. *Proceedings of the Royal Society B: Biological Sciences*

Giam XL, Ng TH, Yap V, and Tan HTW (2010b) The extent of undiscovered species in Southeast Asia. *Biodiversity and Conservation* 19: 943–954.

Gibson L, Lee TM, Koh LP, *et al.* (2011) Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* 478: 378–381.

Giesen W, Wulffraat S, Zieren M, and Scholten L (2006) *Mangrove Guidebook for Southeast Asia*. Bangkok: FAO and Wetlands International.

Gillison AN (2001) Vegetation Survey and Habitat Assessment of the Tesso Nilo Forest Complex. In: *WWF-US*, p. 76.

Hannah L, Midgley G, Andelman S, *et al.* (2007) Protected area needs in a changing climate. *Frontiers in Ecology and the Environment* 5: 131–138.

Harrison ME, Cheyne SM, Darma F, Ribowo DA, Limin SH, and Struebig MJ (2011) Hunting of flying foxes and perception of disease risk in Indonesian Borneo. *Biological Conservation* 144: 2441–2449.

Harrison RD (2000) Repercussions of El Nino: Drought causes extinction and the breakdown of mutualism in Borneo. *Proceedings of the Royal Society of London Series B: Biological Sciences* 267: 911–915.

Humphries P and Winemiller KO (2009) Historical impacts on river fauna, shifting baselines, and challenges for restoration. *BioScience* 59: 673–684.

Huong SL and Sodhi NS (1997) Stats of the Oriental Magpie Robin *Copsychus saularis* in Singapore. *Malayan Nature Journal* 50: 347–354.

Igarashi S and Fukuda H (1997) *The Life Histories of Asian Butterflies*, vol. I. Tokyo, Japan: Tokai University Press.

Igarashi S and Fukuda H (2000) *The Life Histories of Asian Butterflies*, vol. I. Tokyo, Japan: Tokai University Press.

Inger RF and Chin PK (1962) The fresh-water fishes of North Borneo. *Fieldiana Zoology* 45: 1–268.

IUCN (2011) IUCN Red List of Threatened Species 2011.1. URL www.iucnredlist.org

Iwata T, Nakano S, and Inoue M (2003) Impacts of past riparian deforestation on stream communities in a tropical rain forest in Borneo. *Ecological Applications* 13: 461–473.

Johns AG (1996) Bird population persistence in Sabahan logging concessions. *Biological Conservation* 75: 3–10.

Johnson LA and Dearden P (2009) Fire ecology and management of seasonal evergreen forests in mainland Southeast Asia. *Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamics* 289–310.

Jones D, Susilo F, Bignell D, Hardiwinoto S, Gillison A, and Eggleton P (2003) Termite assemblage collapse along a land-use intensification gradient in lowland central Sumatra, Indonesia. *Journal of Applied Ecology* 40: 380–391.

Jutagate T, Krudpan C, Ngamsnae P, Lamkom T, and Payooha K (2005) Changes in the fish catches during a trial opening of sluice gates on a run-of-the-river reservoir in Thailand. *Fisheries Management and Ecology* 12: 57–62.

Jutagate T, Lamkom T, Satapornwanit K, Naiwinit W, and Petchuay C (2001) Fish species diversity and ichthyomass in Pak Mun Reservoir, five years after impoundment. *Asian Fisheries Science* 14: 417–424.

Kettle CJ (2010) Ecological considerations for using dipterocarps for restoration of lowland rainforest in Southeast Asia. *Biodiversity Conservation* 19: 1137–1151.

Kettle CJ, Burslem DFRP, and Ghazoul J (2011a) An unorthodox approach to forest restoration. *Science* 333: 336.

Kettle CJ, Ghazoul J, Ashton P, *et al.* (2011b) Seeing the fruit for the trees in Borneo. *Conservation Letters* 4: 184–191.

Kingston T (2010) Research priorities for bat conservation in Southeast Asia: a consensus approach. *Biodiversity and Conservation* 19: 471–484.

Klorvuttimontara S, McClean CJ, and Hill JK (2011) Evaluating the effectiveness of protected Areas for conserving tropical forest butterflies of Thailand. *Biological Conservation* 144: 2534–2540.

Knöppel H-A (1970) Food of central Amazonian fishes: Contribution to the nutrient ecology of Amazonian rainforest streams. *Amazoniana* 2: 257–352.

Koh LP (2007) Impending disaster or sliver of hope for Southeast Asian forests? The devil may lie in the details. *Biodiversity Conservation* 16: 3935–3938.

Koh LP, Gibbs HK, Potapov P, and Hansen M (2011a) REDDcalculator.com: A web-based decision-support tool for implementing Indonesia's forest moratorium. *Methods in Ecology and Evolution* 3: 310–316.

Koh LP, Miettinen J, Liew SC, and Ghazoul J (2011b) Remotely sensed evidence of tropical peatland conversion to oil palm. *Proceedings of the National Academy of Sciences United States of America* 108: 5127–5132.

Koh LP and Sodhi NS (2004) Importance of reserves, fragments and parks for butterfly conservation in a tropical urban landscape. *Ecological Applications* 14: 1695–1708.

Koh LP and Wilcove DS (2008) Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letter* 1: 60–64.

Kottelat M and Whitten T (1996) Freshwater biodiversity in Asia, with special reference to fish. *World Bank Technical Paper No. 343*. Washington, DC: World Bank.

Kusrini MD, Skerratt LF, Garland S, Berger L, and Endarwin W (2008) Chytridiomycosis in frogs of Mount Gede Pangrango, Indonesia. *Inter-Research Disease of Aquatic Organisms* 82: 187–194.

Lamb D (2011) *Regreening the Bare Hills*. Dordrecht: Springer.

Laurance WF (1999) Reflections on the tropical deforestation crisis. *Biological Conservation* 91: 109–117.

Laurance WF (2007) Forest destruction in tropical Asia. *Current Science* 93: 1544–1550.

Laurance WF, Goosem M, and Laurance SGW (2009) Impacts of roads and linear clearings on tropical forests. *Trends in Ecology and Evolution* 19: 654–660.

Lee RJ, Gorog AJ, Dwiyahreni A, *et al.* (2005) Wildlife trade and implications for law enforcement in Indonesia: A case study from North Sulawesi. *Biological Conservation* 123: 477–488.

Lee TM and Jetz W (2008) Future battlegrounds for conservation under global change. *Proceedings of the Royal Society of London Series B: Biological Sciences* 275: 1261–1270.

Lee TM, Sodhi NS, and Prawiradilaga DM (2007) The importance of protected areas for the forest and endemic avifauna of Sulawesi (Indonesia). *Ecological Applications* 17: 1724–1741.

Lemly AD, Kingsford RT, and Thompson JR (2000) Irrigated agriculture and wildlife conservation: Conflict on a global scale. *Environmental Management* 25: 485–512.

Leprieur F, Beauchard O, Blanchet S, Oberdorff T, and Brosse S (2008) Fish invasions in the world's river systems: When natural processes are blurred by human activities. *PLoS Biology* 6: e28.

Linkie M, Haidir IA, Nugroho A, and Dinata Y (2008) Conserving tigers *Panthera tigris* in selectively logged Sumatran forests. *Biological Conservation* 141: 2410–2415.
Liow L, Sodhi NS, and Elmqvist T (2001) Bee diversity along a disturbance gradient in tropical lowland forests of south-east Asia. *Journal of Applied Ecology* 38: 180–192.
Lohman D, Bickford D, and Sodhi NS (2007) The burning issue. *Science* 316: 376.
Lorion CM and Kennedy BP (2009) Relationships between deforestation, riparian forest buffers and benthic macroinvertebrates in neotropical headwater streams. *Freshwater Biology* 54: 165–180.
Losos EC, Leigh J, Egbert, and Giles (eds.) (2004) *Tropical Forest Diversity and Dynamism: Findings From a Large-Scale Plot Network*. Chicago, IL: University of Chicago Press.
Luck GW, Chan KMA, and Fay JP (2009) Protecting ecosystem services and biodiversity in the world's watersheds. *Conservation Letters* 2: 179–188.
Lynam AJ (2010) Securing a future for wild Indochinese tigers: Transforming tiger vacuums into tiger source sites. *Integrative Zoology* 5: 324–334.
Maas B, Dwi Putra D, Waltert M, Clough Y, Tscharntke T, and Schulze CH (2009) Six years of habitat modification in a tropical rainforest margin of Indonesia do not affect bird diversity but endemic forest species. *Biological Conservation* 142: 2665–2671.
Macintosh DJ, Ashton EC, and Havanon S (2002) Mangrove rehabilitation and intertidal biodiversity: A study in the Ranong Mangrove Ecosystem. *Estuarine, Coastal and Shelf Science* 55: 331–345.
Maddox T, Priatna D, Gemita E, and Salampessy A (2007) The conservation of tigers and other wildlife in oil palm plantations. Jambi Province, Sumatra, Indonesia. *Conservation Report No. 7*. London: The Zoological Society of London (ZSL).
Marsh ST, Brummitt NA, de Kok RPJ, and Utteridge TMA (2009) Large-scale patterns of plant diversity and conservation priorities in South East Asia. *Blumea* 54: 103–108.
McShea WJ, Stewart C, Peterson L, *et al.* (2009) The importance of secondary forest blocks for terrestrial mammals within an Acacia/secondary forest matrix in Sarawak, Malaysia. *Biological Conservation* 142: 3108–3119.
MEA (2005) *Ecosystems and Human Well-being: Biodiversity Synthesis*. Washington, DC: Millennium Ecosystem Assessment. World Resources Institute.
Meijaard E, Sheil D, Marshall AJ, and Nasi R (2008) Phylogenetic age is positively correlated with sensitivity to timber harvest in Bornean mammals. *Biotropica* 40: 76–85.
Mekong River Commission (2007) Diagnostic study of water quality in the Lower Mekong Basin. In: *MRC Technical Paper No. 15*. Vientiane.
Mekong River Commission (2010) *State of the Basin Report 2010*. Vientiane.
Miettinen J, Shi C, and Liew SC (2011a) Deforestation rates in insular Southeast Asia between 2000 and 2010. *Global Change Biology*.
Miettinen J, Shi C, and Liew SC (2011b) Two decades of destruction in Southeast Asia's peat swamp forests. *Frontiers in Ecology and the Environment*.
Ministry of Natural Resources and Environment (2006) *The Environment Report of Vietnam 2006: The State of Water Environment in 3 River Basins of Cau, Nhue-Day, and Dong Nan River System*. Ha Noi, Vietnam: Ministry of Natural Resources and Environment.
Morley RJ (2000) *Origin and Evolution of Tropical Rain Forests*. Chichester, UK: John Wiley & Sons, Ltd.
Morrison JC, Sechrest W, Dinerstein E, Wilcove DS, and Lamoreux JF (2007) Persistence of large mammal faunas as indicators of global impacts. *Journal of Mammalogy* 88: 1363–1380.
Nabhan G and Buchmann S (1997) Services provided by pollinators. In: Daily GC (ed.) *Nature's Services: Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.
Nadler T, Vu NT, and Streicher U (2007) Conservation status of Vietnamese primates. *Vietnamese Journal of Primatology* 1: 7–26.
Newton P, Nguyen VT, Robertson S, *et al.* (2008) Pangolins in peril: Using local hunters' knowledge to conserve elusive species in Vietnam. *Endangered Species Research* 6: 41–53.
Ng J and Nemora (2007) *Tiger trade revisited in Sumatra, Indonesia*. Petaling Jaya, Malaysia: TRAFFIC Southeast Asia.
Nguyen TT and de Silva SS (2006) Freshwater finfish biodiversity and conservation: An Asian perspective. *Biodiversity and Conservation* 15: 3543–3568.
Nijman V (2005) *Hanging in the Balance: An Assessment of Trade in Orang-Utans and Gibbons in Kalimantan, Indonesia*. Petaling Jaya, Malaysia: TRAFFIC Southeast Asia.
Nijman V (2010) An overview of international wildlife trade from Southeast Asia. *Biodiversity and Conservation* 19: 1101–1114.
Nijman V and Shepherd C (2007) Trade in non-native, CITES-listed, wildlife in Asia, as exemplified by the trade in freshwater turtles and tortoises (*Chelonidae*) in Thailand. *Contributions to Zoology* 76: 207–212.
Normile D (2009) Restoring a 'biological desert' on Borneo. *Science* 325: 557.
Normile D (2010) Saving forest to save biodiversity. *Science* 329: 1278–1280.
Nyhus P and Tilson R (2004) Agroforestry, elephants, and tigers: Balancing conservation theory and practice in human-dominated landscapes of Southeast Asia. *Agriculture, Ecosystems and Environment* 104: 87–97.
Osunkoya OO, Othman FE, and Kahar RS (2005) Growth and competition between seedlings of an invasive plantation tree, *Acacia mangium*, and those of a native Borneo heath-forest species. *Melastoma beccarianum*. *Ecological Research* 20: 205–214.
Pacini N and Harper DM (2008) Aquatic, semi–aquatic and riparian vertebrates. In: Dudgeon D (ed.) *Tropical Stream Ecology*, pp. 147–198. London: Academic Press.
Page S, Hoscilo A, Langner A, *et al.* (2009) Tropical peatland fires in Southeast Asia. *In Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamics* 263–287.
Page SE, Rieley JO, Shotyk OW, and Weiss D (1999) Interdependence of peat and vegetation in a tropical peat swamp forest. *Philosophical Transactions of the Royal Society B: Biological Sciences* 354: 1885–1897.
Pantel S and Chin SY (2010) *Proceedings of the Workshop on Trade and Conservation of Pangolins Native to South and Southeast Asia*. Singapore Zoo, Singapore, 30 June–2 July 2008, p. 277. Petaling Jaya, Malaysia: TRAFFIC Southeast Asia (www.traffic.org/species-reports/traffic_species_mammals51.pdf)
Patz JA, Daszak P, Tabor GM, *et al.* (2004) Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. *Environmental Health Perspectives* 112: 1092–1098.
Pauwels OSG, Bauer AM, Sumontha M, and Chanhome L (2004) *Cyrtodactylus thirakhupti* (Squamata: Gekkonidae), a new cave-dwelling gecko from southern Thailand. *Zootaxa* 772: 1–11.
Peh KSH (2007) Potential effects of climate change on elevational distributions of tropical birds in Southeast Asia. *Condor* 109: 437–441.
Peh KS-H (2010) Invasive species in Southeast Asia: The knowledge so far. *Biodiversity and Conservation* 19: 1083–1099.
Peh KS-H, de Jong J, Sodhi NS, Lim SLH, and Yap CA-M (2005) Lowland rainforest avifauna and human disturbance: Persistence of primary forest birds in selectively logged forest and mixed-rural habitats of southern Peninsular Malaysia. *Biological Conservation* 123: 489–505.
Peh KSH, Sodhi NS, de Jong J, Sekercioglu CH, Yap CAM, and Lim SLH (2006) Conservation value of degraded habitats for forest birds in southern Peninsular Malaysia. *Diversity and Distributions* 12: 572–581.
Peh KS-H, Soh MCK, Sodhi NS, Laurance WF, Ong DJ, and Clements R (2011) Up in the clouds: Is sustainable use of tropical montane cloud forests possible in Malaysia? *Bioscience* 61: 27–38.
Peterjohn WT and Correll DL (1984) Nutrient dynamics in an agricultural watershed: Observations on the role of a riparian forest. *Ecology* 65: 1466–1475.
Peters HA (2001) Clidemia hirta invasion at the Pasoh Forest Reserve: An unexpected plant invasion in an undisturbed tropical forest. *Biotropica* 33: 60–68.
Philpott SM, Bichier P, Rice RA, and Greenberg R (2008) Biodiversity conservation, yield, and alternative products in coffee agroecosystems in Sumatra, Indonesia. *Biodiversity and Conservation* 17: 1805–1820.
Phoonjampa R and Brockelman WY (2008) Survey of pileated gibbon Hylobates pileatus in Thailand: Populations threatened by hunting and habitat degradation. *Oryx* 42: 600–606.
Poff NL, Allan JD, Bain MB, *et al.* (1997) The natural flow regime: A paradigm for river conservation and restoration. *BioScience* 52: 659–668.
Poff NL and Hart DD (2002) How dams vary and why it matters for the emerging science of dam removal. *BioScience* 52: 659–668.
Poffenberger M (2009) Cambodia's forests and climate change: Mitigating drivers of deforestation. *Natural Resources Forum* 33: 285–296.
Polprasert C (1982) Heavy metal pollution in the Chao Phraya River estuary, Thailand. *Water Research* 16: 775–784.
Posa MRC and Sodhi NS (2006) Effects of anthropogenic land use on forest birds and butterflies in Subic Bay, Philippines. *Biological Conservation* 129: 256–270.
Posa MRC, Wijedasa LS, and Corlett RT (2011) Biodiversity and conservation of tropical peat swamp forests. *BioScience* 61: 49–57.
Prance GT, Beentje HJD, and Johns R (2000) The tropical flora remains under collected. *Annals of the Missouri Botanical Gardens* 87: 67–71.
Ramírez A, Pringle CM, and Wantzen KA (2008) Tropical stream conservation. In: Dudgeon D (ed.) *Tropical Stream Ecology*, pp. 285–304. London: Academic Press.
Rao M, Rabinowitz M, and Khaing ST (2002) Status review of the protected-area system in Myanmar, with recommendations for conservation planning. *Conservation Biology* 16: 360–368.
Rayan MD and Shariff WM (2009) The importance of selectively logged forests for tiger conservation: A population density estimate of tigers *Panthera tigris* in Peninsular Malaysia. *Oryx* 43: 1–4.

Rickart EA, Balete DS, Rowe RJ, and Heaney LR (2011) Mammals of the northern Phillipines: Tolerance for habitat disturbance and resistance to invasive species in an endemic insular fauna. *Diversity and Distributions* 17: 530–541.
Roberts TR (2001) On the river of no returns: Thailand's Pak Mun Dam and its fish ladder. *Natural History Bulletin of Siam Society* 49: 189–230.
Rodrigues ASL, Pilgrim J, Lamoreux J, Hoffman M, and Brooks TM (2006) The value of the IUCN Red List for conservation. *Trends in Ecology and Evolution* 21: 71–76.
Rowley J, Brown R, Bain R, *et al.* (2010) Impending conservation crisis for Southeast Asian amphibians. *Biology Letters* 6.
Rudel TK, DeFries R, Asner GP, and Laurance WF (2009a) Changing drivers of deforestation and new opportunities for conservation. *Conservation Biology* 23: 1396–1405.
Rudel TK, Schneider L, Uriarte M, *et al.* (2009b) Agricultural intensification and changes in cultivated areas, 1970–2005. *Proceedings of the National Academy of Sciences of the United States of America* 106: 20675–20680.
Saatchi SS, Harris NL, Brown S, *et al.* (2011) Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America* 108: 9899–9904.
Sakagami S, Inoue T, and Salmah S (1990) Stingless bees of Sumatra. In: Sakagami S, Ohgushi R, and Roubik D (eds.) *Natural History of Social Wasps and Bees in Equatorial Sumatra*. Sapporo, Japan: Hokkaido University Press.
Samoilys MA, Martin-Smith KM, Giles BG, *et al.* (2007) Effectiveness of five small Philippines' coral reef reserves for fish population depends on site-specific factors, particularly enforcement history. *Biological Conservation* 136: 584–601.
Sanderbrook C, Nelson F, Adams WM, and Agrawal A (2010) Carbon, forests and the REDD paradox. *Oryx* 44: 330–334.
Sanderson BL, Barnas KA, and Rub AMW (2009) Nonindigenous species of the Pacific Northwest: An overlooked risk to endangered salmon? *BioScience* 59: 245–256.
Shanker K and Pilcher NJ (2003) Marine turtle conservation in south and southeast Asia. *Marine Turtle Newsletter* 100: 43–51.
Shepherd C and Nijman V (2007) The trade in bear parts from Myanmar: An illustration of the ineffectiveness of enforcement of international wildlife trade regulations. *Biodiversity and Conservation* 17: 35–42.
Shepherd C and Nijman V (2008) *Elephant and Ivory Trade in Myanmar*. Petaling Jaya, Malaysia: TRAFFIC Southeast Asia.
Shepherd C, Sukumaran J, and Wich SA (2004) *Open Season: An Analysis of the Pet Trade in Medan, Sumatra 1997–2001*. Petaling Jaya, Malaysia: TRAFFIC Southeast Asia.
Schipper J, Chanson JS, Chiozza F, *et al.* (2008) The status of the world's land and marine mammals: Diversity, threat, and knowledge. *Science* 322: 225–230.
Simberloff D (2010) Invasive species. In: Sodhi NS and Erlich PR (eds.) *Conservation Biology for All*. New York: Oxford University Press.
Sodhi NS and Brook BW (2006) *Southeast Asian Biodiversity in Crisis*. Cambridge, UK: Cambridge University Press.
Sodhi NS, Butler R, Laurance WF, and Gibson L (2011) Conservation successes at micro-, meso-, and macroscales. *Trends in Ecology and Evolution* 26: 585–594.
Sodhi NS, Koh LP, Brook BW, and Ng PKL (2004) Southeast Asian biodiversity: An impending disaster. *Trends in Ecology and Evolution* 19: 654–660.
Sodhi NS, Koh LP, Prawiradilaga DM, *et al.* (2005b) Land use and conservation value for forest birds in central Sulawesi (Indonesia). *Biological Conservation* 122: 547–558.
Sodhi NS, Lee TM, Koh LP, and Brook BW (2009) A meta-analysis of the impact of anthropogenic forest disturbance on Southeast Asia's biotas. *Biotropica* 41: 103–109.
Sodhi NS, Lee TM, Koh LP, and Dunn RR (2005c) A century of avifaunal turnover in a small tropical rainforest fragment. *Animal Conservation* 8: 217–222.
Sodhi NS, Lee TM, Koh LP, and Prawiradilaga DM (2006a) Long-term avifaunal impoverishment in an isolated tropical woodlot. *Conservation Biology* 20: 772–779.
Sodhi NS, Posa MRC, Lee TM, Bickford D, Koh LP, and Brook BW (2010a) The state and conservation of Southeast Asian biodiversity. *Biodiversity and Conservation* 19: 317–328.
Sodhi NS, Soh MCK, Prawiradilaga DM, Darjono, and Brook BW (2005a) Persistence of lowland forest birds in a recently logged area in central Java. *Bird Conservation International* 15: 173–191.
Sodhi NS, Wilcove DS, Lee TM, *et al.* (2010b) Deforestation and avian extinction on tropical landbridge islands. *Conservation Biology* 24: 1290–1298.
Songer M, Aung M, Senior B, DeFries R, and Leimgruber P (2009) Spatial and temporal deforestation dynamics in protected and unprotected dry forests: A case study from Myanmar (Burma). *Biodiversity and Conservation* 18: 1001–1018.
South N and Wyatt T (2011) Comparing illicit trades in wildlife and drugs: An exploratory study. *Deviant Behavior* 32: 538–561.
Spalding M, Kainuma M, and Collins L (2010) *World Atlas of Mangroves*. London: Earthscan.
Steinmetz R, Chutipong W, Seuaturien N, Chirngsaard E, and Khaengkhetkarn M (2010) Population recovery patterns of Southeast Asian ungulates after poaching. *Biological Conservation* 143: 42–51.
Struebig MJ, Christy L, Pio D, and Meijaard E (2010) Bats of Borneo: Diversity, distributions and representation in protected areas. *Biodiversity and Conservation* 19: 449–469.
Stuart SN, Chanson JS, Cox NA, *et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786.
Styring A and Ickes K (2001) Woodpecker abundance in a logged (40 years ago) vs unlogged lowland dipterocarp forest in Peninsular Malaysia. *Journal of Tropical Ecology* 17: 261–268.
Sutherland AB (2007) Effects of increased suspended sediment on the reproductive success of an upland crevice-spawning minnow. *Transactions of the American Fisheries Society* 136: 416–422.
Symington CF (2004) Foresters' Manual of Dipterocarps (revised). *Malayan Forest Records*. Kuala Lumpur, Malaysia: Forest Research Institute Malaysia and Malaysian Nature Society.
Tan SK and Clements RG (2011) *Limicolaria flammea*: Another potentially invasive African land snail in tropical Asia. *Tropical Conservation Science* 4: 97–102.
Taylor D (2010) Biomass burning, humans and climate change in Southeast Asia. *Biodiversity and Conservation* 19: 1025–1042.
Thailand State of Pollution Report Working Group (2011) *Thailand State of Pollution Report 2009*.
The World Bank (2005) Going, Going, Gone: The Illegal Trade in Wildlife in East and Southeast Asia. *Environment and Social Development East Asia and Pacific Region Discussion Paper*. Washington, DC: The World Bank.
Then S (2010) Illegal wildlife traders making a killing. *The Star* 7th October 2010.
Todd PA, Ong X, and Chou LM (2010) Impacts of pollution on marine life in Southeast Asia. *Biodiversity and Conservation* 19: 1063–1082.
Trainor CR (2007) Changes in bird species composition on a remote and well forested Wallacean Island, South-East Asia. *Biological Conservation* 140: 373–385.
UNFCCC (2010) Draft decision (-/CP.16): Outcome of the work of the ad hoc working group on long-term cooperative action under the convention. Available at http://unfccc.int/files/meetings/cop_16/application/pdf/cop16_lca.pdf. Accessed 22 October 2011.
Venter O, Meijaard E, Possingham H, *et al.* (2009) Carbon payments as a safeguard for threatened tropical mammals. *Conservation Letters* 2: 123–129.
Waggener TR and Lane C (1997) Pacific rim demand and supply situation, trends and prospects: Implications for forest products trade in the asia-pacific region. International Forestry Sector Analysis (IFSA), Seattle, WA. *Report Prepared for FAO Asia Pacific Forestry Sector Outlook Study*. Working Paper APFSOS/WP/02. Rome, Italy: Food and Agriculture Organization of the United Nations.
Waltert M, Mardiastuti A, and Mühlenberg M (2004) Effects of land use on bird species richness in Sulawesi, Indonesia. *Conservation Biology* 18: 1339–1346.
Wanger TC, Iskandar DT, Motzke I, *et al.* (2010) Effects of land-use change on community composition of tropical amphibians and reptiles in Sulawesi, Indonesia. *Conservation Biology* 24: 795–802.
Wells K, Kalko EKV, Lakim MB, *et al.* (2007) Effects of rain forest logging on species richness and assemblage composition of small mammals in Southeast Asia. *Journal of Biogeography* 34: 1087–1099.
Whitmore TC (1984) *Tropical Rain Forests of the Far East*, 2nd edn. Oxford: Oxford University Press.
Whitten T, Damanik SJ, Anwar J, and Hisyam N (2000) *The Ecology of Sumatra*. Singapore: Periplus Editions.
Williams RN (2006) *Return to the River: Restoring Salmon to the Columbia River*. New York: Academic Press.
Willott SJ, Lim DC, Compton SG, and Sutton SL (2000) Effects of selective logging on the butterflies of a Bornean rainforest. *Conservation Biology* 14: 1055–1065.
Wilson EO (1987) The little things that run the world (the importance and conservation of invertebrates). *Conservation Biology* 1: 344–346.
Winemiller KO, Agostinho AA, and Caramaschi EP (2008) Fish ecology in tropical streams. In: Dudgeon D (ed.) *Tropical Stream Ecology*, pp. 107–146. London: Academic Press.
Wong M (1986) Trophic organization of understory birds in a Malaysian dipterocarp forest. *Auk* 103: 100–116.
Wootton JT, Parker MS, and Power ME (1996) Effects of disturbance on river food webs. *Science* 273: 1558–1561.

Wright SJ, Sanchez-Azofeifa GA, Portillo-Quintero C, and Davies D (2007) Poverty and corruption compromise tropical forest reserves. *Ecological Applications* 17: 1259–1266.

Wurstera CM, Bird MI, Bull ID, *et al.* (2010) Forest contraction in north equatorial Southeast Asia during the Last Glacial Period. *Proceedings of the National Academy of Sciences of the United States of America* 107: 15508–15511.

www.cbd.int/gspc/ (2010) Global strategy for plant conservation. URL www.cbd.int/gspc/

Yap CA-M, Sodhi NS, and Peh KS-H (2007) Phenology of tropical birds in Peninsular Malaysia: Effects of selective logging and food resources. *Auk* 124: 945–961.

Yule CM (2010) Loss of biodiversity and ecosystem function in Indo-Malayan peat swamp forests. *Biodiversity and Conservation* 19: 393–409.

Yule CM, Boyero L, and Marchant R (2010) Effects of sediment pollution on food webs in a tropical river (Borneo, Indonesia). *Marine and Freshwater Research* 61: 204–213.

Zeuner PE (1941) Geology, climate and faunal distribution in the Malay Archipelago. *Proceedings of the Royal Entomological Society of London* 16: 10–12.

Ziegler AD, Negishi S, Sidle RC, *et al.* (2005) Reduction of stream sediment concentraton by a riparian buffer: Filtering of road runoff in disturbed headwater basins of montane mainland Southeastern Asia. *Journal of Environmental Quality* 35: 151–162.

Biofuels and Biodiversity: The Implications of Energy Sprawl

Bruce A Robertson, Smithsonian Conservation Biology Institute, Washington, DC, USA, and Michigan State University, East Lansing, MI, USA
Patrick J Doran, The Nature Conservancy, Lansing, MI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Agricultural intensification The use of advanced tools and techniques, including pesticide and fertilizer applications and the removal of hedgerows that separate fields, that allow for an increasing amount of food or biomass to be produced from a given amount of land.
Agroenergy The production of biofuels as plant crops in agriculture, as distinct from biofuels produced synthetically or from algal culture.
Biofuel Any fuel that is in some way derived from biomass, including firewood, methane, and petroleum. Globally, the production of plant-derived liquid fuels for use in transportation is the primary focus of biofuel production.
Cellulosic ethanol production Ethanol production that employs the conversion of cellulose, the primary structural molecule of most plants, into glucose that can then be fermented into ethanol.
Cropping system An integrated system of agricultural production, including the crop selection, management, harvest, distribution, processing and marketing.
Energy sprawl An increase in the agricultural land area required for the production of dedicated biofuel crops that replaces food crops, natural and seminatural habitats, and other land uses, occurring directly from land-use change or indirectly due to biofuel-related market conditions.
Feedstock The plants or plant-derived products (e.g., pellets) that can be used to produce biofuels. *First-generation* feedstocks are those that are currently under wide cultivation for the production of ethanol and biodiesel, including corn, soy, sugarcane and palm oil. *Second-generation* feedstocks (e.g., switchgrass) are those crops that are currently not widely cultivated for biofuels but have high potential yields.

"Biofuels and Biodiversity: The Implications of Energy Sprawl" refers to the potential trade-off between the increase in development of agricultural crops for biofuels and the ability of agricultural landscapes to support biological diversity. Evaluations of the impacts of biofuels on biodiversity have received much attention in recent years due to an increasing demand for the development of alternative energy sources focusing on the fundamental question of whether the development of biofuels is at odds with the conservation of biodiversity.

The Promise of Biofuels

Although humans have relied on biofuels like firewood for nearly half a million years, today the term is primarily used to refer to plant-derived liquid fuels, like ethanol, that are most commonly used for transportation. Biofuels, including their development and production, have become a central component of the national energy strategies of many countries, especially in North America and Europe, because they have shown promise for augmenting energy security, supplementing existing energy supplies, and providing a new source of income for rural economies, all with a reduced environmental impact when compared with fossil fuels (see Figure 1). And even though the balance of greenhouse gas emissions associated with biofuel production has become a major environmental concern, liquid biofuels have already become a significant source of transportation fuel supplies in many countries due to production subsidies, import tariffs, and other legislation enacted by producing countries. The US, for example, enacted the Energy Independence and Security Act of 2007 that mandated that the volume of renewable fuel blended into transportation fuel should increase from 9 billion gallons in 2008 to 36 billion gallons by 2022. Interest and investment in biofuels is occurring on a global scale.

It is the production of liquid biofuels from plants grown as dedicated agricultural crops (agroenergy) that has the greatest potential to impact biodiversity on a global scale. More than 30 plant species have been proposed or are currently grown as bioenergy crops. Those more commonly mentioned are listed in Table 1. The most economically important bioenergy crops in the temperate zone are corn, soybeans, and canola – annually sewn plant monocultures. In tropical regions, biofuel production is dominated by sugarcane and African oil palm, a perennial tree whose biodiversity consequences are given more extensive treatment in one of the chapters of this encyclopedia. Yet, so-called cellulosic biofuel technology is projected to expand the range of plant species (i.e., second-generation feedstocks) that can be used to produce biofuels. By facilitating the use of perennial native plants that can be grown in polycultures and that require less crop management activities and chemical inputs, second-generation biofuels may provide more stable, seminatural environments with significant potential to maintain biodiversity in agricultural landscapes.

Determining whether or not growth in bioenergy can or will benefit biodiversity, help mitigate ongoing biodiversity losses, or only contribute to ongoing population declines in populations of native species depends on a number of factors. These include the relative suitability of biomass feedstocks as habitat, the relative habitat value of crops and other land uses that biomass crops replace, the impacts of practices used to grow and harvest crops, and the importance of landscape context in shaping local habitat quality. Therefore, we summarize the predicted biodiversity responses to three aspects of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00341-5

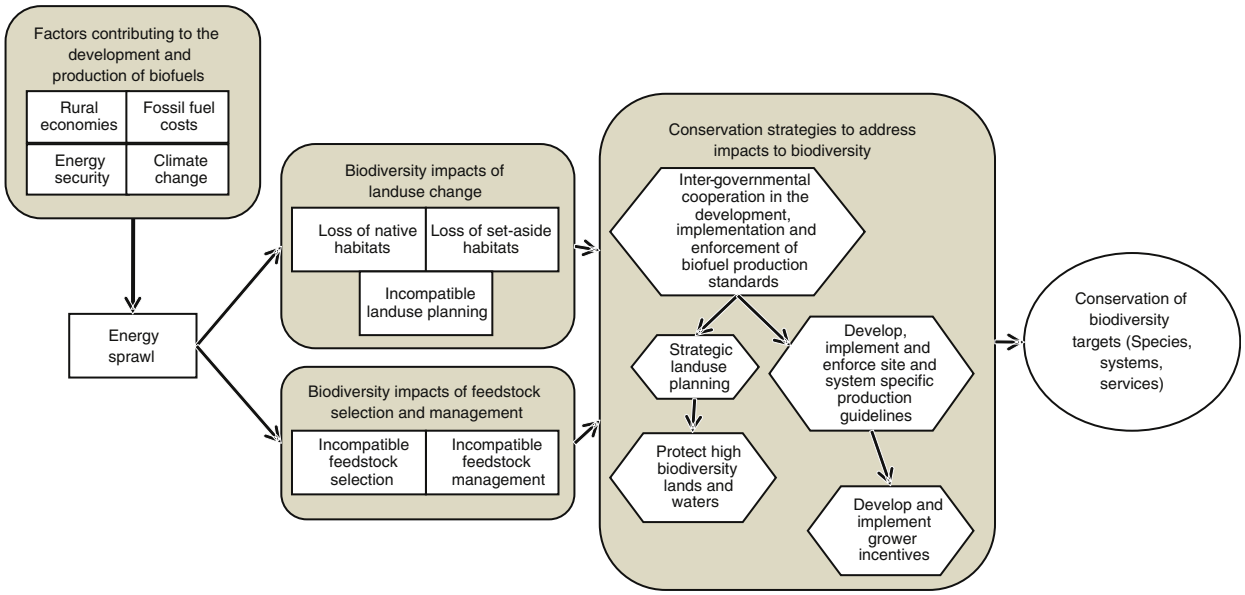


Figure 1 Conceptual model for the implementation of conservation strategies to abate the negative impacts of energy sprawl on biodiversity. A complex set of drivers will ultimately determine which crops are selected for production in various agricultural regions, how they will be managed, and what land-use impacts will follow (left side). Conservation strategies designed to mitigate the impacts of agroenergy production to biodiversity targets can focus on crop selection and management or agroenergy-driven land-use change (right-center). Broad cooperation will be required to development productions standards designed to meet conservation goals for biodiversity (far right). Guidelines for the production of agroenergy crops and biodiversity monitoring must be tied to strategic land-use planning within the context of producer incentive programs while incorporating provisions for the conservation of critical habitats. The impact of conservation strategies on feedstock selection and management and its direct and indirect impacts on agroenergy-related land-use change will need to be monitored for its effectiveness in meeting conservation goals. Failure to meet goals can lead to revisions of conservation strategies within an adaptive management framework.

Table 1 Candidate plant species for dedicated agricultural biofuel production

Annuals	*Perennials*
Corn	Willow
Sunflower	Hybrid poplar
Canola (rapeseed)	Castor bean
Barley	Sweetgum
Sunflower	Reed canary grass
Sugarbeet	Miscanthus
Camelina	Switchgrass
Peanut	Jatropha
Lesquerella	Babassou palm
Mustard	African oil palm
Alfalfa	Sugarcane[a]
Cuphea	

[a]Requires multiple plantings.

bioenergy expansion: (1) which crops are selected for production, (2) how crops are managed, and (3) how land-use changes at landscape and regional scales affect local biodiversity. Each factor may have both positive at negative impacts on biofuel production and biodiversity conservation.

Biodiversity Responses to Crop Selection and Management

Within the context of field-grown biofuel crops, biodiversity may comprise the plants used as feedstocks themselves, plant species able to successfully coexist with crops (e.g., weeds), soil microorganisms that cycle nutrients and decompose organic matter, and the various forms of vertebrate and invertebrate animal life that can utilize biofuel crops as habitat. Research into biodiversity responses to existing and new types of biofuel crops has lagged behind their production. There is a distinct lack of information demonstrating the relative biodiversity value of alternative bioenergy crops and how various management activities mitigate or exaggerate any differences that might influence decisions about crop selection that incorporate biodiversity concerns. As such, our summary of biodiversity responses reflects an inherently incomplete understanding of biodiversity responses to bioenergy production informed by general principals of biodiversity responses to agriculture in general.

Biodiversity responses to the intensification of agricultural practices are generally negative (Matson *et al.*, 1997; Donald *et al.*, 2001; Tscharntke *et al.*, 2005). As crops are sewn more densely, leaving fewer hedgerows and adjacent natural or seminatural habitats, and farmers increase the frequency of disturbance to cultivated lands by applying fertilizers and herbicides and increasing the frequency of harvest, fewer species are able to persist in agricultural fields. Because plant species' diversity is also linked to animal species' richness both within agricultural settings and among more natural habitats (Haddad *et al.*, 2001, 2009), vegetation diversity within a field and at local and landscape scales should influence biodiversity responses. Both the relative habitat stability (annual vs. perennial) and floristic diversity of a bioenergy crop should then

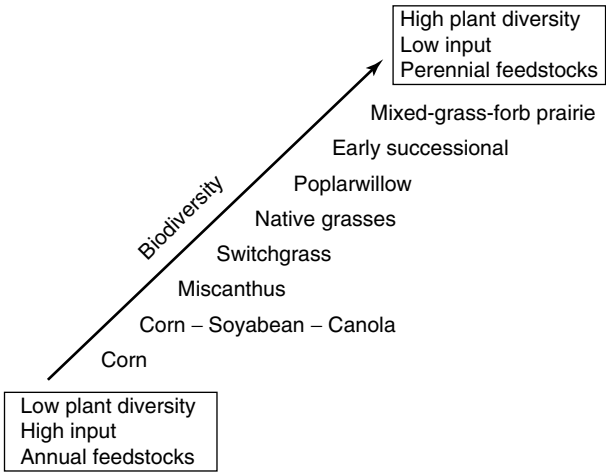


Figure 2 The relationship between crop plant special richness, perenniality, and biodiversity. Monocultures of plant species that must be replanted annually and that require higher chemical inputs generally support lower local biodiversity. Crop rotations reduce the need for inputs. Longer-lived woody species are not harvested annually. Biodiversity should be highest in bioenergy crops consisting of diverse, largely perennial plants that experience less disturbance associate with tilling or the application of chemical inputs.

influence biodiversity responses. Selection of crops across these axes is expected to have predictable consequences for biodiversity as dedicated bioenergy crops enter agricultural landscapes (**Figure 2**).

First-Generation Biofuel Crops

Many of the most economically important biofuel crops in the world (e.g., corn, soybeans, and canola) are annual monocultures that must be resewn by farmers each year (**Table 1**). Corn is the dominant bioenergy crop throughout much of North America, and its production has grown dramatically in response to US legislation financially encouraging its production and the blending of ethanol with gasoline. Ethanol is produced from corn through fermentation of the starchy portion of the corn kernel. It is commonly rotated with soybeans, a feedstock also used for the production of biodiesel.

Feedstock Selection

Corn is commonly rotated with soybeans, and this rotation system supports a relatively species-poor biological community (Gardiner *et al.*, 2010; Fletcher *et al.*, 2011). Only a few common species of breeding birds commonly nest in cornfields. Mammal and arthropod diversity and abundance are similarly reduced in corn and soybean when compared to other seminatural habitats (e.g., set-aside lands) and alternative crops. As a rule, species that are most likely to be found breeding in corn and soybean fields are generalists with relatively wide geographic ranges and stable populations and subsequently of low conservation priority.

Canola, or rapeseed, is a flowering member of the Brassicacea family grown in monoculture primarily for its seed, which has been historically used to produce cooking oil and as animal feed. Its high oil content (38–39%) in comparison to soybeans (18–19%) has made it an increasingly important feedstock for the production of biodiesel. This crop is increasingly common in China, Canada, and Europe, which produces more than 3/4 of the world's supply. Despite its wide geographic extent and economic importance, canola's ability to support biodiversity relative to other crops and habitats has not been well quantified.

Feedstock Management

Crop management can profoundly affect biodiversity responses independently of the type of biofuel crop selected for production. It is the intensity with which these annual crops are managed that negatively impacts their inherent ability to support biodiversity. Perhaps the two most important types of local management that will influence the magnitude and direction of effects of biofuel crops on biodiversity are the application of chemicals and the harvest regime (timing, type, and frequency of harvesting) (Fletcher *et al.*, 2011).

Expansion of corn, soybean, and canola production resulting from energy sprawl is almost certain to result in increased use of pesticides and fertilizers, which is, for example, the single largest input for US corn production (Hill *et al.*, 2006). Fertilizer runoff causes nutrient eutrophication in streams and rivers that reduces habitat availability and diversity for native organisms and directly kills them by creating anoxic conditions. Most notable among these is the great "dead zone" in the Gulf of Mexico, in which nutrient-enriched runoff from the Mississippi river results in oxygen depletion of seawater, killing most sessile and slow-moving organisms in a near-shore area of 13,000–21,000 km^2 (the size of the US state of New Jersey). Fish have some ability to swim away from anoxic areas, but anoxia and other consequences of eutrophication from runoff have resulted in reproductive problems in fish that involve decreased size of reproductive organs, low egg counts, and lack of spawning. Consequently, fertilizer use associated with the production of annual biofuel crops represents an important biodiversity threat to aquatic and marine ecosystems.

An estimated 70–90% of ground-applied pesticides and only 25–50% of aerially applied pesticides actually reach their target area (WWF, 1999). Pesticides have direct effects on target and nontarget organisms, including native bees and hoverflies, which are important pollinators of both crops and native plants (Kearns and Inouye, 1997). Endocrine disruptors impact bird, reptile, fish, and amphibian populations through their ability to cause physical abnormalities (WWF, 1999; Guillette *et al.*, 1999), reduce fertility, compromise immune systems, and trigger metabolic and behavioral abnormalities. Magnification of these effects through food chains can lead to community-level drops in biomass and diversity that may be irreversible in the near term (WWF, 1999) and that extend far beyond the borders of feedstock plantations.

Genetically modified herbicide-tolerant (GMHT) varieties of corn, soybean, and canola of these crops are in widespread use and have further reduced the biodiversity value of these feedstocks. GMHT varieties impact the richness and abundance of plant species in and around arable fields due to the efficacy of the herbicides applied in controlling weeds and reducing their seed banks (Firbank *et al.*, 2008). Loss of plant diversity leads to reduced insect diversity and longer-term reductions of weed populations, which in turn affect animal

populations higher up the arable food chain by altering the quality and quantity of forage resources (e.g., Watkinson *et al.*, 2000). In this way, maximizing crop production through elimination of fast-growing plant competitors creates a bottom-up trophic cascade that reduces the diversity and abundance of organisms at higher trophic levels.

Machinery used in chemical applications and in the harvest of crops can directly impact plants and animals in crop fields. Bird nests are commonly trampled or destroyed by passing machinery and adults may be killed as well when they are caught unaware or are incubating eggs or protecting young. Mortality and reproductive failure is also a problem for a host of mammal and arthropod species and lead to local population declines. These losses can easily lead to population declines. By eliminating the use of pesticides and reducing disturbance the use of conservation tillage, cover crops and organic farming regimes for corn and other crops can increase the diversity and reproductive success of birds and otherwise enhance biodiversity (Warburton and Klimstra, 1984; Lokemoen and Beiser, 1997; Beecher *et al.*, 2002; Henderson *et al.*, 2004; Hole *et al.*, 2005) and may be useful approaches in the production of bioenergy crops where biodiversity conservation is a priority.

Collectively, first-generation biofuel crops consisting of annual monocultures associated with high levels of chemical inputs and disturbance are poor reservoirs for biodiversity in agricultural landscapes. Expansion of the acreage of these crops at the expense of other crops and natural and semi-natural habitats certainly have negative impacts on biodiversity with strong potential to impact a broad taxonomic range of organisms extending far beyond the borders of biofuel plantations and to affecting terrestrial, marine, and aquatic biological communities.

Second-Generation Biofuel Crops

Conventional biofuels are produced only from the starchy or oily portions of the fruit and seeds of first-generation feedstock species, representing only a fraction of the overall biomass of the plant. In addition, these are the same portions of plants required by the food industry. Cellulosic technology is a new method by which lignocellulose, a primary structural molecule in most terrestrial plants and a large portion of the plants overall biomass, is converted to sugar that can be converted to liquid fuels via pyrolysis, gasification, or fermentation. As one of the most plentiful organic molecules on Earth, lignocellulose represents an enormous untapped source of energy that could be produced from a vast range of feedstock species including nonfood plant materials such as cereal straws and corn stover (leaves and stems) as well as dedicated energy crops. Although a large array of second-generation feedstocks have been proposed and are being developed for production (**Table 1**), short-rotation woody crops and perennial grasses are currently considered to be the most efficient and sustainable feedstock for bioenergy production in temperate regions (Adler *et al.*, 2007; Karp and Shield, 2008; Russi, 2008; Williams *et al.*, 2009; Bringezu *et al.*, 2009).

Second-generation biofuels show promise for benefitting biodiversity on a global scale where they replace first-generation crops or other annual cash crops. Fuel can be processed from native plant species that have adapted to cope with and thrive in the face of regionally important pests, diseases, and weather conditions. As a corollary, native plants and animals that have coevolved with native feedstock species can provide benefits (e.g., nitrogen fixing, pest resistance) that increase plant yields while providing habitat for native plants, animals, fungal partners, and other microorganisms. This same reasoning opens the possibility that native plants can be economically produced in polyculture that should further enhance within-field biodiversity at other trophic levels by adding floristic and structural diversity. By supporting native species, such crops have the potential to enhance ecosystem services such as natural pest control, and subsequent economic benefits can extend outward into agricultural landscapes to benefit other nonbiofuel crops. In addition, most candidate feedstocks are perennial species that neither require replanting on an annual basis nor require substantial fertilizer and herbicide inputs. In this way, second-generation crops house great potential to experience substantially reduced machinery-related disturbance and represent increased habitat stability that benefits the persistence of diverse taxa in agricultural landscapes of which biofuels are a part. These more diverse, low-input, perennial feedstocks are more likely to benefit biodiversity (**Figure 2**).

Cellulosic technology is currently available, but further research is required to increase its economic efficiency so that second-generation crop production systems can effectively compete with other cash crops. It is generally thought that the process will become commercially viable and begin to industrialize and expand on a global scale in the years 2017–2023.

Agricultural wastes that are currently in heavy supply will very likely be the first major feedstock for the production of cellulosic ethanol. Foremost among these are the corn leaves and stalks remaining after harvest (e.g., corn stover). Stover has already come into demand as a cellulosic biomass feedstock. However, because stover removal negatively impacts soil nutrient content, expansion in its use as a fuel is predicted to increase demands on conventional farmland (Wilhelm *et al.*, 2007). Unfortunately, the financial incentives for stover removal are projected to be much greater than those to leave residues on the land to protect soil and water resources or retain as stubble and ground cover for the benefit of grassland birds and other wildlife (Cruse and Herndl, 2009). And even though the intensification of land already in high yield production will likely impact biodiversity, these costs must be weighed against the biodiversity cost of putting additional lands (i.e., energy sprawl) into production, even if low-intensity or organic farming methods are used. These same concerns will affect the extraction of wastes from other agricultural operations (e.g., forestry).

Short-rotation Woody Biomass

Short-rotation woody cropping systems include silvicultural operations designed to produce woody biomass over relatively short harvest cycles (1–15 years) using intensive silvicultural techniques and high-yielding tree varieties. Species like cottonwood and poplar (*populus* spp.) as well as willow (*salix* spp.) have come into production in North America and Europe, primarily for the pulp and paper industries. Yet this

same pulp can be converted to liquid fuel via cellulosic technology and so these crops have garnered attention as next-generation agroenergy crops.

Like first-generation biofuel crops, short-rotation woody crops are typically grown using intensive techniques, including fertilizers, pesticides, irrigation, and the control of competitive species. Because several species regenerate by producing multiple new shoots from stumps left after harvest, they are often coppice harvested. This is a key ecological benefit of these plant feedstocks in that it eliminates the need for more extensive disturbance of soils associated with annually resewing crops.

Short-rotation woody biomass systems may be considered intermediate to managed forest plantations and contemporary agrosystems in terms of their associated management inputs and implications for biodiversity (Burger, 2002). Indeed, diversity of birds and mammals and plants are typically lower on short-rotation woody plantations compared with slower-growing managed and unmanaged forests. Fauna associated with stands of fast-growing trees are extremely dynamic and transition from species assemblages that are more similar to those of open fields to those that are more akin to shrubland. In general, these woody plantations tend to favor shrub-associated vertebrates at the expense of cavity nesting and other species associated with mature forests (Riffell *et al.*, 2011). These crops tend to also accumulate and maintain more diverse communities of butterflies, spiders, and other invertebrates than annually harvested row crops (Dauber *et al.*, 2010). However, populations of some herbivorous arthropods (e.g., leaf-eating beetles, Coleoptera: Chrysomelidae) can increase to levels that cause significant crop damage, suggesting pesticide use could become necessary, which could negatively impact the diversity of nonpest invertebrates. Replacement of other managed or native forests by short-rotation crops will have significant negative biodiversity consequences (Archaux and Martin, 2009). Collectively, research on the relative biodiversity value of these crops in comparison to the agricultural, seminatural, and natural land-cover types they replace is extremely limited. Furthermore, information regarding the biodiversity associated with plantations of most of the remaining candidate woody species – for example, loblolly pine (*Pinus taeda*), alder (*Alnus* sp.), and black locust (*Robinia pseudoacacia*) – is largely absent.

Perennial Monocultures: Miscanthus and Switchgrass

Commonly referred to simply as *Miscanthus*, *Miscanthus giganteus* is a hybrid of two perennial grass species native to parts of Asia (*M. sinensis* and *M. sacchariflorus*). This large rhizomatous grass is endemic at latitudes up to 45° N and produces dense stands of very high biomass even under temperate climatic conditions (Lewandowski *et al.*, 2000; Hastings *et al.*, 2009). It appears to have the largest energy yield and energy use efficiency of all potential bioenergy crops (Sims *et al.*, 2006; Heaton *et al.*, 2004) in part due to reduced fertilizer requirements. For this reason, many consider *Miscanthus* to be the top candidate for production as an agricultural crop dedicated to cellulosic ethanol production. Some estimates suggest that even though a typical acre of corn yields around 7.6 t of biomass per acre and 756 gallons of ethanol, *Miscanthus* is capable of producing up to 20 t of biomass and 3250 gallons of ethanol.

In Europe, *Miscanthus* can support a relatively higher diversity of birds, arthropods, and weedy plants (Semere and Slater, 2007b; Bellamy *et al.*, 2009), including sensitive species when compared to other managed perennial crops and grasslands. Yet its spring harvest creates a tall overwinter stand structure that is relatively unsuitable for many vertebrates. This harvest schedule and *Miscanthus*'s relatively low chemical input requirements may reduce disturbance and mortality, which affects plant, bird, and mammal breeding in *Miscanthus* stands. This very limited information is also tempered by the fact that *Miscanthus* is an exotic and potentially invasive species in North America even as it is being considered for large-scale production. Its ability to support native biodiversity in this novel setting and its potential impacts on biodiversity should it become invasive are unknown.

Switchgrass (*Panicum virgatum*), another perennial grown in monocultures, is a warm-season grass native to the North American tallgrass prairie whose expansion in cultivation has been assumed will enhance biodiversity in agricultural landscapes. It has been selected as a model energy crop by the US Department of Energy also due to its high water and nitrogen use efficiency and its regionally adapted genotypes (McLaughlin and Walsh, 1998). It can form dense stands in monoculture and is often grown by farmers enrolled in government-sponsored agricultural set-aside programs to benefit biodiversity.

In some regions of the Midwestern US, switchgrass is currently under limited cultivation as a biomass feedstock for coal-fired power plants (e.g., Roth *et al.*, 2005), but it is not yet in large-scale production for purposes of cellulosic biofuel production. Switchgrass monocultures or near-monocultures support a larger and more enhanced diversity of breeding, wintering, and migratory birds, arthropods, and mammals than traditional corn and soybean row crops (Robertson *et al.*, 2011a, b; Fletcher *et al.*, 2011). This includes an enhanced diversity and biomass of functional groups of arthropods important in pollination, pest control, and nutrient cycling (Gardiner *et al.*, 2010; Landis and Werling, 2010). Compared to contemporary biomass crops, switchgrass also appears to enhance the diversity of the bird species of the highest conservation priority (Fletcher *et al.*, 2011).

Native Polycultures

Evidence indicates that diverse grasslands such as restored or reconstructed prairies containing as little as 60% grass can act as sustainable sources of cellulosic biomass (Tilman *et al.*, 2006; Garlock *et al.*, in preparation). Although high-diversity grasslands require more land area to produce a given amount of energy (Fargione *et al.*, 2009), they also require only a single planting, minimal cultivation, and low chemical inputs, which may improve their economic feasibility as a fuel sources.

The relative field-scale biodiversity value of such agrosystems has been explored largely in North America by comparing biodiversity responses to prairie remnants and grassland reconstructions to switchgrass monocultures and corn–soybean rotations. Diversity and abundance of breeding and migratory birds and terrestrial arthropods all are greatest in native prairie remnants, similar but lower in diverse prairie reconstructions, and generally elevated relative to switchgrass monocultures and cornfields (Gardiner *et al.*, 2010; Bakker

and Higgins, 2009; Landis and Werling, 2010; Robertson *et al.*, 2011a, b). This advantage also extends to arthropod groups effective in pollution and pest control.

New Approaches to Biodiversity Management

Collectively, candidate cellulosic biomass feedstocks all show positive effects on species richness for all taxa yet studied (Dauber *et al.*, 2010). These differences are due to pest- and weed control measures, vegetation structure, harvesting patterns and bottom-up impacts of plant diversity and productivity on higher trophic levels. Because many second-generation cellulosic crops are not yet under commercial cultivation, management regimes associated with field studies are commonly unrepresentative of those that will impact biodiversity in industrially managed fields. Consequently, the apparent biodiversity advantages of perennial monocultures and polycultures are likely overestimates.

The selection of genets and management techniques that maximize agricultural yields, for example, are known to reduce biodiversity in agricultural settings in general (Green *et al.*, 2005) and bioenergy crops specifically (Robertson *et al.*, 2011a). For example, perennial native plants like switchgrass may require little or no fertilizer but economic interests in maximizing crop yields suggest that fertilizers may still be a part of even these types of production systems (Fike *et al.*, 2006; Scott and Tiarks, 2008; Varvel *et al.*, 2008). Pesticide applications will obviously reduce the diversity and abundance of arthropod and weed populations, but may also affect vertebrate populations and are known to correlate with biodiversity loss on a global scale (Mozumdera and Berrens, 2007). Expanded research into the application of natural or integrated pest control in next-generation bioenergy crops may play an important role in mitigating biodiversity losses associated with crop management. For example, the introduction of a stable perennial ground flora may be useful in controlling competitive weeds in woody-biomass plantations and native monocultures and polycultures tend to sponsor greater arthropod-mediated pest-control services than row crops (Werling *et al.*, 2011).

Harvesting times and patterns are known to be of importance to some birds (Roth *et al.*, 2005), and novel harvest strategies, including partially harvesting fields and staggering harvests among fields, may be important strategies for managing biodiversity at the field scale. These techniques could enhance structural diversity and complexity, which are known to enhance species richness and abundance of birds in perennial grasslands and short-rotation plantations (Robertson *et al.*, 2011a, b). Such approaches will also benefit biodiversity in biomass plantations by enhancing species turnover and providing a continuous reservoir from which arthropods, mammals, and other less-mobile taxa can repopulate new biomass plantings (Christian *et al.*, 1997; Hanowski *et al.*, 1997).

Landscape-Scale Biodiversity Considerations

Understanding the dynamic responses of biodiversity to changing agricultural landscapes is fundamental to understanding how best to guide conservation efforts. Like other modern agricultural crops, the production of biofuels will have impacts on biodiversity by altering the structure and composition of the landscapes in which they are produced (Firbank, 2008; Landis *et al.*, 2008), with impacts on biological communities depending on their ecological traits (Burel *et al.*, 2004). Spatial heterogeneity is associated with biological diversity in biofuel crops (Dauber *et al.*, 2010) and agricultural landscapes in general (Benton *et al.*, 2003). Increasing this diversity may be a fundamentally important contribution of biofuel production to biodiversity conservation as biofuels enter existing agricultural landscapes that are commonly dominated by a few intensively managed crop types (Cook and Beyea, 2000). As a corollary, biomass plantations are likely to have adverse effects on biodiversity where they enter landscapes of high conservation value (Anderson and Fergusson, 2006; Schulz *et al.*, 2009) or where the biodiversity they harbor is only a subset of that already found in the surrounding landscape (Britt *et al.*, 2007; Baum *et al.*, 2009; Schulz *et al.*, 2009).

Spatial Considerations and Landscapes

The spatial extent and configuration of biomass crops within a landscape will profoundly shape their suitability for particular taxa. Avian diversity and abundance, for example, commonly increase as patches of habitat become larger and more habitat is available for area-sensitive or edge-avoiding species (Ribic *et al.*, 2009b; Johnson and Temple, 1990; Winter and Faaborg, 1999). North American bird species associated with switchgrass and mixed grass and forb plantings such as prairie reconstructions are commonly those that evolved to exploit various types of grassland habitats (Murray *et al.*, 2003; Roth *et al.*, 2005). Larger patches of these feedstocks will not only hold more species but also increase bird diversity per unit area as well as increase the likelihood of rare and sensitive species colonizing these fields (Herkert *et al.*, 1996; Robertson *et al.*, 2011a, b). Habitat patches within landscapes with a greater proportion of suitable habitat are also more likely to attract rare and sensitive species (Bakker *et al.*, 2002). In this way, production on large spatially concentrated fields of grass-based biomass crops could enhance populations of highly imperiled group of birds where they replace other crops (Robertson *et al.*, 2011a). Interspersion of grass-based and woody-biomass crops could have negative effects on grassland birds if grown in a predominantly agricultural and grassland landscape (Bakker and Higgins, 2009; Fletcher and Koford, 2002; Coppedge *et al.*, 2001; Renfrew and Ribic, 2008). Insomuch as the economics of ethanol production are predicted to cause aggregation of crops in the vicinity of processing plants (Graham *et al.*, 2000), there is the potential for altering agricultural landscape structure, with positive consequences for grassland bird communities. Similar benefits may accrue on other continents or in woody-biomass systems, but research is lacking.

Plant diversity appears to be maximized at a very small patch size in modern short-rotation woody plantations (Kroiher *et al.*, 2008). Because plant, mammal, and bird species richness commonly decreases from edges toward more central and insular portions of habitat patches, the

edge-to-area ratio of a patch can affect the diversity and abundance of species in a patch (Christian *et al.*, 1998; Weih *et al.*, 2003; Ribic *et al.*, 2009b), and this pattern may accumulate over time, for example, as woody biomass stands develop more complex habitat structure (Cunningham *et al.*, 2004).

In contrast to birds, arthropod diversity and abundance are less often linked to patch size but commonly increase as the diversity of habitats and land-use types in the landscape surrounding a patch increases (Simberloff, 1976; Summerville and Crist, 2003). This suggests that there exists a trade-off such that crop placement favoring one biodiversity component may negatively impact other important taxa. When crop placement and landscape-scale management needs of multiple taxa come into conflict, conservationists and land managers will require strategies that prioritize biodiversity components whose populations need the most help in particular regions or landscapes.

Finally, noncrop areas such as headlands, waterway buffer zones and harvest accesses to woody crops are important features for biodiversity that enhance populations of noncrop plants, birds, butterflies, native bees, and other arthropods (Anderson *et al.*, 2004; Marshall *et al.*, 2006; Haughton *et al.*, 2009; Bellamy *et al.*, 2009). Ultimately, the strategic placement of biomass plantings within existing agricultural landscapes has the potential to buffer disturbance regimes and edge effects in remnant patches (Brockerhoff *et al.*, 2008) and increase the connectivity of natural habitats for animal populations (Hanowski *et al.*, 1997; Firbank, 2008), therefore enhancing local and landscape-scale biodiversity. However, because landscape-scale biodiversity responses are, in part, emergent properties of field-scale responses of organisms to crop selection and management, the benefits of landscape-scale management will be contingent on biodiversity-friendly agricultural practices by individual farmers.

Agroenergy Sprawl and Biodiversity

Land-use change is considered one of the major drivers of biodiversity loss on a global scale. Perhaps the greatest threat that bioenergy production represents to biodiversity is its role in driving farmers and governments to convert biodiversity rich natural and seminatural habitats to biomass production. Economically important contemporary agroenergy crops like corn, soybeans, sugarcane and African oil palm have all been singled out for their negative biodiversity impacts associated with land-use change (Groom *et al.*, 2008; Robertson *et al.*, 2008; Eggers *et al.*, 2009; Fargione *et al.*, 2009; Fletcher *et al.*, 2011). European and North American countries and Brazil have legally mandated sustainable energy standards that require an increasing component of liquid biofuel in their national energy budget. Their respective governments have created financial mechanisms that have encouraged rapid and large-scale expansion in biofuel production. In the US, it is estimated that at least 206,000 km^2 of new cultivated land will be needed to meet US energy demand by 2030, an area roughly equal to the size of the state of Kansas (West *et al.*, 2009; McDonald *et al.*, 2009). The suitability of particular biomass feedstocks will vary according to geographic region (Evans *et al.*, 2010), and some biodiversity rich regions will be largely inappropriate for the production of bioenergy from ecological and economic points of view. The land area needed for biofuels is expected to disproportionately reduce the extent of some habitats more than others (e.g., temperate grasslands; McDonald *et al.*, 2009), thereby increasing the vulnerability of species that currently rely on these land-cover types. Biodiversity associated with temperate grasslands has already been drastically reduced through habitat conservation to agricultural use, so further changes associated with bioenergy production have potential to be particularly damaging.

A fundamental question regarding the land-use impacts of bioenergy on a global scale is whether bioenergy crops will replace existing crops or lead to the loss of natural and seminatural habitats with a resulting net loss in biodiversity. To a large degree, this question remains unanswered. American sustainable energy policies have largely been drawn based on the assumption that most energy-crop production will occur on sites already in use for agricultural production, and so far most of the expansion in corn production has come from land previously used for other crops, especially soybeans (Secchi and Babcock, 2007; USDA, 2009). Yet increasing production mandates are rapidly increasing demands on agricultural lands and favoring conversion of native prairie or seminatural grasslands associated with set-aside programs to corn-ethanol production (Secchi and Babcock, 2007; Stephens *et al.*, 2008; Fargione *et al.*, 2009). By one estimate, current trends in the US are projected to lead to losses of 121 million ha (30 million acres) in the prairie pothole region alone (Rashford *et al.*, 2011).

Government policies that provide economic incentives for farmers to produce bioenergy crops lead to more indirect forms of land-use change. For example, in the US farmers have reduced their rotation of corn with soybean plantings in response to federal ethanol subsidies (Secchi and Babcock, 2007). Reduced soybean supply has led to higher global soybean prices and increased clearing of Amazonian rainforest to grow this crop. The European Union's goal to produce 10% of its transport fuel from renewable sources by 2020 is projected to drive farmers to convert 69,000 km^2 of wild lands into fields and plantations, an area twice the size of Belguim (Bowyer, 2010). However, a recent analysis has indicated that indirect land-use changes induced by increased production of biofuels in the US are negligible or nonexistent (Kim and Dale, 2011). Such conflicting results highlight the newness of energy sprawl issues as well as the need for further investigations of land-use changes.

Potentially mitigating the direct and indirect losses of native and seminative habitat supporting important biodiversity components is the possibility of growing perennially based feedstocks on degraded lands or those with relatively nutrient-poor soils, so-called marginal lands (Fuentes and Taliaferro, 2002). Targeting biomass production on abandoned farmland or other lands unsuitable for the production of more traditional crops are considered a way to satisfy the acreage demands required for next-generation fuels (e.g., Hoogwijk *et al.*, 2005; Field *et al.*, 2008). It remains unclear if sufficient marginal lands exist to satisfy a significant fraction of current biofuel production goals in producing countries. More

important, marginally productive lands in Europe and the US are frequently enrolled in government-sponsored set-aside programs that provide payments to farmers for not producing crops on easily erodable or otherwise unused lands.

Regardless, if the biofuel production benchmarks of producing countries are to be met, it will mean large-scale, highly significant land-use changes will take place over relatively short timescales (EEA, 2006; Tuck *et al.*, 2006; McDonald *et al.*, 2009; West *et al.*, 2009; Fischer *et al.*, 2010). Projected biofuel-driven land-use change in Europe in the near future is predicted to negatively impact a broad range of taxa, including reptiles, butterflies, and birds (Louette *et al.*, 2010). These biodiversity impacts are unlikely to be uniformly distributed across agricultural landscapes. Transportation costs dictate a need for economic supply basins surrounding production plants that could easily lead to large-scale monocultures and the segregation of landscapes for energy production from those for food production and biodiversity conservation (Dauber *et al.*, 2010). In this way, economic forces may preclude the ability of strategic landscape design to manage biodiversity in biofuel production landscapes (Bellamy *et al.*, 2009). Consequently, and in the absence of government intervention, whether or not bioenergy production can be effectively integrated into multifunctional landscapes that are socioeconomically and ecologically sustainable will depend on economic markets and the distribution of crop production and processing infrastructure (Berndes *et al.*, 2008; Porter *et al.*, 2009; Dauber *et al.*, 2010).

In summary, our framework of identifying the biodiversity consequences associated with crop selection and management and land-use change is an attempt to parse the individual contributions of these variables to biodiversity loss. Clearly, though, crop traits (e.g., perenniality) do much to dictate important biodiversity-related management regimes associated with crops (e.g., mechanical disturbance type and frequency). Moreover, crops traits like perenniality influence the economics of crop-selection decisions for producers because these crops often take at least a year to establish themselves before they can be harvested. Producer decisions about crop selection influence land-use change at large spatial scales that, in turn, shape the biodiversity consequences of local crop management. These connections among impact categories and spatial scales are common in agricultural systems and are intertwined with the evolving socioeconomic and technological advancements that shape the economics of crop selection and the efficiency of management techniques. If biodiversity conservation is to succeed in the face of biofuel-driven global change, we will need to explicitly address and exploit these connections and systematic inputs in designing sustainable production systems.

Strategies for Biodiversity-Friendly Energy-Crop Production

Strategies to implement biodiversity-friendly energy-crop production arise from a variety of contributing factors (**Figure 1**). The design and implementation of landscapes that provide for both biofuels production and biodiversity conservation may yet be achieved with careful consideration and design as highlighted by an approach that follows a mitigation hierarchy (Fargione *et al.*, 2010). As a first step, areas that primarily provide biodiversity and valued ecosystem services must be avoided (Keeney and Nanninga, 2008). Increases in biofuel production have many consequences, including the removal of lands from conservation set aside programs that provide biodiversity benefits, a movement toward producing crops in marginal areas (e.g., degraded or abandoned cropland), an increase in the diversity of feedstocks that can be used in producing biofuels, potential development of lands that were previously never considered as production landscapes. Therefore, as a first step in avoiding negative biodiversity and ecosystem service impacts, areas that provide such services must be identified and protected.

A second step of the mitigation hierarchy focuses on minimizing detrimental impacts through the use of best management practices not only within the fields that produce biofuel feedstocks but also across broader biofuel production landscapes. Such management can be accomplished through crop selection, novel management strategies, and strategic land-management approaches. Although economic concerns will drive much of crop selection, some crops will better diversify the land-cover composition of agricultural landscapes, thereby accruing more substantial biodiversity benefits. Asynchronous harvesting regimes among field and partial harvesting of focal fields have potential to create mosaics of agroecological habitat differing in their structure that will directly benefit taxa like birds that select breeding and migratory habitat based largely on structural considerations (Roth *et al.*, 2005; Robertson *et al.*, 2011a, b) but also benefit a diverse range of taxa (Sage and Robertson, 1996; Baum *et al.*, 2009; Fargione *et al.*, 2009; Gardiner *et al.*, 2010) and the ecosystem services they provide (Isaacs *et al.*, 2009; Landis and Werling, 2010). Such strategies may also have the secondary effect of broadening within- or between-field habitat structural diversity the following spring and summer by increasing standing dead vegetation where fall regrowth has occurred (Benton *et al.*, 2003; Murray and Best, 2003; Roth *et al.*, 2005).

A strategic landscape-scale approach to biofuel crop placement could also enhance biodiversity in ways that have not traditionally been considered in agricultural systems (Paine *et al.*, 1996). Set-aside programs in North America and Europe already provide for the planting of perennial grassland or other plants along streams or as buffer strips that are known to benefit birds, pollinators, and other species. Biomass crops may also act as ecotones, buffers around natural areas, or act to enhance connectivity between structurally similar natural habitat types (Cook and Beyea, 2000; Dauber *et al.*, 2010). These options can further enhance more straightforward biodiversity benefits of diversifying the land-cover in agricultural landscapes.

Additional consideration must also be given to factors that drive the development and production of biofuels (**Figure 1**). Primarily, there must be consideration of both increases in fuel efficiency as well as increased development and production of alternative-energy vehicles (Wiens *et al.*, 2011). Such conservation measures can have a significant impact on the demand for biofuels as a major contributing factor to the potential impacts of energy sprawl (McDonald *et al.*, 2009). Additional consideration must also be paid to the full suite of

impacts across the entire life cycle of biofuel production. Life-cycle analysis (LCA), an approach designed to evaluate economic, environmental, and social impacts from the development to the delivery of a product, can be used to evaluate the potential costs and benefits of different feedstocks (Jenkins and Alles, 2011; Wiens *et al.*, 2011). Ideally, such analyses are tailored to the specific feedstock as well as the specific landscape in which it is produced.

The Role of Policy Development Inconserving Biodiversity

Ultimately, the prospects for capitalizing on the biodiversity benefits of such options are remote in the absence of government intervention. Although the preceding recommended strategies primarily focus on land management, the drivers of energy sprawl will require policy-level solutions. In the US, where production targets, subsidies, and blending targets have been used to encourage corn production and are now being used to encourage the production of cellulosic crops, environmental standards for the production of biomass crops have not yet been created. In contrast, the European Union (EU) has begun to address bioenergy production issues through the revised Fuel Quality Directive (European Commission Environment, 2009) and the Renewable Energy Sources Directive (EU-RES-D, 2009). These standards prohibit the production of biomass from land with high biodiversity value (including native grasslands), promote the use of contaminated or marginal lands through a bonus system, mandate that biomass be produced in accordance with EU standards (e.g., cross compliance, good agricultural and environmental conditions), and require that practices outside of the EU be monitored and specified through agreements with producing countries. Although indirect land-use impacts are not fully considered (Hennenberg *et al.*, 2009) and production of biomass remains allowed on lands enrolled in set-aside programs (which otherwise cannot be used for production under EU agricultural rules), European standards remain the most stringent and well developed in the world. Still, 0.9 million ha (2.2 million acres) of the set-aside area have been used in recent years for nonfood production. In addition, by focusing on species richness and abundance as measures of programmatic success, agri-environmental schemes have largely overlooked the habitat requirements and management needs of specialist species (Filippi-Codaccioni *et al.*, 2010). Moreover, the EU has neither yet set minimum sustainability standards for all bioenergy against negative environmental and social impacts nor established a robust and verifiable system of certification for biofuels. As leading global producers of biofuels, the US lags far beyond in nearly all of these critical steps necessary to maintain biodiversity as bioenergy production expands.

As biofuels become a global commodity, there will become an increasing urgency for global-scale cooperation in the development, implementation, and enforcement of biofuel production standards. Given the already serious implications of indirect land-use change that cross international borders and the eventual likelihood of international markets for biofuel production, there will be a critical need for international cooperation for the creation of methods that encourage crop selection, management, and production systems that mitigate biodiversity losses associated with land-use change, maximize biodiversity in agricultural landscapes through landscape-scale strategic planning and incentive program, develop rigorous biodiversity standards for producers, and create a role for the enforcement of those standards. Such standards will be insufficient unless they are broadly implemented through production regions (Hennenberg *et al.*, 2009) via grower incentives and effective land-use planning that will need to be implemented by the governments of producer nations. Economic incentives associated with land set-aside programs appear to be a particularly opportune means of encouraging that productions standards are met.

Because responses to biomass crops and their landscape context will vary across biodiversity components (Dauber *et al.*, 2010), specific production guidelines will be needed for rare and sensitive species within a larger policy that more broadly targets the conservation of biodiversity and ecosystem services (e.g., pollination, pest control). In this regard, a strategic land-use planning approach that incorporates the value of private (unprotected) lands (e.g., Wilson *et al.*, 2010) may be useful. Such strategies are capable of revealing not only where to invest but also which strategies to invest in to effectively and efficiently conserve biodiversity. Finally, the effectiveness of programs in accomplishing conservation goals will need to be monitored. Mechanisms such as ecological cross compliance, in which farmers have to meet environmental standards to qualify for acreage-related direct payments, may be a potentially powerful instrument in this regard (Aviron *et al.*, 2009).

Summary and Open Questions

Will energy sprawl necessarily have a negative impact on biodiversity? Are bioenergy production and biodiversity conservation compatible or mutually exclusive? The answer to these questions clearly depends on the type of crop in question, the previous use of the land on which it is being cultivated, whether its management includes provisions to mitigate biodiversity losses or actually enhance its value as wildlife habitat, and on the details of national and international policies that mandate its ecological sustainability.

Further hindering a clear picture of how bioenergy production will impact biodiversity is a broad array of knowledge gaps. First, we still know too little about the biodiversity associated with various biofuel crops, especially important candidate next-generation crops (e.g., *Miscanthus*). Some species are being proposed as crops outside of their native range and where their potential to become invasive and to displace native biodiversity is unknown. Second, our understanding of how next-generation crops are to be managed is largely speculative as these crops are only recently entering agricultural production. Because crop management can profoundly impact field-scale biodiversity, it will be critical to understand how the evolving economics of bioenergy production systems will allow for flexibility in management that mitigates biodiversity impacts. Our ability to develop realistic and sustainable producer guidelines and best practices will be

contingent on an improved understanding of the economics and logistics of crop management. Further confusing the relationship of crop selection and management and biodiversity are the unknown effects of genetic modifications of feedstocks that lead to altered crops structure and pesticide and herbicide resistance.

There exists a complex interplay between socioeconomic and environmental factors in determining the impacts of bioenergy production on biodiversity. In addition to the role of national bioenergy policies and market forces in affecting direct and indirect land-use change, for example, biofuel production has the potential to accelerate global warming or even reduce available freshwater supplies (Evans and Cohen, 2009). Developing scientifically based criteria for the sustainable production of bioenergy will be difficult unless we develop a framework for assessing the risk to biodiversity associated with socioeconomic and environmental conditions that constantly incorporates new knowledge. Bioenergy production should be considered a major threat to the persistence of biodiversity and ecosystem services on a global scale. Yet next-generation crops have substantial promise to benefit biodiversity should they be strategically integrated into existing agricultural landscapes. Polycultural feedstocks such as mixed grass and forb prairie, for example, could have profoundly important benefits for rare and declining species associated with temperate grasslands while providing fuel and a range of other ecological benefits such as pest control (Werling *et al.*, 2011). Ultimately, economically viable and sustainable bioenergy production, integrated with food production, may yet be feasible. Realizing the biodiversity potential associated with cellulosic bioenergy will be dependent on some profound changes in agricultural policies in producer countries, a renewed focus on interdisciplinary research, the ultimate economics of industrialized bioenergy production systems, and international cooperation to ensure that environmental and biodiversity costs of biofuel production are not simply exported elsewhere.

Appendix

List of Courses

- Agroecology
- Sustainable Agriculture
- Landscape Ecology
- Bioenergy Production Systems
- Sustainable Biofuel Production

Acknowledgments

This work was supported by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02–07ER64494) and DOE OBP Office of Energy Efficiency and Renewable energy DE-AC05–76l01830 and US Fish and Wildlife Service, Migratory Bird Joint Ventures Midwest Region Grant no. 30181AG045. BAR was supported by a fellowship from the Smithsonian Conservation Biology Institute. PJD was supported by The Nature Conservancy's Great Lakes fund for Partnership in Conservation Science and Economics and the US National Science Foundation LTER Program.

See also: Biodiversity and Ecosystem Services. Biodiversity and Pest Control Services. Indirect Land Use and Greenhouse Gas Impacts of Biofuels. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production. Life Cycle Analysis of Biofuels. Oil-Palm Plantations in the Context of Biodiversity Conservation

References

Adler PR, Del Grosso SJ, and Parton WJ (April 2007) Life-cycle assessment of net greenhouse-gas flux for bioenergy cropping systems. *Ecological Applications* 17(3): 675–691.

Anderson GQA and Fergusson MJ (2006) Energy from biomass in the UK: Sources, processes and biodiversity implications. *Ibis: The International Journal of Avian Science* 148: 180–183.

Anderson GQA, Haskins LR, and Nelson SH (2004) The effects of bioenergy crops on farmland birds in the United Kingdom – a review of current knowledge and future predictions. In: Parris K and Poincet T (eds.) *Biomass and Agriculture: Sustainability, Markets and Policy*, pp. 199–218. Paris, France: OECD.

Archaux F and Martin H (2009) Hybrid poplar plantations in a floodplain have balanced impacts on farmland and woodland birds. *Forest Ecology and Management* 27: 1474–1479.

Aviron S, Nitsch H, Jeanneret P, *et al.* (2009) Ecological cross compliance promotes farmland biodiversity in Switzerland. *Frontiers in Ecology and the Environment* 7: 247–252.

Bakker KK and Higgins KF (2009) Planted grasslands and native sod prairie: Equivalent habitat for grassland birds? *Western North American Naturalist* 69: 35–242.

Bakker KK, Naugle DE, and Higgins KF (2002) Incorporating landscape attributes into models for migratory grassland bird conservation. *Conservation Biology* 16: 1638–1646.

Baum S, Weih M, Busch G, Kroiher F, and Bolte A (2009) The impact of short rotation coppice plantations on phytodiversity. *Landbauforschung – vTI Agriculture and Forestry Research* 59: 163–170.

Beecher NA, Johnson RJ, Brandle RJ, Case RM, and Young LJ (2002) Agroecology of birds in organic and nonorganic farmland. *Conservation Biology* 16: 1620–1631.

Bellamy PE, Croxton PJ, Heard MS, *et al.* (2009) The impact of growing Miscanthus for biomass on farmland bird populations. *Biomass Bioenergy* 33: 191–199.

Benton TG, Vickery JA, and Wilson JD (2003) Farmland biodiversity: Is habitat heterogeneity the key? *TREE* 18: 182–188.

Berndes G, Börjesson P, Ostwald M, and Palm M (2008) Multifunctional biomass production systems – an overview with presentation of specific applications in India and Sweden. *Biofuels, Bioproducts and Biorefining* 2: 16–25.

Bowyer C (2010) Anticipated Indirect Land Use Change Associated with Expanded Use of Biofuels and Bioliquids in the EU – An Analysis of the National Renewable Energy Action Plans London, UK: *Institute for European Environmental Policy.*

Bringezu S, Schutz H, O'Brien M, Kauppi L, Howarth RW, and McNeely J (2009) Towards sustainable production and use of resources: Assessing biofuels. *United Nations Environment Programme*, p. 119, http://www.unep.fr/scp/rpanel/pdf/Assessing_Biofuels_Full_Report.pdf (accessed 1/7/2010).

Britt CP, Fowbert J, and McMillan SD (2007) The ground flora and invertebrate fauna of hybrid poplar plantations: Results of ecological monitoring in the PAMUCEAF project. *Aspects of Applied Biology* 82: 83–90.

Brockerhoff EG, Jactel H, Parrotta JA, *et al.* (2008) Plantation forests and biodiversity: Oxymoron or opportunity? *Biodiversity Conservation* 17: 925–951.

Burel F, Butet A, Delettre YR, and Millàn de la Peña N (2004) Differential response of selected taxa to landscape context and agricultural intensification. *Landscape and Urban Planning* 67: 195–204.

Burger JA (2002) Soil and long-term site productivity values. In: Richardson J, Björheden R, Hakkila P, Lowe AT, and Smith CT (eds.) *Bioenergy From Sustainable Forestry: Guiding Principles and practice*, 1st edn., pp. 165–189. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Christian DP, Collins PT, Hanowski JM, *et al.* (1997) Bird and small mammal use of short-rotation hybrid poplar plantations. *The Journal of Wildlife Management* 61: 171–182.

Christian DP, Hoffman W, Hanowski JM, *et al.* (1998) Bird and mammal diversity on woody biomass plantations in North America. *Biomass and Bioenergy* 14: 395–402.

Cook J and Beyea J (2000) Bioenergy in the United States: Progress and possibilities. *Biomass and Bioenergy* 18: 441–455.

Coppedge BR, Engle DM, Masters RE, and Gregory MS (2001) Avian response to landscape change in fragmented southern Great Plains grasslands. *Ecological Applications* 11: 47–59.

Cruse RM and Herndl CG (2009) Balancing corn stover harvest for biofuels with soil and water conservation. *Journal of Soil and Water Conservation* 64: 286–291.

Cunningham M, Bishop JD, McKay HV, and Sage RB (2004) *The Ecology of Short Rotation Coppice Crop. – ARBRE Monitoring: B/U1/00727/00/REPORT. ETSU.* Oxford: DTI.

Dauber JS, Jones MB, and Stout JC (2010) The impact of biomass crop cultivation on temperate biodiversity. *GCB Bioenergy* 2: 289–309.

Donald PF, Green RE, and Heath MF (2001) Agricultural intensification and the collapse of Europe's farmland bird populations. *Proceedings of the Royal Society London: Biological Sciences* 268: 25–29.

EEA (2006) *How much Bioenergy can Europe Produce Without Harming the Environment? EEA Report 7/2006.* Copenhagen: European Environment Agency.

Eggers J, Tröltzsch K, Falcucci A, *et al.* (2009) Is biofuel policy harming biodiversity in Europe? *GCB Bioenergy* 1: 18–34.

European Commission Environment Directive 2009/28/EC of 2009 on European renewable energy sources. http://ec.europa.eu/environment/air/transport/fuel.htm (accessed 1/20/2011).

European Commission Environment Directive 2009/30/EC of 2009 on fuel quality amendment. http://ec.europa.eu/environment/air/transport/fuel.htm (accessed 1/20/2011).

Evans JM and Cohen MJ (2009) Regional water resource implications of bioethanol production in the Southeastern United States. *Global Change Biology* 15: 2261–2273.

Evans JM, Fletcher Jr. RJ, and Alavalapati JI (2010) Using species distribution models to identify suitable areas for biofuel feedstock production. *GCB Bioenergy* 2: 63–78.

Fargione JE, Cooper TR, Flaspohler DJ, *et al.* (2009) Bioenergy and wildlife: Threats and opportunities for grassland conservation. *BioScience* 59: 767–777.

Fargione JE, Plevin RG, and Hill JD (2010) The ecological impact of biofuels. *Annual Review of Ecology, Evolution, and Systematics* 41: 351–377.

Field CB, Campbell JE, and Lobell DB (2008) Biomass energy: The scale of the potential resource. *TREE* 23: 65–72.

Fike JH, Parrish DJ, Wolf DD, *et al.* (2006) Long-term yield potential of switchgrass-for-biofuel systems. *Biomass Bioenergy* 30: 198–206.

Filippi-Codaccioni O, Devictor V, Bas Y, and Julliard R (2010) Toward more concern for specialization and less for species diversity in conserving farmland biodiversity. *Biological Conservation* 143: 1493–1500.

Firbank LG (2008) Assessing the ecological impacts of bioenergy projects. *Bioenergy Research* 1: 12–19.

Firbank LG, Petit S, Smart S, Blain A, and Fuller RJ (2008) Assessing the impacts of agricultural intensification on biodiversity: A British perspective. *Philosophical Transactions of the Royal Society: Biological Sciences* 363: 777–787.

Fischer G, Prieler S, van Velthuizen H, *et al.* (2010) Biofuel production potentials in Europe: Sustainable use of cultivated land and pastures, Part II: Land use scenarios. *Biomass and Bioenergy* 34: 173–187.

Fletcher Jr. RJ and Koford RR (2002) Habitat and landscape associations of breeding birds in native and restored grasslands. *Journal of Wildlife Management* 66: 1011–1022.

Fletcher Jr. RJ, Robertson BA, Evans J, *et al.* (2011) Biodiversity conservation in the era of biofuels: Risks and opportunities. *Frontiers in Ecology and the Environment* 9: 161–168.

Fuentes RG and Taliaferro CM (2002) Biomass yield stability of switchgrass cultivars. In: Janick J and Whipkey A (eds.) *Trends in New Crops and New Uses*, pp. 276–282. Alexandria, Virginia: ASHS Press.

Gardiner M, Tuell J, Isaacs R, *et al.* (2010) Implications of three model biofuel crops for beneficial arthropods in agricultural landscapes. *Bioenergy Research* 3: 6–19.

Garlock R, Bals B, Jasrotia P, Balan V, and Dale BE (in preparation) Influence of botanical classification on the saccharification of cellulosic mixed-species feedstocks with comparison to corn stover. *Biomass Bioenergy.*

Graham RL, English BC, and Noon CE (2000) A geographic information system-based modeling system for evaluating the cost of delivered energy crop feedstock. *Biomass Bioenergy* 18: 309–329.

Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–555.

Groom MJ, Gray EM, and Townsend PA (2008) Biofuels and biodiversity: Principles for creating better policies for biofuel production. *Conservation Biology* 22: 602–609.

Guillette LJ, Brock JW, Rooney AA, and Woodward AR (1999) Serum concentration of various environmental contaminants and their relationship to sex steroid concentrations and phallus size in juvenile American Alligators. *Archive of Environmental Contamination and Toxicology* 36: 447–455. Available at http://link.springer-ny.com.

Haddad NM, Crutsinger GM, Gross K, *et al.* (2009) Plant species loss decreases arthropod diversity and shifts trophic structure. *Ecology Letters* 12: 1029–1039.

Haddad NM, Tilman D, Haarstad J, Ritchie M, and Knops JMH (2001) Contrasting effects of plant richness and composition on insect communities: A field experiment. *American Naturalist* 158: 17–35.

Hanowski JM, Niemi GJ, and Christian DP (1997) Influence of within plantation heterogeneity and surrounding landscape composition on avian communities in hybrid poplar plantations. *Conservation Biology* 11: 936–944.

Hastings A, Clifton-Brown JC, Wattenbach M, *et al.* (2009) Future energy potential of Miscanthus in Europe. *Global Change Biology – Bioenergy* 1: 180–196.

Haughton AJ, Bond AJ, Lovett AA, *et al.* (2009) A novel, integrated approach to assessing social, economic and environmental implications of changing rural land-use: A case study of perennial biomass crops. *Journal of Applied Ecology* 46: 315–322.

Heaton E, Voigt T, and Long SP (2004) A quantitative review comparing the yields for two candidate C4 perennial biomass crops in relation to nitrogen, temperature and water. *Biomass and Bioenergy* 27. 21–30.

Henderson IG, Vickery JA, and Carter N (2004) The use of winter bird crops by farmland birds in lowland England. *Biological Conservation* 118: 21–32.

Hennenberg KJ, Dragisic C, Haye S, *et al.* (2009) The power of bioenergy-related standards to protect biodiversity. *Conservation Biology* 24: 412–423.

Herkert JR, Sample DW, and Warner RE (1996) Management of midwestern grassland landscapes for the conservation of migratory birds. In: Thompson FR (ed.) *Management of Midwestern Landscapes for the Conservation of Neotropical Migratory Birds*, pp. 89–116. St. Paul, Minnesota: USDA Forest Service North Central Forest Experiment Station. USDA Forest Service Gen. Tech. Rep. NC-187.

Hill J, Nelson E, Tilman D, Polasky S, and Tiffany D (2006) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11206–11210.

Hole DG, Perkins AJ, Wilson JD, Alexander PV, and Evans AD (2005) Does organic farming benefit biodiversity? *Biological Conservation* 122: 113–130.

Hoogwijk M, Faaij A, Eickhout B, de Vries B, and Turkenburg W (2005) Potential of biomass energy out to 2100, for four IPCCSRES land-use scenarios. *Biomass Bioenergy* 29: 225–257.

Isaacs R, Tuell J, Fiedler A, Gardiner M, and Landis D (2009) Maximizing arthropod-mediated ecosystem services in agricultural landscapes: The role of native plants. *Frontiers in Ecology and the Environment* 7: 196–203.

Jenkins R and Alles C (2011) Field to fuel: Developing sustainable biorefineries. *Ecological Applications* 21: 1096–1104.

Johnson RG and Temple SA (1990) Nest predation and brood parasitism of tall-grass prairie birds. *Journal of Wildlife Management* 54: 106–111.

Karp A and Shield I (2008) Bioenergy from plants and the sustainable yield challenge. *New Phytologist* 179: 15–32.

Kearns CA and Inouye DW (1997) Pollinators, flowering plants, and conservation biology: Much remains to be learned about pollinators and plants. *BioScience* 47: 297–307.

Keeney D and Nanninga C (2008) *Biofuel and Global Biodiversity.* Minneapolis, MN: The Institute for Agriculture and Trade Policy.

Kim S and Dale BE (2011) Indirect land use change for biofuels: Testing predictions and improving analytical methodologies. *Biomass and Bioenergy* 35: 3235–3240.

Kroiher F, Bielefeldt J, Bolte A, *et al.* (2008) Die Phytodiversität in Energieholzbeständen – erste Ergebnisse im Rahmen des Projekts NOVALIS. *Archiv für Forstwesen und Landschaftsökologie* 42: 158–165.

Landis DA, Gardiner MM, van der Werf W, and Swinton SM (2008) Increasing corn for biofuel production reduces biocontrol services in agricultural landscapes. *Proceedings of the National Academy of Science* 105: 20552–20557.

Landis DA and Werling BP (2010) Arthropods and biofuel production systems in North America. *Insect Science* 17: 1–17.
Lewandowski I, Clifton-Brown JC, Scurlock JMO, and Huisman W (2000) Miscanthus: European experience with a novel energy crop. *Biomass Bioenergy* 19: 209–227.
Lokemoen JT and Beiser JA (1997) Bird use and nesting in conventional, minimum-tillage, and organic cropland. *Journal of Wildlife Management* 61: 644–655.
Louette G, Maes D, Alkemade JRM, *et al.* (2010) BioScore – cost effective assessment of policy impact on biodiversity using species sensitivity scores. *Journal for Nature Conservation* 18: 142–148.
Marshall EJP, West TM, and Kleijn D (2006) Impacts of an agrienvironment field margin prescription on the flora and fauna of arable farmland in different landscapes. *Agriculture, Ecosystems and Environment* 113: 36–44.
Matson PA, Parton WJ, Power AG, and Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 5325: 504–509.
McDonald RI, Fargione J, Kiesecker J, Miller WM, and Powell J (2009) Energy sprawl or energy efficiency: Climate policy impacts on natural habitat for the United States of America. *PLoS One* 4: e6802.
McLaughlin SB and Walsh ME (1998) Evaluating environmental consequences of producing herbaceous crops for bioenergy. *Biomass Bioenergy* 14: 317–324.
Mozumdera P and Berrens RP (2007) Inorganic fertilizer use and biodiversity risk: An empirical investigation. *Ecological Economics* 62: 538–543.
Murray LD and Best LB (2003) Short-term bird response to harvesting switchgrass for biomass in Iowa. *Journal of Wildlife Management* 67: 611–621.
Murray LD, Best LB, Jacobsen TJ, and Braster ML (2003) Potential effects on grassland birds of converting marginal cropland to switchgrass biomass production. *Biomass Bioenergy* 25: 167–175.
Paine LK, Peterson TL, Undersander DJ, *et al.* (1996) Some ecological and socio-economic considerations for biomass energy crop production. *Biomass Bioenergy* 10: 231–242.
Porter J, Costanza R, Sandhu H, Sigsgaard L, and Wratten S (2009) The value of producing food, energy, and ecosystem services within an agro-ecosystem. *Ambio* 38: 186–193.
Rashford BS, Walker JA, and Bastian CT (2011) Economics of grassland conversion to cropland in the prairie pothole region. *Conservation Biology* 25: 276–284.
Renfrew RB and Ribic CA (2008) Multi-scale models of grassland passerine abundance in a fragmented system in Wisconsin. *Landscape Ecology* 23: 181–193.
Ribic CAR, Koford R, Herkert JR, *et al.* (2009b) Area sensitivity in North American grassland birds: Patterns and processes. *The Auk* 126: 233–244.
Riffell S, Verschuyl J, Miller D, and Wigley TB (2011) A meta-analysis of bird and mammal response to short-rotation woody crops. *GCB Bioenergy* 3: 313–321.
Robertson BA, Doran PJ, Loomis ER, *et al.* (2011a) Perennial biomass crops enhance avian biodiversity. *Global change Biology Bioenergy* 3: 235–246.
Robertson BA, Doran PJ, Loomis ER, *et al.* (2011b) Avian use of perennial biomass feedstocks as post-breeding and migratory stopover habitat. *PLoS One* 6: e16941.
Robertson GP, Dale VH, Doering OC, *et al.* (2008) Sustainable biofuels redux. *Science* 322: 49–50.
Roth AM, Sample DW, Ribic CA, *et al.* (2005) Grassland bird response to harvesting switchgrass as a biomass energy crop. *Biomass and Bioenergy* 28: 490–498.
Russi D (2008) An integrated assessment of a large-scale biodiesel production in Italy: Killing several birds with one stone? *Energy Policy* 36(3): 1169–1180.
Sage RB and Robertson PA (1996) Factors affecting songbird communities using new short rotation coppice habitats in spring. *Bird Study* 43: 201–213.
Schulz U, Brauner O, and Gru H (2009) Animal diversity on short rotation coppices – a review. *Landbauforschung – vTI Agriculture and Forestry Research* 59: 171–182.
Scott DA and Tiarks A (2008) Dual-cropping loblolly pine for biomass energy and conventional wood products. *Southern Journal of Applied Forestry* 32: 33–37.
Secchi S and Babcock B (2007) Impact of high crop prices on environmental quality: A case of Iowa and the Conservation Reserve Program. Center for Agricultural and Rural Development (CARD) at Iowa State University Publications 07-wp447.
Semere T and Slater FM (2007b) Invertebrate populations in miscanthus (*Miscanthus giganteus*) and reed canary-grass (*Phalaris arundinacea*) fields. *Biomass and Bioenergy* 31: 30–39.
Simberloff D (1976) Species turnover and equilibrium. *Island Biogeography Science* 194: 572–578.
Sims REH, Hastings A, Schlamadinger B, Taylor G, and Smith P (2006) Energy crops: Current status and future prospects. *Global Change Biology* 12: 2054–2076.
Stephens SE, Walker JA, Blunck DR, *et al.* (2008) Predicting risk of habitat conversion in native temperate grasslands. *Conservation Biology* 22: 1320–1330.
Summerville KS and Crist TO (2003) Determinants of lepidopteran community composition and species diversity in eastern deciduous forests: Roles of season, eco-region and patch size. *Oikos* 100: 134–148.
Tilman D, Hill J, and Lehman C (2006) Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science* 314: 1598–1600.
Tscharntke T, Klein AM, Kreuess A, and Steffan-Dewenter I (2005) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* 8: 857–874.
Tuck G, Glendining MJ, Smith P, *et al.* (2006) The potential distribution of bioenergy crops in Europe under present and future climate. *Biomass and Bioenergy* 30: 183–197.
USDA (2009) *Conservation Reserve Program, monthly summary October 2009.* http://www.fsa.usda.gov/Internet/FSA_File/oct2009.pdf.
Varvel GE, Vogel KP, Mitchell RB, Follett RF, and Kimble JM (2008) Comparison of corn and switchgrass on marginal soils for bioenergy. *Biomass and Bioenergy* 32: 18–21.
Warburton DB and Klimstra WD (1984) Wildlife use of no-till and conventionally tilled corn fields. *Journal of Soil and Water Conservation* 39: 327–330.
Watkinson AR, Freckleton RP, Robinson RA, and Sutherland WJ (2000) Predictions of biodiversity response to genetically modified herbicide-tolerant crops. *Science* 289: 1554–1557.
Weih M, Karacic A, Munkert H, Verwijst T, and Diekmann M (2003) Influence of young poplar stands on floristic diversity in agricultural landscapes (Sweden). *Basic and Applied Ecology* 4(2): 149–156.
Werling BP, Meehan T, Robertson BA, Gratton C, and Landis DA (2011) Biocontrol potential varies with changes in biofuel-crop plant communities and landscape perenniality. *Global Change Biology Bioenergy.*
West T, Dunphy-Guzman K, Sun A, *et al.* (2009) *Feasibility, Economics, and Environmental Impact of Producing 90 Billion Gallons of Ethanol per Year by 2030.* Livermore, CA: Sandia National Laboratories.
Wiens J, Fargione J, and Hill J (2011) Biofuels and biodiversity. *Ecological Applications* 21: 1085–1095.
Wilhelm WW, Johnson JME, Karlen DL, and Lightle DT (2007) Corn stover to sustain soil organic carbon further constrains biomass supply. *Agronomy Journal* 99: 1665–1667.
Williams BK, Szaro RC, and Shapiro CD (2009) *Adaptive Management: The U.S. Department of the Interior Technical Guide.* Washington, DC: Adaptive Management Working Group, US Department of the Interior.
Wilson KA, Meijaard E, Drummond S, *et al.* (2010) Conserving biodiversity in production landscapes. *Ecological Applications* 20: 1721–1732.
Winter M and Faaborg J (1999) Patterns of area sensitivity in grassland-nesting birds. *Conservation Biology* 13: 1424–1436.

Biofuels and Biodiversity, Wildlife Habitat Restoration

Lisa A Schulte, Todd A Ontl, and GL Drake Larsen, Iowa State University, Ames, IA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Agricultural lands, degraded Land currently or formerly used for the cultivation of crops or pasture in which the productivity has been reduced as a result of human actions.

Agricultural lands, prime Land that has the best combination of physical (e.g., moisture status, soil depth, or slope condition) and chemical (e.g., nutrient status and salinity) characteristics for producing food, forage, fiber, and oilseed crops; "food" is used comprehensively to include crops grown for direct human consumption, processed for human consumption, and crops fed to livestock raised for human consumption.

Biodiversity The variety of living organisms considered at all levels, from genetic variants within the same species to the whole range of species and ecological complexes.

Bioenergy Forms of energy that can be extracted from biomass, including biodiesel from fats and oils; bioethanol from sugars, starches, and lignocellulosic material; and, bioelectricity and bioheat from the combustion of lignocellulosic material.

Bioethanol Alcohol made from sugar (e.g., sugarcane), starch (e.g., corn grain), or cellulosic (e.g., corn stalks, wheat straw, prairie grasses, and trees) biomass crops that can be used as a biofuel; sugar and starches are converted to ethanol through fermentation, while cellulosic material is converted via enzymatic hydrolysis or synthesis gas fermentation.

Biofuels Liquid fuels derived from biomass, including biodiesel from fats and oils and bioethanol from sugars, starches, and lignocellulosic material; contemporary or "first-generation" biofuels are derived from sugar, starch, or vegetable oils, whereas emerging "second-generation" biofuels are derived from lignocellulosic materials.

Ecosystem services The benefits that humans derive from ecosystems; for example, timber and crop production, water purification, pest control, pollination, and outdoor recreation.

Feedstock, bioenergy Raw, unprocessed biomass slated for conversion to bioenergy.

Reserves Areas in which the primary management objective is to fully protect native species and ecosystems from direct human modification.

Targeted conservation Strategic land allocation and application of management practices, such as buffers, corridors, constructed wetlands, stepping stones, and reserves, to maintain species populations and ecosystem services in concert with human land use.

Wildlife All nonhuman and nondomesticated animals.

Introduction

In an effort to reduce dependence on fossil fuels and enhance energy security, many countries are mandating the use of bioenergy and heavily investing in its research, development, and production. The European Union, for example, established a directive on the use of renewable energy in 2008 requiring that 20% of the energy consumed by member states be derived from renewable sources with biofuels providing at least 10% of all transportation fuels by 2020 (Commission of the European Communities, 2008). Brazil, China, and the USA have adopted similarly aggressive bioenergy policies in recent years. At first glance, these mandates and the bioenergy production they spark may appear to be a boon for the environment, including wildlife; society is attempting to make much needed shifts from fossil fuels, curb greenhouse-gas emissions, and reduce the impacts of future climate change. Closer scrutiny, however, reveals, that large-scale bioenergy production may constitute an environmental boondoggle (Field *et al.*, 2008; Robertson *et al.*, 2008; Eggers *et al.*, 2009; Tilman *et al.*, 2009), largely because of shifts in land use that are occurring to support it. Ultimately, whether future bioenergy production is helpful or harmful to wildlife depends on society's choice of bioenergy production systems and their configuration at a variety of scales.

The demand for bioenergy feedstocks is already affecting ecosystems around the globe. As of 2010, 8.3 million ha (Mha) of biodiversity-rich tropical lowland forest in Borneo, Peninsular Malaysia, and Sumatra had been converted to oil palm plantations to serve both food and biofuel markets (Koh *et al.*, 2011), with biofuels predominantly driving this expansion in the past decade. Furthermore, Indonesia plans to double its production of oil palm by 2020 to supply the European and East Asian markets (Koh and Ghazoul, 2010). Starting in 1975, Brazil began a large-scale program to produce biofuel from sugarcane to reduce dependency on oil imports. With nearly 7 Mha of agricultural land currently devoted to sugarcane production, Brazil is second only to the USA in bioethanol production, reaching 19 billion liters in 2007 (Monteiro *et al.*, 2010). Expansion of sugarcane production has historically resulted in the direct clearing of the biodiversity-rich Cerrado and Atlantic forest habitats (Silviera *et al.*, 2003), and indirectly to deforestation of Amazonian forests through the displacement of soybean production and pasturelands (Martinelli and Filoso, 2008). In the Brazilian state of Alagoas, sugarcane production has led to the loss of 97% of the original forest cover (Power and de Araujo, 1993). In the USA, national biofuel policies are the partial cause of shifts from a traditional corn–soybean biennial crop rotation to corn–corn–soy rotation or continuous corn cropping and the

conversion of grasslands, including native prairies and other previously untilled areas (Fargione *et al.*, 2009; Babcock and Fabiosa, 2011). Although the exact extent of grassland loss is unknown, the US Department of Agriculture projects the amount of land enrolled in the Conservation Reserve Program, which are predominantly grasslands, will decline from a peak of 14.9 Mha in 2007 to 12.2 Mha in 2013 as these lands are converted to row-crop production (Fargione *et al.*, 2009).

Whether bioenergy will ultimately help or harm wildlife is yet to be determined and, as is often the case regarding human interactions with the environment, the answer will depend on how bioenergy is produced. The environmental impact of bioenergy production is dependent on the chosen production systems – as defined by the types of crop(s) grown, where they are grown, how they are grown, and their transport, storage, and conversion to energy (Dauber *et al.*, 2010). Assessments that holistically consider the environmental tradeoffs associated with different bioenergy production systems are rare. Here the authors review and synthesize global trends, constraints, and opportunities in the development of bioenergy with respect to their impacts on wildlife. In particular, they address the potential for the positive use of bioenergy crops to restore wildlife habitat within agricultural landscapes. Their assessment is based on a combination of theory and literature review. Although the perspective of their synthesis is global in character and representative of all nondomesticated animal species, the information available to support this assessment is limited in breadth, depth, and distribution. Empirical data on wildlife response to bioenergy development have been collected for only a limited number of taxa – mostly birds – and bioenergy crops in Europe, North America, and Southeast Asia (Koh and Wilcove, 2008; Dauber *et al.*, 2010). Although much more research is needed on the impacts of specific bioenergy cropping systems on wildlife across regions of the globe, existing knowledge and established ecological and conservation concepts suggest some guidelines for bioenergy development to be sustainable with respect to wildlife.

Impacts of Changing Land Use – Native Ecosystems

The negative environmental impacts of converting native ecosystems to other uses have been well established within the conservation literature (Sala *et al.*, 2000). Thus, most conservation ecologists agree that bioenergy development that requires conversion of native ecosystems to other land uses, whether directly or indirectly, will result in a net loss for biodiversity and wildlife (Aratrakorn *et al.*, 2006; Bies, 2006; Koh and Wilcove, 2008; Fargione *et al.*, 2009; Edwards *et al.*, 2010). Quantitative studies tracking species richness and abundance with land-use conversion represent only a small number of bioenergy crops, taxa, and regions; results from the few studies that compare native ecosystems and bioenergy systems however report declines for most wildlife species. Specifically:

- Koh and Wilcove (2008) report that bird and butterfly richness are lower – 77% and 83%, respectively – in oil palm plantations compared to primary forest in Peninsular Malaysia and Borneo.
- Edwards *et al.* (2010), also working in Borneo, found imperiled bird species to be 60 times less abundant in forest fragments and 200 times lower in oil palm plantations compared to contiguous native forest.
- Aratrakorn *et al.* (2006) show that conversion of native lowland forest to commercial oil palm and rubber plantations has resulted in a greater than 60% reduction in bird species in Thailand. Furthermore, 94% of bird species of regional conservation concern were only found in native lowland forests.
- Martin and Catterall (2001) record very low avian biodiversity in sugarcane fields relative to residential areas and heathland remnants in eastern Australia.
- Petit *et al.* (1999) show that sugarcane fields supported the lowest numbers of total and migratory bird species compared to 10 other land-cover types in central Panama, with 84% fewer species than lowland forests.
- Riffell *et al.* (2011), in a meta-analysis of eight North American studies, found bird and small mammal diversity and abundance to be consistently lower in short-rotation woody crop (SRWC) plantations relative to native forest cover types.

These types of impacts are not equally distributed across the globe (Hansen *et al.*, 2008) nor limited to the terrestrial environment. Land-use conversion within terrestrial ecosystems tends to severely alter the habitat quality of adjacent aquatic ecosystems through increased sedimentation, turbidity, and water temperatures (Tilman, 1999) and pollution with agricultural chemicals (Relyea, 2005). Furthermore, land-use change acts synergistically with other stressors on ecosystems, such as climate change, altered nutrient cycles, and invasive species, to affect species losses. The authors of the Millennium Ecosystem Assessment predict that these changes may cause global biodiversity declines of 12–16% by 2050 (MEA, 2005). Up to 80% of the species lost – approximately 30,000 species – are expected to come from biodiversity hotspots, locations with exceptionally high concentrations of endemic species such as tropical forests and woodlands, savannas, and warm mixed forests (Bradshaw *et al.*, 2009; MEA, 2005). Accordingly, the clear negative impact of the conversion of native ecosystems suggests that, for bioenergy development to minimize its impacts on wildlife, it should avoid conversion of native ecosystems, especially in biodiversity hotspots.

Further negative environmental and social impacts are associated with the conversion of native ecosystems for bioenergy production, thereby strengthening the support for this assertion. These include:

- reductions in air quality (Hill *et al.*, 2007), pest control (Landis and Werling, 2010), and pollination services (Öckinger and Smith, 2007; Landis and Werling, 2010);
- potential increases in greenhouse-gas emissions (Fargione *et al.*, 2008; Scharlemann and Laurance, 2008); and
- the displacement of existing landholders and/or native peoples (Coutla *et al.*, 2008; Bekunda *et al.*, 2009; Schoneveld *et al.*, 2010).

Whether the displacement of fossil fuels with bioenergy results in a net reduction of greenhouse-gas emissions is particularly noteworthy, because bioenergy is often touted as a

"green" solution and greenhouse-gas emission standards are linked to bioenergy policies in the Europe Union, USA, China, and elsewhere. Although many studies calculate a reduction in greenhouse-gas emissions associated with the combustion of biofuels in comparison to fossil fuels, comprehensive assessments of bioenergy production are not always favorable. For example, Searchinger *et al.* (2008) compute that, when land-use change is factored in, biofuels result in more than doubled greenhouse-gas emissions compared to fossil fuels. Fargione *et al.* (2008) estimate that the conversion of natural ecosystems to food crop-based biofuels causes a net total release of 17–420 times more CO_2 than that annually displaced by using biofuels instead of fossil fuels. Some of the biofuel systems considered in these and similar analyses include sugarcane bioethanol, corn bioethanol, palm biodiesel, and soybean biodiesel. If projections like these are correct, many of the systems being widely implemented today, when coupled with direct or indirect conversion of native ecosystems, actually exacerbate climate change – and its expected impacts on biodiversity and wildlife (Dawson *et al.*, 2011) – compared to fossil fuels (Zah *et al.*, 2007 (as reviewed by Scharlemann and Laurance, 2008); Fargione *et al.*, 2008; Lapola *et al.*, 2010). These results suggest that sustainable bioenergy production should result in overall lower greenhouse-gas emissions compared to the fossil energy. The determination of the greenhouse-gas emissions from a given biofuel requires a full life cycle analysis of direct and indirect impacts.

Native ecosystems are often highly productive; many produce copious amounts of woody and herbaceous biomass, a portion of which could be sustainably harvested as bioenergy feedstocks (Hoogwijk *et al.*, 2003; Perlack *et al.*, 2005). Indeed, biomass harvested from native ecosystems is the original bioenergy. In the forestry sector, low-value woody material generated in the milling process (e.g., small or deformed material, bark, and sawdust) has long been used to cogenerate heat and power to operate mills, offsetting energy costs, and effectively using what would otherwise be considered a waste product. As energy costs rise, could harvesting, collecting, and hauling additional material from forests savannas, or grasslands become more economically viable? Perhaps not, as woody and herbaceous biomass constitutes food and cover for biodiversity and wildlife when kept in place. In some cases, however, biomass harvest and removal can be helpful in meeting ecological restoration goals. Examples include managing woody encroachment into prairies and savannas where prescribed fire cannot be used because of proximity to buildings or roads; when mechanical thinning of forests is necessary due to decades of fire suppression; or as a means to control an exotic plant that has thoroughly invaded an ecosystem. In these instances, biomass harvests may be a source of revenue that helps to pay for or offset the cost of ecological restoration. Several European nations, Canadian provinces, and US states have articulated specific guidelines as to where, when, and how woody biomass harvests can be done sustainably (Evans *et al.*, 2010). Many of the guidelines relate to soil quality, water resources, and aesthetic resources (Evans *et al.*, 2010). Guidelines that specifically relate to biodiversity and wildlife include:

- the overall quantity and quality of habitat should be maintained or improved;
- communities of native plants and wildlife should remain intact; and
- the introduction of invasive or nonnative species should be avoided (Raghu *et al.*, 2006).

These considerations suggest that, to be sustainable, biomass harvests should not compromise the long-term functioning of native ecosystems.

Impacts of Changing Land Use – Crop and Pasturelands

Social concerns regarding the ethical use of agricultural lands have led to the assertion that bioenergy feedstocks should not displace food production on prime agricultural lands, but rather be grown on other portions of the global land base. For ethical and food security reasons, many people believe agricultural lands should produce food and not fuel (note: the authors use "food" to comprehensively include crops grown for direct human consumption, processed for human consumption, and fed to livestock raised for human consumption). Yet, growing proportions of the world's cassava, corn, rapeseed oil, palm oil, soybean oil, and sugarcane crops are being routed to bioenergy production streams (**Figure 1**). Additionally, energy crops are now being grown on lands traditionally used to produce food. These shifts are partially responsible for global increases in food prices and food shortages in some countries (Naylor *et al.*, 2007; FAO, 2009). Undernourishment has been on the rise since the mid-1990s, but rose sharply between 2006 and 2009, peaking at just more than a billion people (FAO, 2009). In addition to the human cost, lack of food or rise in food prices can be costly for wildlife. People struggling with poverty sometimes turn to bushmeat as an alternative, though unsustainable, food source when prices of agricultural goods rise (Bennett, 2002; Milner-Gulland and Bennett, 2003). As with land-use change, the incidence and effects of bushmeat hunting are not uniformly distributed. Bushmeat consumption tends to be disproportionately greater in tropical regions, which support some of the world's poorest people and two-third of all species (Milner-Gulland and Bennett, 2003; Bradshaw *et al.*, 2009).

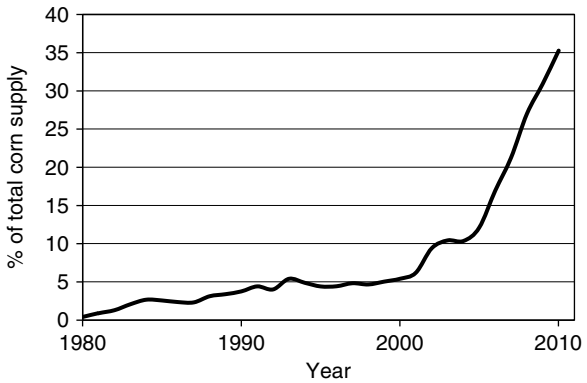


Figure 1 Percent of annual US corn crop used for bioethanol production (USDA ERS, 2011), an example of the increasing use of food crops to generate biofuels.

Furthermore, poverty and undernourishment are correlated with agricultural methods that generate immediate increases in production at the expense of long-term environmental degradation, such as overgrazing, slash-and-burn deforestation, and unsustainable use of soil and water resources (Reynolds and Stafford Smith, 2002).

The food shortages and volatile food prices of recent years may be more of a result of trade distortions, rising energy costs, and a lack of infrastructure than lack of food (Naylor *et al.*, 2007; Heady and Fan, 2008; FAO, 2009; Thurow and Kilman, 2009). The increase in foreign holdings in major food-importing countries and attempts made by food-exporting countries to mitigate their own food price inflation have played key roles in both facilitating and hindering food distribution (Anderson *et al.*, 2006; Mitchell, 2008; Thurow and Kilman, 2009). Furthermore, many regions are not as productive as they could be, and food shortages could be overcome by developing the agricultural capacity of these areas, a goal that has proven to be highly elusive for the past 50 years. If resource (e.g., inability to access fertilizer and seeds) and infrastructural (e.g., inability to move food to markets) obstacles to food production and dissemination in sub-Saharan Africa could be overcome, this might alleviate regional hunger with little competition from bioenergy (FAO, 2010). Consequently, if such inefficiencies were overcome and food production were focused on regionally prime agricultural lands, there might be land available to offset a meaningful portion of global energy demand.

Marginal, degraded, or abandoned agricultural lands offer more socially, and thus ecologically, viable options for bioenergy production (**Box 1**; Field *et al.*, 2008; Robertson *et al.*, 2008; Tilman *et al.*, 2009). A recent assessment by Campbell *et al.* (2008) suggests there is approximately 385–472 Mha of abandoned farmland in the world, which could annually produce 1.4–2.1 billion tons of biomass. If converted to biofuel, this biomass could meet up to 8% of current global energy demand. More efficient and strategic use of the global land base is needed, however, to concomitantly meet food, energy, and environmental pressures (Robertson *et al.*, 2008; Liebman *et al.*, 2011). The next two sections, respectively, review the ability of bioenergy systems to cogenerate these benefits and how these systems may be strategically allocated for this purpose.

Box 1 Prime, marginal, degraded, and abandoned lands

Prime agricultural lands are those possessing the best combination of characteristics, both physical (e.g., moisture status, soil depth, and slope condition) and chemical (e.g., nutrient status and salinity), for producing crops and livestock. Bioenergy crops are often slated for "marginal," "degraded," or "abandoned" lands. Marginal is more variably defined than prime (Dale *et al.*, 2010), but marginal agricultural lands are generally considered to have one or more conditions that make them less conducive to crop or livestock production. These conditions include excessive slopes with high susceptibility for soil erosion when tilled or grazed, excessive wetness, poor soil texture for water and nutrient retention, and shallow depths to bedrock or water table. Degraded lands are current or former crop or pasturelands where the productivity has been semipermanently reduced because of human actions. Common causes of degradation include the removal (i.e., land-cover conversion) or mismanagement (e.g., overgrazing) of natural vegetation; improper agricultural management (e.g., poor tillage or irrigation practices); and industrial activities (e.g., acid rain). According to the Global Assessment of Human-induced Soil Degradation (GLASOD), 15% of the global land base is degraded – an area about twice the size of the USA. Continentally, this breaks down to 23% of Europe, 18% of Asia, almost 17% of Africa, almost 14% of South America, almost 12% of Australasia, and 5% of North America (Oldeman *et al.*, 1991; **Figure 2**). Abandoned agricultural lands are those once used to grow crops or as pasture, but no longer used because of farmer relocation or degradation due to intensive use. Field *et al.* (2008) estimated the global extent of abandoned agricultural lands to be 386 Mha, or an area comparable to about half the size of Australia, though with a sizable margin of error. Assessing the extent and location of abandoned lands is difficult because of shifting cultivation practices of native peoples, who may periodically use lands that appear to be deserted, for some other reasons.

Comparison of Bioenergy Systems

Numerous bioenergy systems are being considered for meeting present and future renewable energy needs. These systems vary in terms of their productivity, management, and impacts on wildlife. The impacts of bioenergy systems on the richness and abundance of wildlife species involve reciprocal interactions between crops and wildlife – these interactions influence crop productivity as well as the suitability and environmental implications of bioenergy production (Dale *et al.*, 2010). The authors group and evaluate these systems according to two gradients; the first is associated with the level of agricultural inputs (e.g., chemicals, equipment, seed, and fuel) required to maintain their productivity and the second is the level of plant diversity of composition. Together these gradients form four contrasting categories of bioenergy systems: high-input low-diversity (HILD), low-input low-diversity (LILD), high-input high-diversity (HIHD), and low-input high-diversity (LIHD) (**Figure 3**).

Current food-based crops, so-called first-generation biofuels, are largely composed of HILD systems. Bioethanol and biodiesel are produced from highly productive annual crops rich in sugars, starches, or oils such as corn (*Zea mays* L.), sugarcane (*Saccharum* spp.), oil palm (*Elaeis guineenis*, *Elaeis oleifera*), rapeseed (*Brassica* spp.), and soybean (*Glycine max* (L.) Merr.), among others. First-generation bioenergy crops are most often grown in annual monocultures, which require high levels of fuel, fertilizer, and other petrochemical inputs to maintain their productivity. This reliance on fossil fuels fundamentally compromises the net energy efficiency of conventional bioenergy crops (Kim and Dale, 2005; Hill *et al.*, 2007). For example, Rathke *et al.* (2007), using field experiments comparing corn and soybean production with different tillage treatments and crop rotations, found that crops grown in monocultures had a lower net energy gain than crops grown in rotation. The reason for differences included the higher level of inputs required by monocultures, whereas reduced tillage treatments maximized the energy efficiency from lower fuel costs.

Agricultural practices within HILD systems deliberately maintain ecosystems in a simplified and disturbed state to maximize crop productivity by minimizing resource competition from noncrop plants. These ecosystems provide habitat for only a narrow suite of wildlife, primarily generalist

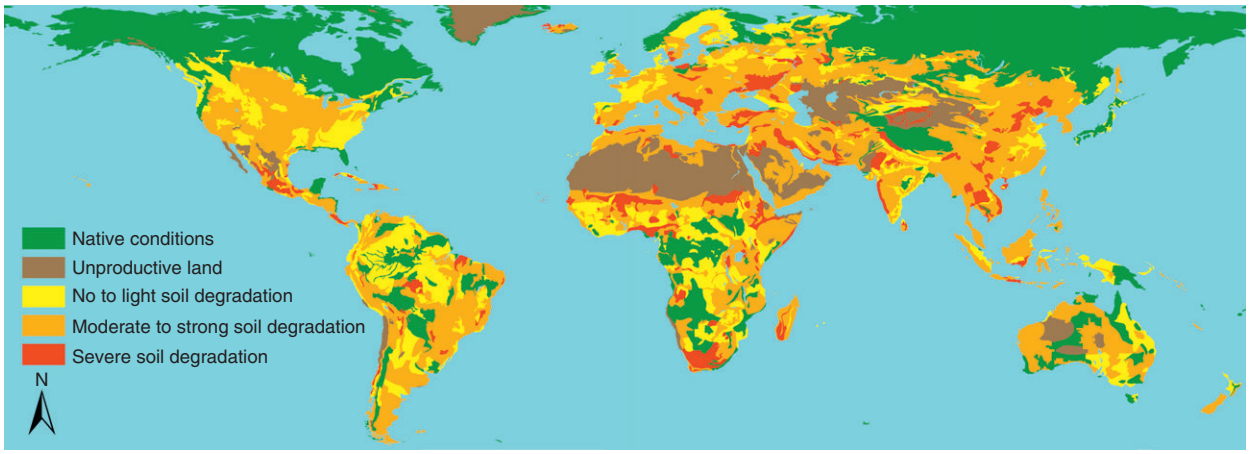


Figure 2 Global distribution of degraded soils. Data from the Global Assessment of Human-induced Soil Degradation (GLASOD); Oldeman LR, Hakkeling RTA, and Sombroek WG (1991) *World Map of the Status of Human-Induced Soil Degradation*, 2nd edn. Wageningen: ISRIC.

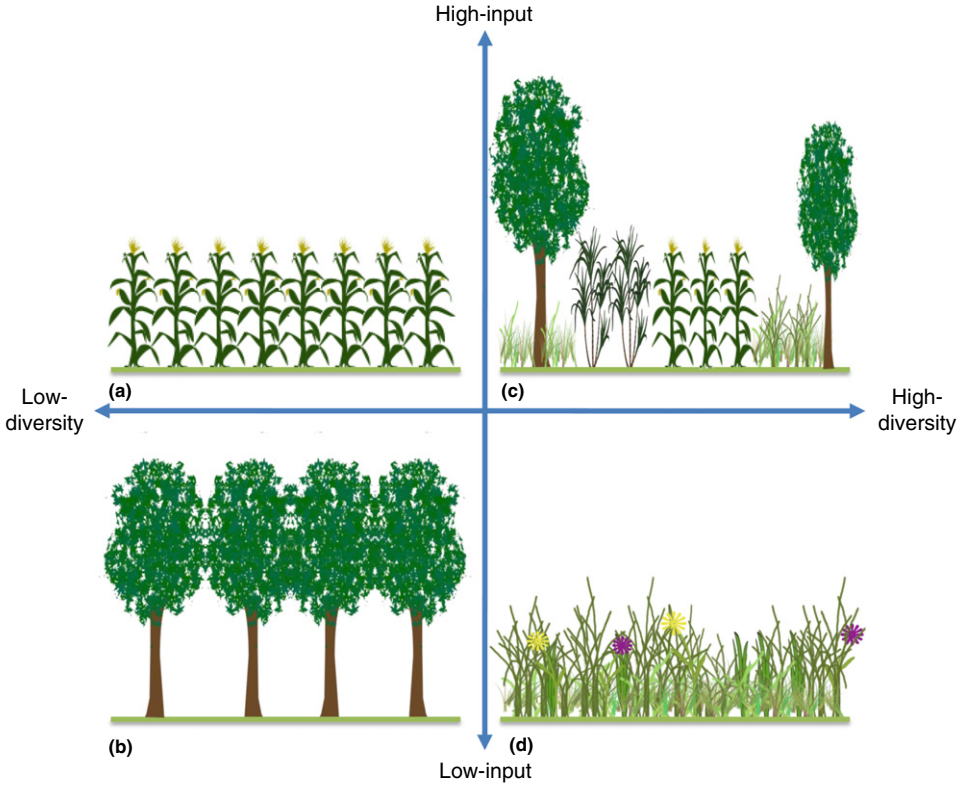


Figure 3 A visual comparison of the composition and structure of example cropping systems that can be used as feedstocks for bioenergy production: (a) corn, (b) hybrid poplar, (c) strip cropping with hybrid poplar, miscanthus, corn, and (d) prairie.

species (Robertson *et al.*, 2011) and sometimes favor crop pests (Landis and Werling, 2010). The simplified vegetative structure of monocultures equates to low habitat complexity, fewer species, and altered foodwebs (Landis and Werling, 2010). Often feedbacks manifest where crop pests favoring the simplified structure trigger greater use of tillage and chemical herbicides or pesticides (Matson *et al.*, 1997), thereby further reducing habitat quality for wildlife. Increased intensification through agricultural inputs also tends to have nontarget impacts, producing a myriad of stressors on wildlife. For example, biodiversity in aquatic systems can be affected by agricultural practices through: direct lethal effects of pesticides (Relyea, 2005); altered hydrological flows from crop irrigation (Roberts *et al.*, 1997); and nutrient or sediment pollution of freshwater and marine ecosystems, leading to harmful algal blooms and reducing aquatic species richness (Carpenter

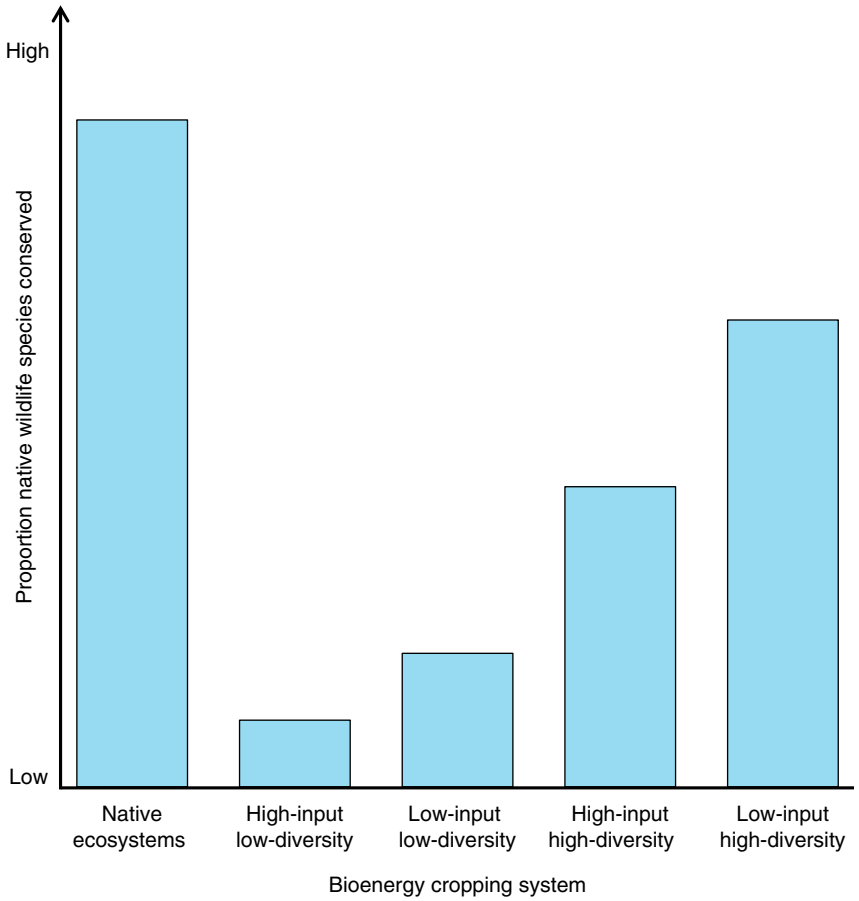


Figure 4 Total number of native wildlife species likely to be conserved in native ecosystems vs. different bioenergy cropping systems. Although the actual proportions are hypothetical and dependent on the specific systems under consideration, existing knowledge supports the relative pattern shown.

et al., 1998). In the USA, agricultural practices are the source of impaired water quality on approximately 72% of river miles (EPA, 1994), and are the leading cause of freshwater species declines (Richter *et al.*, 1997). In sum, when considering all potential types of bioenergy systems, HILD systems provide the least benefit to wildlife (**Figure 4**).

Cellulosic materials left over from the harvest of HILD crops (e.g., corn stover; sugarcane bagasse; rice, wheat, or oat straw) have significant potential as second-generation bioenergy feedstocks. However, these crop residues are important for maintaining soil quality through recycling plant nutrients, storing carbon, and improving water holding capacity. Maintaining residues on site is especially important for reducing soil erosion on sloping and erosion-prone terrain (Blanco-Canqui and Lal 2009). Although indiscriminant removal of crop residues would accelerate soil degradation of prime agricultural lands, removal of a portion of residues could provide a substantial source of bioenergy feedstock while minimizing environmental risks – Nelson (2002) estimated that up to 30% of crop residue can be removed without exceeding tolerable soil erosion levels. Forty percent residue removal could produce 442 billion liters of bioethanol annually, potentially replacing 32% of global gasoline consumption (Kim and Dale, 2005). Beyond the impacts on soil quality, removal of crop residues affects the wildlife habitat value of bioenergy systems. Residues provide cover and a food source for deer, waterfowl, upland game birds (e.g., turkey, grouse, quail, and pheasant), songbirds, small mammals, and other animals (WHMI, 1999), while providing habitat for insects, an important food source for birds and mammals. For example, Best (1986) found bird species richness was higher in fields with high cover of crop residue (15 species) compared to those with low residue cover (six species). Although crop residues may appear to have some potential as bioenergy feedstocks, the negative impacts on soil resources and the subsequent feedbacks on wildlife further reduce the already meager habitat value of HILD systems.

By comparison, LILD systems benefit a slightly greater number of wildlife species (**Figure 4**), due to the reduced level of inputs required to maintain productivity. Many, but not all, of these systems are composed of second-generation bioenergy crops; they rely on the conversion of lignocellulosic biomass into bioethanol. Low-input systems are typically composed of herbaceous or woody nonfood perennial species that have low fertilizer requirements because of high nutrient-use efficiency of the plants. Low fertilizer requirements lead to decreased plant-soluble nitrogen concentrations, reducing pathogen, pest insect, and weed populations and thus minimize the need

for harmful herbicide and pesticide inputs (Matson *et al.*, 1997). In a comparison of high-input to low-input systems, Gardiner *et al.* (2010) found that low-input switchgrass and mixed-grass prairie fields both supported a higher diversity of beneficial insect, including several rare or declining invertebrate species, compared to fields of high-input corn. Moreover, a perennial growth habit reduces energy inputs needed for annual tillage and planting, which minimizes soil erosion and leads to reductions in both the inputs needed to maintain productivity and their associated impacts on aquatic biodiversity. Fewer annual field activities in combination with the longer rotation periods of perennial systems result in fewer mechanical disturbances to terrestrial wildlife, such as ground-nesting bird species (Murray *et al.*, 2003).

There is a growing interest in LILD systems because of their increased sustainability and reduced cost of production. Many of these systems can be highly productive with few inputs on both degraded and marginal croplands (**Box 1**), reducing conflicts between bioenergy and food or fiber production (Somerville *et al.*, 2010). Examples of LILD systems that are being vetted for production on degraded or marginal lands in Europe and North America include perennial grasses such as miscanthus (*Miscanthus giganteus*) and switchgrass (*Panicum virgatum*); SRWC systems such as willow (*Salix* spp.) and poplar (*Populus* spp.); and pine (*Pinus* spp.) species grown in tree plantations. Compared to HILD systems, LILD systems show positive effects on nearly all taxa studied in temperate regions of USA and Europe (Dauber *et al.*, 2010). In an analysis of published studies comparing wildlife responses among native habitats and bioenergy systems, Fletcher *et al.* (2011) found that SRWC plantations affected bird species abundance less negatively than did arable crops. Moreover, the vertical structure found in plantations provides habitat to a broader suite of species, including ground-, canopy-, and cavity-nesting birds (Hanowski *et al.*, 1997). In comparison, row crops provide habitat for only ground-nesting birds that require little cover (Best *et al.*, 1997).

The key to simultaneously boosting the productivity and wildlife habitat benefits of LILD bioenergy systems is matching the suitability of the crop to the regional climate. For example, miscanthus may achieve higher yields of biomass than annual crops like corn, but it requires as much as 60% more water per year. The sustainable production of this feedstock is thus limited to areas with high natural water availability (Hickman *et al.*, 2010), and it is not suitable for arid and semiarid climate zones. Yet, these regions often have abundant degraded land due to unsustainable use in the past; they thus, present vital opportunities for bioenergy production for both economic and environmental benefits. Alternatively, plants in the genus *Agave*, which have traditionally been grown for fiber and alcoholic beverages in Central America, are highly productive and require far less water, making them an attractive feedstock in arid and semiarid regions worldwide (Davis *et al.*, 2011). Likewise, jatropha (*Jatropha curcas* L.), a long-lived tropical tree that is productive on marginal soils, has evoked much interest as a biodiesel feedstock for degraded lands in Africa, Asia, and South America. Jatropha requires few inputs due to its tolerance of drought, low nutrient requirements, and the lack of known pests owing to the toxicity of the leaves and fruit – although toxicity may significantly reduce jatropha's value for wildlife. Little work has been done that assesses the wildlife habitat value of these and other LILD systems suited for marginal or degraded agricultural lands in dry climates. Although LILD systems provide meager direct benefits to wildlife (**Figure 4**), other benefits include reduced indirect land-use change resulting from the displacement of food crops, especially in biodiversity hotspots such as tropical ecosystems.

HIHD and LIHD are more wildlife friendly, owing to the greater level of plant diversity contained within them (**Figure 3**). Wildlife respond positively to the greater habitat complexity associated with more heterogeneous vertical and horizontal structure contained within these systems. High-diversity systems produce bioenergy crops using polycultures where multiple species are grown either in mixtures or mosaics, rather than the production of a single species over large areas as in monocultures. Polycultures can take many different forms, and are considered one of the hallmarks of wildlife-friendly farming (Green *et al.*, 2005). HIHD systems include strip cropping, where rows of different crops are alternated (**Figure 3(c)**); relay cropping, where a second crop is seeded before the first is harvested; and double cropping, where a second crop is planted after harvest of the first to reduce fallow periods. Advantages of these production methods can include higher overall yields, greater energy efficiency, and reduced inputs compared to HILD systems because the greater biodiversity results in fewer crop pests due to biological pest control (Matson *et al.*, 1997). Examples of HIHD cropping approaches for bioenergy include the incorporation of strips of perennial vegetation, such as hedgerows of fast-growing trees and filter strips of perennial grasses, into areas of annual row crops. For example, Hinsley and Bellamy (2000) showed that enhanced plant diversity and structural heterogeneity associated with increasing area of hedgerows and herbaceous strips on crop field margins is associated with greater biodiversity in the UK. Agroforestry systems in the tropics that consist of multiple canopy layers, such as mature trees, shrubs, and annual crops have been shown to have high biodiversity value for birds, butterflies, and plant species compared to annual crops (Schulze *et al.*, 2004).

HIHD systems for bioenergy have not been widely implemented across agricultural landscapes, so their impacts on wildlife have not been directly assessed; yet, the value of these systems can be predicted based on the observed changes in biodiversity with agricultural intensification during the second half of the twentieth century. In a review of the literature, Benton *et al.* (2003) discussed the effects of agricultural intensification across Europe and North America, and the consequent decline in habitat heterogeneity and biodiversity. For example, Fuller *et al.* (1995) found that 86% of farmland bird species had reduced ranges and 83% had declined in abundance in the UK between 1970 and 1990, as agriculture intensified. The same trends are well documented in North America.

HIHD systems are expected to provide greater benefits to wildlife compared to low-diversity systems; however, LIHD systems trump all others, because of the synergistic positive effects of increased habitat complexity and fewer agricultural inputs (**Figures 3** and **4**). They result in fewer indirect negative impacts and provide the greatest habitat benefits for wildlife. LIHD systems consist of polycultures of perennial grasses and

forbs, and forestry systems that produce a variety of tree and herbaceous species. Tilman *et al.* (2006) showed that polycultures of prairie species produce large quantities of biomass on marginal or degraded sites in the USA with few inputs. Although similar herbaceous polyculture systems have not been tested elsewhere, marginal and degraded lands in other regions have high potential for LIHD systems that benefit wildlife. In a comparison of the effects of plant diversity in low-input grassland systems, Robertson *et al.* (2011) found bird abundance, especially that of grassland specialists, was higher in LIHD mixed-grass prairie compared to LILD switchgrass stands, highlighting the importance of structural heterogeneity of the habitat patch. In general, the sustainability and habitat value of bioenergy systems, like other agricultural production systems, may be maximized when systems mimic both the structure and function of the native ecosystems they replace (Jackson and Jackson, 2002).

Global, Regional, Landscape, and Local Allocation

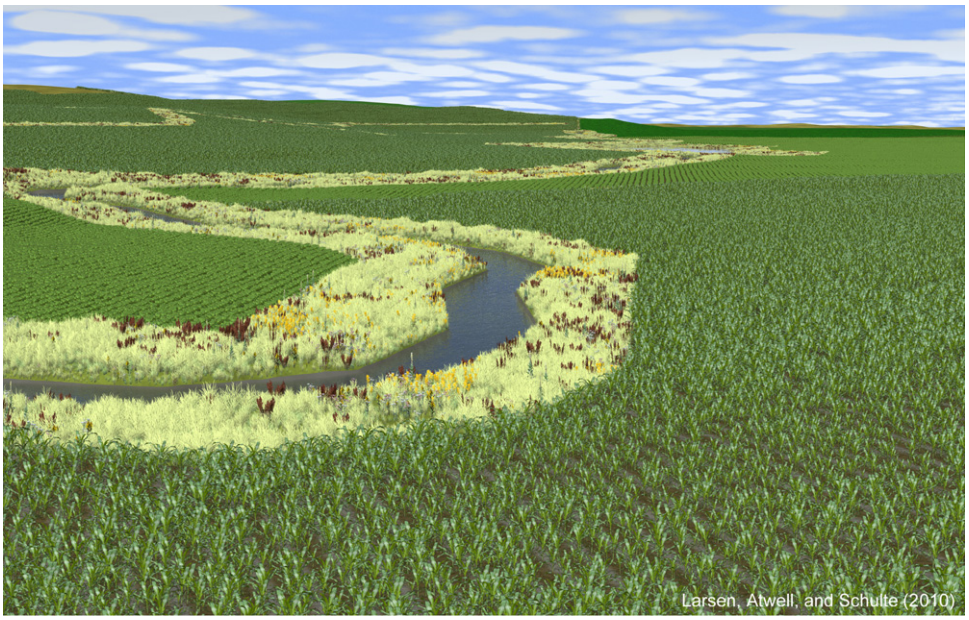
All the systems reviewed here – native ecosystems, food systems, and bioenergy systems – can be a part of a sustainable future that includes bioenergy production. In comparison to the status quo, however, these systems need to be implemented more strategically from global to local scales in order to be a boon. Multidimensional planning, when combined with multistakeholder collaboration, offers a powerful pathway for achieving the coproduction of food, fuel, and environmental benefits, including wildlife conservation. Dauber *et al.* (2010) provided a comprehensive assessment, but here the authors provide a few key recommendations for collaborative groups to consider at each scale: global, regional, landscape, and local. Given that any particular system or configuration of systems may benefit some wildlife species and be detrimental to others (Landis and Werling, 2010), they make these recommendations in light of conserving the largest number of wildlife species possible. Defining system-specific wildlife-related objectives is a critical part of the planning process (Groves, 2003).

Because anthropogenic climate change is manifested globally, global scale assessments are needed to determine whether bioenergy production systems actually meet the goal of reduced greenhouse-gas emissions in comparison to fossil fuels (Fargione *et al.*, 2008; Dauber *et al.*, 2010). Although evaluations at finer-scales are necessary for designing and implementing bioenergy production systems specific to regional socioeconomic and environmental context, only assessments global in breadth will reveal whether bioenergy development will be a boon or boondoggle for wildlife. Bioenergy systems that reduce regional greenhouse-gas emissions at the expense of releasing disproportionate amounts of stored carbon through indirect land-use change – for example, through conversion of tropical rainforests, drainage of wetlands, and tillage of unbroken sod – are clearly harmful to wildlife.

Regarding the preservation of native ecosystems, multidimensional assessment, planning, and regulation can be useful for developing and implementing a network of biodiversity reserves to achieve adequate protection of target species and ecosystems (Koh, 2007). Biodiversity reserves are areas in which the primary management objective is to fully protect existing ecosystems and species populations from direct human modification. Human activities, such as crop production or biomass harvesting, are often prohibited and at least severely restricted within reserves. In 2002, the United Nations Environment Programme's Convention on Biological Diversity adopted a resolution requiring member countries to target at least 10% of each ecological region for placement in reserves. This percentage should be far greater in global biodiversity hotspots (Myers *et al.*, 2000), such as Southeast Asia, Mesoamerica, the Caucasus, and the Horn of Africa. Although delineating biodiversity hotspots is a global matter, reserve networks – which consider the movement of genes, individuals, and species among individual reserves – are generally constructed within regions, which are often defined by political boundaries rather than ecological regions per se. Arranging reserves in networks is an essential strategy for adapting to and mitigating the impacts of global climate change, especially given that reserves exist within a matrix of highly altered landscapes (Ellis and Ramankutty, 2008). Although reserves are an important conservation strategy (Noss and Cooperrider, 1994; Groves, 2003), sensible bioenergy policies ought to discourage the conversion of native ecosystems in any location. Yet, native ecosystems may still play a role in bioenergy production if biomass harvests can assist in meeting local restoration goals and do not compromise the long-term functioning of native ecosystem.

Reserves alone however are unlikely to fully protect wildlife in the face of climate change and other stressors (Fischer *et al.*, 2006); they are too few in number and too widely spaced. Consequently, efforts to protect wildlife must be integrated into all land uses. Agricultural landscapes need to be redesigned and to be more hospitable to wildlife by providing more structurally complex habitat than is common within many contemporary agricultural landscapes, which tend to utilize HILD systems, the least desirable for wildlife (**Figure 4**).

The production of bioenergy feedstocks that do not displace food production on prime agricultural lands is manifested over regional and landscape (1–100 km^2) scales. Within regions, some locations are more suited for agricultural production than others, and most existing biodiversity reserves are already relegated to less productive portions of the land base (Noss and Cooperrider, 1994). For example, within the western US, productive river valleys are often under intensive agricultural land use, foothills comprise grazing lands, and the least productive high elevation lands are in reserves (i.e., national parks). Given such spatial arrangements, LIHD bioenergy systems that make use of native species can be targeted to locations surrounding biodiversity reserves; thus, buffering reserves from the impacts of high production agriculture and potentially increasing the core area of wildlife habitat within the reserve while providing a marketable crop. Moreover, all agricultural landscapes have portions that are inherently less productive due to their edaphic conditions; for example, low-lying wet areas, high dry or stony areas, or highly erosive areas along steep slopes. On these portions of the landscape, bioenergy cropping systems could be strategically deployed in schemes that facilitate the coproduction of biomass and environmental benefits as well as serve as buffers, corridors, and stepping stones for wildlife. Buffers protect sensitive habitats,



(a)



(b)

Figure 5 Visualizations of alternative agricultural landscapes that support, food, bioenergy production, and wildlife. (a) Dominated by a HILD corn system in which corn grain is used for both food and to produce bioethanol. In this scenario, the streamside has been set aside and covered with LIHD prairie to protect the stream from the impacts of corn production. This homogenous landscape is expected to support only a small portion of the wildlife species native to the region. (b) Includes a mixture of agricultural crops including HILD corn for food, LILD switchgrass for cellulosic bioethanol production (foreground), a LIHD reconstructed prairie for forage or cellulosic bioethanol production (streamside), and a LIHD tree and grass intercropping system (back left) for fiber or cellulosic bioethanol production. In this scenario, all sensitive (streamside) and degraded (hillside) agricultural lands have been converted to perennial crops. This heterogeneous landscape is expected to support the majority of wildlife species native to the region.

whereas corridors and stepping stones are designed to facilitate the movement of organisms and genetic material (Fischer *et al.*, 2006). **Figure 5** provides a hypothetical example of how targeting might look different from the status quo (**Figure 5(a)**). A redesigned agricultural landscape might include an LIHD system that is based on a native vegetation type used in combination with an LILD system to form a gradient in habitat suitability between an HILD food system and sensitive aquatic habitats (**Figure 5(b)**). An HIHD system that combines SRWCs and an herbaceous bioenergy crop in an ally cropping system may be targeted to highly erodible lands, providing marketable crops, while also introducing

compositional and structural heterogeneity into an area otherwise dominated by HILD crops (**Figure 5(b)**). Employing a variety of bioenergy cropping systems across a landscape may allow the bioenergy production systems to mimic both the structure and function of the native ecosystems they replace.

Finally, bioenergy systems also have an important role to play in improving the ecological functioning and wildlife habitat at the local, field scale (<1 km^2) on prime agricultural lands. Although the goal at this scale might not be enhanced wildlife habitat per se, biodiversity can provide ecosystem services that are critical for maintaining agricultural productivity, lowering the cost of farming (Higginbotham *et al.*, 1996; Heal and Small, 2002), and enhancing the resilience of agricultural production (Fischer *et al.*, 2006). Ecosystem services that agriculture depends on include nutrient cycling, soil stabilization, pest suppression, and pollination (Kremen *et al.*, 2002; Edwards and Arancon, 2004; Gurr *et al.*, 2004), among others. Moreover, Walk *et al.* (2010) showed that even very small patches of more compositionally and structurally heterogeneous vegetation cover can provide habitat for grassland birds.

Conclusions

Whether bioenergy will help or harm is in the hands of policy makers, land managers, and ultimately the global citizenry. The documented and anticipated changes in wildlife habitat quality, species interactions, and populations suggest the future will be perilous if nothing is done to combat, mitigate, and adapt to climate change (Dawson *et al.*, 2011). Society clearly needs to strive for enhanced energy conservation and the development of renewable energy sources. Although it is widely accepted that bioenergy can and should be a part of the renewable energy portfolio, several important questions regarding its potential contribution remain.

First, how much can bioenergy contribute to the renewable energy portfolio? The science suggests that society needs to lower its expectations with regard to the magnitude of bioenergy's contribution. Factors to be considered in assessments of potential include the overall productivity of the Earth's ecosystems (net primary production), competing demands for biomass (i.e., food, fiber, and habitat), and the efficiency of converting the diffuse calories stored in biomass into concentrated, transportable, and usable energy. Sharlemann (2008) estimated that four times as much land area as is currently used to grow food would be required for fuel crops if all energy for human use was provided through bioenergy. Analysis by Campbell *et al.* (2008) suggested that bioenergy derived from abandoned agricultural lands could maximally offset 8% of current global energy demand (Campbell *et al.*, 2008).

Secondly, how should bioenergy crops be configured on the landscape? The current trajectory of utilizing HILD systems, especially those based on food crops, may not be sustainable at present or over the long term. The land-use changes currently being undertaken to support HILD-based bioenergy have clear negative impacts on wildlife, both through direct habitat degradation or loss, and indirectly by exacerbating climate change (Martinelli and Filoso, 2008; Koh *et al.*, 2011; Searchinger *et al.*, 2008). The high levels of inputs required to maintain these systems make them less feasible in a resource-limited world (Tilman, 1999; Schiesari and Grillitsch, 2011). Bioenergy based on lignocellulosic materials, especially biomass systems that can be grown with low levels of agricultural inputs, offer more sustainable solutions (Robertson *et al.*, 2008; Tilman *et al.*, 2009; Sommerville *et al.*, 2010). However, implementing these systems requires major technical, economic, cultural, social, and political shifts, including:

- Technical – the development of efficient, scalable methods for converting diverse biomass feedstocks, especially lignocellulosic feedstocks, into concentrated bioenergy (Sommerville *et al.*, 2010; Richard, 2010).
- Economic – the adoption of methods of full cost accounting, which explicitly consider the environmental impacts associated with different forms of energy, including habitat loss and greenhouse-gas emissions (Pretty *et al.*, 2001; Ruhl *et al.*, 2007).
- Cultural – a change in cultural norms such that farmers are compelled to manage for the coproduction of environmental benefits in addition to food and fuel (NRC, 2010; Atwell *et al.*, 2011).
- Social – the redesign of institutions to foster collaboration in meeting targets for producing food, fuel, and environmental benefits over local, landscape, regional, and global scales (NRC, 2010; IEA, 2011).
- Political – an overhaul of existing policies and development of new ones to facilitate the shifts above (Ruhl *et al.*, 2007; NRC, 2010).

In conclusion, despite the paucity of research on wildlife response to bioenergy development (Dauber *et al.*, 2010), the conversion of native systems to bioenergy production is clearly harmful. In contrast, the replacement of HILD crops on non-prime agricultural lands with high-diversity systems could be helpful. Otherwise, whether bioenergy will help or harm wildlife will vary substantially depending on the species under consideration, the land-use bioenergy production is replacing, the bioenergy crops being deployed, and the landscape to regional context the bioenergy crops are grown within. The concepts summarized here provide some principles that could help provide renewable bioenergy while sustaining wildlife over local to global scales, and fostering the substantial technical, economic, cultural, social, and political shifts that can make it so.

Appendix

List of Courses

Agriculture and Natural Resources
Biorenewable Energy Systems
Current Topics in Environmental Science
Ecology and Society
Restoration Ecology

See also: Agriculture, Sustainable. Agrobiodiversity. Biofuels and Biodiversity: The Implications of Energy Sprawl. Climate Change and

Wild Species. Ecology of Agriculture. Ecosystem Services. Energy Use, Human. Restoration of Biodiversity, Overview

References

Anderson K, Martin W, and Valenzuela E (2006) The relative importance of global agricultural subsidies and market access. *World Trade Review* 5: 357–376.

Aratrakorn S, Thunhikorn S, and Donald PF (2006) Changes in bird communities following conversion of lowland forest to oil palm and rubber plantations in southern Thailand. *Bird Conservation International* 16: 71–82.

Atwell RC, Schulte LA, and Westphal LM (2011) Tweak, adapt, or transform: Policy scenarios in response to emerging bioenergy markets in the U.S. Corn Belt. *Ecology and Society* 16: 10.

Babcock BA and Fabiosa JF (2011) The impact of ethanol and ethanol subsidies on corn prices: Revisiting history. *CARD Policy Brief 11-PB 5*. Ames, IA: Center for Agricultural and Rural Development, Iowa State University.

Bekunda M, Palm CA, de Fraiture C, *et al.* (2009) Biofuels in developing countries. In: Howarth RW and Bringezu S (eds.) *Biofuels: Environmental Consequences and Interactions with Changing Land Use. Proceedings of the Scientific Committee on Problems of the Environment (SCOPE) International Biofuels Project Rapid Assessment*, pp. 249–269. Ithaca, NY: Cornell University. Available at: http://cip.cornell.edu/biofuels/ (accessed 13 May 2011).

Bennett EL (2002) Is there a link between wild meat and food security? *Conservation Biology* 16: 590–592.

Benton TG, Vickery JA, and Wilson JD (2003) Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology and Evolution* 18: 182–188.

Best LB (1986) Conservation tillage: Ecological traps for nesting birds. *Wildlife Society Bulletin* 14: 308–317.

Best LB, Campa H, Kemp KE, *et al.* (1997) Bird abundance and nesting in CRP fields and cropland in the Midwest: a regional approach. *Wildlife Society Bulletin* 25: 864–877.

Bies L (2006) The biofuels explosion: Is green energy good for wildlife? *Wildlife Society Bulletin* 34: 1203–1205.

Blanco-Canqui H and Lal R (2009) Crop residue removal impacts on soil productivity and environmental quality. *Critical Reviews in Plant Sciences* 28: 139–163.

Bradshaw CJA, Sodhi NS, and Brook BW (2009) Tropical turmoil: A biodiversity tragedy in progress. *Frontiers in Ecology and the Environment* 7: 79–87.

Campbell JE, Lobell DB, Genova RC, and Field CB (2008) The global potential of bioenergy on abandoned agriculture lands. *Environmental Science and Technology* 42: 5791–5794.

Carpenter SR, Caraco NF, Correll DL, *et al.* (1998) Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications* 8: 559–568.

Commission of the European Communities (2008) *Proposal for a Directive of the European Parliament and of the Council on the promotion of the use of energy from renewable sources*, COM (2008) 19 final. Brussels, Belgium: Commission of the European Communities.

Coutla L, Dyer N, and Vermeulen S (2008) *Fuelling Exclusion? The Biofuels Boom and Poor People's Access to Land*. London: IIED.

Dale VH, Kline KL, Wiens J, and Fargione J (2010) Biofuels: Implications for land use and biodiversity. *Biofuels and Sustainability Reports*. Washington, DC: Ecological Society of America.

Dauber J, Jones MB, and Stout JC (2010) The impact of biomass crop cultivation on temperate biodiversity. *Global Change Biology Bioenergy* 2: 289–309.

Davis SC, Dohleman FG, and Long SP (2011) The global potential for *Agave* as a biofuel feedstock. *Global Change Biology Bioenergy* 3: 68–78.

Dawson TP, Jackson ST, House JI, Prentice IC, and Mace GM (2011) Beyond predictions: Biodiversity conservation in a changing climate. *Science* 332: 53–58.

Edwards CA and Arancon NQ (2004) Interactions among organic matter, earthworms, and microorganisms in promoting plant growth. In: Magdoff F and Weil RR (eds.) *Soil Organic Matter in Sustainable Agriculture*, 1st edn., pp. 327–376. Boca Raton, FL: CRC Press.

Edwards DP, Hodgson JA, Hamer KC, *et al.* (2010) Wildlife-friendly oil palm plantations fail to protect biodiversity effectively. *Conservation Letters* 3: 236–242.

Eggers J, Tröltzsch K, Falcucci A, *et al.* (2009) Is biofuel policy harming biodiversity in Europe. *Global Change Biology Bioenergy* 1: 18–34.

Ellis EC and Ramankutty N (2008) Putting people in the map: Antropogenic biomes of the world. *Frontiers in Ecology and the Environment* 6: 439–447.

Environmental Protection Agency (EPA) (1994) *The Quality of Our Nation's Water: 1992. EPA 841-S-94-002*. Washington, DC: EPA Office of Water.

Evans AM, Perschel RT, and Kittler BA (2010) *Revised Assessment of Biomass Harvesting and Retention Guidelines*. Santa Fe, NM: Forest Guild. Available online at: http://www.forestguild.org/publications/research/2009/biomass_guidelines.pdf (accessed 2 May 2011).

FAO (Food and Agricultural Organization of the United Nations) (2009) *The State of Food Insecurity in the World 2009: Economic Crises – Impacts and Lessons Learned*. Rome: FAO.

FAO (Food and Agricultural Organization of the United Nations) (2010) *The State of Food Insecurity in the World 2010: Addressing Food Security in Protracted Crises*. Rome: FAO.

Fargione JE, Cooper TR, Flaspohler DJ, *et al.* (2009) Bioenergy and wildlife: Threats and opportunities for grassland conservation. *BioScience* 59: 767–777.

Fargione JE, Hill J, Tilman D, Polasky S, and Hawthorne P (2008) Land clearing and the biofuel carbon debt. *Science* 319: 1235–1238.

Field CB, Campbell JE, and Lobell DB (2008) Biomass energy: The scale of the potential resource. *Trends in Ecology and Evolution* 23: 65–72.

Fischer J, Lindenmayer DB, and Manning AD (2006) Biodiversity, ecosystem function, and resilience: Ten guiding principles for commodity production landscapes. *Frontiers in Ecology and the Environment* 4: 80–86.

Fletcher Jr. RJ, Robertson BA, Evans J, *et al.* (2011) Biodiversity conservation in the era of biofuels: Risks and opportunities. *Frontiers in Ecology and the Environment* 9: 161–168.

Fuller RJ, Gregory RD, Gibbons DW, *et al.* (1995) Population declines and range contractions among lowland farmland birds in Britain. *Conservation Biology* 9: 1425–1441.

Gardiner MA, Tuell JK, Isaacs R, Gibbs J, Ascher JS, and Landis DA (2010) Implications of three biofuel crops for beneficial arthropods in agricultural landscapes. *Bioenergy Research* 3: 6–19.

Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–555.

Groves CR (2003) *Drafting a Conservation Blueprint: A Practitioner's Guide to Planning for Biodiversity*. Washington, DC: Island Press.

Gurr GM, Wratten SD, and Altieri MA (eds.) (2004) *Ecological Engineering for Pest Management: Advances in Habitat Manipulation for Arthropods*. Victoria: CSIRO.

Hanowski JM, Niemi GJ, and Christian DC (1997) Influence of within-plantation heterogeneity and surrounding landscape composition on avian communities in hybrid poplar plantations. *Conservation Biology* 11: 936–944.

Hansen MC, Stehman SV, Patopov PV, *et al.* (2008) Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9439–9444.

Headey D and Fan S (2008) Anatomy of a crisis: The causes and consequences of surging food prices. *Agricultural Economics* 39: 375–391.

Heal GM and Small AA (2002) Agriculture and ecosystem services. In: Gardner BL and Rausser GC (eds.) *Handbook of Agricultural Economics*, 1st edn, vol. 2a, pp. 1341–1369. Amsterdam: Elsevier.

Hickman GC, Vanloocke A, Dohleman FG, *et al.* (2010) A comparison of canopy evapotranspiration for maize and two perennial grasses identified as potential bioenergy crops. *Global Change Biology Bioenergy* 2: 157–168.

Higginbotham S, Noble L, and Joice R (1996) The profitability of integrated crop management, organic and conventional arable regimes. Rotations and cropping systems. *Aspects of Applied Biology* 47: 327–333.

Hill J, Nelson E, Tilman D, Polasky S, and Tiffany D (2007) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11206–11210.

Hinsley SA and Bellamy PE (2000) The influence of hedge structure, management and landscape context on the value of hedgerows to birds: A review. *Journal of Environmental Management* 60: 33–49.

Hoogwijk M, Faaig A, van den Broek R, *et al.* (2003) Exploration of the ranges of the global potential of biomass for bioenergy. *Biomass and Bioenergy* 25: 119–133.

IEA (International Energy Agency) (2011) *Technology Roadmap: Biofuels for Transport*. Paris, France: Organization for Economic Cooperation and Development and the International Energy Agency.

Jackson D and Jackson LL (eds.) (2002) *The Farm as Natural Habitat: Reconnecting Food Systems with Ecosystems*. Washington, DC: Island Press.

Kim S and Dale BE (2005) Life cycle assessment of various cropping systems utilized for producing biofuels: Bioethanol and biodiesel. *Biomass and Bioenergy* 29: 426–439.
Koh LP (2007) Potential habitat and biodiversity losses from intensified biodiesel feedstock production. *Conservation Biology* 21: 1373–1375.
Koh LP and Ghazoul J (2010) Spatially explicit scenario analysis for reconciling agricultural expansion, forest protection, and carbon conservation in Indonesia. *Proceedings of the National Academy of Sciences of the United States of America* 107: 11140–11144.
Koh LP, Miettinen J, Liew SC, and Ghazoul J (2011) Remotely sensed evidence of tropical peatland conversion to oil palm. *Proceedings of the National Academy of Sciences of the United States of America* 108: 5127–5132.
Koh LP and Wilcove DS (2008) Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letters* 1: 60–64.
Kremen C, Williams NM, and Thorp RW (2002) Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences of the United States of America* 99: 16812–16816.
Landis DA and Werling BP (2010) Arthropods and biofuel production systems in North America. *Insect Science* 17: 220–236.
Lapola DM, Schaldach R, Alcamo J, *et al.* (2010) Indirect land use changes can overcome carbon savings from biofuels in Brazil. *Proceedings of the National Academy of Sciences of the United States of America* 107: 3388–3393.
Liebman M, Helmers MJ, and Schulte LA (2011) Integrating conservation with biofuel feedstock production. In: Nowak P and Schnepf M (eds.) *Managing Agricultural Landscapes for Environmental Quality II: Achieving More Effective Conservation*, pp. 131–142. Ankeny, IA: Soil and Water Conservation Society.
Martin TG and Catterall CP (2001) Do fragmented coastal heathlands have habitat value to birds in eastern Australia? *Wildlife Research* 28: 17–31.
Martinelli LA and Filoso S (2008) Expansion of sugarcane ethanol production in Brazil: Environmental and social challenges. *Ecological Applications* 18: 885–898.
Matson PA, Parton WJ, Power AG, and Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 277: 504–509.
Millennium Ecosystem Assessment (MEA) (2005) Living Beyond our Means: Natural Assets and Human Well-Being. Available online at: http://www.maweb.org/documents/document.429.aspx.pdf (accessed 20 April 2011).
Milner-Gulland EJ and Bennett EL (2003) Wild meat: The bigger picture. *Trends in Ecology and Evolution* 18: 351–357.
Mitchell D (2008) A note on rising food process. *Policy working paper 4682*. Washington, DC: The World Bank Development Prospects Group. Available online at: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2008/07/28/000020439_20080728103002/Rendered/PDF/WP4682.pdf (accessed 12 May 2011).
Monteiro JM, Elebras Veiga LB and Coutinho HLC (2010) Sugarcane crop for biofuel production, demand on soil resource and food security in Brazil. *19th World Congress of Soil Science: Soil Solutions for a Changing World*. Brisbane, Australia: International Union of Soil Sciences.
Murray LD, Best LB, Jacobsen TJ, and Braster ML (2003) Potential effects on grassland birds of converting marginal cropland to switchgrass biomass production. *Biomass and Bioenergy* 25: 167–175.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
Naylor RL, Liska AJ, Burke MB, *et al.* (2007) The ripple effect: Biofuels, food security, and the environment. *Environment* 49: 30–43.
Nelson RG (2002) Resource assessment and removal analysis for corn stover and wheat straw in the Eastern and Midwestern United States – Rainfall and wind-induced soil erosion methodology. *Biomass and Bioenergy* 22: 349–363.
Noss RF and Cooperrider AY (1994) *Saving Nature's Legacy: Protecting and Restoring Biodiversity*. Washington, DC: Island Press.
NRC (National Research Council) (2010) *Toward Sustainable Agricultural Systems in the 21st Century*. Washington, DC: The National Academies.
Öckinger E and Smith HG (2007) Semi-natural grasslands as population sources for pollinating insects in agricultural landscapes. *Journal of Applied Ecology* 44: 0–59.
Oldeman LR, Hakkeling RTA, and Sombroek WG (1991) *World Map of the Status of Human-Induced Soil Degradation*, 2nd edn. Wageningen: ISRIC.
Perlack RD, Wright LL, Turhollow AF, *et al* (2005) Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply. *GO-102005-2135*. Oak Ridge, TN: U.S. Department of Energy and U.S. Department of Agriculture.
Petit LJ, Petit DR, Christian DG, and Powell HDW (1999) Bird communities of natural and modified habitats in Panama. *Ecography* 22: 292–304.
Power S and de Araujo LM (1993) Nature conservation with reference to the state of Alagoas, Brazil. *Environmentalist* 13: 297–302.
Pretty J, Brett C, Gee D, *et al.* (2001) Policy challenges and priorities for internalizing the externalities of modern agriculture. *Journal of Environmental Planning and Management* 44: 263–283.
Raghu S, Anderson RC, Daehler CC, *et al.* (2006) Adding biofuels to the invasive species fire? *Science* 313: 1742.
Rathke G-W, Wienhold BJ, Wilhelm WW, and Diepenbrock W (2007) Tillage and rotation effect on corn-soybean energy balances in eastern Nebraska. *Soil and Tillage Research* 97: 60–70.
Relyea RA (2005) The impact of insecticides and herbicides on the biodiversity and productivity of aquatic communities. *Ecological Applications* 15: 618–627.
Reynolds JF and Stafford Smith DM (2002) Global desertification: Do humans cause deserts? *Dahlem Workshop Report 88*. Berlin: Dahlem University Press.
Richard TL (2010) Challenges in scaling up biofuels infrastructure. *Science* 329: 793–796.
Richter BD, Braun DP, Mendelson MA, and Master LL (1997) Threats to imperiled freshwater fauna. *Conservation Biology* 11: 1081–1093.
Riffell S, Verschuyl J, Miller D, and Wigley TB (2011) A meta-analysis of bird and mammal response to short-rotation woody crops. *Global Change Biology Bioenergy* 3: 313–321.
Robertson BA, Doran PJ, Loomis ER, Robertson JR, and Schemske DW (2011) Avian use of perennial biomass feedstocks as post-breeding and migratory stopover habitat. *PLoS ONE* 6: e16941.
Robertson GP, Dale VH, Doering OC, *et al.* (2008) Sustainable biofuels redux. *Science* 322: 49–50.
Ruhl JB, Kraft S, and Lant C (2007) *The Law and Policy of Ecosystem Services*. Washington, DC: Island Press.
Sala OE, Chapin FS, Armesto JJ, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.
Scharlemann JPW and Laurance WF (2008) How green are biofuels? *Science* 319: 43–44.
Schiesari L and Grillitsch B (2011) Pesticides meet megadiversity in the expansion of biofuel crops. *Frontiers in Ecology and the Environment* 9: 215–221.
Schoneveld G, German L, Andrade R, *et al* (2010) The role of national governance systems in biofuel development: A comparative analysis of lessons learned. *CIFOR infobriefs* 35. Available at: http://www.cifor.cgiar.org/publications/pdf_files/infobrief/3308-infobrief.pdf/ (accessed 12 May 2011).
Schulze CH, Waltert M, Kessler PJA, *et al.* (2004) Biodiversity indicator groups of tropical land-use systems: comparing plants, birds, and insects. *Ecological Applications* 14: 1321–1333.
Searchinger T, Heimlich R, Houghton RA, *et al.* (2008) Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land use change. *Science* 319: 1238–1240.
Sharlemann JPW (2008) Can bird research clarify the biodiversity benefits and drawbacks of biofuels? *Ibis* 150: 640–642.
Silveira LF, Olmos F, and Long AJ (2003) Birds in Atlantic Forest fragments in north-east Brazil. *Cotinga* 20: 32–46.
Somerville C, Youngs H, Taylor C, Davis SC, and Long SP (2010) Feedstocks for lignocellulosic biofuels. *Science* 329: 790–792.
Thurow R and Kilman S (2009) *Enough: Why the World's Poor Starve in an Age of Plenty*. New York: PublicAffairs.
Tilman D (1999) Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences of the United States of America* 96: 5995–6000.
Tilman D, Hill J, and Lehman C (2006) Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science* 314: 1598–1600.
Tilman D, Socolow R, Foley JA, *et al.* (2009) Beneficial biofuels – The food, energy, and environment trilemma. *Science* 325: 270–271.
United States Department of Agriculture Economic Research Service (USDA ERS) (2011) *Feed grains database. Updated September*. Washington, DC: United States Department of Agriculture. Available at: http://www.ers.usda.gov/Data/FeedGrains/ (accessed 3 May 2011).
Walk JW, Kershner EL, Benson TJ, and Warner RE (2010) Nesting success of grassland birds in small patches in an agricultural landscape. *Auk* 127: 328–334.
Wildlife Habitat Management Institute (WHMI) (1999) Conservation tillage and terrestrial wildlife. USDA Natural Resources Conservation Service. *Fish and Wildlife Literature Review* 1: 1–10.
Zah R, Böni H, Gauch M, *et al.* (2007) *Ökobilanz von energieprodukten: ökologische bewertung von biotreibstoffen*. St. Gallen, Switzerland: Empa.

Biogeochemical Cycles

Paul G Falkowski, Rutgers University, New Brunswick, NJ, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 437–453, © 2001, Elsevier Inc.

Glossary

Anion Negatively charged atom or molecule.
Cation Positively charged atom or molecule.
Chemoautotroph Capable of growing on inorganic substrates without light energy.
Eolian Wind blown.
Felsic rock Continental crustal rocks relatively rich in silicon and aluminum.
Fixation To make nonvolatile (forming organic molecules from inorganic gases).
Heterotroph Growth dependent on the oxidation of organic matter.
Mafic rock Oceanic crustal rocks relatively rich in magnesium and iron.
Orogenesis Formation of mountains.
Photoautotroph Growth dependent on the conversion of light to chemical energy (photosynthetic).
Spallation Thermally induced neutron and proton ejection following high-energy proton collision.

Introduction

Like life itself, biogeochemical cycles are far from thermodynamic equilibrium, have evolved over hundreds of millions of years, and are interdependent, forming biogeochemical systems replete with feedback controls (Schlesinger, 1997). Biogeochemical cycles depend on, and co-evolved with, specific metabolic pathways. Hence, biogeochemical cycles depend on and are a selective force in metabolic (i.e., biological) diversity. Over geological time, biogeochemical cycles are responsible for altering the chemistry of the ocean, atmosphere, and terrestrial ecosystems such that rate-limiting reactions within key cycles modify the tempo and mode of evolution. Here, we examine some of the key biogeochemical cycles in the context of their evolution and biological diversity.

The Distribution of Elements on Earth

The Origin of the Elements

The relative abundance of the crustal elements reveals a distribution of much more abundant light elements to lesser abundant heavier elements (**Figure 1**). Although the two lightest elements, H and He, were formed approximately 16 billion years ago in the "big bang," all the heavier elements result from fusion of ^{2}He nuclei or fusion/spallation (proton or neutron loss) reactions in stars (Degens, 1989). The fusion reactions involving ^{2}He with itself, H, O, C, or N tend to form even-numbered atomic nuclei, whereas spallation and proton capture leads to odd-numbered nuclei. Additionally, because the nuclei of elements with paired protons are slightly more stable than those with an odd number, there is generally a larger relative abundance of even-numbered elements.

In the origin of our solar system approximately 4.6 billion years ago, elemental composition and planetary accretion were strongly influenced by the gravitational forces of the sun. Planet bodies closer to the sun contain relatively higher proportions of heavy elements than bodies further away. The four innermost planets are approximately three times denser than the outer planets and have solid-rock surfaces that contain a relatively high proportion of metals, especially iron and aluminum. In the accretion process, a further gravitational distillation occurred within the planets. On Earth, the abundant heavier elements, such as nickel and iron, tended to migrate toward the center of the internal gravitational field, whereas lighter elements tended to float above the metal core and, upon cooling, accumulated as a solid surface and crust. The lightest elements formed a gaseous phase. Almost all of the two lightest gases, H_2 and He, escaped the gravitational field and diffused into interplanetary space during this initial period of Earth's history.

The composition of the gases in Earth's atmosphere following accretion is not completely resolved but almost certainly contained high concentrations of CO_2, N_2, and H_2O, HCl, and H_2SO_4. This gas composition is similar to that currently on Venus. Precipitation of minerals and formation of felsic rocks led to the condensation and upward migration of liquid, precipitable water that overlies vast regions of denser, mafic rocks. Additional water may have been provided by meteoritic bombardment. Based on thermodynamic equilibrium calculations with crustal elements, the acidic gases or hydrated equivalents (e.g., HCl and H_2SO_4) would have solubilized mineral cations in the primordial ocean leading to a seawater dominated by Na^+, K^+, Mg^{2+}, and Ca^{2+}. The anion balance would have been supplied by vulcanism and outgassing from deep crustal sources. The dominant anions were Cl^- and, to a lesser extent, SO_4^{2-} and HCO_3^-. PO_4^{2-} was probably present but to a lesser extent, and fixed inorganic N (as NH_3 or NO_3) was almost certainly very scarce. The pH of the early oceans was probably close to neutral or slightly acidic (6.8–7.0). Additionally, radiogenically produced heat, the high concentrations of greenhouse gases, and geothermal activity would have provided a source of heat to maintain elevated temperatures in Earth's early ocean history.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00013-7

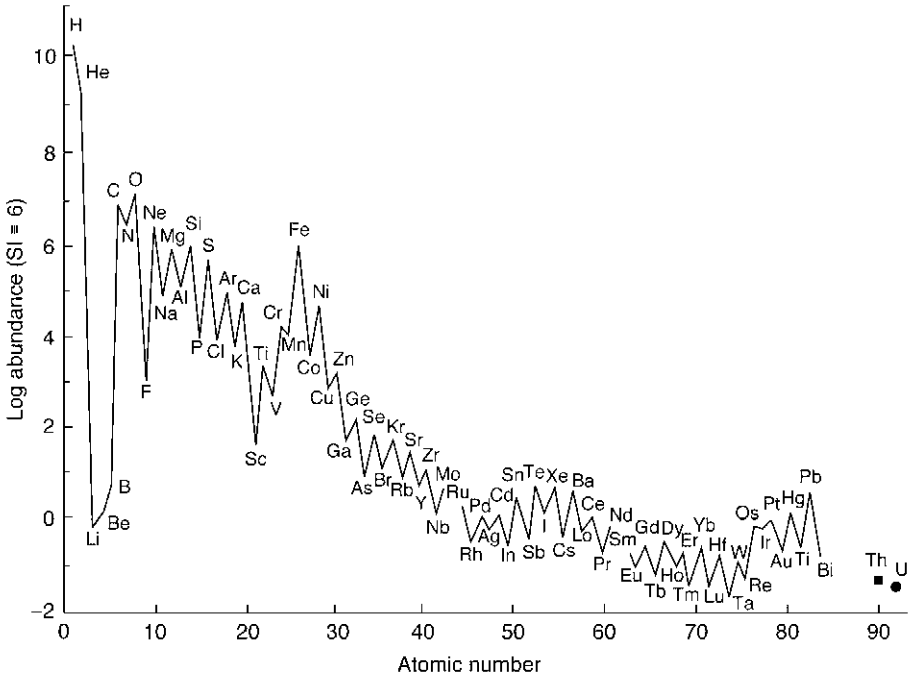


Figure 1 Relative abundances of the elements in the universe (based on log abundance of Si = 6). •, even atomic numbers; ○, odd atomic numbers. Reproduced with permission from Williams R and Frausto da Silva J (1996) *The Natural Selection of the Chemical Elements*. Oxford: Clarendon, with permission from Oxford University Press.

From these conditions, life arose and evolved. Over the course of Earth's history, there has been a continuous trajectory from the mildly reducing conditions that prevailed at the time of origin of the planet to the highly oxidizing conditions that prevail in the contemporary geological epoch. This oxidation trajectory is driven by biological processes but has also led to the increased diversity of life forms.

Phase State Transitions and Elemental Partitioning

The distribution of elements between the atmosphere, lithosphere, and hydrosphere depends, to first order, on the phase state of the element, its chemical reactivity, and its partitioning coefficients between water, the atmosphere, and the lithosphere (**Table 1**). Gases such as N_2 and O_2 partition between the atmosphere and liquid phase according to the solubility coefficient. CO_2, H_2S, and NO_2 undergo hydration to form carbonic, sulfuric, and nitric acids. These acids deprotonate, according to the pH of the aqueous environment, forming anions. Consequently, the total solubility of these species in the aqueous phase exceeds that predicted from the vapor pressure alone, assuming the gases are chemically inert and obey idealized gas laws. Similarly, the distribution of elements between the aqueous phase and the lithosphere is dictated, to first order, by the solubility coefficient of the element, its reactivity or ionic form, and the saturation level. For example, all the alkali and alkali earth metals (i.e., Li, Na, K, Cs, Ca, and Mg) are highly reactive with water and are highly soluble. Precipitation reactions of alkali earth metal cations with carbonate, sulfate, or phosphate anions result in mineral deposits in the solid phase. The movement of elements between the lithosphere, atmosphere, and hydrosphere is mediated by both purely physical/chemical processes and biological processes.

Physical/Chemical Transport Processes: Weathering

The flux of elements from continental rock sources to the atmosphere and oceans is promoted by weathering, which is the physical and chemical degradation of rocks to smaller physical pieces and soluble chemical constituents. The physical processes that promote weathering include erosion by water and wind, cracking through freeze–thaw cycles, and active abrasion by the movement of ice (e.g., glaciers), other rock formations (e.g., earthquakes and tectonic uplifting), or rocks (e.g., tidal abrasion). Ultimately, these processes lead to the formation of sediments that are transported by rivers to the ocean. In addition to physical weathering, chemical weathering is promoted by weak acids (e.g., carbonic and sulfuric) in precipitation as well as microbial and plant growth. Over geological time, the weak acids in precipitation remove alkaline cations, such as Mg and Ca, from carbonates and silicates, leading to the mobilization of these elements. Moreover, the production of weak organic acids by lichens, terrestrial plant roots, and the microbial degradation of decaying organic matter greatly promote chemical weathering. Thus, the invasion of land by terrestrial plants led to increased mobility of numerous elements and enhanced the fluxes of materials to the oceans. This results in a feedback between biogeochemical processes on land and in the ocean.

Table 1 Major chemical constituents of the Earth's Crust, Sediments, Ocean Water, and Atmosphere[a]

Element	*Crystal ionic charge and radius*		*Continental crust*		*Oceanic crust*		*Average sediments*		*Ocean water*		*Atmosphere*	
		r(Å)	*(wt%)*	*(Vol.%)*	*(wt%)*	*(Vol.%)*	*(wt%)*	*(Vol.%)*	*(wt%)*	*(Vol.%)*	*(wt%)*	*(mol%)*
O	−2	1.32	46.40	93.04	43.80	92.57	47.61	91.32	86.0	99.0 as H_2O	23.15	20.95 (O_2)
Si	+4	0.42	28.15	1.04	24.00	0.93	24.40	0.86				
Al	+3	0.51	8.23	0.56	8.76	0.63	6.03	0.40				
Fe	+3	0.64	5.63	0.46	8.56	0.74	3.79	0.30				
	+2	0.74										
Ca	+2	0.99	4.15	1.40	6.72	2.39	7.86	2.54	0.04	0.025		
Na	+1	0.97	2.36	1.31	1.94	1.13	1.36	0.72	1.08	0.11		
Mg	+2	0.66	2.33	0.38	4.5	0.78	2.44	0.39	0.13	0.04		
K	+1	1.33	2.09	1.75	0.83	0.73	2.00	1.61	0.04	0.062		
Ti	+4	0.68	0.54	0.05	0.90	0.09						
Mn			0.095		0.15							
H			0.14		0.2				10.7	See O		
P	+5	0.35	0.105		0.14		0.16	0.003				
S	+6	0.30	0.026		0.025		0.62	0.007	0.09	0.0002		
C	+4	0.16					2.91	0.013	0.28	0.002	0.046	0.03 (CO_2)
Cl	−1	1.81					0.83	1.85	1.94	0.833		
N											75.53	78.09 (N_2)
Ar											1.28	0.93 (Ar)

[a]Adapted from Lerman A (1979) *Geochemical Processes Water and Sediment Environments*. New York: John Wiley & Sons.

The Hydrological Cycle

Weathering is strongly dependent on the hydrological cycle, which is one of the earliest and most critical cycles on Earth. Water is both a solvent and a vehicle for the transport of elements. It is also critically necessary for life. Hence, the distribution of biota is critically related to the availability of water. By far, the oceans are the largest reservoir of water on Earth (**Figure 2**). Flux of water into and from the oceans is driven by evaporation and precipitation, processes that are in turn driven by the heat budget of the planet. Over the oceans, evaporation exceeds precipitation; the net difference leads to precipitation over land. Because precipitation exceeds evaporation over land, there is a flux of water from land to the oceans in rivers. The atmospheric lifetime of water is extremely short—on the order of ca. 10 days. This lifetime compares with the average for the oceans of ca. 3000 years, whereas deep ground water can have lifetimes exceeding 10,000 years. These differences in lifetimes are critical determinants of fluxes of elements and climate feedbacks.

Vulcanism and Orogenesis

The riverine and eolian (wind-blown) transport of sediments from land to the oceans would, over geological time, erode the continents to the peneplane—the lowest level achievable. Vulcanism and other orogenic processes counterbalance erosion and supply new crustal rock and atmospheric sources of minerals. In the subduction regions of converging tectonic plates, sedimentary rocks are injected back into the deeper crust, where they are reprocessed by heat and pressure, emerging again in uplifting regions and through volcanic eruptions. In addition to the formation of igneous rock, vulcanism provides significant sources of iron, sulfate, and CO_2 to the atmosphere. These elements and materials are transported through the atmosphere and deposited over both the oceans and on land.

Biological Transport

Biological processes influence the distribution of elements primarily through phase state transitions. The chemical reduction of inorganic carbon to organic matter is a conversion of a gas to solid phase; the latter is often dissolved in an aqueous phase of the cell cytoplasm. For oxygenic photoautotrophs, the evolution of molecular oxygen leads to the formation of a gas from the liquid phase of water. Biological mediation of phase state transitions is critical for the movement of elements between the atmosphere, lithosphere, and hydrosphere. For example, $CaCO_3$ is by far the largest reservoir of carbon on Earth (**Table 2**). The formation of $CaCO_3$ is the consequence of the biological precipitation by marine organisms and, over geological time, represents an important sink for atmospheric CO_2. The oxidation of organic matter leads to the formation of CO_2, which equilibrates with the atmosphere. Similarly, the biologically catalyzed oxidation of H_2S to S or SO_4 is a phase state transition that influences the partitioning of sulfur between aquatic ecosystems and the atmosphere. It should be noted that the biologically catalyzed phase state transitions need not be direct but can result from the indirect modification of the redox state of the environment. For example, the oxidation of H_2S occurs spontaneously in the presence of O_2; however, the O_2 is formed from oxygenic photoautotrophs.

Biological Assimilation

To accomplish phase state transitions, biological processes catalyze chemical and/or physical reactions. Reaction sequences require an assimilation of the elements. The

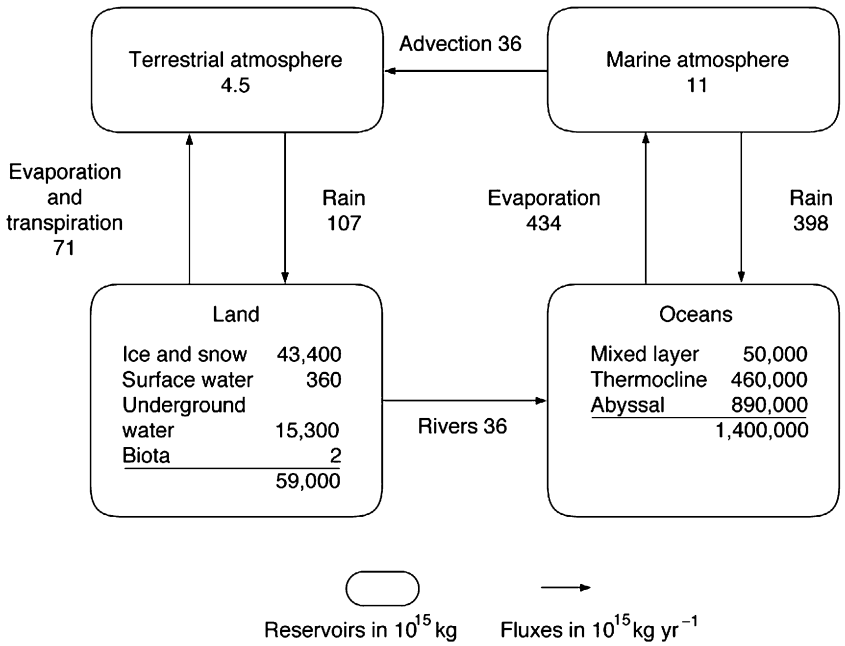


Figure 2 Estimates of the global water cycle and its reserves. The accuracy of several of its components is poor, resulting in a closure error for the whole cycle of about a factor of 2. The obvious interactive nature of the cycle makes it impossible to reduce current closure errors without studying the whole cycle. Reprinted from Chahine M (1992) The hydrological cycle and its influence on climate. *Nature* 359: 373–380.

Table 2 Carbon pools in the major reservoirs on Earth[a]

Pools	*Quantity ($\times 10^{15}$ g)*
Atmosphere	720
Oceans	38,400
Total inorganic	37,400
Surface layer	670
Deep layer	36,730
Total organic	1,000
Lithosphere	
Sedimentary carbonates	>60,000,000
Kerogens	15,000,000
Terrestrial biosphere (total)	2,000
Living biomass	600–1,000
Dead biomass	1,200
Aquatic biosphere	1–2
Fossil fuels	4,130
Coal	3,510
Oil	230
Gas	140
Other (peat)	250

[a]Reproduced from Falkowski P (1997) Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* 387: 272–275, with permission from Nature.

elemental composition of organisms is dictated, to first order, by the biochemical machinery required for maintenance, growth, mobility, reproduction, and defenses against predation. The four fundamental biochemical constituents that comprise all organisms, namely, proteins, nucleic acids, lipids, and carbohydrates, are primarily synthesized from six light elements: H, C, N, O, P, and S. Of these, N, C, and O are generally reduced, P is oxidized, and S can be both reduced and oxidized. These six elements comprise approximately 98% of all living biomass. The remainder consists of approximately 20 ion-forming elements, including Na, Mg, Si, Cl, Ca, V, C, Mn, Fe, Co, Ni, Cu, Zn, Se, Mo, Sn, and I. The transition metals are frequently used in the mediation of electron transfer (redox) reactions.

Key Biogeochemical Reactions and Cycles

Hydrolysis and Oxidation-Reduction Reactions

Two biologically mediated chemical reactions are key to understanding biogeochemical cycles, namely, hydrolysis and oxidation-reduction processes. Hydrolytic reactions chemically remove from or add water to specific elements or molecules; the reactions are almost always accompanied by the transfer of protons and thus are pH-dependent processes. Hydrolytic reactions can be described in a general form by

$$AH^+ + OH^- \rightarrow A + H_2O$$

$$AOH^- + H^+ \rightarrow A + H_2O$$

Hydration/dehydration reactions can thermodynamically store or discharge energy and determine the reactivity and solubility of elements in the first series (H, Li, Na, etc.) and of phosphorous, (inorganic) carbon, and silica in the environment.

Redox reactions are characterized by the addition or loss of electrons, hydrogen atoms (but not simply a proton), or molecular oxygen. Redox reactions are the primary engines of life, and the consequences of redox chemistry strongly impact

the distributions of oxygen, carbon, nitrogen, sulfur, and transition metals in the environment. Redox reactions are always paired; there must be an oxidant and a reductant. The reaction process can be generalized by half-cells:

$$A_{(ox)} + e^- \rightarrow A_{(red)}$$

$$B_{(red)} \rightarrow B_{(ox)} + e^-$$

where the overall reaction is

$$A_{(ox)} + B_{(red)} \leftrightarrow A_{(red)} + B_{(ox)}$$

For each half-cell, the tendency for a substrate to be oxidized or reduced is described by the Nernst equation:

$$E = (E_O + 2.3\ RT)/(nF \log_{10}[A_{ox}]/[A_{red}])$$

where E is the redox potential (in volts), E_o is an arbitrarily accepted standard redox potential, F is the Faraday's constant ($=96{,}487\ C = 1$ mol electrons), n is the number of moles of electrons (Faraday's) transferred in the half-cell reaction, R is the Boltzmann gas constant, T is temperature in Kelvin, and A_{ox} and A_{red} are the activities (or more commonly the concentrations) of the oxidized and reduced forms of the molecules, respectively. The Nernst equation describes the equilibrium condition for both electronic and ionic processes.

To maintain perpetual redox reactions, reductants consumed in one reaction must be regenerated in another (**Table 3**). Hence, by the very fact that redox reactions are paired, a cycle is potentiated. However, to sustain a cycle of redox chemistry requires an input of energy (e.g., solar or thermal) that can be used to drive the initial condition far from thermodynamic equilibrium. Biological reduction of inorganic carbon is one such process.

Table 3 Thermodynamic sequence for reduction of inorganic substances by hydrogen at pH 7.0 and 25 °C[a]

Reaction	$E_h(V)$	ΔG^b
Reduction of O_2 $O_2 + 4H^+ + 4e^- \leftrightarrow 2H_2O$	0.812	−29.9
Reduction of NO_3^- $2NO_3^- + 6H^+ + 6e^- \leftrightarrow N_2 + 3H_2O$	0.747	−28.4
Reduction of Mn^{4+} to Mn^{2+} $MnO_2 + 4H^+ + 2e^- \leftrightarrow Mn^{2+} + 2H_2O$	0.526	−23.3
Reduction of Fe^{3+} to Fe^{2+} $Fe(OH)_3 + 3H^+ + e^- \leftrightarrow Fe^{2+} + 3H_2O$	−0.047	−10.1
Reduction of SO_4^{2-} to H_2S $SO_4^{2-} + 10H^+ + 8e^- \leftrightarrow H_2S + 4H_2O$	−0.221	−5.9
Reduction of CO_2 to CH_4 $CO_2 + 8H^+ + 8e^- \leftrightarrow CH_4 + 2H_2O$	−0.244	−5.6

[a]Reproduced from Schlesinger (1997) p. 234), calculated from Stumm and Morgan (1981, p. 459).

[b]Kcal mole^{-1} per e^-, assuming coupling to the oxidation reaction 1/4 CH_2O + 1/4 $H_2O \rightarrow$ 1/4 $CO_2 + H^+ + e^-$ and $\Delta G = -RT \ln (K)$.

Fundamental Biogeochemical Processes Related to Carbon

The biogeochemistry of carbon is regulated by both hydrolytic and redox reactions. On geological time-scales (>1 million years), the concentration of CO_2 in the atmosphere is a balance between crustal outgassing primarily through vulcanism which provides a source of CO_2 to the atmosphere, and chemical weathering of silica-rich rocks on the earth's surface. These reactions can be simplified as

$$CaSiO_3 + CO_2 \underset{\text{metamorphism}}{\overset{\text{chemical weathering}}{\rightleftharpoons}} CaCO_3 + SiO_2$$

The weathering process is stimulated by the hydration of CO_2 in the atmosphere in rain, forming carbonic acid. This is a self-limiting cycle; as CO_2 rises, weathering accelerates leading to a depletion of the atmospheric inventory, which reduces weathering. In the oceans there is an excess of Ca^{2+}, and an equilibrium in the inorganic carbon system is reached. This reaction series can be specified as

$$H_2O + CO_2 \rightleftarrows H_2CO_3 \rightleftarrows H^+ + HCO_3^- \rightleftarrows 2H^+ + CO_3^{2-}$$

and

$$Ca^{2+} + 2HCO_3^- \rightleftarrows CaCO_{3(\downarrow)} + CO_2 + H_2O$$

Note that the precipitation of $CaCO_3$ leads to the formation of CO_2. In the absence of any biological activity, the geochemical processes would lead to a cycle of inorganic carbon between the three major reservoirs, namely, the atmosphere, oceans, and continents; however, biological processes interact in two critical ways to affect carbon chemistry. First, biological pathways lead to the reduction and oxidation of carbon. Photochemical and chemoautotrophic organisms oxidize alternative substrates to reduce carbon to the equivalent of carbohydrate. These reactions can be broadly classified in the reaction sequence

$$2(H_2A) + CO_2 \rightarrow CH_2O + A_2$$

where A can be sulfur, oxygen, or organic compounds. This is an example of redox reactions applied to the formation of organic material.

Under mildly reducing conditions, photosynthetic carbon fixation is catalyzed by bacteria that couple the oxidation of H_2S or organic compounds to the reduction of CO_2 to form organic compounds. The thermodynamic energy gradient in such systems is relatively low, requiring inputs of ~0.4 eV of photon energy. Such energy demands are satisfied in the infrared spectrum.

Biomineralization and its Biogeochemical Inferences

Many organisms precipitate carbonates, phosphates, or silicates to produce minerals. In the marine environment, the precipitation of $CaCO_3$ by phytoplankton, corals, and several genera of marine invertebrates leads to the formation of aragonite and calcite, two forms of carbonate that are relatively resistant to degradation. Virtually all the carbonate

precipitated on Earth is the consequence of biological mineralization in the ocean. Over geological time, the production of carbonates produced vast beds of chalk, limestones, and marbles. The deposition of carbonates accounts for the largest reservoir of carbon on Earth. The isotopic fractionation of the stable isotopes of oxygen (^{18}O and ^{16}O) in carbonate sediments is strongly influenced by the temperature of the ocean when and where the carbonates are formed. The isotopic fractionation of the stable isotopes of carbon (^{13}C and ^{12}C) is related to the atmospheric/oceanic inventory of CO_2. Hence, isotopic analyses of carbonates provide clues about both oceanic temperatures and carbon dioxide concentrations over geological time (**Figure 3**).

Some groups of marine planktonic organisms, including diatoms and radiolaria, as well as several genera of higher plant grasses, precipitate silica to form hydrated opal. Silica is supplied to the oceans from the weathering of continental rocks and the subsequent flux of soluble silicic acid via rivers. Like carbonates, the precipitation of silicates over geological time leads to sedimentary deposits. In terrestrial ecosystems, several groups of higher plants incorporate silica from soils into stems and shoots.

Under basic conditions, calcium and phosphate ions will spontaneously precipitate to form apatites, which are often mined as a source of phosphate for fertilizers and chemical feedstock. Vertebrates biologically control the precipitation of calcium phosphate to form bones and teeth. Approximately 85% of the phosphate and 95% of the calcium assimilated by vertebrates is incorporated in bones. The fossil record of these biomineral components provides a significant signature of vertebrate evolution but is not a significant component of crustal minerals.

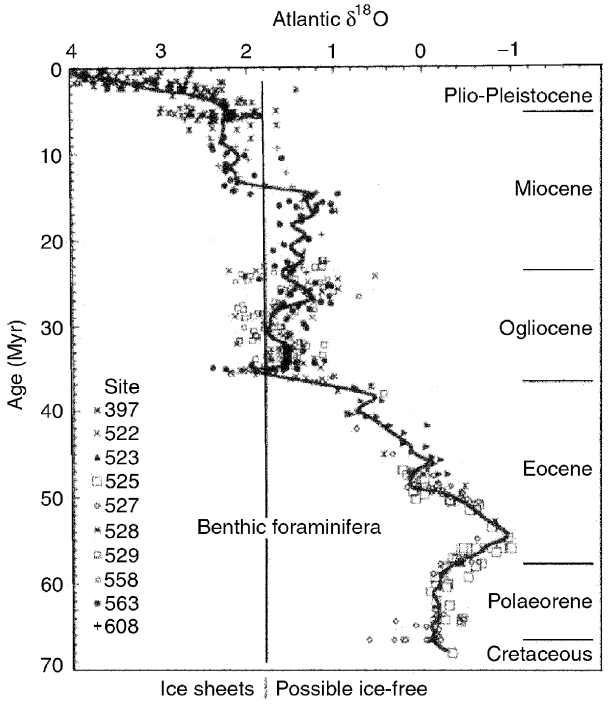


Figure 3 Compilation of benthic $\delta^{18}O$ measurements from Deep Sea Drilling Program sites, spanning the past 70 Ma. The long-term increase in $\delta^{18}O$ values reflects cooling of the deep ocean and growth of ice sheets at high latitudes. $\delta^{18}O = [(^{18}O/^{16}O)_{standard} - 1]$, where standard is PDB. Reprinted from Raymo ME and Ruddiman WF (1992) Tectonic forcing of late Cenozoic climate. *Nature* 359: 117–122.

The Oxygen Cycle

One of the most important characteristics of Earth's atmosphere is the relatively high concentration of free molecular oxygen (O_2). The mere presence of high concentrations of free molecular oxygen suggests that the chemistry of the planet is far from thermodynamic equilibrium. For all practical purposes, all the oxygen in the atmosphere originated from the photosynthetically mediated oxidation of water by oxygenic photoautotrophs. The oxidation of water to form free oxygen requires an oxidant with a potential of at least 0.83 V at pH 7. The bacterial photosynthetic apparatus could not achieve such a high oxidation potential without evolutionary modification. These modifications included the incorporation of Mn as the transient electron acceptor in the photosynthetic reaction center, the alteration of proteins that permitted the oxidation of water, and the raising of the potential of the primary electron donor by ~0.4 V. The latter was achieved by altering the pigment involved in the photosynthetic reaction from the relatively low energy levels that characterize bacteriochlorophylls to the higher energy levels found in chlorophyll *a*, a pigment that distinguishes cyanobacteria (i.e., blue-green algae) from all other photosynthetic bacteria.

The cyanobacteria are the oldest extant oxygenic photoautotrophs. The exact timing of the origin of cyanobacteria is uncertain, but micropaleological evidence suggests that organisms with features remarkably similar to those of modern cyanobacteria existed in the Archean ocean at least 3.5 billion years before present. Based on geochemical evidence (primarily the oxidation of iron by molecular oxygen), it appears that oxidation of Earth's atmosphere occurred over a relatively short interval of ~100 million years, beginning ~2.2. billion years before present (**Figure 4**).

Prior to the photosynthetic production of oxygen, all heterotrophic organisms oxidized organic carbon under anaerobic conditions. The biochemical oxidation of carbon under such conditions is thermodynamically inefficient, often leading to fermentation. For example, in the absence of oxygen, molecules such as acetate are potentially oxidized to methane and CO_2:

$$H_3CCOOH \rightarrow CH_4 + CO_2$$

where the methane is liberated to the environment.

The generation of oxygen led to thermodynamically favorable conditions for the oxidation of organic matter via aerobic respiration. The basic redox sequence for aerobic respiration is

$$(CH_2O)_n + O_2 \rightarrow CO_2 + H_2O.$$

This thermodynamic coupling between the photosynthetic reduction of organic carbon and the oxidation of water is reversed in mitochondria, in which oxygen is used as a sink for electrons and the oxidation of organic carbon is coupled to

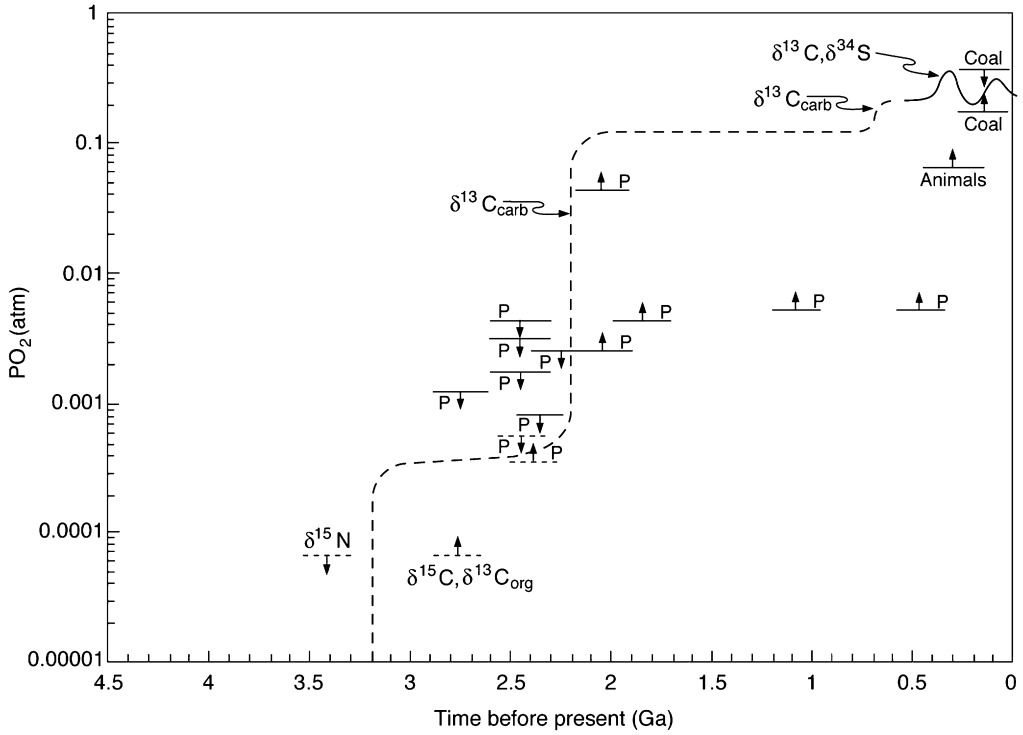


Figure 4 The evolution of oxygen on Earth (reproduced from Rye R and Holland H (1998) Paleosols and the evolution of atmospheric oxygen: a critical review. *American Journal of Science* 298(8): 621–672).

biochemical energy. Initially, the oxygen could be utilized by single-celled organisms via diffusion; subsequently, however, proteins and structures evolved that facilitated the diffusion and transport of O_2 and led to the evolution of metazoans.

Sequestration and Burial

From a biogeochemical perspective, the net evolution (i.e., accumulation) of molecular oxygen in the atmosphere means that oxidation of water and the (bio)geochemical reduction of free oxygen are not balanced, i.e., photosynthetic oxygen evolution must have exceeded oxygen consumption. Given the vast quantities of water on the surface of the planet, oxygenic photosynthesis could produce enormous quantities of oxygen without running out of substrate. When can such conditions occur and what are the oxygen-consuming reactions?

Prior to the net accumulation of oxygen in the atmosphere, oxygen produced by oxygenic photoautotrophs in the oceans would have been exposed to several potential reductants, two of which are especially biogeochemically important, namely, sulfur and iron. After Cl^-, SO_4^{2-} is the most abundant anion in the ocean. Under the mildly reducing conditions of the Archean ocean, dissolved aqueous sulfur in the ocean would have been in equilibrium between a reduced form, probably sulfide (H_2S), and the hydrated, oxidized form H_2SO_4. The latter (sulfuric acid) would equilibrate with the alkaline or earth alkaline elements of the first and second series, resulting in the formation of ionized sulfate salts. The precipitation of sulfate minerals (e.g., gypsum) could occur in shallow seas when evaporation leads to super-saturation and during periods when the influx of Ca from continental sources is high. The latter occurred during periods of rapid invasion of land by terrestrial plants (e.g., during the Triassic Period), where plant growth promoted weathering. Thus, photosynthetic evolution of oxygen led to the formation of vast quantities of sulfate via the oxidation of reduced sulfur anions; sulfur oxidation acted as a sink for molecular oxygen.

Under reducing conditions, iron, as Fe(II), is relatively soluble in seawater, whereas its oxidized counterpart, Fe(III), is virtually insoluble. The photosynthetic oxidation of the Archean ocean not only resulted in the oxidation of sulfur but also led to the oxidation and subsequent precipitation of iron oxides (rust) in ocean sediments. Thus, the biologically catalyzed oxidation of iron served as a second electron sink for molecular oxygen—a sink that ultimately came to represent the largest reservoir for the transition metal on Earth. The deposition of iron oxides would proceed until virtually no soluble iron remained in the ocean. The deposition of iron in Archean sediments is recorded in several deposits, the dating of which provides a record of the oxidation of the atmosphere (**Figure 5**).

When the electron sinks in reduced iron and sulfur were consumed, oxygen could be consumed in respiratory reactions involving (photosynthetically) reduced carbon. If, however, there was a net evasion of photosynthetically derived oxygen to the environment (i.e., the oceans and atmosphere), then a stoichiometric amount of organic carbon must have avoided oxidation, i.e., organic carbon must have accumulated. The primary sink for organic carbon is marine sediment in which, in the absence of oxygen, carbon adsorbs to inorganic sedimentary particles and becomes buried in the sedimentation process. The sequestration and burial process occurred over

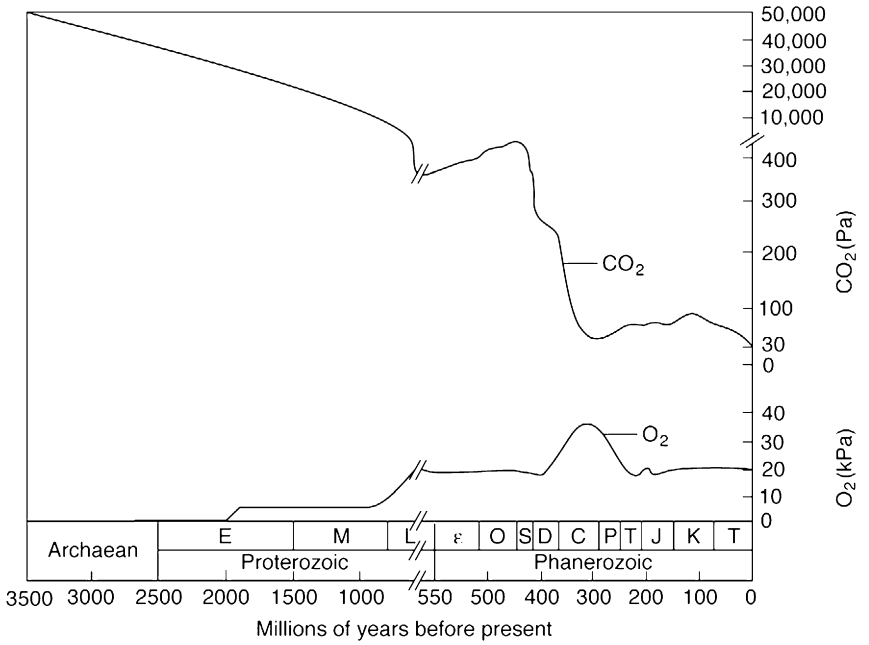


Figure 5 A reconstruction of variations in the partial pressures of CO_2 and O_2 in the atmosphere through geological time using data from Berner (1990, 1993) and Berner and Canfield (1989) for the post-Cambrian epochs (i.e., the Phanerozoic). The absolute values and timing for the evolution of oxygen are not constrained in the Proterozoic Epoch.

hundreds of millions, if not billions, of years, and it resulted in the accumulation of massive amounts of organic carbon associated with sedimentary and lightly metamorphosed rocks (e.g., shales and cherts). A very small fraction of the buried organic carbon underwent alterations due to heat and pressure (i.e., diagenesis) to become fossilized carbon that literally fuels the industrialized world in the current geological epoch (**Table 2**).

Primary Production in the Contemporary World

The carbon remaining in photosynthetic organisms following their own respiratory costs is potentially available for consumption by other organisms. This remaining carbon, or net primary production (NPP), provides the energy requirements of all ecosystems. The distribution of NPP between the terrestrial and marine ecosystems is given in **Table 4**. These data suggest that terrestrial ecosystems, containing approximately 99% of the plant biomass, account for approximately 55% of the global NPP, whereas marine ecosystems, containing approximately 1% of the plant biomass, account for 45% of the productivity. The huge discrepancy between NPP per unit biomass between these two ecosystems suggests that, on average, the lifetime of marine primary producers is less than 1 week, whereas the lifetime of terrestrial primary producers is on the order of a decade or more. On land, NPP is primarily limited by the availability of water, whereas in the oceans NPP is primarily limited by the availability of essential nutrients, especially fixed forms of nitrogen.

During the past several hundred thousand years, global NPP has remained relatively constant, whereas the distribution of NPP between terrestrial and marine ecosystems has waxed and waned with climatic shifts. Briefly, every 100,000 years or so, changes in Earth's orbital relation to the sun lead to a small reduction in surface temperature, which is amplified (through unknown processes) to produce glacial periods. The advance of ice sheets in the Northern Hemisphere, combined with a reduction in liquid precipitation at low latitudes, leads to the loss of terrestrial NPP by approximately 30%. Simultaneously, however, enhanced up-welling fluxes of nutrients in the oceans combined with a strengthened eolian flux of iron stimulate oceanic NPP, such that the change between the terrestrial loss and oceanic gain is approximately balanced. During interglacial periods, the situation reverses, with a retreat of the ice sheets and the relaxation of nutrient enrichment of the upper ocean to lead to the situation shown in **Table 4** for the contemporary world.

Oceanic Solubility and Biological "Pumps"

The oceans contain approximately 50-fold more organic carbon than the atmosphere, and on timescales of centuries the former reservoir controls the concentration of CO_2 in the latter. The vertical gradient of inorganic carbon in the oceans is "inverted"—that is, there is a higher concentration of inorganic carbon in the ocean interior than at the surface (**Figure 6**). Because the ocean surface tends to equilibrate with the atmosphere, the vertical gradient is maintained by processes other than simple equilibrium diffusion. There are two major processes responsible for the observed gradient. First, cooling of waters at high latitudes increases the solubility of CO_2 while simultaneously increasing the density of the waters. The cold, CO_2-rich waters sink into the ocean interior, where they are transported by density discontinuities and the motion of the earth throughout the world oceans. The round-trip takes approximately 1000 years before the waters are again in

Table 4 Annual and seasonal NPP of the major units of the biosphere, from CASA-VGPM[a]

	Ocean NPP		*Land NPP*
Seasonal			
April to June	10.9		15.7
July to September	13.0		18.0
October to December	12.3		11.5
January to March	11.3		11.2
Biogeographic			
Oligotrophic	11.0	Tropical rain forests	17.8
Mesotrophic	27.4	Broadleaf deciduous forests	1.5
Eutrophic	9.1	Broadleaf and needleleaf forests	3.1
Macrophytes	1.0	Needleleaf evergreen forests	3.1
		Needleleaf deciduous forest	1.4
		Savannas	16.8
		Perennial grasslands	2.4
		Broadleaf shrubs with bare soil	1.0
		Tundra	0.8
		Desert	0.5
		Cultivation	8.0
Total	48.5		56.4

[a]Ocean data are averages from 1978 to 1983. The land vegetation index is from 1982 to 1990. All values are in petagrams of carbon (1 pg = 10^{15} g). Oligotrophic <0.1 μg Chl/L; mesotrophic 0.1–1.0 μg Chl/L; eutrophic >1.0 μg Chl/L. The macrophyte contribution to ocean production is not included in the seasonal totals. Reprinted from Field C, Behrenfeld M, *et al.* (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281: 237–240.

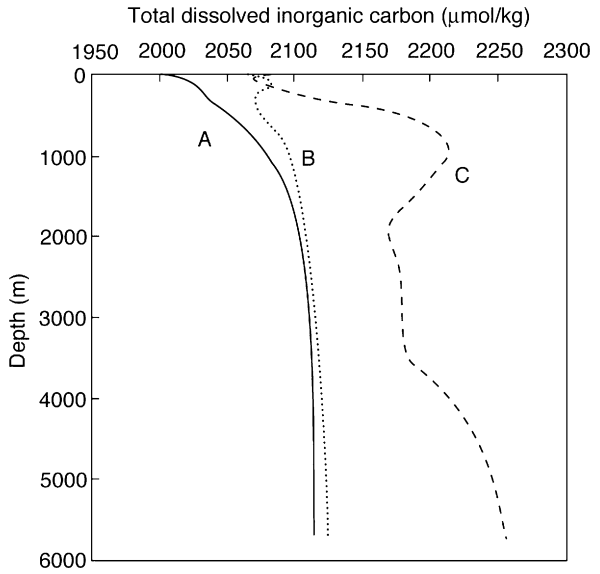


Figure 6 Total dissolved inorganic carbon (DIC) concentration in the ocean. (A) Modeled values prior to 1765; (B) modeled values from 1990; (C) observed totals in 1990. The difference between curves C and B represents the contribution of the biological pump to the sequestration of DIC in the ocean interior.

contact with the atmosphere in the high latitude from whence they originated the journey. This "solubility" pump is responsible for approximately 50% of the inorganic carbon gradient. The other 50% is a consequence of biological activity. In the upper, lighted portion of the ocean, photosynthetic fixation of inorganic carbon by phytoplankton leads to the formation of organic matter, largely in the form of particulate materials. Gravity, acting on the particles, leads to a persistent "rain" of organic matter into the ocean interior. The organic matter is oxidized by heterotrophic bacteria and other organisms, leading to the production of inorganic carbon. Thus, this vertical flux of organic carbon and its subsequent oxidation effectively pumps up the ocean interior with inorganic carbon. The biological pump is maintained by the upward flux of inorganic nutrients from the ocean interior, especially fixed nitrogen and phosphate. Such nutrients are essential to supporting the photosynthetic activity of phytoplankton in the upper ocean.

The Nitrogen Cycle

All organisms contain nitrogen that is reduced to the equivalent of NH_3 (ammonia). The most abundant form of nitrogen in the environment is N_2, which at the temperature and pressure on Earth's surface is a gas and forms 78% by volume of the planet's atmosphere. Conversion of N_2 to NH_3 requires the addition of three protons and three electrons per atom of N; the balanced equation is

$$N_2 + 6H^+ + 6e^- \rightarrow 2NH_3$$

This reaction is kinetically limited at low pressure and the physiological temperatures, and without some catalytic enhancement there would be no ammonia in the environment. A small subset of bacteria evolved with the capability to "fix" N_2. Nitrogen fixation, or diazotrophy ("eating nitrogen"), is catalyzed by a protein complex (nitrogenase) which is irreversibly inactivated by oxygen. The enzyme complex normally contains an Fe subunit and a Mo, Fe subunit; however, V and Fe can substitute for Mo as mediators of electron transfer in some organisms. Gene sequence analysis suggests that nitrogenase evolved from a single common ancestor, and the enzyme is primarily found in eubacteria, although some

Archea, such as methanogens, can also fix nitrogen. *In vivo* nitrogen fixation consumes approximately 16 ATP molecules per mole of N_2 fixed; this high-energy demand requires significant energy supply. In eubacteria, the supply of energy is either light (for N_2 fixing cyanobacteria), or reduced carbon substrates (for heterotrophic bacteria). For methanogens it is H_2. Whatever the energy source, N_2 fixation is a metabolically expensive process, and hence, in the presence of high concentrations of fixed nitrogen, it usually is repressed. Many nitrogen-fixing bacteria live symbiotically with a host organism. One classical symbiosis is found in the Legumacea, which are higher plants such as peas and clover in which nitrogen-fixing bacteria are found in nodules attached to the plant root. In a second symbiosis, some insects, such as termites, contain nitrogen-fixing bacteria within their guts to supply essential amino acids to the host.

In most oxidized environments, fixed inorganic nitrogen is found in the form of nitrate (NO_3^-) or, to a lesser extent, nitrite (NO_2^-). These two oxidized forms of nitrogen are formed from ammonium via the action of nitrifying bacteria. Nitrification is a metabolic process in which the oxidation of ammonium can be coupled to the reduction of inorganic carbon to form organic molecules without light energy. This type of metabolic process is called chemoautotrophy. For example,

$$2NH_3 + 2CO_2 + O_2 \rightarrow 2(CH_2O) + 2NO_2^- + 2H^+$$

The pathway for ammonium oxidation is broken down between organisms that can oxidize the substrate to nitrite and those that further oxidize nitrite to nitrate. Nitrification is a metabolically inefficient process; one mole of NH_3 is oxidized to produce each mole of organic carbon. Hence, the growth rate of nitrifying bacteria is generally low, whereas in the presence of O_2 the rate of ammonium oxidation is generally rapid. Nitrification is confined to a small number of eubacteria; the gene sequence analysis of these organisms suggests the pathways evolved from single ancestral genes and are relatively conserved. Nitrate (NO_3^-) is a thermodynamically stable molecule under the oxidizing conditions of the contemporary world, and as such it is the most prevalent form of fixed nitrogen in the world oceans.

The cycle of nitrogen is completed when oxidized species of nitrogen, especially NO_3^-, are subsequently reduced to form N_2O and N_2. This anaerobic process, called denitrification, is mediated by a wide range of diverse bacteria and Archea. Denitrification is a respiratory pathway in which the terminal electron acceptor is NO_3^- or NO_2^-. Organisms generally reduce these two substrates "opportunistically," that is, when they are deprived of oxygen. The reduction of nitrate to nitrous oxide and N_2 represents a large loss of fixed nitrogen in the environment and largely occurs in anaerobic sediments of marine and lacustrine environments and in areas of soil rich in organic carbon that are periodically flooded (e.g., rice paddies and river deltas).

Although approximately half of all the nitrogen fixed on Earth is due to natural biological activity in the environment, half is due to deliberate fixation by human activities (**Figure 7**). Developed in the latter part of the nineteenth century by a German chemist, the Haber reaction permits the formation of ammonium from N_2 by a high-temperature, high-pressure reduction in the presence of H_2. This reaction is the primary source of fertilizer nitrogen for commercial crops, without which agricultural production would be much smaller than it is currently. However, approximately 25% of the applied nitrogen to agricultural crops is lost through denitrification, leading to an increase in N_2O in the atmosphere. A significant fraction of the applied nitrogen in terrestrial ecosystems is further solubilized in waters, flowing either into groundwater or through surface waters to the sea. Most of this fixed nitrogen seldom makes its way to the open oceans; rather, due to the high rates of denitrification on continental margins (especially in sediments on the margins), the fixed nitrogen tends to be lost to the atmosphere through denitrification.

Profiles of Elements in The Sea

In the oceans there are three basic profiles to describe the distribution of elements. If an element is conservative, and is either extraordinarily abundant or does not interact with particulate matter (i.e., biological material), it will be homogeneously distributed. The soluble salts in seawater represent a vast quantity of alkaline metals, halides, and sulfates. These elements are so abundant relative to biological demands and biomass that they behave as conservative tracers; that is, their distributions are effectively uniform through the oceans. If an element is biologically reactive and can be absorbed in surface waters and subsequently transported by sinking to the ocean interior, the interior ocean can be enriched in the element relative to the surface. Such elements are said to have "nutrient-like" behavior. Examples of such elements include C, P, fixed N, Si, Fe, Cu, Ni, Cd, Hg, and the rare earth elements (REE's). It is important to note that simply because an element has a nutrient-like behavior it is not necessarily a true nutrient and vice versa. A REE, such as La^{3+}, is highly reactive with sulfhydryl ($-SH$) groups found in most proteins and hence can be "scavenged" from the surface ocean by particles. In contrast, sulfate is so abundant in the ocean that even though it is used as a nutrient in formation of proteins, the level of utilization is so small in comparison with its concentration that it behaves more like a conservative tracer than a nutrient. The third category of elemental distributions leads to an elevation in surface water and is primarily associated with an upper ocean source. One example is the vertical profile of O_2, which is generated by photosynthetic activity in the upper ocean and consumed in the sea interior by respiration.

Sulfur

Sulfur is an integral element in two amino acids, methionine and cysteine, and hence is an essential element for protein synthesis in all organisms. The biogeochemical cycle of sulfur is primarily dependent on microbial metabolism. Under anaerobic conditions, sulfate is used as an electron sink for a wide variety of anaerobic heterotrophic microbes, leading to the formation of H_2S. This same gaseous product can be photochemically oxidized by anaerobic photosynthetic bacteria to produce reductants for inorganic carbon. Thus, a simple cycle of sulfur can be generated under totally anaerobic

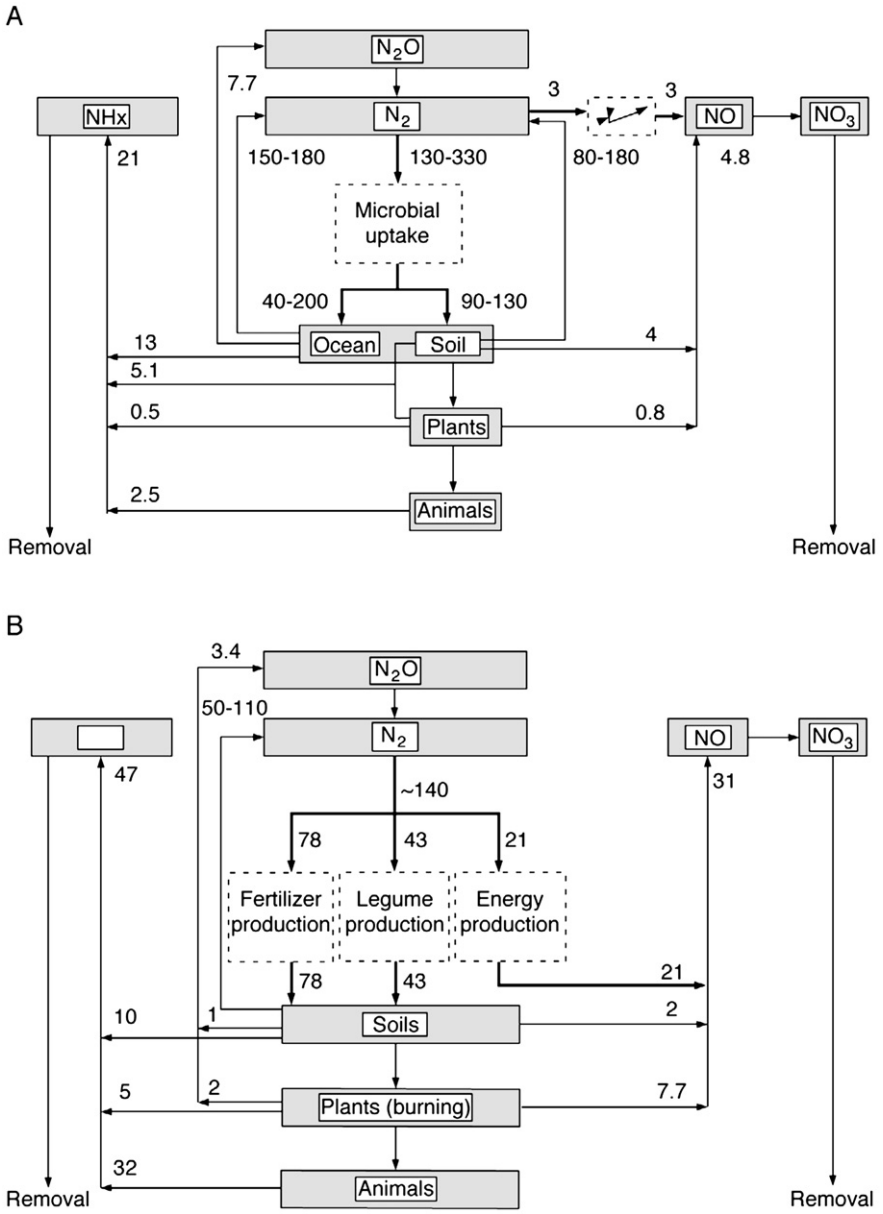


Figure 7 Global N-fixation rates and emission to the atmosphere (Tg N year $^{-1}$) from (A) preindustrial and (B) anthropogenic sources. The boxes with dashed lines represent processes that convert unreactive N to reactive N. Reproduced from Gallaway, *et al.* (1995) N-fixation, anthropogenic influence. *Global Biogeochem. Cycles* 9: 235–252.

conditions, as long as an external energy source (e.g., light) is available to keep the system far from thermodynamic equilibrium. H_2S is also produced from geochemical reduction of sulfur in the earth's mantle and is emitted to the atmosphere and oceans through vulcanism and geothermal springs (e.g., hyperthermal deep-sea vents). The reduced sulfur flux at deep-sea vents provides a reductant for carbon fixation in symbiotic non-photosynthetic (i.e., chemoautotrophic) bacteria found in the guts of a variety of invertebrates that colonize vent regions. This metabolic pathway is also sustained by photosynthesis, albeit indirectly. The thermodynamic gradient that allows chemoautotrophic sulfur-oxidizing bacteria to extract reducing power from H_2S depends on the oxidation of the ocean by oxygenic photoautotrophs. Under aerobic conditions, sulfur is primarily found in an oxidized form, usually as a sulfate salt, that is assimilated and reduced by photoautotrophs and microbes to a sulfhydryl ($-$SH). Hence, sulfur is both a source and a sink for metabolic electrons.

In the oceans, sulfate is the second most abundant anion following chloride. The biological assimilation and reduction of sulfur by marine microbes and plankton leads to the formation of volatile sulfur gases, notably dimethylsulfide and, in anaerobic aquatic systems, H_2S. Upon exposure to oxygen in the atmosphere, these reduced forms of sulfur are oxidized and hydrolyzed to form sulfate salts that are transported as aerosol particles. Additional sources of aerosol sulfate include

volcanic eruptions and the combustion by-products of fossil fuels, especially coal. Aerosol sulfate plays two important roles in climate feedbacks. First, the particles are hygroscopic, and upon absorption of water vapor they enhance the scattering of shortwave solar radiation (i.e., the planetary albedo). This light-scattering process tends to reduce the heating of the earth by the sun (the so-called "white house effect"). Second, as hygroscopic particles, aerosol sulfate particles provide a nucleation site for water precipitation in clouds, i.e., the particles act as cloud condensation nuclei. The atmospheric lifetime of sulfate particles is short, on the order of weeks. As water vapor builds in clouds, or the clouds are advected aloft and cool, precipitation of liquid water (i.e., rain) washes sulfate from the atmosphere. When aerosol sulfate loading is excessive (e.g., when fossil fuel combustion is high), a small fraction of hydrated sulfate in rain is not buffered by cations, and the sulfate reacts with water to form a dilute solution of sulfuric acid (H_2SO_4), producing acidic rain. Note that the transport of sulfur from continental and aquatic sources to and from the atmosphere is a consequence of phase state transitions, redox reactions, and hydrolysis.

Phosphorus Cycle

Phosphorus is an essential component of all cells; the element forms the backbone of nucleic acids, without which cells cannot reproduce. Phosphorus is also incorporated in nucleotides, sugars, proteins, and lipids. In vertebrates, almost all the phosphorus is combined with calcium to form hydroxylapatite, which is the central molecule to build bones.

In the environment, phosphorous is found almost invariably in the oxidized form as hydrated phosphate. Mineral phosphate salts erode and weather from continental sources and are carried to aquatic ecosystems as soluble ions. There is no eolian phosphorus source and no significant transfer of phosphate from aquatic or terrestrial environments to the atmosphere. Unlike carbon, nitrogen, phosphate, or oxygen, phosphorus does not undergo significant biologically mediated redox reactions; the chemistry of phosphorus is solely based on hydrolysis.

Trace Metals

Transition metals play a key role in biological electron transfer (redox) reactions and biological processes in turn affect the availability of metals in the environment. The metals are incorporated into protein "scaffolds" in specific orientations and associated with specific ligands such that the electron transfer reactions are highly specified and optimized. The transfer of electrons to given substrates is often keyed to specific metals. For example, electron transfers involving O_2 frequently are mediated by Mn, whereas electron transfers involving N frequently are mediated by Mo. The availability of transition metals, in turn, is dictated by solubility and redox conditions. Under anaerobic conditions, Mn and Mo are often sequestered as sulfide precipitates, but they become more available under oxidizing conditions. In contrast, Fe, which is a required electron transition metal in all organisms, is abundant under reducing conditions but scarce under oxidizing conditions.

As previously noted, electron transfers through iron are essential to both anaerobic and oxygenic photosynthesis, heterotrophic respiration, and nitrogen fixation. In some cases, anaerobic bacteria can oxidize Fe(II) to Fe(III), thereby precipitating iron as a ferric hydroxide in the sediment. Although iron is the most abundant transition metal in the earth's crust, its concentration in the open ocean is extremely low, averaging no more than ~ 1 nM. To satisfy iron fluxes, eolian sources of iron are critical. Wind-blown dust from major desert areas supplies much of the iron to the open oceans. Consequently, open ocean areas, far removed from continental dust regions, have extremely low iron concentrations that result in a direct limitation of photosynthetic electron transport. Three such areas are the Southern Ocean, the eastern Equatorial Pacific, and the sub-Arctic Pacific. Ice core records suggest that during glacial periods eolian iron fluxes were much greater. The variations in dust supply appear to correspond with changes in the hydrological cycle; warm, wet climate regimes tend to reduce the eolian supply and hence may limit the biological sequestration of carbon in the oceans, whereas cold, dry periods may stimulate oceanic primary production.

Mn is essential for oxygen evolution and, like iron, its concentration is extremely low in the oceans. The Mn requirements for biological activity are lower than those for Fe; consequently, there is no evidence of Mn limitation in either the oceans or the terrestrial ecosystems. Interestingly, Mn and Fe can exchange electrons spontaneously and, in the presence of a redox gradient established in sediments, can cycle between oxidized and reduced states. Mn has two main oxidation states, namely, Mn(II) and Mn(IV). Organic carbon and NH_4 can be oxidized by Mn(IV), forming CO_2 and N_2, respectively, whereas reduced Mn can be oxidized by O_2 and possibly by NO_3.

Many other transition metals have been co-opted for various biologically mediated electron transfers. Mo is often utilized biologically in electron transfer reactions involving nitrogen. Most nitrogenases contain Mo, as do all nitrate reductases. Copper is frequently found in hydrophyllic proteins that either catalyze single electron transfer reactions (e.g., plastocyanin) or coordinate oxygen binding (in the case of hemocyanin). The availability of these metals is inverse to those of Mn and Fe, namely, under oxidizing conditions, Mo and Cu are more abundant than under reducing conditions (Williams and Frausto da Silva, 1996).

Evolution of Biogeochemical Cycles

By definition, biogeochemical cycles are mediated in part by living organisms. The origins of biogeochemical cycles are presumably entwined with the origins of life. Although there is no consensus on the definition of life, to sustain life all living organisms must grow and reproduce. Growth is dependent on the availability of substrates and energy. C, N, P, H, O, and S are required substrates; often these substrates are combined (e.g., all organisms require phosphorus in the form of the phosphate anion, $H_2PO_4^{2-}$). All life forms either oxidize or reduce carbon; hence, carbon is either an electron sink or a source, depending on the thermodynamic basis of the

organism's metabolism. Given an external energy source, the redox reactions can run "uphill," where an environmental redox gradient is used to thermodynamically provide the free energy for the subsequent carbon reduction. This metabolic strategy characterizes chemoautotrophs. In photoautotrophs, the energy is externally provided by light. All carbon-oxidizing organisms require an alternative electron sink, such as O_2, NO_3, and SO_4.

All the basic biogeochemical cycles evolved more than 2 billion years ago in unicellular organisms. The evolution of the cycles is interdependent; that is, the requirement by all organisms to obtain reduced carbon for energy and growth created a selection pressure. Under moderately reducing conditions of the early earth, processes that led to N_2 fixation and methane production were more favorable than under oxidizing conditions; hence, these biogeochemical processes are among the earliest. Phylogenetic trees, constructed from 16S RNA sequence homology, suggest that the earliest organisms were nonphotosynthetic, thermophyllic chemoautotrophs that are placed at the root branch between the Archea and Eubacterial kingdoms. These early organisms could have used inorganic substrates, such as H_2, H_2S, and Fe^{2+}, to reduce CO_2 to carbohydrate. Indeed, such organisms persist and thrive in deep-sea vents, in volcanic hot springs deep in Earth's crust, and in other "extreme" environments in which liquid water and suitable oxidizable inorganic substrates are available. Chemoautotrophs are almost certainly the precursors of photosynthetic cells. The evolution of a photosynthetic process in a chemoautotroph forces consideration of both the selective forces responsible and the mechanism of evolution.

Reductants for chemoautotrophs are generally deep in the earth's crust. Vent fluids are produced in magma chambers connected to the athenosphere. Therefore, the supply of vent fluids is virtually unlimited. In the contemporary ocean, the chemical dissequilibria between vent fluids and bulk seawater (which is highly oxidized) provides a sufficient thermodynamic gradient to continuously support chemoautotrophic metabolism. However, the redox gradient in early Earth's oceans did not exist prior to oxygenic photosynthesis. Moreover, magma chambers, vulcanism, and vent fluid fluxes are tied to tectonic subduction regions, which are transient features of Earth's crust and hence only temporary habitats for chemoautotrophs. To colonize new vent regions, the chemoautotrophs would need to have been dispersed throughout the oceans by physical mixing. This same dispersion process would have helped ancestral chemoautotrophs exploit solar energy near the ocean surface. The evolution of photosynthetic organisms led to a highly redundant process; that is, many organisms are capable of reducing inorganic carbon to form organic materials utilizing solar energy. The ecological (i.e., "functional") redundancy ensures a continuity of photosynthetic carbon fixation in all environments in which light and liquid water are available. Currently, there are approximately 20,000 species of aquatic oxygenic photosynthetic organisms belonging to 10 divisions; there are approximately 250,000 species of terrestrial (vascular) plants belonging to 3 divisions (Falkowski and Raven, 1997).

The evolution of oxygenic photosynthesis provided a niche for the co-evolution of ammonium and sulfide-oxidizing bacteria. The action of these organisms led to the chemoautotrophic production of NO_3^{2-} and SO_4^{2-}, which are the most common forms of N and S found in the oceans and most terrestrial ecosystems today. The cycles of these elements are completed under anaerobic conditions, in which the oxidized forms of the elements are used as electron sinks in respiratory pathways. Under many conditions, fixed nitrogen limits primary production in both terrestrial and aquatic ecosystems. In contrast to photosynthesis, N_2 fixation is constrained to a relatively small number of bacteria, and hence the process is "functionally singular." The lack of diversity among N_2 fixers and the subsequent low supply of fixed N to many ecosystems represents a bottleneck in biogeochemical cycles. In the oceans, N_2 fixers appear to be limited by the availability of Fe and P.

See also: Atmospheric Gases. Carbon Cycle. Diversity, Organism Level. Geologic Time, History of Biodiversity in. Greenhouse Effect. Nitrogen, Nitrogen Cycle. Thermophiles, Origin of. Vents

References

Berner EK and Berner RA (1987) *The Global Water Cycle: Geochemistry and Environment*. Englewood Cliffs, N.J: Prentice Hall.

Berner R and Canfield D (1989) A model for atmospheric oxygen over Phanerozoic time. *Am. J. Sci.* 289: 333–361.

Chahine M (1992) The hydrological cycle and its influence on climate. *Nature* 359: 373–380.

Charlson RJ, Schwartz SE, *et al.* (1992) Climate forcing by anthropogenic aerosols. *Science* 255: 423–430.

Degens ET (1989) *Perspectives on Biogeochemistry*. New York: Springer-Verlag.

Falkowski P (1997) Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* 387: 272–275.

Falkowski PG and Raven JA (1997) *Aquatic Photosynthesis*. Oxford: Blackwell.

Fenchel T, *et al.* (1998) *Bacterial Biogeochemistry: The Ecophysiology of Mineral Cycling*. New York: Academic Press.

Field C, Behrenfeld M, *et al.* (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281: 237–240.

Holland HD (1984) *The Chemical Evolution of the Atmosphere and Oceans*. Princeton, NJ: Princeton Univ. Press.

Raymo ME and Ruddiman WF (1992) Tectonic forcing of late cenozoic climate. *Nature* 359: 117–122.

Sarmiento JL and Bender M (1994) Carbon biogeochemistry and climate change. *Photosyn. Res.* 39: 209–234.

Schlesinger WH (1997) *Biogeochemistry: An Analysis of Global Change*. New York: Academic press.

Williams R and Frausto da Silva J (1996) *The Natural Selection of the Chemical Elements*. Oxford: Clarendon.

Biogeographical Models

Thiago F Rangel and JAF Diniz-Filho, Federal University of Goiás, Goiás, Brazil

© 2013 Elsevier Inc. All rights reserved.

Glossary

Empirically derived correlative model A model that generalizes the observed relationship between patterns of species richness and one or more predictors by means of a linear or curvilinear equation. Inference drawn from these models is often limited to qualitative evaluation of the explanatory power of predictors.

Species distribution model A model that attempts to describe and understand the geographic distribution of species, usually through interpolation and/or extrapolation of empirically derived correlations between the environmental factors and the presence/absence of the species.

Species richness The number of taxa found at a given location, at a given time. Measures of species richness are strongly affected by spatial and temporal scale.

Stochastic simulation model A spatial and/or temporal dynamic model that characterizes discrete agents (usually species), and the mechanisms hypothesized to affect properties of these agents. Unknown mechanisms or known mechanisms not included in the model are represented by stochastic processes.

Theoretically derived correlative model A model that is derived from fundamental and lower-level scientific principles in the form of a mathematical equation, and is evaluated through a correlation between its prediction and observed data.

Introduction

Biogeography is concerned with the causes of geographical distribution of species and biological entities at high hierarchical levels, and with their dynamics through evolutionary time (Lomolino *et al.*, 2009). Biogeographical dynamics of multiple species in space and time generate emergent patterns, especially of species richness and turnover, across a wide spectrum of organizational scales. These patterns are also influenced by a complex set of extrinsic (i.e., ecological or environmental, and evolutionary or historical) and intrinsic (i.e., life history, physiological, behavioral, and morphological) factors and their interaction.

Because of the broad geographical and temporal scales usually involved in determining patterns of biodiversity, it is difficult to test hypotheses about each particular ecological and evolutionary mechanism that is thought to drive the biogeographical patterns. The principal existing approach to this problem relies on statistical analysis of patterns in relation to environmental characteristics. This approach provided the first important evidence about these hypotheses and has since been widely employed in biogeographical studies. However, as explained below, correlative modeling is not a strong tool for understanding or discriminating the underlying mechanisms behind biodiversity patterns. Controlled experiments offer a more powerful way to understand complex phenomena, but unfortunately the spatial and temporal scales on which hands-on experiments are feasible (and the relatively constrained circumstances under which they are ethical) place such experiments out of reach for many compelling and longstanding questions. Scientists in other fields whose objects of study cannot be manipulated, such as astrophysicists interested in the internal structure of stars, have taken advantage of high-capacity computers to carry out virtual experiments on simulated systems designed from first principles. This strategy has only recently begun to be applied to understand biogeographical and macroecological patterns and processes (Gotelli *et al.*, 2009).

Models can be built to explain many different biogeographical patterns, but here we focus on models for species richness patterns, because they illustrate how biogeographical models have progressed during the recent decades and have been a focal point in the development of the field. In as early as the eighteenth and nineteenth centuries, naturalists and biogeographers described and documented what today we call geographical (latitudinal) gradients in taxon diversity (species richness), in which diversity decreases from the tropics towards temperate and polar regions for most taxa (e.g., Willig *et al.*, 2003). The number of species of organisms varies greatly over the continents of the Earth, from species-poor habitats at high latitudes, high elevations, and in arid ecosystems to species-rich habitats in the wet tropics (Hawkins *et al.*, 2003b; Willig *et al.*, 2003). Surprisingly, even after 200 years of research in biogeography and ecology and some important advances mainly in the last 20 years, there is still no clear understanding or consensus on how these patterns arise and are maintained (Hawkins, 2001). What no one doubts is that many interacting causes are at work, and this in turn demands the development of increasingly complex approaches to deal with the problem (**Figure 1**).

The recent conceptual development in the new field of macroecology (Brown, 1995; Gaston and Blackburn, 2000) triggered a demand for more theoretically oriented analyses of biogeographical patterns, in which ecological and evolutionary mechanisms underlying the patterns are simultaneously incorporated into the model, rather than simply documenting or describing the patterns themselves. Biogeographical models have evolved from the simple correlative approach, with increasing degrees of complexity in their predictions, to analytical and theoretically oriented approaches. Finally, during the recent years, biogeographers have begun to develop complex, spatially and temporally explicit computer simulations,

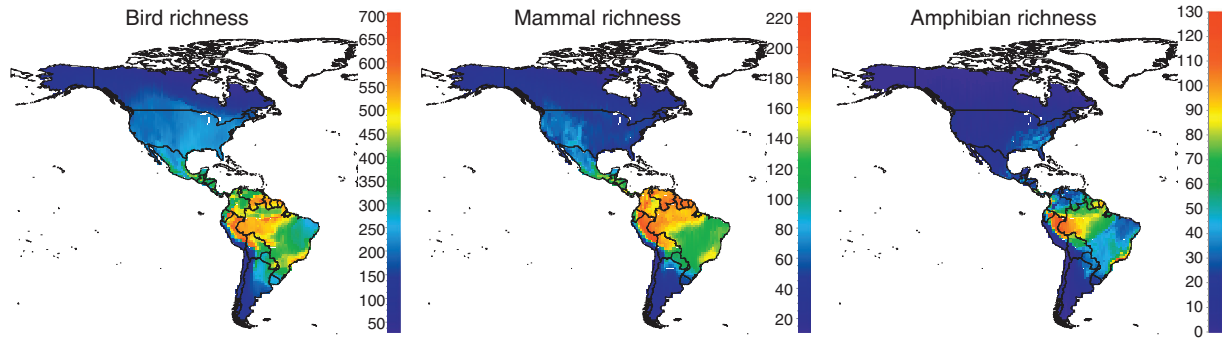


Figure 1 Maps of species richness of bird, mammal, and amphibian species richness in the Western Hemisphere.

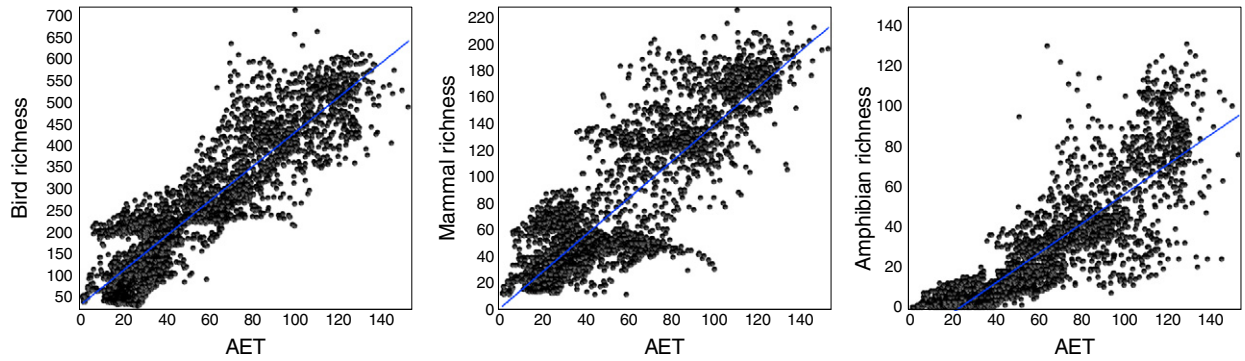


Figure 2 Relationship between species richness of birds, mammals, and amphibians and actual evapotranspiration (AET) in the Western Hemisphere.

designed to account for multiple interactive causes and the stochastic nature of biological entities.

Correlative Models

Initial hypotheses explaining geographical patterns in species richness were deduced solely by observing and describing nature, and were based on nothing more rigorous than an intuitive and loose correspondence between a factor (usually climatic) and biological patterns. Influential early examples include analyses by A.G. Fischer (Fischer, 1960), by G.G. Simpson for North American mammals (Simpson, 1964), and a conceptual review by E.R. Pianka (Pianka, 1966; see Hawkins, 2001 regarding contribution by even earlier researchers). Many other hypotheses were developed, but most of them were restricted to particular groups of organisms, and were empirically derived or untestable (Rhode, 1992). Whittaker *et al.* (2001), Rahbek and Graves (2001), and Willig *et al.* (2003) reduced the large number of initially available hypotheses to a much smaller number of testable and general set of hypotheses, including energy availability, evolutionary time, and habitat heterogeneity, as well as area and geometric constraints, providing a better framework for testing each hypothesis. Of course, testing each of these groups of hypotheses requires establishing environmental or evolutionary surrogates for each of them. For instance, temperature is usually used as a surrogate of energy hypothesis, and actual evapotranspiration (AET) is used when one wants to test the coupled effects of water and energy availability (e.g., Rodriguez *et al.*, 2005). This correspondence is not of course undisputable and may be scale-dependent. A good example is using elevational range (within a species geographical range or habitat) as proxy of habitat heterogeneity, when, in fact, elevation can also be considered a surrogate of isolation and speciation rates.

Until recently, the most common framework used in such investigations still relied primarily on statistical measurements of the concordance between the broad scale spatial patterns in species richness and multiple environmental factors (e.g., Currie *et al.*, 2004). The generally strong relationship between species richness and some of these environmental factors, especially particular balances between water–energy measured by AET, has led many researchers to conclude that environment (e.g., climate) is the main driver of species richness, based on a purely correlative view (**Figure 2**; Hawkins *et al.*, 2003a,b). However, although the correlative approach provides the most straightforward description of the relationship between climate and patterns of species richness, it falls short at disentangling causal links. Nonetheless, many researchers have derived more complex predictions and employed, elegantly, correlative tools to quantify and test biogeographical hypotheses. Improving these tests, however, may require the development of auxiliary hypotheses based on different variables, or the deconstruction of observed patterns. For example, one may study how energy and productivity affect the patterns

of species abundances, as these two environmental factors are usually invoked as the mechanism driving richness gradients, assuming that higher number of individuals promote higher species richness (e.g., Currie *et al.*, 2004). Another example is the comparison of richness patterns between current and past modeled environments (Araújo *et al.*, 2008; Hortal *et al.*, 2011), or the deconstruction of richness patterns into basal and derived taxa to test for niche conservatism (Hawkins *et al.*, 2007a, b; Davies *et al.*, 2011). The deconstruction of biogeographical patterns by range size quantiles (at the simplest, small- vs. large-ranged species) has yielded both new understanding and some surprises (Lees *et al.*, 1999). Meta-analytical techniques can also be used to better understand in which situation (i.e., groups of organisms, spatial scales) gradients tend to appear or are being driven by a particular set of environmental factors (e.g., Field *et al.*, 2009).

Although climate surely has a strong effect on biodiversity patterns, as the geographic distribution of individual species is limited by the overall bioclimatic envelope within which the species is able to survive (Soberón, 2007), there are many pathways by which climate can be linked to diversity. For instance, from a population ecology perspective, it is not difficult to understand why increasing levels of water–energy allows higher carrying capacity in a given locality, but there are still large conceptual gaps to be filled before we understand how a larger population (number of individuals) promotes an increase in the number of species (biodiversity; Currie *et al.*, 2004). Filling these gaps requires the understanding of additional ecological and evolutionary mechanisms that act in concert, such as species interactions and their evolution, niche dynamics, dispersal mechanisms, and the balance between speciation and extinction as functions of mutation rates (Ricklefs, 2006).

An important class of correlative models has been widely used to model species distribution, roughly based on niche concepts (Soberón, 2007; Colwell and Rangel, 2009). These models, generally called Species Distribution Models (SDMs; or niche models; Peterson, 2006), are based on the overall reasoning of modeling the presence (and/or absence) of species as a function of environmental variation. Today there is a wide array of SDMs, ranging from a very simple bioclimatic envelope model (in which the species is assumed to be present everywhere within the environmental limits of its distribution) to complex artificial intelligence-machine learning models, such as neural networks and maximum entropy models. There are many discussions about how well these models determine species' distributions (Elith *et al.*, 2006; Elith and Graham, 2009), but it is more or less clear that they tend to represent potential distributions in the absence of other factors, such as dispersal limitation, historical constraints to distributions (barriers), and biotic interaction (Guisan and Thuiller, 2005). SDMs are currently quite widely applied, and have been primarily used in more applied biodiversity and conservation studies, i.e., to understand the effects on diversity patterns of climate change or invasive species.

Guisan and Rahbek (2011) recently developed an interesting framework, called Spatially Explicit Species Assemblage Models (SESAMs), in which richness maps obtained from what they called macroecological or top-down methods (i.e., empirically derived from observed range maps) are compared with maps derived from bottom-up methods (obtained by summing SDM maps). The framework consists of evaluating if broad scale diversity patterns can be considered a consequence of ecological processes acting at the level of species' ranges, and evaluating how other processes (i.e., dispersal limitation, historical constraints, or biotic interaction) can explain deviations between top-down versus bottom-up mapped patterns. For example, if a region is expected to have a high richness based on SDMs but there is actually a smaller number of species, this may be due to isolation and systematic lack of dispersal towards this region (a geographical dispersal constraint), or perhaps there are assemblage rules that reduce the coexistence of species in the region (biotic interaction such as competition). However, if a region harbors more species than expected by SDMs, one could evaluate if that region is prone to speciation (or buffered against extinction) independently of baseline environmental drivers of species distributions. Although further applications and tests of the protocol are needed, we believe SESAM certainly opens a significant new research avenue in macroecology, especially on how broad scale patterns emerge from species distributions, and hopefully the strategy will allow better inferences about mechanisms and will provide a more comprehensive view of how diversity patterns evolve at different hierarchical levels.

Indeed, one of the important conceptual problems of the "purely" ecological (contemporary) drivers of biodiversity patterns is the time-lag response of biodiversity to climate, and how climatic changes throughout ecological and evolutionary time have driven current diversity patterns. In fact, the protocol proposed by Guisan and Rahbek (2011) can be potentially useful to evaluate if the responses of biodiversity to climate are structured in time. As pointed out by Ricklefs (2006), it is surprising that, despite these advances, the evolutionary components underlying widespread correlations between climate and species richness remain so poorly understood. With regard to short ecological time scales, recent developments in climatic modeling allow us to understand changes in climate during the last few thousand years, which in turn opens the possibility of understanding how climatic stability is related to species richness. The biogeography of species in the Northern Hemisphere after the last glacial maximum provides important clues on the temporal rate at which biotas reach "equilibrium" with climate (Hawkins and Porter, 2003; Araújo *et al.*, 2008; Hortal *et al.*, 2011). Moreover, it has been demonstrated that diversity tends to accumulate in more climatically stable (usually warmer) areas, by allowing more time for adaptation and specialization, whereas unstable regions (usually cooler) are typically occupied by good dispersers that have evolved and maintained adaptations to these cooler regions (Dynesius and Jansson, 2000).

At deeper time scales, there is a growing consensus that correlations between diversity and climate arise as a byproduct of the combined evolution of climate and niche conservatism, which is a version of the environmental stability–richness mechanism (i.e., Wiens and Donoghue, 2004; Wiens *et al.*, 2010). As temperatures in the currently temperate regions declined in the last 30 million years, species previously adapted to widespread tropical climates contracted their geographic ranges or became extinct, creating a gradient between tropical and temperate regions. As evolutionary novelties

arose, eventually adaptations to cooler environment allowed species to colonize and diversify in the temperate regions. Notice that this explanation explains the evolutionary history of clades adapted to Miocene tropical environments worldwide, but fails to explain the biogeography of clades that evolved in a time when the planet was cooler, which by the same logic would produce a richness gradient inverse to what we see at present. If this concept is correct, for any given clade there should be a correlation between the direction of the latitudinal gradient of species richness and the time and place of origin of the clade.

The scenarios described above can be tested if dated phylogenies (and reliable fossil records) are available for a clade, as they can be used to draw predictions about the distribution of species' ages across space or the degree of phylogenetic clustering of species in the tropics compared to temperate regions (Hawkins *et al.*, 2007a, b; Hof *et al.*, 2010; Davies *et al.*, 2011). Alternatively, it is also possible to try to determine, using similar phylogenetic information, if the tropical–temperate gradient is driven by distinct balances between speciation and extinction rates (Weir and Schluter, 2006). However, most phylogenetically explicit tests of niche conservatism have been done at species or clade level, treating environment as a cross-species trait (e.g., Wiens *et al.*, 2006; Diniz-Filho *et al.*, 2007; Hof *et al.*, 2010). Algar *et al.* (2009) proposed an explicit test for these patterns of niche conservatism by using combined metrics for phylogenetic diversity applied to distinct regions across an environmental gradient, and this may be a helpful starting point for exploring a more explicit link between niche conservatism and richness variation. It is very likely that alternative balances between the relative importance of processes more related to geographical patterns of diversification and extinction or niche conservatism are found for different groups of organisms in different parts of the world, but still better methods to quantify this balance need to be developed.

The correlative approach provides a straightforward framework to test hypotheses when they are derived from conceptual predictions, but there are pitfalls that require caution. Correlative models do not allow the dissection of multiple causal links when the surrogate factors (i.e., regression predictors) are collinear, or when multiple factors drive patterns at the same time but at different geographical scales. Thus, it is difficult to integrate local (biotic interactions) with regional (climatic) factors into correlative models (but see Gotteli *et al.*, 2009) or to disentangle the role of each factor in driving species richness on a broad spatial scale.

A fundamentally different class of models for broad scale latitudinal gradients has been developed during recent years. These models are based on the theoretical understanding of the effect of environmental variables on organismal metabolism and physiology, or on population dynamics (e.g., Allen *et al.*, 2002; Field *et al.*, 2005; O'Brien, 2006). These models are important from an epistemological view-point, as they make reasonably precise predictions about the relationship between broad scale patterns of species richness and temperature, and therefore provide hypotheses that are falsifiable from a Popperian perspective (but see Hawkins *et al.*, 2007b). However, as explained below, models of this class use the same statistical background, and therefore are still considered "correlative".

Some of these pioneering models, such as Interim General Models (IGMs) applied to plant diversity, try to understand, based on overall knowledge of plant physiological and ecological responses, which variables (mainly in the sense of water balance and habitat heterogeneity) account for broad scale patterns of species richness. O'Brien (1993, 1998) developed a mechanistic model (the "biological relativity to water–energy dynamics" model) in which there is a linear relationship between richness and rainfall, and a parabolic relationship between richness and insolation. Field *et al.* (2005) developed and compared a series of IGMs by fitting empirical models under the same physiological–ecological reasoning at a local or regional spatial scale and then projected these models to a global scale, showing a good match with macroecological models developed for plants (Currie and Paquin, 1987).

One of the most widely discussed correlative models derived from "first biological principles" in the context of broad scale richness patterns is based on what has been called the Metabolic Theory of Ecology (see Brown *et al.*, 2004 for a summary, and Hawkins *et al.*, 2007a,b for a critique). Allen *et al.* (2002) proposed that richness (S) is a function of temperature:

$$\ln(S) = -(E/1000k)(1000/T) + C$$

where C is a constant, E is the activation energy (0.78 eV), k is Boltzmann's constant (8.62×10^{-5} eV K^{-1}), and T is temperature measured in Kelvin units. This model predicts that richness, on a logarithmic scale, decreases 9 species as $1000/T$ increases (see Brown *et al.*, 2004 for a different parameterization). Allen *et al.* (2006) further expanded this model and tried to merge it conceptually with neutral evolutionary dynamics, using similar analytical predictions for the relationship between evolutionary rates and temperature. However, this more complex version of the model is difficult to test empirically due to lack of data. Hawkins *et al.* (2007a) tested the simple model of Allen *et al.* (2002) using several macroecological datasets and showed a poor fit in almost all cases, opening a wide-ranging discussion about the predictive and explanatory generality of such a model (Hawkins *et al.*, 2007b; Latimer, 2007; Gilloly and Allen, 2007). However, even if generality and explanatory power of the model is dismissed, there may still be important lessons to be learned before a full rejection of the metabolic theory as applied to latitudinal gradients. In fact, Cassemiro and Diniz-Filho (2010) showed that the model of Allen *et al.* (2006) is highly sensitive to model assumptions, and when they are met, and other factors driving gradients are ruled out, model fit may improve significantly. Obviously, this finding shows that the model may not be as general and powerful as originally proposed by Gilloly and Allen (2007), but clearly the model still provides important clues on how and why species respond to environment.

Lessons learned from correlative models suggest that it is reasonable to assume that patterns of biodiversity are driven by a wide array of factors, which interact across spatial scales, including speciation, adaptation, dispersal, species interactions, extinction, and habitat heterogeneity. The model of Allen *et al.* (2002) has many strong assumptions on how biogeographical patterns arise and are maintained, and to a

great extent these assumptions are the reason that the model is unable to predict empirical patterns. Typically, analytical models for massively complex natural systems have a limited utility and applicability, even for specific cases (Wiegand *et al.*, 2003). In addition, correlative models are typically unable to deal with the effects of interactions among multiple ecological mechanisms and to deal with contingencies or stochasticity throughout the evolution of species and niches. This leads us to our last class of models, those based on computer simulations.

Stochastic Simulation Models

Ecologists (Peck, 2004; Grimm *et al.*, 2005) have recently followed philosophers of science (Winsberg, 1999, 2001, 2003), physicists, climate modelers, and others in arguing that experimental methods need not be limited to concrete systems but can be legitimately applied to computer simulations of reality that have been carefully constructed to reflect known, underlying processes. Experiments on simulated systems, just like experiments on real ones, are constantly in dialog with theory and abstract models on the design side, and with statistics on the results side.

Based on the above epistemological reasoning, it is now possible to take advantage of high-performance computational capacity to investigate richness patterns by carrying out virtual experiments on simulated systems. This approach allows the designing and building a sequence of increasingly realistic, spatially and temporally explicit computer models of biogeographical patterns based on well-understood evolutionary and ecological processes, including species origination, adaptation of niches to environments, expansion and contraction of geographic ranges, and extinction. By design, these models are intended to be simple enough to be understood and controlled, but realistic enough to support meaningful simulated experiments and test hypothesized biological mechanisms that are thought to drive patterns of species richness. In the most realistic models, simulated species populate maps of real regions and continents and the resulting patterns of richness can be directly compared with corresponding patterns based on high quality biogeographical data, to discover which model variations come closest to matching nature.

The application of simulation approaches to macroecological analysis of richness patterns was triggered by the discussion of the null models appropriate for testing such patterns, within a more statistically oriented conceptual view. Colwell and Lees (2000) argued that the geographical ranges of species are geometrically constrained by hard boundaries within a region (i.e., a domain), forcing an increasing overlap of the geographical ranges towards the center and, consequently, a mid-domain peak of species richness. They originally suggested that this Mid-Domain Effect (MDE) could be treated as a null model for gradients in species richness with respect to environmental factors, as the model assumes that species' geographical ranges are randomly placed over the domain in the absence of (or despite the presence of) environmental gradients. Perhaps needless to say, the most used implementation of such models involves computer simulations (e.g., Colwell and Hurtt, 1994), although some analytical versions of MDE model have also been developed (Bokma *et al.*, 2001; Connolly, 2005). Of course, defining the correct null expectation for the gradients of species richness in the absence of environmental factors, against which correlative and mechanistic models could be contrasted, is an important leap forward in the study of broad scale geographical patterns. Later, Colwell *et al.* (2004) suggested that geometric constraints imposed by hard boundaries was not only a null model strategy to study spatial patterns of species richness in the absence of any effect of environmental drivers, but could also be treated as an additional factor directly influencing spatial patterns in species richness, and therefore should be evaluated as a competing or contributory explanation for those patterns.

Jetz and Rahbek (2001) designed an ingenious strategy to generate expected species distributions in two-dimensional maps in the absence of the influence of environmental factors. The algorithm is analogous to placing a drop of dye onto a bounded and homogeneous surface. The dye naturally spreads towards all directions within the surface, but the spreading of the dye is limited by the boundaries of the surface. The so-called "spreading dye" model starts by listing species that are present in a given discrete domain, noting their respective range size (number of occupied grid cells for each species). Then, for each and every species, a "starting" position is chosen randomly and independently among all possible grid cells of the domain. The simulated distribution for a species is created by sequentially expanding species distribution towards unoccupied grid cells that are adjacent to those already occupied by the species, until the predefined range size is reached. The process iterates until all species are "spread" in the domain, and the entire procedure is then replicated to estimate the expected mean and variance of species richness per cell. This algorithm has been widely used by macroecologists, and has also served as the basis for more complex and advanced models that incorporate multiple interacting mechanisms, such as evolutionary niche dynamics, speciation, extinction, and range fragmentation. In their original work, Jetz and Rahbek (2001) showed that the pattern of species richness expected under the spreading dye model is somewhat similar to the observed pattern of African bird richness.

As is usual for the most innovative ideas, the MDE sparked a vigorous debate among biogeographers. The validity of MDE as a model for gradients of species richness was challenged on both empirical and conceptual grounds (Hawkins and Diniz-Filho, 2002; Zapata *et al.*, 2003), followed by a point-by-point response to arguments against the MDE (Colwell *et al.*, 2004, 2005; Jetz and Rahbek, 2001). Regardless of the best design for a model that is "null" with respect to the influence of environmental factors, and of the validity of geometric constraints as a significant driver of biogeographical patterns, biogeographers realized that a simulation framework was a useful approach to be included in their research toolboxes (Gotelli *et al.*, 2009).

Pushing forward the simulation approach, many authors developed new, spatially explicit stochastic models to explain spatial patterns in species richness at continental scales. As a consequence of a growing emphasis on exploring the mechanisms behind species distributions, while taking account of

constraints, contingencies, and chance events, these authors pioneered the incorporation of environmental factors known to affect species richness, along with stochastic effects, such as dispersal and centers of origin. The core idea behind such models is that, rather than correlating species richness with environmental or evolutionary factors directly, these factors are incorporated into the simulation framework following the hypothesized mechanism. In addition, because species distributions are explicitly simulated, the model yields predictions in units of species richness per location. The spatial pattern in species richness predicted (expected) by simulation models can be compared against empirical (observed) patterns found in nature as a measure of model fit. Comparing explanatory power among different simulation models is a straightforward statistical task, especially when compared with correlative and curve-fitting model strategies (see Gotelli *et al.*, 2009 for a review).

In an early contribution to this literature, Bokma *et al.* (2001) designed a simulation model to study how species would be distributed over irregularly shaped domains under a purely stochastic process. Their study promoted a significant conceptual advance in biogeographical models by incorporating several spatially explicit mechanisms that are known to affect patterns of biodiversity, such as speciation, migration, and extinction. They found that when these mechanisms occur at random over space, the predicted pattern of species richness must also be random. However, if species migration is constrained by space, and speciation and extinction are also spatially patterned, their model predicts that the center of major landmasses should have higher species richness. However, the model predictive power of real-world pattern of species richness of New World mammals was always very poor, regardless of how the biogeographicl mechanisms incorporated in the model were parameterized.

To explore the role of phylogenesis in range-based models on a bounded domain, Davies *et al.* (2005) developed a simple, stochastic model of the origination and spread of a clade in two-dimensional space (a 100×100 grid). They showed that the spatial pattern of richness across the domain (once 500 species have arisen) depends crucially on (1) phylogenetic conservatism of range location (ranges either split perapatrically, *in situ*, or split and one of the two daughter species is then placed at random within the domain) and (2) asymmetry (imbalance) of phylogenetic trees (parameterized by setting the degree of heritability of speciation rate). With random placement of daughter ranges, with or without tree asymmetry, a classic mid-domain peak of species richness (MDE pattern) arises, as expected. However, when range location is conserved, the MDE pattern is somewhat weakened. Finally, when tree asymmetry is allowed and range location is conserved, the MDE pattern is no longer evident and location of richness peaks becomes increasingly idiosyncratic, suggesting that the MDE is an appropriate ecological null model only when phylogenetic influence on range location is known to be low or nonexistent.

Taking a different perspective, and explicitly modeling the effect of environment on the distribution of species, Rangel and Diniz-Filho (2005a) designed a spatially explicit simulation to model speciation, range expansion, extinction, and species richness on a simple environmental gradient. In their model, the one-dimensional domain is a bounded grid, which is divided into equal-sized cells. Along this transect, a linear environmental gradient is established by assigning a "suitability" value to each cell, with suitability increasing linearly from one end of the transect to the other. The difference in suitability between the least and most suitable cells, at the ends of the spatial domain, determines the steepness of the gradient, and defines the limits of the single axis of niche space.

In the model of Rangel and Diniz-Filho (2005a), an initial species is seeded at a location (cell) along the gradient that is chosen as a stochastic function of environmental suitability, so that a cell toward the more suitable end of the gradient is more likely to be chosen than a cell towards the less suitable end. The value of environmental suitability of the seed cell is defined as the niche center of the species, and a symmetrical niche deviation around the niche center (environmental tolerance) is stochastically defined from a normal distribution. The range of values between maximum and minimum environmental tolerances determines the range of environmental conditions that can be tolerated by the species (its niche breadth), and thus, the boundaries of its niche space. Once the niche is defined, it is projected back onto the geographic space, and the species spreads its geographic range into all suitable cells (all map cells inside its niche, in niche space). The model algorithm iterates in time, and at each step in time, each existing species may speciate (with a specified speciation probability). In the event of speciation, a cell at the periphery of the parent species' range is chosen as the seed cell for a daughter species, as a stochastic function of environmental suitability, so that a cell towards the more suitable end of the range is more likely to be chosen than a cell towards the less suitable end. The environmental condition of the cell where the daughter species arises becomes its niche center, a new niche breadth is sampled from the same normal distribution, and the new species colonizes all adjacent cells in which the environmental factor lies within its niche.

As opposed to the conventional prediction that the peak of species richness should always occupy the most "suitable" region of the domain, in the model of Rangel and Diniz-Filho (2005a) the pattern of species richness depends crucially on the strength of the environmental gradient. In their model, a strong and steep environmental gradient forces species to have small ranges, given their environmental tolerances, and the evolutionary dynamics promotes higher extinction in the less suitable region of the map, therefore promoting high species richness in the region of highest environmental suitability. However, as the strength of the environmental gradient weakens, the average and variance of species range size become larger, and thus species increasingly overlap in the middle of the domain as a consequence of geometric constraints on the range location. The richness peak gradually shifts toward the center of the domain as environment weakens, until, at the limit, the MDE disappears entirely with the absence of environmental gradient, when the model predicts flat pattern of species richness. The model of Rangel and Diniz-Filho (2005a) provides support for the view of geometric constraints as a mechanism driving species richness in relatively uniform environment, with well-defined boundaries.

Storch *et al.* (2006) and Rahbek *et al.* (2007) adapted the spreading dye model to incorporate the influence of environmental factors on the distribution of species. These authors realized that the bounded surface in which the dye spreads is never homogeneous. In fact, the real-world variation of environmental factors across space, as well as the interaction among these factors at different spatial scales, creates habitats that differ with respect to suitability. According to the algorithm independently designed by those authors, species should have higher probability of expanding their distribution towards regions with "more suitable" environment, including warmer and wetter areas known to harbor higher species diversity. In addition, they also relaxed the assumption of range cohesion, and evaluated how different degrees of long distance dispersal could affect the predictive power of those models. Storch *et al.* (2006) evaluated how such a model could predict global patterns of avian diversity, and found that best predictions are achieved when the probability of a range spreading into an area is directly proportional to the AET of the area. Rahbek *et al.* (2007) studied the pattern of species richness of birds in South America, and also showed that model predictive power increases when (i) water–energy environmental factors affect species range expansion, (ii) range cohesion is imposed in the model algorithm, and (iii) only large-ranged species are considered. Perhaps the single most important conceptual advance in those models was moving beyond the debate on the best null model for spatial patterns in species richness by directly incorporating ecological "mechanisms" into the model algorithm (Gotelli *et al.*, 2009).

A natural next step for simulation models in biogeography was to add more realism, and more complexity, by replacing the static environmental gradient of the previous models by a heterogeneous, multivariate, and fluctuating environment. Of course, the complexity of species' niches must also be incorporated, by adding the corresponding dimensions of species environmental niche, which are continuously responding to evolutionary forces and interacting directly with the fluctuating environment. As in previous models, Rangel *et al.* (2007) treated geographic space as defined as a spatially bounded set of square grid cells on a plane. Each cell was characterized by an array of environmental variables, which collectively defined a spatially heterogeneous pattern within the geographic space. In this model, a "seed species" colonizes a randomly chosen grid cell on the map, and the environment in that cell defines the niche center of the seed species in niche space. The species is assigned a niche breadth, stochastically from a normal distribution, separately for each environmental niche axis, and the species' geographical range is then defined by projecting its niche volume back into the geographic space to fill all contiguous cells around the seed cell that map inside the niche volume, so that ranges are always in equilibrium with environment.

In this simulation (**Figure 3**), the environment is not constant through time, but fluctuates sinusoidally. Because each species' geographic distribution is an interactive consequence between its niche (defined by tolerance range for each environmental variable in niche space) and the geographic space (actual environmental conditions on the map, which change through time), environmental fluctuations may cause contraction, expansion, reshaping, or fragmentation of the species' geographic distribution. Range fragmentation may yield vicariant speciation: each isolated "population" (range fragment) may go extinct as function of its size, or it may speciate and become an independent new species. Each incipient daughter species inherits its parent species' niche center, but with a specified degree of adaptive evolution permitted. Adaptation moves the niche, through niche space, toward the average environmental conditions of the daughter species' geographic distribution, which in general will differ from the average environment of its parent species. Adaptive shifts in daughter species' niches in niche space are projected back onto the geographic space to reshape ranges accordingly. The niche breadth of the daughter species is also inherited from the parent species, although the daughter species' niche breadth can be stochastically broadened or narrowed. In this model, niche conservatism is simulated when a daughter species inherits parent species' niche center and niche breadth without much change, whereas niche evolution implies adaptation of the daughter species' niche center towards its own local environmental conditions, and increase or decrease of niche breadth (**Figure 3**).

Rangel *et al.* (2007) used the model to show how the interplay between evolutionary niche shifts and fluctuating environment drives broad scale pattern in species richness. When model parameters are set to strong niche conservatism, species cannot readily adapt to local environmental conditions, and thus they accumulate in regions that are environmentally similar to those occupied by their ancestors. Because of the spatial autocorrelation in the environmental factors, the geographic region of highest diversification rates tends to be closer to the location where the lineage arose. In contrast, if model parameters allow niche evolution during speciation events, species niches readily evolve through niche space, allowing descendent species to exploit environmental conditions that were unsuitable for their ancestors. In this case, the geographic distributions of descendant species tend to be different in size, shape, and location compared to the distributions of their ancestors. Of course, this model shows that niche evolutionary dynamics and center of origin strongly affect spatial patterns in species richness if the temporal dimension is included in the model. Rangel *et al.* (2007) designed an empirical test of the model by contrasting the model predictions against real-world patterns of species richness of South American birds, and they found that the model has very high predictive power when center of origin is tropical and niche conservatism is strong. This was the first attempt to use an explicit evolutionary model to investigate how species richness patterns are influenced by niche conservatism (Wiens and Donoghue, 2004). In fact, their simulation allowed evaluation of the roles of both evolutionary and ecological processes, revealing the great potential of the link between ecology and historical biogeography under integrated theoretical and methodological frameworks.

Recently, Colwell and Rangel (2010) applied a similar stochastic simulation model to study how quaternary glacial–interglacial temperature cycles affected contemporary geographical and phylogenetic patterns of species richness along tropical elevational gradients. In their model, key ecological and evolutionary mechanisms are explicitly simulated,

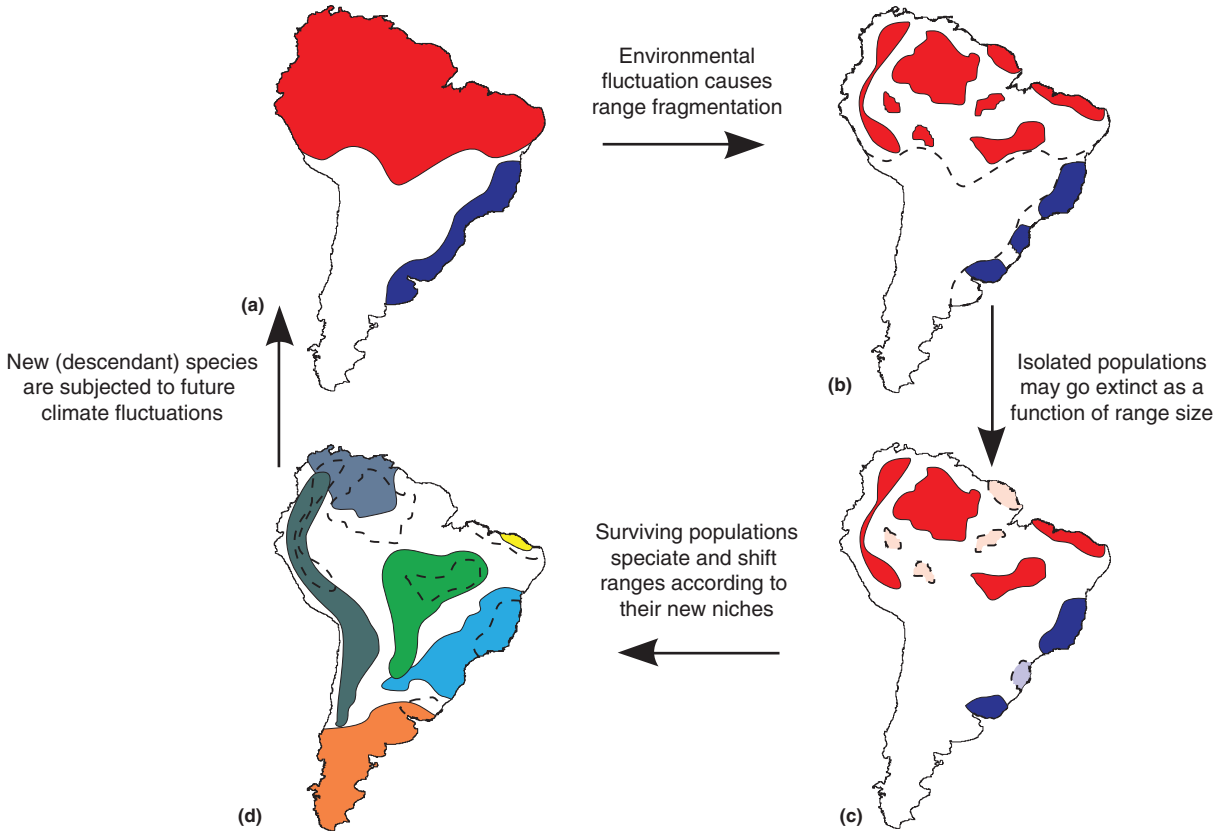


Figure 3 The stochastic model proposed by Rangel *et al.* (2007), in which species are subject to range expansion, displacement, fragmentation, extinction, speciation, and niche evolution in a spatially complex and fluctuating environment (South America). (a) Geographical ranges of two hypothetical, ancestral species (red and blue areas), defined by their respective environmental niches, projected into the geographic space. The environmental factors that characterize the map are subjected to fluctuation, which causes (b) range fragmentation (red and blue areas, solid outlines), with variable fragment size. (c) Smaller fragments go extinct as a stochastic function of their geographical extent (light red and blue areas, dashed outlines), whereas larger fragments survive (dark red and blue areas, solid outlines) as the initial distributions of daughter species. (d) After speciation, evolution of the niche of each new species is modeled in niche space (not shown), with a specified degree of niche conservatism. The geographical distribution of each descendant species is a function of its evolved niche, adjusting the size and placement of its geographical range (original fragments are shown as white areas, dashed outlines) to colonize contiguous areas in the geographical space (colored areas, solid lines) that correspond to its new ecological niche.

including speciation, anagenetic evolution, niche conservatism, range shifts, and extinction. Their models suggested that elevational patterns of species richness arise from the differential survival of founder lineages, consolidated by speciation and the inheritance of thermal niche characteristics. The model yielded a surprisingly rich variety of realistic patterns of phylogenetic and species richness, including some that closely match contemporary elevational richness profiles from an elevational transect in Costa Rica, which were chosen for the wide variety of elevational richness patterns that they represent.

Directions and Future Work

Biogeographical models that strictly rely on the correlative approach are problematic because most such models do not make quantitative predictions. In fact, the application of these models to the study of biogeographical patterns usually leads no further, once the researcher has qualitatively described a good match between spatial patterns in species richness and some measures of contemporary climate. Owing to the lack of *a priori* expectations, parameters estimated by such models are commonly left with little or no biological interpretation. The dynamic nature of evolutionary processes through time is usually left out of the model, although no biogeographer would reject historical factors, such as climate fluctuation, speciation, and extinction as important drivers of the biogeography of living organisms. As noted by Currie *et al.* (2004), most correlative models are not founded on solid theoretical principles of ecology and evolutionary biology, nor do they incorporate biogeographical mechanisms thought to affect the distribution of species. In fact, they are so vaguely specified that they do not predict anything more precise than a high correlation between species richness and measures of contemporary climate. However, for practical reasons, it will be difficult to totally abandon correlative models, so efforts must be made to better define surrogates for the multiple hypotheses, especially for dealing with evolutionarily dynamic mechanisms, such as niche conservatism and geographically

structured diversification rates. To employ correlative models to test niche conservatism as a mechanism for patterns of species richness, it is important to incorporate in the analysis the phenotypic traits that may be linked to the emergence and maintenance of these patterns, which may be under stabilizing selection processes that constrain the dynamics of geographic ranges through time. Of course, when the evolutions of species traits are taken into account in biogeographical studies, the full scope of analytical methods from phylogenetic comparative analysis should be considered as the first resource (e.g., Hof *et al.*, 2010). It is also important to implement better designs for correlative studies of patterns in species richness, especially with regard to multiple geographic scales, both in terms of grain size and geographical extent.

Correlative models that are designed based on theoretical first principles provide quantitative predictions that can be contrasted against real-world patterns. Models of this category are still in their infancy, as the biological mechanisms at lower organizational scales (from molecules to populations) that affect broad scale ecological patterns are not well understood. However, such models may nonetheless bring a dramatic conceptual improvement over correlative models, pushing biogeography towards a more precise definition of patterns, mechanisms, and scales. As pointed out by Cassemiro and Diniz-Filho (2010), there is still room for improving the models designed under the Metabolic Theory of Ecology that are applied to richness gradients, especially regarding the incorporation of more complex macroecological patterns (i.e., geographic variation in body size and abundance) to avoid critical and potentially unrealistic assumptions, such as energy equivalence in ecological communities. The challenge is, obviously, to derive a mathematical prediction of how these additional factors act, in concert and interacting with other mechanisms, to drive ecological patterns from the bottom up.

In contrast with correlative modeling approach, simulation modeling has the potential to explicitly incorporate the processes believed to be affecting the geographical ranges of species. In principle, simulation models can be designed to test specific ecological mechanisms. However, general simulation models are usually preferred because they incorporate multiple interacting mechanisms, and can be both temporally and spatially explicit. Of course, the more the mechanisms incorporated in the model, the more complex the model becomes, and therefore testing and validating the model may become cumbersome. However, the cost of the additional complexity in the model could be balanced, in principle, by a wide array of quantitative predictions that can be compared against empirical patterns during model validation, providing a more robust evaluation of model predictive power. No doubt, the use of simulation models in biogeography is still in its infancy, and so are the statistical techniques available for model test and validation. Gotelli *et al.* (2009) have proposed a measure of goodness-of-fit, along with a corresponding hypothesis test, to evaluate the predictive power of simulation models with regard to patterns of species richness. However, simulation models in biogeography are evolving fast, and models are now extending their predictions to the new aspects of biodiversity patterns, such as the evolution of functional trait diversity (Stegen and Hurlbert, 2011), which thus requires better empirical data and the development of additional statistical techniques to measure model predictive power.

Levins (1966), in his influential paper on "the strategy of model building in population biology," pointed out that "the difference between legitimate and illegitimate (model) simplifications depends not only on the reality to be described but also on the state of the science." Although biogeography is not a young science (Lomolino *et al.*, 2009), the complex and heterogeneous nature of the phenomena that biogeographers struggle to understand slows down the pace of conceptual progress, as well as the development of models with higher predictive power. Still, the strategies of model building laid out by Richard Levins 45 years ago are also applicable to biogeography. According to him, "it is of course desirable to work with manageable models which maximize generality, realism, and precision toward the overlapping but not identical goals of understanding, predicting, and modifying nature". The balance and emphasis on each of these three goals is reflected in the modeling approaches used by biogeographers:

1. Sacrificing precision and realism for the sake of generality. Within this category are the correlative models that are empirically oriented, which are mainly concerned with qualitative rather than quantitative results. Hypothesis testing or theoretical interpretation of estimated parameters is usually not the focus of studies that employ this approach. Instead, the form and direction of the relationship between species richness and each environmental factor is described. Predictions can rarely be drawn based on these models.
2. Sacrificing realism and generality for the sake of precision. Correlative models that are derived from ecological principles fall within this category. These models yield precise predictions on how species richness is affected by the hypothesized factor. One of the main advantages of these models is that their predictions can be tested and validated using robust statistical methods, and if the model is considered reliable enough, it can also be used to project species richness into an area outside the range of environmental values. Derived correlative models vary with regard to the level of generality they incorporate, but certainly the more general they are, the less precise they become.
3. Sacrificing generality and precision for the sake of realism. Stochastic simulation models are arguably the most "mechanistic" and "realist" models used by biogeographers, but their predictions are less precise and usually more general than correlative models. One of the main advantages of the simulation approach is to make explicit assumptions about how the processes believed to be affecting the distribution of species operate. The drawback of this approach to modeling is that the realism comes at the cost of additional free parameters, which are not straightforward to quantify or parameterize. Also, documenting and communicating the entire model algorithm is not a trivial task, a challenge that has motivated some authors to suggest a standard protocol for describing agent-based models (Grimm *et al.*, 2005).

Finding the optimal balance between model simplicity and complexity is not a trivial task, but should be considered as a central problem in biogeographical modeling. The principle of parsimony suggests that when choosing a modeling approach, a simpler model should always be preferred, unless simplicity can be traded for increased explanatory power. On one hand, a model that is too simple tends to ignore important mechanisms known to affect the real system, which would limit its potential to provide scientific understanding and testable predictions. On the other hand, despite the higher predictive power of complex models, the analysis of a model that is too complex tends to be cumbersome, potentially risking the advance of understanding of the problem it addresses. The famous quote from George E. P. Box is particularly true for biogeographical models: "essentially, all models are wrong, but some are useful".

See also: Biodiversity, Evolution and. Global Species Richness. Latitudinal Gradients of Biodiversity. Species Distribution Modeling. Species Diversity, Overview

References

Algar AC, Kerr JT, and Currie DJ (2009) Evolutionary constraints on regional faunas: Whom, but not how many. *Ecology Letters* 12: 57–65.

Allen AP, Brown JH, and Gillooly JF (2002) Global biodiversity, biochemical kinetics and the energy equivalence rule. *Science* 297: 1545–1548.

Allen AP, Gillooly JF, Savage Van M, and Brown JH (2006) Kinetic effects of temperature on rates of genetic divergence and speciation. *Proceedings of the National Academy of Sciences of the United States of America* 103: 9130–9135.

Araújo MB, Nogués-Bravo D, Diniz-Filho JAF, Haywood AM, Valdes PJ, and Rahbek C (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography* 31: 8–15.

Bokma F, Bokma J, and Monkkonen M (2001) Random processes and geographic species richness patterns: Why so few species in the north? *Ecography* 24: 43–49.

Brown JH (1995) *Macroecology*. Chicago: University of Chicago Press.

Brown JH, Gillooly JF, Allen AP, Savage, Van M, and West GB (2004) Toward a metabolic theory of ecology. *Ecology* 1771–1789.

Cassemiro FAS and Diniz-Filho JAF (2010) Deviations from predictions of the metabolic theory of ecology can be explained by violations of assumptions. *Ecology* 91: 3729–3738.

Colwell RK and Hurtt GC (1994) Nonbiological gradients in species richness and a spurious rapoport effect. *American Naturalist* 144: 570–595.

Colwell RK and Lees DC (2000) The mid-domain effect: Geometric constraints on the geography of species richness. *Trends in Ecology and Evolution* 15: 70–76.

Colwell RK, Rahbek C, and Gotelli NJ (2004) The mid-domain effect and species richness patterns: What have we learned so far? *American Naturalist* 163. E-Article.

Colwell RK, Rahbek C, and Gotelli NJ (2005) The mid-domain effect: there's a baby in the bathwater. *American Naturalist* 166: E149–E154.

Colwell RK and Rangel TF (2009) Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences of the United States of America* 106: 19651–19658.

Colwell RK and Rangel TF (2010) A stochastic, evolutionary model for range shifts and richness on tropical elevational gradients under Quaternary glacial cycles. *Philosophical Transactions of the Royal Society B* 365: 3695–3707.

Connolly SR (2005) Process-based models of species distributions and mid-domain effect. *American Naturalist* 166: 1–11.

Currie DJ, Mittelbach GG, Cornell HV, *et al.* (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters* 7: 1121–1134.

Currie DJ and Paquin V (1987) Large-scale biogeographical patterns of species richness of trees. *Nature* 329: 326–327.

Davies TJ, Buckley LB, Grenyer R, and Gittleman JL (2011) The influence of past and present climate on the biogeography of modern mammal diversity. *Philosophical Transactions of the Royal Society B* 366: 2526–2535.

Davies TJ, Grenyer R, and Gittleman JL (2005) Phylogeny can make the mid-domain effect an inappropriate null model. *Biology Letters* 1: 143–146.

Diniz-Filho JAF, Rangel TFLVB, Bini LM, and Hawkins BA (2007) Macroevolutionary dynamics in environmental space and latitudinal diversity gradient in New World birds. *Proceedings of the Royal Society B* 274: 43–52.

Dynesius M and Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences of the United States of America* 97: 9115–9120.

Elith J and Graham CH (2009) Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32: 66–77.

Elith J, Graham CH, Anderson RP, *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.

Field R, Hawkins BA, Cornell HV, *et al.* (2009) Spatial species-richness gradients across scales: A meta-analysis. *Journal of Biogeography* 36: 132–147.

Field R, O'Brien EM, and Whittaker RJ (2005) Global models for predicting woody plant richness from climate: Development and evaluation. *Ecology* 86: 2263–2277.

Fischer AG (1960) Latitudinal Variation in Organic Diversity. *Evolution* 14: 64–81.

Gaston K and Blackburn T (2000) *Patterns and Process in Macroecology*. Wiley-Blackwell.

Gillooly JF and Allen AP (2007) Linking global patterns in biodiversity to evoluation dynamics using metabolic theory. *Ecology* 88: 1890–1894.

Gotelli NJ, Anderson MJ, Arita HT, *et al.* (2009) Patterns and causes of species richness: A general simulation model for macroecology. *Ecology Letters* 12: 873–886.

Grimm V, Revilla E, Berger U, *et al.* (2005) Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science* 310: 987–991.

Guisan A and Rahbek C (2011) SESAM – A new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography* 28: 1433–1444.

Guisan A and Thuiller W (2005) Predicting species distributions: Offering more than simple habitat models. *Ecology Letters* 8: 993–1009.

Hawkins BA (2001) Ecology's oldest pattern? *Endeavour* 25: 133.

Hawkins BA, Albuquerque FS, Araújo MB, *et al.* (2007a) A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. *Ecology* 88: 1877–1888.

Hawkins BA, Diniz-Filho JAF, Bini LM, *et al.* (2007b) Metabolic theory and diversity gradients: Where do we go from here? *Ecology* 88: 1898–1902.

Hawkins BA and Diniz-Filho JAF (2002) The mid-domain effect cannot explain the diversity gradient of Nearctic birds. *Global Ecology and Biogeography* 11: 419–426.

Hawkins BA, Field R, Cornell HV, *et al.* (2003a) Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84: 3105–3117.

Hawkins BA and Porter EE (2003) Relative influences of current and historical factors on mammal and bird diversity patterns in deglaciated North America. *Global Ecology and Biogeography* 12: 475–481.

Hawkins BA, Porter EE, and Diniz-Filho JAF (2003b) Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. *Ecology* 84: 1608–1623.

Hof C, Rahbek C, and Araújo MB (2010) Phylogenetic signal in the climatic niches of the world's amphibians. *Ecography* 33: 242–250.

Hortal J, Diniz-Filho JAF, Bini LM, *et al.* (2011) Ice age climate, evolutionary constraints and diversity patterns of European dung beetles. *Ecology Letters* 14: 741–748.

Jetz W and Rahbek C (2001) Geometric constraints explain much of the species richness patterns in African birds. *Proceedings of the National Academy of Sciences of the United States of America* 98: 5661–5666.

Latimer AM (2007) Geography and resource limitation complicate metabolism-base predictions of species richness. *Ecology* 88: 1895–1898.

Lees DC, Kremen C, and Andriamampianina L (1999) A null model for species richness gradients: Bounded range overlap of butterflies and other rainforest endemics in Madagascar. *Biological Journal of Linnean Society* 67: 529–584.

Levins R (1966) The strategy of model building in population biology. *American Scientist* 54: 421–431.

Lomolino MV, Riddle BR, and Brown JH (2009) *Biogeography*. Sunderland, MA: Sinauer.

O'Brien EM (1993) Climatic gradients in woody plant species richness: Towards an explanation based on an analysis of southern Africa's woody flora. *Journal of Biogeography* 20: 181–198.
O'Brien EM (1998) Water–energy dynamics, climate, and prediction of woody plant species richness: An interim general model. *Journal of Biogeography* 25: 379–398.
O'Brien EM (2006) Biological relativity to water–energy dynamics. *Journal of Biogeography* 33: 1868–1888.
Peck SL (2004) Simulation as experiment: A philosophical reassessment for biological modeling. *Trends in Ecology and Evolution* 19: 530–534.
Peterson A (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics* 3: 59–72.
Pianka ER (1966) Latitudinal gradients in species diversity: A review of concepts. *American Naturalist* 100: 33–46.
Rahbek C and Graves GR (2001) Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences of the United States of America* 98: 4534–4539.
Rahbek C, Gotelli NJ, Colwell RK, Entsminger GL, Rangel TFLVB, and Graves GR (2007) Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society B* 274: 165–174.
Rangel TFLVB and Diniz-Filho JAF (2005a) An evolutionary tolerance model explaining spatial patterns in species richness under environmental gradients and geometric constraints. *Ecography* 28: 253–263.
Rangel TFLVB, Diniz-Filho JAF, and Colwell RK (2007) Species richness and evolutionary niche dynamics: A spatial pattern-oriented simulation experiment. *American Naturalist* 170: 602–616.
Rhode K (1992) Latitudinal gradients in species diversity: The search for the primary cause. *Oikos* 65: 514–527.
Ricklefs RE (2006) Evolutionary diversification and the origin of the diversity–environment relationship. *Ecology* 87: S3–S13.
Rodriguez MA, Belmontes JA, and Hawkins BA (2005) Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. *Acta Oecologica* 28: 65–70.
Simpson GG (1964) Species Density of North American Recent Mammals. *Systematic Zoology* 13: 57–73.
Soberón J (2007) Grinnellian and Eltonian niches and geographic distribution of species. *Ecology Letters* 10: 1115–1123.
Stegen JC and Hurlbert AH (2011) Inferring ecological processes from taxonomic, phylogenetic and functional trait β-diversity. *PLOS* One 6: e209006.
Storch D, Davies RG, Zajícek S, *et al.* (2006) Energy, range dynamics and global species richness patterns: Reconciling mid-domain effects and environmental determinants of avail diversity. *Ecology Letters* 9: 1308–1320.
Weir JT and Schluter D (2006) The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* 315: 1574–1576.
Wiegand T, Jeltsch F, Hanski I, and Grimm V (2003) Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. *Oikos* 100: 209–222.
Wiens JJ, Ackerly DD, Allen AP, *et al.* (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters* 13: 1310–1324.
Wiens JJ and Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution* 19: 639–644.
Wiens JJ, Graham CH, Moen DS, Smith SA, and Reeder TW (2006) Evolutionary and ecological causes of the latitudinal diversity gradient in Hylod frogs: Treefrog trees unearth the roots of high tropical diversity. *American Naturalist* 168: 579–596.
Willig MR, Kaufman DM, and Stevens RD (2003) Latitudinal gradients of biodiversity: Patterns, process, scale and synthesis. *Annual Reviews of Ecology, Evolution and Systematics* 34: 273–309.
Whittaker RJ, Willis KJ, and Field R (2001) Scale and species richness: Towards a general, hierarchical theory of species diversity. *Journal of Biogeography* 28: 453–470.
Winsberg E (1999) Sanctioning models: The epistemology of a simulation. *Science in Context* 12: 247–260.
Winsberg E (2001) Simulation, models, and theories: Complex physical systems and their representations. *Philosophy of Science* 68: S442–S454.
Winsberg E (2003) Simulated experiments: Methodology for a virtual world. *Philosophy of Science* 70: 105–125.
Zapata FA, Gaston KJ, and Chown SL (2003) Mid-domain models of species richness gradients: Assumptions, methods and evidence. *Journal of Animal Ecology* 72: 677–690.

Biogeography, Overview

Mark V Lomolino, Oklahoma Natural Heritage Inventory, and University of Oklahoma, Norman, OK, USA

Published by Elsevier Inc.

This article is reproduced from the previous edition, volume 1, pp 455–469, © 2001, Elsevier Inc.

Glossary

Biogeography Study of the geographic variation of nature, including variation in any biological characteristics (e.g., body size, population density, or species richness) on a geographic scale.

Continental drift Model first proposed by Alfred Wegener that states that the continents were once united and then were displaced over the surface of the globe.

Plate tectonics Study of the origin, movement, and destruction of the plates and how these processes have been involved in the evolution of Earth's crust.

Pleistocene Geologic period from 2 million to 10,000 years before the present, which was characterized by alternating periods of glaciation events and global warming.

Species composition Types of species that constitute a given community or sample.

Species richness Number of species in a given community or sample.

Introduction

Traditionally, biogeography has been defined as the study of patterns in distributions of geographic ranges (Brown and Gibson, 1983). During the past three decades, however, this field has experienced a great surge in development and sophistication, and with this development the scope of the field has broadened to include an impressive diversity of patterns. Simply put, modern biogeographers now study nearly all aspects of the "geography of nature." Biogeography now includes studies of variation in any biological feature (genetic, morphological, behavioral, physiological, demographic, or ecological) across geographic dimensions such as distance among sites or along gradients of area, elevation, or depth (see Brown and Lomolino, 1998).

Fundamentals of Biogeography

Despite the sometimes overwhelming complexity of the natural world, all biogeographic patterns ultimately derive from two very general features of nature. First, as we move along any dimension of the geographic template, environmental conditions tend to vary in a predictable manner. For example, more distant sites tend to be more dissimilar than adjacent sites, environments at higher elevations tend to be cooler and wetter than those at lower elevations, and larger areas tend to capture more solar energy and a greater diversity of environmental conditions than smaller areas. Second, all forms of life differ in their abilities to adapt to geographic variation in their environment. These differences, while including a great diversity of responses (e.g., physiological, behavioral, developmental, and evolutionary), ultimately influence the three fundamental processes of biogeography: *immigration*, *extinction*, and *evolution*. All the biogeographic patterns we study derive from nonrandom variation in these processes across geographic gradients and across individuals, populations, and species.

Early History of the Field

Biogeography has a long and distinguished history, and one inextricably woven into the historical development of evolutionary biology and ecology. The historical development of biogeography had its origins coincident with the Age of World Explorations by Europeans during the eighteenth and nineteenth centuries. Yet the study of geography of nature must be an ancient one. The European explorers were not the first to ask "Where did life come from, and how did it diversify and spread across the earth?" Aristotle asked these same questions, as did many others before and after him, when faced with accounts of strange forms of life from foreign lands.

The development of biogeography into a mature and respected field of science, however, required a much better understanding of variation in what we now call the geographic template and the associated variation in the natural world. It is by no minor coincidence that both evolutionary biology and biogeography developed in earnest during the Age of Exploration. Prior to this time, biologists had "discovered" and described less than 1% of plant and animal forms that we know today. Each new voyage or expedition added to the accumulated information on the earth's environments and life-forms, and would eventually provide the raw material for the disciplines of evolution and biogeography. These disciplines are interconnected by the knowledge that selective pressures vary across space, and that all life-forms and their distributions are the product of natural selection.

The early explorers and naturalists did far more than just label and catalog their specimens. They soon, perhaps irresistibly took to the task of comparing their collections across regions, elevations, and other gradients of the geographic template. At the same time, they began to develop explanations for the similarities and differences among the biotas they studied. In fact, most of the persistent themes of the field of biogeography (**Table 1**) were well established during the eighteenth and nineteenth centuries. To be sure, biogeography has made great strides during

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00014-9

Table 1 Persistent themes of biogeography

1. Comparing and classifying geographic regions based on their biotas.
2. Reconstructing the historical development of biotas, including their origin, spread, and diversification.
3. Explaining the differences in numbers as well as types of species among geographic areas.
4. Explaining geographic variation in the characteristics of individuals and populations of closely related species, including trends in morphology, behavior, and demography.

Source: Reproduced with permission from Brown JH and Lomolino MV (1998) *Biogeography*. Sunderland, Massachusetts: Sinauer Associates.

the twentieth century to become a mature and sophisticated science. It is important to acknowledge, however, that we owe a great deal to the many visionary explorers and naturalists who shared our fascination and asked the same questions about the geographic variation of nature.

Historic Explorations of the Eighteenth Century

While motivated to a large degree by the quest for money and power, the Age of European Exploration was also fueled by the call to serve God. It was widely believed that the Creator had placed on this earth a still unknown diversity of organisms—a divine zoo or garden of life. Accordingly, early European explorers believed that there was perhaps no greater way to worship God than to unlock the mysteries of creation. Yet with each new account of some distant biota came information that challenged the prevailing views of creation. Eventually, the growing body of knowledge would overturn the long accepted view that the earth, its climate, and its species were immutable, unchanging in both space and time. More immediately, however, biologists of the eighteenth century were struck by the astounding diversity of species. Such diversity presented them with two serious problems, one practical and the other conceptual. First, biologists urgently needed a systematic and generally accepted scheme for classifying the burgeoning wealth of specimens, one that would reflect the similarities and differences among the species. Second, it quickly became clear that there were just too many species to be carried by Noah's Ark. How could all the forms of life, now adapted to many distant and distinct regions across the globe, have originated and then spread from that one landing point?

Carolus Linnaeus (1707–1778), certainly one of the most prominent biologists of all time, took on both of these challenges. In fact, his system of binomial nomenclature is the system we continue to use today to classify organisms. Linnaeus also attempted to rectify the Biblical doctrine with what he and his contemporaries knew about the diversity and geography of nature. This was especially challenging because, like most of his colleagues, Linnaeus was sure that species were immutable. Given this, how could species adapted to a single site and climate (Noah's landing) have spread and become perfectly adapted to a suite of different environments (e.g., alpine tundra, coniferous forests, lowland forests, and grasslands)? Linnaeus' answer: Noah's landing had occurred along the slopes of Mount Ararat, a high mountain near the border of Turkey and Armenia. This mountain is so tall (reaching 16,853 ft above sea level) that along its slopes could be found a succession of environments and communities ranging from subtropical grasslands at the lower elevations to forests and alpine tundra at its summit. According to Linnaeus' hypothesis, each elevational zone harbored a distinct assemblage of animals, each immutable but perfectly adapted to their local environment. When the Flood finally receded, these animals then dispersed to eventually colonize their respective environments across the globe.

One of the foremost challenges to Linnaeus' views came from his contemporary Comte de Buffon (1707–1788), who believed that not only were climates mutable, but species were as well. How else could animals disperse across what are now inhospitable habitats to occupy their present ranges in such isolated regions of the globe? Buffon's theory of the origin and spread of life stemmed largely from his studies of living and fossil mammals, especially those of the Old and New World tropics. He was the first to realize that different regions of the globe, even those with the same environmental conditions, had distinct biotas. This observation was so fundamental that it eventually became biogeography's first law: Buffon's Law.

Like Linnaeus, Buffon also concluded that there was one "landing point," one site where all animals originated. However, this site, or region, was located far to the north of Mount Ararat, somewhere in the Arctic Circle where the early animals and their descendants could gain ready access to both the Old and New Worlds. This is where these life-forms survived the Flood during some earlier period when the earth's climate was much warmer, warm enough such that tropical environments could extend far poleward. Once the floods receded, animals spread southward into the continents and began to diverge in form as they became increasingly isolated on different landmasses.

Other biologists of the eighteenth century, including Joseph Banks and Johan Reinhold Forster, both of whom served as naturalist on voyages of Captain James Cook, were quick to confirm the generality of Buffon's Law: environmentally similar but isolated regions have distinct assemblages of plants and animals. Forster also discussed the relationship between regional floras and environmental conditions and, in turn, between plant and animal associations: two cornerstones of the field now known as ecology. Forster was also one of the first scientists to report that plant diversity increases as we move toward the equator, that islands have fewer plants than the mainland, and that the diversity of insular plants increases with island size and available resources. Later in the eighteenth century, Karl Wildenow (1765–1812) and one of his students, Alexander von Humboldt (1769–1859), confirmed and further generalized both Buffon's Law and Forster's. Toward the end of the century, Augustin P. de Candolle added the important insight that, not only is the distribution of organisms influenced by geographic variation in environments, but they also compete for limiting resources such as food, light, and water.

Therefore, by the beginning of the nineteenth century, biogeographers already had their first "law," they described and tested the generality of a number of related patterns about the geography of nature, and they offered some testable theories regarding those patterns. They were actively working on at least three of the four persistent themes of biogeography (Themes 1–3 in **Table 1**). A number of biogeographers had abandoned the notion that species and climates were immutable. But for the field to advance and become a mature science, two additional, fundamental

insights were needed. First, it required a mechanism for the mutability and adaptation of species. As many of us realize, Candolle's observations about competition and the struggle for existence were central to the development of the theory of evolution by natural selection. These advances were to come in the latter part of the nineteenth century. Second, scientists had to recognize that the geographic template (i.e., the foundation for all of these patterns) also was mutable. That is, the size and relative positions of the continents and ocean basins have changed throughout the history of our planet.

Advances of the Nineteenth Century

Many of the most fundamental advances in biogeography, and evolutionary biology as well, have required advances in geology. Until the nineteenth century, the age of the earth was typically assumed to be just a few thousand years, way too brief to allow what we now know to be the requisite time for evolution of its plates and the species that have rafted on those plates. The collective work of nineteenth-century paleobiologists would push the age of the earth back hundreds of thousands and eventually millions of years before the present. Legendary geologists and paleobiologists such as George Lyell (1797–1875) and Adolphe Brongniart (1801–1876), through their studies of fossils, provided incontrovertible evidence for extinction and for changes in regional climate and the elevation of land. How else could they explain the existence of fossils that have no contemporary forms, of fossils from tropical speies found in regions that are now temperate, and of shells and other marine fossils on present-day mountains? A theory of floating and drifting continents (now known as plate tectonics) would await discoveries of twentieth-century marine geologists, but their nineteenth-century forerunners understood that the earth was very old indeed, and it was mutable. Furthermore, if species (and many thousands of them) went extinct, then there had to be multiple periods of creation (or evolution) to compensate for those losses.

Again, these views of a mutable earth, mutable climate, and mutable species were essential for those attempting to classify biogeographic regions based on their respective assemblage of species (Theme 1), to reconstruct the origin, spread, and diversification of life (Theme 2), and to explain differences in numbers and types of species among geographic regions (Theme 3). However, well into the middle of the nineteenth century many scientists held stubbornly to the idea that not only were species immutable, but so were their distributions. Perhaps the most distinguished champion of this static view of biogeography was Louis Agassiz (1807–1873), who argued that species remain essentially unchanged at or near their sites of creation.

The static view was eventually overturned by the passionate and persuasive arguments of no one less than Charles Darwin. Not only did he propose a general theory for the diversification and adaptation of biotas (i.e., natural selection), but he was one of the world's first and foremost dispersalists and champions of dynamic biogeography. Through his observations during his circumnavigation of the globe on the *HMS Beagle* (1831–1836), his later experiments on dispersal of seeds by animals, and his general synthesis on the origin and distribution of life, Darwin convinced many of his colleagues that long-distance dispersal could account for many of the otherwise perplexing patterns of biogeography. Once he was joined by the likes of Asa Gray and Alfred Russell Wallace, Darwin and his colleagues were able to pull off a major paradigm shift in the field—from the static view of the earth and its species to the dynamic view of biogeography.

Yet among all of these visionary scientists it is Wallace who is generally recognized as the father of zoogeography, and biogeography in general. While Darwin argued passionately regarding long-distance dispersal (even to the point of soundly criticizing his mentor, Charles Lyell), most of his energies were devoted toward developing and substantiating his theory of natural selection. On the other hand, biogeography was Wallace's life's work. Brown and Lomolino (1998) listed 17 tenets of the field that were developed by Wallace and included in his seminal monographs *The Malay Archipelago* (published in 1869 and dedicated to Darwin), *The Geographic Distribution of Animals* (1876), and *Island Life* (1880). Five of these tenets of biogeography are listed here:

1. Climate has a strong effect on the taxonomic similarity between two regions, but the relationship is not always linear.
2. The present biota of an area is strongly influenced by the last series of geological and climatic events.
3. Competition, predation, and other biotic factors play determining roles in the distribution, dispersal, and extinction of animals and plants.
4. When two large landmasses are united after a long period of separation, extinctions may occur because many organisms will encounter new competitors.
5. To analyze the biota of any particular region, one must determine the distributions of its organisms beyond that region as well as the distributions of their closest relatives.

Using the latter approach and information provided by over a century of naturalists, Wallace developed a scheme of biogeographic regions (**Figure 1**) that accurately reflected the similarities and differences among biotas. This same scheme, largely unchanged, is still used today.

For obvious reasons, exploration and biogeographic study of the marine realm have always lagged far behind that of terrestrial systems. Yet by the middle of the nineteenth century, biogeographers had made some significant strides in studying this new frontier. Charles Lyell discussed patterns of distribution of marine algae in his seminal work *Principles of Geology*, first published in 1830. Edward Forbes wrote the first comprehensive monograph on marine biogeography in 1856, in which he divided the marine realm into zoogeographic regions based on latitude, depth, and animal assemblages. In 1897 the great British ornithologist and biogeographer Philip Sclater, who produced a predecessor to Wallace's biogeographic scheme, also developed a scheme for the marine realm based on distributions of marine mammals. Following the lead of earlier biogeographers and also based on his own extensive field studies in southwestern North America, C. Hart Merriam (1894) developed a system of what he termed "life zones" that confirmed earlier observations that elevational changes in vegetation were equivalent to those along latitudinal gradients.

Finally, the countless specimens collected during the late eighteenth and early nineteenth centuries enabled others to begin to analyze geographic variation in characteristics of



Figure 1 Alfred Wallace's (1876) scheme of biogeographic regions, which attempts to divide the landmasses into classes reflecting affinities and differences among terrestrial biotas. The regions shown are still widely accepted today. Numbers identify subregions. (Reproduced from Wallace AR (1876) *The Geographic Distribution of Animals*. London: Macmillan.)

individuals and populations (Theme 4). C. L. Gloger reported in 1833 that, within a species, individuals from more humid habitats tend to be darker than those from drier habitats (Gloger's Rule). C. Bergmann (1847) found that in birds and mammals, populations from cooler environments tended to have larger bodies than those from warmer environments (Bergmann's Rule). Also, J. A. Allen reported in 1878 that birds and mammals inhabiting cooler environments also tend to have shorter appendages (Allens' Rule).

Biogeography in the Twentieth Century

Dynamics of the Geographic Template

Even the earliest human explorers appreciated the fact that abiotic conditions vary as one moves from one point on the globe to another. On land, precipitation, temperature, seasonality, prevailing winds, soil conditions, and a host of other important factors vary as we move along transects of latitude, longitude, or altitude. Similarly, in the aquatic realm, temperature, currents, pressure, solar radiation, and concentrations of oxygen and dissolved nutrients vary markedly within and among ecosystems. Together, the variation in all of these environmental characteristics combine to form the geographic template, which influences all biogeographic patterns.

Although a complete understanding of all aspects of the geographic template may be a daunting and truly impossible challenge, at a regional to global scale, geographic variation in environmental conditions is quite regular and interpretable. On land, climatic conditions vary in an orderly manner with latitude, elevation, and proximity to mountain ranges or oceans (**Figure 2**). Major soil types (**Figure 3**) also vary in a similar fashion, partially because soil development is strongly influenced by local climatic conditions, especially precipitation and temperature. In the aquatic realm, although the great volume of ocean waters tends to buffer variation in temperature, surface waters still exhibit a latitudinal gradient in temperature (**Figure 4**). In addition, throughout most of the world's oceans light availability and water temperatures tend to decrease while pressure increases with increasing depth.

As we shall see in later sections, such regular variation in environmental characteristics translates into nonrandom variation in biogeographic patterns of organisms, with each one adapted to slightly different environmental conditions. Such adaptations are, of course, the product of a long and

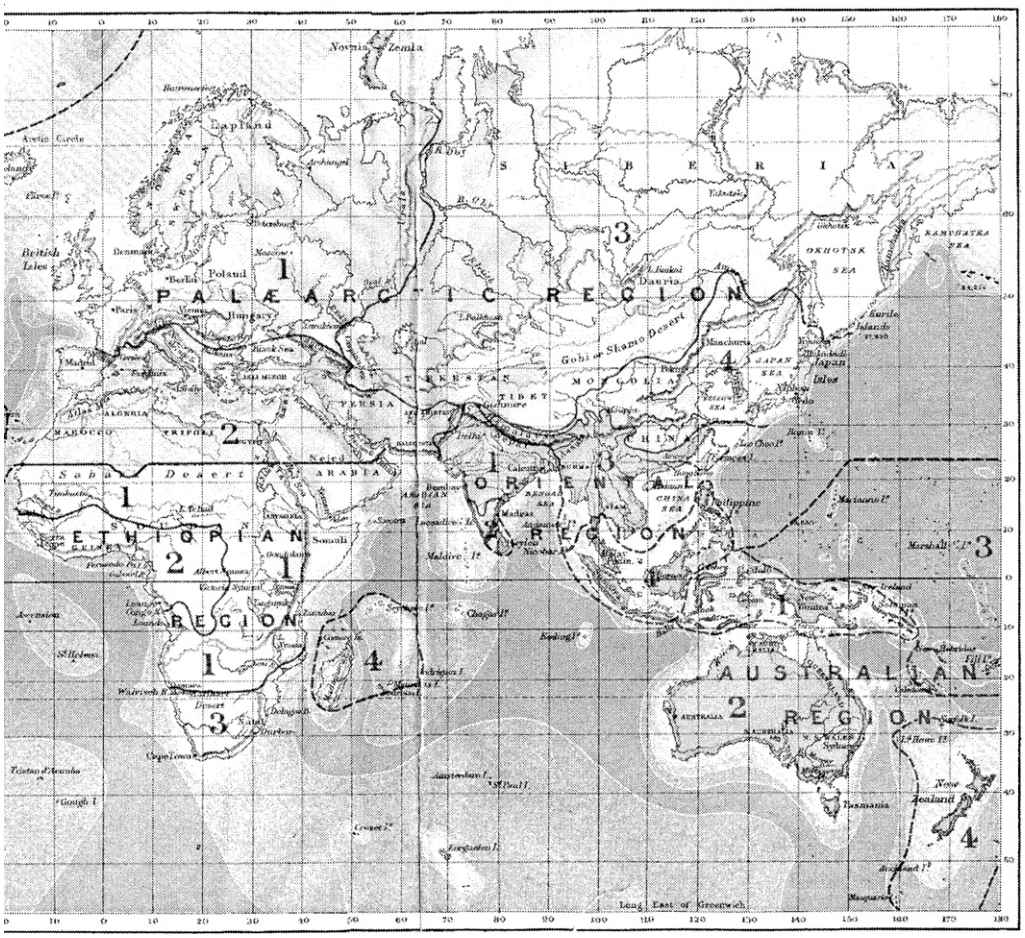


Figure 1 (Continued.)

complex evolutionary history: a series of innumerable interactions between organisms and their environments. With each successive generation, the abilities of descendants to respond and adapt to local environmental conditions change. Evolutionary change, however, is part of a never-ending battle because environmental conditions include other species, which are also evolving. Just as important, the geographic template has evolved throughout earth's 4.5-billion-year history. Because species distributions and other aspects of their geographic variation are influenced by their interaction with the geographic template, a thorough understanding of its dynamics in space and time is essential if we are to understand any major biogeographic patterns.

By the opening of the twentieth century, each of the four persistent themes of biogeography was well established. Explanations for the major biogeographic patterns could now draw on insights from the rapidly growing field of evolutionary biology, as well as our knowledge of the other two fundamental biogeographic processes—immigration and extinction. In addition, biogeographers of the early twentieth century could tap a great wealth of information on geographic variation of biotas and of the environments that they inhabited. Obviously, a thorough knowledge of this underlying geographic template was essential for understanding patterns in distribution and variation among regions and isolated ecosystems. Yet to develop a more accurate and more comprehensive understanding of the major patterns and processes of biogeography, another major scientific revolution was required.

Biogeographers and most other natural scientists knew a great deal about contemporary environments, and most of them appreciated the fact that climatic conditions had changed, sometimes dramatically, during earlier periods of earth's history. Yet until the 1960s, most biogeographers clung to the belief that earth's landforms and ocean basins remained fixed. During the twentieth century, acceptance of the theory of continental drift and plate tectonics revolutionized the field of biogeography as much as acceptance of the theory of natural selection and evolution had in the previous century.

Continental Drift and Plate Tectonics

Although imperceptible to most of us, the earth's continents have moved, colliding at times and drifting apart at others: mountain ranges have formed and eroded away, seas have expanded and contracted, and islands have appeared and disappeared. These changes must have had profound effects on local and regional climates and, in turn, on the geographic distributions and variations of all forms of life on earth. As we will see in a subsequent section, the theory of continental drift and plate tectonics is a relatively recent advance. Yet, with the

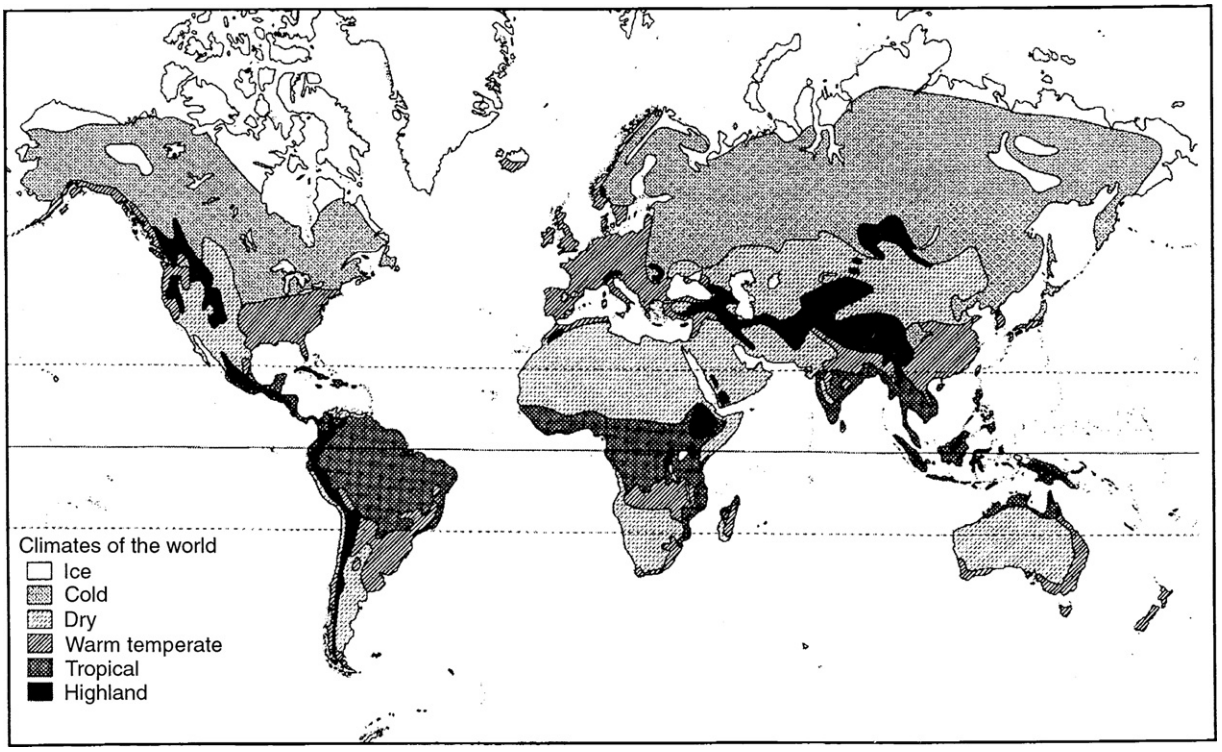


Figure 2 Major climatic regions of the world. Note that these regions occur in distinct patterns with respect to latitude and the positions of continents, oceans, and mountain ranges. (Reproduced from Strahler AN (1973) *Introduction to Physical Geography*, 3rd edn., 468 pp. New York: Wiley.)

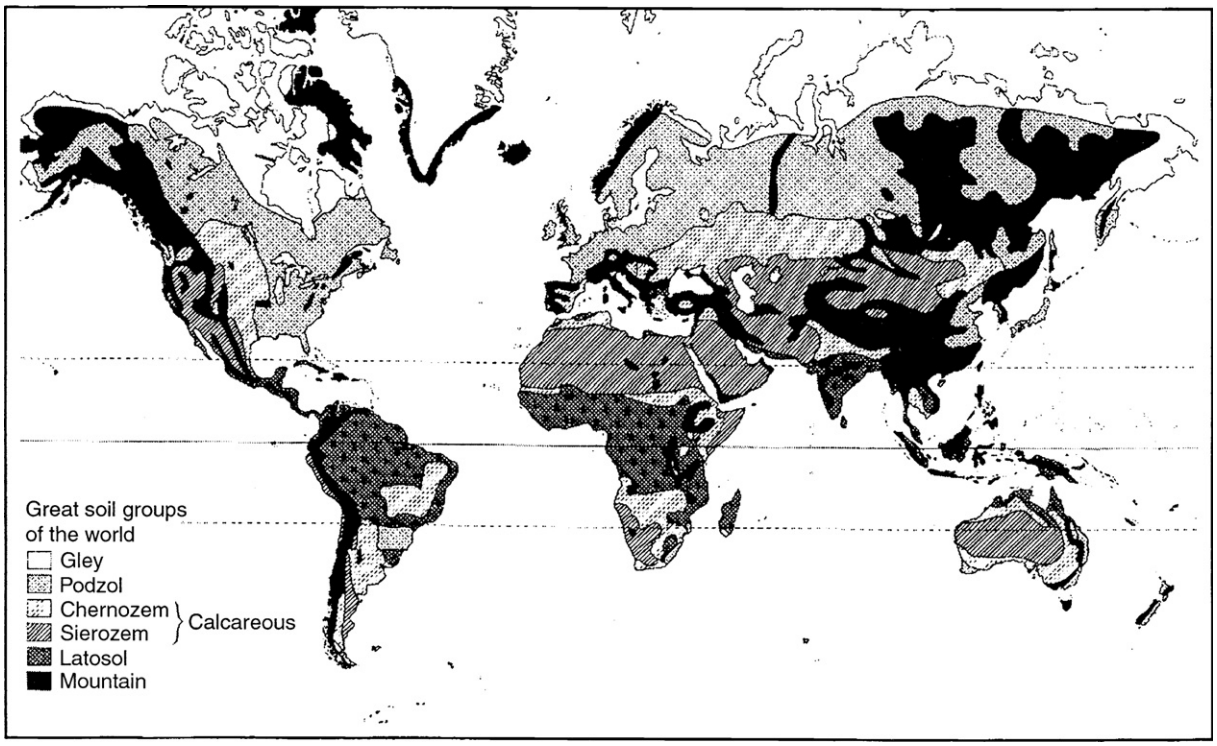


Figure 3 World distribution of major soil types. Note the close correlation of these soil types with the climatic zones shown in **Figure 2**, reflecting the influence of temperature and precipitation on soil formation.

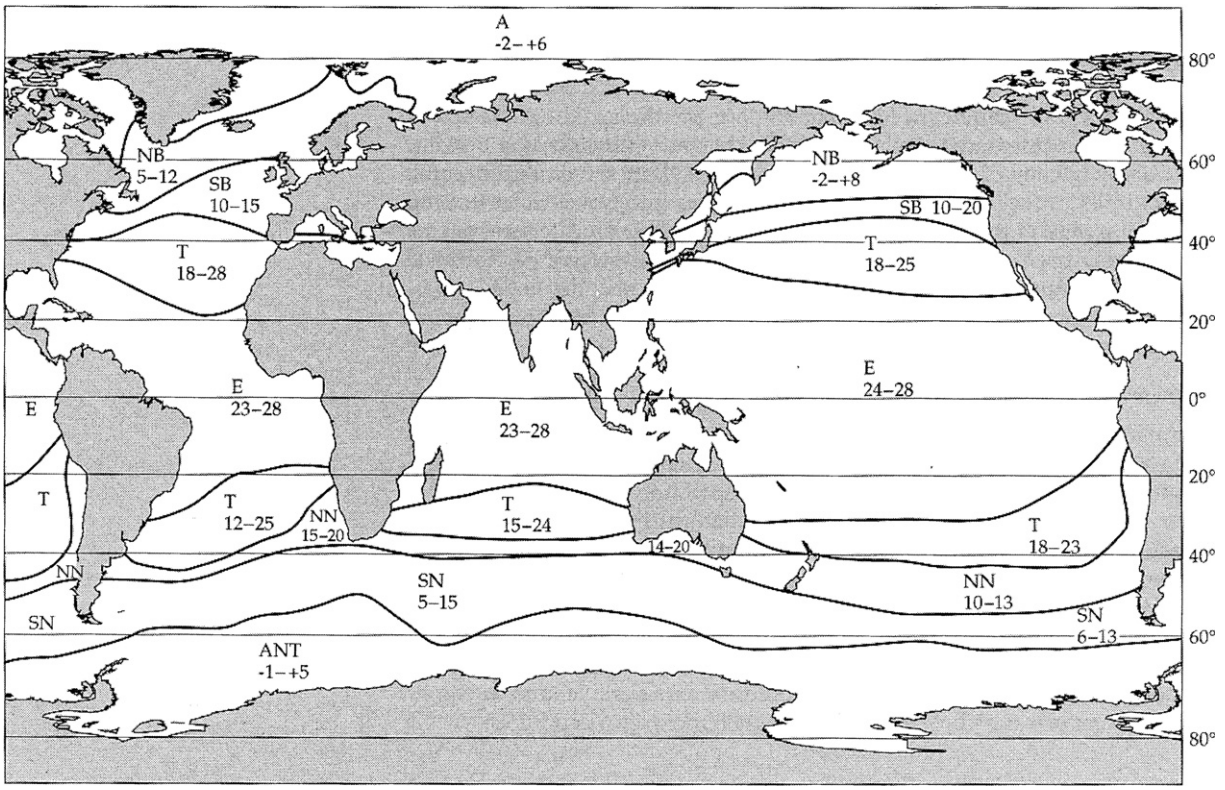


Figure 4 Climatic regions based on mean monthly water temperatures: A, arctic; NB, northern boreal; SB, southern boreal; T, tropical waters; E, equatorial region; NN, northern notal; SN, southern notal; ANT, antarctic. (Reproduced from Rass TS (1986) Vicariance icthyogeography of the Atlantic ocean pelagial. In: *Pelagic Biogeography*, pp. 237–241. UNESCO. Technical Papers in Marine Science, 49.)

possible exception of the acceptance of the theory of natural selection, no other contribution has had more of an impact on the field of biogeography.

Plate tectonics is defined as the study of the origin, movement, and destruction of the earth's plates and how these processes have been involved in the evolution of the earth's crust. The theory of plate tectonics has achieved general acceptance among nearly all scientists and reigns as a unifying paradigm of both geology and biogeography. Yet until just three decades ago, relatively late in the development of these fields, champions of this theory were viewed as oddballs and heretics.

As with any other revolutionary theory in science, it is extremely difficult to pinpoint the origins of the theory of plate tectonics. The great geologist Charles Lyell entertained the idea during the 1830s and 1840s, but then abandoned it in favor of the accepted doctrine of the fixity of the continents and ocean basins. In their attempts to explain the affinities of biotas of now isolated continents, Lyell, Joseph Dalton Hooker, and other "extensionists" of the nineteenth century hypothesized the periodic emergence of great land bridges that then allowed biotic exchange. Darwin, Wallace, and other members of the dispersal camp soundly criticized such views: nothing vexed Darwin more than those extensionists who created land bridges "as easy as a cook does pancakes." In an uncharacteristically critical passage in one of his letters, Darwin complained to Charles Lyell of "the geological strides which many of your disciples are taking…. If you do not stop this, if there be a lower region of punishment for geologists, I believe, my great master, you will go there."

As it turns out, neither the extensionists nor the dispersalists were correct. In most cases, the similarities among now isolated biotas were instead the result of "dispersal" of the continents themselves. Perhaps the first important evidence for what was at first referred to as the theory of continental drift was the configuration of the continents. That is, once geographers had developed relatively accurate maps, it became clear to some that opposite coastlines seemed to fit. In 1858 one of Lyell's contemporaries, Antonio Snider-Pelligrini, may have been the first to demonstrate the geometric fit of the coastlines of continents on opposite sides of the Atlantic Ocean. Yet it wasn't until 1908 and 1910 that an American geologist, F. B. Taylor, and a German meteorologist, Alfred L. Wegener, independently developed models describing the movements of the earth's crust, along with the formation of mountain chains, island arcs, and related geologic features. Wegener continued to develop his model into a more comprehensive theory of continental drift, publishing his treatise in the 1920s. Wegener's theory, however, included too many assumptions about geologic processes and patterns that would not be well established for another three or four decades. His theory also included factual errors, such as overestimating the rate of movement of the earth's plates by perhaps two orders of magnitude. Finally, although he speculated on a potential mechanism, Wegener's theory really lacked a plausible one that could somehow drive the massive plates about the earth

like bits of ice on a pond in spring. It is perhaps one of history's most tragic ironies that, in his quest to discover this mechanism by exploring a volcanically active region of Greenland, Wegener perished in a snow storm.

Wegener's insights would not be widely appreciated for another three decades. Acceptance of the theory of continental drift and its maturation to become the more comprehensive theory of plate tectonics would require many additional insights from geographers, paleontologists, and especially marine geologists during the 1940s and 1950s. These scientists found that when they "rejoined" the continents based on their geometric fit, not only did their biotas seem to match up, but so did topographic features such as mountain chains, rock strata, and fossil and glacial deposits. Perhaps most critical to the acceptance of the theory of continental drift were the efforts by marine geologists following World War II to map the surface of the ocean basins. It soon became clear that beneath each ocean lay a system of ridges that were situated far offshore. As one moved away from these ridges, the seafloor became deeper and more ancient as well. Provided with these and related clues, Herman Hess and his colleagues developed the theory of seafloor spreading: continental drift finally had an underlying mechanism (**Figure 5**). Eventually, paleomagnetic evidence would allow marine geologists to estimate the previous positions of the continents and develop reconstructions of the sequences of their movements and creation and the dissolution of previous continents. Biogeographers were now armed with not just the evidence, but also the mechanisms responsible for the dynamics of biotas and the geographic template itself (i.e., immigration, extinction, evolution, and plate tectonics).

Glacial Cycles of the Pleistocene

The great shifting, collision, and separation of earth's plates profoundly affected the distribution of its biota, both directly and indirectly. Not only did plate tectonics alter major dispersal routes among biotas, but it substantially changed both global and regional climates. As plates shifted across different latitudes, their local biota was exposed to major shifts in climatic conditions. Areas of what is now tropical Africa, South America, and Australia once were situated over the south pole and exposed to severe antarctic climates.

On a global scale, drifting continents also triggered great shifts from periods of global warming to those dominated by glacial conditions. Land absorbs substantially more solar energy than does water. Thus, global climates tended to be warmer when landmasses were situated near the equator, but cooled as they shifted poleward.

Yet global climates can change substantially even during periods too short for substantial shifting of earth's plates. For example, during the Pleistocene (roughly the past 2 million years), earth experienced many climatic upheavals. Rather than being caused by any shifts in plates (which must have been minor given the relatively short period), these climatic shifts were caused by periodic changes in characteristics of the earth's orbit (referred to as Milankovitch cycles; **Figure 6**). These changes significantly altered the total amount of solar energy intercepted by the earth, ultimately causing the climatic reversals of the Pleistocene. During full glacial periods, global temperatures dropped by as much as 6 °C and most landmasses beyond 45° latitude were covered with glaciers often 2 to 3 km thick. Because so much water was tied up in the glaciers, sea levels dropped by 100 to 200 m, thus uniting long-isolated biotas via temporary land bridges. For example, during the last glacial maximum, the region of Southeast Asia and Malaysia was united with Sumatra, Java, and Borneo to form Greater Sunda, while Australia and New Guinea formed the island continent of Sahul (**Figure 7**).

Winds, ocean currents, and precipitation patterns also changed substantially between interglacial and glacial periods. With each climatic upheaval, environmental regimes shifted across both latitudes and elevations. Regions such as the American Southwest, which is now dominated by desert and xeric grasslands, were once covered with coniferous forests. Warming and drying conditions that led to the current interglacial period dramatically reduced these forests and caused them to shift toward the mountain peaks, where cool and relatively humid conditions prevail.

These and other events must have profoundly influenced the distributions of most if not all biotas. As Brown and Lomolino (1998) summarize, however, all the complex biogeographic dynamics of the Pleistocene were triggered by three fundamental changes in the geographic template:

1. Changes in the location, extent, and configuration of principal habitats.
2. Changes in the nature of environmental regimes (combinations of temperature, seasonality, precipitation, and soil conditions).
3. The creation and dissolution of barriers associated with changes in sea level or elevational shifts in habitats.

The responses of both terrestrial and aquatic biotas, while no doubt complex, also were of three types:

1. Some species shifted geographically with their optimal habitats.
2. Some species remained and adapted to the altered local environment.
3. Other species, unable to modify their ranges or ecological associations, suffered range contraction and eventual extinction.

The biogeographic dynamics of the Pleistocene remains one of the field's most active and interesting study areas. Recent advances in analyzing and dating fossil material continue to add to our ability to reconstruct the historical development of biotas (Theme 2) and, in turn, understand major episodes of biotic interchange and recent extinctions, as well as the current distributions of species.

Current Trends in Biogeography

Gradients in Species Diversity and Composition

In addition to biogeographic reconstructions, modern biogeographers continue to study an impressive diversity of patterns encompassing each of the four persistent themes of the field. During the middle of the twentieth century, many biogeographers focused on more general questions. Rather than dissecting and reconstructing the range of selected species,

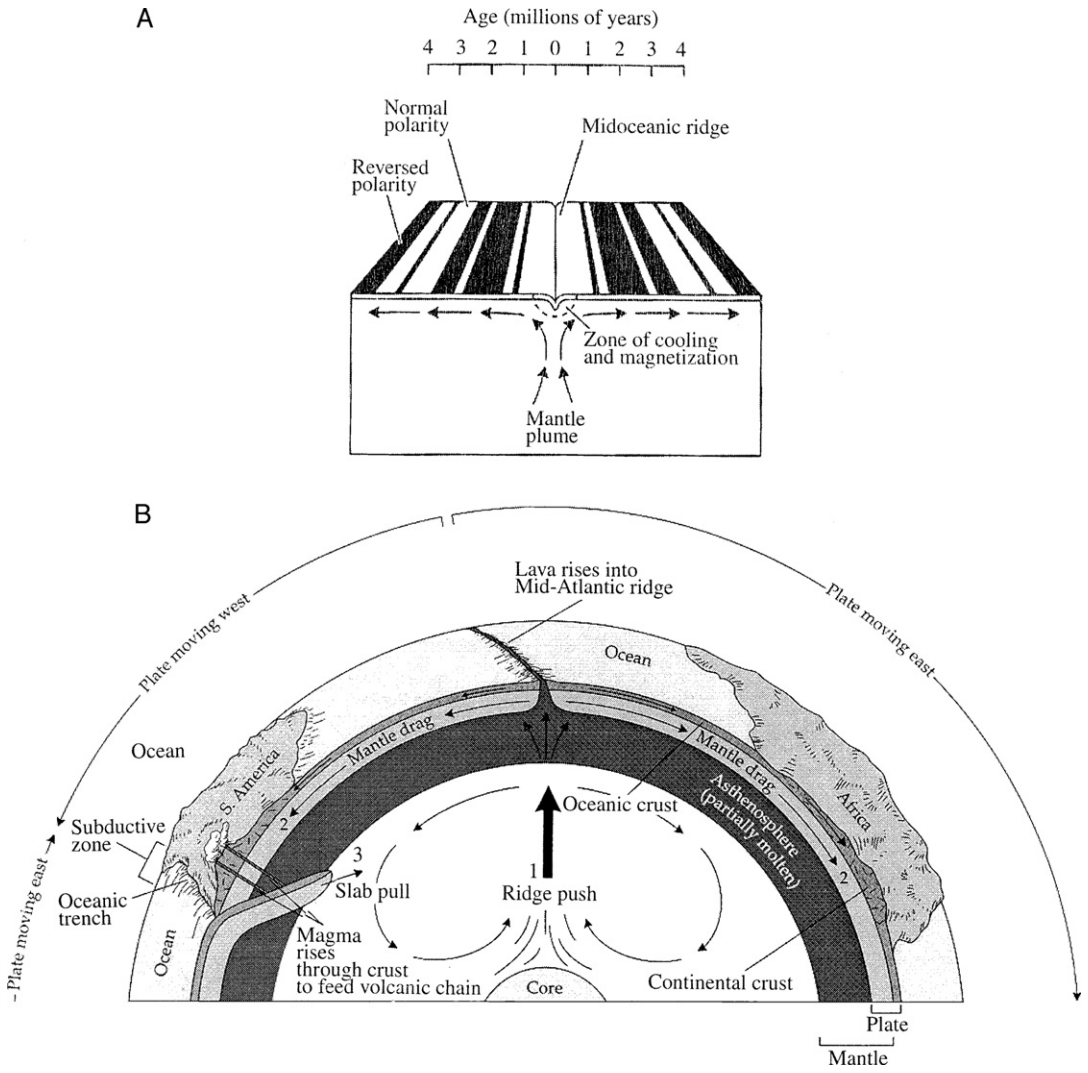


Figure 5 (A) During seafloor spreading, reversals in the earth's magnetic field are recorded as the magnetically sensitive, iron-rich crust cools. Differences in the widths of the magnetic stripes reveal differences in the duration of these polarity episodes and in the rate of seafloor spreading over time and among regions. (B) The current model of plate tectonics includes the possibility that at least three forces may be responsible for crustal movements: (1) ridge push, or the force generated by molten rock rising from the earth's core through the mantle at the midoceanic ridges; (2) mantle drag, the tendency of the crust to ride the mantle much like boxes on a conveyor belt; and (3) slab pull, the force generated as subducting crust tends to pull trailing crust after it along the surface. (Reproduced from Stanley SM (1987) *Extinction*. New York: Scientific American Books Inc. with permission from Nature.)

they examined trends in the total number of species, or what is often termed species richness.

Though many patterns in richness have been studied, two have received the lion's share of attention: the species–area and species–latitude relationships. Early explorers and naturalist of the seventeenth and eighteenth centuries noted the tendencies for species richness to increase with area of a region or island, and be higher for tropical versus temperate, subarctic, and arctic biotas. Armed with data from many hundreds of additional surveys and with a battery of sophisticated statistical tools, twentieth-century biogeographers confirmed the great generality of these patterns and developed some relatively simple models to explain those patterns. Often these models were equilibrial, assuming that species richness resulted from the combined but opposing effects of processes such as immigration into an area (which added species) and extinction (which decreased species richness). MacArthur and Wilson's equilibrium theory of island biogeography is perhaps the prototypic example of such a theory, and one that has dominated the field since its first articulation in the 1960s. Their theory was developed to explain both the species–area relationship and the species–isolation relationship (i.e., the tendency for species richness to decrease as one moves from near to more isolated islands). Simply stated, because immigration rates (the number of species new to an island) should decrease while extinction rate (loss of species already present) should increase as the island accumulates species, the island should eventually reach a level of richness at which immigrations balance extinctions. This equilibrial level of richness should vary among islands: decreasing with isolation because

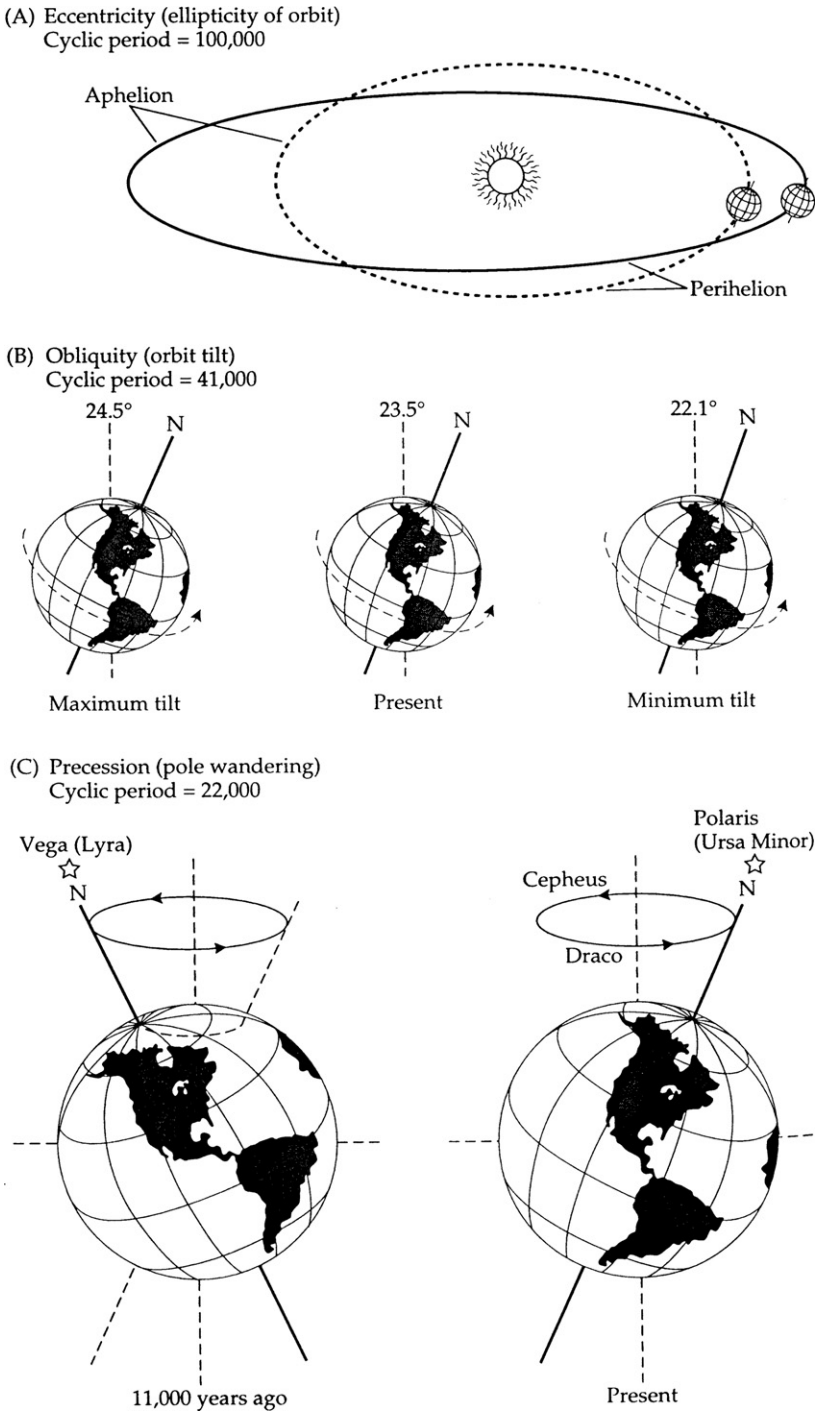


Figure 6 Milankovitch cycles are periodic changes in the eccentricity, obliquity, and precession of the earth's orbit. Each of these changes influences the earth's interception of solar radiation; therefore, these cycles may have been largely responsible for the glacial cycles of the Pleistocene. (Reproduced from Gates DM (1993) *Climate Change and Its Biological Consequences*. Sunderland, Massachusetts: Sinauer Association, with permission from Sinauer Associates.)

immigration rates are lower for more isolated islands, and increasing with island area because populations on larger islands should be less prone to extinction.

Biogeography in the Twenty-first Century

The equilibrium theory stimulated many studies in biogeography and related fields of ecology, and has served as the paradigm of island biogeography for some four decades. Yet an increasing number of biogeographers are beginning to question its utility as a modern paradigm. Species richness is often influenced by speciation and disturbances (e.g., major storms and tectonic events), processes not included in MacArthur and Wilson's original theory. Either the theory has to be expanded to include these processes, or it will be

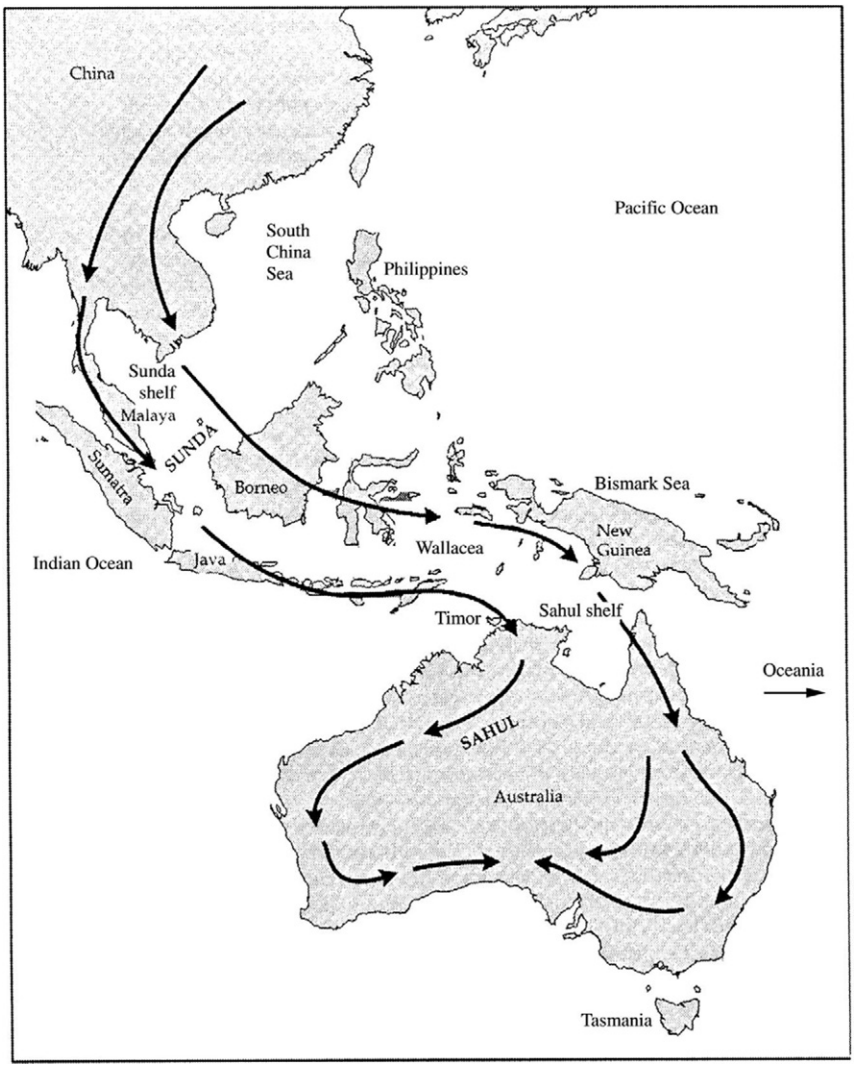


Figure 7 The lowering of sea levels during glacial maxima of the Pleistocene caused the exposure of continental shelves and the formation of dispersal routes across four regions of the eastern Pacific: Sunda, Wallacea, Sahul, and Oceania. (White areas = land exposed during glacial maxima; dark shading = deep water (>200 m); possible dispersal routes are indicated by arrows). Drawn by Simon SS. Driver from The Journey from Eden by Brian M. Fagan, Thames & Hudson Ltd., London.

replaced by an alternative model—one that may eventually become the new paradigm of the field.

Whatever form such a model takes, it must be sophisticated enough to address the growing complexity of questions and patterns that we now study. MacArthur and Wilson's model was primarily developed to explain patterns in richness along gradients of area and isolation. Yet modern biogeographers are now searching for a theory to explain patterns across other geographic gradients and, perhaps more important, to explain geographic trends in the types rather than just the numbers of species. Why do the proportions of large versus small, endothermic versus ectothermic, herbivorous versus carnivorous animals, woody versus herbaceous, or annual versus perennial plant species vary across geographic gradients? How are biogeographic reconstructions related to phylogenies? In what manner do gene frequencies vary with isolation or across other geographic gradients? How do the size and shape of geographic ranges vary with latitude and among taxonomic groups, and how do population density and other demographic parameters vary across the range of a species?

We may have reached a point at which our questions and appreciation for the complexity of nature have become too sophisticated for the relatively simple models that have dominated the field since the 1960s. If this is true, biogeography may be on the verge of a major scientific revolution, one that may well rival those triggered by the seminal insights of scientists such as Charles Darwin, Alfred Russell Wallace, Alfred Wegener, Robert H. MacArthur, and E. O. Wilson.

Biogeography and the Conservation of Biodiversity

Biogeographers study both the patterns and processes influencing the geographic variation of nature. We study not just how many species occur in a particular area, but why more are there than somewhere else and which ones are likely to be

shared among areas. We study and attempt to develop explanations for what are now termed "hotspots," regions of relatively high numbers and high endemicity of species. Biogeographers also study variation in the geographic template, including that associated with anthropogenic disturbances such as the spread of exotic species or the spatial patterns of deforestation. Many biogeographers study extinction and have demonstrated that it has a geographic signature: loss of species tends to be highest for the smallest and most isolated sites, namely, oceanic islands and fragments of once expansive habitats on the mainland. Given this, it becomes obvious that few disciplines could be any more relevant to understanding and preserving biological diversity than biogeography.

Our task, however, is far from a simple exercise of just applying what we already know. Indeed, only a small fraction (perhaps just 2 or 3%) of all extant species have been described, and we know precious little about the geographic distributions of most of those. What we do know often comes down to just general patterns for common species, but conservation biologists require detailed information on the rare species—those that may be the exceptions to most rules.

We have, however, made great progress in recent years in mapping and measuring the intensity (number of endemic species) of hotspots of biological diversity. In theory, these hotspots of biodiversity should receive the highest priority from conservation biologists, especially when they coincide with high levels of human activity. Yet even these approaches, generated by existing survey information and sophisticated geographic analyses, are based on a relatively limited number of surveys.

To develop more effective strategies for conserving global diversity, we still require a much more thorough understanding of the geographic variation of nature. A number of distinguished ecologists and biogeographers, including E. O. Wilson, have called for greatly accelerated efforts to map the diversity of life. With adequate support, within just a few decades we could greatly expand our knowledge of the distributions of most life-forms and, eventually, contribute to their conservation as well.

See also: Darwin, Charles (Darwinism). Dispersal Biogeography. Historical Awareness of Biodiversity. Hotspots. Island Biogeography. Species–Area Relationships. Vicariance Biogeography

References

Briggs JC (1987) *Biogeography and Plate Tectonics*. Amsterdam: Elsevier.

Brown JH (1995) *Macroecology*. Chicago: Chicago University Press.

Brown JH and Lomolino MV (1998) *Biogeography*. Sunderland, Massachusetts: Sinauer Associates.

Carlquist S (1974) *Island Biology*. New York: Columbia University Press.

Darlington PJ Jr. (1957) *Zoogeography: The Geographic Distribution of Animals*. New York: John Wiley & Sons.

Fagan BM (1990) *The Journey from Eden: The Peopling of Our World*. London: Thames and Hudson.

Gates DM (1993) *Climate Change and Its Biological Consequences*. Sunderland, Massachusetts: Sinauer Association.

Guilaine J (1991) *Prehistory: The World of Early Man*. New York: Facts on File Publishers.

MacArthur RH and Wilson EO (1963) *The Theory of Island Biogeography*. Monographs in Population Biology. Princeton, New Jersey: Princeton University Press.

Maurer B (1994) *Geographic Population Analysis: Tools for Analysis of Biodiversity*. London: Blackwell Scientific Publications.

Rass TS (1986) Vicariance icthyogeography of the Atlantic ocean pelagial. In: *Pelagic Biogeography*, pp. 237–241. UNESCO. Technical Papers in Marine Science, 49.

Rosenzweig ML (1995) *Species Diversity in Time and Space*. New York: Cambridge University Press.

Sauer JD (1988) *Plant Migration: The Dynamics of Geographic Patterning in Seed Plants*. Berkeley: University of California Press.

Sclater PL (1858) On the general geographical distribution of the members of the class Aves. *Zoology* 2: 130–145.

Simberloff DS (1988) Contribution of population and community biology to conservation science. *Annu. Rev. Ecol. Systematics* 19: 473–511.

Simpson GG (1980) *Splendid Isolation: The Curious History of Mammals in South America*. Oxford, United Kingdom: Pergamon Press.

Stanley SM (1987) *Extinction*. New York: Scientific American Books Inc.

Strahler AN (1973) *Introduction to Physical Geography*, 3rd edn., 468 pp. New York: Wiley.

Udvardy MDF (1969) *Dynamic Zoogeography*. New York: Van Nostrand-Reinhold.

Wallace AR (1869) *The Malay Archipelago: The Land of the Orangutan, and the Bird of Paradise*. New York: Harper.

Wallace AR (1876) *The Geographic Distribution of Animals*. London: Macmillan.

Wallace AR (1880) *Island Life, or the Phenomena and Causes of Insular Faunas and Floras*. London: Macmillan.

Wegener A (1966) *The Origin of the Continents and the Oceans* (translation of the 1929 edition by J. Biram). New York: Dover Publications.

Bioprospecting

Paul Alan Cox, Institute for Ethnomedicine, Jackson, WY, USA
Steven King, Napo Pharmaceuticals, Inc., San Francisco, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Matteo, N., Nader, W., Tamayo, G., volume 1, pp 471–488, © 2001, Elsevier Inc.

Glossary

ALS Amyotrophic Lateral Sclerosis, a devastating paralytic disease known in Europe as motor neuron disease and in North America as Lou Gehrig's disease.
AMPA-kainate receptors Glutamate receptors often found in motor neurons.
Archaea Nonbacterial microorganisms that often live in extreme environments, comprising (in addition to eukarya and bacteria) the third domain of life.
Biomimetics An artificial compound, device, or structure that mimics or was inspired by nature.
Denticles Small tooth-like projections such as those that occur on shark skin.
DNA fingerprinting Techniques used to amplify small amounts of DNA to identify individuals, often for forensic purposes.
Extremophile An organism that lives and thrives in a habitat characterized by extremes of temperature, acidity, alkalinity, salinity, light, or pressure that would prove lethal to most other organisms.
Herbal A book recording accumulated knowledge of medicinal plants.
Indigenous knowledge Traditional understandings of villagers which are often orally transmitted from generation to generation.
Patent Government grant of the right to inventors to exclude others for a period from manufacturing or selling an invention.

Bioprospecting refers to the assessment of life forms for development of unique molecules, biological entities, or structures that have potential utility in the economic sphere. Examples of bioprospecting include the screening of microorganisms for the ability to metabolize oil spills, the screening of marine organisms for antifouling compounds to prevent the attachment of barnacles to ship hulls, the analysis of archael species in hot springs to discover new temperature resistant DNA polymerases, or the search for new pharmaceutical compounds from rain forest plants for the treatment of disease.

Commercial Development of Biodiversity

Although the term bioprospecting is often used in the press to characterize the search for commercially viable new molecules, genomes, and structures from nature, the imputed analogy to destructive, extractive enterprises such as mining is incorrect because the biological resources are not typically consumed or destroyed during the development process. Thus bioprospecting is considered by many scientists to be a pejorative term that does not properly describe the recourse to nature that is being increasingly used to generate important new discoveries. Although bioprospecting programs have been successfully used to discover new agriculturally useful compounds such as herbicides and insecticides, the development of new or improved agricultural cultivars or livestock animals is typically not considered to be a bioprospecting activity. Implicit in a bioprospecting exercise is the goal of discovering a novel molecule, biological structure, or development of a unique genotype which can, in theory, be produced in the laboratory, rather than be gathered from nature. If replicable synthesis of the molecule or genotype cannot be produced in the laboratory due to molecular or genetic complexity, laboratory extraction of the useful entities from agricultural production is employed. In the absence of such repeatable synthesis or process, the commercial requirements for dependable sourcing and intellectual property protection are unlikely to be achieved. Unlike unsustainable wild harvesting that sometimes occurs in the herbal trade, bioprospecting programs are firmly rooted in biological conservation. When successfully conducted, bioprospecting may better be termed "sustainable development from nature" particularly in biodiverse-rich countries which share the cultural, educational, and financial benefits of discovery.

History

Ethnobotanical and ethnozoological studies document an awareness of the natural world by indigenous peoples that many individuals resident in industrialized countries would find remarkable. For example, in the late 1920s, evolutionary biologist Ernst Mayr was surprised to discover that villagers in New Guinea had names for all of the bird species that he collected. "I realized that every bird the natives saw and that they had shot, they knew the name. So I recorded that Latin name that I knew and the names the natives gave me . . . I found that I had in this locality collected 137 species of birds, and the natives had given me the names of 136 of these." In Samoa, villagers have names for over 46 different breadfruit cultivars (Ragone *et al.*, 2004). Indigenous names are typically

binomials that often describe salient features or useful properties of the plants and animals in their environment.

Although usually transmitted orally, traditional knowledge of biodiversity was sometimes compiled in early herbals and bestiaries, including useful aspects of different species. Gerard's *Herball* of 1597 includes a description of the pharmacological efficacy of *Digitalis purpurea* (Scrophulariaceae): "Foxe glue boiled in water or wine, and drunken, doth cut and consume the thicke toughness of grosse and slimie flegme and naughtie humours; it openeth also the stopping of the liver, spleene and milt and other inward parts." The development of digitalis as a crude drug for the treatment of congestive heart failure awaited further investigations by William Withering in 1785. Today, over 1500 kg of digoxin are prescribed to heart patients throughout the world. A total of 18 different pharmaceuticals, including 16 used today have been derived from Gerard's *Herball* (Cox, 1998). Indigenous observations of animals have also yielded extraordinary uses. For example, in Sweden, certain species of snails were used to lubricate the axles of wagons, while the viscera of fugu puffer fish (*Fugu rubripes*) in Japan was found to produce in individuals eating the fish an unusual and potentially dangerous tingling of the tongue and lips due to the presence of tetrodotoxin. Similarly, the hairs of the velvet bean *Mucuna pruriens* (Fabaceae) in Africa were used to tip poison arrows in Nigeria and as a murder drug in Senegal due to the presence of serotonin and its analogs including dimethyltryptamine and 5-methoxy-dimethyltryptamine. This same plant yields levodopa, a powerful drug used in the treatment of Parkinson disease. More recently, Eric Buenz and his colleagues at the Mayo Clinic have referenced the *Herbarium Amboense*, an early eighteenth century compilation of botanical knowledge from Indonesia by Rumphius, as a source of discovering new pharmaceuticals (Buenz *et al.*, 2004).

These discoveries exemplify the often serendipitous nature of discovery of useful products from biodiversity, a process that has been accelerated with the rise of modern science.

Targets, Screens, and Structural Determination

Historically, most innovative uses of biodiversity have been based on serendipity rather than on a systematic approach to discovery such as those used in modern bioprospecting programs. Often nature, which through evolutionary time has provided solutions to physical and biological challenges, is referenced by scientists in the search for novel solutions to physical or medical problems. Sometimes systematic searches for desired properties or entities from nature are carried out. Regardless of the scale and complexity of the search, three components typically compromise modern analyses of nature for new commercial uses: (1) a target of desired physical, biological, or molecular activity, (2) a screen to determine if a particular species or biological sample possesses the desired activity, and (3) isolation, the attempt to remove the desired activity from the species or sample in order to employ it in other settings. Such searches can be a single reference to nature for inspiration, others can involve multiple, iterative approaches to biodiversity. Because screening and isolation strategies are similar for many new entities from biodiversity, it is perhaps useful to categorize bioprospecting programs not on their process, but instead on their targets. These targets can be divided into those used to discover useful entities for agriculture (e.g., herbicides and pesticides), biomimetics (such as velcro), diagnostic procedures (e.g., the development of new enzymes for DNA fingerprinting), disease prevention (e.g., antiseptics, molecules to control disease vectors), environmental remediation (such as removal of consumer waste, oil spills, and toxic mine tailings), genetic analysis (e.g., discovery of new structural and regulatory gene sequences of commercial value), new materials (including antifouling compounds for ships, dirt resistant paints for houses, etc.), and the discovery of new pharmaceuticals (e.g., antibiotics, anticancer compounds, treatments for chronic pain, etc.). Given the extraordinary molecular diversity of life on earth, novel uses of biodiversity will only increase through time as both innovation and scientific techniques for biodiversity analysis continue to increase. Thus, this list of targets should be seen as illustrative rather than comprehensive.

Insecticides, Pesticides, and Herbicides

Fish Poisons as Insecticides

In the islands of Samoa, reef fish are sometimes harvested by the use of plants which have a chemical agent that stuns them. Both *Tephrosia patulaial* (Fabaceae) and *Barringtonia asiatica* (Barringtoniaceae) are used as fish poisons (Cox, 1979). The roots of *T. patulaial* are macerated and placed in a leaf bundle which is immersed in the shallow reef, while the seeds of the box-like *B. asiatica* fruits are scraped into tidal pools. After about five minutes, the stunned fish rise to the surface and are gathered in woven baskets. The poison does not affect the edibility of the fish. *B. asiatica*, known in most Polynesian languages as "*futu*" is used throughout the southwest Pacific, while in Melanesia a legume related to *T. patulaial*, *Derris elliptica* is known in some languages, such as the Geri language of Guadalcanal, as "*vutu*." Obviously, this term refers to use rather than taxonomic affinity. Chemically, the active agent in all of these fish poisons is rotenone, originally discovered in 1895 by Emmanuel Geoffroy from *Lonchocarpus nicou* (Leguminosae) used by villagers in Guyana to capture river fish.

Fish and some insects, such as parasitic mites on chickens, are exquisitely vulnerable to rotenone since it is quickly absorbed through gills and trachea. It is of low toxicity to mammals, including humans, since it is absorbed only slowly through the gastrointestinal system. Rotenone has been used to kill a variety of garden pests, and since it is biodegradable, has only a short window of toxicity. There is some concern that rotenone use may be linked to Parkinson disease in farmers.

Neem, A Natural Insecticide

The neem tree of India, *Azadirachta indica* (Meliaceae), is considered to be one of the world's most useful trees since it provides shade, fuel, fodder, and an extremely important insecticide, Azadirachtin (National Research Council, 1992). For centuries, villagers in India have placed the leaves and branches of the neem tree in granaries and pantries where it serves to actively deter feeding by insects. The active compound,

azadirachtin, is a potent inhibitor of the insect hormone ecdysome and interferes with metamorphosis and molting. At concentrations as low as 10 ppm, it is lethal to moth and butterfly larvae.

One of the promises of neem is its insecticidal efficacy as a crude extract or even as dried plant materials, particularly since azadirachtin is biodegradable. It therefore is suitable for exploitation as an insecticide by the poorest of the poor as well as by major agribusiness industries. In a remarkable case of parallel chemical discovery by indigenous peoples from different continents, the leaves of the Chinaberry tree, *Melia azedarach* (Meliaceae), were used by the Cherokee and Navajo tribes to keep insects away from their homes.

The potency of neem as an insecticide led several multinational agricultural firms including W.R. Grace to seek patent protection for pesticides derived from neem. However, in May 2000, the European Patent Office revoked a patent granted for a neem-based fungicide. The patent rejection responded to a challenge by Vandana Shiva of the India-based Research Foundation for Science, Technology, and Ecology, who argued that the patent was based in part on prior art, for example, folk uses of the plant.

Antifungal Crop Treatments

Higher plants fight an ongoing chemical battle against fungal hyphae that are ubiquitous in the environment and which quickly infect new plant growth. Extracts of the broad bean, *Vicia faba* (Fabaceae), were found to have potent antifungal activity. This activity was initially attributed to a natural acetylene, wyerone, by Fawcett at the University of London. Further studies led to the identification of a reduced form of wyerone, a phytoalexin, as an even more active component. Wyerone is important because it demonstrates wide spectrum activity against many fungal pathogens and has been successfully tested on a variety of crops.

Biomimetics

Velcro, An Alternative Fastener

Velcro, developed for the US space program, was invented in 1941 by Swiss Engineer George de Mestral, who lived in the village of Commugny near Lake Geneva. He was fascinated by the spiney fruits of burdock plants (*Artium tomentosum* (Compositae)) and the way they adhered to his clothes and his dog's fur. Fruits which are dispersed by adherence to the fur or feathers of their unwitting vertebrate hosts are termed to be epizoochorous, and often are equipped with mechanical or chemical adhesives that allow them to be carried long distances. Under the microscope de Mestral noticed that the hooked hairs of burdock fruits ensnare loops of fabric or hair. He realized that a similar pattern of hook and loops could be used to attach different objects together. In his Swiss patent of 1951 and his US patent no. 2,717,437 in 1953 entitled, "Velvet Type Fabric and Method of Producing Same," de Mestral diagramed a hook and loop arrangement for attaching fabrics that was later trademarked as Velcro. In the search for effective fasteners that are impervious to temperature and which cannot freeze nor be fouled like zippers, Everest climbers quickly sought to have Velcro as a backup fastener on their down parkas and trousers. Eventually, Velcro became an important product for astronauts who work in zero gravity. During space missions, Velcro is used to hold down tools, tablets, pens, and other instruments.

Self-Cleaning Lotus Paint

The lotus plant, *Nelumbo speciosa* (Nelumbonaceae), is sacred to three different major world religions. Lotus is considered in the Hindu faith to have sprung from the navel of Lord Vishnu. The rhizomes are boiled in China, Japan, and other Asian countries as a delicacy, and the young leaves are eaten as a vegetable. A remarkable feature of fresh lotus leaves, known to small children, is the inability of the leaves to be wetted. A droplet of water placed on a lotus leaf assumes the appearance of a drop of mercury placed on a glass pane. By swiveling the leaf, the water drop can be spun about the leaf surface like a ball bearing without diminishing in size. Microscopic examination of the leaves reveal a pubescent, rather than smooth surface, with numerous small hairs spaced at a distance smaller than the diameter of a small water drop. The water drop is in essence held aloft by these hairs which allow only 2–3% of the surface area of the water drop to contact the leaf surface. As a result, lotus leaves are self-cleaning, and the slightest vibration in calm morning air causes condensed dew to roll down into the center of the leaf.

These features of the lotus leaf were observed by Wilhelm Barthlott, Director of the Botanical Gardens at the University of Bonn, Germany, who was inspired to create a self-cleaning paint based on the lotus leaf effect. Marketed by the German company STO as Lotusan™, these hydrophobic paints cause water to bead so that mineral, cement, and concrete surfaces painted with them are self-cleaning. Not only does this hydrophobic paint prevent dirt and grime from accumulating, but it also reduces attachment of algae, bacteria, and fungi to buildings in damp and dirty environments.

Architecture Based on Termite Mounds

Designers and architects have studied and utilized the incredibly strong and beautiful designs in natural forms for centuries. The Spanish architect Antoni Gaudi utilized many nature-inspired forms in his buildings in Barcelona, Spain. Among his most famous projects, the Familia Sagrada Church, contains numerous forms from nature, even though the building was only 25% completed at the time of his death in 1926. The church contains tree-like pillars that support the roof in several sections and in the basement of the Familia Sagrada there is an exhibit demonstrating the connection between the forms of trees, shells, leaves, and other forms found in nature in this building that is still under construction in 2011. Similarly, the Sydney Opera House, while considered to be an expressionist modern design by Danish architect Jørn Utzon, is composed of four large precast concrete shells that resemble cockle shells attached to a rocky marine coast.

Another example of architecture inspired by nature is the Eastgate Centre in Harare, Zimbabwe. This building, designed by architect Mick Pearce, has no conventional air conditioning or cooling systems but still maintains a regulated temperature year round. This is achieved in part by mimicking the self-cooling mounds – some which are 9 m tall – built by African termites. The mounds are passively thermo regulated

to maintain a temperature of 31 °C in savannah environments where external temperatures vary from 2 °C at night to 40 °C in the daytime. Such exquisite temperature regulation is required for cultivation of the fungus the termites depend on. Compass termites of the genus *Amitermes* build their mounds with the long axis oriented north to south, and which have a series of heating and cooling vents throughout the mound. The termites are constantly digging new vents and plugging up old ones as part of the process to regulate the temperature, thus generating a series of convection currents that moves air up and down. By mimicking these termite mounds, the Eastgate Centre uses less than 10% of the energy of a conventional building of the same size. This nature-inspired design saved the owners of the building more than $3 million in construction costs.

Gecko Tape

Many visitors to Polynesia have been amazed by the ability of geckos of the genus *Lepidodactylus* to climb upside down on the rafters of village huts, or the ceilings of hotel rooms. Suspended upside down, the geckos chirp to attract mates and capture unwary insects. The ability of geckos to adhere to even smooth surfaces is produced by countless rows of small setae, tiny hair-like projections on their toes. The density of these structures on the gecko footpad is approximately 500,000 per mm. Each setae then branches into hundreds of patula of 100 nm length. Lacking an electrical charge, these projections from the gecko footpad are attracted to surfaces with a small charge by van der Waals forces. This attraction is so strong that geckos walk by curling their toes for attachment and subsequent peeling.

Engineers have attempted to duplicate this approach by producing tiny nanofibers which can attach to a substrate. Andre Geim, a Nobel Prize winning physicist at the University of Manchester, UK, has been adapting the principles of gecko acrobatics to produce gecko tape, which can be used to allow robots to climb cliffs or difficult surfaces, and which may provide new surgical adhesives.

Shark Swimsuits

Predatory sharks, such as the Great White sharks off the coast of northern California, need to move with speed and agility through the marine environment to capture highly predator-averse marine mammals. They are facilitated in their speed by skins with tiny tooth-like scales called denticles that significantly reduce hydrodynamic drag as they swim through the water.

Amy Lang at the University of Alabama, USA, discovered that sharks that swim at high speeds have smaller denticles than shark species that swim at slower speeds. Her data have indicated that the denticles of faster sharks, much like the dimples on golf balls, "bristle" as they swim creating a distinct surface pattern which significantly reduces hydrodynamic drag.

This feat of evolution has profound implications for aircraft and boat design. Reducing the overall drag on large vessels by 1 or 2% can result in huge energy savings in fuel. Swimsuit design has also been influenced. Swimmers at the Beijing Olympic Games used FS11 "fast swimsuits" that mimic some of the morphology of shark skin to enhance their ability to move through the water with minimum resistance (Eadie and Ghosh, 2011).

Diagnostic Procedures and Techniques

Litmus Paper

Lichens were discovered as useful indicators of acidity by the French chemist Gay-Lussac in the early nineteenth century. The word "litmus" is derived from a Middle English term *litemose* related to the Old Norse term for dyer's herbs *litmosi* (dyer's herbs *litr* + *mosi* moss). Lichens including *Roccella tinctoria* (Roccellaceae) from the Mediterranean and *Lecanora tartarea* (Lecanoraceae) from The Netherlands are combined with wood cellulose to manufacture litmus paper. Under neutral conditions, the paper is purple, but turns red in the presence of acids and blue in the presence of bases. While unsuitable for titrations, litmus paper is still used today for rapid assessment of the acidity of environmental and laboratory samples.

Lichens as Air Pollution Monitors

It has been long known that lichens – symbiotic organisms composed of a fungal and an algal or cyanobacterial partner – are exquisitely sensitive to air pollution, probably because they rapidly absorb chemicals from air and water. The nature of the symbiosis means that when the most vulnerable partner to the pollutant is damaged, the symbiosis is destroyed and the lichen dies.

Since lichen species vary in their sensitivities to different pollutants – heavy metals such as lead, zinc, or copper; ozone, hydrocarbons, or acid rains – surveys of lichen species present in a city or near a mine can give a good indication of air quality. Such lichen surveys can be compared to historical surveys of the same area or compared to the lichen floras of nearby regions with pristine air. In northern Sweden, Mats Karström has developed an innovative system for assessing forest quality merely by surveying the lichen flora.

The long filamentous strands of Methuselah's Beard lichen (*Usnea longissma* (Parmeliaceae)) were once used in Europe as tinsel on Christmas trees. But now this lichen, which was once the source of the antibiotic usnic acid, has largely disappeared from industrial areas of Europe where air quality has been compromised. Lichen surveys based on a 10-point scale for sulfur dioxide concentrations were used as early as 1977 by David Hawksworth at the University of Gloucestershire, UK, and Francis Rose at Kings College, London, to map air pollution levels in England, Wales, and Ireland. More recently, Larry St. Clair and his son Sam at Brigham Young University, USA, have used lichen tissues from the Chiricahua National Monument in Arizona as passive filters for the quantification of copper levels in the air.

Green Fluorescent Protein from Jellyfish

In 2008, the Nobel Prize in Chemistry was awarded to Osamu Shimomura, Martin Chalfie, and Roger Tsien for development of the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria*. As a scientist at Princeton University in 1960, Shimomura began studies of bioluminescence in the jellyfish. From specialized organs, jellyfish release calcium ions to the

protein aequorin which glows with blue photoluminescence. This blue light is then absorbed by GFP, which in turn fluoresces green. In 1987, Douglas Prasher at Woods Hole Oceanographic Institution realized that if the genes for GFP were inserted before the stop codon into the genes for any other protein, they could function as a fluorescent tag. In 1992, Prasher cloned the 238 amino acid protein from the jellyfish and in 1994 sent it to Martin Chalfie at Columbia. Chalfie later said about Prasher, who was left off of the Nobel Prize citation, that his "work was critical and essential for the work we did in our lab. They could've easily given the Prize to Douglas and the other two and left me out." Chalfie gave the GFP clone to his graduate student Ghia Euskirchen who inserted it into *E. coli* bacteria. The bacteria were found to fluoresce green under UV light. Subsequently, Roger Tsien and colleagues at the University of California, San Diego, discovered how GFP works, and produced a variety of mutants that fluoresce with different colors and brightness.

These mutant forms of GFP have facilitated the imaging of a variety of cellular processes and have revolutionized fluorescent microscopy. Since many fluorescent stains are toxic to cells or their components, the nontoxicity of GFP allows many cellular processes to be directly observed in living cells for the first time. While GFP mutants are being optimized in Tsien's laboratory to guide cancer surgery, more prosaic uses have now been commercialized, including glowing zebra fish for the aquarium trade and even the production of fluorescent pigs in Taiwan.

Disease Prevention

Seaweeds to Deworm Children

In southern Japan, children are traditionally treated for internal parasites with a red seaweed called "kainin-sou" (*Digenea simplex* (Ceramiales)). In 1953, an excitotoxic amino acid, kainic acid, was isolated from the alga. Laboratory animals dosed with kainic acid went into seizures. This observation led to the discovery of AMPA-kainate receptors on motor neurons. Currently, kainic acid is used to study Alzheimer's disease, epilepsy, and neurodegenerative illness. Recently, an unusual pattern of epilepsy among California sea lions was studied by Goldstein at the University of California, Davis. Intoxication of these marine mammals was traced to a structurally similar amino acid, domoic acid, produced by harmful algal blooms. Models of kainic acid activity were used to analyze excitotoxicity in the sea lions. Today, careful studies of AMPA-kainate receptors in motor neurons may lead to new cures for amyotrophic lateral sclerosis and other motor neuron diseases.

African Trees and Anti-Schistosomaisis Soap

Millions of people in Africa are deprived of healthy lives by schistosomaisis. Characterized by fever, enlargement of the liver, fatigue, and chronic diarrhea, schistosomaisis is a debilitating chronic illness caused by freshwater parasites that burrow into the skin and then travel to the liver. There they produce eggs which are expelled in feces. The parasites require snails to complete their life cycle. Although drugs such as praziquantel can treat schistosomaisis, such pharmaceuticals are often unavailable to poor villagers. Even when entire villages are treated, reinfection from contaminated streams and rivers can quickly reoccur.

Control efforts for schistosomaisis depend on eliminating the snail host. At the University of Lausanne, Switzerland, Kurt Hostettmann has been studying saponins from African tree species that are toxic to schistosomiasis-carrying snails. The late Ethiopian physician Aklilu Lemma and Ethiopian horticulturalist Legesse Wolde-Yohannes received the Right Livelihood Award, considered an alternative Nobel Prize, for their studies of the vine *Phytolacca dodecandra* (Phytolaccaceae) as a means of controlling the snails. Known as "*endod*" or "African soapberry," these vines are used as fish poisons as well as soaps and shampoos. If crude extracts of such plants could be included in washing soap sold to villagers, the lives of millions of poor people could be dramatically improved.

Guam Cycads and Motor Neuron Disease

Soon after World War II, US Army physicians recorded an extremely high rate of a devastating paralytic illness among the Chamorro people of Guam. Now known as ALS/PDC (Amyotrophic Lateral Sclerosis/Parkinsonism Dementia Complex), this disease is believed to result from a traditional diet rich in the neurotoxic amino acid *β*-*N*-methylamino-L-alanine (BMAA). BMAA is found in the cycad seed flour used by the Chamorros as well as in animals they hunt which forage on the cycad seeds.

Symbiotic cyanobacteria in specialized roots of the cycad plants produce BMAA, and there is now concern that human beings outside of Guam may be exposed to BMAA from cyanobacterial blooms in lakes, estuaries, and marine systems as well as from cyanobacterial dust in desert areas. Double-blinded studies of Alzheimer's and ALS patients show BMAA present in their tissues, but not in healthy controls. Currently, new drugs are being designed to remove or block BMAA in the body as a possible treatment for ALS and Alzheimer's disease.

Environmental Remediation

Oil-Eating Bacteria

With the ongoing specter of oil spills from tankers, ocean-based drill rigs, and pipelines, the need for remediation strategies has never been greater. Enter *Alcanivorax borkumensis*, a small bacterium that thrives on petroleum. Able to feed on both hydrocarbons and alkanes abundant in oil, this bacterium occurs naturally near oil seeps in marine environments. Vitor Martins dos Santos, now at the University of Wangeningen, sequenced the large genome of *A. borkumensis* in search for gene sequences that can be inserted into synthetic organisms designed to degrade oil spills.

Terry Hazen at Lawrence Berkeley Laboratories found that soon after the 2010 Deepwater Horizon oil spill in the Gulf of Mexico, populations of *A. borkumensis*, which are typically rare in the ocean, exploded. Hazen has continued to assay microbial diversity near deep-water seeps in order to expand the repertoire of possible organisms and genes that can be used to develop novel strategies for oil spill remediation.

Biodiesel from Algae

World energy consumption is growing while the availability of fossil fuels is diminishing. Efforts to produce corn-based ethanol and soybean-based biodiesel cannot meet all energy demands due to competing needs for prime agricultural land. Salt-tolerant algae can potentially meet this need as some species can grow in desert ponds using marginal water with high salinity. In the Great Salt Lake, Utah, the green alga *Dunaliella salina* (Dunaliellaceae) exists at high salinities – up to 27% – which may be significant in pointing the way for the biofuels industry to produce biodiesel using marginal water supplies. Development of *Dunaliella salina* as an alternative source for fuel will reduce the need for freshwater for bioreactor development while eliminating the possibility of contaminating freshwater supplies from bioreactor effluents. Today, one gallon of algal biodiesel sufficient to power an automobile, truck, or tractor, costs $33. As a result, significant introduction of algal biofuel into the energy mix of the United States must wait until the cost of fossil fuels exceeds the production costs of algal biofuels.

Genetic Analysis

DNA Fingerprinting and Taq Polymerase

In 1993, Kary Mullis at Cetus Corporation shared the Nobel Prize for his discovery of the polymerase chain reaction (PCR) which, when combined with thermal cycling, allows rapid amplification of DNA from a single strand. The PCR technique is used in DNA fingerprinting for forensic purposes to amplify even trace amounts of DNA. The advent of PCR has resulted in the exoneration of scores of death row prisoners who were falsely convicted of capital crimes.

The first PCR techniques depended on a DNA polymerase extracted from the bacterium *Thermus aquaticus*. *T. aquaticus* was harvested from a thermal pond in the Lower Geyser Basin in Yellowstone near the Great Fountain Geyser. In 1969, Thomas Brock and Hudson Freeze provided a sample of this thermophilic bacterium to the American Type Culture Collection (ATCC). This bacterium thrives at 70 °C, a temperature which would split apart the double-stranded DNA of most organisms. Using a *T. aquaticus* sample from the ATCC, a public repository, in 1976, Alice Chien and colleagues at the University of Cincinnati isolated a new DNA polymerase from the bacterium.

Mullis and his colleagues at Cetus used Taq polymerase to develop the polymerase chain reaction. Since this DNA polymerase is heat resistant, it is possible to break apart DNA strands at low temperature and then reassemble them at a higher temperature. Because the molecule was derived from an organism originally collected from a US National Park, the profits derived by the patent owners have resulted in considerable controversy. Cetus eventually sold the Taq patents for $330 million. While the National Park Service claims federal ownership of the original *T. aquaticus* specimen, it does not have legal rights to the subsequent research results. Yellowstone National Park now seeks to enter into benefit-sharing agreements with researchers prior to discovery of novel molecules.

DNA Fingerprinting and Pfu DNA Polymerase

One of the problems with Taq polymerase is that it has a relatively low replication fidelity, with an error rate of 1/9000 nucleotides. The discovery of Taq polymerase therefore led to a search for other DNA polymerases that have proofreading activities. Karl Stetter at the University of Regensburg, Germany, studied geothermally heated marine sediments on Vulcano Island, Italy, and discovered the archeal species *Pyrococcus furiosus*. This microorganism can withstand considerable fluctuations in temperature, and thrives at temperatures around 100°C. Studies by scientists at the US biotech company Stratagene resulted in the isolation of Pfu DNA polymerase, which has a replication error rate of only 1/1.3 million nucleotides. It has, unlike Taq polymerase, a 3′–5′ exonuclease proofreading capability, and thus has now exceeded Taq polymerase in popularity for PCR cloning.

Novel Materials

Seagrass Insulation

From 1907 to 1960, farmers collected leaves of the seagrass *Zostera marina* from beaches in Nova Scotia. Sandy Wyllie-Echeverria at the University of Washington, USA, has analyzed the commercial gathering of seagrass leaves, and finds that at its peak in 1929, the seagrass industry of Nova Scotia sold 462,744 kg of leaves. These leaves were purchased by Samuel Cabot, Inc. to produce high-quality insulation for houses, hotels, hospitals, and lecture halls. By 1930, 350 apartment buildings in New York City and 200 commercial buildings were insulated with seagrass, including Carnegie Hall. The seagrass insulation proved to have extraordinary insulation and acoustic values and to be resistant to fire. However, the onset of World War II interrupted the seasonal gathering of seagrass, and the advent of fiberglass insulation resulted in the demise of the seagrass industry. Currently, there is industrial interest in basing a biomimetic acoustic insulation for concert halls and recording studios on the morphology of seagrass leaves.

Pharmaceuticals and Global Public Health

Malaria

Artemesia annua (Asteraceae) has been used in China to treat malarial fevers for 2000 years. This species is the source of the antimalarial drug artemisinin, known as *qinghaosu*. This compound is combined with other antimalarial drugs in artemisinin-based combination therapy (ACT.) One such ACT, Coartem, developed by the Swiss pharmaceutical company Novartis is used on multiple continents to assist with the battle to control malaria. An important advance has also been made in the creation of sweetened pediatric dispersible tablets for children, as the adult formulation of Coartem is bitter.

One of the major challenges in the creation, distribution, and utilization of ACT is the low yield of the compound from *Artemesia annua* plants. Although *A. annua* is a fast growing species, the global demand has caused the price and availability to hinder inexpensive versions of ACTs. One innovative solution to the challenge has been successful production via synthetic biology of artemisinin, pioneered by Jay Keasling of the University of California, Berkeley, USA. The large scale production of lower cost ACT combination therapies is being executed as a result of the fruits of biospecting, traditional

knowledge and state of the art scientific advances in synthetic biology.

Diarrhea

The current cornerstone of management of diarrhea and dehydration, oral rehydration solution (ORS), has no effect on the severity or duration of diarrhea. Crofelemer is an investigational drug being developed by Napo Pharmaceuticals in California for the treatment of secretory diarrhea from various causes. This antidiarrheal agent is an isolated and purified compound from the latex of the Amazonian tree species, *Croton lechleri* (Euphorbiaceae). Crofelemer is a partial antagonist of the cystic fibrosis transmembrane receptor (CFTR) chloride channel and has activity on the calcium activated chloride channel (CACC). This compound inhibits chloride ion secretion into intestinal lumen and the resulting gastrointestinal fluid accumulation. Based on its mechanism of action of blocking chloride ion secretion through CFTR and CACC, crofelemer has been investigated as an agent for the treatment of secretory diarrhea in patients with *Vibrio cholera* induced diarrhea, traveler's and infectious diarrhea, HIV-associated diarrhea, and irritable bowel syndrome.

To make crofelemer universally accessible within low income countries, following the drug's regulatory approval, the *Crofelemer Access Program* (CAP) by Napo Pharmaceuticals will make a pediatric formulation of crofelemer available at cost to collaborating multilateral, nongovernmental and nonprofit organizations such as Direct Relief International.

HIV

Current state of care for patients with HIV/AIDS is highly active antiretroviral therapy (HAART) which consists of a cocktail of drugs, including reverse transcriptase inhibitors and CD4-site blockers. While effective in extending the life expectancy of HIV/AIDS patients, HAART has some negative side effects. In addition, it is costly, particularly for people in developing countries, and does not eliminate the virus from the body; for example, if medication ceases, a high viral titer quickly returns. One promising drug on the horizon is Prostratin, isolated from a plant, *Homalanthus nutans* (Euphorbiaceae), the bark of which is used by Samoan healers to treat hepatitis. As a phorbol, Prostratin is unusual in that it activates protein kinase C but is not a tumor promoter. Prostratin is able to clear laboratory animals of virus.

The product of an ethnobotanical research program of Paul Cox, now at the Institute for Ethnomedicine, and investigators at the National Cancer Institute (NCI), Prostratin is being developed by the AIDS Research Alliance (ARA) in California. Negotiations between the healers, village chiefs, and the Government of Samoa resulted in a landmark benefit-sharing agreement which guarantees stakeholders in Samoa 20% of ARAs profits. A second agreement with Jay Keasling and the University of California, Berkeley resulted in a 50% share of profits from the gene sequence for Prostratin being returned to Samoa.

Pain and Addiction

The roots of iboga (*Tabernanthe iboga* (Apocynaceae)) are used throughout central Africa to induce a hallucinogenic experience that serves as a rite of passage into adulthood and which also can treat female sterility. These hallucinations are claimed to render insight as to the causes of sterility, and to facilitate contact with a villager's ancestors. The indole alkaloid ibogaine as well as its metabolite noribogaine are now under investigation by Deborah Mash at the University of Miami Brain Bank for treatment of chronic pain and to help deal with addiction. Initial clinical trials in St. Kitts with heroin addicts have been highly successful (Mash *et al.*, 2000). The Miami biotech firm DEMERx is seeking FDA approval for an ibogaine treatment for chronic pain.

Cancer

Bioprospecting has yielded multiple drugs approved by the FDA to treat various forms of cancer including vincristine and vinblastine isolated from the Madagascar periwinkle *Catharanthus roseus* (Apocynaceae). Vincristine and vinblastine are drugs of choice for the treatment of pediatric leukemia. Taxol, a drug marketed by Bristol-Meyers Squibb, is used to treat ovarian and breast cancer. Taxol was isolated from the Pacific yew tree *Taxus brevifolia* (Taxaceae) during a systematic search by the National Cancer Institute. Currently, Lars Bohlin at Uppsala University and Samantha Gerlach at Tulane University are studying cyclotides, small cyclic peptides from the Violaceae and Rubiaceae, as a new category of biodiversity derived anticancer compounds.

Some innovative research programs are analyzing complex mixtures of plant and other naturally occurring substances used in traditional Chinese medicine and Ayurvedic medicine from India for new anticancer treatments. New Earth Biomed (NEBM), a not-for-profit cancer research organization, is currently evaluating such alternative treatments.

Intellectual Property

Patents and Natural Products

Patents and Exclusionary Rights

A patent is the right granted to an inventor by a government to exclude others for a limited time from selling an invention in the marketplace. As government grants, patents are only enforceable within a limited jurisdiction; for example, US patents are enforceable only within the US, Japanese patents are enforceable only within Japan, etc. Thus a US patent cannot prevent any individual or firm from producing or marketing a pharmaceutical in Costa Rica, China, or Indonesia. Instead, patents would have to be granted by each of these countries to exclude unauthorized sellers. As a result, most pharmaceutical patents are filed only in large industrialized countries where significant revenues are anticipated. Patents do not, therefore, typically exclude traditional noncommercial uses, or uses in biodiverse developing countries. Once a patent expires, anyone can make or sell the invention without restriction. Patent law has been developed to encourage research and innovation by allowing inventors, for a time, to enjoy the financial proceeds of their invention, before the invention becomes public property.

Patents are particularly important in the development of new drugs, since few pharmaceutical firms would risk the $700–850 million in research and development costs needed

to bring a new drug to the market if they did not have a reasonable expectation of recouping these costs via patent protection. In the absence of patents, pharmaceutical development and the creation of new drugs would largely cease.

Patentability of Natural Products

There is an incorrect perception that naturally occurring molecules cannot be the subject of patents. If a new chemical structure is isolated from a species found in nature, and it has not been described in the scientific literature, a patent can be applied for a new "composition of matter" patent. If that novel composition of matter has a use or utility, that is not obvious, a patent may be granted by a patent office, depending on the scientific data presented and the rules of a specific country's patent office. If a naturally occurring compound that has previously been described in the literature, i.e., it is already in the public domain, but there is a novel, as yet never described use for that compound, an inventor may apply for a use patent. The novel use for this compound however must not be obvious or previously described in the scientific literature.

Myriad Genetics and Bioprospecting Patents

Patent law prohibits protection for inventions that are obvious, such as sunlight or pollination, or that do not involve significant innovation on the part of the inventor. As a result, pharmaceuticals based on biological diversity can face special challenges under patent law.

Myriad Genetics, a Utah-based biotech firm, recently lost a major court case to defend their patent on the use of human BRCA1 and BRCA2 genes for predicting risk of breast cancer in women. The Myriad testing procedure currently costs about $3000 per patient, so the Association for Molecular Pathology and the American Civil Liberties Union sought to overturn Myriad's US patent. They argued that patenting human genes is in violation of § 101 of 35 US Code. The initial court ruling, which was against Myriad Genetics, stated that "purification of a product of nature (in this case, DNA from patients), without more, cannot transform the product into patentable subject matter." Because Myriad failed to prove that there were "markedly different characteristics" between natural BRCA1 and BRCA2 genes and the isolated DNA from test subjects, the court awarded summary judgment against Myriad.

Renowned patent attorney John Wetherell notes that this case appears to echo a 1948 ruling in *Funk Bros. Seed v. Kalo Inoculant*. In this case Funk Bros. Seed company obtained a patent for a natural *Rhizobium* inoculant. In *Funk Bros. Seed v. Kalo Inoculant*, the court overturned the patent because the bacterial mixture was held to be "obvious."

If the case against Myriad Genetics is upheld on appeal, there is concern that many current biodiversity-based patents could be at risk, including those for naturally occurring compounds such as DNA polymerases, enzymes, antibodies, microbes, cell lines, stem cells, and gene replacement therapies. Any patents from natural products that are not naturally occurring or which have a different use from their use in nature, such as mutant DNA or drugs from nonhuman sources, such as Taxol, will probably remain valid. Given the massive investment required to bring new biodiversity-based products to market, loss of the Myriad Genetics' appeal will likely be the death knell of the nascent bioprospecting industry.

Patents and Traditional Knowledge Systems

Ayahuasca – the Vine of the Soul

International debates continue on the ethics of patents, drug development, public health, and human rights. Many private and public biotech and pharmaceutical firms rely on investors who are seeking a return on their investment. One of the primary ways that companies protect their large capital investments in research and development is through filing, prosecution and maintenance of patents. This has become a controversial topic in the international development, human rights, environmental and indigenous rights community. The controversy is based in part on the historical lack of appropriate access and benefit-sharing (ABS) with the people, cultures, and countries that have been the origin of multiple new therapeutic agents.

The controversy revolves around traditional knowledge, particularly historic uses of plants. In 1986, Loren Miller, a California entrepreneur, obtained US plant patent 5751 for a cultivar of ayahausca, *Banisteriopsis caapii* (Malpighiaceae), a hallucinogenic plant used ritually in South America. He termed his cultivar "Da Vine," and claimed he collected it growing in a domestic garden in the Amazon. When indigenous tribes in the Amazon learned of the patent, they protested. In this they were supported by CIEL (the Center for International Environmental Law) in Washington, DC. Representing the tribes, CIEL argued in 1991 that the patent on the "Da Vine" cultivar should be rejected because it was not new, the plant was obtained in an uncultivated state, and that issuing a patent on a plant that is sacred to indigenous people "violated notions of public policy and morality." On May 28, 1999, the US Patent and Trade Office (PTO) rejected Miller's patent on the narrow grounds that the cultivar was indistinguishable from herbarium specimens at the Field Museum made prior to the filing date of the patent. However, Miller appealed the ruling. On April 17, 2001 the PTO affirmed that both the cultivar "and its medicinal properties" were patentable and reinstated the patent for the remaining two years of the patent life. The plant patent protected only Miller's original specimen and plants produced from its cuttings.

Fortunately national patent offices are being provided links to existing databases on the traditional use of biodiversity so that there are fewer inappropriate cultivar patents issued. In addition, detailed national databases on the traditional knowledge and use of plants are being created by governments in areas with a rich and ancient history of traditional use of plants, such as India and China.

Patent Limits and Costs

In general patents are valid for 17–20 years. Companies sometimes attempt to file additional patents to extend the time line of the patent estate, a process often referred to as "evergreening a patent position" but those efforts are not always successful. Many indigenous cultures object on ethical grounds to patenting any compounds, plants, or processes that are linked in any way to their traditional knowledge and

cultures, especially if there is no prior informed consent or benefit sharing linked to the patented material. It is important to note as well that many indigenous cultures focus on the well being of their people, communities, and ecosystems on timelines of 500 years or more. Once a patent has expired after 20 years it is forever in the public domain and can be used freely by anyone. In that sense a 20-year period of time seems to be a short time span when compared to the larger time horizon of indigenous communities.

Patent Pools

Application and maintenance of a patent or several related patents often costs several hundred thousand dollars in legal and government fees over a period of many years. The ownership of a patent in and of itself does not bestow or have an actual value unless it protects a product or process that has some commercial utility. This means that many companies, large or small, nonprofits or individuals may expend large sums of money to obtain patents that never have any use or value. In patents related to bioprospecting, few products actually are ever developed or commercialized. There is an incorrect perception that once a patent is issued that there will be a rapid return of financial benefit, which is usually not the case.

Fortunately there is a movement among many for profit and nonprofit companies to make existing patents on natural and synthetic products available in "patent pools" for free access in order to facilitate the development of drugs and vaccines for neglected diseases. Many universities, nonprofit research foundations and companies are utilizing these patent pool resources to accelerate the development of therapeutics for diseases such as malaria, TB, chagas disease, cholera, and many other debilitating infectious diseases.

Bioprospecting and Ethics

International Agreements

The Convention on Biodiversity

The sustainable development of genetic resources and equitable sharing of benefits are core principles of the Convention on Biological Diversity (CBD) that was ratified in 1993 by most countries of the world. The purpose of the CBD was to incentivize biological conservation by encouraging partnerships between biodiversity-rich nations and technologically rich countries. The major challenges related to sustainable development and bioprospecting revolve around the implementation of these core principles at the national and international level.

Use of Indigenous Knowledge under the Convention on Biodiversity

Article 8(j) of the Convention on Biodiversity obligates the parties to preserve indigenous knowledge, find broader applications for its use, and to ensure that the benefits of the knowledge are equitably shared with the traditional custodians of that knowledge. Yet the rights and benefits that should be accorded to indigenous and local communities when scientists, whether academic or corporate based, conduct research in foreign nations has become a contentious topic (Brown 2004).

One of the unintended consequences of the CBD was a more complicated and slow process for taxonomists, ecologists, and other biologists to obtain permits to conduct basic research in biodiversity-rich nations. In many countries it is far easier to obtain permits to clear cut entire forests than to collect a 50 mg sample of a plant for study. Some individuals in biodiversity-rich countries feared that all scientific researchers were essentially agents of bioprospecting organizations seeking to plunder and exploit their natural resources. This was largely a false obsession. Individual scientists, companies, or government consortia such as Shaman Pharmaceuticals, Diversa, the National Cancer Institute and the International Cooperative Biodiversity Groups (ICBGs) sought even prior to the CBD to establish ethical research and benefit-sharing relationships.

An Example of Pre-CBD Benefit-Sharing Agreements: Prostratin

As previously described, Prostratin is being developed by the AIDS Research Alliance as a possible treatment for HIV/AIDS. Originally discovered in the 1970s in species of *Pimelia* (Thymelaceae) which had proved toxic to livestock in New Zealand and Australia, Prostratin was again isolated in the mid-1980s from *Homalanthus nutans* in Samoa. The stem bark of *H. nutans* was recorded by ethnobotanist Paul Cox (then at Brigham Young University (BYU)) to be used by two separate healers to treat hepatitis. After first negotiating benefit-sharing agreements with the healers, the village chiefs, and the government of Samoa, Cox provided samples of *H. nutans* to the National Cancer Institute (NCI) to run against an *in vitro* test for HIV/AIDS. These negotiated agreements were formalized in 1989 (four years before the CBD) as "The Falealupo Covenant." With permission of the village chiefs and the Samoan Prime Minister, a use patent for Prostratin as an antiviral compound was filed by the NCI and BYU with the proviso that any commercial developer of Prostratin would first be required to negotiate a fair and equitable return with the Samoan people before obtaining a license from NCI. Publication of this requirement in the *Federal Register* by NCI resulted in benefit-sharing agreements negotiated between the AIDS Research Alliance and Samoa in 2001 and the subsequent agreements between UC Berkeley and Samoa in 2004, all of which were signed by the village chiefs and the Prime Minister of Samoa, after review by the Samoan Attorney General. In accord with the agreements, water tanks, a small medical clinic, a trail system, three schools, and an aerial rain forest walkway, with aggregate value exceeding $500,000, were constructed in the village (Cox, 2001). All of the tourist revenues from the walkway are retained by the village. Cox was honored with a chiefly title, Nafanua, by the villagers. His efforts to conserve the Falealupo rain forest were recognized by a Goldman Environmental Prize shared by Cox and village chief Fuiono Senio in 1997.

In 2009, Cox was attacked in the Samoan press by a US NGO (nongovernmental organization), M-Cam, who claimed that the Prostratin use patent was illegal since the Samoan healers were not listed as co-inventors. M-Cam also argued that the benefit-sharing agreements signed by the villagers and

the Prime Minister were of no value. Although the NGO claimed to represent the healers, neither healer families nor the village chiefs said they were familiar with the NGO, which persisted in misspelling the healer's name in their press releases. Finally, in 2010 the village chiefs issued an extraordinary declaration in both Samoan and English, which was published in the national newspaper and signed by 87 village chiefs. In this statement, the chiefs reaffirmed the validity of the original benefit-sharing agreements, declared that Cox held a chiefly title and enjoyed their trust, and asked that outsiders cease from interfering with what they considered to be village matters.

Bioparasites and Indigenous Knowledge

Implementation of the principles of article 8(j) of the CBD when working with indigenous communities and local communities is not always easy – not because local communities do not want to partner with researchers in ways that can benefit their communities and families – but because of the emergence of a new form of NGO, termed by some as "bioparasites." Bioparasitic NGOs, supported by funds from foundations, governments, and the private sector ostensibly seek to protect indigenous peoples and local communities from exploitation by scientists, who they claim will continue the cycle of colonial exploitation of the last century. The irony is that these NGOs seeking to represent the interests of indigenous communities often function in a paternalistic and colonial manner, the presumption being that indigenous peoples cannot make their own choices and decisions, nor negotiate effective relationships with outside researchers. However, in Latin America, Africa, Southeast Asia, and the South Pacific, there are strong examples of indigenous communities that know what they want and are effective at negotiating, communicating, and managing partnerships with biodiversity researchers. Some bioparasitic NGOs have urged countries to totally close their doors to scientific research. In countries with rapid loss of biodiversity due to logging, mining, and agricultural expansion such closure would in essence guarantee that their biodiversity will be lost before it is ever screened for useful compounds.

Problems became more pronounced when the bioparasites began to siphon sufficient resources to travel all over the world "on behalf of and representing" indigenous and local communities at meetings focusing on access and benefit sharing. When queried, many of the bioparasitic NGOs are unable to speak the local language, are unacquainted with the local customs, and are unable to name or even spell the name of a single villager.

Concomitant with the rise of bioparasitic NGOs, the actual indigenous community leaders were, at least in the first 10 years after the ratification of the CBD, often not included at important international forums on ABS. To complicate matters further, many national governments did not accord the indigenous and local communities a legal voice or stake in the sharing of what was deemed under the CBD as sovereign national resources.

Ethnobiologists, with long standing relationships and respect for indigenous and local communities, have become new targets for bioparasitic NGOs. As a result several very sound and long standing effective relationships and programs, such as the ICBG program of the highly respected husband and wife team of Brent and Elois Ann Berlin were terminated by the US Government, due to a successful bioparasite effort that divided indigenous communities in Southern Mexico.

Effective Indigenous Negotiators

Bioparasitic NGOs do not always prevail, however. For example, the indigenous Ingano people, in an effort to produce a sustainable harvest and management program for *Croton lechleri* latex with tangible benefits for their communities, successfully negotiated (with local government approval) with the Colombian FARC. These negotiations with a militarized group represented an extraordinarily sophisticated ability on the part of the Ingano people to effectively pursue community benefits despite significant risk. In this case the bioparisitic NGOs did not interfere, possibly due to the strength of the indigenous communities own management of their choices, fear of the FARC or out of appropriate respect for the decision making process of the Ingano people.

The 2010 COP(10) Meeting in Nagoya

In October 2010, an important milestone was achieved at the 10th meeting of the conference of the parties (COP(10)) of the Convention on Biological Diversity, in Nagoya, Japan. During this meeting an agreement on ABS, called the "Nagoya Protocol on Access and Benefit Sharing" was approved, along with several other measures to help halt the loss of global biodiversity. The Nagoya ABS protocol provides greater details and specific requirements on benefit sharing and access agreements. Likewise, indigenous groups have sought stronger national and international recognition of their rights to resources on their ancestral lands and to their traditional knowledge. Hopefully, the new ABS protocol will facilitate achievement of these goals. Scientists will now have a simpler process to gain access for academic research as national governments streamline the permit process as part of the Nagoya ABS protocol. Each country is required to have a focal point for the process of requesting biodiversity access permits. The requirements for prior informed consent and the establishment of benefit-sharing agreements have been reinforced.

There was a desire by some parties to have the Nagoya protocol become retroactive but that was not approved. Instead as a compromise there was a call for a "Global multilateral benefit-sharing mechanism" to address genetic resources that were acquired prior to the new agreement.

Data Return and Public Health

One of the other ways bioprospecting can enhance local communities is through data return programs when scientists send back data on the results of their research. Providing data on the safety and efficacy of plant medicines to integrated clinics and rural health programs, can enhance traditional medicine, which is especially important in areas with limited access to "western medicine." In one example described by Scheinman (2000) in Tanzania, data provided by Shaman Pharmaceuticals helped the Tanga Aids Working Group expand the use of a plant that was demonstrated to have *in vitro* efficacy for herpes. In this case, traditional healers and western trained medical doctors collaborated in the treatment and management of people living with HIV/AIDS within and

outside the hospital setting. Shaman Pharmaceuticals also returned data on the *in vivo* activity of plants used to manage type 2 diabetes to healers and collaborating research groups in Asia, Africa, and Latin America (Richter and Carlson, 1998).

Books, manuals, and school materials on plant medicine have also been generated and returned to collaborating healers, clinics, and research groups in Belize and Micronesia by Michael Balick at the New York Botanical Garden, while Cox has provided copies of Manandhar's *Plants and People of Nepal* to every school and community library in Nepal. Studies of traditional medicine and health care practices can provide long-term documentation on the botanical, cultural, and medical uses that can be integrated in rural and urban health care and conservation programs (Lee *et al.*, 2010; Balick and Arvigio, 1998).

Benefit Sharing in Practice

There are examples of benefits sharing programs that have been practiced for several decades, in some cases, as part of collaborative research and development activities with communities, indigenous peoples and governments.

Shaman Pharmaceuticals

In the early 1990s, Shaman Pharmaceuticals sought to develop new drugs through ethnobotanical research collaborations with indigenous communities in 25 countries in Latin America, Africa, and South East Asia. Early in the formation of this company, a mechanism for the return of benefits was established: a nonprofit organization named the "Healing Forest Conservancy" (HFC) (King, 1994). The collaborating communities agreed that benefits from any drug that might be developed should be shared equally among the collaborating cultural groups, as it would be impossible to predict in advance where a successful drug lead might come from. It was also agreed that the time line for benefit sharing in the drug development process, at times 10–15 years or longer, was unacceptable to indigenous communities. Providing immediate and medium term reciprocity was therefore incorporated into the benefit-sharing process, in addition to the long-term benefits resulting from drug discovery (King *et al.*, 1996).

The company committed 10–15% of each field research budget to fulfill benefits as determined by the local community. Examples included clean water systems/piping for groups in Ecuador and Indonesian Borneo, beds and mosquito nets for a maternity clinic in Tanzania, a boat and engine for emergency medical evacuations for healers in the Sese, Islands of Uganda in Lake Victoria, funds for land demarcation for the Matses people in Peru, visits of dentists to a remote Quichua village in Ecuador, the extension of a dirt airstrip in the same village to facilitate visits of dentists and emergency evacuations of critically ill people, and support for an Ingano medicinal plant community garden in Colombia (Moran *et al.*, 2001).

The HFC also conducted workshops with local people, supported conservation projects and executed a pilot return of resources to test the long-term benefit-sharing process in Nigeria. The pilot benefit-sharing process was published as a case study by the Secretariat of the Convention on Biological Diversity (Moran, 1998). Shaman Pharmaceuticals did not succeed but the research programs and lead clinical drug assets, such as the antidiarrheal compounds previously described, were purchased by Napo Pharmaceutical Inc. in 2001. The long-term benefit-sharing commitment of the Healing Forest Conservancy and Shaman Pharmaceuticals was formally adopted by the board and management of Napo Pharmaceuticals and that drug is in the final stages of development and commercialization in 2011.

Conservation, Sustainable Development, Fair Trade, and the Management of Products Derived from Biological Diversity and Traditional Knowledge

Hundreds of consumer products worldwide have been developed and marketed that were derived from traditional knowledge. There is a need to ensure that the harvesting and cultivation of these products is conducted in a manner that is sustainable and provides fair trade income for individuals, families, communities, and cooperatives. There are now multiple international organizations that evaluate and certify such practices so consumers can elect to utilize products with maximum benefit to ecosystems and communities.

These general principles are evolving to be the guiding structure of ethical trade and are articulated in the Biotrade Initiative of the United Nations Conference on Trade and Development (UNCTAD):

- To work to conserve biological and cultural diversity
- To maximize the sustainable use of biological diversity
- To facilitate the fair and equitable sharing of benefits from the use of biological diversity
- To focus on socio-economic sustainability, product, financial, and market management
- To comply with national and international regulations
- To conduct all work with respect for the rights of all stakeholders involved in the sustainable management
- To respect land tenure, use and access to natural resources and knowledge

Companies are asked to reaffirm their commitment to the following objectives:

- To create social and environmentally responsible management systems
- To support long-term ecosystem conservation
- To maximize wildlife protection and water conservation
- To provide fair treatment and good working conditions for all, with a focus on the occupational health and safety of local communities
- To focus on long-term community relations
- To create, manage and maintain integrated plant and animal production systems
- To enhance and improve soil management and conservation of microbial diversity
- To focus on maximizing the sequestration of carbon by utilizing whatever methods available to capture and retain CO_2 as part of the sustainable management and harvesting of any and all commercially traded plant or animal species.

Companies engaged in marketing products that are wild harvested are also encouraged to meet Fair Wild Standards:

- To maintain wild plant resources
- To prevent negative environmental impacts
- To comply with laws, regulations and agreements associated with wild harvesting
- To respect customary rights
- To engage and promote fair contractual relationships between collectors and the companies
- To limit participation of children in wild collection activities
- To help facilitate benefits for collectors and their communities
- To apply, develop and continually evaluate environmentally responsible ecosystem management practices
- To apply socially responsible business practices and promote buyer commitment.

Intellectual Credit for Discoveries Based on Traditional Knowledge

There is a consistent message from healers, village leaders, and other indigenous peoples. They ask that they be accorded the proper intellectual credit for their traditional knowledge. Unfortunately, most pharmacy or medical textbooks do not provide credit to traditional knowledge as the origin of many important medicines. High school and lower grade books also do not identify the tremendous contributions that traditional knowledge has provided to the western pharmacopoeia. Such recognition can also enhance the integration of traditional medical systems with national public health programs, since the World Health Organization estimates that 80% of the world's populations rely on traditional health care practices.

Bioprospecting and the Sacred

Indigenous leaders and ethnobiologists, such as Maui Solomon, have expressed a deep concern about the codification of traditional knowledge. One of their primary concerns is that the profound spiritual linkage of traditional peoples to the earth and their indigenous cosmologies are being unduly commoditized. For example, in Australia, indigenous communities have protested the appropriation of their iconic symbols for the manufacture of trendy clothing and handbags; they have sought in a type of trademark protection to prevent profaning their sacred symbols. Similar concerns are arising among Native American communities in the American southwest, where petroglyphic symbols considered by some to be sacred, such as Kokopelli, now adorn beer can holders. Recognition of the sacred among indigenous cultures by researchers from an increasingly secularized west is one of the most profound issues facing biodiversity research. Although there is tremendous diversity among indigenous cultures, all appear to agree that the earth and its biodiversity are sacred. Individuals involved in bioprospecting programs should show proper respect for the spiritual values of traditional peoples. What is required is a sincere respect for the cultures, people, and spiritual values embodied in traditional knowledge and ecosystems. As articulated by noted botanist and conservationist Peter Raven "we need to show love, compassion and respect for all creatures on this planet."

See also: Agriculture, Sustainable. Biodiversity and Human Health. Biodiversity as a Commodity. Biodiversity, Human Well-Being, and Markets. Biofuels and Biodiversity: The Implications of Energy Sprawl. Diversity, Molecular Level. Economic Value of Biodiversity, Measurements of. Ethical Issues in Biodiversity Protection. Ethnobiology and Ethnoecology. Government Legislation and Regulations in the United States. Human Impact on Biodiversity, Overview. Indigenous Peoples and Biodiversity. Justice, Equity and Biodiversity. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production. Market Economy and Biodiversity. Pharmacology, Biodiversity and. Plant Sources of Drugs and Chemicals. Property Rights and Biodiversity. Religious Traditions and Biodiversity. Sustainability and Biodiversity. The Value of Biodiversity

References

Balick MJ and Arvigio R (1998) *Rainforest Remedies*. Twin Lakes, Wisconsin: Lotus Press.

Brown MF (2004) *Who Owns Native Culture?* Cambridge, MA: Harvard University Press.

Buenz EJ, Schnepple DJ, Bauer BA, *et al.* (2004) Techniques: Bioprospecting historical herbal texts by hunting for new leads in old tomes. *Trends in Pharmacological Sciences* 25: 494–498.

Cox PA (1979) Use of indigenous plants as fish poisons in Samoa. *Economic Botany* 33: 397–399.

Cox PA (1998) The promise of Gerard's Herball: New drugs from old books. *Endeavour* 22: 51–53.

Cox PA (2001) Ensuring equitable benefits: The Falealupo Covenant and the isolation of anti-viral drug Prostratin from a Samoan medicinal plant. *Pharmaceutical Biology* 39: 33–40.

Eadie L and Ghosh TK (2011) Biomimicry in textiles: Past, present and potential. *Journal of the Royal Society.* Interface published online 16 February 2011. doi: 10.1098/rsif.2010.0487.

King S, Carlson T, and Moran K (1996) Biological diversity, indigenous knowledge, drug discovery and intellectual property rights. In: Brush S and Stabinsky D (eds.) *Valuing Local Knowledge*, pp. 186–208. Washington, DC: Island Press.

King SR (1994) Establishing reciprocity: Biodiversity, conservation and new models for cooperation between forest-dwelling peoples and the pharmaceutical industry. In: Greaves T (ed.) *Intellectual Property Rights for Indigenous Peoples, A Sourcebook*, pp. 69–82. Oklahoma City, OK: The Society for Applied Anthropology.

Lee R, Nieve Shere LA, Balick MJ, *et al.* (2010) *Pohnpei Primary Health Care Manual. Health Care in Pohnpei, Micronesia: Traditional Uses of Plants for Health and Healing*. Charleston, SC: CreateSpace.

Mash DC, Kovera CA, Pablo J, *et al.* (2000) Ibogaine: Complex pharmacokinetics, concerns for safety, and preliminary efficacy measures. *Annals of the New York Academy of Sciences* 914: 394–401.

Moran K (1998) Mechanisms for benefit-sharing: Nigerian case study. In: *Case Studies on Benefit-Sharing Arrangements, Secretariat, Convention on Biological Diversity, Fourth Conference of the Parties (COP 4)*. Bratislava, Slovakia, pp. 1–20.

Moran K, King SR, and Carlson TJ (2001) Biodiversity prospecting: Lessons and prospects. *Annual Review of Anthropology* 30: 505–526.

National Research Council (1992) *Neem: A Tree for Solving Global Problems*. Washington, DC: National Academy Press.

Ragone D, Tavana G, and Stevens JM (2004) Nomenclature of breadfruit cultivars in Samoa: Saliency, ambiguity, and monomiality. *Journal of Ethnobiology* 24: 33–49.

Richter RK and Carlson TJ (1998) Reporting biological assay results on tropical medicinal plants to host country collaborators. *Journal of Ethnopharmacology* 62: 85–88.

Scheinman D (2000) An integrated program for developing medicinal plants: A case study from Tanga, Tanzania. In: *Proceedings of Medicinal Plants Forum of Commonwealth Africa*, December 5, 2000. Cape Town, South Africa, pp. 1–15.

Birds, Biodiversity of

Jeremy JD Greenwood, British Trust for Ornithology, Thetford, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 489–520, © 2001, Elsevier Inc.

Glossary

Biomass The total dry mass of a population or community of animals or plants.
Congener A species that belongs to the same genus as the one under discussion.
Elfin woodland Small, stunted tree growth characteristic of forests at higher elevations in warm, moist regions.
Endothermic An animal whose body temperature is largely determined by its own metabolism rather than by the temperature of the environment.
Epiphyte A plant that uses another plant, usually a tree, for support but not for nourishment.
Niche The unique position occupied by a particular species in terms of the function that it performs within the community.
Polyandrous A female that has more than one (male) mate.
Polygynous A male that has more than one (female) mate.
Trophic Relating to feeding.

Introduction

The life of birds is based on flight, which has both constrained their diversification and opened ecological opportunities. It affects their size and their abundance. Their ecology is based on the exploitation of patchy and variable resources, migration being the most obvious manifestation of this. A few birds are flightless. The taxonomic diversity of birds is limited. Across all continents, bird species diversity is greater in wooded than in forested habitats, generally being related to habitat complexity and productivity. Coexisting species eat different foods and often use different parts of the habitat. Bird species diversity is greater where there is greater habitat diversity; it is less at higher altitudes; and it is generally less as one moves from the tropics toward the poles, with strong local modifications. It may be correlated with that of other animals, but not always. Bird data have been used to test hypotheses about the causes of the latitudinal gradient in biodiversity. There is evidence that it is partly dependent on tropical regions providing more ecological opportunities, more habitable area, and more utilizable energy. No species is found all over the world. Some closely related species replace each other geographically. Some parts of the world, mainly tropical, have unusually large numbers of endemic, restricted-range species. Geological and evolutionary history has resulted in the avifaunas of different regions of the world differing in richness and taxonomic origin. Unrelated species living in different places but similar habitats have often undergone evolutionary convergence. Birds are good colonists of islands, where they subsequently evolve in the presence of few (if any) other land vertebrates. More species are found on larger islands, on those that once formed parts of larger landmasses, and on those closer to sources of colonists. Islands that become cut off from the mainland or which are fragmented lose species because extinctions are not then sufficiently matched by colonizations. Humans have caused the extinction of many birds, in both historical and prehistorical times, especially on islands. The chief current threats to them are habitat loss, hunting, agricultural intensification, and the introduction of alien species to islands. Reducing these threats will be difficult. Avian biodiversity continues to be lost both by large-scale declines of many species in some parts of the world and by the introduction of species without their native ranges, which diminishes the differences in the bird communities of different lands.

Flight: Constraints and Opportunities

Birds as Flying Machines

The feather is one of the most remarkable structures ever to have evolved. It has allowed birds to become highly effective flying machines and to exploit resources in ways impossible for other vertebrates, although it has also placed constraints on them, considerably restricting their morphological diversification and thus limiting the ecological opportunities open to them.

Feathers provide highly effective insulation. Simply stated, the downy bases of the contour feathers trap a layer of air close to the body, and the integrity of this layer is preserved by the vanes of the feathers, which form a largely windproof and waterproof covering, like the tiles on a roof. Weight for weight, birds' plumage is considerably warmer and more waterproof than the hair and wool of mammals. Like mammals, therefore, birds are enabled to be endothermic, usually maintaining body temperatures well above their surroundings and thus leading active and independent lives.

The structure of feathers allows them to form strong but lightweight aerofoils, thus facilitating the evolution of wings. Since the wings are made up of many feathers, they are very adjustable, allowing effective and maneuverable flight. Because feathers have no blood supply, birds do not lose inordinate quantities of heat through their wings, as bats are in danger of doing. The bat wing is difficult to repair; in contrast, a damaged feather is renewed at the next molt. The strength of feathers allows birds to have long tails, which are used for both aerodynamics and social signaling, even though the

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00015-0

bony and fleshy tail on which those feathers are mounted is reduced to a short stub.

The shape of wings varies considerably. Long, pointed wings provide high speed but little maneuverability and lift; they are characteristic of birds that need to fly fast or migrate long distances and which live in open places, such as plovers *Pluvialis*. Short, rounded wings give maneuverability and lift but little speed; they are characteristic of woodland birds such as tits *Parus*. Large, broad wings allow birds to soar on rising air, covering long distances while patrolling the landscape for food; they are characteristic of large hawks, eagles, and vultures. Long, straight wings allow birds to glide on ocean winds; they are characteristic of albatrosses and other such seabirds.

Most vertebrates tend to have long, flexible bodies. Not so with birds: Flight requires a short, rigid trunk. This is achieved not only through the trunk being short and broad but also through the locking together of the skeletal elements, providing a firm base for the operation of the flight muscles while keeping skeletal weight down. Long bones are hollow, reducing weight. The compactness of the body is increased by having the major muscles of both wings and legs close to the center of mass, rendering locomotion mechanically efficient. Because the trunk is rigid, the neck must be long so that the head can be moved around. (The length of the neck is generally not apparent because of the covering of feathers.) Were birds to have a tidal airflow ventilating the lungs, this would be a major disadvantage, because in such a system there is a "dead space" lying between mouth and lungs into which air is drawn but never used. Birds have a through-flow so that all inspired air passes through the lungs. This is why they do not need to breathe as often as mammals of comparable size, even though they operate at the high metabolic rate that flight demands. Birds' hearts are larger than those of similarly sized mammals to underpin their rapid metabolism.

The long neck means that it is particularly important for the head to be light. Modern birds are toothless and do not have the heavy jaws typical of many reptiles and mammals. Although their beaks and tongues are capable of remarkable feats of manipulation (such as finches husking seeds in seconds), this means that food is largely swallowed without chewing. To compensate, most species have part of the stomach modified as a muscular gizzard within which the food is ground. Indigestible elements of the food, such as bones, hair, arthropod exoskeletons, and husks, tend to be regurgitated in pellets.

As bats demonstrate, flight does not absolutely preclude viviparity; however, birds are the only vertebrate class in which no species bears live young. Weight restrictions result in most bats producing only one young per year, which prevents them from adopting "r-selected" lifestyles characterized by rapid reproductive rates. Some birds, in contrast, may produce well over 10 young each year. In those many species in which the young remain in the nest until fully grown, the parents' ability to fly allows them nonetheless to forage over comparatively large areas.

Being so mobile in three dimensions, birds use vision and hearing as their dominant senses; they communicate mainly by sight and sound.

Size and Abundance

The constraints imposed by flight are nowhere as apparent as in the upper limit of bird size. The lifting power of wings depends on their surface area and on the cross-sectional area of the muscles powering them; however, the weight to be lifted depends on the volume of the body, so the ratio of weight to power increases with the size of the animal, imposing an upper limit on the size of flying animals. This appears to be approximately 15 kg for modern birds, exemplified by some swans, pelicans, and bustards, although some extinct species seem to have been appreciably larger. Flightless birds may be larger—more than 100 kg in the case of the ostrich *Struthio camelus* and approximately 450 kg in the extinct elephant bird *Aepyornis maximus;* however, even these are dwarfed by African elephants *Loxodonta africana* at 7 tons and blue whales *Balaenoptera musculus* at 100 tonnes. Flightless birds have no particular advantage over mammals and there are few of them.

Metabolic rate seems to impose similar lower limits on the size of birds and mammals: Kitti's hog-nosed bat is 1.5–2.0 g, shrews such as the pygmy white-toothed shrew *Suncus etruscus* 2–3 g, and the bee hummingbird *Calypte helenae* 1.6–3.0 g. The problem is that metabolic rates of small endotherms are disproportionately great, partly because their large surface to volume ratio results in disproportionate heat loss. Not only is it difficult for them to ingest enough food to maintain such high rates but also there is an upper limit to the rates imposed by the rate at which their hearts can beat. The 1200–1400 contractions per minute observed in the smallest species may be the maximum achievable.

Most birds are small—half of the species weigh less than 38 g (**Figure 1**). They are larger toward the poles (**Figure 2**), perhaps because it is easier for large birds to keep warm or because they are more able to migrate long distances. Within the same region, smaller birds are more common in wooded than in open habitats; the former give them better protection from weather and predators as well as provide many foraging opportunities among twigs and foliage, which heavy birds cannot exploit. Waterbirds tend to be large to aid heat conservation.

In many birds, males are slightly larger than females; they are particularly large in many polygynous species, which suggests that the reason for the difference is the advantage that larger males have in competing for females. Females tend to be larger in polyandrous species and also in species in which males perform aerobatic displays as part of courtship.

On a logarithmic scale, the abundances of bird species in a region are approximately normally distributed, with some skew because of a tail of very scarce breeders (**Figure 3**). The same is probably true on a world scale. The Seychelles warbler *Acrocephalus sechellensis* was once confined to Cousin Island (29 ha), its population reaching a low of 50 pairs. Numerous other species confined to small, remote islands may have maintained populations of no more than a few hundred throughout their history. Various species have total world populations much less than this but usually as a result of human interference. Wilson's petrel *Oceanites oceanicus* is probably the most common seabird, at 100 million or more. On land, the red-billed dioch *Quelea quelea*, of which individual flocks may number over 1 million, is almost certainly even more numerous; house sparrow *Passer domesticus* and starling *Sturnus vulgaris* are so

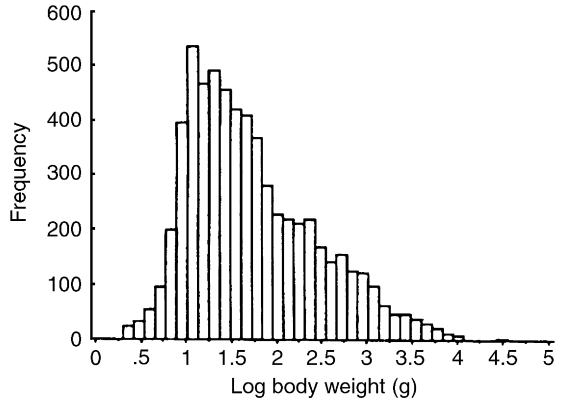


Figure 1 The frequency distribution of the logarithm of body mass of the 6209 species of birds for which data are available (reproduced from Blackburn TM and Gaston KJ (1994) The distribution of body sizes of the world's bird species. *Oikos* 70: 127–130).

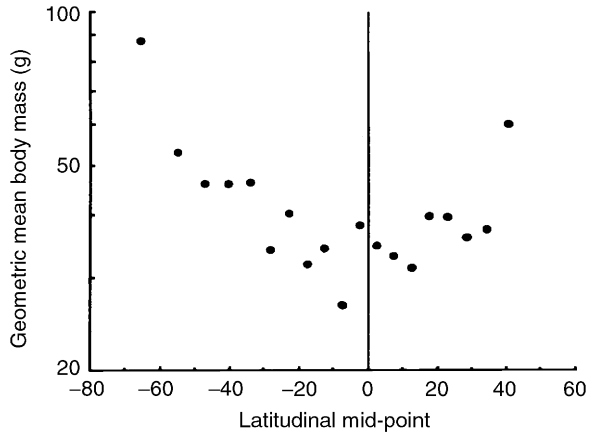


Figure 2 The relationship in New World birds between body mass (geometric mean) and latitude (negative values are in the Northern Hemisphere) (reproduced from Blackburn TM and Gaston KJ (1996a) Spatial patterns in the body sizes of bird species in the new world. *Oikos* 77: 436–446).

widespread, partly through human introduction, that they may have total populations even greater than those of *Q. quelea*, which is confined to the more arid parts of sub-Saharan Africa.

At both local and regional scales, smaller birds are more abundant than larger ones (**Figure 4**). Species that are more abundant in places where they occur occupy more sites within regions (**Figure 5**) and have more extensive geographical ranges (**Figure 6**), though on both scales colonial species (which occupy few sites but at high densities) weaken the relationship. The correlation tends to be weaker outside the breeding season, perhaps because many noncolonial species then form flocks. Various reasons for this relationship have been suggested, but the evidence favors none in particular.

In Britain at least, summer visitors tend to be smaller, on average, than residents, perhaps reflecting difficulties that small birds have in keeping warm in winter. They are substantially less abundant than residents of the same size (**Figure 4**). (In North America, summer visitors are more abundant than residents; this is also true in Europe if one does not consider size.)

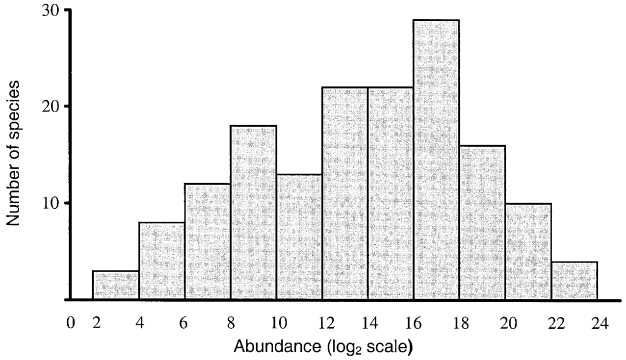


Figure 3 The frequency distribution of the population sizes of the 157 nonmarine native birds regularly breeding in Britain (British Trust for Ornithology data).

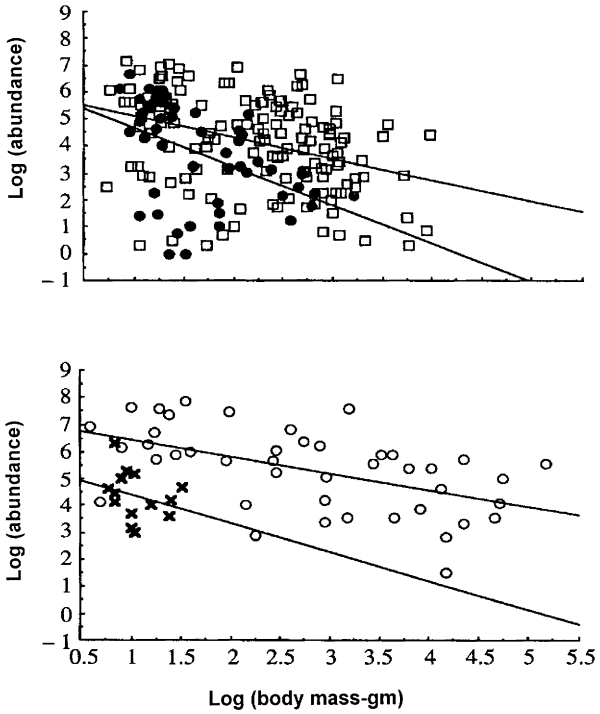


Figure 4 Relationships between abundance and body mass in non-flying mammals (○), bats (crosses), and resident (□) and migrant (●) nonmarine birds breeding in Great Britain. Note that both scales are logarithmic so that the mean difference in abundance between birds and nonflying mammals of similar size is approximately 45-fold (redrawn from Greenwood JJD, Gregory RD, Harris S, Morris PA, and Yalden DW (1996) Relations between abundance, body size and species number in British birds and mammals. *Philos. Trans. R. Soc. London B* 351: 265–278).

The Ecological Diversity of Birds

A 1-kg mammal can travel at sustained speeds of 6 km/hr, expending 118 kJ/km in doing so. A bird of the same size can travel at 47 km/hr, expending only 49 kJ/km. The power of flight thus allows birds to cover larger areas in the courses of their lives and this is the key to their ecology: They tend to use scattered and variable resources much more than

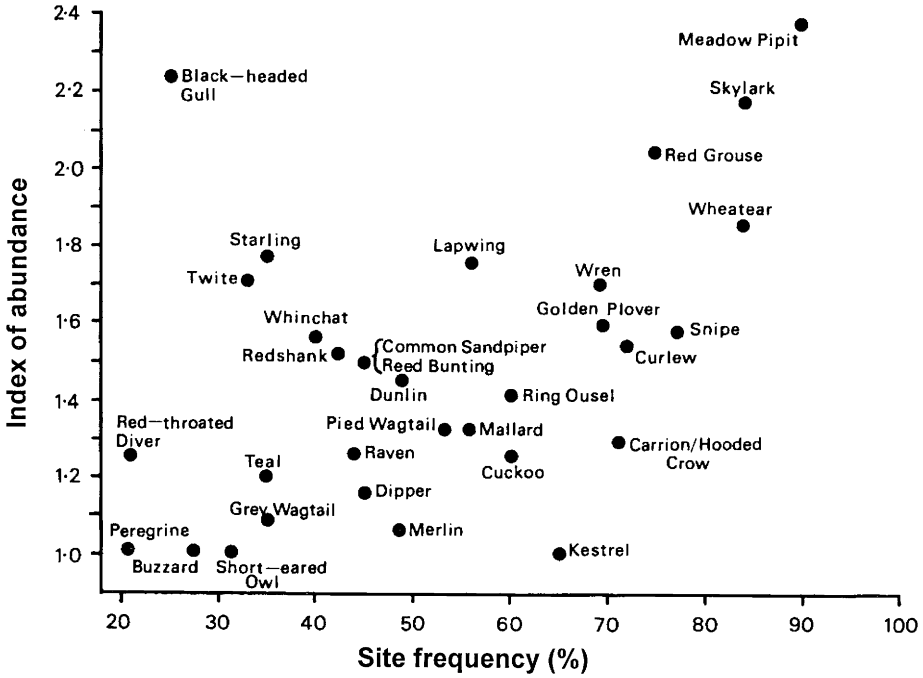


Figure 5 The relationship between the abundance of British upland birds in the sites at which they do occur and their frequency of occurrence at sites (reproduced from Fuller RJ (1982) *Bird Habitats in Britain*. Calton, UK: Poyser).

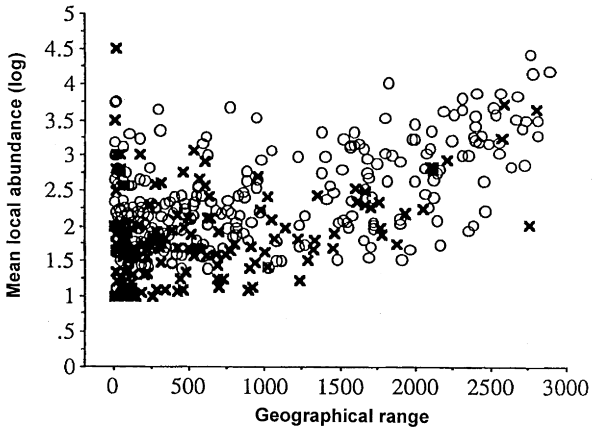


Figure 6 The relationship of mean local abundance and range size in European nonmarine breeding birds. The crosses mark species of special conservation concern (reproduced from Gregory RD, Greenwood JJD, and Hagemeijer EJM (1999) The EBCC atlas of European breeding birds: A contribution to science and conservation. *Biol. Conserv. Fauna* 102: 38–49).

do earth-bound animals. As a result, they are able to occupy significantly narrower ecological niches, subdividing resources more finely. This may be the reason why, within the same geographical area, the abundance of individual species of nonflying mammals is very much greater than that of birds of the same size (**Figure 4**). This, plus the fact that the average mammal is several times larger than the average bird, means that birds make up much less of the total biomass of terrestrial animals than do mammals (e.g., 13,000 tons for British resident birds versus 158,000 tons for mammals).

In contrast to such local restrictions on numbers, birds are able to occupy large geographical ranges because their powers of flight allow them to discover areas that provide them with the resources they require. Thus, local areas tend to be occupied by a higher percentage of the whole continental species pool of birds than is the case for other well-studied groups, such as mammals, amphibians, lizards, and butterflies.

Many waterbirds make seasonal use of water bodies that are impermanent or which freeze over in winter. On a shorter timescale, some species of birds (but few nonflying animals) move onto seashores in large numbers as tides recede, exploiting resources that are often rich but which are only briefly available. Pelagic seabirds use the efficiency of flight in another way, covering large areas in search of food that may be constantly available but is sparsely distributed. Outside the breeding season, they may move over thousands of kilometers, often along definite tracks determined by winds, currents, and seasonal availability of food. Seabirds quickly discover shoals of fish that are close to the surface, such as at upwelling fronts of cold water, which shift unpredictably. Even those land birds that are resident in a small area are constantly moving about that area in ways that allow them to subsist on food that is either sparsely distributed (such as seeds in temperate grasslands in winter) or which occurs in patches, especially temporary patches (such as carcasses, fruiting trees, nectar-bearing flowers, ant and termite swarms, insects disturbed by ungulate herbivores, and small animals disturbed by fire). Outside the breeding season they may temporarily use habitat patches that are inadequate to support them while breeding.

Birds are able to exploit aquatic habitats not just because of their powers of flight but also because of the effectiveness of their plumage for insulation. This is particularly important for smaller animals (because of their high surface to volume ratio)

Table 1 Numbers of bird families feeding on different food categories[a]

Food category	*Primarily*	*Regularly*	*Sparingly*	*Infrequently, if at all*
Green plants, buds	2	13	13	140
Seeds	4	31	24	109
Fruit	19	34	21	94
Nectar	3	5	7	152
Insects	50	58	32	28
Other terrestrial invertebrates	2	10	23	133
Littoral and benthic invertebrates	6	10	13	139
Small vertebrates (<5 kg)	4	14	29	121
Large vertebrates (5 kg +)	0	0	3	165
Fish, crustaceans, squid, etc.	20	10	11	127
Carrion	1	4	6	157

[a]Reproduced from Morse DH (1975) Ecological aspects of adaptive radiation in birds. *Biol. Rev. Cambridge Philos.* Soc. 50: 167–214.

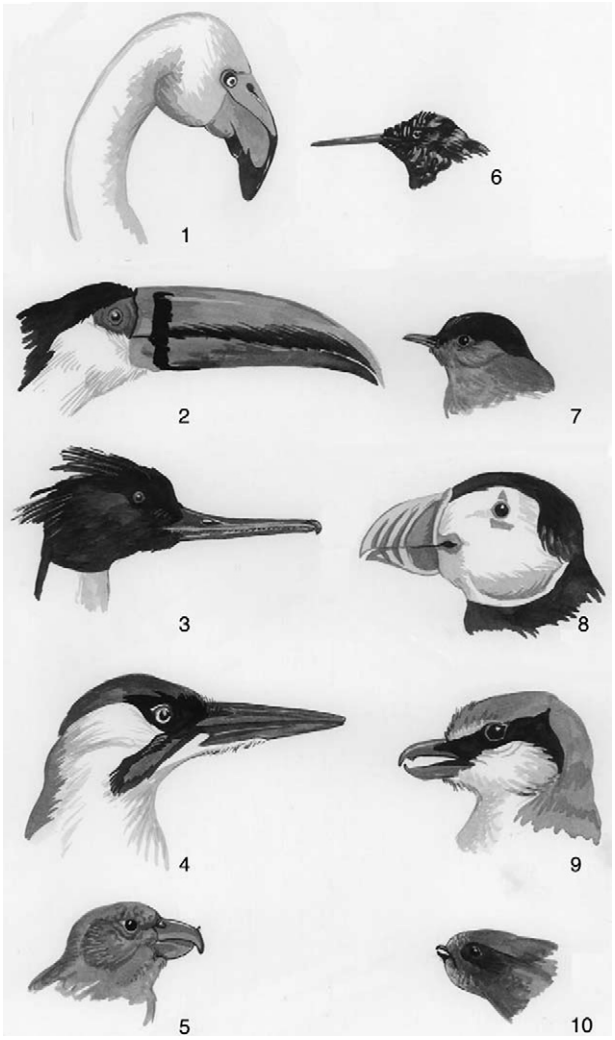


Figure 7 Examples of beaks used for various ways of feeding: 1, flamingo, modified for filter feeding; 2, toucan, elongate for reaching suspended fruit and secondarily decorated for social signaling; 3, merganser, serrated for gripping fish; 4, woodpecker, chisel-shaped for probing wood; 5, crossbill, specialized for extracting seeds from cones; 6, hummingbird, elongate for reaching nectar; 7, warbler, fine for picking up insects; 8, puffin, deepened for carrying several fish at once and secondarily decorated for social signaling; 9, shrike, hooked for tearing prey; 10, swift, wide for catching insects in flight (illustration by Su Gough).

and it is noteworthy that there are few aquatic mammals weighing less than several kilograms. Those that do must frequently leave the water to dry out their fur. Insulation is less important for larger animals, and terrestrial mammals have given rise to cetaceans, sirenians, and seals, apparently excluding birds from occupying these ecological zones.

The Australian central desert, lying between two rainfall belts, may experience occasional rain at any time, anywhere; approximately 30% of Australian birds are essentially nomadic, settling to breed quickly in places where rain has fallen. Although some desert birds can survive without drinking much free water, others cannot; rather, they are able to live in deserts by using their powers of flight to reach isolated drinking places. Sandgrouse (Pteroclididae) may make round-trips of over 100 km daily between their nest sites and water holes, with males transporting water back to the chicks that is soaked into the modified breast feathers.

Although flight allows birds to exploit patchy resources, it also restricts what they feed on. To keep weight down, birds have short guts, through which food passes quickly; therefore, they tend to feed on particularly nutrient-rich items, generally eschewing food that needs to be taken in bulk. Insects (and other terrestrial invertebrates when available), littoral invertebrates, fruit, seeds, fish, and other small vertebrates are taken by many species but, despite leaves making up such a large proportion of the biomass on land, comparatively few birds are specialist leaf eaters (**Table 1**). Some grouse, geese and some ducks, some pigeons, the New Zealand ground parrot *Strigops habroptilus*, and the hoatzin *Opisthocomus hoazin* are exceptions: Most tend to feed selectively on the most nutritious leaves and to move their food through the gut rapidly.

Carrion is very sparsely distributed; therefore, only a few birds specialize on it, although there has been convergent evolution on this specialism in New and Old World vultures. Similarly, nectar is only found in sufficient quantities for birds to use in some parts of the world and efficient nectar-feeding requires anatomical and behavioral specializations so there are few families for which nectar is the primary food (and, as with the vultures, there has been a convergence of New and Old World forms: humming birds and sunbirds, respectively).

The various diets of birds are associated with a variety of bill shapes (**Figure 7**). There are subtler variations: Frugivores with wider gapes take larger fruits, and finches with deeper bills take larger and harder seeds. Even within species, differences between individuals in bill size may be reflected in diet, the most extreme example being the African finch *Pyrenestes ostrinus* in which there are two genetically distinct forms specializing on sedge seeds that differ in hardness. The extinct

Huia *Heteralocha acutirostris* of New Zealand fed on insects in dead wood, with the female probing into borings with her long, pliant bill and the male chiseling the wood open with his shorter, stouter bill.

Migration

The most dramatic manifestation of birds' use of varying resources is migration. Huge areas of North America, Europe, and Asia become inhospitable to life in the winter, primarily because temperatures are low, the ground is snow covered, and waters are frozen; endothermic animals that could perhaps survive the physical conditions are often unable to survive in these regions because their food supplies are inadequate. Therefore, there are relatively few permanently resident birds. In the summer, however, there is an abundance of food and this presents an opportunity that is taken by hundreds of millions of birds from south temperate, subtropical, and tropical regions that move north to breed. Approximately 87% of the species breeding at 80° N in eastern North America are summer migrants, the proportion declining to 12% at 25° N. The same north–south pattern is seen in Europe but the proportion is consistently about 17% lower at equivalent latitudes because the warmer winters of western Europe allow more species to be year-round residents than in eastern America.

Not all movements are north–south. The comparatively warm (and damp) winter conditions on the western edge of Europe allow many waterfowl and shore-birds to winter there (such that British wetland sites tend to be more species rich in winter than in summer); these birds not only move north in summer but also move east, into continental regions where winter condition are intolerable. Summer visitors are generally more common in eastern than western Europe. Other species show altitudinal migrations, moving up into mountains to breed during short seasons of relative plenty. Within the tropics, savanna birds often have migrations that track the seasonally moving rainfall belts.

The seasonal variations in bird communities in particular areas may be complex. In the more benign parts of temperate regions, for example, such as the southern United States and Mediterranean Europe, winter visitors from the north and summer visitors from the south effectively replace each other. Thus, 52% of the wintering species at 25° N in eastern North America are visitors that have bred in the north; the proportion declines further north, to zero at 70° N. A similar pattern is seen in Europe, but there are about 10% fewer northern migrants at equivalent latitudes.

Summer migrants to North America mostly come from Mexico, Central America, the Caribbean, and northern South America rather than beyond the equator. These places are, if not forested, well wooded. Pale-arctic migrants come from similar latitudes, with 35% wintering wholly north of the equator and only 3% wholly south of it. In contrast to America, African migrants come almost entirely from savannas; these are not only much more extensive than forests in tropical Africa but also much more seasonal.

In much of Europe, summer visitors seem to occur particularly in grasslands and scrublands. In British and Irish woodlands, migrant passerines are most common in early successional stages and also in the upland oak-woods of Wales, where residents are less abundant than in lowland woods; migrant species are generally less abundant and less widespread than residents and use a less diverse array of nest sites. These patterns fit with residents, where they can survive the winter, preempting habitats, with migrants utilizing the summer surplus that the residents cannot fully take up.

Table 2 Percentages of birds (excluding seabirds) breeding in the western palearctic according to the way in which their range changes in winter[a]

	%		
	Land	*Freshwater*	*Coastal*
Range is unchanged	44	25	8
North edge of range moves south (range contracts)	4	8	0
South edge of range moves south (so range expands)	4	5	0
Whole range moves south but summer and winter ranges overlap	18	46	46
Whole range moves south, summer and winter ranges do not overlap	30	16	46
Total number of species	399	63	26

[a]Reproduced from Newton I and Dale LC (1997) Effects of seasonal migration on the latitudinal distribution of west Palearctic bird species. *J. Biogeogr.* 24: 781–789.

There is much less land in the Southern Hemisphere than in the north. As a result, in both Africa and South America, although there is some migration between tropical and south temperate regions, it is less dramatic than the northern migration. As in the north, relatively few of the migrants cross the equator.

Although generalizations about migration can be made, there is much variation between species in details, as is apparent even from broad considerations of patterns of migration (**Table 2**). Most migration is within-tropical, within-temperate, or tropical-temperate but there are a few species, such as American golden plover *Pluvialis dominica*, that move between temperate zones in the two hemispheres. At the extreme, some arctic terns *Sterna paradisaea* breed in the Arctic but winter in the Antarctic.

There is variation even within species: Populations (and even individuals within populations) may differ in whether they are resident or migratory and, if migratory, how far they migrate and by what routes. "Leapfrog" migration is not uncommon. In the fox sparrow *Passerella iliaca*, for example, populations of Vancouver Island and the nearby mainland are resident; birds wintering in Oregon breed in the far southeast of Alaska and northern British Columbia; and those wintering in California breed further north and west, along the Alaskan peninsula.

Flightless Birds

A few bird species are flightless. Given that it is the demands of flight that are responsible for the restricted morphological diversity of birds, it is not surprising that some of these

flightless species show marked departures from the typical avian body plan.

Many species that have colonized remote oceanic islands have become flightless. Such islands hold few, if any, other terrestrial vertebrates so there is a variety of vacant ecological niches available for birds to exploit, including those based on resources that, being neither patchy nor unpredictable, do not demand the power of flight for their exploitation. Nor is flight needed as a means of escape in the absence of predators. Among the flightless species on remote islands, rails (Rallidae) feature prominently. Rails generally lead a sedentary existence, using their powers of flight mainly for long-distance migration and to find isolated habitat patches. Island rails have typically achieved flightlessness, as have some other species (ratites in particular), by neoteny, i.e., the retention of juvenile features into adulthood. The key retained features are that the keel on the sternum (on which the massive flight muscles are inserted in flying birds) is absent (or much reduced), the legs are comparatively large, and the plumage is often looser than in typical birds, because the barbs that make up the vane of the feather lack the barbules that, in typical feathers, link adjoining barbs together to make the vane a continuous plate.

Other highly modified island forms, now extinct, were the dodo (*Raphus cucullatus*) of Mauritius and the two species of solitaire (*Pezophaps*) on Rodriguez and Réunion. Turkey-sized birds descended from doves, these species are typically pictured as exceptionally ponderous, although it has been argued that they displayed a marked annual cycle of fattening and that for much of the year they were far slimmer than usually portrayed. They had massive legs and bills and apparently occupied the niche of generalist herbivores, taking large seeds and foliage.

Apart from bats, New Zealand remained mammal free from its time of isolation from Australia 75–80 million years ago until colonized by man about 1000 years ago. In its forested landscape, moa (Dinornithidae) evolved: These were moderate to large birds, somewhat ostrich-like but much more sturdily built, and they usually held their necks horizontal rather than upright; they roamed scrubland and forests in search of seeds, fruit, and shoots. There were about 11 species, all extinguished by man. Kiwis (*Apteryx*), however, survive in small numbers. They are behaviorally and anatomically highly modified for feeding on soil invertebrates, with the nostrils on the tip of the elongate bill (unique among known birds). They are nocturnal. Apart from their locomotion being bipedal, they look remarkably mammal-like when rooting along the forest floor or across pastures in search of food. Another remarkable New Zealand species is the kakapo (*Strigops habroptilus*), an herbivorous flightless parrot; although a few other parrots are largely terrestrial, this is the only flightless species.

The moa-like elephant birds (Aepyornithidae) of Madagascar also evolved in isolation, with their final extinction apparently coinciding with the arrival of man. What of the surviving large flightless birds—the ostrich (*Struthio camellus*) of Africa, the two rheas (Rheidae) of South America, the emu *Dromaius novaehollandia* of Australia, and the three cassowaries (*Casuarius*) of Australia, New Guinea, and nearby islands? It is possible that all except the ostrich evolved in the absence of predators, although they have encountered them since; ostriches and rheas, living on open plains, can run at great speed. All have continued the bird habit of feeding on scattered food (fruit, seeds, flowers, shoots, and the more nutritious leaves, plus small animals in the case of emu and rheas), covering large foraging areas on their long legs and using the reach of their long necks to pick out food items.

The ecological gap left by the extinction of carnivorous bipedal dinosaurs led to the evolution of various giant flightless predatory birds but these seem to have been unable to survive competition from modern carnivorous mammals.

Waterbirds use two different techniques to swim under water, both of which result in a predisposition to the evolution of flightlessness. Some use their feet to drive themselves through the water. Thus, in addition to retaining the massive wing muscles required for flight, these birds must also have particularly powerful hindlimbs. More important, efficient movement through the water requires an elongate body, conflicting with the benefit for flight of a short, compact body. The powers of flight are therefore reduced in many foot-swimmers; species in isolated situations have become flightless, such as the Galapagos cormorant (*Nannopterum harrisi*) and the Lake Titicaca grebe (*Centropelma micropterum*).

The second underwater swimming technique is to use the wings, as do the auks (Alcidae) of northern seas. Since water is denser than air, the wing feathers need to be short if they are not to bend too much. Indeed, not only can the whole wing be much smaller but also it must be; otherwise, the wing muscles would not be strong enough to move it. It is not surprising that the Great Auk (*Alca impennis*) and other large species now extinct were flightless. In the Southern Hemisphere, even greater specialization for swimming occurred in the penguins (Spheniscidae). Their flippers are mostly made of muscle and bone, with the feathers on them being as short as those on the rest of the body. To prevent the bird from being too buoyant, the downy layer is not as well developed as one might expect for animals living in cold water; instead, penguins may carry much subcutaneous fat. Their bones tend to be solid. The elongation of their bodies is so great that they have to stand upright when on land. It can clearly be seen how giving up flight has liberated them from the constraints of body form that apply to most birds.

The Taxonomic Diversity of Birds

Approximately 9700 species of birds are currently recognized, which is more than twice as many as there are species of mammals. The great number of avian species is remarkable given the restrictions on morphological diversity imposed by flight, but it is undoubtedly a consequence of flight. As already considered, the mobility that flight provides allows birds to occupy finely divided niches. It also allows them to colonize isolated islands and habitat patches, where the original colonists may evolve into new species. (Note that bats, similarly equipped with the powers of flight, resemble birds in terms of their abundance/size relationships (**Figure 4**) and they are the second most species-rich order of mammals, with approximately 950 species).

Table 3 shows the orders of birds recognized in traditional classifications of birds and those determined on the basis of studies of DNA hybridization. Many of the differences are

Table 3 Traditionally recognized orders of birds and those based on DNA hybridization, with numbers of species in the latter[a]

	Traditional	*Sibley and Ahlquist*	*Species*
Rheas	Rheiformes	Struthioniformes	10
Ostriches	Struthioniformes		
Cassowaries	Casuariiformes		
Kiwis and moa	Dinornithiformes		
Tinamous	Tinamiformes	Tinamiformes	47
Grouse, pheasants, guinea fowls, and quails	Galliformes	Galliformes	214
Guans and megapodes		Craciformes	69
Swan, geese, ducks, and screamers	Anseriformes	Anseriformes	161
Woodpeckers and barbets	Piciformes	Piciformes	355
Jacamars		Galbuliformes	51
Hornbills	Coraciiformes	Bucerotiformes	56
Rollers, motmots, knigfishers, and bee-eaters		Coraciiformes	152
Hoopoes		Upupiformes	10
Trogons	Trogoniformes	Trogoniformes	39
Mousebirds	Coliiformes	Coliiformes	6
Cuckoos and hoatzin	Cuculiformes	Cuculiformes	143
Parrots	Psittaciformes	Psittaciformes	358
Swifts	Apodiformes	Apodiformes	103
Hummingbirds		Trochiliformes	319
Turacos and plantain-eaters	Musophagiformes	Musophagiformes	23
Owls	Strigiformes	Strigiformes	291
Nightjars and oilbird	Caprimulgiformes		
Doves and dodos	Columbiformes	Columbiformes	313
Sunbittern, bustards, cranes, Kagu, rails, etc.	Gruiformes	Gruiformes	196
Button-quails		Turniciformes	17
Sandgrouse	Pteroclidiformes	Ciconiiformes	1027
Shorebirds, gulls, terns, and auks	Charadriiformes		
Grebes	Podicipediformes		
Cormorants, pelecans, gannets, etc.	Pelecaniiformes		
Hawks, eagles, falcons, secretary bird, etc.	Falconiformes		
Storks, herons, ibises, etc.	Ciconiiformes		
Petrels, shearwaters, and albatrosses	Procellariiformes		
Divers	Gaviiformes		
Penguins	Sphenisciformes		
Flamingos	Phoenicopteriformes		
Perching birds	Passeriformes	Passeriformes	5712

[a]Reproduced from Sibley GC and Ahlquist JE (1990) *Phylogeny and Classification of Birds*. New Haven, CT: Yale Univ. Press, with permission from Yale University Press.

trivial: The four orders united into the Struthioniformes on the basis of the DNA evidence have always been recognized as closely related; the splits of several orders generally correspond with long-recognized subordinal divisions. The most fundamental difference is not just that 10 apparently diverse orders are united into the Ciconiiformes by the DNA evidence but also that some of these orders have not retained their integrity as major subdivisions within the combined order. Further work is needed before these changes can be accepted or rejected.

The classification based on DNA hybridization groups the orders of modern birds into two "infra-classes," Struthioniformes plus Tinamiformes (Eoaves) and all the others (Neoaves). This is a long-recognized division based on the peculiar structure of the palate and other anatomical characteristics of the first group. Eoaves are often referred to as ratites, although strictly speaking this term should not include the tinamous since it refers to the condition of the sternum in the struthioniform birds, which is flat, without the keel on which the flight muscles of most birds (including tinamous) originate.

Both the traditional and the DNA-based classifications recognize more orders of birds than the 21 usually recognized for mammals. Given that birds are substantially less diverse than mammals, this is paradoxical. It reflects the fact that taxonomic categories at the same formal level in the hierarchy are not comparable in different groups and that ornithologists have adopted classifications that specialists in other disciplines would regard as inflated. (Entomologists would probably reduce most ornithological families to the rank of genus.)

The outstanding feature of avian taxonomic diversity is that 59% of the species belong to one order, the Passeriformes. In contrast, the most species-rich order of mammals (the rodents) comprises only 42% of the species. The passerines are "typical birds." They are characterized by a particular form of the foot, well suited to perching, but have no obvious single characteristic that would appear to open up ecological opportunities in the way that the special gnawing incisors seem to have done for the rodents. Perhaps their numerical dominance is simply related to their small size, given that most birds are small (*see* Flight: Constraints and Opportunities – Size and Abundance). An additional peculiarity is that the two

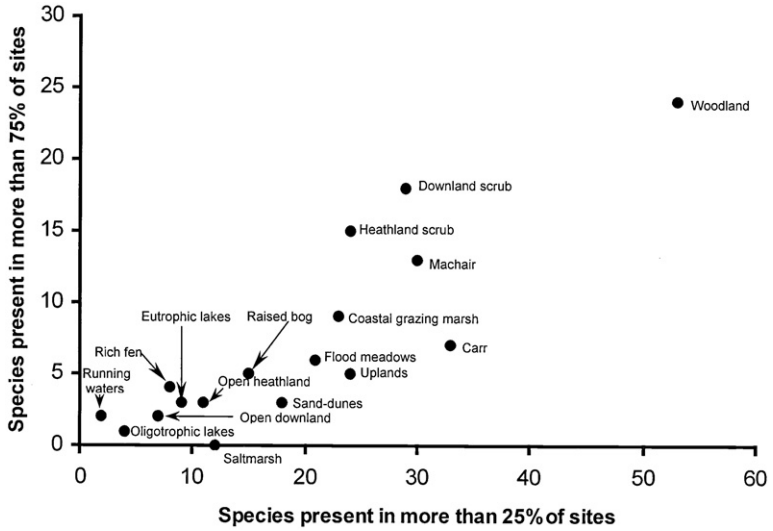


Figure 8 The numbers of species in British bird communities, according to how many are found on more than 25% of sample sites for each habitat and how many on more than 75% (from data in Fuller RJ (1982) *Bird Habitats in Britain*. Calton, UK: Poyser).

suborders of passerines are disparate in size, with 80% of the species falling into the suborder Passeri (or oscines), the true songbirds, and only 20% in the suborder Tyranni (or suboscines). Speciation in songbirds often seems to involve the evolution of song differences between formally conspecific populations. Perhaps the complex vocal apparatus of oscines has promoted high rates of speciation, leading to this suborder (and thus the order to which it belongs) being remarkably species rich.

Another peculiarity of birds, particularly if one considers that avian orders are comparable to taxa of lower rank in other animal groups, is that only a few of the orders contain very few species. The Coliiformes is the only order with less than 10 species, whereas six mammalian orders have fewer than 10. The most taxonomically isolated bird species, generally given families of their own, include the secretary bird *Sagittarius serpentarius*, a large ground-living raptor of African savannas; the hoatzin *Opisthocomus hoazin*, a leaf-eating oddity related to cuckoos which inhabits permanently flooded forests in South America, with an enormous specialized crop for processing its unusual diet; the kagu *Rhynochetos jubatus*, a strange, short-legged flightless gruiform endemic to the forests of New Caledonia; and the oilbird *Steatornis caripensis*, a Neotropical caprimulgid that feeds on fruit by night using sonar to navigate around the dark caves into which it retreats by day. Despite their peculiarities, none of these is as divergent from the rest of their class as is the aardvark *Orycteropus cafer*, placed in an order all its own among the mammals.

Habitat and Bird Species Diversity

Differences between Habitats in Bird Species Diversity

It is a matter of common observation that habitats in the same region differ in the species diversity of the bird communities living in them. Habitats with higher diversity also tend to hold larger numbers of individual birds. However, there have been few systematic surveys of bird communities across all the habitats in a region and even fewer with the extent of replication used for the data in **Figure 8**, in which each habitat was represented by at least 12 sites (with a mean of 72 sites/habitat). This confirms the following generalization: Woodlands, scrub, and carr contain large numbers of species, whereas open downland contains very few. Machair, flat plains based on deposits of calcareous sand fringing parts of the western seaboard of northern Scotland, is a conspicuous exception, having large numbers of birds despite being completely unwooded. This may be the result of the extreme small-scale heterogeneity of machair lands: There may be a transition from mobile sand dunes to stabilized grassland, from very dry to very wet conditions, and from undisturbed to cultivated ground within distances of tens of meters.

The distributions of the points on the two axes of **Figure 8** are correlated: Woodland tends to have many of the more ubiquitous species (occurring in more than 75% of the sites) and the less ubiquitous (occurring in only 25–75% of sites), whereas open downland holds few of either. The scatter around the correlation, however, reveals further differences between habitats. For example, scrub is an ephemeral habitat, typically occupied by species that are good colonists and which therefore tend to be found in the majority of sites; as a result, two-thirds of the species found in more than 25% of scrub sites are actually found in more than 75% of sites.

Bird Species Diversity in Similar Ecosystems on Different Continents

If habitat is a major determinant of bird species diversity, one would expect to find similar levels of diversity in similar vegetation and landscape types on different continents. Often one does, but sometimes one does not (**Table 4**). Even where total species numbers are similar, the numbers of species

within individual feeding categories may differ. Thus, although numbers of bird species in most feeding categories are similar in Mediterranean scrub in Provence (France), Chile, and California, Provence has no nectarivores and only 4 granivores, in contrast with 3 and 14, respectively, in California and 2 and 6 in Chile. Again, the species richness of semiarid habitats is similar in North America and Australia, but the bird communities are clearly different in detail, partly perhaps as a result of different evolutionary histories but partly because rainfall is more erratic in Australia than in North America. Thus, the habitats are similar but the overall environmental conditions differ significantly. Despite differences in species number, the number of species in the various feeding categories tends to depend more on habitat itself than on what continent the habitat is in (**Table 5**).

Table 4 Numbers of bird species in similar habitats on different continents[a]

Habitat	*No. of species*	
Peatlands	33 (Finland)	18 (Minnesota)
Desert	57 (Arizona)	61 (Argentina)
Shrub desert	5.5 (Australia)	6.3 (N. America)
Mediterranean scrub	30 (California)	28 (S. Africa)

[a]From a compilation by Schluter D and Ricklefs RE (1993) Convergence and the regional component of bird species diversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 230–240. Chicago: Univ. of Chicago Press, with permission from University of Chicago Press.

Table 5 Numbers of bird species recorded from mediterranean scrub and wet tropical forests in Africa and South America, grouped into foraging categories[a]

Foraging category	*Forest*		*Scrub*	
	Africa	*South America*	*Africa*	*South America*
Arthropods from foliage	77	55	4	3
Seeds and fruit	20	49	3	1
Arthropods from sallies	21	21	1	1
Trunk, bark, etc.	13	31	0	1
Ground	13	21	7	10
Raptors and scavengers	15	13	2	4
Nectar	1	12	1	1
Aerial and crepuscular	5	3	1	1
Total	165	205	19	22

[a]Data collated by Schluter D and Ricklefs RE (1993) Convergence and the regional component of bird species diversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 230–240. Chicago: Univ. of Chicago Press.

Bird Species Diversity and Habitat Complexity

Although it is not a universal rule, most studies have tended to confirm the discovery by R. H. and J. W. MacArthur that bird species diversity is positively correlated with foliage height diversity, at least in scrub and woodland habitats (**Figure 9**). That is, bird species diversity is greatest where there is an equal volume of foliage at all heights in the woodland. An early finding was that data points from Australia fell on the same line as those from the original study in North America. However, MacArthur discovered that data from Panamanian and Puerto Rican sites only fell on the same line if the measure of foliage height diversity was adjusted by changing the height subdivisions on which it was based. His interpretation was that Panamanian birds were effectively treating the forests as being divided into more layers than were North American birds, whereas Puerto Rican birds were treating them as being more coarsely divided. Similar variations have been found elsewhere and may be related to the productivity of the different areas. Where there are more resources available, species may divide them into narrower ecological niches; perhaps more simply, more resources allow more individual birds to live in an area and thus allow even relatively scarce species to maintain viable populations.

Along an altitudinal gradient from lowland rain forest at 400 m to elfin woodland at 3600 m in the Peruvian Andes, bird species diversity and foliage height diversity decline approximately in parallel. However, the magnitude of the decline

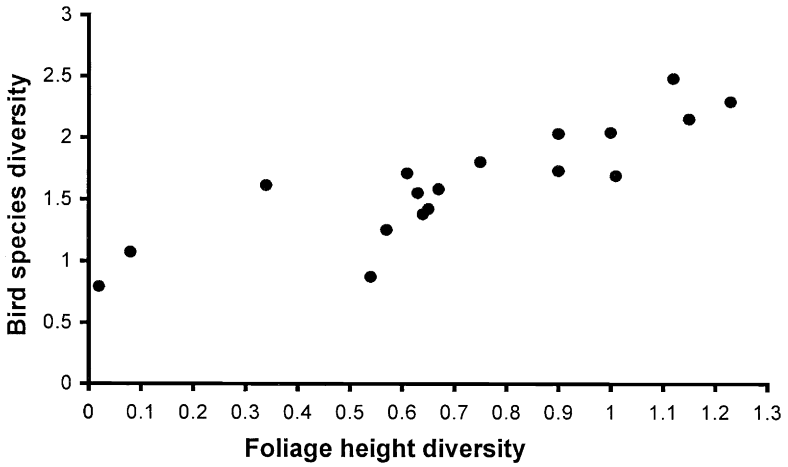


Figure 9 The relationship between songbird species diversity and foliage height diversity in Scottish woodlands. Some points are averages over 2 or 3 years. Both diversities are measured using the Shannon index, (redrawn from Moss D (1978) Diversity of woodland songbirds populations. *J. Anim. Ecol.* 47: 521–527).

differed according to the diets of the birds in question. It was more than fivefold for insectivores but only two- or threefold for frugivores, whereas the species diversity of nectarivores was similar at all altitudes. These differences may be explained in terms of the resources available to different groups: Whole groups of insects are absent at higher altitudes; less fruit is produced at higher altitudes but a higher proportion is available to the birds because there are fewer mammalian competitors than there are at lower altitudes; and because of the distribution of flowers on the vegetation, the availability of nectar only increases slightly with foliage height diversity and, although the annual production of nectar is greater at lower altitudes, it is highly seasonal and therefore can support no more birds than the less abundant but more constantly available nectar of the high-altitude forests. Thus, foliage height diversity is a crude measure of the diversity of ecological niches available to birds but it may become less satisfactory when one considers individual trophic guilds.



Figure 10 The relationship between bird species diversity and the percentage of the basal area of the forest comprising standing dead wood in Swedish woodlands (reproduced from Nilsson SG (1979) Density and species richness of some forest bird communities in south Sweden. *Oikos* 33: 392–401).

Since foliage height diversity is only a rough measure of niche diversity, it is easy to find other features of the habitat that influence bird species diversity. Thus, a survey of various sites in the Americas showed that although there was a good correlation between bird species diversity and foliage height diversity, the percentage vegetation cover of the ground also influenced the number of bird species. In other words, the total volume of vegetation was important, not just how it was structured. In hardwood forests of New Hampshire, for example, tree species differ in their value to birds: Some hold more food than others and some are easier to search (e.g., yellow birch *Betula alaghe-niensis*, with small leaves and short petioles); white ash *Fraxinus americanus* has a distinct foliage structure that allows fewer foraging methods to be used than do other species. Furthermore, floristic diversity is important: Even a few conifers in a hardwood forest allow in birds that specialize on such trees. In south Swedish forests the amount of standing dead timber (which is largely determined by how the forest is managed) chiefly determines bird species richness (**Figure 10**), with the density of low trees and shrubs having only a slight effect.

The most extreme case of the influence of factors other than foliage height diversity on bird species diversity occurs in forests in Patagonia, in which the relationship between birds and foliage diversity appears to be reversed (**Figure 11**). The reason for this is that the forests are of two quite different types and the type with the greater foliage height diversity (*Nothofagus* forest) has relatively few birds because it comprises patches that are isolated from other such forests by 2000 km of steppes and forests; furthermore, these forests are at high elevations, with harsh climates. The abundance of bamboo, which appears to support few insects, in the understory of some of the *Nothofagus* forests may also be important. Note that even these species-poor forests hold

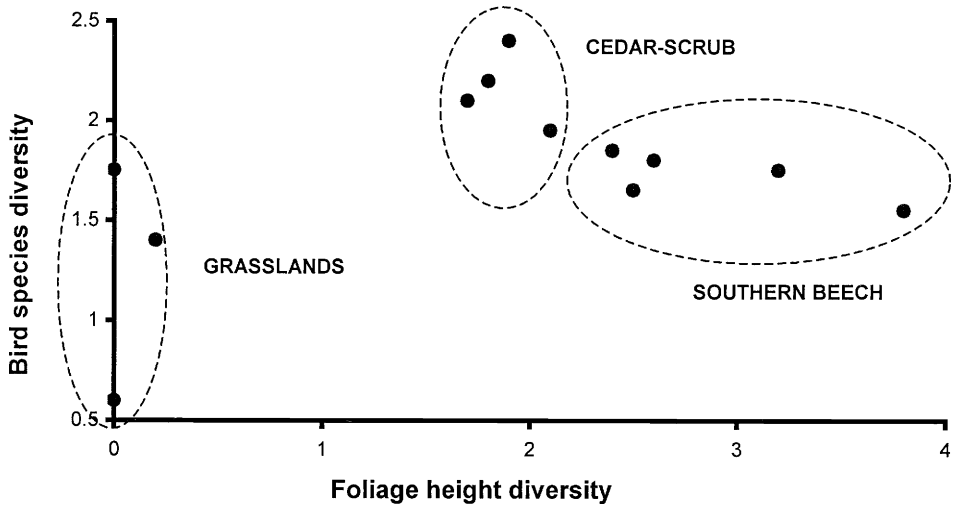


Figure 11 The relationship between bird species diversity and foliage height diversity in three contrasting habits in northern Patagonia (reproduced from Ralph CJ (1985) Habitat association patterns of forests and steppe birds of northern Patagonia, Argentina. *Condor* 87: 471–483).

more species of birds than grasslands in the same region (Figure 11), fitting the usual pattern.

Horizontal complexity is also important, with more varied places holding more bird species. Edges between habitats may be bird rich for this reason, although "edge effects" are variable both in nature and in cause. In Britain, woodland birds are generally more abundant and more diverse at edges, probably because foliage height diversity is greater there than deep in the woodlands, where the lower levels of the vegetation tend to be shaded out. On Finnish islands, bird species diversity is lower on the edges of woodlands than deeper within perhaps because of the importance of physical protection during the early nesting period, when most of the trees have not yet begun to leaf.

Foraging Niches and Habitat Structure

Classic studies of the foraging behavior of tits in broad-leaved woodlands in southern England were conducted by David Lack and colleagues. In summer, all species tend to feed largely on the glut of caterpillars in the canopy; in winters during which the mast of beech *Fagus sylvatica* is abundant, they descend to the ground to feed on it. However, during the average winter their feeding niches are distinct. The great tit *Parus major* feeds mainly on the ground on beech mast, other large seeds, and larger insects. The blue tit *Parus caeruleus*, only about half the weight of the great tit and thus much more agile, frequently hanging upside down to reach its food, feeds high up in the trees, searching for small insects among the twigs, buds, and leaves. The less common marsh tit *Parus palustris*, although no heavier than the blue tit, has a larger beak and feeds on middle-sized insects, which it gathers in the shrub layer or from the twigs and branches of trees, lower down than where the blue tit feeds. Similar studies of other birds in many parts of the world have given similar results: Related species coexisting in the same forests and using broadly similar foods tend to differ in respect of the parts of the trees in which they forage. However, foraging height is not the only factor. Thus, the coal tit *Parus ater* in English broad-leaved woodland tends to forage at approximately the same height as the marsh tit, but it is a smaller species with a finer bill and it feeds on tiny insects that it finds in crevices on the bark. In addition, it is scarce in broad-leaved woods because it is a conifer specialist; its numbers are significantly enhanced when conifers are present—another example of the importance of floristic diversity for bird species diversity.

It is such differences in foraging site and food taken that allow different species to coexist in the same habitat. Introduced birds on Hawaii provide an unplanned experimental demonstration of how differences allow coexistence: Some species persisted for more than 10 years but then became extinct; extinction was particularly likely for species that had morphologically similar relatives (presumably taking similar diets) living in the same forest. Foraging niches sometimes, but not always, shift in apparent response to the presence or absence of competitors: Willow tits in northern Finland avoid the foraging positions of great tits *P. major* and crested tits *Parus cristatus* when in mixed flocks, but if either of these is absent from a flock the willow tits use the vacated niche.

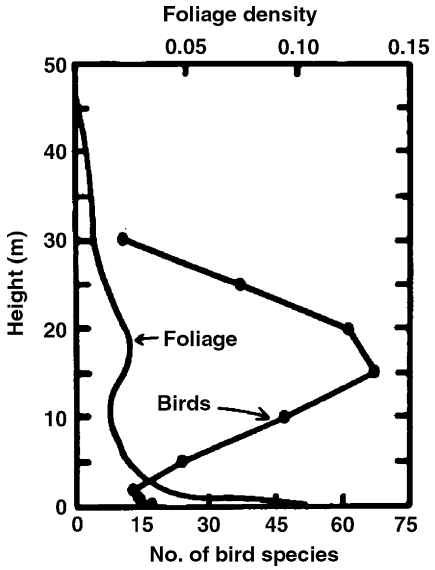


Figure 12 The relationship between bird species richness, foliage density, and height above ground within a Peruvian forest (reproduced from Terborgh J (1980) Vertical stratification of a neotropical forest bird community. *Acta XVII Congr. Int. Ornithol.* 2: 1005–1012).

Table 6 Number of species of different feeding guilds found at various heights in Peruvian forests[a]

Guild	*Height class (m)*					
	Ground	*0–5*	*5–10*	*10–15*	*15–20*	*>20*
Mast	3					3
Fruit	3	2	2	5	3	7
Nectar		6	2		2	1
Insects						
Glean	5	10	5	2	7	2
Sally	1	2	5	7	3	4
Bark			1	7	4	1
Fruit/predator				1	3	1
Fruit/insects				5	9	2
Fruit/insects/nectar					4	4

[a]Reproduced from Terborgh J (1980) Vertical stratification of a neotropical forest bird community. *Acta XVII Congr. Int. Ornithol.* 2: 1005–1012.

Different strata in forests differ not just in which species occupy them but also in their overall species richness. In the example shown in Figure 12, the forest floor, relatively uniform and simple in structure, supports few species—as does the high canopy, which is also relatively simple in structure. The middle layers support many more species because they are more complex, presenting opportunities for foraging on leaves, branches, bark, dead wood, epiphytes, termite nests, and accumulations of dead leaves in tree forks and for sallying from branches to capture insects flying in the open spaces below the canopy. Underlying the overall pattern of species richness are marked differences between feeding classes (Table 6): Mast is found either high in the canopy or fallen to the forest floor, so

mast feeders occur only in these layers; gleaners are also mostly found low down or high up, working over the continuous substrates of the forest floor or the canopy; and salliers tend to occupy middle levels, where the open structure allows them to see and pursue flying insects. The canopy is a much patchier habitat than the understory because the crowns of individual trees occupy large areas, whereas the understory is composed of much larger numbers of plants; thus, the canopy presents particular opportunities to adaptable generalists that can switch, for example, from feeding on fruit to feeding on insects depending on what is immediately available.

Bird Species Diversity and Succession

Where sheep grazing is light enough to allow it, scrub invades open hill land in Wales, eventually developing into mature sessile oak (*Quercus petraea*) woodland. This vegetational succession is accompanied by an almost monotonic increase in the numbers of both individual birds and bird species (**Figure 13**). Most of the change involves increased vertical structure, although the latest stages result from gaps developing in the woodland canopy, with small patches of open ground and scrub. As chalk downland in southern England is succeeded by ever-denser scrub, species numbers increase monotonically from 1 or 2 per study site to approximately 15 (Fuller, 1982). More generally in Britain, the scrub that develops on downland and heaths and the carr that develops on rich fens hold more species overall and more of these species are found on a high proportion of individual sites (**Figure 8**).

The increase in bird species diversity may slow in later stages of succession. In a subalpine northern American sere comprising meadow, aspen (*Populus tremuloides*), fir (*Abies lasiocarpa*), and spruce (*Picea engelmanni*), the number of bird species averaged over 3 years was 3, 17, 22, and 21, respectively. Indeed, although as yew *Taxus baccata* woodlands in southern England mature they gain species steadily (10 species at 60–80 years old and 20 species at 100–200 years), the succession from open scrub with small stands of trees to the young closed yew woodland is marked by a decline from 26 to 10 bird species. Polish deciduous forest stands developing after clear-cutting reach their maximum bird species diversity when 100–150 years old but do not do so monotonically: Following a peak at 15–20 years, diversity declines for the next 15–20 years. Local factors may be important in determining the relationship between bird species diversity and habitat gradients, such as the gradient from very low matorral (0.5 m high) to mature forest of holm oak (*Quercus ilex*): Whereas the number of bird species increases along the vegetation gradient in mainland France, it decreases in Corsica (**Table 7**) because Corsican forests are very restricted and the island simply lacks many forest species found on the mainland, where they occur in large blocks of continuous forest. An additional complication to the story of Mediterranean holm oak forests is that, although the number of bird species in an area is reduced by fire, the reduction is less than expected from the change in vegetation structure. Some characteristic forest birds persist in the open areas and matorral created by the fire either because individual birds simply occupy the same territories or because they spill over from populations in nearby unburned forest.

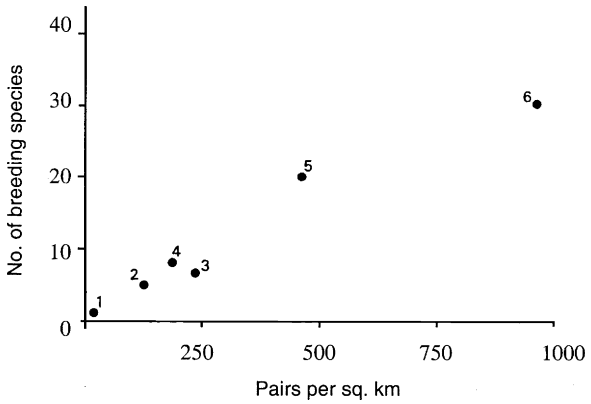


Figure 13 The relationship between the number of breeding bird species and the total population density in sessile oak woods in Wales in various stages in the succession: 1, open ground; 2, low scrub; 3, medium scrub; 4, high scrub; 5, mature woodland; 6, mature woodland with gaps and scrub (reproduced from Jones PH (1972) Succession in breeding bird populations of sample Welsh oak woods, Br. *Birds* 65: 291–299).

Table 7 Numbers of bird species breeding in Corsica and in comparable areas in mainland France in relation to the stage of succession, from 1 (very low matorral, 0.5 m high) to 6 (mature forest of holm oak *quercus ilex*, 2.5 m high)[a]

Stage	*1*	*2*	*3*	*4*	*5*	*6*
No. of species						
Corsica	29	24	32	24	23	18
France	11	15	11	19	26	23

[a]Reproduced from Blondel J (1991) Birds in biological isolates. In: Perrins CM, Lebreton J-D, and Hirons GJM (eds.) *Bird Population Studies*, pp. 45–72. Oxford: Oxford Univ. Press.

Human management can produce habitats in which the normal pattern is disrupted. For example, when English woodlands are coppiced (the trees are cut to the ground but allowed to regenerate from the stumps), the dense, bushy regrowth is colonized by numerous species, but many of these are lost as the coppice ages and the tall regrowth forms a dense canopy that shades out the majority of the lower layers, producing a uniform, structurally simple habitat.

Even where the relationship between bird species diversity and habitat follows the typical pattern, this may not be simply because structurally less complex habitats provider fewer niches. Thus, bird species richness in cleared areas in Cameroon is much less than that in the forest (**Figure 14**), but this may be largely because the sources of open-ground species are savannas and grasslands that lie well beyond the surrounding forest so that the cleared areas may not have been fully colonized. (Note that this case also illustrates that just as different trophic groups of birds differ in their response to vegetation structure so do different taxa of animals generally: Butterfly species richness is affected by forest clearance in much the same way as that of birds but leaf-litter ants are more species rich in partially cleared and second-growth forests.)

In contrast to the absolute level of species diversities, the turnover of species during succession has been little studied.

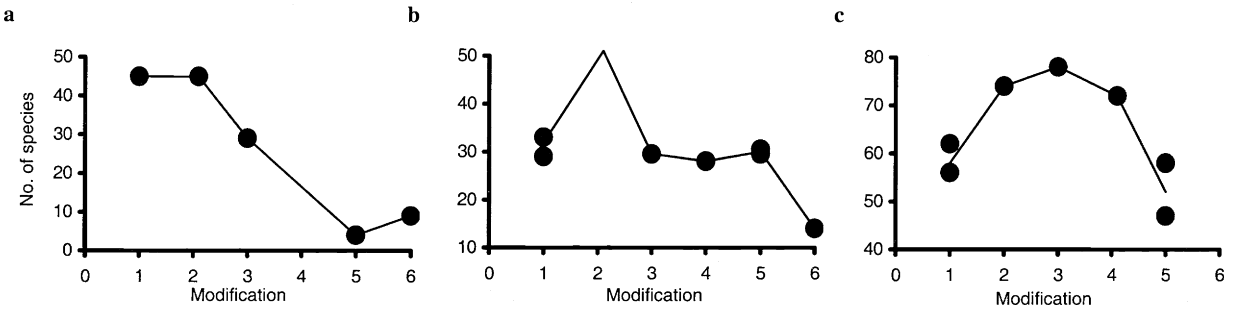


Figure 14 Species richness of (a) birds, (b) butterflies, and (c) leaf-litter ants along a gradient of increasing habitat modification in Cameroon. 1, Near-primary forest; 2, old secondary forest; 3, partial manual clearance, plus plantation; 4, partial mechanical clearance, plus young plantation; 5, complete clearance, with young plantation; 6, manually cleared farm fallow (redrawn from Lawton JH, Bignell DE, Bolton B, *et al.* (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature* 391: 72–76).

In Polish oak-hornbeam forests, there is a rapid shift in the early stages from species characteristic of open ground to those characteristic of scrub, but the rate at which species drop out of or enter the community slows as the succession progresses. In Finnish coniferous forests, in contrast, turnover peaks not at the start of succession but at 5–25 years.

Bird Species Diversity and Productivity

Many of the habitat differences that seem to influence bird species diversity are probably associated with differences in total biological productivity. Even with habitats of uniform physical complexity, such as the lakes of a region, differences in productivity are associated with large differences in the numbers of both individual birds and species present.

Geographical Variation in Bird Species Diversity

The Relationship Between Bird Species Diversity and Habitat Diversity

Different habitats have different bird communities. The greatest differences within geographical regions are between marine and terrestrial habitats since life on land and at sea requires different specializations. On land, the greatest differences are between bird communities of wooded areas and those of open country, just as these broad habitat categories show marked differences in overall species richness. However, the differences go beyond species richness: Grasslands differ from woodlands not just in that most of the woodland species are absent but also in that they have their own grassland specialists. The suites of species occurring in different habitats may be similar in size (number of species) but quite different in composition. For example, the numbers of nonmarine species occurring in more than 25% of sites in three British coastal habitats are 18 in sand dunes, 15 on rocky coasts, and 12 on salt marshes, but only 4 species are common to all three habitats and 6 common to sand dunes and rocky shores, whereas there are 7, 10, and 2 unique to sand dunes, rocky shores, and salt marshes, respectively.

Related species often differ in the habitats they occupy. For example, in Britain great tits *P. major* live in both deciduous and open conifer woods, blue tits *P. caeruleus* and marsh tits *P. palustris* only in deciduous woods, and coal tits *P. ater* almost only in conifer woods. In central Sweden, marsh tits again occupy deciduous woods and willow tits *P. montanus* live in conifers but on the nearby Åland Islands, from which marsh tits are absent, willow tits occupy both.

Given such differences, one would expect bird species diversity to be correlated with habitat diversity in an area. Although this is undoubtedly the case, it has rarely been documented. However, a study of the boreal birds occurring on isolated mountaintops and mountain chains in western North America showed that 91% of the variation in species number could be accounted for by differences in habitat diversity.

Bird Species Diversity and Altitude

The decline in bird species diversity with increasing altitude in the Peruvian Andes (*see* Habitat and Bird Species Diversity – Bird Species Diversity and Habitat Complexity) is typical not just of forested regions but also of open habitats and not just of the tropics but also of temperate and boreal regions. Everywhere, there are fewer species of birds at higher altitudes—unless the higher levels have escaped habitat destruction wrought by man in the lowlands or unless the contrast is between different habitats, such as lowland savanna and montane forest. One cause of such gradients is that the structural complexity and the floristic diversity of the habitat are reduced at higher altitudes. This is true of both forests and open habitats; indeed, forest gives way to open habitats at sufficiently high elevation.

The cooler, windier, and wetter conditions at high altitudes are no doubt responsible for the less luxuriant vegetation and probably also affect the birds directly so that only species capable of withstanding them can survive. Some mountains may also be so isolated from others that they resemble islands, lacking some species because they have never reached these mountains or because they have not recolonized following a chance extinction (*see* Diversity on Islands).

Bird Species Diversity and Latitude

As with so many other groups, there are more bird species in the tropics than toward the poles (**Figure 15**), although some groups do not conform to this generalization: Only 3 of the 17 species of penguins (Spheniscidae) breed within the

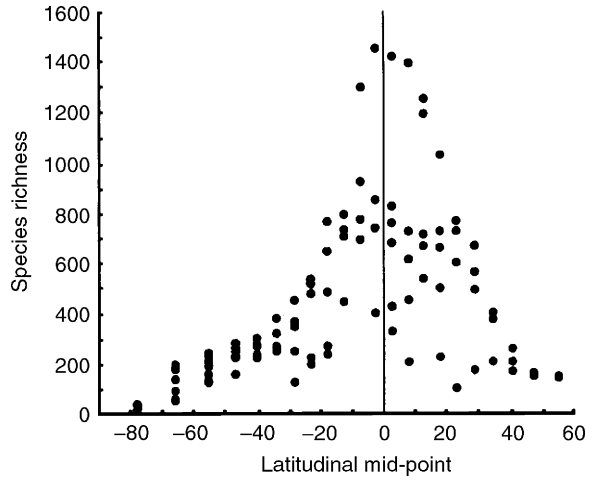


Figure 15 The relationship in the New World between the species richness of birds breeding in a square and the latitude of the square (redrawn from Blackburn TM and Gaston KJ (1996b) Spatial patterns in the species richness of birds in the New World. *Ecography* 19: 369–376).

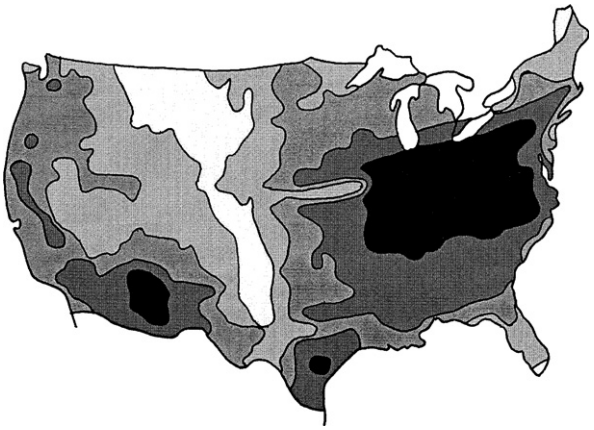


Figure 16 The distribution of species richness of insectivorous birds breeding in the United States. The depth of shading represents richness, from less than 70 species (pale) to more than 100 (dark) (redrawn from Schall JJ and Pianka ER (1978) Geographical trends in numbers of species. *Science* 201: 679–686, © 1978 American Association for the Advancement of Science).

tropics (where marine productivity is low) and sandpipers (Calidridinae) breed mainly in the boreal tundra and taiga (utilizing the summer flush of productivity, especially in wetter areas). The trend occurs in all habitats but may be steeper in forests than in savannas and prairies.

The broad trend is modified by local contingencies. In Argentina, the trend is steepest where forest gives way to savanna. In Africa, the reduction in species number as one moves away from the equator is steeper to the north than to the south, in accord with gradients of aridity and vegetation diversity. Northern Finnish peat bogs hold more breeding bird species (and individuals) than southern ones, probably because they have willows (*Salix*) and therefore are more heterogeneous than the willow-free southern bogs. In Sweden, breeding passerine species are scarcer in northern than in southern bogs but waders are more diverse, with the northern bogs having more food-rich pools. In heavily urbanized locations in Finland, there is no latitudinal gradient in species diversity in winter because the birds are wholly dependent on man for food in such places. Mountain regions tend to be species rich because of their topographical and habitat diversity. Coastal areas differ from inland areas in diversity of nonmarine birds for reasons that are not clear: In North America they are comparatively poor in terms of breeding species, but in parts of Europe they are richer than inland squares.

Even at the continental scale, there are usually major departures from simple north–south trends. The pattern of aridity (and associated vegetation) largely determines bird species diversity in both Australia and Africa. In North America (**Figure 16**), the east is richer in breeding bird species than the west, although the far west is richer than the Midwest, especially in the mountains and the southwestern desert; in the east, there is a midlatitude peak in richness. The latter is also seen in Europe. It is not simply a gross pattern that might be produced by differences in the proportion of richer and poorer habitats because it is observed within individual habitats.

Migration has major impacts on bird species diversity in many regions and habitats: It is substantially lower in the Arctic and much of the north temperate zone in winter than in summer and must vary considerably in African savannas because of migration both between Africa and the Palearctic and within Africa. This can modify geographical patterns of species richness. Thus, the difference between African forest and savanna is diminished, although by no means eliminated, in the northern winter because almost all the migrants move into the savanna; however, within the savanna the geographical pattern is maintained because areas rich in resident species are also areas to which most migrants move. In both Europe and eastern North America, the midlatitude peak of species richness in summer is eliminated in winter, when there is a simple north–south gradient.

Comparisons with Other Animals

There have been few direct comparisons between the geographical patterns of species richness of birds and those of other animals. In North America, a midlatitude peak in richness is also seen in mammals and amphibians, although the peak is further south, especially for amphibians. Reptiles show a simple north–south gradient. The decline in species richness in the north is steeper in mammals than in breeding birds, even steeper in amphibia, and steepest in reptiles. (Note, however, that many birds move out of the north in winter.)

Over North America as a whole, the species richnesses of birds and mammals are positively correlated but birds are not correlated with lizards and are negatively correlated with turtles and snakes. In Australia, there is a similar divergence: Birds are positively correlated with marsupials but negatively with lizards (which are species rich in hot, dry conditions, in which only specialist birds live).

Thus, although birds resemble other animals in showing a general tropical–polar gradient in species richness, local factors may affect them differently.

Causes of the Latitudinal Gradients in Species Diversity

Avian data have been widely used to illuminate possible causes of the tropical–polar gradient in species diversity. It is possible that it depends in part on there being more niches available in the tropics. Tropical vegetation is floristically more diverse and, in forests at least, structurally more complex. In North America, the midlatitude peak in bird species diversity is approximately matched by tree species diversity, as are the broad east–west trends. However, the relationship between the species diversity of birds and that of trees in North America is asymptotic; above a certain level, the richness of tree species appears not to influence that of birds. In Amazonian forests, 15% of bird species depend on habitats arising through dramatic fluctuations in river levels which do not occur in temperate regions. Furthermore, 34% of Amazonian birds belong to foraging guilds that are not found in temperate regions, such as obligate dead-leaf gleaners, obligate ant followers, and year-round frugivores; the rain forest's richness and comparative constancy provide opportunities lacking in temperate forests. However, this is not the whole story because even where guilds are found in both Amazonian and temperate forests they contain fewer species in the latter.

The North American midlatitude peak in breeding bird species diversity has been interpreted in terms of migration. At low latitudes, there are many year-round residents but at midlatitudes there are fewer; therefore, the summer provides many opportunities for migrants to pour in and to exploit the seasonal abundance of productivity. There is indeed evidence that insect food, although seasonally much more restricted in midlatitudes than in the south, is much more abundant when it does occur. Furthermore, it is migrants that feed on the insects that are active in summer that contribute most to the midlatitude peak; resident species such as woodpeckers, which have less seasonally variable food supplies, are not more numerous in midlatitudes than in the south. The hypothesis explains the decline in species richness in the far north (where there is also a summer flush of resources) in terms of constraints imposed by the shortness of the summer and the long journeys that migrants need to make to get there.

Another possibility is that the tropics are species rich because they include greater land areas, allowing more opportunities for speciation and reducing the likelihood of extinction. Although there is no correlation between the species richness of New World birds and land area at a given latitude outside the tropics, there is such a correlation if one excludes those species that are primarily tropical. It appears, therefore, that there is an underlying relationship between species richness and land area in the New World but that the very high richness of the tropics superimposes a wider gradient because some tropical species extend into subtropical and temperate regions.

An alternative hypothesis is that species richness is correlated with available energy in an area, either directly (warmer places allow animals to function more effectively) or indirectly, via productivity (higher productivity supports more individuals and therefore more species). Avian species richness in North America is correlated with potential evapotranspiration, which is a measure of available energy, but less well with actual evapotranspiration, which is more a measure of primary productivity. This supports the direct energy hypothesis. In Britain, temperature gradients are north–south in summer but, because of the influence of the North Atlantic Drift, east–west in winter. Data on the distributions of resident and wintering birds have been used to test the species energy hypothesis, predicting different patterns of species richness in summer and in winter. Early analyses provided support for the direct effect of energy, including indications that correlations with temperature were greater for smaller birds (which are more susceptible to low temperatures). Recent work has not confirmed this result. Nor has it found seasonal changes in distribution that match the predictions of the direct energy hypothesis. Summer temperature is the best predictor (of various climatic and topographical variables) of bird species diversity in Britain, in both summer and winter. This suggests that birds in winter depend largely on productivity stored from the summer (in the form of seeds, nuts, and invertebrates).

Species Replacements Throughout the World

The Ranges of Birds

The barn owl *Tyto alba* is found in every continent of the world except Antarctica, and it is widespread in most of them. Most birds are much more restricted in terms of habitats, climate zones, and continents. Some occur on single small islands or isolated mountaintops. Differences in which species occur in which places are just as much a part of biodiversity as is the species diversity of individual places.

Species Replacements

Two species may have different geographical ranges because they live in different habitats that are found in different parts of the world. Alternatively, they may occupy similar habitats but simply have evolved in different places. Thirdly, they may originate in isolation from each other, subsequently come together, but be ecologically so similar that competition prevents them from coexisting, except perhaps in limited areas (**Figure 17**). If they differ sufficiently in feeding habits or habitat, coexistence is possible but the extent of this varies between different sorts of birds, between breeding and winter quarters, and (within a taxon) between continents (**Table 8**). The occurrence of closely related species pairs that overlap only partially, if at all, such as illustrated in **Figure 17** is common in Europe and is probably the result of the range of an original species being split in two by the great climatic fluctuations of the Pliocene and Pleistocene. The Great Plains of North America are another zone in which such pairs of species meet.

Centers of Endemism

The ranges of 27% of the species of land birds in the world extend less than 50,000 km^2. Such restricted-range species tend to occur in particular areas: Some 221 "endemic bird areas" (EBAs), each with two or more restricted-range species, have been identified. The concentration is such that 2% of the earth's land surface contains 20% of the bird species. The

tropics contain 76% of the EBAs but this may be no more than a reflection of the generally high species diversity in the tropics. Reflecting the importance of isolation in speciation, almost half of the EBAs are on islands, half of which are small and oceanic. Continental EBAs are often in mountainous regions, where both mountains and valleys can form barriers to the dispersal of species that are limited to particular altitudes. Other centers of endemism are less obviously isolated and it is often postulated that they are places that formed habitat islands in the past as climatic changes fragmented their currently more continuous habitat. Since it is wooded rather than open habitats that tend to have been broken into such habitat islands, this fits with 71% of the restricted-range species being forest birds (and another 13% being birds of dry woodland and scrub). Alternatively, rivers and hills may be sufficient barriers to dispersal that they create the boundaries to centers of endemism.

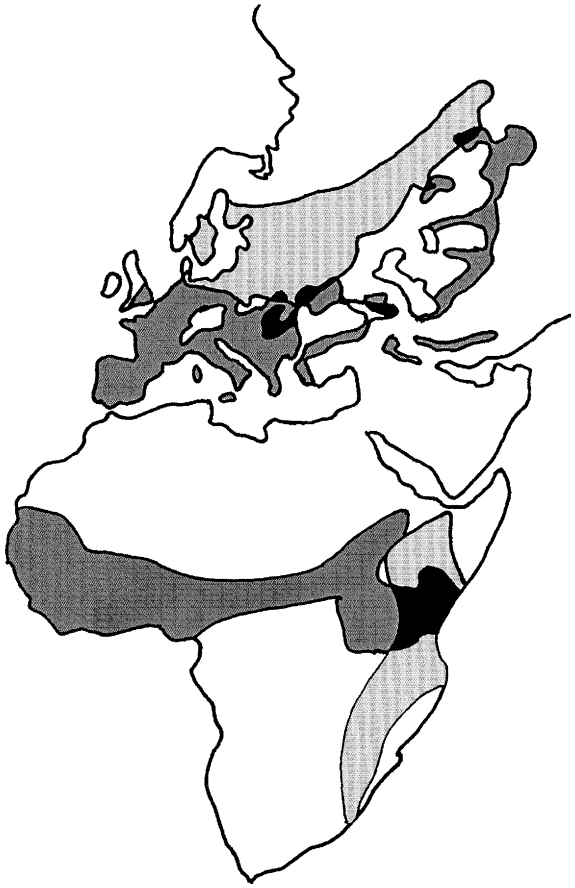


Figure 17 Ranges of nightingale *Luscinia megarhynchos* (dark shade) and sprosser *L. luscinia* (pale shade) in the breeding season (northern) and in winter (southern). Black areas show where the ranges overlap (redrawn from Cramp S (ed.) (1988) *The Birds of the Western Palearctic*, vol. 5: Oxford: Oxford Univ. Press).

Biogeographic Regions

Information on birds was important in the original identification of the six biogeographic regions of the earth— regions characterized by individual peculiarities of their faunas and floras as a result of long periods of separation in the past. Some bird families are found in several regions but others are confined to just one. The level of endemism in a region corresponds to the extent of its past isolation.

The least isolated regions, in that they form two great landmasses that have been joined by a broad land bridge (at the Bering Straits) for much of their recent history, are the Nearctic (North America) and Palearctic (Europe, most of Asia, and North Africa). The first has no endemic bird families and the second only one; many families occur in both, probably as a result of invasions from the Palearctic to the Nearctic via the Bering land bridge. The climate is markedly seasonal in large parts of each of these regions and they support comparatively few species (approximately 750 in the Nearctic and approximately 1100 in the Palearctic), many of which are migrant insectivores that move south in winter. Many species were probably lost from these regions as a result of deterioration of the climate during the Pleistocene. Subsequently, the Caribbean Sea, the Mediterranean Sea, and the Sahara desert, although crossed by many migrants, have been partial barriers to recolonization from Africa and South America.

The Oriental region (the Indian subcontinent and Southeast Asia) has only 1 endemic bird family, although it remains somewhat isolated from the Palearctic by mountains. It has approximately 1000 species. Much of its avifauna originated largely in the Palearctic, although it has some affinities with the Ethiopian region (sub-Saharan Africa). The latter has 13 endemic families among its approximately 1500 species. It receives many Palearctic migrants during the northern winter and some of its resident species may have ancestrally originated in the Palearctic. The richness of forest birds in Africa is

Table 8 Percentage of bird species that are ecologically separated from congeneric species primarily by feeding, habitat, or geographic range (i.e., having largely separate but adjoining ranges) or by having no contact[a]

	Parus *species (%)*		*Trans-Saharan passerine migrants (%)*		*European birds (%)*		
	N. America	*Europe*	*In Europe*	*In Africa*	*Passerines*	*Other land birds*	*Waders and waterbirds*
Feeding	4	43	10	2	19	43	48
Habitat	11	32	62	23	55	20	18
Range	29	8	8	64	7	11	9
No contact	55	17	18		18	14	19
Unknown	—	—	2	10	1	11	7

[a]Data from Lack D (1971) *Ecological Isolation in Birds*. Oxford: Blackwell Scientific Publications.

less than that of Central and South America, perhaps reflecting the disruption to African forests caused by dry periods in the Pliocene and Pleistocene.

The Australian region has a distinct avifauna—16 endemic families and approximately 1000 endemic species among the approximately 1600 bird species living there. Many of these belong to the Corvida, a major division of the Passeriformes revealed by DNA analysis, which appears to have originated in Australia. New Guinea and other nearby islands belong to this region, the somewhat unclear dividing line between the Australian and the Oriental regions lying between Lombok and Bali and between Celebes and Borneo, along a deep strait that has probably always separated these regions even when sea levels have been much lower than they are today. The ancestors of some Australian species have arrived from the north across this barrier, but others are derived from the south, at a time when Australia lay with the other Southern Hemisphere continents in the great landmass of Gondwanaland.

The most distinctive and most species-rich region is the Nearctic (South America), inhabited by 31 endemic families and more than 3000 species living in a variety of habitats ranging from the tropical to the far south temperate and from sea level to extensive mountains. It has a comparatively small proportion of oscines among its avifauna and seems to have had a long evolutionary independence from the rest of the world, although its passerines may be largely derived from North America.

Evolutionary Convergence

When unrelated forms occupy similar habitats in different parts of the world, they may evolve similar characteristics in adaptation to those habitats. Such evolutionary convergence between unrelated forms is widespread among birds: for example, the great auk *A. impennis* and penguins (large, flightless seabirds using wings to swim under water); the little auk *Plautus alle* and diving petrels *Pelecanoides* (small seabirds that plunge into the depths to catch planktonic prey, using wings to both swim and fly); the New World vultures (related to storks) and the Old World vultures (related to hawks); the horn-bills and toucans of the Old and New Worlds, respectively (both with long bills for reaching fruit in the canopy, with secondary use of the decorated bill for social signaling); and the sunbirds and the hummingbirds, also of the Old and New Worlds, respectively (small, brightly colored nectar feeders, with specialized tongues). More generally, the Australian Corvida include many groups of birds that have been considered in the past to be more closely related to various groups of passerines elsewhere in the world than to each other. Such conclusions depended on similarities that they shared with these other groups as a result of occupying similar ecological niches and thus being molded by similar selection pressures. These similarities are now recognized as convergent.

The extent of morphological convergence among whole communities has rarely been examined. A comparison of birds from Mediterranean communities in France, California, and Chile showed that they were no more similar to each other than each was to bird communities from a temperate part of France. Perhaps the environmental differences between temperate and Mediterranean France are too slight to promote significant morphological divergence.

Diversity on Islands

The Colonization of Islands by Birds

Seabirds abound on islands partly because they are safer places to nest than mainlands (generally having fewer predators) but also because the oceans present no barriers to seabirds. In contrast, land and freshwater birds are less diverse on islands than on mainlands, although they have been much more successful in reaching islands than have other nonmarine vertebrates. Continental islands, which were once joined to larger land-masses before being separated by continental drift or sea-level rise, may retain some of their original species (perhaps much modified) but the evidence indicates that the more ancient of them have been colonized by many other species across the sea. Recent genetic evidence that kiwis are related to emus suggests, for example, that their ancestors may have arrived from the north, perhaps by flying over the Tasman Sea. (Moa, in contrast, probably originate from an original Gondwanaland stock.) Oceanic islands, never joined to a mainland, must have been colonized over the sea. Vagrant individuals of a variety of species are commonly observed even on the most remote islands. The efficacy of such colonization is shown by the fact that those islands of Krakataua that were devastated by volcanic eruption in 1883 held as full a complement of species as undisturbed islands within less than 100 years. As another example, if one allows for differences in area (small islands tend to hold fewer species), oceanic islands are as rich in bird species as continental islands in the Gulf of California.

Even for birds, however, isolation appears to result in substantial loss of diversity (**Figure 18**). Islands 250 km from New Guinea have only about half the number of bird species expected from their size, and those more than 800 km away have fewer than 10%. However, is this because birds do not reach remote islands? Perhaps they get there but do not find rich enough habitats (because a sufficient wealth of plants and other animals has not arrived). The answer undoubtedly depends on the species and the island in question. Studies that focus on individual species, rather than just on counts of species number, show clearly that species differ markedly in colonizing ability. J. M. Diamond conducted extensive analyses of the bird faunas of the many satellite islands of New Guinea and found, in addition to simple distance effects, that islands that were connected by land bridges during periods of low sea level in the Pleistocene had more bird species than those that had not had such connections. Some species, apparently poor at crossing water, are found only on land-bridge islands, even though the never-connected islands have the same habitats available. Diamond divided the birds according to apparent colonizing ability. At one extreme are "supertramps," often seen as vagrants, found on the most isolated islands that few other species have colonized, living at high densities, breeding rapidly, and having unspecialized habitat requirements. At the other extreme are species that are found only on islands that many other species have also reached;

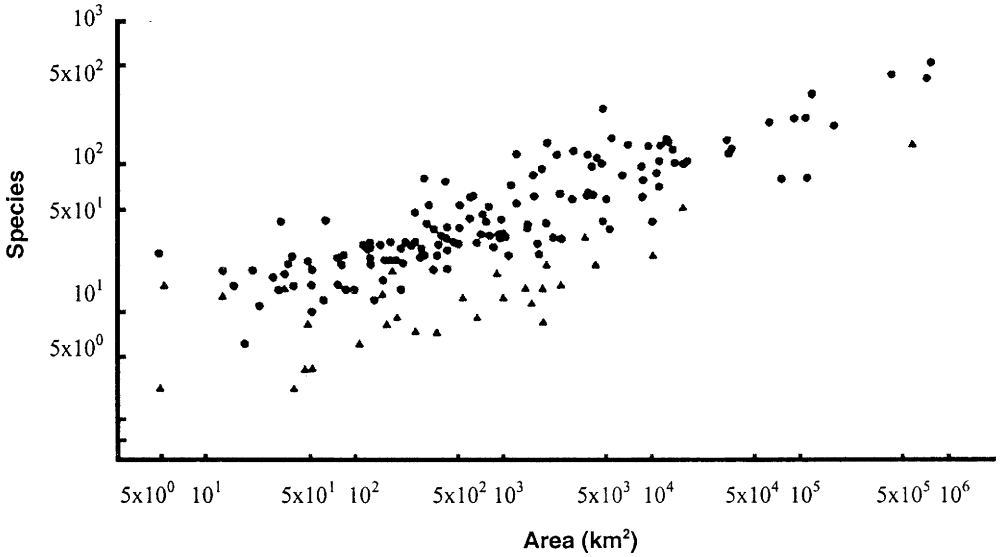


Figure 18 Relationship between the number of species on individual islands in the warmer seas and their areas. Circles represent islands more than 300 km from the next largest landmass or in the Hawaiian or Galapagos archipelagos (redrawn from Williamson M (1981) *Island Populations*. Oxford: Oxford Univ. Press).

they are habitat specialists and live at low densities. When they reach an island these specialists appear to reduce the resources to such a degree that the generalist super-tramps cannot survive; the latter are therefore absent from species-rich islands. Thus, the supertramp dove *Ptilinopus solomonensis*, found on many small islands of the Bismarck Archipelago, is replaced on New Britain and New Ireland by three congeners—one in mountains, one in lowland forests, and the third in open lowland habitats.

Evolution of Island Birds

Island populations are genetically somewhat isolated (sometimes wholly isolated) from the mainland populations from which they sprang. The environment on the island is likely to differ from that on the mainland both physically and biotically (having fewer competitors and predators). As a result, there may be rapid evolutionary adjustment. Flightlessness has already been mentioned. Niches may be broadened. Thus, in most of western Europe, treecreepers *Certhia familiaris* are largely confined to conifers (and generally to high altitudes) by the presence of the closely related *C. brachydactyla*, which is a broad-leaved specialist, but in Britain and Ireland (and also in far eastern Europe and Asia), from which *C. brachydactyla* is absent, they inhabit woods of all sorts. Where studies of the relationship between bird species diversity and foliage height diversity have been conducted on islands, it has been found that the birds appear to divide the foliage into fewer layers than on the mainland; individual species have broader niches when there are fewer competitors. Another evolutionary change in some groups, such as ducks, is that island forms may lack distinctive male breeding plumages, either because there is less chance of hybridization when related species are absent or because demographic changes lead to behavioral modifications.

Island forms are often recognized as subspecies as a result of such changes; they may even become distinct species.

The chaffinch *Fringilla coelebs* is widespread in Europe and North Africa, inhabiting both deciduous and coniferous woods. It has the same habits on the Canary Islands of Hierra and La Palma, but on Tenerife and Gran Canaria it is restricted to deciduous woods; coniferous woods on those islands are occupied by the related *F. teydea*. These islands appear to have been colonized by *F. coelebs*, which then, adapting to the widespread coniferous woodlands there, evolved so much (into *F. teydea*) that when another *F. coelebs* invasion occurred the two species not only failed to interbreed but also each had a habitat to which it was better suited than the other and in which it could therefore persist. Such species pairs are known elsewhere. They represent a rare case of island bird communities being more diverse than those of mainlands.

Studying the birds of the Lesser Antilles, R. E. Ricklefs and G. W. Cox modified the idea of a "taxon cycle," earlier propounded by E. O. Wilson in respect of Australasian ants to describe the distributional and evolutionary changes when a species colonizes an archipelago. *Molothrus bonairiensis*, progressively colonizing the southern islands (from South America), represents an early stage of such a cycle; *Tyrannus dominicensis*, found throughout the islands, represents the next; and then *Loxigilla noctis*, with distinct subspecies on some islands. Finally, some of the islands have actually lost the colonizing species, either through chance extinction or through competition from later invaders: *Dendroica adelaidae*, restricted to St. Lucia and Barbuda, is one such species.

Isolated archipelagos, with very depauperate faunas but diverse habitats and with some isolation between islands, are hotbeds of evolution for those species that do arrive. The adaptive radiation of the Hawaiian honey-creepers (Drepanididae) is a well-known example. Comprising 23 extant species (more than half the endemic bird species on Hawaii) and at least 8 extinct ones, they are derived from a single fringillid

ancestor but with remarkably divergent feeding adaptations: typical finch beaks for seed-eating, parrot-like beaks, thin insect-eating beaks, and beaks elongated up to one-third the body length for sucking nectar. The akiapolaau *Hemignathus wilsoni* has a stout lower mandible with which it chips into decaying wood, holding out of the way its long, curved upper mandible, which it then uses to probe for insects in the exposed galleries. Darwin's finches (Geospizinae) are even better known. The diversity of the Galapagos species of this subfamily is in sharp contrast to the existence of just 1 species on Cocos Island, about 700 km northeast. The latter, although it has a diversity of habitats, has not presented the opportunities for speciation that the many separate islands of the Galapagos have done; as a result, the Cocos finch is behaviorally flexible, utilizing a wide range of foraging opportunities.

The Species–Area Relationship in Birds

It is well-known that larger islands tend to hold more species (see **Figure 18**). Birds have been used in many studies of the species-area relationship because they have reached most islands and their distributions are well-known. An especially steep relationship has been observed for islands in a Minnesota lake, apparently because the smaller islands studied were just too small (<0.5 ha) to support even single territories of many species. However, this is unlikely to account for the relationship observed in many studies of much larger islands. This has been widely interpreted in terms of the MacArthur–Wilson theory of island biogeography: As more species become established on an island, the rate at which new species colonize is bound to decrease (fewer of the arrivals are of new species); as the number of established species increases, the number going extinct also increases (there are more to go extinct); and at some point the colonizations and extinctions therefore reach an equilibrium. Since extinctions are less likely on larger islands (because populations are larger), the equilibrium number of species will be greater on larger islands. On well-studied islands, such as those around the British coast, there are certainly frequent colonizations and extinctions. These British islands are close to the source of colonists and are small, so this result is scarcely surprising, but it illustrates the dynamic nature of island avifaunas. The data also show that species differ in their probabilities of extinction; therefore, species-specific approaches are needed for a complete understanding.

The slope of the species–area relationship tends to be flatter for more isolated island groups because these are colonized only by the species with better powers of dispersal, which can quickly reoccupy islands from which they have gone extinct. On Bahamian islands, the slope is shallower for migrant species than for residents, the former being better colonists (and the slope is steeper for lizards, which are poor colonizers).

For species late in the taxon cycle, where extinctions have begun to take effect, the slope is steeper than for species early in the cycle, as the equilibrium theory would predict. Conversely, the slope was found to flatten in a group of Finnish islands following reduction in disturbance by man; this had previously led to many extinctions, especially on small islands.

Although these observations are in agreement with the equilibrium theory, they do not rule out the importance of habitat. David Lack argued that larger islands held more bird species because they were more diverse, both physically and in terms of floristic richness. Statistical analyses designed to untangle the effects of habitat diversity and those of area *per se* found the former to be the chief cause of the species–area relationship in birds on the Isles of Scilly and on montane "islands" in the American Great Basin. Species richness of birds on the Lesser Antilles correlates independently with both habitat and area; in divergent contrast, bats correlate only with area and butterflies, reptiles, and amphibians only with habitat. It has been suggested that the difference is the result of birds (and bats *a fortiori*) having lower population densities than the other animals and therefore being more vulnerable to chance extinction. Similarly, the low species richness of the most isolated Bahamian islands seems to depend on both their isolation (many birds do not get there) and their habitat poverty (many arrivals find nowhere to live). Studies of forest fragments, which show the same sort of species–area relationship, indicate that part of the reason for the small number of species in small fragments is that they lack those species that are only found deep in forests.

Competition and resources mold island avifaunas. There is evidence that the species–area relationship is less steep for individual trophic groups than for whole avifaunas, which is in agreement with the idea that competition within the group reduces the ability of new species to colonize when an island already has many species or increases the likelihood of a new colonization leading to the extinction of an established competitor. Competition may also explain the paradoxical finding that migrant birds are more species rich on more remote Bahamian islands; the migrants, it is hypothesized, can reach all the islands but are less likely to establish on islands with residents already established (i.e., the less isolated islands).

Fragmentation as a Cause of Extinction

Nature has conducted experiments that reveal the importance of the area of islands in determining extinction rates and thus the numbers of species present at equilibrium. Thus, islands that were connected to New Guinea as a result of lower sea levels approximately 10,000 years ago now have fewer species than expected if they were part of mainland New Guinea (although they still have more species than do islands of comparable size that were never connected, suggesting that equilibrium has not yet been reached). Three New Guinean islands that were part of a single island approximately 10,000 years ago have fewer species than that single island would be expected to have, given its estimated size (although again they have more than expected for islands of their current size). The number of species that each has apparently lost is inversely proportional to its area, as the equilibrium theory predicts. The same is true for a suite of islands off the north coast of South America and the coast of Central America.

Man formed Barro Colorado Island during the construction of the Panama Canal early in the twentieth century. It has subsequently lost 50–60 species. Some of the losses may be the result of increased predation because man no longer hunts small and medium predators. It is striking, however, that the species that have gone extinct are those whose numbers appear to be particularly affected by the occasional periods of very dry weather that occur in the region. Average low population densities, in contrast, seem not to be a good predictor of extinction from Barro Colorado: As long as a population does not fluctuate too much, it can persist even if its average numbers are small.

Forests in most parts of the world have been fragmented by man and numerous studies have shown that fragmentation has resulted in the loss of species. This is not just the effect of random extinction, however, because it particularly involves the loss of forest-interior specialists. Indeed, direct observation in the highly fragmented woodlands of eastern England confirms that many more woodland species pass through small woodlots than stay to breed, showing that they would have no difficulty in colonizing such places if they were suitable. Blue tits *P. caeruleus* and great tits *P. major* breeding in such small woods produce fewer young, partly as a result of breeding later in the season, than do the same species breeding in larger woods. This may be because the small woods provide less shelter from the weather or because they lack sufficiently rich vegetation to produce adequate insect food for the birds. Whatever the cause, these fragmented woods support fewer birds than would a solid block of forest of the same total area.

In North America, Neotropical migrants also tend to avoid fragmented forests, much more than do resident species or temperate migrants. The fragmentation of forests may be the reason why so many Neotropical migrants have declined in numbers in recent decades.

Introduced Species as Colonists: A Challenge to the Equilibrium Theory?

Many birds have been introduced outside their native ranges and have flourished in their new homes. For example, 34 introduced land and freshwater species are established on New Zealand. Many of these have colonized offshore islands—up to 10 on a single island. This does not mean that those islands were not in equilibrium before these species arrived. What has happened is that the pool of species available to colonize the islands has been increased, raising the colonization rate and thus setting a new equilibrium.

If New Zealand were capable of absorbing these species, does this mean that it was not at equilibrium? With only 60 resident, nonmarine indigenous species it was certainly depauperate, so one could argue that ecological niches were vacant. However, in undertaking the introductions, man was enhancing the colonization rate, thus shifting the equilibrium upward. Furthermore, many native species have been driven close to extinction in New Zealand, with some surviving only because of special protection measures. Although habitat loss and the devastations of introduced predators have been major causes of such near losses, competition from the introduced species may also have played a part.

Human Causes of the Loss of Avian Biodiversity

Natural Causes of the Loss of Avian Biodiversity

The paleontological record shows that many birds have become extinct as part of the natural flux of life on Earth. The recent extinctions of an unnamed megapode on the Kermadec Islands (1876) and of the San Benedicto rock wren *Salpinctes obsoletus exsul* (1952) were also natural, resulting from volcanic eruptions. That element of biodiversity that is made up of the differences between regions in the species that inhabit them must also have been eroded by natural colonizations. These natural processes, however, are slow—slow enough for extinctions to be approximately balanced by the origins of new species. Man has accelerated them manyfold, causing serious losses of avian biodiversity.

The History of Bird Extinctions

It is difficult to be sure whether a species has gone extinct, especially if it is scarce and lives in remote places. The Fiji petrel *Pterodroma macgillivrayi*, for example, was known from but a single specimen taken in 1855 until a second was found in 1984. With that caveat, **Table 9** shows the avian extinctions recorded in the past 400 years. These represent more than 1% of known birds, a proportion similar to that for mammals but otherwise greater than that for any other group of animals or plants.

Two things stand out from **Table 9**. First, extinctions peaked between 1850 and 1950 but have still continued at a high rate. Second, approximately 90% of them have been extinctions of island species, even though only 20% of bird species live on islands. One–third of the birds present on Hawaii 200 years ago are now extinct and two-thirds of the remainder are endangered. Two–thirds of the historically indigenous birds of the Mascarenes have been lost and more than 40% of the rest are endangered. Island species are partly vulnerable because their ranges are so small and some of the

Table 9 Chronology of know extinctions of bird species since 1600[a]

	1601–1650	*1651–1700*	*1701–1750*	*1751–1800*	*1801–1850*	*1851–1900*	*1901–1950*	*1951–present*	*Total*
Continental	0	0	0	0	2	2	5	2	11
Island	4	8	5	8	14	31	30	15	115
Total	4	8	5	8	16	33	35	17	126

[a] Reproduced from Newton I (1998) *Population Limitation in Birds*. San Diego: Academic Press.

extinctions on continents have been of species with particularly small ranges. Because birds that occur in few places tend to be uncommon even where they do occur (**Figures 5** and **6**), they are especially vulnerable.

Hunting has caused many extinctions. As recently as 1945, Japanese troops eliminated the Wake Island rail *Rallus wakensis*. The case of the passenger pigeon *Ectopistes migratorius* is classic. Once numbered in hundreds of millions, it occurred in huge flocks that "darkened the skies." Although habitat loss may have contributed to its decline, there is little doubt that it was extinguished by massive commercial hunting. Such uncontrolled exploitation contrasts with situations in which a community has been able to control the hunting, such as was the case on St. Kilda, where seabirds were a major resource whose exploitation was regulated. The great auk *A. impennis* and Dodo *Raphus cucullatus*, extinguished by visitors to their uninhabited island homes, provide a clear contrast and also illustrate other aspects of the vulnerability of many island birds: They were large, flightless, and had limited antipredator responses, thus making them easy and profitable to hunt. They probably had low reproductive rates and therefore were incapable of sustaining high levels of hunting.

The introduction of alien species has been a major cause of island extinctions. Worldwide, there is a clear correlation between the number of indigenous birds that have become extinct in a locality and the number of avian introductions established there. Some introduced species, especially mammals such as rats, cats, dogs, mongooses, and pigs, are important as predators. In the early 1960s, black rats *Rattus rattus* eliminated the last known population of the bush wren *Xenicus longipes* within 3 years of reaching Big South Cape Island, New Zealand, along with the island's populations of four other birds. In 1894, the lighthouse keeper's cat caught 15 Stephen Island wrens *Xeniscus lyalli*; these were the first specimens discovered—and also the last. Introduced monkeys are important predators at the nests of endemic birds on Mauritius. Only one introduced snake is known to have caused serious problems, but its effects have been dramatic: The brown tree snake *Boiga irregularis* seems to have arrived on Guam in the 1940s as an accidental hitchhiker. By 1986 it had extinguished six of the indigenous forest birds and reduced the other four to populations of less than 100; of these, the endemic Guam rail *Rallus owstoni* had been reduced to 100 in 1983 and it was extinct in the wild by 1987, although it survives in captivity. In contrast, none of the seven introduced birds on Guam has been obviously affected by the snake: They have effective antipredator behavior and live in habitats in which the snake is scarce. (Only one of the eight native species in such habitats has gone extinct.) It is noteworthy that few of the species extinguished by introduced predators lived on islands with indigenous rats or land crabs, which would have caused them to maintain their antipredator responses.

Introduced birds may out-compete natives, especially in the modified habitats that man has created on most islands since they may be adapted to such habitats in their own native ranges. Thus, the familiar birds in suburban and agricultural habitats of New Zealand are the familiar birds in suburban and agricultural habitats of western Europe. Competition may be genetic as well as ecological: Part of the cause of the extinction of the Atitlán grebe *Podilymbus gigas* may have been hybridization with the introduced pied-billed grebe *P. podiceps*.

The loss of natural habitats has been another major cause of extinctions. Much of it has been deliberate (the clearance of land for agriculture or human settlement), but much has been an accidental by-product of the introduction of alien herbivores, especially to islands where the native plants have evolved in the absence of such animals. Thus, the rabbit *Oryctolagus cuniculus* destroyed much of the vegetation on Laysan in the 20 years following its introduction in 1903; three of the five species of land birds disappeared with the plants. On some islands, soil erosion, following deforestation, has removed nest sites for burrowing seabirds.

Finally, the impact of introduced diseases is largely unknown but native Hawaiian species, previously free of avian malaria, have been infected via introduced birds.

The Recent Prehistory of Bird Extinctions

It is well established that there were mass extinctions of large mammals following the colonization of continents by man the hunter. Vulnerable birds may have been extinguished at the same time, as may have birds dependent on the mammals. Fossils show the Californian condor *Gymnogyps californianus* once to have been widespread across North America, its restriction to the west presumably consequent on the reduction in numbers of large mammals brought about by man, dating from approximately 11,000 years ago. It now survives only through a captive breeding program, its final historical collapse also being a by-product of human activity—in this case, poisoning by the lead in carcasses abandoned by hunters and by poisons used to kill coyotes *Canis latrans*.

Colonization of islands mostly occurred later. In the Pacific, people had gone as far as the Solomon Islands by 30,000 years ago but did not reach western Melanesia and Micronesia until 4000 years ago, moving then through Fiji and Samoa, the Marquesas, Hawaii, and Easter Island, with New Zealand only colonized about 1000 A.D. European colonists followed in the sixteenth-eighteenth centuries. Some of these human colonizations eventually failed, perhaps because of their dramatic impact on the islands' ecosystems. The early colonists hunted the animals they found, generally introduced alien species, and usually cleared forests. Their activities certainly led to loss of bird populations: On Hawaii, 60 endemic species are known only from bones; at just one site in New Zealand, 29 endemics are known only from bones; on Easter Island, bones of about 30 species of seabirds have been found from the early years of human occupation but only 1 is now found there (with about 10 others surviving on offshore islets) and all the indigenous land birds seem to have died out following the total deforestation of the island. D. W. Steadman reached the following conclusions about Polynesian land birds:

> (i) The ranges of most living species are much smaller today than they were at the time of first human contact. (ii) Few volant species are naturally endemic to only one or two islands. (iii) Most species have become extinct in

the past 3000 years. (iv) Most or all islands supported one to four endemic species of flightless rails, virtually all now extinct. (v) In many cases, individual islands supported two or three species within a genus, unlike the situation today. (vi) At least four formerly widespread genera now are gone from East Polynesia. (vii) Although modern distributions of Polynesian land birds continue to be analyzed as if they were natural, they do not furnish unbiased data for proposing or testing ecological models. (viii) Although some of the range losses of extant species could be restored with conservation efforts, we are centuries too late to preserve any true likeness to the original Polynesian avifauna.

On the basis of the fossil evidence, Steadman suggests that the 800 Polynesian islands inhabited by man have lost an average of 10 species or populations of birds, including 2 or 3 species of flightless rails endemic to each. Other authors have suggested that, because of the incompleteness of the fossil record, even these data may be underestimates. Whatever the details, the avifauna that survived to be further devastated by Europeans, Americans, and Japanese was a small part of that which the earlier colonists found. The marked deficiency of seabirds in the tropics, although probably partly a result of the low productivity of tropical oceans, may be a legacy of this destruction.

Current Threats to Birds

On the basis that they have undergone (or appear likely soon to undergo) rapid declines, that they have small and declining populations, that they leave small and declining ranges, or that they have very small populations, 11% of the species of birds in the world are classed as "threatened." The percentage for mammals is similar; no other group of animals reaches this level of threat. Forest birds make up 65% of this total, and those of scrub and of wetlands each comprise 9%.

Most of the endangered species are tropical. At least for the New World, this is not just because there are more species in the tropics because the proportion of the avifauna that is threatened is also highest in the tropics (**Figure 19**). This appears to be in part because the tropics hold many species with restricted ranges. (Throughout the world, about one-fourth of the threatened species are endangered simply by having very small populations or ranges.) In the New World, it is also because of the intensity of tropical deforestation. Of course, areas that hold relatively few threatened species may do so because those species that were especially vulnerable to the impact of man have already gone extinct; this may explain why there are fewer threatened species north of the equator in the Americas than in the south (**Figure 19**), given the longer history of human interference in the north.

Habitat loss, especially the loss of tropical broad-leaved forest, is a threat to about half of the world's endangered birds. It is caused not only by the expansion of agriculture but also by mining and the construction of dams. In many parts of the world, afforestation with exotic species is a major problem: Being exotics, the introduced species offer homes to only some of the native animals. Furthermore, they are managed in ways that reduce the floristic and structural diversity of the

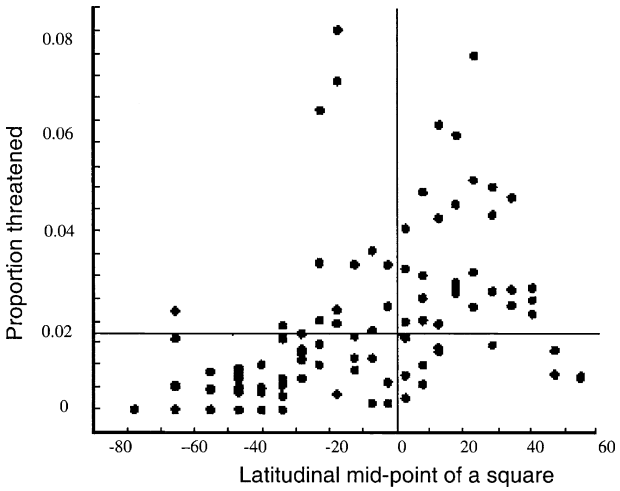


Figure 19 Proportion of New World birds classed as threatened, in relation to latitude (negative values are in the Northern Hemisphere). The horizontal line is the mean proportion threatened (redrawn from Gaston KJ and Blackburn TM (1996) The spatial distribution of threatened species: macroscales and New World birds. *Proc. R. Soc. London B* 263: 235–240).

plantations. Drainage of wetlands is a major problem in many regions, although one that can often be relatively easily reversed if there is the will to do so. About 8% of the world's endangered species are threatened by hunting and capture for the pet trade. In Europe, more than 30% of the species that have declined in recent years are thought to be affected by hunting or by related persecution (such as the killing of raptors that are viewed to be competitors with man). Much hunting is now recreational rather than for subsistence, but it is often more destructive than formerly because there are more people, because more of them have the leisure time to engage in hunting, because there is better technology (mist nets, guns, and motor transport), and because more people can afford the technology. Coastal habitats in sunny places are steadily being destroyed by the development of tourist facilities. Even where beaches are left in a natural state, birds breeding on them suffer from increased disturbance.

In Europe, the greatest threat, affecting more than 40% of declining species, is the intensification of agriculture. Birds of the steppes are threatened by irrigation, whereas those of wet meadows are threatened by drainage. Fewer resources are available for wildlife when hay meadows are converted to silage and when stocking rates on pastures are increased. Insecticides and herbicides remove much potential bird food from crops. Plant breeding, the use of pre-emergence herbicides, and more effective machinery may allow autumn cultivation in areas where fields used to be left in stubble over winter, such that granivores could feed on grain spilled at harvest and on the seeds of weeds. Harvesting is more efficient. Fundamentally, more of the primary productivity is being harvested, so less is available for wildlife. Most of these changes depend on technological developments but their adoption has been promoted by systems of agricultural support, such as the Common Agricultural Policy of the European Union (EU), that concentrate on production rather than also taking account of social and conservation needs. Expansion of the EU to include countries

of Eastern Europe, where agriculture is still relatively less intensive, is a major threat to birds and other wildlife.

Further agricultural changes may flow not only from technology and policy but also from global climate change. This will also affect natural habitats. Although vegetation and wildlife have adapted to the natural climate changes of the past 10,000 years, often apparently within decades, there are concerns that the changes now occurring will be substantially more disruptive because of their speed.

Reducing the Threats

Species have been pulled back from the brink. Use of DDT reduced the Mauritius kestrel to 4 birds in 1974 but 20 years later, through a program of captive breeding and reintroduction, there were more than 200. There have been several similar cases, but they have required very intensive care. More extensive measures are needed if all threatened birds are to be conserved.

There are two broad approaches to such conservation. One is simply to control all of the threats listed previously, which would have a widespread impact. Some actions would be neither politically very difficult nor economically very costly, such as banning deliberate introductions, controlling accidental introductions, and eliminating introduced species that have become established. Most, however, would entail a direct conflict between the benefits to wildlife and the material wealth of burgeoning human populations.

The other remedy is to secure individual sites as biodiversity reserves. This would also be expensive because, although 20% of bird species occupy just 2% of the land surface of Earth, it would take a much larger area to include the other 80%. (Furthermore, because hot spots of diversity for different taxa do not closely coincide, even greater areas would be needed if more than birds were to be considered.) In any case, diversity hot spots are peculiar places; conserving them might maintain many species but it would not conserve the natural diversity of ecosystems. There is also the practical difficulty of ensuring that the correct locations are chosen. For example, after decades of trying to build up numbers of the red kite *Milvus milvus* in the Welsh fastnesses to which the British population had become restricted by nineteenth century persecution, conservationists concluded that it was probably in Wales not because the most suitable habitat was there but just because it had not been shot out there; populations subsequently introduced to eastern England and Scotland have flourished, indicating that the Welsh birds were probably living in relatively poor conditions.

Other Reductions in Local Diversity

Extinction is a particularly obvious way in which biodiversity is diminished but is merely the tip of an iceberg. For every species extinguished by man, there are many others whose numbers and ranges are substantially reduced. (There are also a few whose numbers have been much increased by human activity, such as the house sparrow *Paaser domesticus* and starling *S. vulgaris*.) Changes in numbers are less easy to assess than extinctions and little is known about them in many parts of the world. There are good data, however, for much of North America, where it is clear that recent decades have seen major declines of many Neotropical migrants. It is not clear to what extent this is the result of habitat loss in the wintering areas or of fragmentation of woodlands on the breeding grounds. In Europe, changes in agriculture have been closely followed by declines in the populations of farmland bird specialists (**Figure 20**). Studies of individual species are beginning to reveal the causes of these losses, which are different in detail for different species but all fundamentally related to agricultural intensification.

It is specious to argue that such changes are but a reflection of the natural flux of populations. Although systematic counts are not available for most species, except for recent years, it is clear from the extensive historical literature that British farmland birds have probably never before suffered losses on the

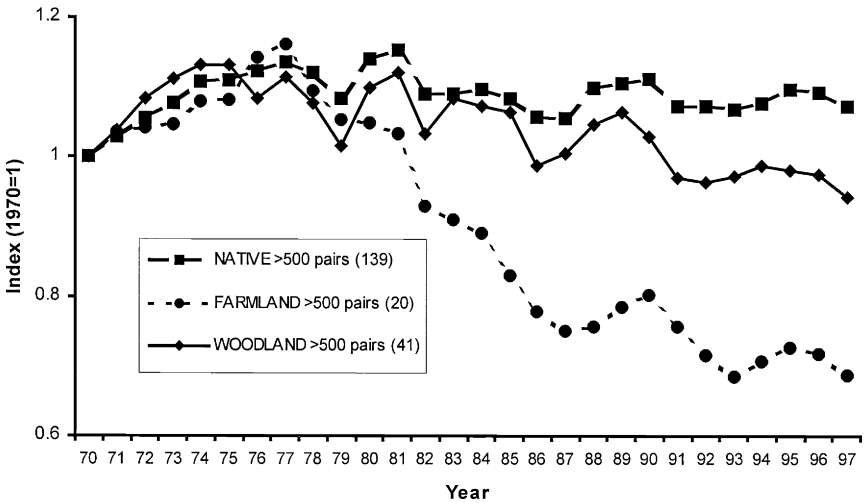


Figure 20 Index of bird populations in the United Kingdom based on geometric means of indices for individual species. Introduced species and those with fewer than 500 pairs are excluded (reproduced from the British Trust for Ornithology and the Royal Society for the Protection of Birds).

scale of those recorded in the past few decades, even during the periods of enclosure and other intensive habitat modification near the end of the eighteenth century and the development of high agriculture in the nineteenth century. The argument that the species characteristic of farmland are much more common than they were before much of western Europe was cleared of its forests is correct but incomplete. Forest clearance led to open-country birds replacing woodland birds, but the farmland birds now being lost are not being replaced. As a result, the total biomass of nonmarine birds breeding in Great Britain declined between 1968 and 1988 by 10% (39% if the pheasant *Phasianus colchicus*, which is widely and increasingly stocked for shooting, is omitted).

Losses of Geographical Diversity

Part of the diversity of life on Earth is the difference between places in their animal and plant communities. Introductions not only endanger biodiversity at the local scale (*see* Human Causes of the Loss of Avian Biodiversity – The History of Bird Extinctions *and* Human Causes of the Loss of Avian Biodiversity – The Recent Prehistory of Bird Extinctions) but also, however harmless they are to the indigenous fauna and flora, lessen biodiversity by reducing the difference between places.

More than 130 species of birds have been naturalized outside their native ranges, i.e., they have been introduced and have established breeding populations. As a result, the range of some species has increased dramatically. The pheasant *Phasianus colchicus*, native in only a band stretching from the Caucasus to Japan, now occurs in most of Europe, temperate North America, and New Zealand, with scattered populations in Australia and various islands.

Sport is a major reason for introductions (even in regions with plenty of native game birds) and one-fifth of established introductions are of pheasants (Phasianidae). Some species have been introduced for food (such as the feral pigeon *Columba livia*), many to remind settlers of "home" (such as European songbirds across the globe), and some because they are beautiful (e.g., peafowl *Pavo cristatus*). Some have been introduced as agents of biological control, such as the house sparrows *P. domesticus* transported to the United States in the nineteenth century to control defoliating moths. The Chimango caracara *Milvago chimango* was introduced to Easter Island as a scavenger. Conservationists have occasionally introduced threatened species beyond their native range because the latter was too badly affected by habitat destruction or by introduced predators or sometimes, perhaps less justifiably, because the native range was considered unsafely small.

Many species have established themselves following accidental escape from captivity; therefore, wildfowl, parrots, doves, and small seed eaters make up a high proportion of naturalized species. In the summer of 1991, Britain (which, like most of western Europe, has only one native species of goose) held established breeding populations of four other geese; 10 other species were at large, 2 of them recorded breeding, and there were 18 different sorts of hybrid geese identified. Some of these avicultural species may be deliberately released because their owners tire of caring for them or wish to see them free-flying. Similarly, farmed species may be released if markets collapse: This is the origin of the Australian population of the ostrich *Struthio camellus*.

Species that are closely assosciated with man, such as house sparrows and house crows *Corvus splendens*, have reached some places by riding on ships.

Introductions of birds have been less damaging than those of mammals and fishes. Nonetheless, each one of them represents a direct reduction in the diversity of life throughout the world.

See also: Amphibians, Biodiversity of. Fishes, Biodiversity of. Mammals, Biodiversity of. Migration. Reptiles, Biodiversity of

References

Blackburn TM and Gaston KJ (1994) The distribution of body sizes of the world's bird species. *Oikos* 70: 127–130.
Blackburn TM and Gaston KJ (1996a) Spatial patterns in the body sizes of bird species in the new world. *Oikos* 77: 436–446.
Blackburn TM and Gaston KJ (1996b) Spatial patterns in the species richness of birds in the New World. *Ecography* 19: 369–376.
Blondel J (1991) Birds in biological isolates. In: Perrins CM, Lebreton J-D, and Hirons GJM (eds.) *Bird Population Studies*, pp. 45–72. Oxford: Oxford Univ. Press.
Campbell B and Lack E (eds.) (1985) *A Dictionary of Birds*. Calton, UK: Poyser.
Collar NJ, Crosby MJ, and Stattersfied AJ (1994) *Birds to Watch 2. The World List of Threatened Birds*. Cambridge, UK: BirdLife Int.
Cramp S (ed.) (1988) *The Birds of the Western Palearctic*, vol. 5: Oxford: Oxford Univ. Press.
Fuller RJ (1982) *Bird Habitats in Britain*. Calton, UK: Poyser.
Gaston KJ and Blackburn TM (1996) The spatial distribution of threatened species: macroscales and New World birds. *Proc. R. Soc. London B* 263: 235–240.
Gaston KJ and Blackburn TM (2000) *Pattern and Process in Macroecology*. Oxford: Blackwell Science.
Grant P (1999) *Ecology and Evolution of Darwin's Finches*, 2nd ed. Princeton, NJ: Princeton Univ. Press.
Gregory RD, Greenwood JJD, and Hagemeijer EJM (1999) The EBCC atlas of European breeding birds: A contribution to science and conservation. *Biol. Conserv. Fauna* 102: 38–49.
Greenwood JJD, Gregory RD, Harris S, Morris PA, and Yalden DW (1996) Relations between abundance, body size and species number in British birds and mammals. *Philos. Trans. R. Soc. London B* 351: 265–278.
Hochberg ME, Clobert J, and Barbault R (eds.) (1996) *Aspects of the Genesis and Maintenance of Biological Diversity*. Oxford: Oxford Univ. Press.
Jones PH (1972) Succession in breeding bird populations of sample Welsh oak woods, Br. *Birds* 65: 291–299.
Lack D (1971) *Ecological Isolation in Birds*. Oxford: Blackwell Scientific Publications.
Lawton JH, Bignell DE, Bolton B, *et al.* (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature* 391: 72–76.
Morse DH (1975) Ecological aspects of adaptive radiation in birds. *Biol. Rev. Cambridge Philos. Soc.* 50: 167–214.
Moss D (1978) Diversity of woodland songbirds populations. *J. Anim. Ecol.* 47: 521–527.
Newton I (1998) *Population Limitation in Birds*. San Diego: Academic Press.
Newton I and Dale LC (1997) Effects of seasonal migration on the latitudinal distribution of west Palearctic bird species. *J. Biogeogr.* 24: 781–789.
Nilsson SG (1979) Density and species richness of some forest bird communities in south Sweden. *Oikos* 33: 392–401.
Ralph CJ (1985) Habitat association patterns of forests and steppe birds of northern Patagonia, Argentina. *Condor* 87: 471–483.

Schall JJ and Pianka ER (1978) Geographical trends in numbers of species. *Science* 201: 679–686.
Schluter D and Ricklefs RE (1993) Convergence and the regional component of bird species diversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 230–240. Chicago: Univ. of Chicago Press.
Sibley GC and Ahlquist JE (1990) *Phylogeny and Classification of Birds*. New Haven, CT: Yale Univ. Press.
Terborgh J (1980) Vertical stratification of a neotropical forest bird community. *Acta XVII Congr. Int. Ornithol.* 2: 1005–1012.
Wiens JA (1989) *The Ecology of Bird Communities: Volume 1. Foundations and Patterns*. Cambridge, UK: Cambridge Univ. Press.
Williamson M (1981) *Island Populations*. Oxford: Oxford Univ. Press.

Boreal Forest Ecosystems

Jennie R McLaren, University of Texas at Arlington, Arlington, TX, USA
Roy Turkington, University of British Columbia, Vancouver, BC, Canada

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Roy Turkington, volume 1, pp 521–532, © 2001, Elsevier Inc.

Glossary

Boreal A term derived from the Greek word for north reflecting the fact that this vegetation type, the boreal forest, occurs only in the Northern Hemisphere.
Permafrost That part of the soil profile that is permanently frozen and which forms a barrier to water drainage often resulting in a wet surface condition.
Podzol A soil type highly characteristic of boreal forests and developed as a consequence of podzolization.
Podzolization The process of acid leaching whereby clay, organic particles, and mineral ions (primarily iron and aluminum) are carried downwards and deposited in the B soil horizon, leaving an impoverished and leached A horizon. This occurs as a consequence of low temperatures and precipitation in excess of the needs of evapotranspiration.
Taiga Mostly used as a synonym for boreal forest, but more precisely it is a Russian word applied to Eurasian conifer forests described as damp and almost impenetrable. It is also described as a coniferous forest with no nonconiferous tree species except *Betula* and *Populus*.

Geography

The boreal forests are confined to the Northern Hemisphere. In North America the forest is a continuous vegetation belt stretching across the continent and spanning more than 10° latitude. Its northern limit is defined by transition to the treeless tundra; this limit extends from about 68° N in Alaska to 58° N on the west coast of Hudson Bay, and it reaches the Labrador coast at about 58° N (Larsen, 1980). In Fennoscandia it ranges from 56° N to 69° N (Esseen *et al.*, 1997), with the northern limit fringing the northern coast of Norway. Across Russia, except for a small border, the northern boundary more or less follows the Arctic coast. In parts of Siberia the boundary is up to 500 km inland (Larsen, 1980) (**Figure 1**). The southern limits of the boreal forest are more difficult to define because the boundaries are rarely sharp. In more oceanic areas the forest is bounded to the south by broad-leaved deciduous forest, and in continental areas transition is to parkland, dry grasslands, and semideserts. For example, in western Canada the southern transition is to subalpine forest, in central Canada it is to prairie grasslands, and in easterly regions it is to mixed deciduous forest.

Climate

There are strong relationships between the climate and the soils of the boreal regions. The characteristic nature of the climate to a large extent dictates the nature of the soils, including the permafrost, and these ultimately determine the plants and animals that live there.

General Limits

The climate of the boreal forest has been extensively documented by, for example, Hare and Ritchie (1972), Larsen (1980), Elliott-Fisk (1988), and Bonan and Shugart (1989). The boreal forest grows where winters are too long and summers too short to support temperate forests. This typically occurs when the growing season is less than 6 months and the frost-free period less than 4 months (or fewer than 4 months with temperatures higher than 1 °C). The boundary between boreal forest and tundra corresponds approximately with (1) the line south of which the temperature is higher than 6 °C for 4.5 months (Hare and Ritchie, 1972), (2) the position of the July 13 °C isotherm with marked departures in regions with montane or oceanic influences (Larsen, 1980), or (3) where there are less than 30 days with a daily mean temperature higher than 10 °C and where the cold season lasts 8 months (Walter, 1973).

Temperature

In more general terms, Elliott-Fisk (1988) describes the climate as cool, humid microthermal, with very cold winters of 7–9 months allowing persistence of snow cover during all but the brief (3 or 4 months), relatively cool summer season. For more than 6 months of the year, the mean temperature is below 0 °C and net radiation is negative. Maximum summer temperatures generally reach the low 20s and winter minimums in the −50s. However, yearly variation can be extreme such as in Verkhoyansk (30 °C in summer and −70 °C in winter), but these extremes are moderated near the coast, for example, in Umeå, Sweden (**Figure 2**).

Precipitation

Generally, the boreal zone is characterized by having a high proportion of the annual precipitation falling as snow. Throughout most of the boreal forest annual precipitation is low, varying between 250 and 1000 mm (**Figure 2**). There are extremes, however, with Fort Yukon in Alaska and

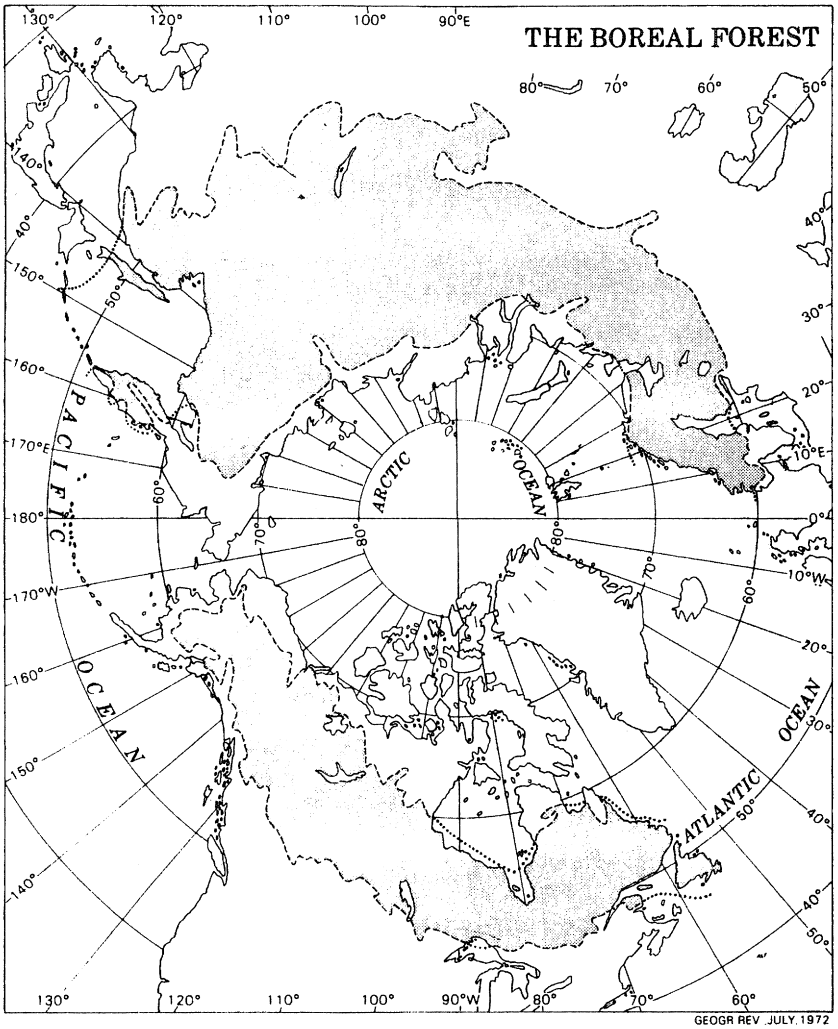


Figure 1 The circumpolar range of the boreal forest. About two-thirds of the area is in Eurasia. The sector in eastern Canada lies furthest from the North Pole (Reprinted from Hare FK and Ritchie JC (1972) The boreal bioclimates. *Geographical Review* 62: 333–365.).

Verkhoyansk in Siberia each recording 180 mm annually. Some areas of western Norway with up to 2000 mm of precipitation annually have been classified as boreal rain forest. Because of low temperatures and short growing seasons, evaporation rates are low and drought is uncommon. When drought does occur, forest fires can ravage vast areas of boreal forest.

Soils

Boreal forest soils are typically low in fertility and acidic, with a thin A horizon. The most characteristic soils are podzols. However, podzols occur in a wide range of climates, not only in boreal regions, and not all boreal regions are underlain by podzols. Climate, vegetation type, chemical composition of the substrate material, and topography are the major environmental influences that produce the typical boreal podzol. The combination of low temperatures and low pH impede decomposition processes and slow the rate of soil development. The soil surface may be covered by a mat of needles up to 3–7 kg m^{-2}. This mat of acidic, partly decomposed plant material is the mor litter layer. The slow but gradual decomposition of this layer continually releases a supply of organic acids that contribute to the leaching of organic particles and mineral ions (primarily iron and aluminum) from the surface soils and to the weathering of the parent material. In addition, soluble materials such as sodium, potassium, and calcium are washed out of the soil by water movement. As a consequence, the surface soils are relatively infertile and high in silica. In many cases in which the process of podzolization is prominent, the upper soil layers are gray or whitish in color. Beneath the leached layer is a zone in which materials leached downward by water accumulate, chiefly in iron–humus complexes. These deposits may be cemented into a hardpan, sometimes thick and strong enough to prevent root penetration to the lower soils. In the more extreme boreal forest climates the subsoil is permanently frozen (the permafrost). The combination of nutrients being largely tied up in the litter layer, an infertile A horizon, hardpans and permafrost results in most boreal forest trees having a shallow root network.

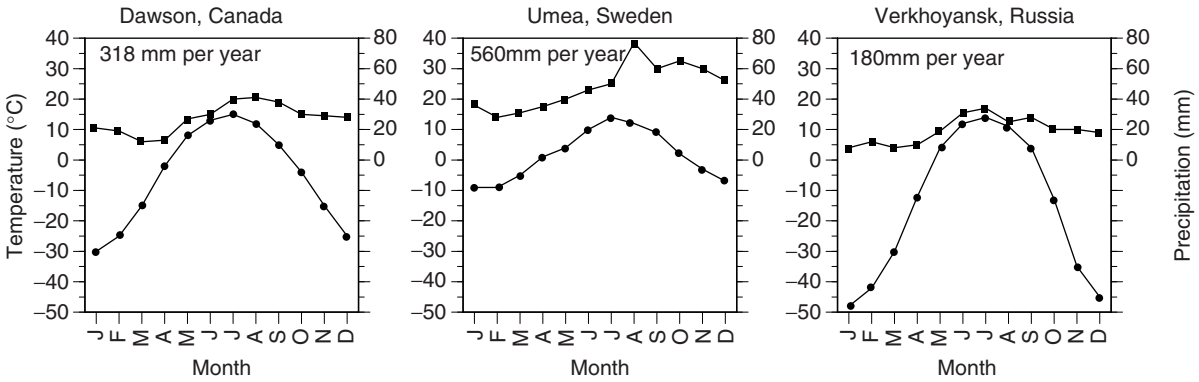


Figure 2 The annual variation of average temperature (●) and precipitation (■) of three locations in the boreal forest zone; Dawson in the Yukon Territories, Canada (altitude, 324 m), Umeå, Sweden (7 m), and Verkhoyansk, Russia (137 m).

Fire

Individual trees or local stands may be killed by windstorms, landslides, snow avalanches, erosion, ice storms, flooding, or insect attack, but the most important natural disturbance in the boreal forest is fire. Longer periods of unseasonably drier and warmer weather lead to a greater probability of fire, and it is the few major fires during extreme fire years that account for the majority of forests burned; this inevitably has enormous consequences for the wildlife that inhabit the forests, particularly those unable to take refuge in water bodies. In northwestern Canada and Alaska, between 60 and 80% of all fires are less than 5 ha in area, but in extreme fire years individual fires can burn up to 200,000 ha. Because fires often burn all or most of the forest floor, they influence organic matter accumulation, soil temperature, and soil moisture, and through these they impact major ecosystem processes such as nutrient cycling, energy flow, and productivity. The natural cycle of fire frequency in boreal forests of North America ranges from an average of 50–200 years to up to 500 years in wetter parts of eastern Canada. Fire frequencies in northern Sweden average from 110 to 155 years but may be as high as 270 years. Fire frequency varies considerably, both within and between regions, resulting in a mosaic landscape with different forest patches in various stages of succession. Factors contributing to the probability of a fire include tree species composition, stand structure, soil conditions, the amount and moisture content of the fuel, exposure, topography, and climate; there is much disagreement regarding whether the probability of a fire increases with stand age. Local fire chronologies indicate that fire frequency is highest in the interior, more continental areas due to a higher frequency of thunderstorms and lightning strikes; most natural fires are caused by lightning strikes. In the boreal forest of Fennoscandia there is apparently a north to south gradient with a generally lower fire frequency toward the north.

Vegetation

The boreal forest is the most continuous and extensive forest in the world. The North American and Eurasian forests are remarkably uniform in their appearance throughout their range, both in their physiognomic structure and in species composition. Typically, there is a simple canopy layer (15–20 m high) in which numerical dominance is maintained nearly everywhere by coniferous tree species belonging to four genera: spruce (*Picea*), pine (*Pinus*), fir (*Abies*), and larch (*Larix*) (**Figure 3**). Species of juniper (*Juniperus*), cedar (*Thuja*), and hemlock (*Tsuga*) also occur. Next is a shrub layer (typically 1 or 2 m tall) supporting mainly broad-leaved deciduous species that are also frequently present as successional components of the forest; deciduous species rarely achieve dominance, except in some postfire successions and in the mountain birch forest in Fennoscandia. Generally, the deciduous species belong to the genera willow (*Salix*), birch (*Betula*), poplar (*Populus*), and alder (*Alnus*). The herb layer typically is poorly developed but it is enmeshed in a well-developed ground layer of mosses, liverworts, and lichens. Despite the uniformity of appearance, the tree species are unique to a particular continent, i.e., there are no circumboreal trees (**Table 1**). The species making up the shrub layer and ground layer may be more wide ranging, and many of the mosses and lichens are circumboreal. Bryophytes and lichens are typically a more common component of the forest floor vegetation than vascular plants. Bryophytes usually dominate on mesic and moist sites and their diversity in boreal forest is higher than in most temperate or tropical forests. Lichens also contribute significantly to diversity in boreal forests, especially on drier and northern sites, and particularly in the ground layer of pine heaths and on rocky ground.

Because there are few species of trees, the boreal forest gives an impression of monotony, but this is misleading even though they are floristically impoverished compared to most other vegetation formations of the earth. Nevertheless, these forests are a complex mosaic of different plant communities, and they caused Cumming *et al.* (1996) to title their paper "Boreal Mixedwood Forests May Have No 'Representative' Areas." Some of the tree species are widely distributed (e.g., *Picea glauca* and *P. mariana*), but many other boreal species have more limited ranges, resulting in local and regional changes to forest composition. Throughout the North American boreal forest, a vegetation group dominated by P. *glauca* (sometimes codominated by *Abies balsamea*, *A. bifolia*, *Betula papyrifera*, *or Populus tremuloides*) is typically the climax

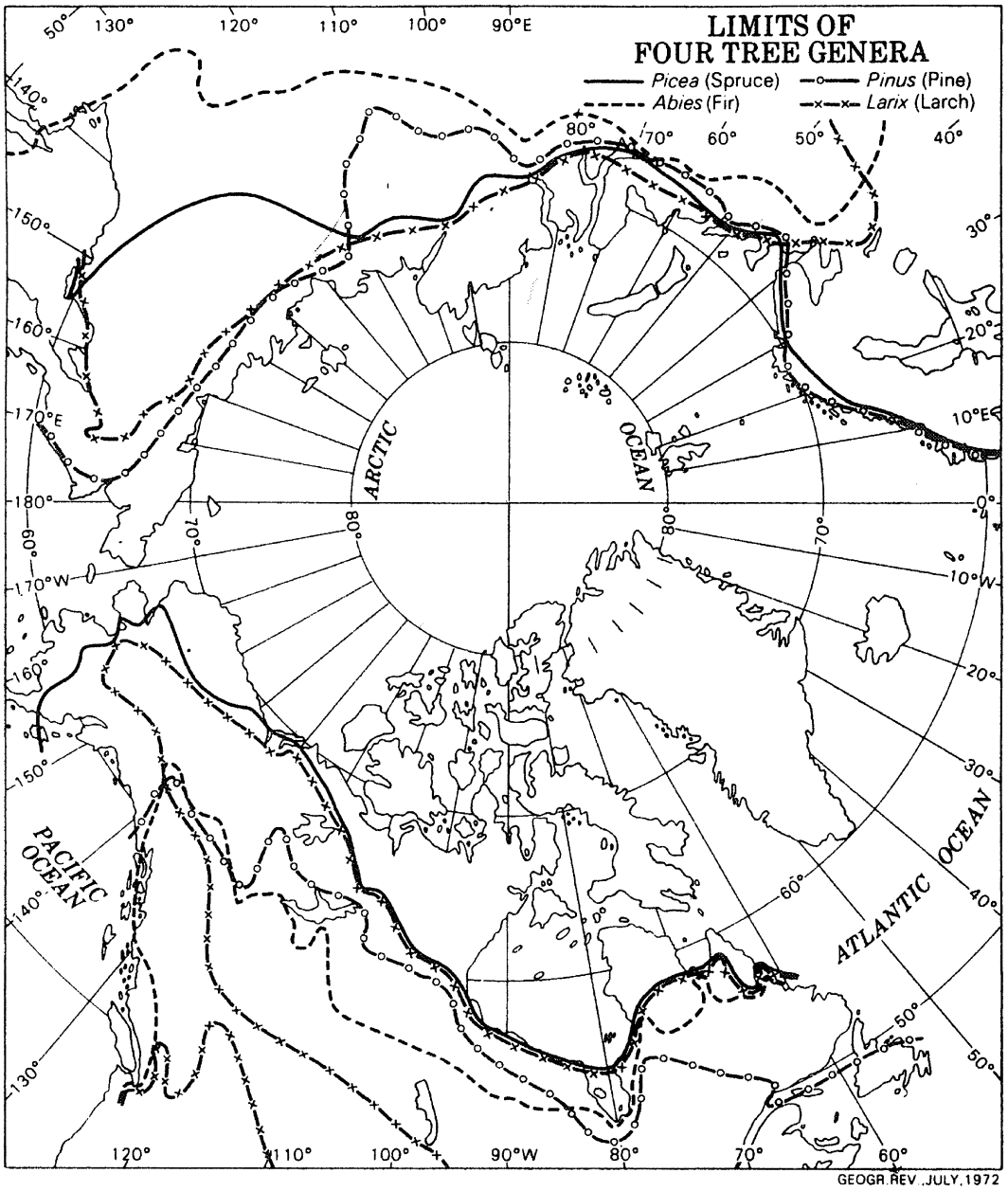


Figure 3 Limits of *Picea*, *Abies*, *Pinus*, and *Larix*, without regard to species. Limits are greatly generalized, being envelope curves around numerous disjunct areas. In particular, the central Alaskan distribution of *Larix* is probably disjunct from the Mackenzie Valley distribution (Reprinted from Hare FK and Ritchie JC (1972) The boreal bioclimates. *Geographical Review* 62: 333–365.).

community in relatively dry areas. A second group dominated by *P. mariana* occurs where soils are wetter, such as in bogs and muskegs. *Picea mariana* is not strictly characteristic of wet areas because it also flourishes after fire. In the European forests the Scots pine (*Pinus sylvestris*) is frequently dominant on drier soils and the Norway spruce (*Picea abies*) occurs on moister sites.

Although there are no abrupt transitions in the vegetation, changes in the dominant tree species, in the subdominant shrubs, and in the herbaceous layers can be detected. Various attempts have been made to survey and describe latitudinal and longitudinal gradients, based primarily on climatic distinctions, and to describe the associated vegetation assemblages. For example, Larsen (1980) described seven regions in the boreal forest of North America, and Rowe (1972) described 35 regions. However, rather than describe each of these, we focus on some of the recent classifications.

Latitudinal Classification

Hare and Ritchie (1972) recognized three latitudinal zonal divisions of the boreal forest in North America: a southern closed forest, a central lichen woodland, and a northern forest–tundra ecotone.

Table 1 Major species of boreal forest trees by longitudinal section[a]

Genus	*North America, (55° W–160° W)*	*Northern Europe (5° E–40° E)*	*Western Siberia (40° E–120° E)*	*Eastern Siberia (120° E–170° E)*
Conifers				
Picea (spruce)	*glauca* *mariana*	*excelsa*[b]	*obovata*	*obovata* *jezoënsis*
Abies (fir)	*balsamea*		*sibirica*	*nephrolepis*
Pinus (pine)	*banksiana*	*sylvestris*	*sibirica* *sylvestris*	*sylvestris* *pumila* *cembra*
Larix (larch)	*laricina*		*sibirica* *sukachzewski*	*dehurica*
Hardwoods				
Populus (popular)	*tremuloides* *balsamifera*	*tremula*	*tremula*	*tremula* *suaveolens*
Betula (birch)	*papyrifera* *kenaica*	*pubescens* *verrucosa*[b] *kusmisscheffi*	*verrucosa*[b] *pubescens*	*ermani*
Alnus (alder)	*tenuifolia* *crispa* *rugosa*	*incana*	*fruticosa*	*fruticosa*
Salix (willow)	*Salix* species	*Salix* species	*Salix* species	*Salix* species

[a]Certain species of nonboreal type range into southern parts of the forest. In North America, for example, these include the white and red pines (*Pinus strobus* and *P. resinosa*), *Thuja occidentalis*, and species from *Ulmus*, *Fraxinus*, and *Acer*. In Europe and Asia, species of *Quercus*, *Tilia*, and *Fraxinus* also have a limited boreal range. The taxonomies *of Betula Alnus* and *Salix* are highly confused, and this table is hence only an approximate guide.

[b]*Picea excelsa* should be *Picea abies* (Norway or European spruce) and *Betula verrucosa* should be *Betula pendula* (European birch).

Source: Reprinted from Hare FK and Ritchie JC (1972) The boreal bioclimates. *Geographical Review* 62: 333–365.

Closed Forests

The closed forest dominates vast areas of the southern boreal forest, and it occurs on a wide range of soil types and topographies. These forests have a closed crown with a moist, deeply shaded floor. The spruce–feathermoss community is characteristic of this zone, with either white or black spruce dominant. The *P. mariana*–feathermoss forests have a fairly uniform, moderately dense tree stratum, with an almost continuous ground cover of bryophytes. In contrast, the *P. glauca* forests are more irregular and open, having an understory of broad-leaved shrubs and herbs and a patchy bryophyte distribution. The most common moss in the black spruce forests is *Pleurozium schreberi*, whereas *Hylocomium splendens* is more characteristic of the white spruce and mixed woodlands.

Lichen Woodland

These forests are more open with a discontinuous shrub layer and abundant lichens. The transition from the southern closed forests to these lichen woodlands is usually smooth or mosaic-like but can sometimes be quite abrupt as in northern Manitoba and Saskatchewan. In some cases, a Western *Stereocaulon paschale*- and an Eastern *Cladonia stellaris*-dominated woodland have been identified. *Picea mariana* and *P. glauca* are the dominant trees in these forests, with *P. glauca* declining to the north. Because these forests are more open, light penetration to ground level is much greater than in closed forests. Although these trees may be as tall (but less dense) as those in closed forests, the additional light promotes branching to ground level. The additional light often promotes a discontinuous layer of health plants such as crowberry (*Empetrum nigrum*) and bilberry (*Vaccinium myrtillus*).

Forest–Tundra Ecotone

Here, there are scattered and isolated trees, often deformed or prostrate, in a tundra landscape. Although this type of forest is located north of the limit of continuous forest, it is still considered boreal forest. The ecotone is more than 300 km wide in Quebec, up to 225 km wide in Central Canada, and it narrows at both its eastern and western ends.

A Longitudinal Classification

There is a remarkable similarity between the vegetation of the boreal forests in eastern North America and that in eastern Asia, with many identical or closely allied genera and sometimes species. These forests contain many different species, whereas those of the Euro-Siberian region contain few. Of all the tree species in the North American boreal forest, only *P. glauca* extends from the Bering Straits across Alaska and Canada to Newfoundland. *Picea mariana*, usually found only

on poor or wet soils, is found at the timberline toward the Arctic, and *Larix laridna* is found in the continental regions. Only two species, *P. sylvestris* and *P. abies*, are of any real importance in the boreal zone of Europe. Only in eastern regions of Europe is *P. abies* replaced by the closely related *Picea obovata*, whereas additional species are being added to the forest in Siberia (*Abies sibirica*, *Larix sibirica*, and *Pinus sibirica;* **Table 1**). Moving east, spruce gradually declines from the forest until it is entirely absent in eastern Siberia.

Peinado *et al.* (1998) analyzed the vegetation of the North American boreal forests. They identified three major sections (**Table 2**) and classified them as eight major groups (**Table 3**). The three sections coincide well with La Roi's (1967) classification for the same regions. Qian *et al.* (1998) examined longitudinal patterns of plant diversity in the North American boreal forests, focusing specifically on the southern closed forests. The central section has a higher species and genera diversity than the western and eastern sections (**Table 4**). White spruce forests are always more diverse than black spruce forests; this is a reflection of the diversity of herbaceous plants and bryophytes and not the diversity of woody plants (**Figure 4**). The diversity of white spruce forests is rather similar between western and eastern sections, but the diversity of black spruce forests is much higher in the west than in the east (**Table 4**). The diversity of bryophytes is remarkably consistent across the continent, but again there are more bryophytes in white spruce than in black spruce forests (**Figure 4**).

Animals

The boreal forest is home to many animals. It is the winter home of the migratory caribou and reindeer and the permanent home of many others. The wolf and lynx are the major predators of the boreal forest. Some of the best examples of population cycles in animals are described from the boreal forest regions; for example, lynx (*Lynx canadensis*), snowshoe hares (*Lepus americanus*), arctic ground squirrels (*Spermophylus parryi*), red squirrel (*Tamiasciurus hudsonius*), and boreal red-backed vole (*Clethrionomys rutilus*) in northern Canada and microtine rodents, owls, capercaillie (*Tetrao urogallus*), black grouse (*Tetrao tetrix*), mountain hare (*Lepus timidus*), and red fox (*Vulpes vulpes*) in Eurasia. The causal relationships of these cycles have not been fully explained but many of the North American examples are synchronized with the snowshoe hare cycle, and Eurasian examples are synchronized with microtine rodent cycles.

In North America, other inhabitants of the boreal forest include moose, black bear, grizzly bear, deer, wolverine, coyote, marten, beaver, porcupine, sable, voles of the genus *Microtus*, chipmunks, shrews, and bats. Typical animals of the boreal forest vary slightly more from location to location than do the plants. The moose which browse on willow, birch, alder, and water plants, and the beaver which feeds on aspen, are widespread. Many birds also inhabit the boreal forest; for example, great horned owl, goshawk, spruce grouse, ruffed grouse, nuthatchers, juncos, and warblers. Brown bears inhabit the boreal forest in Eurasia.

In the boreal zone of Eurasia, the diversity of mammalian herbivores is highest in the interior of the continent and declines to the east. Across Eurasia, species richness of mammalian herbivores is positively correlated to warm climate, the number of hardwood species, and the area of the boreal forest. Across North America, species richness of mammalian herbivores increases as the length of the growing season and the number of coniferous tree species increase (**Figure 5**). Given this information, it appears that indirect measures of primary productivity as well as the number of tree species can accurately predict species richness of mammalian herbivores. Bird diversity decreases from west to east across both the North American boreal forests and the Eurasian boreal forests. In Fennoscandia the

Table 2 Major dominant, codominant, or abundant species identified in the three major longitudinal sections in Peinado *et al.* (1998) and La Roi's (1967) classifications of North American boreal forests

Classification	
Peinado *et al.* (1998)	La Roi (1967)
The western section (northeastern British Columbia, southern Yukon, and Alaska)	
Abies lasiocarpa	Populus/Salix/Shepherdia
Pinus contorta var. *latifolia*	
Picea engelmanii	
The central section (eastern British Columbia, Alberta, Saskatchewan, and Manitoba)	
Abies lasiocarpa	*Lonicera/Rubus pubescens/Lathyrus ochroleucus*
Abies bifolia	*Lonicera/Vaccinium vitis-idaea-Geocaulon*
Abies balsamea	
The eastern section (all forests east of Lake Winnipeg, Manitoba)	
Abies balsamea	*Kalmia-Picea mariana*
Pinus banksiana	*Abies balsamea*

Source: Reproduced from Peinado M, Aguirre JL, and de la Cruz M (1998) Aphytosociologidal survey of the boreal forest (*vaccinio-piceetea*) in North America. *Plant Ecology* 137: 151–202, and La Roi GH (1967) Ecological studies in the boreal spruce–fir forests of the North American taiga. 1. Analysis of the vascular flora. *Ecological Monographs* 37: 229–253.

Table 3 Major dominant, codominant, or abundant species identified in the eight groups in Peinado *et al.*'s (1998) classification of North American boreal forests

Group	*Western*	*Central*	*Eastern*
I	*Picea glauca* and/or *Pinus contorta* var. *latifolia* *Populus tremuloides*	*Abies bifolia* *Abies lasiocarpa*	*Picea glauca* *Picea mariana* *Abies balsamea* *Betula papyrifera*
II	*Picea mariana*	*Abies balsamea*	*Picea rubens* *Abies balsamea* *Betula alleghaniensis*
III			*Alnus rugosa*
IV			*Pinus banksiana*

Source: Reproduced from Peinado M, Aguirre JL, and de la Cruz M (1998) Aphytosociologidal survey of the boreal forest (*vaccinio-piceetea*) in North America. *Plant Ecology* 137: 151–202.

Table 4 Mean number of species and genera in White spruce and Black spruce-dominated ecosystems in three geographic sectors of the boreal forest of North America[a]

	Western North America		*Central North America*		*Eastern North America*	
	Species	*Genera*	*Species*	*Genera*	*Species*	*Genera*
White spruce ecosystems						
No. of plots	13		10		11	
Woody plants	13	11	18	14	13	10
Herbaceous plants	28	25	33	30	25	24
Bryophytes	23	19	25	20	26	22
Vascular plants	41	35	51	44	38	33
All plants	65	54	77	63	65	55
Black spruce ecosystems						
No. of plots	7		10		9	
Woody plants	14	10	14	11	12	11
Herbaceous plants	23	19	28	24	19	16
Bryophytes	15	13	15	13	14	12
Vascular plants	37	30	41	36	31	27
All plants	53	42	57	49	46	39

[a]The boreal forest can be latitudinally divided into three subzones: a wide zone of closed forest in the south, a narrow band of forest–tundra ecotone in the north, and a lichen woodland in the middle. These data cover almost the entire longitudinal range but focus on the southern closed forest (Modified and adapted from Qian H, Klinka K, and Kayahara GJ (1998) Longitudinal patterns of plant diversity in the North American boreal forest. *Plant Ecology* 138: 161–178.).

diversity of forest birds decreases northwards; in Finland, this occurs only in pine forests and not in spruce.

It seems that the boreal forests of Canada, and possibly Russia, differ from those in northern Fennoscandia in that small herbivore biomasses reach much higher levels and are dominated by species of hare rather than voles. In addition, the densities of many fewer species in the boreal forests of Canada are correlated with the dominant herbivore relative to the situation in Fennoscandia.

Tree death and decaying wood provide a variety of habitats for an enormous number of invertebrates. For example, in Sweden approximately 1000 species of beetle are dependent on dead trees. The most diverse fauna on snags is found during the first 2 years after the tree has died. Spruce logs have a more diverse invertebrate fauna than pine, but many invertebrates can inhabit both. Four typical stages in the succession of invertebrates on spruce logs in boreal forests have been described. Initial colonization is by bark beetles and other primary cambial eaters along with their associated parasitoids, predators, and detritivores. Subsequent stages have been described in detail by Esseen *et al.* (1997).

Succession

Many studies of the boreal forest have shown a remarkably close fit between the plant community and the environment, especially with reference to moisture conditions and fertility level. This relationship is not static and communities are always in a process of change. The sequence of primary successional change is typically described as beginning in fens and then progressing as follows:

$$\text{Fen} \rightarrow \text{swamp} \rightarrow \text{bog} \rightarrow \text{muskeg}$$

$$\rightarrow \text{invading trees} \rightarrow \text{forest}$$

Fens develop under alkaline conditions on a peat substrate usually with standing water. Bogs are specialized communities of shrubs and herbs growing on a wet, acidic, peat substrate. The sequence becomes drier and more mineral poor as muskeg develops, and this is followed by tree invasion.

The usual sequence of secondary succession is described as

Disturbance → annual plants → perennial herbs and shrubs → larger shrubs and trees → forests

However, these descriptive sequences mask a dynamic system under the influence of frequent disturbances, and the boreal forest is better considered as a disturbance forest. This means that the concept of the vegetation climax is probably not applicable because most of the tree species in these forests are incapable of self-perpetuation in climax state. They are instead, adapted to fire, and the boreal forest is usually maintained in a pre-climax state by frequent fires. Fire is the primary disturbance factor interrupting the successional process, and the frequency and intensity of fires will determine the nature and stage of succession. Local topographic, soil, moisture, and microclimate conditions and tree composition will determine the local fire regime, and if the local regional environment is patchy so also will be the frequency of fires. The consequence will be a vegetation mosaic, with each patch being at a different stage of succession and representing a different degree of recovery since the last fire. In this sense, the boreal forest climax is not a stand dominated by white or black spruce, but is a mosaic in which all stages of a successional or regeneration cycle are represented. It has been argued that if fires could be excluded from a lichen woodland for at least 200 years, then it is likely, that a closed-forest spruce–-feathermoss would develop; however, fires usually occur more frequently and so the lichen woodland is perpetuated. In the absence of fire, however, it is also argued that the spruce–moss forests are able to maintain themselves for thousands of years (Pollock and Payette, 2010).

Nevertheless, immediately following a fire some species are characteristic pioneer species. Among herbaceous plants, *Epilobium angustifolium* and *E. latifolia* (fireweeds) are the most obvious. Density and cover of *Salix*, *Betula*, and *Alnus* are typically high in early pioneer stages and greatly decrease in

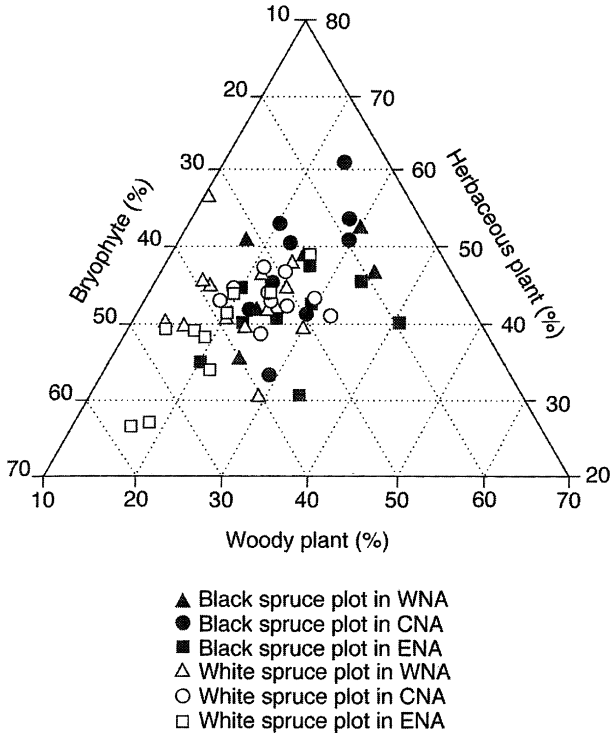


Figure 4 Triangular ordination showing sample plots in the white spruce and black spruce ecosystems of the North American boreal forest. Each point represents a single sample plot of 9 ha (300 × 300 m). Thirty-four of the plots were sampled in *Picea glauca* forest and 26 in *Picea mariana*. CNA, central North America; ENA, eastern North America; WNA, western North America (Reprinted from Qian H, Klinka K, and Kayahara GJ (1998) Longitudinal patterns of plant diversity in the North American boreal forest. *Plant Ecology* 138: 161–178.).

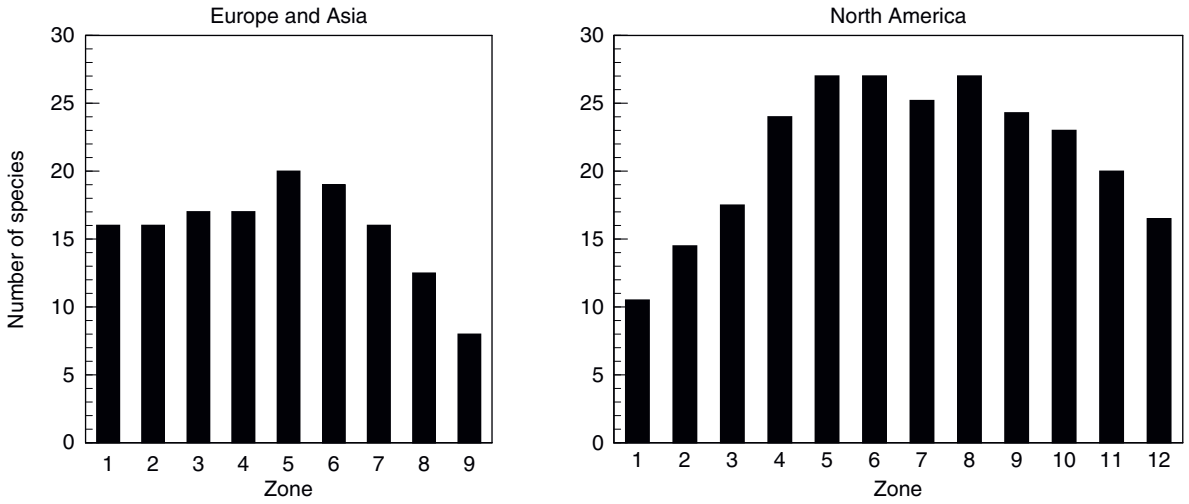


Figure 5 The number of mammalian herbivore species along a gradient from west to east (zones 1–9 across Europe and Asia and zones 1–12 across North America). Note the decline in species number as the continents meet at the Bering Strait (Adapted from Danell K, Lundberg P, and Niemela P (1996) Species richness in mammalian herbivores: Patterns in the boreal zone. *Ecography* 19: 404–409.).

later stages when they are intolerant of the shady conditions imposed by the establishing *Picea*, *Abies*, *Firms*, and *Populus*. Many questions remain concerning the nature of succession in the boreal forest. There are many regional differences, and local site conditions exert considerable influence of the successional process – both on its path and on its "final" form.

Conservation

In both Canada and Russia, vast areas of boreal forest are being cut annually, and the situation is worse in Fennoscandia. Although wood cutting is the biggest danger to biodiversity, there are additional threats from mining, pollution, road building, and dam construction. Clear-cutting vast areas of forest has profoundly altered the landscape structure in northern Sweden and in central Canada, resulting in habitat loss, habitat alteration, and fragmentation. Clearly, this will have mostly negative influences on both animal and plant biodiversity, even though some species will benefit in the short-term, i.e., weedy species such as *Epilobium angustifolium*, *Deschampsia flexuosa*, and *Calamagrostis purpurea* that are adapted to fire and other disturbances. Few data are available on the long-term effects of clear-cutting in the boreal forest, but studies from other forests suggest that herb communities do not recover in logging cycles of 40–150 years. Rotational cutting and clear-cutting will inevitably influence, or eliminate, natural fire regimes and will lead to changes in tree species composition and forest structure, a general reduction in stage age, and reduced input from coarse woody debris (Esseen *et al.*, 1997), all of which will have profound impacts on the natural biodiversity of the boreal forest.

Biodiversity must be preserved at all scales from the genetic variation within a population to heterogeneity occurring at the landscape level. There are two major approaches to preserving biodiversity in these areas: (1) sustainable management and (2) reserves of natural areas. The first approach is more applicable to areas that are already subject to intense management such as Fennoscandia, whereas the second approach can be applied where vast, relatively untouched areas still exist such as in parts of Canada and Russia.

Sustainable Management

To maintain biodiversity, we must also preserve or simulate the processes, mostly natural disturbances that produced the heterogeneity in the first place. Microscale heterogeneity may be enhanced by gap disturbance and by coarse woody debris. Larger scale heterogeneity may be enhanced by fire or insect outbreak. The management of forests itself may have effects on biodiversity, where species dependent on continuous forest cover (bryophytes, lichens, fungi) and carabids are negatively affected by management, but vascular plants positively affected (Paillet *et al.*, 2010). The emulation of natural disturbance in management techniques has been proposed as promising new technique for sustainable forest management, but the lack of consistent views on the meaning of the emulation of natural disturbance raises concerns about its use in policy making (Klenk *et al.*, 2008).

Reserves of Natural Areas

Reserves alone are not sufficient to conserve biodiversity in forests, but any such initiative for maintaining biodiversity in the boreal forest should include at least four components:

1. Large areas representative of southern closed-canopy forest, lichen woodland, the forest–tundra ecotone and each of the longitudinal elements within these must be protected.
2. In selecting these areas, attention must be given to the animal inhabitants, their abundance or rarity, and their migratory behavior, if any.
3. The minimum reserve size should be large enough to incorporate natural disturbance and maintain ecological processes.
4. The decision-making process should consider the "floating reserve" strategy (Cumming *et al.*, 1996) in which portions of a protected area could be periodically replaced in response to aging of components, unexpected large-scale disturbance, or refinements in conservation objectives.

Although much of the world's boreal forest still remains intact with little impact from man, the impact from man has been immense in many areas. In the new millennium, there will continue to be an increasing demand on the world's resources, including the boreal forests. Indeed, man's impact has already been substantial in Fennoscandia (less than 5% of the boreal forest remains in a natural or seminatural state) and is increasing in the North American forests. It is tempting to argue that these forests are the world's largest vegetation type, occupying vast areas (approximately 14 million km^2) and account for approximately 12% of the world's biomass, so the incursion of man is likely to have little impact. As described in this article, although the boreal forest is less diverse than most of the other world's vegetation types, it is nevertheless a complex mosaic of patches with changing species composition both longitudinally and latitudinally, with no representative areas, and home to many species of mammals, birds, and other animals.

See also: Arctic Terrestrial Ecosystems. Fires, Ecological Effects of. Latitudinal Gradients of Biodiversity. North America, Patterns of Biodiversity in. Timber Industry

References

Bonan GB and Shugart HH (1989) Environmental factors and ecological processes in boreal forests. *Annual Review of Ecological Systematics* 20: 1–28.

Cumming SG, Burton PJ, and Klinkenberg B (1996) Boreal mixedwood forests may have no "representative" areas: Some implications for reserve design. *Ecography* 19: 162–180.

Danell K, Lundberg P, and Niemela P (1996) Species richness in mammalian herbivores: Patterns in the boreal zone. *Ecography* 19: 404–409.

Elliott-Fisk DL (1988) The boreal forest. In: Barbour MG and Billings WD (eds.) *North American Terrestrial Vegetation*, pp. 33–62. Cambridge, UK: Cambridge University Press.

Esseen PA, Ehnstrdm B, Ericson L, and Sjöberg K (1997) Boreal forests. *Ecological Bulletin* 46: 16–47.

Hare FK and Ritchie JC (1972) The boreal bioclimates. *Geographical Review* 62: 333–365.

Klenk N, Bull G, and Cohen D (2008) What is the "END" (emulation of natural disturbance) in forest ecosystem management? *Canadian Journal of Forest Research* 38: 2159–2168.

La Roi GH (1967) Ecological studies in the boreal spruce–fir forests of the North American taiga. 1. Analysis of the vascular flora. *Ecological Monographs* 37: 229–253.

Larsen JA (1980) *The Boreal Ecosystem*. New York: Academic Press.

Paillet Y, Berges L, Hjalten J, *et al*. (2010) Biodiversity differences between managed and unmanaged forests: Meta-analysis of species richness in Europe. *Conservation Biology* 24: 101–112.

Peinado M, Aguirre JL, and de la Cruz M (1998) A phytosociologidal survey of the boreal forest (*vaccinio-piceetea*) in North America. *Plant Ecology* 137: 151–202.

Pollock SL and Payette S (2010) Stability in the patterns of long-term development and growth of the Canadian spruce-moss forest. *Journal of Biogeography* 37: 1684–1697.

Qian H, Klinka K, and Kayahara GJ (1998) Longitudinal patterns of plant diversity in the North American boreal forest. *Plant Ecology* 138: 161–178.

Rowe JS (1972) *Forest regions of Canada, Canadian Forestry Service, Publ. No. 1300*. Ottawa, Ontario: Department of the Environment.

Walter H (1973) *Vegetation of the Earth in Relation to Climate and the Eco-Physiological Conditions*. Berlin: Springer-Verlag. (Wieser Trans J).

Breeding of Animals

David R Notter, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
Beate Scherf and Irene Hoffmann, Food and Agriculture Organization of the United Nations, Rome, Italy

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by David R. Notter, volume 1, pp 533–545, © 2001, Elsevier Inc.

Glossary

Breed "Either a subspecific group of domestic livestock with definable and identifiable external characteristics that enable it to be separated by visual appraisal from other similarly defined groups within the same species or a group for which geographical and/or cultural separation from phenotypically similar groups has led to acceptance of its separate identity" (FAO, 1999).

Breeding line Any distinctive livestock population. The term may be synonymous with "breed" but is also often applied to somewhat distinct subpopulations within a breed or to breeding populations, either purebred or crossbred, that are of fairly recent origin. Breeding lines may exist within a breed, and newly created lines may become recognized breeds.

Conservation Generally defined as the planned management of a natural resource to prevent exploitation, destruction, or neglect. Here, further specified as applying to livestock genetic resources and encompassing both *in situ* and *ex situ* maintenance of livestock genetic resources as well as the wise use and active development of those resources in structured breeding programs to meet appropriate development objectives.

Crossbreeding/crossbred Crossbreeding is the mating of animals of different breeds to produce commercial market animals or breeding animals. The resulting crossbred animals often exhibit improved fitness and performance compared to their purebred parents.

Cryopreservation Long-term storage of gametes, embryos, or somatic cells in liquid nitrogen.

***Ex situ* conservation** Maintenance of a breed outside the environment and agricultural production system(s) in which it was developed and used. *Ex situ* conservation often involves cryopreserved gametes and/or embryos but may also involve live animals.

Genome The hereditary information contained in the DNA of eukaryotic organisms, including both genes and noncoding sequences of the DNA.

Germplasm Sources of hereditary material. In animal breeding the term is commonly applied to breeding animals and to fresh or frozen sperm cells, ova, and embryos.

***In situ* conservation** Maintenance of a breed in the environment and agricultural production system(s) in which it was developed and used.

Phenotype The observed characteristics of an individual, such as the color, milk production, body weight, etc. The phenotype is a reflection of both the genome and its interactions with the environment during development.

Purebred An animal that is a member of a recognized livestock breed. Breed membership may be determined by pedigree records, geographical location, or knowledge of the breeding structure of the herd or flock. Animals that possess the full range of characteristics commonly associated with a specific breed are sometimes also designated as purebreds, but the designation must be recognized as subjective in these cases.

Introduction

Domestic animals make crucial contributions to human well-being as sources of food, fiber, power, and many other products and services. Worldwide, domestic animal production centers on a small number of globally distributed species, including cattle, pig, sheep, goat, horse, donkey, rabbit, chicken, duck, goose, and turkey. These are augmented by a similar array of regionally important food- and fiber-producing species, such as water buffalo, dromedary and Bactrian camel, llama, alpaca, yak, reindeer, and guinea pig. Species-level domestic animal biodiversity is thus very limited, representing *in toto* approximately 20 completely domesticated, and perhaps an additional 25 partially domesticated, species. Genetic diversity within domestic species, however, is vast especially for the globally distributed species. Domestic animal genetic diversity evolved in synchrony with human dispersal as domestic animals accompanied their owners from their centers of domestication to every corner of the globe. Importantly, humans also provided these animals with the marginal levels of care, shelter, protection, and supplemental food that facilitated their adaptation to the diverse environments that they encountered. The resulting environmental adaptations, coupled with selection to meet a variety of human needs and preferences, resulted in the creation of an incredibly rich domestic animal fauna.

In the past three to four decades, an increasingly intense process of industrialization and concentration of production has occurred, particularly in egg- and meat-producing chickens and, to a smaller extent, in pork and milk production (Gura, 2008; FAO, 2010c). As a consequence, substantial erosion of between- and within-breed genetic diversity has occurred, not only in developed countries but also increasingly in developing countries (FAO, 2010b). Thus the potential of humankind to adapt production to changing demands and climate increasingly relies on a declining pool of farm animal genetic diversity.

Overview of Livestock Genetic Diversity

Breeds of Livestock

Populations of domestic livestock are traditionally subdivided and isolated from other subpopulations by both distance and controlled breeding. These subpopulations are generally known as "breeds." A single, widely accepted definition of a livestock breed does not exist, but livestock breeds are approximately analogous to the subspecies or races of wild species. Thus, the Food and Agriculture Organization of the United Nations (FAO) defines a livestock breed as:

> Either a subspecific group of domestic livestock with definable and identifiable external characteristics that enable it to be separated by visual appraisal from other similarly defined groups within the same species or a group for which geographical and/or cultural separation from phenotypically similar groups has led to acceptance of its separate identity (FAO, 1999).

The concept of a livestock breed arose in nineteenth-century England, where influential livestock breeders, including the noted Robert Bakewell, began to improve the native livestock of the region by controlled mating and selection of presumed superior individuals. Significantly, their activities included the establishment of "herd books" for most mammalian breeds which enumerated the pedigrees of animals within each breed so that breed membership was confirmed by records of ancestral ties to presumed elite foundation animals. These so-called "purebred" livestock thus had pedigree documentation of breed identity. Similar activities in poultry resulted in establishment of many avian breeds, although individual pedigree recording was not usually practiced for these smaller, more prolific species. This breed concept became widespread in Western Europe and in areas that were settled or influenced by Western Europeans. By the early twentieth century, a central part of livestock breeding in these areas was the identification and use of distinct, ostensibly superior, genetic types and the establishment of breeds to propagate these types.

The Western European definition of a breed had little meaning in most of Asia and Africa, where extensive pedigree recording does not appear to have arisen. However, selective breeding certainly did occur in these areas, and regionally distinct livestock types arose and were propagated. Isolation of these types was generally by distance or selective matings based on adaptation to local environmental conditions or obvious phenotypic traits such as color or horn shape rather than on pedigree information. However, pedigree information was occasionally accorded great value, as in the establishment of the Arabian horse. Today, the term breed has become globally pervasive and is used to describe any identifiable subpopulation possessing relatively distinct characteristics and maintained with a reasonable level of reproductive isolation.

Genetic Structure of Livestock Populations

Livestock breeds represent a partitioning of the genetic diversity within each species. Through many generations of selection and relative reproductive isolation, each of these subpopulations has become a distinctive genetic entity possessing specific combinations of genes and associated phenotypic characteristics. The core value of these breeds lies in the reliable access they provide to these combinations of genetically defined characteristics. For example, ewes of the prolific Russian Romanov sheep breed are capable of producing three to six lambs at a time. Sheep breeders interested in improving prolificacy can thus rely on this breed as a source of useful genes.

The traditional (i.e., mid-twentieth century) genetic structure of most livestock species was thus characterized by many regionally isolated subpopulations. In addition, several distinct subpopulations could often be identified within each region, maintained in more or less pure form by breeders who favored animals of a specific type. These dual mechanisms for maintaining reproductive isolation fueled development of levels of observed genetic diversity that are generally far in excess of those observed in wild species.

It would be incorrect, however, to view the livestock breeds as fixed, immutable genetic entities. Migration, or the exchange of genetic material, among breeds is not uncommon. Even among breeds that maintain pedigree records, decisions to "open" the herd book occasionally occur to allow incorporation of desirable animals from outside the breed. Breeds wax and wane in popularity. Some breeds merge, others are absorbed by more popular breeds, and a few simply lose favor and gradually disappear. Breed evolution is thus an accepted part of livestock breeding. This traditional genetic structure of livestock breeds is remarkably similar to that proposed by Wright (1940) as optimal for evolution in natural populations, with relatively high levels of reproductive isolation, modest effective population sizes, diverse selection pressures, and levels of migration that provide for periodic infusions of genes without seriously compromising the genetic integrity of the subpopulations.

Livestock breeds in most cases possess very significant amounts of within-breed genetic diversity. Although breeds can rightly be viewed as mildly inbred lines, they do not approach the levels of genetic uniformity commonly found in plant breeding lines in which self-fertilization and clonal propagation can strongly limit, and in some cases completely eliminate, genetic diversity. Thus, essentially all livestock breeds retain significant evolutionary potential and can undergo significant genetic changes in response to selection, either natural or artificial. Livestock breeds retain the capacity to change, sometimes radically, in response to changes in breeders' preferences or market demands without losing fundamental characteristics of the breed.

Finally, within each of the domestic species, a very large number of individuals exist that do not exhibit or possess unambiguous breed ancestry. Worldwide, many domestic animals commonly display apparent breed affinity, manifested by the color or morphology of recognized local breeds, but lack either the pedigree documentation or the full array of phenotypic characteristics necessary to confirm breed identity. For example, more than 90% of the dairy cows in the United States would be visually classified as Holsteins, but at most one-third of these would be recorded in herd books as purebred animals. The proportion of animals recorded in herd

books is much lower, rarely exceeding 10%, in cattle used for meat production, sheep, and goats.

Many other animals are obvious crossbreds, showing mixtures of specific breed characteristics that suggest a somewhat predictable breed ancestry. Also, large numbers of animals are often truly nondescript, exhibiting a collage of various breed characteristics that indicate a more diverse ancestry. The value of crossbred or nondescript animals as a genetic resource is hotly debated. Populations are often large and well adapted to prevailing environmental conditions. Although adapted local breeds are often viewed as "contaminated" by crossing with less well adapted, imported breeds, these populations also provide opportunity for creation of new genetic combinations and for selective elimination of undesirable breeds. Thus, these animals represent a potentially useful genetic resource but lack the predictability of the pure breeds for use in commercial livestock production.

Current Trends in Livestock Diversity

The most significant event in global genetic resource utilization in the late twentieth century was the development of highly productive livestock breeds which are now widely distributed globally. These breeds, for the most part, arose in the developed nations of the temperate zones and have had a remarkable impact on livestock production worldwide. They include the Holstein dairy cow, capable of producing up to 60 kg of milk per day for 300 days; the White Leghorn layer chicken, which can produce up to 325 eggs per year; the modern broiler chicken, which can reach a market weight of 2 kg in less than 35 days and on less than 3.5 kg of feed; and Merino sheep with wool fiber diameters of 13–15 μm.

The development and continued improvement of these global breeds is viewed in many quarters as the capstone of modern animal breeding. Large population sizes, detailed recording of performance, and intensive selection fueled their development. Well-organized breeders' groups and, more recently, multinational corporations are involved in their propagation. If we extrapolate figures from Steinfeld *et al.* (2006) and assume that increases in production between the early 2000s and 2009 are exclusively attributable to expansion of industrial systems, then we conclude that industrial systems based on commercial breeding lines provide 79% of global poultry meat, 73% of eggs, and 63% of global pork production (Hoffmann, 2011). The emergence of corporate animal breeding, in particular, allowed merging of genetic improvement, production, processing, and marketing activities and provided resources for aggressive international marketing of germplasm. Within the developed nations, the global breeds have made an important contribution to maintain low food costs, but many of the traditional livestock breeds in these areas have declined because they cannot compete with these global breeds.

Because of their tremendous production potential, these highly developed global breeds also caught the attention of government planners and officials in developing nations. Their global dissemination was seen as a magic bullet to improve animal productivity, analogous to the improvement in crop yields associated with the Green Revolution. Widespread importations of these breeds into developing nations had a negative impact on indigenous breeds, sometimes through outright breed replacement but more often through extensive and generally unregulated crossbreeding between imported and indigenous breeds (**Box 1**). In rural areas and subsistence production systems, the contributions of these highly productive global breeds were often disappointing. When denied the high levels of feeding, housing, and veterinary care under which they were developed and forced to grapple with unfamiliar levels of disease, low-nutrient feeds, and climatic stress, the highly productive global breeds and their crosses were often unable to forage, survive, and reproduce at acceptable levels and were ultimately less productive than the adapted indigenous types they were intended to replace.

Improved technology for transfer and use of genetic material contributed to the demise of many indigenous breeds. Use of frozen semen became commonplace in cattle in the latter half of the twentieth century and is now widespread for many livestock species. Use of frozen embryos is also widely practiced for cattle, sheep, and goats. These developments allowed efficient global exchange of genetic material and effectively ended the genetic isolation of many breeds, especially in Asia and Africa, where the concept of breed identity was less strong than in Europe and North America. Greater ease of transport of live animals also facilitated establishment of populations of foreign breeds and replacement or contamination of native breeds, especially for the smaller or more fecund species such as poultry and pigs. Thus, modern reproductive technologies today permit foreign breeds to have a much greater and more rapid impact, and in some cases to completely replace native breeds.

Urbanization and economic development have also opened important niches for use of global breeds in developing nations. Demand for animal products in these countries has soared in recent years and is expected to increase continuously, resulting in movement from subsistence production to market-driven production and changes in production systems. In particular, urbanization has concentrated animal production in periurban areas to allow convenient access to population centers and has led to opportunities for more intensive production. These changes have occurred despite concerns over economic disenfranchisement of farmers in rural areas, air and water pollution and waste disposal in periurban production, and the long-term sustainability of these production systems. Intensive animal production systems in periurban areas have also expanded the use of highly productive global breeds at the expense of adapted native types. The competitive advantage of these intensive periurban systems may decline however, once the full costs of waste disposal and environmental impacts are incorporated into production costs, and if costs of harvested feeds increase due to competition for land and crops for human food and biofuel, thereby providing justification for a return to more traditional mixed-farming systems and breeds.

Animal welfare concerns may also influence the use of livestock breeds. In Europe, concerns over issues of animal welfare in intensive production systems have prompted a return to less-intensive conditions and a corresponding need for animals that produce well under those conditions. Increased interest in consumption of locally produced foods, natural

Box 1 Crossbreeding: Threat or boon to animal genetic resources?

The production of crossbred animals by mating parents of different breeds is common in livestock breeding. Crossbred animals are often superior to their purebred parents and therefore more desirable for commercial farmers. The desirability of crossbred livestock arises from two sources: heterosis and complementarity.

Heterosis, also known as hybrid vigor, is normally manifested as an increase in fitness of crossbreds relative to their purebred parents. Thus, crossbred animals commonly are more fertile, more disease resistant, and better able to cope with environmental stresses than would be predicted from the average fitness of their purebred parents. As a result, productivity traits such as growth and milk or egg production are also usually improved in crossbred animals. In simple terms, purebred individuals commonly show modest levels of inbreeding as a result of the restricted matings required to genetically fix the defining characteristics of the breeds (e.g., colors and horn shape), and inbreeding commonly results in associated modest reductions in fitness. Crossing of purebred animals relieves accumulated inbreeding and provides an economically significant "kick" in performance. Significantly, however, effective use of hybrid vigor requires the crossing of unrelated parent breeds and is maximally expressed only in the first-generation cross. Thus, hybrid vigor can be maximally exploited and reliably captured only when parents of the original breeds are maintained as purebreds and mated in specific ways.

Complementarity is a characteristic of the production system which arises when different breeds play different and appropriate roles in crossbreeding systems. In low-input or extensive-production systems, adaptation of breeding females to the production environment is critical. If the environment is harsh, the productive capacity of indigenous breeds is often low, thereby allowing demands for nutrients and other inputs to be synchronized with their limited supply. In meat production, mating of females of indigenous breeds to males of more productive and heavily muscled breeds can increase the value of the offspring while maintaining high levels of adaptation in the breeding females. Additionally, the benefits of having a well-adapted mother and of hybrid vigor often permit the crossbred offspring to perform at acceptable levels and increase overall productivity. However, successful use of complementarity again requires maintenance of the adapted, indigenous breeds. Replacement of the original indigenous breeding females with crossbreds results in losses in both adaptation and hybrid vigor, often with serious negative effects on productivity.

Proper use of crossbreeding thus relies on breed diversity. Access to arrays of both adapted and highly productive breeds allows synchronization of the genetic characteristics of both breeding females and crossbred offspring to diverse production environments and markets. However, proper use of crossbreeding also requires that breeds be maintained in their proper roles in the system. In developing nations, in which control over matings is often limited or nonexistent, use of highly structured crossbreeding systems has proven difficult, and initial improvements in productivity in first-generation crosses, as well as the genetic integrity of the indigenous breeds, have been lost through indiscriminate mongrelization of adapted and unadapted breeds. Thus, crossbreeding has significant benefits when an appropriate management infrastructure exists but can have devastating effects on indigenous breeds when not adequately managed.

food products, and regionally or breed-based branded food products likewise has potential to create niche markets for traditional breeds. This phenomenon is most advanced in countries with relatively high disposable incomes, but premiums for locally produced poultry meat from traditional breeds are also becoming common in developing countries.

Conservation Planning: Coordinating the Use and Conservation of Livestock Genetic Resources

Animal breeders are often confronted by an apparent dilemma in the design of livestock breeding programs. On the one hand, maximizing livestock productivity relies on the identification and propagation of superior genetic types. Rates of genetic improvement are generally proportional to the intensity of selection. Thus, propagation of small numbers of elite parents yields greater rates of progress than retention of larger numbers of less-desirable individuals. Indeed, most of the accomplishments of modern animal breeding have resulted from improving the accuracy of genetic evaluation of prospective parents and increasing the number of offspring of these selected parents by use of artificial insemination (AI), embryo cloning and transfer, and other advanced reproductive technologies.

Successes of modern animal breeding programs have, in some cases, been spectacular, as exemplified by the emerging global breeds of poultry, swine, and dairy cattle. The global proliferation of industrial production systems has relied primarily on these species because of their greater efficiency of use of harvested feeds. In other species, such as the grazing ruminants, successes have been more modest, largely because these animals are generally produced under extensive conditions and required to interact closely with their environment to harvest forages, but still have been substantial. Despite these successes, demands for further increases in rates of animal production to supply an expanding human population continue to increase, and pressure to utilize the highly productive global breeds to meet these demands is tremendous.

Improvement-oriented animal breeders view global genetic diversity as a resource to be used in pursuit of improved animal productivity. The existing array of livestock breeds is recognized as potentially valuable but primarily as a consumable resource to be integrated into elite populations as appropriate and molded into ever-improving commercial populations. This viewpoint, described as the "utilizationist view" by the US National Academy of Sciences (NRC, 1993), emphasizes the potential value of the unique genes and gene combinations found in the various breeds as raw material for breed evolution and emphasizes the value in maximizing that evolution. In the utilizationist view, there may well be a need for a few dozen different pig breeds to meet the demands of different global production environments and markets, but a commitment to maintain all of the 656 currently recorded pig breeds (FAO, 2010b) would be seen as excessive and counterproductive.

In contrast to this position is a more "preservationist view" which tends to accord individual breeds a position similar to that of an endangered species and proposes similar management strategies to sustain them. The preservationist view emphasizes the unique history and presumed genetic distinctness of individual breeds and takes breed preservation *per se* as its goal. Blending of breeds to generate improved, adapted commercial populations is acknowledged to be necessary and desirable, but only against a backdrop of secure and relatively stable populations of the contributing breeds. Retention and continued use of existing breeds in their traditional

environments and production systems is promoted, even if public subsidies are required to ensure that use.

Synthesis and rectification of the utilizationist and preservationist views is badly needed but has yet to fully occur in the animal breeding community. The utilizationists have, to some extent, become victims of their successes, especially in poultry, dairy cattle, and pig production. Aggressive and widespread sampling and comparative evaluation of breeds in the mid-twentieth century led to the establishment of today's global breeds and to the replacement of many of the mid level, often dual-purpose, breeds that were widely represented in commercial production a few decades ago. Although large numbers of locally adapted, relatively low-productive breeds of poultry and swine still exist globally, the elite global breeds have become so differentiated from these stocks that there is now thought to be minimal opportunity for these breeds to contribute genetic material to the elite global breeds through traditional breeding methods. Thus, industrial stocks of chickens and turkeys have experienced no outside contributions from other breeds in more than 35 years. A similar situation exists for modern pig populations, and, in the temperate regions, the Holstein dairy cow is rapidly becoming the preeminent, and in many cases exclusive, dairy breed. Increasingly, the elite global breeds are viewed as dependent on existing reservoirs of within-breed genetic diversity for future adaptation and improvement. In contrast, the preservationist position has been made untenable by the events of global economic and social homogenization. In an increasingly cosmopolitan, interdependent, and rapidly changing world, continued use of nearly 7000 local livestock breeds (FAO, 2010b) will not occur and is not necessary to maintain adequate reservoirs of genetic diversity. However, a responsible fallback position which can ensure the identification, retention, and ready access to the core genetic diversity of each domestic species is badly needed.

A reasonable synthesis and integration of conservation and genetic improvement activities have been achieved in the plant breeding community, as described elsewhere in this encyclopedia. Facilities for long-term seed storage exist and have both a broad mandate and the budgetary support to acquire and preserve samples representing the full range of diversity within individual species of crop plants and their wild relatives. The importance of these programs is acknowledged by both public agencies and private breeding companies. Even though there are fundamental differences between plants and animals in the population structure of breeding materials, technologies that can be used for genetic resource conservation, and use of wild relatives in commercial breeding (**Box 2**), a more assertive and proactive approach to animal germplasm conservation is still needed.

Management Strategies for Sustainable Use and Conservation of Livestock Genetic Resources

Leadership in global management of conservation of livestock genetic resources has come from the FAO, beginning with many *Animal Production and Health Papers* (http://www.fao.org/ag/againfo/resources/en/pubs_gen.html), continuing with the *Global Strategy for the Management of Farm Animal Genetic Resources* (FAO, 1999), and recently culminating in the *State of the World's Animal Genetic Resources* (FAO, 2007a) and the *Global Plan of Action for Animal Genetic Resources* (FAO, 2007b). FAO also leads the program of the Convention on Biological Diversity on agricultural biodiversity (CBD, 2006, 2008a, b). The Domestic Animal Diversity Information System (DAD-IS; http:www.fao.org/dad-is) is a comprehensive multilingual web-based clearinghouse mechanism for animal genetic resources and is being further developed by FAO to allow modeling of the impacts of climate change on breeds of livestock. The 2007 *State of the World's Animal Genetic Resources for Food and Agriculture* was the first comprehensive global assessment of the roles, values, status, and trends of animal genetic resources; the capacity at both country and international levels to manage these resources; and the state of the art in management of animal

Box 2 Use of wild relatives in livestock breeding

In contrast to plants, use of wild relatives as a source of genetic material for industry breeding programs in livestock is limited. In most cases, domestic animals represent distinctive species and wild progenitors of these species are no longer found. However, several recent developments suggest that wild relatives may be an useful source of genetic material and that the pool of useful wild relatives may be broader than originally thought.

In bovines, hybrids of domestic cattle with gaur (*Bos gaurus*), banteng (*B. javanicus*), and yak (*B. mutus*) are common in Asia and make important contributions to livestock production. In North America, hybrids of cattle and bison (*Bison bison*) were evaluated in Canada in the 1960s but with limited success because of near-complete infertility of hybrid males. Crosses of bison and cattle in the United States, however, yielded enough fertile hybrid males to allow establishment of herds of animals, known as beefalo, possessing a mixture of genes from cattle and bison. These animals were claimed to possess some desirable adaptational and meat quality traits but are today primarily a novelty.

In western China, genes from the wild ibex (*Capra ibex*) were introduced into domestic cashmere goats. The yield of cashmere from ibex is very low, but the cashmere fibers are very fine, averaging approximately 12 μm in diameter. Typical cashmere goat fibers are 16 or 17 μm in diameter, and the value of the fiber is related to its fineness. Introgression of genetic material from ibex, followed by selection for yield and fineness of cashmere, led to development of animals with 12.5–25% ibex genes that possess increased fineness of the cashmere with near-normal yield. Similar crossing programs have been proposed using the wild vicuña to improve fiber quality in domestic alpaca.

In the Republic of South Africa, there is a need for hardy and well-adapted livestock for use in communal and other low-input production systems. In pigs, crosses between an endangered domestic pig breed, the Kolbroek, and the wild bush pig (*Potamochoerus porcus*) led to production of fertile offspring with improved hardiness, foraging ability, and capacity to digest fiber. This inter-generic cross also indicates the potential to use more distant crosses as a source of genetic material.

Wild relatives of domestic livestock have also contributed to development of genetic maps. Crosses of domestic pig with wild boar (*Sus scrofa*) were used to develop gene maps for pig, and crosses with red junglefowl (*Gallus gallus*) were used in mapping the chicken genome.

In plant breeding, wild relatives of domestic crop plants have long been recognized as a source of genes for disease and pest resistance. Studies involving wild relatives of tomato and rice suggest similar opportunities to improve production traits. Use of gene maps to identify useful genetic material and of improved methods to introduce this material into domestic stocks using either biotechnology or conventional breeding methods now exist for plants and will be developed for animals.

genetic resources (Hoffmann *et al.*, 2011). The report was the result of a country-driven process based on169 country reports (ftp://ftp.fao.org/docrep/fao/010/a1250e/annexes/CountryReports/CountryReports.pdf), nine reports from international organizations, scientific literature and FAO's DAD-IS, and statistical databases. The State of the World report was launched at an international conference at Interlaken, Switzerland, in September 2007. At this conference, governments agreed on the *Global Plan of Action for Animal Genetic Resources*, the first-ever international framework for the promotion of the wise management of animal genetic resources for food and agriculture, and the Interlaken Declaration. Both were later endorsed by the 191 member nations of the FAO.

The Global Plan of Action (**Box 3**) contains 23 Strategic Priorities for Action, clustered into four priority areas covering all areas of management of farm animal genetic resources:

1. Characterization, inventory, and monitoring of trends and associated risks;
2. Sustainable use and development;
3. Conservation; and
4. Policies, institutions, and capacity-building.

The main responsibility for implementation of the Global Plan rests with national governments. However, governments (and especially governments of developing countries) are supported by the international community, including FAO, bilateral and multilateral donors, and international research and development institutions.

Box 3 FAO's Role in Implementing the Global Plan of Action

FAO, in partnership with governments, international organizations, research organizations, and nongovernmental organizations, facilitates global and regional collaboration and networks; supports the convening of intergovernmental and technical meetings; maintains and further develops the Domestic Animal Diversity Information System (DAD-IS); develops communication products; provides technical guidelines and assistance; coordinates capacity-building and training programs; and promotes the transfer of technologies related to sustainable use, development, and conservation of AnGR. Guidelines to assist countries in the implementation of the Global Plan of Action are being developed, including:

- preparation of national strategies and action plans for animal genetic resources (FAO, 2009a); and
- breeding strategies for sustainable management of animal genetic resources (FAO, 2010a).

Further guidelines are under development and will complete the set and provide guidance to policy makers as well as technicians comprehensively in all four priority areas. These include:

- Development of the institutional framework for the management of animal genetic resources (http://www.fao.org/docrep/meeting/021/am137e.pdf).
- Surveying and monitoring (http://www.fao.org/docrep/meeting/021/am133e.pdf).
- Phenotypic characterization (http://www.fao.org/docrep/meeting/021/am134e.pdf).
- Molecular genetic characterization (http://www.fao.org/docrep/meeting/021/am135e.pdf).
- Cryoconservation (http://www.fao.org/docrep/meeting/021/am136e.pdf).

Guidelines on animal identification and recording and *in situ* conservation are being developed, and individual countries are developing national strategies and action plans (Hoffmann and Scherf, 2010). In celebration of the International Year of Biodiversity 2010, FAO produced a special issue of its journal *Animal Genetic Resources* (http://journals.cambridge.org/action/displayJournal?jid=AGR) focusing on the four priority areas of the Global Plan and invited national coordinators for the Management of Animal Genetic Resources (http://dad.fao.org/cgi-bin/EfabisWeb.cgisid=-1,contacts) and international organizations to prepare posters describing implementation of the Global Plan (http://www.fao.org/ag/againfo/programmes/en/genetics/natcord.html). Most of the FAO documents referenced in this article can be accessed via the DAD-IS library (http://dad.fao.org/) or the AGA homepage at http://www.fao.org/ag/againfo/resources/en/pubs_gen.html.

Inventory, Description, Characterization, and Assessment of Degree of Endangerment

Understanding the status and trends of domestic animal genetic diversity is prerequisite to its successful management, and the inventory, description, and comparative characterization of livestock breeds is a key management activity. Part 1 of *The State of the World's Animal Genetic Resources* (FAO, 2007a) describes the state of livestock biodiversity with annexes listing breeds categorized in different risk statuses (see http://www.fao.org/docrep/010/a1250e/a1250e00.htm). Two additional status and trend reports have also been prepared and published (FAO, 2009b, 2010b).

The 2010 status and trend report catalogs 8054 breeds, including 5833 mammalian and 2221 avian breeds representing 37 species used for food and agriculture and reported by 182 countries (**Figure 1**). Of these, 631 breeds are classified as extinct and are not included in **Figure 1**. More than 1000 breeds are represented in more than one country, and the total number of country-level breed populations is 10,507 for mammals and 3414 for birds. However, data on population size and status are available for only 54% of mammalian and 47% of avian breeds. The remaining breeds have only cursory information on breed characteristics and distribution. Reporting of breed information is likewise not consistent throughout the world. European breeds dominate the list, with 4127 mammalian and 1609 avian entries, primarily because Europe was the cradle for the concept of breed formation and because of a long history of concern over genetic diversity. African breeds, in contrast, are almost certainly underreported, with only 1669 mammalian and 421 avian breed populations listed.

To be useful, inventory information must be augmented with some objective assessment of degree of endangerment of the populations involved. The current FAO system for determining breeds at risk is shown in **Table 1** (ftp://ftp.fao.org/docrep/fao/010/a1250e/a1250e.pdf). Based on this classification, approximately 18% of mammalian and 31% of avian breeds with population data are considered to be at risk (**Figure 2**). More detailed risk-status assessments are shown by species for mammalian and avian breeds in **Figures 3** and **4**, respectively. Furthermore, only 19% of mammalian breeds and 21% of avian breeds that are at risk have active programs in place to promote their conservation. Because many of these endangered breeds are found in harsh production

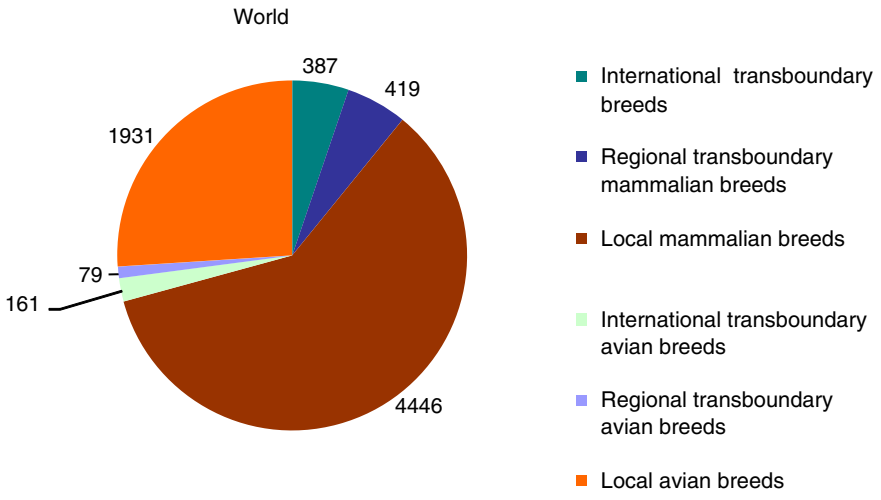


Figure 1 Numbers of local and transboundary breeds at global level (excluding extinct breeds).

Table 1 FAO risk status classification

Category	*Criteria*
Not at risk	Total number of breeding females is greater than 1000 and total number of breeding males is greater than 20; or population size approaches 1000, the percentage of females being bred to males of the same breed is near 100%, and the overall population size is increasing.
Endangered	Total number of breeding females is greater than 100 and less than or equal to 1000 or the total number of breeding males is less than or equal to 20 and greater than five; or the overall population size is greater than 80 and less than 100 and increasing and the percentage of females being bred to males of the same breed is above 80%; or the overall population size is greater than 1000 and less than or equal to 1200 and decreasing and the percentage of females being bred to males of the same breed is below 80%, and it is not assigned to any of above categories.
Endangered-maintained	Endangered populations for which active conservation programs are in place or populations are maintained by commercial companies or research institutions.
Critical	Total number of breeding females is less than or equal to 100 or the total number of breeding males is less than or equal to five; or the overall population size is less than or equal to 120 and decreasing and the percentage of females being bred to males of the same breed is below 80%, and it is not classified as extinct.
Critical-maintained	Critical populations for which active conservation programs are in place or populations are maintained by commercial companies or research institutions.
Extinct	No breeding males or breeding females remaining. Nevertheless, genetic material might have been cryoconserved which would allow recreation of the breed. In reality, extinction may be realized well before the loss of the last animal or genetic material.
Unknown	Information on population size is not available.
Breed at risk	A breed that has been classified as critical, critical-maintained, endangered, or endangered-maintained.

Source: ?ftp://ftp.fao.org/docrep/fao/010/a1250e/a1250e.pdf

environments with minimal external inputs and management, they can reasonably be anticipated to possess unique genetic adaptations and disease-resistance characteristics. However, in common with many endangered species, they are at risk of disappearing before they can be adequately evaluated and characterized.

Conservation Strategies

Conservation of livestock genetic resources will almost certainly involve both *in situ* conservation of breeding herds in agricultural production systems and *ex situ* conservation of animals, gametes, embryos, and somatic cells in reserves and cryogenic repositories. Issues of current utility, degree of endangerment, and cost of conservation will determine the method(s) of choice.

In Situ Strategies

Strategies that are based on effective use of livestock genetic diversity in commercial and/or subsistence production systems are most likely to provide the security necessary to allow breeds to remain viable and maintain the population sizes necessary to ensure continued improvement. Thus, a key activity in livestock genetic resource conservation is the comprehensive appraisal of the productive merit of the various breeds under current and anticipated future production conditions.

Most of the world's domestic animals are found in developing countries, and nearly two billion people obtain at least a part of their livelihood from livestock. Rapid changes in production conditions have occurred in some of these countries, but the number of animals used in subsistence production systems remains very large. Levels of veterinary care,

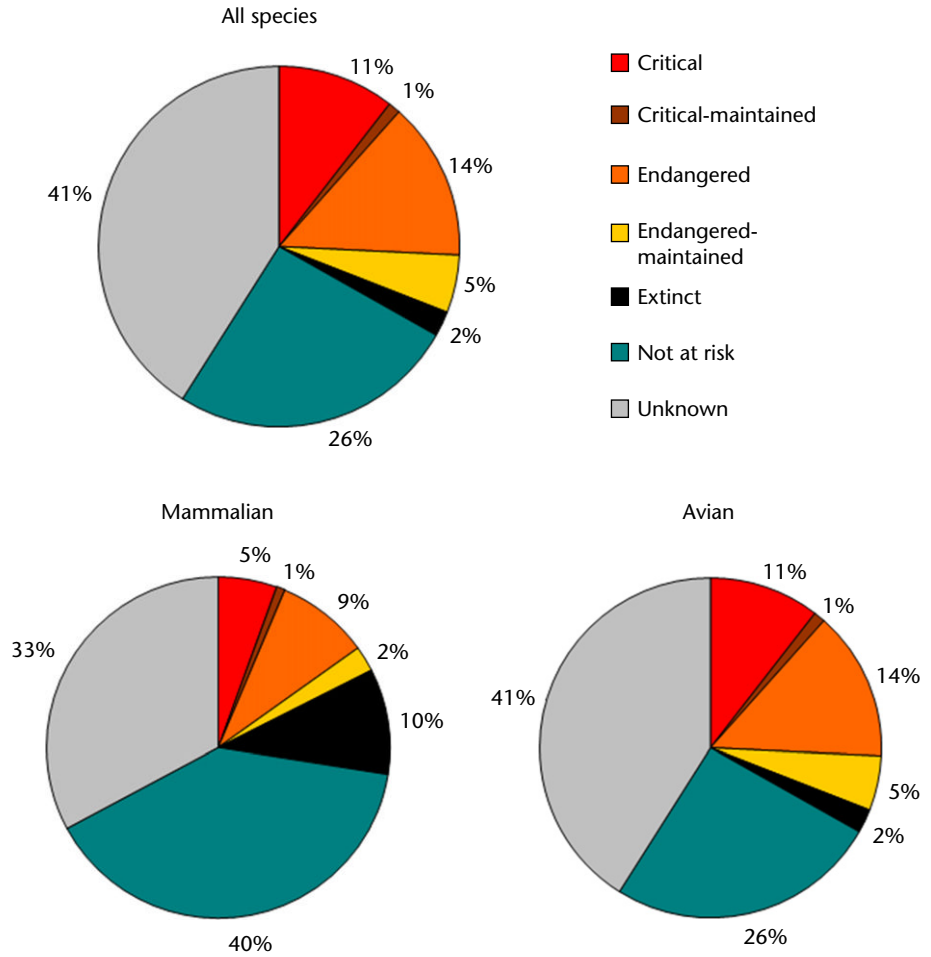
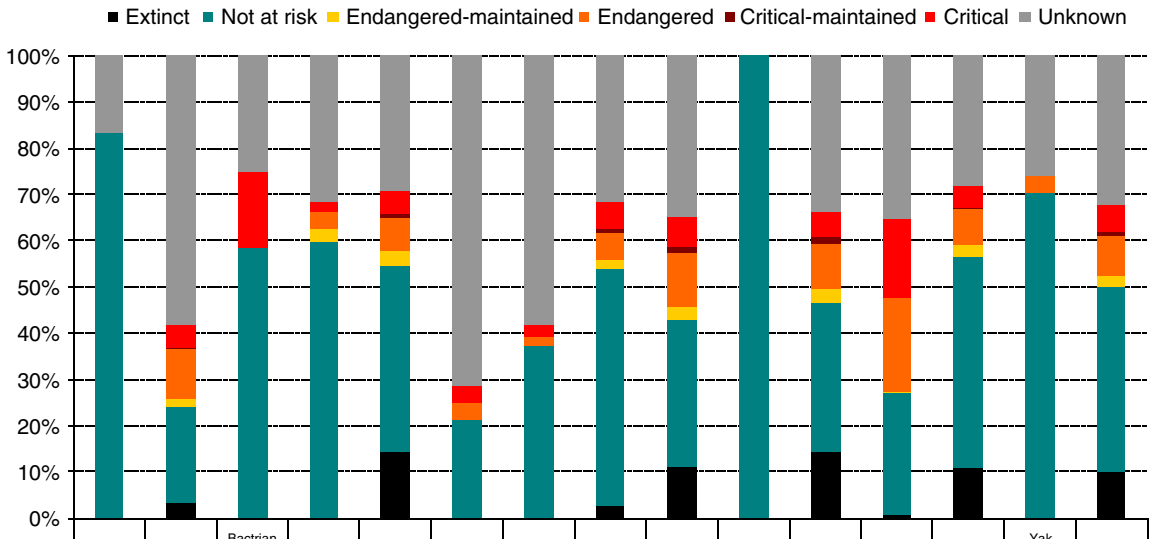


Figure 2 Proportion of the world's breeds at risk by status risk status category.

supplemental feeding, and shelter from climatic stresses are often limited, and adaptational characteristics of indigenous livestock remain important to many farmers. Studies throughout the world have confirmed that if environmental conditions are harsh, imported temperate breeds of high-production potential often experience increased levels of mortality and morbidity and reduced reproductive success, and therefore may be less productive per unit of time or per unit of inputs than indigenous breeds. Advisors often recommend changes in the production environment, even though farmers commonly do not have the financial resources or market incentives to do so. A more responsible strategy, and one that is more consistent with conservation of livestock diversity, is to focus on retention and improvement of local breeds (Madalena, 2008). The implementation of modern livestock improvement procedures in indigenous breeds can allow them to be improved in parallel with gradual improvements in production conditions while retaining important adaptational characteristics.

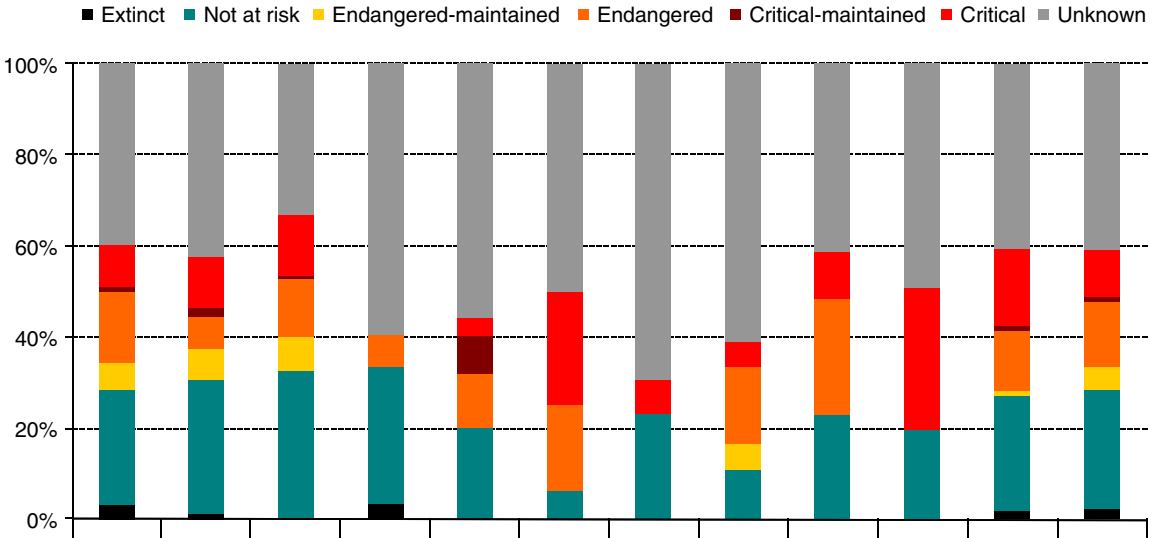
An example of such an approach is found in India, where Operation Flood was begun more than 40 years ago to enhance milk production and marketing opportunities for smallholders in western India. The goal was to improve the nutrition of the people of India by increasing supplies of buffalo and cattle milk. Certainly, Operation Flood and its successor projects have been a quantitative success. India is now the world's number one producer of milk. Much of this success was achieved without recourse to establishment of high-volume Western dairy production systems or a rapid infusion of poorly adapted foreign germplasm. Instead, emphasis was placed on establishing marketing conditions that would reward efficient producers and provide incentives for improved production practices. Only in the past decade has the identification of superior germplasm within indigenous breeds of cattle and buffalo and strategic use of imported breeds become a part of this program. In Latin America, Brazil has likewise been a leader in the utilization of adapted temperate and more recently introduced tropical genetic resources to create highly productive, optimally adapted populations of dairy cattle (Madalena *et al.*, 1990).

In other cases, however, the pace of socioeconomic change has been so rapid that indigenous breeds in the developing nations cannot meet the challenges. Under these conditions, pressures on indigenous breeds can become extreme very quickly, and careful analysis is required to identify appropriate genetic resource development activities. Developments in pig production in China provide an example. China has the world's most diverse array of pig breeds, with 142 reported



	Alpaca	Ass	Bactrian Camel	Buffalo	Cattle	Deer	Dromedary	Goat	Horse	Llama	Pig	Rabbit	Sheep	Yak (domestic)	Total
Unknown	1	99	3	44	393	20	52	205	282	0	258	92	409	7	1865
Critical	0	8	2	3	66	1	2	36	52	0	42	44	67	0	323
Critical-maintained	0	1	0	0	13	0	0	6	11	0	12	0	4	0	47
Endangered	0	18	0	5	95	1	2	39	96	0	74	53	116	1	500
Endangered-maintained	0	3	0	4	46	0	0	10	21	0	23	1	32	0	140
Not at risk	5	35	7	83	541	6	33	331	257	4	247	68	666	19	2302
Extinct	0	6	0	0	194	0	0	18	91	0	110	2	160	0	581

Figure 3 Risk status of the world's mammalian breeds by species in October 2010 in terms of absolute numbers (table) and percentages (chart).



	Chicken	Duck (domestic)	Goose (domestic)	Guinea fowl	Muscovy duck	Ostrich	Partridge	Pheasant	Pigeon	Quail	Turkey	Total
Unknown	563	104	62	32	14	8	9	11	29	25	43	900
Critical	130	27	25	0	1	4	1	1	7	16	18	230
Critical-maintained	16	4	1	0	2	0	0	0	0	0	1	24
Endangered	226	18	24	4	3	3	0	3	18	0	14	313
Endangered-maintained	83	16	14	0	0	0	0	1	0	0	1	115
Not at risk	357	72	61	16	5	1	3	2	16	10	27	570
Extinct	43	3	0	2	0	0	0	0	0	0	2	50

Figure 4 Risk status of the world's avian breeds by species in October 2010 in terms of absolute numbers (table) and percentages (chart).

breeds (including 10 now thought to be extinct). Almost all evolved in situations in which grain was primarily used for human food. The pigs therefore became adapted to the use of fibrous by-product feeds from grain and vegetable production. They also were relatively slow growing (to synchronize their nutrient requirements to their limited food supply), quite fat (because animal fat was an important energy source for peasant farmers), and, in some cases, highly prolific. In the 1980s, government policies in China were changed to provide greater access to grain for animal feed, and rising incomes since that time have further supported that decision. Under these conditions, the indigenous pig breeds were markedly inferior to the faster growing, leaner, and more feed-efficient Western breeds. Government-sponsored importations and entry of multinational pig breeding companies into the market resulted in widespread crossbreeding of indigenous pigs with Western breeds. An extensive network of government AI centers for pigs facilitated the use of imported breeds, and the development of more sophisticated urban markets heightened preferences for leaner pork. By the early 1990s, many of the AI stations no longer provided boars of the local breeds.

The events in Chinese pig breeding were fully rational responses to changed government policy and to changes in the economy and the nature of society. Both farmers and consumers benefited from the changes. However, the result has been to endanger a whole array of local breeds, many of which possess globally unique characteristics.

In developed nations, focus on a declining number of elite breeds in relatively intensive production systems has likewise endangered many traditional breeds. The situation is particularly acute in Europe, and greater cognizance of the problem exists there. FAO (2010b) lists 3537 breeds of European livestock. Of these, 505 are considered extinct and no information on population size is available for an additional 758. Of the remaining 2274 breeds, fewer than two-thirds (1151 breeds) were considered secure or listed as "maintained."

In response to this contraction in livestock genetic diversity, many nations of Western Europe have established programs to maintain endangered breeds. Although often using a combination of *in situ* and *ex situ* techniques, high priority is placed on *in situ* conservation. The importance of livestock breeds as a cultural and historical resource reflecting the heritage of the nation or as a component of unique and perhaps themselves endangered agroecosystems is particularly recognized in Europe. The contribution of livestock breeds to public welfare is sometimes referred to as "landscape value," and public resources are being directed toward breed conservation activities. These programs may involve maintenance of endangered breeds on public facilities or direct payments to farmers who maintain endangered breeds.

Throughout the developed nations, "grassroots" organizations play a significant role in *in vivo* conservation of endangered livestock breeds. These nongovernmental organizations (NGOs) have for many years provided leadership for conservation of livestock diversity. Examples include the Rare Breeds Survival Trust in the UK (http://www.rbst.org.uk), the American Livestock Breeds Conservancy in the United States (http://www.albc-usa.org), the Safeguard for Agricultural Varieties in Europe (http://www.save-foundation.net), and the Canadian Foundation for the Conservation of Farm Animal Genetic Resources (http://www.cfagrf.com). The success of these organizations attests to the potential to involve committed private individuals in breed conservation. They have been most successful in wealthy countries in which significant private resources can be exploited, but the identification of private patrons to aid in breed conservation can also occur in developing nations. The success of these organizations relies on members who are committed to, and educated about, the technical aspects of genetic conservation. Otherwise, well-meaning but misdirected breeding policies may result in losses of genetic diversity. Public–private partnerships to provide members of these NGOs with technical expertise can be particularly beneficial.

Ex Situ Strategies

Ex situ conservation usually involves cryopreservation of gametes, embryos, tissues, or somatic cells or the storage of DNA, but may also involve live animals kept in farm parks, research farms, or other noncommercial settings. In the developed nations, farm parks have increased in popularity and often contribute to meaningful maintenance of biodiversity through their association with responsible grassroots organizations. Live-animal, *ex situ* conservation can also occur at public-funded facilities, generally under conditions that at least approximate those found in commercial agriculture but which often do not replicate the particular conditions under which the breeds were evolved and traditionally used. In these situations, selection for unique adaptational characteristics is relaxed, but with proper breeding management and adequate population size, key genetic characteristics of the breed can be retained for many generations. Costs of *ex situ* live-animal conservation are high, however, involving feeding and daily care of breeding animals, and disease outbreaks or localized disasters may threaten endemic breeds (Carson *et al.*, 2009). The risks associated with live-animal *ex situ* programs are therefore also high, especially in developing nations in which sustained funding for long-term conservation programs may not be available and risks of social and political upheaval are greatest.

Use of cryopreserved gametes, embryos, and tissues is a more common form of *ex situ* conservation. In farm animals, sperm cells can be successfully frozen and stored for future use in all species, although success rates from use of frozen sperm vary considerably among species. Techniques for collection, cryopreservation, and subsequent use of sperm cells are relatively well developed; a single collection provides a relatively large number of gametes, and multiple samples are relatively easy to obtain. For these reasons, cryopreserved sperm cells are the most common material used for *ex situ* conservation of endangered breeds. Cryopreserved sperm are ideally suited to support *in situ* conservation activities. Storage of sperm from a wide sample of males of a breed provides future access to the genetic material of these representative foundation animals. Losses of genetic diversity in living populations can thus, if necessary, be restored by use of sperm from males of past generations.

However, cryopreserved sperm cells are not particularly efficient for regeneration of a breed that has become extinct. Sperm cells contain a sample of only half of the animal's DNA; therefore, restoration of an extinct breed from cryopreserved sperm requires a "grading-up" process in which sperm

is used on females of a different breed over several generations to eventually create animals that have a majority of their nuclear genes from the cryopreserved breed. The efficiency of use of sperm cells to restore a breed depends on the generation time and the fecundity of the species. For example, restoration of 93.75% of the nuclear genes (i.e., four generations of upgrading) can be accomplished in approximately 3.5 years in pigs, and use of 100 sows initially could conservatively result in production of 1000 breeding females by generation four. In contrast, in cattle a minimum of 10 years would be required using conventional breeding techniques to produce even a small number of animals possessing 93.75% of the nuclear genes of the preserved breed. Although nuclear genes can be adequately preserved using frozen sperm, cytoplasmic DNA found in animal mitochondria are contributed only via the ovum and can thus be preserved only by storage of cryopreserved ova or embryos. Cryopreservation and subsequent *in vitro* fertilization of ova are not yet practical for domestic species. Embryo cryopreservation, however, is practical in cattle, sheep, and goat and permits conservation of the full genome, both nuclear and cytoplasmic. However, collection of embryos for cryopreservation is more difficult and expensive than collection of sperm cells and yields of embryos are much lower, often on the order of only 2–8 embryos per mating. Also, cryopreservation of embryos is difficult in pigs and not practiced for poultry.

Recommendations for *ex situ* conservation programs thus generally focus on extensive use of frozen sperm cells. However, access to modest numbers of cryopreserved embryos (in species for which this is possible) and/or small populations of breeding females preserved under either *in situ* or *ex situ* conditions is recommended to provide a source of cytoplasmic genes and allow efficient regeneration of the breed. Cryopreserved embryos produced by mating 25 pairs of unrelated parents and bolstered by cryopreserved sperm from 25 unrelated males can effectively capture the genetic diversity of a breed for long-term conservation and future use.

Cryopreservation of somatic cells is also becoming common in *ex situ* conservation programs as an economically efficient approach to back up *in situ* populations and insure animal genetic resources against extinction. Several companies offer cloning services on a commercial basis, but with a focus on unique, high-value individuals rather than populations. Use of cloning in breed conservation thus relies on the assumption that continuing research will eventually result in substantial increase in efficiency of cloning new individuals from cryopreserved adult cells. Use of cloning to support conservation of rare breeds will be discussed in the section Advances in Cloning.

Impacts of New Technologies

Advances in Cryopreservation

Technological advances in cryopreservation continue to enhance our ability to preserve biodiversity. Key research areas in cryopreservation include:

- Improved cryopreservation of pig embryos to allow capture of the full genome including mitochondrial DNA. This is especially important in Asia, where the number of endangered pig breeds is large.
- Improved cryopreservation of poultry embryos and semen. In chickens, success rates from use of cryopreserved semen remain low. Better success rates would prompt greater use of frozen semen in the poultry industry, which would then facilitate storage and access to frozen gametes for conservation purposes. Cryopreservation of fertilized eggs will likely remain difficult, but extraction and cryopreservation of the germinal disk of cells from the fertilized egg separate from the egg itself appears possible. Recent studies have also successfully cryopreserved both male and female gonadal tissue in poultry, with subsequent successful transplantation into live recipients.
- Improved understanding of why semen from some males and embryos from some breeds do not survive freezing well. Large individual differences exist in success rates for cryopreservation of semen and embryos. In commercial cattle breeding, there is considerable selection for "freezability" of semen, and because the populations are large this selection has little impact on the genetic diversity within the breed. However, endangered breeds are commonly represented by only a few males, and if semen from several of these cannot be satisfactorily cryopreserved, a significant segment of the biodiversity of the breed may be lost. Success rates for frozen embryos are also often lower in poorly characterized endangered breeds, and these breeds sometimes respond poorly to the hormonal regimen required for embryo collection. More reliable and robust techniques for semen and embryo recovery and cryopreservation are thus needed.
- Reproductive biotechnologies and cryobiologies for minor livestock species. Most developments in reproductive biotechnologies have targeted the most common and commercially important livestock species, and *Bos taurus* cattle in particular. For many other species, these technologies remain far behind, if investigated at all.

Advances in Molecular Characterization and Gene Mapping

The emergence of molecular tools for characterization of animal DNA has revolutionized the study of livestock genetic diversity. Techniques developed for sequencing of the human genome in 2003 were rapidly applied to genomes of a variety of livestock species, most notably cattle, pig, and chicken, but also horse, dog, sheep, turkey, etc., and with increasing emphasis on regionally important species such as water buffalo, camel, and alpaca (http://www.ncbi.nlm.nih.gov).

Molecular analyses have increasingly contributed to understanding patterns of domestication of livestock species and confirmed the existence of multiple centers of domestication for several species (Groeneveld *et al.*, 2010). Objective methods to describe genetic relationships among livestock breeds now provide information for setting conservation priorities. The establishment of effective conservation programs for the thousands of livestock breeds shown in **Figures 3** and **4** is neither possible nor necessary to protect a high proportion of the diversity present within the species. However, identification of high-priority candidates for conservation requires objective

information on their genetic uniqueness relative to other breeds. Thus in 1993, FAO developed recommended sets of microsatellite markers that are to be used to assess genetic relationships among breeds within each of the major livestock species (FAO, 1998). These markers have been widely used to clarify relationships among breeds within each of the domestic species (Baumung *et al.*, 2004). More sensitive molecular tools, such as high-density SNP (single-nucleotide polymorphism) arrays and increasingly powerful software packages now provide a range of options for setting conservation priorities (Boettcher *et al.*, 2010) and predicting breed composition for individuals that lack pedigree documentation of ancestry. However, to date, the proportion of the global livestock breeds that has been characterized at a molecular level remains small. Genetic distances derived from microsatellite markers have been used for reconstruction of phylogenetic trees but are rarely used for conservation decisions.

In common with the primary immediate use of genome sequence information in humans, rapid genetic diagnosis of simply inherited genetic defects, diseases, and other conditions in livestock is now relatively straightforward for species with published genome sequences and high-density SNP arrays. Resulting DNA-based diagnostic tests provide necessary information to identify asymptomatic carriers of the defect and manage breeding decisions. In the past, discoveries of serious genetic defects in popular breeding lines in livestock often resulted in extended periods of confusion and recriminations, required extensive and expensive progeny testing of potentially affected sires, and resulted in unnecessary losses of elite breeding stock. However, with today's improved diagnostic capabilities, the identification and management of such defects becomes straightforward and creates minimal disruptions in breeding programs. A recent example is discovery of the causative mutation for the lavender foal defect in Arabian horses (Brooks *et al.*, 2010).

Most of the economically important characteristics of farm animals are under polygenic control, involving several to many generally as-yet-unidentified genes scattered across the genome. Simply inherited mutations with large effects on phenotype have been identified for a few characteristics in livestock. Notable among these are mutations in several genes that promote increases in frequency of twin and triplet births in sheep (Davis, 2005) and a number of point mutations in the myostatin gene that result in increased muscle mass in several different species (Rodgers and Garikipati, 2008). Most of these mutations are associated with failures of developmental regulation and would have been deleterious in nature where litters of three or more lambs or the overproduction of muscle tissue or milk is not adaptive, but may be advantageous for intensively managed domestic species. However, most of the genetic variation used in practical breeding programs remains largely uncharacterized, and is often portrayed as consisting of some number of "quantitative trait loci,"(QTLs), influencing production traits in as-yet-unknown ways. The use of molecular techniques to access these QTLs and to develop diagnostic tests that can identify animals with favorable sets of QTLs for specific traits is a central area of current study in the breeding of animals. A variety of techniques exist for detection and localization of QTLs. All are based on detection of associations between production levels and various sorts of genomic DNA markers such as microsatellite or, increasingly, SNP variants. The Animal Quantitative Trait Locus database (http://www.genome.iastate.edu/cgi-bin/QTLdb/index) lists 13,931 reported QTLs for 1369 different characteristics from 739 publications in cattle, pigs, sheep, and chickens. However, relatively few of these reported QTLs have been successfully validated and shown to have unambiguous effects in other populations and breeds or successfully fine-mapped to identify causal mutations, and only a few have been used in marker-assisted breeding programs.

Current focus for use of molecular information in animal breeding applications involves high-density SNP arrays that can be used to probe the genome for 50,000–750,000 individual genetic markers and detect associations between these markers and accurate and objective measures of animal performance (Hayes and Goddard, 2010). The ability of these arrays to detect QTLs of large effect and facilitate identification of favorable mutations or other genetic modifications remains an area of interest. However, current research in livestock is often based on the assumption that relatively small individual effects of relatively large numbers of markers can be summed across the genome to provide an informative picture of the average genetic merit of potential breeding animals for various economically important characteristics, even if individual gene effects and mechanisms of action remain, for the moment, unknown. Once the necessary associations have been documented, these genome-wide association studies allow generations of molecular estimates of genetic merit in descendants of the original research population. This process has been proposed to allow direct prediction of genetic merit from genomic, rather than performance-based, information. However, the large numbers of phenotypic measurements needed to establish these associations are often not available, especially in small breed populations or breeds kept in extensive systems. The likely long-term impact of these research efforts for practical animal breeding practices and the conservation of animal genetic resources thus remain controversial, especially in developing countries (Hayes *et al.*, 2009a, b; Hoffmann, 2010; Marshall *et al.*, 2011).

Advances in Cloning

Cloning of farm animals from adult cells, if perfected and commercialized, could have far-reaching effects on both use of biodiversity and the efficiency of its conservation. Widespread use of genetically identical cloned individuals in commercial production systems would reduce the biodiversity present within those systems and increase the vulnerability of the animals to specific diseases or other environmental stressors. However, the ability to efficiently produce clonal offspring from adult animals whose performance has already been characterized has potential to facilitate synchronization of genetic resources to specific production conditions. In this scenario, breeding animals that have demonstrated exceptional adaptation and productivity in a particular environment could be clonally reproduced for use in that specific environment.

The ability to produce clonal offspring from adult cells would have tremendous implications for conservation of

endangered genetic resources. Large numbers of cells could be harvested from individual animals using minimally invasive biopsy techniques and cryopreserved for future use. Production of clonal offspring from these cryopreserved cells would recreate the full nuclear genome of the preserved animals, though the mitochondrial genome would presumably be lost during the process of nuclear transplantation to a donor cell. Furthermore, because sampling of somatic cells for cloning is relatively easy, large numbers of founder animals could be sampled, increasing the biodiversity present within the sample. Complications involved in harvesting and conservation of gametes and embryos would be circumvented by somatic cloning and efficiency of storage of animal genetic material would approach that achieved in plants by storage of seeds.

Techniques for clonal propagation of adult animals remain less efficient than use of frozen semen or embryos, but early results hold great promise. Dolly, the sheep produced by cloning of adult mammary cells at the Roslin Institute in Scotland, represented a watershed event in this technology. Many other livestock species have now been successfully cloned from adult cells, several companies offer animal cloning on a commercial basis, and cloned pigs can be obtained for a few thousand US dollars. In New Zealand, one of the last individuals of an endangered Enderby Island cow was cloned to attempt to save the breed (RBCSNZ, 2002).

Prospects and Threats

An assessment of likely trends in global livestock genetic diversity must recognize that population growth, economic development, and urbanization will require significant increases in animal productivity worldwide. Increasing demand for animal-based foods, as well as increasing resource constraints involving use of farm land for biofuel production, environmental degradation of grazing lands, effects of climate change, and strategies for climate change mitigation, may also have an impact on breed and species utilization. In the short to intermediate term, the result will likely be an even greater emphasis on monogastric livestock species and on those breeds that are the most efficient converters of feed into meat, milk, and eggs. Long-term impacts are, however, difficult to predict.

The normal processes of breed evolution have accelerated during the past 50 years. Breeds are disappearing at more rapid rates than in the past, and the rate of new breed formation has not correspondingly increased. Many indigenous breeds are ill-suited to contribute in meeting the challenges of sustainable intensification of livestock production in a global economy in a timely manner, and these breeds will likely become endangered or restricted to subsistence production systems, especially in ecologically marginal areas (Hoffmann, 2011). The contribution of highly productive, highly selected global breeds to world food production will continue to expand, placing further pressure on currently less-productive indigenous breeds.

Climate change is an additional challenge for the livestock sector. Highly productive global breeds managed in controlled production environments may be effectively buffered from direct physiological effects of higher temperatures and changing precipitation patterns, but only if costs for feed, energy, and water remain low. Also, the consequences of climate change will likely include a variety of indirect effects that may be felt via ecosystem-wide changes that influence supplies of feed, alter the incidence of various diseases, or lead to political instability. All these factors will, in the short term, increase the risk of decline of indigenous breeds. However, the adaptational characteristics of local breeds have value in synchronizing animal genetic resources with local stressors and should contribute to development of new, improved breeds that can better cope with changing production conditions. Current breeding goals have begun to place increasing weight on functional traits, but further adjustments may be required in the future to address higher temperatures, lower-quality diets, and greater disease challenges (Hoffmann, 2010). Species and breeds that are well adapted to such conditions may become more widely used.

Given the potential for significant future changes in livestock production conditions and objectives, a synthesis of the utilizationist and preservationist views on genetic resource management is needed. There is need to utilize genetic resources from a variety of sources and manage them in ways that will maximize livestock productivity in sustainable production systems. Such a synthesis must likewise be accompanied by programs to protect breeds which may be endangered or replaced, and requires better characterization of breeds, production environments, and associated knowledge; the compilation of more complete breed inventories; improved mechanisms to monitor and respond to threats to genetic diversity; more effective *in situ* and *ex situ* conservation measures; genetic improvement programs targeting adaptive traits in high-output and performance traits in locally adapted breeds; increased support for developing countries in their management of animal genetic resources; and wider access to genetic resources and associated knowledge. Also, additional research in the valuation of nonmarket products and services, including ecosystem services, provided by diverse livestock breeds is needed in order to reduce the disparity between private and public interest in farm animal genetic resource conservation.

Appendix

List of Courses

1. Animal Breeding and Genetics
2. Livestock Conservation Genetics

See also: Agriculture, Industrialized. Agriculture, Sustainable. Agriculture, Traditional. Agrobiodiversity. Cattle, Sheep, and Goats, Ecological Role of. Climate Change and Wild Species. Ecology of Agriculture. Gene Banks. Genetic Diversity. Inbreeding and Outbreeding. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon

References

Baumung R, Simianer H, and Hoffmann I (2004) Genetic diversity studies in farm animals – a survey. *Journal of Animal Breeding and Genetics* 121: 361–373.

Boettcher PJ, Tixier-Boichard M, Toro MA, *et al.* The GLOBALDIV Consortium (2010) Objectives, criteria and methods for using molecular genetic data in priority setting for conservation of animal genetic resources. *Animal Genetics* 41(supplement 1): 64–77. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2052.2010.02050.x/abstract

Brooks SA, Gabreski N, Miller D, *et al.* (2010) Whole-genome SNP association in the horse: Identification of a deletion in myosin va responsible for Lavender Foal Syndrome. *Public Library of Science Genetics* 6(4): e1000909. http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1000909

Carson A, Elliot M, Groom J, Winter A, and Bowles D (2009) Geographical isolation of native sheep breeds in the UK – Evidence of endemism as a risk factor to genetic resources. *Livestock Science* 123: 288–299.

CBD (2006) *CBD COP 8 Decision VIII/23 Agricultural biodiversity, D. In-Depth Review of the Programme of Work on Agricultural Biological Diversity.* Montreal, Canada: Secretariat of the Convention on Biological Diversity.

CBD (2008a) *COP 9 Decision IX/1. In-Depth Review of the Programme of Work on Agricultural Biodiversity.* Montreal, Canada: Secretariat of the Convention on Biological Diversity.

CBD (2008b) *UNEP/CBD/SBSTTA/13/2. In-Depth Review of the Implementation of the Programme of Work on Agricultural Biodiversity.* Montreal, Canada: Secretariat of the Convention on Biological Diversity.

Davis GH (2005) Major genes affecting ovulation rate in sheep. *Genetics Selection Evolution* 37(supplement 1): 11–23.

FAO (1998) Technical Annex 4: Project MoDAD – the FAO Global Project for the Measurement of Domestic Animal Diversity. *Primary Guidelines for Development of National Farm Animal Genetic Resources Management Plans.* Rome: Food and Agriculture Association of the United Nations. http://dad.fao.org/cgi-bin/getblob.cgi?sid=-1,50006248

FAO (1999) *The Global Strategy for the Management of Farm Animal Genetic Resources: Executive Brief.* Rome: Food and Agriculture Association of the United Nations. http://lprdad.fao.org/cgi-bin/getblob.cgi?sid=-1,50006152

FAO (2007a) In: Rischkowsky B and Pilling D (eds.) *The State of the World's Animal Genetic Resources for Food and Agriculture.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/010/a1250e/a1250e00.htm

FAO (2007b) *Global Plan of Action for Animal Genetic Resources and the Interlaken Declaration.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/010/a1404e/a1404e00.htm

FAO (2009a) Preparation of national strategies and action plans for animal genetic resources. *FAO Animal Production and Health Guidelines, No. 2.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/012/i0770e/i0770e00.htm

FAO (2009b) *Status and Trends Report on Animal Genetic Resources – 2008.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/ag/againfo/programmes/en/genetics/documents/CGRFA_WG_AnGR_5_09_Inf_7.pdf

FAO (2010a) Breeding strategies for sustainable management of animal genetic resources. *FAO Animal Production and Health Guidelines, No. 3.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/012/i1103e/i1103e00.htm

FAO (2010b) *Status and Trends Report on Animal Genetic Resources – 2010.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/meeting/021/am131e.pdf

FAO (2010c) *The State of Food and Agriculture 2009. Livestock in the Balance.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/012/i0680e/i0680e00.htm

Groeneveld LF, Lenstra JA, Eding H, *et al.* (2010) The GLOBALDIV Consortium. Genetic diversity in farm animals – a review. *Animal Genetics* 41(supplement 1): 6–31. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2052.2010.02038.x/abstract

Gura S (2008) Industrial livestock production and its impact on smallholders in developing countries. Consultancy Report, League for Pastoral Peoples and Endogenous Livestock Development, Ober-Ramstadt, Germany. http://www.pastoralpeoples.org/docs/gura_ind_livestock_prod.pdf

Hayes BJ, Bowman PJ, Chamberlain AJ, and Goddard ME (2009a) Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science* 92: 433–443.

Hayes BJ, Bowman PJ, Chamberlain AJ, *et al.* (2009b) A validated genome wide association study to breed cattle adapted to an environment altered by climate change. *PLoS ONE* 4(8): e6676. http://www.plosone.org/article/info:doi:10.1371/journal.pone.0006676

Hayes B and Goddard M (2010) Genome-wide association and genomic selection in animal breeding. *Genome* 53: 876–883.

Hoffmann I (2010) Climate change and the characterization, breeding and conservation of animal genetic resources. *Animal Genetics* 41(supplement 1): 32–46. http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2052.2010.02043.x/abstract

Hoffmann I (2011) Livestock biodiversity and sustainability. *Livestock Science* 139: 69–79.

Hoffmann I, Boerma D, and Scherf B (2011) The global plan of action for animal genetic resources – the road to common understanding and agreement. *Livestock Science* 136: 7–14.

Hoffmann I and Scherf B (2010) Implementing the global plan of action for animal genetic resources. *Animal Genetic Resources* 47: 1–10. http://www.fao.org/docrep/013/i1823t/i1823t00.pdf

Madalena FE (2008) How sustainable are the breeding programs of the global main stream dairy breeds? The Latin American situation. *Livestock Research for Rural Development* 20. Article #19. http://www.lrrd.org/lrrd20/2/mada20019.htm

Madalena FE, Teodoro RL, Lemos AM, Monteiro JBN, and Barbosa RT (1990) Evaluation of strategies for cross-breeding of dairy cattle in Brazil. *Journal of Dairy Science* 73: 1887–1901.

Marshall K, Quiros-Campos C, van der Werf JHJ, and Kinghorn B (2011) Marker-based selection within smallholder production systems in developing countries. *Livestock Science* 136: 45–54.

NRC (1993) *Managing Global Genetic Resources: Livestock.* Washington, DC: National Academy Press.

RBCSNZ (2002) *Enderby Island Cattle: A New Zealand Rare Breed Society Rescue Project.* Christchurch: Rare Breed Conservation Society of New Zealand. http://www.rarebreeds.co.nz/endcattlepro.html

Rodgers BD and Garikipati DK (2008) Clinical, agricultural, and evolutionary biology of myostatin: A comparative review. *Endocrine Reviews* 29: 513–534.

Steinfeld H, Gerber P, Wassenaar T, Castel V, Rosales M, and de Haan C (2006) *Livestock's Long Shadow.* Rome: Food and Agriculture Association of the United Nations. http://www.fao.org/docrep/010/a0701e/a0701e00.HTM

Wright S (1940) Breeding structure of populations in relation to speciation. *American Naturalist* 74: 232–248.

Butterflies

Philip J DeVries, Milwaukee Public Museum, Milwaukee, WI, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 559–573, © 2001, Elsevier Inc.

Glossary

Aposematic Describing an organism that is rendered less susceptible to predation by advertising its obvious unpalatability.

Batesian mimicry A form of mimicry in which the target organism is rendered less susceptible to predation by its resemblance in morphology or coloration to a different species that is unpalatable.

Cryptic Describing an organism that is concealed or obscured by the similarity of its appearance to the surrounding environment.

Müllerian mimicry A form of mimicry in which two or more unpalatable species resemble each other, with the effect that predators are more likely to avoid any species with this appearance.

Myrmecophily Ability to form symbiotic associations with ants.

Vibratory papillae Mobile, grooved, rod-like appendages arising from the distal edge of the first thoracic segment, used for communicating.

Overview of Butterfly Taxonomic Diversity

Although butterflies may be the best known group of insects, our understanding of their taxonomic diversity has two fundamental weaknesses. The first regards the recent decline of professional butterfly taxonomists and species level revisions in the past 50 years. This requires most species diversity estimates (from family to genus) to be derived from a literature that is out-of-date. The second is that the number of families and subfamilies of butterflies varies among classifications because, despite current interest in phylogenetic systematic methods and analyses, the relationships within the major groups are unresolved, particularly within Nymphalidae and Lycaenidae. Such is the state of butterfly taxonomy. Nevertheless, the "true" butterflies (superfamily Papilionoidea) may be placed conservatively into four families (Papilionidae, Pieridae, Nymphalidae, and Lycaenidae) that together include between 12,900 and 15,819 species. One useful framework for organizing butterfly taxonomic diversity is P. R. Ackery's (1984) synthesis of butterfly classification, which forms the basis of the following synopsis:

Family Papilionidae (Swallowtails, **Figure 1**): a group of 500–600 species in three subfamilies, distributed worldwide but with most species being tropical; adults medium to large sized, both sexes have six walking legs bearing nonbifid tarsal claws; most are brightly colored, may be Batesian or Müllerian mimics, all feed on flower nectar, and males drink at wet soil (**Figure 1**); caterpillars are herbivores, body smooth without hardened spines, and possess extrusible glands (osmeteria) that are unique among butterflies; pupa typically with a silk girdle at third thoracic segment.

Subfamily Baroniinae: one species (*Baronia brevicornis*) endemic to Mexico; caterpillars unique among the family by feeding on Fabaceae.

Subfamily Parnassinae: 40–50 species, mainly in north temperate mountains; host plant families include Crassulaceae, Fumariaceae, and Zygophyllaceae.

Subfamily Papilioninae (**Figure 1**): more than 500 species placed into three or more tribes; distributed worldwide; most species are tropical. Many species with long hindwing tails; many with sexes strongly dimorphic; many involved in mimicry; species range from palatable mimics (*Papilio* and *Eurytides*) to unpaltable models (*Parides, Pachliopta*, and *Troides*), and some are among the largest butterflies (*Trogonoptera* and *Ornithoptera*); host plant families include Annonaceae, Apiaceae, Aristolochiaceae, Fumariaceae, Hernandicaeae, Lauraceae, Magnoliaceae, Rosaceae, and Rutaceae.

Family Pieridae (whites, sulphurs, jezebels, cabbage butterflies; **Figure 1**): a group of more than 1100 species in four subfamilies; adults small to medium-sized, both sexes have six walking legs and distinctly bifid claws; most species are white or yellow or orange (or combinations thereof) derived from pterin pigments, some with red and black patterning; adults feed on flower nectar (**Figure 1**), and males visit wet soil; caterpillars are herbivores, body smooth without hardened spines, often covered in granulations; pupa suspended at 45° angle from substrate with silk girdle at the first abdominal segment.

Subfamily Pseudopontinae: one species (*Pseudopontia paradoxa*) endemic to West Africa; early stage biology unknown.

Subfamily Dismorphiinae: approximately 100 species, nearly all Neotropical. Most *Dismorphia* (**Figure 1**) are astonishingly precise mimics of certain aposematic Nymphalidae (*Heliconius, Mechanitis*, and *Oleria*), and represent the original example from which Batesian mimicry theory was derived; host plants are in the Fabaceae.

Subfamily Pierinae (whites, jezebel, and cabbage butterflies): approximately 700 mostly tropical species, especially in Africa and Asia; some species have spectacularly bright, contrasting colors of white, red, and black (e.g., *Delias, Mylothris, Pereute*, and allies) suggesting unpalatability and warning coloration; some groups with dimorphic sexes in which females resemble unpalatable nymphalids and papilionids; some species undergo spectacular mass migrations (e.g., *Ascia*,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00018-6

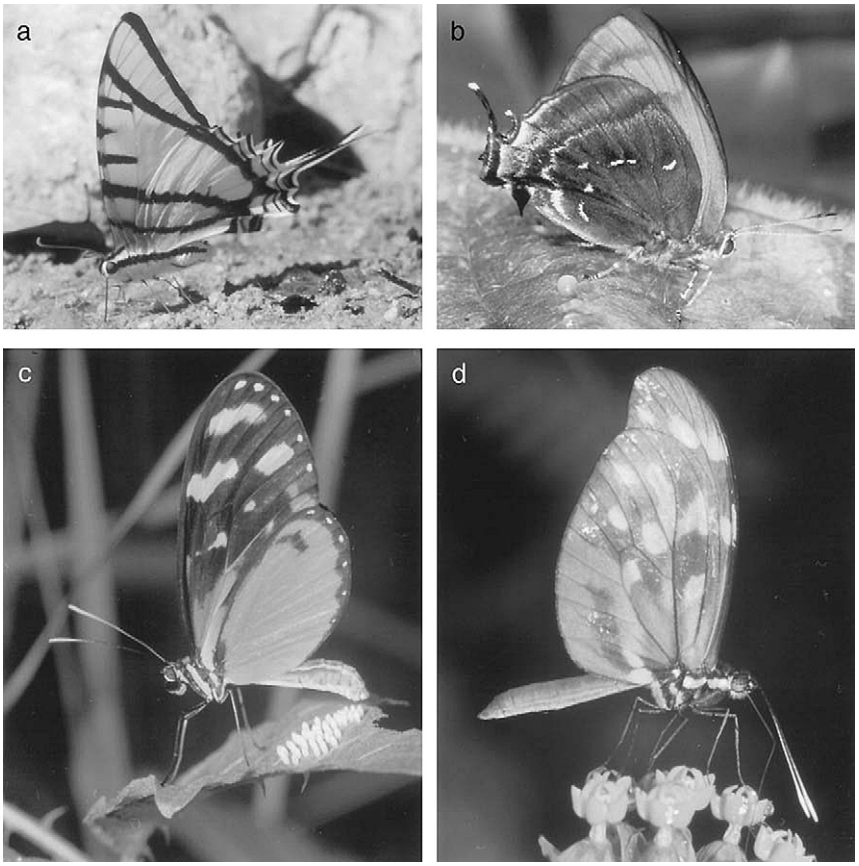


Figure 1 Adult butterflies. Clockwise from upper left: (a) newly eclosed male *Eurytides protesilaus* (Papilionidae) drinking from wet soil; (b) male *Theritas* nr. *hemon* (Lycaenidae: Theclinae) perching on a leaf; (c) male *Dismorphia amphiona* (Pieridae: Dismorphinae) visiting flowers; (d) female *Mechanitis isthmia* (Nymphalidae: Ithomiinae) ovipositing a cluster of eggs. (All photos copyright P. J. DeVries.)

Appias, and *Belenois)*, and a few species are crop pests (*Pieris*); host plant families include Capparidaceae, Brassicaceae, Santalaceae, and Loranthaceae.

Subfamily Coliadinae (sulphurs): 400–600 species well represented in temperate and subtropical regions; generally yellow or white with short, thickly scaled antennae; common in open areas, some migrate in large numbers and have been recorded out at sea far from land (*Colias, Phoebis, Catopsilia, Aphrissa*, and *Eurema*); many species with patterns visible only in the ultraviolet.

Family *Lycaenidae* (hairstreaks, blues, coppers, and metal marks; **Figure 1**): a group of 6000–6500 mostly tropical species in 10 subfamilies (the total number may change with future study), and accounts for nearly 50% of all butterflies; most species are small to very small, both sexes have six walking legs (except Riodininae), and frequently with alternating black and white bands on the antennae; the group displays a tremendous diversity of form, color, and life histories; adults feed on flower nectar, fruits, carrion, and honeydew, and a few species do not feed as adults; some African and Southeast Asian groups are involved in mimicry complexes; caterpillars most often slug-like, without hard spines or projections, but the Riodininae shows an extensive variety of form; as a group, lycaenids have the widest diet breadth of all butterflies, and depending on the group caterpillars may feed on plants, other insects, or insect secretions; many caterpillars form intimate and complex symbiotic associations with ants and produce acoustical calls similar to ant calls, secretions that are harvested by ants, and chemicals that alter ant behaviors; pupae are typically round, seed-shaped, unadorned with projections, and may produce clicking or whirring sounds when stimulated.

Subfamily Riodininae (metalmarks): 1200–1400 species that are almost exclusively Neotropical; arguably the most diverse of all lycaenoid groups with respect to adult and larval forms, and they are often treated as a distinct family (Riodinidae) divided into five subfamilies; adults feed primarily on flower nectar, some drink at wet soil, others at carrion; overall their life histories are poorly known, but as a group riodinid caterpillars appear to be mainly herbivores (*Euselasia, Mesosemia, Ancyluris, Metacharis, Mesene, Symmachia, Emesis, Anteros*, and *Helicopis*), with a few carnivorous species on Homoptera (*Alesa* and *Setabis*); about one-third of the riodinids form symbioses with ants and have vibratory papillae as sound producing organs (*Thisbe, Audre, Lemonias, Synargis, Nymphidium*, and *Setabis* (**Figure 2**)); riodinid myrmecophily is entirely Neotropical; host plant families include Araceae, Asteraceae, Bromeliaceae, Bombacaceae, Cecropiaceae, Clusiaceae, Dilleniaceae, Euphorbiaceae, Fabaceae, Lecythidaceae, Loranthaceae, Malpighiaceae, Marantaceae, Melastomataceae, Myrtaceae, Orchidaceae, Rubiaceae, Sapindaceae, and Zingiberaceae plus mosses, liverworts, and lichens.

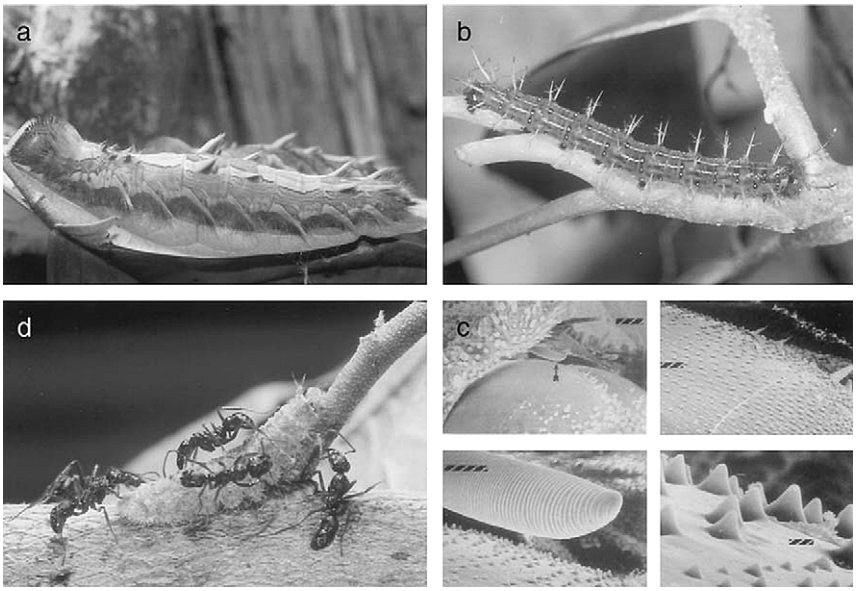


Figure 2 Butterfly caterpillars. Clockwise from upper left: (a) final instar caterpillar of *Morpho achilles* (Nymphalidae: Morphinae); (b) final instar caterpillar of *Tigridia acesta* (Nymphalidae: Nymphalinae); (c) Four frames showing details of the call production mechanism of *Thisbe irenea* (Lycaenidae: Riodininae). Upper left—top of caterpillar head, a vibratory papilla (arrow) and anterior margin of first thoracic segment. Upper right—granulations on top of caterpillar head. Lower left—detail of vibratory papilla. Lower right—head granulations at high magnification; (d) final instar caterpillar of *Thisbe irenea* (Lycaenidae: Riodininae) being tended by ants. (All photos copyright P. J. DeVries.)

Subfamily Styginae: one enigmatic species (*Styx infernalis*) that is endemic to the Peruvian high Andes; the early stage biology is unknown.

Subfamily Lipteninae: a small African group of 30–40 species; some are Batesian mimics of aposematic nymphalids (*Mimacraea* and *Ornipholidotos*); caterpillars feed on lichens and microscopic fungi (*Durbania* and *Deloneura*).

Subfamily Poritiinae: a small group restricted to the Oriental regions; caterpillars may be gregarious, feed on plants in the Fagaceae (*Poritia*), and may not associate with ants.

Subfamily Liphyrinae: a small group found in Africa, Australia, and Asia; adults have the proboscis partly or entirely atrophied; depending on the species, caterpillars feed entirely on insects (Homoptera) or on ant brood within ant nests (*Liphyra* and *Aslauga*).

Subfamily Miletinae: a moderate-sized group with most species in tropical Africa and the Orient and one in North America; caterpillars mainly feed on nymphs of Homoptera (*Feniseca*, *Spalgis*, and *Allotinus*), with a few species feeding on secretions of Homoptera (*Allotinus*) or ant regurgitations (*Thestor*); adults feed mainly on honeydew secretions of Homoptera.

Subfamily Curetinae: approximately 40 species restricted mainly to tropical Asia, all in the genus *Curetis*; caterpillars with tentacle organs as conspicuous, rigid cylindrical tubes; caterpillars feed on Fabaceae and have loose associations with ants.

Subfamily Theclinae (hairstreaks; **Figure 1**): this very large and diverse group is found worldwide, but most species are tropical; caterpillars are mainly herbivores, some carnivores on Homoptera, and many associate with ants, but the life histories of most species remain unknown; host plant families include Anacardiaceae, Annonaceae, Asteraceae, Bromeliaceae, Clusiaceae, Cycadaceae, Euphorbiaceae, Fabaceae, Fagaceae, Geraniaceae, Lecythidaceae, Loranthaceae, Malpighiaceae, Malvaceae, Melastomataceae, Meliaceae, Myrtaceae, Oleaceae, Orchidaceae, Sapindaceae, Sapotaceae, Solanaceae, Sterculiaceae, Ulmaceae, and Verbenaceae.

Subfamily Lycaeninae (coppers): a group of 20–40 species found mainly in temperate regions, a few are tropical; caterpillars are herbivores on Polygonaceae, and some form loose associations with ants.

Subfamily Polyommatinae (blues): a large group with a worldwide distribution, most are pale, reflective blue above; caterpillars appear to always associate with ants, and their diets range from herbivores to carnivores, and some (*Maculinea*) feed on ant larvae and are clearly complex parasites and predators within ant nests; host plant families include Crassulaceae, Euphorbiaceae, Fabaceae, Lamiaceae, Myrsinaceae, Oxalidaceae, Primulaceae, Rhamnaceae, Rutaceae, Santalaceae, Sapindaceae, Saxifragaceae, Selaginaceae, Sterculiaceae, Verbenaceae, Zingiberaceae, and Zygophyllaceae.

Family Nymphalidae (brush-footed butterflies; **Figures 1** and **2**): a group of 4800–6000 species in 14 subfamilies embracing a prodigious variety of forms and sizes, with both sexes having four walking legs; the forelegs are greatly reduced (hence "brush-footed"); adults may be dull brown (Satyrinae and Nymphalinae), brightly colored (Nymphalinae and Brassolinae), brilliantly iridescent due to physical properties of the wing scales (Morphinae and Apaturinae), or transparent (Ithomiinae and Satyrinae); some groups are entirely palatable, others highly distasteful, and some (Danainae, Ithomiinae, Heliconiinae, and Acraeinae) are extremely important mimetic models; some species (Charaxinae and Nymphalinae) mainly inhabit tropical forest canopies; adults may feed on flower nectar, pollen, rotting fruits, carrion, or do not feed at all; the caterpillars may bear many spines on the body, some

also have head spines, whereas others are devoid of spines (Figure 2); caterpillars are entirely herbivorous, and the particular groups exhibit strong associations with particular plant families; the pupae are typically suspended.

Subfamily Charaxinae: a group of 350–400 mainly tropical species; adults fly very fast with robust bodies; underside of wings typically camouflaged and leaf-like, some with brilliantly colored upperside (*Agrias*, *Prepona*, and *Charaxes*); all are palatable, with few Batesian mimics (*Consul* and *Euxanthe*); adults feed primarily on juices of rotting fruit, dung, and/or carrion (rarely flower nectar), and most inhabit the forest canopy; caterpillars have smooth bodies, often bearing a corona of head spines or projections; host plant families include Annonaceae, Celastraceae, Convolvulaceae, Euphorbiaceae, Fabaceae, Flacourtiaceae, Lauraceae, Myrtaceae, Piperaceae, Poaceae, Rhamnaceae, Rutaceae, Santalaceae, and Sapindaceae.

Subfamily Apaturinae: a small group of medium to large species that are mainly tropical (often placed within the Nymphalinae); males often with brilliant iridescence on wing upperside, the Neotropical species (*Doxocopa*) with proboscis and forelegs lime-green; adults are strong fliers and feed entirely on carrion and putrefying fruits; caterpillars are herbivores on Ulmaceae and have a smooth body, bifid tail, and head with a pair of strong horns.

Subfamily Satyrinae (satyrs, wood nymphs, and browns): a cosmopolitan group of 1500–2000 mainly tropical species that are generally dull brown (some notable exceptions) with well-developed eyespots on the wings; most feed on juices of rotting fruits, some on fungi, and some on flower nectar in temperate regions; all are palatable, with only few clear examples of Batesian mimicry from Asia (*Penthema*, *Zethera*, and *Elymnias*); caterpillars are smooth with bifid tails, and some bear paired head projections; host plant families include Arecaceae, Araceae, Cyperaceae, Heliconiaceae, Poaceae, and Selaginellaceae.

Subfamily Brassolinae (owls): 70–80 entirely Neotropical species ranging from medium to some of the largest butterflies known (*Caligo*) that fly at dawn and dusk, most with characteristic, large eyespots on the hindwing underside, and they are common in butterfly conservatories; most species feed entirely on rotting fruit juices, but a few with a strongly reduced proboscis (*Brassolis* and *Dynastor*) may not feed; caterpillars have smooth bodies, often with dorsal pseudospines, bifid tails, and multiple horn-like projections on the head; host plant families include Arecaceae, Bromeliaceae, Heliconiaceae, Musaceae, and Poaceae, and they can be pests in banana and palm plantations.

Subfamily Amathusiinae (fauns and duffers): an Indo-Australian group of approximately 80 species (sometimes placed in Brassolinae or Morphinae); medium to large butterflies that fly at dawn and dusk and feed on rotting fruit juices; the group is palatable (except for perhaps *Taenaris*) and shows little or no mimicry; caterpillars have smooth bodies, some with long, downy setae, bifid tails, often with paired head horns; host plant families include Arecaceae, Musaceae, Poaceae, and Smilacaceae.

Subfamily Morphinae (morphos; Figure 2): a group of 40–50 Neotropical species of medium to very large butterflies; the large reflective blue species (*Morpho*) are immediately noticeable in nature and in butterfly conservatories, but others that fly in the forest understory (*Antirrhea* and *Caerois*) are seldom observed; all feed on rotting fruits, none are considered to be unpalatable, and none are mimics; a favorite of collectors and butterfly conservatories, surprisingly little is known of their natural history; caterpillars are covered with red, yellow, and green patterns, bear tufts of dorsal and lateral setae, possess bifid tails, and bear short, paired projections on the head (Figure 2); host plant families include Arecaceae, Bignoniaceae, Fabaceae, Menispermaceae, Poaceae, and Sapindaceae.

Subfamily Calinaginae: a group represented by two to five species in the genus *Calinaga* restricted to the Himalayan regions; it is apparent that *Calinaga* forms mimicry complexes with *Parantica* (Danainae), but it is not clear if it is Batesian or Müllerian mimicry; caterpillars are smooth with short bifid tails and stout head horns and feed on Moraceae.

Subfamily Nymphalinae: a diverse, cosmopolitan group of more than 3000 species (sometimes split into Limenitinae and Nymphalinae) containing a tremendous range of size and color patterns; some are migratory (*Vanessa*, *Eunica*, and *Sallya*), some are unpalatable models, some are palatable mimics, some show tremendous seasonal polymorphism (*Junonia*), some feed on flower nectar, others feed on rotting fruits and carrion; some groups inhabit tropical forest canopies (*Euphaedra*, *Cymothoe*, *Panacea*, *Epiphile*, and *Baeotus*), some pass north temperate zone winters as adults (*Nymphahlis*); caterpillars usually covered in spines, many with well-developed pairs of head spines (Figure 2); pupa often with bifid projections on head; host plant families include Acanthaceae, Caprifoliaceae, Convolvulaceae, Euphorbiaceae, Fagaceae, Flacourtiaceae, Lamiaceae, Loranthaceae, Moraceae, Plantaginaceae, Poaceae, Rubiaceae, Rutaceae, Salicaceae, Sapindaceae, Scrophulariaceae, Urticaceae, and Verbenaceae.

Subfamily Acraeinae: a group of approximately 250 tropical African species, with a few in Asia and the neo-tropics; ranging from small to large, all are slow-flying unpalatable species that contain cyanogenic compounds, forming aposematic models for a variety of other groups, and are involved in Müllerian mimicry (*Acraea*, *Bematistes*, and *Actinote*); one species (*Acraea encedon*) is almost entirely female and reproduces via parthenogenesis; caterpillars (many which feed gregariously in communal nests) are densely covered in spines, but lack head spines; host plant families include Asteraceae, Passifloraceae, Sterculiaceae, Tiliaceae, and Urticaceae.

Subfamily Heliconiinae (passionflower butterflies): a group of 70–80 mainly Neotropical species with a few in Asia; most with well-developed eyes that are wider than the thorax, elongate forewings, and most serve as models in Batesian and Müllerian mimicry complexes; they feed on flower nectar, and some on pollen (*Heliconius* and *Laparus*) from which they can derive some of their unpalatability; caterpillars are densely spiny, bear paired head spines, and feed entirely upon plants in the Passifloraceae and allies—hence the name passionflower butterflies.

Subfamily Danainae (milkweed butterflies, tigers, and crows): a cosmopolitan group of approximately 150 species, with most in tropical Africa and Asia; ranging from medium to large size, most are slow flying, conspicuously colored, and models in mimicry complexes (*Danaus*, *Amauris*, *Idea*,

Tirumala, and *Parantica*); their distasteful nature derives from feeding as caterpillars on milkweeds, but they also acquire other chemical defenses by feeding as adults on flowers and wounds in certain plants containing particular alkaloids; males often have well-developed, brush-like scent hairs they use during courtship; caterpillars are smooth, often patterned with alternating bands of black, white, and yellow, and many have dorsal pairs of fleshy tubercles; the pupae are frequently reflective gold or silver; host plant families include Apocynaceae, Asclepidaceae, and Moraceae.

Subfamily Ithomiinae (ithomiines and glass wings; **Figure 1**): approximately 300 entirely Neotropical species; adults typically with a very small head, elongate wings, and species vary in color from transparent (*Greta*, *Ithomia*, and *Pteronymia*) to bright tiger-striped patterns (*Mechanitis, Tithorea,* and *Melinaea* (see **Figure 1**)); they are involved in both Müllerian and Batesian mimicry, representing unpalatable models for many other groups; their unpalatable properties derive from larval host plants and chemicals acquired by adult feeding, and males typically possess a tuft of scent hairs between the wings that disseminate pheromones; caterpillars are smooth, often with fleshy tubercles, and may be brightly colored or cryptic; The pupa is often reflective gold or silver and squat; host plant families include Apocynaceae, Gesneriaceae, and Solanaceae.

Subfamily Tellervinae: a group of 6–10 Australasian species all in the genus *Tellervo*; they serve as models for nymphaline and satyrine Batesian mimics; this group has been included within the Ithomiinae; the caterpillars resemble some Danainae and feed on Apocynaceae.

Subfamily Libytheinae (beaks and snout butterflies): a cosmopolitan group of approximately 10 species recognized by the large erect palpi that form a "snout"; most species are well-known to periodically undergo spectacular migrations; the caterpillars somewhat resemble those of the Pieridae, and all feed on Ulmaceae.

Early Stages and Host Relationships

Like all members of Lepidoptera, butterflies have four discrete stages in the life cycle (egg, caterpillar, pupa, and the adult), each with particular characteristics, behaviors, and requirements. Furthermore, to complete their life cycle, butterflies require a plant or insect host to feed on.

Egg

Butterfly eggs are laid either singly or in small to large clusters, either on or off the host, and the location where the egg is laid is typically important (**Figure 2**). The eggshell frequently has an elaborate sculpturing that plays a role in respiration, and each major group of butterfly has its own form of egg.

Caterpillars

Depending on the group, the appearance of butterfly caterpillars may range from cryptic to aposematic, they may be covered with spines and/or hairs or appear to be naked, and their diets may include plant tissue or the flesh of other insects or they may feed entirely on secretions produced by other insects (**Figure 2**). All caterpillars consist of three major sections: the head, thorax, and abdomen. The hardened head houses mandibles that function to shear off bites from their food. The head bears a gland that lays down silk that is grasped as the caterpillar moves forward, helping the caterpillar grip the substrate and also to secure rolled leaves in which they may shelter. Each thoracic segment bears a pair of true legs, whereas the 10 abdominal segments form the bulk of the body, housing the long gut. Caterpillars walk by using the prolegs (segments 3–6 and segment 10), and these function by hydraulic pressure and muscles. Segments 1–8 bear the external ring-like orifices (spiracles) that allow gas exchange with the atmosphere.

After reaching a particular size each caterpillar instar stops eating and undergoes a molt to the next instar; this is how caterpillars grow. Generally, but not always, there are five larval instars followed by a molt to the pupa, or chrysalis. Caterpillar growth is not a steady increase in weight from first to final instar; rather, there is a dramatic fluctuation of weight between each molt. Newly molted caterpillars may weigh about the same as or even less than previous instars, but their weight will quickly increase and exceed that of previous instars.

Caterpillar feeding behavior often differs among instars. In many groups, late-instar caterpillars may stop feeding during the day and then feed entirely at night. Alternatively, some lycaenid caterpillars may start life feeding on plants and, after molting to a later instar, they fall off the plant, are picked up by ants, and are carried into their nest where the caterpillars feed as carnivores on ant brood.

Pupa

When the final-instar caterpillar is fully grown, it stops eating and enters its molt to the pupa, or chrysalis. Within the pupa the caterpillar tissues are reconstructed into the adult by the process known as metamorphosis. The pupa attaches to a substrate by a series of hooks on the last segment (cremaster), and major groups typically have characteristic manners of pupation. For example, pupating with the head downward attached only by the cremaster is typical of Nymphalidae, but pupae of Papilionidae and some Pieridae attach by the cremaster with head upward and are restrained by a silk girdle. Pupae of Lycaenidae produce a whirring, clicking, or buzzing sound with a rasp and file system on the abdomen that may be a defense against predators. Sound production is also known in some nymphalid pupae. When mature, the pupa splits along its dorsal surface, and the adult ecloses. After eclosion, the adult normally hangs from the pupal shell or nearby with the wings suspended downward so that it can expand and dry its wings. If dried in a crumpled manner the wings are useless for flying, and the butterfly is effectively dead. As a rule, female butterflies are mated soon after or even before eclosion, exemplifying one of the most potent laws of evolution—nature abhors a virgin.

Adult

The adult butterfly (**Figure 1**) is incapable of additional growth but is capable of flight, mating, and reproduction. Like

all insects, the butterfly body is composed of the head, thorax, and abdomen.

The most obvious features of the butterfly head are the large compound eyes composed of numerous facets (ommatidia) that cannot focus but are sensitive to movement, light, and certain colors. A pair of distally thickened antennae arise from between the eyes that vary in shape according to the group. The antennae function as sensory organs for finding food, mating, and balance during flight and are sensitive to volatile chemicals. Between the eyes there is a pair of appendages called labial palpi, and between them lie the proboscis, a hollow tube composed of two interlocking halves that is coiled like a watch spring when not in use and can be extended for feeding. By virtue of having a proboscis, butterflies are restricted to a liquid diet that may include flower nectar, the juices of rotting fruit, carrion, dung, or semidigested pollen. Proboscis length may vary according to the group; in some species it is nearly vestigial, thus precluding feeding as adults (*Brassolis* and *Liphyra*), whereas in others the proboscis measures more than 1.5 times the length of the body (*Eurybia*), allowing them to take nectar from a wide range of flowers.

The thorax is composed of three fused segments bearing the wings and legs, and it contains the muscles for locomotion and various internal organs. As in all insects, the adult butterfly has six legs, one pair per segment. Butterflies have four wings (two forewings and two hindwings) typically covered in scales that give butterflies their characteristic colors and patterns. The color patterns of butterflies result mainly from the covering of scales that are arranged like overlapping roof tiles. There are three notable types of scales. The pigmentary scales are colored by the deposition of melanin, pterins, or other chemicals. Structural scales generate blue, violet, copper, or green colors by reflecting particular wave lengths of incident light. Androconial scales store and disseminate chemical odors (pheromones) that are used in mating, and some of these scales may be physically transferred to the female during mating.

The abdomen houses the digestive and reproductive tracts, terminating in the reproductive organs (genitalia). Except for segments housing the genitalia, the abdomen can stretch when the gut becomes filled with food, and abdominal distention may be considerable in groups feeding on rotting fruits (Charaxinae). The penultimate abdominal segment of males bears two appendages (claspers) that open to expose the aedeagus (penis) and serve to grip the female's abdomen during mating. The female abdomen terminates in three openings: the anus, egg pore, and copulatory pore. The configuration of genitalia is used extensively in butterfly taxonomy.

Host Relationships

An important aspect in the butterfly life cycle is the ability of ovipositing females to find, and caterpillars to feed on, particular plants. The liaison with plants is so strong that many groups of butterflies only associate with particular taxonomic groups of plants; other plants are unacceptable to both caterpillars and ovipositing females. In the Lycaenidae this association may extend to particular species of ants or Homoptera. For example, caterpillars of milkweed butterflies (Danainae) only feed on plant families containing milky latex, and those of the Heliconiinae use plants in the Passifloraceae. Such patterns of host association in butterflies gave rise to Ehrlich and Raven's classic paper that developed the concept of coevolution. On the whole, host association records for the Papilionidae, Pieridae, and Nymphalidae are much more complete than for the Lycaenidae.

Butterfly–Ant Symbioses

The ability to form symbiotic associations with ants (myrmecophily) occurs only within the Lycaenidae. Here, caterpillars provide food secretions to ants in exchange for protection against insect predators such as social and parasitic wasps. To form these symbioses, caterpillars have a suite of unique adaptations that may include organs (collectively known as ant-organs) to produce food secretions, volatile chemicals, and sound, all of which work in concert to modify ant behavior and enhance the protective attitude of ants toward caterpillars. Recent studies on ant-organs indicate that myrmecophily evolved at least twice in the butterflies: once in the Riodininae and independently in other lycaenid subfamilies.

The widespread trait of myrmecophily within the Lycaenidae and the fact that lycaenids account for approximately 50% of all butterfly species led Pierce (1984) to suggest that myrmecophily has amplified speciation rates in this group. Indeed, the diversity of life histories in myrmecophilous butterflies can be exceedingly complex, encompassing herbivores, carnivores, and those that feed as caterpillars only on secretions, and the associations with ants range from mutualistic to completely parasitic or predatory.

Food Secretions

Ants pay close attention to particular abdominal segments bearing ant-organs that produce food secretions, which in some species are known to have high concentrations of amino acids and sugars. In some Riodininae, these consist of a pair of organs (tentacle nectary organs) on the eighth abdominal segment that can be extruded individually or simultaneously. In all other lycaenid subfamilies, this organ is a single dorsal pore on segment 7 (dorsal nectary organ). Ants are so intent on obtaining the secretions that they constantly antennate the caterpillar to solicit more, and in many cases, this is a good example of a general rule among participants in symbiotic associations— "you scratch my back and I'll scratch yours."

Semiochemical Production

Myrmecophilous caterpillars may have extrusible glands that seem to produce volatile chemicals (semiochemicals or pheromones) that alter the behaviors of attending ants. Some Riodininae have a pair of extrusible glands on the third thoracic (anterior tentacle organs), whereas other lycaenids have a pair of glands on the eighth abdominal segment (tentacle organs). In both cases, when extruded from the body these organs do not

produce a liquid secretion but rather the tip is modified with spines that gives the appearance of a tiny feather duster. These spines likely provide a larger surface area to disseminate volatile chemicals that may be similar to ant alarm pheromones. Instead of anterior tentacle organs, a small group of Riodininae caterpillars (*Theope*) possess a corona of inflated setae around the head that appear to disseminate semiochemicals.

Call Production

The idea that caterpillars produce acoustic calls might seem unlikely. We now know, however, that myrmecophilous caterpillars produce substrate-borne calls that function in the formation and enhancement of their symbioses with ants, and these calls bear similarities to those produced by ants for communicating among themselves. In most Riodininae caterpillars calls are produced by a pair of mobile, grooved, rod-like appendages (vibratory papillae) arising from the distal edge of the first thoracic segment. An acoustical signal is produced when the grooves on the vibratory papillae grate against head granulations (**Figure 2**). In other lycaenid caterpillars the call is produced by thickened bumps located ventrally between abdominal segments. Most concepts of insect communication suggest that acoustical calls evolved in a sexual context. However, caterpillar calls provide an example showing that, by forming symbiotic associations, the call of one species may evolve to attract another, unrelated species.

Ants and Caterpillar Associations

In general, myrmecophily in caterpillars occurs with a particular type of ant. A basic element among myrmecophilous caterpillars is that they typically form symbioses only with ant species that depend heavily on secretions as food—ants that also form symbioses with Homoptera and plants. Therefore, secretion-harvesting ants likely played a key role in the evolution of myrmecophily, whereas those ants that are predators or herbivores did not. An ecological consequence of the evolution with secretion-harvesting ants is that in any suitable contemporary habitat, a suite of caterpillar, plant, and Homoptera species all share the same species of ant symbionts.

Among myrmecophilous caterpillars there are two main categories of ant association. The most widespread category comprises an association in which a particular caterpillar species may be tended by a suite of secretion-harvesting ant species. The other category comprises an association in which a caterpillar has an obligate association with a single species of ant. In this case, female butterflies may require the presence of a particular ant species to lay their eggs since caterpillars are unable to form symbioses with any other ant species. Furthermore, it is in these types of associations in which some caterpillars are adopted by the ants, taken into the nest, and become parasites or predators of their hosts.

Butterfly Mimicry and Diversity

When several species of butterflies share conspicuously bright color patterns and fly together in the same habitat, this is called mimicry—a widespread and important antipredator defense. Fundamentally, mimicry occurs when one species closely resembles another species, and based on outward appearance one or both are avoided by predators. Virtually all major taxonomic groups of butterflies exhibit mimicry, but the warningly colored (aposematic) models are found predominately in particular groups of Papilionidae, Pieridae, and Nymphalidae, whose caterpillars feed on poisonous plants. Mimicry is a complex and subtle topic, but there are two basic types of mimicry in butterflies, both involving species that advertise conspicuous color pattern to predators: Batesian and Müllerian mimicry.

Batesian Mimicry

This is the phenomenon whereby one or more palatable species resemble one or more unpalatable model species, and the palatable species gain protection by duping predators. Here, predators avoid mimics because they resemble unpalatable models. The best known basic example of Batesian mimicry involves the palatable viceroy (*Limenitis archippus*) and the unpalatable monarch butterfly (*Danaus plexippus*) in North America. However, Batesian mimicry occurs in all biogeographic regions, and one of the most intriguing examples involves the sex-limited mimicry of the wide-ranging African swallowtail, *Papilio dardanus*. Here, only the females are mimics. In this case, females are polymorphic with respect to their appearance, and their color patterns change geographically when flying with different model species. In other words, *P. dardanus* females from Kenya may look entirely different from those from Zaire or elsewhere, and in each instance they precisely mimic a different unpalatable model species. Thus, the type of female color pattern evolves in response to and depends on the local community diversity of the model species.

Müllerian Mimicry

This is the phenomenon whereby several unpalatable species (co-mimics) fly in the same area and share the same color pattern. In this case, the association of similarly patterned unpalatable species is thought to reinforce the nasty experience predators have when eating butterflies of such a color pattern, thus educating predators to avoid it.

Mimicry and Diversity

Aposematic species are an important element of butterfly diversity, particularly in the tropics. For example, approximately 20–25% of all Papilionidae, Pieridae, and Nymphalidae in Costa Rica and Kenya are clearly aposematic models. Although accounting for a large percentage of all butterfly species, there are few documented examples of aposematic Lycaenidae.

Müllerian mimicry is primarily a tropical phenomenon and is an important part of tropical faunal diversity. This can be appreciated by considering the diversity of color patterns shown by many co-mimetic species and races of ithomiine and heliconiine butterflies (plus many co-mimic moths and other butterfly groups) that converge across large areas of the

neotropics. Even when the tightly overlapping distribution of only two co-mimetic species of *Heliconius* and their many precisely convergent color patterns are mapped over the Neotropical region, one cannot doubt the power mimicry plays in the evolution, organization, and diversification of tropical butterfly faunas. The numerous co-mimetic species of Acraeinae butterflies throughout the African region also serve as another potent example of the influence of mimicry on the diversity of butterfly communities. It is clear that butterfly mimicry evolved in the context of multispecies interactions which involved butterflies, their predators, and the plant and insect hosts they feed on as adults and caterpillars. Thus, mimicry serves to remind us that habitat disturbance and/or destruction may have important implications for the continued survival and future evolution of interacting species.

Geographical Patterns of Butterfly Diversity

One of the most frequently asked questions about butterflies is the following: How many are there? The answer is thought to be between 12,900 and 15,819 species. This is hardly a satisfying answer since the question has little biological context. In contrast, questions framed by comparing different geographical areas impart a sense of dynamics to butterfly diversity. For example, one might ask: Does butterfly species diversity change with respect to latitude? or How do different areas of Africa compare with respect to numbers of butterfly species? Indeed, comparing species numbers among areas shows that latitude, biogeographical region, area size, and taxonomic affinity all contribute to global butterfly species diversity. A caveat is in order about comparative species numbers: Sampling effort is almost always unequal among different areas, and the taxonomic accuracy of species counts or lists often varies among sites. Thus, it is prudent to use care when interpreting the generality of species diversity patterns among areas.

Table 1 Approximate numbers of American and world butterfly species

	North America	*Neotropics*	*World*
Papillionidae	35	169	621–645
Pieridae	63	347	1086–1105
Lycaenidae	171	2300–3000	6000–6900
Nymphalidae	202	1850–2500	5175–5975

Latitudinal Gradients of Species Diversity

As is the case for most groups of terrestrial organisms, butterfly species diversity increases toward the equator. Comparing the North American and Neotropical faunas suggests that the neotropics has an estimated 10–12 times more butterfly species compared to North America (**Table 1**). Comparing site diversity at different latitudes, however, provides a richer perspective of this phenomenon. A latitudinal transect using five well-known sites from the Americas makes it obvious that species numbers increase when moving from the north latitudes toward the equator (**Table 2**), even though the smallest areas being compared are at the equator. The dramatic equatorial increase in species is exemplified by the fact that a mere 500 ha of lowland Ecuadorian forest has more butterfly species than all of North America.

The relative contribution of each butterfly family to site diversity also varies with latitude. From **Table 2**, it is evident that the relative contribution of Papilionidae and Pieridae to total species diversity is about the same regardless of latitude, but the contributions of Nymphalidae and Lycaenidae increase dramatically between 20 and 0° latitude (**Figure 3**). The increase in species of Nymphalidae is accounted for by subfamilies that reach their greatest diversity in the Amazon (Morphinae, Brassolinae, Charaxinae, Satyrinae, and Ithomiine), whereas the increased contribution of Lycaenidae is due mainly to the subfamily Riodininae, which is also an Amazonian group.

Variation Among Neotropical Sites

Species diversity is not equal among Neotropical forest sites. Nine well-known sites that have many species in common vary strongly in numbers of species and in the contribution of each family to species diversity (**Table 3**). Here, species numbers may range from 212 in Mexico to 1199 at one Brazilian site. Among six Amazonian sites, numbers range from 676 to 1199 species. When averaged over the entire **Table 3**, the contribution of each family to species diversity is as follows: Papilionidae, 3.8%; Pieridae, 4.8%; Nymphalidae, 45.2%; and Lycaenidae, 46.1%. It may be significant that these proportions approximately reflect those observed when comparing latitude (**Figure 3**).

Comparisons among sites from the Americas point to the general relationship of the higher species richness that occurs at lower latitudes and to the marked variation among sites that share many species in common. These examples also show that the relative contribution of each family to species richness

Table 2 Latitudinal transect showing changes in American butterfly species diversity

Site	*Latitude*	*Area*	*Family*				*Total species*
			Papilionidae	*Pieridae*	*Nymphalidae*	*Lycaenidae*	
H. J. Andrews Forest, Oregon	44°14′	6400 ha	5	7	27	23	62
Los Angeles Basin, California	34°4′	>1 million ha	6	15	26	32	79
Estacion Los Tuxlas, Mexico	18°35′	700 ha	14	19	113	66	212
La Selva, Costa Rica	10°26′	1000 ha	16	26	219	181	442
Garza Cocha, Ecuador	0°29′	500 ha	24	25	314	312	676

is not equal; the proportion of nymphalid and lycaenid species increases toward the equator, but the proportion of papilionid and pierid species appears relatively constant. At this point, we might ask if there are similar patterns outside the neotropics.

Regional Patterns of Species Diversity

Although most butterfly species are tropical, the number of species is not distributed equally among tropical regions (Table 4). As one might expect, those areas closest to the equator have a greater number of species, but there is also an effect of size on species diversity (Table 4). In this comparison, the continent of Africa has by far the most species, but it also encompasses the greatest geographical area and habitat types (deserts, mountains, forests, and savannas) of any area. Within Africa there are more species in Zaire than in Kenya or Southern Africa, highlighting the ecological diversity in African climate and habitat types. Thus, the entire continent of Africa should contain more species than a geographic subset.

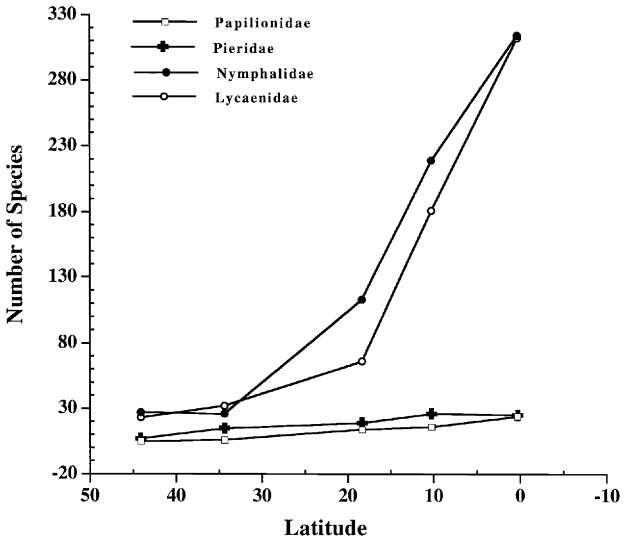


Figure 3 Latitudinal gradient of American butterfly species diversity. Each line represents species numbers per family contributed to five sites distributed from 44° north to the equator.

As in the neotropics, the relative contributions of families to total species numbers differ among areas of Africa (Table 4). Here, the Papilionidae and Nymphalidae contribute equally to total richness of both Kenya and Zaire, but Kenya shows an increased proportion of Pieridae and a decreased proportion of Lycaenidae. On the other hand, the proportional contributions of Papilionidae and Pieridae are reversed in the Zaire fauna. The Papilionidae and Nymphalidae in the Southern African fauna contribute less to the total than other mainland African areas, but nearly 55% of all Southern African butterflies are Lycaenidae. Overall, these proportional differences (Table 4) may reflect the ecological responses of taxonomic groups to differences in habitat types and/or the differences among regional taxonomists. Finally, in comparison to mainland Africa the fauna of Madagascar is interesting. Here, only 17% of all butterflies are lycaenids, but nymphalids constitute 67% of the total fauna, most of which are Satyrinae (i.e., 42% of all the butterflies). This example demonstrates how historical colonization and subsequent radiation of an island by one group can produce a fauna distinctly different in composition from that of the mainland.

Only a broad-brush comparison among areas of different sizes and regions is necessary to appreciate that, overall, butterfly diversity is highest in the neotropics (Tables 1–5). Despite the great disparity in geographical area in the comparison, the small country of Costa Rica has a greater percentage of the world's butterfly species than does North America, Australia, New Guinea, or the peninsula of Malaysia (Table 5). Since our taxonomic understanding of African, Australian, and Oriental butterflies is more thorough than that of the neotropics, it is likely that the total Neotropical butterfly species count will increase in the future.

Species Diversity in Space and Time

When estimating species diversity in any habitat, the variables of time, space, and sample size can be profoundly important. For example, consider two samples taken in the same Wisconsin prairie—one for 7 days in mid-July and the other for 7 days in mid-January. Despite equal sampling effort, the number of species tallied in January would certainly be zero

Table 3 Comparative diversity among neotropical sites[a]

Site	*Family*				*Total species*	*Total area (ha)*
	Papilionidae	*Pieridae*	*Nymphalidae*	*Lycaenidae*		
Estacion Los Tuxlas, Mexico	14 (6.6)	19 (8.9)	113 (53.3)	66 (31.1)	212	700
Finca La Selva, Costa Rica	16 (3.6)	26 (5.8)	219 (49.5)	181 (40.9)	442	1,000
Garza Cocha, Ecuador	24 (3.5)	25 (3.7)	315 (46.6)	312 (46.1)	676	500
Jatun Sacha, Ecuador	25 (3.6)	23 (3.2)	306 (43.7)	345 (49.3)	699	600
Paquitza, Manu, Peru	25 (2.9)	31 (3.6)	369 (43.3)	427 (50.1)	852	3,900
Tambopata, Peru	25 (3.1)	26 (3.3)	337 (42.3)	409 (51.3)	797	>2000
Cacaulandia, Brazil	30 (2.5)	31 (2.6)	423 (35.3)	715 (59.6)	1199	2,000
Alto Jurua, Brazil	38 (3.7)	37 (3.6)	467 (45.0)	496 (47.8)	1038	>500
Cerro do Japi, Brazil	19 (4.7)	36 (8.9)	193 (47.8)	156 (38.6)	404	>400

[a]The percentage of the total species at a particular site is in parentheses.

Table 4 Variation in the contribution of each family to species richness among different faunas[a]

Fauna	*Family*				*Total*
	Papilionidae	*Pieridae*	*Nymphalidae*	*Lycaenidae*	
All of Africa	80 (2.9)	145 (5.3)	1107 (40.6)	1397 (51.2)	2729
Zaire	48 (3.7)	100 (7.6)	607 (46.5)	551 (42.2)	1306
Kenya	27 (3.7)	87 (12.1)	335 (46.5)	271 (37.6)	720
Southern Africa	17 (2.3)	54 (7.2)	265 (35.6)	409 (54.9)	745
Madagascar	13 (4.9)	28 (10.7)	175 (66.8)	46 (17.5)	262
Australia	18 (6.5)	35 (12.6)	85 (30.6)	140 (50.3)	278
New Guinea	41 (5.2)	146 (18.6)	222 (28.3)	376 (47.9)	785
Malaysia	44 (5.8)	44 (5.8)	273 (35.9)	400 (52.5)	761
Costa Rica	42 (4.0)	71 (6.8)	438 (41.9)	493 (47.2)	1044

[a]The percentage of total faunal richness by each family is given in parentheses.

Table 5 Percentage of world species by geographical region[a]

	Region						
	North America	*Neotropics*	*Costa Rica*	*Africa*	*Australia*	*New Guinea*	*Maylasia*
Total species	471	5341	1044	2729	278	785	761
% World fauna	3.4	38.8	7.6	19.8	2.0	5.7	5.5

[a]Total number of butterfly species for world fauna (13, 753.5) and the neotropical region (5341) represent averages.

because no butterfly is known to fly in a prairie during mid-winter. Here, it is easy to see that time of sampling is important for estimating species diversity.

The problem of sampling in space and time becomes more important in forest areas, especially tropical ones. Consider a recent study of fruit-feeding nymphalid butterflies conducted on 600 ha of lowland Ecuadorian rain forest that investigated how species were distributed in space and time. To ensure equal sampling, butterflies were simultaneously trapped in the canopy and understory for the first week of every month for a period of 1 year. The results showed that both time of sampling (month) and position of trap (canopy or understory) were extremely important in estimating species diversity at this site (**Figure 4**). In other words, had the study sampled for only a few months, or only in the understory, species diversity would have been greatly underestimated. A similar study from a different site in Ecuador also demonstrated that a significant proportion of species diversity was accounted for by time of sampling and the forest canopy fauna. This study further showed that species diversity varied significantly over short distances within the same, seemingly uniform rain forest, and it highlights the importance of sample size in comparisons of species diversity (**Figure 5**). Such studies illustrate the critical nature of sampling methods, space, time, and sample sizes in comparisons of butterfly diversity.

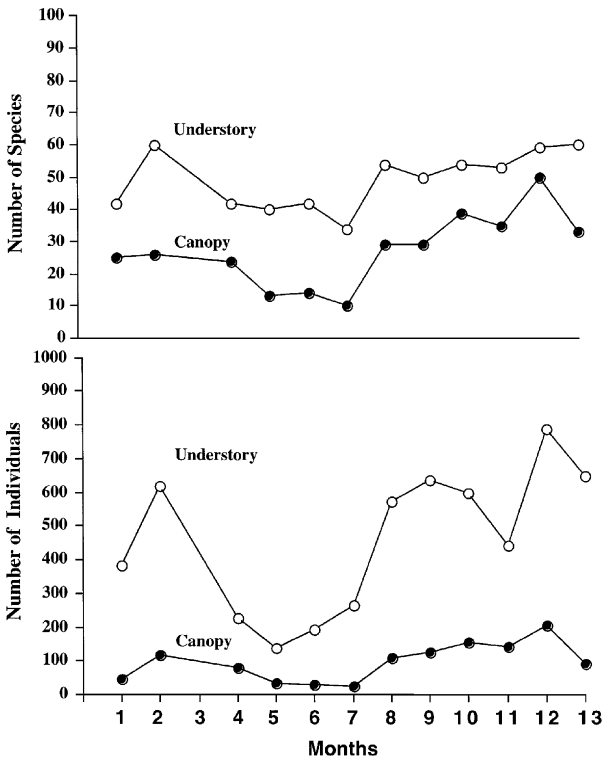


Figure 4 Monthly variation in species richness (top) and abundance (bottom) of fruit-feeding Nymphalidae in forest canopy and understory (reproduced from DeVries PJ (1997) *The Butterflies of Costa Rica and Their Natural History. II. Riodinidae*. Princeton, NJ: Princeton Univ. Press).

Butterfly Diversity and Habitat Destruction

It is obvious that the evolution of butterfly diversity is based on historical and contemporary interactions with many

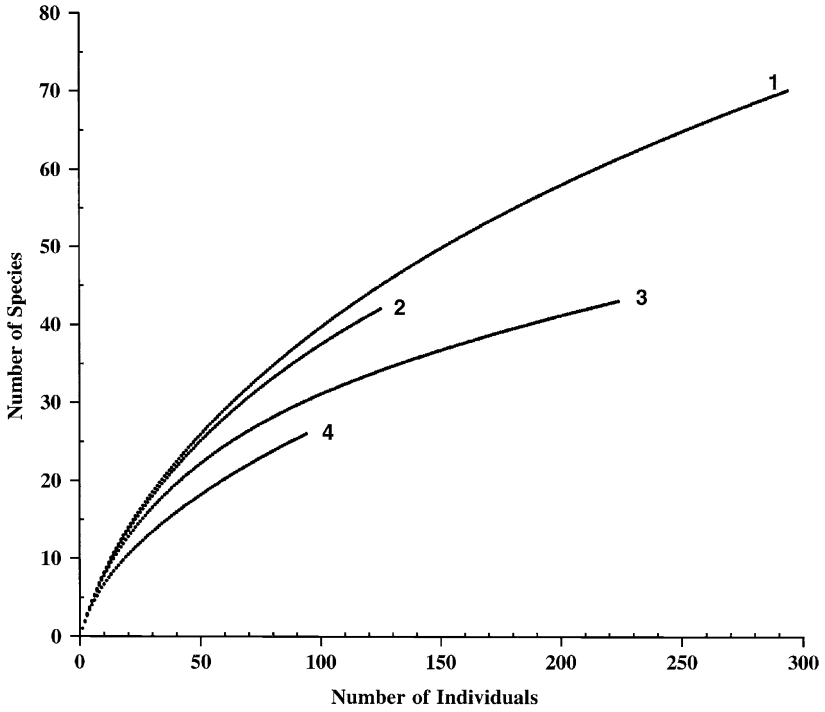


Figure 5 Species diversity of fruit-feeding Nymphalidae showing the relationship to space and sample-size. Each numbered line represents a plot sampled from within a single tract of continuous rainforest (reproduced from DeVries PJ, Walla TR, and Greeney H (1999) Species diversity in spatial and temporal dimensions of fruit-feeding butterflies from two Ecuadorian rainforests. *Biol. J. Linnean Soc.* 68: 333–353).

species. These biological interactions include plant and/or insect hosts, co-mimics in Batesian and Müllerian mimicry complexes, predators, and parasites. Butterflies have also evolved within and adapted to a great many biomes, habitats, and microhabitats, ranging from the multilevels within lush tropical rain forests to starkly dry deserts and subarctic tundra.

Habitat destruction always has profound effects on the biological communities that inhabit them, and butterflies are no exception. Like all organisms, butterflies live, evolve, and diversify within dynamic biological systems, and as such they cannot be studied as art objects or protected as inventoried stock. To date, butterflies have served as tools for understanding the diversification of life on Earth and the fundamental interactions among species. However, our future understanding of butterfly diversity will depend on a renewed interest in studying them in the natural world and valuing the habitats in which they occur.

Acknowledgments

The author is indebted to P. Ackery, G. Austin, G. Beccaloni, S. Collins, H. Greeney, A. Heath, R. Hill, R. Lande, T. Larsen, D. Lees, J. Mallet, N. Pierce, K. Rozema, T. Walla, and especially C. M. Penz for providing information, encouragement, assistance, and boundless enthusiasm about butterfly diversity.

See also: Hymenoptera. Insects, Overview. Moths. Species Coexistence

References

Ackery PR (1984) Systematic and faunistic studies on butterflies. *Symp. R. Ent. Soc.* 11: 1–21.

Ackery PR (1988) Host plants and classification: A review of nymphalid butterflies. *Biol. J. Linnean Soc.* 33: 95–203.

Ackery PR and Vane-Wright RI (1984) *Milkweed Butterflies: Their Cladistics and Biology.* London: British Museum (Natural History).

Beccaloni GW (1997a) Vertical stratification of ithomiine butterfly (Nymphalidae: Ithomiinae) mimicry complexes: The relationship between adult flight height and larval host-plant height. *Biol. J. Linnean Soc.* 62: 313–341.

Corbet AS, Pendlebury HM, and Eliot JN (1992) *The Butterflies of the Malay Peninsula.* Kuala Lumpur: Malayan Nature Society.

Dejong R, Vane-Wright RI, and Ackery PR (1996) The higher classification of butterflies (Lepidoptera): Problems and prospects. *Entomol. Scand* 27: 65–101.

DeVries PJ (1987) *The Butterflies of Costa Rica and Their Natural History. I. Papilionidae, Pieridae & Nymphalidae.* Princeton, NJ: Princeton Univ. Press.

DeVries PJ (1992) Singing caterpillars, ants and symbioses. *Sci. Am.* 267: 76–82.

DeVries PJ (1997) *The Butterflies of Costa Rica and Their Natural History. II. Riodinidae.* Princeton, NJ: Princeton Univ. Press.

DeVries PJ, Walla TR, and Greeney H (1999) Species diversity in spatial and temporal dimensions of fruit-feeding butterflies from two Ecuadorian rainforests. *Biol. J. Linnean Soc.* 68: 333–353.

Ehrlich PR and Raven P (1964) Butterflies and plants: A study in coevolution. *Evolution* 18: 596–604.

Larsen T (1991) *Butterflies of Kenya and Their Natural History.* Oxford: Oxford Univ. Press.

Mallet JBL and Joron M (1999) Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. *Annu. Rev. Ecol. Syst.* 30: 201–233.

Neild AFE (1996) *The Butterflies of Venezuela. Part 1: Nymphalidae I.* London: Meridian.

Parsons M (1999) *The Butterflies of Papua New Guinea.* San Diego: Academic Press.

Pierce NE (1987) The evolution and biogeography of associations between lycaenid butterflies and ants. *Oxford Surv. Evol. Biol.* 4: 89–116.

Pierce NE (1995) Predatory and parasitic lepidoptera: Carnivores living on plants. *J. Lepidopterists Soc.* 49: 412–453.
Pringle ELL, Henning GA, and Ball JB (eds.) (1994) *Pennington's Butterflies of Southern Africa.* Capetown: Struik Winchester.
Scriber JM, Tsubaki Y, and Lederhouse RC (eds.) (1995) *Swallowtail Butterflies: Their Ecology and Evolutionary Biology.* Gainesville FL: Scientific Publishers.
Turner JRG and Speed MP (1996) Learning and memory in mimicry. I. Simulations of laboratory experiments. *Philos. Trans. R. Soc. B* 251: 1157–1170.
Tyler HA, Brow KS Jr., and Wilson KH (1994) *Swallowtail Butterflies of the Americas.* Gainesville, FL: Scientific Publishers.
Vane-Wright RI and Ackery PR (eds.) (1984) The biology of butterflies. *Symp. R. Entomol. Soc.* 11: 1–429.

C

Captive Breeding and Reintroduction

Katherine Ralls and Jonathan D Ballou, Smithsonian Conservation Biology Institute, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

This article is revision of the previous edition article by Katherine Ralls and Robin Meadows volume 1, pp 599–607, © 2001, Elsevier Inc.

Glossary

Allele One of two or more alternative forms of a gene.
Founder Wild-caught individual that contributes genetically to the captive population.
Founder genome equivalent Number of equally contributing founders that would have produced the same genetic diversity found in an existing captive population if there had been no random loss of founder alleles.
Genetic drift Variation of allele frequency from one generation to the next that occurs due to chance. Genetic drift leads to the loss of genetic variation in small populations due to the random loss of founder alleles during reproduction.
Heterozygosity Average proportion of loci that are heterozygous (have two different alleles in an individual) in a population.
Mean kinship This value, calculated for every living member of a captive population, is the average kinship between that individual and all members of the population (including itself). Typically, living founders are excluded in the calculation of mean kinships. A population's average mean kinship is the average of the mean kinships of all the individuals in the population.
Reintroduction Releasing individuals of a species into an area where that species no longer occurs in an effort to reestablish a wild population. Reintroduced individuals may be captured from a healthy wild population in another area or may be derived from a captive population if there are no healthy wild populations remaining.
Studbook List of all the living and dead individuals in a captive population that contains information on the mother, father, date of birth, location, and other topics for each individual.

General Aspects of Captive Breeding and Reintroduction Programs

Captive breeding is the only choice for species that are extinct or nearly extinct in the wild. Nearly one-fourth of mammals, 12% of birds, and almost one-third of amphibians are threatened with extinction, according to the International Union for the Conservation of Nature (IUCN) Red list of threatened species. The need for captive breeding will undoubtedly increase as loss and degradation of wildlife habitat continue as a consequence of human population growth. However, captive breeding and reintroduction programs can play only a minor conservation role in comparison to protecting and improving habitat. Owing to limited space, staff, and funds, zoos will not be able to preserve populations of all animal species likely to become extinct in the wild.

The major goal of most captive breeding programs is to develop a selfsustaining captive population. Even if a species is never reintroduced, a successful captive breeding program will supply zoos with animals to exhibit, thereby minimizing the need to collect them from the wild. Captive breeding programs also have considerable educational value because they are used to inform zoo visitors of the value of conserving biodiversity and to increase public interest in conservation issues. Animals maintained in captive breeding programs also support a variety of research programs.

Several organizations provide help with captive breeding and reintroduction efforts. The International Species Information System (ISIS) has about 825 member institutions in almost 80 countries around the world and maintains a global database on over 10,000 species of captive animals dating back to 1829. ISIS also develops and provides software to maintain and analyze data on captive animals, such as Single Population Analysis and Record Keeping System (SPARKS) and ZIMS (Zoological Information Management System). The World Association of Zoos and Aquariums (WAZA) manages 118 international studbooks for rare or endangered species or subspecies. About 1150 other populations are cooperatively managed at the regional level. In the United States alone, the Association of Zoos and Aquariums (AZA) manages 115 populations through interzoo captive breeding programs known as Species Survival Plans (SSPs). More than 300 other

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00268-9

populations are covered by less formal population management plans. Similar regional programs exist in Europe, Africa, Central America, Australia, and New Zealand. Some species may be maintained in captivity for long periods without the possibility of reintroduction. For example, Pére David's deer and the Mongolian wild horse survived in captivity many decades after their extinctions in the wild before reintroduction began.

The IUCN maintains two international committees that offer help with population management, and reintroduction: the Conservation Breeding Specialist Group and the Reintroduction Specialist Group. In the US, the AZA maintains two similar committees: the Small Population Management Advisory Group and the Reintroduction Advisory Group. Both groups provide guidelines for captive breeding and reintroduction programs and maintain other committees that focus on particular groups of species such as carnivores, primates, or parrots.

Difficulties with Captive Breeding and Reintroduction Programs

Not all species breed well in captivity. The species that are easiest to breed in captivity are those that have management requirements similar to those for domestic animals or for other species with which zoos have had considerable experience. For example, zoos suddenly faced with the challenge of breeding California condors experienced little difficulty because they had been breeding Andean condors for many years. However, species with which zoos have had little prior experience may breed very poorly at first until zoos develop appropriate husbandry techniques. Thus, because husbandry techniques tend to be species specific, new captive breeding programs often require substantial research programs on behavior, reproductive biology, nutrition, genetics, or disease. Research on closely related species is also often helpful.

Poor reproduction in captivity is often due to behavioral problems caused by inadequate husbandry techniques. Because different zoos often have different degrees of success in breeding a particular species, important insights can often be gained by comparing the behavior and reproductive success of individuals kept under different conditions at different zoos. More invasive research work can also make important contributions. For example, studies on black-footed ferret reproductive biology have improved captive breeding success for this species. Asa *et al.* (2011) have suggested that allowing individuals to mate with preferred partners might improve reproductive success. If true, it will be important to develop ways to allow mate choice without compromising genetic goals for captive populations.

A captive population's risk of extinction is increased by inbreeding and loss of genetic variation, and inbred animals with little genetic variation are less likely to survive when reintroduced to the wild than more outbred individuals. However, both inbreeding and loss of genetic variation are unavoidable in small, closed populations because all individuals in the population eventually become related to each other. Captive populations are routinely managed to minimize inbreeding and loss of genetic variation. Another threat to captive populations is that selective pressures in captive habitats are different than those in wild habitats; thus captive populations tend to adapt to captivity, which can make it more difficult to re-establish a population in the wild. Genetic management can reduce but not prevent adaptation to captivity. Thus, it is advisable to minimize the number of generations in captivity before animals are reintroduced when possible.

Learned behavioral traits can degenerate in captivity even more rapidly than genetic diversity. Traits that may degenerate in the captive environment include foraging skills, detection and avoidance of predators, and fear of humans. When captive-bred and wild-born individuals have been experimentally released in the same location, the captive-bred individuals have tended to survive for shorter periods due to lack of appropriate behaviors. Behavioral problems tend to be species-specific and research on how to minimize these problems and thus maximize survival of the reintroduced individuals is often needed when starting a reintroduction program.

Genetic and Demographic Management of Captive Populations

In contrast to husbandry of captive animals and reintroduction techniques, genetic and demographic management methods are similar for all captive populations (Ralls and Ballou, 1992; Ballou *et al.*, 2010; Frankham *et al.*, 2010). Genetic and demographic management of captive populations focuses on maintaining genetic diversity in order to minimize undesirable genetic changes due to selection in the captive environment, avoid deleterious effects of inbreeding depression, and maintain future options for genetic management.

Ideally, the first step in the development of a captive breeding is agreement among all concerned parties, such as agency personnel, nongovernmental conservation groups, and outside scientific advisers that such a program would benefit a particular species. Once in place, captive breeding programs have three phases. In the founding phase, the captive population is started. In the growth phase, the population rapidly increases to the final "target" population size specified by its managers. In the carrying capacity phase, the population is maintained at its target size and excess individuals may be reintroduced into the wild (**Figure 1**). Management concerns change as the captive population progresses through these phases.

The main management concerns during the founding phase are removing individuals for the captive population with minimal impact on the wild population, acquiring enough founders from the wild to achieve genetic goals, getting the species to breed reliably in captivity, and setting general goals and plans for the captive population. Ways to capture animals for the captive population with minimal impact on the wild population include removing eggs from nests, using orphaned or injured animals, and capturing dispersing juveniles. Many birds (e.g., condors) will usually lay another egg to replace an egg that has been removed, and dispersing juvenile mammals often have high mortality rates

in the wild. If the species is one that zoos do not know how to breed reliably in captivity, it is best to solve husbandry problems with only a few wild individuals or even animals of a closely related species.

Genetic goals for a captive population are usually specified in terms of the proportion of genetic variation (measured as heterozygosity) to be maintained for a specified time. A common goal is to maintain 90% of the genetic diversity of the source population for 100 years. However, some programs use other time frames. For example, the Guam rail and black-footed ferret programs are using the goal of "90% for 50 years" because of the short generation times for these species (**Table 1**) and plans for the rapid establishment of several wild populations.

Once a genetic goal has been set, population genetics theory enables calculation of the number of founders needed for the captive population (the number of wild animals that must be captured and successfully bred) and the target population size (the number of individuals that the population needs to grow to achieve its genetic goals). Planning to retain a higher proportion of genetic variation usually increases the target population size. Increasing the number of founders reduces the size of the target population needed to reach a particular goal. Twenty to thirty unrelated individuals are generally a sufficient number of founders. Unfortunately, many existing captive breeding programs were begun after it was already too late to acquire this many founders. For example, the ferret population had only 10 founders. If it had been possible to obtain 25 founders, the target population size could have been reduced from 500 to 200 individuals. Although a small number of founders reduces the probability that a captive breeding program will be successful, it does not doom it to failure. Thus, the lack of an ideal number of founders does not justify abandoning or failing to initiate a captive breeding program.

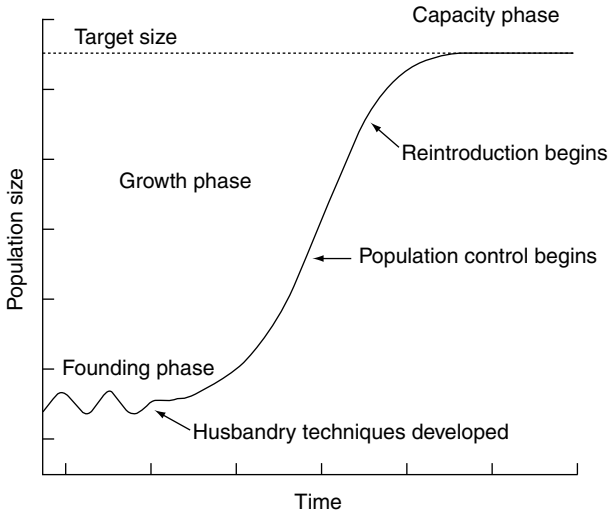


Figure 1 The development of a captive breeding and reintroduction program from the founding to the capacity phase.

The target population size also depends on the rate of species reproduction and generation length. A smaller target population will be required to reach the genetic goal if the species can grow more rapidly each generation or if it has a long generation time (because genetic variation is lost due to genetic drift each time individuals reproduce). The target population size may also be limited by practical considerations, such as the number of spaces available in zoos. Fewer zoos may be willing to participate in the program if the species is not attractive as an exhibit. Thus, the target population size may be a compromise between genetic and demographic factors and the limited resources available.

Once husbandry problems have been solved and the species is breeding well in captivity, the rest of the founders should be obtained as soon as possible. Unfortunately, the number of animals that must be captured from the wild is usually greater than the number of founders needed. Wild-caught animals may be related or fail to breed, or their descendants may fail to reproduce. For example, although 25 wild black-footed ferrets were captured, some died of distemper, some were known to be parents and offspring, and some failed to reproduce. Although ideally each founder would contribute an equal number of offspring to the captive

Table 1 The goals and number of founders of captive breeding programs with reintroduction components[a]

	Species			
	California condor[b]	*Black-footed ferret*	*Guam rail*	*Golden lion tamarin*
Heterozygosity goal (%)	90	90	90	90
Length of program (years)	200	50	50	200
Number of generations	10	20	22	33
Target population size	150	500	150	550
Number of wild-caught	14	18	21	69[e]
Number of contributing founders[c]	14	10	13	45
Founder genome equivalents[d]	8	5	5	12

[a]Species are listed in order of increasing number of generations in the program length (reproduced from Ralls and Ballou, 1992).
[b]Heterozygosity goal and program length have not been officially adopted by program managers; target population size is the minimum size of the captive population (which, with the two reintroduced populations in California and Arizona, should total no less than 450–500). Other data derived from Mace (2010).
[c]Founders with currently living descendants.
[d]The number of theoretically ideal founders taking into consideration loss of genetic diversity in the current captive population (Lacy, 1989).
[e]Includes the number of wild-caught tamarins acquired after the captive program was initiated in 1981 in addition to the number of founders and wild-caught individuals alive at the initiation of the program.

population, those ferrets that did reproduce did so unequally, skewing their genetic contributions to the captive population. Ultimately, the ferret population was founded by the theoretical genetic equivalent of only five ferrets; that is, five founder genome equivalents.

Management efforts during the growth phase center on getting the population to increase as rapidly as possible. Rapid growth has two benefits: it increases the captive population's chances of survival and it retains as much of the founders' genetic diversity as possible. Small captive populations are at higher risk of extinction due to many factors, including random demographic events (such as a succession of male births), inbreeding depression, and unpredictable events that can kill numerous individuals such as diseases, fires, hurricanes, and other catastrophes.

The standard SSP breeding strategy used in the US is designed to maximize the retention of genetic diversity. This is accomplished by minimizing mean kinship among the members of the captive population. Breeding pairs are formed based on mean kinship, beginning with the individuals with lowest mean kinships, until the desired number of pairs is attained. Efforts are also made to avoid mating closely related individuals when forming new pairs. During the growth phase, this strategy is modified slightly to choose new pairs to minimize mean kinship as much as possible but breed all individuals in the population.

Species living in groups, where individuals cannot be individually managed (i.e., fish in tanks, flocks of birds, some antelope herds), need to be managed differently. Here genetic management focuses on maintaining multiple groups of individuals, periodically transferring individuals between groups, and for some species, restricting the number of offspring any one individual can have (e.g., limiting the tenure of herd sires in an antelope herd).

Although managers attempt to minimize mean kinship and inbreeding during the growth phase, rapid population growth takes priority over genetic concerns, particularly when the population is very small and the risk of extinction outweighs the risk of a few less-than-ideal matings. For instance, if a female rejects the genetically ideal mate, she may be allowed to mate with another male she prefers.

At some point during the growth phase, the captive population usually is divided into subpopulations housed in different breeding facilities. This reduces the risk that a catastrophe such as disease or fire will decimate the entire captive population. To ensure that each subpopulation is as genetically diverse as possible, each should have individuals descended from each of the founders.

Once the population has reached the target size, relatively few offspring may be needed each year to maintain it at that level. Thus, genetic concerns become more important as managers need to determine exactly which individuals to breed in the population.

The number of offspring needed to maintain the captive population can be calculated by standard demographic techniques. Any "extra" offspring can be used for reintroduction. If there are more offspring than are needed for reintroduction, managers can prevent some adults from breeding either by

Table 2 The elements of a successful captive breeding and reintroduction program

Captive population
Ongoing research in behavior, genetics, physiology, nutrition, reproduction, and pathology
Genetic and demographic management of the population
Selfsustaining viable captive population
Sufficient size to remove animals for reintroduction without impacting the viability of the captive population
Field studies
Regular censuses of the size, distribution, and genetics of the wild population
Behavioral ecology studies (home range size, movements, habitat preferences, social organizations, mating system, feeding, and antipredator adaptions)
Locating existing suitable habitat containing critical resources for reintroduction
Habitat preservation and management
Protection of habitat from degradation and exploitation
Restoration and management of degraded habitats
Increase in or maintenance of the number of preservation areas
Conservation education for long-term support
Professional training through academic studies, workshops, internships, courses, and fellowships
Determining the most appropriate public relations and educational strategies through surveys
Public relations educational efforts using appropriate mass media (e.g., television, radio, magazines, and newspapers)
Local community education, both formal and informal
Preparation and reintroductions of animals
Choice of candidates and assessment of their characters for retrospective correlation with postrelease survival
Training in survival techniques, including foraging and feeding, antipredator tactics, locomotion, and orientation
Adaption to local conditions at release site (food, climate and temperature, and disease)
Release and long-term monitoring to evaluate causes of death and basis for survival

Source: Reproduced from Kleiman DG (1989) Reintroduction of captive mammals for conservation: Guidelines for reintroducing endangered species into the wild. *Bioscience* 393: 152–161.

using contraceptives or by housing them individually or in same sex groups.

There are two general strategies for producing the individuals to be reintroduced. If the date of a reintroduction effort can be scheduled well in advance and the species has a predictable breeding pattern (such as breeding once a year during the spring), males and females can be paired up for the specific purpose of producing excess young for that particular reintroduction. However, if the date of a reintroduction effort is difficult to predict in advance (this may occur due to difficulties with funding or permits), animals for reintroduction can be selected from the existing population and breeding pairs can be set up to replace the reintroduced individuals.

In the early stages of a reintroduction program, reintroduction techniques are still being refined and mortality may be high. Thus, initially the most genetically expendable individuals are usually released. Later, emphasis will gradually shift to choosing individuals that are not closely related to the individuals already present in the wild population. This maximizes the genetic diversity of the wild population. The final genetic goal is to make the wild population as genetically diverse as the captive population.

Reintroducing Captive-Bred Animals to the Wild

Ideally, the goals of all captive breeding plans would include reintroduction back to the wild. However, some species may be impossible to reintroduce due to lack of habitat or other problems. Furthermore, some species will be easier to reintroduce than others. The elements of a successful reintroduction program are shown in **Table 2**. These elements include research on both the captive and wild populations, habitat preservation and management, conservation education to ensure long-term support of the program, and careful management and monitoring of the reintroduced individuals (Serena, 1995; Seddon *et al.*, 2007; Griffiths and Pavajeau, 2008; Earnhardt, 2010). The IUCN and AZA have developed guidelines that discuss the biological, socioeconomic, and legal requirements for successful reintroduction.

Most important, reintroduction is a realistic goal only when habitat protection is an integral part of the species' overall conservation plan. A species should not be reintroduced unless the factors that led to its decline in the wild have been identified and eliminated – or at least greatly reduced – and suitable legally protected habitat exists. In addition, the release site should be within the species' historic range. Occasionally, however, a species must be "reintroduced" into areas of suitable habitat outside of its historic range. For example, the Guam rail was reintroduced to the nearby island of Rota because the non-native brown tree snake, which led to the bird's extinction in the wild, has invaded its entire historic range on Guam. Similarly, a variety of species, including birds, reptiles, and invertebrates, from New Zealand and Australia are being reintroduced to offshore islands that are free of the non-native predators that led to their extinction on the mainland.

There are many other factors to consider when reintroducing captive-bred individuals. For instance, the release of captive-bred animals can spread diseases to an existing wild

Table 3 Checklist for deciding whether or not appropriate conditions exist for beginning a reintroduction program applied to three species of lion tamarins

	Lion Tamarin Species		
	Golden	*Golden-headed*	*Black*
Condition of species			
1. Need to augment wild population	Yes	No	Yes(?)
2. Available stock	Yes	Yes	No
3. No jeopardy to wild population	?	?	?
Environmental conditions			
4. Causes of decline removed	?	No	No
5. Sufficient protected habitat	Yes(?)	No	Yes
6. Unsaturated habitat	Yes	Yes(?)	?
Biopolitical conditions			
7. No negative impact for locals	No	?	?
8. Community support exists	5	2	4
9. GOs/NGOs supportive/involved	Yes	Yes	Yes
10. Conformity with all laws/regulations	Yes	?	?
Biological and other resources			
11. Reintroduction technology known/in development	4	3	3
12. Knowledge of species' biology	5	1.5	3
13. Sufficient resources exist for program	Yes	No	No
Recommended reintroduction/translocation?	Yes	No	No

Some conditions were rated on a numerical scale from 0 = worst to 5 = best.

Source: Reproduced from Kleiman DG (1990) Decision-making about are introduction: Do appropriate conditions exist? *Endangered Species UPDATE* 8(1): 18–19.

population or disrupt its social organization. Thus, it is advisable to screen captive-bred individuals for diseases prior to release and to release them in habitat that has no wild individuals.

When reintroduced to the wild, captive-bred individuals are likely to suffer high mortality rates due to inappropriate behavior. For example, they may have difficulty finding enough food or fail to avoid predators. It has proven very difficult to help orphaned sea otter pups raised in captivity develop appropriate foraging skills and teach them to avoid humans. Captive-bred condors develop appropriate foraging skills fairly easily but often fail to avoid humans and human structures. Substantial research programs are often needed to develop husbandry and reintroduction techniques that will promote behaviors needed for survival in the wild.

The conditions under which captive-bred individuals are raised can be critical. The development of appropriate survival skills may require a skilled parent or a particular stimulus at some critical period during development. For example, adult ferrets prefer eating whatever they were fed when they were 2 or 3 months old, which is when ferrets develop their permanent teeth. Therefore, at the age of 2 or 3 months, captive ferrets should be fed prairie dogs, their exclusive prey in the wild. Methods of reintroduction may also require research. For example, should the animals be released as social groups or as individuals? Should they be fed after they are released and, if so, for how long? The answers to such questions depend on the particular species being reintroduced. Reintroduction programs using captive-bred individuals are usually expensive, lengthy, complex, and difficult. Thus, the decision to begin such a program should not be made lightly. A short checklist of the major factors should be considered when deciding whether or not to reintroduce a species is illustrated in **Table 3** with respect to three species of lion tamarins. Answers to the questions in the checklist indicate that reintroduction is appropriate for golden lion tamarins but not the two other species because the causes of their decline have not been eliminated and funds to support a reintroduction program are not available.

There are no generally accepted guidelines for declaring the success of a reintroduction effort. Beck and colleagues suggested two possible criteria: if 500 wild individuals survive without human support or if a formal population viability analysis predicts that the population would be selfsustaining (Beck *et al.*, 1994). Successful captive breeding and reintroduction programs require sustained long-term, adequately funded efforts. Research can solve many problems involved in successfully breeding a species in captivity and reintroducing it to the wild. However, the ultimate success of many programs, such as that for the black-footed ferret, will depend on whether or not we are able to preserve enough suitable habitat to sustain viable wild populations of the species.

See also: Breeding of Animals. Biodiversity in Plant Breeding. Conservation Efforts, Contemporary. Endangered Amphibians. Endangered Mammals. Endangered Reptiles. Zoos and Zoological Parks

References

Asa CS, Traylor-Holzer K, and Lacy RC (2011) Can conservation breeding programs be improved by incorporating mate choice? *International Zoo Yearbook* 45: 203–212.

Ballou JD, Lees C, Faust LJ, *et al.* (2010) Demographic and genetic management of captive populations. In: Kleiman DG, Thompson KV, and Kirk-Baer C (eds.) *Wild Mammals in Captivity*, 2nd edn., pp. 219–252. Chicago: University of Chicago Press.

Beck BB, Rapaport LG, Stanley Price MR, *et al.* (1994) Reintroduction of captive-born animals. In: Olney PJS (ed.) *Creative Conservation: Interactive Management of Wild and Captive Animals*, pp. 265–286. London: Chapman & Hall.

Earnhardt JM (2010) The role of captive populations in reintroduction programs. In: Kleiman DG, Thompson KV, and Kirk-Baer C (eds.) *Wild Mammals in Captivity*, 2nd edn., pp. 268–280. Chicago: University of Chicago Press.

Frankham R, Ballou JD, and Briscoe DA (2010) *Introduction to Conservation Genetics*, 2nd edn. Cambridge: Cambridge University Press. (See especially ch. 19, Genetic management of captive populations, and ch. 20, Genetic management for reintroduction.)

Griffiths RA and Pavajeau L (2008) Captive breeding, reintroduction, and the conservation of amphibians. *Conservation Biology* 22: 852–861.

Kleiman DG (1989) Reintroduction of captive mammals for conservation: Guidelines for reintroducing endangered species into the wild. *Bioscience* 393: 152–161.

Kleiman DG (1990) Decision-making about a reintroduction: Do appropriate conditions exist? *Endangered Species UPDATE* 8(1): 18–19.

Lacy RC (1989) Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zoo Biology* 8: 111–123.

Mace M (2010) *California Condor Studbook*. San Diego, California: San Diego Zoo.

Ralls K and Ballou J (1992) Managing genetic diversity in captive breeding and reintroduction programs. *Transactions of the 57th North American Wildlife and Natural Resources Conference*, pp. 263–282.

Seddon PJ, Armstrong DP, and Maloney RF (2007) Developing the science of reintroduction biology. *Conservation Biology* 21: 303–312.

Serena M (ed.) (1995) *Reintroduction Biology of Australian and New Zealand Fauna*. New South Wales, Australia: Surrey, Beatty and Norton.

Captive Breeding and the Evolutionarily Significant Unit

Robin S Waples, NOAA Fisheries, Seattle, WA, USA

Published by Elsevier Inc.

Glossary

Captive breeding Human-mediated breeding and rearing of wildlife under controlled conditions.

Distinct population segment A subspecific unit that can be afforded legal protection under the US Endangered Species Act.

Ecological exchangeability Occurs if individuals can be moved interchangeably among populations and perform similar ecological functions.

Evolutionary legacy Genetic resources that are the product of past evolutionary events and also represent the raw material for future evolution.

Hatchery A facility in which aquatic species are spawned and resulting juvenile progeny are reared for at least part of their life cycle.

Monophyly A monophyletic unit (one that includes a common ancestor and all of its descendants) is said to exhibit monophyly.

mtDNA A small, maternally inherited molecule of DNA found only in the mitochondria.

Pacific salmon Any of several species in the genus *Oncorhynchus* that always die after spawning (and hence are semelparous). Five Pacific salmon species occur in North America: Chinook, coho, sockeye, chum, and pink. Western trouts in the genus *Oncrorhynchus* are closely related to Pacific salmon; however, like the Atlantic salmon (*Salmo salar*), these species can spawn in more than 1 year. Steelhead, the anadromous for of rainbow trout (*O. mykiss*), is also discussed in this paper.

A Variety of Evolutionarily Significant Unit (ESU) Concepts

Those interested in applied conservation biology face a daunting task, given that biological diversity occurs along a continuum from individuals to populations, species and higher taxa, together with all the interactions among groups of organisms that characterize functioning ecosystems. Inevitably, triage is necessary to select from this overwhelming array of biodiversity those segments that should be the focus of conservation efforts. The concept of ESUs can help guide this process of triage. Originally intended as a means of helping to determine how to apportion the limited resources available in zoos for captive breeding (Ryder, 1986), the ESU concept has subsequently been applied much more broadly to the general problem of defining units for conservation.

Although ESUs have caught the attention of conservation biologists and evolutionary biologists, a consensus has not emerged regarding how best to identify them (for a review, see Fraser and Bernatchez, 2001). The original description of the ESU concept (Ryder, 1986) provided few details regarding implementation but did identify a broad range of types of information that should be considered, including natural history, morphology, distribution, and genetics. The first practical applications used the framework developed by Waples (1991, 2006), which has been used to guide conservation of Pacific salmon (*Oncorhynchus* spp.) under the US Endangered Species Act (ESA). Under this framework, to be considered an ESU a population or group of populations must meet two criteria: (1) it must be substantially isolated reproductively from other populations of the species and (2) it must make a substantial contribution to the evolutionary legacy of the species as a whole. Tagging data and inferences about migration based on neutral genetic markers are important sources of data for evaluating degree of reproductive isolation. The second criterion focuses on adaptive genetic differences and is trickier to evaluate because typically one has to rely on proxies (e.g., information on genetically based life history traits (which can also be influenced by environmental factors) and on physical and biotic features of the habitat (which collectively determine the pattern of natural selection experienced by the population)).

Quite a different approach was proposed by Moritz (1994), who wanted to direct conservation attention to lineages that had been isolated for long periods of time. Accordingly, he proposed that ESUs must exhibit reciprocal monophyly of mitochondrial DNA – that is, based on its mtDNA genotype, every individual within an ESU must be more closely related to other individuals within the ESU than to any individual outside the ESU. Even after populations become isolated, it can take a great deal of time before reciprocal monophyly is achieved, so this approach is effective in identifying divergent, isolated lineages; it is also easy to apply and has been widely used for a variety of organisms. However, it is a stringent criterion for assessing intraspecific diversity (many full species do not show monophyly for mtDNA), is less applicable to plants and other taxa than it is to animals, and makes no attempt to assess the importance of adaptive differences. In response to this latter (perceived) limitation, Crandall *et al.* (2000) proposed a framework based on genetic and ecological exchangeability. Rather than specifying a fixed criterion for identifying ESUs, their framework attempts to identify some important places along the spectrum of population differentiation. Those populations that are less exchangeable (replaceable) are stronger candidates to be ESUs and could be afforded higher conservation priority. Populations that are not ecologically exchangeable are likely to harbor important local adaptations. Unfortunately, ecological exchangeability is

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00265-3

difficult to assess directly, so in most cases inferences must be drawn from other types of information.

Although a single, unified theory of ESUs has not emerged, some common elements can be identified. First, two major axes of biodiversity should be considered: one related to isolation, and one related to adaptation. The published ESU concepts differ in the relative importance they accord to each of these axes, as well as in specifics of how to evaluate each axis. Second, there is a general recognition that it is not enough to identify (and hopefully conserve) populations and groups of populations (i.e., the pattern that is the product of evolution); rather, it is important also to conserve ecological and evolutionary processes that provide the context for promoting local adaptations, heterogeneity, and viable populations (Moritz, 2002).

Captive Breeding

Captive breeding has become an important component of conservation planning for many endangered species (see review by Ralls and Meadows, print version of the *Encyclopedia of Biodiversity*), and captive breeding programs now exist for a wide range of species. It is generally recognized that captive breeding is most effective when integrated into a comprehensive conservation program that addresses problems faced by the species in the wild (most frequently, these problems involve loss or degradation of habitat). Ralls and Meadows describe three phases of a typical captive breeding program: (1) a founding phase, during which managers attempt to balance the need to use an adequate number of founders against risks such collections would pose to the remaining wild population; (2) a growth phase, during which captive abundance is increased as fast as reasonably possible; and (3) a capacity phase, during which the captive population has approached carrying capacity and a portion can be used for reintroduction into the wild.

The goals of captive breeding programs are diverse and often include both short- and long-term objectives. The short-term objective is generally to maintain the gene pool, with minimal losses of diversity, while efforts are made to address the root causes for the species' decline. Although a new species often presents unexpected challenges for successful husbandry, the empirical record demonstrates that many captive breeding programs have been successful in achieving their short-term objectives, over periods of a few decades. Ideally, the long-term goal includes re-establishing a self-sustaining population in the wild. This is much more difficult to achieve, for two major reasons. For some species there are no realistic prospects for securing sufficient habitat to support reintroductions in the foreseeable future; therefore, simply maintaining the gene pool in captivity becomes a long-term objective. Second, even when habitat conditions are more favorable for reintroduction attempts, they might not be successful. Environments experienced by captive populations differ dramatically from those experienced in the wild, and as a consequence captive populations are subject to behavioral and genetic changes (domestication) that can compromise their ability to survive and reproduce in the wild. Although encouraging progress is being made to resolve some of these complications, too few empirical examples exist to provide a sound basis for predicting the likelihood of long-term success of any given reintroduction effort.

Captive breeding of fish and other aquatic populations has been conducted for thousands of years. Although some recent captive programs have conservation objectives, historically these efforts were designed either to provide a source of food (as in domesticated livestock) or to enhance harvest opportunities. To achieve the latter objective, it is necessary to release captively reared individuals into the wild. This unusual feature of aquatic species – that large numbers of captive individuals are released into the wild where they can interact with wild populations (Lorenzen *et al.*, 2010) – presents some unusual challenges in defining conservation units.

An Example from Pacific Salmon

Pacific salmon have an unusual life cycle (**Figure 1**). Adults spawn in fresh water; the eggs incubate buried in gravel in streams or lakes; juveniles spend a few days to a year or more in fresh water before migrating to sea as smolts; subadults grow rapidly for 1 or more years in the ocean before maturing and returning to fresh water (generally to their natal stream) to spawn and die.

Major anthropogenic changes to salmon habitats in the Pacific Northwest and California started before the US Civil War and have continued to the present day. As a result of a series of status evaluations conducted in the 1990s, about half of the remaining salmon populations in the contiguous US are now listed as threatened or endangered "species" under the US ESA. These listings have taken advantage of a provision of the ESA that allows listing of "distinct population segments" (DPSs) of vertebrates – a provision that has been used to list US populations of alligators, grizzly bears, and bald eagles, among others. The ESU framework developed by Waples (1991) has guided the Pacific salmon listing determinations; groups of populations have been considered to be DPSs if they met the two criteria to be ESUs (reproductive isolation and

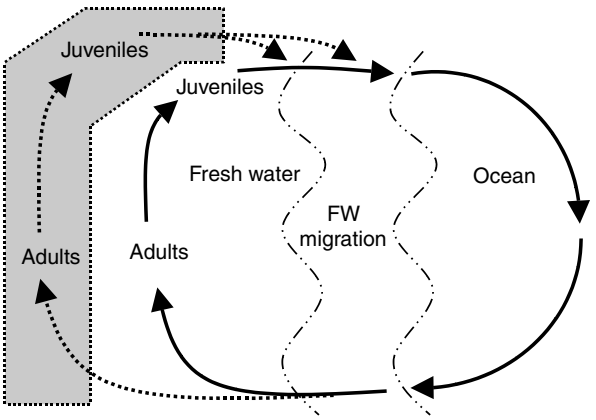


Figure 1 Life cycle of Pacific salmon. Solid lines with arrows indicate the life cycle in the wild; dashed lines with arrows and shaded polygon indicate life cycle in captivity. Most Pacific salmon hatcheries release juveniles near the hatchery when they are ready to undergo their downstream migration to the sea; some hatcheries transport juveniles and release them near tidewater.

contribution to evolutionary legacy). Most Pacific salmon ESUs include quite a few (typically 20–30) separate populations or stocks that differ somewhat in ecology, life history, and genetics, but differences among populations within ESUs are typically much less than those among ESUs.

Captive propagation also has a long history with Pacific salmon; the first salmon hatcheries on the West Coast were built in the 1870s. By far the most common form of propagation is known as sea ranching, where adults are spawned in captivity and their progeny reared as juveniles until they are released into the wild. Today, hundreds of facilities around the Pacific Rim release several billion juvenile Pacific salmon into the wild each year. Historically, the objective of most of these programs has been to provide more adult salmon for harvest than can be produced by wild populations alone. More recently, some hatchery programs have developed conservation goals, and some have dual goals of fishery enhancement and conservation. The scale of these programs has raised considerable concern, as well as considerable uncertainty, about long-term effects on natural populations (Fraser, 2008; Araki and Schmid, 2009).

As natural salmon populations became listed as threatened or endangered, it became important to determine what, if any, legal protections hatchery fish should receive under the ESA. To address this need, NMFS developed a two-step process. In the first step, biologists reviewed stock histories and other pertinent information for each hatchery population associated with a listed salmon ESU and determined which hatchery populations are part of the ESU. **Figure 2** illustrates how ESU determinations were made for hatchery populations using the lineage-based ESU framework of Waples (1991). Stock histories and genetic analyses were used to determine which population/ESU the hatchery population in question was derived from, and then all available information was considered to determine whether the hatchery population had diverged so much from its natural progenitors that it no longer should be considered part of the same ESU. In most cases, absent information for substantial divergence or mixing with other populations, hatchery populations recently derived from a natural population were considered to be part of the ESU. The second step involved an administrative decision whether to list (and therefore afford legal protection to) the hatchery populations that are part of listed ESUs. In general, such populations were not listed because doing so would provide little or no additional conservation benefit but would create considerable administrative and bureaucratic burdens for the general public as well as the agency. Exceptions were made for hatchery populations considered "essential" for recovery; such populations were listed to ensure that the remnant gene pool received the necessary legal protections to ensure its persistence. A key element of this approach was that hatchery populations that were included in the ESU were only viewed as potential contributors to eventual recovery and delisting; status evaluations of (and recovery planning for) the listed ESUs focused on demonstrating that the natural populations are self-sustaining without relying on continual influx of hatchery individuals (McElhany *et al.*, 2000).

The legal framework for this approach was called into question by a federal court ruling (*Alsea Valley Alliance v. Evans*, 2001) that held that the smallest unit that can be listed under

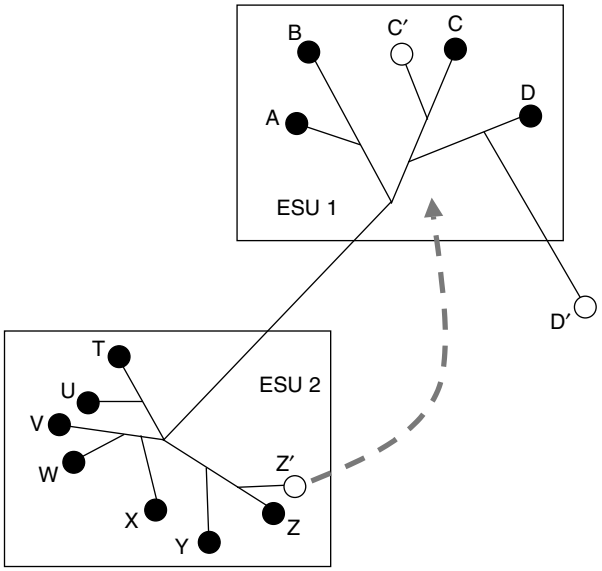


Figure 2 A lineage-based approach for evaluating evolutionary relationships between wild (black circles) and hatchery (open circles) populations in two ESUs. In ESU 1, hatchery population C′ is recently derived from wild population C and has diverged relatively little so is still considered to be part of the ESU. Hatchery population D′ was derived from wild population D in the more distant past and/or has undergone so much genetic change due to domestication or stock mixing that it is no longer considered part of the ESU. In ESU 2, the relationship between populations Z and Z′ is similar to that between C and C′ in ESU 1, so hatchery population Z′ is included in ESU 2. However, if fish from Z′ are released into the geographic area of ESU 1, they do not become part of ESU 1, even if they successfully reproduce there, because they are from a different lineage.

the ESA is a "DPS" (which, for salmon, is legally equivalent to an ESU). A consequence of this ruling is that the agency must either list or not list an entire ESU and all its component members, whether hatchery or wild. This ruling caused considerable controversy and concern because the agency was seemingly faced with a dilemma: either (1) include in the salmon listings millions of hatchery produced fish, which would greatly (and unnecessarily) increase the regulatory burden on the agency and the public; or (2) forego legal protection for entire ESUs whose wild populations were at serious risk. An influential paper (Myers *et al.*, 2004) concluded that hatchery fish should categorically be excluded from ESUs but provided few details about how this might be accomplished in a biologically sound framework. A subsequent workshop considered this topic in detail and produced a report (Hey *et al.*, 2005) that outlines a novel approach for considering ESUs and captive breeding.

Two Ways of Thinking about Captive Populations and ESUs

The salmon ESU framework (as well as all other published ESU concepts) begin by identifying evolutionary lineages – groups of individuals that have a shared ancestry. To merit separate ESU consideration, different lineages must be isolated

strongly enough and for enough time to allow evolutionarily important differences to accrue, though a process of selection and local adaptation, random genetic drift, or (more likely) a combination of the two processes. Because considerable time is required for different units to become ESUs, it follows that loss of "evolutionary significance" or "evolutionary legacy" will also, in general, be a gradual process. Of course, the process of genetic change can be accelerated by anthropogenic factors, particularly in captive breeding programs where selective pressures are very different from those experienced in the wild. Nevertheless, under the published ESU frameworks it would be difficult to conclude that a recently derived hatchery population, operated according to conservation guidelines, had diverged so much that it was no longer part of the ESU it was derived from. This would be especially true for fully integrated hatchery programs (in which wild fish are taken into the hatchery every generation to spawn and hatchery produced fish are allowed to spawn naturally every year), since in this case there is no separate hatchery population but instead a single hatchery-wild system. Within such a hatchery-wild system, all individuals share a common evolutionary lineage, differing at most by one or a few generations in the number of hatchery versus wild ancestors each has.

Hey *et al.* (2005) concluded that considering evolutionary lineages is a necessary but not always sufficient condition for defining conservation units. They argued that, in some cases at least, the current ecological context can be more important than evolutionary lineage in determining conservation status. A salmon hatchery (or any captive breeding program) represents a very different ecological context from that experienced in the wild. As illustrated in **Figure 1**, in a typical salmon hatchery adults are collected from the wild and spawned in captivity, where their progeny are reared until they are ready to smolt (i.e., undergo the physiological transition to seawater). In some cases, to minimize mortality through anthropogenically altered freshwater migration corridors, hatchery smolts are transported and released close to tidewater. This radically different ecological context, in turn, means that captive individuals experience a profoundly different suite of evolutionary processes during their life cycle. Not only does the hatchery actively intervene in the process of reproduction, it also serves as a substitute for most or all of the freshwater life cycle of the juvenile salmon. In the view of Hey *et al.* (2005), removing wild individuals from their natural ecological context disrupts the dynamic processes that have shaped evolution of the ESU, and so captive-origin individuals can be considered not to be part of the ESU. Under the framework proposed by Hey *et al.* (2005), determining ESU membership is as easy as determining whether an individual was reared for part of its life cycle in a hatchery (hence, all the hatchery populations in **Figure 2** would be considered not to be part of the ESUs). Importantly, this conclusion based on current ecological context is not a judgment about the value of hatchery fish or their potential for use in recovery of depressed wild populations. Instead, it is a statement about the role they are currently playing in their ecosystems and the evolutionary forces they are experiencing during their life cycle. In this framework, progeny of hatchery-produced fish that spawned naturally could be considered part of the ESU, provided that their parents were still part of the evolutionary lineage that characterizes the ESU.

Although the ecological-context approach to define ESUs was developed to address a practical problem in applied conservation biology involving captive breeding, the principles underlying the approach could be applied beyond hatchery-wild issues. For example, a situation somewhat analogous to the captive-wild issue arises within the species *Oncorhynchus mykiss* in considering the ESU relationships between resident (rainbow trout) and anadromous (steelhead) life history forms. Three generic types of relationship between the two forms can be identified (Good *et al.*, 2005): no physical barriers to interbreeding between the two forms, either historically or presently (Case 1); resident fish isolated above a long-standing natural barrier (e.g., a waterfall) (Case 2); resident fish isolated above a recent, man-made barrier (e.g., a dam) (Case 3). Under the lineage-based ESU concept, Case 1 resident fish are considered to be part of the anadromous ESU, Case 2 resident fish are generally considered not to be part of the ESU, and Case 3 situations are evaluated individually (in a way that is analogous to evaluation of individual hatchery populations) to assess the degree of divergence between the life history forms. As outlined by Hey *et al.* (2005), application of the ecological context approach would lead to the same conclusions as above for Cases 1 and 2, but Case 3 resident fish would be excluded from ESUs because human intervention had disrupted natural evolutionary processes and placed the resident fish in a novel ecological context.

Each of these two frameworks has certain advantages and disadvantages. The lineage-based approach is consistent with most published thinking regarding ESU delineations, and it also provides a framework for legal protection of endangered gene pools that exist primarily or entirely in captivity. A disadvantage is that the decision as to how divergent hatchery populations must be before they no longer are part of the ESU is necessarily somewhat arbitrary, and once a standard is established it can be difficult to determine whether the divergence of any particular population meets or exceeds this threshold. A significant advantage of the ecological context approach is its simplicity: To determine the ESU status of an individual, it is only necessary to know whether it was born in a hatchery or has spent a significant amount of time in captivity. Results of applying this framework are also quite predictable, which can be a boon to conservation management and planning. A potential disadvantage is that application of this framework would commonly result in placing parents and offspring in different ESUs – a curious result for a concept based on evolutionary principles.

In 2005, NMFS revised its hatchery listing policy to be consistent with the Alsea decision (NMFS, 2005a) and updated the ESA listing status of Pacific salmon DPSs, based on the new policy (NMFS, 2005b). In one case, although assessments concluded that the wild populations qualified as endangered, the DPS as a whole was listed as threatened because of relatively abundant hatchery fish that were considered part of the DPS. This determination thus followed the stipulation of *Alsea Valley Alliance v. Evans* that hatchery fish must be considered in ESA listing determinations, but assigned a higher priority to the status of wild populations.

In revising the ESA steelhead listings, NMFS had to deal with the relationship of resident and anadromous *O. mykiss*. In previous listing determinations, NMFS had concluded that

all Case 1 resident fish (and some Case 3 fish) were part of the steelhead ESUs; however, the US Fish and Wildlife Service (USFWS), which has ESA jurisdiction over terrestrial and freshwater species, refused to list the resident fish. As a consequence, the previously listed steelhead ESUs had the same legal vulnerability that *Alsea Valley Alliance v. Evans* identified for hatchery fish: resident fish were in the ESUs (and hence part of ESA-recognized DPSs) but were not listed. In finalizing the steelhead listing determinations, NMFS adopted an extreme version of the ecological-context approach, concluding that in general resident rainbow trout and steelhead are in different DPSs, even when they interbreed (NMFS, 2006). A key factor in this approach was the decision to consider steelhead under the joint USFWS–NMFS policy for defining DPSs of vertebrate species (USFWS and NMFS, 1996), rather than under the Pacific salmon ESU policy. The joint policy has a similar two-part test for identifying DPSs but does not use the term "ESU" and potentially provides more flexibility in making DPS determinations. Furthermore, as the USFWS and NMFS share ESA jurisdiction for *O. mykiss*, a logical argument can be made for using the joint (1996) policy for this species.

Subsequent court decisions have supported these listing determinations involving hatchery fish and resident fish. In Trout Unlimited v Lohn (2009), the court held that it was reasonable to focus primarily (but not entirely) on the status of wild populations in determining whether a DPS is threatened or endangered. In Modesto Irrigation District vs Gutierrez (2010), the court found that the ecological-context approach for placing steelhead and rainbow trout in separate DPSs was a reasonable interpretation of existing statutes and the joint vertebrate DPS policy.

Conclusions

Collectively, these two frameworks provide complementary ways of thinking about captive populations and ESUs. In most pristine natural systems, the two frameworks would be congruent. Ecological context to a large extent shapes the evolutionary processes that occur, and evolutionary lineages are the product of these processes. As a consequence, in most natural systems considerations of evolutionary lineages and ecological context would lead to the same conclusions about conservation units. In artificial systems, however, a disconnection often occurs between ecological context and evolutionary lineage. Individuals from the same evolutionary lineage can be abruptly placed in very different ecological contexts, resulting in major differences in the evolutionary processes they experience. Which of the two frameworks is more appropriate for any particular application involving captive populations? There is no single scientific answer to this question. Instead, the appropriate framework should be selected after careful consideration of the advantages and disadvantages of each approach, in light of the goals one is trying to accomplish.

It is also possible to envision the ecological-context approach being applied more broadly defining conservation units in natural systems that have experienced anthropogenic change. Before this happens, however, it would be necessary to consider in more detail how extreme the change in ecological context must be before the affected individuals can no longer be considered to be part of the ESU. In Pacific salmon hatcheries, direct human intervention plays such a large role in the life cycle that the distinction between hatchery and wild fish is quite clear, at least conceptually. Similarly, a dam that completely blocks migration also severs direct ecological or evolutionary interactions and creates two discrete groups of individuals. But what about a hatchery program that returns fertilized eggs to the wild to hatch on their own? Are the progeny that emerge part of the ESU, since almost their entire life cycle will be spent in the wild? What about dams that are leaky to migration and gene flow in both directions? More generally, what about anthropogenic activities that modify natural environments and hence change the ecological context of individuals and the evolutionary processes they experience? Many of these anthropogenic changes mimic natural processes. For example, habitat fragmentation, which disrupts evolutionary processes within and among populations, is a common result of anthropogenic activity but also occurs in natural systems. Should populations that suffer fragmentation due to anthropogenic causes be excluded from ESUs based on more pristine systems? Human activities have also led to the spread of many species outside their native ranges; many of these flourish in the novel environments, and this changes the ecological context and evolutionary processes experienced by native species. Under the ecological-context framework for defining ESU membership, how would such populations be considered? It is likely that scientists and policy makers will wrestle with these and other questions in the future as they continue the dialogue about how best to conserve biodiversity in human-dominated landscapes.

References

Alsea Valley Alliance v. Evans (2001) 161 F. Supp. 2d 1154 (D. Or. 2001).

Araki H and Schmid C (2009) Is hatchery stocking a help or harm? Evidence, limitations and future directions in ecological and genetic surveys. *Aquaculture* 308(supplement 1): S2–S11.

Crandall KA, Bininda-Emonds ORP, Mace GM, and Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution* 15: 290–295.

Fraser DJ (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evolutionary Applications* 1: 535–586.

Fraser DJ and Bernatchez L (2001) Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Molecular Ecology* 10: 2741–2752.

Good TP, Waples RS, and Adams P (eds.) (2005) *Updated Status of Federally Listed ESUs of West Coast Salmon and Steelhead*. Seattle, WA: U.S. Department of Commerce, NOAA Technical Memorandum. NMFS–NWFS-C66.

Hey J, *et al.* (2005) Considering life history, behavioral, and ecological complexity in defining conservation units for Pacific Salmon. *Report of an independent panel to NOAA Fisheries*. Available at (http://www.nwfsc.noaa.gov/trt/regarding_salmon_esus.pdf).

Lorenzen K, Leber KM, and Blankenship HL (2010) Responsible approach to marine stock enhancement: An update. *Reviews in Fisheries Science* 18: 189–210.

McElhany P, Ruckelshaus MH, Ford MJ, Wainwright TC, and Bjorkstedt EP (2000) *Viable Salmon Populations and the Recovery of Evolutionarily Significant Units*, p. 156. Seattle, WA: U.S. Department of Commerce. NOAA Technical Memorandum. NMFS–NWFSC-42.

Modesto Irrigation District vs Gutierrez (2010) No. 09-15214 (9th Cir 20 August 2010).

Moritz C (1994) Defining "evolutionarily significant units" for conservation. *Trends in Ecology & Evolution* 9: 373–375.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology* 51: 238–254.

Myers RA, Levin SA, Lande R, James FC, Murdoch WW, and Paine RT (2004) Hatcheries and endangered salmon. Policy Forum. *Science* 303: 1980.

NMFS (2005a) Policy on the consideration of hatchery-origin fish in Endangered Species Act listing determinations for Pacific salmon and steelhead. *Federal Register* 70(123): 37204–37216. 28 June 2005.

NMFS (2005b) Final listing determinations for 16 ESUs of West Coast Salmon, and final 4(d) protective regulations for threatened salmonid ESUs. *Federal Register* 70(123): 37160–37204. 28 June 2005.

NMFS (2006) Final listing determinations for 10 distinct population segments of West Coast Steelhead. *Federal Register* 71(3): 834–862. 5 January 2006.

Ryder OA (1986) Species conservation and systematics: The dilemma of subspecies. *Trends in Ecology & Evolution* 1: 9–10.

Trout Unlimited v Lohn (2009) Nos. 07-35623, 07-35750, 2009 WL 650534 (9th Cir 16 March 2009).

USFWS and NMFS (1996) Policy regarding the recognition of distinct vertebrate population segments under the Endangered Species Act. *Federal Register* 61(26): 4722–4725. 7 February 1996.

Waples RS (1991) Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. *Marine Fisheries Review* 53(3): 11–22.

Waples RS (2006) Distinct population segments. In: Scott JM, Goble DD, and Davis FW (eds.) *The Endangered Species Act at Thirty: Conserving Biodiversity in Human-Dominated Landscapes*, pp. 127–149. Washington, DC: Island Press.

Carbon Cycle

John Grace, University of Edinburgh, Edinburgh, UK

Published by Elsevier Inc.

Glossary

Autotrophic respiration The respiration of photosynthetic organisms.

Biomass The mass of biological material after the removal of water by oven-drying at 70–100 °C, often expressed as mass per area of ground surface.

Flux The flow of any material measured as quantity per area per unit of time.

Gigatons of carbon (Gt C) Billion (10^9) tonnes of carbon or 3.67 billion tonnes of CO_2.

Hectare (ha) 1 ha = 10^4 m^2 = 0.01 km^2

Heterotrophic respiration The respiration of all those life forms that feed on organic matter, including bacteria, fungi, and animals.

Sink A component of the carbon cycle that is a net accumulator of carbon annually, such as the ocean or the forests.

The carbon cycle is the circulation of carbon atoms between the ocean, land, and atmosphere by physical, chemical, and biological processes. Living organisms are of great importance in the carbon cycle, because life on planet Earth is comprised of carbon compounds such as proteins, carbohydrates, and lipids. Thus, the cycle is one of many *biogeochemical cycles*, whereby organisms re-use vital elements such as carbon, nitrogen, sulfur, and phosphorus.

The important natural processes that circulate carbon between the atmosphere and the surface of the planet by living organisms are photosynthesis and respiration, whereas the significant nonliving processes (physical and chemical) are dissolution of gases in water, and the weathering of rocks.

The carbon cycle may be considered over either geological time frames (millions of years), or over recent years, decades, and centuries. In this article, we focus on the recent period and we speculate about the future behavior of the carbon cycle over the coming decades.

The carbon cycle has been perturbed by human activities. Since the Industrial Revolution (1760–1830), the widespread use of fossil carbon for energy, such as coal, oil and gas, has led to a substantial increase in the CO_2 concentration in the atmosphere. Moreover, biomass carbon is returned to the atmosphere as CO_2 when forests are cleared and replaced by agriculture. There is now widespread realization among national governments and among scientists that these anthropogenic emissions of CO_2 and other trace gases are increasing the greenhouse effect and causing global warming. At the same time, it has been discovered that the land surface and the ocean absorb a substantial fraction of the fossil fuel emissions. However, these are uncertain sinks, and their uptake of CO_2 may not persist for much longer. Consequently, there has been intense research activity in this area of science, and efforts to manage the carbon cycle to limit the greenhouse effect and thus protect the climate.

Outline

The carbon cycle is changing as a result of an increasing rate of fossil fuel burning and a continuing high rate of deforestation. The Kyoto Protocol has failed to limit the rise in carbon dioxide in the atmosphere. The author outlines some of the critical processes in the carbon cycle, and the nature of and future of the natural sinks for CO_2. The author discusses how the carbon cycle might be managed to limit the rise of CO_2.

General Features

Figure 1 summarizes current knowledge of the carbon cycle, showing the size of the carbon reservoirs and the magnitude and direction of the annual exchanges between them. The carbon cycle is said to be *out of balance* because the concentration in the atmosphere is currently increasing by about 4 Gt C annum^{-1}. Records of CO_2 trapped as bubbles in old ice show that in the half-million years before the Industrial Revolution, the CO_2 concentration varied only slightly in the range 180–325 ppm in cycles that correspond to periods of glaciation. However, in the past 50 years, when precise physical measurements have been possible, the concentration has increased rapidly and inexorably from 315 ppm to the present-day (2011) value of 390 ppm (Figure 2), diverging from the pattern that is observed so clearly in the ice cores. Here, the author discusses the processes involved in the flows (fluxes) represented in Figure 1.

The Land Sink

In a steady-state world, before industrialization, it is presumed that the uptake of CO_2 from the atmosphere by the process of photosynthesis was more or less in balance with the losses from respiration and fire. If we simplify the gas exchanges by allowing glucose, $C_6H_{12}O_6$, to represent biomass, we may write an equation to express this steady state:

$$6CO_2 + 12H_2O \leftrightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O \qquad [1]$$

Here, the double arrow ↔ represents a reversible reaction. Photosynthesis proceeds from left to right. In photosynthesis, green plants absorb the sun's energy and use it to split water. The hydrogen from the water is the reducing power to turn

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00306-3

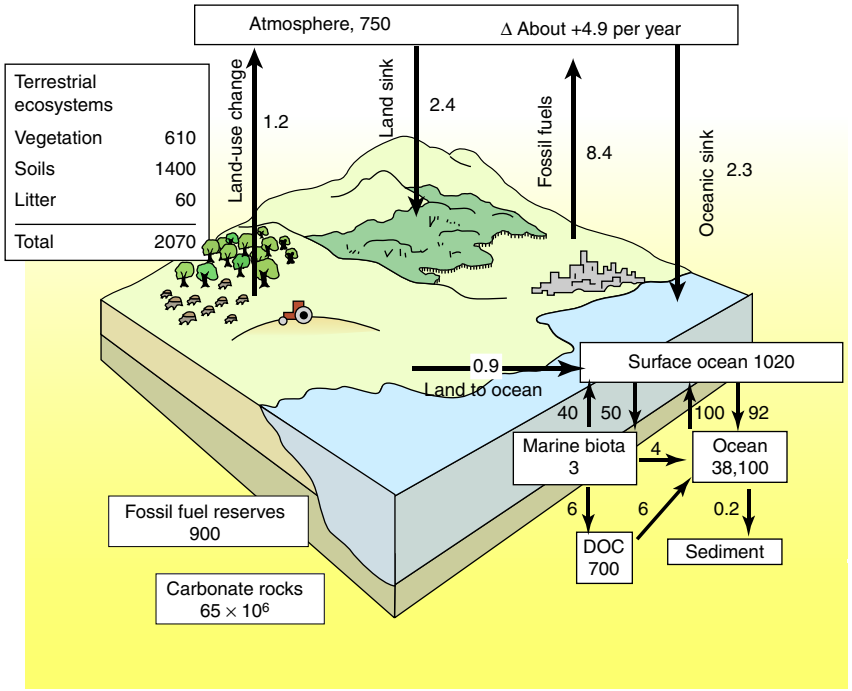


Figure 1 Global carbon cycle, updated for 2010. The term for fossil fuels includes emissions from cement manufacture. Data for emissions have the following uncertainty: fossil fuels 8.4±0.5 and land-use change 1.2±0.7. Land and ocean sinks may vary from year to year, especially the land sink. Data for stocks on the land and in the ocean are very uncertain.

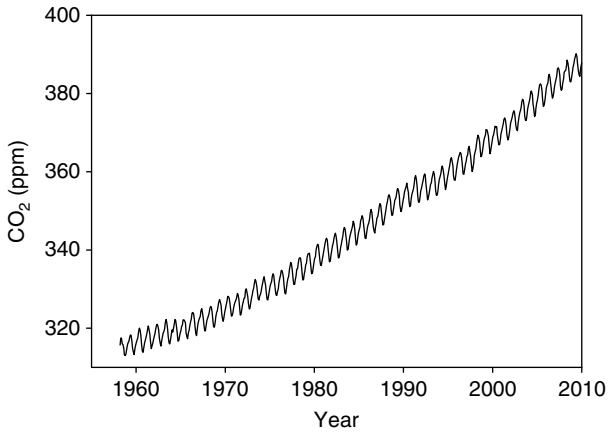


Figure 2 Carbon dioxide concentrations in parts per million, measured at Mauna Loa. Data source: Dr. Pieter Tans,NOAA/ESRL(www.esrl.noaa.gov/gmd/ccgg/trends/).

CO_2 into glucose. The oxygen escapes into the atmosphere as a byproduct, and over some 4 billion years, this is the process that gives Earth such an unusual atmosphere, with a high oxygen concentration. Likewise, respiration and fire are represented by running the equation from right to left. Equation [1] is merely a summary: in practice, photosynthesis and respiration consist of many reactions, biophysical and biochemical.

We know there is currently a land sink from various observational studies. First, application of the micrometeorological method of eddy covariance, at sites all over the world, has shown that most ecosystems, and forests in particular, are taking up more CO_2 by photosynthesis than they are losing in respiration (Luyssaert *et al.*, 2007). This sink effect seems to be most significant in tropical rain forests and in the north temperate zone. Surprisingly, even old forests have been shown to be active carbon sinks (Luyssaert *et al.*, 2008). The conclusion that the terrestrial ecosystems are a sink may also be drawn from analysis of the CO_2 concentrations in the atmosphere (Rodenbeck *et al.*, 2003): forested areas (north America and the humid tropics) tend to have lower concentrations of CO_2 in the air above them. After making due allowance for atmospheric mixing and biomass burning, it is possible to infer the sink strengths over the land and sea. Tropical and north temperate forests are the strongest sinks (Rodenbeck *et al.*, 2003). Further evidence for the existence of a terrestrial carbon sink comes from repeated inventories of the biomass in relatively undisturbed forests, both tropical and temperate (Lewis *et al.*, 2009; Luyssaert *et al.*, 2008). The biomass inventory method tells us nothing about CO_2 from the soil, but it does show that more carbon is being stored as biomass per hectare than in former times, and does provide an estimate that is broadly consistent with the other data sources.

There are a number of explanations for the imbalance in the carbon cycle. First, CO_2 is rising rapidly and it is known from many other studies that photosynthesis is usually stimulated by high CO_2 (Ellsworth *et al.*, 2004). Second, the deposition from atmosphere to land of active nitrogen (NH_4^+, NO_3^-) acts as a fertilizer and increases the productivity of vegetation (Magnani *et al.*, 2007). This active nitrogen is widely regarded as a pollutant, as it may change the species composition of ecosystems, and it may impact negatively on biodiversity because N-fertilization always favors the more

competitive plant species against others. Moreover, nitric acid is formed from NO_3^- and this contributes to the acidification of water. However, alongside these undoubted disbenefits, it does seem that nitrogen deposition has a significant benefit in terms of its effects on the land sink. 'Natural' or preindustrial N-deposition is just a few kg ha^{-1} annum^{-1} whereas industrialized parts of the world receive up to 40 kg ha^{-1} annum^{-1}. Typical agricultural applications of ammonium nitrate fertilizer, for rapid plant growth, are 80–150 kg ha^{-1} annum^{-1}. There have been many attempts to model this interplay between fertilization by CO_2 and active nitrogen (Lloyd and Farquhar 1996), although the reasons for the sharp increase in sink strength are not fully understood and still represent an active field of research.

Photosynthesis is believed to remove as much as 120 Gt C annum^{-1} from the atmosphere, but around 60 of this is "lost" again as plant (autotrophic) respiration. The difference between these two numbers represents the rate at which the world's vegetation is producing organic matter (the NPP, Net Primary Productivity). However, as soon as it is produced, it begins to be consumed by animals and microbes (known as the heterotrophs). Approximately 60 is emitted as heterotrophic respiration (the respiration of heterotrophic organisms, including microbial decomposition and animal respiration). Data from a large number of actual sites may be seen in Luyssaert *et al.* (2007).

The land sink occurs because photosynthesis exceeds the sum of autotrophic and heterotrophic respiration. Annually, this excess is now about 2.4 Gt C. This means that the terrestrial ecosystems of the world are removing about 29% of the carbon released in fossil fuel burning.

Transport from Land to Ocean

Not all the carbon captured by terrestrial systems is retained on the land. Estimates of the mass of carbon carried away are hard to obtain from the limited amount of work carried out so far, and the huge diversity of river systems to be considered. Here, the author uses a recent review paper to provide the estimate in **Figure 1**. According to Cole *et al.* (2007), inland waters receive 1.9 Gt carbon annually from the land, of which 0.2 is buried as sediment and perhaps 0.8 is returned from rivers to the atmosphere as CO_2 and CH_4. The remaining 0.9 Gt carbon enters the ocean, where some of it remains for a very long time. This fraction of plant material that has survived the journey by river is likely to contain the chemically and biologically resistant parts of plants, notably plant cuticle, pollen, and wood fibers.

The Ocean Sink

We may consider two types of ocean sink: the solubility sink and the biological pump. The solubility sink is simply the dissolution of CO_2 in water; at 10 °C, 1.16 m^3 of the gas dissolves in a cubic meter of water. In solution, there is an equilibrium mixture of carbonic acid and bicarbonate and carbonate ions:

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \leftrightarrow 2H^+ + CO_3^- \quad [2]$$

The process of dissolution is temperature sensitive, according to Henry's Law. Gases are less soluble in warm water than in cold: at 0 °C 3.4 g CO_2 dissolves in 1 kg of water, whereas at 10 °C, the corresponding value is 2.5 and at 20 °C the value is 1.7.

In those parts of the ocean where the water surface has a low concentration of dissolved CO_2 and is also cold and down-welling, the reaction proceeds strongly from left to right, i.e., CO_2 from the atmosphere dissolves in unsaturated water. Obviously, this rate will increase as the atmospheric CO_2 concentration increases. In other parts of the ocean where surface waters are warm, upwelling and supersaturated with CO_2, as in the tropics, dissolution occurs (i.e., the reaction runs from right to left).

From eqn [1], it is clear that dissolution tends to increase the acidity of the ocean. Observational data show the scale of this effect. From 1700 to the present, the northern and southern oceans have decreased in pH by 0.1 units of pH. There are fears that this may impact on the physiology of marine organisms, especially those that form carbonate shells.

Superimposed on the chemical solubility is the ecology of marine organisms. They remove dissolved carbon to synthesize $CaCO_3$ as shells in the case of invertebrates and bones in the case of vertebrates. These organisms are connected in food chains. When the organisms die, the shells and bones sink to the sea bed under the influence of gravity, thus removing carbon from the surface layers of the ocean and maintaining the concentration gradient, so that eqn [1] tends to proceed from left to right because the products (carbonates) are removed. This is referred to as the biological pump.

The ocean sink removes about 2.3 Gt C from the atmosphere each year, about 27% of the carbon released in fossil fuel burning.

Deforestation Flux

The world was once much more heavily forested than it is today. For example, England's forest cover was reduced from 90% to 20% by Roman times, and then declined to 5% by 1350, the time of the Black Death. European deforestation was rapid in the period of population growth before the Industrial Revolution. Then, land was needed for agriculture and timber was needed for dwellings and ships. Empire-building and the colonization of north America and Australia initiated further great waves of deforestation.

Deforestation in the last half-century has been in the humid tropics. Peninsular Malaysia's forest cover declined from 74% to 40% in the period 1958–1990. Currently, the deforestation rate in Indonesia and South America is a cause of great concern, especially because of the biodiversity these forests contain.

The causes of deforestation vary geographically. In Latin America and Africa, the main driver has been the planned clearance of forest to make way for agriculture, and sometimes simply the unplanned expansion of agriculture onto forested lands (Geist and Lambin, 2002). In Southeast Asia the drivers are somewhat different. There, the Dipterocarp forests contain excellent timber and have often been destroyed or degraded by logging. However, the new driver in SE Asia is clearance of

forest to make way for oil palm, *Elaeis guineensis*, a lucrative crop. Its product, palm oil, is incorporated into diesel fuel for vehicles.

Deforestation is often defined as the removal of a forest or stand of trees, where the land is thereafter converted to a nonforest use, frequently agriculture. Once an agricultural system is formed, the biomass carbon declines to a few tonnes per hectare, and the soil organic carbon also declines, although slowly. However, in addition to complete deforestation, the degradation of forest is highly significant, both for carbon and for biodiversity. In its degraded state, the structure and microclimate is changed and so is the species composition. Moreover, degraded forest is inherently more susceptible to complete deforestation.

The pressure on forests is related to a country's stage of economic development (**Figure 3**). In this framework, we trace a pattern of deforestation from the initial point where forest cover is high and human populations are low. As the population grows, forest is cleared to provide land for crops. For economic reasons, such countries encourage logging (or these days, perhaps biofuel production from oil palm plantations), which accelerates deforestation. At a later stage, when development of urban centers is more advanced, forests become scarcer and more valued so that a transition point is reached. This seems to occur when the GDP is greater than US $5000 per person. At this point, deforestation is reduced and may even be reversed.

Tropical deforestation is high and its measurement remains controversial (Achard *et al.*, 2002; Hansen *et al.*, 2010); country-scale estimates such as those oft-quoted from the Food and Agriculture Organization (FAO) are sometimes disputed and do not always agree with data from satellites. Ultimately, the best estimates will be those from satellite remote sensing, which detects the conversion from forest to nonforest, pioneered by the Brazilian PRODES project from the 1970s. Most conventional satellite data, using optical remote sensing, fail to record forest degradation produced by logging, but advances under radar remote sensing provide prospects of measuring the biomass of any forest in any weather conditions and operating by night and by day, thus achieving good imagery many times per year and detecting quite slight disturbances (Mitchard *et al.*, 2009). Recent assessments (van der Werf *et al.*, 2009, 2010) suggest that CO_2 emissions from deforestation and forest degradation (excluding peatland emissions) are now about 1.2 Gt C annum^{-1}. The global value is found by multiplying the area of forest lost by the estimate of the changes in carbon stock when the land is covered from forest to agriculture. Both the areas and the stock change have associated error in the order of 20%, and so the figure of 1.2 is uncertain.

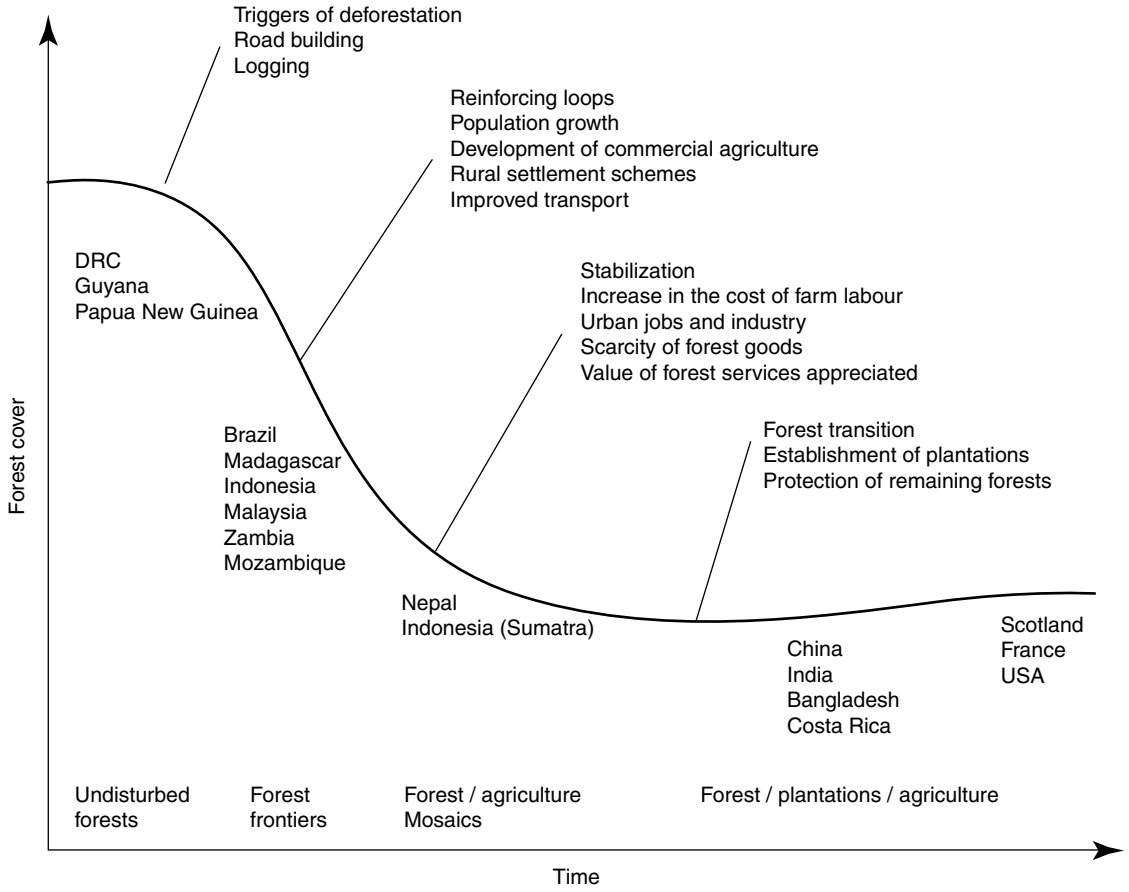


Figure 3 Hypothetic decline in forest cover associated with national development, showing where a few specific countries lie on the curve.

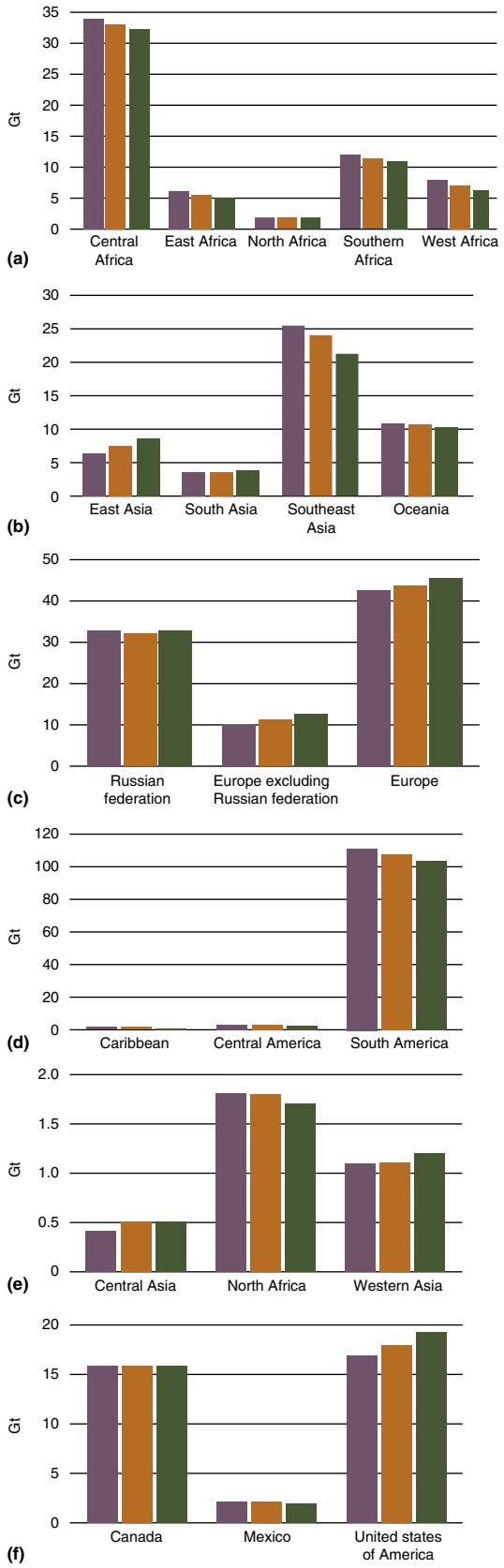


Figure 4 Regional carbon stocks in forests (Gt of carbon) in 1990 (purple), 2000 (brown) and 2010 (green). Regions are: a, Africa; b, Asia and the Pacific; c, Europe; d, South America; e, Near East; f, North America.

The FAO data set (FAO, 2011) reveals major geographical variations in both the rate of deforestation and the reasons for land-use change (**Figure 4**). The data are assembled annually from information returned by most countries of the world. It should be noted that the FAO definition (forest cover more than 10% and an area of more than 0.5 ha) is different from that which would generally be used in the remote-sensing community, and that most of the respondent countries perform forest inventories infrequently, and may thus submit estimates that will subsequently be shown to be unreliable (**Table 1**).

This 1.2 Gt C annum^{-1} from forest clearance constitutes 14% of the total anthropogenic emissions. It is sometimes said that this percentage is much higher. It used to be, but fossil fuel burning has increased very rapidly since 2000, and estimates of global deforestation rates have been reduced somewhat, partly because Brazil and Indonesia have reduced their deforestation rate.

Fossil Fuel Burning

Fossil fuel emissions have increased sharply since the year 2000 (**Figure 5**), partly because of the failure of most signatories to the Kyoto Protocol to limit their CO_2 emissions, and partly because of the extremely rapid industrial growth of China and India (Friedingstein *et al.*, 2010). The world's largest emitting country is now China, followed by the USA and then India (Peters *et al.*, 2012). The population of both China and India exceeds 1 billion, and their per capita emissions are increasing rapidly. However, they are still far behind other countries on emissions per head of population. China's per capita emissions are about 1.2 tC capita^{-1} annum^{-1} whereas Saudi Arabia, USA, Australia and Canada all exceed 4 tC capita^{-1} annum^{-1} (Peters *et al.*, 2012).

Known stocks of fossil fuels on a world scale are enough to sustain current emissions for about a century (**Figure 1**). The prices of coal, oil, and gas are likely to rise abruptly as stocks dwindle, although fresh supplies may be found as exploration continues. Gas and oil reserves are substantially less than those of coal. Very large stocks of coal remain in Australia, China, and Russia. Added to the conventional fossil fuels (coal, oil, gas) we should not overlook other possible supplies – tar sands and methane clathrate.

Tar sand is a mixture of clay, sand, water, and highly viscous oil, and is especially abundant in Canada and Venezuela. World stocks are thought to exceed two trillion barrels (about 170 Gt C). These stocks are dwarfed by methane clathrate. The term 'methane clathrate' refers to a family of loosely bound molecular associations of methane with water. These compounds exist in stable form in marine sediments, but when brought to normal atmospheric pressure, they decompose to release methane. Some estimates of the quantity of carbon as clathrate are as high as 5000 Gt C, but the resource is widely dispersed, unlike coal, oil and gas, and thus only a small fraction is ever likely to be extracted.

Carbon as clathrates and tar sands is not considered in the fossil fuel stocks in **Figure 1**, which is why this pool is shown as being substantially smaller than illustrated in other representations of the carbon cycle (IPCC 2007).

Table 1 Deforestation rates by region assessed over two decades, according to data collected from national forestry agencies (FAO, 2011). Units are thousands of hectares

	1990	*2000*	*2010*	*% annual change 2000–2010*
Africa	749,238	708,564	674,419	−0.49
Asia Pacific[a]	733,364	726,339	740,383	+0.19
Europe[b]	989,471	998,239	1,005,001	+0.07
Latin America	978,074	932,735	890,782	−0.46
Near East	126,612	121,431	122,327	+0.07
North America	676,760	677,080	678,958	+0.03
Total	**4,168,399**	**4,085,063**	**4,032,905**	**−0.13**

[a]Includes China.
[b]Includes Russia.
Source: Data collected from national forestry agencies (FAO (2011) *The State of the World's Forests 2011*. Rome: FAO).

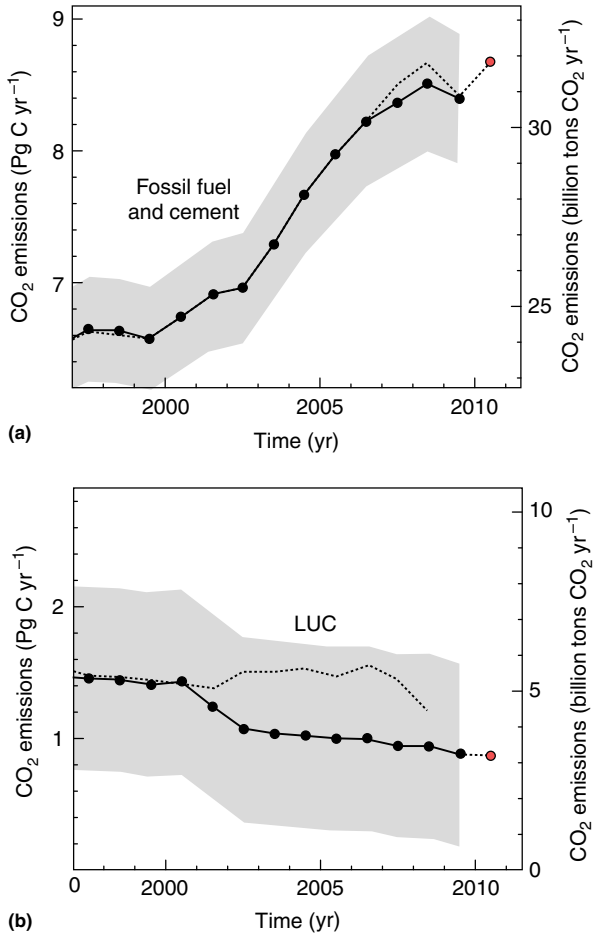


Figure 5 Trends in global emissions of CO_2 from fossil fuel burning and land use change (LUC). The solid line shows figures applicable at the end of 2010; dotted lines represent estimates made in a similar report one year earlier. The final data point for 2011 is speculative. Reproduced with permission from Friedlingstein P, Houghton RA, Marland G, *et al.* (2010) Update on CO_2 emissions. *Nature Geoscience* 3: 811–812.

Other Sources and Sinks of CO_2

If the carbon cycle were redrawn to cover the whole of the Earth's history, it would include slower processes that become highly significant when accumulated over geologic time scales.

Carbonic acid reacts slowly with minerals in several ways, as illustrated in the following examples. Such processes are collectively called *weathering*. In this case, biocarbonate reacts with feldspar to form the clay kaolinite:

$$2NaAlSi_3O_8 + 2H_2CO_3 + 9H_2O \Rightarrow$$
$$2Na^{2+} + 2HCO_3^- + 4H_4SiO_4 + Al_2Si_2O_5(OH)_4 \quad [3]$$

In the second example calcium carbonate is weathered and once more bicarbonate is produced:

$$CaCO_3 + H_2CO_3 \Rightarrow Ca^{2+} + 2HCO_3^- \quad [4]$$

In warm shallow seas and elsewhere, large deposits of calcium and magnesium carbonate were formed over geological time scales from bicarbonate, most of them having been deposited by living organisms. Many groups of microscopic and macroscopic organisms form these carbonates. In the Paleozoic (251–542 million years ago), the most abundant carbonate formers were corals, trilobites, brachiopods, bryozoans, and crinoids. Later, the dominant organisms involved in this process were coccolithophores, foraminifera, red algae, bivalves, and gastropods. Some very large deposits of carbonate rocks have uncertain origins, for example, dolomite, calcium magnesium carbonate $CaMg(CO_3)_2$, which gives its name to the Dolomite mountain range in north-east Italy.

Carbonate rock is subducted over geological time scales to become part of the oceanic crust. This carbonate may react with silica (SiO_2) to form CO_2, which is released back into the atmosphere in volcanic eruptions. In this way, the geologic carbon cycle is completed. Volcanic activity has not been constant, but over recent times, the rate of CO_2 emission estimated by scaling up data from a small number of volcanoes is very small: only 0.07 Gt C per year (Gerlach, 1991). During periods of intense vulcanism, volcanoes are capable of influencing the climate system because of the large quantities of sulfate (and other) aerosols they emit. The aerosols remain in the atmosphere for a few years; this may cause cooling, and the enhanced aerosols scatter solar radiation, thus increasing the diffuse fraction of solar radiation that reaches the land and enhancing the rate of photosynthesis.

Other Volatile Carbon Compounds (VOCs)

Methane is a greenhouse gas emitted from both natural and anthropogenic sources. The natural sources are all those places

where organic matter is microbially decomposed in the absence of oxygen, including wetlands and the guts of animals. At times, there may also be efflux from ocean to atmosphere from the decomposition of clathrates. Significant methane emissions are reported from leaves under aerobic conditions, produced by the action of sunlight on pectin, the material that binds plant cells together (McLeod *et al.*, 2008). The anthropogenic sources include fossil fuel burning, mining, burning, and livestock. The gas is 72 times more active than CO_2 in absorbing outgoing thermal radiation, when assessed over 20 years. Its lifetime in the atmosphere is only about 10 years, as it is consumed by various sinks, whereby it is converted to CO_2. The concentration in the atmosphere increased rather rapidly from 1850, when it was about 0.8 ppm, but since 1995 the growth rate has slowed, and now the concentration is beginning to level off at about 1.8 ppm.

The balance between sources and sinks of methane is still uncertain (**Table 2**). The primary sink in the atmosphere is the hydroxyl radical OH, which is formed by photochemical reactions in the atmosphere. In the soil and probably also on leaves, communities of methanotrophic bacteria consume methane and release CO_2.

Plants emit many volatile organic compounds, some of which give them their characteristic smells, well known in the perfume industry.

Table 2 Possible sources and sinks of CH_4 (Tg methane annum^{-1}) according to IPCC Fourth Assessment Report, section 7.4.1. Tg = 10^{12} g = 10^6 tonne

	Wuebbles and Hayhoe (2002)	*Range of estimates since given in IPCC (2007)*
Natural sources	145	145–260
Wetlands	100	100–231
Termites	20	20–29
Ocean	4	4–15
Hydrates	5	4–5
Geological sources	14	4–14
Wild animals		15
Wildfires	2	2–5
Anthropogenic sources	358	307–428
Energy		74–77
Coal mining	46	30–48
Gas, oil, industry	60	36–68
Landfills and waste	61	35–69
Ruminants	81	76–198
Rice agriculture	60	39–112
Biomass burning	50	14–88
C3 vegetation		27
C4 vegetation		9
Total sources	503	503–598
Imbalance		
Sinks		
Soils	30	26–34
Tropospheric OH	445	445–511
Stratospheric loss	40	30–45
Total sink	515	492–581

Isoprene (C_5H_8) is the predominant VOC emitted by many broadleaved species including poplar, oak, willow, sycamore, and eucalyptus. The emissions occur in the daylight, especially when plants are stressed, as for example by drought. Global production may be as high as 0.5 Gt C per year (Guenther *et al.*, 2006). Isoprenes are the building blocks that are responsible for the biosynthesis by plants of the familiar terpenes, which can be represented as $(C_5H_8)_n$. Isoprene and terpenes are implicated in the seeding of clouds, and so may play a role in regulating global temperatures (Meir *et al.*, 2006). When $n = 2$, the compounds are volatile and include many of the compounds that give coniferous species their characteristic smell. When n is large, the compounds are resinous and when n is very large (molecular weight 100,000 to 1,000,000) we have rubber.

There are many anthropogenic VOCs, for example, compounds that are used as solvents in paint and in cleaning fluids. Poorly tuned automobile engines emit VOCs in significant quantities. VOCs become highly reactive in sunlight, so that in cities in sunny climes such as Los Angeles and Rome we experience photochemical smog when aerosol particles, ground-level ozone and oxides of nitrogen from motor cars coexist.

Lifetime of CO_2 in the Atmosphere

Estimation of the lifetime of CO_2 in the atmosphere is a significant challenge to our understanding of the basic processes. Although photosynthesis removes CO_2 at a rate of some 120 Gt C per year from the atmospheric pool of 700 Gt C, much of this carbon is released almost immediately as autotrophic respiration. Some is stored in relatively long-lived materials such as wood and soil organic matter, and it will be released in hundreds or perhaps thousands of years for some compounds (e.g., wood trapped in bogs decomposes very slowly to release methane). Likewise, when CO_2 dissolves in the ocean, it may be released rather quickly from the surface layers, or it may be circulated to the deep abyss to reemerge a century or more later. A small fraction is incorporated into carbonates and enters a geological reservoir, which is long term but also not necessarily a permanent store. An analysis by Archer *et al.* (2009) suggests that 20–35% of any emissions would remain in the atmosphere after equilibration with the ocean (a period of 2–20 centuries). A similar argument applies to any attempt to remove CO_2 by geo-engineering; efflux from the ocean would continue for some time after CO_2 emissions are eliminated and tend to slow down any attempt to "scrub" the atmosphere.

Managing the Carbon Cycle

To avoid further global warming, it is essential to find other sources of energy. The world's energy demand increased substantially every year until 2009, when it declined, along with CO_2 emissions (**Figure 5**). The decline was the result of the economic recession and is likely to be temporary (Peters *et al.*, 2012). The challenge is to enable an increase in energy supply to countries that are still developing economically, whilst reducing or eliminating the emission of greenhouse gases

including CO_2. This seemingly impossible challenge has been approached by the United Nations and national governments in a number of ways.

Responding to concerns that human activities are increasing atmospheric concentrations of greenhouse gases such as carbon dioxide and methane to levels that threaten to cause damaging global warming, most nations of the world signed the United Nations Framework Convention on Climate Change in 1992. They promised to work toward the stabilization of emissions. Following subsequent negotiations, the Kyoto Protocol (in late 1997) committed nations to legally binding reductions in emissions of six greenhouse gases: carbon dioxide, methane, nitrous oxide, hydrofluorocarbons, perfluorocarbons, and sulfur hexafluoride. These reductions were to be counted from 1990 to the "commitment period" between 2008 and 2012, and should have limited the emissions of 37 industrialized countries and the European community. However, most of these countries have increased their emissions despite their "legally binding" pledge to reduce them. Thus, the Protocol has failed, and at the time of writing (April 2012) nothing has been agreed to replace it.

Various countries have, however, passed legislation to limit their own emissions. For example, the UK government's Climate Change Act states: "It is the duty of the Secretary of State to ensure that the net UK carbon account for the year 2050 is at least 80% lower than the 1990 baseline". Other European countries have similar legislation, passed or in preparation. However, in the USA, no such legislation exists, nor in China and India.

The issue of how to reduce emissions has been considered by the Intergovernmental Panel on Climate Change (IPCC) in its Fourth Assessment Report published in 2007. This 850-page document is freely available on the internet.

To mitigate climate change by moving from fossil- fuel-based energy to other forms of energy will be expensive and the task requires long-term planning and investment. The *Stern Review on the Economics of Climate Change* was an attempt by the UK government to analyze the problem and to outline an appropriate course of action. Most importantly, the study estimated the scale of the investment needed to limit CO_2 concentrations so as to avoid dangerous climate warming. In June 2008, Stern increased the estimate for the annual cost of achieving stabilization between 500 and 550 ppm CO_2 to 2% of the global GDP to account for faster than expected climate change (Stern, 2006). Since then, other economists have challenged the underlying assumptions made by Stern, and some place the cost as high as 10%.

Carbon Capture and Storage

Carbon Capture and Storage is a geo-engineering approach that applies to large carbon dioxide sources, such as power stations and some chemical facilities. The idea is to capture the CO_2 at the source of emission so that it cannot enter the atmosphere. This might be done by injecting the gas into suitable geological reservoirs, usually envisioned as sedimentary rocks that have previously been filled with oil or natural gas (IPCC, 2005). The engineering effort required for this is substantial, but much of the technology has already been developed by the extraction industry. Another approach would be to inject CO_2 directly into the deep ocean, where it is expected to remain for millennia. A third approach currently being investigated is to use large-scale industrial chemistry to form carbonate rocks.

The stability of the storage is an environmental concern, and also a health and safety issue. CO_2 is highly toxic to humans and other animals when its concentration exceeds 10%, as we see at locations on the Earth where CO_2 escapes from geological sources such as Lake Nyos in NW Cameroon, where 1800 people were killed in 1986 when a "bubble" of 1.6 million tonnes of CO_2 was ejected by the lake. Capture from a coal or a gas-fired power station is estimated to cost US $15–75 per tonne of CO_2 captured. This high price (compared with a few dollars per tonne for forest plantations) is partly because of the "energy penalty" associated with compressing and piping the gas to storage reservoirs. Nevertheless, by 2010, nine industrial-scale storage projects were in operation, and many more are planned. It is generally considered that CCS has the potential to sequester more CO_2 than any other single man-made activity.

Iron Fertilization of the Ocean

The productivity of many parts of the ocean is limited by the availability of iron. This has been shown by experimental addition to the ocean of soluble salts of iron and subsequent evaluation of the phytoplankton populations. Thus, it might be possible to enhance the ocean's biological pump using iron additions. Probably, one-third of the ocean is iron limited, and might be expected to sink more carbon if it could be fertilized. This prospect has attracted considerable attention, especially in the commercial sector, as substantial quantities of carbon could be captured, at least in theory. These quantities might be tradable in the carbon offset market. It is not clear, however, how well this would work, despite several major international research projects having been completed. The issues are: (i) if there is a sudden stimulation of fast-growing phytoplankton, it does not necessarily follow that other biota feeding on the plankton would increase. If not, the phytoplankton might simply die and decompose, returning much of the carbon to the atmosphere; (ii) some parts of the ocean are known to be deficient in other nutrients and iron. Silicon is sometimes deficient, essential for some phytoplankton; (iii) measurement of the amount of carbon that is indeed sequestered for a long term would be very difficult.

The iron fertilization option is still of interest, but probably not as easy as was thought in the 1990s when experimentation began (Watson *et al.*, 2008).

Managing Soil Respiration

Soil is a complex and highly variable mixture of mineral and organic material, formed naturally over thousands of years. The minerals are derived from particles of rock, altered by chemical, and physical processes. They include chemically rather active materials such as clay, which stores cations and exchanges them with the soil solution and coats particles of organic material. Much of the organic material is derived from dead plant and animal material, and is subject to biological

decay through the activities of fungi and bacteria that feed on them. Living organisms such as these produce carbon dioxide through respiration; it accumulates in the pores of the soil, eventually diffusing to the atmosphere. The other source of respired CO_2 in the soil is plant roots. Thus, CO_2 diffuses from the soil surface to the atmosphere. Its rate can be measured using sensitive instruments and is found to be between 2 and 10 μmol CO_2 m^{-2} s^{-1}, depending on the location and the environment (temperature and moisture are the main variables). Globally, the heterotrophic part of soil respiration alone is thought to be many times greater than fossil fuel emissions, but as we have seen (*see* The Land Sink), the respiratory "loss" terms are approximately balanced by the photosynthetic "gains."

Soil respiration increases more or less exponentially with temperature, but only as long as substrate is continuously added to the soil. The exponential relationship is best seen in short-term experiments (lasting hours or days). In long-term experiments (years), the enhanced rate of respiration in warmed plots is not maintained, presumably because the substrate (readily decomposable organic matter) becomes depleted.

The management of soils may considerably influence their carbon storage capacity. One of the longest experiments in soil management ever undertaken is at the Rothamsted Experimental Station in central England. Wheat has been grown in the Broadbalk field since 1843, while the land has been subject to different applications of fertilizer or manure (Gregory *et al.*, 2009). It is found that soil that has been fertilized by adding farmyard manure has three times more stored carbon than soil that has been dressed with a chemical fertilizer. The reason for this difference is that plant residues take a long time to decompose; farmyard manure contains residues of grasses, some of which are highly resistant to decay. These residues remain in the soil for decades, or in cold climates, centuries.

Another way to store carbon in soils is to bury charred biomass, known as biochar. This is mostly composed of elemental carbon (charcoal) but contains many other elements derived from the biomass, including nitrogen. Charcoal is particularly resistant to decay and remains in the soil for a very long time. It is widely believed that the dark soils found in several parts of S America (in Brazil, *Terra Preta do Indio*) were mostly created by the use of biochar by pre-Columbian Indians from 500 to 2500 years B.P.

A third way to manage soil as a carbon store is to avoid too much cultivation. When soil is disturbed by a plough, roots are cut and organic material that has been long buried is brought to the surface, where it may be more easily decomposed. Moreover, the increased drainage capacity of the cultivated soil may increase microbial growth. However, this expectation is not always borne out by field trials. West and Wilfred (2002) reviewed 67 long-term experiments on the effect of no-tillage on soil carbon storage and found cases where the results were the other way round. However, on average, the storage of carbon was increased by 57 ± 14 gC m^{-2} $annum^{-1}$. If we assume that the sites in that study were truly representative of arable land, and further assume that arable land of the world is 30 million km^2, we see that the potential exists to sequester a global total of 1.7 Gt C each year. In some soils, however, this practice would reduce the agricultural yield substantially.

Planting and Managing Forests

In temperate regions of the world, plantations of fast-growing trees such as willows and poplars are capable of accumulating carbon in biomass and soil at rates from 2 to 10 tonnes of carbon per hectare per year. In tropical regions, Eucalyptus and probably many other trees can exceed this as long as there is adequate water supply. However, generally, these are rates quoted for the early phase of growth and without subtracting carbon losses from heterotrophic respiration (mostly 'soil respiration'). Over the life time of the plantation, and taking into account carbon losses at establishment and harvesting (Kowalski *et al.*, 2004), the rate is much less. The fast-growing species, *Picea sitchensis* (Sitka spruce), has been introduced into the UK from North America and is widely grown on poor soils in wet and windy places. Its rate of carbon accumulation over a crop rotation of 40 years is 2–4 tC ha^{-1} $annum^{-1}$ even without fertilizer.

The rate of growth of trees can often be increased by adding fertilizer. In temperate and boreal regions, nitrogen generally limits growth and most tree crops in these zones respond to nitrogen fertilizer. In the tropics, nitrogen may sometimes limit growth, but in the humid tropics phosphorus is often rate-limiting. When nitrogen fertilizer is used anywhere, there is likely to be enhanced emissions of nitrous oxide from the soil. This gas has a Global Warming Potential, which is 289 times higher than carbon dioxide when assessed over 20 years, and so where the purpose of the tree crop is to reduce the rate of accumulation of greenhouse gases in the atmosphere, fertilizer should be avoided. This issue has loomed large in the debate about biofuels. In 'first-generation biofuels', it is usual to stimulate plant growth by using nitrogenous fertilizer, but heavy dosage results in emissions of N_2O from the soil. The contribution this makes to warming the atmosphere may be more than that saved from avoiding the use of fossil fuels.

There have been several estimates of the sink that could be formed by a world-wide attempt to maximize tree planting: Cannell (2003) considered that 2–5 Gt C $annum^{-1}$ was 'theoretically possible', but 0.2–1.0 Gt C $annum^{-1}$ was 'achievable'. Plantations have so far been established in many parts of the world, including the humid tropics, where the process is much more difficult (and expensive) because of the challenge of weed control.

The idea of carbon forestry supposes that trees will be planted with the primary aim of capturing and storing carbon rather than providing timber or fiber for immediate use. When this becomes the aim, new silvicultural techniques will be applied to minimize the losses at planting and harvesting, and usually the best solution will be some type of continuous cover forestry where clear-felling does not occur.

Reduced Deforestation

In the 1997 Kyoto Protocol, policies were developed to encourage reforestation and afforestation, but deforestation and degradation were excluded as it was considered not practical to measure accurately the rate of these processes with existing satellite surveillance. Later, in 2005, Papua New Guinea led a consortium known as the *Coalition of Rainforest Nations* to consider "the urgent need to take further meaningful action to

reduce emissions from deforestation and forest degradation." The various steps to achieve forest protection are collectively described as Reduced Emissions from Deforestation and Degradation (REDD). This strategy for countries to unite to protect forests requires substantial funds to be paid from international sources to the countries (and the people) who own and currently use the vulnerable rain forests. The scale of the funding can be estimated as follows: suppose the carbon losses from deforestation are presently 1.2 billion tonnes of carbon per year, as shown in **Figure 1**. This quantity of carbon on combustion will release some 4.4 billion tonnes of CO_2 into the atmosphere. In the last few years, carbon has been traded at about US $14 per tonne of CO_2, and so the relevant value (assuming that prices are unaffected by introducing a new element into carbon trading) is 61.5 billion dollars per year, which is a mere 0.083% of the global economy.

Binding international agreement to support REDD was supposed to be reached in 2009, but countries failed to agree on terms. However, a number of international and national agencies and governments have pledged substantial money for specific REDD projects, amounting to billions of dollars.

REDD is sometimes linked to the sustainable management of forest and also enhancement of forest cover, in which case the term REDD + is applied.

Barriers that have prevented the development of REDD include the problem of achieving consistent satellite remote sensing of deforestation, but since 1997, when this seemed a major issue, a great deal more research has been carried out. The limitations of optical remote sensing are better understood, and radar remote sensing is expected to overcome many of the perceived problems. Another barrier is the issue of weak governance and unstable government in some of the beneficiary countries.

Meanwhile, there has long been widespread recognition that rain forests provide a multiplicity of environmental services. These go far beyond merely carbon capture. Foremost is the realization that more than half of the world's estimated 10 million species of plants, animals and insects live in the tropical rain forests. Less widely appreciated (less visible) is the role of rain forests in water and soil protection, rainfall generation, climate regulation, and food supplies. It is not surprising that an informal carbon market developed spontaneously in the wake of the Kyoto Protocol, much of it in the forestry sector, and as a result, it is possible for companies and private individuals to purchase certified carbon credits to offset their own emissions, for example associated with air travel. Forest protection projects have been conducted in both temperate and tropical countries (Jindal *et al.*, 2008), and their successes and failures provide useful lessons for any future regulatory carbon market.

Future of the Carbon Cycle, 100-Year Perspective

The exploitation of fossil fuels is likely to continue as long as easy-to-obtain stocks of oil, gas and coal remain. As an aid to thinking of the future, and managing the carbon cycle over the next century, Pacala and Socolow (2004) introduced the carbon wedge concept (**Figure 6**). In this diagram, it is envisaged that energy demand will continue to increase, but a growing fraction of the requirement will be met by savings made by making energy utilities more efficient, by Carbon Capture and Storage, and by alternative energy sources. These measures cannot be brought in immediately. Some depend on technologies that are already developed, like nuclear energy, and others require much greater investment of research and development before they can be implemented at full scale. Thus, they are shown as a series of wedges. In Pacala and Socolow's (2004) original IPCC (2005) paper these wedges represent efficient vehicles, reduced use of vehicles, efficient buildings, efficient coal plants, CO_2 capture, replacing coal energy by nuclear energy, wind power, solar power, biomass power, reducing deforestation and soil management. Application of all of these measures (and perhaps more, as several are excluded) might stem the increase of CO_2 to 500 ± 50 ppm and thus limit global warming.

Currently, the scenario of replacement of carbon energy by wedges of efficiency and alternatives seems optimistic. It is likely that CO_2 concentrations and global temperatures will continue to rise, despite the 2009 economic recession, more or less in line with the pessimistic scenarios drawn up by IPCC (2007).

A crucially important question becomes: how will the existing sinks respond to global warming? A number of early modelling investigations suggested that quite small amounts of warming would cause carbon to be lost from the Amazon forests because a rise in temperature tends to increase respiration more than it increases photosynthesis (Grace *et al.*, 1995; Cox *et al.*, 2000). Increases in temperature are often

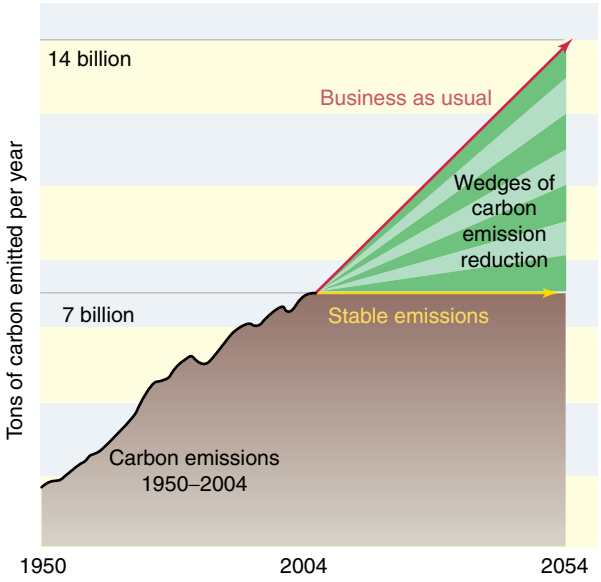


Figure 6 The Pacala–Socolow concept of stabilization wedges. In this diagram, business-as-usual would lead to a near-linear increase in CO_2 emissions to the mid-century. Emissions can, however, be stabilized by decarbonizing the energy supply and capturing carbon dioxide before it reaches the atmosphere. In this drawing, seven wedges are used to show how the gradual introduction of remedial measures might allow the business-as-usual energy demand to be satisfied while stabilizing the atmospheric CO_2 concentration.

accompanied by drought, and now we know from observational data that hot dry weather in both temperate and tropical forest is accompanied by substantial failure of the land carbon sink (Ciais *et al.*, 2005; Lewis *et al.*, 2011). Indeed, it has been known for some time that the CO_2 concentration of the atmosphere increases more in El Niño years, when drought prevails in several parts of the world (Heimann and Reichstein, 2008). Most climate models show that in future, we may expect higher variability in the climate, and some of them predict that El Niño will become more frequent and extreme. However, this remains a topic of research.

The strength of the ocean sink may also decrease in response to climate warming (Schuster & Watson, 2007). Failure of both the land and the ocean sink would mean that the atmospheric concentrations of CO_2 would rise much more rapidly in response to increasing emissions, thus leading to a positive feedback and potentially cataclysmic warming.

However, there may be negative feedbacks, as yet unknown, which act to stabilize the world's climate. In the meantime, the Precautionary Principle should be applied and the energy supply should be decarbonized as soon as possible.

See also: Biogeochemical Cycles. Climate Change and Ecology, Synergism of. Deforestation and Land Clearing. Ecosystem Services. Endangered Marine Invertebrates. Greenhouse Effect. Nitrogen, Nitrogen Cycle. Rainforest Loss and Change

References

Achard F, Eva HG, Stibig H-J, *et al.* (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297: 999–1002.

Archer D, Eby M, Brovkin V, *et al.* (2009) Atmospheric lifetime of fossil fuel carbon dioxide. *Annual Review of Earth and Planetary Sciences* 37: 117–134.

Cannell MGR (2003) Carbon sequestration and biomass energy offset: Theoretical, potential and achievable capacities globally, in Europe and the UK. *Biomass & Bioenergy* 24: 97–116.

Ciais P, Reichstein M, Viovy N, *et al.* (2005) Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* 437: 529–533.

Cole JJ, Prairie YT, Caraco NF, *et al.* (2007) Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems* 10: 171–184.

Cox PM, Betts RA, Jones CD, Spall SA, and Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* 408: 184–187.

Ellsworth DS, Reich PB, Naumburg ES, Koch GW, Kubiske ME, and Smith SD (2004) Photosynthesis, carboxylation and leaf nitrogen responses of 16 species to elevated *p*CO(2) across four free-air CO_2 enrichment experiments in forest, grassland and desert. *Global Change Biology* 10: 2121–2138.

FAO (2011) *The State of the World's Forests 2011*. Rome: FAO.

Friedlingstein P, Houghton RA, Marland G, *et al.* (2010) Update on CO_2 emissions. *Nature Geoscience* 3: 811–812.

Geist H and Lambin E (2002) Proximate causes and underlying driving forces of tropical deforestation. *Bioscience* 52: 143–150.

Gerlach TM (1991) Present-day CO_2 emissions from volcanoes. *EOS, Transactions, American Geophysical Union* 72: 249. 254–255.

Grace J, Lloyd J, McIntyre J, *et al.* (1995) Net carbon dioxide uptake by an undisturbed tropical rain forest in South-West Amazonia during 1992/3. *Science* 270: 778–780.

Gregory AS, Watts CW, Whalley WR, *et al.* (2009) The effect of long-term soil management on the physical and biological resilience of arable and grassland soils in England. *Geoderma* 153: 172–185.

Guenther A, Harley KT, Wiedinmver P, *et al.* (2006) Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmospheric Chemistry and Physics* 6: 3181–3210.

Hansen MC, Stehman SV, and Potapov PV (2010) Quantification of global gross forest cover loss. *Proceedings of the National Academy of Science* 107: 8650–8655.

Heimann M and Reichstein M (2008) Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* 451: 289–292.

IPCC (2005) *IPCC Special Report on Carbon Dioxide Capture and Storage*. Cambridge University Press.

IPCC (2007) *Climate Change 2007: Mitigation of Climate Change*. Cambridge University Press.

Jindal R, Swallow B, and Kerr J (2008) Forestry-based carbon sequestration projects in Africa: Potential benefits and challenges. *Natural Resources Forum* 32: 116–130.

Kowalski AS, Loustau D, Berbigier P, *et al.* (2004) Paired comparisons of carbon exchange between undisturbed and regenerating stands in four managed forests in Europe. *Global Change Biology* 10: 1707–1723.

Lewis SL, Brando PM, Phillips OL, Van der Heilden GMF, and Nepstead C (2011) The 2010 Amazon Drought. *Science* 331: 554.

Lewis SL, Lopez-Gonzalez G, Sonke B, *et al.* (2009) Increasing carbon storage in intact African tropical forests. *Nature* 457: 1003–1008.

Lloyd J and Farquhar GD (1996) The CO_2 dependence of photosynthesis, plant growth responses to elevated atmospheric CO_2 concentrations and their interaction with soil nutrient status. I. General principles and forest ecosystems. *Functional Ecology* 10: 4–32.

Luyssaert S, Inglima I, Jung M, *et al.* (2007) CO_2 balance of boreal, temperate, and tropical forests derived from a global database. *Global Change Biology* 13: 2509–2537.

Luyssaert S, Schulze E-D, Borner A, *et al.* (2008) Old-growth forests as global carbon sinks. *Nature* 455: 213–215.

Magnani F, Mencuccini M, Borghetti M, *et al.* (2007) The human footprint in the carbon cycle of temperate and boreal forests. *Nature* 447: 848–850.

McLeod AR, Fry SC, Loake GJ, *et al.* (2008) Ultraviolet radiation drives methane emissions from terrestrial plant pectins. *New Phytologist* 180: 124–132.

Meir P, Cox P, and Grace J (2006) The influence of terrestrial ecosystems on climate. *Trends in Ecology & Evolution* 21: 254–260.

Mitchard ETA, Saatchi SS, Woodhouse IH, *et al.* (2009) Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different African landscapes. *Geophysical Research Letters* 36: L23401.http://dx.doi.org/10.1029/2009GL040692.

Pacala S and Socolow R (2004) Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. *Science* 305: 968–972.

Peters GP, Marland G, LeQuéré C, *et al.* (2012) Rapid growth in CO_2 emissions after 2008–2009 global financial crisis. *Nature Climate Change* 2: 2–4.

Rodenbeck C, Houweling S, Gloor M, *et al.* (2003) CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics* 3: 1919–1964.

Schuster U and Watson AJ (2007) A variable and decreasing sink for atmospheric CO_2 in the North Atlantic. *Journal of Geophysical research – Atmospheres* 112: C11006.

Stern N (2006) *Stern Review on the Economics of Climate Change*. London: HM Treasury.

van der Werf GR, Morton DC, DeFries RS, *et al.* (2009) CO_2 emissions from forest loss. *Nature Geoscience* 2: 737–738.

van der Werf GR, Randerson JT, Giglo L, *et al.* (2010) Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics* 10: 11707–11735.

Watson AJ, Boyd PW, Turner SM, Jickells TD, and Liss PS (2008) Designing the next generation of ocean iron fertilization experiments. *Marine Ecology – Progress Series* 364: 303–309.

West TO and Wilfred M (2002) Post soil organic carbon sequestration by tillage and crop rotation. *A Global Data Analysis Soil Science Society of America Journal* 66: 1930–1946.

Wuebbles D and Hayhoe K (2002) Atmospheric methane and global change. *Earth Sciences Review* 57: 177–210.

Carnivores

Hans Kruuk, Institute of Terrestrial Ecology, Huntingdon, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 629–640, © 2001, Elsevier Inc.

Glossary

Basal rate of metabolism The minimum amount of energy spent by an adult animal that has not eaten recently, at normal body temperature during rest (usually sleep).
Eutherian Mammal in which the embryo is attached to the mother by a placenta.
Fissipedia Suborder of the Carnivora with divided toes.
Guild Species in a community that use similar resources.
Marsupial Mammal without placenta and with a pouch to carry the young.
Monophyletic Derived from a single ancestor.
Pinnipedia Suborder of the Carnivora with fin-like limbs (seals, sea lions, and walrus).
Plantigrade Walking on the soles of the feet.

Species Diversity

Among the mammals that are broadly referred to as Carnivores there is a weasel of approximately 45 g and a polar bear of up to 700 kg, approximately 15,000 times heavier. There are species living almost permanently in water (sea otter), in trees (palm civet), or in deserts (fennec fox); some eat buffaloes, some eat beetles, and some eat bamboo. It is not surprising, therefore, that with such large variation and divergent adaptive features there is disagreement among researchers about the evolution and classification of species. Classification is based on morphological evidence (dental characteristics, anatomy of the skull base, and other morphological features) and on molecular genetic information.

This article will deal with the approximately 230 Fissiped carnivores, i.e., the terrestrial species of the order Carnivora and excluding the seals, sea lions, and walrus (Pinnipeds). It will also discuss a group of ecologically rather similar species of marsupials in Australia belonging to the families Dasyuridae and Thylacinidae.

The Carnivora are a monophyletic order, descended from the family Miacidae approximately 60 million years ago in the Palaeocene. The order has two main branches, the dog-like Canoidea and the cat-like Feloidea. The 33 seals and sea lions and walrus belonging to the Pinnipeds are sometimes included in the order Carnivora and sometimes given separate status. They evolved from the Canoidea, but there is disagreement about which terrestrial carnivores are their closest relatives, the main candidates being the bears and the mustelids.

The Carnivora are generally divided into seven families, although some taxonomists recognize more. The Canoidea include the Canidae (dogs) with 35 species, the Mustelidae (martens) as the largest carnivore family with 67 species, the Ursidae (bears and pandas) with 9 species, and the Procyonidae (raccoons) with 15 species. The Feloidea comprise the Felidae (cats) with 35 species, the Hyaenidae with 4 species, and the Viverridae (mongooses and genets) with 66 species. The main taxonomic disagreements are over the Ursidae and the Viverridae. The two species of panda are often thought not to belong to the bears but to deserve a separate family or they may be included with the Procyonidae, and the Viverridae are often divided into two families—the Herpestidae (mongooses, 31 species) and the Viverridae (genets and civets, 35 species).

The dogs and foxes (Canidae) constitute a highly monomorphic family. All are very dog-like with nonretractile claws, and all are coursing predators, such as wolves, coyote, jackals, foxes, wild dog, bush dog, and maned wolf. All species have much in common in their ecology and social behavior, although they may vary from solitary to gregarious. Sizes vary between that of the large gray wolf (up to 80 kg) to that of the tiny fennec fox of little more than 1 kg. Canids occur on all continents, and with the dingo they even fielded an early introduction in Australia. The gray wolf is the ancestor of all domestic dogs.

Eleven canids (31%) are endangered or vulnerable, and all the large species are threatened in at least some parts of the world (the two wolves, African wild dog, and the Asian dhole).

The members of the marten family or Mustelidae occur in both the New and the Old World, with weasels, martens, mink, polecats, skunks, otters, badgers, and others. There are relatively few species in Africa, perhaps because of competition with the similar viverrids. In Britain, there are three times as many mustelid species as members of all other carnivore families combined, and in the United States there are also more mustelids than others. Almost all are quite small animals (45 g to 45 kg), with the largest species being some of the 13 species of otters. They are usually slimly built, some with retractile claws, and with a very distinctive bouncing gait, but the family is quite variable and some species are stocky, such as the badgers. Also, in feeding and social behavior they vary greatly, from solitary predators on mammals to group-living animals feeding on earthworms or fish.

Stoats and weasels have been introduced on several islands outside their normal range, and the North American mink is now an abundant exotic in many places throughout Eurasia and South America. Some of the mustelids are almost extinct (e.g., the black-footed ferret and the European mink), some were exterminated recently (sea mink), and several are in trouble (otters) or have just returned from the brink of extinction (such as the largest of them all, the sea otter).

However, only seven species (10%) are listed as endangered or vulnerable.

The bears and pandas (Ursidae) are comparatively large and some are huge (weighing up to 700 kg), stockily built, plantigrade with nonretractile claws, and very short tails. The exception is the very aberrant red panda (a small arboreal species with a long tail), a species often classified as a procyonid. All but the polar bear are mostly vegetarian; they are solitary animals with a fairly simple social organization. The bears evolved more recently than the other carnivores and successfully colonized areas as far apart as the drift ice in the Arctic and the dense forests of the Old World and neotropics. Bears of colder regions hibernate. More than half of the species (five or 56%) are endangered or vulnerable, and all are in trouble in at least part of their geographic range.

Members of the raccoon family Procyonidae are all rather long-bodied animals, relatively small (up to 8 kg), with long, ringed tails, and they walk plantigrade with nonretractile claws. They occur naturally only in the New World. In addition to the various species of raccoon throughout the Americas, there are also other Neotropical species, such as the coatis, the kinkajou, and the cacomistle. They are carnivorous and insectivorous as well as herbivorous. Their social behavior is quite variable; most are solitary but some, for example, female coatis, live in large packs. Also, common raccoons may occupy winter dens in groups. None are classified as endangered, although some raccoon species are known from only one or two islands, on which they may be rare. In several areas of Eurasia common raccoons are now frequent as an introduced species.

The Viverridae are a large Old World family of small animals, occurring almost entirely in Africa and Asia. They are the most ''primitive'' carnivores, i.e., the closest relatives of the carnivores' original forebears, the Miacidae. The family includes many species of genet and civet, linsang, binturong, a large range of mongooses, meerkat, fossa, fanaloke, and others. Some taxonomists recognize the mongooses as a separate family—the Herpestidae (otherwise classified as a subfamily Herpestinae). The mongooses are long-bodied and look somewhat mustelid-like, with nonretractile claws; in contrast, the rest of the viverrids look and walk much more cat-like, with retractile claws. The fossa from Madagascar is unusually large (up to 20 kg), but all other Viverrids are much smaller, especially the mongooses. Most are solitary, but some mongooses and meerkats live in large packs. Most of their food is invertebrate, but they also take small mammals, etc. and occasional vegetable food.

Few viverrid species appear to be endangered (four or 6%); of these four, three are on Madagascar, an island beset by deforestation problems.

The Hyaenidae have only four species left, despite a fossil record showing large numbers in the geological past, and they occur only in the Old World. Three are relatively large species of 50 kg or more and are typical coursing animals with a dog-like build and nonretractile claws. Hyaenids are carnivorous; they often scavenge, although some also take vegetable food, and the smallest, the aardwolf, is a termite specialist. The largest species, the spotted hyena, is a gregarious hunter of large mammals, but the others are solitary. One of the four, the brown hyena, is classified as vulnerable as a species, but in many countries the others are also endangered.

Finally, there are the Felidae, the proper cats, a highly monomorphic family with many species. All are built as stalking predators, with retractile claws, and they range in size from fairly small to large, from serval and flat-headed cat and many other small species to lynx and tiger (up to more than 300 kg), various leopards, cheetah, and lion (up to 250 kg). All are carnivorous, but small species also eat many invertebrates. The social organization of all species is very similar—solitary and territorial, with the lion as the only group-living exception. Felids occur naturally on all continents except Antarctica and Australia. The wild cat *Felis lybica* has been domesticated and it has been introduced virtually everywhere in the world; it is now the most widespread carnivore.

There is concern about the conservation status of almost all cat species, especially the ones with desirable fur. Twelve (34%) are endangered or vulnerable, including all the large ones such as tiger, cheetah, jaguar, and various leopards, but not the lion.

The mammalian carnivore families outside the Carnivora proper are the Australian marsupials Dasyuridae and Thylacinidae. There are four dasyurid species that are very similar to Carnivora: three quolls or, as the Australians call them, ''native cats'' and the Tasmanian devil. They are relatively small, with the quolls similar in size and appearance to martens or mongooses and the Tasmanian devil more like a small badger in size and shape, and they are typically carnivorous, solitary, nonterritorial animals. The thylacine was the only species in its family and it was also the largest of these marsupial carnivores (comparable in size to a large jackal or coyote). It was exterminated very recently. The other, dasyurid species now have very much reduced geographical ranges and they are vulnerable.

Size and Ecology

There is a large diversity of sizes among carnivores, and the ecological implications of this variation are considerable. The basal rate of metabolism of carnivores tends to be greater than that of same-sized herbivores, and a large predator expends more energy than a small one; therefore, it has to capture more prey. However, the increase in metabolism with size is not linear, and when expressed as energy requirements per kilogram of body weight a large predator is more efficient than a small one (**Figure 1**).

The body mass of a carnivore, when combined with the group size in which it lives, is correlated with the size of its home range, although there is much variation (**Figure 2**). Much of this residual variation is due to effects of habitat and to differences in diet (predators of vertebrates having larger ranges than insectivorous ones). There is no significant correlation between body mass and gregariousness, i.e., large carnivores are no more likely to live in groups than are small ones.

Brain size of carnivores increases with body size. Irrespective of this, a meat-eating, hunting carnivore or an omnivorous species generally has a relatively larger brain than an insectivorous one. Bears have relatively large brains, whereas Viverrids and Hyaenids have relatively small ones and the others have intermediate-sized brains.

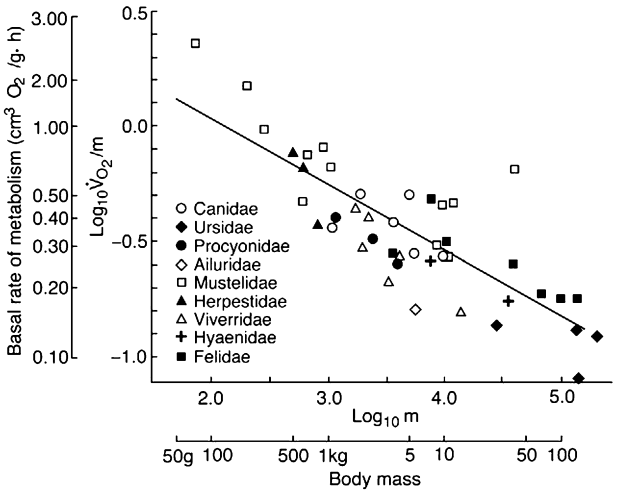


Figure 1 Basal rate of metabolism (oxygen consumption per gram per hour at rest) in carnivores as a function of body mass. Reprinted from McNab BK (1989) Basal rate of metabolism, body size, and food habits in the order Carnivora. In: Gittleman JL (ed.) *Carnivore Behavior, Ecology and Evolution*, pp. 335–354. Ithaca, NY: Cornell Univ. Press.

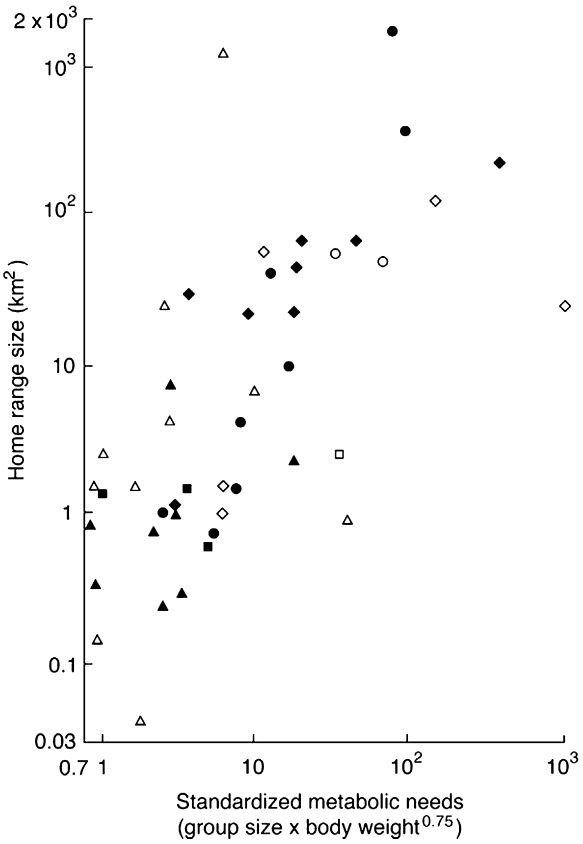


Figure 2 Home range size of carnivores as a function of metabolic needs. ●, Canidae; ○, Ursidae; ■, Procyonidae; □, Ailuridae; ▲, Viverridae; △, Mustelidae; ◆, Felidae; ◇, Hyaenidae. Reproduced from Figure 1 in Gittleman JL and Harvey PH (1982). Carvinore home-range size, metabolic needs and ecology. *Behav. Ecol. Sociobiol.* 19: 57–63.

In general, large carnivores take larger prey than do small ones, but there are important variations. First, some of the largest (bears) are mostly vegetarian. Second, the trend varies between carnivore families. Within the mustelids there is a nonsignificant negative correlation between predator body mass and prey size, so the predator–prey size principle does not apply. Both felids and canids show a strong positive correlation, but the slope of the relationship is different: The increase in prey size with predator body mass is much steeper in felids (Figure 3).

There are many specific deviations from the previously discussed general patterns, but the trends are significant. Greater energetic demands on larger carnivores are met by a dependence on larger prey, which is likely to make the large predators more vulnerable. Many of the larger carnivores are threatened globally or in parts of their range, e.g., many of the large spotted cats and the tiger, wolves, and hyenas.

Foraging and Phylogeny

The diet of carnivores in general, and that of their marsupial equivalents in Australia, consists of vertebrate prey but also invertebrates and vegetable matter. The order Carnivora shows dental specializations enabling easy digestion of vertebrates, such as large canines and the carnassial shear, but in several species the molars have been further adapted to a more grinding function when eating plants.

Foraging or hunting behavior of carnivores consists of variations on a general theme. A search leads to detection and selection of a potential victim or other source of food, and it is followed by an approach which may contain elements of stalking and/or chasing. The actual capture of prey may include seizure, immobilizing, and killing, and this is followed by eating, taking food to cubs, or sometimes quietly caching it for later consumption. Parts of this sequence may be absent, and in fact most carnivores will just search and then eat small food items without much further ado. However, even if there is a full-blooded hunt, some species never stalk (e.g., dogs or hyenas), others never chase (e.g., cats), and some may not show any specific killing behavior but just eat (e.g., hyenas).

The chain of events aimed at the capture of prey may be broken off at any stage depending on circumstances; it can

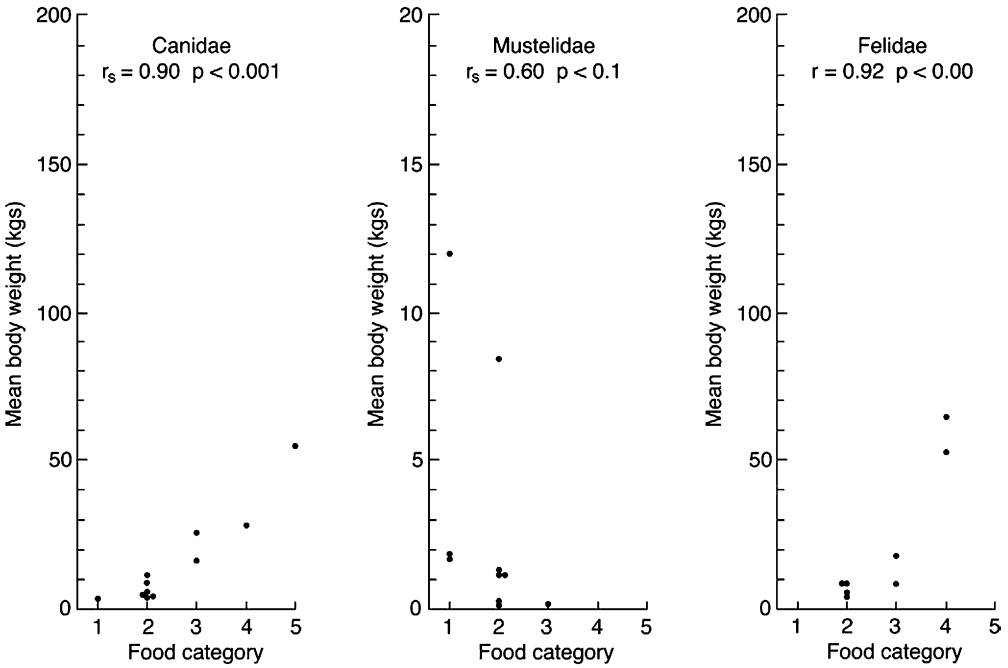


Figure 3 The relationship between size of predator and prey. For each species of carnivore, mean body weight and main food category are given. 1, invertebrates; 2, small rodents; 3, rabbits and hares; 4, larger mammals less than 50 kg; 5, mammals more than 50 kg. Reproduced from Kruuk H (1986) Interactions between Felidae and their prey species: A review. In: Miller SD and Everell DD (eds.) *Cats of the World: Biology, Conservation and Management*, pp. 353–373. Washington, D.C.: National Wildlife Federation.

also be started at any stage. It is highly adaptive, depending on prey and environment, with the predator's own motivation (its degree of hunger) apparently affecting especially the early searching stages of the hunting sequence.

Conspicuous in the predatory behavior of many carnivores is the phenomenon of surplus killing, i.e., killing more than is required for immediate consumption. Classic examples are the fox in the hen house or gull colony or a lion among a cattle herd; it has been described for hyenas, polar bears, wolves, leopard, and others. Large numbers of animals were killed without being eaten. All these situations have in common a lack of defense by the prey: The prey may be immobilized by particular weather conditions, it may be penned in, or it may have lost its antipredator defense through domestication. In these cases, the predator is sated and no longer hungry, but hunger normally affects only the early stages of the carnivore hunting sequence, especially the search. If, for some reason, no search or stalk or chase is needed because the hunter suddenly finds itself close to the quarry, then the rest of the hunting sequence is put in train irrespective of hunger, and there is no inhibition to capturing and killing. Functionally, such events are wasteful from the carnivore's point of view because they reduce prey availability without the predator getting the benefits.

Food caching (i.e., the storage of food in a hidden place) is a behavior pattern that limits this waste and utilizes the consequences of surplus killing. Many species do it, and do it in many different ways. In all canids it is highly stereotyped: A small hole is dug, and one single prey is dropped into it and covered with earth or vegetation by sweeps of the snout. In other families there are many variations on the caching theme: Leopards take a carcass high into a tree; spotted hyenas cache chunks of food in shallow water; brown and striped hyenas push it into a dense bush; stoats, mink, and other mustelids may make large stores by dragging numbers of prey into a single hole; wildcats may put remains of their quarry under a log; and some of the larger cats may civer a carcass with vegetation. The methods are consistent within the species, for canids within the family, and for martens within the subfamily, suggesting that caching has evolved in carnivores on several different occasions. There is no evidence of caching for the procyonids or viverrids, nor is it done by the marsupial carnivores.

Although food caching may use some of the surplus kills, it still does not utilize all the apparent waste. Not only are some of the caches never revisited by the perpetrator but also in the larger surplus kills only a small proportion of the victims are stored. There may be scores of dead gulls left by a fox and of gazelle left by spotted hyenas. In some carnivores caching is particularly highly developed. For instance, foxes are able to remember where they stored what, and they return preferentially to the more desirable cached items. Foxes also use some kind of bookkeeping system for their caches, leaving a drop of smelly, long-lasting urine near those caches which they have emptied.

Foraging and hunting behaviors are features that differ with the phylogeny of the predator. Canids invariably are coursing and running predators, with an occasional semistalking approach of prey, and they often forage by "sniff and search." At the other extreme, felids search almost entirely dependent on vision and they approach their quarry in a highly concealed stalk or ambush. Most members of the other families, including the marsupial carnivores, show the canid sniff and search behavior pattern. Only the genets and civets have a stalking behavior that is similar to that of the felids.

When a relatively small vertebrate prey is caught, the felids, many of the mustelids, and civets will kill it by severing the spinal cord with their canines. Canids also have a specific killing method—violently shaking the prey. Felids kill large prey with a throat bite, or they suffocate it with a bite over the nose and mouth. In almost all other predator–prey interactions there is no specific killing or immobilizing behavior—the predator just eats the prey. Most species of carnivore have highly specific ways of dealing with a prey and its carcass, and it is often possible to distinguish afterwards which predator was responsible for a kill.

Some subfamilies have evolved extreme foraging specializations. The Lutrinae (otters) dive for fish and crabs using their tactile senses, but essentially their hunting is also based on the canid pattern. Several Melinae and Mellivorinae (badger) species pursue their prey by digging after it. Species such as the otters and African wild dogs use energetically demanding behavior to catch their quarry (a high investment and high reward strategy), which makes them especially vulnerable to fluctuations in prey density and to food loss to scavengers.

As a result of the variation in foraging methods between families and subfamilies, there are also phylogenetic patterns in the diet. Compared with many other species of mammals and birds, diet analysis for carnivores is relatively easy, although it has some serious problems. Food usually consists of clearly discrete items, which can be recognized and quantified. Furthermore, scats are often easy to find, no matter how elusive the animals may be, and scat analysis has become a major tool despite the difficulty of relating scat content to diet in a quantitative manner. Also, for many species it has proved possible to obtain direct, quantitative observations of predation and foraging behavior. The result is an extensive body of knowledge of carnivore diet, the important link in the relation between the predators and their environment.

One summary showed that of 111 species of carnivore (from all families), only 36% could be classified as predominantly meat eaters, i.e., taking more than 60% of their diet in the form of other mammals or birds. Indeed, in that analysis among representatives of families such as the bears, raccoons, and viverrids, there were no proper meat eaters at all, and they were found to feed on insects, vegetation, or a mixture of various food categories. Many species were called omnivorous if their diet did not include 60% of any one category of food. However, such a limit is quite arbitrary, and many important species were left out of the analysis because of lack of information. Therefore, it is an oversimplification, and one could also summarize the diet differently (**Figure 4**): The majority of carnivore species will eat meat and will prey on other mammals at some time or other. However, the point had to be made that many other kinds of food are involved.

There is a clear importance of phylogeny in the diet: The food of a felid tends to be more like that of another felid than that of a canid and vice versa, and this applies to several families (**Figure 4**). All families except the felids are intensive exploiters of vegetable and invertebrate food sources: The Felidae are the most exclusively carnivorous—the ultimate predators. Only the felids, hyaenids, and canids feed significantly on large mammals, with an occasional exception in the other families. The viverrids, which are probable ancestors of felids and hyaenids, do not kill larger mammalian prey but they are either insectivorous or have a very mixed diet. The diets of the four hyenas are more different from each other than between the members of the other families, and their specializations range from wildebeest to termites or melons or carrion. The bears and pandas are vegetarians. Species in the raccoon family all have a mixed diet, including much vegetable matter. Canids have mostly varied diets, composed of insects, fruits, and mammals. Many of them are proper meat eaters, but even these often include some vegetable matter, quite unlike the cats.

Some of these family-specific trends are further refined in the subfamilies. The mostly meat-hunting mustelids also include two subfamilies with 9 badgers, which almost all feed on invertebrates and vegetation, and a subfamily of 13 otters that subsist on fish or crabs.

Not only are there differences between families and subfamilies in the kinds of prey or vegetation which they select, but also the degree of specialization varies (**Figure 5**). Specialization can affect an animal's vulnerability to environmental change. Among the canids, each species uses an average of 6.5 (out of 10) major prey categories, each constituting at least 1% of its food. On the other end of the scale, each bear only uses 3.7 prey categories and each felid 4.0, so they are much more specialized. The other carnivore families are intermediate. This specialization (dependence on a few resources) may make bears and felids more vulnerable to environmental change than canids.

The degree of specialization can be described only in the broadest of terms because it is difficult to measure and quantify. Terms such as omnivore, opportunist, generalist and specialist have no absolute values, and they refer to animals which may be selecting from a variable set of availabilities. There is also a problem with definitions. For instance, the Eurasian badger is highly focused in its food selection in any one area, concentrating entirely on earthworms in northwestern Europe, on rabbits in southern Spain, and on olives in northern Italy. There is no doubt that in each of these areas badgers are highly specialized, but their specializations are different in different places and overall they could be described as opportunists. Currently, there is no good quantitative descriptive term for such a pattern.

Nevertheless, despite inadequate terminology it is recognized that in any one place some species use many more different prey categories than others. For instance, a cheetah on an African savanna takes almost only small or medium-sized ungulates on open grasssland, whereas a leopard in the same area is much more catholic in its tastes. It eats the same mammals but also much smaller mammals as well as birds and snakes, stalking them in the open and in dense bush or between the rocks. Along European streams a mink will eat small mammals, frogs, fish, birds, and insects, whereas along the same banks the (larger) otters feed almost exclusively on fish and frogs, clearly much more focused. Such a broad comparative indication of specialization suggests the dependence of a predator on few or many prey categories, even

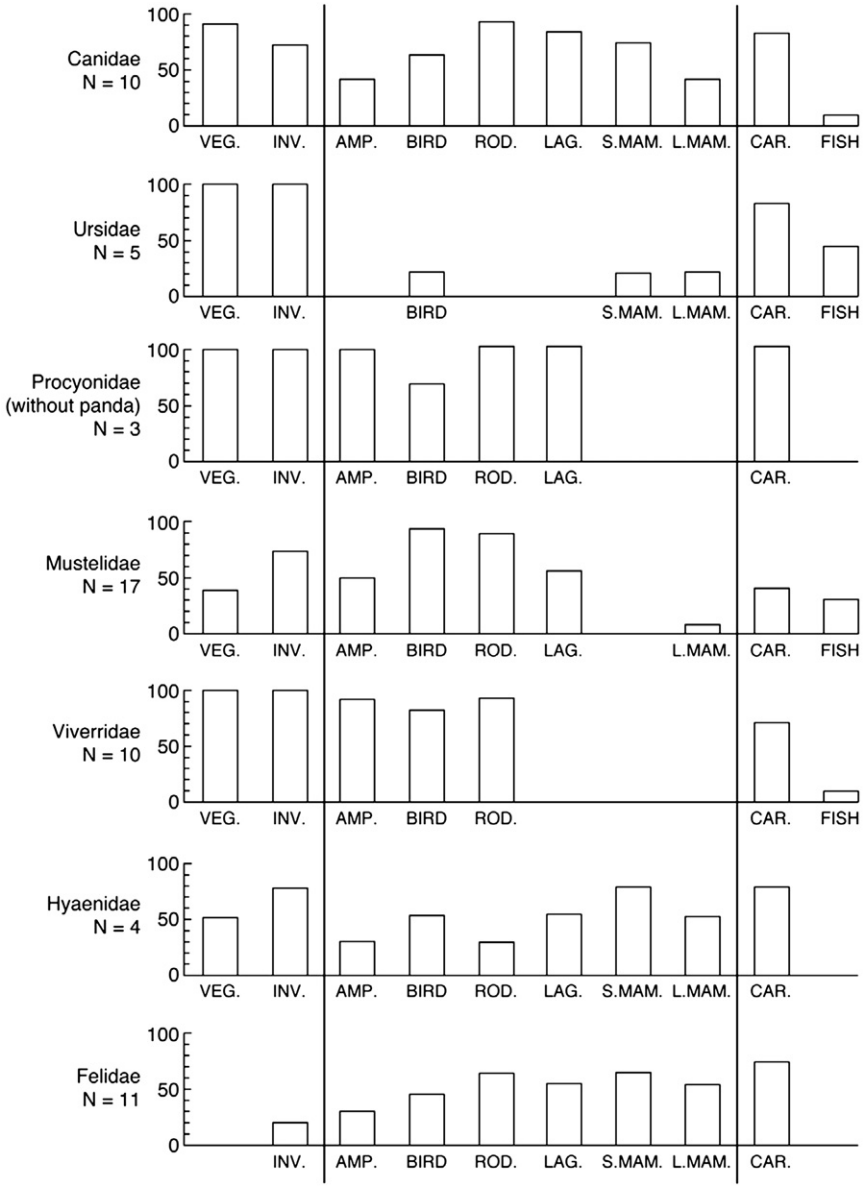


Figure 4 The use of different foods by species from different families. Percentage of species in each family in which a particular food category constitutes more than 1% of diet. Veg, vegetable foods; inv, invertebrates; amp, amphibia and reptiles; rod, small rodents; lag, rabbits and similar size prey; s.mam, large mammals smaller than 50 kg; l.mam, large mammals larger than 50 kg; car, carrion. Reproduced from Kruuk H (1986) Interactions between Felidae and their prey species: A review. In: Miller SD and Everell DD (eds.) *Cats of the World: Biology, Conservation and Management*, pp. 353–373. Washington, D.C.: National Wildlife Federation.

though there may be difficulties with labeling as specialist or opportunist.

In general, because closely related animals often have similar food habits, food selection may be termed a conservative characteristic in the evolution of carnivores. There are many exceptions; for instance, among the meat-hunting felids there is a fishing cat, and the hyena family includes the aardwolf, which feeds on nothing but termites, the spotted hyena, which is an exclusive large ungulate hunter or scavenger, and the striped and brown hyenas, which are as catholic in their tastes as possible. However, these exceptions do not invalidate the overall importance of ancestry in the animals' environmental relationships.

Carnivore Guilds in Ecosystems

Closely related members of the same carnivore family tend to exclude each other, although there are important and puzzling exceptions. Of course, competition between natural populations of different carnivores is rarely observed, probably because those cases in which it occurred have long ago come to their natural conclusion, i.e., the demise of one of the contestants. However, when perturbations occur competition may be obvious. For instance, the introduction of the American mink in Europe totally removed the European mink from most of its range. The famous wolf population of Isle Royal arrived there in the early 1950s, and it completely replaced a

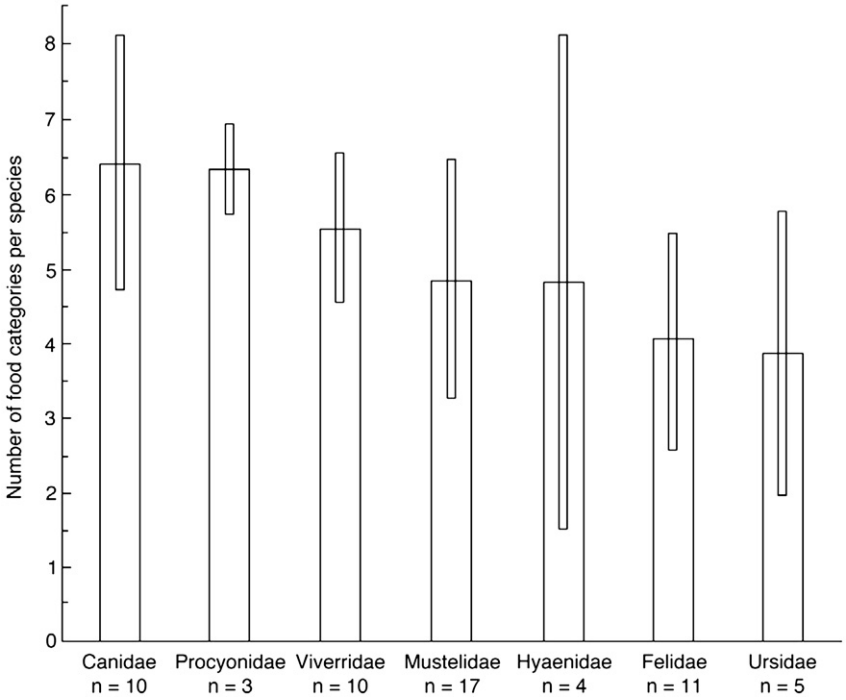


Figure 5 The number of food categories used by species of different carnivore families. *N*, number of species studied, with mean and standard deviation of the number of major food categories (see also **Figure 4**). Reproduced from Kruuk H (1986) Interactions between Felidae and their prey species: A review. In: Miller SD and Everell DD (eds.) *Cats of the World: Biology, Conservation and Management*, pp. 353–373. Washington, D.C.: National Wildlife Federation.

population of coyotes. There are several observations of wolves killing coyotes, apparently without the coyotes being eaten. In many other areas in North America, just one of the two species is found, although usually these places would appear to be suitable for both. Coyotes tend to replace red foxes.

Red foxes are aggressive to arctic foxes and exclude them when their areas overlap, but arctic foxes can feed and survive at much lower temperatures. Recent red fox increases (e.g., in northern Scandinavia) caused the demise of arctics over large areas. In the 1970s, when the black-backed jackals of the Serengeti were decimated by disease, the very similar and previously rare side-striped jackal population increased dramatically. Tigers are reported to often exclude leopards, and different species of otters exclude each other.

Also, species that are not closely related but have overlapping ecological niches may affect each other. For instance, the African wild dog disappeared from the Serengeti between the 1960s and 1990s probably partly because of an increase in populations of hyenas and lions.

Exceptions to this pattern of exclusions are the three species of African jackal. They are very similar and closely related; nevertheless, their geographical ranges overlap considerably and they can be seen eating from the same carcass.

There are other examples of perturbations or species introductions causing the disappearance of competing carnivores, but such events are relatively rare. Thus, it is often assumed that populations of predators each have their own ecological niche and rarely affect each other. However, the ecosystems that we see are end results, and often what we see may be the status quo long after earlier populations have been wiped out or have been prevented from moving in.

One way in which predator species adapt to an ecological niche in a particular habitat is through morphological or behavioral variation. When potentially competing species share a range, differences between them (e.g., in size) may be greater than those from animals that live in ranges not shared with the other species. This is known as character displacement. One of its consequences is that in a guild of predators in any one area a character, such as body size or size of the canines, shows a fairly evenly spaced stepwise distribution among species (**Figure 6**). The obvious effect of this is to minimize competition. Such divergence of morphological characters does not always occur, however, and its striking demonstration in British mustelids is contrasted with an almost absence in Serengeti jackals. The marsupial carnivores of Tasmania also show a clear character divergence.

The existence of such structured mechanisms, to avoid competition within predator guilds in ecosystems, suggests that resources for the carnivores may be at a premium. This in turn presents the possibility that the predators depress prey populations.

It is often argued that predators have little or no effect on prey numbers, as demonstrated by ecosystems such as the Serengeti with more than 25 carnivore species that coexist with the many prey in apparently stable populations. However, such an apparent stability is the result of interactions over a long period of time, and it is possible that many prey species were extinguished in the past and many predator-prey relationships have not had a chance to become established (as first arrivals in the ecosystems were eliminated).

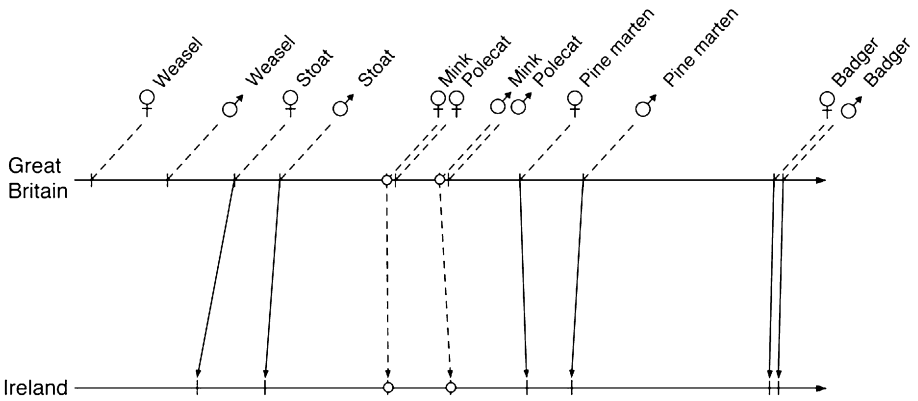


Figure 6 Even spacing of mean lengths of skulls of Mustelid species in Britain and Ireland, log scale. The mink has recently been introduced. Reproduced from Dayan T and Simberloff D (1994) Character displacement, sexual dimorphism, and morphological variation among British and Irish mustelids. *Ecology* 75: 1063–1073.

The deliberate or accidental introduction of carnivores into ecosystems in which they were previously absent provides an opportunity to demonstrate effects of predation. Introduced predators (cat, dog, red fox, mongoose, American mink, stoat, and ferret) have caused extinctions and declines of endemic prey species throughout the world, including many birds, mammals, and reptiles. There is little doubt, therefore, that carnivores cause extinctions or prevent populations from becoming established, and they must have done so in the evolutionary past. On all continents the fauna would look very different from what it is now if there had been no carnivore predators.

Also in the predator–prey systems which currently exist, the effects of carnivore predation can be far-reaching. Predator removal often dramatically increases numbers in prey populations compared with those in suitable control sites, as demonstrated by foxes preying on partridges. Pest species such as rabbits can be limited in numbers by carnivores operating in conjunction with other mechanisms such as disease. Wolves may exercise major effects on populations of caribou, moose, and white-tailed deer, and spotted hyenas can have a major effect on wildebeest. Often, other limiting factors are involved simultaneously, e.g., predators may be the means whereby a herbivore population is maintained at the carrying capacity of the vegetation.

Carnivore Social Systems

The majority of carnivores are solitary and territorial. Their spatial organization is maintained by females (with or without offspring) defending a range against other females of the same species, and by males defending a (larger) range against other males. Male ranges may overlap with several females ranges. Territorial behavior involves scent marking and visual displays as well as direct aggression.

There are some important variations to this generalization. The marsupial carnivores are not territorial; therefore, instead of a regular spacing between individuals, they are more or less randomly distributed. This begs the question, not satisfactorily answered to date, why eutherian carnivores spend so much time and effort on risky territorial behavior if other species with similar ecology can do without.

The organization of canids is pair based rather than solitary; a male and a female share the same range. This is also the only family in which males usually assist in providing the offspring (by regurgitating food). The fact that such pair behavior is tied with the species' phylogeny suggests that in individual species it is not necessarily adaptive.

From the simple, territorial arrangement a group organization has evolved independently in several species in all carnivore families except the bears, and such group living has complicated land tenure considerably. Among the canids the most striking examples are the wolf, the wild dog, and the dhole in Asia, but in several other canid species offspring may also remain for 1 or 2 years, overlapping in time with subsequent litters of the basic pair. This has been described for several jackals, foxes and others.

African wild dogs live in extremely tight packs, often composed of 20 or more individuals, almost always close together, with on average twice as many adult males as females. The wolf has a very different organization. It also lives in sometimes large packs, but it does so in a much more "fission–fusion" type of society, with individuals coming and going, sometimes hunting in groups and sometimes alone, within the pack territory. In both species usually only one female per pack breeds.

Group organizations in other carnivore families show as many different patterns as there are species, with, for instance, lions in permanent prides of up to 20 related females joined by small groups of males which are replaced every few years. Spotted hyenas live in female-dominated clans of up to 80 in group territories, and individual members may hunt or join with others or they may be solitary within the group territory. Cubs are cared for almost entirely by their own mothers. Both banded and dwarf mongooses occur in dense packs, with the sexes mixed, and all members of the pack care for all offspring, Eurasian badgers forage on their own and look after their own cubs only, but they live in group territories defended by both sexes. Female Eurasian otters, on the other hand, may occupy individual core areas within a group territory of five or six females, whereas males remain in their own ranges.

For the phenomenon of group living in carnivores, there are no striking, simple phylogenetic trends and no substantial effects of body size, climate, prey size, predation, or various other factors to explain the occurrence of packs, bands, or prides. It does not appear that social species have done any better or worse than solitary ones in terms of numbers or densities, nor is their future survival more or less endangered. There is, however, one set of environmental factors which appear to affect gregariousness, i.e., the distribution of food or of resources in general.

Currently, there is only one general hypothesis to explain grouping of carnivores—the resource dispersion hypothesis, first explicitly presented by Macdonald (1983). In principle, it relates the size of territories and the numbers of animals inhabiting each territory to the spatial and temporal pattern of resources. For instance, it suggests that badgers in northwestern Europe live in groups because their main food, earthworms, occurs in well-spaced patches, each of which is available only at certain times. To be able to feed at all times a badger needs a large area with several patches, with the size of the area determined by the scatter of the patches. Once a badger has such a territory it can accommodate several more individuals with only limited or no competition. The hypothesis is attractive, but it does not explain all the intricacies of spatial organization. To date, there is no comprehensive alternative.

Carnivores living in groups may or may not hunt cooperatively, depending on species or conditions. Clear advantages of cooperation, in terms of hunting success or energy balance, have been demonstrated for wild dogs and spotted hyenas. Lions collaborate on a well-organized basis, with each pride member occupying a preferred place in the hunting formation and deviations of this pattern resulting in lowered success.

Changes in Diversity

When the Carnivora first appeared, approximately 60–40 million years ago, there were other, similar animals already well established. For instance, the extinct order Creodonta included families such as the Hyaenodontidae, and there were various large marsupial predators such as the Thylacoleo or pouch lion. Proper carnivorous feeding and predation have evolved several times independently—at least twice among the marsupials (in the Borhyaenidae in South America and some Dasyuridae in Australia) and twice among placental mammals (in Creodonta and Carnivora).

The large carnivorous expansion from the Palaeocene onwards coincided with the evolution of angiosperms, flowering plants and grasses (evolving away from the ferns and gymnosperms such as conifers). This resulted in a large floral diversification, including savanna-type vegetations, and enabled the extensive diversity of ungulates and rodents to evolve. This in turn enabled the evolution of specialist predators.

Most of the other, noncarnivore large predators have now gone extinct, with the last creodonts occurring approximately 8 million years ago and the last really large marsupial carnivores occurring 2 million years ago, when species of Homo were already well established. Of the smaller marsupial carnivores, just a few small dasyurids still occur in Australia. However, while the other predators slowly disappeared, the order Carnivora diversified into a multitude of different families, genera, and species.

Why the creodonts disappeared, while at least initially carnivores thrived and probably replaced the creodonts, is a mystery. Creodonts and carnivores were closely related and the skeletal remains, such as vertebrae and the locomotory system, were similar. The difference between creodonts and carnivores was not much greater than the variation within these groups. However, obviously there is more to an animal than its skeleton, and the reasons for extinction may well have lain in other aspects of morphology, physiology, or behavior.

In the early stages of evolution the advantages appear to have lain with the creodonts, and only after 20 million years did the balance tip in favor of the Carnivora. The Carnivora have been four times as successful as the creodonts: The latter are known from 45 genera spanning approximately 45 million years, whereas Carnivora, excluding living and aquatic genera, are know from 218 genera over a span of more than 55 million years.

Carnivore species are also fewer now than they were in the geological past. In fact, most of the Carnivora have become extinct, and although there is still a rich complement of species, there were many more in the past. For instance, we know of 333 genera in the seven extant families of carnivores, of which 237 (71%) are extinct. Many complete carnivore families have also disappeared, just like the creodonts and large marsupial predators.

An interesting phenomenon was the occurrence of sabertooth species. Saber teeth evolved several times independently in different species, families, and even orders and included the Megantereon, Homotherium, and Machairodus, which were at least as big as a lion, and formidable felids such as the North American saber-tooth cat Smilodon. Some of these were present at the same time as early species of Homo, but all of them are extinct. They pose a difficult problem to paleontologists. Large canines are used by today's carnivores for killing prey and for social purposes such as fighting opponents over territorial claims. However, were the extra-large saber teeth—the huge, flat daggers which were seemingly far too large for any jaw—used to kill extra-large prey, for opening carcasses, or what? In fossil assemblages it was always the very largest ones, the top predators, which sported saber teeth.

There is no likely explanation for saber teeth in the acquisition of food. The fragile, sharp weapons, often with serrated inner edges, must have been quite useless against thick skin or on large bodies, with the gape of the owner being insufficient to use these canines effectively. However, saber teeth were obviously effective weapons; otherwise, they would not have evolved several times independently. However, they also disappeared again in all these cases. It is likely that saber teeth made use of some Achilles heel in their prey (of which we have no evidence today), but that in response the prey species evolved means of protection. It was an arms race which was eventually lost by the sabers, but we do not know who conquered and why.

Homo is a highly successful hunter, and mankind was and is in competition with Carnivora at several different levels. Homo also preys on and is preyed on by Carnivora. Therefore, are people the cause of carnivore extinctions? The question is

often asked, and answers are far from straightforward. Many carnivore species extinctions occurred approximately 4 million years ago, at the time that hominids arrived on the scene, and many followed during the next 2 million years.

However, many extinctions occurred before man arrived, and at the time Homo's arrival was not the only event which changed the environment. There were drastic changes in climate, for instance, approximately 3.2, 2.4, and 0.8 million years ago. The Pleistocene Period started at 0.8 million years ago with a massive climatic shift that coincided with the appearance in Eurasia of African species such as the lion, leopard, spotted hyena, and perhaps the major (although not the first) movement out of Africa of Homo. If such major dispersal events occurred in conjunction with climatic changes, it is equally likely that extinctions would have occurred.

For many of the carnivore extinctions, we cannot blame our own species with any conviction. Homo may have been closely involved or not at all, and perhaps environmental change rendered species more vulnerable to competition and predation by mankind. Of course, recent extinctions have been fairly well documented, and here mankind's guilt is in no doubt. The Tasmanian wolf or Thylacine was exterminated by sheep farmers in the 1920s and 1930s. The North American sea mink was obliterated in the late nineteenth century for its fur, and the sea otter almost followed it into oblivion; it was barely saved and it has recovered fairly well. The "wolf" of the Falkland Islands was still there when Charles Darwin visited, but sheep farmers exterminated it and since the 1880s it has existed only in museums. Several carnivores were totally eradicated from Britain, including the brown bear and the wolf, whereas wildcat, polecat, and pine marten have only just managed to survive. There are long lists of carnivore extinctions from throughout the world, and there is no doubt that most of these were man induced.

Whatever caused the demise of carnivore species in the past, it is important to prevent it from occurring in the future. Several species are on the brink of extinction in the wild, including tiger, panda, European mink, and black-footed ferret, and many more face local extermination. In geological terms, the diversity of carnivores is decreasing extremely rapidly.

See also: Food Webs. Mammals, Biodiversity of. Predators, Ecological Role of

References

Creel S and Creel NM (1995) Communal hunting and pack size in African wild dogs, *Lycaon pictus*. *Anim. Behav.* 50: 1325–1339.

Dayan T and Simberloff D (1994) Character displacement, sexual dimorphism, and morphological variation among British and Irish mustelids. *Ecology* 75: 1063–1073.

Gittleman JL (ed.) (1989) *Carnivore Behavior, Ecology and Evolution*. Ithaca, NY: Cornell Univ. Press.

Gittleman JL (ed.) (1996) *Carnivore Behavior, Ecology and Evolution*, Vol. 2. Ithaca, NY: Cornell Univ. Press.

Gittleman JL and Harvey PH (1982) Carnivore home-range size, metabolic needs and ecology. *Behav. Ecol. Sociobiol.* 19: 57–63.

Kruuk H (1986) Interactions between Felidae and their prey species: A review. In: Miller SD and Everell DD (eds.) *Cats of the World: Biology, Conservation and Management*, pp. 353–373. Washington, D.C.: National Wildlife Federation.

Kruuk H (1989) *The Social Badger*. Oxford: Oxford Univ. Press.

Kruuk H (1995) *Wild Otters: Predation and Populations*. Oxford: Oxford Univ. Press.

Macdonald DW (1983) The ecology of carnivore social behaviour. *Nature London* 301: 379–384.

McNab BK (1989) Basal rate of metabolism, body size, and food habits in the order Carnivora. In: Gittleman JL (ed.) *Carnivore Behaviour, Ecology and Evolution*, pp. 335–354. Ithaca, NY: Cornell Univ. Press.

Mills MGL (1990) *Kalahari Hyaenas*. London: Unwin-Hyman.

Stander PE (1992) Cooperative hunting in lions: The role of the individual. *Behav. Ecol. Sociobiol.* 29: 445–454.

Woodroffe R, Ginsberg J, and Macdonald D (1997) *The African Wild Dog*. Gland, Switzerland: IUCN.

Carrying Capacity, Concept of

Gregg Hartvigsen, State University of New York, Geneseo, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Density dependence The condition that environmental factors influence population growth rate in relation to population size. Density dependence usually is seen as a linear, inverse relationship between population growth rate and population density (i.e., population growth decreases as density increases) and may occur if individuals compete or predators are more effective as a prey population increases.

Density independence The absence of environmental factors that influence population growth as a function of density. This may occur if mortality removes a fixed percentage of a population, independent of population size.

Logistic growth Regulated population growth that follows the logistic equation $dN/dt = rN(1 - N/K)$. Populations growing according to this equation increase rapidly at low densities and the growth rate decreases as they approach carrying capacity (K).

Population A group of individuals of a particular species that live in a region. A population is usually a subset of the entire species.

Population regulation The constraint of positive population growth. The study of population regulation deals with the factors that cause this constraint, such as competition for food or predation.

Population stability The tendency for populations to return to a previous size after a disturbance, such as reductions due to hunting or disease or increases due to immigration. Stable populations may be locally stable (return after small disturbances) or globally stable (return after severe or catastrophic disturbances).

Introduction

Populations, or groups of individuals within a species, change over time. There is a general agreement among ecologists that population growth is bounded by biotic and abiotic environmental factors that result in approximate, maximum numbers of organisms that can be supported in different habitats. A population's carrying capacity is difficult to measure and likely varies over time and through space. The concept of carrying capacity has played an important role in the fields of basic ecological research, wildlife management, and conservation biology.

The concept of carrying capacity also involves determining how many people Earth can support. Dialogs about human carrying capacity are often quite contentious and illustrate the difficulty surrounding the concept of carrying capacity.

The concept of carrying capacity is alive and well, although some have argued that it should be abandoned altogether. On the surface, the concept is easy to understand and intuitive and, therefore, is likely to stay with us for some time. However, recent developments in our understanding of the dynamics of population change over time have, greatly modified what is considered an area's carrying capacity. Therefore, we need to recognize both the strengths and the weaknesses of the concept of carrying capacity.

Origin of the Concept of Carrying Capacity

Recognition of carrying capacity probably occurred long before written history began. It is likely that the earliest agriculturists, perhaps 10,000 years ago, were keenly aware of the number of mouths that an area could sustainably feed and that increasing numbers of people required increases in food production (i.e., area of land in cultivation). Long before agriculture, hunter–gatherer groups likely were aware of the sustainable number of members that regions could support, although it may be argued that high mortality rates inhibited these early populations from pushing the limits of sustainability. During difficult times populations reached or exceeded what we might think of as a carrying capacity, possibly imposing nomadic lifestyles in which groups had to intermittently move after local resources were depleted. Speculating on past population dynamics hints at the trouble with the concept of carrying capacity: It must represent a dynamic value that changes over time and is highly dependent on many interacting factors, such as environmental variability and, for early humans, what hungry animals awaited their forays. This makes the concept difficult to use.

Charles Darwin (1859), on page 116 of his book *Origin of Species*, quoted in his third chapter on the struggle of existence from Malthus' 1798 "An Essay on the Principle of Population as it Affects the Future Improvement of Society:"

> A struggle for existence inevitably follows from the high rate at which all organic beings tend to increase. Every being, which during its natural lifetime produces several eggs or seeds, must suffer destruction during some period of its life, and during some season or occasional year, otherwise, on the principle of geometrical increase, its numbers would quickly become so inordinately great that no country could support the product. Hence, as more individuals are produced than can possibly survive, there must in every case be a struggle for existence, either one individual with another of the same species, or with the individuals of distinct species, or with the physical conditions of life. It is the doctrine of Malthus applied with manifold force to the whole animal and vegetable kingdoms; for in this case there can be no artificial increase of food, and no prudential restraint from marriage. Although

some species may be now increasing, more or less rapidly, in numbers, all cannot do so, for the world would not hold them.

Malthus' clear recognition of the importance of limitations of growth in populations helped Darwin to lay the foundation for his theory of natural selection, which is built on the premise that populations are regulated primarily by competition which leads to differential reproduction.

The earliest concise description of carrying capacity is derived from Pierre Francois Verhulst, a Belgian who lived in the mid-nineteenth century. Verhulst, perplexed on account that the human population appeared to be increasing exponentially, derived a mathematical formula which he called the "logistic" equation that would account for a slowing in the population growth rate as a function of population size. The same relationship was rediscovered by Pearl and Reed (1920), who used the logistic equation to predict the population of the US based on census data collected from 1790 to 1910 (apparently neglecting the fact that the area of the US increased more than threefold during this time). The resulting application of the logistic equation to US census data led Pearl and Reed to greatly underestimate the US population, predicting it would level off at about 197 million in the year 2050 (**Figure 1**).

The failure of Pearl and Reed to accurately predict the population of the US, currently at about 275 million, reveals at least one important aspect of mathematical models. It indicates that, for the US human population, some aspect of the logistic equation must be wrong. Two candidate problems include violations of the assumptions that the area of the US remained constant and that the US population lacked immigration. These differences between data and model predictions can help us to better understand the problem at hand. In this example, the violations of model assumptions leads us to re-evaluate the factors that influence population growth and ultimately regulate it. The prediction from the model also clearly did not hold for a variety of socioeconomic and human health reasons and may have been wrong for statistical reasons. Using the logistic equation (based on the filled data points in **Figure 1**) to predict population size well into the future, without providing confidence limits, is clearly tenuous at best. These problems have led us to be more careful in our predictions of how populations change over time and whether a carrying capacity can be predicted from such data or even exists.

Definitions of Carrying Capacity

Carrying capacity is the maximum population that a given area can sustain. The measures commonly used include the number of individuals or the total biomass of a population, which are each highly dependent on differences in physiology and age structure among species and across large taxonomic groups. The use of the term carrying capacity has changed over time, but most models suggest that population growth is rapid when density is low and decreases as populations increase towards some maximum. In addition, any definition of this concept improves as we narrow the time and area for the population that we are studying. Population descriptions, therefore, are often depicted as densities, accounting for the number of individuals per unit area. Population density usually varies over time and from place to place. In practice, we generally use population size or density to describe carrying

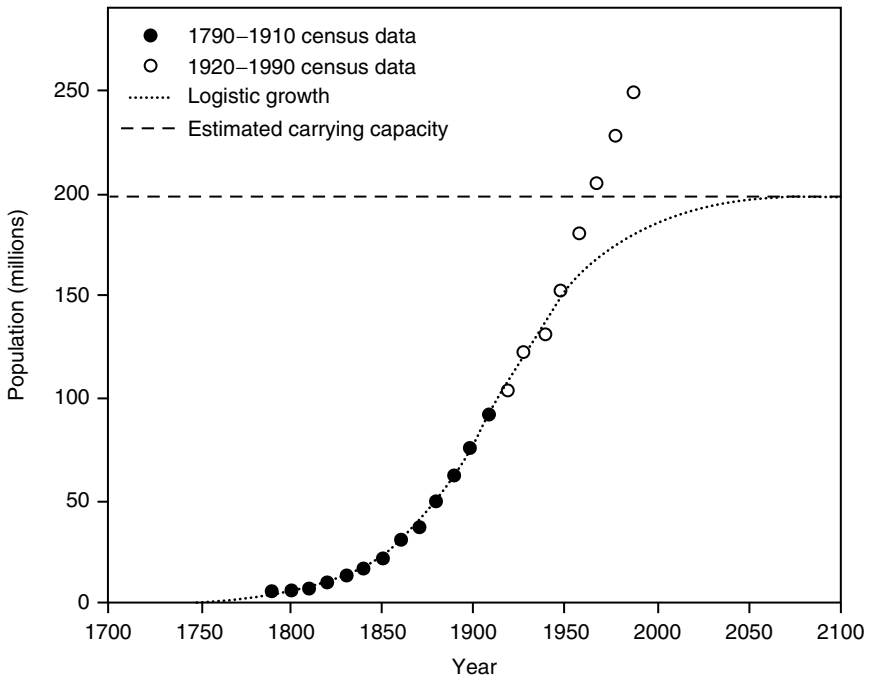


Figure 1 Data on the growth of the US population from 1790 to 1910 (●) and from 1920 to 1990 (○). The dotted line represents Pearl and Reed's fit of the logistic equation, yielding a carrying capacity (K) of about 197 million people, estimated to be reached in about the year 2050. The population in the year 2000 is about 275 million (not shown).

capacity, which is determined either by resource availability or by the influence of enemies (predators and/or pathogens).

Various definitions of carrying capacity arose in the twentieth century, ranging from the suggestion that carrying capacity is that level below which predators have no effect on a population to the population size which can be maximally supported in a given region (previously referred to as the "saturation level"). There also has been a distinction made between "ecological carrying capacity," which refers to the limitation of a population due to resources, and a management oriented, maximum sustainable yield for a population, referred to as an "economic carrying capacity," which is usually lower than ecological carrying capacity. These definitions clearly lead to difficulty for wildlife managers who have been preoccupied with attempting to determine whether populations are either too high or too low. These debates continue, as exemplified by range management decisions in Yellowstone National Park and issues regarding the increasing frequency of reintroduction programs of top predators.

Carrying capacity may best be expressed mathematically. One of the simplest forms of population change over time can be represented as the differential equation $dN/dt = rN$, where dN/dt represents the instantaneous change in a population over a short time period, r is the intrinsic growth rate of the population, and N is the size of the population. This yields, what is often referred to as a "J" curve, or exponential growth (**Figure 2**). In discrete time this relationship is referred to as geometric growth.

In 1838, Verhulst modified the exponential growth equation and derived the logistic equation that depicted population growth rate as being inversely related to population size. To slow population growth he added an additional term yielding $dN/dt = rN(1 - N/K)$, where K is the population carrying capacity. The term $1 - N/K$ slows growth rate linearly toward zero as the population (N) approaches the carrying capacity (K). This results in a sigmoidal S-shaped curve for an increasing population over time (**Figure 2**). If the population exceeds $K (N > K)$, then $1 - N/K$ is negative, causing growth rate dN/dt to be negative and the population to decline monotonically toward K.

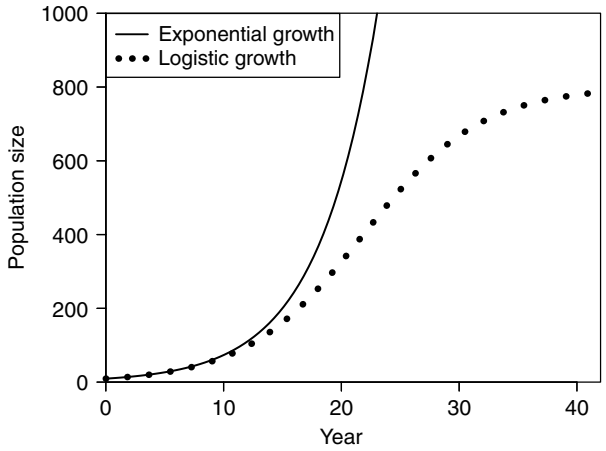


Figure 2 Comparison of unregulated exponential growth (solid line) with regulated logistic growth (dotted line).

An important attribute to bear in mind is that the logistic equation is deterministic, meaning that if we use the equation to predict population size, at the end of a fixed amount of time we will derive the same population each time we start the population over. This assumption is usually violated in field conditions in which random effects, such as accidental deaths, failure to find mates, or fluctuations in environmental conditions, are common. Therefore, it has been argued that we should not expect real populations to behave according to the logistic equation.

This simple equation has been challenged repeatedly by critics without apparent damage. This resilience of a theory is rather rare in science, which is a discipline that prides itself on being able to quickly dispel hypotheses (or equations), given even a small amount of contradictory data. However, the intuitive nature of the idea that populations are regulated by factors such as food supply helps the logistic equation to remain a staple in ecological texts and classrooms. The reason why this equation and carrying capacity (K) endure is that the equation's shortcomings help us better understand the dynamics of real populations, ensuring its utility for many years to come.

The discrete, or difference, form of the logistic equation yields a different prediction of population behavior compared to the previously described continuous version. In particular, the discrete form was the equation used by Robert May to first describe how a simple, deterministic equation could produce chaotic population dynamics, a pattern that emerges when intrinsic growth is relatively high. This chaotic behavior appears to mimic realistic changes in populations over time. Several long-term data records conform to chaotic dynamics, including the change in the number of lynx captured over time in Canada (**Figure 3**).

Do Populations Have Carrying Capacities?

This question has been addressed using a variety of techniques, including observational types of studies that rely on long-term time series data sets such as the number of lynx captured over time (**Figure 3**), highly controlled laboratory experiments (**Figure 4**), and mathematical models to determine potential mechanisms through time series reconstruction. The answer therefore seems to be that carrying capacities do exist for most species but that determining these at any one point in time or space is quite difficult.

Many controlled laboratory experiments have been published that show populations behaving in a fashion consistent with the logistic model (i.e., populations reach a carrying capacity). One of the earliest studies was completed by Gause (1934) (**Figure 4**). Most laboratory populations tend to increase and then reach some level at which they fluctuate around what might represent a carrying capacity. It is interesting to note that populations rarely exhibit a smooth transition between a growth phase and gradual reductions in growth rate ending in stable populations, despite the controlled environmental conditions. An additional caveat to consider is that prior studies that concluded population growth patterns differed significantly from the expected

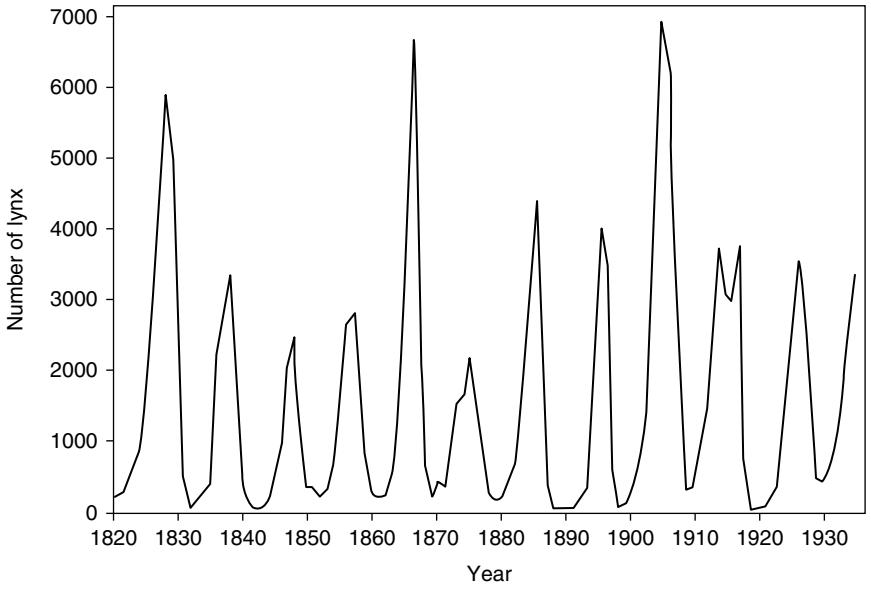


Figure 3 The number of lynx trapped in the Mackenzie River district from 1821 to 1934. Reproduced from Elton C and Nicholson M (1942) The ten-year cycle in numbers of lynx in Canada. *The Journal of Animal Ecology* 11: 215–244, with permission from British Ecological Society.

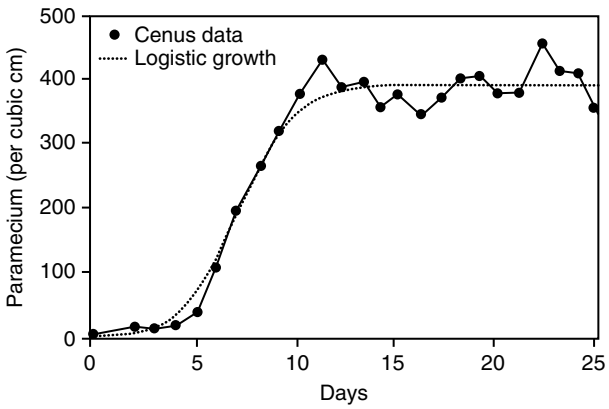


Figure 4 The change in the density of *Paramecium caudatum* over time in the laboratory. The dotted line represents the best fit logistic equation. Reproduced from Gause GF (1934) *The Struggle for Existence*. New York: Macmillan.

logistic growth likely suffered disproportionately during the review process and failed to make it into print.

Determining how natural populations change over time is surprisingly difficult. The first requisite information necessary to determine whether a population is at or near a constant size, considered here to be a proxy for the habitat's carrying capacity, is simply the population's size over time. This often has to be determined over long periods of time in ways that are accurate, reliable, and repeatable. In field studies, it is rare to have the luxury to repeatedly estimate population size, a technique allowing us to assess the accuracy of our estimates. Determining population size fluctuations may represent real changes in population size or may represent either natural variability (statistical "errors") or actual errors in our estimates.

Assumingly, we overlook these shortcomings in our data, what do populations do? In general, populations usually fluctuate over time. We may be able to correlate these changes with biotic or abiotic factors or some function of the two with time. Sometimes the fluctuations cannot be distinguished from random noise. Some populations, including the classic examples of lynx, hares, and lemmings, cycle periodically (**Figure 3**). The persistence of population cycles over long periods of time has led to great speculation regarding the factors that might lead to periodicity. Recent work suggests that simple causal mechanisms of cycling are unlikely and that a combination of random environmental factors and nonlinear, density-dependent factors influence populations.

Determining the Carrying Capacity of a Population

Determining the carrying capacity of any particular population at a particular time is not trivial. Many different techniques have been suggested and tested, including three primary techniques that can be used to attempt to detect a change in population growth rate as a function of population size: Mathematical modeling of specific mechanisms, tactical experimental tests in the laboratory and field, and statistical analysis of time series data. Ultimately, a combination of these techniques will enable us to understand the importance of regulation in populations and the degree to which populations appear to be governed by a carrying capacity. The author has already discussed the importance of the logistic equation and will briefly introduce the empirical approaches.

There has been much interest in the work that is completed to determine what factors regulate the change in growth of populations over time. The factors that slow population growth rate change over time and from location to location and differ for different species. Regulating factors also are likely to interact with each other, thus complicating the determination of a population's carrying capacity.

In a classic study, Davidson and Andrewartha (1948) used a partial regression technique to analyze an experiment designed to test the relative influence of biotic and abiotic factors on regulating a small herbivorous insect population. They concluded that 78% of the population variance was due to abiotic or weather-related factors. In particular, the number of individual thrips in the spring was related mostly to the preceding autumn's climate. This study was influential because it provided strong evidence that this population of thrips was regulated not by biotic factors such as competition or predation but rather by abiotic factors.

A second method used to detect the presence of density dependence on population regulation is the analysis of time series data. The best data are those that have been collected over consecutive years and that exceed the periodicity of both observed environmental and population fluctuations (generally >10 years). These data can be subjected to tests that investigate the relationship of change from year to year as a function of the population during the previous year or years in order to detect whether the population appears to be regulated. Such analyses, however, are unable to provide any information on the underlying mechanisms that might lead to population regulation. Therefore, time series analysis is an excellent exploratory tool that can be used to investigate the possibility that a population is regulated. This information can then suggest experiments designed to partition variance among potential candidate mechanisms.

Determination of a population's carrying capacity is best done through a combination of modeling, experimentation, and time series analysis. Research efforts, however, need to be directed toward investigating the underlying mechanisms that govern population regulation. Without an understanding of the relative importance of these regulating factors, it will be difficult to determine whether populations are regulated and whether we can detect a population's carrying capacity.

Current Research on Carrying Capacity

Two main areas of research continue to drive our quest to understand population regulation and the strength and importance of carrying capacity. The persistence of these questions indicates the need to clarify the mechanisms that influence population change over time.

Determining the Relative Strengths of Factors that Regulate Populations

Although some researchers have argued that populations are unregulated, most agree that negative feedback mechanisms operate on populations, resulting in decreased growth at high densities. This may occur through changes in the abundance of food, through increased predation or disease, or through a combination of these biotic factors and abiotic factors such as local climate. These factors may reduce birth rates or increase death rates, or both. Although there are circumstances in which these rates change at low population densities (e.g., the Allee effect, which states that very small populations are likely to decrease due to such factors as difficulty in finding mates or pollen limitation), their regulation at high densities is likely to be common. This change in birth and death rates as a function of density is referred to as "density dependence."

A population that is regulated has intrinsic, extrinsic, or a combination of these factors that slows population growth. Under such conditions a population's per capita growth rate decreases with increasing population size through reduced birth rates and/or increased death rates. This relationship, in logistic growth, is assumed to be linear. The existence of a carrying capacity, however, is not dependent on the shape of this function, so the violation of this linearity assumption does not weaken the concept of carrying capacity. A better understanding of this relationship, generally determined through carefully designed experiments, will help us understand the importance of regulation on population dynamics.

Krebs *et al.* (1995) suggested that hare and lynx cyclic population dynamics are likely influenced by different sets of factors, including food availability and predation driving the dynamics of hare populations while the lynx population is driven primarily by changes in the number of hares. In a more highly controlled experiment using three trophic levels, Hartvigsen *et al.* (1995) determined that plant performance was controlled by the interaction of top-down and bottom-up factors, including the level of plant resource availability and the presence or absence of herbivores and herbivore predators. These studies suggest that complex, interacting biotic and abiotic factors likely influence population dynamics.

Determining Population Carrying Capacity

The logistic growth equation attempts to model, regulated population change over time and relies on several important assumptions, including the absence of time lags (population dynamics is independent of prior events), migration or immigration, genetic variability or selection, population age structure, and the fact that density dependence is linear (each individual added to the population has a similar effect on the population's per capita growth rate). Violations of these assumptions have been found in various populations and have led to more refined, realistic, and complicated forms of the logistic equation. In addition, the model assumes that the carrying capacity (K) is constant over time and space. This assumption occasionally may be valid in situations in which a population is regulated by habitat availability. This might occur, for example, where the number or area of nesting sites is fixed. It is easy to conjure up situations, however, where this assumption would be violated over very small spatial or temporal scales. It is not likely that K would be constant since populations are usually limited by resources, competitors, enemies, and often combinations of these factors that vary with the environment over time. Under these conditions, changes in resource availability can influence population size directly or indirectly through its often nonlinear effect on the population of competitors and/or predators.

In addition, there is great concern about the stability and persistence of threatened and endangered species (*see* The Importance of the Concept of Carrying Capacity to Biodiversity). Work in this area has begun to recognize the importance of species interactions, immigration and emigration

among subpopulations, the introduction of exotic species, and other factors that violate the assumptions of simple logistic growth. The movement of individuals among subpopulations enables the possibility of increased long-term persistence of populations by reducing large-scale fluctuations and the spreading of risk that a species will become extinct in the event that a single local population disappears (becomes extirpated). This area of research, referred to as "metapopulation biology," involves determining long-term viability of these subdivided species, and there is currently much research activity in this area.

The Importance of the Concept of Carrying Capacity to Biodiversity

The concept of carrying capacity suggests that species are likely to have some upper limit to their population. If the upper limit is "hard," then we expect populations to achieve this state and remain relatively constant. Populations, however, as demonstrated in **Figures 3** and **4**, do not behave in such a simple fashion, but have rather "soft" limits. As such, populations usually exhibit random, cyclic, or chaotic dynamics. These dynamics generally lead to increased chances that populations will reach the absorbing state of zero (become either locally extirpated or globally extinct).

We must be concerned about the dynamics of relatively small populations over time. The probability that a population will go extinct is generally related to the degree to which it fluctuates (population amplitude and frequency). Therefore, processes that cause populations to increase fluctuations are likely to lead to species loss and associated reductions in biodiversity. Thus, conservation efforts may be needed that will buffer populations and associated habitats from extreme fluctuations. Conservation efforts are often directed toward increasing a population's carrying capacity. It should be kept in mind, however, that constant environments also may lead to species losses. The intermediate disturbance hypothesis has gained much empirical support and suggests that the maximum number of species that an area can support occurs when disturbances are intermediate in either frequency or impact. We should be concerned that our management efforts do not reduce the carrying capacity of target species.

The Human Carrying Capacity

The best estimates of human population indicate that it has continued to grow exponentially over recorded history, although the actual growth rate has changed over time. Attempts to fit data on the human population to the logistic equation have failed (**Figure 1**), and current indications are that no human carrying capacity can be predicted from simple population statistics. However, we might ask whether our population growth rate is likely to slow down in the foreseeable future and, ultimately, reach a stable carrying capacity or whether it will overshoot its carrying capacity and eventually collapse.

Cohen (1995) found that estimates of the human carrying capacity have ranged between 1 billion and 1 trillion people, with the majority of estimates falling between 4 and 16 billion (the current population is about 6 billion). These estimates suggest that we are approaching an apparent limit for our species. Regardless of which estimate seems most appropriate as an upper limit for humans on Earth, the growth rate of our population will eventually slow to ≤ 0. This can occur as a result of increasing death rates and/or decreasing birth rates. It is predicted that as our population grows in the coming decades there will be an increase in mortality due to diseases. The effect of disease agents on controlling population growth will likely increase due to increases in human contact rates and rapid transit, increasing evolution of drug resistance, and increasing virulence rates. These factors also may reduce birth rates, which of course presents a more pleasant alternative to slow population growth.

Can a population crash be avoided? The author ventures the guess that one cannot. Any long-term stabilization of the human population will require a decrease in the current global birth rate. We certainly cannot hope to achieve a relatively stable population without invoking a substantially higher death rate than the current rate, which is not a comforting thought. It is difficult to imagine, however, that the influence of disease will operate in a simple density-dependent fashion. Instead, it seems more plausible that diseases will "break out" more often with increasing population size and with larger-scale consequences, bringing about a strong reduction in our population – a response seen in many other populations that have increased beyond their carrying capacities.

One last area of hope is that individuals will lower their consumption rates, thereby adjusting the human carrying capacity. It is unlikely that Earth can support tens of billions of people with lifestyles matching those of people in the developed nations such as the US. Therefore, there remains a chance that changes in human behavior will allow our population to gently transition toward a sustainable, zero growth rate population.

Conclusions

The concept of carrying capacity has a history that spans at least thousands of years. The formal definition is approximately 150 years old and is generally coupled to the asymptotic population in the logistic growth equation (K). Critics argue that because of ongoing confusion and the multitude of definitions attached to the concept we would be better off simply abandoning the term. One must also be concerned that the term not be used to advance any particular political agenda associated with determining how large populations of any particular species "should" be in particular areas. This entry cautiously suggests that the concept remains useful. Since most populations are likely to be at least occasionally limited by factors that depend on the population's density, we need to continue advancing our knowledge of how populations behave and use this information to guide the design of laboratory and field experiments aimed at determining the mechanisms that regulate populations. Only by using the combination of field and laboratory techniques, grounded in a theoretical framework that has roots going back to the simple logistic equation, will we hope to understand and

conserve populations, including our own, over long periods of time.

Appendix

List of Courses

1. Principles of Ecology
2. Population Biology
3. Community Ecology
4. General Biology

See also: Population Density. Population Dynamics. Population Stabilization, Human. Population Viability Analysis

References

Cohen JE (1995) *How Many People Can the Earth Support?* New York: Norton.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection*. London: Murray.

Davidson J and Andrewartha HG (1948) The influence of rainfall, evaporation, and atmospheric temperature on fluctuations in the size of a natural population of *Thrips imaginis* (Thysanoptera). *The Journal of Animal Ecology* 17: 200–222.

Elton C and Nicholson M (1942) The ten-year cycle in numbers of lynx in Canada. *The Journal of Animal Ecology* 11: 215–244.

Gause GF (1934) *The Struggle for Existence*. New York: Macmillan.

Hartvigsen G, Wait DA, and Coleman JS (1995) Tri-trophic interactions influenced by resource availability: Predator effects on plant performance depend on plant resources. *Oikos* 74: 463–468.

Krebs CJ, Boutin S, Boonstra R, *et al.* (1995) Impact of food and predation on the snowshoe hare cycle. *Science* 269: 1112–1115.

Pearl R and Reed LJ (1920) On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences of the United States of America* 6: 275–288.

Cattle, Sheep, and Goats, Ecological Role of

Andreas Troumbis, University of the Aegean, Lesvos, Greece

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 651–663, © 2001, Elsevier Inc.

Glossary

Grazing intensity Frequency and closeness of grazing.
Grazing pressure Stocking rate, the units of grazing animals per land area.
Overcompensatory growth In grazed plants, the reaction to tissue removal by enhanced primary production compared to in undefoliated controls.

Domestication of Ruminants: The Process that Changed the Face of the Earth

The natural capacity of ruminants to digest hard plant tissue and to transform it into animal biomass is a service that has been provided by biodiversity to human societies for millennia. The cultural evolution of humans is strongly related to the exploitation of this service, which has influenced—and in turn has been influenced by—the pace, the pattern, and the distribution of societies. The first traces of domestication have been found in the Iranian mountains, where there are paleolithic settlements of sheep and goat herders dating from ca. 10,000 B.P., suggesting that domestic ruminant keeping precedes proper agriculture. Evidence of organized livestock breeding culminates between 8000 and 6000 B.P., at several sites in the Middle East and Anatolia. Ali Kosh, located at the head of the Persian Gulf in the ancient kingdom of Susiana (7000 to 6500 B.P.), is such a site where villages of farmers-breeders are found. Animals and plants were originally domesticated in a specific region, and then diffused during a 10,000 year process to many other parts of the world.

In the past, nomadic tribes and their herds of domesticated animals roamed freely to avoid drought, in search of good grazing land. In the rangelands of North America, Native Americans followed the wandering herds of buffalo, and it was not until domesticated cattle and sheep were introduced by European colonists that the vegetation and soils of the plains and mountains were put under stress. The rise of national boundaries has created a barrier to many such movements. Seasonal transhumance is still a practice in Mediterranean areas, where especially flocks of sheep are moved between winter quarters in the mild rainy lowlands and summer quarters in rich alpine pastures.

On a global scale, diffusion was drastically accelerated by European expansion from the fifteenth century onward. At present, rangelands cover a large part of the world's land surface (40%). Eighty percent of these rangelands are in arid and semiarid areas, where most of the 50 million humans who are directly dependent on livestock breeding for their subsistence live. In Africa, livestock breeding—mostly cattle and goats—is the main source of income south of the Sahara, especially in countries like Sudan, Chad, and Somalia. This is also the case in much of Asia, in particular in Mongolia, the central Asian republics of the former USSR, and Tibet. Cattle and sheep are also economically important domestic animals in New Zealand, Australia, Europe, and parts of the United States.

The domestication of animal species, especially of ruminants, and the associated land management and breeding practices are historically important ecosystem disturbance agents through the effects of grazing, trampling, and digging. They have transformed, shaped, and selected landscape and vegetation types, and have had significant, often dramatic, effect on many ecosystem properties, including plant and animal diversity. Uncontrolled grazing on shrublands and the creation of pastures have gone hand in hand with the creation of agricultural land. Native plant and animal communities have been greatly modified by the introduction of domestic animals, either directly or through habitat alteration and landscape change. There are many well-documented examples from all around the world, but those that occurred on islands are the most informative for our purposes.

The direct effects of the introduction of domestic grazers on native faunas since prehistoric times are well described for the Mediterranean islands, where original faunas have been affected by species extinction and introductions promoted by humans. The modern mammalian faunas on the big islands of the Mediterranean (e.g., Corsica, Sicily, Crete, and Cyprus) replaced the entirely endemic Holocene faunas that existed before human arrival. After the spontaneous extinctions of the late Pleistocene, the surviving endemic herbivores (e.g., *Episoriculus* spp; *Prolagus sardus*, *Rhagamys orthodon*, *Meridiopitymys henseli*) were rapidly eliminated, even before A.D. 7000, either intentionally by hunting or indirectly through the effects of commensal species. The faunas of today are the result of a severe human-induced selection in favor of species belonging to the geographic and cultural universe of the human groups that immigrated to the islands, within which domestic livestock predominate.

In southern Greenland, the arrival of the Norsemen in ca. A.D. 1000, and the introduction of cattle have disturbed the fragile dwarf-shrub heaths, which had established slowly after the deglaciation period. The activities of Norsemen, but especially of their cattle, destroyed the equilibrium between climate and vegetation, causing severe erosion of the unstable soil. When the Norsemen disappeared in the fifteenth century, a new phase of soil stabilization began, but this is now being severely disrupted with the immigration of more humans, now as sheep-breeders, in the beginning of the twentieth century.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00021-6

Madagascar's highland region was once covered with evergreen forests dominated by about 20 endemic tree species. Beginning around A.D. 600, Indonesian settlers started to remove forests to create swidden fields. At about A.D. 1000, zebu cattle were introduced from Africa, further increasing the need to expand grassland at the expense of forests. These forests were being permanently replaced by a floristically impoverished steppe vegetation on ferrolitic soils. By A.D. 1600, the highland forest had mostly disappeared: tree and humus removal had led to massive erosion, floods, water shortages, and faunal extinctions or endangerment.

Although much poorer in species richness, the faunas of the French subantarctic islands (Amsterdam, Saint Paul, Kerguelen, and Crozet Islands) show an analogous history following their discovery in the sixteenth and seventeenth centuries. Sheep, mouflon, and cattle are among the nine introduced species that thrive owing to lack of competitors and predators, and despite the small number of founding individuals. Herbivores have induced particularly significant changes in the nature and the structure of plant communities, leading to the extinction of endemic plants (e.g., *Phylica nitida*) or the degradation of the fragile peat-bogs that constitute the nesting sites of the rare Amsterdam albatross (*Diomedea amsterdamensis*).

In North America's Great Basin Desert, changes in plant communities that occurred after the introduction of domestic livestock in the late 1800s resulted in unusual, unforeseen cascade effects on the interactions between native mammalian species. The establishment and dissemination of cheatgrass (*Bromus tectorum*) in the Great Basin has played a central role in this process. Evidence suggests that cheatgrass was introduced accidentally as a grain contaminant at the end of the nineteenth century, at the same time that large-scale domestic animal grazing began. Imported from Mediterranean Europe and central and southwestern Asia, seeds of cheatgrass exploited an ecological niche, as no native annual was dominant in the Great Basin. Cattle, sheep, and feral horses facilitated its establishment, for they spread the seeds in the same areas that they disturbed. Once established, cheatgrass promoted the likelihood of fire to the detriment of the native species. In addition, other factors, such as the effects of the lack of vesicular-arbuscular mycorrhizae and selective lagomorph grazing, have worked in concert to further establish cheat grass dominance. The ecological consequences of this establishment have been an increase in fire frequency and intensity, a decrease in species diversity, and a landscape that is susceptible to severe erosion.

Ecosystem change resulted in the subsequent irruption of mule deer (*Ocodoileus hemionus*) and the expansion of mountain lions (*Felis concolor*). Domestic sheep depredation is currently increasing in western North America and is related to the expansion of suitable mountain lion habitat, and consequently of the lion's distribution and abundance. Furthermore, the expansion of this predator's population has caused the severe reduction of the porcupine (*Erethizon dorsatum*), another North American native species.

Although environmental determinism is not the dominant explanation of agricultural geography, the limits of the distributions of domestic animals, especially of their races, are often attributed from a physiological point of view to the physical environment, temperature, and water availability on a regional scale. However, it is also quite clear that the driving forces behind some of these ecological changes stem largely from cultural and social factors, including the size and growth of human populations, which determine the scale and the practice of stock-keeping activities.

Ecosystem-Level and Global Effects of Large Domestic Animals

The ecological role of cattle, sheep, and goats, that is, the effects they have on ecosystem processes and biodiversity, is far more complex than the fact that they remove biomass by eating grass and often browse, or feed on, twigs, shoots, and leaves of other plant species. The floristic composition and productivity of the grazed ecosystem and the persistence of vegetation and of plant and animal species, as well as physical and biogeochemical processes, may suffer substantial changes because of grazing animal-related factors, such as the severity and frequency of grazing, species of animal, method of prehension, treading, excreta deposition, and even the saliva deposited on plants. In some environments, the effects of grazing are quite predictable, whereas in others such effects are much more difficult to predict. Under certain conditions, grazing can act as a completely density- and species-independent disturbance, and under other conditions the effects on plants can be very selective and effects on soil properties and chemistry very heterogeneous in space.

Impacts on Vegetation

The impact of domestic ruminants on vegetation is typically studied through comparisons between grazed and ungrazed ecosystems. The dependent variables in these studies are descriptors of vegetation structure, commonly diversity and productivity. The independent variables are ecosystem or environmental parameters as well as grazing variables. Grazing variables are regulated by humans. Depending on different grazing regimes, the range of these variables extends from the unregulated (by humans) grazing of native species, to uncontrolled- or free-grazing of domestic ruminants, to overgrazing situations, and to optimized stocking rates for profitable animal mass gain per unit area of grazed land.

A meta-analysis of data on the effects of grazing on vegetation and soils, from more than 230 studied sites worldwide, has shown that the typical symptoms of the disturbances caused by domestic ruminants are changes in species composition, changes in dominant species, life-forms, and growth forms, changes in aboveground net primary production, and, finally, effects on the relationship between species, aboveground net primary production, root mass, and soil nutrients. In general, herbivory prohibits the most productive species from dominating and suppressing through competition the less productive ones. The conceptual model of Milchunas, Sala, and Lauenroth (1988) predicts the variation in plant diversity in relation to grazing intensity along gradients of moisture and evolutionary history of grazing (**Table 1**). According to this model, grazing should have a greater effect on species composition in more humid areas because adaptations of tall growth forms capable of competing for light in a dense

Table 1 The plant diversity–grazing intensity relationship under various moisture and grazing history conditions, according to the predictions of the conceptual model of Milchunas, Sala, and Lauenroth (1988)

Moisture gradient			
		Semi arid	*Subhumid*
Evolutionary history of grazing	long	monotonic (linearly decreasing)	unimodal symmetric (parabolic)
	short	unimodal with steeply decreasing part	left skewed unimodal

Source: Reproduced from Milchunas DG, Sala DE, and Lauenroth WK (1988) A generalized model of the effects of grazing by large herbivores on grassland community structure. *Am. Nat.* 132(1): 87–106, with permission from University of Chicago Press.

canopy are opposite to those that provide resistance to or avoidance of grazing. In contrast, plant adaptations to frequent loss of organs from drought or herbivory are similar, and under arid and semiarid conditions, competition is primarily for belowground resources. The explanation of the evolutionary history of grazing gradient in the conceptual model is that increasing grazing history over evolutionary time results in greater capacities for regrowth following herbivory, and thus favors prostrate growth forms. In communities of long evolutionary history and high aboveground net primary productivity, grazing causes rapid shifts in the dynamic balance between suites of species adapted to either grazing avoidance/tolerance or competition in the canopy.

Field research offers evidence that herbivory by domestic ruminants may cause either an increase or a decrease in plant diversity. For example, in salt marsh habitats, heavy grazing eliminates sensitive species and produces a dense cover of graminoids in upper marsh coastal habitats. In other marshes, grazing produces bare patches that allow annuals and other low marsh species to invade upper marsh zonal communities. A retrogression in plant succession may occur in salt marshes and salt deserts because of heavy grazing. Intermediate levels of grazing by sheep, cattle, and horses could produce communities with the highest species richness and heterogeneity. At the other extreme of the aridity gradient, in a seasonally dry tropical savanna, species diversity was higher in a grazed area than in the neighboring area where grazers were permanently excluded. Furthermore, similarity in species composition between the grazed and ungrazed areas was very low.

The introduction of domesticated animals to a pasture or rangeland invariably causes changes in species composition, leading to an increase in the abundance of those species that are less palatable to animals. Low-growing, prostrate growth forms are selected by grazing, and annuals and shrubs appear to be favored. Tall perennials decrease in abundance because this growth form, which offers competitive advantages in dense canopies free from grazing conditions, exposes the plants to selective herbivory pressure in grazed ecosystems.

Between sites, aboveground primary productivity is negatively affected by grazing in the most productive ecosystems. Within sites, lowlands are more affected than uplands. More productive lands are more likely to be preferred by grazing animals and breeders.

Grazing has negative impacts on root systems. Intensive grazing removes the product of vegetative growth, which in turn reduces root growth. Under drought conditions, shallow roots cannot fully exploit subsoil moisture reserves, and the plant becomes stressed.

Erosion, Desertification, and Land Degradation

Cattle are important agents of geomorphological change. The animals' impact on the landscape is to create bare soil by weakening the vegetation cover and then by breaking this cover down by trampling. On uplands, heavy grazing compacts the soil, reduces infiltration, increases runoff, and increases erosion and sediment yield. However, light and moderate grazing has effects that are much less significant. In riparian zones, grazing decreases erosional resistance by reducing vegetation and exposing more vulnerable substrate. Trampling directly erodes riverbanks, thus increasing turbulence and consequent erosion. Trampling also maintains and expands the area of bare soil upon which frost, rain, and wind act.

Livestock keeping when practiced in a nonsustainable way is one of the major causes of desertification. Desertification is defined as land degradation in arid, semiarid, and dry subhumid areas resulting from temporary climatic crises, especially droughts that occur periodically, and harmful human activities in vulnerable ecosystems. Land degradation leads to a reduced capacity of dryland areas to produce useful outputs—crops, fodder for grazing livestock, bush and tree cover—or to sustain wildlife. Degradation of dryland areas involves a range of processes: the erosion of soils through water and wind, falling levels of soil fertility and damage to soil structure, loss of vegetation cover and change in species composition, reduced availability and decline in the quality of water supplies, loss of wildlife, and a decline in the biological diversity of plant and animal life. Such processes reduce the productivity of crops and livestock systems in dryland areas, and increase the vulnerability to food crisis of populations depending on these resources.

The processes of desertification were first recognized in the Sahelian region of West Africa. Overstocking and overgrazing, firewood collection, and cultivation of unsuitable soils are responsible for 80% or more of the desertized lands in Africa. However, dryland degradation is now accepted as a worldwide problem. Over the last decade, surveys have been carried out on a global level to assess the extent of desertification and soil degradation by region, and they have produced significantly differing results. The GLASOD survey (Global Assessment of Soil Degradation), commissioned by the United Nations Food and Agriculture Organization (FAO), showed that 19.5% of drylands worldwide were suffering from desertification. In contrast, the United Nations Environment Programme (UNEP) usually quotes an estimate of 70% of the world's dryland areas as suffering from some degree of desertification, with an estimated 900 million people worldwide at risk from

problems of degradation. This survey includes not only areas affected by soil erosion, but also where a change in vegetation had occurred (e.g., where perennial grasses had been replaced by annuals). All regions of Africa have been affected by drought conditions and human pressures on land, as have parts of Mediterranean Europe, North America, Asia, and Latin America. In Spain alone, data from 1993 suggest that almost 1 million hectares (ha) of land are already considered as desert lands and another 7 million ha have been identified as being at high risk of desertification. In the United States, 90 million ha are considered to be affected by desertification.

The underlying factors that cause such adverse effects are many and various, and operate at different levels. At the local level, for example, inappropriate range, water, and livestock management practices may accelerate rates of erosion, especially on sloping land. Information continues to accumulate on the effects of digging deep boreholes in several regions of Africa. These boreholes discharge several liters of water per second, without the enforcement of any range management or land use policy, and have too often resulted in large concentrations of livestock (20,000 to 40,000 head) during the dry season. This destroys the range in the vicinity of the well over a radius of 20 km (125,000 ha) in one or two seasons (the stock rate being 10–15 times the carrying capacity).

At the national level, government policies on land tenure and use determine whether production processes will be sustainable or not. During the second half of the twentieth century, the Autonomous Russian Republic of Kalmykia has undergone severe desertification. Under Soviet rule, rangelands were increasingly devoted to animal production, and pastures were converted to cropland in a campaign to increase crop productivity. Pastures were grazed at rates that were two or three times their sustainable production, saiga antelope (*Saiga tatarica*) populations and habitat greatly decreased, more than 17 million ha were subjected to wind erosion, 380,000 ha were transformed into moving sands, and 106,000 ha were ruined by secondary salinization and waterlogging. By the 1990s, almost 80% of the Republic had undergone desertification, and 13% had been transformed into a true desert.

At the international level, rising prices provide encouragement to breeders and farmers to produce more and, at the same time, raise revenues that can be invested to increase capacity. The African arid zone harbors about 55% of that continent's 550 million head of livestock, and these livestock numbers have increased by 75% from 1950 to 1985, in spite of severe droughts that occurred in most arid zones in the early 1970s and 1980s. This is exponential growth of about 0.7% per year, compared to the 1.0–1.5% demographic growth rate of the African pastoralist population.

Nutrient Cycling and Over-Enrichment in Grazed Ecosystems

Nutrient cycles in a grazed ecosystem do not differ from those in an ungrazed one in terms of individual element pools, flows, potential input and output pathways, and interrelationships among various nutrient pools (**Figure 1**). On the contrary, the sizes of the pools, the rates of biogeochemical processes, the flows, and the residence time in the various pools of elements and nutrients are greatly affected by herbivores and the breeding practice and stock size. For example, in natural pastures, animals remove aboveground plant biomass (shoots, stems, and leaves) and deposit excrement; in managed pastures, where grazing intensity exceeds the natural carrying capacity, additional herbage and other fodder is provided to the animals and, thus, additional inputs of nutrients are entering the system. The cycle of nitrogen, most likely the most perplexing of all nutrients, is certainly disturbed under grazing conditions. In natural pastures, less litter returns to the soil, the topsoil organic matter declines, and the

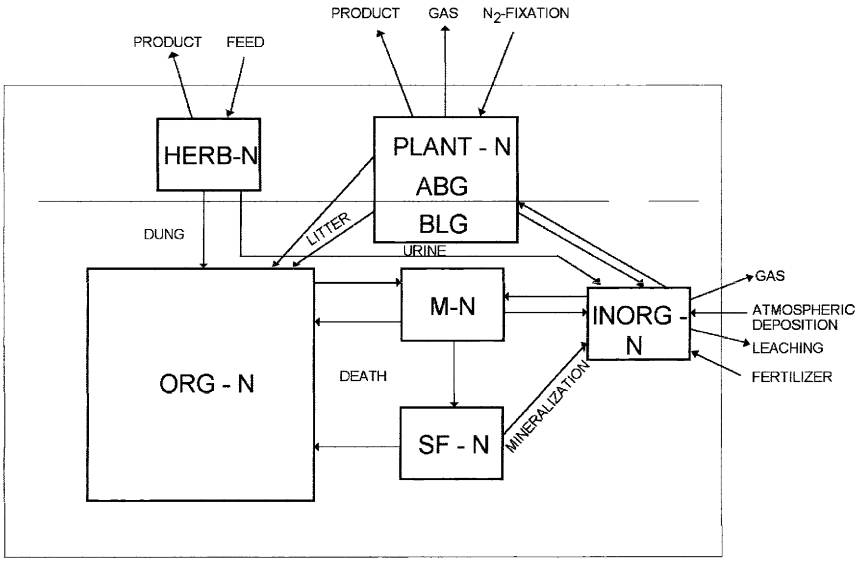


Figure 1 Pools and fluxes of nitrogen (N) in a typical pasture. Nitrogen pools are shown in their relative average sizes: ORG-N, organic soil N (about two orders of magnitude larger than other pools); M-N, microbial biomass N; SF-N, soil faunal biomass N; INORG-N, inorganic (NH_4^+, NO_2^-, and NO_3^-) soil N; PLANT-N, plant biomass N (ABG: shoots; BLG: roots); and HERB-N, large herbivore biomass N. The large rectangle represents field boundaries. For clarity, not all directions of flux are shown.

nutrient cycles are quantitatively disturbed. Grazers shorten the N cycle, increase the rate of N cycling, and cause significant redistribution of N, among pools and in space. On a daily basis, cattle excrete a somewhat constant amount of N in feces (8 lb/1000 lb of dry herbage consumed), the remainder being excreted in urine. Because they excrete more at night than during the day, bedding or watering places receive higher quantities of excreta-carried nutrients. Cattle and sheep retain a small fraction of the N consumed, and thus they have an impact on the control of N cycling that is out of proportion to their low biomass on an areal basis.

Areas of pasture that receive dung and urine may undergo marked changes in botanical composition. Urine spots and fecal pats contain the equivalent of 500 to 1000 lb N/acre, which is of different availability to plants: the N in dung is mainly in organic form, with a slow overall mineralization, whereas in urine N is present as urea, which is hydrolyzed rapidly to forms available to plants. The plant recovery of nutrients from excreta spots is rarely greater than 30%. Generally, the nitrogen in urine stimulates grass growth and the phosphorus in dung stimulates legume growth, especially on P-deficient soils. However, plants immediately beneath dung pats may be killed and urine occasionally scorches them.

The cycles of nitrogen, phosphorus, and other nutrients are not generally closed at the field or farm level. For nitrogen, six main pathways of loss have been described in grazed ecosystems: NH_3 volatilization, denitrification, wind and water erosion, NO_3^- leaching, fire, and incorporation into animal biomass (exported or retained *in situ*). The relative importance of these pathways depends on environmental conditions and breeding practices. In humid and subhumid ecosystems, N losses are related to leaching, whereas volatilization is most important in semiarid regions, and the export of animal products and erosion in deserts.

Runoff from agricultural land and livestock feedlots is among the major "nonpoint" sources of nutrients entering marine and freshwater ecosystems, causing pollution and eutrophication. This has been shown on very large scales, as in the United States, where these widely dispersed activities are the major source of water pollution, as well as at the level of local ecosystems, such as lakes, rivers, and estuaries. In aquatic ecosystems, overenrichment with P and N causes a wide range of problems, including toxic algal blooms, loss of oxygen, fish kills, loss of seagrass beds and other submerged aquatic vegetation, degradation of coral reefs, and loss of biodiversity—including species important to commercial and sport fisheries and shellfish industries.

Lake studies allow contemporary sediment and nutrient dynamics to be placed in a historical context so that trends and rates of change in catchment inputs may be calculated. Surface runoff from heavily grazed grassland has a high suspended sediment, ammonium nitrogen, and particulate phosphorus load. The combined effect of the long-term increase in the organic loading from livestock and the inorganic N and P load from fertilizers may be the source of nutrient enrichment in lakes. The literature offers several examples of eutrophication studies in lakes and streams; for example, in the already extinct karstic Jastericie Lake in Slovakia, in the coastal freshwater systems of Slapton Ley and Loe Pool in southwestern England, and in the streams draining the Mount Lofty ranges in South Australia, livestock breeding is the main source of organic N and P that affects the trophic status of surface waters.

Intensive animal production generally involves feeding large numbers of animals in small areas. For example, 4% of the cattle feedlots in the United States produce 84% of the cattle. Nutrients in manure can be recycled by applying the manure to cropland. However, the amount of manure generated by concentrated livestock operations often far exceeds the capacity of nearby croplands to use and retain the nutrients. Thus, excess fertilization and manure production cause a surplus of N and P, which accumulate in the soil. Some of this surplus is transported in soil runoff to aquatic ecosystems. Especially for N, the surplus is mobile in many soils, and much leaches into surface waters or percolates into groundwater. Indeed, the density of animals on land is directly related to nutrient outflow to aquatic ecosystems. Surplus N can also volatilize in nitrous (NO_2) and ammonia (NH_3) forms to the atmosphere. The emission of ammonia from stored and land-applied manure can result in a significant loss of nitrogen for crop production. A high atmospheric concentration of ammonia can result in the acidification of land and water surfaces, cause plant damage, and reduce plant biodiversity in natural systems.

Nitrous oxide, together with carbon dioxide and methane, are considered to be major greenhouse gases. Besides emissions of these gases from the plant–soil compartment of the pasture soil system, CO_2 and CH_4 emissions from cattle and their excreta are significant components of the global fluxes. Methane emitted from dung is a very important contributor to the global methane budget; the corresponding figure is as high as 20%. Grazing animals on managed pastures and rangelands have also been identified recently as significant contributors to the global N_2O budget. This occurs because of the concentration of herbage N in urine and dung patches, and by the compaction of the soil due to treading and trampling. The limited amount of experimental data indicates that 0.1 to 0.7% of the N in dung and 0.1 to 3.8% of the N in urine is emitted to the atmosphere as N_2O. Integral effects of grazing animals have been obtained by comparing grazed pastures with mown-only grassland. Grazing-derived emissions, expressed as a percentage of the amount of N excreted by grazing animals in dung and urine, range from 0.2 to 9.9% with an overall mean of 2%. Using the emission factor and data statistics from FAO for numbers of animals, the global contribution of grazing animals is estimated at 1.55 Tg N_2O-N per year.

Biotic Relationships of Large Domestic Animals

Reactions of Plants to Grazing

The direct act of grazing represents a loss of organs or parts of organs to individual plants and an alteration of canopy structure to the community. Resistance of plants to grazing involves avoidance and/or tolerance mechanisms.

Plant tolerance reactions to the grazing event will depend on the capacity of individual plants to compensate for lost organs and the relative impact of the removal on competitive

relationships in the canopy. Various species of animals graze differently because of their prehension organ anatomy and method, behavior, and diet preferences. For instance, cattle graze individual plant parts (leaves or stems) less selectively than do sheep and goats. The literature offers examples of strong interactions among intensity, severity, and other properties of grazing of different animal species and pasture (species) in terms of various components of plant fitness. Especially for grass species, the height, tiller number, survival, and reproduction are the fitness components that are most usually monitored in grazed plots/individuals and ungrazed controls.

Three alternative hypotheses on the effects of grazing intensity on plant growth and fitness have been proposed (**Figure 2**). The first hypothesis predicts that net primary productivity shows a consistent decline as the intensity of grazing increases. The second hypothesis predicts that plants compensate for tissue removal up to some level, beyond which productivity begins to decline. The third hypothesis, known as "overcompensatory growth," states that within some levels of herbivorous feeding, plant productivity may be enhanced before declining beyond a threshold of grazing intensity. This hypothesis has generated an ongoing debate on the controversy of plant response to grazing, such as herbivore optimization and overcompensation.

Observations of mixed cattle and American elk grazing in high-elevation rangeland conditions in southwestern North America show that vegetation has deteriorated. Experiments in the North American tallgrass prairie suggest that overcompensation is a nonequilibrium plant response to grazing. Photosynthate that would be stored as reserves and used for root growth and flower and seed production is instead used to replace lost leaf area, thereby resulting in higher foliage productivity. However, under chronic grazing or mowing, vegetation is prevented from maintaining high nutrient and water uptake capacity (large root biomass) and accumulating reserves that allow overcompensation responses.

Herbivory by large animals is known to function as a selection pressure to increase herbivory resistance within plant populations by decreasing the frequency of genotypes possessing large erect canopies. Data on the trade-off between herbivory resistance and competitive ability in *Schizachyrium scoparium* confirm that herbivory by domestic cattle may function as a selection pressure to induce architectural variation in grass populations within an ecological time frame (ca. 25 years).

Comparative studies between species suggest that there is no consistent pattern of grazing effects on survival, reproduction, recruitment, and regeneration. In one study, five abundant native Kansas tallgrass prairie perennial forbs (*Baptisia bracteata*, *Oenothera speciosa*, *Vernonia baldwinii*, *Solidago missouriensis*, and *Salvia azurea*) were chosen to examine the effects of native (bison) and domestic (cattle) ungulates on plant growth and reproduction. The results show that their responses are complex and vary significantly among plant species, ungulate species, and plant life-history stages. *Baptisia bracteata*, *O. speciosa*, and *V. baldwinii* increased in growth and reproduction in grazed sites, indicating competitive release in response to selective grazing of the dominant warm-season matrix grasses. Species with reduced performance in grazed sites are likely to be negatively affected by disturbances generated by ungulate nongrazing activities, because none of the forbs studied was directly consumed by bison or cattle. Furthermore, the native and domestic ungulates differ significantly in their effects on forb growth and reproduction.

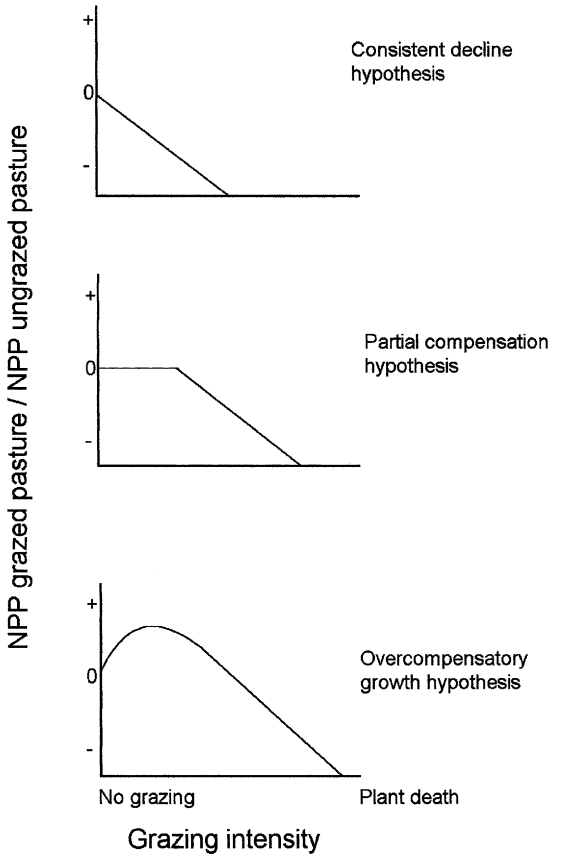


Figure 2 Theoretical relationships between the variation of net primary productivity (NPP) of grazed and ungrazed grassland pastures and the grazing intensity. Lines represent possible trends in NPP as grazing intensity increases, according to the three examined hypotheses: consistent decline hypothesis, partial compensation hypothesis, and overcompensatory growth hypothesis.

Among 18 species of shrubs and trees in southeastern Australia, 10 showed significant negative effects on recruitment and/or regeneration from present or past sheep grazing. In this case it was shown that the negative effect of sheep grazing on recruited seedlings must exceed that of natural thinning before overall regeneration is affected. Evidence that grazing by sheep and goat increases the total inclusive fitness in the Mediterranean shrub *Anthyllis cytisoides* has been reported: in this species, moderate grazing promotes growth, stability of vegetative structures, and adult survival, and a drop in seed production. Direct consumption of reproductive organs has significant effects on the overall fitness of the grazed species. In the annual wild wheat *Triticum dicoccoides*, removal of maturing inflorescences by cattle reduces its fitness by 50%, estimated as the number of seeds produced per seedling. In this species, protection from grazing leads to a significant increase in the number of individuals that produce mature inflorescences as well as in the number of spikelets per

inflorescence. In the clonal species *Yucca elata*, cattle browsing of inflorescences may reduce reproductive effort, but the most significant changes in population structure are due to the browsing of small caudices, including both genets and ramets. Thus, the expansion and dominance of ''resistant'' species on heavily grazed land may be a result not only of reduced competition, as is commonly assumed, but also of enhanced performance of some of these species.

Grazing avoidance mechanisms reduce the likelihood of defoliation by animals. For example, leaf surface chemistry of plants is related to the extent of defoliation by livestock. Data on cattle, sheep, and goat herbivory on tarbush (*Flourensia cernua*), an abundant but generally unpalatable Chihuahuan Desert shrub, support the hypothesis that secondary chemicals—mostly terpenes and phenolic compounds—in its resinous leaves may influence the diet selected by the ruminants. Although individual leaf surface compounds do not appear to greatly affect the degree of use of the plant by livestock, collectively these compounds may partially explain the differential herbivory on tarbush plants by livestock. Mimicry, both chemical and morphological, has also been suggested as a grazing avoidance mechanism for plants. Experimental tests with sheep showed that odor alone is not persistently effective in preventing herbivory, but that both taste and odor must be similar for one plant to successfully mimic another.

Spines and thorns are considered to be defense structures against herbivory by both wild and domestic animals. Experimental evidence from *Acacia drepanolobium* suggests that spine length is an inducible defense, with longer spines being produced by branches experiencing a greater level of herbivory. Examination of *Acacia* trees protected from herbivory for several years suggests that reduction in spine length eventually exceeds 70%. The effectiveness of spines and thorns as an antiherbivore defense in several plants in arid Australia did not vary with the evolutionary history of the herbivores (i.e., wild versus domestic). Furthermore, additional evidence shows that the interaction of ants of the genus *Crematogaster* and thorns of *A. drepanolobium* is a means of defense against browsing goats. This interaction causes the animal to stop feeding almost immediately, therefore keeping the amount of foliage lost to a minimum. It is hypothesized that the acacia–ant relationship evolved partly because of the pressure from browsing herbivores.

Seed Dispersal and Germination

Throughout the world, the presence of wild or domestic ruminant herbivores is correlated with the maintenance of high levels of plant diversity in natural and semi-natural grazed systems, where the grazing pressure does not exceed the carrying capacity of the vegetation. Zoochorous dispersal by herbivorous mammals has been verified repeatedly and its possible influence on the structure, function, and diversity of plant communities, mainly herbaceous, has been suggested.

Cattle, sheep, and goats are important vectors for endo- and epi-zoochorous seed dispersal. On large scales, seed dispersal systems associated with domestic ruminants have been proved to be particularly favorable for the introduction of alien plant species, and herbivores have facilitated the naturalization and spread of many alien herbaceous species from their initial points of introduction. The case of central Iberian herbaceous species introduced into the Mediterranean-type zone of Chile is very relevant: almost 15% of the central Iberian herbaceous species are naturalized in Chile, with the endo- and epi-zoochorous species representing 21 and 23% of them, respectively. The expansion of the distribution of individual species is often correlated with the effects of cattle ingestion on the viability and germination rates of seeds. For example, the calden (*Prosopis caldenia* Burkart) is the dominant tree of the xerophytic open forest in the semi-arid pampa of Argentina. Calden has gradually increased its distribution throughout the region during the past century as a result of livestock grazing in the pampa forest. Excreted calden seeds display a range of delayed germination responses. This variation would increase the probability of seed germination for a variety of environmental and site conditions.

The potential dispersal distances for adhesive seeds attached to the fur of cattle range from tens of meters to kilometers. The morphology of the seed's structure and its position on the animal's body influences the length of time that it remains attached to the fur.

Endozoochorous dispersal selects for traits that enhance ingestion and passage of viable seeds through the animal. Passage of buffalo grass (*Buchloe dactyloides*) seeds, one of the two dominant grasses of the North American shortgrass prairie, through cattle has a positive effect on germination and seedling growth from intact diaspores; the damage due to mastication is minimal, and the retention time is from 1 to 5 days. This combination of retention time and movements of the animals influences the spatial expansion, distribution, and abundance of this species. Clear experimental evidence that the germinability of seeds increases significantly following their passage through the cattle gut has been provided for the legume *Biserrula pelecinus*, which is greatly favored by its dispersal through cattle dung.

The amounts and diversity of viable seeds contained in the feces of domestic ruminants grazing in nature may be quite high. Based on germination trials, the number of seed species germinated in the dung of feral cattle feeding in savannas and floodplain wetlands in Rajasthan, India, was about 450 seeds per m^2, belonging to 35 species. Similar results have been obtained in Mediterranean open woodlands, where cow dung may yield as much as 70 seeds per gram of dry dung from a large number of species (about 75 species). Although ruminant species differ in their traits and feeding habits, dispersal seems to be mainly determined by seed production of the plant community.

Large isolated trees are a common feature of the agricultural landscape in humid tropical regions that were originally covered by rain forest. These isolated trees are used as a source of shade for cattle, but they also function as nursery plants for rain forest species by facilitating the establishment of zoochorous species whose seeds are deposited under their canopies by cattle, frugivorous birds, and bats. The same pattern is observed in Botswanan savannas: the analysis of seed pools under the canopy trees *Dichrostachys cinerea* showed that epizoochorous species such as *Tragus berteronianus* were dominant at the cattle resting sites under trees.

Patch dynamics are also affected by seed zoochorous dispersal. In Mediterranean pastures, cattle-dispersed endozoochorous seeds germinate in manure and colonize the dung patches. The micro-succession involved is independent of the type of pasture. A small-scale spatial pattern results in which gaps of old dung are dominated by endozoochorous species. Thus, dung patches enhance the similarity between different communities when they are grazed, but also increase the variation within communities. In Australian sub-alpine grassland vegetation, the density of shrub seedlings in gaps varied considerably in space, but it is related to the dispersal of seeds and the trampling and browsing effects of domestic cattle.

Large seeds such as *Quercus suber* acorns are rapidly predated by large herbivores. However, in this case, experimental data seem to suggest that seedling emergence rate is inversely related to the intensity of predation on the acorn bank.

Domestic Animals as a Food Source for Wild Predators

Domestic ruminants have largely replaced wild grazing animals over large areas on all five continents. Natural predators and scavengers of large mammals have easily adapted to this alternative food source. Population densities of cattle, sheep, and goats are equivalent and often higher than the densities of the wild grazers that have been replaced, as the animal keepers' tendency is to maximize herd or flock size and food may be supplemented at times of natural fodder shortage. Furthermore, domestication has usually produced heavier, slower-moving, less alert, and less experienced animals, which often have their ability to move impaired by various means and structures such as pens, fences, and tethering. Thus, domesticated animals are generally speaking easy prey for most natural predators. The way of life of pastoral societies around the world has been profoundly influenced by the constant strive for protection of their flocks and herds from natural predators.

Wolves (*Canis lupus*) are legendary in Western culture as predators of cattle, goats, and especially sheep. Today they have been exterminated or reduced to small populations over much of their former range in the Northern Hemisphere, but where they do exist they are still important predators of livestock. In regions like parts of southern Europe, where stock-raising is prevalent and wild prey are rare, their presence is still connected with the raising of stock animals. The wolves either prey directly on live individuals, most often stray ones, or feed on carcasses and offal. Similarly, dingos (*Canis familiaris*) can be serious predators of sheep in Australia, as are coyotes (*Canis latrans*) in North America. In Africa, Asia, and North and South America, large cats such as leopards (*Panthera pardus*), tigers (*Panthera tigris*), and pumas (*Puma concolor*) commonly prey on cattle, sheep, and goats. Big cats and the other canids differ from wolves in that their existence does not seem to rely to any significant extent on domestic animals. The big cats are especially dependent on extensive tracts of undisturbed habitat where natural prey is available, and consequently domestic animals are consumed opportunistically and usually only by certain individuals.

In many parts of Eurasia, both golden eagles (*Aquila chrysaetos*) and lammergeiers (*Gypaetus barbatus*, German for "lamb vulture") have traditionally been regarded as significant predators of lambs, and this has been the reason for their local extermination. In reality they kill very few lambs but commonly feed on dead lambs or the carcasses of older sheep. Both Old and New World vultures also include dead cattle, sheep, and goats in their diet. As in the case of the wolf, in parts of the world where wild ungulates are now rare, such as around the Mediterranean, these sources of food are essential for the survival of vulture populations. Although griffon vultures (*Gyps fulvus*) are not migratory, they perform movements between summer and winter quarters, closely following flocks of sheep as they move from high-altitude summer pastures to lowland grazing land in winter. For the larger eagles and vultures, open range grazing has a dual functionality in providing a food source (essential or supplemental) and by indirectly maintaining open habitats where such species are able to forage.

Cattle, sheep, and goats appear in the diet of the majority of other medium or large natural predators such as bears (*Ursus* spp.), foxes (*Vulpes* spp.), jackals (*Canis* spp.), other eagles (*Aquila* spp.), and buzzards (hawks, *Buteo* spp.) to a varying extent, depending on the region, the species of predator, and the method of stock-keeping. In most such cases, carcasses, offal, and still borns are consumed rather than healthy, full-grown adults.

Domestic Animals and Invertebrate Fauna

Both positive and negative effects of cattle, sheep, and goats on invertebrate communities of grazed ecosystems have been reported. In grazing systems where the carrying capacity of vegetation is not exceeded by grazing, as in the traditionally managed European mountain meadows or mixed woodlands, the maintenance of a landscape mosaics favors the persistence of high plant diversity and abundance of flowers and has positive consequences on insect population dynamics. Bumblebees, butterflies, syrphids, and other insect groups foraging on flowers depend on the landscape management in which domestic grazers play an important role. In central Spain, traditional landscape management for non-intensive grazing by goats and cattle favors the conservation status of the butterfly *Euphydryas aurinia*, which depends on patches of open oak woodlands mixed with open areas, where important nectar sources and larval foodplants are present. However, indirect effects on pollinators may appear from the removal of flowers by grazing: in *Yucca elata*, cattle preferentially consume inflorescences, which are found to be highly nutritious. This implies the possible local extinction of the yucca moth *Tegeticula yuccasella*, the exclusive pollinator of the plant.

In the shortgrass prairie of Colorado, in the United States, heavy grazing negatively affects the colony density and distribution of harvester ants (*Pogonomyrmex occidentalis*) on a broad scale, with important consequences to behavioral, community, and ecosystem processes. Browsing by goats has been proven to increase the damages from both tunneling (Diptera) and "blotch-making" (Lepidoptera) leaf miners on trees growing on the Aldabra Atoll.

Insects are particularly sensitive indicator taxa of land use (especially the Cicindellidae, Staphylinidae, and Carabidae).

Domestic cattle influence the arthropod diversity through trampling, its intensity being more important than the type of trampling. In semi-arid tropical soils in Queensland, Australia, heavy grazing significantly affects Acari populations as well as the diversity and activity of termite species. The deterioration of soil hydraulic properties associated with cattle trampling at high stock rates is responsible for these negative effects on detritivorous termite activity in the topsoil.

A wide range of invertebrates, as well as bacteria and fungi, are involved in the breakdown and eventual decomposition of dung. Some are facultative generalist decomposers of organic matter, but there are also a large number of specialist dung consumers, such as dungflies (Scathophagidae) and dung beetles (Geotrupidae, Scarabaedae). The array of dung-feeding invertebrates present in a given area is mainly dependent on the characteristics of the dung, that is, on the taxonomy of the dung producer and its diet rather than on a distinction between wild and domestic stock.

Domestic Animals and Birds and Small Mammals

Excluding predators and scavengers, the main influence of cattle, sheep, and goats on other animals occurs indirectly through the creation and maintenance of open pasture land, in particular in areas where the natural vegetation cover would be forest or scrub. Tall-grass meadows, short turf on stabilized sand dunes, alpine meadows, dry stony pastures, and all other forms of grazing land that exist on different soils and under different climatic regimes are all attractive to species that would otherwise be restricted to steppe, desert, or arctic habitats on a global scale or to the margins of wetlands, beaches, steep hill slopes, and woodland clearings on a more local scale. The creation of pastures has gone hand in hand with the creation of agricultural land, which has also provided vast areas of open habitat. The expansion of both pasturelands and croplands began several thousand years ago, and the process of colonization and spread of open-country species into such newly created habitats is lost in prehistory and is largely a matter of conjecture.

Animal species closely connected with grazing land include mammals such as moles (*Talpa* spp.) and voles (Microtinae), birds such as larks (Alaudidae) and plovers (Charadriidae), and insects such as grasshoppers (Acrididae) and ants (Formicoidea). Grazing is often employed as a management tool, for example, to create the very short turf required by wintering white-fronted geese (*Anser albifrons*) in Britain.

Other species may use cattle, sheep, or goats more directly in feeding. Two well-known cases are cattle egrets (*Bubulcus ibis*) and yellow wagtails (*Motacilla flava*). Cattle egrets commonly use cattle, as well as other domestic and wild ungulates, as a vantage point and vehicle, capturing large insects disturbed by the cattle as they move. Yellow wagtails may also do the same or they may feed on flies and other insects attracted to the animals themselves or their dung.

Negative impacts on other animal species may also occur. For example, when natural grazing land is used, domestic ruminants may compete with wild herbivores. Although wild herbivores may suffer, human interest is usually focused on the reverse impact of wild herbivores on domestic ones [e.g., rabbits (*Oryctolagus cuniculus*) and sheep in Australia].

Another negative impact on wildlife is the trampling of eggs and young of ground-nesting birds. This problem is more frequent in wetland habitats, where most bird species nest on the ground, and may be particularly damaging for colonially breeding species such as terns (Sternidae) and avocets (*Recurvirostra* spp.). Although these birds generally select islets or other sites that are inaccessible to ground predators (e.g., foxes) and thus also inaccessible to livestock, problems may be caused when water levels change or when a herd or flock is intentionally transferred to such a site.

Domestic Ruminants as Hosts for Parasites and Pathogens

Domestic ruminants harbor significant numbers of internal and external parasites. Helminths, protozoa, and arthropods (especially ticks and mites) are the most common components of the parasitic communities, although many other species of occasional or temporary parasites, such as blood-sucking Diptera, also infect domesticated animals. Endoparasites, mainly gastrointestinal parasites, have received most attention in veterinary research because of their prevalence as disease-causing organisms that are economically important in rearing commercial livestock. Within a species, the taxonomic diversity of parasites varies between the various organs of the hosting organism. Thus, the digestive tract, the liver, the circulatory system, the respiratory tract, the skin and the subcutaneous tissue, the muscles and the tendons, the eyes, the central nervous system, and the serous cavities are affected by quite distinct species of parasites such as trematodes, cestodes, nematodes, protozoa, and arthropods. For example, in sheep and goats, the list of gastrointestinal parasites comprises 15 trematodes, 9 cestodes, 88 nematodes, and 25 protozoa; the parasitic fauna in the skin and the subcutaneous tissues comprises only 6 nematodes but 78 arthropods. Parasitic communities also differ between host animal species, even closely related races. Likely causes are differences in host suitability and feeding differences that affect the probability of transmission.

Besides the specificity of the host–parasite interactions, the diversity of the domestic ruminants' parasite communities is strongly influenced by species–area and species–climate relationships. As an example, for nematodes it has been demonstrated that (a) large areas of permanent pasture include a variety of microenvironments that are favorable to the development of the free-living stages of various species of nematodes and (b) free-living stages are very susceptible to dryness and survive better in areas with heavy rainfall.

The widespread use of veterinary chemicals against endoparasites, pests, and other pathogens of domestic animals characterizes the high-input livestock breeding and production strategy. Avermectins, a relatively new class of broad-spectrum pesticides, are used widely to control livestock parasites. Following veterinary treatment, avermectins are eliminated in the livestock feces. The dung mesofauna potentially exposed to avermectin (or other active compound) residues includes insects, earthworms, springtails, mites, and nematodes. The effects range from acute toxicity in larvae and

adults, through disruption of metamorphosis, to interference with reproduction. For example, at high drug concentration in the dung, larvae of the dipteran Cyclorrhapha are killed or paralyzed, while at lower levels their metamorphosis is inhibited. At very low concentrations of avermectins, well below levels occurring in feces after routine treatment, adult emergence is reduced and a significant number of imagines show morphological abnormalities. Nematoceran Diptera are less sensitive than Cyclorrhapha, but larval and pupal development are affected at higher dose levels.

Dung mesofauna occupy a variety of different niches within the ecosystem and the faunal composition changes as the pats age. Some members of this fauna act in concert with soil microbial flora and assist in the breakdown of dung and consequent nutrient recycling on pasture-lands. A retardation in the rate of loss of biomass of dung pats from avermectin-treated cattle has been observed following the various forms of drug administration. Rare insects could be put at risk by the use of avermectins, especially those that breed exclusively in the dung of the herbivores on which avermectins are used.

The use of avermectins may also indirectly affect some species of vertebrates by depleting the quality and quantity of important food resources. The effects of any reduction in invertebrate food in livestock dung would be expected to be especially severe if it occurred at critical times for the vertebrates, such as during the breeding season or when newly independent young animals were foraging and fending for themselves. Insects that develop in livestock dung therefore have important, additional roles in the ecology of pasture-lands other than aiding dung degradation processes. Livestock dung is an important feeding habitat for a number of vertebrate species. The potential for direct poisoning of vertebrates through the accumulation of avermectins following consumption of invertebrates containing residues would, on present knowledge, appear to be limited, but it should not be disregarded.

This issue has generated much controversy regarding the extent of the effects on pasture and rangeland biodiversity and ecosystem functioning. It has been suggested that ecotoxicological studies commonly disregard the veterinary use patterns of drugs and consequently they overestimate their negative effects at a populational level. Avermectins administration patterns in temperate regions indicate that peak periods of insect activity and peak times of avermectin use are often asynchronous. When avermectin usage and insect activity do coincide, the heterogeneous patterns of administration to livestock and the focus of treatment on young animals result in the deposition of feces that are predominantly free of avermectin residues. Results of large-scale, long-term studies indicate that, even under conditions of relatively high levels of avermectin use in cattle, the impact on non-target insect populations and their function is limited.

See also: Agriculture, Traditional. Desertification. Grazing, Effects of. Greenhouse Effect. Range Ecology, Global Livestock Influences

References

Clarholm M and Berstrom D (eds.) (1989) *Ecology of Arid Lands: Perspectives and Challenges*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Evans R (1998) The erosional impacts of grazing animals. *Progr. Phys. Geogr.* 22(2): 251–268.

FORUM (1993) Grazing theory and rangeland management. *Ecol. Appl.* 3(1): 1–38. (Papers by SA Levin; EL Painter and AJ Belsky; MI Dyer, CL Turner, and TR Seastedt; SJ McNaughton; MJ Trlica, and LR Rittenhouse; DD Briske; JW Bartolome; DL DeAngelis and MA Huston; I Noy-Meir; DT Patten; and PG Risser.)

Milchunas DG and Lauenroth WK (1993) Quantitative effects of grazing on vegetation and soils over a global range of environments. *Ecol. Monogr.* 63(4): 327–366.

Milchunas DG, Sala DE, and Lauenroth WK (1988) A generalised model of the effects of grazing by large herbivores on grassland community structure. *Am. Nat.* 132(1): 87–106.

Pain DJ and Pienkowski MW (eds.) (1996) *Farming and Birds in Europe*. San Diego: Academic Press.

Popay I and Field R (1996) Grazing animals as weed control agents. *Weed Technol.* 10: 217–231.

Russelle MP (1992) Nitrogen cycling in pasture and range. *J. Production Agric.* 5: 13–23.

Snaydon RW (ed.) (1993) *Managed Grasslands; Analytical studies. Ecosystems of the world*, vol. 17b. Amsterdam: Elsevier.

Spratt DM (1997) Endoparasite control strategies: Implications for biodiversity of native fauna. *Int. J. Parasitol.* 27(2): 173–180.

Stuth JW and Lyons BG (eds.) (1993) *Decision Support Systems for the Management of Grazing Lands*. Man and the Biosphere Series, vol. 11. Rome: UNESCO/The Parthenon Publishing Group.

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF
Simon A Levin
Princeton University, Princeton, NJ, USA

VOLUME 2



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xv

Contributors to Volume 2 xxi

Foreword to the First Edition xxv

Preface to the First Edition xxvii

Preface to the Second Edition xxix

How to use of the Encyclopedia xxxi

VOLUME 2

Census of Marine Life 1
Darlene Trew Crist and Ronald O'Dor

Central America, Ecosystems of 23
Rodolfo Dirzo and Maria Argenis Bonilla

Cladistics 33
Ian J Kitching, Peter L Forey, and David M Williams

Cladogenesis 46
Christopher J Humphries

Climate Change and Ecology, Synergism of 58
Stephen H Schneider and Terry L Root

Climate Change and Extinctions 73
Joanne L Isaac and Stephen E Williams

Climate Change and Wild Species 79
Terry L Root, Stephen H Schneider, Rachel Warren, Jeff R Price, and Patricia R Mastrandrea

Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species 100
Joshua J Lawler, Carrie A Schloss, and Ailene K Ettinger

Climate, Effects of 115
GF Midgley and FI Woodward

Coastal Beach Ecosystems 128
Anton McLachlan and Omar Defeo

Coevolution 137
Douglas J Futuyma and André Levy

Commons, Concept and Theory of 149
Colin W Clark

Commons, Institutional Diversity of 155
Elinor Ostrom

Comparing Extinction Rates: Past, Present, and Future 167
Vânia Proença and Henrique Miguel Pereira

Competition, Interspecific 177
Bryan Shorrocks

Complementarity 192
Paul Williams and Lisa Manne

Complexity Versus Diversity 201
Shahid Naeem

Computer Systems and Models, Use of 213
Louis J Gross

Conservation and People 221
Michelle Marvier

Conservation and the World's Poorest of the Poor 230
Craig Leisher and M Sanjayan

Conservation Biology, Discipline of 238
Andrew P Dobson, Katarzyna Nowak, and Jon Paul Rodríguez

Conservation Efforts, Contemporary 249
Kristiina A Vogt, Jason J Scullion, Lloyd L Nackley, and Maura Shelton

Conservation Genetics 263
Katie Elizabeth Frith and A Rus Hoelzel

Conservation Movement, Historical 278
Curt Meine

Conserving Biodiversity Outside Protected Areas 289
Belinda Reyers

Convention on Biological Diversity 306
H David Cooper and Kieran Noonan-Mooney

Cooperation, Evolution of 320
Egbert Giles Leigh Jr.

Corals and Coral Reefs 330
Nancy Knowlton and Jeremy Jackson

Countryside Biogeography 347
Chase D Mendenhall, Carrie V Kappel, and Paul R Ehrlich

C_4 Plants 361
Rowan F Sage and Tammy L Sage

Crop Mixtures and the Mechanisms of Overyielding 382
Long Li, Lizhen Zhang, and Fusuo Zhang

Crustaceans 396
Marjorie L Reaka-Kudla

D

Darwin, Charles (Darwinism) 419
Michael T Ghiselin

Defenses, Ecology of 426
Phyllis D Coley and John A Barone

Defining, Measuring, and Partitioning Species Diversity 434
Hanna Tuomisto

Deforestation and Land Clearing 447
Jaboury Ghazoul

Desert Ecosystems 457
Exequiel Ezcurra and Eric Mellink

Desertification 479
James F Reynolds

Development by Design: Using a Revisionist History to Guide a Sustainable Future 495
Joseph M Kiesecker, Kei Sochi, Michael Heiner, Bruce McKenney, Jeffrey Evans, and Holly Copeland

Differentiation 508
Nicholas H Barton

Dinosaurs, Extinction Theories for 516
J David Archibald and Norman MacLeod

Diseases, Conservation and 523
Barbara A Han and Sonia Altizer

Dispersal Biogeography 539
Ran Nathan

Disturbance, Mechanisms of 562
Frank W Davis and Max A Moritz

Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems 568
Robert Keane

Diversity and Conservation of Neotropical Mammals 582
Ricardo A Ojeda

Diversity, Community/Regional Level 595
Howard V Cornell

Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea 608
Matthew Church and Daniela Böttjer

Diversity, Molecular Level 626
Carlos A Machado and Marcos A Antezana

Diversity, Organism Level 638
Daniel R Brooks and Deborah A McLennan

Diversity, Taxonomic versus Functional 648
John C Moore

Domestication of Crop Plants 657
Daniel Zohary

Dormancy and Diapause 665
Nelson G Hairston Jr. and Jennifer A Fox

Dynamic Global Vegetation Models 670
Iain Colin Prentice and Sharon A Cowling

E

Ecogenomic Sensors 690
Christopher A Scholin

Ecological Footprint, Concept of 701
William E Rees

Ecological Genetics 714
Beate Nürnberger

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 2

Sonia Altizer
University of Georgia, Athens, GA, USA

Marcos A Antezana
Universidade de Lisboa, Lisboa, Portugal

John A Barone
Mississippi State University, Mississippi State, MS, USA

Nicholas H Barton
IST Austria, Klosterneuburg, Austria

Maria Argenis Bonilla
Universidad Nacional de Colombia, Bogotá, Colombia

Daniela Böttjer
University of Hawaii, Honolulu, HI, USA

Daniel R Brooks
University of Toronto, Toronto, Canada

Matthew Church
University of Hawaii at Manoa, Honolulu, HI, USA

Colin W Clark
University of British Columbia, Vancouver, BC, Canada

Phyllis D Coley
University of Utah, Salt Lake City, UT, USA

Holly Copeland
The Nature Conservancy, Wyoming, WY, USA

Howard V Cornell
University of Delaware, Newark, DE, USA

Sharon A Cowling
University of Toronto, Ontario, Canada

Darlene Trew Crist
University of Rhode Island, North Kingstown, RI, USA

J David Archibald
San Diego State University, San Diego, CA, USA

H David Cooper
Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada

Frank W Davis
University of California, Santa Barbara, CA, USA

Omar Defeo
Faculty of Sciences, UNDECIMAR, Montevideo, Uruguay

Rodolfo Dirzo
Stanford University, Stanford, CA, USA

Andrew P Dobson
Princeton University, Princeton, NJ, USA

Paul R Ehrlich
Stanford University, Stanford, CA, USA

Ailene K Ettinger
University of Washington, Seattle, WA, USA

Jeffrey Evans
The Nature Conservancy, Wyoming, WY, USA

Exequiel Ezcurra
University of California, Riverside, CA, USA

Peter L Forey
The Natural History Museum, London, UK

Jennifer A Fox
Georgetown University, Washington, DC, USA

Katie Elizabeth Frith
Durham University, Durham, UK

Douglas J Futuyma
State University of New York at Stony Brook, New York, NY, USA

Jaboury Ghazoul
Department of Environmental Sciences, ETH Zurich, Switzerland

Michael T Ghiselin
California Academy of Sciences, San Francisco, CA, USA

Louis J Gross
University of Tennessee, Knoxville, TN, USA

Nelson G Hairston Jr.
Cornell University, Ithaca, NY, USA

Barbara A Han
University of Georgia, Athens, GA, USA

Michael Heiner
The Nature Conservancy, Colorado, CO, USA

Christopher J Humphries
The Natural History Museum, London, UK

Joanne L Isaac
James Cook University, Townsville, QLD, Australia

Jeremy Jackson
Smithsonian Institution, Washington, DC, USA

Carrie V Kappel
University of California Santa Barbara, Santa Barbara, CA, USA

Robert Keane
Missoula Fire Sciences Laboratory, Missoula, MT, USA

Joseph M Kiesecker
The Nature Conservancy, Colorado, CO, USA

Ian J Kitching
The Natural History Museum, London, UK

Nancy Knowlton
Smithsonian Institution, Washington, DC, USA

Joshua J Lawler
University of Washington, Seattle, WA, USA

Egbert Giles Leigh Jr.
Smithsonian Tropical Research Institute, Panama, Republic of Panama

Craig Leisher
The Nature Conservancy, Millburn, NJ, USA

André Levy
State University of New York at Stony Brook, New York, NY, USA

Long Li
China Agricultural University, Beijing, China, and Shihezi University, Shihezi City, China

Carlos A Machado
University of Maryland, College Park, MD, USA

Norman MacLeod
Natural History Museum, London, UK

Lisa Manne
University of Toronto at Scarborough, Toronto, ON, Canada

Michelle Marvier
Santa Clara University, Santa Clara, CA, USA

Patricia R Mastrandrea
Stanford University, Stanford, CA, USA

Bruce McKenney
The Nature Conservancy, Virginia, VA, USA

Anton McLachlan
The University of Sydney, Sydney, NSW, Australia

Deborah A McLennan
University of Toronto, Toronto, Canada

Curt Meine
Aldo Leopold Foundation, Baraboo, WI, USA

Eric Mellink
CICESE, Carretera Ensenada-Tijuana, Ensenada, BC, México

Chase D Mendenhall
Stanford University, Stanford, CA, USA

GF Midgley
South African National Biodiversity Institute, Cape Town, South Africa

John C Moore
Colorado State University, Fort Collins, CO, USA

Max A Moritz
University of California, Santa Barbara, CA, USA

Lloyd L Nackley
University of Washington, Seattle, WA, USA

Shahid Naeem
Columbia University in the City of New York, New York, NY, USA

Ran Nathan
The Hebrew University of Jerusalem, Jerusalem, Israel

Kieran Noonan-Mooney
Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada

Katarzyna Nowak
Princeton University, Princeton, NJ, USA

Beate Nürnberger
The University of Edinburgh, Edinburgh, UK

Ronald O'Dor
Dalhousie University, Halifax, Canada, and Consortium for Ocean Leadership, Washington, DC, USA

Ricardo A Ojeda
Institute for Aridland Research (IADIZA-CONICET), Mendoza, Argentina

Elinor Ostrom
Indiana University, Bloomington, IN, USA

Henrique Miguel Pereira
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Iain Colin Prentice
Macquarie University, North Ryde, Australia

Jeff R Price
World Wildlife Fund, Washington, USA

Vânia Proença
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Marjorie L Reaka-Kudla
The University of Maryland, College Park, MD, USA

William E Rees
University of British Columbia, Vancouver, BC, Canada

Belinda Reyers
Council for Scientific and Industrial Research, Stellenbosch, South Africa

James F Reynolds
Duke University, Durham, NC, USA

Jon Paul Rodríguez
Instituto Venezolano de Investigaciones Científicas, Maracaibo, Venezuela

Terry L Root
Stanford University, Stanford, CA, USA

A Rus Hoelzel
Durham University, Durham, UK

Rowan F Sage
The University of Toronto, Toronto, ON, Canada

Tammy L Sage
The University of Toronto, Toronto, ON, Canada

M Sanjayan
The Nature Conservancy, Missoula, MT, USA

Carrie A Schloss
University of Washington, Seattle, WA, USA

Stephen H Schneider†
Stanford University, Stanford, CA, USA

Christopher A Scholin
Monterey Bay Aquarium Research Institute, CA, USA

Jason J Scullion
University of Washington, Seattle, WA, USA

Maura Shelton
University of Washington, Seattle, WA, USA

Bryan Shorrocks
University of Leeds, Leeds, UK

Kei Sochi
The Nature Conservancy, Colorado, CO, USA

Hanna Tuomisto
University of Turku, Turku, Finland

Kristiina A Vogt
University of Washington, Seattle, WA, USA

Rachel Warren
University of East Anglia, Norwich, UK

David M Williams
The Natural History Museum, London, UK

Paul Williams
Natural History Museum, London, UK

Stephen E Williams
James Cook University, Townsville, QLD, Australia

FI Woodward
University of Sheffield, Sheffield, UK

Fusuo Zhang
China Agricultural University, Beijing, China

Lizhen Zhang
China Agricultural University, Beijing, China

Daniel Zohary
The Hebrew University, Jerusalem, Israel

†Deceased.

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

C

Census of Marine Life

Darlene Trew Crist, University of Rhode Island, North Kingstown, RI, USA
Ronald O'Dor, Dalhousie University, Halifax, Canada, and Consortium for Ocean Leadership, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Ronald O'Dor, Darlene T. Crist and Andrea Ottensmeyer, pp 1–17, © 2001, Elsevier Inc.

Glossary

Abyssal plain Large areas of flat areas referred to as abyssal plain from ~4000 to 6000 m in water depth.
Archaea Microorganisms that are similar to bacteria in size and simplicity of structure but radically different in molecular organization. They are now believed to constitute an ancient intermediate group between the bacteria and eukaryotes.
Biodiversity The variety of life in the world or in a particular habitat or ecosystem.
Biome A large naturally occurring community of flora and fauna occupying a major habitat.
Cryptic species Morphologically indistinguishable but taxonomically distinct species.
Cyanobacteria A division of microorganisms that are related to the bacteria but are capable of photosynthesis. They are prokaryotic and represent the earliest known form of life on the earth.
Eukaryote An organism consisting of a cell or cells in which the genetic material is DNA in the form of chromosomes contained within a distinct nucleus.
Prokaryote A microscopic single-celled organism, including the bacteria and cyanobacteria, that has neither a distinct nucleus with a membrane nor other specialized organelles.
Protist A kingdom or large grouping that comprises mostly single-celled organisms such as the protozoa, simple algae and fungi, slime molds, and (formerly) the bacteria.
Taxon A taxonomic group of any rank, such as a species, family, or class.

Goals, Scope, and Strategy

In 2000, the Census of Marine Life was launched to assess and explain the diversity, distribution, and abundance of marine organisms throughout the world's oceans. It was organized around three questions: (1) What lived in the oceans? (2) What lives in the oceans? (3) What will live in the oceans? The answers required an ambitious international collaborative effort that ultimately took 10 years to complete, involved some 2700 scientists from more than 80 nations, 670 institutions, 540 expeditions, and a staggering 9000 days at sea. The first Census of Marine Life, released on 4 October 2010, reported life in the ocean was richer, more connected, and more impacted than previously thought. Its tangible achievements were many and included the creation of a baseline against which future change can be measured, establishment of the world's largest online geospatially referenced marine life database of nearly 30 million records, advancement of technology used to explore the ocean, and a foundation for how such large scientific investigations can be structured to succeed. This cooperative $650 million scientific endeavor also systematically defined for the first time both the known and the vast unknown, unexplored ocean (Ausubel *et al.*, 2010).

To accomplish its far-reaching agenda, the Census had to be global in nature, encompass all ocean realms from top to bottom, from the tropics to the poles, and include all ocean animals from microbes to whales. Three years of discussion and dozens of planning meetings preceded the first expedition and official launch of the Census in 2003.

The Census was divided into four component parts: a project that investigated historical fish populations, one that modeled future populations using a centralized database in which data were made accessible and permanently stored for future use based on input from 14 field projects that explored previously unexplored or little explored ocean areas. The oceans were divided into six realms (with respective subzones as necessary), with Census field projects developing efficient approaches for the exploration of each.

An International Scientific Steering Committee representing seven countries guided the science over the decade.

National and Regional Implementation Committees provided leadership in 13 regions around the world and extended access to expertise and resources and enhanced the sharing of knowledge and data (Figures 1 and 2).

The Census began with the question "What lived in the oceans?" Researchers undertook the challenge of constructing the history of marine animal populations (HMAP) since human predation became important, roughly the past 500

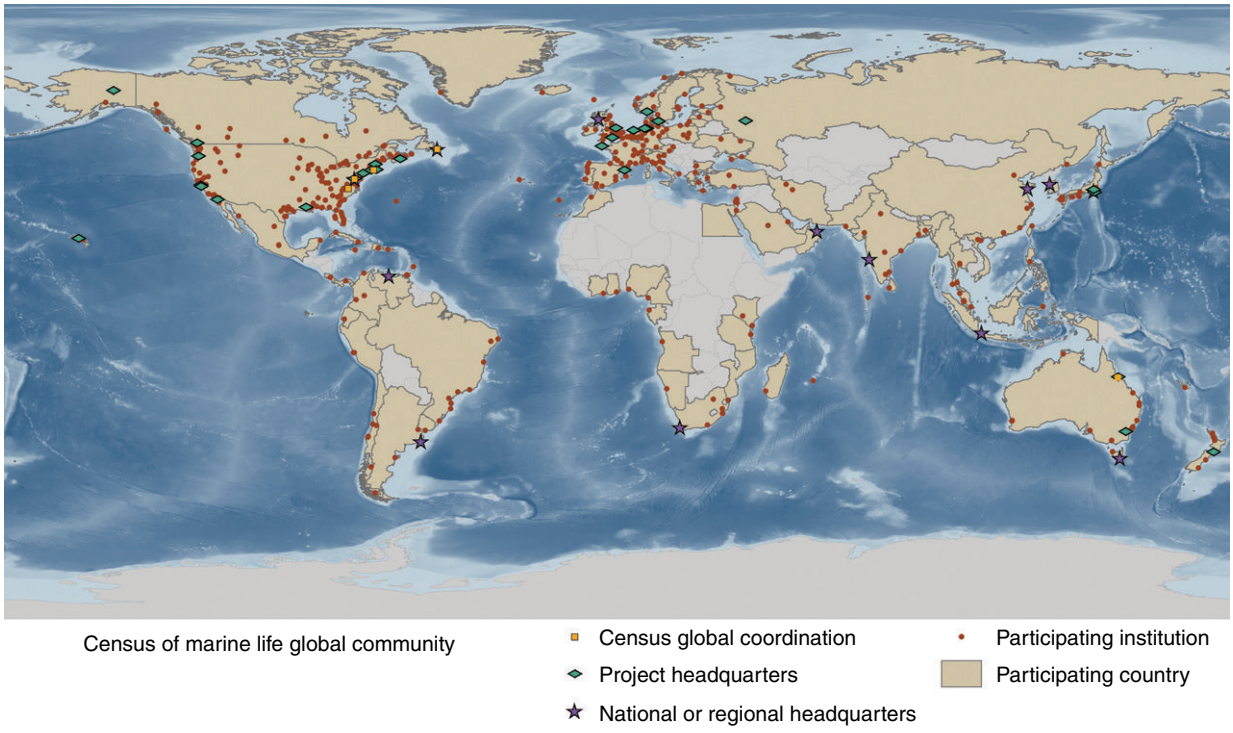


Figure 1 Shading shows the global span of 80-plus nations that participated in the Census of Marine Life. Two-thousand seven-hundred scientists participating came from more than 600 institutions and worked in 14 field projects spanning ocean realms from near shore to abyss and latitudes from poles to tropics. The global span of National and Regional Implementation Committees, headquarters of the field projects, and participating institutions such as museums ensured access to expertise and seas.

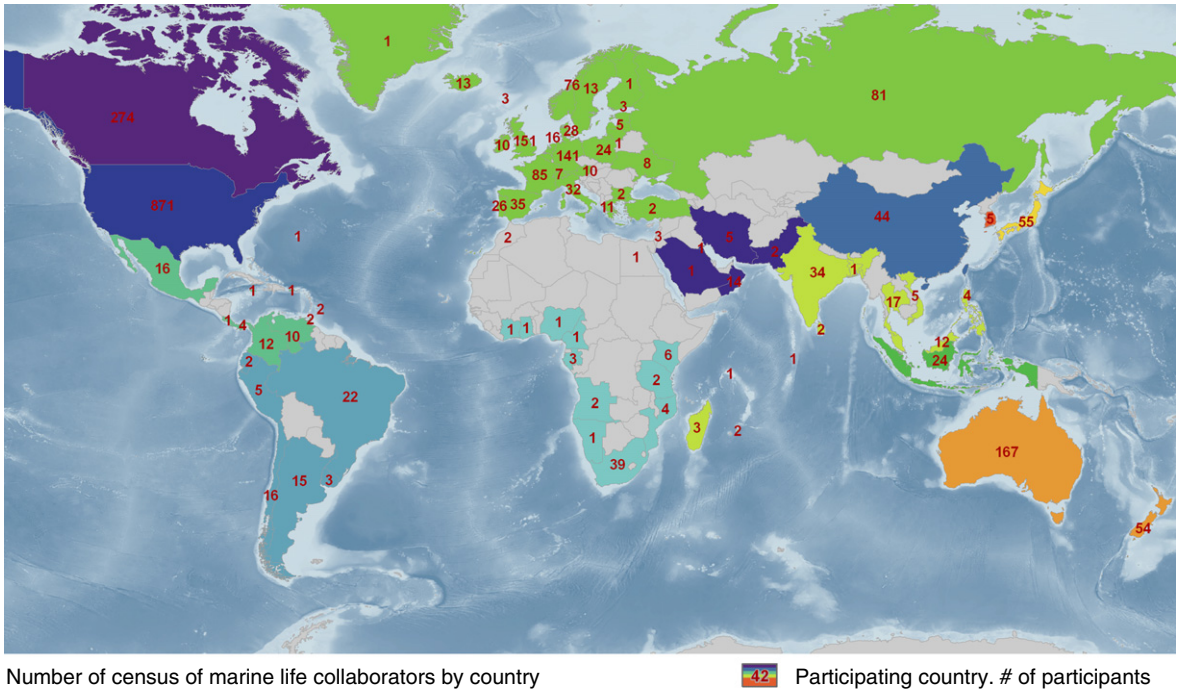


Figure 2 More than 80 nations participated in the global Census of Marine Life. Shown here is the number of participants by country.

years. This program component called the HMAP consisted of teams of fisheries scientists, historians, and economists, who initiated case studies in southern Africa, Australia, and about a dozen other regions. These case studies created the first reliable picture of life in the oceans before fishing, served to help distinguish the contributions of natural fluctuations in the environment from the effects of human activities, and reported that human impacts across-the board happened much earlier and were more pronounced than imagined previously.

Such a systematic collaboration between humanities and sciences bred the establishment of a new scientific discipline – marine environmental history – that will carry on the investigation of historic marine life populations well past the duration of the Census (**Figure 3**).



(a)



(b)

Figure 3 While recreational fishing removes much less sea life than commercial fishing, recreational fishers do sample marine life, especially large animals. In 1958, after an excursion on a Key West, Florida (USA) charter boat, a family of recreational fishers displayed their trophies, especially large groupers, subfamily Epinephelinae. In 2007, the size of the fish displayed by the same charter enterprise had plummeted, and the mix of species had changed. These images graphically illustrate many of the findings of the Census project that studied historic fish populations. Reproduced from Monroe County Library (top) and McClenachan L (2009) Documenting loss of large trophy fish from the Florida Keys with historical photographs. *Conservation Biology* 23(3): 636–643.

The largest component of the Census involved answering the question "What lives in the oceans?" Fourteen field projects investigated the major habitats and groups of species in the global ocean to assess the present marine life. Of these, 11 explored habitats, such as seamounts, vents, and coral reefs, or regions such as the Arctic and Southern Oceans, the Mid-Atlantic Ridge, or the Gulf of Maine. The remaining three projects surveyed animals from the top predators to zooplankton and microbes. Details of these interesting and varied field projects are provided in Section "Legacies" (**Figure 4**).

To answer the question "What will live in the oceans?" required numerical modeling and simulation. This component program was organized under the rubric of the future of marine animal populations, which focused on integrating data from many different sources and created new statistical and analytical tools to make predictions for marine populations and ecosystems in the future.

To complete the initiative and provide permanent access to data, the Census created a database, the Ocean Biogeographic Information System (OBIS), which in 2010 contained nearly 30 million records. OBIS will serve researchers, students, and policymakers with its vast library of spatially referenced marine life data for years to come as part of the Intergovernmental Oceanographic Commission (UNESCO) global data system. The data help define the unknown ocean, delineating areas that remain unexplored and where more research might be targeted (**Figures 5** and **6**).

The Census of Marine Life 2010 was unique in its global perspective, its breadth and resolution of taxonomic detail, and the diversity of ecosystems explored. It significantly enhanced knowledge about what is known about life in the ocean by exploring previously unexplored or little explored areas, following journeys of marine animals as they plied the open waters, and providing new windows into what was an opaque ocean. Along the way, it created analytical tools for predicting the future of ocean ecosystems, and contributed to a more accurate understanding of ocean life in the past. Each of these advances was aided by the core foundation of Census work: A commitment to collaboration and sharing of knowledge and data among scientists, to pushing the envelope either through adaptation of time-tested techniques or development of new technology, and to exploring places where "no one had gone before." This approach advanced discovery, increased understanding, provided a baseline of ocean life against which future change can be measured, and will serve as a model for how future scientific endeavors might be structured to conduct research on a global level.

What was Learned About Diversity

To accomplish a biological census of the global ocean and compile an inventory of diversity, the Census focused on species. The Census compiled tens of millions of species-level observations obtained before and outside the Census and added millions of its own. The result of this species inventory is that the Census increased the estimate of known marine species from about 230,000 to nearly 250,000, found more than 6000 potentially new species, and completed formal descriptions of more than 1200 of them.

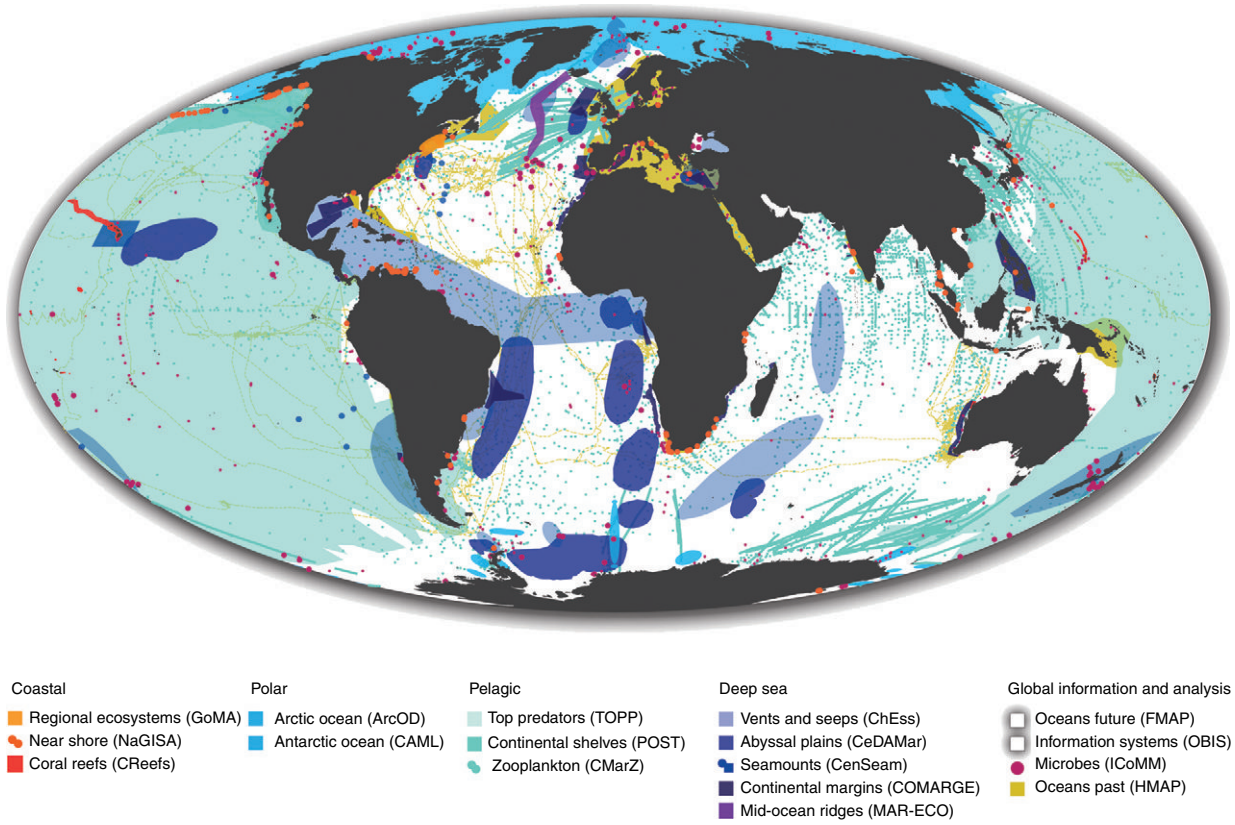


Figure 4 Census field projects sampled all major zones and realms of the oceans. Reproduced from Census of Marine Life Mapping and Visualization Team.

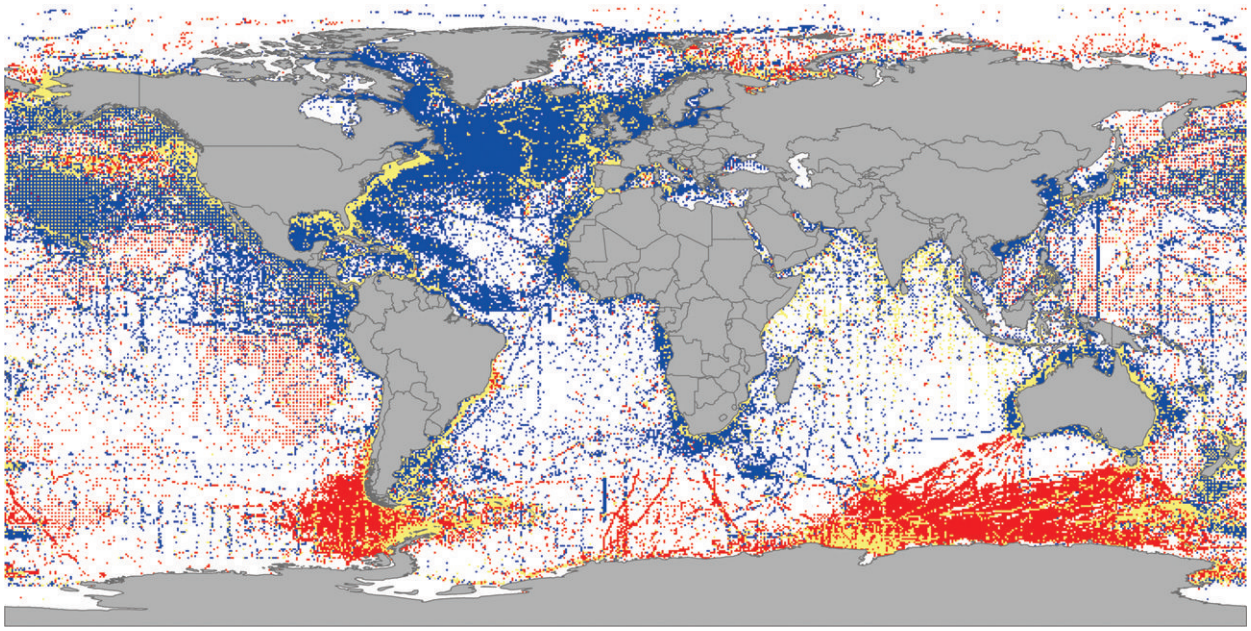


Figure 5 A global map of the nearly 30 million OBIS records of 120,000 species from more than 800 datasets shows the known and unknown ocean in half-degree squares by latitude and longitude. In blue areas, the Census aggregated data from before the Census began and from partner programs and institutions, often assembled by OBIS regional and thematic nodes. Yellow indicates regions with data both from Census partners and from the Census's own expeditions. Red indicates regions with data from Census expeditions where there were no prior data. The records attributable to and supplemented by the Census are numerous in the Southern Ocean, North Pacific, and along the Mid-Atlantic Ridge. Wide regions, such as the Eastern Pacific, remain largely unexplored. Reproduced from Ocean Biogeographic Information System.

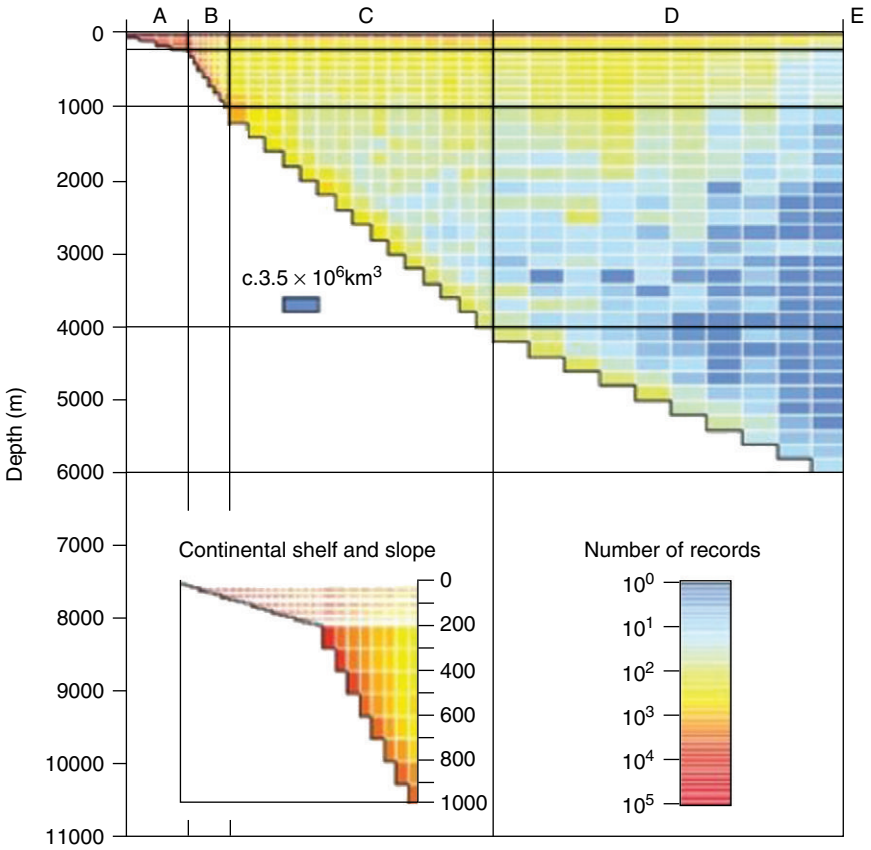


Figure 6 OBIS exposes the still-to-be-explored ocean by depth as well as latitude and longitude. On a cross-section of the global oceans, the spectrum from red to blue extends from many to few or no records. The records are concentrated near shores and in shallow waters, while the largest habitat on Earth, the vast middle waters, is largely unexplored. Reproduced from Ocean Biogeographic Information System.



Figure 7 Species thought to be extinct have been rediscovered. For example, Census scientists found a Jurassic shrimp, *Neoglyphea neocaledonica* thought to have become extinct 50 million years ago. Photo: Richer de Forges B and La J (2007).



Figure 8 Another species thought to be extinct but rediscovered: a living Caribbean fossil, *Pholadomya candida*, the only remaining species of a genus of deepwater clams that flourished worldwide for more than 100 million years and was thought during the 1800s to have vanished long ago. Photo: Juan Manuel Díaz.

The Census found a shrimp species thought to be extinct, a clam deemed a living fossil, worms that live up to 600 years on the sunless bottom, cataloged the longest living mollusk (age up to 500 years), and identified more than 38,000 different kinds of bacteria in a liter of sea water. Interestingly, the Census did not have to go to distant places to discover new species. Large and conspicuous species in relatively well-known shallow waters still await discovery. Sampling in relatively well-known shallow waters such as the Aleutian Islands (Alaska, USA) and the Great Barrier Reef Australia revealed many species presumed new to science, with some already formally identified as such (**Figures 7–9**).

The Census also found that rare species are common. During an 11-month study in 2007, for example, scientists sequenced the genes of more than 180,000 specimens from the western English Channel. One in every 25 readings yielded a new genus of bacteria (7000 genera in all), quantitatively demonstrating that less is known about the small than about the large. Many rare species in a sample contradict the notion of a few species predominating. Wherever Census researchers looked, they found many kinds of microbes in a sample represented by less than one in 10,000, including individuals that occur only once. They hypothesized that many microbes, now rare, could become dominant if environmental changes favor them. And the rare make up most of the microbial diversity in the oceans (**Figure 10**).



Figure 9 Sampling near the shore of the Aleutian Islands (Alaska, USA) in 2008 uncovered a new species of large brown seaweed, the Golden V kelp, *Aureophycus aleuticus*, distinct enough to be assigned to a new genus. This kelp with characteristic V-shaped blades, or "leaves", can grow as long as 3 m. Even the relatively well-known nearshore environment still harbors unknown species. Photo: Max K. Hoberg.

Via its digital database, the Census made possible the first regional and global comparisons of marine species diversity. OBIS allowed mapping of the levels of biodiversity in different regions of the oceans and for different taxa, and then analysis of what environments favor diversity. Two major patterns emerged from a global study of 13 taxa from zooplankton to mammals: Diversity in the open oceans peaked in midlatitude or in subtropical "strips" in all oceans, while coastal species were most diverse in tropical areas such as Indonesia, southeast Asia, and the Philippines. Sea surface temperature was the only environmental predictor highly related to diversity across all 13 taxa. An unprecedented review of all known marine biodiversity in 25 regions confirmed the coastal pattern.

The Census also helped to create the first comprehensive list of the known marine species, cataloged by the World Register of Marine Species, which surpassed 188,000 names in August 2010. Members of the Census team also helped to compose Web pages for more than 80,000 marine species in the online catalog, the Encyclopedia of Life.

The Census advanced genetic analysis, barcoding, and identification of marine species, compiling an impressive dataset of 35,000 marine species. The dataset allowed researchers to graph the proximity and distance of relations among distinct species, painting a new picture of the genetic structure of marine diversity, aptly named for artist Paul Klee (**Figure 11**).

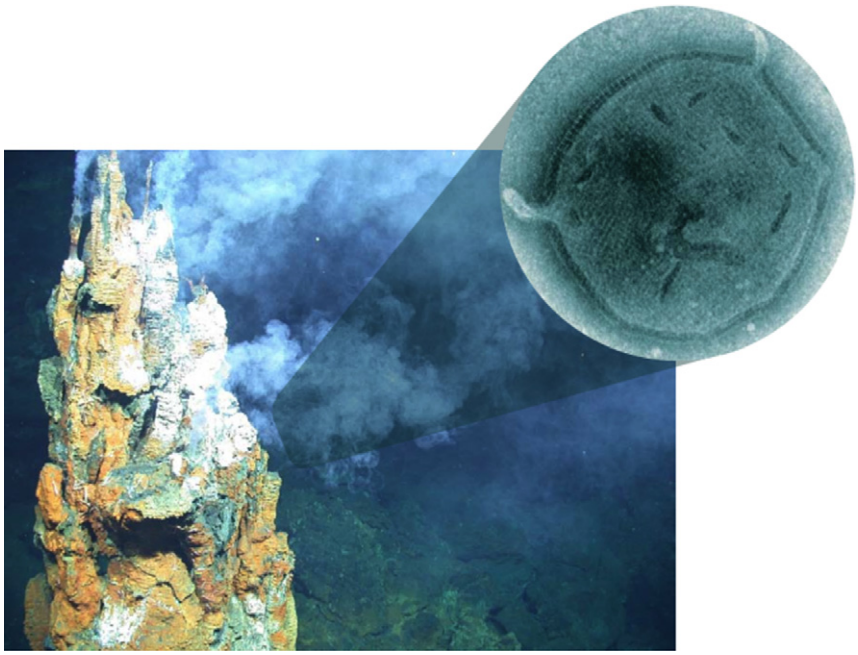


Figure 10 This newly described microbe, a thermoacidophilic archaeon found in a hydrothermal vent deposit from the Eastern Lau Spreading Center in the South Pacific near Fiji, exemplifies the rare biosphere. Wherever Census researchers looked, they found many species in a sample represented by less than one in 10,000 of all individuals, including those that occurred only once. Photo: Reysenbach AL *et al.* (2006).

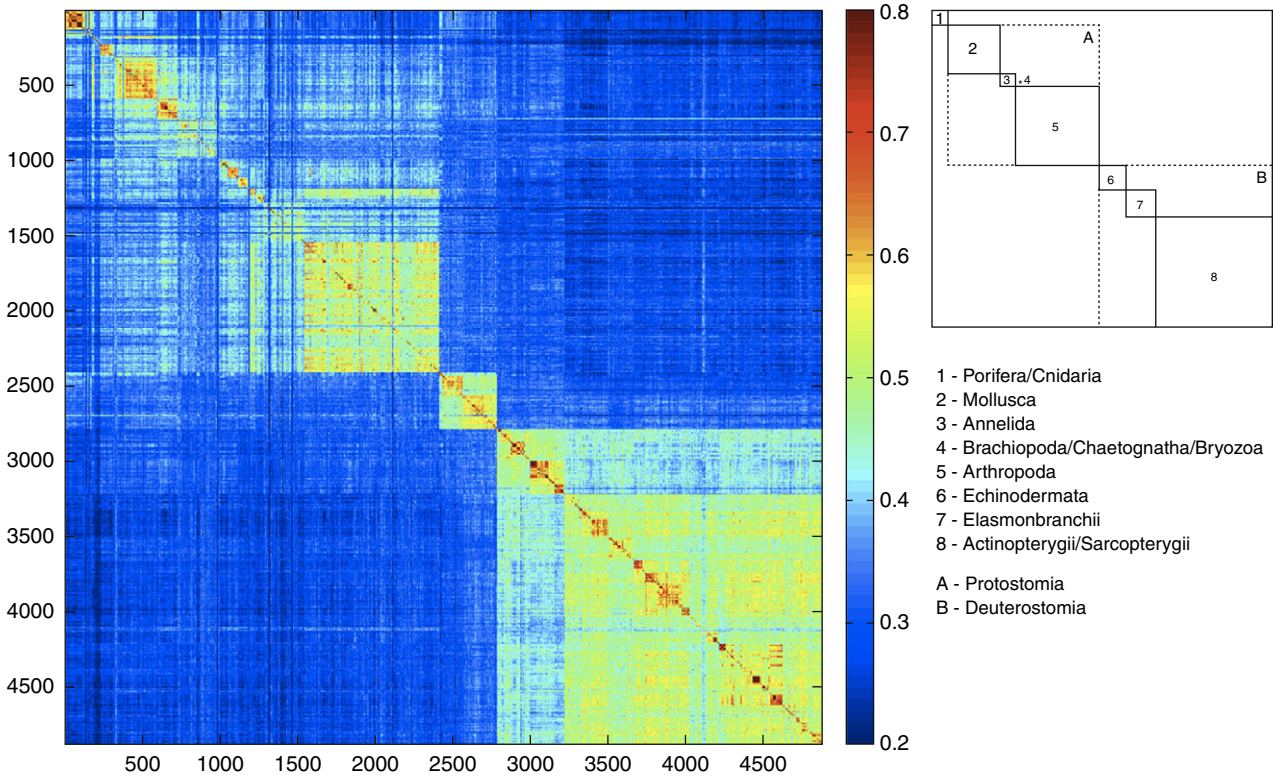


Figure 11 Klee diagrams, named for their analogy to the paintings of Paul Klee, display the genetic separation of species and the intensity of relations among species. Blocks of high correlation (red or yellow) on the diagonal reflect affinities within groups of species, corresponding to taxonomic divisions. Each species contributes a line, and a group of closely related species, such as fishes, becomes a group of lines that form a box. Klee diagrams illuminate broad patterns in the genetic structure of biodiversity. The color map depicts correlations for 5000 marine species from 10 phyla. Elasmobranchii, Actinopterygii, and Sarcopterygii are, respectively, cartilaginous, ray-finned, and lobe-finned taxa of fishes. The correlations coalesce into a coherent picture of categories of marine life. Photo: Dirk Steinke.

At the end of the decade, it was clear that much work remains to be done on diversity, including further refinement of the total number of marine species. Researchers agreed that at least one million species live in the world's ocean, possibly many more, and for microbes, there may be tens or even hundreds of millions of different kinds, awaiting discovery.

What was Learned About Distribution

The Census found life everywhere it looked – a virtual riot of species – and in unexpected places – everywhere from at the site of a hydrothermal vent at temperatures that would melt lead to within seafloor sediments in the deep Mediterranean where Census scientists were the first to discover anaerobic animals living without oxygen. The millions of OBIS records assembled during the Census quickly display where species have been reported, and allow for creation of species range maps. Known habitats and ranges in which species are known to exist were expanded during the decadal study and Census scientists confirmed that extreme is normal for life in the global ocean (**Figure 12**).

Using sound, satellites, and electronics, the Census tracked thousands of animals, mapping their migratory routes and charting hotspots and breeding places across the interconnected ocean, creating a map of "blue highways." In some cases, the tags measured the physical parameters of the water in which the animals traveled, providing for a richer database as well as increased understanding of parameters that influence animal movements. The Census identified migratory routes and breeding areas of many marine animals. Some of the distances traveled were quite remarkable. Atlantic bluefin tuna, *Thunnus thynnus*, for example, migrate east to west and west to east across the North Atlantic about 6000 km between Europe and North America, and Pacific bluefin tuna, *T. orientalis*, make multiple crossings of the Pacific. Equally impressive as marathon swimmers, humpback whales, *Megaptera novaeangliae*, complete annual north–south migrations of more than 8000 km. Circumnavigating the entire Pacific, a small seabird, the sooty shearwater, *Puffinus griseus*, made a 64,000-km round trip from New Zealand to Japan to Alaska to Chile and back to the Southern Ocean in 1 year. This trip, averaged about 40 km h^{-1}, and marked the longest-ever electronically recorded migration. Knowing how, when, where, and for what purposes, marine animals travel throughout the global ocean is invaluable and can be used to help preserve areas and routes for important activities such as migrating and breeding (**Figures 13** and **14**).

Curtains of acoustic receivers on the seabed were another means to follow marine animals. Using such an array allowed



Figure 12 Anaerobic animals living without oxygen have been imagined by science fiction writers but were not proved to exist on Earth until 2010. In the deep Mediterranean, a Census team found and added three new species to the few known species of the group called loriciferans. Unknown until the 1980s, loriciferans can live their entire lives hidden in sediment on the seafloor without oxygen. The *Nanaloricus cinzia* grows to about 0.3 mm long and 0.1 mm wide, comparable to the head of a pin. Reproduced from Danovaro R, Dell'Anno A. Pusceddu A, Gambi C, Heiner I, and Kristensen RM (2010) The first metazoa living in permanently anoxic conditions. *BMC Biology* 8: 30.

Census researchers to track 18 species of acoustically tagged fish more than 3000 km from California to Alaska, year-round. The searchable public database from this network in August 2010 contained more than 9 million detections in from more than 15,000 tagged animals. The exercise demonstrated the means for tracking even small marine animals (20 g), such as salmon, over long distances and through dissimilar waters. Among the many tracked, two juvenile Chinook salmon, *Oncorhynchus tshawytscha*, survived a 2500-km trip that took more than 3 months – from the Rocky Mountain headwaters of the Columbia River to the Pacific Ocean and north along the continental shelf of Canada's British Columbia to the coast of Alaska. The research provided insight about life cycles, movement, and behavior from the point of view of the fish and showed how the mortality of young salmon is apportioned among the habitats they encounter during their early life (**Figure 15**).

In addition to mapping blue highways, the Census also observed traffic patterns with hosts of animals commuting at regular hours, moving back and forth to the surface from hundreds of meters below. Using new types of echo sounders that reached the seafloor as well as sonars that looked upward in deep water, the Census observed a summer rush hour when animals rose to the surface to feed. The echo sounder moored 1000 m below the surface near the Mid-Atlantic Ridge recorded fishes and plankton rising about 400 m at about 21.00 h to the zone where sunlight and photosynthesis had produced food during the day. At about 06.00 h, they descended to the deep water below to avoid visual predators.

A new biogeography summarizing the distribution of marine life was also charted. Census researchers compared the distributions of 65,000 species of marine animals and plants and produced the first all-taxa map of global marine species diversity. These analyses distinguished 30 biogeographic regions worldwide (**Figure 16**).

Census research also identified animal movements resulting from changing ocean conditions, such as melting ice and temperature changes, with attendant implications for both policymakers and those who make their livings from the sea. Deep or shallow, changes in ocean temperatures, currents, and chemistry would redistribute much marine life. Census researchers predict a decline in diversity in a tropical ocean that becomes yet warmer, and an increase of diversity at latitudes of about 50–70° in both hemispheres.

Census data also revealed the unknown, unexplored ocean. For more than 20% of the ocean's volume, the Census database has no or very few records, providing a guidepost for where further exploration is needed (**Figure 17**).

What was Learned About Abundance

Measuring abundance of swimming and drifting life in the three dimensions of the interconnected global ocean created particular challenges for the Census. These obstacles were overcome through review of historical archives, calculation of primary production of food, and adoption of new technology to scan vast schools of fish that allowed the Census to increase knowledge about abundance of life in the global ocean.

Historic baselines of abundance were garnered from catch records, monastery records, fish bones, and other credible materials from which a snapshot of abundance was constructed. History showed people began catching marine life long ago, and their extractions were far broader in scope than imagined. Overfishing and habitat destruction were, and continue to be, the top threats to marine life associated with human activities. Declines in numbers and sizes were also documented, with about 90% of top predators in decline from historic levels. Recovery of some species where conservation efforts were implemented was also reported (**Figure 18**).

Many examples abound of declining size and quantity of catches. Off the North American coast after the mid-1800s, for example, commercial fishers caught swordfish, *Xiphias gladius*, in numbers varying by a factor of 10. The fish population and effort expended in fishing raised and lowered the numbers. However, the average weight of the fish, caught by either harpoons or lines, declined steadily from as much as 270 kg in 1860 to near or less than 100 kg during the decade of the Census (**Figure 19**).

Extractions of target species sometimes cause fast changes, as the trawl fishery off Southeast Australia between 1914 and 1950 demonstrated. First tiger flathead (*Neoplatycephalus richardsoni*) plummeted from early, high catch-rates, and then Chinaman leatherjacket (*Nelusetta ayraudi*) and latchet (*Pterygotrigla polyommata*), species abundant in the initial years

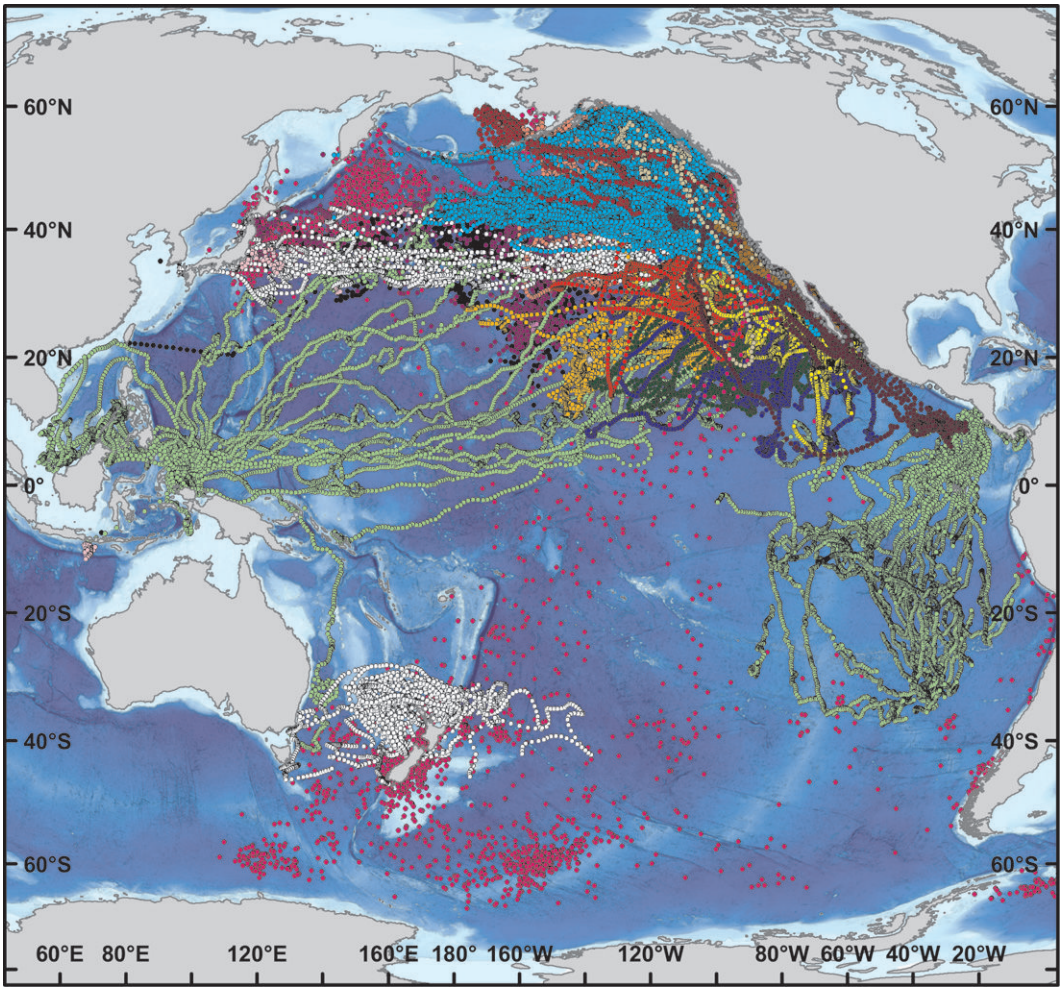


Figure 13 Reports from birds, fish, whales, and other animals carrying small tags reveal highways and neighborhoods of the vast Pacific. Census biologgers followed bluefin tuna, *Thunnus orientalis*, commuting between Japan and California and leatherback turtles, *Dermochelys coriacea*, between Borneo and Mexico. Traveling animals connect all the oceans. Reproduced from Tagging of Pacific Predators, and McIntyre AD (ed.) (2010) *Life in the World's Oceans – Diversity, Distribution, and Abundance*. London: Wiley-Blackwell.

of the fishery, virtually disappeared from catches in later years. Similar declines were measured for the once abundant Atlantic bluefin tuna, *Thunnus thynnus*, off the coast of northern Europe. Industrialized fishing and booming catches helped strip the Atlantic bluefin population in a blink of time – 1910–1950. The species virtually disappeared from the region in the early 1960s and is still rare today.

In contrast, recoveries tend to be slow. Increases are most marked for species whose exploitation ended at least 100 years ago and species that became protected in the early to mid-twentieth century.

Beyond catch statistics, Census scientists directly observed the abundance of marine life by deploying a new sensor system that tracked vast fish populations, or shoals, as well as small schools, over tens of thousands of square kilometers along the continental shelf. It was used to identify a shoal of tens of millions of Atlantic herring – comparable in size to Manhattan Island – off the coast of New Jersey in the United States. This new advanced sonar system had a range over an expanse of ocean a million times wider than could previously be studied. New sonar images, updated every minute, so scientists could continuously monitor the size and density of the fish in shoals and the shoal's shape as they changed from hour to hour. Beginning with a density of less than one fish per square meter, one shoal grew until it contained as many as eight fish per square meter.

The Census affirmed that by weight most marine life is microbial, up to 90%. (The equivalent weight is about 35 elephants for every living person on the planet.) Marine microbes are the tiniest cogs essential to planetary functioning and maintaining the planet's habitability. Responsible for more than 95% of respiration in the oceans, their importance is hard to overstate. By turning atmospheric carbon dioxide absorbed by the ocean back into carbon to be sunk to the depths (and doing likewise with nitrogen, sulfur, iron, manganese, and more), microbes regulate the composition of Earth's atmosphere, influence climate, recycle nutrients, create extractable resources, and decompose pollutants. Scientists are testing the sensitivity of marine microbes to such changes as the creeping acidification of the seas from rising atmospheric carbon dioxide levels, to understand which microbes thrive in different environments.

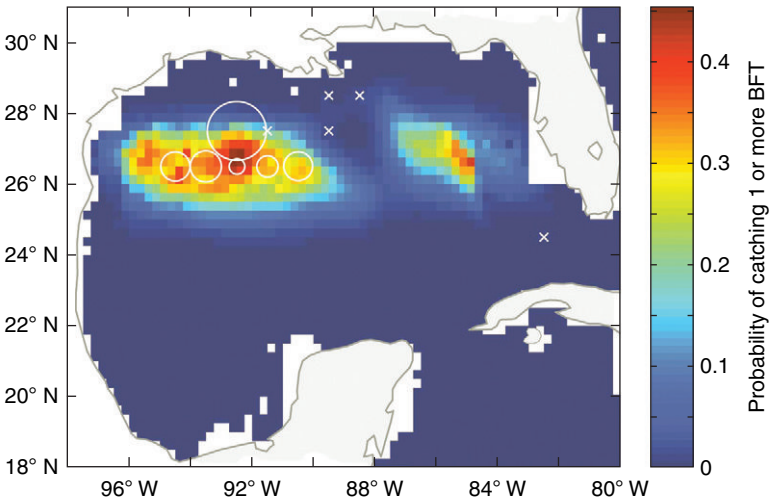


Figure 14 Conservation requires knowledge of mating areas and nurseries. Spawning Atlantic bluefin tuna (BFT), *Thunnus thynnus*, swim along the slopes in the western and eastern Gulf of Mexico. Red and blue colors on the map indicate high and low probabilities of encountering a bluefin when fishing with longlines in the Gulf of Mexico during 2002 and 2005. A long line caught no tuna at the X's on the map. The narrow band of 24–27 °C where most tuna swam prompts the speculation that a small change in temperature would change the timing and location of spawning.



Figure 15 The continental-scale acoustic curtains arrayed by the Census tracked migrations of 18 species, including juvenile Chinook salmon, *Oncorhynchus tshawytscha,* thousands of kilometers along the Pacific coast of North America and in rivers, too. The red spots on the map indicate the curtains through which animals swim, signal their passage, and give clues where and how they survive in the oceans. The orange line shows migration of the salmon the size of a banana spanning 2500 km in 3 months. The acoustic curtains also revealed an unexpected, long migration of the threatened green sturgeon, *Acipenser medirostris*, along the continental shelf from US to Canadian waters. A surprising group swam north for the winter.

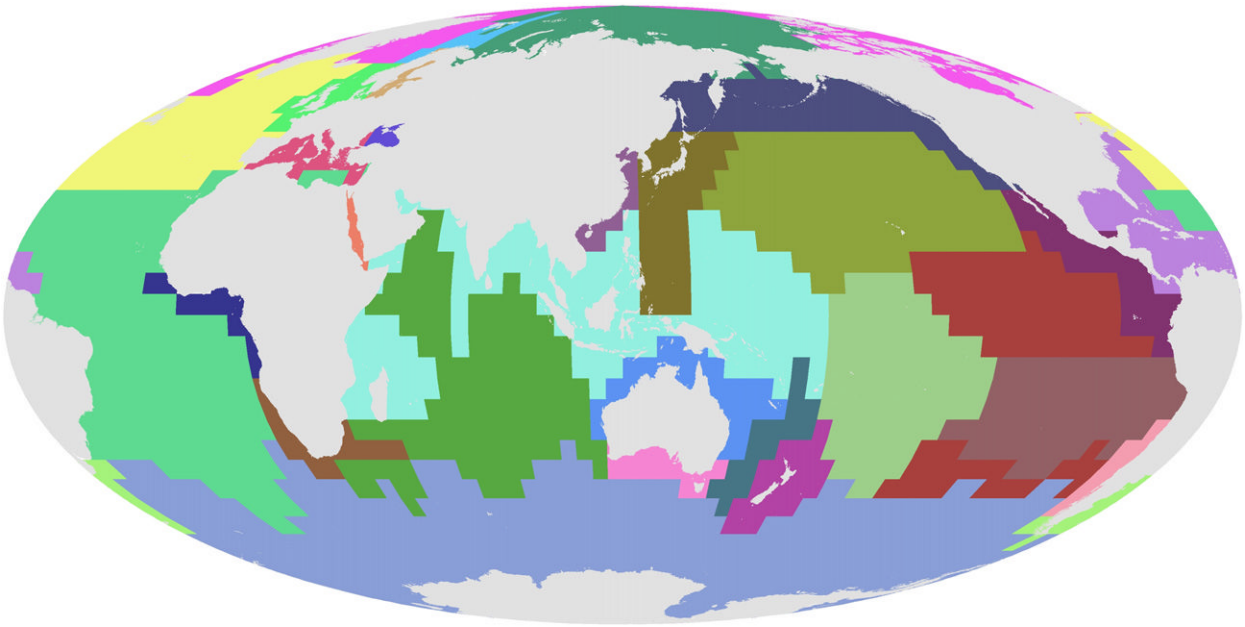


Figure 16 Spanning microscopic plant-like organisms to whales, the Census produced maps of global marine species using the distributions of more than 65,000 species from the Ocean Biogeographic Information System. Thirty marine biomes summarize the distributions of many species, each biome representing a region with a distinctive fauna and flora.

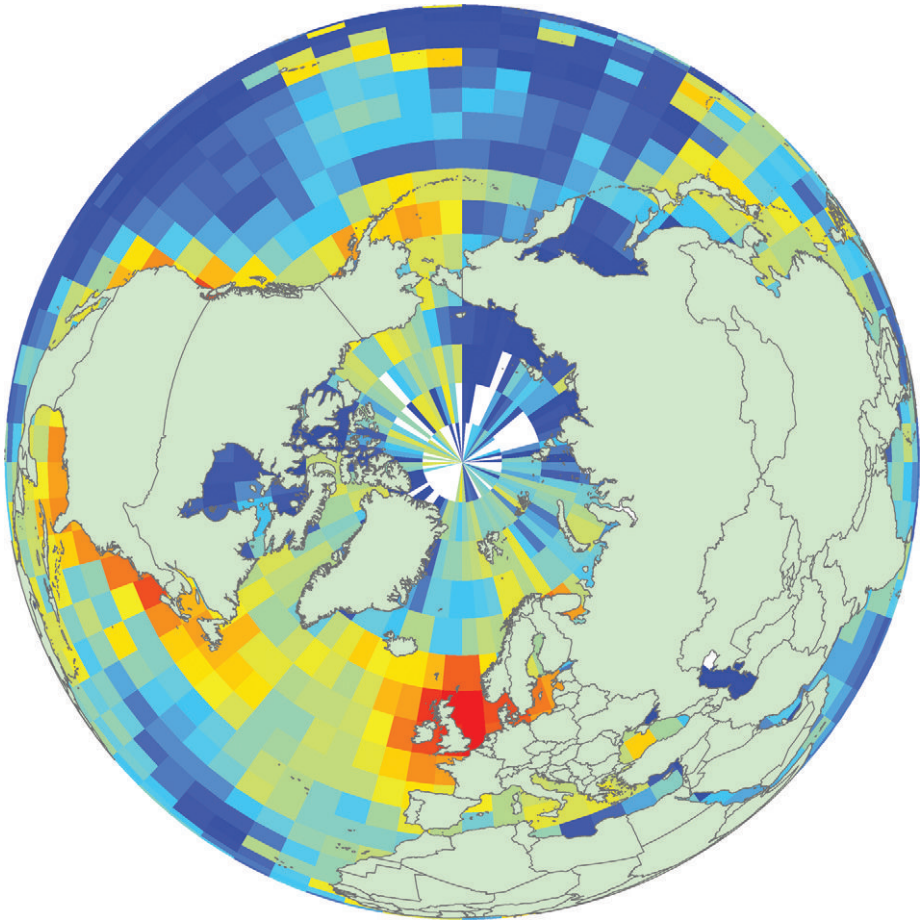


Figure 17 The global marine life database built by the Census exposes gaps in our knowledge of the known and unknown. The color spectrum from red to blue ranges from many to few records. Blank areas have no records of reliably identified marine life. Explorers have yet to explore much of the high Arctic Ocean for life.

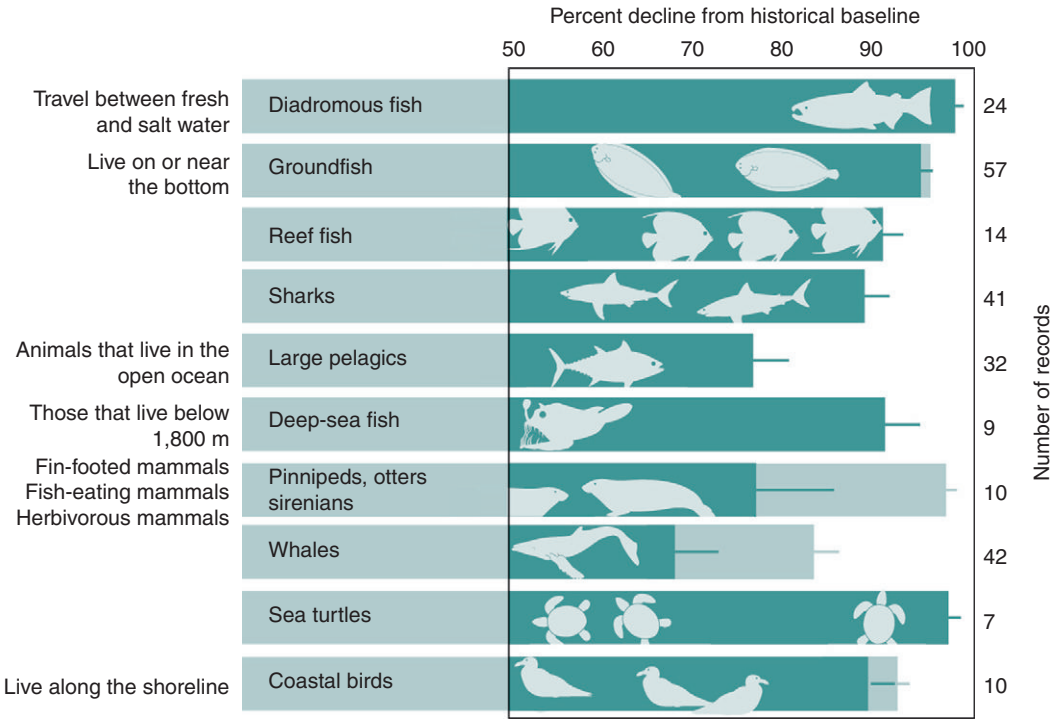


Figure 18 Dark-shaded bars indicate that the estimated declines in populations of large marine animals from their historical levels average about 90%. Some recovery (shown by light-shaded bars) has been achieved by four groups, including seals, whales, birds, and such bottom-dwelling fish species as flounder and sole. These estimates for exploited species in particular, rather than all marine animals, confirm the feared declines and offer a few hopeful instances of recovery.

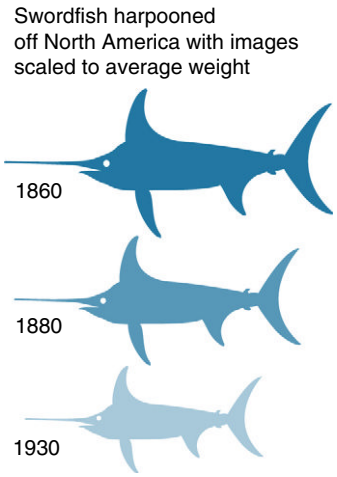


Figure 19 From 1860 to the present, the average weight of swordfish, *Xiphias gladius*, caught off the North American coast by either harpoons or lines declined from as much as 270 to near or less than 100 kg. A big drop occurred from 1860 to 1930.



Figure 20 Small life can form large structures. Off Chile, filamentous bacteria of the genus *Thioploca* wove mats covering an area the size of the nation of Greece. Census scientists estimated the mats to weigh about 14 million tons, about one-seventh of the weight of the annual world harvest of commercial fisheries.

Analyzing indirect observations from oceangoing vessels since 1899, Census researchers also discovered that the food-producing phytoplankton near the surface has declined, globally. How this decline in phytoplankton, the basis of the food web, relates to changes in abundance and biomass at higher levels of marine life, which are all dependent on it, is a major area of ongoing research.

Delivery of food in a "snow" from water above controls the mass of living things on the seafloor. The Census reported the quantity of life on the seafloor peaks toward polar regions, along continental margins where cool currents well up toward the surface, and where equatorial currents diverge. On the deep-sea margins, the Census unexpectedly discovered mats of bacteria extending hundreds of kilometers (**Figure 20**).

The question of whether the total weight of life in the ocean is changing remains unanswered after 10 years of investigation.

Technological Innovation

The Census focused new "binoculars" of technology into the ocean, providing a sharper focus on the world below the waves. Novel and reconfigured tools were used to explore many previously unexplored locations and helped to provide some better definition to what lives where in the huge global ocean. These new vision-expanding tools included the genetic barcoding techniques that automate and speed identification of species helping to compensate for a shortage of marine taxonomists, tagging and tracking methodologies that provide an insiders' view of this watery world, and acoustic innovations providing abundance assessments of marine populations heretofore not possible. While advancing technology, scientists also strove to standardize sampling protocols making it possible to make comparisons across the global ocean. These tools and strategies have provided a path forward for the next generation of ocean explorers. Some of the innovations not mentioned earlier are reviewed below (**Figure 21**).

DNA barcoding technology has dramatically increased the accuracy and speed of known species identification, discovery, and mapping. At the end of the decade, the Census had barcoded 35,000 marine species, which will prove invaluable for future research, as well as practical applications. The Census also championed integrated morphological and molecular genetic approaches to the analysis of zooplankton species diversity, and especially to the analysis of cryptic (i.e., morphologically indistinguishable but taxonomically distinct) species. The Census also pioneered shipboard DNA identification technology. Using environmental barcoding (i.e., sequencing DNA barcodes from an unsorted bulk sample), Census scientists detected 189 species of zooplankton in a sample from the western equatorial Pacific Ocean. This approach may dramatically accelerate the rate of sample and data analysis for ocean monitoring and management. The Census also spearheaded the use of "Pyrotag sequencing," a DNA technique used to ascertain the diversity of bacteria, archaea, and protists, while providing information about their relative abundance.

The Census demonstrated new approaches and technologies for zooplankton research using technologies that provide higher spatial resolution sampling and more accurate *in situ* depictions of animal behavior. One example, the Plankton Investigatory Collaborative Autonomous Survey System Operon (PICASSO) ROV system was deployed from ships of opportunity to study deep-sea gelatinous plankton below 1000 m. This system provided many striking images of organisms never before observed alive and allowed discovery of new zooplankton species. A second novel approach was using a 10-m Multiple Opening and Closing Net, with an Environmental Sensing System (MOCNESS) trawl equipped with fine-mesh nets to sample the deep sea (below 5000 m). This allowed high-volume sampling for rare species and capture of living specimens of zooplankton and fish species – some never before seen alive.

Other innovations involved data collection and standardization. With existing data plus its own, the Census built OBIS, the largest archive of data about marine species, which is now under the aegis of the International Oceanographic

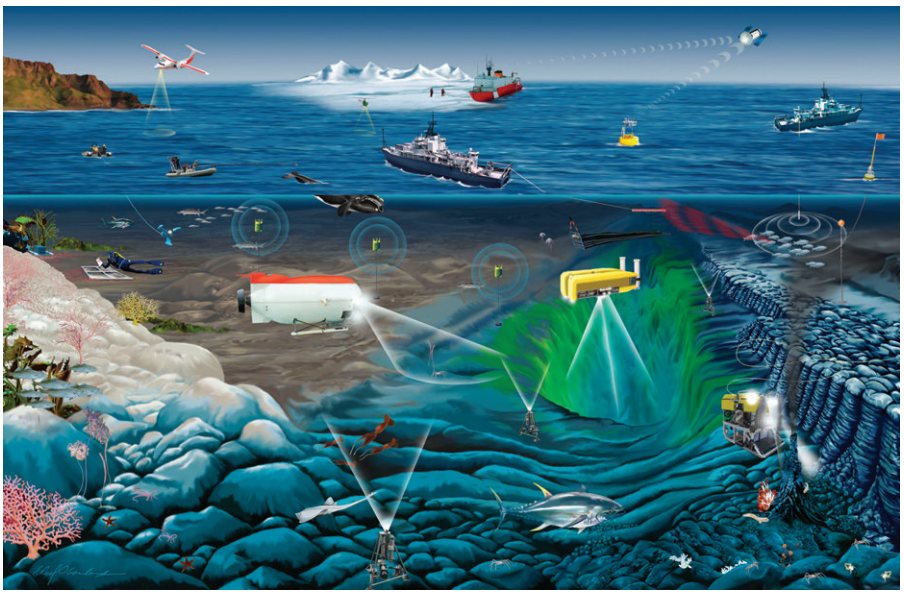


Figure 21 Census investigators explored on and beneath polar ice. Their aircraft remotely sensed animals through properties of scattered light. Marine animals themselves carried tags that stored records of their travels and dives and communicated with satellites. Fish carried tags that revealed their migration as they passed acoustic listening lines. Sounds that echoed back to ships portrayed schools of fish assembling, swimming, and commuting up and down. Standardized frames and structures dropped near shores and on reefs provided information for comparing diversity and abundance. Manned and unmanned undersea vehicles plus divers photographed sea floors and cliffs. Deep submersibles sniffed and videotaped smoking seafloor vents. And nets and dredges still caught specimens, shallow and deep, for closest study.

Commission (IOC) and International Oceanographic Data and Information Exchange (IODE) to ensure its continuation and viability for the long term. By making OBIS publicly available and connecting it to the Encyclopedia of Life and the World Register of Marine Species, and other elements of the "e-Biosphere," the Census created an infrastructure for future research.

To standardize the global assessment of life on coral reefs and render worldwide analysis of distributions and measurement of change, the Census innovated Autonomous Reef Monitoring Structures and standardized methods for appraising marine life near shores and in the deep sea. Collectively, Census technological innovations will make ocean exploration easier, more efficient, and possible during the decades ahead.

Legacies

At the end of its decade, the Census had increased knowledge, advanced technology, changed the way marine science is conducted, and helped advance public understanding and curiosity about what lives below the surface.

Regarding knowledge, the Census reported its findings in more than 2600 papers, many freely accessible online, a report by Trew Crist *et al.* (2009) (http://www.coml.org/highlights-2010), two books summarizing the scientific findings by Snelgrove (2010) (http://www.coml.org/discoveries-census-marine-life), and *Life in the World's Oceans: Diversity, Distribution, and Abundance* edited by Alasdair McIntyre (Blackwell Publishing, 2010) (http://www.coml.org/life-worlds-oceans), a third explaining the biology: *Citizens of the Sea, Discoveries of the Census of Marine Life* by Nancy Knowlton (National Geographic, 2010) (http://www.coml.org/citizens-sea), and a fourth explaining the Census research objectives and preliminary findings by Trew Crist *et al.* (2009) (http://www.coml.org/results-publications/worldoceancensus), and a map produced by National Geographic http://www.coml-maps.org/oceanlifemap, as well as other Web resources including videos (http://www.coml.org/). The Census built the largest repository of data about marine species by compiling observations and adding its own, and then made it a publicly accessible infrastructure for future research, which governments have committed to sustain. The Census drew baselines to help nations and the international Convention on Biological Diversity select areas and strategies for greater protection of marine life. Its baselines will help assess habitat changes such as warmer water or damage from oil spills.

One prime example of the practical application of such baselines is for the Gulf of Mexico, where 140 taxonomic experts from 80 institutions and 15 countries associated with the Census established a baseline for marine life, fortuitously before the April 2009 Deepwater Horizon oil spill. A complete inventory of marine life in the Gulf of Mexico before April 2009 is available in print and digital form (http://www.tamupress.com/product/Gulf-of-Mexico-Origin-Waters-and-Biota,5338.aspx), which can be used to assess the impacts of the oil spill on marine life in the Gulf as well as the recovery after cleanup (**Figure 22**).

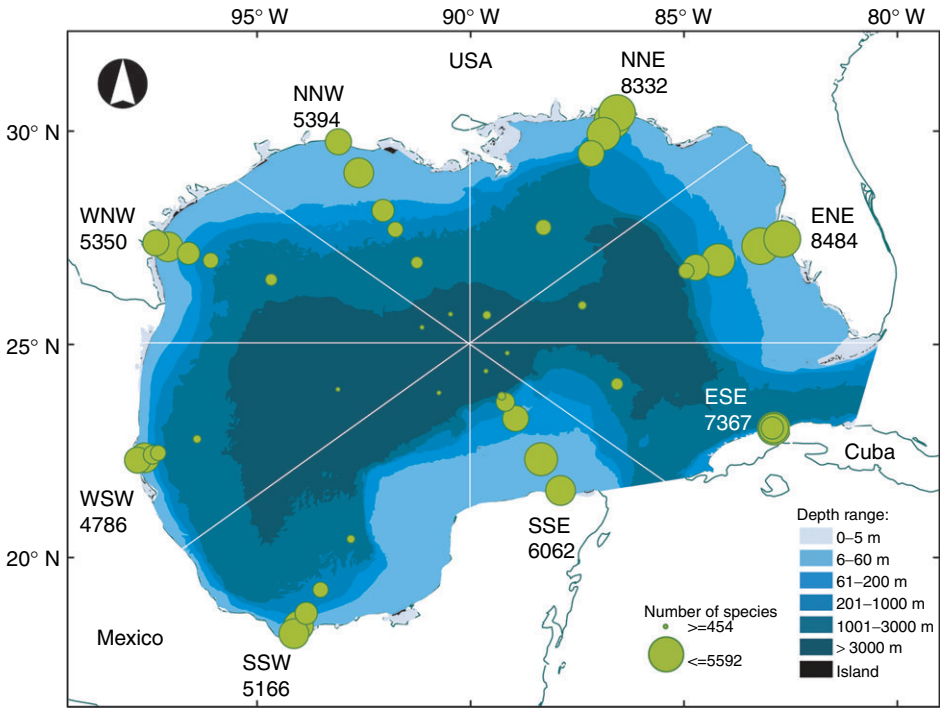


Figure 22 A map of marine life establishes a baseline for evaluating the outcome of an oil spill in the Gulf of Mexico. One-hundred and forty taxonomic experts from 80 institutions and 15 countries associated with the Harte Research Institute, a Census affiliate, established the baseline of known species in the Gulf of Mexico in 2009 and made it available online shortly thereafter. The data answer graphically how many species of each taxonomic group live in different sections of the Gulf of Mexico, at what depth, and on what substrate. The new database sets a baseline of 8332 species in the north-northeast sector of the Gulf, where the Deepwater Horizon rig tragedy occurred in 2010.

Summarizing technology, the Census proved new technology, such as DNA barcoding for the identification of marine life. It arrayed microphones from California past Canada to Alaska to pioneer a global ocean tracking network for animals, invented Autonomous Reef Monitoring Structures to standardize global assessment of reef life, and developed sound systems to measure abundances over tens of thousands of square kilometers. Together, these technologies show that the IOC Global Ocean Observing System (GOOS) can observe life as well as measure physical parameters.

Regarding scientific initiatives, the Census brought scientists with different interests from different nations together under one umbrella, to use standard protocols for sampling marine life from the deep sea to the near shore, to speed the adoption of good techniques, to build capacity economically, and by so doing jump started initiatives in marine research. It strengthened partnerships of scholars in the humanities and natural and social sciences to use archival research to build the picture of life in oceans past and assess changing diversity, distribution, and abundance.

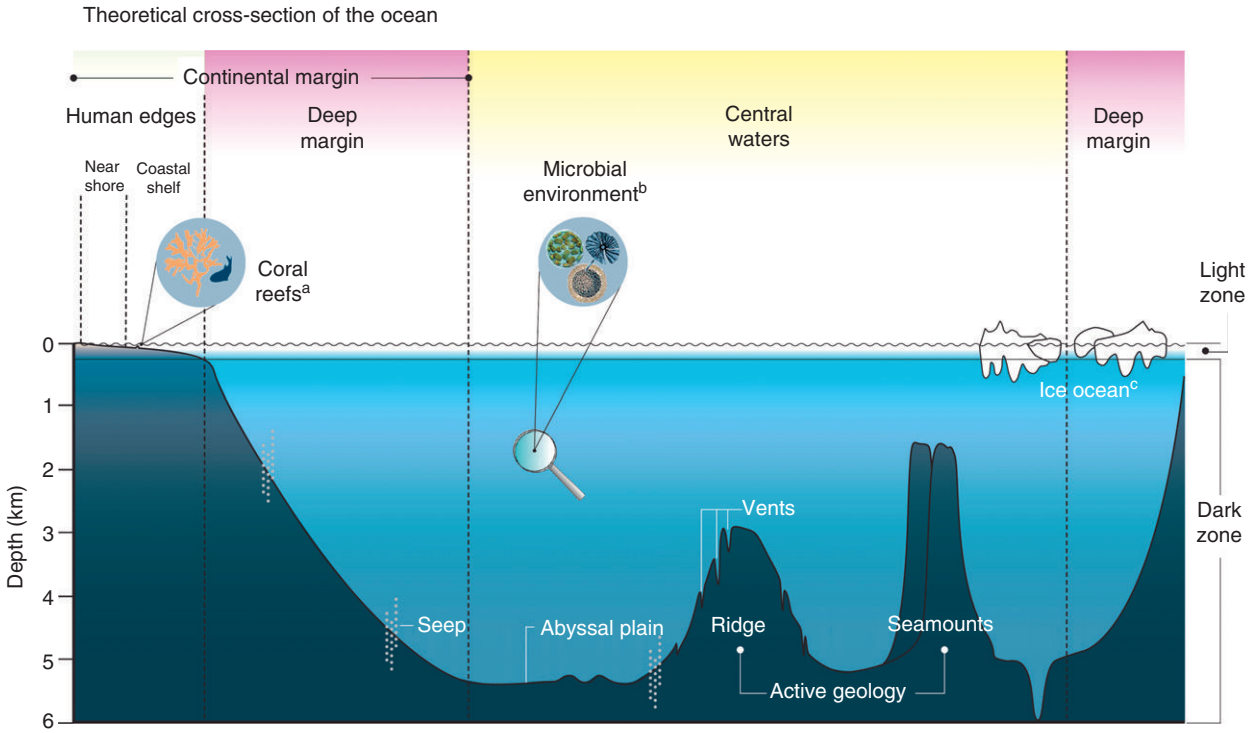
A truly unique aspect of the Census was that sharing its findings with the public was an integral part of the program. It accomplished this through regular news releases, video news releases, and even "creature features" that were posted on YouTube to inspire the public with the beauty, wonder, and magic of some of the newly discovered or rediscovered marine animals. Working with National Geographic, the Census created a new two-sided wall map of ocean life that visually depicted what was learned about diversity, distribution, and abundance and the past, present, and future of ocean life. A partnership was established with Galatee Films to ensure the scientific accuracy of the content of its feature-length documentary *Oceans*. And, the Census served as inspiration for artists around the world, from everything from sculptors to painters to cartoonists, including being featured in the syndicated cartoon *Sherman's Lagoon* and the long-running cartoon *Mark Trail*. The release of its final results in October 2010 reached an estimated 2 billion people worldwide, who as a result of the Census knew more about life in the world's oceans than they did a decade before.

How it was Accomplished

To successfully undertake a task as massive as a global census of marine life, the oceans were divided into six realms (with respective subzones as necessary), with Census field projects developing efficient approaches for the exploration of each. Realms were intended to encompass all major ocean systems and taxa, but were also selected to take advantage of the best available technologies. The approaches adopted to study these realms and the major findings in each realm are described in **Figure 23**.

Human Edges

The continental shelves are very gradually sloping borders between the edges of continents and ocean basins. While the



[a]Coral reefs are found in the warm waters of the Atlantic, Pacific, and Indian oceans.
[b]Microbial environment encompasses the entire world ocean.
[c]Ice oceans occur at both poles.

Figure 23 Six realms of the global ocean studied by Census of Marine Life scientists.

shelves comprise only 10% of ocean areas, they contain most of the known marine biodiversity and lie mainly within the exclusive economic zones of nations. For the purposes of the Census, these "Human Edges," from the high tide line to the edge of the continental shelf, were divided into nearshore and coastal zones.

Nearshore Zone

The Census defines the nearshore as the region between the high tide line and 10 m water depth. The accessible nearshore area has been studied in minute detail in many locales around the world. The nearshore, however, stretches for more than a million kilometers around all oceans and across latitudes and climates. Testing nearshore ecological hypotheses required that researchers use similar methods in nearshore areas across all latitudes, climates, and ecosystems. This linkage was accomplished in the field project Natural Geography in Shore Areas (NaGISA – "coastal environment" in Japanese).

NaGISA initiated the first global nearshore biodiversity inventory, sampling more than 200 sites and establishing more than 40 long-term monitoring sites around the world. The project discovered new species and documented species and habitats in areas not previously found, developed sampling protocols that can be applied globally, engaging communities and increasing public awareness of the importance of nearshore biodiversity (**Figure 24**).

Coral reefs were another nearshore system investigated. The global Census of Coral Reefs (CReefs) field project was national (Australia, USA), regional (Caribbean), and international in scope. CReefs conducted the largest concerted exploration of life in the world's coral reefs, discovered thousands of new species and records, filled data gaps for lesser-known species, and collected thousands of DNA samples, many new, for the Barcode of Life. It also developed a standardized tool, Autonomous Reef Monitoring Structures, to assess spatial patterns of cryptic biodiversity globally and monitor biodiversity impacts of climate change and ocean acidification. Its Coral Reef Biodiversity Resource Locator aids scientists and managers globally (**Figure 25**).



Figure 24 Census scientists studying the near shore developed standardized protocols to establish baseline information and long-term monitoring of nearshore sites and allowed global comparison of biodiversity along the coastline.

Coastal Zone

Recent crises in fisheries have forced a re-examination of how single species are managed and new strategies for management have evolved, including for trans-boundary species in coastal nations. The Census program for assessing and explaining diversity, distribution, and abundance – especially of non commercial species – may help speed evolution of new management strategies. It accelerated and increased knowledge for the management of commercial species, while adding to general knowledge about and understanding of the global fisheries.

One coastal field project, the Gulf of Maine Area Program (GoMA) was designed to identify and collect the biological knowledge necessary for ecosystem-based management in a large marine environment. GoMA mined many data sources to assemble a list of >4000 named species in a new register and other databases, and estimated the abundance and diversity of microbes. Novel statistical analyses show that habitat explains about a third of the variation in distribution and abundance of the abundant fish and invertebrates, which can contribute to future management decisions. New technologies increase researchers' ability to sample and record patterns of species distributions, including synoptic views of very large schools of fish (**Figure 26**).

Another coastal project was the Pacific Ocean Shelf Tracking Project (POST). The POST deployed the first continental-scale array to acoustically track marine animals, and followed juvenile salmon thousands of kilometers from river to ocean, estimated survival of some ocean-migrating salmon stocks, and discovered surprisingly vast migrations of sturgeon and other species across national boundaries. With an expanding library of 16,000 tracks from 18 different species from salmon to jumbo squid, the project built a clearinghouse for sharing telemetry data, contributed to the designation of critical habitat for threatened green sturgeon, and served as a prototype for the IOC GOOS Ocean Tracking Network (**Figure 27**).

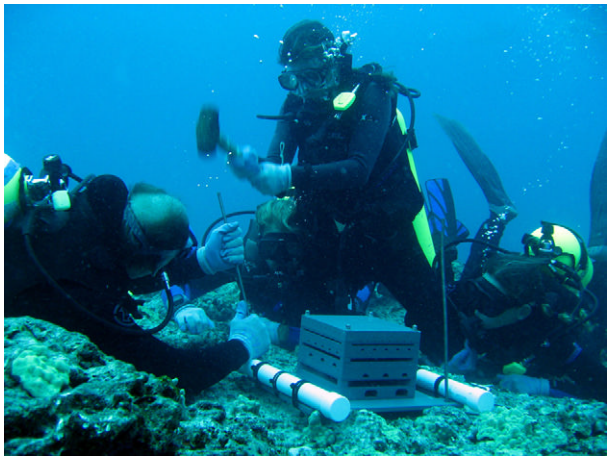


Figure 25 At the end of the decade, the Census had installed more than 600 Autonomous Reef Monitoring Structures (ARMS), to assess spatial patterns of cryptic biodiversity globally and monitor biodiversity impacts of climate change and ocean acidification. The ARMS is the box-like structure in the foreground of the image.

Hidden Boundaries

The Census also investigated regions for which there was little prior biological data. The slope region of continental margins that begins at the edge of the continental shelf and extends to ocean basins were considered in the Hidden Boundaries category because they were difficult to sample and little studied. The deep sea floor below the base of the continental slope covers 54% of the Earth's surface, more than the total continental mass (29%). This region, from ~4000 to 6000 m in water depth, has large areas of flat areas referred to as abyssal plain, where only very small areas have been sampled and very little was known about life. Hence, they were also included in the Hidden Boundaries realm.

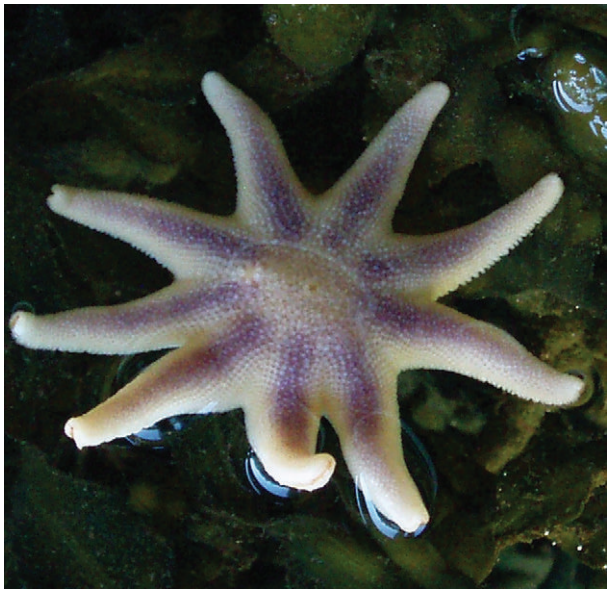


Figure 26 The 4000 plus species in the Gulf of Maine searchable register include the pictured sea star, *Solaster endeca*, found along the shore of Cobscook Bay, Maine. The Census project found twice as many species and thus greater biodiversity in the Gulf than previously estimated.

Slope Zone

Exploration of continental slopes has been inhibited by distance from shore, depth, complex current systems, and geomorphology. Sonar and seismic images of the lower margins have recently revealed that apparently uniform slopes hide mixtures of rock, sand, mud, and methane hydrates. The Census project, Continental Margin Ecosystems on a Worldwide Scale (COMARGE), explored slopes of all continents during more than 60 expeditions. A meager rain of detritus was found to sustain ribbons of life along mud bottoms with highest diversity discovered at mid-slope depths. Off Mauritania, deep-sea coral structures greater than 400 km long were discovered. Elsewhere vast mats of microbe-based ecosystems lived off seafloor methane. The project documented the sensitivity of continental slopes to global changes and to exploitation of deposits of petroleum and gas (**Figure 28**).

Abyssal Plain

Exploration of the ocean's abyssal plains has been hindered by time, distance, and the resources required to investigate this hidden realm. The Census of Diversity of Abyssal Marine Life (CeDAMar) was launched to study the abyssal plain biodiversity of the endo-, epi-, and hyper-benthic organisms (those living in, on, or directly above the sediment).

CeDAMar explored the abyssal plains of major ocean basins, particularly the southern Atlantic and the Southern Ocean, described more than 500 new abyssal species from unicellular animals to large squid, investigated the feeding of abyssal organisms, and mapped the distribution of deep-sea life. The project compiled data for establishing protected seafloor outside national jurisdiction and documented that climate change, human debris, and seabed mining affect abyssal life (**Figure 29**).



Figure 27 With acoustic tags implanted in five species of young Pacific salmon and detected by curtains of receivers along the coast, scientists reconstructed the direction, speed, and timing of migrating individuals. Signals from fish like those in the photograph showed the coastal regions and seasons that either favor or discourage survival.



Figure 28 A close-up photograph of the sea cucumber *Enypniastes* caught at 2750 m on the continental margin in the Celebes Sea between Indonesia and the Philippines reveals its mud-filled intestine through its transparent body.



Figure 29 The tiny copepod *Ceratonotus steiningeri*, about three or four times as long as a hair is wide, was found in the Angola Basin.

Central Waters

Between the continental margins are the ocean basins and open deep water. This region was characterized by the Census as Central Waters. These large open ocean areas are home to at least 40% of the world's primary production of biomass. The Census divided the Central Waters into the Light Zone, from the ocean surface to 200 m water depth, and the Dark Zone, from 200 m to the ocean floor.

The focus of study on the organisms that drift in the Light Zone was the question of whether this community was consistent globally. This question required molecular tools and global assessment of populations throughout the ocean. In contrast, there was no question about the basin-wide or even global connections among the large swimming pelagic animals in the Light Zone. New technologies used by the Census made it possible to provide realistic estimates of the global distribution and abundance of animals in this realm. Although many of the tracked species also spend time in the Dark Zone, they were categorized in the Light Zone realm where they typically feed.

Light Zone (Drifters and Swimmers)

The foundation of life in the oceans is the photosynthetic activity of phytoplankton, including cyanobacteria, and microscopic algae. This requires sufficient light intensity, present to ~200 m water depth depending on water conditions. In this top 200 m live the phytoplankton that comprise almost 95% of total marine productivity. The open ocean phytoplankton are consumed by a community of extremely diverse zooplankton dominated by planktonic crustaceans called copepods.

The Census of Marine Zooplankton (CMarZ) was undertaken to accomplish global-scale analysis of all marine zooplankton groups using new and emerging technologies including molecular, optical, and acoustical imaging, and remote detection. Efforts were initially focused on DNA barcoding of existing specimen collections to identify cryptic species among cosmopolitan groups. In the end, the Census of Marine Zooplankton project produced a global view of diversity and distribution of holo-zooplankton, the animals that drift in ocean currents throughout their lifespan. Researchers collected more than 10,000 samples from every ocean basin. Training 250 new taxonomists, they discovered more than 85 new species, seven new genera, and two new families, many of which are already formally described. They used genetic profiles to develop characteristic fingerprints for all the zooplankton life in any volume of water (**Figure 30**).

At the other end of the spectrum was the Tagging of Pacific Predators (TOPP), which worked with marine animals to create a view of vast open ocean habitats as seen by the animals themselves, especially top predators. Knowing the behavior of the top predators allows inferences about the distribution and abundance of many other organisms that live in the ocean, such as where prey species accumulate. TOPP used electronic tags to follow 4300 marine predators, representing 23 species, on their journeys throughout the Pacific Ocean. These animals undertook basin-scale migrations, moving from the poles to the tropics and from continent to continent, and followed migratory corridors to ocean "hotspots," where large numbers of individuals – and multiple species – congregate for extended periods. Tags on these "animal oceanographers" collect

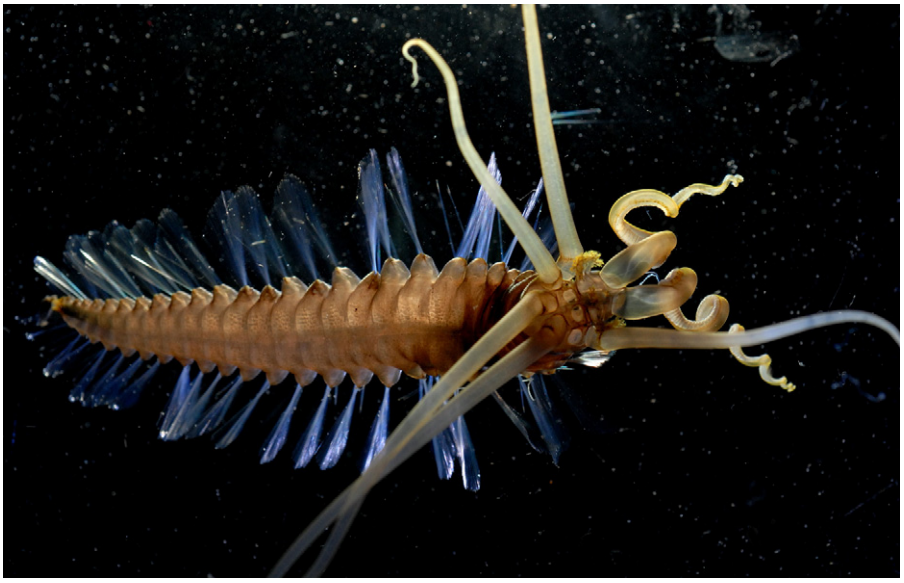


Figure 30 When US and Filipino zooplankton scientists searched the Celebes Sea for new species in its deep water, they discovered and temporarily named this worm, about 10 cm long with tentacles, squidworm. One of the small animals that floats or swims weakly, the squidworm turned out to be a new species of polychaete in the class of annelid worms.

invaluable environmental data, yielding new insights into their requirements from the structure of the ocean itself (**Figure 31**).

Dark Zone (Mid-Water and Bottom-Water)

The challenge of investigating the Dark Zone was taken on by a multinational group of researchers who formed the Patterns and Processes of the Ecosystems of the Northern Mid-Atlantic Project (MAR-ECO). The objective of MAR-ECO was to explore and understand the distribution, abundance, and trophic relationships of the organisms inhabiting the middle and deep waters of the mid-oceanic North Atlantic, and identify and model ecological processes that cause variability in these patterns. A multidisciplinary trans-Atlantic team of researchers, led by a steering group with representatives from the United States, Germany, the United Kingdom, Iceland, Portugal, France, and Norway, carried out investigations between Iceland and the Azores using ships and submersibles. MAR-ECO explored marine life along the Mid-Atlantic Ridge, the world's longest mountain range, which rises from the seafloor 4500 m deep. Explorers found approximately 1000 species, from small crustaceans to whales, including at least 30 new species. They found both mid-water animals from deep basins and bottom-living animals along slopes. Abundance generally peaked where cool and warm waters met. Project findings supported protection of areas of the Mid-Atlantic Ridge against harmful bottom fishing totaling 330,000 km^2, larger than Italy (**Figure 32**).

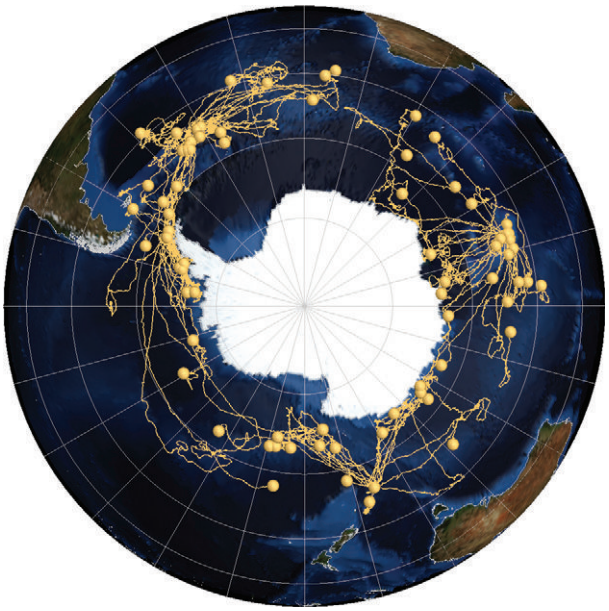


Figure 31 Biologging, the use of miniaturized electronic tags to track animals in the wild, revealed unknown behaviors, movements, physiology, and environmental preferences of a variety of ocean animals. Biologging has shown that communities of southern elephant seals, *Mirounga leonina*, from different islands surround the harsh environment of Antarctica, demonstrating the principle that animals inhabit even the farthest reaches of the ocean.

Active Geology

Seamounts, hydrothermal vents, and cold seeps are geologically active areas, grouped together in the Active Geology realm. The Census project, the Biogeography of Deep-Water Chemosynthetic Ecosystems (ChEss), was launched to discover new hydrothermal vents and cold seeps, to assess the diversity, distribution, and abundance of their fauna in relation to other chemosynthetic ecosystems, and to explain the differences and similarities at the global scale. The ChEss



Figure 32 This remarkable red worm, between 20 and 30 cm long, discovered by scientists investigating the Mid-Atlantic Ridge turns out to belong to a completely new family. In fact, it is not a worm, as they initially thought, but an "enteropneust" that has more in common with vertebrates than with worms.

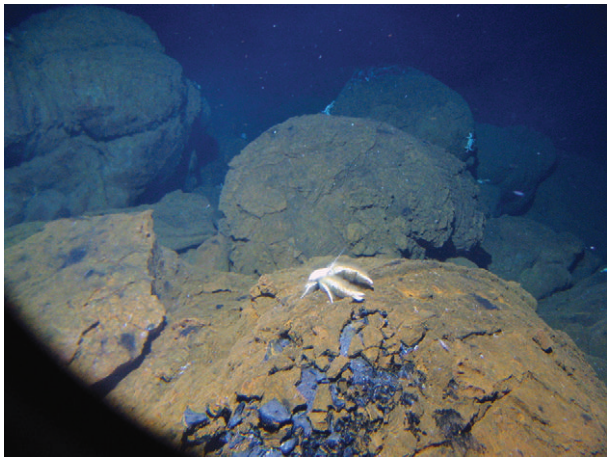


Figure 33 In the inhospitable environment on a vent in the southeast Pacific near Easter Island, Census explorers discovered the yeti crab, *Kiwa hirsuta*. Given a species name for its hirsute appearance, its white, hairy look has prompted the nickname yeti for the abominable snowman, or yeti. The remarkable crab belongs to a new species and new genus and family, too.

scientific steering group identified a number of specific locations and three major priority regions – the Equatorial Atlantic Belt, the Southeast Pacific, and New Zealand regions along with the development of a database on deep water chemosynthetic species. ChEss also worked to establish close collaboration with other ocean science programs that will venture to sites of vents, helping to make future research and exploration affordable.

ChEss explored life at seeps, vents, and whale falls, where abundant life thrives, fueled by chemical reactions rather than sunlight. Using advanced robots, the project extended the known limits to such life – farther north (72°N) and farther south (60°S), deeper than 4900 m and hotter than 407 °C.



Figure 34 The variety that seamounts contribute to ocean topography can also contribute to its biodiversity. In the photograph, a bright orange brisingid sea star is prominent among corals on a seamount. The sea star is feeding by raising its arms to capture passing food from the Antarctic Circumpolar Current.

Project scientists described about 200 new species, bringing to more than 1000 the number of described species known to live without the sun's energy (**Figure 33**).

The project focusing on seamounts, Census of Seamounts (CenSeam) synthesized what was known about biodiversity, which will help direct future field efforts toward a comparative ecology of seamounts. CenSeam encompassed up to 100,000 seamounts rising 1000 m or more from the seafloor. The project discovered new seamounts, species, and communities and found that seamount inhabitants broadly resemble those on adjacent continental slopes. Yet, owing to isolation and slow growth of deep animals, life on seamounts can struggle to recover from disturbances such as fishing. The project's central repository of global seamount data is informing international negotiations about protecting areas in the high seas (**Figure 34**).

Ice Oceans – Arctic and Antarctic

Two Census field projects were designed to study the ocean areas at opposite poles of the Earth. They shared the need for specially equipped ice-breaking ships, forcing highly integrated sampling schedules for the entire water column from surface ice to abyssal plains. The first Ice Ocean field project launched was Arctic Ocean Diversity (ArcOD). Its goal was to assemble existing knowledge of biodiversity in the least known ocean, direct new international explorations using new technologies, and create a framework for understanding and predicting biological changes in this rapidly melting ocean.

ArcOD established the baseline of marine life in the region. Its search on, in, and under sea ice, in deep basins, and along continental shelves cataloged more than 7000 animal species, of which at least 70 are new, plus thousands of microbes. Project researchers documented recent northward extensions of ranges of invertebrate and fish species, and a rising ratio of warm- to cold-water species. Making more than 250,000 records accessible online, the project allows measuring changes



Figure 35 Deep in the frigid Canada Basin, a remotely operated vehicle of the Arctic Ocean project captured the pictured jellyfish of the genus *Crossota*.

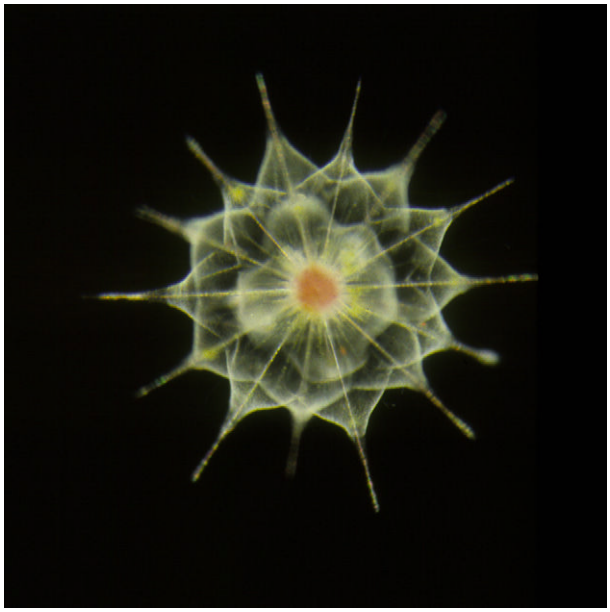


Figure 36 The acantharians are one of the four types of large amoebae (microbial eukaryotes) known to occur in marine open waters. The 2-mm-wide specimen, about the size of a flea, was captured off Bermuda.

in Arctic life caused by humanity and natural forces (**Figure 35**).

Like its northern counterpart, the Census of Antarctic Marine Life (CAML) assembled rich biological data on the Southern Ocean, and encouraged biodiversity sampling on all cruises in the region, particularly during the focal period of the International Polar Year, IPY, (March 2007–March 2009). The goal was to couple this new understanding of biology with the complex current dynamics that control gene flow through the world's oceans. CAML's scientific steering group represented seven nations with research interests in the region, and extended to other South American countries participating in IPY projects in benthic biology.

CAML coordinated 18 research voyages during IPY, the widest-ever survey of the region to date, and provided the baseline for marine life in the Southern Ocean. CAML also helped establish the monitoring of its change by playing a key role in the development of the Southern Ocean Observing System. The project documented more than 16,500 taxa, including hundreds of new species. The researchers identified a single region of seafloor life under the fastest current on the planet, found colonists from deeper water beneath melting ice shelves, and showed how Antarctic waters may incubate new species.

Microbes – The Microscopic Ocean

The number of microbes in marine environments is staggering – with estimates of tens or even hundreds of millions of them – and they have a dominant role in ocean processes. To complete the Census of Marine Life, the International Census of Marine Microbes (ICOMM) was established to develop a highly resolved biodiversity database for marine microbes and to understand how microbial populations – the oldest life on the planet – evolve, interact, and redistribute on a global scale.

Microbiologists from 25 countries joined forces and determined the range of genetic diversity, distribution, and abundance of oceanic microbes. Investigators found diversity at least 10–100 times the previous estimate. A liter of sea water contained about 38,000 kinds of bacteria. In more than 1200 widely distributed marine samples from water, sediment, and around animals, researchers found some microbes live everywhere, while others live in only a few places. They discovered a rare biosphere in which "rare microbes" account for most ocean diversity (**Figure 36**).

See also: Ecosystem Function Measurement, Aquatic and Marine Communities. Endangered Marine Invertebrates. Freshwater Ecosystems, Human Impact on. Invertebrates, Marine, Overview. Marine and Aquatic Communities, Stress from Eutrophication. Marine Ecosystems. Marine Mammals, Extinctions of. Marine Sediments

References

Ausubel JH, Trew Crist D, and Waggoner PE (2010) *First Census of Marine Life: Highlights of a Decade of Discovery.* Washington, DC.

Bergstad OA and Godø OR (2003) The pilot project "Patterns and processes of the ecosystems of the northern Mid-Atlantic": Aims, strategy and status. *Oceanologica Acta* 25: 219–226.

Buddemeier RW (2002) *Climate Change and Coral Reefs.* Presentation on Capital Hill Ocean Water.

Danovaro R, Dell'Anno A, Pusceddu A, Gambi C, Heiner I, and Kristensen RM (2010) The first metazoa living in permanently anoxic conditions. *BMC Biology* 8: 30.

Falkowski PG and de Vargas C (2004) Shotgun sequencing in the sea: A blast from the past? *Science* 304: 58–60.

Knowlton N (2010) *Citizens of the Sea, Discoveries of the Census of Marine Life.* Washington, DC: National Geographic.

Products (2006) http://www.kgs.ku.edu/ (accessed December 2006).

Richer de Forges B (2006) Découverte en mer du Corail d'une deuxième espèce de glyphéide (Crustacea, Decapoda, Glypheoidea). *Zoosystema* 28(1): 17–29.

Rutherford S, D' Hondt SD, and Prell W (1999) Environmental controls on geographic distribution of zooplankton diversity. *Nature* 400: 749–753.

Seamounts Online (2006) http://seamounts.sdsc.edu (accessed December 2006).

Schmitz WJ (1995) On the interbasin-scale thermohaline circulation. *Reviews of Geophysics* 33: 151–173.
Snelgrove PVR (2010) *Discoveries of the Census of Marine Life-Making Ocean Life Count*. London: Cambridge University Press.
Trew Crist D, Scowcroft G, and Harding Jr JM (2009) *World Ocean Census*. Toronto: Firefly Books.
Weaver PPE, Billett DSM, Boetius A, Danovaro R, Freiwald A, and Sibuet M (2004) Hotspot ecosystem research on Europe's deep-ocean margins. *Oceanography* 17(4): 132–143.
Welch DW, Boehlert GW, and Ward BR (2002) POST – The Pacific Ocean salmon tracking project. *Oceanologica Acta* 25: 243–254.

Central America, Ecosystems of

Rodolfo Dirzo, Stanford University, Stanford, CA, USA
Maria Argenis Bonilla, Universidad Nacional de Colombia, Bogotá, Colombia

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Dirzo, R., volume 1, pp 665–676, © 2001, Elsevier Inc.

Glossary

Biogeography The study of the distribution of plant and animal life in the earth's environment and of the biological and historical factors that produced this distribution.
Geomorphology The study of the surface configuration of the earth, especially the nature and evolution of current land forms, their relationships to underlying structures, and the history of geological activity as represented by such surface features.
Land-use change The changes in the way land is used at a given location, for example, when a forested land is converted to agricultural fields or when a forestry plantation is changed to other types of use.
Páramo This term refers to a cold and humid landscape located above the tree line.
Plate tectonics A modern geological theory of tectonic activity according to which the earth's crust is divided into a small number of large, rigid plates whose independent movements relative to one another cause deformation, volcanism, and seismic activity along their margins.

Central America as a Biogeographic Region

The definition of Central America varies considerably among sources. The most popular notion of what constitutes Central America is based on a geopolitical criterion. According to this, Central America is that region of the Western Hemisphere that constitutes the isthmus lying between Mexico and South America, encompassing the countries of Guatemala, Belize, Honduras, El Salvador, Costa Rica, and Panama. Some geographers also include at least one (Jamaica) or most (Cuba, Haiti, the Dominican Republic, and Barbados) of the Antillean Islands as part of Central America, but the most accepted views only consider the continental part of the region as Central America. However, even this widely accepted geopolitical definition is not entirely satisfactory for the purposes of this encyclopedia. A description of the ecosystems of Central America requires a more natural definition. Thus, this article will consider a biogeographical and geomorphological approach, according to which Central America is more naturally defined as an elongated, tapering isthmus that begins in the narrowest part of southern Mexico, at the Isthmus of Tehuantepec (**Figure 1**). This narrow region divides the area of volcanic rocks to the northwest from the folded and considerably faulted structures of the more conventional Central America to the east. However, a more complicated delimitation is that of the Mexican sector that extends into the Yucatan Peninsula. Although this sector naturally extends northward from Chiapas, into the States of Campeche, Quintana Roo, and Yucatan, with the vegetation following a gradient of precipitation, it reaches a very high latitude, more northerly than what typically would be considered Central America. For that reason, we have included a portion of the Peninsula, somewhat arbitrarily cut at the latitude of 20° N, leaving out the drier part of the peninsula. The southernmost limit of Central America is the valley of the Atrato River in Colombia, located just to the east of the Panama–Colombia border, where the massive Darién jungles begin.

An important biogeographic aspect derived from the location of the region is that this narrow strip of land currently connects the neotropics of South America and the Nearctic zone of North America, two major biogeographic realms of Earth. An important additional implication is that Central America currently operates as a corridor for terrestrial organisms from both realms and as a barrier to marine organisms from the Caribbean and the Pacific seas. However, such situation has not always been the case. A variety of hypotheses have been suggested regarding the possible configuration of the area in the past. That said, a consensus exists that from the latter Mesozoic, until approximately 5 or 6 million years ago in the Pliocene, no continuous terrestrial connection existed between South and North America *via* Central America. Moreover, throughout the largest part of the Cenozoic era (65 to approximately 5 Ma), the region comprised from Nicaragua to the northernmost part of Colombia was probably an archipelago similar to the Lesser Antilles of today, providing only, perhaps, an occasional pass for terrestrial life forms. Most likely, the continuous connection that Central America currently provides between the north and the south has been in existence for only approximately the past 3 million years, and the mountainous backbone of the region did not reach its present elevation until the latter part of the Cenozoic. This means that the variety of habitats that currently characterize the region is a rather recent phenomenon and that the ecosystems and biodiversity of Central America in general are a recent blend that has resulted from a complex paleobiogeographic history. A review of the paleobiogeographic history of the region is beyond the scope of this article. However, some salient aspects that, in addition to what we described previously, impinge on the current biodiversity of Central America include the following (Rich and Rich, 1983). Until approximately 3 Ma the mammalian faunas of Central America were composed predominantly of North American elements and only after this time did a north–south interchange begin with

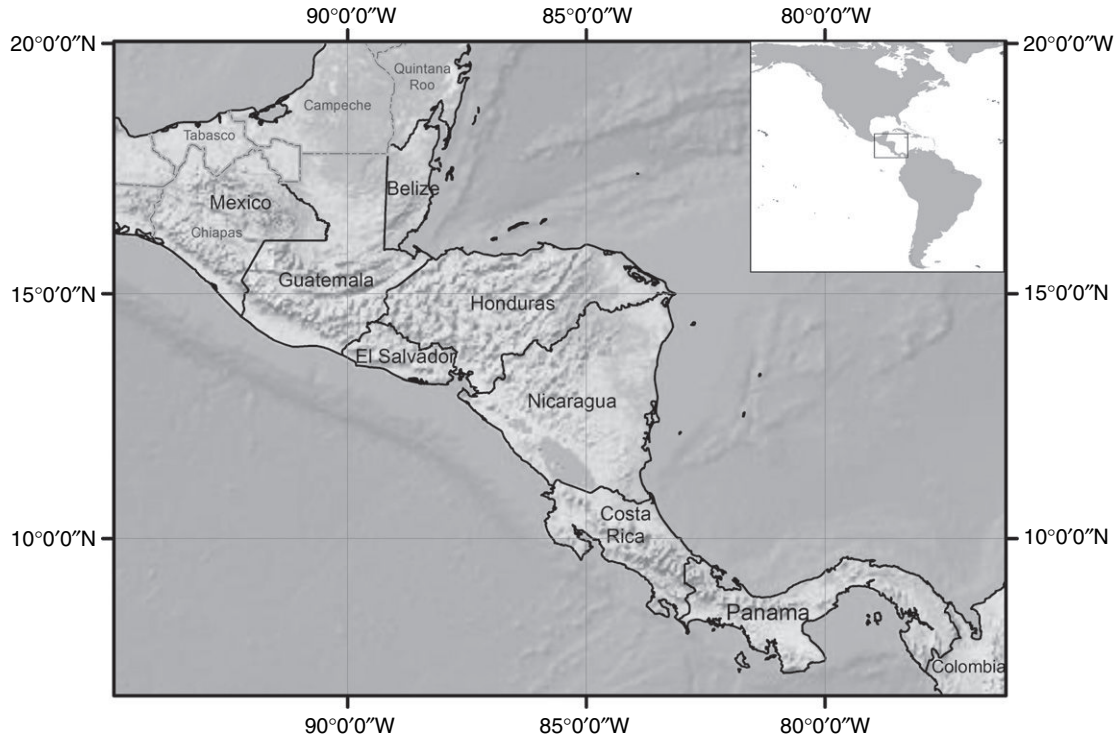


Figure 1 Geographic location of Central America, and delimitation of the region considered in this article, including the southern part of Mexico. The Mexican portion of Central America starts at the Isthmus of Tehuantepec and includes the states of Tabasco, Chiapas, and a sector of the Yucatan Peninsula (see text for more details). Reproduced from Meffe GK and Carroll CR (1997).

a significant incursion of vertebrates from South America. It is possible that an endemic small-mammal fauna might have developed at about this time in Central America and southern North America and dispersed toward the south. The greatest affinity of Central American angiosperm (flowering plants) floras with South American ones indicates that dispersal from the south was more pronounced for plants than for animals, although Pleistocene cooling periods promoted dispersal of typically montane plants toward the south. In addition, at the height of the North American glaciation period in the Pleistocene, climatic conditions favored the dispersal of grassland- and savanna-adapted forms. Nevertheless, these ecosystems must have become as fragmented as they are today by the increasing precipitation of the latter part of this epoch. Finally, it was not until close to the end of the Pleistocene that the tropical rain forest ecosystem became widespread in Central America. This has the implications that this ecosystem, geologically speaking, is very young in the region and that its expansion decreased the effectiveness of the terrestrial connection between North and Central America. This has played a determining role in the modern composition of the biodiversity and ecosystems currently seen in Central America.

Biodiversity of Central America: An Overview

An important facet of the biodiversity of Central America is its diversity of natural ecosystems. Because of its complex geological history and intricate topography, the region of Central America includes a mosaic of ecological conditions, in which several of the major natural ecosystems of the earth are represented. Despite the fact that the region has a restricted latitudinal range, of approximately 12° (from about 8° to 20° N), and the overall climate is of a marked tropical affinity, significant climatic variations occur due to the topographic complexity rather than to latitude. Accordingly, the ecosystems of Central America are related to three major elevational/climatic situations. The lowlands, at elevations of sea level to about 900 m, with hot climates of average temperatures of 24 °C, provide the conditions for the development of tropical forests, both seasonally dry and evergreen. The temperate zones, from about 900 to 1800 m above sea level (masl) and average temperatures of 18–24 °C, are occupied by the exuberant tropical cloud forest and temperate zone-related coniferous and broadleaf (mostly oak) forests. The highlands, from 1800 to 3400 masl and with temperatures of 12–18 °C, sustain mostly high-elevation grasslands and páramos. Superimposed to this elevational gradient, the topography and the patterns of wind circulation of the region determine the existence of marked contrasts in rainfall. The cold California current of the Pacific coast chills the air, preventing it from absorbing much water vapor from the ocean and thus reducing the chances of precipitation on this side of the region. In contrast, the easterly winds of the Atlantic coast absorb much of the moisture, which is then deposited on this side of Central America as these winds flow up and over the mountains of the region. As a result, the plains and eastern mountain slopes of the Caribbean coast receive approximately twice as much precipitation as the plains and western mountain slopes of the Pacific coast.

Such a mosaic of environmental conditions determines the existence, within a restricted geographical range, of a dramatic variety of ecosystems within a given latitudinal position. For example, within a latitudinal range of only approximately 2° and an area of approximately 51,000 km^2, the country of Costa Rica harbors 12 distinct life zones (vegetation types), one more than in the eastern United States, even though Costa Rica is only the size of West Virginia. This sample of ecosystems ranges from the seasonally dry tropical forests in the Guanacaste province, to the high-elevation grasslands (páramos) of the central part of the country and the wet tropical forests of the Atlantic coast. At a more local scale, a southwest–northeast transect of approximately 200 km in a straight line (i.e., not considering the ruggedness of the terrain) across the mountainous region of the state of Chiapas, in southern Mexico, harbors a collection of 12 distinct vegetation types: mangroves, palm communities, savannas, seasonally dry tropical forests, tropical moist and wet forests, cloud forests, deciduous sweet gum (i.e., *Liquidambar*)-oak forests, oak forests, pine forests, grasslands, and high-elevation grasslands.

The conglomeration of natural ecosystems present in Central America is responsible for the fact that this region contains a significant proportion of the biodiversity of the planet in terms of species richness. The mere fact that the tropical rain forest, the most diverse biome of the planet, is widely distributed in the region implies that the biodiversity of Central America is of great significance. The presence of tropical rain forest is in fact one of the distinctive criteria defining the countries of megadiversity. Another important criterion for the definition of megadiversity is the presence of coastline and the ratio of coastline to land surface. The Central American coastline is approximately 5700 km (2900 km on the Pacific coastline and 2800 km on the Caribbean side); thus, there is a coastline : surface area ratio of 0.011, which is comparable to that of some of the most important megadiversity countries. Undoubtedly, marine ecosystems and the presence of coral reefs (arguably the most species-rich marine ecosystem of Earth), including some of the most important coral reefs of the Western Hemisphere (located on the Caribbean coast of Central America), add to the diversity of the region.

The well-established gradient of decrease in species richness with latitude implies that, overall, the tropical ecosystems of Central America are not expected to be as species-rich as their more equatorial counterparts. However, a compilation of the scattered and, most undoubtedly, incomplete information on species richness in some of the better known organisms of the region's countries, suggests that, despite its latitudinal position, species richness is outstanding (**Table 1**; see also INBIO's web page). The range of species richness, cumulative for plants and vertebrates, is 4270 (El Salvador) to 13,091 (Costa Rica). Furthermore, the Mexican States that form part of the region are also considerably diverse, the most notable of which is Chiapas. With an area of 73,887 km^2, it amounts to 5665 species of those groups. The other Mexican States, with lower values (Campeche, 1353; Quintana Roo, 1358; Tabasco,1373), nevertheless increase the species richness of the region (CONABIO, 2008). The diversity of ecological conditions, and therefore of distinct ecosystems, dictates that the identity of species and species composition change considerably among localities within this restricted area. Thus, species turnover (or beta diversity) is likely to be very high in this region. The information available for some areas of Central America (e.g., southern Mexico) suggests that indeed this is the case. However, biological inventories are still limited and species turnover and the overall species richness of the region remain unknown. Although knowledge for some groups, primates and birds, for example, can be regarded as adequate, knowledge is particularly limited for other groups, especially invertebrates. A few isolated cases, however, suggest a significant species richness, as exemplified by the known number of species of Lepidoptera in Costa Rica,1114 (DeVries, 1983), and 20,000, but projected to reach 40,000 species of Hymenoptera (Hanson and Gauld, 1995).

In addition to the quantitative aspects of the biodiversity of Central America, a qualitative feature deserves special consideration. This refers to the combination of floristic and

Table 1 Species richness for six major groups of organisms reported for seven Central American countries

	GUA[a]	BEL[b]	HON[c]	ELS[d]	NIC[e]	COS[f]	PAN[g]
Plants	7745	3750	7525	3411	9000	11,451	10,444
Mammals	192	152	229	147	187	239	259
Birds	722	574	715	542	673	857	957
Reptiles	245	126	212	98	170	226	229
Amphibians	142	41	121	32	92	183	179
Fish (fresh water)	240	116	130	40	157	135	206
TOTAL	9295	4759	8932	4270	10,279	13,091	12,274

GUA, Guatemala; BEL, Belize; HON, Honduras; ELS, El Salvador; NIC, Nicaragua; COS, Costa Rica; PAN, Panama.

[a]Jolon-Morales MR (2005) *Recopilación de información sobre biodiversidad en Guatemala*. Nueva Guatemala de la Asunción, Guatemala: INBio, CONAP.

[b]Meerman JC (2005) *Compilation of Information on Biodiversity in Belize*. Belmopan, Belize: INBio, Belize Forest Department.

[c]Portillo H (2007) *Recopilación de la información sobre la biodiversidad de Honduras*. Informe Final de consultoría. Tegucigalpa, Honduras: INBio, DiBio.

[d]Gallo M (2005) *Estado del Conocimiento de la Biodiversidad en El Salvador*. San Salvador, El Salvador: INBio, MARN.

[e]Rueda-Pereira R (2007) *Recopilación de información sobre la biodiversidad de Nicaragua*. Managua, Nicaragua: INBio, Universidad Nacional Autónoma de Nicaragua-León, Nicaragua.

[f]Obando V (2007) *Biodiversidad de Costa Rica en cifras*. Santo Domingo de Heredia, Costa Rica: Editorial INBio.

[g]Fundación PANAMA (2007) *Informe sobre el Estado del Conocimiento y Conservación de la Biodiversidad y de las Especies de Vertebrados de Panamá*. Panamá, República de Panamá: INBio, ANAM.

faunistic elements of the two major biogeographic zones that Central America is responsible for bringing together. This is readily noticeable in the case of flora. The elements that compose the flora of Central America belong, to a great extent, to two major floristic realms: the Arctoterciary, comprising the extra tropical territories of the Northern Hemisphere with representative genera, such as *Pinus* spp. (pines) and *Quercus* spp. (oaks), and the Neotropical, which includes the territories of tropical America, with representative genera such as *Swietenia* (caoba or mahogany) and *Cedrela* spp. (tropical cedar). Some Andean elements that belong to the paleooceanic realm, which comprises territories of the Andes, South America, South Africa, Australia, and New Zealand, are also represented in the Central American flora. These Andean elements include genera such as *Podocarpus* and the spectacular herbaceous plants of *Gunnera*.

The elements of the Neotropical realm are best represented in the tropical forests of the lowlands, whereas the Arctoterciary and Andean elements are predominant in the vegetation of the temperate and cold highlands. Frequently, however, there is no clear demarcation between these types of floristic elements and in several localities of Central America there is a mixture of them producing, for example, remarkable tropical rain forests at elevations of 600–900 masl, which combine the typical tall trees, lianas, palms, and tree ferns with gigantic oaks and other trees of a clearly northern affinity (e.g., *Chaetoptelea*). Conversely, some cloud forests and deciduous forests, with a predominance of Nearctic elements, may include trees of clear tropical affinity such as *Cecropia* or *Nectandra*. Another remarkable feature of the Nearctic elements of Central American forests is that several of them comprise plants characteristic of the eastern United States, with a major interruption in their distribution in Texas but that penetrate into eastern Mexico and further south into Chiapas and Central America. Some of these are taxa as representative as *Liquidambar styraciflua* (sweet gum), *Nyssa sylvatica* (black gum), and *Ostrya virginiana* (hophornbeam).

However, elements typical of the drier areas located toward the north of the Tehuantepec Isthmus are absent in the state of Chiapas (and in most of tropical Central America), but they reappear further south in the dry lands of Guatemala (Departments of Zacapa, El Progreso, and Chiquimula). This is congruent with the assumption, presented previously, that the climatic changes leading to the recent expansion of tropical rain forests must have fragmented the distribution of elements of arid and semiarid ecosystems such as the arborescent cactus *Myrtillocactus* and trees such as *Juliania adstringens* and *Apoplanesia paniculata*.

To this amalgamation of elements of tropical, temperate, and Andean affinity, an additional contingent of endemic elements must be added. Although some elements are distributed in most of Central America, others have a more restricted distribution within the region. For example, although several species and genera typical of the floras of Guatemala and Belize extend their distribution into Chiapas, other genera of the rain forests of these two countries are not present in Chiapas or any other tropical forests of Mexico. This also indicates that for a portion of the flora of Central America, the northernmost limits of distributions are Belize and northern Guatemala. However, the most important genera of trees of the temperate and cold ecosystems of Chiapas, including *Pinus*, *Quercus*, *Liquidambar*, *Carpinus*, *Abies*, *Cupressus*, *Juniperus*, and *Taxodium*, all of which are of northern affinity, find their southernmost distribution in Guatemala or Nicaragua. The exception is the oaks, which extend well into the south, in the Colombian Andes. Likewise, several species of oak that in all probability evolved in Central American forests (e.g., *Quercus benthami*, *Quercus policaulis*, and *Quercus crispifolia*) extend their distribution into Chiapas, which is their northernmost distribution.

A final aspect related to endemism has to do with the relatively recent discovery of new taxa as a result of biological surveys and taxonomic research in the region. In the past 20 or so years, two new endemic plant families have been discovered. One, the family Ticodendraceae, is a tree from Costa Rica, and the other is a herbaceous plant in the family Lacandoniaceae from the Lacandon forests in Chiapas. The latter is a truly revolutionary plant family in which the parasitic plants lack chlorophyll and the hermaphroditic flowers present the reproductive organs in a morphological disposition which is unknown among the flowering plants (male organs in the center of the flower and female organs in the periphery).

In summary, despite the limited knowledge of the region's biodiversity, the available information provides good evidence that the biological diversity of Central America is of special planetary significance. This is not only due to its quantitative aspects (e.g., diversity of species and ecosystems) but also because of the remarkable combination of life forms of different biogeographic affinity, together with endemic elements.

Life Zones and Ecosystems

The most synthetic and revealing way of describing the ecosystems of Central America is through the description of the vegetation, since it is the most obvious physiognomic descriptor of the ecosystem, is the base of the food chain, and provides the structural matrix on which most communities and populations of animals live or indirectly depend on.

A very popular classification system is widely used to describe the ecosystems of Central America. The system, based on the life zone concept elaborated by Holdridge (1967), combines three climatic characteristics of a given region – temperature, precipitation, and an index of the availability of moisture, potential evapotranspiration, as a predictive model to forecast a corresponding vegetation type, or life zone. Although the Holdridge system is widely used, particularly in Costa Rica, the classification system developed by Miranda and Hernández-X (1963) is the one most used for the Mexican vegetation. This system is based on the physiognomy (i.e., the appearance: life forms and height) and phenological character (e.g., whether the predominant plants are deciduous or evergreen) of the vegetation. The conceptual basis for this system is that vegetation physiognomy reflects the adaptive response of plants to the environmental characteristics where vegetation develops. Given the facts that these two systems do not coincide in vegetation nomenclature and the number of ecosystem types resulting from both is very large and detailed, we use here a system that attempts to capture the essential aspects

of the two systems and provides a more simplified classification. In addition, this simplified system attempts to present a nomenclature that is comparable to that used more generally in the literature.

Tropical Rain Forests

The most distinctive aspect of this type of forest in Central America is, as in other regions, their great diversity of species. For example, the Lacandon forest of Chiapas includes approximately 4300 plant species, and the La Selva field station in Costa Rica harbors 1500 plant species and approximately 411 species of birds, 116 species of mammals, and 479 species of butterflies (McDade *et al.*, 1994).

Regarding physiognomy and structure, these forests are characterized by the presence of very tall trees (30–50 m) and by their evergreen vegetation – most of the plants retain their leaves year-round. Most of the trunks of the large trees, besides being straight, have diameters at breast height of between 30 and 150 cm. However, a few, such as the *Ceiba* trees, can have even thicker trunks and their height can reach beyond the forest canopy. Another distinctive feature of these forests is the presence of irregular or undulating contours of the tree trunks of many species, particularly at the base, where they become extended protrusions, approximately triangular in profile, which may play the role of support for the tree. The structures, called buttresses, can adopt several shapes or designs that can often serve as distinctive characteristics of particular species. The bark on the trunks, in general, is either smooth or scaly, with color ranging from light tones in some figs, to dark ones in the mahogany tree and shiny reddish hues in the gumbolimbo (*Bursera*).

Even when not clearly defined, from the canopy down, there is a succession of layers of vegetation (stratification), from the subcanopy to the understory, with a spectacular profusion of life forms. The understory includes a layer of plants, most notably short palms, from less than 1 m up to 8 m. At ground level, the undergrowth is composed of ferns, several herbaceous plants such as gingers, some trailing plants, and the saplings or seedlings of the species found at higher levels.

A direct consequence of this stratification and profusion of plants is that light diminishes from the top down to ground level, where only 1–5% of the available light reaches through. Such limited light availability leads to spectacular morphological and functional adaptations in the plants to compensate for the limitation of this critical resource. Such adaptations are in turn responsible for the existence of other physiognomic and structural features of the forest, including the occurrence of a great variety of climbing plants, both woody (lianas) and herbaceous (vines, mainly in the family Araceae), together with a variety of epiphytic plants (i.e., plants that live on top of other plants). By far, lianas are the most conspicuous climbers since they can reach over 50 cm in girth and, depending on the locality, they can contribute one to three of every 10 trunks. In Central America, even some species of palms have evolved the habit of liana, as is the case of *Desmoncus* spp. and *Chamaedorea elatior*. The Araceae provide the most notable example of nonwoody climbers, not only because they cover the trunks of the trees on which they climb and their abundance on the canopy, but also because of their abundance as trailing plants at ground level, where they spread in search of trunks to climb. Vascular epiphytes, the most evident of which are the orchids, constitute a significant component of the biodiversity, given that the orchid family is commonly the most species-rich family in Central American forests. In addition to orchids, bromeliads constitute another important element of the epiphytic life form of these forests. Although bromeliads are not as specious, they are very abundant so that these two families constitute one of the most distinctive features of the physiognomy of Central American rain forests.

A much less conventional type of climbing plant is that of the hemiepiphytic trees, represented chiefly by the strangler figs, which also constitute a distinctive feature of these forests. Stranglers germinate and take root on the branches of other large trees and then grow downward at an impressive rate, producing many strong trunks that hold on to the ground while covering the tree on which are established, forming a strangling network that eventually kills the initial support tree.

In addition to this diversity of morphological adaptations, most woody plants of the tropical rain forest have a series of ecological and functional attributes that allow them to deal with the problem of light limitation. Such attributes comprise adaptations that fit the dynamic nature of the tropical rain forest. In essence, the winds that occur in these regions, particularly during the rainiest season and winter months (November–February), overturn one or more of the giant trees of the canopy, creating so-called light gaps of up to 800 m^2. Depending on the locality, the incidence of such gap formation can be very intense, ranging from turnover rates (i.e., the number of years elapsed between two successive gaps) of 60 (e.g., in some Mexican forests) to 100 (e.g., in some Costa Rican Forests) years. The result is the production of natural clearings with high light availability and higher temperature than in the shady understory. Many plant species produce dormant seeds equipped with photosensitive mechanisms that allow them to detect the amount and type of light of the clearing and germinate in these conditions. This type of plant constitutes the group of the pioneer species, the most representative of which are *Cecropia* spp. and *Ochroma*, the balsa tree. Pioneer species display a very rapid growth and recolonize gaps by growing together with some of the young plants of the typical shade-tolerant species of the mature forest that might have been growing in the understory. The gap gradually begins to close, and is colonized by other species adapted to increasing levels of shading. These conditions make it impossible for the survival of the pioneer species that did not manage to grow and reproduce on time. The shade-tolerant species, provided with morphological, physiological, and growth mechanisms suited to scarce illumination, make their way to the canopy. This process of opening and closure of the forest creates a dynamic mosaic of vegetation with varying degrees of regeneration. In the tropical rain forests of Central America the mature phase is predominant and the most representative of the described physiognomy. In addition, in this phase, the leaves of the plants are predominantly dark green and sometimes shiny on the upper surfaces, tough in texture, and with sizes that range from medium

(5–10 cm in length) to large (up to 1 m or more). As in other tropical rain forests, leaves have tapering extensions on their tips, which are thought to serve to drain off the excess of water. Colorful, fleshy fruits are very abundant in the plants of the mature forest, but flowers tend to be mostly small and there is a predominance of light shades of white and green; bright and shiny colors are less abundant.

The complex and diverse vegetation serves as a matrix in which a highly diverse fauna is found, particularly insects. Given the great diversity of plants and animals, a distinctive aspect of these forests is the complex network of interactions among different species, particularly between plants and animals: pollination, dispersal (the transportation of fruits and seeds by animals), and herbivory (the consumption of plant tissue by animals). As a result, the shapes, colors, and scents of flowers and fruits and seeds, and a wealth of toxic or attractive substances, determine the behavior, feeding, and sensory patterns of animals. These animals defoliate, seek out seeds, consume nectar and other fluids, guard and defend plants against herbivores, and even attempt to copulate with orchids that resemble and smell like female bees.

Seasonally Dry Tropical Forests

Located at altitudes from sea level up to 1500 m, this kind of forest is characterized by a long dry period (from 4 to 6 months). As a result, physiognomically, they are typified by the deciduous nature of their vegetation: Green and luxuriant in the rainy season and grayish and leafless when it is dry. The highest trees rarely exceed 25 m, and the dominant flora comprises smaller trees, shrubs, and long-branching trees, although vines are also found. Commonly, two vegetation strata can be defined: the canopy, with trees 15–25 m tall, and the understory, with shrubs and treelets 5–15 m tall. The ground layer is sparse in general. The scarcity of vines and epiphytes is notable, although they can be found in microhabitats with favorable conditions and along riverbeds. Some epiphytes, such as bromeliads, can be found in abundance on trees growing on some suitable slopes. Another typical feature of these forests is the presence of numerous lichens attached to the branches of trees and shrubs which acquire the colors of these lichens. In general, most woody plants have relatively small diameters, except in areas with greater soil moisture, such as in the vicinity of permanent or semipermanent bodies of water. As a consequence, Central American dry forests present two characteristic physiognomies: The seasonally deciduous forest associated with the hills, which is the most widespread type, and the riparian forest, smaller in extent, with larger trees that do not shed their leaves during the dry season. Frequently, riparian sectors of the dry forests include species of trees that are typically present in evergreen forests, including several large figs. In addition to these two major variants of the dry forest, in rocky outcrops or as they extend into even drier climates or toward their northernmost distribution in Mexico, average tree height decreases and there is an increased presence of spiny plants, mainly of the legume and cactus families. Another major distinctive feature of the trees in these forests is their bark, which in many species is smooth and shiny and frequently exfoliating (i.e., it peels off), as is the case of *Jatropha* or *Bursera*. In addition, several species present prominent thorns, as is the case of several members of the Bombacaceae family (*Ceiba* and *Bombacopsis*). Another distinctive aspect of the Central American dry forests is the abundance of species with small leaves, frequently of less than 5 cm in length, and many of the species present divided leaves composed of leaflets, although plants with entire or undivided leaves are also common. Several members of the legume family are representative of species with divided leaves. In addition, in marked contrast to the rain forest, dry forest plants typically present showy flowers, and many of the species produce their flowers in remarkable synchrony, particularly during the dry season when leaves are absent. These reproductive and vegetative rhythms are the norm among the plants of these forests, but notable exceptions occur, the most spectacular of which are those related to the patterns of leafing out. Although some species keep their leaves during the dry season, others such as *Jacquinia nervosa* do not merely maintain their foliage during the harsh season but also lose it during the rainy season. The reasons for this leafing pattern are not fully understood by ecologists. Likewise, several species do not flower in mass or otherwise during the season of drought but rather at other times, a remarkable example of which is the *Ipomoea* tree.

The presence of endemic taxa in tropical dry forests is particularly high in these forests. Endemic taxa in Central American forests include species and genera of a marked affinity with the dry lands of northern Mexico. In addition, at least one endemic life form is known to the tropical dry forests of Mexico – the *Ipomoea* tree referred to previously. This remarkable species is a member of the morning glory family, but one that has evolved as an arborescent, not a vine, life form.

In stark contrast with the rain forest, dry forests have been studied to a much poorer extent throughout Central America. Nevertheless, the only available reviews of this type of forest (Bullock *et al.*, 1995; Dirzo *et al.*, 2011) suggest that at least some aspects of the ecological complexity of these ecosystems are comparable to those of their rainy counterparts. Many complex plant–herbivore interactions, for example, are the same as those known for rain forests. Some of them, such as the famous mutualistic interaction between the bull's thorn *Acacia* shrubs/trees and their defending *Pseudomyrmex* ants, can be even more prominent in dry forests. In this remarkable interaction, the plants provide a home (the bull's thorns) and food (sugar produced in extrafloral nectaries and specialized, lipid-rich structures) for the ants. These, in turn, are very aggressive and attack potential herbivores of the plant and even remove the vegetation in the immediate vicinity of the plant, thus avoiding potential competition.

Tropical Cloud Forests

Although with considerable variation depending on the locality, elevation, and aspect, these forests owe their physiognomy, in general, to their enormous trees. Although not very rich in species, this group of trees is very abundant. Some of these may exceed 50 m, with trunks of up to 2 m in diameter at breast height, such as is the case of some oaks and the spectacular *Talauma mexicana* in the cloud forests of

Chiapas. In the mature forest areas the cloud forest is clearly dominated by woody plants. Among these, the lianas, although not completely absent, are considerably scarce. A few of the trees lose their leaves during the driest time of the year, but because a high proportion of the trees do not lose their foliage or, if they lose it, they replace it very quickly, the forest as a whole can be regarded as an evergreen ecosystem. Although the cloud forests located in the southern parts of Central America are dominated by species of tropical affinity, toward their northern distribution, particularly in Chiapas, elements of boreal affinity become prevalent and intermingle with the tropical elements. This is the case for the oaks (e.g., *Quercus candicans* or *Quercus skinnerii*) and especially for the sweet gum tree *L. styraciflua*. This noticeable tree extends in the Central American cloud forests from Honduras to Chiapas. Although the dominant trees of the more southern forests, such as in Costa Rica, are taller, and a stratification is more clearly defined than in their more northern counterparts, the rest of the physiognomic characteristics are very similar. Although buttresses are present, they are not as developed as in the tropical rain forest trees. The understory and lower strata of the cloud forest are notable for the great profusion of palms and ferns, the most conspicuous of which are the spectacular tree ferns. In some areas these can reach up to 20 m. The ferns, including smaller species, tree ferns, and their immature forms, achieve a degree of abundance in the cloud forests that is hardly comparable to that of any other Central American ecosystem. Understory trees can be 8–15 m tall and frequently have leaning, crooked trunks and relatively long crowns. The shrub layer is fairly dense and can reach 2 or 3 m. The ground layer is considerably covered by ferns, *Selaginella*, and several broad-leaved herbs. Bluish coloration in several plants of these lower strata is quite common. Another distinctive feature of these forests is the great profusion of epiphytes and semi-epiphytic plants. Most of the trees are covered by dense associations of ferns, moss, orchids, bromeliads, peperomias, herbaceous vines, and other vascular plants such as the gesneriads. A spectacular fauna of characteristic birds, including quetzals, tanagers, and horned guans, accompanies such variety of plants.

Throughout Central America, this type of forest is of a very restricted extension and patchy distribution. As a result, and because of the relatively poor knowledge of its ecology and also because of its crucial role in retaining water, this Central American ecosystem requires attention and formal protection before it is converted into plantations and grasslands.

Mangrove Forest

Although geographically widespread, mangrove forests are locally restricted to coastal areas, called estuaries, where fresh water meets marine water and salinity is high. This demands the presence of plant species that are tolerant of high levels of salinity: the mangrove trees, which are the dominant species of these ecosystems. In Central America mangroves grow on the Caribbean, Pacific and Atlantic coasts, with the greatest coverage in Panama (174,435 ha), Honduras (78,668 ha), and Belize (78,511 ha) (FAO, 2007). Only a few tree species occupy these forests, typically three, with the richest localities reaching up to nine species. The overall physiognomy of these forests varies locally ranging from 40 m tall forests in areas of high precipitation such as Panama, Belize and several localities of Mexico, to 1.5 m in Nicoya, Costa Rica, where precipitation is low. Mangroves can be considered as a keystone ecosystem given the significant provision of environmental services they provide. Such services include their role as nurseries and refuge for marine and terrestrial wildlife and, most noticeably, their role in protecting the coastline against the effects of tides, storms and tsunamis.

Temperate Forests

Temperate-like forests of Central America include two major types (oak forest and pine forest), but in some parts, in which the dominant species overlap, a mixed oak–pine forest can be distinguished. Oak forests range from the tropical montane and premontane forest dominated by *Quercus* spp. and Lauraceae in Costa Rica, to the clearly oak-dominated forests (encinar) of Chiapas, Mexico. The Costa Rican oak forests are evergreen communities of low to intermediate height, with trees 20–25 m tall and unbuttressed trunks and rough bark. The understory is open, with trees 4–15 m tall and with slender trunks and highly ramified crowns. The branches and trunks of trees are heavily covered with herbaceous epiphytes and moss. The oak forests are best represented in Chiapas, where they form distinct associations ranging from woodlands intermingled with tropical rain forest (with *Quercus oocarpa*) or cloud forest (*Q. candicans* and *Q. skinneri*), to dense monodominated oak forests at elevations of approximately 3000 m, with trees 35 m in height (*Quercus acatenanguensis*). The most widespread oak forests of Chiapas are those of intermediate elevations (700–2500 masl). In the drier areas near the central depression of the state, these are low forests with grasses in the understory and they are sometimes adjacent to seasonally dry forests and savannas. In contrast, in areas of higher elevation and where winds are still loaded with water vapor, they are similar to cloud forests: dense, and loaded with epiphytes, particularly bromeliads and orchids. The most impressive forests of this type are those near the spectacular Montebello Lagoons. Other associations include the low oak forest of 2–4 m in height, which resembles a chaparral, near the city of Comitán.

The pine forests have their greatest distribution in the territory of Chiapas and extend, more sparsely and with fewer associations and diversity, toward the south to Honduras. With the exception of some fragmentary information for the pine forests of Chiapas, these Central American ecosystems are very poorly known scientifically. In Chiapas, these forests occupy most of the surface of temperate and cold areas, to which pines are very well adapted. This is due to the reduced surface of their leaves (needles), which are also well protected on their surface, allowing them to resist long periods of drought and low temperatures. In addition, the resin present in these plants' tissue provides resistance to several pests and the thick bark of the trunk and branches help to reduce the effects of fire. In general, these forests are dominated by one or a few tree species. They extend from intermediate elevations up to the timberline, and where rainfall does not exceed

1200 mm y^{-1}. However, even in areas of more precipitation, pine forests can be occupied by species with thin and flexible needles (*Pinus strobus* and *Pinus ayacahuite*). The most widespread pine forests, at elevations of 750–3000 m, are dominated by *Pinus oocarpa*, whereas *Pinus hartwegii* and *Pinus rudis* dominate the more restricted pine forests of the coldest regions (at elevations of 2900–4000 m).

In most cases, the understory is dominated by dense grassland. A significant, although seldom considered, aspect of the biodiversity of both oak and pine forests is the presence of a rich flora of shrubs and herbs, many of northern affinity but including several species endemic to these ecosystems.

This variety of Central American oak and pine forests, of peculiar and restricted distribution, is very poorly studied ecologically and from a biodiversity point of view, but these forests are disappearing at alarming rates, and most of what remains is heavily fragmented. It is likely that some significant aspects of their ecology and biodiversity have been very strongly disrupted or lost.

Other High-Elevation Ecosystems

Two distinct ecosystems are present at the highest elevations in Central America: the páramo and the high-elevation grassland or zacatonal. The latter is an association of tall and dense grasses that extend in the clearings of pine and oak forests of the highest elevations. This indicates that their current extension is most likely secondary; that is, derived from human perturbation due to the considerable resistance of the grasses to the fires that sweep these regions. The height of these systems is 0.5–2 m and the dominant grasses include *Epicampes macrura*, *Festuca amplissima*, and *Stipa ichu*. The Central American páramo is the northernmost occurrence of the Andean páramo. In Costa Rica it is dominated by shrubs where drainage is adequate and by bogs where drainage is poor. Another portion of páramo occurs in Chiapas, at elevations of 4000 m or higher (above the timberline of the Tacaná Volcano), where it is physiognomically similar to the zacatonal but the dominant grasses (*Calamagrostis* and *Festuca*) are shorter, sometimes forming a prairie. Frequently, the distance among grasses is considerable and there is much open space for the establishment of other plants, many of which barely rise above the level of the ground. This group of plants includes several members of the rose (*Alchemilla* and *Potentilla*), daisy (*Senecio* and *Gnaphalium*), and cabbage (*Draba*) families. The small size and prostrate habit of the plants of this ecosystem are advantageous. They are able to take advantage of the heat in the layer of air close to the ground during daylight since the higher layers, even during daylight, are extremely cold and incompatible with their functioning and survival.

Central American Ecosystems in Light of Global Environmental Change

Several aspects of global environmental change are predicted to have a profound effect on the biodiversity of the planet. Among these, climatic change has received the greatest attention, but recent information indicates that, in addition, the current patterns of land use (leading to deforestation and habitat fragmentation) may significantly affect biodiversity. Although information is limited, particularly for this region, it is of interest to review the available models and the information regarding the situation of Central American ecosystems in the light of climatic change and land-use patterns.

Climatic Change

General circulation models have been developed to forecast the changes in temperature and rainfall and their geographic variation under a given set of assumptions, mostly related to CO_2 concentrations. These models suggest that the greatest changes of climate are to be expected at higher latitudes. Consequently, given the relatively low latitude of Central America, it could be expected that the natural ecosystems of this region would be affected to a minor degree. Indeed, the maps resulting from these models depict Central America as a region of relatively minor climatic change. Although this is an optimistic result, it is important to bear in mind that such maps and forecasts are, by necessity, of very gross resolution. Unfortunately, no detailed analyses are available to assess the situation at a finer geographical scale. The IPCC report for 2007 (Christensen *et al.*, 2007) suggests that for the Central America region (including Mexico), annual precipitation is likely to decrease in most of the region. The multimodel data set models suggest a general decrease in precipitation over most of Central America, and the median annual change by the end of the twenty-first century is projected to be −9% under the A1B scenario (temperature rise of 2.8 °C, range 1.7–4.4 °C, and sea level rise likely range 21–48 cm). Likewise the report suggests that the projected precipitation decrease will be accompanied by more frequent dry extremes in all seasons.

A local study on the expected changes in vegetation coverage resulting from climatic change in the forests of Mexico (Villers and Trejo, 1998) provides some useful insights, given that the study included the northern part of Central America. In addition, the study considered the major Central American ecosystems described previously. In synthesis, the results show that, in comparison to their current potential distribution, tropical forests of the warm lowlands would be the least affected: Seasonally dry forests are expected to remain about the same, whereas rain forests are even predicted to increase. In contrast, temperate forests appear to be affected the most: Pine forests are predicted to disappear, whereas cloud forests would experience reductions of 45–75%. In summary, whereas global models forecast little climate change impact on the ecosystems of the region, more detailed models suggest that this is likely the case for the tropical ecosystems of the warm/lowland climates but the temperate ecosystems are expected to be significantly impacted.

Land Use Change

The best available descriptive index of the patterns of land use is the rate of deforestation. Deforestation rates are available from the FAO annual reports, which include data for regions and the globe at large (**Figure 2a**), as well as countries,

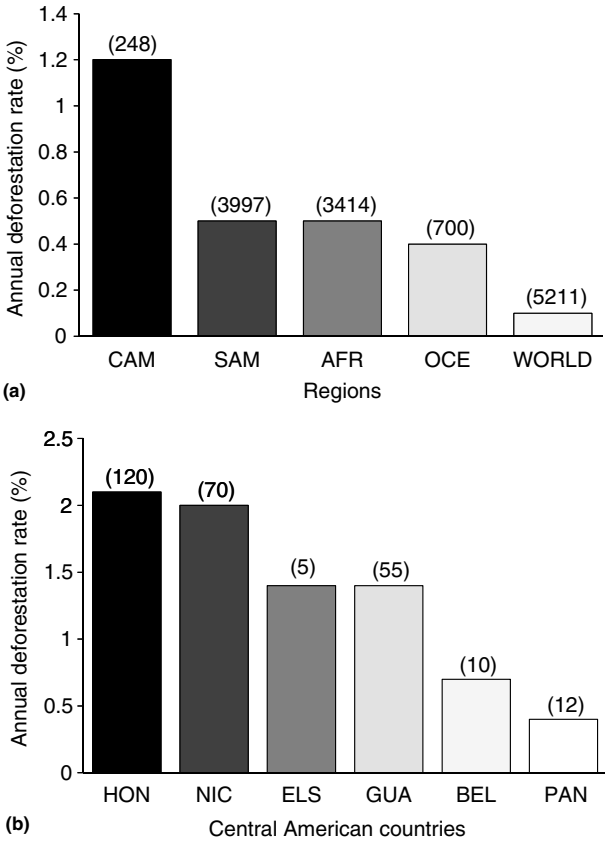


Figure 2 The estimated rate of deforestation (% forest area lost per year) in Central America (CAM) in the context of other regions of the world, including South America (SAM), Africa (AFR), Oceania (OCE) and the overall estimate for the world (a), and annual deforestation rate in six Central American countries (b). Data from Food and Agriculture Organization (FAO) (2011) *State of the World's Forests*. Rome: Food and Agriculture Organization of the United Nations. Codes for the countries are given in **Table 1**. Numbers in parentheses above each bar indicate absolute deforestation rates (ha × 1000).

including Central American ones (**Figure 2b**) (FAO, 2011). In a global and regional context (**Figure 2a**) the statistics indicate that although in absolute terms deforestation in the Central American region, for the period 2000–2010 was low (248,000 ha y^{-1}), yet its proportional rate of deforestation was 2.4 times higher than that of South America and Africa, and 12 times higher than the global rate. Within the region (**Figure 2b**), there is considerable variation among countries, with the highest rates in Honduras and Nicaragua, and the lowest in Belize and Panama, and with El Salvador and Guatemala in the middle. The situation in Costa Rica is particularly revealing, given that it is a country that has made a significant effort to conserve its natural ecosystems. While studies from the 1980s to early 1990s indicate that most of the country's ecosystems were heavily deforested, with percentages of area of habitat lost ranging from 32 (montane rain forest) to more than 90 (dry deciduous forest, lowland moist forest, premontane moist forest, and lower montane moist forest), national deforestation rate decreased to 0.8% in the period 1990–2000 and in the period 2000–2010 it actually became a positive change, with 23,000 ha of forest regrowth.

In the case of Mexico, studies on deforestation in the tropical rain forest of Lacandonia, Chiapas (Mendoza and Dirzo, 1999) show a considerable degree of variation in deforestation rates among localities, with a range of 0–8% and an average of 1.6% per year. Even if these values seem low, the estimated absolute deforestation rate for this area is 6286 ha y^{-1}.

Finally, it is worth considering the deforestation trajectories of mangrove forests in Central America, given their special ecological and ecosystem services importance (described above). While these forests have been dramatically impacted throughout Central America, recent proportional deforestation rates are relatively low, ranging from zero (Guatemala, Nicaragua), to 0.1–0.5 (Belize, El Salvador, Costa Rica, and Panama), to an extreme 3.1% y^{-1} (Honduras). With the exception of Honduras, the low deforestation rates in the rest of the countries are indicative of the fact that some areas are being protected from anthropogenic impact, and other areas have been impacted to the extent that only the most inaccessible areas remain susceptible to being deforested.

In synthesis, the information presented in this section indicates that even if climatic change will have a moderate impact on the ecosystems of Central America, forests of all types seem to be seriously threatened by the current patterns of land use. Moreover, the available information suggests that such patterns of land use exacerbate the negative effects of climatic change. Finally, the current patterns of land use and the expected effects of climatic change on the temperate forests might have a significant effect on the functioning of the ecosystems and overall structure and composition of the biodiversity of Central America. The relevant information to assess the consequences of the loss or degradation of the Central American ecosystems is very limited and is beyond the scope of this article. Nevertheless, our knowledge of their biodiversity and the threats they are currently experiencing indicates that a significant effort toward their conservation and wise management is a major agenda for the scientific community and society at large.

See also: Biodiversity-Rich Countries. Ecosystems of South America. North America, Patterns of Biodiversity in. Rainforest Ecosystems, Animal Diversity. Rainforest Ecosystems, Plant Diversity

References

Bullock SH, Mooney HA, and Medina E (eds.) (1995) *Seasonally Dry Tropical Forests*. Cambridge, UK: Cambridge University Press.

Christensen JH, Hewitson B, Busuioc A, *et al.* (2007) Regional climate projections. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis*, pp. 847–940. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge*, UK/New York, NY, USA: Cambridge University Press.

CONABIO (2008) *Capital Natural de México: Conocimiento Actual de la Biodiversidad*. México D.F., México: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

DeVries PJ (1983) Checklist of Costa Rican butterflies. In: Janzen DH (ed.) *Costa Rican Natural History*, pp. 659–678. Chicago: University of Chicago Press.

Dirzo R, Young H, Mooney HA, and Ceballos G (eds.) (2011) *Seasonally Dry Tropical Forest. Ecology and Conservation*. Washington, USA: Island Press.

Food and Agriculture Organization (FAO) (2007) *The World's Mangroves, 1980–2005: A Thematic Study in the Framework of the Global Forest Resources Assessment 2005*. Rome: Food and Agriculture Organization of the United Nations.

Food and Agriculture Organization (FAO) (2011) *State of the World's Forests*. Rome: Food and Agriculture Organization of the United Nations.

Hanson PE and Gauld ID (eds.) (1995) *The Hymenoptera of Costa Rica*. Oxford, UK/New York, NY, USA: Oxford University Press.

Holdridge LR (1967) *Life Zone Ecology*. San Jose, Costa Rica: Tropical Science Center.

McDade LA, Bawa KS, Hespenheide HA, and Hartshorn GS (1994) *La Selva: Ecology and Natural History of a Neotropical Rain Forest*. Chicago: University of Chicago Press.

Mendoza E and Dirzo R (1999) Deforestation in Lacandonia (southeast Mexico): Evidence for the declaration of the northernmost tropical hot-spot. *Biodiversity and Conservation* 8: 1621–1641.

Miranda F and Hernández-X E (1963) Los tipos de vegetación de México y su clasificación. *Boletín de la Sociedad Botánica de México* 28: 29–179.

Portillo H (2007) *Recopilación de la información sobre la biodiversidad de Honduras*. Informe Final de consultoría. Tegucigalpa, Honduras: INBio, DiBio.

Rich PV and Rich TH (1983) The Central American dispersal route: Biotic history and paleogeography. In: Janzen DH (ed.) *Costa Rican Natural History*, pp. 12–34. Chicago: University of Chicago Press.

Villers L and Trejo I (1998) Assessment of the vulnerability of forest ecosystems to climate change in Mexico. *Climate Research* 9: 87–93.

Cladistics

Ian J Kitching, Peter L Forey, and David M Williams, The Natural History Museum, London, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 677–692, © 2001, Elsevier Inc.

Glossary

Apomorphy A derived character or character state; if two or more taxa share apomorphies, these are referred to as synapomorphies.
Clade Group of taxa diagnosed as monophyletic by the discovery of homologies (or synapomorphies).
Cladogram Branching diagram specifying hierarchical relationships among taxa.
Cladogram support Tests that permit some evaluation of how well data fit a cladogram.
Consensus cladogram (tree) Branching diagram that summarizes the common branching patterns from two or more cladograms.
Homology Two characters passing the similarity, conjunction, and congruence tests are termed homologous; in cladistics, homology is synonymous with synapomorphy.
Homoplasy A character or character state acquired by parallel or convergent evolution that bears resemblance to a character in a different group.
Monophyly Relationship between taxa united by a synapomorphy.
Optimization Procedure for reconstructing the most parsimonious sequence of character change on a cladogram.
Parsimony General scientific principle that given alternative explanations or hypotheses for a set of observations or data, the most corroborated is that requiring the fewest ad hoc (ancillary or additional) hypotheses.
Plesiomorphy An apomorphic character or character state that specifies a more inclusive group than that under consideration.

Relationships

The basic concept of cladistics is that genealogical connections among organisms are expressed in relative terms. Consider three taxa, A, B, C, whose genealogical relationships are as given in **Figure 1a**. Taxa B and C are more closely related to each other than either is to taxon A because they share a common ancestor, x (which lived at time t_1), that is not shared with taxon A. Similarly, taxon A is more closely related to the group (B + C) because A, B, and C together share a unique common ancestor, y, that lived at an earlier time (t_0). In a real example (**Figure 2**), the human and turkey are considered to share a unique common ancestor (w) that lived at t_3. Similarly, the frog, turkey, and human are more closely related to each other than to either the perch or dogfish because these three taxa uniquely share a common ancestor x that lived at time t_2. The human and turkey are called sister-groups. Likewise, in this example, the frog is the sister-group of (human + turkey). The aim of cladistic analysis is to infer the sister-group hierarchy of life-forms by analysis of characters

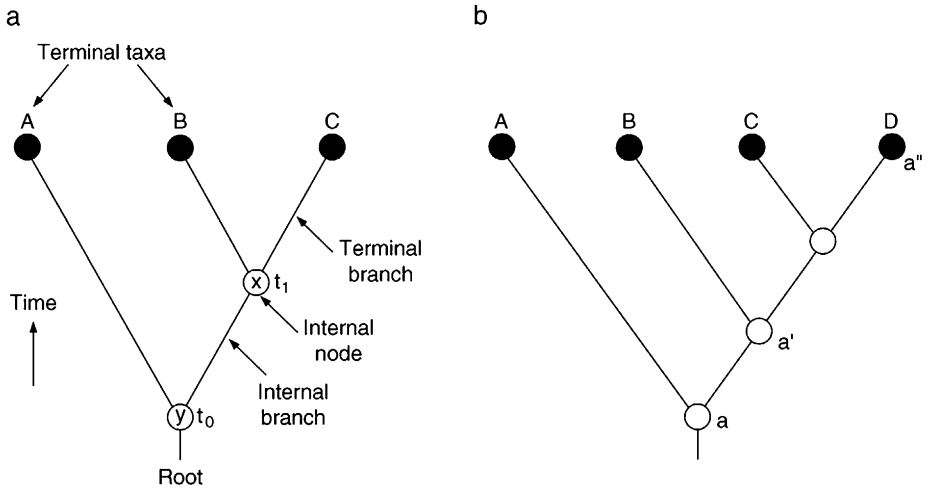


Figure 1 (a) The phylogenetic tree and the cladistic meaning of relationship. Taxa B and C are considered more closely related to each other than either is to taxon A because they share a unique common ancestor (x) that is not shared by taxon A. (b) Plesiomorphy and apomorphy are relative terms. On this phylogenetic tree, a′ is apomorphic with respect to a but plesiomorphic with respect to a″.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00022-8

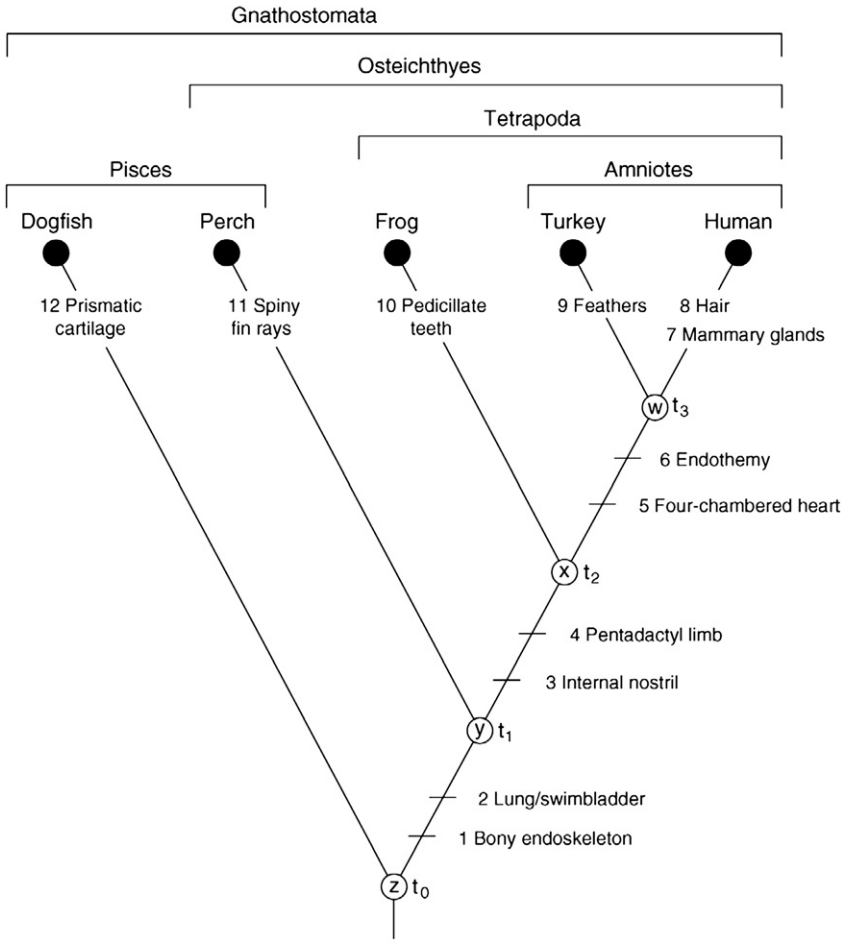


Figure 2 A phylogenetic tree for five taxa of vertebrates. Three monophyletic groups are established using characters 1–6, while autapomorphic characters 7–12 diagnose the terminal taxa. The group Pisces is paraphyletic because one of its included members (the perch) is cladistically more closely related to Tetrapoda.

and to express the results as branching diagrams. These diagrams are called "cladograms" because they identify a hierarchical arrangement of taxa based on homologies termed "clades."

Types of Characters

Two types of characters are recognized based on where they occur in the inferred phylogenetic history of a group (**Figure 1b**). The character that occurs in the ancestor is termed "plesiomorphic" (near to the ancestral morphology) and the derived character is "apomorphic" (away from the ancestral morphology). Apomorphic and plesiomorphic are relative terms, that is, relative to a particular systematic problem. In **Figure 1b**, character a′ is apomorphic with respect to character a, but plesiomorphic with respect to character a″.

Cladistic analysis proceeds by identifying shared apomorphic characters or "synapomorphies." In **Figure 2**, a four-chambered heart and endothermy are synapomorphies that suggest the human and the turkey share a unique common ancestor w. The cladogram implies that these two characters arose in ancestor w and were then inherited by both the human and the turkey. Synapomorphies may therefore be considered as evolutionary homologies. In contrast, the shared possession of internal nostrils and pentadactyl limbs by the human and turkey does not imply that they share a unique common ancestor because these attributes are also found in the frog. These shared primitive characters (or "symplesiomorphies") are inherited from an ancestor more remote than the most recent common ancestor of the human and the turkey. They are thus irrelevant to the hypothesis of a relationship between the human and the turkey. However, with respect to the more inclusive three-taxon problem comprising the frog, turkey, and human, internal nostrils and pentadactyl limbs are relevant. At this level, they are synapomorphies suggesting that these three taxa form a group with a common ancestry at x. Apomorphies occurring in only a single terminal taxon are termed "autapomorphies." In **Figure 2**, these are prismatic cartilage (dogfish), spiny fin rays (perch), pedicillate teeth (frog), feathers (turkey), and hair and mammary glands (human). However, if a terminal taxon is itself a group, then its autapomorphies are also synapomorphies of its component taxa.

Thus, as with the cladistic meaning of relationship, characters are also relative, depending on the systematic problem under consideration. Furthermore, it should be stressed that characters are observations of the features occurring in organisms and the explication of their hierarchical distribution need not imply a particular theory of evolution.

Parsimony

Relationships among three taxa (as in **Figure 1a**) can be resolved in three ways—A (B C), B (A C), and C (A B)—whereas for four taxa (as in **Figure 1b**) there are fifteen possible fully resolved cladograms. In cladistic analysis, parsimony is the universal criterion for selecting among alternative hypotheses of character distribution. Characters are fitted onto alternative topologies and the cladogram that accounts for the greatest number of characters in the simplest way is chosen as the best hypothesis of relationships.

Suppose six characters are distributed among four taxa as shown in the taxon/character matrix in **Figure 3**. Taxon A has none of the characters but the other three taxa each have a different complement. Characters 2 and 4 are autapomorphies because they are each present in only one of the taxa. They are uninformative for grouping taxa (they serve only to diagnose the terminal taxa). Characters 1, 3, 5, and 6 are potentially useful because they occur in more than one taxon. Given the three taxa that have potentially informative information, there are three ways in which these taxa can be arranged dichotomously (**Figure 3b–d**). If the characters are now placed onto each possible cladogram, according to the groups they specify, then three different results are obtained (**Figure 3e–g**). In **Figure 3e**, all characters except one appear only once. In this solution, character 6 must be assumed to appear twice, once in taxon B and once in taxon C, which are not sister-groups. In this example, the behavior of character 6 is homoplastic, that is, it occurs more than once on the cladogram and is said to be a homoplasy. In contrast, in the other two topologies (**Figures 3f** and **3g**), we must assume that two or more characters appear more than once. Hence, the cladogram in **Figure 3e** accounts for the distribution of the characters in the most economical way and is thus the preferred solution.

The distribution of characters can also be regarded as the number of steps on a cladogram, which, in **Figure 3**, is the number of instances where a character is gained. In **Figure 3e**, this is seven, while the other cladograms (**Figures 3f** and **3g**) are more costly, requiring nine and eight steps, respectively. The concept of steps is actually a little more subtle than the sum of character gains because a character may appear at one point on a cladogram and then disappear at another point. For example, another explanation for the distribution of character 6 in **Figure 3e** is to assume that it is gained by the group (B + C + D) and then lost in taxon D. Each change, whether gain or loss, is considered a step. In this example,

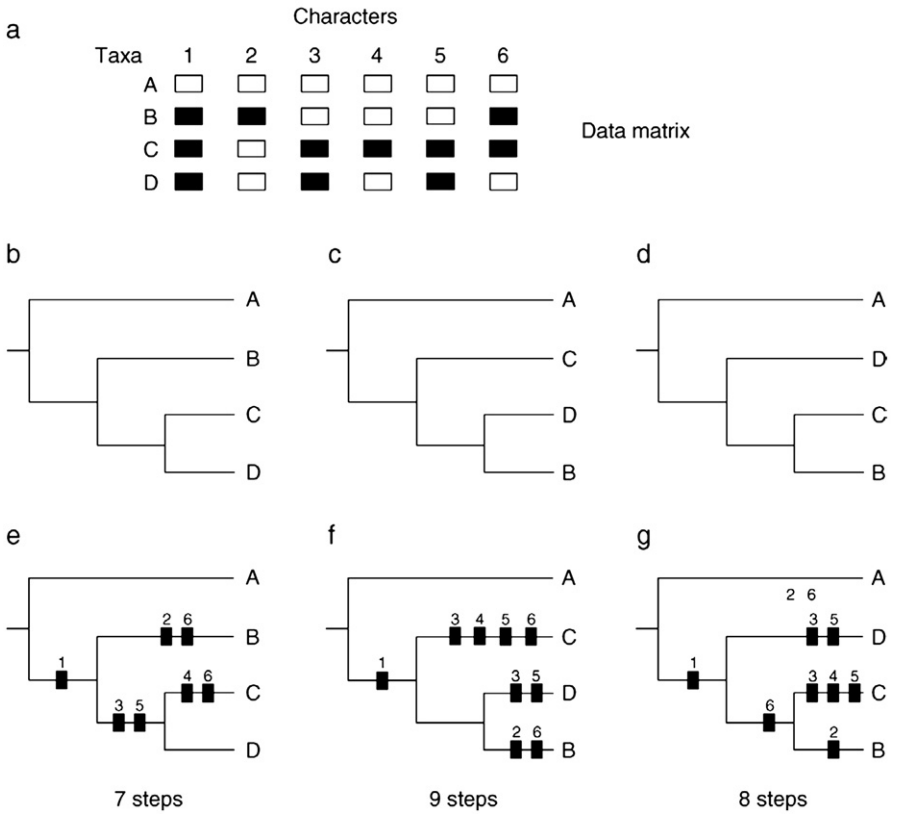


Figure 3 Parsimony. (a) A data matrix of six characters distributed among four taxa, A–D. Plesiomorphic states are shown by open boxes and apomorphic states by black boxes. Taxon A is totally plesiomorphic. (b–d) The three possible resolutions of taxa B–D relative to A. Placing the characters on these three topologies results in (e) being selected as the optimal cladogram while (f) and (g) are suboptimal.

both accounts of character change demand two steps and therefore both hypotheses of character change are equally parsimonious. The sum of the number of steps on a cladogram is termed the length of the cladogram, irrespective of whether the changes are gains or losses. The most parsimonious solution is also known as the optimal cladogram and the other cladograms (i.e., those requiring more than the minimum number of steps to explain the character distributions) as suboptimal.

It is possible for a given set of characters to yield two or more equally most parsimonious cladograms. Then, we may prefer to accept one of the solutions based on other criteria, such as a closer agreement with the stratigraphic record or by differential weighting of one type of character change relative to another. Alternatively, we may simply accept that the conflict in the data is such that we cannot derive a unique most parsimonious solution. For certain purposes, we may choose to combine those components common to the different solutions to form a consensus cladogram.

Groups

Cladistics recognizes only monophyletic groups of organisms, which are those based on synapomorphies. Monophyletic groups are the only groups that can be circumscribed by objective boundaries. In evolutionary terms, monophyletic groups comprise the most recent common ancestor and all of its descendants. In **Figure 2**, Amniota, Tetrapoda, Osteichthyes, and Gnathostomata are all monophyletic. Two other types of "groups" are sometimes referred to but these are not groups in the same sense as monophyletic groups. Paraphyletic "groups" are based on symplesiomorphy; in evolutionary terms, their members are linked by common ancestry but one or more of the descendants of the most recent common ancestor are excluded. In **Figure 2**, Pisces (fishes) is a paraphyletic assemblage. Many taxa traditionally regarded as ancestral, such as fishes, reptiles, and green algae, are paraphyletic. Polyphyletic "groups" are based on homoplasy, that is, characters that are considered convergently derived and that cannot be inferred to have been present in the most recent common ancestor of the included taxa. In **Figure 2**, an assemblage comprising the dogfish and the turkey (perhaps based on the observation that both lay eggs surrounded by a shell, although no one would claim such a homology) would be a polyphyletic group.

Cladograms and Phylogenetic Trees

A cladogram is a diagram that summarizes a pattern of character distribution. Usually, a cladogram is drawn as a branching diagram (e.g., **Figure 1**). The nodes denote a hierarchy of synapomorphies but there is no necessary implication of ancestry and descent. Cladograms may also be written in parenthetical notation or illustrated as a Venn diagram (**Figure 4a**), which conveys the same grouping information as a branching diagram. In contrast, phylogenetic trees include a time axis and embody concepts of ancestry and descent with modification. In phylogenetic trees, the nodes denote ancestors (known or hypothetical) and the branches imply character change. Several phylogenetic trees may be compatible with the pattern of character distribution implied by a cladogram (**Figure 4b**). Some of these trees allow the possibility that one or more taxa are ancestral to others. Only the phylogenetic tree that assumes all nodes represent hypothetical ancestors has the same topology as the cladogram. Thus, cladograms are more general than phylogenetic trees, which are precise statements about ancestry and descent.

Characters and Coding

Opinions differ over the nature and discovery of taxonomic characters. One view holds that characters are properties of organisms that provide quantifiable variation. Alternatively, characters may be viewed as theories concerning two (or more) attributes, which may look different but are nevertheless

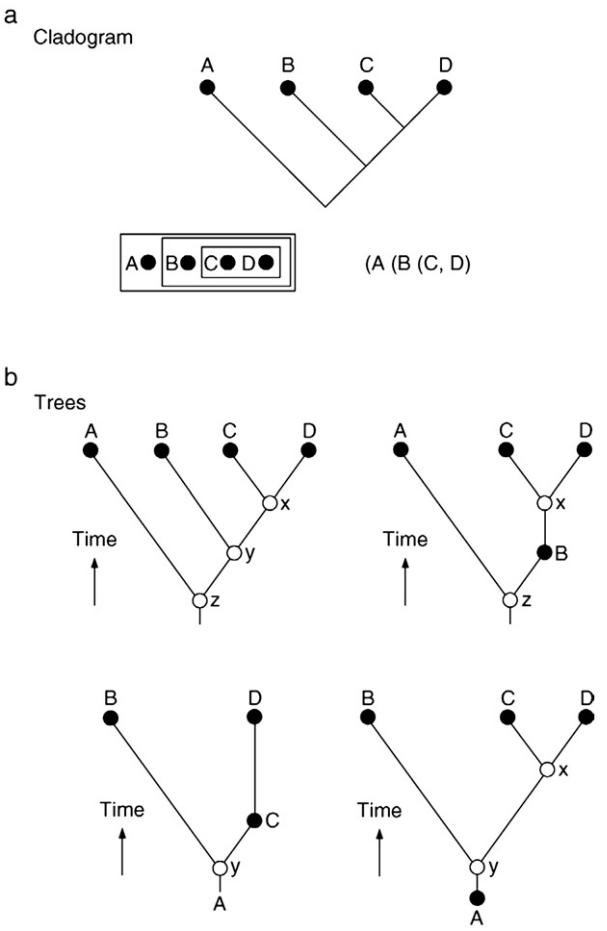


Figure 4 Cladograms and trees. (a) A cladogram simply shows the hierarchical distribution of characters. It has no time axis and may be drawn as a branching diagram, a Venn diagram, or in parenthetical notation. Given this cladogram, several phylogenetic trees may be inferred from the same data. Some of these are shown in (b). All except that on the top left invoke the concept of one of the taxa being ancestral. The tree at the top left assumes only hypothetical ancestors and is the only one that is synonymous with the cladogram.

considered the same. This latter view is embraced within a general understanding of homology, such that characters may be understood in the same manner as homology. There is a lack of agreement over what indicates the discovery of a character. However, all definitions of homology suggest that it concerns features that are similar in different taxa sharing a recent common ancestor. Such definitions satisfy as explanations but do not aid discovery.

Homology

Within cladistics, various tests assist in establishing homology. One view of characters is that they are identical in meaning and discovery to homology, and homology may be conceived as a series of three tests that apply to methods of comparison.

The similarity "test" suggests that without evidence to allow direct comparison of one feature with another, there would be no proposition of homology and, consequently, no concept of a character. This "test" is not exact and cannot be taken to imply "identity." For example, comparisons may consider the detailed similarity of any two stamens or the inferred similarity of mammalian stapes with gnathostome hyoid arches.

The conjunction test suggests that two features that co-occur in the same organism cannot be considered homologous. A familiar example, albeit contrived, is an angel with both wings and forearms. The two kinds of limbs in the same individual cannot be considered homologous. Many comparisons fail this test and are often associated with "homomony" or serial homology (e.g., the individual vertebrae of a single vertebral column or the abdominal appendages of arthropods).

The congruence test is considered the most exacting and refers to the support afforded to one homology by others. In other words, homologies are considered to have passed the test if there are other homologies that specify the same taxon. Congruence is actually an analytical procedure and is usually considered in terms of parsimony. However, it also points to another property of homology, that is, homology can never be proven. As data are accumulated, previously supported homologies may be overturned and new theories of homology established in their place.

Character Recognition

It is generally agreed that characters, however conceived, are based on observations. Stated simply, a feature (e.g., stamens, shoulder girdles, wings) is observed in a particular specimen and directly translated into the character. This approach may initially seem useful and would eventually lead to enumeration of all features of the specimen. However, the final list would not consist of "characters" but would be an inventory of "features," each being a descriptive element of the specimen and implying that such descriptions apply to all specimens of the same taxon. Each descriptive element contains a notion of theory.

Suppose the specimen examined is a rat. Initially, it would be straightforward to describe: head, body, limbs, tail, and so on. More detailed examination would reveal, for example, a vertebral column. We identify the vertebral column by drawing on knowledge of previous studies of rat anatomy and are able to confirm its detailed similarity to other vertebral columns. In so doing, we assimilate what is already known of vertebrates: they are animals with a vertebral column. If this process were performed for all features, then it would seem that all attributes of this single rat could uncover its place in the hierarchy of life, at every inclusive level. In this sense, taxonomic characters are very much like homologies: features that (potentially) specify a particular taxon. A vertebral column does not tell us it is a rat; it tells us that it is a vertebrate. In this sense, the vertebral column is a *feature* of any particular rat but only a *character* of vertebrates. This distinction identifies the general task of systematics: to identify the level at which various attributes are homologies (characters) diagnosing taxa.

Kinds of Characters

Characters are often thought of as comprising different types. Some refer to different numbers of a feature, and others refer to differences in structure. For example, variation in stamen structure in angiosperms encompasses both different forms of anthers and filaments and differences in their numbers. This exemplifies the distinction between quantitative and qualitative characters, the former usually being counts ("meristic characters") or measurements ("biometric characters"), the latter relating to structural differences. Quantitative characters may be problematic for cladistics insofar as it can be difficult to render measurements and counts as meaningful homology statements. This is not to say that such characters are not useful, for they can serve to identify particular specimens. However, their use may be limited because they are not always amenable to cladistic analysis.

Characters as Phylogenetic Evidence

Although structural evidence is sought for cladistic purposes, the observed features themselves are not necessarily the characters. For example, some organisms have fins, others have arms, and yet others have wings. Studies of fins, arms, and wings show that they have certain parts in common as well as certain parts that are unique. These common properties might suggest an initial proposition that fins, arms, and wings are all examples of a single character, in this case "paired appendages." However, further details are needed to confirm this hypothesis.

Suppose the "wings" examined were from an insect such as a housefly. It might still seem reasonable to call them "paired appendages," but in this example there is nothing in the detailed anatomy to suggest that housefly wings and mammalian arms (or even bird wings) are in any way "the same." One conclusion is that the term "wings" is ambiguous when describing attributes of organisms but not when describing the function of these attributes (here, flight). Wings may indeed be considered as a part of an organism but flight is their assigned function (usually). A more reasonable conclusion is that "wings" is not a character at all but a functional attribute. Thus, the wing of a bird is better considered as a modified "paired appendage," modified for flight. The problem,

however, is not simply semantic. If the comparison was made between a bird and a bat, then detailed anatomy does suggest that both are indeed "paired appendages" and that both are modified for flight. Our current understanding of vertebrates suggests that birds and bats do not form a monophyletic group. Hence, bird wings and bats wings would be considered two characters rather than one when interpreted on the cladogram (**Figure 5**). Their "sameness" is captured as "forelimbs," their differences as wings of birds and wings of bats.

Consideration of the wings of birds, the forelimbs of tetrapods, and the fins of fishes together identifies a well-defined character ("paired appendages") with various manifestations. These manifestations might suggest particular taxonomic groups. There might be a taxon with "fins," a taxon with "arms," and a taxon with "wings." This, of course, was precisely the situation for many years: fishes (Pisces) have fins, tetrapods have arms, and birds (Aves) have wings (**Table 1**). Derived from these observations is the implication that fins, wings, and arms are in some way connected, other than by being "paired appendages":

$$\text{fins} - \text{arms} - \text{wings}$$

Furthermore, that connection might be viewed in an evolutionary context, such that it represents the transformation of one manifestation (e.g., fins) into another (e.g., arms):

$$\text{fins} \rightarrow \text{arms} \rightarrow \text{wings}$$

In more general terms, features could be represented as modifications of other features, such that they are hierarchically related:

Fins
Standard fins (fins)
Modified fins (arms)
Modified arms (wings)

The features are no longer structured as a series of alternatives, as in the first example (fins–arms–wings), but now specify a nested set of relationships. The proposition is that wings are really kinds of arms, and that arms are really kinds of fins, and fins represent the entire set of animals with "paired appendages." It is possible to interpret all of these taxa (fishes, tetrapods, and birds) as having fins. Hence "fins" is not a character of fishes but rather of gnathostomes (in this example, fishes + mammals + birds). Consequently, both the character "fins" and the taxon Pisces disappear. This view confirms the notion that characters are hypotheses drawn from observations rather than simply the observations themselves. Such hypotheses are identical to those made for general statements of homology. Superficially, the relationship between fins, arms, and wings may be considered identical to "fins → arms → wings," as was implied by Hennig (1966) in his concept of "transformation series." However, it is possible to view characters as more general, specifying particular relationships in terms of a definitive statement connecting to a taxon.

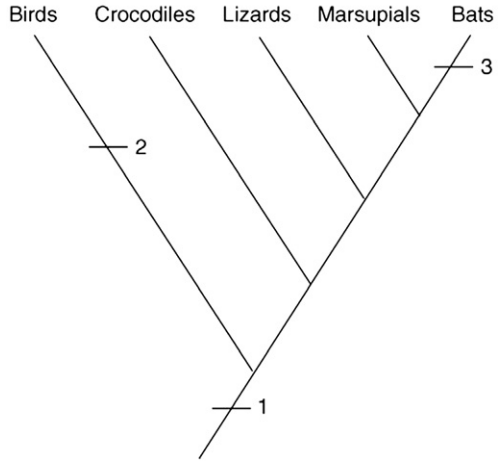


Figure 5 Relationships of vertebrates; 1 ≡ "forelimbs"; 2 ≡ "wings of birds" 3 ≡ "wings of bats."

Table 1

Taxon	*Character*
Pisces	Fins
Tetrapods	Arms
Birds	Wings

Character Coding

One significant outcome of theories relating to characters is how they might be represented numerically for cladistic analysis. For example, one might code each "character" (fins, wings, arms) separately (**Table 2**, characters 2–4). This scheme reflects the "uniqueness" of each attribute but contains no information relevant to recognizing that the three observed forms are connected as "paired appendages." This approach is referred to as "absence/presence binary coding," because a positive value (usually 1) is assigned to the presence of a feature and a negative value (usually 0) is assigned to the absence of the feature. Alternatively, one might represent the same series of observations in a single column to signify their connection (as "paired appendages"), then assign each unique feature a separate value (**Table 2**, character 1). This is "multistate coding" and considers the character to be composed of discrete states that bear some (usually unspecified) relationship to one another. Hence different values appear in the same column and are treated as dependent on the other values. This might not be seen as completely sufficient, as additional information would be needed to specify the exact nature of the connection. For instance, one might wish to specify that the "characters" are connected but that the nature of that connection is unknown. Choices of this nature relate to character optimization (see Cladistic Analysis: Optimization).

To summarize, characters have their origin, but not their identity, in observations. Characters are what lead us to suspect that taxa exist (vertebral column ≡ vertebrates ≡ taxon Vertebrata) and hence are identical to conjectures of homology derived from empirical investigation of specimens. Homology is the relation that specifies taxa and that implies

Table 2

Taxon	*Characters*			
	1	*2*	*3*	*4*
Fins	1	1	0	0
Wings	2	0	1	0
Arms	3	0	0	1

an intimate relationship between characters and taxa. Both are the results of analyses and are discovered by our investigation of features.

Cladistic Analysis

Cladogram Construction

The original method of cladogram construction was proposed by Hennig (1950, 1966) and is thus known as Hennigian argumentation. In this approach, characters are first polarized into plesiomorphic and apomorphic states. The groups thus diagnosed by synapomorphies are then organized manually into a cladogram. However, this procedure can only find the most parsimonious cladograms when the data are free or nearly free of homoplasy (i.e., the fit of data to most parsimonious cladogram is perfect or nearly so). For larger and more complex data sets, computerized algorithmic methods become a necessity.

There are two main computerized approaches to cladogram construction. Exact methods guarantee to find the most parsimonious cladograms. The simplest exact method is "exhaustive search." First, three taxa are chosen and connected to form the only possible unrooted, fully resolved cladogram for these taxa. Then, a fourth taxon is selected and added to each branch to yield the three possible fully resolved, partial, unrooted cladograms for four taxa. A fifth taxon is then selected and added to each of the five branches on these three partial cladograms to yield the fifteen possible fully resolved unrooted topologies for five taxa. This process is continued, following every possible path of taxon addition, until all taxa have been added and all possible fully resolved cladograms have been found. The lengths of these cladograms are then calculated and the shortest is chosen as the optimal solution(s). However, as the number of taxa increases, the number of cladograms to be examined rises exponentially and the time required for exhaustive search soon becomes unreasonable.

One exact method that does not require every possible cladogram to be evaluated is "branch-and-bound" analysis. In this approach, a preliminary cladogram is constructed and its length is set as the upper bound for subsequent searches. A procedure similar to an exhaustive search is then undertaken but at each step the length of the partial cladogram is recorded. Whenever this length exceeds the current upper bound, that partial cladogram is rejected (and so, consequently, are those complete topologies that would be derived from it by adding the remaining taxa). By this means, the number of topologies to be examined is reduced. Once all taxa have been added, the length of the complete cladogram is examined and if it is equal to the upper bound, then that topology is retained as a most parsimonious cladogram. However, should this cladogram be shorter than the current upper bound, then its length is substituted as a new upper bound. This important procedure allows subsequent partial cladograms to be rejected quickly and thus speed up analysis. This process continues until all possible paths have been examined, whence the set of optimal cladograms will have been found.

For large data sets (more than 30 taxa), even branch-and-bound analysis can be too time-consuming. In this case, approximate or "heuristic" methods are used. These approaches examine only a subset of all possible topologies and thus are not guaranteed to find the most parsimonious cladogram(s). However, they are faster than exact methods for large numbers of taxa and thus certainty of finding the optimal cladogram(s) is sacrificed for decreased computational time.

Heuristic analysis comprises two stages. In the initial building phase, a cladogram is constructed using a process of "stepwise addition." The order in which taxa are added is termed the "addition sequence" and there are various ways in which taxa may be added. Once a complete cladogram has been constructed, attempts can be made to improve upon it by performing a series of rearrangements called "branch-swapping." The cladogram is cut into two or more partial cladograms, which are then recombined in order to try to find new, shorter topologies. The efficiency of current branch-swapping algorithms in finding most parsimonious cladograms is very high, but it is always possible that they can become trapped in a local optimum. Thus, one should always be aware with heuristic analyses that shorter topologies than those reported may exist.

Character Polarization and Cladogram Rooting

Manually implemented cladistic methods, such as Hennigian argumentation, require synapomorphies to be identified in advance of cladogram construction. The process through which plesiomorphic and apomorphic characters are distinguished is termed "character polarization." Numerous criteria for polarizing characters have been proposed, but only two are now considered valid.

The first criterion, called "outgroup comparison," was classified by Nelson (1973) as an "indirect" method, because it draws upon evidence from a source (the "outgroup") that is external to the taxa under investigation (the "ingroup"). In its most basic form, polarization using outgroup comparison can be defined as follows: "For a given character with two or more states within a group, the state occurring in related groups is assumed to be the plesiomorphic state" (Watrous and Wheeler, 1981: 5). This definition is adequate when all outgroup taxa share the same state, but it is insufficient if the outgroup taxa are heterogeneous. Maddison *et al.* (1984) further noted that it is inappropriate to estimate the state in the most recent common ancestor of the ingroup (the "ingroup node"). Rather, it is the state at the next most distal node, linking the ingroup to the first outgroup (the "outgroup node"), that should be estimated if the solution is to be

globally optimal, and they described an algorithmic approach to such reconstruction.

In contrast, Nelson (1973) classified the "ontogenetic criterion" as a "direct" method, because its implementation relies on evidence from the ingroup taxa alone. It is defined as follows: "Given an ontogenetic character transformation from a character observed to be more general to a character observed to be less general, the more general character is primitive (plesiomorphic) and the less general character derived (apomorphic)" (Nelson, 1978: 327). For example, the embryos of both sharks and frogs have cartilaginous skeletons. However, this condition persists into the adult shark, but in frogs the cartilage is replaced by bone during ontogeny. In this example, a bony skeleton is observed to be less general (occurring only in the frog) and is thus interpreted as apomorphic. In this context, "more general" is defined as that occurring earlier in ontogeny. As such, the more general character is not simply the more common (although this may often be the case) and the ontogenetic criterion does not equate with commonality. What is important is that the less general character is nested within the observed distribution of the more general character. This requirement is violated by ontogenetic sequences that are secondarily simplified through paedomorphosis or neoteny. Such an ontogeny cannot be interpreted as proceeding from the more general to the less general, and Nelson's criterion will not allow us to distinguish a secondarily reduced ontogeny from the plesiomorphic sequence. Then, we rely on congruence with other characters to make the distinction.

Most recent cladistic studies do not actually include a priori polarization of characters, but undertake "simultaneous, unconstrained analysis": "simultaneous" because both ingroup and outgroup taxa are analyzed together, and "unconstrained" because outgroup taxon relationships are unspecified before analysis. Cladograms are then rooted between the outgroup node and the remaining outgroup taxa, at which point character polarities are established.

Optimization

Optimization is the process of determining the sequence of character state changes on a cladogram in order to test hypotheses of transformation. If the data include characters coded as multistate, then these may be interpreted according to many different optimality criteria, of which the best known are Wagner and Fitch optimization.

Wagner optimization (**Figure 6a–b**) is used for "ordered" or "additive" multistate characters, in which transformations between successive states are considered as incremental. Thus, the changes $0 \leftrightarrow 1$ and $1 \leftrightarrow 2$ each "cost" the same number of steps (usually one), whereas the change $0 \leftrightarrow 2$ is considered to "pass through" state 1 and thus costs two steps. Costs are symmetrical, so that the changes $0 \rightarrow 1$ and $1 \rightarrow 0$ both constitute a single step (termed "free reversibility").

Wagner optimization is implemented in two stages. First, the minimum number of steps for a character is determined. On a pass down the cladogram, from the most distal taxa toward the root, each of the internal nodes is assigned a "state set" (**Figure 6a**), which is defined as the intersection of the two derivative state sets. If the intersection is empty, then the smallest closed set that includes an element of each derivative state set is assigned. For example, the intersection of the state sets of taxon F (2) and taxon G (4) is empty. Thus, the smallest closed set, (2–4), is assigned to the node linking these two taxa. In contrast, the intersection between this state set and that of taxon E is not empty. Both contain state 2, and this value is assigned to the node joining taxa E, F, and G. Unambiguous states are then assigned to internal nodes by a second pass, up the cladogram (**Figure 6b**), to produce a "most parsimonious reconstruction." If the state set of a node contains more than one value, then the node is assigned the value that is closest to that of the node of which it is a derivative. Thus, the node joining taxa F and G is assigned state 2, because this is the value of the next most distal node. The most parsimonious reconstruction has six steps because the character is ordered. As a result, the change from state 2 to state 4 along the branch leading to taxon G counts as two steps.

Fitch optimization (**Figure 6c–d**) is used for "unordered" or "nonadditive" multistate characters in which transformations between any two states are considered equal. Thus, the changes $0 \leftrightarrow 1$, $1 \leftrightarrow 2$, and $0 \leftrightarrow 2$ all cost a single step. As with Wagner optimization, costs under Fitch optimization are freely reversible.

Fitch optimization follows similar procedures to Wagner optimization but with two differences. First, the state set assigned to an internal node is the union of the two derivative state sets (**Figure 6c**). Second, when determining the most parsimonious reconstruction, a node with an ambiguous state set is assigned the value of the next most distal node if that value is an element of the ambiguous state set. Otherwise, an element is selected arbitrarily. This most parsimonious reconstruction (**Figure 6d**) has four steps because the character is unordered. As a result, the change from state 1 to 3 along the branch leading to taxon D and the change from 2 to 4 along the branch leading to taxon G each count as a single step.

These procedures do not necessarily yield a unique most parsimonious reconstruction. For example, it is equally parsimonious to optimize the ordered character using the nodal state reconstructions shown in **Figure 6d** for the unordered character. State 1 is assigned to the node joining taxa C–G, and is followed by a two-step change, $1 \rightarrow 3$, on the branch leading to taxon D and a one-step change, $1 \rightarrow 2$, on the branch joining taxa E–G. The steps on the branches joining taxa B–G and leading to taxon C would be lost, thus maintaining the most parsimonious length of six steps. This type of optimization, in which changes are placed onto the cladogram as far from the root as possible, is called "delayed" or "slow" transformation. In contrast, "accelerated" or "fast" transformation places changes onto the cladogram as close to the root as possible, as in **Figure 6b**. When alternative most parsimonious reconstructions are possible, character optimization may lead to "spurious resolution," in which groups appear to be resolved but have no unambiguous support in the data. Such groups are not strong hypotheses of relationship, and Nixon and Carpenter (1996) suggested that they should be eliminated wherever possible (without violating the minimum length requirement). Those cladograms that remain, which are both of minimum length and have all groups

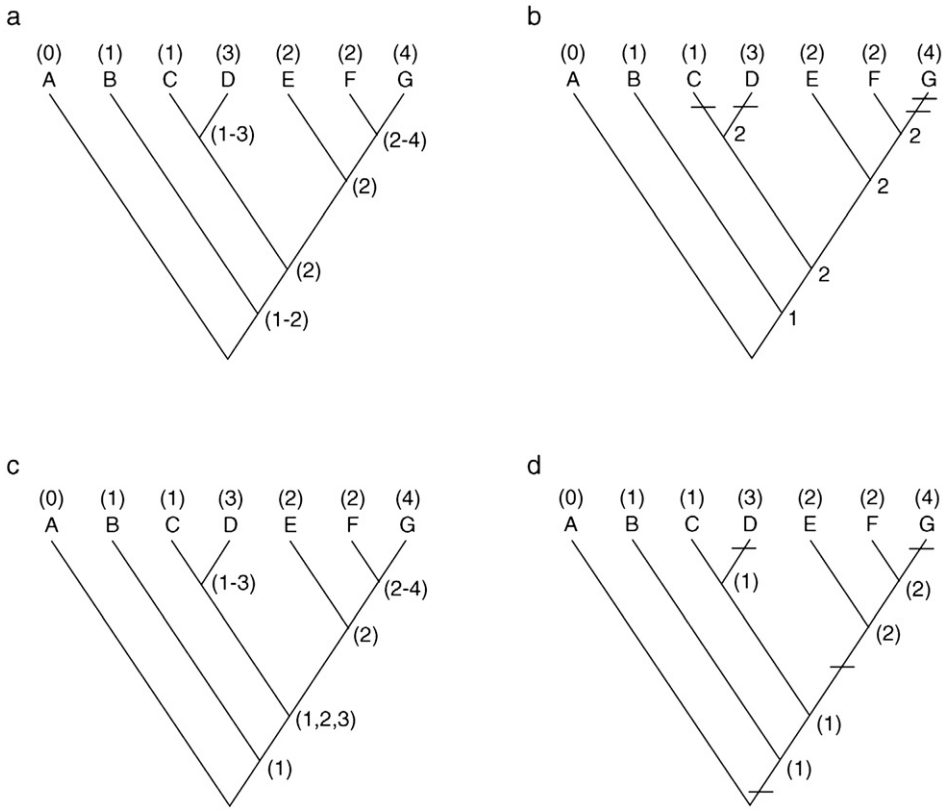


Figure 6 Character optimization, (a) Wagner optimization (ordered multistate characters). State sets are assigned to internal nodes on the first pass from the terminal taxa toward the root. (b) States are assigned to the internal nodes on the second pass from the root to the terminal taxa. (c) Fitch optimization (unordered multistate characters). State sets are assigned to internal nodes on the first pass from the terminal taxa toward the root, (d) States are assigned to the internal nodes on the second pass from the root to the terminal taxa.

supported unambiguously by data, are termed "strictly supported cladograms" and are the preferred topologies.

Cladogram Evaluation

Character Fit

The preferred solution to a cladistic analysis is the most parsimonious cladogram because this represents the simplest explanation of the data with the number of ad hoc hypotheses of homoplasy kept to a minimum (see **Figure 3e**). Consequently, the most basic measure for assessing the fit of data to a cladogram is cladogram length. The most parsimonious cladogram has best fit because it is the shortest; longer cladograms have poorer fit. However, the length of the most parsimonious cladogram is partly dependent on the absolute size of the data set from which it is derived. Larger data sets will necessarily yield longer cladograms than smaller data sets. A binary character will display perfect fit when it is placed on a cladogram with a single step. Homoplasy is manifest as an increase in the number of steps. The amount of homoplasy implied by a character on a cladogram is measured by the "consistency index" (ci), which is the ratio of the minimum number of steps required by the character ($m = 1$ for a binary character) to the observed number (s). In **Figure 3e**, character 5 occurs only once and hence its ci = 1 ($m/s = 1/1$), whereas character 6 shows two steps and thus its cI = 0.5 ($m/s = 1/2$). The amount of homoplasy implied by the whole data set can be measured using the "ensemble consistency index" (CI), which is the ratio of the minimum number of steps implied by all characters (M) to the length of a cladogram (S). For the cladogram in **Figure 3e**, the CI = 0.86 ($M/S = 6/7$).

There are three perceived problems with the consistency index as a measure of homoplasy. First, although uninformative characters do not add any grouping information to a cladogram, they will inflate its CI. However, this is of significance only when different data sets are being compared. Second, CI can never attain a zero value. A data set in which all possible informative characters occur in equal numbers (an "undecisive" matrix) provides no evidence for preferring one cladogram to any other. Nevertheless, these cladograms will all have positive, nonzero CI values. Third, it has been observed empirically that CI decreases as the number of taxa increases, irrespective of change in the information content of the data. However, this is a recognized and expected property of the CI.

Although the consistency index is useful as a measure of the amount of homoplasy in a character or data set, it is indifferent to the pattern of fit on a cladogram. A binary character that occurs on two separate terminal branches will have the same ci value (0.5) as one that supports two separate groups of taxa. However, in the former case, the character

contains no grouping information, while in the latter it is a synapomorphy (albeit homoplastic) for two groups of taxa. The amount of similarity in a character that is interpreted as synapomorphy is measured by the "retention index" (ri). This is defined as $(g-s)/(g-m)$, where s and m are the same variables as for ci, and g is the maximum number of steps that a character can show on any cladogram. For character 5 in **Figure 3e**, the minimum and observed number of steps is one, and the maximum number is two. Hence its ri = 1 ($(g-s)/(g-m) = (2-1)/(2-1)$) and all similarity is interpreted as synapomorphy. In contrast, for character 6, the minimum number of steps is one, and the observed and maximum number is two. Hence its ri = 0 ($(g-s)/(g-m) = (2-2)/(2-1)$) and none of the similarity is interpreted as synapomorphy. The method can be extended to the whole data set as the "ensemble retention index" (RI), which uses the summed values of g, s, and m (G, S, and M, respectively). For the cladogram in **Figure 3e**, the RI = 0.67 ($(G-S)/(G-M) = (9-7)/(9-6)$).

Character Weighting

The application of differential weights to characters has a long history in systematics. Methods of weighting can be divided into a priori and a posteriori procedures, depending on whether they are applied before or after cladogram construction.

A priori approaches to character weighting generally invoke beliefs that some characters are more important than others or use a particular model of evolution or character change, under which certain types of transformation are considered more or less likely than others. For example, it is common when analyzing nucleotide sequence data to downweight transition substitutions relative to transversions. Alternatively, changes in third codon positions may be disregarded (i.e., accorded zero weight) because they are considered much more likely than changes in first or second positions as a result of the redundancy of the genetic code. Numerous other models have been proposed and they are particularly frequent in the field of molecular systematics. However, such weighting schemes are justifiable only insofar as their underlying model is justifiable.

A posteriori weighting schemes are based on "cladistic consistency" (i.e., the fit of characters to a cladogram) and characters with greater fit are accorded greater weight. One indication of a character's fit is the amount of homoplasy it shows on a cladogram. However, homoplasy does not imply that all similarity is uninformative and the proportion of similarity interpreted as synapomorphy also needs to be taken into account. Hence, both the consistency index (homoplasy) and the retention index (synapomorphy) can be used to estimate character weights. Farris (1989) suggested using the product of these two measures, the "rescaled consistency index" (rc). By combining the ci and ri in this way, characters in which none of the similarity is synapomorphic (ri = 0) receive zero weight, irrespective of their level of homoplasy. All other characters, which contain some amount of grouping information (ri > 0), are differentially weighted according to their level of homoplasy. Using this approach, in **Figure 3e** character 5 would receive a weight of 1 (ci = ri = 1), but character 6 would receive a weight of 0 (ci = 0.5, ri = 0). These weights are applied in a new analysis and the most parsimonious cladogram(s) obtained are used to estimate a new set of weights. This procedure is repeated until a stable set of both weights and most parsimonious cladograms is achieved; hence the name "successive approximations character weighting."

The level of homoplasy of a character may also be viewed as the number of extra steps required to fit it to a cladogram. If all extra steps in all characters are considered equal, then a linear fitting function is being applied to their relative cladistic consistency. For example, in **Figure 3e**, the single step of character 5 is considered equal to each and either of the two steps of character 6. However, intuitively, we might consider that characters showing fewer extra steps are "better" than those showing more. The former can be assigned higher weights than the latter using a concave fitting function of relative cladistic consistency. This approach was implemented by Goloboff (1993) as "implied weighting," in which the weight (W) accorded to a character is calculated as $W = K/(K + ESi)$. ESi is the number of extra steps shown by the character and K is the "constant of concavity." The value of K can be varied to weight more or less strongly against those characters with the most extra steps. As K decreases, these characters will receive progressively lower weights. For example, in **Figure 3e**, character 6 will receive a weight of 0.85 when $K = 6$, but a weight of only 0.5 when $K = 1$ (the "perfect" character 5, which has no extra steps, receives the maximum weight of 1). The optimal cladogram is that for which the summed weights for all characters has the largest value.

Consensus Cladograms (Trees)

A cladistic analysis will often produce more than one most parsimonious cladogram as a result of contradictory signal in the data (homoplasy). The agreement between such "fundamental cladograms" (so-called because they are generated directly from the analysis of data) can be conveniently summarized by means of a consensus cladogram (usually referred to as a consensus tree). Several consensus methods have been proposed, of which the most widely used are "strict," "combinable components" (or "semistrict"), "Adams," and "majority-rule."

The strict consensus tree is the most conservative, because it includes only those groups (often referred to as "components") that are common to all the fundamental cladograms. For example, in the two cladograms shown in **Figures 7a** and **b**, only groups ABC and DEF occur in both, and thus these are the only groups that appear in the strict consensus tree (**Figure 7d**). Groups EF, AB, and BC are excluded because the first is lacking from **Figure 7b** (where it is unresolved) and the other two are contradictory.

However, it is possible for a group to be lacking from one or more fundamental cladograms and yet be uncontradicted. For example, group EF in **Figure 7a** does not conflict with the cladogram in **Figure 7b** because it is one of the three resolutions possible for the group DEF. Combinable components consensus allows such non-replicated, but non-conflicting, groups to be included in the consensus tree, in addition to those groups in common (**Figure 7e**). When all fundamental

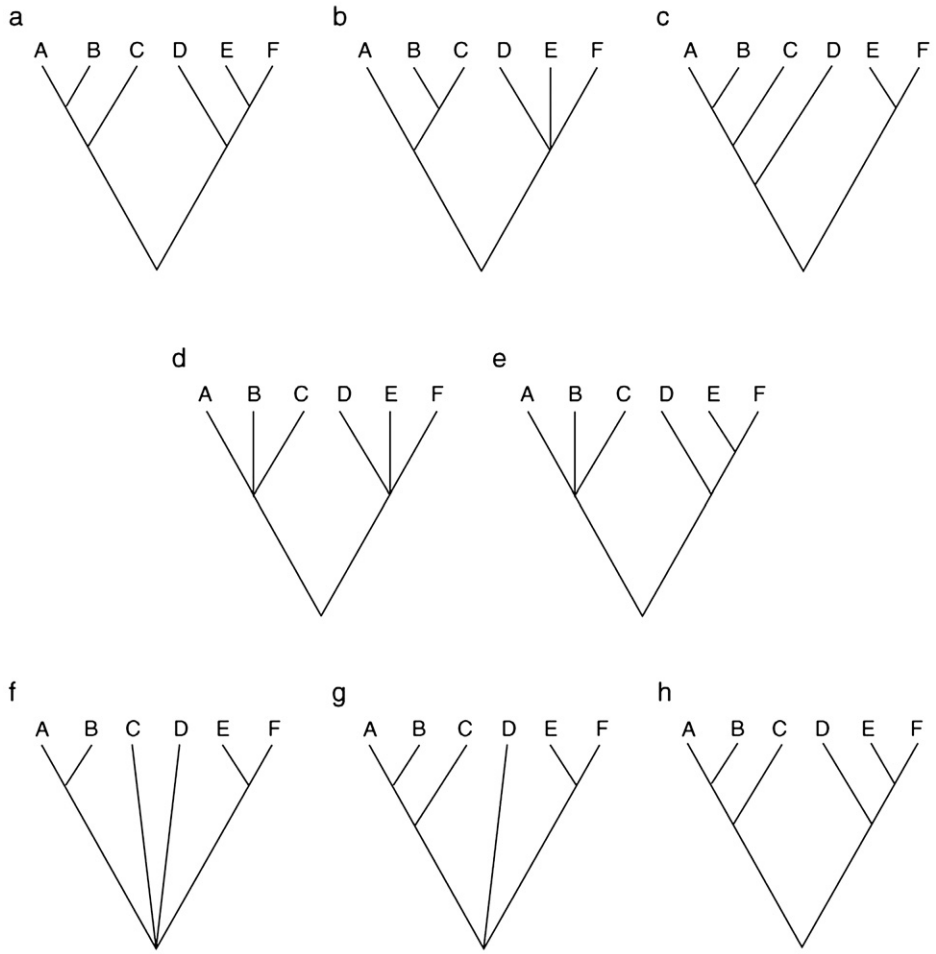


Figure 7 Consensus analysis. (a–c) Three cladograms for six taxa, A–F. (d) Strict consensus tree of cladograms 7a and 7b. (e) Combinable components consensus tree of cladograms 7a and 7b. (f) Strict consensus tree of cladograms 7a and 7c. (g) Adams consensus tree of cladograms 7a and 7c. (h) Majority-rule consensus tree of cladograms 7a, 7b, and 7c.

cladograms are fully resolved, with no spurious resolution due to ambiguous optimization, then the strict and combinable components consensus trees will be the same.

A problem with both strict and combinable components consensus is that a single taxon appearing in highly disparate positions on two cladograms is sufficient to collapse all intervening resolution. For example, taxon D in **Figures 7a** and **c** appears as the sister-group to two different terminal taxon-pairs. Consequently, the strict consensus tree (**Figure 7f**) is relatively unresolved. However, examination of **Figures 7a** and **c** shows that taxon D is acting as a "rogue" taxon; that is, apart from its differing positions, the resolution of the remaining taxa is identical in the two cladograms. Such rogue taxa can be identified using Adams consensus analysis. On an Adams consensus tree, taxa in conflicting positions on the fundamental cladograms are placed at the most inclusive node they have in common; in other words, the consensus contains all intersecting sets of taxa common to the fundamental cladograms. However, as a result, it is possible for an Adams consensus tree to contain groups that are not found on any of the fundamental cladograms and thus they need to be interpreted with care. Sometimes taxa such as D would simply be deleted to give the "largest common pruned tree" as a consensus.

When the number of fundamental cladograms is large, strict consensus trees are often very poorly resolved and can be viewed as too restrictive. One method of increasing resolution of the consensus is to retain those groups that occur in a prespecified number of the fundamental cladograms. Typically, such majority-rule consensus trees will comprise those groups that occur in more than 50% of cladograms. The majority-rule consensus tree of the cladograms in **Figure 7a–c**, shown in **Figure 7h**, is fully resolved, despite the marked topological differences in its fundamental cladograms.

Regardless of their number, the most parsimonious cladograms found by cladistic analysis remain our best estimate of the relationships among the taxa under study. Because resolution is lost, most consensus trees are less parsimonious than their fundamental cladograms. However, if the topological differences among the fundamental cladograms are due solely to ambiguous optimization, the strict consensus tree will also be of minimum length. Then, the strict consensus tree is also the strictly supported cladogram (Nixon and

Carpenter, 1996) and is the preferred most parsimonious topology because it is the only cladogram that is both of minimum length and has all groups supported unambiguously by data.

Cladogram and Group Support

A number of statistics attempt to assign levels of support or confidence to the results of cladistic analyses. They can be divided into two categories: methods that address support for an entire cladogram and aim to determine whether there is any "significant" structure in the data, and methods that examine the support afforded to individual groups on a cladogram and attempt to distinguish those groups that are well supported from those that are not.

Methods aimed at assessing support for an entire cladogram all use the same general principle. The length of the most parsimonious cladogram obtained from the observed data set is compared with those derived from a large number of "phylogenetically uninformative" data sets, with the expectation that the former will be substantially shorter than any of the latter. Several definitions of "phylogenetically uninformative" data have been proposed, including "statistically random" (scores in a data matrix are allocated at random), "undecisive" (all possible informative characters occur in equal numbers), and "randomly co-varying." The last of these forms the basis of the "permutation tail probability" (PTP) test. Covariation here is the degree to which characters are explicable by the same cladogram (i.e., congruence). A most parsimonious cladogram that contains a large amount of homoplasy may be derived from data exhibiting such poor covariation that randomly co-varying characters could produce a cladogram of equal length or shorter. The PTP test uses pseudoreplicate data sets in which character codes are randomly reassigned to taxa, with the restriction that the proportion of each code is maintained, and with each character treated independently. For example, for three taxa (A, B, and C) originally coded 0 and two (D and E) coded 1, permutation may reassign 0 to A, C, and D and 1 to B and E. This procedure is repeated to create a large number of such pseudoreplicates, for which the most parsimonious cladograms are then found. The PTP is defined as the proportion of all data sets (original plus permuted) that yield most parsimonious cladograms at least as short as the original data set and may be interpreted as the probability that a cladogram of this length could have arisen by chance.

The simplest measure of support for a particular group on a cladogram is branch length. However, homoplasy makes the assessment of branch support difficult and groups may appear to be better supported than they actually are. "Bremer support" is a more precise measure of clade support and is defined as the number of extra steps required to lose a clade from the strict consensus tree of near-minimum-length cladograms. When there is no homoplasy in the data, the Bremer support of a group is the same as its branch length. Otherwise, support is reduced to the extent that there are alternative equally parsimonious groupings. To calculate Bremer support, first those cladograms that are one step longer than minimum are found and the strict consensus tree formed from them and the most parsimonious cladogram. The process is repeated, adding a step at a time, until the group in question is lost from the consensus. The number of extra steps required to achieve this is the Bremer support for the group. If more than one most parsimonious cladogram is found initially, then the procedure begins with the consensus tree of these cladograms. Any group that may be a potential resolution of the consensus will have a Bremer support equal to 0.

The bootstrap seeks to estimate group support using pseudoreplicate data sets, which are formed by randomly sampling characters with replacement. The effect is to weight some characters and delete others, with the constraint that the total weight equals the original number of characters. A large number of such pseudoreplicates are generated and their most parsimonious cladograms are found. Conflict among these cladograms is assessed using a majority-rule consensus tree and the support for a group is estimated as the proportion of most parsimonious cladograms on which it is recovered.

However, both the PTP test and bootstrap are questionable in a cladistic context. The PTP test's null hypothesis of randomly co-varying characters is contrary to the basis of cladistics. Characters as nested homology statements are intrinsically hierarchical and thus it is inappropriate to measure their performance against randomized characters from which this hierarchy has been removed. The bootstrap assumes that a data set represents a random sample of all possible characters. However, taxonomic characters are generally carefully selected with the aim of resolving the relationships of the taxa under study. Furthermore, it only requires a single synapomorphy to diagnose a clade. However, the random nature of the bootstrap process means that such a character may be represented in only a few pseudoreplicates and thus the group will not appear in the majority-rule consensus tree despite being uncontradicted. Thus, bootstrap values are a one-sided test; recovered groups have some measure of support in the data, but groups that are not recovered cannot be rejected.

Simultaneous and Partitioned Analysis

It is generally recognized that data from many sources may be used in a cladistic analysis (e.g., morphological, physiological, behavioral, ecological, or molecular sequences) and that analyses of these data can yield different hypotheses of relationships. Simultaneous analysis (sometimes called a total evidence approach) combines all data, from whatever source, into a single taxon $\times$ character matrix for analysis. The resulting hypothesis of relationships is thus determined by character congruence (**Figure 8a**). Alternatively, different classes of data may be analyzed separately and the resulting cladograms added together using a consensus method to extract the common phylogenetic signal. This is called the partitioned evidence approach (**Figure 8b**) and the result is determined by taxic congruence. The reasoning underlying this approach is that different classes of data may reflect different evolutionary processes and so should be analyzed separately. A third alternative, known as conditional data combination, attempts to distinguish those conditions under which it would be best to keep data sets separate and conduct partitioned evidence analysis from those conditions under

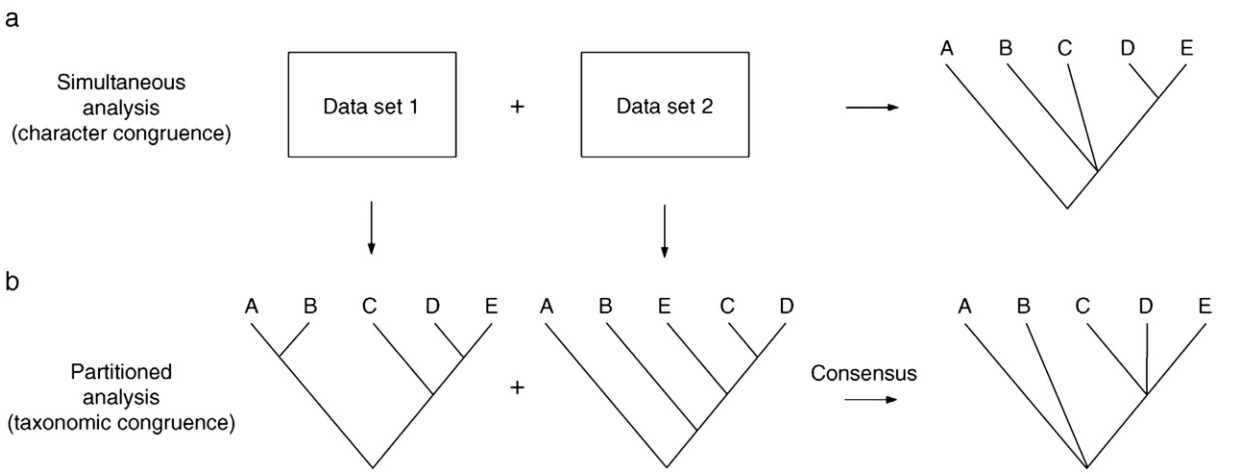


Figure 8 Simultaneous and partitioned analysis. In simultaneous analysis, all data are combined into a single matrix before analysis. In partitioned analysis, each data set is analyzed separately and the resulting cladograms are then "added" together using a consensus method.

which it is more appropriate to conduct a simultaneous analysis. This approach estimates the degree of heterogeneity of phylogenetic signal among data sets, and if the heterogeneity is greater than might be explained by sampling error, then the data sets are analyzed separately.

See also: Cladogenesis. Diversity, Taxonomic versus Functional. Evolution, Theory of. Systematics, Overview. Taxonomy, Methods of

References

Farris JS (1989) The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.

Goloboff PA (1993) Estimating character weights during tree search. *Cladistics* 9: 83–91.

Hennig W (1950) *Grundzüge einer Theorie der phylogenetischen Systematik*. Berlin: Deutsche Zentralverlag.

Hennig W (1966) *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Kitching IJ, Forey PL, Humphries CJ, and Williams DJ (1998) *Cladistics: The Theory and Practice of Parsimony Analysis*, 2nd ed. Systematics Association Publication 11, Oxford, United Kingdom: Oxford University Press.

Maddison WP, Donaghue MJ, and Maddison DR (1984) Outgroup analysis and parsimony. *Systematic Zool.* 33: 83–103.

Nelson GJ (1973) The higher-level phylogeny of the vertebrates. *Systematic Zool.* 22: 87–91.

Nelson GJ (1978) Ontogeny, phylogeny, paleontology, and the biogenetic law. *Systematic Zool.* 27: 324–345.

Nixon KC and Carpenter JM (1996) On consensus, collapsibility, and clade concordance. *Cladistics* 12: 305–321.

Watrous LE and Wheeler QD (1981) The outgroup comparison method of character analysis. *Systematic Zool.* 30: 1–11.

Cladogenesis

Christopher J Humphries, The Natural History Museum, London, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 693–707, © 2001, Elsevier Inc.

Glossary

Adaptation Any genetically controlled characteristic that increases an organism's fitness, usually by ensuring the organism to survive and reproduce in the environment it inhabits.

Adaptive radiation The evolution of one or a few forms into many different species that occupy different habitats within new geographical areas or habitats.

Anagenesis A pattern of evolutionary change involving the transformation of an entire population, sometimes to a state different enough from the ancestral population to justify renaming it as a separate species; also called phyletic transformation of an un-branched lineage of organisms which changes to such an extent as to be justifiably called new species or taxa. This is considered "biological improvement" (sensu Huxley, 1958), covering all types of change from detailed adaptation to generalized organizational advance.

Clade The branch between the nodes on a cladogram or phylogenetic tree; the unit of cladogenesis.

Cladogenesis Branching evolution; the origin of diverging new forms from an ancestral lineage; the form of divergence and phyletic splitting, including speciation; adaptive radiation of species to major divergence of families and phyla. Evidence for cladogenesis is from those fossil groups that increase in the number of recognizable taxa over time.

Genotype The genetic constitution of individuals, species and taxa.

Grade The product and "unit" of anagenesis; paraphyletic group.

Linnean system/hierarchy The hierarchical arrangement of inclusive categorical ranks in which, apart from the lowest rank, subordinates members of all lower ranks; it contrasts with exclusive classifications (e.g., *scala naturae*).

Monophyly (monophyletic group) A natural group that includes a most recent common ancestor and all of its descendants.

Natural classification Systems of classification portraying as accurately as possible the entire pattern of life and its relationships.

Ontogeny The developmental history of an organism from egg to adult. Includes embryogenesis that describes the generative phase and the allometric phase of growth and maturity.

Paraphyly A group that includes a most recent common ancestor and only some, not all, of its descendants.

Phenotype The characteristics of an individual and all its parts as an interaction between the genotype and the environment.

Phylesis Line(s) of descent between ancestors and descendants.

Phylogeny The inferred lines of descent showing genealogical relationships of organisms; often used to describe nested series of monophyletic taxa both in cladograms and phylogenetic trees.

Polyphyly (polyphyletic group) A group that does not include the most recent common ancestor of all of its members.

Punctuated equilibrium The morphological stasis of fossil species over long periods of time punctuated by seemingly instantaneous speciation and morphological change.

Scala naturae (Great Chain of Being) One of the most pervasive ideas in western thought derived from two concepts Plato's principle of plenitude and the linear series of Aristotle and Plotinus (Panchen, 1992); in plain English, the belief in a linear progression from the simplest forms of life to the most perfect.

Stasis (stasigenesis) The persistence of organisms through long geological periods of time that greatly resemble their fossil forebears. The "process" is used to recognize delimitable anagenetic units, or "grades."

Synapomorphy Shared derived characters; homologies of monophyletic groups.

Tree of porphyry A tree classification constructed as a dichotomous key that at any rank divides on the basis of differential characters to give two taxa at the rank below. The net result is a comblike branching diagram intended to classify individuals within a general scheme.

Diversification—Patterns and Process

The terms anagenesis (divergent evolution), cladogenesis (branching evolution or diversification), stasis (stasigenesis; constancy), grades, and clades are all associated with mechanisms and patterns of phyletic evolution. Anagenesis is evolutionary change in lineages through time, cladogenesis is diversification of clades by branching of lineages through time (Rensch, 1959), and stasigenesis (stasis) refers to persistence of lineages through time (Huxley, 1958). The distinctions came about through the belief that in evolution there were three main processes at work that led to "biological improvement," to "diversification," and to "persistence." Extinction, although seen by some to be of importance in evolution,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00023-X

is perceived as a failure to persist and is now treated as an important part of diversity assessments of life on earth. The concepts of phyletic evolution came to prominence in the neo-Darwinian synthesis of genetics and evolution, but their meanings changed when viewed through the lenses of different systematic perspectives. Grades, for example, once seen as extremely important with respect to the pattern of evolution in the scala naturae, are now seen as artifacts, described as paraphyletic and polyphyletic groups in the literature of cladistics where the emphasis lies in the analysis of clades and the recognition of sister groups. Furthermore, the processes of anagenesis and cladogenesis refer either to microevolutionary processes (speciation) or to patterns and rates of evolution through geological time calculated as rates of morphological, taxic, and molecular change.

The centerpiece of evolutionary theory is that the environment varies with time and from place to place. Heritable variations that suit a particular environment are selected *in situ*, and so with time populations diverge and differentiate as each becomes adapted to its own conditions (Patterson, 1999). Darwin (1859) considered that such adaptive change led to diversity of form, with features thought to result from the process of diversification. Cetaceans and sea grasses, for example, are thought to be marine-adapted mammals and flowering plants, respectively, with land-borne ancestors. With the impact of tremendous changes in systematic thinking and population biology, Darwin's theory of organic evolution is now considered by most people to be the only worthy explanation of the diversity and form of life on earth. Darwin's theory has two quite different and distinct aspects. Evolution by common descent is used to describe the patterns of changes in diversity through time, and evolution by natural selection describes the processes leading to diversity as a measure of change over time.

The language of micro- and macroevolution emerged at a time when fossils were considered as of cardinal importance for understanding the patterns and processes of evolution, when evolutionary theory, which was making great strides by amalgamating with genetics and evolutionary systematics, had yet still to mature. Evolution is the key. However, the greatest strides over the past 30 years have been made in systematics, where fossils have been analyzed together with recent taxa. The relationships have changed, and systematics is the key. Progress in evolutionary theory has largely been in mechanisms. The concepts of cladogenesis and anagenesis have consequently changed. For example, as interpreted in the light of phylogenetic systematics (or cladistics) grades pertain to polyphyly and paraphyly, and clades to monophyly. The origin of species or higher taxa by anagenesis, cladogenesis, or phylogenesis, and persistence through stasis, is associated more with process biology in microevolution (see speciation processes) and particularly the assessment of patterns and rate of macroevolution.

The Synthetic Theory

The synthetic theory of natural selection (often described as neo-Darwinism) incorporates the original theory of natural selection of Darwin and Wallace (Darwinism) but also incorporates models of population genetics and the underlying bases of morphological and genetic heterogeneity within organisms. Classic genetic studies on model organisms such as *Drosophila*, *Cepaea*, and mimetic butterflies (see Panchen, 1992) have largely been concerned with genetic variation on different alleles in heterogeneous populations and the consideration of adaptive significance, which led to the rather stifling view that all change is adaptive. This has led to the idea that natural selection is essential for adaptive, anagenetic change and is a necessary precondition for cladogenesis (speciation). The processes of anagenesis, cladogenesis, and stasis are used to describe changes (or the lack thereof) in phenotypes and genotypes through changes in the environment both in space and time and by mutations and heritable variations within phylogenetic lineages. Patterns of diversity are described in phylogenetic schemes or genealogical trees through systematic analysis.

Speciation

As Huxley (1958) noted, Darwin (1859) recognized three processes at work "leading respectively to biological improvement, to diversification, and to persistence." Improvement covers adaptation to specialized lifestyles, through phyletic transformation of lineages, resulting in the derivation of descendent species directly from ancestors. These are recognized as "transformational" or "evolutionary" species. New species arise as recognizable transformations within a lineage, by-products of the genetic properties of the lineage, and are thus recognized as arbitrarily divided segments of that lineage. Diversification covers all patterns and processes from micromorphological changes and subspeciation (microevolution), adaptive radiation (cladogenesis with adaptation), and macroevolution, the division into major phyla and higher taxa. Thus, the processes of divergence include lineage splitting, speciation, and adaptive radiation where genetic isolating mechanisms result in two or more descendants derived from a common ancestor. Persistence refers mostly to the idea that organisms of great geological age survive unchanged as living fossils into contemporary biotas. Little or no detectable change takes place through time such that the modern organisms resemble fossils from early geological periods. Examples include animals, such as *Latimeria* (Coelacanth), *Nautilus*, and *Sphenodon*, and plants, such as *Ginkgo* (Maidenhair tree), *Wollemia* (Wollemi Pine or "dinosaur tree"), and the *Metasequoia* (dawn redwood).

Anagenesis, Cladogenesis, and Stasis

New species can either become modified through gradual change in an entire lineage in response to a changing environment or can emerge through diversification into two or more species through formation of internal and external isolating mechanisms. Anagenesis is gradual change in an entire lineage (**Figure 1**). Division into two or more species is termed cladogenesis (**Figure 2**). There are many modes of population differentiation and considerable debate as to their roles in speciation. Likewise the mechanisms for anagenesis and cladogenesis (speciation) are many and varied. Nevertheless, they

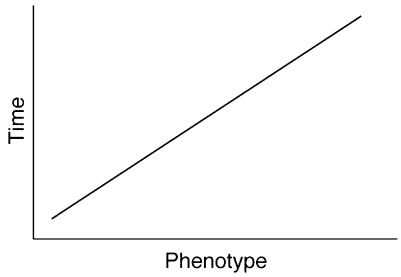


Figure 1 Anagenesis. The transformation of a species or taxon on an unbranched lineage of organisms. Divergence has occurred to such an extent that it is justifiably called a new species or taxon.

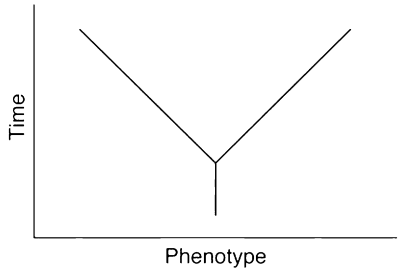


Figure 2 Cladogenesis. The transformation of a species or taxon into two (or more) species (or taxa) by branching along a lineage.

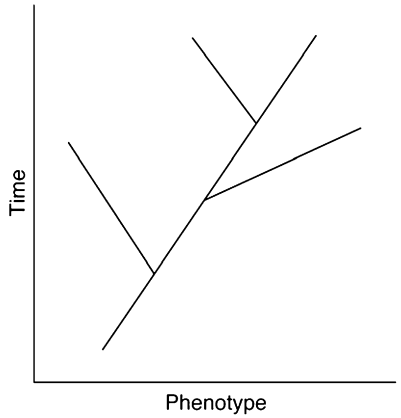


Figure 3 Phyletic gradualism. A fairly constant rate of change through time.

all include some form of population differentiation, either by gradual change of genetic differences and natural selection (phyletic gradualism), or by abrupt punctuational changes, involving chromosome inversions or translocations and rapid isolation between populations, or by historical accidents, such as allopatric speciation by vicariance and isolation.

There is considerable controversy among the theories of diversification and how evolution proceeds. The Darwinian hypothesis is phyletic gradualism, whereby the same micro-evolutionary processes that lead to population differentiation cause ever-increasing divergence between populations (**Figure 3**). Eventually, sufficient divergence has occurred for differentiation at species level to be recognized. Differentiation continues at a steady rate and new species originate by

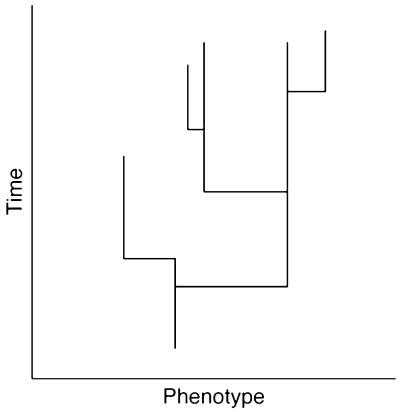


Figure 4 Punctuated equilibrium. Morphological stasis of fossil taxa over long periods apparently punctuated with occasional instantaneous change.

slow, gradual changes of ancestral species. The neo-Darwinian perspective is that evolutionary transformation takes place within species, or lineages, and the branching process of cladogenesis accounts for diversification but relatively small amounts of evolutionary change.

An alternative theory of evolutionary rates and speciation, punctuated equilibrium, was proposed by Eldredge and Gould (1972; **Figure 4**). Evidence from paleontology on well-preserved fossils indicate long periods of stasis, where "species" remain relatively little changed over long periods of time. At other times there appears to have been rapid evolution and great morphological differentiation. Eldredge and Gould (1972) thus concluded that rates of evolution were not constant over time. They hypothesized that little evolutionary change occurs within species and that genetic changes within populations do not account for different species. Instead, the events of speciation account for evolutionary change, and short periods of "punctuation" were interspersed with little of no evolutionary divergence.

Phyletic gradualism and punctuated equilibrium represent the opposite extremes of a continuum. Recent research, especially on rates of molecular change, would suggest that evolutionary rates are clocklike, or at least "clocky," caused by periods of slow rates of phylesis and rapid periods of cladogenesis. This is reflected in research of the 1980s and 1990s, which has concentrated on uncovering the patterns of divergence through systematic analysis (particularly cladistics). Sustained application of cladistic methods has determined nature's hierarchy and techniques such as maximum likelihood have been used to assess rates of change amongst phylogenetic trees.

Clades, Grades, Gossils, and Ancestors

Systematic classifications are represented in branching diagrams, variously know as trees, phenograms, dendrograms, and cladograms depending on which systematic philosophy is being used. For most practicing biologists of the 20th century, modern systematics dates back to Linnaeus. He recognized that nature's hierarchy could be arranged into a series of

named categories. His system of categories, although still in use today, is somewhat arbitrary, based on the predefined levels of *Regnum* (Kingdom), *Classis*, *Ordo*, *Genus*, and *Species*, to which Phylum and family were later added (Panchen, 1992). All the members of the same level were given the same rank. As distinct from exclusive systems, such as military ranks, biological ranks are all inclusive and genuinely hierarchical, whether viewed agglomeratively from the bottom or divisively from the top, such that eventually all taxa belong to one group, life itself. The hierarchy is divergent, so that a taxon of, say, a specific rank, belongs to a taxon of higher rank and the taxon of this rank belongs to a taxon of even higher rank, and so on. The hierarchy is not necessarily symmetrical, because taxa of high rank can have everything from one isolated member to a great number of members. Elements of Linnean system are still in use today, especially nomenclature, but its avowedly artificial category hierarchy, expressed as a preordained ranking system, has become replaced by a whole succession of "natural" classification systems portraying as accurately as possible the entire pattern of life and its relationships.

Natural Classification

Almost a century of debate has questioned the meaning of "natural" classification systems but by the time that the *Origin* was published (Darwin, 1859), the form of representation had become a branching diagram of hierarchical clusters. To his credit, Darwin made the tree metaphor explicit in what was the only diagram published in the *Origin* (**Figure 5**). This was not a diagram produced by clear-cut rules but a tree rooted in the past—a genealogical tree to denote phylogeny in a vertical direction and classification of relative distances of present-day organisms in the branching hierarchy. One of the most important issues that came from Darwin's evolution by common descent was that phylogeny is the theory that underpins natural classification today.

Fairly soon after the *Origin* appeared, "forests" of genealogical trees began to appear in the European literature. Perhaps the most memorable rendition of all schemes ever published is that of Haeckel (1874) who used literally the European oak (*Quercus robur*) to synthesize his interpretation of phylogeny and classification and the origin of man into a natural classification using Darwin's tree metaphor (**Figure 6**). Apart from its shape, the named parts of the tree embody two quite separate concepts. The shape is simple enough in that it depicts a multifurcating inclusive genealogy with the names of the branches indicating monophyletic groups (groups derived from a common ancestor). However, the names on the vertical axis, horizontal lines or "grades" of evolutionary achievement, embody a watered down version of the *scala naturae*, a belief in the progression from the lowliest simplest forms of life

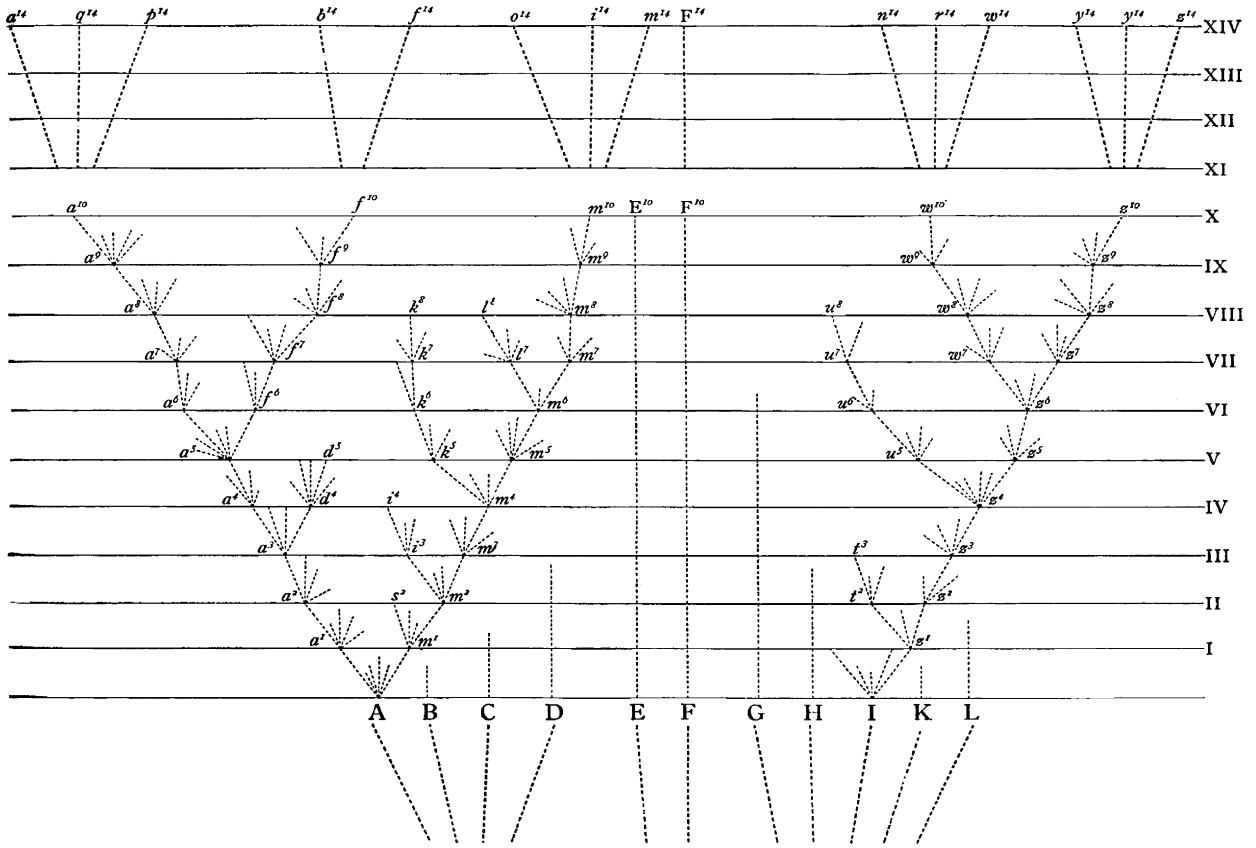


Figure 5 Darwin's (1859) branching diagram from the *Origin*. Note that A–L represent the species of a genus. The small letters show degrees of divergence through time. The horizontal lines represent 1000 or more generations. The diverging lineages are shown with extinct populations at most generations in every lineage. Reproduced from Darwin C (1859) *The Origin of Species by Means of Natural Selection*. London: Murray.

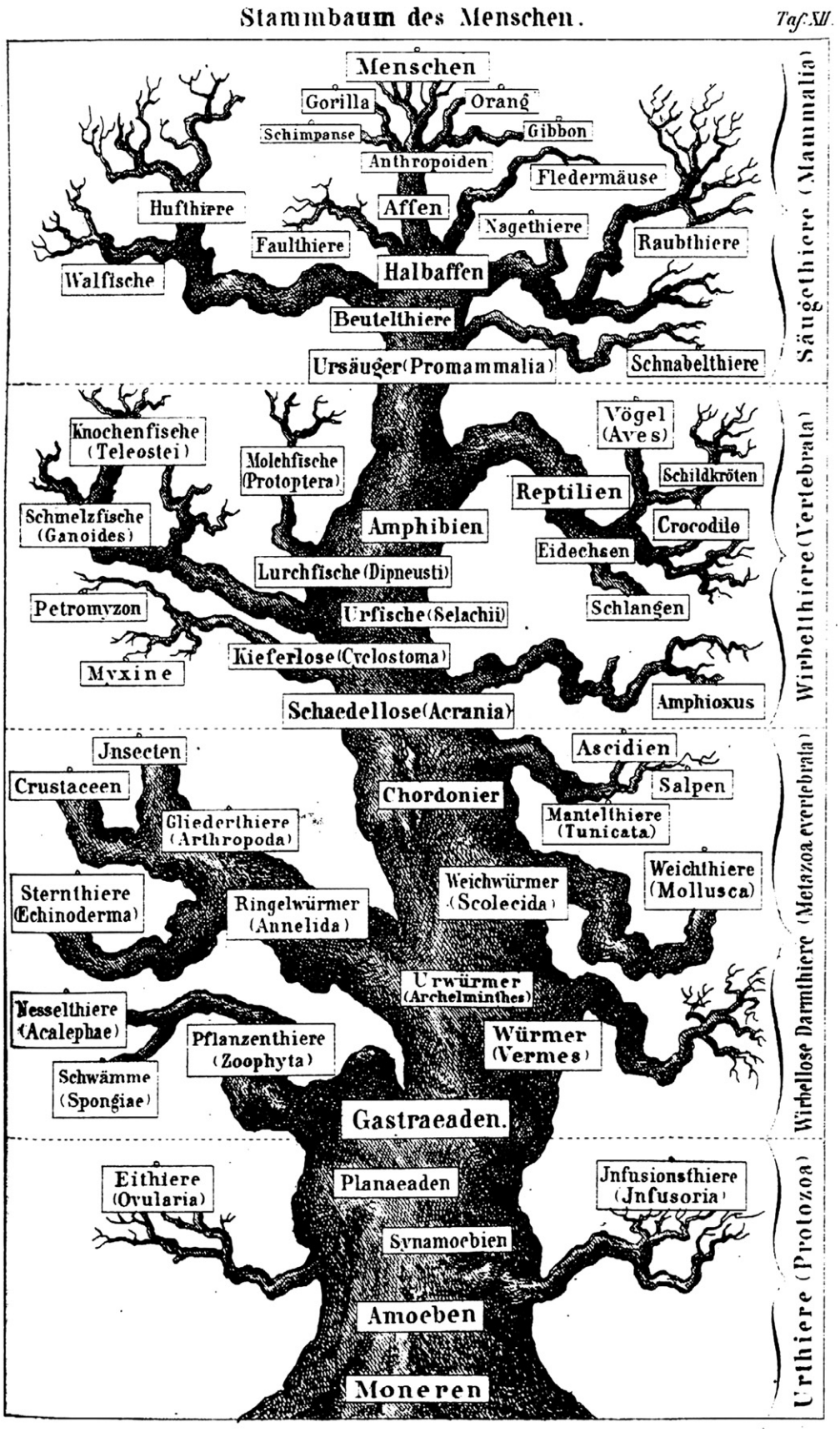


Figure 6 Haeckel's (1874) "literal tree" (European oak) depicting human ancestry as interpreted by Darwin's theory of evolution by common descent. The bifurcating branches indicate groups determined on embryology and the vertical axis (conflated with time) represent grades. Reproduced from Haeckel E (1874) *Anthropogenie: Keimes- und Stammes-Geschicte des* Menschen, 1st ed. Leipzig: W. Engelmann.

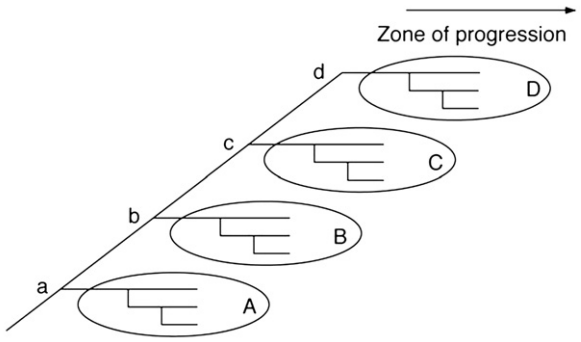


Figure 7 Abel's concept of Stufenreihe (clades A–D) and Ahnen-reihe (grades a–d). See text for explanation. Reproduced from Figure 3.6 in Panchen AL (1992) *Classification and Evolution, and the Nature of Biology*, 55 pp. Cambridge: Cambridge University Press.

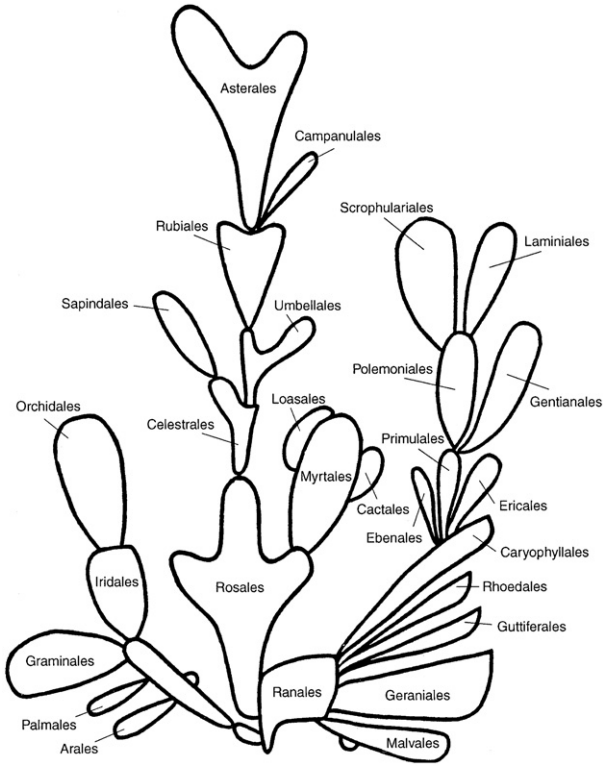


Figure 8 "Bessey's cactus". Diagram to show the relationships of orders. The size of each order is represented by the shape and dimensions of each component of the diagram. For an explanation of a, b, and c, see text. Reproduced from Figure 1 in Bessey CE (1915) The phylogenetic taxonomy of flowering plants. *Annals Missouri Botanical Garden* 2: 109–164.

(Monera) to the most "perfect" (human) through the series protozoa—metazoa—evertebrata—vertebrata—mammalia. The distinction was based on Haeckel's twin concepts of the evolution of the foetus, ontogeny, and recapitulation in the evolution of the stem (phylogenesis). Arising out of former was the phylogenetic interpretation of the *Tree of Porphyry* as a step series (Stufenreihe) and distinguished from the ancestor series (Ahnenreihe) in palaeobiology (see Panchen, 1992).

By the turn of the 20th century, branching diagrams were largely depicted as phylogenetic trees to represent "Ahnenreihe" as "grades" in ancestor-descendant series and "Stufenreihe" to represent evolution by cladogenesis into a succession of adaptive zones (Panchen, 1992; **Figure 7**). There have been endless variations of this theme in comparative biology and different emphases placed on grades and clades. For example, Bessey's cactus (1915), deliberately emphasized grades and even interpreted extant supraspecific taxa, Ranales (**Figure 8a**), as giving rise to other supraspecific taxa, Rosales (**Figure 8b**), which in turn gave rise to Myrtales (**Figure 8c**). At the other extreme, Janchen's (1932) interpretation of Wettstein's classification clearly emphasizes clades with grades added almost as an afterthought to indicate the distribution of floral evolution through gymnospermae/angiospermae, apetalae/sympetalae, and monocotoledony/dicotoledony (**Figure 9**).

Grades and Clades in Evolutionary Systematics

These days we know that Bessey's approach is untenable—one supra-specific taxon cannot give rise to another. The Ranales cannot be the ancestor of the Rosales, and even if we did consider them to have an ancestor-descendant relationship then the Ranales and Rosales would have to be classified in the same group rather than as depicted in **Figure 8**. Nevertheless, grades and clades are still established practice and appear in textbooks even today. The 20th century has seen an enormous output of different phylogenetic schemes and the persistence of grades has pervaded the methods of "evolutionary systematics." The idea that grades are important in systematics so as to express patterns of anagenesis is the hardest notion to expunge. Grades are characteristic of classifications involving "satellite" groups considered sufficiently divergent to separate them at a high rank from their nearest relatives. The more "primitive" taxa are similarly rendered as a grade, a paraphyletic group in the language of phylogenetic systematics or cladistics.

To illustrate, perhaps the two most famous examples include human (*Homo sapiens*) considered superior to its primate relatives, and birds (Aves) considered greatly diverged from relatives in the class Reptilia (**Figure 10**; see Mayr, 1974; Panchen 1992). The net effect of such schemes is that the phylogenetic reconstruction, based on character analysis, differs greatly from the written classification. **Figure 10** illustrates this result in both examples. Thus for the diapsid reptiles the phylogenetic diagram shows ancestor A ancestral to two groups B (Diapsid reptiles) and C + D (crocodiles + birds). The classification differs in that group D (Aves) shows such considerable divergence along the horizontal character axis (the anagenetic component), that to express the information in a classification puts B + C together (diapsid reptiles + crocodiles) and D is kept separate (Aves). A similar story can be portrayed for orangutan + gorillas/chimps and *Homo sapiens*. The phylogenetic diagram shows C + D (gorillas/chimps and humans) as sister group to B (orangutan) but the classification puts B and C together and humans in a group D of their own. Mayr (1974) claimed that the anagenetic component should carry more weight because of many genes shared in common among birds and not found in the nearest relatives and because of the unique

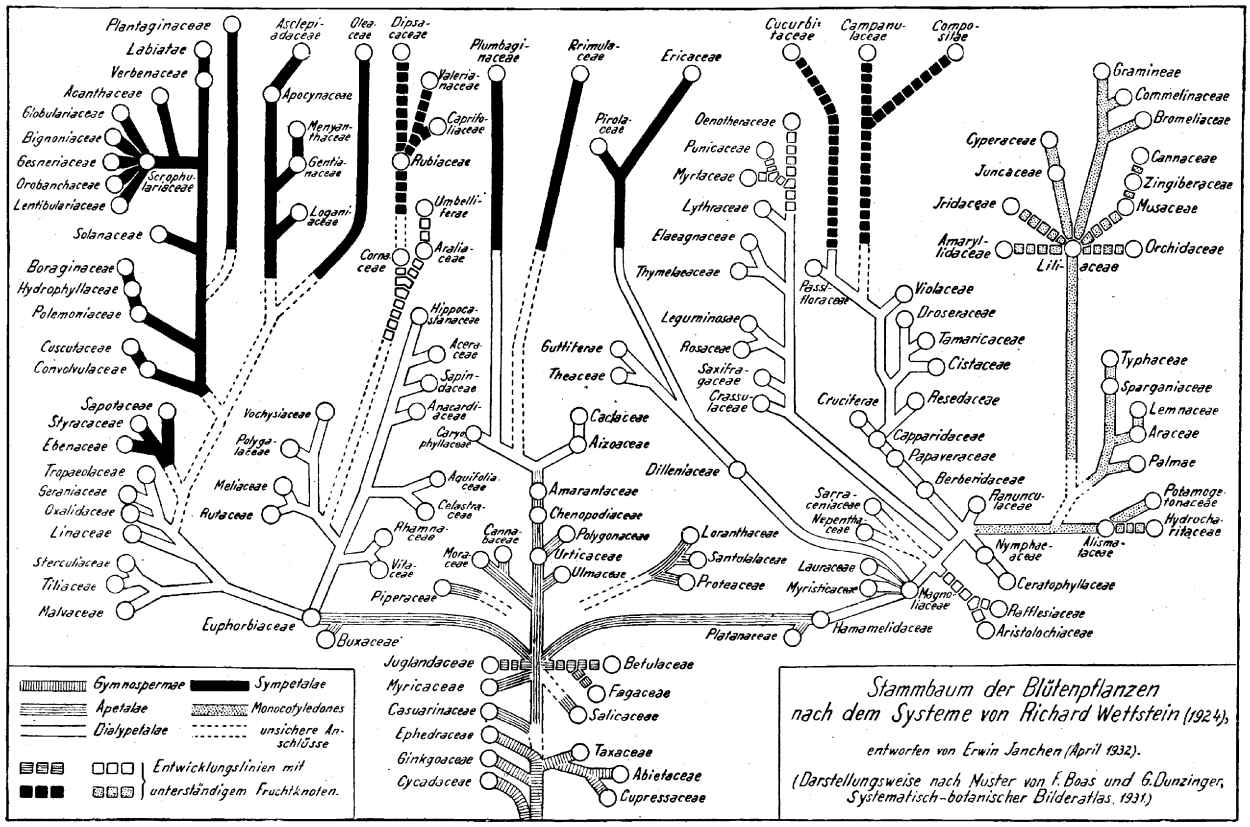


Figure 9 Grades and clades, Janchen's (1932) phylogeny. The phylogeny of the seed plants is shown as a branching diagram. Grades are depicted by irregular, thick black lines crossing the branches of the phylogeny. Reproduced from Janchen E (1932) Entwurf eines Stammbaumes der Blütenpflanzen nach Richard Wettstein. *Öster. Bot. Zeit* 81(3): 161–166.

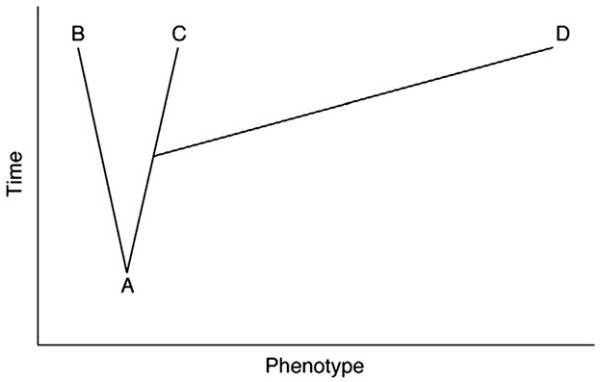


Figure 10 Grades and satellite groups. A = ancestor; B = diapsid reptiles, or orangutan; C = crocodiles, or gorillas/chimps; D = birds, or humans. See text for explanation.

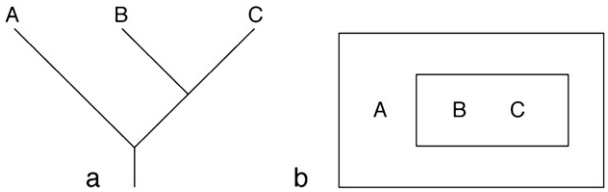


Figure 11 Definition of relationship in (a) a cladogram and (b) a Venn diagram.

behavioral characters in man. Thus, groups represented by D are placed in higher grades and groups represented by B + C in lower grades.

Paraphyly and Monophyly in Cladistics

Cladistics demonstrates that grades are impossible to characterize as they do not express relationships between organisms. There is no way of knowing when one grade starts and another stops. How many characters along a branch are required before it is possible to draw a line and say that one side represents a lower grade and the other side a higher grade? Related to this question is how many taxa belong to a grade group?

Hennig (1996) described a method to implement evolution by common descent by reconstructing phylogenies based on assessments of speciation (cladogenesis) and transformation of characters (now known as cladistics; see Methods of Systematics, and Cladistics). His most important contribution was to offer a precise definition of relationship and a technique for those relations to be discovered. A minimum of three taxa is necessary to express a relationship. For example, in **Figure 11a** taxa B and C are more closely related to each other, than either is to A, because they share a common ancestor not shared by A. Cladistic analysis finds monophyletic groups on the basis of uniquely derived, shared characters (synapomophies).

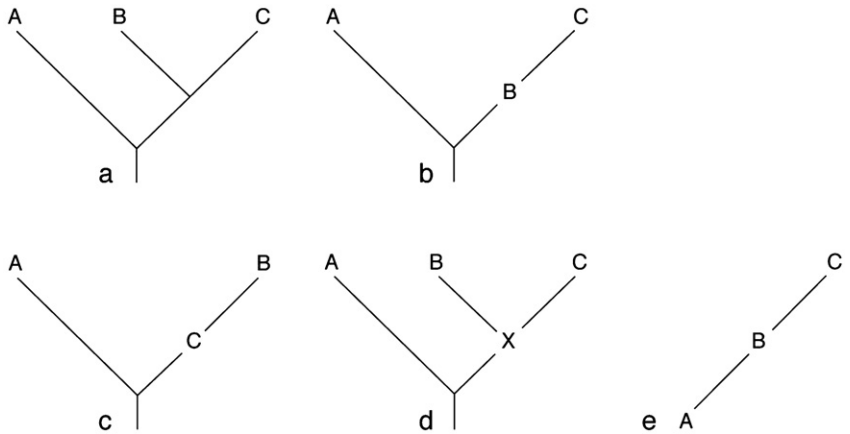


Figure 12 Cladograms and trees: (a) cladogram, (b–e) a selection of phylogenetic trees consistent with the cladogram in (a).

Hennig showed that monophyletic groups are "natural groups," that branching nodes expressed relations, and that synapomorphies are the only measurable quantities for determining pattern. Hence neither horizontal nor vertical branches (anagenesis) are meaningful for expressing relationships on cladograms or for the determination of groups. Vertical branches say nothing about time and the relative nesting of nodes on the cladogram provide only relative rather than absolute estimates of ordinal time. Relationships could just as easily be represented by nested sets or Venn diagrams (**Figure 11b**).

Cladograms are synapomorphy schemes, induced from the most parsimonious distribution of characters to show sister-group relationships. In **Figure 12a**, taxa B and C represent one sister group nested in larger sister group, A and B + C. Cladograms are different from phylogenetic trees because they rely entirely on empirical data, taxa, and characters. They express only the general branching pattern of life because that is all that is available from analysis of taxa and form. The strongest support for this idea is that many phylogenetic trees can be hypothesized for the same cladogram irrespective of branch length (**Figures 12b–e**). Consequently, anagenesis, ancestors, and ancestor-descendant relationships are not directly available from character analysis but require models of one kind or another to arrive at answers to questions of rates of divergence. Cladograms are different from phylograms or phylogenetic trees because they are isomorphic with the classification. Cladograms are consistent with name hierarchy and they can be recovered from written classifications. On the basis of this property, Hennig (1966) justifiably claimed that phylogenetic systematics provided the only truly general reference system consistent with the theory of evolution by common descent.

The task of systematics for Hennig was to understand natural relationships (monophyly, monophyletic groups) and rid the general reference system of polyphyly and paraphyly. Although there has been considerable proliferation in the methods and sources of information in systematics, the main effect has been to put intense effort into cladogenesis by the discovery of clades or monophyletic groups. For the past 25 years or so systematics has concentrated on determining the pattern of life from its earliest beginnings to the highest nodes on the tree, especially as a result of massive strides in molecular biology. Programs in pattern analysis range from the minutiae of phylogeography within species and populations to the discovery of monophyletic clades throughout the entire history of life. On the process side have been intense efforts to discern the rates of macroevolution by calibrating what is known about fossil history and morphological evolution with what is known about base substitution rates in ubiquitous molecules.

Patterns of Life

The emphasis over the past 30 years has switched from worrying so much about whether one species or taxon gave rise to another species or taxon to emphasizing phylogenetic trees and cladistic patterns. The emphasis has switched entirely from ancestor-descendant stories to expressions of sister-group relationships. The emphasis on the overall morphology and anatomy of the phenotype has comparably given way to the use of molecular characters. Furthermore, the discovery that sequences of amino acids of protein chains and nucleic acids contain detailed sources of information has given impetus to the discovery of family trees across the entire hierarchy of life. By comparing homologous amino acid or gene sequences it has been possible to compile family trees using computerized cladistic analyses with thousands of taxa in one calculation. Through a massive empirical enterprise, involving literally thousands of published investigations, the stage has been reached when comparisons between bacteria, flowering plants, and mammals in the same analysis are a reality. The potential for observing the patterns on both a grand scale and at species level is dominating comparative biology in all disciplines. The output of data and cladograms is so vast and the information coming from a variety of museums and laboratories around the world that dedicated web sites are being built to makes sense of it all (see, for example, The Tree of Life web page at http://phylogeny.arizona.edu/tree/phylogeny.html).

The overriding message from all of this empiricism is that large-scale differences between humans, mice, elephants, and oak trees have the same sort of causes as closely related species (Patterson, 1999). It appears that at all scales of investigation the differences between organisms related by common ancestry owe their differences to divergence, the accumulation of mutations over great periods of time. The evidence for this homology comes from the fact that organisms share homologous features at many levels, molecular, structural, and

physiological. At face value it would seem that gradual phylesis would explain changes in both micro- and macroevolutionary events. Several evolutionists have different opinions about this. They can accept point mutations, gene duplications, and chromosome mutations to explain the smaller divergences in birds and butterflies but require other mechanisms to suggest the massive differences between fishes and land organisms or the original appearance of crustacea and green plants. The main reason for requiring macromutations for macroevolution is that many features of major groups cannot be imagined to have come into being by small adaptive changes. The point often made is the existence of complex organs. Familiar questions are of what use is half an eye or half a lung? The answers are not readily apparent because the evidence is unclear. The intermediate steps cannot be imagined. By improving the accuracy of reconstruction of cladograms for the whole of life the potential exists for assessing whether the story is one of gradual phylesis or not.

Rates of Evolution

Rates of evolution have always been of interest to systematists and evolutionists alike, and one of the popular myths is that the fossil record is the only direct evidence to track the course of evolution. During the history of paleontology many arguments have been constructed to link the fossil record to evolution because we know that fossils are the remains of once living organisms and they predate the rock sediments in which they lie (Panchen, 1992). Furthermore, the idea that the stratigraphic record encapsulates the historical record, and that most fossils are extinct but related to modern groups that extend a little way back in time, is the underlying justification for schemes to correlate the pattern of fossil record with evolution through time. It has long been considered possible to calculate different rates for a number of reasons. "Living fossils" that have survived from the geological past to the present-day show that most organisms in the fossil record were once living but are now extinct. Similarly, "Lazarus" taxa, found in an early geological period, which then disappear for millions of years to reappear in more recent periods, were considered to be a good line of evidence to support phylogenetic reconstructions with absolute, rather than relative time, intervals. The fact that many modern groups are modified older groups and can be traced back in time has reinforced evidence for the transmutation of fossil organisms into modern forms and hence the belief that rates of evolution can be calculated.

Morphological Evolution

Perhaps the most cited examples of evolution between fossils and modern taxa are the so-called missing links. Fossils, such as *Archaeopteryx*, were considered to show links between different major groups. The apparent intermediate nature of *Archaeopteryx* by having teeth, as well as derived features of feathers and wishbones, as found in all birds, was used to suggest that birds are descendants of dinosaur-like creatures. No exact sequence of fossils or ancestors could ever be found and it has been known for a long time that no particular group of fossils could give rise to a particular group of modern taxa. The search for ancestors has gradually switched from reconstructing history sequences of grades in a *scala naturae* to the modern approach of comparing dated fossils in stratigraphic sequences with well-corroborated cladograms (Smith, 1994).

There have been many reconstructions of rates of evolution through geological time in fossil and modern taxa of plants, vertebrates, and invertebrates (see Smith, 1994). The oft cited examples are mammals, particularly horses, from the Eocene to the present, because of their good fossil record (see Panchen, 1992; Simpson, 1951, 1953). In the case of horses, the technique was to see how well morphoclines of characters from the most primitive characters in the oldest taxa gradually changed into the derived characters of modern taxa. These were then examined to see how well these matched the chronocline (the age of fossils as determined by their appearance in the stratigraphic sequence). Simpson's (1951) pioneering work traced lines of descent from the fossil *Eohippus* to *Equus* by creating a linked time series, or *scala naturae*, to measure the rates of morphological changes in skull, forelimb, and upper molar characters in a time period of 60 million years. The approach used was to literally measure the amount of change over a generalized phylogenetic tree after dates were assigned to the horse genera in the sequence (**Figure 13**). Despite the great

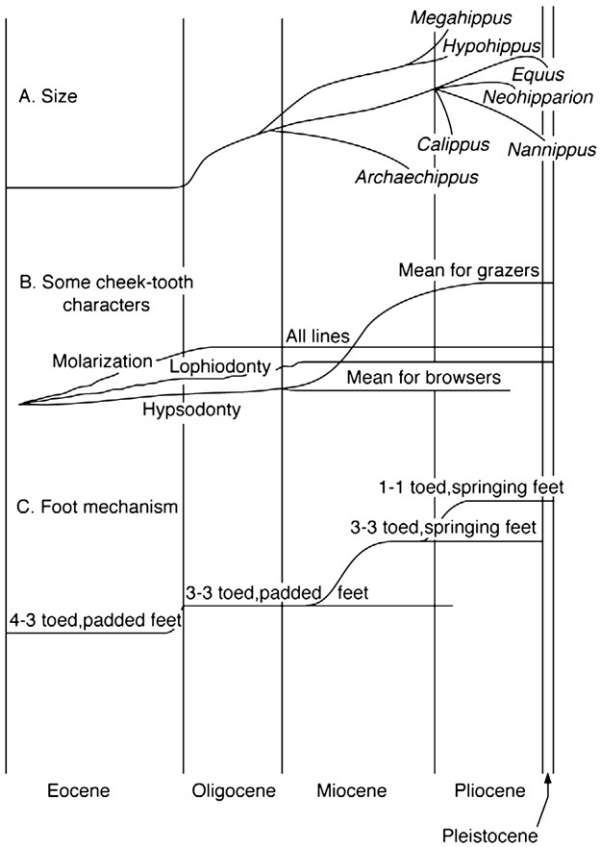


Figure 13 Evolution of fossil horses. (A) represents a phylogram of increasing size (reproduced with permission from Figure 34 in Simpson GG (1953) *The Major Features of Evolution*, 265 pp. New York: Columbia University Press), (B) changes in cheek teeth characters, (C) represents number of toes and foot mechanisms.

Table 1 Differences (distances) in the alpha hemoglobin chains of a selection of vertebrates

	Human	*Elephant*	*Platypus*	*Ostrich*	*Starling*	*Crocodile*	*Lungfish*	*Coelacanth*	*Goldfish*	*Shark*
Human	—	26	40	43	41	47	83	70	68	71
Elephant	18	—	45	45	48	50	84	72	63	74
Platypus	28	32	—	54	52	51	89	74	70	76
Ostrich	30	32	38	—	26	36	91	75	68	73
Starling	29	34	37	18	—	47	91	77	67	70
Crocodile	33	35	36	26	33	—	85	78	70	77
Lungfish	59	59	62	64	64	59	—	90	94	86
Coelacanth	49	51	52	53	54	55	63	—	83	78
Goldfish	48	44	49	48	47	49	66	58	—	88
Shark	50	52	54	52	50	55	63	55	62	—

The alpha chain contains 141 amino acids in most vertebrates, but it contains 142 in the goldfish and the lungfish has 143. Numbers to the upper right half are differences and those in the lower left are percentages. (Reproduced from Table 9.1 in Patterson C (1999) *Evolution*, 2nd ed. London: The Natural History Museum.)

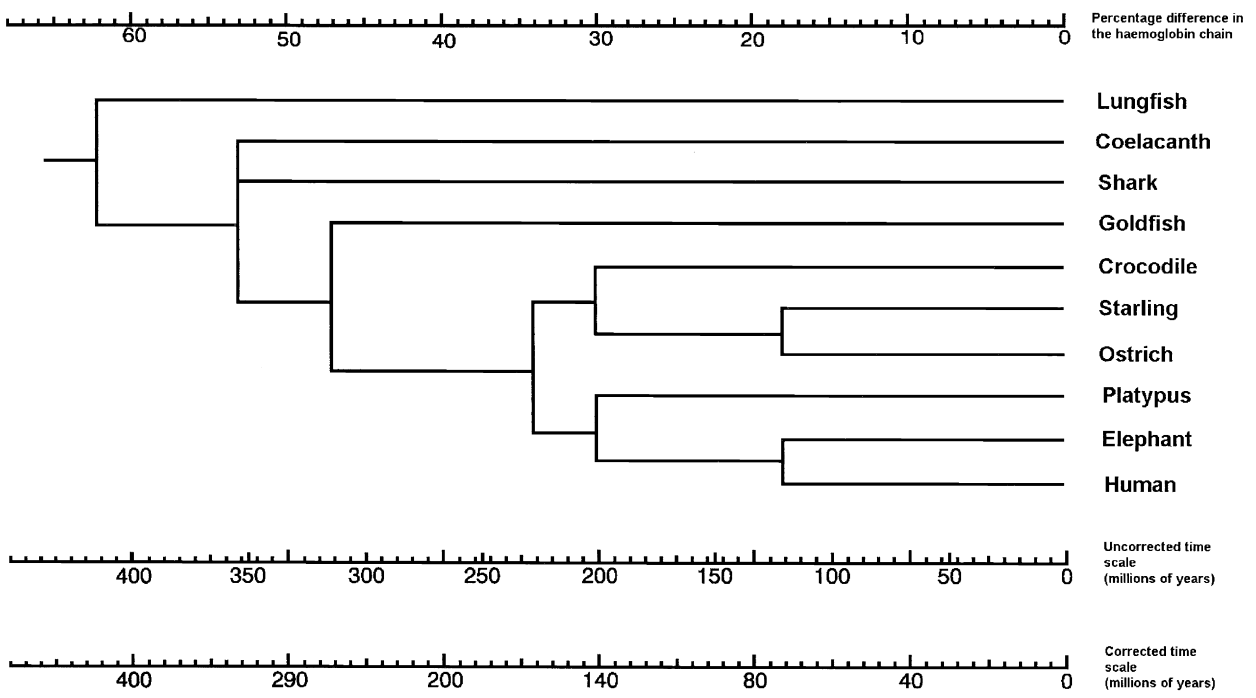


Figure 14 Hemoglobin family tree based on differences between the alpha hemoglobins in **Table 1**. The two time scales are calibrated from the assumption (based on the earliest known fossils) that sharks, lungfishes, and coelocanths diverged about 400 mybp with 60% amino acid divergence calibrated at 400 myrs. Reproduced with permission from Figure 9.1 in Patterson C (1999) *Evolution*, 2nd ed. London: The Natural History Museum.

effort expended by Simpson, his results give only general approximations on the rates of divergence and the absolute dates are hard to accept because of a lack of a well-corroborated cladogram.

Rates of Molecular Evolution—Hemoglobin

The flood of molecular data accrued since the mid-1960s offers the prospect of more accurate assessments in the calculation of rates of evolution. Absolute rates of molecular evolution have been estimated from branching diagrams by calculating trees from various forms of sequence data and alignment of the cladograms to fossil or biogeographic events as fixed starting points. By calculating pairwise distances, the minimum ages of cladogenetic events across pairs of taxa can be estimated. These calculations assume that a "clock" is operating within the molecules by gradual and regular substitution of nucleotides or amino acids constantly over time and that the clock can be calibrated to external events (e.g., the age of a fossil), which then reflect accurately the initial cladogenetic events.

Early estimates in the 1960s were made on chains of amino acids and earliest occurrence of fossils to calculate rates of divergence in proteins. Typical results for alpha haemoglobins (**Table 1**) comparing recent organisms with their fossil relatives suggest that sharks, coelocanths, and lungfishes have hardly changed over 350 million years or more but mammals

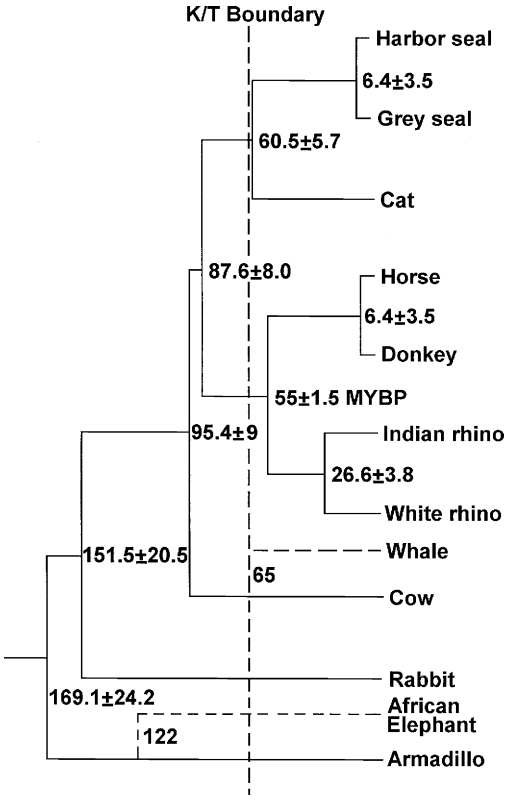


Figure 15 A clock-constrained maximum likelihood model mitochondrial DNA tree for mammals. The dates at each node represent millions of years before present (reproduced from Figure 1 in Waddell PJ, Cao Y, Hasegawa M, and Mindell D (1999) Assessing the Cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. *Systematic Biology* 48(1): 119–137, with permission from Oxford University Press).

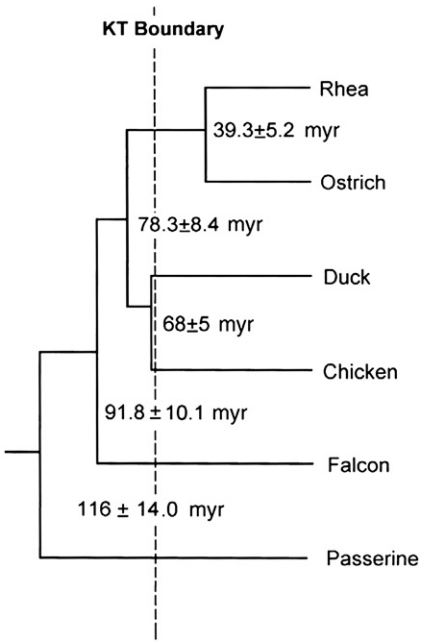


Figure 16 A clock-constrained maximum likelihood mitochondrial DNA tree for birds. The dates at each node represent millions of years before present (reproduced from Figure 2 in Waddell PJ, Cao Y, Hasegawa M, and Mindell D (1999) Assessing the Cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. *Systematic Biology* 48(1): 119–137, with permission from Oxford University Press).

and birds have transformed totally from their fishy ancestors (Patterson, 1999). Patterson showed these calculations indicate that hemoglobin in humans and elephants differ by about 18% as does that of the starling and ostrich. Also, humans and elephants diverged from the ancestral mammals at roughly the same time that ostrich and starling diverged from ancestral birds. In terms of amino acid differences, crocodiles differ from birds by about 33% and mammals by about 35%.

From the rough phylogenetic tree (**Figure 14**), Patterson calculated from the distance data with a corrected time scale relative rates of amino acid substitution (**Table 1**). Because there were more than 60% differences in the hemoglobin of sharks, as compared with humans, Patterson corrected for the error of multiple hits, because logically the real number of differences between the different taxa must be greater than those observed. The adjusted divergence times gave 140 myrs for the Platypus/mammal and crocodile/bird splits and about 70 myrs for the elephant/human and ostrich/starling divergence times.

It was from these kinds of results that questions of how natural selection could keep constant rates of divergence through time were asked. One such theory was the Red Queen hypothesis. It takes its name from *Alice in Wonderland* because every species has to keep on running just to keep in the same place. In other words, to compete and to survive all of nature's vicissitudes during the evolutionary changes from shark to human, and between any other pair of species, it is necessary for hemoglobin, and presumably all other features, to adapt at a constant rate (Patterson, 1999).

Rates of Molecular Evolution—DNA

With the advent of nuclear and plastid sequences in the late 1970s and a routine availability of sequences becoming available by the 1990s, it has become increasingly popular to infer divergence time estimates based on nucleotide sequences directly. Through a series of trials and errors it had become obvious by the mid-1990s that single point estimates of divergence times without any assessments of standard error were hardly worth calculating, and some greater degree of sophistication was needed (Waddell *et al.*, 1999). Recent calculations have attempted to improve results by incorporating differential substitution rates for molecular sequence change, a wider range of reliable calibration points based on fossils and other sources of information, and incorporation of standard error values to factors that can be quantified. Waddell *et al.* (1999) applied these techniques to estimate rates using maximum likelihood models and the origins of major lineages of mammals and birds using all available published mitochondrial DNA protein sequences (including unpublished ones for elephants and birds (**Figures 15** and **16**).

For mammals (**Figure 15**) the results show that the well-studied horse/rhinoceros split diverged about 55 million years before present. The splitting time among carnivores is confidently shown to be more than 50 million years, and splits within the placentals are relatively old, in the case of armadillo and other mammals 169 mybp but within 2 standard errors of other estimates. Waddell *et al*. reported also that the whale/cow divergence at 65 mybp may be much older than previously assumed. One of the interesting findings is that the sampled splits between the main groups of fereuungulates (carnivores, cetartiodactyls, perissodactyls, and pholidotes) took place prior to the critical Cretaceous/Tertiary boundary. The date of the vicariant separation of Africa and South America is coincident with the age of the cladogenetic event of the armadillo/elephant close relationship at around 122 mybp and could be a causal factor.

The data for birds (**Figure 16**) was considered more controversial by Waddell *et al*. than for mammals. This was because the deeper calibration points normally used by other researchers between either crocodiles or mammals and birds were considered controversial in their assumptions that rates of evolution in birds had not changed much. They used the fossil anseriform, *Prebyornis*, and placed it with the ducks (geese and swans) using a constraint of anseriform monophyly to have at least one calibration point within birds. Using this information together with more detailed studies of anserifom relationships gave an anseriform stem lineage date of 58–78 (midpoint 68) mybp. The divergence times are indicated in **Figure 16**. The interesting findings are that the passerines seem to have split off from the rest of the birds at a much earlier date than normally believed and that the divergence time of ratites (rhea and ostrich) is much younger than generally believed.

Conclusions

In an online *Encyclopaedia of Genetics*, Ernst Mayr recently considered cladogenesis and anagenesis as the two great phylogenetic processes of biology. He said that cladogenesis is the study of the origin and of the nature of the branching pattern of the phylogenetic tree. It deals with the various different methods by which the phylogenetic tree is reconstructed. Also, it includes the process of speciation, because every act of speciation adds a branch, no matter how short, to the phylogenetic tree. It is difficult to argue with such points of view because the definitions encapsulate so much of what we know and do in the modern synthesis. They are catchall statements with little or no precision in the contemporary scheme of things. The position taken here is that cladogenesis, and all of its related terminology, sits in some kind of neo-Darwinian time warp, which has become overtaken by huge developments in both systematics and evolutionary analysis. As Mayr points out, cladogenesis in a very particular sense refers to microevolution and speciation, both topics dealt with elsewhere. In macroevolution the interests are centered around the reconstruction of cladograms and in the discovery of monophyletic groups, topics considered elsewhere in this encyclopedia. In cladistics all effort is concentrated on the patterns of clade distribution and it is generally accepted that grades and branch lengths are of no value in phylogenetic classification but are of interest for determining rates of change. Cladogenesis, anagenesis, and stasis are general descriptors of studies of rates in evolution and extinction, as expressed here in relatively recent calculations of ages of taxa and deeper evolutionary splits. In conclusion, the importance of cladogenesis and related terms have become uncritical in the light of relatively recent studies in systematics; rather they are words in an arcane language largely consigned to the history books.

See also: Adaptation. Adaptive Radiation. Cladistics. Evolution, Theory of. Fossil Record. Phylogeny. Systematics, Overview; Taxonomy, Methods of

References

Bessey CE (1915) The phylogenetic taxonomy of flowering plants. *Annals Missouri Botanical Garden* 2: 109–164.

Darwin C (1859) *The Origin of Species by Means of Natural Selection*. London: Murray.

Eldredge N and Gould SJ (1972) Punctuated equilibria: An alternative model to phyletic gradualism. In: Schopf TJM (ed.) *Models in Paleobiology*, pp. 82–115. San Francisco: Freeman, Cooper & Co.

Haeckel E (1874) *Anthropogenie: Keimes- und Stammes-Geschicte des Menschen*, 1st ed. Leipzig: W. Engelmann.

Hennig W (1966) *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Huxley JS (1958) Evolutionary processes and taxonomy with special reference to grades. *Uppsala Univ. Årsskr* 6: 21–39.

Janchen E (1932) Entwurf eines Stammbaumes der Blütenpflanzen nach Richard Wettstein. *Öster. Bot. Zeit* 81(3): 161–166.

Mayr E (1974) Cladistic analysis or cladistic classification. *Z. Zool. Syst. Evolforsch* 12: 94–128.

Panchen AL (1992) *Classification and Evolution, and the Nature of Biology*. Cambridge: Cambridge University Press.

Patterson C (1999) *Evolution*, 2nd ed. London: The Natural History Museum.

Rensch B (1959) *Evolution Above the Species Level*. London: Methuen.

Simpson GG (1951) *Horses*. New York: Oxford University Press.

Simpson GG (1953) *The Major Features of Evolution*. New York: Columbia University Press.

Smith AB (1994) *Systematics and the Fossil Record; Documenting Evolutionary Patterns*. Oxford: Blackwell Scientific Publications.

Waddell PJ, Cao Y, Hasegawa M, and Mindell D (1999) Assessing the Cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. *Systematic Biology* 48(1): 119–137.

Climate Change and Ecology, Synergism of

Stephen H Schneider[†] **and Terry L Root,** Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Stephen H. Schneider, volume 1, pp 709–725, © 2001, Elsevier Inc.

Glossary

Climatic models Mathematical descriptions of the flows of energy, momentum, and materials around the atmosphere and oceans, typically run on large computers to stimulate the Earth's climate naturally and when disturbed by greenhouse gases or other so-called forcings or disturbances.

Global-change disturbances Alterations or disturbances to the natural conditions of the global atmosphere (e.g., greenhouse gases produced by human activities) or to enough regions (e.g., large-scale deforestation or the widespread introduction of exotic, invasive species in many places throughout the world) to have global impacts.

Global warming The additional heating of the Earth's climate from the incremental injection of greenhouse gases to the atmosphere from human activities such as deforestation or fossil fuel consumption.

Greenhouse gases Gases in the atmosphere, which selectively let more of the solar energy into and out of the atmosphere than they permit the transmission of long-wave infrared radiation. They contribute to the greenhouse effect, which traps radiant energy in the lower atmosphere and makes the Earth warmer.

Synergisms Interactions among many factors, which collectively may have a much larger or smaller effect than the sum of the effects of each factor acting independently.

"Climate Change and Ecology, Synergism of" is adapted from Schneider and Root (1998).

Synergisms of Climate Change and Ecology

The Earth's climate is vastly different now from what it was 100 million years ago (Mya) when dinosaurs roamed the planet and tropical plants thrived closer to the poles. It is different from what it was 20,000 years ago when ice sheets covered much of the Northern Hemisphere. Although the Earth's climate will surely continue to change, climatic changes in the distant past were driven by natural causes, such as variations in the Earth's orbit or the carbon dioxide (CO_2) content of the atmosphere. Today and in the future, climatic changes have had and will quite likely continue to have another source as well – human activities. Humans cannot directly rival the power of natural forces driving the climate, for example, the immense energy input to the Earth from the sun that powers the climate. We can, however, indirectly alter the natural flows of energy enough to create significant climatic changes. The best-known way people inadvertently modify climate is by enhancing the natural capacity of the atmosphere to trap radiant heat near the Earth's surface – the so-called greenhouse effect. This natural phenomenon allows solar energy to reach the Earth's surface and warm the climate. Gases such as water vapor and CO_2 trap a large fraction of long-wavelength radiant energy called terrestrial infrared radiation near the Earth's surface. This causes the natural greenhouse effect to be responsible for approximately 33 °C (60 °F) of surface warming. Thus, seemingly small human-induced changes to the natural greenhouse effect are typically projected to result in a global warming of 1.1–6.4 °C or higher by 2100 (IPCC, 2007a). It is highly likely that this could result in ecologically significant changes, which is why climatic considerations are fundamental in the discussion of the status and trends of ecological conditions.

We are already feeling the climatic effects of us having polluted the atmosphere with gases such as CO_2 (IPCC, 2007a). Many activities associated with human economic development have changed the physical and chemical environment in ways that modify natural resources. As these changes, such as burning fossil fuels that release CO_2 or using land for agriculture or urbanization that causes deforestation, have become large enough, significant global (worldwide) changes have occurred and are expected to continue. Such modifications can disturb the natural flows of energy in Earth's systems and thus can force significant climatic changes. These disturbances are also known as global-change forcings. Quantitative evaluations of the potential effect of human activities in creating global change are needed. Such evaluations are also central to potential policy responses to mitigate global changes (Schneider, 1997; IPCC, 2001c, 2007c).

Synergisms

One of the most potentially serious global-change problems is the synergistic or combined effects of habitat fragmentation and climate change. People fragment natural habitats for farmland, settlements, mines, roads, and other developmental activities. As climate changes, individual species of plants and animals are forced to adjust if they can, as they have in the past by tracking their preferred temperatures by moving into historically cooler locations. It seems unlikely that all of the species that moved into historically warmer locations to survive the Ice Age will be able to safely reach refuges due to the

[†]Deceased.

necessity of migrating across freeways, agricultural zones, industrial parks, military bases, and cities of the twenty-first century. An additional complication arises with the imposition of the direct effects of changes in CO_2, which can change terrestrial and marine primary productivity as well as alter the competitive relations among photosynthesizing organisms.

One representative instance of synergism is that of the Kirtland's warbler in northern Michigan; this species is restricted to a narrow area of jack pines that grow in sandy soil (Botkin *et al.*, 1991). Forest gap models of growth and decline of jack pines indicate that this tree will move north with warming, but the Kirtland's warbler will not likely survive the transition. This bird nests on the ground under relatively young pines, and the soil to the north is not generally sandy enough to allow sufficient drainage for successful fledging of young (Anich *et al.*, 2011). Consequently, global warming could well doom the warbler to extinction in 30–60 years. This potential for extinction indicates how the already high rate of extinctions throughout the world could be exacerbated by climatic changes occurring more rapidly than species can adapt (Pimm *et al.*, 2006; Pereira *et al.*, 2010; Barnosky *et al.*, 2011).

The synergism question raises a controversial management problem of anticipating global-change risk and responding by setting up interconnected nature reserves to ensure against some species becoming extinct as the climate changes. Alternatively, we could simply let the remnants of relatively immobile wildlife and natural plant communities remain in isolated reserves and parks as now exist, which could doom them to extinction in the isolated areas. If we do opt for more environmental safeguards by interconnecting our parks, the question then becomes how do we interconnect the nature reserves? Priorities must be set and money made available for constructing natural corridors through which species can travel. For example, elevated sections of highways may be needed to allow for migration routes, similar to what was done for the caribou in the Arctic when the Alaskan pipeline was built. To examine such questions as these in scientific detail, it is first necessary to make a multidisciplinary examination of the subcomponents of the various aspects of climatology and ecology. We begin with a background discussion of climatic history, processes, modeling, and validation, as a prelude to focus on ecological processes, that need to be examined in order to project possible synergisms between ecology and climate change.

Climate History: What Has Happened?

Scientists can reconstruct the cyclical expansion and contraction of polar caps and other ice masses using ice core samples taken from Greenland and Antarctica. When snow falls on high, cold glaciers the air trapped between snow grains is eventually transformed into air bubbles as the snow is compressed into ice from the weight of subsequent accumulations. The ratio of two oxygen molecules with different molecular weights (O^{16} and O^{18} isotopes) is a proxy record for the temperature conditions that existed when the snow was deposited. From this, scientists have been able to determine that the ice buildup from 90,000 to 20,000 years ago was quite variable and was followed by a (geologically speaking) fairly rapid 500- to 10,000-year transition to the (current) climatically very stable Holocene Period. The Holocene is the 10,000-year interglacial period in which human civilization developed and modern plant and animal distributions evolved to their current states. These ice cores also provide information on the presence of CO_2, an important greenhouse effect gas. Carbon dioxide was in much lower concentrations during cold periods than in interglacials (which is also similar for the greenhouse gas methane, CH_4). This implies an amplifying effect, or a positive feedback, because lower amounts of these gases during cold glacial periods means less trapped infrared radiative heat, thereby amplifying the cooling, and vice versa, during warm interglacial periods. The ice cores also show that concentrations of CO_2 and CH_4 and temperature were remarkably constant for about the past 10,000 years (but before AD 1700), particularly when compared with the longer record. That relative constancy in chemical composition of the greenhouse gases held until the industrial age during the past two centuries.

The transition from Ice Age to the Holocene took 5000–10,000 years, during which time the average global temperature increased 5–7 °C and the sea-level rose 100 m. Thus, we estimate that natural rates of warming on a sustained global basis are about 0.5–1 °C per thousand years. Such changes were large enough to have radically influenced where species lived (Graham, 1992) and to have potentially contributed to the well-known extinctions of woolly mammoths, sabertooth cats, and enormous salamanders.

A large interdisciplinary team of scientists, including ecologists, palynologists (scientists who study pollen), paleontologists (scientists who study prehistoric life, especially fossils), climatologists, and geologists, formed a research consortium (Cooperative Holocene Mapping Project, 1988; Wright *et al.*, 1993) to study the dramatic ecological changes accompanying the transition from Ice Age to the recent interglacial period. One group of these researchers used a variety of proxy indicators to reconstruct vegetation patterns over the past 18,000 years for a significant fraction of the Earth's land areas. In particular, cores of fossil pollen from dozens of sites throughout North America clearly showed how boreal tree pollen, now the dominant pollen type in the boreal zone in central Canada, was a prime pollen type during the last Ice Age (15,000–20,000 years ago) in what are now the mixed hardwood and Corn Belt regions of the US During the last Ice Age, most of Canada was under ice; pollen cores indicate that as the ice receded, boreal trees moved northward chasing the ice cap. One interpretation of this information was that biological communities moved intact with a changing climate. In fact, Darwin (1859) asserted as much:

> As the arctic forms moved first southward and afterward backward to the north, in unison with the changing climate, they will not have been exposed during their long migrations to any great diversity of temperature; and as they all migrated in a body together, their mutual relations will not have been much disturbed. Hence, in accordance with the principles inculcated in this volume, these forms will not have been liable to much modification.

If this were true, the principal ecological concern over the prospect of future climate change would be that human land-use patterns might block what had previously been the

free-ranging movement of natural communities in response to climate change. The Cooperative Holocene Mapping Project, however, incorporated multiple pollen types into its analyses, including not only boreal species but also herbs and more arid (xeric) species as well as oaks and other mesic species. They discovered that during the transition from the last Ice Age to the present interglacial, nearly all species moved north, as expected. During a significant portion of the transition period, however, the distribution of pollen types provided no analog associations to today's vegetation communities (Overpeck *et al.*, 1992), that is, whereas all species moved, they did so individual by individual, not as community groups. Consequently, the groupings of species during the transition period were often dissimilar to those present today (**Figure 1**). The relevance of this conclusion is that in the future ecotypes will not necessarily move as a unit as climate changes (assuming there is enough time and space for such a migration). Past vegetation response to climatic change at an average rate of 1 °C per millennium indicates that credible predictions of vegetation changes from comparable or even more rapid climatic changes projected from human activities cannot neglect transient (i.e., time-evolving) dynamics of the ecological system. Furthermore, caution should be exercised against relying too much on past conditions in forecasting future patterns resulting from global-change forcings because the forecasted global average rate of temperature increase far exceeds those rates typical of the past 120,000 years.

Future climates may not only be quite different from more recent previous climates but may also be different from those inferred from paleoclimatic data and from those to which some existing species are evolutionarily adapted. Therefore, possible future changes inferred from past changes can be taken only as a guide or a means to verify aspects of the forecasts of models of climate or ecosystem dynamics (Crowley, 1993; Schneider, 1993a). Such verification exercises may provide more credible forecasts of the effects of climatic change on animals.

Forces of Climate Change

The two basic categories of causes of climatic change are external and internal. However, these terms are defined relative to the focus of study; stating which components are external or internal to the climatic system depends on the time period and spatial scale being examined, as well as the phenomena being considered. External causes of climate change do not have to be physically external to the Earth (such as the Sun) but do occur outside of the climate system. If our focus is on atmospheric change on a 1-week time scale (i.e., the weather), the oceans, land surfaces, biota, and human activities that produce CO_2 are all external (i.e., they are not influenced much by the atmosphere in such a short time). If our focus is on 100,000-year Ice Age interglacial cycles, however, the oceans, ice sheets, and biota are all part of the internal climatic system and vary as an integral part of the Earth's environmental systems. On this longer scale we must also include as part of our internal system the solid earth, which really is not solid but viscous and elastic.

Fluctuations in heat radiated by the Sun – perhaps related to varying sunspots – are external to the climate system. Influences of the gravitational tugs of other planets on the Earth's orbit are also external. Human-caused changes in the Earth's climate could not perceptibly alter either one of these cycles.

Carbon dioxide and methane levels rise and fall with Ice Age cycles, which are clearly internal on a 10,000-year time scale. However, on a 20-year scale these greenhouse gases become largely an external cause of climatic change, because small changes in climate have little feedback effect on, for example, humans burning fossil fuel.

Changes in the character of the land surface, such as those caused by human activities, are largely external. If vegetation cover changes because of climatic change, however, land surface change then becomes internal because changes in plant cover can influence the climate by changing albedo (reflectivity to sunlight), evapotranspiration, surface roughness, and relative humidity (Henderson-Sellers *et al.*, 1993).

Snow and ice are important factors in climatic change because they have higher albedo (reflectivity) than warmer surfaces and, in the instance of sea ice, can inhibit transfer of heat and moisture between air and wet surfaces. Salinity, which affects changes in sea ice and in the density of seawater, which helps control where ocean waters sink, may also be an internal cause of climatic variation. The sinking and upwelling of ocean waters are biologically significant because the upwelling waters are often nutrient-rich.

Unusual patterns of ocean surface temperature, such as the El Niño, demonstrate the importance of internally caused climatic fluctuations because the atmospheric circulation can change simultaneously with ocean surface temperatures. When the atmosphere rubs on the ocean, the ocean responds by modifying its motions and temperature pattern, which forces the atmosphere to adjust, which changes the winds, which changes the way the atmosphere rubs on the ocean, and so forth (Trenberth, 1992). As a result, air and water interact internally in this coupled system like blobs of gelatin of different size and stiffness connected by elastic bands or springs, all interacting with one another while also being pushed from the outside (by solar, volcanic, or human-caused change).

Climate Change Projections: What Will Happen?

To predict the ecologically significant ways the climate might change, one must specify what people do that modifies how energy is exchanged among the atmosphere, land surface, and space because such energy flows are the driving forces behind climate. Air pollution is an example of such a so-called societal forcing of the climate system. Estimating societal forcing involves forecasting a plausible set of human (or societal) activities affecting pollution from now until 2100. The next step is to estimate the response of the various components of the earth system to such societal forcings. The earth system consists of the following interacting subcomponents: atmosphere, oceans, cryosphere (snow, seasonal ice, and glaciers), and land-surface (biota and soils) systems.

Research in the field and in laboratories provides an understanding about various processes affecting the subcomponents

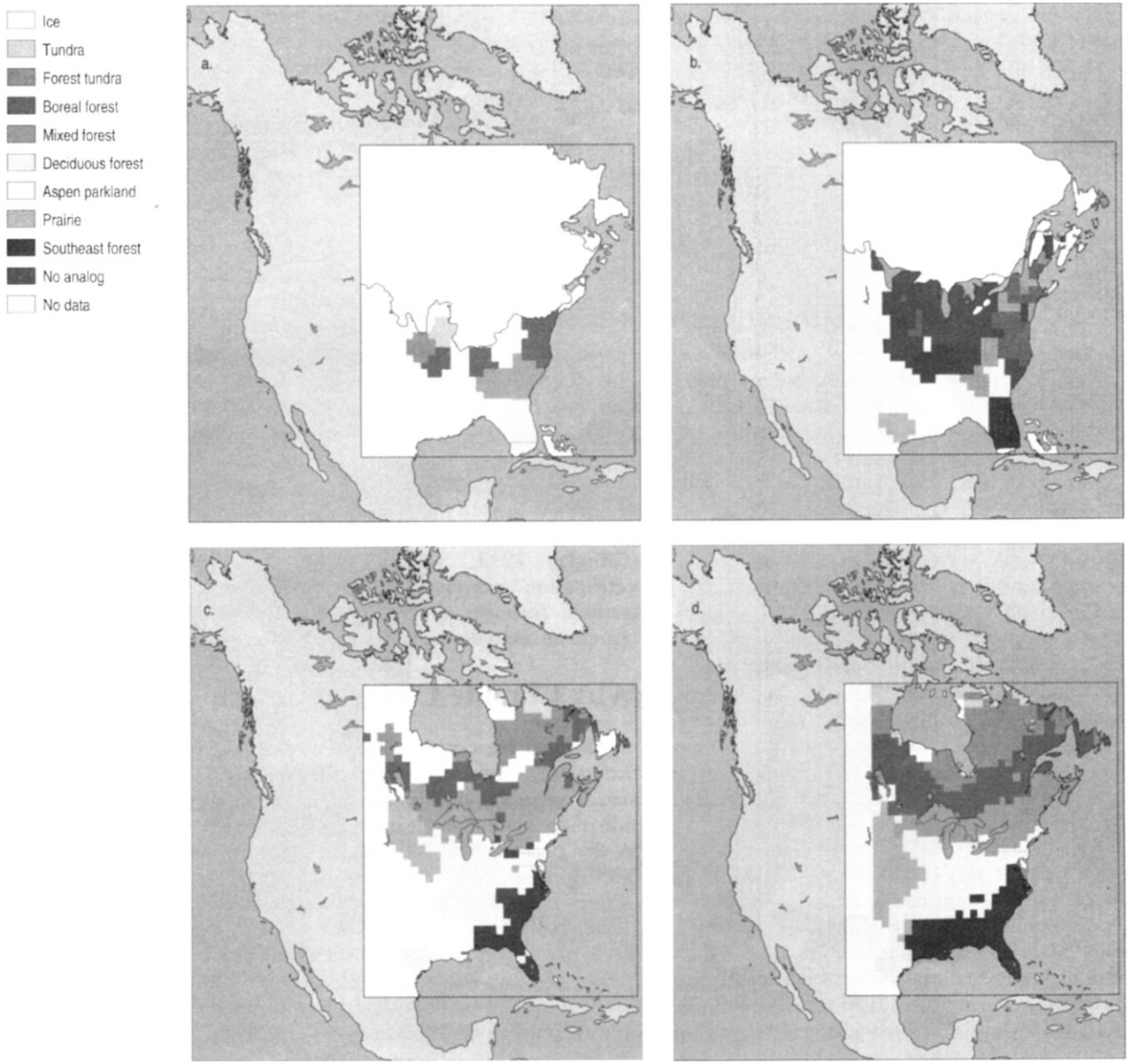


Figure 1 Selected paleovegetation maps reconstructed using the method of modern analogs and more than 13,000 samples of fossils and modern pollen: (a) 18,000 BP, (b) 12,000 BP, (c) 6000 BP, and (d) modern. No vegetation is mapped in areas without fossil pollen data, and no analog refers to vegetation without any modern analog. Reproduced from Overpeck JT, Webb RS, and Webb III T (1992) Mapping eastern North American vegetation change over the past 18,000 years: No analogs and the future. *Geology* 20: 1071–1074, with permission from Geological Society of America Bulletin.

of the earth system. This understanding can be put into mathematical expressions that, when combined, form a model of the behavior of particular components of the earth system. In practice, models of the atmosphere are connected to models of the oceans, ice, biota, and land surfaces to simulate the consequences of a scenario of societal forcing on climate and ecosystems. Controversy arises because both the societal forcing that will actually occur and the scientific knowledge of each subsystem are incomplete. Because models cannot be perfect replicas of the actual natural system, scientists must expend considerable efforts to test their models against the expanding base of field and laboratory data. This not only allows them to assess the robustness of current simulations but it also reveals improvements for the next generation of models.

Elements of Global Warming Forecasts

The societal driving forces behind global-warming scenarios are projections of population, consumption, land use, and technology. Typical twenty-first-century projections for human population size and affluence for less highly developed countries and more highly developed countries show drastic increases in population and wealth. When these factors are multiplied by the amount of energy used to produce a unit of economic product (the so-called energy intensity) and the amount of CO_2 emitted per unit of energy (the technology factor called carbon intensity), carbon emissions are predicted to increase several fold throughout the twenty-first century. It is very difficult to make such projections credibly; therefore, analysts disagree by as much as a factor of 10 about how much CO_2 will be emitted by 2100 (IPCC, 2007a). Specific scenarios are debatable because the amount of carbon emitted through human activities will significantly depend on what kinds of energy systems will be developed and deployed globally and on what the standards of living will be during the next several decades, not to mention population growth.

To turn estimates of CO_2 emissions into estimates of CO_2 concentrations in the atmosphere – the variable needed to calculate potential climate changes – one must estimate what fraction of CO_2 emitted will remain in the atmosphere. This airborne fraction is most simply estimated at 45% because the amount of CO_2 buildup in the atmosphere each year is about half the fossil fuel-injected CO_2 with the other part being absorbed by ocean and biosphere. The atmospheric concentration of CO_2, however, should be computed using carbon-cycles models, which account for the time-evolving amounts of carbon in vegetation, soils, and oceanic and atmospheric subcomponents (IPCC, 2007a). The estimated CO_2 concentration can then be fed into computerized climatic models to estimate its effects on climate.

Climate prediction, like most other forecasts involving complex systems, generally involves educated guesses. Those attempting to determine the future behavior of the climate system from knowledge of its past behavior and present state basically can take two approaches. One approach, the empirical–statistical, uses statistical methods such as regression equations that connect past and present observations statistically to obtain the most probable extrapolation. The second approach, usually called climate modeling, focuses on first principles, which are equations representing laws believed to describe the physical, chemical, and biological processes governing climate. Because the statistical approach depends on historical data, it is obviously limited to predicting climates that have been observed or are caused by processes appropriately represented in the past conditions. The statistical method cannot reliably answer questions such as what would happen if atmospheric CO_2 increased at rates much faster than in the known past. Thus, the more promising approach to climate prediction for conditions or forcings different from the historic or ancient past is climate modeling. A significant component of empirical–statistical information, though, is often embedded into these models (Root and Schneider, 1995). This often makes modelers uncomfortable about the validity of predictions of such models on unusual or unprecedented situations unless a great deal of effort is expended to test the models against current and paleoclimatic baseline data.

Climate models vary in their spatial resolution – that is, the number of dimensions they simulate and the spatial detail they include. The simplest model calculates only the average temperature of the Earth, independent of the average greenhouse properties of the atmosphere. Such a model is zero-dimensional: It reduces the real temperature distribution on the Earth to a single point, a global average. In contrast, three-dimensional climate models produce the variation of temperature with latitude, longitude, and altitude. The most complex atmospheric models, the general circulation models (GCMs), predict the time evolution of temperature plus humidity, wind, soil moisture, sea ice, and other variables through three dimensions in space.

Verifying Climate Forecasts

The most perplexing question about climate models is whether they can be trusted as a reliable basis for altering social policies, such as those governing CO_2 emissions or the shape and location of wildlife reserves. Even though these models are fraught with uncertainties, several methods are available for verification tests. Although no method is sufficient by itself, several methods together can provide significant, albeit circumstantial, evidence of a forecast's credibility.

The first validation testing method involves checking the model's ability to simulate the current climate. The seasonal cycle is a good test because temperature changes in a seasonal cycle are larger on a hemispheric average than the change from an Ice Age to an interglacial period (i.e., 15 °C seasonal range in the Northern Hemisphere versus 5–7 °C glacial–interglacial cycle). GCMs map the seasonal cycle well. This supports the scientific consensus about the plausibility of global warming of several degrees in the twenty-first century. The seasonal test, however, does not indicate how well a model simulates slow processes such as changes in deep-ocean circulation, ice cover, forests, or soil carbon storage, which may have important effects on the decade- to century-long timescales over which atmospheric CO_2 is expected to double.

A second verification technique involves isolating individual physical components of the model and testing them

against actual data. A reasonable model should reproduce the flow of thermal energy among the atmosphere, the surface, and space with no more than about a 10% error. Together, these energy flows make up the well-established natural greenhouse effect on Earth and constitute a formidable and necessary test for all models. A model's performance in simulating these energy flows is an example of physical validation of model components.

A third validation method involves the model's ability to reproduce the diverse climates of the past. This method is aided by recorded instrumental observations made during the past few centuries and paleo-records that serve as a proxy for climatic conditions of the ancient earth, or even include testing the models' ability to simulate climates of other planets (Kasting *et al.*, 1988). Paleoclimatic simulations of the Mesozoic (Age of the Dinosaurs), glacial–interglacial cycles, or other extreme past climates help scientists understand the coevolution of the Earth's climate and living things (Schneider and Londer, 1984). As verification tests of climate models, they are also crucial in predicting future climates and changes in biological systems.

Using these techniques, much has been learned from examining the global climatic trends of the past century. Northern Hemisphere surface temperature increased more during the twentieth century than during any other century in the past 1000 years. The 1990s was the warmest decade of the millennium, and the first decade of the twenty-first century has been even warmer, with nine of the 10 warmest years on record occurring in the twenty-first century. These data are consistent with an enhanced greenhouse effect signal that might be anticipated from the greenhouse gas injections during the past 250 years, which saw a 36% increase in CO_2, a $\sim$160% increase in CH_4, and the introduction of human-generated heat-trapping chemicals such as chlorofluorocarbons and halons. Industrial activities since the 1950s have contributed to the increase of sulfur dioxide and other aerosol particles into the atmosphere, which reduce surface temperature by reflecting sunlight back to space. Although such cooling effects may have counteracted global warming by only several tenths of a degree, the hazes occur regionally and could be producing ecologically significant, unexpected regional changes in climate patterns (Schneider, 1994).

Although the 0.7 ± 0.2 °C surface warming in the twentieth century is consistent with the human-induced greenhouse gas buildup, it is possible that the 0.7 °C warming trend was wholly natural and that there was little or no contribution from the buildup of greenhouse gases (for estimates of the probability of human-induced global warming amounts, see Morgan and Keith (1995)). However, it is also not possible to rule out the counterfactual that, independent of the enhanced greenhouse effect due to human activity, there was a natural cooling fluctuation taking place during the twentieth century. If so, the world would then have warmed up much more than observed had we not had a fortuitous natural cooling trend. One could even speculate that the dramatic temperatures since the 1970s with global high temperature records reflect the termination of a natural cooling trend combined with the rapid establishment of the expected enhanced greenhouse effect.

Studies (Santer *et al.*, 1996) suggest that when aerosols and greenhouse gas forcings are combined, climate models more closely match 30 years of observations. Nevertheless, there is wide variability in predictions for a doubling of CO_2: Temperature changes as low as a 0.5 °C warming to as high as a 5.0 °C warming (Wigley and Raper, 1991) are all consistent with current observations. Several reasons exist for such a wide range of uncertainty: difficulty in knowing how to model delays in global warming because of the large heat capacity of the oceans; not knowing what other global-change forcings may have opposed warming, for example, sulfate aerosols from burning high-sulfur coal and oil or undetectable changes in the Sun's light output before 1980; and large, unknown, internal natural climatic fluctuations. As mentioned previously, though, the ecologically important forecasts of time-evolving regional climatic changes are much less credible and require that ecologists use many alternative scenarios of possible climatic changes.

In summary, the warming of the globe is unequivocal and this is very likely due to human activities as its primary driver (IPCC, 2007a). This conclusion has emerged clearly from studies done since 1995 when the Intergovernmental Panel on Climate Change (1996a) carefully weighed the uncertainties and concluded that "the balance of evidence suggests that there is a discernible human influence on global climate."

Relevance of Climate Modeling to Regional Climate Change and Ecosystem Studies

Scientists who estimate the future climatic changes that are relevant to ecosystems have focused on the GCMs that attempt to represent mathematically the complex physical and chemical interactions among the atmosphere, oceans, ice, biota, and land. As these models have evolved, increasingly more information has become available, and more comprehensive simulations have been performed. Nevertheless, the complexities of the real climate system still vastly exceed the GCMs and the capabilities of even the most advanced computers (Intergovernmental Panel on Climate Change, 1996a, 2007a). Simulating a 50-year projection of climate with a grid size of 100 km on a side and 20 vertical layers – about 2,500,000 grid points – takes several months on a supercomputer. This level of resolution, however, cannot resolve the Sierra Nevada of California and the Rocky Mountains as separate mountain chains. Refining the resolution to 50-km grid squares would so dramatically increase the number of computations that it would take many more months of computer time to simulate weather statistics for 1 year.

Even the highest-resolution, three-dimensional general circulation model will not have a grid with nodes much less than 100-km apart within the foreseeable future; individual clouds and most ecological research (to say nothing of cloud droplets) occur on scales far smaller. Therefore, GCMs will not be able to resolve the local or regional details of weather affecting most local biological communities or the importance of regional effects of hills, coastlines, lakes, vegetation boundaries, and heterogeneous soil (Root and Schneider, 1993). Nonetheless, it is important to have climatic forecasts and ecological-response analyses on the same physical scales.

What is most needed to evaluate potential biological effects of temperature change is a regional projection of climatic

changes that can be applied to ecosystems at a regional or local scale. Analyses of large, prehistoric climatic changes (Schneider, 1987; Cooperative Holocene Mapping Project, 1988) and historical weather analogs (Shabalova and Konnen, 1995) provide some insights into such changes. Historical weather analogs, however, since they are empirically and statistically based, rely on climatic cause-and-effect processes that probably differ from those that will be driven by future greenhouse gas radiative effects (Schneider, 1984; Crowley, 1993). Consequently, ecologists turn to climatic models to produce forecasts of regional climatic changes for the decades ahead. How credible are such forecasts?

Regional Changes

Although the consensus among researchers about the plausibility of significant human-induced global climatic change is growing, no assessment (Intergovernmental Panel on Climate Change, 1996a) has suggested the existence of a strong consensus about how that global climatic change might be distributed regionally. The world is undergoing a steady increase in greenhouse gas forcing. Because that increase is heating the Earth in a reasonably uniform way, one might expect a uniform global response, although this is far from likely. For example, the centers of continents have relatively low heat-retaining capacity, and the temperatures there would move relatively rapidly toward whatever their new equilibrium climate would be compared with the centers of oceans, which have high heat-retaining capacity. Tropical oceans, though, have a thin (about 50 m) mixed layer that interacts primarily with the atmosphere. It takes about 10 years for that mixed layer to substantially change its temperature, which is still much slower than the response time of the middle of the continents but is much faster than that of the oceans closer to the poles. At high latitudes, in places such as the Weddell or Norwegian seas, waters can mix down to the bottom of the ocean, thereby continuously bringing up cold water and creating a deep-water column for which a century or more is required to substantially change its temperature.

During the transient phase of climate change over this century, therefore, one would expect the middle of continents, the middle of oceans, and the polar and subpolar oceans all to change toward their new equilibrium temperatures at different rates. Thus, the temperature differences from land to sea and equator to pole will evolve over time, which in turn implies that the transient character of regional climatic changes could be very different from the expected long-term equilibrium (Schneider and Thompson, 1981; Washington and Meehl, 1986). This does not imply that transient regional changes are inherently unpredictable, only that currently they are very difficult to predict credibly.

Even more uncertain than regional averages, but perhaps more important to long-term ecosystem responses, are estimates of climatic variability during the transition to a new equilibrium, particularly at the regional scale. These include estimates of events such as the frequency and magnitude of severe storms, enhanced heat waves, temperature extremes, sea-level rises (Titus and Narayanan, 1996), and reduced frost probabilities (Rind *et al.*, 1989). For example, there is a physical principle that evaporation increases dramatically as surface-water temperature increases. Because hurricanes are powered by evaporation and condensation of water, if all other factors are unchanged, the intensity of hurricanes and the length of the hurricane season could increase with warming of the oceans (Emanuel, 2003, 2006; Trenberth, 2005). Such changes would significantly affect susceptible terrestrial and marine ecosystems.

Downscaling Climate Predictions to Regional Effects

Empirical Mapping Techniques

Techniques exist that can translate the output of climate models so that it is closer to that of most ecological scales. One method that uses actual climatic data at both large and small scales can help provide maps that may allow small-scale analysis of large-scale climate change scenarios. For example, the Sierra Nevada of California or the Cascades in the northwestern US are north–south mountain chains whose east–west dimensions are smaller than the grid size of a typical general circulation model. In the actual climate system, onshore winds on the Pacific coast would produce cool upslope and rainy conditions on the western slope and a high probability of warmer and drier conditions associated with that flow pattern on the downslope or eastern slope.

Regional maps have been generated for Oregon (Gates, 1985) in which a high-resolution network of meteorological stations was used to plot temperature and precipitation isopleths based on observed climatic fluctuations at large (e.g., state-sized) scales. These maps show that the dominant mode of variation for this area is warm and dry on one side of the mountains and cold and wet on the other side. Although this empirical mapping technique seems appropriate for translating low-resolution, grid-scale climate-model forecasts to local applications, a strong caveat must be offered, that is, the processes in the climate system that give rise to internal variability or natural fluctuations are not necessarily the same processes that would give rise to local deviations from large-scale patterns if the climate change were driven by external forces rather than an internal variation of the system. For example, the Oregon maps would indicate that if the grid box average temperature were warmer on the eastern slope, then it should be cooler and wetter on the western slope. This condition is the most probable regional situation for today's naturally fluctuating climate. However, if 50 years from now the warming on the eastern slope were, for example, a result of doubled atmosphere CO_2 causing an enhanced downward infrared radiative heating, then both eastern and western slopes would probably experience warming. Although the degree of warming and associated precipitation changes would not necessarily be uniform, an entirely different climatic change pattern would probably occur as opposed to that obtained from the empirical mapping technique if one used the naturally varying weather conditions existing today rather than the anthropogenically forced conditions of the twenty-first century (Schneider, 1993b).

Therefore, techniques to downscale climate forecasts that use current distributions of environmental variables at local scales and correlate them with current large-scale regional

patterns may not necessarily provide a good guideline about how large-scale patterns would be distributed regionally. This is especially true if there are changes in prevailing winds, seasonality, or cloud cover. The reason is that the causes of the future change may be physically or biologically different from the causes of the historical fluctuations that led to the empirical maps in the first place. This caveat warns scientists to use caution before adopting such empirical techniques for global-change applications.

There are practical limits to downscaling, especially using global techniques. Downscaling to a weather station, assuming that the weather-station data are accurate (a dubious assumption in many parts of the world and the above caveats are heeded), provides useful information at the level of the point. Gridded weather data are often used when weather-station data are not available. These grids are typically $0.5^\circ \times 0.5^\circ$ of latitude and longitude in size, but the uncertainties in gridding areas with sparse data and with large elevational changes must be recognized. The standard gridded data set used in many studies is the CRU TS data (University of East Anglia Climate Research Unit). In some areas with large numbers of weather stations (e.g., US, Europe, and Australia) smaller grids (approximately 10 km^2) have been developed (e.g., VEMAP, OzClim). Although it is theoretically possible to downscale to such small scales and even smaller ones (1 km^2 based on digital elevation model data), caution must be taken when using any of these gridded data. At a large scale (50 km^2) the differences in slopes of mountains are rarely picked up. At 1 km^2, elevation is important, but slopes, aspect and, especially wind direction and water transport, are essential when calculating lapse rate to estimate temperature and vital for precipitation estimates.

Regional-Scale Models with GCM Inputs

Other techniques can still translate large-scale patterns to smaller scales, but these techniques are based on known processes rather than empirical maps for today's conditions. One such technique is to drive a high-resolution, process-based model for a limited region with the large-scale patterns produced by a GCM. In essence, this approach uses a mesoscale model (i.e., 10- to 50-km^2 grid cells) based on physical laws to solve the problem of translating GCM grid-scale averages into a finer scale mesh much closer to the dimensions of most ecological applications. Of course, even this mesoscale grid will still be too coarse to assess many impacts, necessitating further downscaling techniques. Neither are the problems of GCMs entirely eliminated by mesoscale grids because they are still bigger than individual clouds or trees. However, such methods do bring climate-model scales and ecological-response scales much closer.

Examples of Ecological Responses to Climate Changes

Bringing climatic forecasts to ecological applications at local and regional scales is one way to bridge the scale gap across ecological and climatological studies (IPCC, 2001b, 2007b). Ecologists, however, have also analyzed data and constructed models that apply over large scales, including the size of climatic model grids. A long tradition in ecology has associated the occurrence of vegetation types or the range limits of different species with physical factors such as temperature, soil moisture, land–sea boundaries, or elevation. Biogeography is the field that deals with such associations, and its results have been applied to estimate the large-scale ecological response to climate change.

Predicting Vegetation Responses to Climate Change

The Holdridge (1967) life zone classification assigns biomes (e.g., tundra, grassland, desert, or tropical moist forest) according to two measurable variables – temperature and precipitation. Other more complicated large-scale formulas have been developed to predict vegetation patterns from a combination of large-scale predictors (e.g., temperature, soil moisture, or solar radiation); vegetation modeled includes individual species (Davis and Zabinski, 1992), limited groups of vegetation types (Schimel *et al.*, 2001), or biomes (Steffen *et al.*, 1996). These kinds of models predict vegetation patterns that represent the gross features of actual vegetation patterns, which are an incentive to use them to predict vegetation change with changing climate, but they have some serious drawbacks as well, that is, they are typically static, not time-evolving dynamic simulations, and thus cannot capture the transient sequence of changes that would take place in reality. In addition, such static biome models occasionally make "commission errors;" that is, they typically predict vegetation types to occur in certain zones where climate would indeed permit such vegetation, but other factors such as soils, topography, or disturbances such as fire or land-use change actually preclude it. However, newer dynamic general vegetation models (DGVMs) have started to include disturbance regimes. These DGVMs, which work at the level of the plant functional type (PFT, similar to biomes), have been used to look at potential changes in net primary productivity as well as changes in plant groups over time. One of the problems with using DGVMs in a predictive sense is that they are typically on or off – that is, they show what the current PFT might be and what the future PFT might be but they lack a transition state. This can be misleading for policymakers as it misses the critical information both as to the longevity of some of these PFTs in the absence of disturbance (i.e., reproduction stops but individuals persist for the lifetime of the organisms, potentially several hundred years) as well as the slow rate of dispersal for many plant species. Thus a transition state, matching neither the current nor future PFT is likely. This can be overcome, to some extent by looking at individual species but the large number of species involved, until recently, has made it difficult to look at changes at this scale. The access of large clusters of computers and global data sets, however, has meant that large-number-of-species models can now be run. For example, one particular effort called the Wallace Initiative has modeled almost 50,000 terrestrial plant and animal species. Furthermore, local patterns may influence vegetation dynamics at scales not captured in some simulations, and seed germination and dispersal mechanisms are also either not explicitly simulated or simulated only crudely with such models. It is remarkable that they are still able to produce

generalized maps of vegetation types that do indeed resemble current or even paleoclimatic patterns in a broad sense, but their details do not provide confident projections for future vegetation states. Fortunately, progress is being made to include some of the deficiencies mentioned previously, and the so-called dynamical global vegetation models are being developed to treat the transient nature of vegetation change that would likely accompany climatic change.

Predicting Animal Responses to Climate Change

Scientists of the US Geological Survey, in cooperation with Canadian scientists, conduct the annual North American Breeding Bird Survey, which provides distribution and abundance information for birds throughout the US and Canada. From these data, collected by volunteers under strict guidance from the US Geological Survey, shifts in bird ranges and abundances can be examined. Because these censuses were begun in the 1960s, these data can provide a wealth of baseline information. Price (1995) used these data to examine the birds that breed in the Great Plains. Using the present-day ranges and abundances for each of the species (**Figure 2(a)**), Price derived large-scale, empirical–statistical models based on various climate variables (e.g., maximum temperature in the hottest month and total precipitation in the wettest month) that provided estimates of the current bird ranges and abundances (**Figure 2(b)**). Then, using a general circulation model to forecast how doubling of CO_2 would affect the climate variables in the models, Price applied the statistical models to predict the possible shape and location of the birds' ranges and abundances (**Figure 2(c)**).

Significant changes were found for nearly all birds examined. The ranges of most species moved north, up mountain slopes, or both. The empirical models assume that these species are capable of moving into these more northerly areas, provided habitat is available and no major barriers exist. Such shifting of ranges and abundances could cause local extinctions in the more southern portions of the birds' ranges, and, if movement to the north is impossible, extinctions of entire species could occur. We must bear in mind, however, that this empirical–statistical technique, which associates large-scale patterns of bird ranges with large-scale patterns of climate, does not explicitly represent the physical and biological mechanisms that could lead to changes in birds' ranges. Therefore, the detailed maps should be viewed only as illustrative of the potential for very significant shifts with different possible doubled CO_2 climate change scenarios. More refined techniques that also attempt to include actual mechanisms for ecological changes are discussed later. Many think that birds may be able to keep up with rate of climate change because they can fly. However, rates of natal dispersal should not be confused with the expansion capability of populations. A review of the literature on dispersal rates (both historical and current) suggests that the range dispersal rate of birds and mammals is about the same – approximately 1–4 km per year.

Reptiles and amphibians are different from birds in many ways that are important to our discussion. First, because amphibians and reptiles are ectotherms – meaning that their body temperatures adjust to the ambient temperature and radiation of the environment – they must avoid environments in which temperatures are too cold or too hot. Second, amphibians must live near water not only because the reproductive part of their life cycle is dependent on water but also because they must keep their skin moist because they "breathe" through their skin as well as their lungs. Third, neither reptiles nor amphibians are able to disperse as easily as birds because they must crawl rather than fly, and the habitat through which they crawl must not be too dry or otherwise impassable (e.g., high mountains or superhighways).

As the climate changes, the character of extreme weather events, such as cold snaps and droughts, will also change (NOAA), necessitating relatively rapid habitat changes for most animals. The role of extreme events and climate variability is still not adequately taken into account in models of species ranges. Although the climate change extinction-rate estimates are often criticized, it needs to be borne in mind that these are based on MEAN climates. Even the variables taken as extreme are normally not true extremes, but are averages over more than a month. Mortality in species is often measured in x number of days exceeding a certain threshold (the same as for coral bleaching). Thus, because the return rates of these threshold events will increase as the climate warms, then extinctions, or at least local extirpations, may very well happen sooner than projected by changes in the mean.

In general, animals most likely to be affected earliest by climatic change are those for which populations are fairly small and limited to isolated habitat islands. As a result of human-generated landscape changes, many reptiles now fall into this category, as do many other animals. Indeed, temperature-dependent sex-determined species are especially likely to suffer from extreme sex ratio biases, and therefore their sensitivity to rapid climate change appears potentially more severe than that of most other animals.

There are estimates that many small mammals living near isolated mountaintops (which are essentially habitat islands) in the Great Basin would become extinct given typical global-change scenarios (Beever *et al.*, 2011). Recent studies of small mammals in Yellowstone National Park show that statistically significant changes in both abundances and physical sizes of some species occurred with historical climate variations (which were much smaller than most projected climate changes for the next century), but there appear to have been no simultaneous genetic changes (Hadly *et al.*, 2004). Therefore, it is likely that climate change in the twenty-first century could cause substantial alteration to biotic communities, even in protected habitats such as Yellowstone National Park.

Top-Down Approaches

The biogeographic approach previously summarized is an example of a top-down technique (e.g., Holdridge life zone classification), in which data on abundances or range limits of vegetation types or biomes are overlain on data of large-scale environmental factors such as temperature or precipitation. When associations among large-scale biological and climatic patterns are revealed, biogeographic rules expressing these correlations graphically or mathematically can be used to forecast changes in vegetation driven by given climate changes.

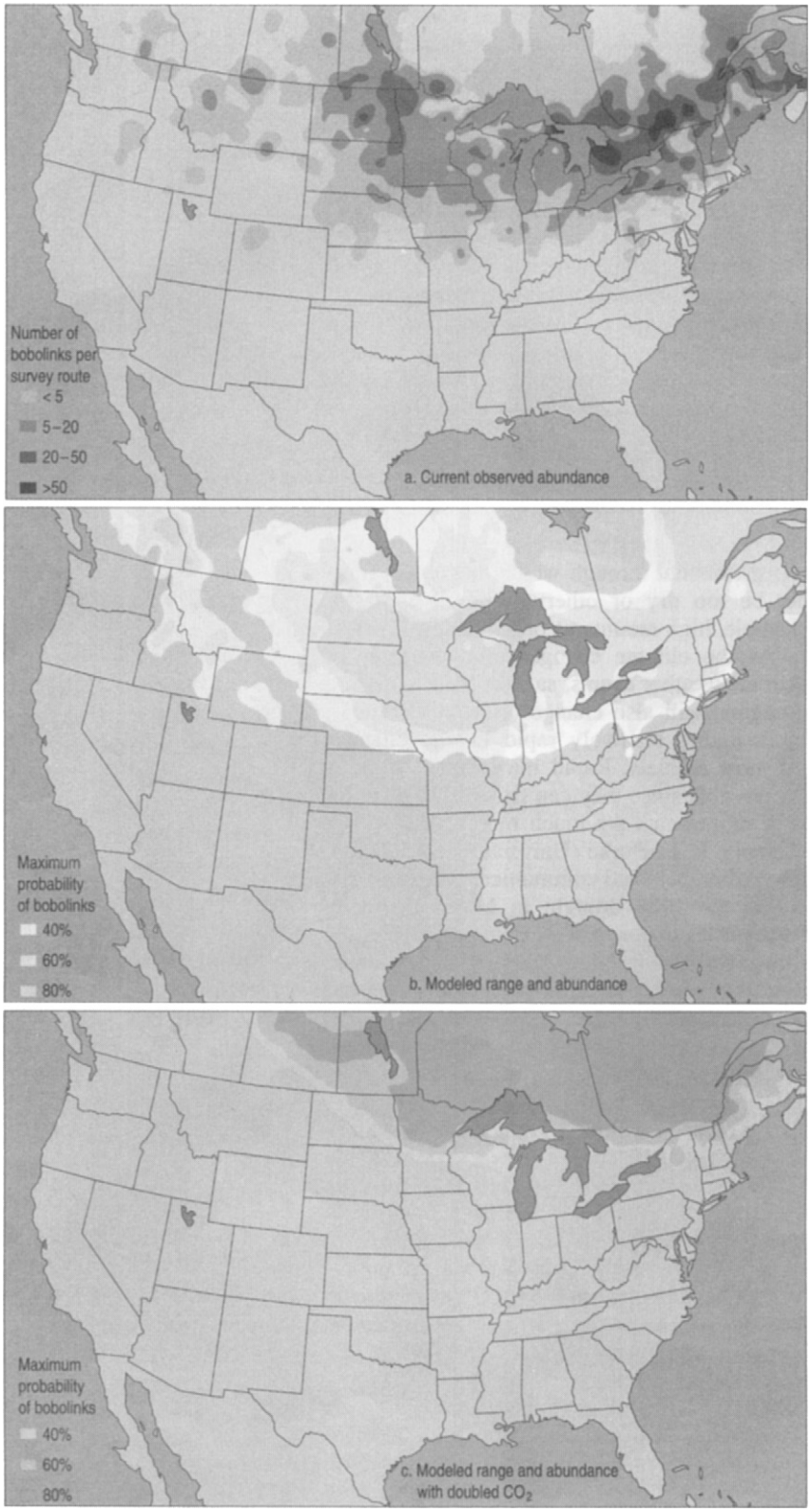


Figure 2 (a) Map of current range and abundance of the bobolink as determined from actual observations during the US Geological Survey Breeding Bird Survey and (b) map of current range and abundance of the bobolink as estimated from the empirical–statistical model. The high correspondence in patterns between the maps in (a) and (b) suggests that this model reliably captures many of the features of the actual observed range and abundance of this species as depicted in the map in (a). (c) Map of the forecasted range and abundance of the bobolink for climate change response of a model with doubled CO_2. This map illustrates the potential for very significant shifts that doubled CO_2 would cause.

Bottom-Up Approaches

Another traditional analysis and forecasting technique is often referred to as bottom-up. Small-scale ecological studies have been undertaken at the scale of a plant or even a single leaf (Idso *et al.*, 1993) to understand how, for example, increased atmospheric CO_2 concentrations might directly enhance photosynthesis, net primary production, or water-use efficiency. Most such studies indicate increases in all these factors – increases that some researchers have extrapolated to ecosystems (Ellsaesser, 1991).

However, at the scale of a forest, the relative humidity within the canopy, which significantly influences the evapotranspiration rate, is regulated by the forest. In other words, if an increase in water-use efficiency decreased the transpiration from each tree, the aggregate forest effect would be to lower relative humidity. This, in turn, would increase transpiration, thereby offsetting some of the direct CO_2/water-use efficiency improvements observed experimentally at the scale of a single leaf or plant. Regardless of the extent to which this forest-scale feedback effect will offset inferences made from bottom-up studies of isolated plants, the following general conclusion emerges: The bottom-up methods may be appropriate for some processes at some scales in environmental science, but they cannot be considered credible without some sort of validation testing at the scale of the system under study.

Combined Top-Down and Bottom-Up Approaches

To help resolve the deficiencies of the top-down biome forest models mentioned previously, more process-based, bottom-up approaches such as forest gap models have been developed (Ostle *et al.*, 2009). These models include individual species and can calculate vegetation dynamics driven by time-changing climatic change scenarios. However, the actual growth rate calculated in the model for each species has usually been determined by multiplying the ideal growth-rate curve by a series of growth-modifying functions that attempt to account for the limiting effects of nutrient availability, temperature stress, and so forth. These growth-modifying functions for temperature are usually determined empirically at a large scale by fitting an upside-down U-shaped curve, with maximum temperature midway between the average temperature of the species' northern range limit and the average temperature of its southern range limit. Growing degree-days (related to temperature but not temperature *per se*) are used in this scenario.

In essence, this technique combines large-scale, top-down empirical-pattern correlations into an otherwise mechanistic bottom-up modeling approach. Although this combined technique refines both approaches, it has been criticized because such large-scale, top-down inclusions are not based on the physiology of individual plant species and lead to confusion about the fundamental and realized niches (Pacala and Hurtt, 1993). (The fundamental niche is the ecological space in which a given species could theoretically survive – for example, if its competitors were absent – and the realized niche is where it actually exists.) The question is then, what limits the realized niche, particularly at the southern boundary? Furthermore, more refined models should include factors such as seed dispersal so that plant recruitment is related to the preexisting population and is not simply the result of a random number generator in the computer code.

Studies of More Refined Approaches

As noted previously, problems with the singular use of either top-down or bottom-up methods have led to well-known criticisms. For bottom-up models, the primary problem is that some of the most conspicuous processes observable at the smaller scales may not be the dominant processes that generate large-scale patterns.

Top-down approaches suffer because of the possibility that the discovered associations at large scales are statistical artifacts that do not, even implicitly, reflect the causal mechanisms needed for reliable forecasting. As Jarvis (1993, p. 121) stated, "A major disadvantage of a top-down model is that predictions cannot be made safely outside the range of the variables encountered in the derivation of the lumped parameter function."

Many studies (e.g., Wright *et al.*, 1993; Root, 1994; Harte *et al.*, 1995) provide examples of a refined approach to analyzing across large and small scales, which Root and Schneider (1995) labeled strategic cyclical scaling. This method builds on the combined techniques in which top-down and bottom-up approaches are applied cyclically in a strategic design that addresses a practical problem: in our context, the ecological consequences of global climatic change. Large-scale associations are used to focus small-scale investigations; this helps ensure that tested causal mechanisms are generating the large-scale relations. Such mechanisms become the laws that allow more credible forecasts of the consequences of global-change disturbances. Levin (1993) observed:

> Although it is well understood that correlations are no substitute for mechanistic understanding of relationships, correlations can play an invaluable role in suggesting candidate mechanisms for (small-scale) investigation (p. 14)

Strategic cyclical scaling, however, is intended not only as a two-step process but also as a continuous cycling process between large- and small-scale studies, with each successive investigation building on previous insights from all scales. This approach is designed to enhance the credibility of the overall assessment process, which is why strategic is the first word in strategic cyclical scaling.

If the rate at which humans are injecting greenhouse gases into the atmosphere is not greatly decreased, Earth's climate will most likely warm by several degrees Celsius by 2050. With this in mind, Root (1988a, 1988b) examined the biogeographic patterns of all wintering North American birds. Root chose this group of species because birds are important parts of ecosystems and because of the availability of the necessary data. The National Audubon Society and the US Geological Survey have volunteer forces amassed to aid in the collection of Christmas Bird Count data and Breeding Bird Survey data, respectively. Using Christmas Bird Count data, Root determined that for a large proportion of species average distribution and abundance patterns are associated with various environmental factors (e.g., the northern range limits

of some species are apparently limited by average minimum January temperature; Root (1988a, 1988b)).

The following is the scaling question: What mechanisms (such as competition or thermal stress) at small scales may have given rise to the large-scale associations? Root first tested the hypothesis that local physiological constraints may be causing most of the particular large-scale, temperature-range boundary associations. She used published small-scale studies on the wintering physiology of key species to determine that about half of the songbird species wintering in North America extend their ranges no farther north than the regions where, to avoid hypothermia during winter nights, they need not increase their metabolic rates more than approximately 2.5 times their basal metabolic rate (Root, 1988b). Root embarked on a larger, regional study to determine whether the longer nights – hence, fewer hours of daylight available for foraging – or the colder temperatures in the more northerly locations are relatively more important. Preliminary results indicate that changing temperatures are more likely than day length to explain this effect (Root, unpublished data): thus, global temperature changes would probably cause a rapid range and abundance shift, at least by selected bird species. Indeed, Root found significant year-to-year shifts in ranges and abundances; these shifts are apparently associated with year-to-year changes in winter temperatures. This example permits a clear demonstration of refined methods for cycling across scales to estimate ecological responses to climatic change.

Three-Way Linkages and Community Ecology

The anticipated changes in plant ranges will probably have dramatic effects on animals, both on the large biogeographic scale and on the local regional scale. The ranges of many animals are strongly linked to vegetation. For example, red-cockaded woodpeckers are endemic to mature longleaf pine and pine–oak forests (Csuti, 1988), and the winter range of Sprague's pipit is coincident with bluestem, a grass (Root, 1988c). Consequently, the ranges of various animals that rely on specific vegetation will change as the ranges of these plants shift, assuming that some other factor is not limiting them. If the climate changes more rapidly than the dispersal rates of the plants, it could result in contraction of plant ranges due to die-offs in the south or downslope before individuals can disperse and become established in the north or upslope. Thus, the ranges of animals relying on these plants could also become compressed, and in some instances both the plants and the animals could become extinct. For instance, the red-cockaded woodpecker needs mature, living trees for nesting sites (Csuti, 1988), and if increasing temperature causes most large trees to die before the newly established dispersing trees mature enough, then this woodpecker, federally listed as endangered, could easily become extinct.

Many animal species have ranges that are not directly limited by vegetation but are instead restricted by temperature. This is true for most ectotherms (insects and related arthropods, amphibians, and reptiles) as well as some endotherms (fish, mammals, and birds). For example, the eastern phoebe, a North American songbird, winters in the US in areas with average minimum temperatures warmer than −4 °C (Root, 1988b). As the Earth warms, those species directly limited by temperature will be able to expand northward as rapidly as their dispersal mechanisms will allow, again assuming other factors are not limiting them. The animals limited by vegetation will be able to expand their ranges only as rapidly as the vegetation changes. Consequently, the potential for significant disruption among communities is high. For instance, some animals may no longer be able to coexist because an invading species disrupts the balance between competing species or between predator and prey. Therefore, to understand the ecological consequences of global climatic change on animals, the three-way linkages among animals, plants, and climate must be understood. It is critical to realize that this is not simply a one-way process whereby climate influences biota but rather a three-way process because animals and plants affect each other and are affected by climate. At the same time, altered surface vegetation can affect climate because mid-continental summer precipitation is significantly influenced by water vapor from evapotranspiration (Salati and Nobre, 1991).

Climate Forecasts, Ecosystem Responses, and Synergistic Effects

Improve Regional Analysis, Study Transients, and Include Many Variables

The most reliable projections from climatic models are for global-scale temperature changes. Ecological impact assessments, however, need time-evolving (transient) scenarios of regional to local-scale climate changes: included are changes in precipitation; severe storm intensity, frequency, and duration; drought frequency, intensity, and duration; soil moisture; frost-free days; intense heat waves; ocean currents; upwelling zones; near-ground ozone; forest canopy humidity; and ultraviolet radiation and total solar radiation reaching the surface, where photosynthesis is important. Data gathered at many scales and by coordinated volunteer and professional sources are needed for archives of these regional and local variables, which in turn can be used to develop and test models or other techniques for climatic forecasting.

Abrupt Climatic Changes

We have argued that sustained globally averaged rates of Earth and ocean surface temperature changes from the past Ice Age to the present were about 1 °C per 1000 years. Alarmingly, this is a factor of approximately 10 slower than the expected changes of several degrees Celsius per 100 years typically projected for the twenty-first century due to human effects. We emphasize the words sustained globally averaged because comparably rapid regional variations have occurred. For example, about 13,000 years ago, after warm-weather fauna had returned to northern Europe and the North Atlantic, there was a dramatic return to Ice Age-like conditions in less than 100 years. This Younger Dryas miniglacial lasted about 1000 years before the stable recent period was established. The Younger Dryas was also accompanied by dramatic disturbances to plants and animals in the North Atlantic and Europe

(Ruddiman and McIntyre, 1981). During the same period, dramatic shifts can be found outside of the North Atlantic region, but no significant climate change is evident in Antarctic ice cores. However, studies of fossils in the North Atlantic show that the warm gulf stream current deviated many degrees of latitude to the south and that the overall structure of deep-ocean circulation may have returned to near Ice Age form in only decades – a weakening of the vertical circulation known as the conveyor-belt current (Broecker *et al.*, 1985).

Plausible speculations about the cause of the Younger Dryas center on the injection of fresh melt water into the North Atlantic presumably associated with the breakdown of the North American ice sheet (Boyle and Weaver, 1994). Could such a rapid change to the conveyor-belt current be induced today by pushing the present climatic system with human disturbances such as greenhouse gases or sulfur oxides? The potential for this is speculative, of course, but its possibility has concerned many scientists (Broecker, 1998; Rahmstorf, 1999). The prospect of climatic surprises in general is chilling enough to lend considerable urgency to the need to speed up the rate of our understanding, slow down the rates at which we are forcing nature to change, or both.

Adaptability

Our current inability to credibly predict time-evolving regional climatic changes has many implications, one of which concerns the adaptability of agricultural ecosystems, that is, any experience farmers might have with anomalous weather in, for example, the 2020s may not help them adapt to the evolving climate change in the 2030s because a transient climate change could differ dramatically over time. This would inhibit learning by doing, creating a potential lack of adaptability associated with the difficulty of reliably predicting regional climatic consequences (Schneider *et al.*, 2000). Such rapid climate changes would be especially difficult for natural ecosystems to adapt to because habitats do not have the luxury of choosing to plant new seeds or change irrigation systems, soil tillage practices, or other agricultural practices.

Ecological Applications-Driven Climatic Research

Regional projections of climatic change arising from a variety of greenhouse gas and sulfur oxide emissions scenarios are essential for ecological applications. Such studies must stress the climatic variables most likely to have significant effects on biological resources. For example, extreme variability measures, such as high temperature and low relative humidity, are important for evaluating the risk of forest fires (Torn and Fried, 1992; Westerling *et al.*, 2006). Identifying such variables of ecological importance and communicating this information to climate scientists require close interdisciplinary, multi-institutional, and cross-scale research efforts to ensure that combinations of variables relevant to ecological applications receive research priority by climatologists. A focus of climate research toward changing climatic variability (Folland *et al.*, 2002) and extremes (Parmesan *et al.*, 2000) might be more useful for ecological impact assessments than the current focus among climatic modelers on climatic means.

Interactive, Multiscale, Ecological Studies Needed

Most ecological studies project the response of one species at small scales or shifts in biomes at large scales to an equilibrium, CO_2-doubled climate model (e.g., the Vegetation/Ecosystem Modeling and Analysis Project, 1995; Stralburg *et al.*, 2009). What is needed for more realistic and useful ecological impact assessments is a multiscale, multispecies, multitaxa analysis driven by regionally specific, transient climatic change forecasts. The construction of ecological forecasts models first requires long-term, large-scale data sets gathered locally by professional (e.g., US Geological Survey landcover data sets) and volunteer (e.g., National Audubon Society Christmas Bird Count) workers. Without such data sets, virtually no credible progress is possible in determining large-scale patterns of associations among ecological and climatic variables. Small-scale studies informed by large-scale patterns are then needed to refine causal mechanisms underlying such large-scale associations, thereby testing the formulas used to make projections of various species or biome responses to hypothesized global changes. For example, Pacala and Hurtt (1993) suggested small- to medium-scale experiments to improve forest gap models. Their criticisms suggest that largely first principles, bottom-up models may still be unrealistic if some top-down parameters (i.e., growth-modifying functions in the instance of gap models) are not appropriately derived from data at the scale at which the models are being applied (Root and Schneider, 1995).

One obvious truism emerges: Credible modeling required for forecasting across many scales and for complex interacting systems is a formidable task requiring repeated testing of many approaches. Nevertheless, tractable improvements in refining combined top-down and bottom-up techniques can be made. It will, however, take more than one cycle of interactions to reliably address the cross-scale and multicomponent problems of ecological assessment – what we (Root and Schneider, 1995) have labeled strategic cyclical scaling.

See also: Carbon Cycle. Climate, Effects of. Diversity, Community/Regional Level. Energy Use, Human. Geologic Time, History of Biodiversity in. Greenhouse Effect

References

Anich NM, Trick JA, Grveles KM, and Goyette JL (2011) Characteristics of a red pine plantation occupied by Kirtland's Warblers in Wisconsin. *Journal of Ornithology* 123: 199–205.

Barnosky AD, Matzke N, Tomiya S, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57.

Beever EA, Chris R, Wilkening JL, Brussard PF, and Mote PW (2011) Contemporary climate change alters the pace and drivers of extinction. *Global Change Biology* 17: 2054–2070.

Botkin DB, Woodby DA, and Nisbet RA (1991) Kirtland's warbler habitats: A possible early indicator of climatic warming. *Biological Conservation* 56: 63–78.

Boyle E and Weaver A (1994) Ocean circulation: Conveying past climates. *Nature* 372: 41–42.

Broecker WS (1998) Paleocean circulation during the last deglaciation: A bipolar seesaw? *Paleoceangraphy* 13: 119–121.
Broecker WS, Peteetand DM, and Rind D (1985) Does the ocean–atmosphere system have more than one stable mode of operation? *Nature* 315: 21–26.
Cooperative Holocene Mapping Project (COHMP) (1988) Climatic changes of the last 18,000 years: Observations and model simulations. *Science* 241: 1043–1052.
Crowley TJ (1993) Geological assessment on the greenhouse effect. *Bulletin of the American Meteorological Society* 74: 2363–2373.
Csuti B (1988) Red-cockaded woodpecker. *Conservation Biology* 2: 136–137.
Darwin C (1859) *The Origin of Species*. New York: Oxford University Press.
Davis MB and Zabinski C (1992) Changes in geographical range resulting from greenhouse warming effects on biodiversity in forests. In: Peters RL and Lovejoy TE (eds.) *Global Warming and Biological Diversity*, pp. 297–308. New Haven, CT: Yale University Press.
Ellsaesser HW (1991) The climate effect on CO_2: A different view. *Atmospheric Environment* 18: 431–434.
Emanuel K (2003) Tropical cyclones. *Annual Review of Earth and Planetary Sciences* 31: 75–104.
Emanuel K (2006) Hurricanes: Tempests in a greenhouse. *Physics Today* 59: 74–75.
Folland CK, Karl TR, and Salinger MJ (2002) Observed climate variability and change. *Weather* 57: 269–278.
Gates WL (1985) The use of general circulation models in the analysis of the ecosystem impacts of climatic change. *Climate Change* 7: 267–284.
Graham RW (1992) Late Pleistocene faunal changes as a guide to understanding effects of greenhouse warming on the mammalian fauna of North America. In: Peters RL and Lovejoy TE (eds.) *Global Warming and Biological Diversity*, pp. 76–87. New Haven, CT: Yale University Press.
Hadly EA, Ramakrishnan U, Chen YL, *et al.* (2004) Genetic response to climate change: Insights from ancient DNA and phylochronology. *PLOS Biology* 2: 1600–1609.
Harte J, Torn MS, Chang FR, *et al.* (1995) Global warming and soil microclimate: Results from a meadow warming experiment. *Ecological Applications* 5: 132–150.
Henderson-Sellers A, Dickinson RE, Durbidge TB, *et al.* (1993) Tropical deforestation: Modeling local-scale to regional-scale climate change. *Journal of Geophysical Research-Atmospheres* 98: 7289–7315.
Holdridge LR (1967) Life Zone Ecology. San Jose, Costa Rica: Tropical Science Center.
Idso SB, Kimball BA, and Hendrix DL (1993) Air temperature modifies the size-enhancing effects of atmospheric CO_2 enrichment on sour orange tree leaves. *Environmental and Experimental Botany* 33: 293–299.
Intergovernmental Panel on Climate Change (1996a) Climate Change 1995 – The Science of Climate Change. In: Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, and Maskell K (eds.) *The Second Assessment Report of the IPCC: Contribution of Working Group I*. Cambridge, UK: Cambridge University Press.
Intergovernmental Panel on Climate Change (IPCC) (2001b) Impacts. In: McCarthy JJ, Canziani OF, Leary NA, Dokken DJ, and White KS (eds.) *Adaptation, and Vulnerability – Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. 1032 pp.
Intergovernmental Panel on Climate Change (IPCC) (2001c) In: Metz B, Davidson O, Swart R, and Pan J (eds.) *Mitigation – Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. 752 pp.
Intergovernmental Panel on Climate Change (IPCC) (2007a) The Physical Scientific Basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyl KB, Tignor M, and Miller HL (eds.) *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
Intergovernmental Panel on Climate Change (IPCC) (2007b) Impacts. In: Perry ML, Canziani OF, Palutikof JP, van der Linden PJ, and Hanson CE (eds.) *Adaptation, and Vulnerability – Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
Intergovernmental Panel on Climate Change (IPCC) (2007c) In: Metz B, Davidson O, Bosch PR, Dave R, and Meyer LA (eds.) *Mitigation of Climate Change – Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
Jarvis (1993) pp. 121.
Kasting JF, Toon OB, and Pollack JB (1988) How climate evolved on the terrestrial planets. *Scientific American* 258: 90–97.
Levin SA (1993) Concepts of scale at the local level. In: Ehleringer JR and Field CB (eds.) *Scaling Physiological Processes: Leaf to Globe*, pp. 7–19. New York: Academic Press.
Morgan MG and Keith DW (1995) Subjective judgments by climate experts. *Environmental Science and Technology*. 29: 468A–477A.
Ostle NJ, Smith P, Fisher R, *et al.* (2009) Integrating plant–soil interactions into global carbon cycle models. *Journal of Ecology* 97: 851–863.
Overpeck JT, Webb RS, and Webb III T (1992) Mapping eastern North American vegetation change over the past 18,000 years: No analogs and the future. *Geology* 20: 1071–1074.
Pacala SW and Hurtt GC (1993) Terrestrial vegetation and climate change: Integrating models and experiments. In: Kareiva P, Kingsolver J, and Huey R (eds.) *Biotic Interactions and Global Change*, pp. 57–74. Sunderland, MA: Sinauer.
Parmesan C, Root TL, and Willig MR (2000) Impacts of extreme weather and climate on terrestrial biota. *Bulletin of the American Meteorological Society* 40: 443–450.
Pereira HM, Leadley PW, Proenca V, *et al.* (2010) Scenarios for global biodiversity in the 21st Century. *Science* 330: 1496–1501.
Pimm S, Raven P, and Peterson A, Sekercioglu CH, and Ehrlich PR. (2006) Human impacts on the rates of recent, present and future bird extinctions. *Proceedings of the National Academy of Sciences* 103: 10941–10946.
Price JT (1995) *Changes in Birds Distributions in the Great Plains with Double CO_2*. Dissertation at Wayne State University.
Rahmstorf S (1999) Shifting seas in the greenhouse? *Nature* 399: 523–524.
Rind D, Peteet D, and Kuklag G (1989) Can Milankovitch orbital variations initiate the growth of ice sheet in a general circulation model? *Journal of Geophysical Research* 94: 12851–12871.
Root TL (1988a) Environmental factors associated with avian distributional limits. *Journal of Biogeography*. 15: 489–505.
Root TL (1988b) Energy constraints on avian distributions and abundances. *Ecology* 69: 330–339.
Root TL (1988c) Atlas of Wintering North American Birds. Chicago, IL: University of Chicago Press.
Root TL (1994) Scientific/philosophical challenges of global change research: A case study of climatic changes on birds. *Proceedings of the American Philosophical Society* 138: 377–384.
Root TL and Schneider SH (1993) Can large-scale climatic models be linked with multiscale ecological studies? *Conservation Biology* 7: 256–270.
Root TL and Schneider SH (1995) Ecology and climate: Research strategies and implications. *Science* 269: 334–341.
Ruddiman WF and McIntyre A (1981) The North Atlantic Ocean during the last deglaciation. *Palaeogeography, Palaeoclimatology and Palaeoecology* 35: 145–214.
Salati E and Nobre CA (1991) Possible climatic impacts of tropical deforestation. *Climatic Change* 19: 177–196.
Santer BD, Taylor KE, Wigley TML, *et al.* (1996) *A search for human influences on the thermal structure of the atmosphere* 382: 39–46.
Schimel DS, House JI, Hibbard KA, *et al.* (2001) Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature* 414: 169–172.
Schneider SH (1984) On the empirical verification of model-predicted CO_2-induced climatic effects. In: Hansen J and Takahashi T (eds.) *Climate Processes and Climate Sensitivity*, pp. 187–201. Washington, DC: American Geophysical Union.
Schneider SH (1987) Climate modeling. *Scientific American* 256: 72–84.
Schneider SH (1993a) Can paleoclimatic and paleoecological analyses validate future global climate and ecological change projections? In: Eddy JA and Oeschger H (eds.) *Global Changes in the Perspective of the Past*, pp. 317–340. Chichester, England: Wiley.
Schneider SH (1993b) Scenarios of global warming. In: Kareiva P, Kingsolver J, and Huey R (eds.) *Biotic Interactions and Global Change*, pp. 9–23. Sunderland, MA: Sinauer Associates.
Schneider SH (1994) Detecting climatic-change signals: Are there any fingerprints? *Science* 263: 341–347.
Schneider SH (1997) Laboratory Earth: The Planetary Gamble We Can't Afford to Lose. New York: Basic Books.
Schneider SH and Londer R (1984) The Coevolution of Climate and Life. San Francisco: Sierra Club Books.
Schneider SH and Root T (1998) *Climate Change*. US Reston, VA: Department of the Interior, US Geological Survey. [Reprinted from: Mac MJ, Opler Puckettt, Haecker PA, and Doran PD. *Status and Trends of the Nation's Biological Resources*.]

Schneider SH and Thompson SL (1981) Atmospheric CO_2 and climate: Importance of the transient response. *Journal of Geophysical Research: Oceans and Atmospheres* 86: 3135–3147.

Schneider SH, Easterling WE, and Mearns L (2000) Adaptation: Sensitivity to natural variability, agent assumptions and dynamic climate changes. *Climatic Change* 45: 203–221.

Shabalova MV and Konnen GP (1995) Climate change scenarios: Comparisons of paleoreconstructions with recent temperature changes. *Climatic Change* 29: 409–428.

Steffen WL, Cramer W, Plochl M, and Bugmann H (1996) Global vegetation models: Incorporating transient changes to structure and composition. *Journal of Vegetation Science* 7: 321–328.

Stralburg DD, Jongsomjit D, Howell CA, *et al.* (2009) Re-shuffling of species with climate disruption: A no-analog future for California birds. *Public Library of Science One* 4(9): e6825.

Titus JG and Narayanan V (1996) The risk of sea level rise. *Climatic Change* 33: 151–212.

Torn MS and Fried SJ (1992) Predicting the impacts of global warming on wildland fire. *Climate Change* 21: 257–274.

Trenberth KE (ed.) (1992) *Climate System Modeling*. Cambridge, UK: Cambridge University Press.

Trenberth K (2005) Uncertainty in hurricanes and global warming. *Science* 308: 1753–1754.

Vegetation/Ecosystem Modeling and Analysis Project (1995) Vegetation/Ecosystem Modeling and Analysis Project (VEMAP): Comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO_2 doubling. *Global Biogeochemical Cycles* 9: 407–437.

Washington WM and Meehl GA (1986) General circulation model CO_2 sensitivity experiments: Snow-sea ice albedo parameterizations and globally averaged surface air temperature. *Climatic Change* 8: 231–241.

Wright HE, Kutzbach JE, Webb, III T, Ruddiman WE, Street-Perrott FA, and Bartlein PJ (eds.) (1993) *Global Climates Since the Last Glacial Maximum*. Minneapolis: University of Minnesota Press.

Westerling AL, Hidalgo HG, Cayan DR, and Swetnam TW (2006) Warming and Earlier Spring Increases Western US Forest Wildfire Activity. *Science* 313: 940–943.

Wigley TML and Raper SCB (1991) Natural variability of the climate system and detection of the greenhouse effect. *Nature* 344: 324–327.

Climate Change and Extinctions

Joanne L Isaac and Stephen E Williams, James Cook University, Townsville, QLD, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Advective precipitation When moist air passes over cool ground by advection (wind) and is cooled.
Biodiversity hotspot An area with exceptionally high biodiversity but threatened by habitat loss.
Endemic A species restricted to one place only.
Niche The function of a particular species in an ecological community.
Orographic cloud layer Clouds that are formed as moist air rises over mountains.
Phenology The timing of seasonal activities such as timing of flowering, migration, or breeding.
Stochastic event A random event such as a hurricane or wildfire.

Climate change is predicted to result in a worldwide loss of biodiversity, with extinctions becoming increasingly common in the future. The climate is changing at an ever more rapid rate, affecting individuals and populations and disrupting community interactions and ecosystem functions. Climate change will likely reduce worldwide biodiversity through negative effects on already extinction-prone species such as endemics and those at the top of the food chain. This will likely result in a more homogeneous fauna and flora worldwide that is dominated by generalist taxa that are able to adapt to changing environmental conditions.

Climate Change

Worldwide, the climate is changing at an unprecedented rapid rate. In general, many regions have become warmer and wetter than in the past. Average temperatures have already risen approximately 0.74 °C globally, and the Intergovernmental Panel on Climate Change (IPCC, 2007) reports that recent trends are greater than those of the IPCC 2001 report (0.6 °C). Temperature increases are currently tracking the "worst case scenario" predicted by the (IPCC, 2007), which predicts further increases of between 1.8 °C and 4.0 °C by the end of this century. There has been an increase in the number of heat waves around the world, along with warming of the oceans, coral-bleaching events, retreating glaciers and sea ice, fewer frosts, and heavier rainfall events. The sea level has risen an average of 3.1 mm yr^{-1} since 1993, and the IPCC predicts a further rise of 0.18–0.59 cm by 2100. Severe weather and catastrophic events are also becoming more common, including El Niño, droughts, floods, and wildfires. Climatic changes are attributed due to increased greenhouse gas (GHG) emissions such as carbon dioxide. Human GHG emissions continue to increase, with an increase of 70% between 1970 and 2005 (IPCC, 2007), due to human activities including the use of coal, oil, and gas for energy and also the burning and conversion of forests for other land uses (IPCC, 2007).

Over history, Earth has experienced much climatic variability, but current climatic changes differ in two important ways in terms of their negative impact on biodiversity and extinction rates. First, the rate of temperature change in the late 20th century, and that predicted in the near future, is generally considered to be unparalleled in the past 10,000 years. Second, many of Earth's ecosystems are already stressed by other detrimental human impacts, such as land clearing and habitat fragmentation, resulting in small and isolated animal populations that are highly susceptible to the types of stochastic events that climate change will bring. Climate change is predicted to have a profound effect on individuals and populations in animal communities, and its effect on biodiversity and ecosystem function is of worldwide interest and concern.

Extinction Risk

Intrinsic Risks

The risk of extinction is rarely random across species as particular traits make some species more extinction-prone than others (McKinney, 1997). A number of life history traits appear to consistently predispose species to a high risk of extinction across animal groups, especially those associated with a slow life history, including large body size, a long life span, slow maturity, few reproductive events, and the production of a few slow-growing offspring at each reproductive event. These traits, such as low fecundity, mean populations will be slow to recover if reduced to small sizes and therefore more likely to go extinct if a stochastic event increases the rate of mortality (Bennett and Owens, 1997). Larger species are also commonly at the top of the food chain, such as large carnivores, and are predicted to be at high risk of extinction as they are more vulnerable to the cumulative effects of disturbance to species lower down the food chain – so-called chains of extinction (**Figure 1**).

Other intrinsic factors that have been related to extinction risk include population isolation, such as the populations of endemic species restricted to islands or mountaintops, and ecological specialization, including specialized diet and habitat requirements. Population models also generally predict that there will be an increased extinction risk in populations that show high variability in growth and dynamics, because if a population commonly experiences large fluctuations in size and a stochastic event occurs at a time when that population

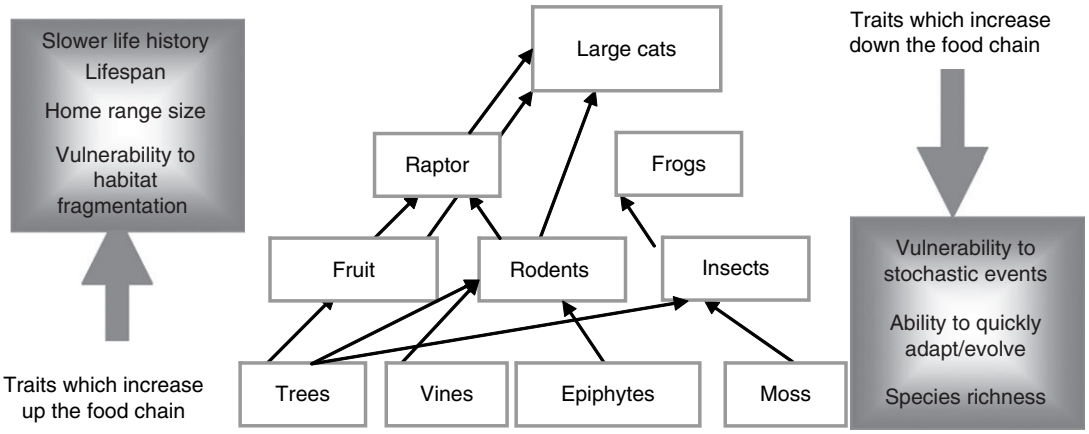


Figure 1 A simplified forest food web demonstrating how loss of biodiversity differs, depending on where a species lies in the food chain. Many of the traits that make species vulnerable to extinction tend to increase up the food chain, culminating at the top predator.

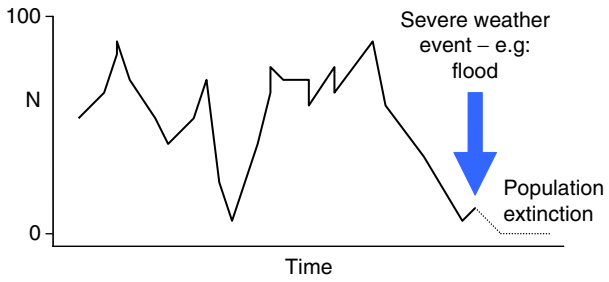


Figure 2 A graphical illustration of how population variability can influence extinction risk. This population undergoes great variability in numbers (*N*) over time; if a stochastic weather event such as a severe flood occurs at a population low, then the population will be unable to recover and will likely go extinct.

is experiencing a crash, extinction is a likely outcome (Figure 2). Poor dispersal ability has also been related to increased extinction risk in a variety of taxa such as insects and frogs. The dispersal ability of such species may have evolved under a less-fragmented landscape structure, and thus individuals emigrating from habitat fragments in search of new suitable habitat patches may be unable to cross nonhabitable areas.

Extrinsic Risks

There are also many extrinsic factors that can contribute to the extinction or decline of a species. Susceptibility to disease is an important determinant in extinction risk, and disease has been implicated as a primary factor in the decline and extinction of amphibians worldwide. Introduced species have also been implicated in the extinction and decline of a variety of taxa in recent years. In Australia, the cane toad (*Bufo marinus*) is toxic to predators such as native cats and snakes and may compete with native amphibians for resources. Anthropogenic activities that currently threaten numerous species include habitat modification, habitat destruction and pollution, overfishing and hunting. These extrinsic risks commonly interact with intrinsic factors to increase extinction risk in the most vulnerable species (Figure 3).

Climate Change, Extinctions, and Biodiversity Loss

Geographic and Landscape Scale Variation in Expected Biodiversity Loss

The degree of climatic change, and also the impact of climate change on biodiversity, is predicted to differ both at a large geographic scale and also at smaller scales such as across altitudinal and vegetation zones. However, data from 75 studies to date that show significant change in biological and physical systems – nearly 90% are consistent with changes expected as a response to warming. Land areas have warmed faster than oceans, and the largest increases in temperature are found at higher northern latitudes, suggesting that terrestrial species at higher latitudes will be most at risk of extinction (IPCC, 2007). However, due to rising sea levels and flooding events, species inhabiting coastal wetland habitats are expected to be at extremely high risk; the IPCC predicts that approximately 33% of coastal wetlands could become open ocean by 2080. Rainfall is projected to increase at high latitudes and also across the equatorial zone but decrease in the subtropics. Although areas at high latitude are predicted to experience greater climatic change in terms of increases in mean annual temperature, the high biodiversity and high proportion of endemic species found in tropical areas mean that even small changes in the tropics could be catastrophic for global biodiversity.

Conservation International recognizes 34 international hotspots of biodiversity – areas across the globe that support generally high species richness and diversity and also high levels of endemism. These areas, although covering only 2.3% of Earths land surface, are estimated to account for 77% of world's vertebrate biodiversity. Any negative effects of climate change on these areas would thus be extremely detrimental to global biodiversity as a whole. Malcolm *et al.* (2006) modeled extinctions of endemic species from biodiversity hotspots under predicted climatic changes and found an average of 11.6% extinctions in 100 years – vulnerable hotspots were distributed across a variety of biomes, including Mediterranean, tropical, and dry savanna areas. In tropical regions, extinctions induced by climate change exceeded those predicted

through deforestation. Tropical areas account for a disproportionate number of hotspot areas (54%), and tropical montane cloud forest (high altitude forest with persistent cloud and mist) is particularly important in terms of biodiversity and endemic species. However, tropical mountain systems, and especially tropical montane cloud forests, are likely to be extremely vulnerable to climatic changes due to their unique and fragmentary nature and association with elevational and temperature gradients. The possible effects of climate change on tropical mountain areas include a general increase in temperature, increased water loss through evaporation, a rise in the average basal altitude of the orographic cloud layer, changing rainfall input, decreased soil moisture, increases in fire and drought, and increased invasion potential for introduced pest species. Furthermore, forest types at high elevation have been shown by to be more vulnerable to climate change than other tropical forest types (Enquist, 2002), possibly due to the lower migration rates found in tropical plant species compared to those in boreal or temperate areas, as demonstrated by Jay Malcolm and co-workers.

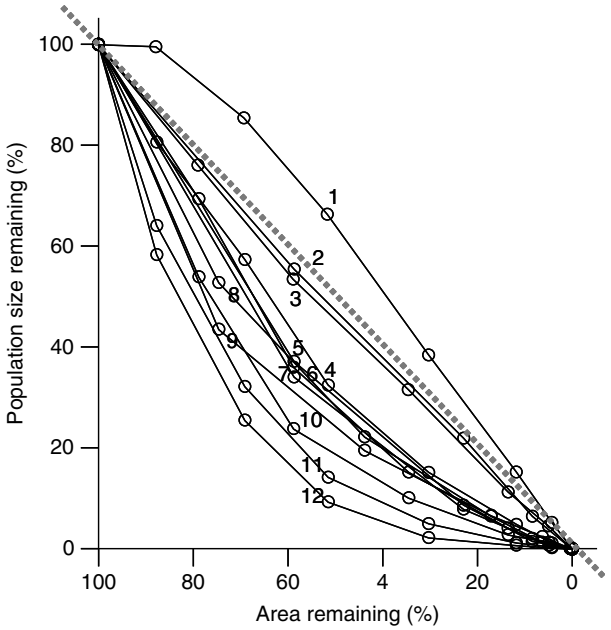


Figure 3 Projected change in the relative population size and distribution area of rainforest birds following altitudinal range shifts associated with climate warming. The relationship is linear where population size declines in direct proportion to distribution area (i.e., dotted gray line) and curvilinear where population size declines more rapidly (below the dotted gray line) or more slowly (above the dotted gray line) than distribution area with increasing temperature. Number labels refer to individual species of rainforest birds: (1) Macleay's honeyeater (*Xanthotis macleayana*); (2) tooth-billed bowerbird (*Scenopoeetes dentirostris*); (3) pied monarch (*Arses kaupi*); (4) Victoria's riflebird (*Ptiloris victoriae*); (5) golden bowerbird (*Prionodura newtoniana*); (6) chowchilla (*Orthonyx spaldingii*); (7) mountain thornbill (*Acanthiza katherina*); (8) fern wren (*Oreoscopus gutturalis*); (9) Atherton scrub wren (*Sericornis keri*); (10) bridled honeyeater (*Lichenostomus frenatus*); (11) gray-headed robin (*Heteromyias albispecularis*); (12) Bower's shrike-thrush (*Colluricincla boweri*).

Effects of Climate Change on Individuals And Populations

Research over the past decade has demonstrated that many of the intrinsic and extrinsic factors that are related to extinction risk within a species can be affected by climate change. The observed impacts of climate change on individuals and populations are increasing rapidly, although the magnitude and direction of effects appear to vary considerably between species. The most commonly observed impact is shifts in distribution and range of animals and plants; the most common pattern has been shifting toward the poles for terrestrial and marine species and also higher in altitude for mountain species. These shifts have resulted in range expansions for some species but contractions for others. A key review by Chris Thomas and colleagues was among the first to show range retractions are occurring as a result of climate change, highlighting research conducted on 16 butterfly species in Spain by Robert Wilson and co-workers, who document that range boundaries have risen by an average of 212 m in the past 30 years, accompanying a 1.3 °C rise in mean annual temperature. More recently, Chen *et al.* (2011) estimate, using a meta-analysis approach, that species are moving up a median of 11 m in altitude per decade and to higher latitudes at 11 km per decade; these rates are around two and three times greater than earlier predictions. Furthermore, Shoo *et al.* (2005) incorporated spatial patterns of abundance into a predictive model and demonstrate that for many endemic wet tropics birds in Australia, as temperature increases population size is likely to decline more rapidly than distribution area (**Figure 4**). This means that extinction risk associated with climate change will be more severe than expected from decline in distribution area alone.

Recently, mass deaths of individuals as a result of extreme weather events, particularly heat waves, have also been noted in a variety of species and locations. In Australia, more than 3500 flying foxes were seen to perish when temperatures rose to over 42 °C (Welbergen *et al.*, 2008). Heat-related deaths

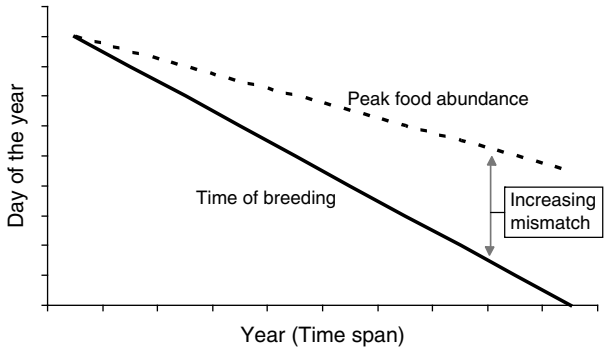


Figure 4 A graphical representation demonstrating how different responses to climate change can result in the decoupling of predator–prey relationships. In this case, as the climate has changed over time, breeding of a predator (for instance, a migratory bird) is occurring earlier in the year. The peak food abundance of the predators main prey (for example, caterpillars) has also shifted to earlier in the year – but not to the same degree. Over time, this has resulted in an increasing mismatch between predator food requirements for breeding and prey biomass availability, likely resulting in breeding failures for the predator.

have also been noted in birds and are expected to rise as heat waves become more common (McKechnie and Wolf, 2010). For restricted species, a mass die-off event could result in local or total extinction. Ocean warming has also resulted in mass coral-bleaching events; a study by Carpeneter *et al.* (2008) predicts that one-third of reef building corals are threatened with extinction primarily as a result of climate change.

Shifts in phenology and variations in reproductive behavior and success have also been documented for a variety of taxa. For example, many migratory birds are arriving at their breeding grounds earlier and reproducing earlier (Van Buskirk *et al.*, 2009); the impacts on reproductive success appear mixed, with both positive and negative outcomes (Møller *et al.*, 2010). Research suggests shifts in the timing of reproduction may often be in response to changes in food availability as climate change affects species further down the food chain. For example, Jaume Forcada and co-workers show that pup production in Antarctic fur seals, South Georgia, has declined considerably since 1985 with increased sea surface temperatures, likely associated with low availability of food (krill). A negative effect of climatic warming on reproductive parameters is also the likely cause of extinction in endemic Costa Rican amphibians such as the golden toad (*Bufo periglenes*). Crump *et al.* (1992) hypothesized that climatic changes, specifically warmer water temperatures and less-advective precipitation, created adverse breeding conditions for the toads. These species also displayed many of the traits purported to predispose species to high extinction risk, in particular restricted distributions: the golden toad was found only in an area ~0.5 km by 8 km and elevations between 1500–1620 m. Global warming is accompanied by an increase in ambient ultraviolet b (UV-B) radiation, and Michael J. Adams and colleagues have shown that a north temperate frog species, *Rana cascadae*, chooses breeding sites with low UV-B attenuation; it is likely that high UV-B exposure results in lower survival and high deformity in offspring.

Finally, there is also evidence that climatic changes have resulted in significant changes in body mass and size in a number of vertebrates. According to Bergmann's rule, animals in warmer areas should be smaller than those in cooler areas, as the relatively smaller body surface of larger animals may assist in heat conservation in cooler climates. Thus, we might predict a decrease in body size or mass or both with global warming. A number of authors have shown that this is indeed the case for a number of species, including plants, birds, and mammals (review in Sheridan and Bickford, 2011). Decline in size and mass within a species have strong implications for other life history traits and extinction risk and reduced size may negatively impact on reproductive success and survival. However, contrary to the predictions of Bergmann's rule, the converse pattern has also been found in lizards and mammals, particularly in very cold areas. In isolated populations of the common lizard (*Lacerta vivipara*) in the French Alps, an increase in temperatures resulted in an increase in body size, with an associated increase in reproductive success (clutch size and reproductive output) (Chamaille-Jammes *et al.*, 2006). Similarly, Yorav and Jonathan Yom-Tov demonstrate an increase body mass in masked shrews in Alaska. In these environments with extreme winters, global warming may have resulted in an increase in food availability in these insectivorous species or allowed them a longer growth period – resulting in size or mass increases that may benefit these species, at least in the short term.

Effects of Climate Change on Communities and Ecosystems

As the previous section highlights, the response of different species to climate change varies considerably, which may disrupt the coupling between species, with cascading effects up the food chain resulting in extinctions and biodiversity loss. For example, the fecundity and reproductive success of one species is commonly dependent on the abundance and survival of their food source, such that timing of births usually coincides with peak food availability. An important concept in the effort to predict biotic responses to climate change is the *temporal match–mismatch* (TMM) hypothesis, which suggests that climate change can create a mismatch between timing of peak breeding of predator and prey (**Figure 4**). Studies indicate that trophic mismatch is already occurring. For example, mistiming of avian reproduction and caterpillar biomass appears to be relatively widespread and has resulted in failed breeding and population crashes in many birds (reviewed in Visser and Both, 2005), and mismatch between plants and pollinators have also been noted (Hegland *et al.*, 2009). The decoupling of phenological relationships between species within a community will have important implications for trophic interactions, leading to changes in food-web structure and ecosystem-level changes. Other recent empirical studies highlight the potential for climate change to affect keystone predators, with impacts throughout the food chain and effects touching community structure and diversity (Harley, 2011).

A number of authors have also demonstrated that extreme weather events, predicted to become more common as climate change progresses, and general global warming can result in changes to species assemblages and community structure through differential effects on species. Following Hurricane Georges, which hit the Caribbean in 1998, there was a decline in abundance and species richness in bat communities, with large species more affected than smaller species, and frugivorous and nectarivourous species more affected than insectivorous species (Jones *et al.*, 2001). Similarly, following Hurricane Isis in 1998, there was a change in the composition of cloud forest birds in Mexico, whereby generalists were favored following the storm (Tejeda-Cruz and Sutherland, 2005). In Australia, Luke Shoo and colleagues predict that 74% of rainforest birds in the wet tropics will become threatened as a result of projected warming in the next 100 years. However, extinction risk varies according to where on the altitudinal gradient a species is currently most abundant. Upland birds are predicted to be most affected and are likely to be immediately threatened by small increases in temperature, whereas there may be a capacity for the population size of lowland species to increase, at least in the short term. Plant communities are likely to be similarly affected – Nadlini Nadkarni and Rodrigo Solanao demonstrate that climate change will result in negative effects on the productivity and longevity of rainforest epiphytes in tropical cloud forest in Costa Rica, resulting in subsequent emergence of terrestrial components into the canopy community.

Interactions of Climate Change and Existing Risk of Extinction

For many species, it is likely that the factors that underlie their existing risk of extinction will be exacerbated by climate change. For example, climate change may increase the occurrence, range, or prevalence of diseases, parasites, and other introduced pest species (reviewed by Brook *et al.*, 2008). The chytrid fungus, *Batrachocytrium dendrobatidis*, has been implicated in widespread amphibian extinctions. Evidence suggests that large-scale warming exacerbates outbreaks of chytrid fungus by producing conditions optimal for the growth of the fungus (Pounds *et al.*, 2006). Further evidence that warming can expand the range of pest species comes from North America, where the aggressive mountain pine beetle, *Dendroctonus ponderosae*, has expanded its natural range and now occupies areas naïve to this species. Cudmore *et al.* (2010) demonstrate that reproductive success for the beetles is higher in areas that have not seen previous outbreaks of the pest; this increased beetle productivity is likely responsible for massive tree death in affected areas.

Endemic species commonly display many traits that predispose them to high extinction risk, including small geographical ranges, specific ecological requirements, and little capacity for long-distance migration. Thus, the already intrinsically high extinction risk for endemic species is also expected to increase as the climate changes. For example, Thomas *et al.* (2004) demonstrate a strong negative relationship between geographical range size and risk of extinction under predicted climate-warming scenarios. They found that 15–37% of species in their data set will be "committed to extinction" by 2050, the majority being species with already small geographic range sizes. More recently, Isaac *et al.* (2008) demonstrate that the endemic species of Australia's wet tropics have both lower resistance (the ability of species to withstand an environmental perturbation) and lower resilience (ability to recover from an environmental change); this means these species are at greater risk of extinction under climate change, and other environmental change, than more widespread, generalist species.

Conclusions

There is no doubt that a changing climate has already had significant impacts on global biodiversity patterns and extinctions and will continue to do so. The research reviewed here indicates that extinctions due to climate change will not be random and that species with more generalist ecological requirements will be better able to cope with and adapt to climatic changes. Those species that display traits that already predispose them to high extinction risk, such as specialists and range-restricted species, will be negatively affected. We will likely see the decline and extinction of many of these species and many more becoming threatened. As generalists are favored and introduced species expand their ranges, we are also likely to see a more homogeneous, less diverse fauna and flora as generalist species become more dominant.

The capacity to adapt to a changing climate will depend on all of the intrinsic and extrinsic factors discussed and the actual realized degree of climate change. We would suggest that there is not enough time to evolve novel traits and extinction probability will largely depend on existing plasticity in relevant ecological and evolutionary characteristics – with some possibilities for active adaptation management, such as habitat restoration, wildlife corridors, and assisted migration – to help increase resilience. Natural resource management should focus on maintaining ecosystem resilience while putting pressure on governments to reduce emissions. Research effort should focus on the factors discussed in this chapter to allow more robust predictions and better evaluation of the relative extinction proneness of species and ecosystems. This knowledge will enable informed decisions and ensure the efficient and effective allocation of scarce management resources.

Appendix

List of Courses

1. Life History Theory
2. Climate Change and Biodiversity
3. Conservation Biology

See also: Biodiversity-Rich Countries. Climate Change and Ecology, Synergism of. Conservation Efforts, Contemporary. Endemism. Global Declines of Amphibians. Impact of Past Global Warming on Biodiversity. Introduced Species, Impacts and Distribution of. Loss of Biodiversity, Overview. Modern Examples of Extinctions

References

Bennett PM and Owens IPF (1997) Variation in extinction risk among birds: Chance or evolutionary predisposition? *Proceedings of the Royal Society of London, Series B: Biological Sciences* 264: 401–408.

Brook BW, Sodhi NS, and Bradshaw CJA (2008) Synergies among extinction drivers under global change. *Trends in Ecology and Evolution* 8: 453–460.

Carpeneter KE, Abrar M, Aeby G, *et al.* (2008) One-third of reef building corals face elevated extinction risk from climate change and local impacts. *Science* 321: 560–563.

Chamaille-Jammes S, Massot M, Aragon P, and Clobert J (2006) Global warming and positive fitness response in mountain populations of common lizards Lacerta vivipara. *Global Change Biology* 12: 392–402.

Chen I-C, Hill JK, Ohlemuller R, Roy DB, and Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. *Science* 333: 1024–1026.

Crump ML, Hensley FR, and Clark KL (1992) Apparent decline of the golden toad: Underground or extinct? *Copeia* 1992: 413–420.

Cudmore TJ, Bjorklund N, Carroll AL, *et al.* (2010) Climate change and range expansion of an aggressive bark beetle: Evidence of higher beetle reproduction in naive host tree populations. *Journal of Applied Ecology* 47: 1036–1043.

Enquist CAF (2002) Predicted regional impacts of climate change on the geographical distribution and diversity of tropical forests in Costa Rica. *Journal of Biogeography* 29: 519–534.

Harley CDG (2011) Climate change, keystone predation, and biodiversity loss. *Science* 334: 1124–1127.

Hegland SJ, Nielsen A, Lazaro A, *et al.* (2009) How does climate warming affect plant–pollinator interactions? *Ecology Letters* 12: 184–195.

IPCC (2007) In: Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, and Hanson CE (eds.) *Contribution of Working Group II to the Fourth Assessment*

Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press.

Isaac JL, VanDerWal JJ, Johnson CN, *et al.* (2008) Resistance and Resilience: Quantifying relative extinction risk in a diverse assemblage of Australian rainforest vertebrates. *Diversity and Distributions* 15: 280–288.

Jones KE, Barlow KE, Vaughan N, Rodriguez-Duran A, and Gannon MR (2001) Short-term impacts of extreme environmental disturbance on the bats of Puerto Rico. *Animal Conservation* 4: 59–66.

Malcolm JR, Canran L, Neilson RP, Hansen L, and Hannah L (2006) Global warming and extinctions of endemic species from biodiversity hotspots. *Conservation Biology* 20: 538–548.

McKechnie AE and Wolf BO (2010) Climate change increases the likelihood of catastrophic avian mortality events during extreme heat waves. *Biology Letters* 6: 253–256.

McKinney ML (1997) Extinction vulnerability and selectivity: Combining ecological and paleontological views. *Annual Review of Ecology and Systematics* 28: 495–516.

Møller AP, Flensted-Jensen E, Klarborg K, Mardal W, and Nielsen JT (2010) Climate change affects the duration of the reproductive season in birds. *Journal of Animal Ecology* 79: 777–784.

Pounds JA, Bustamante MR, Coloma LA, *et al.* (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439: 161–167.

Raffaelli D (2004) How extinction patterns affect ecosystems. *Science* 306: 1141–1142.

Sheridan JA and Bickford D (2011) Shrinking body size as an ecological response to climate change. *Nature Climate Change* 1: 401–406.

Shoo LP, Williams SE, and Hero JM (2005) Potential decoupling of trends in distribution area and population size of species with climate change. *Global Change Biology* 11: 1469–1476.

Tejeda-Cruz C and Sutherland WJ (2005) Cloud forest bird responses to unusually severe storm damage. *Biotropica* 37: 88–95.

Thomas CD, Cameron A, Green RE, *et al.* (2004) Extinction risk from climate change. *Nature* 427: 145–148.

Van Buskirk J, Mulvihill RS, and Leberman RC (2009) Variable shifts in spring and autumn migration phenology in North American songbirds associated with climate change. *Global Change Biology* 15: 760–771.

Visser ME and Both C (2005) Shifts in phenology due to global climate change: The need for a yardstick. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 272: 2561–2569.

Welbergen JA, Klose SM, Markus N, and Eby P (2008) Climate change and the effects of temperature extremes on Australian flying foxes. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 275: 419–425.

Climate Change and Wild Species

Terry L Root and Stephen H Schneider†, Stanford University, Stanford, CA, USA
Rachel Warren, University of East Anglia, Norwich, UK
Jeff R Price, World Wildlife Fund, Washington, USA
Patricia R Mastrandrea, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Stephen H. Schneider, Terry L Root, Patricia R Mastrandrea, pp. 1–26 © 2007, Elsevier Inc.

Glossary

Climatic models Mathematical descriptions of the flows of energy, momentum, and materials around the atmosphere and oceans, typically run on large computers to stimulate the Earth's climate naturally and when disturbed by greenhouse gases (GHGs) or the other so-called forcings or disturbances.

Global change disturbances Alterations or disturbances to the natural conditions of the global atmosphere (e.g., GHGs produced by human activities) or to enough regions (e.g., large-scale deforestation or the widespread introduction of exotic, invasive species in many places throughout the world) to have global impacts.

Global warming The additional heating of the Earth's climate from the incremental injection of GHGs to the atmosphere from human activities such as deforestation or fossil fuel consumption.

Greenhouse gases Gases in the atmosphere, which selectively let more of the solar energy into and out of the atmosphere than they permit the transmission of long-wave infrared radiation. They contribute to the greenhouse effect, which traps radiant energy in the lower atmosphere and makes the Earth warmer.

Synergisms Interactions among many factors, which collectively may have a much larger or smaller effect than the sum of the effects of each factor acting independently.

Forces Impacting Wild Species

Biological diversity, which includes various species of plants, animals, fungi, and microorganisms, genetic differences within species, and the variety of ecosystems and habitats, is the result of billions of years of evolution shaped by natural processes and, increasingly, by human influences. Humans are an integral part of this web of complex communities interacting with one another and their physical environment, and we are entirely dependent on them for our well-being and survival. This complex, interdependent web of ecosystems either directly or indirectly provides the basic materials for life (e.g., food and water), delivering essential life-support goods and services. The Millennium Ecosystem Assessment (2005) identifies 24 ecosystem services that maintain the essential life-support processes of the planet, including primary production and nutrient recycling, pollination, air and water purification, waste detoxification and decomposition, soil fertility regeneration, protection from natural disasters and disease (e.g., regulating climate, floods, and pests), and production and maintenance of biodiversity. Of the 24 ecosystems, 15 have become endangered over the last 50 years and five more are threatened with disruption. This array of services is generated by a complex interplay of natural cycles powered by solar energy and operating across a wide range of space and timescales. Daily (1997) identifies such ecosystem goods as seafood, fodder, fuel wood, timber, and pharmaceutical products that are of great importance to human populations as well as an integral part of essential ecosystems that support other species. Ecosystems also provide the basis for such important aspects of human culture as spiritual needs, knowledge systems, and traditional use. Biodiversity is essential for the sustained provision of these support goods and services that are provided at all levels – local, regional, and global – each one making crucial contributions to human well-being (UNEP, 2005).

Recognizing the fundamental role of biodiversity in supporting human life, the Convention on Biological Diversity (CBD), a legally binding landmark treaty, was opened for signature at the Rio Earth Summit and entered into force in 1993. The CBD's three main objectives are the conservation of biodiversity, the sustainable use of its components, and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources. In 2002, 10 years after the Convention's entry into force, Parties acknowledged that there is a continued threat to biodiversity from human activities, reaffirming that biodiversity is the living foundation for sustainable development, that the rate of loss is still accelerating, that threats must be addressed, and that the Convention continues to be a key tool for sustainable development. At this time, Parties adopted a Strategic Plan whereby they committed to a more effective and coherent implementation of the Convention objectives so as to achieve a significant reduction of the current rate of biodiversity loss at all levels by 2010. There are five main threats to biodiversity, which are commonly recognized in the programs of work of the Convention: habitat change, invasive alien species, climate change, nutrient loading and pollution, and overexploitation. These work synergistically together to amplify negative impacts on wild species.

†Deceased.

Threats to Biodiversity: Habitat Change, Invasive Species, Overexploitation, Pollution, and Nutrient Loading

Human activities are responsible for much of the accelerating loss of biodiversity. Humans are adversely impacting biodiversity by changing habitat, generating pollution, and loading nutrients into the natural environment, introducing invasive species, overharvesting species, and affecting climate.

Habitat Change

The natural landscape has been divided or fragmented for urban and suburban development and agricultural production disrupting interactions among organisms and destroying, reducing, or seriously altering essential habitats. Habitat fragmentation not only affects species, but also affects the processes that drive biodiversity. Fragmentation of habitat causes large populations to be broken into smaller populations that may be isolated from one another. Eventually, loss of habitat can affect the genetic variability and viability of populations that can lead to local extinctions. If local extinction of species occurs, fragmentation cuts off the potential for repopulation because no intact populations are nearby. If a species experiences local extinction in many locations, this can eventually lead to global or permanent extinction. Maintaining habitat is important for annual animal migrations. For example, the Mississippi River, which crosses through the Midwest, serves as a main flyway for neotropical migrating birds. If the habitat that these birds require at stopover sites, where they replenish their energy along their migration route, are significantly degraded, the species are in jeopardy of not having enough energy to complete their trip.

Poor agricultural practices degrade soil quality and promote the loss of topsoil. Furthermore, agriculture has resulted in local reductions and extirpations of flora and fauna associated with agricultural lands (e.g., grassland and scrubland plants, birds, and wild pollinating insects).

Overharvesting/Overexploitation

Overharvesting has a great effect on biodiversity and often occurs with habitat loss, as removal of an organism from its environment can have irreversible impacts on the environment itself. Humans have historically exploited plant and animal species to maximize short-term profit, at the expense of sustainability of the species or population. The unsustainability of biological-resource extraction is rooted in three main problems:

1. too many people who want the resource,
2. short-term return on investment goals of extractors, and
3. lack of information about the ecology and life histories of populations being harvested.

Overexploitation follows a predictable pattern: initially, a species harvested from the wild can provide a substantial profit, which encourages more extraction and leads to the development of more large-scale and cost-effective methods of harvesting the species, which eventually depletes the resource. Quota systems are often put in place, leading to more competition, decreased earnings and the need for government subsidies to support the extraction industry. The fishing industry, the logging industry, and grazing of cattle on public lands have all been examples of this process of overexploitation. When the rate of increase in demand for the resource far outstrips the reproductive rate of the population, demand outstrips supply, and drives the value of that resource higher, increasing the incentive to extract the resource, and eventually leading to the collapse of the population. The results for the resource are population crashes, which can ultimately lead to global extinction.

Pollution and Nutrient Loading

Human activities produce pollution, which adversely affects ecosystems. Human-generated pollution impacts water resources in the form of toxic discharges of metals, organic chemicals, and suspended sediments usually found in industrial and municipal effluents, which can inversely impact the biota (living organisms) in an ecosystem; bacterial contamination from human waste that can cause disease; and nutrient buildup of, for example, phosphorus and nitrogen that often originate as runoff from fertilizers applied on agricultural fields. Nutrient loading can result in excesses that stimulate rapid growth of algae and aquatic plants, ultimately limiting the amount of oxygen and light available to other organisms in the ecosystem.

Air pollution generated by human activities causes significant damage to ecological resources by directly damaging plants, acidifying lakes, streams, and soil, and leading to concentrations of poisons in the food chain. Acid precipitation, usually caused by sulfur dioxide and nitrogen dioxide combining with moisture in the air, is extremely harmful to terrestrial plants as well as organisms that live in rivers and lakes. Such pollution directly affects crops. Ground-level ozone over time reduces plant growth and fruit production and renders plants more vulnerable to disease and insect attack. Particulate matter can stunt plant growth by scattering the Sun's rays and reducing sunlight.

Introduction of Exotic Species

Exotic species are organisms that have been introduced into an area outside their normal distribution. Human individuals or agencies have both intentionally or accidentally introduced alien species into ecosystems, sometimes with devastating unintended consequences. The presence of such species can alter habitat, profoundly affect food chains with the introduction of new predators, affect the ability of native species to compete for food and habitat acquisition due to the lack of local controls (disease and predators) to keep populations of exotic species in check, promote hybridization – the interbreeding of native and nonnative species – resulting in the decline and extinction of native species, and introduce pest species that can cause disease or parasitic infestations. These impacts all combine to impact biodiversity by decreasing the number of native species, thus causing regional homogenization of ecosystems.

Climate

Scientists have documented the important role that climate plays in the geographic distribution of the world's ecosystems. Climate variability and change can affect plants and animals

in a number of ways, including their distributions, population sizes, physical structure, metabolism, and behavior. The Earth's climate is vastly different now from what it was 100 million years ago when dinosaurs roamed about the planet and tropical plants thrived closer to the poles. It is different from what it was 20,000 years ago when ice sheets covered much of the Northern Hemisphere. Although the Earth's climate will surely continue to change, climatic changes in the distant past were driven by natural causes, such as variations in the Earth's orbit, sunspots or volcanic eruptions. In 2007, the IPCC reported that the warming of our climate is unequivocal and that this increased warming is very likely due to human-caused increases in GHGs (Intergovernmental Panel on Climate Change (IPCC), 2007a). Current and future climatic changes thus have an additional source of change – human activities. Many activities associated with human economic development have changed our physical and chemical environment in ways that modify natural resources. These include activities that burn fossil fuels that release CO_2, use land for agriculture, or urbanization that causes deforestation. We are already feeling the climatic effects of having polluted the atmosphere with GHGs such as CO_2. Humans cannot directly rival the power of natural forces driving the climate – for example, the immense energy input to the Earth from the Sun that powers the climate. We can, however, indirectly alter the natural flows of energy enough to create significant climatic changes. The atmosphere has a natural capacity to trap radiant heat near the Earth's surface – the so-called greenhouse effect. This natural phenomenon acts as an atmospheric blanket, allowing solar energy to warm the climate by trapping long-wavelength radiant energy called terrestrial infrared radiation near the Earth's surface. The natural greenhouse effect is responsible for $\sim 33\,^{\circ}C$ ($60\,^{\circ}F$) of surface warming. But, human-induced emissions of GHGs are inadvertently modifying the climate by enhancing the natural greenhouse effect, and are projected to result in a global warming of 1.1–6.4 °C or higher in the twenty-first century (Intergovernmental Panel on Climate Change (IPCC), 2007a). Such an increase in global mean temperature, even at the lower end of the range, will likely result in ecologically significant changes, which is why climatic considerations are fundamental in the discussion of the status and trends of ecological conditions. Quantitative evaluations of the potential effect of human activities in creating global change are needed. Such evaluations are also central to potential policy responses to mitigate global changes (Schneider, 1997; Root and Schneider, 2005; Intergovernmental Panel on Climate Change (IPCC), 2001c, 2007c).

Synergisms

With continued and more severe changes in the climate, the ability of wildlife to adapt through migration and physiological change will be increasingly limited. The challenge is even greater when considered along with the broad range of other environmental threats currently affecting wildlife, such as habitat fragmentation and loss, environmental contamination, and invasive species. People fragment natural habitats for farmland, settlements, mines, roads or other developmental activities. As the climate changes, individual species of plants and animals will be forced to adjust if they can, as they have in the past. It seems unlikely that all of the migrating species that survived the Ice Age would be able to safely reach refuges by migrating across the freeways, agricultural zones, industrial parks, and cities of the twenty-first century. An additional complication arises with the imposition of the direct effects of changes in CO_2, which can change terrestrial, freshwater, and marine primary productivity, drop the pH of the oceans significantly, and alter the competitive relations among photosynthesizing organisms and among plants and animals.

One representative instance of synergism is that of the woodland caribou (Boutin *et al.*, 2011). This species is restricted to large undisturbed tracks of Boreal Forest. As humans invade the Boreal Forest, land-use changes have forced a contraction of the caribous' current range to about half the size of the historic range. Now with climate change, there is added concern about the increasing of frequent and severe forest fires. The contraction of the Boreal Forest may be faster than expected due to both human land-use change and severe fires, which in turn will cause the contraction of the range of the woodland caribou. If protection of the Boreal Forest is not implemented, then in the extreme situation this caribou species could face extinction. This potential for extinction indicates how the already high rate of extinctions throughout the world could be exacerbated by climatic changes occurring more rapidly than species can adapt (Pimm *et al.*, 1995, 2006; Pereira *et al.*, 2010; Barnosky *et al.*, 2011). Thomas *et al.* (2004) conclude that despite the many uncertainties that remain in projecting extinctions, "the consistent overall conclusions across analyses establish that anthropogenic climate warming is likely to be the greatest threat in many if not all regions," and that the most severe climate change impacts are likely to result from interactions among threats such as habitat loss, fragmentation, and competition with invasive species.

The synergism question raises a controversial management problem of anticipating global change risk and responding. One such response is setting up interconnected nature reserves or networks of adequate habitat allowing species to move as the need along habitat corridors to ensure against some species becoming extinct due to climatic changes. Alternatively, we could simply let the remnants of relatively immobile wildlife and natural plant communities remain in isolated reserves and parks as they exist now, hoping the species can survive in a steadily warming environment. If we opt for more environmental safeguards by interconnecting our parks and refuges, the question then becomes how we interconnect the nature reserves. The nature of corridors or networks is itself an interesting scientific question, for the network needs to invite species to move to another area but not be of sufficient quality to encourage breeding behavior within the corridor, which could cause populations settling in corridors to become sinks (i.e., not producing enough young to sustain a population) and not serve the purpose of sustaining biodiversity. Priorities must be set and money made available for constructing adequate networks through which species can travel. For example, elevated sections of highways may be needed to allow for migration routes, similar to what was done for the caribou in the Arctic when the Alaskan pipeline was built. To examine

such questions as these in scientific detail, it is first necessary to make a multidisciplinary examination of the subcomponents of the various aspects of climatology and ecology. The authors begin with a discussion of climatic history and climate processes important to wild species, followed by a discussion of examples of the interactions between climate and biological systems, particularly the changes in climate induced by human activities. Next, they discuss climate and ecological modeling and validation, focusing on possible synergisms among ecology and climate change. The authors conclude with recommendations for future research.

Climate History: What Has Happened

Scientists can reconstruct the cyclical expansion and contraction of polar caps and other ice masses from ice core samples taken from Greenland and Antarctica. When snow falls on high, cold glaciers, the air trapped between snow grains is eventually transformed into air bubbles as the snow is compressed into ice from the weight of subsequent accumulations. The ratio of two oxygen molecules with different molecular weights (O^{16} and O^{18} isotopes) is a proxy record for the temperature conditions that existed when the snow was deposited. From this, scientists have been able to determine that the ice buildup from 90,000 to 20,000 years ago was quite variable and was followed by a (geologically speaking) fairly rapid 10,000-year transition to the (current) climatically very stable Holocene Period. The Holocene is the 10,000-year warm interglacial period in which human civilization developed and modern plant and animal distributions evolved to their current states (Eddy and Oeschger, 1993). These ice cores also provide information on the presence of CO_2, an important greenhouse gas. Carbon dioxide as well as methane (CH_4), another greenhouse gas, were in much lower concentrations during cold periods than in interglacials. This implies an amplifying effect, or a positive feedback, because lower amounts of these gases during glacial periods means less-trapped infrared radiative heat, amplifying the cooling, and vice versa during interglacials. The ice cores also show that concentrations of CO_2 and CH_4 and temperature were remarkably constant for about the past 10,000 years (before AD 1700), particularly when compared with longer records. That relative constancy in chemical composition of the atmosphere held until the industrial age of the past two and half centuries.

The transition from Ice Age to the Holocene took 5000–10,000 years, during which time the average global temperature increased 5–7 °C and the sea level rose 100 m. Thus, we estimate that natural rates of warming on a sustained global basis are about 0.5–1 °C per thousand years. Such changes were large enough to have radically influenced where species live and to have potentially contributed to the well-known extinctions of woolly mammoths, sabertooth cats, and enormous salamanders.

A large interdisciplinary team of scientists, including ecologists, palynologists (scientists who study pollen), paleontologists (scientists who study prehistoric life, especially fossils), climatologists, and geologists formed a research consortium (Cooperative Holocene Mapping Project (COHMP), 1988; Wright *et al.*, 1993) to study the dramatic ecological changes accompanying the transition from Ice Age to the recent interglacial period. One group of these researchers used a variety of proxy indicators to reconstruct vegetation patterns over the past 18,000 years for a significant fraction of the Earth's land areas. In particular, cores of fossil pollen from dozens of sites throughout North America clearly shows how boreal tree pollen, now the dominant pollen type in the boreal zone in central Canada, was a prime pollen type during the last Ice Age (15,000–20,000 years ago) in what are now the mixed hardwood and Corn Belt regions of the US. During the last Ice Age, most of Canada was under ice; pollen cores indicate that as the ice receded, boreal trees moved northward chasing the ice cap. One interpretation of this information was that biological communities moved intact with a changing climate. In fact, Darwin (1859) asserted as much.

> "As the arctic forms moved first southward and afterward backward to the north, in unison with the changing climate, they will not have been exposed during their long migrations to any great diversity of temperature; and as they all migrated in a body together, their mutual relations will not have been much disturbed. Hence, in accordance with the principles inculcated in this volume, these forms will not have been liable to much modification."

If this were true, the principal ecological concern over the prospect of future climate change would be that human land-use patterns might block what had previously been the free-ranging movement of natural communities in response to climate change. The COHMP, however, incorporated multiple pollen types into its analyses, including not only boreal species but also herbs and more arid (xeric) species as well as oaks and other mesic species. They discovered that during the transition from the last Ice Age to the present interglacial, nearly all species moved north, as expected. During a significant portion of the transition period, however, the distribution of pollen types provided no analog or similar associations to today's vegetation communities (Overpeck *et al.*, 1992). Late-Quaternary global warming caused individual species distributions to change along environmental gradients in different directions, at different rates, and over different periods (Graham and Grimm, 1990). That is, whereas all species moved, they did so individual by individual, not as linked groups of species. Consequently, the groupings of species during the transition period were often dissimilar to those present today (**Figure 1**). Graham and Grimm (1990) observe that in light of the paleoecological record, which reveals that nonanalog climates produce nonanalog assemblages, individualistic changes can be anticipated in the future.

The relevance of this conclusion is that in the future, ecotypes will not necessarily move as a unit as climate changes (assuming there is enough time and space for such a migration). Credible predictions of vegetation changes from climatic changes projected from human activities cannot neglect transient (i.e., time-evolving) dynamics of the ecological system. Furthermore, caution should be exercised against relying too much on past conditions in forecasting future patterns resulting from global change forcings because the forecasted global average rate of temperature increase exceeds those rates typical of the last 120,000 years. Recently, ice core researchers have extended the analyses of the cores back over several glacial/interglacial cycles, finding very similar results to that of

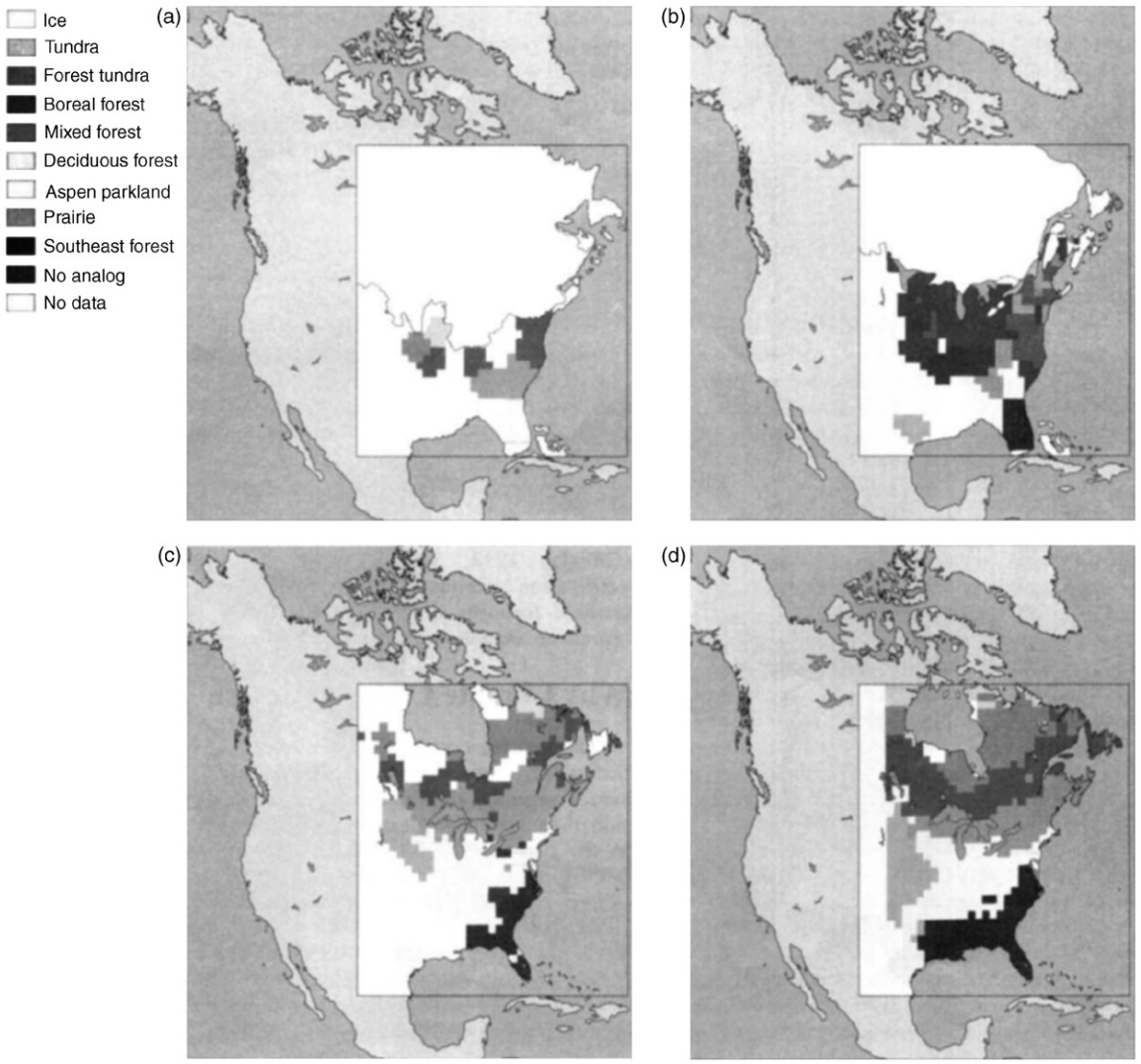


Figure 1 Selected paleovegetation maps reconstructed using the method of modern analogs and more than 13,000 samples of fossils and modern pollen: (a) 18,000 BP, (b) 12,000 BP, (c) 6000 BP, and (d) modern. No vegetation is mapped in areas without fossil pollen data, and no analog refers to vegetation without any modern analog. Data from Overpeck JT, Webb RS, and Webb III T (1992) Mapping eastern North American vegetation change over the past 18,000 years: No analogs and the future. *Geology* 20: 1071–1074.

the past cycle. If, as predicted, the rate of warming caused by the "greenhouse effect" is greater than in past events, then the individualistic responses may be even more profound. Therefore, possible future changes inferred from past changes can be taken only as a guide or a means to verify aspects of the forecasts of models of climate or ecosystem dynamics (Crowley, 1993; Schneider, 1993a, b). Such verification exercises may provide more credible forecasts of the effects of climatic change on animals.

Processes of Climate Change Important to Wild Species

Trends (Over Last 100 Years)

The Earth's climate system has demonstrably changed on both global and regional scales since the preindustrial era. Key observed twentieth century changes in the Earth's atmosphere, climate, and biophysical system include increases in atmospheric concentrations of greenhouse gases (GHGs), global mean temperature, precipitation, extreme weather events, and sea levels as well as diminishing ice, snow cover, and permafrost (Intergovernmental Panel on Climate Change (IPCC), 2001d).

CO_2 has increased ~36% above 1000–1750 levels, methane by ~160%, nitrogen dioxide by ~20%, tropospheric ozone by 35% (varying by region), with increases as well in hydrofluorocarbons (HFCs), and perfluorocarbons (PFCs).

Global mean surface temperature has risen 0.85 °C, with land areas warming more than the oceans. Northern Hemisphere surface temperature increased more during the twentieth century than during any other century in the last 1000 years. The 1990s was the warmest decade of the millennium, and the first decade of the twenty-first century has been even

warmer, with nine of the ten warmest years on record occurring in the twenty-first century. The diurnal temperature range has decreased over the years 1950–2000 over land, and it is probable (with a 66–90% confidence level) that at nighttime minimum temperatures increased at twice the rate of daytime maximum temperatures. Hot days increased whereas cold days decreased. It is very likely that continental precipitation increased by 5–10% over the twentieth century in the Northern Hemisphere, although it decreased in some regions. Heavy precipitation events increased at mid- and high-northern latitudes. Summer drying and associated incidence of drought have likely increased in a few areas, and in some regions including parts of Asia, Australia and Africa, there has been an observed increase in the frequency and intensity of droughts in recent decades.

Global mean sea level increased at an average annual rate of 1–2 mm during the twentieth century. It is very likely that the duration of ice covering rivers and lakes decreased by about 2 weeks over the twentieth century in mid- and high latitudes of the Northern Hemisphere. In September the Arctic sea-ice extent and thickness thinned as much as ~40% in the recent decade when compared to the thickness from 1979 to 2000, according to the National Snow and Ice Data Center. As well, there was a widespread retreat of nonpolar glaciers during the last century. Snow-cover area very likely decreased by 15% since global observations became available from satellites in the 1960s. Permafrost thawed, warmed, and degraded in parts of the polar, subpolar, and mountainous regions. El Niño events became more frequent, persistent, and intense during the last 20–30 years compared with the previous 100 years. Growing seasons lengthened by about 1–5 days per decade during the last 40 years in the Northern Hemisphere, especially at higher latitudes. Plant and animal ranges for many insects, birds, mammals, and fish have shifted poleward and up in elevation. The breeding, flowering, and migration of some plants, birds, and insects have shifted: Certain plant species are flowering earlier, many migrating birds are arriving in their spring locales and breeding earlier, and numerous insects are emerging earlier in the Northern Hemisphere. Coral reef bleaching has increased in frequency, especially during El Niño events.

Global inflation-adjusted weather-related economic losses have risen an order of magnitude over the last 40 years – an observed upward trend, which is linked to both socio economic factors (e.g., population growth, increased wealth, and urbanization in vulnerable areas) as well as to regional climatic factors (e.g., changes in precipitation and flooding events).

Importance of Extremes

Parmesan *et al.* (2000) state that climate drives biotic systems by affecting individual fitness, population dynamics, distribution and abundance of species, and ecosystem structure and function. In the absence of humans, broad-scale, long-term consequences of climatic warming on wild organisms are generally predictable. For example, evidence from Pleistocene glaciations indicates that most species responded ecologically by shifting their ranges poleward and upward in elevation, rather than through local adaptation (e.g., morphological changes). However, such broad patterns do not reveal the important part that extreme weather events play as mechanistic drivers of broad ecological responses when compared with the effects of gradual climatic trends. Parmesan *et al.* (2000) report evidence of the effects of extreme weather events on shaping the processes of biotic systems. Thus, they must be included in biological models used to project future changes.

With regard to Lepidoptera (butterflies and moths) and birds, extreme weather events appear to drive local population dynamics. Evidence of direct effects of extreme weather events include many Edith Checkerspot butterfly population extinctions associated with particular climatic events such as severe drought over California (Ehrlich *et al.*, 1980; Singer and Ehrlich, 1979; Ehrlich and Hanski, 2004) and extreme winter weather events in the California Sierra Nevada (Singer and Thomas, 1996; Thomas *et al.*, 1996). Two Sierra Nevada weather events produced a low winter snow pack, which triggered an early emergence of adult butterflies well before flowers were in bloom, which led to mass starvation of the adult population. During a third event, temperatures fell to $-5\,^{\circ}C$ on 16 June without insulating snowfall, which killed the host plants and led to the starvation of the young caterpillars. Infrequent and severe climatic events, which produced short-term responses at the population level, appear to have driven a gradual range shift in this butterfly species, which has shifted its range northward and upward during the twentieth century. According to Parmesan (1996), this shift is the result of increased population extinctions at its southern-range boundary and at lower elevations, with a symmetrical tendency toward population stability along the northern-range boundary and at the highest elevations.

Migratory bird populations have exhibited a shift in abundances in relation to weather – their abundances are higher at their southern-range areas during cold winters (in which there is an increase in the number of days below a given temperature threshold) and higher at their northern-range areas when the weather is warmer (Root, 1999).

Extreme weather events can also affect morphology, influencing shape, size, and color of individuals (Hadly, 1997). Impacts can affect more than one tropic level in a food web, as has been documented in the case of the seed-feeding finches in the Galapagos Islands, whose populations are dependent on changes in food availability and composition due to variable precipitation (Boag and Grant, 1984). For one finch species, wet years result in smaller billed individuals and drought years select for large ones, which leads to frequent shifts in the distribution of bill sizes.

Change in the incidence of extreme high or low temperatures through time can lead to highly skewed population sex ratios in the case of reptiles because sex is often determined during a critical phase of embryonic development (reviewed by Bull, 1980). However, sex ratios may no longer be self-adjusting, as evidenced by the current trends in many wild turtles toward smaller population sizes, fewer and more isolated protected nesting habitats that lead to less-successful dispersal, as well as limited choice of sites for nesting within a habitat. Thus, impacts of shifting climatic distribution in the past cannot be simply extrapolated to the future.

Weather events may provide a cue that affects the timing of breeding. For example, the beginning of the rainy season may signal the onset of a window of time for breeding. Climate variability can influence which group of individuals within a species breeds. In especially wet years when food resources are high, subordinate Galapagos mockingbird females breed opportunistically. In the case of African elephants, dominant males breed in the wet season and subordinate males in the dry season of a given year. A change in the intensity or duration of rainy versus dry seasons may change relative breeding rates and genetic structures within populations (Poole, 1989; Rubenstein, 1992).

Extreme weather can have major impacts on wild populations. Drought has caused crashes in diverse taxa of insect populations (Hawkins and Holyoak, 1998) and booms in populations of other insects (Mattson and Haack, 1987). Heat and cold stresses also can lead to population crashes, including aquatic systems (Huntsman, 1946; Bailey, 1955). Changes at the community level due to extreme events may also have large impacts. The 1982/83El Niño event caused widespread loss (bleaching) of coral reef communities in the eastern and central Pacific (Coffroth *et al.*, 1990), and hurricanes and typhoons have changed the composition of local communities with regard to the identity and abundance of species (Willig *et al.*, 1998; Willig and McGinley, 1999). Large weather events including hurricanes and typhoons, tornadoes, wildfires, floods, droughts, landslides, and even tree-falls can trigger ecosystem-level disturbances, affecting the organization (species composition and diversity) and functional attributes of systems (Walker, 1999; Walker and Willig, 1999). The initial degree of response of a system and the extent to which it can return to its original conditions after such a disturbance depends on the frequency, intensity, duration, and extent of the disturbance as well as the evolutionary history and extent of exotic invaders (Willig and Walker, 1999; Myers, 1996; Zimmerman *et al.*, 1996; Pickett and White, 1985).

In light of the reality of today's modern human-dominated landscape, in which natural ecosystems are increasingly confined to smaller and more isolated fragments and population sizes of wild nonweedy species have generally declined (Groombridge, 1992), natural systems have limited options by which to contend with the predicted rapid changes in climatic extremes or in frequency and intensity of disturbances. Genetic variation is often diminished by reduced population sizes, limiting the potential for local adaptation. Increases in distances between natural habitats, decrease successful dispersal, and hinder simple shifts of species' distributions. Increased fragmentation also lowers the probability of colonizing devastated areas after catastrophic disturbances. Colonists must travel farther, and their source populations within impoverished communities are smaller.

Modern systems have lowered resiliency to the types of nonlinear dynamic processes predicted by scenarios of global climate change (Schneider and Root, 1996; Easterling *et al.*, 2000; Meehl *et al.*, 2000). We need a more precise understanding of the relationships between biotic processes and extreme weather events to predict the impacts of future changes in the shape and position of the climatic distribution with more than modest confidence.

Examples of Ecological Responses to Climate Changes

Distribution Shifts

As the climate has changed a wide range of taxonomic groups have shifted their ranges because they are attempting to stay in areas where they are exposed to the same regional ambient temperatures, which generally means moving toward the pole and up in elevation. For example, the ranges of 39 butterfly species in Europe and North America have shifted up to 200 m northward over the last three decades (Parmesan *et al.*, 1999). Twelve bird species in Britain have moved an average of 18.9 km northward (Thomas and Lennon, 1999). Some range shifts involve community-level changes. Walther *et al.* (2002) and others point out that changes in distribution are often asymmetrical with species invading faster from lower elevations (Pounds *et al.*, 1999) or latitudes (King and Harangozo, 1998) than resident species shifting their ranges poleward or upslope, resulting in a transient species richness due to the different rates at which ranges shift.

In the early 1900s Joseph Grinnell systematically surveyed a 3000 m elevation gradient in Yosemite National Park. The same area was resurveyed in the early 2000s. Roughly half of the 28 mammals surveyed moved up in elevation by an average of around 500 m. For example, one of the species shifting up in elevation in response to climatic warming is the American pika (*Ochotona princeps*) (Moritz *et al.*, 2008). In 1900 Grinnell found it up to 2400 m, but in 2004 it was recorded as high as 2900 m. The same type of analyses with similar results has also been done on birds (Tingley *et al.*, 2009).

Abundance Shifts

The abundance or density of a species at a given location can change. For example, in Antarctic terrestrial ecosystems, macroscopic plants (largely mosses) are colonizing previously bare or newly exposed ground and the two higher plants on the continent are rapidly expanding their populations as well as their ranges (Convey, 2001; Smith, 2001; Kennedy, 1995). In tandem with the changes in vegetation, soil invertebrates are also colonizing (Walther *et al.*, 2002).

The Living Planet Index aggregates trends of some 3000 wild populations of species (from published sources), which are the result of multiple impacts associated with human activities. It shows a consistent decline in average species abundance of about 40% between 1970 and 2000; inland water species declined by 50%, whereas marine and terrestrial species both declined by around 30% (CBD, 2005). Similar trends have been observed for abundant and widespread bird species breeding on farmland throughout Europe, and for populations of species threatened by extinction. With regard to the abundance and distribution of selected species, most species with limited population size and distribution are being further reduced whereas some invasive species become more common.

Phenology

The timing of different seasonal activities (phenology) of many species is shifting in concert with the changing climate.

Common changes in the timing of spring activities include the timing of migration in various species such as birds, mammals, fish, and insect species (e.g., MacMynowski and Root 2007; Ogdan *et al.*, 2008). Meta-analyses have been performed to look for consistent patterns around the globe over a large number of species. For all northern hemisphere species reported in the literature with more than a 10 year record showing phenological changes in the spring, the average number of days changed in the spring observed over the last 30 years of the twentieth century was ~15.5 or ~5 days per decade earlier timing for those species showing a phenological change in the spring (Root *et al.*, 2003).

Over evolutionary time species have formed predator–prey relationships that are being disrupted because the predator and prey do not necessarily shift at the same time. For example, the common cuckoo (*Cuculus canorus*) is a migrant and lays its eggs in the nests of other birds that migrate in Europe to its breeding grounds from both short and long distances. With climate change the short-distance migrants are arriving significantly earlier now than in the past and they are nesting before the cuckoo has arrived. Consequently, the short-distant migrants seem to have experienced a way by which to reduce brood parasitism (Saino *et al.*, 2009).

Changes in Physical Structure, Genetics, Metabolism, and Behavior

Physical characteristics that are sensitive to climate include body size, hairiness, length of limbs, and skin or hair color. The average body trait for any given population can be genetically based; however, variation in the body trait of an individual animal can also be caused by variation in the environment experienced by the animal as it develops. Thus, the development of physical traits may involve genetic evolution or merely an immediate, plastic response to a changing environment. Many such traits appear to be affected by past climate changes (Parmesan and Galbraith, 2004).

Hadly (1997), Sullivan and Best (1997), Badgley (1998), and Smith and Betancourt (1998) have documented body size changes of small mammals, lizards, and zooplankton during historical (natural) climate shifts. In a phenomenon first associated with latitudinal trends in temperature now known as "Bergman's rule," bodies tend to become smaller with general warming and larger with cooling (Bergmann, 1847) with relatively few exceptions. For example, temporal fluctuations of body size in wood rats are so precise that Smith and Betancourt (1998) labeled them "paleothermometers" because changes in their body sizes are so closely aligned with changes in temperature (Parmesan and Galbraith, 2004). A study by Wolf *et al.* (2009) on two rodents that were captured in southern New Mexico from 1991 to 1998 found strong correlation with a time-appropriate climate variable and hind leg, ear, and body length.

Extinction

As a group the background extinction rate for mammals was estimated using the fossil record, and it is less than one species each millennium (May 2010). Both mammals and birds have been well studied throughout the past century and their extinction rate over that time period has been around one extinction per year. This rate is up roughly 1000 times the background rate for mammals (May 2010). Attempts have been made to try to understand how many and which species may face extinction in the future.

Range expansions of species at high elevations and latitudes in response to climate-change-generated impacts have been widely documented, but Wilson *et al.* (2005) observe that there has been surprisingly little evidence for contractions at warm margins (Parmesan *et al.*, 1999; Thomas and Lennon, 1999; Hill *et al.*, 2002), even though the first expected symptoms of a climate-induced biodiversity crisis are range contractions and extinctions at lower elevational and latitudinal limits to species distributions. Wilson *et al.* (2005) show that lower elevational limits for 16 butterfly species in the Sierra de Guadarrama (central Spain), a mountain range that represents the lower latitudinal or elevational boundary for many species of butterfly (Gómez de Aizprua, 1987; Kudrna, 2002; García-Barros *et al.*, 2004) have risen on average by 212 m (± SE 60) in 30 years. This rise, which accompanies a 1.3 °C rise (equivalent to about 225 m) in mean annual temperature, signifies, on average, a 33% reduction in habitable area. They point out that these declines might be underestimated if there is a time lag in extinctions from sites that are outside the long-term climatic tolerances of the species (Hill *et al.*, 2002). In addition, they project regional habitat losses of 50–80% for the twenty-first century, given maintenance of the species climatic associations and continued regional warming – losses that will, in all probability, be exacerbated by habitat loss resulting from land-use changes. The magnitude of these declines in available habitat may possibly be typical for species that are restricted to high elevations at their warm range margins (Stefanescu *et al.*, 2004), or for endemic taxa that are entirely restricted to mountainous regions (Pounds *et al.*, 1999; Williams *et al.*, 2003). Because suitable high elevations for these species exist in spatial isolation, they are unable to colonize new areas as the climate warms (Grabherr *et al.*, 1994; Pounds *et al.*, 1999; Midgley *et al.*, 2002; Klanderud and Birks, 2003; Konvicka *et al.*, 2003; Williams *et al.*, 2003). Unless climate change and habitat loss can be stopped, widespread extinctions may follow.

Another recent study (Pounds *et al.*, 2006) that implicates climate change as a factor in species extinctions proposes that there is a link between recent climate changes and the spread of a pathogenic chytrid fungus, *Batrachochytrium dendrobatidis*, which appears to have caused a mass extinction of frog and toad species in the mountains of Costa Rica. Seventeen years ago, the Monteverde harlequin frog (*Atelopus* sp.) and the golden toad (*Bufo periglenes*) vanished. According to Pounds *et al.*, this pathogenic chytrid fungus is implicated in the disappearance of an estimated 67% of the 110 or so species of Atelopus, which are endemic to the American tropics. After analyzing the timing of species losses in relation to changes in sea-surface temperature and air temperature, Pounds *et al.*, posits with very high confidence that large-scale warming is a key factor in the disappearances. They observed two patterns with regard to a link between climate-induced temperature rise and the chytrid fungus: Firstly, just as the lowlands are often

too warm for the chytrids during the day, the highlands are often too cool for them at night. Secondly, most Atelopus extinctions have occurred at elevations where the minimum temperature is shifting toward the growth optimum for these pathogens. Thus, Pounds *et al.* propose the chytrid-thermal-optimum hypothesis, in which daytime cooling (local or microscale) and nighttime warming accelerate disease development, observing that the impacts at night may explain the association with warm years and thereby resolve the climate–chytrid paradox. These findings have garnered mixed reviews from experts on amphibian disease and ecology with some saying that they provide important evidence of another impact that climate change is having on wildlife, whereas others say that there are uncertainties that have not been sufficiently addressed. In this, however, as with all complex problems, in the early phases of learning one looks for multiple lines of argument and when they converge with a basic theory, it increases confidence in a connection.

Meta-Analyses

Meta-analyses provide methods for combining results from various studies, whether statistically significant or not. Results from meta-analyses determine whether there is a consistent "signal" or "fingerprint" among the studies. The balance of evidence from two such meta-analyses done on species from many different taxa examined at disparate locations around the globe (Root and Schneider, 2002; Walthers *et al.*, 2002; Parmesan and Yohe, 2003; Root *et al.*, 2003) suggests that a significant impact from recent climatic warming is discernible in the form of long-term, large-scale alteration of animal and plant populations. Though these studies had some significant methodological differences, their basic conclusions were very similar.

Root *et al.* (2003) gathered information on species and global warming from 143 studies, each of which spanned at least 10 years, showed a trait of at least one species that exhibited change over time, and that found either a temporal change in temperature at the study site or a strong association between the species trait and site-specific temperature. To document a strong role for climate change in explaining many of the observed changes in animal and plant populations, Root *et al.* looked for repeated examples occurring over long temporal and broad spatial scales that showed unidirectional changes predicted by the current understanding of the physiological tolerances of species to temperature. The predicted result, or fingerprint, of an underlying consistent shift in a large-scale pattern shown by many species around the globe, coupled with an understanding of the possible causal mechanisms, provides confidence in attributing observed species changes to climatic change.

The Root *et al.* study shows that recent temperature change has already had a marked influence on many species. Their meta-analyses revealed a consistent temperature-related shift in species ranging from mollusks to mammals and grasses to trees with more than 80% of the species showing changes shifting in the direction that would be expected based on species' known physiological constraints. For example, their analyses showed an average shift in spring phenology (timing) of events, such as breeding or blooming, for temperate-zone species of 5.1 ± 0.1 days earlier in a decade and strongly suggest that species at higher altitudes are reacting more strongly to more intense changes in temperature.

The conclusion that significant impact from recent climatic warming is discernible in the form of long-term, large-scale alteration of animal and plant populations supported by the meta-analyses discussed above was extended by the IPCC (Intergovernmental Panel on Climate Change (IPCC), 2001b, Chapter 4) to include "environmental systems" – sea- and lake-ice cover and mountain glaciers. Clearly, if such climatic and ecological signals are now being detected above the background of climatic and ecological noise for a twentieth century warming of only 0.6 °C, it is likely that the expected impacts on ecosystems of temperature increases up to an order of magnitude larger by 2100 will be dramatic.

Joint Attribution

The meta-analyses discussed above have established that many plant and animal species are responding to regional climatic changes, but can the extent to which such regional warming is natural variation as opposed to attributable to human activities be teased apart? If there is a discernible impact of human activities on climate (Intergovernmental Panel on Climate Change (IPCC), 1996a, b, c, 2001a, b, c, 2007a, b, c) and a discernible impact of climate on plants and animals (e.g., Intergovernmental Panel on Climate Change (IPCC), 2007b), can it be asserted that there is a discernible impact of human-induced climate change on plants and animals, the so-called joint attribution? Root *et al.* (2005) correlated Northern Hemispheric phenological responses of plants and animals reported to be changing over the past 50 years to spring temperature data produced by the HadCM3 climate model. This model was driven by three sets of observed potential causes: (1) only natural forcings (e.g., solar activity and volcanic dust veils), (2) only anthropogenic forcings (e.g., CO_2 and aerosol increases), and (3) combined forcings (i.e., natural and anthropogenic forcings together). Given the many uncertainties and missing factors, it is not expected that most of the variance of observed phenological changes in the past 50 years can be attributed to anthropogenic forcings. Thus, the key question is as follows: Did the correlations between observed plant and animal phenological records improve when anthropogenic forcing was driving the models relative to the correlations when only natural forcings produced the climate-model records? The Root *et al.* (2005) results show a clear signal that, despite known uncertainties and missing factors, springtime temperatures driven by anthropogenic forcing are much more highly correlated with observed phenological changes in plants and animals than are temperatures driven by natural forcing. Human activities are highly likely to be contributing to the changes in regional surface temperatures, and these human-influenced temperature patterns are significantly associated with discernible changes in plant and animal phenological traits.

Climate Modeling: Validation and Forecasts

Constructing Climate Models

The evidence outlined above paints a compelling picture of widespread human-induced climate change impacts on plant and animal species. As mentioned above, these changes are expected to continue and probably accelerate during the twenty-first century. Projections of future ecological changes, informed by projections of future climate changes, are essential to meet the challenge posed by the multiple stressors on biodiversity outlined above.

To predict the ecologically significant ways in which the climate might change, one must estimate a plausible set of assumptions about what people do that modifies how energy is exchanged among the atmosphere, land surface, and space, because such energy flows are the driving forces behind climate. Such driving forces include projections of population, consumption, land use, and technology. The next step is to estimate the response of the various components of the earth system to such societal forcings. The earth system consists of the following interacting subcomponents: atmosphere, oceans, cryosphere (snow, seasonal ice, and glaciers), and land-surface (biota and soils) systems.

Research in the field and in laboratories provides an understanding about various processes affecting the subcomponents of the earth system. This understanding can be put into mathematical expressions that, when combined, form a model of the behavior of particular components of the earth system. In practice, models of the atmosphere are connected to models of the oceans, ice, biota, and land surfaces to simulate the consequences of a scenario of societal forcing on climate and ecosystems. Controversy arises because both the societal forcing that will actually occur and the scientific knowledge of each subsystem are incomplete. Because models cannot be perfect replicas of the actual natural system, scientists must expend considerable efforts to test their models against the expanding base of field and laboratory data. This not only allows them to assess the credibility of current simulations but it also reveals improvements for the next generation of models.

Scientists who estimate the future climatic changes that are relevant to ecosystems have focused on general circulation models (GCMs) that attempt to represent mathematically the complex physical and chemical interactions among the atmosphere, oceans, ice, biota, and land. As these models have evolved, increasingly more information has become available, and more comprehensive simulations have been performed. Nevertheless, the complexities of the real climate system still vastly exceed the GCMs and the capabilities of even the most advanced computers (Intergovernmental Panel on Climate Change (IPCC), 1996a, 2001a, 2007a).

Verifying Climate Forecasts

The most perplexing question about climate models is whether they can be trusted as a reliable basis for altering social policies, such as those governing CO_2 emissions or the shape and location of wildlife reserves (other than those based on no-regrets strategies). Even though these models are fraught with uncertainties, several methods are available for verification tests. Although no method is sufficient by itself, several methods together can provide significant, albeit circumstantial, evidence of a forecast's credibility.

The first validation testing method involves checking the model's ability to simulate the current climate. The seasonal cycle is a good test because temperature changes in a seasonal cycle are larger on a hemispheric average than the change from an ice age to an interglacial period (i.e., 15 °C seasonal range in the Northern Hemisphere vs 5–7 °C glacial–interglacial cycle). GCMs map the seasonal cycle well. This supports the scientific consensus about the plausibility of global warming of several degrees in the twenty-first century. The seasonal test, however, does not indicate how well a model simulates slow processes such as changes in deep ocean circulation, ice cover, forests, or soil carbon storage, which may have important effects on the decade- to century-long timescales over which atmospheric CO_2 is expected to at least double.

A second verification technique involves isolating individual physical components of the model and testing them against actual data. A reasonable model should reproduce the flow of thermal energy among the atmosphere, the surface, and space with no more than about a 10% error. Together, these energy flows make up the well-established natural greenhouse effect on Earth and constitute a formidable and necessary test for all models. A model's performance in simulating these energy flows is an example of the physical validation of model components.

A third validation method involves the model's ability to reproduce the diverse climates of the past. This method is aided by recording instrumental observations made during the past few centuries and paleorecords that serve as a proxy for climatic conditions of the ancient Earth, or even include testing the models' ability to simulate climates of other planets (Kasting *et al.*, 1988). Paleoclimatic simulations of the Mesozoic (Age of the Dinosaurs), glacial–interglacial cycles, or other extreme past climates help scientists understand the coevolution of the Earth's climate and living things (Schneider and Londer, 1984). As verification tests of climate models, they are also crucial to predicting future climates and changes in biological systems.

The joint attribution study mentioned earlier (Root *et al.*, 2005) also served as a verification exercise, since that work tested the correlation of the model-generated temperature records versus phenological patterns of plants and animals locally, regionally, and hemispherically. Root *et al.* (2005) found that, even at the local level, the HadCM3 computer model predicted much more of the variance of the observed species records for anthropogenic forcing than for natural forcing, and even more of the variance for a combination of natural and anthropogenic forcing. This is a powerful proxy technique for model verification at even a local scale – the level for which climate modelers are most cautious to claim skill levels.

Downscaling Climate Predictions to Regional Effects

Even the highest resolution, three-dimensional GCMs do not have a grid with nodes less than 100 km apart; individual clouds and most ecological research (to say nothing of cloud

droplets) occur on scales far smaller. Therefore, GCMs will not be able to resolve the local or regional details of weather affecting most local biological communities or the importance of regional effects of hills, coastlines, lakes, vegetation boundaries, and heterogeneous soil (Root and Schneider, 1993). Nonetheless, it is important to have climatic projections and ecological-response analyses on the same physical scales.

What is most needed to evaluate potential biological effects of temperature change is a regional projection of climatic changes that can be applied to ecosystems at a regional or local scale. Analyses of large, prehistoric climatic changes (Barron and Hecht, 1985; Budyko *et al.*, 1987; Schneider, 1987; Cooperative Holocene Mapping Project, 1988) and historical weather analogs (Pittock and Salinger, 1982; Jager and Kellogg, 1983; Lough *et al.*, 1983; Shabalova and Können, 1995) provide some insights into such changes. Historical weather analogs, however, since they are empirically and statistically based, rely on climatic cause-and-effect processes that probably differ from those that will be driven by future greenhouse gas radiative effects (Schneider, 1984; Mearns *et al.*, 1990; Crowley, 1993). Consequently, ecologists turn to climatic models to produce forecasts of regional climatic changes for the decades ahead. How credible are such forecasts?

Statistical Downscaling Techniques

Techniques exist that can translate the output of climate models so that it is closer to that of most ecological scales. One method that uses actual climatic data at both large and small scales can help provide maps that may allow small-scale analysis of large-scale climate change scenarios. For example, the Sierra Nevada of California or the Cascades in the northwestern US are north–south mountain chains east–west dimensions smaller than the grid size of a typical GCM. In the actual climate system, onshore winds on the Pacific coast would produce cool upslope and rainy conditions on the western slope and a high probability of warmer and drier conditions associated with that flow pattern on the downslope or eastern slope.

Researchers have developed statistical downscaling techniques that link large-scale climatic patterns to small-scale climatic conditions. For example, the analysis of Hayhoe *et al.* (2004) relied on a deterministic method in which grid-cell values of temperatures and precipitation from two GCMs were used to produce weather-station-level data for California, by rescaling the data using simple monthly regression relations to ensure that the overall probability distributions of the simulated daily values closely approximated the observed probability distributions at selected long-term weather stations (Dettinger *et al.*, 2004). The same regression relations were then applied to future simulations, such that rescaled values share the weather statistics observed at the selected stations. These methods allow the reproduction of the dominant mode of variation for this area (warm and dry on one side of the mountains and cold and wet on the other side). Although such a statistical downscaling technique seems appropriate for translating low-resolution, grid-scale climate-model forecasts to local applications, a strong caveat must be offered. That is, the processes in the climate system that give rise to internal variability or natural fluctuations are not necessarily the same processes that would give rise to local deviations from large-scale patterns if the climate change were driven by external forces rather than an internal variation of the system. The conditions produced using such a method reflect the most probable regional situation for today's naturally fluctuating climate. However, if 50 years from now, as the result of ongoing emissions producing an enhanced downward infrared radiative heating, then both eastern and western slopes of the Sierra Nevada would probably experience warming, and an entirely different climatic change pattern linking large- and small-scale processes would probably occur. This pattern would be different than that obtained from statistical downscaling techniques using the naturally varying weather conditions existing today rather than the anthropogenically forced conditions of the twenty-first century.

Therefore, techniques to downscale climate forecasts that use current distributions of environmental variables at local scales and correlate them with current large-scale regional patterns may not necessarily provide a good guideline about how large-scale patterns would be distributed regionally. This is especially true if there are changes in prevailing winds, seasonality or cloud cover. The reason is that the causes of the future change may be physically or biologically different from the causes of the historical fluctuations that led to the empirical maps in the first place. This caveat warns scientists to use caution before adopting such empirical techniques for global change applications.

There are practical limits to downscaling, especially using global techniques. Downscaling to a weather station, assuming that the weather-station data are accurate (a dubious assumption in many parts of the world and the above caveats are heeded), provides useful information at the level of the point. Gridded weather data are often used when weather-station data are not available. These grids are typically $0.5^\circ \times 0.5^\circ$ of latitude and longitude in size, but the uncertainties in gridding areas with sparse data and with large elevational changes must be recognized. The standard gridded data set used in many studies is the Climate Research Unit (CRU) TS data (University of East Anglia Climate Research Unit). In some areas with large numbers of weather stations (e.g., US, Europe, Australia) smaller grids (approximately 10 km^2) have been developed (e.g., VEMAP, OzClim). Although it is theoretically possible to downscale to such small scales and even smaller ones (1 km^2 based on digital elevation model data), caution must be taken when using any of these gridded data. At a large scale (50 km^2) the differences in slopes of mountains are rarely picked up. At 1 km^2, elevation is important, but slopes, aspect and, especially wind direction and water transport are essential when calculating lapse rate to estimate temperature and vital for precipitation estimates.

Regional-Scale Models with GCM Inputs (Dynamical Downscaling)

Other techniques can still translate large-scale patterns to smaller scales, but these techniques are based on known processes rather than empirical maps for today's conditions. One such technique is to drive a high-resolution, process-based model for a limited region with the large-scale patterns produced by a GCM. In essence, this approach uses a

mesoscale model (i.e., 10–50-km^2 grid cells) based on physical laws to solve the problem of translating GCM grid-scale averages into a finer scale mesh much closer to the dimensions of most ecological applications. Of course, even this mesoscale grid will still be too coarse to assess many impacts, necessitating further downscaling techniques. Neither are the problems of GCMs entirely eliminated by mesoscale grids because they are still bigger than individual clouds or trees. However, such methods do bring climate-model scales and ecological-response scales much closer.

Ecological Modeling Based on Future Climate Projections

Bringing climatic forecasts to ecological applications at local and regional scales is one way to bridge the scale gap across ecological and climatological studies. Ecologists, however, have also analyzed data and constructed models that apply over large scales, including the size of climatic model grids. A long tradition in ecology has associated the occurrence of vegetation types or the range limits of different species with physical factors such as temperature, soil moisture, land–sea boundaries, or elevation. Biogeography is the field that deals with such associations, and its results have been applied to estimate the large-scale ecological response to climate change.

Methods

Top–Down Approaches

The biogeographic approach is an example of a top–down technique (e.g., Holdridge life zone classification), in which data on abundances or range limits of vegetation types or biomes are overlain on data of large-scale environmental factors such as temperature or precipitation. When associations among large-scale biological and climatic patterns are revealed (Root, 1988b), biogeographic rules expressing these correlations graphically or mathematically can be used to forecast changes in species driven by given climate changes.

Top–down approaches suffer because of the possibility that the discovered associations at large scales are statistical artifacts that do not, even implicitly, reflect the causal mechanisms needed for reliable forecasting. As Jarvis (1993, p. 121) stated, "A major disadvantage of a top–down model is that predictions cannot be made safely outside the range of the variables encountered in the derivation of the lumped parameter function."

Bottom–Up Approaches

Another traditional analysis and forecasting technique is often referred to as bottom–up. Small-scale ecological studies have been undertaken at the scale of a plant or even a single leaf (Idso and Kimball, 1993) to understand how, for example, increased atmospheric CO_2 concentrations might directly enhance photosynthesis, net primary production, or water-use efficiency. Most such studies indicate increases in all these factors – increases that some researchers have extrapolated to ecosystems (Idso and Brazel, 1984; Ellsaesser, 1990).

However, at the scale of a forest, the forest regulates the relative humidity within the canopy, which significantly influences the evapotranspiration rate. In other words, if an increase in water-use efficiency decreased the transpiration from each tree, the aggregate forest effect would be to lower relative humidity. This, in turn, would increase transpiration, thereby offsetting some of the direct CO_2/water-use efficiency improvements observed experimentally at the scale of a single leaf or plant. Regardless of the extent to which this forest-scale feedback effect will offset inferences made from bottom–up studies of isolated plants, the following general conclusion emerges: the bottom–up methods may be appropriate for some processes at some scales in environmental science, but they cannot be considered credible without some sort of validation testing at the scale of the system under study. For bottom–up models, the primary problem is that some of the most conspicuous processes observable at the smaller scales may not be the dominant processes that generate large-scale patterns. In this connection, multifactorial studies were conducted at Jasper Ridge in California, in which a grassland ecosystem was subjected to enhanced CO_2, enhanced CO_2 and temperature, enhanced CO_2 and temperature and moisture stress, and enhanced CO_2 and temperature and moisture stress and nutrient stress. What was observed was that this multifactorial analysis greatly reduced the benefits of CO_2 enhancements observed in single species, one-factor experiments. Although it is premature to over generalize, the implication is clear – single species or single-factor experiments are not likely to be very predictive for actual responses of ecosystems to climate and CO_2 changes in the decades ahead.

Combined Top–Down and Bottom–Up Approaches

To help resolve the deficiencies of the sole use of top–down biome forest models mentioned previously, more process-based, bottom–up approaches such as forest gap models have been developed (Botkin *et al.*, 1972; Pastor and Post, 1988; Smith *et al.*, 1992). These models include individual species and can calculate vegetation dynamics driven by time-evolving climatic change scenarios. However, the actual growth rate calculated in the model for each species has usually been determined by multiplying the ideal growth rate curve by a series of growth-modifying functions that attempt to account for the limiting effects of nutrient availability, temperature stress, and so forth. These growth-modifying functions for temperature are usually determined empirically at a large scale by fitting an upside-down U-shaped curve, with maximum temperature midway between the average temperature of the species' northern-range limit and the average temperature of its southern-range limit. Growing degree-days (related to temperature but not temperature *per se*) are used in this scenario.

In essence, this technique combines large-scale, top–down empirical pattern correlations into an otherwise mechanistic bottom–up modeling approach. Although this combined technique refines both approaches, it has been criticized because such large-scale, top–down inclusions are not based on the physiology of individual species and lead to confusion about the fundamental and realized niches (Pacala and Hurtt, 1993). (The fundamental niche is the ecological space in which a given species could theoretically survive – e.g., if its

competitors were absent – and the realized niche is where it actually exists.) The question is then, what limits the realized niche, particularly at the southern boundary? Furthermore, more refined models should include factors such as seed dispersal so that plant recruitment is related to the preexisting population and is not simply the result of a random number generator in the computer code.

Studies of More Refined Approaches

As noted previously, problems with the singular use of either top–down or bottom–up methods have led to well-known criticisms. A search of the literature (Wright *et al.*, 1993; Root, 1994; Harte and Shaw, 1995) provides examples of a refined approach to analyzing across large and small scales, which Root and Schneider (1995, 2003) labeled strategic cyclical scaling. This method builds on the combined techniques in which top–down and bottom–up approaches are applied cyclically in a strategic design that addresses a practical problem: in our context, the ecological consequences of global climatic change. Large-scale associations are used to focus small-scale investigations; this helps ensure that tested causal mechanisms are generating the large-scale relations. Such mechanisms become the laws that allow more credible forecasts of the consequences of global change disturbances. Levin (1993, p. 14) observed: "Although it is well understood that correlations are no substitute for mechanistic understanding of relationships, correlations can play an invaluable role in suggesting candidate mechanisms for (small-scale) investigation."

Strategic cyclical scaling, however, is intended not only as a two-step process but also as a continuous cycling process between large- and small-scale studies, with each successive investigation building on previous insights from all scales. This approach is designed to enhance the credibility of the overall assessment process (Vitousek, 1993; Harte and Shaw, 1995), which is why strategic is the first word in strategic cyclical scaling.

If the rate at which humans are injecting GHGs into the atmosphere is not greatly decreased, the Earth's climate will most likely warm by several degrees Celsius by 2050 (Intergovernmental Panel on Climate Change (IPCC), 2007a). With this in mind, Root (1988a, b) examined the biogeographic patterns of all wintering North American birds. She chose this group of species because birds are important parts of ecosystems and because of the availability of the necessary data. The National Audubon Society and the US Geological Survey have volunteer forces amassed to aid in the collection of Christmas Bird Count data and Breeding Bird Survey data, respectively. Using Christmas Bird Count data, Root determined that for a large proportion of species, average distribution and abundance patterns are associated with various environmental factors (e.g., the northern-range limits of some species are apparently limited by average minimum January temperature; Root, 1988a, b, 1989; Repasky, 1991).

The following is the scaling question: What mechanisms (such as competition or thermal stress) at small scales may have given rise to the large-scale associations? Root first tested the hypothesis that local physiological constraints may be causing most of the particular large-scale, temperature-range boundary associations. She used published small-scale studies on the wintering physiology of key species to determine that about half of the songbird species wintering in North America extend their ranges no farther north than the regions where, to avoid hypothermia during winter nights, they need not increase their metabolic rates more than ~ 2.5 times their basal metabolic rate (Root, 1988b). Root embarked on a larger, regional study to determine whether the longer nights – hence, fewer hours of daylight available for foraging – or the colder temperatures in the more northerly locations are relatively more important. Preliminary results indicate that changing temperatures are more likely than day length to explain this effect (Root, unpublished data): thus, global temperature changes are predicted by this set of analyses to cause a rapid range and abundance shift, at least by selected bird species. Indeed, Root initially found significant year-to-year shifts in ranges and abundances; these shifts are apparently associated with year-to-year changes in winter temperatures.

Then, much more recently, the meta-analyses of Parmesan and Yohe (2003) and Root *et al.* (2003), as well as the joint attribution study of Root *et al.* (2005) – and the more recent phenological meta-analysis of Menzel *et al.* (2006) – clearly demonstrate a consistent shift of species in response to the warming of the late twentieth century, where the vast majority of species observed to be changing in the direction expected from physiological ecology: poleward, up mountains, and advances in spring phenology.

Progress toward a way of accounting for the impact of biological interactions on species distribution was recently made by incorporating projections of marine primary productivity into a dynamic bioclimatic envelope model (Cheung *et al.*, 2009). Inclusion of indirect impacts of climate change on a species' distribution by factoring in the direct impacts on prey species is likely to greatly improve their accuracy due to the more responsive nature of lower trophic levels to environmental variation (Ottersen *et al.*, 2010). A new approach using combined behavioral and biophysical modeling has also drawn attention to the high risks to terrestrial ectotherms such as insects, frogs, lizards, and turtles in tropical and desert regions such as Australia (Kearney *et al.*, 2009).

Studies

Predicting Vegetation Responses to Climate Change

The classical Holdridge (1967) life zone classification assigns biomes (e.g., tundra, grassland, desert, or tropical moist forest) according to two measurable variables – temperature and precipitation. Other more complicated large-scale formulas have been developed to predict vegetation patterns from a combination of large-scale predictors (e.g., temperature, soil moisture, or solar radiation); vegetation modeled includes individual species (Davis and Zabinski, 1992), limited groups of vegetation types (Box, 1981), or biomes (Prentice *et al.*, 1992; Melillo *et al.*, 1993; Neilson, 1993). These kinds of models predict vegetation patterns that represent the gross features of actual vegetation patterns, which are incentives used to predict vegetation change with changing climate, but they have some serious drawbacks as well. That is, they are typically static, not time-evolving dynamic simulations, and thus cannot capture the transient sequence of changes that

would take place in reality. In addition, such static biome models occasionally make "commission errors;" that is, they typically predict vegetation types to occur in certain zones where climate would indeed permit such vegetation, but other factors such as soils, topography, or disturbances such as fire actually preclude it. However, newer dynamic general vegetation models (DGVMs) have started to include disturbance regimes. These DGVMs, which work at the level of the plant functional type (PFT, similar to biomes), have been used to look at potential changes in net primary productivity as well as changes in plant groups over time. One of the problems with using DGVMs in a predictive sense is that they are typically "on" or "off" – that is, they show what the current PFT might be and what the future PFT might be but they lack a transition state. This can be misleading for policy makers as it misses the critical information both as to the longevity of some of these PFTs in the absence of disturbance (i.e., reproduction stops but individuals persist for the lifetime of the organisms, potentially several hundred years) as well as the slow rate of dispersal for many plant species. Thus a transition state, matching neither the current nor future PFT is likely. This can be overcome, to some extent by looking at individual species but the large number of species involved, until recently, has made it difficult to look at changes at this scale. The access of large clusters of computers and global datasets, however, has meant that large-number-of-species models can now be run. For example, one particular effort called the Wallace Initiative has modeled almost 50,000 terrestrial plant and animal species. Furthermore, local patterns may influence vegetation dynamics at scales not captured in some simulations, and seed germination and dispersal mechanisms are also either not explicitly simulated or simulated only crudely with such models. It is remarkable that they are still able to produce generalized maps of vegetation types that do indeed resemble current or even paleoclimatic patterns in a broad sense, but their details do not provide confident projections for future vegetation states. Fortunately, progress is being made to include some of the deficiencies mentioned previously, and the so-called dynamical global vegetation models are being developed to treat the transient nature of vegetation change that would probably accompany climatic change.

Using the integrated assessment model IMAGE to link different ecological impacts to different levels of climate change, Leemans and Eickhout (2004) have predicted the impacts of climate change on terrestrial species, ecosystems, and landscapes over a range of increasing global mean temperatures and the corresponding temperature and precipitation patterns for the twenty-first century. Although there are large regional differences, their analysis shows that even small increases in global mean temperatures will considerably impact many species, ecosystems, and landscapes – especially such temperature-limited ecosystems as tundra. At an increase of between 1 °C and 2 °C in global mean temperatures, most species, ecosystems, and landscapes will be affected and adaptive capacity will become limited.

At a rate of warming of 0.1 °C per decade (i.e., 1 °C global mean temperature increase by 2100), Leemans and Eickhout's analysis predicts that 50% of all ecosystems will be able to adapt within a century, but only 36% of all affected forests. At a rate of warming of 0.3 °C per decade (i.e., 3 °C global mean temperature increase in 2100), only 30% of all affected ecosystems will be able to adapt and only 17% of all affected forests. Leemans and Eickhout observe that with the current rate of climate change (which is >0.2 °C per decade), a decline in biodiversity will accelerate and simultaneously many ecosystem services will become less abundant.

Predicting Animal Responses to Climate Change

Scientists of the US Geological Survey, in cooperation with Canadian scientists, conduct the annual North American Breeding Bird Survey, which provides distribution and abundance information for birds throughout the US and Canada. From these data, collected by volunteers under strict guidance from the US Geological Survey, shifts in bird ranges and abundances can be examined.

Because these censuses were begun in the 1960s, these data can provide a wealth of baseline information. Price (1995) used these data to examine the birds that breed in the Great Plains. Using the present-day ranges and abundances for each of the species (**Figure 2(a)**), Price derived large-scale, empirical–statistical models based on various climate variables (e.g., maximum temperature in the hottest month and total precipitation in the wettest month) that provided estimates of the current bird ranges and abundances (**Figure 2(b)**). Then, using a GCM to forecast how doubling of CO_2 might affect the climate variables in the models, he applied the statistical models to predict the possible shape and location of the birds' ranges and abundances (**Figure 2(c)**).

Significant changes were found for nearly all birds examined. The ranges of most species moved north, up mountain slopes, or both. The empirical models assume that these species are capable of moving into these more northerly areas, provided habitat is available and no major barriers exist. Such shifting of ranges and abundances could cause local extinctions in the more southern portions of the birds' ranges, and, if movement to the north is impossible, extinctions of entire species could occur. We must bear in mind, however, that this empirical–statistical technique, which associates large-scale patterns of bird ranges with large-scale patterns of climate, does not explicitly represent the physical and biological mechanisms that could lead to changes in birds' ranges. Therefore, the detailed maps should be viewed only as illustrative of the potential for very significant shifts with different possible doubled CO_2 climate change scenarios. More refined techniques that also attempt to include actual mechanisms for ecological changes are discussed later. Many think that birds may be able to keep up with rate of climate change because they can fly. However, rates of natal dispersal should not be confused with the expansion capability of populations.

Reptiles and amphibians are different from birds in many ways that are important for our discussion. First, because they are ectotherms – meaning that their body temperatures adjust to the ambient temperature and radiation of the environment – they must avoid environments in which temperatures are too cold or too hot. Second, amphibians must live near water not only because the reproductive part of their life cycle is dependent on water but also because they must keep their skin moist because they 'breathe' through their skin as well as their lungs. Third, neither reptiles nor amphibians are able to disperse as easily as birds because they must crawl rather than fly,

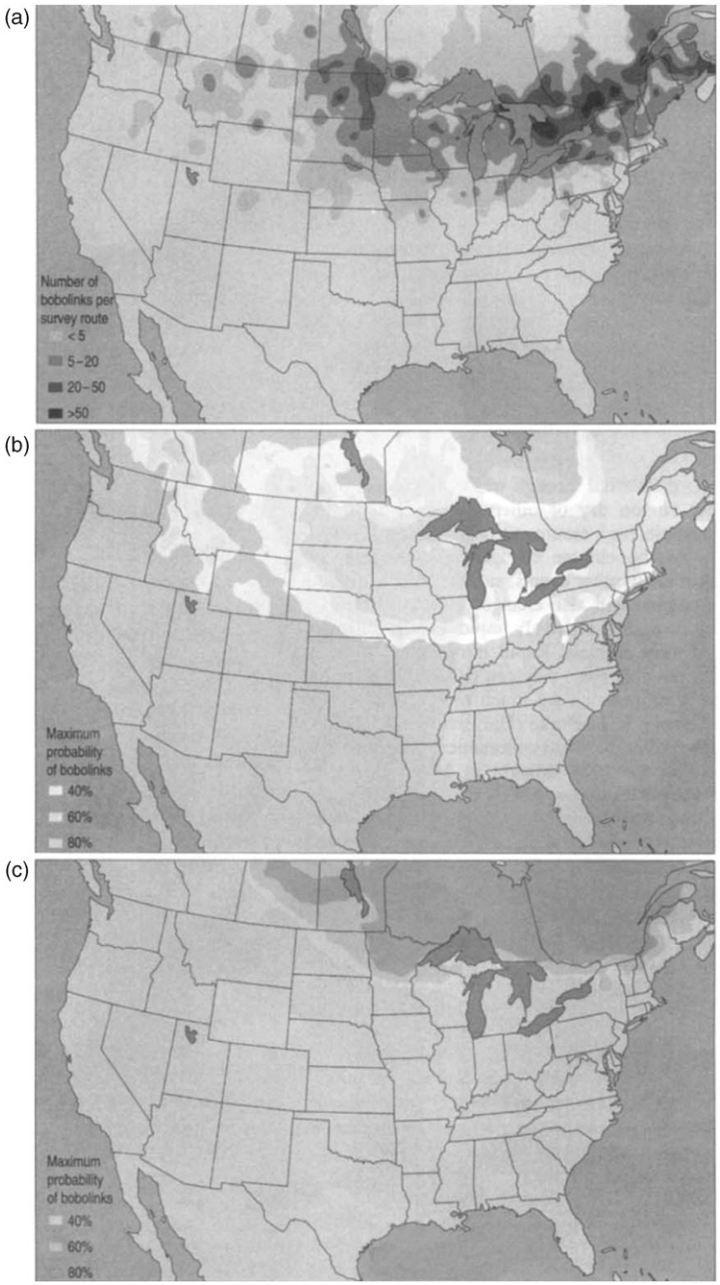


Figure 2 (a) Map of current range and abundance of the bobolink as determined from actual observations during the US Geological Survey Breeding Bird Survey and (b) map of current range and abundance of the bobolink as estimated from the empirical–statistical model. The high correspondence in patterns between the maps in (a) and (b) suggests that this model reliably captures many of the features of the actual observed range and abundance of this species as depicted in the map in (a). (c) Map of the forecasted range and abundance of the bobolink for climate change response of a model with doubled CO_2. This map illustrates the potential for very significant shifts that doubled CO_2 would cause.

and the habitat through which they crawl must not be too dry or otherwise impassable (e.g., high mountains or superhighways). Work by Araujo (2004) looking at patterns of richness of these species in Europe shows that they essentially do not disperse.

As the climate changes, the character of extreme weather events, such as cold snaps and droughts, will also change (Karl *et al.*, 1995), necessitating relatively rapid habitat changes for most animals. There is increasing evidence of experienced and projected effects on animals of changes in climate variability. For example, population reductions in flying foxes and other species (Welbergen *et al.*, 2008; Thibault and Brown, 2008) and displacement (Kelly and Goulden, 2008) have been detected. Species will need to move into areas that are cooler/wetter (or both), much as they would in a warming climate. This means that barriers, even roads, which often cause high mortality, can be obstacles. For example, the role of extreme events and climate variability is still not adequately taken into account in models of species ranges. Recent work based on three different methods has suggested that species in the

tropics are more in danger of extinction because they already live close to their thermal tolerance. Wright *et al.* (2009) compare distances to cool refuges, finding that 20% of tropical mammal species would have to travel more than 1000 km to a cool refuge under a moderate climate change scenario compared with only 4% of small-ranged extratropical mammal species. Deutsch *et al.* (2009) show how, due to physiological constraints, tropical insect abundance may decline rapidly for quite small climate changes, whereas in temperate regions, insect abundance might be expected to increase with warming. Recent projections for impacts in marine systems show dramatic turnovers of 60% of fish assemblages under climate change, which could disrupt marine ecosystem functioning (Cheung *et al.*, 2009).

Three-Way Linkages and Community Ecology

Anticipated changes in plant ranges will probably have dramatic effects on animals, both on the large biogeographic scale and on the local regional scale. The ranges of many animals are strongly linked to vegetation. For example, red-cockaded woodpeckers are endemic to mature longleaf pine and pine-oak forests (Mengel and Jackson, 1977), and the winter range of Sprague's pipit is coincident with bluestem, a grass (Root, 1988c). Consequently, the ranges of various animals that rely on specific vegetation will change as the ranges of these plants shift, assuming that some other factor is not limiting them. If the climate changes more rapidly than the dispersal rates of the plants, it will very likely result in extensive plant die-offs in the south or downslope before individuals can disperse and become established in the north or upslope. Thus, the ranges of animals relying on these plants could become compressed, and in some instances both the plants and the animals could become extinct. For instance, the red-cockaded woodpecker needs mature, living trees for nesting sites (Jackson, 1974), and if increasing temperature causes most large trees to die before the newly established dispersing trees mature enough, then this woodpecker, federally listed as endangered, could easily become extinct.

Many animal species have ranges that are not directly limited by vegetation but are instead restricted by temperature. This is true for most ectotherms (insects and related arthropods, amphibians, and reptiles) as well as some endotherms (mammals and birds). For example, the Eastern Phoebe, a North American songbird, winters in the US in areas with average minimum temperatures warmer than $-4\,^{\circ}C$ (Root, 1988a, b). As the Earth warms, those species directly limited by temperature will be able to expand northward as rapidly as their dispersal mechanisms will allow, again assuming other factors are not limiting them. The animals limited by vegetation will be able to expand their ranges only as rapidly as the vegetation changes. Consequently, the potential for significant disruption among communities is high. For instance, some animals may no longer be able to coexist because an invading species disrupts the balance between competing species or between predator and prey. Therefore, to understand the ecological consequences of global climatic change on animals, the three-way linkages among animals, plants, and climate must be understood. It is critical to realize that this is not simply a one-way process whereby climate influences biota but rather a three-way process because animals and plants affect each other and are affected by climate. At the same time, altered surface vegetation can affect climate because mid-continental summer precipitation is significantly influenced by water vapor from evapotranspiration (Ye, 1989; Salati and Nobre, 1991). There is increasing evidence of climate-change-induced alterations in ecosystem composition and function (e.g., nutrient cycling, productivity, and community dynamics; Bertin, 2008; Olofsson *et al.*, 2009). In particular shifts have already occurred in several marine food webs (Byrnes *et al.*, 2007; Alheit, 2009) as a result of (observed) changes in sea-surface temperature and (separately) of natural climate variability. This shows how likely future climate change will analogously affect species composition and hence ecosystem functioning and potentially biogeochemical cycles. Recent studies simulate the impacts of combinations of climate changes and their impacts on ecosystems and generally show larger impacts than previously anticipated. For example, an increased die-off of tree species is expected in response to the combination of increased temperatures and increased drought frequency (Adams *et al.*, 2009), and climate-induced shifts in key aquatic species can reduce the resilience of the process that maintains lakes in a clear state (Mooij *et al.*, 2009).

Predicting Extinction

As discussed above, impacts from climate change have already been observed on animal and plant species' distributions, abundances, and phenologies over the last $\sim$30 years (Parmesan and Yohe, 2003; Root *et al.*, 2003). A large number of independent studies now project that many species ranges will be reduced so much by climate change that they will be committed to extinction. For example, there are estimates that many small mammals living near isolated mountaintops (which are essentially habitat islands) in the Great Basin would become extinct given typical global change scenarios (MacDonald and Brown, 1992). Studies of small mammals in Yellowstone National Park show that statistically significant changes in both abundances and physical sizes of some species occurred with historical climate variations (which were much smaller than most projected climate changes for the next century), but there appear to have been no simultaneous genetic changes (Hadley *et al.*, 1997). Therefore, it is likely that climate change in the twenty-first century could cause substantial alteration to biotic communities, even in protected habitats such as Yellowstone National Park.

Thomas *et al.* (2004) used projections of 1103 species' distributions for future climate scenarios to assess extinction risks for sample regions that cover some 20% of the Earth's terrestrial surface. These scientists explore three approaches in which the estimated probability of extinction shows a power-law relationship with geographical range size. They predict, on the basis of midrange climate-warming scenarios for 2050, that 15–37% of species in their sample of regions and taxa will be "committed to extinction." When they averaged the three methods and two dispersal scenarios, minimal climate-warming scenarios produced lower projections of species committed to extinction ($\sim$18%) than midrange ($\sim$24%) and maximum change ($\sim$35%) scenarios. However, even minimal climate-warming scenarios, which produced lower percentages of species expected to be committed to extinction, show the importance of reducing GHGs. Despite many

unknowns that exist in projecting extinctions, including a lack of information on the time lag between climate change and species-level extinctions, the effects of including land use in the analyses of extinction risk and the uncertainty about which species will inhabit parts of the world that are projected to have climates for which no current analog exists, Thomas *et al.* report that the consistent overall conclusions across analyses establish that anthropogenic climate warming at least ranks alongside other recognized threats to global biodiversity and is likely to be the greatest threat in many if not most regions. However, they caution that most severe impacts of climate change are likely to result from the interactions of various threats rather than from climate acting in isolation, synergisms that their calculations did not take into account.

Using different methods Malcolm *et al.* (2006) provides an overall estimate of extinctions of endemics in biodiversity hotspots, and generally supports the findings of Thomas *et al.* (2004). Fischlin *et al.* (2007) summarize extinction risks under various levels of global mean temperature rise, estimating that 20–30% of species studied would be at risk of extinction if global temperature rise exceeds 2–3 °C. A meta-analysis of some 80 studies linking quantitative changes in climate change with projected quantified extinction risks for species found increased impacts on ecosystems with increased annual global mean temperature rise (Warren *et al.*, 2011). Substantial increases in extinction risks are commonly projected as the global average temperature reaches and exceeds 2 °C, especially in biodiversity hotspots.

Sekercioglu *et al.* (2008) also discuss the vulnerability of highland species to climate change, in this case, highland bird species. To approximate the effects of elevational distribution on avian extinction risk, they define an elevational area index that is the number of square kilometers of cone surface area a bird species, with its current elevational range, would occupy on an idealized mountain. This is analogous to actual land-surface area available to a bird and makes it possible to compare, for all land bird species, the limitations of elevation on area. Based on this index, which is not sensitive to the model-mountain's shape or size, Sekercioglu *et al.* show that differences in elevational distribution explain 94% of the variation in the proportions of bird species facing a risk of extinction, based on the International Union of Conservation of Nature and Natural Resources (IUCN) definitions of near threatened and threatened. This means that by knowing the elevational distribution of a bird species, one can correctly predict its likelihood of risking extinction 94% of the time, which is not only ecologically interesting, but also has significant conservation implications for birds and other terrestrial taxa. Furthermore, 92% of the variation is explained after various sensitivity analyses regarding the choices of parameters and species. The fact that this relationship is insensitive to large variations in uncertain criteria supports their confidence in the robustness of the results. Equally importantly, by utilizing a modeling approach that combines this index with a wide range of temperature lapse rates, elevational-limitation scenarios, and IPCC surface-warming projections for 2100, Sekercioglu *et al.* estimate that intermediate climate change effects may result in ~600 land bird extinctions and threaten an additional 330 species with extinction. The worst case scenario projects nearly 3200 extinctions – 38% of all land birds. As well as testing the robustness of the relationship between elevational distribution and extinction risk, they also bracketed the uncertainty for climate change scenarios.

Although the climate change extinction-rate estimates are often criticized, it needs to be borne in mind that these are based on MEAN climates. Even the variables taken as 'extreme' are normally not true extremes, but are averages over a month. Mortality in species is often measured in x number of days exceeding a certain threshold (the same as for coral bleaching). Thus, because the return rates of these threshold events will increase as the climate warms, then extinctions, or at least local extirpations, may very well happen sooner than projected by changes in the mean.

As mentioned above, the ability of species to reach new climatically suitable areas will be hampered by habitat loss and fragmentation, and their ability to persist in appropriate climates is likely to be affected by new invasive species. Indeed a recent study by Barnosky *et al.* (2011) indicates that we may be at the beginning of a mass extinction event. Five such extinction events, during which ~75% of species go extinct per event, have occurred on our planet, but this would be the first one caused by only one species.

Recommendations

Improve Regional Analysis, Study Transients, and Include Many Variables

The most reliable projections from climatic models are for global-scale temperature changes. Ecological impact assessments, however, need time-evolving (transient) scenarios of regional to local-scale climate changes: included are changes in precipitation; severe storm intensity, frequency, and duration; drought frequency, intensity, and duration; soil moisture; frost-free days; intense heat waves; ocean currents; upwelling zones; near-ground ozone; forest canopy humidity; and ultraviolet radiation and total solar radiation reaching the surface, where photosynthesis is important. Data gathered at many scales and by coordinated volunteer and professional sources are needed for archives of these regional and local variables, which in turn can be used to develop and test models or other techniques for climatic forecasting.

Adaptability

Our current inability to credibly predict time-evolving regional climatic changes has many implications, one of which concerns the adaptability of agricultural ecosystems. That is, any experience farmers might have with anomalous weather in, for example, the 2020s may not help them adapt to the evolving climate change in the 2030s because a transient climate change could differ dramatically over time. This would inhibit learning by doing, creating a potential lack of adaptability associated with the difficulty of reliably predicting regional climatic consequences (Schneider *et al.*, 2000). Such rapid climate changes would be especially difficult for natural ecosystems to adapt to because habitats do not have the luxury of 'choosing' to plant new seeds or change irrigation systems, soil tillage practices, or other agricultural practices.

There is a growing debate in the natural resources management field over the degree to which natural ecosystems are able to adapt to current and future changes, given that they have been exposed to past climatic changes. Paleoecological studies, and studies on current range and phenological shifts, indicate that the natural response of many species has been, and is, to move or shift to track suitable climatic conditions. Some have argued that the amount of change that some communities have survived in the past might leave them with the ability to survive equivalent changes in the future. This is very difficult to demonstrate. An approach would be to identify, using paleoclimatology, the amount of change ecological communities might have been exposed to in the past and infer that they would be able to survive an equivalent amount of change in the future. Such analyses present many difficulties that include the following: (1) having a paucity of needed large-scale, very long-term appropriately continuous datasets for most species, (2) making risky assumptions on the rate at which species are able to autonomously adapt (behaviorally, physiologically, or evolutionarily) relative to the rate of change not only in the climate but also in other members of the ecological and human-modified landscapes and communities, and (3) overlooking that such changes in the past occurred in much more intact landscapes – lacking barriers of roads, cities, and agricultural fields. Given that the current velocity of climate change is generally thought to exceed already the rate of autonomous adaptation (including movement rates) of most species, then the precautionary principle would dictate that an assumption of autonomous adaptability not be made. The precautionary principle would dictate that efforts to conserve biodiversity in the face of climate change, except perhaps in species with very high birth rates, rely on changing natural resource management practices and for integrating natural resource protection with livelihood protection in the face of a rapidly warming world.

Ecological Applications-Driven Climatic Research

Regional projections of climatic change arising from a variety of greenhouse gas and sulfur oxide emissions scenarios are essential for ecological applications. Such studies must stress the climatic variables most likely to have significant effects on biological resources. For example, extreme variability measures such as high temperature and low relative humidity are important for evaluating the risk of forest fires (Torn and Fried, 1992). Identifying such variables of ecological importance and communicating this information to climate scientists require close interdisciplinary, multi-institutional, and cross-scale research efforts to ensure that combinations of variables relevant to ecological applications receive research priority by climatologists. A focus of climate research toward changing climatic variability (Mearns *et al.*, 1984, 1990; Rind *et al.*, 1989) might be more useful for ecological impact assessments than the current focus among climatic modelers on climatic means.

Interactive, Multiscale, Ecological Studies Needed

Conservation biologists and others face a number of challenges regarding climate change over the next several decades, both in predicting and documenting its patterns and in dealing with its effects. Most ecological studies project the response of one species at small scales or shifts in biomes at large scales to an equilibrium, CO_2-doubled climate model (e.g., the Vegetation/Ecosystem Modeling and Analysis Project (VEMAP), 1995). What is needed for more realistic and useful ecological impact assessments is a multiscale, multispecies, and multitaxa analysis driven by regionally specific, transient climatic change forecasts. The construction of ecological forecast models first requires large-scale datasets gathered locally by professional (e.g., US Geological Survey landcover datasets) and volunteer (e.g., National Audubon Society Christmas Bird Count) workers. Without such datasets, virtually no credible progress is possible in determining large-scale patterns of associations among ecological and climatic variables. Small-scale studies informed by large-scale patterns are then needed to refine causal mechanisms underlying such large-scale associations, thereby testing the formulas used to make projections of various species or biome responses to hypothesized global changes. For example, Pacala and Hurtt (1993) suggested small- to medium-scale experiments to improve forest gap models. Their criticisms suggest that largely first principles, bottom–up models may still be unrealistic if some top–down parameters (i.e., growth-modifying functions in the instance of gap models) are not appropriately derived from data at the scale at which the models are being applied (Root and Schneider, 1995).

One obvious truism emerges: credible modeling required for forecasting across many scales and for complex interacting systems is a formidable task requiring repeated testing of many approaches. Nevertheless, tractable improvements in refining combined top–down and bottom–up techniques can be made. It will, however, take more than one cycle of interactions to reliably address the cross-scale and multicomponent problems of ecological assessment. The strategic cyclical scaling paradigm has two motivations: (1) better explanatory capabilities for multiscale, multicomponent interlinked social–natural systems (e.g., climate–ecosystem interactions or behavior of adaptive agents in economic models responding to the advent or prospect of climatic changes) and (2) more reliable impact assessments and problem-solving capabilities – predictive capacity – as requested by the policy community (Root and Schneider, 2006).

For the purposes of assessing progress toward the target to achieve by 2010 a significant reduction in the current rate of biodiversity loss, biodiversity loss is defined as the long-term or permanent qualitative or quantitative reduction in components of biodiversity and their potential to provide goods and services, to be measured at global, regional, and national levels (decision VII/30, paragraph 2 of Strategic Plan). The "current" rate is taken to be the rate in 2002, when the Strategic Plan was adopted.

See also: Carbon Cycle. Climate, Effects of. Diversity, Community/Regional Level. Energy Use, Human. Geologic Time, History of Biodiversity in. Greenhouse Effect

References

Adams HD, Guardiola-Claramonte M, Barron-Gafford GA, *et al.* (2009) Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. *Proceedings of the National Academy of Sciences USA* 106: 7063–7066.

Alheit J (2009) Consequences of regime shifts for marine food webs. *International Journal of Earth Science* 98: 261–268.

Araujo SP (2004) An evaluation of methods for modeling species distributions. *Journal of Biogeography* 31(10): 1555–1568.

Badgley C (1998) Vertebrate indicators of paleoclimate. *Journal of Vertebrate Paleontology* 18(supplement 3): 25A.

Bailey RM (1955) Differential mortality from high temperature in a mixed population of fishes in southern Michigan. *Ecology* 36: 526–528.

Bergmann C (1847) Ueber die Ve rhältnisse der Wärmeö konomie der Thiere zu ihrer Grösse. *Göttinger Studien* 1: 595–708.

Bertin RI (2008) Plant phenology and distribution in relation to recent climate change. *Journal of the Torrey Botanical Society* 135: 126–146.

Boag PT and Grant PR (1984) The classical case of character release: Darwins' finches (Geospiza) on Isla Daphne Major, Galapagos. *Biological Journal of the Linnean Society* 22(3): 243–287.

Botkin DB, Janak JR, and Wallis JR (1972) Some ecological consequences of a computer model of forest growth. *Journal of Ecology* 60: 849–872.

Bull JJ (1980) Sex determination in reptiles. *Quaternary Review of Biology* 55: 3–21.

Byrnes JE, Reynolds PL, and Stachowicz JJ (2007) Invasions and extinctions reshape coastal marine food webs. *Public Library of Science ONE* 2: e295.

Caldeira K and Wickett ME (2003) Anthropogenic carbon and ocean pH. *Nature* 425: 365.

Coffroth MA, Lasker HR, and Oliver JK (1990) Coral mortality outside of the eastern Pacific during 1982–1983: Relationship to El Niño. In: Glynn PW (ed.) *Global Ecological Consequences of the 1982–1983 El Niño-Southern Oscillation*. Amsterdam: Elsevier.

Cheung WL, Lam VWY, Sarmeinto JL, Watson R, and Pauly D (2009) Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries* 10: 235–251.

Cooperative Holocene Mapping Project (COHMP) (1988) Climatic changes of the last 18,000 years: Observations and model simulations. *Science* 241: 1043–1052.

Crowley T (1993) Use and misuse of the geologic "analogs" concept. In: Eddy JA and Oeschger H (eds.) *Global Changes in the Perspective of the Past*, pp. 17–27. Chichester, England: Wiley. Dahlem Workshop Report ES12.

Daily GC (1997) *Nature's Services: Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection*, 432 pp. London: John Murray.

Davis MB and Zabinski C (1992) Changes in geographical range resulting from greenhouse warming effects on biodiversity in forests. In: Peters RL and Lovejoy TE (eds.) *Global Warming and Biological Diversity*, pp. 297–308. New Haven, CT: Yale University Press.

Dettinger MD, Cayan DR, Meyer MK, and Jeton AE (2004) Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900–2099. *Climatic Change* 62: 283–317.

Deutsch CA, Tewsbury JJ, Huey RB, *et al.* (2009) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences USA* 105: 6668–6672.

Easterling DR, Evans JL, Groisman PY, Karl TR, Kunkel KE, and Ambenje P (2000) Observed variability and trends in extreme climate events: A brief review. *Bulletin of the American Meteorological Society* 81: 417–425.

Eddy JA and Oeschger H (eds.) (1993) *Global Changes in the Perspective of the Past*, 383 pp. New York: Wiley.

Ehrlich PR and Hanski I (2004) *On the Wings of Checkerspots: A Model System for Population Biology*. New York: Oxford University Press.

Ehrlich PR, Murphy DD, Singer MC, Sherwood CB, White RR, and Brown IL (1980) Extinction, reduction, stability and increase: The responses of checkerspot butterfly (*Euphydryas editha*) populations to the California drought. *Oecologia* 46: 101–105.

Ellsaesser HW (1990) A different view of the climatic effect of CO_2-updated. *Atmósfera* 3: 3–29.

Fischlin A, Midgely G, Price J, *et al.* (2007) Ecosystems, their properties, goods and services. In: Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, and Hanson CE (eds.) *Climate Change 2007: Impacts, Adaptation and Vulnerability*, pp. 211–272. Cambridge, UK: Cambridge University Press. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC).

García-Barros E, Munguira ML, Martín Cano J, Benito Romo H, Garcia-Pereira P, and Maravalhas ES (2004) *Atlas of the Butterflies of the Iberian Peninsula and Balearic Islands (Lepidoptera: Papilionoidea & Hesperioidea)*. Zaragoza, Spain: Sociedad Entomológica Aragonesa.

Gómez de Aizprua C (1987) *Atlas Provisional Lepidopteros de Madrid*. Madrid, Spain: Comunidad de Madrid.

Grabherr G, Gottfried M, and Pauli H (1994) Climate effects on mountain plants. *Nature* 369: 448.

Graham RW and Grimm EC (1990) Effects of global climate change on the patterns of terrestrial biological communities. *Trends in Ecology and Evolution* 5: 289–292.

Groombridge B (ed.) (1992) *Global Biodiversity: Status of the Earth's Living Resources*, 585 pp. London: Chapman & Hall.

Hadly EA (1997) Evolutionary response of pocket gopher (*Thomomys talpoides*) to Late-Holocene climatic change. *Biological Journal of Linnean Society* 60: 277–296.

Harte J and Shaw R (1995) Shifting dominance within a montane vegetation community: Results of a climate-warming experiment. *Science* 267: 876–880.

Hawkins BA and Holyoak M (1998) Transcontinental crashes of insect populations? *American Naturalist* 152: 480–484.

Hayhoe K, Cayan D, Field CB, *et al.* (2004) Emissions pathways, climate change, and impacts on California. *Proceedings of the National Academy of Sciences USA* 101(34): 12422–12427.

Hill JK, Thomas CD, Fox R, *et al.* (2002) Responses of butterflies to twentieth century climate warming: Implications for future ranges. *Proceedings of the Royal Society of London B* 269: 2163–2171.

Holdridge LR (1967) *Life Zone Ecology*. San Jose, Costa Rica: Tropical Science Center.

Huntsman AG (1946) Heat stroke in Canadian maritime stream fishes. *Journal of the Fisheries Research Board of Canada* 6: 476–482.

Idso SB and Brazel AJ (1984) Rising atmospheric carbon dioxide concentrations may increase streamflow. *Nature* 312: 51–53.

Idso SB and Kimball BA (1993) Tree growth in carbon dioxide enriched air and its implications for global carbon cycling and maximum levels of atmospheric CO_2. *Global Biogeochemical Cycle* 7: 537–555.

Intergovernmental Panel on Climate Change (IPCC) (1996a) In: Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, and Maskell K (eds.) *Climate Change 1995 – The Science of Climate Change – Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, 572 pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (1996b) In: Watson RT, Zinyowera MC, and Moss RH (eds.) *Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change – Contribution of Working Group II: Scientific Technical Analysis to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, 878 pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (1996c) In: Bruce J, Lee H, and Haites E (eds.) *Climate Change 1995 – Economic and Social Dimensions of Climate Change – Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, 448 pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (2001a) In: Houghton JT, Ding Y, Griggs DJ, *et al.* (eds.) *The Scientific Basis, Summary for Policy Makers – Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 881pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (2001b) In: McCarthy JJ, Canziani OF, Leary NA, Dokken DJ, and White KS (eds.) *Impacts, Adaptation, and Vulnerability – Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 1032 pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (2001c) In: Metz B, Davidson O, Swart R, and Pan J (eds.) *Mitigation – Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 752 pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (2001d) In: Watson RT and the Core Writing Team (eds.) *Synthesis Report 2001 – Contribution of Working Groupe I, II, and III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 397 pp. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) (2007a) In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis*. New York, NY, USA: Cambridge University Press. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Intergovernmental Panel on Climate Change (IPCC) (2007b) Climate change 2007: Impacts adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Intergovernmental Panel on Climate Change (IPCC) (2007c) Climate change 2007. Mitigation of climate change. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Jackson JA (1974) Gray rat snakes versus red-cockaded woodpeckers: Predator prey adaptation. *Auk* 91: 342–347.

Jacobson MZ (2005) Studying ocean acidification with conservative, stable numerical schemes for nonequilibrium air–ocean exchange and ocean equilibrium chemistry. *Journal of the Geophysical Research Atmosphere* 110: D07302. doi: 10.1029/2004JD005220.

Jarvis PG (1993) Prospects for bottom-up models. In: Ehleringer JR and Field CB (eds.) *Scaling Physiological Processes: Leaf to Globe*, pp. 117–126. New York: Academic Press.

Karl TR, Knight RW, Easterling DR, and Quayle RG (1995) Trends in U.S. climate during the twentieth century. *Consequences* 1: 3–12.

Kearney M, Shine R, and Porter WP (2009) Potential for behavioural thermoregulation to buffer 'cold-blooded' animals against climate warming. *Proceedings of the National Academy of Sciences USA* 106: 3835–3840.

Kelly AE and Goulden ML (2008) Rapid shifts in plant distribution with recent climate change. *Proceedings of the National Academy of Sciences USA* 105: 11823–11826.

Kennedy AD (1995) Antarctic terrestrial ecosystem responses to global environmental change. *Annual Review of Ecological System* 26: 683–704.

King JC and Harangozo SA (1998) Climate change in the western Antarctic peninsula since 1945: Observations and possible causes. *Annual Glaciology* 27: 571–575.

Klanderud K and Birks HJB (2003) Recent increases in species richness and shifts in altitudinal distributions of Norwegian mountain plants. *Holocene* 13: 1–6.

Konvicka M, Maradova M, Benes J, Fric Z, and Kepka P (2003) Uphill shifts in distribution of butterflies in the Czech Republic: Effects of changing climate detected on a regional scale. *Global Ecology and Biogeography* 12: 403–410.

Kudrna O (2002) The distribution atlas of European butterflies. *Oedippus* 20: 1–342.

Leemans R and Eickhout B (2004) Another reason for concern: Regional and global impacts on ecosystems for different levels of climate change. *Global Environmental Change* 14: 219–228.

Levin SA (1993) Concepts of scale at the local level. In: Ehleringer JR and Field CB (eds.) *Scaling Physiological Processes Leaf to Globe*, pp. 7–19. New York: Academic Press.

MacDonald KA and Brown JH (1992) Using montane mammals to model extinctions due to global change. *Conservation Biology* 6: 409–425.

Malcolm JR, Liu CR, Neilson RP, Hansen L, and Hannah L (2006) Global warming and extinctions of endemic species from biodiversity hotspots. *Conservation Biology* 20: 538–548.

Mattson WJ and Haack RA (1987) The role of drought in outbreaks of plant-eating insects. *BioScience* 37: 110–118.

Mearns LO, Katz RW, and Schneider SH (1984) Changes in the probabilities of extreme high temperature events with changes in global mean temperature. *Journal of Climate and Applied Meteorology* 23: 1601–1612.

Mearns LO, Schneider SH, Thompson SL, and McDaniel L (1990) Analysis of climate variability in general circulation models: Compared with observations and changes in variability in $2xCO_2$ experiments. *Geophysical Research Letters* 95: 20469–20490.

Meehl GA, Zwiers F, Evans J, Knutson T, Mearns L, and Whetton P (2000) Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bulletin of the American Meteorological Society* 81: 427–436.

Melillo JM, McGuire AD, Kicklighter DW, Moore III B, Vörösmarty CJ, and Schloss AL (1993) Global climate change and terrestrial net primary production. *Nature* 363: 234–240.

Mengel RM and Jackson JA (1977) Geographic variation of the red-cockaded woodpecker. *Condor* 79: 349–355.

Midgley GF, Hannah L, Rutherford MC, and Powrie LW (2002) Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot. *Global Ecology and Biogeography* 11: 445–451.

Mooij WM, De Senerpont Domis LN, and Janse JN (2009) *Ecological Modelling* 220(21): 3011–3020.

Myers N (1996) Two key challenges for biodiversity: Discontinuities and synergisms. *Biodiversity and Conservation* 5: 1024–1034.

Neilson RP (1993) Transient ecotone response to climatic change: Some conceptual and modelling approaches. *Ecological Applications* 3: 385–395.

Ottersen G, Kim S, Huse G, Polovina JJ, and Stenseth NC (2010) Major pathways by which climate may force marine fish populations. *Journal of Marine Systems* 79(3–4): 343–360.

Overpeck JT, Webb RS, and Webb III T (1992) Mapping eastern North American vegetation change over the past 18,000 years: No analogs and the future. *Geology* 20: 1071–1074.

Pacala SW and Hurtt GC (1993) Terrestrial vegetation and climate change integrating models and experiments. In: Kareiva P, Kingsolver J, and Huey R (eds.) *Biotic Interactions and Global Change*, pp. 57–74. Sunderland, MA: Sinauer Associates.

Parmesan C (1996) Climate and species' range. *Nature* 382: 765–766.

Parmesan C, and Galbraith H (2004) Observed impacts of global climate change in the US. *Pew Center for Global Climate Change Report*.

Parmesan C, Root TL, and Willig MR (2000) Impacts of extreme weather and climate on terrestrial biota. *Bulletin of the American Meteorological Society* 40: 443–450.

Parmesan C, Ryrholm N, Stefanescu C, *et al.* (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* 399: 579–583.

Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Pastor J and Post WM (1988) Response of northern forests to CO_2-induced climate change. *Nature* 334: 55–58.

Pickett STA and White PS (eds.) (1985) *The Ecology of Natural Disturbance and Patch Dynamics*. New York: Academic Press.

Poole JH (1989) Announcing intent: The aggressive state of musth in African elephants. *Animal Behaviour* 37: 140–152.

Pounds JA, Bustamante MR, Coloma LA, *et al.* (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439: 161–167.

Pounds JA, Fogden MPL, and Campbell JH (1999) Biological response to climate change on a tropical mountain. *Nature* 398: 611–615.

Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA, and Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19: 117–134.

Repasky RR (1991) Temperature and the northern distributions of wintering birds. *Ecology* 72: 2274–2285.

Rind D, Goldberg R, and Ruedy R (1989) Change in climate variability in the 21st century. *Climatic Change* 14: 5–37.

Root TL (1988a) Environmental factors associated with avian distributional limits. *Journal of Biogeography* 15: 489–505.

Root TL (1988b) Energy constraints on avian distributions and abundances. *Ecology* 69: 330–339.

Root TL (1989) Energy constraints on avian distributions: A reply to Castro. *Ecology* 70: 1183–1185.

Root TL (1994) Scientific/philosophical challenges of global change research: A case study of climatic changes on birds. *Proceedings of the American Philosophical Society* 138: 377–384.

Root TL (1999) In: Ernst G (ed.) *Earth Systems: Processes and Issues*, chapter 21. New York: Cambridge University Press.

Root TL, MacMynowski DP, Mastrandrea MD, and Schneider SH (2005) Human-modified temperatures induce species changes: Joint attribution. *Proceedings of the National Academy of Sciences USA* 102: 7465–7469.

Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, and Pounds JA (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421: 57–60.

Root TL and Schneider SH (1995) Ecology and climate research strategies and implications. *Science* 269: 334–341.

Root TL and Schneider SH (eds.) (2002) *Wildlife Responses to Climate Change: North American Case Studies*. Washington, DC: Island Press.

Root TL and Schneider SH (2003) Strategic cyclical scaling: Bridging five orders of magnitude scale gaps in climatic and ecological studies. In: Rotmans J and Rothman DS (eds.) *Scaling Issues in Integrated Assessment*, pp. 179–204. Lisse, The Netherlands: Swets and Zeitlinger Publishers.

Root TL and Schneider SH (2006) Conservation and climate change. *Conservation Biology* 20(3): 706–708.
Rubenstein DI (1992) The greenhouse effect and changes in animal behavior: Effects on social structure and life-history strategies. In: Peters RL and Lovejoy TE (eds.) *Global Warming and Biological Diversity*, pp. 180–192. New Haven: Yale University Press.
Salati E and Nobre CA (1991) Possible climatic impacts of tropical deforestation. *Climatic Change* 19: 177–196.
Schneider SH (1984) On the empirical verification of model-predicted CO_2-induced climatic effects. In: Hansen J and Takahashi T (eds.) *Climate Processes and Climate Sensitivity*, pp. 187–201. Washington, DC: American Geophysical Union.
Schneider SH (1993a) Can paleoclimatic and paleoecological analyses validate future global climate and ecological change projections? In: Eddy JA and Oeschger H (eds.) *Global Changes in the Perspective of the Past*, pp. 317–340. Chichester, England: Wiley.
Schneider SH (1993b) Scenarios of global warming. In: Kareiva P, Kingsolver J, and Huey R (eds.) *Biotic Interactions and Global Change*, pp. 9–23. Sunderland, MA: Sinauer Associates.
Schneider SH (1997) *Laboratory Earth: The Planetary Gamble We Can't Afford to Lose*. New York: Basic Books.
Schneider SH, Easterling WE, and Mearns L (2000) Adaptation: Sensitivity to natural variability, agent assumptions and dynamic climate changes. *Climatic Change* 45: 203–221.
Schneider SH and Londer R (1984) *The Coevolution of Climate and Life*, 563 pp. San Francisco: Sierra Club Books.
Schneider SH and Root T (1998) *Climate Change*. Reston, VA: US Department of the Interior, US Geological Survey. Reprinted from: Mac MJ, Opler PA, Puckett Haeker CE, and Doran PD. *Status and Trends of the Nation's Biological Resources*. Washington, DC: US Geological Survey.
Schneider SH and Root TL (1996) Ecological implications of climate change will include surprises. *Biodiversity and Conservation* 5: 1109–1119.
Sekercioglu CH, Schneider SH, Fay JP, and Loarie SR (2008) Climate change, elevational range shifts, and bird extinctions. *Conservation Biology* 22: 140–150.
Singer MC and Ehrlich PR (1979) Population dynamics of the checkerspot butterfly *Euphydryas editha*. *Fortschritte der Zoologie* 25: 53–60.
Singer MC and Thomas CD (1996) Evolutionary responses of a butterfly metapopulation to human and climate-caused environmental variation. *American Naturalist* 148: S9–S39.
Smith FA and Betancourt JL (1998) Response of bushy-tailed woodrats (*Neotoma cinerea*) to late Quaternary climatic change in the Colorado Plateau. *Quaternary Research* 50: 1–11.
Smith RIL (2001) Plant colonization response to climate change in the Antarctic. *Folia Faculty of Science of National University Masarykianae Brunensis, Geographia* 25: 19–33.
Smith TM, Shugart HH, Bonan GB, and Smith JB (1992) Modeling the potential response of vegetation to global climate change. In: Woodward FI (ed.) *Advances in Ecological Research: The Ecological Consequences of Global Climate Change*, pp. 93–116. New York: Academic Press.
Stefanescu C, Herrando S, and Páramo F (2004) Butterfly species richness in the north-west Mediterranean Basin: The role of natural and human-induced factors. *Journal of Biogeography* 31: 905–915.
Sullivan RM and Best RL (1997) Effects of environment on phenotypic variation and sexual dimorphism in *Dipodomys simulans* (Rodentia: heteromyidae). *Journal of Mammalogy* 78: 798–810.
Thibault KM and Brown JH (2008) Impact of an extreme climatic event on community assembly. *Proceedings of the National Academy of Sciences USA* 105: 3410–3415.
Thomas C, Cameron A, Green RE, *et al*. (2004) Extinction risk from climate change. *Nature* 427: 145–148.
Thomas CD and Lennon JJ (1999) Birds extend their ranges northwards. *Nature* 399: 213.
Thomas CD, Singer MC, and Boughton DA (1996) Catastrophic extinction of population sources in a butterfly metapopulation. *American Naturalist* 148: 957–975.
Torn MS and Fried SJ (1992) Predicting the impacts of global warming on wildland fire. *Climatic Change* 21: 257–274.
UNEP (2005) *Convention on Biological Diversity: Second Global Biodiversity Outlook*. Montreal, Canada: Draft Executive Summary.
Vegetation/Ecosystem Modeling and Analysis Project, (VEMAP) (1995) Comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO_2 doubling. *Global Biogeochemical Cycles* 9: 407–437.
Vitousek PM (1993) Global dynamics and ecosystem processes: Scaling up or scaling down? In: Ehleringer JR and Field CB (eds.) *Scaling Physiological Processes: Leaf to Globe*, pp. 169–177. New York: Academic Press.
Walker LR (ed.) (1999) *Ecology of Disturbed Ground*, 868 pp. Amsterdam: Elsevier Science.
Walker LR and Willig MR (1999) An Introduction to Terrestrial Disturbance. *Ecology of Disturbed Ground*. Amsterdam: Elsevier Science. pp. 1–16.
Walther G-R, Post E, Menzel A, *et al*. (2002) Ecological responses to recent climate change. *Nature* 416: 389–395.
Warren R, Price J, de la Nava Santos S, Fischlin A, and Midgley G (2011) Increasing impacts of climate change upon ecosystems with increasing global mean temperature rise. *Climatic Change* 106: 141–177.
Welbergen JA, Klose SM, Markus N, and Eby P (2008) Climate change and the effects of temperature extremes on Australian flying-foxes. *Proceedings of the Royal Society B: Biological Sciences* 275: 419–425.
Williams SE, Bolitho EE, and Fox S (2003) Climate change in Australian tropical rainforests: An impending environmental catastrophe. *Proceedings of the Royal Society of London B* 270: 1887–1892.
Willig MR and McGinley MA (1999) Animal responses to natural disturbance and roles as patch generating phenomena. *Ecology of Disturbed Ground*, pp. 667–689. Amsterdam: Elsevier Science.
Willig MR, Secrest MF, Cox SB, *et al*. (1998) Long-term monitoring of snails in the Liquillo Experimental Forest of Puerto Rico: Heterogeneity, scale, disturbance, and recovery. In: Dallmeier F and Comisky J, (eds.) *Forest Biodiversity in North, Central, and South America and the Caribbean: Research and Monitoring* vol. 21. pp. 293–322. Carnforth, Lancashire, UK: UNESCO and Parthenon Press. Man and the Biosphere Series.
Willig MR and Walker LR (1999) Disturbance in terrestrial ecosystems: Salient themes, synthesis, and future directions. In: Walker LR (ed.) *Ecology of Disturbed Ground*, pp. 747–767. Amsterdam: Elsevier Science.
Wilson RJ, Gutiérrez D, Gutiérrez J, Martınez D, Agudo R, and Monserrat VJ (2005) Changes to the elevational limits and extent of species ranges associated with climate change. *Ecology Letters* 8: 1138–1146.
Wright HE, Kutzbach JE, Webb III T, Ruddiman WE, Street-Perrott FA, and Bartlein PJ (1993) *Global Climates Since the Last Glacial Maximum*, 569 pp. Minneapolis: University of Minnesota Press.
Wright SJ, Muller-Landau HC, and Schipper J (2009) The future of tropical species on a warmer planet. *Conservation Biology* 23: 1418–1426.
Ye D (1989) Sensitivity of climate model to hydrology. In: Berger A, Dickinson RE, and Kidson JW (eds.) *Understanding Climate Change*, pp. 101–108. Washington, DC: American Geophysical Union.
Zimmerman JK, Willig MR, Walker LR, and Silver WL (1996) Introduction: Disturbance and Caribbean ecosystems. *Biotropica* 28: 756–765.

Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species

Joshua J Lawler, Carrie A Schloss, and Ailene K Ettinger, University of Washington, Seattle, WA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptation Actions that ameliorate the impacts of climate change on systems or species or that allow for systems and species to take advantage of climatic changes.

Adaptive management An iterative process of management and monitoring that treats management actions as experiments, the results of which inform changes in management actions.

Assisted colonization The translocation of species outside of their native range to allow them to track changes in climate.

Co-benefits Beneficial outcomes for human or natural systems of adaptation or strategies designed to address one or the other system.

Mitigation Actions designed to reduce the amount that the climate will change. These actions generally involve reducing greenhouse gas emissions and sequestering carbon dioxide (CO_2).

Phenology The timing of ecological events.

Resilience The ability of a system or species to return to its initial condition after perturbation by climate change.

Resistance The ability of a system or species to remain largely unchanged in the face of climate change.

Triage A system of prioritization for making decisions when resources are scarce and the need for responses is widespread.

Introduction

Climate change is poised to significantly alter ecological systems. Recent climatic changes have resulted in clear shifts in species distributions and the timing of ecological events (Parmesan, 2006). Although range shifts and changes in phenology are the two most well documented effects of recent climatic changes, there are a myriad of other ways in which changes in climate have affected and will likely affect terrestrial species. Among other things, climate change has the potential to alter population processes, interspecific interactions, and the impact of diseases and parasites. All of these effects have been documented to some degree.

The magnitude and rate of climatic changes projected for the coming century will provide challenges for many terrestrial species (IPCC, 2007a, b). Addressing these challenges will require some understanding of how species will likely respond to projected changes. There are several tools that can help scientists, managers, and planners anticipate the impacts of climate change on terrestrial species. These tools range from understanding historical patterns and conducting experiments to developing complex simulation models to forecast potential impacts.

An understanding of how species will likely respond to climate change is critical for developing management strategies and policies to maintain populations, protect species, and sustain ecosystem functions in a changing climate. Such approaches are generally referred to as adaptation strategies. The Intergovernmental Panel on Climate Change (2007a) has defined adaptation as "the adjustment in natural or human systems in response to actual or expected climatic stimuli or their effects, which moderates harm or exploits beneficial opportunities." Many adaptation strategies have been proposed for terrestrial ecological systems and specifically for terrestrial species. The majority of these strategies are broad recommendations or general concepts, but some are more specific actions.

The following sections provide both an overview of some of the impacts of climate change and how they can be anticipated and a review of the adaptation strategies that have been proposed for addressing these impacts.

Anticipating Impacts

There are many potential impacts of climate change on terrestrial species. The following sections concentrate on six types of impacts – impacts on phenology, range shifts, population processes, interspecific interactions, diseases and parasites, and interactions with other, nonclimatic stressors. Each section provides an overview of impacts and approaches for anticipating those impacts.

Phenology

One of the most widely observed ecological effects of climate change is a shift in phenology – the seasonal timing of life history events, such as flowering, egg hatching, migration, and senescence. For centuries, people have been observing phenology for agricultural and religious reasons, as well as simply to record changing seasons. For example, grape harvest dates have been tracked for the past 500 years across Europe (Menzel, 2005) and the appearance of spring cherry blossoms has been recorded since the fifteenth century in Japan (Menzel and Dose, 2005). In general, these and other historical data suggest a shift in the seasonal timing of biological events as temperatures have risen over the past century. Spring and summer events, including frog spawning, bird nesting, and

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00327-0

leaf unfolding, are occurring earlier, and the vegetative growing season has lengthened (Parmesan, 2006; Feehan *et al.*, 2009; Thackeray *et al.*, 2010). These changes have occurred rapidly and have been particularly strong at high latitudes, where warming has been the greatest (Parmesan, 2006; Feehan *et al.*, 2009). From 1976 to 2005, phenology of plants, invertebrates, and vertebrates in terrestrial, marine, and freshwater environments advanced, on average, by 11.7 days (0.38 days per year) (Thackeray *et al.*, 2010).

Although there has been an overall advancement in the timing of biological events, significant variation exists in patterns of phenological shifts across ecosystems, species, trophic levels, and functional groups. Of terrestrial organisms, plants have shown the most rapid rate of change (0.58 days per year from 1976 to 2005), and vertebrates have shown the slowest (0.25 days per year) (Thackeray *et al.*, 2010). These varying rates of change suggest that asynchronies may develop between interacting species, such as predators and their prey or plants and their pollinators. Indeed, asynchronies, or "phenological mismatches," have already been observed and, in some cases, have resulted in both reductions in individual fitness and declining population sizes (Parmesan, 2006; Forrest and Miller-Rushing, 2010; Miller-Rushing *et al.*, 2010).

By combining historical relationships between phenologies and climate with future climate projections, research efforts have begun to forecast how future changes in climate will affect phenology and the consequences of these potential phonological shifts (Memmott *et al.*, 2010; Ogden *et al.*, 2008a, b; Tobin *et al.*, 2008; Caffarra and Eccel, 2011). These forecasts project continued shifts toward earlier arrival of spring events, particularly at high altitudes (Caffarra and Eccel, 2011). Forecasts suggest that climate-induced phenological shifts can have major impacts on species interactions and communities, even when complete phenological mismatches do not occur (Fabina *et al.*, 2010).

Forecasting future phenological shifts is difficult, because the understanding is based largely on correlative observational studies. In many cases, there is no mechanistic understanding of controls over phenology and the degree to which it is controlled by climate. Genetics and nonclimatic cues such as photoperiod also affect phenology (Forrest and Miller-Rushing, 2010; Valtonen *et al.*, 2011). Furthermore, observational data have been biased toward plants, and the knowledge of trends in other organisms is limited. Even when extensive data and understanding are present, it may be problematic to assume that future trends will follow those in the past. For example, in the Tibetan Plateau, from 1982 to 2006, trends in spring vegetation phenology initially advanced concurrent with warming patterns, but started retreating in the mid-1990s in spite of continued warming (Yu *et al.*, 2010). The authors conclude that warm winter conditions caused a delay in spring phases due to chilling requirements. Species- and site-level variations in the magnitude and direction of phenological responses to changes in temperature highlight the need for further research on climate-induced shifts in phenology. In particular, mechanistic studies that determine the influence of climate on phenology and research on mammals, amphibians, fungi, and other understudied organisms are needed to improve the ability to anticipate future phenological changes.

Range Shifts

Climatic factors broadly determine species distributions, and therefore, climatic changes can cause associated changes in species distributions, or range shifts. Over the past century, both altitudinal and latitudinal range shifts in temperate- and tropical–terrestrial species have accompanied climatic changes (Parmesan and Yohe, 2003; Parmesan, 2006; Walther *et al.*, 2002; Root *et al.*, 2003). As the climate changes, species expand their ranges to occupy previously climatically unsuitable areas (Parmesan, 2006). Conversely, climatic changes may reduce the climatic suitability of species' current distributions, resulting in range contractions through local population extinctions (Parmesan *et al.*, 1999). The combination of range expansions into regions of newly suitable climate and range contractions away from regions of unsuitable climate can result in overall shifts in species ranges (Pearson and Dawson, 2003; Parmesan *et al.*, 1999).

Range shifts have been documented for plants, insects, birds, mammals, amphibians, and reptiles in tropical, temperate, and arctic regions. Overall, the magnitude and direction of these shifts have been consistent with climatic changes and tend to be poleward and/or upward in elevation, with an average shift of 6.1 km poleward or 6.1 m upward per decade (Parmesan and Yohe, 2003). Despite the general trend of upward and poleward range shifts, distributional responses to climate change are largely individualistic, with some species showing no change in their distributions despite warming and others showing shifts inconsistent with climatic changes (Parmesan and Yohe, 2003; Parmesan, 2006). The inconsistent shifts likely reflect the different degrees to which nonclimatic factors influence species distributions (Hellmann *et al.*, 2008b). For example, local-scale patterns, species interactions, habitat requirements, or proximity to source populations may be more important than climatic conditions in determining range shifts and establishment for some species (Elliott, 2011; Matthews *et al.*, 2010; Melles *et al.*, 2011). Many factors that interact with climate change to influence species distributions make projecting climate impacts on species ranges particularly difficult.

Nonetheless, many studies have used predictive models to project the potential effects of climate change on species distributions. Most of these studies have used species distribution models (often referred to as niche models or climate-envelope models in the context of climate change) that relate current climatic patterns to current species distributions and use this relationship along with projections of future climate from general circulation models to project future species distributions (Pearson and Dawson, 2003; Guisan and Zimmermann, 2000). These projections are often consistent with the direction of observed range shifts. Bioclimatic model projections can provide useful approximations of the magnitude and location of climate-induced changes to continental biodiversity. Projections have been made for a wide array of plants and animals in most regions of the world.

Despite widespread use of bioclimatic models, there are many uncertainties associated with the projections they produce. For example, projections from bioclimatic models are sensitive to the uncertainties associated with forecasts of future climate and the type of bioclimatic model employed.

Disparate projections can be summarized using ensembles of different types of species distribution models that incorporate projections from multiple climate models (Buisson *et al.*, 2010). In addition to the uncertainties, bioclimatic models have other limitations. For example, they generally do not incorporate the influence of nonclimatic factors (e.g., species interactions, dispersal potential, or habitat requirements), the potential for species to adapt to new climatic conditions rather than move to regions of suitable climate, and the possibility that species distributions are not at equilibrium with respect to current climatic conditions (Pearson and Dawson, 2003; Yates *et al.*, 2010; Dormann, 2007). Some bioclimatic models have addressed these limitations by incorporating dispersal potential, biotic interactions, or habitat data (Carroll, 2010; Vos *et al.*, 2008; Midgley *et al.*, 2006). Process-based models further incorporate detail by including biological processes parameterized by observations or measured species physiological tolerances and may provide more accurate forecasts of species distributions (Morin *et al.*, 2008; Morin and Thuiller, 2009; La Sorte and Jetz, 2010; Morin and Lechowicz, 2008).

Population Processes

Although projected range shifts provide a coarse view of how species will likely respond to climate change, they do not capture detailed changes that will occur at finer spatial scales. For example, climate change affects population processes such as birth and death rates, individual growth and reproductive capacity, life expectancy, immigration, and emigration. All of these processes affect short- and long-term changes in the size and age-structure of populations and determine whether populations increase, decrease, are able to establish in unoccupied regions, or go extinct. By observing population sizes and demographic parameters through time and by modeling population dynamics using observed vital rates, the authors have gained tremendous insight into climate-sensitive population processes.

Climate change affects population dynamics directly when changes in temperature, precipitation, or other climate factors alter vital rates in a population. Diverse population responses have been observed and forecasted, with some populations increasing due to climate change, others decreasing, and still others remaining essentially stable. Many insect populations, including mosquitoes and some beetles, are expected to increase in a warmer world, often because reproductive success (specifically, hatching and larva survival) is positively correlated with temperature (Morin and Comrie, 2010; Jonsson *et al.*, 2007; Estay *et al.*, 2009). Alternatively, populations of other organisms, such as tree species in western North America and polar bears in the arctic, are expected to decrease because of increased mortality rates associated with climate change (van Mantgem *et al.*, 2009; Molnar *et al.*, 2010). Because, for some species (e.g., many reptiles), the sexes of offspring are determined by temperature, changes in climate are expected to alter sex ratios, which may decrease long-term population viability (Mitchell *et al.*, 2010).

Climate change also indirectly affects population processes through interspecific interactions. Climate-sensitive dynamics in one species may lead to altered population processes of another species. For example, in alpine areas of Colorado, USA, the growing season has lengthened over the past 30 years, leading to declines in yellow-bellied marmot mortality, which triggered increases in their population sizes (Ozgul *et al.*, 2010).

Within communities and ecosystems, population-level responses to climate change vary greatly between and within species. For example, forecasted temperature increases are expected to cause rising soil temperatures in much of the world, potentially differentially affecting the longevity and dynamics of persistent soil seed banks of plants. In an arid region of Australia, some plant species showed significantly greater levels of germination after exposure to predicted increases in soil temperatures, whereas others experienced dramatic decreases in seed viability (Ooi *et al.*, 2009). Even within a species, population-level responses to climate change may differ. For example, American beaver populations are predicted to expand modestly at their northern range limits as a result of climate change, but population densities are likely to increase more dramatically in the interior portions of the beaver's range (Jarema *et al.*, 2009).

Even small alterations to vital rates due to climate change can have large consequences for population trajectories (McRae *et al.*, 2008). As computing power has increased, the capacity to model the effects of climate change on population dynamics has improved, and in recent years, studies have combined climate forecasts or bioclimate envelope models with population viability analyses and spatially explicit stochastic population models (Ooi *et al.*, 2009; Keith *et al.*, 2008; Mitchell *et al.*, 2010). Such forecasts are data and computationally intensive, but enable one to understand and better prepare for ecological impacts of climate change.

Interspecific Interactions

Climate-change-induced shifts in ranges, phenology, and population dynamics can lead to altered interspecific interactions and the formation of novel communities. Even strongly interacting species, such as specialist pollinators and the plants they pollinate, may have large differences in their physiological tolerances, life history strategies, and dispersal abilities. Differences in sensitivities and adaptive capacities can lead to decoupling of even the strongest relationships. Species' individualistic responses to climate change make forecasting altered species' interactions enormously challenging, and scientists often combine observational or experimental studies with modeling.

Climate change affects species interactions both within and across trophic levels. Effects on plant interactions are the best documented, but recent studies suggest that climate change alters competitive interactions among animals as well and causes shifts in other interspecific relationships across trophic levels. Increasing temperature and carbon dioxide may intensify pathogen infection rates, weaken mutualisms involving plants (e.g., pollination and seed dispersal), and enhance herbivory, particularly by insects (Parmesan, 2006; Tylianakis *et al.*, 2008; Traill *et al.*, 2010). Studies predict largely negative ramifications for important ecosystem services provided by species interactions, such as natural pest control by consumers

of insects, as well as the potential for increases in species coextinction rates (Traill *et al.*, 2010).

It is challenging to anticipate future effects of climate change on species interactions, but the ability to accomplish this is improving with the development of new forecasting tools and with an expanded knowledge base and data. Bioclimatic models can be fit separately for interacting species to evaluate the likelihood of altered interactions under future climates (e.g., interaction between a monophagous butterfly and its larval host plant) (Schweiger *et al.*, 2008). Bioclimatic models can also be nested, with the output for one species becoming a predictor variable for another species, to incorporate species interactions into projections of range shifts. Including these interactions is generally thought to improve explanatory and predictive power (Preston *et al.*, 2008; Sutherst *et al.*, 2007; Araújo and Luoto, 2007). Mechanistic models that explicitly incorporate species interactions can also be used to forecast range shifts (Ponti *et al.*, 2009), but these models are data intensive. Thus, observational and experimental studies are also needed to improve the understanding of species interactions and effects of climate and carbon dioxide concentrations on these interactions, particularly for mammals, amphibians, fungi, and other understudied organisms.

Diseases and Parasites

Diseases and parasites are specific instances of interspecific interactions that will likely be affected by climate change. The impacts of climate change on terrestrial species (i.e., phenological changes, range shifts, and changes in population processes) also affect parasites, diseases, disease vectors, the susceptibility of hosts, and the interactions between all of these organisms. Therefore, climate change will both directly and indirectly affect the emergence and spread of parasites and disease (Canto *et al.*, 2009). The impacts of climate change on parasites, diseases, vectors, and hosts are individualistic, and interactions between these impacts are complex (Moller, 2010; Luck *et al.*, 2011; Lafferty, 2009; Garrett *et al.*, 2011). However, the frequency of parasite and disease outbreaks will likely increase in a changing climate (Canto *et al.*, 2009; Brooks and Hoberg, 2007). These outbreaks have the potential to negatively impact plants and wildlife, agriculture, and human health (Reid and Gamble, 2009; Luck *et al.*, 2011; Patz *et al.*, 2007; Fussel, 2008; Garrett *et al.*, 2006).

Physiological tolerances to climatic conditions often determine disease and parasite distribution and abundance. Therefore, climate change will directly impact diseases with free-living life stages and diseases that require ectothermic vectors or hosts (Mas-Coma *et al.*, 2009; Patz *et al.*, 2008; Polley and Thompson, 2009). For example, the ability for parasites or disease vectors to overwinter requires a specific range of climatic conditions (Garrett *et al.*, 2006). Also, incubation time and the number of generations per year for some vectors and parasites are sensitive to temperature and humidity, and therefore, outbreaks of diseases and parasites will be impacted by climate change (Patz *et al.*, 2008; Jaramillo *et al.*, 2009). In general, higher precipitation and temperatures correspond with a higher disease transmission rates and higher diversity of diseases (Lafferty, 2009; Froeschke *et al.*, 2010). However, the responses of parasites and diseases to climatic changes are species specific, and so the resultant impact on hosts may be positive, negative, or neutral (Garrett *et al.*, 2006). For example, for a single host species, multiple parasites responded differently to changes in different climatic variables, resulting in no change to the fitness of the host species (Moller, 2010).

Because of the individualistic responses of parasites, diseases, vectors, and hosts to climate change and the complexity of the interactions of these responses, forecasting the impact of parasites and disease is difficult. Despite these complexities, projections are important to identify regions that are most susceptible to disease emergence or parasite outbreaks to facilitate proactive responses. Process-based models are often used to forecast the response of diseases to climatic changes by modeling climatic tolerances for survival, transmission, and reproduction (Rosenthal, 2009). For example, plague levels in black-tailed prairie dogs are forecasted to decrease due to inhibited transmission from higher temperatures (Snall *et al.*, 2009). Likewise, a simulation of host–parasite dynamics forecasts reduced transmission rates from stochastic events in regions of host expansion (Phillips *et al.*, 2010).

Phenological changes will also impact disease and parasite transmission and abundance. For example, increases in the length of flying seasons of disease vectors and parasites may increase disease transmission and the spread of the disease (Canto *et al.*, 2009). Conversely, phenological changes may also reduce the impact of parasites and disease by causing mismatches with hosts. Process-based models can also forecast phenological changes and the effects of those changes on population dynamics and pathogen–host dynamics (Ogden *et al.*, 2008a).

As climates change, new regions may become climatically suitable for a parasite, disease, or disease vector. Diseases and parasites may expand into these previously unsuitable, uninhabited regions. Species distribution models have been used to project range shifts for diseases, parasites, vectors, and hosts. For example, species distribution models for a tick, *Rhipicephalus appendiculatus,* and several host species forecasted overall range reductions for the tick and hosts, but an increase in tick–host assemblages in certain regions (Olwoch *et al.*, 2009). As for all species, nonclimatic factors such as dispersal limitations, land use, and interspecific interactions may limit climate-induced range expansions (Lafferty, 2009). However, disease and parasite distributions may be even more sensitive to nonclimatic distributional determinants because of their complex interactions with vectors and hosts. Therefore, forecasts from species distribution models may not be as effective as process-based models for anticipating the impacts of climate change on parasites and disease. For parasites and diseases, in particular, host availability may influence range expansion (Lafferty, 2009; Rosenthal, 2009). For example, if parasite or disease distributions are limited by host availability, distributional shifts of host species may correspondingly cause shifts in disease and parasite distributions. Also, climate change increases the potential for host switching, which may cause disease outbreaks in previously unaffected species that may be difficult to anticipate (Brooks and Hoberg, 2007). Pest and disease control may also have a large influence

on the distribution of a disease (Rosenthal, 2009). Diseases that affect humans in particular are sensitive to nonclimatic distributional determinants due to public health programs that are often influenced by socioeconomic distributions (Rosenthal, 2009). The extensive influence of nonclimatic factors on the distributions of diseases and parasites may overwhelm the impact of climate change, making impacts somewhat difficult to forecast.

Interactions with Other Stressors

Climate change will not affect species in a vacuum, but instead will interact with other factors that currently affect population dynamics or species distributions. These other factors include, but are not limited to, invasive species, land-use change, disturbance, and human responses to climate change.

Invasive Species

Invasive species are also experiencing changes in phenology and distributions, which may exacerbate the threats of climate change to native species. Climatic changes are likely to result in increases in invasive species' survival, abundance, and range expansions. Experiments and field observations provide evidence of the tendency for invasive species to outcompete native species in a changing climate (Verlinden and Nijs, 2010; Willis *et al.*, 2010). First, invasive species have a higher propensity than native species to adjust their phenology in accordance with climatic changes. The more adaptable phenologies of invasive species facilitate community invasions and also lead to increases in the abundance of invasive species (Willis *et al.*, 2010). Moreover, characteristics common to invasive species such as high dispersal abilities, high growth rates, short generation times, and broad climatic tolerances facilitate rapid range expansion in accordance with the rapid changes in climatic conditions (Schweiger *et al.*, 2010; Hellmann *et al.*, 2008a). Bioclimatic models can forecast the extent of a species' invasive potential in a changing climate by projecting the distribution of suitable climate for the invader. Forecasts of potential invasions from climate-change-induced expansions comprise most of the recent research on the interaction of invasive species and climate change. Generally, these models predict expansions of the invaders' ranges (Bradley *et al.*, 2010; Jarnevich and Stohlgren, 2009). However, in regions where they anticipate contractions, forecasts can provide useful guidance for restoration of sites that are no longer suitable for an existing invader (Bradley and Wilcove, 2009). Incorporating invasive ranges into bioclimatic models may also be useful for improving the forecasts of species distributions by more accurately approximating fundamental climatic niches (Beaumont *et al.*, 2009).

Climate-change-induced range shifts of native species may challenge the traditional definitions of nonnative and invasive species as ranges expand beyond species' historical distributions. Previously noninvasive species have the potential to become invasive in a new region without the biological controls provided by interspecies interactions present in the previous community (Hellmann *et al.*, 2008a). Bioclimatic models do not predict these changes in fitness or species interactions and so may underestimate the potential for ecosystem invasions from previously noninvasive species.

Land-Use Change

Changes in land use, and the corresponding habitat destruction, are currently the greatest threats to biodiversity (Hoffmann *et al.*, 2010; Dawson *et al.*, 2011). The interaction between climate change and land use may exacerbate the impacts of both stressors to flora and fauna. Land use may limit species' range expansions by inhibiting population establishment or impeding movement to climatically suitable regions (Feeley and Silman, 2010). This inability to realize range expansions may result in reductions in range size and decreases in species richness. Climate projections and associated response models can be used to assess these potential interactions and compounding impacts. For example, a process-based dynamic global ecosystem model forecasted a shift in the climatic conditions that are associated with high species richness in northern South America from less impacted to highly modified landscapes that cannot support as many species, resulting in a reduction in overall richness (Higgins, 2007).

Not only will land use inhibit range expansions, but it may also cause distributional shifts (Hockey *et al.*, 2011; Gehrig-Fasel *et al.*, 2007). These range shifts either augment climate-induced shifts or result in shifts inconsistent with the direction of climatic changes. Climate projections can also be coupled with land-use projections to anticipate species responses to these combined threats. For example, in Switzerland, the broad-scale changes forecasted in the distributions of nonvascular plants were attributed to climate change, whereas more local scale changes were attributed to land-use change (Nobis *et al.*, 2009). An individual-based population model anticipated that projected land-use changes will have a larger impact on habitat quality than climate change, but that climate change is still likely to impact the population dynamics of two bird species in the Willamette National Forest (McRae *et al.*, 2008).

Land use may also act as an additional driver of phenological advancement and may further exacerbate the impacts associated with phenological changes (Neil *et al.*, 2010). For example, urbanization may create warmer microclimates resulting in further advancement of phenology (Neil *et al.*, 2010). Together, land-use change and climate change also lead to increases in the prevalence of invasive species and diseases, further challenging floral and faunal communities (Crowl *et al.*, 2008; Patz *et al.*, 2008). Modeling approaches that anticipate impacts will be important for proactive management to address these potential compounding threats.

Disturbance

Climate change has the potential to alter the frequency and severity of disturbances, such as fire, resulting changes in ecosystem structure and/or function. Climate change has already altered fire regimes through increases in fire severity and area burned, and these impacts are projected to increase in the future (Littell *et al.*, 2009; Marlon *et al.*, 2009; Nitschke and Innes, 2008; Podur and Wotton, 2010). Increases in fire severity alter community composition through the loss of nonvascular plants, shrubs, and drought-sensitive species and

the rapid species establishment that succeeds severe burns (Colombaroli *et al.*, 2007; Bernhardt *et al.*, 2011). Climatic changes may also increase the fire susceptibility of communities that have historically lacked fire regimes. Ignition in these newly susceptible regions will further alter community composition (Malhi *et al.*, 2009).

Many studies have forecasted changes in fire severity and the area burned. These projections are significantly different from historical time-series simulations (Keane *et al.*, 2008) and generally include longer fire seasons, increases in fire severity, shifts toward full crown fires, and increases in the area burned (Podur and Wotton, 2010; Nitschke and Innes, 2008; Malevsky-Malevich *et al.*, 2008). However, far fewer studies forecast the impacts of these fire-regime changes on species and community composition.

The Human Response to Changes in Ecosystem Services

Through impacts on species, communities, and ecosystems, climate change will alter the goods and services provided by Earth's ecosystems. Ecosystem services include a wide array of benefits that people derive from ecosystems including provisioning, regulating, cultural, and supporting services. Provisioning services are those that deliver products that humans use (e.g., water, food, fiber for clothes, and shelter). Regulating services are those that control the states and rates of physical and biotic systems and processes in ways that are beneficial to humans (e.g., the reduction of storm-surge damage by mangrove forests, flood control by riparian systems, and carbon sequestration and storage by plants). Cultural services increase societal and spiritual well-being (e.g., the provision of recreational amenities, spiritual experiences, esthetics, and human health). Supporting services are those that assist in the provision of all other services such as nutrient and water cycling, pollination, and nitrogen fixation.

There are many ways in which climate change will alter the quantity or quality of the four different types of services. For example, shifts in species distributions have the potential to directly affect the provision of many food resources, particularly for human communities that rely on wild-caught foods. Climate change will also affect regulating services such as the ability to grow specific crops in particular locations (Lobell and Asner, 2003; Lobell and Field, 2007; Battisti and Naylor, 2009) and the quantity and quality of water for drinking and irrigation (Vörösmarty *et al.*, 2000). Changes in phenology, pollinator communities, soil microbial communities, the distribution of pest and pathogens, and invasive species all have the potential to alter supporting services for food production. Changes in microbial communities also have the potential to alter nutrient cycling, water purification, and timber and other fiber production.

Understanding how ecosystems services will change in the future and how humans will respond will be a critical step in developing adaptation strategies for species and systems. Crop failure, water shortages, and sea-level rise will force human migrations and likely result in conflicts between humans and terrestrial species, particularly in developing countries. These migrations and conflicts will further displace some terrestrial species and potentially provide opportunities for others. Shifts in the modes of energy production also have the potential to negatively impact species as solar arrays, wind farms, and biomass production replace oil and gas extraction. Efforts to increase carbon sequestration, however, have the potential to benefit species depending on the approach taken.

Studies have begun to examine the potential effects of climate change on specific ecosystem services; however, few studies have attempted to evaluate the impacts of climate change on suites of ecosystem services (Hayhoe *et al.*, 2004; Metzger *et al.*, 2008; Schröter *et al.*, 2005; Alcamo *et al.*, 2005). Many of the methods that are applicable to projecting changes in ecosystem services are those that can be used to assess the potential effects of climate change on species and ecological systems. In addition to these approaches, methods for quantifying the services themselves will also be needed. Many models have been devised for evaluating ecosystem services (Kareiva *et al.*, 2011), and some of these have been coupled with climate projections to investigate the impacts of climate change on these services (Lawler *et al.*, 2011). Estimating the human response to these changes in services requires going beyond the ecosystem service model projections and forecasting human migrations, land-use changes, water use, and other human actions. These predictions can, like the projections of ecological responses, be derived from predictive models or be based on historic patterns or the results of surveys and other socioeconomic studies.

Adaptation

There are two types of actions that humans can take to reduce the impacts of climate change. Mitigation actions reduce the amount that the climate will change. Such actions include approaches for reducing greenhouse gas emissions and sequestering atmospheric carbon dioxide. Adaptation actions are those that reduce the impact of a given magnitude of climate change and include approaches such as protecting more land, restoring riparian areas to reduce stream temperatures, and moving species to more suitable climates. There are several general concepts that have guided the development of adaptation strategies. These are discussed in the section, General Concepts and Principles, below. This section is followed by sections describing specific Adaptation Strategies, Challenges to Adaptation, and the co-benefits of adaptation (in Multidimensional Strategies).

General Concepts and Principles

Many of the commonly recommended climate-change adaptation strategies for species and ecological systems tend to be general concepts for altering the way species and systems are managed. These concepts include assessing vulnerability, managing for resistance and resilience, planning at broader spatial scales, increasing cooperation and coordination in planning and management, and employing adaptive management and triage when necessary. These general concepts are discussed briefly.

Vulnerability

Climate-change adaptation is deeply tied to the concept of vulnerability – the extent to which a species, an ecosystem, or

an area is likely to be harmed as a result of climate change and associated stresses (Brooke, 2008). Vulnerability can be defined as having three components: sensitivity, exposure, and adaptive capacity (Dawson *et al.*, 2011). Sensitivity refers to innate characteristics of an organism or ecosystem (e.g., tolerance to changes in temperature) that predispose it to being more or less susceptible to climate change. Exposure refers to the amount of change either in climate or in climate-driven factors that a species or system will face. Adaptive capacity is a measure of the ability of a species or system to respond in a way that reduces the impacts of, or to takes advantage of, climate change. Vulnerability assessments inform adaptation planning by documenting aspects of all three of these components of vulnerability and using them to identify which species or systems will be affected by projected climate changes and the causes of those impacts Glick *et al*. (2011).

Resistance, Resilience, and Fostering Change

There are three general approaches to managing systems in the face of climate change. One can attempt to manage a system to resist change, to make it resilient to climate change, or to foster change to a new state. Much discussion has focused on managing for resilience. Resilience can be defined in several ways. Here, resilience is defined as the ability of a species or system to return to its current state following a perturbation (e.g., Holling, 1973). Systems that are resilient to climate change will be able to maintain ecosystem functions and processes and avoid a transition to a new state as climates change. Managing for resilience is an attractive concept, because, if successful, it allows continued delivery of ecosystem services and, possibly, the persistence of plant and animal populations.

Many of the other general concepts as well as most of the more specific adaptation strategies focus on increasing resilience. For example, reducing other (nonclimatic) threats, increasing genetic diversity, and restoring riparian vegetation are all approaches that will likely increase the resilience of a species or system in a given place. Removing barriers to inland migration of coastal species is also a potential strategy for increasing the resilience of a species, albeit by increasing the adaptive capacity of species to address change.

Given the magnitude of change that is projected for the coming century, it may be impossible to foster resilience in many systems. Managing these systems may require promoting change to a new state. Promoting change in human-dominated systems will likely be more straightforward (e.g., planting new crops) than doing so in more natural systems. Strategies for fostering change include assisted colonization (see below), shifting management efforts from one species to another and more specifically promoting newly arriving species. The outcomes of these more aggressive and forward-looking strategies will be more uncertain and their implementation (as is already evident with assisted colonization, Ricciardi and Simberloff, 2009) will be more controversial.

Spatial and temporal scales

Climate change is a global phenomenon that will result in long-term changes in ecological systems. The spatial and temporal scales of the changes one is likely to see far exceed the scale of the traditional 1 m^2 ecological sample plot and the time-span of an ecological study. More importantly, these changes exceed the spatial and temporal scales over which most planning and management occur. Managing for species and systems in a changing climate will require taking a broader spatial and temporal perspective (Welch, 2005; CCSP, 2008). Developing climate-change adaptation strategies for species will require planning at regional and, perhaps in some instances, continental scales (Hannah, 2010). Because ecological systems will be undergoing relatively rapid changes for the foreseeable future, managers will have to manage for moving targets. The whole concept of ''restoration'' must be reconsidered (Harris *et al.*, 2006). Returning ecosystems to a previous state (e.g., before human settlement) with a complement of species that were historically native to the site might be counterproductive in a changing climate.

Similarly, the time frame for planning and management actions will, in many cases, need to shift to address climate change. Planning horizons have traditionally been relatively short – for example, 3- to 5-year plans. Planning in such small increments is likely to result in short-sighted management strategies that are incapable of addressing climate change. Managers and scientists need to strike a balance in terms of the temporal and spatial scales over which climate impacts are projected and the scales at which strategies and plans are developed. Many of the ecological climate-impact forecasts are for periods of 50–100 years in the future. Although these projections will be useful for developing management strategies, mangers will also need 10- and 20-year projections.

Cooperation and Coordination

Given the broad spatial scales that will need to be considered to address climate change, managers, planners, and policy makers will need to work together and to coordinate efforts across jurisdictions and traditional management units (Kareiva *et al.*, 2008). This will mean developing regional instead of local efforts in which strategies are designed to reach across state, province, or country borders. It will mean collaboration across different governmental agencies that are currently focused on specific systems (atmospheric, marine, freshwater, and terrestrial). It will also mean increasing cooperation among different types of groups such as nongovernmental organizations, local and regional governments, citizen groups, and industry. The US Climate Change Science Program (CCSP) is one such organization that reaches across federal agencies. Other regional and national efforts have been put in place in other parts of the world, but such efforts will need to be increased to address climate change.

Managing for Process Over Composition

Because changes in climate will drive species range shifts resulting in new communities, it will be exceedingly difficult to manage for specific assemblages of species. Many have suggested that such changes will necessitate a shift from managing for species and communities to managing for ecosystem processes and shifting baselines (Harris *et al.*, 2006; Heller and Zavaleta, 2009; Hobbs *et al.*, 2009). In many systems, managing for ecosystem function may mean taking a forward-looking approach in which new species are planted or introduced into a system. As discussed below in the context of assisted colonization, these more manipulative approaches

that tend toward engineering ecosystems and moving species have raised significant concerns and debates about the unintended consequences of such actions.

Triage

Many species and systems will be highly vulnerable to climate change. Given the relative scarcity of funds for conservation, it will be impossible to address the needs of all species. Thus, conservation planners and managers will have to make difficult decisions about how to allocate limited funds. They will likely have to choose which populations and species to try to save and which to let go extinct. Triage is one approach to making such difficult prioritizations.

Triage is a medical ranking system developed for the treatment of patients in emergency situations. Traditionally, the system has had four levels based on the severity of the injuries and the likelihood that the patient will survive. The levels are generally deceased or expectant, critical, severe, and minor. A similar system can be applied to species or populations at risk of extinction in the face of climate change (Lawler, 2009). In such a system, some species would be classified as likely to experience such large changes that management actions will do little to prevent their decline or loss. Other species would be deemed as critically threatened by climate change, but their decline or extinction could be prevented by immediate and intense management efforts. A third class of species would be threatened by climate change but would not be at risk of extinction in the short term, and thus management effort could wait. A final class of species would face relatively minor threats from climate change and could be monitored over time and addressed later if necessary. Although there are both ecological and ethical considerations that make triage unpopular (Kareiva and Levin, 2003), such a system will undoubtedly become necessary as climates change and ecological systems respond.

Adaptive Management

Adaptive management is an iterative process of management and monitoring in which management actions are treated as experiments (Holling, 1978; Walters and Hilborn, 1978). The outcomes of these experiments are used to inform the next round of management actions, which are again treated as experiments. Adaptive management was conceived of for managing uncertain systems, making it potentially a highly useful approach to deal with climate change (Arvai *et al.*, 2006). To address climate change, adaptive management approaches will have four basic steps that will be iteratively repeated (Kareiva *et al.*, 2008). The first step will involve assessing the potential impacts of climate change on the system in question. In the second step, management actions, in the form of experiments with testable hypotheses or predictions, will be designed to address one or more of these potential impacts. The third step involves monitoring the system for climatic changes and system responses. Finally, the results of the monitoring can be used to evaluate the effectiveness of the management actions and to redesign them as necessary before the four steps are repeated.

Adaptation Strategies

Climate-change adaptation strategies are practices and adjustments that enhance resilience and/or reduce vulnerability to changes in climate (IPCC, 2007a). These strategies have been implemented across the globe on a limited but increasing basis, in both developed (Krysanova *et al.*, 2010; Olesen *et al.*, 2011; Tompkins *et al.*, 2010; Wheeler, 2008) and developing countries (Krysanova *et al.*, 2010; Mertz *et al.*, 2009; Ziervogel and Zermoglio, 2009). The following section describes some of the adaptation strategies that have been proposed for protecting terrestrial species in a changing climate.

Protected Areas: Reserve Selection and Design

By providing species with refuge from many threats, protected areas have the potential to increase the resilience of species and populations to climate change (Hunter *et al.*, 2010). However, climate change brings into question the ability of current protected lands to provide for the biodiversity of the future. Typically, reserve boundaries are static and reserve networks are designed based on current biodiversity. As species' ranges shift with a changing climate, some species may lose protection as their ranges shift out of reserve boundaries, whereas others may move into reserves (Peters and Darling, 1985). Reserves across the globe are anticipated to experience changes in biodiversity composition, and several species may lose protection entirely from the current reserve network (Araújo *et al.*, 2004; Hole *et al.*, 2009). One adaptation response that aims to provide protection for anticipated climate-induced changes in species distributions has been to augment the reserve network by adding reserves to increase the total area protected across a landscape.

Several general rules of thumb have been proposed for addressing climate change in the reserve selection (i.e., the placement of new reserves) and reserve design (i.e., the size and shape of reserves) processes. Most simply, many have suggested increasing the total area protected and increasing the size of existing reserves. All else being equal, larger reserves will be more likely to maintain populations of species as climates change by providing the space to facilitate within-reserve range shifts. Thus, designating larger reserves or designating buffer zones around reserves may help to protect more species for longer periods of time in a changing climate (Peters and Darling, 1985). Similarly, reserves that span strong environmental gradients will also provide more opportunities for within-reserve range shifts by providing future niches at different altitudes as the climate changes. Strategies for the placement of new reserves include placing reserves at the elevational or poleward range limit of key species (Peters and Darling, 1985), at major transitions between vegetation formations (Halpin, 1997) or at the core of species environmental distributions (Araújo *et al.*, 2004). Yet another simple suggestion involves increasing the redundancy in the reserve system. Protecting the same species in multiple places provides multiple opportunities for the species to weather or adapt to climate change.

Others have suggested more sophisticated methods for identifying the best places for new protected areas to address climate change. For example, projected shifts in species distributions can be used to identify areas that are likely to protect

species today as well as into the future under multiple climate-change scenarios (Hannah *et al.*, 2007; Vos *et al.*, 2008). The increase in specificity potentially makes this approach more effective for the targeted species; however, it incorporates higher levels of uncertainty inherent in forecasting future climatic changes and the biotic responses to those changes. For example, the most-often-applied approach for forecasting species range shifts, species distribution or niche modeling, is better suited to describe general patterns and trends in range shifts than to identify particular locations for protecting specific species in the future (Pearson and Dawson, 2003).

Another approach to locate reserves to address climate change involves selecting areas that protect the underlying environmental gradients that largely determine patterns of biodiversity at broader scales, but that will likely remain stable as the climate changes. It has been argued that protecting these underlying gradients essentially protects the stage on which biodiversity will play out as climates change. Some proposed strategies based on abiotic features focus on selecting sites that span elevational gradients or that represent heterogeneity in soils (Peters and Darling, 1985). Other have stressed geologic variability in a reserve network (Anderson and Ferree, 2010), preserving a range of current climatic conditions – with the assumption that future climates may change, but that many climatic gradients will be preserved (Pyke and Fischer, 2005) – and conserving the regional diversity of land facets, or unique combinations of abiotic conditions (i.e., topographic and edaphic features) (Beier and Brost, 2010). These strategies seek to maintain the elements of the landscape that are responsible for the distribution of regional biodiversity, may be more effective due to more specificity in reserve placement, and are robust to the uncertainties of forecasts.

Other abiotic-based strategies target climate conditions by placing reserves in regions projected to be climate refugia. Planners refer to climate refugia both as areas that are projected to experience minimal climatic changes and as areas projected to have cooler microclimates (Hansen *et al.*, 2010; Shoo *et al.*, 2011; Saxon *et al.*, 2005). The former aims to reduce the potential changes in species composition of new reserves so that these reserves will be more effective over a longer period of time for currently protected species. The latter aims to reduce the distance necessary for distributional shifts by locally protecting cooler microclimates. Locally cooler microclimates may occur at slightly higher elevations or be in areas with more vegetative cover or potential for vegetative cover. Although these strategies may depend on uncertain climate projections, they avoid the compounding uncertainty of using climate projections to forecast biotic responses and still incorporate specificity into the planning process.

Connectivity

Range shifts have occurred in the past as species have responded to historical changes in climate. However, during these periods of historical climate change, species movements were not hindered by anthropogenic barriers. Fragmentation of habitat, roads, and other barriers can inhibit species movement or survival in regions of or between newly suitable climates. Not surprisingly, increasing landscape connectivity is the most commonly recommended adaptation approach for addressing climate impacts on biodiversity (Heller and Zavaleta, 2009).

Until recently, suggested strategies for improving connectivity to address climate change have been somewhat general – for example, placing new reserves between existing reserves, creating a system of stepping stone reserves, or adding reserves in proximity to existing reserves. Recently, studies have highlighted more sophisticated approaches to connecting landscapes to address climate change. One such approach uses climate projections to orient corridors and expand existing reserves in the direction of anticipated climatic changes (Ackerly *et al.*, 2010). Other approaches involve using projected shifts in species distributions to identify potential routes that species would need to take to move from currently suitable climates to places where climates will likely be suitable in the future (Williams *et al.*, 2005; Rose and Burton, 2009). Beier and Brost (2010) recommend using abiotic elements or land facets (described above) to define corridors. Similar to the multiple reserve placement approaches, these connectivity approaches vary in the level of uncertainty in the data, models, and model projections that they draw upon.

Although it will be important to identify potential movement corridors, increasing the permeability of the landscape in general will also be critical (Franklin and Lindenmayer, 2009). It may be possible to manage lands to facilitate species' movements in response to climate change. For example, selective harvest or retention cuts, tree planting, and rotational grazing may provide habitat to facilitate range expansions (Manning *et al.*, 2009; Kohm and Franklin, 1997). It is possible to prioritize areas for these incentives and management actions using approaches similar to those discussed above for the placement of reserves or corridors.

Species- and Place-Specific Approaches

The adaptation approaches discussed above tend to be general concepts or broad-scale actions (reserve selection and promotion of landscape connectivity). Despite the fact that climate change is a global issue and landscape-level actions are therefore necessary, species- and place-specific approaches will also likely be critical components of climate-change adaptation. Species-specific adaptation strategies aimed at preventing extinction of a particular organism may be appropriate for threatened and endangered species for which legal frameworks, such as the US Endangered Species Act and the International Union for Conservation of Nature (IUCN) Red List, may require species-specific protection and for organisms of economic importance to humans. Place-based approaches, such as efforts focused on a particular park or preserve or within a particular city, are important because adaptation strategies are likely to be implemented at local scales, by the private and public landowners and managers. Proactive management can address the vulnerability of an area or a species by reducing exposure through the manipulation of microclimates, by reducing sensitivity, and by increasing adaptive capacity.

There are several ways in which the degree of local exposure to climate change can be altered. For example, species-based approaches may include supplemental watering of key plant species in drought years (e.g., Pavlik *et al.*, 2002) and altering microclimates of artificial nest boxes for rare

birds by painting boxes white or locating them on north-exposed slopes (Catry *et al.*, 2011). Place-based approaches can rely on microclimatic variations that provide refugia from large-scale changes in climate (Mosblech *et al.*, 2011). Place-based approaches to reduce exposure may therefore include protection of these natural refugia, as well as planting trees and other vegetation to provide shade and reduce temperatures (Wilby *et al.*, 1998; Wilby and Perry, 2006). These relatively fine scale adaptation actions are often successful, at least in the short term, but may become increasingly expensive, time consuming, and impractical as climate-change progresses.

Adaptation strategies can also address the vulnerability of an area or a species by reducing sensitivity to the effects of climate change. Diverse systems are generally less sensitive to disturbances (Kareiva *et al.*, 2008; Fargione and Tilman, 2005; Schindler *et al.*, 2010). This pattern is true for genetic diversity within a population of a single species (Schindler *et al.*, 2010), as well as for species richness at the ecosystem level (Chapin *et al.*, 2000). Thus, promoting and protecting diversity at multiple levels, from within a species or a population to across different species and functional groups in an ecosystem, is an important adaptation strategy (Glick *et al.*, 2009).

Adaptation strategies can address vulnerability of an area or a species by reducing sensitivity to climate change or by strengthening its adaptive capacity to respond to the effects of climate change. Other stressors, such as land use, invasive species, pathogens, fragmentation, and pollutants, interact with climate change to further harm species and natural systems (Schweiger *et al.*, 2010; Walther, 2010). Reducing these other stressors can increase adaptive capacity and decrease sensitivity to climate change (Glick *et al.*, 2009). Species-specific approaches include reducing harvest levels of focal organisms, limiting their exposure to pollutants and pathogens, reducing habitat loss, and improving connectivity between populations through corridors (Heller and Zavaleta, 2009; Glick *et al.*, 2009). Worldwide conservation efforts have already reduced species' extinctions, particularly by mitigating threats from invasive species on birds and mammals, but efforts need to be augmented in order to abate many substantial threats to global diversity (Hoffmann *et al.*, 2010).

Assisted colonization – the translocation of species outside their native range to facilitate movement in response to climate change – is another strategy for increasing the adaptive capacity of a species. Assisted colonization may be most appropriate in cases in which current species' ranges become inhospitable with climate change and connectivity to suitable regions is severely limited. This approach is controversial because of the potential for negative, ecological, evolutionary, and economic impacts, as well as ethical concerns (Ricciardi and Simberloff, 2009; Sax *et al.*, 2009). Some potentially less controversial strategies include planting climate-resistant species or ecotypes (Glick *et al.*, 2009) and establishing "neo-native forests" in restoration efforts, that is, planting species where they were in the past, but are not found currently (Millar *et al.*, 2007). Many current restoration and forestry practices adhere to strict rules about using only "local" species and ecotypes. To prepare for climate change, one may need to broaden the genetic and species diversity used in restoration and forestry (Glick *et al.*, 2009). In some cases, even aggressive strategies such as assisted colonization may fail as climate change and other stresses threaten the existence of rare species. These extreme cases will require *ex situ* conservation strategies, including seed banking and captive breeding to ensure the long-term survival of the species.

Protected areas, corridors for movement, and assisted colonization may all help species track changes in climate. However, given the magnitude of projected changes, these measures may be insufficient. It may be necessary to implement policies that facilitate species movements. Such policies may prohibit the destruction of habitat or removal of species in places through which those particular species will likely need to move in response to climate change (Kostyack *et al.*, 2011).

Challenges to Developing Successful Adaptation Strategies

There are several challenges to adaptation planning and implementation. One of the most frequently mentioned challenges to developing climate-change adaptation strategies is the lack of certainty about future climatic conditions. Although climate-change projections and projected climate impacts are necessarily uncertain, some projections are less uncertain than others and many projections can be used to inform the development of adaptation strategies. In general, projected changes in temperature are less uncertain than projected changes in precipitation and projections for the near-term are less uncertain than projections for the more distant future (IPCC, 2007b). Projected climate impacts on terrestrial species will necessarily be more uncertain than climate-change projections as they generally incorporate the uncertainties of climate models and the additional uncertainties about how species or habitats will respond to climate change and/or the uncertainty associated with the modeling approaches used to forecast those responses (Thuiller, 2004).

A second, often-cited challenge to developing adaptation strategies is the lack of projected climatic changes at a spatial resolution that is meaningful for management. The general circulation models that have been used to project changes in climate have, to date, generated projections at sales of degrees of latitude and longitude (IPCC, 2007b). These projections can be downscaled to finer resolutions; however, there are uncertainties associated with downscaling process (Dettinger, 2005; Wilby *et al.*, 1998). Despite the additional uncertainties, downscaled projections are important for understanding how climatic changes may manifest at finer scales. Downscaled data sets are currently available for many regions of the globe, and tools have been developed that make those data sets readily accessible (Girvetz *et al.*, 2009).

A third limitation to developing adaptation strategies is the general lack of information on species-specific responses to climate. Responses to interacting climate variables vary greatly across and within species, and for many terrestrial plants and animals, the relationship between climate variables and performance (i.e., growth, survival, and reproduction) is unknown (Parmesan, 2006). Climate-change effects on species are further complicated by their interactions with biotic factors, such as competition, predation, mutualisms, and other stressors, such as land use, invasive species, pathogens, and pollutants (Schweiger *et al.*, 2010; Walther, 2010). Thus,

for many species, an important first step for climate-change adaption is simply to improve the understanding of responses to climate change, including physiological, behavioral, and demographic responses (Heller and Zavaleta, 2009)

Planning can be difficult in the face of these uncertainties, and currently implemented adaptation strategies are limited in scope and are insufficient to fully address climate-change impacts (Wheeler, 2008; Reyer *et al.*, 2009). One of the primary concerns expressed by land and resource mangers is a lack of "downscaled" studies on which to base their decisions, in terms of both climate projections and species responses (Glick *et al.*, 2009). However, there have been significant advances in model development at regional scales. As one's knowledge base and expertise grow, a key challenge is to improve communication and cooperation across the many groups and individuals involved in developing climate-change adaptation strategies. Species- and place-specific approaches are practical, because of existing legal, ownership, and management frameworks. Nonetheless, these approaches will be far more effective if there is substantial cooperation and communication within and between groups implementing climate-change adaptation strategies.

Table 1 Example approaches for anticipating and adapting to climate impacts on terrestrial species

Impacts	*Anticipating impacts*	*Adaptation strategies*
Phenology		
Shift in grape flowering and ripening times in montane Australia (Caffarra and Eccel, 2011)	• Project changes in temperature and precipitation • Apply phenological models of budburst, flowering, and fruit ripening	• Alter grape varieties grown • Focus on high elevation areas, where climate may become more suitable
Range shifts		
Shift in the distribution of wolverine in the US	• Project changes in snowpack • Apply species distribution models • Map historical ranges and climatic conditions	• Enhance connectivity both northward in latitude and upward in elevation
Population processes		
Shift to male-biased sex ratios of tuatara reptiles in New Zealand (Mitchell *et al.*, 2010)	• Conduct population viability analysis under different sex ratios	• Use *ex situ* techniques to increase females • Translocate populations to cooler environments or sites with more microclimate complexity
Interspecific interactions		
Spatial mismatch of a monophagous butterfly (*Boloria titania*) and its larval host plant (*Polygonum bistorta*) (Schweiger *et al.*, 2008)	• Project changes in monthly temperature • Apply species distribution models for two interacting species • Map future ranges of two species	• Increase likelihood of plant dispersal by enhancing connectivity between habitat patches • Facilitate migration of host plant through habitat restoration or planting efforts
Diseases and parasites		
Change in the prevalence of East Coast Fever (ECF) in sub-Saharan Africa	• Project changes in distribution of a tick and 10 of its host species • Assess the probability of occurrences of tick–host assemblages • Apply dilution effect model to determine potential ECF transmission	• Regulate cattle movement to reduce interaction between cattle and wild host species
Interactions with other stressors		
Impact of multiple stressors on two bird species with varying habitat requirements	• Develop spatially explicit individual-based models • Apply projected changes in land use • Apply projected changes in climate	• Evaluate and implement management strategies that increase population viability

Multidimensional Strategies (Co-benefits)

Climate change will have profound effects on many human populations (IPCC, 2007a). These effects include water shortages, crop reductions, and failures resulting in malnutrition, displacement due to sea-level rise, emerging diseases, and armed conflict (Costello *et al.*, 2009). Humans will develop adaptation strategies to reduce these impacts. Dams will be built to store water; more water will be extracted from streams and rivers for agriculture, drinking, and manufacturing; new areas will be cultivated; migrations will occur and settled areas will expand; and wetlands will be drained and pesticides applied to reduce crop pests and disease vectors. Many of these adaptation strategies will have negative consequences for nonhuman terrestrial species. Conversely, some of these strategies can be designed to benefit nonhuman terrestrial species.

Co-benefits can be defined as the positive effects that an adaptation strategy designed to address a specific human need has on nonhuman species, communities, or ecosystems, or conversely the positive effects on human well-being of adaptation strategies designed to help nonhuman species and systems adapt to climate change. Co-benefits have largely been described for mitigation strategies that, in addition to reducing greenhouse gas emissions, benefit human health (Haines *et al.*, 2007; Jack and Kinney, 2010). Such co-benefits of adaptation strategies have not been as well documented, but these win–win outcomes will likely arise in many instances.

There are clear co-benefits for human health and well-being that are likely to result from adaptation strategies that restore coastal habitats, riparian areas, and flood plains. Removing sea walls, restoring mangrove forests and dune systems, and protecting barrier beaches are all adaptation strategies that will increase the resilience of coastal plant and animal communities to sea-level rise and storm surges. These same strategies, if well designed and well sited, have the potential to protect human communities from these threats as well. Conversely, addressing these threats to humans by building sea walls will not offer the co-benefit to nonhuman species and systems. As a second example of a co-benefit, riparian restoration has the potential to reduce stream temperatures through shading and may also increase water quality and water storage for human use.

To develop strategies that maximally benefit both nonhuman and human systems and communities, it will be necessary to develop approaches for evaluating and comparing co-benefits. As of yet, little work has been done in this area, in part because it involves measuring and weighing disparate outcomes and values and because it involves collaboration and coordination across diverse disciplines. Nonetheless, such comparative metrics and analyses will be quite useful for developing strategies and prioritizing among them.

Conclusions

Addressing climate-change impacts on terrestrial species will require at least a basic understanding of how climate change will affect species. Much is already known about climate effects on phenologies, distributions, populations, interspecific interactions, and diseases and pathogens. Although there are significant gaps in the knowledge of these effects, detailed studies of potential climate impacts on specific species are not likely to provide managers and policy makers with the most critical information for developing adaptation strategies for terrestrial species. On the contrary, research focused on developing actionable adaptation strategies from the many general concepts and basic principles that have been proposed for addressing climate change has the potential to inform the management of a wide range of species and systems. For example, research focused on approaches to adaptive management, refining triage systems, connecting landscapes to facilitate range shifts, and measuring and valuing the co-benefits of adaptation strategies for human and nature systems will be critical components of a research agenda for addressing climate impacts on terrestrial species and systems (**Table 1**).

See also: Climate Change and Ecology, Synergism of. Climate Change and Wild Species. Climate, Effects of. Comparing Extinction Rates: Past, Present, and Future. Conservation Efforts, Contemporary. Evolution in Response to Climate Change. Landscape Corridors. Latitudinal and Elevational Range Shifts Under Contemporary Climate Change. Mammals, Conservation Efforts for. Phenological Shifts in Animals Under Contemporary Climate Change. Plant Phenology Changes and Climate Change. Species Distribution Modeling. Wildlife Management

References

Ackerly DD, Loarie SR, Cornwell WK, *et al.* (2010) The geography of climate change: Implications for conservation biogeography. *Diversity and Distributions* 16: 476–487.

Alcamo J, Vuuren DV, Ringler C, *et al.* (2005) Changes in nature's balance sheet: Model-based estimates of future worldwide ecosystem services. *Ecology and Society* 10. Article 19.

Anderson MG and Ferree CE (2010) Conserving the stage: Climate change and the geophysical underpinnings of species diversity. *PLoS ONE* 5: e11554.

Araújo MB, Cabeza M, Thuiller W, Hannah L, and Williams PH (2004) Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology* 10: 1618–1626.

Araújo MB and Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16: 743–753.

Arvai J, Bridge G, Dolsak N, *et al.* (2006) Adaptive management of the global climate problem: Bridging the gap between climate research and climate policy. *Climatic Change* 78: 217–225.

Battisti DS and Naylor RL (2009) Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* 323: 240–244.

Beaumont LJ, Gallagher RV, Thuiller W, *et al.* (2009) Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions* 15: 409–420.

Beier P and Brost B (2010) Use of land facets to plan for climate change: Conserving the arenas, not the actors. *Conservation Biology* 24: 701–710.

Bernhardt EL, Hollingsworth TN, and Chapin FS (2011) Fire severity mediates climate-driven shifts in understorey community composition of black spruce stands of interior Alaska. *Journal of Vegetation Science* 22: 32–44.

Bradley BA and Wilcove DS (2009) When invasive plants disappear: transformative restoration possibilities in the western United States resulting from climate change. *Restoration Ecology* 17: 715–721.

Bradley BA, Wilcove DS, and Oppenheimer M (2010) Climate change increases risk of plant invasion in the eastern United States. *Biological Invasions* 12: 1855–1872.

Brooke C (2008) Conservation and adaptation to climate change. *Conservation Biology* 22: 1471–1476.

Brooks DR and Hoberg EP (2007) How will global climate change affect parasite–host assemblages? *Trends in Parasitology* 23: 571–574.
Buisson L, Thuiller W, Casajus N, Lek S, and Grenouillet G (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology* 16: 1145–1157.
Caffarra A and Eccel E (2011) Projecting the impacts of climate change on the phenology of grapevine in a mountain area. *Australian Journal of Grape and Wine Research* 17: 52–61.
Canto T, Aranda MA, and Fereres A (2009) Climate change effects on physiology and population processes of hosts and vectors that influence the spread of hemipteran-borne plant viruses. *Global Change Biology* 15: 1884–1894.
Carroll C (2010) Role of climatic niche models in focal-species-based conservation planning: Assessing potential effects of climate change on Northern Spotted Owl in the Pacific Northwest, USA *Biological Conservation* 143: 1432–1437.
Catry I, Franco AMA, and Sutherland WJ (2011) Adapting conservation efforts to face climate change: Modifying nest-site provisioning for lesser kestrels. *Biological Conservation* 144: 1111–1119.
CCSP (2008) Preliminary review of adaptation options for climate-sensitive ecosystems and resources. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. In: Julius, SH, West JM (ed.), Washington, DC: U.S. Environmental Protection Agency.
Chapin FS, Zavaleta ES, Eviner VT, *et al.* (2000) Consequences of changing biodiversity. *Nature* 405: 234–242.
Colombaroli D, Marchetto A, and Tinner W (2007) Long-term interactions between Mediterranean climate, vegetation and fire regime at Lago di Massaciuccoli (Tuscany, Italy). *Journal of Ecology* 95: 755–770.
Costello A, Abbas M, Allen A, *et al.* (2009) Managing the health effects of climate change. *Lancet* 373: 1693–1733.
Crowl TA, Crist TO, Parmenter RR, Belovsky G, and Lugo AE (2008) The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment* 6: 238–246.
Dawson TP, Jackson ST, House JI, Prentice IC, and Mace GM (2011) Beyond predictions: Biodiversity conservation in a changing climate. *Science* 332: 53–58.
Dettinger MD (2005) From climate-change spaghetti to climate-change distributions for 21st century California. San Francisco Estuary and Watershed Science, 3, Article 4.
Dormann CF (2007) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology* 8: 387–397.
Elliott GP (2011) Influences of 20th-century warming at the upper tree line contingent on local-scale interactions: evidence from a latitudinal gradient in the Rocky Mountains, USA. *Global Ecology and Biogeography* 20: 46–57.
Estay SA, Lima M, and Labra FA (2009) Predicting insect pest status under climate change scenarios: Combining experimental data and population dynamics modelling. *Journal of Applied Entomology* 133: 491–499.
Fabina NS, Abbott KC, and Gilman RT (2010) Sensitivity of plant-pollinator-herbivore communities to changes in phenology. *Ecological Modelling* 221: 453–458.
Fargione JE and Tilman D (2005) Diversity decreases invasion via both sampling and complementarity effects. *Ecology Letters* 8: 604–611.
Feehan J, Harley M, and Van Minnen J (2009) Climate change in Europe. 1. Impact on terrestrial ecosystems and biodiversity. A review. *Agronomy for Sustainable Development* 29: 409–421.
Feeley KJ and Silman MR (2010) Land-use and climate change effects on population size and extinction risk of Andean plants. *Global Change Biology* 16: 3215–3222.
Forrest J and Miller-Rushing AJ (2010) Toward a synthetic understanding of the role of phenology in ecology and evolution. *Philosophical Transactions of the Royal Society B – Biological Sciences* 365: 3101–3112.
Franklin JF and Lindenmayer D (2009) Importance of matrix habitats in maintaining biological diversity. *Proceedings of the National Academy of Science of the United States of America* 106: 349–350.
Froeschke G, Harf R, Sommer S, and Matthee S (2010) Effects of precipitation on parasite burden along a natural climatic gradient in southern Africa – Implications for possible shifts in infestation patterns due to global changes. *Oikos* 119: 1029–1039.
Fussel HM (2008) Assessing adaptation to the health risks of climate change: What guidance can existing frameworks provide? *International Journal of Environmental Health Research* 18: 37–63.
Garrett KA, Dendy SP, Frank EE, Rouse MN, and Travers SE (2006) Climate change effects on plant disease: Genomes to ecosystems. *Annual Review of Phytopathology* 44: 489–509.
Garrett KA, Forbes GA, Savary S, *et al.* (2011) Complexity in climate-change impacts: An analytical framework for effects mediated by plant disease. *Plant Pathology* 60: 15–30.
Gehrig-Fasel J, Guisan A, and Zimmermann NE (2007) Tree line shifts in the Swiss Alps: Climate change or land abandonment? *Journal of Vegetation Science* 18: 571–582.
Girvetz EH, Zganjar C, Raber GT, *et al.* (2009) Applied climate-change analysis: The climate wizard tool. *PLoS One* 4: e8320.http://dx.doi.org/10.1371/journal.pone.0008320.
Glick P, Staudt A, Stein B (2009). A new era for conservation: review of climate change adaptation literature. Washington DC: National Wildlife Federation.
Glick P, Stein BA, and Edelson NA (eds.) (2011) Scanning the Conservation Horizon: A Guide to Climate Change Vulnerability Assessment. Washington, DC: National Wildlife Federation.
Guisan A and Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.
Haines A, Smith KR, Anderson D, *et al.* (2007) Energy and health 6 – Policies for accelerating access to clean energy, improving health, advancing development, and mitigating climate change. *Lancet* 370: 1264–1281.
Halpin PN (1997) Global climate change and natural-area protection: Management responses and research directions. *Ecological Applications* 7: 828–843.
Hannah L (2010) A global conservation system for climate-change adaptation. *Conservation Biology* 24: 70–77.
Hannah L, Midgley GF, Andelman S, *et al.* (2007) Protected area needs in a changing climate. *Frontiers in Ecology and the Environment* 5: 131–138.
Hansen L, Hoffman J, Drews C, and Mielbrecht E (2010) Designing climate-smart conservation: Guidance and case studies. *Conservation Biology* 24: 63–69.
Harris JA, Hobbs RJ, Higgs R, and Aronson J (2006) Ecological restoration and global climate change. *Restoration Ecology* 14: 170–176.
Hayhoe K, Cayan D, Field CB, *et al.* (2004) Emissions pathways, climate change, and impacts on California. *Proceedings of National Academy of Sciences* 101: 12422–12427.
Heller NE and Zavaleta ES (2009) Biodiversity management in the face of climate change: A review of 22 years of recommendations. *Biological Conservation* 142: 14–32.
Hellmann JJ, Byers JE, Bierwagen BG, and Dukes JS (2008a) Five potential consequences of climate change for invasive species. *Conservation Biology* 22: 534–543.
Hellmann JJ, Pelini SL, Prior KM, and Dzurisin JDK (2008b) The response of two butterfly species to climatic variation at the edge of their range and the implications for poleward range shifts. *Oecologia* 157: 583–592.
Higgins PAT (2007) Biodiversity loss under existing land use and climate change: An illustration using northern South America. *Global Ecology and Biogeography* 16: 197–204.
Hobbs RJ, Higgs E, and Harris JA (2009) Novel ecosystems: Implications for conservation and restoration. *Trends in Ecology & Evolution* 24: 599–605.
Hockey PAR, Sirami C, Ridley AR, Midgley GF, and Babiker HA (2011) Interrogating recent range changes in South African birds: Confounding signals from land use and climate change present a challenge for attribution. *Diversity and Distributions* 17: 254–261.
Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the World's vertebrates. *Science* 330: 1503–1509.
Hole DG, Willis SG, Pain DJ, *et al.* (2009) Projected impacts of climate change on a continent-wide protected area network. *Ecology Letters* 12: 420–431.
Holling CS (1973) Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics* 4: 1–23.
Holling CS (1978) *Adaptive Environmental Assessment and Management.* New York, NY: Wiley.
Hunter M, Dinerstein E, Hoekstra J, and Lindenmayer D (2010) A call to action for conserving biological diversity in the face of climate change. *Conservation Biology* 24: 1169–1171.
IPCC (2007a) *Climate change 2007: Impacts, Adaptation and Vulnerability, Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Cambridge University Press.
IPCC (2007b) Climate Change 2007: The Physical Science Basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL (eds.), *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Cambridge University Press.
Jack DW and Kinney PL (2010) Health co-benefits of climate mitigation in urban areas. *Current Opinion in Environmental Sustainability* 2: 172–177.

Jaramillo J, Chabi-Olaye A, Kamonjo C, *et al.* (2009) Thermal tolerance of the coffee berry borer *Hypothenemus hampei*: Predictions of climate change impact on a tropical insect pest. *Plos One* 4.

Jarema SI, Samson J, Mcgill BJ, and Humphries MM (2009) Variation in abundance across a species' range predicts climate change responses in the range interior will exceed those at the edge: A case study with North American beaver. *Global Change Biology* 15: 508–522.

Jarnevich CS and Stohlgren TJ (2009) Near term climate projections for invasive species distributions. *Biological Invasions* 11: 1373–1379.

Jonsson AM, Harding S, Barring L, and Ravn HP (2007) Impact of climate change on the population dynamics of *Ips typographus* in southern Sweden. *Agricultural and Forest Meteorology* 146: 70–81.

Kareiva P, Daily G, Ricketts T, Tallis H, and Polasky S (2011) *The Theory and Practice of Ecosystem Service Valuation in Conservation.* Oxford, UK: Oxford University Press.

Kareiva P, Enquist C, Johnson A, *et al.* (2008) Synthesis and Conclusions. In:Preliminary review of adaptation options for climate-sensitive ecosystems and resources. In: Julius SH and West JM (eds.), A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. Washington, DC: U.S. Environmental Protection Agency.

Kareiva P and Levin SA (2003) *The Importance of Species: Perspectives on Expendability and Triage.* Princeton, NJ: Princeton University Press.

Keane RE, Holsinger LM, Parsons RA, and Gray K (2008) Climate change effects on historical range and variability of two large landscapes in western Montana, USA. *Forest Ecology and Management* 254: 375–389.

Keith DA, Akcakaya HR, Thuiller W, *et al.* (2008) Predicting extinction risks under climate change: Coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters* 4: 560–563.

Kohm KA and Franklin JF (eds.) (1997) *Creating a Forestry for the 21st Century: The Science of Ecosystem Mangement.* Washington, DC: Island Press.

Kostyack J, Lawler JJ, Goble DD, Olden JD, and Scott JM (2011) Beyond reserves and corridors: Policy solutions to facilitate the movement of plants and animals in a changing climate. *Bioscience* 61: 713–719.

Krysanova V, Dickens C, Timmerman J, *et al.* (2010) Cross-comparison of climate change adaptation strategies across large river basins in Europe, Africa and Asia. *Water Resources Management* 24: 4121–4160.

La Sorte FA and Jetz W (2010) Avian distributions under climate change: Towards improved projections. *Journal of Experimental Biology* 213: 862–869.

Lafferty KD (2009) The ecology of climate change and infectious diseases. *Ecology* 90: 888–900.

Lawler JJ (2009) Climate change adaptation strategies for resource management and conservation planning. *Annals of the New York Academy of Sciences* 1162: 79–98.

Lawler JJ, Nelson E, Conte M, *et al.* (2011) Modeling the impacts of climate change on ecosystem services. In: Kareiva, P, Daily, G, Ricketts, T, Tallis, H & Polasky, S (eds.), *The Theory & Practice of Ecosystem Service Valuation in Conservation*, pp. 323–338. Oxford, UK: Oxford University Press.

Littell JS, Mckenzie D, Peterson DL, and Westerling AL (2009) Climate and wildfire area burned in western US. ecoprovinces, 1916–2003. *Ecological Applications* 19: 1003–1021.

Lobell DB and Asner GP (2003) Climate and management contributions to recent trends in U.S. agricultural yields. *Science* 299: 1032.

Lobell DB and Field CB (2007) Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters* 2: 14002–14007.

Luck J, Spackman M, Freeman A, *et al.* (2011) Climate change and diseases of food crops. *Plant Pathology* 60: 113–121.

Malevsky-Malevich SP, Molkentin EK, Nadyozhina ED, and Shklyarevich OB (2008) An assessment of potential change in wildfire activity in the Russian boreal forest zone induced by climate warming during the twenty-first century. *Climatic Change* 86: 463–474.

Malhi Y, Aragao L, Galbraith D, *et al.* (2009) Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences of the United States of America* 106: 20610–20615.

Manning AD, Gibbons P, and Lindenmayer DB (2009) Scattered trees: a complementary strategy for facilitating adaptive responses to climate change in modified landscapes? *Journal of Applied Ecology* 46: 915–919.

Marlon JR, Bartlein PJ, Walsh MK, *et al.* (2009) Wildfire responses to abrupt climate change in North America. *Proceedings of the National Academy of Sciences of the United States of America* 106: 2519–2524.

Mas-Coma S, Valero MA, and Bargues MD (2009) Climate change effects on trematodiases, with emphasis on zoonotic fascioliasis and schistosomiasis. *Veterinary Parasitology* 163: 264–280.

Matthews A, Spooner PG, Lunney D, Green K, and Klomp NI (2010) The influences of snow cover, vegetation and topography on the upper range limit of common wombats *Vombatus ursinus* in the subalpine zone, Australia. *Diversity and Distributions* 16: 277–287.

McRae BH, Schumaker NH, Mckane RB, *et al.* (2008) A multi-model framework for simulating wildlife population response to land-use and climate change. *Ecological Modelling* 219: 77–91.

Melles SJ, Fortin MJ, Lindsay K, and Badzinski D (2011) Expanding northward: influence of climate change, forest connectivity, and population processes on a threatened species' range shift. *Global Change Biology* 17: 17–31.

Memmott J, Carvell C, Pywell RF, and Craze PG (2010) The potential impact of global warming on the efficacy of field margins sown for the conservation of bumble-bees. *Philosophical Transactions of the Royal Society B – Biological Sciences* 365: 2071–2079.

Menzel A (2005) A 500 year pheno-climatological view on the 2003 heatwave in Europe assessed by grape harvest dates. *Meteorologische Zeitschrift* 14: 75–77.

Menzel A and Dose V (2005) Analysis of long-term time series of the beginning of flowering by Bayesian function estimation. *Meteorologische Zeitschrift* 14: 429–434.

Mertz O, Halsnaes K, Olesen JE, and Rasmussen K (2009) Adaptation to climate change in developing countries. *Environmental Management* 43: 743–752.

Metzger MJ, Schröter D, Leemans R, and Cramer W (2008) A spatially explicit and quantitative vulnerability assessment of ecosystem service change in Europe. *Regional Environmental Change* 8: 91–107.

Midgley GF, Hughes GO, Thuiller W, and Rebelo AG (2006) Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. *Diversity & Distributions* 12: 555–562.

Millar CI, Stephenson NL, and Stephens SL (2007) Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* 17: 2145–2151.

Miller-Rushing AJ, Hoye TT, Inouye DW, and Post E (2010) The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society B – Biological Sciences* 365: 3177–3186.

Mitchell NJ, Allendorf FW, Keall SN, Daugherty CH, and Nelson NJ (2010) Demographic effects of temperature-dependent sex determination: Will tuatara survive global warming? *Global Change Biology* 16: 60–72.

Moller AP (2010) Host–parasite interactions and vectors in the barn swallow in relation to climate change. *Global Change Biology* 16: 1158–1170.

Molnar PK, Derocher AE, Thiemann GW, and Lewis MA (2010) Predicting survival, reproduction and abundance of polar bears under climate change. *Biological Conservation* 143: 1612–1622.

Morin CW and Comrie AC (2010) Modeled response of the West Nile virus vector *Culex quinquefasciatus* to changing climate using the dynamic mosquito simulation model. *International Journal of Biometeorology* 54: 517–529.

Morin X and Lechowicz MJ (2008) Contemporary perspectives on the niche that can improve models of species range shifts under climate change. *Biology Letters* 4: 573–576.

Morin X and Thuiller W (2009) Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology* 90: 1301–1313.

Morin X, Viner D, and Chuine I (2008) Tree species range shifts at a continental scale: New predictive insights from a process-based model. *Journal of Ecology* 96: 784–794.

Mosblech N aS, Bush MB, and Van Woesik R (2011) On metapopulations and microrefugia: Palaeoecological insights. *Journal of Biogeography* 38: 419–429.

Murphy D and Pavlik B, Tahoe Yellow Cress Technical Advisory Group (2002) *Conservation Strategy for Tahoe Yellow Cress (Rorippa subumbellata).* Zephyr Cove, NV: Tahoe Regional Planning Agrency.

Neil KL, Landrum L, and Wu JG (2010) Effects of urbanization on flowering phenology in the metropolitan phoenix region of USA: Findings from herbarium records. *Journal of Arid Environments* 74: 440–444.

Nitschke CR and Innes JL (2008) Climatic change and fire potential in South-Central British Columbia, Canada. *Global Change Biology* 14: 841–855.

Nobis MP, Jaeger JAG, and Zimmermann NE (2009) Neophyte species richness at the landscape scale under urban sprawl and climate warming. *Diversity and Distributions* 15: 928–939.

Ogden NH, Bigras-Poulin M, Hanincova K, *et al.* (2008a) Projected effects of climate change on tick phenology and fitness of pathogens transmitted by the North American tick *Ixodes scapularis*. *Journal of Theoretical Biology* 254: 621–632.

Ogden NH, St-Onge L, Barker IK, *et al.* (2008b) Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *International Journal of Health Geographics* 7.
Olesen JE, Trnka M, Kersebaum KC, *et al.* (2011) Impacts and adaptation of European crop production systems to climate change. *European Journal of Agronomy* 34: 96–112.
Olwoch JM, Reyers B, and Van Jaarsveld AS (2009) Host–parasite distribution patterns under simulated climate: implications for tick-borne diseases. *International Journal of Climatology* 29: 993–1000.
Ooi MKJ, Auld TD, and Denham AJ (2009) Climate change and bet-hedging: Interactions between increased soil temperatures and seed bank persistence. *Global Change Biology* 15: 2375–2386.
Ozgul A, Childs DZ, Oli MK, *et al.* (2010) Coupled dynamics of body mass and population growth in response to environmental change. *Nature* 466: 482-U5.
Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics* 37: 637–669.
Parmesan C, Ryrholm N, Stefanescu C, *et al.* (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* 399: 579–583.
Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.
Patz J, Gibbs H, Foley J, Rogers J, and Smith K (2007) Climate change and global health: Quantifying a growing ethical crisis. *EcoHealth* 4: 397–405.
Patz JA, Olson SH, Uejio CK, and Gibbs HK (2008) Disease emergence from global climate and land use change. *Medical Clinics of North America* 92: 1473–1491.
Pearson RG and Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12: 361–371.
Peters RL and Darling JDS (1985) The greenhouse effect and nature reserves: Global warming would diminish biological diversity by causing extinctions among reserve species. *BioScience* 35: 707–717.
Phillips BL, Kelehear C, Pizzatto L, *et al.* (2010) Parasites and pathogens lag behind their host during periods of host range advance. *Ecology* 91: 872–881.
Podur J and Wotton M (2010) Will climate change overwhelm fire management capacity? *Ecological Modelling* 221: 1301–1309.
Polley L and Thompson RCA (2009) Parasite zoonoses and climate change: Molecular tools for tracking shifting boundaries. *Trends in Parasitology* 25: 285–291.
Ponti L, Cossu QA, and Gutierrez AP (2009) Climate warming effects on the Olea europaea-Bactrocera oleae system in Mediterranean islands: Sardinia as an example. *Global Change Biology* 15: 2874–2884.
Preston K, Rotenberry JT, Redak RA, and Allen MF (2008) Habitat shifts of endangered species under altered climate conditions: Importance of biotic interactions. *Global Change Biology* 14: 2501–2515.
Pyke CR and Fischer DT (2005) Selection of bioclimatically representative biological reserve systems under climate change. *Biological Conservation* 121: 429–441.
Reid CE and Gamble JL (2009) Aeroallergens, allergic disease, and climate change: Impacts and adaptation. *EcoHealth* 6: 458–470.
Reyer C, Guericke M, and Ibisch PL (2009) Climate change mitigation via afforestation, reforestation and deforestation avoidance: and what about adaptation to environmental change? *New Forests* 38: 15–34.
Ricciardi A and Simberloff D (2009) Assisted colonization is not a viable conservation strategy. *Trends in Ecology & Evolution* 24: 248–253.
Root TL, Price JT, Hall KR, *et al.* (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421: 57–60.
Rose NA and Burton PJ (2009) Using bioclimatic envelopes to identify temporal corridors in support of conservation planning in a changing climate. *Forest Ecology and Management* 258: S64–S74.
Rosenthal J (2009) Climate change and the geographic distribution of infectious diseases. *EcoHealth* 6: 489–495.
Sax DF, Smith KF, and Thompson AR (2009) Managed relocation: A nuanced evaluation is needed. *Trends in Ecology & Evolution* 24: 472–473.
Saxon E, Baker B, Hargrove W, Hoffman F, and Zganjar C (2005) Mapping environments at risk under different global climate change scenarios. *Ecology Letters* 8: 53–60.
Schindler DE, Hilborn R, Chasco B, *et al.* (2010) Population diversity and the portfolio effect in an exploited species. *Nature* 465: 609-U102.
Schröter D, Cramer W, Leemans R, *et al.* (2005) Ecosystem service supply and vulnerability to global change in Europe. *Science* 310: 1333–1337.
Schweiger O, Biesmeijer JC, Bommarco R, *et al.* (2010) Multiple stressors on biotic interactions: How climate change and alien species interact to affect pollination. *Biological Reviews* 85: 777–795.
Schweiger O, Settele J, Kudrna O, Klotz S, and Kuhn I (2008) Climate change can cause spatial mismatch of trophically interacting species. *Ecology* 89: 3472–3479.
Shoo LP, Storlie C, Vanderwal J, Little J, and Williams SE (2011) Targeted protection and restoration to conserve tropical biodiversity in a warming world. *Global Change Biology* 17: 186–193.
Snall T, Benestad RE, and Stenseth NC (2009) Expected future plague levels in a wildlife host under different scenarios of climate change. *Global Change Biology* 15: 500–507.
Sutherst RW, Maywald GF, and Bourne AS (2007) Including species interactions in risk assessments for global change. *Global Change Biology* 13: 1843–1859.
Thackeray SJ, Sparks TH, Frederiksen M, *et al.* (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology* 16: 3304–3313.
Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology* 10: 2020–2027.
Tobin PC, Nagarkatti S, Loeb G, and Saunders MC (2008) Historical and projected interactions between climate change and insect voltinism in a multivoltine species. *Global Change Biology* 14: 951–957.
Tompkins EL, Adger WN, Boyd E, *et al.* (2010) Observed adaptation to climate change: UK evidence of transition to a well-adapting society. *Global Environmental Change-Human and Policy Dimensions* 20: 627–635.
Traill LW, Lim MLM, Sodhi NS, and Bradshaw CJA (2010) Mechanisms driving change: Altered species interactions and ecosystem function through global warming. *Journal of Animal Ecology* 79: 937–947.
Tylianakis JM, Didham RK, Bascompte J, and Wardle DA (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters* 11: 1351–1363.
Valtonen A, Ayres MP, Roininen H, Poyry J, and Leinonen R (2011) Environmental controls on the phenology of moths: Predicting plasticity and constraint under climate change. *Oecologia* 165: 237–248.
Van Mantgem PJ, Stephenson NL, Byrne JC, *et al.* (2009) Widespread increase of tree mortality rates in the western United States. *Science* 323: 521–524.
Verlinden M and Nijs I (2010) Alien plant species favoured over congeneric natives under experimental climate warming in temperate Belgian climate. *Biological Invasions* 12: 2777–2787.
Vörösmarty CJ, Green P, Salisbury J, and Lammers RB (2000) Global water resources: Vulnerability from climate change and population growth. *Science* 289: 284–288.
Vos CC, Berry P, Opdam P, *et al.* (2008) Adapting landscapes to climate change: Examples of climate-proof ecosystem networks and priority adaptation zones. *Journal of Applied Ecology* 45: 1722–1731.
Walters CJ and Hilborn R (1978) Ecological optimization and adaptive management. *Annual Review of Ecology and Systematics* 9: 157–188.
Walther G-R, Post E, Convey P, *et al.* (2002) Ecological responses to recent climate change. *Nature* 416: 389–395.
Walther GR (2010) Community and ecosystem responses to recent climate change. *Philosophical Transactions of the Royal Society B – Biological Sciences* 365: 2019–2024.
Welch D (2005) What should protected areas managers do in the face of climate change? *The George Wright Forum* 22: 75–93.
Wheeler S (2008) State and municipal climate change plans: The first generation. *Journal of the American Planning Association* 74: 481–496.
Wilby RL and Perry GLW (2006) Climate change, biodiversity and the urban environment: A critical review based on London, UK. *Progress in Physical Geography* 30: 73–98.
Wilby RL, Wigley TML, Conway D, *et al.* (1998) Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research* 34: 2995–3008.
Williams PH, Hannah L, Andelman SJ, *et al.* (2005) Planning for climate change: identifying minimum-dispersal corridors for the Cape Proteaceae. *Conservation Biology* 19: 1063–1074.
Willis CG, Ruhfel BR, Primack RB, *et al.* (2010) Favorable climate change response explains non-native species' success in Thoreau's Woods. *Plos One* 5.
Yates CJ, Elith J, Latimer AM, *et al.* (2010) Projecting climate change impacts on species distributions in megadiverse South African Cape and Southwest Australian Floristic Regions: Opportunities and challenges. *Austral Ecology* 35: 374–391.
Yu HY, Luedeling E, and Xu JC (2010) Winter and spring warming result in delayed spring phenology on the Tibetan Plateau. *Proceedings of the National Academy of Sciences of the United States of America* 107: 22151–22156.
Ziervogel G and Zermoglio F (2009) Climate change scenarios and the development of adaptation strategies in Africa: Challenges and opportunities. *Climate Research* 40: 133–146.

Climate, Effects of

GF Midgley, South African National Biodiversity Institute, Cape Town, South Africa
FI Woodward, University of Sheffield, Sheffield, UK

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by F.I. Woodward, volume 1, pp 727–740, © 2001, Elsevier Inc.

Glossary

Disturbance regime The nature, frequency and intensity of events or agents that cause a temporary or permanent loss of biomass.
Fynbos Shrub and heath vegetation type that is characteristic of the Cape Floral Kingdom in South Africa.
Herbaceous species Plant species that do not form woody tissue, and are often much shorter lived than woody species such as trees and shrubs.
Net primary production Plant photosynthesis less plant respiration.
Species diversity Species richness adjusted to account for abundance.
Species migration For plants this refers to the directional change in a species geographic range over time, as opposed to the seasonal bi-directional movement of individual organisms such as migratory birds.
Species range The geographic area within which a species can be found.
Species richness The number of species recorded for a particular area, with each species given an equal weighting.

Observations

One of the most general apparent relationships between climate and the diversity of species is reflected in increases in species richness from each pole toward the equator, a pattern that has very few exceptions (Hillebrand, 2004). The general term diversity used here indicates a measurement of the number of species which exist in a particular area, a measure more correctly termed species richness. The measurement of species richness quantifies the number of species recorded for a particular area, with each species given an equal weighting. Strictly speaking, species diversity is a measure that accounts both for species richness and different species' abundances (termed equitability), but the terms are often used interchangeably. Because data on species abundances globally are limited and, at best, may only include species lists, with some notional area of observation, this general analysis will substitute species richness for species diversity. The latitudinal gradient identifies a potentially important role for climate in limiting diversity, mainly through climate control of energy availability (temperature and radiation), but several other factors are also seen as important in explaining the pattern, including the role of spatial area, evolutionary time, and ecological factors (species interactions).

Species to Area Relationships

The horizontal surface area within which species are counted plays a critical role in quantifying species richness. This relationship has been termed one of the few fundamental laws of ecology (Lawton, 1999). Rosenzweig (1995) provides an extensive account and examples of the relationships between species richness and area. One example is presented to illustrate the typical nature of the relationship for the flora of the British Isles (**Figure 1**). **Figure 1(a)** shows the relationship for untransformed scales, with a particularly notable and rapid increase in species richness following small initial increments of area. **Figure 1(b)** presents the same data for log-transformed scales, in which the relationship is linear. **Figure 1(b)** indicates that the rapid increase in richness to about 1000 species occurs as the scale of observation increases from 0.1 ha to 1 million ha (10,000 km^2 or 100 km $\times$ 100 km).

All spatial scales contain data of ecological interest; however, in the context of the climatic control of diversity, the small areas are of less direct relevance. For example, the smallest scale in **Figure 1** of 0.1 ha (approximately 32 m $\times$ 32 m) is too small to include many individual plants such as mature trees, and this would also apply to data for animal species that are often locally rare (e.g., birds and mammals). This is therefore a sampling problem. The problem lies in deciding when this species to area relationship might indicate a primary control by climate, over and above any problems of sampling. It can also be envisaged that if an area under study has high landscape diversity, or habitat heterogeneity (e.g., it is a flat plain vs. a mountainous area), then this will also influence the species to area relationship, as has been reported (Cowling *et al.*, 1994; Báldi, 2008). Such environmental features are incorporated into general species lists and cannot easily be dissected out. Nonetheless, a comparison of species richness with scale for the British Isles and the fynbos of South Africa (**Figure 2**) indicates very large differences in diversity, but the untransformed scale (**Figure 2(a)**) shows that the differences between the two floras decrease much more slowly above areas of approximately 2×10^7 ha (2×10^5 km^2 or 450 km $\times$ 450 km). It is well known that fynbos diversity has been strongly enhanced by landscape heterogeneity (Linder *et al.*, 2010), and therefore the decreasing divergence in trend between these two data sets above 450 km $\times$ 450 km indicates a useful order of scalar magnitude for comparing floras in the context of examining climatic control.

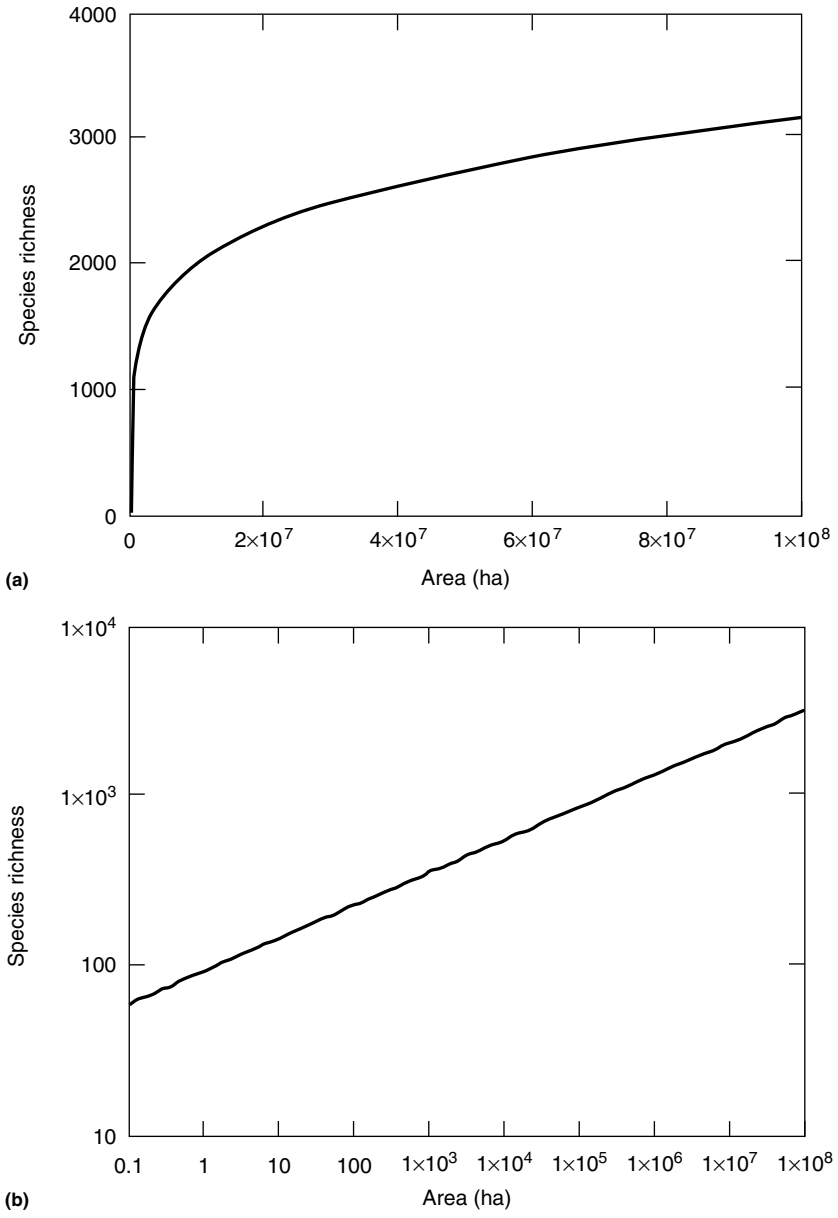


Figure 1 Species richness of the British Isles: (a) untransformed axes and (b) log-transformed axes.

Time Dependency

The species composition of an area endures changes with time, through mortality, disturbance, and migration. Disturbance of a moderate frequency can maximize species richness, the so-called "intermediate disturbance hypothesis," but analysis shows that this pattern is far from universal (Mackey and Currie, 2001). Large-scale extreme disturbances produce landscapes in various stages of recovery, accompanied by increasing diversity. An extreme disturbance event was the 1883 volcanic eruption which completely denuded the flora of the Krakatau Islands. A recent enumeration of higher plant species (**Figure 3**; Whittaker *et al.*, 1990) showed a monotonic increase in the species complement during the past 100 years. It is clearly not possible to assess the final species richness of these islands until the species complement reaches some asymptote. This is an example of primary succession (where the biophysical environment establishes *de novo*) and is extreme in the context of contemporary time scales, although smaller-scale recoveries from natural disturbance and larger-scale recoveries from human disturbance are current widespread features. Human disturbance is generally considered to deplete species richness and push natural ecosystems to early stages of secondary succession (where the biophysical environment, e.g., soil structure, remains intact); however, over the large spatial scales considered for this article, it is assumed that species richness has not yet been markedly depleted by human activities.

The last ice age was the most recent natural global-scale disturbance to the biosphere. At that time, glaciers spread over

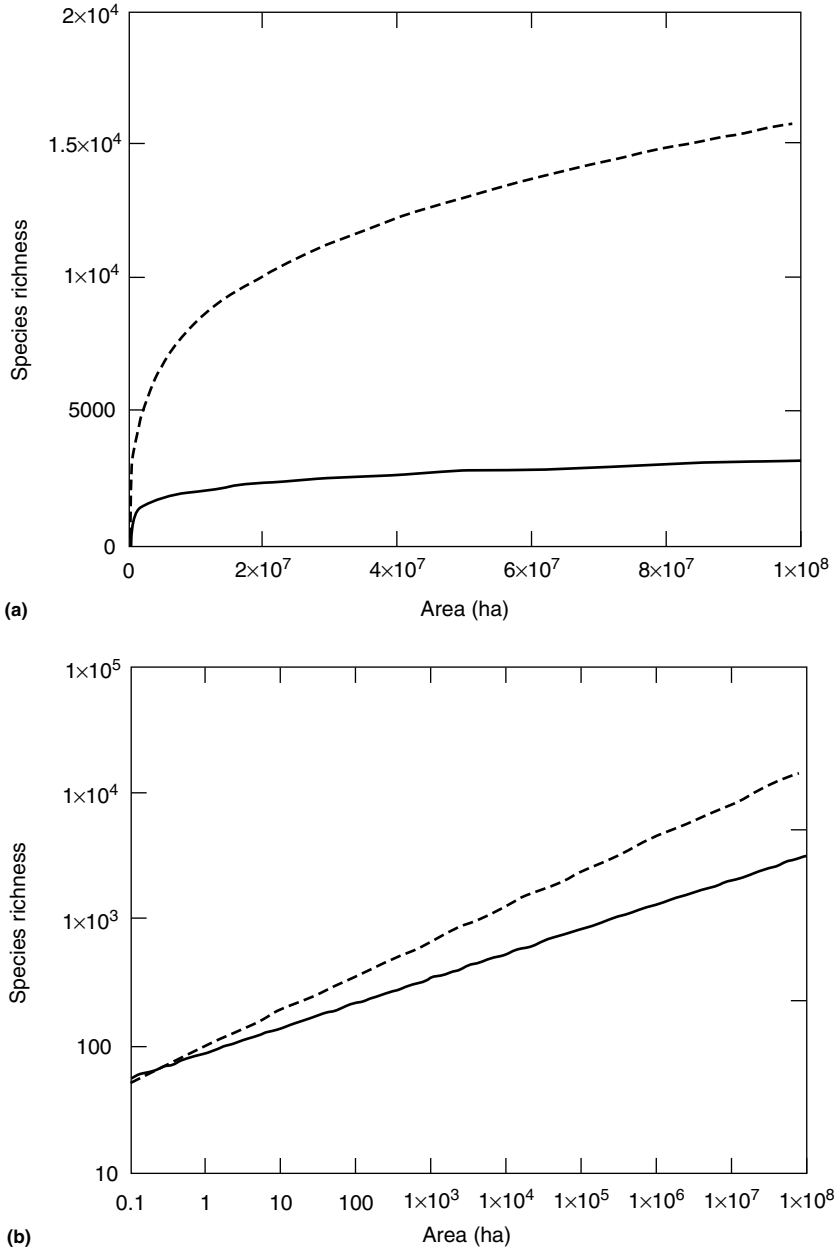


Figure 2 Species richness of the British Isles compared with the fynbos of South Africa: (a) untransformed axes and (b) log-transformed axes.

the large areas of North America, Europe, and Russia and clearly removed all species in their paths. The cooler glacial climate, in addition to the spread of glaciers, led to the equatorial migration of most species, especially in the Northern Hemisphere. However, in Europe, for example, the Alps prevented significant migration further toward the equator, with significant extinctions at this barrier. The climatic amelioration to the current interglacial climate reversed the process, which occurred during the ice age. However, reduced richness follows as a consequence of the irreversible species extinctions. The degree of this loss of richness can be addressed in part not only from paleoecological information but also from comparative subcontinental biogeographic and phylogenetic studies (Wiens and Donoghue, 2004). An example of the latter is that of eastern Asia and eastern North America, where eastern Asia shows regional richness approximately 1.3 times higher than eastern North America after the effects of climate are accounted for (Ricklefs *et al.*, 2004), and may be due to greater latitudinal connectivity between temperate and tropical floras in eastern Asia, which lowered regional extinction rates under glacial–interglacial climatic changes. These long-term effects, and the shorter-term effects of disturbance, are incorporated in regional assessments of species richness. Despite this, the effect of contemporary climate on diversity remains relatively strongly discernible within these regions, with richness positively related to winter temperature.

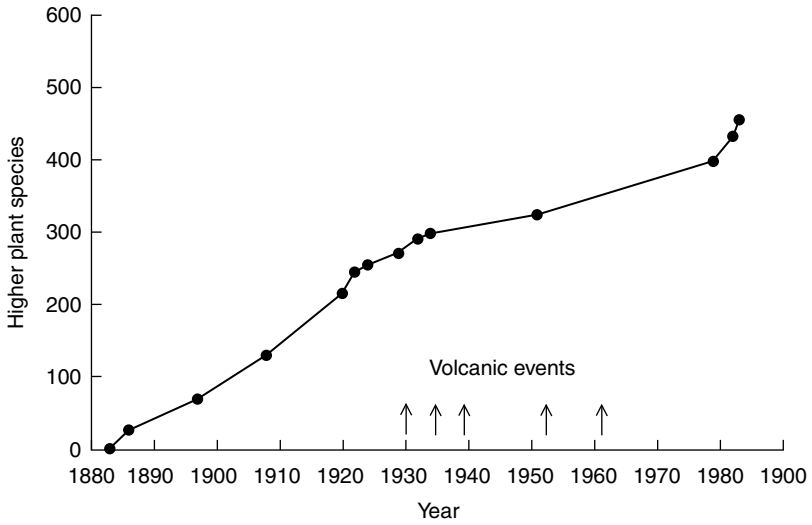


Figure 3 Measurements of species richness on the Krakatau Islands, Indonesia, since the major volcanic eruption of 1883. Subsequent and lesser volcanic events are also indicated by arrows.

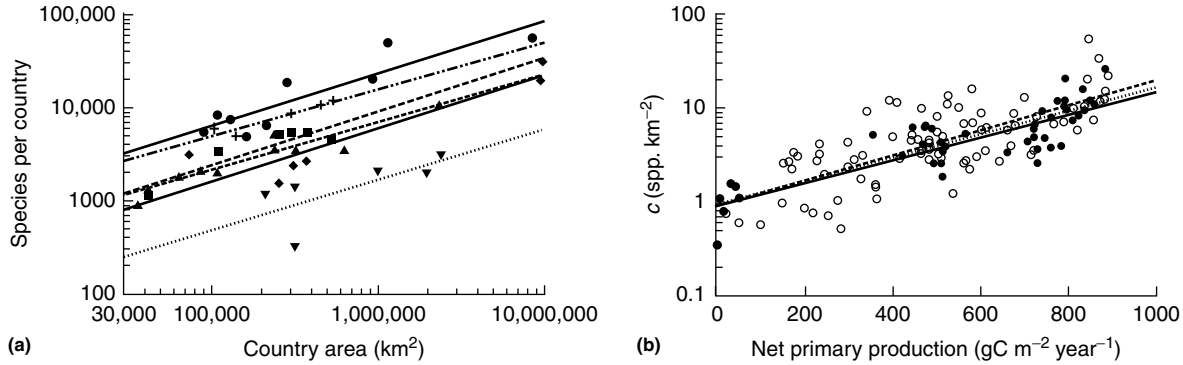


Figure 4 Species diversity in relation to area and productivity. (a) Species area curves for six groups of countries (42 countries in total) selected for similarity of temperature and NPP, using data for country area and mean country NPP. (b) Relationship between c (plant species per square kilometer) and NPP (gC m^{-2} yr^{-1}) for 42 countries selected for similarity of NPP, mean annual temperature, and geographical location; for an additional 86 countries (- - -); and for all 128 countries combined (....), where $\log_{10} c = 0.01233 \times \text{NPP(gC m}^{-2}\text{ yr}^{-1}) - 0.01475$, $r^2 = 0.58$. Reproduced from Kelly CK and Woodward FI (2008) Responses of global plant diversity capacity to changes in carbon dioxide concentration and climate. *Ecology Letters* 11: 1229–1237, with permission from Wiley.

Observations by Area

The impacts of area on species enumeration and on the relationship between climate and species richness are well illustrated (**Figure 4**) using data synthesized by Kelly and Woodward (2008). They showed that plant species richness per square kilometer is significantly positively related to net primary productivity (NPP), a measure that reflects climatic drivers of plant production (temperature and water availability). Previous work also showed that per area absolute plant species richness is lower in colder environments. Such observations conform with observations and modeling by Woodward (1987), which demonstrate a decreasing capacity for species survival with decreasing temperatures.

Global-Scale Observations

The global-scale distribution of species richness is increasingly better documented, and a comprehensive atlas, for example, is available for download at: http://quin.unep-wcmc.org/resources/publications/biodiversityatlas/presspack/index.htm

Species Richness and Latitude

The first cut at assessing species richness has traditionally been conducted by latitude. This spatial variable includes many environmental features, particularly solar radiation receipt, day length, and temperature, but it is a useful variable for assessing broad-scale controls on vegetation. **Figure 5** indicates a spread of richness with latitude. In the northern latitudes, the maximum observed richness decreases with increasing latitude. A similar relationship can be defined for the Southern Hemisphere, except that there are significant exceptions. These cases include the southeast coastal region of Brazil, extending from Porto Allegre (**Figure 5**, Brazil 3) in the south through Rio de Janeiro (**Figure 5**, Brazil 1) to Vitoria (**Figure 5**, Brazil 2) in the north. Additional sites of high diversity have been sampled for South Africa (The Cape) and

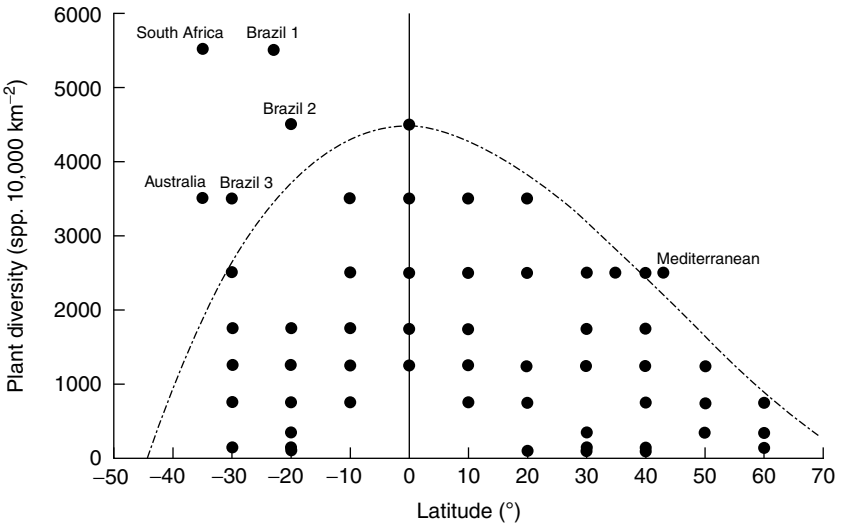


Figure 5 Changes in species richness with latitude. Details of the identified sites are discussed in the text. Data from Barthlott W, Lauer W, and Placke A (1996) Global distribution of species diversity in vascular plants: Towards a world map of phytodiversity. *Erdkunde* 50: 317–327.

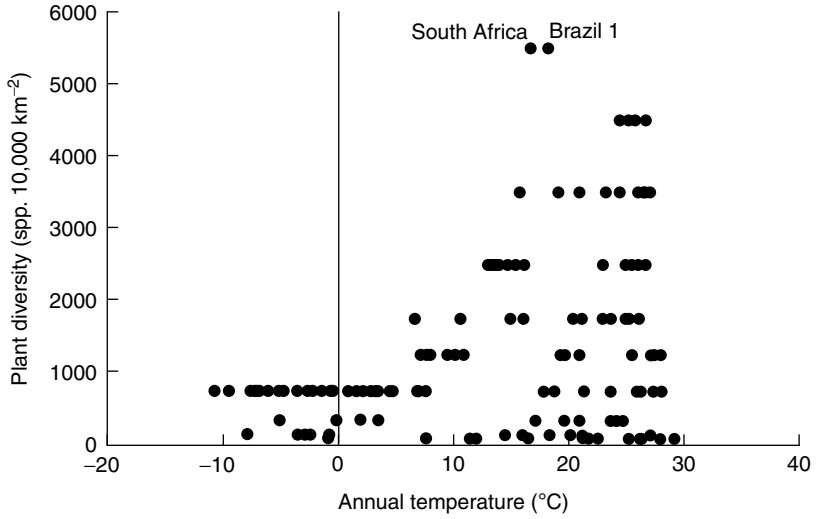


Figure 6 Changes in species richness with annual mean temperature.

Australia (Perth). The sampling scheme for global diversity does not cover all areas of high diversity, but these five sites in the Southern Hemisphere and, to a lesser extent, the Mediterranean region in the Northern Hemisphere clearly indicate that areas of unusually high diversity do exist. These areas are defined as unusual because they seem outside the typical latitudinally dependent maximum extents of diversity. Using the increasingly resolved global data set for plant species richness, it has been found that potential evapotranspiration (a measure of energy availability), the number of wet days per year, and measurements of topographical and habitat heterogeneity are powerful predictors of species richness (Kreft and Jetz, 2007). This analysis explains a remarkably high 70% of the global variation in plant species richness, and reproduces the latitudinal gradient in richness. Only one major exception is found – that of the Cape Floristic region in South Africa, which has about double the expected number of plant species! Like plant species richness, animal species richness shows a strong latitudinal trend that could indicate co-dependence on plant species richness, direct mechanisms driven by climate, or a combination of these (Midgley *et al.*, 2010).

Species Richness and Climate

Temperature

The observations in sections Observations by Area and Species Richness and Latitude suggest that cooler and drier climates are less diverse. This same result is observed at the global scale (**Figure 6**), when mean annual climate (for the 1970s and sampled at a scale of 10,000 km^2) is compared with a map of standardized species diversity (also adjusted to a 10,000 km^2 scale) (Barthlott *et al.*, 1996). At mean temperatures greater

than approximately 5 °C, species richness both increases markedly and becomes significantly more variable between individual sample points. The high-temperature, low-diversity sites are typically deserts or very arid, where it is clear that water supply is the critically limiting factor for plant survival. The areas of unusually high diversity in South Africa and Brazil are clear extremes in the temperature relationship, but the other sites shown in **Figure 5** are not distinct.

Precipitation

The relationship between richness and temperature is not simple and is also influenced by other climatic features, particularly precipitation. In contrast, 42% of the variance in richness (**Figure 7**) is accounted for by a linear correlation between richness and annual precipitation. This association defines the obvious observation that diversity increases toward the tropics and is highest when warmth and high precipitation coincide. Four of the five recognized areas of high diversity remain as outliers from the remainder of the data set.

Net Primary Productivity

Precipitation is often taken as a surrogate for vegetation productivity in order to investigate any relationships between productivity, usually NPP and diversity or richness. Although precipitation exerts a large effect on NPP, this is not the only influence; other climatic features such as solar radiation receipt, day length, growing season temperature, and annual minimum temperature can exert significant effects on NPP (Woodward, 1987). Simulation models of NPP typically incorporate the effects of all these climatic variables to provide an integrated effect, seen as simulated NPP (Woodward *et al.*, 1995).

In diversity studies, productivity is used and assessed in two different ways. For small field plot experiments, productivity is measured directly as the accumulation of biomass and is generally viewed as a consequence of diversity. On larger scales, intensive measurements of NPP are impossible, and in this case productivity is generally simulated by some type of process model; it is more appropriate to consider diversity as a consequence of simulated NPP because this simulated value accounts for all variations in climate but is generally independent of any aspects of diversity.

In this case, a typical process model (Woodward *et al.*, 1995) has been used. Annual NPP is the net fixation of carbon into plant biomass, after plant respiration is decremented from annual photosynthetic fixation. In ecological terms, NPP is a measure of the rate of production of biomass (usually over an annual interval) and which is available for consumers and for creating new plant structure. Simulation models can often provide more accurate estimates of annual production, particularly in multiseasonal environments, which would require many harvests for accurate assessments of NPP in the field.

Simulated NPP provides a tighter fit than annual precipitation to the observed variations in species richness at the global scale (**Figure 8**). However, three of the high-diversity sites are notable outliers from the line of best fit. The site in South Africa is the most extreme, as previously recognized, with a very high richness associated with a low NPP. Much of this area of vegetation is known as fynbos, a heathland on very impoverished soils. A further analysis of species richness in this area, and for the whole of southern Africa northwards to a latitude of 15° S, has been completed by O'Brien *et al.* (1998). In this case, only the richness of woody species (native trees and shrubs with heights of greater than 1.5 m) was considered. Woody species richness is very strongly correlated with NPP (**Figure 9**), and the high-diversity site of **Figure 8** (arrow in **Figure 9**) is not now distinguishable from the general trend. The high richness at this site is clearly due to herbaceous species and subshrubs, and explanations for their high diversity must be sought in factors other than current climate, such as disturbance frequencies and types, the low nutrient status of the soils, and factors related to biogeographic history, particularly rates of speciation and probabilities of species' migrations.

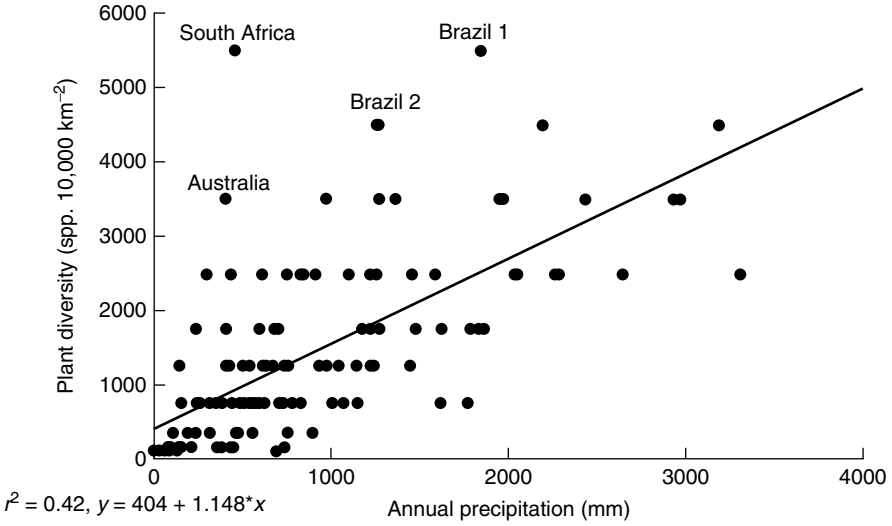


Figure 7 Changes in species richness with annual precipitation.

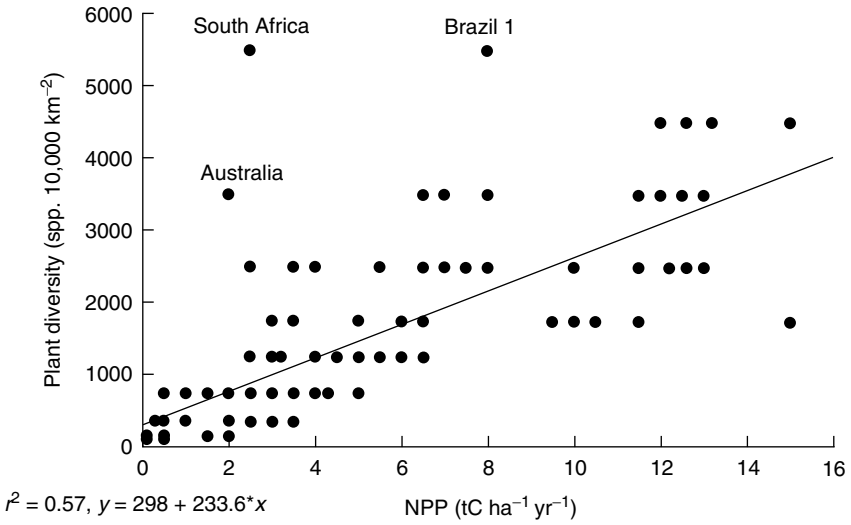


Figure 8 Changes in species richness with annual NPP.

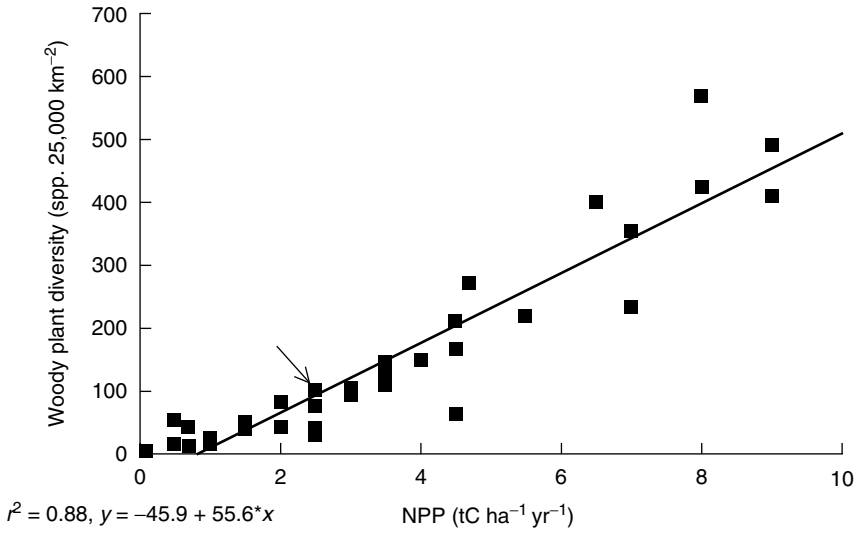


Figure 9 Changes in woody species richness of southern Africa with annual NPP. The arrow indicates the NPP and woody plant richness for the site in South Africa indicated in **Figures 5–8**. Data from O'Brien EM, Whittaker RJ, and Field R (1998) Climate and woody plant diversity in southern Africa: Relationships at species, genus and family levels. *Ecography* 21: 495–509.

Theories

The strong latitudinal gradient in species richness has long been recognized and has stimulated a wide range of theories to explain the observations. At the same time, relevant new data sets are continually being produced and analyzed in new ways. These data have been used to derive empirical or observational hypotheses about the controls of richness, mainly by seeking statistical support for relationships between diversity and multiple explanatory variables, with the recognition that correlation does not imply causation. In principle, an empirically derived hypothesis is supported if it can be tested using first principles. However, nature does not yield its secrets so easily, and ambiguity seems to be the rule in that it is often difficult to differentiate between different theories and hypotheses when the differences in their consequences are often by degree rather than absolute. It is also clear that several explanations of what determines diversity patterns may be valid and not mutually exclusive, but rather interact to yield the global, regional, and local patterns observed. This article presents current theories and hypotheses which address the relationships between climate and species richness.

Narrowing the Questions Using Observations

The defined area for observing species richness exerts a major impact on the quantity of richness. A less obvious, but also

important, consideration is that the context within which areas exist requires consideration. For example, **Figure 10** indicates that islands show a greater increase in richness with area than do areas of the same size but nested within a region with no dispersal and migratory barriers. This observation highlights that the species richness of any area is the net result of species immigration, extinction, and evolution. For an island, the rate of species immigration will be less than that for an area embedded in contiguous land surface, and as a consequence the island richness will be less because local species extinctions are not readily matched by migration. Such insights have led to the theory of "island biogeography." With increasing island size, the potential for the arrival, establishment, and persistence of migrants increases, and so extinctions will be fewer and richness will increase to values similar to that of the contiguous land sites. This has also been elegantly demonstrated for islands artificially created through the damming of the Vaal River in South Africa (Dean and Bond, 1990). An important contemporary corollary to the change in diversity with area is that any process, such as human land clearance, which effectively creates isolated fragments will inevitably lead to a loss of diversity.

The response shown in **Figure 10** also scales to larger areas (**Figure 11**). In this case, the species-to-area relationship for the British Isles is much flatter than that for separate provinces and continents, from New Zealand, at the smallest, through Australia and Africa to the world. The flattest curve for the British Isles will include some component due to the net balance of extinctions and immigrations, plus a capacity for species to migrate throughout the British Isles, thus evening

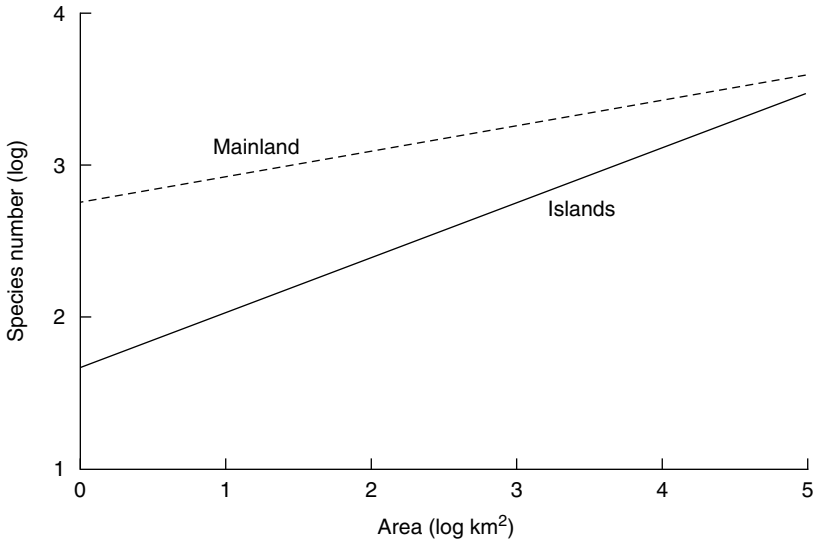


Figure 10 Changes in species richness with area for islands and mainland California.

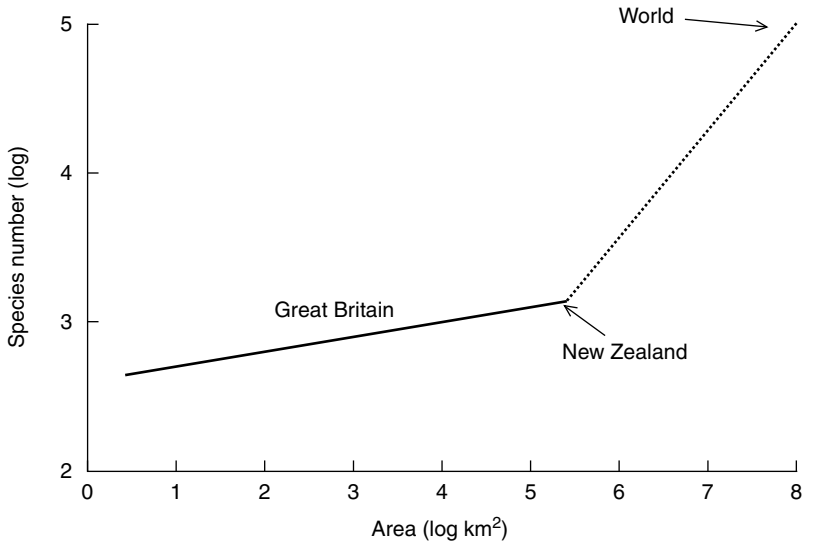


Figure 11 Changes in species richness with area within the British Isles (solid line) and between different floras (dashed line).

out, to some degree, the scale dependence of species richness. Different provinces and continents may be more or less isolated for immigration and also possess species complements determined by local effects such as habitat heterogeneity, historical effects such as past ice ages, and even longer-term speciation events. Climate will play and will have played a part in conjunction with these influences, therefore, in the contemporary species richness of these large areas. The climatic range of the British Isles – for example, in temperature and precipitation – is small compared with that for a continent, but the range of soil conditions is wide, from fertile alluvial soils to very low-nutrient bogs and chalk downs. Therefore, the species area curve (**Figure 11**) for the British Isles also displays the influence of habitat diversity, particularly soil nutrient status on species richness. In this case the effect is small, and suggests, therefore, that latitudinal gradients in soil nutrient conditions may not play a major role in determining the gradients of richness seen, for example, with latitude (**Figure 5**).

Observations by Rapoport (1982) introduced an alternative twist to the understanding of the latitudinal pattern of species richness. Rapoport noted that, on average, the geographical range of species increased with latitude, and surmised that there are fewer species at high latitudes and they are on average more widely dispersed. Subsequent analyses (Gaston, 1999) have indicated that this relationship is neither universal nor simple. However, there is some consistency in the pattern from the tropics to high latitudes, particularly in the continental landmasses of the Northern Hemisphere, although the pattern is far weaker at the global level, possibly reflecting different biogeographical processes in the Northern and Southern Hemispheres (Ruggiero and Werenkraut, 2007).

Theories and Hypotheses

Theories and hypotheses for explaining the documented decline in diversity with latitude are based on one or more of the following fundamental controls:

- Competition for necessary resources.
- Gap creation in vegetation by disturbance.
- Energy availability.
- Climatic extremes.
- History.

Competition

The competitive exclusion of species, which require the same resource, has been a theoretical underpinning to much of ecology, not just the climate control of diversity. Originating from simple experiments, it has proved easy to demonstrate the importance of competition in species interactions in response to changes in climate and in very small-scale experiments. It has not been so easy to demonstrate the importance of competition in defining the distributional limits of species on a larger field scale. Often, it seems that environmental heterogeneity in time and space can allow coexistence to occur, apparently circumventing direct competition for resources (Huston, 1979). However, competition is still considered an important mechanism in controlling survival and therefore diversity, but it does not appear to be easy to demonstrate competition in action.

It does seem that if competition is to play a major role in the control of species distributions, then it must do so as a secondary action to the primary action of climate. For example, consider the distribution of a boreal forest species in the cold of the high latitudes and a tropical species from the warmth of low latitudes. If competition plays a role in these distributions, then it must favor boreal species at high latitudes and tropical species at low latitudes (i.e., competition is latitude dependent). Competition can only be latitude dependent through its impact on species' interactions, and this would either be through a climatic dependency of species' contemporaries or through historical responses. Therefore, it seems that competition is an integral and probably secondary part of a climatic control of diversity.

Disturbance

Creating gaps in vegetation or clearing vegetation above ground by disturbance appears to provide a mechanism for avoiding direct competition and for providing bare ground and the opportunity for new species immigration. If some species are long lived and can survive disturbance events, then they can co-occur with shorter-lived species, which depend on disturbance, and therefore diversity can be maintained. If the rate of disturbance changed systematically with latitude, then this could provide an explanation for the decline in diversity with latitude. However, the major global disturbance agent, wildfire, has been shown to peak in frequency at midlatitudes (Midgley *et al.*, 2010), though this affects ecosystem structure more than it does biodiversity. It has been suggested that midlatitude grassland fires reduce the potential global area of dominance of trees by a considerable fraction (Bond *et al.*, 2005), but ancient fire-maintained grasslands have accumulated significant levels of biodiversity in low-growing plant growth forms such as bulbous plants that tolerate fire. Nonetheless, there is some evidence that at local and regional spatial scales, an intermediate level of disturbance supports maximum diversity (Huston, 1979).

Energy Availability

Hutchinson (1959) suggested that the availability of energy may constrain species diversity. Studies have since shown strong positive correlations between diversity and evapotranspiration, which is often taken as a surrogate for both energy and NPP. **Figure 8** shows a clear relationship between NPP and species richness, but in this case climates with high energy inputs from solar radiation but low precipitation will lead to low NPP. O'Brien *et al.* (1998) describe the same response where water availability tempers a direct response of diversity to energy. Therefore, the relationship between diversity and energy is not simple, whereas the relationship between NPP and richness is in general quite linear (**Figure 4**). However, the mechanistic nature of the relationship is quite unclear in either case.

In some cases, particularly in small-scale experiments with variable resource supply, NPP and diversity are often negatively correlated, with just a small number of rank-growing species dominating energy capture and production. Such a

response is opposite to that observed at the larger, global scale (**Figure 8**) and, based on **Figure 1(a)**, indicates that climatic control of diversity is difficult to determine when the area of observation is on the steeply rising curve of the species–area relationship. Nevertheless, small-scale experiments have often influenced ideas at the global scale. It is generally considered that sites of high NPP are highly competitive, with only a small number of dominant species. However, highly productive tropical forests are also highly diverse in terms of both species and vegetation structure, even with limiting soil nutrient resources. The vegetation canopy may have many strata, with different species in each strata. High NPP requires a high supply rate of resources, particularly soil nutrients, water, and solar radiation, with no significant temperature limitations. On low-nutrient soils, the vegetation retains the majority of the soil nutrients, whereas water is in abundant supply. In these forests, turnover of plant parts (leaves, twigs, and small branches) is quite rapid, as is the decomposer system and thus also the rate of nutrient recycling. Therefore, the high NPP goes hand in hand with a high requirement and supply of nutrients to all component species. This appears to be the major mechanism by which high NPP can sustain high diversity. In addition, large areas of forest also transpire large quantities of water vapor, which returns as rain and also increases local cloudiness. The high NPP vegetation therefore exerts some control over its own climate, making it more equable and moderating energy supply.

It is clear that the energy availability, as represented by evapotranspiration, provides a good prediction for species richness globally, and this has been linked with NPP, which is modeled as peaking at low latitudes. However, the widely accepted view that NPP increases toward the equator has been questioned recently (Huston and Wolverton, 2009), a development that is likely to call into question many assumptions about the mechanisms that underpin the climate–species richness relationship. This development may well provide a new impetus to the question of climate control of species diversity.

Climatic Extremes

Climate appears to influence diversity by a range of mechanisms such as by influencing turnover times of soil nutrients and through constraining the length of the growing season. However, it is also well established that species have a range of minimum temperature tolerances which is much greater than the range of maximum temperature tolerances (Woodward, 1987). At the global scale, absolute minimum temperatures in vegetated areas of the world vary from approximately -80 to 20 °C and species are appropriately adapted to the local absolute minimum temperature, even if this occurs only every 10–30 years. Therefore, minimum temperature tolerance increases with latitude. There also appears to be a cost involved in possessing a particular degree of temperature tolerance, with the cost probably reducing growth capacity and competitive ability compared with those of species possessing a lower minimum temperature tolerance (Woodward, 1987). As the absolute minimum temperature decreases, so also does the number of species that can survive (Woodward, 1987).

The absolute minimum temperature decreases with latitude; therefore, this mechanism alone could account for the latitudinal trend in species richness (**Figure 5**), if fewer species have evolved the capacity to endure increasingly lower temperatures. However, this does not account for the absence of low-temperature-tolerant species from warm climates; in addition, low absolute temperatures are associated with shorter and cooler growing seasons. Therefore, the impact of climatic extremes on richness almost certainly occurs in parallel with other climatic limits.

It is possible to use simple models to investigate the impacts of changes in absolute minimum temperature and competition on distribution and diversity. **Figure 12** shows a typical output of a model in which species are designed

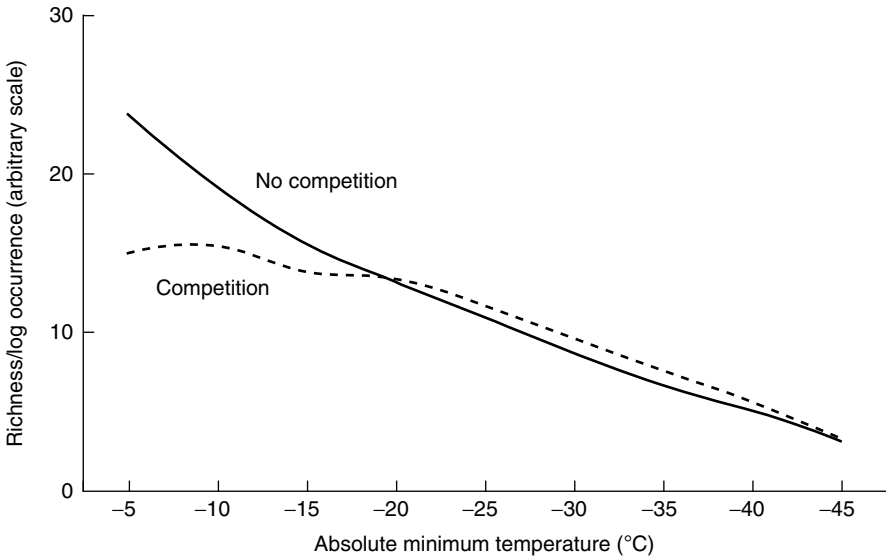


Figure 12 Simulated changes in species richness with absolute minimum temperature, without competition (solid line) and with competition (dashed line).

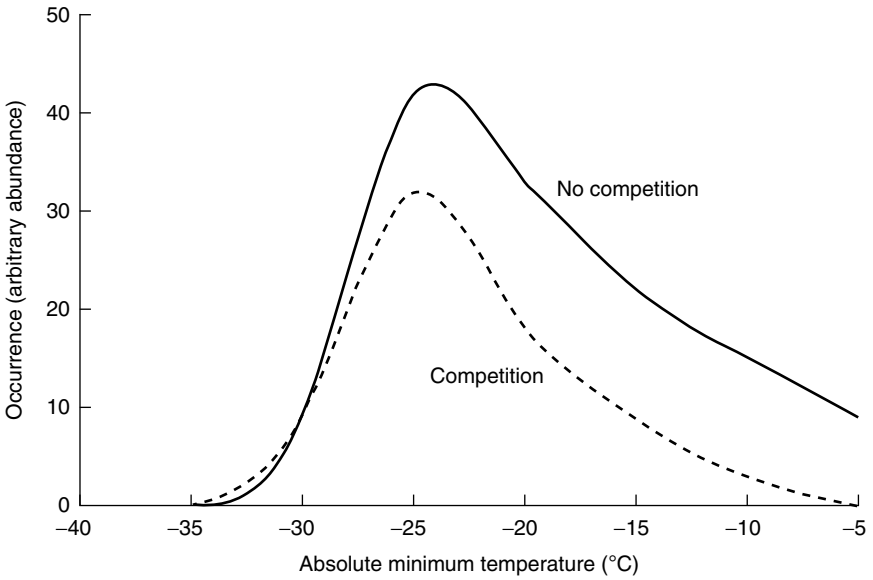


Figure 13 Simulated occurrence of a species with a minimum temperature tolerance to −35 °C, without competition (solid line) and with competition (dashed line).

with the capacity to survive to predetermined minimum temperatures but have the capacity to survive at all higher temperatures. In a second case, the competitive ability of the species declines as temperatures increase above the low temperature cutoff for survival. In cases of both competition and no competition, richness increases with temperature. However, in the simulation with competition, diversity increases less at temperatures higher than −20 °C. This occurs because competition excludes those species with the lowest temperature tolerances from the highest temperatures. This does not occur in the simulation with no competition. **Figure 13** illustrates the case for a species with a minimum temperature tolerance to −35 °C. Competition exerts no effect on abundance to −30 °C, but at higher temperatures the abundance or occurrence of this species decreases consistently to zero at −5 °C. Without competition, there is still a reduction in the occurrence of this species with increasing temperature, and this occurs simply because the chance of occurrence decreases as the pool of species increases with temperature. **Figure 12** demonstrates that richness decreases with decreasing temperature; however, the occurrences of the tolerant species increase, in line with the suggestion by Stevens (1989).

The model is simple but it readily predicts changes in diversity with temperature and also indicates that a reduction in richness will be associated with an increase in the abundance of species. Both examples (**Figures 12** and **13**) also indicate that the effect of competition is primarily one of degree, and therefore it will prove very difficult to differentiate between direct competitive effects and a simple lottery model as both temperature and the species pool increase.

Biogeographic History

It appears that competition, disturbance regime, energy availability, and climatic extremes can all influence the way in which climate affects species diversity. It also seems likely that we cannot readily differentiate between these controls at the global scale. However, none of the models are capable of explaining the sites that are outliers in the general relationships between richness and latitude (**Figure 5**), temperature (**Figure 6**), precipitation (**Figure 7**), and NPP (**Figure 8**). Further analysis indicates that these areas all have high frequencies of endemic species, and at the global scale species richness is positively correlated with the degree of endemism. The high degree of endemism in the fynbos of South Africa (*see* Species Richness and Latitude) and kwongan vegetation of South West Australia appears to be strongly correlated with the degree of soil infertility. In these areas, richness decreases with increasing soil fertility. However, both landscapes are very infertile and ancient in geological terms, and it seems likely that the infertile soil serves as a barrier to immigrants. The coastal Atlantic forests of Brazil are isolated from other areas by an inland mountain chain, indicating another method of limiting migration and cross-fertilization and enhancing species richness.

Processes associated with history and regional landscape structure seem to be the most parsimonious explanations for the high species richness in these areas. In the Cape floral kingdom of South Africa, there are more than 500 species in the genus *Erica* (Bond, 1989). The species are structurally very similar and it seems unlikely that their ecological characteristics are very different. Such situations suggest that the species concept is not an ideal measure of richness, particularly of ecological richness. The indicated areas of high diversity (**Figure 5**) certainly indicate high taxonomic richness, but is this relevant to the climatic controls of richness? A growing body of work has identified the strong role of climate history, particularly during the Pleistocene (ice ages) on species diversity, and this is providing insights into these species diversity outlier regions. Analysis shows that these regions may have escaped exposure to the extreme climatic shifts that were experienced over much of the rest of the planet during the

Pleistocene (Jansson, 2003), allowing species to diversify with low rates of extinction.

Experiments

Numerous experiments have investigated the relationship between species richness and productivity. However, through practical necessity, these experiments have been carried out on a small scale, well within the range of rapidly increasing richness with area, as shown in **Figure 1**. Tilman (1999), a major contributor to the productivity–diversity debate with Grime (1997), has indicated that this understanding, developed from scales where sites are populated from the same regional species pool, is inappropriate for addressing larger-scale latitudinal patterns. Although this scale may not be appropriate for the larger-scale considerations discussed in this article, it is interesting to explore briefly its possible relevance.

One experimental approach which can, in part, be used to address the relationship between diversity and climate is that of a herbaceous productivity–diversity multisite experiment distributed across Europe (Hector *et al.*, 1999). Eight sites were distributed between 39° N and 64° N latitude and 8° W and 27° E longitude. In all cases, these grassland experiments were seeded from a bare soil state with a known number of species, from one to at least eight. At all sites, aboveground biomass, a measure of productivity, increased with species richness, in keeping with the correlation shown in **Figure 8**. Although the expected reduction in biomass with short growing season (either through drought or low temperatures) was clearly demonstrated, it was not possible, in a 2-year experiment, to determine the effects of climate on the equilibrium species richness of the plots. More recently, strong evidence has emerged that there appears to be no relationship between species diversity and productivity at this level of spatial scale (Huston *et al.*, 2000). Indeed, a large number of experiments have been completed since the 1990s, and a meta-analysis of the results is now revealing that biodiversity effects on aspects of ecosystem function vary between ecosystem types, show idiosyncratic patterns (Schmid *et al.*, 2009), and are therefore very difficult to generalize.

Genome Size

In addition to resolving the mechanisms by which climate controls species richness, it is also important to be predictive. For example, given a particular climate, what will be the diversity and what will be the characteristic behavior of the local species? An alternative question is, given a particular metric of a species, where might the species be distributed naturally and in how diverse a mixture of species? These questions cannot be resolved; however, there has been a continued interest in relating the size of the nuclear genome of a species in this way (Grime, 1998). Relationships between genome size and critical climatic tolerances of frost resistance, precipitation supply, and growing season have been found, as have relationships between genome size and plant growth dynamics. The understanding of this issue has developed with the proposal of the "large genome constraint" hypothesis. In essence, it is now emerging from the data available that plant species with large genome sizes are constrained in terms of their photosynthetic rates, their diversification rates are lower, and they have restricted ecological ranges (Knight *et al.*, 2005).

Conclusions

The impacts of climate on plant biodiversity are best studied on the large scale, on the order of hundreds of kilometers. This approach escapes the currently insoluble problems of disentangling the effects of scale, landscape structure, and negative small-scale relationships between productivity and diversity. Freed from these constraints, the large-scale nature of the study indicates that observations and correlations are the primary mechanisms for understanding the processes by which climate controls diversity. The processes are still not fully clear, but primary productivity, disturbance, length of the growing season, and competition all appear to play important roles. There are some locations where species richness is much larger than the average that would be expected based on current climate. These areas are rich in endemic species which have probably originated and persisted because of migratory and reproductive barriers, and benefited from less variable Pleistocene glacial/interglacial climatic shifts. In these areas, it is likely that species differ very little if at all in ecological characteristics, and reflect a globally rare phenomenon of species accumulation by extraordinary diversification and persistence that partly reflects climate history. The widely accepted view that NPP increases toward the equator has been recently questioned, a development that should provide a new impetus to the question of climate control of species diversity.

Appendix

List of Courses

1. Environmental Science
2. Ecology
3. Terrestrial Ecosystems
4. Biogeography

See also: Climate Change and Ecology, Synergism of. Competition, Interspecific. Disturbance, Mechanisms of. Energy Flow and Ecosystems. Species–Area Relationships

References

Báldi A (2008) Habitat heterogeneity overrides the species–area relationship. *Journal of Biogeography* 35: 675–681.

Barthlott W, Lauer W, and Placke A (1996) Global distribution of species diversity in vascular plants: Towards a world map of phytodiversity. *Erdkunde* 50: 317–327.

Bond WJ (1989) Describing and conserving biotic diversity. In: Huntley BJ (ed.) *Biotic Diversity in Southern Africa: Concepts and Conservation*, pp. 2–18. Cape Town: Oxford University Press.

Bond WJ, Woodward FI, and Midgley GF (2005) The global distribution of ecosystems in a world without fire. *New Phytologist* 165: 525–537.
Cowling RM, Esler KJ, Midgley GF, and Honig MA (1994) Plant functional diversity, species diversity and climate in arid and semi-arid southern Africa. *Journal of Arid Environments* 27: 141–158.
Dean WRJ and Bond WJ (1990) Evidence for rapid faunal changes on islands in a man-made lake. *Oecologia* 83: 388–391.
Gaston KJ (1999) Why Rapoport's rule does not generalise. *Oikos* 84: 309–312.
Grime JP (1997) Biodiversity and ecosystem function: The debate deepens. *Science* 277: 1260–1261.
Grime JP (1998) Plant classification for ecological purposes: Is there a role for genome size? *Annals of Botany* 82(Supplement A): 117–120.
Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.
Hillebrand H (2004) On the generality of the latitudinal diversity gradient. *The American Naturalist* 163: 192–211.
Huston M (1979) A general hypothesis of species diversity. *The American Naturalist* 113: 81–101.
Huston MA, Aarssen LW, Austin MP, *et al.* (2000) No consistent effect of plant diversity on productivity. *Science* 289: 1255a.
Huston MA and Wolverton S (2009) *The Global Distribution of Net Primary Production: Resolving the Paradox*. Washington, DC: Ecological Society of America.
Hutchinson GE (1959) Homage to Santa Rosalia, or why are there so many kinds of animals? *The American Naturalist* 93: 145–159.
Jansson R (2003) Global patterns in endemism explained by past climatic change. *Proceedings of the Royal Society of London Series B – Biological Sciences* 270: 583–590.
Kelly CK and Woodward FI (2008) Responses of global plant diversity capacity to changes in carbon dioxide concentration and climate. *Ecology Letters* 11: 1229–1237.
Knight CA, Molinari NA, and Petrov DA (2005) The large genome constraint hypothesis: Evolution, ecology and phenotype. *Annals of Botany* 95: 177–190.
Kreft H and Jetz W (2007) Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences* 104: 5925–5930.
Lawton JH (1999) Are there general laws in ecology? *Oikos* 84: 177–192.
Linder HP, Johnson SD, Kuhlmann M, Matthee CA, Nyffeler R, and Swartz ER (2010) Biotic diversity in the Southern African winter-rainfall region. *Current Opinion in Environmental Sustainability* 2: 109–116.
Mackey RL and Currie DJ (2001) The diversity–disturbance relationship: Is it generally strong and fixed? *Ecology* 82: 3479–3492.
Midgley GF, Bond WJ, Kapos V, Ravilious C, Scharlemann JPW, and Woodward FI (2010) Terrestrial carbon stocks and biodiversity: Key knowledge gaps and some policy implications. *Current Opinion in Environmental Sustainability* 2: 264–270.
O'Brien EM, Whittaker RJ, and Field R (1998) Climate and woody plant diversity in southern Africa: Relationships at species, genus and family levels. *Ecography* 21: 495–509.
Rapoport EH (1982) *Aerography: Geographical Strategies of Species*. Oxford: Pergamon.
Ricklefs RE, Qian H, and White PS (2004) The region effect on mesoscale plant species richness between eastern Asia and eastern North America. *Ecography* 27: 129–136.
Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.
Ruggiero A and Werenkraut V (2007) One-dimensional analyses of Rapoport's rule reviewed through meta-analysis. *Global Ecology and Biogeography* 16: 401–414.
Schmid B, Balvanera P, Cardinale BJ, *et al.* (2009) Consequences of species loss for ecosystem functioning: Meta-analyses of data from biodiversity experiments. In: Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing*, pp. 14–30. Oxford: Oxford University Press.
Stevens GC (1989) The latitudinal gradient in geographical range: How so many species coexist in the tropics. *The American Naturalist* 133: 240–256.
Tilman D (1999) The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80: 1455–1474.
Whittaker RJ, Bush MB, and Richards K (1990) Plant recolonization and vegetation succession on the Krakatau Islands, Indonesia. *Ecological Monographs* 59: 59–123.
Wiens JJ and Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution* 19: 639–644.
Woodward FI (1987) *Climate and Plant Distribution*. Cambridge, UK: Cambridge University Press.
Woodward FI, Smith TM, and Emanuel WR (1995) A global primary productivity and phytogeography model. *Global Biogeochemical Cycles* 9: 471–490.

Coastal Beach Ecosystems

Anton McLachlan, The University of Sydney, Sydney, NSW, Australia
Omar Defeo, Faculty of Sciences, UNDECIMAR, Montevideo, Uruguay

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by McLachlan, A., volume 1, pp 741–751, © 2001, Elsevier Inc.

Glossary

Accretionary Accumulating sand.
Dissipative beaches Wide, flat beaches.
Effluent line Level on the beach face where the water table intersects the surface.
Macrotidal Describing a spring tide range of more than 4 m.
Microtidal Describing a spring tide range of less than 2 m.
Morphodynamics Interactions between the physical structure and the water and sediment movement in a beach and surf zone environment.
Psammophilic Sand-loving.
Reflective beaches Narrow, steep beaches.
Supralittoral Immediately above the intertidal zone.
Transect A linear series of samples, for example, across the intertidal zone.

Introduction

Sandy beaches dominate the ocean shorelines of all temperate and tropical continental coasts. These ecosystems are devoid of any biological structures and their morphology and dynamics can be defined in terms of three interacting factors: tides, waves, and sand particle size. Although still less studied than most other coastal systems, beaches in the new millennium are much better characterized. A significant body of knowledge has accumulated and beach studies have become increasingly theory driven. Here we describe the global range of wave-exposed, sandy beach types that can occur in response to changes in waves, tides, and sand. We also summarize the current ecological status for sandy beach macrofauna, focusing on key processes operating on different scales and how they interact to yield the patterns observed by ecologists.

Sandy Beach Types

Ocean beaches are defined by the interactions between tidal regimes, the wave energy they experience and the nature of the sand available for sorting and transport by the tides and waves. These interactions produce a range of beach morphodynamic types which span a continuum from microtidal reflective beaches, narrow and steep, to macrotidal dissipative systems, which are wide and flat and, under conditions of large tides, grade into tidal flats (Short, 1996). The simplest overall index of beach state is the beach slope, which is a product of the interaction among all three of these variables: beach face slopes flatten as wave energy increases, tide range increases or particle size decreases, if other factors are kept constant. Thus the flattest beaches occur in macrotidal regions of high wave energy and fine sand, and the steepest beaches occur in microtidal regions with low wave energy and coarse sand. A range of beach morphodynamic types can be distinguished between these extremes.

In microtidal regimes, where beaches are wave dominated, three beach states can be recognized: reflective, intermediate, and dissipative. The reflective beach, characterized by a steep face and absence of a surf zone, occurs under a combination of coarse sand and gentle waves. The shoreward transport of sand, which occurs under these conditions, causes all sediment to be stored on the subaerial beach face; the reflective beach thus represents the accretionary extreme in beach states. Here, waves surge up the beach face, where they may break before being reflected back to sea (**Figure 1**). Dissipative beaches, in contrast, are a product of large waves moving over fine sand. This results in a flat beach face and wide surf zone. Waves break far out and dissipate their energy while traversing the surf zone as bores before expiring as swash on the beach face. Dissipative beaches, with their sand spread out over extensive surf zones, thus represent the erosional extreme in beach states. Between these two extremes, intermediate beaches are distinguished by the presence of surf zones smaller (20–100 m wide) than in the dissipative situation. The intermediate surf zone has well-developed bars (sandbanks) and channels with rip currents (see **Figure 1**).

Beaches are not locked into a single morphodynamic state and respond to changes in wave energy by moving towards dissipative conditions during storms (and spring tides) and towards reflective conditions during calm weather (and neap tides): thus, sand erodes or accretes on the beach face as wave height (and tide range) rises or drops. Dean's parameter (Ω), also known as the dimensionless fall velocity, is a useful index that describes the state of a microtidal beach, that is, the extent to which wave energy is dissipated or reflected:

$$\Omega = \frac{H_b}{W_s \cdot T}$$

where wave energy is given by modal breaker height (H_b, cm) divided by modal wave period (T, seconds) and sand fall velocity is the sinking rate (W_s cm s^{-1}) of the mean sand particle size on the beach. Ω values <2 characterize reflective beaches, whereas $\Omega>5$ define dissipative ones and $2<\Omega<5$ characterize intermediate beach states.

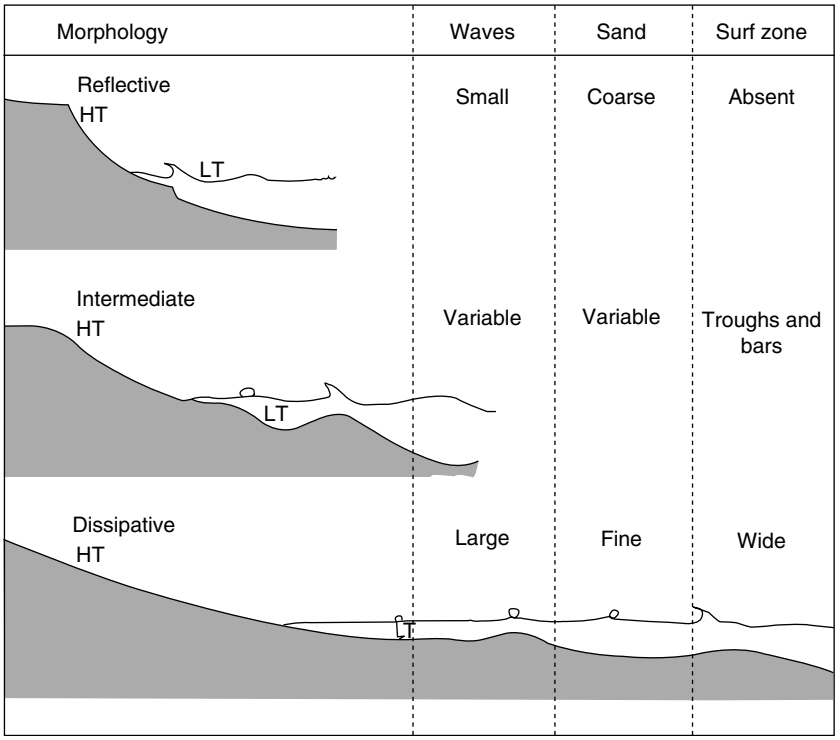


Figure 1 Three morphodynamic states of microtidal beaches (HT: high tide, LT: low tide).

The foregoing description of beach types adequately covers most microtidal situations, but increasing tide range tends to make beaches even more dissipative, because it allows the surf zone to work back and forth over a wider area (Short, 1996). Indeed, fully reflective beaches will not occur when tide range exceeds 1–1.5 m. On beaches with larger tides, reflective conditions can only occur at the top of the shore, between the neap and spring high-water swash lines: this area is reached by swash (but not surf), it is controlled by swash processes and is accretionary and steep, in contrast to the rest of the shore, where shoaling and breaking wave and bore processes operate.

Beaches are generally tide dominated under large tidal regimes (mean spring range >4 m), whereas in intermediate situations (tide range 2–4 m) they are mixed and either waves or tides can dominate. A useful index of the relative importance of waves and tides is the relative tide range (RTR), given by the mean spring tide range divided by the modal breaker height. By combining Ω and RTR, a two-dimensional model of beach states can be produced (**Figure 2**).

Ecologists have developed various indices as simple measures of overall beach state. Among these is the beach index (BI, McLachlan and Dorvlo, 2005), which can be used to compare beaches of differing tide range:

$$\mathrm{BI} = \log_{10}\left(\frac{M_z \cdot \mathrm{TR}}{S}\right)$$

where M_z is the mean grain size (in phi units $+1$, to avoid negative values), TR is the maximum spring tide range (m) and S is beach face slope (dimensionless). BI ranges 0–4, with values <1.5 representing microtidal reflective beaches and values >3 representing macrotidal dissipative beaches.

Sampling Beach Macrofauna

Community studies of beach macrofauna make use of standard beach transect surveys. These typically involve quantitative sampling across the intertidal zone by excavating quadrats and passing the sand through a screen of 0.5- or 1-mm mesh. The finer mesh size is more effective in sampling larval stages of macrofauna and some larger meiofauna, but is not practical where sand is coarse. Besides quadrat sampling, pitfall traps have been used to sample the supralittoral fauna. These are less quantitative but useful for collecting sparse fauna that move over large areas. In transect sampling, beaches are compared as if they were units, each transect survey representing one datum point. Species richness in beach surveys is summed for all samples in a transect and abundance is usually calculated per running meter of transect, so results are expressed as the number of species per transect and the number of individuals per linear meter of transect (ind m^{-1}), which can be estimated as follows:

$$\mathrm{IST_r} = \frac{\sum_{i=1}^{n} q_i}{n_\mathrm{r}} w_\mathrm{r}$$

where the mean density q (ind $\cdot\,\mathrm{m}^{-2}$ or g m^{-2}) of all n samples pertaining to transect r is multiplied by the corresponding width (w) of the species distribution across the beach. This equation is used to avoid biased results from using mean

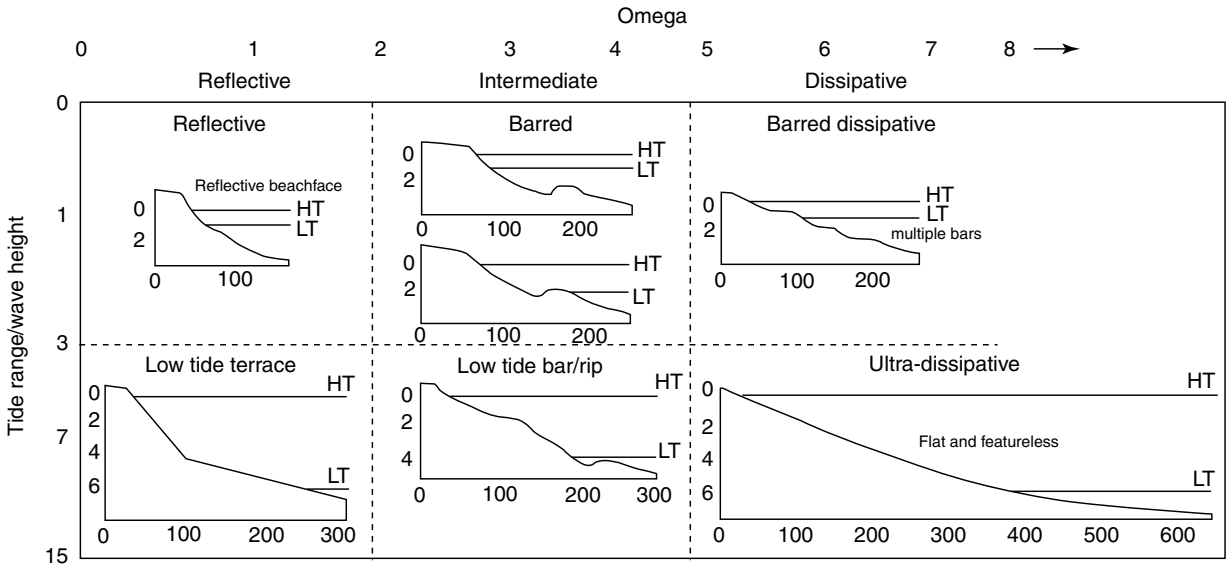


Figure 2 Two-dimensional model of beach states. Reproduced from Short AD (1996) The role of wave height, period, slope, tide range and embaymentisation in beach classification: A review. *Revista Chilena de Historia Natural* 69: 589–604. Distances are in meters.

density per quadrat ($ind \cdot m^{-2}$) as an abundance estimate, after which beach profiles have been impacted during either rough or calm conditions. Despite the above, density estimates are still being used in sandy beach ecology.

The width of sandy beaches ranges from as little as 10 m in reflective, microtidal situations to hundreds of meters in dissipative tidal flats, thus representing an order of magnitude range in the length of an intertidal transect. The total area sampled per transect (assessed as total area excavated, i.e., levels × replicates × quadrat size), however, has fallen within a much smaller range (2–4.5 m^2), with most cases being around 3 m^2. A total sampling effort per beach <2 m^2 introduces problems of undersampling (Jaramillo *et al.*, 1995) and species–area curves (**Figure 3**) indicate that a sample area of 3 m^2 recovers 90% of the species on microtidal beaches but would result in serious undersampling on wider beaches, such as macrotidal flats. Simulation and field studies indicate that sampling 4 m^2 of sediment from three systematic across-shore transects is sufficient for intermediate beaches, which are the most common morphodynamic type globally (Schoeman *et al.*, 2008).

Two different sampling designs are usually conducted on sandy beaches (Defeo and Rueda, 2002): (1) species driven and (2) environmentally driven. In the former, sampling points are systematically allocated along transects perpendicular to the shoreline to cover the entire across-shore distribution of the species; at least two quadrats with zero abundances are sampled in both the upshore and downshore directions before terminating a transect. The second sampling procedure is also based on shore-normal transects, with quadrats randomly allocated at fixed beach levels defined by the position of reference points, such as tidal marks (meso-/macrotidal beaches), swash levels (microtidal beaches), or drift and effluent lines. The validity of this sampling design

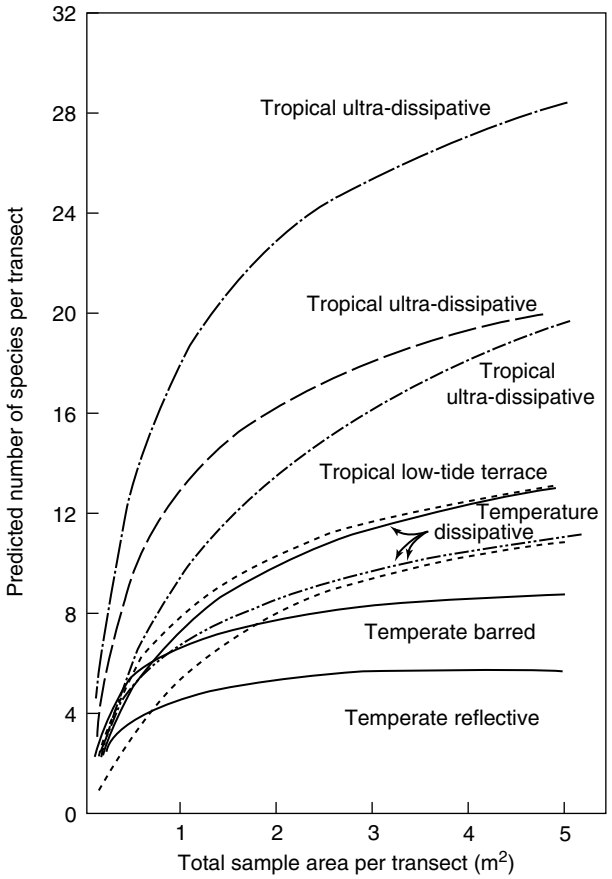


Figure 3 Species/area curves for sandy beach macrofauna. Reproduced from Jaramillo E, McLachlan A, and Dugan J (1995) The species/area relationship on sandy beaches. *Marine Ecology Progress Series* 119: 311–314.

has been questioned, because short- (hours, days) and medium- (seasons) term variations in macrofauna distribution might not necessarily follow swash levels or effluent or drift line positions (James and Fairweather, 1996; Brazeiro and Defeo, 1996). This diminishes the applicability of an environmentally driven sampling design defined by *a priori* fixed strata. A full discussion of sampling issues in sandy beaches is presented in Schoeman *et al.* (2008).

General Patterns

Macroscale

Latitudinal Trends

Species richness on sandy beaches increases from temperate to tropical regions for beaches of similar types (**Figure 4**). Comparison of latitudinal trends in species richness using beach face slope as covariate (ANCOVA, with slope log-transformed to fulfill statistical assumptions), shows that, for the same slope, the number of species is significantly greater for the tropics than for the subtropical, warm temperate and cold temperate zones. However, abundance and biomass tended to increase from tropical to temperate beaches, although these trends are less clear for polychaetes and molluscs than for crustaceans (McLachlan and Dorvlo, 2005).

Greater abundance and biomass on temperate dissipative beaches could be attributed to greater food availability, due to greater productivity. The presence of surf diatom blooms has led to the definition of temperate dissipative beaches as semiclosed ecosystems, in contrast to reflective beaches, defined as interfaces with low productivity, subsidized by organic inputs from the sea (McLachlan and Brown, 2006). Although increased wave energy in temperate areas seems to cause increased productivity and biomass, stranded seagrass or algae (Colombini and Chelazzi, 2003), or the presence of upwelling can also add to sand beach energetics and thus biomass.

Morphodynamic Effects

Species richness increases in response to increasing tide range, increasing wave energy, decreasing sand particle size and in flatter and wider beaches. This means that macrofaunal species richness increases as beaches become more dissipative (McLachlan and Dorvlo, 2005). For ocean beaches, the number of species recorded in a single transect survey ranges from 1 to 40 (if insects are excluded). More dissipative beaches and flats may harbor 20–40 species at least, as the shore flattens towards the dissipative or tidal flat extreme.

The best models fitted to 194 beaches based on data from five continents shows that the relationships between species richness and morphodynamic features of beach type are not linear. Summarizing global trends for a diverse range of exposed sandy beaches, **Figure 5** shows that species richness: (1) increases asymptotically from reflective to dissipative beaches (**Figure 5(a)**), with increasing tide range (**Figure 5(f)**), and with a measure of intertidal area obtained by dividing tide range by the beach face slope (**Figure 5(e)**); and (2) decreases exponentially with increasing sand particle size (**Figure 5(c)**) and slope (**Figure 5(d)**). Moreover, species richness increases linearly with the Beach Index (**Figure 5(b)**), i.e. from microtidal reflective to macrotidal dissipative beaches.

Following the same trend as for species richness, macrofaunal abundance and biomass decrease exponentially with increasing sand particle size and slope (**Figure 6**). Thus, species richness in sandy beaches increases from microtidal reflective to macrotidal dissipative conditions, whereas abundance and biomass increase exponentially, leading to greater densities of organisms.

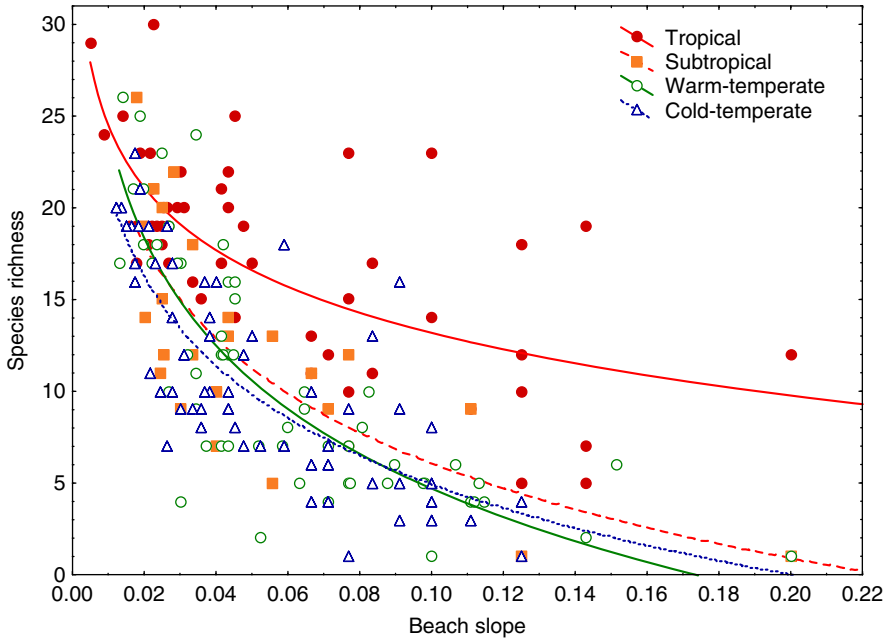


Figure 4 Latitudinal trends in species richness (excluding insects), based on 194 beaches in 10 countries using beach face slope as covariate. The model fitted for all the four regions ($Y = a + b \cdot \log X$) was highly significant ($p \ll 0.001$).

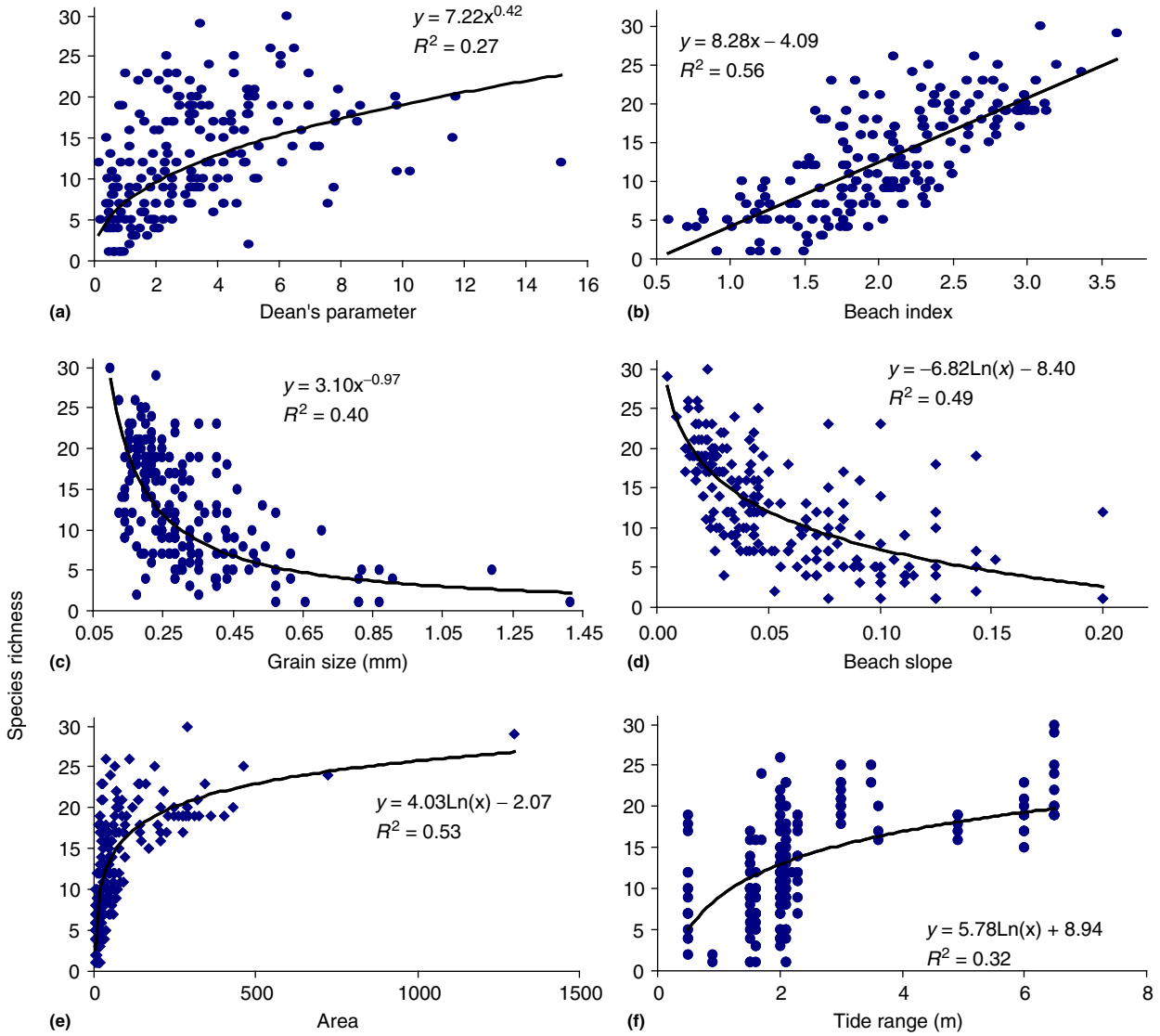


Figure 5 Global patterns of species richness on sandy beaches from five continents ($n = 194$). Relationships between species richness and (a) Dean's parameter Ω; (b) Beach Index; (c) grain size; (d) beach face slope; (e) the composite index Area; and (f) tide range. All models included in each panel were highly significant ($p \ll 0.001$).

Mesoscale: Zonation and Alongshore Distribution

The mesoscale concerns variations within a single beach, i.e., an uninterrupted sandy shore, both in the alongshore and across-shore directions (Defeo and McLachlan, 2005). The intertidal macrofauna of ocean sandy beaches is usually dominated by crustaceans, molluscs, and polychaetes, with other groups, such as insects, nemertean worms, echinoderms, anemones, and fishes, being of minor importance or restricted to the extreme upper or lower beach levels. Crustaceans tend to be most successful in reflective conditions, where their great mobility enables them to cope with turbulence. Among the crustaceans, ocypodid crabs, hippid crabs, cirolanid isopods, a variety of amphipods, and psammophilic mysids are the most typical. Molluscs, both gastropods and bivalves, are successful over a wide range of beaches. Clams, in particular, may form large populations on high-energy dissipative beaches, where they can support commercial fisheries. Polychaetes are the group that is most sensitive to beach state and are absent or scarce on reflective or coarse sand beaches. On low-energy shores of fine sand, polychaetes can be particularly abundant and include predator/scavengers, deposit, and suspension feeders.

Sandy beach populations and communities exhibit a well-defined structure with significant across- and alongshore variations. Macrofaunal species tend to be aggregated in elliptical patches from m to km, with the major axis parallel to the shore, and varying according to the different susceptibility of each species and year group (i.e., recruits vs. adults) to environmental factors (Defeo and McLachlan, 2005). In this setting, sandy beaches display zonation of their macrofauna (McLachlan and Jaramillo, 1995). A three-zone model was first developed by Erik Dahl: a supralittoral zone characterized by ocypodid crabs in warm regions and talitrid amphipods in

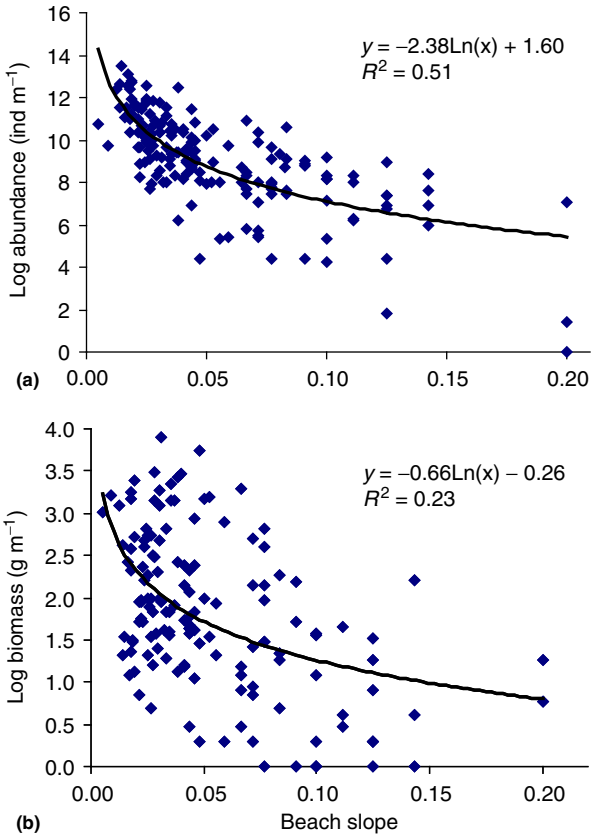


Figure 6 Best models ($p<0.001$) fitted to the relationships between beach face slope and total (a) abundance and (b) biomass, based on data gathered from 194 sandy beaches on five continents.

temperate areas; a littoral or midshore zone characterized by cirolanid isopods and spionid and opheliid polychaetes; and a lower shore or sublittoral fringe with many groups, including hippid crabs, mysids, haustoriid and phoxocephalid amphipods, donacid clams, and nephtyid worms. However, the number of recognizable zones depends on beach type. The lower zone tends to reduce or even disappear in very reflective situations, but it can expand to a broad, species-rich terrace under dissipative conditions. The supralittoral zone, above the high-water mark or drift line, is present on all shores. Thus, on most shores there is a clear gradient of increasing species richness as one moves downshore, typically from one or two species in the supralittoral to many species on the lower shore.

Zonation patterns and patchiness respond to an environment spatially and temporally structured by sharp gradients. Aggregations persist in time, but in contrast to sessile species in, e.g., rocky shores, the across-shore position of species, and their population structure vary according to the susceptibility of each species or population component (i.e., recruits and adults) to variations in beach topography. Aperiodic variations related to short-term increases in sea level generated by storm surges and barometric tides will cause the fauna to follow the high-water level; alternatively, seasonal variation occurs when species occupy higher levels during spring and summer than in autumn and winter. As faunal zones are dynamic, temporal studies are needed for a full picture of zonation patterns, requiring intensive sampling to provide unbiased estimates (Brazeiro and Defeo, 1996).

Species are mostly concentrated in the lower intertidal zone with richness decreasing upshore. Reflective beaches with coarse sand, dynamic swash action, and rapid drainage may be devoid of intertidal species and harbor only supralittoral forms, whereas loss of species under reflective beach conditions occurs mainly in the lower intertidal, not the supralittoral zone. The Habitat Favorability Hypothesis (Defeo and McLachlan, 2005) postulated that the center of the across- and alongshore range of a species, under optimal conditions of sand moisture, penetrability, and temperature, may be occupied by larger and dominant intraspecific competitors for space or food, whereas small individuals (recruits) may be displaced towards suboptimal conditions on one or both extremes of the distribution range. The same holds true in the case of interspecific competition.

Communities and populations in sandy beaches also vary along a beach (alongshore axis). Species richness within a single beach stretch tends to increase with beach length, decreasing sand particle size and in flatter and wider beaches, as well as from disturbed sites affected by freshwater discharges to undisturbed ones. Gradients in physical factors invoked to explain alongshore patterns have included changing exposure from sheltered to exposed beach ends, changing sand particle size, swash climate, morphodynamics, salinity, a combination of biotic (intra and interspecific interactions) and abiotic factors, and also passive sorting by the swash and active alongshore movements. Variations in human activities along a beach can also be causative factors of alongshore faunal variability (Defeo and McLachlan, 2005).

Microscale

The microscale concerns variations from mm to m and covers interactions in high-density patches and encompasses the area of influence of an individual (Defeo and McLachlan, 2005). Small-scale environmental gradients, together with intra- and interspecific interactions and behavioral factors, may affect community composition at this fine spatial scale on dissipative beaches where densities are much higher than on reflective beaches. Most of the macrofaunal biodiversity is concentrated within the top 20 cm of sediment, adults of large hippid crustaceans and mollusks being the few exceptions to this pattern. The third dimension, vertical distribution in the sand, is an additional way to partition space in order to ameliorate the strength of biological interactions. Investigations about the importance of physical and biological factors on this three-dimensional structure of the sandy beach habitat deserve future research.

Causative Factors: Main Hypotheses

Several models have been proposed to explain patterns in species richness on sandy beaches. Because of the physically controlled nature of beaches and the relative simplicity of the beach environment, these models have focused on explanations revolving around physical rather than biological control. This supports the general applicability of the

Autecological Hypothesis of Noy-Meir as adapted to sandy shores (McLachlan and Brown, 2006), which states that inhabitants of physically controlled environments respond independently to the physical environment. This appears to apply across the range of beach types, but decreasing towards dissipative conditions where biological factors may become more important.

The physical environment impinging directly on beach fauna differs primarily in two main variables: (1) the movement of water over and through the beach face, which has been termed the swash climate, and (2) the sediment particle size range. On most beaches, waves do not break directly in the intertidal but rather in the surf zone. It is only after transformation to bores and crossing the surf zone that wave energy reaches the beach face as swash. Thus, intertidal fauna experience the effects of waves, tides, and the surf zone transformation of these forces as the swash "climate" on the beach face. Sand particle size also changes with beach type, generally being coarser towards reflective states.

There is a consistent relationship between beach type and swash climate features (McArdle and McLachlan, 1991). Dissipative beaches are characterized by swash with extended periods and lengths, variable speeds, most swash activity below the effluent line (water table outcrop), and laminar flow. In general, conditions of fine sediment and long swash periods, with limited swash activity above the effluent line, appear most conducive to developing rich faunas. The key characteristic of this type of swash may, in fact, be its low degree of turbulence and laminar flow over fine, saturated sand, enabling even delicate forms to survive. Reflective beaches display the opposite swash features: the swash climate is extremely harsh; there are high swash speeds throughout the tidal cycle; and waves break directly in the intertidal, resulting in considerable turbulence and increased probability of animals being stranded above the effluent line where unsaturated sand makes burrowing difficult. Swash drainage through the sand is rapid on reflective beaches and there is little inundation time for feeding. Physical stress in the swash zone on the beach face thus increases from dissipative to reflective beaches.

McLachlan *et al.* (1993) refined these ideas to define the Swash Exclusion Hypothesis: the swash climate associated with dissipative beaches is considered sufficiently accommodating and varied to enable most psammophilic macrofauna species to maintain viable populations, but towards reflective conditions harsh swash climate increasingly excludes species. Swash also has important indirect effects on the macrofauna since animals move, feed, and reproduce in the water moving over the beach face. Swash patterns and percolation of water through the beach face above the effluent line are closely coupled.

Swash is not the only factor to be considered: sand also plays a role. Small body sizes tend to be excluded in coarse sediments, and sand particle size influences burrowing rate in a variety of sandy beach species – in most cases coarse sand makes burrowing difficult or impossible. Further, coarse sand can cause vicious abrasion, especially of bivalve shells, and it reduces sand saturation by raising permeability and hastening drainage. Thus, both sediment particle size and swash climate may directly influence the number of species of macrofauna on wave-exposed sandy beaches. At the population level, the Habitat Harshness Hypothesis postulates that in harsh reflective beach environments, organisms must divert more energy towards maintenance and they therefore have less scope for reproduction, with consequent higher mortality (reviewed in Defeo and McLachlan, 2005).

Bringing all this together, McLachlan and Dorvlo (2005) proposed two levels of control of intertidal sandy beach macrofauna biodiversity: (1) primary control is by tide range and latitude, the former defining the vertical dimension of the intertidal habitat and the latter determining the size of the species pool available to colonize a beach; and (2) secondary control is by harsh swash climates and coarse sand excluding species from colonizing beaches towards the reflective end of the spectrum. According to this hypothesis the primary factors determine the number of species that could occur on dissipative beaches in a particular region, whereas secondary factors limit the number of species that manage to establish populations across the range of beach types by excluding less well adapted species from the harsher conditions on reflective beaches. However, supralittoral species respond differently.

As beach type changes toward reflective conditions, the increasingly inhospitable swash climate and coarse sand exclude more species until, in the fully reflective situation, only supralittoral forms (talitrid amphipods, ocypodid crabs, insects), which live "outside" the area affected by swash, remain. The Habitat Safety Hypothesis separates intertidal and supralittoral forms, taking into account both swash and sand effects and differences in life histories. This hypothesis suggests that supralittoral species (in contrast to intertidal forms) are safer on reflective beaches, which usually have a distinct berm that is not inundated (Defeo and Gómez, 2005).

Threats

Sandy beaches are subject to many pressures, natural and anthropogenic, and these have recently been reviewed (Defeo *et al.*, 2009). A variety of human activities impact on beaches: recreation, pollution, development, mining, etc. However, the overarching problem is coastal squeeze: beaches are squeezed between rising sea level associated with climate change and increased coastal development associated with expanding human populations and coastal economies. There is considerable inertia in both these seaward and landward forces, so beaches will continue to be threatened for some time, even in the face of actively addressing these issues.

There are great challenges for ecologists studying beaches to design long-term monitoring and experimentation to properly understand the dynamics of key ecological attributes of these systems on different spatial and temporal scales. There are also challenges for coastal managers needing to conserve coastal biodiversity. The ideal management scenario is to have zoning of the shoreline, avoiding strip development and with large areas set aside for conservation. Also essential are generous setback lines that keep development well landward of beaches and dunes and allow space for these systems to retreat naturally in the face of sea level rise. This requires legislation

and enforcement. Sound coastal zone management supported by appropriate legislation and strong enforcement is key.

Discussion and Conclusions

Macrofauna populations and communities in sandy beach ecosystems respond primarily to the abiotic environment. This is a restatement of the idea of a physically controlled community and the Autecological Hypothesis. Thus, composite indices (Dean's parameter, Beach Index, Area) based only on the three physical variables that define a beach (tides, waves, and sand) have high predictive power to estimate the species richness likely to be encountered on ocean beaches. Latitude, beach length, and biological interactions may contribute additionally to explain the remaining variability in the data.

Tropical beaches support more species than do temperate beaches of the same type (McLachlan and Dorvlo, 2005), and long beaches appear to support greater abundance and more species than do short (or pocket) beaches. The former trend may be due to the greater species pool available to colonize beaches in the tropics, whereas the latter trend reflects the more favorable surf circulation patterns on long beaches, which promote greater retention of particulate primary production, better recruitment of planktonic larvae, and larger and therefore more resilient adult populations. Long beaches are also more persistent, that is, unlikely to erode away completely during storms, as can happen on pocket beaches.

Species richness and abundance in sandy beaches increase from microtidal reflective systems through high-energy dissipative systems to ultradissipative tidal flats. Species richness in wave-exposed beaches is therefore predictable on the basis of the physical nature of individual beach environments. Beaches that tend toward the ultra dissipative/tidal flat extreme – the product of large tide ranges, vigorous wave action, and fine sand – harbor most species likely to be encountered in a region. Progressing from this end of the beach continuum towards reflective conditions, that is, decreasing tide range or wave energy or increasing sand particle size, results in beaches becoming steeper and narrower and the increasing exclusion of species. On highly reflective beaches, intertidal macrofauna may be absent and only supralittoral forms remain.

Two physical factors define the immediate environment experienced by the beach macrofauna and may therefore control species richness: the swash climate and sand particle size. Reflective beaches have harsh swash climates in the sense of high turbulence, short swash periods, and rapid swash drainage, resulting in low effluent lines; beaches at the other extreme have more benign swash climates in the form of long-period swash/tidal bores with laminar flow over the beach face and low turbulence, keeping much of the beach saturated because of the slow drainage. All beach fauna can burrow into saturated sand and, in terms of grain size, fine to medium sands seem optimal for most beach fauna. Coarse sand appears to exclude small or delicate forms by crushing and abrasion and most species experience decreasing burrowing efficiency in coarse sands. Thus harsh swash climates and coarse sand appear to exclude many species, the degree of exclusion increasing from the dissipative/macrotidal to the reflective/microtidal end of the continuum of beach morphodynamic types.

The dynamic, three-dimensional nature of wave-exposed ocean beaches, their shifting populations, and the absence of biological structures (they are even devoid of permanent burrows in most cases) leave limited scope for biological controls of community structure. However, biological interactions (intra- and interspecific) may be more important regulatory agents that previously thought in dissipative beaches, which have higher species richness and densities of organisms and hence greater potential for interactions between individuals (Defeo and McLachlan, 2005).

Beaches are widely valued for their scenic vistas and recreational potential but are threatened by coastal squeeze, being caught between sea level rise on the marine side and expanding development along the shoreline. Strong coastal zone management based on good science is essential to conserve the biodiversity present on sandy beaches.

Several lines of research could improve our understanding of patterns of sandy beach biodiversity and their control. Collaborative work between ecologists and geomorphologists could lead to refinement of our understanding of the coupling between physical and biological process on sandy beaches on various temporal and spatial scales. Key areas for the future will be investigating the importance of physical and biological factors on finer scales, through population studies and experimental work on ecophysiology and behavior. Only by deciphering cause and effect relationships at the individual and population levels we will truly begin to understand the processes that underlie the large scale community patterns that have become clear. Once these processes and mechanisms are better understood, more accurate and useful predictions can be made of global beach diversity patterns and their causes. This capability would be useful in predicting and identifying areas of high diversity and in planning coastal conservation strategies.

See also: Estuarine Ecosystems. Intertidal Ecosystems. Lake and Pond Ecosystems. Marine Ecosystems. River Ecosystems

References

Brazeiro A and Defeo O (1996) Macroinfauna zonation in microtidal sandy beaches: Is it possible to identify patterns in such variable environments. *Estuarine, Coastal and Shelf Science* 42: 523–536.

Colombini I and Chelazzi L (2003) Influence of marine allochthonous input on sandy beach communities. *Oceanography and Marine Biology: An Annual Review* 41: 115–159.

Defeo O and Gómez J (2005) Morphodynamics and habitat safety in sandy beaches: Life history adaptations in a supralittoral amphipod. *Marine Ecology Progress Series* 293: 143–153.

Defeo O and McLachlan A (2005) Patterns, processes and regulatory mechanisms in sandy beach macrofauna: A multi-scale analysis. *Marine Ecology Progress Series* 295: 1–20.

Defeo O, McLachlan A, Schoeman DS, *et al.* (2009) Threats to sandy beach ecosystems: A review. *Estuarine, Coastal and Shelf Science* 81: 1–12.

Defeo O and Rueda M (2002) Spatial structure, sampling design and abundance estimates in sandy beach macroinfauna: Some warnings and new perspectives. *Marine Biology* 140: 1215–1225.

James RG and Fairweather PG (1996) Spatial variation of intertidal macrofauna on a sandy ocean beach in Australia. *Estuarine, Coastal and Shelf Science* 43: 81–107.

Jaramillo E, McLachlan A, and Dugan J (1995) The species/area relationship on sandy beaches. *Marine Ecology Progress Series* 119: 311–314.

McArdle S and McLachlan A (1991) Dynamics of the swash zone and effluent line on sandy beaches. *Marine Ecology Progress Series* 76: 91–99.

McLachlan A and Brown AC (2006) *The Ecology of Sandy Shores*, 2nd edn. Amsterdam: Elsevier.

McLachlan A and Dorvlo A (2005) Global patterns in sandy macrobenthic communities. *Journal of Coastal Research* 21: 674–687.

McLachlan A and Jaramillo E (1995) Zonation on sandy beaches. *Oceanography and Marine Biology, Annual Review* 33: 305–335.

McLachlan A, Jaramillo E, Donn TE, and Wessels F (1993) Sandy beach macrofauna communities and their control by the physical environment: A geographical comparison. *Journal of Coastal Research (Special Issue)* 15: 27–38.

Schoeman DS, Nel R, and Goulart Soares A (2008) Measuring species richness on sandy beach transects: Extrapolative estimators and their implications for sampling effort. *Marine Ecology* 29: 134–149.

Short AD (1996) The role of wave height, period, slope, tide range and embaymentisation in beach classification: A review. *Revista Chilena de Historia Natural* 69: 589–604.

Coevolution

Douglas J Futuyma and André Levy, State University of New York at Stony Brook, New York, NY, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 753–767, © 2001, Elsevier Inc.

Glossary

Character displacement Pattern of geographic variation in which a character differs more between sympatric than between allopatric populations of two species.

Competition Interaction between individuals of the same or different species, whereby resources (e.g., food, space, mating partners) used by one are made unavailable to others.

Mutualism Symbiotic relation in which each of two species benefits from the interaction.

Resistance An individual's capacity to reduce the damage inflicted upon it by an enemy.

Specialization Evolutionary adaptation in a particular mode of life or habitat.

Tolerance An individual's capacity to sustain damage by an enemy with limited decrease in fitness.

Trade-off Negative correlation between traits, such that a benefit due to changes in the value of one trait is associated with a cost produced by changes in the value in another trait.

Virulence Degree to which a parasite reduces the probability of survival or the reproductive capacity of an individual host. A relatively avirulent (benign) parasite has little impact on its host's fitness.

Meaning and Varieties of Coevolution

(1) *Patterns of correlated evolution*, especially *phylogenetic congruence*, among species or other lineages are usually detected by phylogenetic analysis. In extreme cases, the phylogeny of one clade of species, such as host-specific parasites, is congruent with that of a clade with which it interacts such as host species (**Figure 1(a)**). Such congruence would require *coincident speciation* of hosts and associated parasites. Deviations from perfect congruence might be caused by speciation of a lineage but not its associate, by extinction of one but not the other, or by a parasite switching from one host to another (colonization) (**Figure 1(b)**).

In some cases, the coincident speciation that underlies concordant phylogenies may be caused by the interaction between associated species. However, concordance might stem solely from a shared history of geographic isolation, without the interaction having played any causal role. Although phylogenetic concordance does not in itself provide evidence of reciprocal adaptation of the interacting species, it does imply extended historical opportunity for adaptation to have occurred. The greatest phylogenetic concordance has been described for some maternally transmitted endosymbiotic bacteria and their eukaryote hosts, and for certain ectoparasites (e.g., lice of gophers) that are primarily transmitted by contact among conspecific hosts. In many groups of parasites and herbivorous insects, related species are associated with related hosts, that is, members of a clade, but the phylogenies of the two clades are not congruent. This pattern often implies a long history of association, but one that has included shifts of parasite lineages among host lineages.

(2) Coevolution consisting of *reciprocal adaptive responses* has been rather arbitrarily classified as *specific* (or pairwise, or tight) coevolution and *diffuse* (or guild) coevolution. Specific coevolution is reciprocal adaptation of two species, independent of their interactions with other species, whereas diffuse coevolution occurs when evolutionary change in one species affects its interaction with two or more other species (e.g., a change in a host that affects two parasite species). Specific coevolution is easier to visualize, model, and detect. Diffuse coevolution is most likely to occur when a species' response to selection by one interacting species is accompanied by genetically correlated effects on its interactions with other species. The term "diffuse coevolution" has also been used to describe nonadditive selection on one species by two or more other species, owing to an interaction between them or between their effects on the recipient species.

(3) *Cospeciation* may be induced by the interaction between species, but this probably requires special conditions. A likely case is in figs (*Ficus*) and the highly host-specific seed-eating wasps that pollinate them. Divergence between population of either a fig species or a wasp species in features governing the host preference of the wasp may engender divergence in the figure resulting in reproductively isolated populations of both. Such cospeciation may result in concordant phylogenies.

(4) *Escape-and-radiate coevolution* describes a scenario in which an evolutionary lineage diversifies after it evolves a defense that breaks its association with enemies. Such a scenario was first proposed by P. R. Ehrlich and P. H. Raven for plants and herbivorous insects. In this hypothesis, a plant lineage that evolves a novel defense sheds most of its herbivores, and therefore (by an unspecified causal inference) diversifies into many species. Only later do some herbivore species adapt to one or more of the plant species, and diversify in turn.

Collectively, these hypothetical scenarios suggest that coevolution may increase both the numbers of species and the phenotypic diversity (disparity) of species by selecting for the many characteristics that affect their interactions. Coevolution may therefore be an important engine of biological diversity, but we shall shortly note that coevolution may also cause extinction.

Often, the adaptation of one species to others is obvious, but reciprocal adaptation of the others is not. Because we seldom know a priori whether or not adaptation has been

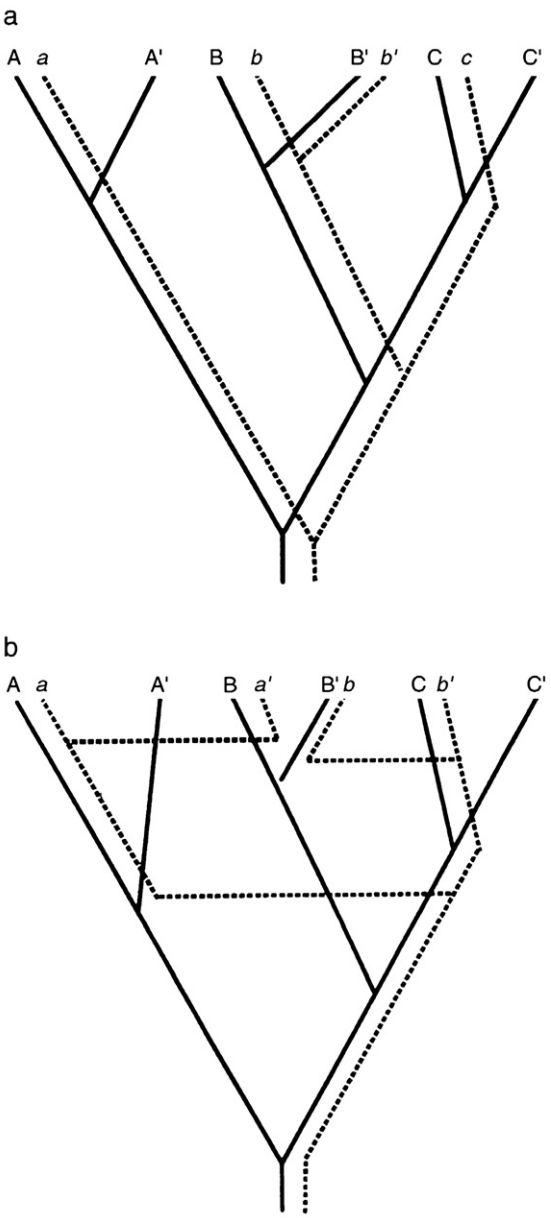


Figure 1 Congruent (a) and incongruent (b) phylogenies of hosts (solid lines) and parasites (dashed lines). Letters indicate most closely related taxa. Each parasite taxa is specialized on the host with which it is closely associated in the diagram.

reciprocal, adaptations to ecological interactions are often loosely referred to as "coevolutionary," even if evidence for reciprocity is slight or lacking. Certainly the *study* of coevolution includes adaptations to other species, which may or may not prove to have adapted in turn.

Evidence of Coevolution

Sources of Evidence for Coevolution

The study of coevolution comprises the same approaches as studying evolution in general. As in the broader field, the first forms of evidence for coevolution consisted of detailed natural history observations, descriptions of the diversity of adaptive structures that mediate ecological interactions, and comparison among populations and species.

Charles Darwin published the first comprehensive illustration of how intricately species are adapted to one another and how structural complexity can be explained by coevolution in his description *The Various Contrivances By Which Orchids Are Fertilized By Insects* (1877). By comparing the shape of different orchid flowers and their associated pollinators, Darwin demonstrated that certain moth features are needed to successfully obtain nectar from the flower, features that are indeed exhibited by their specialized pollinators. By restricting nectar collection to a few pollinators, orchids increase the likelihood of cross-fertilization. Such is the correspondence between flower and pollinator shape that on observing the 29-cm-long nectar-bearing spur of the Madagascan orchid *Angraecum sesquipedale*, Darwin predicted the existence of a pollinating moth with a proboscis of that length. Such a moth, *Xanthopan morgani praedicta*, was indeed discovered 40 years later.

The description of patterns in plant use by lepidopteran larvae preceded the concept of *escape-and-radiate coevolution*. Higher taxa of butterflies often feed upon a single group of flowering plants. While some feed on more than one plant family, these tend to be closely related or have similar biochemistries. For instance, larvae of the butterfly subfamily Pierinae, or whites, feed predominantly on the families Capparaceae and Brassicaceae, which are closely related. Some whites also feed on members of the family Tropaeolaceae that share with the other families the production of mustard oil glycosides and a rare fatty acid. These regularities imply an important role for plant secondary metabolites in determining butterfly host use. Given that these compounds affect herbivore behavior, acting often as deterrents, secondary chemistry may have constituted the key feature that allowed plant escape.

Comparisons among conspecific populations have also been suggestive of coevolution. The coloration pattern of the butterfly *Heliconius erato*, thought to be a signal to predators indicating distastefulness, varies among populations in Central and South America. Strikingly, the wing coloration of *H. melomene*, an equally distasteful congener with distinct life history and host preference, varies geographically in parallel with *H. erato*. This pattern is thought to be an example of coevolution of mimicry between prey species that share a predator. Fritz Müller, a contemporary of Darwin, first suggested this particular model of coevolution to explain similarities in wing pattern among unpalatable butterfly species belonging to two distinct genera (*Ituna* and *Thyridia*).

Müller also introduced the use of mathematical models to study the coevolutionary process. Modern mathematical and computer simulation models may incorporate population genetics, quantitative genetics, evolutionary game theory, and optimality theory. Mathematical modeling has proven useful in describing the dynamic of the interactions between species and in determining which conditions favor coevolution.

Although ecological interactions usually do not "fossilize," the analysis of paleontological records has provided some evidence of coevolution. For instance, the appearance in the

Ordovician of predaceous cephalopods is associated with the simultaneous appearance of several defensive strategies on the part of their prey (e.g., strong sculpture and coiling in gastropods and shell-bearing cephalopods, spines in echinoderms), suggestive of diffuse coevolution between predators and their prey. The antiquity of certain interactions may also be determined by inspecting fossils of extant species. Several plant families possess structures (domatia) that harbor mites, which attack plant enemies. Domatia similar to the modern form have been discovered in fossilized leaves from the Eocene, 55 million years ago!

The relative age of clades of associated taxa is relevant for demonstrating correlated coevolution or co-speciation. These processes would be necessarily excluded if one group were much older than the other. The age of an association, or of interaction-related adaptations, can often be estimated from phylogenies with time calibration (e.g., using approximate molecular clocks or stem-group fossils). Molecular evidence from deep-sea vesicomyid clams and the sulfur-oxidizing endosymbiotic bacteria on which they depend for nourishment indicates that the interacting clades are both approximately 100 million years old. These two lineages appear to have been in close association since their origin and to have cospeciated, as indicated by the remarkable level of congruence between their estimated phylogenies. Phylogenetic information also becomes relevant in testing whether a character is an adaptation for an ecological interaction or an ancestral feature that exists in the absence of the interaction.

In some instances it has been possible to document the particular genes that affect a species' interaction. H.H. Flor found several genes in flax (*Linum usitatissimum*) that provide resistance to the rust *Melampsora lini*. Rust virulence is determined by a set of complementary genes, in a one-to-one relationship. This study inspired the gene-for-gene model (*see* Special Features in Parasite-Host Coevolution: Gene-for-Gene Systems), which has become a paradigm of phytopathology. Most traits, however, have a complex genetic basis, involving many genes. Such complexity requires a quantitative genetic approach, which partitions the trait variation into genetic and environmental components. This approach has demonstrated that many of the traits relevant to interactions have genetic variability, that is, there is potential for coevolution. For example, the wild parsnip (*Pastinaca sativa*) and its most important associated herbivore, the parsnip webworm (*Depressaria pastinacella*), are thought to be engaged in coevolution mediated by the evolution of furanocoumarins and the insect's detoxifying mechanisms. May R. Berenbaum has documented genetic variation both in the production of furanocoumarins and in the webworm's ability to metabolize this group of plant toxins.

Quantitative genetics is also used to measure correlations between traits. The detection of negative genetic correlations is indicative of trade-offs between traits, such that selection for the increase in value of one trait leads to a decrease in value of the correlated trait. Tradeoffs are of particular relevance in explaining evolutionary constraints, and particularly why species are specialized. Pea aphid (*Acyrthosiphon pisum*) clones collected from two crop plants (alfalfa and red clover) exhibited higher fitness when reared on the plant from which they had been collected, suggesting local adaptation. The negative genetic correlation in fitness across crops may constrain the evolution of generalist clones, as these would be outcompeted on either plant by crop-specialized clones.

Measurement of genetic variation and correlation between traits offers information on the genetic context in which selection can act. Correlations between traits and fitness suggest the form and direction of selection. In a greenhouse study, the wild parsnip exhibited a negative genetic correlation between concentration of several furanocoumarins and seed set, suggesting that the production of the chemicals may impose a cost to reproduction in the absence of the parsnip webworm. These negative correlations were not detected in the field, indicating that presence of furanocoumarins increases fitness in the presence of the herbivore. Ideally such studies are performed in a natural setting, as our ultimate interest is understanding how natural selection works in the wild, but one can use model systems in the laboratory, such as evolving populations of bacteria and bacteriophage.

Finally, studies of interacting species are commonly based on the analysis of single communities. However, most species are composed of many local populations, and increasing importance has been attributed to the geographical structure of species and their interactions. Across the distribution of an interaction one is likely to observe a mosaic of selection pressures as a result of variation in abiotic and biotic factors, and the particular demographic and genetic histories of local populations. Some localities may be coevolutionary hot spots, that is, sites of reciprocal coevolution, whereas in others selection may be unidirectional or act on neither species. The geographic variation in outcomes is further modified by gene flow among populations. Consequently, different degrees of coadaptation are to be expected among populations. Clearly our understanding of the dynamics of a species interaction requires the study of many communities and interpopulation processes.

A few cases that satisfy the requirements of long-term multipopulation studies have emerged recently and have reinforced our need for a *geographic mosaic theory of coevolution*. For instance, resistance and virulence structures of *Linum* and *Melampsora* (referred to earlier), studied in New South Wales, sometimes vary dramatically across populations and time. The frequency of susceptible genotypes of flax will affect the local frequency of a particular strain of flax rust, but additional factors were found to play a role, namely, drift, extinction, and migration from neighboring populations. The geographic structure of flax and rust proved to be an essential factor in explaining the persistence of the interaction.

Rates of Adaptation

Phylogenetic reconstructions and the fossil record provide some examples of ancient interactions between groups of organisms. Character mapping of feeding habits onto beetle phylogeny indicates that conifers and cycads are the ancestral host plants. The oldest fossils of herbivorous beetle taxa, from the Jurassic, are indeed ancestral forms of modern conifer and cycad feeders, suggesting that this association may have evolved 200 million years ago, making it the oldest insect/plant association known. Lichens, which are an intimate

association between fungi and green algae, may be 100–200 million years old. Yet how often do new interactions establish themselves, how rapidly do species adapt to one another, and does that process occur on a timescale that is relevant to ecological processes?

The evolution of interspecific interactions can occur very rapidly, that is, during the course of 100 years. Examples include cases where recorded changes in the environment led to measurable changes in phenotype. In 1977, the Galápagos Islands suffered a severe drought during an El Niño event, causing an increase in the median size of available seeds. During the same event, the medium ground finch (*Geospiza fortis*) suffered strong selection for an increase in bill size, as only larger-billed birds could effectively feed on the available large seeds. This morphological shift occurred in a single generation. Many cases involve species of animals and plants introduced into new habitats. Populations of the apple maggot (*Rhagoletis pomonella*) that fed exclusively on native hawthorn have in the course of the past 100 years become adapted to a novel host, the introduced apple. The loss of interaction-related adaptations has also been observed to occur on a short timescale. Within the course of a 34-year study, guppies (*Poecilia reticulata*) artificially introduced into drainages in Trinidad, where natural predators are absent, lost defensive antipredator behaviors exhibited in the source populations.

Although these examples unequivocally demonstrate that interspecific interactions can evolve rapidly, they do not provide examples of coevolution. Almost all examples describe the evolution of *one* species adapting to another, with little reference to reciprocal changes. One of the few exceptions is the effort to biologically control the rabbit (*Oryctolagus cuniculus*) in Australia. The introduction of the myxoma virus into Australia initially reduced the rabbit population drastically. Such strong selection led to the evolution in just a few generations of increased resistance in rabbits. This evolutionary change was accompanied by a decline in the virulence of the virus to an intermediate level. Although in this case evolutionary change was observed in both interacting species, this does not necessarily constitute a case of coevolution. The reduction in virulence did not occur as a result of changes in rabbit resistance, but rather because the initial strains of the virus were so virulent that they quickly killed the host, lowering the likelihood of transmission by the disease's vector. More recent data, however, seem to indicate that the initially more virulent strains are presently increasing in frequency, perhaps as a result of increased rabbit resistance.

Rapid evolution is not, however, unconstrained. Each species is influenced by its phylogenetic history, and even populations of the same species will possess different evolutionary potential. To grasp how often new interactions are formed, one needs to understand these constraints. Rapidly evolving interspecific interactions can potentially affect the dynamics of community structure. Models of community ecology should not rest on the assumption that evolutionary dynamics occur on a larger timescale that is irrelevant to ecological processes. But it has yet to be determined under which conditions coevolution drives community dynamics and whether it has a stabilizing role.

Competition Between Species

Hypotheses about the coevolution of competing species have long played a major role in both evolutionary biology and ecology. Most of the hypotheses are variations on the theme that species should evolve differences in resource utilization, thus reducing competition between them. For Darwin, the principle of "divergence of character" in response to competition, described in *The Origin of Species*, explained why an ancestral species should give rise to multiple, diverse descendant species. W. L. Brown and E. O. Wilson cited reduction of competition as a major cause of *character displacement*, the term they coined for a greater difference between sympatric than allopatric populations of two species. D. Lack and R. H. MacArthur provided famous examples of differences in diet and foraging behavior or trophic morphology among closely related species of birds, interpreting the patterns as stemming from avoidance of competition. G. E. Hutchinson cited instances of apparently constant ratios of the dimensions of trophic structures among sympatric congeners. MacArthur and R. Levins elaborated on this theme, developing a theory of limiting similarity that concerned the expected degree of niche differentiation among sets of potentially competing species, as part of a theory of the evolution of community structure that has had lasting influence. The importance of competition and of coevolution of competitors was the subject of vigorous debate in the 1970s and 1980s.

The most detailed recent models of coevolution of competitors (Taper and Case, 1992) assume that the size of a phenotypic feature, such as body size or beak depth, is correlated with the mean food type (e.g., prey size) that an individual uses. The evolution of the character, and thus of resource use, is modeled using quantitative genetic models for polygenic traits. (The character is approximately normally distributed, with specified additive genetic variance and heritability.) It is assumed that "a jack of all trades is master of none," that is, that there exist trade-offs in adaptation so that each phenotype performs effectively over only a relatively small range of resources (its within-phenotype niche width, WPNW), and that a specialized phenotype (with smaller WPNW) is more efficient than a generalist. (For example, studies of finches have shown that the rate of energy intake from smaller versus larger seeds is greater for small-billed versus large-billed birds, respectively.) In the simpler models, the competitive effects of phenotypes on each other are symmetrical and are proportional to the degree of resource overlap, whether the phenotypes be the same or different species. Thus competition is more intense when two phenotypes are more similar.

The size–frequency distribution of resources is assumed to be unimodal. Thus the rarity of resources in the tails of the distribution selects against extreme phenotypes. Consequently, the mean phenotype of a solitary species evolves to match the peak of the resource curve. Frequency-dependent selection maintains variant phenotypes, each adapted to a slightly different resource. However, the higher WPNW is, the less genetic variation is maintained. Because of recombination among the many loci affecting the trait, the frequency distribution of phenotypes remains approximately normal. It does not necessarily fit the frequency distribution of resources, some of which may be underutilized.

When applied to two or more species, these models generally support earlier suppositions. In each species, phenotypes that do not bear the combined burden of interspecific and intraspecific competition have higher fitness. Hence, the species evolve so that their phenotypic means are spaced apart (character displacement). Moreover, the phenotypic variance within each species at equilibrium is somewhat smaller than in a solitary species. At equilibrium, some overlap remains in the species' resource utilization, for otherwise abundant food items of intermediate size would be left unused, creating selection for genotypes that can use them. Under some conditions, especially if two species initially overlap greatly in resource use, they may converge. Competitive exclusion, resulting in extinction, then becomes more likely. Moreover, if competitive effects are asymmetrical (e.g., if larger individuals reduce the fitness of smaller ones more than the converse), a species may converge toward another and "chase" it to extinction.

Facile invocation of evolutionary responses to competition to explain ecological patterns was severely criticized in the 1980s, resulting in more critical evaluation of evidence. Even under the new scrutiny, many data strongly support the theory of coevolution of competitors. For example, character displacement has been documented in many instances, such as sticklebacks in British Columbian postglacial lakes. Some lakes harbor two species, recently derived from a common ancestor, that differ in microhabitat and in morphological features associated with feeding. Other lakes have only one species, with intermediate behavior and morphology. Closely related sympatric species often differ more consistently than would be expected by chance. For instance, body size in bird-eating hawks (*Accipiter*) is correlated with average prey size, and throughout the world sympatric species of these hawks differ more in size than do random pairs of *Accipiter* species.

Such nonrandomness of sympatric assemblages might arise from either coevolution or ecological assembly (colonization followed by extinction of excessively similar species). Phylogenetic analyses of anoline lizards in the West Indies have provided evidence of both processes. Many islands in the Lesser Antilles have both a large and a small species. These form distinct clades, so there is no evidence that the size of any one of these species have evolved in response to a sympatric congener. But each of the Greater Antilles has a monophyletic group of species that differ ecologically and morphologically from each other in a parallel pattern that is almost the same on each island. For example, each island has at least one species of "trunk anole," one "twig anole," and one "crown giant." Coevolution has taken much the same historical course in each case.

Release from competition is thought to have important evolutionary consequences. "Ecological release" at a microevolutionary level is illustrated by cases in which only a single member of a genus occurs in a region, and there occupies a broader ecological niche than do species that coexist with congeners elsewhere. The single species of finch on isolated Cocos Island, northeast of the Galápagos archipelago, feeds on a much wider variety of items than does any of its many Galápagos relatives, and the sexual dimorphism in beak size and forging mode is greater in the Hispaniolan woodpecker than in its continental relatives that are sympatric with other woodpecker species. At a macroevolutionary level, the spectacular diversification of African lake cichlids in the virtual absence of other fishes, the explosive diversification of mammals after the K-T mass extinction of large reptiles, the flowering of modern turtles after the extinction of primitive amphichelydian turtles, and many other examples are thought to show how diversification can be released when incumbent competition is alleviated.

Responses to Common Predators

Escape Space and Divergent Defenses

Competition is not the only possible explanation of resource partitioning and community assembly. For instance, species may interact indirectly via a shared predator. Predator density is highest where prey species coexist, and consequently a particular prey species will be maintained at lower density if it coexists with other prey species than if it is the sole victim. This form of predator-mediated *apparent competition* creates a pattern identical to that of direct competition. Coexistence is possible if prey species partition the resource they compete for, in this case an ecological space of reduced predation, known as an *escape or enemy-free space*.

Predators tend to increase their search intensity with increasing densities of prey. Search behavior can be improved by specialization in a foraging strategy, for instance, by attacking prey of a certain shape (i.e., similar search image) or prey that occupy particular sites, regardless of whether they belong to the same or different species. If faced with a diversity of prey patterns, they are most likely to form a search image for the most common pattern. This has been demonstrated for many bird and mammalian predators that use primarily visual cues. For example, if wild thrushes are presented with different combinations of two morphs of the snail *Cepaea nemoralis*, they will invariably preferentially attack the most common morph. The rarer individuals that differ from the ecological characteristics that the predator is attuned to will consequently experience lower predation rates. Such frequency-dependent selection within species, or *apostatic selection*, could lead to the divergence among prey species and an eventual increase in the diversity of forms and coloration patterns, that is, *aspect diversity*.

Evidence for this process is mostly circumstantial. R. E. Ricklefs and K. O'Rourke compared morphology, color, and behavioral characteristics of the moth species of a tropical and two temperate communities. Despite a higher number of moth species in the tropical site, the average similarity among moths within a community was the same in all three sites. They suggested that predation influences competition among moths for "escape space," limiting the similarity of appearance in different moth communities. *Cepaea nemoralis* and *C. hortensis* are both polymorphic for shell coloration pattern, and they share similar patterns. Bryan Clarke studied mixed colonies of *Cepaea* and found a negative correlation between the frequencies of visually similar shell patterns in the two species.

Mimicry

Another form of coevolution among prey species subject to common predators is the coevolution of mimicry. Two

common forms of mimicry are typically distinguished. Batesian mimicry refers to the convergence of palatable mimic species on distasteful models. Predators learn to avoid certain prey shape and color patterns they experienced as distasteful and mimics of such patterns can profit from this aversion. Monarch butterfly larvae, *Danaus plexippus*, feed almost exclusively on milkweed, from which they sequester cardiac glycosides. These toxic compounds are retained in the adult and vertebrate predators quickly learn to avoid both monarch adults and the more palatable mimetic viceroy butterfly, *Limenitis archippus*. Cleaner fish provide a variation on Batesian mimicry. In coral reefs in the Pacific, many fish allow cleaner fish, such as the sea swallow (*Labroides dimidiatus*), to feed on parasites on their bodies and even in the interior of their mouths. The sabre-toothed blenny (*Aspidontus taeniatus*) mimics the white-and-black-striped coloration and swimming pattern of *Labroides*. By taking advantage of the passive behavior of fish toward the model, it is able to approach fish and bite off pieces of tissue. *Labroides* and *Aspidontus* show parallel variation in color patterns across different geographic areas, which strongly suggests that indeed the mimic is converging on the model. It is a matter of discussion, however, whether mimics will instill evolution in the model, which might be expected to evolve differences that lessen the resemblance.

Müllerian mimicry refers to the convergence toward a similar pattern among unpalatable species. Faced with several undesirable species that look alike, a predator must learn a lower number of patterns to avoid. Evolution in all prey species leads toward a common pattern, and so warrants the designation of coevolution. One of the most striking cases of Müllerian mimicry, mentioned earlier, is the convergence between the neotropical butterflies *Heliconius erato* and *H. melpomene*. Despite differences in life history, these species share a common wing color pattern that varies geographically in parallel. One of the species, *H. erato*, is usually the most abundant where both species co-occur, raising the possibility that parallel evolution occurred by mere convergence of the rarer *H. melpomene* toward a common model. However, comparison between sympatric and allopatric populations of *H. erato* in Central America revealed that the width of the *H. erato* yellow hind-wing bar converges upon that of *H. melpomene* when in sympatry, suggesting that both species converge on each other.

Enemy–Victim Interactions

The general heading of enemy–victim interaction includes a variety of antagonistic interactions, ranging from those in which the enemy (or predator) usually kills several individuals (prey) to sustain itself, to scenarios in which the enemy (or parasite) restricts its negative effects to one victim (host) during most of its lifetime. The interaction between plants and herbivores is multiform, depending strongly on the herbivore taxon, and does not conform neatly to the previous division. Herbivores can be more akin to predators or parasites depending on whether they consume several individuals, as do seed harvesters, or complete their life cycle on a single plant, as do many leaf-feeding insects. They may also be quite unlike either category, as are many grazers and browsers that feed on parts of many individual plants without necessarily killing any of them. Differences in the degree of specialization of the enemy and victim may enhance the asymmetry of the interaction. Most plants, for instance, are attacked by numerous herbivores and parasites, yet many of their enemies are relatively specialized on particular plant taxa. These distinctions and characterizations are relevant as the nature of the interaction determines the selection pressures imposed on both species, and the resulting coevolutionary process.

Enemy–victim coevolution has been envisioned as an arms race in which the exploiter evolves offensives that increase the strength (i.e., frequency, intensity) of the interaction, and the victim evolves counterdefenses to decrease it. In predator–prey interactions, for instance, one would focus on traits that in some way affect the predator's functional response, that is, its rate of predation. This includes everything from the predator's detection ability and pursuit speed to the prey's aposematic or cryptic coloration and escape speed. An *evolutionary arms race* would correspond, for instance, to a continuous improvement of a prey's escape speed and the predator's pursuit speed.

Clearly there are limits to such continuous escalation. There are physical limitations to improving an organism's features, for example, to how fast an animal can run. Furthermore, investment in an interaction-related trait may imply costs. Development of the trait may require resources that could otherwise be invested in reproduction and other functions. For instance, elaborate morphological defenses (e.g., a thick shell or thorns) may imply an *allocation cost* expressed as a slower growth rate. Chemical defenses can act as feeding deterrents, but at high concentrations they may be autotoxic, implying a *physiological cost*. Additionally, some chemicals that act as deterrents toward generalists attract specialist enemies. *Ecological costs* produced by such genetic correlations limit a victim's potential to adapt to its complete array of enemies. Traits tend to affect more than one function of the organism, and a compromise must be reached between its role in the different functions. For instance, Geerat Vermeij suggests that such an adaptational dilemma could have been responsible for the extinction of most ammonite cephalopods. The external shell of cephalopods, other than providing passive protection, also affects the speed of locomotion and the ability to compensate for changes in pressure during vertical movements. Toward the end of the Cretaceous, as predation and competition became more intense, no further improvement in these functions could be reconciled by evolution of the shell.

Theoretical work has shown that the existence of costs and density dependence (the effect of a species' density on fitness of an individual of that species or of an interacting species) can hamper the occurrence of continuous reciprocal responses implicit in an arms race. Neither the prey nor the predator population is expected to always increase its investment in predation-related traits as a response to an increased investment by its partner. On the one hand, depending on the shape of the predator's cost–benefit function, the benefits reaped by an individual predator by investing more in predation can be insufficient to compensate for any additional costs. The predator's capture rate may remain unaltered if the evolution of greater prey elusiveness is accompanied by an increase in prey density. On the other hand, models have shown that the evolution of greater prey-capture abilities can lead to a decrease in the predator's equilibrium population size as a result

of overexploitation of prey (Abrams, 1990). A smaller predator population can imply reduced selection for antipredator traits, and if the costs of defense become a more important selective factor than predation, the prey population may evolve a lower level of defense. So rather than a protracted arms race, one might expect enemy and victim to coevolve toward an intermediate stable state.

The models referred to previously focused on character values and population dynamic parameters. A different class of models focuses on the genetics that underlie interaction-related characters, while usually ignoring changes in population densities. For example, gene-for-gene models of host–parasite interactions examine changes in genotypic frequencies at a virulence locus in the parasite and at a resistance locus in the host. Coevolutionary cycles are expected under a broad range of conditions. As the frequency of a virulent genotype increases, selection for resistant host genotypes becomes more intense. Frequency-dependent selection will generate fluctuations in genotypic frequencies, and both populations will remain polymorphic. Some polygenic models also produce stability, as either stable limit cycles or equilibrium points, particularly if the additive genetic variances of the characters are high in both interacting populations. Stability is most likely if the victim has a larger genetic variance and is under stronger stabilizing selection than the enemy is. A higher additive genetic variance allows the prey to respond to selection more rapidly than the predator.

The "Red Queen" model extends the notion of evolutionary arms race to the level of a community. According to these models, evolution of co-occurring species can lead to the continuous deterioration of a species' environment, forcing it to constantly evolve just to avoid extinction. The name owes its origin to the character in Lewis Carroll's *Through the Looking Glass* who explained to Alice that one must run as fast as possible just to stay in place. Analogously, a community at evolutionary equilibrium determined by the Red Queen principle contains a set of interacting species continually coevolving at rates that exactly balance each other. It has been hypothesized that Red Queen coevolution has selected for recombination and sexual reproduction, as only those populations with available genetic variation could have sustained continuous coevolution for extended periods.

The most plausible models of enemy–victim coevolution predict either evolution to an equilibrium or an oscillatory coevolutionary "chase" (rather than indefinite escalation of offensive and defensive properties). The clearest evidence for these models' predictions should come from direct observation of long-term coevolution, which to date has been possible only with laboratory cultures of organisms with very short generations, such as bacteria and virulent bacteriophage. The most extensive coevolution observed in such experiments consisted of fixation of a resistance mutation in the bacterial population, then of a countervailing "virulence" mutation in the phage, and finally of a second resistance mutation in the bacteria, resulting in stable coexistence and an apparent genetic equilibrium. The resistance mutations of the bacteria carried costs that reduced fitness in the absence of phage, but evolution at modifier loci later reduced the cost. The failure of the phage population to coevolve greater virulence was attributed to architectural constraints on these very simple organisms.

Although data from natural populations reveal little about the long-term dynamics of coevolution, they do show that it occurs. Perhaps the most famous instance is the evolution of greater resistance of rabbits to myxoma virus, released as a rabbit-control measure in Australia. This was coupled with evolution of somewhat lower virulence in the virus (see earlier discussion). Comparisons among geographic populations and species also provide evidence of coevolution. Parasites are commonly better adapted to their local host population than to other populations, as in the case of a microsporidian that infects the cladoceran *Daphnia magna*. Brood-parasitic cuckoos are polymorphic for egg color: females generally specialize on one or another species of host, and lay eggs that mimic those of the favored host. Some species of birds, especially those most frequently parasitized, reject foreign eggs if they can detect them. In some such species, rejection behavior is stronger in populations that experience parasitism by cuckoos than in those that do not. Thus, both the cuckoo and some of its hosts have evolved responses to the interaction.

Geographic populations of interacting species sometimes vary in parallel, although the causes of variation can be difficult to identify. For example, the parsnip webworm (*Depressaria pastinacella*), the sole specialized herbivore that has accompanied wild parsnip (*Pastinaca sativa*) from Europe to North America, exerts selection on the plant's profile of several toxic furanocoumarins. Variation in furanocoumarin profile among populations of the plant is paralleled by variation in the differential capacity to metabolize these compounds in the associated populations of webworm. Although geographic variation in the insect is surely an adaptation to local plant characters, it is harder to show that variation among plant populations has evolved in response to selection by the insect.

Examples of adaptations of predators to prey or of the reverse abound. Yet in few instances has it been demonstrated that these adaptations have evolved in response to any one species. Many characteristics have a similar effect on several or many species of antagonists, and probably represent diffuse coevolution. For instance, the "Mesozoic marine revolution," in which the evolution of diverse shell-crushing crustaceans and fishes was mirrored by the evolution of spines, thicker shells, and other defensive features in many lineages of molluscs, doubtless represents diffuse coevolution. It is difficult in such cases to show that an evolutionary change in any one species stimulated coevolutionary change in another.

Special Features in Parasite–Host Coevolution

Several topics loom large in the study of parasites and hosts that seem to play a lesser role in coevolution of animal predators and their prey.

Gene-for-Gene Systems

Whereas features that mediate predator–prey interactions, such as size, trophic structures, and cryptic coloration, are usually polygenic (quantitative) characters, specific loci for resistance have been identified in several species of plants. For each such resistance gene, a corresponding locus in a fungal

pathogen confers "virulence," meaning here the ability to attack and develop on a host with a specific resistance allele. (Most well-studied gene-for-gene systems involve crop plants, which has led some authors to suggest that resistance genes with large effects may be a result of the methods used to breed for resistance.) In the absence of pleiotropic costs, resistance alleles at all loci would be fixed in the host population and virulence alleles would be fixed in the parasite. If these alleles do have costs, then frequency-dependent selection can result in long-term fluctuations in allele frequencies, so that each species remains polymorphic and is continually adapting to the changing genetic constitution of the other. (This would be a clear instance of the Red Queen principle.) Populations of Australian wild flax differ in the frequencies of phenotypes resistant to various strains of flax rust, and associated rust populations likewise vary in virulence. The variation among populations may be an effect of Red Queen dynamics.

Selection for Sex

The supposition that parasite populations continually evolve has suggested to many authors that a major advantage of recombination and sexual reproduction may reside in the ability to generate novel, parasite-resistant genotypes. In a species of freshwater snail, for example, sexually reproducing females are more frequent than parthenogenetic females in habitats where the risk of infection by trematodes is greatest. Furthermore, among the several theories of sexual selection by female choice is the proposal that it is advantageous for females to choose males whose elaborate ornaments and behavior indicate that they are not debilitated by parasites, and thus are likely to father parasite-resistant offspring. Evidence from a few studies has been interpreted as support for this hypothesis, illustrating the broad ramifications that coevolution may have.

Evolution of Virulence and Avirulence

Often, though by no means always, parasites have less virulent effects on their normal host species than on novel, recently invaded species of hosts. This pattern might result from (1) evolution of resistance in a host species, (2) extinction of highly virulent species of parasites ("species selection"), because they drive their hosts to extinction, and/or (3) evolution within parasite species toward lower virulence. The evolutionarily naive often do not distinguish between the latter two causes, and imagine that parasites evolve avirulence because this will safeguard the host population and thus the perpetuation of the parasite species. However, natural selection is not prescient, and selection within populations can easily lead to fixation of "selfish" genotypes that enhance the risk of extinction. Evolution of the level of virulence depends on many factors.

For a parasite such as a virus, bacterium, or protozoan that reproduces on or in its host, each individual host carries a temporary group of parasites. Consider first a case in which each host carries only a single, clonally propagating parasite genotype. The fitness of this genotype is measured by the number of uninfected hosts it infects, relative to other genotypes. This number is often proportional to the number of offspring produced within the host and capable of transmission. But greater numbers of parasites take a greater toll on a host, resulting in greater virulence. The fitness of a parasite genotype is thus set by a balance between the rate of reproduction within the host and the reduction in transmission due to death or debilitation of the individual host. If a parasite population evolves less than maximal virulence, it is only because this enhances transmission of progeny to new individual hosts, not because it preserves the host population for future use.

The optimal level of virulence depends partly on the mode of transmission, especially whether it is mostly vertical (from parent host to offspring) or horizontal (as in sexually transmitted parasites). The transmission rate of a vertically transmitted parasite is directly proportional to the number of offspring of its carrier, so such parasites should be relatively benign. However, a horizontally transmitted parasite, given opportunity for transmission, does not profit from its host's further survival or reproduction. Many such parasites kill their hosts (e.g., baculoviruses of insects) or castrate them (e.g., some plant-pathogenic fungi).

The evolution of virulence is still more complex if individual hosts are typically coinfected by multiple genotypes of the parasite. *Within* each host, the genotype with highest growth rate r (and highest virulence) has an advantage, but if individual hosts bearing this genotype die before transmitting it, groups of parasites dominated by less virulent genotypes will transmit the most offspring. Thus group selection favors lower virulence, and individual selection favors higher virulence. At equilibrium, the level of virulence is likely to be the higher than in the case of singly infecting parasites. Comparison of the virulence of species of nematodes that infect fig wasps either singly or multiply has supported this theory, and phage in laboratory cultures evolved lower virulence if they were vertically than if they were horizontally transmitted (Bull, 1994).

Special Features in Herbivore–Plant Coevolution

Resistance, Tolerance, and Overcompensation

Herbivore impact on plant fitness can be intense, imposing selection for traits that reduce the effects of herbivory. One class of traits simply reduces the incidence of herbivory by escaping herbivores in space and/or time, that is, by decreasing the plant's "apparency." The idea is that abundant, long-lived plants in communities of low diversity will be more easily tracked down and fed upon by herbivores than emphemeral plants that establish short-lived populations in more diverse communities. Size of the plant may also be a factor, as small plants with simple architecture generally have lower insect diversity, perhaps by offering little protection to insects. Consistent with this hypothesis is the observation that among Capparales in Morocco, less damage by pierid butterfly larvae is incurred by species that fluctuate more in density. This suggests that greater constancy in population size and density, that is, greater predictability and apparency, increases the likelihood of herbivory. Fluctuations in density could, however, be a result of adaptation of germination to specific

environmental cues or other factors rather than pressures from herbivory.

Most plants must deal with the presence of herbivores, and possess a variety of mechanisms that either reduce herbivory (*resistance*) or ameliorate its effects (*tolerance*). The ability to tolerate damage, that is, to suffer a reduced impact of damage on fitness, depends both on extrinsic factors, such as resource availability, the timing of herbivory, and the parts affected, and on characteristics of the plant. Thus it is thought that plants with intrinsically higher growth rates, storage capacity, or higher number of active meristems will have greater opportunity for regrowth in face of biomass loss to herbivory. It is as yet unclear whether these traits are adaptive defenses or whether they evolved because of other selective forces, such as competition and resource assimilation.

Some experiments have suggested that under ideal conditions, damage at certain phenological stages and the presence of abundant resources could render plants more fit than had they not incurred damage, that is, they would overcompensate. For example, artificially damaged scarlet gilia (*Ipomopsis aggregata*) produced approximately twice as many flowers and fruits as undamaged plants. The root biomass of clipped plants was greater than that of controls, suggesting that by increasing root size plants were able to take up the extra nutrients necessary for the overcompensation response. This implies that herbivores would benefit plants by feeding on them. There is controversy over the design and results of these experiments, and overcompensation has not been observed in enough species to be considered a general phenomenon.

Greater attention has been given to plant resistance, namely, to mechanisms that actively repel herbivores or that make the plant unpalatable or undesirable. Resistance mechanisms differ with respect to when they are expressed: many defenses are constitutive, that is, expressed regardless of herbivore presence, whereas others are absent until induced in the presence of herbivores. Induced responses curtail the costs associated with defenses, as these are incurred only when strictly necessary.

Some plants have morphological defenses, such as hard seed coats, thorns, or stinging hairs. However, a central role in defense has been attributed to secondary metabolites. These chemicals play no part in primary metabolism and groups of compounds are usually restricted to a few plant families. They have been found to affect herbivore behavior by decreasing a herbivore's feeding rate, deterring oviposition, slowing development, and reducing fecundity. Yet many of these compounds have also been found to affect bacteria, fungi, viruses, other plants, and other herbivores. Secondary compounds might just as well be an adaptive defense to these organisms, or even serve other ecological or physiological functions (e.g., UV protection, drought tolerance). Convincing evidence that insects are a driving force in the evolution of plant chemical defenses comes from studies that compare the pattern of natural selection on putative resistance factors in the presence and absence of herbivores. For example, in the presence of herbivores a population of *Arabidopsis thaliana* was shown to be under selection for increased levels of mustard oils (or glucosinolates) and trichome densities.

In contrast to resistance, plant tolerance does not affect the performance or fitness of the herbivore. Herbivores will not be under selection pressure to overcome this plant defense and so coevolution should not be expected, unless resistance and tolerance are correlated. Theoretical models do indeed predict a trade-off between these two forms of defense, but the limited evidence so far is inconclusive.

Specialist and Generalist Herbivores

Most herbivorous vertebrates and many insects are generalists, feeding on plants in many families. A more specialized diet, of plants in only a single family, genus, or even species, characterizes the majority of herbivorous insects and a few vertebrates (such as the koala, a specialist on *Eucalyptus* foliage). The advantages of a broad diet seem clear: an individual can feed despite spatial or temporal variation in the availability of any one food type. Theoretically, the possible advantages of specialization are clear, but evidence is often more ambiguous.

Commonly, closely related species of insects specialize on different plants. Assuming their common ancestor fed on one or more of these plants, the problem is to explain why some of its descendants, instead of expanding their diet, abandoned their ancestral host to which they were presumably well adapted in favor of new hosts. Several ecological advantages of such a switch have been suggested. The new host might provide escape from predators or parasites. Certainly specialization is maintained, if not originally caused, by the protection that many insects, such as milkweed-feeding monarch butterflies, gain by sequestering defensive plant compounds. Competition for resources, as we have noted earlier, may also select for specialization. In many species, mating occurs only on the host plant, and use of a single host may be the most reliable way to find mates. Host specialization will then be a consequence of positive frequency-dependent selection. There is indirect evidence for each of these hypotheses.

The most general explanation for specialization is likely to be trade-offs in the ability to find, handle, digest, or detoxify different kinds of plants. The secondary compounds that characterize higher taxa of plants are thought to play an especially important role, since many are toxic and/or deter feeding by nonadapted insects. Conversely, host-specific insects typically use some such compounds as feeding or oviposition stimuli. The hypothesis that related specialist insects usually feed on related plants because of the plants' chemical similarity has been supported by a study of *Blepharida* flea beetles: the beetles' phylogeny is more closely mirrored by chemical similarity than by phylogenetic affinity among the species of their *Bursera* host plants.

Tests for trade-offs have had mixed results. There is little evidence that closely related generalist and specialist species differ in their efficiency of digestion of plant species that they both naturally feed on (Futuyma and Wasserman, 1981). There is some evidence that specialist species can find and recognize host plants faster than generalist species, owing to properties of the nervous system, and abundant evidence that closely related species are physiologically better adapted to their own than to each other's host plant species. However,

since such differences may have evolved after these species became specialists, most research on trade-offs has compared genotypes within species. Only a few such tests have found clear evidence of trade-offs; different clones of pea aphids, for example, have high fitness on either pea or alfalfa. In no case is the biochemical or physiological basis for such trade-offs fully understood.

Mutualism

In a mutualistic interaction, individuals of each of two species obtain a fitness benefit from the interaction between them. Long-term physical associations are referred to as symbioses. These include endosymbioses, in which microorganisms such as bacteria reside within their hosts. In some cases, endosymbionts have become, in effect, part of the host. Mitochondria and chloroplasts, for example, evolved from endosymbiotic bacteria early in eukaryote evolution.

Some mutualisms have doubtless evolved from casual, fortuitously beneficial interactions; such may be the origin, for instance, of the nectar glands of many plants that attract generalist ants, which attack herbivores and other intruders. Some such mutualisms have become obligate, as in the case of some *Pseudomyrmex* ants that inhabit and defend only certain species of *Acacia* trees that provide food bodies and special domiciles. In many cases, especially symbioses, mutualisms have arisen from parasitic interactions; for example, a complete spectrum from parasitism to mutualism is found in endophytic fungi (Clavicipitaceae). The conditions that favor evolution of parasitism into mutualism are much the same as those favoring the evolution of avirulence (see earlier). Specifically, the fitness of vertically transmitted symbionts depends on the fitness of their hosts, so a benefit provided to the host will enhance the symbiont's fitness. Even without vertical transmission, long-term (e.g., lifelong) association between individual partners satisfies the conditions for "reciprocal altruism." The survival and reproduction of a colony of *Acacia*-dwelling ants depends on the health of its host tree, so it pays to defend it, just as the fitness of the tree depends on nurturing its defending ant colony.

Many mutualisms are rather nonspecific, in that at least one of the species of partners holds a mutualistic relationship with several or many species. Most mycorrhizal fungi, for example, are not highly host-specific, and most species of pollinating or seed-dispersing animals visit many species of plants. There are, however, some exceptions, such as the fig wasps, each species of which develops as a seed predator of one of the several hundred species of fig and is the sole pollinator of that species. The vertically transmitted bacterial endosymbionts (*Buchnera*) of aphids are so intimately tied to their hosts that the phylogeny of the bacteria mirrors faithfully that of the aphids. Unique associations of lineages of aphids and bacteria have persisted for more than 100 million years.

Mutualism can generally be viewed as reciprocal exploitation. Mutualists do not provide benefits to their partners out of altruism; the benefit is either wrested from them (e.g., pollen consumed by pollinating bees) or it is a "payment" that ensures a reciprocal benefit. Thus potential conflict pervades most mutualisms, for selection often favors "cheater" genotypes that do not repay the benefit they receive. For instance, bees often rob nectar by chewing through the bases of flowers, so that they do not pollinate. Many orchids deceive pollinating insects, providing neither nectar nor consumable pollen, but instead produce scents that mimic the insect's sex pheromone and induce "pseudocopulation" by male insects that effect pollination. Some orchids produce no reward whatever. Selection for cheating can destabilize a mutualistic relationship. For example, yucca moths (*Tegeticula*) oviposit in the ovaries of *Yucca*, where the larvae feed on some of the developing seeds. Most species of *Tegeticula* stereotypically (and "carefully") place pollen on the stigmas of flowers in which they lay, and so both the host plant and the moth's offspring benefit. However, several lineages of "cheater" *Tegeticula* have evolved from pollinating ancestors; they do not pollinate, but lay eggs in flowers that have been pollinated by those species that do.

Major Consequences of Coevolution

If it is difficult to document and characterize coevolution between pairs or among small numbers of species, it is harder still to demonstrate the effects that coevolution has had on the properties of communities, ecosystems, or the history of biological diversity. For the most part, questions on so grand a scale have, so far, only tentative or even speculative answers.

Phenotypic Diversity

Although no quantitative data are available, coevolution, or at least adaptation to interspecific interactions, has unquestionably enhanced the phenotopic diversity ("disparity") among species. The diverse trophic structures displayed by related species of fishes, birds, and other animals partly reflect adaptation to the characteristics of the different arrays of prey species on which they feed, and partly arise from selection due to interspecific competition, resulting in divergence and specialization in resource use. The diverse morphologies in such adaptive radiations as the African lake cichlids illustrate that ecological and morphological diversification of a clade may be promoted by competition among its member species, but may be constrained by competition from members of other clades, until such competition is relieved.

Predation and parasitism have doubtless enhanced phenotypic diversity among prey. There is much circumstantial support for Ehrlich and Raven's hypothesis that chemical diversity among plant taxa has been driven by herbivores (and, we would add, pathogens). Intraspecific genetic diversity at major histocompatibility loci has almost certainly been maintained by selection imposed by pathogens, and the extraordinarily complex immune system of vertebrates is testimony to the potent selection that pathogens generate. The Mesozoic evolution of shell-crushing predators called forth an immense proliferation of defensive features in the shells of molluscs.

The cooperation and conflict inherent in mutualism have likewise had immense effects on phenotypic diversity. The mycetocytes of aphids, the light organs in some fishes and

cephalopods, and the root nodules of legumes are among the many special structures that house symbiotic bacteria. The phenotypic diversity of the Orchidaceae, the largest family of plants, lies largely in the extraordinary variety of flower forms that attract and manipulate diverse pollinating animals.

Although such examples provide undeniable evidence that biotic interactions have had great evolutionary consequences, phenotypic disparity has seldom been measured. Moreover, although the role of interactions in these examples is obvious, the importance of reciprocity (the "co" in coevolution) is not. Rigorous tests of the effects of coevolution remain to be done.

Species Diversity

Variation in species diversity among clades, communities, ecosystems, and biomes and through geologic time is affected by many factors. The relative importance of coevolution on rates of speciation and extinction, which together determine diversification rate, is unclear. Only in exceptional cases, such as the divergence of figure and their species-specific agaonid pollinators, is it likely that coevolution directly induces speciation, that is, reproductive isolation. It has often been suggested that animal-mediated pollination and seed dispersal may enhance rates of plant speciation by establishing reproductively viable isolated populations. However, species diversity of angiosperms is only slightly greater in animal- than wind-pollinated families, although it is greater in families that include both biotic and abiotic seed dispersal than in families that use one mode exclusively.

In principle, extinction rates can be increased or decreased by coevolution, although either is very difficult to document in practice. If coevolution has effected a net increase in diversification rate, it has probably done so chiefly in three ways: (1) by shifts in resource use, thus reducing competition; (2) by evolutionary escape from predation, followed by radiation; and (3) by evolution of specialized predators, parasites, and mutualists following the evolution of new prey or hosts. That is, diversification of one group of organisms begets diversification of others, as Ehrlich and Raven postulated for plants and herbivorous insects.

These hypotheses are supported by some modest evidence. Farrell *et al.* (1991) found that clades of plants with canals that deliver defensive latex or resin to sites of herbivore attack are consistently more diverse than their sister groups that lack these features. Similar sister-group comparisons among insects showed that the evolution of herbivory has consistently been followed by greater diversification. However, it is not known if the diversity of herbivorous insects has been enhanced in clades that have adapted to chemically novel, diverse plant clades, as Ehrlich and Raven proposed. In western Europe, the species richness of aphids and certain other insects that feed on different plant families is correlated with the number of species of plants in the family. Whether or not such a correlation holds on a global scale is not known.

Community Structure

The convergent evolution of ecologically "equivalent" species, such as many Australian marsupials and their placental counterparts in other regions, suggests that independently evolving communities might converge toward similar structure. Certainly, properties such as vegetative morphology and the architecture of vegetation converge; for example, "Mediterranean" vegetation (chaparral, matorral, maquis) is dominated by sclerophyllous, small-leaved shrubs in several parts of the world. Such similarities, however, arise simply from each species' independent adaptation to physical environmental factors. Whether or not community properties such as stability, food-web structure, or species diversity converge as a result of coevolution is a different, much more difficult question. Simple models of interspecific competition and of food webs suggest that the coevolutionarily stable equilibria might be rather few, so that coevolution might yield some predictable structure (MacArthur and Levins, 1964). Whatever the ideal applicability of these models might be, however, the opportunity for such pervasive effects of coevolution may generally be rather slight. Gene flow among conspecific populations that interact with different ensembles of species may prevent finely tuned coevolution. Moreover, paleoecological studies have shown that throughout the Pleistocene, species have had highly individual histories of change in geographic distribution, so that many of today's assemblages of species are very recent. Except for specialized associations, as of host-specific parasites that have moved about with their hosts, there has been little time for coevolutionary adjustments in many of today's species assemblages. Nevertheless, paleontologists have documented rather steady levels of diversity at both global and local levels over vast periods of time (10^7–10^8 years), despite turnover of taxa (Jablonski and Sepkoski, 1996). These observations suggest that whatever convergence or constancy of community structure exists may be attributable more to purely ecological rules of community assembly rather than to coevolution.

An unusually clear example of convergent multispecies assemblages is provided by the anoline lizards mentioned earlier, which form monophyletic groups of morphologically and ecologically equivalent species on each of the Greater Antilles islands. However, similar habitats in different parts of the world generally are not very similar in species richness. For example, lizard diversity in Australia exceeds that in corresponding habitats in southern Africa. Nevertheless, in some instances variation in species richness among habitats shows similar patterns; lizards are more diverse in deserts than in wetlands in both Australia and Africa. Thus, habitats seem to have consistent effects on species coexistence, and perhaps on the evolution of resource partitioning (Ricklefs and Schluter, 1993).

Coevolution may also affect community properties such as stability, that is, the property of returning to equilibrium after a disturbance. The study of artificially constructed food webs has identified some patterns of interaction among component species that are conducive to stability. One such aspect is connectance, that is, the ratio of actual to potential interspecific links. High levels of connectance are associated with reduced stability, as a disturbance to one species will affect many other species in the community. Thus processes that reduce connectance, or generate compartments within a community, such as a subset of species that interact predominantly with one another, contribute to stability. To the

extent that coevolution between parasites and hosts and between plants and herbivores leads to increased levels of specialization (and thus compartmentalization), it can thus indirectly enhance community stability.

See also: Competition, Interspecific. Darwin, Charles (Darwinism). Evolution, Theory of. Parasitism. Plant–Animal Interactions. Recombination. Species Coexistence

References

Abrams PA (1990) The evolution of anti-predator traits in prey in response to evolutionary change in predators. *Oikos* 59: 147–156.

Bull JJ (1994) Perspective—Virulence. *Evolution* 48: 1423–1437.

Farrell BD, Dussourd D, and Mitter C (1991) Escalation of plant defenses: Do latex and resin canals spur plant diversification? *Am. Nat.* 138: 881–900.

Futuyma DJ and Slatkin M (1983) *Coevolution*. Sunderland, Massachusetts: Sinauer Associates.

Futuyma DJ and Wasserman SS (1981) Food plant specialization and feeding efficiency in the tent caterpillars *Malacosoma disstria* and *Malacosoma americanum*. *Entomol. Exp. Appl.* 30: 106–110.

Jablonski D and Sepkoski JJ (1996) Paleobiology, community ecology, and scales of ecological pattern. *Ecology* 77: 1367–1378.

MacArthur RH and Levins R (1964) Competition, habitat selection and character displacement in a patchy environment. *Proc. Nat. Acad. Sci.* 51: 1207–1210.

Ricklefs RE and Schluter D (eds.) (1993) *Species Diversity in Ecological Communities*. Chicago: University of Chicago Press.

Taper ML and Case TJ (1992) Coevolution among competitors. *Oxford Surveys in Evol. Biol.* 8: 63–109.

Thompson JN (1994) *The Coevolutionary Process*. Chicago: University of Chicago Press.

Vermeij GJ (1987) *Evolution and Escalation: An Ecological History of Life*. Princeton, New Jersey: Princeton University Press.

Commons, Concept and Theory of

Colin W Clark, University of British Columbia, Vancouver, BC, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bionomic equilibrium In a biological common property resource, an equilibrium between resource production rate and exploitation rate, characterized by the dissipation of economic rents.

Commons (or common property resource) Any resource asset that is not privately owned and controlled.

Discount rate Rate (usually annual) at which future revenues are discounted to calculate a present value. Personal discount rates often exceed market rates by a wide margin.

Economic rent The flow of net economic benefits (revenue minus costs) derived from the exploitation of a resource asset.

Externality Cost or benefit imposed on others as the result of some economic activity.

Pareto efficiency An economic equilibrium is Pareto efficient if any change will decrease the economic well-being of at least one participant.

Tragedy of the commons Process whereby a commons is overexploited because individual users cannot expect to realize the potential benefits of resource conservation.

In medieval England the term "commons" referred to an unfenced area of land that, although privately owned (usually by the local manor lord), was by tradition available to specific persons (the commoners) for specific activities, such as pasturage or removal of building materials. The subsequent enclosures of the commons, for the sole benefit of their owners, reduced many former users to penury. In recent times, and especially following the publication of the famous article "The Tragedy of the Commons" (Hardin, 1968), the term has been used to refer to any resource asset that is not privately owned and controlled. This is the sense in which the term commons, synonymous with "common property resource," will be used here.

Introduction

Important contemporary examples of common property resources include the global atmosphere, the oceans, large lakes, rivers, forests, and fish and wildlife populations, including birds. Though not inevitable, the overexploitation of common property resources is always a potential threat, and often a frightening reality. Many current environmental problems can be traced to the working of a so-called tragedy of the commons.

The classification of resources as either common or private property is an oversimplification. Many gradations exist between a total lack of access restriction (i.e., open-access) and complete individual control. For example, coastal marine resources within 200-mile fishing zones are now recognized as being under the sole jurisdiction of the coastal state, with foreign fishermen either excluded or subject to payment of fees. These resources are thus limited access, but (unless privatized) still common property by our definition. Also, private landowners seldom have full rights to do as they wish with their property – the state may retain mineral rights, for example, and zoning regulations may restrict land use and development.

The logic of the tragedy of the commons is straightforward, as will be explained later. Unfortunately, however, much of what has been written about common property resources suffers from vagueness and imprecision, leading to unnecessary confusion and controversy. The following hypothetical example can be useful as a basis for a general understanding of the economics of a commons.

Imagine a small lake, completely surrounded by the properties of two landowners. The lake contains a population of fish, and both landowners are avid fishermen. What will be the long-term outcome? The answer is not obvious and depends on various factors.

The first possibility is that the lake is so large, and the fish population so productive, that no amount of fishing by the two landowners has any appreciable effect on the population. In this ideal situation, by assumption nothing happens – the tragedy of the commons does not occur.

The opposite case is more instructive. Suppose, for example, that upon first taking up their properties, the two neighbors are able to catch all the fish they want. After a few years, however, they observe that fish are becoming more scarce. Also, the big ones that used to put up such a fight seem to have disappeared. Subsequent developments now depend on how the two landowners react. If they meet on amicable terms and agree to restrict their catches, or to return some of them to the lake, the dynamics of the exploited fish population may stabilize, permitting sustained if limited catches by both owners. It may take several years of trial and error before the maximum sustainable catch is determined – or they may decide to forego the maximum catch in favor of maintaining a large population of large fish.

An alternative solution might be to subdivide the lake into two fishing zones, one for each landowner. But this assumes cordial relations between the landowners, with each cooperating in not fishing in the other's zone. If fish are able to swim freely between zones, such an agreement may fail to prevent overfishing. Privatization of property or resources is not always feasible, and when it is feasible it presumes mutual trust and cooperation, or external enforcement.

Instead of such amicable agreements, however, the owners may begin blaming each other for overfishing. Relations may

then deteriorate, with both determined to catch all the fish they can find. One may buy a faster boat, hoping to outdo his neighbor, who then responds in kind, and so on. Such behavior between two neighboring individuals may seem petulant. But what if instead of two property owners, there were 100 cottages on the lake, all owned by avid fishermen? Besides increasing the pressure on the fish population, this would make the problem of agreeing over and enforcing catch restrictions much more difficult. Cheating by individuals would be hard to detect, and even harder to control.

This simple example illustrates many of the characteristic features of common property resources, including potential overexploitation, the necessity of mutual cooperation, and the increased difficulties of control as the number of joint owners increases. The author returns to these questions later.

Theory

The theory of a common property renewable resource will be discussed here in terms of a specific example, a commercial fishery. The results and predictions apply, however, to most other biological resources, with minor modification.

The theory of bionomic equilibrium in an unregulated, open-access, common property fishery was initiated by Gordon (1954) (**Figure 1**). Fishermen exert fishing effort E on a certain fish population; fishing effort is a measure of fishing intensity, for example, the number of vessel-days per year, or (in a rod-and-reel fishery) the number of rod-hours per year. **Figure 1** shows the total sustained annual revenue R and total cost C as functions of effort E. Cost increases in proportion to E, but revenue rises to a peak ("maximum sustainable yield") and then declines at higher levels of effort. An input–output function having this inverted-U shape is unconventional in economic theory; it reflects the underlying biology of the resource, which becomes increasingly unproductive as it is depleted by high levels of fishing effort.

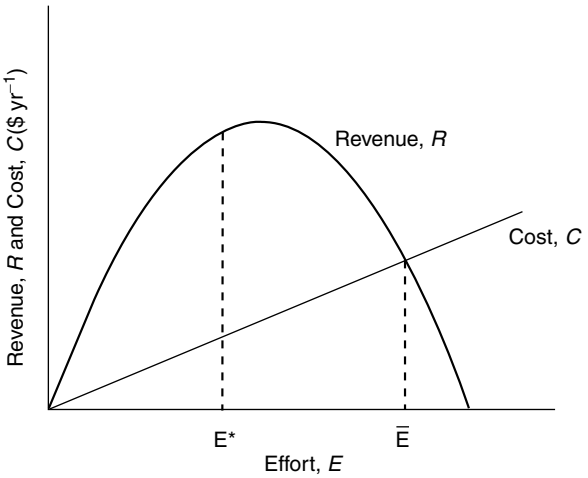


Figure 1 Gordon's (1954) model of equilibrium revenue R and cost C in a fishery; $\bar{E}$ = bionomic equilibrium; E^* = effort level for maximum sustainable economic rent. Reproduced from Gordon HS (1954) The economic theory of a common property resource: The fishery. *Journal of Political Economy* 62: 124–142.

Gordon's prediction is that the fishery will reach an equilibrium, called the bionomic equilibrium, at the effort level $\bar{E}$ where revenues and costs are equal. The argument is as follows. First, for $E < \bar{E}$, revenue exceeds costs and fishermen are making money in excess of their opportunity costs. This positive net revenue attracts additional entrants to the fishery, resulting in an increased effort level. On the other hand, if $E > \bar{E}$, fishermen are losing money and some will leave the fishery. Equilibrium can only occur where $E \equiv \bar{E}$. At this bionomic equilibrium the fishermen are barely breaking even financially. (See Clark 2010, for mathematical representation of these ideas.)

For levels of E below $\bar{E}$, revenues R exceed costs C; the excess R–C is called economic rent. Gordon's prediction can then be stated: in an open-access unregulated fishery all potential economic rents will be dissipated in long-run bionomic equilibrium. Additional predictions are that if the price of fish increases, or if the cost of fishing decreases, the effort level $\bar{E}$ will increase, resulting in a more depleted fish population. What would normally be recognized as benefits to the fishing industry (higher prices, lower costs) are dissipated because of the tragedy of the commons.

What is not clear in the Gordon model is whether the fish population will survive in the long run, or be driven ultimately to extinction. In practice, very few fish species have been fished to extinction, although some marine mammals have been. On land, America's most abundant bird species, the passenger pigeon, was hunted to extinction in the nineteenth century. Today species such as rhinoceroses, tigers, and certain birds are endangered by market hunters. If we agree that the Gordon theory applies to any such commons, then any species is likely to become endangered if the value of the last specimen caught exceeds the cost of capturing it. In a world of extremely affluent consumers and destitute hunters, this condition applies to an ever-widening class of wild species, both marine and terrestrial.

Gordon also identified the effort level E^* yielding maximum economic rent (see **Figure 1**). He pointed out the apparent anomaly that if the fishermen would only work less hard – reduce their effort – they would be better off! This argument may seem compelling, but the reality is that many, if not most, marine fisheries today remain closer to bionomic equilibrium than to rent maximization. It often seems as if the fishing industry is determined to suffer the tragedy of the commons even though impoverishment is the inevitable outcome. What explains this phenomenon?

To address this question we must consider the dynamics of the fish population (recall that **Figure 1** pertains only to equilibrium situations). For example, suppose that a certain fishery is currently at bionomic equilibrium $\bar{E}$. A reduction of effort is needed in order to improve the situation. Unfortunately, such a reduction of effort will not lead to an immediate increase in catches – quite the reverse. At first catches will decrease, more or less in proportion to the decreased effort. The decreased catches will (it is hoped) allow the fish stock to increase over time. Eventually a new equilibrium will be reached, combining reduced effort with increased catches and increased rents. A simulated illustration of this process is shown in **Figure 2**.

The principal message from the argument is that resource conservation necessarily involves a sacrifice of potential

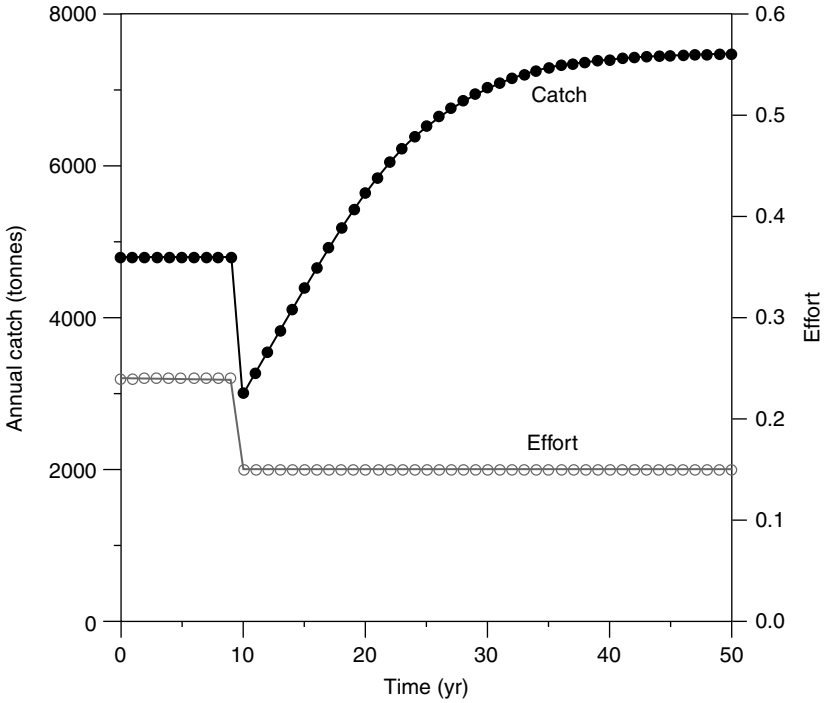


Figure 2 Simulated annual catches following a reduction in effort at time $t_0 = 10\ \text{yr}^{-1}$. At first catches are reduced, but eventually catches increase to a higher level than before, as the fish population recovers from overharvesting. (Discrete logistic model; $r = 0.3\ \text{yr}^{-1}$, $K = 100{,}000$ metric tons, $E_{t<t_0} = 0.24\ \text{yr}^{-1}$, $E_{t<t_0} = 0.15\ \text{yr}^{-1}$.)

current benefits in order to achieve greater future benefits. From the economic point of view, conservation is thus a form of investment in the future.

We also see now why management of a common property resource is likely to be difficult: the resource users must agree to make the necessary sacrifice (investment). Moreover, all resource users must then somehow be prevented from overexploiting the protected resource, even though such activities would be individually profitable. The more successful the management program, in terms of protecting the resource, the greater will be the incentive for excessive exploitation by individuals. This dilemma is the real tragedy of the commons.

Although formulated here in terms of fish populations, the concepts of bionomic equilibrium and rent dissipation in common property resources are widely applicable. Overgrazing of grasslands, desertification, deforestation, soil erosion, and pollution are other important examples of overexploitation of the commons. Everywhere that resource stocks are used as common property, the incentive for the commons tragedy exists.

The participants in a tragedy of the commons are sometimes said to be motivated by short-term financial considerations – with "everyone following their own short-term interest" to quote from a recent publication. This portrayal of such resource users as greedy, short-sighted souls is often both unfair and inaccurate (although there probably is a tendency for such individuals to gravitate toward unregulated resource industries). In a commons situation, it is in each participant's best interests to get as much as possible out of the resource before it is depleted by others. Any personal sacrifice for the sake of resource conservation is simply dissipated by the other users.

Regulated Bionomic Equilibrium

How can the problems of common property resource exploitation be overcome? Let us first consider an approach that often does not, and perhaps cannot, succeed. Again, imagine a fishery initially at bionomic equilibrium (this is not essential for the discussion). A government agency is set up to manage the fishery. Realizing that the problem is excessive effort in the fishery, the managers introduce a limited-entry program, in which only a fraction of the previous participants are licensed to catch fish. (It goes without saying that, unless compensated in some way, the excluded fishermen will tend to resist this program.) At first the new system may appear to be successful – effort decreases, the fish population begins to recover, catch rates go up; the remaining fishermen are making money.

But now the economic dynamics of the commons begin to operate once again. By increasing their own levels of effort, individual fishermen can increase their current income. Some of the fishermen buy bigger boats, forcing others to do the same in order to compete. The managers must now bring in regulations to restrict the size of fishing boats. The owners then respond by "capital stuffing"– installing larger engines, more powerful winches, better electronic devices, and so on, all of which increase the fishing power of their vessels. These items in turn must then also be regulated. Eventually either

the regulations become so cumbersome that economic rents are dissipated simply by the complexity of the management system, or else the regulations fail to prevent overexpansion of fishing capacity. In either case, the fishery is now in regulated bionomic equilibrium.

Such regulated bionomic equilibrium can nowadays be seen in many managed marine fisheries. Where the regulations are effective and strongly enforced, depletion of fish stocks may be prevented. In practice, however, the economics of the commons has often proved more powerful than the regulations. In a number of cases, intensively managed fisheries have collapsed from persistent overfishing, cod and groundfish on the Atlantic Coast of the United States and Canada being notorious examples. In most of these cases, the fishing industry itself persistently and adamantly opposed the imposition of the regulations that were required to preserve the fishery.

While all of this may seem quite irrational (unless you understand the economics of the commons), in actuality the situation is often more complicated – and more perverse – than described here. To be specific, resource industries are often highly subsidized. Individuals can only benefit from the subsidy if they participate in exploiting the resource. The effect of the subsidy is to reduce the opportunity cost *C*, in extreme cases to near zero. The Gordon theory explains what happens: bionomic equilibrium corresponds to virtual resource extinction in such a situation. Exploiters become dependent on the subsidy, not the resource. Bionomic equilibrium is more devastating with subsidies than without them.

In the case of Canada's Atlantic cod fishery, the system of regulated, subsidized bionomic equilibrium proceeded to the ultimate conclusion with the almost complete disappearance of cod stocks in 1991. This collapse destroyed the basis of the economy of the entire province of Newfoundland; since 1991 the Canadian government has been paying over $1 billion per year as welfare to displaced Newfoundland fishermen and plant workers. Whether the Newfoundland cod will ever return is uncertain at the present time (2010).

The management of fisheries that take place outside the 200-mile zones of any coastal state is even more difficult, for lack of jurisdiction. International agreements have sometimes succeeded in preventing overfishing, but the overall picture remains bleak for these truly common property resources.

Managing Incentives versus Managing Resources

Two general approaches to the management of common property resources can be identified: those directed toward management of the resource stock and those that alter the incentives of resource users. Methods aimed at managing the resource include total annual catch quotas, controlled access to the resource, restrictions on the technology of harvesting the resource, and so on. Methods that rely on altering the economic incentives of common property resource exploiters include user fees or royalties and individually allocated quotas (see the following). These methods may also be used in combination. As indicated by the foregoing discussion, management strategies that do not alter users' basic incentives tend to result in regulated bionomic equilibrium, with excessive inputs and persistent rent dissipation. Methods that alter incentives have the potential of overcoming these difficulties, although the persistent incentive to cheat must be recognized, and countered by monitoring and strict enforcement of the rules.

A third approach, privatization of the resource – usually involving geographical subdivision into individually manageable units – may be considered the ideal solution, since it allows users to become independent of each other and of excessive government interference. This is historically the usual method for dealing with nonfugitive resources, such as land and minerals. Such subdivision and privatization is not feasible, however, for many of today's remaining common property resources. In such cases, some form of centralized management seems necessary. We first consider the case in which the resource is located within the jurisdiction of a single state; the case of international common property resources is discussed in the section International Common Property Resources.

Allocated Quotas

One method of altering the incentives of resource users is a system of allocated quotas, often sold under auction to the highest bidders. This method is commonly used in logging in national forests and in offshore oil drilling.

Having obtained such a quota, the users' incentives are to harvest the quota at minimum cost. Persistent incentives toward cheating are thwarted by rigorous enforcement of the quotas. Resource rents are preserved, and divided between resource users and the government. The resource remains common (state) property, but specific user rights are privatized through the system of quotas. Multiple uses of the resource, as in national forests used for both timber production and recreational activities, can be achieved, although different user groups may attempt to influence management decisions in their favor.

The establishment of 200-mile national fishery zones in the 1980s rendered coastal marine resources subject to quota-based management systems, and individual catch quotas are indeed now in use in many such fisheries. Marine populations, however, provide unusual management difficulties, not the least of which is the fact that these resources are largely hidden from view until they are brought to the surface by net or hook. Also, fish populations undergo major natural fluctuations, which are poorly understood and hard to predict. In addition, many fish populations migrate, with migration routes that change in response to oceanic conditions. These features introduce large, irreducible levels of uncertainty into fisheries management; the author discusses this topic further in the section Risk and Uncertainty.

International Common Property Resources

As an important example of an international common property resource, let us briefly consider the atmosphere. Although partitioning of the atmosphere is clearly not possible, policies at the national level have had significant effects on

atmospheric pollutant emissions. Norway, Sweden, Denmark, France, and Japan have employed taxes on the emission of sulfur dioxide since the 1970s. Some countries now use tradable quotas to control atmospheric emissions. Such policies affect incentives, and are thought to be far more efficient than direct quantitative regulations.

At the international level, over 100 nations have now ratified the 1987 Montreal Protocol requiring the phasing out of ozone-depleting chlorofluorocarbons and related substances. This example shows that, when evidence of severe harm is strong, international cooperation can be achieved. However, the example also illustrates the difficulties, especially at the international level, associated with common property resource management, as noncompliance, failure to ratify, and the proliferation of illegal trade in ozone-depleting substances have persisted. Although no international agency exists to enforce the Protocol, trade sanctions (or the threat thereof) have been used to encourage adherence to the rules.

Similar agreements, and methods of encouraging nations to follow them, apply to other international resource problems, such as greenhouse gas emissions, marine resource exploitation, ocean pollution, and the conservation of migratory birds. Although international management may lag behind the need, it is encouraging that the general principles of managing the commons do appear to be having an effect on the way in which the rules are being written.

Discounting the Future

The view is sometimes expressed that privatizing a commons will automatically result in its conservation. Some economists have gone so far as to assert that the private owner of a renewable resource asset would never overexploit it, except perhaps as a result of miscalculation. This assertion is entirely false and is based on a failure to understand the economic dynamics of natural resource stocks.

To summarize the argument in a nutshell (see Clark, 2010 for details), the hypothetical owner of a renewable resource will calculate the total discounted present value of various options, including sustainable harvesting and depleting the resource for an immediate gain. Managing a certain resource sustainably might, for example, produce an annual income I, whereas eliminating the resource stock could yield an immediate gain P. The present value of the perpetual income stream I is equal to I/i, where i is the annual discount rate. If it happens that P is larger than I/i, it will be more profitable to deplete the resource. By investing the proceeds at interest rate i, the owner will obtain a greater income than by harvesting it sustainably.

Many biological resources, such as whales, trees, and soil, have low annual rates of productivity, often less than 5% per annum. Removing more than 5% of the stock each year will deplete the resource, eventually to extinction. If the resource owner discounts future revenues at more than 5% per annum, he or she may deliberately choose to eliminate the resource for short-term gain.

Resource users may have reasons for discounting future income at rates higher than market interest rates. For example, they may need, or desire, an immediate cash flow. In addition, the future of the resource may be uncertain, enhancing the relative value of current harvests. Whatever the reason, a resource owner may decide to practice less conservation than may be consistent with sustainable use of the resource.

The same calculation applies to multiple users of a managed common property resource. It is a notorious fact, for example, that the fishing industry almost always demands larger quotas than recommended by government managers. The managers' calculations may be based tacitly on zero discounting, but the industry almost always prefers to receive its income now.

Community Control

A subsidiary theme in the common property resource literature proposes that common property can be a good thing, not necessarily the stage of a tragedy. The idea is that if all so-called "stakeholders" can be assured a voice in the management of the resource, then agreement to achieve the optimum exploitation should follow. Doubtlessly this could be true in a tightly knit society in which defectors face ostracization or other severe penalties. Recent attempts to reach compromise agreements between users of forests or fisheries, however, have proven frustrating. The hypothetical example of the lake with two adjacent property owners suggests that resource conservation can depend strongly on individual behavior. Although decentralization can sometimes bring benefits, the likelihood that community control of biological resources will improve the prospects for conservation needs to be assessed critically in each instance.

As described earlier, individual community members may discount the future to such an extent that the results differ little from the tragedy of the commons. Also, if there is any likelihood that some users will try to take more than their share of the resource, community control can quickly degenerate into a destructive commons scramble. Always a master of the rapier phrase, Hardin (1994) labeled those who preach the doctrine of community control as "commonists." In some cases, community control might resolve a local tragedy of the commons, but the number of individuals involved has to be fairly small, and all outsiders must be excluded. Furthermore, the community probably needs to be homogeneous in tradition and religion.

Risk and Uncertainty

Risk and uncertainty are important in the management of many resources, including common property resources. Here, the following terminology is used: Risk occurs when the outcome of some process is not known in advance, but the probabilities of the various possible outcomes are known from past experience. Uncertainty occurs when these probabilities are not known, or not known accurately. Uncertainty can be reduced through experience and research, but in most situations it cannot be eliminated entirely, that is, it cannot be transformed into pure risk. Examples of irreducible uncertainty are common in resource management. To mention one example, we do not know for certain whether current global warming is the result of increased levels of greenhouse

gases, nor do we know what impact greater concentrations of these gases will have.

In the case of fishery resources, uncertainty persists regarding current stock levels, relationships between stock and recruitment, and the overall effects of exploitation on future stock levels. Furthermore, marine ecosystems are not well understood, meaning that the implications of harvesting certain species cannot be predicted with certainty. In addition, uncertainty pertaining to future environmental conditions, such as global warming, only adds to the difficulty of foreseeing future trends in marine populations.

Reducing uncertainty (e.g., through scientific research), or reducing the effects of uncertainty (e.g., by adopting conservative harvest strategies), can be considered as forms of investment in maintaining the future viability of a renewable resource. The tragedy of the commons operates here as well – under common property conditions, resource exploiters ignore future uncertainties as they attempt to maximize current revenues from the resource. Management strategies that limit resource exploitation on the basis of scientific uncertainty are widely unpopular with fishermen and other common property resource users. Indeed, the argument is often advanced that uncertainty implies that the need to restrict resource use has not been proved. In truth, exactly the opposite is probably true – uncertainty implies that resource use should be limited in a precautionary way.

Biological Reserves

One response to the ever-growing pressure on biological commons is the establishment of protected biological reserves. While the benefits of reserves are apparent in terms of protecting biodiversity, the potential direct economic benefits of reserves have until recently been less appreciated. For example, large-scale marine reserves could help in preventing severe overexploitation, or collapse of important marine resources. Such marine reserves should be thought of as complementary to, rather than as substitutes for, other types of resource management. Unlike other management strategies, however, reserves can provide a positive, potentially "fail-safe" backup to traditional resource management.

Related Concepts

Closely related to the concept of a commons is the economist's concept of externalities (sometimes called spillover effects). An externality is a cost or benefit imposed on others (without compensation) as the result of some economic activity. Externalities can be positive (e.g., a homeowner paints her house to protect it from the weather, and neighbors enjoy the color scheme) or negative (smoke from the homeowner's fireplace chokes her neighbor). Users of common property resources impose negative externalities by reducing the stock of the resource available to other users. Environmental pollution is another common example of a negative externality. In this case, the quality of a resource, rather than the quantity, is reduced by pollution.

Another related economic concept is that of social cost. When a negative externality occurs, the agent that causes the externality incurs a personal cost (his or her private cost) that is less than the total social cost of the activity.

Yet another notion is that of free riders. When the users of a common resource agree to limit their individual impacts, others – the free riders – may continue to exploit the resource. Examples include poachers who illegally slaughter wildlife in protected areas, or nations that fail to ratify or honor international agreements. The dilemma here is that the greater the success in managing the commons, the greater is its attraction to free riders. The problem of free riders explains the dual difficulties inherent in community control of the commons: exclusion of outsiders, and control of cheating by insiders. The failed experiments of national communism serve as strong warnings against undue optimism in these situations.

A basic theoretical prediction of economics, that competitive equilibria are Pareto efficient, is valid only in the absence of externalities. How to remedy the effects of negative externalities is the subject of welfare economics. The immense literature in this field cannot be succinctly summarized, but it can be stated that no fully satisfactory solution to the tragedy of the commons (negative externalities, if you prefer) has been discovered, and it may be that none exists. If not, the commons dilemma promises to become ever more serious as the world's population continues to grow and exert increasing pressures on the biological systems that are essential to our very survival.

See also: Commons, Institutional Diversity of. Environmental Ethics. Fish Stocks. Natural Reserves and Preserves. Property Rights and Biodiversity. Resource Exploitation, Fisheries. Resource Partitioning

References

Andelson RV (1991) *Commons without Tragedy: Protecting the Environment from Overpopulation – A New Approach*. Savage, MD: Barnes and Noble.

Berkes F (ed.) (1989) *Common Property Resources: Ecology and Community-Based Sustainable Development*. London: Belhaven Press.

Clark CW (2010) *Mathematical Bioeconomics: The Mathematics of Conservation*, 3rd edn. New York: Wiley-Interscience.

Gordon HS (1954) The economic theory of a common property resource: The fishery. *Journal of Political Economy* 62: 124–142.

Hardin G (1968) The tragedy of the commons. *Science* 162: 1243–1247.

Hardin G (1994) The tragedy of the unmanaged commons. *Trends in Ecology and Evolution* 9: 199.

Lauck T, Clark CW, Mangel M, and Munro GR (1998) Implementing the precautionary principle in fisheries management through marine reserves. *Ecological Applications* 8(supplement): 872–878.

Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.

Shackell NL and Willison JHM (1995) *Marine Protected Areas and Sustainable Fisheries*. Wolfville, Nova Scotia, Canada: Centre for Wildlife and Conservation Biology, Acadia University.

Commons, Institutional Diversity of

Elinor Ostrom, Indiana University, Bloomington, IN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Common-pool resources Resources that include all ecosystems that are large enough such that excluding potential beneficiaries from their use is a nontrivial task and each individual's consumptive use (e.g., harvesting of a boatload of fish or a truckload of forest products) reduces what is available to others.

Governance systems Sets of rule configurations used to govern human–ecosystem relationships at operational, collective-choice, and constitutional levels of analysis.

Polycentric governance arrangements Complex, multitiered governance systems in which there is no single center of authority.

Public goods Outcomes that are widely shared and the benefit received by one individual does not reduce benefits available to others.

Rules Commonly understood, normative statements that specify who must, must not, or may take some action or affect some outcome at a particular node in a decision tree.

Introduction

Ecosystems characterized by high levels of biodiversity are complex systems with interactions occurring at multiple spatial and temporal scales. The task of enhancing and sustaining biodiversity through the design of institutions needs to take into account the complexity of the natural systems in which biodiversity occurs. Without effective institutions to limit who can use and regulate diverse harvesting and management practices, common-pool resources can be overharvested and even destroyed. When common-pool resources containing endemic or threatened species are destroyed, the amount of biodiversity in the world is reduced.

When complexity characterizes the nature of the systems to be governed, it is essential to think seriously about the complexity of the proposed governance systems. Without sufficient concern for sustaining complex governance systems, the very processes of regulating behavior to preserve biodiversity will produce the tragic and unintended consequence of destroying the complexity we are trying to protect. In his book *Design for a Brain: The Origin of Adaptive Behavior*, Ashby (1960), an eminent biologist of an earlier era, developed the Law of Requisite Variety. This law can be roughly stated as: Any regulative system needs as much variety in the actions that it can take as exists in the system it is regulating. Translated into the subject of this Encyclopedia, the law of requisite variety could be stated as: Any governance system that is designed to regulate complex biological systems must have as much variety in the actions that it can take as there exists in the systems being regulated. This is a tall order, but it is one to which we need to pay serious attention.

Unfortunately, much of contemporary policy analysis does not base recommendations on the law of requisite variety. Policy proposals related to the preservation of biodiversity tend to focus on two entirely different solutions: (1) the creation of effective market institutions and (2) the creation of national agencies to manage all common-pool resources within a large domain. Both markets and the state are important institutional mechanisms that should be part of the solutions of protecting biodiversity, but not the exclusive mechanisms. A third option, that of combining large numbers of relatively autonomous local governance systems for specific common-pool resources with larger scale governance regimes – in other words, a polycentric system – has not been given sufficient attention.

This article focuses on monocentric and polycentric forms of regulation. The section Centralized Regimes and Biodiversity discusses the underlying presumption of substantial policy proposals, which is that strong, centralized regimes are essential for managing common-pool resources. Substantial empirical work challenges the validity of this presumption. The section Using Rules as Tools to Change the Structure of Incentives in Common-Pool Resources explores the complexity of using rules as tools to change the structure of commons dilemmas, presents the seven clusters of rules that affect the components of any action situation, and then describes the specific rules that are used in field settings by resource users and government agencies. An examination of the types of rules used in the field yields several important findings. First, the number of rules actually used in field settings is far greater than generally recognized. Second, the type of rules is also different. Boundary rules tend to include as coappropriators of a resource those who are more likely to be trustworthy because they live permanently nearby and have a long-term stake in keeping a resource sustainable. Choice rules define rights and duties that are easy to understand, directly related to sustaining the biophysical structure of the resource, and easy to monitor and enforce. Some rules recommended in the policy literature are not found among the rules used by self-organized systems. This leads to the difficult search for optimal rules, which is discussed in the section The Daunting Search for Better Rules.

Given the complexity of the process of designing rules to regulate the use of common-pool resources, the author argues in the section Experimenting with Rule Change that all policy proposals must be considered as experiments. No one can possibly know whether a proposed change in rules is among the more optimal rule changes, or even whether a rule change will lead to an improvement. All policy experiments have a positive probability of failing. In the section Self-Organized Resource Governance Systems as Complex Adaptive Systems, the author draws on recent research by Holland and

colleagues at the Santa Fe Institute to discuss the attributes and mechanisms of a different form of general organization – a complex adaptive system – that is not the result of central direction. Complex adaptive systems cannot be understood if one tries to fit these systems into the image of an organization with a central director.

The section The Advantages and Limits of Parallel Sets of Local Users in Policy Experiments then shows how the parallel efforts by a large number of local resource users to search out and find local rule configurations may find better rule combinations over the long term, whereas top–down design processes are more limited in their capacities to search and find appropriate rules. All forms of decision making have limits. In the section The Capabilities of Polycentric Systems in Coping with Tragedies of the Commons, the author discusses the limits of a series of completely independent resource governance systems and the importance of building polycentric governance systems with considerable overlap to combine the strengths of parallel search and design processes with the strengths of larger systems in conflict resolution, acquisition of scientific knowledge, monitoring the performance of local systems, and the regulation of common-pool resources that are more global in their scope. The resulting polycentric governance systems are also not directed by a single center. They, too, are complex adaptive systems requiring policy analysts to change their fundamental views of organization in order to cope more effectively with tragedies of the commons and many of the other problems facing modern societies.

Centralized Regimes and Biodiversity

A substantial number of contemporary policy proposals stress the desirability of central regimes for the regulation of natural resource systems. The scientific management of natural resources, for example, which is regularly taught in many universities to future regulators of natural resources, presents fisheries and forests as relatively homogeneous systems that are closely interrelated across vast domains. This approach, as it has been applied to fisheries management, is described by Acheson *et al.* (1998: 391–392):

> For those trained in scientific management, it is also an anathema to manage a species over only part of its range… . [I]t is not rational to protect a species in one zone only to have it migrate into another area where it can be taken by other people due to a difference in regulations. Lobsters, for example, extend from Newfoundland to the Carolinas… . From the point of view of the National Marine Fisheries Service, it makes sense to have a set of uniform regulations for the entire US coast rather than one for each state.

Contemporary policy analysts also share a belief in the feasibility of designing close-to-optimal rules to govern and manage species over a large domain by utilizing top–down direction. Since natural resources are viewed as relatively homogeneous, and interlinked and simple models exist of how they work, officials acting in the public interest are considered capable of devising uniform and effective rules for an entire region. Prescriptions calling for central governments to impose uniform regulations over most natural resources are thus consistent with much of contemporary policy analysis (but see the extensive evidence presented in Poteete *et al.*, 2010).

These prescriptions are not, however, supported by extensive research. An extraordinary number of field studies have found that local groups of resource users, sometimes by themselves and sometimes with the assistance of external institutional arrangements, have created a wide diversity of institutional arrangements for coping with common-pool resources when they have not been prevented by central authorities. These empirical studies document successful self-organized resource governance systems in diverse sectors in all parts of the world. Examples also exist of commons dilemmas that have continued unabated. One conclusion that can firmly be made in the light of extensive empirical evidence is that overuse and destruction of common-pool resources is not a determinant and an inescapable outcome when multiple users face a commons dilemma. Scholars have begun to identify the conditions of a resource, and of the users of a resource, that are most conducive to local users self-organizing to find solutions to commons dilemmas. Furthermore, the broad design principles that characterize robust self-organized resource governance systems that have resolved commons dilemmas for long periods of time have been identified (Ostrom, 1990) and found basically sound by other scholars (Cox *et al.*, 2010).

Another important set of findings is that national governmental agencies have been notably unsuccessful in their efforts to design an effective and uniform set of rules to regulate important common-pool resources across a broad domain. Many developing countries nationalized all land and water resources during the 1950s and 1960s. The institutional arrangements that local users had devised to limit entry and harvesting lost their legal standing, but the national governments lacked funds and personnel to monitor the use of these effective resources. Thus, common-pool resources were converted to de jure government-property regimes, but reverted to de facto open-access regimes. The perverse incentives of such open-access regimes are often accentuated, since local users have specifically been told that they will not receive any benefits if they adopt a long-term perspective and regulate their use of the resource. When resources that were previously controlled by local participants have been nationalized, state control has usually proved to be less effective and efficient, if not disastrous, than control by those directly affected. The harmful effects of nationalizing forests that had earlier been governed by local user groups have been well documented in Asia and in Africa. Similar results have occurred with regard to inshore fisheries taken over by state or national agencies from local control by the inshore fishermen themselves.

These findings challenge the theoretical foundations of much of contemporary policy analysis. A foundational belief underlying this analysis is that designing rules to change the incentives of participants is a relatively simple analytical task best done by objective analysts not intimately related to any specific resource. Analysts view most resources in a particular sector as relatively similar and sufficiently interrelated, and thus conclude that they need to be governed by the same set of rules. This is bolstered by the view that an organization itself requires central direction. Consequently, the multitude of self-organized resource governance systems are viewed as mere collections of individual agents each out to maximize their

own short-term returns. The groups who have actually organized themselves are frequently invisible to those who cannot imagine organization without rules and regulations issued by a central authority.

To understand why designing rules to regulate the use of common-pool resources so as to sustain their long-term productivity and biodiversity is not a simple task, one has to somewhat change the perspective. One must begin to think through the process of using rules as tools to change the structure of incentives facing users of a common-pool resource.

Using Rules as Tools to Change the Structure of Incentives in Common-Pool Resources

With this change in perspective, we can think of users trying to understand both the biophysical structure of a common-pool resource and how to affect each other's incentives in order to increase the probability of sustainable use over the long term. Instead of being endowed with complete knowledge of how a common-pool resource reacts to use patterns over time, they have to explore and discover the biophysical structure of a particular resource that will differ on key parameters from similar resources in the same region. Further, they have to cope with considerable uncertainty related to the weather, complicated growth patterns of biological systems that may at times be chaotic in nature, and external price fluctuations that affect the costs of inputs and the value of outcomes. In addition to the physical changes that they can make in the resource, seven clusters of rules directly affect the components of their own action situations. This list of rules is the result of many years of theoretical and empirical work on a framework for conducting institutional analyses (see Ostrom *et al.*, 1994: Ch. 2). Specifically, the rules they can change include:

1. Boundary rules affect the characteristics of the participants.
2. Position rules differentially affect the capabilities and responsibilities of those in positions.
3. Choice rules affect the actions that participants in positions may, must, or must not do.
4. Scope rules affect the outcomes that are allowed, mandated, or forbidden.
5. Aggregation rules affect how individual actions are transformed into final outcomes.
6. Information rules affect the kind of information present or absent in a situation.
7. Payoff rules affect assigned costs and benefits to actions and outcomes.

Given the nonlinearity and complexity of common-pool resource situations, it is rarely easy to predict what effect a change in a particular rule will produce. For example, a change in a boundary rule to restrict the entry of users simultaneously reduces the number of individuals who are tempted to break choice rules, but it also reduces the number of individuals who monitor others or contribute funds toward hiring a guard. Thus, the opportunities for rule breaking may increase. Further, the cost of a rule infraction will be spread over a smaller group of users and, thus, the harm to any individual may be greater. Assessing the overall effects of a change in boundary rules is a nontrivial analytical task. Instead of conducting such a complete analysis, users are more apt to use their intuitive understanding of the resource and each other to experiment with different rule changes until they find a combination that seems to work in their setting.

To understand the types of tools that are available to users somewhat better, the kinds of boundary, choice, payoff, and position rules used in field settings are examined in some detail. These four clusters of rules are the major tools used to affect the management of common-pool resources, whereas information, scope, and aggregation rules are utilized to complement changes induced by these four rules.

For the past two decades, colleagues at or associated with the Workshop in Political Theory and Policy Analysis at Indiana University have studied a very large number of fisheries, forests, irrigation systems, and groundwater basins, as well as other common-pool resources. Colleagues have collected an immense archive of original case studies written by many scholars on all sectors in all parts of the world (Hess, 1999). Colleagues have developed structured coding forms to help them identify the specific kinds of common-pool situations faced in the field, as well as the types of rules that users have evolved over time to try to govern and manage these resources effectively. To develop standardized coding forms, colleagues read hundreds of cases describing how local common-pool resources were or were not regulated by a government agency, by the users themselves, or by a nongovernmental organization (NGO).

Affecting the Characteristics of Users through Boundary Rules

The most frequent recommendation concerning boundary rules in the policy literature is to limit the number of persons allowed to appropriate from a common-pool resource so that the level of appropriation is reduced or to require users to obtain a license before harvesting. Boundary rules affect the types of participants with whom other participants will be interacting. If contingent cooperation is perceived to be a possibility, then one of the most important ways to enhance the likelihood of using reciprocity norms is to increase the proportion of participants who are well known in a community, have a long-term stake in that community, and would find it costly to have their reputation for trustworthiness harmed in that community. Reducing the number of users while opening the resource to strangers who are willing to pay a license fee, but who lack a long-term interest in the sustainability of a particular resource, may reduce the level of trust and willingness to use reciprocity and thus increase enforcement costs substantially.

Colleagues have identified 27 boundary rules as having been used in at least one common-pool resource somewhere in the world (Ostrom *et al.*, 1989). Although some systems use only a single boundary rule, many use two or three of these rules in combination. Boundary rules can be broadly classified in three general groups defining how individuals gain authority to appropriate resource units from a common-pool resource. The first type of boundary rule relates to an individual's citizenship, residency, or membership in a particular organization. Many forestry and fishing user groups require

members to have been born in a particular location. A second broad group of rules relates to individuals who have ascribed or acquired personal characteristics. Other user groups may require that appropriation depend on ethnicity, clan, or caste. A third group of boundary rules relates to the relationship of an individual with the resource itself. Using a particular technology or acquiring appropriation rights through an auction or a lottery are the examples of this type of rule. About half of the rules relate to the characteristics of the users themselves. The other half involves diverse relationships with the resource.

In a systematic coding of those case studies for which sufficient information existed about rules related to inshore fisheries in many parts of the world, Schlager coded 33 user groups out of the 44 groups identified as having at least some rules regarding the use of the resource. All 33 groups depended on some combination of 14 different boundary rules (Schlager, 1994: 258). None of these groups relied on a single boundary rule. Thirty out of 33 groups (91%) limited fishing to those individuals who lived in a nearby community, whereas 13 groups also required membership in a local organization. Consequently, most inshore fisheries organized by the users themselves restrict fishing to those individuals who are well known to each other, have a relatively long-term time horizon, and are connected to one another in multiple ways.

After residency, the next most frequent type of rule, used in two-thirds of the organized subgroups, involves the type of technology that a potential fisher must use. These rules are often criticized by policy analysts, since gear restrictions tend to reduce the efficiency of fishing. Gear restrictions have many consequences, however. Used in combination with choice rules that assign fishers using one type of gear to one area of the fishing grounds and fishers using another type of gear to a second area, they solve conflicts among noncompatible technologies. Many gear restrictions also place a reduced load on the fishery itself and thus help to sustain longer term use of the resource.

Other rules were also used. A scattering of groups used ascribed characteristics (age, two groups; ethnicity, three groups; race, five groups). Three types of temporary use-rights included government licenses (three groups), lottery (five groups), and registration (four groups). Seven groups required participants to have purchased an asset such as a fishing berth, whereas three groups required ownership of nearby land. Schlager did not find that any particular boundary rule was correlated with higher performance levels across the groups. She did find, however, that the 33 groups who had at least one boundary rule tended to be able to solve common-pool problems more effectively than the 11 groups who had not crafted boundary rules.

Thus, many of the rich diversity of boundary rules used by users in the field attempt to ensure that the users will be relating to others who live nearby and have a long-term interest in sustaining the productivity of the resource. One way of coping with the commons is, thus, changing the composition of who uses a common-pool resource to increase the proportion of participants who have a long-term interest, who are more likely to use reciprocity, and who can be trusted. Central governments tend to use a smaller set of rules and some of these may open up a resource to strangers without a longer term commitment to the resource.

Affecting the Set of Allowable Actions through Choice Rules

Choice rules are also a major type of rule used to regulate common-pool resources. Colleagues identified a diversity of choice rules used in field settings. Some rules involve a simple formula. Many forest resources, for example, are closed to all forms of harvesting during one portion of the year and open for extraction by all who meet the boundary rules during an open season. Most choice rules, however, have two components: an assignment and a basis. A fisher, for example, might be assigned to a fixed location (a fishing spot) or to a fixed rotational schedule, or a member of the founding clan may be assigned (allowed) to cut timber anywhere in a forest. In addition to the assignment used in a choice rule, most rules required a basis for the assignment. For example, a fisher might be assigned to a fixed location based on a number drawn in a lottery, on the purchase of that spot in an auction, or on the basis of his or her historical use.

If all of the bases were likely to be combined with all of the assignment possibilities, there would be 112 different choice rules (8 assignment formulas times 14 bases). A further complication is that the rules for one product may differ from those of another product in the same resource. In regard to forest resources, for example, children may be authorized to pick fruit from any tree located in a forest so long as it is for their own consumption, women may be authorized to collect so many headloads of dead wood for domestic firewood and certain plants for making crafts, whereas *shamans* are the only ones authorized to collect medicinal plants from a particular location in a forest. Appropriation rights to fish are frequently related to a specific species. Thus, the exact number of rules that are actually used in the field is difficult to compute since not all bases are used with all formulas, but many rules focus on specific products. A further complication is that the rules may regularly change over the course of a year, depending on resource conditions.

Schlager (1994: 259–260) found that all 33 organized subgroups used one of five basic assignments in their choice rules. Every user group included in her study assigned fishers to fixed locations using a diversity of bases, including technology, lottery, or historical use. Thus, spatial demarcations are a critical variable for inshore fisheries. Nine user groups required fishers to limit their harvest to fish that met a specific size requirement, whereas seven groups allocated fishers to fishing spots using a rotation system and seven other groups only allowed fishing locations to be used during a specific season. Four groups allocated fishing spots for a particular time period (a fishing day or a fishing season).

An important finding – given the puzzles addressed in this article – is that the choice rule most frequently recommended by policy analysts is not used in any of the coastal fisheries included in Schlager's study. No attempt was made

> by the fishers involved to directly regulate the quantity of fish harvested based on an estimate of the yield. This is particularly surprising given that the most frequently recommended policy prescription made by fishery economists is the use of individual transferable quotas based on estimates on the economically optimal quantity of fish to be harvested over the long run (Schlager, 1994: 265).

In an independent study of 30 traditional fishery societies, Acheson, Wilson, and Steneck also noted the surprising absence of quota rules:

> All of the rules and practices we found in these 30 societies regulate [how] fishing is done. That is, they limit the times [when] fish may be caught, the locations where fishing is allowed, the technology permitted, and the stage of the life cycle during which fish may be taken. None of these societies limits the [amount] of various species that can be caught. Quotas – the single most important concept and tool of scientific management – is conspicuous by its absence (Acheson *et al.*, 1998: 397).

Local inshore fishers, when allowed to manage a riparian area, thus use rules that differ substantially from those recommended by advocates of scientific management. Fishers have to know a great deal about the ecology of their inshore region, including spawning areas, nursery areas, the migration routes of different species, and seasonable patterns, just to succeed as fishers. Over time, they learn how

> to maintain these critical life-cycle processes with rules controlling technology, fishing locations, and fishing times. Such rules, in their view, are based on biological reality (Acheson *et al.*, 1998: 405).

The diversity of rules devised by users greatly exceeds the limited choice rules recommended in textbook treatments of this problem. Users thus cope with the commons by a wide variety of rules affecting the actions available to participants and, thus, their basic set of strategies. Given this wide diversity of rules, it is particularly noteworthy that rules assigning users a right to a specific quantity of a resource are used so infrequently in inshore fisheries. (They are used more frequently when allocating forest products, where the quantity available, as well as the quantity harvested, is much easier to measure.) To assign a user a specific quantity of a resource unit requires that those making the assignment know the total available units.

Affecting Outcomes through Payoff and Position Rules

One way to reduce or redirect the appropriations made from a common-pool resource is to change payoff rules so as to add a penalty to actions that are prohibited. Many user groups also adopt norms that those who are rule breakers should be socially ostracized or shunned and individual users tend to monitor each other's behavior rather intensively. Three broad types of payoff rules are used extensively in the field: (1) the imposition of a fine, (2) the loss of appropriation rights, and (3) incarceration. The severity of each of these types of sanctions can range from very low to very high and tends to start out on the low end of the scale. Inshore fisheries studied by Schlager relied heavily on shunning and other social norms and less on formal sanctions.

Passing rules that impose costs is relatively simple. The real difficult task is monitoring behavior to ascertain if rules are being broken. Self-organized fisheries tend to rely on self-monitoring more than the creation of a formal position of guard. Most inshore fishers now use shortwave radios as a routine part of their day-to-day operations, allowing a form of instant monitoring to occur. An official of a West Coast Indian tribe reports, for example, that:

> it is not uncommon to hear messages such as 'Did you see so-and-so flying all that net?' over the shortwave frequency – a clear reference to a violation of specified gear limits (Singleton, 1998: 134).

Given that most fishers will be listening to their shortwave radio,

> such publicity is tantamount in creating a flashing neon sign over the boat of the offender. Such treatment might be preceded or followed by a direct approach to the rule violator, advising him to resolve the problem. In some tribes, a group of fishermen might delegate themselves to speak to the person (Singleton, 1998: 134).

Among self-organizing forest governance systems, creating and supporting a position as guard is frequently essential, since resource units are highly valuable and a few hours of stealth can generate substantial illicit income. Monitoring rule conformance among forest users by officially designated and paid guards may make the difference between a resource in good condition and one that has become degraded. In a study of 279 forest *panchayats* in the Kumaon region of India, Agrawal and Yadama (1997) found that the number of months a guard was hired was the most important variable affecting forest conditions. The other variables that affected forest conditions included the number of meetings held by the forest council (a time when infractions are discussed) and the number of residents in the village.

It is evident from the analysis that the capacity of a forest council to monitor and impose sanctions on rule-breakers is paramount in maintaining the forest in good condition. Also, the presence of a guard should not be taken simply as a formal mechanism that ensures greater protection. It is also an indication of the informal commitment of the *panchayat* and the village community to protect their forests. Hiring a guard costs money… If there was scant interest in protecting the forest, villagers would have little interest in setting aside the money necessary to hire a guard (Agrawal and Yadama, 1997: 455). Recent findings from the International Forestry Resources and Institutions (IFRI) research program have also shown that the outcome of users themselves monitoring forests is extremely effective (Coleman and Steed, 2009; Chhatre and Agrawal, 2008, 2009).

Boundary and choice rules also affect how easy or difficult it is to monitor activities and impose sanctions on rule infractions. Closing a forest or an inshore fishery for a substantial amount of time, for example, has multiple impacts. It protects particular plants or fish during critical growing periods and allows the entire system time to regenerate without disturbance. Further, during the closed season, rule infractions are highly obvious to anyone as any user in the resource is almost certainly breaking the rules. Similarly, requiring users to use a particular technology may reduce the pressure on the resource, help to solve conflicts among users of incompatible technologies, and also make it very easy to ascertain if rules are being followed. Changing payoff rules is the most direct way of coping with commons dilemmas.

Affecting Outcomes through Changes in Information, Scope, and Aggregation Rules

Information, scope, and aggregation rules tend to be used in ways that complement changes in boundary, choice, payoff, and position rules. Individual systems vary radically in regard to the mandatory information that they require. Many smaller and informal systems rely entirely on a voluntary exchange of information and on mutual monitoring. Where resource units are valuable and the size of the group is large, more and more requirements are added regarding the information that must be kept by users or their officials. Scope rules are used to limit harvesting activities in some regions that are being treated as refugia. By not allowing any appropriation from these locations, the regenerative capacity of a system can be enhanced. Aggregation rules are used extensively in collective-choice processes and less extensively in operational settings, but one aggregation rule that is found in diverse systems is a requirement that harvesting activities be done in teams. This increases the opportunity for mutual monitoring and reduces the need to hire special guards.

It is important to note that repeated studies have not yet found any particular rules to have a statistically positive relationship to performance. The absence of any boundary or any choice rules is consistently associated with poor performance. Relying on only a single type of rule for an entire set of common-pool resources is also negatively related.

The Daunting Search for Better Rules

It should now be obvious that the search for rules that improve the outcomes obtained in commons dilemmas is an incredibly complex task involving a potentially infinite combination of specific rules that could be adopted. To ascertain whether one has found an optimal set of rules to improve the outcomes achieved in a single situation, one would need to analyze how diverse rules affect incentives, strategies, and outcomes. Since there are multiple rules that affect incentives, conducting such an analysis would be an incredibly time- and resource-consuming process. For example, if only five changes in rules per component were considered, there would be 5^7 or 75,525 different rule configurations to analyze. This is a gross simplification however, since some of the important rules used in field settings include more than 25 rules (in the case of boundary rules) and even more than 100 variants (in the case of choice rules). Further, how these changes affect the outcomes achieved in a particular location depends on the biophysical characteristics of that location and the type of community relationships that already exist. No set of policy analysts (or even all of the policy analysts in the world today) could ever have sufficient time or resources to analyze more than 75,000 combinations of rule changes and resulting situations, let alone all of the variance in these situations due to biophysical differences.

Experimenting with Rule Changes

Instead of assuming that designing rules that approach optimality or that even improve performance, is a relatively simple analytical task that can be undertaken by distant objective analysts, we need to understand the policy design process as involving an effort to tinker with a large number of component parts (see Jacob, 1977). Those who tinker with any tools – including rules – are trying to find combinations that work together more effectively than other combinations. Policy changes are experiments based on more or less informed expectations about potential outcomes and the distribution of these outcomes for participants across time and space (Campbell, 1969). Whenever individuals agree to add a rule, change a rule, or adopt someone else's proposed rule set, they are conducting a policy experiment. Further, the complexity of the ever-changing biophysical world combined with the complexity of rule systems means that any proposed rule change faces a nontrivial probability of error.

When there is only a single governing authority, policymakers have to experiment simultaneously with all of the common-pool resources within their jurisdiction with each policy change. Once a change has been made and implemented, further changes will not be made rapidly. The process of experimentation will usually be slow, and information about results may be contradictory and difficult to interpret. Thus, an experiment that is based on erroneous data about one key structural variable or one false assumption about how actors will react can lead to a very large disaster (see Wilson *et al.*, 1999). In any design process where there is substantial probability of error, having redundant teams of designers has repeatedly been shown to have considerable advantage (see Landau, 1969).

Self-Organized Resource Governance Systems as Complex Adaptive Systems

As discussed in the section Centalized Regimes and Biodiversity, the very concept of organization is closely tied for many scholars to the presence of a central director who has designed a system to operate in a particular way. Consequently, the mechanisms used by noncentrally directed systems are not always well understood. Many self-organized resource governance systems are invisible to the officials of their own country or those from donor agencies. A classic example of this occurred in the Chitwan valley of Nepal several years ago when an Asian Development Bank team of irrigation engineers recommended a very large loan to build a dam across the Rapti River to enable the farmers there to irrigate their crops. What the engineering design team did not see was the 85 farmer-managed irrigation systems that already existed in the valley and that had achieved relatively high performance. Most farmers in the Chitwan valley already obtained three irrigated crops a year as a result of their participation in the activities of these irrigation systems (see Benjamin *et al.*, 1994). In fact, a study of more than 200 irrigation systems in Nepal found that the farmer-managed irrigation systems outperformed the well-financed agency-managed irrigation systems in regard to economic and technical efficiency as well as equity in terms of delivering substantial water to the tail end of the systems (Joshi *et al.*, 2000).

In contrast to forms of organization that are the result of central direction, most self-organized groups – including the

types of locally organized users of fisheries and forests discussed in this article – are better viewed as complex adaptive systems. These systems are composed of a large number of active elements whose rich patterns of interactions produce emergent properties that are not easy to predict by analyzing the separate parts of a system. Holland (1995: 10) views complex adaptive systems as:

> systems composed of interacting agents described in terms of rules. These agents adapt by changing their rules as experience accumulates. Such systems exhibit coherence under change, via conditional action and anticipation, and they do so without central direction (Holland, 1995: 38–39).

Holland points out that complex adaptive systems differ from physical systems that are not adaptive and that have been the foci of most scientific effort. It is the physical sciences that have been the model for many aspects of contemporary social science. Thus, the concepts needed to understand the adaptivity of systems are not yet well developed by social scientists.

Properties and Mechanisms of Complex Adaptive Systems

No general theory of complex adaptive systems yet exists to provide a coherent explanation for processes shared by all complex adaptive systems. Biologists have studied many different adaptive systems but within separate fields of biology. Thus, even biologists have not recognized some of the similarities of structures and processes that characterize the central nervous system, the immune system, and the evolution of species. Recent work at the Santa Fe Institute has begun to identify central attributes, mechanisms, and processes used by all complex adaptive systems including both biological systems as well as markets and other social systems that are not centrally directed.

It appears that all complex adaptive systems share four basic properties: nonlinearity, flows, diversity, and aggregation. The first three properties are self-evident and clearly characterize the types of self-organized resource governance systems discussed in this article. Aggregation refers to the

> emergence of complex larger-scale behaviors from the aggregate interactions of less complex agents (Holland, 1995: 11).

For example, many irrigation systems are divided into several tiers and multiple units at each of these tiers. The farmers on a field irrigation channel are responsible for distributing the water to this small channel as well as keeping it in good repair. All farmers whose field channels are served by a branch canal may send a representative to a branch canal organization that focuses its attention on the distribution of water among all branches and on the maintenance of the distribution canals. The branch canal organization may send a representative to a central committee who is responsible for the headworks that divert the water from a river into the system in the first place. The rules used on one branch canal or one field channel may be quite different than on others. There is no single center of authority for these systems that makes all relevant decisions on how to get water from the river to a farmer's field, but in many farmer-organized systems the water is distributed in an organized fashion and all of the waterworks are maintained as a result of the aggregation of decisions and actions at multiple levels.

In addition to these four attributes, complex adaptive systems also use three mechanisms that are key to the adaptive process itself. These include the use of tags, internal models, and building blocks.

The Use of Tags

Tagging is a universal mechanism for boundary formation and aggregation of units in complex adaptive systems.

> Tags are a pervasive feature of (complex adaptive systems) because they facilitate selective interaction. They allow agents to select among agents or objects that would otherwise be indistinguishable (Holland, 1995: 14).

All of the types of boundary rules discussed here involve the specification of the tags that will be used to determine who is authorized to be a co-user from a common-pool resource. Residency, prior membership, and personal characteristics are attributes that already exist and are easy to use as boundary tags. Local governance systems rely heavily on tags that identify individuals who are already known to each other, who have a long-term stake in the sustainability of a resource, who have an incentive to build a reputation for being trustworthy, and who are thus likely to extend reciprocity rather than recalcitrance in dealing with joint problems.

Tags are also used extensively to mark locations in a resource, to warn rule infractors, and even to mark individual organisms that need to be treated in a special way. An example of the latter occurs along the Maine coast, where it is forbidden to harvest berried lobsters (those with eggs). Such lobsters are V-notched and returned back to the sea. Any other fisher who captures a V-notched lobster is also supposed to return it to the sea.

Internal Models

The users from a common-pool resource build internal models of the resource, of the relationships among the components of the resource, and of where their own actions positively or negatively affect one another and the resource. Among the shared lore for most fishing villages is a clear understanding of where fish breed, where young fish tend to cluster, the length of time it takes for fish to be mature and reproduce, the migration patterns of fish, the food chain in a location, and other information. Many inshore fishers develop their own maps of all of the fishing spots in their grounds. In an effort to reduce the interference of one boat with another boat's fishing, these are frequently defined so that if all spots are filled, all boats are still able to have a good chance to catch fish. These maps are then used in a variety of allocation rules that specify the basis for how any particular boat is assigned to a particular fishing spot. Users of forests also map their forests and may create refugia – sometimes as sacred forests – for sections of a forest that are particularly rich in biodiversity. By not harvesting from these refugia, they serve as a source of regeneration to other nearby locations that are disturbed through harvesting.

Building Blocks

Building blocks are ways of breaking down complex processes into small chunks that can be used in multiple ways and combined and recombined repeatedly and at diverse levels. Once a choice rule that allocates resource units on some basis is determined for example, using the same basis again to allocate responsibility for maintenance work is considered to be a fair allocation of benefits and costs in many cultures and is relatively easy to remember. On the large Chhattis Mauja farmer-organized system in Nepal, for example, water was originally allocated by the land area served. In the 1950s, the formula used for maintenance work was that each branch canal was responsible to send one person to work on the main canal for each 17 ha of area it irrigated.

> The term used for a person-day of labor for canal maintenance was *kulara*. Since the share of water a branch canal is entitled to receive is the same as the resource mobilization requirement, water allocation is now also referred to as 'so many *kulara* of water' (Yoder, 1991: 7).

As the system has grown, the total number of kularas has now been set at 177 shared among 44 branch canals. Voting rights are now also set in terms of *kularas*.

> Therefore, a branch canal with five *kulara* was entitled to 5/177 of the water in the main canal, responsible to supply 5/177 of the resources mobilized for the irrigation, and had five of the total 177 votes in all important decisions (Yoder, 1991: 7).

Changing Rules as an Adaptive Process

Given the logic of combinatorics, it is impossible – as we showed earlier – to conduct a complete analysis of the expected performance of all of the potential rule changes that could be made by the individuals who are served by a self-organized resource governance system and are trying to improve its performance. A similar impossibility also exists for many biological systems. Let us explore these similarities.

Self-organizing resource governance systems have two structures that are somewhat parallel in function to the concepts of a genotype and a phenotype in biology. Phenotypic structures characterize an expressed organism – how bones, organs, and muscles develop, relate, and function in an organism in a particular environment. The components of an action situation characterize an expressed situation – how the number of participants, the information available, and their opportunities and costs create incentives, and how incentives lead to types of outcomes in a particular environment. The genotypic structure characterizes the set of instructions encoded in DNA to produce an organism with a particular phenotypic structure. A rule configuration is a set of instructions on how to produce the structure of relationships among individuals in an action situation, which is also affected by the biophysical world and the kind of community or culture in which an action situation is located.

The evolution of social systems does not follow the same mechanisms as the evolution of species. As an evolutionary process, of course, there must be the generation of new alternatives, selection among new and old combinations of structural attributes, and retention of those combinations of attributes that are successful in a particular environment. In evolving biological systems, genotypic structures are changed through mechanisms such as crossover and mutation, and the distribution of particular types of instructions depends on the survival rate of the phenotypes that they produce in given environments. Instead of blind variation, however, human agents do try to use reason and persuasion in their efforts to try to devise better rules, but the process of choice always involves experimentation.

Rule configurations can be represented as a string of symbols that describe which rules are in effect in a particular location. In **Table 1**, the left-hand column lists a small subset of the rules typically found in the field. A "1" has been entered in a cell of a column to indicate when a specific rule is in use, and a "0" has been entered when the rule is not in use. Thus, the eight columns represent the simplified rule configurations for eight hypothetical locations.

Location A would be characterized as one at which the participants are required both to live in a local community and to use geographic assignments or a rotation system, but also at which there are no formal sanctions used or guards hired to monitor conformance to the entry rules and rotation system. Further, no records are kept, there are no refugia, and users can work individually or in teams. In contrast, Location C creates a refugium in which no appropriation may be undertaken, and does not allocate the remaining space to individual users. If there is a very good breeding and nursery area in this fishery, and it can be well demarcated so that it is obvious if someone has ventured into the refugium by mistake (or by intent), a rule system that does not limit fishing in any

Table 1 Rule configurations and appropriation situations[a]

Rule configurations	*Locations*							
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Boundary rules								
Local residency	1	1	1	1	0	0	0	0
Member of local coop	0	1	0	1	1	1	1	1
Choice rules								
Spatial assignment	1	0	0	0	1	0	1	0
Rotate in time	1	1	0	0	1	1	0	1
Payoff rules								
Fines	0	0	1	1	0	0	1	1
Loss of rights	0	0	0	0	0	0	1	1
Position rules								
All symmetric	1	1	1	1	0	0	0	0
Guard created	0	0	0	0	1	1	1	1
Information rules								
Infractions are recorded	0	0	0	1	1	0	0	1
Appropriation is recorded	0	0	0	0	0	0	0	1
Scope rules								
Forbidden access to refugia	0	0	1	0	1	0	1	1
Aggregation rules								
Must appropriate in teams	0	0	0	0	0	0	0	1

[a]A "1" indicates when a specific rule is in use, and a "0" indicates when the rule is not in use. See text for details.

other way may be a very effective system for protecting the regenerative capabilities of a fishery. It is certainly less expensive than naming specific fishing spots, making an allocation of each fishing team to particular spots, and enforcing a more complicated rule system. At Location C, users also need to live in the local area, and they do face fines if a fellow user notices that they have appropriated from the refugium. Though a guard is not created, records are kept of any infractions noted through the mutual monitoring process.

Location H utilizes a much fuller array of rules. Users must belong to a local cooperative, even though they do not have to live in the local community. Both a refugium and a system for allocating time in which to conduct appropriation have been created, although, with the exception of the forbidden area of the refugium, the latter is not spatially defined. Payoff rules have been modified so that there are both fines for infractions as well as the possibility of losing rights to appropriate, if infractions are substantial. Records of infractions and of the time when someone appropriates are required, and individuals are constrained to working in teams rather than being allowed to appropriate as individuals.

Whether any or all of these rule configurations structured the incentives of users so that participants appropriate from their resource in a sustainable and efficient manner cannot be ascertained from learning about the rules alone. Whether a set of rules enhances performance depends on the structure of the biophysical system in that location and the willingness of participants to follow the rules. Whether the strategy of creating a refugium, as in C and H, works better than the geographic assignment, as in A, depends on how large the refugium is, how rapidly and extensively regeneration from the refugium spreads through the rest of the resource, and how well users understand each of these systems and try to make them work. All three rule systems are likely to work well in some locations and poorly in others.

With this form of representation, we can begin to see how experimentation with rules may be similar to structural changes affecting other types of complex adaptive systems. Most systems are likely to start with one or two very simple rules. An obvious first candidate is to use tags to close the boundary to outsiders so that the likelihood of contingent cooperation and conformance to agreements will be enhanced. By only changing a few rules at the beginning, everyone can come to understand those rules while they are evaluating how they work. A second obvious candidate is to use the shared model of the environment built up through years of interaction in an environment to refine where harvesting should be undertaken and when. Space and time are obvious candidates for allocating access to resources in a manner that is relatively low in cost to sustain. If the community is small enough and shares common norms at a high enough level, creating formal sanctions, guards, records, and other rules may not be necessary. Thus, one can imagine a process in which a rule configuration of mostly zeros slowly converts over time to a scattering of ones and potentially to a rule configuration with many ones.

Changes in specific rules may come about through accident (such as forgetting a rule) or through specific collective-choice processes in which considerable time and effort are devoted to considering why the performance needs to be enhanced and which rules might be changed. Since many users will have experience with more than one product (e.g., mushrooms from a forest), rules tested with regard to one product may also be applied to others if they are successful. Migration of individuals into a community brings individuals with repertoires of different rules used in other locations. Commerce with other groups lets users see and learn about others who may be doing better (or worse) than them in regulating a sustainable biodiverse and efficient resource system. Thus, a self-organized resource governance system with a higher level of in-migration or greater communication with other localities is more likely to adapt and change rules over time than one in which few new ideas concerning how to use rules as tools are brought into the system. Trial-and-error processes may give relatively rapid feedback about rules that obviously do not work in a particular environment, but this is not always the case when the effect of human action on the environment has a long time delay. If all self-organized resource governance systems are totally independent and there is no communication among them, then each has to learn through its own trial-and-error process. Many will find that rules that they have tried do not work, whereas some will fail entirely.

The rate of change will differ among self-organized resource governance systems. As with all learning, the rate of change is an important variable that affects performance over time. If change occurs too rapidly, little is learned from each experiment before another is launched. Respect for tradition and even religious mystification has been used to increase the retention of rules that are considered by at least some participants to work better. If the heavy hand of tradition is, however, too heavy and squelches innovation, a system that may have been well adapted to a past environment may find itself faltering as external changes occur without internal changes also occurring.

The Advantages and Limits of Parallel Sets of Local Users in Policy Experiments

Let us now discuss why a series of relatively autonomous, self-organized, resource governance systems may do a better job in policy experimentation than a single central authority. Among the advantages of authorizing the users of smaller-scale common-pool resources to adopt policies that regulate the use of these resources are:

- Local knowledge. Users who have lived and appropriated from a resource system over a long period of time have developed relatively accurate mental models of how the biophysical system itself operates, since the very success of their appropriation efforts depends on such knowledge. They also know others living in the area well and what norms of behavior are considered appropriate.
- Inclusion of trustworthy participants. Users can devise rules that increase the probability that others are trustworthy and will use reciprocity. This lowers the cost of relying entirely on formal sanctions and paying for extensive guarding.
- Reliance on disaggregated knowledge. Feedback about how the resource system responds to changes in actions of users

is provided in a disaggregated way. Fishers are quite aware, for example, if the size and species distribution of their catch is changing over time.

- Better-adapted rules. Given the preceding, users are more likely to craft rules that are better adapted to each of the local common-pool resources than any general system of rules.
- Lower enforcement costs. Since local users have to bear the cost of monitoring, they are apt to craft rules that make infractions obvious so that monitoring costs are less. Further, by creating rules that are seen as legitimate, rule conformance will tend to be higher.
- Redundancy. The probability of failure throughout an entire region is greatly reduced by the establishment of parallel systems of rule making, interpretation, and enforcement.

There are, of course, limits to all ways of organizing the governance of common-pool resources. Among the limits of a highly decentralized system are:

- Some users will not organize. Although the evidence from the field is that many local users do invest considerable time and energy into their own regulatory efforts, other groups of users do not do so. There appear to be many reasons why some groups do not organize, including the presence of low-cost alternative sources of income and, thus, a reduced dependency on the resource, considerable conflict among users along multiple dimensions, lack of leadership, and fear of having their efforts overturned by outside authorities.
- Some self-organized efforts will fail. Given the complexity of the task involved in designing rules, some groups will select combinations of rules that generate failure instead of success. They may be unable to adapt rapidly enough to avoid the collapse of a resource system.
- Local tyrannies. Not all self-organized resource governance systems will be organized democratically or rely on the input of most users. Some will be dominated by a local leader or a power elite who only change rules that they think will give them further advantage. This problem is accentuated in locations where the cost of exit is particularly high and is reduced where users can leave when local decision makers are not responsible to a wide set of interests.
- Stagnation. Where local ecological systems are characterized by considerable variance, experimentation can produce severe and unexpected results, leading users to cling to systems that have worked relatively well in the past and to stop innovating long before they have developed rules likely to lead to better outcomes.
- Inappropriate discrimination. The use of identity tags is frequently an essential method for increasing the level of trust and rule conformance. However, tags based on ascribed characteristics can be the basis for excluding some individuals from access to sources of productive endeavor that has nothing to do with their trustworthiness.
- Limited access to scientific information. Although time and place information may be extensively developed and used, local groups may not have access to scientific knowledge concerning the type of resource system involved.
- Conflict among users. Without access to an external set of conflict–resolution mechanisms, conflict within and across common-pool resource systems can escalate and provoke physical violence. Two or more groups may claim the same territory and may continue to make raids on one another over a very long period of time.
- Inability to cope with larger scale common-pool resources. Without access to some larger scale jurisdiction, local users may have substantial difficulties regulating only a part of a larger scale common-pool resource. They may not be able to exclude others who refuse to abide by the rules that a local group would prefer to use. In this situation, local users have no incentives to restrict their own use and watch others take away all of the valued resource units that they have not appropriated.

The Capabilities of Polycentric Systems in Coping with Tragedies of the Commons

Many of the capabilities of a parallel adaptive system can be retained in a polycentric governance system. A polycentric system is one in which citizens are able to organize not just one but multiple governing authorities at differing scales (see McGinnis, 1999; Ostrom, 1997). Each unit may exercise considerable independence to make and enforce rules within a circumscribed scope of authority for a specified geographical area. In a polycentric system, some units are general-purpose governments and others may be highly specialized. Self-organized resource governance systems, in such a system, may be special districts, private associations, or parts of a local government. These are nested in several levels of general-purpose governments that also provide civil, equity, and criminal courts.

In a polycentric system, the users of each common-pool resource would have some authority to make at least some of the rules that are related to how that particular resource will be utilized, and thus would achieve most of the advantages of utilizing local knowledge, and the redundancy and rapidity of a trial-and-error learning process. However, problems associated with local tyrannies and inappropriate discrimination can be addressed in larger, general-purpose governmental units that are responsible for protecting the rights of all citizens and for the oversight of appropriate exercises of authority within smaller units of government. It is also possible to make a more effective blend of scientific information with local knowledge where major universities and research stations are located in larger units but have a responsibility to relate recent scientific findings to multiple smaller units within their region. Because polycentric systems have overlapping units, information about what has worked well in one setting can be transmitted to others who may try it out in their settings. Associations of local resource governance units can be encouraged to speed up the exchange of information about relevant local conditions and about policy experiments that have proved particularly successful. And, when small systems fail, there are larger systems to call on – and vice versa.

Polycentric systems are themselves complex adaptive systems without one central authority dominating all of the

others. Thus, there is no guarantee that such systems will find the optimal combination of rules at diverse levels that are optimal for any particular environment. In fact, one should expect that all governance systems will be operating at less-than-optimal levels, given the immense difficulty of fine-tuning any very complex, multitiered system.

In the US, there are many examples of dynamic, polycentric resource governance systems that display strong evidence of high performance. One example is the Maine lobster fishery, which is noteworthy because of the long-term, complementary roles adopted by both local and state governance systems. Maine is organized into riparian territories along most of the coast. Boundary rules and many of the day-to-day fishing regulations are organized by "harbor gangs".

In order to go fishing at all, one must become a member of a harbor gang, the group of fishermen who go lobstering from a single harbor. Once one has gained admittance into such a group, one can only set traps in the traditional territory of that particular harbor gang. Members of harbor gangs are expected to obey the rules of their gang concerning fishing practices, which vary somewhat from one part of the coast to another.

> There is strong statistical evidence that the territorial system, which operates to limit the number of fishers exploiting lobsters in each territory, helps to conserve the lobster resource (Acheson *et al.*, 1998: 400).

At the same time, the state of Maine has long established formal laws that protect the breeding stock and increase the likelihood that regeneration rates will be high.

> At present, the most important conservation laws are minimum and maximum size measures, a prohibition against catching lobsters with eggs, and a law to prohibit the taking of lobsters which once had eggs and were marked – i.e., the 'V-notch' law (Acheson *et al.*, 1998: 400).

Neither the state nor any of the harbor gangs has tried to limit the quantity of lobster captured. The state does not make any effort to limit the number of fishers since this is already done at a local level. However, the state has been willing to intercede when issues exceed the scope of control of local gangs. In the late 1920s, for example, when lobster stocks were at very low levels and many local areas appear to have had substantial compliance problems, the state took a number of steps – including threats to close the fishery – that supported informal local enforcement efforts. By the late 1930s, compliance problems were largely resolved and stocks had rebounded (although it cannot be shown that these two results are related, just correlated).

Recently, in response to changes that were breaking down the harbor gang system, the state formalized the system by dividing the state into zones with democratically elected councils. Each council was given authority over rules that have principally local impacts – trap limits, days and times fished, and so on. Interestingly, the formalization of local zones was followed, almost immediately, by the creation of an informal council of councils to address problems at a greater than local scale. It is expected that this council of councils will be formalized soon.

Today the state uses only approximately six patrol officers on the water to police the activities of 6800 lobstermen, all the other fisheries, and coastal environmental laws. Since the early 1990s, the lobster fishery has been growing substantially with increased yields. At the same time, there is strong evidence that the number of females of reproductive age in Maine waters is very large and that the recruitment will continue at a high level (Wilson *et al.*, 2007).

Conclusions

Widespread concern for conserving biodiversity is a reflection of the inherent and ecological value of such diversity. Those who advocate the necessity of sustaining biodiversity have frequently called for policy reforms that reduce the diversity of institutional arrangements for governing and managing complex common-pool resources. Yet, as Ashby long ago established, regulators need as much variety in their response capabilities as the systems they are regulating. To achieve biodiverse resource systems, we need to protect and enhance the institutional diversity of evolved governance regimes within the context of a broader multitiered, polycentric governance regime.

Acknowledgments

Support for writing this entry has been provided by a subcontract with the University of Gothenburg that is funded by FORMAS. My thanks to Patty Lezotte for editing this manuscript.

See also: Commons, Concept and Theory of. Government Legislation and Regulations in the United States. Market Economy and Biodiversity

References

Acheson JM, Wilson JA, and Steneck RS (1998) Managing chaotic fisheries. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience*, pp. 390–413. Cambridge, UK: Cambridge University Press.

Agrawal A and Yadama GN (1997) How do local institutions mediate market and population pressures on resources? Forest panchayats in Kumaon, India. *Development and Change* 28(3): 435–465.

Ashby WR (1960) *Design for a Brain: The Origin of Adaptive Behavior*. New York: Wiley.

Benjamin P, Lam WF, Ostrom E, and Shivakoti G (1994) Institutions, incentives, and irrigation in Nepal. *Decentralization: Finance & Management Project Report*. Burlington, VT: Associates in Rural Development.

Campbell DT (1969) Reforms as experiments. *American Psychologist* 24(4): 409–429.

Chhatre A and Agrawal A (2008) Forest commons and local enforcement. *Proceedings of the National Academy of Sciences* 105(36): 13286–13291.

Chhatre A and Agrawal A (2009) Tradeoffs and synergies between carbon storage and livelihood benefits from forest commons. *Proceedings of the National Academy of Sciences* 106(42): 17667–17670.

Coleman E and Steed B (2009) Monitoring and sanctioning in the commons: An application to forestry. *Ecological Economics* 68(7): 2106–2113.

Cox M, Arnold G, and Villamayor Tomás S (2010) A review of design principles for community-based natural resource management. *Ecology and Society* 15(4): 38. [online].

Hess C (1999) *A Comprehensive Bibliography of Common Pool Resources (CD-ROM)*. Workshop in Political Theory and Policy Analysis. (http://dlc.dlib.indiana.edu/cpr/index.php). Bloomington: Indiana University.

Holland JH (1995) *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Addison-Wesley.

Jacob F (1977) Evolution and tinkering. *Science* 196(4295): 1161–1166.

Joshi NN, Ostrom E, Shivakoti G, and Lam WF (2000) Institutional opportunities and constraints in the performance of farmer-managed irrigation systems in Nepal. *Asia-Pacific Journal of Rural Development* 10(2): 67–92.

Landau M (1969) Redundancy, rationality, and the problem of duplication and overlap. *Public Administration Review* 29(4): 346–358.

McGinnis M (ed.) (1999) *Polycentric Governance and Development: Readings from the Workshop in Political Theory and Policy Analysis*. Ann Arbor: University of Michigan Press.

Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Ostrom E, Agrawal A, Blomquist W, Schlager E, and Tang SY (1989) *CPR Coding Manual*. Workshop in Political Theory and Policy Analysis. Bloomington: Indiana University.

Ostrom E, Gardner R, and Walker JM (1994) *Rules, Games, and Common-Pool Resources*. Ann Arbor: University of Michigan Press.

Ostrom V (1997) *The Meaning of Democracy and the Vulnerability of Democracies: A Response to Tocqueville's Challenge*. Ann Arbor: University of Michigan Press.

Poteete A, Janssen M, and Ostrom E (2010) *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*. Princeton, NJ: Princeton University Press.

Schlager E (1994) Fishers' institutional responses to common-pool resource dilemmas. In: Ostrom E, Gardner R, and Walker JM (eds.) *Rules, Games, and Common-Pool Resources*, pp. 247–265. Ann Arbor: University of Michigan Press.

Singleton S (1998) *Constructing Cooperation: The Evolution of Institutions of Co-management in Pacific Northwest Salmon Fisheries*. Ann Arbor: University of Michigan Press.

Wilson JA, Low B, Costanza R, and Ostrom E (1999) Scale misperceptions and the spatial dynamics of a social-ecological system. *Ecological Economics* 31(2): 243–257.

Wilson JA, Yan L, and Wilson C (2007) The precursors of governance in the Maine lobster fishery. *Proceedings of the National Academy of Sciences* 104(39): 15212–15217.

Yoder RD (1991) Peer training as a way to motivate institutional change in farmer-managed irrigation systems. In: *Proceedings of the Workshop on Democracy and Governance*. Decentralization: Finance & Management Project Report, pp. 53–67. Burlington, VT: Associates in Rural Development.

Comparing Extinction Rates: Past, Present, and Future

Vânia Proença and Henrique Miguel Pereira, Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

© 2013 Elsevier Inc. All rights reserved.

Glossary

Background extinction rate The natural rate of extinction over geological time, measured from the fossil record and independent of anthropogenic influence.

Extinction rate The number or proportion of taxa becoming extinct per unit time.

Mass extinction Rare events in Earth's history when extinction rates escalated and extinction reached devastating magnitudes in a relatively short period of geological time. Mass extinction events affected species from a wide variety of taxa and had a global impact, affecting all regions. There were five mass extinction events in Earth history with magnitudes of 75% or more of species loss.

Species committed to extinction Species whose populations have already experienced, or will probably experience, severe population declines that compromise the viability of their genetic and demographic base and anticipate the eventual extinction.

The IUCN Red List of Threatened Species™ A list built by the International Union for Conservation of Nature (IUCN) that provides information on the global conservation status of assessed species of plants and animals.

Why to Assess and Compare Extinction Rates?

Current estimates of biodiversity range from 5 million to 30 million species (a recent estimate of Mora *et al.* (2011) suggests a value of 8.7 million terrestrial eukaryotes), but living species may represent only a small fraction, 2–4%, of all species that have ever existed on Earth (Mace *et al.*, 2005). The reasons for this small fraction of remaining species include species evolving, with species giving origin to others (this is called pseudo-extinction), species becoming extinct not only due to competitive exclusion or failing to cope in their environment, but also the occurrence of discrete extinction events associated with abrupt peaks of environmental stress (Raup, 1992; Green *et al.*, 2011).

Species extinction, like species origination, is a natural process in species evolution and diversification. The natural disappearance of species has occurred on a rather constant and low rate through geological time (Raup, 1986). This natural rate of extinction is called the background extinction rate. The background extinction rate was interrupted by five events of mass extinction (in the Ordovician, Devonian, Permian, Triassic, and Cretaceous) when extinction rates escalated and extinction reached devastating proportions of at least 75% loss of all existent species in a relatively limited time interval at the geological scale (Barnosky *et al.*, 2011). Mass extinction events affected species from a wide variety of taxa and had a global impact, affecting all regions. Biodiversity recovery after each mass extinction was a slow and long process as new species evolved during millions of years (Kirchner and Weil, 2000).

Current and recent past extinction rates, measured from records of species loss from the last few centuries, and future extinctions rates, estimated using scenarios and modeling approaches, are much higher than the background extinction rate (**Figure 1**), or in other words than the expected natural rate of species loss (Mace *et al.*, 2005; Pereira *et al.*, 2010a). This is very worrying because species extinction is a definitive and irreversible form of biodiversity loss. The increase of extinction rates is directly or indirectly driven by human activities and may have negative consequences for humankind. The impoverishment of natural communities and the disruption of ecosystem functioning are caused by several drivers, in particular habitat change, climate change, invasive species, pollution, and unsustainable use of biological resources (MA, 2005; SCBD, 2010). These drivers act together, creating combined pressures over ecosystems and showing complex dynamics at different temporal and spatial scales, which pose huge challenges to their understanding and control (Pereira *et al.*, 2010a).

The comparison between past extinction rates, estimated from the fossil record before human-era disturbances, and anthropogenic extinction rates, derived from documented extinctions and from quantitative models, is a way of discriminating human-induced extinction from natural extinctions. This allows us to put into perspective the extent of the modern biodiversity crisis and to evaluate the severity of human impacts on biodiversity. This information is essential for the prioritization and design of conservation measures that may reduce extinctions and prevent harmful and severe consequences for natural systems and humankind.

Metrics of Extinction

Metrics of Magnitude and Rate

Extinction can be measured using different metrics (**Table 1**) that can be classified in measures of magnitude or measures of rate. The first includes the number or proportion of extinctions and the latter correspond to the number or percentage of extinctions over a time interval. Using a measure of magnitude or rate depends on the purpose of the assessment. Measures of magnitude provide information on the extent of the extinction. The proportion of extinctions normalizes the number of extinctions over the total number of species in the species pool, thus controlling the effect of sample size (i.e., the species pool). However, this metric, like other measures of magnitude, lacks information on the duration of the time interval along which extinctions occurred, being therefore less suitable for comparing past and modern extinctions due to the disparity in

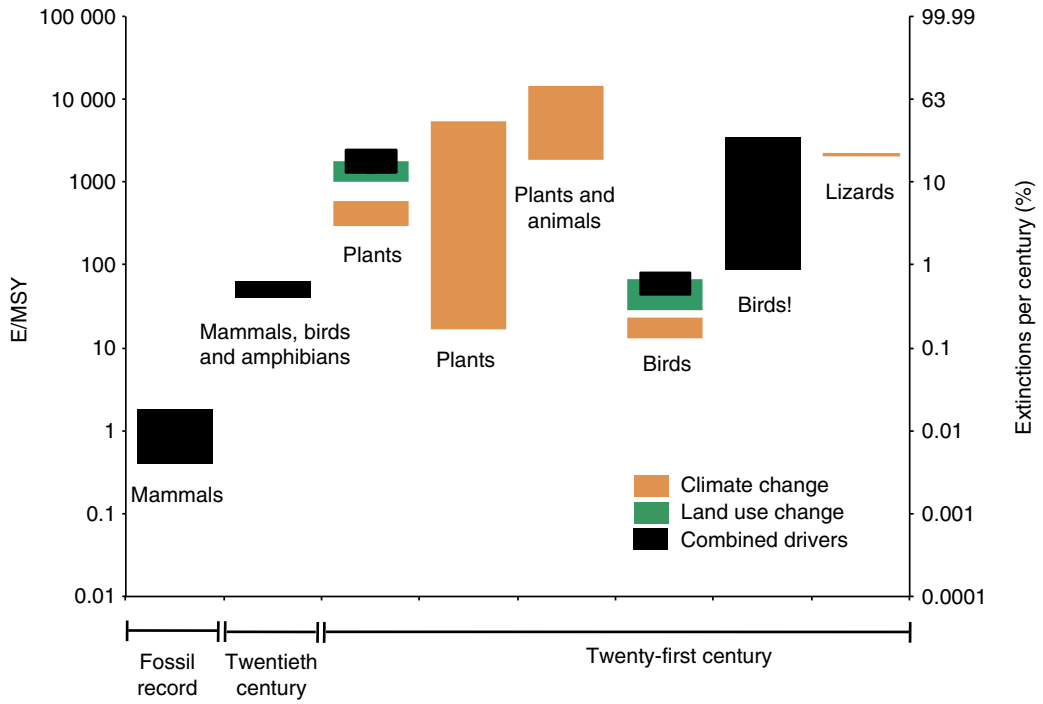


Figure 1 Past, recent, and future extinction rates. Past extinction is represented by the extinction rate of mammals in the Cenozoic fossil record (upper bound from Barnosky *et al.* (2011) and lower bound from Foote and Raupe (1996); recent extinction is represented by the extinction rate of mammals (upper bound), birds, and amphibians (lower bound) in the twentieth century, using documented extinctions (Mace *et al.*, 2005); future extinction is represented by projections of species committed to extinction in the twenty-first century, according to different global scenarios: vascular plants (Malcolm *et al.*, 2006; van Vuuren *et al.*, 2006), plants and animals (Thomas *et al.*, 2004), birds (Jetz *et al.*, 2007; Sekercioglu *et al.*, 2008), and lizards (Sinervo *et al.*, 2010). Extinction rates driven by land-use change or by climate change are discriminated when possible. Modified from Pereira HM, Belnap J, Brummitt N, *et al.* (2010b) Global biodiversity monitoring. *Frontiers in Ecology and the Environment* 8: 459–460, and Pereira HM, Leadley PW, Proença V, *et al.* (2010a) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496–1501.

Table 1 Extinction metrics and their definition

Extinction metric	*Definition*
Number of extinctions	E
Proportion of extinctions	E/S
Van Valen metric[a]	E/(S − E/2 − O/2)
Total extinction rate	E/t
E/MSY	
Proportional extinction rate (linear model)	(E/S)/t
Proportional extinction rate (exponential model[b])	− ln (1 − (E/S))/t
Van Valen metric per unit of time	(E/(S − E/2 − O/2))/t

[a]The denominator (S − E/2 − O/2) in the Van Valen metric is an estimate of standing diversity. It corresponds to the estimated diversity at the midpoint of a time interval, assuming a linear change in diversity along the interval (see Foote, 1994).
[b]For a given time interval the proportion of species going extinct is $E/S = 1 - e^{-\lambda t}$, where $\lambda = E/MSY$.
E, number of extinctions; S, total diversity in the time interval; O, number of originations; t, time interval, conventionally expressed in million years.

the length of the respective time intervals. Extinction rates constitute a more appropriate metric to compare background extinction with current extinction because extinction is expressed per time unit, thus informing on the pace of species loss. However, extinction rates may be misleading because the normalization by time assumes that extinctions are distributed evenly along the time interval, which might not be true (Foote, 1994).

Measuring Extinction in E/MSY Units

The E/MSY metric, that is, extinctions per million species-years, has been widely adopted to measure extinction rates (Pimm *et al.*, 1995; Mace *et al.*, 2005). It is a proportional extinction rate measure (**Table 1**), having the advantage of normalizing the rate of extinction both by time and size of the species pool, allowing the comparison of extinction rates between different periods of time and between species pools of different sizes. For instance, studies suggest an average background extinction rate of terrestrial mammals of 0.4 E/MSY (Foote and Raupe, 1996) to 1.8 E/MSY (Barnosky *et al.*, 2011). A value of 1 E/MSY can be interpreted as one extinction per thousand species per thousand years. Hence, when evaluating current extinction rates, a natural rate of extinction would be, for example, one species extinct in a total of 10,000 species during a time interval of 100 years.

There are several approaches for calculating E/MSY (**Table 1**): the linear model, the exponential model (also known as per capita extinction rate), and the Van Valen metric

per unit time. The linear model assumes that the proportion of extinctions, E/S, increases linearly with time. This assumption tends to be violated for long time intervals or very high extinction rates (Foote, 1994). Therefore, extinction rate estimates using the linear model are sensitive to the size of the time interval over which they are calculated, tending to decrease as the time interval increases (Foote, 1994; Regan *et al.*, 2001).

A better model of the process of extinction is the exponential model (Foote, 2000). It assumes that the lifetime of each species is given by an exponential distribution. The parameter of the exponential distribution is the per capita rate of extinction and corresponds to the value of E/MSY. Therefore, the inverse of E/MSY is the mean life expectancy of a species, which for a value of 1 E/MSY corresponds to 1 million years. The exponential model is particularly appropriate to compare different estimates of future extinction rates, as several studies report values that are extremely high (Pereira *et al.*, 2010a), which would lead to the saturation of the linear model (i.e., extrapolations of the extinction rates for the future could lead to magnitudes greater than 100% extinction).

Another alternative to the linear model is the Van Valen metric (Foote, 1994). It accounts for both speciation and extinction events (Foote, 1994), and therefore it is commonly used in paleontological studies. The exponential model can also be modified to account for both speciation and extinction (Foote, 2000).

Projected Extinction and Realized Extinction

Although past and recent extinctions refer to the disappearance of a species on the death of its last surviving individual, most projections of future extinction are based on estimates of species committed to extinction. Species committed to extinction are those whose populations have already experienced, or will probably experience, severe population declines that compromise the viability of their genetic and demographic base and anticipate the eventual extinction (Heywood *et al.*, 1994). The time needed for extinction is not certain (Sala *et al.*, 2005; Leadley *et al.*, 2010), because the time lag between impacts, population decline and realized extinction may range from decades to millennia, and in some cases declining trends may even be reverted if measures are taken to halt the causes of endangerment and restore populations.

Currently, 3663 species, from a total of 56,135 species of vertebrates, insects, mollusks, and plants that have already been assessed by the International Union for Conservation of Nature (IUCN) Red List are considered as critically endangered (IUCN, 2011), which means that these species are facing an extremely high risk of extinction in the wild, and will be committed to extinction if the present conditions are maintained or worsened (Heywood *et al.*, 1994). Some of these critically endangered species are species "possibly extinct" (Butchart *et al.*, 2006). "Possibly extinct" species are species that are likely to be extinct because they have not been seen for many years, but for which there is still a small chance that they may be extant (Butchart *et al.*, 2006). These species will only be considered as extinct after adequate efforts to prove their extinction in the wild (Butchart *et al.*, 2006).

Projections of future extinction may include the species that are currently committed to extinction (Barnosky *et al.*, 2011) or the species that might be committed to extinction in the future (Pereira *et al.*, 2010a). The number of species committed to extinction in the future may be assessed by using trends of degradation of species conservation status to extrapolate how many species will be extinct or become critically endangered in the future, or by estimating how many species will become endangered due to projected trends in drivers of change (*see* Future Extinction).

Overall, extinction rates in the future will probably be lower than what is suggested by projected extinction rates because not all species committed to extinction will become extinct. However, current extinction rates, based on recent extinction data, correspond to conservative estimates because some "possibly extinct" species will eventually be classified as extinct.

Present Extinction

Human-Induced Extinction Rates

Human-induced extinctions followed the growth and expansion of human populations across the globe. The extinction of megafauna species in the late Pleistocene (ca. 50,000 BP–12,000 BP) was at least partly a consequence of overexploitation and habitat change, although climate change may have also contributed to extinctions, especially in the northern hemisphere (Barnosky *et al.*, 2004). Human migration through land and sea lead humans to regions where biota had evolved without human disturbance. The coincidence between records of first human arrival and records of widespread extinction suggest that humans were responsible for or prompted species disappearance (Steadman, 1995). Islands provide a paradigmatic example: after the first human arrivals, 1500–2000 years ago, between 70 and 90 endemic birds were extinct in the Hawaiian archipelago, in an estimated pool of about 150 species, and 40% of large mammals were extinct in Madagascar (Pimm *et al.*, 2006). Overall, at the global level around one-third of autochthonous mammals were extinct from oceanic islands and oceanic-like islands, such as the Mediterranean islands (Alcover *et al.*, 1998).

The IUCN Red List (www.iucnredlist.org) includes more than 800 records of extinctions since the 1500s, including nearly 200 extinctions of birds and mammals (Baillie *et al.*, 2004; Mace *et al.*, 2005). Most records refer to realized extinctions and less than 10% refer to extinctions in the wild (i.e., species that disappeared from their natural habitat but still survive in captivity). Considering mammals, birds and amphibians all together, there were about 100 documented extinctions since the start of the twentieth century (Mace *et al.*, 2005), in a total pool of nearly 21,000 described species, which corresponds to an extinction rate of 48 E/MSY, which is 30–120 times greater than the background extinction rate of 0.4–1.8 E/MSY for mammals. If possibly extinct species are included, then the number of extinctions rises to 215, resulting in an extinction rate of 102 E/MSY (Mace *et al.*, 2005).

There are several sources of uncertainty that affect the assessment of current extinction magnitudes and rates. A basic

problem is that we do not know with an adequate certainty how many species exist in the world today and we only know a very small fraction of species. In addition, systematic species description only started during the eighteenth century and only about 1.7 million species have been described up to the present day (IUCN, 2011). That means that many taxa may have become extinct, or will become extinct, before being described. Another problem is that existing data are biased toward terrestrial vertebrates, which are better known and monitored than all other taxa but represent only around 2% of described species (Mace *et al.*, 2005). Moreover, several species are rare or have secretive habits, which constrains the assessment of their current status, including extinction. Finally, the time lag between impacts and extinction also limits an accurate estimation of extinction rates.

Therefore, estimates of the magnitude of contemporary extinctions, derived from documented extinctions, are very conservative, because only a small fraction of species has been described and monitored. In addition, estimates of current extinction rates are probably very conservative as well, since only realized extinctions are considered (Baillie *et al.*, 2004; Pimm *et al.*, 2006). Still, estimates of current extinction rates are already very high when compared with background extinction rates (**Figure 1**). This value will probably rise in the near future due to the continuous degradation of species conservation status, as revealed by the Red List Index (RLI) (**Figure 2**), which measures trends of species extinction risk across time (Hilton-Taylor *et al.*, 2009; Butchart *et al.*, 2010; Hoffmann *et al.*, 2010).

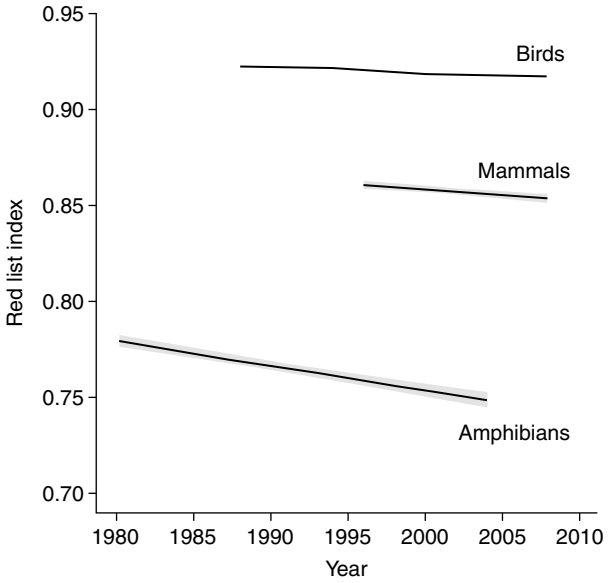


Figure 2 Changes in the Red List Index (RLI) for the period of 1980–2010, for birds, mammals, and amphibians. The RLI varies between 0 and 1: 1 corresponds to all species having a conservation status of Least Concern, and 0 corresponds to all species being extinct. A decrease in the RLI indicates that species' conservation status are worsening. Shaded areas correspond to 95% confidence intervals. Reproduced from Figure 3A. Trends in the Red List Index (RLI) for the world's birds, mammals and amphibians in Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the World's vertebrates. *Science* 330: 1503–1509.

Species extinction rates can also be estimated indirectly using data on drivers of species extinction, such as habitat loss. For example, species extinction along time after habitat fragmentation can be modeled by assessing species loss in habitat fragments of different size and age since fragmentation (for more details see Pimm and Brooks (1999)). These modeling studies tend to deliver higher rates of extinction than the conservative approach, on the range of three to four orders of magnitude higher than a background extinction rate of 1 E/MSY (Pimm and Brooks, 1999). True values of current extinction are probably found between the estimates delivered by both approaches, but nevertheless they will always be much higher than the background extinction rate.

Patterns of Modern Extinction

Patterns of contemporary extinction are not homogenous across the globe and across species groups. Islands have been the stage of most documented extinctions so far (**Figure 3**), encompassing 62% of mammal extinctions, 88% of bird extinctions, 54% of amphibian extinctions, and 68% of mollusc extinctions (Baillie *et al.*, 2004). The main causes of extinction were habitat loss, the introduction of invasive species, which includes the introduction of predators, competitors, and pathogens, and overexploitation (Baillie *et al.*, 2004). In the case of large birds and mammals, overexploitation and direct persecution were important causes of extinction. Species extinction in continents has become more frequent during the past century (**Figure 3**). Habitat change has been the main driver of species extinction in continental areas, but climate change is also becoming increasingly important (Leadley *et al.*, 2010).

The extinction record of terrestrial species is still very incomplete; even so this is the group for which present conservation status and recent extinctions are better studied. In contrast, the conservation status of freshwater and marine

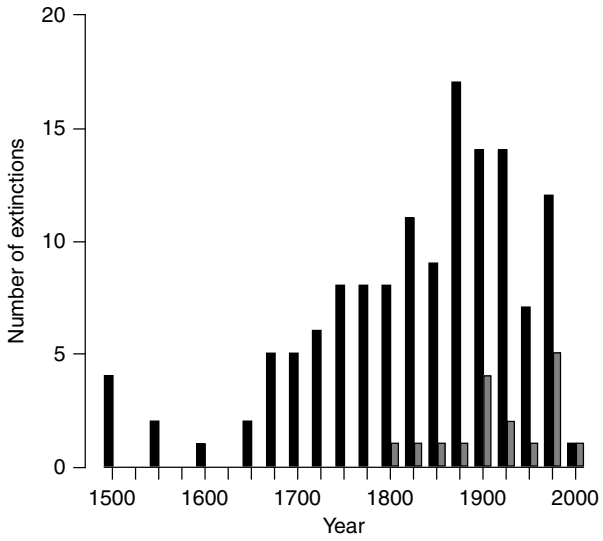


Figure 3 Bird extinctions on islands (black) and continents (gray) since 1500 AD. Reproduced from Baillie JEM, Hilton-Taylor C, and Stuart SN (2004) *2004 IUCN Red List of Threatened Species. A Global Species Assessment.* Gland: IUCN.

species is still a major unknown. There are still few records of global extinction in the marine realm, which may be due to the lack of data and adequate detection methods (Dulvy *et al.*, 2003; Baillie *et al.*, 2004). Marine species appear to be as vulnerable to human action as terrestrial species, and given the long history of human use of oceans, especially in coastal areas, one might expect a higher number of extinctions than the extinctions documented (Dulvy *et al.*, 2003). Overexploitation has been identified as the main cause of extinction of local populations, and it may cause the collapse of several world fisheries in the next decades (Worm *et al.*, 2009). Habitat destruction, pollution, and climate change are also important drivers, with severe impacts in coastal regions and in coral reefs (Dulvy *et al.*, 2003; Leadley *et al.*, 2010).

Species extinctions in freshwater systems were best documented in the USA (Baillie *et al.*, 2004). Data reveal a high magnitude of extinction, with 105 documented extinctions due to habitat destruction, pollution, and invasive alien species (Baillie *et al.*, 2004). In addition, data of freshwater fishes conservation status on the IUCN Red List indicate that more than 50% of species in the Mediterranean and Madagascar and nearly 50% of species in Europe may be threatened (Darwall *et al.*, 2009). These values are alarming, because they suggest that freshwater species could be even more threatened than terrestrial species.

A characteristic shared by many extinct species and seriously threatened species is their low total abundance, that is, a relatively small number of individuals at the global level (Pimm and Jenkins, 2010). This includes species that have limited distribution ranges and species that have large distribution ranges but are rare. In the first case, examples include the endemic species on islands, which could be rare or have high local population densities, but always have a low global abundance due to their limited distribution, and also endemic species in continental areas whose distribution is also limited. The second case includes species with large distribution ranges but with low densities and therefore with a low total abundance, this is the case with top predators, such as large carnivores.

When analyzing regional patterns of extinction, the regions that concentrate a high number of endemics are also the ones where species loss is more probable. This is the case of the biodiversity hotspots that concentrate a high diversity of endemics, and which were exposed to extensive habitat change and continue to face a severe threat of further habitat loss (Myers *et al.*, 2000; Brooks *et al.*, 2002). Extinctions may be less pronounced in regions with few endemics even when the level of impacts is considerable, such as in eastern North America, where intense deforestation since the 1600s caused the extinction of four species only from a total of 162 forest species (Pimm and Jenkins, 2010).

Future Extinction

Projecting Extinction Rates

Future extinction rates will depend on several factors, including demographic and socio-economical variables, trends in drivers of environmental change (e.g., climate change) and ecosystems and species responses to drivers of change. Research approaches to estimate future extinction rates generally include three main components: socio-economic scenarios, models to project the future state of environmental change drivers, and models to project ecosystems and species response to those drivers (Pereira *et al.*, 2010a).

Socio-economic scenarios are used to characterize changes in demographic and socio-economic variables, such as population growth, fossil fuel use, or implementation of environmental policies. Demographic and socio-economic variables act like indirect drivers of biodiversity change affecting the trajectory of environmental change drivers (e.g., climate change and land use change), which will have a direct effect on species and ecosystems. Finally, projections on the future levels of direct drivers are used to assess ecosystems and species responses to environmental changes. The response of species to drivers of change is usually estimated using phenomenological models or process-based models.

Phenomenological models are built on empirical relationships between environmental variables, such as area of natural habitat or precipitation, and biodiversity metrics, such as species richness or presence/absence of particular species. Examples of phenomenological models to assess species extinction include species–area relationships, niche-based models, and dose-response models (Leadley *et al.*, 2010). Species–area relationships are based on the known correlation between habitat area and species richness (Sala *et al.*, 2005; van Vuuren *et al.*, 2006). These models derive species extinction from the extent of predicted habitat loss. Niche-base models (or species distribution models) are based on the effect of environmental variables, namely climatic variables, on species distribution. These models use species' current distributions to define their environmental niche, or their "climate envelope" in the case of climatic variables (Thuiller *et al.*, 2008). Changes in environmental conditions will have effects on species distribution ranges, thus the future distribution of a species is estimated by combining current environmental niches with future environmental conditions. Species extinction may occur if species fail to change their distribution range in order to adapt to environmental changes, for example, due to limitations to migration, such as physical barriers or species inability to migrate. Dose-response models establish the relations between the intensity of a global change driver, which is the "dose", and the species responses (Alkemade *et al.*, 2009). An example is the decrease in plant species richness with increasing levels of nitrogen deposition.

Process-based models simulate biological processes such as population growth or mechanisms such as ecophysiological responses (e.g., changes in climatic variables affecting rates of photosynthesis, respiration and evapotranspiration in plants, or affecting demography of animal populations) to mimic species or population responses to environmental changes (Thuiller *et al.*, 2008).

Future Extinction Rates

Future global extinction rates have been projected for different taxa. Existing studies are mainly directed at terrestrial species (**Figure 1**, **Table 2**) and very few address the marine and

Table 2 Summary of global scale studies projecting terrestrial species extinction for the twenty-first century

	Plants	*Plants*[a]	*Plants and animals*	*Birds*	*Birds*	*Lizards*
E/MSY[b]						
UB (CD/LU/CC)	2489/1775/589	–/–/17	–/–/1886	51/28/23	89/–/–	–/–/2125
LB (CD/LU/CC)	1326/1003/290	–/–/5534	–/–/14,679	79/67/13	3500/–/–	
Global drivers	LU & CC	CC	CC	LU & CC	LU & CC	CC
Socio-economic scenarios	MA socio-economic scenarios[c]	GHG emission scenarios[f]	GHG emission scenarios	MA socio-economic scenarios[c]	MA socio-economic scenarios[c]	IPCC SRES emission scenarios
Drivers models	CC based on IMAGE model.[d] Habitat loss from IMAGE model based on LU and global vegetation response to climate.	Several CC scenarios.[f] Habitat loss based on two global vegetation models	Hadley centre climate model	CC based on IMAGE model.[d] Habitat loss based on LU and IMAGE model for global vegetation response to climate.	CC based on IPPC SRES scenarios.[g] Habitat loss based on LU and IMAGE model[d] for global vegetation response to climate.	Several CC scenarios. Habitat loss not considered
Species distribution model and/or species extinction model	Species–area curves	Species–area curves	Synthesis of regional projections of species ranges changes using niche-based models. Species extinction from species–area curves and IUCN status	Overlap between habitat loss and species distribution (extinction defined by 100% habitat loss/overlap)	Extinction risk based on species tolerance to CC and habitat availability.	Physiological model based on species thermal behavior
Time interval	1995–2100[e] 1995–2050	100 years	2000–2050	1985–2100[e] 1985–2050	2000–2100	1975–2080
Study	Van Vuuren *et al.* (2006)	Malcolm *et al.* (2006)	Thomas *et al.* (2004)	Jetz *et al.* (2007)	Sekercioglu *et al.* (2008)	Sinervo *et al.* (2010)

[a]Malcolm *et al.* also present projections for vertebrate species but only projections for plants are presented in Figure 1.

[b]Values of E/MSY were derived from data available in each study (see Pereira *et al.* 2010a). The upper (UB) and lower (LB) bounds correspond to the worst and best case scenarios analyzed in each study, except for Sinervo *et al.* that present a unique value for global extinction rate. Values are presented for combined drivers (CD), land use change (LU), and climate change (CC) according to available data.

[c]MA, 2005 scenarios (Carpenter *et al.*, 2005).

[d]IMAGE 2.2. model (Alcamo *et al.*, 1998, IMAGE-team, 2001).

[e]Extinctions rates presented in **Figure 1** were based on projections for the longest time interval.

[f]See Neilson *et al.* (1998).

[g]See IPCC (2007).

LU, Land use change; CC, Climate change; GHG, Greenhouse gas. Projections of species extinction are graphed in Figure 1. Studies are presented in the same order as in Figure 1.

Source: Reproduced from Leadley P, Pereira HM, Alkemade R, *et al.* (2010) *Biodiversity Scenarios: Projections of 21st Century Change in Biodiversity and Associated Ecosystem Services*. Montreal: Secretariat of the Convention on Biological Diversity, and Pereira HM, Leadley PW, Proença V, *et al.* (2010a) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496–1501.

freshwater realms (Pereira *et al.*, 2010a). Studies on terrestrial species are focused on two main drivers: land-use change and climate change. Each study uses a set of interconnected scenarios, including socio-economic, climatic and land-use change scenarios, to model extinction rates and bound estimates between potential best- and worst-case scenarios (**Figure 1**, **Table 2**). All projected extinction rates are much higher than the background extinction rate (fossil record), with several studies projecting species becoming committed to extinction at rates as high as thousands of extinctions per million species-years (**Table 2**, **Figure 1**). Even best-case scenarios, that is, the lower bounds of the intervals of variation of projected extinction rates, are projected to be above recent extinction rates in most studies (**Figure 1**).

All studies in **Table 2** except Sinervo *et al.* (2010) use phenomenological models to project the number of species committed to extinction. These studies estimate extinctions based on habitat loss due to land-use changes and/or climatic changes (e.g., van Vuuren *et al.*, 2006), or due to shifts in species ranges driven by climatic changes (e.g., Thomas *et al.*, 2004). One study, Sinervo *et al.*, 2010, uses a physiological model (process-based model) to project the number of species committed to extinction given climatic changes. The authors project lizard extinction by modeling species thermal tolerance and adapting behavior in the face of increasing environmental temperatures.

Two studies, by van Vuuren *et al.* (2006) and Jetz *et al.* (2007), analyze the separate and combined effects of land-use change and climate change on the rate of species extinction of plants and birds, respectively. Both studies predict that land-use change will cause more extinctions than climate change during the twenty-first century. However, the effects of those two drivers will not be uniform across the globe: climate change will mostly affect natural communities in high latitudes and elevated regions, whereas land-use change will be the prevailing driver of species loss in temperate and tropical latitudes, where a higher diversity of endemic and specialist species is also found.

Global studies on future extinction risk of marine species are less common. Cheung *et al.* (2009) investigated the effect of climate change on distribution ranges of marine species, fish and invertebrates, and predicted patterns of local extinction and invasion. Results point to numerous events of local extinction particularly in the subpolar regions, tropics and semi-enclosed seas, and invasions in the Artic and the Southern ocean. Species turnover will occur at a global scale with potential deleterious effects for ecosystem functioning. The authors found a global average local extinction rate of 3% by 2040–2060 of the initial species richness in 2001–2005. Despite being local extinction values, they are below most projected global extinction rates for terrestrial species. This may be partially explained by a higher dispersal ability of marine species in relation to terrestrial species.

The future extinction risk of freshwater species has been addressed in a study on the consequences of changes in river discharge, due to a decline in run-off caused by climate change and water withdrawal from human activities, in selected river basins across the globe (Xenopoulos *et al.*, 2005). Species extinction was determined at the basin scale and results suggest that 15% of the world's rivers may have more than 20% of their fish species committed to extinction during the next century.

Uncertainty in Projected Extinction Rates

The large variation in projected extinction rates, within studies and among studies, raises questions about their accuracy and suitability for comparison with past and recent rates. Pereira *et al.* (2010a) identify three main sources of uncertainty in projected extinction rates: (1) variation in the intensity of the drivers of change, (2) lack of knowledge on species ecology and species response to drivers, and (3) differences in modeling approaches and in model sensitivity.

The first factor (1) underlies much of the variation within studies. Studies test different scenarios, characterized by different degrees of land-use or climate change, to encompass a variety of possible futures and assess the effects of each scenario on biodiversity.

The second factor (2) is also a significant cause of variation in projected extinction rates within studies. Species may be able to adapt or show different levels of tolerance to changes. When the level of adaptation is weakly understood, the option is to test several alternatives that represent species responses. For example, Thomas *et al.* (2004) include species ability to migrate in their model and found a large difference in projected extinction depending on that characteristic (38% species loss with unlimited migration vs. 58% species loss with no migration).

The last factor (3) should explain most of the variation between studies, and it is related with differences in the modeling approaches used by the studies and on the criteria used to define species extinction. For example, while Jetz *et al.* (2007) estimate extinctions based on the number of species projected to lose 100% of their initial habitat from biome changes caused by climate change and land-use change, Thomas *et al.* (2004) use niche-based models to project decreases in species ranges caused by climate change and estimated extinctions based on species–area relationships.

Reducing uncertainty in projected extinction rates is critical for a more accurate comparison of extinction rates and to better inform decision makers. The evaluation of model projections, by validating models against data and using indicators to compare models outputs, and the harmonization of modeling methods are essential steps toward this objective (Pereira *et al.*, 2010a). Also important is the development of models, particularly process-based models, that integrate the effects of a broader range of drivers, and which capture the feedback of the biodiversity changes to the human system and the drivers of environmental change.

Comparing Past, Present, and Future Extinction Rates

Are We Driving the Sixth Mass Extinction?

Recent extinction rates are up to two orders of magnitude higher than the background extinction rate and future extinction rates are projected to be at least as high as current rates and likely one or two orders of magnitude higher. A

relevant question is whether mankind is driving a new mass extinction. Although past extinction rates provide a useful benchmark to evaluate the extent and severity of modern extinction rates, events of mass extinction are usually defined as the loss of at least 75% of species in a relatively short period in geological time, which provides a measure of magnitude but it is elusive regarding duration. Mass extinction events may have occurred in less than a year, along thousands of years or along millions of years (Raup, 1992; Barnosky *et al.*, 2011).

A way of dealing with this limitation is to compare simultaneously measures of rate and magnitude (Barnosky *et al.*, 2011). With this approach it is possible to assess if current rates of extinction may lead to a magnitude of extinction identical to the magnitude of past mass extinctions. Barnosky *et al.* (2011) calculated the range of variation of possible extinction rates for each event of mass extinction, using data on their magnitude and the interval of their possible duration, and then compared those values with modern rates of extinction of vertebrates considering not only species extinct in the past 500 years, but also all critically endangered and all threatened species (critically endangered, endangered and vulnerable species) that could became potentially extinct in the near future (**Figure 4**). Current rates of observed extinctions are already above the rates associated with all mass extinction events but the Cretaceous mass extinction. However, with the exception of the Cretaceous event that may have spanned from a period of less than a year to 2.5 Myr, all past events spanned for periods much larger than 500 years. If a hypothetical and fixed period of 500 years is set for comparison purposes, then modern rates of extinction will be slower than past rates, but will became dangerously similar to the rates of past mass extinctions if the (hypothetical) extinction of currently threatened species is also considered in that time span.

It is worth noting that future extinction rates derived from scenarios (**Figure 1**) are in many cases higher than hypothetical extinction rates calculated considering threatened species as inevitably extinct in a period of 500 years (**Figure 4**).

Following a different approach, one can also ask if modern extinction rates may lead, and in how much time, to a magnitude of 75% species loss. This question may be answered by testing different rates of extinction based on hypothetical assumptions, namely by assuming the loss of all threatened species or of just critically endangered species, in a period of 100 years or of 500 years, and then using those extinction rates to determine the interval of time needed to reach a magnitude of 75% loss. According to Barnosky *et al.* (2011) reaching 75% of (mammals, birds, and amphibians) species loss may take from a few centuries in a worst-case scenario, if all threatened species are lost in the next century, to about ten millennia, if only critically endangered species become extinct in the next five centuries.

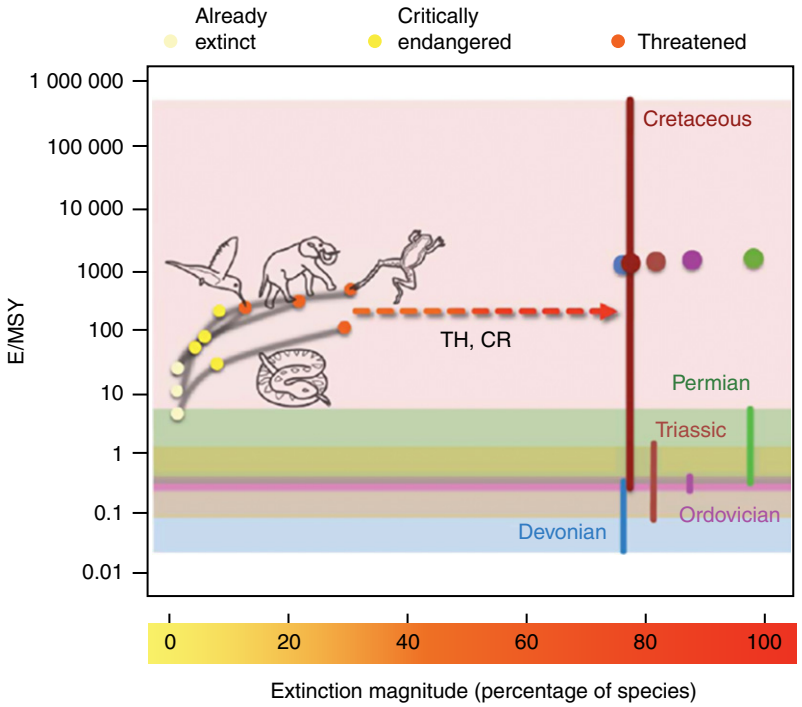


Figure 4 Extinction rate vs. extinction magnitude of past mass extinction events and of modern extinction (past 500 years). Vertical lines on the right illustrate the range of extinction rates for each event of mass extinction that would result in the magnitude associated with each event. The range of extinction rates varies accordingly with the interval of possible duration of each event. Each colored dot on the right represents the extinction rates associated with past events of mass extinction if they have occurred (hypothetically) over only 500 years. Dots and lines on the left represent modern extinction rates computed for a period of 500 years for vertebrates, using data from already extinct species (light yellow) and of species committed to extinction: critically endangered species (dark yellow) and all 'threatened' species (orange). If all threatened (TH) species became extinct in 100 years and that rate was maintained in the future, a magnitude of 75% species loss would be reached within 240–540 years. A similar scenario using the rate of critically endangered (CR) species yields an estimate of 890–2270 years. Reproduced from Figure 3. Extinction rate versus extinction magnitude in Barnosky AD, Matzke N, Tomiya S, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57, with permission from Nature.

Caveats to Comparison

Differences in criteria and in the quality of the data used to assess past, present, and future extinction rates are major sources of uncertainty limiting their comparison.

Estimates of past extinction are based on the fossil record, which is a very restricted sample of past diversity. Only a very limited (but still impressive) fraction of species left a trace of their existence in the fossil record. Overall around a quarter of a million species have been identified from a total of tens of millions of well-preserved fossils (Raup, 1992). Near-shore aquatic species are among the most common, because sediments in riverbeds, lakebeds, and in the seabed tend to provide a good medium for fossil preservation (Raup, 1992; Barnosky *et al.*, 2011). In addition, organisms with hard body parts, such as bivalves, fossilize better than soft-body organisms, such as insects. The action of decomposers and erosion erased most signs of past terrestrial organisms, and existent fossils usually correspond to individuals that were trapped and died in good preservation settings, such as volcanic ashes or tar pits (Raup, 1992). The best records of fossil vertebrates are for temperate terrestrial mammals (Regan *et al.*, 2001; Barnosky *et al.*, 2011).

The assessment of present and future extinction rates mostly relies on terrestrial vertebrate data for which there is more data available. With the exception of terrestrial mammals, the differences in the composition of past and recent extinction records limit the comparison between past and current extinction rates. Therefore, it is important to improve our knowledge on the conservation status of groups that are well represented in the fossil record but currently less known, such as near-shore marine invertebrates with shells (Barnosky *et al.*, 2011). A caveat would be that selected groups might not be good indicators of other taxa, ascribing a limited utility to comparisons.

A second problem relates with the taxonomic levels used to assess extinction rates and with the concept used to define species (Barnosky *et al.*, 2011). Fossil diversity and extinction rate are usually assessed at the genus level, and species diversity and extinction rate are then extrapolated using species-to-genus ratios, based on proportions of well-known groups. However, modern extinction rates are directly based on species data. Moreover, species in the fossil record are defined using the morphological species concept, whereas modern species are often defined using other criteria, especially the phylogenetic species concept, based on genetic similarity. The type of concept used may affect the number of species in a taxon and influence extinction rate metrics, adding uncertainty to extinction rate comparisons.

A third source of uncertainty when comparing extinction rates is that past extinction rates represent a measure of realized extinction, current extinction rates represent a conservative measure of recent extinction, due to the omission of "possibly extinct" species, and projected extinction rates may represent an overestimated measure of extinction rate, due to the inclusion of species committed to extinction.

Key Messages

Human activities and intervention over ecosystems have been causing the loss of species at rates much higher than the natural rate of extinction. Earth is experiencing a biodiversity crisis and while a sixth mass extinction is still avoidable, modern rates of extinction are already too high and alarming. Although developed regions, namely Europe, have already lost an important share of their biodiversity during millennia of human occupation and resources exploitation, other regions, namely species-rich tropical regions are now experiencing intense human pressure, which may lead to very high rates of biodiversity loss. The risks of biodiversity loss, from species loss to ecosystem degradation, are real and may jeopardize human well-being by affecting the delivery of critical ecosystem services, including food, clean water, and regulation of natural disasters.

Despite these risks our knowledge of biodiversity and of the complex interactions between drivers of change and between them and natural systems is still very incomplete, and limits our capacity of action to target and develop measures to halt biodiversity loss. There is a need for a better understanding of the impacts of drivers on biodiversity, especially of drivers that are still weakly integrated in projections of future biodiversity change (i.e., land-use change, overexploitation, pollution, and invasive species), but there is also a need for data from regions where biodiversity is now under more severe pressure. The implementation of a global network for biodiversity monitoring (Pereira *et al.*, 2010a, 2010b) that delivers data to fill knowledge gaps and to validate and calibrate current models of biodiversity change may contribute to reduce uncertainty in projections and hence to better assess impacts of drivers on biodiversity and relevant conservation measures.

See also: Extinction, Causes of. Extinction in the Fossil Record. Human Impacts on Ecosystems: An Overview. Mass Extinctions, Concept of. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Modern Examples of Extinctions

References

Alcamo J, Leemans R, and Kreileman E (eds.) (1998) *Global Change Scenarios of the 21st Century. Results from the IMAGE 2.1 Model*. Oxford, UK: Pergamon and Elsevier Science.

Alcover JA, Sans A, and Palmer M (1998) The extent of extinctions of mammals on islands. *Journal of Biogeography* 25: 913–918.

Alkemade R, van Oorschot M, Miles L, Nellemann C, Bakkenes M, and ten Brink B (2009) GLOBIO3: A Framework to investigate options for reducing global terrestrial biodiversity loss. *Ecosystems* 12: 374–390.

Baillie JEM, Hilton-Taylor C, and Stuart SN (2004) *2004 IUCN Red List of Threatened Species. A Global Species Assessment*. Gland: IUCN.

Barnosky AD, Koch PL, Feranec RS, Wing SL, and Shabel AB (2004) Assessing the causes of late pleistocene extinctions on the continents. *Science* 306: 70–75.

Barnosky AD, Matzke N, Tomiya S, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57.

Brooks TM, Mittermeier RA, Mittermeier CG, *et al.* (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology* 16: 909–923.

Butchart SHM, Stattersfield AJ, and Brooks TM (2006) Going or gone: Defining "Possibly Extinct" species to give a truer picture of recent extinctions. *Bulletin-British Ornithologists Club* 126: 7.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328: 1164–1168.

Carpenter SR, Pingali PR, Bennett EM, and Zurek MB (eds.) (2005) *Ecosystems and Human Well-Being: Scenarios*. Washington, DC: Island Press.

Cheung WWL, Lam VWY, Sarmiento JL, Kearney K, Watson R, and Pauly D (2009) Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries* 10: 235–251.

Darwall W, Smith K, Allen D, *et al.* (2009) Freshwater biodiversity: A hidden resource under threat. In: Vié J-C, Hilton-Taylor C, and Stuart SN (eds.) *Wildlife in a Changing World: An Analysis of the 2008 IUCN Red List of Threatened Species*. Gland, Switzerland: IUCN.

Dulvy NK, Sadovy Y, and Reynolds JD (2003) Extinction vulnerability in marine populations. *Fish and Fisheries* 4: 25–64.

Foote M (1994) Temporal variation in extinction risk and temporal scaling of extinction metrics. *Paleobiology* 20: 424–444.

Foote M and Raupe DM (1996) Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 22: 121–140.

Foote M (2000) Origination and extinction components of taxonomic diversity: General problems. *Paleobiology* 26: 74–102.

Green WA, Hunt G, Wing SL, and DiMichele WA (2011) Does extinction wield an axe or pruning shears? How interactions between phylogeny and ecology affect patterns of extinction. *Paleobiology* 37: 72–91.

Heywood VH, Mace GM, May RM, and Stuart S (1994) Uncertainties in extinction rates. *Nature* 368: 105.

Hilton-Taylor C, Pollock CM, Chanson JS, Butchart SHM, Oldfield TEE, and Katariya V (2009) State of the world's species. In: Vié J-C, Hilton-Taylor C, and Stuart SN (eds.) *Wildlife in a Changing World: An Analysis of the 2008 IUCN Red List of Threatened Species*. Gland, Switzerland: IUCN.

Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the World's vertebrates. *Science* 330: 1503–1509.

IMAGE-team (2001) *The IMAGE 2.2 implementation of the SRES scenarios: a comprehensive analysis of emissions, climate change and impacts in the 21st century.* RIVM CD-ROM Publication 481508018. Bilthoven: National Institute of Public Health and the Environment.

IPCC (2007) *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.

IUCN (2011) *IUCN Red List of Threatened Species. 2011.1*. Summary Statistics.

Jetz W, Wilcove DS, and Dobson AP (2007) Projected impacts of climate and land-use change on the global diversity of birds. *PLoS Biology* 5.

Kirchner JW and Weil A (2000) Delayed biological recovery from extinctions throughout the fossil record. *Nature* 404: 177–180.

Leadley P, Pereira HM, Alkemade R, *et al.* (2010) *Biodiversity Scenarios: Projections of 21st Century Change in Biodiversity and Associated Ecosystem Services*. Montreal: Secretariat of the Convention on Biological Diversity.

MA – Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press.

Mace GM, Masundire H, and Baillie JEM (2005) Biodiversity. In: Scholes R and Hassan R (eds.) *Ecosystems and Human Well-Being: Current State and Trends*, pp. 77–122. Washington, DC: Island Press.

Malcolm JR, Liu C, Neilson RP, Hansen L, and Hannah L (2006) Global warming and extinctions of endemic species from biodiversity hotspots. *Conservation Biology* 20: 538–548.

Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B (2011) How many species are there on Earth and in the ocean? *PLoS Biology* 9: e1001127.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Neilson RP, Prentice IC, Smith B, Kittel TGF, and Viner D (1998) Simulated changes in vegetation distribution under global warming. In: Watson RT, Zinyowera MC, Moss RH, and Dokken DJ, (eds.) *The Regional Impacts of Climate Change: An Assessment of Vulnerability*, pp. 439–456. Cambridge: Cambridge University Press.

Pereira HM, Belnap J, Brummitt N, *et al.* (2010b) Global biodiversity monitoring. *Frontiers in Ecology and the Environment* 8: 459–460.

Pereira HM, Leadley PW, Proença V, *et al.* (2010a) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496–1501.

Pimm S, Raven P, Peterson A, Sekercioglu CH, and Ehrlich PR (2006) Human impacts on the rates of recent, present, and future bird extinctions. *Proceedings of the National Academy of Sciences* 103: 10941–10946.

Pimm SL and Brooks TM (1999) The sixth extinction: How large, how soon, and where? In: Raven P (ed.) *Nature and Human Society: The Quest for a Sustainable World*. Washington, DC: National Academy Press.

Pimm SL and Jenkins CN (2010) Extinctions and the practice of preventing them. In: Sodhi NS and Ehrlich PR (eds.) *Conservation Biology for All*. New York: Oxford University Press.

Pimm SL, Russell GJ, Gittleman JL, and Brooks TM (1995) The future of biodiversity. *Science* 269: 347–350.

Raup D (1986) Biological extinction in earth history. *Science* 231: 1528–1533.

Raup DM (1992) *Extinction: Bad genes or bad luck?*. New York: WW Norton & Company.

Regan HM, Lupia R, Drinnan AN, and Burgman MA (2001) The currency and tempo of extinction. *The American Naturalist* 157: 1–10.

SCBD – Secretariat of the Convention on Biological Diversity (2010) *Global Biodiversity Outlook 3*. Montreal: Secretariat of the Convention on Biological Diversity.

Sala OE, Van Vuuren DP, Pereira HM, *et al.* (2005) Biodiversity across scenarios. In: Carpenter SR, Pingali PL, Bennett EM, *et al.* (eds.) *Ecosystems and Human Well-Being: Scenarios*. Washington, DC: Island Press.

Sekercioglu CH, Schneider SH, Fay JP, and Loarie SR (2008) Climate change, elevational range shifts, and bird extinctions. *Conservation Biology* 22: 140–150.

Sinervo B, Méndez-de-la-Cruz F, Miles DB, *et al.* (2010) Erosion of lizard diversity by climate change and altered thermal niches. *Science* 328: 894–899.

Steadman DW (1995) Prehistoric extinctions of Pacific island birds: Biodiversity meets zooarchaeology. *Science* 267: 1123–1131.

Thomas CD, Cameron A, Green RE, *et al.* (2004) Extinction risk from climate change. *Nature* 427: 145–148.

Thuiller W, Albert C, Araújo MB, *et al.* (2008) Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics* 9: 137–152.

van Vuuren D, Sala O, and Pereira HM (2006) The future of vascular plant diversity under four global scenarios. *Ecology and Society* 11: 25.

Worm B, Hilborn R, Baum JK, *et al.* (2009) Rebuilding global fisheries. *Science* 325: 578–585.

Xenopoulos MA, Lodge DM, Alcamo J, Marker M, Schulze K, and Van Vuuren DP (2005) Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology* 11: 1557–1564.

Competition, Interspecific

Bryan Shorrocks, University of Leeds, Leeds, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 793–811, © 2001, Elsevier Inc.

Glossary

Carrying capacity Maximum number of organisms that can be supported by a given habitat, based on the amount of resources available (such as food, nutrients, shelter, and space).

Exploitation competition Interaction among two or more species that use a common resource that is limited, in which one species benefits more than the other.

Interference competition Interaction among two or more species that use a common resource that is not limiting, in which one species is harmed by having its access to the resource restricted.

Resource partitioning Ecological arrangement in which two or more species use different, nonoverlapping resources in a given habitat, such as warblers foraging for insects in different locations within a tree or canopy.

Introduction

Interactions between species can be classified and defined in various ways. A very useful method of classification is to consider the effect that individuals of one species have on the population growth of another species and vice versa. We ask the question: In the presence of species A (+) does species B, (1) increase its numbers (+), (2) not change its numbers (0), or decrease its numbers (−) relative to when species A is absent (−)? The same question is asked of species A in the presence of species B. The answers are conveniently summarized in **Table 1**.

Because of the symmetry there are only six types of interaction. These are frequently called (1) 00 neutralism, (2) +0 commensalism, (3) + − predator/prey, parasite/host, or herbivore/plant interactions, (4) 0− amensalism, (5) − − competition, and (6)+ + symbiosis. This article focuses on interactions of types (4) and (5) between individuals of different species. Amensalism is just an extreme example of asymmetric competition. This is competition, usually between a pair of species, in which the adverse effect of one species on the other is much greater than the reciprocal effect.

The negative effects on numbers are produced by individuals directly competing for essential resources (frequently food and space) that are in short supply. If the negative effect is caused by not getting enough of the limited resource, we call it exploitation competition. If the resource is not limiting, but individuals nonetheless harm each other in the process of obtaining it, we call it interference competition. An ecological situation can arise where two species appear to show the reciprocal negative effects associated with interspecific competition, but this is in fact the result of predation by a third species. This is called apparent competition.

Consider the situation shown in **Figure 1**. A single species of predator attacks two species of prey. The predator/prey interactions are of − + type and therefore both species are adversely affected by the predator, and the predator is positively affected by both species of prey. This means that the positive effect that prey 1 has on the predator will, in turn, increase the negative effect on prey 2, and vice versa. The overall consequence of this is that there will appear to be a − − interaction between prey 1 and prey 2, even if they are not competitors for any essential and limiting resource.

How have ecologists studied competition in order to see the patterns that this ecological process produces? In general, there are four approaches (in common with all studies in ecology and, in fact, all science). These are (1) mathematical models (equations, graphs, or computer simulations), (2) laboratory experiments (laboratory models), (3) field experiments (field models), and (4) field observations. Each approach has its advantages and disadvantages. As you progress from (1) through (2) and (3) to (4), the unit of study becomes more complex but more realistic. Approach (1) helps us to understand a process (or system) and suggests what kind of (2) and (3) studies may be useful. In turn, these may direct us toward particular kinds of field observation that will help confirm or

Table 1 A classification of two species interactions

		Effect of A on B		
		+	0	−
Effect of B on A	+	++	+0	+ −
	0	0+	00	0−
	−	− +	−0	− −

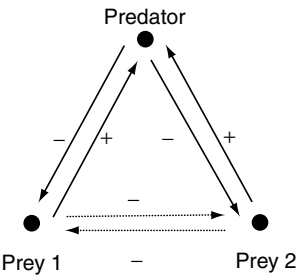


Figure 1 Diagram showing apparent competition between two prey species. The solid lines indicate direct interactions between prey and predator, and the dotted lines show indirect (apparent) interactions between the two prey species.

reject the patterns suggested by (1), (2), and (3). In fact the study of competition (its historical development) progressed in part from (1) to (4) and it is convenient to do the same in this account of interspecific competition. The remaining sections illustrate these four approaches with three examples.

Competition Models

The Lotka–Volterra Model

Vito Volterra was an Italian mathematician who, inspired by his son-in-law Umberto D'íAncona (an eminent hydrobiologist), published work on competitive interactions in the 1920s and 1930s. A. J. Lotka was an American mathematician who published on similar topics, also in the 1920s and 1930s. Their names are associated with a basic model of interspecific competition that has had a major impact on how ecologists think about this species interaction.

The Lotka–Volterra model of interspecific competition is based on two other models of population growth, the exponential growth model and the logistic sigmoid growth model. Let us look first at exponential growth, that is, when there is no competition of any kind. Under such conditions the rate at which numbers change with time (dN/dt) can be represented by a per capita (per head) rate of increase multiplied by the number of individuals (N) in the population. The rate of increase used is the intrinsic rate of natural increase (r) and is equal to the difference between the per capita birth rate and the per capita death rate. In mathematical symbols this can be written as

$$\frac{dN}{dt} = rN$$

In this simple model, the rate of increase does not change with density, that is, there is no density dependence and growth is unlimited. However, in most real populations, growth is a function of density (fN) and therefore the growth equation should be written as

$$\frac{dN}{dt} = rN(fN)$$

The simplest relationship that this function could have with density is shown in **Figure 2**. We know that the end points of this relationship are quite sensible. At very low population size ($N \cong 0$), fN must be $\cong 1$ because this would give unlimited exponential growth. At the equilibrium population size or carrying capacity, $N=K$, $fN=0$ and growth stops ($dN/dt=0$). Joining these two known points by a straight line assumes the simplest relationship possible and is not based on any empirical evidence. (Scientists, including ecologists, always develop their descriptions of the natural world by initially assuming the simplest relationships. Only when this proves inadequate do they make the explanation more complex. *See* Francisco Ayala's experiments discussed in the section Drosophila.) The slope of the line is equal to $1/K$ and the equation of this straight line is therefore $fN=1-N/K$. Substituting the value of fN into the previous equation now produces

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

which is the well-known logistic equation that produces sigmoidal limited population growth. It is important to remember that the term in brackets represents intraspecific density-dependent regulation. The inhibitory effect of one individual on population growth is $1/K$ and the inhibitory effect of N individuals on population growth is N/K.

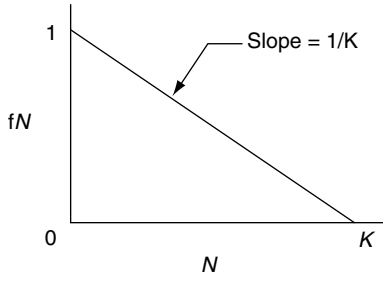


Figure 2 The density-dependent function used to produce the logistic equation.

Now let us consider two species, each of which when living separately has its growth described by a logistic equation. However, when living together this logistic equation must be modified. In the following equations the two species are denoted by the subscripts 1 and 2. Let us consider first of all the population growth of species 1 when it competes with species 2. In terms of its own population growth, as described by the logistic equation, the inhibitory effect of one individual of species 1 on its own population growth is $1/K_1$, where K_1 is the carrying capacity of species 1. The inhibitory effect of every individual (N_1) of species 1 on its own population growth is therefore N_1/K_1. If species 2 was identical to species 1 in its effect on the population growth of species 1, then we could simply write the effect of every individual of species 2 on the population growth of species 1 as N_2/K_1. However, this is unlikely. It is more realistic to imagine that each individual of species 2 would have a greater or lesser effect on the population growth of species 1, and therefore we have to multiply N_2/K_1 by a constant that expresses this difference. This constant (usually called the competition coefficient and denoted by α_{12}) describes the relative effect of an individual of species 2 compared to an individual of species 1. If an individual of species 2 was identical (from an ecological point of view), then $\alpha_{12}=1$. Similarly, the effect of every individual of species 2 on its own population growth will be N_2/K_2 and the effect of every individual of species 1 on the population growth of species 2 will be N_1/K_2 multiplied by α_{21} (note the reversal of the subscripts).

To understand the competition coefficients (α_{12} and α_{21}) more clearly, let us take an imaginary example in which competition is solely for a single food resource. Let us imagine that two species of desert mouse eat seeds. Mouse 1 consumes 100 seeds per day, and mouse 2 consumes 200 seeds per day. In terms of resource units (seeds), an individual of species 1=half an individual of species 2, or two individuals of species 1=one individual of species 2. Therefore, $\alpha_{12}=2$ (the effect of adding one more individual of species 2 is like adding two more individuals of species 1) and $\alpha_{21}=1/2$.

We can now rewrite the two logistic growth equations for the two species (which incorporate intraspecific competition)

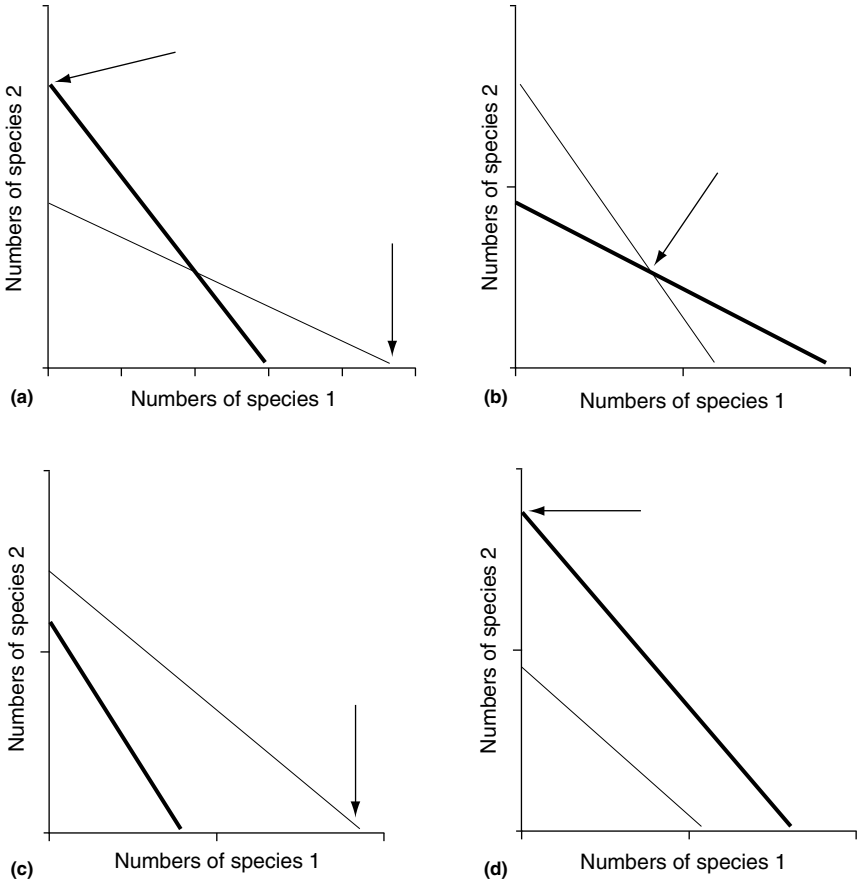


Figure 3 Graphical representation of the four predicted outcomes of the Lotka–Volterra model. Arrows indicate equilibrium points. The thin line = species 1 zero isocline and the thick line = species 2 zero isocline.

to take into account interspecific competition:

$$\frac{dN_1}{dt} = r_1 N_1 \left[1 - \left(\frac{N_1}{K_1}\right) - \left(\frac{\alpha_{12} \cdot N_2}{K_1}\right)\right]$$

$$= r_1 N_1 \left[\frac{K_1 - N_1 - \alpha_{12} \cdot N_2}{K_1}\right]$$

$$\frac{dN_2}{dt} = r_2 N_2 \left[1 - \left(\frac{N_2}{K_2}\right) - \left(\frac{\alpha_{21} \cdot N_1}{K_2}\right)\right]$$

$$= r_2 N_2 \left[\frac{K_2 - N_2 - \alpha_{21} N_1}{K_2}\right]$$

This system of two species (the Lotka–Volterra model of two-species competition) will be at equilibrium when $dN/dt = dN/dt = 0$, that is, when both species are at their carrying capacities and population growth has stopped. Just considering for the moment species 1, this means that

$$r_1 N_1 \left[\frac{K_1 - N_1 - \alpha_{12} N_2}{K_1}\right] = 0$$

Since r_1, N_1, and K_1 cannot = 0 (mathematically possible but ecologically uninteresting if species 1 cannot reproduce, is not present, or cannot exist in that habitat), this means that for species 1, $K_1 - N_1 - \alpha_{12} N_2 = 0$ at equilibrium. Rearranging this equation we get $N_1 = K_1 - \alpha_{12} N_2$, which is the equation of a straight line and is called the species 1 zero isocline. On a graph of the numbers of species 1 (N_1) plotted against the numbers of species 2 (N_2), it represents all those combinations of N_1 and N_2 for which the population growth of species 1 (dN_1/dt) = 0.

Following a similar procedure for species 2 will produce the species 2 zero isocline, which is $N_2 = K_2 - \alpha_{21} N_1$. On a graph of the numbers of species 2 (N_2) plotted against the numbers of species 1 (N_1), it represents all those combinations of N_1 and N_2 for which the population growth of species 2 (dN_2/dt) = 0.

There are four ways that these two straight lines can be placed on a graph of N_2 versus N_1, and these are illustrated in **Figure 3**. These are known as phase-plane graphs/diagrams. Each point on the graph represents a joint density of the two species. To each point in this graph we can attach an arrow, or vector, whose length and direction will indicate the dynamics of the system. Below its zero isocline, a species can increase in numbers. Above its zero isocline, a species will decline in numbers. Exactly on its zero isocline, a species will neither increase nor decrease. This means that the four graphs predict different outcomes for interspecific competition. In (a) either species can win depending on which area between the zero isoclines the population trajectory reaches. This depends on starting numbers or relative growth rates (r_1 and r_2). The point where the two lines cross represents an unstable equilibrium point. In (b) the two species coexist at the values of

N_1 and N_2 were the two lines cross. This represents a stable equilibrium point. In (c) species 1 always wins and in (d) species 2 always wins.

Laboratory examples of one species winning (c and d) and coexistence (b) are provided by G. F. Gause's experiments (*see* Yeast and *Paramecium*). Examples of one species winning (c and d) and either species winning (a) are provided by *T. Park's* experiments (*see* Flour Beetles). It is important to make two comments about the Lotka–Volterra model. First, it contains hidden variables. The competition coefficients (α) are simply conversion factors to allow one species to be represented in the other species numerical equivalents. Species may compete for food and space and poison each other, but all of these mechanisms of competition are summarized and hidden within α. Second, because the real world (unlike most laboratory environments) is heterogeneous and patchy, even (a), (c), and (d) may result in coexistence. If within a habitat patch (a), (c), and (d) apply, there may only be one species per patch. But looking over many patches, both species will be seen to apparently coexist. This is the difference between what ecologists call local and regional coexistence.

What do these graphs in **Figure 3** imply for competition and coexistence? Let us take graph (a) as an example (species 1 or species 2 wins depending on initial concentrations and/or the rates of increase). Using the end points of the lines, this graph can be defined as

$$\alpha_{12} > K_1/K_2 \quad \alpha_{21} > K_2/K_1$$

or by dividing the two sides of the left-hand inequation by K_1 and the two sides of the right-hand inequation by K_2

$$\alpha_{12}/K_1 > 1/K_2 \quad \alpha_{21}/K_2 > 1/K_1$$

In competition terms this means that the

effect of one individual of species 2 on species 1	>	species 2 on species 2	species 1 on species 2	>	species 1 on species 1

In other words, each species reduces the dN/dt of the other species more than its own dN/dt, which we can summarize as

$$2 \text{ on } 1 > 2 \text{ on } 2 \quad 1 \text{ on } 2 > 1 \text{ on } 1$$

Interspecific effects for both species are stronger than intraspecific effects.

For graph (b) (coexistence) these inequalities are

$$2 \text{ on } 1 < 2 \text{ on } 2 \quad 1 \text{ on } 2 < 1 \text{ on } 1$$

which is the opposite of graph (a). Here each species reduces its own dN/dt more than that of the other species, and therefore interspecific effects for both species are weaker than intraspecific effects. This case has always intrigued ecologists, and the traditional explanation as to why interspecific effects for both species should be weaker than intraspecific effects is resource partitioning (*see* Resource Utilization Curves). However, there are other mechanisms that will produce this effect (*see* The Aggregation Model).

For graph (c) (where species 1 always wins) the inequalities are

$$2 \text{ on } 1 < 2 \text{ on } 2 \quad 1 \text{ on } 2 > 1 \text{ on } 1$$

which implies that the two species now do different things. Species 2 reduces its own dN/dt more than that of species 1, and species 1 reduces the dN/dt of species 2 more than its own dN/dt.

For graph (d) (where species 2 always wins) the inequalities are

$$2 \text{ on } 1 > 2 \text{ on } 2 \quad 1 \text{ on } 2 < 1 \text{ on } 1$$

and we have the opposite of graph (c).

Resource Utilization Curves

The Lotka–Volterra competition model predicts that two potential competitors will coexist if they each affect the growth of the other more than they affect their own growth. That is, intraspecific effects are greater than interspecific effects for both species. This is summarized by the inequations

$$\alpha_{12}/K_1 < 1/K_2 \quad \alpha_{21}/K_2 < 1/K_1$$

Traditionally this led to the concept of resource separation (different niches) as a mechanism of coexistence. Clearly, if two species use different resources they will interact (compete) only with their own species, and therefore the preceding inequalities will be true. But partial resource partitioning has also been used widely as an explanation for coexistence. Field observations in particular (*see* Field Observations) have frequently found partial resource separation between species and accumulated these differences over several niche dimensions to convince the reader that enough separation exists to explain the coexistence observed. Unfortunately, most of these field studies do not calculate α_{12}, α_{21}, K_1, or K_2 and therefore we do not know if sufficient separation exists between the species to make $\alpha_{12}/K_1 < 1/K_2$ and $\alpha_{21}/K_2 < 1/K_1$. However, one attempt to quantify the amount of separation required is provided by a simple model, initially developed by Robert MacArthur (1972).

Imagine two species competing for a resource that can be visualized as varying continuously in one dimension, for example, insects or seeds of different sizes or foraging sites up a tree. Imagine also that the two species' utilization of this resource follows a unimodal distribution, with a preferred area of use in the center and less preferred resource items on either side. **Figure 4** shows this idea for two species using a resource j, in which their resource utilization curves (u_{1j} and u_{2j}) are described by a normal distribution. Each species uses a

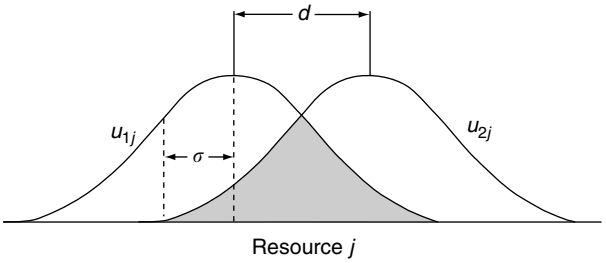


Figure 4 Resource utilization curves used in the MacArthur model. D is the difference between the means of the two normal distributions and σ is the standard deviation for the two normal distributions.

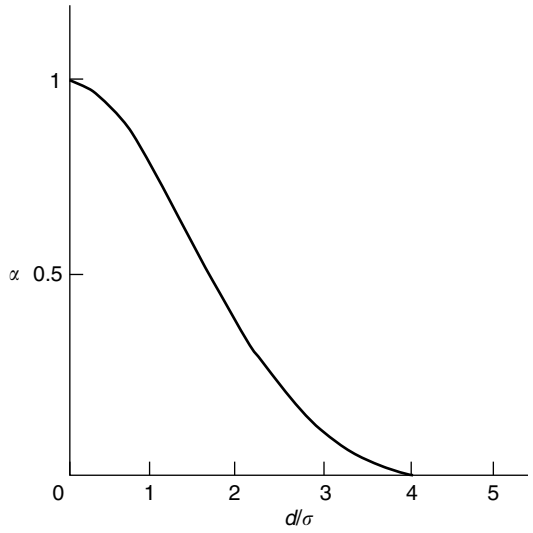


Figure 5 Relationship between α and the ratio d/σ predicted from the MacArthur model.

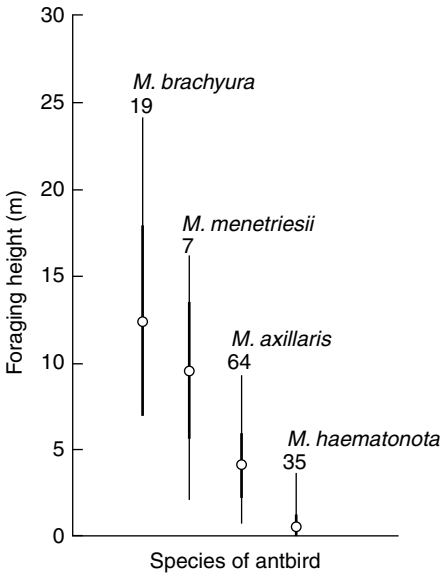


Figure 6 Foraging height of antbirds (Formicariidae). Notice that the means are separated by about one standard deviation, as predicted by the MacArthur model.

different part of the resource, with some overlap. The distance between the two means is d, and σ is the standard deviation. The two species are assumed to have the same carrying capacity and the intensity of competition is related to the area of overlap between the two distributions. With two normal distributions this is given by

$$\alpha = e^{-d^2/4\sigma^2}$$

and the relationship between α and d/σ is shown in **Figure 5**. For stable coexistence between two species, d/σ should be approximately greater than unity, that is, d (the distance between means) should be greater than σ (the standard deviation of resource utilization). A field example that agrees with this amount of limiting similarity between species is provided by the foraging height relationships within antbirds (*Myrmotherula*), shown in **Figure 6**. Unfortunately, it has been pointed out that $d/\sigma > 1$, and promoting coexistence may be a property of this type of model only with a one-dimensional resource. With competition in several dimensions and alternative resource utilization curves, lower values of d/σ may be compatible with stable coexistence.

The Aggregation Model

Many insects exploit resources that are patchy, consisting of small, separate units, and that are ephemeral in the sense that they persist for only one or two generations. Such resources can include fruit, fungi, sap flows, decaying leaves, flowers, dung, carrion, seeds, dead wood, and small bodies of water held in parts of terrestrial plants (phytotelmata). This general view of insect ecology inspired the aggregation model of competition (Shorrocks *et al.*, 1984; Atkinson and Shorrocks, 1981), which allows a competitively inferior species to survive in probability refuges. These are patches of resource (a single fungus, fruit, etc.) with no or a few superior competitors that arise because the competing stages (usually larvae) have an aggregated distribution across the patches. These probability refuges are a permanent feature of such systems because patches, such as fungi, are ephemeral and aggregation increases mean crowding. Regional population density is limited by strong intraspecific competition in patches with high local density whereas low-density patches still exist (e.g., population size within a wood is limited by high density in some fungi, whereas other fungi still contain no or a few individuals). As with resource partitioning, coexistence is promoted because aggregation of the superior species increases its intraspecific competition and reduces interspecific competition.

In the aggregation model, the eggs of both insect species are independently distributed over the patches according to a negative binomial distribution, which has an exponent, k, inversely related to the degree of intraspecific aggregation. The use of the negative binomial and the assumption of independence have been justified for drosophilid flies. In the first version of the model, the parameter k (level of aggregation) was constant and independent of density. This is not valid for real populations, but relaxing this assumption does not prevent coexistence. Within each patch, competition is modeled by a difference form of the Lotka–Volterra equations.

The predictions of the aggregation model are that with k of the negative binomial < 1 (strong aggregation), it is virtually impossible for the competitively superior species to eliminate the competitively inferior species. **Figure 7** shows the model's results as a graph of "critical α" against k of the negative binomial. Critical α is the competition coefficient that the superior species must have in order to exclude the inferior species. Also shown on the graph are distributions of α and k for drosophilid flies. For these flies it is clear that k of the negative binomial is usually < 1 and that competition coefficients are not sufficiently large to prevent competitive exclusion. For many animals exploiting ephemeral and patchy resources, this model therefore provides a viable alternative to

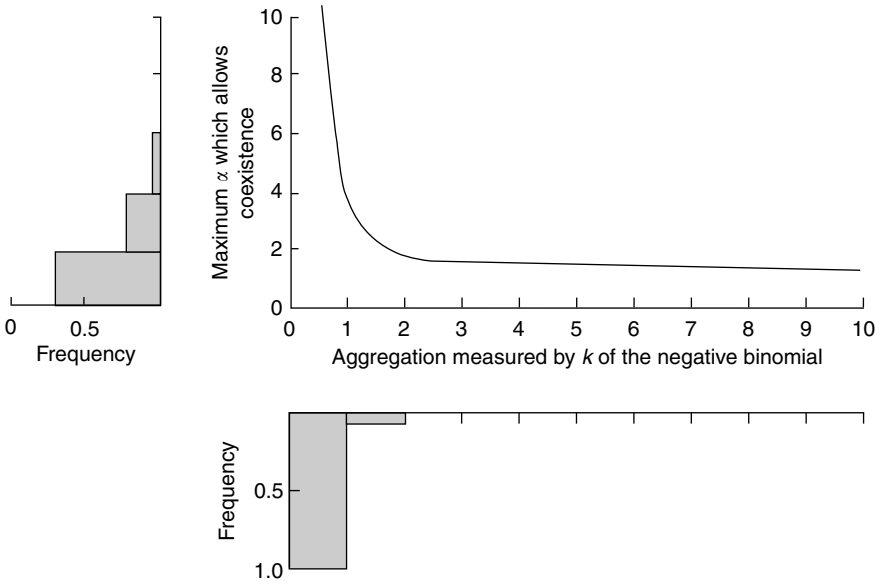


Figure 7 Relationship between "critical α" and aggregation (k of the negative binomial) predicted by the aggregation model.

traditional resource partitioning as an explanation for the coexistence of species. The two-species model has been extended to a many-species model (Shorrocks and Rosewell, 1987) and predicts average group sizes of about seven species coexisting on identical resources.

Laboratory Experiments

Yeast and Paramecium

In the early part of this century, G. F. Gause carried out a series of simple experiments with yeasts and protozoa that are instructive, are of historic interest, and set the experimental protocol for many of the so-called "bottle" experiments conducted in the 1940s, 1950s, and 1960s. His early experiments are detailed in his classic book *The Struggle for Existence* (Gause, 1934). He was mainly interested in testing the Lotka–Volterra model that had been put forward as a description of the competitive process, between species, a few years earlier.

The experimental design was simple. He grew two species, each alone and then together in mixed culture, to determine the course of competition under laboratory conditions. In his first experiments he used two species of yeast, which he called *Saccharomyces* and *Schizosaccharomyces*. However, we are not now sure of the exact taxonomic identity of the two species. He grew these species on an extract of brewer's yeast, water, and sugar. The growth medium was not renewed, but growth ceased before the sugar resource was exhausted, apparently because of the accumulation of alcohol. These experiments with yeast are therefore an example of interference competition rather than exploitation competition. The main interest of these experiments lies in the calculation of the competition coefficients (α), which Gause did by estimating all the variables in the Lotka–Volterra competition equations except α. The intrinsic rate of natural increase (r_1 and r_2) and the carrying capacity (K_1 and K_2) were estimated by fitting a logistic equation to the single-species growth curves. From the mixed cultures, he estimated both the numbers of each species (N_1 and N_2) and the growth rates (dN_1/dt and dN_2/dt). By rearranging the Lotka–Volterra equations he obtained

$$\alpha_{12} = \frac{K_1 - \dfrac{dN_1/dt \cdot K_1}{r_1 \cdot N_1} - N_1}{N_2}$$

$$\alpha_{21} = \frac{K_2 - \dfrac{dN_2/dt \cdot K_2}{r_2 \cdot N_2} - N_2}{N_1}$$

Since all quantities on the right-hand side are known, he could estimate the competition coefficients (α_{12} and α_{21}). Since N_1, N_2, dN_1/dt and dN_2/dt all vary with time, Gause selected three occasions during his experiments to estimate α. He then took the mean of these three calculations. His experiments were also repeated under two environmental conditions, which he called anaerobic (yeasts grown in test tubes) and aerobic (yeasts grown in flasks). Since α is a measure of the effect that one individual of species 1 has on the population growth of species 2 (and vice versa), and with these species competition is thought to be mainly the result of alcohol poisoning, it should be possible to calculate α directly from the relative production of alcohol by the two species. **Table 2** shows the values of α calculated from population experiments and relative alcohol production, under both anaerobic and aerobic conditions.

Two important points come out of these yeast experiments. First, the competition coefficients are not constant; they vary (quite markedly) with different environmental conditions. Second, relative alcohol production and its effect on growth are sufficient to explain the interference competition between these two species (particularly under aerobic conditions).

Therefore, in this experiment, the competition coefficients are relatively simple, pure quantities and $\alpha_{12} \cong 1/\alpha_{21}$

In his later and perhaps more famous experiments, Gause used two species of single-celled protozoa called *Paramecium*. Initially he used what he called *P. caudatum* (1) and *P. aurelia* (2), although once again we are not sure about the precise taxonomy because these species are now known to be closely related groups of species. The *Paramecium* were grown in 5 cm^3 of Osterhaut's medium (a mixture of salts) in test tubes at 26 °C. Each day fresh food was added to the test tubes. The food was a bacterium grown on agar plates and added to the cultures using a platinum loop (to standardize the amount given). The bacterium used did not grow in Osterhaut's medium. Each experiment was started with 20 *Paramecium* individuals on Day 0, and Gause estimated the number of individuals each day by sampling 10% of the medium. Since an individual of *P. aurelia* occupies only 39% of the volume of a single *P. caudatum*, he estimated the volume of a typical cell for each species and converted the estimated numbers to biomass (in volume). Following the same protocol as the yeast experiments, the two species were grown alone (four cultures of *P. caudatum* and three cultures of *P. aurelia*) and in mixed culture (three cultures); **Figure 8** shows the mean volumes, over time, for these experiments. In single culture, both species appear to grow to an equilibrium at around $K = 200$. In mixed culture, *P. caudatum* appears to suffer more from interspecific competition than does *P. aurelia*, although the latter also grows less well. Using the same procedure as for the yeast experiments, Gause was able to estimate the competition coefficients (α_{12} and α_{21}), and these are given in **Table 3**.

Notice that in the first few days the competition coefficient for the effect of *P. aurelia* on *P. caudatum* is negative. This implies that the effect of *P. aurelia* is actually positive at this stage. Given that $K_1 = K_2 = 200$, $\alpha_{12} = 1.64$, and $\alpha_{21} = 0.61$, then we can calculate $K_1/\alpha_{12} = 122$ and $K_2/\alpha_{21} = 328$ and therefore construct the Lotka–Volterra phase-plane for this experiment.

Table 2 Competition coefficients (α) calculated from yeast population experiments and relative alcohol production, under both anaerobic and aerobic conditions

	Experiments	*Alcohol*
Anaerobic		
α_{12}	3.15	2.19
α_{21}	0.44	0.46
Aerobic		
α_{12}	1.25	1.25
α_{21}	0.85	0.80

Table 3 Competition coefficients calculated from Gause's *Paramecium* experiments

	α_{12}	α_{21}
First day	−1.00	+0.50
Greater than fifth day	+1.64	+0.61

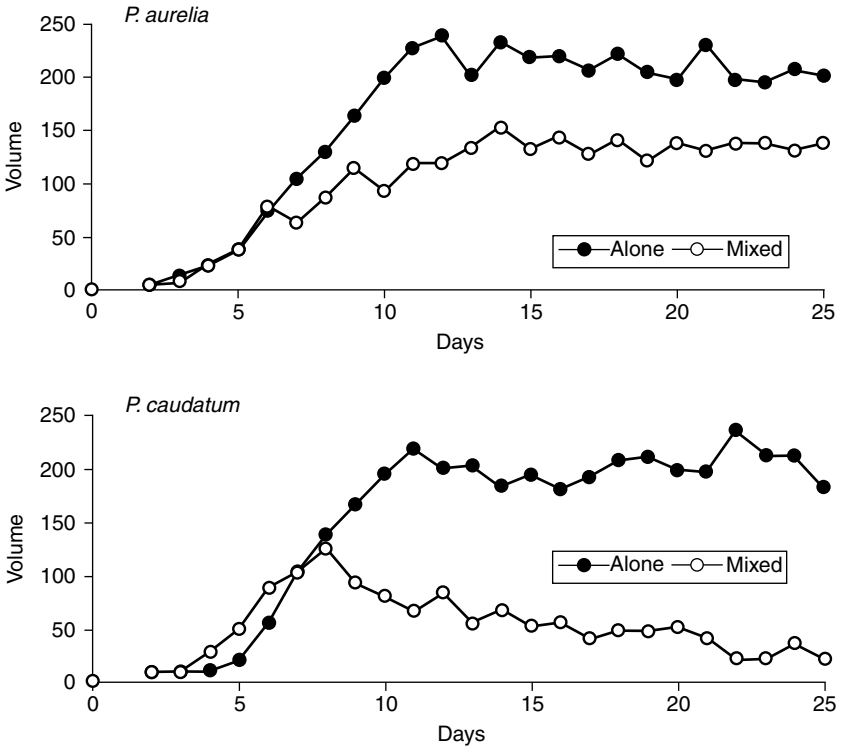


Figure 8 Growth in volume of *Paramecium caudatum* and *P. aurelia* in both single and mixed culture (from Gause, 1934). Volumes are converted from number of individuals per 0.5 cm^3 in the culture. Since the volume of *P. caudatum* was set at 1.00 and that of *P. aurelia* at 0.39, the volume for *P. caudatum* is equal to the number of individuals per 0.5 cm^3.

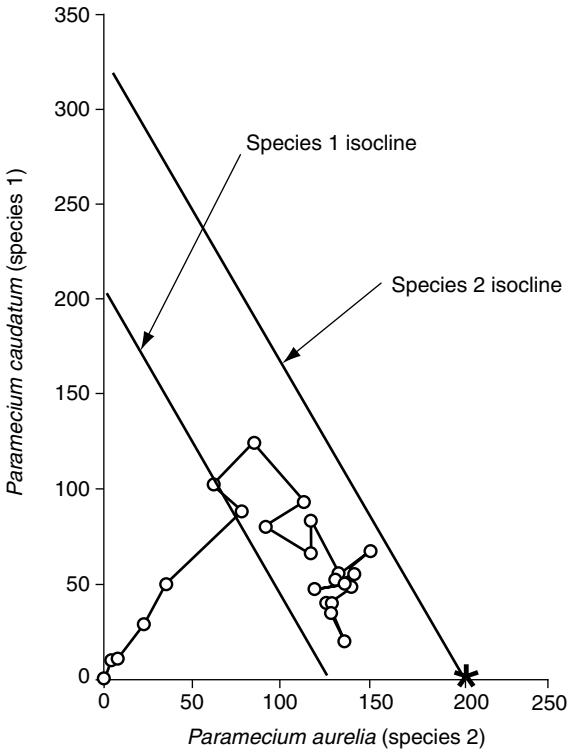


Figure 9 Phase-plane graph showing the experimental trajectory from **Figure 8**. Numbers are actually the "volume" of each species. Intercepts (and therefore zero isoclines) are worked out from Gause's experimental data (see text). The predicted equilibrium is indicated by an asterisk.

This is shown in **Figure 9**, along with the two-species population trajectory for the results in **Figure 8**. These experiments with *Paramecium* appear to broadly confirm the predictions of the Lotka–Volterra model of interspecific competition, although the mortality imposed by the sampling procedure (10%) and the premature termination of the experiments result in the predicted equilibrium not being reached in 25 days. However, later experiments in which Osterhaut's medium was buffered to a pH of 8 and the daily food concentration was reduced resulted in *P. caudatum* being excluded from the cultures in just 16 days.

In another set of experiments using *P. aurelia* and *P. bursaria*, Gause obtained population trajectories that suggested coexistence, although again the experiments were not run for long enough to reach an actual equilibrium. However, irrespective of the starting numbers, the two-species population trajectories converge on the same joint densities. *Paramecium bursaria* is interesting since it contains within its cell body unicellular green plants. The suggestion has been made that *P. bursaria* was able to exploit the bacterial food that had settled at the bottom of the experimental tube. This area of the microhabitat is depleted of oxygen (by the bacteria) and cannot be used by *P. aurelia*. However, *P. bursaria* carries its own oxygen supply (produced by its unicellular plants) and can use this part of the bacterial resource. Heterogeneity within this simple habitat, produced by the bacterial food, may have been sufficient to allow resource partitioning.

Table 4 Average densities of beetles observed in Park's single species experiments

	34 °C	*29 °C*	*24 °C*
T. confusum (1)			
70%	41	33	29
30%	24	30	30
T. confusum (2)			
70%	30	50	45
30%	10	19	3

Table 5 Comparison of Park's single-species beetle experiments

	34 °C	*29 °C*	*24 °C*
70%	$1>2$	$2>1^*$	$2>1^*$
30%	$1>2^*$	$1>2^*$	$1>2^*$

Asterisks indicate that *T. confusum* and *T. castaneum* maintained significantly different population sizes.

Flour Beetles

In a series of studies, Thomas Park carried out competition experiments between two species of flour beetle, *Tribolium confusum* (1) and *T. castaneum* (2). Both are frequently found as stored-product pests and they can be conveniently raised under laboratory conditions. Populations were kept in glass tubes with 8 g of medium (95% flour and 5% yeast). All adults, pupae, and larvae were counted every 30 days and both single- and mixed-species populations were kept. An interesting aspect of the experimental design, which proved instructive, was that populations were maintained at a number of temperatures (24, 29, and 34 °C) and relative humidities (30% and 70%), and each type of population (single or mixed), under all six environmental conditions, was replicated 30 times.

Single-Species Results

Apart from *T. castaneum* at 24 °C and 30% humidity, all single-species populations persisted, and **Table 4** shows the average population densities during the course of the experiment. The value of 3, for *T. castaneum* at 24 °C and 30% humidity, is the average population density until extinction.

Using these average densities as an indication of success, Park compared the relative performance of these two species (alone) for each temperature/humidity combination. The results of this comparison are shown in **Table 5**, where an asterisk indicates a statistically significant difference between the two species. Except for 34 °C and 70% humidity, each of the physical environments significantly favors one species over another. These results are due to intraspecific competition and the effect of the physical environment on it.

Two-Species Results

Park found that while single-species populations persisted (with the exception of *T. castaneum* at 24 °C and 30% humidity), mixed-species populations did not. In each population only one species persisted and the other was eliminated. However, unlike most previous bottle experiments

(e.g., Gause's experiments), it was not always the same species that was eliminated, even under the same environmental conditions. This "indeterminate" outcome of competition is one of the predicted possible outcomes of the Lotka–Volterra model. **Table 6** shows the percentage of replicates that *T. castaneum* won.

These experiments illustrate two interesting and possibly universal features of competition. First, the outcome of two-species competition can sometimes be predicted from the performance of each species on its own (i.e., 29 °C/70% and 24 °C/30%), but also *sometimes it cannot* (i.e., 34 °C/70% and 24 °C/70%). Second, the outcome of competition between two species can vary along an environmental gradient. Park's experiments go from hot–humid conditions (34 °C/70%) to cool–dry conditions (24 °C/30%). At the hot–humid end, *T. castaneum* eliminates *T. confusum* each time. At the cool–dry end, *T. confusum* eliminates *T. castaneum* each time. In the middle region either species can win, but *T. castaneum* wins more often toward the hot–humid end, and *T. confusum* wins more often toward the cool–dry end. If we imagine that the isoclines in the Lotka–Volterra phase-plane graphs can move their position depending on environmental conditions, then these *Tribolium results are still understandable in terms of the Lotka–Volterra model*. **Figure 10** presents a series of phase-plane graphs, going from hot–humid (left) to cool–dry (right). The arrow(s) in each diagram indicate the equilibrium point. Notice that the species predicted to win moves from always *T. castaneum* (left), to either species, to always *T. confusum*.

Drosophila

In the late 1960s and early 1970s, Francisco Ayala conducted a series of laboratory experiments, each using a pair of *Drosophila* species. Although five species were used in different combinations, the basic design and outcome of all of these experiments were very similar. One experiment, between *D. serrata* and *D. pseudoobscura*, will be sufficient to illustrate the apparent dilemma that emerged.

Drosophila serrata comes from Australia and *D. pseudoobscura* comes from North America. Notice therefore that since neither of these flies comes from the same continent, let alone the same community, Ayala was simply using the flies like analog computers to find out what might happen between flies in real communities.

At 25 °C, *D. serrata* eliminates *D. pseudoobscura* in a few generations; at 19 °C, *D. pseudoobscura* eliminates *D. serrata*. Ayala performed his experiments at 23.5 °C, at which temperature the two species coexisted. Populations were started with 300 adult flies of each species and were maintained in a series of milk bottles with standard *Drosophila* medium (a kind of yeasty porridge). Adult flies feed off the yeasty surface of the medium, female flies lay their eggs onto the surface, and larvae feed within it. Third instar larvae pupate on the surface of the medium or on the side of the bottle. When an apparent equilibrium was reached in the mixed populations, single-species populations were established and maintained under the same conditions. From the population data, and assuming that the Lotka–Volterra model of competition describes what is happening between the two species, Ayala was able to calculate the competition coefficients, α_{12} and α_{21}. He did this in the same manner as Gause, by rearranging the Lotka–Volterra equations:

$$\alpha_{12} = \frac{K_1 - \dfrac{dN_1/dt \cdot K_1}{r_1 \cdot N_1} - N_1}{N_2}$$

$$\alpha_{21} = \frac{K_2 - \dfrac{dN_2/dt \cdot K_2}{r_2 \cdot N_2} - N_2}{N_1}$$

However, by assuming that the two-species populations had reached an equilibrium ($dN_1/dt = dN_2/dt = 0$), these equations can be reduced to

$$\alpha_{12} = \frac{K_1 - N_1}{N_2}$$

$$\alpha_{21} = \frac{K_2 - N_2}{N_1}$$

where K_1 and K_2 equal the equilibrium population densities of each species on their own and N_1 and N_2 equal the

Table 6 The percentage of replicates that *T. castaneum* won in Park's beetle experiments

	34 °C	*29 °C*	*24 °C*
70%	100	86	29
30%	10	13	0

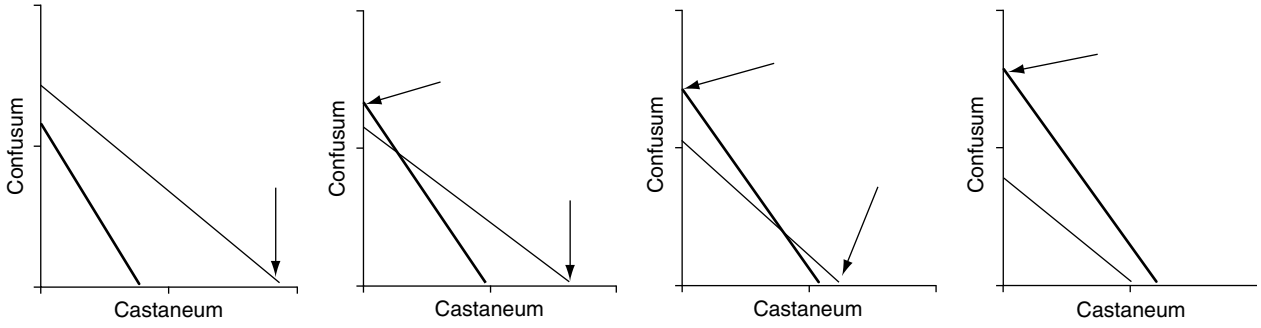


Figure 10 A series of phase-plane graphs along an environmental gradient from hot–humid (left) to cool–dry (right). The arrow(s) in each diagram indicate the equilibrium point.

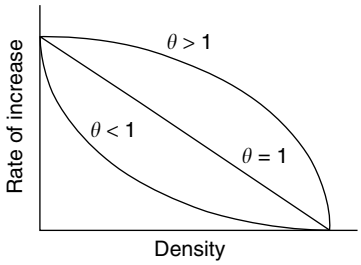


Figure 11 The nature of density dependence in Ayala's model.

equilibrium population densities for each species in the mixed populations.

All the pairs of species that Ayala studied coexisted for many generations, however, in every case he found that

$$\alpha_{12} > \frac{K_1}{K_2} \text{ and } \alpha_{21} > \frac{K_2}{K_1}$$

This is not the "case for coexistence" in the Lotka–Volterra model of competition. Intriguingly, Ayala did not conclude that the model was incorrect, but rather that the species could coexist even if they utilized the same resources.

Of course, the Lotka–Volterra model incorporates the logistic equation, which assumes (as a first approximation) that the relationship between rate of increase and density is a straight line (see **Figure 2**). However, Gilpin and Justice examined the relationship between birth/death and density in Ayala's experiments and found that for these *Drosophila* it was frequently a curve. It is not surprising therefore that an apparent conflict had occurred between theory and practice. This led Gilpin and Ayala (1973) to develop a new Lotka–Volterra model that incorporated such "nonlinear" relationships. Their new model of competition has four parameters, one more than the Lotka–Volterra model, and explains 95% of the variance in Ayala's experiments. For two species, the model can be written as

$$\frac{dN_1}{dt} = r_1 N_1 \left[1 - \left(\frac{N_1}{K_1}\right)^{\theta_1} - \left(\frac{\alpha_{12} \cdot N_2}{K_1}\right)\right]$$
$$\frac{dN_2}{dt} = r_2 N_2 \left[1 - \left(\frac{N_2}{K_2}\right)^{\theta_2} - \left(\frac{\alpha_{21} \cdot N_1}{K_2}\right)\right]$$

The new parameter (θ) describes the nonlinear nature of density dependence. When $\theta = 1$, the new model is identical with the Lotka–Volterra model. **Figure 11** shows the nature of density dependence when $\theta \neq 1$, < 1, and > 1. These modified Lotka–Volterra equations give zero isoclines, on an N_1-by-N_2 graph, that are curves rather than straight lines. This leads to the intriguing possibility that more than one equilibrium may exist for a pair of competing species (**Figure 12**).

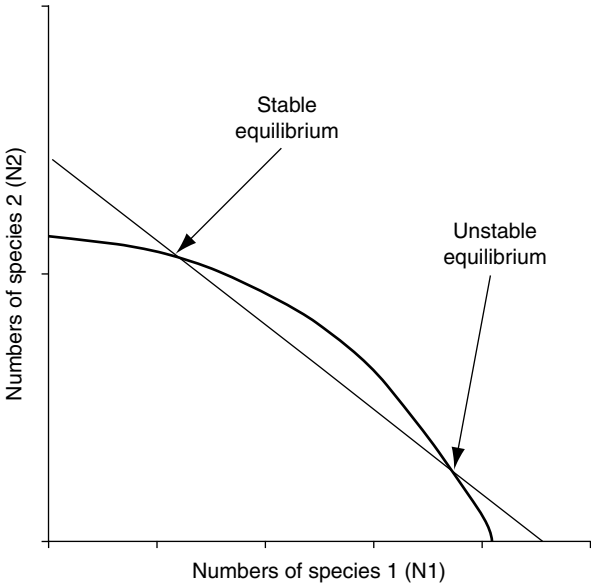


Figure 12 Two equilibrium points resulting from curved zero growth lines derived from the Gilpin–Ayala model.

Field Experiments

Barnacles

Barnacles compete for space in the intertidal zone of rocky shores. Because they are attached to the rock, they are ideal animals for experimental manipulation in the field. In the 1950s, Joseph Connell conducted a classic set of removal experiments to test for competitive interactions between two species of barnacle, *Chthamalus stellatus* and *Balanus balanoides*. The experiments were carried out on the Isle of Cumbrae in the Firth of Clyde, Scotland.

These two species occupy two separate horizontal zones (with a small area of overlap), with *Chthamalus* (which is more resistant to desiccation) higher up the shore. *Chthamalus* larvae settle on the shore in September/October, whereas *Balanus* settle in April/May. Connell selected eight areas for study, on different parts of the shore, and used old glass lantern slides (10.7 cm × 8.2 cm) as quadrats on which he could mark the position of all individuals of both species of barnacle.

After the settlement of *Balanus* had stopped in early June, a census of the *Chthamalus* was made. Each quadrat was divided into two halves and from one randomly chosen half all the *Balanus* individuals were removed. Survival of the *Chthamalus* individuals in both halves of the quadrats was then monitored for 12 months. **Figure 13** shows three representative results from three of the quadrats. What is quite clear is that survival of *Chthamalus* is much better in those halves of the quadrats without *Balanus*. Direct observation showed that this was due to interference competition from *Balanus* individuals who smothered, undercut, or crushed *Chthamalus* individuals. Those *Chthamalus* individuals in the *Balanus* half of the quadrat that survived the competition were smaller than those that had not been subjected to this interspecific competition, illustrating that competition affects fecundity as well as survival.

Ants and Mice

Seeds play a major role in the ecology of desert regions and the seeds of annual plants are the primary food of several distantly related taxa of specialized granivores, such as rodents, birds, ants, and beetles. In particular, rodents and ants are very similar in their utilization of seed resources.

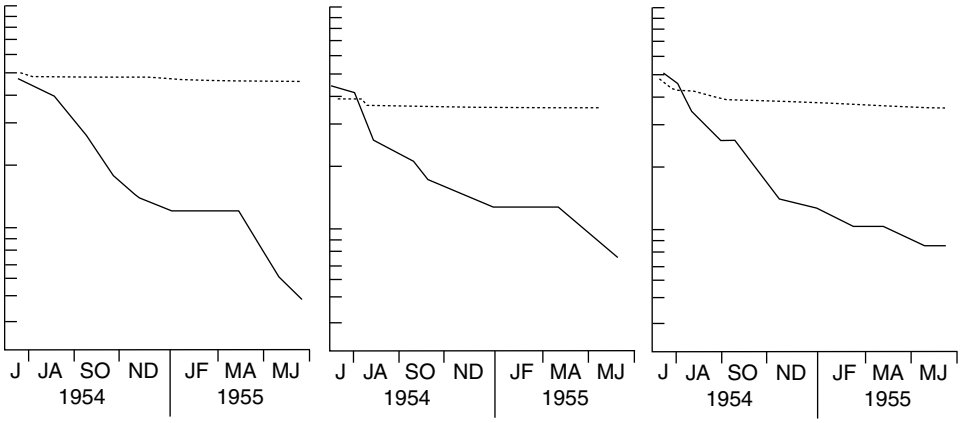


Figure 13 Survivorship curves for *Chihamalus stellatus*, with (solid line) and without (dashed line) *Balanus balanoides*.

Table 7 Numbers of ants and mice in the exclusion experiments of Brown and Davidson

	Rodents removed	*Ants removed*	*Control*	*% Increase*
Ant colonies	543	–	318	70.8
Rodent numbers	–	144	122	18.0

In the desert near Tucson, Arizona, experiments demonstrated that ants and rodents take most of the seeds, harvest the same size and species, and collect them from the same microhabitat. Brown and Davidson performed a set of exclusion experiments to demonstrate competition between these two taxa. They established replicated, circular plots (36 m in diameter) in relatively level, homogeneous desert scrub, approximately 60 km northwest of Tucson. Two plots were subjected to each of the following treatments. (1) Plots were fenced to exclude seed-eating rodents. Those rodents present were removed by trapping. (2) Granivorous ants were removed by repeated application of insecticide to individual colonies. (3) Both of the preceding treatments were done. (4) None of the treatments were done (control experiment).

Over a period of several months, the ants were censused five times (colonies counted) and the rodents were censused twelve times (trapping). **Table 7** shows the "numbers" of ants and rodents in the different treatments at the end of the experimental period. In the "control" experimental plots both ants and rodents showed a decrease in "numbers" compared to their single-taxon plots. Both groups showed a negative effect, which was interpreted as competition for seeds. This suggestion is strongly supported by the fact that on plots from which both ants and rodents were removed, there was a 5.5 times increase in density of seeds. However, there was no difference in seed density between plots with ants + rodents, ants only, and rodents only.

Anolis Lizards

Anolis lizards are an important component of island terrestrial communities in the eastern Caribbean. In these communities, anoles substitute for the ground-feeding insectivorous bird guild and are major components of the animal biomass. Joan Roughgarden and her colleagues have demonstrated the presence of competition between two species, *Anolis gingivinus* and *Anolis wattsi*. Both species are insectivorous, territorial as adults, and somewhat arboreal in habit.

Table 8 Lizard introductions to a Caribbean Cay

Grove		*Date*	A. gingivinus	A. wattsi
SK (559 m^2)	Original	August 1979	196	0
	Start	August 1979	196	103
	End	May 1980	205	14
MK (263 m^2)	Original	March 1980	116	0
	Start	March 1980	75	55
	End	October 1980	116	15
SK (559 m^2)	Original	March 1981	194	10
	Start	March 1981	92	110
	End	October 1981	153	37
LU (296 m^2)	Original	March 1981	104	0
	Start	March 1981	104	48
	End	October 1981	127	7

On the island of St. Maarten, distributional evidence suggests that there is present-day competition. *Anolis wattsi* occurs only in the central hills of St. Maarten, whereas *A. gingivinus* occurs throughout the island including the central hills. This observation is interesting because all the relatives of *A. wattsi* on nearby islands occur throughout all elevations and habitats, including sea-level habitats. Two other observations are suggestive of a competitive interaction between these two species. *Anolis gingivinus* shows a lower abundance and shifts its perch position where it co-occurs with *A. wattsi*, relative to when it occurs alone. However, these observational data are not conclusive evidence of competition.

Roughgarden and colleagues therefore carried out some manipulative field experiments on a small cay close to St. Maarten. This very small island was essentially a limestone platform (100 m × 400 m), 15 m above sea level, with several vegetation groves consisting of sea grape, Manchaneel, and perennial grasses. This cay had a resident population of *A. gingivinus*, but lacked any *A. wattsi*. **Table 8** shows the results

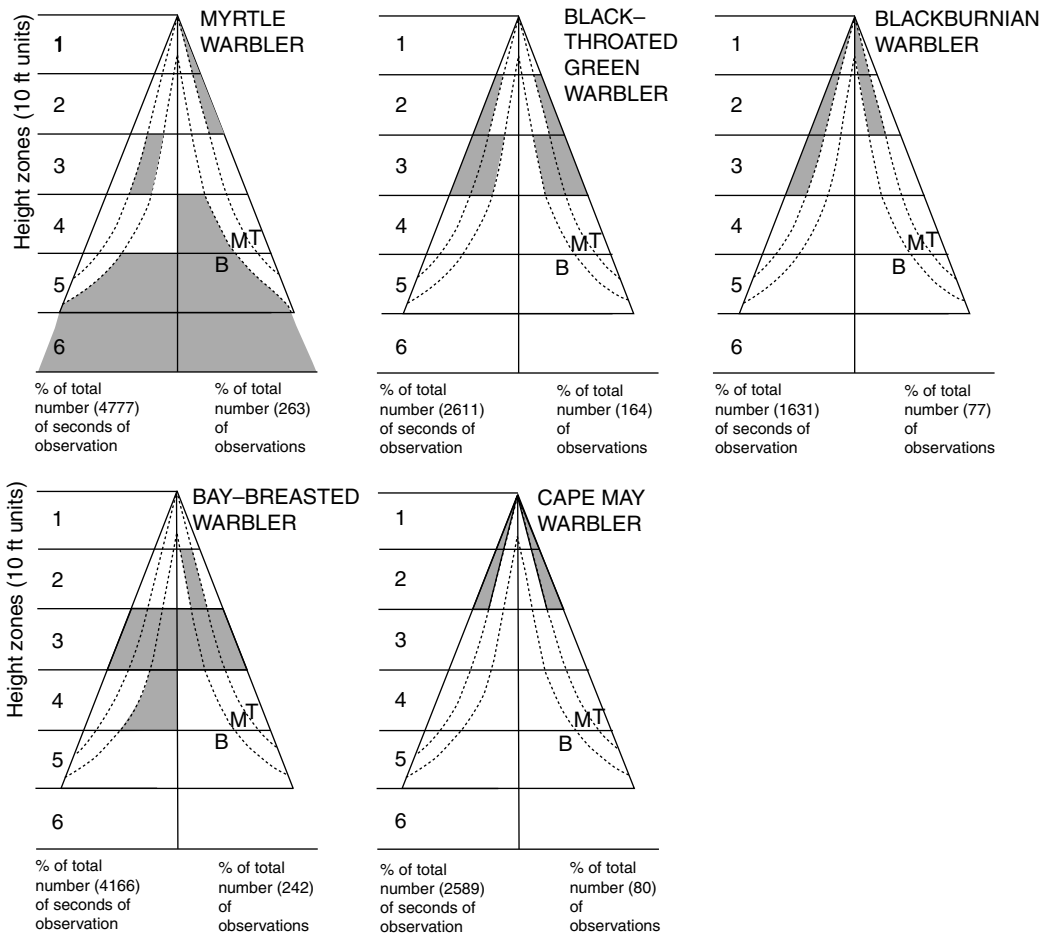


Figure 14 Foraging positions of five species of warbler in the coniferous forests of the northeastern United States. Shaded zones represent 50% of recorded time.

of four experiments carried out over a three-year period in three of the vegetation groves. In the experiments in grove SK (1979) and LU, marked individuals of *A. wattsi* were released and the resident *A. gingivinus* were left undisturbed. Survival of *A. wattsi* was low; most of the introduced individuals disappeared and those that were left did not establish territories in the center of the grove, only at the periphery. In the experiments in grove SK (1981) and MK, between 40% and 50% of the resident *A. gingivinus* were removed prior to introducing *A. wattsi*. In these experiments, survival was twice that in the other experiments and some surviving *A. wattsi* did establish territories in the center of the grove. It seems likely therefore that the restricted distribution of *A. wattsi* (to the central hills) observed on St. Maarten is due to competitive exclusion of *A. wattsi* by *A. gingivinus*.

Field Observations

New England Warblers

Many field studies of competition actually look for resource partitioning (niche separation) that would reduce interspecific competition, relative to intraspecific competition, and therefore promote coexistence. Furthermore, these field studies frequently do not examine directly the resources used (types of seed or types of insect) but instead use surrogate measurements (bill size or foraging area). An example of this type of study is Robert MacArthur's work on New England warblers.

Five species of warbler (myrtle, black-throated green, blackburnian, bay-breasted, and Cape May) inhabit the spruce forests of Maine and Vermont in the US. All of these species nest in the forests of New England and winter in the Caribbean. The spruce forests in which they live and breed are uniform, without obvious variety. Their beaks are all the same size and shape and their stomach contents are approximately the same. In his study, MacArthur recorded how long each species spent foraging in each part of a tree. A tree was divided into 10-ft zones vertically and branches were divided into three horizontal zones: (1) the bare or lichen-covered base, near the trunk, (2) the middle zone of needles, and (3) the terminal zone of new needles or buds. He produced diagrams of trees showing the zones of foraging activity for each species of warbler (**Figure 14**). The resulting foraging distributions suggested a rather subtle kind of resource partitioning (assuming that foraging in different parts of the tree gives access to different sections of the insect resource being used by these birds).

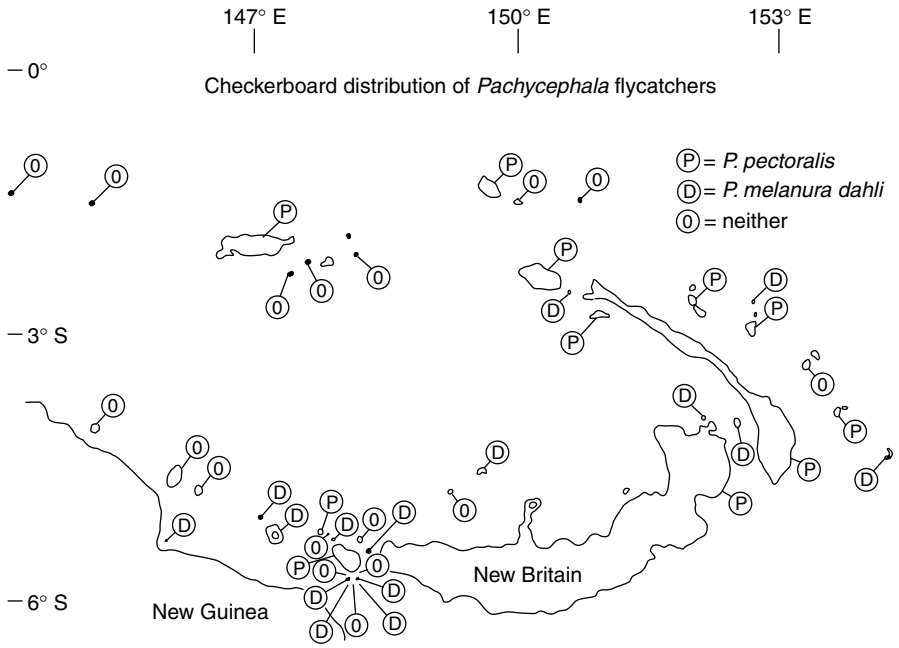


Figure 15 Checkerboard distribution of Pachycephala flycatcher species on the Bismarck Islands off Papua New Guinea. *P*=*P. pectoralis*, *D*=*P. melanura dahli*, and 0=neither.

In addition to these zone differences, MacArthur also suggested differences in foraging behavior between the five species: "subjectively, the black-throated green appeared nervous, the bay-breasted slow and deliberate." He quantified this by recording the length of time between "use of wings" and was able to show that the black-throated green had a shorter time interval than either the blackburnian or myrtle, which in turn had a shorter time interval than the Cape May and bay-breasted. He also recorded the direction in which each species predominantly moved while foraging. This could be vertical (up and down the tree), tangential (around the tree), or radial (from the center of the tree outward and vice versa). Again he found differences between the species.

MacArthur's work suggests that even when two species overlap in foraging space, they examine that space differently, and therefore use a different part of the available resource. The five warbler species are clearly searching for their insect resources in different spaces and in different ways, but whether this is sufficient to promote coexistence between potential competitors is difficult to tell. In fact, we have no direct evidence that these warblers are even competing for limited resources.

Pacific Island Birds

A number of field studies have used patterns in spatial distribution as evidence of the effects of competition on species co-occurrence. This type of field observation is illustrated by Jared Diamond's study (1975) of land-breeding birds on the Bismarck Islands off the coast of Papua New Guinea. Several pairs of ecologically similar species (good candidates for potentially strong competitors) showed what Diamond called

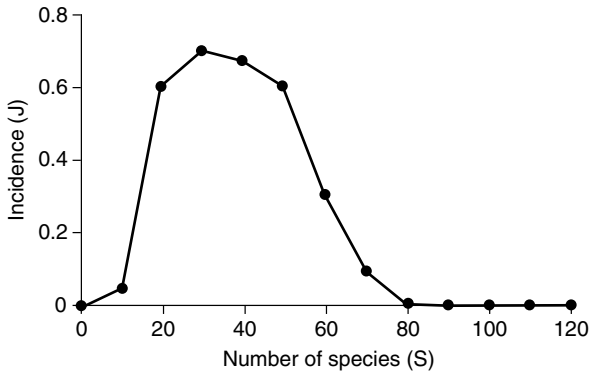


Figure 16 Incidence function for the cuckoo-dove *Macropygia mackinlayi* on the Bismarck Islands.

checkerboard distributions, for example, the two flycatchers *Pachycephala pectoralis* and *P. melanura dahli* (**Figure 15**). These two species have mutually exclusive but interdigitating distributions, with only one of the two flycatchers present on any one island (some islands have neither species). These checkerboard distributions are consistent with the idea that two ecologically similar species are excluding each other from islands because of strong competitive interactions.

Although Diamond did not find many examples of checkerboard distributions among his island birds, he did find other evidence of a more diffuse type of competition by plotting what he called incidence functions. These are graphs in which the probability of occurrence on an island (J=(# of islands with the species present)/(total # of islands)) is plotted against the number of species (S) on that island. Because of the species–area effect, S is also a measure of island size. **Figure 16** shows the incidence function for the cuckoo-dove

Table 9 Population estimates for the spotted hyena population of the Serengeti

Year	*Population estimate*
1967	2207 ± 120
1977	3306 ± 432
1986	5214 ± 828
1991[a]	9500

[a]Estimate includes hyenas of the Mara.

Macropygia mackinlayi. This is an example of a species that Diamond calls a "supertramp." These species are excellent colonizers but are poor at persisting in the diverse communities of large islands. This absence of supertramps from large islands implies strongly that competition is important in structuring these island bird communities.

African Wild Dogs and Hyenas

Many field observations of "competition" have relied heavily on changing species numbers over time. For example, one species arrives in an area (or a resident species increases its numbers), while another species declines in abundance. An early case was that of two species of crayfish in Russian streams, first commented upon by Gause in his book that reported his early *Paramecium* experiments (Gause, 1934). Another well-cited example is the invasion and spread of the American gray squirrel (*S. carolinensis*) in Great Britain, followed by the decline and contraction of the native red squirrel (*Sciurus vulgaris*). Yet another example is provided by the increase in abundance of the spotted hyena (*Crocuta crocuta*) in the Serengeti-Mara ecosystem of Kenya and Tanzania and the subsequent decline of the African wild dog (*Lycaon pictus*).

Spotted hyenas are the most abundant large predator in the Serengeti-Mara savanna ecosystem. Between 1969 and 1976, the hyena population increased by 50%, probably as a response to the herbivore increase following the elimination of the rinderpest virus. Some estimates for the subsequent population of hyenas in the Serengeti are shown in **Table 9**. Over this period the wild dog showed a decline in numbers, with all remaining dog packs finally disappearing from the Serengeti in the early 1990s (**Figure 17**). During the early 1970s, part of this decline in wild dog numbers was certainly due to hunting, as wild dogs were regarded as vicious vermin to be exterminated. However, the reason for the later decline is less certain, but competition with hyenas may have been important.

Both of these predators hunt in packs and run down their prey (mainly Thomson's gazelle, wildebeest, and zebra). Hyenas frequently follow dog packs when they go hunting from a den (during the breeding season) and steal their kill. Dogs can defend a kill, but about four dogs are required to keep off one hyena. Although food on the "hoof" is probably not a limiting resource for large carnivores in the Serengeti, food in the form of "kills" may be. Particularly during the breeding season, dog packs are hunting not only for themselves but also for the dominant bitch and her pups back at

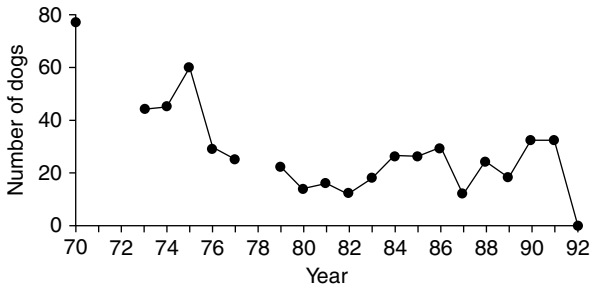


Figure 17 Number of wild dogs in the Serengeti ecosystem.

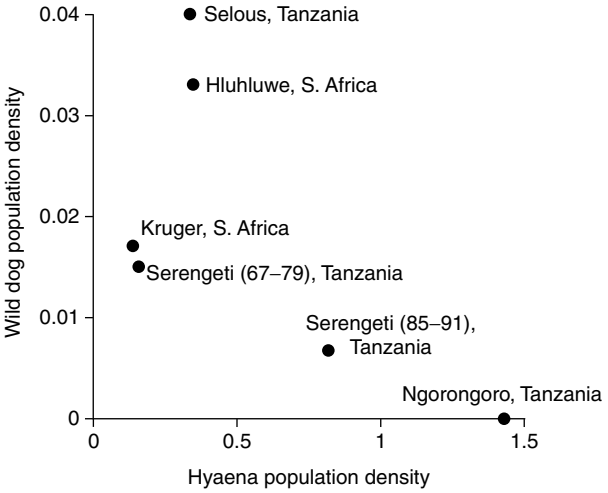


Figure 18 Correlation between the population densities of wild dogs and hyenas over a number of African ecosystems.

the den. Losing kills to competitors such as hyenas may have a serious effect on pup survival. Because the ability of wild dogs to defend kills depends on the ratio of dogs to hyena, this effect may have become gradually more serious over the past 30 years. Additional evidence supporting this story comes from the correlation ($r = -0.92$, $P = 0.01$) between hyena density and dog density over a series of ecosystems in eastern and southern Africa (**Figure 18**).

Conclusions

The Lotka–Volterra model can theoretically predict the outcome of interspecific competition between two species. Depending on initial population size (N_1 and N_2), carrying capacity (K_1 and K_2), and competition coefficient (α_{12} and α_{21}), either species can be the sole survivor, one species is always the sole survivor (competitive exclusion), or the two species will coexist. What is more, coexistence is assured if $\alpha_{12}/K_1 > 1/K_2$ and $\alpha_{21}/K_2 > 1/K_1$, which implies that intraspecific competition must be greater than interspecific competition for both species. These predicted outcomes are local outcomes (within a habitat patch such as a fungal fruiting body, a freshwater pool, a coniferous wood, or an island

depending on the scale of the environmental grain appropriate to the ecology of the organism). Both the aggregation model of competition and resource partitioning predict "local" competitive exclusion, with global coexistence using different local mechanisms that ultimately have the same effect on competitive interactions.

There is considerable evidence, from laboratory experiments, that these models could describe what happens to some real organisms. Graphs of N_1 against N_2 (using K and α) predict the correct competitive outcome for *Paramecium aurelia* and *P. caudatum*, while a small amount of environmental heterogeneity allows sufficient resource partitioning for *P. bursaria* to coexist with *P. aurelia*. The *Tribolium* experiments show that initial numbers can influence which species will win the competitive battle, and these same experiments, plus those with *Drosophila*, show that more complex relationships can still be incorporated into the same theoretical framework.

Lastly, there is some evidence, from field experiments and field observations, that these mechanisms might operate outside the computer and laboratory. We see patterns of replacement reminiscent of competitive exclusion (Bismarck birds and Serengeti carnivores) and niche differences consistent with resource partitioning (New England warblers). We also see numerical changes, after removal or introduction, that are consistent with competitive effects of one species on another (desert ants and mice, Scottish barnacles, and Caribbean lizards). However, the field data tend to fit the patterns predicted by theory rather than providing detailed parameter values for the models. Nonetheless, the ideas encapsulated within theoretical models, such as the Lotka–Volterra model, provide a believable framework within which to interpret complex species interactions in the real world.

See also: Carrying Capacity, Concept of. Coevolution. Predators, Ecological Role of. Resource Partitioning. Species Coexistence

References

Atkinson WD and Shorrocks B (1981) Competition on a divided and ephemeral resource. A simulation model. *Journal of Animal Ecology* 50: 461–471.

Diamond JM (1975) Assembly of species communities. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*, pp. 342–444. Cambridge, Massachusetts: Belknap.

Gause GF (1934) *The Struggle for Existence*. Baltimore: Williams and Wilkins.

Gilpin ME and Ayala FJ (1973) Global models of growth and competition. *Proceedings of the National Academy of Science* 70: 3590–3593.

MacArthur RH (1972) *Geographical Ecology*. New York: Harper & Row.

Shorrocks B and Rosewell J (1987) Spatial patchiness and community structure: Coexistence and guild size of drosophilids on ephemeral resources. In: Gee JHR and Giller PS (eds.) *Organisation of Communities: Past and Present*, pp. 29–51. Oxford, United Kingdom: Blackwell Scientific.

Shorrocks B, Rosewell J, Edwards K, and Atkinson WD (1984) Interspecific competition is not a major organising force in many insect communities. *Nature* 310: 310–312.

Complementarity

Paul Williams, Natural History Museum, London, UK
Lisa Manne, University of Toronto at Scarborough, Toronto, ON, Canada

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 1–12, © 2001, Elsevier Inc.

Glossary

Efficiency The degree of success in reaching a conservation goal relative to the cost (or to some surrogate for cost, such as number of areas). For goals such as maximizing the number of species represented for a particular cost, a distinction should be made between merely selecting areas on the basis of any records of species and choosing areas that maximize persistence of species in the long term.

Flexibility The degree to which alternatives exist for one or more selected areas in the context of reaching a particular conservation goal. When seeking a set of areas to represent maximum diversity, flexibility for a selected area may be absent, incomplete (replacing the selected area while still reaching the conservation goal would require substitution of two or more areas, or one or more areas of greater cost), or complete (other areas could be substituted, one-for-one by number or by cost, with the current choice).

Goals An explicit and precise statement of conservation aims. Goals should express the values of those people who provide the mandate for conservation. Making the goal explicit allows efficiency to be measured as an aid to accountability. A statement should include which attributes are valued (such as genetic diversity), which surrogates for this value are actually surveyed (such as higher taxa, species, and threatened species), which areas are to be considered (such as land management units or grid cells), and how constraints of viability, threat, and cost are to be measured and accommodated. Goals are not universal, but depend on people's values and their situations. Consequently, different goals may conflict, and areas necessary to meet one goal may be insufficient to meet broader goals.

Irreplaceability A property of areas that include species (or other valued attributes) restricted to so few areas that all such areas would be needed to meet a conservation goal.

Priority areas Areas that need the most urgent management, e.g., action to avert threat to meet a conservation goal (such as to increase the probability of persistence of valued species or other attributes). Ranking of areas by priority may differ from their ranking by value.

Representation The occurrence of species (or other attributes) within a set of selected areas. A distinction must be made between records of a species and areas with high probability of persistence for the species in the long term.

Values Values are interpreted here in the broadest sense, to include monetary and nonmonetary values. Biodiversity attracts many different values, which are not universal, and which depend on social and economic situations. Consequently different values may conflict.

Origins: The Biological Basis of Complementarity

Figure 1 shows a simple example of species complementarity for the mammal faunas of two areas. These faunas share one species, the lion, but each fauna also has species that are not shared. It is the species in one fauna that are not shared with a second fauna which are said to be "complementary" in relation to the second fauna.

The biological causes of complementarity are the processes that lead to differences among biotas. Fundamentally, the causes are descent with modifications, leading to differences among genes, species, and biotas. At the species level, the processes affecting this are familiar from differences in ecology and history. In ecology, differences among species in their responses to the environment, including responses to other species, lead to differences in species distributions near equilibrium. Unique events in history, particularly the gain or loss of barriers that prevent any approach to equilibrium with governing ecological factors,

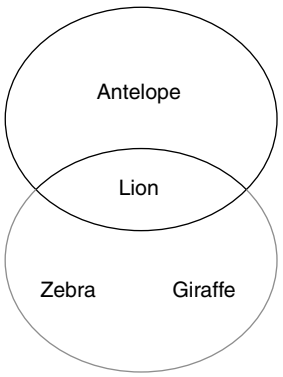


Figure 1 Complementarity. The two ellipses represent two areas, one with a fauna of two and the other with a fauna of three mammal species. They share one species, the lion. The zebra and giraffe in the lower area are complementary to the fauna of the upper area, whereas the antelope of the upper area is complementary to the fauna of the lower area.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00151-9

also contribute to species distributions, as studied in vicariance biogeography.

The description of patterns of complementarity at the species level is closely related to the description of spatial turnover, also known as beta or gamma diversity, depending on the spatial scale being studied. Similarly, patterns of complementarity in biodiversity may be expected to be highly dependent on spatial scale. This is because the size of area patches used to study patterns in the distribution of species has a strong effect on measures of their co-occupancy of these patches.

Importance for Cost-Effective Biodiversity Conservation

Need for Complementarity in Area Selection

Complementarity is an apparently simple idea that has considerable power to aid area-selection methods for conserving as much biodiversity as possible. While decisions on area selection will always benefit from experienced judgment, quantitative methods provide a framework of rules designed to make such complex decisions more rigorous and efficient, and more open to public accountability (Margules and Pressey, 2000). Complementarity value is a particular case of the idea in land-use planning of "marginal value," which is the incremental gain accrued when the area is managed in a particular way as part of a land management plan.

Accountability is important because conservationists should be acting in the broader interests of society, so people ought to be able to see that their values are being acted upon, and that limited resources are not being wasted. Accountability makes it possible to trace precisely why one area is chosen in preference to another, and if necessary, to challenge the data or rules.

Efficiency is important because the area of land (or volume of sea and air) available for conservation is limited, as there is often competition between conservation and other incompatible land uses. Efficiency ought to be considered in relation to the goal of ensuring effective long-term persistence of biodiversity.

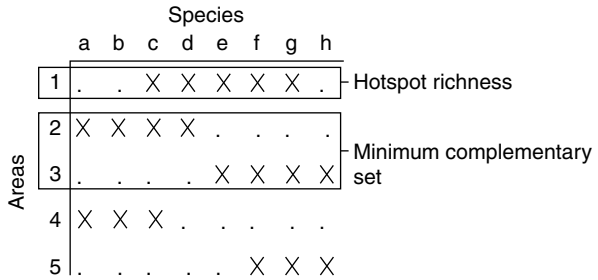


Figure 2 Hotspots and complementarity. Five areas have a biota of eight species. Area 1 is the hotspot of highest species richness (five species). However, this area is not part of the minimum fully representative set for all eight species, which is the combination of areas 2 and 3.

As an example of the efficiency and accountability that comes from using complementarity in area selection, suppose we could choose only two areas from **Figure 2** to represent as much biodiversity as possible, it might be tempting to begin by choosing area 1, the "hotspot," because it has the largest number of species (five). The next richest area would be area 2. Areas 1 and 2 would then represent a total of seven species (*a*–*g*). However, we can see from **Figure 2** that areas 2 and 3 have completely different (highly complementary) biotas, so if chosen together, these two areas alone could succeed in representing all eight species (*a*–*h*).

History of Complementarity in Area Selection

Complementarity-based methods were first described by Kirkpatrick (1983). The term "complementarity" was coined specifically in the context of area selection by Vane-Wright *et al.* (1991). The importance of the term is that it identifies a property essential to a group of techniques that have proven to be more efficient (Pressey and Nicholls, 1989).

The impetus for the development of complementarity methods in conservation has been strengthened in the last 10 years by Agenda 21 and by the Biodiversity Convention. These policy initiatives gave a new emphasis to conserving the value of broader biodiversity, in addition to the long-recognized need to conserve the relatively few, particularly threatened or vulnerable species. Conserving broader biodiversity presented several new challenges, particularly in keeping track of the very large numbers of components of which biodiversity is comprised.

Complementarity-based methods have been used more in the Southern Hemisphere (particularly in Australia and South Africa), where knowledge of the biota comes more often from new surveys, some designed specifically for identifying potential conservation areas. In contrast, complementarity methods have been less popular in countries with a long natural history tradition and with a strong local knowledge base, where many locally important areas may already be well known (although rigorous prioritizing among them or adding to them efficiently may still be difficult without using complementarity). Nonetheless, there is a growing interest in complementarity methods by the US Gap Analysis Program, The Nature Conservancy, Conservation International, and by biological recording schemes in Europe.

Application to Simple Species Representation

Setting Goals

A common goal of biodiversity conservation projects is to ensure the persistence of as much biodiversity as possible for the future. The basic form of the problem is one of optimizing the number of valued components of biodiversity (usually species) represented within a set of areas (whether these are to be protected or not), subject to constraints including species viability, threat, and cost.

A simple representation goal used to illustrate the basic principles in many academic studies has been to achieve at least one representation of every included species within the

reserve network. Alternatively, the goal could be to achieve any required number of representations (this number could also differ among species), or could be expressed in terms of population size, or probability of occurrence. For many studies to date, information on constraints has been less-widely quantified (although perhaps because less effort has been invested). Viability and threat are often either ignored or are treated only by using crude rules.

In the simplest formulation of conservation goals, the cost constraint on area selection can be represented by the number of areas, treating all areas as being of equal cost. A few studies have used the extent of areas of differing size as a measure of cost. Even when financial cost data have been used, they have rarely included a thorough consideration of all the appropriate cost components. These may include cost of area acquisition, costs of management to reduce threats, the number of people impacted if the area should be set aside for conservation, and (crucially) the opportunity costs that arise from income from any alternative uses of an area (e.g., forestry and agriculture) that is foregone when it is managed for biodiversity conservation (Faith and Walker, 1996c).

An important part of specifying an area-selection goal must be to include the contribution of an existing conservation network. If some species can be considered to be adequately protected by the existing network, then only the complementary species need be targeted for additional representation. This is the idea behind "gap analysis," which is essentially about finding the complement to the existing coverage.

One kind of goal (the minimum-cost problem) is of the form: "which is the minimum set of areas required to represent the diversity of species?" to achieve at least one representation (or any other required number) for species within the area set. Another kind of goal (the maximum-coverage problem) takes the form: "how can we choose 1% of the total area to represent the greatest diversity of species?"

Complementarity-Based Techniques

Achieving the representation goals described in the previous section is not entirely straightforward. For example, the size of the problem when seeking a minimum-cost set of areas cannot be calculated simply from the numbers of species and areas (this kind of problem is described as not-polynomial complete or "n-p complete").

Fortunately, the problem is familiar for many kinds of similar applications in Operations Research, where a range of techniques for dealing with it have gained acceptance. First, exhaustive search can test every possible combination of areas to find the optimum, although this is usually prohibitively slow. Second, branch-and-bound algorithms can restrict the search to only a subset of all possible combinations of areas, but a subset within which the optimal solution is expected to lie. They have become a more practical option, but may be slow, may require specialist software, and may not be able to solve all problems (Pressey *et al.*, 1996). Third, heuristic algorithms can use relatively simple sets of rules that have been found from the experience with previous datasets, to give good approximations to optimal solutions. In practice, heuristic suboptimality may be trivial compared with concerns about data quality and practical implementation for conservation (Pressey *et al.*, 1996). Heuristics are also easy for non-specialists to implement, with or without computers.

Flexibility, Priority, and Irreplaceability

Quantitative area-selection techniques frequently identify several alternative combinations of areas that meet their representation goals. This allows for flexibility in the final area choices (Pressey *et al.*, 1993). In general, more constraints placed upon the optimization problem mean less flexibility in the solution. When using exhaustive or branch-and-bound area-selection techniques, flexibility may be identified directly as alternative optimal sets. When using heuristic techniques, however, it is done indirectly, by identifying which species each selected area contributes uniquely to the representative set. Any other area that shares these same "goal-essential" species (Rebelo, 1994) could be substituted in its place, and is therefore a fully flexible alternative.

Priority is regarded here as the urgency with which areas require management action for conservation to avert threats to the persistence of their biodiversity value. Therefore, priority ranking of areas may differ from their ranking by biodiversity value or by their flexibility. The distinction between value and priority is important but often ignored.

Areas within a representative set for which there are no flexible alternatives (to meet a particular goal) have been described as "irreplaceable," an all or nothing property. However, "shades of irreplaceability" has also been used to refer to the relative number of alternatives that may exist for an area choice, depending on how widespread the species occurring there may be (Pressey *et al.*, 1994). Shades of irreplaceability concerns the frequency with which an area is required among all possible fully representative area sets, whatever the size or cost of the set. This frequency usually depends on the rarest species in each area, because a rare species has few alternative areas available to represent it. Using the definitions at the beginning of this article, shades of irreplaceability is a measure of flexibility rather than a measure of the conservation value of an area (which might be measured in terms of its complementary biodiversity value).

Assessing Efficiency

Much of the emphasis on using complementarity in area selection is on how it can be used to represent more biodiversity with limited resources. The relationship between representation and cost has been identified with concepts of efficiency. However, efficiency ought to be considered in relation to not just representation of species records, but in the context of all the other constraints (viability, threat, etc.) necessary to reach the goal of ensuring the effective long-term persistence of biodiversity.

For minimum-set goals, the efficiency of species representation was measured by Pressey and Nicholls (1989) as the proportion of all areas that is "not" required to meet the goal. This is an absolute measure, in the sense that efficiency is judged relative to selecting all areas or none.

Because the absolute efficiency that can be achieved in any particular case varies with patterns in the data, it may be more useful to measure efficiency relative to what might actually be expected with those particular data. In the past, efficiency of species representation has often been assessed informally by comparing species-accumulation curves for ordered sequences of selected areas. The upper bound to this is the maximum representation within the chosen constraints, the maximum-coverage set ($s_{\max,c}$) for the given constraints (c). Some studies have also used simulated random draws of areas (without replacement), to estimate whether the representation achieved by an area set is significantly greater than would be expected by chance, when selecting the same number of areas (**Figure 3a, c**). One useful lower bound for comparison is the representation that might be expected if areas were selected at random ($s_{\text{rand},c}$), without regard to their biota. Therefore, for the representation within an observed set of areas ($s_{\text{obs},c}$) for the same constraints, the relative efficiency of any set of areas ($E_{\text{obs},c}$) for constraints c is given by

$$E_{\text{obs},c} = (s_{\text{obs},c} - s_{\text{rand},c})/(s_{\max,c} - s_{\text{rand},c})$$

In the simplest form of this measure, the number of areas can be used as a measure of the cost constraint, in which case efficiency is measured relative to what might be expected from the data when selecting a particular number of areas (**Figure 3b, d**).

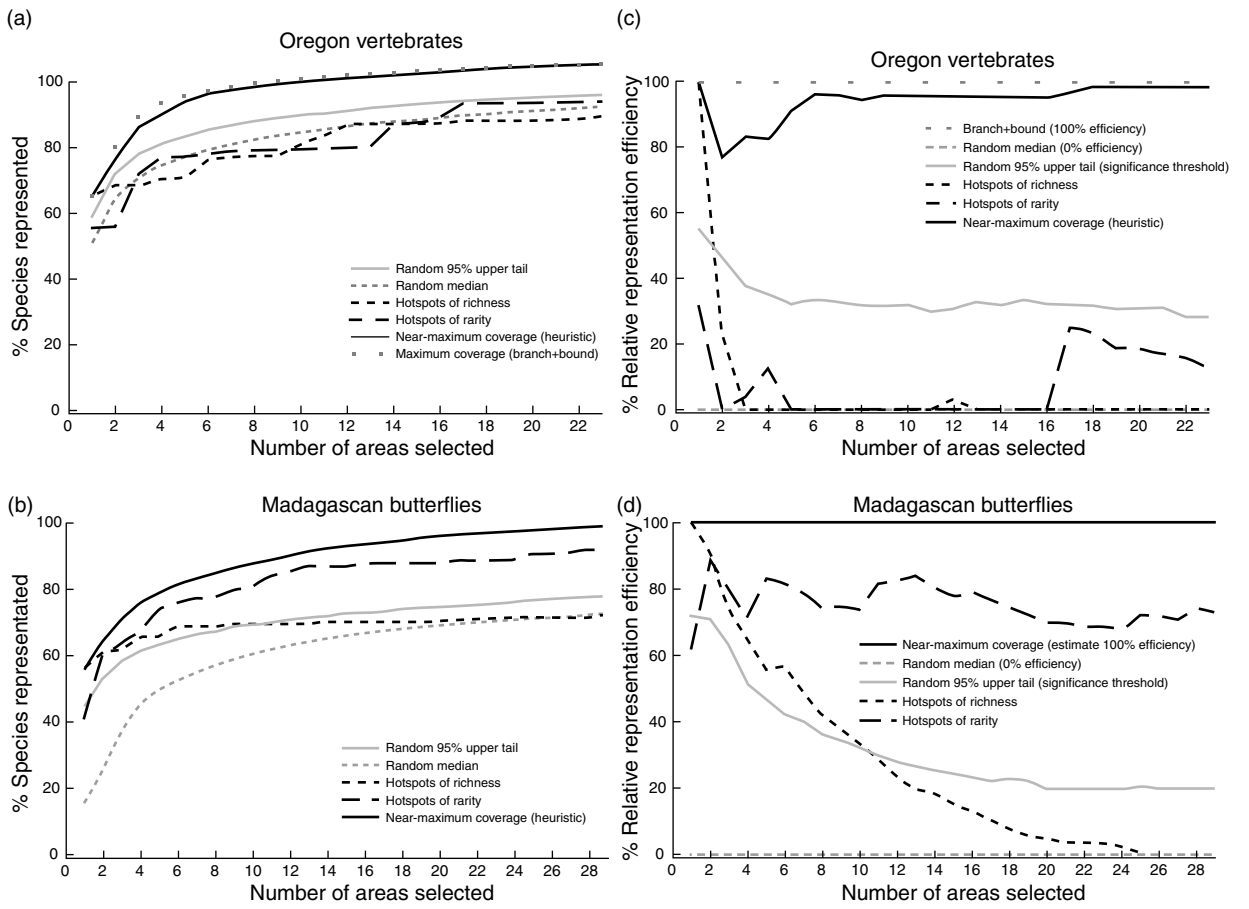


Figure 3 Species representation and relative efficiency. (a) Cumulative percentage of a group of 426 species of Oregon vertebrates represented within increasing numbers of hexagonal grid cells (each 635 km^2, as developed for the Oregon Gap Analysis Program) selected using maximum-coverage (branch-and-bound algorithm), near-maximum coverage (heuristic algorithm), hotspots of richness, and hotspots of rarity (richness in the rare-quartile of species), and areas chosen at random. Scores above the solid gray line (the percentage threshold score to the top 5% within 1000 randomly drawn scores) are significantly better than expected when choosing areas at random. Data reanalyzed from Csuti *et al.* (1997). (b) Cumulative percentage of a group of 321 species of Madagascan butterflies represented within increasing numbers of quarter-degree grid cells (each ~170 km^2) selected using near-maximum coverage (heuristic algorithm), hotspots of richness, and hotspots of rarity (richness in the rare-quartile of species), and among areas chosen at random. (c) Results from (a), rescaled for relative efficiency between the maximum that can be achieved (using the branch-and-bound algorithm), and the median representation expected from selecting areas at random (relative representation efficiency: $E_{\text{obs},n} = 100(s_{\text{obs},n} - s_{\text{rand},n})/(s_{\max,n} - s_{\text{rand},n})$). Scores above the solid gray line are significantly better than expected when choosing areas at random. Data courtesy of David Lees. (d) Results from (b), rescaled for relative efficiency between an approximate estimate of the maximum coverage that can be achieved (using a heuristic algorithm), and the median representation expected from selecting areas at random. Because the heuristic algorithm is likely to be further from the optimum when very small numbers of areas are selected, the corresponding estimates of relative efficiency may be less reliable.

The maximum absolute efficiency in species representation that can be achieved using complementarity is determined by patterns in the data of co-occupancy of areas by species (i.e., "nestedness"). For example, if all species were found in one area, then only one area would be needed as a minimum to represent them; whereas if none of the species were found together, then as many areas as species would be needed (Vane-Wright *et al.*, 1991).

Application to More Complex Problems

Complementarity and Probability of Persistence or Occurrence

Ultimately, effective biodiversity conservation is about ensuring the long-term persistence of valued biodiversity. One way to improve area selection is to estimate the probability of persistence of species in different areas.

Probabilities of occurrence can be estimated by fitting explanatory models to distribution data. A popular approach is to use niche-based models, which assume a unimodal response by species to variation in environmental variables, such as climate. Probabilities of occurrence are sometimes also used alone to overcome the problem of incomplete species occurrence data, without attempting to adjust for viability and threat.

The simplest approach to treating viability in area selection is to include only those populations that are considered to be viable. Unfortunately, the data required to decide this are rarely available. To deal with threat, areas where species are known to be under severe threat might be excluded from consideration because they are deemed to be irretrievable. Alternatively, the cost of averting threats might be included in the area cost.

At least in principle, probabilities of occurrence (at the time when data were obtained) could be converted to estimated probabilities of persistence (the probability of continued occurrence at some specified time in the future) if threats and species' responses to those threats could be predicted. There are usually substantial uncertainties associated with this process. Alternatively, if it can be assumed that conservation management can ameliorate all threats in any selected areas, then probabilities of persistence can be expected to be related positively to probabilities of occurrence. Either way, techniques are available to seek complementary areas to reach a conservation goal that is expressed in terms of probabilities with efficiency (Williams and Araujo, 2000).

Complementarity and Changing Distributions

The effect of any anthropogenic environmental change on distributions can be estimated using niche-based models of

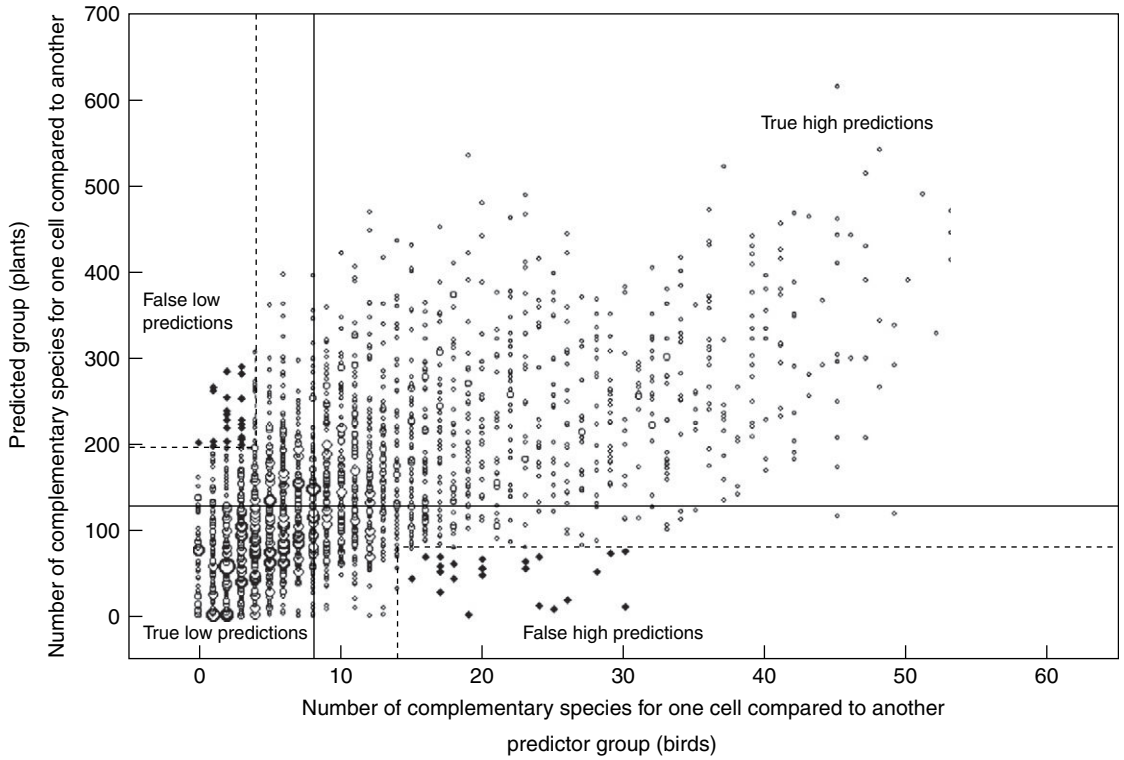


Figure 4 Assessing indicators of complementarity. Scatter frequency plot of complementarity scores for a sample of 2776 pair-wise comparisons drawn at random from among 2776 British 10 × 10 km grid cells with presence data for both selected birds (105 species) as indicators of vascular plants (1353 species). Higher frequencies of scores are shown by larger circles. Medians on each axis are shown by solid lines. Quartiles are shown by dashed lines, defining opposing quartile boxes, with points within these boxes shown as diamonds. Reproduced from Williams PH, Faith D, Manne L, Sechrest W, and Preston C (2006) Complementarity analysis: Mapping the performance of surrogates for biodiversity. *Biol. Conserv.* 128: 253–264.

habitat suitability, for example by fitting habitat models across a series of time slices using the expected changes in predictive environmental variables. These equilibrium models of habitat suitability need to be combined with species dispersal models to constrain species expected distributions to those areas that species are likely to be able to reach by dispersal. Some species may have to disperse if they are to persist at all. Consequently, conservationists are interested in the potential to identify corridors or stepping stones to aid dispersal of these species. Efficiency has required techniques that seek complementarity in the selection of chains of areas rather than among single static sites (Williams *et al.*, 2005). Further developments in this area can be anticipated.

Finding Indicators of Complementarity within Broader Biodiversity

Data are rarely available for all of the valued parts of biodiversity. One attempt to deal with this problem is to use data for one group of (related or unrelated) species as a surrogate or indicator for another group (Ryti, 1992), or ideally for all of biodiversity. The success of this approach relies on the congruence in distributions between the indicator and indicated groups, not just in patterns of richness, but also in patterns of complementarity at different spatial scales. Reliable indicators of complementarity have proven difficult to find, with some indicators resulting in no more species being represented than when areas were chosen at random. Unfortunately, many attempts to assess the value of indicators as predictors of complementarity within biodiversity have substantial flaws or limitations.

First, some authors have claimed that complementarity is a kind of dissimilarity, and have used symmetrical dissimilarity metrics, particularly Bray-Curtis, as a measure of complementarity. This approach is seriously flawed, because complementarity between pairs of biotas is usually asymmetrical, and may be strongly so. For example, a small coastal island may have only a small subset of the biota of the neighboring mainland, without additional species. In this case, the mainland has high complementarity with respect to the island, but the island has no complementarity with respect to the mainland. Another common weakness of this first approach is that using the mean (or median) value for a complementarity measure to assess prediction between groups could be misleading. For example, **Figure 4** shows a scatter plot of complementarity scores (numbers of complementary species of birds, x, and plants, y) for pairs of areas. If **Figure 4** had similar ranges of values on both axes, and if the complementarity values were concentrated both near the origin and in the upper left and lower right corners, but not in the middle, then

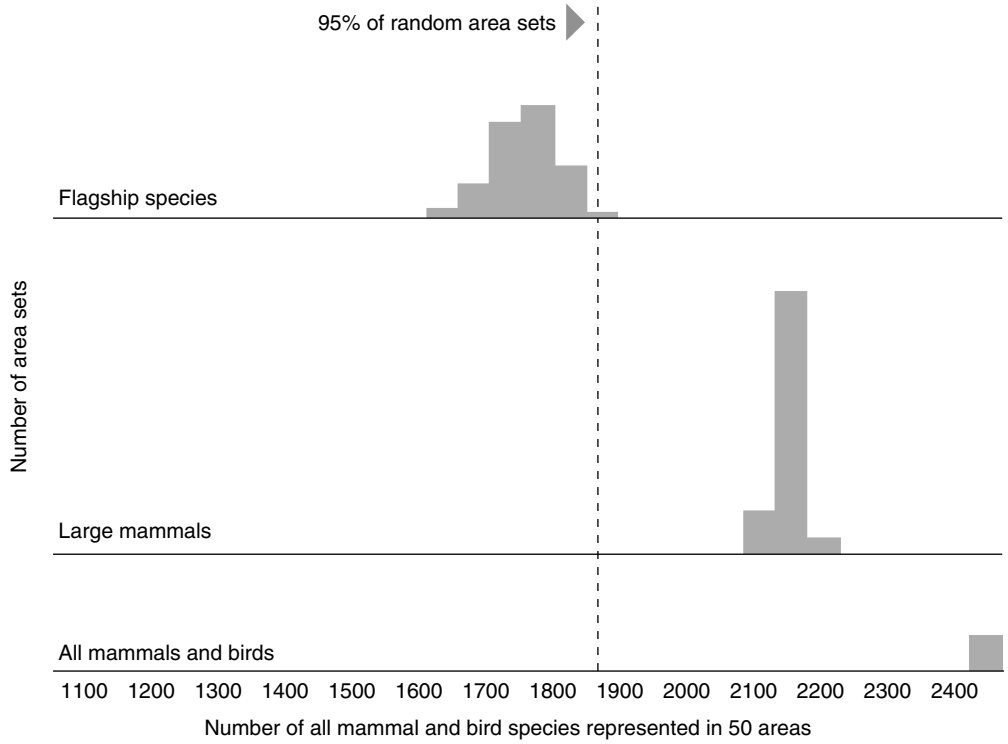


Figure 5 Flexibility and representation of other species. Histograms of representation of all mammals and breeding birds within flexible sets of 50 one-degree grid cells (each $\sim$11,025 km^2) in sub-Saharan Africa, each selected for maximum coverage of three different groups of species: (upper) six flagship mammals; (middle) five orders of larger mammals; and (lower) all mammals and breeding birds. With the exception of the histogram at the bottom, which shows all 120 flexible sets of 50 cells, scores are from a randomly drawn sample of 1000 alternative sets from totals of 2.12×10^{86} flexible sets of 50 cells for flagship species and 6.54×10^{41} flexible sets of 50 cells for larger mammals. Representation scores to the left of the dotted line are within the range expected when choosing 50 grid cells at random. Redrawn from Williams PH, Burgess ND, and Rahbek C (2000) Flagship species, ecological complementarity, and conserving the diversity of mammals and birds in sub-Saharan Africa. *Anim. Conserv.* 3: 249–260.

mean (or median) complementarity for the two groups could be very similar, even though true prediction of complementarity would be very poor (with most predictions being falsely high or falsely low).

A second approach has been to apply complementarity-based area selection to both an indicator and the group to be predicted, and then examine the spatial overlap between the two area sets. This approach is flawed because only one of a large number of alternative flexible area sets is considered in each case, which may be unrepresentative of the entire population of sets, and it says nothing about whether or not areas within the subset of overlapping areas include most of the complementarity, or make only minor contributions. It is possible for the complementary contribution of an overlap area to differ strongly between the predictor and predicted groups of organisms.

The third and more common approach recently has been to apply complementarity-based area selection using an indicator and then sum the representation of another group within the area set. In **Figure 5**, all the 120 flexible sets of 50 areas selected for maximum coverage of all mammals and birds must by definition have the same number of mammals and birds. However, when sets of 50 areas are selected for maximum coverage of just the flagship species, or just the large mammals, then any flexible sets may vary in the numbers of all mammals and birds represented. This approach, although a good start, has limited value because whether the results include only one area set, a large sample of flexible sets (**Figure 5**), or even all area sets, they do not include the pattern of complementarity among all areas.

To assess the indicators for broader biodiversity more broadly, we need methods for assessing whether they are good predictors of complementarity across all areas (Williams *et al.*, 2006). The fundamental measure (**Figure 4**) is the complementary richness of a "focal biota" for one area with respect to a "reference biota" of other area(s) (which might be under some form of conservation management, referred to here for convenience as a "reserve"). This complementarity measure is then plotted for the indicator or predictor on the *x*-axis, against the corresponding complementarity for the same area comparison but for overall biodiversity on the *y*-axis. The reference biota in this comparison can be contributed by any number of areas: from 0 (whereupon the calculation simplifies to one of species richness) to a single area, to a set of areas,

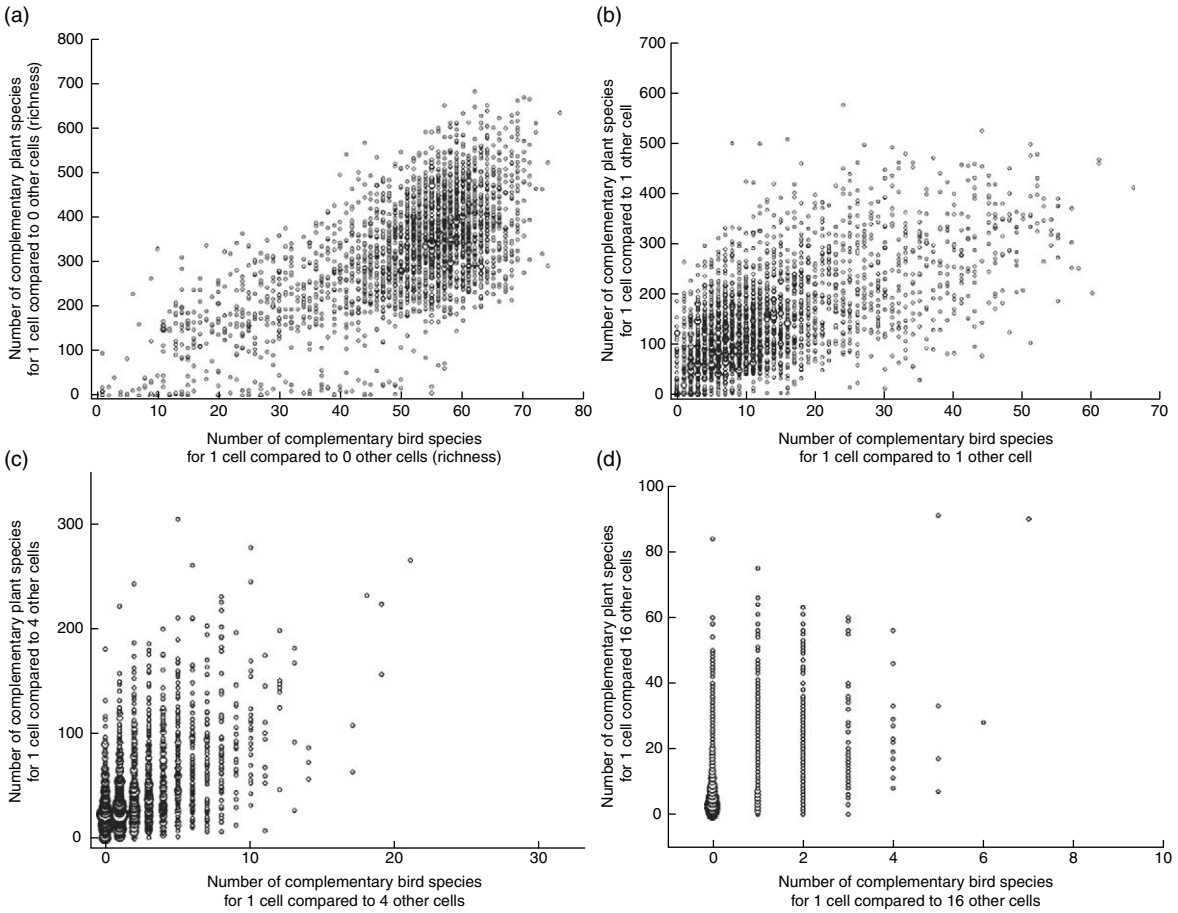


Figure 6 Declining predictivity of complementarity as reserves are selected. Frequency plot of complementarity scores for a sample of 2776 comparisons drawn at random from among 2776 British 10 × 10 km grid cells with presence data for both birds (105 species) as indicators of vascular plants (1353 species). In each case, one focal cell is compared with a group of reference cells drawn at random: (a) 0 reference cells; (b) 1 reference cell; (c) 4 reference cells; and (d) 16 reference cells. Reproduced from Williams PH, Faith D, Manne L, Sechrest W, and Preston C (2006) Complementarity analysis: Mapping the performance of surrogates for biodiversity. *Biol. Conserv.* 128: 253–264.

and up to the total of all other areas with records (whereupon the calculation simplifies to one of "endemism" at the scale of the analysis). Because when dealing with large numbers of grid cells the number of comparisons can become very large, the focal area and a single reference area can both be selected at random to make a sample of comparisons (Figure 4). To address the problem of interdependence among pair-wise data points, segments of the frequency distribution can be compared. The number of points falling within the opposing upper and lower quartiles on the plot (dotted lines in Figure 4), which correspond to false high and false low predictions are counted. To test for significant predictivity of complementarity by possible indicators, the numbers of these departures from a predictive relationship must be significantly lower than would be expected from a null model, which is used to reassign local biotas at random.

Indicator relationships change as the number of reference "reserve" cells increases. Initially, when there are no reference cells, the scores for each focal cell are simply the species richness scores for those cells. When there is one reference cell, then the scores for the focal cell are complementary with respect to that single cell. As the number of reference cells (which could be the number of cells in an existing reserve network) goes up, the score for the focal cell becomes increasingly dependent on some of the most narrowly distributed species in the data. There is a corresponding increase in the scatter of the predictive relationship (Figure 6), so that the strength of the predictive relationship decreases with an increasing percentage of false predictions. However, the predictivity of complementarity can remain significant.

The spatial pattern in false predictions can also be mapped. Areas on these maps associated with many false high predictions are overvalued and would be an inefficient investment for scarce conservation resources. Areas associated with many false low predictions are undervalued and unlikely to attract conservation action, so we need to know whether they are particularly likely to be highly threatened. Any geographical patterns could be used to identify habitat-associated biases in the performance of indicator groups.

Patterns of complementarity might be expected to differ among groups of organisms because they differ in the ecological and historical processes that have shaped their distributions. For example, some groups share closer ecological similarities in diet, habitat, and climatic tolerances. If some of them have diversified at the same time and in the same areas, they may also have shared similar phylogenetic and biogeographic patterns, as envisaged by vicariance biogeography. When seeking groups of species to use as indicators for use in area selection to represent other groups or broader biodiversity, what will be needed is shared patterns of complementarity within ecological or phylogenetic patterns.

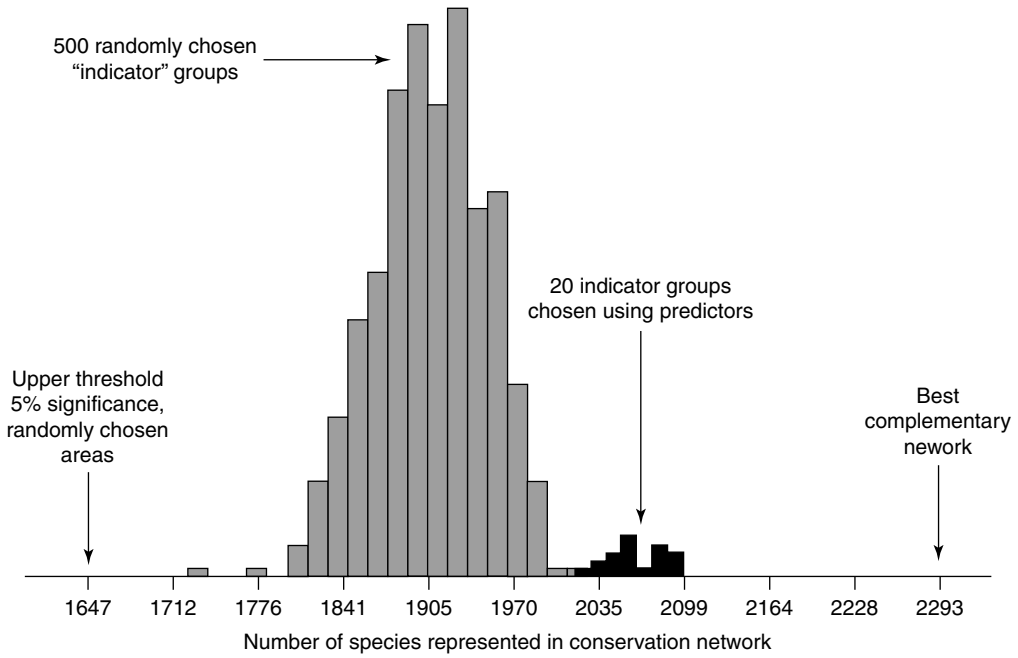


Figure 7 Improving indicators of complementarity by selecting species with particular range properties. The gray bars are results from trial "indicator" groups of 155 species constructed at random and used to choose a set of 24 areas: the height of the bars shows the relative frequency and the *x*-axis shows the number of nonindicator species represented (from among 2937 remaining species) in each area set selected using the randomly selected "indicator" groups. At the far left of the *x*-axis is the upper 5% threshold of significance for species representation from choosing 24 areas at random. At the far right is the best 24-area set possible from using all 3092 species to make the area selection. The black bars show the improvement in species representation above the gray bars when groups of 155 indicator species are chosen using (i) species with low variance in geographic ranges among ecoregions; (ii) species with low mean range sizes; (iii) species showing low mean species richness (among grid cells containing indicator species); and (iv) a high proportion of plants or restricted plants (which tended among the species examined to have smaller range sizes). Reprinted from Manne LL and Williams PH (2003) Building indicator groups based on species characteristics can improve conservation planning. *Anim. Conserv.* 6: 291–297.

Can groups of species be chosen to improve their performance as indicators of complementarity in biodiversity? Manne and Williams (2003) showed that even indicator groups comprising of randomly chosen, taxonomically unrelated species performed better than randomly chosen areas. Furthermore, they demonstrated that area-selection performance could be improved when using groups of indicator species that had been chosen on the basis of the characteristics they found to be associated with good indicators (**Figure 7**).

Pattern Complementarity

Complementarity, as well as being applicable to discrete entities such as species, can also be applied to patterns that might be used to predict their distributions (Faith and Walker, 1996b). This extension of the indicator idea has been applied in two areas.

First, complementarity has been used in seeking sets of genotypes (usually species) to represent genetic diversity efficiently (Vane-Wright *et al.*, 1991). In this case, phylogenetic trees are seen as patterns that can be used to predict genetic differences between terminal genotypes (species), because the genes are expected to follow a hierarchical model of descent with modification along the branches of the tree (Williams *et al.*, 1995). This led to the development of a subtree-length measure of phylogenetic diversity (or PD) (Faith, 1994), which is widely accepted.

Second, complementarity has been used in seeking sets of areas to represent ecological diversity (ED) efficiently. The pattern used is the evenness with which species or areas are distributed within a multivariate ordination space of environmental variables. The unimodal response model for how each species' ecological niche occupies a part of this environmental space means that this pattern of species' distributions within the space can be used to predict their occupancy of candidate areas for selection. Consequently, particularly in cases of incomplete species distribution information, this approach may represent ED with greater efficiency (Faith and Walker, 1996a). This approach holds the promise of being able to predict complementarity within broader biodiversity, but in practice it is weakened by the strong departures of many real species distributions from the assumption of equilibrium with their ecological governing factors. Many aspects of the approach remain to be explored.

Conclusions

Using complementarity in area selection for biodiversity conservation is not only accountable and efficient, but also can be much more flexible for dealing with diverse and heterogeneous goals and data than simple exercises may imply. Complementarity can be applied in combination with any available and appropriate information on the distribution of diversity: on its viability, on threats, on costs, on its distribution within multidimensional environmental space, and on any other ecological or social constraints that are currently being used to make decisions. Consequently, it should be widely applicable with even very complex data to improve upon *ad hoc* decisions. The major challenges for the immediate future are how best to predict probability of persistence and threat from limited information and how best to integrate this information into complementarity-based area-selection methods.

See also: Biodiversity, Definition of. Biodiversity, Origin of. Conservation Biology, Discipline of. Environmental Ethics. Hotspots. Indicator Species

References

Csuti B, Polasky S, Williams PH, *et al.* (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biol. Conserv.* 80: 83–97.

Faith DP (1994) Phylogenetic pattern and the quantification of organismal biodiversity. *Phil. Trans. Roy. Soc. Lond. B* 345: 45–58.

Faith DP and Walker PA (1996a) Environmental diversity: On the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiv. Conserv.* 5: 399–415.

Faith DP and Walker PA (1996b) How do indicator groups provide information about the relative biodiversity of different sets of areas? On hotspots, complementarity and pattern-based approaches. *Biodiv. Lett.* 3: 18–25.

Faith DP and Walker PA (1996c) Integrating conservation and development: Effective trade-offs between biodiversity and cost in the selection of protected areas. *Biodiv. Conserv.* 5: 431–446.

Kirkpatrick JB (1983) An iterative method for establishing priorities for the selection of nature reserves: An example from Tasmania. *Biol. Conserv.* 25: 127–134.

Manne LL and Williams PH (2003) Building indicator groups based on species characteristics can improve conservation planning. *Anim. Conserv.* 6: 291–297.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature (Lond.)* 405: 243–253.

Pressey RL and Nicholls AO (1989) Efficiency in conservation evaluation: Scoring versus iterative approaches. *Biol. Conserv.* 50: 199–218.

Pressey RL, Humphries CJ, Margules CR, Vane-Wright RI, and Williams PH (1993) Beyond opportunism: Key principles for systematic reserve selection. *Trends Ecol. Evol.* 8: 124–128.

Pressey RL, Johnson IR, and Wilson PD (1994) Shades of irreplaceability: Towards a measure of the contribution of sites to a reservation goal. *Biodiv. Conserv.* 3: 242–262.

Pressey RL, Possingham HP, and Margules CR (1996) Optimality in reserve selection algorithms: When does it matter and how much? *Biol. Conserv.* 76: 259–267.

Rebelo AG (1994) Using the Proteaceae to design a nature reserve network and determine conservation priorities for the Cape Floristic Region. In: Forey PL, Humphries CJ, and Vane-Wright RI (eds.) *Systematics and Conservation Evaluation*, pp. 375–396. Oxford: Oxford University Press.

Ryti RT (1992) Effect of the focal taxon on the selection of nature reserves. *Ecol. Appl.* 2: 404–410.

Vane-Wright RI, Humphries CJ, and Williams PH (1991) What to protect? Systematics and the agony of choice. *Biol. Conserv.* 55: 235–254.

Williams P, Hannah L, Andelman S, *et al.* (2005) Planning for climate change: Identifying minimum-dispersal corridors for the Cape Proteaceae. *Conserv. Biol.* 19: 1063–1074.

Williams PH and Araujo MB (2000) Using probabilities of persistence to identify important areas for biodiversity conservation. *Proc. Roy. Soc. B* 267: 1959–1966.

Williams PH, Burgess ND, and Rahbek C (2000) Flagship species, ecological complementarity, and conserving the diversity of mammals and birds in sub-Saharan Africa. *Anim. Conserv.* 3: 249–260.

Williams PH, Faith D, Manne L, Sechrest W, and Preston C (2006) Complementarity analysis: Mapping the performance of surrogates for biodiversity. *Biol. Conserv.* 128: 253–264.

Williams PH, Gaston KJ, and Humphries CJ (1995) Do conservationists and molecular biologists value differences between organisms in the same way? *Biodiv. Lett.* 2(1994): 67–78.

Complexity Versus Diversity

Shahid Naeem, Columbia University in the City of New York, New York, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biotic interaction When the growth or behavior of one species affects those of another species; such interactions can be antagonistic (e.g., competition for limiting resources, predation, herbivory, or parasitism) or facilitory (e.g., pollination or other forms of mutualism).

Community Set of coexisting species in an ecosystem.

Food web A set of trophic relationships among species in a community; schematically, the pattern often resembles a web of species each connected by trophic interactions with other species.

Functional dependency Dependency of one species on another to complete a particular ecosystem process or function; an example would be the dependency of plants on decomposers and decomposers on plants to recycle nutrients.

Phylogeny A pattern of evolutionary descent among species, generally illustrated as a tree in which nodes represent common ancestors and the lengths of branches represent time since divergence; time since divergence is generally assumed to correlate with genetic relatedness.

Process intricacy Complexity of temporal or spatial patterns of ecological processes such as population dynamics or production.

Structural intricacy Complexity of patterns of links or connections among species in a community that are created by biotic interactions, shared pathways of nutrient and energy flow, or phylogenetic relationships.

Trophic interaction Feeding relationship between two species; these include predation, herbivory, parasitism, bacterivory, frugivory, or any other interaction that involves individuals of one species consuming individuals or parts of individuals from another species; trophic interactions represent a subset of all biotic interactions in a community.

Ecological complexity and biological diversity are often presumed to be strongly correlated or even synonymous with one another, but there is no necessary relationship between these terms. Ecological complexity is a scale by which we compare ecosystems. At one end of the scale, systems that exhibit unpredictable behavior or many connections among its species (e.g., connected by biotic interactions such as predation, parasitism, facilitation, or competition, that influence the population dynamics or flow of energy and nutrients among species in an ecosystem) are considered "complex." At the other end of the scale, systems that exhibit highly predictable behavior or consist of largely independent or unconnected species are considered "simple." In contrast, biological diversity is a measure of the extent of genetic and phenotypic variation found within a community. Although these two characteristics of ecosystems are relatively distinct, they are, at a certain level, fundamentally related. For example, the number of potential connections (i.e., potential complexity) is a function of the number of species (i.e., diversity), even if the actual number of connections among species (i.e., realized complexity) may range from none to the full complement of possible connections (i.e., every species is connected to every other species).

Widespread changes in diversity experienced by most ecosystems have generated considerable interest in quantifying complexity and assessing whether current changes are affecting levels of ecological complexity in natural ecosystems. The resolution of issues surrounding the relationship between complexity and diversity is an active area in contemporary ecological research.

The Complexity of Ecologically Diverse Systems

> Hence, if certain insectivorous birds (whose numbers are probably regulated by hawks or beasts of prey) were to increase in Paraguay, the flies would decrease – then cattle and horses would become feral, and this would certainly greatly alter (as indeed I have observed in parts of South America) the vegetation: this again would largely affect the insects; and this, as we just have seen in Staffordshire, the insectivorous birds, and so onwards in ever increasing circles of complexity.
>
> – Charles Darwin, 1859

> It is reasonable to believe that in following the numbers of orders in the fossil record we are indeed following the approximate overall course of ecological complication and diversification even though in a necessarily loose way.
>
> – Simpson, 1969

> Roughly speaking, we here take complexity to be measured by the number and nature of the individual links in the food web.
>
> – May, 1974

> The complexity of a system is in the eye of the beholder. It is measured by how well we understand causes, expect behaviours

> and, in praxis, achieve purposes. Hence, large numbers of variables, nonlinear relationships among them, and the open nature of a system are important only to the degree they present barriers to understanding.
>
> – Holling, 1985

The Complexity of Nature and the Entangled Bank

Darwin provided us with a wonderful metaphor for what is intuitively meant by "ecological complexity" in his famous closing passage to *The Origin of Species*. In this passage, Darwin envisioned a riverbank teeming with a diverse and entangled array of organisms all involved in a complex web of biotic interactions. At the heart of this seemingly complex system, however, lies evolutionary and ecological processes that produce and maintain such remarkable diversity. For Darwin, biodiversity and the complexity that accompanies it are what distinguish nature from the inanimate features and processes of our world.

The entangled bank metaphor highlights two main ingredients of ecological complexity. First, there is biological diversity itself, with a clear implication that greater diversity provides a greater potential for complexity. Second, and more importantly, there is the web of interactions among species, with an equally clear implication that more complex webs (e.g., greater numbers and kinds of biotic interactions among species) are indicative of more diverse communities. Biotic interactions, however, describe only one way in which species are linked to one another. We can expand our entangled bank metaphor to include others linkages or relationships among species such as phylogenetic relationships, functional relationships, and exchanges of material and energy among species.

This entangled bank metaphor also highlights two distinct meanings of complexity. First, complexity refers to the intricacy of the pattern of linkages among species. This is the meaning used in many studies that have related ecological complexity to ecological stability beginning with MacArthur's (1955) classic work examining ecosystem structure and May's (1974) work on ecological communities, to innumerable studies that examine food-web and community structure (McCann, 2000) or more generally the topology of ecological networks (Jordán and Scheuring, 2004; Montoya *et al.*, 2006). The second meaning applied to the term complexity is when nature defies easy characterization and cannot be reduced to a finite set of processes and principles that govern its order. This second meaning is used from the "macroecological" perspective (Brown, 1995), the hierarchical perspective of nature (O'Neill *et al.*, 1986), the perspective of nature as complex adaptive system (Levin, 2005), or when nature is viewed as a self-organizing system (Allen and Holling, 2010).

Figure 1 illustrates these different kinds of complexities. The top row of **Figure 1** shows the dynamics of an ecosystem or evolutionary process as an example of an ecosystem property that may exhibit simple or complex behavior. This process could be production, decomposition, or any other process representing the aggregate activities of species in the community. It could also be an evolutionary process such as origination or extinction. The remainder of **Figure 1** shows three fundamental kinds of linkages among species in communities: ecological networks (i.e., the network of biotic interactions among species), phylogenetic structure or the pattern of phylogenetic relationships among species, and functional structure or the pattern of functional dependencies among species.

The pattern of temporal or spatial variation of processes is known as process intricacy. The pattern of linkages, whether referring to an ecological network, phylogenetic structure, or functional structure, is known as structural intricacy.

Whether considering process or structural intricacy, exhibiting either highly ordered states or highly irregular or disordered states is considered complex. As shown in **Figure 1**, a temporal pattern for a specific ecological or evolutionary process that exhibits highly ordered, predictable, repeated series of oscillations would be considered complex. Yet a highly irregular, unpredictable, chaotic pattern of oscillation would also be considered complex when compared to a simple, mildly oscillating process. Similarly, **Figure 1** shows that a pattern of biotic interactions in which every species is linked to every other species in the community by exactly the same kind of interaction with exactly the same strengths of interactions creates a symmetrical, highly ordered structure that would be considered complex. At the same time, a community that has a jumbled network of biotic interactions that vary in type and strength would also be considered complex when compared to a simple set of independent, noninteracting species. Finally, if all species contribute identically to ecosystem function, the functional structure is simple, but if species contribute in different ways, ranging from partially redundant to completely unique, then the structure is complex.

Although ecological complexity is difficult to define precisely given the many different ways the term is used, when we consider the basic elements of ecological complexity it becomes apparent that diversity is not a necessary part of the definition of complexity. Note in **Figure 1**, for example, that all nine structures shown are made up of six species. In spite of having the same number of species, these structures differ quite dramatically in complexity.

Despite the fact that diversity is not a necessary part of the definition of complexity, diversity and complexity are related in an elementary yet fundamental way. Few would debate that extremely depauperate ecosystems such as monoculture plantations of bananas are less complex than the extremely species-rich tropical rain forests they replaced, but for the bulk of ecological communities that lie somewhere between these extremes of biodiversity, determining just how complex a given system is can be quite difficult. For example, if a community contained 25 species of grasses whereas another contained only 10 species, but was made up of a mixture of three grasses, four forbs, and three shrubs, which community would be more complex? If a community consisted of 25 noninteracting species and another contained five species involved in a complex web of interactions, which would be more complex? If two communities each contained 10 species, but one exhibited chaotic dynamics whereas the other hovered close to a constant number of individuals per year in spite of repeated perturbations, which would be more complex?

Another way to look at the problem is to plot, simply for illustrative purposes, complexity against biodiversity as done in **Figure 2**. Aside from the trivial points where there is no

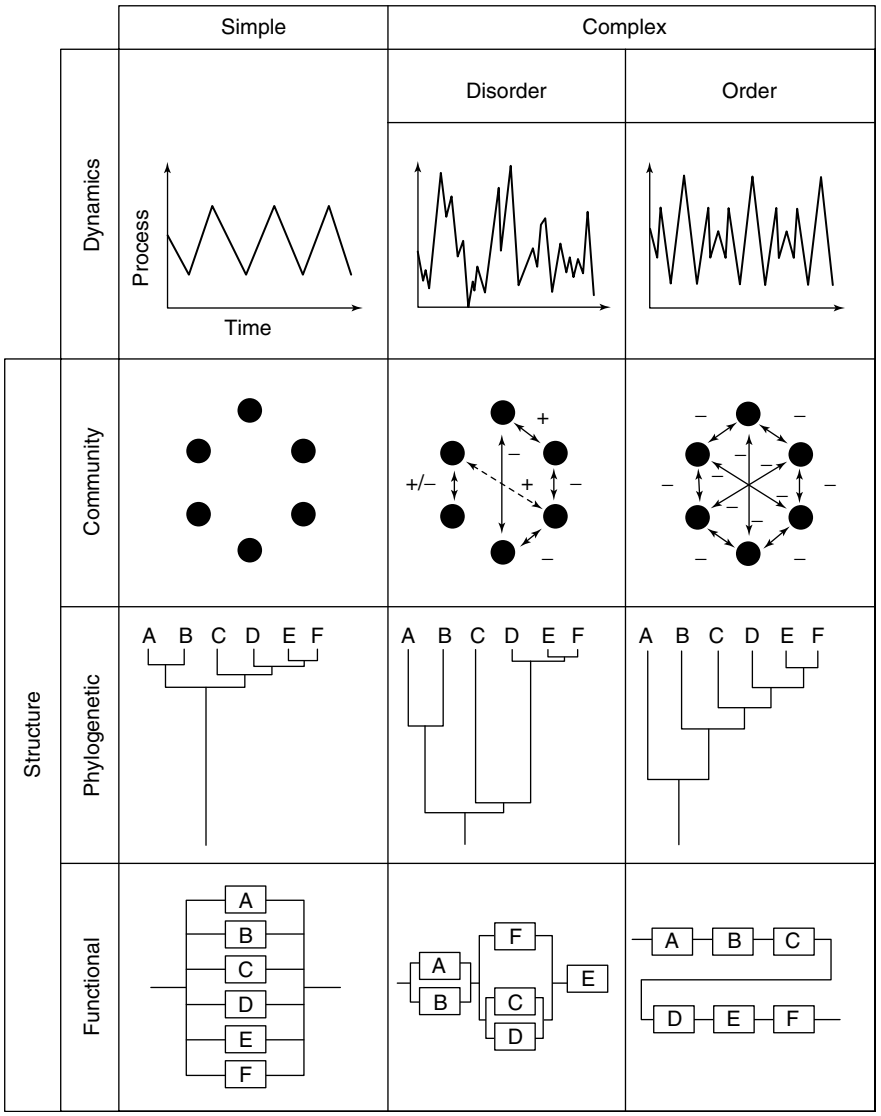


Figure 1 Simplicity and complexity in ecological systems. Dynamics are illustrated as temporal patterns in an ecological or evolutionary process, such as production, standing biomass, energy or nutrient flow, or rates of diversification. Community or network structure is illustrated as the number, type (e.g., competition or facilitation, shown as $-$ or $+$, respectively), strengths (thickness of arrows), and arrangement of interactions (arrows) among species (circles). If the only interactions portrayed are trophic interactions, then the network diagram would be that of a food web. Phylogenetic structure is illustrated as phylogenetic trees in which every two species share a common ancestor. The vertical lengths of branches indicate distances, assumed to be a measure of genetic similarity or time since divergence. The complex phylogenies of more distantly related species contain more genetic information than would be found in the simple phylogeny of many closely related species. Functional complexity is illustrated as species whose contributions to ecosystem processes are either largely dependent (serially linked) on the activities of other species or exhibit a diverse array of interdependencies among one another. A simple system is one in which all species contribute to the functioning of an ecosystem in the same way (parallel redundancy) such that the loss of a species from a system does not affect its functioning.

diversity, and therefore no ecological complexity, and where there is some diversity and there is some complexity, a variety of relationships could be drawn between the two points depending on one's perspective or experience. That is, complexity and diversity could either increase hand-in-hand or show any possible relationship that is bounded by a positive, asymptotic relationship between diversity and complexity on one end and an exponential relationship between diversity and complexity on the other end (see **Figure 2**). Of interest would be whether some general relationship could be described or whether every ecosystem would show a unique relationship. If the latter is true, then knowing a system's diversity would provide no means for predicting its complexity and whatever ecological behaviors might be a function of complexity.

Because ecological complexity is believed to be an important part of biologically diverse ecosystems and because diversity is currently rapidly declining, the role of ecological

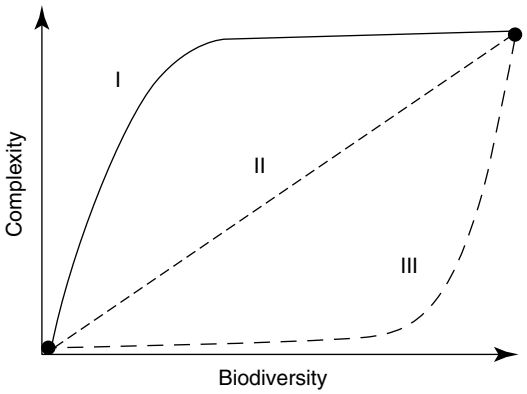


Figure 2 The potential relationships between complexity and biodiversity. Three possible relationships are shown. (I) The upper bound represented by an asymptote in complexity as biodiversity increases. (II) A linear relationship in which biodiversity is a direct measure of complexity in an ecosystem. (III) The lower bound represented by an exponential increase in complexity as diversity increases. Curve II represents the common perception of biodiversity and complexity being intimately related, but I and III represent the actual bounds for where the true relationship for a given ecosystem lies.

complexity is of considerable concern to ecologists. Does biodiversity loss mean loss of ecological complexity? If so, what are the ecological consequences of losing ecological complexity? This article reviews ecological complexity, its definition, its measurement, its relationship to biodiversity, and what is known about its role in ecosystem functioning.

Concepts and Definitions

In 2004, the journal *Ecological Complexity* became the first scientific journal focused, as its name implies, on ecological complexity. Li (2004), the managing editor, on behalf of the editorial board, provided the following definition of ecological complexity:

> "… the complex interplay between all living systems and their environment, and emergent properties from such an intricate interplay. The concept of ecological complexity stresses the richness of ecological systems and their capacity for adaptation and self-organization. The science of ecological complexity seeks a truly quantitative and integrative approach towards a better understanding of the complex, nonlinear interactions (behavioral, biological, chemical, ecological, environmental, physical, social, cultural) that affect, sustain, or are influenced by all living systems, including humans."

In contrast, the most widely used definition of biodiversity is taken from Article 2 of the 1992 United Nations Convention on Biological Diversity (CBD) which defines biodiversity as:

> "…the variability among living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic systems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems."

Li does not refer to biodiversity in his definition of ecological complexity nor does the CBD refer to ecological complexity in its text. Li does, however, refer to the richness (i.e., diversity) of ecological systems which is related to the CBD's reference to the variability among ecological complexes. The two definitions thus overlap in seeing complexity and diversity as relevant parts of each of their definitions.

Chapman (1985), in his review of a symposium on the concept of complexity, identified four distinct meanings of "complex" used in the conference and in the literature. First, there is what Chapman called "complication," or the process by which a system becomes naturally complex, but such complexity is ultimately "reducible" to fundamental principles or mechanisms. Second, there is "complification," which, like complication, concerns reducible systems, but involves purposeful design (e.g., a watch or a banana plantation has a different kind of complexity than the solar system or a tropical rain forest). Third, there is "complexification," which concerns irreducible systems (systems that are not composites of simpler parts). Finally, the fourth usage is "complexity" itself, which is simply a state of uncertainty concerning whether a system is reducible or not (irrespective of it being natural or of purposeful design). These decidedly ugly terms are not likely to catch on, but they do point to the variety of ways that ecologists, like other scientists, use complexity.

Community, Phylogenetic, and Functional Complexity

Typically, the web of interactions within a community, characterized by the number, nature, strengths, and structural pattern of biotic interactions among species, is often what is treated as ecological complexity. This can be the web of all interactions (community web, including competition, facilitation, parasitism, predation, and other interactions) or a subset of the community web, such as just trophic or feeding interactions (food web, including predation, herbivory, disease, or other consumer–resource species interactions).

In contrast to community complexity, however, phylogenetic complexity may be of interest when a community contains a diverse array of taxa. From the perspective of systematists or paleontologists, simply having many species that are phylogenetically diverse is evidence for complexity (Raven, 1996). A system of 15 species of grasses from the same genus might be considered less phylogenetically complex than a system of 10 species from 10 genera of grasses, which in turn might be less complex than a system with five species each from a different plant family.

Finally, complexity of an ecological system may be tied to its functioning as a biogeochemical system rather than to its population or biomass dynamics. That is, the flow of nutrients and energy through the system may be of interest rather than the dynamics of the many populations found within it. Note that ecosystems function in the sense that they exhibit activity or processes such as nutrient and energy flow. "Function" need not imply purpose or design.

From the foregoing, it is apparent that ecological complexity is a composite feature of ecosystems consisting of an extraordinary number of different factors, including population, phylogenetic, and functional characteristics of the

species in the community. The full measure of the complexity of a single ecosystem would therefore include a large number of factors, most of which are difficult to quantify. In general, increases in any of the properties listed in **Table 1**, whether associated with structural or process intricacy, tend to increase ecological complexity, whereas decreases tend to simplify them. **Table 1** demonstrates that measures of biotic diversity per se, such as species richness or evenness, are likely to be inadequate measures of complexity, though they may serve as crude proxy measures. The next section considers some basic measures of complexity, but not surprisingly, it will be evident that such measures of complexity generally deal with only single dimensions of complexity.

Measures and Consequences of Ecological Complexity

Measures of Complexity

Because complexity is a scale by which we compare communities and ecosystems, quantifying or measuring complexity is important in providing accurate means of making comparisons. Measuring complexity, however, is difficult because complexity is difficult to define. Boulding (1985) nicely described the problem at a United Nations conference held in Tokyo on the theory of complexity. He stated:

> "Complexity is a very complex concept, which is not altogether surprising. We do have a vague concept of it as a somewhat linear property, both of structures and processes, in which we make judgments of "more" or "less." I am pretty sure I am more complex than an amoeba... . Exact measurement of the concept eludes us and probably always will, for although we do make these estimations of it in terms of more or less, it represents a highly multidimensional reality that cannot be reduced to a linear measure."

In spite of the obvious difficulties Boulding outlined, ecologists and evolutionary biologists have nevertheless provided a number of measures of complexity, which are reviewed in the following sections. These measures can be used in conjunction with measures of diversity to assess the relationship between complexity and diversity.

Measures of Community Complexity

Measures of ecological complexity are closely allied with measures of diversity, which reinforces the common practice among ecologists of treating biodiversity and complexity as one and the same. At its simplest, one might consider diversity to simply be the number of species (S), or species richness.

Table 1 Measures of ecological complexity

Term	*Measure*
Structural intricacy	Numbers of populations
	Numbers of species
	Numbers of biotic interactions
	Kinds of interactions
	Strengths of interaction
	Higher order effects
	Indirect effects
	Connectedness
	Number of patches
	Arrangement of patches
	Immigration probabilities
	Emigration probabilities
	Number of functional groups
	Number of functionally singular species (inverse of redundancy)
	Number of links among functional groups
	Number of compound functional groups (e.g., omnivores)
	Distribution of species' functional traits (means and variances)
	Covariance among species' functional traits
	Distances or overlaps among species' functional traits
	Trait-space volume occupied by species in an ecosystem
	Phylogenetic distances among species
	Phylogenetic moment
Process intricacy	Invasibility
	Stability (Resistance, Resilience, Variability)
	Predictability
	Reliability
	Degree of nonlinearity
	Succession
	Population growth
	Ecosystem processes
	Extinction
	Speciation

Such a measure is often felt to be inadequate because it does not reflect differences in commonness and rarity nor the degree of connectedness among species. Thus, many other measures of diversity are used that attempt to take linkages or potential linkages into account. For example, the Shannon–Weaver index (H) considers both species richness and the relative abundances of species (evenness) as the important components of diversity. H quantifies diversity by assuming that richness and evenness directly represent the potential number of community configurations that a particular assemblage of species can produce. The original index quantified information, summing logged probabilities of a system having particular configurations.

The configuration of a community is seldom known, let alone the variety of possible configurations it can occupy or the probabilities of taking on these configurations. One therefore uses proportional abundances (p) of species as a means of approximating probabilities of configurations, a strategy that some consider fairly unsound.

The formula using this approach is

$$H = -\sum_{i-1}^{S} p_i \ln p_i \qquad [1]$$

where i refers to the ith species in the community.

A related measure of potential complexity that compares the full potential with the realized potential for configurations is called connectivity (C), or the proportion of actual interactions (links or connections) to total possible interactions, which is distinct from food-web connectance (see later). The formula for connectivity is

$$C = \sum_{i \neq j} a_{ij} p_i p_j \Big/ \sum_{i \neq j} p_i p_j \qquad [2]$$

in which p is the probability of a species being present, and a_{ij} is the probability of interaction when two individuals, each from a different species (i and j), encounter one another. In reality, identifying interactions in biological communities is very difficult and cannot be deduced from abundance data. Indeed, Margalef and Gutiérrez (1983) explored connectance by examining components and "interactions" in functional plans for construction toys because information of equivalent clarity from real communities was unavailable.

The Shannon formalism serves as the basis for the preceding and other measures of complexity and diversity, but this practice is questionable. The Shannon formalism quantifies complexity by estimating an upper bound for the number of configurations possible for a given set of entities as if all configurations would be functional. In reality, however, rules of assembly ultimately determine the actual limits of information that a set of entities can store and retrieve efficiently. For example, if one had a box of 26 unique letters (i.e., A–Z), each equally abundant, and a box of four unique nucleotides, each equally abundant, the Shannon formalism would consider the box of letters to have a greater capacity to store and retrieve information. Assembly rules for letters, however, must be followed to produce sentences containing useful information just as assembly rules for nucleotide sequences must be followed to produce useful proteins. It is quite possible that four nucleotides could code for a greater diversity of proteins in nature than 26 letters can code for words in English, the difference being attributable to the rules used in each system. Without knowing these rules, it is difficult to assess the actual information content for a collection of entities such as letters, nucleotides, or species. Although assembly rules for nucleotides and letters in human languages are well understood, assembly rules for species in communities, although well studied, are still poorly understood.

Appraisals of complexity based on static patterns of the distribution and abundance of species ((1) and (2)) ignore flows of energy or nutrients and the dynamics of interacting populations, all of which are considered key elements of community complexity (see **Figure 1**). Rutledge *et al.* (1976) adapted approaches based on information theory, as discussed earlier, to consider flows among species within a community. Their approach usefully distinguishes between and quantifies both flavors of complexity – structured versus disordered patterns of flows among species.

Appraisals of community complexity may also be made using the pattern of biotic interactions that links species in a community. The notion that biotic interactions are the primary elements of structural intricacy (complexity) that determine community dynamics has been the dominant theme in community ecology for several decades, inspired by the original contributions of Lotka and Volterra. Mathematical treatments of complexity have explored the consequences of varying interactions by constructing model communities whose species richness, interaction coefficients, and signs and strengths of interaction vary randomly and by examining their stability features. The heart of this approach is to use the "community matrix" as a measure of complexity, a concept formalized by Richard Levins in 1968 and still very much in use today. The general approach is to assume that all populations can be modeled by differential equations of the form

$$\frac{dx_i}{dt} = r_i x_i \left(K_i - x_i - \sum \alpha_{ij} x_i \right) \Big/ K_i$$

where r_i is the intrinsic rate of increase of the population of the ith species, K_i is the carrying capacity of the ith species, x_i is the population size of the ith species, and α_{ij} is the effect of the jth species on the ith species. For a community at equilibrium, where

$$K_i = x_i + \sum \alpha_{ij} x_j$$

the system of equations can be written as

$$AK \equiv K$$

Here, x is the column vector of species population sizes, K is the column vector of species-specific carrying capacities, and A is the community matrix

$$A = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdot & \cdot \\ \alpha_{21} & 1 & \alpha_{23} & \cdot & \cdot \\ \alpha_{31} & \alpha_{32} & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \qquad [3]$$

The properties of this matrix can then be used in analyses of complexity. Although wonderfully appealing as a logical measure of community complexity, identifying and quantifying interactions in a community is a formidable task and requires that a community be at or near equilibrium, rarely the case in nature.

One concern over measures of structural complexity based on biotic interactions is that they must account for "keystone" species. Defined originally by Paine (1966) as a predator species that fosters the coexistence of competing prey species and thereby promotes structural complexity in ecosystems, it has become colloquially used to describe any species that has a disproportionate impact on the community in comparison to other resident species (Liebold, 1996) or species that serve as ecological engineers (Jones *et al.*, 1994).

Power *et al.* (1996) proposed that the distribution of "community importance values" (CIVs) may serve as a tool for comparing communities for this element of structural complexity. Formally, the *i*th species in a community has a community importance value (V) defined as

$$V_i = [(m_N - m_D)/m_N](1/p_i) \qquad [4]$$

where $m_N \equiv$ the quantitative property (e.g., production or rate of decomposition) of the intact community, $m_D \equiv$ the quantitative property of the community with the *i*th species removed, and p_i represents the proportional abundance of the *i*th species before it was removed. In practice, one would sequentially delete one species at a time, each time measuring the community response to such a loss. Eventually a distribution of CIVs would be obtained and statistics (e.g., the mean or modal CIV and its variance) could be used to compare one community with another.

Communities with mean CIVs furthest away from 0 would be more complex than those whose distributions were not significantly different from a normal distribution with a mean of 0. Of course, such an experiment would be very difficult to conduct and Power *et al.* (1996) could identify only two such studies.

One approach to the intractability of contending with all species interactions is to focus on a subset of interactions. An example of such a focus is the study of feeding or trophic interactions. Typically, food-web statistics serve to quantify food-web structural complexity. These statistics include food chain length, connectance (the ratio of existing trophic links to possible links), compartmentation (the number of compartments or isolated subnetworks of interacting species in a food web), the ratio of predator species to prey species, and interval versus noninterval state (presence or absence of complex overlaps in resource sharing). The difficulty with all of these statistics is that they rest entirely on the quality of the set of food webs, many of which were not constructed for such analyses, and thus produce artifacts when subjected to such analyses.

Measures of Functional Complexity

From a functional perspective, complexity has been measured simply as the number of functional groups, but this approach is steadily being abandoned in favor of measures based on species' traits. Functional groups are collections of species that are related by their ecosystem activities, though many other definitions have been used. For example, if a functional group were defined as the group of species sequestering carbon from the atmosphere by photosynthesis, then all macrophytes, algae, and cyanobacteria found in a lake would be considered part of this group, in spite of their very different ecologies and evolutionary histories. This degree of aggregation is much larger than what is typically found in contemporary food webs (trophic groups or species), community matrices (interaction coefficients from species pairs), or high-resolution phylogenies (orders, families, genera, species, etc.). Species' traits are the phenotypes, characteristics, or properties of species that determine how a species responds to or influences their environment. For example, the average metabolic rate of individuals of a single species of mammal for a given set of conditions tells us how that species responds to its environment whereas the amount of carbon dioxide released by that species tells us its influence on the environment (i.e., carbon cycling).

In the same way that tallying up numbers of species fails to capture complexity, tallying up functional groups is unlikely to serve as an adequate measure of functional complexity. More important would be the degree of parallel versus serial dependency among species in an ecosystem. As in community complexity, network architecture is the focus for measuring complexity, but rather than linkages of biotic interactions, only linkages of dependencies among species and functional groups become important.

Ecosystem reliability is an example of an index that combines measures of serial and parallel dependency (Naeem, 1998). Serial dependency occurs when every species or every functional group is dependent on the other for the ecosystem to function (e.g., decomposers depend on producers for energy whereas producers depend on decomposers for nutrient cycling). Parallel dependency occurs when one or more species within a functional group substitute for each other on the extinction of one of the members of the group. In its simplest form, ecosystem reliability may be formally represented as

$$H(t) = \prod_{j=1}^{F} \left[1 - \prod_{i=1}^{s_j} (1 - e^{-\lambda t}) \right] \qquad [5]$$

where F is the number of functional groups, S_j, is the number of species in the *j*th functional group, and λ is the probability of extinction over a small increment in time, t. Note that the reliability of an ecosystem is defined as 1 ($H(0) \equiv 1$) and eventually decays in the absence of replacement of species by recolonization ($H(\infty) \equiv 0$).

As in other measures of complexity, the information needed for this index is difficult to obtain. Species presence or absence, local extinction rates, and functional roles are often difficult to determine without considerable empirical work. Focusing on redundancy demonstrates, however, that if the properties being measured are the result of the collected activities of groups of similar (i.e., redundant) species, such as production, having many species could actually mean more steady levels of production (less complex behavior). Thus, the association between biodiversity and complexity of ecosystem

functioning is dependent on the degree of species redundancy in the system.

Because functional groups are sometimes seen as arbitrary (Wright *et al.*, 2006), metrics based on species' traits have multiplied dramatically since Petchey and Gaston (2002) introduced the first index that became widely used in ecology. In Schluter *et al.*'s (2010) review of metrics of functional diversity they identified three classes of functional diversity metrics; measures of functional richness, divergence, and evenness. If one considers each trait of a species to be a different axis (e.g., metabolic rate, body mass, and temperature where a species is found), then each species occupies a specific volume in the space described by these trait axes. Three axes describe a three dimensional space, or trait space, which is easy to visualize, but the space can, of course, range from unidimensional to hyperdimensional depending on how many traits the researcher decides are important in their study. Functional richness concerns how much of that space is occupied by all the species in an ecosystem. Functional evenness concerns how uniformly the trait space is filled by species whereas divergence concerns how distant species are from one another in trait space.

Measures of functional diversity, because they are based on traits, their means, their variances, and their correlations with one another, are much more difficult to estimate than other measures of biodiversity, such as H, which are often based solely on the relative abundance of species. For illustration, we can consider Schleuter *et al.*'s measure of multidimensional functional richness (FR_{Im}) as an example of how one might quantify functional diversity. One version of this measure is

$$FR_{\mathrm{Im}} = \int \max_{s \in S_c}\left[\exp\left(-0.5(\boldsymbol{Z}-\boldsymbol{X}_s)^T\sum\nolimits_s^{-1}(Z-X_s)\right)\right]\mathrm{d}\boldsymbol{Z} \quad [6]$$

where s is the sth species, c is the cth community, S_c is the number of species in the cth community, T is the number of traits, $\boldsymbol{Z}$ is the vector of trait values, $\boldsymbol{X_s}$ is the vector of trait means for the sth species, $\max_{s \in S_c}$ is the maximum value of the bracketed function for each species in the set of species in the cth community, and $\sum_s^{-1}$ is the variance/covariance matrix of traits. Essentially, this metric estimates the space occupied in a trait space of T dimensions (one dimension for each trait used) by integrating over the distribution of maxima of every species occupancy (based on their means and variances for each trait) while correcting for covariation among traits (e.g., if two traits correlate perfectly then they should influence the metric as if they were one trait). This is neither easy to comprehend or calculate, though in principle the data needed are relatively straightforward to collect. Measures of functional evenness and functional divergence are equally complex.

It is important to note that although measures of functional diversity based on traits obviate the arbitrariness of assigning species to groups, which traits to use can wind up being another source of arbitrariness. In principle, one does not use any and all traits available or some arbitrary set, but uses the set of traits relevant to the ecosystem function of interest. Determining which traits are relevant requires modeling an ecosystem process or property as a function of species' traits.

Measures of Phylogenetic Complexity

Although community and functional complexity focus on structural aspects of ecosystems, phylogenetic complexity examines the structure of the phylogenies that relate species to one another within a community by their evolutionary history. As with community and functional complexity, the number of species is an important starting point, but the architecture of the evolutionary "tree" or "bush," as opposed to network architecture, serves as the primary basis for assessing phylogenetic complexity.

A straightforward measure of phylogenetic complexity might be simply the sum of the lengths in the tree that unites all species in the set being examined or a simple modification of the Shannon–Weaver index in which species are weighted by their taxonomic distinctness. More informative, however, is the "phylogenetic moment" (defined in eqn [6]), which provides a statistic by which communities might be compared. A difficult necessity for using this measure is that one has to have the complete phylogenetic tree at hand (referred to as tree T), including those species that might have gone extinct and might be missing from the local community being examined. The current community will consist of some subset of T (referred to as set R), and the local community that one is measuring will invariably be missing some species and contain an even smaller set of species (referred to as set r). If we assume that the tree (its nodes and branches) can be represented by a series of points in time, with a total of PT points for the whole tree, then the phylogenetic moment for this subtree (M_R) is

$$M_R = \sum_{p \in P_T} d_{pr_p} \quad [7]$$

where d_{pr_p} is the distance between a point on the tree and the nearest species in the set r. The smaller the phylogenetic moment is, the more closely the community represents the full phylogenetic structure (found in T). For example, in **Figure 3**, although the total distances among the species in trees R_1 and R_2 are the same, using the phylogenetic tree T, which includes extinct species F and G and missing but extant species B, shows that R_1 would have a smaller phylogenetic moment than R_2 because C would have more near neighbors than D. That is, A, C, and E will presumably contain genetic information more representative of the bulk of the species in the full tree (T) than A, D, and E.

Metrics of phylogenetic diversity can also include the relative abundance of species (Cadotte *et al.*, 2010) which can make them more comparable to indices such as the H, C, V_i, or FD_{Im}.

Utility of Measures of Ecological Complexity

Although several measures of complexity or proxies for complexity based on measures of diversity have been developed and explored ((1) through (7)), in reality no single measure is likely to adequately describe ecological complexity in all its

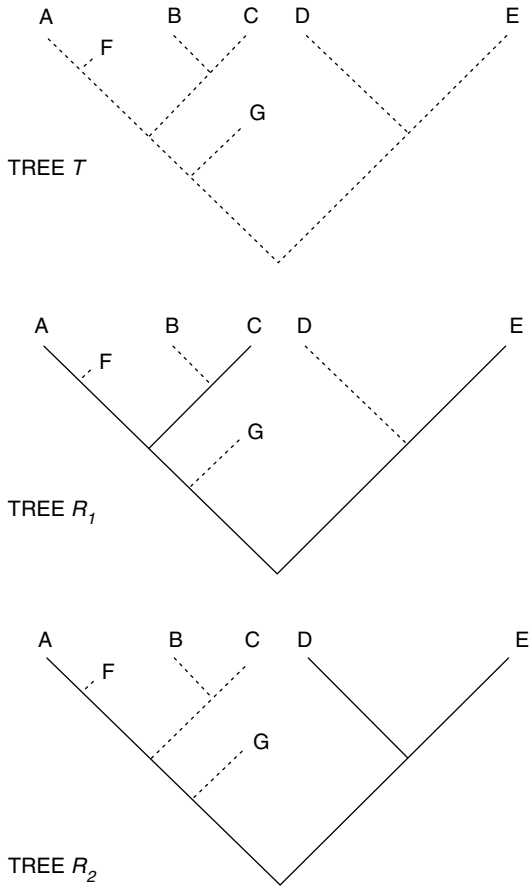


Figure 3 Hypothetical phylogenetic trees. Tree T represents the tree for all taxa in the community, including extinct species F and G. Two current local communities, each missing species B and one other extant species (dashed lines), are illustrated as R_1 and R_2. Solid lines show trees that would be used in calculating the total phylogenetic distance among species. Note that they are the same for R_1 and R_2. The phylogenetic moments for these trees would be the sum of the distances from each point on the tree T to the nearest species present. The moment would be smaller for R_1.

forms. Virtually any index based on a subset of the factors listed in **Table 1** is likely to produce rankings in complexity that will not agree with other methods, much in the same way that different diversity indexes may disagree in how they rank communities. The quantification of complexity is hindered by the difficulty in uncovering all species and all their interactions, their functional roles, and their phylogenetic relationships. As the next section shows, this means that we currently have only a partial picture of how complexity and biodiversity are related.

The Significance of Ecological Complexity

Complexity in Nature

Although diversity and complexity are practically universal features of ecosystems, we have much better information on diversity than on complexity. The estimation of existing diversity is, to be sure, challenging, but determining the distribution and abundance of species would only be the starting point for investigations of ecological complexity.

Unfortunately, the components of complexity are extraordinarily difficult to measure and for this reason empirical treatments have lagged behind the substantially greater progress made in theoretical explorations of the significance of ecological complexity. An ecologist never has a complete inventory of all species within a region, let alone their relative abundance, biotic interactions (direct, indirect, and higher order), phylogenetic relationships (at the regional level, including extinct species), functional roles, functional traits, and population and process dynamics, or other details listed in **Table 1**.

In most cases, however, the question to hand is more focused and thus more tractable. If the question concerns population stability, then quantifying only community complexity may suffice. If the question concerns what sets of species are most likely to contain the bulk of genetic information stored within a regional biota, then quantifying phylogenetic complexity may be enough. If the question concerns the magnitude or rate of ecosystem functioning, then quantifying functional complexity may do.

Community Complexity

An early and influential work that examined community complexity was MacArthur's (1955) study, inspired by Odum's theory that the "amount of choice which the energy has in following the paths up through the food web is a measure of the stability of the community." Using the Shannon–Weaver formula to model the number of alternative pathways and relate this to stability, MacArthur concluded that "stability increases as the number of links increases," though he was primarily considering resistance stability (remaining intact in spite of perturbation).

The theoretical relationship between complexity and stability is treated elsewhere in this volume, but it is worth noting here that this theory stimulated further investigations into the relationship between complexity and ecological processes, such as population dynamics. Later mathematical studies of local resilience stability (the ability to recover after perturbation) showed an inverse relationship between community complexity and resilience stability. Studies of connectivity and connectance (the nature and degree of connections among species) showed that resilience stability could either increase or decrease, depending on which trophic level was examined and the degree of connectance. Pimm (1979) also showed that complexity and stability could be positively related if one considered the deletion of species rather than resilience stability. Thus, those biologists who observed a positive association and those who observed a negative association between diversity and complexity more than likely differed in how they defined stability or how they defined complexity. Suffice it to say that there was agreement that complexity and stability were indeed related, but the nature of the relationship was dependent on one's definitions.

Although food-web studies initially showed much promise as a means for examining the relationship between complexity (community structure) and diversity by focusing solely on trophic interactions, early findings did not fare well in the test

of time. Polis and Winemiller (1996) summarize four patterns derived from early analyses of topographic or static food webs. First, there is constancy in the number of links per species in a food web irrespective of number of species. Second, the numbers of trophic levels seem to be invariably small (three or four). Third, the ratio of predator to prey species is relatively constant. Fourth, omnivory is rare. All of these patterns, however, have been challenged as artifacts.

Though potentially artifacts, food-web patterns at least touched on an obvious truth about communities – there are likely to be constraints in the structures of food webs. For example, at the simplest level, one cannot have consumers without resource species (i.e., no predators without prey) and there are likely to be stability or energy constraints that limit food chain lengths (Kaunzinger and Morin, 1998). Currently, food-web biology has gone far beyond the sorts of patterns that initially stimulated research in this field, but new patterns, if they exist, await the completion of a new set of higher resolution food webs.

The main observation to be drawn from studies of community complexity is that complexity is never at its maximum in diverse ecosystems. For example, although the potential number of pairwise interactions is large for any community, few of these interactions may ever be realized. As Margalef (1985) states, "the considerations of artifacts (from mathematical studies of diversity and complexity) provide an elementary but effective approach to the assertion that no system is completely interlocked or connected." We may infer from this fact that complexity asymptotes with diversity, an important consideration in our quest to understand the relationship between diversity and complexity (see **Figure 2**).

Phylogenetic Complexity

The relationship between phylogenetic complexity and biotic diversity is not well understood. Globally, estimates of origination and extinction of species suggest more or less steady increases in biotic diversity punctuated by aperiodic mass extinctions (Kirchner, 2002). This steady rise may be taken to mean that origination rates, more than millions of years, have been outpacing extinction, except during periods of mass extinction, and that recovery from mass extinctions has always occurred. Such a history of the biota suggests to some that the biota can tolerate even very deep "prunings" of the tree of life. At a global level, this suggests that random sets of species drawn from regional pools are likely to contain equivalent amounts of genetic information and that it takes severe levels of extinction to reach the point where insufficient genetic information is left in the remaining taxa to regenerate lost biodiversity. Because large-scale patterns in diversity across environmental gradients are similar across many taxa (Currie, 1991), we may expect that random samples from communities would result in phylogenies of similar complexity. Of course, since only a fraction of all species have been described, it is difficult to know exactly what the distribution and abundance of species are. As in community complexity and functional diversity, the true relationship between diversity and complexity is unknown, but it is likely to be asymptotic. That is, highly diverse systems may experience little change in phylogenetic complexity in the face of moderate levels of extinction.

Ecosystem Functioning

Biodiversity and ecosystem functioning are treated elsewhere in this Encyclopedia, but here the author examines its relationship to complexity and functioning. May (1974) noted in his "afterthoughts" to the second edition of his influential book that aggregate properties of communities, such as production (i.e., ecosystem functions), could show the opposite relationship with complexity that dynamic stability showed. Theoretical support comes from different approaches, but experimental tests of the idea are few. Support for this idea can be found in Tilman and Downing's (1994) study of experimental grassland plots. In their study, production recovery from a drought (resilience stability) increased with plant diversity. Two laboratory microcosm experiments using freshwater microbial communities by McGrady-Steed *et al.* (1997) and Naeem and Li (1997) showed that predictability and reliability of ecosystem functions, such as production and community respiration, can indeed be associated with biological diversity and complexity. The study of biodiversity and ecosystem functioning began in the 1990s and the number of studies has grown exponentially since then, but how complexity affects ecosystem functioning remains dominated by theoretical rather than empirical studies (Griffin *et al.*, 2009; Loreau *et al.*, 2002).

The shift away from taxonomic diversity to functional and phylogenetic diversity, which better relate to ecological complexity, has been on the rise in biodiversity and ecosystem functioning research (Cadotte *et al.*, 2008). Understanding the distinctions between functional, phylogenetic, and taxonomic diversity (Swenson, 2011) offers much promise for advancing our understanding of how ecological complexity influences ecosystem functioning.

Conclusions: The Disentangled Bank

Common usage of the term "complexity" in ecology implies that ecological complexity is a function of the phylogenetic, functional, and community structure of ecosystems in relation to temporal and spatial patterns in the distribution and abundance of species. Most natural processes, such as origination, local and global extinction, biogeography, demography, dispersal, biotic interactions, niche relationships, and energy and nutrient flows, to name but a few, contribute to overall complexity, which makes it difficult to define precisely or quantitatively. Biodiversity, which is a function of the spatial and temporal extent of genetic, phylogenetic, taxonomic, and functional diversity, is similarly difficult to treat with precision. These two constructs in environmental science are very real and perhaps best defined in the way the United States Supreme Court Justice Potter Stewart famously defined pornography; "… I know it when I see it." Some ecosystems are clearly more complex and some more diverse than others even if we have difficulty being precise about what it is we mean.

On the surface, ecological complexity and diversity seem so closely related that it is tempting to consider "biodiversity" as synonymous with "complexity." If we consider complexity to be the structural and process intricacies of a system, however, it is clear that diversity, in terms of numbers and types of organisms in a community, shows no necessary relationship

with any measure of complexity aside from the trivial fact that without any diversity there is no complexity. Thus, two communities identical in numbers of species could be quite different in community, phylogenetic, or functional complexity, as illustrated in **Figure 1**.

Studies of complexity in nature suggest asymptotic relationships between complexity and species richness, where species richness represents one way to define biotic diversity independent of complexity. It is hard to imagine, for example, that tropical rain forests, where species often number in the thousands, will show much reduction in complexity in the face of the loss of a few species. Current rates of extinction, however, are estimated to be staggeringly high and many diverse systems, such as tropical ecosystems, are not suffering extinction of species within their boundaries so much as they are being transformed to depauperate, managed ecosystems.

Thresholds in the response of ecosystems to the loss of diversity are particularly of interest (Levin, 1997) and these may be hypothesized from current knowledge about the relationship between diversity and complexity. Sometimes referred to as tipping points, critical transitions, regime shifts, or catastrophic change, the possibility of a rapid change in the dynamics, properties, or behavior of an ecosystem is of considerable concern in a rapidly changing world (Scheffer *et al.*, 2009). Habitat modification such as conversion to agriculture, habitat degradation due to pollution or mining, over exploitation of biotic resources such as deforestation or overharvests that lead to collapsed fisheries, and ecological homogenization due to widespread transport of species among communities, to name but a few aspects of global change, all contribute to ecological simplification, or the loss of ecological complexity. From what is known about nature, both theoretically and empirically, it is reasonable to postulate that both community and phylogenetic complexity show asymptotic relationships with diversity. Because functional complexity aggregates many species into small numbers of functional groups (e.g., all plants may be considered producers), we can postulate that the asymptote for functional complexity will be lower than that for community or phylogenetic complexity. When these are plotted together, as shown in **Figure 4**, we can identify two thresholds of interest. At some point (admittedly difficult to define precisely without accurate measures of community complexity or phylogenetic complexity), one will continue to experience no appreciable loss in functional complexity, but ecosystem properties, such as dynamic stability, may change and genetic resources will decline. As diversity declines even further, functionality is eventually lost after the second threshold (the leftmost dotted line in **Figure 4**) is crossed.

Managed ecosystems, such as aquaculture systems, farms, or monoculture tree plantations, most likely reside at the lower threshold, where the managers maintain the system at its minimal diversity. In general, these ecosystems may be functionally complete, containing all necessary functional groups to ensure that the desired ecosystem functions are maintained or, if not functionally complete, humans provide missing functions, such as fertilizer addition where nitrogen fixers have been removed. Complex systems are always more difficult to manage, but there are likely costs and

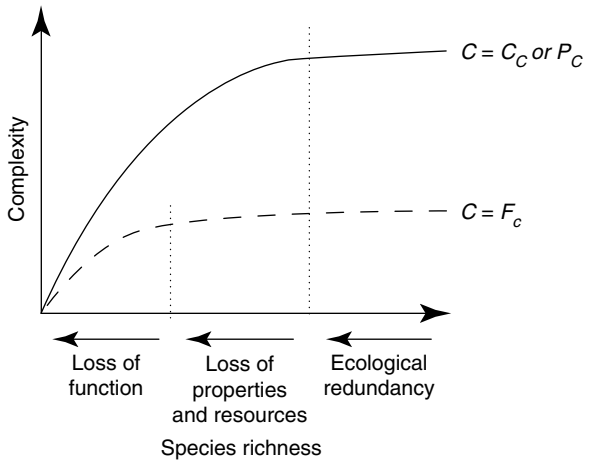


Figure 4 Hypothetical relationships between complexity and species richness. $C \equiv$ complexity. Vertical dotted lines indicate where the threshold in ecological consequences of declining ecological complexity might be identified. $C_c \equiv$ community complexity; $F_c \equiv$ functional complexity; $P_c \equiv$ phylogenetic complexity.

consequences for resorting to simple systems, many of which have to do with system stability, system utility, and system efficiency, as the material in this article reveals.

The appropriate metaphor for the process of human transformations of ecological systems may be the disentangled bank, or the converse of Darwin's entangled bank. As the earth's ecosystems are steadily transformed to agro-ecosystems and aquaculture systems, there is no question that average levels of biodiversity in ecosystems will continue to experience sharp declines in local biotic diversity, which eventually leads to losses in complexity that can, in turn, lead to transitions in ecosystem and Earth-system functioning that threaten all of humanity (Rockstrom *et al.*, 2009). Fortunately, as our methods for studying and understanding the nature and significance of ecological complexity and biodiversity improve, solutions will emerge that can assure a sustainable future and avoid unwanted change in the functioning of our ecosystems or our crossing safe planetary boundaries.

Appendix

List of Courses

1. Community Ecology
2. Conservation Biology
3. Sustainable Development
4. Biodiversity
5. Human Ecology

See also: Diversity, Community/Regional Level. Food Webs. Genetic Diversity. Landscape Diversity. Measurement and Analysis of Biodiversity. Species Diversity, Overview

References

Allen CR and Holling CS (2010) Novelty, adaptive capacity, and resilience. *Ecology and Society* 15. Article 24.

Boulding KE (1985) Learning by simplifying complexity: How to turn data into knowledge. *The Science and Praxis of Complexity*, pp. 25–34. Tokyo: United Nations University.

Brown JH (1995) *Macroecology*. Chicago: University of Chicago Press.

Cadotte MW, Cardinale BJ, and Oakley TH (2008) Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences* 105: 17012–17017.

Cadotte MW, Davies TJ, Regetz J, Kembel SW, Cleland E, and Oakley TH (2010) Phylogenetic diversity metrics for ecological communities: Integrating species richness, abundance and evolutionary history. *Ecology Letters* 13: 96–105.

Chapman GP (1985) The epistomology of complexity and some reflections on the symposium. In: Aida S (ed.) *The Science and Praxis of Complexity*, pp. 357–374. Tokyo: United Nations University.

Currie DJ (1991) Energy and large-scale patterns of animal- and plant-species richness. *The American Naturalist* 137: 27–49.

Griffin JM, O'Gorman EJ, Emmerson MC, *et al.* (2009) Biodiversity and the stability of ecosystem functioning. In: Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing*, pp. 78–93. Oxford: Oxford University Press.

Holling CS (1985) Perceiving and managing the complexity of ecological systems. In: Aida S (ed.) *The Science and Praxis of Complexity*, pp. 217–227. Tokyo: United Nations University.

Jones CG, Lawton JH, and Shachak M (1994) Organisms as ecosystem engineers. *Oikos* 69: 373–386.

Jordán F and Scheuring I (2004) Network ecology: Topological constraints on ecosystem dynamics. *Physics of Life Reviews* 1: 139–172.

Kaunzinger CMK and Morin PJ (1998) Productivity controls food-chain properties in microbial communities. *Nature* 395: 495–497.

Kirchner JW (2002) Evolutionary speed limits inferred from the fossil record. *Nature* 415: 65–68.

Levin SA (1997) Biodiversity: Interfacing populations and ecosystems. In: Abe T, Levin SA, and Higashi M (eds.) *Biodiversity: An Ecological Perspective*, pp. 277–288. New York: Springer-Verlag.

Levin SA (2005) Self-organization and the emergence of complexity in ecological systems. *BioScience* 55: 1075–1079.

Li BL (2004) Editorial. *Ecological Complexity* 1: 1–2.

Liebold MA (1996) A graphical model of keystone predators in food webs: Trophic regulation of abundance, incidence, and diversity patterns in communities. *The American Naturalist* 147: 784–812.

Loreau M, Downing AL, Emmerson MC, *et al.* (2002) A new look at the relationship between divesity and stability. In: Loreau M, Naeem S, and Inchausti P (eds.) *Biodiversity and Ecosystem Functioning: Synthesis and perspectives*, pp. 79–91. Oxford: Oxford University Press.

MacArthur R (1955) Fluctuations of animal populations and a measure of community stability. *Ecology* 36: 533–536.

Margalef R (1985) Ecosystems: Diversity and connectivity as measurable components of their complication. In: Aida S (ed.) *The Science and Praxis of Complexity*, pp. 228–244. Tokyo: United Nations University.

Margalef R and Gutiérrez E (1983) How to introduce connectance in the frame of an expression for diversity. *The American Naturalist* 5: 601–607.

May RM (1974) *Stability and Complexity in Model Ecosystems*. Princeton: Princeton University Press.

McCann KS (2000) The diversity-stability debate. *Nature* 405: 228–233.

McGrady-Steed J, Harris PM, and Morin PJ (1997) Biodiversity regulates ecosystem predictability. *Nature* 390: 162–165.

Montoya JM, Pimm SL, and Solé RV (2006) Ecological networks and their fragility. *Nature* 442: 259–264.

Naeem S (1998) Species redundancy and ecosystem reliability. *Conservation Biology* 12: 39–45.

Naeem S and Li S (1997) Biodiversity enhances ecosystem reliability. *Nature* 390: 507–509.

O'Neill RV, De Angelis DL, Waide JB, and Allen TFH (1986) *A Hierarchical Concept of Ecosystems*. Princeton: Princeton University Press.

Paine RT (1966) Food web complexity and species diversity. *American Naturalist* 100: 65–75.

Petchey OL and Gaston K (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters* 5: 402–411.

Pimm SL (1979) Complexity and stability: Another look at Mac-Arthur's original hypothesis. *Oikos* 33: 351–357.

Polis GA and Winemiller KO (1996) Food webs: What do they tell us about the world? In: Polis GA and Winemiller KO (eds.) *Food Webs*, pp. 1–24. London: Chapman and Hall.

Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *BioScience* 46: 609–620.

Raven PH (1996) Biological complexity. In: Pullman B (ed.) *The Emergence of Complexity in Mathematics, Physics, Chemistry, and Biology*, pp. 251–265. Vatican City: Pontificia Academia Scientiarvm.

Rutledge RW, Basore BL, and Mulholland RJ (1976) Ecological stability: An information theory viewpoint. *Journal of Theoretical Biology* 57: 355–371.

Rockstrom J, Steffen W, Noone K, *et al.* (2009) A safe operating space for humanity. *Nature* 461: 472–475.

Scheffer M, Bascompte J, Brock WA, *et al.* (2009) Early-warning signals for critical transitions. *Nature* 461: 53–59.

Schleuter D, Daufresne M, Massol F, and Argillier C (2010) A user's guide to functional diversity indices. *Ecological Monographs* 80: 469–484.

Swenson NG (2011) The role of evolutionary processes in producing biodiversity patterns, and the interrelationships between taxonomic, functional and phylogenetic biodiversity. *American Journal of Botany* 98: 472–480.

Tilman D and Downing JA (1994) Biodiversity and stability in grasslands. *Nature* 367: 363–365.

Wright JP, Naeem S, Hector A, *et al.* (2006) Conventional functional classification schemes underestimate the relationship with ecosystem functioning. *Ecology Letters* 9: 111–120.

Computer Systems and Models, Use of

Louis J Gross, University of Tennessee, Knoxville, TN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aggregation Combining several potentially separate components of a system to simplify analysis.
Dynamic model Mathematical description of a system with components that vary in time.
Model A construct, physical, mathematical, or computational, that is a simplification of reality developed to meet certain objectives.
Multimodel Single integrated model that links together models utilizing different mathematical or computational approaches.
Parameter Constant in a model that must be estimated from data, or assumed to be of a particular value.

What Is a Model?

Just as storytellers can take their audience on trips to faraway places and provide a glimpse of life in different cultures, scientists tell their stories about the way the world works by making models. Such models never provide a complete view of how the world works, but do give us glimpses that help us to piece together interactions between different parts of the world and the processes that connect them. These models take many forms, some being mostly verbal, others mostly qualitative and graphical, some phrased in various mathematical forms, and still others set up as collections of rules within a computer program.

Models provide maps of varying levels of complexity to help us understand the topography of science. There are coarse road maps that provide merely the outline of major arteries for traffic, telling us nothing about buildings or other features of the landscape, but providing an overview of the linkages between key components of a system. More elaborate models show us the buildings and the infrastructure that links these buildings – the sewer and power lines. Even more complex models would indicate the humans in each building, their occupations, and the flow of money or capital goods between them. Similarly, models at many levels of complexity can be useful for addressing different questions in the study of biodiversity. A very coarse model might analyze the effect of land use change worldwide on total species richness. A more complex model could consider particular regions and analyze differentially the changes in land use within them and the associated changes in species richness. A still more complex model could consider the local dynamics of a population of a particular species and from this elaborate the dynamics of the population's abundance as well as the abundances of populations of other interacting species. A further level of complexity in model formulation would incorporate the feedbacks of land use on climate change, connect this to socioeconomic models that account for ecosystem services, and project how these forces impact biodiversity from local to global scales.

The Purposes of Modeling

General Objectives of Models

Which type of model one constructs depends on the questions being asked and the availability of data to construct a reasonable model. Though there are many specific purposes for constructing models (Haefner, 2005), these may be grouped into a few general objectives: description, mechanism, prediction, and control. These objectives are not mutually exclusive, so that descriptive and mechanistic approaches may be used to aid prediction and control.

Description

Sometimes all that might be desired is a simple description of a collection of data. For example, an arithmetic or geometric average provides a single value to summarize a list of numerical data. The single descriptor used may depend on the type of data – for example, a geometric mean is typically more appropriate as a metric for summarizing varying growth rates than an arithmetic mean. A single descriptor may be sufficient for some purposes that do not require knowledge of how much variation is in the data. To assess variation, a dispersion measure such as variance would be needed. These summary statistics are coarse, ignoring many of the details in the data. Yet they do allow us to easily comprehend major differences between different data sets.

Extensive species lists within certain taxa from two locations may be compared by considering just the total numbers of species in the two locations and the number of species in common between them. Such a summary may be sufficient for a comparison of the two locations, while ignoring details such as the diversity within the taxa included. Descriptive approaches may be much more complex than simply providing averages and variances. Options for analysis of ecological data have grown tremendously due to the ready availability of computational platforms and open-source tools such as R. Bayesian methods and exploratory statistics approaches are applied to analyze complex multidimensional data (Whitlock and Schluter, 2008). Spatiotemporal analysis as an outgrowth of traditional time-series approaches is now feasible. In these the histories of species richness or abundances might be compared between locations or correlated with the histories of anthropogenic actions in the locations. Geographic Information Systems provide the functionality to analyze biodiversity across multiple regions, accounting for landscape metrics such as distance and connectedness.

Mechanism

If the objective is to provide an understanding of how a particular system operates, then it is necessary to take account of the processes that govern the system. While all such mechanistic models are descriptive at some level, the point is to deal with the basic physical, chemical, and biological processes operating in the system. This requires including those processes that operate at a spatial and temporal extent appropriate for the problem one is addressing, and ignoring others. Thus, analysis of how alternative global warming predictions would affect worldwide biodiversity might include the geographic variation in the temperature predictions at a spatial extent of hundreds of square kilometers, but would no doubt ignore the microclimate variation of every square meter. Even if it were possible to characterize the meter-by-meter temperature differences predicted to occur according to alternative warming trends, the lack of available detail on the species present at this detailed spatial resolution limits the utility of including such detail. Despite the lack of data at fine spatial resolution, it is feasible to consider hybrid methods that incorporate mechanistic biophysical models for constraints on species' presence linked to either statistical models based on climate data or down-scaled climate models. (Buckley *et al.*, 2010). A discussion of scaling issues is included elsewhere in these volumes.

Prediction

Predictive models are of two general types: those that attempt to project the behavior of the system based on certain explicit assumptions, and those that attempt to forecast the future behavior. The difference is between what might be true in the future if certain assumptions hold (projection) and what will be true in the future (forecasting; Caswell, 2001). In many biological situations, the forecasting problem is not even attempted, as it would involve taking account of a wide variety of unpredictable abiotic phenomena (e.g., hurricanes and droughts). It is often feasible to construct a model to project over some time period the future dynamics of a system based on current observations and particular reasonable assumptions about the interactions in the system. The majority of population models (discussed elsewhere in these volumes) are of this form, in which abiotic influences are not included. These models can project the future behavior of the population based on the biotic forces of demographics, genetics, and social structure within the population. Similarly, species distribution or niche models can project shifts in spatial patterns of species' presence based on assumptions on spatial climatic patterns. Uncertainties associated with unpredictable phenomena can be taken into account by attempting to project just the mean and variance of the variables of interest (e.g., population size), or by developing a stochastic model that incorporates assumptions about the probabilities of alternative future conditions. In the latter situation, projections would typically arise from numerous alternative simulations accounting for the assumed probabilities of different conditions, and the projections are therefore not single values or single spatial maps, but rather probability distributions for the values of interest.

Control

When a system has one or several components that are under human control, either completely or in part, then a model can be used to help determine how to apply such a control to meet certain objectives. Examples of controls are harvest quotas, limits on fertilization or pesticide application, flows from a dam, limiting importation of potentially harmful invasive species, and land use zoning regulations. Examples of objectives are maximizing biodiversity, minimizing population extinction probability, reducing the spread of nonnative species, and maintaining population size above some determined threshold (such as a minimum viable population size). Control models mostly focus on the dynamics of the system, with the simplest form of control being bang/bang, meaning on/off, such as allowing harvest in certain years and not allowing harvest in other years. A related objective is for control models to produce a relative comparison of alternatives in order to rank these alternatives according to some criteria (DeAngelis *et al.*, 1998). Still other control models are used to analyze the physiological responses of individual organisms to varying environments and the homeostasis that can arise through these responses. Spatial control involves choosing what action to carry out, where to do it, and when to apply the action. Many problems in spatial management of natural systems can be phrased as optimization problems in which the actions taken at different locations are constrained, often due to some overall budget limitations, with an objective to maximize some characteristic of the entire landscape, such as a diversity metric (Hof and Bevers, 2002).

Types of Models

Models can be physical, such as animal models used in drug testing and airplane models used in wind tunnels. In biodiversity contexts, microcosms and mesocosms, which are limited biological systems built in a laboratory setting, play this role. These are meant to mimic the key biotic forces interacting within a natural system, but are constructed at a spatial extent that allows for easy observation and controlled experimentation. They cannot include all of the components of the real system, but do allow for projection of how the real system might respond under particular perturbations. A wide array of model systems have been proposed as appropriate to evaluate responses of organism behavior and species interactions. Physical models are clearly limited, particularly to organisms that are mostly sessile or have very short distance movements.

Mathematical models come in a wide variety of forms. Some are simply graphical relations that show the qualitative relationship between certain components of a system, mainly to demonstrate the shape of response and whether one component increases or decreases with another. An example would be the increase and then decrease in species diversity along a gradient from low to high frequency of disturbance (**Figure 1**). Here, there is no attempt to predict at exactly what disturbance frequency the exact peak in species diversity occurs. Rather, the objective is to illustrate the qualitative behavior of diversity, representing the "intermediate disturbance hypothesis," which

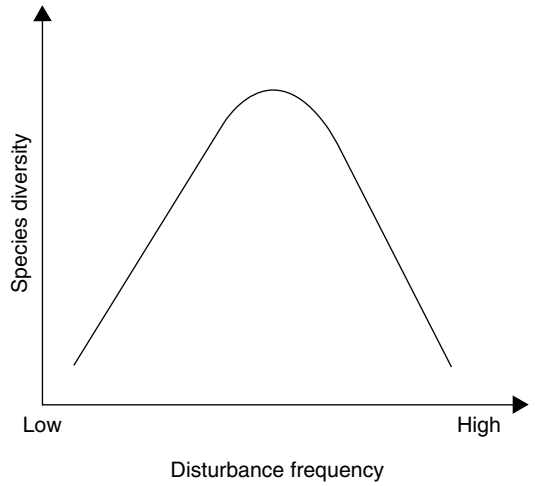


Figure 1 Illustration of the intermediate disturbance hypothesis.

posits that high diversity occurs at intermediate disturbance frequencies.

The majority of mathematical models in ecology deal with the dynamics of populations and communities. Such models consider the basic processes of birth and death, immigration and emigration, and competition and predation to elucidate general theories of population dynamics. Described using differential or difference equations, these models allow for projection of the long-term behavior of populations, and provide methods to project the within-population structure (age, size, genetic, etc.) over time and space. These models can be expanded in numerous ways to incorporate between-species interactions, account for demographics of births and deaths, use discrete spatial units (e.g., islands or patches) as well as continuous spatial spread, and deal with the stochastic aspects arising particularly in small populations (Murray, 2007).

Computational approaches have become a standard tool in every area of ecology. Computers are used to analyze complex data, interpret the implications of mathematical models by carrying out numerical schemes to solve the mathematical models, and to create models that are not readily described through any other means than through computer code. A computer system is made up of the various components of a computer, including the central processing units, graphics processing unit, monitor, hard drive, and input/output devices, as well as software and peripheral devices such as external storage, along with network connections to other computers and external data sources. Computer systems designed for ecological monitoring and real-time analysis, as are being constructed for the National Ecological Observatory Network (www.neoninc.org), also incorporate specially designed workflow to visualize, manage, and analyze a variety of incoming data streams, and provide remote access to diverse users. Cyberinfrastructure projects, such as the iPlant Collaborative (www.iplantcollaborative.org), develop the computer systems and associated products and tools necessary to link information at various scales of investigation, from the genome to the whole organism, to the ecosystem. This article emphasizes the use of computers to model biological processes, with the realization that technology affects our capacity to develop and investigate natural systems. The expectation is that our ability to mine large data sets arising from field-based and remote sensors, construct computational models from this, and evaluate the utility of these models will continue to expand. The advent of parallel computation, in which computational tasks are split among numerous processors, has greatly expanded the speed of computation, the range of problems we can address, and the methods we can use to address them. The flourishing use of Markov Chain Monte Carlo methods to carry out complex Bayesian modeling is one example of the continuing potential for advances in computational power to affect the science of biodiversity.

Computer models are quite varied in structure. All the standard mathematical models of populations and communities, constructed using differential or difference equations, may be implemented on computers. Indeed, since it may be quite difficult to develop analytical solutions for such models, analysis of their behavior often requires the use of numerical solution methods implemented on a computer. There are many computer models that, although they may have a description that is essentially mathematical in form, are really described by the computer code itself rather than an explicit set of mathematical equations. An example would be cellular automata models, one type of which consists of a two-dimensional lattice, with each point on the lattice having one of a number of states. The simplest situation would be each lattice point being occupied (e.g., in the 1-state) or unoccupied (in the 0-state). The model is then described by a set of rules that determines how the state of a lattice point changes from one time step to another, based on the states of surrounding lattice points. Such a cellular automaton may be used to mimic the spatial dynamics of populations, in which each lattice point represents a possible location of an individual. Alternatively, each lattice point can be interpreted as a local population, and the entire lattice can then follow the collection of such populations, called the metapopulation. It has been argued that the major questions in science are feasible to be investigated using cellular automata (Wolfram, 2002).

System simulation models are more elaborate computer models that attempt to include most of the major biotic and abiotic factors that affect the system. Many agricultural system models are of this type, and include the crop, its pests, soil nutrients, and weather conditions, among other factors. Some other types of computer models are described in later sections. In all cases, though the model is in essence specified by the code itself, it is very useful to have some graphical description of the major components of the model. One example is shown in **Figure 2**. A number of general modeling software packages are designed explicitly to aid construction of computer models through the use of graphical elements. These packages provide users with the capability to rapidly develop and compare models with differing numbers of components, for example, differing species and associated trophic interactions between them, by automatically generating the computer code for the models and allowing users to vary the basic model parameters such as birth rates. Each of these tools has limitations, for example, in the solution methods they employ, but as long as users are aware of these the tools can greatly

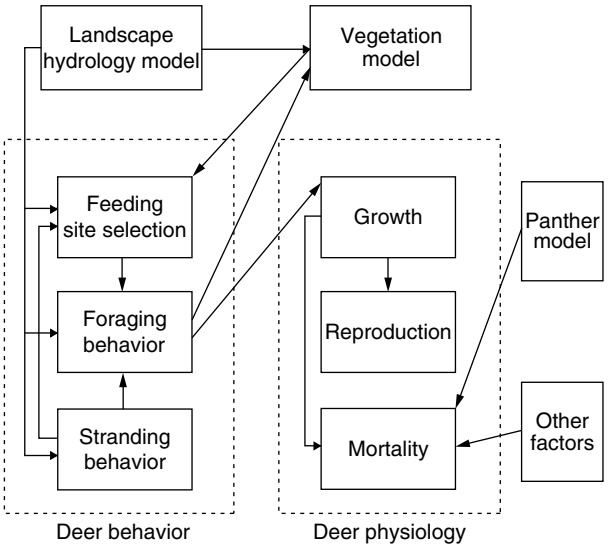


Figure 2 Graphical depiction of the major components of an individual-based model for deer.

reduce the time needed to construct computer models with many interacting components.

Limitations of Models

Trade-offs: Generality, Precision, Realism

No single model can do everything. In the process of deciding what components of a system to include, what processes to consider, and what spatial and temporal extent is appropriate, the model excludes part of reality. Modeling is a process of selective ignorance. We decide what to include and what to exclude. Part of the art of modeling is coming to grips with the issue of which details are important and which ones are not. In most cases the process is iterative, with a sequence of different models being tried until a model is arrived at that includes just the essential details necessary to address the problem of concern. Inherent in the process of modeling is that there are criteria by which the model will be evaluated. These criteria might be best chosen prior to model development, based on the objectives for which the model is being constructed.

One view of the trade-offs in constructing a model is that no single model can be simultaneously general, precise, and realistic (Levins, 1968). As **Figure 3** illustrates, these properties may be viewed as points of a triangle. Generality implies that the model may be useful in many different natural systems. A realistic model is one with components, parameters, and variables that are all possible to estimate from observations. A precise model is one that produces quantitative, accurate descriptions of the natural system. Models for theory development, including most of the classical population and community models, are quite general and somewhat realistic, but lack in precision. Descriptive models designed to mimic the response of particular systems tend to be quite precise, slightly realistic, and not at all general. Most regression models

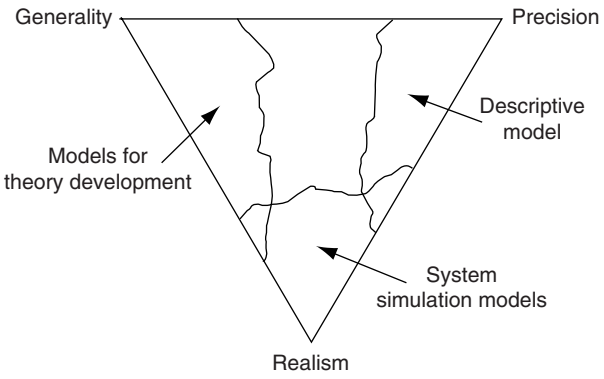


Figure 3 Trade-offs in modeling.

are of this type. They may provide an accurate portrayal of a particular system, for example, winter wheat growth in Nebraska, but are not transferable to other situations such as winter wheat growth in eastern Russia. System simulation models tend to be quite realistic and somewhat precise, but not very general. Control models take up various positions in the figure, depending on the level of precision desired.

Aggregation and Loss of Detail

A factor that affects where a particular model fits into the scheme shown in **Figure 3** is the amount of aggregation included. Natural systems consist of many components that can be lumped together or disaggregated. Population models that use a single variable to represent the whole population inherently ignore the within-population structure (e.g., age and size) or equivalently incorporate assumptions about the distribution of this structure within the population (e.g., stable age distribution). Such a model would not be able to discriminate between a population with mostly small individuals and one with mostly large individuals, unless this within-population structure was assumed to affect the population's growth characteristics in some manner. Aggregation implies a loss of detail that allows consideration of quite general models. Often this then incorporates model parameters that are not at all easily estimated from observations. Examples would be the growth rates in population models and the competition coefficients in community models. The less the aggregation in the model, the more the parameters there tend to be and more data are required in order to estimate them.

Uncertainties: Mechanisms, Data for Parameterization, Biotic and Abiotic Forcing Functions

The modeling process is limited by the information available. There may not be basic agreement on the mechanisms that are critical in the system of concern, so that any particular model includes only some of the mechanisms, or just partial information on these mechanisms. Many population models applied to vertebrates ignore social structure, despite evidence that it is often present in such populations. Excluding some mechanisms in a model may occur due to lack of understanding of the effect of such mechanisms at the scale of

interest for the model. This can give rise to another set of models designed specifically to investigate these effects. So there might be a hierarchy of models, each of which disaggregates some components in order to investigate the potential impacts of uncertainties in these components on the overall model results of particular interest.

Under situations in which the mechanisms are well understood, it may not be possible to accurately estimate model parameters because adequate data are lacking. A variety of statistical methods are designed specifically to determine the optimal choice of parameters in such situations (Hilborn and Mangel, 1997). These methods may take account of parameters estimated either for similar species or for similar locations other than the one being considered. For example, it may be necessary to use observed clutch size distributions for one bird species in application to a similar species about which such information is lacking. Another uncertainty associated with natural system models arises due to the unpredictability of forcing functions such as weather and disturbance. If historical information is available on these functions, this may be used to estimate the stochastic effects of such forcing. Simple ways to incorporate unpredictability of forcing functions would be to determine how the model responds to the maximum and minimum observed values of these forcing functions, and compare this to the model results when using the mean value of the forcing function. This is one form of uncertainty analysis in which a protocol is used to evaluate the impact of uncertainties in model inputs and model structure on the results of the model. Computational methods make it very feasible to evaluate model response to more elaborate and realistic assumptions on forcing by making numerous model simulations under different forcing function assumptions. For example, Fuller *et al.* (2008) evaluate the robustness of Everglade restoration scenarios under diverse assumptions about future rainfall conditions. Various uncertainties limit the detail at which models may be constructed, and thus limit the types of questions that may be addressed using models.

Some Tools of the Trade

Statistical Approaches

Statistical models usually have a descriptive objective rather than a mechanistic one. The parameters within these models are directly estimated by choosing them in a manner that best fits a certain dataset. Thus any particular statistical model is typically not very general in application to different systems. The structure of such models may be useful in a wide variety of different contexts. Regression models, which assume a particular mathematical relationship between variables and assume that errors in the data take a particular form, are widely applied. Numerous regressions have been estimated for species richness as a function of latitude, altitude, and rainfall (Huston, 1994). Discussion of statistical methods applied to estimation of population sizes and densities may be found elsewhere in these volumes.

An alternative view of statistical models is that they are far more than simply descriptive, by providing a method to tease apart potential impacts of complex interactions from limited observations. This is one of the objectives of hierarchical Bayesian methods in which data on uncontrolled variables are used to estimate the interactions at different levels or scales in a system. This might include determining the relative genetic variability within and between populations, developing methods to combine data from quite different experiments, and analyzing how global change impacts plant diversity through its various effects on CO_2, nitrogen deposition, and changes in precipitation. The idea is to consider a hierarchy of processes in which the probability distribution at one level of the process depends on that of components at other levels. Using conditional probability arguments several times allows a Bayesian updating to build up a complex model from simpler conditional relationships. The posterior distribution can then be quite complex, but is possible to obtain using a Markov Chain Monte Carlo method in which the Markov chain being simulated has the same stationary distribution as the desired posterior distribution from the hierarchical Bayes model. See Clark and Gelfand (2006) for a variety of applications of this methodology to environmental problems.

Dynamic Models

Although many of the traditional models for populations and communities are in the form of dynamical systems (e.g., collections of linked differential or difference equations), often the types of analyses performed for these models are based around equilibrium assumptions. The objective is to find long-term asymptotic behavior. This may be a static equilibrium (e.g., population sizes approach a constant value through time) or a dynamic equilibrium (e.g., population sizes follow repeatable patterns through time; Murray, 2007). While these situations may arise, many models produce behavior that does not have a long-term equilibrium structure. Another key objective is to determine the stability characteristics of any equilibria that arise, in the mathematical sense of determining whether a model that is perturbed from an equilibrium condition will return to it. The dynamics arising in all these models takes account of the basic demographics of the population, as well as interactions with other populations. Adding abiotic conditions such as temperature and rainfall, or adding spatial components, often requires that the analysis be done using numerical simulations. The availability of rapid solution methods for even quite complex dynamical systems has allowed for extensive investigation of the behavior of these models when parameters are not well known. "Parameter sweeps" are particularly simple to perform in parallel using cluster computers, and provide a means to determine how a system is projected to move from one type of dynamical behavior to another as parameters change.

Geographical Information Systems (GIS)

Remote-sensing methods have opened new possibilities for following and modeling the responses of the earth's biota. A key tool that allows the use of such materials is GIS, which enables computers to graphically display the remote-sensing data as two-dimensional maps. Each image may represent one

aspect of an underlying landscape, such as landcover or vegetation type. The image value at any particular location (or pixel) in the map is estimated using models that classify the output of the cameras or the multispectral scanners on satellites into types appropriate for the objective. These models require ground-truthing to ensure that the estimated value for a particular location matches what is actually present. GIS methods allow various spatially explicit components of a landscape to be combined by looking at different map layers (different images measuring different aspects of the landscape). A mathematical function is then applied that averages or applies thresholds to these various components. Estimates of regional and worldwide carbon uptake are obtained using such methods applied to vegetation maps, in which different carbon assimilation values are assigned to different vegetation types, linked with weather maps supplying temperature and rainfall patterns. Many specialized tools are available within standard GIS systems, for example, to follow the movements of individual organisms with remote-sensing tools. Though not yet standardized, temporal GIS methods provide a means to analyze dynamic models across landscapes, and there are 3-D extensions of GIS.

Applications

Habitat Suitability Indices

Habitat Suitability Evaluation Procedures (HEP) are a formalized methodology for impact assessment on wildlife habitat. These are based on Habitat Suitability Index (HSI) models (Verner *et al.*, 1986), which attempt to summarize the site characteristics that affect the utilization of particular habitats by a variety of wildlife species. Numerous HSI models have been constructed, typically consisting of very simple regression-type models. The key habitat variables are often some measure of canopy cover in a variety of classes, diameter classes of trees and shrubs, tree stem densities, area of open water, and distance to forest cover, among others. The objective is to combine these variables, based on extensive field observations done in a correlative manner, to provide overall indices of suitability. HSIs are always indexes with values between zero and one, and they are assumed to be proportional to carrying capacity.

HSIs are based only on local habitat variables; they completely ignore any effects due to species interactions, except those due to indirect effects on related habitat variables. The models ignore the spatial interactions of habitat types across a landscape. This leads to difficulty in situations for which the size, shapes, edge effects, and neighborhood relationships have a greater effect on habitat preference than local forest composition and structure variables. The models also do not take account of the issue of presence/absence of a species, and thus ignore any historical influence on potential local abundance. HSIs are inherently static entities, so any dynamics they produce are driven completely by changes in habitat variables and not by the inherent dynamics and demography in the species being considered. Despite these criticisms, HSIs are among the most commonly used set of ecological models, in part because users realize their limitations and view them as a simplified tool to summarize a very complicated situation by a single number. Such simplification must result in a loss of information, but a key issue regarding HSIs is whether they indeed can be used as a useful predictor for abundance.

Metapopulation Models

HEP is an attempt to account for the spatial nature of biodiversity by including explicit maps of basic habitat variables. An intermediate approach between models that include all the spatial detail available from maps and those that ignore all spatial aspects of a system is metapopulation models. These models consider a landscape to be split up into a number of localized populations, called subpopulations, with the entire collection of these called a metapopulation. Most of the biotic interactions driving population dynamics occur within the localized subpopulations, but there are exchanges of individuals between these subpopulations. This allows for differing environmental, demographic, genetic, and disease situations to be present in the subpopulations. Depending on the assumptions about transport of individuals between the subpopulations, these can be relatively isolated or closely coupled.

A clear advantage of metapopulation models is the ability to derive analytical results, such as equilibrium and stability behavior, as a function of the within-subpopulation characteristics and the between-subpopulation factors such as movement. The models are particularly appropriate for cases in which a landscape can be reasonably viewed as containing discrete patches of habitat suitable for the population, with the intervening regions not suitable. The level of detail in these models can be quite variable, with the simplest versions just treating subpopulations as either present or absent. More complicated models take account of demographics within each subpopulation or explicit details on the relative spatial locations of each subpopulation that affect movement between them. These can be used for a population viability analysis, in which the probability that the overall population will survive for varying time periods is estimated.

Individual-Based Approaches

Classical ecological modeling approaches include various levels of aggregation in order to simplify the model. An alternative reductionist approach is to take account of differences between individuals within a population, allow the individuals to feed, grow, and interact, and from the aggregated behavior of these individuals build an understanding of population-level responses (Grimm and Railsback, 2005). These individual-based approaches are increasingly common; thanks to advances in computers and the increasing availability of data on behavior and physiology of species of particular interest. The advantages of these approaches include the ability to consider the effect of abiotic factors on populations through their direct impact on individual behavior and growth, to take account explicitly of spatial variation in habitat factors, and to deal with small populations in which individual differences within the population can have great impacts on population-level responses. Disadvantages of

individual-based approaches include the requirement for a great amount of detailed data to realistically simulate individual behaviors, and the typical necessity of making numerous simulations to evaluate any particular response that may arise because of the stochastic nature of the models. The methodology to describe and interpret these models is now becoming standardized and there are readily accessible general tools for constructing these models, NetLogo being one of the more accessible ones.

Future Trends

Multimodeling and Regional Assessment

Natural systems have many interacting components operating at a variety of temporal and spatial extents and requiring differing levels of detail to describe the interactions between them. One historical approach in ecology to model such systems is to break it into a number of compartments (often for different trophic levels, and sometimes with grouping within each trophic level) and consider the dynamics of each compartment with movements of energy, biomass, or nutrients among them. It is quite difficult to make these systems analysis approaches spatially explicit or to link them to GIS. Yet it is now becoming possible to link together a variety of different modeling approaches in order to best utilize the available data, with different resolutions at different trophic levels, and carry this out in a spatially explicit manner. Such multimodels may use very simple models similar to HSIs for some trophic components, more complex dynamical systems for certain populations with mostly very localized interactions, and individual-based models for organisms that move great distances and average over the spatial heterogeneity. One example of such an approach is the Across Trophic Level System Simulation (ATLSS) Project, a multimodel used to estimate the biotic impacts of alternative water management plans on the Everglades of south Florida (DeAngelis *et al.*, 1998).

Building multimodels requires extensive landscape data obtainable from remote-sensing and ground efforts. The approach is inherently dynamic, and thus requires methods to estimate the spatial dynamics of key environmental drivers, or else have available a history of this spatially that can be analyzed statistically. In the Everglades, the major driver is water, and both historical data and detailed hydrologic models are available to provide estimates for scenario evaluations. Without such data, assumptions must be made about the dynamics of the landscape. Continued enhancement of remote-sensing methods and sensor data on individual organisms is providing extensive time series of remotely sensed data to both calibrate multimodels and provide the opportunity to iteratively improve their predictive abilities. For problems at regional spatial extents, such multimodels are a rational method to aid planning while taking account of the best scientific data at the variety of resolutions available.

Behavioral Dynamics of Species of Special Concern

Great strides are being made in improving our understanding of the conservation biology of rare, threatened, and endangered species throughout the world. In an effort to better estimate the responses of populations of these species, numerous remote-sensing methods have been employed to track the movements of individual organisms. These include tracking devices implanted within or fixed on sampled individuals, which allow explicit location and physiological data (e.g., body temperature) to be obtained regularly and remotely throughout the individual's life. When done for many individuals within a population, it is becoming possible to follow the behavioral dynamics of mixtures of individuals. This includes the ability to ascertain details of mating, territoriality, and aggressive interactions. It is likely in the future that managers concerned with a particular species will be able to observe in real time the movements of many individuals within a population. Then they might apply spatially explicit modeling methods to project the response of the population to particular management alternatives and compare these with optimization models that project the "best" that is possible to do according to some criteria.

Acknowledgments

The author appreciates the long-term support of the US Geological Survey through various awards for the ATLSS project, and the support of the National Institute for Mathematical and Biological Synthesis, an Institute sponsored by the National Science Foundation (NSF), the US Department of Homeland Security, and the US Department of Agriculture through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the USGS or the NSF.

See also: Biogeographical Models. Dynamic Global Vegetation Models. Indicator Species: Computation. Landscape Ecology and Population Dynamics. Landscape Modeling. Measurement and Analysis of Biodiversity. Metapopulations. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Population Dynamics. Remote Sensing and Image Processing. Species Assemblages, Macroecology, and Global Change. Species Distribution Modeling

References

Buckley LB, Urban MC, Angilletta MJ, Crozier LG, Rissler LJ, and Sears MW (2010) Can mechanism inform species distribution models? *Ecology Letters* 13: 1041–1054.

Caswell H (2001) *Matrix Population Models*, 2nd edn. Sunderland, MA: Sinauer Associates.

Clark JS and Gelfand AE (eds.) (2006) *Hierarchical Modelling for the Environmental Sciences*. Oxford: Oxford University Press.

DeAngelis DA, Gross LJ, Huston MA, *et al.* (1998) Landscape modeling for Everglades ecosystem restoration. *Ecosystems* 1: 64–75.

Fuller MM, Gross LJ, Duke-Sylvester SM, and Palmer M (2008) Testing the robustness of management decisions to uncertainty: Everglades restoration scenarios. *Ecological Applications* 18: 711–723.

Grimm V and Railsback SF (2005) *Individual-Based Modeling and Ecology*. Princeton, NJ: Princeton University Press.

Haefner JW (2005) *Modeling Biological Systems: Principles and Applications*, 2nd edn. Berlin: Springer.
Hilborn R and Mangel M (1997) *The Ecological Detective: Confronting Models with Data*. Princeton, NJ: Princeton University Press.
Hof J and Bevers M (2002) *Spatial Optimization in Ecological Applications*. New York, NY: Columbia University Press.
Huston MA (1994) *Biological Diversity*. Cambridge, UK: Cambridge University Press.
Levins R (1968) *Evolution in Changing Environments*. Princeton, NJ: Princeton University Press.
Murray JD (2007) *Mathematical Biology: I. An Introduction*. Berlin: Springer-Verlag.
Verner J, Morrison ML, and Ralph CJ (eds.) (1986) *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates*. Madison, WI: University of Wisconsin Press.
Whitlock MC and Schluter D (2008) *The analysis of biological data*. Greenwood Village, CO: Roberts and Company Publishers.
Wolfram S (2002) *A New Kind of Science*. Champaign, IL: Wolfram Media.

Relevant Websites

www.iplantcollaborative.org
iPlant Collaborative.
www.neoninc.org
NEON.
ccl.northwestern.edu/netlogo/
Netlogo.
www.NIMBioS.org
NIMBioS.

Conservation and People

Michelle Marvier, Santa Clara University, Santa Clara, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Anthropocene Current era in which human impacts on land cover, biogeochemical cycling, water quality and availability, and other major features of Earth now rival nonanthropogenic forces.

Conservation refugee Person who, as a result of the establishment of a protected area, has been displaced without fair compensation from the areas in which she or he lived, hunted, foraged, or grazed livestock.

Ecosystem services Goods and services ranging from medicines and building materials to oxygen, clean water, and flood control that natural ecosystems deliver to people.

Integrated conservation and development project Project that seeks to simultaneously protect nature and improve human well-being.

Kuznets curve Hypothesized trend in which pollution and other environmental degradation initially rise, then level off, and eventually decrease as incomes increase and the national economy develops.

Payments for ecosystem services Financial rewards for individuals who protect habitats so that the provision of natural goods and services can continue; payments may be made by private parties that benefit from the delivery of those ecosystem goods and services or by a governmental or other agency.

Conservation is a response to the extinctions, habitat loss, and ecosystem degradation caused by human population growth and environmentally damaging or resource-extractive human activities. Thus, people are the reason why conservation, as both a scientific and a sociopolitical enterprise, is necessary. At the same time, conservation is fundamentally an expression of human values, with motivations for conservation ranging from religious or ethical to material and economic. Thus, conservation exists because of the effects of people on the nonhuman biosphere, and the scope, goals, and implementation of conservation are shaped by people. However, conservation is arguably conducted not only because of people but also for people. Conservation seeks to protect biodiversity as well as ecosystems that provide highly valuable and potentially irreplaceable products and services for both current and future generations of humans. Although conservation exists, at least in part, for the benefit of humanity, many of the actions performed in the name of conservation have harmed people and their interests. Displacements of human communities and restrictions on human activities are enacted presumably for the greater good of both nature and humanity. However, the distribution of benefits and harms among different human communities has been highly uneven, causing some observers to question the ethical basis for certain conservation practices. This article explores four dimensions of the complex relationship between people and conservation: (1) people as a conservation threat; (2) people as the creative force, shaping conservation's goals and actions; (3) people as victims of conservation; and (4) people as beneficiaries of conservation.

People as a Conservation Threat

The cumulative effect of people on the natural world is conservation's *raison d'être*, and one dimension of conservation research entails the measurement of human threats and the development of strategies to reduce the negative effects of people on nature.

Measuring and Mapping Human Threats to Conservation

The effects of people on the planet, sometimes cumulatively called the "human footprint," increased greatly over the last century, largely because of explosive human population growth combined with vastly increased per capita consumption of natural resources. The human population is now approaching 7 billion and is continuing to grow at a rapid pace. The United Nations predicts a population of 9–10 billion people at the end of the current century. The growth of the human population in developed nations has slowed, and in some wealthy nations the population is shrinking. However, the population in developing nations continues to grow rapidly.

Feeding the anticipated 9 billion plus people will require enormous increases in agricultural production. Production of crops, including those used for fiber and biofuels, is and will continue to be a major pressure for conversion of habitats to agricultural uses. Unless agricultural practices are radically improved, crop production will cause even greater land conversion, water withdrawals, environmental pollution, and loss of topsoil. In addition to agricultural products, the additional billions of people will increase demand for fish and seafood, freshwater, wood products, fossil fuels, and just about every other natural resource imaginable. As a result, many nonhuman species will face even greater loss of their habitats and increased mortality stemming from direct harvest, pollution, and altered ecological interactions due to, among other factors, species introductions.

The spatial coincidence of tropical areas of high biological diversity and future human population growth (Cincotta *et al.*, 2000) is a matter of great concern for conservationists. However, scientific assessments of the planet find negative effects of

human activities everywhere. Birds, fish, and whales in remote arctic oceans have flesh contaminated with chemical pesticides. The nitrogen cycle and hydrological cycle are now dominated by humans – human activities produce 60% of all the fixed nitrogen deposited on land each year (Kaiser, 2001), and people appropriate more than half of the annual accessible runoff (Postel *et al.*, 1996), leaving little freshwater for other species. Human modifications of habitats include deforestation, draining of wetlands, impoundment of rivers, urbanization, pollution, species introductions, climate change, and overharvest of plants, fish, and other wildlife. Scientists have coined a name for this era – the Anthropocene – to emphasize that we have entered a new geological era in which human influences on land cover, biogeochemical cycling, water quality and availability, and other major features of the world now rival or even surpass nonanthropogenic forces (Steffen *et al.*, 2007).

Various researchers have produced maps to illustrate the distribution and intensity of the human footprint. An especially well-known assessment led by scientists at the Wildlife Conservation Society defined the human footprint on the basis of population density, land transformation, accessibility, and electrical power infrastructure (Sanderson *et al.*, 2002). These efforts consistently reveal that few places remain relatively unaffected by humans. One outcome of the growing human footprint is species extinction. Although the rate of extinction has been overestimated by many scientists (He and Hubbell, 2011), it is clear that large numbers of species are at risk.

Analyses of the oceans, which might seem from a distance to be unspoiled, reveal a similar picture of widespread human impacts. In fact, a team of scientists concluded that more than 40% of the world's oceans are strongly altered by human actions, and no ocean area is unaffected by people (Halpern *et al.*, 2008). Ocean impacts – including dredged ocean bottoms, overexploited fish stocks, and pollution – are not as immediately visible as are terrestrial changes such as deforestation, but their effects on biodiversity and the delivery of ecosystem services are just as dramatic. Humans influence virtually everyplace on the planet, no matter how remote. Adding atmospheric pollutants and global climate change to the tally, it is safe to say that human activities have altered all of the planet's lands and waters in profound ways.

It is easy to attribute the negative effects of people on the natural world to ignorance and greed. However, many of the negative outcomes are simply trade-offs from activities meant to improve the human condition. People purposefully try to kill off predators such as wolves to reduce predation on their livestock, people build dams to improve navigation and irrigate crops during dry periods, and people drain wetlands to reduce mosquito-transmitted disease. These activities are attempts to tame or domesticate nature for the safety and health of people. Although the domestication of nature has pushed many nonhuman species to the brink of extinction and sometimes beyond, humans have enjoyed profound benefits including improved nutrition, health, and longevity (Kareiva *et al.*, 2007).

Given the negative effects of people on the natural environment, it is of the utmost importance, not only for conservation but also for human welfare, that society move quickly to reduce the rate of human population growth. However, the raw number of people on the planet is not the only factor determining the human impact on nature. *Per capita* consumption of natural resources also plays an important role. Heightened *per capita* demand for everything from indoor plumbing and refrigeration to meat, automobiles, and portable electronic devices, translates into more use of natural resources and therefore more pressure on ecosystems and nonhuman species. There is hope that economic development may predispose societies to take better care of the natural environment. For example, the amount of certain pollutants initially increased but later decreased as a function of economic growth. This so-called Kuznets relationship has been documented in some cases, but there is so far no evidence that it holds for conservation-related activities or outcomes (Dietz and Adger, 2003). The other big hope for reducing human impacts on the planet lies in technological developments. If crops can be produced using less water and less fertilizer, if alternative energy sources can be adopted on a large enough scale, if use of recycled materials can reduce the reliance on wood, cotton, and other natural resources, then there is yet hope for reducing the human footprint, despite the foreseeable growth and increasing prosperity of the human population.

Conservation Strategies to Reduce Human Threats

Beyond measuring and mapping human impacts, conservationists design strategies to reduce the negative effects of humans on ecosystem function and biological diversity. A major conservation strategy has been the establishment of protected areas. The International Union for Conservation of Nature (IUCN) describes a protected area as a "clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values" (Dudley, 2008). Conservationists typically seek to identify and protect places that harbor long lists of species, or species of particular concern, or habitats that are relatively undisturbed by human activities. The degree to which human activities are curtailed within protected areas varies from strict exclusion of people to community-managed areas that incorporate traditional levels and modes of resource use.

Conservationists initially set a goal of protecting 10% of the world's land area. That goal has since been surpassed, and approximately 13% of the world's land area is now under some level of conservation protection (Jenkins and Joppa, 2009). Establishment of marine protected areas (MPAs) has lagged behind that of terrestrial protected areas. Conservationists now seek to protect 10% of the planet's marine area, or at least 10% of marine areas within 200 nautical miles of land. These goals for global protection were set in the absence of reliable data on the number and extent of MPAs. Thanks to a concerted data-collection effort, researchers documented the existence by 2006 of nearly 4500 MPAs covering only 0.65% of the world's oceans (Wood *et al.*, 2008). This rather paltry coverage means that conservationists have a long way to go to meet even the most modest goal of 10% coverage of areas within 200 nautical miles of land.

In addition to setting aside particular areas for protection, conservationists often seek legal limits on human activities that might harm species at risk of extinction or particularly important habitats, such as wetlands. The US Endangered Species Act, Canada's Species at Risk Act, and Australia's Environment Protection and Biodiversity Conservation Act are but a few examples of national laws that restrict human activities in the name of conservation. The degree to which activities are limited varies from country to country, but many nations have adopted the approach pioneered in the USA of compiling a list of at-risk species that on listing gain additional protections.

In the USA, the presence of a listed animal species on privately owned land can greatly curtail the activities of landowners. On federally owned lands, the presence of a listed animal or plant can precipitate such restrictions. In addition to prohibiting direct harm to individuals of the listed species, the law may prohibit economic activities such as farming, grazing, harvesting trees, and the construction of roads and buildings, if these activities will reduce the amount or quality of habitat for the species. Even very large, economically important activities, such as the construction of a dam, may be halted. Such was the case when construction of the Tellico Dam on the Little Tennessee River was stopped because the proposed dam would submerge and destroy the habitat of a small endangered fish known as the snail darter (*Percina tanasi*). Litigation of this case went to the US Supreme Court (*Tennessee Valley Authority* v. *Hill* 1978), and the Court ruled that indeed the dam could not be completed because to do so would extirpate the snail darter. Many people were astonished by this decision because $110 million had already been spent on the project (Mann and Plummer, 1995). The Court's decision made the snail darter a symbol in the highly polarized pro-development versus conservation controversy. The prodevelopment sector ridiculed the snail darter as a tiny (less than 8.5 cm in length), useless fish that should never trump the jobs and economic benefits associated with a dam. The conservationists argued that the dam was not needed and would not benefit people as claimed.

Conservationists have implemented a broad range of strategies to protect biodiversity and ecosystem services. The two examples discussed here – protected areas and legal restrictions on human activities that could harm at-risk species – illustrate that the primary focus of conservation is eliminating or at least mitigating the negative effects of human activities. Because such restrictions can have economic consequences, conservation actions can be controversial.

People Define Conservation Goals and Priorities

Broadly speaking, conservationists seek to protect nature. But what exactly is "nature," and what places, species, or other entities are considered worthy of conservation attention and investment? Debate about these issues makes clear that the concept of nature is a human construction, shaped and designed for human ends. For example, portrayals of nature as either a fearsome, dark power or as fragile and feminine vary across cultures and are subject to change over time. Similarly, the targets of conservation attention and the goals by which conservationists measure their success are human constructions, reflecting societal values and whims. Three issues – shifting conservation baselines, the wilderness ideal, and the disproportionate attention given to large-bodied vertebrate species – illustrate some of the ways in which conservation priorities reflect human values.

Shifting Baselines

The phenomenon of shifting baselines illustrates just how malleable conservation goals can be (Pauly, 1995). Shifting baselines refers to how people's perception of the "normal" state for an ecosystem changes across generations. For example, as one generation of fishers replaces the previous one, any sense of how abundant or large the fish once were is lost (Sáenz-Arroyo *et al.*, 2005). Or consider that until recently there was little recognition that Yellowstone National Park once harbored large populations of jackrabbits (*Lepus townsendii*; Berger, 2008). The past presence of these rabbits in Yellowstone, along with their recent decline, is much more than a historical curiosity; it could help explain the sometimes intense conflict between today's wolf population and domestic livestock. Wolves might not have such a large impact on livestock if an ample number of jackrabbits were still available as prey.

Shifting baselines are potentially problematic because, once the memory of past conditions is lost, it becomes all too easy to accept today's depleted conditions as normal. And since everyone, including conservation managers, is susceptible to this phenomenon, the state of the natural world is ratcheted ever downward. Moreover, with so little knowledge of what has been lost, there is less urgency and motivation for sound management of today's resources.

Attitudes toward introduced species are also susceptible to the phenomenon of shifting baselines phenomenon. Conservationists generally seek to prevent species introductions, and they often try to control or eliminate certain harmful introduced species because these invaders can cause extinction of native species and disruption of ecosystem services. However, the desire to control or eliminate invasive species is not universal. The picture can become murky when the invader provides valuable ecosystem services. For example, plantations of nonnative trees can sometimes be more productive, retain nutrients better, and even have greater total diversity (albeit nonnative diversity) than native forests.

The Nile perch (*Lates niloticus*) provides an example where groups of people have very different attitudes about a highly invasive species. Most scientists view the introduction of the Nile perch to Lake Victoria as an ecological nightmare. This fish, intentionally introduced in 1954 to provide food and income for local communities, caused the extinction of possibly hundreds of fish species native to the lake. However, the Nile perch did indeed provide boom fishery harvests and was at least temporarily of great benefit to the local people, particularly because it increased employment. The perch itself is now overfished, and its economic benefits have declined. Concordant with fishing-induced declines in the perch, many native species, including some that were thought to have been lost from the lake, are enjoying a resurgence (Balirwa *et al.*, 2003). The ultimate fates of the Nile perch and of native fish

diversity in Lake Victoria are still uncertain, but local support for eradication or control programs aimed at Nile perch is minimal and unlikely to gain traction. To the contrary, the perch fishery is viewed as needing better management. The most likely scenario for Lake Victoria is that a new "ragamuffin" (Marris, 2009) ecological community will develop, with some extinctions of native species (but not as many as originally thought), and with the Nile perch and other species representing a commercially important fishery. In other words, the perch will become integrated into the lake's ecosystem, and what is "natural" will be redefined to include it.

One does not have to look hard to find invasive species that many people are fond of and would hate to see removed. Human values play an important role in determining which species introductions are considered calamitous and which are considered indispensable. In fact, intentional introductions may become an important tool in conservation, potentially used to move species threatened by climate change to areas with more favorable conditions and to replace the ecosystem functions previously performed by now extinct species (Marris, 2011). Clearly, conservation goals are a reflection of both scientific and societal understandings of what conditions are normal or desirable in nature.

The Wilderness Ideal

In western conservation, "pristine" wilderness, relatively untouched by humans, is given high priority for protection. The US Wilderness Act defines wilderness as "an area where the earth and community of life are untrammeled by man, where man himself is a visitor who does not remain." This definition reflects the view that humans and wild nature are incompatible. Whether or not one agrees with this point largely depends on whether one views people as a part of nature or an unnatural force, external to nature (Adams and Hutton, 2007; Ereshefsky, 2007). Surveys in the USA indicate that people overwhelmingly see themselves as part of nature, yet the majority paradoxically describe a natural environment as one that is free from human presence or influences (Vining *et al.*, 2008).

The human-nature dichotomy has profoundly influenced conservation thinking and practice. The creation of at least some protected areas via the displacement and exclusion of human communities – what some people term "fortress conservation" (Wilshusen *et al.*, 2002) – is one important manifestation of this worldview. However, given the extensive human effects on the planet and that protecting wilderness may first require moving people out of the area, some question whether wilderness truly exists or is merely a mental construct. Even the supposedly "virgin" rainforests have in fact been subjected to a long history of slash and burn agriculture and were dotted with surprisingly large human settlements (Willis *et al.*, 2004).

Preferential Treatment for Large-Bodied Vertebrate Species

The role of human values in shaping conservation priorities also manifests itself in the set of species that have become the icons of conservation. Entire organizations are devoted to the conservation of certain animal species such as cheetahs, elephants, and gorillas. Other species may not have their own organization looking out for their interests, but they receive a disproportionate amount of conservation funding. Species such as the giant pandas, rhinos, whales, and lions that adorn the magazines and calendars of conservation organizations are not representative of biological diversity. In fact, the majority of species are various kinds of insects. In contrast, the species that have been elevated to the status of conservation icons tend to be large-bodied vertebrate animals.

The concept of biodiversity is in some ways an attempt to level the playing field and give more or less equal value to all types of species. According to this view, if the goal is simply to protect the longest lists of species, a nematode or beetle species would count as equivalent to a bird or mammal species. But this impartial treatment of species is not reflected in actual conservation practice. Invertebrate animals comprise 84% of all animal and plant species in the USA and more than one-third of all at-risk US species (Goble *et al.*, 2006), but they account for only 14% of the species listed under the US Endangered Species Act. Funding for conservation of at-risk species also disproportionately favors vertebrates, especially iconic species. For example, in the 2004 fiscal year, 20 listed vertebrate species received more than 50% of US federal and state conservation expenditures that could be attributed to particular species (USFWS, 2006). Moreover, although plants make up about half of the federally listed species, they receive less than 10% of the funds for recovery (Schemske *et al.*, 1994).

Conservation goals reflect human ideas of what nature "should" look like. These goals change over time as a function of generational amnesia about past conditions and changing societal preferences for different types of landscapes and species.

People as Victims of Conservation

Conservation can sometimes seem like a movement that is antipeople, entranced by sentimental longings for pristine wilderness untrammeled by humans and unmarred by cities. Conservation's misanthropic reputation has arisen, in part, because millions of people have been forced off their land so that animals and habitats could be preserved (Dowie, 2009). Hunters and farmers in some Asian and African nations contend that protected areas limit their diet and earnings. In the USA, farmers, fishermen, and loggers are angry about losing their water privileges or their jobs because of fishing restrictions and other endangered species protections. These problems are discussed in the following sections in the context of conservation refugees and the negative economic effects of conservation.

Protected Areas and Conservation Refugees

In the process of creating protected areas, conservation has harmed people by displacing communities, stripping them of their access to wildlife and other natural resources, and potentially aggravating poverty and hunger by limiting opportunities for agricultural and other economic development

(Adams and Hutton, 2007; Dowie, 2009). The frequency with which the establishment of protected areas has displaced people and the exact number of people affected have not been well quantified (Brockington *et al.*, 2006), but at a minimum the history of such events traces back to the forced relocation of Native Americans during the 1872 establishment of Yellowstone, the world's first national park.

One example illustrating the conflict between protected areas and people concerns the Maasai. The Maasai are pastoralists who traditionally migrate with their cattle herds across the savannahs of eastern Africa. In 1947, the Amboseli ecosystem in southern Kenya and eastern Tanzania was declared a National Reserve and grazing was restricted. The Maasai who had grazed their cattle in Amboseli for many generations were angered. The 1971 declaration of Amboseli National Park was accompanied by additional restrictions including a ban on hunting. The Maasai were promised compensation for their losses, but they received little and began pressuring the government to return the lands (Dowie, 2009). In 2005, Kenyan president Mwai Kibaki returned Amboseli National Park to its original Maasai inhabitants. This controversial decision reflects a growing global discontent with such preemptory displacements.

Forced eviction of local people is not just a social injustice; it can also accentuate human threats to conservation if the local people fight back. When the Amboseli Maasai did not receive compensation they had been promised, they fought back by killing wildlife and leaving the carcasses of rhinos and leopards near popular camping spots (Dowie, 2009). In another example, the 1982 creation of Uganda's Mburo National Park was followed in 1983 by the violent and uncompensated eviction of approximately 4500 families from the area. In 1986, under a new government, the dislocated people were encouraged to resettle their former homelands. On their return, they slaughtered the park's wildlife and vandalized its facilities with the goal of pressuring the government to degazette the entire park. The government did return approximately 60% of the original park area to the people. Even after that concession, conservationists realized that the entire park would fail without the support of surrounding communities. Thus conservationists made concerted and relatively successful efforts to improve relations with the local people via regular meetings, education programs, and sponsorship of community initiatives such as the construction of schools and clinics (Turyaho and Infield, 1996).

Critics have pointed out that it is poor, often indigenous, and always politically weak groups that are uprooted from their homes. In contrast, the people who reap the benefits of protected areas and get to enjoy their scenic beauty and wildlife often are tourists, wealthier and typically whiter than the displaced people. Although modern conservationists do not overtly endorse a racist agenda, early conservation heroes John Muir, Teddy Roosevelt, and others openly expressed their disdain for indigenous groups. This is not just a historical matter – conflicts between conservation and the rights of indigenous and other local communities are ongoing. Some, and perhaps many, of conservation's successes in terms of land protection entail a dark side of racism and oppression.

Conservationists are increasingly acknowledging these injustices and are working to improve the social equity of their strategies and actions. For example, a group of major international conservation organizations recently formed a Conservation Initiative on Human Rights. This consortium is working to develop and implement a set of principles concerning human rights in the context of conservation.

Negative Economic Effects and Unemployment

Conservation strategies, including the establishment of protected areas and legal protections for at-risk species, seek to prohibit or limit certain human activities. Resource-extractive activities such as logging, fishing, mining, and withdrawal of freshwater are popular targets for conservation restrictions. Other activities that alter habitats, such as agriculture, grazing, and construction of roads and buildings, also may be limited. These limitations can come at the expense of food production, economic gain, and jobs.

Conservation laws in some nations, such as Canada's Species at Risk Act allow regulators to weigh the economic consequences of conservation restrictions. Others, such as the US Endangered Species Act call for decisions based solely on the best science available with no consideration of economic factors. As described earlier, the use of the US Endangered Species Act to halt construction of the Tellico Dam is one example of how conservation regulations may upend development plans, regardless of the economic effects.

Other examples of conflicts between conservation and jobs abound. In the 1990s, the US federal government halted timber harvesting on publicly owned land in Oregon and Washington because logging jeopardized the continued existence of the northern spotted owl (*Strix occidentalis caurina*). As a result, sawmills were shut down, and rural towns across the Pacific Northwest suffered severe economic hardship. Out-of-work loggers expressed their displeasure by displaying bumper stickers proclaiming, "Save a logger, eat an owl!" Another flashpoint for conservation involves restrictions on the amount of water allocated to farmers. In these cases, there tend to be powerful economic lobbies on both sides of the debate – farming on one side and salmon fisheries on the other. But, at least in the USA, the decisions about how to allocate water between farmers and endangered fish species are supposed to remain blind to the economic consequences.

Some conservationists believe that economic growth is in direct conflict with conservation because increased resource consumption directly translates into increased pressure on wildlife (Czech, 2000). However, others see economic development as a necessary prerequisite to achieving conservation goals. Regardless of one's views on economic development, it is undeniable that the negative economic consequences of conservation actions, loss of jobs, and limitations on the rights of private landowners have evoked significant public backlash against conservation. In the USA, there have been numerous attempts to repeal or weaken the Endangered Species Act, an Act that was unanimously approved by the US Senate in 1973. As a result of public backlash, the US conservation policy has evolved to rely more heavily on conservation easements, habitat conservation plans, and other negotiated agreements that provide economic incentives for landowners who pursue conservation-friendly practices.

Rather than hurting people's pocketbooks, these policies financially reward landowners through direct payments, tax breaks, or freedom from future additional restrictions if they adhere to certain agreed on practices or restrictions.

People as Beneficiaries of Conservation

Conservation of nature is motivated not just by ethical arguments regarding human-caused extinction; it is also conducted for the explicit benefit of humans. Many people worry that their children and grandchildren might not be able to enjoy large open spaces and see species ranging from large mammals and birds to orchids and butterflies. Conservation has succeeded in protecting vast expanses of land and water from deforestation, overexploitation, and development (Nagendra, 2008). Across the globe, national parks and other protected areas are hugely popular tourist destinations. And despite their sometimes controversial beginnings, these parks often become a source of national pride, clearly beloved by a majority of citizens. Conservation has also aided the recovery of species such as bald eagles, peregrine falcons, and gray whales. Thanks to a ban on dichlorodiphenyltrichloroethane (DDT) (a pesticide particularly harmful to predatory bird species), restrictions on the harvest of whales, and other such limitations on human activities, future generations of people will get to witness these truly awe-inspiring species.

Beyond the aesthetic value of nature, casual observers do not always see links between conservation and human well-being. But such connections abound. Grasslands and coral reefs provide food and income, forests store carbon and regulate water supply, wetlands and mangrove stands protect people from lethal storms. For example, field studies conducted after the Indian Ocean tsunami of 2004 found that shorelines with intact mangrove forests suffered almost no devastation because the trees dissipated the energy of the waves, whereas neighboring shorelines with destroyed or degraded mangroves suffered tremendous human and economic losses (Dahdouh-Guebas *et al.*, 2005). Based on a retrospective analysis of deaths and damages caused by a 1999 super cyclone in eastern India, researchers estimated that every 68 ha of remaining mangrove forest saved one human life (Das and Vincent, 2009).

Moreover, when the economic costs and benefits of conservation are properly assessed, the positives often outweigh the negatives. For example, unsustainable logging practices in the Togean Islands of Indonesia produce silt and sediment that damage coral reefs. The income lost because of the resulting declines in tourism and fishing far exceeds that generated by the logging. In such a case, adopting more sustainable logging practices or even protecting the forest from all logging would be a net win for society. But this is not obvious without a full accounting of the gains and losses caused by logging versus habitat protection (Pabon-Zamora, 2009).

Protected Areas Can Benefit People

Protected areas can be compatible with and even benefit local people. Conservationists have become increasingly interested in the idea of community-based conservation, which can mean anything from sharing the financial benefits of protected areas with local communities to actively relying on local people to monitor and manage an area's natural resources. This interest is motivated in part by the realization that it is difficult and expensive to guard and manage large areas that are devoid of people. Such places can become targets for poachers and others looking to make money by clearing trees or extracting other natural resources. Parks devoid of people can also become havens for paramilitary groups and illegal activities such as cultivation of drug-producing plants (West and Brockington, 2006). In light of these challenges, conservationists began to realize that it is better to have the local people present and monitoring the situation (Hutton *et al.*, 2005).

Community-based conservation has had some success. In Bolivia, for instance, all of the conservation areas are inhabited by indigenous people and peasant communities, and the success of the nation's protected-area system is due largely to the fact that these people are custodians of these conservation areas (Eguino and Pabon-Zamora, 2009). Indigenous people are often more open to the creation of protected areas on their lands, as long as the areas can continue to provide livelihoods for local communities. At the same time, conservationists are struggling to move beyond merely involving local communities as participants in conservation debates to truly empowering local communities to make and implement management decisions (Brown, 2003).

Another positive is that, on average, protected areas seem to reduce poverty in the surrounding communities. Besides creating a relatively small number of jobs as conservation managers or park rangers, protected areas can attract large numbers of tourists who spend money on lodging, food, guided tours, rentals of vehicles and recreation equipment, and gifts to bring home. Measuring the social impacts of protected areas is difficult because of confounding factors and the fact that protected areas are often associated with centers of poverty that existed before the areas were established. Using data from Costa Rica and Thailand and quasi-experimental matching methods, Andam *et al.* (2010) asked what the socioeconomic outcomes near protected areas would have been if the protected areas had not been established. After controlling for preprotection characteristics associated with poverty and development potential, they found that poverty declined as a result of nature protection. This does not mean that protected areas never produce undesirable impacts on local populations. But at least there is clear evidence that protected areas can increase the well-being of local communities.

Payments for Ecosystem Services

Beyond intrinsic and aesthetic value, many conservationists argue that conservation of nature can improve human health, nutrition, and material prosperity (Daily, 1997; Kareiva and Marvier, 2007). For example, both human and nonhuman species benefit from clean air and water, so conservation efforts to protect environmental health directly benefit people as well. In addition, nature provides highly valuable ecosystem services including timber, fish, protection from storms and floods, and

sequestration of climate-warming carbon dioxide (Millennium Ecosystem Assessment, 2005).

In many cases the people who benefit from the delivery of ecosystem services live far away from the ecosystems that provide them. For example, a large coastal city may enjoy the freshwater and flood control provided by an upstream forest hundreds of kilometers away. Yet the city dwellers may not recognize the benefits that the forest provides, particularly because they seldom pay for them. Given this, the owner of the forest has little economic incentive to protect the ecosystem services provided by the forest, but he or she may have a big economic incentive to clear the land and sell the timber. To remedy this situation, payments for ecosystem services (PES) have been implemented in hundreds of places around the world (Jack *et al.*, 2008). Through PES programs, governments or other parties financially reward landowners for conserving and restoring the flow of ecosystem services.

Costa Rica is home to the longest-standing PES program. Established in 1995, this program has made payments to more than 4000 farmers for reforestation, forest conservation, and sustainable forest management. The Costa Rican PES program recognizes four services: carbon sequestration, watershed protection, biodiversity protection, and landscape beauty for recreation and tourism (Russo and Candela, 2006). Payments are larger for reforestation than for forest conservation. After 5 years of payments, landowners can renegotiate prices, but they are committed to managing or protecting their contracted forests for 15–20 years, depending on whether their activities are primarily conservation or reforestation. The payments are funded through international donations and a nationwide tax on fuel. This PES scheme is widely thought to be one of the most effective in operation, but it is not without problems. For one thing, the prices paid to landowners are not based on any quantification of what the lands actually produce and do not account for different values based on where the land is situated in a watershed.

The largest PES program is China's Natural Forest Protection Program. Between 1998 and 2005, this program distributed more than $8 billion to peasants and workers to cease logging and to reforest previously logged lands (Liu *et al.*, 2008). The ecological outcomes were impressive – the annual rate of logging declined by more than 40%, the amount of carbon sequestered was 44.1 Tg (one Tg is a million metric tons), and the area suffering soil erosion was reduced by 6%. The carbon sequestered as a result of this program is equivalent to removing approximately 17.5 million cars from US roads. China has also instituted a Grain to Green Program, which compensates farmers for converting cropland on steep slopes back to forests in the Yangtze and Yellow River basins. The farmers are given grain worth $300–$400 per year per hectare of farmland converted back to trees. The main aim of the program is restoration of ecosystem function, especially prevention of flood disasters; secondary aims are poverty alleviation and economic development. This program has directly benefited farmers in more than 30 million households, as assessed by household surveys. Many of the farmers have shifted to the construction, transportation, and restaurant businesses.

Both the Costa Rican and the Chinese examples are considered successes, although program evaluation has been modest at best. In fact, no PES program to date has convincingly demonstrated that it caused the delivery of additional environmental services relative to other possible conservation approaches (Ferraro, 2011). One of the challenges for PES stems from the voluntary nature of these programs. Given a choice, landowners are most likely to volunteer lands that have little potential for other economic activities. Thus the lands that end up being protected by PES might not have been converted to other uses even in the absence of PES schemes. Although these pioneering schemes represent an important new direction in conservation – one relying on financial incentives rather than command-and-control regulations – it is important that future programs be designed to yield measures of their effectiveness (Ferraro, 2011).

Conservation and Poverty Reduction

The relationship between conservation and human poverty is complicated, and debate about this relationship has evolved over time (see Roe, 2008). Adams *et al.* (2004) provide a helpful guide to the range of views that exists within the conservation community. Some view poverty as an issue completely separate from conservation. These individuals feel that conservation organizations must stick to their primary mission of protecting nature. Poverty, food security, and related concerns should fall strictly within the realm of development and humanitarian organizations. Other conservationists see poverty as a constraint on conservation success. After all, a poor person will be more concerned with immediate needs such as food and shelter than with loftier goals of long-term sustainability and biodiversity. Thus, these people see improving economic security as a necessary ingredient for conservation success. A third camp holds the view that conservation should do no harm, or in other words should be careful to not exacerbate poverty as protected areas are established. Finally, a fourth group believes that poverty reduction depends on the conservation of living resources. This view is consistent with a 2005 report to the United Nations concerning strategies to meet the Millennium Development Goals (an ambitious international agreement to, among other things, halve the proportion of people living in extreme poverty by 2015). This report argues that ecology will be the key to relieving poverty for the world's 750 million rural poor (Giles, 2005).

Conservation goals and goals for improving human well-being do often align. For example, deforestation in Tanzania has increased water runoff and caused water shortages such that more than half the population lacks access to safe drinking water and agricultural productivity has been undermined (Madulu, 2002). Proper management of the water-provisioning services of Tanzanian forests could benefit both people and biodiversity. More generally, the world's poorest people and populations, with the lowest food security, stand to benefit greatly from the protection of nature. Although the economies of wealthier nations are not as directly tied to the goods produced by nature, developing nations derive substantial income from naturally occurring building materials, fuel, and food. Forest products and fisheries, looked at as components of national economies, are 5–10 times more

important for developing nations than for European nations or the USA.

Integrated conservation and development projects (ICDPs) are an attempt to simultaneously protect nature and meet the basic needs of humans (McShane and Wells, 2004; Kaimowitz and Sheil, 2007). The idea behind ICDPs is that because extreme poverty is linked to environmental degradation, improving the lives of local people will improve the success of conservation efforts. The approaches used in ICDPs have focused on ecotourism, game hunting, and nontimber forest products to generate income for local people living in or near protected areas (McShane and Wells, 2004). Reviews of ICDPs suggest that many early efforts were unsuccessful (Adams *et al.*, 2004; Agrawal and Redford, 2006). The reasons for the failures include everything from poor project conception (such as unrealistic projections of markets for nontimber products) to civil strife and corruption to poor project management and overly simplistic understandings of both biodiversity and poverty. One example of a successful ICDP is the creation of the Il'Ngwesi Eco Lodge. Cattle rustling, wildlife poaching, and government pressure to subdivide and develop land threaten both the Maasai culture and the wildlife of Kenya (Christ, 2002). Subdivision would have fragmented some of the largest remaining tracts of habitat in Kenya and ended traditional pastoral life for the Maasai. In 1995, the Maasai community at Il'Ngwesi established an eco-lodge and began promoting tourism to the area. Since then, security in the region has improved, income from the eco-lodge is funding the local school, poaching has been halted, thirteen of the region's nineteen large mammals – including the endangered Grevy's zebra (*Equus grevyi*) – have returned to the area, and some wildlife populations have increased by as much as 500% (Christ, 2002). Ecotourism projects do not always turn out so well, but this example demonstrates how conservation and human interests can sometimes be aligned.

Conclusion

The relationship between conservation and people is complex and dynamic. The negative consequences of an enormous, rapidly growing, and increasingly prosperous human population are why conservation is needed. In response, conservation is a grand intervention in both the natural world and human society, the goals of which largely reflect human likes and dislikes. Conservationists may seek to modify or prohibit certain economic activities, and conservation has at times caused relocations of human communities, disruption of traditional resource management practices, and job loss. Although conservation has earned a reputation in some circles of valuing nature above people, conservation also provides profound benefits to people by protecting open space, magnificent wildlife, and ecosystems that provide invaluable services for both current and future generations. Conservation attempts to improve environmental conditions for the benefit of both human and nonhuman species, and it seeks to increase the sustainability of resource use so that the prosperity and health enjoyed by today's human population may continue into the future. In light of criticisms regarding the uneven distribution of harms and benefits across human groups, conservationists are becoming more sensitive to the rights of human communities and are developing new approaches that aim to better advance the well-being of both people and nature.

Excerpts from Kareiva and Marvier (2011) have been reprinted or adapted here with the permission of the publisher.

Appendix

List of Courses

1. Conservation Biology
2. Environmental Ethics
3. Environmental Policy

See also: Conservation and the World's Poorest of the Poor. Endangered Marine Invertebrates. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Indigenous Peoples and Biodiversity. Justice, Equity and Biodiversity. Natural Reserves and Preserves. Poverty and Biodiversity

References

Adams WM, Aveling R, Brockington D, *et al.* (2004) Biodiversity conservation and the eradication of poverty. *Science* 306: 1146–1149.

Adams WM and Hutton J (2007) People, parks and poverty: Political ecology and biodiversity conservation. *Conservation and Society* 5: 147–183.

Agrawal A and Redford K (2006) *Poverty, Development, and Biodiversity Conservation: Shooting in the Dark?* New York, NY: Wildlife Conservation Society.

Andam KS, Ferraro PJ, Sims KRE, Healy A, and Holland M (2010) Protected areas reduced poverty in Costa Rica and Thailand. *Proceedings of the National Academy of Sciences of the United States of America* 107: 9996–10001.

Balirwa JS, Chapman CA, Chapman LJ, *et al.* (2003) Biodiversity and fishery sustainability in the Lake Victoria basin: An unexpected marriage? *Bioscience* 53: 703–715.

Berger J (2008) Undetected species losses, food webs, and ecological baselines: A cautionary tale from the Greater Yellowstone Ecosystem, USA. *Oryx* 42: 139–142.

Brockington D, Igoe J, and Schmidt-Soltau K (2006) Conservation, human rights, and poverty reduction. *Conservation Biology* 20: 250–252.

Brown K (2003) Three challenges for a real people-centred conservation. *Global Ecology and Biogeography* 12: 89–92.

Christ C (2002) Il'Ngwesi, Kenya. Grist, March 20, 2002.

Cincotta RP, Wisnewski J, and Engelman R (2000) Human populations in the biodiversity hotspots. *Nature* 404: 990–992.

Czech B (2000) Economic growth as the limiting factor for wildlife conservation. *Wildlife Society Bulletin* 28: 4–15.

Dahdouh-Guebas F, Jayatissa LP, Nitto DD, Bosire JO, Seen DL, and Koedam N (2005) How effective were mangroves as a defence against the recent tsunami? *Current Biology* 15: 443–447.

Daily GC (ed.) (1997) *Nature's Services: Societal Dependence on Natural Ecosystems.* Washington, DC: Island Press.

Das S and Vincent JR (2009) Mangroves protected villages and reduced death toll during Indian super cyclone. *Proceedings of the National Academy of Sciences of the United States of America* 103: 7357–7360.

Dietz S and Adger WN (2003) Economic growth, biodiversity loss and conservation effort. *Journal of Environmental Management* 68: 23–35.

Dowie M (2009) *Conservation Refugees: The Hundred-Year Conflict Between Global Conservation and Native Peoples.* Cambridge, MA: MIT Press.

Dudley N (ed.) (2008) *Guidelines for Applying Protected Area Management Categories.* Gland, Switzerland: IUCN.

Eguino S and Pabon-Zamora L (2009) *Valuing Nature: Why Bolivia's Protected Areas Matter for Economic and Human Well-Being*. Arlington, VA: The Nature Conservancy.
Ereshefsky M (2007) Where the wild things are: Environmental preservation and human nature. *Biology & Philosophy* 22: 57–72.
Ferraro PJ (2011) The future of payments for environmental services. *Conservation Biology* 25: 1134–1138.
Giles J (2005) Ecology is key to effective aid, UN told. *Nature* 437: 180.
Goble DD, Scott JM, and Davis FW (eds.) (2006) *The Endangered Species Act at Thirty: Renewing the Conservation Promise*, Vol. 1. Washington, DC: Island Press.
Halpern BS, Walbridge S, Selkoe KA, *et al*. (2008) A global map of human impact on marine ecosystems. *Science* 319: 948–952.
He F and Hubbell SP (2011) Species–area relationships always overestimate extinction rates from habitat loss. *Nature* 473: 368–371.
Hutton J, Adams WM, and Murombedzi JC (2005) Back to the barriers? Changing narratives in biodiversity conservation. *Forum for Developmental Studies* 2: 341–370.
Jack BK, Kousky C, and Sims KRE (2008) Designing payments for ecosystem services: Lessons from previous experience with incentive-based mechanisms. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9465–9470.
Jenkins CN and Joppa L (2009) Expansion of the global terrestrial protected area system. *Biological Conservation* 142: 2166–2174.
Kaimowitz D and Sheil D (2007) Conserving what and for whom? Why conservation should help meet basic human needs in the tropics. *Biotropica* 39: 567–574.
Kaiser J (2001) The other global pollutant: Nitrogen proves tough to curb. *Science* 294: 1268–1269.
Kareiva P and Marvier M (2007) Conservation for the people. *Scientific American* 297: 50–57.
Kareiva P and Marvier M (2011) *Conservation Science: Balancing the Needs of People and Nature*. Greenwood Village, CO: Roberts and Company.
Kareiva P, Watts S, McDonald R, and Boucher T (2007) Domesticated nature: Shaping landscapes and ecosystems for human welfare. *Science* 316: 1866–1869.
Liu J, Li S, Ouyang Z, Tam C, and Chen X (2008) Ecological and socioeconomic effects of China's policies for ecosystem services. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9477–9482.
Madulu NF (2002) *Poverty, conflicts and water resource use in Tanzania*. Paper presented at the Third WaterNet/Warfsa Symposium on Water Demand Management for Sustainable Development. Dar es Salaam.
Mann C and Plummer M (1995) *Noah's Choice: The Future of Endangered Species*. New York, NY: Knopf.
Marris E (2009) Ragamuffin Earth. *Nature* 460: 450–453.
Marris E (2011) *Rambunctious Garden: Saving Nature in a Post-Wild World*. New York, NY: Bloomsbury.
McShane TO and Wells MP (2004) *Getting Biodiversity Projects to Work: Towards More Effective Conservation and Development*. New York, NY: Columbia University Press.
Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press.
Nagendra H (2008) Do parks work? Impact of protected areas on land cover clearing. *Ambio* 37: 330–337.
Pabon-Zamora L (2009) *Valuing Nature: Why Indonesia's Marine Protected Areas Matter for Economic and Human Wellbeing*. Arlington, VA: The Nature Conservancy.
Pauly D (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology & Evolution* 10: 430.
Postel SL, Daily GC, and Ehrlich PR (1996) Human appropriation of renewable fresh water. *Science* 271: 785–788.
Roe D (2008) The origins and evolution of the conservation–poverty debate: A review of key literature, events and policy processes. *Oryx* 42: 491–503.
Russo RO and Candela G (2006) Payment of environmental services in Costa Rica: Evaluating impact and possibilities. *Tierra Tropical* 2: 1–13.
Sáenz-Arroyo A, Roberts CM, Torre J, Cariño-Olvera M, and Enríquez-Andrade RR (2005) Rapidly shifting environmental baselines among fishers of the Gulf of California. *Proceedings of the Royal Society Biological Sciences Series B* 272: 1957–1962.
Sanderson EW, Jaiteh M, Levy MA, Redford KH, Wennebo AV, and Gillian W (2002) The human footprint and the last of the wild. *Bioscience* 52: 891–904.
Schemske DW, Husband BC, Ruckelshaus MH, Goodwillie C, Parker IM, and Bishop JG (1994) Evaluating approaches to the conservation of rare and endangered plants. *Ecology* 75: 584–606.
Steffen W, Crutzen PJ, and McNeill JR (2007) The Anthropocene: Are humans now overwhelming the great forces of nature? *Ambio* 36: 614–621.
Turyaho M and Infield M (1996) *Pastoralists, Fishermen and Farmers in and Around Lake Mburo National Park: Changing Conflict into Awareness and Responsibility*. Nairobi, Kenya: African Wildlife Foundation. Community Conservation Discussion Paper No. 6.
USFWS (2006) *Federal and State Endangered and Threatened Species Expenditures: Fiscal Year 2004*. Washington, DC: US Fish and Wildlife Service.
Vining J, Merrick MS, and Price EA (2008) The distinction between humans and nature: Human perceptions of connectedness to nature and elements of the natural and unnatural. *Human Ecology Review* 15: 1–11.
West P and Brockington D (2006) An anthropological perspective on some unexpected consequences of protected areas. *Conservation Biology* 20: 609–616.
Willis KJ, Gillson L, and Brncic TM (2004) How "virgin" is virgin rainforest? *Science* 304: 402–403.
Wilshusen PR, Brechin SR, Fortwangler CL, and West PC (2002) Reinventing a square wheel: Critique of a resurgent "protection paradigm" in international biodiversity conservation. *Society and Natural Resources* 15: 17–40.
Wood LJ, Fish L, Laughren J, and Pauly D (2008) Assessing progress towards global marine protection targets: Shortfalls in information and action. *Oryx* 42: 340–351.

Conservation and the World's Poorest of the Poor

Craig Leisher, The Nature Conservancy, Millburn, NJ, USA
M Sanjayan, The Nature Conservancy, Missoula, MT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Comanagement Resource users and resource managers share power and responsibilities to cooperatively manage a renewable natural resource, such as a fishery or forest. It may be two-way comanagement between a community and a government agency or three-way comanagement among a community, a bridging organization such as a nongovernmental organization or a university, and a government agency.

Elite capture Benefits from a conservation or a development initiative may be "captured" by local elites who have the economic or political power to determine who will benefit from an initiative. When this happens, an initiative can cause an increase in economic inequality.

Integrated Conservation and Development A project that links biodiversity conservation, usually of a protected area, with local socioeconomic development activities in which alternative incomes or livelihoods are introduced to reduce pressure on local natural resources.

Natural resource management This is the management of the resources that nature provides such as forests, fisheries and grasslands.

Poverty The pronounced deprivation in well-being in which a person lacks opportunities, empowerment, and security. The most common proxy for measuring poverty is per capital income. If a person lives on less than US$1.25 a day at purchasing power parity, it is classified as "extreme poverty." If a person lives on US$2.00 a day or less at purchasing power parity, it is classified as "moderate poverty."

History of Conservation and Poverty Links

Historically, the preservation of natural areas and prevention of anthropogenic impacts has been the primary goal of conservation and conservationists. Others, however, viewed the use of natural areas as economically important for improving human well-being. The conflict between preservation and use is exemplified in USA by John Muir and Gilford Pinchot – two early twentieth century contemporaries who shared the common goal of protecting wilderness in the US food division but had contrasting philosophies on how to do so. Pinchot, for example, believed sheep should be allowed to graze on public lands because it benefited shepherds and kept native grasses in check. Muir disagreed with this. He considered sheep a "hoofed locust" because of the damage they inflicted in the high sierras of California.

Muir as the founder of the Sierra Club and Pinchot as the first leader of the US Forest Service personified the historical dichotomy in conservation of preservation versus use. A century later, this dichotomy still defines how some view conservation, but it is waning in the face of growing evidence that preservation and use are points on a spectrum and there are many shades in between. In developing countries, for example, conservation initiatives may seek to conserve local biodiversity while also providing sustainable goods and services to local people.

One of the drivers of this more nuanced view is that the world's human population is expected to increase by 50% between 2000 and 2100, and meeting the basic needs of 10 billion people will require greater use of natural resources and better stewardship of our natural patrimony. Strict preservation will continue to be important in places, but the sustainable use of natural resources is likely to be central to conservation in the twenty-first century. Perhaps nowhere is this more evident than where poverty and the use of natural resource are entwined.

The poor live disproportionately in the tropics and in rural areas, where much of the world's biological diversity is also found (Redford *et al.*, 2008). The rural poor also frequently have a proportionally greater dependence on local natural resources than the better-off (Balmford *et al.*, 2008). For instance, the extremely poor are often the ones who rely most on nontimber forest products (Neumann and Hirsch, 2000), and the income and well-being of many poor households in forested areas depends greatly on the availability of forest products (Vedeld *et al.*, 2004).

Natural resources clearly matter for many rural poor, but does rural poverty matter to conservation of these resources? The debate on this point has been contentious (Sanderson and Redford, 2003, 2004; Roe and Elliott, 2004; Redford *et al.*, 2008; Roe, 2008).

To better understand the diversity of views and the evolution of thinking about conservation and poverty, it is helpful to examine a conceptual typology of the relations between conservation and poverty reduction. Based on a review of the conservation–poverty literature, Adams *et al.* (2004) identified four different conceptual views.

The first is a predominantly preservationist one. Proponents believe that poverty reduction is best left to governments and organizations with poverty reduction mandates because of its inherent complexity and because the number of poor people in the remote natural areas favored by conservationists is less than 1% of the global poor (Redford *et al.*, 2008). Dividing conservation and poverty reduction, however, may create a split that does not inherently exist in many rural communities. There is evidence that curtailing access to natural resources for local people can lead to extreme poverty such as when people are displaced by the establishment of a new protected area (Brockington *et al.*, 2006).

The second conceptual view began with a seminal United Nations (UN) report. In 1987, a UN commission produced a

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00320-8

report entitled Our Common Future (Brundtland, 1987) that called for "sustainable development." The report argued, *inter alia*, that poverty is a major cause of environmental problems. Building on the thinking of this report, in the early 1990s, many conservationists expected that reducing poverty would help reduce destructive resource use (e.g., Wells and Brandon, 1992). In specific cases where the poor are the drivers of resource degradation, such as fuelwood consumption in parts of India and Tanzania (Baland et al., 2006; Ndangalasi *et al.*, 2007), reducing poverty may in fact reduce pressure on some natural resources. Yet China, Indonesia, and Brazil have greatly reduced rural poverty while continuing to draw down their natural resources. There are two primary reasons why the alleviation of rural poverty is unlikely to break unsustainable resource-use patterns: consumption of many natural resources tends to increase with income (e.g., Narain *et al.*, 2008; Brashares *et al.*, 2011); and better-off people are often the ones who benefit the most from natural resources use because they have the capital to exploit the resources (Cavendish, 2000; Narain *et al.*, 2008).

The third conceptual view seeks to balance preservation and use. From the mid-1980s to the early 2000s, international organizations supported a number of Integrated Conservation and Development Projects (ICDPs) that aimed to both conserve biodiversity and develop the local economy by providing alternative livelihoods and incomes while curtailing local resource use mainly in protected areas. The results, however, were lackluster and much has been written about the problematic nature of ICDPs (Sanjayan *et al.*, 1997; Wells *et al.*, 1999; Adams and Hulme, 2001; McShane and Newby, 2004; Hutton *et al.*, 2005). Although interest in classical ICDPs has waned, the *de facto* integration of conservation and development activities in approaches such as Population, Health, and Environment (PHE) projects continues.

The fourth conceptual view is on the usable portion of the spectrum and focuses on the essential goods and services that natural ecosystems provide to people. Within this conceptual view, ecosystem services (Daily, 1997) and the comanagement of natural resources (Carlsson and Berkes, 2005) are found. Proponents of ecosystem service note that it is the poor who often suffer most when ecosystem services, such as clean water and clean air, decline (Wunder, 2008). Proponents may be willing to trade a reduction in the biodiversity for an ecosystem for sustainable benefits to people and protecting the remaining nature (Kareiva and Marvier, 2010). The emphasis within the study of ecosystem services is often on valuations of ecosystem services to show the importance of an ecosystem to people in general and the poor in particular (e.g., the Natural Capital Project). The implicit assumption is that a better understanding of the value of ecosystem services will improve the political will to protect these services. Yet among the extremely poor, time horizons are often short and economic discount rates are high. Hence, managing ecosystem services for long-term sustainability in places with extreme poverty may be challenging. Within comanagement of natural resources, there is the recognition that local communities have the greatest stake in the sustainable use of the local natural resources, but that community-based natural resource management can just as easily be backwards and fragmented as progressive and cohesive (Blaikie, 2006). By balancing community power with government power in a comanagement arrangement, natural resources can in principle be more equitably and sustainably managed. Comanagement conceptual analysis and a management framework have been developed (Borrini-Feyerabend *et al.*, 2004), but long-term experience with this approach is limited.

Key Players

Understanding the key players in the conservation–poverty field is important for understanding the drivers of conservation–poverty activities. Many players have an institutional bias that emphasizes either conservation or poverty reduction.

The primary implementers of conservation–poverty projects are the international conservation nongovernmental organizations (NGOs) ranked here by 2009 overall budgets: The Nature Conservancy (TNC), the WWF, Conservation International (CI), the International Union for the Conservation of Nature (IUCN), Wildlife Conservation Society, Fauna and Flora International, African Wildlife Fund, the BirdLife International network, and Wetlands International.

Among development NGOs in 2011, those with an active interest in poverty–conservation issues include CARE International members, Pathfinder International, and World Education. CARE in East Africa has a long-time presence in the poverty–conservation field.

Among development agencies working on poverty and conservation issues, the standouts are Department for International Development (DFID), United States Agency for International Development (USAID), United Nations Development Program (UNDP), and the World Bank. DFID was the first government aid agency to extensively research the links between conservation and poverty reduction. USAID fostered community-based natural resource management in the 1990s and PHE projects more recently. UNDP is an important convener of conservation–poverty meetings. The World Bank was a large implementer of ICDPs in the late 1990s and has published influential studies examining the links between environment and poverty. The World Bank is also a key player in reduced emissions from deforestation and degradation that aims to link forest preservation, reduced CO_2 emissions, and poverty reduction. Norway and its bilateral aid agency Norwegian Agency for Development Cooperation have joined with the World Bank in this work. Other bilateral agencies with a history in poverty and conservation projects include Danish International Development Agency (Denmark), Swiss Agency for Development and Cooperation (Switzerland), Swedish International Development Cooperation Agency (Sweden), Canadian International Development Agency (Canada), Finnish International Development Agency (Finland), Kreditanstalt für Wiederaufbau (Germany), Deutsche Gesellschaft für Internationale Zusammenarbeit (Germany), and Fonds Français pour l'Environnement Mondial (France).

Between the conservation and development organizations is the London-based International Institute for Environment and Development. This organization has been at the center of conservation–poverty issues since the early 2000s, as a convener, publisher, and an analytic force.

Other international organizations important to conservation and poverty reduction include Consultative Group on

International Agriculture Research members, the Center for International Forestry Research, World Agroforestry Center, and the World Fish Center. The Global Environment Facility and the Convention on Biological Diversity Secretariat are also influential multilateral players in conservation and poverty issues.

Does Conservation Benefit the Poor?

Evidence Showing Conservation has Benefited the Poor

Perhaps the central question in conservation and poverty studies is to what extent does the conservation of biodiversity actually benefit the poor and the extreme poor? Much of the existing conservation–poverty knowledge is based on anecdotal evidence for a variety of reasons (Ferraro and Pattanayak, 2006). There is, however, suggestive empirical evidence that conservation initiatives can measurably benefit the poor and the extremely poor (see Leisher *et al.*, 2010 for the full review). The best evidence of conservation initiatives that have reduced poverty comes from forestry, fisheries, grasslands, and nature-based tourism and is summarized here.

Forestry

Within the forestry sector, evidence of conservation benefiting the poor is largely found in community-based timber enterprises. Timber is harvested in such enterprises at levels that ensure the ecological integrity of the forest, whereas the income generated from the timber sales contributes to local poverty reduction. Such enterprises are frequently based on contractual arrangements between communities and companies to supply fiber, pulp, or construction timber.

A study of 14 community-forestry timber enterprises in developing countries found that they can be profitable (Molner *et al.*, 2007), and in Mexico, community-forestry enterprises have helped reduce local poverty (Antinori and Bray, 2005). For nature, community forestry can also be a benefit. There is evidence from Nepal (Gautam *et al.*, 2002), Mexico (Bray *et al.*, 2003), and Vietnam (Nguyen *et al.*, 2009) that community forestry has led to increases in forest cover.

It is not only timber that can benefit poor communities but also the forest itself. A meta-analysis of 51 case studies from 17 countries found that forest environmental income represents on average 22% of the total rural household income (Vedeld *et al.*, 2004). Fuelwood and nontimber forest products are often an integral part of the livelihood strategies of the poor and extremely poor and maintaining an intact forest has benefits to the rural poor that are often overlooked (Naughton-Treves *et al.*, 2011).

Fisheries

Near-shore fisheries provide a high percentage of animal protein for people in poor countries, but many near-shore fisheries are fished beyond the limits of sustainable yields (FAO, 2010). One strategy to address this is the creation of no-take zones. Such zones provide space for fish to grow bigger, and bigger fish usually have exponentially more offspring than smaller fish (PISCO, 2007). After several years of protection, fish may spillover into the area outside the no-take zone where they can be caught (Gell and Roberts, 2003). Greater fish catches generate more income for fishers and hence help reduce poverty, whereas the no-take zone provides protected habitat for marine life.

In places where many people are poor, and fishing is in crisis, there may be medium-term poverty reduction benefits from no-take zones. WRI (2005) and Leisher *et al.* (2007) found that spillover from two community-managed marine areas in Fiji roughly doubled local incomes within 5 years of establishing the no-take zone compared with control sites, and women were the primary beneficiaries.

The indirect benefits to the poor may also be important. Organizing a community to manage a no-take zone often strengthens the social fabric of the community, giving them a ready decision-making body and a more unified voice to address other community issues (Leisher *et al.*, 2007). The stronger social cohesion also improves local security and empowers local decision making, two key elements of poverty reduction (World Bank, 2001).

There is evidence that no-take zone management is more effective where there is active participation of local communities in the resource management (McClanahan *et al.*, 2006). There is also evidence that size matters. If a no-take zone is too large, spillover will not offset the losses to fishers from closing sections of the fishing grounds (PISCO, 2007). From the perspective of fisheries, networks consisting of a number of smaller marine protected areas may be preferable to a few very large marine protected areas (IUCN-WCPA, 2008).

Grasslands

Grasslands can have greater grass productivity with grazing than without grazing (Guo, 2007), and some types of grassland are ecologically dependent on grazing to maintain their biodiversity (Fratkin and Mearns, 2003). Within grasslands, livestock is often the primary form of wealth for local communities (Eriksen and Watson, 2009), and these communities can be among the extreme poor when measured by a multidimensional definition of poverty (e.g., Glew *et al.*, 2010). The relationship among grazing, biodiversity, and local livelihoods is what underpins grasslands conservation projects that have benefited the poor. Pastoralism, especially in drylands, may be one of the few land uses that is genuinely compatible with "formal" nature conservation (WISP, 2008).

Studies in a South African community-managed grassland (Leisher *et al.*, 2011a), in the Mongolian Gobi grasslands (Leisher *et al.*, 2011b), and in northern Kenya's grasslands (Glew *et al.*, 2010) used household surveys and remote sensing to measure differences before and after and between impacted and matched control areas using a variety of socioeconomic and ecological indicators. These studies identified a number of benefits to the poor and extreme poor as well as greater net primary productivity of project grasslands compared with before the project and matched control sites. The evidence from these studies suggests that better grasslands management can benefit both the poor and the grasslands themselves.

Nature-Based Tourism

Nature-based tourism can include community-based operations on one end and all-inclusive international ecolodges and safari operations on the other. Tourism can offer opportunities for reducing poverty via jobs in the formal tourism sector, such as accommodation and guiding, and in new

markets for local services and products, such as sales of crafts, cultural services, food, and drinks. Tourism may also be associated with infrastructure development, such as roads, telecommunications, and health care facilities, which when provided for the benefit of tourists, benefit the poor as well.

A meta-review of 27 tourism case studies in Asia found income gains for all economic levels but with those already better-off gaining most (Shah and Gupta, 2000). In Zambia, a World Bank Study (2007) found that nature-based tourism had reduced poverty, but those with most assets benefit up to 50% more than the extremely poor. The indirect benefits may be important to the poor as well. There is a multiplier effect from tourism that creates opportunities and downstream effects for casual laborers, crafters, and small businesses (Ashley and Roe, 2002). The indirect benefits, such as better infrastructure, can also benefit poor people (e.g., Shackleton *et al.*, 2007). Where tourism operators commit to hire and train local people, the financial benefits to a local community can be substantial (Davis, 2005).

Although there is suggestive empirical evidence that a conservation initiative can benefit the poor from community-forestry timber enterprises, no-take zones for fish, grasslands management, and nature-based tourism, there are other areas where the evidence of poverty reduction is even more limited but may have potential for benefiting the poor, including nontimber forest products, payments for environmental services, mangrove restoration, protected area jobs, and agroforestry (Leisher *et al.*, 2010). Some, such has nontimber forest products and payments for environmental services, have been well studied but the evidence is inconclusive. Others remain to be studied rigorously.

Evidence Showing Conservation has Hurt the Poor

The evidence that conservation initiatives have hurt the poor comes primarily from the establishment of new protected areas that displaced local people (Brockington *et al.*, 2006). This can be either physical displacement, such as eviction from their homes, or economic displacement by excluding people from an area that is important to their livelihoods (Cernea, 2005). Regardless of the type of displacement, the impacted people are often worse-off economically, culturally, and socially. For indigenous people displaced by a conservation initiative, the negative impacts can continue for decades after the displacement (Colchester, 2003).

The number of people and the number of places where local people have been displaced by a new protected area are disputed (Borgerhoff-Mulder and Coppolillo, 2005; Wilkie *et al.*, 2006), but most of the cases occurred before 1980. In recent decades, the impacts of displacing people from a new protected area have become better understood, and most donors and conservation organizations are reluctant to participate in such projects.

Other Key Issues

In addition to the issue of whether conservation initiatives have helped or hurt the poor, there are several other key issues within conservation and poverty.

Perhaps foremost is the considerable heterogeneity among poverty indicators used to measure poverty impact. The choice of indicators is often driven by differing definitions of who is poor (Agrawal and Redford, 2006). Existing definitions of "poor" may be qualitative or quantitative, temporal or points-in-time, and may be multidimensional or one dimensional (Ravallion, 2008). The poverty reduction impact of a conservation initiative depends greatly on how poverty is defined and the indicators used to measure it.

Conservation and poverty reduction links are often dynamic and locally specific (Kepe *et al.*, 2004; Sanderson and Redford, 2004; BirdLife, 2007). Thus, universal models applicable to multiple local contexts for jointly addressing conservation and poverty are unlikely. Instead, habitat types where the biological diversity itself is fundamental to the benefits to people may hold more promise for conservation and poverty reduction. Examples of this are fisheries in which greater biodiversity produces greater biomass (Worm *et al.*, 2006) and grasslands that have coevolved with grazers and now depend on grazing to maintain plant diversity (Fratkin and Mearns, 2003).

Elite capture of benefits is also an issue. The wealthier segments of a population participating in a conservation or development initiative may "capture" the majority of the benefits (e.g., Jagger, 2008; Bandyopadhyay and Tembo, 2009). Part of the reason is because the better-off generally have greater human and financial capital, which allows them to benefit more from an initiative (Weber *et al.*, 2011), and they are also more likely to participate in an initiative (Groom *et al.*, 2010). Where local leaders are not downwardly accountable to their constituents, the local elites may shape a conservation initiative to benefit themselves (Jumbe and Angelsen, 2006).

A key issue often overlooked is that generating benefits from conservation to the poor and especially the poorest of the poor is frequently more about increasing biomass than increasing biodiversity (Balmford *et al.*, 2008). For example, the volume of fish is usually more important to a small-scale fisher than the variety. Although in a few natural ecosystems, biomass productivity may be linked to species biodiversity, in other cases, a managed and less diverse ecosystem may provide greater production of the biomass that poor people value than a natural ecosystem. Examples of this include woodlots for firewood production and grasslands converted to agriculture.

A final issue is the limited political will to improve the well-being of people in hard-to-reach areas. Rural, impoverished populations tend to be politically marginalized and expensive to reach with services. Conservation initiatives are frequent in remote areas, and the national government's willingness to work in these areas on conservation–poverty issues may be lacking. This has implications for the sustainability of externally driven initiatives, especially those with exit strategies dependent on the government assuming a greater role in service delivery.

Trends Within Conservation and Poverty

Within conservation and poverty reduction, what is happening at the global level influences what happens at the local level. Herein, we look at the macro- and microlevel trends within conservation and poverty.

Although the proportion of people under US$1.25 a day is dropping, the absolute number of the extreme poor continues to climb in most countries (**Table 1**). Much of the growth in the number of poor is in urban areas, but it is in rural areas where poverty is most often a chronic condition. It is also in rural areas where the extreme poor tend to be most dependent on the goods and services nature provides. Given existing demographic trends, the number of poor people is likely to increase in many of the remote, rural places where conservation organizations work.

Poverty reduction is also becoming more challenging. The Millennium Development Goals signed by 192 of the world's governments call for reducing poverty by 50% between 1990 and 2015 as the top priority. Even if this goal is achieved, this would leave more than 2 billion people still in poverty. Moreover, the easier gains would be past because those just below the poverty line will have moved over the line, and the ones left will require greater effort. Inside and outside of conservation, poverty will remain one of the greatest challenges globally.

Table 1 Number of people living on less than US$ 1.25 a day (millions)

Region	*1990*	*2005*	*Change*
South Asia	579	595	+16
Sub-Saharan Africa	295	387	+92
China	683	208	−475
East Asia and Pacific (minus China)	190	109	−81
Other	68	72	+4

The World Database on Protected Areas shows that the rate of establishing new protected areas has dropped sharply, declining from a historic high of 2920 in 2002 to 40 in 2006 (UNEP-WCMC, 2009). For a variety of reasons, the focus within conservation has shifted away from establishing new protected areas. Protected areas will remain the core conservation strategy of most governments, but much of the conservation efforts in the twenty-first century are likely to be outside protected areas. This is especially true for conservation–poverty issues where efforts outside of designated protected areas are likely to yield greater conservation gains with greater poverty reduction benefits than inside protected areas.

Another macro trend is the devolution of power. The locus for decision making in many countries has shifted to the subnational level. The principle is that power should be devolved to the lowest feasible level. The assumption is that the closer decision making is to the problem, the more likely the problem will be addressed successfully. For conservation, this may entail working more with local-level governments. For conservation and poverty reduction, this may result in focusing more on restoring degraded habitats than preserving intact habitats given the presumed development focus on most local-level governments.

At the microlevel, conservation organizations have responded by making people more central to their mission. TNC and CI both changed their mission statements to highlight the benefits to people from conservation. Three of the largest US foundations that fund conservation organizations (Moore, Packard, and MacArthur) have shifted their funding toward more people-centered conservation.

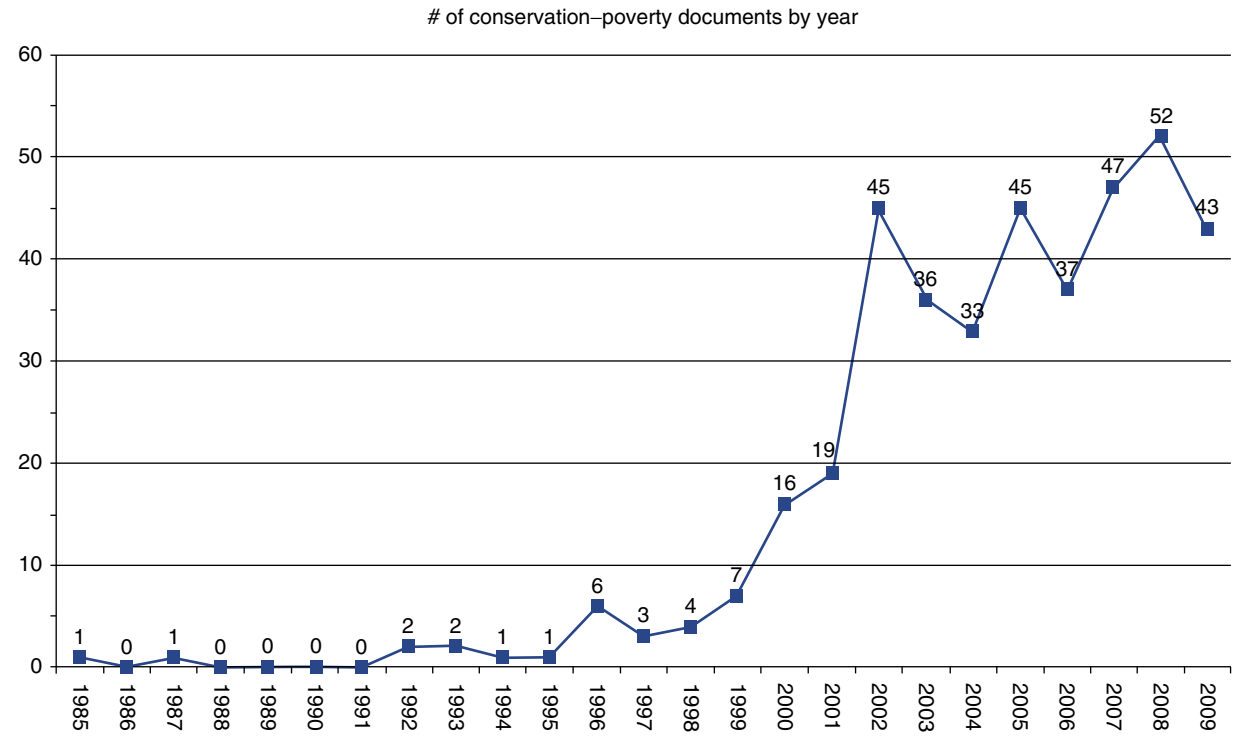


Figure 1 Number of conservation–poverty publications by year.

There is also a growing recognition within international conservation organizations that the social sciences are as important as the natural sciences. After several decades of defining where the Earth's maximal biodiversity is located and where it is most threatened, conservation organizations know where they need to work. What is lacking is a greater understanding of how they need to work. One promising development comes from the Nobel Prize-winning economist Elinor Ostrom. She identified how communities that self-organize to manage common-pool resources, such as fisheries and forests, can avoid the tragedy of the commons. Ostrom found 10 factors that are predictive of communities self-organizing to manage socioecological systems (Ostrom, 2009). Gutierrez *et al.* (2011) did similar work for marine areas and highlighted the success factors for fisheries comanagement. Yet these success factors were determined via ex-post quantitative analysis. Identifying and measuring the level of key factors before a conservation initiative begins (*ex-ante*) could improve conservation success. Turning the known conditions for success into site-screening tools that can be used by social scientists is a nascent trend in conservation.

Another micro trend is the greater focus on measuring the impact of a conservation initiative on local people and the poor. Ferraro and Pattanayak (2006) provided a good summary of why this is needed and approaches that can separate what happened because of a conservation initiative and what would have happened without the conservation initiative. The authors call for greater use of baselines, measure of covariates, and matched control groups. Several international conservation organizations have been working together on improving conservation impact measures (www.conservationmeasures.org).

A final trend is the growing interest in conservation–poverty issues as evidenced by the growth in the literature on the topic (**Figure 1**). A comprehensive literature review in June 2010 found 430 articles, books, and publications on conservation–poverty topics (Leisher *et al.*, 2010). Few were published before 1998, but the field began to grow rapidly around 2000. In recent years, the average is about 50 new publications a year on the topic.

Conclusions

As a former rancher in North Dakota, President Teddy Roosevelt believed in using the land to benefit people but wanted to balance this with protection of special places. He wanted both preservation and use of natural areas, and that is why he gave his political support to both Muir and Pinchot. Roosevelt recognized that preservation and use need not be an either/or choice but a spectrum with strict preservation on one end and destructive use on the other. Usage that is sustainable can enable biodiversity preservation. It also addresses the reality of several billion people globally needing resources to move out of poverty.

The knowledge base of conservation–poverty issues continues to grow and the rigor of new studies has increased such that there is far more evidence of impacts than a decade ago. On balance, however, the evidence base is still limited and largely suggestive of how the poorest of the poor may benefit from conservation.

Appendix

List of Courses

1. Conservation Beyond Reserves – Oxford University http://www.geog.ox.ac.uk/graduate/msc-bcm/structure.html#core4
2. Management of Ecosystems and Landscapes – Columbia University http://www.columbia.edu/cu/e3b/graduate_offerings.html
3. Social Science of Development and Conservation – Yale School of Environment http://environment.yale.edu/courses/detail/535/
4. Human Dimensions of Biological Conservation – University of Florida http://www.wec.ufl.edu/entities/pstc/courses.php

See also: Biodiversity and Human Health. Biodiversity, Human Well-Being, and Markets. Conservation and People. Economic Value of Biodiversity, Measurements of. Indigenous Peoples and Biodiversity. International Organizations and Biodiversity. Justice, Equity and Biodiversity. Market Economy and Biodiversity. Poverty and Biodiversity

References

Adams WM, Aveling R, Brockington D, *et al.* (2004) Biodiversity conservation and the eradication of poverty. *Science* 306: 1146–1149.

Adams WM and Hulme D (2001) If community conservation is the answer in Africa, what is the question? *Oryx* 35: 193–200.

Agrawal A and Redford K (2006) *Poverty, Development and Biodiversity Conservation. Shooting in the Dark?*. New York, NY: Wildlife Conservation Society. WCS Working Paper No. 26.

Antinori C and Bray DB (2005) Community Forest enterprises as entrepreneurial firms: Economic and institutional perspectives from Mexico. *World Development* 33: 1529–1543.

Ashley C and Roe D (2002) Making tourism work for the poor: Strategies and challenges in southern Africa. *Development Southern Africa* 19: 61–82.

Baland J, Bardhan P, Das S, Mookherjee D, and Sarkar R (2006) Managing the environmental consequences of growth: Forest degradation in the Indian mid-Himalayas. In: Bery S, Bossworth B, and Panagariya A (eds.) *India Policy Forum*. New Delhi, India: Sage Publications.

Balmford A, Rodrigues ASL, Walpole M, *et al.* (2008) *The Economics of Ecosystems and Biodiversity: Scoping the Science*. Cambridge, UK: European Commission.

Bandyopadhyay S and Tembo G (2009) *Household Welfare and Natural Resource Management Around National Parks in Zambia*. Washington, DC: The World Bank. Policy Research Working Paper No. 4932.

BirdLife (2007) *Livelihoods and the Environment at Important Bird Areas: Listening to Local Voices*. Cambridge, UK: BirdLife International.

Blaikie P (2006) Is small really beautiful? Community-based natural resource management in Malawi and Botswana. *World Development* 34: 1942–1957.

Borgerhoff-Mulder M and Coppolillo P (2005) *Conservation. Linking Ecology, Economics and Culture*. Princeton, NJ: Princeton University Press.

Borrini-Feyerabend G, Pimbert M, Farvar T, Kothari A, and Renard Y (2004) *Sharing Power: Learning by Doing in Co-management of Natural Resources Throughout the World*. London, UK: IIED.

Brashares JS, Golden CD, Weinbaum KZ, Barrett CB, and Okello GV (2011) Economic and geographic drivers of wildlife consumption in rural Africa.

Proceedings of the National Academy of Sciences of the United States of America 108: 13931–13936.
Bray DB, Merino-Pérez L, Negreros-Castillo P, Segura-Warnholtz G, Torres-Rojo JM, and Vester HFM (2003) Mexico's community-managed forests as a global model for sustainable landscapes. *Conservation Biology* 17: 672–677.
Brockington D, Igoe J, and Schmidt-Soltau K (2006) Conservation, human rights, and poverty reduction. *Conservation Biology* 20: 250–252.
Brundtland G (ed.) (1987) *Our common future: The World Commission on Environment and Development*. Oxford, UK: Oxford University Press.
Carlsson L and Berkes F (2005) Co-management: Concepts and methodological implications. *Journal of Environmental Management* 75: 65–76.
Cernea M (2005) Restriction of access is displacement: A broader concept and policy. *Forced Migration Review* 23: 48–49.
Cavendish W (2000) Empirical regularities in the poverty–environment relationship of rural households: Evidence from Zimbabwe. *World Development* 28: 1979–2003.
Colchester M (2003) *Salvaging Nature: Indigenous Peoples, Protected Areas and Biodiversity Conservation*. Moreton-in-Marsh, UK: World Rainforest Movement, Forest Peoples Program.
Daily G (ed.) (1997) *Nature's Services. Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.
Davis, T (2005) *Local and Semi-Local Economic Impacts of Dive Tourism in Bunaken National Park, North Sulawesi, Indonesia*. Master Thesis, School of Marine Affairs, University of Washington, Seattle.
Eriksen SEH and Watson HK (2009) The dynamic context of southern African savannas: Investigating emerging threats and opportunities to sustainability. *Environmental Science and Policy* 12: 5–22.
FAO (2010) *State of the World's Fisheries and Aquaculture*. Rome, Italy: FAO.
Ferraro PJ and Pattanayak SK (2006) Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *Public Library of Science Biology* 4: e105.
Fratkin E and Mearns R (2003) Sustainability and pastoral livelihoods: Lessons from east African Maasai and Mongolia. *Human Organization* 62: 112–122.
Gautam AP, Webb EL, and Eiumnoh A (2002) GIS assessment of land use/land cover changes associated with community forestry implementation in the middle hills of Nepal. *Mountain Research and Development* 22: 63–69.
Gell FR and Roberts CM (2003) Benefits beyond boundaries: The fishery effects of marine reserves and fishery closures. *Trends in Ecology and Evolution* 18: 448–455.
Glew L, Hudson MD, and Osborne PE (2010) *Evaluating the Effectiveness of Community-Based Conservation in Northern Kenya*. Southampton, UK: University of Southampton. Unpublished report.
Groom B, Grosjean P, Kontoleon A, Swanson T, and Zhang S (2010) Relaxing rural constraints: A "win-win" policy for poverty and environment in China? *Oxford Economic Papers* 62: 132–156.
Guo Q (2007) The diversity-biomass-productivity relationship in grasslands management and restoration. *Basic and Applied Ecology* 8: 199–208.
Gutierrez NL, Hilborn R, and Defeo O (2011) Leadership, social capital and incentives promote successful fisheries. *Nature* 470: 386–389.
Hutton J, Adams WM, and Murombedzi J (2005) Back to the barriers? Changing narratives in biodiversity conservation. *Forum for Development Studies* 2: 341–370.
IUCN World Commission on Protected Areas (IUCN-WCPA) (2008) *Establishing Marine Protected Area Networks – Making It Happen*. Washington, DC: IUCN-WCPA, National Oceanic and Atmospheric Administration, and The Nature Conservancy.
Jagger P (2008) *Forest Incomes After Uganda's Forest Sector Reform: Are the Rural Poor Gaining?*. Washington, DC: International Food Policy Research Institute. CAPRi Working Paper No. 92.
Jumbe C and Angelsen A (2006) Do the poor benefit from devolution policies? Evidence from Malawi's forest co-management program. *Land Economics* 82: 562–581.
Kareiva P and Marvier M (2010) *Conservation Science: Balancing the Needs of People and Nature*. Greenwood Village, CO: Roberts and Company.
Kepe T, Saruchera M, and Whande W (2004) Poverty alleviation and biodiversity conservation: A South African perspective. *Oryx* 38(2): 143–145.
Leisher C, Brouwer R, Boucher TM, Vogelij R, Bainbridge WR, and Sanjayan M (2011a) Striking a balance: Socioeconomic development and conservation in grassland through community-based zoning. *Public Library of Science One* 6(12): e28807.http://dx.doi.org/10.1371/journal.pone.0028807.
Leisher C, Hess S, Boucher TM, van Beukering P, and Sanjayan M (2011b) *Measuring the Impacts of Community-Based Grasslands Management in Mongolia's Gobi*. Arlington, TX: The Nature Conservancy.
Leisher C, Sanjayan M, Blockhus J, Kontoleon A, and Larsen SN (2010) Does conserving biodiversity work to reduce poverty. In: Roe D (ed.) *Linking Biodiversity Conservation and Poverty Alleviation: A State of Knowledge Review*. Montreal, Canada: CBD Secretariat. CBD Technical Series No: 55.
Leisher C, Van Beukering P, and Scherl LM (2007) *Nature's Investment Bank: How Marine Protected Areas Contribute to Poverty Reduction*. Arlington, TX: The Nature Conservancy.
McClanahan T, Marnane M, Cinner J, and Kiene W (2006) A comparison of marine protected areas and alternative approaches to coral-reef management. *Current Biology* 16: 1408–1413.
McShane TO and Newby SA (2004) Expecting the unattainable: The assumptions behind ICDPs. In: McShane TO and Wells MP (eds.) *Getting Biodiversity Projects to Work: Towards More Effective Conservation and Development*. New York, NY: Columbia University Press.
Molner A, Liddle M, Bracer C, Khare A, White A, and Bull J (2007) Community-based forest enterprises in tropical forest countries: Status and potential. *Report to the ITTO*. Washington, DC: Forest Trends/Rights and Resources Group.
Narain U, Gupta S, and Van't Veld K (2008) Poverty and resource dependence in rural India. *Ecological Economics* 66: 161–176.
Naughton-Treves L, Alix-Garcia J, and Chapman CA (2011) Lessons about parks and poverty from a decade of forest loss and economic growth around Kibale National Park, Uganda. *Proceedings of the National Academy of Sciences of the United States of America* 108: 13919–13924.
Ndangalasi HJ, Bitariho R, and Dovie D (2007) Harvesting of non-timber forest products and implications for conservation in two montane forests of East Africa. *Biological Conservation* 134: 242–250.
Neumann RP and Hirsch E (2000) *Commercialization of Non-Timber Forest Products: Review and Analysis of Research*. Bogor, Indonesia: Center for International Forestry Research.
Nguyen QT, Tran NT, Hoang HT, Yasmi Y, and Enters T (2009) *Red Books for Greener Trees: Strengthening Community Forestry in Vietnam*. Hanoi, Vietnam: Forest Governance Learning Group.
Ostrom E (2009) A general framework for analyzing sustainability of social-ecological systems. *Science* 325: 419–422.
PISCO (2007) *The Science of Marine Reserves*. Corvallis, OR, USA: Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO). www.piscoweb.org.
Ravallion M (2008) Poverty lines. In: Blume L and Durlauf S (eds.) *The New Palgrave Dictionary of Economics*. London, UK: Palgrave Macmillan.
Redford KH, Levy MA, Sanderson EW, and deSherbinin A (2008) What is the role for conservation organizations in poverty alleviation in the world's wild places? *Oryx* 42: 512–528.
Roe D (2008) The origins and evolution of the conservation poverty debate: A review of key literature, events and policy processes. *Oryx* 42: 491–503.
Roe D and Elliott J (2004) Poverty reduction and biodiversity conservation: Rebuilding the bridges. *Oryx* 38: 137–139.
Sanderson S and Redford K (2004) The defense of conservation is not an attack on the poor. *Oryx* 38: 146–147.
Sanderson SE and Redford K (2003) Contested relationships between biodiversity conservation and poverty alleviations. *Oryx* 37: 389–390.
Sanjayan M, Shen S, and Jansen M (1997) *Experiences with Integrated-Conservation Development Projects in Asia*. Washington, DC: World Bank. World Bank Technical Paper 388.
Shackleton C, Shackleton S, Buitenb E, and Bird N (2007) The importance of dry woodlands and forests in rural livelihoods and poverty alleviation in South Africa. *Forest Policy and Economics* 5: 558–577.
Shah K and Gupta V (2000) *Tourism, the Poor and Other Stakeholders: Experience in Asia*. London, UK: Odi Fair Trade in Tourism Project.
UNEP-WCMC (2009) *World Database on Protected Areas*. Cambridge, UK: UNEP-WCMC.
Vedeld P, Angelsen A, Sjaastad E, and Kobugabe Berg G (2004) *Counting on the Environment: Forest Incomes for the Rural Poor*. Washington, DC: World Bank. Environmental Economics Series No. 98.
Weber JG, Sills EO, Bauch S, and Pattanayak SK (2011) Do ICDPs Work? An empirical evaluation of forest-based microenterprises in the Brazilian Amazon. *Land Economics* 87: 661–681.
Wells M and Brandon K (1992) *People and Parks: Linking Protected Area Management with Local Communities*. Washington, DC: The World Bank.
Wells MP, Guggenheim S, Khan A, Wardojo W, and Jepson P (1999) *Investing in Biodiversity: A Review of Indonesia's Integrated Conservation and Development Projects*. Washington, DC: World Bank.
Wilkie D, Morelli G, Demmer J, Starkey M, Telfer P, and Steil M (2006) Parks and people: Assessing the human welfare effects of establishing protected areas for biodiversity conservation. *Conservation Biology* 20: 247–249.

World Bank (2001) *Attacking Poverty*. Oxford, UK: Oxford University Press.
World Bank (2007) *Zambia Economic and Poverty Impact of Nature-Based Tourism*. Washington, DC: The World Bank. Report No. 43373-ZM.
World Initiative for Sustainable Pastoralism (WISP) (2008) *Policies that Work for Pastoral Environments: A Six-Country Review of Positive Policy Impacts on Pastoral Environments*. Nairobi, Kenya: WISP.
Worm B, Barbier EB, Beaumont N, *et al*. (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.
WRI (2005) *World Resources 2005: The Wealth of the Poor – Managing Ecosystems to Fight Poverty*. Washington, DC: World Resources Institute.
Wunder S (2008) Payments for environmental services and poor: Concepts and preliminary evidence. *Environment and Development Economics* 13: 279–297.

Conservation Biology, Discipline of

Andrew P Dobson and Katarzyna Nowak, Princeton University, Princeton, NJ, USA
Jon Paul Rodríguez, Instituto Venezolano de Investigaciones Científicas, Maracaibo, Venezuela

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Andrew P. Dobson and Jon Paul Rodríguez, volume 1, pp 855–864, © 2001, Elsevier Inc.

Glossary

Allee effect Biological phenomenon whereby the survival of the individuals of a population increases by spatial aggregation.
Beta diversity Spatial variation of species diversity across sites within a region community. All forms of life that coexist and interact with each other in a particular habitat.
Community All forms of life that coexist and interact with each other in a particular habitat.
Ecological footprint Cumulative impact of all people.
Ecosystem Complex dynamic system that arises by the integration of a biological community with the physical environment.
Ecosystem services Benefits obtained by people from (fully functioning) ecosystems.
Edge effects Penetration of abiotic and biotic conditions including wind, microclimate, and ecological traps from the matrix into forest interiors that effectively reduces the size of a patch.
Extinction vortex Positive feedback loops that increase the risk of extinction of a population as it declines in size.
Gene bank Facility where genetic material is stored in the form of seeds, pollen, embryos, or semen, or, in some cases, as live plants or animals living in a field, a greenhouse or other installation.
Habitat corridors or connections Strips of habitat that connect isolated habitat patches in a landscape transformed by human land use. Connections can be achieved through the conservation of existing habitat or by ecological restoration.
Habitat restoration Active modification of the current state of a degraded habitat in order to return it to a former, preferred state.
Matrix A heterogeneous area composed of a dynamic mosaic of interacting habitats, some natural and some anthropogenic.
Maximum sustainable yield Policy of maximizing the harvest of natural resources without reducing their future potential which often overestimates yield.
Population A group of individuals with common ancestry that are much more likely to mate with one another than with individuals from another such group.
Systems theory Framework for understanding the dynamics of interconnected components trophic cascade changes in a community as a result of changes in the abundance of a species or group of species.
Trophic cascade Changes in a community as a result of changes in the abundance of a species or group of species.

Definition

The current global crisis in the loss of biodiversity is the result of the immense success of one species, *Homo sapiens*, at the expense of the majority of most other species. Concern about the loss of biodiversity arises from spiritual, moral, and esthetic motives, through economic rationales, to purely selfish reasons. Moreover, the growth of these concerns has fueled the rapid expansion and awareness of conservation biology into academia, government, industry, and other sectors of society worldwide, earning it the label "discipline with a deadline" (Wilson, 2002). The wider field of conservation relies on efforts across scales, from the vital grassroots at which small-scale solutions to local conservation problems are piloted and where local people act as stewards of natural biodiversity; through the large international conservation NGOs, capable of land purchases, of acting as watchdogs, and of establishing enduring partnerships with governments and private land-owners to governmental agencies that pass and enforce conservation directives. The "formulae" that underpin these efforts come from the science – Conservation Biology – which aims to translate into practice and action outcomes of prioritization and horizon scanning exercises, and optimize research and monitoring to ultimately influence conservation policy and ensure that it is evidence-based. Although the first conservation biology textbooks and scientific journals date back to the early 1960s, conservation biology's growth as a scientific discipline during the last three decades has been rapid and global, and is evident in the academic training programs in conservation biology that now span the world.

Conservation biologists have an enormous agenda. Initially, they sought to understand the impact on biodiversity of the complex interactions between human societies and the natural world. They hope this understanding will provide guidelines for minimizing the negative effects of human actions on the persistence of biodiversity. Increasingly, the discipline has grown to recognize the great extent to which humans depend on biodiversity, and has begun to explore the services that natural ecosystems provide to humans. Ultimately, conservation biology aims to guarantee the persistence of natural landscapes, healthy ecosystems, species,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00272-0

populations, and genes as well as the complex interactions among them and the dynamic processes that characterize them. This pursuit is necessarily trans-disciplinary, integrating the principles and practices of a variety of disciplines, principally ecology, population genetics, economics, and biogeography, but also increasingly anthropology, psychology, education, engineering, law, and public policy.

Ecosystems and Communities

The mandate of conservation biology covers a spectrum that runs from concern about the conservation of large, intact, functioning ecosystems, through the maintenance of viable indigenous human and ecological communities, to the preservation of the last surviving individuals of species in zoos and botanical gardens. Arguably, this spectrum may be extended to include the DNA of currently extant or recently extinct species in museums. This spectrum includes both processes that are complex and only partially understood and ones that are more completely characterized.

The most economical and effective way to conserve biodiversity is through *in situ* protection that guarantees the persistence of all the processes that characterize biodiversity in a fully functioning ecosystem. Biogeochemical and hydrologic cycles, the flow of energy through food webs, and evolution by natural selection all require a complex web of interacting biological and physical elements. It is in general unrealistic to assume that these organisms and the processes they mediate can be removed from the natural setting and recreated in a zoo, botanical garden, theme park, or gene bank. For long-term persistence, they must be conserved where they naturally occur. The most popular instruments for *in situ* conservation are nature reserves and protected areas. According to the International Union for Conservation of Nature (IUCN), a protected area is a section of land or sea especially dedicated to the protection and maintenance of biodiversity and of natural and associated cultural resources, which is managed through legal or other effective means. Protected areas come in many kinds and differ mainly in size and the degree of human intervention that is allowed to occur within them. At one extreme are large, strict nature reserves and wilderness areas, where little or no obvious modification of the landscape is permitted. Examples of these include the Bob Marshall wilderness in Montana, US, and the Parima-Tapirapecó national park in southern Venezuela. At the other extreme are "managed resource" protected areas, which are devoted to the sustainable use of natural ecosystems, such as the networks of game reserves in Botswana or Spain, or community-run "Wildlife Management Areas" (WMAs) in Tanzania. In all cases, the maintenance of ecological and evolutionary processes is recognized as a key motivation for the creation of such reserves, even in the case of WMAs where the proximate goal is to ensure that local people benefit from conservation. Transboundary parks are meanwhile becoming more prevalent, which may be an evidence of improved high-level political cooperation in conservation and natural resource management, and the recognition that animal home ranges often intersect political boundaries.

Ecosystem studies play an increasingly major role in understanding complex ecological systems, and they provide important guidelines for the management of protected areas.

Although there are very few examples of nature reserve systems that were based from their inception on ecological principles, conservation biologists have undertaken many *post hoc* analyses of reserve systems that were originally designated for political, esthetic, and occasionally scientific reasons. For example, Yellowstone National Park in the US, formerly the world's first national park, was founded for the benefit of the people so that they could appreciate its spectacular and unique geohydrothermic features. By chance, it also happens to contain a nearly complete set of the fauna and flora found in the North American Northern Rockies, and thus is also valuable from the perspective of the conservation biologist.

The first systematic approaches to examining the design of nature reserve systems originated in the application of R. H. MacArthur and E. O. Wilson's "theory of island biogeography" to fragmented landscapes. This theory postulates that the number of species on islands is the result of a dynamic equilibrium between the processes of extinction and immigration, such that as island size increases, extinction rate decreases, and as distance from the mainland increases, immigration rate decreases. As a result, a small island located far away from the mainland will achieve a smaller equilibrium number of species than a large island located near the mainland (**Figure 1**). This led Jared Diamond to suggest a series of geometric arrangements that compared preferred to less optimal distributions of patches of protected habitat (**Figure 2**).

The creation of a nature reserve may reduce or completely halt habitat conversion within its borders, but is not likely to prevent land use changes in surrounding areas. As time passes, nature reserves become habitat "islands," and may effectively

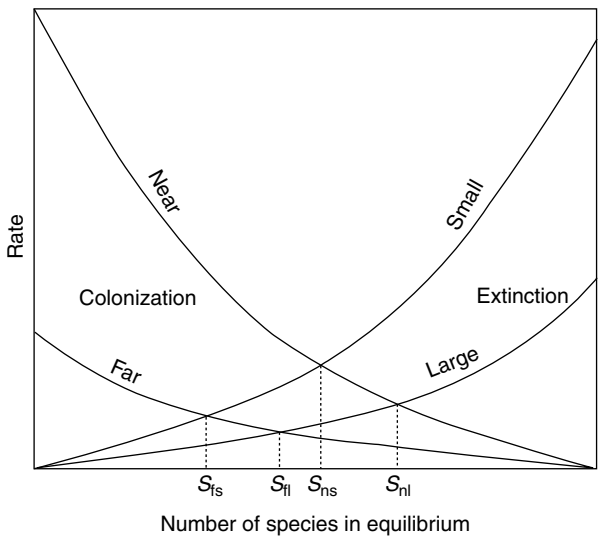


Figure 1 Theory of island biogeography predicts that the number of species in equilibrium for a small island located far away from the mainland is smaller than for a large island located near the mainland. Redrawn with permission from Wilcox BA (1980) Insular ecology and conservation. In: Soulé ME and Wilcox BA (eds.) *Conservation Biology: An Evolutionary-Ecological Perspective*, pp. 95–118. Sunderland, Massachusetts: Sinauer Associates.

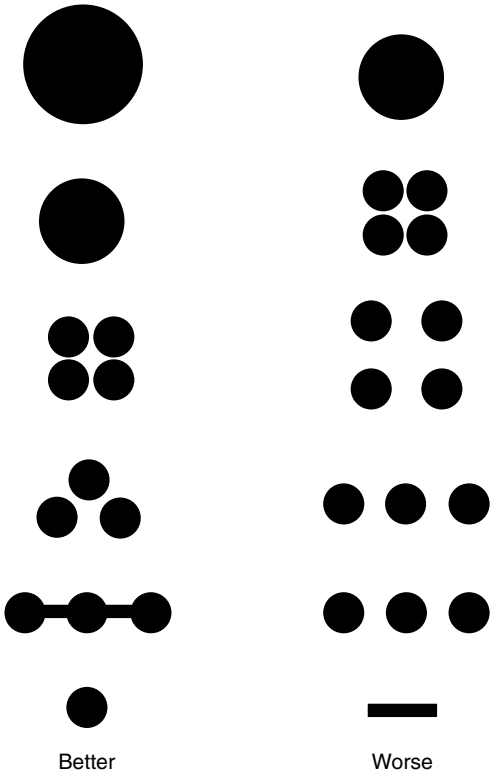


Figure 2 According to the theory of island biogeography, large, round, connected, or closely located protected area networks are better for assuring the persistence of a larger proportion of the species present in an area. Reprinted from Diamond JM (1975) The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation* 7: 129–146.

decrease in size while increasing in degree of isolation from the matrix of habitat that once included them. The theory of island biogeography predicts that the equilibrium number of species in habitat islands should decrease until either the effective size of the nature reserve stops decreasing or the effective distance from sources of immigrants stops increasing. Thus the effectiveness of many parks in the long-term conservation of biodiversity is compromised by phenomena such as edge effects, which effectively reduce the size of a patch. At the same time, conventional habitat fragmentation studies have recently recognized the importance of studying fragments within the larger landscape mosaic, paying attention to the vegetation types found between forest patches in the "matrix," or modified habitats surrounding the fragments. The influence of this matrix on animal and plant dispersal, diversity, abundance, and persistence can have a significant effect on the magnitude of edge effects and will simultaneously modify the effective dispersal rate of different species between patches (Ricketts, 2001).

The SLOSS Debate

With time, conservation biologists have become aware that the theory of island biogeography might offer additional insights to questions related to the design of nature reserve networks. For example, what sort of actions would the theory prescribe to minimize the rates of species loss? Alternatively, how can a reserve network maximize the number of species preserved? Is there an optimal shape for nature reserves, or can they be arranged in optimal geographic designs?

In attempting to answer some of these questions, conservation biologists initiated a controversy that came to be known as the single large or several small or SLOSS debate. The main point of contention was whether the conservation of biodiversity in a particular area was better served by one large reserve or by several small ones. Proponents of the "single large" approach argued that large reserves were able to house large population sizes, thus minimizing the risk of extinction of the species present. In small reserves, high extinction rates would result in changes of the species assemblages in the reserve and lead to habitats that were very different from those initially targeted for conservation. Proponents of the "several small" approach demonstrated that several small reserves whose combined size equaled that of one single large reserve could include a larger proportion of the species present in a given region. Furthermore, several small reserves generate a set of independent duplicate subpopulations, which decreases the likelihood of extinction of the population as a whole from diseases or catastrophes.

Spatial considerations also play a major role in the size of areas set aside for endangered species. Indeed, the problem of incomplete consideration of spatial scale was at the heart of much of the SLOSS debate. For example, consider a small town that decides to set aside an area of habitat as a nature reserve. At this spatial scale, there is little difference in the species composition of similar-sized patches; however, a relatively small increase in the size of the patch is likely to capture a significantly larger proportion of the local biodiversity. In contrast, as we increase the scale of parks to the state, national, or even semicontinental scale, similar small increases in the area preserved will not significantly affect the species composition in the patch but similarity between patches of habitat of similar sizes will decline. At this national or continental scale, it will be better to purchase, set aside, or otherwise preserve a variety of different patches of habitat. The importance of "beta diversity," a measure of biodiversity, which represents this variety and can be defined as the change in species composition from place to place, has been increasingly recognized in parallel with the appreciation of spatial scales. Thus, the answer to the SLOSS debate was seen to depend on spatial scale.

Spatial and temporal considerations of scale are also central to our understanding of how ecosystems respond to natural and anthropogenic disturbances. The relationship between the spatial scale of a disturbance, the frequency of a disturbance, and the length of time it takes to recover from a disturbance (the "disturbance regime") will determine how readily degraded land can be recovered and used to increase the effective size of a preserved area. In habitats that require disturbances such as fire to maintain all their natural components of biodiversity, such as the New Jersey pine barrens in the eastern US, the Fynbos of South Africa, or the Banksia woodlands in western Australia, successful management will have to mimic the scale and frequency of natural disturbances. Some preliminary synthesis suggests that the recovery time

from disturbance, scales roughly with the square root of the area disturbed.

Gap Analysis

Modern reserve design methods seek to maximize the representation of the components of biodiversity in protected area networks while minimizing the resources needed to do so. For this approach to be effective, comprehensive information is needed on the geographical distribution of species from a wide range of different taxa.

The first step in a reserve design process is to select a series of surrogate measures of biodiversity and assess the existing network of nature reserves for how well it samples this biodiversity. This allows managers and researchers to identify components of biodiversity that are unprotected or not well represented in the current reserve network. The spatial configuration of these missing components defines priorities for future protected area designation. This technique, known as *gap analysis*, was pioneered by Australian scientists in the mid-1970s and popularized by the US governmental agencies in the decades that followed, particularly Michael Scott and colleagues in Hawaii and Idaho.

Gap analyses typically focus on well-known species or groups of species, and assume that their presence in a nature reserve network can be used as an indicator of the presence of other components of biodiversity as well. Gap analyses may be carried out on the basis of climatic zones, landscape properties, or any subset of features in a geographical region and for any proportion of population size or geographical extent. In fact, given that for most regions of the world knowledge of the detailed distribution and abundance of species is quite limited (although camera traps have recently improved our ability to pick up rare and elusive species), the best available data may be a combination of abiotic environmental parameters that are known to correlate with the biological species of interest. These data can then be integrated into a program of biological surveys that constantly adds new information to the database.

Figure 3 depicts the steps in a gap analysis. This process frequently employs "remotely sensed" data obtained from satellites and requires the use of geographical information systems (GIS). Maps of the distribution and abundance of plants and animals, vegetation types, topography, climate, and soil properties (among other possible data sources) are combined to generate overlay maps that quantify landscape heterogeneity. A map that combines existing protected areas is contrasted with the target features overlay map and the presence of these attributes in the nature reserve system is assessed. Thus, a final map with the distribution of those attributes that are not included in nature reserves allows for the identification of gaps in protection.

The definition of what constitutes adequate representation is the next challenge. Assessors may be interested in ensuring that all species, communities, or habitats considered important are present at least once in a network of protected areas. After achieving this most basic objective, other criteria may be added that minimize costs or future threats, or that maximize the total area of the reserve system by considering alternative configurations of land. It may also be possible to select

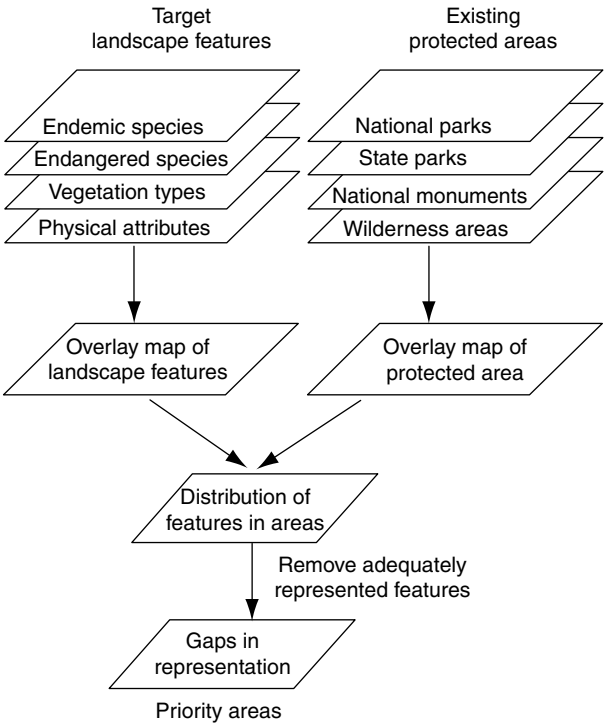


Figure 3 Schematic of the gap analysis process.

alternative configurations to reduce conflicts with alternative land uses for the area under consideration.

An important problem that arises is that even if a species is present in a protected area at the time of designation, this does not ensure that this species will persist in the long term. Furthermore, the presence of a set of target species does not ensure that their parasites, predators, prey, or mutualists are also included, or that other unrelated species will be present as well. The protection of a target species may not provide an adequate measure of the status of ecosystem services, such as the prevention of erosion or the pollination of crops. However, these are not necessarily weaknesses of the gap analysis process itself, but of the way in which it is implemented. The use of proxy species (indicator, umbrella, keystone, flagship; see Tim Caro's Conservation by Proxy) must thus be done with care, taking species interactions and food webs into account. Their use may be most appropriate under conditions of uncertainty when the number of species being protected is uncertain and the spatial scale is of intermediate size (Caro, 2010; see **Figure 4**).

Reserve Selection Algorithms

The costs of administering a reserve network increase proportionally to the area protected. Two fundamental concerns of protected area designers are to efficiently allocate available resources, and to minimize the area required to protect as many natural features as possible. The efficiency of a reserve network can be described by the efficiency (E) index

$$E = 1 - (x/t)$$

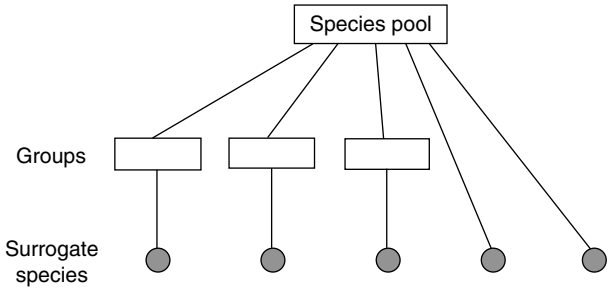


Figure 4 Use of surrogate species or species groups for conservation planning rests on the assumption that surrogates represent an entire pool of species or subsets of this pool. Reprinted from Wiens JA, Howard GD, Holthausen RS, *et al.* (2008) Using surrogate species and groups for conservation planning and management. *Bioscience* 58: 241–252, with permission from University of California press.

where x is the number or area of the sites required to achieve a particular representation target and t is the total number or area of sites to choose from. Thus, a network of protected areas is more efficient as the amount of land required to achieve the target decreases.

Reserve selection algorithms provide systematic means for designing efficient reserve networks. They follow a step-wise process for choosing a subset of sites from a list of proposed sites, such that the efficiency of the chosen subset is maximized. In other words, these algorithms allow for the selection of the minimum amount of land that is required to include all target species (or other attributes) at least once. There are three types of reserve selection algorithms. "Richness-based" algorithms begin the process by selecting the site containing the greatest number of species or attributes and sequentially adding those sites that include the greatest number of new species, until all species have been accounted for. "Rarity-based" algorithms select sites on the basis of rarity scores, favoring sites that are richest in species with restricted ranges. The third type of algorithm depends on linear-programming techniques, which sequentially search the database of available sites for replacement sites that increase the efficiency of the current set.

Although richness-based algorithms are the least efficient, they are the easiest to implement and have the advantage of including the principle of complementarity, i.e., at each step of the process, the selected site is most complementary in terms of the features that it adds to the set. Rarity-based algorithms tend to be more efficient than richness-based algorithms; by focusing on species with restricted ranges, more wide-ranging species tend to be included by default, and fewer sites are required to achieve fairly complete levels of coverage. Though linear-programming methods are ultimately the most efficient, they rely heavily on intensive computing and are prohibitive for very large data sets.

Reserve selection algorithms allow for the selection of efficient reserve networks, but focus only on the adequate representation of a set of species or habitats within the landscape. They do not provide answers to questions related to the size, shape, or number of protected areas that are required to ensure viable populations of different species, nor do they guarantee ecosystem functions and services. These more ambitious targets may only be achieved by expanding reserve networks as an interconnected system that allows the interaction of biological and physical components in a continuous landscape.

Corridors and Connectivity across Landscapes

Although overexploitation is responsible for the extinction of some species, the majority of species have declined because of the loss and fragmentation of their natural habitats, combined with the introduction of alien competitor, predator, pest, and pathogen species. It is these indirect effects of human development that have had the biggest impact on biodiversity. The crucial, and perhaps the only, way to minimize future impact is to reduce current and future rates of habitat loss and to explore ways of expanding current systems of reserves through the protection, restoration, or development of large-scale connections across the landscape.

Isolation of nature reserves and protected areas can often disrupt or break migration corridors that are used by seasonal residents of the park, and often increase human–wildlife conflict outside of reserve boundaries. The regular movements will range from the diurnal movement of snakes, birds, and many other organisms between feeding and nesting or brooding sites, to the larger annual migrations of large ungulates such as wildebeest and megaherbivores such as elephants. In some instances, it is possible to maintain needed connectivity between habitat fragments by protecting natural corridors or stepping stones of appropriate habitat. In other cases, natural habitat and stopover points may be lost and enhancement of these converted habitats through ecological restoration may be the only option. The interactions between protected areas and developing or developed areas of human use outside them are often subtle and complex. Increased isolation is likely to lead to changes in the flow of nutrients and pollutants into and out of the protected area. Habitat conversion in the surrounding matrix facilitates the invasion of alien species that interfere with the species that the reserve was designed to protect.

Corridors and connections in the landscape have two major functions at the species level. First, they permit regular daily or seasonal movements, helping to ensure that different sub-sections of larger populations may have access to all the resources they require, while also maintaining the potential for all individuals in the population to successfully interbreed. Second, connectivity facilitates the dispersal of animals from their place of birth to their adult home range where they breed. At the regional or landscape level, this latter function is usually the most important, justifying, for example, the achievement of connectivity between mountain ranges in the western US for the long-term conservation of fully functioning ecosystems.

The restoration of connectivity must occur at many scales. Most attention to date has been paid to local and regional connectivity as a way to redress the various forces threatening small populations. For example, amphibians must be able to move safely across a country lane during their annual migration to a breeding pond. However, many regional and inter-regional corridors have additional goals, such as accommodating the need for grizzly bears (*Ursus arctos*) to

disperse safely between the Canadian Rockies and the Northern Rockies of the US. In the face of global climate change and other major environmental changes, a substantial system of landscape connections is a major prerequisite for ensuring species persistence. Some species may be more suitable indicators of landscape connectivity than others, for example, wolves managed to initially recolonize Yellowstone National Park by dispersing down the chain of protected lands and National Forests before they were experimentally reintroduced into the park. Once they were formally reintroduced and monitored with radio-collars, they have provided vital information on remaining natural corridors that connect Yellowstone to other parks and protected areas.

Species and Populations

Although conservation at the ecosystem level is the most effective way of conserving biodiversity, a significant proportion of research within the discipline of conservation biology has focused on conservation at the population level. This emphasis on species and populations partly reflects the constraints contained within legal mechanisms to preserve biodiversity; for example, both the Endangered Species Act in the US and the global IUCN Red Lists of Threatened Species focus on conserving species and populations. However, the species also offers a relatively well-defined (although not unambiguous) biological unit of study. For example, trophic cascade effects from the loss of single species can occur particularly if the species lost are strongly interacting ones (e.g., fig trees considered to be keystone, and deer, known to be dominant). Conserving distinct populations of a species is also justifiable when behavioral diversity is taken into account (i.e., the loss of one population means the loss of an entire set of adaptive behaviors). Variety buffers the species "from the ups and downs" of life. Although it is understood that species diversity keeps ecosystems healthy, this approach emphasized that population diversity is also key. Research on several hundreds of discrete populations of salmon in Alaska illustrates how each population can do better or worse according to certain environmental and random conditions. At any one time, a high diversity will provide enough populations experiencing a productive time, thus compensating for those populations doing less well.

Finally, thinking about a population of populations or a metapopulation, has direct relevance to reserve planning when incorporating linkages that connect isolated fragments into the design.

The global biodiversity crisis is often most effectively communicated in terms of increasing extinction rates or numbers of threatened species. Therefore, awareness building and fund-raising tend to be more successful when focused on a species, rather than an ecosystem or a particular environmental problem. It is often easier to raise funds or public support for charismatic species, such as elephants (*Loxodonta* sp.), chimpanzees (*Pan troglodytes*), or giant pandas (*Ailuropoda melanoleuca*), than for a generic habitat, such as a tropical or bamboo forest. More pragmatically, it is often possible to use the requirements of a single species to maintain the viable complexities of entire biological communities or ecosystems. By focusing on either a target species that occupies a keystone position in an ecosystem or a species that requires large tracts of land to persist (often referred to as an "umbrella" species), many other additional species and their natural habitats are also protected. This is effectively the strength of the "jeopardy amendment" in the US Endangered Species Act, which provides legislation that is designed to conserve habitat in order to ensure the viability of one specific species – as occurred with the northern spotted owl (*Strix occidentalis caurina*) and "old-growth forest" in the north-western US. Similar species-centered approaches led to the creation of the Cosk-somb Basin jaguar (*Panthera onca*) preserve in Belize, the Arabian oryx (*Oryx leucoryx*) sanctuary in Oman, the Laguna Brava vicuña (*Vicugna vicugna*) reserve in Argentina, the Wolong and Wanglang nature reserves for panda habitat protection in China, and the significant expansion of the El Guácharo national park in Venezuela to secure feeding habitat for a large oilbird (*Steatornis caripensis*) colony.

Population Viability Analysis

The central theme of population ecology is represented in a large canon of work that examines how populations respond to intrinsic and extrinsic factors and how these affect their long-term dynamics and persistence. Although deterministic, stochastic, environmental, demographic, and genetic processes can all have complex and interacting effects on extinction risk, these processes tend to operate in a hierarchical fashion determined by the size of the remnant population. Furthermore, habitat destruction and overexploitation can exacerbate intrinsic risk factors and accelerate population declines. A population viability analysis (PVA) explores the interactions between different known factors and assesses the likelihood that a population will become extinct within a specified time frame and under particular circumstances. PVAs generally require the use of computer simulations and simple mathematical models. The work of M. A. Burgman, S. Ferson, and H. R. Akçakaya (1993) provides the definitive introduction to this area.

Population viability analyses are a valuable tool for conservation biologists concerned with studying the long-term persistence of particular species or populations. In the early 1980s, Mark Shaffer introduced the concept of the minimum viable population (MVP), as the minimum size required for an isolated population to have a 99% chance of remaining extant for the following 1000 years in a particular habitat. Shaffer's work provided a quantitative framework for thinking about species conservation objectives and the timescales within which human actions should be considered. While focusing on the MVP concept, researchers began to realize that the key to conservation at the population level was understanding the impact of different stochastic and deterministic processes that determine extinction risk. By examining the sensitivity of the population to changes in these processes, it was argued that management actions could be taken to increase population persistence. Although there was an unfortunate tendency for many early analyses of viability to emphasize captive breeding as a means of enhancing persistence, this is now perceived as an option of last resort. Moreover, the notion of an MVP was progressively assimilated into

a series of techniques that are known as "population viability analysis."

PVAs focus on individual species, but they can also examine the potential consequences of different ecological factors, including interactions with other species, rates of habitat conversion, or management interventions that may be taken to improve population persistence. When integrated sensibly into wildlife or endangered species monitoring programs, PVA can be a powerful tool to assist in decision making within an adaptive management framework, but to be complete it requires years of data. The degree of complexity of a PVA will depend on the number of elements that a manager decides to consider, yet more complex models do not necessarily mean more useful results. Understanding the model structure and its sensitivity to varying parameter values may be more important than the modeling process itself. Numerous PVA software packages are available. Unfortunately, the different assumptions hidden within different packages mean that they can produce very different results from the same set of data. Thus, it is crucial to bear in mind that PVAs should not be used to estimate extinction risks *per se*. Instead, they are most useful for ranking different management options to determine those that increase the likelihood of persistence and that can be implemented at lower costs. They are also invaluable for indicating areas of ignorance about a species and thus focusing future research. The precautionary principle indicates that ignorance should always imply increased protection until a more complete set of demographic data on the species is available.

The Demography of Small and Declining Populations

The recognition that current losses of biodiversity constitute a global crisis has led to a growing interest in the problems faced by small populations. By definition, all populations that go extinct first become small. Research in this area has emphasized the positive feedback between a number of mechanisms, each capable of increasing extinction risk as populations further decrease in size. The four principal mechanisms are (1) chance demographic events (demographic stochasticity), (2) environmental stochasticity, (3) inbreeding depression, and (4) allee effects due principally to the breakdown of the social benefits accrued by living in groups. Unfortunately, in the 1980s this "cult of the extinction vortex" led to an emphasis on crisis management, whereas relatively less attention was focused on understanding the processes that initially caused the decline in a previously viable population. In an insightful review of population-level approaches to conservation biology, Graeme Caughley (1994) proposed that the discipline has developed along two relatively independent but complementary lines: the "small population paradigm," which has received more emphasis and addresses problems related to the persistence of small populations, and the "declining population paradigm," which considers the causes of population declines and how to reverse them.

In broad terms, "small population" conservation biologists have focused on finding ways of increasing the size of small populations and achieving the maintenance of a regular flow of individuals between fragmented populations. The application of these principles to the conservation of endangered species in the wild has been limited, but they have played a major role in influencing captive breeding practices, while providing some insights into the theory of how nature reserve systems might be designed.

In contrast, an increasing number of conservation biologists have focused their research on identifying the causes that lead to declines in species abundance or the contraction of the range of a species. The primary forces operating here are habitat loss and fragmentation, although threats posed by alien species, over-exploitation, and pollution are also important, as are species traits that may predispose them to extinction such as rarity, island endemism, and ecological specialization. By examining populations at an earlier stage in their decline, it is hoped that strategies may be found for halting and reversing the decline. This paradigm is concerned with prescribing early, preventative actions that require minimal intervention that may be taken before a population becomes seriously endangered. At present, this approach has not succeeded in generating elegant theoretical principles to guide management decisions. This is mainly because population declines can be caused by numerous factors operating simultaneously on a community of populations. Not surprisingly, these effects are difficult to quantify. The key to future success in this area is carefully designed experimental and monitoring programs that systematically seek causes for population declines.

The most likely future for these two paradigms is eventual fusion. Clearly, both approaches are useful and contribute complementary insights to the advancement of the discipline.

Individuals and Genes

All forms of life carry genetic information. In the long term, the capacity of any species to adapt to anthropogenic and nonanthropogenic changes in the environment will be related to the existence of genetic variation. *In situ* conservation programs are almost without exception the best option for conserving the genetic diversity that is a component of biodiversity. The protection of an ecosystem and its ecological processes provides the necessary conditions for evolution and natural selection to proceed. Additionally, protecting species in the wild is much more cost-effective than maintaining them in captivity. In the cases of some species, however, such as varieties of domesticated plants and animals and highly endangered species, *ex situ* conservation technologies are required. In these cases, the maintenance of genetic variation becomes a central topic of concern, especially if an endangered species is taken into captivity for future reintroduction into the wild.

Measuring Genetic Variability

Genetic information is encoded in genes, which are mainly composed of long chains of deoxyribonucleic acid, or DNA. Genes aggregate in the cells of every organism to form chromosomes. Variation in genetic information can be measured at any step of the biological process that transforms

the sequence encoded in genes into its protein products. This information can then be used to quantify differences between organisms, populations, species, or any other taxa.

Methods for measuring genetic diversity focus on the variation of "markers," much in the way that gap analysis uses species or topographic variables to characterize the degree of spatial heterogeneity in a landscape. These markers are assumed to correlate in a broad sense with the genome as a whole; the more variable a population is at marker loci, the more variable it is overall. In the past, scientists were limited to examining indirect or "phenotypic" measures of genetic diversity through markers such as variable external morphology or coloration, intracellular concentration of secondary compounds such as phenolics, cell pigments, or hydrocarbons, and variable protein products called "allozymes," which may be visualized using electrophoresis. In recent years, however, biologists have developed techniques for examining the DNA itself. This permits the use of markers such as the presence or frequency of specific base-pair sequences in DNA molecules. DNA fingerprinting, for example, allows the characterization of differences between individuals in a population on the basis of patterns in the variation of minisatellites, or relatively short (10–50 base-pairs long) repeat sequences that are distributed in tandem arrays throughout the nuclear genome. Variation in organelle DNA (such as mitochondria or chloroplasts) may also be examined. Different versions of these techniques are used to assay genetic differences within and between populations, species, or other taxonomic levels.

DNA fingerprinting is being increasingly used in the field of wildlife forensics to trace, for example, the source of elephant ivory seized on the black market.

The Role of Zoos, Aquaria, Botanical Gardens, and Gene Banks

In extreme cases, such as when the last few remaining individuals of a species are under imminent threat in the wild, or when a wild variety of a particular crop is being rapidly replaced by modern monocultures, the maintenance of populations in captivity may be the only short-term action to prevent their extinction. Zoos have played an important role in preventing the extinction of the Arabian oryx, the peregrine falcon (*Falco peregrinus*), and the Guam rail (*Rallus owstoni*), but in all cases the captive facilities were used to buy time while the agents of decline were being dealt with. Gene banks may be the last resource for hundreds of threatened varieties of domestic animals and plants. The Royal Botanical Gardens in Kew (UK) hold about 38,000 species in their collection; Brazil and Colombia are the only countries in the world that have more species within their territories.

The primary functions of these captive facilities are not to save endangered species, however, to promote education and research; for example, many captive facilities now maintain web-based repositories of data (e.g., eTaxonomy projects) updateable in real time. Kew Gardens' e-monocot site, for which Kew employed the emerging science of Biodiversity Informatics, is a database of the world's monocot plants and facilitates their identification and classification. While conservation objectives are clearly most effectively and efficiently achieved by *in situ* rather than by *ex situ* approaches, zoos, aquaria, and botanical gardens provide visitors with an opportunity for close interaction with a wide variety of plants and animals that are otherwise out of reach for most people. Zoos attract millions of people worldwide and are good facilities for educating them about the challenges and opportunities created by the present biodiversity crisis. They are also potentially important facilities for raising funds that can then be invested in conserving habitat in the regions where the animals and plants on display originally came from.

Future Challenges

The future of conservation biology lies in teamwork, bold vision, and allocation of more time to action and less to assessment. The complexity of issues surrounding the conservation of biodiversity requires a trans-disciplinary approach that integrates knowledge and expertise from both the natural and social sciences. Conservation biology will also become increasingly cross-institutional as governments, academia, and nongovernmental organizations work together on the same issues but with complementary approaches and points of view.

Of the 47,677 species assessed in 2009 by the IUCN, nearly 40% were found to be threatened with extinction. Meanwhile, conservation biology is becoming less species-oriented, more holistic, and systems-oriented. The expansion of protection from species to ecosystem springs in part form a transition to thinking about systems as dynamic. A new Red List aimed at ecosystems rather than species was recently conceived (Rodriguez *et al.*, 2011; **Figure 5**).

In recognizing the dynamic nature of ecosystems, the discipline is moving away from arbitrary numbers – both minimum and maximum. For example, with regards to MVP, there is now consensus that no single rule of thumb exists for a safe minimum size of populations that differ in the levels of environmental stochasticity that they confront. Whereas, capping populations of large mammals (elephants, wild horses) at specific numbers on the basis of human-defined carrying capacities (based on, e.g., viewing conversion by elephants of woodland into bush as disturbance rather than ecosystem engineering) is increasingly seen as an old-fashioned management strategy.

One requirement of conservation biology in the future will be the design of biodiversity management programs and decision-making mechanisms that incorporate, rather than reject, uncertainty. For example, if we were to wait until all the desired data on a country's corridors were collected and compiled before protecting them, we may lose vital, quickly-disappearing corridor routes. Rapid assessment exercises combined with local knowledge ("citizen science") may therefore, in some instances, do more for conservation than detailed priority-setting exercises. We must ask "how much evidence is enough for "evidence-based conservation?," and also "how can we balance concern over collecting enough data with ensuring that we do not duplicate efforts?"

Although necessary, the establishment of parks and reserves may become viewed as "desperate action" by which we cannot ensure by law that what happens on one side of the

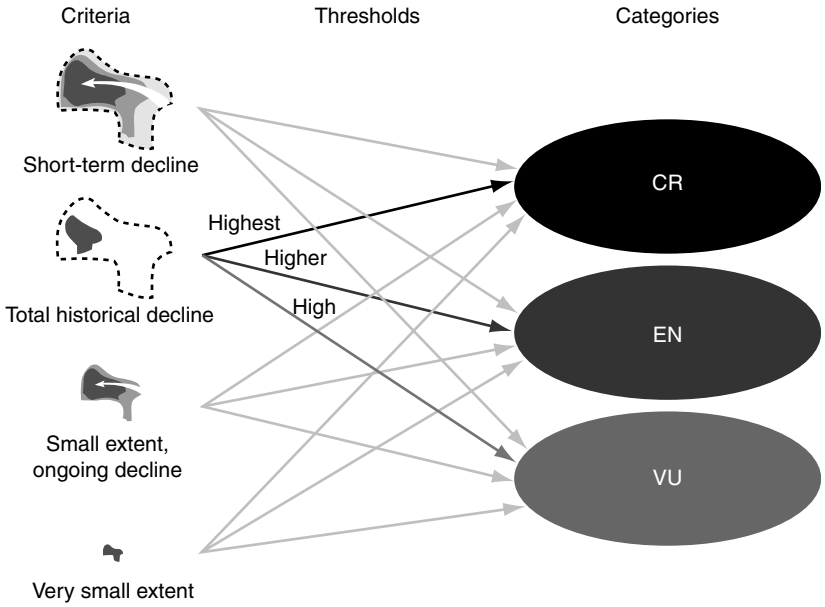


Figure 5 Ecosystem extinction-risk assessment entails evaluating data on one or more quantitative proxy risk indicators (criteria) against thresholds to assign a threat category (critically endangered (CR), endangered (EN), or vulnerable (VU)) to the ecosystem. Reprinted from Rodríguez JP, Rodríguez-Clark KM, Baillie JEM, *et al.* (2011) Establishing IUCN Red List criteria for threatened ecosystems. *Conservation Biology* 25: 21–29.

line does not happen on the other (Adams, 2006). With growing agricultural expansion, pollution, overharvesting and climate change, many of these protected areas are failing in their stated aims, losing these target species or habitats (Graham *et al.*, 2008; Mora, 2008; Craigie *et al.*, 2010). Coupled with growing pressures from expanding populations and the need for more agricultural land, many protected areas are going to be under pressure for conversion and, with their conservation targets in decline or lost it will be hard to justify their existence. Lastly, the development of roads and more optimistically railways will significantly improve the infrastructure of many Southern Nations, particularly in sub-Saharan Africa and South America; however, these routes will open up huge areas of wilderness for exploitation, while simultaneously allowing access to invasive alien plant and animal species as well pathogens. How can we strengthen laws and economic arguments that recognize that natural ecosystems require protection while at the same time moving away from fortress conservation to multi-use management schemes such as biosphere reserves and transboundary parks? Multiple-use zones such as the ones around Masoala National Park in Madagascar have worked well to buffer the park and support local livelihoods. Concurrently, the biodiversity value of plantations and agro-forests is increasingly acknowledged.

There is better recognition of the need to more accurately map the spatial distribution of the natural ranges of species and to monitor the abundance of as many species as possible. This will allow us to more accurately assess how species will respond to the twin threats of climate change and habitat loss. Simultaneously, the authors are concerned that the interest in climate change has detracted from habitat loss as the major threat to biodiversity. The impact of climate change over the next 25 years is most likely to be strongest in polar and temperate regions, there will be less of an impact in the Tropics and indeed, the loss of habitat in the tropics due to agricultural expansion may leave very little biodiversity to respond to climate change (Jetz *et al.*, 2007).

New conservation biology books will be aimed not only at intellectual communities but at laymen; notable recent ones include Jonathan S. Adams', *The Future of the Wild: Radical Conservation for a Crowded World*, which urges combining science with traditional knowledge and economic sense and "saving some of everything" in the spirit of thinking big; and *Conservation Biology for All*, available online (http://www.dbs.nus.edu.sg/staff/sodhi.htm).

The conservation biologist's toolbox is growing to include emerging technologies such as camera traps and established ones such as cellular phones; and also stable isotopes; genetics and genomics, DNA fingerprinting; modeling; Bayesian inference; geographic information systems and remote sensing (**Figure 6**).

Four important recent developments that will help buttress the efforts of conservation biologists include the Millenium Ecosystem Assessment, Convention on Migratory Species, the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES), and Reducing Emissions from Deforestation and Degradation (REDD). The authors highlight IPBES and REDD here. Established in 2010 (the International Year of Biodiversity), IPBES is aimed at peer reviewing biodiversity research to make certain that governments receive the most up to date information and evidence-based recommendations. IPBES has been dubbed the "IPCC for Nature" as it bears resemblance to the Intergovernmental Panel on Climate Change (IPCC), a global organization that has helped raise awareness and promote action on climate change. The IPCC, although a good model, has not been free of flaws and criticisms, but it has succeeded in identifying issues on which there is consensus (for example, that much of the warming

Mobile Phones: Readily accessible and affordable, mobile phone use has exploded in developing countries and improved voice communications. Applications in field biology range from tracking animals to providing early warning of environmental crises, facilitating rapid responses by rangers, park managers and communities to threats such as poaching, fire, and encroaching wildlife (Banks and Burge, 2004). New devices combining GPS and mobile phone technology are leading to low-cost alternatives for animal tracking equipment. See: http://www.kiwanja.net/wildlive!.htm

Camera Traps: Self-contained surveying tool that provides data on the presence and movement of elusive and rare wildlife species. New cameras capture video, which can provide clues about the behavior of such species. Combined with computer software, individuals can be identified enabling the use of capture–recapture analysis to estimate species abundance and density.

Global Positioning System and Geographic Information Systems: GIS is used to process geo-referenced data from a GPS (hand-held units and wildlife collars) and integrate it into a map project to plot species movements, home ranges and habitat use. Combined with modeling techniques such as habitat suitability and CorridorDesigner, researchers can model species distributions and design linkages between habitat patches. Other uses include anti-poaching and land-use planning.

Stable Isotopes Analysis: The abundances of stable isotopes, atoms that do not undergo radioactive decay found in most elements, are used as a recorder to reconstruct ecological processes (e.g., climate from tree rings) and trace ecological activities (elephant migration routes and dietary patterns from stable isotopes found in teeth and hair). http://isoscapes.org

Genotyping: New molecular techniques may offer a way to obtain more reliable estimates of population size. Animal faeces are typed for microsatellite loci using genotyping. Mark-recapture models can then be used to analyze faecal genotyping data treating the same multilocus genotype as a recapture.

Mathematical Modeling: A variety of models of harvesting, predator–prey and host–parasite dynamics, and optimization, are widely used in conservation ecology and are becoming more accessible to field biologists (see Kokko, 2007).

Network Analysis: Focuses on patterns of relations within a system. A popular application is network representations of ecological interactions such as food-webs to analyze diversity, stability and response to perturbations; another is spatial networks to represent species' dispersal pathways between fragments. See Cumming *et al.*, 2010.

Web-based Data Repositories: Ranging from species cataloging projects to compilations of images and references, web based databases are increasingly a part of conservation projects aimed at making biological data readily and freely accessible. For a comprehensive annotated list of biodiversity databases, see http://www.biodiverselife.com/biodiversitydata.html

Media Training: Prepares biologists for interactions with journalists by helping them to communicate research findings accurately, often in everyday language and in a way that grabs attention.

Figure 6 The conservation biologist's expanding toolbox. http://www.corridordesign.org/

observed over the last five decades can be attributed to human activities). Given that 20% of global greenhouse gas emissions is due to deforestation, REDD was established to provide incentives (in the form of forest carbon stocks) for developing countries to reduce emissions. Some scientists (e.g., Sutherland *et al.*, 2010) have expressed concern over REDD's emphasis on forests, which could possibly undervalue the high carbon sequestration potential in non-forested ecosystems such as savannas, mangroves, and peat.

Although our understanding of ecosystem function and human dependence on nature's services has greatly increased, it is still far from perfect. We still do not have the equivalent to the chemists' Michelis–Menten relationships that allow us to say what proportion of habitat loss will lead to a 50% decline in the rate at which a specific ecosystem service is delivered. Although there are logical reasons to assume that ecosystem services may be lost in a consistent order as species from sequentially lower trophic levels go extinct, there have been very

few attempts to systematically organize studies of ecosystem services along these lines (Dobson *et al.*, 2006). As in the case of global climate change, we might never understand every detail of the numerous processes that influence biodiversity and the delivery of ecosystem services, but the need to understand their relationship within a quantitative analytical framework is increasingly urgent. This ignorance cannot remain an excuse for inaction.

Ultimately, conservation of the world's remaining biodiversity will require the adoption of an adaptive management framework that consistently embraces the precautionary principle in combination with the best science available to decide on issues of land stewardship. One of the most difficult issues may be the revision of current resource consumption patterns, particularly in the developed world, and the design of new paradigms and technologies to attend to the needs of a growing and increasingly hungry global population. This retains its status as the major driver of biodiversity decline and the reason for the conservation biologist's task becoming a juggling act of science and advocacy. After particular actions are taken, their success may be monitored and adjusted in response to changing conditions, which arise because of either abrupt, unanticipated events or slowly evolving processes. With time, our understanding of biodiversity and our management decisions will both improve, generating creative solutions to reduce the impact of humans on Earth's systems while assuring the quality of life of all those who inhabit the planet.

See also: Conservation Movement, Historical. *In Situ*, *Ex Situ* Conservation. Island Biogeography. Keystone Species. Natural Reserves and Preserves. Population Viability Analysis. Restoration of Biodiversity, Overview

References

Adams JS (2006) *The Future of the Wild: Radical Conservation for a Crowded World*. Boston: Beacon Press.

Avise JC and Hamrick JL (eds.) (1996) *Conservation Genetics: Case Histories from Nature*. New York: Chapman and Hall.

Banks K and Burge R (2004) *Mobile Phones: An Appropriate Tool for Conservation And Development?* Cambridge, UK: Fauna and Flora International. Available online at: http://www.kiwanja.net/ICT_Report.pdf.

Burgman MA, Ferson S, and Akçakaya HR (1993) *Risk Assessment in Conservation Biology*. London: Chapman and Hall.

Caro T (2010) *Conservation by Proxy: Indicator, Umbrella, Keystone, Flagship, and Other Surrogate Species*. Washington DC: Island Press.

Caughley G (1994) Directions in conservation biology. *Journal of Animal Ecology* 63: 215–244.

Craigie ID, Baillie JEM, Balmford A, *et al.* (2010) Large mammal population declines in Africa's protected areas. *Biological Conservation* http://dx.doi.org/10.1016/j.biocon.2010.06.007.

Cumming GS, Bodin O, Ernstson H, *et al.* (2010) Network analysis in conservation biogeography: challenges and opportunities. *Diversity and Distributions* 16: 414–425.

Csuti B, Polasky S, Williams PH, *et al.* (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80: 83–97.

Dobson AP (1995) *Conservation and Biodiversity*. New York: Scientific American Library, Freeman.

Dobson AP, Lodge DM, Alder J, *et al.* (2006) "Habitat loss, trophic collapse and the decline of ecosystem services." *Ecology* 87: 1915–1924.

Graham NAJ, McClanahan TR, MacNeil MA, *et al.* (2008) Climate warming, marine protected areas and the ocean-scale integrity of coral reef ecosystems. *PLOS ONE* 3: 9.

Jetz W, Wilcove DS, and Dobson AP (2007) "Projected impacts of climate and land-use change on the global diversity of birds." *PLOS Biology* 5(6): e157.http://dx.doi.org/10.1371/journal.pbio.0050157.

Kokko H (2007) *Modeling for Field Biologists and Other Interesting People*. Cambridge: Cambridge University Press.

Margules C, Higgs AJ, and Rafe RW (1982) Modern biogeographic theory: Are there any lessons for nature reserve design? *Biological Conservation* 24: 115–128.

Meffe GK and Carroll CR (eds.) (1997) *Principles of Conservation Biology*. Sunderland, Massachusetts: Sinauer Associates.

Mora C (2008) A clear human footprint in the coral reefs of the Caribbean. *Proceedings of the Royal Society B: Biological Sciences* 275: 767–773.

Pressey RL, Humphries CJ, Margules CR, *et al.* (1993) Beyond opportunism: Key principles for systematic reserve selection, Trends in Ecology. *Evolution* 8(4): 124–128.

Ricketts TH (2001) The matrix matters: Effective isolation in fragmented landscapes. *American Naturalist* 158: 87–99.

Rodríguez JP, Rodríguez-Clark KM, Baillie JEM, *et al.* (2011) Establishing IUCN Red List criteria for threatened ecosystems. *Conservation Biology* 25: 21–29.

Scott JM, Davis F, Csuti B, *et al.* (1993) Gap analysis: A geographic approach to protection of biological diversity. *Wildlife Monographs* 123: 1–41.

Sodhi NS and Ehrlich PR (eds.) (2010) *Conservation Biology for All*. Oxford: Oxford University Press.

Soulé ME and Terborgh J (eds.) (1999) *Continental Conservation: Scientific Foundations of Regional Reserve Networks*. Washington, DC: Island Press.

Sutherland WJ, Clout M, Cote IM, *et al.* (2010) A horizon scan of global conservation issues. *Trend in Ecology and Evolution* 25(1): 1–7.

The Human Footprint Project. Available online at: http://www.wcs.org/humanfootprint/

United Nations Environment Program (UNEP) (1995) *Global Biodiversity Assessment*. Cambridge, UK: Cambridge University Press.

Wiens JA, Howard GD, Holthausen RS, *et al.* (2008) Using surrogate species and groups for conservation planning and management. *Bioscience* 58: 241–252.

Wilson EO (2002) *The Future of Life*. Boston: Little, Brown.

Conservation Efforts, Contemporary

Kristiina A Vogt, Jason J Scullion, Lloyd L Nackley, and Maura Shelton, University of Washington, Seattle, WA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Vogt, K., Schmitz, O., Beard, K., O'Hara, J., Booth, M., volume 1, pp 865–881, © 2001, Elsevier Inc.

Glossary

Ecological integrity The balanced state of an ecosystem under normal environmental conditions, including the capacity of an ecosystem to absorb disruption and also to recover from it.

Ecosystem engineers Animals (e.g., beavers) that modify an ecosystem by physically changing it (e.g., building dams) or by being present in such a great abundance (e.g., large grazing animals) that they have a disproportionately large impact on the ecosystem functions.

Ecosystem management Approach to management that attempts to maintain high ecosystem integrity while providing the services, uses, values, and products from that system for the long term. It explicitly integrates social, economic, and natural system sustainability.

Invasive species Native and exotic species that are capable of displacing indigenous species or spreading into habitats where they were not common previously.

Keystone species Species whose presence and abundance controls the integrity of a community or ecosystem and allows that system to persist within its natural range of environmental conditions.

Legacy ecosystems Ecosystems that carry an intact "memory" of a change in some physical, chemical, or structural attribute or a change in species composition that modifies the resistance and resilience characteristics of the ecosystem.

Resilience Rate at which an ecosystem is able to recover to conditions similar to those that existed prior to the imposition of a disturbance.

Restoration Reestablishment of similar structures and functions of an ecosystem or parts of an ecosystem that are no longer present because of past land uses or disturbances.

Sustainability State that defines the biogeophysical, economic, social, cultural, and political thresholds in between which, it is acceptable to continue to use and obtain services or products from a given piece of land. This definition requires all five factors to be considered when determining whether future generations will be able to acquire the same natural resources as the current generation.

Sustainable development Meeting human needs without damaging the ecosystems that produces these resources and at the same time providing equitable distribution and access to these resources around the world.

System functions Processes that occur in an ecosystem that are typically measured as changes in the carbon or nutrient cycles regulated by decomposers, consumers, or primary producers.

Umbrella species Species that have either large habitat needs or have other requirements whose conservation results in many other species being conserved at the ecosystem or landscape level.

Introduction: New Concepts in Conservation

The discipline of conservation biology arose from the need to apply science to the protection of species that are threatened or on the brink of extinction. In its early years, biologists and ecologists had to worry about refining the science of protecting endangered species on particular tracts of land. This resulted in a conservation science which was very species focused. Nowadays, conservation biology has to resolve many conflicting social, ecological, and economic demands that are simultaneously being placed on ecosystems services provided by particular tracts of land. Saving endangered species is typically only one component of a larger task to protect the entire biotic community (biodiversity) that make up the whole ecosystem. This ecosystem perspective has placed an emphasis on putting into practice many of the holistically based philosophical ideas that were previously articulated by John Muir, Aldo Leopold, and other visionary thinkers. The mounting worldwide conflicts over the conservation of biodiversity and other uses of the same land base make it imperative to refine and test new models of conservation that are ecosystem based, socially applicable, and culturally compatible. The following example regarding tropical conservation illustrates the contemporary efforts occurring at the ecosystem level that address the human environment as well.

As documented by the United Nations, fully one-half of the world's biodiversity is found in tropical ecosystems where conflicts over resource uses are particularly intense. At the same time, the human population growth rates in these areas are double the world average placing increased pressures on consuming finite bio-resources and their ecosystem services. Such prodigious growth typically leads to a high level of poverty that brings conflicts over resource use and their protection particularly polarizing for resource managers and policymakers who are unable to satisfy all the stakeholders simultaneously. In these circumstances, even the best scientific knowledge about the ecology of groups of species will not provide the needed insight to protect biodiversity, since human demands on the natural resources will constrain the application of science to conservation. Therefore, the

traditional model of conservation, whereby scientists continually refine and apply knowledge to the cause of species preservation, will be insufficient to meet the challenges of the future.

This knowledge has made us realize that the use of species demographic data is generally inadequate by itself for gauging conservation success and therefore for maintaining a species in the landscape. Species demographic data are effective in conservation when the species are keystone or umbrella species, or when they perform the ecosystem engineering functions. Even then, knowledge of changes in population dynamics of a species does not necessarily provide a mechanistic understanding of the role that species play in the ecosystems. But when species data can be linked more effectively through geographical information systems and comparative global data bases, species data can be combined with other aspects (ecological, hydrological, societal etc.) to display dynamic interactions. Conservation has thus been moving towards holistic analyses of systems (i.e., ecosystem management) in which factors other than species biology dictate the approaches used to maintain species in their habitats. A holistic analysis also recognizes that humans are an integral part of the system and not external to conservation projects.

Managing the impacts of human behavior in conservation landscapes continues to be the most challenging and is being aggravated by increased societal and natural environment vulnerability to climate change. This challenge has stimulated the development of many new approaches and tools in conservation. Five major philosophical changes are commonly mentioned while discussing the need to adapt old models and to find new models for conservation:

- First, there is the recognition that a focus only on species is inadequate to conserve species within habitats in perpetuity.
- Second, there is the realization that the human dimensional aspects of ecosystems cannot be ignored if conservation efforts are to succeed.
- Third, ecosystems are dynamic, so the maintenance of a particular type of habitat in the landscape will require several examples of that habitat existing elsewhere within that landscape but at different successional stages. This approach will require a landscape that is sufficiently large to include these stages. If the appropriate conservation area has been selected, when a particular habitat is lost at one location, its replacement will be developing at the same time in another part of the landscape.
- Fourth, conservation approaches will have to incorporate the tools of restoration ecology to re-establish the structural and functional aspects of native habitats required to maintain species. The restoration of damaged habitats is central to the conservation practice since centuries of human modification of the natural environment have resulted in an inadequate habitat area or have significantly modified ecosystems.
- Fifth, disturbances (e.g., hurricanes, and droughts or floods associated with El Niño) at long time-scales can eliminate or modify the habitat set aside in nature reserves so that the species being conserved are unable to survive in that diminished or altered habitat.

These five changes in philosophy are causing a reassessment of how conservation biologists approach the practice of conservation. These changes are partially being driven by new scientific understandings that have been integrated into conservation during the last 20 years. In this article, we present a case for the need to increase the acceptance of these concepts so that conservation practices can become more effective in human dominated landscapes. These five major philosophical shifts have already changed the practice and implementation of conservation projects (the most common changes are summarized in **Table 1**). These shifts will be discussed in greater detail in the next sections because they are critical for increasing conservation efficacy (the reader is encouraged to consult the Reference list for further readings on these topics). The Section, Analysis of Evolving Thinking in Conservation Biology will present an analysis of two critical drivers contributing to our evolving thinking in how to practice conservation. The section, Tools for Managing and Conserving Species will present a brief discussion of the contemporary tools and approaches used in conservation in the context of these philosophical changes. The Sections, Conservation in Human Altered Landscapes and Future Directions and Emerging Challenges in Conservation will provide brief discussions of the continuing difficulty for integrating the social and natural environments in conservation projects and the emerging challenge posed by climate change.

The past difficulties experienced in implementing conservation in human landscapes has clearly shown that sustainable economies must be created at the local level for conservation projects to be successful. This recognition influenced the explicit linking of conservation with international sustainable development projects, especially in the humid tropics. Projects that meld these dual objectives are commonly called Integrated Conservation and Development Projects (ICDPs). These projects vary in scale and scope but typically include one or more of the following topics as a central theme: biosphere reserves, eco-tourism ventures, nontimber forest product harvesting, and regional land use plans. Such projects are attempting to address the impacts of higher human population growth rates that result in the development of economies which in turn demand increasingly greater use of natural resources to produce viable communities. The ultimate result of all of these changes is that humans are leaving a much larger "ecological footprint" on the land-base and making it more difficult to practice conservation. This is especially relevant in landscapes experiencing higher deforestation rates.

One aspect of a higher deforestation concomitant with larger ecological footprint has been a reduction in ecological legacies. The influence of ecological legacies on species survival is often not accounted for in current conservation management, because many policy decisions have been based on science collected in non-legacy ecosystems. However, it is important for conservation management to account for ecological legacies, because of their influence on ecosystem function, including how ecosystems respond to disturbance events. Contemporary conservation efforts can incorporate ecological legacies into management planning by identifying and conserving factors that contribute to the formation of

Table 1 Summary of pre-1990 and post-1990 practices used to implement conservation efforts

	Pre-1990 practices	*Post-1990 practices*
Conservation focus	Single species based approach. *Ex situ* conservation for single species (e.g., zoos, expensive reintroduction programs, captive breeding programs).	Umbrella species approach. Ecosystem based. *In situ* conservation model preferred. Restoration of species and ecosystems, elimination of invasive species linked to conservation failures.
Integration of humans and human impacts on ecosystems in conservation	Protected areas management or reserves could be designed to exclude humans.	Conservation is human ecosystem based – explicit integration of people as primary drivers of the success of protected areas or reserve management and a need for people to access conservation areas.
	Conservation practices strongly based on managing spatial distribution of ecosystems or habitat scales in landscapes since the major human impact has been fragmentation and loss of ecosystems.	Recognition of the temporal scales of human land-use activities and how they produce legacy ecosystems and landscapes that control the success of protected areas or reserves needed in conservation (states that few nonhuman-impacted ecosystems exist).
Human values	Species values dominant and used to drive conservation practices.	Recognize species is one of multiple values existing in a system and a species focus may miss the ''real'' driving variables determining whether an established protected area or reserve will be successful. Incorporation of and determining trade-offs of all values desired out of a land as part of conservation.
	Incorporate short-frequency disturbances.	Inclusion of decade-scale disturbances (e.g., hurricanes) for species conservation.
Disturbances	Conservation plan to exclude or limit human extraction of resources from nature reserves. Humans classified as disturbance agents who are negative to conservation efforts.	Conservation plan includes human extraction of resources from nature reserves.
		Explicitly integrates invasive species as a major disturbance factor affecting conservation success.
	Reserve design and size varied with territory needs of each species. Landscape matrix analysis recognizing influences outside of habitat fragment will impact conservation.	Reserve design and size based on the ecosystem and landscape and not on species territory needs. Recognition that linking habitat fragments with corridors is inadequate.
Spatial scale of analysis	Identify edge and interior regions of reserves and their influence on species composition. Assumed ability to transfer data from plot to produce regional maps of endangered species.	Knowledge that human land-use legacies and disturbances modify edge and interior habitat quality. Inability to transfer plot data to produce regional maps of endangered species unless characterizes ecosystems into their diverse types within a vegetative community.
	Use of corridors and buffer zones to link habitat fragments and reserve networks. Reserve design focus – single large or several small, its size, where to place it.	Explicit link to GIS or remote sensing tools to integrate lessons learned from localized research sites.
Models	Small-scale, data-intensive species and community models.	Large-scale biodiversity models linked to spatial analysis tools under a climate change envelop. Ecosystem models geared to predicting human land-use activities and the sustainability of the ecosystems needed for conservation under a climate change envelop. Use of multi-objective decision-making models to assess trade-offs between conservation and sustainable development.
Economic evaluation	Development of nonmarket values for species.	Need to address all the values in each environment. Recognition that conservation efforts should not be driven just by economic analyses. Recognition that every landscape has multi-objectives and values compete so trade-offs need to be made.

legacy systems. Some factors contributing to the formation of legacy systems are:

- Past land uses and management practices, by fragmenting the conservation landscape and changing species composition, may be unable to provide sufficient and suitable habitat for species surviving in that landscape. For example, the insufficiency of habitat area for managing threatened grizzly bears has been well documented for the Greater Yellowstone Ecosystem. In this case, the logical approach would be to restore a sufficient area of suitable habitat for the grizzly bear. However, habitat alone will not ensure grizzly bears' survival since an exotic fungus is causing the decline of *Pinus albicaulis* whose seeds are an important food source for bears. This underscores the need to recognize that conservation efforts are complicated by other factors regulating the functioning of a system that are not considered because they do not appear to have a direct bearing on a species.
- Introduction of invasive plant species that alters the function of ecosystems, so they are less conducive for native species to grow. In Hawaii, an invasive nitrogen-fixing plant changed the abiotic soil environment of the newly invaded ecosystem or warm-season grasses arrested succession, thus the native species were unable to grow. Forest invasions by perennial grasses have been shown to increase the frequency and intensity of fire disturbances and thus shift the annual timing of these events. Invasive trees on forested watersheds have been known to transpire much more water than native trees, significantly altering forest hydrology.
- Existence of disturbances (e.g., hurricanes) that affect succession occurring at longer temporal frequencies than a manager's typical time frame. Recently the ecologists have come to understand the importance of disturbances that occur at long time frequencies (e.g., hurricanes) and how they affect the structure and drive the functioning of a system. Knowledge of these temporal dynamics is relevant for conservation, since the design of conservation efforts based only on consideration of intervals between disturbances can endanger an already threatened species. For example, when Hurricane Hugo swept over the tropical forests on Puerto Rico, the already small population of Puerto Rican parrots was dramatically reduced while the amphibian populations increased significantly. These hurricanes are a reoccurring phenomenon there and have been recorded at a frequency of 21 years. The effect of these hurricanes on determining the structure and functioning of these forests during the intervals between hurricanes is well documented. Consequently, if ecosystem processes are studied only during the inter-disturbance intervals, we may overlook the important causal drivers for that system.

Analysis of Evolving Thinking in Conservation Biology

Species to Ecosystem Shift

Because of the historical emphasis on saving threatened and endangered species, conservation biology has had a largely singular focus on protecting those species and their habitats. This focus was found on a simple ethic that humankind has a moral obligation to protect all living things on the earth. This species-centric ethos seemed to be relevant in the early phase of conservation because anthropogenic changes to the land base were geographically localized. However, current approaches in conservation are strongly based on the ecosystem, landscape, and human system. This shift resulted because of the failure of many conservation efforts that were based on the protection of species. Yet it is important to recognize that the single-species approach should not be eliminated from the suite of tools available in conservation in the push to assess systems more holistically.

In some ecosystems, the single-species approach will continue to be a powerful tool for conservation biologists. For example, the single-species approach may still be appropriate as a primary conservation tool when a single species functions as a keystone species, an ecosystem engineer, or an umbrella species or a coarse-filter. In some ecosystems, keystone species can play an essential role in regulating ecosystem structure or function, and the maintenance of community diversity. In the event that keystone species are lost from the system, the result may be a disproportionately larger effect on some property of the system so that certain species can no longer survive in the landscape. However, the concept of keystones may be difficult to apply in conservation because these species are not always obvious or we have not developed a clear mechanism for identifying them in the system, and some are only keystones part of the time and under certain habitat conditions. Once keystone species have been identified, the ability to use the single-species approach simplifies conservation efforts and facilitates integrating conservation with restoration efforts. When a single species has not been identified that provides a key structural or functional role, species are poor indicators of whether a system is going to degrade or if the habitat can be maintained for the species needing to be conserved.

The use of a single-species focus in conservation is greatly facilitated when species function as ecosystem engineers. In any ecosystem, the existence of ecosystem engineers is easy to recognize because they modify some structural part of the ecosystem that can be readily detected. They do not have to be present in large numbers for the impact to be registered. In many examples, a few individuals will have a disproportionately large impact on directly changing plant successional processes and indirectly modifying carbon and nutrient cycles (as has been documented for beavers and moose). In addition, ecosystem engineers may play keystone roles at particular stages of the development of a vegetative community (i.e., their impact would not be measurable at all times).

In conservation biology, an important concept that has been discussed for more than a decade is the identification of single species that can function as an umbrella species or a coarse-filter. Umbrella species are those species that have large habitat needs (e.g., spotted owl, red-cockaded woodpecker, and grizzly bear) or have other requirements such that when the system is managed to conserve that species, many other species will consequently also be conserved. The single-species approach has been most often justified through the concept of umbrella species.

Unfortunately, many ecosystems do not have a single species that can be monitored to provide an indication of how that system is functioning and whether a conservation project will be successful. Based on our current scientific understanding, many species are merely passive inhabitants of ecosystem in which they live. Instead, entire communities of species may play integral roles in ecosystem function. Biodiversity may then be a critical determinant of ecosystem productivity and stability–including the tendency to resist anthropogenic impacts and the ability to recover from impacts (i.e., resilience). However, the relationships between biodiversity and ecosystem productivity or stability cannot be generalized to all ecosystems. Strong relationships between these variables have been found in certain ecosystems (e.g., grasslands), but biodiversity cannot be automatically used as an indicator of the ecosystem integrity. For example, the biodiversity at the species level is a poor indicator of the resistance and resilience characteristics of many forests to disturbances since the response of the system occurs at levels other than species diversity (e.g., genetic diversity within a species).

The growing realization that biodiversity may be inextricably linked to ecosystem function has given rise to a more holistic, system-centric perspective for conservation, now known as ecosystem management. This new approach was codified in 1996 by Mangel and others, when they defined a set of principles to which conservationists must adhere to successfully conserve wild living beings. The principles of ecosystem management are:

> (1) Focus on the sustainability of ecosystems, not on the output of products; (2) Adopt a holistic understanding of the way all the parts are linked together in an ecosystem and the feedbacks among those linkages; (3) Incorporate a long-term perspective and examine issues at a scale relevant to the functioning of the ecosystem; and (4) Recognize that human values shape ecosystem structure and function in myriad ways that can constrain, promote, or reduce sustainability.

Because the primary goal of conservation is to sustain the habitat or ecosystem where species live, conservation biologists have already espoused the first principle of ecosystem management as articulated here. Most of the tools developed in conservation biology address this first principle but not the other three principles. Similar to what is still happening in ecosystem management, conservation biologists are currently assessing the practicality of the different tools to implement the other three principles in conservation. Out of the four principles listed here, the incorporation of people and their values into a holistic framework of analysis that is not solely human-oriented is one of the greatest challenges for both ecosystem management and conservation.

Contemporary conservation also faces similar challenges to those previously experienced by natural resource managers attempting to implement ecosystem management under conditions of high scientific uncertainty. For example, natural resource management shifted from a resource-based emphasis (e.g., forest stands, fish stocks) to a more holistic, ecosystem-based approach that better enfranchises all stakeholders of a particular land or resource base. In the face of such a shift, managers were forced to embrace new management models (i.e., adaptive management) to deal with large-scale management problems under a scenario of high scientific uncertainty about the response of ecosystems to human activities. Because our scientific knowledge base is still fairly limited, few guidelines exist to help resolve conflicts arising from the need to determine the tradeoffs between different natural resource uses. These conflicts are especially difficult to resolve since conservation biologists and ecosystem managers have to deal with resource scarcities.

The current view of Gordon is that ecosystem and adaptive management are both approaches attempting to respond to resource scarcities. Ecosystem management is attempting to respond to a scarcity of land. This land scarcity means that the exclusive uses of any given land area for one purpose is no longer possible because the production capacities of ecosystems are finite. This contrasts with adaptive management, in which managers are responding to scarcity of scientific information and knowledge. This scarcity makes it difficult to develop consensus on which tools are most appropriate to use for resolving the conflicts over resource uses. In a similar vein, conservation biologists are being forced to respond to scarcity and the resulting conflicts that can arise over the use of finite resources. However, conservation has yet to adopt a model that can adequately handle conflicts resulting from the scarcity of resources and simultaneously achieve natural and social system sustainability.

Direct and Indirect Influences of Humans on Conservation Efforts

Humans are strong influences on whether conservation projects are successful. A major deterrent to conservation has been the inability of most governments to satisfy all the resource uses, services, and values that are expected from the same land base by local, regional, national, and international stakeholders. Under these circumstances, the scarcity of land area and resources often results in conflicts over the uses of natural resources that typically polarize the people inhabiting the environment against conservation organizations articulating a need to establish reserves to save species. These conflicts often occur in environments where the land base constraints the options that are available to satisfy all the needs of all stakeholders. Some of the direct impacts of humans on the environment appear to be easier to manage, since alternative solutions are possible through substitution of products, technological breakthroughs that make resource extraction more efficient, shifting to the use of different resources, alternative employment opportunities, and so on.

Some of the direct and indirect impacts of humans are much more difficult to manage in conservation since we are just beginning to understand these effects. Many of these effects occur at longer temporal and spatial scales, making it more difficult to explicitly link the cause-and-effect relationships needed to identify the tools which will indicate when changes to species and ecosystems may occur. The impacts of humans on species loss and changing ecosystem structure and function will be further discussed next, especially as it relates to sustainable development.

Humans and Sustainable Development as an Integral Part of the Conservation Formula

Within the last two decades, there has been an increasing understanding that conservation projects cannot be successful when implemented in isolation from neighboring human populations. Management of nature reserves has undergone a major strategy shift from trying to keep people out of reserves to allowing some economic development by local communities as an integral part of the conservation strategy. The response of the international community to deal with this issue has been to design projects that explicitly link international aid to local communities to the establishment of nature reserves to conserve species. The goal of these projects has been to improve the livelihood of people living close to the natural resources so that they would not be dependent on extracting species or resources from the protected area. The strategy behind this approach was to find a balance between development and conservation by providing alternative approaches for people to survive and at the same time provide zones of lower human impact for conservation purposes.

This ecosystem-based model of conservation attempts to reconcile human pressures on resources in areas rich in biodiversity and those identified as high priority areas for conservation. This represents a major shift from the traditional conservation model, whose main focus was to lock up resources by removing humans already living in nature reserves without considering how these communities would continue to survive when denied access to these areas. Conservationists now recognize that successful conservation efforts will require explicit recognition and alleviation of human dependence on natural resources at the local level. It has become quite apparent that it was insufficient to demarcate parks and assume that people with few options would stay out of these areas.

In fact, humans have had an integral hand in shaping many of the landscapes and ecosystems that we now wish to conserve. It is important to recognize that present forest conditions may be strongly influenced by management and land use activities of the indigenous and rural communities. Thus by eliminating indigenous land managers from the landscape will likely alter the forest composition to a state unacceptable or undesirable for conservation purposes. For instance in wet tropical forests, humans have strongly influenced the resident biodiversity; indeed, the ''virgin forests'' and exceptionally rich biodiversity that are the focus of conservation efforts in the Amazon River basin are in part relicts of old agricultural practices and management.

The importance of the human dimension was especially reinforced for conservationists working in the tropics, where desperate people continued to harvest resources from forests set aside as nature reserves. Many of the conservation organizations funding the establishment of protected areas quickly recognized that nature reserves could not be maintained if they were designed to exclude local communities without providing alternative sources of income. This recognition stimulated the development of conservation and sustainable development projects called Integrated Conservation and Development Project (ICDP) which will be further discussed in the Section, Economic Evaluations and Conservation Tradeoffs with Competing Resource Demands. Most of the current projects designed to conserve biodiversity are encouraging local communities to use areas outside of reserves or at least outside of the core conservation zones. This strategy resembles the UNESCO Man and the Biosphere (MAB) design for reserves, which is based on having concentric zones representing different intensities of human use–with the center or core area designated for conservation without human presence. However, shifting or confining people's use of forest to particular zones without their cooperation is impossible in most developing countries. Often, land tenure or ownership does not exist for these local communities but they are still dependent on extracting resources from these same lands for their survival. Furthermore, these communities are accustomed to adjusting their resource acquisition efforts within the landscape both seasonally and by migrating to different locations.

In more industrialized countries, conservation organizations historically did not have to deal with indigenous or rural communities. The reason for this is that newly designated conservation areas had already mostly excluded or removed communities who lived and survived off the natural resources contained on these lands. However, this does not mean that the more industrialized countries do not incorporate the human dimensions into conservation; rather indigenous people have not driven the approaches being used in conservation. For example, Native Americans used fire to maintain grasslands for ungulates or to manage forests to increase the dominance of trees with cultural value such as cedar or pines. However, when parks were originally established in present day US, wildfire control was the accepted practice and fire was not deemed as an appropriate park management tool.

In stark contrast to the less industrialized countries, industrialized countries have placed greater emphasis on developing the tools to protect and restore habitats and species, and in eradicating invasive species that threaten indigenous non-human species composition and alter succession. The spread of invasive species and the concomitant loss of indigenous species is a problem driven by international trade and the globalization of world trade. The changing transportation networks and development of a global economy are contributing to an accelerated rate of the spread of invasive species that has not been documented prior to this time. In addition, the spread of invasive species frequently occurs in ecosystems that have been significantly modified by human land-use activities and few good control mechanisms have been produced. Humans have been modifying their environments at unprecedented rates during the last several decades. Because of the role that humans are playing in spreading species and in altering the health of ecosystems, one of the greatest challenges to the successful implementation of conservation projects may be our ability to control the spread of invasive species. The impact of invasive species on conservation efforts is further discussed in the next section.

Humans and Invasive Species

Within the last two decades, many conservation organizations have implemented programs to reduce or eradicate exotic invasive species. This was driven by numerous cases of invasive species disrupting or altering native ecosystem functions and structures, as well as reducing regional biodiversity. The significant economic costs of controlling invasive species divert resources from other areas of biological conservation. Invasive

species challenge conservation efforts because they limit the options available for the conservation of indigenous species and their intact ecosystems. The spread of invasive species will probably become more critical in the future as humans create conditions or change disturbance cycles sufficiently to give invasive species the opportunity to dominate more habitats. Climate change is also expected to increase the ability of invasive species to spread into landscapes under duress (*see* the Section, Contemporary Conservation and Climate Change).

Humans have played a dominant role in the spread of invasive species around the world. Humans function as vectors of invasive species and alter the landscape in a way that enables invasive species to become better competitors than native species. Many weeds, insects, and other animal pests are transported accidentally with produce and goods; food crops and horticultural varieties are often introduced to new environments intentionally for agriculture and landscaping. Likewise, there is a long history of humans introducing game animals to new territories. Species already associated with areas of human development are consequently more likely to be introduced to new areas by human vectors, and the areas to which they are introduced tend to be more readily overwhelmed. Observations suggest that alien plant species richness correlates positively with proximity to human development and transport centers. For example, stocks of seed and produce shipped overseas often carry with them seeds of many weed species that germinate in their new environment. These shipments commonly include alien microbes and animal pests as well. Intentional introductions of species imported for agriculture and other human uses are common, as are introductions of organisms employed as biological control agents to stanch the tide of previously introduced and invasive species. Today, the production of biofuels using grasses is stimulating new concerns that potentially invasive plants will be introduced to areas where grasses have not previously grown and where they are more competitive compared to native species.

Although the invasion of ecosystems by alien species certainly has occurred throughout human history, the problems of alien invasion remain particularly challenging both for scientists studying patterns and processes of invasion and ecosystem managers attempting to eradicate or control invaders. In many parts of the world, particularly on islands, even the greatest conservation efforts are ineffective at countering the proliferation of exotic species. Where invaders are well established, such as rats in New Zealand and *melastome* plants in Hawaii, invasive species are nearly impossible to control and conservation strategies need to explicitly manage invasive species that currently cannot be eliminated. In less severe cases, conservation practices may be able to protect biodiversity and limit the establishment and spread of exotic species.

Because of the uncertainty in our understanding of invasive species, conservation biologists will have to consider whether invasive species are present in areas targeted for conservation. If invasive species are present, it will be important to consult the biological literature and determine whether and what types of effects may be attributed to that invasive species. If these invasive species are shown to negatively affect indigenous species and ecosystems, it will be critical to eradicate small pockets of invasive species when they are found to prevent their establishment and spread to the point where they cannot be controlled. This suggests that conservation practices will have to aggressively manage invasive species even when their impacts have not been conclusively shown to detrimentally affect the target system. It will be critical for conservation to establish priorities and develop an effective strategy for preventing and handling invasive species.

Tools for Managing and Conserving Species

When Michael Soulé, the "father of conservation biology," was interviewed in 1994 and asked about the developments in conservation biology in the last decade, he pointed out that there has not been a revolution in how conservation biology is practiced. Rather, he stated that conservation has been experiencing a gradual refinement in the principles and tools of the field. Soulé suggested that the major tools have been the application of theory from the fields of island biogeography, population biology, and community ecology, namely, consideration of the matrix landscape within which a habitat is embedded, development of biological corridors, development of new tools in molecular biology, and development of integrated population viability analysis frameworks that include demographic, environmental, and genetic stochasticity. All of these are commonly still used in conservation today.

Since the early 1990s, however, there has been a revolution in the philosophical basis of conservation biology. As discussed earlier, five philosophical changes have occurred in conservation; however, two have arisen to the forefront of discussions during the last two decades. These two philosophical changes have also been the most difficult to convert into practical applications for natural resource management. One is the acceptance of the ecosystem approach as the management tool to resolve conflicts arising from the public perception that communal lands are poorly managed and not conserving biodiversity. The second philosophical change is the acknowledgment of the importance of human dimensions in natural resource management. Because both philosophical changes determine what tools and approaches to use in conservation, they have made the science and practice of conservation more challenging and have required conservation biologists to make more complicated and difficult choices. For the purposes of this review we define conservation approaches as specific frameworks or underlying structures used to benefit biodiversity indirectly, and conservation tools as devices or processes designed to benefit biodiversity directly. Contemporary conservation tools and approaches will be explored next.

Contemporary Conservation Tools and Approaches

Based on discussions in the Sections, Introduction: New Concepts in Conservation and Analysis of Evolving Thinking in Conservation Biology, it is easy to see that the practice of conservation has changed dramatically since its inception in the early twentieth century. In fact, conservation has evolved to such a great rate in recent decades that many of the

approaches, tools, and foci of contemporary conservation are largely different from those implemented 10 years ago. Yet, despite a number of broad and fundamental developments in the field of conservation, perhaps the clearest trend has been the shift away from the dominant approaches developed by early conservationists. Early conservation was primarily focused on the creation of parks for recreational and aesthetic reasons, and the conservation of a small number of species, particularly beautiful birds and enigmatic mammals. To achieve these objectives, such an approach to conservation developed tools that were effective at designing and monitoring.

However, over the course of the twentieth century, conservationists' focus has expanded from the objective of establishing beautiful parks and conserving select species towards a more holistic goal of ecosystem integrity; a goal that goes well beyond the conservation of individual species and beautiful landscapes to include the protection of the existing diversity of species, natural habitats, and ecosystem processes. While protecting particular species and areas of outstanding natural beauty remain an element of contemporary conservation, advocates for the protection of ecosystem integrity argue that a focus on single species or beautiful landscapes inherently limits biodiversity conservation. In other words, to focus conservation efforts solely on a limited number of species makes the conservation of earth's biodiversity almost impossible, as the sheer number of species and the lack of positive attention to many species (e.g., insects, snakes, fungi) leaves them and the ecosystems unprotected. Likewise, by focusing conservation efforts only on natural areas of high aesthetic value (e.g., temperate mountainous regions), other ecosystems critical to the continued provision of ecosystem and cultural services may be lost.

Taken together, as a result of these persuasive arguments and broader changes in thinking across the conservation movement, contemporary conservation has shifted its focus to protecting biodiversity in all its forms. One of the primary outcomes of this shift is the development of a host of new tools and approaches to achieve conservation. These developments have undoubtedly improved the efficacy and efficiency of conservation, but they have not eliminated conservationists' need to answer fundamental questions about their goals and strategies, particularly (1) What biodiversity should be conserved; e.g., should the focus be particular species, ecosystems, or ecosystem services? (2) Where does the targeted biodiversity occur, and where is the best place to protect it? and (3) Given the variety of conservation tools available, which is the most effective method to achieve conservation objectives? Each of these questions should be asked by a conservation planner because they require managers to address the key elements and challenges integral to the development of an ecosystem-based toolkit to guide conservation efforts. These questions are often difficult to solve, but they cannot remain unanswered, as they provide essential guidance to conservation efforts. And while the science of conservation has matured dramatically over the last few decades, the field of conservation still remains immature, leaving conservationists to work with a body of knowledge and set of tools that remain incomplete. However, despite a limited understanding of how conservation efforts interact with surrounding socio-environmental systems, the field of conservation has amassed an impressive base of knowledge, as well as a robust set of tools and approaches to achieve its goals.

Identifying Biodiversity Conservation Goals

In recent decades, the field of conservation has coalesced around three overarching methods for conserving and protecting biodiversity: (1) *in situ* conservation (i.e., onsite conservation), (2) *ex situ* conservation (i.e., offsite conservation), and (3) restoration (i.e., the reestablishment of native species and ecosystems). With the growing depth of conservation science in mind, we now turn to the prominent conservation approaches and tools used in *in situ* conservation.

When conservationists focus on *in situ* activities the first important question they often face is "what biodiversity should be conserved?" Among the many options available to answer this question, one of the more prominent methods used is known as the "hotspots" approach. The concept of biological hotspots, originally developed by Myers, is designed to focus conservation efforts by identifying geographic regions with exceptionally high levels of species richness, endemism, or rarity. The hotspots approach has proven widely popular and has been embraced across a variety of conservation groups and initiatives, including the Critical Ecosystems Partnership Fund, whose members include Conservation International and the Global Environmental Facility, the Endemic Bird Areas initiative organized by Birdlife International, and the Global 200 Ecosystems initiative developed by World Wildlife Fund. Generally speaking, each hotspot initiative generates maps and strategies to prioritize conservation efforts for ecosystems with exceptionally high levels of biodiversity, and depending on the hotspot methodology used, areas defined as hotspots have undergone a substantial loss of habitat and face growing threats from human disturbance.

Another approach often used to define the focus of conservation efforts is known as the coarse-filter/fine-filter approach that was originally developed by the US Nature Conservancy. The coarse-filter/fine-filter approach is designed to achieve biodiversity conservation by ensuring that a given landscape's naturally occurring species and ecological communities are protected. At its most basic, the coarse-filter approach is a broad scale analysis designed to identify and protect existing plant and animal communities, whereas, the fine-filter approach is used to complement the coarse-filter approach by identifying individual species, features, or processes missed by the coarse-filter analysis. Together, the coarse-filter/fine-filter approach enables conservationists to protect representative samples of native species, plant and animal communities, and ecosystem processes occurring across a targeted landscape, while ensuring that conservation activities include unique or underrepresented species and features.

Designing and managing conservation networks using the coarse-filter/fine-filter approach has become commonplace in many organizations involved in conservation efforts across the US and beyond, including land management agencies, such as the US Forest Service, large conservation organizations, particularly the Nature Conservancy, and a variety of small conservation groups, including many local land trusts. Conservationists frequently use the hotspot approach or the

coarse-filter/fine-filter approach, but in other cases conservation priorities are determined by existing species, ecosystems, or ecosystem services, as well as undeveloped green spaces, or cultural and historic features.

After conservation priorities have been determined, the next set of questions often includes, "where does the targeted biodiversity occur, and where is the best place to protect it?" By asking this question, scale considerations will become the focus of a conservation planner. Today, the shift from conserving species to an ecosystem-based approach complicates the demarcation of the habitat borders and the extent of land area needed by conservation projects. Because identifying the scale needed to protect biodiversity is critical, a brief discussion is given of the evolution of our approach to species and scale before the ecosystem approach began to dominate.

Identifying the Appropriate Conservation Landscape Scale

Species-Based Conservation Scale

Much of the original theoretical work by conservation biologists was based on determining the relationships between species and their space (as first articulated by MacArthur and Wilson). The suggestion that habitat area was positively correlated to species richness provided arguments for designing larger natural reserve areas for maximizing the conservation of species. The size of a reserve area needed for effective conservation has been argued in the literature for more than 20 years without the development of a consensus. Part of the reason for a lack of consensus is that we know that area by itself is insufficient to characterize how many species can be maintained at a site.

Even though space has been an important theoretical issue in conservation, the practical application of theory in conservation appears to have no dominant spatial scale of analysis. Therefore, theories and applications have not been linked at the spatial scale. A lack of defined spatial scale in applied conservation has been strongly influenced by the focus on specific animals as important species to conserve. The spatial scale at which conservation has been practiced has varied according to the territorial needs of the species of interest. This trend is apparent from a survey of key journals published in 1998 by Vogt and others who tabulated the scales of analysis used by conservation biologists, community and population ecologists, and social ecologists. An amazingly high proportion of the articles published in the journal *Conservation Biology* did not even present the scale being used in the study; this suggests that spatial scale was not considered relevant by the authors for understanding their system. When a scale of analysis was presented, the smallest scale (<0.01 ha) was as equally represented as the largest scale ($>10{,}000$ ha). Again, this reflects the focus on animals with different habitat area requirements.

This contrasts with much of the information collected from community and population ecology journals, where small-scale research is predominantly used to understand species relationships to ecosystem parameters. For example, plant ecologists who developed the theory behind species dynamics focused on scales that were less than 0.01 ha in size. Some studies are being conducted at larger scales, but these larger scales of analyses are applied in the tropics and are less common in other regions of the world. In the humid tropics, the presence of rare plant species has necessitated the establishment of permanent plots that are 50 ha in size.

Furthermore, because conservation biologists have accepted the ecosystem management paradigm, many more scales of analyses are included in conservation. For example, the watershed has been accepted as the relevant scale in ecosystem management, for it includes the drainage of a freshwater footprint of a bounded-land area. In ecosystem management, the watershed defines the scale and boundaries of interaction of the social and natural environments. A watershed scale is only one of the many scales that could be used in conservation but it will only be the appropriate scale for some animals, e.g., small, non-migrating animals. Selecting one scale of analysis for all systems is problematic because scale is site dependent and must be chosen to reflect the driving variables controlling the functioning and resilience of that system.

Landscape Ecosystem-Based Conservation Scale

Today, there are a number of ways to identify and prioritize the location of conservation efforts using the ecosystem approach, including "gap analysis," site selection algorithms, and expert opinion. A gap analysis is a spatial tool used for conservation planning designed to identify conservation gaps in conservation areas. The objective of the gap analysis is similar to the coarse-filter/fine-filter approach in that it seeks to ensure that all components of biodiversity on a given landscape are included in conservation areas. In addition to the gap approach, soliciting expert opinion and using site selection algorithms or a combination of expert opinion and site selection algorithms can be used for reserve planning and expansion. Experts can range from local community members familiar with the location of local flora and fauna to research scientists and land managers. Generally, site selection algorithms tend to result in more cost efficient conservation networks, as they tend to rely on objective measures, such as the price of land parcels. However, few algorithms have been used in applied conservation because of their complexity. Nevertheless, a combination of expert knowledge and algorithms may result in robust conservation planning, especially in areas with limited or non-existent datasets.

The selection of specific sites to focus conservation efforts is essential to the practice of conservation, but site selection is not limited to small geographic areas. Rather, because many areas of natural habitat are often islands in large seas of human landscapes, one emerging approach to focusing conservation efforts is to ensure landscape connectivity between protected areas. To connect dispersed conservation areas into an integrated conservation network, often termed landscape-scale conservation, conservationists work to identify the best locations to create habitat corridors and enlarge protected areas. A common underlying goal of this approach is to increase habitat connectivity for species that require large areas of continuous habitat, such as many large mammals and migratory species. Additionally, another catalyst for the push to connect natural habitats across large geographic areas is the emergence of climate change as an increasing threat to biodiversity. Put simply, as the global climate continues to

warm, many species will be forced to migrate to more suitable habitats, but if the habitats they need are inaccessible, the likelihood of their extirpation or extinction is magnified (see the Section, Contemporary Conservation and Climate Change).

Today, landscape-scale conservation initiatives are becoming a popular approach to conservation, and numerous projects are underway around the world. One innovative proposal known as the Sky Islands Wilderness Network (SIWN) seeks to conserve and connect natural habitats across forty arid mountains ranges of the American Southwest and northern Mexico. Another proposal in North America, even larger than the SIWN, is a plan to link habitat from Yellowstone National Park in the US to the northern Yukon region of Canada. This ambitious plan, known as the "Yellowstone to Yukon" initiative, aims to ensure connectivity between lands spread across more than 1.3 million square kilometres. Another ambitious landscape-scale proposal worthy of note is the proposed Mesoamerican Corridor in Central America, which is designed to create a continuous habitat corridor between parks in southern Mexico to other protected areas throughout Central America (Trombulak and Baldwin, 2010). The bulk of these plans are on public lands where landscape scale initiatives are easier to implement in contrast to private lands because of their competing and often conflicting land-use objectives. The long-term goal of these projects is to expand the conservation value of public lands by buying private lands when these lands become available to increase the landscape connectivity for biodiversity.

Conservationists now generally agree that landscape-scale approaches are essential to conserve biodiversity. Arguably the greatest challenge facing many landscape-scale conservation initiatives is managing a greater number of complex issues than those typically faced in smaller conservation projects. Increased complexity for larger landscape-scale projects stems largely from the greater diversity of human demands and governance structures inherent across larger areas. However, given that the core challenge of conservation is ultimately about managing people, the fundamental focus of conservation is often remarkably similar across both large and small scale conservation projects. So whether conservation efforts are focused on a small city park or across vast heterogeneous landscape, conservationists must ultimately confront the needs and preferences of people and balance them with the needs of biodiversity. In considering the many challenges inherent in balancing competing needs, the third and perhaps the most important question facing the practice of conservation is "given the variety of conservation tools available, what is most effective method to achieve conservation?" This question will be briefly discussed next.

Need for Multiple Conservation Operational Tools

From establishing parks to enacting environmental legislation, policy-makers and conservationists have a wide variety of conservation tools available to them. Yet, while there are many possible options, no tool is a "silver bullet" because contextual factors are highly variable across time and space. Contextual factors most pertinent to the effectiveness of conservation tools include: political and financial considerations, the specificities of the conservation target, and institutional capacity for implementing conservation strategies. Also, the sheer number of tools available can create difficulties in selecting the most appropriate tool for each situation. In short, determining the right tool to achieve conservation objectives can be difficult.

To simplify the selection process, it is helpful to evaluate available options based on the operational theories underlying the development and implementation of each tool. By using operational theories as a guide, conservation tools can be organized into three categories: (1) *Governance-based* tools, (2) *Market-based* tools, and (3) *Civil society-based* tools (see **Table 2**). These are summarized below:

Governance-based tools are often premised on the assumption that conservation objectives are most *effectively* achieved when developed and conducted by a government – local, regional, or national. The effectiveness of these tools is thought to stem from the inherent powers vested in governments to create and enforce the rule of law. Governance-based tools are often defined as a "top-down" or "command-and-control" approaches to conservation.

Market-based tools are commonly identified as the most *efficient* approach to conservation. Market-based tools are believed to be efficient because they are designed to use cost advantages often achieved through the practice of free market economics. Market-based tools are sometimes described as "free-market" approaches to conservation. Market-based approaches can achieve conservation goals because people generally act in their own self-interest. Given this reality, a host of tools have been developed to incentivize human actions that also benefit biodiversity.

Civil society-based tools are often considered to be the most *durable* conservation solutions because they are found on the strength of civic consensus and democratic principles. The tools of civil society embrace a wide variety of methods developed and implemented by individual citizens and civic organizations. These tools are often defined as a "bottom-up approach" to conservation.

Table 2 The three categories of tools and how they have been implemented in conservation landscapes

Governance-based approaches	*Market-based approaches*	*Civil Society-based approaches*
Parks, domestic regulations and legislation, international treaties, criminal prosecution and fines, debt-for-nature swaps, species reintroduction, and management	Payment for environmental services, forest products certification, sustainable products, ecotourism, and conservation investment banking	Privately owned parks, civil suits, media and education campaigns, civil disobedience, conservation easements, backyard habitat areas, and fundraising and grant-making

Economic Evaluations and Conservation Trade-offs with Competing Resource Demands

In the past, economics was the dominant tool used to build a market-based rationale for conserving species and ecosystems. Economics was used to assess whether conservation, and the resulting protection of ecosystem services, is more valuable than extraction or harvesting of resources. However, this overreliance on economics as the primary tool to build a rationale for conservation has produced mixed results. It has been extremely difficult for economists to provide a monetary value to non-market goods obtained from an ecosystem, such as clean water, or to value goods that are locally traded but where there is no transfer of money. This quandary is apparent from the brief discussion that follows of how economic evaluation developed as an approach to link conservation with sustainable development.

In the late 1980s, the need to address the trade-offs between conservation and the sustainable livelihoods of communities living in the same tropical landscape, conservation biologists and economists began to combine their expertise to make conservation efforts more robust. Because these rural communities were dependent on resource extraction to survive, they needed new economic opportunities to replace the resources to which they no longer had access when their traditional resource gathering grounds were placed under protected status. These economic analyses suggested several options for generating income for rural communities, but mostly they suggested producing economic value for previously non-market-valued products.

Historically, tropical forests had been undervalued when only timber and agricultural uses were used to estimate their net worth, so that there was no clear economic incentive to conserve these forests. If market-value could be generated for non-timber forest products, it would diversify the economic options for rural communities while maintaining habitats for species. Economists had suggested that the lack of proper economic valuation had contributed to accelerating the rate of harvesting and conversion of forests for other uses. Therefore, a tactic to decrease deforestation rates was to determine the "real" value of the forest by including all other products that could be extracted from it (e.g., non-timber forest products and services like watershed protection). Using an ecosystem based economic analysis, the forest was no longer undervalued, and alternative products could be extracted that would provide income to local communities. This strategy was widely promoted by the international development community to conserve forests and biodiversity in reserves while also generating income for local communities from the same reserve or surrounding areas.

A strategy of combining conservation with sustainable development, as previously discussed and known as ICDP, was supported by economic studies reporting that the extraction of non-timber forest products would return greater value in the marketplace than timber harvesting. In addition, this approach was promoted as a new alternative to reduce the dependence of the local community on natural resources that were being protected in a newly formed conservation area. International funding organizations were interested in developing these non-market values from forests because it suggested that the goals of conservation were compatible with developing an economic base for local communities. The idea was that local communities could forgo harvesting some resources (trees, wildlife, etc.) within designated reserve areas if sufficient income could be generated from pursuing alternative income generating activities in the same forest. As part of the new ICDP approach that began in the 1990s, ecotourism and the harvesting of non-timber forest products were presented as two viable alternatives for local communities to generate income. However, these options have not as yet proven to be effective in meeting the dual goals of biodiversity conservation and rural economic development. Unfortunately, it is not clear whether the goals of conservation and sustainable development are compatible under this rubric. Dove and others have conducted research suggesting that this strategy should be approached with caution, for developing a total value for a forest does not protect it from being overharvested.

Some economic analyses have been criticized for producing unrealistic expectations of the value of what can be sustainably extracted from a forest. These earlier economic analyses used biological inventory data to calculate an optimal amount of non-timber forest resources to be collected and the potential economic value of harvesting non-timber forest resources. These economic analyses did not consider the dependence of local communities on these same resources for their survival or their role in maintaining ecosystem function or landscape biodiversity. Biological inventory data also do not indicate the amount of a product that can be harvested in an ecologically sound manner. These data produced results that significantly overestimated the expected income generated annually from extracting these products. Indeed, placing an economic value on a resource can increase the demand for that product so that it is quickly overharvested; when this occurs, the species may be lost from the landscape where it previously was quite common.

Conservation projects have also struggled with how to achieve the dual goals of conservation and maintenance of indigenous communities living in the same land. Frequently the species to be conserved was also an important food source for local communities. Therefore, rural communities were perceived to be a threat to conservation. There are abundant examples from around the world of the competing objectives causing strife between conservationists and rural communities. This has made it difficult for conservationists to balance the goals of conservation projects with the continued presence of indigenous communities living on these same lands. This inability to make choices, and the assumption that these communities threatened conservation efficacy, drove the forced removal of indigenous communities from their traditional lands. These people have been called 'conservation refugees'. For the indigenous communities, the consequences were dire. This loss of access to resources from their traditional lands caused them to become dependent on government subsidies to survive or to have few survival options because of the difficulty of adapting to landless based employment opportunities.

There is greater recognition today that alternative and diverse approaches are needed to increase the economic opportunities for local communities to become self-sustainable. Recent efforts to allow local communities to manage and

make the trade-off decisions on how resources are protected or consumed have facilitated local communities' contributions towards regional sustainability. For example, forest certification in Indonesia had the unintended consequence of facilitating indigenous communities to gain more control over managing and controlling resources on lands that they traditionally owned, i.e., today they have no land ownership rights for these lands. A new emerging trend in the US is for indigenous and rural communities to become stewards and comanagers of public lands for economic benefits and conservation even though they do not own these lands.

Conservation in Human Altered Landscapes

The solutions to the problems facing conservation biologists will not be found solely by adopting new paradigms such as ecosystem management or sustainable development. Blending the theories developed in conservation biology with these new paradigms will potentially allow for better management of species within their landscapes. However, conservation biologists will not automatically find the tools that they need from other disciplines. These new paradigms are currently struggling to develop tools that will be able to take theories and make them applicable to, for example, forest certification. Furthermore, conservation biologists have not developed structures similar to ecosystem management that would allow them to develop an adaptive management model (i.e., a way to allow scientific data to continuously modify policy links as better information became available). Without such a process or framework to effectively link policy to science and management, it is difficult to integrate scientific findings into management itself. As stated by Gordon in 1999, "Often, policy is made with 'old' science, while interest groups confront the manager of public resources with 'new' science." One could speculate that "old" science is used because the tools have not been developed to implement or make the "new" science work at the ground level or in policy.

Conservation biologists will also have to confront the fact that humans have altered the environment on a global scale. On this globe, it is difficult to find land that has not experienced some form of human land-use activity or extraction of resources. For example, 30 years ago scientists were shocked to find pollution generated in distant countries on glaciers in the European Alps. In many cases, the current ecosystem we want to conserve may be an artifact of past human activities. If this is not recognized, it may be extremely difficult to conserve species because we really do not understand the driving variables that may control the resistance and resilience characteristics of the system. As protected area managers are increasingly called upon to generate funds through development projects (e.g., ecotourism, timber harvest, and nontimber forest product extraction), it is important that one of the original goals of setting up these reserves as "genetic libraries" is not compromised.

Restoration is frequently suggested as a way to re-establish desired ecosystems and at the same time maintain the biodiversity of indigenous species and exclude the invasive species. This is especially relevant considering the problems that conservation managers are facing with invasive species and altered landscapes. However, the science of restoration ecology is still evolving, and standard tools are not readily available for utilization by conservation biologists. A current practice has been to allow the development and elimination of a specific habitat with the assumption that an equivalent habitat will be recreated at another location. This approach has been commonly used in wetlands. Unfortunately, the successful restoration of a habitat that can then be used by an endangered species rarely results. For example, it is not unusual that a restored marsh that was supposed to support an endangered species was not suitable habitat for that species or did not support other species found in a reference marsh. Since restoration ecology is still a young science, it is not surprising that restoration tools still need to be tested in a variety of systems. Many restoration efforts are currently dealing with eradicating invasive species that have begun to dominate much of our landscape. Yet we have a minimal understanding of the impact of invasive species within landscapes and of the long-term ecological effects that they may exert in these newly modified environments.

Conservation biology has done well in developing the science of understanding individual species in their habitat, performing spatial scale analyses of individuals, and modeling their activity within the landscape. It has not done well in the following areas:

- Incorporating the roles and needs of humans in conservation situations, especially at the local level. The values and principles of national and international organizations have largely driven conservation efforts.
- Developing frameworks for moving beyond the species focus and incorporating the dynamic nature of ecosystems into these frameworks.
- Focusing on the use of "easy" tools (i.e., GIS models) has restricted the diversity of conservation approaches and created a dependence on a small set of tools, thereby excluding the use of a more varied and innovative suite of tools.
- Adapting conservation planning to future environmental (e.g., land conversion, loss of forests, loss of soil productive capacity, climate change, etc.) and social change (e.g., technology in converting bio-resources to new products, higher population density, increased industrialization of less developed countries, etc.).

One of the greatest challenges to conservation efficacy is climate change. It can potentially decrease environmental as well as social resilience since it is not clear how species will adapt to climate change and its altered disturbance cycles. This topic deserves further discussion because of its potential impact on conservation planning.

Contemporary Conservation and Climate Change

There are many potential impacts of climate change on the ecology of conservation landscapes. The predicted ecological consequences and impacts on biodiversity due to climate change include significant restructuring of coastal and estuarine systems, shifting species distributions (often hindered by habitat fragmentation or geographic isolation), as well as

significant changes in ecosystem processes including altered fire and hydrological cycles. Furthermore, temperature induced changes in plant phenology may disturb plant–pollinator relationships. In addition, temperature changes are also predicted to trigger trophic cascades, parasite and pathogen pandemics and general population declines especially of sensitive species. Most of these impacts will be expressed across larger landscape scales making it challenging for conservation planners managing at the scale of a reserve.

The most challenging impact of climate change may be the alteration of habitat area and quality needed by conservation projects. Species populations depend principally on the area of suitable habitat and the quality of this habitat. Therefore, conservation planning may need to include strategies to maintain and increase the areas of high quality habitats, prioritize centers of species endemism, and also prioritize areas that have high environmental heterogeneity. Maintaining conservation areas with fixed boundaries, however, is problematic given the vagaries of climate change. The artificial and even natural boundaries designating zones of conservation may be rendered obsolete by rising water tables, persistent droughts, and other climate induced alterations. Increasing the connectivity of conserved regions is a contemporary tool to hedge against the climatic realignments thereby providing conserved species a potential exit strategy.

Projected temperature changes may shift the ranges of diverse taxa poleward by an average of 61 km over the next hundred years. This distance is an order of magnitude larger than the average north-south extent of protected areas in North America. Therefore, this dramatic shift would require current conservation zones to be increased 10-fold, which is an unlikely situation given the challenges of implementing and managing conserved land. This projected increase assumes exiting conservation zones will remain habitable for many species which is a difficult assumption to make since many ecosystems are decreasing in area (e.g., tropical cloud forests) or no longer provide the refugia (e.g., Sky Islands of the south western US) for species due to land-use changes that are isolating ecosystems in a landscape.

Since climate change models address scales that are global or at regional scales, it may be difficult to justify using them in conservation planning being implemented at a local scale. For example, managers of parks or nature reserves may find it difficult to manage for changes in species distribution that may occur beyond their management area. In other cases, vegetation shifts along moisture gradients and temperature isoclines are predicted at large geographic scales, so the management area may only encompass a small portion of the area needing to be managed. Coordinated efforts of adjacent land management units will likely be required to facilitate these shifts. Restructuring of plant communities to "preferred microclimates" may be possible within a management unit and is already being implemented by managers as a hedge against the loss of species. It has been observed that by altering plant communities, the fundamental unit of many conservation plans, these large and small scale changes undermine contemporary conservation.

Under climate change, the scale of the plant communities may not be relevant to the designated new conservation area. Plant communities are often used as surrogates for habitat availability for one or many species. However, plant community composition and distribution is dynamic as has been demonstrated by paleoclimatic evidence showing individuals, and not communities, responding to climate change over time. These changes to plant communities will also disrupt contemporary efforts to measure ecosystem functional connectivity. Individualistic responses to climate change effects connectivity metrics because functional connectivity is typically inferred from structural connectivity of vegetation cover types. As structure changes with vegetation shifts, the challenge facing managers is to determine if functions change as well. Furthermore, contemporary conservation planning needs to incorporate shifts in populations due to climate change and consider that areas with smaller endemic populations may provide significant genotypes as well as larger genetic pools.

Although climate change is affecting ecosystems today, the uncertainty and the unpredictability of future climate impacts makes it difficult to develop plans and projects to mitigate its impacts at the scale of a contemporary conservation project. Furthermore, a climate-based conservation plan will probably be more expensive to implement than a traditional plan because the climate based plan will require more land area to buffer species against uncertain environmental changes. Despite these uncertainties in global climate models, conservation planning has to consider the future repercussions of climate change because it will likely decrease the future efficacy or eliminate all the benefits derived from contemporary conservation projects.

Future Directions and Emerging Challenges in Conservation

With the widespread and increasing loss of earth's biological diversity, conservationists face an ever-increasing list of challenges that they need to consider when developing conservation plans. Some of the most urgent and important issues facing contemporary conservationists include: (1) The development of robust place-based conservation strategies, (2) achieving conservation objectives in the face of short election cycles and increasingly crowded political agendas, (3) getting social and economic incentives aligned with conservation, (4) developing institutions to manage for ecosystem resilience and sustainable development, (5) managing for climate change, (6) restoring damaged ecosystems, (7) generating the economic and political capital necessary to achieve conservation objectives, (8) mitigating the impacts and reducing the spread of invasive species, (9) developing a stronger constituency for conservation through media and education campaigns, and (10) strengthening intergovernmental and international institutions for the conservation of biodiversity. These issues vary from the species to global scales even though conservation planners will need to focus on local to regional scales to achieve their goals of increasing conservation efficacy. They will need to understand these larger scale issues since decisions made at these scales will have repercussions on where conservation is possible and what can be practiced at the conservation project scale.

Conservation planners need to make many trade-offs, from what to conserve and where to establish conservation projects,

so conservation goals are realistic to pursue. Trade-off choices need to be made since we recognize that there are few new unexplored lands that can be converted solely to conservation. UNEP World Conservation Monitoring Centre (UNEP-WCMC) reported that 13% of the terrestrial landscape was in protected status in 2010. The competing demands for using the same land suggest that it will be difficult to significantly increase the amount of land set aside for biodiversity protection. Conservation planning has to accept that the land area is finite and must be shared with many other stakeholders.

Today the practice of conservation will need practitioners who are familiar with the knowledge and tools to integrate cultural resources, environments resilience, and economic constraints to address these competing demands. Since no conservation project can satisfy the competing demands of all of the stakeholders, conservation planning is increasingly becoming more sophisticated and now requires expertise of modelers using numerous algorithmic and optimization tools to help reconcile trade-offs among competing alternatives. Therefore, conservation planners must have tools allowing them to explore the repercussions and the trade-offs needed to be made among multiple competing demands and identify the places where resources are finite. Reaching compromise solutions may require considering a multitude of options. Economic evaluation is one of the tools that may facilitate making conservation choices but its use has to be placed in the cultural and environmental context. Today multi-objective modeling tools are being developed to examine the tradeoffs associated with maintaining ecosystem resilience (e.g., biodiversity habitat) and social resilience (e.g., whether a sustainable livelihood is possible). These models also need to factor in the vulnerability of the environments and society to climate change since these will impact conservation efficacy.

Many of the challenges mentioned already are interconnected so that as solutions are developed for one challenge, it will facilitate resolving problems faced by the other challenges. It is imperative that conservation planners are able to overcome these challenges since conservation efficacy hinges on producing the right solutions. Some of the major challenges for conservation efficacy are social. The pressures of human activity and population growth will probably be the major determinants of whether a conservation plan can be designed to successfully conserve a given species. Sometimes the scientific knowledge of how much area is needed by a species cannot be implemented in the face of these social pressures. In these cases, it is appropriate to use science to determine the necessary size of the conservation reserve, but managers must keep in mind that social and economic variables are exerting more control on whether the reserve will be able to protect its resident biodiversity.

See also: Biodiversity-Rich Countries. Conservation Movement, Historical. Ecosystem, Concept of. Hotspots. Introduced Species, Impacts and Distribution of. Keystone Species. Land-Use Issues. Resource Partitioning. Sustainability and Biodiversity. The Value of Biodiversity

References

Caughley G and Gunn A (eds.) (1995) *Conservation Biology in Theory and Practice*. Cambridge, Massachusetts: Blackwell Science.

Daily GC (ed.) (1997) *Nature's Services. Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.

Hunter Jr. ML (ed.) (1999) *Maintaining Biodiversity in Forest Ecosystems*. Cambridge, United Kingdom: Cambridge University Press.

Jianguo L and Taylor WM (eds.) (2000) *Integrating Landscape Ecology into Natural Resource Management*. Cambridge, United Kingdom: Cambridge University Press.

Loehman ET and Kilgour MD (eds.) (1998) *Designing Institutions for Environmental and Resource Management*. New Horizons in Environmental Economics: Edward Elgar Publishing.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being Synthesis. Findings of the Condition and Trends Working Group of the Millennium Ecosystem Assessment*. Washington, DC: Island Press.

Myers N (2001) Hotspots. In: Levin SA (ed.) *Encyclopedia of Biodiversity,* vol. 3. pp. 371–381, San Diego: Academic Press.

NRC (1992) Global environmental change: Understanding the human dimensions In: Stern PC, Young OR, and Druckman D (eds.) *Committee on the Human Dimensions of Global Change*. Washington, DC: National Academy Press.

Primack RB, Bray D, Galletti HA, and Ponciano I (eds.) (1998) *Timber, Tourists, and Temples: Conservation and Development in the Maya Forest of Belize, Guatemala, and Mexico*. Washington, DC: Island Press.

Trombulak SC and Baldwin RF (2010) Introduction: Creating a context for landscape-scale conservation planning. In: Trombulak SC and Baldwin RF (eds.) *Landscape-scale Conservation Planning*, 1st edn. pp. 1–16, New York: Springer.

Vogt KA, Gordon J, Wargo J, *et al.* (1997) *Ecosystems: Balancing Science with Management*. New York: Springer-Verlag.

Vogt KA, Larson BC, Vogt DJ, Gordon JC, and Fanzeres A (1999) *Forest Certification: Roots, Issues, Challenges and Benefits*. Boca Raton, Florida: CRC Press.

Conservation Genetics

Katie Elizabeth Frith and A Rus Hoelzel, Durham University, Durham, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allele Variant of a gene at a given locus.
Chain-termination sequencing Method for DNA sequencing that employs four reactions, one for each nucleotide, and terminates the copy at each occurrence of a given base.
Coalescent Looking back in time through a genealogy to where two genetic lineages come together in a common ancestor.
DNA slippage Polymerase error leading to a change in the length of a simple repetitive sequence.
Genetic drift Stochastic changes in allele frequencies in a finite population.
Genetic load Extent to which an individual is inferior to the best possible individual in the population as a consequence of the genes it contains.
Microarray Collection of microscopic DNA spots attached to a solid surface; used in the analysis of gene expression or genetic variation.
Microsatellite locus Sequence of 2–4 base pairs repeated end on end.
Neutral theory That the majority of nucleotide substitutions result from the random fixation of neutral or nearly neutral mutations.
Phenotype Observable characteristics or traits of an organism.

Beginnings

Charles Darwin understood that the motive force of evolution must be a process that reflected the diversity of individuals within populations of a given species. Inspired by the writings of economist Thomas Robert Malthus, Darwin saw that the exponential potential of reproductive growth would be kept in check by limited resources. As part of that process, some individuals would be better suited to survive and reproduce than others. These ideas are at the core of what we think of as conservation genetics today. An essential belief is that diversity in populations must be retained to allow adaptation to changing environments by the process Darwin called "natural selection." Diversity at the population level is the raw material of evolution by natural selection, and the confirmation of this has been evident to animal and plant breeders for a long time. Even so, it was not until the mid-1960s that the reality of intraspecific diversity at the population level became widely appreciated. The turning point was a study by Hubby *et al.* (1966) in which they described the gel electrophoresis of enzyme variants, referred to as "allozymes." Recognizing that proteins (including enzymes) have charge properties that affect their migration through an electrical field, and that through the manipulation of the catalytic properties of enzymes a dye can be produced to reveal the presence of a specific gene product (a "polypeptide"), they provided a simple and reproducible method that could reveal molecular genetic diversity at the population level.

In no time, keen population geneticists were grinding up pretty much anything they could get hold of and producing a bourgeoning record of diversity in natural populations for an expanding range of taxa. Soon the first studies that explicitly considered the importance of conserving genetic diversity were published. In the late 1960s and early 1970s, these studies were predominantly about conserving genetic resources for their economic value. For example, Orrewing (1969) wrote of developing a program of "genetic improvement" for commercial forests in British Columbia, and Turner (1971) wrote about the conservation of genetic resources in domestic animals. These studies extended to various aspects of plant and animal breeding and included studies associated with fisheries (e.g., Moller, 1969). However, even at this early stage, there were studies that concerned themselves primarily with the management of natural populations – for example, a study on the genetic diversity of declining peregrine falcon populations (White, 1969). A significant contributor to this early work was Otto Frankel, by then a septuagenarian. He had devoted his earlier career to plant breeding and initially built on that experience. But in 1974, Frankel published a paper that spelled out in the clearest of terms a dominant ethos of modern conservation genetics. In a paper titled "Genetic Conservation: Our Evolutionary Responsibility," he wrote, "Wild species, increasingly endangered by loss of habitats, will depend on organized protection for their survival. On a long-term basis this is feasible only within communities in a natural state of continuing evolution, hence there is an urgent need for exploration and clarification of the genetic principles of conservation." (Frankel, 1974).

This set the stage and the tone for the next decade or so. The emphasis was very much on endangered species and on documenting levels of diversity toward a better understanding of the implications of lost diversity. Proponents of the neutral theory of evolution were describing not only the role of genetic drift in the loss of diversity but also the relationship between population size and the rate at which diversity was lost (Kimura, 1968; Nei *et al.*, 1975). Small populations lose diversity quickly (at a rate of $1/2N$ each generation), so exploited populations were investigated for evidence of an anthropogenic impact. In 1974, Michael Bonnell and Robert Selander investigated genetic diversity at 24 allozyme loci for the northern elephant seal (*Mirounga angustirostris*) and found no diversity at all. Elephant seals were abundant off California in the 1970s and their populations were growing fast, but 100 years earlier the species had almost vanished. These seals had

been hunted for their blubber to produce oil for lamps among other products, and their nature made them easy to approach and therefore easy prey. The last few were thought to have been taken by collectors from the Smithsonian Institution in 1892 when an expedition found eight on Guadalupe Island after no recorded sighting for the preceding 8 years. Most of those found there that day were killed, but one survived due to increasing inclement weather that forced the expedition to abandon the site. A later study calculated through simulation analyses that there had been about 20 individuals left at the species nadir, and it further demonstrated the extent of lost diversity at various genetic markers (Hoelzel *et al.*, 1993, 2002).

Bonnell and Selander (1974) and later studies showed the potential impact of human exploitation on genetic diversity (though it was not until later that the direct impact was demonstrated through the comparison of samples before and after bottleneck events using ancient DNA; e.g., Groombridge *et al.*, 2000; Hoelzel *et al.*, 2002). The next question was, what difference does it make? Are species or populations depauperate of variation really at a disadvantage? Apart from the need to retain adaptive variation, there was the possibility that low diversity could affect individual fitness. In the early 1980s, a very influential study by Ralls and Ballou (1983) illustrated the impact of inbreeding depression (sexual recombination of deleterious alleles during consanguineous matings, exposing homozygous recessive traits) on a range of mammalian species based on data from zoological collections. However, the potential harm from inbreeding depression when close relatives mate had been recognized for some time. For example, Darwin compared self- and cross-fertilization among 57 species of plants and found that self-fertilization reduced seed production by 41% and plant height by 13% on average (Darwin, 1876). Even so, the paper by Ralls and Ballou (1983) inspired a search for evidence of inbreeding depression in "natural" populations, and there was ample evidence reviewed by Crnokrak and Roff (1999).

In the early 1980s, a number of papers illustrated the correlation between diversity at allozyme markers and various indicator measures of fitness. For example, Koehn and Shumway (1982) showed that the number of heterozygous loci correlated with O_2 uptake in the American oyster (*Crassostrea virginica*) per unit time, such that greater heterozygosity correlated with lower uptake of O_2 (greater efficiency). Around this time, the indirect measure of fitness based on fluctuating asymmetry (FA; van Valen, 1962) was also being assessed in this context, and a number of studies looked for correlations with genetic diversity. For example, Leary *et al.* (1983) assessed FA in the rainbow trout (*Oncorhynchus mykiss*) and showed that the number of asymmetric characters decreased with increasing number of heterozygous loci. To the extent that FA, the unbiased asymmetry of bilateral traits, reflects developmental stability (such that the expected symmetrical state has been disrupted), it could reflect the disruption of intergenic associations due to stochastic processes when populations are small and inbred. These and other measures could have reflected reduced fitness due to inbreeding, but the relative fitness of the heterozygote compared to the homozygote could also be a factor. The latter could come about by balancing selection, either through the advantage of the heterozygote, spatial, or temporal variation in adaptive advantage or through frequency-dependent selection (well-established examples include resistance to sickle cell anemia and selection favoring diversity at immune system genes; e.g., Penn *et al.*, 2002).

Therefore, within a decade or so, a principle had been established that recognized threats to population size (and therefore to genetic diversity), together with a good foundation of work indicating a negative impact on fitness (and by implication species survival) due to the loss of genetic diversity. In this context, Franklin (1980) proposed the "50, 500 rule." The "50" and "500" referred to effective population sizes (N_e; an evolutionarily relevant measure of population size that represents the size of an idealized population that would show the same rate of decay of heterozygosity as the observed population). The essential idea was that an N_e of $\sim$50 should be enough to avoid the short-term effects of inbreeding depression, whereas an N_e of $\sim$500 would be required to avoid the long-term erosion of genetic variation in quantitative traits (phenotypic traits involving multiple genes) with high heritability. Franklin chose to consider quantitative traits because it is a phenotypic variation that is exposed to selection, and so this type of variation could allow a population to respond to a changing environment. The genetic variance at quantitative traits can be divided among "additive" and "nonadditive" components, but most quantifiable phenotypic variation will be determined by additive variance (V_A), a correlated effect seen at multiple genes, and so he focused on this aspect. In small populations, the rate of change in V_A should be determined by the rate of lost diversity through genetic drift and by the gain through mutation. Since the rate of loss in diversity is related to N_e, and at equilibrium (between genetic drift and mutation) the change in V_A becomes zero, an approximate value for a minimum N_e in this context could be calculated from data available at the time. This figure was $N_e = 500$, though various higher and more recent estimates have been made (some as high as $N_e = 5000$).

In the 1980s the 50, 500 rule was used quite extensively, especially by managers promoting the protection of populations under the US Endangered Species Act (ESA; enacted in 1973). Two quite high-profile occasions involved bird species in protected habitats. The northern spotted owl (*Strix occidentalis*) is found especially in old growth forest in western North America from California to Canada. In 1984, an estimated 2500 breeding pairs remained, and at that time the US Forest Service instituted a management plan based on the preservation of at least 500 pairs to maintain genetic diversity. This species was eventually listed as threatened in the states of California, Oregon, and Washington under the ESA in 1990. The red-cockaded woodpecker (*Picoides borealis*) lives in colonies in pinewood forests that are at least 80 years old. It was listed under the ESA in 1973 when there were an estimated 6000 birds. In 1985, the US Fish and Wildlife Service proposed a recovery plan under which a minimum of 500 birds would be retained to maintain genetic diversity. In each case, it was the number of individual birds that was to be managed (>1000 for the owl and >500 for the woodpecker), not N_e, which can be an order of magnitude or more smaller than the census number (*see* Low Diversity and Small Populations).

Largely in response to the use of this shorthand measure in the ongoing development of management policy (though not specifically in response to the mistaken focus on census number rather than N_e), Russ Lande published a highly influential paper in 1988 titled "Genetics and Demography in Biological Conservation," in which he focused on a number of key points. First was the concern that a convenient metric (e.g., $N_e > 500$ to conserve genetic diversity) could lead to management plans that neglect essential factors, indicating the need for protecting a larger population. His focus was primarily on demographic factors such as Allee effects (when small populations experience low viability and reproduction for nongenetic reasons, such as difficulty in finding a mate, in addition to genetic reasons such as inbreeding), demographic stochasticity, edge effects, and local extinctions. The issue of N_e compared to the census size is again important here since an N_e of 500 could mean a census population size of 5000, 50,000, or more, and it is the census population size that is relevant to processes like demographic stochasticity. Lande made an example of both the northern spotted owl and the red-cockaded woodpecker. In the case of the former, he pointed out that models based on stochastic demography and habitat occupancy suggest a trajectory toward extinction if only 500 pairs were preserved due to the sparseness of the habitat. He cited habitat fragmentation as the particular problem with the red-cockaded woodpecker as well, and he predicted population decline under the proposed management plan. At the same time, if the managers working with the northern spotted owl, for example, had planned to preserve an N_e of 500 (perhaps 5000 birds) rather than a census number of 1000, this would have made a difference.

Another aspect of Lande's argument was the suggestion that a gradual reduction in population size, which may often be the case, would result in relatively little inbreeding depression since selection would purge deleterious recessive alleles when they become homozygous. However, more recent studies suggest that purging is not a very efficient process and that low diversity does correlate to greater extinction risk (see Frankham, 2005; and *see* Low Diversity and Small Populations). Lande's conclusion – that demography may be more important in assessing minimum viable population sizes than population genetics – formed the basis for a long discussion on the subject and raised questions for some at the time about the utility of conservation genetics. The idea that species are typically driven to extinction before genetic factors could have an impact (referred to as the "Lande effect") is addressed in a meta-analysis by Spielman *et al.* (2004), which suggests that this is not the case.

However, the broader issue about the utility of conservation genetics has been addressed in other ways as well. As pointed out in the editorial of the first issue of the journal *Conservation Genetics*, the potential applications of genetic methods in support of more-effective conservation are diverse. Beyond the simple identification of levels of diversity or the estimation of effective population size, genetics can provide taxonomic identification (e.g., in support of wildlife forensics), assess introgression (e.g., between introduced and native species), identify management units for stock management (and the rate and direction of gene flow among stocks), assess behavior relevant to management (e.g., kinship and reproductive behavior), track demographic histories, assist the practical management of captive populations, and aid in the assessment of local adaptation, among numerous other aspects. The remainder of this essay will focus on modern approaches to these various aspects and begin with our current understanding of the earlier question about the impact of lost diversity on everything from organismal health to species survival.

Low Diversity and Small Populations

The effect of genetic drift on diversity in small populations, confirmed in various laboratory and natural experiments, focused minds on the urgency associated with reduced population size. As Lande (1988) reminded us, conservation biology is a "crisis discipline" with negative impacts (especially anthropogenic impacts) mounting all the time. Although he was more concerned with the extinction of small populations for demographic reasons, there is growing evidence that the loss of genetic diversity also plays a crucial role (e.g., Spielman *et al.*, 2004). However, the loss of genetic diversity is related to the effective population size (N_e) in particular. Therefore, one of the most useful roles for conservation genetics in addressing this problem has been to provide methods for the estimation of N_e. At the same time, N_e is difficult to measure using genetics because of the stochastic nature of the key relevant processes (inbreeding and genetic drift) and the various factors that influence those processes (e.g., mutation, selection and migration).

Given full information about population demographic histories, reproductive behavior (especially sex-biased skew in reproductive success), variance in family size, and various other factors, it would be possible to calculate N_e without genetic methods, but these details are rarely fully known. Instead, genetic markers can be used to estimate current N_e if the "systemic" forces of mutation, migration, and selection can be ruled out. In that case, changes in allele frequency over time will be due to the stochastic process of genetic drift, and since there is a known relationship between genetic drift and N_e, N_e can be estimated. The systemic forces can be controlled for by making some basic assumptions. First, for estimates of current N_e we assume a closed population (and so no migration). We also assume a short time frame (so no mutation, since mutation is a very slow process) and base analysis on neutral markers (so no selection). The first assumption is often most problematic, but the implications of violating this assumption can be assessed and the results interpreted accordingly.

Two common methods for estimating current N_e are the "temporal method" and linkage disequilibrium (LD). The temporal method further assumes random mating and discrete generations (though some models allow overlapping generations), and it considers the change in allele frequencies between generations. Allele frequency changes will be due to genetic drift and sampling effects, so accounting for the latter allows for an estimate of the effects of genetic drift and therefore N_e. LD (a nonrandom association of alleles at different loci in gametes) can be generated in finite populations by migration, selection. and genetic drift (even without physical linkage on the chromosomes). Therefore, the same assumptions again allow us to focus on genetic drift and

estimate N_e. A rather different approach is to assume a state of equilibrium between mutation and genetic drift and estimate long-term, average effective population size. This is based essentially on a simple relationship between diversity (designated as θ) and N_e, such that $\theta = 4N_e\mu$, where μ in the mutation rate. A challenge is to estimate the mutation rate accurately.

In general, estimates have shown that N_e can be much smaller than the census population size, and this is important especially if the apparent census size is large (giving the appearance of a population not at risk). Although in a review Frankham (1995) suggested a typical value of about 10%, in some cases N_e can be several orders of magnitude smaller than N (common for pelagic marine fish species; see White *et al.*, 2011) or quite close to *N* (as in some natural populations of Drosophila; see Shapiro *et al.*, 2007). Some populations apparently have naturally low levels of N_e; for example, Bourke *et al.* (2010) proposed that golden eagles in the British Isles have always had a low N_e despite once being more widely geographically distributed. A more extreme case would be naturally inbreeding populations, including selfing species of plants and some eusocial species such as some species of Hymenopteran insects (especially bees and ants). This raises the question of how such populations can survive the concomitant loss of diversity, especially the impact of inbreeding depression.

Because inbreeding exposes deleterious recessive alleles and expresses the associated phenotype, natural selection should purge them from the population with the loss or reduced fitness of those individuals. However, research into purging (which has largely focused on laboratory species such as *Drosophila*) suggests that the positive effects are variable and often small (see Boakes *et al.*, 2007). For example, Fox *et al.* (2008) performed serial inbreeding experiments (full-sibling mating over a few generations) in laboratory populations of a beetle (*Stator limbatus*). They found that the genetic load of recessive alleles was rapidly and efficiently purged but that despite this, 61.5% of the populations became extinct due to the stochastic fixation of deleterious alleles via genetic drift. Purging is likely to be more efficient in populations where alleles have a large deleterious effect (because selection quickly removes individuals expressing such alleles) when alleles are completely recessive (so they are not masking other deleterious alleles), when inbreeding is slow (so the population does not become extinct before purging can occur), when the population is large (so selection is more effective), when the population is naturally inbreeding (because long-term inbreeding will have reduced the genetic load), and where the population is sufficiently isolated such that immigration does not reintroduce purged deleterious mutations. For a review, see Charlesworth and Willis (2009).

Therefore, in spite of the potential for purging to reduce the risk of inbreeding, there remains strong evidence that low diversity and inbreeding constitute a significant conservation risk. For example, a meta-analysis of the effect of inbreeding depression on extinction risk in 30 mammal and bird species found that inbreeding depression decreases the time to extinction by an average of 37% (O'Grady *et al.*, 2006). Inbreeding depression has the potential to reduce any component of reproductive fitness, including birth weight, growth rate, survival, parasite resistance, and tolerance to environmental stresses (Keller and Waller, 2002). It is also likely to be more severe in harsh and stressful environments; for example, Keller *et al.* (2002) found that inbreeding depression in Galapagos finches was five times more severe in years with low food availability and high population densities. As described earlier, factors such as frequency-dependent selection or heterozygote advantage are also important in some cases and best described for the evolution of immune system genes (e.g., Nei and Rooney, 2005). In this case, individual heterozygosity can itself be positively correlated to fitness.

Assessments of heterozygosity and fitness correlations (HFCs) have revealed variable results, although there is general agreement that genetic diversity does correlate positively with fitness (Chapman *et al.*, 2009). However, a correlation may be supported in several ways. There may be some direct association between the locus and fitness, as may have been the case for some of the allozyme studies showing HFCs, since diversity at these loci reflects function. It is also possible that there are relevant functional genes linked to the neutral markers under investigation. Perhaps more common, though, is the assumption that neutral markers will reflect the general diversity of the genome and so imply genomic diversity by inference. The extent to which this may be true is not yet fully known. In a meta-analysis by Reed and Frankham (2001), there was a weak correlation between molecular and quantitative measures of genetic variation but no significant relationship between genetic diversity and life-history traits. This could be because selection acts to maintain diversity at functionally important loci or conversely to purge deleterious alleles at those loci or because of the sometimes complex interactions among genes associated with phenotypic traits. For example, Olano-Marin *et al.* (2011) tested HFCs in blue tit populations using markers located at both neutral and functional loci. For neutral markers, they found a positive relationship between heterozygosity and fitness (measured as female hatching success and recruitment), but for markers located in functional genomic regions the relationship was negative. They attributed this to "outbreeding depression" at the functional genes. This can happen when local adaptation has generated complexes of genes that are not well-adapted elsewhere. In the blue tit study, there may have been a local effect such that diversity associated with these functional genes was negatively associated with fitness, whereas the neutral markers showed a positive correlation due to the effects of inbreeding.

In general, there are data in support of the perceived risks associated with small population size and lost diversity and therefore an interest in strategies for mitigation. This is often associated with controlling patterns of connectivity. For example, populations of conservation concern are often found in highly fragmented habitats. Small habitat fragments may only support a relatively small population; depending on the species and nature of the habitat, fragmentation may inhibit gene flow. The usual mitigation policy involves either the establishment of corridors among fragments or the translocation of individuals. The former may be more likely to succeed since differentiation between neighboring populations may be relatively low (avoiding outbreeding depression). Translocations need to be carefully managed with respect to levels of

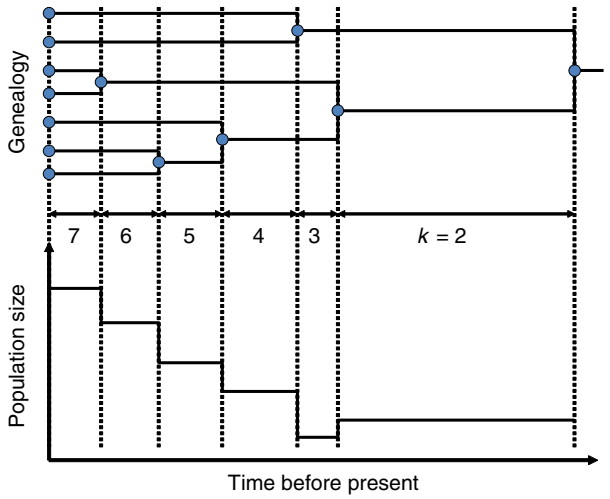


Figure 1 A "skyline plot" showing how coalescent intervals from a genealogy can be used to estimate changes in effective population size over time. The number of lineages represented at each coalescent, interval is given as *k*. Reproduced from illustrations in Pybus OG, Rambaut A, and Harvey PH (2000) An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* 155: 1429–1437.

differentiation (or adaptation to a captive environment if relevant). The idea is to "rescue" genetic diversity. Laboratory studies of genetic rescue in inbred *Drosophila* lines found that the short-term rescued populations increased in fitness, but that this effect may break down after a number of generations, especially if the N_e of the population remains low (Bijlsma *et al.*, 2010). Genetic rescue has been practiced in managed populations with mixed results. For example, cross-lineage breeding between captive Mexican wolves (an endangered subspecies of the gray wolf) increased fitness (measured by an increase in offspring number) in cross-lineage wolves (Hedrick and Fredrickson, 2010). However, for the Scandinavian wolf, genetic rescue resulted in an initial increase in fitness but then an increase of inbreeding depression in subsequent generations (Liberg *et al.*, 2005).

Another important consideration is the demographic trajectory of populations, and genetic methods can help to assess these trends as well. Wright (1931) noted that the probability that two gene copies come from the same gene copy in the previous generation is $1/2N_e$, so every generation there is a $1/2N_e$ chance of coalescence (the lineage coming together as you look backward in time). John Kingman later generalized this idea to consider multiple gene copies and showed that the interval between coalescent points is related to N_e (see Kingman, 2000), and so changes in N_e over time can be estimated based on gene genealogies. This method of tracking the changes in N_e was initially referred to as a "skyline plot" (**Figure 1**; Pybus *et al.*, 2000), and it has since been developed into quite a powerful tool that can track demography even over relatively short timescales, given the application of multiple genetic markers (see Ho and Shapiro, 2011 for a review). The method is greatly enhanced when data from ancient DNA are incorporated, providing a fuller representation of historical coalescent points.

Therefore genetic methods can facilitate conservation through the identification of populations that have small effective sizes by tracking the dynamics of populations – for example, by allowing mitigation for populations in decline – and by quantifying the impact of low diversity and the potential contribution that may have on extinction risk. This has been a central component of the growing field of conservation genetics since its inception. However, in recent years the dominant body of work has instead been on the next step up in the hierarchical structure of diversity: differentiation among populations.

Population Structure

The initial variation documented by allozyme studies at the population level was already a revelation, but the full details were appreciated only after DNA sequencing data became available. For example, Kreitman (1983) sequenced part of the alcohol dehydrogenase (ADH) locus in five natural *Drosophila melanogaster* populations and revealed 42 silent (i.e., no amino acid change) polymorphisms. A well-known two-allele allozyme system (comprising a "fast" and a "slow" allele) was thereby shown to reflect considerable neutral variation as well, and the importance of variation among populations has since become increasingly apparent. Species are subdivided into populations connected to one another to a greater or lesser degree by gene flow and differentiated by the processes of genetic drift and natural selection. This means that the preservation of the potential to evolve (retaining diversity as the raw material for evolution) entails conserving diversity both within and among populations. A rapidly expanding human population and increasing habitat degradation complicates this objective since these activities lead to populations becoming further subdivided and isolated (a process of "fragmentation"). In addition, natural population structure can be cryptic (e.g., without obvious barriers to gene flow between them) representing diversity that may be important as populations adapt to changing environments over time but may not be intuitive from the physical structure of the environment. As a consequence of these and related factors, quantifying the level and patterns of differentiation between natural populations has become a major focus of conservation genetics.

Genetically differentiated populations can evolve in isolation ("allopatry"), and this is the most obvious and easy to predict pattern of population structure. There can also be differentiation in sympatry (same geographic range) or parapatry (populations next to each other), though examples are less common and more controversial. Populations often show a pattern of isolation by distance (increasingly differentiated with greater geographic distance), though not always, and isolation by distance may also reflect a continuous pattern, not subdivided into populations. A useful definition of a *population* toward more effective conservation needs to incorporate the evolutionary forces of gene flow, mutation, drift, and selection. A commonly used definition is "a group of individuals living in sufficiently close proximity that any member of the group can potentially reproduce with any other member" (Waples and Gaggiotti, 2006). Such groups are "panmictic"

(free flowing) with respect to gene flow within the group, separate from other similar groups, and distinct from an "ideal" population (*see* Beginnings).

Highly variable genetic markers are now readily available for essentially any nonmodel organism, and increasingly easy and inexpensive to screen (e.g., see Pearse and Crandall, 2004). In practice, there are two main approaches for identifying genetically distinct populations. The first is based on determining the degree of genetic differentiation between predefined putative populations (the statistical rejection of panmixia). By this method, putative populations are identified a priori based on factors such as geographic isolation or apparent barriers to gene flow. The second is based on using information from multi-allelic markers to cluster individuals into groups by some methods without a priori assumptions about putative population boundaries. Investigating population structure using the first approach has been largely based on *F*-statistics (or a derivative of *F*-statistics), which are "fixation indices" that describe the partitioning of diversity in a structured population at different hierarchical levels (e.g., for the individual within a subpopulation, as F_{IS}, the inbreeding coefficient, and for subpopulations in the total population as F_{ST}; Wright, 1931). Population structure can be assessed using F_{ST}, which quantifies differentiation among populations (expressed as a value between 0 and 1, where 0 implies panmixia and 1 implies no gene flow). Although the magnitude of F_{ST} has some meaning, the primary initial question is typically about whether or not panmixia can be rejected (whether or not F_{ST} is significantly greater than zero). Wright's (1943) original "island" model considered a complete set of subpopulations, however it is as common to see pairwise population comparisons as a single measure considering the structuring of all sampled populations, and the inference about significant structure remains valid.

In some cases, the apparent level of structure will be small and the interpretation in the context of conservation strategy somewhat equivocal. For example, a study on population structure in sockeye salmon sampled from 11 different spawning sites and applied the criteria that a gene-flow rate between populations of less than 10% would be enough to render the populations independent. Because the N_e of each population was estimated to be approximately 1000, it was estimated that there would be panmixia if F_{ST} was below 0.0025 (thus reflecting gene-flow rates of less than 10%). However, F_{ST} was calculated to be 0.007, suggesting that even though differentiation was very low it was strong enough to reject panmixia and warrant the separate management of the 11 spawning-site populations (Ramstad *et al.*, 2004). At the extreme there is some question about the meaning of these very low values of F_{ST}, a question addressed directly in a study that combined population genetic assessment with an extensive capture–mark–recapture study of Norwegian cod (*Gadus morhua*; Knutsen *et al.*, 2011). The authors concluded that although the F_{ST} values were very low (average was 0.0037), they were nevertheless consistent (over 10-year replicate samples and among loci) and supported by the mark–recapture data, where individuals were most likely to be recaptured near their initial site of capture. For this example at least, the small but significant measure of genetic structure did appear to provide useful inference toward effective conservation and management. More generally, the "rejection of panmixia" method has been criticized by some because low rates of gene flow (theoretically just one migrant per generation, though usually more than this in real populations) can be enough to prevent divergence, and conversely populations may be demographically independent even when gene flow is high (see Palsboll *et al.*, 2007). White *et al.* (2011) modeled the latter question and showed that sufficient migration to make demographic trends match among populations may be too high to be detectable using genetic methods.

The "individual assignment" approach can also be based on the initial identification of putative populations, in which case the relative likelihood of assignment to sample site compared to alternative populations is assessed based on individual genotype profiles and the distribution of allele frequencies in each putative population. However, assignment to populations does not require populations to be predefined and can instead use information contained within multi-allelic markers to cluster individuals into populations based on either minimizing deviation from equilibrium assumptions (concerning the equilibrium between migration and genetic drift) or grouping populations based on similar allele frequencies (using one of several "principle components" clustering analyses). These methods are especially useful for populations that are difficult to define a priori (especially when individuals have a continuous geographic distribution) or when there is cryptic population structure between sympatric populations. For example, the European gray wolf is a highly mobile animal, with a relatively continuous distribution and no obvious geographic barriers to gene flow. However, investigation of the population structure in Eastern Europe using neutral microsatellites found two distinct population clusters, suggesting restricted gene flow and cryptic population structure (Pilot *et al.*, 2006; **Figure 2**) and implying an important change in conservation strategy for this species. Clustering approaches that quantify the degree of assignment can also allow estimates of the degree of admixture (historical gene flow between different populations), which can help with the determination of the history and extent of isolation.

Conservation managers are interested in the practical utility of population structure for defining the population segments that require management (coined *management units*, or MU, by Moritz, 1994). Although the exact definition of an MU varies (and has been refined over the years, see below), the idea is that genetically distinct populations should be managed separately to preserve locally adapted population genetic variation and independent evolutionary processes. A deeper level of divergence was defined as an *evolutionarily significant unit* (ESU) and contrasted with an MU by Moritz in his 1994 paper. The general idea has been officially recognized at the policy level for some time in the US as *distinct population segments* under the ESA, which are afforded similar protection as species. However, the application of DPSs has not been without controversy. One famous example is that of the red wolf. The red wolf was once widely distributed across the southern US, but hybridization with coyotes was thought to have resulted in the loss of almost all pure red wolf individuals. Consequently, the red wolf was afforded protection under the ESA. However, later genetic analyses suggested that

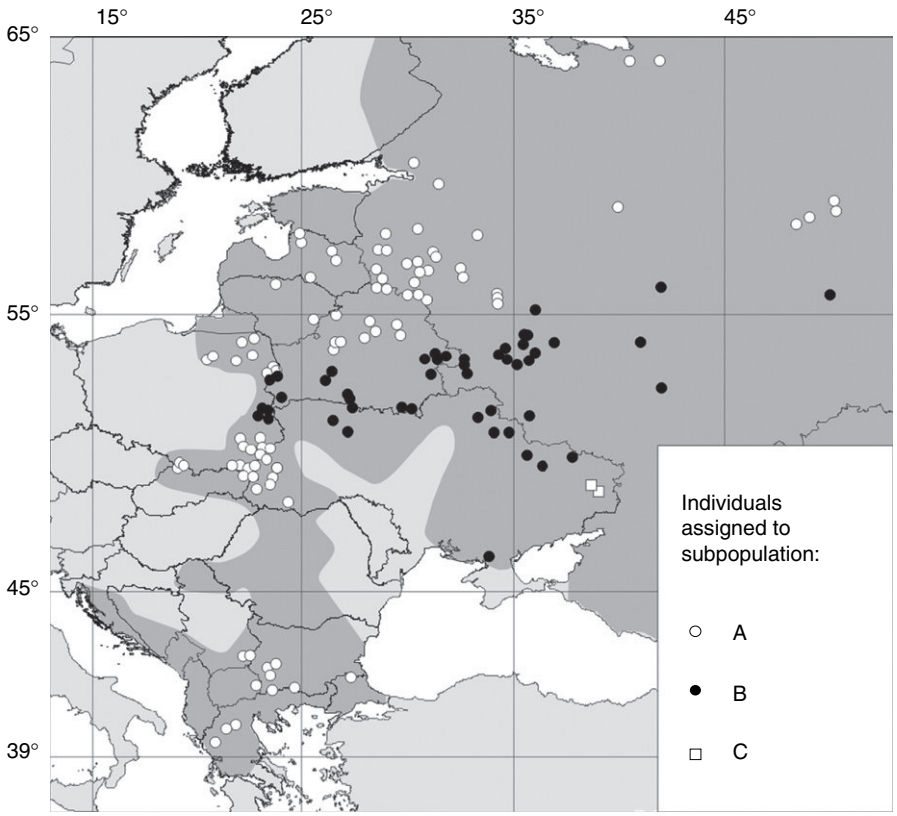


Figure 2 Assignment of individual wolves to subpopulations identified from microsatellite DNA clustering analyses. Note that distinct populations were identified despite the continuous range of wolves (dark gray area) in the region.

the red wolf was formed as a hybrid between the gray wolf and the coyote (neither of which were listed as endangered; Wayne and Jenks, 1991; Reich *et al.*, 1999), and the protection status of the red wolf was subsequently hotly disputed (see Allendorf *et al.*, 2001).

The routine assessment of MUs has been common, exemplified by fisheries management. For example, it is recognized that many salmon fisheries are comprised of different population units that have either been historically isolated or have functionally significant adaptive differences. Consequently, the need to define and manage such populations as separate stocks has been effectively applied in salmon fisheries across the globe (e.g., Beacham *et al.*, 2004). Later revisions to the concept of an MU have included the suggestion that adaptive variation and ecological specialization are more important than genetic differentiation at neutral markers (Crandall *et al.*, 2000). The debate about the relative importance of historical isolation (reflected in neutral markers) compared to functional differences is likely to be resolved eventually by our increasing ability to address both (see below) but can also be addressed in part by protecting distinct populations across a range of different habitats, thus preserving both past evolutionary processes and selective pressures (Moritz, 2002).

An essential aspect of understanding population structure is the further understanding of the pattern of connectivity among populations. Fixation indices and clustering methods allow us to understand population structure in terms of population differentiation (which is a function of gene flow), but we can also use genetics to specifically quantify the level and direction of gene flow among populations. Traditional population genetic methods for estimating gene flow used Wright's island model to calculate the effective number of genetic migrants per generation (Nm). Because of a simple relationship with F_{ST}, Nm is relatively straightforward to calculate. However, Wright's island model makes some simplifying assumptions that may not be biologically realistic (such as equilibrium between migration and genetic drift, equal subpopulation sizes, equal migration rates between subpopulations, no mutation, and no selection), and it only gives us an estimate of the magnitude of gene flow between subpopulations, providing no information about the directionality (Whitlock and Macaulay, 1999).

With increasingly powerful computers, we can now estimate migration rates much more effectively using modeling approaches based on coalescent theory (see above) to provide us with estimates of asymmetrical migration rates between subpopulations (directional gene flow). The basic principle is as follows. The coalescent models a gene genealogy for a population and works backward in time to find points on the genealogy where lineages come together (coalesce) and share a common ancestor. Separate populations, reflecting different lineages, reveal directional gene flow when the historical coalescent for an individual is in a population different from the one it was sampled in (**Figure 3**). The specific approach used depends on whether (1) the subject

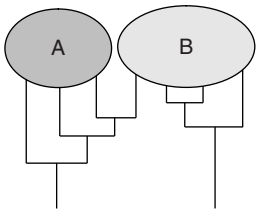


Figure 3 Illustration of an individual belonging to one lineage found in a population dominated by another lineage to show how coalescent analysis can provide information on directional migration (in this case, from population A to population B).

populations diverged from one another a long time ago such that they exchange migrants at a constant rate over evolutionary time (Beerli and Felsenstein, 2001) or (2) they diverged more recently and are still exchanging migrants with one another at a variable rate (Hey and Nielsen, 2004). Both methods have been applied successfully to determine directional migration rates, but the second class of methods is especially useful for looking at more recent patterns of gene flow. These latter methods employ an isolation with migration (IM) model, which assumed that two populations (N_1 and N_2) have diverged from a single ancestral population (N_A) at some time point (t) in the past (though note that more recent versions allow for multiple populations; Hey, 2010). Hoelzel *et al.* (2007) used the IM model to determine that ongoing low-level gene flow was occurring between distinct populations of killer whales (*Orcinus orca*) despite there being considerable population structure and different resource specializations suggesting possible isolation. This type of inference is important as it indicates the need for continuing to facilitate connectivity, even when populations appear to be strongly differentiated. Although methods that assume divergence a long time ago also assume an equilibrium between genetic drift and migration, IM is in this sense a "nonequilibrium" model, since it simultaneously estimates a point of divergence and the migration rate after that point in time (although it is an average rate following the time of divergence).

Contemporary dispersal between populations can also be estimated using assignment tests, though the relevant time frame is very short and may not reflect gene flow (unless historical admixture is being assessed). One complication is that migrant individuals are easier to detect when there is strong population structure (and hence different allele frequencies between populations), but strongly differentiated populations will not be exchanging many migrants. Conversely, migrants are difficult to detect when population structure is weak, despite migration occurring at a higher rate. Using genetics to assess contemporary population assignment is especially useful in the fight against wildlife crime – wildlife forensics. Before the widespread use of genetic methods it would be extremely difficult to identify the origin of products in trade or to determine whether specimens had been poached illegally. Genetic assignment tests can identify biological samples to species and in some cases populations. For example, Wasser *et al.* (2008) used assignment tests to determine that a large volume of elephant ivory traded in Asia and Africa originated from a common geographic area, suggesting that elephant populations in some regions were being hit particularly hard by ivory poaching.

Another factor with potential relevance to conservation is the possibility of sex bias in gene flow, in which one sex may disperse further or more frequently than the other (for a review see Lawson-Handley and Perrin, 2007). This can be assessed either by comparing biparentally inherited autosomal markers between the sexes (whereby postdispersal individuals of the dispersing sex should show weaker population assignment and population structure) or by using markers with sex-specific modes of inheritance, notably mitochondrial DNA or the W chromosome (for species like birds with a ZZ/ZW sex determination system) in females and the Y chromosome in males. For example, Kerth *et al.* (2002) investigated population structure in the Bechstein's bat using both autosomal and mitochondrial markers and found mitochondrial differentiation to be more than 60 times higher than autosomal differentiation, suggesting female philopatry and male-biased dispersal in this species.

Describing the structure of populations is useful one species at a time, but broader inference requires an appreciation of the underlying mechanisms. One approach to learning more about this is to consider how geographic and environmental variables may have differentially affected gene flow across a landscape (for a review, see Segelbacher *et al.*, 2010). Most such landscape genetic studies have two key steps: first to identify genetic discontinuities, and second to correlate these discontinuities with landscape and environmental features. Methods that identify discontinuities have generally focused on global statistics that identify large-scale patterns within a landscape. These include matrix correlations such as Mantel tests, which test for a significant statistical association between genetic and geographic distance (isolation by distance); spatial autocorrelation analyses, which test whether individuals adjacent in space are more genetically similar than those farther apart; and clustering methods, which use genetic and spatial information to group individuals into populations. In the second step, local landscape or environmental features that may act as barriers to gene flow are assessed. Various methods have been applied but typically are individual-based analyses (like assignment tests), which makes them well-suited to investigating fine scale processes. For example, Perez-Espona *et al.* (2008) used least-cost analyses to investigate the impact of landscape features on gene flow in red deer populations in the Scottish highlands. They found that landscape features significantly affected gene flow (sea lochs, slopes, roads, and forests acted as barriers whereas inland lochs and rivers promoted gene flow), explaining more of the variation than geographic distance alone. Essential components of these analyses are an accurate assessment of environmental variables and an appropriate choice of geographic scale for the focal species.

Population structure also has a temporal dimension, which may reflect historical processes or events. A common approach to investigating this has been "phylogeography," a method described in detail by its main proponent, John Avise, in his book of the same title (Avise, 2000). Phylogenies based on samples from populations distributed across a geographic range (often constructed from mtDNA sequences) are interpreted in the context of both geography and lineage history.

One system about which considerable insight was revealed from this approach was the dynamic population structure of various species during and since the last glacial maximum in Europe. The ice age forced species to relocate further south into three principal areas of refuge – Anatolia, Italy, and the Iberian peninsula. The signature of this period of isolation and patterns of mixing to various extents following the retreat of the ice and during migrations north were evident in the phylogenetic reconstructions. These processes have played an important role in the structuring of natural populations in Europe (see review in Hewitt, 2000). More recently, population dynamics and historical connectivity have been investigated using methods based on the coalescent and simulation modeling, providing even greater insight into the processes that have shaped population structure (see review in deBruyn *et al.*, 2011).

Investigating population structure using neutral markers has provided a wealth of information on demographic processes associated with genetic drift (especially related to gene flow and population size). These data on connectivity, population history, and the rate and direction of gene flow are central to the development of effective conservation strategies. However, the potential for local adaptation is also of conservation concern. Neutral markers are only informative about natural selection when linked to a functional gene. For example, White *et al.* (2010) investigated population structure in a deep sea fish (*Coryphaenoides rupestris*) in the North Atlantic and found significant population structure only at a locus that showed deviation from neutral expectations. Allele frequencies at this locus were strongly correlated with the depth at which the fish had been living, suggesting linkage to a functional gene associated with adaptation to depth. However, phenotypic traits often involve multiple genes; some effects may be associated with the control of other genes in the system, and the key differences in different environments may relate to gene expression rather than allelic variation. Therefore, although examples of linkage to neutral markers suggesting local adaptation exist, this is not a very reliable way to detect selection at functional genes.

Given a "candidate" gene that may be under selection, various tests can be applied to look for relevant signals. Selection acts on phenotypes and genotypes, and it can increase the frequency of a favorable allele (or decrease the frequency of an unfavorable one) through "directional" selection. It can also favor diversity by frequency-dependent selection on different alleles, spatial or temporal pattering to selection, or selection for heterozygous genotypes ("balancing" selection). Selection therefore affects allele frequencies, and so methods that detect departures from a neutral distribution of alleles can be used to detect selection, though departures from neutrality may also be due to demographic effects (e.g., Tajima, 1989). Another potential signature of selection can be assessed by investigating the pattern of change in codons (the three-base code that determines the amino acid sequence in a protein) within coding genes. Only some base pair changes result in a change of amino acid ("nonsynonymous" changes), whereas the rest are silent" (synonymous changes). Because the rate of nonsynonymous substitution under neutrality is expected to be equal to the mutation rate, the ratio of nonsynonymous to synonymous change (d_N/d_S) should be 1. Under negative directional selection, d_N/d_S will be less than 1, whereas under positive directional selection and under balancing selection it should be greater than 1 (see Nielsen, 2005). Screens of many genes at once have identified outliers as anonymous, possible candidates for being under selection, though this approach is controversial.

Another possibility is to screen genes at the population level and test the expectation that there should be a predictable relationship between diversity within populations and genetic differentiation between them (typically assessed using F_{ST}). If there is more differentiation than expected for a given level of diversity, then this suggests directional selection, whereas lower suggests balancing selection (see Foll and Gaggiotti, 2008). For example, the marine snail species *Littorina saxitalis* has two divergent ecotypes found in different rocky shore habitats, and there are phenotypic differences between the ecotypes. Galindo *et al.* (2009) compared multiple loci from across the genome to identify outlier loci associated with ecotype differences, and they compared population structure using outliers versus selectively neutral loci. They found population structure to be far stronger using the outliers, which suggested that selection was promoting divergence between the ecotypes despite ongoing gene flow.

One set of genes stands out as having special conservation relevance, and those are the genes involved in immune response. There is ample evidence that diversity is selected for at these loci, in response to pressure from the diversity of environmental pathogens (e.g., Klein, 1987), and this extends over a broad range of studies involving natural populations (*see* Bernatchez and Landry, 2003). There was some initial enthusiasm for a focus on these genes in particular in support of conservation (e.g., Schreiber and Tichy, 1992), though also the realization that this is only one aspect of a more complicated storey (e.g., Miller and Hedrick, 1991).

Assessing the patterns of genetic structure at neutral and selected loci in natural populations allows us to determine best strategies for conserving the natural state and to understand the processes that generate the observed patterns of structure and connectivity. However, another key objective is to help address problems that we generate as we exploit the natural world and construct roads and cities around it. Habitat fragmentation and degradation are particular concerns, because the consequent reduction in gene flow can result in increasingly isolated populations and the loss of genetic diversity (and therefore evolutionary potential) through genetic drift and inbreeding (*see* Beginnings). For example, Reed *et al.* (2009) compared historic and contemporary levels of gene flow between nine populations of a wolf spider (a species that disperses by a form of wind-driven dispersal called "ballooning"). Historically, they found that gene flow between populations occurred at a rate of 1.5 migrants per generation but that this dropped to 0.2 migrants per generation in extant populations. The dramatic decline in gene flow coincided with an increase in habitat fragmentation in the area in which the spiders were sampled, suggesting that anthropogenically mediated habitat degradation was having a negative impact. However, some studies have found habitat fragmentation to have no effect or even to increase gene flow in some circumstances, likely a consequence of differences in life history. In a review on the genetic effects of fragmentation in plants, Young *et al.* (1999) found that factors such as longevity, pollination system, and mating system all

had the potential to influence whether or not a population was adversely affected by fragmentation.

Human disturbance can also work the opposite way and force hybridizations (the interbreeding of individuals from genetically distinct populations or species) that are maladaptive (see Allendorf *et al.*, 2001). One well-known example is that of the white-headed duck, a globally threatened species that has been negatively affected by hybridization with the invasive ruddy duck. Using genetic markers, Munoz-Fuentez *et al.* (2007) were able to show that the hybridization process had begun in the last pure stronghold of the ruddy duck in Europe, but there was no evidence for widespread introgression (incorporation of foreign genes), suggesting that conservation measures could be implemented in time to save this population. Both of these examples of the potential impact of human influence show how understanding more about the behavior of individual organisms can be valuable toward the development of more-effective conservation strategies.

Individual Genotypes

In the previous sections (*see* Beginnings, Low Diversity and Small Populations, *and* Population Structure), we describe how individual fitness is correlated with genetic diversity, how genotypes can facilitate individual assignment to populations, and how this can facilitate wildlife forensic investigations. Here we consider how knowing individual genotypes can provide information on behavior or inform about processes that can influence the level and distribution of diversity within populations relevant to the development of effective conservation strategies. For example, populations may show genetic substructuring around kin groups, or a skew in reproductive success may affect levels of diversity. Dispersal behavior that shows a directional or sex bias may provide important information about population trends or habitat requirements, and food resource requirements can be assessed through the molecular identification of food species.

The identification of kin relationships is important due to the possible clustering of kin (e.g., in social groups) and because reproductive skew reduces effective population size. The most straightforward application is the identification of parent–offspring relationships. This involves subtracting the genotype of one diploid parent from the profile of the offspring in order to identify the other parent, though it is necessary to have a sufficient number of polymorphic loci to reduce the probability of chance identifications, and it is often necessary to compensate for possible genotyping errors (see Queller and Goodnight, 1989). Although parentage can sometimes be estimated by observing mating events between individuals, there are obvious limitations and logistical considerations (e.g., if there is a poor correlation between association and mating success, failed fertilization). For species such as mammals, where there is a dependent association between mother and offspring, the identification of the mother–offspring pair can serve as the basis for the identification of paternity.

The apparent extent of skew can sometimes be very much at odds with the reality. For example, the reed bunting is a monogamous bird in which parental care is shared by the sexes; however genetic paternity tests revealed that up to 55% of chicks were sired as the result of extra pair fertilizations (EPF; Dixon *et al.*, 1994). That EPFs occur under monogamy is likely to be the rule rather than the exception, though the extent of EPFs and any consequent skew will vary among populations and species (Hughes, 1998). In addition, for some species in which reproductive skew was thought to be the norm, genetics has revealed that a significant portion of paternity can be assigned to sneaky copulations with subordinate or outsider males. Arguably the most extreme polygyny is displayed by the elephant seal. Single males can defend groups of hundreds of females in harems at annual breeding colonies. The males are several times as massive as the females and fight exhaustively with competing males for access. For the northern elephant seal, paternity testing indicated that some males showed a surprisingly poor correspondence between apparent success (based on observational behavioral data) and their lower actual success based on paternity testing (Hoelzel *et al.*, 1999). For this species, it is possible that severe depletion by overhunting is in part to blame, as genetic diversity was largely lost (Bonnell and Selander, 1974), and in other species this has been shown to affect reproductive health (e.g., Wildt *et al.*, 1987). At the same time, work with the southern elephant seal (*Mirounga leonina*) suggests that alpha male success may also be related to environment. Harem holders at the Falkland Islands where the tidal range is small were more successful than in Argentina where the tidal cycle is large, possible due to the greater access a large tidal range avails peripheral males as estrous females leave the beach (Fabiani *et al.*, 2004). These differences matter toward effective conservation, since skew can affect the relationship between apparent abundance and the effective population size.

Genetics can also be used to determine kinship beyond parentage based on shared alleles that are identical by descent (IBD) as measured by the coefficient of relatedness (r). For example, in a sexually reproducing diploid species, we would expect parents and offspring or full siblings to have an $r=0.5$, half siblings or cousins 0.25, and so on. However, based on genetic markers, unless the number of markers used is very large, these estimates provide a broad distribution of values around the expected mean and so provide approximate kinship with decreasing accuracy as the level of kinship diminishes. Expectations that social groups represent associated kin are sometimes shown to be correct (e.g., African elephants, *Loxodonta africana*; Wittemyer *et al.*, 2009) and sometimes not, even when social structure is very similar (e.g., sperm whales, *Physeter macrocephalus*; Ortega-Ortiz *et al.*, 2011). Each of these species shows a hierarchical social structure where the primary association is between mother and dependent calf, followed by stable female–subadult groupings, and then temporary associations of the stable groups. However, although the stable "level 2" elephant groups show average kinship at about the "cousin" level, this is more variable for the sperm whale, and average kinship among whales did not differ within compared to between groups.

Dispersal behavior is the primary driver of population connectivity and inbreeding, and it can be determined in part by social or mating behavior. For example, a mating system such as promiscuity fosters outbreeding, and in some species it

may have evolved as a mechanism to avoid mating with kin. In a harem or hierarchical system in which only a few males breed, daughters may need to disperse to avoid inbreeding. Although purging deleterious recessives is possible (*see* Low Diversity and Small Populations), inbreeding avoidance is a powerful force. One striking example involves killer whales: both males and females are known to remain in some social groups for life. This leads to fixed mtDNA lineages within regional groups and high levels of kinship within social groups. However, Pilot *et al.* (2010) found that male-mediated gene flow was occurring during temporary associations among social groups, including among social groups whose core home ranges were quite distinct. In this way, outcrossing was being promoted even without dispersal.

Mechanisms for inbreeding avoidance in plants reflect the complexity of their breeding systems, where self-fertilization is in some cases common. Some plants are obligate outcrossers because they possess self-incompatible (S) alleles preventing self-fertilization or that from a plant with a similar S-allele genotype. However, although this promotes diversity, it can limit productivity, especially in small, isolated populations when the diversity of S alleles becomes low. The various implications of plant breeding systems in the context of conservation strategy are reviewed by Barrett (2003).

The timing, direction and any sex bias associated with dispersal can also be assessed using individual genetic profiles in much the same way that this is done with tags or natural markings. Methods for estimating the extent and direction of gene flow based on population parameters were addressed in the previous section, but in some cases it is useful to use the permanent, unique nature of genetic identity to supplement those data by tracking individuals (or their propagules). For example, Riley *et al.* (2006) used a combination of radio tracking and genetic assignment tests (*see* Population Structure) to determine that freeways acted as significant barriers to dispersal for the bobcat and coyote, despite these species possessing strong dispersal abilities. In a study investigating sex-biased dispersal behavior, observational evidence suggested that both male and female bottlenose dolphins in southeastern Australia were philopatric. However, genetic assignment tests revealed that females were correctly assigned to their natal population more often than males (Moller and Beheregaray, 2004), suggesting male-biased dispersal.

Finally, individual genotypes can be used as a tool to help understand aspects of animal ecology (and habitat requirements), for example with respect to foraging behavior. This is most commonly based on genetic analyses to determine the dietary content of stomachs or scats. For example, little information was known about the dietary ecology of the Barbastelle bat due to its nocturnal habits and the degradation of soft prey tissues during digestion. However, Zeale *et al.* (2011) used DNA barcoding to identify 37 different taxa in scats, revealing broad resource requirements. For a review of molecular dietary analysis, see Symondson (2002).

Ex Situ Conservation

Sometimes habitats become so degraded or a species so rare in the wild that it becomes necessary to implement ex-situ conservation measures. Ex-situ measures entail the captive breeding of organisms, often with a view to one day reintroducing their descendants into the wild. There are two primary objective that should be met during these efforts: first to retain as much of the natural diversity as possible, and second to avoid adaptation to the captive environment (at the expense of fitness in the wild). Various methods are employed to minimize inbreeding (and thereby maximize N_e), including equalizing family size and establishing pedigrees so that matings can be arranged that minimize kinship for the mating pair. Genetic methods can help both through the choice of initial founders and in the establishment of mating programs when pedigree data are not available.

A key challenge has often been the nature and size of the initial founder group. Especially for endangered species, there may be little option about which individuals are chosen to establish the breeding population. Genetic assessment of kinship can help early on so that the least-related pairs are bred (though the resolution is limited, *see* Individual Genotypes). Leaving them to their own devices can lead to substantial skew. For example, the golden lion tamarin (*Leontopithecus rosalia*) captive population was founded from 48 individuals and increased in size to 500, but pedigree analyses revealed that almost two-thirds of genes in the population had descended from a single breeding pair, revealing a need for managed breeding to reduce inbreeding (Ballou and Lacy, 1995). Subsequent managed breeding to minimize kinship between mated pairs greatly reduced the level of inbreeding for the golden lion tamarins.

Another concern is the sampling effect that occurs when the founder group is chosen (or during the bottleneck that reduced the size of the native population), potentially distorting allele frequencies. In this way, a rare recessive trait could become quite common in the captive population just by chance. A well-known example is the California condor (*Gymnogyps californianus*), which declined dramatically in the wild to a population of just 14 individuals. All individuals were brought into captivity and entered into a breeding program. The breeding program was a success, and numbers recovered sufficiently to begin reintroducing birds to the wild. However, managers noticed that a few chicks had begun to hatch with lethal deformities. Further investigation revealed the deformities to be consistent with chondrodystrophy – a form of dwarfism with an autosomal recessive genetic basis in the condor (Ralls *et al.*, 2000). Worryingly, the frequency of this recessive allele in the captive population was estimated to be ~9% – probably a consequence of the initial founder effect. Selective breeding can be used to reduce the frequency of the deleterious allele in the captive population, though as in this case, such a strategy has to be balanced against the risks associated with removing too many individuals from the breeding pool.

To avoid inbreeding and to minimize sampling effects, large founder sizes and large N_e in captive populations are key objectives. However, selection is a stronger force when N_e is large. Therefore, since adaptation to the captive environment has the potential to reduce fitness for organisms subsequently released into the wild, an intermediate effective population size may be most appropriate in captivity

(see Frankham, 2008). Araki *et al.* (2007) compared lifetime reproductive success of steelhead trout bred in captivity and released into the wild over three generations. They found that reproductive success declined by ~40% each generation for individuals from domestic stock, a process they attributed to the absence of selective pressures in captivity relevant to natural conditions. This has now been documented in a wide range of taxa, including plants, birds, amphibians, and insects (Frankham, 2008). Estimates of N_e using genetic markers may help breeders maintain an optimal captive population size. Experimental studies on populations of *Drosophila* (Margan *et al.*, 1998) compared the efficacy of a single large effective population to several smaller subpopulations for maintaining diversity and reducing adaptation to captivity. They found that the smaller populations collectively had higher diversity, higher reproductive potential, higher fitness under competitive conditions, and lower inbreeding coefficients than the equivalent effective size managed as a single larger population. Although this strategy has yet to be tested on a large scale in captive populations of conservation concern, such experimental studies suggest it may be a viable management option (Frankham, 2008).

More generally, it is important to assess the compatibility between the organisms to be reintroduced (or translocated) and the proposed destination population. Apart from concerns associated with habitat suitability, genetic compatibility is important to avoid outbreeding depression (a reduction in fitness when individuals with divergent adaptive or genomic characteristics interbreed). Sagvik *et al.* (2005) tested the effect of outbreeding depression on the common frog by crossing individuals from two different populations (one from a large pond and the other from a small, isolated pond). They found reduced body size and increased deformities in F_1 tadpoles produced by crosses between females from the large and males from the small pond. The genetic component leading to reduced fitness may or may not be reflected in diversity at neutral markers, and ultimately it would be ideal to be able to investigate relevant functional gene systems directly. However, gross chromosomal compatibility can be assessed (e.g., with respect to chromosome number) and combining populations that are strongly divergent at nuclear markers can be avoided. For example, Piertney *et al.* (2005) assessed mitochondrial sequence variation among water vole populations in the UK and found divergent lineages between Scottish and English and Welsh regions, and consequently recommended against translocation among these sites.

For the longer term, cryopreservation of animal gametes and seed banks can help preserve diversity and permit a controlled management of diversity and reintroductions based on individuals of known genotype. These collections have the potential to be substantial, given the reduced space requirements and ability to store materials in a suspended state (dormant seeds or cryopreserved eggs and sperm). For example, the Millennium seed bank project at Kew Gardens, London, currently houses seeds from 10% of the world's wild plant species and aims to increase this figure to 25% by the year 2020. This will provide a valuable resource for promoting plant genetic diversity in the future, which is clearly important given the current rates of habitat destruction and climate change.

The Future

Substantial inference has been gained toward the more effective conservation of biodiversity based on the assessment of neutral genetic diversity (or diversity that was assumed to be neutral), greatly facilitated by advancing technologies for the more-efficient analysis of molecular markers. Allozymes reflected partial information because only coding regions were assessed and only some DNA changes were reflected in the different allozyme charge properties that could be detected as different alleles. "Chain-termination" sequencing allowed all base pairs to be analyzed for a given locus, but sequencing large regions by this method was time consuming. Something of a revolution in resolution was achieved through the analysis of repetitive DNA regions, which evolve by evolutionary processes (especially "DNA slippage") that are faster than point mutational change altering DNA sequence. These markers are also relatively easy to analyze, and "microsatellite DNA" loci especially remain widely used.

However, a new era began once the first entire human nuclear genome was sequenced (Venter *et al.*, 2001) based on traditional sequencing technology, and both the potential of having these data and the great expense and difficulty of generating it by existing technologies became evident (the Human Genome Project cost billions of dollars and took 10 years to complete). The new objective became the analysis of diversity across the genome, and at first this was facilitated by methods that looked at random loci defined by largely uncharacterized variation (amplified fragment length polymorphisms, or AFLP) and then by the more precise assessment of single nucleotide polymorphisms (SNP). Screening hundreds or thousands of SNP markers (e.g., using "microarray" technology; see Stoughton, 2005) allowed an exact assessment of genotype but still only represented a small proportion of the total genome. However, in the 21st century, the difficulty and cost associated with sequencing whole genomes was dramatically decreased by the advent of new technologies, referred to now as "next generation sequencing" (for a review, see Mardis, 2008). These methods typically sequence many fragments in parallel and use a new type of chemistry associated with the quantification of light released when DNA strands are extended ("pyrosequencing"; see Mardis, 2008).

The applications of these genomic technologies in conservation genetics are largely in two categories (see review in Allendorf *et al.*, 2010). The first simply takes advantage of the much greater power availed by the very large number of neutral polymorphic sites revealed (e.g., to obtain more accurate estimates of N_e, migration rate, population structure (units of management), and populations dynamics). It would also permit better assessment of introgression toward management against breeding hybrid organisms. Although these applications permit greatly improved resolution, the other category offers something new. Although work on candidate genes has allowed tests for selection at functional genes, sequences from whole genomes will eventually allow functional gene systems important to conservation to be specifically identified and assessed. However, single gene effects based on mutational differences in that gene are likely to be relatively rare. Instead, phenotypes are encoded by multiple genes, genes may play different roles in different genomic or environmental

contexts, and phenotype is usually determined to some extent by environment. Further, some genes control the expression of other genes (so called nonadditive interactions), and these expression profiles may be the relevant measure for assessing local adaptation. In some cases, differences in phenotype will be due to the ability of organisms with a given genotype to express different phenotypes, known as "phenotypic plasticity". Genomic methods make the analysis of all of these aspects possible but not less complicated. The researcher investigating a whole genome is typically confronted with billions of bases of DNA, and one of the greatest challenges for the future will be finding the means to efficiently and effectively analyze those data. In spite of the limitations so far, the potential is great, and already significant advances are being made. Hundreds of nonhuman whole genome sequences are now available or are in progress at different levels of resolution, and a consortium of scientists is promoting the sequencing of 10,000 species (the Genome 10K project; Haussler *et al.*, 2009). A major objective of the Genome 10K project is to promote biodiversity conservation.

Work to promote the conservation of functional diversity will be able to employ methods that permit the identification of loci under selection on a comparatively large scale. As described previously, there is a predictable relationship between diversity within populations and genetic differentiation between them (see Foll and Gaggiotti, 2008). When thousands of loci can be investigated at the sample time, then an assessment of this relationship can be undertaken for all loci at once, and the significance of outliers assessed, indicating candidate genes that may be under selection (**Figure 4**). So, rather than starting with the gene based on known function, it is possible to search for genes showing a signal for selection and then consider their function in the context of quantifiable differences in environmental or ecological factors. One of the first studies to apply this approach using genomic technologies involved a study of marine and freshwater populations of the three-spine stickleback (*Gasterosteus aculeatus*; Hohenlohe *et al.*, 2010). The authors used a technique that sequences a subset of DNA from across the genome (but still including millions of base pairs) to identify SNP loci that can be screened among individuals at the population level. Hohenlohe and co-authors identified more than 45,000 SNPs and screened them among 100 fish from both marine and freshwater populations. They identified numerous loci that showed evidence for selection and found parallel changes that suggest a common signature for adaption to freshwater, indicating that large random mating marine populations had repeatedly given rise to freshwater populations, and that the adaptive mechanisms permitting this were comparable for independent events. This and similar studies in future will provide conservation biologists with key information about how biodiversity can best be protected. It will likely be possible to identify genes' relevant major processes such as climate change, and it is already clear that populations showing little differentiation at neutral genetic markers may be differentiated at loci that are under selection (e.g., White *et al.*, 2010), reflecting functional diversity that should be conserved. Conservation genetics is a field that is especially well-served by advancing technologies and is of increasingly critical importance as anthropogenic impacts increase and natural populations decline.

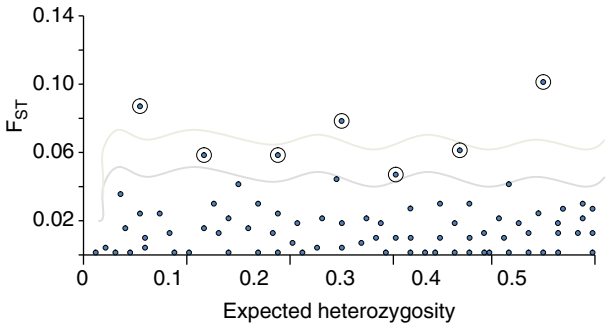


Figure 4 Under neutral expectations, there is a predicted relationship between the diversity within and the genetic distance between populations. This illustration shows how multiple loci can be screened for this relationship, where outliers outside 95% (light gray line) or 99% (dark gray line) confidence limits indicate loci likely under selection (ringed blue dots in the figure).

Appendix

List of Courses

1. Conservation Biology
2. Molecular Ecology
3. Population Genetics
4. Evolutionary Biology
5. Biogeography

See also: Captive Breeding and Reintroduction. Captive Breeding and the Evolutionarily Significant Unit. Conservation Biology, Discipline of. Diversity, Molecular Level. Ecological Genetics. Evolution in Response to Climate Change. Evolution, Theory of. Genetic Diversity. Habitat Loss and Fragmentation. *In Situ*, *Ex Situ* Conservation. Inbreeding and Outbreeding. Loss of Biodiversity, Overview. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms. Population Genetics. Restoration of Biodiversity, Overview. Species Diversity, Overview. Threatened Species: Classification Systems and Their Applications

References

Allendorf FW, Hohenlohe PA, and Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics* 11: 697–709.

Allendorf FW, Leary RF, Spruell P, and Wenburg JK (2001) The problems with hybrids: Setting conservation guidelines. *TREE* 16: 613–622.

Araki H, Cooper B, and Blouin MS (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science* 318: 100–103.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press.

Ballou JD and Lacy RC (1995) Identifying genetically important individuals for management of genetic diversity in pedigreed populations. *Population Management for Survival and Recovery: Analytical Methods and Strategies in Small Population Conservation*. New York: Columbia University Press. pp 76–111.

Barrett SCH (2003) Mating strategies in flowering plants: The outcrossing-selfing paradigm and beyond. *Philosophical Transactions of the Royal Society Series B* 358: 951–1004.

Beacham T, Lapointe M, Candy JR, Miller KM, and Withler RE (2004) DNA in action: Rapid application of DNA variation to sockeye salmon fisheries management. *Conservation Genetics* 5: 411–416.

Beerli P and Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *PNAS* 98: 4563–4568.

Bernatchez L and Landry C (2003) MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? *Journal of Evolutionary Biology* 16: 363–377.

Bijlsma R, Westerhof MDD, Roekx LP, and Pen I (2010) Dynamics of genetic rescue in inbred *Drosophila melanogaster* populations. *Conservation Genetics* 11: 449–462.

Boakes EH, Wang J, and Amos W (2007) An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity* 98: 172–182.

Bonnell ML and Selander RK (1974) Elephant seals: Genetic variation and near extinction. *Science* 184: 908–909.

Bourke BP, Frantz AC, Lavers CP, Davison A, Dawson DA, and Burke TA (2010) Genetic signatures of population change in the British golden eagle (*Aquila chrysaetos*). *Conservation Genetics* 11: 1837–1846.

de Bruyn M, Hoelzel AR, Carvalho GR, and Hofreiter M (2011) Faunal histories from Holocene ancient DNA. *TREE* 26: 405–413.

Chapman JR, Nakagawa S, Coltman DW, Slate J, and Sheldon BC (2009) A quantitative review of heterozygosity fitness correlations in animal populations. *Molecular Ecology* 18: 2746–2765.

Charlesworth D and Willis JH (2009) The genetics of inbreeding depression. *Heredity* 10: 783–796.

Crandall KA, Bininda-Emonds ORP, Mace GM, and Wayne RK (2000) Considering evolutionary processes in conservation biology. *TREE* 15: 290–295.

Crnokrak P and Roff DA (1999) Inbreeding depression in the wild. *Heredity* 83: 260–270.

Darwin C (1876) *The Effects of Cross and Self Fertilization in the Vegetable Kingdom.* Cambridge University Press

Dixon A, Ross D, O'Malley SLC, and Burke TA (1994) Paternal investment inversely related to degree of extra-pair paternity in the reed bunting. *Nature* 371: 698–700.

Drummond AJ, Rambaut A, Shapiro B, and Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22: 1185–1192.

Fabiani A, Galimberti F, Sanvito S, and Hoelzel AR (2004) Extreme polygyny among southern elephant seals on Sea Lion Island, Falkland Islands. *Behavioural Ecology* 15: 961–969.

Foll M and Gaggiotti OE (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180: 977–993.

Fox CW, Sceibly KL, and Reed DH (2008) Experimental evolution of the genetic load and its implications for the genetic basis of inbreeding depression. *Evolution* 62: 2236–2249.

Frankel OH (1974) Genetic conservation: Our evolutionary responsibility. *Genetics* 78: 53–65.

Frankham R (1995) Conservation genetics. *Annual Review of Genetics* 29: 305–327.

Frankham R (2005) Genetics and extinction. *Biological Conservation* 126: 131–140.

Frankham R (2008) Genetic adaptation to captivity in species conservation programs. *Molecular Ecology* 17: 325–333.

Franklin IR (1980) Evolutionary change in small populations. In: Soule ME and Wilcox BA (eds.) *Conservation Biology: An Evolutionary Ecological Perspective.* Sunderland Massachusets: Sinauer Associates.

Galindo J, Moran P, and Rolan-Alvarez E (2009) Comparing geographical genetic differentiation between candidate and noncandidate loci for adaptation strengthens support for parallel ecological divergence in the marine snail *Littorina saxatilis*. *Molecular Ecology* 18: 919–930.

Groombridge JJ, Jones CG, Bruford MW, and Nichols RA (2000) Conservation biology: "Ghost" alleles of the Mauritius kestrel. *Nature* 403: 616.

Haussler D, O'Brien SJ, Ryder OA, *et al.* (2009) Genome 10 K: A proposal to obtain whole-genome sequence for 10,000 vertebrate species. *Journal of Heredity* 100: 659–674.

Hedrick PW and Fredrickson R (2010) Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservation Genetics* 11: 615–626.

Hewitt G (2000) The genetic legacy of the quaternary ice ages. *Nature* 405: 907–914.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution* 27: 905–920.

Hey J and Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167: 747–760.

Ho SYW and Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources* 11: 423–434.

Hoelzel AR, Fleischer RC, Campagna C, Le Boeuf BJ, and Alvord G (2002) Direct evidence for the impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *Journal of Evolutionary Biology* 15: 567–575.

Hoelzel AR, Halley J, Campagna C, *et al.* (1993) Elephant seal genetic variation and the use of simulation models to investigate historical population bottlenecks. *Journal of Heredity* 84: 443–449.

Hoelzel AR, Hey J, Dahlheim ME, Nicholson C, Burkanov V, and Black N (2007) Evolution of population structure in a highly social top predator, the killer whale. *Molecular Biology and Evolution* 24: 1407–1415.

Hoelzel AR, Le Boeuf BJ, Reiter J, and Campagna C (1999) Alpha-male paternity in elephant seals. *Behavioural Ecology and Sociobiology* 46: 298–306.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, and Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* 6: e1000862.

Hubby JL and Lewontin RC (1966) A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics* 54: 577–594.

Hughes C (1998) Integrating molecular techniques with field methods in studies of social behavior: A revolution results. *Ecology* 79: 383–399.

Keller LF, Grant PR, Grant BR, and Petren K (2002) Environmental conditions affect the magnitude of inbreeding depression in survival of Darwin's Finches. *Evolution* 56: 1229–1239.

Keller LF and Waller DM (2002) Inbreeding effects in wild populations. *TREE* 17: 230–241.

Kerth G, Mayer F, and Petit E (2002) Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*). *Molecular Ecology* 11: 1491–1498.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature* 217: 624–626.

Kingman JFC (2000) Origins of the coalescent 1974–1982. *Genetics* 156: 1461–1463.

Klein J (1987) Origin of major histocompatibility complex polymorphism: The trans species hypothesis. *Human Immunology* 19: 155–162.

Knutsen H, Olsen EM, Jorde PE, Espeland EH, Andre C, and Stenseth NC (2011) Are low but statistically significant levels of genetic differentiation in marine fishes "biologically meaningful"? A case study of coastal Atlantic cod. *Molecular Ecology* 20: 768–783.

Koehn RK and Shumway SE (1982) A genetic/physiological explanation for differential growth rate among individuals of the American oyster, *Crassostrea virginica*. *Marine Biology Letters* 3: 35–42.

Kreitman M (1983) Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* 304: 412–417.

Lande R (1988) Genetics and demography in conservation biology. *Science* 241: 1455–1460.

Lawson-Handley LJ and Perrin N (2007) Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology* 16: 1559–1578.

Leary RF, Allendorf FW, and Knudsen KL (1983) Developmental stability and enzyme heterozygosity in rainbow trout. *Nature* 301: 71–72.

Liberg O, Andren H, Pedersen HC, *et al.* (2005) Severe inbreeding depression in a wild wolf *Canis lupus* population. *Biological Letters* 1: 17–20.

Mardis ER (2008) Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics* 9: 387–402.

Margan SH, Nurthen RK, Montgomery ME, *et al.* (1998) Single large or several small? Population fragmentation in the captive management of endangered species. *Zoo Biology* 17: 467–480.

Miller PS and Hedrick PW (1991) MHC polymorphism and the design of captive breeding programs – simple solutions are not the answer. *Conservation Biology* 5: 556–558.

Moller D (1969) The relationship between arctic and coastal cod in their immature stages illustrated by frequencies of genetic characters. *Fiskeridirektoratets Skrifter Serie Havundersokelser* 15: 220–233.

Moller LM and Beheregaray LB (2004) Genetic evidence for sex-biased dispersal in resident bottlenose dolphins (*Tursiops aduncus*). *Molecular Ecology* 13: 1607–1612.

Moritz C (1994) Defining "evolutionarily significant units" for conservation. *TREE* 9: 373–375.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematics Biology* 51: 238–254.

Munoz-Fuentez V, Vila C, Green AJ, Negro JJ, and Sorenseon NG (2007) Hybridization between white-headed ducks and introduced ruddy ducks in Spain. *Molecular Ecology* 16: 236–269.
Nei M, Maruyama T, and Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution* 29: 1–10.
Nei M and Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics* 39: 121–152.
Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics* 39: 197–218.
O'Grady JJ, Brook BW, Reed DH, Ballou JD, Tonkyn DW, and Frankham R (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation* 133: 42–51.
Olano-Marin J, Mueller JC, and Kempenaers B (2011) Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. *Molecular Ecology* 20: 4028–4041.
Orrewing AL (1969) Development of a program for genetic improvement of douglas-fir in British Columbia. *Forestry Chronicle* 45: 395.
Ortega-Ortiz JG, Engelhaupt D, Winsor M, Mate BR, and Hoelzel AR (2011) Kinship of long-term associates in the highly social sperm whale. *Molecular Ecology* http://dx.doi.org/10.1111/j.1365-294X.2011.05274.x.
Palsboll PJ, Berube M, and Allendorf FW (2007) Identification of management units using population genetic data. *TREE* 21: 11–16.
Pearse DE and Crandall KA (2004) Beyond FST: Analysis of population genetic data for conservation. *Conservation Genetics* 5: 585–602.
Penn DJ, Damjanovich K, and Potts WK (2002) MHC heterozygosity confers a selective advantage against multiple-strain infections. *PNAS* 99: 11260–11264.
Perez-Espona S, Perez-Barberia SJ, McLeod JE, Jiggins CD, Gordon IJ, and Pemberton JM (2008) Landscape features affect gene flow of Scottish highland red deer (*Cervus elaphus*). *Molecular Ecology* 17: 981–996.
Piertney SB, Stewart WA, Lambin X, Telfer S, Aars J, and Dallas JF (2005) Phylogeographic structure and postglacial evolutionary history of water voles (*Arvicola terrestris*) in the United Kingdom. *Molecular Ecology* 14: 1435–1444.
Pilot M, Dahlheim ME, and Hoelzel AR (2010) Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). *Journal of Evolutionary Biology* 23: 20–31.
Pilot M, Jedrzejewski W, Branicki W, *et al.* (2006) Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology* 15: 4533–4553.
Pybus OG, Rambaut A, and Harvey PH (2000) An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* 155: 1429–1437.
Queller DC and Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution* 43: 258–275.
Ralls K and Ballou J (1983) Extinction: Lessons from zoos. In: Schonewald-Cox C, Chambers S, MacBryde B, and Thomas L (eds.) *Genetics and Conservation*, pp. 164–184. Manlo Park, CA: Benjimin/Cummings.
Ralls K, Ballou JD, Rideout BA, and Frankham R (2000) Genetic management of chondrodystophy in California condors. *Animal Conservation* 3: 145–153.
Ramstad KM, Woody CA, Sage GK, and Allendorf FW (2004) Founding events influence genetic population structure of sockeye salmon (*Oncorhynchus nerka*) in Lake Clark, Alaska. *Molecular Ecology* 13: 277–290.
Reed DH and Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55: 1095–1103.
Reed DH, Teoh V-H, Stratton DE, and Hataway RA (2009) Levels of gene flow among populations of a wolf spider in a recently fragmented habitat: Current versus historical rates. *Conservation Genetics* 12: 331–335.
Reich DE, Wayne RK, and Goldstein DB (1999) Genetic evidence for a recent origin by hybridization of red wolves. *Molecular Ecology* 8: 139–144.
Riley SPD, Pollinger JP, Sauvajot RM, *et al.* (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* 15: 1733–1741.
Sagvik J, Uller T, and Olsson M (2005) Outbreeding depression in the common frog *Rana temporaria*. *Conservation Genetics* 6: 205–211.
Schreiber A and Tichy H (1992) MHC polymorphisms and the conservation of endangered species. In: Moore HD, Holt WV, and Mace GM (eds.) *Symposia of the Zoological Society of London; Biotechnology and the Conservation of Genetic Diversity*, pp. 103–121. Book Series: Symposia of the Zoological Society of London.
Segelbacher G, Cushman SA, Epperson BK, *et al.* (2010) Application of landscape genetics in conservation biology: Concepts and challenges. *Conservation Genetics* 11: 375–385.
Shapiro JA, Huang W, Zhang C, *et al.* (2007) Adaptive genic evolution in the *Drosophila* genomes. *PNAS* 104: 2271–2276.
Spielman D, Brook BW, and Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *PNAS* 101: 15261–15264.
Stoughton RB (2005) Applications of DNA microarrays in biology. *Annual Review of Biochemistry* 74: 53–82.
Symondson WOC (2002) Molecular identification of prey in predator diets. *Molecular Ecology* 11: 627–641.
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
Turner HN (1971) Conservation of genetic resources in domestic animals. *Outlook Agriculture* 6: 254–260.
van Valen L (1962) A study of fluctuating asymmetry. *Evolution* 16: 125–142.
Venter JC, Adams MD, Myers EW, *et al.* (2001) The sequence of the human genome. *Science* 291: 1304–1351.
Waples RS and Gaggiotti OE (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15: 1419–1439.
Wasser SK, Clark J, Drori W, *et al.* (2008) combating the illegal trade in African elephant ivory with DNA forensics. *Conservation Biology* 22: 1065–1071.
Wayne RK and Jenks SM (1991) Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature* 351: 565–568.
White CM (1969) Is there a genetic continuity concerned in eyrie maintenance? In: Hickey JJ (ed.) *Peregrine Falcon Populations – Their Biology and Decline*, pp. 391–397. Madison: University of Wisconsin Press.
White TA, Fotherby H, and Hoelzel AR (2011) Comparative assessment of population genetics and demographic history of two congeneric deep sea fish species living at different depths. *Marine Ecology Progress Series* 434: 155–164.
White TA, Stamford J, and Hoelzel AR (2010) Local selection and population structure in a deep-sea fish, the roundnose grenadier (*Coryphaenoides rupestris*). *Molecular Ecology* 19: 216–222.
Whitlock MC and McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity* 82: 117–125.
Wildt DE, Bush M, Goodrowe KL, *et al.* (1987) Reproductive and genetic consequences of founding isolated lion populations. *Nature* 329: 328–331.
Wittemyer G, Okello JBR, Rasmussen HB, *et al.* (2009) Where sociality and relatedness diverge: The genetic basis for hierarchical social organization in African elephants. *Proceedings of the Royal Society Series B: Biological Sciences* 276: 3513–3521.
Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97–159.
Wright S (1943) Isolation by distance. *Genetics* 28: 114–138.
Young A, Boyle T, and Brown T (1999) The population genetic consequences of habitat fragmentation for plants. *TREE* 11: 413–418.
Zeale MK, Butlin RK, Barker GLA, Lees DC, and Jones G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources* 11: 236–244.

Conservation Movement, Historical

Curt Meine, Aldo Leopold Foundation, Baraboo, WI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Anthropocentric Human-centered in perspective, especially with regard to the value of the natural world.

Biocentric Valuing the existence and diversity of all biological species.

Conservation biology An integrative approach to the protection and management of biological diversity that uses appropriate principles and experiences from the natural sciences, the social sciences, and various resource management fields.

Preservation The school of conservation philosophy that emphasizes the protection of natural features and landscapes from human exploitation.

Progressive Era The period in American political history, especially coinciding with the presidency of Theodore Roosevelt (1901–1909), in which the conservation movement gained definition.

Sustainable In general, capable of meeting economic goals in a manner that does not degrade the quality of the underlying ecosystem.

Sustained yield The management and harvesting of renewable resources in a manner that does not exceed their rate of replacement or reproduction.

Utilitarian The school of conservation philosophy that emphasizes the value of the natural world in contributing to human well-being, especially in terms of economic standard of living.

The Background of Conservation

As a conscious goal of citizen action, public policy, scientific research, and professional endeavor, conservation first gained coherence in the late nineteenth and early twentieth centuries. Over the decades, the conservation movement has evolved in response to varied forces: Emerging findings in the natural sciences, new environmental threats, shifts in philosophical assumptions and esthetic standards, novel technologies, expanding legal mandates, and constantly changing social, economic, and political conditions. The complex interplay of these forces has produced a movement that remains in flux and whose goals continue to evolve, even as the world's ecosystems face increasing challenges to their diversity and integrity.

Prehistoric Precedents

The term conservation acquired its modern meaning in the early twentieth century with the rise of the Progressive Era conservation crusade in the USA. However, as an expression of cultural commitment to an enduring and resilient relationship with the natural world, conservation has much deeper historic and prehistoric roots.

Prehistoric peoples did not live in a simple or constant state of peaceful coexistence with and within their natural surroundings. Converging lines of evidence from paleontology, paleoecology, archeology, and anthropology suggest in fact a sobering picture of the human past: That the dispersal of the human population over the past 120,000 years has been accompanied by spasms of continental and insular extinction (among other forms of environmental degradation). The best known of these prehistoric extinction events involves the disappearance of mastodons, ground sloths, camels, giant beaver, and other large North American mammal species coincident with a major wave of human colonization between 12,500 and 10,000 years ago. The same general pattern of extinction has been traced back over the centuries as humans moved out of Africa and Eurasia into Australia, Polynesia, and island groups around the world. The implications of this pattern of "dreadful syncopation" are controversial, and research continues on the exact sequence of cause and effect, the interaction of probable causal factors, and the precise mechanisms of extinction (MacPhee, 1999).

However, in contrast to this record of prehistoric anthropogenic extinction, there is the countervailing experience of many cultures that have achieved relatively sustainable ways of life. Native peoples in landscapes throughout the world developed sophisticated belief systems and resource-use traditions that recognized their connections to and dependence on the nonhuman world. In the past, these traditions allowed people to adapt to even extreme environments throughout the world. For example, the Maasai of East Africa, practiced livestock herding in a manner that allowed them to coexist with the native megafauna of their semiarid homeland. Arctic dwellers coexisting with caribou herds and marine mammals, shifting cultivators and agroforesters in the humid tropics, small-scale farmers in temperate zones, fishing villages in coastal, riverine, and wetland environments: These and other traditional cultures developed social systems that balanced human needs against the capacity of the natural environment to meet those needs. Some of these traditions have survived into the modern era, though they must contend with increasing pressures from population growth, diminished resources, altered land tenure systems, climate change, and rapid economic and technological change.

The resource-use practices of native peoples had significant impacts on the biota, though these impacts varied in intensity over time and according to ecosystem type and taxonomic group. People employed fire to modify vegetation and to concentrate populations of game animals. They fished,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00029-0

hunted, and gathered in ways that affected species populations, animal behavior, and the dynamics of interspecific relationships. They domesticated plant and animal species, and altered natural hydrological systems through irrigation. In some cases – for example, the deforestation of Mediterranean watersheds and the development of hydraulic cultures in arid and semiarid lands – large-scale resource exploitation caused fundamental changes in biological diversity and ecosystem function. In other cases, the limited rate and scale of exploitation resulted in relatively minor long-term ecological impact.

Humanized landscapes, in sum, have experienced varying degrees of anthropogenic environmental change (Ellis, 2011). However, in many if not most native cultures, social mechanisms evolved to reinforce an attitude of respect and deference toward nature, to evoke nature's bounty, and to sanction appropriate human use of that bounty. These mechanisms included hunting, planting, and herding rituals; stories and myths involving social and ecological relationships; fertility and birth control practices; the recognition of sacred spaces; and the invocation of taboos and totems. The conservation movement might be characterized as modern society's conscious effort to develop and exercise analogous social practices and restraints to guide its relations with the nonhuman world.

Historic Precedents

Contemporary environmental dilemmas have prompted modern scholars to reexamine conservation traditions in texts and stories from animist, Native American, Buddhist, Hindu, Islamic, Judeo-Christian, and other faith traditions (Taylor, 2005). For example, in the Judeo-Christian tradition, the biblical injunction to "fill the earth and subdue it, and have dominion over every living thing that moves on the earth" has often been cited as providing justification and rationale for centuries of environmental exploitation in the Western experience. More recently, scholars have returned to such textual sources to identify alternative traditions of stewardship and respect for creation.

Since ancient times, observers have recorded instances of environmental deterioration due to human action. For example, Plato compared the deforested mountains of Attica to the "bones of a wasted body … the richer and softer parts of the soil having fallen away, and the mere skeleton being left" (Dubos, 1980). Evidence of early conservation practices can be found in the history of both Western and non-Western cultures. These include efforts to protect particular species and special lands, to maintain populations of wild plants and animals, and to sustain the productivity of agroecosystems. For example, Leopold (1933) cited Marco Polo's narrative account of hunting restrictions and provisions for game bird populations in thirteenth-century China as "the first clear record of a well-rounded system" of wildlife conservation.

European traditions of forestry and game-keeping date back to the Middle Ages and beyond. The establishment of royal game preserves and forests on the land estates of feudal Europe led to the development of customs and formalized laws regulating hunting and use of the forests. In the eleventh century, William the Conqueror set up in the newly conquered England "many deer preserves and also enacted laws/That whoever killed a hart or hind/Should be blinded./He also placed a ban on harts, also on boars" (Harrison, 1992). The high-handedness of the Norman conquerors in establishing privileged use of the forests gave rise to a venerable tradition of local resentment over centralized resource management. However, as codified in forest law, these measures allowed England's forests to survive and regenerate. On the continent, Germany and France, in particular, developed silvicultural systems and techniques that prevented wholesale destruction of the forest estate.

Europeans carried these "proto-conservation" traditions to their expanding colonial empires. The change in jurisdiction over natural resources, from native peoples to colonial and state governments, had profound implications for social systems and ecosystems alike (Cronon, 2003). Colonial domination, coupled with rising populations and industrialization, disrupted traditional patterns of land tenure and resource use. In many regions, native peoples became increasingly alienated from their landscapes, whereas Western conservation ideas and practices were slow to adapt to the new environments.

In the New World, European colonists encountered a landscape of abundance that admitted profligate resource use. Unchecked resource exploitation was the norm as European settlement of North America proceeded. Despite this record, seeds of the later conservation movement were apparent in early efforts to protect wild game populations, forests, soils, and special natural features. Early measures to conserve game include decrees to protect the cahow (*Pterodroma cahow*) and green turtle (*Chelonia mydas*) in Bermuda (1616–1622); the first closed hunting season on white-tailed deer (*Odocoileus virginianus*) in Massachusetts (1694); closed seasons on several species of game birds in New York (1708); and the institution of game warden systems in Massachusetts (1739) and New York (1741) (Matthiessen, 1959). The first federal game law, enacted in 1776, mandated closed seasons on deer in all the American colonies except Georgia. Through the 1800s, the new American states intermittently passed laws establishing closed seasons, prohibiting hunting of nongame birds, and placing bounties on predators.

Exploitation of North America's extensive native forests – for conversion of land to agriculture, for construction and ship-building materials, for charcoal and domestic fuel supplies – drove economic development through much of the colonial and early American era. In the late seventeenth and early eighteenth centuries, the British crown instituted the Broad Arrow policy, which reserved for the crown the right to cut the strategically important white pine (*Pinus strobus*) mast trees of the New England colonies. Although serving to restrict exploitation, these laws were intended more to secure property against unauthorized timber cutting than to protect or perpetuate the forests *per se*. Closer in intent to later forest conservation measures was William Penn's 1681 dictum that, in the process of land development, one acre of forest should be maintained for every five acres cleared.

Little heed was paid to such recommendations. By 1745, Benjamin Franklin was noting the scarcity and high cost of fuelwood in Philadelphia. President James Madison in 1818 drew attention to the "injurious and excessive destruction of timber and firewood" as the most important and regrettable force shaping the rural economy of the USA. Such

destruction continued through the nineteenth century. With the depletion of the eastern forests and the continuing settlement of interior North America, the focus of forest exploitation shifted to the Great Lakes region. In 1831, Alexis de Toqueville, while visiting Michigan, noted that Americans were generally "insensible to the wonders of inanimate nature, and may be said not to perceive the mighty forests that surround them till they fall beneath the hatchet." The swift destruction of the Great Lakes forests in fact marked a turning point in forest conservation, and in the conservation movement generally (Williams, 1989).

A parallel pattern of resource degradation and early conservation response marked the process of agricultural development in North America. By 1776, the problem of soil erosion was evident to the new nation's leading statesmen. Thomas Jefferson experimented with contour plowing to retard soil erosion. Patrick Henry, speaking before the Virginia Assembly, stated that "Since the achievement of our independence, he is the greatest patriot who stops the most gullies." George Washington noted in a 1797 letter, "We ruin the lands that are already cleared and either cut down more wood if we have it, or emigrate into the western country" (Jackson, 1985). Over the next century, destructive agricultural techniques continued to affect widely varied ecosystems across North America (and other parts of the world as well).

The preservation impulse in conservation found special expression in the North American setting, as European settlers encountered scenic landscapes and natural features unlike any in their prior experience. Jefferson celebrated Virginia's Natural Bridge in his Notes on the State of Virginia. Niagara Falls, the Hudson River valley, and the prairies, mountains, and canyons of the American West became emblematic of the esthetic quality of New World landscapes, a quality seized on by artists, promoters, and early conservationists alike. The artist George Catlin suggested after his first excursions in the West that portions of the continent's undeveloped lands were "worthy of our preservation and protection." As early as 1833, Catlin envisioned government reservation of "a magnificent park … a Nation's park, containing man and beast, in all the wild[ness] and freshness of their nature's beauty" (Nash, 1982). Such protected areas would not come into being until later in the nineteenth century, although the USA did move as early as 1832 to protect special natural features when it reserved Hot Springs in Arkansas as a federal holding.

What these disparate efforts had in common was not the conservation of biological diversity in the modern sense, but the protection or regulated exploitation of economically and esthetically important components of the landscape or ecosystem. As such, they provided the foundation on which a more coherent conservation movement began to take shape in the latter half of the nineteenth century.

Modern Origins of Conservation

Prelude to a Movement

The conservation movement emerged in response to accelerated changes in demographic and environmental conditions, and in human perceptions and understanding of the natural world, through the eighteenth and nineteenth centuries. Exploration of the world's diverse ecosystems – from the earliest voyages of discovery through the New World scientific expeditions of Alexander von Humboldt, John and William Bartram, Meriwether Lewis and George Rogers Clark, John James Audubon, and others – contributed to a golden age of natural history studies. Establishment and adoption of the Linnean system of binomial nomenclature in the 1700s allowed for an unprecedented flourishing of taxonomic research (Mayr, 1982). This, in turn, provided critical foundations for the development of evolutionary theory in the work of Alfred Russel Wallace and Charles Darwin. Although the biological sciences had not yet developed field methods for comprehending the full diversity of life, they had begun to reveal the fundamental processes through which life diversified and maintained itself.

These gains in systematics, biogeography, and evolutionary theory occurred even as the Industrial Revolution and the expansion of market economies accelerated the scale, pace, and character of human environmental impacts. Through the 1800s, the advent of more efficient technologies resulted in increasingly intensive exploitation of forests, game populations, fisheries, agricultural lands, and river systems. Traditional resource management practices and established land tenure systems were abandoned or changed to fit the emerging economies of scale. Industrial pollution and the global spread of invasive species became widespread problems for the first time.

Coincident with these scientific, cultural, and environmental changes, the Enlightenment and Romantic movements were altering Western conceptions of order, value, and beauty in the natural world. The natural philosophers of the Enlightenment stressed the smooth workings and stability of a mechanistic natural order. The Romantic philosophers and poets emphasized the unity and wholeness to be found in a spontaneously creative organic nature. Although offering different conceptions of nature, both encouraged human comprehension of natural objects and processes, and so laid the foundation for greater appreciation of human impacts on the natural world. The writings of Thomas Malthus and other early economic philosophers provided the basic framework for considering the interwoven fate of the human population, human economies, and natural resources.

In Europe, the Romantic Movement drew heavily on the experiences of New World explorers and settlers, the encounter with native peoples there and elsewhere, and the exposure to wild landscapes. In turn, adaptation of the Romantic impulse in the North American setting provided important literary and philosophical foundations for conservation. For example, the strong American identification with wild nature found early expression in the essays of the transcendentalists Ralph Waldo Emerson and Henry David Thoreau, the poems of William Cullen Bryant, and the novels of James Fenimore Cooper (Nash, 1982).

George Perkins Marsh's *Man and Nature: Or Physical Geography as Modified by Human Action* (1864) is widely considered the first great landmark in modern conservation literature. A native Vermonter, Marsh saw in the destruction of New England's forests, the latest expression of an ancient human tendency to "[derange] the original proportions between

different orders of organic life." Drawing also on his extensive personal observations of long-term landscape change in New England and the Mediterranean, Marsh argued that the human actions had caused widespread disruption of the "harmonies" of the natural world. Marsh's reasoning followed lines that would sound familiar to later generations of ecologists and biodiversity conservationists. "All nature," he wrote, "is linked together by invisible bonds, and every organic creature, however low, however feeble, however dependent, is necessary to the well-being of some other among the myriad forms of life with which the Creator has peopled the earth" (Marsh [1864], 1965).

Gaining Definition

Publication of Marsh's book provided direction and definition to the nascent conservation movement through the remainder of the 1800s and the first decade of the 1900s. In North America, dispossession of the Native American tribes, enactment of liberal land distribution policies, and the flow of settlers and capital into "virgin" landscapes resulted in an unprecedented wave of exploitation of natural resources. The effects were visible in ecosystems across the continent. These decades saw the virtual depletion of the extensive pine forests of the upper Great Lakes; overhunting of many game bird and mammal species by market hunters; rapid conversion of the Midwest's extensive prairies to agriculture; degradation of aquatic systems through overfishing, pollution, and hydrological changes; and widespread overgrazing of the semiarid western rangelands. Several well-recognized cases of species depletion and extinction – of the white pine, of waterfowl, of the bison (*Bison bison*), the Carolina parakeet (*Conuropsis carolinensis*), and the passenger pigeon (*Ectopistes migratorius*) – came to symbolize this era of unmitigated resource extraction and decline (Barrow, 2009).

However, this period also saw the first concerted efforts to address the causes and consequences of these massive changes. At first, these actions tended to focus on particular issues, species, problems, or landscapes. Sportsmen led the campaign to rein in market hunting and to institute stronger game laws at the state and local levels. Many of the nation's most prominent sportsmen banded together in 1887 to form the Boone and Crockett Club, one of the first and most effective nongovernmental organizations to become involved in shaping national conservation policy. Exploitation of plume-bearing egrets and other birds for the millinery trade mobilized sportsmen, scientists, and nature lovers to work together for reform, and prompted formation of state Audubon Societies beginning in 1896. These moves culminated in passage of the federal Lacey Act of 1900. The first important piece of modern wildlife conservation legislation, it barred interstate shipment of wild animal species taken in violation of state laws, and soon succeeded in ending the plume trade (Dunlap, 1988).

Efforts to reform American land policy proved more difficult. Through the 1800s, the nation's land allotment policies had encouraged rapid immigration, settlement, and land conversion, often accompanied by corruption and land speculation. In the midcontinent's mesic landscapes, the system of land surveying and distribution facilitated the conversion of millions of acres of native savanna and grassland to crop and livestock production. However, these policies were less tenable in the progressively drier lands of the American West, where farming was possible only near permanent watercourses, and livestock ranching required large land holdings. The geologist and explorer John Wesley Powell attempted in the 1870s and 1880s to devise a more appropriate tenure system that recognized the inherent environmental limitations and social requirements of the landscape, but the speculative momentum proved intractable. Only decades later did the US government institute more careful land management policies on the nation's remaining public domain. However, Powell's innovations, especially his commitment to a strong role for science in land management, laid important foundations for future private and public conservation planning (Worster, 2001).

The rapid destruction of forests in the upper Great Lakes between 1865 and 1900 stimulated the most forceful and effective response among early American conservation reformers. As this period began, forestry as a profession had not yet gained a foothold in the USA. However, through the 1870s and 1880s, voices opposed to the destructive "cut and run" logging practices began to emerge. These reformers included botanists Charles S. Sargent, Asa Gray, and William H. Brewer; naturalists John Muir and Increase Lapham; the European-trained forester Bernard Fernow; Secretary of Interior Carl Schurz; and Franklin B. Hough, appointed as the nation's first government forestry official in 1876.

The forestry movement gained momentum quickly in the last decade of the nineteenth century. Forest activists working through professional scientific and forestry organizations campaigned for national legislation to reform public land laws and prevent further forest destruction. These efforts culminated in 1891 with the enactment of legislation that included a provision allowing the US president to set aside public land forests as forest reserves, the germ of the nation's current system of national forests. Within weeks of its passage, President Benjamin Harrison used the law to designate forest reserves in lands adjacent to Yellowstone National Park. In the next 10 years, 47 million acres of forest land were placed in reserves.

During this period a parallel movement arose to protect and preserve special scenic landscapes, natural features, and wild spaces. Like the forestry movement, this channel of activism gained momentum in the decades following the American Civil War. In 1872, the US Congress established the world's first national park at Yellowstone. In 1885, New York created a state forest preserve in the Adirondack Mountains in order to protect its wild character as well as its watershed values. That same year, Canada designated Banff National Park, its first. The impetus toward preservation drew heavily on the enthusiasm of nature writers such as John Muir, who led the political effort that in 1890 resulted in designation of an enlarged Yosemite National Park in California and who in 1892 founded the Sierra Club (Fox, 1981).

At the end of the nineteenth century, the varied strands of conservation concern – over depleted wildlife, widespread deforestation and watershed degradation, inappropriate land development, industrial pollution, and the loss of esthetic

quality – were connected only loosely. In the American experience, political corruption and inordinately concentrated wealth had typically accompanied these environmental changes. As the new century began, these overlapping social, economic, political, and environmental concerns brought forth a more consolidated movement that gave new meaning and power to the term *conservation*. However, the growing prominence of conservation also revealed inherent tensions in the emerging movement.

Progressive Era Conservation and the Utilitarian/Preservationist Split

When Theodore Roosevelt assumed the American presidency in 1901, the stage was set for a revolution in conservation policy. Roosevelt brought to the office a formidable command of the natural sciences, extensive experience as a rancher and outdoorsman in the American West, a long-standing commitment to political reform, and boundless personal energy. In his first presidential address to the US Congress, Roosevelt spoke at length of the importance of the nation's forests, stating that "We have come to see clearly that whatever destroys the forest … threatens our well-being" (Pinchot, 1947). Conservation became a cornerstone of the progressive political movement and, for the first time in history, a high national priority.

Roosevelt's partner in political innovation was forester Gifford Pinchot (Miller, 2001). Pinchot, the first American to receive formal training in forestry, was a friend of Roosevelt's who had been active in the complex politics of the American forestry movement through the 1890s. At the time Roosevelt became president, Pinchot was in charge of the Division of Forestry in the US Department of Agriculture. The two quickly forged a close alliance to strengthen the government's role in forestry. Their partnership during Roosevelt's presidency resulted in an increase in the total acreage of the forest reserves (renamed national forests in 1905) from 60 million to 151 million acres; transfer of these lands to the Department of Agriculture under Pinchot's jurisdiction; and the creation, in 1905, of the US Forest Service to administer the new national forests.

Pinchot's Forest Service embodied the Progressive Era spirit and its approach to bureaucratic responsibilities. In contrast to Muir and others in the preservationist wing of the conservation movement, Pinchot conceived of the forest in utilitarian terms. The forests were not regarded as "reserves" to be "locked up," but as lands to be worked "for the greatest good of the greatest number for the longest time" (as the utilitarian credo put it). In practice, this meant that the forests were to be managed by a trained, professional workforce; that scientific principles were to guide the efficient exploitation and processing of forest resources; and that the wealth derived from the forests was to be equitably distributed for the common good (Hays, 1959). Applied not only to forests but also to natural resources in general, the "resource conservation ethic" provided the dominant paradigm of the early movement (Callicott, 1991). In Pinchot's words, "The first great fact of conservation is that it stands for development" (Pinchot [1910], 1967).

At the core of utilitarian conservation was the concept of sustained yield. As Aldo Leopold later (1933) observed, under the progressive conservation banner "wild life, forests, ranges, and waterpower were conceived … to be renewable organic resources, which might last forever if they were harvested scientifically, and not faster than they reproduced. "Conservation" had until then been a lowly word, sleeping obscurely in the dictionary. The public never heard of it. It carried no connotation of woods or waters. Overnight it became the label of a national issue". At the leading edge of this movement, the Forest Service became a model not only for other resource management agencies but for Progressive Era government agencies in general.

Theodore Roosevelt's conservation activism extended beyond the public forests. During his presidency, he established the nation's first wildlife refuge, at Florida's Pelican Island (1903); designated 50 additional wildlife refuges; used new presidential powers under the Antiquities Act of 1906 to set aside 18 national monuments (including the Grand Canyon); convened in 1908 a national Governors' Conference on the Conservation of Natural Resources; and created, also in 1908, a National Conservation Commission to provide continuing advice to the president and to work with the states on conservation policy. Early in 1909, Roosevelt and Pinchot sought to broaden the scope of conservation through a North American Conservation Conference and laid still more ambitious plans for a World Conservation Conference. Though the latter was unrealized, Pinchot and Roosevelt had succeeded in securing a place for conservation within both the political and actual landscape.

Absent from many of the progressive political initiatives were the voices of the increasingly influential preservationists and nature protectors. In contrast to the utilitarian view that Pinchot espoused, adherents of the "romantic-transcendental preservation ethic" emphasized wild nature's esthetic and spiritual values and the need to safeguard those values for future generations through strict prohibitions on development and manipulation (Callicott, 1991). The tensions between the utilitarian and preservationist approaches intensified as conservation assumed center stage politically. These tensions surfaced in the changing relationships among the principal players. Pinchot and John Muir had been friends, but by the late 1890s, their differing approaches to management of the forest reserves caused a rift between them that would never heal. Few from the preservation camp participated in the landmark 1908 governors' conference on conservation.

Muir maintained a respectful relationship with Roosevelt. They camped together in Yosemite in 1903 and worked together in gaining protection for the Grand Canyon as a national monument in 1906. But the schism between the conservation factions continued to widen. It came to a head in the bitter political struggle over plans to dam the Tuolumne River in Yosemite National Park's Hetch Hetchy Valley. Waged over a period of 6 years, the battle culminated in 1913 with the adoption of national legislation providing federal support for the dam. Although Muir and his colleagues lost the battle, they had aroused a national constituency in favor of protection. The growing popular and political acceptance of preservation led directly to creation of the US National Park Service in 1916 (Fox, 1981).

For all of the profound developments in conservation during the Progressive Era, scant attention was given to the state of the ecological processes and biological diversity characteristic of either wild or more humanized landscapes. That attention would come only slowly, as conservation science, philosophy, policy, and practice coevolved through the twentieth century.

The Evolution of Twentieth-Century Conservation

The Progressive Era advances fixed conservation permanently in the public mind and in public policy. However, the innovations of the period served only to highlight the fundamental questions that would challenge conservationists throughout the twentieth century: What are we conserving? To what end? For whom? How? With what scientific information and understanding? With what social adjustments, economic mechanisms, and legal tools? Moreover, the conservation movement emerged from the Progressive Era with the tension between preservationist and utilitarian approaches still unresolved. The continuing efforts to answer those questions and to resolve that tension framed the century's technical and political advances in conservation.

Institutionalizing Conservation

In the USA, the first four decades of the twentieth century saw the consolidation of the Progressive Era's conservation gains in the government, the private sector, the professions, and the academy. The spirit of reform that drove the conservation movement was now redirected toward the development of sound conservation policy and administration. However, new issues arose to keep the movement invigorated and expand its philosophical dimensions.

As conservation became institutionalized, it tended to follow the tenets of Pinchot's resource conservation ethic. By the late 1930s, the basic principles of utilitarian resource conservation had been applied not only to forests but also to other "useful" components of the biota and the landscape: rangelands, game populations, sport and commercial fisheries, scenic lands and recreational areas, agricultural soils, and river systems. New policies, laws, bureaucracies, academic disciplines, research and training programs, and professional societies arose to promote sustained yields of these various "resources." Foresters, range managers, wildlife managers, recreation planners, and soil conservationists achieved their own professional identities in these years. Amateur conservationists, meanwhile, became increasingly involved in decision making through the formation of new citizen organizations.

Forestry

As an established field, forestry provided the model and foundation for the other emerging resource management professions in the USA. With those fields still lacking formal training and employment opportunities, forestry tended to attract professionals with broad conservation interests. However, institutionalization tended over time to narrow the scope of the field to silviculture and timber harvesting. This trend was magnified during these decades as the supply of timber taken from private lands diminished and the pressures on public forests increased.

Range Management

In the American West, rapid settlement and exploitation had degraded vast stretches of arid and semiarid grassland in the late 1800s and early 1900s. Research and academic training programs followed in the wake of deteriorating range conditions, bringing scientific management to rangelands much as forestry had brought such management to forests. By the late 1930s, changes in land policy resulted in stricter regulation of the public rangelands, cessation of privatization of the public domain, establishment of a new federal bureau (the Bureau of Land Management) to oversee these lands, and the emergence of range management as a discrete field.

Wildlife Management

Although concern about the fate of game populations had often galvanized the young conservation movement, wildlife management did not carve out its own professional niche until the 1930s. Continuing declines in game populations provided the initial impetus. To arrest these declines, game protectors at first focused on enactment of stricter hunting regulations, sporadic establishment of refuges, development of artificial propagation programs, control of predators, and the introduction of exotic game species. Only in the early 1930s did a revolutionary new approach, focused on the protection and restoration of habitats so that populations might sustain themselves, begin to take hold. (Fisheries management followed a roughly parallel pattern with regard to game fish species and habitats.) As principles from ecology filtered into the field, the term *wildlife* itself came into common usage, and the idea of management quickly expanded to encompass the perpetuation of "nongame," rare, and threatened species. Although the field tended to focus primarily on game species, it provided important foundations for efforts later in the century to conserve biological diversity in general.

Recreation

As the American population became increasingly urban and mobile, outdoor recreation became a more important component of conservation. Camping, fishing, hunting, hiking, and other outdoor activities continued to grow in popularity, even after the prosperity of the 1920s gave way to the Great Depression of the 1930s. Park agencies, at all levels of government, responded by establishing and developing areas of significant esthetic and recreational value. Other land management agencies, such as the US Forest Service, began to devote greater attention to the recreational value of the lands under their jurisdiction. Although there were exceptions, recreational developers tended to override or ignore problems associated with the protection or management of the biota, and rarely incorporated scientific information in planning for increased tourism.

Soil Conservation

The rapid expansion and mechanization of agriculture in the early decades of the twentieth century brought widespread

disruption to the agricultural landscape and economy. The disaster of the Dust Bowl in the Southern Great Plains during the mid-1930s was the most visible consequence of forces that had degraded soils across North America. Intensified exploitation of farm resources through land clearing, "clean farming," poor grazing management, and wetland drainage led to accelerated rates of soil erosion, degradation of pasture and forage vegetation, loss of wildlife habitat, and aggravated cycles of siltation and flooding. In response, the US Congress in 1935 created the Soil Conservation Service (SCS) to work with farmers and other private landowners to adopt conservation practices. Under the leadership of Hugh Hammond Bennett, the SCS importantly emphasized watershed-scale approaches that integrated soil protection measures with other resource conservation goals (Egan, 2006).

Although increasingly defined by this enhanced structure of agencies, disciplines, and professions, the conservation movement remained a battleground for competing approaches and intense political activity. Even as conservation became professionalized, growing numbers of citizens became active as members of nongovernmental conservation organizations. New national groups, including the Izaak Walton League, the National Wildlife Federation, and Ducks Unlimited, joined such older organizations as the Sierra Club and the Audubon Association as significant players in shaping conservation policy (Fox, 1981).

Meanwhile, the preservationist approach found renewed vigor as the campaign to protect wildlife, natural areas, and wildlands reemerged in the 1920s and 1930s. In these years, field zoologists such as Joseph Grinnell and Olaus Murie became leaders in the effort to reverse government antipredator policies. Early ecologists, under the leadership of Victor Shelford, created in 1917 a Committee on the Preservation of Natural Conditions within the Ecological Society of America. This committee provided the nucleus for what became, decades later (1950), The Nature Conservancy. In 1924, the US Forest Service designated the nation's first "wilderness area" on a large roadless expanse of the Gila National Forest in New Mexico. In 1935, proponents of wilderness protection in the USA banded together as The Wilderness Society, the first conservation organization dedicated solely to the cause of wildland protection. Although the wilderness protection movement still emphasized primarily the esthetic, spiritual, and recreational value of wildlands, these moves signaled a shift toward greater recognition of the scientific and ecological value of undeveloped lands.

At the international level, the years prior to World War II saw increased recognition of the global scope of conservation challenges and halting moves to institutionalize a response to those challenges. The Western powers – particularly the USA, Britain, France, and Germany – continued to export forestry and national park programs to their colonies and protectorates. Early international conventions and treaties addressed the problems of protecting fur seals, sea otters, whales and other marine species, and migratory birds (Matthiessen, 1959). In 1913, an International Conference on the Protection of Nature in Switzerland attracted representatives from 16 nations. World War I prevented this effort from gathering momentum, but in the 1920s and 1930s intermittent steps were taken toward greater international cooperation: a second international conservation conference in Paris in 1923; establishment in 1928 of an International Office for the Protection of Nature in Brussels; an unprecedented international bird conservation conference in Paris in 1933; a Conference for the Protection of African Fauna and Flora in London, also in 1933; the first North American Wildlife Conference in 1936; establishment of the Pan-American Union, dedicated largely to conservation issues in the Western Hemisphere, in 1940 (Nash, 1982). These set the stage for the expanded international programs that emerged in the aftermath of World War II.

Conservation Amid Crisis: The 1930s and 1940s

The social and environmental convulsions of the 1930s and 1940s fundamentally altered perspectives and priorities within the conservation movement. The Depression forced conservationists and nonconservationists alike to consider the connections between human economic systems and the sources of wealth in nature. In practical terms, the New Deal programs in the USA vastly increased governmental support for conservation programs, though often with little consideration of ultimate conservation goals. The experience of the Dust Bowl focused the attention of conservationists and many natural scientists on the consequences of indiscriminate exploitation of sensitive lands, highlighting especially the systemic nature of the soil erosion problem. With the outbreak of World War II, conservation issues fell into the background of concerns, proving the difficulty of maintaining mindfulness of human–nature relationships when social crises erupt. As in no time since the Progressive Era, these years of crisis demonstrated the interwoven nature of social, economic, and environmental problems.

As a movement, conservation remained prominent in North America and elsewhere, with growing popular and political acceptance. But for all of the recent advances, the conservation movement still suffered from its own internal crisis. Beneath the very active surface of conservation, there remained the underlying philosophical rift between the utilitarian and preservationist approaches. Neither approach adequately addressed such extensive problems as soil erosion and exhaustion, disruption of hydrological cycles, endangerment and loss of wildlife species, and the degradation of biotic communities (both terrestrial and aquatic). For a small but growing group of conservationists, the Dust Bowl (among other factors) stimulated a new effort to address the roots of conservation problems, one that would draw on elements of both the utilitarian and preservationist schools. Embracing the preservationist critique of human hubris in manipulating the natural world, the esthetic appreciation of wild nature, the traditional commitment to wise use and stewardship of resources, and the value of science in the service of conservation, this cohort of conservationists sought to define a more robust synthesis to guide conservation action.

The incipient synthesis reflected a basic shift in conservation's scientific foundations. Theory and application – that is, the science and practice of conservation – were meeting in ways they had not before. Developments in ecology and evolutionary biology over the previous decades had begun to inform

conservation issues, even as contemporary problems forced conservationists to reexamine their scientific assumptions. Ecology was revolutionizing scientific understanding of the functioning of biological communities, landscapes, and systems. Evolutionary biology provided new perspectives on, for example, the adaptations, roles, and interactions of all forest species (in contrast to the basic descriptive botany, dendrology, timber physics, and forest mensuration on which silviculture and forestry had previously rested).

For at least some conservationists, these scientific advances suggested a new need: to marry ecology and the various fields of resource management in the effort to sustain, not just the yields of particular commodities, but the healthy functioning of the systems generally. The preeminent voice of this emerging approach was the American forester and wildlife ecologist Aldo Leopold, who in the course of his career had applied ecological principles first to the conservation of forests, then to soils, watersheds, game and wildlife species, then ultimately to the land "as a whole." Writing in 1939, Leopold noted that ecology provided "a new fusion point for all the natural sciences" and that its emergence had placed the economic biologist in a peculiar dilemma: With one hand, he points out the accumulated findings of his search for utility, or lack of utility, in this or that species; with the other, he lifts the veil from a biota so complex, so conditioned by interwoven cooperations and competitions, that no man can say where utility begins or ends. … The only sure conclusion is that the biota as a whole is useful, and [the] biota includes not only plants and animals but also soils and waters as well (Leopold, 1939, p. 727).

Leopold's expanded conservation philosophy, as finally expressed in his landmark essay "The Land Ethic" in *A Sand County Almanac* (1949), placed greater emphasis on the ecological integrity, resilience and beauty of what he called "the biotic community" and rejected the view of nature as merely a collection of disaggregated natural resources. It shifted the role of human beings "from conqueror of the land community to plain member and citizen of it" (Leopold, 1949). This proposed shift, away from an anthropocentric and toward a more biocentric approach to the human role in nature, suggested the need for fundamental changes within the various fields of resource management. Resource managers could no longer regard timber trees, game animals, water, soils, scenic vistas, or any other economically or esthetically significant resource as separate entities. Rather, resources had to be seen as components within diverse systems, connected to and interacting in complex ways with other parts of the system (including human beings). It followed that the conservation professions had to develop more integrated approaches to resource management.

As an attempt to address conservation's underlying schism, Leopold's "evolutionary-ecological land ethic" would confront continuing challenges in the dramatically altered postwar environment (Callicott, 1991). However, it would also provide those in the conservation movement with new conceptual tools with which to meet those challenges.

From Conservation to Environmentalism

The conservation movement emerged from World War II into a changed world. The war altered conservation's social and ecological context. In the world at large, the bonds of empire dissolved in the postwar decades as colonial nations gained their independence. The former colonial powers remained involved in the fate of their erstwhile colonies as donors and providers of development assistance. This in turn altered the shape of the global conservation movement, as aid providers, international agencies, and nongovernmental agencies initiated the long struggle to somehow reconcile international development and conservation priorities.

These changes in international relations occurred even as human population growth, land degradation, air and water pollution, and overexploitation of marine resources first emerged as global-scale conservation issues. The war had demonstrated the interrelated and global nature of modern conservation problems. Fairfield Osborn's *Our Plundered Planet* (1948) and William Vogt's *The Road to Survival* (1948) were among the first books to address in an integrated manner issues of global conservation and development. Reflecting this broader scale of concern, a wide spectrum of national and international governments and organizations met in 1948 and formed the International Union for the Protection of Nature with the ambitious goal of preserving "the entire world biotic environment." Shortly thereafter, the newly established United Nations became involved in international conservation programs.

In the USA, the war years constituted a demographic divide, with important implications for the conservation movement. As veterans and their families resumed their lives at home, the postwar population boom fueled an economic boom that placed increasing demands on natural resources. To meet those demands, war-spawned technologies and industrial processes were quickly adapted for use in the peacetime economy. Over the next several decades, rings of suburbs expanded away from cities into rural lands and, along with dams, highways, and other infrastructural developments, continually reconfigured the landscape.

The postwar years brought a new generation of professionals into agriculture, forestry, wildlife management, and the other conservation fields. Conservation was not immune to the larger social and economic forces in the postwar environment. As the pressure to meet rising demands increased, the resource management professions became more specialized, more focused on commodity outputs, and more inclined to adopt narrow technical solutions to conservation problems. Importantly, the war also changed the ways in which scientific research was supported and conducted, reinforcing the tendency toward specialization and utilitarian applications. The traditional foundations of conservation in systematics, organismal biology, population biology, ethology, and other field-oriented sciences suffered from relative neglect. Collectively, these changes overwhelmed the integrated approach that had begun to emerge prior to the war, requiring conservationists to confront complex environmental problems within an increasingly rigid framework of disciplines and institutions.

However, the postwar years also gave rise to forces that countered these trends. New tools in the earth sciences provided greater scientific understanding of the interrelations within and among terrestrial, aquatic, marine, and atmospheric systems. In the 1950s and 1960s, revolutions in fields ranging from genetics and evolutionary biology to atmospheric

chemistry and geology began to rapidly reshape our understanding of the global biosphere and the human place within it. New communications technologies allowed such information to be shared more efficiently. Meanwhile, the indiscriminate adoption of other technologies – especially in the production of agricultural pesticides, industrial chemicals, and nuclear power – created new concerns over their impacts on human health and ecological systems. With the publication in 1962 of Rachel Carson's landmark book *Silent Spring*, examining the biological impacts of DDT (dichlorodiphenyltrichloroethane) and other pesticides, the modern environmental movement began to assume an identity of its own, distinct from but still connected to the older conservation tradition.

Efforts to protect wildlands in the USA also intensified in the years following World War II. As pressures from urban and suburban expansion, resource extraction, water development, and motorized tourism increased, wilderness protection became a unifying goal among diverse conservation organizations. In 1955, a proposal to build the Echo Park Dam in Dinosaur National Monument in Colorado was defeated, indicating that the conservation movement had gained not only in cohesion but also in popular and political power. That unity led to the passage of the Wilderness Act in 1964, providing for a strengthened national system of wilderness areas on the public lands of the USA.

The Wilderness Act was only one of the many laws adopted during this period of environmental awakening (much as earlier laws had institutionalized the conservation movement). Over the next decade, the US Congress would enact a series of important environmental statutes, including the National Environmental Policy Act (1970), the Clean Air Act (1970), the Clean Water Act (1972), and the Endangered Species Act (1973). These far-reaching changes in environmental policy reflected a rising wave of popular support and organized political activism, symbolized by the observance of the first Earth Day on 22 April 1970.

The Reintegration of Conservation

With the rise of the modern environmental movement, the context of conservation once again changed dramatically. The years following Earth Day saw increasing acceptance, within the conservation professions and within society at large, of environmental values. However, those values were not easily translated into effective conservation action. The resource management professions, in seeking to address broader environmental concerns, found the tendency toward overspecialization difficult to overcome. Conservation programs tended still to focus on single species, or particular economic resources, or separate jurisdictions within a given landscape. As the long-term ecological impacts of such fragmented approaches to resource management became clear, the traditional conservation fields came under increasing public scrutiny. Although many environmentalists pressed for change through legal means, the evolution of management philosophy, policy, and practice would require years of incremental change and adjustment.

At the international level, differences in perspective between the wealthier, developed countries of the North and the poorer, developing countries of the South likewise proved difficult to overcome. International conservation continued to make important gains through a series of conventions and treaties, including the Convention on Wetlands of International Importance (the "Ramsar Convention") (1971), the Convention on International Trade in Endangered Species (1975), the International Convention for the Conservation of Migratory Species of Wild Animals (the "Bonn Convention") (1978), the United Nations Convention on the Law of the Sea (1982), and the United Nations Convention on Biological Diversity (1992). Again, however, international development policies would incorporate stronger conservation and environmental provisions only gradually, as the need to connect economic development and long-term environmental security became increasingly evident.

As these broad patterns of change in the human dimensions of conservation unfolded in the 1970s and 1980s, the scientific foundations of conservation were again shifting. Taxonomy and systematics provided more robust estimates of the extent of species diversity and of its actual and potential loss. The subfield of island biogeography revealed principles governing the spatial distribution, persistence, and extinction of species, with important implications for land use and the establishment and management of protected areas. Genetics became an increasingly important component of conservation science as attention focused on the reproductive success of rare and endangered species and the viability of their populations, both in captivity and in the wild. Ecology moved away from its "classical paradigm," which emphasized single, stable, deterministic equilibria, and toward a view of ecosystems that emphasized flux, uncertainty, and contingency. Disturbances such as fire, flooding, and drought were incorporated into our understanding of long-term ecosystem dynamics and were shown to be important factors affecting species diversity. Increasingly, conservation strategies required the integration of knowledge from the many branches of biological science, involving various levels of biological organization.

The need to rethink conservation across disciplinary lines was driven not only by changes in the foundational sciences but also by changes in the environment itself. By the late 1970s, scientists and conservationists were alarmed by the escalating loss of genetic, species, and ecosystem diversity at the global scale. Of special concern was the accelerated destruction of the species-rich forests of the humid tropics. The advances in island biogeography revealed that ecosystems of all types were being fragmented by human activity, and that even the most effective protected areas were at risk due to their inadequate size and their isolation. Wildland managers increasingly understood that preservation alone was an inadequate management strategy, and that the loss of diversity and the disruption of ecological functions were intimately associated. Agricultural scientists, foresters, and other resource managers, too, were increasingly concerned about environmental degradation, the breakdown of ecosystem processes, the loss of diversity in humanized landscapes, and the human costs associated with these changes.

These concerns prompted the emergence in the mid-1980s of a new synthetic discipline, conservation biology, specifically devoted to the integration of scientific disciplines and other fields in the effort to understand, maintain, and restore

biological diversity (Meine *et al.*, 2006). Conservation biology has sought to address conservation problems within an evolutionary and ecological context, and to stimulate the traditional conservation professions to reassess their management methods and goals accordingly. The term *biodiversity* itself was coined in 1987 and has since been widely adopted by conservationists. The concepts of *sustainability* and *sustainable development* came into general usage during the same period, reflecting the complex challenge of integrating long-term social, economic, and environmental factors in assessing human demands and impacts on ecosystems. The integrating impulse was also evident in the recent emergence of restoration ecology, landscape ecology, sustainable agriculture, environmental history, environmental ethics, ecological economics, and other interdisciplinary fields.

Meanwhile, these changes in conservation science, philosophy, and policy prompted many in the traditional resource management fields to adopt more integrated approaches. In practical terms, this meant that as the conservation professions approached the new century, they sought to move away from narrow economic criteria for success and toward broader ecological standards; away from the mathematical desideratum of sustained yield of commodities and toward the more complex challenge of sustaining healthy ecosystems; away from the management of discrete components of the system and toward integrated ecosystem management; and away from an exclusive focus on the human goods and services provided by ecosystems and toward a focus on the biological diversity and ecological processes that sustain those goods and services (Minteer and Manning, 2003). These changes also began to redefine the meaning and value of wilderness, and to reintegrate wildlands within a broader approach to whole-landscape conservation. Whereas wilderness as traditionally defined existed apart from the humanized landscape (even while being preserved to meet human recreational and esthetic needs), wildlands at all spatial scales have become increasingly valued as repositories of biological diversity, as core protected areas and corridors within greater ecosystems, and as "controls" against which to compare the human impact on more intensively utilized lands.

These examples of increasing integration in conservation hearken back in many ways to the ecology-based synthesis that Aldo Leopold and others sought to articulate prior to the onset of the Cold War. Since then, the science of ecology has evolved, analytical tools and technologies have grown vastly more sophisticated, environmental threats have intensified, and the social and economic context of conservation has changed continually. Nevertheless, there is continuity between contemporary efforts to expand conservation's cultural and natural connections and Leopold's efforts to "enlarge the boundaries of the community to include soils, waters, plants, and animals" (Leopold, 1949).

Emerging Themes in Conservation

As in the past, conservation remains responsive to new demands, new information, and new realities. The sweeping changes that have reshaped the conservation movement since the mid-1980s will continue to alter both its social and political dimensions and its practical field methods. Although we are still too much embedded in the present to recognize all these recent developments, we can identify important contrasts between emerging themes and conservation's past practices.

In terms of its scientific foundations, conservation now seeks greater collaboration among the natural and social sciences, even while reaching beyond the sciences to build connections with philosophy, economics, history, literature, and the arts. Ecology continues to serve as a "fusion point" for the conservation-related sciences. In its conservation applications, ecology is bringing attention to a broader spectrum of species and to the processes that maintain diversity. To improve its chances of success, conservation continues to integrate the insights of ecology into the resource management professions, and to do so at various spatial and temporal scales. Conservationists increasingly focus on the resilience of coupled human and natural systems, and the recognition of ecosystem services as a means of encouraging and rewarding practices that enhance such resilience. The integrating fields of conservation biology, landscape ecology, restoration ecology, sustainable agriculture, environmental ethics, and ecological economics are now more solidly established (even while resisting the tendency toward specialization and reductionism to which even interdisciplinary fields are prone).

Emerging threats are making the conservationist's task even more precarious. These include the rapid spread of invasive exotic species (especially as the human economy continues to globalize), the gathering momentum of global climate change, the increased presence of genetically modified organisms in the landscape, and the seemingly relentless fragmentation of habitats and ecosystems. At the same time, the sciences are being called on to illuminate the connections among biodiversity conservation, human health, and the well-being of human communities. Accordingly, the active weaving of conservation science into community-based ecological restoration programs promises to redefine the role and methodology of science in conservation.

The conservation movement continues to be shaped by overarching social and demographic trends. The growth in human population and in resource consumption rates remains a profound factor determining conservation's future. To better inform that future, conservationists are seeking novel ways to integrate sustainable economies with effective protection, management, and restoration programs. Conservation planning efforts are expanding their reach across jurisdictional boundaries and across the landscape, recognizing the connections among wildlands, semiwild lands, "working" landscapes, and suburban and urban environments. Such cross-boundary forays have few precedents in conservation history and represent an important departure from the past.

In the past, conservation has also neglected human and cultural relationships to its own detriment. Increasingly conservationists recognize the need to integrate natural science, cultural traditions, and social relationships in the effort to protect and restore particular places. In seeking a better fit between the built environment and surrounding landscapes, conservationists have also begun to work more closely with architects, landscape architects, planners, and engineers to incorporate the emerging principles of conservation biology.

Meanwhile, the conservation movement itself continues to increase in diversity as individuals from varied backgrounds come into conservation as both professionals and amateurs. The environmental justice movement has, since the early 1990s, sought to broaden the political purview and demographics of the conservation movement. Local nongovernmental organizations are now regarded as key conservation players, and their role in conservation continues to grow. Although action at the international level remains necessary in addressing global threats to biodiversity and in supporting local initiatives, the assumption of greater local responsibility for ecosystem health is increasingly urgent. Community-based conservation projects, watershed restoration programs, land trusts, and other mechanisms are emerging to meet that need. As these trends continue, they seek to create not just an enhanced conservation movement, but an enduring culture of conservation. In doing so, they extend the long history of increasing citizen involvement in conservation.

Appendix

List of Courses

1. Conservation Biology
2. Introductory Environmental Science
3. Survey of Environmental History

See also: Ecosystem, Concept of. Environmental Ethics. Historical Awareness of Biodiversity. Indigenous Peoples and Biodiversity. Literary Perspectives on Biodiversity. Religious Traditions and Biodiversity. Stewardship, Concept of

References

Barrow MV (2009) *Nature's Ghosts: Confronting Extinction from the Age of Jefferson to the Age of Ecology*. Chicago: University of Chicago Press.

Callicott JB (1991) Whither conservation ethics? *Conservation Biology* 4: 15–20.

Cronon W (2003) *Changes in the Land: Indians, Colonists, and the Ecology of New England*, Twentieth Anniversary Edition. New York: Hill and Wang.

Dubos R (1980) *The Wooing of Earth*. New York: Charles Scribner's Sons.

Dunlap T (1988) *Saving America's Wildlife*. Princeton, NJ: Princeton University Press.

Egan T (2006) *The Worst Hard Time: The Untold Story of Those Who Survived the Great American Dust Bowl*. New York: Houghton Mifflin Harcourt.

Ellis EC (2011) Anthropogenic transformation of the terrestrial biosphere. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 369: 1010–1035.

Fox S (1981) *John Muir and his Legacy: The American Conservation Movement*. Boston: Little, Brown.

Harrison RP (1992) *Forests: The Shadow of Civilization*. Chicago: University of Chicago Press.

Hays SP (1959) *Conservation and the Gospel of Efficiency: The Progressive Conservation Movement, 1890–1920*. Cambridge, MA: Harvard University Press.

Jackson W (1985) *New Roots for Agriculture*. Lincoln and London: University of Nebraska Press.

Leopold A (1933) *Game Management*. New York: Charles Scribner's Sons.

Leopold A (1939) A biotic view of land. *Journal of Forestry* 37: 727–730.

Leopold A (1949) *A Sand County Almanac and Sketches Here and There*. New York: Oxford University Press.

MacPhee RDE (ed.) (1999) *Extinctions in Near Time: Contexts, Causes, and Consequences*. New York: Plenum Press.

Marsh GP (1864, 1965) *Man and Nature: Or Physical Geography as Modified by Human Action*. Cambridge: Cambridge University Press.

Matthiessen P (1959) *Wildlife in America*. New York: The Viking Press.

Mayr E (1982) *The Growth of Biological Thought: Diversity, Evolution and Inheritance*. Cambridge, MA: Harvard University Press.

Meine C, Soulé M, and Noss R (2006) "A mission-driven discipline": The growth of conservation biology. *Conservation Biology* 20: 631–651.

Miller C (2001) *Gifford Pinchot and the Making of Modern Environmentalism*. Washington, DC: Island Press.

Minteer B and Manning R (eds.) (2003) *Reconstructing Conservation: Finding Common Ground*. Washington, DC: Island Press.

Nash R (1982) *Wilderness and the American Mind*, 3rd edn. New Haven and London: Yale University Press.

Pinchot G (1910, 1967) *The Fight for Conservation*. Seattle and London: University of Washington Press.

Pinchot G (1947) *Breaking New Ground*. New York: Harcourt, Brace and Co.

Taylor B (2005) *The Encyclopedia of Religion and Nature*. London and New York: Continuum.

Williams M (1989) *Americans and their Forests: A Historical Geography*. Cambridge: Cambridge University Press.

Worster D (2001) *A River Running West: The Life of John Wesley Powell*. New York: Oxford University Press.

Conserving Biodiversity Outside Protected Areas

Belinda Reyers, Council for Scientific and Industrial Research, Stellenbosch, South Africa

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alien species A species introduced by humans (deliberately or accidentally) outside its native distributional range.

Bioregional management An approach to management which encompasses whole ecosystems so as to protect and restore their components sustainably.

Biosphere reserves Sites established by countries and recognized under UNESCO's Man and the Biosphere (MAB) Programme to promote sustainable development by combining core protected areas with zones where sustainable development is fostered.

Conservancy A collection of mostly agricultural properties which are cooperatively managed under a single management plan, bound with a voluntary agreement between the landowners and the state.

Conservation planning The systematic and spatially explicit identification of areas important to achieving particular biodiversity conservation targets.

Conservation targets Quantitative estimates of the amount of each species or ecosystems that should be included in a conservation plan.

Driver of change A natural or human-induced factor that directly or indirectly causes a change in an ecosystem.

Easement An agreement which transfers certain specified rights (e.g., development, timber harvesting, mining, etc.) from the landowner to a third party. These transfers can be attached to the property title deed and are either acquired by donation or purchase.

Ecologically sustainable management Management which focuses on the wise use, conservation and enhancement of natural resources so that ecological processes, on which life depends, are maintained, and the total quality of life, now and in the future, can be improved.

Ecosystem services Benefits that people receive from ecosystems including provisioning, regulating, supporting and cultural benefits.

Gap analysis A tool used in conservation to identify gaps (e.g., ecosystems or species not represented in the current network of protected areas).

Multifunctional landscapes Landscapes which provide multiple environmental, social and economic functions and are able to achieve multiple societal needs including energy and food production, management of waste, conservation of biodiversity, and management of water quantity and quality across the landscape; the improvement of landscape heterogeneity and therefore resilience; and the provision of recreational opportunities.

Off-reserve conservation Conservation efforts which focus on the management of whole landscapes, complementing formal protected areas with other (potentially cheaper) property rights mechanisms including the creation of management agreements and the designation of conservation easements.

Payment for ecosystem services Contractual or negotiated agreements where ecosystem service beneficiaries compensate ecosystem stewards for enhancing, maintaining, reallocating or offsetting damage to ecosystem services.

Property rights instruments Mechanisms which support conservation on private land and vary from voluntary to contracted agreements between the conservation agency and the landowner(s). They can include management agreements and conservation easements.

Protected area A clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values.

Protected Areas: History, Purpose and Gaps

Protected areas have long been the mainstay of humanity's attempts to conserve areas deemed to be of natural, cultural, or ecological importance. As long as 2000 years ago ancient societies in Greece, Rome, Asia, and Africa are known to have set aside areas as sacred groves or sites (Chandrashekara and Sankar, 1998), while European societies protected hunting grounds for the use of royalty and the wealthy as long as 1000 years ago. Some of these areas are still in existence today. But it was only in the nineteenth century that the modern protected areas movement took hold globally with the establishment of large protected areas like Yellowstone National Park (1872) in the USA, Royal National Park in Australia (1879), and the Kruger National Park in South Africa (1898) (**Box 1**).

The motivations behind these protected areas historically were multiple, for example, the protection of game species in colonial Africa or scenic wilderness landscapes in North America (Chape *et al.*, 2005). Until recently, in fact, the motivations have seldom been the protection of biodiversity *per se*, and have usually been based on culturally valued aspects of biodiversity and the broader landscape, for example, charismatic megafauna, attractive habitats, important watersheds, recreational areas, or endangered species. It is therefore important to note that the current set of protected areas serve many functions besides biodiversity conservation; a fact recognized by the IUCN (1994), which lists multiple functions of protected areas including: scientific research, wilderness protection, preservation of species and genetic diversity, maintenance of environmental services, protection of specific

Box 1 What is a protected area?

The IUCN, the International Union for Conservation of Nature, defines a protected area as: "a clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values." It has created a system (the IUCN Protected Areas Categories System) that includes six distinct categories of protected areas:

Ia. Strict Nature Reserves are strictly protected areas set aside to protect biodiversity and also possibly geological/geomorphical features, where human visitation, use, and impacts are strictly controlled and limited to ensure protection of the conservation values.

Ib. Wilderness Areas are usually largely unmodified or slightly modified areas, retaining their natural character and influence without permanent or significant human habitation, which are protected and managed so as to preserve their natural condition.

II. National Parks are large natural or near natural areas set aside to protect large-scale ecological processes, along with the complement of species and ecosystems characteristic of the area, which also provide a foundation for environmentally and culturally compatible, spiritual, scientific, educational, recreational, and visitor opportunities.

III. Natural Monuments or Features are set aside to protect a specific natural monument, which can be a landform, sea mount, submarine cavern, geological feature such as a cave or even a living feature such as an ancient grove.

IV. Habitat/Species Management Areas aim to protect particular species or habitats and their management reflects this priority.

V. Protected Landscape/Seascape is an area where the interaction of people and nature over time has produced an area of distinct character with significant, ecological, biological, cultural, and scenic value: and where safeguarding the integrity of this interaction is vital to protecting and sustaining the area and its associated nature conservation and other values.

VI. Protected area with sustainable use of natural resources is an area that conserves ecosystems and habitats together with associated cultural values and traditional natural resource management systems. They are generally large, with most of the area in a natural condition, where a proportion is under sustainable natural resource management and where low-level nonindustrial use of natural resources compatible with nature conservation is seen as one of the main aims of the area

In this article the term protected area refers to Categories I and II unless otherwise stated.

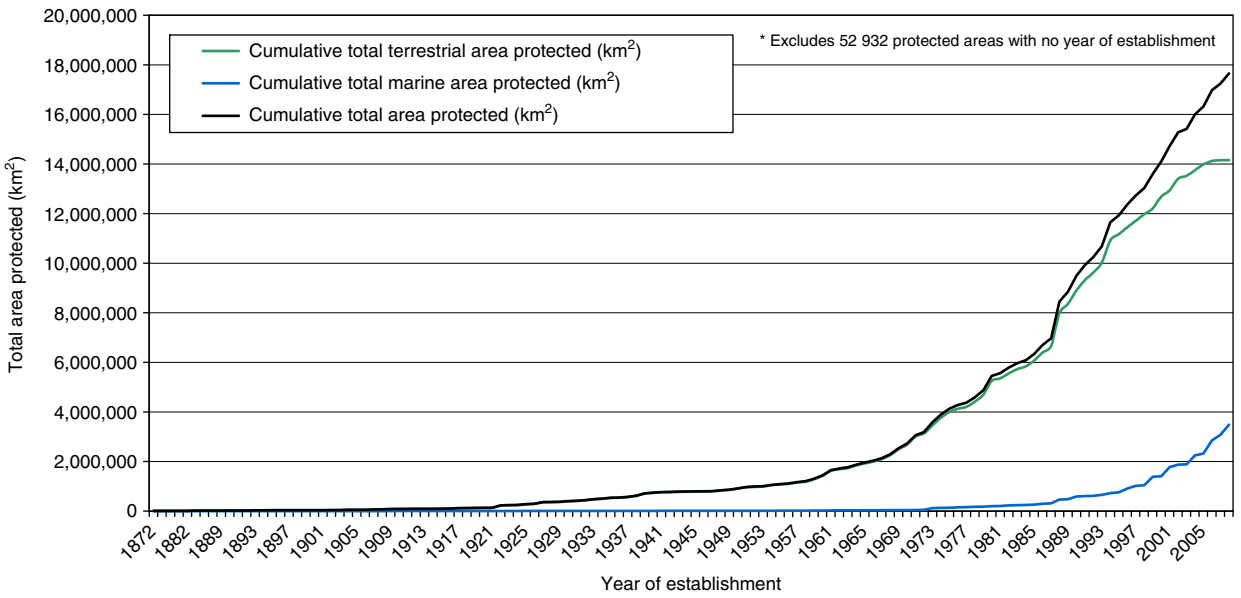


Figure 1 Global growth in protected areas. Reproduced from IUCN and UNEP-WCMC (2009) *The World Database on Protected Areas (WDPA): January 2009*. Cambridge, UK: UNEP-WCMC.

natural and cultural features, tourism and recreation, education, sustainable use of resources from natural ecosystems, and maintenance of cultural and traditional attributes. When viewed from this perspective, the growth in protected areas and their role in society has been substantial. From a mere 10,000 protected areas in 1962, today the World Database on Protected Areas totals 114,000 sites comprising 12.9% of the earth's land surface contributing in many varied direct and indirect ways to economies, societies, and individuals (Dixon and Sherman, 1991; Chape *et al.*, 2008; **Figures 1** and **2**).

However, when viewed through the lens of biodiversity conservation, the bottom up and rather *ad hoc* approach to protected area establishment, further exacerbated by the fact that protected areas were often only established in remote and unproductive areas that no-one else wanted (Pressey *et al.*, 1996; Norton, 2000; Scott *et al.*, 2001), has resulted in an inefficient and ineffective global network of protected areas (Brooks *et al.*, 2004; Rodrigues *et al.*, 2004; Hoekstra *et al.*, 2005).

The first bias in this global coverage is seen when distinguishing between terrestrial, inland water, and marine



Figure 2 Protected areas of the world. Reproduced from World Database on Protected Areas (WDPA), UNEP-WCMC, July 2011.

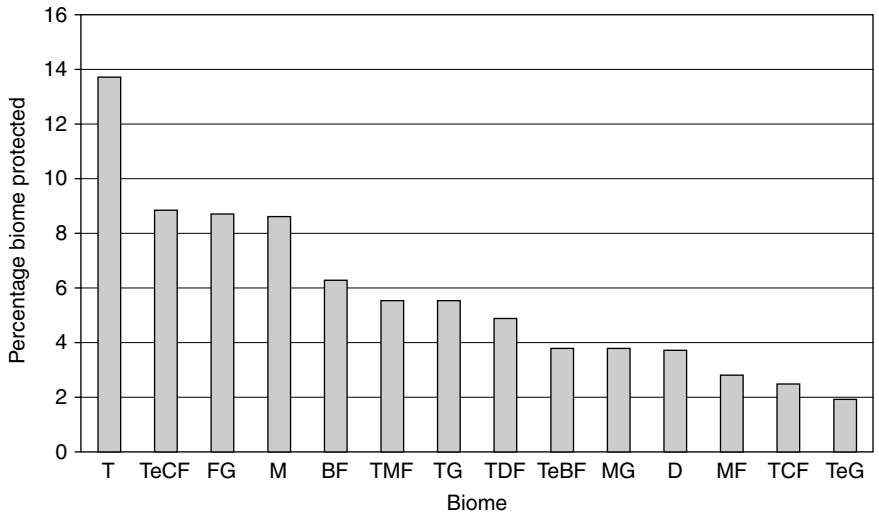


Figure 3 Protected area coverage of the world's biomes. TMF, Tropical and subtropical moist broadleaf forests; TDF, Tropical and subtropical dry broadleaf forests; TCF, Tropical and subtropical coniferous forests; TeBF, Temperate broadleaf and mixed forests; TeCF, Temperate coniferous forests; BF, Boreal forests/taiga; TG, Tropical and subtropical grasslands, savannas, and shrublands; TeG, Temperate grasslands, savannas, and shrublands; FG, Flooded grasslands and savannas; MG, Montane grasslands and shrublands; T, Tundra; MF, Mediterranean forests, woodlands, and scrub or Sclerophyll forests; D, Deserts and xeric shrublands; M, Mangrove. Reproduced from Mace GM, Masundir H, Baillie J, *et al.* (2005) *Biodiversity. Millennium Ecosystem Assessment: Current State and Trends*. Washington, DC: Island Press. ISBN 1-55963-228-3.

systems. While 10–12% of the land surface area and inland water areas fall within protected areas, marine systems have less than 1% of their area protected (Chape *et al.*, 2008). Within terrestrial systems the biases are even more apparent with poorly protected biomes like temperate grasslands, Mediterranean forests, and tropical coniferous forests (around 2% of their area is conserved) contrasting with (the less economically productive) flooded grasslands, tundra, temperate coniferous forests, mangroves, and boreal forests, which have the highest percentage area under protection (7–14%; **Figure 3**; Mace *et al.*, 2005). Global data at the ecosystem level for freshwater and marine systems are currently unavailable, but at a regional scale Chape *et al.* (2008) showed that only Australasia and North Eurasia have more than 2% in marine protected areas, while data seem to indicate that Australasia, the Caribbean, East Asia, and Eastern and Southern Africa have around 15% of their inland water systems protected. Protected area coverage of inland waters is low in North America and North Eurasia – areas with large areas of inland waters. In a national study of freshwater protected areas Nel *et al.* (2007) found that more than 90% of all main rivers in South Africa fall completely outside statutory protected areas, while half of the remaining rivers form boundaries of protected areas and therefore less than 5% of main rivers in the country are core rivers within protected areas, receiving protection on both sides of their river bank.

At the species level, using data on terrestrial vertebrates, Rodrigues *et al.* (2004) found that 12% of mammals, birds, turtle, and amphibians do not fall into the global network of protected areas. This number doubled if small protected areas were excluded. The number was highest for amphibians, for threatened species, and for areas with many endemic species. The authors point out that their numbers are certainly an underestimate of terrestrial species falling between the cracks of protected areas. In the freshwater and marine realms major gaps in species distribution data (e.g., Lévêque and Balian, 2005) imply that little is known about the contribution of protected areas to biodiversity conservation. But considering the low percentage of area protected in the case of the world's oceans, marine species are probably worse off than their terrestrial counterparts. This is further supported by growing evidence that freshwater systems are more threatened than terrestrial ones (Dudgeon *et al.*, 2006) and the fact that many rivers actually form the boundaries of protected areas and cannot necessarily be considered conserved by the protected area (Nel *et al.*, 2011). The gaps in freshwater species conservation are potentially large.

Results from these gap assessments in global protected areas, as well as improvements in the data availability and the science of conservation planning, have all added to calls for a more systematic approach to the establishment of protected areas (**Box 2**).

Beyond Protected Areas

Despite the proliferation of new data, methods and results from conservation planning helping to plug some of the gaps in the global protected area network, it is unlikely that protected areas alone will stem the tide of biodiversity loss. In fact recent evidence suggests that despite increases in the extent of protected areas, biodiversity losses continue (Butchart *et al.*, 2010). While some of this has to do with the limitations of protected areas, for example, limited size, area and resources for expansion, and a lack of spatial congruence with biodiversity conservation needs (Lindenmayer and Burgman, 2005), at the center of this conundrum lies conservation's inability to secure societal commitment and adequate responses to address the drivers of change in biodiversity both inside and outside of protected areas.

Box 2 The rise of systematic conservation planning

In their review of systematic conservation planning, Margules and Pressey (2000) argue that protected areas should ideally fulfill two roles with respect to biodiversity conservation: (1) they should *represent* or sample the full variety of biodiversity at all levels of organization (genes, species, and ecosystems); and (2) they should ensure the persistence or long-term survival of this biodiversity that they represent through maintaining the ecological and evolutionary processes necessary to this persistence.

As Margules and Pressey (2000) reflect, it is clear that a "structured systematic approach to conservation planning provides the foundation needed to meet (the) objectives" of representation and persistence of biodiversity within the constraints of limited resources and competition for land. They propose a six-step framework for systematic conservation planning, which is based on spatial biodiversity data, as well as quantitative targets – for example, 10% of each habitat, 200 individuals of a species – which explicitly set out the goals of the conservation plan.

Conservation planning deals with the identification of geographic priority areas that contain species, habitats or processes essential to achieving the conservation goals of a region or country (Margules and Pressey, 2000). It uses techniques concerned with the maximal coverage problem including complementarity-based reserve selection and linear programing algorithms. Systematic conservation planning is said to be effective at achieving the twin objectives of representation and persistence due to its cost efficiency, its defensibility and flexibility in the face of competing land and resource requirements, and its accountability in allowing the outputs of plans to be critically reviewed (Margules and Pressey, 2000). These past two decades have witnessed a proliferation of tools, techniques, software, databases, and research outputs devoted to the task of systematically identifying spatial priority areas for conservation action – see Moilanen *et al.* (2009) for a review.

These drivers, for example, habitat loss, overexploitation, alien species introductions, and climate change have resulted in significant losses to biodiversity, especially over the past 50 years (MA, 2005) and are set to increase into the future (Butchart *et al.*, 2010; CBD, 2010). While their impact on biodiversity outside of protected areas is evident (e.g., MA, 2005; Butchart *et al.*, 2010), more subtle changes have been wrought by these drivers on the biodiversity inside protected areas. Chape *et al.* (2008) highlight the large range of physical (e.g., fire), biological (e.g., alien species), social (e.g., community opposition), political (e.g., political support), economic (lack of resources), and managerial (e.g., lack of planning) threats faced by biodiversity within protected areas. Climate change poses a new and largely uncertain threat to these areas with some projections highlighting major biodiversity losses (e.g., Erasmus *et al.*, 2002). Pressey and Logan (1997) add to these constraints the limited effectiveness that protected areas can have on some drivers of change inside and outside of protected areas, for example, salinization and airborne chemical pollution.

If the conservation community is to be successful in addressing the drivers of change in biodiversity then two things are clear: (a) a reliance on protected areas will not be enough; and (b) conservation will have to start engaging with the areas and sectors from which the drivers of biodiversity loss originate. Both of these realizations signal a move for the conservation community beyond protected areas – a move that begun over two decades ago (e.g., Hale and Lamb, 1997). There has since been significant effort invested in the development of conservation interventions focused outside protected areas. This article reviews the current set of conservation approaches using a framework to categorize approaches and then explores each category to highlight developments, challenges, and opportunities.

Conserving Biodiversity Outside of Protected Areas

A Typology of Conservation Approaches

Biodiversity conservation is a broad and diverse initiative ranging from local site scale interventions to global political interventions. In the past few decades approaches to conservation have increased in number, scope, and complexity in terms of what the approaches aim to conserve, as well as how, when, and where they go about conserving biodiversity (Redford *et al.*, 2003). The focus of conservation efforts has evolved over time and today ranges from specific species, through landscapes, to people, and their well-being. As this focus has evolved so too have the approaches with which to conserve these elements. Redford *et al.* (2003) reviewed 13 conservation organizations and 21 approaches used to identify where or how conservation should be done. Their results illustrate the enormous diversity of approaches, scales, and principles underpinning the approaches.

This article, using a framework based on several existing classifications of conservation approaches, presents and explores the current diversity of approaches to conservation beyond protected areas. **Table 1** shows the nine approaches used by McNeely *et al.* (2005) who focused on the most widely used and discussed responses to biodiversity loss. Their list includes protected areas; wild species management; regional planning, agricultural/forestry and fishery policy; governance; multilateral environmental agreements; education/communication; local capture of benefits; and private sector activities. Mawdsley *et al.* (2009) identified four approaches in exploring conservation options for climate change. They classified biodiversity conservation responses into (a) land and water protection and management; (b) direct species management; (c) monitoring and planning; and (d) law and policy. Based on these two categorizations it appears that there are six broad approaches to conservation (**Table 1**). None of these approaches is mutually exclusive and in fact conservation usually relies on a combination of approaches to be successful (Young and Gunningham, 1997).

Examining the approaches with respect to their type of action, it is possible to divide them into two main types of actions: direct or enabling. Direct actions are those that involve on-site and on-the-ground management of areas and species and would include approaches 1 and 2 in **Table 1**. Enabling actions are those that happen mostly off-site and make it possible for the direct action to take place and include approaches 3, 4, 5, and 6. While acknowledging the importance of enabling actions to conserving biodiversity outside of protected areas, this article does not explore them in detail, rather focusing on approaches to land and water protection

Table 1 Classification of conservation approaches based on McNeely *et al.* (2005) and Mawdsley *et al.* (2009)

Class	*Approaches to conservation*	*McNeely* et al. *(2005)*	*Mawdsley* et al. *(2009)*	*Focus of approach*
I	Land and water protection and management	Protected areas	Land and water protection and management	Direct action
II	Species management	Wild species management	Direct species management	Direct action
III	Monitoring and planning	Regional planning	Monitoring and planning	Enabling action
IV	Law, governance and policy	Agricultural/forestry & fishery policy Governance Multilateral environmental agreements	Law and policy	Enabling action
V	Incentives and partnership	Local capture of benefits Private sector activities		Enabling action
VI	Outreach and education	Education/Communication		Enabling action

Source: Reproduced from McNeely J, Albers H, Faith DP, *et al.* (2005) *Biodiversity. Policy Responses of the Millennium Ecosystem Assessment, Ecosystems and Human Well-being, Vol. 3 Millennium Ecosystem Assessment.* Washington, DC: Island Press. ISBN 1–55963-270-4, and Mawdsley JR, O'Malley R, and Ojima DS (2009) A review of climate-change adaptation strategies for wildlife management and biodiversity conservation. *Conservation Biology* 23: 1080–1089.

and management and some approaches to *in situ* species management, referring only to the other approaches where relevant in supporting these first two approaches.

Land and Water Protection and Management

There are many conservation approaches currently in use in the protection and management of land and water for biodiversity conservation (**Box 1**). In differentiating these approaches it is useful to distinguish their primary management objective. For example, the range of protected areas recognized by the IUCN considers that categories I–III are mainly concerned with the conservation of natural areas where human intervention and modification of the environment are limited. Categories V and VI generally involve greater intervention and modification. Category IV sits somewhere in the middle with its primary goal the interventionist management of certain natural conditions or processes. However, there are many areas, which fall outside of these categories, that can play a valuable role in conserving biodiversity. These include small areas, areas without formal management arrangements or proper security of protection, or areas that are too transformed to be listed (Pressey and Logan, 1997). The value of these areas was recognized in early work by proponents of bioregional management – an approach to management, which sought to encompass the whole ecosystems so as to protect and restore their components sustainably (WRI, 1996; UNESCO, 1995; Wells and Brandon, 1993). Work in this field included the UNESCO Man and the Biosphere Reserve Programs, as well as Integrated Conservation and Development Projects. Since then there has been much scientific and practical progress in the area of whole ecosystem management, which this article summarizes and explores.

From Strict Protection to Living Landscapes: A Framework for Review

While the world has historically relied on state-owned reserves and land-use regulations to conserve biodiversity; the right combination of beyond protected area conservation mechanisms and incentives has shown tremendous potential in conserving biodiversity outside of protected areas. The framework depicted shows the evolution of conservation approaches to managing biodiversity on land and in the water (**Figure 4**). Starting at the bottom right hand corner the framework moves from the strict traditional approach to protected areas managed almost exclusively for biodiversity conservation, through the less stringent IUCN Category V and VI areas where some level of human activity and resource

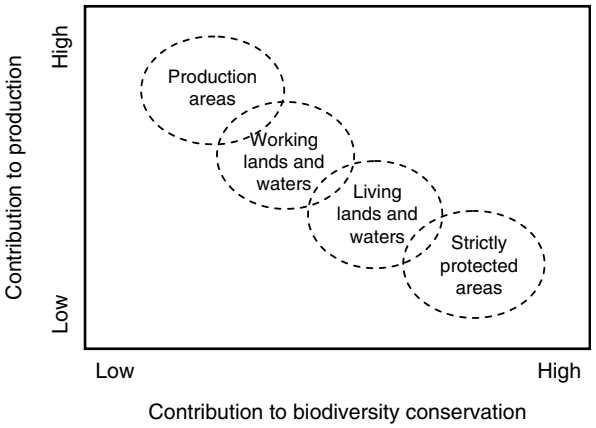


Figure 4 A framework for reviewing the contribution of areas of land and water to biodiversity conservation. Starting at the bottom right hand corner the framework moves from "strictly protected areas," reflecting the more traditional approach to protected areas managed almost exclusively for biodiversity conservation, The next category is "living lands and waters", which are areas managed primarily for biodiversity conservation with some extractive uses limited to the ecologically sustainable management of areas of land and water to support life of all forms. "Working lands and waters" are mostly agricultural lands managed primarily for extractive uses while attempting to conserve biodiversity at the same time. The final category is "production areas" of land and water where the management focuses exclusively on maximizing extractive and productive uses and biodiversity conservation is not an objective.

extraction is permissible, through areas managed for both conservation and production purposes, to a set of final areas of land and water where the management focuses exclusively on maximizing extractive and productive uses. The evolution also signals a shift from mostly state-owned land to mostly private and communal land. As illustrated in this framework there are overlaps in this transition from strict protected areas to intensive extraction and it is likely that in all areas some contribution to biodiversity conservation, as well as resource use and benefits is possible. While the opposite extremes of this framework are relatively well understood and classified in terms of their contribution to biodiversity conservation and production uses, the transition zone is less clear and as yet no single classification exists for the different types of conservation approaches within this zone. There are many terms currently in use to describe the conservation approaches including off-reserve, biosphere, conservancy, easement, bioregion, etc. The framework proposes two main categories within this transition zone differentiated based on primary management objectives as either: biodiversity conservation with some extractive uses (living lands and waters) or extractive uses, which try to conserve biodiversity at the same time (working lands and waters). In the former category one would expect a more intact or natural landscape with low levels of extractive use, for example, grazing, wildlife farming, or fuelwood harvesting. In the other category the landscape would be converted into production lands, but with an attempt to limit losses to biodiversity, for example, biodiversity-friendly agriculture and sustainable forestry.

Living Lands and Waters

The idea behind this broad approach to conservation is the ecologically sustainable management of areas of land and water to support life of all forms (e.g., Driver *et al.*, 2003). Ecologically sustainable management focuses on the wise use, conservation and enhancement of natural resources so that ecological processes, on which life depends, are maintained, and the total quality of life, now and in the future, can be improved (Knight and Cowling, 2003). This concept has also found traction in the management of river systems. Richter *et al.* (2003) define ecological sustainable management in the context of water management as "management that protects the ecological integrity of affected ecosystems while meeting the intergenerational human needs for water and sustaining the full array of other products and services provided by natural freshwater ecosystems. Ecological integrity is protected when the compositional and structural diversity and natural functioning of affected ecosystems are maintained." While there is as yet no commonly accepted definition of "sustainable management" or "ecological integrity" it is clear that such management approaches focus on maintaining functional ecosystems.

These areas of land and water can be under many forms of tenure, but are largely privately or communally owned and used to derive income or livelihoods for the landowner(s). Many vast tracts of land can be said to exist in this state even in the absence of direct conservation interventions, for example, the large rangelands areas in Africa, Australia and North America or remote mountain catchment areas, forests and their rivers (e.g., Wallis de Vries *et al.*, 1998; Schwartzman *et al.*, 2000; Kessler and Thomas, 2006). However, the contribution of such areas to biodiversity conservation, while currently high, is uncertain and almost certainly declining if one considers the increases in drivers of biodiversity loss. Overgrazing, changes in fire management practices, climate change, and habitat loss all threaten to reduce the contribution made by these areas and have motivated many conservation organizations to work on ensuring the sustained contribution of the areas to conservation.

In an effort to secure the contribution of areas of land and water (especially those with private or communal tenure) to biodiversity conservation, conservation organizations have developed and used a range of conservation instruments developed for managing biodiversity outside of protected areas. Young *et al.* (1996) introduce the wide array of instruments available, which Pence *et al.* (2003) group into five categories: motivational, educational, and information instruments; voluntary instruments; property-rights instruments; financial mechanisms; and regulation.

Motivational, educational, and information instruments include a wide variety of approaches, which aim to encourage attitudinal change and the adoption of an ethic of conserving nature on private land and can also be combined with the other instruments to help support conservation on private land. Voluntary instruments allow for conservation areas to be set aside on private land and in return some technical assistance may be offered. They are often favored over binding contractual arrangements or compensatory measures as a mechanism for conserving biodiversity on private property. At the other end of the spectrum are the regulatory instruments, which in some countries regulate what private landowners may and may not do on their land (e.g., clearing native vegetation). The other categories of property rights and financial mechanisms are currently receiving a lot of attention in biodiversity conservation efforts – particularly those outside of protected areas.

Property rights mechanisms aim to support the management of these areas and their biodiversity and vary from relatively voluntary to contracted agreements between the conservation agency and the landowner(s). The most popular forms of agreement include management agreements and conservation easements. A management agreement is usually a short term (<30 year) legally binding contract setting out the use and management of the land and is an agreement between a landowner and a conservation agency (Crompton, 1990). Such agreements are currently being used in schemes like the Common Agricultural Policy of the European Union and conservation in the UK (Pence *et al.*, 2003). Management agreements are flexible and often coupled with financial mechanisms. A conservation easement transfers certain specified rights (e.g., development, timber harvesting, mining, etc.) from the landowner to a third party. These transfers can be attached to the property title deed and are either acquired by donation or purchase (Boyd *et al.*, 1999). In the USA conservation easements are used by private, nonprofit land trusts and by government agencies.

Financial mechanisms include a large range of instruments, which conservation organizations are putting into use in



Figure 5 This canopy walkway built in 1995 in Kakum National Park in Ghana has become a major attraction for tourists. Tourist numbers have grown from <1000 visitors per year in 1991 to almost 90,000 visitors in 2002, >80% of which are domestic. This has helped to create an effectively managed and sustainable national park (http://www.conservation.org/learn/culture/ecotourism/destinations/Pages/kakum_canopy_walkway.aspx) (Photo B Reyers).

motivating private and communal landowners to manage and conserve biodiversity on their land, and in helping landowners defer some of the costs associated with managing their land in this way. These instruments include grants, tax policies, commercialization, business instruments, and removal of perverse subsidies. Changes to tax policy have been quite pervasive and include changes to income tax, capital gains tax, rate relief as well as exemptions or deductions via property taxes, federal and state income taxes, and estate taxes (Fowler *et al.*, 1998). Examples of the commercialization of biodiversity and the business opportunities it offers can be seen in the growth of the wildlife industry and ecotourism sector (**Figure 5**) and more recently in the payments for ecosystem services arena. This growth has created many opportunities for private and communal land to contribute to biodiversity conservation objectives.

Private Conservation Areas

Conservationists have recently begun to realize the value of and need to strengthen the network of private conservation areas (Hale and Lamb, 1997; Knight, 1999; Langholz and Lassoie, 2001; Theobald *et al.*, 2000; Langholz and Krug, 2004). For example, Pence *et al.* (2003) and Gallo *et al.* (2009) demonstrate the biodiversity contributions and cost savings of private conservation areas in South Africa. Pence *et al.* (2003) found that if private conservation areas were used in conjunction with statutory protected areas for conserving biodiversity, then the state would save 80% in acquisition costs. Private conservation areas are known by many other names, for example, private reserve, game reserve, private protected areas, but generally they meet four characteristics: (1) owned by freehold or long-term leasehold by a private investor(s) or syndicate; (2) funded and/or run by a private investor or syndicate; (3) managed for biodiversity and possibly for nature tourism, wildlife-based ventures or leisure; and (4) owned with the intent of preserving the land in a predominantly undeveloped state (Pasquini, 2007).

For the most part, these areas have been left out of conservation statistics, national conservation-planning frameworks, and until recently, generally out of academic research. However, the few studies performed show that thousands of private conservation area owners have demonstrated a willingness and capacity to conserve several million hectares of land (Langholz, 1996; Thackway and Olsson, 1999; Langholz and Lassoie, 2001; Chacon, 2005). Gallo *et al.* (2009) demonstrated that not only do private conservation areas significantly increase the achievement of biodiversity conservation targets (in their case target achievement tripled), but also that private land significantly complemented statutory conservation areas in the types of biomes, elevation classes, and threat status classes conserved. Private conservation areas were especially important in conserving productive lower-elevation habitat, and by association, endangered vegetation. In recognition of the contribution that private lands can make to conservation, several conservation agencies have begun to develop biodiversity stewardship programs to formalize and support the contribution of these areas to conservation (**Box 3**).

In growing and supporting private conservation areas both property rights and financial mechanisms have proved very important. While property rights have helped secure and enhance the contributions from private lands, financial instruments have proved very useful in bringing more landowners into conservation. The wildlife industry has played a substantial role in incentivizing landowners to switch from livestock grazing and other land-use practices to more conservation-aligned land-uses. As Young *et al.* (1996) point out, if clear and enforceable property rights and obligations are created for wildlife with commercial value, then it can be argued that an incentive (for those holding the property rights) will exist to maintain the species and their habitats. But as Scholes *et al.* (2008) argue, even in the absence of clear legal mechanisms, the incentive (financial and cultural) to switch to wildlife-associated land-use is large.

While in many cases hunting was a major trigger of the conversion of uneconomic livestock farms into mixed or wildlife-only farms, later it was 'nonconsumptive' uses such as recreational wildlife viewing, which took over as the major driver. The large expansion of foreign tourism into Africa and other popular wildlife areas has also played a role in maintaining the growth of this sector. **Table 2** below from Scholes *et al.* (2008) provides some indications of the contribution of private wildlife reserves to the numbers of wildlife species in

Namibia. Scholes *et al.* (2008) estimate that in South Africa state-owned protected areas constituted 6% of the land area while the estimate of private wildlife ranches amounts to about 14%.

At a global level, wildlife and nature-based tourism has become the leading foreign exchange earner in several countries (Fillion *et al.*, 1992; The Ecotourism Society, 1998). These studies suggest that between 40% and 60% of international tourists were nature tourists, and that 20–40% of these were wildlife-related tourists. The Ecotourism Society (1998) further suggests that in 1994 there were between 106 million and 211 million wildlife-related tourists worldwide.

Community-Based Natural Resource Management

Another approach to biodiversity conservation outside of protected areas is the communal management of natural resources. The Community-Based Natural Resources Management (CBNRM) approach focuses on the collective management of ecosystems to improve human well-being. It aims to devolve authority for ecosystem management to the local (community) level, thereby empowering communities to manage their own resources without permanently damaging, depleting, or degrading them (Fabricius and Collins, 2007; Roe *et al.*, 2006).

A good number of reviews of CBNRM are critical of the input requirements and costs versus the household level benefits to community members and the benefits to conservation, as well as the high frequency of collapse (Shackleton *et al.*, 2002; Fabricius, 2004; Magome and Fabricius, 2004). Fabricius and Collins (2007) point out that because most CBNRM initiatives are in impoverished, remote rural areas, where human capital (skills and education), financial capital and physical capital (infrastructure and services), are in short supply, CBNRM initiatives, and especially the institutions that are critical to CBNRM's functioning (Agrawal, 2001), are vulnerable to external change like conflicts, climate change, and political change.

In Southern Africa there has been a substantial investment in trying to address some of these shortcomings by making

Table 2 Some trends relating to livestock and wildlife on private land in Namibia

	Changes over time			
Livestock				
Livestock populations				
Year	1971	1981	1991	2001
Cattle ('000)	1800	1400	1300	910
Decline (1971–2001)				−49%
Small stock ('000)	4550	4350	3500	2700
Decline (1971–2001)				−41%
Wildlife				
Huntable game populations				
Year	1972	1982	1992	1997
Huntable game ('000)[a]	565	701	1489	1161
Increase (1972–1997)				105%
Game auctions[b]				
Year		2000	2001	2002
Numbers of game sold		400	1954	3122
Increase (2000–2002)				681%

[a]Numbers for 1972–1992 based on questionnaire survey, 1997 based on aerial survey.
[b]All of game auctioned sold to private landholders in Namibia and South Africa.
Source: Reproduced from Unpublished statistics (2006) Ministry of Agriculture, Water and Forestry, Namibia.

Box 3 Biodiversity stewardship as a means of growing a nation's conservation estate

In many parts of the world the concept of stewardship is helping conservation agencies in working with landowners outside of formal protected areas. Stewardship is basically defined as: "The wise use, management and protection of that which has been entrusted to you as a landowner or is rightfully yours". Biodiversity stewardship is therefore "the practice of effectively managing land-use outside the existing state-managed protected area system to ensure that natural systems, biodiversity and the ecosystem services they provide are maintained and enhanced for present and future generations" (www.stewardship.co.za). In South Africa, if a willing landowner has land in an area of high conservation-value, then an agreement between the land owner and government can be signed and the degree of protection is determined. This degree of protection is determined by national legislation and can include:

- Conservancies
- Protected Environments
- Nature Reserves
- Parts of land to be protected under biodiversity stewardship agreements

A nature reserve is the highest protection possible under this scheme, and the requirements are the most stringent. The Conservancy option is the least onerous on the landowner, and is also the lowest protection. These agreements:

- require commitment from landowners;
- provide long-term security to biodiversity;
- require that sites be managed for conservation;
- require monitoring and auditing of sites; and
- reward landowners for committing land for the public good.

In Australia, Environmental Stewardship uses competitive auctions where private land managers can bid to participate in protecting, rehabilitating, and improving particular ecological communities. Successful land managers are contracted to manage these communities on their land and can receive funding for activities (http://www.nrm.gov.au/stewardship/index.html).

management information available to communities. For example, the Namibian Association of Community-Based Natural Resource Management Support Organizations (NACSO) is an association comprising 15 Non-Government Organizations (NGOs) and the University of Namibia, which aims to provide quality services to rural communities seeking to manage and utilize their natural resources in a sustainable manner (http://www.nacso.org.na/). There are similar programs in Botswana (http://www.cbnrm.bw/) and in the region broadly (http://sacfnet.org/).

Many remain positive about the contribution and need for the devolution of natural resource management to communities. CBNRM's greatest contribution to biodiversity conservation is likely to be through management of habitats and ecosystems in a manner that makes them suitable for the maintenance of wildlife populations. A review of CBNRM in Southern Africa (Nunes *et al.*, 2008) found evidence of the positive impacts of CBNRM on wildlife habitats and populations. For example, in Zimbabwe, 12 rural district councils maintain wildlife habitats ranging from 500 to 5000 km^2 (Roe *et al.*, 2006). In Namibia CBNRM was shown to have contributed to the recovery of wildlife populations, especially large mammals, across large parts of the north (NACSO, 2004). Furthermore, between 1993 and 2003, approximately 3000 head of mixed plains game species had been successfully reintroduced in communal conservancies in this country (Roe *et al.*, 2006). In South Africa a number of communities granted legal ownership of land currently under conservation management during the national land restitution process, have chosen to maintain these conservation areas (Grossman and Holden, 2007). However, beyond these few studies, true impacts are poorly quantified and are largely assumed based on changes in land management practices. This makes definitive conclusions on the the contributions and costs of these areas currently not possible.

In a review of CBNRM in Southern Africa, Fabricius and Collins (2007) list 11 strategies necessary for increasing the success of these initiatives ranging from an improved understanding of the system to be managed; the involvement of local organizations with a clear vision; good planning, monitoring and rule enforcement; management, incentives and financial capacity; and supportive legal frameworks.

While many of these strategies link to the financial, legal, motivational, and other instruments listed above for growing the contribution of areas outside of protected areas, there are some unique challenges when it comes to applying these instruments in communal areas. What is often termed communal or customary tenure is a mix of colonial and customary law and is "the joint creation of colonial officials and African leaders… a reflection of the [then] contemporary situation" (Colson, 1971). Peters (2007) argues that the formation of customary law (as it applies to communal land tenure) in communal areas was shaped by colonial thinking and often served state, European, and elitist African interests. Though individuals do not have private tenure, this does not mean that individuals do not have defined rights to resources. These tenure rights are complex and in many instances can be considered as nested or layered, with different individuals having different rights (Peters, 2007). This makes the application of motivational, legal, or financial instruments quite complex as opposed to the case of a single private owner for whom they were designed.

Payments for Ecosystem Services

While many of the approaches listed above have the potential to contribute substantially to the conservation of biodiversity outside of protected areas, this potential is often constrained by the lack of financial incentives for landowners to set aside land for conservation. Putting private land into conservation can often incur management costs, as well as limit or even lose income from other production opportunities (Frazee *et al.*, 2003). While in some cases conservation agencies can provide some assistance in reducing these costs, for example, through providing management assistance, or through other financial mechanisms like tax rebates, this is often not enough. There has been a recent emergence of other approaches that attempt to grow the range and amount of financial incentives (Fowler *et al.*, 1998). Chief amongst these approaches are payment for ecosystem service (PES) schemes, which compensate individuals or communities for undertaking actions that increase the provision of ecosystem services. Ecosystem services are the benefits that people receive from ecosystems, including flood protection, carbon sequestration, and water flow regulation and purification (MA, 2005). In many cases, land that is managed to conserve biodiversity will also provide many ecosystem services and thus landowners could potentially receive payment for the ecosystem services their land provides (but see Redford and Adams, 2009). As Jack *et al.* (2007) highlight, PES schemes, because they incentivize behavioral change, can be seen as part of the broader class of incentive- or market-based mechanisms. These schemes can bring more private and public funding to conservation, thus expanding available budgets (Jenkins *et al.*, 2005; Turpie *et al.*, 2008).

Working Lands and Waters

Given that agriculture (including crop cultivation, forestry, and aquaculture) now accounts for the greatest proportion of the earth's land surface (~40%), uses most of the world's water (66%) and continues to expand (Foley *et al.*, 2005; Scanlon *et al.*, 2007), a better understanding and engagement with the agriculture-conservation interface is clearly required (Matson *et al.*, 1997). Agricultural activities have had a widespread impact on a variety of habitats resulting in a loss of biodiversity across many taxa (Benton *et al.*, 2003), and simultaneously affecting the production of ecosystem services through ecosystem simplification (Björklund *et al.*, 1999; Tilman *et al.*, 2002; Gordon *et al.*, 2010). Agriculture has substantially modified the global hydrological cycle in terms of both water quality and quantity (Gordon *et al.*, 2010).

In the past 40–50 years agriculture has become a technical process, in which outputs are dependent on external inputs such as pesticides, fertilizer, and fossil-fuel driven technology (Altieri, 1999; Björklund *et al.*, 1999; Edwards and Abivardi, 1998; Pretty, 1995). Ecological processes, which help to promote fertility or control pests and disease, have generally not been considered important in maintaining agricultural production (Edwards and Abivardi, 1998). Much agriculture is therefore currently sitting in the top left hand corner of **Figure 4** with a focus on production uses to the exclusion of any contributions to biodiversity conservation.

Determining the links, potential trade-offs and synergies between conservation and agriculture is now seen as crucial in formulating future conservation strategies, stemming large-scale biodiversity loss, ensuring ecosystem service flows and moving agricultural lands into the center of the framework – as working lands producing food, fiber and other materials, but not to the exclusion of biodiversity. From the current evidence it appears that there are a number of changes necessary in the way agriculture is carried out, if food and fiber security are to be achieved without negative impacts to biodiversity and ecosystem services. The chief of these changes are agricultural management practices and the current focus on monocropping.

Management Practices

While the rate of agricultural expansion has slowed in the past three decades, yields have increased dramatically and have outpaced global human population growth resulting in more food available per capita today than in recent human history (MA, 2005). This progress has been made possible by the use of high-yielding crop varieties, chemical inputs, and a focus on monoculture, often on land already under agricultural management (Matson *et al.*, 1997). Intensification has substantial impacts at all scales including impacts on water quantity, water quality, biodiversity loss, and soil quality. Some of these declines in biodiversity and ecosystem services are of importance to agriculture itself and can have negative effects on agricultural productivity, thereby undermining the long-term sustainability of agricultural intensification. Soil salinization is an example where the regulating ecosystem service (water table regulation) is compromised in a way that feeds back on agriculture (Gordon *et al.*, 2010). This then results in either a collapse of the agricultural system, or as is most often the case, more technological and external inputs in the system, for example, more fertilizers, irrigation, or more pesticides, further exacerbating the problem.

In a recently published review of the impacts of agricultural activities, Rockström *et al.* (2009) point to several "planetary boundaries" or global thresholds, towards which agriculture has forced the global Earth system. Many ecosystems of Earth react in a nonlinear, often abrupt, way, and are particularly sensitive around threshold levels of certain key variables. If these thresholds are crossed, then important subsystems, such as a monsoon system, could shift into a new state, often with deleterious or potentially even disastrous consequences for humans. These boundaries include climate change; rate of biodiversity loss; interference with the nitrogen and phosphorus cycles; global freshwater use; change in land-use; chemical pollution; stratospheric ozone depletion; ocean acidification; and atmospheric aerosol loading nitrogen levels. Of these nine the first six are inextricably linked to agricultural practices.

While agricultural practices have resulted in reductions in biodiversity and ecosystem services, it is also evident that some agricultural ecosystems can be critical repositories of biodiversity, and that the particular type of agricultural practice is a determinant of the biodiversity within that ecosystem (Perfecto *et al.*, 2003; Moguel and Toledo, 1999; Donald, 2004). This biodiversity includes both planned biodiversity (the species of importance to agricultural production, e.g., food crops) and associated biodiversity (Vandermeer *et al.*, 2002). Intensification of agriculture is taken to be the transition from ecosystems with high planned biodiversity to low planned biodiversity and an increase in the use of agrochemicals and machinery (Vandermeer *et al.*, 2002; Perfecto *et al.*, 2003; Perfecto and Vandermeer, 2010). In many cases the application of agrochemicals is to substitute the functions or ecosystem services of some of the associated biodiversity that is eliminated. The impacts on associated biodiversity are less certain and can either start off slowly and increase as intensification increases, or disappear immediately as intensification starts. This uncertainty lies behind much of the debate between intensification versus extensification.

Currently there is a confusing array of options to explore and little consensus on which management practices are best. Some suggest intensification in high-yield areas will conserve more biodiversity in the remaining areas (Green *et al.*, 2005), others suggest low-intensity agriculture over wider areas is better for biodiversity (Donald, 2004). For more discussion on this debate see Perfecto and Vandermeer (2008). In essence this debate rests on the assumption that more productive systems are worse for biodiversity but that is not always the case. Bengtsson *et al.* (2005) demonstrate that organic/agroecological production systems can be just as productive in terms of crop yields but better for biodiversity. Perfecto and Vandermeer (2008) conclude that "it is not only the conversion from a native habitat to agriculture that matters for biodiversity conservation, but also the conversion of agriculture from a biodiversity-friendly type to a biodiversity-unfriendly type that accounts for most biodiversity loss within agricultural landscapes".

Biodiversity-friendly Agriculture

Biodiversity-friendly agricultural practices are basically those that are based on the principles of ecological design. Matson *et al.* (1997) suggest that "ecologically designed agricultural systems that reintegrate features of traditional agricultural knowledge and add new ecological knowledge into the intensification process can contribute to meeting this challenge" (of increased production while avoiding the side effects on biodiversity and other ecosystem services). Such management approaches include integrated nutrient–organic matter management, precision agriculture and pest management approaches. For example, integrated pest management through host resistance and biological control could replace the damaging use of pesticides. No-till agriculture and riparian and wetland restoration are all examples of the recent development of ecologically informed management practices, which can help promote production in agricultural systems while limiting the damaging effects on nontarget species and systems.

The costs and benefits of biodiversity-friendly management practices are not yet clear. But in the case of high costs, the range of financial and other instruments (*see* Living Lands and Waters) could be useful in addressing some costs to

Box 4 South Africa's biodiversity and wine initiative

http://www.wwf.org.za/what_we_do/outstanding_places/fynbos/biodiversity___wine_initiative/

The biodiversity and wine initiative (BWI), which emerged from intense cooperation and collaboration between the South African wine industry and the conservation sector, is a good example of the successes which can be achieved in the area of biodiversity-friendly agriculture. The central tenet of this initiative is to prevent the further loss of biodiversity in the areas actively engaged in wine production by increasing the area within these regions under conservation (Kotze, 2008). This initiative has introduced a stewardship system where producers who establish a contract nature reserve qualify for between 80% and 100% tax rebates. To date a total of 116 farmers formally participate in the BWI (Kotze, 2008). The total area of land conserved by these landholders as a result of this program is 70,411 ha of pristine natural vegetation, incorporating threatened lowland Fynbos, Renosterveld, and Succulent Karoo (Kotze, 2008). This is approximately equivalent to 70% of the footprint of the wine industry in the Western Cape. This initiative has also developed biodiversity best practice guidelines for the wine industry. Guidelines are aimed at changing practices within the industry to enhance suitability of vineyards for biodiversity conservation, for example, guidelines for buffer areas around watercourses, clearing of invasive alien plants, indigenous wildlife introductions, and fire management. The initiative assists farmers in assessing the conservation value of their land and implementing these guidelines.

The opportunities for certification and product differentiation based on conserved biodiversity features are clear, and the BWI has developed a two-pronged approach in this regard. First, the BWI markets itself in conjunction with the farms that are members of this initiative, advertising the farms and their conservation credentials widely. Second, the BWI in association with their stakeholders and members have developed an 'on bottle' labeling system, which allows consumers to easily identify 'conservation wines,' BWI members, and their conservation contributions. Close cooperation and on-farm biodiversity assessments by the BWI ecologists also provide landholders with biodiversity information that can be used in their labeling and marketing strategies.

The initiative is currently undertaking a spatial review assessing the contribution of conserved areas to meeting conservation targets within the broader Cape Floristic Region. This will focus future efforts in prioritizing conservation and stewardship efforts (Kotze, 2008). The BWI has assisted the wine industry in creating and implementing a sustainable production strategy, and the biodiversity guidelines prescribed as part of the industry's Integrated Production of Wine Scheme, have been accepted as part of the international sustainable production guidelines for the global wine industry (Kotze, 2008). The BWI is clearly making great strides and serves as a model for other agriculture-conservation partnerships. Their experiences are already guiding the initiation and development of other schemes including potato and *rooibos* tea farming, as well as the production of honey (Petersen, 2007).



landowners, but also in encouraging consumers to pay more for these products. Certification schemes, labeling, niche markets, monitoring, and evaluation are all areas requiring improved understanding and investment if biodiversity-friendly agriculture is to have the desired impact (O'Farrell *et al.*, 2008). **Box 4** showcases an example of the development of biodiversity-friendly viticulture through a collaboration between the wine industry and the conservation sector in South Africa.

Multifunctional Landscapes

In addition to management practices, which reduce the impact of intensification processes on biodiversity and ecosystem services, there is a need to adopt a more integrated approach when considering food/fiber production or biodiversity conservation in agricultural lands. As Matson *et al.* (1997) point out, one of the main flaws of the traditional practices of intensification is monocropping, which can ultimately undermine productivity

in these areas. Similarly one of the main flaws of traditional approaches to biodiversity conservation is the focus on alpha (local) diversity rather than beta (species turnover across patches in a landscape) diversity, which can ultimately lead to extinctions if species are prevented from moving between habitat patches. As Perfecto and Vandermeer (2008) point out, farming systems, if managed as diverse, organic agroecosystems, can actually help conserve beta diversity by providing the matrix through which species can move. This shift in thinking to a multifunctional diverse landscape has heralded exciting opportunities for conservation in agricultural systems.

The notion of landscape multifunctionality moves away from the traditional management of a single-function landscape manipulated to, for example, either produce food or serve as a recreation area to a landscape offering multiple environmental, social, and economic benefits (de Groot, 2006; Lovell and Johnston, 2009; Wiggering *et al.*, 2006). The design (Nassauer and Opdam, 2008) and management of landscapes with multiple goals including sustainable food production, biodiversity conservation, water production, and job creation holds the potential to improve both production and ecological functions and therefore the longer-term resilience or sustainability of the landscape (McNeely and Scherr, 2003). This is an appealing prospect and will require the consideration of the inherent contributions of various landscape features to multiple goals (Lovell and Johnston, 2009). Furthermore, it requires a thorough understanding of synergies and trade-offs between multiple goals (De Fries *et al.*, 2004; Reyers *et al.*, 2009; Rodríguez *et al.*, 2006), a good knowledge of the social context in terms of stakeholders, institutions, and incentives (Cowling *et al.*, 2008) and the ability to transfer all of this knowledge into the design and establishment of multifunctional landscapes (Nassauer and Opdam, 2008).

In many ways the multifunctional landscape approach embraces the entire spectrum in **Figure 4** in that it supports areas devoted to multiple objectives through a clear regional plan, which takes into account the contribution all areas make to the development and conservation goals and arranges them in the landscape in a way that ensures achievement of multiple societal goals while taking into account interconnections and trade-offs between goals (e.g., Polasky *et al.*, 2008; Lovell and Johnston, 2009; O'Farrell *et al.*, 2010).

Multiple Uses

In many cases examples of living and working lands approaches to conservation include multiple land-uses and tenures in their design and allow for a network of areas and corridors managed for varying uses and conservation goals. These networks can even be associated with formal protected areas in the region.

Biosphere Reserves

Biosphere reserves are sites established by countries and recognized under UNESCO's Man and the Biosphere (MAB) Program to promote sustainable development by combining core protected areas with zones where sustainable development is fostered. Biosphere reserves aim to achieve integrated management of land, fresh and marine waters and living resources by putting in place bioregional planning schemes based on integrating conservation into development through appropriate zoning of core, buffer, and transition zones. There are currently 580 such sites in the world.

Conservancies

There are several other kinds of conservation approach, which mimic the approach of biosphere reserves, but are usually smaller and more local in their management.

There are many names by which these multiple use conservation areas are known including conservancy or mega-conservancy. But in essence they are a collection of mostly agricultural properties, which are cooperatively managed under a single management plan, bound with a voluntary agreement between the landowners and the state. They promote the conservation of certain natural resources, the extension of the regional conservation area network, and increased public awareness regarding the wise management of the natural environment.

They aim to achieve conservation goals over a contiguous patchwork of properties of various tenures and land uses, maximizing landscape heterogeneity and the management of capital flows (e.g., natural, financial, social) (Knight and Cowling, 2003). This can be achieved only if the component properties are managed in a coordinated, cooperative, and integrated way (Pierce *et al.*, 2005). These areas ideally cover important sensitive natural areas so as to ensure that landscape-scale ecological processes, such as the movement of animals and water flows, function effectively. This better ensures that the goals of agricultural production, water use, and biodiversity conservation can be aligned and achieved in a cooperative way. The result is more viable environmental, social, and economic systems – or in other words, living landscapes (**Figure 6**).

As Knight and Cowling (2003) point out the management of natural resources and biodiversity in conservancies is usually informed by integrated instruments e.g., integrated catchment plans, integrated development plans. These approaches ensure shared goals across these resource management fields in order to ensure that objectives are aligned and integrated. Integrated management requires that the condition of resources is regularly monitored and evaluated and new information from this process is used for refining land management activities. A wide range of instruments are used to achieve such conservancy goals and to encourage participation by stakeholders. Examples of instruments include tenure (e.g., protected areas), zoning categories, tax relief, legislation, land-use policy, market incentives, and conservation and land-use programs.

The Contribution of Living and Working Lands to Conservation

As many conservationists would point out: land managed for wildlife or for tourism or even for grazing is not synonymous with biodiversity conservation. Many have expressed concerns about these activities and their negative impacts on

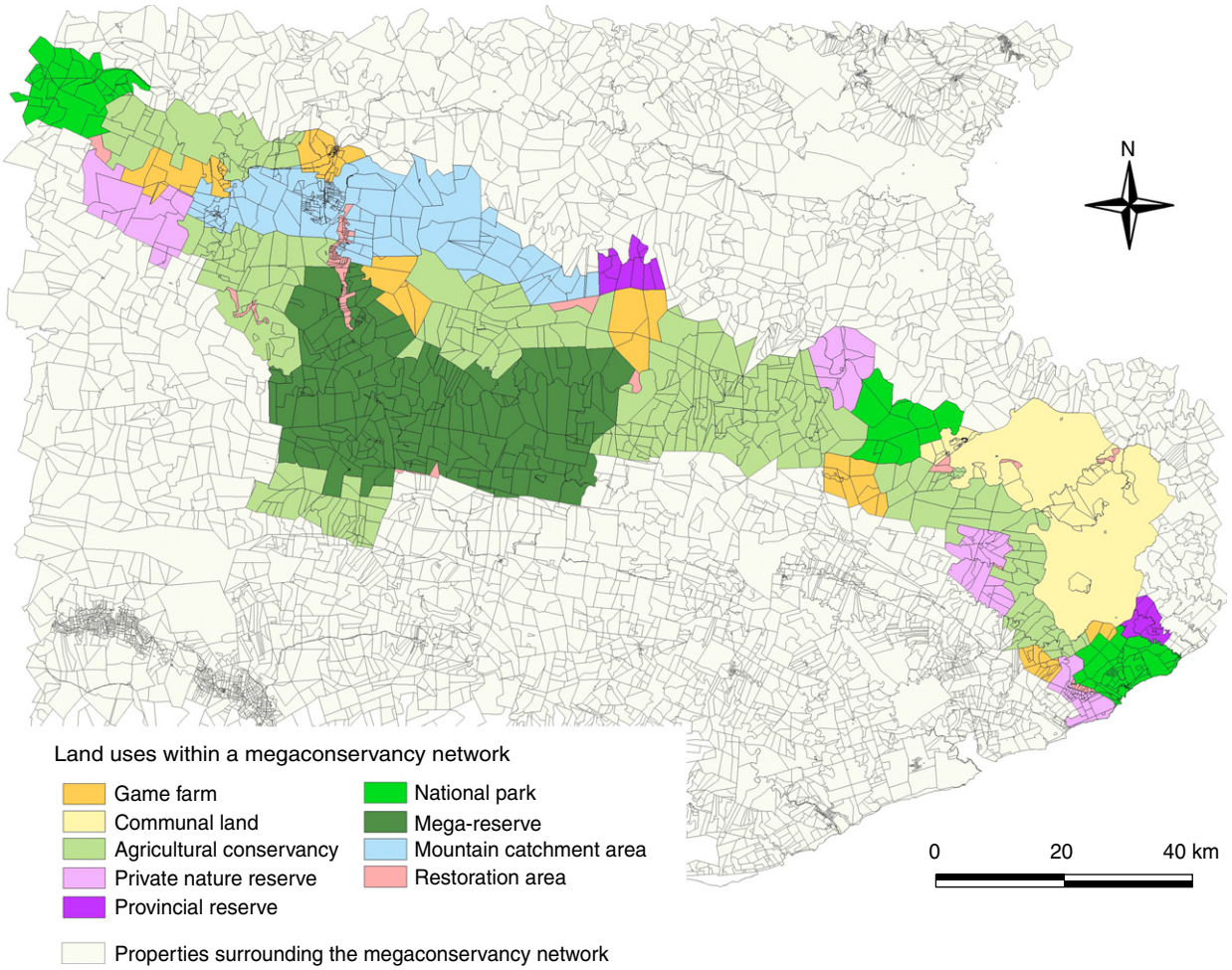


Figure 6 Hypothetical conservancy map showing multiple land uses possible within such an area. Reproduced from Knight AT and Cowling RM (2003) *The Megaconservancy Network concept: 'Keeping people on the Land in Living Landscapes'*, Terrestrial Ecology Research Unit Report No. 45. University of Port Elizabeth.

biodiversity (e.g., Dalal-Clayton *et al.*, 1997) as well as their lack of security as a means for conservation. But while these concerns might be true in some cases, they also apply to some formal protected areas where processes like land reform, land development, and the degazetting of protected areas threatens biodiversity and the security of these areas (Brandon *et al.*, 1998; Bruner *et al.*, 2001). The benefits of such areas may also outweigh their supposed disadvantages as is evidenced by the contribution they can make to biodiversity conservation, often conserving biodiversity that is not protected in formal protected areas. Gallo *et al.* (2009) found that while private conservation areas were more degraded and transformed than formal protected areas, these minor shortcomings were offset by the overall representation and complementarity benefits provided by private areas. Further, in regions with poor enforcement of protected area boundaries (e.g., Liu *et al.*, 2001), private and communal conservation areas (which are enforced by the landowner themselves) may actually be more pristine than other protected areas, as well as helping constrained conservation budgets do more with less (e.g., Yuan *et al.*, 2008). Several of the motivational and educational instruments can also help contribute to improving the longevity of the contribution made by private and communal protected areas (Chacon, 2005; Winter *et al.*, 2005).

It would appear that the contribution of private and communal living lands to conservation is potentially large, not only in the extent of land and water, people and resources it could offer to conservation, but also in the complement of biodiversity found in these areas usually absent from state owned and formal protected areas. But in order to realize this potential there is a need for more information on these areas where monitoring and evaluation are currently rare. This information would allow for the assessment of their contribution to biodiversity conservation, as well as other development goals, and for improved learning on the effects of a wide variety of management practices on biodiversity. All reviews highlight the need for improvements in the reliability of information on the location, extent, and management practices on private and communal land. There is a further need to ensure that information and necessary skills are made available to the managers of biodiversity on private and communal land to improve the track record of these areas in the arena of biodiversity conservation. Finally, conservationists need to engage more with the wide range of instruments

available for extending the conservation effort outside of protected areas and begin to engage with other disciplines that better understand issues like incentives, motivation, and tax reform.

In distilling what appears to be critical in ensuring the success of these living and working lands, a few key ingredients emerge from a review of the literature (O'Farrell *et al.*, 2008; Knight and Cowling, 2003; Donaldson, 2002; Turpie *et al.*, 2008):

- A vibrant conservation sector with a strategically located champion integrally linked to the emerging initiative.
- A strong collaboration between the conservation and production (e.g., agriculture, wildlife, or forestry) sectors.
- Knowledge dissemination and extension services to landowners to promote the understanding of trade-offs associated with management decisions.
- A critical mass of support for the initiative coupled to the establishment of an 'anchor tenant/s' whom the initiative can be built around.
- Clearly perceived benefits from the maintenance of biodiversity, which outweigh the costs of conserving biodiversity.
- Enabling mechanisms such as subsidies, extension services, and research. These have been found to often be more effective at achieving a change in behavior compared with legal instruments.
- Clearly distinguishable products emerging from the intervention or initiative, for example, labeling or certification.
- Well-assessed market identification and public/consumer demand.
- Restoration programs (often partnered with job creation programs) which focus on biodiversity and development targets.
- A sound legislative and policy framework (essential, but by itself not enough).
- Promotion of incentives, tourism, and other biodiversity-based business opportunities as a means to offset costs.
- Monitoring and evaluation incorporated into interventions, and geared toward assessing and evaluating both the biological and socioeconomic implications of interventions at appropriate stages.

See also: Biodiversity-Friendly Farming. Biodiversity, Human Well-Being, and Markets. Commons, Concept and Theory of. Ecosystem Services. Feeding the World and Protecting Biodiversity. Identifying Conservation Priorities Using a Return on Investment Analysis. Natural Reserves and Preserves. Stewardship, Concept of. Tourism, Role of

References

Agrawal A (2001) Common property institutions and sustainable governance of resources. *World Development* 29: 1649–1672.

Altieri MA (1999) The ecological role of biodiversity in agroecosystems. *Agriculture Ecosystems and Environment* 74: 19–31.

Bengtsson J, Ahnström J, and Weibull AC (2005) The effects of organic agriculture on biodiversity and abundance: A meta-analysis. *Journal of Applied Ecology* 42: 261–269.

Benton TG, Vickery JA, and Wilson JD (2003) Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology and Evolution* 18: 182–188.

Björklund J, Limburg KE, and Rydberg T (1999) Impact of production intensity on the ability of the agricultural landscape to generate ecosystem services: An example from Sweden. *Ecological Economics* 29: 269–291.

Boyd J, Caballero K, and Simpson RD (1999) The law and economics of habitat conservation: Lessons from an analysis of easement acquisitions. *Resources for the Future Discussion Paper* 99: 1–42.

Brandon K, Sanderson S, and Redford K (1998) *Parks in Peril: People, Politics, and Protected Areas*, 532 pp. Nature conservancy. Island Press.

Brooks TM, Da Fonseca GAB, and Rodrigues ASL (2004) Protected areas and species. *Conservation Biology* 18: 616–618.

Bruner G, Gullison R, Rice R, and da Fonseca G (2001) Effectiveness of parks in protecting tropical biodiversity. *Science* 29: 125–128.

Butchart SH, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328: 1164–1168.

Chacon CM (2005) Fostering conservation of key priority sites and rural development in Central America: The role of private protected areas. *Parks* 15: 39–47.

Chandrashekara UM and Sankar S (1998) Ecology and management of sacred groves in Kerala, India. *Forest Ecology and Management* 112: 165–177.

Chape S, Harrison J, Spalding M, and Lysenko I (2005) Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society B – Biological Sciences* 360: 443–455.

Chape S, Spalding M, and Jenkins MD (2008) *The World's Protected Areas: Status, Values and Prospects in the 21st Century.* Berkeley, USA: University of California Press. Prepared by the UNEP World Conservation Monitoring Centre.

Colson E (1971) The impact of the colonial period on the definition of land rights. In: Turner V (ed.) *Colonialism in Africa*, pp. 1870–1960. Cambridge: Cambridge University Press.

Convention on Biological Diversity (2010) *Global Biodiversity Outlook 3.* http://www.cbd.int/gbo/gbo3/.

Cowling RM, Egoh B, Knight AT, *et al.* (2008) An operational model for mainstreaming ecosystem services for implementation. *Proceedings of the National Academy of Sciences* 105: 9483–9488.

Crompton J (1990) Protecting park and natural areas without purchasing term: A review of methods adopted in the USA. *Journal of Society and Leisure* 13: 431–453.

Dalal-Clayton B, Leader-Williams N, and Roe D (1997) *Take Only Photographs, Leave Only Footprints: The Environmental Impacts of Wildlife Tourism.* London: International Institute for Environment and Development. IIED Wildlife and Development Series.

De Fries RS, Asner GP, and Houghton R (2004) Trade-offs in landuse decisions: Towards a framework for assessing multiple ecosystem responses to land-use change. *Ecosystem Land Use Change* 153: 1–9.

de Groot R (2006) Function-analysis and valuation as a tool to assess land use conflicts in planning for sustainable, multifunctional landscapes. *Landscape Urban Plan* 75: 175–186.

Dixon JA and Sherman PB (1991) *Economics of Protected Areas: A New Look at Benefits and Costs.* East-West Centre: Island Press.

Donald PF (2004) Biodiversity impacts of some agricultural commodity production systems. *Conservation Biology* 18: 17–38.

Donaldson JS (2002) Biodiversity and conservation farming in the agricultural sector. In: Pierce SM, Cowling RM, Sandwith T, and MacKinnon K (eds.) *Mainstreaming Biodiversity in Development. Case Studies from South Africa*, pp. 43–55. Washington: The World Bank.

Driver A, Cowling RM, and Maze K (2003) *Planning for Living Landscapes: Perspectives and lessons from South Africa.* Center for Applied Biodiversity Science at Conservation International, Washington DC and the Botanical Society of South Africa, Cape Town. http://www.botanicalsociety.org.za/default.php?pageID=ccu/ccu_downloads.htm

Dudgeon D, Arthington AH, Gessner MA, *et al.* (2006) Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews* 81: 163–182.

Ecotourism Society (1998) *Ecotourism Statistical Fact Sheet: General Tourism Statistics.* Washington, DC.

Edwards PJ and Abivardi C (1998) The value of biodiversity: Where ecology and economy blend. *Biological Conservation* 83: 239–246.

Erasmus BFN, Van Jaarsveld AS, Chown SL, Kshatriya M, and Wessels KJ (2002) Vulnerability of South African animal taxa to climate change. *Global Change Biology* 8: 679–693.

Fabricius C (2004) The fundamentals of community-based natural resource management. In: Fabricius C, Koch E, Magome H, and Turner S (eds.) *Rights, Resources and Rural Development: Community-based Natural Resource Management in Southern Africa*, pp. 3–43. London: Earthscan.

Fabricius C and Collins S (2007) Community-based natural resource management: Governing the commons. *Water Policy* 9: 83–97.

Fillion FL, Foley JP, and Jaquemot AJ (1992) *The Economics of Global Tourism*. Paper presented at the Fourth World Congress on National Parks and Protected Areas. Caracas. Venezuela, 10–12 February 1992.

Foley JA, DeFries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.

Fowler L, Neuhauser H, and Baker B (1998) *A Landowner's Guide: Conservation Easements for Natural Resource Protection*, 2nd edn. Georgia Environmental Policy Institute, Georgia Land Trust Service Center Resource Paper No. 1.

Frazee SR, Cowling RM, Pressey RL, Turpie JK, and Lindenberg N (2003) Estimating the costs of conserving a biodiversity hotspot: A case-study of the Cape Floristic Region, South Africa. *Biological Conservation* 112: 275–290.

Gallo JA, Pasquini L, Reyers B, and Cowling RM (2009) The role of private conservation areas in biodiversity representation and target achievement within the Little Karoo region, South Africa. *Biological Conservation* 142: 446–454.

Gordon L, Finlayson CM, and Falkenmark M (2010) Managing water in agriculture for food production and other ecosystem services. *Agricultural Water Management* 97: 512–519.

Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–555.

Grossman D and Holden P (2007) *Case Studies on Successful Southern African NRM Initiatives and their Impacts on Poverty and Governance: South Africa*. IUCN/USAID FRAMEwww.frameweb.org

Hale P and Lamb D (eds.) (1997) *Conservation Outside Nature Reserves*. Centre for Conservation Biology. The University of Queensland.

Hoekstra JM, Boucher TM, Ricketts TH, and Roberts C (2005) Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters* 8: 23–29.

IUCN (1994) *Guidelines for Protected Area Management Categories*. Gland and Cambridge: IUCN.

Jack BK, Kousky C, and Sims KR (2007) Designing payments for ecosystem services: Lessons from previous experience with incentive-based mechanisms. *Proceedings of the National Academy of Sciences* 28: 9465–9470.

Jenkins M, Scherr S, and Inbar M (2005) Markets for biodiversity services. *Environment* 46: 32–42.

Kessler WB and Thomas JW (2006) Conservation biology from the perspective of natural resource management disciplines. *Conservation Biology* 20: 670–673.

Knight AT and Cowling RM (2003) *The Megaconservancy Network concept: 'Keeping people on the Land in Living Landscapes'*, Terrestrial Ecology Research Unit Report No. 45. University of Port Elizabeth.

Knight RL (1999) Private lands: The neglected geography. *Conservation Biology* 13: 223–224.

Kotze I (2008) *BWI Progress Report*. May 2008. In. Biodiversity and Wine Initiative, Stellenbosch. www.wwf.org.za/bwi.

Langholz J (1996) Economics, objectives, and success of private nature reserves in sub-Saharan Africa and Latin America. *Conservation Biology* 10: 271–280.

Langholz JA and Krug W (2004) New forms of biodiversity governance: Non-state actors and the private protected area action plan. *Journal of International Wildlife Law and Policy* 9: 7.

Langholz JA and Lassoie JP (2001) Perils and promise of privately owned protected areas. *Bioscience* 51: 1079–1085.

Lévêque C and Balian EV (2005) Conservation of freshwater biodiversity: Does the real world meet scientific dreams? *Hydrobiologia* 542: 23–26.

Lindenmayer DB and Burgman MA (2005) *Practical Conservation Biology*. Melbourne: CSIRO Press. 609 p. ISBN 0 643 09089 4.

Liu J, Linderman M, Ouyang Z, An L, Yang J, and Zhang H (2001) Ecological degradation in protected areas: The case of Wolong nature reserve for giant pandas. *Science* 292: 98–101.

Lovell ST and Johnston DM (2009) Creating multifunctional landscapes: How can the field of ecology inform the design of the landscape? *Frontiers in Ecology and the Environment* 7: 212–220.

Mace GM, Masundire H, Baillie J, *et al.* (2005) *Biodiversity. Millennium Ecosystem Assessment: Current State and Trends*. Washington, DC: Island Press. ISBN 1-55963-228-3.

Magome H and Fabricius C (2004) Reconciling biodiversity conservation with rural development: The Holy Grail of CBNRM? In: Fabricius C, Koch E, Magome H, and Turner S (eds.) *Rights, Resources and Rural Development: Community-based Natural Resource Management in Southern Africa*, pp. 93–114. London: Earthscan.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.

Matson PA, Parton WJ, Power AG, and Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 277: 504–509.

Mawdsley JR, O'Malley R, and Ojima DS (2009) A review of climate-change adaptation strategies for wildlife management and biodiversity conservation. *Conservation Biology* 23: 1080–1089.

McNeely J, Albers H, Faith DP, *et al.* (2005) *Biodiversity. Policy Responses of the Millennium Ecosystem Assessment, Ecosystems and Human Well-being, Vol. 3 Millennium Ecosystem Assessment*. Washington, DC: Island Press. ISBN 1–55963-270-4.

McNeely J and Scherr S (2003) *Ecoagriculture: Strategies to feed the World and Save Wild Biodiversity*. Washington, DC: Island Press.

Millennium Ecosystem Assessment (MA) (2005) *Ecosystems and Human Well-being: Current State and Trends*, vol. 1. Washington, USA: Island Press.

Moguel P and Toledo VM (1999) Biodiversity conservation in traditional coffee systems of Mexico. *Conservation Biology* 12: 1–11.

Moilanen A, Wilson KA, and Possingham HP (2009) *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools*. Oxford: Oxford University Press.

NACSO (2004) *Namibia's Communal Conservancies: A Review of Progress and Challenges*. Windhoek: Namibia Association of CBNRM Support Organisation.

Nassauer JI and Opdam P (2008) Design in science: Extending the landscape ecology paradigm. *Landscape Ecology* 23: 633–644.

Nel JL, Reyers B, Roux DJ, Cowling RM, and Impson ND (2011) Incorporating natural processes into freshwater conservation planning. *Freshwater Biology* 56: 106–124.

Nel JL, Roux DJ, Maree G, *et al.* (2007) Rivers in peril inside and outside protected areas: A systematic approach to conservation assessment of river ecosystems. *Diversity and Distributions* 13: 341–352.

Norton DA (2000) Conservation biology and private land: Shifting the focus. *Conservation Biology* 14: 1221–1223.

Nunes I, von Maltitz G, and Zeidler J (2008) A natural resource tenure response in southern Africa – taking stock of biodiversity conservation and human-well being contributions of the approach In: *Promoting Human Well-Being while Protecting Biodiversity: Some Innovative Approaches from Southern Africa*. Report for Netherlands Environmental Assessment Agency. CSIR/NRE/ECO/ER/2008/0122/.

O'Farrell PJ, Anderson PML, and van Wilgen BW (2008) Biodiversity interventions in South African agricultural landscapes. In: *Promoting Human Well-Being while Protecting Biodiversity: Some Innovative Approaches from Southern Africa*. Report for Netherlands Environmental Assessment Agency. CSIR/NRE/ECO/ER/2008/0122/.

O'Farrell PJ, Reyers B, Le Maitre DC, *et al.* (2010) Multi-functional landscapes in semi arid environments: Implications for biodiversity and ecosystem services. *Landscape Ecology* 25: 1231–1246.

Pasquini L (2007) *Privately-owned Lands and Biodiversity Conservation: Analysing the Role of Private Conservation Areas in the little Karoo, South Africa*. PhD Dissertation. University of Sheffield.

Pence GQK, Botha MA, and Turpie JK (2003) Evaluating combinations of on-and off-reserve conservation strategies for the Agulhas Plain, South Africa: A financial perspective. *Biological Conservation* 112: 253–273.

Perfecto I and Vandermeer J (2008) Biodiversity conservation in tropical agroecosystems: A new conservation paradigm. *Annals of the New York Academy of Science, (The Year in Ecology and Conservation Biology 2008)* 1134: 173–200.

Perfecto I and Vandermeer J (2010) The agricultural matrix as an alternative to the land-sparing/agricultural intensification model: Facing the food and biodiversity crises. *Proceedings of the National Academy of Science* 107: 5786–5791.

Perfecto I, Mas A, Dietsch TV, and Vandermeer J (2003) Species richness along an agricultural intensification gradient: A tri-taxa comparison in shade coffee in southern Mexico. *Biodiversity and Conservation* 12: 1239–1252.

Peters PE (2007) *Challenges in Land Tenure and Land Reform in Africa: An Anthropological perspective*. CID Working Paper No. 141.

Petersen C (2007) *The Business Case for Biodiversity and Good Biodiversity Practice in the Republic of South Africa*. Report for submission to the Secretariat of the Convention on Biological Diversity in preparation for the ninth meeting of the Conference of the Parties Ref.: SCBD/OMG/NB/57847. Prepared for the

Department of Environmental Affairs and Tourism by the South African National Biodiversity Institute. 23 March 2007.
Pierce SM, Cowling RM, Knight AT, Lombard AT, Rouget M, and Wolf T (2005) Systematic conservation assessment products for land-use planning: Interpretation for implementation. *Biological Conservation* 125: 441–548.
Polasky S, Nelson E, Camm J, *et al.* (2008) Where to put things? Spatial land management to sustain biodiversity and economic returns. *Biological Conservation* 141: 1505–1524.
Pressey RL, Ferrier S, Hager TC, Woods CA, Tulley SL, and Weiman KM (1996) How well protected are the forests of north-eastern New South Wales analyses of forest environments in relation to tenure, formal protection measures and vulnerability to clearing. *Forest Ecology and Management* 85: 311–333.
Pressey RL and Logan VS (1997) Inside looking out: Findings of research on reserve selection relevant to "off-reserve" nature conservation. In: Hale P and Lamb D (eds.) *Conservation Outside Nature Reserves*, pp. 407–418. Brisbane: University of Queensland, Centre for Conservation Biology.
Pretty JN (1995) Participatory learning for sustainable agriculture. *World Development* 23: 1247–1263.
Redford KH and Adams WM (2009) Payment for ecosystem services and the challenge of saving nature. *Conservation Biology* 23: 785–787.
Redford KH, Coppolillo P, Sanderson EW, *et al.* (2003) Mapping the conservation landscape. *Conservation Biology* 17: 116–131.
Reyers B, O'Farrell PJ, Cowling RM, Egoh BN, Le Maitre DC, and Vlok JHJ (2009) Ecosystem services, land-cover change, and stakeholders: Finding a sustainable foothold for a semiarid biodiversity hotspot. *Ecology and Society* 14: 38. http://www.ecologyandsociety.org/vol14/iss1/art38/.
Richter BD, Mathews R, Harrison DL, and Wigington R (2003) Ecologically sustainable water management: Managing river flows for ecological integrity. *Ecological Applications* 13: 206–224.
Rockström J, Steffen W, Noone K, *et al.* (2009) A safe operating space for humanity. *Nature* 461: 472–475.
Rodrigues ASL, Andelman SJ, Bakarr MI, *et al.* (2004) Effectiveness of the global protected area network in representing species diversity. *Nature* 428: 640–643.
Rodríguez JP, Beard Jr. TD, Bennett EM, *et al.* (2006) Trade-offs across space, time, and ecosystem services. *Ecology and Society* 11: 28. http://www.ecologyandsociety.org/vol11/iss1/art28/.
Roe D, Jones B, Bond I, and Bhatt S (2006) *Local Action, Global Aspirations: The Role of Community Conservation in Achieving International Goals for Environment and Development*. London, UK: International Institute for Environment and Development.
Scanlon BR, Jolly I, Sophocleous M, and Zhang L (2007) Global impacts of conversions from natural to agricultural ecosystems on water resources: Quantity versus quality. *Water Resources Research* 43: W03437.http://dx.doi.org/10.1029/2006WR005486.
Scholes RJ, Grossman D, and Barnes J (2008) The shift from domestic livestock to wildlife-based land use in South Africa and Namibia, 1970–2008. In: *Promoting Human Well-being While Protecting Biodiversity: Some Innovative Approaches from Southern Africa*. Report for Netherlands Environmental Assessment Agency. CSIR/NRE/ECO/ER/2008/0122/.
Schwartzman S, Nepstad D, and Moreira A (2000) Arguing tropical forest conservation: People versus parks. *Conservation Biology* 14: 1370–1374.
Scott JM, Davis FW, McGhie RG, Wright RG, Groves C, and Estes J (2001) Nature reserves: Do they capture the full range of America's biological diversity? *Ecological Applications* 11: 999–1007.
Shackleton S, Campbell B, Wollenberg E, and Edmunds D (2002) *Devolution and community-based Natural Resource Management: Creating Space for Local People to Participate and Benefit?*. London: Overseas Development Institute.
Thackway R and Olsson K (1999) Public/private partnerships and protected areas: Selected Australian case studies. *Landscape and Urban Planning* 44: 87–97.
Theobald DM, Hobbs NT, Bearly T, Zack JA, Shenk T, and Riebsame. WE (2000) Incorporating biological information in local landuse decision making: Designing a system for conservation planning. *Landscape Ecology* 15: 35–45.
Tilman D, Cassman KG, Matson PA, Naylor R, and Polasky S (2002) Agricultural sustainability and intensive production practices. *Nature* 418: 671–677.
Turpie JK, Marais C, and Blignaut JN (2008) The working for water program: Evolution of a payments for ecosystem services mechanism that addresses both poverty and ecosystem service delivery in South Africa. *Ecological Economics* 65: 788–798.
UNESCO (1995) *The Vision from Seville for the 21st century: The Seville Strategy for Biosphere Reserves*. Paris: United Nations Educational, Scientific, and Cultural Organization.
Unpublished Statistics (2006) Ministry of Agriculture, Water and Forestry, Namibia.
Vandermeer J, Lawrence D, Symstad A, and Hobbie S (2002) Effect of biodiversity on ecosystem function in managed ecosystems. In: Loreau M, Naeem S, and Inchausti P (eds.) *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*, pp. 221–235. Oxford, UK: Oxford University Press.
Wallis de Vries MF, Bakker JP, and Van Wieren SE (1998) *Grazing and Conservation Management*. Chapman and Hall.
Wells M and Brandon K (1993) The principles and practice of buffer zones and local participation. *Ambio* 22: 157–162.
Wiggering H, Dalchow C, Glemnitz M, *et al.* (2006) Indicators for multifunctional land use – linking socio-economic requirements with landscape potentials. *Ecological Indicators* 6: 238–249.
Winter SJ, Esler KJ, and Kidd M (2005) An index to measure the conservation attitudes of landowners towards Overberg Coastal Renosterveld, a critically endangered vegetation type in the Cape Floral Kingdom, South Africa. *Biological Conservation* 126: 383–394.
WRI (1996) *Balancing the Scales: Guidelines for Increasing Biodiversity's Chances Through Bioregional Management*. USA: World Resources Institute.
Young M and Gunningham N (1997) Mixing instruments and institutional arrangements for optimal biodiversity conservation. In: Hale P and Lamb D (eds.) *Conservation Outside Nature Reserves*, pp. 123–135. Brisbane: Centre for Conservation Biology, University of Queensland.
Young MD, Gunningham N, Elix J, *et al.* (1996) *Reimbursing the Future: An evaluation of Motivational, Voluntary, Price-based, Property-right, and Regulatory Incentives for the Conservation of Biodiversity*. Department of the Environment, Sport and Territories Biodiversity Unit Biodiversity Series Paper No. 9.
Yuan J, Dai L, and Wang Q (2008) State-led ecotourism development and nature conservation: A case study of the Changbai Mountain Biosphere Reserve, China. *Ecology and Society* 13(2): 55. http://www.ecologyandsociety.org/vol13/iss2/art55/.

Convention on Biological Diversity

H David Cooper and Kieran Noonan-Mooney, Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ad-Hoc Open-ended Working Group on Review of Implementation A Convention body established to consider progress in the implementation of the Convention, to review the impacts and effectiveness of existing processes under the Convention, national focal points and the Secretariat, as part of the overall process for improving the operations of the Convention.

Aichi Biodiversity Targets Twenty internationally-agreed outcome-oriented targets included in the Strategic Plan for Biodiversity 2011–2020 setting the agenda for achievements at the global level and providing a framework for the establishment of national or regional targets.

Article 8(j) An article of the text of the Convention which calls for the respect, preservation and maintenance of knowledge, innovations and practices of indigenous and local communities embodying traditional lifestyles relevant for the conservation and sustainable use of biological diversity.

Biodiversity The Convention on Biological Diversity defines biodiversity as the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems, and the ecological complexes of which they are part; this includes diversity within species, between species, and of ecosystems.

Conference of the Parties The Conference of the Parties is the governing body of the Convention on Biological Diversity, and advances implementation of the Convention through the decisions it takes at its periodic meetings. It is composed of the countries that are Party to the Convention.

Convention on Biological Diversity The Convention on Biological Diversity is a framework for national and cooperative action for the conservation of biodiversity, the sustainable use of its components, and the equitable sharing of benefits arising out of the utilization of genetic resources. It requires contracting Parties to ensure that actions to achieve its objectives are undertaken at all levels and in all sectors, in a coordinated and crosscutting fashion.

Ecosystem approach A strategy for the integrated management of land, water and living resources that promotes conservation and sustainable use in an equitable way. It is based on the application of appropriate scientific methodologies focused on levels of biological organization which encompass the essential processes, functions and interactions among organisms and their environment. It recognizes that humans, with their cultural diversity, are an integral component of ecosystems.

Global Biodiversity Outlook The flagship publication of the Convention on Biological Diversity and currently in its third edition. The Outlook draws on a range of information to summarize the status and trends of biodiversity and draws conclusions for the further implementation of the Convention.

Global Environment Facility The financial mechanism of the Convention on Biological Diversity which provides grants to developing countries and countries with economies in transition for projects which benefit the global environment.

National Biodiversity Strategies and Action Plans The process by which countries can plan to address the threats to their biodiversity and implement the Convention on Biological Diversity and related instruments.

Party A country that has ratified, acceded to, approved or accepted the Convention is considered a Party to it.

Programmes of Work Programmes of Work establish a vision for, and basic principles to guide work under the Convention by setting out key issues for consideration, identifying potential outputs and means for achieving these.

Strategic Plan for Biodiversity 2011–2020 The Strategic Plan promotes the coherent and effective implementation of the three objectives of the Convention on Biological Diversity and serves as a flexible framework for the establishment of national and regional targets.

Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA) An intergovernmental scientific advisory body that provides the Conference of the Parties to the Convention on Biological Diversity and its other subsidiary bodies with advice relating to the implementation of the Convention. Multidisciplinary and open to participation by all Parties, SBSTTA comprises government representatives competent in the relevant field of expertise.

Tipping points A situation in which an ecosystem experiences a shift to a new state, with significant changes to biodiversity and the services to people it underpins, at a regional or global scale.

Introduction

The Convention on Biological Diversity (CBD) is the key international treaty for the conservation and sustainable use of biodiversity and for the fair and equitable sharing of the benefits arising out of the use of genetic resources. One-hundred and ninety-two countries, and the European Union are Party to the Convention which is implemented primarily through action at the national level. The governing body of the Convention reviews implementation and develops policy to guide national implementation and promote international cooperation. Two main subsidiary legal agreements have been adopted under the convention: the Cartagena Protocol on Biosafety and the Nagoya Protocol on Access and Benefit

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00418-4

Sharing. In 2010, a new Strategic Plan for Biodiversity 2011–2020 was adopted, providing a 10-year framework of action by all stakeholders. The Plan includes 20 outcome-oriented Aichi Biodiversity Targets. The major challenges for the implementation of the Plan relate to the need to address the pressures on biodiversity and the underlying causes of biodiversity loss by mainstreaming biodiversity across government and society.

The CBD: Main Instruments

Concern over the rapid loss of biodiversity and the realization that it plays a fundamental role in supporting human life motivated the creation of the CBD, a legally binding global treaty. Opened for signature at the Earth Summit in Rio de Janeiro in 1992 and entering into force in 1993, the Convention arose from an international dialogue begun a decade earlier by the World Commission on Environment and Development (known as the Brundtland Commission).

The word biodiversity, a contraction of the synonymous phrase "biological diversity", is defined by the CBD as "the variability among living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems" (United Nations, 1992). As such the Convention is holistic, covering all aspects of biodiversity, and was the first international treaty to acknowledge the role of biodiversity in sustainable development.

The Convention objectives are "the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding." The Convention is composed of 42 Articles that provide for principles (e.g., sovereign rights over national resources and the responsibility to ensure that activities within their jurisdiction or control do not cause damage to the environment of other States or of areas beyond the limits of national jurisdiction), substantive obligations (e.g., for *in situ* and *ex situ* conservation, respect for indigenous knowledge, sustainable use, use of environmental impacts assessment, use of incentive measures) and procedural arrangements (e.g., meetings of the governing body and the operations of its Secretariat). These articles form the basis of the structure and function of the Convention.

The Convention on Biological Diversity has near-universal participation of countries and is recognized by the United Nations General Assembly as the primary international instrument for biodiversity. To date, 192 countries and the European Union have ratified the Convention, constituting the 193 Parties. These form the governing body of the Convention referred to as the Conference of the Parties (COP) which meets periodically (currently, every 2 years). The United States which has signed the Convention but not yet ratified it, participates as an observer. International organizations and representatives of indigenous and local communities and civil society also participate in the work of the Convention.

The Conference of the Parties reviews implementation of the Convention on the basis of national reports and the Global Biodiversity Outlook – a periodic publication that assesses the state of biodiversity, actions taken under the Convention to achieve its objectives and future prospects for biodiversity (*see* The Global Biodiversity Outlook: Assessing Progress and Providing the Basis for Future Action). The COP also advances implementation of the Convention through the decisions it takes at its periodic meetings. In its first 10 meetings, the Conference of the Parties has taken a total of 299 decisions. Among these are key decisions that outline the major areas of work of the Convention, establish principles and guidelines for action, and set out a plan for the more effective and coherent implementation of the Convention as a whole (*see* The CBD Regime: Supporting Action at the International Level).

The Conference of the Parties has held 10 ordinary meetings, and 1 extraordinary meeting (the latter, to adopt the Cartagena Protocol on Biosafety (see **Box 1**) was held in two parts). At its 10th meeting, the Conference of the Parties adopted the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization (see **Box 2**). The 11th meeting of the Conference of the Parties takes place in Hyderabad, India, in October 2012.

The Conference of the Parties is supported by several additional Convention bodies. One of these bodies is an open-ended intergovernmental scientific advisory body known as the Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA). SBSTTA comprises government representatives of all Parties as well as observers and its functions

Box 1 Cartagena Protocol on Biosafety to the CBD

The Cartagena Protocol on Biosafety to the CBD is an international treaty governing the movements from one country to another of living modified organisms resulting from modern biotechnology (more commonly known as genetically modified organisms). It was adopted on 29 January 2000 as a supplementary agreement to the CBD and entered into force on 11 September 2003. As of the end of 2011, there are 162 Parties to the Protocol.

The Protocol establishes an advance informed agreement procedure for ensuring that countries are provided with the information necessary to make informed decisions before agreeing to the import of such organisms into their territory. The Protocol contains reference to a precautionary approach and reaffirms the precaution language in Principle 15 of the Rio Declaration on Environment and Development. The Protocol also establishes a Biosafety Clearing-House to facilitate the exchange of information on living modified organisms and to assist countries in the implementation of the Protocol.

During a meeting of the governing body of the Cartagena Protocol on Biosafety (known as the COP serving as the meeting of the Parties to the Protocol), a new international treaty, the Nagoya – Kuala Lumpur Supplementary Protocol on Liability and Redress to the Cartagena Protocol on Biosafety, was adopted. The supplementary Protocol provides international rules and procedure on liability and redress for damage to biodiversity resulting from living modified organisms. The new treaty will enter into force 90 days after being ratified by at least 40 Parties to the Cartagena Protocol on Biosafety.

Box 2 Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization: the Convention and the Nagoya Protocol

The principles governing access to genetic resources and the fair and equitable sharing of benefits

In its article 15, the Convention reaffirms national sovereignty over natural resources and establishes that the right to determine access to genetic resources rests with national governments and is subject to national legislation. At the same time, the Convention requires Parties to facilitate access and "not to impose restrictions that run counter to the objectives of the Convention." Article 15 introduces two key concepts: access shall be on "mutually agreed terms" (MAT) and subject to "prior informed consent" (PIC).

For access to be on MAT, both supplier and recipient must agree to the terms, thus providing an opportunity for the providing country to negotiate a share of the benefits derived from the use of the genetic resources. Usually, this implies a contractual arrangement, executed on a bilateral basis. The contract often takes the form of a "Material Transfer Agreement" (MTA) setting out the agreed terms on which the genetic material is transferred. The MTA may specify the permitted or prohibited uses of the genetic resources provided, including whether or not it may be commercialized; any rights which may or may not be taken out over the resource or its derivatives; and the benefits that are to be shared.

Unless the Party determines otherwise, access is also subject to PIC. This means that the responsible authority of the providing country can decide to grant or refuse access following a request from the applicant. In the application, the applicant may be required to provide information concerning the genetic resources required and the purpose for which they are required, as well as any proposal for benefit sharing. As part of the PIC procedure, the responsible authority may consult with indigenous and local communities, or other stakeholders, concerned.

The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity

The Nagoya Protocol on Access and Benefit Sharing is a supplementary agreement to the CBD. It provides a transparent legal framework for the effective implementation of one of the three objectives of the CBD: the fair and equitable sharing of benefits arising out of the utilization of genetic resources.

The Nagoya Protocol was adopted on 29 October 2010 in Nagoya, Japan, and will enter into force 90 days after the 50th instrument of ratification is received. Its objective is the fair and equitable sharing of benefits arising from the utilization of genetic resources, thereby contributing to the conservation and sustainable use of biodiversity. The Nagoya Protocol will create greater legal certainty and transparency for both providers and users of genetic resources by establishing more predictable conditions for access to genetic resources and by helping to ensure benefit-sharing when genetic resources leave the contracting party providing the genetic resources. The Nagoya Protocol applies to genetic resources that are covered by the CBD, and to the benefits arising from their utilization. The Nagoya Protocol also covers traditional knowledge associated with genetic resources that are covered by the CBD and the benefits arising from its utilization.

By helping to ensure benefit-sharing, the Nagoya Protocol creates incentives to conserve and sustainably use genetic resources, and therefore enhances the contribution of biodiversity to development and human well-being. The Protocol sets out core obligations related to access, benefit sharing, and compliance for its contracting Parties.

include providing assessments of the status of biological diversity, providing assessments of the types of measures taken in accordance with the provisions of the Convention, and responding to questions that the COP may put to the body. By mid 2012, SBSTTA had met 16 times.

The COP is further supported by the Ad Hoc Open-ended Working Group on Review of Implementation of the Convention. The Working Group is another intergovernmental body and was established to consider progress in the implementation of the Convention and to review the impacts and effectiveness of existing processes under the Convention. By mid 2012, the Working Group on the Review of Implementation had met four times. Another Ad Hoc Open-ended Working Group addresses Article 8(j) of the Convention and its related provisions which call for the respect, preservation and maintenance of the knowledge, innovations and practices of indigenous and local communities relevant to the conservation and sustainable use of biodiversity and serves to help reflect these in the Convention.

The Convention's broad scope makes translating its provisions into policy and practice extremely challenging. In the first 10 years following its entry into force, the Conference of the Parties has developed a comprehensive body of guidance relating to the conservation and sustainable use of biodiversity and the equitable sharing of the benefits from the use of genetic resources. The new Strategic Plan for Biodiversity 2011–2012 and the Aichi Biodiversity Targets (*see* The Strategic Plan for Biodiversity 2011–2020 and the Aichi Biodiversity Targets) provide an agreed framework for implementation of the Convention through this decade. It superseded the earlier strategic plan and the "2010 Biodiversity Target". The major challenge now is to ensure implementation of the Convention and the decisions of the Conference of the Parties in order to achieve the Aichi Targets. This requires increased commitment from Parties, enhanced technical and scientific cooperation and the mobilization of additional financial resources. To ensure progress, the Conference of the Parties will need to review implementation more effectively than in the past.

The Convention on Biological Diversity is supported by a Secretariat based in Montreal, Canada since 1996. As a neutral organization staffed by international civil servants, the Secretariat is accountable to the COP and its subsidiary. Its head, the Executive Secretary, is appointed by the Secretary-General of the United Nations in consultation with the Parties to the Convention.

The Secretariat assists and provides administrative support to the COP and other Convention bodies. It represents the day-to-day focal point for the Convention, organizes all meetings under the Convention, prepares background documentation for those meeting and facilitates the flow of authoritative information on the implementation of the Convention. The Secretariat plays a significant role in coordinating the work carried out under the Convention with that of other relevant institutions and conventions, and represents the Convention at meetings of relevant bodies. The Secretariat also facilitates support to Parties for the implementation of the Convention, for example, by providing technical and practical guidelines and by organizing capacity-building activities.

The Global Biodiversity Outlook: Assessing Progress and Providing the Basis for Future Action

In 2002, the Conference of the Parties adopted a Strategic Plan. Specifically Parties committed themselves "to a more effective and coherent implementation of the three objectives of the Convention, to achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level as a contribution to poverty alleviation and to the benefit of all life on earth". This target came to be known as the 2010 Biodiversity Target and was subsequently endorsed by heads of state and government at the World Summit on Sustainable Development in Johannesburg, 2002, and by the United Nations General Assembly. It was also incorporated as a new target under one of the Millennium Development Goals – Ensure Environmental Sustainability. The 2010 biodiversity target therefore represented a commitment from all governments; including those not party to the CBD.

The third edition of the Global Biodiversity Outlook (GBO-3), published in 2010, assessed progress toward the 2010 Biodiversity Target. In its final assessment it found that the 2010 biodiversity target had not been met despite an increase in conservation efforts. It concluded that globally the state of biodiversity continued to decline, according to most indicators, largely because the pressures on biodiversity continue to increase. It also found that there was neither indication of a significant reduction in the rate of decline in biodiversity, nor of a significant reduction in pressures on it (see **Box 3**). However, it also observed that there were several indications that responses to biodiversity loss were increasing and improving, although not yet on a scale sufficient to affect the overall negative trends in the state of biodiversity or the pressures on it.

The continued decline of biodiversity risks pushing the Earth's ecosystems beyond certain thresholds or tipping points (see **Box 4**). A tipping point refers to a situation in which an ecosystem experiences a shift to a new state, with significant changes to biodiversity and the services to people it underpins, at a regional or global scale. Tipping points are a major concern for scientists, managers, and policy–makers, because of their potentially large impacts on biodiversity, ecosystem services, and human well-being. It can be extremely difficult for societies to adapt to rapid and potentially irreversible shifts in the functioning and character of an ecosystem on which they depend (Leadley *et al.*, 2010; Pereira *et al.*, 2011). While it is highly likely that tipping points will occur in the future if effective action to protect biodiversity is not taken, the dynamics in most cases cannot yet be predicted with precision. Accordingly, to avoid tipping points a precautionary approach is warranted.

The Global Biodiversity Outlook arrived at the following conclusions. These guided the development of the Convention's new strategic plan:

- There are greater opportunities than previously recognized to address the biodiversity crisis while contributing to other social objectives.
- Well-targeted policies focusing on critical areas, species, and ecosystem services are essential to avoid the most

Box 3 Biodiversity Indicators

There is no single measurement that captures the current status or trends in global biodiversity. Owing to its complexity a range of indicators has been developed by the CBD to provide scientifically rigorous assessments of trends in the state of the various components of biodiversity. The Convention adopted a set of 15 headline indicators which relate to the different components of biodiversity (genes, populations, species, and ecosystems) as well as the pressures being imposed on it and the responses being adopted to address biodiversity loss.

The GBO-3 found that 10 of the 15 headline indicators show trends unfavourable for biodiversity. Overall the state of biodiversity shows negative trends, with no significant reduction in the rate of decline. There is no evidence of a slowing in the increase of pressures on biodiversity, based on the trend shown by indicators of humanity's ecological footprint, nitrogen deposition, alien species introductions, overexploited fish stocks and the impact of climate change on biodiversity. In contrast, all indicators of the responses to address biodiversity loss are moving in a positive direction. More areas are being protected for biodiversity, more policies and laws are being introduced to avoid damage from invasive alien species, and more money is being spent in support of the CBD and its objectives. The overall message from these indicators is that despite the many efforts taken around the world to conserve biodiversity and use it sustainably, responses so far have not been adequate to address the scale of biodiversity loss or reduce the pressures (**Figure 1**).

Box 4 Examples of Tipping Points

There is a high risk of dramatic biodiversity loss and accompanying degradation of a broad range of ecosystem services if ecosystems are pushed beyond certain thresholds or tipping points. The poor would face the earliest and most severe impacts of such changes, but ultimately all societies and communities would suffer. Examples include:

The Amazon forest, due to the interaction of deforestation, fire and climate change, could undergo a widespread dieback, with parts of the forest moving into a self-perpetuating cycle of more frequent fires and intense droughts leading to a shift to savanna-like vegetation. While there are large uncertainties associated with these scenarios, it is known that such dieback becomes much more likely to occur if deforestation exceeds 20–30% (it is currently above 17% in the Brazilian Amazon). It would lead to regional rainfall reductions, compromising agricultural production. There would also be global impacts through increased carbon emissions, and massive loss of biodiversity.

The build-up of phosphates and nitrates from agricultural fertilizers and sewage effluent can shift freshwater lakes and other inland water ecosystems into a long-term, algae-dominated (eutrophic) state. This could lead to declining fish availability with implications for food security in many developing countries. There will also be loss of recreation opportunities and tourism income, and in some cases health risks for people and livestock from toxic algal blooms. Similar, nitrogen-induced eutrophication phenomena in coastal environments lead to more oxygen-starved dead zones, with major economic losses resulting from reduced productivity of fisheries and decreased tourism revenues.

The combined impacts of ocean acidification, warmer sea temperatures, and other human-induced stresses make tropical coral reef ecosystems vulnerable to collapse. More acidic water – brought about by higher carbon dioxide concentrations in the atmosphere – decreases the availability of the carbonate ions required to build coral skeletons. Together with the bleaching impact of warmer water, elevated nutrient levels from pollution, overfishing, sediment deposition arising from inland deforestation, and other pressures, reefs worldwide increasingly become algae-dominated with catastrophic loss of biodiversity and ecosystem functioning, threatening the livelihoods and food security of hundreds of millions of people.

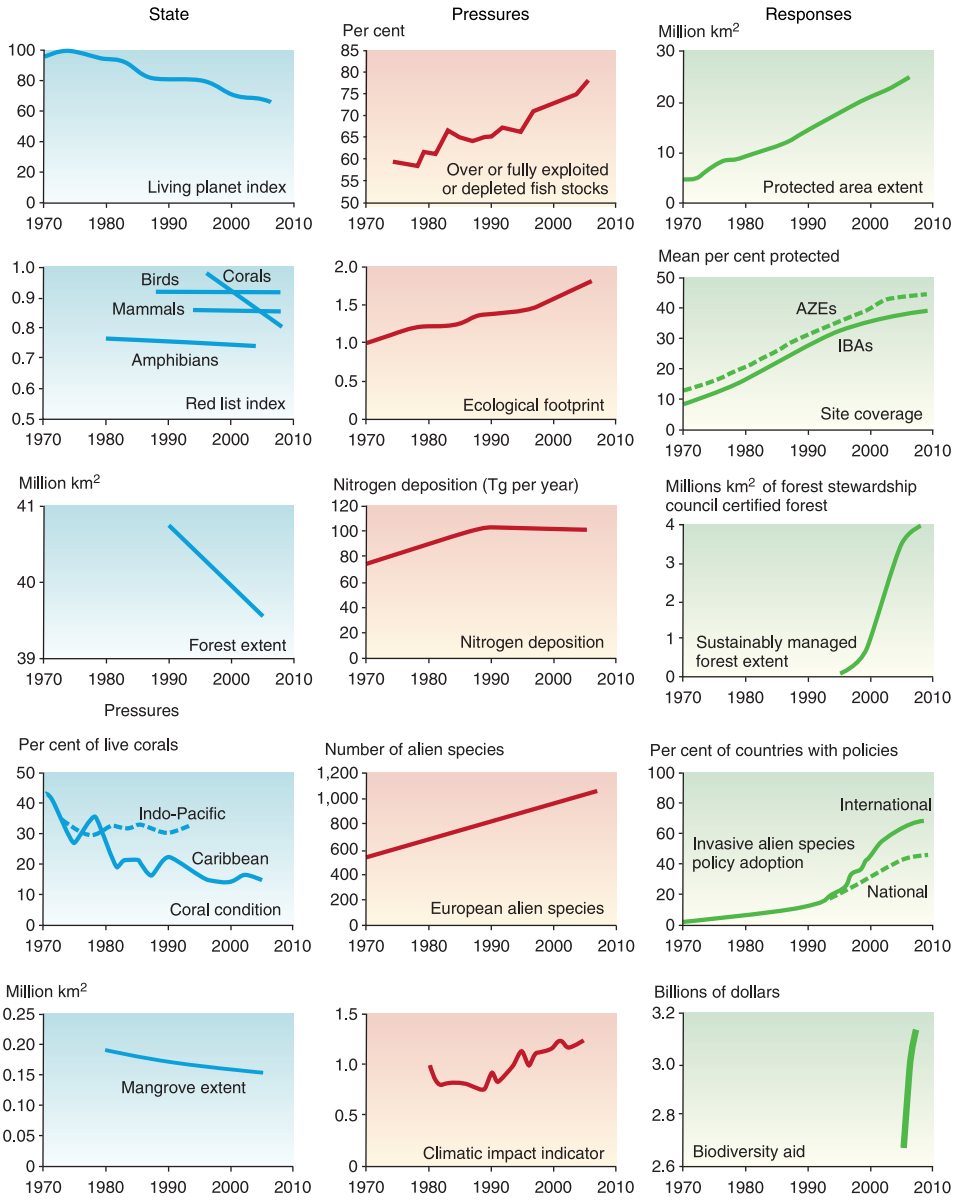


Figure 1 These graphs illustrate that the state of biodiversity is declining and the pressures on it are increasing but that our responses to its loss are also increasing. Reproduced from Secretariat of the Convention on Biological Diversity (2010a) *Action for Biodiversity: Towards a Society in Harmony with Nature*. Secretariat of the Convention on Biological Diversity; Secretariat of the Convention on Biological Diversity (2010b) *Global Biodiversity Outlook 3*, pp. 8–13. Secretariat of the Convention on Biological Diversity; Based on Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328(5982): 1164–1168.

dangerous impacts on people and societies. Preventing further human-induced biodiversity loss for the near-term future will be extremely challenging, but biodiversity loss may be halted and in some aspects reversed in the longer term, if urgent, concerted and effective action is initiated now in support of an agreed long-term vision. While some actions may entail moderate costs or tradeoffs, the gains for biodiversity can be large in comparison.

- Such action to conserve biodiversity and use its components sustainably will reap rich rewards – through better health, greater food security, less poverty and a greater capacity to cope with, and adapt to, environmental change. In fact, placing greater priority on biodiversity is central to the success of development and poverty-alleviation measures.
- The linked challenges of biodiversity loss and climate change must be addressed by policymakers with equal priority and in close coordination, if the most severe impacts of each are to be avoided.
- Better protection of biodiversity should be seen as a prudent and cost-effective investment in risk-avoidance for the global community.
- Scientific uncertainty surrounding the precise connections between biodiversity and human well-being, and the

functioning of ecosystems, should not be used as an excuse for inaction.

- Effective action to address biodiversity loss depends on addressing the underlying causes or indirect drivers of that decline. This will mean: much greater efficiency in the use of land, energy, fresh water and materials to meet growing demand; use of market incentives, and avoidance of perverse subsidies to minimize unsustainable resource use and wasteful consumption; strategic planning in the use of land, inland waters and marine resources to reconcile development with conservation of biodiversity and the maintenance of multiple ecosystem services. While some actions may entail moderate costs or tradeoffs, the gains for biodiversity can be large in comparison; and ensuring that the benefits arising from use of and access to genetic resources and associated traditional knowledge, for example, through the development of drugs and cosmetics, are equitably shared with the countries and cultures from which they are obtained. Enhanced efforts are also needed in the areas of communication, education, and awareness-raising to ensure that as far as possible, everyone understands the value of biodiversity and what steps they can take to protect it, including through changes in personal consumption and behaviour.
- Urgent action is needed to reduce the direct drivers of biodiversity loss. The application of best practices in agriculture, sustainable forest management and sustainable fisheries should become standard practice, and approaches aimed at optimizing multiple ecosystem services instead of maximizing a single one should be promoted. In many cases, multiple drivers are combining to cause biodiversity loss and degradation of ecosystems. Sometimes, it may be more effective to concentrate urgent action on reducing those drivers most responsive to policy changes. This will reduce the pressures on biodiversity and protect its value for human societies in the short to medium-term, while the more intractable drivers are addressed over a longer time scale. For example, the resilience of coral reefs – and their ability to withstand and adapt to coral bleaching and ocean acidification – can be enhanced by reducing overfishing, land-based pollution, and physical damage.
- Increasingly, restoration of terrestrial, inland water and marine ecosystems will be needed to reestablish ecosystem functioning and the provision of valuable services.
- The action taken over the next decade or two, and the direction charted under the CBD, will determine whether the relatively stable environmental conditions on which human civilization has depended for the past 10,000 years will continue beyond this century. If we fail to use this opportunity, many ecosystems on the planet will move into new, unprecedented states in which the capacity to provide for the needs of present and future generations is highly uncertain.

The Strategic Plan for Biodiversity 2011–2020 and the Aichi Biodiversity Targets

At the 10th meeting of the Parties to the CBD, in 2010 in Nagoya, Japan, the Parties considered the conclusions of the GBO-3 and adopted a new Strategic Plan with the purpose of inspiring broad-based action in support of biodiversity over the next decade by all countries and stakeholders (**Box 5**). The Strategic Plan for Biodiversity 2011–2020 is an overarching framework on biodiversity, not only for the biodiversity-related conventions, but also for the entire United Nations system. The Strategic Plan is based on more than 2 years of consultation with Governments and stakeholders, experience in implementation the Convention, as well the conclusions of the GBO-3.

The Strategic Plan is comprised of a shared vision, a mission, and five strategic goals under which 20 ambitious yet achievable targets, known as the Aichi Biodiversity Targets, are organized. The goals and targets comprise both aspirations for achievement at the global level, and a flexible framework for the establishment of national or regional targets. In adopting the Plan Parties committed themselves to setting their own targets within this flexible framework, taking into account national needs and priorities, while also bearing in mind national contributions to the achievement of the global targets.

The Strategic Plan for Biodiversity 2011–2020 will be implemented primarily through activities at the national or subnational level, with supporting action at the regional and global levels. The Programmes of Work on various biomes and crosscutting issues that have been developed under the Convention, and formalized by decisions of the Conference of the Parties, provide detailed guidance on the types of actions and considerations which should be taken to implementation of the Strategic Plan. For the implementation of the Strategic Plan there is a need to broaden political support for the Convention and biodiversity more generally, and to integrate the biodiversity agenda into the plans and policies of other sectors across government, the economy and society.

Progress toward the implementation of the Strategic Plan will be reported by Parties through national reports that they provide to the CBD on a regular basis. Specifically Parties have been asked to provide information on any national targets, commitments, and/or policy instruments they adopt to implement the Strategic Plan, as well as any milestones toward these targets. This information will be considered by the Conference of the Parties at its 11th and subsequent meetings.

National Implementation of the Convention

While the CBD is an international environmental agreement it is primarily implemented through national level action. National biodiversity strategies and action plans (NBSAPs) are the principal instruments for implementing the Convention at the national level. NBSAPs reflect how a country intends to fulfill the objectives of the Convention in light of specific national circumstances, and the steps it plans on taking. The Convention requires countries to prepare a national biodiversity strategy (or equivalent instrument) and to ensure that this strategy is mainstreamed into the planning and activities of all those sectors whose activities can have an impact (positive and negative) on biodiversity.

By the end of 2011, 175 countries and the European Union (over 90% of the total Parties to the Convention) had developed their NBSAPs or equivalent instruments and were implementing them. Some 40 Parties have revised their

Box 5 The Strategic Plan for Biodiversity 2011–2020 and the Aichi Biodiversity Targets

Vision

The vision for the new plan is: "Living in Harmony with Nature" where "By 2050, biodiversity is valued, conserved, restored and wisely used, maintaining ecosystem services, sustaining a healthy planet and delivering benefits essential for all people."

Mission

The mission of the new plan is to "take effective and urgent action to halt the loss of biodiversity in order to ensure that by 2020 ecosystems are resilient and continue to provide essential services, thereby securing the planet's variety of life, and contributing to human well-being, and poverty eradication. To ensure this, pressures on biodiversity are reduced, ecosystems are restored, biological resources are sustainably used and benefits arising out of utilization of genetic resources are shared in a fair and equitable manner; adequate financial resources are provided, capacities are enhanced, biodiversity issues and values mainstreamed, appropriate policies are effectively implemented, and decision-making is based on sound science and the precautionary approach."

Aichi Biodiversity Targets (*The text of the targets in this box has been abridged. For the full official text, please refer to www.cbd.int/sp*)

Strategic Goal A: Address the Underlying Causes of Biodiversity Loss

Target 1 – People are aware of the values of biodiversity and the steps they can take to conserve and use it sustainably.
Target 2 – Biodiversity values have been integrated into national and local development and poverty reduction strategies and planning processes.
Target 3 – Incentives, including subsidies, harmful to biodiversity are eliminated, phased out or reformed and positive incentives are developed and applied.
Target 4 – Governments, business and stakeholders at all levels have taken steps to achieve or have implemented plans for sustainable production and consumption.

Strategic Goal B: Reduce the Direct Pressures on Biodiversity and Promote Sustainable Use

Target 5 – The rate of loss of all natural habitats is at least halved and where feasible brought close to zero.
Target 6 – Overfishing is avoided and fisheries have no significant adverse impacts on threatened species and vulnerable ecosystems.
Target 7 – Areas under agriculture, aquaculture, and forestry are managed sustainably.
Target 8 – Pollution, including from excess nutrients, has been brought to levels that are not detrimental to ecosystem function and biodiversity.
Target 9 – Invasive alien species and pathways are identified, priority species are controlled or eradicated, and measures are in place to manage pathways.
Target 10 – The multiple anthropogenic pressures on vulnerable ecosystems impacted by climate change or ocean acidification are minimized.

Strategic Goal C: To improve the Status of Biodiversity by Safeguarding Ecosystems, Species and Genetic Diversity

Target 11 – At least 17% of terrestrial and inland water, and 10% of coastal and marine areas are conserved through effective, ecologically representative and well connected systems of protected areas.
Target 12 – The extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained.
Target 13 – The genetic diversity of cultivated plants and farmed and domesticated animals and of wild relatives is maintained.

Strategic Goal D: Enhance the Benefits to All from Biodiversity and Ecosystem Services

Target 14 – Ecosystems that provide essential services are restored and safeguarded.
Target 15 – Ecosystem resilience and the contribution of biodiversity to carbon stocks has been enhanced, through conservation and restoration, including restoration of at least 15% of degraded ecosystems.
Target 16 – The Nagoya Protocol on Access and Benefit Sharing is in force and operational.

Strategic Goal E: Enhance Implementation through Participatory Planning, Knowledge Management and Capacity Building

Target 17 – Each Party has developed, adopted as a policy instrument, and has commenced implementing an effective, participatory and updated national biodiversity strategy and action plan.
Target 18 – Traditional knowledge, innovations and practices of indigenous and local communities, and their customary use of biological resources, are respected.
Target 19 – Knowledge relating to biodiversity is improved, shared and transferred, and applied.
Target 20 – The mobilization of financial resources for implementing the Strategic Plan for Biodiversity 2011–2020 increased substantially from the current levels.

NBSAPs, and most others are in the process of doing so, in the light of the new Strategic Plan for Biodiversity 2011–2020. Among other things, revisions allow Parties to identify and meet new challenges and to respond to recent guidance from the Conference of the Parties. Moreover, biodiversity planning is, of necessity, an ongoing process and updating national strategies allows the main actors to learn from experience and improve their polices and activities to implement the Convention.

While a few countries (e.g., Costa Rica, Samoa, Saint Lucia) have implemented most of the activities proposed in their national strategies, the level of implementation of the "first generation" of national strategies was estimated to be in the range of only 30–40%. The main obstacles to implementation reported by countries include: limited financial, technical and human resources and capacities, limited information, low political will, lack of coordination between ministries, poverty, low awareness level of biodiversity issues, and limited incentives for biodiversity conservation and sustainable use. Nonetheless, new legislation has been developed, institutions strengthened and many activities for biodiversity conservation and use carried out (Secretariat of the Convention on Biological Diversity, 2010a). For example, specific biodiversity laws and policies in Brazil, Costa Rica, Japan, Norway, and South Africa have established the institutional framework for coordinated implementation of the Convention in those countries.

NBSAPs have catalysed action in a broader range of issues such as: invasive alien species (beyond those already addressed through plant protection services in the agricultural and forest sectors); sustainable use; incentive measures; protection of traditional knowledge; access and benefit-sharing; biosafety; and agricultural biodiversity. There appears to have been an evolution in the scope and strategic focus of national biodiversity strategies and action plans; recently developed and updated ones tend to be more strategic than the first generation of national biodiversity strategies and action plans and have a stronger emphasis on the mainstreaming of biodiversity into other sectors and cross-sectoral plans and policies (Prip and Gross, 2010).

Mainstreaming of biodiversity needs to take place at various levels: integration into cross-sectoral policies and strategies (finance, national development, poverty eradication); integration of biodiversity into economic sectors (including through different government ministries); and integration into spatial planning, at all levels of government, especially at provincial/state and municipal levels.

To be effective, NBSAPs must reflect broader national development and environment objectives. For example, the Namibian NBSAP is positioned as a contribution to National Development and Vision 2030, and Rwanda has integrated biodiversity issues into its Economic Development and Poverty Reduction Strategy. A number of recently developed or updated NBSAPs are closely linked with the cycle of national planning processes such as 5-year plans (e.g., Indonesia, Malaysia, Thailand), poverty reduction plans (Cambodia, Madagascar, Vietnam), the framework for achievement of the Millennium Development Goals (Cambodia), and development plans (Namibia, Philippines). In Indonesia, the planning authority led the development of the NBSAP which facilitated the later incorporation of the NBSAP into the Medium-Term Development Plan.

Many countries report the integration of biodiversity into the tourism, forestry and agricultural sectors. The integration of biodiversity into other sectors is less common. For example, France has developed sectoral action plans to implement the NBSAP through various ministries. Rwanda reports that it has successfully mainstreamed biodiversity in other sectors besides the environment, such as agriculture, education, health, rural development, forestry, mining, tourism, finance, trade and industry. Most countries are promoting biodiversity activities related to climate change adaptation, and an increasing number of them are taking actions related to climate change mitigation. For example, Samoa's National climate change adaptation program of action recognizes the importance of coral reefs and mangroves in the protection of coastal assets and the plan provides for the establishment of marine and terrestrial conservation areas.

Nearly all countries have laws and mechanisms in place for environmental impact assessment and about one-third of them have policies related to strategic environmental impact assessment. Both figures represent an increase from the situation 5 years previously. Many countries are integrating biodiversity into mechanisms for spatial planning. Brazil has promoted Ecological-Economical Zoning processes at federal, state, and municipal levels as well as for some hydrological basins. South Africa has carried out a national spatial assessment of biodiversity and is integrating biodiversity into spatial planning and economic development at the municipal level in North West and Western Cape Provinces.

The use of economic instruments is seen as an important – but under-used – approach to mainstream biodiversity. Some countries such as Mexico are integrating biodiversity-related issues into national accounts. South Africa has adopted amendments to the tax legislation, effective from 2009, under which specified conservation costs are tax deductible. For example, expenses incurred as a result of controlling alien and invasive vegetation are to be allowed as a deduction for farming purposes. Similar tax deductions are being applied to conservation and maintenance expenses for landowners with contracts to keep part of their property as nature reserves. Costa Rica has introduced a fuel tax that is used to support the costs of managing protected areas. Ecuador provides cash incentives to support sustainable forest management by indigenous and local communities and landowners through its Socio-Bosque Programme. In both Ecuador and Costa Rica, a levy on city water supply helps to protect high-biodiversity water catchment areas.

Brazil provides an example of effective inter-ministerial action to protect biodiversity in its Action plan to protect and control deforestation in the Brazilian Amazon. The plan includes the establishment of protected areas and indigenous reserves, land planning using ecological and economic zoning, use of incentive schemes, monitoring by satellite and aircraft, and strengthened enforcement measures. The plan involves action by all relevant ministries including environment, agriculture, science and technology, justice and defence coordinated through an inter-ministerial working group under the authority of the President of the Federal Republic. Deforestation rates were reduced by some 80% in the period 2002–2012.

These examples demonstrate that success depends on a coherent approach across several ministries and levels of government, sectors of society and business.

Some national biodiversity strategies and action plans, especially the more recently developed or updated ones, encourage biodiversity planning at the subnational level, that is at the level of states or provinces and at the local, district, or municipal levels. This is relevant because many of the decisions affecting biodiversity, such as land-use planning decisions, are taken at local or other subnational levels. Moreover, in some countries, biodiversity planning at the subnational level is being promoted as part of broader programmes of decentralization or increased regional autonomy (e.g., China, Indonesia, Maldives, Myanmar, Pakistan). Some federal countries have promoted the development of state or provincial biodiversity strategies and action plans (e.g., India, which has 71 strategies and action plans including state strategies, local strategies ,and eco-regional strategies; Mexico, where state biodiversity strategies have been adopted by Michoacán and Morales and are in preparation in another seven states). In addition, Peru has developed 17 "regional biodiversity strategies" as biodiversity planning instruments. Japan is promoting biodiversity strategies for its prefectures. The United Kingdom has a large number of local biodiversity action plans.

Many regions or subregions have developed regional biodiversity strategies or action plans: these include: the European Union Biodiversity Action Plan; the Central American Biodiversity Strategy (Central American Commission for Environment and Development (CCAD)); the Cooperative Strategy for the Conservation of Biological Diversity in the Arctic region (Conservation of Arctic Flora and Fauna); the Southern Africa Biodiversity Strategy (Southern Africa Development Community); the Andean Biodiversity Strategy (Andean Community); and the Regional Action Plan for Amazonian Biodiversity (Amazon Treaty Cooperation Organization). In addition, a number of regional mechanisms such as the Association of South-East Asian Nations and its Centre for Biodiversity, the European Union the CCAD, Commission of Forestry in Central Africa, Secretariat of the Pacific Regional Environment Programme, the Committee of Arab Ministers Responsible for the Environment of the League of Arab States, among others, play an important role in supporting implementation of the Convention.

With the adoption of the Strategic Plan for Biodiversity 2011–2020, most countries are updating their national biodiversity strategies and action plans, incorporating national targets in the framework of the Aichi Biodiversity Targets. For example, the new biodiversity strategy for England sets a target to establish and effectively manage an ecologically coherent network of marine protected areas which covers in excess of 25% of English waters by the end of 2016 (relating to Aichi Targets 6 and 11). Norway's new biodiversity strategy includes a comprehensive set of targets for each of Norway's six major biomes. Guatemala's new strategy includes targets for reducing deforestation and for forest restoration. Colombia has developed a new biodiversity policy that focuses on maintaining and increasing ecosystem resilience. Brazil developed a comprehensive set of biodiversity targets in the framework of the 2010 global biodiversity target. This included a target to reduce the rate of deforestation in the Amazon biome by at least 75%, a target that was met, as noted above. Following a series of dialogues among the various ministries of government and sectors of society, Brazil is developing an updated set of targets.

The CBD Regime: Supporting Action at the International Level

Advancing New Concepts

At the time of its inception, the CBD marked a departure from earlier international environmental law instruments by supporting a balance between conservation and sustainable use rather than promoting conservation only. It introduced novel legal concepts such as biodiversity, ecosystems, genetic resources and biotechnology, benefit sharing, and traditional knowledge. It provided an innovative and flexible framework for accommodating developed and developing countries' concerns and capacities and for encouraging partnerships between national and local authorities, local and indigenous communities, and the private sector.

Over time, the Convention, through the SBSTTA and the Conference of the Parties has helped to translate a number of scientific concepts into politically recognized guidance. An example is the ecosystem approach. The ecosystem approach (**Box 6**) is a strategy for the integrated management of land,

Box 6 The 12 principles of the Ecosystem Approach

1. The objectives of management of land, water, and living resources are a matter of societal choices.
2. Management should be decentralized to the lowest appropriate level.
3. Ecosystem managers should consider the effects (actual or potential) of their activities on adjacent and other ecosystems.
4. Recognizing potential gains from management, there is usually a need to understand and manage the ecosystem in an economic context. Any such ecosystem-management program should:
   a. Reduce those market distortions that adversely affect biological diversity;
   b. Align incentives to promote biodiversity conservation and sustainable use;
   c. Internalize costs and benefits in the given ecosystem to the extent feasible.
5. Conservation of ecosystem structure and functioning, in order to maintain ecosystem services, should be a priority target of the ecosystem approach.
6. Ecosystems must be managed within the limits of their functioning.
7. The ecosystem approach should be undertaken at the appropriate spatial and temporal scales.
8. Recognizing the varying temporal scales and lag-effects that characterize ecosystem processes, objectives for ecosystem management should be set for the long term.
9. Management must recognize the change is inevitable.
10. The ecosystem approach should seek the appropriate balance between, and integration of, conservation and use of biological diversity.
11. The ecosystem approach should consider all forms of relevant information, including scientific and indigenous and local knowledge, innovations and practices.
12. The ecosystem approach should involve all relevant sectors of society and scientific disciplines.

water, and living resources that promotes conservation and sustainable use in an equitable way. Application of the ecosystem approach helps to reach a balance of the three objectives of the Convention. It is based on the application of appropriate scientific methodologies focused on levels of biological organization which encompass the essential processes, functions, and interactions among organisms and their environment. It recognizes that humans, with their cultural diversity, are an integral component of ecosystems. The ecosystem approach requires adaptive management to deal with the complex and dynamic nature of ecosystems and the absence of complete knowledge or understanding of their functioning. As described by the Conference of the Parties, the ecosystem approach is the primary framework for action under the Convention. The Conference of the Parties, at its Fifth Meeting, endorsed the description of the ecosystem approach and operational guidance.

The Convention has also highlighted the linkages between biodiversity, ecosystem functioning, ecosystem services and human well-being, building on the work of the 2005 Millennium Ecosystem Assessment, and more recently on The Economics of Ecosystems and Biodiversity (TEEB, 2010). Overall, these scientific and conceptual developments have helped to operationalize the ecosystem approach, albeit in an incomplete manner. They have also shed new light on the CBD's preambular language on biodiversity's importance for meeting the food, health, and other needs of the world's growing population and on biodiversity's social, economic, scientific, cultural, and other values. Conceptualising ecosystem services therefore highlighted the links between biodiversity and human development and, to that extent, "modernised" the concept of sustainable use (Morgera and Tsioumani, 2011).

Developing Guidance and Programs of Work

The CBD programs of work are the main instruments that CBD parties use to achieve the commitments contained in the Convention and the Strategic Plan for Biodiversity 2011–2020. They include guidelines for national implementation, often recommending reforms of national laws, policies, or administrative practices. The COP has established seven thematic work programs, namely on agricultural biodiversity, dry and subhumid lands biodiversity, forest biodiversity, inland waters biodiversity, island biodiversity, marine and coastal biodiversity, and mountain biodiversity; and five crosscutting work programs on incentive measures, the Global Taxonomy Initiative, protected areas, Article 8(j) (traditional knowledge), and technology transfer and cooperation.

Work has also been undertaken on a series of other crosscutting issues, including climate change and biodiversity, the ecosystem approach, and sustainable use of biodiversity. Some cross cutting initiatives directly support work under thematic programs, for example, the work on indicators provides information on the status and trends of biodiversity for all biomes. Others develop discrete products quite separate from the thematic programmes. The work done for these crosscutting issues has led to a number of principles, guidelines, and other tools to facilitate the implementation of the Convention (see **Box 7**).

In addition to the tools and guidance, the Conference of the Parties regularly establishes initiatives and undertakes assessment on developing biodiversity issues. For example, the Convention has ongoing work to examine the impacts on biodiversity of ocean acidification, biofuel production and use, and geoengineering. The findings or outputs from these processes are published as technical reports and provide a basis for policy guidance to Parties to avoid, reduce, or mitigate the impacts.

The work of the CBD on indigenous and local communities is founded on Article 8(j), which calls on parties to encourage the equitable sharing of the benefits arising from the utilization of the knowledge, innovations, and practices of indigenous and local communities embodying traditional lifestyles relevant for the conservation and sustainable use of biological diversity. The Convention's programme of work on Article 8(j) and related provisions has delivered not only guidelines but has also advanced the interests of indigenous people's in the Nagoya Protocol. In fact, the CBD has emerged as the preferred international forum for indigenous and local communities to express their interests and demands for the protection of their traditional knowledge, particularly within the CBD Working Group on Article 8(j), which affords community representatives broad participation rights (Morgera and Tsioumani, 2011).

Promoting Cooperation and Coherence Among International Instruments

The CBD was developed after a number of other biodiversity related agreements, namely the Convention on International Trade in Endangered Species of Wild Fauna and Flora, the Convention on the Conservation of Migratory Species of Wild Animals, the Ramsar Convention on Wetlands, and the World Heritage Convention. The CBD builds on these preexisting species-based and area-based international environmental agreements while establishing a wider context in which such agreements are interpreted and implemented. While the CBD is not formally an "umbrella" convention, because it does not supersede previous agreements, it has the normative character of an umbrella convention in articulating new norms that could also apply to preexisting agreements.

The CBD promotes close collaboration with and among these other international conventions that focus on biodiversity issues as well as the International Treaty on Plant Genetic Resources for Food and Agriculture. The governing bodies of each convention have set out specific mandates for cooperation among the biodiversity-related conventions, and a number of joint work programs have been established.

The Convention has also explored synergies with the other two Conventions born of the 1992 Rio Conference on Sustainable Development: the United Nations Framework Convention on Climate Change and the United Nations Convention to combat desertification. Building on assessments carried out by expert groups under the authority of SBSTTA, the Convention has developed policy guidance for ecosystem-based approaches to climate change mitigation and adaptation that can be used for linked implementation of national plans and strategies under the three Rio Conventions.

Box 7 Principles and Guidelines developed under the Convention

Guidelines	*Description*
Addis Ababa Principles and Guidelines for the Sustainable Use of Biodiversity	A framework for advising stakeholders on how they can ensure that their use of the components of biodiversity will not lead to long-term biodiversity declines, but will instead promote conservation and contribute to poverty alleviation. Applying to both consumptive and nonconsumptive uses of biodiversity, the Principles and Guidelines take into account issues related to policies, laws and regulations; management of biodiversity; socioeconomic conditions; and information, research and education.
Guiding Principles on Invasive Alien Species	The Guiding Principles are intended to assist governments to control invasive alien species, as an integral part of conservation and economic development. They comprise 15 principles on prevention, intentional and unintentional introduction, and mitigation of impacts.
Guidelines for Incorporating Biodiversity related Issues into Environmental Impact Assessment Legislation and/or Processes and in Strategic Environmental Assessment	Impact assessment is a comprehensive process and assessment tool that promotes sustainable development and is used to ensure that projects, programs and policies are economically viable, socially equitable, and environmentally sustainable. These guidelines provide advice on the incorporation of biodiversity-related concerns into new or existing environmental impact assessment (EIA) and strategic environmental assessment (SEA) procedures. Guidelines for EIA and SEA in marine areas, including the high seas, and areas beyond national jurisdiction are being developed under the Convention.
Akwé: Kon Voluntary Guidelines for the Conduct of Cultural, Environmental, and Social Impact Assessment regarding Developments on Lands and Waters Traditionally Occupied or Used by Indigenous and Local Communities	The guidelines provide advice on how to incorporate cultural, environmental (including biodiversity-related), and social considerations of indigenous and local communities into new or existing impact-assessment procedures, to ensure appropriate development. They support the full and effective participation of indigenous and local communities in screening, scoping and development planning exercises, taking into account their traditional knowledge, innovations and practices.
The Tkarihwaié:ri Code of Ethical Conduct to Ensure Respect for the Cultural and Intellectual Heritage of Indigenous and Local Communities	The guidelines contain ethical principles and propose methods for promoting respect for the cultural and intellectual heritage of indigenous and local communities relevant for the conservation and sustainable use of biological diversity.
Guidelines on Biodiversity and Tourism Development	A comprehensive instrument for managing tourism activities in an ecologically, economically and socially sustainable manner. The guidelines emphasize a consultative approach involving multiple stakeholders, and are structured around ten steps, from development of an overall vision to implementation of adaptive management programs.
Proposals for the Design and Implementation of Incentive Measures	Incentive measures serve to correct the failure of markets to properly reflect biodiversity's value to society. These Proposals identify and explain key elements that need to be considered when using incentive measures for the conservation and sustainable use of biodiversity. They also provide advice on the application of complementary measures for the provision of capacity-building, and for management, monitoring and enforcement.
Proposals for the Application of Ways and Means to Remove or Mitigate Perverse Incentives	Perverse incentives induce unsustainable behaviours that destroy biodiversity, often as unanticipated side effects of policies designed to attain other objectives. These Proposals offer a general framework for the removal or mitigation of perverse incentives, based on a three-phase approach: identification of policies and practices generating perverse incentives; design and implementation of appropriate reforms; and monitoring, enforcement and evaluation of these reforms.
Biodiversity and ecosystem-based approached to climate change mitigation and adaptation	The Guidance provides ways to conserve, sustainably use and restore biodiversity and ecosystem services while contributing to climate change mitigation and adaptation, including: reducing the impacts on biodiversity of climate change, and of measures to address climate change.

The Convention has also developed guidance and promoted safeguards to ensure that mechanisms for reducing emissions of greenhouse gases from deforestation and forest degradation also help to protect biodiversity and the rights of indigenous and local communities (Reducing emissions from deforestation and forest degradation (REDD+)).

In cooperation with regional fisheries management organizations, the regional seas conventions and other international organizations, the Convention on Biodiversity is undertaking a process to describe ecologically and biologically sensitive marine areas. The results will be used by countries as well as by the United Nations General Assembly in their work

to establish marine protected areas or other means to protect these areas, including the high seas.

Actions to control and prevent the spread of invasive alien species, including pests of plants and animal diseases, are promoted through various national government departments which relate to various international agreements and organizations including the International Plant protection Convention, the World Organization for Animal Health and the International Maritime Organization as well as the World Trade Organization. A liaison group on invasive alien species brings together these and other organizations to promote cooperation in the development of international policy and standards and to support a coherent approach to implementation at the national level among border control officials, conservation officials, and others.

Strengthening Capacity and Providing Resources for Implementation

A number of processes have also been established as part of the CBD to help ensure that resources needed to carry out work in support of the Convention are available to countries. One of these is the Clearing House Mechanism (CHM). The term "clearing-house" originally referred to a financial establishment where checks and bills are exchanged among member banks so that only the net balances need to be settled in cash. Today, its meaning has been extended to include any agency that brings together seekers and providers of goods, services or information, thus matching demand with supply. In the case of the Convention the CHM serves as a means of sharing information among Parties and partners. In its next phase of work in order to contribute significantly to the implementation of the Strategic Plan for Biodiversity 2011–2020, additional effort will be required to promote technical and scientific cooperation including institutional cooperation and the exchange of expertise among Parties and partners. This will build on the successful work of national institutions such as Mexico's Commission for the Conservation and Use of Biodiversity, Costa Rica's Institute for Biodiversity, and the South African National Biodiversity Institute (SANBI).

The Convention provides for a financial mechanism, operated by the Global Environment Facility (GEF). To 2010, GEF has provided some $3 billion in grants for biodiversity conservation and sustainable use and leveraged an additional $8 billion in cofinancing for almost 1000 projects in 155 countries. For the period 2010–2014, the GEF is providing some $ 300 m annually for biodiversity. About $3 billion per year is provided in biodiversity-related bilateral aid representing less than 3% of overseas aid or development assistance. The largest sources of funding for biodiversity in developed and developing countries are domestic budgets; precise figures are not available but domestic support to biodiversity related activities globally is estimated to be of the order of $100 billion per year. Virtually all countries report that financing for biodiversity is insufficient. To meet the ambitious Aichi Biodiversity targets, the Convention is developing its Strategy for Resource Mobilization. It is agreed that a substantial increase in financial resources from all sources will be required.

Challenges and Prospects for Achieving the Aichi Biodiversity Targets

The Strategic Plan for Biodiversity 2011–2020 with the Aichi Biodiversity Targets provides an ambitious plan for the international community to take urgent action to address biodiversity loss with a view to realizing a long term vision where, "by 2050, biodiversity is valued, conserved, restored and wisely used, maintaining ecosystem services, sustaining a healthy planet and delivering benefits essential for all people". Some of the key challenges in achieving the targets and realizing this Vision are reviewed in this section.

Increasing Pressure for Land

Land-use change remains the biggest driver of biodiversity loss, at least in terrestrial ecosystems. The recent food crisis has thrown into sharp relief the increasingly strong multiple pressures for land: for crops and livestock to feed an expanding population and for meeting the increased consumption by that part of the population that can afford it; for biofuels which are supposed to contribute to the effort to mitigate climate change (but, in fact, frequently increase net emissions of greenhouse gases); and for infrastructure development, especially in the narrow band of 100 km from coastlines where the majority of the population lives and works.

Yet despite these pressures, there are some encouraging signs. One results from yet another demand for land use – for carbon storage and sequestration. Mechanisms such as REDD+ (the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries) are being agreed and put in place to provide incentives for conserving forests and other ecosystems that contain high levels of carbon.

Perhaps more significant is the emerging consensus that biodiversity conservation and, specifically, reducing land conversion, is compatible with increased agricultural productivity. For example, a recent study looking at how agriculture can feed the world in a sustainable manner proposes four strategies that are consistent with the Strategic Plan for Biodiversity 2011–2020 (Foley *et al.*, 2011). These are: halting agricultural expansion; closing "yield gaps"; increasing the efficiency of use of water and nutrients; and shifting diets and reducing waste. Such strategies are consistent with Aichi Targets 5 (half deforestation and other loss of natural habits), 7 (promote sustainable agriculture), 8 (reduce nutrient pollution) and 3 (sustainable consumption), respectively. Similarly, a study on the Future of Food and Farming commissioned by the United Kingdom Chief Government Science Advisor (Foresight, 2011) concludes that "there will hardly ever be a case to convert forests, especially tropical rainforests, to food production." In fact, there is a range of technical and policy tools available to facilitate more rational and efficient land-use planning. Restoration of degraded land will also be increasingly be important – with so many demands for land use, leaving land in a degraded state will no longer be a viable option. A coalition of governments and international organizations is proposing to restore at least 150 million hectares of lost and degraded forests by 2020.

This is not to say that it will be easy to meet the multiple objectives of increasing food production, protecting biodiversity and sequestering carbon. The Foresight study calls for "nothing less … than a redesign of the whole food system to bring sustainability to the fore". Further Foley *et al.* acknowledge that "our analysis focuses on the agronomic and environmental aspects of these challenges, and leaves a richer discussion of associated social, economic, and cultural issues to future work." Among these, the strengthening of rights to land and natural resources of indigenous and local communities will be a high priority.

Inadequate Protection of the Marine Environment

The largest threat to marine and coastal biodiversity has been overfishing and destructive fishing practices. These have resulted literally in the decimation – or more – of fish stocks and the transformation of marine habitats to less productive states, with severe losses on livelihoods and food security. Eutrophication and dead zones from land-based pollution and sedimentation are additional pressures. But in the marine realm too there are signs of progress (Worm, *et al.*, 2009). With well thought-out action, only a 10% reduction in fishing effort may be needed to half the pressure on marine ecosystems. Ending subsidies for fishing and especially for fuel would be a major step to sustainability. As noted above, by identifying ecologically and biologically significant marine areas, the Convention is paving the way for an expansion of marine protected areas in coastal waters and the high seas. Marine protected areas would also help to maintain or increase the resilience of marine ecosystems to global warming and ocean acidification.

The Increasing Threat of Climate Change

Climate change imposes an increasing severe range of threats to biodiversity and ecosystem services, greatly increasing the risk of species extinctions and local losses. Urgent and large-scale action to reduce emissions from fossil fuels and land-use change is needed. Reducing this threat to biodiversity requires progress under the sister Convention – the United Nations Framework Convention on Climate Change (UNFCCC). The Cancun (2010) and Durban (2011) agreements under the UNFCCC are steps forward. However, as noted above (*see* Developing Guidance and Programs of Work), action under the CBD can contribute to climate change mitigation and adaptation. Relevant actions include minimizing other pressures on biodiversity (as articulated in Aichi Target 10), protecting genetic diversity to allow adaptation to future conditions (Aichi target 8), and restoring ecosystems, thereby contributing to both adaptation (though increased ecosystem resilience) and mitigation (though carbon sequestration) (Aichi Target 15).

The Unachieved Goal of a Sustainable Economy

Ultimately achieving the goals of the CBD as well as those of the Climate Convention, will require a genuine shift to sustainable production and consumption patterns. A full discussion of the measures needed in this regard is beyond the scope of this article. However changes in public perception and behaviour (see Aichi Target 1), inclusive wealth accounting (including integration of biodiversity into national accounts as called for in Aichi Target 2), and subsidy reform (Aichi Target 3) are key elements. Changes in consumption patterns are also key (Aichi Target 4). Some changes such as moderation in meat consumption would deliver real health benefits for people and the planet.

The Need for Concerted Action

A study commissioned by the Executive Director of the United Nations Environment Programme (Netherlands Environment Assessment Agency, 2010) for the international year of biodiversity showed that while it would not be possible, given other socioeconomic imperatives, to halt biodiversity loss in the next few decades, even less by 2020, an aggressive implementation of a combination of measures to expand protected areas, reduce deforestation, improve forest management, reduce fishing efforts, increase agricultural efficiency by closing the yield gap, reduce food waste, change diets and mitigate climate change could halve the loss of biodiversity (using mean species abundance as indicator) by 2030. With the addition of aggressive plans for ecosystem restoration, somewhat more ambitious goals may be justifiable (ten Brink, personal communication). Like the other studies referred to above (*see* Increasing Pressure for Land), this reinforces the general messages provided in Global Biodiversity Outlook Reports and broadly support the feasibility of the Strategic Plan for Biodiversity 2011–2020.

As acknowledged in a number of studies, this is a unique time in history – Humanity is now the major force altering the planetary environment and decisions made now and over the next few decades will disproportionately affect the future.

Appendix

List of Courses

1. Environmental Studies
2. Environmental Science
3. Conservation Biology
4. Global Change Studies
5. International Law

See also: Biodiversity, Definition of. Bioprospecting. Conservation and People. Conservation Efforts, Contemporary. Conservation Movement, Historical. Human Impact on Biodiversity, Overview. Indigenous Peoples and Biodiversity. International Organizations and Biodiversity. Loss of Biodiversity, Overview. Priority Setting for Biodiversity and Ecosystem Services

References

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328(5982): 1164–1168.

Foley JA, Ramankutty N, Brauman KA, *et al.* (2011) Solutions for a cultivated planet. *Nature* 478: 337–342.
Foresight (2011) *The Future of Food and Farming: Final Project Report*. London: Government Office for Science.
Leadley P, Pereira HM, Alkemade R, *et al.* (2010) Biodiversity scenarios: Projections of 21st century change in biodiversity and associated ecosystem services. *Technical Series 50*. Montreal: Secretariat of the Convention on Biological Diversity.
Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press.
Morgera E and Tsioumani E (2011) Yesterday, today, and tomorrow: Looking afresh at the convention on biological diversity. *Yearbook of International Law* 21.
Netherlands Environment Assessment Agency (2010) Rethinking Global Biodiversity Strategies: Exploring structural changes in production and consumption to reduce biodiversity loss. http://www.pbl.nl/sites/default/files/cms/publicaties/500197001.pdf
Pereira HM, Leadley PW, Proença V, *et al.* (2011) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496.
Prip C and Gross T (2010) *Biodiversity Planning: An Assessment of National Biodiversity Strategies and Action Plans*, pp. 1–5. United Nations University.
Secretariat of the Convention on Biological Diversity (2010a) *Action for Biodiversity: Towards a Society in Harmony with Nature*. Secretariat of the Convention on Biological Diversity.
Secretariat of the Convention on Biological Diversity (2010b) *Global Biodiversity Outlook 3*, pp. 8–13. Secretariat of the Convention on Biological Diversity.
TEEB (2010) The Economics of Ecosystems and Biodiversity: Mainstreaming the Economics of Nature: A Synthesis of the Approach, Conclusions and Recommendations of TEEB. http://www.teebweb.org/Portals/25/TEEB%20Synthesis/TEEB_SynthReport_09_2010_online.pdf
United Nations (1992) *Convention on Biological Diversity.* Rio de Janeiro, 5 June 1992. United Nations.
Worm B, Hilborn R, Baum JK, *et al.* (2009) Rebuilding global fisheries. *Science* 325: 578.

Cooperation, Evolution of

Egbert Giles Leigh Jr, Smithsonian Tropical Research Institute, Panama, Republic of Panama

© 2013 Elsevier Inc. All rights reserved.

Glossary

Altruism Assistance provided by an agent to other individuals at the expense of its own lifetime reproductive fitness.

Ethology The study of the behavior and social relations of animals.

Genetic terms A locus is a particular place on a chromosome (strictly speaking, an "address" on a chromosomal map) where a gene resides. This gene may be any one of several alternative types or alleles. A diploid cell contains two copies of each autosomal chromosome, and therefore, two genes at each autosomal locus – one from its mother and one from its father. A diploid cell in the "heterogametic sex" also contains two sex chromosomes, which may differ in form (one X, one Y), whereas in the "homogametic sex" it would have two X chromosomes. A haploid organism contains one gene at each locus. A genome is the set of all of an organism's genes.

Inclusive fitness An individual's inclusive fitness is its lifetime reproductive output if not assisted by others, plus the added reproduction it confers on each relative it helps, times the proportion r of genes identical by recent descent it shares with that relative (Hamilton, 1964). In an outbred population of diploids with weak selection, $r = ½$ between full siblings, and between parent and offspring, ¼ between half siblings, between an individual and its grandchild, or between aunt and nephew, etc. Inclusive fitness thus defined represents the common interest of the autosomal genome, and is what natural selection maximizes (Grafen, 2006).

Mutualism A mutually beneficial association or interaction, temporary or permanent, among organisms of different species.

Parasitism An association among members of different species in which members of one species (the parasite) feeds on parts of the other's (the host's) body, or on resources the host has obtained. The parasite usually lives inside or on the host.

Social cooperation A mutually beneficial association, temporary or permanent, among members of the same species.

Symbiosis A mutualism among members of different species in which members of one species (the symbiont) live inside or on a body of the other (the host).

What are the Benefits and Dangers of Cooperation?

Natural selection is a competitive process. The winners are those genes which reproduce most successfully, and the individuals that carry them (Fisher, 1930; Hamilton, 1964). As competition in human economies causes people to cooperate so as to compete better with third parties, so in nature competition often favors the evolution of social cooperation among members of the same species, and mutualism, cooperation among members of different species, to compete better with others. Cooperative endeavor is essential to the productivity and diversity of ways of life in natural ecosystems, as it is to human economies.

The evolution of close cooperation among members of the same species was crucial to several major transitions of evolution, such as the evolution of sexual reproduction, multicellular plants and animals, and insect societies. Yet cooperation only evolves if cheaters, which appropriate the benefits of others' cooperation without contributing, can be suppressed or limited. Many animals live in social groups with other members of their species because they depend in some way on each other's activities. Yet an animal's fellow group members are often its closest competitors – for food, mates, or living space. Life in social groups thus poses three questions: Why live in groups (Maynard Smith and Szathmáry, 1995)? What ecological circumstances, and what social mechanisms, prevent competition among group members from destroying their common interest in cooperating (Moynihan, 1998)? How could such social mechanisms evolve (Leigh and Rowell, 1995)? Indeed, how social cooperation evolves is a central problem of ethology.

This problem of ethology has various analogs. First, human beings benefit from, and indeed depend on, social life. Nevertheless, individuals often appropriate for their private advantage an undue share of the common wealth created by social activity. From the time of Aristotle onward, philosophers and statesmen have struggled to learn how to organize a society so that individual advantage coincides most nearly with the common good.

Second, metazoans are complex multicellular animals. A metazoan's cells usually serve their organism well. However, cancer suggests that conflict between an individual and one of its cell lines is a real, often devastating, possibility (Buss, 1987). How did the harmony among a metazoan's cells evolve? How are conflicts between metazoans and their cells reduced or suppressed?

Mutualism between members of very different species evolves more easily. By definition, members of different species do not compete for mates. Because they have different needs, they compete far less intensely for food or shelter than members of a social group, and they are more likely to have different abilities that can be pooled for mutual benefit. Nonetheless, mutualism, like social cooperation, only evolves if cheating can be controlled or suppressed. Mutualism may involve encounters as brief as the dispersal of fig seeds by a bat that ate the fig containing them, and symbioses as prolonged

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00161-1

as that between an aphid and the bacteria in its gut that synthesize vitamins the aphid cannot live without (Leigh, 2010b; Douglas, 2010).

First, mutualistic partnerships can give rise to true individuals – a major transition in evolution. Eukaryotic cells, whose chromosomes are separated from the cytoplasm by a nuclear membrane, provide the clearest example of this process. Most eukaryotic cells contain organelles, mitochondria and chloroplasts, with genomes of their own that reveal their ancient ancestors – bacteria that invaded ancestors of their host cells about two billion years ago (Margulis, 1993). Nowadays, organelles are essential to their host cells. Cells require mitochondria to obtain energy by oxidizing carbohydrates to release carbon dioxide and water. Leaves cannot photosynthesize unless their cells contain chloroplasts. Despite organelles' crucial importance to their host cells, conflicts between them can arise. Fertilized eggs that receive organelles from both parents may be ruined by the struggle for dominance between maternally and paternally derived organelles. Organelles might multiply faster at the expense of serving their host (Eberhard, 1980). How are these conflicts resolved effectively enough to make a coherent individual out of a cell and its organelles?

Second, genes are the "ultimate" units whose selfinterest drives evolution in the sense that no characteristic evolves nonrandomly unless it serves the selfinterest of some gene (Dawkins, 1982). However, an individual gene, divorced from the rest of its genome and an organism appropriate for the genome's expression, is as useless as a piece of computer program without the rest of the program, a computer that can run the program, and an operator who can run the program on that computer. Thus, these ultimate units of selfinterest are utterly dependent on each other. Chromosomes express the mutual advantage that different kinds of genes derive from each other's presence – their mutual dependence on each other's functions. In particular, chromosomes maintain their genes' continued association by ensuring their simultaneous replication (Maynard Smith and Szathmáry, 1995). Indeed, a genome's genes can be thought of as one of the closest and most intimate of all mutualisms. However, some types of genes, which do nothing to benefit the organisms that carry them, multiply independently of the chromosomes, threatening to fill their genome with selfish DNA. How is this sort of destructive competition prevented from draining away the common wealth of the genome?

Third, conflicts can also occur between chromosomal genes and their genome. Meiosis, the process whereby a gamete is assigned one gene at each locus, is usually as fair a lottery as nature can devise, in the sense that at each locus a gamete has equal chance of inheriting its grandmother's or grandfather's gene. When meiosis is fair, that is say, when a gamete has equal chance of inheriting either allele at a heterozygous locus, natural selection favors an allele only if it enhances the survival or reproduction of the organisms carrying it: thus fair meiosis reflects the common interest of an individual's autosomal genes, its autosomal genome. However, a few alleles that injure their bearers spread by biasing meiosis of heterozygotes in their own favor. Such alleles are called segregation distorters. Some experimental populations have been wiped out by the spread of such alleles (Leigh and Rowell, 1995). Why is meiosis normally so fair when segregation distortion can spread alleles so effectively?

Finally, like human economies, ecosystems are both arenas of competition and webs of mutualistic interdependence (Leigh, 1999). Are ecosystems the greatest of all mutualisms? Human influence on natural ecosystems, especially unintended influence, is usually damaging: ecologists call such influence "disturbance." In accord with this language, natural ecosystems are organized for high productivity and diversity, in the sense that unprecedented "random" disturbance depresses one or the other, just as Aristotle, like Fisher (1930), argued that organisms are organized, adapted to survive, and reproduce because visibly mutant individuals usually do so less well than their normal counterparts. Despite the abundance of natural enemies (predators, parasites, and pathogens), organisms depend on their ecosystems for the necessities of life, even the air they breathe. Therefore, organisms share a common interest in their ecosystem's integrity, as if ecosystems are true "commonwealths." Does the common interest of organisms in the integrity of their ecosystem influence the evolution of individual species? If so, how? These are among the greatest mysteries of biology (Leigh, 1999).

The Evolution and Maintenance of Cooperation

Cooperation evolves only if the potential partners all benefit from it. In other words, there must be advantages which are most easily or effectively obtained by cooperating, and all partners must benefit from cooperating, even if unequally. In sum, cooperating must truly serve a common good.

What can be Gained by Cooperating?

Among members of the same species, the most basic form of cooperation is sexual reproduction, by which two individuals jointly produce offspring of more varied genotype than those that either could produce alone (Maynard Smith and Szathmáry, 1995). Sexual reproduction and the associated exchange of homologous chromosomes, creates the selection that maintains the fairness of meiosis. In turn, fair meiosis ensures that alleles spread only if they benefit their bearers (Leigh, 2010a). Sexual reproduction also allows an allele to be tested in a variety of genotypes, so that its spread is governed by its own contribution to fitness rather than the merit of the genotype in which it first happened to appear (Fisher, 1930; Felsenstein, 1974). Thus, other things being equal, selection favors alleles that are good mixers which benefit all the genotypes that include them, which in turn favors a modular expression of gene action. In other words, sexual reproduction favors genetic systems organized to facilitate adaptive evolution (Fisher, 1930). In many species, cooperation for sexual reproduction is prolonged. For example, among insectivorous birds and some fish, the parents share the labor of, or divide the tasks involved in, feeding, sheltering, and defending the young.

Members of a species join in larger groups, as a rule, if grouping enhances safety or effectiveness (Moynihan, 1998).

They may combine for greater safety against predators because, with more eyes, predators are detected sooner or because members of a group can coordinate activities to confuse or repel predators. A group may also be able to overcome a competitor, or a large prey animal, which could defeat any one of its members. Finally, the advantages of a suitable division of labor drove the evolution of insect societies, some of which attained enormous size and great complexity.

Mutualisms among members of different species usually involve complementation of different functions (Leigh, 2010b). Thus, a coral provides its symbiotic algae with nutrients from small animals that the coral polyps catch. In return, the coral receives carbohydrates from the photosynthesis of these algae. A plant provides its pollinators with nectar or pollen in hopes that these mobile animals will convey some of this pollen to another plant of the same species. Some plants provide food and shelter for specific kinds of ants, which repel most of their host plant's herbivores and destroy the growing tips of vines that would otherwise overgrow their host plant.

What Keeps Cooperation from Becoming Parasitism?

Cooperation involves the pooling of labor to create a common good. Why don't some members of a social group cheat their partners rather than collaborating with them? What keeps an individual from exploiting a good created by a partner in mutualism without offering any benefits in return?

Social Cooperation and Kin Selection

An individual's alleles can spread if the individual helps close relatives reproduce at little cost to itself: this is called kin selection (Hamilton, 1964). A child inherits half its mother's genes, one of the two genes at each of its mother's loci. Thus, alleles of either mother or child can spread if their carrier helps the other relative reproduce. Indeed, kin selection expresses the common interest of an individual's autosomal genes.

Kin selection plays a crucial role in the evolution and maintenance of insect societies. A wasp lays an egg where the hatchling larva has immediate access to food. Often, this is a provisioned nest cell. Sometimes, solitary nesting is futile because unguarded nests are always robbed of their eggs or provisions. Communal nesting makes for better-defended nests but intensifies competition for nest cells and provisions, creating winners and losers. If winners and losers are related, it may be more profitable (as measured by the numbers of copies of the loser's alleles propagated to offspring) for the loser to help the winner reproduce, perhaps by provisioning the winner's nest while the winner guards it. Relatedness to the winner enables the loser to benefit from cooperating with it, but their common interest in a defended nest is equally essential to their cooperation.

In most insect societies with division of labor, a single queen usually produces all the young (Fisher, 1930), and most of her daughters are workers who help the queen reproduce. All such societies descend from societies where a queen mates with a single male, ensuring that workers are closely related to each other as well as the queen (Leigh, 2010a). In very large, crowded, colonies, however, the queen mates with many males and mixes the sperm thoroughly to create the genetic diversity among workers needed to resist disease. At least in honeybees, this genetic diversity is also needed to ensure an adaptive distribution of workers over tasks (Mattila and Seeley, 2007). Workers share a common interest in helping their queen if it is impossible, or at least much less effective, for them to reproduce on their own. Among ants, workers are either sterile or lay eggs so slowly that they do better to help their queen (Hölldobler and Wilson, 1990). Because a honeybee queen mates with many males, most of a worker's colleagues are half-sisters. Because a worker is more closely related to her mother's egg than to a half-sister's egg, workers eat eggs laid by half-sisters, rendering worker reproduction futile (Seeley, 1995). The queen thereby creates selection favoring mutual policing among workers to enforce their common interest in helping their queen. This community of interest among the workers makes beehives models of self-organization, where workers perform their tasks, change from task to task, and adjust appropriately to changed circumstances automatically without the queen exerting direct control (Seeley, 1995). Kin selection is crucial to the evolution of these insect societies, but other arrangements are needed to create a strong enough community of interest among the workers for truly complex insect societies to evolve.

Selection among Groups

Selection among groups is one type of kin selection. Suppose that each individual mates only with fellow members of its group, no migrants are exchanged among groups, and each group is founded by emigrants from a single parent group. Then the groups are endogamous. This endogamy, or inbreeding, eventually makes a group's members so nearly genetically uniform, so much more closely related to each other than to outsiders, that an individual's reproductive advantage coincides almost precisely with its group's good as measured by its reproductive success (Leigh, 2010a). In other words, the raw material for natural selection is genetic variation. If genetic variation among a group's members exceeds variation among groups (variance among group means), selection among individuals within a group usually overrides selection among groups. Only if endogamy is so strong that variation among groups exceeds variation within groups can selection among groups prevail. In the 1950s, selection among groups was often invoked to explain cooperation. In 1966, however, George Williams showed that very stringent conditions were required for the good of the group to override individual advantage, and interest in selection among groups diminished greatly (Leigh, 2010a). Nevertheless, selection among groups sometimes enforces cooperation.

To understand how endogamy increases relatedness among a group's members, try the following experiment. Set 10 pennies on a table, five with heads up and five with tails up. Every minute, choose two pennies at random, one to "die" and the other to "reproduce." This is symbolized by turning over the first penny, if need be, so that it has the same side up as the second. In a short period of time, all pennies will have the same side up. If one uses more pennies, still starting with half heads and half tails, uniformity takes longer to achieve but eventually prevails. Next, represent migration

from outside by choosing a penny at random every 5 or 10 min and tossing it to see which side should be up. Now, uniformity is achieved more rarely and briefly, if ever.

How endogamous must groups be to allow selection among groups to work? Let groups, as well as individuals, have finite lifetimes. The intensity of selection among groups on a characteristic is the proportionate increase in the number of groups founded per preexisting group per group lifetime conferred by its members' possession of this characteristic, just as the intensity of selection among individuals on this characteristic is the proportionate increase that this characteristic confers on an individual's lifetime reproductive success, relative to fellow group members. Selection among individuals balances an equally intense selection among groups on the same characteristic if (1) each group descends from a single parent group, and one migrant is exchanged per two groups per group lifetime, or (2) no migrants are exchanged but one of every N groups is founded by the joining of colonists from two parent groups, where N is the number of mature individuals per group (Leigh, 2010a). Moreover, selection among groups is unlikely to override strong within-group selection unless there are many more groups than individuals per group. In sum, selection among groups is decisive only if groups are definite individuals in their own right.

Selection among groups played an integral role in the evolution of metazoans. All metazoans descend from sexually reproducing ancestors, whose young began life as a fertilized egg or zygote. A zygote grows by dividing mitotically so that an animal's cells all have the same genotype. Each individual is therefore genetically unique, but (mutations excepted) an individual's cells are genetically identical. Moreover, selection favors cells that can distinguish "self" from "nonself," thereby preventing the exchange of migrant cells among individuals. These circumstances identify the reproductive interests of an animal very closely with those of each of its cells (Maynard Smith and Szathmáry, 1995).

Mutants do occur, however, and cancerous cell lineages can run riot at their organism's expense. Multicellular plants and animals have many features which enhance the harmony among an individual's cells (Buss, 1987). In some groups of animals, the "germline" is sequestered, preventing a cancerous lineage from spreading to an individual's offspring. This circumstance is analogous to a honeybee or ant queen preventing worker reproduction to create a common interest among them in helping her reproduce. In some species, the mother's genes control the early stages of the development of her young just as, by judicious distribution of food or hormones, the queen of a complex ant society apportions her workers among different "castes," thereby programming a suitable division of labor for her colony (Hölldobler and Wilson, 1990).

Vertically Transmitted Partners in Mutualism and Group Selection

Intimate mutualisms, where members of one species, the symbionts, live inside members of the other, larger species, the host, are called symbioses (Douglas, 2010). Cooperation within symbioses is maintained if each host's symbionts can only be transmitted to its offspring (vertical transmission). Vertical transmission creates a selection among groups of symbionts for benefitting their hosts, each group consisting of one host's symbionts.

Selection among groups played a crucial role in evolving mutualism between organelles and their host cells. Cells do not exchange organelles. Each cell receives its organelles from a single preexisting cell. This is usually, but not quite always, true even of zygotes. Most kinds of eukaryotic organisms have evolved means to ensure transmission of organelles from one parent, usually the mother. Organelles are therefore subject to a stringent selection among groups (each group being the organelles of given type in one cell) in their host's interest: indeed, organellar reproduction depends utterly on the reproductive success of their hosts (Leigh, 2010b).

How did organelles become susceptible to selection among their hosts? Consider mitochondria. The mitochondrial ancestors which first invaded host cells were parasites (Davidov and Jurkevitch, 2009). In some parasites, selection favors increased dispersal from one host to another, although this injures their current host's welfare. For others, selection favors caring for their current host even if this reduces the effectiveness of dispersal. It depends on how greatly increasing dispersal impairs the current host's usefulness, which differs from case to case. The complementation of functions between a host cell which procures food and an ancestral mitochondrion with aerobic respiration which metabolizes food far more effectively than the host can, presumably favored symbionts which cared for their current hosts. As a house cat defends its owner's house against conspecifics, so ancestral mitochondria defended their valuable hosts against conspecific invaders. Thus, migration of mitochondria between cells was prevented, thereby subjecting the mitochondria to selection among host cells. Selection among hosts favored mechanisms ensuring uniparental transmission of organelles when their hosts reproduced sexually, thereby perfecting the mechanism maintaining mutualism between host cells and their organelles (Leigh, 2010a).

Vertical transmission of symbionts and exclusion of immigrants from other hosts preserve a wide range of mutualisms. For example, aphids depend on bacteria in their guts to synthesize vitamins the aphids need: in turn, the bacteria depend on their hosts for carbohydrates and shelter: mothers pass on these bacteria to their offspring (Douglas, 2010). Similarly, leaf-cutter ants farm a fungus which the ants nourish with leaf fragments they cut from nearby trees. The ants eat certain parts of the fungus they cultivate, and also feed these fragments to their brood. Mutualism between ant and fungus is preserved because the fungus reproduces only when a new leaf-cutter queen takes a mouthful of fungus from its mother's nest to start a fungal culture in her own. Leafcutters carefully remove foreign fungal spores, preventing both infection by pathogenic fungi and immigration of symbiotic fungi from other nests (Hölldobler and Wilson, 1990).

Choosing Cooperative Partners

How can symbioses be maintained when hosts must call their symbionts from their surroundings rather than inherit them from these hosts' mothers? One approach to the problem is to test the usefulness of potential symbionts, and choose only those that pass the test, as a female bowerbird judges the versatility and stamina of her suitors by the size and

attractiveness of the bowers they build (Leigh, 2010b). Nocturnally active bobtail squid, *Euprymnia scolopoides*, are past masters at choosing functional partners. These squid depend on symbiotic bioluminescent bacteria, *Vibrio fischeri*, to adjust their shade to match that of the surrounding seawater, in effect hiding the squid from both its prey and its predators. In return, the squid houses its bacteria and feeds them abundantly. Margaret McFall-Ngai and Edward Ruby have documented how these squid choose appropriate symbionts (Leigh, 2010b). When only two hours old, a hatchling squid calls gram-negative bacteria from the surrounding water, trapping those responding to the call in mucus. Many are called, but few are chosen. After a few hours, only *Vibrio fischeri* remain from all the bacteria that responded to the call. These *Vibrio* then enter ducts leading from the squid's undersurface to its light organs. To reach and colonize these organs, however, these *Vibrio* must run an obstacle course quite as rigorous as any a fairy-tale swain must surmount to win his princess. The *Vibrio* must be able to move, survive strong acidity and potent microbicides, release compounds that make the squid's light-organs a suitable home for them, and produce light, before they can multiply and fill their new home (Leigh, 2010b).

Partner Fidelity and Entangled Fates

Some partners in mutualism are so dependent on each other that injuring one imperils the other. For example, if two men, each capable of wielding only one oar, are in a stormy sea, both must row if either is to reach shore and survive (Maynard Smith and Szathmáry, 1995). Their fates are conjoined: failure of one to cooperate kills both. Similarly, in monkey species such as *Cebus capucinus*, which live in groups that fight each other to maintain territories large enough to supply sufficient food even in the season of shortage, injuring a fellow group member injures the group's capacity to fight. If emigration to other groups is difficult or impossible, a group's monkeys share a common interest in their group's welfare that promotes cooperation (Leigh and Rowell, 1995).

J.L. Wulff (2008) has documented mutualism among three species of branching coral reef sponges. Sponges of each species grow better and live longer if intertwined with a sponge of one of the other two mutualistic species than if it grows alone, or intertwined with another of its own species. This mutualism arises because each partner contributes the ability to resist different hazards to the pair's mutual benefit: each species resists different diseases, different predators, and different abiotic threats such as sedimentation or breakage by storms. When a predator or disease attacks a sponge of a vulnerable species intertwined with one of a resistant species, at least those parts of the vulnerable species attached to the resistant one survive to grow again another day (Wulff, 2008).

Enforcement by Hosts of Cooperation from Live-in Symbionts

In many symbioses, hosts must supply food to their symbionts to enable them to provide benefits in return. The symbionts, however, must often expend copious energy to cooperate in the desired manner. In such symbioses, hosts of many species enforce cooperation by denying uncooperative symbionts resources they need to survive and reproduce (Leigh, 2010b).

In plants of the family Leguminosae, roots release chemicals that attract nitrogen-fixing bacteria, which use carbohydrates supplied by their hosts to convert atmospheric nitrogen, N_2, into ammonia. When the exchange of signals between bacteria and host root is appropriate, the plant lodges these bacteria, called rhizobia, in nodules in their roots, where they are supplied with oxygen and the carbohydrate fuel they need to live, reproduce, and turn atmospheric nitrogen into ammonia. Chemical signals can lie, however: unlike bobtail squid, these plants cannot assess the cooperativeness of their bacteria before lodging them. If, however, a nodule's bacteria do not produce ammonia, the plant host cuts off the nodule's oxygen supply, halving these bacteria's fitness (Kiers and Denison, 2008). Similarly, many corals attract symbiotic algae, zooxanthellae, from the surrounding water, lodging them in their polyps, where the algae supply their host corals with carbohydrates in return for a safe place to live and the nutrients released when their host polyp digests organisms that it has snared. If, however, zooxanthellae cease to produce carbohydrates, their host coral expels them (Leigh, 2010b).

Mutual Enforcement

Mutual enforcement is vital to maintaining many cooperative endeavors, within both social groups and mutualistic partnerships. The author mentioned previously how honeybee workers mutually enforce their common interest in helping their queen reproduce. By eating each other's eggs, they make it pointless for workers to lay eggs on their own rather than helping their queen.

An analogous mechanism enforces the fairness of meiosis. Fair meiosis ensures that selection favors only those alleles that benefit their carriers. Some "distorter alleles," however, bias meiosis in their own favor (Leigh, 2010a). Up to 95% of the sperm of an individual heterozygous for a segregation distorter carry the distorter allele. If individuals homozygous for the distorter die childless, the distorter allele spreads until death of its homozygotes balances its spread from biased meiosis. Imagine a mutant at a locus on another chromosome whose only effect is to restore fair meiosis in all its bearers. Because alleles on different chromosomes assort independently at meiosis, no allele at this mutant's locus can benefit from the distorter's bias. This mutant spreads because it prevents the distorter's spread among its bearers and thus spares some of its bearers from death caused by being homozygous for the distorter. Because mutants which restore fair meiosis are favored at loci on every chromosome but the distorter's, fair meiosis expresses the common interest of the genome as a whole. This selection, which depends on sexual reproduction and recombination of genes on homologous chromosomes, has been demonstrated empirically. Therefore, even though genes are the ultimate units of selfinterest which drive evolution, it is reasonable to analyze the adaptation of whole animals.

The most familiar form of mutual enforcement is reciprocal altruism – help those who help you, and make it stick by retaliating against those who cheat you (de Waal, 1996). Reciprocal altruism is illustrated by the sexual behavior of hamlets, simultaneously hermaphroditic coral reef fish, studied by Eric Fischer (Leigh and Rowell, 1995). Even though a sperm is much smaller and "cheaper" than an egg, a successful

sperm contributes as many genes to future generations as a successful egg. Thus the common interest of autosomal genes favors hermaphrodites that devote as much energy to male functions – making great quantities of sperm, fighting or otherwise competing with each other for mates, and so on – as to the female functions of bearing and raising young. The proportion of a population's reproductive effort devoted to male functions represents the "50% cost of sexual reproduction." Hamlets avoid this cost by pairing off at mating time and alternating sex roles in successive spawns (matings). The fish which releases eggs during one spawn fertilizes the other's eggs during the next. This "egg-trading" short-circuits competition for mates and permits much lower sperm production, greatly reducing the cost of male functions. Reciprocation is favored even though it is cheaper to play the male role. At successive rounds of reciprocal spawns, each partner releases more eggs for the other to fertilize, as if mutual trust were increasing, whereas a fish that tries to play the male role twice in a row is dismissed by its partner to find another mate and begin the process of confidence building all over again.

Cooperation among members of a chimpanzee group is also based in large part on reciprocal altruism (de Waal, 1996). The community of interest among a group's chimpanzees is based on two fundamental features. First, a chimpanzee is in danger outside a group: it needs to belong to one. Second, a group's effectiveness depends on all its members. Thus, all must accommodate, even the dominant, because without his fellow members the dominant would be alone (W.J. Smith as cited in Leigh and Rowell, 1995). In captive groups, chimpanzees help those who have helped them, and they retaliate against those who have attacked them or failed them in time of need. Chimpanzees who are most generous to others when they obtain food are most likely to be given food when they beg for it. Chimpanzees' strategies of deception show that they are sufficiently selfaware that they can predict another's behavior by "putting themselves in the other's shoes." A chimpanzee who has supported another in a fight expects support from that other if it is attacked in consequence, and it will show anger if that support is not forthcoming.

Chimpanzees, however, seem to have a sense of justice (defined as what promotes the group's common good) that transcends reciprocal altruism (de Waal, 1996). Females may gang up on an alpha male who is taking excessive revenge on a subordinate. The group supports alpha males who mediate fairly in disputes, protecting lower ranking animals. If an alpha male always favors his allies when he interferes in fights, females will keep him from interfering. Indeed, alpha males retain their dominance by consent of the dominated. Finally, chimpanzees try to reconcile with animals whom they have just attacked, while the group noisily celebrates such reconciliations. Because fights endanger the welfare of the whole group, reconciliations serve the "common good," including that of the reconciler, who needs its fellow group members.

Cooperation Involving Brief Exchanges

The forms of cooperation discussed so far are long-term partnerships. Many other forms of cooperation involve brief exchanges analogous to those between shoppers and storekeepers in large cities (Leigh, 2010b). In many species, sexual reproduction involves matings between individuals that will never meet again. Among plants, the bee that takes pollen from one plant to another of its species, may move on to different species. Just as shoppers compete for bargains and storekeepers for shoppers, so animals compete for the best mates and flowers for effective pollinators. Cooperation involving such brief exchanges has played a crucial role in evolution. Sexual reproduction made possible the evolution of genetic systems designed to facilitate adaptive evolution. Employing animals to pollinate their flowers and disperse their seeds enabled young to escape their parents' specialized pests and pathogens without sacrificing genetic variation, and to divert resources from defense to fast growth, allowing flowering plants to replace conifers (Leigh, 1999). How do partners in these brief exchanges keep the risk of being cheated low enough to make these exchanges profitable? We do not yet know. This question is particularly urgent since the ancestors of pollinators were pollen thieves or seed predators, and the ancestors of seed dispersers were likewise seed predators (Leigh, 2010b).

Indeed, such brief encounters offer no opportunity to punish partners that cheat, or supply inefficient or substandard goods. To avoid being cheated, individuals must try to choose more cooperative partners and exclude nonperformers. Individuals share a common interest in mating with others of their own species (Fisher, 1930), but not all males of their species are equally suitable mates. Females of promiscuous species judge the suitability of courting males by criteria ranging from the size, beauty, and elaborateness of the structure a male bowerbird builds and defends from other males or the capacity for sustained, vigorous flight of a courting sulfur butterfly to the bright plumage of a healthy cock-of-the-rock. In these brief exchanges, however, a male's only interest is that his mate devote maximum effort on his young, even if this lowers her expected lifetime reproduction. Once copulation begins, a male may inject substances meant to keep her from mating with other males and to drug her into devoting all possible effort to his young. The female must evolve a variety of ways to parry these threats. And indeed, even after copulation starts, females of many species can prevent the copulating male's sperm from fertilizing their eggs (Eberhard, 1996). Despite these conflicts and hazards, all many-celled plants and animals descend from sexually reproducing ancestors, as if sexual reproduction were needed to evolve large, complex organisms. Most species of eukaryote, moreover, reproduce sexually, at least occasionally.

Similarly, plants use their flowers' design to try to choose their pollinators. Different rewards – resin, musk, perfume – attract different pollinators. The location of the reward within the flower (is it easily reached, or are specialized structures required to reach it?) and the flower's shape (Does it offer the pollinator a place to land, or must it hover?) and color also restricts the range of animals attracted. Nonetheless, cheating is rife among both plants and potential pollinators. Large trees, or plants of common species, cannot afford to cheat, for pollinators could easily learn to avoid them, but whole groups of orchids, and many other rare plant species, use a stunning variety of ruses to attract animals to pollinate their flowers without providing them any reward for doing so. Similarly, in at least some plant species, most flowers are robbed of at least

some of their nectar by animals that do not pollinate them. Despite the risks of cheating, most plants, especially in the tropics where competition is most intense, employ animals to pollinate their flowers (Leigh, 2010b).

Mutualisms of brief exchange are therefore essential for both evolutionary progress and ecosystem function. Nonetheless, we do not yet know how participants in these exchanges reduce the risk of being cheated to levels that do not threaten the profitability of participating in them.

Cooperation and the Common Good

Cooperation evolves only if there is genuine community of interest among the potential partners. In his *Politics*, Aristotle expressed this point very clearly: States whose constitution and social organization serve the common good more nearly are less susceptible to overthrow by popular revolution or factional putsch. Often, a common interest permits mutualisms to develop that lack any means of enforcement. However, even coreplication of host and symbiont (whereby symbionts can propagate only to the offspring of the host), such as that maintaining the mutualism between eukaryotic cells and their organelles, will not create a stable mutualism if host and symbiont do not share a stable common interest.

Mixed-species bird flocks are examples of mutualisms without enforcement (Leigh and Rowell, 1995). A typical Neotropical mixed flock contains one pair each of several "nuclear species," some with attendant young. Adults of these nuclear species jointly defend a common territory. As the flock progresses regularly over its territory, certain birds with smaller territories join the flock as it passes through them, whereas other birds with larger territories move from flock to flock. Each species of bird eats different food, or feeds in different types of places, as if to minimize competition among flock members. The advantage of a flock is more eyes to watch for predators and more birds to "mob" a predator when need be (Moynihan as cited in Leigh and Rowell, 1995). Even an alarm signal benefits its giver – not only by helping to keep safe the group on which it depends but also by informing a potential predator that, since it has been seen in time to be avoided, it might as well go elsewhere. This mutualism of seemingly unenforced common interest has caused the evolution of certain colors and behaviors among members of certain mixed flocks (Moynihan, 1998).

One of the most striking mutualisms of the tropical forest is that between fig trees and their pollinating wasps (Herre, 1996). Each fig fruit is a flowerhead turned outside in to form a ball, lined with flowers on the inside, with a hole at one end. Each fruit is pollinated by one or more female wasps. These enter the fruit, pollinate its flowers, and lay eggs in up to half these flowers. Each wasp's larva grows within a single fig seed. When a fruit's adult wasps "hatch," they mate among themselves and the fertilized females fly out in search of new trees to pollinate.

These pollinator wasps are parasitized by nematodes. In some fig species, each fruit is pollinated by several wasps, each carrying nematodes into the fig. The young of these different nematodes compete to enter the fertilized females leaving the fig, reducing the wasps' ability to reach other figs, and their reproductive success should they reach one. In other, "single-foundress" fig species, almost every fruit is pollinated by a single wasp. In these species, nematodes can only infect the young of their current host. Because of this coreplication, their reproductive success depends on that of their host. Selection thus favors nematodes which minimize the damage they inflict on their host. Such nematodes may sometimes enhance their host's reproduction.

Coreplication shaped a stable mutualism between eukaryotic cells and their organelles, which persists in those few species whose members receive organelles from both parents. Nonetheless, coreplication of wasps and nematodes in single-foundress fig species has had no such effect. Some fig species with several pollinators per fruit are descended from single-foundress species. Nonetheless, their nematodes have reverted to the status of damaging parasites. They must lack a stable community of interest with their host wasps (Herre, 1993).

Another striking mutualism involves reef-building corals and their zooxanthellae, the symbiotic algae that supply them with carbohydrates in return for nutrients and a place in the sun. Many of these corals call their zooxanthellae from the surrounding water rather than inheriting them from their mothers by mutualism-enforcing coreplication. Wasps which inherit their nematodes from several hosts apiece suffer grievously from these parasites. Multiple origins of a coral's zooxanthellae do not disrupt the coral–algal mutualism, presumably because the complementation of functions between corals and their zooxanthellae establishes an effective community of interest between these organisms. Community of interest makes the mutualism, not the breeding system.

Similarly, cooperation does not outlive the community of interest among its partners. When a coral ceases to benefit from its zooxanthellae (as often happens when the seawater becomes too warm), it expels them. When the queen dies in a colony of the social wasp *Metapolybia aztecoides*, no one member of the colony can lay enough eggs to maintain the colony's pool of workers. Several would-be queens thus join forces to undertake this task. Practice, however, increases their egg-laying rate. When the reproductives no longer need each other, they fight it out for the profits of their cooperation (West-Eberhard, 1978). In summary, a defensible community of interest among the potential partners is the one thing needful for the evolution of cooperation. Coreplication, social system, reciprocal altruism, etc. can only affect enduring cooperation if there is a genuine community of interest among the potential partners.

Why is Cooperation so Crucial to Progressive Evolution?

Cooperation has allowed the successive formation of more complex and effective wholes from parts already tested by natural selection (Maynard Smith and Szathmáry, 1995). Moreover, wholes composed of modules already adapted to respond "constructively" to challenging circumstances are more capable of further adaptive evolution (Gerhart and Kirschner, 1997). Finally, the conflicts between selection at different levels, and the mechanisms by which these conflicts are controlled or suppressed, represent a series of distinctive footprints of the decisive role played by natural selection in

macroevolution. Indeed, study of the role of mutualism in evolution allows evolutionary history to testify decisively to the mechanisms of evolution.

Cooperation and the Major Transitions of Evolution

Cooperation played a crucial role in all the major transitions of evolution. Perhaps the first of these transitions, almost effaced by the mists of time, was the transformation of nucleic acid molecules which replicated themselves by parasitizing protein-based metabolism into genes programming the metabolism and development of discrete organisms (Dyson, 1985). Whether Dyson is correct or not, the origin of life as we know it did involve the transformation of mutually dependent genes into coherent genomes (Maynard Smith and Szathmáry, 1995). Later transitions include the evolution of bacterial cells containing a variety of smaller parasitic and commensal microbes into genuine eukaryotes (Margulis, 1993); the evolution of sexual reproduction, whereby two individuals cooperate to produce offspring more varied than either could alone (Maynard Smith and Szathmáry, 1995); the evolution of meiosis, which ensures an appropriate apportionment of the genes of a diploid cell to its haploid descendants (Maynard Smith and Szathmáry, 1995); the evolution of complex, multicellular organisms (Buss, 1987); and the evolution of animal societies (Leigh and Rowell, 1995).

All these transitions involve entities cooperating toward achievements which no one entity could accomplish unaided. Many of these transitions, such as the evolution of eukaryotes, metazoans, and complex insect societies, transformed groups of formerly independent entities into coherent, integral wholes. Insofar as evolutionary progress has occurred, the evolution of cooperation has made it possible.

Cooperation, Modularity, and Evolvability

To show how complex systems might evolve, Herbert Simon (1962) compared two watchmakers, Hora and Tempus. Hora put together a watch by making subassemblies of 10 parts apiece and then combining them into 2d-level assemblies of 10 subassemblies apiece, and so forth, until her watch was finished. Tempus would try to put together a watch without benefit of subassemblies. Both were subject to interruptions. An interruption forced Tempus to begin his watch again from scratch, whereas it forced Hora to begin her current subassembly again. As a result of her subassemblies, only Hora could finish watches.

The relevance of this parable to evolution is that cooperation can lead to combining preexisting "subassemblies," each already tested by natural selection, into larger, more effective wholes (Simon, 1962; Dyson, 1985). Cooperation thereby allowed the evolution of complex organisms by manageable stages.

Moreover, when cooperating entities share a common interest strong enough for selection to transform groups of cooperators into better integrated, more effective wholes, the resulting combines benefit from two properties. First, the selfinterested adaptability of the component entities makes for a more effective and adaptable combine. Second, the adaptability of its components enhances the combine's capacity for adaptive evolution.

The first of these propositions is illustrated by the behavior of the oldest of intracellular organelles. The microtubules and centrioles (centrosomes) of eukaryotic cells, responsible for the "asters" and "spindle" that form during mitosis and meiosis, are not programmed by nuclear DNA. Instead, these organelles appear to be remnants of spirochetes which invaded the ancestors of their hosts countless ages ago (Margulis, 1993). When a cell divides in mitosis, the centriole, just outside the nucleus, divides first, and its two descendants move to opposite sides of the nucleus. Microtubules grow out from the centrioles, forming asters. A microtubule quickly dissolves unless it encounters a part of a chromosome called a "kinetochore," which stabilizes a microtubule touching it. This "random exploration" by microtubules allows chromosomes to be attached to their respective asters in a few minutes, accomplishing the essential first step in apportioning a complete set of chromosomes to each daughter cell (Gerhart and Kirschner, 1997). At the level of cells within organisms, the generation and multiplication of a great variety of "T cells" allows selection for and mass production of those antibodies required to protect their organism from disease (Gerhart and Kirschner, 1997). At the level of organisms within societies, the search by foraging honeybees for new sources of food and the communication of successful searches to the hive, thereby recruiting other foragers to the newly found food, allow honeybee colonies to allocate foraging effort effectively (Seeley, 1995). At all three levels, forms of random exploratory search, conjoined with mechanisms rewarding success, play a crucial role in the development of well-adapted morphology and behavior. Search by quasiautonomous organisms often serves these purposes more effectively than more centrally controlled processes (Gerhart and Kirschner, 1997).

Adaptability of cells and tissues greatly increases the probability that a major change, genetic or environmental, will not prove lethal. E.J. Slijper described a goat born without forelegs. This goat learned to hop about like a kangaroo. To suit this goat's peculiar two-legged hop, its hind legs became enlarged, its spine became curved and the sizes and shapes of its vertebrae changed, and the sizes and points of attachment of many muscles and ligaments also changed. These adjustments testify to the adaptability of this goat's cells and tissues. Similarly, N. Smythe (as cited in Leigh, 1999) was able to create a social strain of pacas, otherwise fiercely territorial animals, by rearing one generation of young under special conditions. These young passed on their new social "traditions" to their descendants. The adaptability of animals, developmental and social, allows them to adapt to some very unusual circumstances.

Cooperation and Unmistakable Footprints of Natural Selection

The author has shown how several evolutionary transitions involved smaller parts combining into larger, more integral wholes. Thus, certain parasites and commensals of bacteria were transformed into organelles of eukaryotic cells, certain multicellular aggregates were transformed into coherent

multicellular individuals, and certain insects which lived together in groups were transformed into units of cohesive insect societies with a single reproductive and complex division of labor.

A trace of each of these transitions is left in the potential conflicts between the erstwhile parts and the whole they formed. Competition between organelles inherited from different parents can injure their hosts. The spread of cancerous cells can kill the individual to which they belong. Worker bees can lay eggs of their own rather than help their queen.

How a transition occurred is often revealed by the mechanisms which suppress conflicts between the resulting whole and its parts. Most organisms are designed to ensure uniparental transmission of organelles, thereby preventing competition between organelles from different parents and ensuring that the reproduction of organelles depends utterly on the reproduction of their hosts. In most metazoan species, sexual reproduction ensures that each individual is genetically unique, but that individual's cells are genetically identical save for the accidents of somatic mutation, a circumstance that ensures that selection discriminates among individuals and not the cells within a single individual. The other means by which metazoans reduce the threat of "rebel" cell lineages – sequestration of the germline (so that cancers of the body cannot be passed on to offspring), maternal control of the early stages of development of her young, and so on – are paralleled by the means whereby queens of complex insect societies maintain harmony within their colonies.

Since these conflicts are suppressed by the very mechanisms that joined the parts into larger wholes, the mechanisms involved represent unmistakable footprints of the critical role natural selection played in these transitions – the most important events of macroevolution. The conflicts involved in each such transition, and the means by which they are suppressed, reveal the historical path of evolution and the selective mechanisms that directed it. Thus, study of the role of mutualism in evolution enables evolutionary history to testify to the mechanisms of this evolution (Maynard Smith and Szathmáry, 1995).

Cooperative Endeavor and Ecosystem Evolution

In our concern with how natural selection of random mutation can yield precise, complex adaptation, we easily forget that organisms evolve in ecological communities. The genius of Archimedes, Newton, Darwin, and Einstein could only be expressed in an urbanized, literate civilization with a diversity of occupations and an audience able to appreciate and benefit from their discoveries. Similarly, social insects, energetic warm-blooded mammals, and flowering plants dependent on animal pollinators could only evolve in diverse ecosystems with intense predator and herbivore pressure and a great variety of resources. In Madagascar, a famous orchid, *Angraecum sesquipedale*, attracts high-performance hawkmoths to pollinate its flowers; these moths need a 30-cm-long proboscis to reach its nectar. On the far smaller island of Réunion, where diversity is far lower and competition far less intense, the orchid *Angraecum cadetii*, descended from Malagasy invaders closely related to *Angraecum sesquipedale*, has a much simpler flower, offers a poorer reward, and attracts a flightless cricket as its only pollinator (Leigh, 2010b).

Indeed, just as human society could not progress without individuals' innovations in fields ranging from agriculture and transport to communication and politics, so modern ecosystems could not evolve without the major transitions of evolution. Thanks to a succession of crucial innovations such as those just mentioned, human societies have evolved many different occupations (ways to make a living), high economic productivity, intense competition, and complex forms of interdependence. The array of human laws devoted to preventing cheating on various forms of cooperative endeavor – laws against violating contracts, embezzling, fraud, bribery, and so on – reflect how much economic productivity and social welfare depend on various forms of cooperation. The role in ecosystem evolution of the major transitions of evolution, each centering on a new form or expression of cooperation, suggests that ecosystem productivity and diversity depend utterly on cooperative endeavor (Leigh, 2010b).

Eukaryotes, organisms whose cells have nuclei, evolved when natural selection transformed α-proteobacteria of an aerobic, parasitic species that invaded individuals of an archaeobacterial species into mutualistic mitochondria (Davidov and Jurkevitch, 2009). All eukaryotes descend from this one mutualistic symbiosis. This mutualism led to the evolution of meiosis, mitosis, and orderly sexual reproduction. Mitochondria lost many of their genes, and other genes were transferred to their hosts' nuclei, but they retained their ability to produce energy-rich adenosine triphosphate. These mitochondria provided the energy their hosts needed to replicate nuclear genomes many times larger than any bacterium's, and to express these genomes' genes. These large genomes allowed the evolution of complex, multicellular organisms (Lane and Martin, 2010).

Symbioses between corals, whose polyps snare small animals from the surrounding water, and zooxanthellae living in these polyps and supplying them with carbohydrates in return for the fertilizing wastes from the digestion of the animals the polyps caught, allowed coral reefs to evolve. The first and largest coral reefs formed off tropical shores in the Devonian. They have disappeared and reformed many times since. The sessile bivalves that replaced reefs during the warmest periods of the Mesozoic also harbored symbiotic algae. Finally, symbiotic cyanobacteria allowed filter-feeding sponges to colonize those parts of Australia's Great Barrier Reef furthest offshore, where the water is clearest and the plankton sponges live on is most scarce (Leigh, 2010b). The reefs resulting from this symbiosis are the most diverse, complex, and beautiful of marine ecosystems.

Employing animals as pollinators and seed dispersers was crucial to evolving today's productive, diverse tropical rain forests. Employing animals as pollinators enabled the understory weeds that gave rise to flowering plants to maintain genetic variation even when their species was made rare by specialized pests or pathogens. Moreover, when a species becomes so rare that many of its members escape the attentions of its specialized pests, selection favors diverting resources from antiherbivore defense to fast growth. During the forty million years after flowering plants first appeared, 140 million

years ago, they evolved efficient systems for moving water from the soil through roots and stems to the leaves, enabling them to invade sunny habitats. In the end, leaves of flowering plants averaged four times the vein density (length of veins per mm^2 of leaf) of conifers, which doubled their photosynthetic capacity and potential transpiration rate. Because flowering plants also employed animals to disperse their seeds beyond the reach of their parents' pests, the stage was set for a diverse set of fast-growing flowering trees to replace a far less diverse set of slower-growing, better-defended, wind-pollinated conifers in all but the least favorable habitats. Moreover, the abundant transpiration of flowering plants has increased rainfall and expanded the domain of tropical rain forest. Thus mutualism has made possible the most diverse, productive and magnificent of the earth's terrestrial ecosystems (Leigh, 1999, 2010b; Brodribb and Field, 2010; Lane and Martin, 2010).

See also: Biodiversity, Evolution and. Coevolution. Competition, Interspecific. Diversity, Community/Regional Level. Genetic Diversity. Parasitism. Social Behavior. Species Coexistence

References

Brodribb TJ and Field TS (2010) Leaf hydraulic evolution led to a surge in leaf photosynthetic capacity during early angiosperm diversification. *Ecology Letters* 7: 175–183.

Buss LW (1987) *The Evolution of Individuality*. Princeton, NJ: Princeton University Press.

Davidov Y and Jurkevitch E (2009) Predation between prokaryotes and the evolution of eukaryotes. *BioEssays* 31: 748–757.

Dawkins R (1982) *The Extended Phenotype*. Oxford: Oxford University Press.

Douglas AE (2010) *The Symbiotic Habit*. Princeton, NJ: Princeton University Press.

Dyson FJ (1985) *Origins of Life*. Cambridge, UK: Cambridge University Press.

Eberhard WG (1980) Evolutionary consequences of intracellular organelle competition. *Quarterly Review of Biology* 55: 231–249.

Eberhard WG (1996) *Female Control: Sexual Selection by Cryptic Female Choice*. Princeton, NJ: Princeton University Press.

Felsenstein J (1974) The evolutionary advantage of recombination. *Genetics* 78: 737–756.

Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford: Oxford Univ. Press.

Gerhart J and Kirschner M (1997) *Cells, Embryos and Evolution*. Oxford: Blackwell Scientific.

Grafen A (2006) Optimization of inclusive fitness. *Journal of Theoretical Biology* 238: 541–563.

Hamilton WD (1964) The genetical theory of social behavior I. *Journal of Theoretical Biology* 7: 1–16.

Herre EA (1993) Population structure and the evolution of virulence in nematode parasites of fig wasps. *Science* 259: 1442–1445.

Herre EA (1996) An overview of studies on a community of panamanian figs. *Journal of Biogeography* 23: 593–607.

Hölldobler B and Wilson EO (1990) *The Ants*. Cambridge, MA: Harvard University Press.

Kiers ET and Denison RF (2008) Sanctions, cooperation, and the stability of plant-rhizosphere mutualism. *Annual Review of Ecology, Evolution and Systematics* 39: 215–236.

Lane N and Martin W (2010) The energetics of genome complexity. *Nature* 467: 929–934.

Leigh Jr. EG (1999) *Tropical Forest Ecology: A View from Barro Colorado Island*. New York: Oxford University Press.

Leigh Jr. EG (2010a) The group selection controversy. *Journal of Evolutionary Biology* 23: 6–19.

Leigh Jr. EG (2010b) The evolution of mutualism. *Journal of Evolutionary Biology* 23: 2507–2528.

Leigh Jr. EG and Rowell TE (1995) The evolution of mutualism and other forms of harmony at various levels of biological organization. *Écologie* 26: 131–158.

Margulis L (1993) *Symbiosis in Cell Evolution*, 2nd edn. New York: Freeman.

Mattila HR and Seeley TD (2007) Genetic diversity in honey bee colonies enhances productivity and fitness. *Science* 317: 362–364.

Maynard Smith J and Szathmáry E (1995) *The Major Transitions of Evolution*. Oxford: Freeman/Spektrum.

Moynihan MH (1998) *The Social Regulation of Competition and Aggression in Animals*. Washington, DC: Smithsonian Institution Press.

Seeley TD (1995) *The Wisdom of the Hive: The Social Physiology of Honeybee Colonies*. Cambridge, MA: Harvard University Press.

Simon HA (1962) The architecture of complexity. *Proceedings of the American Philosophical Society* 106: 467–482.

de Waal F (1996) *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

West-Eberhard MJ (1978) Temporary queens in *Metapolybia* wasps: Non-reproductive helpers without altruism. *Science* 200: 441–443.

Wulff JL (2008) Life-history differences among coral reef sponges promote mutualism or exploitation of mutualism by influencing partner fidelity feedback. *American Naturalist* 171: 597–609.

Corals and Coral Reefs

Nancy Knowlton and Jeremy Jackson, Smithsonian Institution, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by James W. Porter, Jennifer I. Tougas, volume 5, pp 73–95, © 2001, Elsevier Inc.

Glossary

Bleaching The loss of symbiotic zooxanthellae from corals. Bleaching is usually caused by elevated sea surface temperatures, but it can also be caused by low temperature, sedimentation, extreme salinities or light levels, and bacterial infection.

Cnidaria The marine invertebrate phylum containing the reef-building corals.

Colony A collection of interconnected coral polyps. Often the polyps are connected with living tissue, but sometimes tissue death means that only skeletal material connects them.

Coral Triangle The region, located in the waters of southeast Asia (e.g., Indonesia and Philippines), that contains the highest diversity of corals and other marine organisms.

Hermatypic Reef-building; more recently, this term has been replaced by the term zooxanthellate to refer to coral species with symbiotic algae.

Holobiont The entity that comprises the coral animal and the diverse assemblage of microbes that live associated with the coral.

Planula (plural – planulae) The ciliated planktonic larval stage of a coral.

Polyp The basic unit of the coral animal, consisting of a simple tubular structure with a mouth and ring of tentacles at one end. The polyp sits in a skeletal cup called the corallite.

Scleractinia The taxonomic order of cnidarians that includes the reef-building corals.

Zooxanthellae A diverse group of symbiotic, photosynthetic dinoflagellate algae associated with corals and other tropical marine invertebrates.

Coral Reefs – Rainforests of the Sea

Distribution in Time and Space

Living reefs have a long history, and the earliest reefs were built by microbes (Wood, 1999). However, not all geological eras have consistently had abundant reef growth, and the reasons for the waxing and waning of reefs through time are not entirely clear (Kiessling, 2009). Modern reefs, that is, those built primarily by stony corals in the order Scleractinia (class Anthozoa, phylum Cnidaria), date back at least to the Triassic, and many of the groups found associated with reefs today were established by 50–25 million years ago in the Eocene or Miocene (Bellwood and Wainwright, 2002; Renema *et al.*, 2008).

From a geological perspective, reefs may be defined as masses of carbonate limestone, built up from the seafloor by the accumulation of the skeletal material of many coral reef plants and animals. Reef growth has shaped the face of Earth by creating limestone structures more than 1.2 km thick (Enewetak Atoll; Szabo *et al.*, 1985) to more than 2500 km long (Great Barrier Reef), although they occupy only approximately 0.6×10^6 km^2, or approximately 0.1% of the surface of the planet (Reaka-Kudla, 1997). Depending on their proximity to land, coral reefs are classified as fringing reefs (closely paralleling the coastline), barrier reefs (paralleling the coastline farther from shore, with a deep lagoon separating the reef from the shore), or atoll reefs (oceanic ring-like reefs without any relationship to continental or island land masses). Patch reefs are smaller accumulations of coral within lagoons. Darwin (1842) wrote an important early geological treatise on the formation of reefs; his explanation for the formation of atolls – the legacy of reefs once fringing now sunk volcanoes – was finally confirmed in 1950.

To a certain extent, coral reefs are an enigma: on the one hand, they are among the most luxuriant ecosystems on Earth, supporting high diversity and high biomass, and yet on the other hand, they achieve this status in some of the least fertile waters on Earth. Coral reefs flourish on stable substrates within a relatively narrow range of physical parameters. These requirements include sunlit depths (typically <50 m), normal oceanic salinities, warm sea surface temperatures (generally, average monthly minimum temperatures $>18\,^{\circ}$C), high oxygen concentrations, and often high water clarity and low nutrient concentrations (Sheppard *et al.*, 2009). Because corals do not tolerate cold temperatures (Lirman *et al.*, 2011; **Figure 1**), coral reefs are restricted to the tropics, mostly between 30° north and south latitudes. For similar reasons, reefs are best developed and most diverse along the western boundaries of the world's oceans in the tropical western Atlantic (especially the Caribbean) and the Indo-West Pacific, although they can also be found in the tropical eastern Pacific and eastern Atlantic.

Some coral species are more able to tolerate suboptimal extremes, and these form lower diversity reefs where conditions are less favorable. The northernmost genuine coral reefs occur in Japan, the Northwestern Hawaiian Islands, and Bermuda, and the southernmost reefs are found in Australia and South Africa (Spalding *et al.*, 2001; Sheppard *et al.*, 2009). Some reefs of the Persian Gulf tolerate extremes of high temperatures, low temperatures, and high salinities. To a

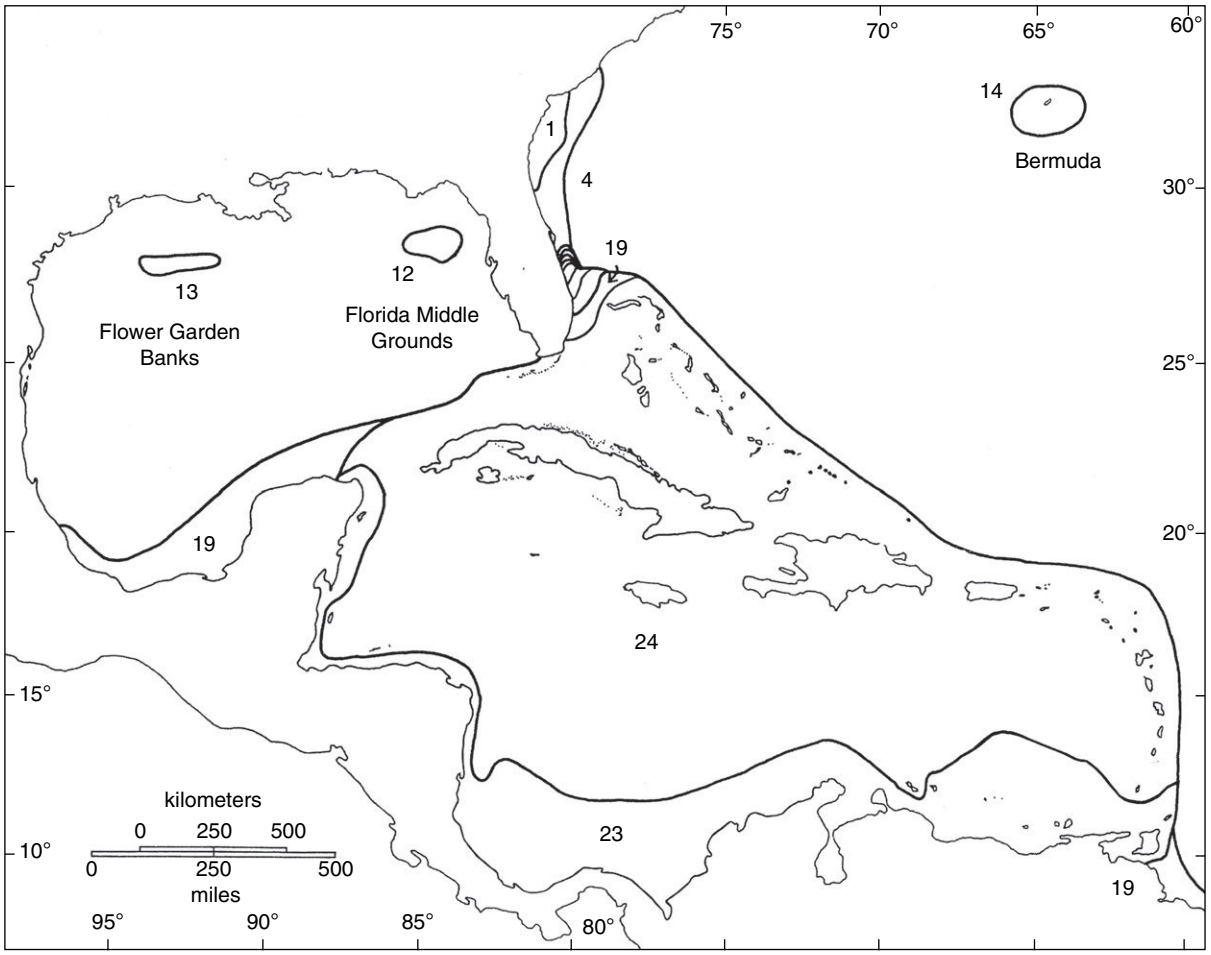


Figure 1 Distribution of coral genera throughout the Caribbean. The dense packing of generic diversity isopleths along the eastern coast of Florida correlates with the frequency and intensity of cold-water disturbances (e.g., see Lirman *et al.*, 2011).

limited extent, corals are able to adapt to ambient conditions; consequently, the upper lethal temperature for a species in the tropics will be higher than that of the same species in the subtropics.

Biodiversity

Diversity of Major Groups

Coral reefs harbor extraordinary biodiversity in terms of the deep branches of the tree of life, a level of diversity that far exceeds any tropical rainforest. For example, commonly found coral reef animals include the phyla Porifera (sponges), Cnidaria (corals, sea anemones, and their relatives), Ctenophora (comb jellies), Platyhelminthes (flatworms), Nemertina (ribbon worms), Brachiopoda (lamp shells), Mollusca (snails, clams, chitons, and cephalopods), Sipuncula (peanut worms), Arthropoda (crabs, shrimps, and their relatives), Annelida (segmented worms), Echinodermata (seastars, brittlestars, crinoids, sea cucumbers, and sea urchins), and Chordata (sea squirts and fishes). In contrast, the conspicuous and common animal phyla found in tropical rainforests are far fewer, primarily the Arthropoda (insects and their relatives), Mollusca (land snails and slugs), and Chordata (tetrapod vertebrates). However, these differences between coral reefs and rainforests are paralleled in most comparisons of marine and terrestrial habitats. Life evolved in the sea and only later emerged on land, so most higher taxonomic groups (phyla and classes) are either restricted to or predominantly found in marine environments (May, 1994).

Species Diversity

It is at the level of species (and genus and family) that the diversity of coral reefs stands out from other marine habitats, and just as insects are far more diverse than trees in tropical rainforests, the bulk of the diversity on reefs lies not with the corals but with the many species associated with them. The species diversity of coral reefs greatly exceeds that of any other marine environment on a per-area basis, and probably overall; the only competitor in terms of total diversity is the deep sea, because of its vast extent. On coral reefs, there are also many cryptic species (species that are hard to distinguish morphologically) (Knowlton and Jackson, 1994; Meyer *et al.*, 2005), although whether cryptic species are proportionately more common on reefs than in other marine habitats is less clear.

The epicenter of marine species diversity today occurs in an area known as the Coral Triangle in southeast Asia. The average diversity for locations in the Coral Triangle for many groups is roughly three times that of the most diverse Caribbean locations; for instance, in an assessment of corals, fishes, snails, and lobsters, the Philippines had nearly 1500 species, whereas the western Caribbean had only 450 (Roberts *et al.*, 2002). In some other groups, the differences in diversity are even more extreme; for example, the tropical Indo-west Pacific has approximately 1200 echinoderm species, whereas the tropical western Atlantic has fewer than 150 (Sheppard *et al.*, 2009). However, the center of diversity for coral reefs has not always been located in the current position of the Coral Triangle (Renema *et al.*, 2008), moving across almost half the globe over the past 50 million years.

At present, no fully comprehensive all-taxa biodiversity inventory has ever been conducted on a coral reef, but it is obvious that were this to be done, the total biodiversity would be extremely high. Several efforts have been made to intensively collect from particular locations or habitats for specific groups, and examples of the extraordinary diversity of reefs are listed in **Table 1**. Bouchet and colleagues have led several such expeditions and found staggering numbers of mollusks and decapod crustaceans. For example, the number of marine mollusk species collected from 42 stations across 295 km^2 in New Caledonia was 2738 (Bouchet *et al.*, 2002). Results from a more recent effort covering 150 km^2 in Panglao, Philippines, are roughly 1200 species of decapod crustaceans and 5000–6000 mollusk species (Bouchet *et al.*, 2009). These data depend on collecting, sorting, and identification by many taxonomic experts and are extremely labor intensive. For example, the collections in New Caledonia and the Philippines involved 400 person-days and six person-years, respectively, but that was just the beginning; despite the fact that the Philippines expedition took place in 2004, six years later, detailed analyses of these collections were still ongoing.

For this reason (and also because of the difficulty of distinguishing cryptic species), efforts to use genetic methods (e.g., barcoding; Hebert *et al.*, 2003) to assess biodiversity have been increasingly employed. For example, Plaisance and colleagues (2009, 2011) using this method found 525 species of crustaceans taken from 6.3 m^2 of reef rubble and artificial substrates; the number of brachyuran crabs found (168) was equal to approximately 80% of the entire described brachyuran crab diversity of all of Europe. Molecular methods are the only way to census microbial diversity in the ocean, and the numbers are staggeringly large. An early study based on cloning and sequencing of just 14 small samples (each 1 cm^2) from corals yielded 430 species of bacteria (Rohwer *et al.*, 2002). More recent analyses using "next-generation" sequencing have made this number seem small (e.g., 300,000 distinct sequences from 10 coral fragments representing bacteria, Archaea, eukaryotic viruses, phage, and fungi (Wegley *et al.*, 2007)).

The numbers presented in the preceding paragraph are absolute counts of what occurred in samples. Statistical methods that examine the prevalence of rare versus common species can be used to estimate the total diversity that would be expected if everything present were counted. Samples from coral reefs typically have many rare species; for example, 20% of the mollusks from the New Caledonia expedition (Bouchet *et al.*, 2002) were represented by a single specimen, as were 38% of the crustaceans analyzed by Plaisance *et al.* (2011) and 23% of the sequences detected by Rohwer *et al.* (2002). The greater the number of rare species, the higher the estimates relative to the observed values; for example, Plaisance and colleagues (2011) estimated that further sampling would have yielded more than 700 crustacean species, and Rohwer and colleagues estimated that further sequencing would have

Table 1 Examples of the extraordinary animal biodiversity of coral reefs

Group	*Number of species*	*Sampling unit*	*Location*	*Source*
Organisms > 0.2 mm (all groups)	309	Colonies of the coral *Oculina arbuscula*	Florida	McCloskey (1970)
Infaunal invertebrates	800	10 m^2	Australia	Poore and Wilson (1993)
	350	10 m^2	Aldabra Atoll	Hughes and Gamble (1977)
Polychaetes	158	6 l of sediment	Oahu, HI	Butman and Carlton (1993)
	103	One colony of living coral	Heron Island, Australia	Grassle (1973)
Motile cryptofauna	776	One reef flat	Moorea	Peyrot-Clausade (1983)
Mollusks	2738	295 km^2	New Caledonia	Bouchet *et al.* (2002)
Crustaceans	525	6.3 m^2	7 sites globally	Plaisance *et al.* (2011)
Boring cryptofauna	220	Dead coral	Solomon Islands	Gibbs (1971)
Cheilostome bryozoans	46	Hard substrates	Jamaica	Jackson (1984)
Hermatypic corals	362	Milne Bay	Papua New Guinea	Werner and Allen (1998)
	350	Great Barrier Reef	Australia	Veron (1985)
	53	Discovery Bay	Jamaica	Wells (1973)
Fishes (all groups)	1500	Great Barrier Reef	Australia	Sale (1977)
	1039	Milne Bay	Papua New Guinea	Werner and Allen (1998)
	496	Bahamas	Caribbean	Bohlke and Chapin (1968)
	517	Alligator Reef	Florida	Starck (1968)
	23	Single coral head, Big Pine Key	Florida	Bohnsack (1979)

yielded 6000 bacterial sequences. However, when there are many rare species, even these statistical methods tend to underestimate diversity (Coddington *et al.*, 2009).

At present across all coral reefs, there are thought to be approximately 93,000 described species of macroscopic coral reef plants and animals, of which almost 66,000 are invertebrates (Reaka-Kudla, 1997). However, as with tropical rainforests, many coral reef species are undescribed, particularly among the tiniest members of the community, which are the most diverse and least known (Reaka-Kudla, 1997; **Table 2**). For example, among the mollusks collected in New Caledonia, 52% were smaller than 8.8 mm, and for one family of small parasitic snails, it was estimated that as many as 80% of the species collected may have been undescribed (Bouchet *et al.*, 2002). In addition, many coral reef species are members of the cryptofauna, small organisms hidden in the cracks and crevices of reefs that are far more difficult to census than the large, spectacular corals and fishes that sit on or swim over the reef (Reaka-Kudla, 1997; Richter *et al.*, 2001).

If one cannot rely on the numbers of described species for most groups, how then does one estimate the global diversity of coral reefs? Several approaches have been tried (Bouchet, 2006; Knowlton *et al.*, 2010). Reaka-Kudla (1997) used estimated species–area relationships (number of species$=$(area)^z), the relative area of coral reefs versus tropical rainforests, and estimated diversity values for rainforests; depending on the figures for rainforest diversity used, she arrived at a low estimate of 618,000 species on coral reefs and a high estimate of 9,477,000 species. Small *et al.* (1998) also used species–area relationships; they counted everything in an aquarium derived from 5 m^2 of Bahamian coral reef and extrapolated upward to reefs globally, producing an estimate of approximately 2.6–3.2 million species. However, extrapolating from 5 m^2 to 600,000 km^2 is problematic, because the exact nature of species–area relationships in the ocean or on reefs is highly uncertain. For example, for the brachyuran crabs censused by Plaisance and colleagues (2011), the estimated number of species ranged from one-third to nearly 600 times the current number of described brachyuran crabs, depending on which end of the range of common marine estimates for z was used. At the other end of the spectrum, Bouchet (2006) estimated that there are 1.4–1.6 million marine species by extrapolating from the well-known brachyuran crab fauna of Europe, and Mora and colleagues (2011) arrived at an estimate of 2.2 million species. Taking the Bouchet (2006) estimate, if coral reefs contained approximately 35% of all marine species (a figure based on the percentage of well-known groups such as fishes that are associated with coral reefs), this would yield an estimate for coral reefs of only approximately 500,000 species (Knowlton *et al.*, 2010). Regardless of which number is closest to the truth, these calculations confirm that with the exception of species in a few showy classes and orders, the vast majority of coral reef species are as yet undescribed. (Mora *et al.* (2011) estimated that for the ocean as a whole, 91% of all species await formal description.) Indeed, most of the DNA sequences obtained from coral reef organisms do not even appear in genetic data banks (Rohwer *et al.*, 2002; Plaisance *et al.*, 2009, 2011).

Factors Responsible for Diversity Patterns

It is far easier to get a clear understanding of patterns of diversity than it is of diversity itself. Typically, well-known groups (especially corals, fishes, and large mollusks for coral reefs) are used as proxies for marine biodiversity as a whole (e.g., Roberts *et al.*, 2002). Although proxies may not accurately reflect detailed diversity patterns for other groups (Mellin *et al.*, 2011), the strongest patterns, such as the concentration of species in the Coral Triangle, are unambiguously robust. Explaining the processes underlying diversity patterns is more of a challenge, however. Explanations for diversity patterns include processes acting at global or regional scales (e.g., area and age), processes acting at local scales (e.g., disturbance from storms, extremes of temperature or salinity, and sedimentation), and patterns explained by null models that require no detailed assumptions about biology.

Biogeographers trying to explain the diversity of the Coral Triangle have long argued over competing (but not mutually exclusive) hypotheses for the region: as a center of origin (a place where species are born), a center of accumulation

Table 2 Estimates of undescribed biodiversity for several tropical marine faunas (Merrell, 1995)[a]

Site	*Taxon*	*Number of undescribed species out of the total collected in the taxon*	*Source*[b]
New Guinea	Corals	14 of 362	Werner and Allen, 1998
New Guinea	Fish	3 of 1039	Werner and Allen, 1998
New Guinea	Snails, sea slugs	310 of 564	Gosliner, 1993
Philippines (one island, multiple sites)	Snails, sea slugs	135 of 320	Gosliner, 1993
New Caledonia (295 km^2)	Eulimid snails (parasitic)	Up to 110 of 138	Bouchet *et al.*, 2002
Hawaii (one island, 6 l of coral reef sediment)	Marine polychaete worms	112 of 158	Dutch, 1988
Great Barrier Reef (two islands)	Marine flatworms (polyclads)	123 of 134	Newman and Cannon, 1994
Gulf of Mexico	Copepods (harpacticoids)	19–27 out of 29	Merrell, 1995

[a]This table has been arranged from larger to smaller body size and suggests that, as with fauna everywhere, especially in the tropics, the smaller the body size, the higher the percentage of undescribed species in the sample.
[b]Sampling effort and number of samples varied among studies.

(a place where species do not go extinct), or a center of overlap (a place where neighboring faunas, e.g., from the western Pacific and Indian Ocean, come together). An analysis of the ages and ranges of species of fishes and cowries (Bellwood and Meyer, 2009) indicated that the youngest endemics (species confined to a single small location) are not concentrated in the Coral Triangle, making the center of origin hypothesis less likely. Rather, the region seems to be one that accumulates species – but why? An analysis by Bellwood and Hughes (2001) suggests that available shallow water area within 600 km is the strongest predictor of the local species diversity, a finding broadly compatible with predictions from the theory of island biogeography. Interestingly, paleontological analyses of coral reef diversity also point to the importance of peripheral locations in generating new species (Budd and Pandolfi, 2010) and the role of large amounts of suitable habitat in supporting biodiverse communities (Renema *et al.*, 2008). However, there is no relationship between reef diversity and productivity as measured by the accumulation of calcium carbonate, in either modern or fossil reefs (Johnson *et al.*, 2008). It has also been argued that the fact that the Coral Triangle sits midway between Africa and the Americas may contribute its diversity due to geometric constraints (Bellwood *et al.*, 2005), but the shifts in the position of this major biodiversity hotspot through geologic time (Renema *et al.*, 2008) argue against the importance of this kind of purely statistical explanation, as do statistically significant differences between observed diversity patterns and the predictions of null models (Connolly *et al.*, 2003; Dornelas *et al.*, 2006).

These explanations and analyses are regional in scale, and indeed the biodiversity found in any particular location is best predicted by the regional species pool rather than by locally operating factors (Karlson *et al.*, 2004). However, local factors undoubtedly play a role as well, particularly in the context of severe habitat degradation. For example, fish diversity (Jones *et al.*, 2004) and invertebrate diversity (Idjadi and Edmunds, 2006) decline in response to losses in living coral cover or reef three dimensionality. The observed increase and then decrease in coral species diversity with depth (Huston, 1985; Karlson *et al.*, 2004) may reflect both biological factors (a shifting balance of competition and disturbance with depth) and geometric constraints (limiting corals to depths between the sea surface and depths lacking sufficient light).

Interspecific Relationships

With so many species on reefs, it is no surprise that feeding and competitive relationships are extraordinarily complex and diverse. Filter-feeding invertebrates, both on the surface of the reef and lining reef crevices (Richter *et al.*, 2001), capture and retain a high proportion of the material that moves over them, hence the description of a reef as a "wall of mouths" (Hamner *et al.*, 2007). Herbivores, carnivores, and apex predators cruise the reef for food, and deposit feeders sift through the sediments. Less visible but no less important, microbes and multicellular parasites play crucial roles in the foodweb of a healthy coral reef. One analysis of a Caribbean foodweb contained more than 3000 trophic linkages (Bascompte *et al.*, 2005). Competitive relationships are similarly diverse, and interspecific competition is the rule rather than the exception in competition for space. Studies of competitive relationships on reefs led to the discovery of competitive networks, where instead of a linear dominance hierarchy, relationships have loop structures where A beats B, and B beats C, but C beats A (Buss and Jackson, 1979), a phenomenon in part made possible by the diverse array of chemicals with specialized effects that organisms deploy in self-defense. As in other marine ecosystems, microbes undoubtedly play crucial roles on coral reefs. Modern molecular techniques allow scientists to quantify the expression of metabolic enzymes in order to understand the functional roles that microbes play (Wegley *et al.*, 2007; Dinsdale *et al.*, 2008).

Like tropical rainforests, coral reefs are renowned for the abundance of specialized mutualisms between organisms, relationships where both partners benefit. In addition to the nutritional partnerships between microbes and corals (as well as other groups, to be described in more detail in the following section), there are a host of intricate partnerships linking larger coral reef organisms. Striking mutualisms include the following:

- Coral reef fishes and shrimps that clean parasites from the tissues of fish (Floeter *et al.*, 2007).
- Partnerships between gobies and snapping shrimps, with the partners sharing the duties of watching for predators and digging, respectively (Thacker *et al.*, 2011).
- Cooperation between moray eels and groupers in hunting prey (Bshary *et al.*, 2006).
- Corals that shelter schools of fish and receive nutritional benefits from the fishes' waste products (Holbrook *et al.*, 2008).
- Crabs and shrimp that live with and defend corals from their predators (Pratchett, 2001).
- Sponge species that grow intertwined and benefit from the complementary characteristics of their partners, such as anti-predator chemicals and resistance to wave damage (Wulff, 1997).

In addition to these relationships with clear benefits to both partners, there are enormous numbers of specialized associations (e.g., shrimp that live only with specific hosts such as crinoids or sea anemones (Bruce, 1977)), where the exact nature of the relationship has never been studied. Some may be mutualistic, others commensal (the host neither suffers nor benefits), and still others parasitic, and sometimes a single relationship can move between these categories (Duris *et al.*, 2011). Whatever the nature of the relationship, the abundance of these partnerships is a clear example of how diversity begets diversity on coral reefs.

Corals as Organisms

Diversity and Evolutionary Relationships

A variety of organisms are referred to as corals. All belong to the phylum Cnidaria, which is characterized by two distinct tissue layers, the inner endoderm and the outer ectoderm, separated by a jelly-like layer called the mesoglea. The basic functional unit of bottom-dwelling cnidarians is the coral polyp,

essentially a cylinder with a central mouth cavity at one end that is surrounded by tentacles armed with stinging cells called nematocysts. All corals that can build reefs secrete a calcium carbonate skeleton, and most contain symbiotic, photosynthetic, single-celled dinoflagellate algae (zooxanthellae) in ectodermal cells that power rapid growth (*see* Coral-Dwelling Microbes *and* Coral Nutrition, Growth, Reproduction, and Recruitment). It is this relationship that results in the bathymetric distribution of reef-building corals being largely restricted to high-light environments, because photosynthetically assisted growth makes it possible for new construction to keep pace with physical and biological disturbances (the exceptions being deep-sea corals, some of which can form reefs despite their very slow growth because disturbances are relatively infrequent).

Two classes within the Cnidaria contain major reef-builders. Fire corals are in the class Hydrozoa and can produce a strong burning sensation when touched (hence their name). The other reef-builders belong to the class Anthozoa, which is divided into two groups, the Hexacorallia and the Octocorallia. Among the octocorals, there are two species that are reef-builders – the blue coral and the organ-pipe coral. True corals belong to the order Scleractinia within the hexacorals. Several other cnidarians are conspicuous on reefs but are not reef-builders: the stylasterine corals (hydrozoans, which build skeletons but lack zooxanthellae), the soft corals and gorgonians (octocorals, which are important space occupiers but lack stony skeletons), and the sea anemones, coralliomorphs, and black corals (all hexacorals, which live on reefs but lack stony skeletons). The rest of this section will focus primarily on scleractinians, as they are the most important of all the reef builders.

The understanding of scleractinian evolutionary relationships has been radically altered, thanks to the application of molecular methods. For example, a recent analysis showed that 11 of 16 families were incorrectly defined and an entire family of corals endemic to the Atlantic had been overlooked (Fukami *et al.*, 2008). Scleractinians with zooxanthellae are sometimes referred to as hermatypic (reef-building) corals, but ahermatypic (non-reef-building or azooxanthellate) corals are comparably diverse (669 and 656 species, respectively; Cairns, 1999) and are scattered across the scleractinian tree of evolutionary relationships (Kitahara *et al.*, 2010). Some scleractinian families are exclusively azooxanthellate, others exclusively zooxanthellate, and a few contain a mixture of zooxanthellate and azooxanthellate species (including a few species that facultatively contain zooxanthellae).

Coral-Dwelling Microbes

Corals cannot be simply understood as animals. Rather, they contain a complex assemblage of microbes, and for that reason the term "coral holobiont" has come into increasing use to refer to this animal-microbial entity (Rosenberg *et al.*, 2007; Bourne *et al.*, 2009). As noted (*see* Species Diversity), a single coral contains a diverse assemblage of endolithic algae, bacteria, Archaea, fungi, and viruses (Wegley *et al.*, 2007). Some are pathogens and others are beneficial, but the roles of most of these microbes are unknown.

The best understood of the microbial associates of corals are the zooxanthellae, because of their importance for coral nutrition. The understanding of the diversity and relationships among the zooxanthellae has also been revolutionized by molecular methods. What was once thought to be a single dinoflagellate species, *Symbiodinium microadriaticum*, is now recognized as a complex assemblage of distantly related groups (types A–F), each of which contains considerable genetic diversity (Baker, 2003). Different species of zooxanthellae, most of which remain identified only by numbers and associated genetic profiles, have different preferences with respect to host, light, temperature, and other environmental characteristics. Importantly, they also have different life histories (e.g., speed of colonization into coral tissues) and different sensitivities to stress, and provide different benefits to their hosts (LaJeunesse *et al.*, 2009; Jones and Berkelmans, 2010).

Coral Nutrition, Growth, Reproduction, and Recruitment

Reef-building corals depend on their zooxanthellae for nutrition via the transfer of photosynthetically fixed carbon, and eventually starve to death and die without them. Although the role of zooxanthellae in this autotrophic nutrition of corals was long emphasized, it is increasingly recognized that shallow-water, reef-building corals also have a diverse array of other nutritional options (Houlbrèque and Ferrier-Pagès, 2009). Heterotrophically they are able to capture food (detritus, bacteria, ciliates, and zooplankton) with their tentacles and mucus nets, as well as absorb dissolved nutrients directly. Heterotrophy accounts for anywhere from 0% to 66% of the fixed carbon that goes into skeleton formation and contributes from 15% to 35% of the metabolic needs of healthy corals. Feeding by corals increases tissue growth and skeleton formation, and benefits the zooxanthellae as well. The importance of heterotrophy varies as a function of the coral species and environmental conditions.

A few coral species typically exist as solitary polyps (e.g., the mushroom corals) that grow by increasing the size of the polyp. Most corals species, however, are colonial, consisting of tens or hundreds to many thousands of polyps. These corals grow by asexually budding new polyps and by extending the height of the skeletal cups (the corallites) in which the polyps sit. The shape of the resulting skeleton is species specific at the corallite level, but the overall shape of the colony is also influenced, within limits, by environmental conditions (Todd, 2008). Although corals grow relatively slowly in general (often less than 1 cm yr^{-1}), under optimal conditions, the growth rates of branching corals in the genus Acropora can exceed 10 cm yr^{-1} (Lirman *et al.*, 2010). In many cases, the polyps of the colonies are interconnected with living tissue, but in some types of growth, the polyps become disconnected because of the death of connecting tissues, even though the skeletal structure connecting the corallites remains intact. In addition, partial mortality (e.g., caused by predators) can break the physiological connection between polyps.

Some species of corals regularly form new colonies – entities that are physically and physiologically separate – through

asexual reproduction. Typically this happens through fragmentation, but asexual production of larvae and even the budding off of polyps without skeletons also occur (Yeoh and Dai, 2010). This form of reproduction can result in concentrations of apparently separate colonies that are genetically identical; the only unambiguous way to determine if this form of asexual propagation has occurred is through genetic methods. The complexity of these processes (asexual reproduction as well as partial mortality) means that it is nearly impossible to tell the age of a coral from its size (Hughes and Jackson, 1985). Also, size rather than age is a better predictor of important life history characteristics, such as when sexual reproduction occurs.

Sexual reproduction takes place in a variety of ways (Baird *et al.*, 2009). Depending on the species, a given colony may be hermaphroditic, producing both eggs and sperm, or gonochoric, producing either eggs or sperm. Fertilization typically occurs in one of two ways: through spawning, in which gametes (eggs and sperm) are released simultaneously into the water column, where fertilization takes place, or by brooding, in which sperm are released into the water column and are taken inside the maternal coral polyp to fertilize the eggs stored there. Hermaphroditic broadcast spawning is the most common mode of reproduction, and most genera exhibit only one reproductive mode. Broadcasting species tend to reproduce once a year, and the eggs and larvae typically lack zooxanthellae; not only do members of one species reproduce in synchrony, but also often multiple species do so, a phenomenon termed "mass spawning." For example, on the Great Barrier Reef, up to 30 species on a single reef may spawn within a few hours of each other. Brooders, however, often reproduce over multiple months in accordance with a lunar cycle, and their larvae often have zooxanthellae. Most broadcasting corals must cross-fertilize, but some brooding species are capable of self-fertilization.

Coral larvae (called planulae) can often settle within a few days of fertilization or release from the brooding colony, but in many species, they are capable of surviving much longer (more than 200 days in the laboratory in some cases) (Graham *et al.*, 2008). It is not easy to estimate the dispersal distances of coral larvae, but many probably settle relatively close to their parents (on the same reef or nearby reefs within a few kilometers). In addition, the dispersal of the larvae of brooders is often more limited, so that inbreeding as well as selfing may occur (e.g., Sherman, 2008), although the presence of zooxanthellae in the larvae may also facilitate occasional long-distance dispersal. Planulae prefer to settle on hard, relatively well grazed surfaces (often coralline algae) (Arnold *et al.*, 2010), after which they metamorphose into a cemented polyp.

The Disappearance of Living Corals and Coral Reefs

Reefs and People

Reefs are subject to influences that come from the land, the air, and the sea, many of them directly related to human activities (**Figure 2**). Approximately 10% of the human population lives within 100 km of coral reefs, and 91% of these people are citizens of developing countries (Donner and Potère, 2007), where environmental regulations are often lacking or poorly enforced. Coastal cities and coast-based tourism (which can dwarf local populations in terms of numbers of people and resource demands) have been growing rapidly. This results in increasing demands for potable fresh water, food and other renewable resources (including seafood and marine curios), power, and cleared land. Globally, demand for all Earth's resources has inexorably increased due to increases in both human population numbers and increasing consumption per capita.

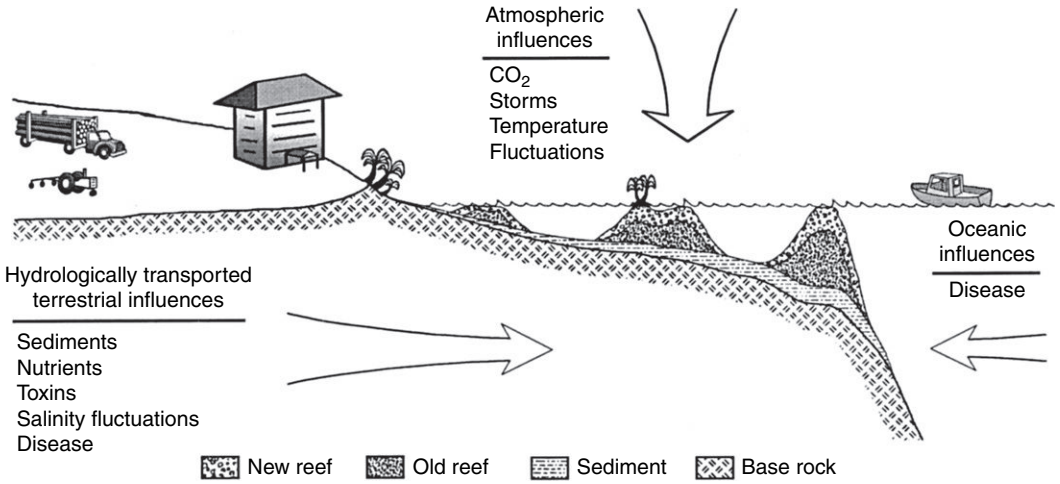


Figure 2 Coral reefs are subject to terrestrial, atmospheric, and oceanic influences. Sediments, nutrients, and toxins, released from activities such as deforestation, agriculture, and industry, are hydrologically transported to coral reefs through local rivers. CO_2 buildup in the atmosphere increases CO_2 concentrations in the ocean and alters climate patterns. Finally, diseases are circulated by ocean currents. Reefs located near human population centers are subjected to multiple stresses simultaneously and so suffer losses in diversity and cover. Adapted with permission from Wilkinson CR and Buddemeier RW (1994) *Global Climate Change and Coral Reefs: Implications for People and Reefs*. Report of the UNEP–IOC–ASPEI–IUCN Global Task Team on the Implications of Climate Change on Coral Reefs. Gland, Switzerland: IUCN.

Meeting local demands can have multiple negative effects on the local environment. Water quality degrades as strains are placed on whatever water treatment systems exist. Poorly executed construction projects and deforestation result in erosion, and the promise of jobs often leads to dredging of ports and filling in of coral reefs for airports and resorts. Meeting food needs leads to increasing fishing, as well as intensified agriculture and aquaculture. The latter two reduce pressure on wild stocks, but can result in elevated nutrients in coastal waters, the introduction of nonnative species, and contamination by pesticides and other toxic materials. Water temperatures and salinities can be locally impacted by power and desalination plants and water diversion projects.

Local populations (permanent or transient) are not the only source of stress on coral reefs. Most notably, increasing greenhouse gas emissions are causing the ocean to be both warmer (**Figure 3**) and more acidic (*see* Greenhouse Gas Emissions). This is already creating conditions unlike those of the past 420,000 years, and projections show a drastically altered ocean by the end of the century (Hoegh-Guldberg *et al.*, 2007). Global demand by people far from reefs for coral reef products also contributes to both overharvesting and destructive harvesting methods, especially the use of dynamite and cyanide to support the live fish food trade (McManus *et al.*, 1997).

Extent and Patterns of Loss of Corals and Coral Reefs

As a consequence of increasing human impacts summarized above and described in more detail later, coral reefs around the world have been declining at an alarming rate. This trend was first starkly outlined for the Caribbean, where the cover of living coral declined by approximately 80%, from approximately 50 to 10% coral cover, between 1997 and 2001 (Gardner *et al.*, 2003). Although the Pacific remains in better condition, it is on the same trajectory, with the average amount of living coral dropping approximately 1% per year since the mid-1980s (Bruno and Selig, 2007). These recent declines came on top of earlier losses, which began centuries ago in some locations (Pandolfi *et al.*, 2003). In 2008, one-third of all coral species were deemed to be at risk of extinction (Carpenter *et al.*, 2008). Because dead coral skeletons are subject to bioerosion, eventually the loss of living coral results in the loss of three-dimensional structure and even reef rock itself (Lewis, 1992; Sheppard *et al.*, 2005; Alvarez-Filip *et al.*, 2009).

Sometimes, the loss of living coral at any one site is rapid, exhibiting the dynamics of thresholds and tipping points. For example, much of the collapse of coral reef ecosystems on the north coast of Jamaica occurred within a decade (Hughes, 1994); chronic and severe overfishing underpinned the collapse of the reefs following the loss of the remaining dominant grazer, the long-spined sea urchin, *Diadema antillarum*, due to disease. Thresholds are often associated with the establishment of alternate stable states, for example, the replacement of corals by macroalgae or other space occupiers, which inhibit the recovery of the corals (Hughes *et al.*, 2010).

Major Agents of Coral Mortality

Corallivore Explosions

Fishes and a variety of invertebrates feed on corals, and a few of these corallivore species can have dramatic impacts on reefs (Rotjan and Lewis, 2008). Indeed, the first scientific awareness of threats to coral reefs came with outbreaks of the crown-of-thorns seastar, *Acanthaster planci*, in the 1960s. Individuals numbering in the hundreds of thousands were observed to

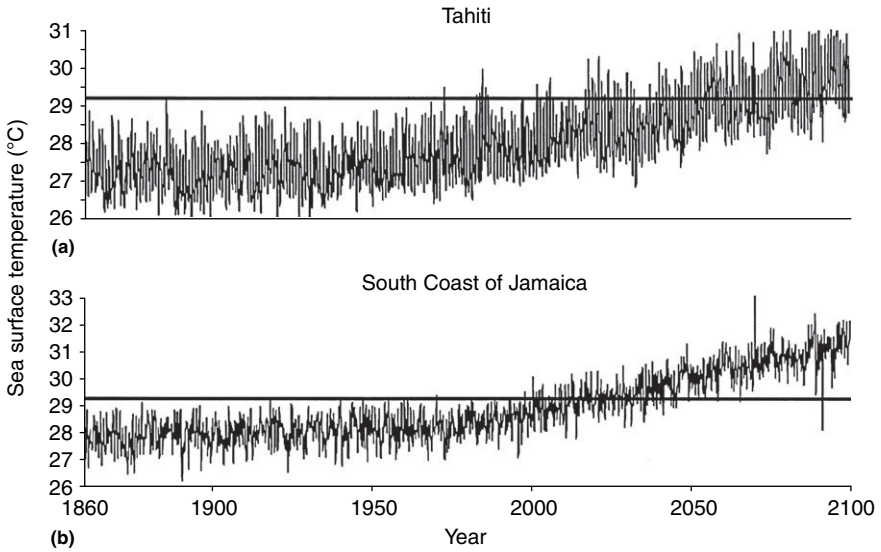


Figure 3 A model of sea surface temperatures based on greenhouse gas concentrations and El Niño Southern Oscillation events predicts temperatures will exceed normal thresholds for many reefs in the very near future. The horizontal lines indicate the temperature thresholds at which corals begin to bleach (*see* Coral Bleaching). Reproduced from Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research* 50: 839–866, with permission from CSIRO.

reduce drastically the amount of living coral cover on affected reefs. Although there was initially debate as to whether this was a natural phenomenon, scientists gradually realized that frequent massive outbreaks could not have occurred in the past, because otherwise the coral communities would have been far different than observed before the outbreaks. Controversy over the relative importance of removal of predators on crown-of-thorns versus nutrients in the water favoring survival of their larvae has persisted for decades, and there is evidence for both hypotheses. Recent work suggests that primary outbreaks are associated with flood waters, pointing to a crucial role for nutrients (Fabricius *et al.*, 2010), but marine protected areas with healthy populations of potential predators on crown-of-thorns also have reduced outbreaks (McCook *et al.*, 2010). Although crown-of-thorns have received the most attention, other coral feeders can have significant effects on coral numbers, and some, like the snail *Drupella*, can also have population outbreaks that cause extensive coral loss.



Figure 4 In 1997 and 1998, severe coral bleaching episodes were caused by dramatically elevated sea surface temperatures worldwide. A vast majority of elkhorn corals (*Acropora palmata*) on Looe Key, in the Florida Keys, bleached stark white (photograph by James W. Porter).

Algal Overgrowth

When corals die from other causes, their bare skeletons are rapidly overgrown by other organisms, often macroalgae. Algae are potentially capable of overgrowing living corals because they grow much faster than corals do. Algae also have direct, negative effects on adjacent corals because of compounds that they release. Some macroalgae release allelochemicals that directly kill corals (Rasher *et al.*, 2011). Algae also release dissolved organic carbon that can fuel the growth of bacteria, which in turn kill coral (Smith *et al.*, 2006).

Both excessive nutrients in the water (eutrophication) and overfishing (particularly the fishing of herbivores) can shift the competitive balance between these two space occupiers and favor algae over corals. The relationships linking overfishing with algal overgrowth are not always clear, however. For example, a simple prediction in this context might be that increased numbers of apex predators might reduce the number of herbivorous fish, and consequently result in more macroalgae on reefs, an example of a trophic cascade. Yet the presence of apex predators does not necessarily result in a decrease in herbivorous fish and a blooming of algae (Mumby *et al.*, 2006; Sandin *et al.*, 2008).

Coral Bleaching

If the relationship between the coral and its symbiotic zooxanthellae is disturbed, bleaching occurs. The term "bleaching" describes the condition in which the zooxanthellae die, exit, or are expelled from the coral, thus showing the stark white skeleton beneath the coral tissue (**Figure 4**). Without the symbiotic algae, corals lose their vital source of nutrition, slow their growth rates, stop reproducing, and sometimes die (Van Oppen and Lough, 2009). When environmental conditions return to normal, zooxanthellae repopulate the coral. Susceptibility to bleaching is influenced by the species of coral in question, the species of zooxanthellae it hosts, and the environmental history of the coral holobiont. It can be caused by a variety of stressors, including excessive heat or cold, low salinity, and excessive light or dark. Of these, warming oceans pose the greatest threat.

Reef-building corals typically live much closer to their upper lethal temperature (the temperature that will kill or disable them) than to their lower lethal temperature. A rise in the water temperature of only 1 °C above the normal summertime average is much more stressful physiologically than a drop of 1 °C below this value. This sensitivity to even small increases in temperature is due to the effect of elevated temperature on the photosynthesis of zooxanthellae. Temperature-induced bleaching occurs in one of two ways, either by brief exposure to moderately increased temperature or by prolonged exposure to slightly elevated temperature. Degree heating weeks (DHWs) are often used to characterize accumulated thermal stress over 12-week intervals; for example, two DHWs are equivalent to 2 weeks at 1 °C above the expected summer maximum temperature or one week at 2 °C above the expected maximum. Widespread bleaching and some mortality occurs when DHW reaches 8 °C weeks.

Over the past 20 years, there has been a dramatic increase in the number of reef provinces bleaching (**Figure 5**) and in the severity of these bleaching episodes. During the 1982–3 bleaching event, Glynn and Feingold (1992) documented up to 95% loss of corals in the Galapagos Islands. Mass mortalities have since also been reported for Australia and the Indian Ocean (Hoegh-Guldberg, 1999) and the Caribbean (Eakin *et al.*, 2010).

Disease

Coral reefs are no exception to the truism that, even in healthy ecosystems, disease is part of the natural environment. Diseases in the ocean, however, are poorly understood, and most coral reef pathogens are unknown. For instance, in a recent review, 18 coral diseases were listed as recognizable by their symptoms, but only six pathogens were identified, the first just in 1996 (Bourne *et al.*, 2009). Stony corals are not the only

reef organisms that are affected by disease. In 1982 and 1983, almost all of the black spined sea urchins, *Diadema antillarum*, in the Caribbean died from an unknown pathogen (Lessios *et al.*, 1984), which triggered the replacement of corals by seaweeds in many locations, as noted above. Sponges have also suffered from diseases, and these diseases are even less well understood (Wulff, 2007).

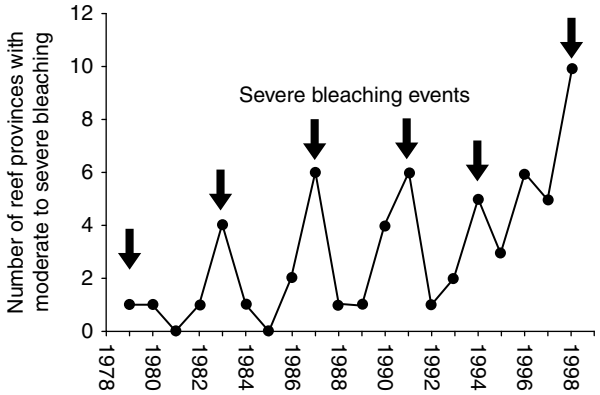
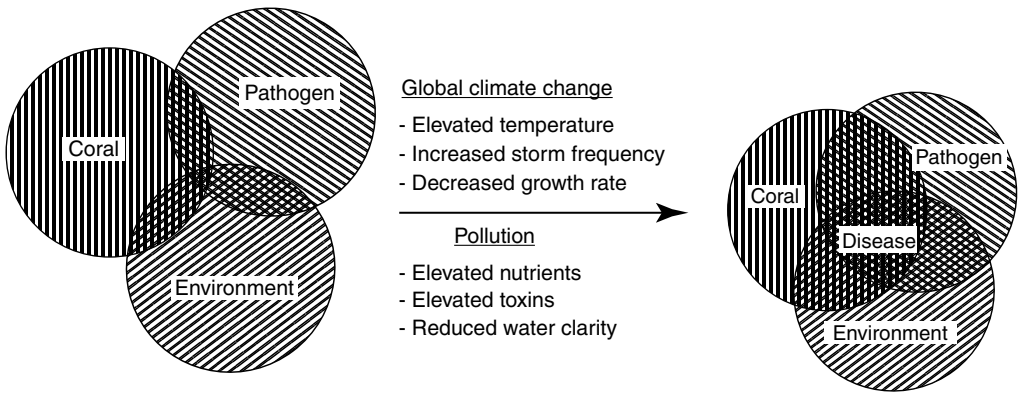


Figure 5 Recently, the frequency, intensity, and geographical extent of bleaching episodes have increased. During the strongest bleaching event to date, 1998, bleached corals were recorded for the first time in many provinces. Reproduced from Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research* 50: 839–866, with permission from CSIRO.

Although the specific agents responsible may be unclear, the ecological consequences are not. The Caribbean has been particularly hard hit, where two of the most important reef-builders, *Acropora palmata* and *A. cervicornis*, have recently been put on the US endangered species list because of catastrophic losses. Because corals grow slowly and live for decades or centuries, diseases have far-reaching impacts on coral reefs, both today and on geological time scales (Aronson and Precht, 1997). Although the lack of baseline data makes it difficult to measure with precision, it seems clear that the incidence of diseases on coral reefs has increased (Coral Disease Working Group, 2007). Paleontological evidence has demonstrated that disease outbreaks in Belize have no historical precedent over the past 5000 years (Aronson and Precht, 1997), lending credence to the idea that disease outbreaks on the present scale are a recent phenomenon.

The reasons for increasing amounts of disease on coral reefs remain debated. Traditionally, diseases have been viewed as the outcome of infection by specialized pathogens, but coral reef scientists have increasingly turned to the hypothesis that environmental stress may be weakening immune systems or causing normally relatively benign microbes to grow out of control (**Figure 6**; Coral Disease Working Group, 2007). Support for this comes from observations correlating water quality with disease severity and outbreaks of disease



(a) Low incidence of disease		**(b)** High incidence of disease
Naturally fluctuating environment	⇨ ⇨ ⇨	Anthropogenically altered environment
Healthy corals	⇨ ⇨ ⇨	Immuno-compromised corals
Low number of pathogens	⇨ ⇨ ⇨	High number of pathogens
Low pathogen virulence	⇨ ⇨ ⇨	High pathogen virulence
Few susceptible species	⇨ ⇨ ⇨	Many susceptible species
Few locations with disease	⇨ ⇨ ⇨	Many locations with disease
Low coral mortality rates	⇨ ⇨ ⇨	High coral mortality rates

Figure 6 Coral disease-stress model. Although natural background levels of disease are expected even in healthy ecosystems, a variety of stresses could lead to the suppression of the disease defense systems in coral. The consequence of reduced health would be an increase in the number of pathogenic organisms, susceptible species, outbreak locations, and mortality rates.

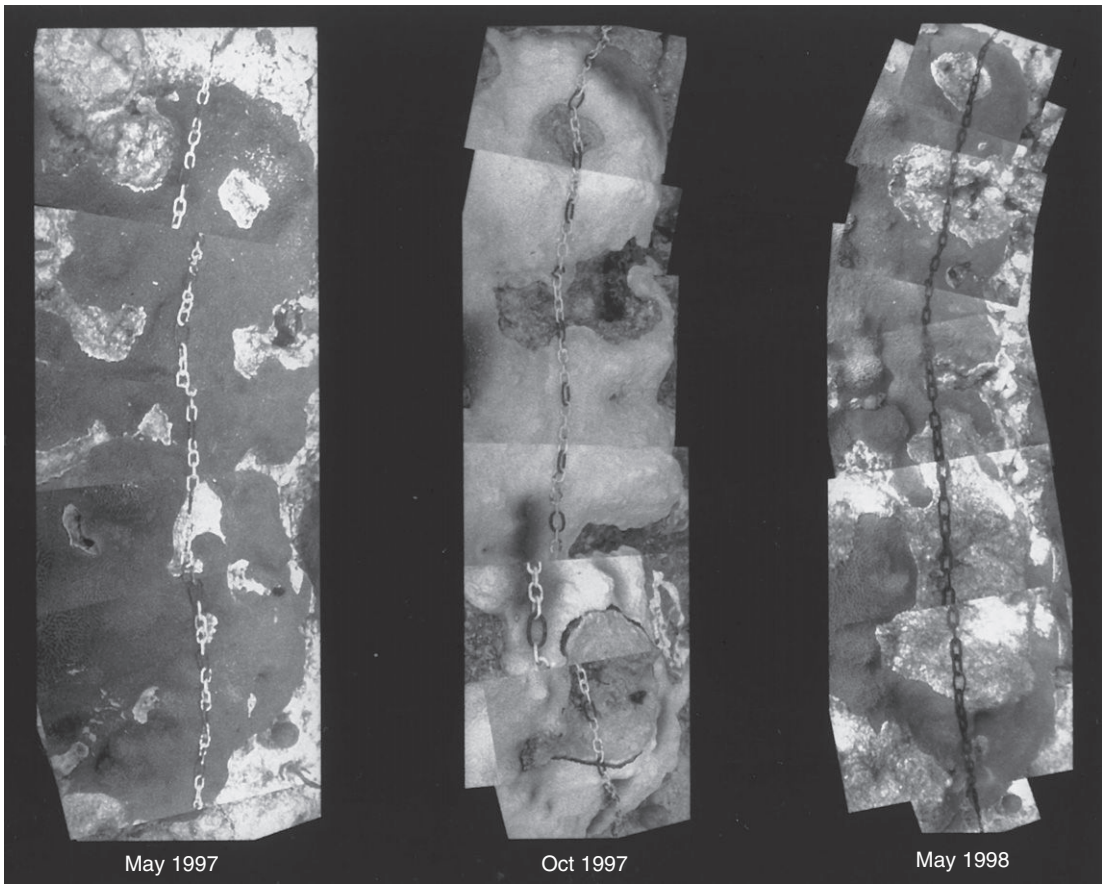


Figure 7 The potential interaction between coral bleaching and disease can be seen in this montage of images from 1997 to 1998. Healthy colonies of *Montastraea annularis* (left (May, 1997)) on Looe Key, in the Florida Keys, bleached in late summer due to elevated sea surface temperatures (middle (October, 1997)). This colony also contracted black band disease (middle, lower part of the image). By May 1998 (right), most of the colony had recovered, but the black band damaged tissue did not.

following bleaching events (**Figure 7**; Miller *et al.*, 2009). Some diseases are caused by normally terrestrial pathogens; the fungus *Aspergillus sydowii* is associated with sea fan disease and is normally found in soil, and the bacterium *Serratia marcescens* attacks elkhorn corals and is found in wastewater (Sutherland *et al.*, 2011).

Underlying Anthropogenic Causes of Coral Mortality and Reef Loss

Direct Destruction

Just as tropical rainforests are threatened by the clearing of land, direct destruction of coral reefs also occurs, although unlike the situation for forests, this is not the most important cause of reef loss globally. Nevertheless, in some locations, direct destruction of coral reefs is a major threat. For example, in southeast Asia, destructive fishing methods associated with the live fish food trade, in particular the use of dynamite or cyanide, have caused extensive reef loss (McManus *et al.*, 1997). Recovery is often lengthy, especially after dynamiting, which produces a highly unstable substratum of small coral fragments. Dredging to expand ports and even the harvesting of corals for building material can locally inflict substantial losses.

Some corals are also collected for the aquarium trade or as curios. All corals have been listed under Appendix II of the Convention on International Trade of Endangered Species (CITES) since 1990. This listing stipulates the regulation of international trade in corals but does not prohibit it, and considerable trade in corals continues (Green and Hendry, 1999).

Poor Water Quality

Because reefs thrive in clear, clean seawater, human activities that lower water quality can cause coral mortality and loss of reefs. Terrestrial runoff associated with deforestation or coastal construction delivers sediments, inorganic nutrients, organic matter, and potential pathogens onto reefs. These can have direct negative effects on the corals (e.g., smothering from sediments and lowered light availability) or indirect effects (e.g., favoring coral predators, competitors, pathogens, or bioeroders) (Fabricius, 2005). Terrestrial runoff associated with agriculture, mining, and manufacturing may also carry heavy metals and pesticides onto reefs, which directly

compromise coral health (Howard and Brown, 1984; Guzmán and Jiménez, 1992). Extreme salinities associated with freshwater diversion and powerplant operations can also locally affect corals. Of these, the most important are the impacts of sediments, nutrients, and organic matter, which threaten 22% of reefs globally (Fabricius, 2005).

Sedimentation influences coral communities through lethal and sublethal mechanisms. Although increased sedimentation causes direct mortality of corals by smothering them, most effects are sublethal. Corals remove sediments by secreting copious amounts of mucous that trap the sediments, and this constitutes a substantial energy drain. Despite this removal process, sediments tend to accumulate in depressions on large, massive colonies and cause death to those patches. Suspended sediments in the water also decrease the amount of light available for photosynthesis, resulting in lower rates of growth and reproduction, and a reduction in the maximum depth at which corals can grow. In addition, coral larvae are not able to settle on loose sediments, and sedimentation negatively affects the crustose coralline algae on which many coral larvae prefer to settle. Consequently, if a fine layer of sediments covers the reef benthos, successful recruitment drops dramatically.

Inorganic nutrients and organic matter primarily affect corals via their impact on other organisms that interact with corals. When nutrients are low, corals are able to outcompete other space occupiers, such as photosynthetic seaweeds and filter feeding sponges and tunicates, because tight nutrient cycling between the coral host and zooxanthellae affords a competitive advantage to the coral. When nutrients are added to the system, the competitive edge shifts to faster growing macroalgae and filter feeders. Coral recruitment is also reduced because algae can prevent coral larvae from settling. The growth of boring organisms is promoted, which weakens the reef structure itself and increases the probability of storm damage. Some coral diseases also are more virulent when the water has high levels of inorganic nutrients, and excess organic carbon can be even more detrimental than inorganic nutrients because it fuels the growth of bacteria (Kline *et al.*, 2006). Coral predators may also increase in number because terrestrial runoff stimulates the growth of phytoplankton on which their larvae feed; the best evidence of this comes from studies of outbreaks by the voracious coral predator *A. planci*, the crown-of-thorns seastar (Fabricius *et al.*, 2010). Finally, increased phytoplankton reduces water clarity and results in a concomitant decrease in light availability for photosynthesis by zooxanthellae.

Overfishing

Reduction in fish density and biomass is typically one of the first consequences of human activities (Jackson *et al.*, 2001). The effects of fishing (which also includes harvesting of invertebrates) stem from the destruction of the reef itself because of the methods used (reviewed earlier), but even more importantly, from the indirect consequences of the removal of fishes from the community. These latter effects on reef communities have been so pervasive that only recently have reef scientists recognized that most of the reefs they had been studying were far from pristine (Jackson, 1997). Expeditions to the remote reefs of the central Pacific have shown that the fish communities of unfished reefs typically are dominated by apex predators, which represent 50% or more of the biomass in these areas (Friedlander and DeMartini, 2002; Sandin *et al.*, 2008). The effects of fish removal can be pervasive across the reef community. A comparison across a gradient of human disturbance in the Northern Line Islands showed that with increasing human densities, apex predators disappeared, percentage cover of corals and coralline algae declined, and microbes (including potential pathogens) increased in density in the seawater (Sandin *et al.*, 2008; Dinsdale *et al.*, 2008).

Loss of herbivores is especially detrimental to reefs, because, as with increasing nutrients, it results in a shift in the competitive balance in favor of algae at the expense of corals. In most locations, the effects of overfishing on coral–algal competitive relationships is much greater than the effects of elevated nutrients. The effects of overfishing can be complex, as herbivores differ in terms of their impact on algal communities (e.g., Burkepile and Hay, 2008), and sometimes the impacts of herbivores have not been anticipated. For example, after a caging experiment that resulted in extensive algal growth was ended, the algae were largely consumed by batfish, a previously largely unstudied member of the fish community (Bellwood *et al.*, 2006). Other fishes also play important and sometimes complex roles. For example, removal of certain predators can contribute to outbreaks of corallivores, such as crown-of-thorns starfish.

Invasive Species

Humans routinely move organisms from place to place, either intentionally (for aquaculture or aquaria) or unintentionally (e.g., in ballast water). Most fail to prosper in their new locations, and invasive species until recently have been largely ignored as a threat to coral reefs. This has changed, however, particularly with the explosive proliferation of lionfish, native to the Indo-Pacific, throughout the Caribbean. A single lionfish on a patch reef can consume nearly 80% of the arriving fish recruits in 5 weeks (Albins and Hixon, 2008), so that their huge numbers greatly exacerbate the effects of overfishing in the region, which may itself have contributed to the initial spread of the invaders. Trophic cascades involving the invasion of lionfish, the loss of herbivores that they prey on, followed by the expansion of algae have already been documented on deep reefs, which were previously less affected by other local anthropogenic stressors (Lesser and Slattery, 2011). Other invasive species, for example, seaweeds introduced in the context of aquaculture, have also caused problems on reefs (Conklin and Smith, 2005).

Greenhouse Gas Emissions

Greenhouse gas emissions have two primary effects, rising temperatures and falling pH, and the consequences of both of these for coral reefs are dire (Hoegh-Guldberg *et al.*, 2007). The most conspicuous consequences of increased CO_2 concentrations to date have been the increase in coral bleaching and coral disease associated with rising temperatures, both discussed earlier. However, coral reefs are also affected in other

ways by rising temperatures and will increasingly be affected by more acidic ocean water.

Rising temperatures have two important effects on coral reefs in addition to their direct effects on coral physiology. First, the intensity of major storms, such as hurricanes, is expected to increase with increasing temperatures. These storms cause direct physical destruction of corals by increased wave action and scouring. Many coral reefs lie squarely within the two belts flanking the equator that are characterized by frequent and intense tropical storms, and healthy coral reefs recover relatively quickly from these disturbances. However, reefs that are impacted by other stressors lack resilience, and a major storm can set the stage for continued declines (Hughes, 1994). Second, and similarly, healthy coral reefs can usually keep up with rising sea levels associated with rising temperatures through the upward growth of coral colonies, but compromised reefs may "drown."

The detrimental effect of lower pH on the ability of corals to secrete their skeletons is now an area of active research, although the threat posed by ocean acidification was recognized only relatively recently (Kleypas *et al.*, 1999). Skeletal secretion by newly settled larvae can be strongly disrupted (Cohen *et al.*, 2009), and in some species, an entire colony can lose its skeleton altogether (Fine and Tchernov, 2007). Over the next century, grossly elevated atmospheric CO_2 concentrations are expected to reduce the ability of corals to calcify; Kleypas *et al.* (1999) argue convincingly that this has already begun (**Figure 8**).

Environments today that naturally have low pH and the response of fossil reefs to past environmental change give a window into what reefs may look like in the future. Studies of reefs with CO_2 seeps in Papua New Guinea exhibiting values of pH ranging from 8.3 to less than 7.7 suggest that CO_2 concentrations of 750 parts per million (ppm; 750 ppm is the level predicted for the end of the century under many business-as-usual scenarios) will cause declines in coral diversity, coral recruitment, and the abundance of structurally complex coral species (Fabricius *et al.*, 2011). For pH values less than 7.7 (equivalent to CO_2 concentrations of more than 1000 ppm), there was no coral reef growth. The fossil record also provides cause for extreme concern, as the four most recent reef crises, including two of the five mass extinctions since the evolution of complex life, occurred during periods of ocean acidification combined with rapid global warming (Pandolfi *et al.*, 2011).

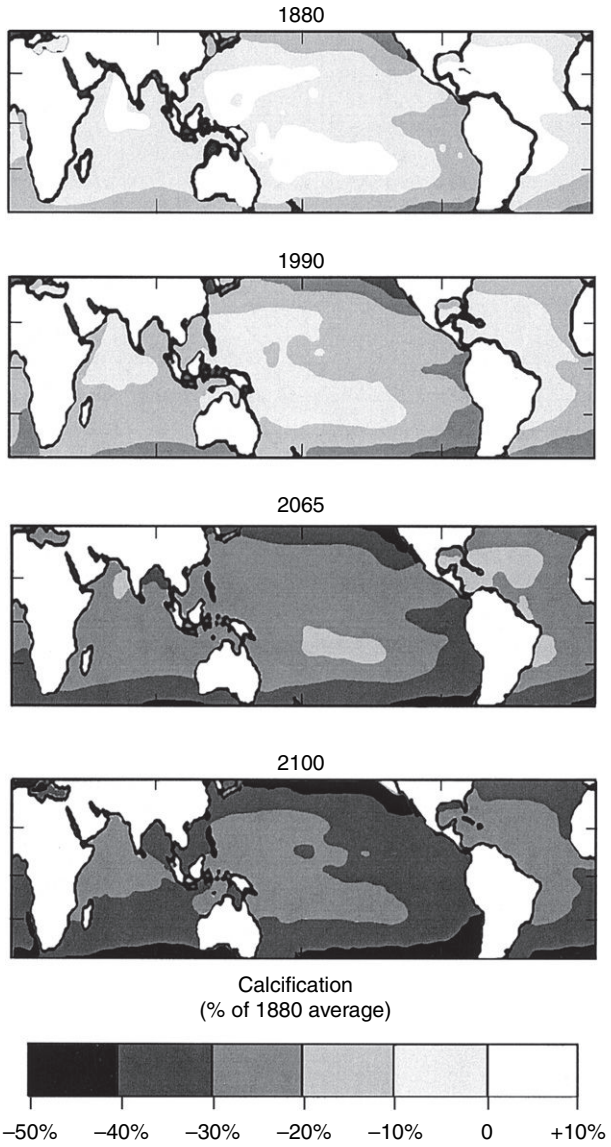


Figure 8 Projected changes in reef calcification rates are depicted as a percentage of conditions from 1880. This model suggests that oceanic conditions in the year 2100 will be substantially less optimal for coral growth than in the nineteenth century. Reprinted from Kleypas JA, Buddemeier RW, Archer D, Gattuso J-P, Langdon C, and Opdyke BN (1999) Geochemical consequences of increased atmospheric carbon dioxide on coral reefs. *Science* 284: 118–120, with permission from American Association for the Advancement of Science.

Consequences of Loss of Corals and Coral Reefs

Species Loss

Both the International Union for Conservation of Nature and Natural Resources (IUCN) and CITES define extinction as occurring when a species is not definitely located in the wild during the past 50 years. With this strict definition, and in the complete absence of monitoring efforts at the appropriate temporal and spatial scales, extinctions in the marine environment in general and on coral reefs in particular are almost impossible to prove. An example of this kind of difficulty can be seen in the announcement of the extinction of an eastern Pacific coral species due to severe El Niño conditions, and an almost immediate retraction when it was subsequently rediscovered alive (Glynn and Feingold, 1992). Carlton (1993), in his review of modern marine invertebrate extinctions, included only one tropical species in his list, the Indo-Pacific mangrove periwinkle, *Littoraria flammea*, which was last seen in the mid-1800s. It has been argued that marine species might generally be less vulnerable to extinction than terrestrial species.

However, it seems likely that some coral reef species have gone extinct without being noticed. Reaka-Kudla (1997) points out that most species on coral reefs are small and that

these smaller species also have much smaller geographic ranges, making them more prone to extinction if their habitats are destroyed. For example, some species of coral reef fish have been observed only in a single bay. In addition, the widespread destruction of coral reef habitat makes future extinctions (an extinction debt; Tilman *et al.*, 1994) more likely because of the correlation between the area of suitable habitat available and the number of species that can be sustained. In general, damaged reefs support fewer species of fishes and invertebrates than do healthy reefs (Jones *et al.*, 2004; Idjadi and Edmunds, 2006), although coral associates are probably less at risk of global extinction than corals themselves (e.g., Graham *et al.*, 2011).

Because of their vulnerability to so many different aspects of human activities, corals are of particular concern when it comes to extinction. For species that must cross-fertilize, scarcity of potential mates may hamper reproductive success. Indeed, a recent analysis (Carpenter *et al.*, 2008) suggested that as many as one-third of all coral species may be at risk of extinction due to the combined effects of global change and local stresses.

Loss of Ecosystem Goods and Services

Humans benefit from the resources, physical protection, and recreational opportunities that coral reefs provide (Hoegh-Guldberg, 1999). In one study, the total global value was estimated at nearly US$30 billion annually (Conservation International, 2008). Tourism, coastal protection, and biodiversity (all nonextractive sources of value) constituted more than 80% of the total.

Corals are valuable just as physical structures. They are used for building materials in areas where there is no viable alternative. More importantly, coral reefs reduce coastal erosion by protecting coastlines from severe storms (Sheppard *et al.*, 2005). This is particularly important in tropical waters where hurricanes and tropical storms occur frequently.

Coral reefs are crucially important for local economies. The reefs support valuable fisheries for local consumption and for the aquarium trade. Throughout the Caribbean and Indo-Pacific, local diets derive nearly 60% of their intake of protein from these reefs. The life cycles of many commercially important fish and shellfish are dependent on the presence of healthy mangrove swamps, coral reefs, seagrass beds, and coastal lagoons. A multibillion dollar tourism industry is supported by tropical coral reefs and is a critical source of income, particularly for small island nations with few alternative resources to exploit. Even in developed countries, the financial benefits of healthy reefs are substantial; the Great Barrier Reef is estimated to bring AUD5.5 billion and more than 53,000 jobs to the Australian economy (McCook *et al.*, 2010).

Finally, pharmaceutical companies have discovered naturally occurring bioactive compounds among the organisms found on coral reefs: antitumor compounds have been found in the mucus of corals, anti-inflammatory agents have been isolated from soft corals, and coral has successfully been used as a bone substitute in reconstructive surgeries. Cone snails represent a classic example; their toxins have many potential and actual uses (e.g., the pain reliever Prialt derived from the toxin of *Conus magus*), but most to date remain unexplored. Many cone snails are found on or near coral reefs and are themselves highly threatened; more than 20% of cone snails have ranges of less than 3500 km^2, and people threaten 50% or more of the ranges of 69% of cone snail species (Chivian *et al.*, 2003).

Coral Reef Conservation

Local Management and Restoration

Traditionally, conservation has operated through the principle of protecting individual species from extinction. On reefs, this approach has been applied to a limited number of commercially harvested species, notably sharks and species considered at risk from the aquarium trade such as seahorses. In the case of the former, outright bans have been implemented; for example, shark finning is now illegal in Palau and the Bahamas. The rationale is strictly economic – these species are worth more as sights to be seen than fins to be consumed. A number of threatened aquarium fishes are increasingly captive bred because of concerns about overharvesting; examples include anemone fishes, cardinal fishes, and seahorses, as well as some corals.

Increasingly, however, conservation has turned to the protection of entire ecosystems (so-called ecosystem-based management), a trend that is particularly welcome for coral reefs given the impossibility of protecting huge numbers of species one by one. For coral reefs, this involves reducing local stressors to increase the resilience of reefs, so that they can recover from disturbances whose global causes (e.g., warming temperatures) cannot be locally managed directly (Hughes *et al.*, 2010).

The two primary local stressors that require tackling are poor water quality and overfishing. Marine protected areas are the primary way that overfishing has been addressed, in part because tropical fisheries are so diverse that management of individual species is not practical. Sometimes this involves modification of fisheries practices – for example, bans on fishing parrot fishes, not because parrot fishes are themselves endangered in most cases (the highly prized bumphead parrot fish being an exception in this regard), but because they play a crucial role in keeping seaweeds under control. In developed countries, no-take marine protected areas have been especially highlighted. The Great Barrier Reef is the gold standard in this regard, with approximately 30% of the reef in no-take status. However, most coral reefs lack effective protection (Mora *et al.*, 2006). In developing countries, greater emphasis is given to incorporating more traditional management methods, such as rotational closures, as they are more likely to be accepted by local communities (Cinner and Aswani, 2007).

Data are slowly emerging that marine protected areas eventually work, especially when they are reasonably large and are well enforced. Fish respond most rapidly to protection, but corals also start to increase with time in some cases (Knowlton and Jackson, 2008; Selig and Bruno, 2010). The mechanisms are not always entirely clear, but often involve improvement in reef resilience, for example, the ability of reefs to recover from bleaching events or severe storms (Hughes *et al.*, 2010).

The ability of reefs to resist threats is also improved; for example, in the Great Barrier Reef, protected areas are less vulnerable to population explosions of crown-of-thorns seastars (McCook *et al.*, 2010), and in the Northern Line Islands, protection appears to reduce the prevalence of coral disease (Sandin *et al.*, 2008; Dinsdale *et al.*, 2008). Marine protected areas work even better when coupled with other management measures, such as efforts to improve water quality.

The restoration of damaged reefs is also an option (e.g., Lirman *et al.*, 2010). Methods typically involve growing coral fragments or recruits and placing them on damaged reefs. Although this can be effective in some circumstances, restoration is doomed if the factors that originally killed the corals are still present.

Dealing with Greenhouse Gas Emissions

Climate change models predict that tropical sea surface temperatures will continue to rise and pH will continue to fall. If these scenarios are correct, then bleaching will be more frequent, more prolonged, and more lethal, and corals will find it harder and harder to lay down their skeletons (Hoegh-Guldberg *et al.*, 2007). These predictions are not for the distant future, but for the near future, only a few decades away. The coral holobiont does have some ability to adapt to these threats (e.g., some corals and zooxanthellae are more heat tolerant), but the potential for adaptation is limited, particularly given the rapid rate of global change. Although coastal zone management practices are critical in protecting the well-being of coral reefs in the near term, especially near population centers, over this century, global climate change and how humans mitigate this anthropogenic stress will determine the long-term survival of the most diverse marine environment on Earth.

See also: Crustaceans. Eutrophication and Oligotrophication. Grazing, Effects of. Invertebrates, Marine, Overview. Marine Ecosystems. Plankton, Status and Role of

References

Albins MA and Hixon MA (2008) Invasive Indo-Pacific lionfish *Pterois volitans* reduce recruitment of Atlantic coral-reef fishes. *Marine Ecology Progress Series* 367: 233–238.

Alvarez-Filip L, Dulvy NK, Gill JA, Côté IM, and Watkinson AR (2009) Flattening of Caribbean coral reefs: Region-wide declines in architectural complexity. *Proceedings of the Royal Society B* 276: 3019–3025.

Arnold SN, Steneck RS, and Mumby PJ (2010) Running the gauntlet: Inhibitory effects of algal turfs on the processes of coral recruitment. *Marine Ecology Progress Series* 414: 91–105.

Aronson R and Precht W (1997) Stasis, biological disturbance and community structure of a Holocene coral reef. *Paleobiology* 23: 326–346.

Baird AH, Guest JR, and Willis BL (2009) Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annual Review of Ecology, Evolution and Systematics* 40: 551–571.

Baker AC (2003) Flexibility and specificity in coral-algal symbiosis: Diversity, ecology and biogeography of *Symbiodinium*. *Annual Review of Ecology, Evolution and Systematics* 34: 661–689.

Bascompte J, Melian CJ, and Sala E (2005) Interaction strength combinations and the overfishing of a marine food web. *Proceedings of the National Academy of Sciences USA* 102: 5443–5447.

Bellwood DR and Hughes TP (2001) Regional-scale assembly rules and biodiversity of coral reefs. *Science* 292: 1532–1534.

Bellwood DR, Hughes TP, Connolly SR, and Tanner J (2005) Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters* 8: 643–651.

Bellwood DR, Hughes TP, and Hoey AS (2006) Sleeping functional group drives coral-reef recovery. *Current Biology* 16: 2434–2439.

Bellwood DR and Meyer CP (2009) Searching for heat in a marine biodiversity hotspot. *Journal of Biogeography* 36: 569–576.

Bellwood DR and Wainwright PC (2002) The history and biogeography of fishes on coral reefs. In: Sale PF (ed.) *Coral Reef Fishes*, pp. 5–32. Amsterdam: Academic Press, Elsevier.

Bohlke J and Chapin C (1968) *Fishes of the Bahamas and Adjacent Tropical Waters*. Wynnewood, PA: Livingston.

Bohnsack JA (1979) *The Ecology of Reef Fishes on Isolated Coral Heads: An Experimental Approach with an Emphasis on Island Biogeographic Theory*. PhD Dissertation, University of Miami, Coral Gables, FL.

Bouchet P (2006) The magnitude of marine biodiversity. In: Duarte CM (ed.) *The Exploration of Marine Biodiversity: Scientific and Technological Challenges*, pp. 31–64. Bilbao, Spain: Fundación BBVA.

Bouchet P, Lozouet P, Maestrati P, and Heros V (2002) Assessing the magnitude of species richness in tropical marine environments: Exceptionally high numbers of molluscs at a New Caledonia site. *Biological Journal of the Linnean Society* 75: 421–436.

Bouchet P, Ng PKL, Largo D, and Tan SH (2009) Panglao 2004 – Investigations of the marine species richness in the Philippines. *Raffles Bulletin of Zoology Supplement* 20: 1–19.

Bourne DG, Garren M, Work TM, *et al.* (2009) Microbial disease and the coral holobiont. *Trends in Microbiology* 17: 554–562.

Bruce AJ (1977) Shrimps and prawns of coral reefs, with special reference to commensalism. In: Jones OA and Endean R (eds.) *Biology and Geology of Coral Reefs*, pp. 38–94. New York: Academic Press.

Bruno JF and Selig ER (2007) Regional decline of coral cover in the Indo-Pacific: Timing, extent, and subregional comparisons. *PLoS One* 8: e711.

Bshary R, Hohner A, Ait-el-Djoudi K, and Fricke H (2006) Interspecific communicative and coordinated hunting between groupers and giant moray eels in the Red Sea. *PLoS Biology* 4: 2393–2398.

Budd AF and Pandolfi JM (2010) Evolutionary novelty is concentrated at the edge of coral species distributions. *Science (New York, NY)* 328: 1558–1561.

Burkepile DE and Hay ME (2008) Herbivore species richness and feeding complementarity affect community structure and function on a coral reef. *Proceedings of the National Academy of Sciences USA* 105: 16201–16206.

Buss LW and Jackson JBC (1979) Competitive networks: Non-transitive competitive relationships in cryptic coral reef environments. *American Naturalist* 113: 223–234.

Butman CA and Carlton JT (1993) *Biological Diversity in Marine Systems*. Washington, DC: National Science Foundation.

Cairns SD (1999) Species richness of Recent Scleractinia. *Atoll Research Bulletin* 459: 1–46.

Carlton JT (1993) Neoextinctions of marine invertebrates. *American Zoologist* 33: 499–509.

Carpenter K, Abrar M, Aeby G, *et al.* (2008) One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* 321: 560–563.

Chivian E, Roberts CM, and Bernstein AS (2003) The threat to cone snails. *Science* 302: 391.

Cinner JE and Aswani S (2007) Integrating customary management into marine conservation. *Biological Conservation* 140: 201–216.

Coddington JA, Agnarsson I, Miller JA, Kuntner M, and Hormiga G (2009) Undersampling bias: The null hypothesis for singleton species in tropical arthropod surveys. *Journal of Animal Ecology* 78: 573–584.

Cohen AL, McCorkle DC, de Putron S, Gaetani GA, and Rose KA (2009) Morphological and compositional changes in the skeletons of new coral recruits reared in acidified seawater: Insights into the biomineralization response to ocean acidification. *Geochemistry Geophysics Geosystems* 10.http://dx.doi.org/10.1029/2009GC002411.

Conklin EJ and Smith JE (2005) Abundance and spread of the invasive red algae, *Kappaphycus* spp., in Kane'ohe Bay, Hawai'i and an experimental assessment of management options. *Biological Invasions* 7: 1029–1039.

Connolly SR, Bellwood DR, and Hughes TP (2003) Indo-Pacific biodiversity of coral reefs: Deviations from a mid-domain model. *Ecology* 84: 2178–2190.
Conservation International (2008) *Economic Values of Coral Reefs, Mangroves, and Seagrasses: A Global Compilation*. Arlington, VA: Center for Applied Biodiversity Science, Conservation International.
Coral Disease Working Group (2007) Coral disease, environmental drivers, and the balance between coral and microbial associates. *Oceanography* 20: 172–195.
Darwin CR (1842) The Structure and Distribution of Coral Reefs. Being the First Part of the Geology of the Voyage of the Beagle, under the Command of Capt. Fitzroy, R.N. during the Years 1832 to 1836. London: Smith Elder and Co. (reprinted by Kessinger Publishing).
Dinsdale EA, Pantos O, Smriga S, *et al.* (2008) Microbial ecology of four coral atolls in the Northern Line Islands. *PLoS One* 3: e1584.
Donner SD and Potère D (2007) The inequity of the global threat to coral reefs. *Bioscience* 57: 214–215.
Dornelas M, Connolly SR, and Hughes TP (2006) Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440: 80–82.
Duris Z, Horká I, Juracka PJ, Petrusek A, and Sandford F (2011) These squatters are not innocent: The evidence of parasitism in sponge-inhabiting shrimps. *PLoS One* 6: e21987.
Dutch M (1988) *A Characterization of Polychaete Assemblages on a Hawaiian Fringing Reef*. Master's Thesis, Zoology Department, University of Hawaii, Honolulu.
Eakin CM, Morgan JA, Heron SF, *et al.* (2010) Caribbean corals in crisis: Record thermal stress, bleaching, and mortality in 2005. *PLoS One* 5: e13969.
Fabricius KE (2005) Effects of terrestrial runoff on the ecology of corals and coral reefs: Review and synthesis. *Marine Pollution Bulletin* 50: 125–146.
Fabricius KE, Langdon C, Uthicke S, *et al.* (2011) Losers and winner in coral reefs acclimatized to elevated carbon dioxide concentrations. *Nature Climate Change* 1: 165–169.
Fabricius KE, Okaji K, and De'ath G (2010) Three lines of evidence to link outbreaks of the crown-of-thorns seastar *Acanthaster planci* to the release of larval food limitation. *Coral Reefs* 29: 593–605.
Fine M and Tchernov D (2007) Scleractinian coral species survive and recover from decalcification. *Science (New York, NY)* 315: 1811.
Floeter SR, Vazquez DP, and Grutter AS (2007) The macroecology of marine cleaning mutualisms. *Journal of Animal Ecology* 76: 105–111.
Friedlander AM and DeMartini EE (2002) Contrasts in density, size, and biomass of reef fishes between the northwestern and the main Hawaiian islands: The effects of fishing down apex predators. *Marine Ecology Progress Series* 230: 253–264.
Fukami H, Chen CA, Budd AF, *et al.* (2008) Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS One* 3: e3222.
Gardner TA, Côté IM, Gill JA, Grant A, and Watkinson AR (2003) Long-term region-wide declines in Caribbean corals. *Science (New York, NY)* 301: 958–960.
Gibbs PE (1971) The polychaete fauna of the Solomon Islands. *Bulletin of the British Museum (Natural History) Zoology* 21: 99–211.
Glynn PW and Feingold JS (1992) Hydrocoral species not extinct. *Science* 257: 1845.
Gosliner TM (1993) *Proceedings of the Seventh International Coral Reef Symposium*, vol. 2, pp. 702–709.
Graham EM, Baird AH, and Connolly SR (2008) Survival dynamics of scleractinian coral larvae and implications for dispersal. *Coral Reefs* 27: 529–539.
Graham NAJ, Chabanet P, Evans RD, *et al.* (2011) Extinction vulnerability of coral reef fishes. *Ecology Letters* 14: 341–348.
Grassle JF (1973) In: Jones OA and Endean R (eds.) *Biology and Geology of Coral Reefs*, pp. 247–270. New York: Academic Press.
Green EP and Hendry H (1999) Is CITES an effective tool for monitoring trade in corals? *Coral Reefs* 18: 403–407.
Guzmán HM and Jiménez CE (1992) Contamination of coral reefs by heavy metals along the Caribbean coast of Central America (Costa Rica and Panama). *Marine Pollution Bulletin* 24: 554–561.
Hamner WM, Colin PL, and Hamner PP (2007) Export-import dynamics of zooplankton on a coral reef in Palau. *Marine Ecology Progress Series* 334: 83–92.
Hebert PDN, Cywinska A, Ball SL, and deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B* 270: 313–321.
Hoegh-Guldberg O (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research* 50: 839–866.
Hoegh-Guldberg O, Mumby PJ, Hooten AJ, *et al.* (2007) Coral reefs under rapid climate change and ocean acidification. *Science* 318: 1737–1742.
Holbrook SJ, Brooks AJ, Schmitt RJ, and Stewart HL (2008) Effects of sheltering fish on growth of their coral hosts. *Marine Biology* 155: 521–530.
Houlbrèque F and Ferrier-Pagès C (2009) Heterotrophy in tropical scleractinian corals. *Biological Reviews* 84: 1–17.
Howard LS and Brown BE (1984) Heavy metals and reef corals. *Oceanography and Marine Biology Annual Reviews* 22: 195–210.
Hughes R and Gamble J (1977) A quantitative survey of the biota of intertidal soft substrata on Aldabra Atoll, Indian Ocean. *Philosophical Transactions of the Royal Society of London, Series B* 279: 324–355.
Hughes TP (1994) Catastrophies, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science* 265: 1547–1551.
Hughes TP, Graham NAJ, Jackson JBC, Mumby PJ, and Steneck RS (2010) Rising to the challenge of sustaining coral reef resilience. *Trends in Ecology and Evolution* 25: 633–642.
Hughes TP and Jackson JBC (1985) Population dynamics and life histories of foliaceous corals. *Ecological Monographs* 55: 141–166.
Huston MA (1985) Patterns of species diversity on coral reefs. *Annual Review of Ecology and Systematics* 16: 149–177.
Idjadi JA and Edmunds PJ (2006) Scleractinian corals as facilitators for other invertebrates on a Caribbean reef. *Marine Ecology Progress Series* 319: 117–127.
Jackson JBC (1984) Ecology of cryptic coral reef communities. III. Abundance and aggregation of encrusting organisms with particular reference to cheilostome Bryozoa. *Journal of Experimental Marine Biology and Ecology* 75: 37–57.
Jackson JBC (1997) Reefs since Columbus. *Coral Reefs* 16(Suppl): S23–S32.
Jackson JBC, Kirby MX, Berger WH, *et al.* (2001) Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293: 629–638.
Johnson KG, Jackson JBC, and Budd AF (2008) Caribbean reef development was independent of coral diversity over 28 million years. *Science* 319: 1521–1523.
Jones A and Berkelmans R (2010) Potential costs of acclimitization to a warmer climate: Growth of a reef coral with heat tolerant vs. sensitive symbiont types. *PLoS One* 5: e10437.
Jones GP, McCormick MI, Srinivasan M, and Eagle JV (2004) Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Sciences USA* 101: 8251–8253.
Karlson RH, Cornell HV, and Hughes TP (2004) Coral communities are regionally enriched along an oceanic biodiversity gradient. *Nature* 429: 867–870.
Kiessling W (2009) Geologic and biologic controls on the evolution of reefs. *Annual Review of Ecology, Evolution and Systematics* 40: 173–192.
Kitahara MV, Cairns SD, Stolarski J, Blair D, and Miller DJ (2010) A comprehensive phylogenetic analysis of the Scleractinia (Cnidaria, Anthozoa) based on mitochondrial CO1 sequence data. *PLoS One* 5: e11490.
Kleypas JA, Buddemeier RW, Archer D, Gattuso J-P, Langdon C, and Opdyke BN (1999) Geochemical consequences of increased atmospheric carbon dioxide on coral reefs. *Science* 284: 118–120.
Kline DI, Kuntz NM, Breitbart M, Knowlton N, and Rohwer F (2006) Role of elevated organic carbon levels and microbial activity in coral mortality. *Marine Ecology Progress Series* 314: 119–125.
Knowlton N, Brainard RE, Fisher R, *et al.* (2010) Coral reef biodiversity. In: Macintyre AD (ed.) *Life in the World's Oceans: Diversity, Distribution, and Abundance*, pp. 65–77. Oxford: Wiley-Blackwell.
Knowlton N and Jackson JBC (1994) New taxonomy and niche partitioning on coral reefs: Jack of all trades or master of some? *Trends in Ecology and Evolution* 9: 7–9.
Knowlton N and Jackson JBC (2008) Shifting baselines, local impacts, and global change on coral reefs. *PLoS Biology* 6: e54.
LaJeunesse TC, Smith RT, Finney J, and Oxenford H (2009) Outbreak and persistence of opportunistic symbiotic dinoflagellates during the 2005 Caribbean mass coral 'bleaching' event. *Proceedings of the Royal Society B* 276: 4139–4148.
Lesser MP and Slattery M (2011) Phase shift to algal dominated communities at mesophotic depths associated with lionfish (*Pterois volitans*) invasion on a Bahamian reef. *Biological Invasions* 13: 1855–1868.
Lessios HA, Robertson DR, and Cubit JD (1984) Spread of *Diadema* mass mortality through the Caribbean. *Science* 226: 335–337.
Lewis J (1992) Evidence from aerial photography of structural loss of coral reefs at Barbados, West Indies. *Coral Reefs* 21: 49–56.
Lirman D, Schopmeyer S, Manzello D, *et al.* (2011) Severe 2010 cold-water event caused unprecedented mortality to corals of the Florida reef tract and reversed previous survivorship patterns. *PLoS One* 6: e23047.
Lirman D, Thyberg T, Herlan J, *et al.* (2010) Propagation of the threatened staghorn coral *Acropora cervicornis*: Methods to minimize the impacts of fragment collection and maximize production. *Coral Reefs* 29: 729–735.

May RM (1994) Biological diversity: Differences between land and sea. *Philosophical Transactions of the Royal Society of London B* 343: 105–111.
McCloskey L (1970) The dynamics of the community associated with a marine scleractinian coral. *Internationale Revue der gesamten Hydrobiologie und Hydrographie.* 55: 13–81.
McCook LJ, Ayling T, Cappo M, *et al.* (2010) Adaptive management of the Great Barrier Reef: A globally significant demonstration of the benefits of networks of marine reserves. *Proceedings of the National Academy of Sciences* 107: 18278–18285.
McManus JW, Reyes Jr. RB, and Nañola Jr. CL (1997) Effects of some destructive fishing methods on coral cover and potential rates of recovery. *Environmental Management* 21: 69–78.
Mellin C, Delean S, Caley J, *et al.* (2011) Effectiveness of biological surrogates for predicting patterns of marine biodiversity: A global meta-analysis. *PLoS One* 6: e20141.
Merrell WJ (1995) *Understanding Marine Biodiversity: A Research Agenda for the Nation.* Washington, DC: National Academy Press.
Meyer CP, Geller JB, and Paulay G (2005) Fine scale endemism on coral reefs: Archipelagic differentiation in turbinid gastropods. *Evolution* 59: 113–125.
Miller J, Muler E, Rogers C, *et al.* (2009) Coral disease following massive bleaching in 2005 causes 60% decline in coral cover on reefs in the US Virgin Islands. *Coral Reefs* 28: 925–937.
Mora C, Andréfouët S, Costello MJ, *et al.* (2006) Coral reefs and the global network of marine protected areas. *Science* 312: 1750–1751.
Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B (2011) How many species are there on earth and in the ocean? *PLoS Biology* 9: e1001127.
Mumby PJ, Dahlgren CP, Harborne AR, *et al.* (2006) Fishing, trophic cascades, and the process of grazing on coral reefs. *Science* 311: 98–101.
Newman LJ and Cannon LRG (1994) Biodiversity of tropical polyclad flatworms from the Great Barrier Reef, Australia. *Memoirs of the Queensland Museum* 36: 159–163.
Pandolfi JM, Bradbury RH, Sala E, *et al.* (2003) Global trajectories of the long-term decline of coral reef ecosystems. *Science* 301: 955–958.
Pandolfi JM, Connolly SR, Marshall DJ, and Cohen AL (2011) Projecting coral reef futures under global warming and ocean acidification. *Science* 333: 418–422.
Peyrot-Clausade M (1983) Transplanting experiments of motile cryptofauna on coral reef flat of Tulear (Madagascar). *Thalassgraphica* 6: 27–48.
Plaisance L, Caley MJ, Brainard RE, and Knowlton N (2011) The diversity of coral reefs: What are we missing? *PLoS One* 6: e25026.
Plaisance L, Knowlton N, Paulay G, and Meyer C (2009) Reef-associated crustacean fauna: Biodiversity estimates using semi-quantitative sampling and DNA barcoding. *Coral Reefs* 28: 977–986.
Poore GBC and Wilson GDF (1993) Marine species diversity. *Nature* 361: 597–598.
Pratchett MS (2001) Influence of coral symbionts on feeding preferences of crown-of-thorns starfish *Acanthaster planci* in the western Pacific. *Marine Ecology Progress Series* 214: 111–119.
Rasher DB, Stout EP, Engel S, Kubanek J, and Hay ME (2011) Macroalgal terpenes function as allelopathic agents against reef corals. *Proceedings of the National Academy of Sciences USA* 108: 17726–17731.
Reaka-Kudla ML (1997) The global biodiversity of coral reefs: A comparison with rain forests. In: Reaka-Kudla M, Wilson D, and Wilson E (eds.) *Biodiversity II. Understanding and protecting our biological resources*, pp. 83–108. Washington, DC: Joseph Henry Press.
Renema W, Bellwood DR, Braga JC, *et al.* (2008) Hopping hotspots: Global shifts in marine biodiversity. *Science* 321: 654–657.
Richter C, Wunsch M, Rasheed M, Kotter I, and Badran MI (2001) Endoscopic exploration of Red Sea coral reefs reveals dense populations of cavity-dwelling sponges. *Nature* 413: 726–730.
Roberts CM, McClean CJ, Veron JEN, *et al.* (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295: 1280–1284.
Rohwer F, Seguritan V, Azam F, and Knowlton N (2002) Diversity and distribution of coral-associated bacteria. *Marine Ecology Progress Series* 243: 1–10.
Rosenberg E, Koren O, Reshef L, Efrony R, and Zilber-Rosenberg I (2007) The role of microorganisms in coral health, disease and evolution. *Nature Reviews Microbiology* 5: 355–362.
Rotjan RD and Lewis SM (2008) Impact of coral predators on tropical reefs. *Marine Ecology Progress Series* 367: 73–91.
Sale PF (1977) Maintenance of high diversity in coral reef fish communities. *The American Naturalist* 111: 337–359.
Sandin SA, Smith JE, DeMartini E, *et al.* (2008) Baselines and degradation of coral reefs in the Northern Line Islands. *PLoS One* 3: e1548.
Selig ER and Bruno JF (2010) A global analysis of the effectiveness of marine protected areas in preventing coral loss. *PLoS One* 5: e9278.
Sheppard C, Dixon DJ, Gourlay M, Sheppard A, and Payet R (2005) Coral mortality increases wave energy reaching shores protected by reef flats: Examples from the Seychelles. *Estuarine and Coastal Shelf Science* 64: 223–234.
Sheppard RC, Davy SK, and Pilling GM (2009) *The Biology of Coral Reefs.* Oxford, UK: Oxford University Press.
Sherman CDH (2008) Mating system variation in the hermaphroditic brooding coral, *Seriatorpora hystrix. Heredity* 100: 296–303.
Small AM, Adey WH, and Spoon D (1998) Are current estimates of coral reef biodiversity too low? The view through the window of a microcosm. *Atoll Research Bulletin* 458: 1–20.
Smith JE, Shaw M, Edwards RA, *et al.* (2006) Indirect effects of algae on corals: Algae-mediated, microbe-induced coral mortality. *Ecology Letters* 9: 835–845.
Spalding MD, Ravilious C, and Green EP (2001) *World Atlas of Coral Reefs.* Berkeley, CA: University of California Press.
Starck W (1968) A list of fishes of Alligator Reef, Florida with comments on the nature of the Florida reef fish fauna. *Undersea Biology* 1: 1–40.
Sutherland KP, Shaban S, Joyner JL, Porter JW, and Lipp EK (2011) Human pathogen shown to cause disease in the threatened elkhorn coral *Acropora palmata. PLoS One* 6: e23468.
Szabo BJ, Tracey JI, and Goter ER (1985) Ages of subsurface stratigraphic intervals in the quaternary of Enewetak Atoll, Marshall Islands. *Quaternary Research* 23: 54–61.
Thacker CE, Thompson AR, and Roje DM (2011) Phylogeny and evolution of Indo-Pacific shrimp-associated gobies (Gobiiformes: Gobiidae). *Molecular Phylogenetics and Evolution* 59: 168–176.
Tilman D, May RM, Lehman CL, and Nowak MA (1994) Habitat destruction and the extinction debt. *Nature* 371: 65–66.
Todd PA (2008) Morphological plasticity in scleractinian corals. *Biological Reviews* 83: 315–337.
Van Oppen MJH and Lough JM (eds.) (2009) *Coral Bleaching: Patterns, Processes, Causes and Consequences.* Berlin: Springer Verlag.
Veron J (1985) *Proceedings of the Fifth International Coral Reef Congress*, vol. 4, pp. 83–88.
Wegley L, Edwards R, Rodriguez-Brito B, Liu H, and Rohwer F (2007) Metagenomic analysis of the microbial community associated with the coral *Porites astreoides. Environmental Microbiology* 9: 2707–2719.
Wells J (1973) New and old scleractinian corals from Jamaica. *Bulletin of Marine Science* 23: 16–58.
Werner T and Allen G (eds.) (1998) *A Rapid Biodiversity Assessment of the Coral Reefs of Milne Bay Province, Papua New Guinea.* Washington, DC: Conservation International.
Wood R (1999) *Reef Evolution.* Oxford, UK: Oxford University Press.
Wulff JL (1997) Mutualisms among species of coral reef sponges. *Ecology* 78: 146–159.
Wulff JL (2007) Disease prevalence and population density over time in three common Caribbean coral reef sponge species. *Journal of the Marine Biological Association of the United Kingdom* 87: 1715–1720.
Yeoh S-R and Dai C-F (2010) The production of sexual and asexual larvae within single broods of the scleractinian coral, *Pocillopora damicornis. Marine Biology* 157: 351–359.

Countryside Biogeography

Chase D Mendenhall, Stanford University, Stanford, CA, USA
Carrie V Kappel, University of California Santa Barbara, Santa Barbara, CA, USA
Paul R Ehrlich, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biogeochemical processes The chemical, physical, geological, and biological processes that govern the composition of the biosphere.
Biogeography The study of life over space and time.
Countryside The increasing fraction of the Earth's surfaces whose ecosystem qualities are strongly influenced by humanity.
Countryside biogeography The study of life over space and time in areas of substantial human activity.
Human-dominated ecosystem An ecosystem where human presence and activities strongly influence living and nonliving ecosystem components.

Introduction

The word "countryside" generally elicits images of Anglo-American cottage communities nestled within a landscape of woodlands, pastures, and gardens. As landscape painting developed in seventeenth-century Dutch art like Bruegel's *The Harvesters* (*c.* 1565), the imagined countryside was epitomized in European representations such as Haytley's *The Brockman Family at Beachborough* (*c.* 1774–1776, **Figure 1**). These artistic genres framed popular western notions of idealized "landscape" and "countryside" as the harmonious coexistence of people, agriculture, and nature. However, "countryside" as an ecological term must be separated from idyllic depictions, cultural nostalgia, and European stately settings to include all of the Earth's surfaces whose ecosystem qualities are strongly influenced by humanity (Daily, 2001). Countryside encompasses different biomes such as tropical forests, arctic tundra, and even underwater seascapes. With virtually all ecosystems impacted to a greater or lesser degree by human activity, it is clear that the planet is human dominated (Ellis *et al.*, 2010; Halpern *et al.*, 2008).

Yet human-dominated ecosystems are rarely continuous regions of urban settlement, agriculture, or fishing grounds; instead, the majority of these ecosystems comprise mosaics of roads and buildings mixed with gardens, agricultural plots, areas grazed or harvested, and remnants of native habitats. The majority of the planet can therefore be classified as countryside, with varying proportions of nearly natural, seminatural, and heavily human-affected habitats. Given the trajectory of human population growth and increasing consumption per capita, the fate of biodiversity will be decided in the countryside.

A fundamental tool for understanding the interplay between living and nonliving ecosystem components is biogeography, the union of biology, geography, and history. In a broader sense, biogeography studies biological variation in space and time at genetic, morphological, physiological, behavioral, demographic, and ecological levels.



Figure 1 Left: *The Harvesters* (1565), Pieter Bruegel the Elder. Oil on wood. *Metropolitan Museum of Art.* http://www.metmuseum.org/toah/works-of-art/19.164 Right: *The Brockman family at Beachborough* (c. 1744–1746), Edward Haytley. Oil on canvas. *National Gallery of Victoria.* http://www.ngv.vic.gov.au/col/work/4021

Clarification of Concepts: From Islands to Countryside

Island Biogeography

Countryside biogeography was developed to update and expand the fields in which it is rooted, primarily MacArthur and Wilson's theory of island biogeography (MacArthur and Wilson, 1967). Island biogeography has made a substantial contribution to theory and practice of ecology, evolution, and conservation biology.

Despite its name, island biogeography has been applied beyond island systems where areas are naturally or anthropogenically isolated. For instance, climatic conditions can isolate biological communities at the peaks of mountains, geologic variation can create "islands" of different soil chemistry, and succession often isolates floral (and therefore faunal) communities. Agriculture transforms landscapes creating natural habitat fragments or "islands" in a "sea" of human-dominated habitats. Marine environments experience variable temperature zones and light regimes based on depth, segregating habitats into underwater "islands" of habitat for some species.

Island biogeography has been critical to the development of species–area relationships and community assembly theories, but has proven insufficient for real-world applications in many situations. Island biogeography and its corollaries immediately introduce biases deep within their theoretical and empirical approaches by dismissing the "sea" or "matrix" as inhospitable habitat for the organisms in question. The consequence has been a focused effort on studying life in areas seemingly untouched by humans. This *a priori* human valuation of ecosystem components, especially as a binary of hospitable/inhospitable habitats, ignores the fact that nearly every cubic meter of the biosphere has been impacted by humans (Vitousek *et al.*, 1997).

The assumption that biodiversity only exists as a shadow of its former state outside of the so-called native habitat was only challenged once biologists began surveying and encountering organisms that were not human inquilines persisting in the long term in areas that might have been thought hostile to them. The existence of populations of numerous species completing their life cycles – possibly in modified or novel states – in the "sea" of agricultural plots and seminatural habitats, to say nothing of urban areas, has called into question the generality of island biogeography and demanded a closer examination of expanding human-dominated landscapes as a possible refuge for a substantial portion of biodiversity (Ricketts, 2001). Island biogeography is a useful theoretical construct that in some situations can predict species equilibria with reasonable accuracy (Bush and Whittaker, 1993) and can help us to better understand patterns, processes, and mechanisms operating in the countryside, and has served as a jumping off point for further investigation.

Vicariance Biogeography

Countryside biogeography is also partially founded on principles from vicariance biogeography. Generally applied to less dispersal-prone taxa, vicariance biogeography attempts to explain the distribution of biodiversity through time and space by examining isolation and reconnection through geologic and climatic phenomena (Nelson, 1974).

Biological variation must now cope with separation and reconnection of habitats by human-created physical barriers (e.g., paved roads, crop rotations, deforestation, and bottom trawling), corridors (e.g., bridges, causeways, and hedgerows), chemical barriers (e.g., spraying for pesticides, use of fertilizers, and endocrine-disrupting pollution in waterways), stochastic barriers (e.g., nuclear disasters, oceanic oil spills, and fires), and biological barriers (e.g., regional and local extinctions of prey and mutualism partners, species invasions, population manipulation through feeding and harvesting, and human-induced changes in animal behaviors). Many barriers for some are bridges for others and permit range expansion and connection between populations.

Conservation Biogeography

The state of the environment has rallied biogeographers to redirect the relevance, aims, and values of modern biogeography toward the conservation of biodiversity by developing the field of conservation biogeography. Conservation biogeography has been defined as "the application of biogeographical principles, theories, and analyses, being those concerned with the distributional dynamics of taxa individually and collectively, to problems concerning the conservation of biodiversity" (Whittaker *et al.*, 2005, p. 4). Conservation biogeography is the product of biogeography and conservation biology and can be generalized as a contemporary field that uses biogeography to calculate and reduce rates of species extinction mainly through the design and placement of natural reserves with the goal of conserving species diversity.

The importance biogeography has played in conservation biology and the environmental movement is undeniable, but scientific visibility of biogeography has waned with the development of new fields. The goal of conservation biogeography is to revitalize the discipline and reaffirm the importance of biogeographical principles to current and future conservation efforts, while attempting to strengthen biogeographic theory and our overall knowledge of life on the planet. Conservation biogeography is different from countryside biogeography by explicitly championing a political goal – preventing species extinctions.

Countryside Biogeography

Countryside biogeography is the study of the distribution of biological variation over space and time in human-dominated ecosystems. Countryside refers to the growing fraction of terrestrial and marine habitats whose ecosystem qualities are strongly influenced by humanity. On land, the countryside includes terrain dedicated to agriculture, rural, and urban settlements, and the pockets of seminatural habitats in a variety of sizes and configurations existing in human-dominated landscapes (**Figure 2**). In water, virtually no area remains pristine, making marine ecosystems, coastal zones, and freshwater systems critical areas of study in countryside biogeography (**Figure 2**).

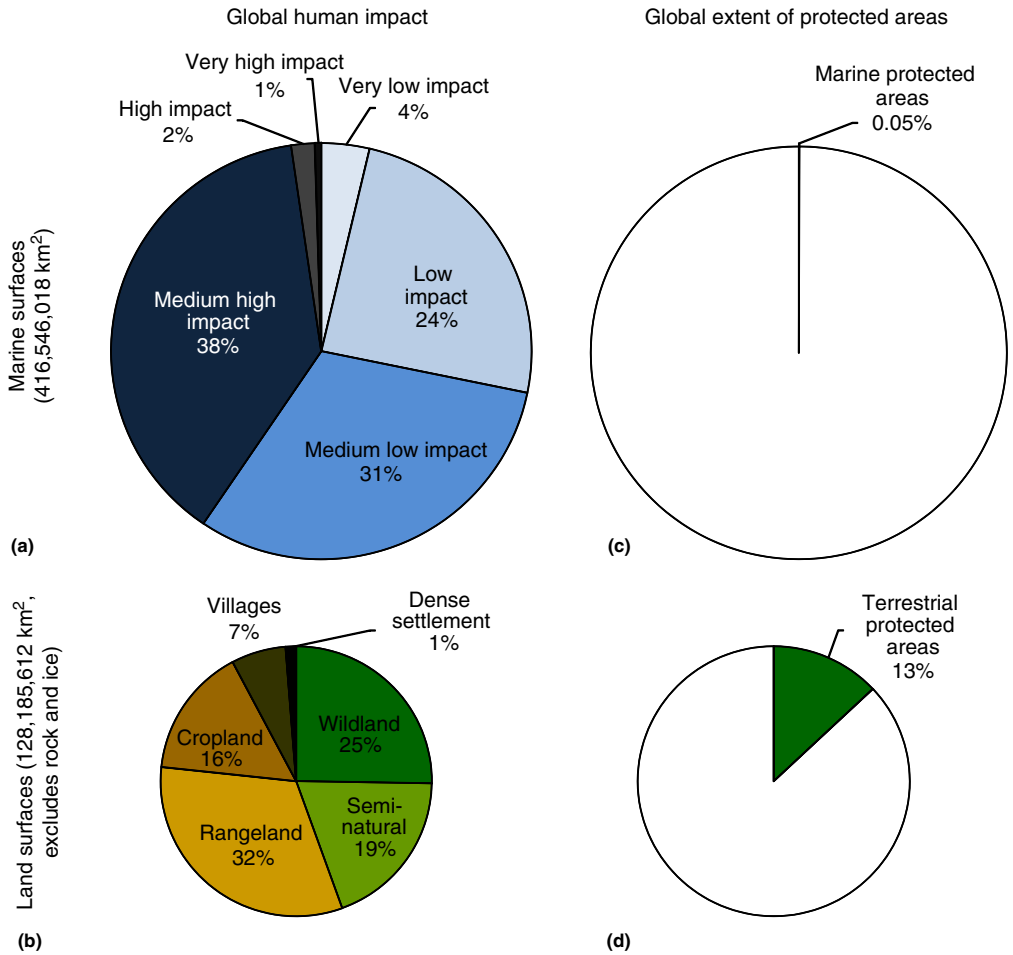


Figure 2 The planet is human-dominated and poorly protected. Global human impact on marine (Halpern, *et al.*, 2008) (a) and terrestrial (Ellis, *et al.*, 2010) (b) ecosystems is extensive, with approximately 75% of the planet's ecosystems strongly influenced by humanity. Moreover the extent of global protected areas is scant (Coad, *et al.*, 2009; Agardy, 1997; Sheppard, 2010) (c) and (d). Relative pie chart sizes for marine and land surfaces are proportional to their global surface areas. Impact score in (a) was calculated by summing 17 anthropogenic drivers including fishing, nutrient input, pollution, transport, and invasive species. Semi-natural label in (b) includes residential woodlands and inhabited treeless lands. Wildland in (b) includes protected areas and treeless lands. Protected areas in (c) and (d) include designated protected areas, many of whose management and effectiveness is unknown.

Countryside biogeography is explicitly inclusive of all ecosystem components (including native habitats, novel ecosystems, domestic and feral plants and animals, human populations, and the services provided by nature to human beings, including quality of life). It focuses on forecasting changes in biogeography in tandem with changes in land use, human populations, and climate.

In the wake of humanity's expansion and domination of the planet, wilderness is fading away on land and in the water, making the objectives of countryside biogeography progressively more relevant. Assuming that human impacts intensify as currently projected, countryside biogeography will be pivotal to answering three increasingly urgent questions. First, what sorts of ecosystems and their components (e.g., species, populations, and biological communities) will exist over the coming decades and centuries? Second, what sorts of ecosystems and their qualities (e.g., ecosystem functions and services) do we need and want – and how do we decide? And third, how, and to what extent, can we achieve our goals? Answering these questions will depend heavily on countryside biogeographic investigation of the following related questions:

- What species or functional traits confer advantages or disadvantages for coping with human-induced environmental change?
- How will human-induced changes in biogeography alter the functioning of ecosystems in the future?
- What role will countryside biota and human-dominated ecosystems play in the provisioning of ecosystem services that support human well-being?
- Are existing tools (ecological theory, remote sensing, interdisciplinary models for forecasting environmental change) up to the task of predicting changes in biodiversity and ecosystem functioning?
- What practical measures can be taken to increase the capacity of human-dominated ecosystems to sustain biodiversity and provide ecosystem services to society?

To investigate these questions, countryside biogeographers utilize empirical and theoretical frameworks in concert with established ecological and evolutionary theory. Mainly through comparison of human-modified habitats with near-pristine baselines, countryside biogeography has revealed that high amounts of biodiversity can thrive in some human-dominated habitats (e.g., Gibson *et al.*, 2011; Mendenhall, 2011; Mendenhall *et al.*, 2011). These findings, along with new theories, provide scope to harmonize human activities with the conservation of biodiversity (Daily, 2001; Pereira *et al.*, 2004; Pereira and Daily, 2006; Mendenhall *et al.*, 2012). Empirical studies in countryside biogeography have shown that biodiversity is affected by a range of human activities, by the configuration and sizes of native and human-modified habitat patches, and by the physical structure of native and seminative ecosystem elements (Mendenhall *et al.*, 2011; Sekercioglu *et al.*, 2007; Daily *et al.*, 2001, 2003; Horner-Devine *et al.*, 2003; Daily, 2001; Ricketts *et al.*, 2001; Ranganathan *et al.*, 2008).

Furthermore, countryside biogeography ideally aims to measure changes in ecosystems beyond species numeration and diversity counts by quantifying ecosystem functioning (Mayfield *et al.*, 2006) and the services provided by nature to society. A growing number of studies are identifying the explicit links between changes in biodiversity, ecosystem functioning, and human well-being, but the complexity inherent in these connections hinders fast or extensive studies (Balvanera *et al.*, 2006; Winfree and Kremen, 2009). Nonetheless, tools that aid in the valuation of these connections in human-dominated ecosystems are being developed and tested around the world as countryside biogeography gains momentum (Kareiva *et al.*, 2011; Watts *et al.*, 2009).

Human-Domination of the Earth's Ecosystems

Human Effects on Ecosystems

Human effects on ecosystems have altered the course of life. Though far too complex a topic to cover in this article the authors summarize the human effects that directly influence biogeography. The main negative actions that change biogeography are habitat alteration, toxification of ecosystems, species invasions, overexploitation of natural resources, and changes in global biogeochemistry. These actions are driven by the expansion and intensification of growing human enterprises, overpopulation and continued population growth, overconsumption, and the use of environmentally damaging technologies. The extent and potency of these actions and drivers are to a greater or lesser degree the determining influence on living and nonliving components of the biosphere, including the most pristine places on the planet.

The Myth of the Untouched Ecosystems

Ecosystems are complex units made of perpetually changing components. Historically, ecosystem change generally happened slowly with few changes observable within a human lifespan. Owing to human activity, the current pace of change appears to be at its greatest in 65 million years – since Earth was struck by an asteroid, extinguishing the dinosaurs and many other organisms. Humans have created a world where rapidly occurring changes cover expansive areas. Despite this, many environmental and political institutions continue to attempt to preserve or restore "nature" in a mythical untouched state.

Currently 3–4% of the planet has been set aside in terrestrial ($\sim$13% of all terrestrial surfaces) or marine-protected areas ($<$0.05% of the ocean's surface area) (Agardy, 1997; Coad *et al.*, 2009; Sheppard, 2010). Similarly, "wildlands" have shrunk to 25% of the total terrestrial cover and less than 4% of the world's oceans have "very low human impact scores" (**Figure 2**). Protected areas, wilderness, and relatively pristine waters have enormous biological, economic, and cultural value, but their extent is dwarfed by the expanding portions of the planet dedicated to the human enterprise of food production and impacted by our other activities. Moreover, protected areas, "wildlands," and "pristine" waters are not exempt from human activity and many, on closer examination, are and have been human-dominated for many decades or centuries.

Human-Dominated Ecosystems

Human-domination implies that the presence and activities of people exert a determining influence on living and nonliving components of an ecosystem, not control over nature. In an attempt to propose a threshold for "human-dominance" we classify an ecosystem as human-dominated if the influence by human activities is obvious and measurable to scientists upon preliminary observation. Beyond this threshold there are many gradients of human activity intensity, and studies that span and account for such gradients will be useful to inform how biodiversity responds to various degrees of human dominance.

Human-Dominated Landscapes

The most obvious characteristics of human-dominated landscapes are the physical alterations (**Figure 3**). In terrestrial ecosystems the majority of human-dominated landscapes are

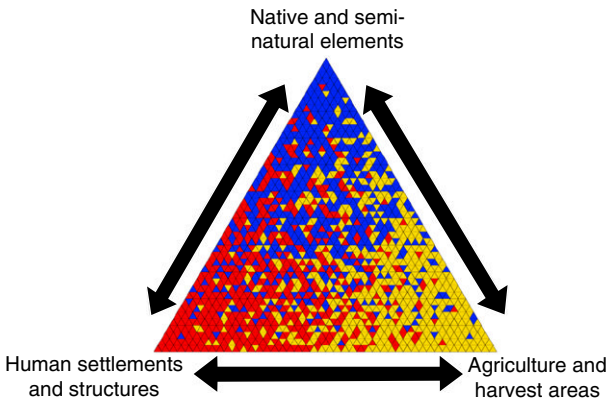


Figure 3 A gradient of major elements in human-dominated ecosystems: native and seminatural elements, agriculture or harvest areas, and human settlements and structures. Very rarely do ecosystems contain absolute regions of a single element type and most are usually heterogeneous, especially when less simplified ecosystem element types are used.

significantly influenced by agriculture. Almost half of the terrestrial land surface is dedicated to human food production. Rangelands account for two-thirds of that and croplands a third (**Figure 2**). As complex variables of population size, available arable land, per capita consumption, maximum yields, and food distribution interact, they result in agricultural intensification and annual net expansion of land dedicated to food production (Tilman *et al.*, 2001; Green *et al.*, 2005; Perfecto *et al.*, 2009).

The second major land cover to consider in human-dominated landscapes is human settlements. The global extent of towns and cities is relatively small, with rural villages covering 6.5% of all land surfaces and urban settlements 1.25% (**Figure 2**), but their influence on ecosystems is extremely potent. Human structures create obvious changes. The construction of Tokyo, Delhi, and New York City obviously displaced many local natural populations and created habitat for rats, pigeons, cockroaches, and other human inquilines, but scattered buildings and homes in rural areas are also important. Human infrastructure also fragments or bridges wild populations with roads and telecommunication lines acting as filters that restrict or enhance the movements of species, populations, and individuals. Moreover, human settlements bring exploitation of resources, invasive species, pollution, and changes in regional climate.

Human-Dominated Seascapes

Humans occupy and influence marine ecosystems differently than landscapes. To most people, the ocean is one big biological soup. We have only fuzzy ideas about how many marine species are out there, where they can be found, and how our actions affect them. This is exacerbated by the fact that marine biology and biogeography are severely understudied compared with their terrestrial counterparts, so that no clear estimate of global marine diversity yet exists (Sala and Knowlton, 2006), though the Census of Marine Life aims to fill this fundamental knowledge gap (Yarincik, 2005). Despite these shortcomings we do know that ecological principles apply underwater and that human influences are changing marine biogeography. Moreover, human activity has reached the farthest corners of the sea and many marine ecosystems (especially coastal ones) are dominated by humans. Characteristics of a human-dominated seascape include eutrophication and pollution, physical alteration of habitat, changes in other biogeochemical processes (e.g., climate change and ocean acidification), and changes in marine food webs because of direct harvesting of sea life.

Human Influence and Biodiversity

The Sixth Extinction Wave

Extinction is the greatest process altering modern biogeography. The number of species today is thought to be the largest in the 3.7-billion-year history of life (Rohde and Muller, 2005). In recent history, negative human impacts on ecosystems have ushered in the "sixth extinction wave" – a mass extinction event whose magnitude cannot yet be calculated because it is ongoing and the first to occur during the existence of *Homo sapiens* (Jablonski and Chaloner, 1994; Barnosky *et al.*, 2011; Ceballos *et al.*, 2010) Current species extinction rates are estimated to be 100–1000 times their prehuman levels in well-known taxa (Pimm *et al.*, 1995; Millennium Ecosystem Assessment, 2005). The magnitude of species extinction levels is greatly affecting biogeography and being played out in human-dominated ecosystems. The degree to which biogeography is altered by extinction is extensive, especially when considering that species extinction is only the cusp of even greater losses of biodiversity at the levels of population diversity, genetic diversity, community structure, and ecosystem function.

Species Richness and Diversity

Biogeography fundamentally relies on the categorization and ordering of biodiversity into units that can be measured, compared, and contrasted. Biologists categorize and order the natural world to make meaning of biogeography. The concept of a species is the oldest, most popular, and most debated biological unit to describe biodiversity and biogeography. Since Aristotle through to the Middle Ages, the Enlightenment, and up to the present, the concept of a species has undergone countless transformations (Wilkins, 2009). In fact, an exact universal working definition is likely to never exist because of the growing number of organisms and populations that continue to defy all proposed taxonomic, reproductive, or genetic definitions. The "species problem" is still stimulating to some, but ultimately trivial – with most fields of biology tailoring definitions for purposes of human understanding on a taxon-by-taxon case and using species or other taxonomic units when it makes sense (Ehrlich, 1997; Kelt and Brown, 2000). Unfortunately, this has been largely miscommunicated, leading to the misconception that biodiversity and species diversity are equivalents. This mishap has been described as "biology's biggest blunder" and has stunted clear understanding of the state of biodiversity and biogeography (Woodwell, 2010).

Despite the artificial difficulties with the species concept, quantifying biogeography using species as units is always the first tool ecologists turn to because of its efficiency – counting species is far faster and easier than identifying and tallying populations or piecing together interaction webs. The species concept is a particularly useful unit for countryside biogeographers who are largely interested in comparing biodiversity's responses to human influences. The first measure is usually species richness, or the number of different species in a given area, which is later followed by an index of species diversity, which incorporates species and their abundances within an area. Species richness and diversity are deployed first in countryside biogeography, but, ideally, countryside biogeography also measures biodiversity using ecosystem components that determine function.

Population Diversity

Along with species, a key unit of biodiversity is the population. Populations are not without their own definitional problems, but generally present a more meaningful measure

of biodiversity than species (Ceballos and Ehrlich, 2002). Populations are generally defined as geographical entities within a species, which are distinguishable either ecologically or genetically (Ehrlich and Daily, 1993). That a species cannot go extinct until all of its populations have gone extinct indicates that there is considerable amplification between these levels of biodiversity. Rough estimations assess population extinction rates to be 200–4000 times greater than global species extinction rates (Hughes *et al.*, 1997). The extinction of populations affects genetic diversity. Genetic diversity is tightly linked to population diversity and refers to the total number of genetic characteristics in the genetic makeup of a species. Population connectivity, size, and distribution greatly affect genetic diversity and a population's potential to adapt to changing environments.

The extinction of populations is sometimes countered by the creation of new populations by human influences. Through human-induced physical, chemical, stochastic, climatic, and biological alterations to ecosystems, populations can be split either ecologically or genetically; they can be manipulated by feeding, hunting, facilitating or preventing disease, or reintroduction; or their movement to new locations can be intentionally or unintentionally assisted. Despite the difficulty in quantifying and monitoring population diversity over space and time, its use as a biological measure is generally more informative for understanding ecosystem change than species diversity.

Community Assemblage and Structure

Species-specific population responses to human activities determine regional species pools or ecological communities. Individual communities within a region are frequently a limited subset of species from regional pools. The notion that communities are limited in their membership was first proposed by Elton using three criteria that excluded an organism or allowed it entry into a community (Elton, 1927; Roughgarden and Diamond, 1986). The first criterion is an organism's physiological ability to persist in the physical environment where a community occurs. These physical components include climate, nutrients, and light; but under the framework of countryside biogeography they can also include infrastructure, noise pollution, chemical compounds, artificial light, and human land use. The second important factor limiting community membership is dispersal capability. The failure to reach potentially suitable habitat limits community membership, but human activity can both facilitate and hinder the movement of organisms across all spatial scales. The third element limiting community membership is species interactions, or community structure. Community structure defines organisms' roles within a community, their ability to persist, and thus their contributions to vital functions.

Classic examples of community structures are food webs, where producer–consumer and predator–prey communities are linked. Existing community structures have the potential to exclude organisms; for example, a predator may be excluded from a community if the prey sources it depends on are absent because of physical, dispersal, or other community structure limitations. By contrast, interactions can also boost community membership. For example, plants and corals that enhance the structural complexity of an area and create habitats can permit other species to join a community. The same structural enhancement can be achieved with crop selection and configuration, such as areca nut, coffee, or fruit and timber plantations (Ranganathan *et al.*, 2008; Mendenhall *et al.*, 2011; Daily *et al.*, 2003). Finally, organisms interact with abiotic factors in their environments that are critical to many ecosystem processes, such as water and nutrient cycling. Countryside biogeography includes community properties (assemblage and structure) and ecosystem function as meaningful biological measures to examine the distribution of life in the context of human activity and land uses.

Ecosystem Function, Ecosystem Services, and Natural Capital

It is the collection of organisms, their environment, and all of their interactions with living and nonliving factors that result in the innumerable processes that determine the functioning of an ecosystem. In an attempt to quantify the most valuable ecosystem components and functions for human life support, ecologists, economists, and policymakers use the term "ecosystem services" to refer to the myriad of goods and services provided by species interactions within an ecosystem. From a countryside biogeography perspective, ecosystem services are the positive products and processes that derive from organisms interacting with each other and the nonliving components of their environments that either directly or indirectly benefit humankind. Species diversity, population diversity, community assemblage, and community structure are the cascading levels of biodiversity that create and maintain the elements of the biosphere necessary to support human life.

Many interactions can be harmful to human life and well-being. For example, forest remnants provide roosting habitat for the common vampire bat (*Desmodus rotundus*) that may spread disease and harm livestock; increased contact with African elephants (*Loxodonota africana*) has resulted in loss of crops, property, and human life; predator populations feed on livestock; and approximately 2.5 million people are envenomed and 100,000 are killed annually by poisonous snakes worldwide (Chippaux, 1998; Graham *et al.*, 2005; Wohlgenant, 1994; Hoare, 1999).

Despite these ecosystem "disservices," most would agree that mass extinction will have an irreversible and overwhelming net negative effect on humanity. The dangers of biodiversity loss include, but are not limited to, reduced production of food, fuel, and fiber; exacerbated climate change and more severe weather events; increased disease; runaway pollution of air and water; and the elimination of the many aesthetic, cultural, and psychological facets that make life worth living (Millennium Ecosystem Assessment, 2005; Bratman *et al.*, 2012). Countryside biogeography and related fields that explicitly consider human components of ecosystems are appropriate tools for measuring the trade-offs between human enterprises, ecosystem disservices, and ecosystem services to inform land use and policy decisions in a sustainable manner (Kareiva *et al.*, 2011).

Countryside biogeography is linked to economics, politics, conflict, and human needs and desires because all these occur within the biosphere. To articulate the value and functioning

of the biosphere's life support systems, the concept of natural capital has gained popularity in today's global capitalist economy (Costanza *et al.*, 1997; Daily, 1997; Wackernagel, 1999). Thinking of ecosystem elements as natural capital from which flow beneficial goods and services has increased understanding of the value of biodiversity and ecosystem processes in political and economic spheres where decisions affecting human activity and, correspondingly, biogeography are made (Clark *et al.*, 2001). In concert with ethical arguments for preserving nature, the valuation of ecosystem services and natural capital has become a key strategy for solving the environmental crisis.

Case Studies of Countryside Biogeography

Early Countryside Biogeography

The first studies to quantify biodiversity in human-dominated ecosystems, beyond understanding the ecology of pests or domesticated plant and animals, were conducted in farmland communities in Australia, Europe, and Mexico. As early as 1925, ecologists were anecdotally describing biodiversity in human-dominated landscapes (Barnard, 1924), but it was not until the early 1990s that rigorous countryside biogeography studies were being conducted and communicated.

Early and influential studies were from Australia. One was a survey in southeast Queensland, Australia conducted over a period of 5 years in diverse coastal farmland. The authors performed 20 standardized surveys and encountered 131 bird species in a landscape "substantially cleared for agriculture, other rural use and residential occupation" (Leach and Hines, 1993). Though the authors do not provide context for comparison with historical frequencies of avifauna, they conclude the study by stating that the human-dominated "district supports a rich diversity of bird species" (Leach and Hines, 1993). Another Australian countryside biogeography study took place in northeast New South Wales in an intensively grazed landscape with less than 20% remaining tree cover, approximately 85% of which was on private land (Barrett *et al.*, 1994). This study surveyed birds, but kept careful track of woodland area size and species observations and linked observed species richness with available woodland habitat. This linking of observations with habitat allowed for an early example of a concave decay function between human land use and a biodiversity index (**Figures 4** and **5**).

Simultaneously across the globe, English ecologists were turning their attentions to the conservation value of semi-natural habitats in the farmed countryside. The European approach to countryside biogeography brought with it political organizations concerned with conservation of biodiversity, food security, and the preservation of a traditional way of life. In the early 1990s the long established Joint Nature Conservation Committee (JNCC) and affiliates began an initiative to engage scientists, policymakers, agriculturists, and the general public over the English and Scottish countryside's biological, cultural, and political importance. The concern put forth by members and affiliates of the JNCC concentrated on the disappearance of European traditional, low-intensity farmland. Numerous countryside biogeography reports and findings

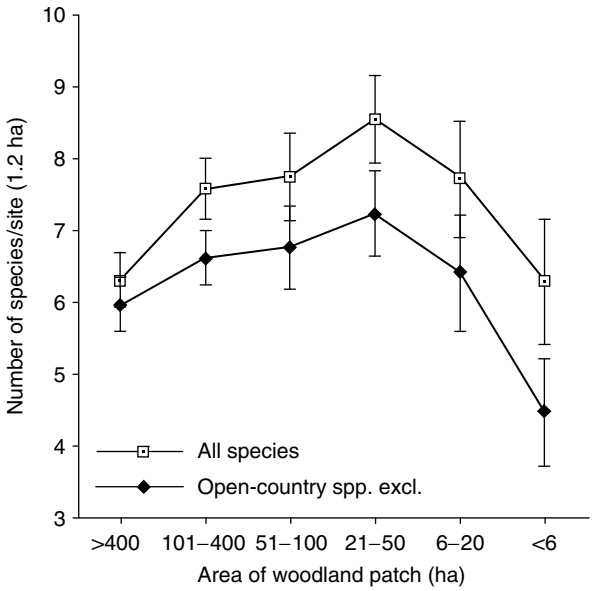


Figure 4 The effect of area of woodland patch size on the mean number (±SE) of bird species (within 1.2 ha transects, 294 in total) in a rural landscape of New South Wales. Woodland patches were divided into six size categories: >400 ha ($n=77$), 101–400 ha ($n=79$), 51–100 ($n=39$), 21–50 ha ($n=43$), 6–20 ha ($n=35$), and <6 ha ($n=21$). There was a significant difference in the mean number of species in different sized patches (Kruskal–Wallis, $H=11.61$, df$=5$, $p<0.05$). This difference was less significant when open-country species were excluded ($H=10.62$, df$=5$, $p<0.06$). Reproduced from Barrett G, Ford H, and Recher H (1994) Conservation of woodland birds in a fragmented rural landscape. *Pacific Conservation Biology* 1(3): 245–256, with permission from Surrey Beatty and Sons Pty Ltd.

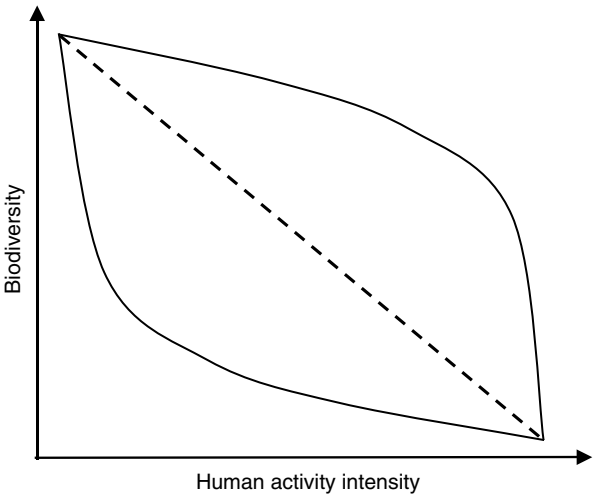


Figure 5 Three hypothetical decay functions are presented between human activity intensity and biodiversity. Top curve is a concave decay function, dashed line is a linear decay function, and bottom curve is a convex decay function. The amount of biodiversity retained under each decay function varies markedly, with the concave decay function (top curve) permitting high biodiversity retention under low and moderate human activity intensities.

from European conservation organizations and various scientific meetings and forums were produced in the early 1990s (Bignal and McCracken, 1996).

In addition to work in Australia and Europe, early studies of biogeography in tropical countryside were being carried out in Mexico. For example, Eric Mellink surveyed birds and rodents in the San Luis Potosi Plateau countryside of Mexico and found agricultural systems to support substantial biodiversity (Mellink, 1991a, b). Shortly thereafter a detailed study of Los Tuxtlas in southern Veracruz, Mexico, was conducted by Alejandro Estrada and his team. They demonstrated that many rainforest species persisted on farmland in biodiverse tropical regions. In two studies by Estrada, the team surveyed for bats and nonflying mammals in 45 and 35 variably sized forest fragments, respectively, and 20 agricultural plots of cocoa, coffee, citrus, allspice, and mixed crops. They estimated that 77% of the bat and 44% of the nonflying mammal species were found outside of tropical rainforest habitat and were at least partially using farmland (Estrada *et al.*, 1993, 1994). These and other earlier comparisons of species richness across variable human land-use intensities in the Mexican countryside set the stage for later works examining biodiversity in other biomes.

Soon after the publication of studies and reviews by Australian, European, and Mexican scientists, countryside biogeography studies started peppering conservation biology literature and were included in many conservation-focused management schemes (Benton, 2003). Moreover, in retrospect, many authors acknowledged that studies seemingly unrelated to the field of countryside biogeography had actually been conducted in human-dominated ecosystems. For example, all of Brown's comparative studies on the population biology of desert rodents took place on sites that had been grazed at some point in time (personal comm., Brown, 1989). Many of these studies were conducted, analyzed, and communicated under the paradigm that biodiversity would generally be greatly reduced at the earliest threshold of human influences (**Figure 5**). Early countryside biogeography studies began to cautiously test the assumption that human-dominated ecosystems were inhospitable. At the same time, scientists began to realize that the human landscape provided interesting settings to investigate ecosystem resilience, resistance, and evolution. Finally, early countryside biogeography studies began to face the reality that not all biodiversity can, or will, persist in the coming decades.

Tropical Countryside Biogeography in Southern Costa Rica

Not until 1997 was "countryside biogeography" formally coined, despite the long-standing interest in Europe and elsewhere. It first appeared in an article in the proceedings of the 1997 Forum on Biodiversity (Daily, 1997). In this article, preliminary results of the tropical countryside biogeography of birds in Coto Brus, Costa Rica were presented.

The in-depth studies performed in the Coto Brus Valley of southern Costa Rica are among the investigations that have had the largest contribution to understanding biodiversity in human-dominated ecosystems (Gardner *et al.*, 2009). In Coto Brus, work has spanned six major taxonomic classes of mammals, birds, reptiles, amphibians, arthropods, and herbaceous plants (Daily *et al.*, 2001, 2003; Ricketts *et al.*, 2001; Hughes *et al.*, 2002; Horner-Devine *et al.*, 2003; Mayfield and Daily, 2005; Brosi *et al.*, 2007; Santos-Barrera *et al.*, 2008).The results from these works have influenced ecological theory and land-use practices across biomes – helping develop countryside biogeography into a fully-fledged field of inquiry to be tested and replicated in other regions.

Because of the depth in which it has been examined there, Coto Brus Valley has become a flagship study system for countryside biogeography. Coto Brus Valley has a long history that includes human influences stretching back at least 3000 years. First it was heavily influenced by indigenous populations until the end of the fifteenth century, at which point it was left to regenerate until the mid-twentieth century (Clement and Horn, 2001). In the early 1950s, 45 Italian immigrants colonized Coto Brus Valley under the leadership of Vito Sansonetti (Sansonetti, 1995). After that, rapid colonization and deforestation ensued in the area. By 2000 the population of Coto Brus had grown to over 43,500 inhabitants and the tree cover in the study area dropped to approximately 30–40% (Daily *et al.*, 2001; Sansonetti, 1995; Mendenhall *et al.*, 2011).

The landscape now mostly consists of small-scale farms that raise cattle and maintain sun coffee, banana, plantain, pineapple, bean, corn, sugar cane, or yucca. Farms are rarely dedicated to a single activity and the hilly topography prevents the use of industrialized methods. A single farm rarely exceeds 5 ha and trees persist on nearly all properties in clusters or patches that weave across the landscape along valleys, hilltops, roads, and property boundaries. Within the landscape is a small Biological Reserve that is home to the Organization for Tropical Studies Las Cruces Biological Station. The reserve protects approximately 260 ha of primary and secondary premontane tropical wet forest (Holdridge, 1979). Work in Coto Brus Valley has concluded that substantial floral and faunal biodiversity exists on farmland at considerable distances from the forest reserve (**Figure 6**).

Like most initial studies of biogeography, the first countryside biogeography in Coto Brus Valley used bird surveys. In 1995 and 1996 surveys conducted in eight variably sized forest fragments and 13 open habitats on farmland agreed with previous work conducted in Australia, Europe, and Mexico that substantial avifaunal diversity was capable of using agricultural habitats in tropical countryside landscapes (Daily *et al.*, 2001). These findings sparked a second, expanded study in 1998 that systematically surveyed birds at 53 independent locations – corroborating the results (Hughes *et al.*, 2002). When 27 surveys of terrestrial mammals returned the same patterns, the message was clear: a high amount of biodiversity persists in the human-dominated landscape of Coto Brus. Additional studies have provided clear evidence that these patterns were not restricted to birds and mammals (**Figure 5**) and continue to shed light on the responses of many taxa at multiple levels of biodiversity, including an optimistic decay function where biodiversity and humanity coexist (**Figure 6**).

Biogeography in Ancient Tropical Countryside

Since the groundwork of the 1990s, numerous studies have been published that show high biodiversity persisting in

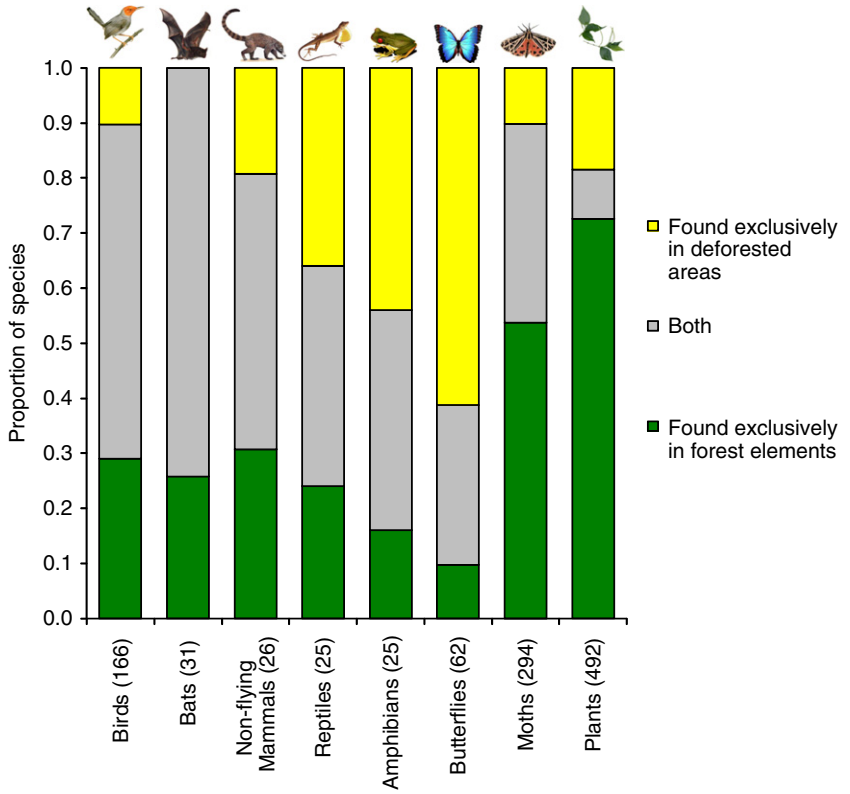


Figure 6 Distribution of species diversity in the Costa Rican countryside compiled from eight studies in Coto Brus, Costa Rica. Forest elements include the Las Cruces Forest Reserve and surrounding forest fragments embedded in the Costa Rican Countryside. Number of species representing each taxon in parenthesis. Images are of a single species represented in each study. Bird data from Mendenhall CD, Sekercioglu CH, Oviedo Brenes F, Ehrlich PR, and Daily GC (2011) Predictive model for sustaining biodiversity in tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America* 108(39): 16313–16316, and Mendenhall C, Ehrlich P, and Daily G (2012) Improving estimates of biodiversity loss. *Biological Conservation* 151(1): 32–34. http://www.sciencedirect.com/science/article/pii/S0006320712001115; bat data from Mendenhall, *et al.*, in progress; non-flying mammal data from Daily GC, Ceballos G, Pacheco J, Suzan G, and Sanchez-Azofeifa A (2003) Countryside biogeography of Neotropical mammals: Conservation opportunities in agricultural landscapes of Costa Rica. *Conservation Biology* 17(6): 1814–1826; reptile and amphibian data from Santos-Barrera G, Pacheco J, Mendoza-Quijano F, *et al.* (2008) Diversity, natural history and conservation of amphibians and reptiles from the San Vito region, southwestern Costa Rica. *Revista de Biología Tropical* 56(2): 755–778, and Mendenhall *et al.*, in progress; butterfly data from Horner-Devine MC, Daily GC, Ehrlich PR, and Boggs CL (2003) Countryside biogeography of tropical butterflies. *Conservation Biology* 17(1): 168–177; moth data from Ricketts TH, Daily GC, Ehrlich PR, and Fay J (2001) Countryside biogeography of moths in a fragmented landscape: Biodiversity in native and agricultural habitats. *Conservation Biology* 15(2): 378, and plant data from Mayfield M and Daily G (2005) Countryside biogeography of Neotropical herbaceous and shrubby plants. *Ecological Applications* 15(2): 423–439, and Mendenhall and Oviedo Brenes, in progress.

countryside, including Indonesia (Waltert *et al.*, 2004), Africa (Naidoo, 2004), and India (Chundawat *et al.*, 1999). Despite the increasing number of studies, the question of extinction debts remains (Tilman *et al.*, 1994). Extinction debt is the extinction of a species because the remaining area of habitat after modification is not sufficient to allow for long-term persistence. In the context of Coto Brus Valley, it is possible that observed patterns of high biodiversity in farmland may be populations slowly going extinct over generations because the remaining area of habitat cannot support them long term. The uncertainty associated with extinction debts is of great concern for the future of biodiversity in human-dominated ecosystems. One way of getting a glimpse into the future is to look to ecosystems that underwent modification in the past.

A 2008 study of the ancient tropical countryside of Karnataka, India, has yielded insight into extinction debt in human-dominated ecosystems. The Indian countryside of Karnataka has been continuously cultivated for more than 2000 years. The landscape is typical of many tropical regions with small-scale farms spatially intermixed with tree elements of various shapes and sizes. The agriculture supports extensive areca nut palm, cashew, rice, and peanut plantations – with land cover configuration remaining relatively constant for more than 200 years (Ranganathan *et al.*, 2008). The study again centered on ecology's best-known taxon, birds, and surveyed for species in 30 independent locations equally distributed among land-use types. From the 114 observed species, 96% were observed outside of intact forest remnants – with some of India's most charismatic bird species, like the Great Hornbill (*Buceros bicornis*) and Malabar Gray Hornbill (*Ocyceros griseus*), found in areca nut palm plantations. Specifically, a strong relationship with increased structural complexity of human-dominated habitats was

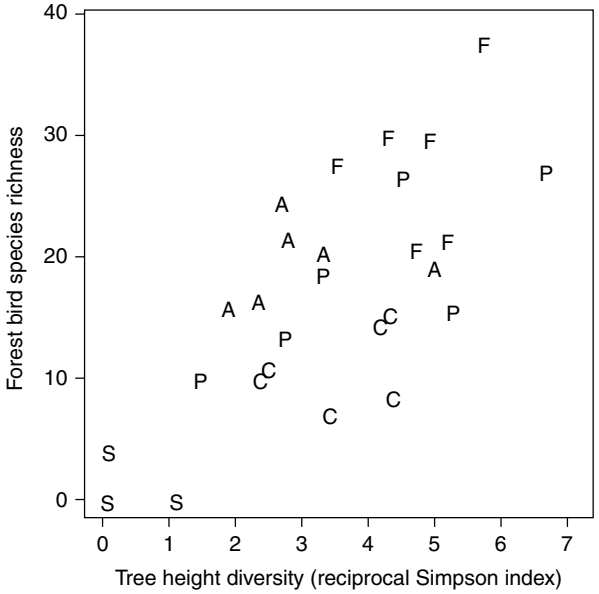


Figure 7 Forest bird species richness vs. tree height diversity in the ancient tropical countryside of Karnataka, India. All species richness measures are estimated using the bootstrap species estimator. Five land covers are specified as follows: A, areca nut palm; C, cashew; F, intact forest; P, production forest; and S, shrub. Reproduced from Ranganathan J, Daniels R, Chandran M, Ehrlich P, and Daily GC (2008) Sustaining biodiversity in ancient tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America* 105(46): 17852–17854.

found to support more forest species (**Figure 7**) (Ranganathan *et al.*, 2008).

The persistence of biodiversity in tropical ancient countryside suggests that patterns observed in the Neotropics and Europe are not artifacts of extinction debts soon to be paid or low regional species pools, respectively. The equilibrium states of newly deforested areas, like the Coto Brus Valley, will continue to be revealed as studies dig deeper into how biodiversity is influenced by human activities, but insights from Karnataka, India, suggest these patterns are not necessarily temporary.

From Land to Water

Terrestrial ecosystems have always been disproportionately represented in ecological studies despite the greater global extent of marine ecosystems. Countryside biogeography has also been subject to ecology's terrestrial bias. Currently we do not know of any studies that directly apply the countryside biogeography approach to a human-dominated seascape. This may be due, in part, to the different ways in which humans use the land and sea. Historical ecology has revealed profound changes in the abundance and diversity of marine life as a result of human activities (Jackson *et al.*, 2001; Pandolfi *et al.*, 2003; Kappel, 2005; Lotze *et al.*, 2006). Unlike agricultural and urban landscapes, the impacts of human activities on the seafloor and the ocean's volume through fishing, bottom-trawling, pollution, and transport are rarely visible. Because the majority of marine systems are generally unseen (and therefore understudied), the principles of countryside biogeography developed in human-dominated landscapes should be applied with care to the oceans.

Even though pioneering investigations of countryside biogeography in marine ecosystems have yet to be conducted, a growing number of studies are underway. Studies of the heavily human-dominated waters of the Mediterranean Sea and Hawaii provide examples of how some of countryside biogeography's concepts may be taken from the land to the water.

The distribution and composition of the Mediterranean Sea's biodiversity have been directly shaped by human activity since ancient times. The sustained influence humanity has exerted on the Mediterranean through habitat alteration and harvesting for thousands of years is now compounded by the modern impacts of industrialized fishing and coastal development, pollution, eutrophication, climate change, and the introduction of invasive species (Coll *et al.*, 2010). Despite these human influences, high biodiversity persists within the Mediterranean with over 12,000 plant and animal species estimated to reside in the region (Coll *et al.*, 2010). We do not, however, have a clear understanding of how these species are responding to increasing human activity.

There is evidence that charismatic and sensitive species like the monk seal, loggerhead and green sea turtles, and many seabirds have undergone range constrictions or abandoned breeding areas, but fishing and hunting records have indicated that other species have remained stable (Coll *et al.*, 2010). In particular, the Mediterranean harvest of a number of small, short-lived species like the spotted mantis squid (*Squilla mantis*), sardines (*Sardina pilchardus*), bogue (*Boops boops*), and red mullet (*Mullus barbatus*) appears to be sustainable (Caddy and Surette, 2005).

In the heavily used waters surrounding the Main Hawaiian Islands, a study has shown that substantial biodiversity persists in heavily fished oceans (**Figure 8**). These accounts of biodiversity in the human-dominated seascapes of Hawaii and the Mediterranean Sea are early results suggesting that as on land, substantial marine life can coexist with humans.

Methods and Future Directions

Species Traits and Human-Dominated Ecosystems

The predictive nature of countryside biogeography encourages ecologists to forecast biodiversity's trajectory in response to human activities. Identifying genetic, physical, or behavioral traits of species or populations that confer fitness (the ability to survive and reproduce) advantages or disadvantages in a human-dominated ecosystem is a powerful tool because we can infer similar responses from other species and populations with shared traits. The ability to generalize findings across biomes and to distantly related taxa is critical to modeling how biodiversity will respond to changes in human activity, rather than experimenting on each species and population individually. For instance, a trait often taken as a measure of fitness is clutch size (number of eggs laid) of a bird or reptile

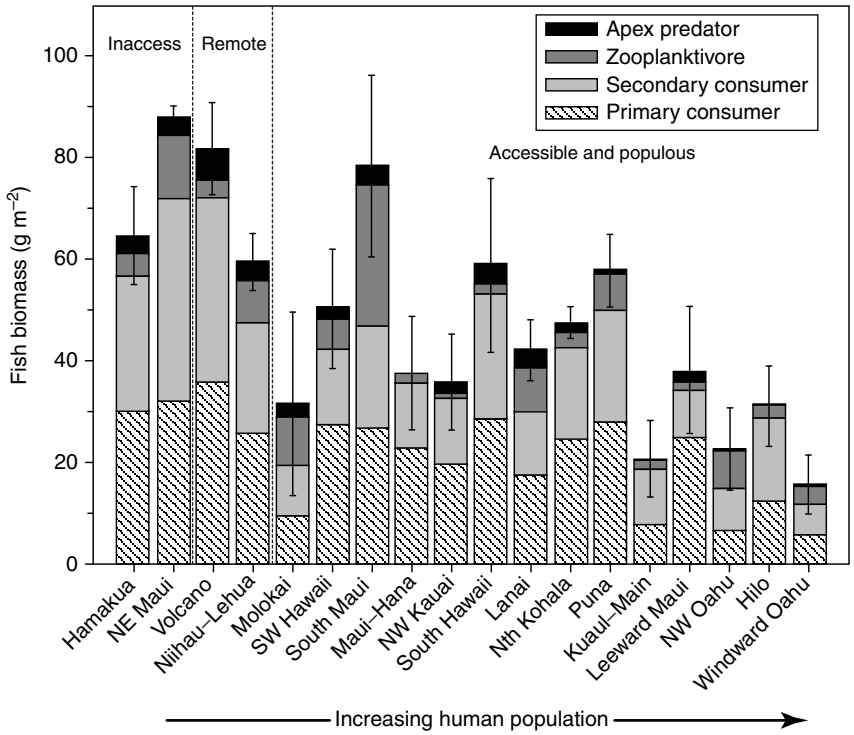


Figure 8 Mean fish biomass by consumer group per location around Main Hawaiian Islands. ''Inaccessible,'' ''remote,'' and ''accessible and populous'' locations are sorted in order of increasing local human population density (lowest left, highest right). Error bars represent SE of total fish biomass at survey sites within each location. Reproduced from Williams ID, Walsh WJ, Schroeder RE, *et al.* (2008) Assessing the importance of fishing impacts on Hawaiian coral reef fish assemblages along regional-scale human population gradients. *Environmental Conservation* 35(3): 261–272, with permission from CUP.

in a single breeding event (Jetz *et al.*, 2008; Shine, 2005). Clutch sizes can be used as a means to compare distantly related species or separate closely related species (or even populations) (Jetz *et al.*, 2008). For example, populations of House Sparrows can be distinguished by clutch size: tropical populations lay 2–3 eggs per clutch, whereas temperate populations lay 4–5 eggs, reflecting general patterns in clutch sizes of the tropical and temperate bird communities to which House Sparrow populations belong.

Such differences and similarities in traits can shed light on species' potential advantages or disadvantages in human-dominated ecosystems. Moreover, species interact with each other, and one species' success can trigger another's success. Such as the case in Ghana, where the small royal or ball python (*Python regius)* is flourishing in concert with rodent infestations (Gorzula *et al.*, 1997). Lumping and linking organisms based on traits, rather than taxonomic relationships, permits a bioassay approach, or the use of multiple organisms or species to sort observed patterns, processes, and mechanisms.

The trait-based approach has been used in ecology for decades to understand how physical, behavioral, and genetic traits affect fitness in variable environments. Traits can include the concept of a guild – usually a group that exploits the same resource in a similar way and similar to the concept of an ecological niche (Simberloff and Dayan, 1991), behavioral syndromes (Sih *et al.*, 2004), trophic levels, morphology and life form, or life history traits (e.g., home range/territory size, reproductive capacity, and activity patterns). Traits can help or hinder fitness in the face of human-induced environmental changes. A primary challenge of countryside biogeography is to identify and understand which traits confer advantages and disadvantages in human-dominated ecosystems. For example, one study found that plants that were bat-, bird-, and bee-pollinated and bird- or monkey-dispersed were uncommon in deforested habitats, but plants with seeds dispersed by clinging to mammal fur (''burrs'') were common in deforested areas (Mayfield *et al.*, 2006). These patterns are indicative of community structure shifts between plants and their mutualistic partners in deforested areas, which may be shaping the ecosystem's vegetation communities.

Another example is a comparative study of moths and butterflies that showed nocturnal behavior of moths to be advantageous. Nocturnality in moths permitted greater use of agricultural plots, foraging opportunity in distant forest fragments, and possibly better dispersal capabilities. This was because microclimate was homogenized across the countryside at night versus relatively extreme daytime differences between areas with and without extensive tree cover, which created barriers for diurnal butterflies (Daily and Ehrlich, 1996). Traits-based approaches may help narrow the possible trajectories of future extinctions, colonizations, and evolutionary paths of organisms in human-dominated ecosystems.

Evolution in the Countryside

Biogeography and human activities shape evolutionary paths of life on the planet. One of the most powerful ways humans affect evolution through their activities is by the removal of species and populations. Extinction of a species is not only the removal of an organism forever, but it is also the removal of its evolutionary potential. All future life will be derived from those able to exist in human-dominated ecosystems. The evolutionary bottleneck of the human-induced sixth extinction wave is currently irreversible – with little hope of sensationalized science fiction ideas of cloning or miraculous reanimation of extinct species. Similarly to the Cretaceous–Tertiary extinction event 65 million years ago that cleared the evolutionary playing field for mammals, our actions are determining the planet's future occupants.

In addition to the grave impact species extinction will have on evolution, population extinctions also cause bottlenecks altering the course of evolution. Human activity is altering the distribution of genes over space and time. For example, the reduction of African cheetah (*Acinonyx jubatus*) populations has caused such a tight genetic bottleneck that virtually no genetic variation exists among extant individuals, which makes inbreeding unavoidable and increases the species' vulnerability to disease and environmental change (O'Brien *et al.*, 1983). Conversely, positive population changes in both size and range of many species can also alter their evolutionary paths by increasing mutation rates and providing new evolutionary contexts for mutations to fail or succeed. Factors relating to gene flow, such as dispersal, temporal and spatial isolation, and changes in population size acutely impact evolution. Unfortunately, our understanding of these deep processes is limited, which makes it a key area of current and future study within countryside biogeography.

Remote Sensing

Perhaps one of the most important tools in modern biogeography is remote sensing. It is increasingly common for ecologists to use aerial and satellite sensor technologies to detect and classify objects on earth – including land and marine surfaces, the atmosphere, subterranean geologic characteristics, and floral and faunal community components. With the accessibility afforded to scientists and nonscientists through Google Earth and other virtual globe, map, and geographical information systems our collective knowledge of how the planet changes is unprecedented. The use of such tools coupled with hyperspectral imaging sensors, advanced spectroscopic and laser imaging, light detection and ranging technology, and geodetic data collection facilitates the rapid assessment of ecosystem change.

Mathematical Models and Theory

Countryside biogeography's theoretical foundations rely heavily on the fields in which it is rooted. Many theories in ecology already account for community responses to invasion, responses to disturbance and succession, and fluctuations in biogeochemical processes. Appropriate modifications of some theories have been conducted. For example, species–area relationships have been adjusted to account for gradients of habitat hospitability (Pereira and Daily, 2006; Pereira *et al.*, 2004) and the permeability of different human-made habitats have been incorporated into metapopulation dynamics (Ray *et al.*, 2002).

There remain, however, considerable portions of ecological, evolutionary, sociological, and economic theory that have yet to incorporate effects of human-dominated ecosystems into their frameworks. For example, there has been very little work looking at how human-dominated ecosystems affect species' predator–prey, host–parasitoid, or mutualistic interactions. Also, the genetic consequences of habitat alteration and toxification of the planet are poorly understood from the perspectives of genetic drift, genetic mutation/recombination rates, and evolutionary consequences. And, finally, there remains considerable scope to develop theories that articulate how fluctuations in biodiversity and biodiversity-driven ecosystem services will interact and change human populations, well-being, societies, and enterprises.

Practical Strategies to Enhance the Conservation Value of Countryside Habitats

Despite the presentation of countryside biogeography in this piece as an apolitical framework, it is our understanding that most will (and should) practice countryside biogeography with the goal of conserving and protecting biodiversity in the long term. We acknowledge that there will be trade-offs and not all species or populations will survive into the future. Perhaps one of the most effective ways to ensure our survival and maintain our ecosystem-driven life-support systems is to control our own population growth through education, women's empowerment, and harnessing economic forces.

An equally vital action is to reduce per-capita consumption by the already rich and promote the conservation value of countryside and other habitats. Enhancement of the countryside can be done both locally and globally. At local scales even the conservation of narrow strips of trees around rivers and property borders, small patches of native grassland, or reduced use of coastal enclaves have been found to abate the decay of biodiversity in human-dominated ecosystems (Mendenhall *et al.*, 2011; Karp *et al.*, 2011). At the global scale, the countryside can be enhanced through the reduction of chemical inputs, drastically reducing carbon emissions, increasing the number of protected areas, and creating national and global policies that account for the full ecological costs of food production practices and other human activities. Finally, strategic land- and ocean-use decision making that adequately weighs the short- and long-term trade-offs between human activities and biodiversity-driven ecosystem services will be critical. Without all of these actions, the society we know may perish in the sixth extinction wave.

Appendix

List of Courses

1. Biogeography
2. Biology

3. Conservation Biology
4. Ecology
5. Natural Resources
6. Wildlife Management

See also: Ecosystem, Concept of. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Island Biogeography. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Vicariance Biogeography

References

Agardy T (1997) *Marine Protected Areas and Ocean Conservation*. San Diego, CA: RG Landes Co. and Academic Press.

Balvanera P, Pfisterer AB, Buchmann N, *et al.* (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9(10): 1146–1156.

Barnard C (1924) A review of birdlife on Coomooboolaroo Station, Duainga District, Queensland, during the past fifty years. *Emu* 24: 252–265.

Barnosky AD, Matzke N, Susumu T, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471(7336): 51–57.

Barrett G, Ford H, and Recher H (1994) Conservation of woodland birds in a fragmented rural landscape. *Pacific Conservation Biology* 1(3): 245–256.

Benton T (2003) Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology and Evolution* 18(4): 182–188.

Bignal EM and McCracken DI (1996) Low-intensity farming systems in the conservation of the countryside. *The Journal of Applied Ecology* 33(3): 413–424.

Bratman GN, Hamilton JP, and Daily GC (2012) The impacts of nature experience on human cognitive function and mental health. In: Ostfeld R and Schlesinger W (eds.) *The Year in Ecology and Conservation Biology*. Annals of the New York Academy of Sciences.

Brosi B, Daily G, and Ehrlich P (2007) Bee community shifts with landscape context in a tropical countryside. *Ecological Applications* 17(2): 418–430.

Brown J (1989) Comparative population ecology of eleven species of rodents in the Chihuahuan Desert. *Ecology* 70(5): 1507–1525.

Bush MB and Whittaker RJ (1993) Non-equilibration in island theory of Krakatau. *Journal of Biogeography* 20(4): 453–457.

Caddy JF and Surette T (2005) In retrospect the assumption of sustainability for Atlantic fisheries has proved an illusion. *Reviews in Fish Biology and Fisheries* 15(4): 313–337.

Ceballos G and Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science* 296(5569): 904–907.

Ceballos G, Garcia A, and Ehrlich PR (2010) The sixth extinction crisis: Loss of animal populations and species. *Journal of Cosmology* 8: 1821–1831.

Chippaux J (1998) Snake-bites: Appraisal of the global situation. *Bulletin of the World Health Organization* 76(5): 515–524.

Chundawat R, Gogate N, and Johnsingh A (1999) Tigers in Pann: Preliminary results from an Indian tropical dry forest. In: Seidensticker J, Christie S, and Jackson P (eds.) *Riding the Tiger: Tiger Conservation in Human-Dominated Landscapes*, pp. 123–129. Cambridge, UK: Cambridge University Press.

Clark JS, Carpenter SR, Barber, *et al.* (2001) Ecological forecasts: An emerging imperative. *Science* 293(5530): 657–660.

Clement RM and Horn SP (2001) Pre-Columbian land-use history in Costa Rica: A 3000-year record of forest clearance, agriculture and fires from Laguna Zoncho. *The Holocene* 11(4): 419–426.

Coad L, Burgess N, Loucks C, *et al.* (2009) *The Ecological Representativeness of the Global Protected Areas Estate in 2009: Progress Towards the CBD 2010 Target*, United Nations Environment Programme World Conservation Monitoring Center, pp. 1–35. Oxford, UK: Oxford University Press.

Coll M, Piroddi C, Steenbeek J, *et al.* (2010) The biodiversity of the Mediterranean Sea: Estimates, patterns, and threats. *PLoS One* 5(8): 1–36.

Costanza R, D'Arge R, De Groot R, *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature* 387(6630): 253–260.

Daily GC (1997) Countryside biogeography and the provision of ecosystem services. In: Raven PH (ed.) *Nature and Human Society: The Quest for a Sustainable World*, pp. 104–113. Washington, DC: National Research Council, National Academy Press.

Daily GC (2001) Ecological forecasts. *Nature* 411(6835): 245.

Daily GC, Ceballos G, Pacheco J, Suzan G, and Sanchez-Azofeifa A (2003) Countryside biogeography of Neotropical mammals: Conservation opportunities in agricultural landscapes of Costa Rica. *Conservation Biology* 17(6): 1814–1826.

Daily GC and Ehrlich PR (1996) Nocturnality and species survival. *Proceedings of the National Academy of Sciences of the United States of America* 93(21): 11709–11712.

Daily GC, Ehrlich PR, and Sanchez-Azofeifa G (2001) Countryside biogeography: Use of human-dominated habitats by the avifauna of southern Costa Rica. *Ecological Applications* 11(1): 1–13.

Ehrlich PR (1997) *A World of Wounds: Ecologists and the Human Dilemma*. Oldendorf/Luhe, Germany: Ecology Institute.

Ehrlich PR and Daily GC (1993) Population extinction and saving biodiversity. *AMBIO* 22(2/3): 64–68.

Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, and Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography* 19: 589–606.

Elton C (1927) *Animal Ecology*. Chicago: University of Chicago Press.

Estrada A, Coates-Estrada R, and Meritt Jr D (1993) Bat species richness and abundance in tropical rain forest fragments and in agricultural habitats at Los Tuxtlas, Mexico. *Ecography* 16(4): 309–318.

Estrada A, Coates-Estrada R, and Meritt Jr D (1994) Non flying mammals and landscape changes in the tropical rain forest region of Los Tuxtlas, Mexico. *Ecography* 17(3): 229–241.

Gardner TA, Barlow J, Chazdon R, *et al.* (2009) Prospects for tropical forest biodiversity in a human-modified world. *Ecology Letters* 12(6): 561–582.

Gibson L, Lee TM, Koh LP, *et al.* (2011) Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* 1–6.

Gorzula S, Nsiah W, and Oduro W (1997) Survey of the status and management of the Royal Python (Python regius) in Ghana. *Report to the Secretariat of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES)*, pp 1–17. Geneva, Switzerland: UNEP-CITES.

Graham K, Beckerman A, and Thirgood S (2005) Human–predator–prey conflicts: Ecological correlates, prey losses and patterns of management. *Biological Conservation* 122(2): 159–171.

Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–554.

Halpern BS, Walbridge S, Selkoe KA, *et al.* (2008) A global map of human impact on marine ecosystems. *Science* 319(5865): 948–952.

Hoare RE (1999) Determinants of human–elephant conflict in a land-use mosaic. *Journal of Applied Ecology* 36(5): 689–700.

Holdridge L (1979) *Life Zone Ecology*. San José, Costa Rica: Tropical Science Center.

Horner-Devine MC, Daily GC, Ehrlich PR, and Boggs CL (2003) Countryside biogeography of tropical butterflies. *Conservation Biology* 17(1): 168–177.

Hughes JB, Daily GC, and Ehrlich PR (1997) Population diversity: Its extent and extinction. *Science* 278(5338): 689–692.

Hughes JB, Daily GC, and Ehrlich PR (2002) Conservation of tropical forest birds in countryside habitats. *Ecology Letters* 5(1): 121–129.

Jablonski D and Chaloner WG (1994) Extinctions in the fossil record (and discussion). *Philosophical Transactions: Biological Sciences* 344(1307): 11–17.

Jackson JB, Kirby MX, Berger WH, *et al.* (2001) Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293(5530): 629–637.

Jetz W, Sekercioglu CH, and Böhning-Gaese K (2008) The worldwide variation in avian clutch size across species and space. *PLoS Biology* 6(12): 2650–2657.

Kappel CV (2005) Losing pieces of the puzzle: Threats to marine, estuarine, and diadromous species. *Frontiers in Ecology and the Environment* 3(5): 275–282.

Kareiva P, Tallis H, Ricketts TH, Daily GC, and Polasky S (2011) *Natural Capital: Theory and Practice of Mapping Ecosystem Services*. Oxford, UK: Oxford University Press.

Karp D, Guy Z, Zook J, Ehrlich P, and Daily G (2011) Resilience and stability in bird guilds across tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America* 108(52): 21134–21139.

Kelt DA and Brown JH (2000) Species as units of analysis in ecology and biogeography: Are the blind leading the blind? *Global Ecology and Biogeography* 9(3): 213–217.

Leach G and Hines H (1993) Frequency of observation of bird species in sub-costal farmland in southeast Queensland. *Memoirs of the Queensland Museum* 33(1): 259–275.

Lotze HK, Lenihan HS, Bourque BJ, *et al.* (2006) Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science* 312(5781): 1806–1809.
MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Princeton, NJ: Princeton University Press.
Mayfield M and Daily G (2005) Countryside biogeography of Neotropical herbaceous and shrubby plants. *Ecological Applications* 15(2): 423–439.
Mayfield MM, Ackerly D, and Daily GC (2006) The diversity and conservation of plant reproductive and dispersal functional traits in human-dominated tropical landscapes. *Journal of Ecology* 94(3): 522–536.
Mellink E (1991a) Bird communities associated with three traditional agroecosystems in the San Luis Potosi Plateau, Mexico. *Agriculture, Ecosystems and Environment* 36(1–2): 37–50.
Mellink E (1991b) Rodent communities associated with three traditional agroecosystems in the San Luis Potosi Plateau, Mexico. *Agriculture, Ecosystems and Environment* 33(4): 363–375.
Mendenhall C, Ehrlich P, and Daily G (2012) Improving estimates of biodiversity loss. *Biological Conservation* 151(1): 32–34. http://www.sciencedirect.com/science/article/pii/S0006320712001115
Mendenhall CD (2011) Tropical forests: Try holistic conservation. *Nature* 479(7372): 179.
Mendenhall CD, Sekercioglu CH, Oviedo Brenes F, Ehrlich PR, and Daily GC (2011) Predictive model for sustaining biodiversity in tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America* 108(39): 16313–16316.
Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Current State and Trends.* Chicago, IL: Island Press.
Naidoo R (2004) Species richness and community composition of songbirds in a tropical forest-agricultural landscape. *Animal Conservation* 7(1): 93–105.
Nelson G (1974) Biogeography historical: An alternative formalization. *Systematic Biology* 23(4): 555–558.
O'Brien S, Wildt D, Goldman D, Merril CR, and Bush M (1983) The cheetah is depauperate in genetic variation. *Science* 221: 459–462.
Pandolfi JM, Bradbury RH, Sala E, *et al.* (2003) Global trajectories of the long-term decline of coral reef ecosystems. *Science* 301(5635): 955–958.
Pereira H and Daily GC (2006) Modeling biodiversity dynamics in countryside landscapes. *Ecology* 87(8): 1877–1885.
Pereira HM, Daily GC, and Roughgarden J (2004) A framework for assessing the relative vulnerability of species to land-use change. *Ecological Applications* 14(3): 730–742.
Perfecto I, Vandermeer J, and Wright A (2009) *Nature's Matrix: Linking Agriculture, Conservation, and Food Sovereignty.* Oxford, UK: Earthscan.
Pimm SL, Russell G, Gittleman J, and Brooks T (1995) The future of biodiversity. *Science* 269(5222): 347–350.
Ranganathan J, Daniels R, Chandran M, Ehrlich Paul, and Daily GC (2008) Sustaining biodiversity in ancient tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America* 105(46): 17852–17854.
Ray N, Lehmann A, and Joly P (2002) Modeling spatial distribution of amphibian populations: A GIS approach based on habitat matrix permeability. *Biodiversity and Conservation* 11: 2143–2165.
Ricketts TH (2001) The matrix matters: Effective isolation in fragmented landscapes. *The American Naturalist* 158(1): 87–99.
Ricketts TH, Daily GC, Ehrlich PR, and Fay J (2001) Countryside biogeography of moths in a fragmented landscape: Biodiversity in native and agricultural habitats. *Conservation Biology* 15(2): 378.
Rohde RA and Muller RA (2005) Cycles in fossil diversity. *Nature* 434: 208–210.
Roughgarden J and Diamond J (1986) Overview: The role of species interactions. In: Diamond J and Case T (eds.) *Community Ecology*, pp. 333–343. New York, NY: Harper and Row.
Sala E and Knowlton N (2006) Global marine biodiversity trends. *Annual Review of Environment and Resources* 31(1): 93–122.
Sansonetti V (1995) *Quemé mis naves en estas montañas: La colonización de la altiplanicie de Coto Brus y la funcadión de San Vito de Java.* San José, Costa Rica: Jiménez and Tanzi Ltda.
Santos-Barrera G, Pacheco J, Mendoza-Quijano F, *et al.* (2008) Diversity, natural history and conservation of amphibians and reptiles from the San Vito region, southwestern Costa Rica. *Revista de Biología Tropical* 56(2): 755–778.
Sekercioglu CH, Loarie SR, Oviedo Brenes F, Ehrlich PR, and Daily GC (2007) Persistence of forest birds in the Costa Rican agricultural countryside. *Conservation Biology* 21(2): 482–494.
Sheppard C (2010) Marine protected areas and pelagic fishing: The case of the Chagos Archipelago. *Marine Pollution Bulletin* 60(11): 1899–18901.
Shine R (2005) Life-history evolution in reptiles. *Annual Review of Ecology, Evolution, and Systematics* 36(1): 23–46.
Sih A, Bell A, and Johnson JC (2004) Behavioral syndromes: An ecological and evolutionary overview. *Trends in Ecology and Evolution* 19(7): 372–378.
Simberloff D and Dayan T (1991) The guild concept and the structure of ecological communities. *Annual Review of Ecology and Systematics* 22(1): 115–143.
Tilman D, Fargione J, Wolff B, *et al.* (2001) Forecasting agriculturally driven global environmental change. *Science* 292(5515): 281–284.
Tilman D, May R, and Lehman C (1994) Habitat destruction and the extinction debt. *Nature* 371: 65–66.
Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of Earth's ecosystems. *Science* 277(5325): 494–499.
Wackernagel M (1999) National natural capital accounting with the ecological footprint concept. *Ecological Economics* 29(3): 375–390.
Waltert M, Mardiastuti A, and Mühlenberg M (2004) Effects of land use on bird species richness in Sulawesi, Indonesia. *Conservation Biology* 18(5): 1339–1346.
Watts M, Ball I, Stewart R, *et al.* (2009) Marxan with Zones: Software for optimal conservation based land- and sea-use zoning. *Environmental Modelling & Software* 24(12): 1513–1521.
Whittaker RJ, Araujo MB, Paul J, *et al.* (2005) Conservation biogeography: Assessment and prospect. *Diversity and Distributions* 11: 3–23.
Wilkins JS (2009) *Species: A History of the Idea.* Ewing, NJ: University of California Press.
Winfree R and Kremen C (2009) Are ecosystem services stabilized by differences among species? A test using crop pollination. *Proceedings of the Royal Society B: Biological Sciences* 276(1655): 229–237.
Wohlgenant TJ (1994) Roost interactions between the common vampire bat (*Desmodus rotundus*) and two frugivorous bats (*Phyllostomus discolor* and *Sturnira lilium*) in Guanacaste, Costa Rica. *Biotropica* 26(3): 344–348.
Woodwell GM (2010) The biodiversity blunder. *BioScience* 60(11): 870–871.
Yarincik K (2005) The census of marine life: Goals, scope and strategy. *Scientia Marina* 69: 201–208.

C_4 Plants

Rowan F Sage and Tammy L Sage, The University of Toronto, Toronto, ON, Canada

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Rowan F. Sage, volume 1, pp 575–598, © 2004 Elsevier Inc.

Glossary

Bundle sheath (BS) tissue A cell layer at the periphery of the vascular bundles in leaves. In C_4 plants, the reactions of the photosynthetic carbon reduction cycle are localized in this layer, in contrast to C_3 plants, where they occur in the mesophyll (M) tissue. In comparison to C_4 species, the BS cells of C_3 species are small, with few chloroplasts.

Kranz anatomy Specialized anatomy of the C_4 leaf in which the BS tissue is enlarged and enriched with chloroplasts, whereas the mesophyll is reduced in size and often forms a lighter green halo around the BS. This produces a wreath-like ('Kranz' in the original German) appearance.

Mesophyll (M) tissue Photosynthetic cells that are located between the arrays of vascular bundles and BS cells of a leaf. In C_4 plants, Phosphoenolpyruvate (PEP) carboxylase and pyruvate–phosphate dikinase are localized in the M tissue, while Rubisco is absent.

NUE Nitrogen-use efficiency, a measure of photosynthesis or growth relative to the nitrogen content of the leaves or plant.

Phosphoenolpyruvate carboxylase PEP carboxylase, the initial carboxylation enzyme in C_4 photosynthesis.

Photorespiration Biochemical process in which O_2 is assimilated during the oxygenation of RuBP, and CO_2 is given off in the metabolism of the products of RuBP oxygenation.

Rubisco Ribulose-1,5-bisphosphate carboxylase/oxygenase, the enzyme in all plants that catalyzes the formation of two phosphoglycerate molecules from RuBP and CO_2, and catalyzes the competing reaction of RuBP oxygenation.

WUE Water-use efficiency, a measure of the amount of photosynthesis or growth per unit of water lost in transpiration.

Introduction

In the past 35 million years, atmospheric CO_2 levels have declined from above 1000 ppm (ppm) to values below the current concentration of 390 ppm. During the peak of the ice age 20,000 years ago, atmospheric CO_2 levels were 180 ppm. The low CO_2 levels of recent geological time substantially impaired photosynthetic carbon gain in the Earth's flora, which relied on a biochemical mode of carbon assimilation termed C_3 photosynthesis. In response to carbon deficiencies brought about by low CO_2, multiple lineages of plants evolved a CO_2-concentrating mechanism termed C_4 photosynthesis. C_4 plants function with much greater efficiency than C_3 plants in warm environments, and contribute a much larger fraction to global primary productivity than their species numbers would indicate. The origin of C_4 photosynthesis also set in motion the evolutionary diversification of grazing and predatory guilds that created the diverse animal assemblages of grassland biomes of recent time. Today, C_4 plants are important contributors to biodiversity in tropical, subtropical, and warm-temperate habitats. C_4 plants currently play a major role in global agriculture and are the leading candidates for the next generation of bioenergy crops (Table 1). Much of the world's C_4 flora may be at risk in the future, because rising CO_2 and intense exploitation of C_4-dominated landscapes will reduce available habitat for native C_4 plants. A handful of invasive C_4 species will do well, however, as they are adapted to human disturbance. These species are already responsible for major reductions in biodiversity in tropical and subtropical ecosystems. In recognition of the significance of C_4 plants, this article provides an overview of C_4 photosynthesis and its contribution to global biodiversity.

What is C_4 Photosynthesis?

C_4 photosynthesis arose as an evolutionary modification of the preexisting C_3 photosynthetic pathway, which is used by approximately 250,000 plants. The C_4 pathway does not replace C_3 photosynthesis; instead, it complements the C_3 pathway, which is relocated to an interior compartment within leaves of C_4 plants. In C_3 plants, CO_2 is first fixed to the acceptor molecule ribulose-1,5, bisphosphate (RuBP), forming phosphogylcerate (PGA), a three-carbon, or C_3, organic acid (Figure 1). PGA is metabolized to sugars, starch, and RuBP by the C_3 photosynthetic carbon reduction (PCR) cycle, or Calvin–Benson cycle named after the team that discovered it. The PCR cycle completely occurs within the same cell, and hence all photosynthetic cells in C_3 plants conduct the same photosynthetic metabolism.

Photorespiration

The fixation of CO_2 by Rubisco and its conversion to starch and sugars using light-dependent energy is the primary source of photosynthetic carbon in all higher plants and algae. Rubisco, however, has a side reaction that is inhibitory for photosynthetic carbon gain. In the presence of O_2, Rubisco can oxygenate RuBP, creating one PGA molecule and one molecule of phosphogylcolate (PG), which has no value to the

Table 1 Economically significant C_4 plants of the World

Grasses	
Global crops	*Zea mays* (maize), *Sorghum bicolor* (sorghum), *Pennisetum glaucum* (pearl millet), *Saccharum officinalis* (sugarcane)
Regional cereals	*Eleusine coracana* (finger millet, Africa), *Setaria italica* (foxtail millets, N. Africa, Arabia), *Digitaria exilis, Digitaria deflexa.* and *Digitaria iburua* (fonios, Africa), *Eragrostis tef* (tef, Ethiopia), *Brachiaria deflexa* (animal fonio, Africa), *Coix lacrymi-jobi* (adlay, India), *Echinochloa colona* (Sawa, India), *Panicum sumatrense* (sama, India), *Setaria pumila* (korali, India),
Forages	*Paspalum* (5 spp.), *Pennisetum* (2 spp.), *Panicum* (2 spp.), *Cenchrus ciliaris, Chloris guyana, Cynodon dactylon, Digitaria decumbens, Melinis minutiflora, Sorghum almum, Setaria anceps*
Weeds	*Cynodon dactylon* (bermuda grass), *Digitaria sanguinalis* (crabgrass), *Echinochloa crusgalli* (barnyard grass), *Echinochloa colonum, Eleusine indica, Imperata cylindrica, Rottboellia cochinchinensis* (itchgrass), *Sorghum halopense* (Johnson grass)
Turf grasses	*Cynodon dactylon* (bermuda grass), *Zoysia japonica* (zoysia), *Pennisetum clandestinum, Paspalum* spp.
Bioenergy plants	*S. officinalis*, *Z. mays*, *S. bicolor*, *Miscanthus x giganteus*, Miscane (*Saccharum* x *Miscanthus* hybrid), *Panicum virgatum* (switchgrass), *Panicum maximum* (guinea grass), *Pennisetum purpureum* (elephant grass)
Sedges (no major crops)	
Weeds	*Cyperus esculentus, Cyperus rotundus*
Other	*Cyperus papyrus* (papyrus, central Africa)
Dicots	
Minor crops	*Amaranthus edulis* (grain amaranth, Americas), *Amaranthus tricolor* (vegetable amaranth, Americas, India), *Portulaca oleracea* (purslane, S.E. Asia)
Weeds	*Amaranthus* (pigweeds; 3 spp.), *Heliotriopium indicum, P. oleracea, Euphorbia hirta, Salsola kali* (tumbleweed), (*Tribulus terrestris*) (caltrop)
Other	*Haloxylon* spp. (saxouls, 4 species provided fuel and fodder along the silk roads of central Asia); *Atriplex* saltbushes.

Source: Modified from Brown RH (1999) Agronomic implications of C_4 photosynthesis. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 473–507. San Diego: Academic Press.

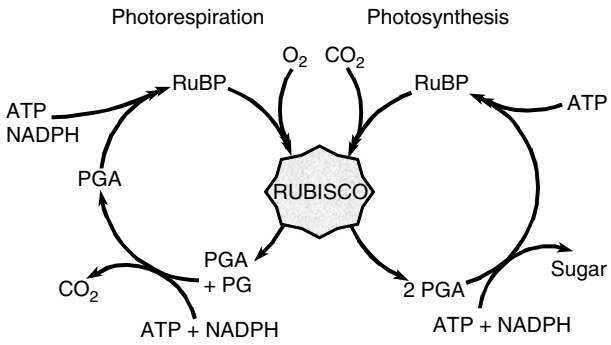


Figure 1 Diagrammatic representation of the photosynthetic carbon reduction cycle and the photorespiratory cycle in C_3 plants. *Abbreviations*: PG, phosphoglycolate; PGA, phosphoglyceric acid; RuBP, ribulose-1,5-bisphosphate. Reproduced from Sage RF (1999) Why C_4 Plants? In: Sage RF and Monson RK (eds.) *C_4 Plant Biology,* pp. 3–16. San Diego: Academic Press.

plant and is toxic if it accumulates in the cell (**Figure 1**). Plants metabolize PG back to PGA, but in the process 25% of the carbon in the pool of PG molecules is given off as CO_2. In addition, for every two RuBP oxygenation events, one adenosine triphosphate (ATP) and the equivalent of one nicotineamide adenine dinucleotide phosphate (NADPH) are required to convert PG to PGA, which can then be used to regenerate RuBP. The oxygenation of RuBP and subsequent PG metabolism are collectively termed photorespiration, because it involves the uptake of O_2 and release of CO_2, as observed in other respiratory processes. Photorespiration is a significant inhibition on C_3 photosynthesis in warm, low CO_2 environments, due to the energy use and loss of previously fixed CO_2 that occurs during PG metabolism. To overcome this inhibition, plants have evolved CO_2 concentrating mechanisms, the most significant of which is C_4 photosynthesis.

The C_4 Photosynthetic Pathway

In C_4 photosynthesis, leaf biochemistry and anatomy are modified from the C_3 pattern to concentrate CO_2 from an outer tissue or cellular compartment into an inner compartment in which Rubisco and the PCR cycle are localized (**Figure 2**). Concentration of CO_2 around Rubisco greatly reduces the rate of photorespiration and allows for greater, more efficient rates of photosynthesis. The outer compartment is commonly (but not always) the mesophyll (M) tissue in a leaf, whereas the inner compartment is either a distinct tissue layer in the middle of the leaf, or a distinct region within a single cell. CO_2 concentration occurs via the operation of a C_4 metabolic cycle, in which atmospheric CO_2 (in the form of bicarbonate) is first assimilated into four-carbon-organic acids in the outer compartment; these acids then move to the inner compartment where they are metabolized to release CO_2 (**Figure 2**). The term 'C_4' refers to the number of carbon molecules in the carboxylation product of the C_4 metabolic cycle.

C_4 photosynthesis has evolved dozens of times via the modification of preexisting enzymes and structures. Each version of the C_4 pathway is unique in some way. Variation is present in the enzymes and isoenzymes that have been co-opted for the C_4 metabolic cycle, in the tissue compartments used for CO_2 concentration, and in the cellular modifications to allow for efficient C_4 function. There are, for example, at least 22 distinct types of leaf anatomy in C_4 leaves. Because of

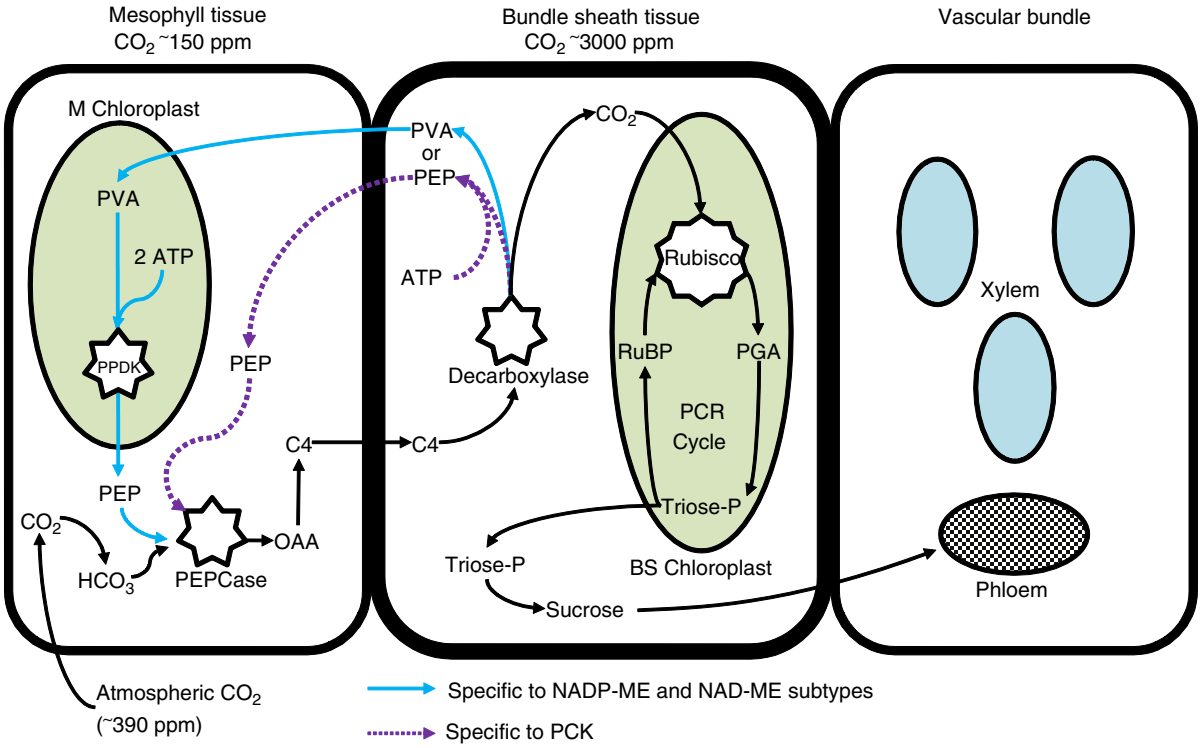


Figure 2 Diagrammatic representation of the C_4 photosynthetic cycle. The product of the PEP carboxylation reaction is oxaloacetic acid (OAA), which is converted into a transportable four carbon acid (C_4) that is either malic acid (NADP-ME type) or aspartic acid (NAD-ME and PCK subtypes). The decarboxylation reactions produce either pyruvic acid (PVA, produced by NADP-ME and NAD-ME) or PEP (produced by PEP carboxykinase). The PVA returns to the mesophyll chloroplasts, where it is converted to PEP using ATP. *Abbreviations*: CHL, chloroplast; triose-P, triose phosphate; PCR, photosynthetic carbon reduction cycle.

this variation, C_4 photosynthesis is best considered a syndrome, where dissimilar features are assembled to give rise to a common function. Although the elements making up the different versions of the C_4 pathway vary, in all cases there are common characteristics that delineate C_4 photosynthesis (Table 2).

C_4 Biochemistry

C_4 photosynthesis begins with the carboxylation of the three-carbon compound phosphoenolpyruvate (PEP) by PEP carboxylase, a cytosolic enzyme that is highly expressed in the M tissue of C_4 plants (Figure 2). PEP carboxylation is the only biochemical step common to all C_4 plants. The product of PEP carboxylation is the four-carbon organic acid oxaloacetate (OAA), which is rapidly converted into the organic acids malate or aspartate. These acids are then transported to the inner compartment where a decarboxylating enzyme breaks them down to CO_2 and either pyruvate (in most species) or PEP. There are three decarboxylases used in the different lineages of C_4 photosynthesis and these define the biochemical subtypes of C_4 photosynthesis. They are NADP-malic enzyme (which predominates in the NADP-ME subtype of C_4 photosynthesis), NAD-malic enzyme (which defines the NAD-ME subtype), and PEP carboxykinase (PCK) (which is predominant in the PCK subtype). NADP-malic enzyme is the most commonly used decarboxylase, being employed in about two-thirds of the C_4 evolutionary lineages, including both monocots and dicots. NAD-ME also occurs in monocots and dicots and is the second most common type, being predominant in about a third of the C_4 lineages. PEP carboxykinase is the least used, being present in only four grass lineages. During catalysis, NADP-ME and NAD-ME release pyruvate and CO_2. PEP carboxykinase releases PEP and CO_2, but requires one ATP to drive the decarboxylation reaction. The released CO_2 accumulates to concentrations 10 times or more above current atmospheric value, whereas the pyruvate and PEP return to the mesophyll (Figure 2). In PCK plants,

Table 2 Requirements for C_4 photosynthesis

1. CO_2 fixation by PEP carboxylase to form four-carbon organic acids.
2. Coupling of photosynthetic energy (ATP, NADPH, or equivalents) to regenerate PEP.
3. An intermediate metabolite pool (generally organic or amino acids) to transport CO_2 from the compartment of PEP carboxylation to the compartment containing Rubisco and the decarboxylating enzymes.
4. A decarboxylating enzyme to release CO_2 from the intermediate metabolite pool.
5. A compartment in which to concentrate the released CO_2 around Rubisco.
6. A diffusion barrier to slow the loss of CO_2 from the site of CO_2 concentration.

Source: Modified from Leegood R (2002) C_4 photosynthesis: Principles of CO_2 concentration and prospects for its introduction into C_3 plants. *Journal of Experimental Botany* 53: 581–590.

PEP can be used directly by PEP carboxylase in the next round of the cycle. In NADP-ME and NAD-ME subtypes, pyruvate is converted to PEP by pyruvate, phosphate-dikinase (PPDK) in the M chloroplasts using the equivalent of two ATP molecules. The ATP expenditure by PPDK or PCK during PEP regeneration is considered to be the principle cost of operating the C_4 metabolic cycle.

Anatomical Modifications (Kranz Anatomy)

In most C_4 plants, the inner compartment where Rubisco is localized and CO_2 concentrated is a layer of cells that ring the vascular bundles in a leaf. This layer is commonly called the bundle sheath (BS) out of convenience. In many C_4 lineages, the BS is derived from anatomically correct BS cells, however, in numerous others the 'BS' is derived from mestome sheath and water storage cells. C_4 BS cells are typically enlarged relative to what occurs in the C_3 leaf, and are half-filled with numerous large chloroplasts (**Figure 3**). The M cells fill much of the volume of C_3 leaves such that there are typically many (5–20) M cells between adjacent veins (**Figure 3(a)**). In C_4 plants, by contrast, the M cells are typically reduced into one or two concentric layers that ring the BS tissue. Together, the ring of tissue formed by BS and M cells create a wreath-like layer of cells around the vascular bundle termed Kranz anatomy.

Single-Cell C_4 Photosynthesis

In more than 99% of C_4 plants, the C_4 metabolic cycle functions between two distinct cell layers, typically the M and BS tissues. However, a dozen or so species have been identified where the C_4 pathway functions completely within a single cell. These are the single-cell C_4 species. Only three species of single-cell C_4 plants are known in the terrestrial flora, from two evolutionary lineages. In the monospecific *Sueada aralocaspica* lineage, CO_2 is pumped from the outer to the inner end of an elongated M cell (**Figure 4(a)** and **4(c)**). In the two species of the *Bienertia* lineage, CO_2 is concentrated into a ball of chloroplasts in the middle of a large, succulent M cell (**Figure 4(b)** and **4(d)**). The remaining single-celled species all occur in submersed, aquatic plants, either in the monocot family Hydrocharitaceae or in various lines of eukaryotic algae. Unlike terrestrial plants, where the C_4 pathway is

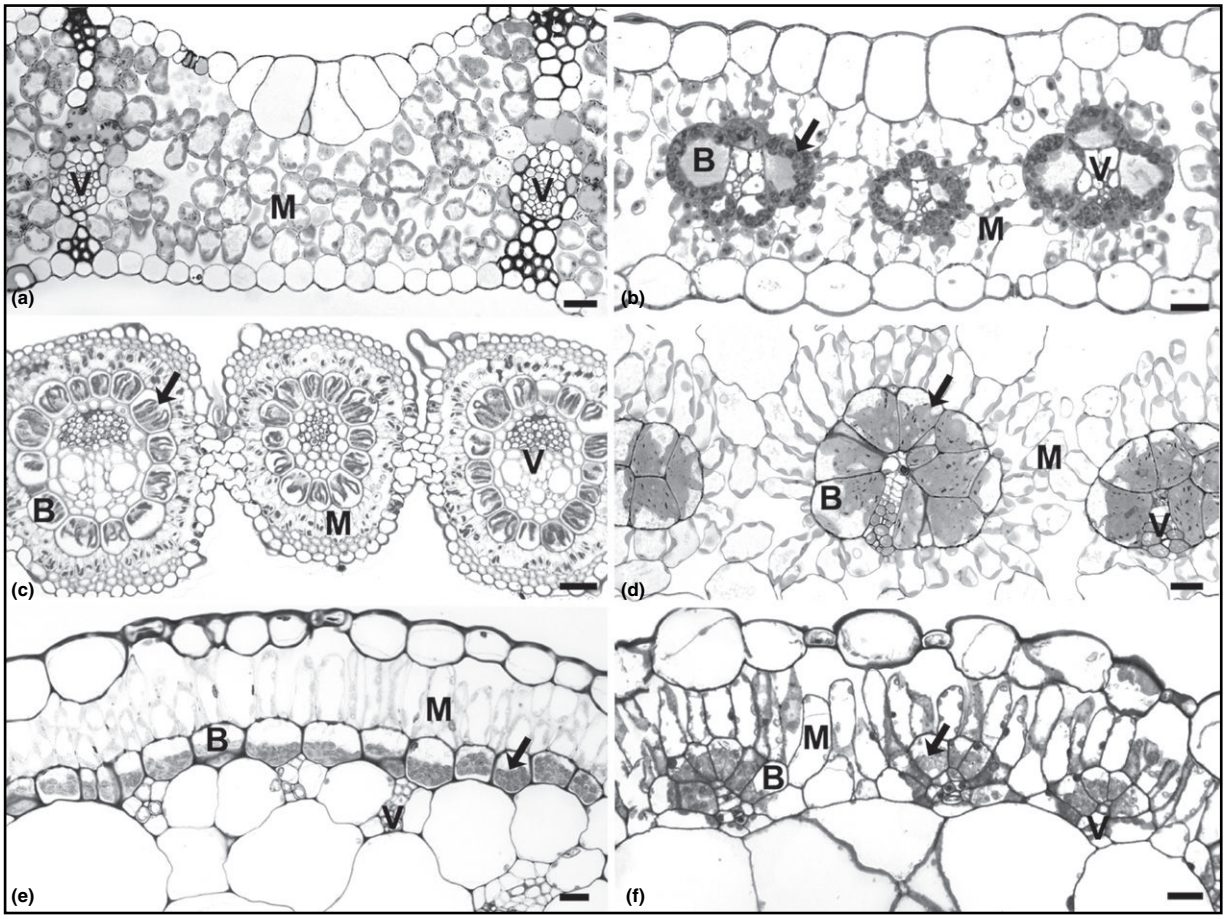


Figure 3 Transverse sections of C_3 and C_4 leaves, demonstrating five of the major anatomical patterns identified in leaves in C_4 plants. (a) The C_3 grass *Brachypodium distachyon*, a model C_3 grass. (b–f) C_4 plants of different evolutionary origins: (b) *Zea mays*, a classic NADP-ME-type grass; (c) *Muhlenbergia montana*, a classic NAD-ME grass; (d) *Atriplex rosea*, an Atriplicoid-type NAD-ME dicot from the Chenopodiaceae; (e) *Salsola kamarovii*, a salsoloid-type NADP-ME dicot from the Chenopodiaceae; (f) *Zygophyllum simplex*, a kochioid-type NAD-ME dicot. Arrows indicate bundle sheath chloroplasts where Rubisco is localized. *Abbreviations*: B, bundle sheath cell; M, mesophyll tissue, V, vascular tissue. The horizontal bars indicate 15 μm.

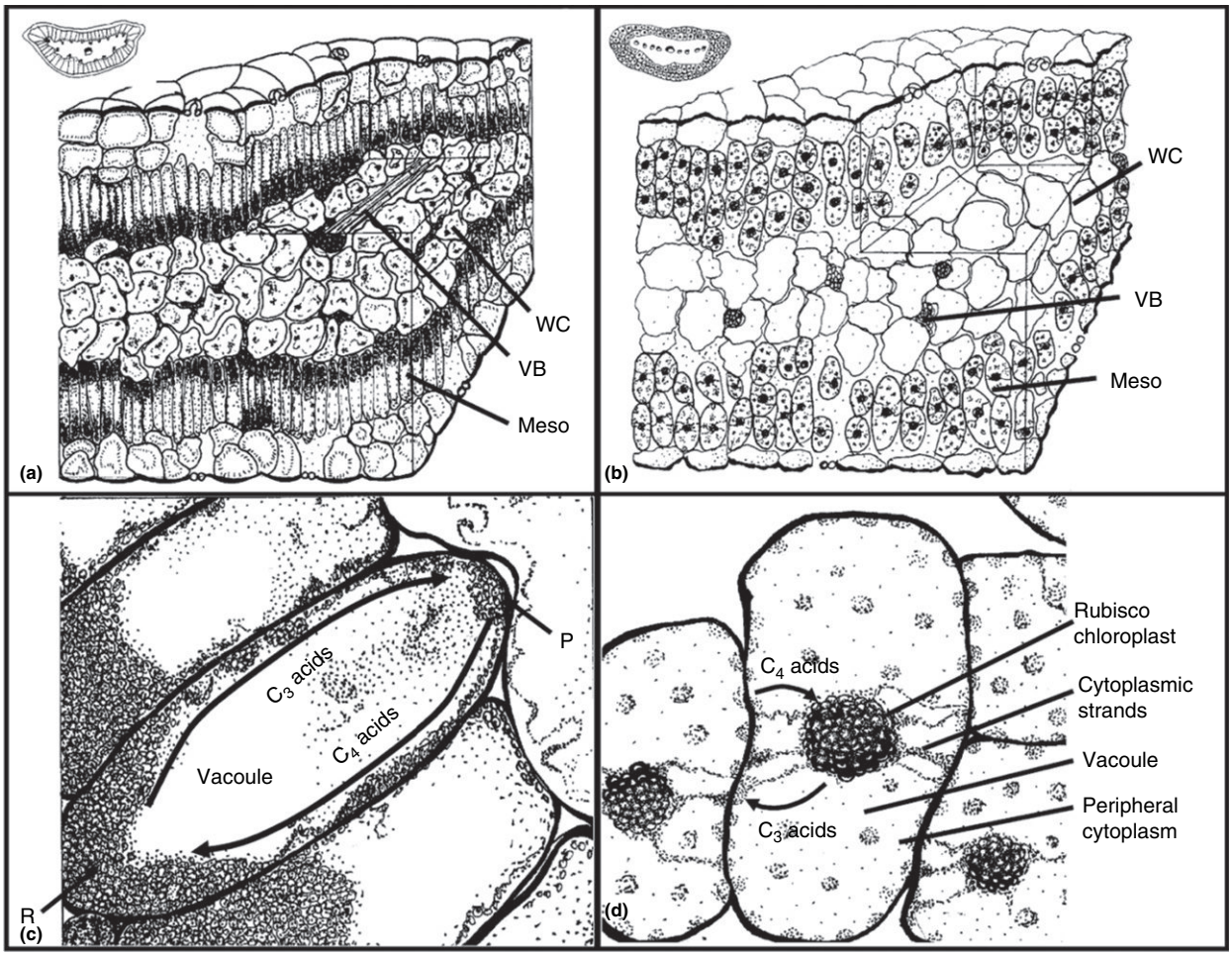


Figure 4 The two forms of leaf anatomy in terrestrial single-cell C_4 plants. (a, c) Leaf anatomy of *Suaeda aralocaspica* showing whole-leaf (panel a) and mesophyll cell (panel c) views. (b, d) Leaf anatomy of *Bienertia cycloptera* showing whole leaf (b) and mesophyll cell (d) views. For *S. aralocaspica*, P indicates the end of the cell where PEP is formed; R indicates where Rubisco is localized. Arrows indicate the movement of C_4 organic acids or the movement of pyruvate (C_3 acids) within the cells. In *S. aralocaspica*, PEP carboxylase forms C_4 acids at the outer (P) end of the cell, whereas decarboxylation occurs at the inner end of the cell (R). In *B. cycloptera*, C_4 acids are formed at the cell periphery, whereas decarboxylation occurs in the chloroplast ball in the middle of the cell. Other abbreviations: meso., mesophyll cells; VB, vascular bundles; WC, water storage tissue. Reprinted from Sage RF (2004) The evolution of C_4 photosynthesis. *New Phytologist* 161: 341–370, with permission from Wiley.

always expressed in aerial tissues, in aquatic C_4 organisms, the C_3 pathway predominates when CO_2 is abundant whereas the C_4 pathway is expressed when CO_2 is deficient. These species are thus considered to be facultative C_4 plants. In the aquatic plants and algae, CO_2 is concentrated into chloroplasts, but there is no distinct anatomical difference from the C_3 mode. For this reason, there has been some debate as to whether they should be termed C_4 species. Since there is CO_2 concentration into a compartment where Rubisco is localized (the chloroplast), a key criterion of C_4 photosynthesis is met and they are now generally considered to be C_4 plants. C_4 algae are generally not included in C_4 discussions, in part because few have been identified and the overall significance of C_4 photosynthesis in algae is poorly known. This review continues the practice of emphasizing C_4 photosynthesis in higher plants (=kingdom Plantae). For a review of C_4 algae, see Bowes (2011).

Diffusion Barriers Preventing CO_2 Loss

Compartmentalization of Rubisco and the decarboxylating enzymes in an inner compartment contributes to CO_2 concentration in two ways. First, the compartmentalization localizes Rubisco to the immediate vicinity of the deacarboxylation reaction. Second, it serves to trap the CO_2 near Rubisco before it can escape back into the mesophyll and intercellular air spaces. This trapping ability requires the presence of a diffusion barrier to slow CO_2 efflux. In numerous grasses, the diffusion barrier is a thick BS wall impregnated with the waxy compound suberin. In other grasses and the dicot C_4 lineages, chloroplasts and the decarboxylating enzymes are located on the centripetal (inner) end of the BS cells, near the vascular tissue, and a large vacuole separates them from the outer edge of the BS cell. CO_2 diffuses 10,000 times slower through water than air, so a large vacuole can serve as an effective diffusion barrier.

In the terrestrial single-celled C_4 species, a large vacuole between the site of decarboxylation and the intercellular air space slows CO_2 efflux (**Figure 4**).

Physiological Significance

Photosynthetic Responses to Atmospheric CO_2 and Temperature

The C_4 pathway has two immediate benefits: first, it raises the CO_2 content in the BS to levels that enable Rubisco to operate close to its maximum activity. Second, the concentration of CO_2 around Rubisco greatly reduces photorespiration. For high rates of photorespiration, CO_2 levels in the atmosphere must be less than about 500 ppm (**Figure 5(a)**). Elevated levels of CO_2 (>1000 ppm) effectively inhibit the rate of RuBP oxygenation, whereas the current level of about 390 ppm allows for modest (20%) photorespiratory inhibition near 25 °C (**Figure 5(a)**). Atmospheric CO_2 levels have been sufficiently high through geological time to effectively suppress photorespiration, with two low CO_2 episodes occurring during the history of land plants (the past 400 million years). A low CO_2 episode is estimated to have occurred during the Carboniferous and Permian periods in Earth's history, centered around 300 mya (million years ago), and recently, in the past 35 million years (**Figure 6(a)**). During these periods, modeled activities of Rubisco oxygenation at 30 °C are predicted to have been significant, particularly in the past 500,000 years, when CO_2 levels were generally below 300 ppm (**Figure 6(b)**). At present, the level of atmospheric CO_2 is anthropogenically enhanced by one-third over the 260–280 ppm that persisted during the Holocene epoch of the past 10,000 years, and are about double the CO_2 levels that predominated in the late Pleistocene (10,000–50,000 years ago). As a result of these lower CO_2 levels in recent geological

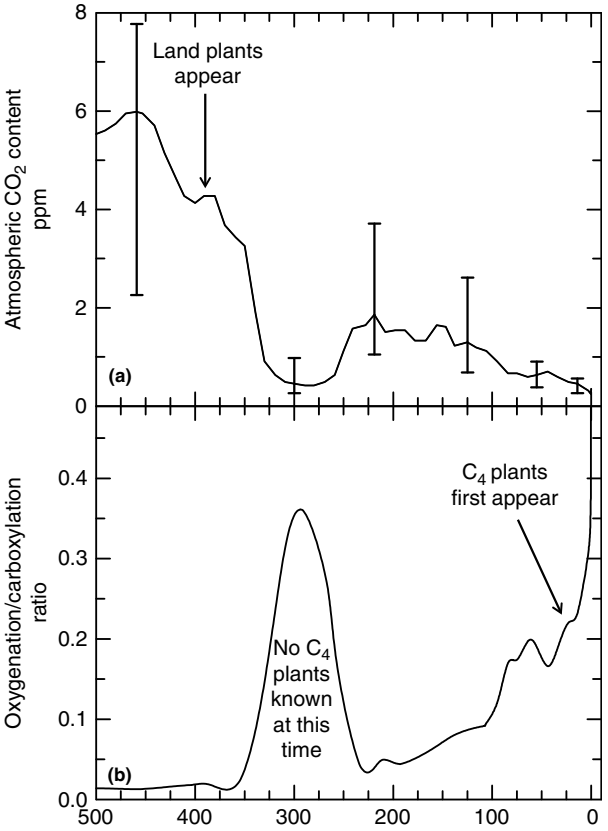


Figure 6 The modeled change in atmospheric CO_2 content (a) and corresponding rates of Rubisco oxygenase activity relative to carboxylase activity over the past 500 million years. The oxygenation/carboxylation ratio was modeled at 30 °C assuming spinach-type of Rubisco and O_2 levels of 21%. Adapted from Sage RF (1999) Why C_4 Plants? In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 3–16. San Diego: Academic Press.

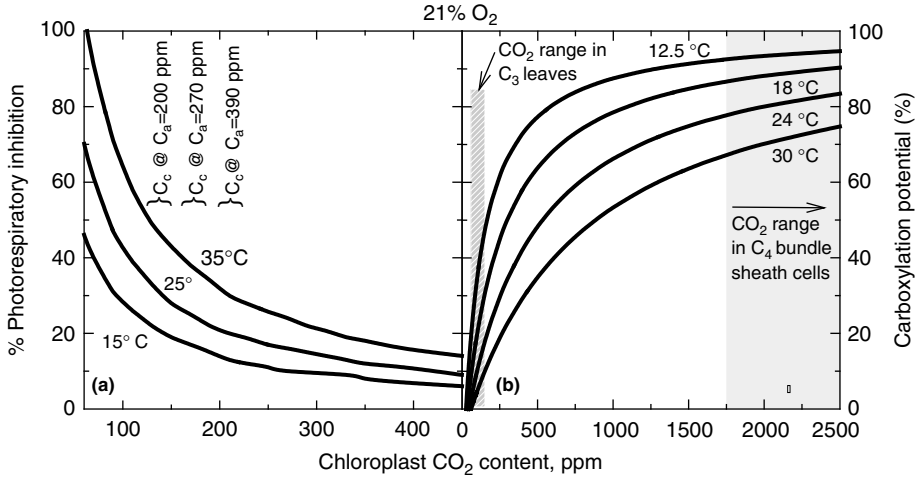


Figure 5 The relationship between photorespiratory inhibition of photosynthesis and chloroplast CO_2 content in C_3 plants at 21% atmospheric O_2. The brackets indicate the range of chloroplast CO_2 levels (C_c) that correspond to atmospheric CO_2 levels (C_a) of the current era (390 ppm), the preindustrial Holocene era (270 ppm), and the late Pleistocene epoch (~200 ppm). (b) The relationship between chloroplast CO_2 and carboxylation potential (V_c–$0.5V_0$) expressed relative to the maximum rate of Rubisco at CO_2 saturation.

time, photorespiration was much higher in C_3 plants than it is today, and this reduced the photosynthetic potential of C_3 relative to C_4 vegetation.

The third requirement for high rates of photorespiration is warm temperatures. At low temperatures ($<15\ ^\circ C$), the rate of photorespiration is low because Rubisco has low specificity for O_2 relative to CO_2. As temperature increases, the specificity of Rubisco for O_2 increases relative to that for CO_2 and the solubility of CO_2 relative to O_2 declines. Together, these responses enhance RuBP oxygenation such that above 35 °C, photorespiration inhibits photosynthesis by more than 30% in current atmospheric conditions and by more than 40% at preindustrial CO_2 levels of 270 ppm (**Figure 5(a)**). By contrast, because of the high level of CO_2 in the BS, photorespiration in C_4 plants is marginal (<6% of the rate of photosynthesis) at all temperatures (**Table 3**). Temperature also affects the kinetics of Rubisco in a manner that has significance for C_4 photosynthesis (**Figure 5(b)**). At low temperature ($<12\ ^\circ C$), Rubisco operates relatively close to CO_2 saturation in the current atmosphere. As temperatures rise, the amount of CO_2 required to saturate Rubisco increases, such in current atmospheric conditions above 30 °C, the enzyme in C_3 plants is operating well below its maximum rate. By raising CO_2 levels in the BS, C_4 plants are able to operate Rubisco near its maximum, substrate-saturated activity (V_{max}) at all temperatures (**Figure 5(b)**).

Together, the ability to reduce photorespiration and operate Rubisco near CO_2 saturation confers substantial photosynthetic advantages for C_4 plants in warm climates and low-CO_2 atmospheres. These differences are well demonstrated by comparing the CO_2 and temperature responses of photosynthesis between C_3 and C_4 species of similar ecological habitat, for example, in prairie grasses or agricultural weeds (**Figures 7** and **8**). Comparisons of the CO_2 response show that C_4 species always have a CO_2 compensation point of photosynthesis (the minimum CO_2 requirement for positive photosynthesis) near 0 ppm; by contrast, the CO_2 compensation point in C_3 plants is near 60 ppm at 31 °C (**Figure 7(a)**). C_4 plants have a steeper CO_2 response of net CO_2 assimilation rate at low CO_2 levels than C_3 species, which combined with the lower CO_2 compensation point, enables them to have a higher photosynthesis rate at low to current CO_2 values (**Figure 7(b)**; **Table 3**). C_4 photosynthesis becomes CO_2 saturated at CO_2 levels near the current atmospheric value, whereas C_3 plants show substantial photosynthetic enhancement with CO_2 enrichment above the current atmospheric value. Thus, at elevated CO_2, C_3 species typically have similar or greater carbon assimilation rates than their C_4 competitors.

The temperature response of photosynthesis in C_3 plants at current CO_2 levels tends to be weaker than that of C_4 species. At current CO_2 levels, C_3 plants have a relatively broad thermal optimum in contrast to the narrow and higher thermal optimum of C_4 photosynthesis (**Figure 8(a)**; **Table 3**). Typically, C_4 plants have substantially greater CO_2 assimilation rates at elevated temperatures ($>30\ ^\circ C$) than C_3 plants of similar ecological habit whereas C_3 species often have higher rates below 20 °C. At double the current CO_2 level, C_3 plants retain their advantage at low CO_2, whereas they have similar photosynthetic capacities at elevated temperature as C_4 species (**Figure 8(b)**). This enhancement of net CO_2 assimilation rate in C_3 species by rising CO_2 is due in large part to the suppression of the high rates of photorespiration that inhibit carbon gain at current levels of CO_2 in C_3 plants. C_4 plants lack this inhibition and so are relatively unresponsive to increasing CO_2 levels above current values (**Figure 8**).

In warm environments, these differences translate into significant photosynthetic and growth advantages for C_4 species in the atmospheric conditions of the past 20,000 years. In crops and weeds, for example, C_4 species in a warm climate typically have greater photosynthetic capacity, higher daily

Table 3 Characteristics of C_3 and C_4 crops and weeds for atmospheric conditions of the late twentieth century[a]

Attribute	*C_3*	*C_4*	*Units*
Maximum photosynthesis rate	20–50	35–75	$\mu mol\ m^{-2}\ s^{-1}$
Maximum daily production rate	10–40	40–80	$g\ m^{-2}\ d^{-1}$
Maximum biomass yield	1–5	3–8	$kg\ DM\ m^{-2}\ year^{-1}$
Photosynthetic thermal optimum	15–30	30–40	°C
Photorespiration/photosynthesis @ 10 °C	8%	1–6%	
@25 °C	25%	1–6%	
@40 °C	40%	1–6%	
$\delta^{13}C$	−20 to −35	−10 to −15	‰ (per mil)
Intercellular CO_2	200–300	100–200	ppm
Maximum leaf nitrogen (N)	150–250	80–160	$mmol\ m^{-2}$
Relative Rubisco content	15–30	5–10	% of leaf nitrogen
Relative PEPCase content	<1	3–6	% of leaf nitrogen
Maximum efficiencies:			
Water use (WUE)	1.5–2.5	3–5	$g\ DM\ kg^{-1}\ H_2O$
Nitrogen use (NUE)	50–280	280–520	$\mu mol\ CO_2\ mol^{-1}\ N$
Radiation use (RUE)	1.7–1.9	2.5	$g\ DM\ MJ^{-1}$

[a]DM indicates dry matter.

Source: Reproduced from Larcher W (1995) *Physiological Plant Ecology*, 3rd edn. Berlin: Springer; Brown RH (1999) Agronomic implications of C_4 photosynthesis. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 473–507. San Diego: Academic Press, and Sage RF and Pearcy RW (2000) The physiological ecology of C_4 photosynthesis. In: Leegood RC, Sharkey TD, and von Caemmerer S (eds.) *Photosynthesis: Physiology and Metabolism*, pp. 497–532. Dordrecht, Netherlands: Kluwer Academic.

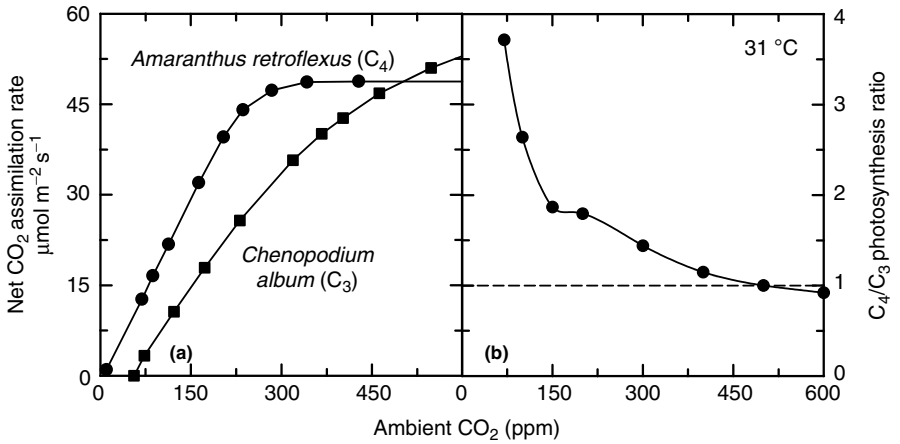


Figure 7 The CO_2 response of photosynthesis in the weedy annuals *Amaranthus retroflexus* and *Chenopodiwn album* at 31 °C. (b) The ratio of C_4 to C_3 photosynthesis for the responses in panel (a).

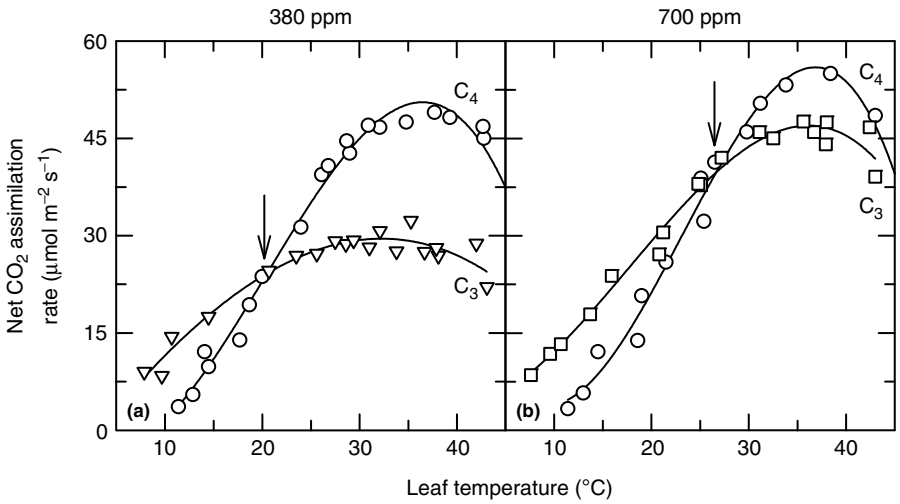


Figure 8 Representative temperature responses of C_3 (*Chenopodium album*) and C_4 (*Amaranthus retroflexus*) rates of net CO_2 assimilation measured at 21% O_2 and either an ambient CO_2 of 380 ppm (panel a) or 700 ppm (panel b). Adapted from Sage RF and Pearcy RW (2000) The physiological ecology of C_4 photosynthesis. In: Leegood RC, Sharkey TD, and von Caemmerer S (eds.) *Photosynthesis: Physiology and Metabolism*, vol. 9, pp. 497–532. Dordrecht, Netherlands: Kluwer Academic, with permission from Springer.

production rates, and yield more biomass during a growing season than do C_3 species (**Table 3**). As a comparison, the world record crop production in terms of dry biomass is close to 100 t per hectare by numerous tropical C_4 grasses; the world record yield from a C_3 crop is near 50 t per hectare.

The advantages of the C_4 pathway do not come without costs. In C_3 plants without photorespiration, assimilating each CO_2 costs 2 NADPH and 3 ATP to fix it into carbohydrate end products. In C_4 plants, 2 NADPH and 4.7 ATP (in PCK subtypes) or 5.7 ATP (in NADP-ME and NAD-ME subtypes) are needed, assuming that 25% of the pumped CO_2 leaks out of the BS (**Table 4**). The higher ATP cost of C_4 photosynthesis reflects the energy required to pump CO_2 into the BS. As the rate of photorespiration increases in C_3 plants, the energy costs of photosynthesis rise, such that they are equivalent to that of C_4 plants at 25–30 °C, and greater above about 30 °C. Differences in the energy requirement of photosynthesis are demonstrated by comparing differences in the light-use efficiency (quantum yield) of C_3 and C_4 plants as a function of temperature (**Table 4**; **Figure 9**). C_3 species at current CO_2 levels have superior quantum yields at cool temperatures, but lower quantum yield than C_4 species at warm temperature. Using quantum yields, it is possible to model the relative performance of C_3 versus C_4 photosynthesis as a function of CO_2 and temperature (**Figures 9** and **10**). As temperature increases, the CO_2 concentration required to favor C_3 over C_4 plants increases as well. Whereas C_3 grasses outperform C_4 grasses below 200 ppm below 15 °C, at 35 °C they need at least 500–600 ppm to show superior performance (**Figure 10**).

Table 4 Energy requirements for carbon metabolism to triose phosphate, and associated quantum yields of C_3 and C_4 plants at 30 °C[a]

	ATP	*NADPH*	*Quantum yield at 30 °C, 350 ppm CO_2*
C_3 photosynthesis			
No photorespiration	3	2	0.078
With 32% photorespiration	5	3.2	0.053
C_4 photosynthesis			
NADP-ME and NAD-ME types			
C_3 cycle	3	2	
C_4 cycle	2	–	
25% leak rate	0.5	–	
3% photorespiration	0.15	0.1	
Total	5.65	2.1	0.055–0.068
PCK types (assuming 25% NAD-ME activity per unit of PCK activity)[b]			
C_3 cycle	3	2	
C_4 cycle	1.25	–	
25% leak rate	0.31	–	
3% photorespiration	0.15	0.1	
Total	4.71	2.1	0.06–0.65

[a]The leak rate refers to fraction of pumped CO_2 that leaks out of the bundle sheath. The percentage photorespiration refers to the rate of photorespiration relative to photosynthesis.

[b]PCK-type plants require one ATP to power each PEP carboxykinase (PCK) reaction in the bundle sheath. This ATP is supplied by oxidation of malate in the bundle sheath mitochondria, with the malate coming from NAD-ME activity at a ratio of one NAD-ME event per three PCK events (Kanai and Edwards, 1999).

Source: Adapted from Kanai R and Edwards GE (1999) The biochemistry of C_4 photosynthesis: Patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 49–87. San Diego: Academic Press.

Enhancement of Water- and Nitrogen-Use Efficiency

Because the C_4 cycle has a high capacity to concentrate CO_2 into the BS, even at low atmospheric CO_2 levels, C_4 plants will have less open stomata than C_3 species with the same photosynthetic rate. This reduces the rate of transpiration in C_4 relative to C_3 plants and leads to significant enhancement in water-use efficiency (WUE). Typically, C_4 species have two to three times the WUE of ecologically similar C_3 species (**Table 3**). Even at very low stomatal apertures that would yield negligible rates of CO_2 assimilation in C_3 plants, C_4 plants are able to have significant rates of photosynthesis. This is most important in hot, arid regions, where evaporative demands are so high that anything but low stomatal apertures could lead to unsustainable rates of water loss. High WUE also assists C_4 species in saline soils, apparently by reducing the volume of transpiration water from which salt must be filtered. In addition to improved WUE, C_4 plants have higher nitrogen-use efficiencies (NUEs) of photosynthesis, typically 1.5–3 times that of C_3 species (**Table 3**). Because C_4 plants operate Rubisco near CO_2 saturation and greatly reduce oxygenase activity, they require about 25% as much Rubisco as C_3 species of equal photosynthetic capacity, and have very low levels of photorespiratory enzymes. As a consequence, C_4 plants have

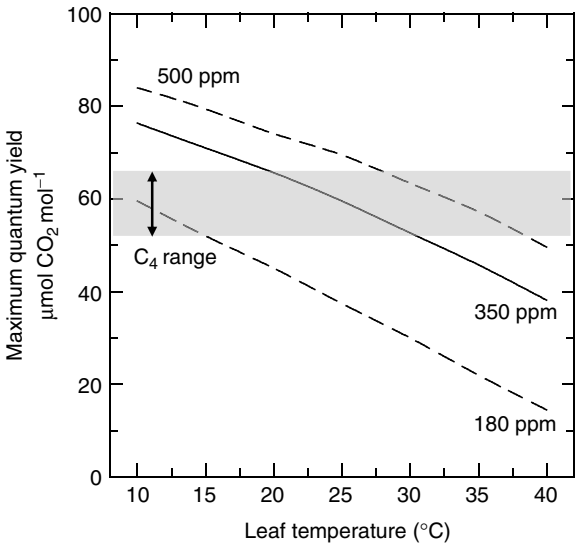


Figure 9 The modeled response of the maximum photosynthetic quantum yield (=light use efficiency) from C_3 plants at atmospheric CO_2 levels equivalent to the last glacial maximum (180 ppm), current (350 ppm, in 1990), and late in this century (500 ppm). The range of quantum yield measured in a variety of C_4 plants is shown for comparison as a gray band. CO_2 variation has little effect on maximum C_4 quantum yield. Adapted from Ehleringer JR, Cerling TE, and Helliker BR (1997) C_4 photosynthesis, atmospheric CO_2, and climate. *Oecologia* 112: 285–299, with permission from Springer.

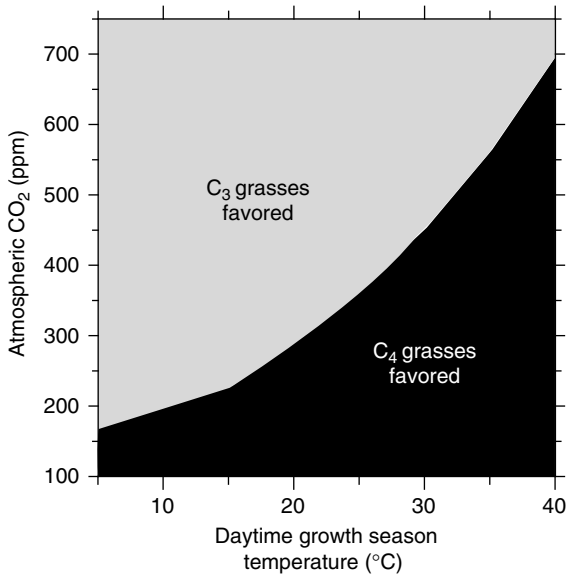


Figure 10 The effect of daytime growth season temperature on the atmospheric CO_2 concentration at which performance of C_3 species equals that of C_4 species. Crossover temperatures were determined as the CO_2 levels needed for equivalent maximum quantum yields at a given temperature. C_4 photosynthetic performance is superior in the black regions of the figure; C_3 photosynthetic performance is superior in the gray region of the figure. Adapted from Ehleringer JR, Cerling TE, and Helliker BR (1997) C_4 photosynthesis, atmospheric CO_2, and climate. *Oecologia* 112: 285–299, with permission from Springer.

about 30% less nitrogen per unit leaf area than ecologically similar C_3 species. Higher NUE promotes superior performance of C_4 grasses in nitrogen-deficient soils, such as those of tropical grasslands. The reduced leaf N arising from higher NUE also reduces the nutritional quality of C_4 leaves, which can have major consequences for grazing herbivores.

C_4 Biogeography

The distinct physiology of C_4 photosynthesis plays a major role in determining where C_4 plants are able to occur and dominate C_3 plants. As a result of their superior photosynthetic performance at high but not cool temperatures, C_4 species can dominate warm environments, but are uncommon in cool environments. They are also uncommon in shaded environments, due in part to their inferior quantum yields below 25 °C. The primary requirements for the success of C_4 plants are therefore warmer temperatures and access to at least moderate light intensities. Aridity, soil nitrogen supply, and salinity also affect the success of C_4 plants, but these are secondary in that they are primarily significant in warm, illuminated habitats but are inconsequential in cold or shaded environments.

Primary Controls: Temperature and Light

The critical temperature parameter affecting C_4 plant distribution is growth season temperature. In the temperate zones, C_4 plants tolerate severe cold outside of the growth season as well as co-occurring C_3 species, and often tolerate low night temperature as well as their C_3 associates. During the growing season, however, daytime leaf temperature must routinely rise above 25 °C for C_4 plants to be present in a community. Where daily temperatures commonly rise above 30 °C during the growing season, C_4 plants generally dominate grass and sedge floras of open landscapes. When plotted on the basis of growth season temperature (which reflects night and morning temperatures as well as the afternoon high), C_4 plants dominate grassland floras and biomass above 20–22 °C. Above growth season averages of 23–24 °C, C_3 grass productivity in open grasslands is very low (typically less than 10% of total biomass). At the other end of the spectrum, C_4 biomass is rare to absent where the temperatures at the peak of the growing season average less than 12–14 °C. Some C_4 species occur in cold climates, but these are restricted to warm microsites where daytime temperatures rise above the regional average. For example, a few dozen C_4 species grow in alpine regions of the world up to 5200 m. These species exploit high solar irradiance in the alpine zone to warm the leaf canopy to >25 °C on sunny days, but they must also tolerate low, often subzero, growth temperatures during cloudy days and at night. These examples demonstrate that cold *per se* does not exclude C_4 species, rather the lack of warm daytime temperatures prevents C_4 plants from ecological success.

The influence of temperature over the distribution of C_4 plants is demonstrated by altitude and latitude transects that examined the proportional contribution of C_4 grasses to regional grass floras. In regions where C_4 grasses dominate lowland grass floras, their representation in the grass flora declines as elevation increases, such that they are uncommon above 2000–3000 m. Along latitude transects, the transition between C_3- and C_4-dominated floras corresponds to temperate latitudes, as is demonstrated by a survey of grass floras of the world's oceanic islands (**Figure 11**). Above >50° latitude, C_4 species are rare or absent, whereas below temperate latitudes (<30°), C_4 grasses dominate the grass flora of any given island. On larger tropical islands, where extensive forests form, a significant number of C_3 forest grasses are present, so that the C_4 contribution may only be 70% of the grass flora. On smaller islands and atolls, where soils tend to be shallow, sandy, and drought-prone, the grasses are often all C_4.

The second requirement for C_4 success is availability of moderate to high light intensity. Typically, this means that C_4 species require 30–50% of full sunlight intensities if they are to dominate a local habitat. In forest canopies where light intensities average less than 20% of full sun on a clear day, C_4 species rarely occur, regardless of latitude. The failure of C_4 plants to dominate in low-light environments results from a combination of factors. Shaded environments tend to be cool (<30 °C at midday), so that high temperatures favoring C_4 species are not frequently encountered. In shade, much of the carbon a plant acquires is obtained from sunflecks (brief episodes of high light). C_4 species are less able to exploit short sunflecks than C_3 species. Also, because shaded environments tend to be cooler and more humid, C_3 plants often have more open stomates than they otherwise would. This allows for higher intercellular CO_2 levels in shade plants, thereby promoting greater photosynthetic efficiency.

The failure of C_4 species to do well in shaded settings has important consequences for the distribution of C_4-dominated life-forms and biomes around the globe (**Tables 5** and **6**). Establishment in woodlands generally requires some level of shade tolerance; if C_4 species do poorly in shade, then it would be difficult for them to establish under the canopy of a forest. This reasoning explains why there are no canopy-forming trees or vines using the C_4 pathway, with the

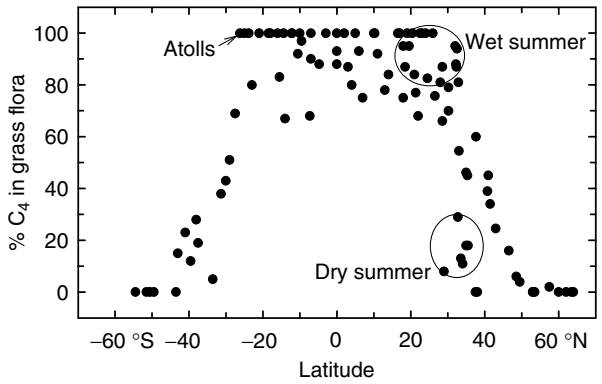


Figure 11 The percentage representation of C_4 grasses in the grass flora of the oceanic islands of the world. The upper oval encloses islands from the western Atlantic/Pacific regions where summers are wet. The lower oval encloses islands from the eastern Atlantic/Pacific where summers are dry. Reproduced from Sage RF, Wedin DA, and Li M (1999b) The biogeography of C_4 photosynthesis: Patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 313–373. San Diego: Academic Press.

Table 5 Principal life-forms of C_4 plants[a]

Life-form	*C_4 families*	*Examples*
Trees	None	
Sub-trees (5–10)[b]	Chenopodiaceae	*Haloxylon* spp. (xerophytic halophytes and psammophytes[c])
	Polygonaceae	*Calligonum* spp. (xerophytic psammophytes)
	Euphorbiaceae	*Euphorbia* (= *Chamaesyce*), (three small Hawaiian trees)
Vines	None	
Shrubs (500)	Amaranthaceae	*Aerva, Gomphrena*
	Chenopodiaceae	*Anabasis, Atriplex, Haloxylon, Salsola, Suada* (mostly xerophytic halophytes and psammophytes)
	Euphorbiaceae	*Euphorbia spp*; (= *Chamaesyce*)
	Polygonaceae	*Calligonum* spp. (xerophytic psammophytes)
Forbs (dicot herbs) (1200)	All dicot C_4 families	Summer desert annuals, xerophytic annual and perennial herbs of low latitudes, tropical and summer-active weeds, psammophytic herbs of low latitudes, tropical beach herbs
Graminoids (5800)	Poaceae	Grasses, mainly of low-latitude and warm temperate summer habitats
	Cyperaceae	Sedges, mainly low-latitude and warm temperate wetlands

[a]Sub-trees refer to species that develop arborescence with age, but never rise above 12 m in height.
[b]Numbers in parenthesis refer to the estimated number of species in each life-form.
[c]Psammophytes are plants of sandy soils.
Source: Reproduced from Sage RF, Wedin DA, and Li M (1999b) The biogeography of C_4 photosynthesis: Patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp 313–373. San Diego: Academic Press.

Table 6 Major biomes of the World with high C_4 plant representation (loosely defined as >25% of the vegetative cover) or low (<1%) C_4 representation[a]

Biome	*Major location*
High C_4 representation	
Tropical, subtropical grassland, and savanna	South and Central America, Africa, India, S.E. Asia, Australia
Warm temperate grassland and savanna	Central and S.E. North America, N. Argentina, Australia
Arid steppe (low to middle latitudes)	S.W. North America, central Asia, Australia, Africa
Beach dunes and bluffs, warm temperate to tropics	Global
Tropical, subtropical wetlands (nonarborescent)	Global, especially South America, central Africa, S.E. Asia
Salt marsh (warm temperate to tropical)	Global, but more temperate owing to mangrove dominance in tropics
Salt desert (<45° latitude)	W. North America, central Asia, central Australia
Hot deserts and semideserts	Southwest North America, Africa, Australia, S.E. Asia
Disturbed ground (low latitudes, low altitude)	Global
Always C_3 dominated (*low C_4 representation*)	
Forests	Global
Tundra	Polar latitudes, high mountains (all latitudes)
Heathlands	N. Europe, Canada, Russia
Cool temperate grasslands and savanna	Canada, N.W. U.S.A., S.E. Europe, S. Russia, Australia, S. Argentina, Tasmania
Montane grasslands	Global, elevations between 2000 and 4000 m
Mediterranean grasslands	California, Chile, S. Europe, Asia Minor, N. Africa, S. Africa, S.W. Australia
Mediterranean-type shrublands (chaparral and mattoral)	California, Chile, S. Europe, Asia Minor, N. Africa, S. Africa, S.W. Australia
Temperate to boreal wetlands, Tiaga	Canada, N. Europe, Russia, New Zealand, Patagonia
Cold deserts and semideserts	W. North America, N. central Asia, Tibet
Salt marshes (>50° latitude)	Global
Mangles (mangrove swamps)	Tropical, sutropical latitudes
Disturbed ground (high latitude and altitude)	Global

[a]Biome list developed from Ecosystems of the World (1977–1990) Goodall, DW ed., vols. 1–16, Elsevier, Amsterdam.
Source: Reprinted from Sage RF, Wedin DA, and Li M (1999b) The biogeography of C_4 photosynthesis: Patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 313–373. San Diego: Academic Press, and Edwards EJ, Osborne CP, Stromberg CAE, Smith SA, and C_4 Grass Consortium (2010) The origins of C_4 grasslands: Integrating evolutionary and ecosystem science. *Science* 328: 587–591.

exception of a few desert shrubs that become arborescent with advanced age, and a few rare species of small trees from Hawaii (**Table 5**). Some C_4 grasses have evolved shade tolerance in warm climates and are found in forest understories; however, these are not dominant and do not occur in the dense shade of old-growth canopies. Thus, in warm to hot climates, the absence of forest and woodland is the single greatest indicator of C_4 abundance and diversity. Where forests

are present, the light environment will typically be too low to support more than a few C_4 species. Where forests are absent, either because of edaphic factors (shallow, nutrient-poor or saline soils), severe abiotic stress (killing drought), mechanical stress (wind, waves, or flooding), or severe disturbance (logging, fire, large animal activity), C_4 plants will be important components of the regional diversity. C_4-rich habitats include beaches along warm temperate to tropical coasts where wave action prevents woody dominance, tropical swamps, and floodplains where flooding prevents dominance by a woody plant canopy, abandoned cropland at low latitudes, salty soils, and warm temperate to tropical grasslands (**Table 6**). For example, in sub-Saharan Africa, the margins of lakes and slow-moving rivers are dominated by papyrus (*Cyperus papyrus*), a C_4 sedge. Here, floodwaters exclude the forest trees. In the African savannas, elephants, giraffes, and fire aid in the reduction of C_3 *Acacia* trees that otherwise would dominate the landscape. Where elephants and fire become common, wooded landscapes are converted to C_4 grassland. Fire control and excessive poaching of elephants permit woodland formation, leading to a loss of the C_4-dominated grasslands.

Secondary Controls: Moisture, Nutrients, and Salinity

Moisture Availability

Aridity is often noted to directly favor C_4 over C_3 biomass, because the higher WUE of C_4 species theoretically promotes their competitiveness in arid situations. In addition, interactions between aridity and other ecological factors, notably fire, are critically important. In the case of fire, aridity slows growth of woody species, thereby increasing their establishment time and the period during which they are vulnerable to fire injury. Aridity also enhances the probability of initial fire establishment, its subsequent rate of spread, and its intensity. In contrast to most woody species, C_4 grasses are typically fire tolerant, as their meristems are protected below ground or in fire-resistant tufts. Forest plants, by contrast, are intolerant of fire. Thus, the success of C_4 grasses in semiarid regions is typically promoted by fire because it reduces the woody competition.

The WUE advantage of C_4 photosynthesis is perhaps most critical in dry locations, such as hot deserts in Africa, Asia, and northern Australia, and semiarid grasslands of low-to-mid latitudes. Here, lack of rainfall combined with high daily temperature creates conditions where the transpiration potential may be greater than what a C_3 plant could support, particularly in the seedling stage. C_4 species not only survive, but they can also thrive under thermal regimes that are harmful to most C_3 species. Consequently, C_4 plants are common if not dominant in the hot arid zones of the Earth. By contrast, the floras of deserts at high latitude or high elevation are exclusively C_3, reflecting the generally poor ability of C_4 species to compete in cold climates, regardless of any aridity that may be present.

Seasonality of precipitation has great importance for the success of C_4 vegetation in mid-latitudes. Temperate grasslands can be almost completely C_3 or C_4 depending on the timing of precipitation. Grasslands of Mediterranean climate zones (southern Europe, North Africa, California, Chile, southwestern Australia, and southwestern Africa) have abundant winter precipitation but lack summer precipitation. The winters are cool and moist, allowing for lush growth of C_3 plants, whereas summers are so dry that plant growth is not possible except near watercourses. If summer precipitation does occur (even a small amount), then plants can be active and many of these are C_4 species. Hence, at similar latitudes, temperate monsoon grasslands with significant summer precipitation are dominated by C_4 grasses, whereas grasslands from Mediterranean-climate zones are dominated by C_3 grasses that are dormant in the summer. For example, the southwest tip of Africa near Capetown has a Mediterranean-type climate and $<10\%$ C_4 species in the grass flora. At Durban, on the eastern coast of South Africa at similar latitudes, the bulk of the precipitation falls during the warm season and the grass flora is 75% C_4. In California, nearly all the precipitation falls in the cool months, and the native grass flora of the valley grasslands is almost completely C_3; by contrast, along the Gulf coast of Texas, in Florida, and the US Atlantic coast of Georgia at similar latitudes, there is an abundance of summer precipitation and the grasslands are C_4 dominated. Oceanic islands show a similar pattern. In the eastern Atlantic and Pacific at temperate latitudes, summer rain is rare and C_4 presence is low (**Figure 11**). On the western edges of these oceans, summer precipitation is abundant as are C_4 grasses.

In southwestern North America, the Mojave desert in California receives 60–90% of its precipitation during the winter, and as a consequence has a high proportion (40%) of C_3 grasses in the grass flora. The Chihuahuan desert in southwestern Texas receives the majority of its precipitation in summer, and C_3 grasses are infrequent (20% of the grass flora). Due in part to the low summer precipitation, the Mojave Desert lacks native grasslands, in contrast to the Chihuahuan desert, where C_4 species dominate the desert grass flora. In addition, two floristic seasons are often present in warm temperate deserts. In the deserts of the American Southwest, winter rains and mild temperatures support a rich flora of C_3 annuals that produce spectacular blooms during wet years. By contrast, summer rains produce a rich flora of C_4 annuals (**Table 7**). Because C_4 annuals typically produce inconspicuous flowers, the blooms of summer annuals fail to attract the attention of tourists that the cool-season C_3 blooms receive.

Salinity

Moderate salinity also favors C_4 species, apparently because the inherently greater WUE reduces the salt load that the plants must deal with in the transpiration stream. In temperate to tropical latitudes, soils of moderate salinity are commonly dominated by C_4 species (**Table 6**). For example, C_4 shrub species in *Atriplex* (saltbush; Chenopodiaceae) can form dense brush on saline soils of central Asia, the Americas, and Australia. In coastal estuaries, salinity establishes a paradoxical trend where low-latitude estuaries are dominated by C_3 species whereas temperate estuaries are dominated by C_4 marsh grasses. In low-latitude estuaries, C_4 grasses can be abundant following a major disturbance; in time, C_3 mangrove trees invade and eventually form dense forests that dominate the estuaries. Mangroves are unique in the plant kingdom in that they are the only tree species that can tolerate both flooding and high salinity. They are, however, freezing intolerant and are excluded from the temperate zone by periodic winter

Table 7 Photosynthetic pathway and mean density over a 10-year period for the 12 most abundant species of summer and winter annuals on the Cave Creek Bajada, near Portal, Arizona (Chihuahan Desert)

Species	*Photosynthetic pathway*	*Density (plants per 0.25 m^{-2})*
Winter annuals		
Haplopappus gracilis	C_3	36.11
Eriastrum diffusum	C_3	30.65
Eriogonum abertianum	C_3	30.25
Descurania pinnata	C_3	4.07
Erodium cicutarium	C_3	3.77
Cryptantha micrantha	C_3	2.55
Gilia sinuata	C_3	2.01
Baileya multiradiata	C_3	1.56
Malacoihrix fendleri	C_3	1.53
Spermolepis echinata	C_3	1.41
Plantago purshii	C_3	1.38
Erigeron divergens	C_3	1.34
Summer annuals		
Bouteloua aristidoides	C_4	6.92
Mollugo verticillata	C_4	3.82
Eriogonum abertianum	C_3	3.60
Mollugo cerviana	C_4	2.46
Pectis papposa	C_4	2.27
Aristida adscensionis	C_4	2.20
Boerhavia spicata	C_4	1.34
Chenopodium fremontii	C_3	0.87
Euphorbia fendleri	C_4	0.84
Baileya multiradiata	C_3	0.72
Tidestromia lanuginosa	C_4	0.65
Panicum arizonicum	C_4	0.44

Source: Reproduced from Guo Q and Brown JH (1996) Interactions between summer and winter annuals in the Chihuahuan desert. *Oecologia* 111: 123–128.

frosts. The resulting lack of mangrove competition allows C_4 grasses to dominate estuary communities in the temperate zone, where they form highly productive saltmarshes. C_4 saltmarshes in the temperate zone typically have low plant diversity because one or two species in the C_4 grass genus *Spartina* effectively exclude most C_3 plants.

Moderate salinity also allows C_4 species to dominate landscapes in climate zones that may otherwise be too cold. *Spartina* marshes predominate in cool-temperate eastern Canada and northern Europe at latitudes where inland there are few if any C_4 species. Along the cool-temperate coastline of northern California, C_4 salt grasses (*Distichilis* spp.) can dominate coastal bluffs, marshes, and cliff ledges because higher salinity allows them to compete with the local C_3 flora for much of the year. Away from the coastline, there are few native C_4 species.

The Evolution of C_4 Photosynthesis

Polyphyletic Origins

Taxonomic Diversity

In the plant kingdom, there are an estimated 7500 C_4 plant species in 10 taxonomic orders and 19 families (**Table 8**). All terrestrial C_4 species are angiosperms, and all occur in the more advanced angiosperm families of the monocots and eudicots. No ferns, gymnosperms, or lower vascular plants are known to employ C_4 photosynthesis. In proportional terms, $<3\%$ of the world's terrestrial flora is C_4; however, this relatively low value in terms of floristic composition is offset by the high productivity and ecological success of many C_4 species across the planet. In terms of net primary productivity (NPP) of land plants, C_4 plants contribute about a fourth of the global terrestrial NPP, far higher than their limited numbers would indicate.

The distribution of C_4 photosynthesis within the angiosperm phylogeny is dispersed, although most C_4 families are clumped in two orders, the Poales and Caryphyllales (**Figure 12**). Most C_4 species are monocots, with half being grasses (Poaceae, with about 400 genera and 4500 C_4 species), and half of the remaining species being sedges (Cyperaceae, with 17 genera and 1300 species) (**Table 8**). Eudicots account for about a quarter of the world's C_4 flora. In the eudicots, most C_4 species occur in the Chenopodiacae (550 C_4 species), Amaranthaceae (250 species), and Euphorbiaceae (250 species).

Evolutionary Lineages

Extensive phylogenetic analysis demonstrates that C_4 photosynthesis is one of the most convergent of complex evolutionary phenomena in the living world. There are currently an estimated 66 distinct evolutionary lineages of C_4 photosynthesis (**Table 8**). The single most prolific family in terms of C_4 origins is the grass family, with 22 distinct origins of the C_4 pathway, followed by the Chenopodiaceae with 10 estimated origins. Sedges evolved the C_4 pathway six times, whereas there are five origins in the Amaranthaceae. Multiple origins are also observed in the Asteraceae (four C_4 lineages), Cleomeaceae (three C_4 origins), and Nyctaginaceae (two C_4 origins). Numerous genera also have evolved the C_4 pathway on multiple occasions. *Suaeda* (Chenopodiaceae) has four distinct lineages, including the two single-celled C_4 lineages found in land plants; *Cleome* (Cleomeaceae) has three C_4 lineages, whereas *Eleocharis* (Cyperaceae), *Mollugo* (Molluginaceae), and *Flaveria* (Asteraceae) have two each. A prolific ability to evolve C_4 is not correlated with species numbers. *Euphorbia* has evolved C_4 photosynthesis just once and has more than 200 C_4 species; whereas *Cleome* has only one C_4 species in each of its three lineages.

With the identification of the many C_4 lineages, it is now possible to delineate the geographic regions where the C_4 pathway evolved by identifying the distributions of the species that branch at nodes between full C_3 and C_4 species on a phylogeny. Centers of origins occur on all low-to-mid latitude continents, largely in semiarid to arid regions having hot summers that receive episodic precipitation from monsoon rains (**Figure 13**). The most prolific centers of origin occur in central Asia, where *Calligonum* (Polygonaceae) and all 10 of the Chenopodiaceae lineages arose, and southwestern North America and adjacent Mexico, where at least eight lineages arose.

When Did C_4 Plants Evolve?

Molecular clock analysis of phylogenetic data also indicates when the C_4 pathway arose. The oldest C_4 lineage estimated

Table 8 The occurrence of C_4 photosynthesis in higher plant orders and families

Order	*Family*	*Number of evolutionary lineages*	*Numberof C_4 genera*	*Number of C_4 species*	*% of all C_4 species*[a]
Dicotyledoneae (subclass)					
Asterales	Asteraceae	4	7	140	2.0
Brassicales	Cleomeaceae	3	1	3	<0.1
Caryophyllales	Aizoaceae	1	5	15	0.2
	Amaranthaceae	5	10	250	3.3
	Caryophyllaceae	1	1	20	0.3
	Chenopodiaceae	10	45	550	7.3
	Molluginaceae	1	1	2	<0.1
	Gisekiaceae	1	1	7	0.1
	Nyctaginaceae	2	3	80	1
	Portulacaceae	1	1	100	1.3
Lamiales	Acanthaceae	1	1	15	0.2
Malpighiales	Euphorbiaceae	1	1	250	3.3
Polygonales	Polygonaceae	1	1	150	2
Uncertain (Solanales?)	Boraginaceae	1	1	100	1.3
Zygophyllales	Zygophyllaceae	2	4	50	0.7
	Scrophulariaceae	1	1	4	<0.1
Total dicot	**16**	**36**	**84**	**~1700**	**23**
Monocotyledoneae					
Alismatales	Hydrocharitaceae	2	2	4	<0.1
Poales	Cyperaceae	6	17	~1300	17
Poales	Poaceae	22	372	~4500	60
Total monocot	**3**	**30**	**389**	**~5800**	**77**
Total C_4	**19**	**66**	**473**	**~7500**	**100**

[a]The percentage of all C_4 refers to the number of C_4 species in a taxonomic group divided by 7500, the estimated total number of C_4 species.

Source: Estimates are from Sage RF, Li M, and Monson RK (1999a) The taxonomic distribution of C_4 photosynthesis: Patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 551–584. San Diego: Academic Press; Sage RF, Sage TL, Pearcy RW, and Borsch T (2007) The taxonomic distribution of C_4 photosynthesis in Amaranthaceae *sensu stricto*. *American Journal of Botany* 94: 1992–2003; Christin PA, Osborne CP, Sage RF, Arakaki M, and Edwards EJ (2011) C_4 eudicots are not younger than C_4 monocots. *Journal of Experimental Botany* 62: 3171–3181; Sage RF, Christin PA, and Edwards EJ (2011) The C_4 plant lineages of planet Earth. *Journal of Experimental Botany* 62: 3155–3169, and Sage (unpublished). Species numbers derived from Sage *et al*. (1999a) and updated using Tropicos.org, Theplantlist.org, the Flora of North America volume 4 (Flora of North America Committee, 2003, Oxford University Press, Oxford, UK), and Bruhl JJ and Wilson KL (2007) Towards a comprehensive survey of C_3 and C_4 photosynthetic pathways in Cyperaceae. *Aliso* 29: 99–148.

by molecular clock methods is the grass subfamily Chloridoideae, which is approximately 30–35 million years old (**Figure 14(a)**). The oldest estimated eudicot lineage is the Chenopodiaceae subfamily Caryoxylonae, estimated to have arisen about 25 mya. The remaining C_4 lineages evolved over the subsequent 25 million years, with a greater number arising 20–25 mya, and during the past 10 million years. The youngest lineages are estimated to have arisen in the genera *Neurachne* (Poaceae), *Eleocharis* (Cyperaceae), and *Flaveria* (Asteraceae) in just the past 2–3 million years, indicating that C_4 evolution may still be an on-going process. *Neurachne* and *Flaveria* contain species that have intermediate characteristics between C_3 and C_4 photosynthesis and thus have become important groups for studying the dynamics of C_4 evolution.

Why Did C_4 Photosynthesis Evolve?

The leading hypothesis explaining why C_4 photosynthesis evolved is the carbon starvation hypothesis, which proposes the C_4 pathway originated in response to reductions in photosynthetic potential associated with high rates of photorespiration and low availability of CO_2. A number of lines of evidence support this hypothesis. First, all the postulated lineages of C_4 photosynthesis arose in the past 30–35 million years, after atmospheric CO_2 is estimated to have declined from levels where photorespiration is nil to levels where it would have been a significant inhibition on C_3 plants in hot climates (**Figures 6** and **14(b)**). CO_2 levels are estimated by geochemical models and proxy measurements to have been high (above 1000 ppm) between 40 and 100 mya, but declined to near current levels between 40 and 25 mya (**Figure 6**). Accompanying this CO_2 decline was a general drying of the planet, which expanded open, dry landscapes at low latitudes by the early Miocene. Second, the postulated centers of origin for the C_4 lineages correspond to regions where growing conditions were hot and episodically dry or saline (**Figure 13**). Examination of the closest C_3 relatives of C_4 lineages, and of C_3–C_4 intermediate species, consistently indicates that C_4 photosynthesis arose in summer active species that exploited episodic monsoon precipitation or salinized water supplies. Drought or elevated salinity are common in the habitats of these species, and evaporative demand is high due to the hot, generally dry climate. The presence of episodic summer moisture is significant because it allows for growth and photosynthesis during the hottest times of the year. If there were no summer moisture, the vegetation would be dormant, as occurs during the dry seasons at low latitude.

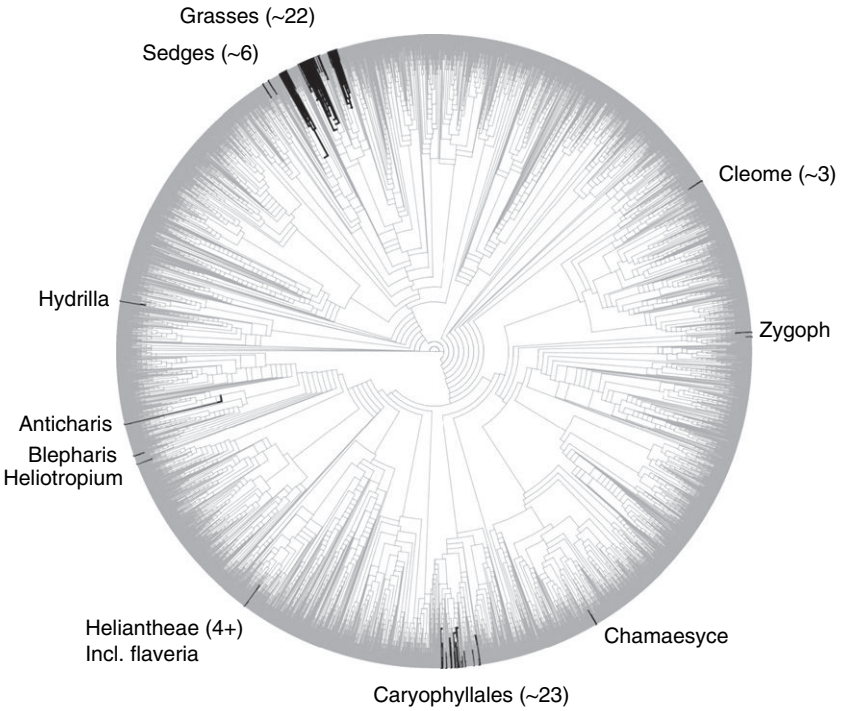


Figure 12 The distribution of C_4 evolutionary lineages in the angiosperm phylogeny. C_4 lineages are indicated in black branches, C_3 lineages by gray branches. The angiosperm phylogeny was generated from 9412 species, including C_4 species from 47 of the estimated 66 lineages. Numbers beside lineages indicate the number of distinct C4 origins in the lineage. Reprinted from Sage RF, Christin PA, and Edwards EJ. (2011) The C_4 plant lineages of planet Earth. Journal of Experimental Botany 62: 3155–3169, with permission from Oxford University Press.

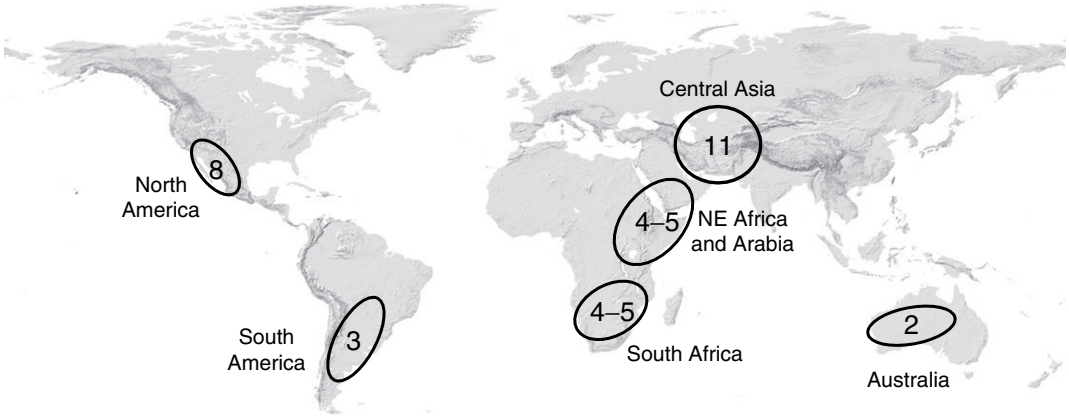


Figure 13 The centers of evolutionary origin for 33 of the 66 known C_4 lineages in the world. Lineages represent 32 eudicot lines and the grass genus *Neurachne*. Centers of origin for the 33 lineages not represented here remain to be determined. Reprinted from Sage RF, Christin PA, and Edwards EJ (2011) The C_4 plant lineages of planet Earth. Journal of Experimental Botany 62: 3155–3169, with permission from Oxford University Press.

Drought stress, salt stress, and high evaporative demand reduce stomata aperture and lower intercellular CO_2 level in the leaf. This, in combination with low atmospheric CO_2 and high temperature ($>35\ ^{\circ}C$), would keep CO_2 levels in C_3 chloroplasts low, thus promoting very high rates of photorespiration ($>50\%$ of the rate of photosynthesis) in atmospheres below 250 ppm CO_2. Such high levels of photorespiration would have inhibited the competitive ability and survival of C_3 plants, thereby opening up landscapes in extreme habitats where evolutionary experiments in carbon scavenging could proceed with minimal competition from C_3 vegetation. More importantly, however, high levels of photorespiration created a resource, photorespired CO_2, that could be exploited to offset the costs of photorespiration. Evolutionary selection for mechanisms to recapture photorespired CO_2 is now thought to have initiated the C_4 evolutionary process.

C_3–C_4 intermediate species from numerous evolutionary lineages demonstrate that C_4 photosynthesis arose from a series of innovations that initially served to capture and refix photorespired CO_2. Key early steps in the evolution of C_4 photosynthesis are the restriction of glycine decarboxylase to

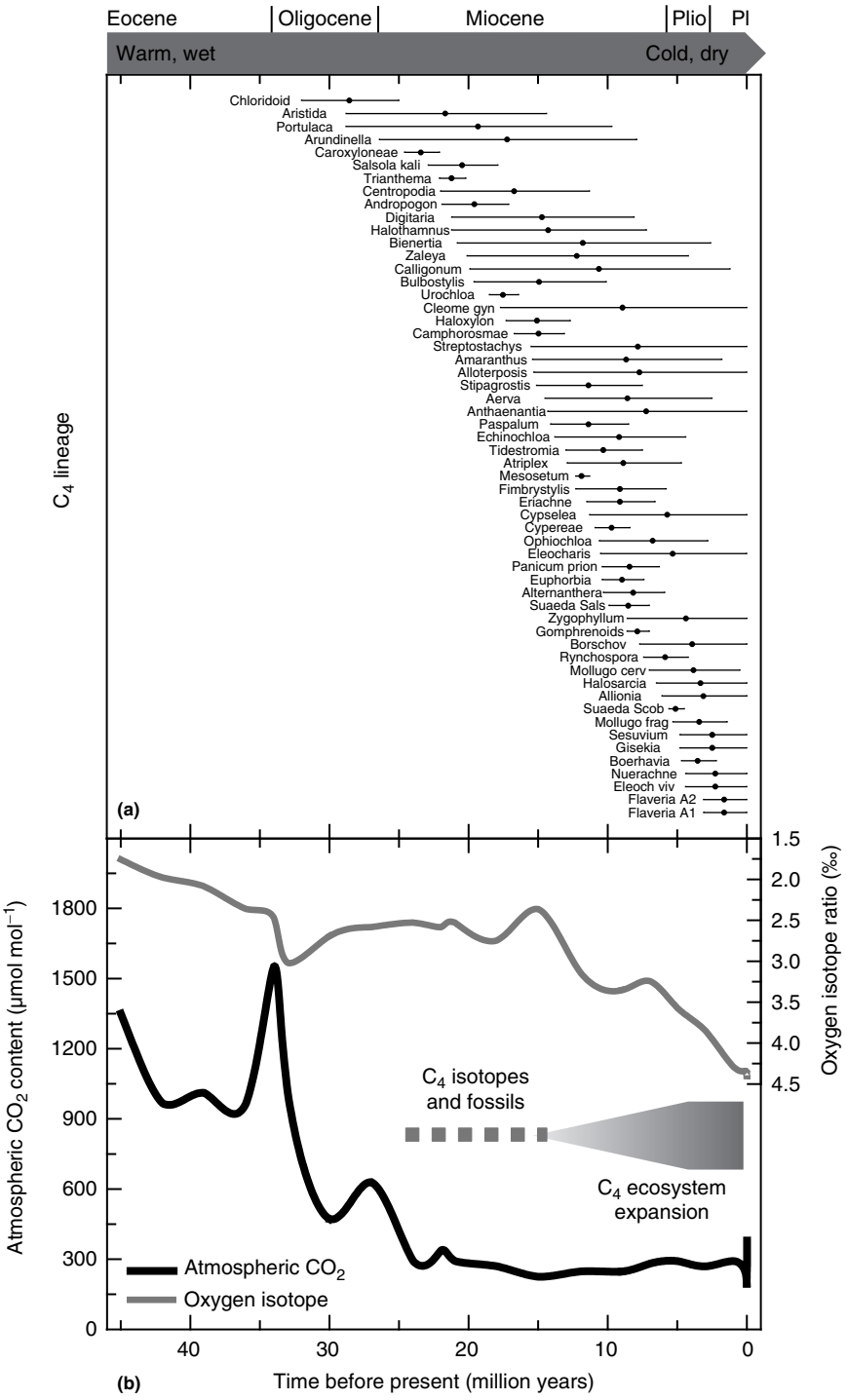


Figure 14 (a) The age of 56 lineages of C_4 photosynthesis estimated from molecular phylogenetic data. The dot in each bar represents the median estimate for a lineage whereas the bars show the estimate ages of the stem and crown nodes in the phylogeny. Names indicate the lineage name as listed in Sage *et al.* (2011). (b) The change in atmospheric CO_2 concentration (black line) and oxygen isotope ratios in foraminifera sediments (gray line) since the early Eocene epoch. The corresponding appearance of C_4 fossil and isotope evidence is shown indicating the presence (dotted line) or rise to dominance (green triangle) of C_4 vegetation at low to mid-latitudes. Larger oxygen isotope ratios indicate a colder (high latitude) and drier (low latitude) planet. Abbreviations: Plio, Pliocene; Pl, Pleistocene. Reprinted with permission from Sage RF, Sage TL, and Kocacinar F (2012) Photorespiration and the evolution of C_4 photosynthesis. *Annual Review of Plant Biology*. DOI: 10.1146/annurev-arplant-042811-105511.

the BS cells and increase in the number of BS chloroplasts. Glycine decarboxylase is the enzyme that releases CO_2 during photorespiration. These changes localize the release of photorespired CO_2 to the BS of C_3–C_4 intermediates, allowing Rubisco in the BS chloroplasts to refix it with high efficiency. The C_4 metabolic cycle is then upregulated to capture photorespired CO_2 leaking out of the BS cells, and full C_4 photosynthesis evolves when M expression of Rubisco and the PCR cycle is curtailed. When the evolutionary sequence indicated by the C_3–C_4 intermediate species is considered, the C_4 pathway is most directly an adaptation to high rates of photorespiration, rather than to high temperature, drought, or salinity. These environmental factors play a contributing role by promoting high rates of photorespiration, but are not directly the reason for C_4 photosynthesis. Habitats where these factors act in concert would have been the most likely to have produced C_4 plants, and consistently, most of the C_4 lineages originate in extreme environments where these factors occur in combination.

Rise to Ecological Dominance

While molecular clock evidence indicates when C_4 lineages began and taxonomically diversified, they do not address when C_4 plants became important elements in ecological communities. This information has to come from either fossilized tissues of identifiable C_4 plants, or the carbon isotope signatures derived from remains of C_4 plants or their herbivores. C_4 photosynthesis discriminates less against heavy isotopes of carbon (^{13}C–CO_2) than does C_3 photosynthesis, such that C_4 tissues have eight to 15 more ^{13}C atoms per hundred thousand ^{12}C atoms than do C_3 plants. Using mass spectrometry, these differences can be readily detected in plant remains, or any material that originated from photosynthetic carbon. Thus, fossilized organic matter, soil carbonates, and animal remains can be assessed to determine whether C_3 or C_4 species were common to dominant on a landscape at the time of fossilization.

Fossilized soil carbonates and grass phytoliths (silica bodies) with C_4-like carbon-isotope ratios appear in Africa and the Americas between 16 and 23 mya, indicating C_4 plants become locally common (but not dominant) in open habitats by the Mid-Miocene epoch. The oldest known C_4 leaf fossils come from the Ricardo formation of southern California, dated from the middle Miocene at approximately 12.5 mya. These fossils show distinct C_4 Kranz-type leaf anatomy and have isotopic ratios indicative of C_4 species. Probable C_4 leaf fossils have also been found in the Fort Ternan formation of Kenya, dated at 14.5 mya.

While C_4 grasses become common by the mid-Miocene, there is no evidence that they dominate landscapes until after 10 mya. Beginning about 10 mya, C_4-dominated grasslands expanded across East Africa, southern Asia, and the interior of both North and South America. Both soil and herbivore remains from these areas show an increase in carbon isotopic ratios from C_3-like to C_4-like values between 10 and 5 mya. In East Africa and Pakistan, isotopic shifts are dramatic, indicating near-complete replacement of C_3 with C_4 vegetation in as little as 2 million years. In temperate latitudes (China, South America, and the North American Great Plains), the increase in the carbon isotope ratio is less on average, indicating partial replacement of C_3 vegetation by C_4 species.

The expansion of C_4-dominated grasslands 5–10 mya appears to have resulted from climate change events that caused an increase in the seasonality of precipitation and overall aridity, such that summer monsoons would have alternated with pronounced dry seasons. There is little evidence for a significant drop in atmospheric CO_2 when the C_4 grasslands expand, so declines in CO_2 are not thought to be important drivers. Atmospheric CO_2 was already low in the late Miocene, however, so it could have been an important contributing factor that allowed other environmental agents such as increased seasonality to trigger C_4 grassland expansion. Increased aridity and seasonality of precipitation for grassland expansion would also have promoted wildfire, which would have favored fire-tolerant C_4 grasses over woody C_3 vegetation. Relatively high productivity of C_4 grasses would have produced an abundance of fuel to support frequent, killing fires during the dry season, whereas low CO_2 would have slowed establishment and succession of the woody vegetation. Consistently, late-Miocene deposits show increased soot and charcoal during the time when C_4-grasslands expanded.

Consequences of C_4 Plant Evolution for Animal Diversity

The evolution of C_4 photosynthesis produced a new plant functional type that could exploit opportunities created by low atmospheric CO_2 and more frequent drought. In a direct sense, this led to the diversification of 7500 new species, which is not much when compared to the estimated 250,000 C_3 species on the planet. As the novel C_4 species expanded across low-to-mid latitudes, however, they created novel conditions that triggered an evolutionary cascade, which in time led to the functional and trophic diversity in the grasslands biomes of the modern biosphere. Amongst the most significant changes was the expansion of open, low stature vegetation where sustained running was possible yet there was limited access to shelter. These open spaces would have favored novel traits in animals, such as herding in large numbers for defense against predators, extended strides for rapid locomotion, and dental structures adapted to grass. In response to the intensification of fire cycles promoted by C_4 grasses, fire-tolerant species evolved with traits such as belowground storage organs and thick, heat-resistant bark. High productivity by the C_4 grasses also created an abundant food supply for grazing herbivores; however, much of the C_4 leaf protein is sequestered in BS tissues that are difficult to digest due to the thick BS wall. In response to this digestive challenge, megaherbivores with specialized digestive physiologies and extensive fermentation evolved to extract sufficient energy and protein from the C_4 forage. Among the megaherbivores that exploited the C_4 grasslands were the ruminant ungulates and the equines, whose large herds were maintained by the productive C_4 grasslands. The diversification of megaherbivores produced a vast resource of meat, which in turn promoted the evolutionary diversification of megacarnivores

such as the large cats. The rise of our own species may also be a response to the expansion of C_4 grasslands. C_4 grasses would have provided little for hominids to directly consume, because C_4 leaves are difficult to digest and the seeds are small and easily dispersed. Grazing herbivores, by contrast, would have been a valuable resource if hominids could have survived on the grassland and developed the ability to hunt or trap them. Expansion of C_4 grasslands in the late Tertiary period between 10 and 3 mya is thus hypothesized to have created an open habitat that selected for upright stature, teamwork, the use of weapons, and improved communication skills, which then led to increased brain size and cognitive ability. It is possible, therefore, that the rise of C_4 grassland formed the nursery for the evolution of rationale intelligence.

Global Change and the Future of C_4 Photosynthesis

Climate Change and Atmospheric CO_2 Enrichment

If C_4 plants evolved in response to past reductions in CO_2, then it follows that future increases in atmospheric CO_2 due to human activities would reduce the advantage of C_4 photosynthesis. Photosynthesis and growth of C_3 plants is enhanced 19–46% by a doubling in atmospheric CO_2; in C_4 plants, the stimulation is typically less than 10% if it occurs at all. As atmospheric CO_2 levels increase, cooler, more temperate environments would be affected initially, whereas C_4 species in warm, arid habitats would lose their advantage as atmospheric CO_2 levels increase 30–50% above the current value (**Figure 10**). There are, however, a number of considerations that indicate many C_4 plants will survive and continue to prosper in a world with higher CO_2.

First, rising CO_2 does not harm C_4 plants but does provide some benefits. Direct photosynthetic enhancement of C_4 photosynthesis by rising CO_2 is minor except during drought stress. The more significant impact of rising CO_2 on C_4 plants is that it reduces stomatal aperture, which lowers transpiration, improves WUE and delays the onset of drought. The degree of stomatal closure is similar in C_3 and C_4 plants, such that in systems where water is limited by high evapotranspiration and drought, similar responses of C_3 and C_4 species may be expected. However, water limitations are more likely in summer and semiarid landscapes where C_4 species currently dominate, so the impact of elevated CO_2 on drought could be more significant when and where C_4 plants are active. Because of this, C_4 plants could benefit as much from rising CO_2 as C_3 species in cooler environments.

Second, climate warming alone will generally favor C_4 more than C_3 plants, because C_4 plants are generally adapted for warmer situations than potential C_3 competitors. C_4 photosynthesis has a higher thermal optimum than C_3 photosynthesis, and C_4 plants are adapted to handle heat stress due to their current existence in generally hot climates. Third, for plants to exploit rising CO_2, they require sufficient supplies of soil nutrients such as nitrogen and phosphorous. On soils' very low nutrient availability, the CO_2 stimulation of C_3 photosynthesis is nil. Nutrients are often limiting in tropical soils, in part because C_4 grasses are effective at sequestering nutrients in their rootstocks and leaf litter, and significant nutrients loss may result from fire. Hence, C_4 grasslands could maintain themselves where soil infertility prevents C_3 species from exploiting elevated CO_2. Finally, drought is a common limitation in ecosystems where C_4 species occur and here, climate change effects on precipitation patterns could be paramount. If summers become wetter, summer-active C_4 species could be favored, as long as increases in precipitation are not extensive enough to allow for woodland establishment. Alternatively, if winter and spring receive more moisture, C_3 species could be favored. Mediterranean zones that develop a summer monsoon could see a dramatic rise in the cover of C_4 species, with dire consequences for native shrub vegetation should the grasses accelerate fire cycles. Warmer temperatures and more frequent aridity will also enhance fires regimes in (sub)tropical systems to the benefit of C_4 vegetation.

Because past climate change has been associated with CO_2 rise, paleoecological studies indicate how future C_3/C_4 dynamics might develop. Between 20,000 and 10,000 years ago, temperate zone climates warmed 5–10 °C and CO_2 levels rose 30%, thereby providing a similar situation to that predicted to occur in coming centuries. Theoretically, the CO_2 rise should have had a dramatic effect on C_4 relative to C_3 distributions, because the low-CO_2 level of 20,000 years ago (180 ppm) would have been far more limiting for the performance of C_3 relative to C_4 plants. Paleoecological studies indicate that in tropical climates, there was a wider distribution of C_4 species 20,000 years ago relative to 10,000 years ago and today. In central Africa and South America, isotopic and pollen data indicate that grasslands were more widespread 20,000 years ago than at present, which is consistent with a low-CO_2 effect favoring C_4 plants. However, these regions were drier than at present, which could also explain the pattern. In temperate zones, little CO_2 effect on C_3 versus C_4 plant distribution is apparent during this time. In southwestern North America, for example, the distribution of C_3- and C_4-dominated grass and shrub vegetation appears to remain stable over the past 20,000 years, possibly because the seasonality of precipitation has also changed little over this time. In contrast to the American Southwest, substantial expansion of C_4 grasses occurred at the end of the last Ice Age on the plains grasslands of central North America. Here, C_4-dominated grasslands expanded northward between 12,000 and 8000 years ago, a time of little CO_2 change, but considerable climate warming.

Land-Use Change and C_4 Plants

Humans currently exploit much of the world's land area where C_4 plants occur; this exploitation will intensify as population and economic growth increase demands for food, fuel, and fiber. Because of this, the future of C_4 plants will largely depend on land-use decisions made by future society. In recent decades, land use decisions have been largely detrimental for native C_4 species, although a few weedy or agronomic species have done very well. Loss of natural C_4 grasslands has been severe in low-to-mid latitude areas due to the conversion of natural landscapes to timber plantations, pasture, and farms. Low-latitude pastures largely comprise C_4 grasses, but typically just a select few that are preferred by cattle. In North America,

the tall grass prairie biome has been largely replaced by row crop agriculture and pasture. Where bluestems and switchgrass (*Schizychyrium*, *Andropogon*, and *Panicum* spp.) once formed a sea of C_4 grasses stretching from Illinois to Kansas, a new C_4 species – cultivated maize – has become dominant. A second great C_4 biome lost in North America is the wire-grass/long-leaf pine savanna that once dominated the coastal plain of the southeastern USA. More than 98% of this biome was destroyed in the 1800s by grazing, logging, fire suppression, agricultural development, and more recently, timber plantations.

Overgrazing by cattle has had a particularly widespread impact on C_4-dominated grasslands in semiarid regions of Africa, the Americas, and Australia. Many of these areas have converted to dense shrublands because cattle selectively removed the C_4 biomass and the associated fuel layer needed for regular fires (**Figure 15(a)** and **15(b)**). As a result, shrubs species that were no longer suppressed by grass competition and fire were able to displace the grassland. Rising CO_2 aids the establishment of shrubs in grasslands, because it accelerates root and shoot growth of young seedlings. Woody species typically grow deep taproots below the more shallow, fibrous root systems of grasses, and once they can do this, below-ground competition between shrubs and grasses is reduced. More rapid shoot growth allows for a faster escape from the shade of the grasses, but more importantly, it elevates the shrub canopy above the kill zones of grass fires. In coming decades, continued growth in atmospheric CO_2 levels and grazing intensity could accelerate loss of low-latitude grasslands, unless humans actively suppress shrub growth and employ management practices that favor C_4 grasses. Rather than restoring native C_4 diversity, however, grassland restoration has often relied on a limited number of selected C_4 grasses in order to improve grazing conditions for cattle.

C_4 species that will do well in the future are agricultural and invasive weeds that adapted to human disturbance. C_4 plants disproportionally make up the world's worst weeds of agricultural landscapes, and in low latitudes, they are severe bioinvaders that are capable of converting a range of landscapes to exotic C_4 grasslands (**Figure 15(c)** and **15(d)**). In the presence of human-caused fires, exotic C_4 grasses are able to intensify wildfire regimes to the detriment of many types of native vegetation. These grasses rapidly colonize disturbed landscapes and with their high productivity, can quickly produce fuel loads that drive hot, killing fires. With more widespread disturbance and natural resource exploitation, future decades will provide greater opportunities for weedy C_4 species to expand and suppress natural vegetation in both C_3- and C_4-dominated ecosystems.



Figure 15 Photographs illustrating the potential fate of native C_4 grasslands and tropical forests in a changing global environment. (a) A C_4 range land with high cover of C_4 desert grasses and sporadic shrub cover, in western Texas. This landscape is representative of much of the rangelands of the southwestern USA prior to 1800. (b) Shrub domination of former C_4 grassland in west-central Texas, USA. (c) Invasion of C_4 grasslands into pristine rainforest in New Caledonia. The C_4 grassland (light area) replaced the forest after a fire created the triangular patch on the left side of the ridgeline. With the next fire, this patch will expand and further reduce the forest area. Reproduced from Sage RF and Kubien DS (2003) *Quo Vadis* C_4*?* An ecophysiological perspective on global change and the future of C_4 plants. *Photosynthesis Research* 77: 209–225, with permission from Springer. (d) An exotic C_4 grassland of weedy *Saccharum* spp. that has spread into forested uplands in northern Oahu, Hawaii, USA. Note how the forest is being broken into patches and isolated trees on the hillside in the middle of the photo.

Exotic C_4 Grasses and Tropical Deforestation

Through the use of fire, humans have been major agents of vegetation change for more than 2 million years. The most significant impact of burning has come in the past few thousand years, when humans increasingly used fire to maintain and expand agricultural lands. As a result, many of the natural C_4 grasslands and savannas of the world are derived from human-caused fires. In recent decades, the introduction and spread of exotic C_4 grasses has become one of the more pervasive threats to biodiversity in tropical forests and savanna. African C_4 grasses that are well adapted to human activities are the leading threats, and many of these are introduced as pasture grasses. In Central and South America, the C_4 grasses *Melinus minutiflora*, *Panicum maximum*, and various *Pennisetum* species have become widespread pests because of their ability to rapidly colonize wet soils cleared for farming or grazing, after which they establish highly flammable swards that kill all other species through crowding or burning. In Southeast Asia and Indonesia, and on many Pacific islands, the C_4 grass *Imperata cylindrica* has become a serious problem for similar reasons. Compared with the original diversity of hundreds to thousands of plant species per hectare in a mature forest, weedy grasslands are dominated by one or two species of C_4 grass.

Although the major cause of tropical deforestation has been clear-cutting for timber and agricultural development, weedy C_4 grasses contribute to deforestation by promoting destructive fires and preventing forest regeneration. The high productive potential of the C_4 pathway allows weedy C_4 grasses to rapidly spread into these disturbed areas and establish swards that facilitate frequent burning. After repeated burn cycles, soils lose fertility due to nutrient volatilization in smoke, leaching, and nutrient sequestration by the C_4 biomass. Climates over grass landscapes are drier than forests (C_4 vegetation transpires less than C_3 forests), and trees become too infrequent to provide seeds for natural reforestation. Establishment of forest seedlings becomes difficult, if not impossible and succession is arrested in the grassland state. To avoid these problems, clear-cutting is discouraged in favor of sustainable, low-intensity harvesting. In the presence of C_4 exotics, however, even small amounts of logging can set in motion autocatalytic grass-fire cycles that cause unintended (cryptic) deforestation (**Figure 16**).

The objective of low-intensity logging is to mimic natural gap dynamics by creating small openings that quickly return to forest. In the absence of C_4 grasses, forest recovery is rapid because seed rain from the surrounding trees quickly establishes seedlings that fill the gap within decades. If exotic C_4 grasses can instead establish in the gap, the successional trajectory can be substantially different because of their ability to quickly generate flammable conditions. In areas with human activity, invasion of C_4 grasses into logged gaps is likely, because weedy grasses can spread from parent populations in previously disturbed areas and along roads and logging paths. Seeds of these grasses can be carried into gaps by logging equipment or even on clothing. Once established, the next critical phase is ignition, which is typically promoted by

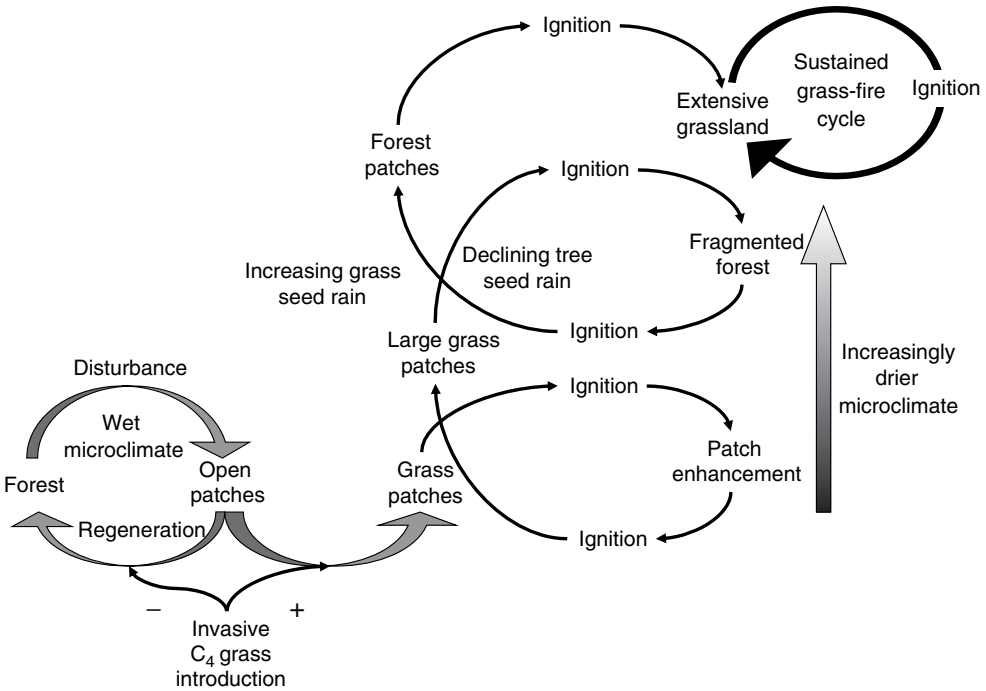


Figure 16 A conceptual model of forest conversion to exotic C_4 grassland following low-intensity logging. In the absence of C_4 grasses and fire, the forest will regenerate in open patches created by logging. In the presence of invasive grasses and fire, autocatalytic grass-fire cycles can become established, which eventually convert the landscape to an exotic grassland. Reprinted from Sage RF and Kubien DS (2003) *Quo Vadis* C_4? An ecophysiological perspective on global change and the future of C_4 plants. *Photosynthesis Research* 77: 209–225, with permission from Springer.

human action such as the discarding of cigarettes. Ignition probabilities are higher in grass patches because they readily dry, unlike forest interiors. The first fire is often small but intense, enhanced by downed woody debris from the logging operation. Seedlings of regenerating trees are typically killed by the fire, and the fire kills forest vegetation on the edge of the patch. Grasses quickly colonize the newly killed forest margins, enlarging the patch (**Figures 15(d)** and **16**). This promotes a second fire causing further patch enhancement, followed by additional fires within a few years of each other as long as humans provide an ignition source. Each successive fire increases the probability of subsequent fires and forest loss, because the microclimate over larger grass swards is warmer and less humid, thereby causing more extensive drying of adjoining forests (**Figure 16**). Also, as grass swards enlarge and coalesce, fires can spread much farther from a single ignition source. Eventually, forests become fragmented and forest patches decline in size, with consequent reduction of the seed rain from forest species. These trends are difficult to reverse, because forest regeneration is inhibited by the lack of tree seeds, reduced soil fertility, and a dryer regional climate, and so the weedy grasslands persist on the landscape. Although forest loss following low-intensity logging is slower than clear-cutting or pasturing, the end result can be the same unless people understand the process and take steps to avoid it by preventing C_4 grass establishment and fire.

Summary and Conclusion

C_4 photosynthesis is a polyphyletic solution to challenges imposed by reductions in atmospheric CO_2 and increasing aridification that occurred over the past 35 million years. In C_3 plants, the low atmospheric CO_2 levels of recent geological time promoted substantial inhibition of photosynthesis by photorespiration in hot conditions, particularly if water was also limiting. C_4 plants overcame these limitations by localizing Rubisco in the BS compartment into which CO_2 is concentrated. As a consequence, C_4 plants are highly productive, aggressive competitors under conditions promoting photorespiration in C_3 plants. This improved performance has enabled native C_4 species to dominate hot, open landscapes of low latitudes, and warm temperate regions receiving summer rain. It has also enabled invasive C_4 grasses adapted to human disturbance to become severe weeds, with important consequences for tropical biodiversity. Where humans disrupt forest cover and allow grass establishment, they establish conditions where invasive C_4 grasses can accelerate fire cycles and alter soil chemistry. In many locations, diverse tropical forests are replaced by grasslands dominated by a few aggressive C_4 species. Because natural C_4-dominated grasslands and savannas often occur on more productive landscapes, they too have been heavily exploited for agricultural purposes, resulting in habitat loss and impoverishment of the native C_4 flora of the Earth. Unfortunately, this loss is disproportionately impacting the global biodiversity of C_4 plants, since there are only 7500 C_4 species on planet Earth.

Appendix

List of Courses:

1. Introductory Botany
2. Plant Physiology
3. Plant Ecology
4. Conservation Biology
5. Ecology
6. Physiological Ecology
7. Global Change Biology
8. Paleoecology

See also: Carbon Cycle. Desert Ecosystems. Mediterranean-Climate Ecosystems. Nitrogen, Nitrogen Cycle. Photosynthesis, Mechanisms of

References

Bobe R and Behrensmeyer AK (2004) The expansion of grassland ecosystems in Africa in relation to mammalial evolution and the origin of the genus Homo. *Palaeogeography, Palaeoclimatology, Palaeoecology* 207: 399–420.

Bond WJ (2008) What limits trees in C_4 grasslands and savanna? *Annual Review of Ecology and Systematics* 39: 641–659.

Bond WJ and Midgely GF (2012) Carbon dioxide and the uneasy interactions of trees and savannah grasses. *Philsophical Transactions of the Royal Society Series B* 367: 601–612.

Bowes G (2011) Single-cell C_4 photosynthesis in aquatic plants. In: Ragavendra AS and Sage RF (eds.) *C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms*, pp. 63–80. Dordrecht: Springer.

Cerling TE, Wynn JG, Andanje SA, *et al.* (2011) Woody cover and hominin environments in the past 6 million years. *Nature* 476: 51–56.

Christin PA, Salamin N, Kellogg EA, Vicentini A, and Besnard G (2009) Integrating phylogeny into studies of C_4 variation in the grasses. *Plant Physiology* 149: 82–87.

Ehleringer JR, Cerling TE, and Dearing MD (eds.) (2005) A history of CO_2 and its effects on plants. *Animals and Ecosystems*. Berlin: Springer.

Kanai R and Edwards GE (1999) The biochemistry of C_4 photosynthesis: patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 49–87. San Diego: Academic Press.

Leakey ADB, Ainsworth AA, Bernacchi CJ, Rogers A, Long SP, and Ort DR (2009) Elevated CO_2 effects on plant carbon, nitrogen and water relations: Six important lessons from FACE. *Journal of Experimental Botany* 60: 2859–2876.

Raghavendra AS and Sage RF (eds.) (2011) *C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms*. Dordrecht: Springer.

Sage RF, Christin PA, and Edwards EJ (2011) The C_4 plant lineages of planet Earth. *Journal of Experimental Botany* 62: 3155–3169.

Sage RF, Li M, and Monson RK (1999a) The taxonomic distribution of C_4 photosynthesis: patterns and controlling factors. In: Sage RF and Monson RK (eds.) *C_4 Plant Biology*, pp. 551–584. San Diego: Academic Press.

Sage RF and Monson RK (eds.) (1999) *C_4 Plant Biology*. San Diego: Academic Press.

Stromberg CAE (2011) Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences* 39: 517–544.

Crop Mixtures and the Mechanisms of Overyielding

Long Li, China Agricultural University, Beijing, China, and Shihezi University, Shihezi City, China
Lizhen Zhang and Fusuo Zhang, China Agricultural University, Beijing, China

© 2013 Elsevier Inc. All rights reserved.

Glossary

Agroforestry A collective name for land-use systems and practices in which woody perennials are deliberately integrated with crops and animals on the same land management unit. The integration can be either in a spatial mixture or in a temporal sequence.

Contour hedgerow intercropping Also called alley cropping; crops are grown in the interspaces between rows of planted woody shrub or tree species that usually are legumes, and in which the woody species are periodically pruned during the cropping season.

Crop mixture Two or more than two crop (tree) species are grown simultaneously on the same field in alternative row(s) or mixture with no distinct row arrangement.

Intercropping Growing two or more crops simultaneously on the same field. According to row arrangement and cogrowth time, intercropping can be divided into three types: mixed, relay, and strip intercropping. Mixed intercropping refers to growing two or more crops simultaneously with no distinct row arrangement. Relay intercropping means growing two or more crops simultaneously during part of the life cycle of each. A second crop is planted after the first crop has reached its reproductive stage of growth but before it is ready for harvest. Strip intercropping means growing two or more crops simultaneously in different strips that are wide enough to permit independent cultivation but narrow enough for the crops to interact agronomically.

Land equivalent ratio (LER) The area needed for sole cropping to produce the same amount of crops as produced in 1 ha of intercropping or mixed cropping.

Light use efficiency (LUE) Crop light use efficiency is defined as the slope of the linear relationship between accumulated biomass and cumulative intercepted light.

Monoculture Growth of the same sole crop on the same field.

Root length density (RLD) Root length per unit soil volume (m m^{-3} or cm cm^{-3}).

Root weight density (RWD) Root fresh or dry weight in unit soil volume (g cm^{-3}).

Sole cropping One crop variety grown alone in pure stand at normal density.

Crop Mixtures

Crop mixtures are of agricultural significance because appropriate mixtures can provide increased yields or decreased need for inputs such as fertilizers and pesticides, as reviewed in later sections of this article. Crop mixtures refer to cropping systems in which different crop (or tree) species or different varieties of the same crop species are grown simultaneously, or partly so, in the same field. Crop mixtures usually include agroforestry, hedgerow intercropping (or alley cropping), mixed intercropping, strip intercropping, and relay intercropping. Crops or varieties are not necessarily sown at the same time or harvested at the same time, but they have cogrowth stages for some period. Agroforestry is a collective name for land-use systems and practices in which woody perennials are deliberately integrated with crops on the same land management unit. Contour hedgerow intercropping (also called alley cropping) crops are grown in the interspaces between rows of planted woody shrub or tree species, usually legumes, and in which the woody species are periodically pruned during the cropping season. Intercropping means growing two or more crops simultaneously either in no distinct row arrangement or in rows or strips on the same field. The strips are wide enough to permit independent cultivation but narrow enough for the crops to interact agronomically. Intercropping can be classified into mixed intercropping and row or strip intercropping based on the row arrangement and the relay intercropping cogrowth time of two crops.

Importance of Crop Mixtures

Crop mixture, mainly as agroforestry or intercropping, is practiced globally, including in Africa, Asia, Europe, and the Americas. Although there is no firm estimate of the global extent of mixed cropping systems, its importance can be illustrated by several examples.

Africa and Latin American

Cassava (*Manihot esculenta* Crantz), the fourth most important energy staple of the tropics, provides food and income for some 750 million people (Leihner, 1983). It is often found in mixed stands together with a variety of other food or cash crops. Recent estimates indicate that 40% and 50% (or more), respectively, of the cassava grown in America and Africa is intercropped (Leihner, 1983). Bananas are very important in Kenya for domestic consumption and export, and are extensively grown, mainly intercropped with short-term crops. There has been an increase in the grower interest in using intercropping to develop new cropping systems for their land (Ouma, 2009). Intercropping constituted 91% of the cropping systems in Ethiopia (Fininsa and Yuen, 2001). Legume–cereal intercropping, especially maize/beans intercropping, is common throughout East and Southern Africa (Mucheru-Muna *et al.*, 2010). Agroforestry is also popular all over the world. For example, there are 2.8 million hectares of jungle rubber agroforests and 3.5 million hectares of all multistrata

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00363-4

agroforests in Indonesia, 7.4 million hectares in India, 5–6 million hectares in Niger, 5.1 million hectares (90% of agricultural land) in Mali, 9.2 million hectares of silvopastural systems and 0.77 million Coffee agroforests in Central America, 6 million hectares of dehasa agroforestry in Spain/Portugal, and 7.8 million hectares of cocoa agroforests worldwide (Zomer *et al.*, 2009).

Asia

There are thousands of intercropping systems in China, including agroforestry, cereals/cereals, legumes/cereals, cereals/vegetables intercropping, etc. The northern plain region, a major agricultural area covering Hebei, Henan, Shandong, Beijing, Tianjin, and parts of Anhui, Jiangsu, Shanxi, and Shaanxi, is typical in agroforestry. Various agroforestry schemes have also been adopted in the region within 4.5 million hectares of the total area (Yin and He, 1997). Paulownia (*Paulownia elongata*)-based intercropping is the most popular, accounting for approximately two-thirds of the total area (Yin and He, 1997). Poplars are also the major tree component of traditional agroforestry systems throughout the south temperate central area of China, which includes all or portions of Jiangsu, Anhui, Zhejiang, Hubei, Henan, Shandong, and Shanxi provinces, an area covering approximately 600,000 km^2 (Fang *et al.*, 2010). Hedgerow intercropping has become increasingly popular in China since the 1970s. There were, for example, 700,000 ha in Shanxi Province and 10,000 ha in northern Shaanxi Province in 1999 (Sun *et al.*, 2008).

Intercropping systems of cereals/cereals, grain legume/cereals, and cereals/vegetables are common in China. A total of 25–28 million hectares of intercropping was distributed across all of the provinces in the country except for Tibet and Qinghai Province in the 1980s, and the area reached 33 million hectares in the 1990s (Zou and Li, 2002). It is estimated that there was 1.5 million hectares of soybean-cultivated area, which was mainly intercropped with various other crops in southern China in 2008 (Zhou *et al.*, 2010). Wheat/maize intercropping (**Figure 1(a)**) has been an important cropping system for raising grain yields per unit of arable land in northwestern China. The area under wheat/maize intercropping was 75,000 ha in Ningxia Hui Autonomous Region in 1995, which produced 43% of total grain yields for the region, and there is 200,000 ha of wheat/maize intercropping annually in Gansu, with more than 12 tons of grain yields per hectare (Li *et al.*, 2001a). In China, intercropping has been developed into a pattern of high output and high input, which is completely different from the traditional pattern of low input and low output. Intercropping featured with high yield (output) and high input plays an important role in Chinese modern agriculture. In tropical regions of Asia, intercropping is much more common.

Main Patterns of Crop Mixtures

Intercropping is practiced in both tropical and temperate areas. In Central Africa, maize, sorghum, or millet is grown with pumpkin (*Cucurbita* spp.), cowpeas (*Vigna unguiculata*), pigeon peas (*Cajanus cajan* L.), or beans (*Phaseolus* spp.). Cocoa (*Theobroma cacao* L.) is grown with yams (*Dioscorea* spp.) or cassava (*M. esculenta* Crantz). In the tropical Americas, maize is grown with beans and squash (*Cucurbita* spp.). In both Africa and Latin America, beans and peas (*Pisum sativum* L.) climb tall cornstalks, whereas pumpkins and squash cover the ground below (Machado, 2009). In Asia, intercropping with similar patterns is practiced by farmers in tropical areas of China, India, Iran, Nepal, Sri Lanka, and Thailand.

Intercropping is also practiced in temperate regions. For instance, there are many kinds of intercropping in northwestern China, such as spring wheat/spring maize (**Figure 1(a)** and **1(c)**), maize/soybean/flax (*Linum usitatissimum* L.) (**Figure 1(b)** and **1(d)**), winter wheat/spring maize, wheat/soybean (**Figure 1(e)**), wheat/faba bean (*Vicia faba* L.), maize/potato (*Solanum tuberosum* L.) (**Figure 1(f)**), wheat/potato, wheat/sunflower (*Helianthus annuus* L.), wheat/vegetables (e.g., tomato (*Lycopersicon esculentum* Mill.), Peking cabbage (*Brassica pekinensis* Rupr.), and onion (*Allium* spp.)), maize/vegetables (e.g., garlic (*Allium sativum* L. (**Figure 1(g)**)), maize/pea (*P. sativa* L.) (**Figure 1(h)**) intercropping, in areas with one cropping season because of the annual thermal limitation in the northwestern part of China (Ministry of Agriculture, China, 2006).

In tropical areas of Indonesia, India, Niger, and Mali, there are jungle rubber-based agroforestry and multistrata agroforestry. There are silvopastoral systems, coffee agroforests in Central America, dehesa agroforestry in Spain/Portugal, and cocoa agroforestry worldwide (Zomer *et al.*, 2009). China practices agrofostery based not only on rubber tree (*Hevea* spp. Aubl.), tea tree (*Camellia* spp.), and citrus (*Citrus* L.) (**Figure 1(o)** and **1(p)**) in its tropical and subtropical regions, but also on fruit trees (e.g., apple (*Malus* Mill.), pear (*Pyrus* L.), apricot (*Armeniaca* Mill.) (**Figure 1(i)**), peach (*Amygdalus* L.), and jujube (*Ziziphus jujuba* Mill.) (**Figure 1(j)**), pepper (*Zanthoxylum bungeanum* Maxim.), and mulberry (*Morus* L.) in its temperate regions (Ministry of Agriculture, China, 2006).

In Europe, too, some kind of intercropping has been identified, such as barley (*Hordeum sativum* L.)/pea (*P. sativum* L.) intercropping in Denmark, the United Kingdom, France, Italy, and Germany, and wheat/pea intercropping in Denmark, popcorn/melon and potato/cabbage (*Brassica oleracea* var. *capitata* L.) intercropping in the United Kingdom, berseem clover (*Trifolium alexandrinum* L.)/barley, common vetch (*Vicia sativa* L.)/wheat, triticale (*Triticosecale* Wittmack), barley, or oat (*Avena sativa* L.) intercropping in Greece, fennel (*Foeniculum vulgare* Mill.)/dill (*Anethum graveolens* L.) intercropping in Italy, maize/bush bean (*Phaseolus vulgaris* L.) intercropping in Spain, leek (*Allium porrum* L.)/celery (*Apium graveolens* L. var. *dulce* (Mill.) Pers.) intercropping in Switzerland, and various vegetables/vegetables intercropping (e.g., cabbage, cauliflower (*B. oleracea* var. *botrytis* L.), or strawberry (*Fragaria xananassa* L.) intercropped with bean, cos lettuce (*Lactuca sativa* L.), leaf lettuce (*L. sativa* L.), onion (*Allium cepa* L.), or radish (*Raphanus sativus* L.) in Turkey.

In the Americas, peas are intercropped with barley or oat, and in Canada, wheat is intercropped with canola or pea and broccoli is intercropped with peas, beans, potatos, oats, cauliflower, or cabbage; in the USA, maize and soybeans are intercropped. In Brazil, there are many kinds of intercropping as in other tropical areas.



Figure 1 Various intercropping systems practiced by farmers in China. (a) Wheat/maize intercropping practiced by local famers in Ningxia, Northwest China. (b) Various intercropping systems of crops (flax, soybean, maize) in Gansu Province, China. (c) Wheat/maize strip intercropping in Gansu Province, China. (d) Maize/soybean/flax intercropping in Gansu Province, China. (e) Wheat/soybean intercropping in Gansu Province, China. (f) Potato/maize intercropping in Gansu Province, China. (g) Maize/garlic (*Allium sativum* L.) intercropping in Gansu Province, China. (h) Maize/pea intercropping in Gansu Province, China.



Figure 1 Continued. (i) Apricot (*Armeniaca* Mill.)/cotton agroforestry intercropping in Southern Xinjiang, China. (j) Jujube (*Ziziphus jujuba* Mill.)/onion intercropping in Southern Xinjiang, China. (k) Hybrid/glutinous rice intercropping in Yunnan Province, China. (l) Tobacco/peanut intercropping in Fujian Province, China. (m) Winter wheat/maize relay intercropping in Sichuan Province, Southwest China. (n) Banana/capsicum intercropping in Guangxi Province, China. (o) Citrus/potato intercropping in Guangxi Province, China. (p) Citrus/ *Raphanus sativus* L. Intercropping in Guangxi Province, China.

Overyielding

Land Equivalent Ratio (LER)

LER is defined as the relative land area with sole cropping required to produce the yields achieved by intercropping (Willey, 1990). It is calculated using the following formula:

$$LER = \frac{Y_{ia}}{Y_{sa}} + \frac{Y_{ib}}{Y_{sb}}$$

where Y_{ia} and Y_{ib} are yields of crops a and b in intercropping, respectively, and Y_{sa} and Y_{sb} are yields of crops a and b in sole cropping, respectively. LER indicates the area needed under sole cropping to produce the same amount of crops as produced in 1 ha of intercropping or mixed cropping. Therefore, there is yield advantage of intercropping if LER is greater than 1, and no yield advantage of intercropping if LER is less than 1.

The literature published in the past 15 years has summarized LER values. In most cases, intercropping has had a significant yield advantage, with LERs greater than 1 indicating overyield compared to sole cropping. However, when some similar crops were intercropped (e.g., faba bean (*V. faba* L.)/pea intercropping), there was no overyield (Li *et al.*, 1999).

Yields

Yield advantages have been reported for many crop mixtures. Overyielding of intercropping can be classified into three types: 1. *Overyielding of both intercropped crop components.* For example, in 4 years of field experiments, maize (*Zea mays* L.) overyielded on average 43% and faba bean overyielded on average 26% when grown on a low-phosphorus but high-nitrogen soil (Li *et al.*, 2007). A similar result was observed for potato/maize intercropping in Yunnan Province, southwestern China (Li *et al.*, 2009). 2. *One component overyields but the other component is comparable to its corresponding sole species.* For instance, there was 33–85% overyielding in tobacco/maize, sugarcane/maize, and wheat/faba bean intercropping systems in Yunnan province, China (Li *et al.*, 2009), in which the overyielding was mainly derived from additional crop maize in tobacco/maize and sugarcane/maize intercropping, and additional faba bean production in wheat/faba bean intercropping (Li *et al.*, 2009). *Faidherbia* is a nitrogen-fixing Acacia species that is indigenous to Africa and is widespread throughout the continent. It goes dormant and sheds its foliage during the early rainy season, at the time when field crops are established. Its leaves regrow only at the end of the wet season. This unusual phenology makes it highly compatible with food crops, since it does not compete with crops significantly for light, nutrients, or water during the growing season. On the contrary, annual crops in the vicinity of *Faidherbia* trees tend to exhibit improved performance and yield (Garrity *et al.*, 2010). 3. *One component crop overyields; the other component is reduced in mixture, but the total yield of the two component crops overyield*. For instance, in a mixture of faba bean and barley grown in Ethiopia, Agegnehu *et al.* (2006) found that increasing faba bean proportion from 12.5% to 62.5% increased faba bean seed yields from 12% to 48% but lowered barley grain yields from 93% to 73% of the respective sole crop yields. Total grain yield and barley yield equivalents increased as the seeding percentage of faba bean in the mixture increased (Agegnehu *et al.*, 2006).

Mechanisms of Overyielding

Relative Importance of Above-Ground and Below-Ground Interactions Between Crop Species

Yield advantages of crop mixtures result from both above- and below-ground interactions between associated species (Semere and Froud-Williams, 2001; Thorsted *et al.*, 2006; Mariotti *et al.*, 2009). Donald (1958) found that root competition had a greater effect than shoot competition on crop growth and that there were positive root interactions among species. In 33 of 47 (70%) cases found in the literature, root competition had a greater effect than shoot competition on plant growth and resource capture (Wilson, 1988). In a maize/faba bean intercropping system, comparisons on the basis of LER clearly showed that the intercropping advantage was significantly decreased by plastic sheet and nylon mesh (37 μm) barriers between root systems. When the roots of the two species intermingled, LER values based on total above-ground biomass and grain yields were 1.21 and 1.34, respectively, but when the roots of the two species were separated completely by plastic sheet or partly by nylon mesh, the LER values declined to 1.06 and 1.19, or 1.12 and 1.26, respectively (Li *et al.*, 1999). Recent reports also showed similar results between intercropped maize and pea (Semere and Froud-Williams, 2001), between intercropped winter wheat and white clover (*Trifolium repens*) (Thorsted *et al.*, 2006), and between faba bean and maize (Li *et al.*, 2007). In combinations of two winter cereals (barley and wheat) and two legumes (white lupin and common vetch), an example of low-input intercropping system, when a nitrogen-fixing species is present, the mixture of the roots of components is important for the utilization of the soil resources, and when a climbing species is also present, the mixture of shoots can result in an increased utilization of light (Mariotti *et al.*, 2009).

Above-Ground Interactions Between Species

When two or more crops are growing together, each must have adequate space to maximize interspecific facilitation and to minimize interspecific competition between the crops. Above-ground interactions between intercropped species can be mainly summarized into three aspects: (1) light competition and utilization, (2) canopy heat resource and crop development, and (3) spatial and temporal compensation of microclimate factors, such as radiation, temperature due to the structures of crop mixtures, space configuration, and time arrangement.

Light Interception (LI) and Light Use Efficiency (LUE)

LI and LUE are useful concepts for characterizing the resource capture and use efficiency of cropping systems, including intercropping. Crop LUE is defined as the slope of the often linear relationship between accumulated biomass and cumulative intercepted photosynthetically active radiation (PAR) (Monteith, 1977; Russell *et al.*, 1989). LI and its conversion to

dry matter by individual components have been quantified in several studies of conventional intercropping systems such as in leucaena/millet (Corlett *et al.*, 1992), in sorghum/groundnut (Harris *et al.*, 1987), in maize/bean (Tsubo *et al.*, 2001), and in wheat/cotton (Zhang *et al.*, 2008a). Improved productivity can result from either greater interception of light, a higher LUE, or a combination of the two (Willey, 1990). LI sometimes increases as a result of mixing two species and growing them together instead of growing each species alone, as a result of either a lengthening of the period of soil coverage in the intercrop (temporal advantage) or a more complete soil cover (spatial advantage) in the intercropping, compared to the ensemble of sole cropping (Keating and Carberry, 1993).

At a small spatial scale, the strip intercropped species capture more light than the canopy of sole crop species per unit of strip area, but not per unit of homogenized land area. In relay intercropping, although the intercropping period is not too long, the direct effect of shading is less than the indirect effect through a lower leaf area index (LAI) (incomplete canopy closure) as a result of retarded growth due to a development delay during the intercropping phase. Thus, to increase light capture, the intercropping systems should be improved by alleviating the delay in crop development by applying film cover or by increasing plant densities to reduce 'spatial' losses in LI (Zhang *et al.*, 2008a).

LUE will not be much affected in intercropping systems with component crops that differ in growing period, since the competition between component crops is less (Fukai and Trenbath, 1993). In most of the previous studies, LUEs of both intercropped species were not greatly affected by the mixture. The LUE of dominant component crops in intercropping systems is generally not affected. Examples include millet in millet/groundnut (Willey, 1990), sorghum in sorghum/groundnut (Matthews *et al.*, 1991), and maize in maize/cowpea (Watiki *et al.*, 1993). For the lower, shade crops, for example, groundnut and cowpea, it was reported that the LUE was sometimes 'slightly' affected, potentially as a result of competition for other resources, that is, water and nitrogen.

Canopy Heat Capture and Crop Development

A shaded intercropping crop always captures less radiation than that captured by taller crop, which causes decreases of canopy ventilation and soil temperatures and delays crop development. In wheat/cotton relay strip intercropping, the cotton was delayed by 9–15 calendar days, compared to cotton in monoculture (Zhang *et al.*, 2008b). In wheat/soybean intercropping, the development of the relay intercropped crop was also affected at an early stage (Wallace *et al.*, 1992). However, when comparing the development delay to the results reported for other intercropping systems (Matthews *et al.*, 1991; Gethi *et al.*, 1993; Bukovinszky *et al.*, 2004), the delay in wheat/cotton intercropping is longer and has more serious consequences.

The most obvious modification of the environment in intercropping is shading of one species by another taller crop; as a consequence, the reduced capture of light results in a reduced growth and yield of shaded crop. Shading also modifies other environmental conditions, for example, ventilation and soil temperature (Midmore *et al.*, 1988; Midmore, 1993). By environmental modification, intercropping can even obtain higher yields than sole cropping (Midmore *et al.*, 1988); for example, under heat conditions, reduced soil temperature and moisture by intercropping may favor multiplication and growth of some soil microorganisms (Singh *et al.*, 1986). The environment can also be modified by management practices, for example, covering the soil by a film or straw (Rana *et al.*, 2004).

The decrease in both air and soil temperatures was mainly caused by shading of taller crops (**Figure 2**), because there were no differences in temperature between intercropping and monoculture when the air temperature was low due to the absence of direct radiation (Zhang *et al.*, 2008b). A soil cover consisting of plastic film effectively ameliorated soil temperature in intercropping. Licht and Al-Kaisi (2005) found an increased soil temperature as a result of tillage. Thus, tillage may also provide an option to raise soil temperature in intercrops. Increase in temperature may also be achieved by measures that diminish shading, such as using of shorter wheat cultivars or planting crops on ridges.

Spatial and Temporal Complementarity

Yield advantages from intercropping are often attributed to complementation between component crops in the mixture, resulting in a better total use of resources rather than growing crops separately. A good example of temporal complementarity of crop mixtures is relay intercropping system, in which there is an increase in the total LI because the

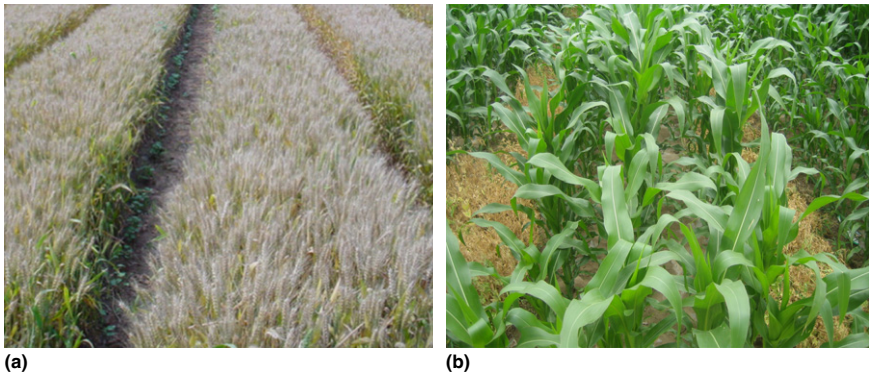


Figure 2 The shading effect in strip wheat/cotton (a) and maize/field pea (b) intercropping. Photos by Lizhen Zhang.

intercropped second crop has built its seedling with relative larger LAI after the first crop is harvested. Relay intercropping is commonly practiced in an area that has insufficient growing degree days for continuous double cropping system.

Plant architecture can be used strategically to allow one member of a mix to capture sunlight that would otherwise be unused. This is considered spatial complementarity. Widely spaced maize plants growing above an understory of beans and pumpkins are a classic example. Different mixture designs are used in practice and have been characterized as to their productivity in relation to the spatial configuration.

In strip intercropping systems, a border row effect has often been found. In maize/soybean strip intercropping, Ghaffarzadeh *et al.* (1994) found that strip intercropping had 20–24% higher maize yields and 10–15% lower soybean yields in adjacent border rows in Iowa, USA. In this case, the higher yield of border rows of the tallest species (maize) suggests increased radiation capture at the expense of the neighboring rows of the smaller species (soybean). In China, Li *et al.* (2001a) found that the yields in border rows of intercropped wheat were significantly higher than those in inner rows, in both wheat/maize and wheat/soybean intercropping systems. Waterworth (1994) suggested that a crop such as cotton, which both tolerates a wide range of population densities and has a late developing leaf canopy, may be particularly well suited to intercropping.

The greater yield in border rows can be attributed to greater LI (Zhang *et al.*, 2008a) and a better acquisition of nutrients in the border rows. The number of border rows is a major factor in determining the advantage of the productivity in intercropping systems. Li *et al.* (2001a) reported that yield increase of intercropped wheat is due to yield increase not only of border rows but also of inner rows in the wheat/maize intercropping systems.

An important aspect of an intercropping system is the extent of competition between the crops. The plant types of the intercropped species should be selected to complement each other (Smith and Francis, 1986; Davis and Woolley, 1993). Inadequate management of plant density can be detrimental to the intercrop (Davis and Garcia, 1987).

Below-Ground Interactions Between Species

Complementarity in Root Spatial and Temporal Distribution

Under field condition, crops in intercropping systems are not necessarily sown or harvested at the same time (Vandermeer, 1989). The growth rate and architecture of roots can differ greatly among different crop species, even if crops are grown and harvested at the same time. There are temporal and spatial niche complementarities in the use of resources, such as nutrients (Midmore, 1993; Morris and Garrity, 1993a, b; Stern, 1993; Li *et al.*, 2001a, 2003) and water (Sekiya and Yano, 2004). The niche complementarity hypothesis implies that plant species or functional groups occupy functionally distinct niches in an ecosystem and use resources in a complementary way (Kahmen *et al.*, 2006). Dynamics of temporal and spatial distribution of roots of different species in soil are directly related to nutrient acquisition and water capture in a mixture of different species. Therefore, the rooting patterns of mixed species are related closely to the above-ground performance of these species in crop mixture.

The root spatial distribution between mixed crop species can be classified into three patterns:

1. Spatial separation of the crop root systems between the associated species, which results from interspecific interactions. Some species change the rooting pattern when associated with other species. For instance, sorghum distributes roots deeper in intercropping than in sole cropping, in which root distribution in spatial scale is separated from sunflower, a crop with a shallow rooting system (Miyazawa *et al.*, 2010). In an alley cropping system with *Acacia saligna* and *Sorghum bicolor* in northern Kenya, Lehmann *et al.* (1998) found that sorghum has more roots in the topsoil and the trees have more roots in the subsoil under alley cropping than in respective monoculture. As a result, the root system of the alley cropped Acacia and sorghum exploited a larger soil volume utilizing soil resources more efficiently than the respective monocultures (Lehmann *et al.*, 1998). By a means of ^{32}P tracer, Hauggaard-Nielsen *et al.* (2001) also found that interspecific interactions made the barley roots grow deeper and enhanced lateral root development in barley/pea intercropping, compared to sole barley cropping (Hauggaard-Nielsen *et al.*, 2001). By both auger and monolith methods, the root distributions were measured up to a depth of 100 cm in the soil profile in wheat (*Triticum aestivum*)/maize (*Z. mays*) intercropping and in sole wheat and maize cropping (Li *et al.*, 2006). Li *et al.* (2006) found that the roots of intercropped wheat spread under maize plants and had much greater root length density (RLD) at all soil depths than sole wheat, and the growth of maize roots was limited laterally by intercropped wheat (Li *et al.*, 2006), indicating that the root distribution of the two crops was changed by interspecific interactions.
2. The associated crops have inherently different traits in rooting patterns, with one being deeper and the other shallower in the soil profile. When grown together, interspecific competition is reduced, leading to overyielding. For example, Crotalaria (*Crotalaria juncea* L.) had a deep rooting system under both sole cropping and intercropping (Miyazawa *et al.*, 2010) that was not influenced by intercropping. The roots' separation in space decreased root competition, thus causing overyielding of aboveground parts compared to sole cropping.
3. Compatibility, or intermingling, between the roots of mixed crops facilitates resource utilization, thus overyielding. In faba bean/maize intercropping, for instance, interspecific interaction was mainly positive (or interspecific facilitation) (Li *et al.*, 1999). The roots of the two species were not distinctly separated from each other, and there was no apparent interspecific competition (Li *et al.*, 1999, 2006). The RLD of maize intercropped with faba bean at different soil depths was influenced by intercropping to a smaller extent, compared with maize intercropped with wheat. The roots of intercropped maize spread together with them (Li *et al.*, 2006). The intermingling of faba bean and maize roots in some zone enhanced nitrogen and phosphorus acquisition by crops

because faba bean roots are able to mobilize insoluble P in soil (Li *et al.*, 2007) and supply some N to maize (Fan *et al.*, 2006).

The temporal complementarity in root distribution in soil profile is another mechanism of crop mixture overyielding. There is some temporal differentiation in root development and growth between crop species because crops are not really sown and harvested simultaneously in many intercropping systems and because of the inherent traits of specific crop root growth. Using a ^{32}P labeling, Hauggaard-Nielsen *et al.* (2001) found that barley roots grow faster than pea roots, approximately 10 days earlier, in both vertical and lateral directions (Hauggaard-Nielsen *et al.*, 2001). In wheat/maize intercropping, the roots of intercropped wheat occupy larger below-ground space during the cogrowth stages of the two crops, compared to sole wheat cropping (Li *et al.*, 2006). After wheat harvesting, the roots of intercropped maize occupy the space under wheat and maize, which extends the maize to the below-ground space at a later growth stage (Li *et al.*, 2011). This explains the observed overyielding of intercropped maize with wheat (so-called recovery growth of maize after wheat harvesting (Li *et al.*, 2001b)). The temporal differentiation in root developments for wheat at an early stage and maize at a late stage partly explains the overyielding of wheat/maize intercropping in Gansu province, northwestern China.

Nitrogen Complementary Utilization in Legume-Based Crop Mixtures

Because nitrogen is an important limiting nutrient in arable soils, legume/cereal intercrops often overyield in tropical and temperate regions. Some of the symbiotic N_2 fixed by legumes is taken up by non N_2-fixing species, increasing their growth and reducing nitrate leaching, which is aided by combinations of species with deeper and shallower roots. Nitrogen transfer from legumes to cereals occurs in most cases of intercropping legume–cereal systems. In a pot experiment, it was shown that N transferred from faba bean to wheat comprised 15% of the total N in wheat (Xiao *et al.*, 2004). Nitrogen transfer from peanut to associated aerobic rice was 11.4%, 6.4%, and 5.5% of the total N for rice in peanut/rice intercropping at N fertilizer application rates of 15, 75, and 150 kg N ha^{-1}, respectively (Chu *et al.*, 2004). Significant N transfer also occurs from sweet-blue lupine (*Lupinus angustifolius*) to Italian ryegrass (*Lolium multiflorum*) but not to intercropped westerwolde ryegrass (*L. multiflorum*) or oats (*A. sativa*) in intercropping systems (Palmason *et al.*, 1992). In maize-based intercropping, the proportion of N derived by maize from the associated legume varies from 7% to 11% for mungbean (*Vigna radiata*), 11–20% for cowpea (*V. unguiculata*), and 12–26% for peanut (*Arachis hypogaea*) (Senaratne *et al.*, 1995). In a soybean/sorghum intercropping system, approximately 7% of the total N accumulated by intercropped sorghum is transferred from the associated soybean (Ofosubudu *et al.*, 1993), whereas another study observed the proportion as 20–55% (Ofosubudu *et al.*, 1995). In contrast, nitrogen transfer does not occur from legumes to cereals in intercropping systems of cowpea/maize (Hardter and Horst, 1991), soybean/maize (Hamel *et al.*, 1991), field pea (*P. sativum*)/barley (Izaurralde *et al.*, 1992; Jensen, 1996), mungbean/rice (Aggarwal *et al.*, 1992), sweet-blue lupin (*L. angustifolius*)/oats (*A. sativa*) (Danso *et al.*, 1993), pea (*P. sativum*)/mustard (Waterer *et al.*, 1994), and cowpea/cotton (Rusinamhodzi *et al.*, 2006).

The occurrence of net nitrogen transfer depends on crop combinations (Fujita *et al.*, 1992; Farnham and George, 1993; Gebhart *et al.*, 1993; Stern, 1993; Farnham and George, 1994a, b, c; Cochran and Schlentner, 1995; Kobes and Simek, 1998; Beschow *et al.*, 2000; Carter and Ambus, 2006; Erkovan *et al.*, 2008; Serin and Erkovan, 2008). More specifically, it depends on the distance between legumes and cereals, because this affects the intermingling of roots (Ofosubudu *et al.*, 1993). Other factors include soil N concentration (Beschow *et al.*, 2000) and legume species (Stern, 1993).

Intercropping usually enhances the proportion of nitrogen derived from the atmosphere (%Ndfa) in the legumes, compared with that of corresponding legumes grown in monoculture. This occurs in sweet-blue lupin (*L. angustifolius*)/ryegrass (*L. multiflorum*) and sweet-blue lupin/oats (*A. sativa*) (Palmason *et al.*, 1992); cowpea (*V. unguiculata*)/maize, mung bean (*V. radiata* L.)/maize, and peanut (*A. hypagaea*)/maize (Senaratne *et al.*, 1995); vetch/barley (Kurdali *et al.*, 1996); dhaincha (*Sesbania aculeata*)/sorghum (*S. bicolor* L.) (Kurdali *et al.*, 2003); peanut/rice under 150 kg N ha^{-1} of N fertilization, but not under lower rates of N-fertilizer application (Chu *et al.*, 2004); rape (*Brassica napus* L.)/pea (Andersen *et al.*, 2004); cotton (*Gossypium hirsutum*)/cowpea (*V. unguiculata*) (Rusinamhodzi *et al.*, 2006); wheat (*T. aestivum*)/faba bean (*V. faba*) (Fan *et al.*, 2006); and pea (*P. sativum*)/oat (*A. sativa*) intercropping (Neumann *et al.*, 2007). The amount of biological N_2 fixation by intercropped legumes is frequently reduced compared with that of legumes in monoculture, in line with a reduction of sown area or planting density for legumes in mixture. However, the amount of biological N_2 fixation increases by intercropping when expressed on the basis of equivalent land.

Cereals associated with legumes use more soil and N fertilizer through interspecific interactions. In a pot experiment of faba bean/wheat intercropping, fertilizer ^{15}N was added to soil. The amount of N derived from ^{15}N added to soil of wheat was 24, 30, and 43 mg per pot in the treatment without root interactions (due to a solid barrier), with mesh barrier (partial roots interactions), and without barrier (complete root interactions), respectively (Xiao *et al.*, 2004). The root contact enhanced ^{15}N acquisition of wheat, but decreased that of faba bean.

These results indicate that interspecific interactions between legumes and cereals made cereals acquire more soil N, which forced the legumes to fix more nitrogen from air.

Reduced nitrogen leaching in the soil profile could play a role in determining nitrogen-acquisition efficiency of the intercropping system. Nitrogen leached from the root zone of N_2-fixing plant could be taken up by nearby roots of non N_2-fixing plants. For instance, nitrate leaching tended to be smaller in the lysimeters originally cropped with pea–barley intercrops than in those of sole pea crop (Hauggaard-Nielsen *et al.*, 2003). In a maize/soybean intercropping system, maize grain yields were increased by 5% and NO_3-N leaching losses reduced by 6% (Kanwar *et al.*, 2005). A study of an oat/pea mixture showed that oat in mixture acquired more soil N from

deeper layers. Therefore, the risk of soil N losses through leaching after mixtures can be lower than that after a monoculture of pea (Neumann *et al.*, 2007).

In Africa, *Faidherbia albida*, a legume tree species, can increase yields of associated crops when they are intercropped together. Garrity *et al.* (2010) reviewed research conducted in various countries on the improvement of soil nutrient content and crop yields under *F. albida* canopies, compared with controls in the open. Increases in nitrogen content ranged from 15% to 156%, but significant increases were also found in carbon, phosphorus, exchangeable potassium, calcium, and magnesium in soil. As a result, increases in millet yields ranged from 49% to 153% and in sorghum from 36% to 169%. In absolute terms, this means, in most cases, an additional cereal yield of 400–500 kg ha^{-1} or more. Yield increases of 37% were observed for groundnut, 200% for sorghum in the north-central Senegal, and 115% for sorghum in Burkina Faso (Garrity *et al.*, 2010).

Phosphorus Facilitation Between P-Efficient Species and P-Inefficient Species

Although plants often perform poorly on soils with low levels of soluble P, some of these soils contain considerable reserves of P that is unavailable to most species, such as sparingly soluble inorganic P and organic P in soil. There is significant difference in P acquisition from soil between different plant species. P-efficient species usually acquire insoluble soil P or organic P much more than P-inefficient species do.

Phosphorus acquisition of P-inefficient species may benefit from the presence of P-efficient species when they grow close together. Interspecific facilitation of P acquisition for wheat (*T. aestivum*)/white lupine (*Lupinus albus*) (Horst and Waschkies, 1987), sorghum (*S. bicolor*)/pigeon pea (*C. cajan*) (Ae *et al.*, 1990), wheat (*T. aestivum*)/chickpea (*Cicer arietinum*) (Li *et al.*, 2003), peanut (*A. hypogaea*)/maize (*Z. mays*) (El Dessougi *et al.*, 2003), and maize (*Z. mays*)/chickpea (*C. arietinum*) (Li *et al.*, 2004) intercropping has been observed in greenhouse pot experiments, and for faba bean (*V. faba*)/maize (*Z. mays*) intercropping under field conditions has been observed (Li *et al.*, 2007). The mechanisms underlying interspecific facilitation include efficient acquisition of organic P (Li *et al.*, 2003, 2004) and poorly available inorganic P (Horst and Waschkies, 1987; Ae *et al.*, 1990; Li *et al.*, 2007) by P-efficient species, which enhances P acquisition of P-inefficient species in intercropping systems.

Faba bean (*V. faba*) facilitated P acquisition of associated maize (*Z. mays*) under both field and greenhouse conditions. Overyielding was an average 46% for maize and an average 26% for faba bean in a 4-year field intercropping experiment; intercropping enhanced P acquisition by faba bean and maize, compared with that by the respective species in monoculture (Li *et al.*, 2007). There is temporal and spatial niche differentiation in resource acquisition for these species, with faba bean being sown and harvested earlier than maize under traditional farming practices. Greenhouse experiments have shown interspecific facilitation between faba bean and maize in terms of P acquisition and growth when temporal and spatial niche complementarity are diminished by sowing both at the same time and when their roots exploit the same soil volume (Li *et al.*, 2007). The physiological mechanism involved in this facilitation is that faba bean acidifies the rhizosphere by releasing protons and mobilizing insoluble soil P by exuding malate and citrate (Li *et al.*, 2007), which are chelating agents that mobilize P bound to calcium in calcareous soil and solubilize phosphate sorbed to soil particles (Hinsinger, 2001). Crop species vary widely in their capacity to access sparingly soluble inorganic P (Pearse *et al.*, 2006). Therefore, the combination of P-efficient and P-inefficient species is able to overyield under P-deficient soils.

Chickpea (*C. arietinum*), a species that effectively accesses organic P, facilitates P acquisition by associated organic P-inefficient wheat (*T. aestivum*) (Li *et al.*, 2003) and maize (*Z. mays*) (Li *et al.*, 2004) when grown in intercropping systems. The physiological mechanism behind this facilitation is the exudation from roots of chickpea of acid phosphatases (Li *et al.*, 2004), which hydrolyze organic P (Li *et al.*, 1997) and benefit both species when growing together. Many crop species (e.g., *L. albus* and *Trifolium subterraneum*) can access nucleic acids, phospholipids, glucose-1-phosphate, and glycerophosphate in soil, due to the release of phosphatases from their roots (Lambers *et al.*, 2008). This may be an important mechanism of facilitation and overyielding in P-limited soils.

Fe/Zn improvement In Mixture

Some iron-inefficient species, for example, guava (*Psidium guajava*), peanut (*A. hypogaea*), and red fescue (*Festuca rubra*) frequently show iron chlorosis when growing in calcareous soils (Kamal *et al.*, 2000; Zuo *et al.*, 2000; Ma, 2005). The condition can be improved greatly by intercropping these plants with Fe-efficient graminaceous species (Kamal *et al.*, 2000; Zuo *et al.*, 2000; Inal *et al.*, 2007). Greenhouse rhizobox studies demonstrated that the Fe nutrition of peanut or guava intercropped with maize was mainly improved by rhizosphere interactions between peanut or guava and maize (Kamal *et al.*, 2000; Zuo *et al.*, 2000). When peanut and maize grow together, phytosiderophores released by maize roots mobilize Fe(III), which may then diffuse from close to the roots of maize to the root surface of peanuts. At the root surface of peanuts, Fe(III) can be reduced, and thus the release of phytosiderophores from maize roots may improve the iron nutrition of Fe-inefficient plants. Intercropping increases Zn extractability in calcareous soils and the Zn content in guava leaves (Kamal *et al.*, 2000). Similar results have been obtained for the peanut/maize intercropping system, increasing both Fe and Zn acquisition (Inal *et al.*, 2007).

Water

Increased water capture, or water use (WU): Water capture (evapotranspiration) was increased to 604–609 mm in wheat/soybean relay intercropping from 313 to 334 mm for sole wheat and 434 to 359 mm for sole soybean in the southeastern Pampas. The relay intercropping used a much larger fraction of annual precipitation than sole crops, the fraction increased from 0.26 to 0.51 in sole crops and from 0.53 to 0.71 in the wheat/soybean relay intercropping (Caviglia *et al.*, 2004). The crop coefficient K_c, an important item for evaluating crop evapotranspiration, is defined as the ratio of actual crop evapotranspiration to reference crop evapotranspiration. The K_c in winter wheat/spring maize strip intercropping was slightly higher than that in the monocultures because the

evapotranspiration in the intercropping system was higher than that in monocultures (Gao *et al.*, 2009). However, even though some other crop treatments used water differently within the soil profile, there were no differences in WU between spring wheat sole crop, canola sole crop, field pea sole crop, wheat/canola intercropping, wheat/pea intercropping, canola/pea intercropping, and wheat/canola/pea intercropping (Szumigalski and van Acker, 2008).

Water use efficiency (WUE): The ratio of yield or biomass to WU differed with crop combination. Some crop combination enhanced WUE, but some did not. Maize/beans intercropping had WUE as high as maize sole cropping, and intercropping WUE were greater than bean sole cropping in a semiarid region of South Africa (Tsubo *et al.*, 2003). Greater WUE in the intercrop system was attributed to differences in canopy structure and plant biomass production. Faba bean and red fescue had lower WUE than barley and the intercrop (Izaurralde *et al.*, 1994). The combination with *A. saligna* and *S. bicolor* in northern Kenya used the soil water from between hedgerows more efficiently than the sole cropped trees or crops, as water uptake of the trees reached deeper and started earlier after flood irrigation than for sorghum, whereas sorghum better utilized topsoil water (Lehmann *et al.*, 1998). Under the experimental conditions, the root system of alley cropped Acacia and sorghum exploited a larger soil volume and utilized soil resources more efficiently than the respective monocultures(Lehmann *et al.*, 1998).

However, in a 5-year-study in Loess Plateau with sole switchgrass, milkvetch, and switchgrass/milkvetch intercropping, Xu *et al.* (2008) found that WUE in both sole switchgrass and milkvetch was greater than that in switchgrass/milkvetch intercropping ($p < .05$) across an average of 5 years (Xu *et al.*, 2008). In a study of intercropping of wheat, canola, and field pea in Manitoba, Canada, the WUE was not different from that in sole cropping. The overyielding in intercrops was inconsistent and seemed to be related more to LI than to water utilization (Szumigalski and van Acker, 2008). Average WUE in winter wheat/spring maize intercropping system in the Huang-Huai-Hai Plain of China was 21.72 kg ha^{-1} mm^{-1}, which was 23% less than that of the sole maize, but 4% greater than that of the sole wheat. Therefore, although winter wheat/spring maize intercropping system does not improve WUE, it may significantly raise yield (Gao *et al.*, 2009).

Hydraulic lift effect: During periods of drought, water may be transported upward through root systems from moister subsurface to dry surface soil by a process known as 'hydraulic lift.' On warm and dry days when plant transpiration peaks, hydraulically lifted water released into soil can support growth and survival of the lifting and neighboring plants. Soil and rhizosphere microorganisms and the soil fauna could also benefit from hydraulic lift-derived water, which eventually increases the availability of nutrients to plants. Although hydraulic lift was examined mainly in the context of terrestrial plant ecology, this biological subterranean sprinkler process may also prove to be a sustainable alternative to conventional engineered irrigation techniques currently used for agronomical purposes (Liste and White, 2008). Introduction of deep-rooted species into intercropping systems can allow shallow-rooted crops to utilize otherwise unavailable water sources deep in the soil layers. Sekiya and Yano (2004) found that pigeon pea supplied water from deeper soil layers to the associated maize plants through hydraulic lift, whereas sesbania did not. The interspecific difference in the sprinkler-like function of hydraulic lift could be attributable to the differences in root structures between the two legumes (Sekiya and Yano, 2004).

Symbiosis with Mycorrhiza

Vesicular-arbuscular mycorrhizal (VAM) fungi interconnect the root systems of adjacent plants and probably mediate the transfer of nutrients between different plant species. In a mixture of wheat and ryegrass, wheat growth and yield were enhanced and ryegrass growth was inhibited in +VAM herbicide-treated associations (Rejon *et al.*, 1997), which was reversal of sole crops. Rejon *et al.* (1997) explained that wheat, as a stronger sink, was interconnected to associated ryegrass via VAM mycelium between two species; nutrients were transferred from ryegrass roots to wheat by the VAM mycelium (Rejon *et al.*, 1997).

Crop Pest and Disease Control

Crops in mixture are often less damaged by pest and disease organisms than when grown as sole crops (Trenbath, 1993), which is an additional mechanism underlying overyielding.

Push–Pull Principle and Biological Pest Control

Push–pull strategies use a combination of behavior-modifying stimuli to manipulate the distribution and abundance of pest and beneficial insects for pest management. Strategies targeted against pests try to reduce their abundance on the protected resource, for example, a crop or farm animal. The pests are repelled or deterred away from this resource (push) using stimuli that mask host apparency or are repellent or deterrent. The pests are simultaneously attracted (pull), using highly apparent and attractive stimuli, to other areas such as traps or trap crops where they are concentrated, facilitating their elimination (Cook *et al.*, 2007). For instance, intercropping cauliflower with noncrucifer host plants, like sunflower (*H. annuus* L.), marigold (*Tagetes erecta* L.), tomato (*L. esculentum* L.), lucerne (*Medicago sativa* L.), and berseem (*T. alexandrinum* L.), significantly reduced aphid, larvae of diamond back moth and cabbage butterfly incidence and had more numbers of aphidivorous coccinellids, a larval parasitoid. As a result, higher yield in intercropping was observed than in the sole crop of cauliflower (Muthukumar and Sharma, 2009).

Genetic Diversity Controls Crop Disease and Improves Lodging-Resistance

Genetic diversity of crop in mixture controls disease through supporting diverse pathogens. Hybrid and glutinous rice intercropping has been popular and has become more and more important in modern agriculture in the Yunnan Province of China, due to its control of rice blast, a serious disease for rice. Zhu *et al.* (2000) found that disease-susceptible rice varieties, glutinous rice, planted in mixtures with resistant varieties, hybrid rice, had 89% greater yield and blast was 94% less severe than when they were grown in monoculture (Zhu *et al.*, 2000). By polymerase chain reaction (PCR) fingerprinting of pathogen, the genetic composition of the pathogen populations

derived from intercropping and monoculture was determined, and results indicated that the mixtures supported diverse pathogen populations with no single dominant strain. In contrast, pathogen populations from monoculture were dominated by one or a few strains (Zhu *et al.*, 2000). The more diverse pathogen population from the mixed stands may have contributed to greater induced resistance from incompatible interactions (Zhu *et al.*, 2000). Recent results in the system showed that prevention of lodging of the tall glutinous rice was identified as an important additional advantage of growing these rice varieties in mixture as a consequence of the presence of the lodging-resistant hybrid rice (Revilla-Molina *et al.*, 2009).

Species Diversity Controls Crop Disease

In common bean-based intercropping systems, the mean bean rust incidence and severity were reduced by 25% and 16%, respectively, compared with sole cropping. Similarly, bacterial blight incidence was reduced by 29% and severity by 20% (Fininsa and Yuen, 2001), which probably resulted from a slower disease progress rate on beans planted with maize or sorghum in row, mixed, and broadcast intercropping than on bean planted alone (Fininsa and Yuen, 2002). Consequently, overyielding in common bean intercropped with maize or sorghum was observed (Fininsa, 2003). As large as 20% of annual yield losses resulted from leaf spot diseases caused by various pathogens, including the septoria leaf spot complex consisting of *Septoria tritici* Roberge in Desmaz. and *Stagnospora nodorum* (Berk.) Castellani & E. G. Germano, which cause septoria and stagnospora blotch, respectively, and *Pyrenophora tritici-repentis* (Died.) Drechs., which causes tan spot (Hummel *et al.*, 2009). Hummel *et al.* (2009) suggest that wheat disease severity in canola/wheat intercropping decreased with increasing canola proportion, suggesting a possible interference of the canola on the infection of wheat plants (Hummel *et al.*, 2009). At the same time, wheat/canola canopies, which enclosed the interrows and maintained higher humidity, may have exacerbated wheat leaf disease infestations compared to those in monocultures (Hummel *et al.*, 2009).

Large-scale intercropping experiments in 15,302 ha of farmland in Yunnan Province of China have provided evidence that intercropping had significant yield advantages in tobacco/maize, sugarcane/maize, potato/maize, and wheat/faba bean intercropping systems and protected intercropped crops from various diseases (Li *et al.*, 2009). The severity of northern maize leaf blight in intercropped plots was decreased by 17.0% and 19.7% in 2006 and 2007, respectively, compared with the sole plots, although tobacco brown leaf spot disease in intercropping and sole cropping was comparable. The severity of sugarcane eye spot disease was comparable in sole and intercropped sugarcane with maize, but northern maize leaf blight in intercropped plots decreased by 55.9% and 49.6% in 2006 and 2007, respectively, compared with the sole cropping. Severity of potato late blight in the intercropping system was decreased by 32.9% and 39.4% in 2004 and 2005, respectively, compared to corresponding sole potato; simultaneously northern maize leaf blight in intercropped plots was decreased by 30.4% and 23.1%. The severity of broad bean chocolate spot disease in the intercrops also decreased by 33.8% and 31.7% in 2004 and 2005, respectively, compared with sole faba bean (Li *et al.*, 2009). The reduced disease severity in maize was associated with less rainfall during the earlier growing period of the intercropped maize with sugarcane, and that in faba bean with reduced disease spread among faba bean plants because they were separated by rows of wheat plants (Li *et al.*, 2009).

Conclusions

Overyielding of crop mixtures was frequently observed with various combinations of crops all over the world. The mechanisms underlying the overyielding are attributed to both above-ground and below-ground interspecific interactions between crop species. Above-ground interactions include increased LI complementarity across temporal and spatial scales. Below-ground interspecific interactions involve reduction of competition by temporal and spatial differentiation in the distribution of the roots of mixed crops and enhanced nutrient (nitrogen, phosphorus, and some microelements) acquisition by interspecific facilitation. In some but not all cases, a similar enhancement of water capture and water use was observed in crop mixtures. Finally, crop mixtures may enhance genetic and species diversity in agroecosystems in ways that help control crop pests and disease.

There is little information on long-term changes in soil fertility under crop mixture, which is a key factor for sustainability of the cropping systems, because overyielding results in more nutrient and water consumption in crop mixture. Crop mixture, with two or more species, adds some complexity of research; therefore, it is important in future study to employ simulation models as tools to determine the dynamic processes of interspecific interactions in the systems.

Appendix

University Courses

1. Biodiversity Science
2. Agro-biodiversity Science
3. Agroecology
4. Agronomy
5. Crop Cultivation

Acknowledgments

This work was supported financially by the National Basic Research Program of China (973 Program) (Project no 2011CB100405 and 2006CB100206) and the National Natural Science Foundation of China (Project no 30821003, 30890133 and 30870406). We thank Prof. G. David Tilman for his comments and assistance.

See also: Agriculture, Sustainable. Agriculture, Traditional. Agrobiodiversity. Biodiversity and Pest Control Services. Feeding the World and Protecting Biodiversity. Land-Use Patterns, Historic. Plant–Soil Interactions. Poverty and Biodiversity. Species Coexistence. Species Interactions

References

Ae N, Arihara J, Okada K, Yoshihara T, and Johansen C (1990) Phosphorus uptake by pigeon pea and its role in cropping systems of the Indian subcontinent. *Science* 248: 477–480.

Agegnehu G, Ghizaw A, and Sinebo W (2006) Yield performance and land-use efficiency of barley and faba bean mixed cropping in Ethiopian highlands. *European Journal of Agronomy* 25: 202–207.

Aggarwal PK, Garrity DP, Liboon SP, and Morris RA (1992) Resource use and plant interactions in a rice–mungbean intercrop. *Agronomy Journal* 84: 71–78.

Andersen MK, Hauggaard-Nielsen H, Ambus P, and Jensen ES (2004) Biomass production, symbiotic nitrogen fixation and inorganic N use in dual and tri-component annual intercrops. *Plant and Soil* 266: 273–287.

Beschow H, Schulze J, and Merbach W (2000) Transfer of symbiotically fixed nitrogen in an alfalfa–grass mixture studied through isotope dilution in a pot experiment. *Isotopes In Environmental and Health Studies* 36: 21–33.

Bukovinszky T, Trefas H, van Lenteren JC, Vet LEM, and Fremont J (2004) Plant competition in pest-suppressive intercropping systems complicates evaluation of herbivore responses. *Agriculture Ecosystems & Environment* 102: 185–196.

Carter MS and Ambus P (2006) Biologically fixed N_2 as a source for N_2O production in a grass–clover mixture, measured by $^{15}N_2$. *Nutrient Cycling in Agroecosystems* 74: 13–26.

Caviglia OP, Sadras VO, and Andrade FH (2004) Intensification of agriculture in the south-eastern pampas – i. Capture and efficiency in the use of water and radiation in double-cropped wheat-soybean. *Field Crops Research* 87: 117–129.

Chu GX, Shen QR, and Cao JL (2004) Nitrogen fixation and N transfer from peanut to rice cultivated in aerobic soil in an intercropping system and its effect on soil N fertility. *Plant and Soil* 263: 17–27.

Cochran VL and Schlentner SF (1995) Intercropped oat and fababean in alaska – dry-matter production, dinitrogen fixation, nitrogen transfer, and nitrogen-fertilizer response. *Agronomy Journal* 87: 420–424.

Cook SM, Khan ZR, and Pickett JA (2007) The use of push–pull strategies in integrated pest management. *Annual Review of Entomology* 52: 375–400.

Corlett JE, Black CR, Ong CK, and Monteith JL (1992) Above- and below-ground interactions in a leucaena/millet alley cropping system. II. Light interception and dry matter production. *Agricultural and Forest Meteorology* 60: 73–91.

Danso SKA, Palmason F, and Hardarson G (1993) Is nitrogen transferred between field crops - examining the question through a sweet-blue lupin (*Lupinus angustifolius* L) oats (*Avena sativa*) intercrop. *Soil Biology & Biochemistry* 25: 1135–1137.

Davis JHC and Garcia S (1987) The effects of plant arrangement and density on intercropped beans (*Phaseolus vulgaris*) and maize. 1. Traits related to dry matter and seed productivity. *Field Crops Research* 16: 105–115.

Davis JHC and Woolley JN (1993) Genotypic requirement for intercropping. *Field Crops Research* 34: 407–430.

Donald CM (1958) The interaction of competition for light and nutrients. *Australian Journal of Agricultural Research* 9: 421–435.

El Dessougi H, Dreele AZ, and Claassen N (2003) Growth and phosphorus uptake of maize cultivated alone, in mixed culture with other crops or after incorporation of their residues. *Journal of Plant Nutrition and Soil Science-Zeitschrift fur Pflanzenernahrung und Bodenkunde* 166: 254–261.

Erkovan HI, Tan M, Halitligil MB, and Kislal H (2008) Performance of white clover-grasses mixtures: Part ii nitrogen fixation and transfer from white clover to associates grasses. *Asian Journal of Chemistry* 20: 4077–4084.

Fan FL, Zhang FS, Song YN, *et al.* (2006) Nitrogen fixation of faba bean (*Vicia faba* L.) interacting with a nonlegume in two contrasting intercropping systems. *Plant and Soil* 283: 275–286.

Fang SZ, Li HL, Sun QX, and Chen LB (2010) Biomass production and carbon stocks in poplar-crop intercropping systems: A case study in northwestern jiangsu, china. *Agroforestry Systems* 79: 213–222.

Farnham DE and George JR (1993) Dinitrogen fixation and nitrogen transfer among red-clover cultivars. *Canadian Journal of Plant Science* 73: 1047–1054.

Farnham DE and George JR (1994a) Dinitrogen fixation and nitrogen transfer in birdsfoot-trefoil orchardgrass communities. *Agronomy Journal* 86: 690–694.

Farnham DE and George JR (1994b) Harvest management effects on dinitrogen fixation and nitrogen transfer in red-clover orchardgrass mixtures. *Journal of Production Agriculture* 7: 360–364.

Farnham DE and George JR (1994c) Harvest management effects on productivity, dinitrogen fixation, and nitrogen transfer in birdsfoot-trefoil orchardgrass communities. *Crop Science* 34: 1650–1653.

Fininsa C (2003) Relationship between common bacterial blight severity and bean yield loss in pure stand and bean-maize intercropping systems. *International Journal of Pest Management* 49: 177–185.

Fininsa C and Yuen J (2001) Association of bean rust and common bacterial blight epidemics with cropping systems in Hararghe highlands, eastern Ethiopia. *International Journal of Pest Management* 47: 211–219.

Fininsa C and Yuen J (2002) Temporal progression of bean common bacterial blight (*Xanthomonas campestris* pv. Phaseoli) in sole and intercropping systems. *European Journal of Plant Pathology* 108: 485–495.

Fujita K, Ofosubudu KG, and Ogata S (1992) Biological nitrogen-fixation in mixed legume-cereal cropping systems. *Plant and Soil* 141: 155–175.

Fukai S and Trenbath BR (1993) Processes determining intercrop productivity and yields of component crops. *Field Crops Research* 34: 247–271.

Gao Y, Duan AW, Sun JS, *et al.* (2009) Crop coefficient and water-use efficiency of winter wheat/spring maize strip intercropping. *Field Crops Research* 111: 65–73.

Garrity DP, Akinnifesi FK, Ajayi OC, *et al.* (2010) Evergreen agriculture: A robust approach to sustainable food security in Africa. *Food Security* 2: 197–214.

Gebhart DL, Call CA, and Weaver RW (1993) Dinitrogen fixation and transfer in legume-crested wheatgrass mixtures. *Journal of Range Management* 46: 431–435.

Gethi M, Omolo EO, and Mueke JM (1993) The effect of intercropping on relative resistance and susceptibility of cowpea cultivars to *Maruca testulalis* geyer when in mono and when intercropped with maize. *Insect Science and Its Application* 14: 305–313.

Ghaffarzadeh M, Prechac FG, and Cruse RM (1994) Grain yield response of corn, soybean, and oat grown in a strip intercropping system. *American Journal of Alternative Agriculture* 9: 171–177.

Hamel C, Furlan V, and Smith DL (1991) N-2-fixation and transfer in a field-grown mycorrhizal corn and soybean intercrop. *Plant and Soil* 133: 177–185.

Hardter R and Horst WJ (1991) Nitrogen and phosphorus use in maize sole cropping and maize cowpea mixed cropping systems on an alfisol in the northern Guinea savanna of Ghana. *Biology and Fertility of Soils* 10: 267–275.

Harris D, Natarajan M, and Willey RW (1987) Physiological basis for yield advantage in a sorghum groundnut intercrop exposed to drought .1. Dry-matter production, yield, and light Interception. *Field Crops Research* 17: 259–272.

Hauggaard-Nielsen H, Ambus P, and Jensen ES (2001) Temporal and spatial distribution of roots and competition for nitrogen in pea–barley intercrops – a field study employing p-32 technique. *Plant and Soil* 236: 63–74.

Hauggaard-Nielsen H, Ambus P, and Jensen ES (2003) The comparison of nitrogen use and leaching in sole cropped versus intercropped pea and barley. *Nutrient Cycling in Agroecosystems* 65: 289–300.

Hinsinger P (2001) Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: A review. *Plant and Soil* 237: 173–195.

Horst WJ and Waschkies C (1987) Phosphorus-nutrition of spring wheat (*Triticum aestivum* L) in mixed culture with white lupin (*Lupinus albus* L). *Zeitschrift fur Pflanzenernahrung und Bodenkunde* 150: 1–8.

Hummel JD, Dosdall LM, Clayton GW, *et al.* (2009) Canola-wheat intercrops for improved agronomic performance and integrated pest management. *Agronomy Journal* 101: 1190–1197.

Inal A, Gunes A, Zhang F, and Cakmak I (2007) Peanut/maize intercropping induced changes in rhizosphere and nutrient concentrations in shoots. *Plant Physiology and Biochemistry* 45: 350–356.

Izaurralde RC, Chanasyk DS, and Juma NG (1994) Soil–water under conventional and alternative cropping systems in cryoboreal subhumid central Alberta. *Canadian Journal of Soil Science* 74: 85–92.

Izaurralde RC, McGill WB, and Juma NG (1992) Nitrogen-fixation efficiency, interspecies N transfer, and root-growth in barley–field pea intercrop on a black chernozemic soil. *Biology and Fertility of Soils* 13: 11–16.

Jensen ES (1996) Grain yield, symbiotic N-2 fixation and interspecific competition for inorganic N in pea–barley intercrops. *Plant and Soil* 182: 25–38.

Kahmen A, Renker C, Unsicker SB, and Buchmann N (2006) Niche complementarity for nitrogen: An explanation for the biodiversity and ecosystem functioning relationship? *Ecology* 87: 1244–1255.

Kamal K, Hagagg L, and Awad F (2000) Improved Fe and Zn acquisition by guava seedlings grown in calcareous soils intercropped with graminaceous species. *Journal of Plant Nutrition* 23: 2071–2080.

Kanwar RS, Cruse RM, Ghaffarzadeh M, Bakhsh A, Karlen DL, and Bailey TB (2005) Corn–soybean and alternative cropping systems effects on NO3-N leaching losses in subsurface drainage water. *Applied Engineering in Agriculture* 21: 181–188.

Keating BA and Carberry PS (1993) Resource capture and use in intercropping – Solar radiation. *Field Crops Research* 34: 273–301.

Kobes M and Simek M (1998) The influence of grass species on the enforcement of red clover in mixture swards. *Rostlinna Vyroba* 44: 177–181.
Kurdali F, Janat M, and Khalifa K (2003) Growth and nitrogen fixation and uptake in dhaincha/sorghum intercropping system under saline and nonsaline conditions. *Communications in Soil Science and Plant Analysis* 34: 2471–2494.
Kurdali F, Sharabi NE, and Arslan A (1996) Rainfed vetch–barley mixed cropping in the Syrian semiarid conditions. 1. Nitrogen nutrition using N-15 isotopic dilution. *Plant and Soil* 183: 137–148.
Lambers H, Chapin III FS, and Pons TL (2008) *Plant Physiological Ecology*, 2nd ed. New York: Springer.
Lehmann J, Peter I, Steglich C, Gebauer G, Huwe B, and Zech W (1998) Below-ground interactions in dryland agroforestry. *Forest Ecology and Management* 111: 157–169.
Leihner D (1983) *Management and Evaluation of Intercropping Systems with Cassava*. Colombia: Centro International de Agricultura Tropical, Cali.
Li CY, He XH, Zhu SS, *et al.* (2009) Crop diversity for yield increase. *Plos One* 4.
Li L, Li SM, Sun JH, *et al.* (2007) Diversity enhances agricultural productivity via rhizosphere phosphorus facilitation on phosphorus-deficient soils. *Proceedings of the National Academy of Sciences of the United States of America* 104: 11192–11196.
Li L, Sun JH, and Zhang FS (2011) Intercropping with wheat leads to greater root weight density and larger below-ground space of irrigated maize at late growth stages. *Soil Science and Plant Nutrition* 57: 61–67.
Li L, Sun JH, Zhang FS, *et al.* (2006) Root distribution and interactions between intercropped species. *Oecologia* 147: 280–290.
Li L, Sun JH, Zhang FS, Li XL, Yang SC, and Rengel Z (2001a) Wheat/maize or wheat/soybean strip intercropping I. Yield advantage and interspecific interactions on nutrients. *Field Crops Research* 71: 123–137.
Li L, Sun JH, Zhang FS, Li XL, Rengel Z, and Yang SC (2001b) Wheat/maize or wheat/soybean strip intercropping II. Recovery or compensation of maize and soybean after wheat harvesting. *Field Crops Research* 71: 173–181.
Li L, Tang CX, Rengel Z, and Zhang FS (2003) Chickpea facilitates phosphorus uptake by intercropped wheat from an organic phosphorus source. *Plant and Soil* 248: 297–303.
Li L, Yang SC, Li XL, Zhang FS, and Christie P (1999) Interspecific complementary and competitive interactions between intercropped maize and faba bean. *Plant and Soil* 212: 105–114.
Li MG, Shinano T, and Tadano T (1997) Distribution of exudates of lupin roots in the rhizosphere under phosphorus deficient conditions. *Soil Science and Plant Nutrition* 43: 237–245.
Li SM, Li L, Zhang FS, and Tang C (2004) Acid phosphatase role in chickpea/maize intercropping. *Annals of Botany* 94: 297–303.
Licht MA and Al-Kaisi M (2005) Strip-tillage effect on seedbed soil temperature and other soil physical properties. *Soil & Tillage Research* 80: 233–249.
Liste HH and White JC (2008) Plant hydraulic lift of soil water – implications for crop production and land restoration. *Plant and Soil* 313: 1–17.
Ma JF (2005) Plant root responses to three abundant soil minerals: Silicon, aluminum and iron. *Critical Review of Plant Science* 24: 267–281.
Machado S (2009) Does intercropping have a role in modern agriculture? *Journal of Soil and Water Conservation* 64: 55A–57A.
Mariotti M, Masoni A, Ercoli L, and Arduini I (2009) Above- and below-ground competition between barley, wheat, lupin and vetch in a cereal and legume intercropping system. *Grass and Forage Science* 64: 401–412.
Matthews RB, Azamali SN, Saffell RA, Peacock JM, and Williams JH (1991) Plant growth and development in relation to the microclimate of a sorghum groundnut intercrop. *Agricultural and Forest Meteorology* 53: 285–301.
Midmore DJ (1993) Agronomic modification of resource use and intercrop productivity. *Field Crops Research* 34: 357–380.
Midmore DJ, Roca J, and Berrios D (1988) Potato (*solanum spp*) in the hot tropics. 4. Intercropping with maize and the influence of shade on potato microenvironment and crop growth. *Field Crops Research* 18: 141–157.
Ministry of Agriculture, China (2006) *Main Multiple Cropping Systems in China* (in Chinese). Beijing: China Agriculture Press.
Miyazawa K, Murakami T, Takeda M, and Murayama T (2010) Intercropping green manure crops-effects on rooting patterns. *Plant and Soil* 331: 231–239.
Monteith JL (1977) Climate and the efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* 281: 277–294.
Morris RA and Garrity DP (1993a) Resource capture and utilization in intercropping – Non-nitrogen nutrients. *Field Crops Research* 34: 319–334.
Morris RA and Garrity DP (1993b) Resource capture and utilization in intercropping – Water. *Field Crops Research* 34: 303–317.
Mucheru-Muna M, Pypers P, Mugendi D, *et al.* (2010) A staggered maize-legume intercrop arrangement robustly increases crop yields and economic returns in the highlands of Central Kenya. *Field Crops Research* 115: 132–139.
Muthukumar M and Sharma RK (2009) Eco-friendly management of insect pests of cauliflower (*Brassica oleracea* var. *Botrytis*) with intercropping and botanicals. *Indian Journal of Agricultural Sciences* 79: 135–137.
Neumann A, Schmidtke K, and Rauber R (2007) Effects of crop density and tillage system on grain yield and N uptake from soil and atmosphere of sole and intercropped pea and oat. *Field Crops Research* 100: 285–293.
Ofosubudu GK, Sumiyoshi D, Matsuura H, and Fujita K (1993) Significance of soil-N on dry-matter production and N-balance in soybean sorghum mixed cropping system. *Soil Science and Plant Nutrition* 39: 33–42.
Ofosubudu KG, Noumura K, and Fujita K (1995) N-2 fixation, N transfer and biomass production of soybean cv bragg or its supernodulating nts1007 and sorghum mixed-cropping at 2 rates of N fertilizer. *Soil Biology & Biochemistry* 27: 311–317.
Ouma G (2009) Intercropping and its application to banana production in East Africa: A review. *Journal of Plant Breeding and Crop Science* 1: 013–015.
Palmason F, Danso SKA, and Hardarson G (1992) Nitrogen accumulation in sole and mixed stands of sweet-blue lupin (*Lupinus-angustifolius* L.), ryegrass and oats. *Plant and Soil* 142: 135–142.
Pearse SJ, Veneklaas EJ, Cawthray GR, Bolland MDA, and Lambers H (2006) Carboxylate release of wheat, canola and 11 grain legume species as affected by phosphorus status. *Plant and Soil* 288: 127–139.
Rana G, Katerji N, Introna M, and Hammami A (2004) Microclimate and plant water relationship of the 'overhead' table grape vineyard managed with three different covering techniques. *Scientia Horticulturae* 102: 105–120.
Rejon A, Garcia-Romera I, Ocampo JA, and Bethlenfalvay GJ (1997) *Mycorrhizal fungi* influence competition in a wheat-ryegrass association treated with the herbicide diclofop. *Applied Soil Ecology* 7: 51–57.
Revilla-Molina IM, Bastiaans L, van Keulen H, *et al.* (2009) Does resource complementarity or prevention of lodging contribute to the increased productivity of rice varietal mixtures in Yunnan, China? *Field Crops Research* 111: 303–307.
Rusinamhodzi L, Murwira HK, and Nyamangara J (2006) Cotton-cowpea intercropping and its N-2 fixation capacity improves yield of a subsequent maize crop under Zimbabwean rain-fed conditions. *Plant and Soil* 287: 327–336.
Russell G, Jarvis PG, and Monteith JL (1989) Absorption of radiation by canopies and stand growth. In: Russell G (ed.) *Plant Canopy: Their Growth, Form and Function*, pp. 21–39. Cambridge: Cambridge University Press.
Sekiya N and Yano K (2004) Do pigeon pea and sesbania supply groundwater to intercropped maize through hydraulic lift? Hydrogen stable isotope investigation of xylem waters. *Field Crops Research* 86: 167–173.
Semere T and Froud-Williams RJ (2001) The effect of pea cultivar and water stress on root and shoot competition between vegetative plants of maize and pea. *Journal of Applied Ecology* 38: 137–145.
Senaratne R, Liyanage NDL, and Soper RJ (1995) Nitrogen-fixation of and N-transfer from cowpea, mungbean and groundnut when intercropped with maize. *Fertilizer Research* 40: 41–48.
Serin Y and Erkovan HI (2008) Nitrogen use efficiency, symbiotic N-2 fixation and transfer in red clover – grass mixtures. *Asian Journal of Chemistry* 20: 2205–2210.
Singh NB, Singh PP, and Nair KPP (1986) Effect of legume intercropping on enrichment of soil nitrogen, bacterial activity and productivity of associated maize crops. *Experimental Agriculture* 22: 339–344.
Smith ME and Francis CA (1986) Breeding for multiple cropping systems. In: Francis CA (ed.) *Multiple Cropping Systems*, pp. 219–249. New York: Macmillan.
Stern WR (1993) Nitrogen-fixation and transfer in intercrop systems. *Field Crops Research* 34: 335–356.
Sun H, Tang Y, and Xie JS (2008) Contour hedgerow intercropping in the mountains of China: A review. *Agroforestry Systems* 73: 65–76.
Szumigalski AR and van Acker RC (2008) Land equivalent ratios, light interception, and water use in annual intercrops in the presence or absence of in-crop herbicides. *Agronomy Journal* 100: 1145–1154.
Thorsted MD, Weiner J, and Olesen JE (2006) Above- and below-ground competition between intercropped winter wheat (*Triticum aestivum*) and white clover (*Trifolium repens*). *Journal of Applied Ecology* 43: 237–245.
Trenbath BR (1993) Intercropping for the management of pests and diseases. *Field Crops Research* 34: 381–405.
Tsubo M, Mukhala E, Ogindo HO, and Walker S (2003) Productivity of maize–bean intercropping in a semiarid region of South Africa. *Water Sa* 29: 381–388.

Tsubo M, Walker S, and Mukhala E (2001) Comparisons of radiation use efficiency of mono-/inter-cropping systems with different row orientations. *Field Crops Research* 71: 17–29.

Vandermeer J (1989) *The Ecology of Intercropping*. New York: Cambridge University Press.

Wallace SU, Whitwell T, Palmer JH, Hood CE, and Hull SA (1992) Growth of relay intercropped soybean. *Agronomy Journal* 84: 968–973.

Waterer JG, Vessey JK, Stobbe EH, and Soper RJ (1994) Yield and symbiotic nitrogen-fixation in a pea mustard intercrop as influenced by N fertilizer addition. *Soil Biology & Biochemistry* 26: 447–453.

Waterworth JV (1994) Intercropping cotton and groundnut in low and high rainfall areas in eastern Zambia. *Experimental Agriculture* 30: 461–465.

Watiki JM, Fukai S, Banda JA, and Keating BA (1993) Radiation interception and growth of maize/cowpea intercrop as affected by maize plant density and cowpea cultivar. *Field Crops Research* 35: 123–133.

Willey RW (1990) Resource use in intercropping systems. *Agricultural Water Management* 17: 215–231.

Wilson JB (1988) Shoot competition and root competition. *Journal of Applied Ecology* 25: 279–296.

Xiao YB, Li L, and Zhang FS (2004) Effect of root contact on interspecific competition and N transfer between wheat and fababean using direct and indirect N-15 techniques. *Plant and Soil* 262: 45–54.

Xu BC, Li FM, and Shan L (2008) Switchgrass and milkvetch intercropping under 2 : 1 row-replacement in semiarid region, northwest China: Aboveground biomass and water use efficiency. *European Journal of Agronomy* 28: 485–492.

Yin R and He Q (1997) The spatial and temporal effects of paulownia intercropping – the case of Northern China. *Agroforestry Systems* 37: 91–109.

Zhang L, van der Werf W, Bastiaans L, Zhang S, Li B, and Spiertz J (2008a) Light interception and utilization in relay intercrops of wheat and cotton. *Field Crops Research* 107: 29–42.

Zhang L, van der Werf W, Zhang S, Li B, and Spiertz J (2008b) Temperature-mediated developmental delay may limit yield of cotton in relay intercrops with wheat. *Field Crops Research* 106: 258–268.

Zhou XA, Nian H, Yang WY, and Han TF (2010) Status quo and countermeasures of soybean production and development intercropping with other crops in South China (in Chinese). *Soybean Bulletin* 3: 1–2.

Zhu YY, Chen HR, Fan JH, *et al.* (2000) Genetic diversity and disease control in rice. *Nature* 406: 718–722.

Zomer RJ, Trabucco A, Coe R, and Place F (2009) *Trees on Farm: Analysis of Flobal Extent and Geographical Patterns of Agroforestry.* Nairobi: ICRAF. Working Paper no. 89. World Agroforestry Centre.

Zou CY and Li ZJ (2002) Intercropping and relay intercropping. In: Shi YC (ed.) *Chinese Academic Aanon in the 20th Century: Agriculture* (in Chinese). Fuzhou: Fujian Education Press.

Zuo YM, Zhang FS, Li XL, and Cao YP (2000) Studies on the improvement in iron nutrition of peanut by intercropping with maize on a calcareous soil. *Plant and Soil* 220: 13–25.

Crustaceans

Marjorie L Reaka-Kudla, The University of Maryland, College Park, MD, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 915–943, © 2001, Elsevier Inc.

Glossary

Antennules, antennae (pl.) Anntennules are the first pair of sensory appendages on the head, and antennae are the second pair of sensory appendages on the head of crustaceans.
Benthic, epibenthic Living at the bottom of the sea; *epi* = upon, thus epibenthic organisms live on top of sediment or hard substratum.
Brackish Less salty than sea water.
Chelate, subchelate, chela (s.), chelae (pl.) Chelate means having a pincher-like claw (usually due to an extension of the second from terminal segment beside the terminal segment, forming the claw). In subchelate forms, the terminal segment merely folds back on the second from terminal segment. The chela is the clawlike appendage.
Commensal, symbiotic Living in association with (e.g., on or in) another organism.
Crepuscular Active in dim light conditions, such as dusk or dawn.
Dorsal, ventral Dorsal part of the organism is the "back," on the opposite side of the body from the mouth; ventral is the same side of the body as the mouth and legs for a crustacean.
Extant Living.
Furca (s.), furcae (pl.) Forked structure, usually associated with the telson in crustaceans.
Gonochoristic Individual is either a male or a female throughout life.
Hypersaline More salty than seawater.
Interstitial Living among sand or silt grains; usually refers to an aquatic benthic environment, the intertidal, or in groundwater.
Invertebrates Animals (usually used in reference to phyla or major taxa) that lack backbones.
Macrobiota Organisms large enough to see, usually greater than 1 mm.
Median, medial Along or toward the midline of the body.
Phyllopod A leaflike appendage.
Pleopod Abdominal appendage in crustaceans.
Polyphyletic A group of species or lineages that have independent origins instead of being descended from a common ancestral group.
Protogynous, protandrous Different types of hermaphroditism (in contrast to gonochoristic). Individuals in protogynous species develop into a reproductive female early in life, then change sex and become a reproductive male later in life. Individuals in protandrous species are functional males first, then females.
Rami (pl.) A cuticular extension (usually relatively long and thin, can be leaflike) from the body (e.g., caudal rami = extention from the posterior segment of the body, usually paired in crustaceans).
Setae (pl.) Bristle- or hairlike extensions of the cuticle that are characteristic of arthropods and related phyla.
Taxon (s.), taxa (pl.) A group of organisms with shared similarities at any level of a taxonomic scheme (could be a species, genus, family, order, etc.).
Telson Most posterior segment of the body in crustaceans, usually flattened and often armored.
Thoracopod Appendage on the thorax (paired in crustaceans).
Uniramous, biramous, triramous, polyramous An extension of the body (usually an appendage) having one, two, three, or many branches (e.g., triramous antennae have three main branches or flagellae).
Vestigial An organ or structure that is present only in rudimentary form (probably lost from a previous condition).

Summary and Definition of Biodiversity in Crustacea

In the last two decades of the twentieth century, it became evident to both the scientific and much of the lay public that the health of the earth's living biota—the foundation of human existence—is seriously threatened by environmental changes, mostly generated by the exponentially increasing activities of humans. This realization was made more poignant by the simultaneous recognition that our knowledge of the diversity of the world's animals, plants, fungi, and mircoorganisms is woefully inadequate even to assess the amount of diversity present and the degree of risk faced by different components of the global biota (Wilson, 1985, 1992).

The term "Biological diversity" and its shortened form "biodiversity" were introduced by Norse and McManus, (1980), Lovejoy (1980), and the National Forum on BioDiversity, held at the Smithsonian Institution (1986). Whereas "biological diversity" was first defined to include the concepts of genetic diversity (variability within species) and ecological diversity (number of species in a community of organisms; Norse and McManus, (1980), the term "biodiversity" was formalized and popularized in the Proceedings of the National Forum (Wilson and Peters, 1988) to include an appreciation for the range of all of the different levels and types of biotic diversity found on earth, including the processes that generate and maintain diversity, and is thus followed here. "Diversity" sometimes implies concepts about the degree of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00165-9

eveness in abundance among species as well as number of species, but, for simplicity and convenience, the terms "diversity" and "biodiversity" will be used interchangeably here to indicate the number of species or higher taxa in a group and to encompass the range of structural variation, habitat characteristics, ecological roles, life history traits, and biogeographical characteristics of particular groups of organisms. The diversity of these features, including the uniqueness of certain adaptations or lineages, often can be comprehended best in the context of the ecological and evolutionary processes that generate these patterns of variation.

To properly evaluate biodiversity in Crustacea, we need to know the range of major ways of life of different groups of crustaceans, as indicated in their body plans. These architectural differences are reflected in the number of higher taxonomic groups within taxa and provide insights into the diversity of ecological processes in which the group participates.

Ecological diversity includes the range of habitats that the group occupies, whether the group lives in unique environments or species assemblages, and the diversity and importance of the ecological roles that taxa within the group play in their natural communities. For example, are the species low on the food chain (thus supporting the diversity of the rest of the community), do they affect community structure as top predators, do they determine diversity of the community by engineering the physical environment? Other important components of biodiversity include the phylogenetic history and uniqueness of lineages. Some groups of Crustacea represent the only surviving remnants of lineages that have persisted for half a billion years. Because of their early derivations, such groups are likely to be unique in their genetic, morphological, physiological, or ecological adaptations. Such a historical legacy merits special conservation attention. Similarly, the diversity of life history patterns is intimately related to genetic variability within and among populations. Some crustaceans show rapid maturation, short life spans, and massive reproductive effort, while others live for extended periods (often many years) and are morphologically and behaviorally adapted for strong interactions with other species. Some produce a few large offspring that develop directly into juvenile stages similar to the adults and do not disperse far from the parental population. Other species release huge numbers of tiny offspring, entirely different in form from the adults, that feed for extensive periods in the plankton and disperse great distances. Many of these patterns track within lineages and are correlated with body size. High variability among individuals within populations, frequently associated with wide dispersal of pelagic offspring and wide geographic ranges of species, may enable a population to persist in spite of environmental changes. However, high variability among populations (often associated with lower gene flow and isolated endemic populations) may increase the probability of losing entire genetic complements if local populations are eliminated by environmental change. Distributional patterns, particularly at the species level, are one of the most important and neglected aspects of biodiversity in most groups of organisms, including the Crustacea. The extent of the geographic range of a group, and of species within the group, are some of the most important predictors of extinction in the fossil record (Jablonski, 1991). Organisms that inhabit highly specialized habitats (especially when these habitats are geographically restricted), occupy high trophic levels, or are of large body size are likely to be vulnerable to extinction because of relatively low total population sizes and their dependence on a reliable food supply or special micro-habitat (both of which fluctuate in many environments). Understanding the paleontological history of crustaceans is essential for understanding the present configuration and future prospects for crustacean biodiversity. The fossil record indicates that recoveries from major mass extinctions require about 10 million years, rendering the process of diversification irrelevant for the survival of humans and other biota over the next 500 years. However, understanding vulnerability to extinction in the past and in the present rapidly changing environment is crucial for assessing present crustacean biodiversity and sustaining it for future generations.

The Setting

Biodiversity in Major Global Habitats, with Special Reference to Marine Environments

Marine environments, the habitat in which crustaceans originated and which continues to sustain the greatest proportion of crustacean species, contain more higher-level biodiversity (different types of organisms represented, for example, as phyla, classes, orders, families, and genera) than other habitats because life arose in the sea, and many of these ancient lineages are still with us. Of plants, animals, and fungi large enough to see, 43 phyla live in the ocean and 28 live on land. Within the 33 animal phyla, 64% inhabit only the sea while only 5% are exclusively land dwellers, and 90% of all known classes are marine (May, 1994; Ray, 1991). It is important for humans to know and understand the diversity of these marine lineages, with all of their bizarre body plans and adaptations, in order to understand how life evolved and how we fit into it. Of more immediate urgency, marine organisms provide a critical source of protein for humans. This fact will become more important in the 21st century and beyond (especially in the world's tropical countries, which are experiencing rapid population growth and associated environmental degradation). Possibly even more significant in the long run, these marine groups provide a repository of genetic and chemical material that has been evolving, each in its own direction, for more than half a billion years. This genetic and chemical material is much more likely to be novel than the more recent, better-known lineages that invaded land.

Despite its huge expanse and the greater number of higher level groups it supports, the ocean contains fewer known species than the land. Only about 274,000 or 14% of the world's 1.9 million described species are known from the ocean (Reaka-Kudla, 1997). Of these, approximately 200,000 are macrobiota (vertebrates, invertebrates, and plants large enough to see) and 180,000 are macroinvertebrates (groups of marine invertebrates sufficiently big to see). Global freshwater environments contain approximately 44,000 total species (2.4% of all those described), 35,000 species of macrobiota and 26,000 species of macroinvertebrates. There are about

950,000 insects, 80,000 noninsect and noncrustacean (mostly chelicerate) arthropods, and 57,000 other terrestrial macroinvertebrates. Globally, about 100,000 species live symbiotically on or in other organisms (5.3% of all described species); roughly 73,000 of these probably are macrobiota (mostly invertebrates). If we exclude the noncrustacean terrestrial arthropods (chelicerates, insects, and their relatives) because of their unique specializations for life on land, there are approximately 336,000 described macroinvertebrate species in the marine, freshwater, terrestrial, and symbiotic environments. Of these, marine macroinvertebrates constitute 53.6%, freshwater macroinvertebrates 7.7%, terrestrial macroinvertebrates 17.0%, and symbiotic macroinvertebrates 21.7% of the total species. We can use these figures to estimate the numbers of crustacean species that occupy different types of habitats.

There are about 43,000 species of described crustaceans—2.3% of the 1.9 million described species and 12.8% of all macroinvertebrates. If crustaceans inhabit major types of environments in about the same proportions as other nonchelicerate and noninsect macroinvertebrate groups, probably about 23,000 described species of crustaceans live in marine and brackish coastal waters (figures rounded to nearest thousand); this represents about 12% of all species of marine macrobiota. Similarly, if crustaceans have invaded fresh water to about the same degree as other nonchelicerate and noninsect invertebrates, at least 3400 crustacean species likely inhabit fresh waters (rounded to nearest hundred); this represents about 10% of all described freshwater macrobiota. If crustaceans are about equally adapted to land as are other nonchelicerate and noninsect invertebrates, about 7300 crustaceans are terrestrial. This is less than 1% of the global terrestrial macrobiota. Since about 22% of all macroinvertebrate species are symbiotic, probably about 9,500 crustacean species are estimated to be symbionts (about 13% of global symbiotic macroinvertebrates). In fact, parasitism is rampant in some groups of crustaceans. Certainly these figures are underestimates—most likely great underestimates—of the true numbers of species in aquatic and particularly marine environments because these environments are so poorly known. If only 10 to 20% of all crustacean species have been described (discussed later), then there likely are about 200,000 to 450,000 total (known and unknown) species of crustaceans on earth.

The following observations suggest that the numbers of crustacean species, as well as those of other marine phyla, usually are underrepresented. Based on samples of 1597 species of soft-bottom marine macrofauna off the east coast of North America (255 to 3494 m), Grassle and Maciolek (1992) calculated that the global deep-sea fauna, primarily because of the huge area it occupies, may include as many as 10 million species (mostly polychaete annelids, crustaceans, and mollusks). Most of these species are small and rare (90% of the species sampled comprised less than 1% of the individuals; 28% of the species in the entire fauna were collected only once). Reaka-Kudla (1997) calculated that global coral reefs contain between 1 and 5 million total species (though only about 100,000 described species).

Experts have estimated that about 21% of global crustacean species have been described and that about 26% of global molluscan species are known (Systematics Agenda 2000, 1994). These phyla represent the two most commercially important and thoroughly documented invertebrate groups. Only 10% of the relatively well-known isopod crustaceans collected from shallow sediments in southern Australia had been described (Poore and Wilson, 1993). Because of the great variation in the extent to which different regions of the oceans have been sampled, these scientists estimated that only 5% of marine invertebrates are known from the oceans overall. Calculations suggest that about 10% of the species on global coral reefs have been described. Scientific descriptions exist for about 17% of the species of algae. The microbiota is even more poorly known. Only 1%, 1 to 10%, 4 to 7%, and 2 to 3% of the total species of the viruses, bacteria, fungi, and nematodes have been described.

Although the ancient major groups (phyla, classes) of marine organisms are more different from each other than those on land, species within marine invertebrate groups often look alike, leading to underestimates of their true numbers. Recent molecular and genetic studies have shown that described species frequently include several different, previously unrecognized species (Knowlton, 1993).

Major new groups of organisms are still being discovered in the sea, whereas this is rare on land. At least three new invertebrate phyla (the largest category of related organisms) and many major new groups (classes, orders, families, genera, which represent increasingly smaller groupings of related organisms) of deep-sea organisms, picoplankton, marine viruses, and archaebacteria (a new kingdom of one-celled organisms that often live in extreme environments) have been described over the past 20 years. In the crustaceans, a new class, the primitive Remipedia, were discovered in caves in the Bahamas and described only in 1981. A new subclass of the Maxillopoda, the Tantulocarida (tiny ectoparasites of deep sea copepods, ostracods, and isopods) was erected in 1983. In the relatively well known malacostracan crustaceans, the Order Spelaeogriphacea was not discovered until 1957, the Order Amphionidacea recognized only in 1973, and the Order Mictacea established as late as 1985 (Schram, 1986). Crustacean biology, and marine biology in general, remains a realm of discovery.

Aquatic environments are difficult to see, sample, and study. The number of scientists who study invertebrates and the amount of scientific effort devoted to invertebrates other than insects is two orders of magnitude less than that devoted to vertebrates and one order of magnitude less than that devoted to plants (Gaston and May, 1992). In addition, tropical environments, far from the location of most biologists, have been much less studied than those at higher latitudes; 80% of ecological researchers are based in North America and Europe in comparison to 7% in Latin America and tropical Africa. Identification of species in aquatic habitats is further reduced because of the inability to observe them easily and the frequent requirement for Scuba diving and extensive logistic support.

Moreover, most crustacean and other aquatic taxa are small in body size. Many aquatic species live within crevices, holes, or burrows in coarse (rocks, coral, rubble) and level (mud, sand) substrata. Many are crepuscular or nocturnal to minimize fish predation (Dominguez and Reaka, 1988), making them difficult to see, collect, and study. Small organisms of

any environment always are poorly known and studied, but this is more extreme in aquatic environments, particularly in tropical marine environments, because of the remote localities and the difficult conditions for observing and collecting.

It seems likely that major advances in our knowledge of the extent of crustacean biodiversity will continue into the next centuries. These advances are imperiled, however, by a crisis in the number of systematists and taxonomists being trained (Feldmann and Manning, 1992; Reaka-Kudla *et al.*, 1997; Wilson, 1985). Although we know a fair amount about the morphology, ecology, behavior, life history, and development of many of the crustacean groups, we know astonishing little about the facets of their biology that are most critical for assessing their biodiversity status, particularly the distribution and abundance patterns of individual species and the life history features that render the species vulnerable or resistant to extinction. More distressing, we are incapable of determining these aspects of the biology of given taxa without a reliable foundation of systematics and taxonomy. The understanding and conservation of crustacean biodiversity will depend on the implementation of major new funding and training programs that are directed toward systematics and taxonomy, and toward natural history museums which hold the collections necessary for such comparative studies.

Ecology of the Arthropod Radiation

Between 3 and 4 billion years ago, biopolymers (proteins, polysaccharides, nucleic acids), which had been formed as a result of physical processes, became enclosed in membranes, and the resulting cells developed the ability to replicate themselves. These duplicated cells became organized into multicellular organisms that could move about on their own about a billion years ago. An accelerating explosion of biotic diversity occurred sometime in the Precambrian (perhaps 700–800 million years ago), but fossil remnants of this early proliferation are few, having been either poorly preserved or destroyed in subduction zones beneath migrating continents. By the beginning of the Paleozoic era (about 570 million years ago (mya)), almost all of the major phyla of animals already were present (by then being larger and more amenable to fossilization or preserved in continental rocks that are still available). The Cambrian marine environment (approximately 570–500 mya) was studded with sessile epibenthic forms such as stromatolites (aggregations of unicellular organisms, their secretions, and sand or other sedimentary particles); sponges and other now extinct spongelike forms; upright branched and laminar bryozoans (“moss animals”); filamentous now extinct organisms called graptolites, brachiopods (“lamp shells”), and other sessile organisms with ciliated feeding arms; various types of cnidarians (hydroids, anemones, primitive types of corals, jellyfish); stalked crinoids (sea lillies) and other primitive echinoderms; and the sessile life stages of primitive forms of chordates (our own ancient forebearers). Most of these sessile organisms fed on suspended particles near the seafloor.

Various other organisms crawled on the seafloor, many of them having developed segmentation (and other architectural features that allowed them to move more rapidly and feed more effectively) as well as external armor that protected them from the increasing tendency for their ilk to eat one another. In addition to many lineages (often with bizarre features unlike any organisms known today) that perished in the race for motility and are known only as fossils, the motile benthic fauna included the early mollusks (whose followers would diversify during the Paleozoic into major grazers, suspension feeders, burrowing sediment feeders, and large extraordinarily active predators). Also present were diverse groups of early echinoderms (that would become grazers, burrowing sediment feeders, and slow-moving predators as the Paleozoic progressed) and the early filter-feeding chordates (many of which would develop into large, increasingly active and important predators as jawless and then jawed fish during the early to mid-Paleozoic). At least as important for biodiversity in the long run, the Precambrian benthic marine environment abounded with a diverse array of unsegmented and segmented worms that provided the prototype for the development of the joint-legged animals, the arthropods.

Whether the arthropod phyla (uniramians (centipedes, millipedes, insects), onycophorans, tardigrades, chelicerates (horseshoe crabs, scorpions, spiders, ticks), pycnogonids, trilobites, pentastomids, and crustaceans) arose from only one group or independently from several groups of the early segmented annelid worms is hotly debated. The annelids and arthropod groups originated in the Precambrian, and all of the major phyla were distinct by the time they were fossilized. The crustaceans, trilobites, and chelicerates appeared first in the fossil record, followed at a later time by the uniramians. Some experts have concluded, on the basis of functional morphology and developmental patterns, that the major arthropod groups probably all arose independently, and that the Crustacea are not closely related to the uniramians, trilobites, or chelicerates. Other scientists have argued, on the basis of morphology, that the trilobites are ancestral to both the chelicerates and crustaceans. New evidence from RNA and DNA sequences has brought forth various views, including the idea that the crustaceans are more closely related to the uniramians than to other arthropod groups. The question of how the arthropod groups evolved, then, remains far from solved at present.

Many workers concur that the onycophorans (velvet worms) and the uniramians have a common origin. Some workers have suggested a possible connection between two Cambrian marine fossils and the onycophorans, who are represented today by 80 species in warm moist terrestrial environments, The first myriapods (centipedes, millipedes) were fossilized in the Silurian (440–405 mya) and probably they or their ancestors were marine. Terrestrial myriapods and wingless insects are known from the Devonian (405–360 mya), and winged insects had arisen by the Carboniferous (360–290 mya). These underwent a continuing spectacular radiation that resulted in almost a million species of described insects that inhabit almost all terrestrial and many aquatic environments today. In addition to 32 living orders of insects, at least 10 other orders arose and went extinct in the late Paleozoic and early Meszoic.

About 400 species of tardigrades (water bears) are known from marine (intertidal to deep-sea) environments, freshwater, hot springs, and semiterrestrial environments in plants

and humus; the most primitive group is marine. A single fossil is known from Cretaceous (135–65 mya) amber. These minute organisms (most 0.1–0.5 mm, a few to 1.7 mm) occupy benthic, interstitial, water film, commensal, and parasitic habitats. They also have been been collected from some of the driest habitats on earth, the Antarctic dry valleys; they are capable of being freeze-dried, after which they resume metabolism when rehydrated.

Represented by the xiphosurans (horseshoe crabs) and the eurypterids (sea scorpions), the early chelicerates were marine and had appeared by the Cambrian. Eurypterids and xiphosurans were crawling scavengers and predators, crushing food between the bases of their legs. The anterior appendages (chelicerae) of some eurypterids were well developed and clawed. Many eurypterids were small but some grew to very large sizes (up to 3 m). These marine forms flourished during the Ordovician, Silurian, and Devonian (500–360 mya) and were abundant until the Permian (290–245 mya), after which the eurypterids became extinct and the xiphosurans were greatly diminished (only five species of horseshoe crabs remain today). Some eurypterids apparently invaded fresh water and may have become semiterrestrial. Related lineages, including the scorpions (the most primitive living arachnids) were present in the Silurian and invaded land very early, at least by the Carboniferous. Diversification of the Chelicerata. resulted in a total of about 80,000 described, mostly terrestrial species living today.

The pycnogonids (sea spiders), considered a sister group to the Chelicerata, are known from the Devonian onward. They are represented by about 1000 species and 86 genera today, all marine. They are suctorial predators and occupy all oceans down to about 7000 m. Most are small (<10 mm), though the bodies of some deep sea forms are 20 to 30 mm with leg spans of up to 700 mm. The "protonyphon" larva of some pycnogonids and the "trilobite" larva of some xiphosurans share some developmental features with the "protaspid" larva of trilobites.

The trilobites, along with the chelicerates, were major components of the motile benthic fauna of the early Paleozoic. The trilobites were exceedingly abundant during the Cambrian and Ordovician (600–440 mya) and persisted in somewhat lower numbers until the end of the Permian (245 mya), after which they became extinct. All of the 4000 species were marine. They were major scavengers and predators, crushing food between the bases of their legs, which operated as gnathobase jaws. Some were possibly sediment or suspension feeders. Most species were small (10–30 mm), some were moderately sized (30–60 mm), and a few giants (600–700 mm) are known. Most species were benthic and epifaunal or shallow infaunal, but some small (<10 mm) forms were planktonic.

The pentastomids, wormlike parasites that inhabit the respiratory system of terrestrial vertebrates, include about 95 species and are thought to be related to crustaceans. They may have been a relatively late offshoot of early arthropod lineages, developing with or after the radiations of terrestrial vertebrates in the Carboniferous (amphibians), Permian (reptiles), and Mesozoic (reptiles, birds, mammals).

In the midst of this evolutionary theater rose the crustaceans. We shall return to the historical development of crustacean diversity after we are introduced to the major characters.

Biodiversity of Fossil and Modern Crustacea

The Major Groups of Crustacea

Centers of Taxonomic Diversity

At least five major radiations in ancestral crustacean lineages produced the remipedes, cephalocarids, branchiopods, maxillopods, and malacostracans (Table 1). Scientists believe, either because of their primitive body forms or their fossil record, that all five major groups have ancient origins extending back at least as far as the beginning of the Paleozoic. The remipedes, cephalocarids, and branchiopods probably are impoverished relicts of ancient great lineages that did not fossilize well. Today, however, these three groups (along with a number of primitive malacostracan lineages) are marginalized in hypersaline, brackish, and fresh water (sometimes hot springs or ephemeral pools), caves, groundwater, or other isolated, seemingly suboptimal habitats. Collectively, these three classes contribute only 14% of the diversity in orders, 5% of the diversity in families, 2% of the generic diversity, and 2% of the species diversity within modern crustaceans. The remaining 2 classes—the maxillopods and especially the malacostracans—are the dominant groups in most major marine environments today and are responsible for much of the diversity in major body plans (e.g., as represented in subclasses; Table 1). The malacostracans alone constitute over half (53%, 59%, and 55%) of all present and past (fossilized) crustacean families, genera and species, respectively. The major groups of malacostracans have been subject to considerable turnover in the fossil record, however, since the lineage is ancient and many of their major evolutionary experiments (body plans at the level of order) have been lost to extinction. Whereas 80 to 100% of all the orders of branchiopods and maxillopods ever known are still extant, only 59% of malacostracan orders are still living. If their order did survive background extinction and the great mass extinctions (e.g., at the ends of the Paleozoic and the Mesozoic), however, malacostracan families have persisted (or are currently diversifying) at a higher rate (88% of all known families are still extant) than those of the maxillopods (76% living) and branchiopods (74% extant). Though only 87% of branchiopod genera still exist, very high percentages of genera in both the malacostracans (97%) and maxillopods (96%) are alive today.

Within the maxillopods, centers of taxonomic diversity are found in the copepods, ostracods (seed shrimp), and, to a lesser extent, cirripeds (barnacles). Copepods include 38% of the orders, 50% of the families, 54% of the genera, and 49% of the species in the maxillopods. Ostracods constitute 29% and 38% of the orders and families, and 36% and 44% of the genera and species in the maxillopods. Barnacles comprise 19% and 11% of the orders and families, and 10% and 6% of the genera and species in the group. Both ostracods and barnacles are ancient lineages whose fossil records extend back at least to the dawn of the Paleozoic.

Within the malacostracans, peracarids and eucarids dominate the group in terms of taxonomic diversity. Peracarids

Table 1 Minimum numbers of taxa for the five major crustacean evolutionary radiations

Class	*Subclasses*	*Orders*	*Families*	*Genera*	*Species*
Class Remipedia	—	2 (1 living), **4%**	1 (1 living), **0.1%**	6 (5 living), **0.1%**	11 (10 living), <**0.1%**
Class Cephalocarida	—	2 (1 living), **4%**	2–3 (1–2 living), **0.2%**	At least: 5 (4 living), **0.1%**	At least: 11 (10 living), <**0.1%**
Class Branchiopoda	—	4–5 (4 living), **7%**	At least: 35 (26 living), **4%**	At least: 108 (94 living), **2%**	At least: 836 (821 living), **2%**
Class Maxillopoda	6 living	21 (18 living), **38%**	355 (270 living), **43%**	At least: 2162 (2077 living), **39%**	At least: 18,388 (18,293 living) **43%**
	Ostracoda	6 (3 living), 29%	136 (55 living), 38%	At least: 776 (695 living), 36%	At least: 8081 (about 8000 living), 44%
	Mystacocarida	1 living, 5%	1 living, 5%	2 living, 0.1%	10 living, 0.1%
	Copepoda	8 living, 38%	177 living, 50%	1167 living, 54%	About 9000 living, 49%
	Branchiura	1 living 5%	1 living, 0.3%	4 living, 0.2%	150 living, 1%
	Tantulocarida	1 living, 5%	1 living, 0.3%	2 living, 0.1%	5 living, <0.1%
	Cirripedia	4 living, 19%	39 (35 living), 11%	At least: 211 (207 living), 10%	At least: 1142 (1138 living), 6%
Class Malacostraca	2–7 (usually 5) living	27 (16 living) (3 extinct groups unassigned to subclasses), **48%**	413 (383 living), **53%**	At least 3175 (3093 living), **59%**	At least: 23,890 (23,800 living) **55%**
	Phyllocarida	5 (1 living), 19%	13 (2 living), 3%	At least: 55 (5 living), 2%	At least: 65 (13 living), 0.3%
	Hoplocarida	3 (1 living), 11%	At least: 19 (13 living), 4%	At least: 78 (68 living), 2%	At least: 413 (400 living), 2%
	Syncarida	3 (2 living), 11%	10 (6 living), 2%	At least: 37 (33 living), 1%	At least: 119 (115 living), 0.5%
	Peracarida	10 (9 living), 37%	268 (255 living), 62%	At least: 1793 (1775 living), 56%	At least: 12,726 (12,706 living), 53%
	Eucarida	3 living, 11%	121 (107 living), 28%	At least: 1213 (1212 living) (fossils mostly not quantified), 38%	At least: (10,567 living) (fossils mostly not quantified), 44%
Totals	–	56–57 (40 living)	At least: 824 (681 living)	At least: 5457 (5273 living)	At least: 43,136 (42,934 living)

In cases where number of taxa is controversial, the lower number was used in the total. In many cases, the number of certain taxa (especially genera and species) were not summarized in sources, so minimum numbers were estimated from the number of higher taxa (e.g., in a group with five families for which data on number of genera were not summarized, it was assumed that at least five genera (one per family) and at least five species (one per genus) were present). Therefore, these data probably considerably underestimate the number of genera and particularly the number of species in these classes (although total species exceed those previously estimated by other authors, e.g., Brusca and Brusca; 1990, who did not include paleontological data). Percentages horizontally allied with classes (bold) represent the total number in that taxon compared to the number for all Crustacea. Percentages horizontally aligned with subclasses (not bold) represent the total number in that taxon compared to the total number in that class. Percentages are based on total (fossil and living) taxa and are rounded to the nearest whole number (except when <1%).

include 37% of the orders and 62% of the families, and at least 56% of the genera and 53% of the species, in the Malacostraca. The eucarids make up only 11% of the orders and 28% of the families, but at least 37% of the genera and 44% of the species within the group.

Architectural Diversity

To understand the diversity of Crustacea, it is necessary to come to grips with their major ways of life as reflected in their body plans. The range of morphological diversity among crustaceans is remarkable, considerably greater than that of the insects. Primitively, each segment supported a pair of appendages. Crustacean appendages have multiple joints along their length and are considered to have evolved from a biramous condition (divided into two branches), although both head and trunk appendages have lost one side of the branch in some groups. The primary mode of evolutionary change in the early diversification of crustaceans was by molding segments into functional units along the length of their body and reducing various morphological features, including both segments and appendages. The resulting fundamental body regions included a five segmented head (cephalon) and a long, segmented trunk that is usually divided into a thorax and abdomen.

The head was formed by telescoping five segments and their associated appendages into a fused unit that bears two pairs of antennae (a distinguishing characteristic of crustaceans), the mandibles (jaws) that flank the mouth, and two pairs of accessory food handling appendages, the maxillules and maxillae. The antennules (first pair of antennae) are

biramous in the primitive remipedes. However, the antennules are uniramous in the cephalocarids, all of the branchiopods, all of the maxillopods (where present), and some of the eumalacostracans, while those of the primitive phyllocarid malacostracans and some of the eumalacostracans are biramous and those of stomatopods triramous, raising the question of whether or not the antennules of crustacean ancestral groups all were biramous. The antennules are reduced or lost in the parasitic branchiurans and tantulocaridans and in many sessile adult cirripeds. The second antennular appendages (antennae) are biramous in most crustaceans but are uniramous in anostracans, some copepods, leptostracans, some but not other fossil phyllocarids, and some eumalacostracans. The antennae are reduced in the parasitic branchiura (modified for attachment) and lost in the notostracans, tantulocarids, and sessile adult cirripeds.

One or more of the anterior thoracic segments may be fused to the head (remipedes, isopods, amphipods), or the anterior thoracic appendages (maxillipeds) may be functionally associated with the mouthparts of the head (fossil kazacharthran branchiopods; fossil lipostracans; ostracod, mystacocaridan, and copepod maxillopods; stomatopod and possibly palaeostomatopod hoplocaridans; anaspidacean syncarids; mysidan, lophogastridan, tanaidacean, spelaeogriphacean, cumacean, mictacean, isopod, and amphipod peracarids; stomatopods and many eumalacostracans).

Reflecting very early divergence among groups of crustaceans, the division of the trunk into units that functioned in feeding (the anterior appendages), reproduction, and locomotion (often with accessories that aided respiration) was much more variable among major crustaceans groups than the head region. The wormlike remipedes usually are considered to represent the most primitive crustacean morphology. Their oarlike appendages are undifferentiated throughout the long trunk, followed by an anal segment with simple caudal rami. The trunk of the cephalocarids, however, is divided into 8 thoracic segments, 11 abdominal segments, and an anal segment with caudal rami. The trunk region of both the branchiopods and maxillopods is variable among groups. Several groups with modified morphology possess simply a reduced trunk (cladocerans, conchostracans, ostracods, mystacocarids, and adult tantulocarids). A thorax is present in the anostracans and notostracans (usually 11 segments), extinct kazacharthrans, copepods (6 segments), branchiurans (4 segments), tantulocaridans (6 segments in juveniles), cirripeds (6 segments, reduced), and malacostracans (8 segments). The number of segments included in the abdomen is variable within several groups of branchiopods (anostracans, notostracans, fossil kazacharthrans, conchostracans) and in juvenile tantulocarids and ascothoracidan barnacles. The abdomen is reduced or lost, or segmentation obscured, in groups that are modified for specialized ways of life (cladocerans, the parasitic branchiurans and tantulocaridans, and most cirripeds, especially the parasitic forms). The primitive phyllocarid malacostracans all possess an abdomen with seven segments, in addition to a terminal anal segment with caudal rami. Adult hoplocaridans and eumalacostracans have lost one of the abdominal segments, however, and bear six abdominal segments plus an additional flattened anal segment termed a telson. Morphological and developmental evidence of a vestigial seventh abdominal segment (either at the front or the posterior of the abdomen) in the stomatopod hoplocaridans and in the mysidan, tanaidacean, and possibly the isopod eumalacostracans reflects another evolutionary tendency toward reduction of segments.

Trunk appendages in the remipedes are oarlike but different from the leaflike or phyllopodous appendages of the primitive branchiopods. Leaflike or phyllopodous appendages predominate in the cephalocarids, all of the branchiopods, and in the primitive malacostracans, including a number of extinct groups and the modern leptostracans. However, greatly modified or reduced numbers of trunk appendages (especially on posterior trunk segments) characterize the maxillopods (e.g., the ostracods, mystacocarids, and copepods). Parasitic forms (e.g., the adult rhizocephalan barnacles) often lose all appendages. The copepods, branchiurans, and juvenile tantulocarids all have biramous thoracic legs. In the malacostracans, however, some of the thoracic appendages are uniramous and some biramous in stomatopods, and all eight pairs of the thoracopods usually are uniramous (with some of these enlarged and clawed) in the eumalacostracans. Even some of these groups, however, show reduction in development of appendages toward the posterior of the thorax. Branched or layered structures arising from the base of the thoracic legs (epipods) serve as gills in the hoplocaridans and eumalacostracans.

Posterior trunk or abdominal appendages are reduced in most cephalocarids, branchiopods, and maxillopods (notostracans, conchostracans) or absent (cephalocarids, anostracans, cladocerans, ostracods, mystacocarids, copepods, branchiurans, tantulocarids, and cirripeds). Like the remipedes, however, most of the malacostracans bear pleopods on the abdomen that often function in swimming (and sometimes facilitate mating). These may be present or absent in different fossil phyllocarids and are present, but reduced toward the posterior, in the leptostracans. Hoplocarids have five pairs of well-developed pleopods that support tufty gills (palaeostomatopods, stomatopods). Some peracarid eumalacostracans (isopods) also bear breathing structures (gills or tracheae) on their pleopods. The eumalacostracans have pleopods, but their number and size often is reduced (particularly toward the posterior) or may be absent (anaspidaceans, some mysidans, some thermosbaenaceans, some cumaceans, and some anomurans). The sixth abdominal segment is fused to the telson in modern (but not ancient) tanaidaceans, as well as some cumacean and isopod eumalacostracans. The hoplocaridans and most groups of eumalacostracans have a well-developed "tailfan" that is composed of uropods (bladelike appendages on the sixth abdominal segment) and a telson (a flattened and often armored posterior segment that may have one or more terminal projections but never bears caudal rami).

The position of the gonadal duct openings (gonopores) on either the ventral part of the body or the bases of the legs varies consistently among groups of crustaceans, reflecting the ancient divergence of different taxa. The female gonopore opens on the sixth thoracic segment, and the male gonopore on the eighth thoracic segment, in all malacostracans.

Both simple (ocelli) and compound (faceted) eyes are present in most taxa. One medial ocellus frequently occurs in larvae, and one or more ocelli are found in adults of some

groups of branchiopods and maxillopods such as the anostracans, notostracans, cladocerans, ostracods, mystacocarids, copepods, and branchiurans. Only ocelli (or ocellus-like structures) are present in the adults of the extinct kazacharthrans, some cladocerans, mystacocarids, and copepods. Compound eyes often are positioned on stalks, as in the anostracan branchiopods, most malacostracans, including many phyllocarids (extinct canadaspidans, leptostracans), the hoplocarids (extinct aeschronectids, stomatopods), extinct belotelsonideans, and most eumalacostracans (extinct palaeocaridaceans, some anaspidaceans, mysidans, lophogastridans, thermosbaenaceans (nonfunctional), some tanaidaceans (on lobes), spelaeogriphaceans (on lobes), cumaceans (on a medial lobe), mictaceans (nonfunctional on lobes), some isopods (on lobes), euphausiaceans, amphionidaceans, dendrobranchiate shrimps, procaridid shrimps (reduced), carideanshrimps, stenopodidean shrimps, thalassinidean shrimps, astacideans, palinurans, anomurans, and brachyurans). Eyes are reduced or absent in members of primitive lineages, sessile lineages, or species that inhabit caves or the deep sea (the remipedes, cephalocarids, many ostracods, tantulocarids, adult cirripeds, tanaidaceans, some spelaeogriphaceans, many cumaceans (especially females), some mictaceans, some isopods, some amphipods, amphionidaceans, procaridid shrimps, some caridean shrimps, some stenopodid shrimps, and some astacideans). The branchiuran maxillopodans have a particularly intriguing pair of compound eyes that project internally into a blood sinus suspended on a stalk composed of the optic nerves. Muscles move the eyes within the blood sinus.

Either a cephalic shield or a carapace may cover the head and part or all of the thorax. Remipedes, cephalocarids, copepods, and tantulocarids bear a cephalic shield, while a "folded" or "bivalved" carapace characterizes the notostracans, extinct kazacharthrans, cladocerans, conchostracans, ostracods, branchiurans, cirripeds, phyllocarids, hoplocaridans, and most eumalacostracans. However, these structures are reduced or absent in other groups (anostracans, mystacocaridans, syncarid and most peracarid eumalacostracans). The carapace of cirripeds is modified into a fleshy, saclike mantle that may envelope most of the body and, in thoracicans, is covered with calcareous plates; eggs are brooded in the mantle cavity for various lengths of time. Conchostracans, most cladocerans, and some ostracods use their carapace as a brood chamber; one cladoceran broods eggs in its molted carapace. Thermosbaenacean females possess a very enlarged, thin, blood-filled carapace that functions both as a brood chamber and for gaseous exchange. The enlarged, very thin carapace in female amphionidaceans may serve a similar function, though this has not yet been examined. The carapace extends laterally in decapods to enclose the base of the legs and gills, increasing the efficiency of water flow over the gills.

The Rise of Crustacean Diversity in the Fossil Record

All of the early continents, including Gondwana (modern South America, Africa, Madagascar, India, Australia, Antarctica), Baltica (Scandinavia and northwestern Europe), Laurussia (North America), Siberia, Kazakhstania, and China, girdled the earth in tropical and subtropical locations. No land extended above 60° N or below 60° S during the Cambrian (570–500 mya), allowing mild warm conditions in most coastal areas. Developing from segmented wormlike forms, the earliest crustaceans must have developed sets of pronglike or leaflike legs to shuffle along the surface of the soft marine sediments. In addition, they secreted a tough, often calcified cuticle to protect them from the other benthic predators (nautiloid molluscs and trilobites, which were the most abundant benthic predators and scavengers in most marine habitats during the Cambrian). The earliest fossilized crustaceans included three orders of phyllocarid malacostracans (phyllopodous detritus feeders, scavengers, and generalized predators known from the Lower Cambrian); the thoracican goose barnacles (suspension-feeding maxillopods known from the Middle Cambrian); and the ostracods (scavenging, detritus-feeding maxillopods known from the Cambrian onward). Following the recent disovery of the predaceous Remipedia, which exhibit a very primitive arrangment of segmentation and appendages, many workers concluded that the remipedes, though not readily fossilized, represent one of the probable forms of the earliest crustaceans and likely were present in the Cambrian. In addition, the primitive morphology of the ascothoracidan barnacles suggests that they also may have been present (though unfossilized) in the very early Paleozoic. Of these early groups, the canadaspid phyllocarid malacostracans did not persist into the Ordovoician.

During the Ordovician (500–440 mya), part of the huge island continent of Gondwana slipped beneath what was then the south pole, but the climate remained generally warm and mild, with the northern reaches of Gondwana and most of the island continents situated in tropical or subtropical waters. Xiphosuran (horseshoe crabs) and eurypterid (sea scorpions) chelicerates joined the exceedingly abundant, ravenous trilobites on the seafloor. Nautiloids continued to diversify morphologically; jawless fish (encased in protective outer layers of bone and still not very motile) proliferated. The ostracod crustaceans underwent their greatest radiation, and phyllocarid malacostracans (hymenostracans, archaeostracans) were abundant; the hymenostracans, however, disappeared after the Ordovician. A mass extinction closed the Ordovician.

The Silurian (440–405 mya) also was characterized by mild climates (except in the parts of Gondwana that rested over the pole), and a circumtropical current wound its way among the island continents and reefs composed of stromatoporoid sponges, rugose corals, and other sessile fauna. Eurypterids flourished, trilobites persisted in lesser numbers, myriapods and primitive arachnids (probably marine) arose, and jawed fish made their appearance. Thoracican goose barnacles (attached to a eurypterid) are recorded, ostracods flourished, and archaeostracan phyllocarid malacostracans were abundant.

During the Devonian (405–360 mya), predation from the flourishing eurypterids, increasingly motile jawed fishes, and nautiloid cephalopod mollusks dominated benthic ecology. Trilobites still occupied many benthic marine environments. Pycnogonids were present. In an era of mild climates and inland seas, fish and invertebrate groups radiated into freshwater environments. A circumtropical current coursed between Gondwana and the island continents. It was in these

warm, sheltered, shallow coastal waters laden with archipelagos that crustaceans began their great radiations. The fossil record shows that the archaeostracans (phyllocarid malacostracans) reached their greatest diversity, and the predaceous hoplocarids (palaeostomatopods, archaeostomatopods) made their first appearance. The ostracods began their second great radiation. The boring acrothoracican barnacles and the first predaceous, lobster-like decapod malacostracans appeared (*Palaepalaemon*, allied either to the astacid crayfish and lobsters or palinurid lobsters). Detritus-feeding and scavenging conchostracan branchiopods originated, invading fresh water, and the detritus-feeding lipostracans (allied to the cephalocarids or branchiopods) flourished in hot springs. Fossils and primitive morphology suggest that other primitive detritus-feeding branchiopods (notostracans, anostracans) likely were present too. Another mass extinction impacted faunas during the Late Devonian.

The taxonomic, morphological, and ecological diversification of the Crustacea accelerated during the Carboniferous (Mississippian, Pennsylvanian; 360–290 mya) of the Late Paleozoic as the marine chelicerates and trilobites began their decline. The continents were beginning to congeal, setting the stage for the global climatic changes that likely contributed to the greatest mass extinction of all time at the end of the Paleozoic. Gondwana had largely completed its migration beneath the south pole and emerged to begin to fuse with Laurussia and Baltica (modern North America, Scandinavia, and northwestern Europe) over the equator. Though the southern reaches of Gondwana (modern South America) were still plagued with bouts of glaciation, the warm moist equatorial continents were awash with inland seas and swamps. Land plants had multiplied in diversity and biomass, their seed ferns, scale trees, and conifer trees producing great forests that decayed in coal swamps or washed out onto the continental shelves, enriching coastal sediments. Insects took flight (including some huge species), amphibians underwent their greatest radiation, and the first reptiles appeared. Marine organisms proliferated in the enriched coastal sediments. Sharks and jawed boney fishes were abundant and active. Nautiloids declined, replaced by the emerging ammonoid cephalopod mollusks. Remipede crustaceans are recorded from the Lower Pennsylvanian. Conchostracan branchiopods radiated anew. Notostracan branchiopods probably were present. Acrothoracican barnacles are known from their borings in clam shells, and lepadomorph thoracican barnacles are recorded from this period. Early malacostracan lineages (the scavenging archaeostracan and hoplostracan phyllocarids; the scavenging aeschronectids and predaceous hoplocarids; the scavenging belotelsonidans, waterstonellideans, eocaridaceans, and paleocaridaceans) were exceedingly abundant and dominated the shallow marine fauna. Pygocephalomorphan and lophogastridan mysidaceans came on the scene, as well as most of the peracarid malacostracans (tanaidaceans, spelaeogriphaceans, cumaceans, isopods, and probably other unfossilized groups). Definitive palinuran decapods appeared, and the earliest lineages leading to the brachyuran crabs evolved. Several of the early malacostracan groups (hoplostracan phyllocarids, aeschronectid and palaeostomatopod hoplocarids, belotelsonids) disappeared after the Carboniferous as the peracarid and eucarid malacostracans took center stage.

The Permian period (290–245 mya) brought one of the most significant geomorphological events in the history of the earth. All of the previous continental landmasses fused into a single gigantic continent, Pangaea, between 260 and 250 mya. The collision of Gondwana, Laurussia, Baltica, Kazakhstania, and Siberia is still visible in the mountain belts stretching across northern South America, northwestern Africa, eastern North America (the Appalachians), Britain and Scandinavia, and Eurasia (the Urals and other ranges). By 250 mya, the ancient island continent of China also had joined the supercontinent. Pangaea now straddled the equator, anchored by Siberia near the north pole and Antarctica over the south pole. The circumtropical seaway was closed. The elevated landmasses produced a more extreme, cooler, continental climate than had been present throughout most of the Paleozoic (whether or not this was solely responsible for the mass extinctions is debated). In the cool, dry conditions, reptiles radiated, replacing the previously dominant amphibians, and the first mammal-like reptiles appeared. Many organisms, including worms, heart urchins, sea cucumbers, and crustaceans, mined the organic riches in the coastal sediments, throwing up sediment that buried and clogged the feeding structures of the epibenthic sessile filter feeders. Conchostracan branchiopods, ostracod maxillopods, primitive malacostracans (archaeostracan phyllocarids, the first leptostracan phyllocarids, waterstonellideans, eocaridaceans, palaeocaridacean syncarids, and pygocephalomorph mysidaceans), at least several peracarid malacostracans (tanaids, cumaceans, isopods), and the first definitive astacid (crayfish- or lobster-like) eucarid malacostracans thrived in benthic marine environments. Brachyurans likely were present. The Paleozoic era ended, however, with the loss of 96% of all species. Many crustacean groups did persist through the Permo-Triassic biotic catastrophe, though, and underwent reradiation during the Mesozoic.

The early Triassic was a relatively barren environment both on the land and in the sea. The landmasses were high, cool, and dry, with widespread deserts. In coastal waters, no corals or reef communities are known. All of the trilobites, eurypterids, most of the nautiloid cephalopods, and many of the primitive fishes had been lost. All of the primitive phyllocarid malacostracans, except for the newly evolved leptostracans, were gone. All of the early hoplocarids had been eliminated except for a thin line, probably present but not fossilized, that would lead to the modern stomatopods. All of the late Paleozoic primitive malacostracans (the belotelsonids, waterstonellideans, eocaridaceans, the early palaeocaridacean syncarids, the pygocephalomorphan mysidaceans) were now absent. The sessile epibenthic suspension feeding community (many of the bryozoans and brachiopods, many primitive echinoderms such as blastoids and most of the crinoids, primitive tabulate, and rugose corals) that had dominated Paleozoic benthic habitats had been largely obliterated. Most of the plankton was gone. As the Triassic progressed, however, mammal-like reptiles and the early groups of dinosaurs appeared, other reptiles diversified, and coniferous trees came to dominate the land. In the sea, the surviving remnants of the cephalopods began to diversify into modern groups, including predaceous squid- and octopus-like forms. Scleractinian corals began to proliferate, even though they were subordinate to

sponges, algae, and bryozoans as reef builders. Conchostracan branchiopods, ostracods and lepadomorph cirripeds, anaspidacean syncarids, lophogastrid mysidaceans, tanaid and isopod peracarids, the first dendrobranchiate decapod shrimps, and astacid and palinurid decapods are recorded. Probably many more lineages were reestablishing themselves, though unfossilized, during this period (remipedes, cephalocarids, all of the branchiopod groups, many of the primitive maxillopods, other primitive eumalacostracans such as syncarids, and the mysidaceans, peracarids, and eucarids). Another mass extinction, however, spread biological devastation in the Late Triassic.

It is thought that huge continental landmasses begin to accumulate heat beneath them, creating convection currents in the mantle that cause landmasses to drift apart approximately every 500 million years. In the Jurassic (210–140 mya), Pangaea began to break apart. About 200 mya, Laurasia (North America and Eurasia) moved away from Gondwanaland, forming a nascent Atlantic Ocean in the west and a gulf between Asia and East Africa in the east. India drifted away from Antarctica toward Asia, and the south Indian Ocean began to form between Australia (still attached to Antarctica), Antarctica, and the joined southern tips of Africa and South America. A widening rift between South America and Africa was flooded with seawater near the end of the Jurassic. The northern drift of Laurasia created a pathway for a circumtropical current between the northern and southern continents. This shallow warm seaway, adorned with archipelagos, would be called the Tethys Sea, and it was the cradle from which most of the modern families of organisms that we know today would emerge.

The early Jurassic climate was cool, becoming warmer. High sea levels encroached on the continents, spreading shallow seas. Although mass extinctions challenged the faunas during the Middle and Late Jurassic, the Jurassic period is known for its spectacular radiation of reptiles, including the dinosaurs and pterosaurs. Modern groups of amphibians (frogs, salamanders) and birds arose. Primitive flowering plants and many of the modern groups of insects (flies, moths, bees, wasps, ants) diversified in the increasingly warm, mild climate. In the sea, marine reptiles must have been fearsome predators, crocodiles and turtles abounded, many of the modern families of boney fish were formed, corals acquired symbiotic algae that allowed them to grow up and away from sedimentation (such as that created by burrowing animals in the late Paleozoic) 10 times faster than they could in the absence of the photosynthetic symbiont. Reefs became more extensive and boring organisms more common. The Jurassic fossil record of Crustacea includes conchostracan and notostracan (extinct Kazacharthrans) branchiopods, and shows a renewed ostracod radiation. The lepadomorph barnacles were present, and the balanomorph (acorn) barnacles made their first appearance. Many of the other groups (remipedes, cephalocarids, other branchiopods, other maxillopods) likely were present but not fossilized. In the Malacostraca, the fossil sculdid hoplocarids (descendents of the Devonian archaeostomatopods and progenitors of the modern stomatopods) arose. The last remaining vestige of the phyllocarids (the leptostracans) and some syncarids may have been present but not fossilized. The first mysidacean eumalacostracans appeared, tanaid and isopod peracarids were present, and many of the other peracarids likely were present but not preserved. The eucarid malacostracans burst on the scene, as shown by the first appearances of fossil caridean, stenopodidean, and thalasinidean shrimps, and anomuran and definitive brachyuran crabs. Dendrobranchiate shrimps, as well as astacidean and palinuran crayfishes and lobsters, also were present in the fossil record, completing the roster of all the major groups of eucarids known today. Despite the mass extinctions, no major lineages were lost during the transition from the Jurassic to the Cretaceous.

The tropical Tethyan seaway continued to girdle the earth and maintain mild conditions during the Cretaceous (140–65 mya), although the pathway was narrowing in several sites (e.g., northeastern and northwestern Africa). North America was still united with Eurasia despite a widening gulf invading northward from the opening Atlantic. India had not yet crashed into Eurasia, Madagascar was sliding northeastward from southern Africa, and Australia had begun to separate from Antarctica, beginning its trek toward the tropics. Africa and South America had left Antarctica behind as they migrated northward and severed their connections to each other. The Andes and Rocky mountains rose as North and South America ploughed westward and the Pacific plates dived beneath them. The continents were awash with spreading inland seas and swamps. The reptilian radiation was at its apex, spawning the largest land and marine reptiles ever known. Grasses appeared, and other flowering plants and insects diversified even further; coniferous gymnosperms declined.

In the sea, reefs began to accrete and proliferate on an unprecedented worldwide scale and reached their highest diversity ever. Scleractinian corals flourished, and in sheltered areas huge rudist bivalve mollusks (which, like the corals, may have obtained photosynthetic symbionts) committed massive deposits of calcium carbonate to reef structure, threatening the dominance of corals in the reef community. The teleostean (boney) fishes underwent major diversification, developing new body shapes that increased their maneuverability, calcareous mouth plates that crushed shelled and hard-bodied prey, and pincer-like jaws that could grab and pluck prey from the substratum. Benthic octopuses with flexible arms and bodies could engulf and subdue prey. Crustaceans developed hardened specialized claws that could crush or cut prey. Motile and sessile organisms seeking refuge from the burrowing organisms that caused extensive movement of benthic sediment (bioturbators) likely colonized the high, hard substratum provided by the corals and rudists. Boring organisms proliferated worldwide, for boring into the calcareous substratum provided protection from the increasingly diverse and dangerous predators. Individual bioeroders eventually died, however, leaving a vacant crypt. This three-dimensional microenvironment provided a safe refuge for a multitude of small sessile and motile invertebrates. The organisms that invaded the holes left by the borers, termed the cryptofauna, were constrained to small body sizes in order to occupy the protective holes, were plagued by predation whenever they emerged from the crevice to feed or mate, and had to develop special adaptations for acquiring and keeping the packed refuge space. All of these processes accelerated morphological and other adaptations in the benthic fauna. Small body size

usually is associated with production of fewer, larger eggs, abbreviated development, and higher rates of speciation and extinction in motile invertebrates. These life history features likely facilitated diversification of other behavioral and morphological features, such as the deadly fighting behavior and armored bodies and appendages that developed in many of the stomatopods and decapods. Special adaptations (including coloniality, unique growth strategies, and chemical warfare) also developed in the sessile cryptofauna, promoting diversification and enabling them to better acquire and defend living space.

Among the crustaceans, the conchostracan branchiopods and ostracod maxillopods again are represented in the Cretaceous fossil record; the remipedes and all other modern groups of branchiopods and maxillopods likely were present but not preserved. Copepods, ascothoracidan barnacles (as trace fossils), verrucomorph (wart) barnacles, and the squat modern cthamaloid acorn barnacles made their first appearance as fossils. Among the lower malacostracans, the leptostracan phyllocarids likely were present but unfossilized, and the predaceous stomatopod hoplocarids came into their own. The extinct sculdids (bathysquilloids) and three of the major modern groups of stomatopods (the reef-dwelling gonodactyloids, mud-burrowing squilloids, and sand-burrowing lysiosquilloids) all appeared for the first time. The syncarids and mysidaceans probably were present but not recorded. Although their distribution suggests a more ancient derivation, the thermosbaenaceans are first known in the Cretaceous. Other peracarids and other eucarid shrimps (stenopodideans, thalassinideans) probably were present, but the tanaidaceans and isopods, as well as the dendrobranchiate and caridean shrimps, definitively inhabited Cretaceous seas. The advanced eucarids (astacids, palinurans, anomurans, brachyurans) were all present and continued their radiations begun in the Jurassic.

Following a mild mid-Cretaceous mass extinction, a significant mass extinction at the end of the Cretaceous eliminated the dinosaurs and many of the great reptiles (including the huge flying and marine reptiles), the ammonoids and many of the older cephalopod lineages, the reef-building rudist bivalves, and many groups of plankton. It is likely that the explosive impact of a giant asteroid contributed to this biotic disaster. The role of additional factors in the faunal turnover between the Mesozoic and the Cenozoic is debated. Mass extinctions often determine the biotic structure of the next era by filtering what major groups survive to radiate again in the postextinction periods. This was true for the mammals and birds, relatively minor groups in the Mesozoic that persisted through the mass extinction and then, in the absence of the great reptiles, diversified and dominated the ecology of terrestrial environments in the Cenozoic. For the crustaceans, however, the biotic upheaval at the end of the Mesozoic had far less impact than had the great extinction at the end of the Paleozoic. The kazacharthran branchiopods, related to the notostracans, had been lost during the Jurassic, but the lineage continued and survives today. The hoplocarids also persevered. Despite the possible loss of the sculdid stomatopods at the end of the Cretaceous, the bathysquilloid lineage persisted (though restricted today to deep continental shelf environments). The new stomatopod superfamilies generated in the Cretaceous endured and are common in warm shallow waters today. No other major crustacean lineages were lost. Their small body sizes (compared to those of some of the giant reptiles and cephalopods, for example) may have allowed crustaceans to persist in a time of great ecosystem disturbance and food limitation. Also, although the extent of the geographic range of an individual species confers resistance to background extinction, a broad distribution for the entire group is one of the most important factors enabling lineages to survive mass extinctions (Jablonski, 1991). Reaping the success of radiations in the warm, reef-studded seas of the Jurassic and Cretaceous, many crustacean groups likely had worldwide distributions that enabled them to survive the end of Mesozoic mass extinction.

The Cenozoic can be divided into the Tertiary (65–1.6 mya, including the Paleocene, Eocene, Oligocene, Miocene, Pliocene) and the Quaternary (1.6 mya–present, including the Pleistocene and Holocene). The Atlantic was now formed and widening. India crashed into Eurasia, creating the Himalayan mountains. Extensive folding, resulting in the Alps, Caucasas, and other mountains, occurred as parts of Africa fused with Eurasia, forming Asia Minor and constricting the ancient Tethys seaway into the isolated, sometimes dry, Mediterranean and Black Sea areas. Mountain building continued in the Rockies, the Coast Range, and the Andes, eventually shifting the flow of the Amazon from westward to eastward. The early Tertiary climate was warm, wet, and mild. The great radiations of placental mammals began. Horses, camels, rhinoceroses, dogs, cats, whales, bats, and primates diversified in Eocene terrestrial environments. A gradual cooling trend began in the Oligocene. Extensive plains, grasslands, and new mammalian lineages (deer, elephants, antelopes, giraffes, and the ancestors of cattle and pigs) developed through the Miocene, and large mammalian carnivores radiated in the Miocene and Pliocene.

In the oceans, there were no extensive coral reefs in the Paleocene, although scleractinian corals had survived the mass extinction at the end of the Mesozoic. The Eocene brought a reradiation of reef-building corals and many families and genera of fish, mollusks, and crustaceans that survive today. Extensive volcanism and tectonic movements in the first half of the Tertiary set the stage for island fringing reefs and the development of atolls in the Pacific.

The Tertiary produced a new radiation of the cirriped barnacles, including the coronuloids (which live as commensals on whales and turtles and appeared in the Paleocene) and the first modern intertidal balanoids (which originated in the Eocene and diversified again in the Miocene). The ostracods radiated once again during the Eocene; they also are represented in the Paleocene and Miocene fossil records. Among the branchiopods, the first definitive anostracans and cladocerans appeared during the Eocene. In the stomatopod hoplocarids, a possible sculdid bathysquilloid, the squilloids, and the first pseudosquilloids are known from the Eocene; in addition, fossil gonodactyloids and squilloids are recorded from the Miocene, and squilloids from the Pliocene. Following a late Paleozoic genesis, the isopod peracarid eumalacostracans radiated during the Eocene, and the first amphipod peracarids appeared; both of these groups are known also from the Oligocene and Miocene. Among the eucarid malacostracans,

the caridean shrimps left fossils in the Oligocene and Miocene, and the thalasinidean shrimps are recorded from the Miocene; both groups hailed from the Jurassic. The astacid crayfish and lobsters, among the first of the Paleozoic decapods, are known from the Eocene. The brachyuran crabs, also from a likely Paleozoic heritage, underwent major radiations in both the Eocene and Pliocene, with many of the modern freshwater and terrestrial lineages developing in the latter epoch. Brachyurans are represented in the Paleocene and Miocene fossil record as well. Anomurans are known from the Oligocene and Miocene.

One of the most significant events in the late Tertiary was the gradual establishment of a complete barrier to circumtropical ocean circulation about 3 mya as mountain building along the west coasts of North and South America created a Central American isthmus that linked the two continents. This removed the ameliorating effect of the circumtropical current on the climate and created a much smaller, basin-like West Atlantic region as well as a long thin East Pacific region. Considerable extinction and divergence of new genera and species ensued in the New World tropical seas.

In spite of a minor mass extinction at the end of the Eocene and deteriorating ocean circulation in the Miocene and Pliocene, however, no major groups of crustaceans were lost during the Tertiary. The ostracod and cirriped maxillopods, the stomatopods, the isopod and amphipod perarcarids, and the brachyuran decapods all underwent major radiations, particularly during the Eocene.

The Quaternary (Pleistocene 1.6–0.01 mya, Holocene 10,000 years ago-present) ushered in four major periods of glaciation, particularly on the northern continents (North America, Europe, North Asia) that had not been extensively glaciated during the Paleozoic and Mesozoic. Early humans appeared in the Pleistocene. Mountain building in the Cascades, Andes, Caucasus, and the Himalayan regions continued, and rifts in eastern Africa, Asia Minor, Baikal, and south-east Asia reflected continuing tectonic shifts. Worldwide sea levels receded with the advance of continental glaciers and advanced during warmer interglacial intervals. The current climate represents emergence from a colder glaciated interval about 10,000 years ago. The extent to which the frigid polar glaciation impacted climate and seawater temperatures at lower latitudes is debated. Most modern coral reefs (including Australia's Great Barrier Reef) are only about 10,000 to 15,000 years old and represent a thin surface veneer on a complex topography inherited from past reef communities as they adjusted to changing sea levels. The anostracan branchiopods, isopod and amphipod peracarids, and brachyuran crabs are among the relatively few crustacean groups that were preserved as fossils during the Pleistocene. Although morphological and distributional evidence suggests that some of the smaller and thinly skeletonized lineages were present much earlier, many groups are recorded for the first time during the Holocene. These include the nectiopodan remipedes, cephalocarids, mystacocarids, branchiurans, tantulocarids, ascothoracidan and rhizocephalan cirripeds, bathynellacean syncarids, thermosbaenaceans, mictaceans, euphausiaceans, and amphionidaceans. Many other groups have radiated or renewed their diversification during the Holocene. Showing remarkable evolutionary resilience since the beginning of the Paleozoic, the ostracods are undergoing a fifth major radiation. The cirripeds show considerable current diversity. The stomatopods, peracarids (especially isopods and amphipods), and decapods (dendrobranchiate shrimps, and especially the caridean shrimps, thalassinidean shrimps, anomurans, and brachyurans) all are undergoing major modern diversification. Although no major lineages have been lost during the Quaternary, today the remipedes, cephalocarids, branchiopods, remaining lower malacostracan lineages (phyllocarids, syncarids) and some of the ancient eumalacostracan lineages such as the palinurans and astacids often inhabit peripheral refugia and appear to have their greatest eras behind them.

Ecological Diversity in Crustaceans

The Habitats That Crustaceans Occupy

The major types of environments inhabited by crustaceans can be separated into 16 major, often overlapping, categories. Of all of these, crustaceans are most commonly found in benthic marine habitats, including both level bottom and coarse (rock and reef) habitats (35 groups). Brackish water (coastal deltas and lagoons, brackish marine caves and tidepools, salty inland pools, and brackish groundwater) is the next most common habitat for which 23 groups (59%) of crustaceans are well suited. Seventeen taxa (44%) are pelagic in the upper part of the ocean (<200 m), and 18 (46%) groups are found in the deep sea (200 to >6000 m), either as pelagic or as benthic species on level bottoms or around hydrothermal vents. Fifteen taxa (38%) inhabit permanent bodies of fresh water (ponds, lakes, streams, rivers), including both benthic and pelagic components. Eight groups (21%) contain species that live in intertidal environments (some of these among sand grains), and 8 (21%) inhabit other semiterrestrial or terrestrial conditions (usually moist conditions or decaying vegetation). Some isopods, some amphipods, and a few brachyurans are truly terrestrial, having achieved independence from water for reproduction. Eight taxa (including 5 highly specialized groups) have adopted parasitism (21%), and 7 groups (18%) include commensals that live on or in other organisms. Seven groups or sections of taxa (18%) flourish at high latitudes, 5 (13%) inhabit either marine or freshwater caves (with various degrees of specialization), and 5 (13%) earn a fragile existence in inland ephemeral freshwater and briney pools. Three taxa (8%) can tolerate hypersaline (very salty) conditions, 3 (8%) are known from hot springs, and 3 (8%) live a specialized existence in groundwater. At least 1 group lives interstitially among the meiofauna (organisms smaller than 0.5mm) in sub tidal benthic marine sediments. The number of major taxa that occupy each of the various types of habitat may but does not necessarily reflect the number of species in that environment. This analysis, instead, demonstrates the diversity of major adaptations (in morphology, physiology, and behavior) that crustaceans have developed to accommodate particular environmental challenges. These adaptations provide the underlying framework, overlain by other specializations, on which evolution can proceed in the long run.

The same environmental categories can be used to examine ecological diversity among the different crustacean groups. It is likely that those groups that occupy the greatest diversity of

habitats are the ones that will be able to survive massive environmental change, since not all habitats are likely to be equally affected by most forms of environmental disturbance. As shown in the discussion of taxonomic diversity, the maxillopods and the eumalacostracans (particularly the peracarids and eucarids) are the champions in terms of ability to occupy a diversity of different environments. Following a history of evolutionary success that stretches from the earliest Paleozoic to the present, the ostracods inhabit seven different types of environments. The comparatively recent (late Mesozoic) but ubiquitous copepods also occupy seven major types of habitats, including exceedingly specialized parasitic ways of life, and the ancient thoracican barnacles live in five major habitat types. Among the lower malacostracans, the anaspidaceans are surprising in the diversity of seemingly marginal types of environments that they inhabit (five: permanent fresh water, ephemeral freshwater pools, caves, semiterrestrial, groundwater), and the mysidans have radiated into seven major habitats and lifestyles. Even more remarkable, the thermosbaenacean peracarids, with only 11 species, inhabit five different types of environment as categorized earliear (benthic marine, inland salt ponds, fresh and salty groundwater, caves, hot springs). The tanaid peracarids and the anomuran eucarids occupy five and six major types of habitats. The caridean shrimps and brachyuran crabs (decapod eucarids) and the isopod and amphipod peracarids, however, inhabit more different types of environment than any other groups of crustaceans (seven, eight, nine, and nine types of habitat, respectively). Although many carideans and some brachyurans live commensally with other organisms, a fully parasitic existence is the only habitat that these eucarids have not exploited as fully as have some isopods and amphipods. Although caridean shrimps have been at least as successful in fresh water and are more diverse in pelagic environments than the brachyurans, they have failed to invade terrestrial habitats as completely as have the brachyurans, amphipods, and isopods. The isopods are the most successful terrestrial crustaceans, and they, along with most other peracarids, flourish in cold, high-latitude environments as well as the deep sea. Amphipods are less specialized for parasitism and terrestrial existence than the isopods. Like the amphipods and isopods, the carideans and brachyurans have radiated into all major marine environments, including the deep sea and its hydrothermal vents. The carideans likely invaded fresh water very early (Mesozoic), whereas the brachyurans undertook major invasions of freshwater and terrestrial habitats in the Late Tertiary and Quaternary.

A number of crustaceans also modify their microenvironment by commandeering empty shells or other structures, engineering tubes or burrows, and camouflaging themselves from pervasive predation in various ways. Anomurans, tanaids, and amphipods all inhabit vacated gastropod shells; anomuran hermit crabs also inhabit tusk shells, hollow twigs and bamboo, calcareous worm and vermetid mollusk tubes, hollow pieces of coral, and even bottle caps. Some anomurans plant anemones that function in camouflage and defense on their gastropod home. Some dromiid brachyuran crabs carry a clam shell on their backs, and lithodid and majid brachyurans culture domes of sponge or decorate their carapace with algae. Tanaid and amphipod peracarids construct pebble tubes, and amphipods modify algal structures for domiciles or camouflage. One group of amphipods bores into wood, one group of isopods bores burrows in silt and clay, and acrothoracican barnacles bore into coral to make protective lodgings. Coral-dwelling stomatopods chip, scrape, and modify preexisting holes in coral or reef substratum to suit their needs, and most groups of crustaceans are well represented in preformed crevices and bioeroded holes in hard, especially calcareous, substrata (coral, limestone and reef basement, oyster and tube worm reefs). Many of these organisms, such as the thalassinideans, line and modify the preexisting crevice with silt and mucous for a better fit. Some groups of caridean snapping shrimps construct smoothly lined networks of burrows within sponges. Large species, such as stone crabs, coral crabs, spider crabs, spiney lobsters, and some stomatopods inhabit refuge space beneath bioeroded ledges and caverns on reefs. Several groups of stomatopods, astacid crayfishes and lobsters, thalassinidean shrimps, and brachyuran land crabs, fiddler crabs, and ghost crabs all excavate large (in some species very complex) burrows in mud and sand environments. All of these habitat modifications reflect the pervasive intensity of predation by epibenthic and benthic predators such as cepahlopods (particularly octopuses), other crustaceans, and especially fishes. By increasing the complexity of the microenvironment, many of these structures increase the potential for commensal associations and heighten the capacity of the environment to support high species diversity.

The Role of Crustaceans in Biological Communities

Examination of the trophic position of crustacean groups indicates that most crustaceans have relatively generalized feeding habits and most are subject to strong predation. Their abundance, moderate size, and generalized carnivory likely exerts top-down control in some communities. Also, as evidenced by almost universal morphological and behavioral adapatations against predation as well as their importance in the diets of many different predators, the availability of crustaceans as prey for other trophic levels probably is essential for the structure, function, and diversity of many aquatic communities. Cladocerans and copepods support higher trophic levels, including fishes, in freshwater habitats. When their ephemeral habitats are located, anostracans and notostracans are heavily consumed by birds. Anaspidacean syncarids are heavily preyed upon by freshwater fishes if the shrimp are dislodged from their refuges. Stomatopods, mysidans, tanaids, cumaceans, isopods, amphipods, dendrobranchiate shrimps, caridean shrimps, anomurans, and brachyurans all are common in both epibenthic and benthic fish stomachs; cumaceans, isopods, and amphipods are particularly important in the deep sea. Dendrobranchiate shrimps and mud-dwelling stomatopods are important in the diets of many large, commercially important fish, and euphausiaceans and pelagic anomurans are a major source of food for whales, squid, tunas, and seabirds.

Most crustaceans are either generalized predators and scavengers (30 taxa or portions of taxa), or detritus and suspension feeders (35 taxa or parts of groups), Many biologists consider the predaceous, scavenging remipedes to exemplify the primitive feeding mode, but detrital suspension feeding, as seen in the phyllopodous cephalocarids, branchiopods, and

some of the primitive phyllocarid malacostracans, obviously can be traced far back in crustacean history as well. Herbivory is not particularly well developed in crustaceans, although many of the phyllopodous groups (cephalocarids, branchiopods, leptostracans) may ingest phytoplankton and plant material along with other detritus. The suspension feeding ascothoracidan, acrothoracican, and thoracican barnacles consume phytoplankton; and some ostracods, isopods, amphipods, and brachyurans feed on micro-or macroscopic plants (some of these species also eat detritus or other organisms as well). Euphausiaceans, however, are important grazers of phytoplankton at high latitudes, and copepods may consume more primary productivity than any other organism on earth because of their ubiquity and consumption of phytoplankton in the largest habitat on earth, the open ocean. The mouthparts of crustaceans are relatively easily modified for piercing and tearing flesh, so strict carnivory and ectoparasitic habits, with the potential for considerable impact on host populations, have evolved in a number of groups, including a cladoceran, branchiurans, tantulocaridans, some ascothoracidan barnacles, isopods, and amphipods. Parasitism is most extremely developed in the endoparasitic copepods and rhizocephalan barnacles; many ascothoracidan barnacles and some isopods also live inside and consume the energy or tissues of their hosts. Large body size and armored claws that smash, crush, or shear hard-bodied prey make stomatopods, palinuran and astacid lobsters and crayfish, and anomuran and brachyuran crabs potent predators that, near the top of their benthic food webs, exert important effect on their local communities.

Life History Patterns of Crustaceans

Although most crustacean groups reproduce sexually, a number of groups, particularly in fresh water and high latitude environments, produce offspring without exchanging eggs and sperm. "Parthenogenesis," a form of asexual reproduction where eggs undergo development without fertilization, is found almost exclusively in the Branchiopoda and Maxillopoda. Other forms of asexual reproduction, such as budding of new individuals from a previous individual, do not appear to be present in the Crustacea, although regeneration of missing body parts is common. In the branchiopods, most anostracans reproduce sexually (and produce fertilized resistant eggs that are dormant during unfavorable conditions), but some populations are exclusively parthenogenetic. Parthenogenesis is common in both notostracans and cladocerans. Some notostracan populations are exclusively parthenogenetic, especially at high latitudes. Notostracans also reproduce sexually, producing fertilized resistant eggs that can withstand harsh conditions. Some species of cladocerans produce males and undergo sexual reproduction only during periods of environmental deterioration, such as autumn; males may be dwarfed and short lived. Some species of conchostracans are exclusively parthenogenetic, and other species may include both parthenogenetic and sexual individuals. In the maxillopods, some freshwater ostracods are parthenogenetic, and other species are both parthenogenetic and sexual. Some harpacticoid copepods are parthenogenetic. Some copepods produce resting eggs that are dormant for extended periods. Of the rare deep-sea tantulocarids, only females are known, though parthenogenesis is not proven. Evidence suggests parthenogenesis in some thoracican barnacles. Some terrestrial isopods are parthenogenetic.

Reduction of the prominence of males, as evidenced by short-lived, usually dwarfed males that frequently live on or in the female, is seen in the freshwater cladocerans, parasitic ascothoracidan barnacles (dwarf male free living or lives on female), boring acrothoracican barnacles (lives on female), sessile thoracican barnacles (lives on hermaphroditic individual or female), parasitic rhizocephalan barnacles (short-lived dwarf male invades extruded female gonad), and parasitic isopods (lives on female). In the deep sea gnathiid isopods, the second maxillipeds are reduced to flaps that cover the mouth of the nonfeeding male. Few males are known (although sexual dimorphism appears in development) in the pelagic deep sea amphionidaceans, and populations of the primitive procaridid shrimps are composed mostly of females; these skewed sex ratios may be the result of differential mortality in males, reduction in the prominence of the male life cycle, or flexible sex determination (as is common in many crustacean groups) rather than a propensity for parthenogenesis.

The reproductive organs of crustaceans are relatively simple, and both testes and ovaries apparently develop readily, with dominance shifted from one to the other depending on hormonal and environmental conditions. The primitive remipedes and cephalocarids are hermaphroditic. Although most sexually reproducing branchiopods are gonochoristic, some species of notostracans are hermaphroditic and likely undergo sex reversal. In the maxillopods, the sex of the sessile cirripeds is notably flexible. Although most of the parasitic ascothoracidan barnacles are gonochoristic, species in one group are simultaneous hermaphrodites and those in another group are protandrous hermaphrodites. In others, the first arrival to a host becomes female, and the next arrival becomes a male. Most of the sessile thoracican barnacles are simultaneously hermaphroditic or mixed hermaphroditic and separate sexes (dwarf males may reside either on females or hermaphrodites). Most acrothoracican and rhizocephalan barnacles, however, are gonochoristic. The other center of sex change is found in the malacostracans. In the peracarids, tanaidaceans usually are protogynous hermaphrodites, or the same population may have both gonochoristic and hermaphroditic individuals. Sex change appears to be influenced by environmental factors, possibly even the ratio of males to females as affected by predation. Most isopods are gonochoristic, but some (especially freshwater) lineages change sex. Both protandrous and protogynous species are known, and both gonochoristic and hermaphroditic individuals may occur in the same population. Particularly in parasitic isopods, sex is determined by the presence or absence of another individual; first arrivals become females and the second arrival becomes a dwarf male that resides on the resident female. Most amphipods also are gonochoristic, but protandrous species are known, and some species are composed of both protandrous and gonochoristic individuals. Although most caridean shrimps have permanent separate sexes, some species are simultaneous hermaphrodites and some are protandrous. Both gonochoristic and hermaphroditic individuals may occur within some populations. Sexual differentiation may be affected by both social and environmental factors. For example,

some species of snapping shrimp live in colonies of up to 350 individuals in a sponge, with only one large reproductive female, the "queen." A host of smaller male or nonreproductive "workers" defend the colony and maintain communal burrows.

Sexual reproduction often is accompanied by morphological structures that have differentiated between males and females, probably as the result of sexual selection. Males often are adapted for greater motility than females. In the ascothoracidan barnacles, the female's body is enveloped entirely within the carapace, while most of the swimming appendages are exposed in males. Male tanaids have well-developed pleopods, but abdominal appendages are absent in females; males also have larger eyes and more antennal sensory structures than females. Male cumaceans bear abdominal pleopods while females do not, the antennae are reduced in females but not males, and the eyes are more reduced in females than males. The carapaces of female thermosbaenaceans and amphionidaceans are very enlarged, while those of the males are reduced, exposing swimming appendages. The last pair of thoracic legs are absent in female amphionidaceans but not in males, and the females, with reduced mouthparts and a vestigial gut, do not feed.

Various appendages have been modified in many groups to facilitate mating. The first pair of maxillae (maxillules) were modified as a clasper in males of the Devonian Lipostracans (allied to the cephalocarids). The antennae are modified for detecting or holding females in male anostracans; some cladocerans, ostracods, and cumaceans; and in many amphipods. The size and shape of the antennules of male leptostracans and euphausiaceans and of some species of copepods are modified, probably for clasping females. Other secondary sexual characteristics include copulatory structures, which are common in crustaceans. Anostracans have a pair of eversible penes on the first two abdominal segments. The leaflike appendages of the first two trunk segments are modified for copulation in conchostracans; the first two trunk limbs of some genera of ostracods are modified to assist mating, and the third trunk appendages often are modified into copulatory organs. The fourth trunk appendages, sixth thoracopods, and second through fourth thoracopods of male mystacocarids, copepods, and branchiurans, respectively, are modified for copulation. Abdominal appendages have been fused into a single long median penis, the only remaining vestige of the abdomen, in thoracican barnacles. Paired penes also are fused into a single median penis in the tantulocarids and some isopods. A pair of copulatory tubes are associated with the eighth thoracopods in male stomatopods, and the eighth thoracopods are modified, probably for copulation, in the bathynellaceans. The bases of the eighth thoracopods in mate thermosbaenaceans bear a pair of well-developed penes. Male mictaceans sport a copulatory process on the eighth abdominal segment, and, although pleopods are otherwise absent in both males and females, male mictaceans retain the second pair of pleopods, probably to assist transfer of sperm. The anterior abdominal pleopods are structurally modified to assist copulation in many male malacostracans, including stomatopods, mysidans, isopods, amphipods, and dendrobranchiate shrimps. The anterior pleopods of both male and female euphausiaceans are elaborately modified for copulation.

Several groups of crustaceans exhibit strong sexual dimorphism in body size and other structures, including the late Paleozoic pygocephalomorph mysidaceans. Sexual dimorphism can be observed among developing juvenile amphionidaceans after stage 9. Male tanaids show pronounced sexual dimorphism in a variety of sensory structures as well as relative size of the chelipeds. The overall morphology of male and female tanaids becomes so highly modified as they molt into copulatory versus noncopulatory phases that they frequently have been misclassified as different taxa. Amphipods and caridean shrimps show similar variability of body size and shape between copulatory and noncopulatory stages, probably reflecting the impact of strong selection for attracting and defending mates against other suitors. Some groups of isopods (especially terrestrial lineages) show extreme dimorphism in size, with males greatly exceeding the size of females. The claws of many decapods (including some caridean shrimps, some anomurans, and particularly some thalassinidean shrimps, astacid crayfish and lobsters, and many groups of brachyuran crabs) are greatly enlarged in males. Either both or only one of the claws may be enlarged for fighting or behavioral displays. Mating or territorial displays often are associated with enlargement of only one claw, as in the fiddler crabs that attempt to attract females to their burrows with colorful choreographic waving displays. Though not sexually dimorphic, stomatopods also seek to attract mates by exposing species-specific colored eyespots and irridescent patterns on their spread maxillipods. Several groups, including land crabs and fiddler crabs, use sound vibrations (usually stridulating or thumping the walls of their burrow) to attract mates. Various other crustaceans stridulate (stomatopods, some isopods, palinurans, other brachyurans), but these sounds may function to deter attacking predators rather than to attract mates. Bioluminescent photophores on euphausiaceans may facilitate sexual identification, although this is not proven. One of the most sensational courtship displays in crustaceans is created by male ostracods that emerge from coral reef substratum about an hour after dark, emitting sequential puffs of bioluminescent material from the maxillary glands in a species-specific pattern as they ascend toward the surface. Different species deposit the discrete puffs of light in uniformly spaced vertical strings, suspend them in a line at a 45° angle from the substratum, or place sequential clusters at different heights. Females do not emit bioluminescent puffs, but emerge from the reef and follow the trail of lights to mate with the displaying males. A horde of nonluminescent males also follow the light trails, apparently to "sneak" matings with the pursuing females.

Most crustaceans produce round nonmotile sperm, although the sperm of some maxillopod groups are flagellated. The sperm, released either in seminal fluid or in packets of spermatophores, are deposited directly into the female genital openings or into a fused medial receptacle. Females may store sperm for long periods (up to several years in some astacid lobsters), although subsequent matings can dilute the sperm of a particular male.

Most crustacean matings are relatively brief (minutes to hours), but some crustaceans undertake long arduous courtships and maintain long-term pair bonds. Since little paternal care or mate provisioning is documented, most pair bonds

probably function to ensure that another male does not also fertilize the female. Males of isopods, amphipods, and brachyuran crabs often sequester females, holding and carrying the smaller female for extended periods prior to copulation, which frequently follows female molting. In the violently aggressive and solitary stomatopods, mating is accompanied by an extended period of courtship (involving both visual displays and tactile stroking) that gradually mediates the tendency to smash the potential mate with the clublike or spearing raptorial appendages. The male and female reside together (either in the female's or male's burrow, depending on the taxonomic group) for a number of days, copulating repeatedly. After as much as a week, the female's aggressive levels rise and she evicts the male before laying the eggs. Commensal pinnotherid crabs ("pea crabs," which inhabit oysters and other organisms' burrows), parasitic isopods, stenopodid shrimps, and many groups of caridean shrimps usually are found in pairs, and probably maintain extended (seasonal or permanent) pair bonds. Some deep-sea stenopodid shrimps apparently colonize glass sponges as postlarvae, and the male and female (perhaps the opposite sex is induced in the later arrival) become permanently imprisoned in the crystalline grid-work of the sponge as they grow. These life-long pair bonds are the source of a romantic oriental legend.

Most crustaceans brood their young, but there is a surprising diversity in methods of treating the young, and some release the embryos free into the environment. The remipedes probably shed their embryos directly into the water (no brooding structures have been observed). In the branchiopods, the notostracans brood their embryos very briefly, then deposit them on the substratum, while most cladocerans release their embryos into the water. In the maxillopods, some ostracods release embryos directly into the water and some deposit them on the substratum. The interstitial mystacocarids apparently lay their embryos free among sand grains. Some copepods release embryos directly into water. Branchiurans leave their host to attach their embryos to a plant or rock. In the malacostracans, the extinct Late Paleozoic belotelsonideans, waterstonellideans, and palaeocaridaceans probably shed their embryos directly into the water, since no structural adaptations for brooding have been seen. Contemporary anaspidacean syncarids, and probably bathynellaceans, release their embryos directly into the environment. Contemporary stomatopods deposit their eggs in a free wheel-shaped disk that the female holds in her maxillipeds, grooming and defending them for several weeks. Most euphausiaceans and most dendrobranchiate shrimps shed their embryos directly into the water.

However, most crustaceans exhibit a variety of structural, physiological, and behavioral modifications to better protect their progeny. Cephalocarids brood one or two embryos at a time on anterior abdominal processes, and the allied Devonian lipostracans had a flap-like egg pouch. Among the branchiopods, the anostracans carry their eggs briefly in an egg sac before releasing them. Some cladocerans shelter their eggs in their carapace; the embryos even obtain nutrients from the sides of the brood chamber in some species. One species of cladoceran broods eggs in the molted carapace. Conchostracans carry their embryos on the thoracopods inside the bivalved carapace. Most ostracods brood their young inside the bivalved carapace. Copepods usually drag an egg sac behind the body, but some brood embryos on the sixth thoracopods. In the ascothoracidan cirripeds, the female's carapace is enlarged (enclosing her body) for brooding the young. Embryos are released at the nauplius stage or retained in the carapace through all larval stages up to the cyprid. The embryos of acrothoracican cirripeds hatch at an early stage but are brooded in the carapace through the cyprid stage. The body of thoracican barnacles is engulfed by a fleshy mantle cavity in which the embryos are brooded through the nauplius stage. In the malacostracans, leptostracan females brood embryos under their carapace until after two larval molts. As noted earlier, female stomatopods undertake extensive care and protection for their free egg mass within their cavities in coral or excavated burrows in sand or mud. Female mysidans bear a brood pouch on the last two or three thoracopods. The enlarged, thin, blood-filled carapace of female thermosbaenaceans serves as a brood chamber as well as for respiration. Similarly, the carapace of female amphionidaceans is thin and greatly enlarged, and it may be used for brooding. In this group, the first pair of pleopods, greatly enlarged and extending anteriorly beneath the carapace, are thought to assist in brooding the young. The brood pouch is particularly well developed in the tanaids, cumaceans, mictaceans, isopods, and amphipods. This "marsupius" is formed from expanded flaps, called oostegites, at the bases of thoracopods 3 to 6 in tanaids and cumaceans. Sometimes the oostegites are restricted to thorapod six in tanaids. Cumaceans, like isopods, bear an additional oostegite projecting toward the posterior from the maxillipeds. Female mictaceans bear oostegites either on thoracopods 2 to 6 or 3 to 7. Female isopods have oostegites on the second to fifth, second to sixth, or all eight thoracopods. One family forms a brood pouch from ventral extensions of the body wall rather than oostegite projections from the base of the legs. Several types of internal brooding are present in isopods, including paired invaginated pockets in the cuticle with narrow openings and a single median pocket extending into the thorax between the sixth and seventh segments. Two groups (the gnathiids and flabelliferans) brood young within the body cavity itself; in the flabelliferans, the enlarged oviducts serve as uteri and the oostegites are reduced to small flaps covering the gonopore. Female amphipods develop oostegites on the third to fifth, third to sixth, or fourth to fifth thoracopods; many amphipod taxa also have simple or elaborate gills that protrude from the base of the thoracopods; in some cases the oostegites are reduced and the gills provide a brood chamber. A few euphausiaceans and one species of dendrobranchiate shrimp brood embryos on their posterior thoracopods. Most other decapods, including the carideans, stenopodideans, thalassinideans, astacids, palinurans, anomurans, and brachyurans, brood embryos on their pleopods.

Crustacean embryos proceed through some form of spiral cleavage, often highly modified by the amount of yolk present, to form a nauplius larva, a triangular shaped organism (larger in the anterior) with a simple median eye (ocellus) and three pairs of biramous appendages (the future antennules, antennae, and mandibles). The embryo may either emerge as a plankton-feeding nauplius when yolk reserves are depleted, or the nauplius may be laden with yolk, reducing the need to feed until a later stage. Larvae hatching as nauplii require long

developmental times and are observed in many of the "less advanced" groups (some anostracans, some notostracans, some cladocerans, conchostracans (yolkey), some ostracods, mystacocarids (develop in interstitial sand), nonparasitic copepods (nonfeeding), some ascothoracidan cirripeds, some acrothoracican cirripeds, some thoracican cirripeds, rhizocephalan cirripeds (nonfeeding), euphausiaceans, and dendrobranchiate shrimps). Evolution appears to have progressed toward hatching at a more advanced stage, either as a metanauplius (more segments but still only three appendages), a protozoea (where a metanaupliar stage undergoes a metamorphic molt within the egg membrane to a larval stage with sessile compound eyes, maxillules, and maxillae) or a zoea (a more advanced larval form with stalked eyes and three pairs of maxillilipeds; these swimming larvae eventually hatch into a postlarval form resembling an immature adult). The cephalocarids, fossil lipostracans, some anostracans, some notostracans, some cladocerans, some ostracods, some ascothoracidan cirripeds, and some acrothoracican cirripeds hatch as metanauplii. The reduction or loss of the free larval stage is accomplished by retaining the naupliar, metanaupliar, or protozoeal stages within the egg membrane during a longer brooding period before hatching or, in some cases, by shortening the total development time by omitting the morphological development of some or all of these larval stages entirely.

Crustaceans have three types of development, including "direct," "anamorphic," and "metamorphic" development; the last two forms sometimes are called less and more extreme types of "mixed" development. In direct development, embryos hatch as juveniles that resemble miniature adults, either because larval stages are passed inside brooded eggs before hatching, because they have been suppressed or lost, or possibly because they never were present. Direct development probably occurs in remipedes and probably was present in the Cambrian canadaspid malacostracans, either because larvae were suppressed or never present. Also, most cladocerans, most ostracods, most branchiurans, leptostracans, anaspidaceans and bathynellaceans, and all of the peracarids show direct development. The young hatch as an adult-like juvenile in some freshwater and terrestrial astacids and brachyurans and in some taxa of very small body size, including some stomatopods and some of the very small caridean shrimps. Most of the instances of abbreviated larval phases probably are explained by passage of larval stages within brooded eggs, but some cladocerans, some ostracods, branchiurans, and the anaspidacean and bathynellacean syncarids do not brood. Either larval stages were suppressed or never were present in these groups.

In anamorphic development, embryos hatch as some form of naupliar, metanaupliar, or later-stage larva, and the adult form is gradually achieved by a progression of morphological changes as the larva acquires more segments and appendages. Anamorphic development is found in many cephalocarids and allied fossil lipostracans, anostracans, notostracans, some cladocerans, conchostracans, some ostracods, mystacocarids, copepods, and one group of branchiurans. In the stomatopods, euphausiaceans and amphionidaceans, the morphology of the larvae generally resembles the form of the adult (thus not requiring a major metamorphosis) so these groups could be considered to have anamorphic development; however, these three groups usually are considered to have metamorphic development because they have protozoeal and zoeal larval forms.

In metamorphic development, the embryo hatches as some form of naupliar, metanaupliar, or later stage larva and develops with a body form that is radically different from that of the adult; thus, a major morphological reorganization (similar to that of holometabolous insects such as caterpillars and butterflies) is required before the adult stage can be attained. All of the eucarids as well as the stomatopods usually are considered to have metamorphic development because they possess protozoeal and zoeal larval forms. Other groups that undergo a major metamorphosis between the larval and adult phases include the cirripeds, although they lack the typical protozoeal and zoeal stages. The cirripeds are unique in developing a morphologically distinct nonfeeding bivalved postlarval form called a cyprid before transforming into the adult phase. In thoracicans, the cyprid attaches to the substratum with its antennule and reorganizes its body into a sessile adult. In addition to a cyprid stage, boring acrothoracicans metamorphose into a postcyprid pupa stage in order to reorganize itself into the adult form, similar to holometabolous insects. Following four feeding naupliar stages and a nonfeeding cyprid stage that attaches to the prospective host, the parasitic rhizocephalans form a hypodermic-like "kentrogon" stage through which the body, completely transformed, injects itself into the host's body. The parasitic tantulocarids and copepods also undergo metamorphoses in order to attain the highly modified adult parasitic form. The naupliar larval forms are suppressed in parasitic copepods, which hatch as copepodites (small juveniles).

The number of eggs produced by crustaceans generally correlates with body size, with larger species within lineages usually producing vastly more eggs than smaller species. This tracks only within lineages, however, and, if multiple lineages are combined, there is no trend. Certain crustacean groups are characterized by "weedy" life history characteristics while others are long lived and slow to reproduce. Anostracans and notostracans usually live only a few weeks and produce many eggs, including resistant eggs that are capable of surviving long periods of unfavorable conditions, similar to the seed banks of desert plants. Some copepods also produce dormant eggs. Cladocerans pass through several generations in one summer season. Amphionidaceans may be short-lived as adults (the female does not feed), and males in several taxa (gnathiid isopods, dwarf males in rhizocephalan cirripeds) are short-lived and do not feed. On the other hand, mysidans reproduce infrequently and may require up to 7 years to mature. Some Arctic cumaceans have delayed maturity and live many years, reproducing little. Some stomatopods, astacid and palinuran decapods, and some brachyurans live many years.

The significance of the type of development and number of eggs produced in crustaceans lies in their relationship to dispersal. Species that produce more eggs and those that are released at early stages so that longer periods are spent in the plankton have the largest geographic ranges and are the most resistant to environmental change and extinction. The resistant eggs of anostracans and notostracans often are dispersed by wind or on the feet of birds. Nonparasitic copepods develop through six naupliar stages and five copepodite stages.

Stomatopods pass several protozoeal stages (advanced naupliar or metamaupliar stages) in the egg while brooded by the female. The protozoeal stage, termed an ''antizoea'' or ''pseudozoea,'' is followed by four to eight or nine zoeal stages, called ''erichthus'' or ''alima'' larvae, in small reef and larger level bottom lineages, respectively. These modified zoeae are pelagic for periods that range from 3 to 4weeks to 9 months in small reef versus larger level bottom species, respectively. A minority of small species of stomatopods produce only a few, unusually large eggs relative to their body size and undergo their entire larval development within the burrow of the female. Stenopodid shrimps, palinuran lobsters, and some anomurans are characterized by many larval stages and long pelagic phases. Most small species, most species that inhabit fresh water and terrestrial habitats, many species that inhabit high latitude or deep cold environments, and species with anamorphic or metamorphic development that hatch at advanced stages are likely to produce fewer, relatively larger eggs with shorter pelagic phases. These species, and those with direct development, are most likely to have extremely restricted, endemic distributions.

The Geographic Distributions of Crustaceans

Many of the geographic distributions for each of the major groups of crustaceans strongly reflect ancient conditions where the group originated, indicating that dispersal (usually by long lived pelagic larvae) has not dominated the historical ecology of many groups of crustaceans. The importance of intermediate dispersal in speciation, however, where larvae travel far enough to establish new populations but not sufficiently frequently to swamp arising genetic variants, is well illustrated in the thoracican barnacles. Also, the long lived pelagic phases of some of the large level bottom stomatopods have been shown to dampen diversification within and among regions, while those lineages with short lived larvae and significantly more restricted geographic ranges rapidly speciate and also suffer higher rates of apparent extinction (Reaka 1980; Reaka and Manning, 1981, 1987). The distributions of fossil taxa are not extensively considered below because, except for the worldwide distributions of the archaeostracan phyllocarid malacostracans, many extinct taxa are recorded from relatively few regions, possibly as a result of preservation bias.

To understand patterns of distribution in contemporary crustaceans, data for all of the major taxa were divided into nested categories where the entire group (subclass, order, or suborder) had a wide ($>$half the circumference of the earth) or narrow ($\leq$half the earth's circumference) global distribution, those with many (>100) or few species (≤ 100), and those in which most of the individual species had wide or narrow geographic ranges, producing a table of eight columns. This procedure revealed the somewhat surprising result that entire lineages of crustaceans with narrow distributions are rare.

Restricted geographic distributions for entire lineages are found only in four taxa (the nectiopodan remipedes, anaspidacean syncarids, thermosbaenacean pancarids/peracarids, and spelaeogriphacean peracarids). Each of these groups also contains few species (10, 15, 11, and 2, respectively) that all have highly restricted individual geographic ranges (one or a few localities). These four lineages likely represent the remnants of once more broadly distributed groups that have experienced extinction outside of the marginal refugia they currently inhabit. There are no lineages with narrow distributions and few but broadly distributed species, and no lineages with narrow distributions and many species with either wide or narrow species ranges.

Among groups where the entire taxon has a broad distribution, only six groups have low species diversity and contain species with restricted geographic ranges (cephalocarids (10 species); tantulocarids (5 species); bathynellaceans (100 species); mictaceans (3 species); procaridid shrimps (3 species); and stenopodidean shrimps (25 species)). Widely distributed lineages with few species but broad species ranges include another seven groups (notostracans (11 species), mystacocarids (10 species), acrothoracicans (most of the 50 species), leptostracans (most of the 13 species), lophogastridans (most of the 40 species), euphausiaceans (more than half of the 90 species), and the single, though variable, amphionidacean species).

Groups in which the entire taxon is widely distributed on a global scale and that contain many species with broad species ranges are not numerous. Only most dendrobranchiate shrimps and most palinurans fit clearly within this category, although whether the majority of copepods, cumaceans, anomurans, and brachyurans have wide or narrow individual species ranges is not well documented.

By far the most common distribution pattern observed in crustaceans includes lineages that are speciose (>100 species), broadly distributed on a global scale, and contain species with relatively restricted geographic ranges (14 groups, 18 if copepods, cumaceans, anomurans, and brachyurans, whose status is not certain, are included). These lineages include most of the ostracods (>8000 species), anostracans (>180 species), probably most of the cladocerans (>450 species), most of the conchostracans, (>180 species), most of the branchiurans (150 species), most of the thoracicans (>800 species), most of the stomatopods (>400 species), probably most of the tanaids (>850 species), most of the isopods (>4000 species), most of the amphipods (>6000 species), probably most of the carideans (2500 species), and probably most of the thalassinideans and astacideans.

In summary, the major lineages of crustaceans are most commonly broadly distributed on a global scale (29 of the 33 taxa documented). However, the dominant pattern for species is to have relatively restricted geographic ranges. In at least 24 of the 33 groups documented, the majority of species have relatively small geographic ranges, in some cases restricted to a single locality. Lineages containing many species and those with few species are about equally represented (16 and 17, respectively).

Vulnerability to Extinction

Extinction in the Fossil Record

Extinction in the fossil record tells us a considerable amount about how prone modern crustaceans are to extinction. Studies show that the breadth of geographic distribution of a species or lineage, which is intimately intertwined with their larval dispersal potential, is the single most important

predictor of extinction. Factors that favor extinction or survival, however, depend on whether one is considering organisms during normal levels of background extinction or during cataclysmic mass extinctions (Jablonski, 1991). During background conditions, lineages with long-lived planktonic larvae persist longer in the fossil record (i.e., have lower extinction rates) than those with abbreviated development, individual species with broad geographic ranges suffer lower extinction than those with narrow endemic distributions, and lineages characterized by high species richness survive longer than those with few species. Species-rich lineages composed of broadly distributed species have especially high survival. During mass extinction episodes, however, all of these traits are ineffectual, and the broad geographic deployment of an entire lineage, regardless of the ranges of the individual species or the number of component species, is the primary characteristic that increases the likelihood that the lineage will persist into the next geological era (Jablonski, 1986, 1991).

Jablonski (1991) designated nine major mass extinctions (Late Ordovician, 61% and 85% of genera and species lost; Late Devonian, 55% and 82% of genera and species lost; Late Permian, 84% and 96% of genera and species lost; Late Triassic, 47% and 76% of genera and species lost; Middle Jurassic, 26% and 53% of genera and species lost; End Jurassic, 21% and 45% of genera and species lost; Middle Cretaceous 26% and 53% of genera and species lost; End Cretaceous 47% and 76% of genera and species lost; Late Eocene, 15% and 35% of genera and species lost). Of 15 crustacean lineages that became extinct, 9 of them did not persist beyond one of these biotic crises. Although this does not prove that the mass extinctions were directly responsible for their demise, the failure of lineages to survive into the next period, especially in taxa that had peristed through several previous geological periods, suggests that the conditions associated with the mass extinction may have contributed to their disappearance. In many cases, representation in the fossil record is too spotty to determine whether the lineage and its component species had broad geographic ranges, but some useful information can be gleaned from what we know about fossil crustaceans. The lipostracans (allied to the cephalocarids) and the archaeostomatopod hoplocarids are known only from the Devonian, and the kazacharthrans (allied to the notostracans) are recorded only from the Jurassic. Mass extinctions closed both of these periods, and it appears that these lineages were not widely distributed, being known only from Laurentia (North America and Europe), Scotland, and Kazachstan, respectively. Also, only a few species are known in each lineage. Known only from equatorial Laurentia in the Cambrian and Ordovician, the hymenostracan phyllocarid malacostracans did not survive beyond the Late Ordovician mass extinction. The extraordinarily successful archaeostracan phyllocarid malacostracans persisted from the beginning until the end of the Paleozoic, making it through both the Late Ordovician and Late Devonian crises but failing to survive the greatest of mass extinctions in the Late Permian. This lineage had a worldwide distribution and included many taxa that occupied diverse habitats; these factors may have facilitated their persistence through several major crises but were not sufficient to allow survival through the End of Paleozoic debacle. Other primitive malacostracans, including the waterstonellideans, eocaridaceans, paleocaridacean syncarids, and pygocephalomorph mysidaceans, arose in the Carboniferous but also perished during the End of Paleozoic extinction. On the other hand, however, one order of the primitive remipedes (the enantiapodans), the canadaspid and hoplostracan phyllocarid malacostracans, the aeschronectid hoplocarids, and the belotelsonid malacostracans, all known only from the Carboniferous, and the palaeostomatopod hoplocarids, which originated in the Devonian and persisted until the Carboniferous, all perished in the absence of a major biotic crisis.

Tropical, especially reef, biotas have throughout history suffered disproportionately greater losses to mass extinctions than those at higher latitudes (Jablonski, 1991). Perusal of the habitats and distributions of the extinct fossil crustaceans suggests that most of those lost occurred in regions that were tropical at the time. The tropics also are the largest source of evolutionary novelties (due to speciation and development of major new adaptive types). This is possibly because the recurring extinction events open new evolutionary opportunities there, because high diversity in the tropics makes origination of new types more probablistic simply because of the large numbers of lineages available for genetic experiments, or because something about the ecology of tropical environments causes greater divergence per speciation event (Jablonski, 1993).

Low density and small total population size also has been shown to render species susceptible to extinction in some taxa (Raup, 1986), and this characteristic has been of overriding concern to many conservation biologists in recent years. Evidence on the importance of population density as a predictor of extinction is equivocal in crustaceans. On one hand, the archaeostracan phyllocarid malacacostracans were one of the most abundant benthic organisms, and they persisted from the Cambrian until the end of the Paleozoic, supporting the thesis that population density is important. On the other hand, however, the archaeostomatopod hoplocarids were abundant during the Devonian but did not survive into the next period. Aeschronectid hoplocarids dominated their benthic communities and belotelsonid malacostracans were abundant during the Carboniferous, but neither persisted beyond that period. Waterstonellidean malacostracans formed abundant aggregations, palaeocaridacean syncarids were very numerous, and pygocephalomorph mysidaceans were among the most important components of their communities during the Late Paleozoic (Carboniferous to Permian); yet all succumbed to extinction at the end of the Permian.

Similarly, when a species occupies a unique or particularly restricted type of habitat, environmental changes are more likely to entirely eliminate these localized populations, and individuals may be unable to find or migrate to a similar specialized environment. This feature of organism-environmental interaction also has been a dominant concern in conservation biology. The evidence from fossil crustaceans on this topic, however, is not clear. An example of a lineage with a very specialized habitat that became extinct is the lipostracans, allied to the cephalocarids, that lived in Devonian hot springs. However, other crustaceans that became extinct were not obviously specialized for unique habitat types and inhabited apparently generalized benthic coastal environments. Enantiopodan remipedes; canadaspid and hoplostracan phyllocarid malacostracans; belotelsonid

and eocaridacean malacostracans; archaeostomatopod, aeschronectid, and palaeostomatopod hoplocarids; palaeocaridacean syncarids; and pygocephalomorph mysidaceans all inhabited shallow benthic coastal environments. Of these, at least the abundant aeschronectids, pygocephalomorphs, and palaeocaridaceans also co-occurred in brackish swamps and deltas, and at least the palaeocaridaceans extended into fresh water. The latter four lineages possibly could be considered to exhibit some habitat specialization (fresh water), but most of the lineages lost to extinction were not habitat specialists.

Broadly compared across phyla, the fossil record shows that species high on the food pyramid (predators or parasites) have sometimes been at risk during environmental perturbations, particularly if they also had achieved large body size and had developed specialized morphology and high levels of activity (e.g., for predation). It is thought that their high metabolic requirements and need for large amounts of food imperils their survival during difficult periods compared to smaller species that usually are lower in the food web. A related concept, known from the 1800s as Cope's Rule, posits that most lineages originate as small generalized forms, evolve toward larger and more specialized morphologies, then become extinct (Stanley, 1973). Body sizes of extinct crustaceans are not synthesized in the literature in most cases, but the few available cast doubt on whether or not crustacean lineages of large body size are particularly succeptible to extinction. Enantiopodan remipedes probably were relatively small, kazaharthrans were small to moderately sized (<30 mm), and palaeocaridaceans and phygocephalomorphs probably were small to moderately sized for crustaceans. Yet all became extinct. In many crustaceans (Reaka and Manning, 1981, 1987; Reaka, 1980) and among invertebrates in general (Menge 1975; Strathmann and Strathmann, 1982), larger species produce many more small offspring that spend considerably longer periods in the plankton and have wider geographic ranges than smaller species. Larval dispersal and wide geographic ranges tend to slow down both extinction and speciation (Jablonski 1986; Jablonski and Lutz, 1983), suggesting that, in marine invertebrates, dispersal at large body sizes usually overpowers susceptibility to extinction. However, not all large-sized species produce large numbers of widely dispersing offspring (especially in brackish, freshwater, and terrestrial environments or in lineages where abrreviated development is inherited from past legacies), and available food may be sufficiently restricted to prevent reproduction (thus dispersal) during particularly stressful environmental conditions. Thus, although Cope's Rule may apply more universally among fossil vertebrates (which were Cope's original source of data for the pattern of greater extinction in species of large body size) than among invertebrates, large body size still may render lineages susceptible to extinction among invertebrates in some situations.

The Prospects of Extinction in Modern Crustaceans: Which Taxa are Most Vulnerable, and which Ones Represent the Most Urgent Conservation Targets?

Table 2 summarizes the major factors that are likely to affect the viability of contemporary crustacean groups during natural or human-induced environmental changes. Many scientists caution that drastic environmental shifts could occur as human populations increase exponentially (perhaps doubling in the next 100 years) and affect the atmosphere, soil, and coastal waters (Wilson, 1989). If such change impacts the biotic world in a manner similar to the mass extinctions of the past, the primary factor that would facilitate survival of the crustacean groups analyzed in this study would be wide distribution of entire groups. Fortunately, most (35 of 39) groups of crustaceans are distributed nearly worldwide, suggesting that at least some of these groups probably would survive. The remipedes, anaspidaceans, thermosbaenaceans, and spelaeogriphaceans, however, would be at high risk.

If the deteriorating environment resembled background conditions rather than a mass extinction event, however, the number of lineages within the group, the size of the geographic ranges of species, whether or not most of the group inhabited tropical (especially reef) environments, and possibly whether the species were characterized by low population densities, occupied specialized or highly restricted types of habitat, were positioned high in the food web of their communities, or were large and active could influence the probability that they would succumb to extinction over the long run.

Of the 34 groups for which data were available on species numbers, 25 groups (74%) were considered to be at risk as a result of relatively few species within the lineage (14 at high risk, 11 at moderate risk). The remipedes, cephalocarids, notostracans, mystacocarids, leptostracans, tantulocarids, anaspidaceans, thermosbaenaceans, and spelaeogriphaceans all have fewer than 15 species and would be at very high risk. Even more notable, the mictacean and procaridid lineages are represented by only 3 species each, and only 1 species is known in the entire Order Amphionidacea. These phylogenetically unique groups could easily be lost.

Of 39 groups, 27 (70%) are composed of species with mostly restricted geographic ranges. Ten of these groups contain species with especially restricted distributions, considered at high risk, and 17 contain some species with restricted ranges that are judged to be at moderate risk. Of the 39 groups, 14 (36%) are considered to be at risk because they are primarily tropical or reef dwelling. Twenty-seven (74%) of the 39 groups are characterized by small population sizes. Eight of these have especially low densities or total populations and are considered to be at high risk, and 19 appear to be at moderate risk. Of the 39 groups, 29 (74%) occupy unique or highly specialized types of habitats; 11 and 18 of these are judged to be at high or moderate risk, respectively, because of their degree of habitat specialization. Although many crustaceans are predators or parasites, relatively few are at the apex of the food web (this position often being usurped by fishes), and relatively few are species specific in their choice of prey or hosts. Of the 39 groups, 23 (59%) are judged to be at moderate risk due to a relatively high position in the food chain. Also, relatively few entire groups of crustaceans are sufficiently large in body size that their persistence is likely to be imperiled. Large body size is unlikely to be associated with extensive extinction for the following reasons. Wherever the distributions of body sizes among species of a group have been compared, there always are far more species of small than large body size (May 1988; Reaka-Kudla, 1991, 1997), and this almost certainly is true throughout the Crustacea. Also, in many crustacean groups, species of large body size

Table 2 Summary of extinction vulnerability and conservation value of different groups of contemporary crustaceans

Taxon	*Few living species (++<100, +<500)*	*Entire group restricted geographic range (<half circumference Earth)*	*Most species restricted geographic range*	*Group primarily tropical*	*Small population size*	*Unique or restricted habitat type*	*High trophic level*	*Large body size (or long lived)*	*Old (+probably since mid-Paleozoic, ++probably since beginning Paleozoic) or phylogenetically unique lineage*	*Total score (sum of +,± counted as 0.5)*
Remipedia	++(10)	++	++	+	++	++(brackish caves)	+		++	15
Cephalocarida	++(10)		++		++				+	7
Anostraca	+(≥180)		+			+(ephemeral pools)			+	4
Notostraca	++(11)					+(ephemeral pools)			+	4
Cladocera	+(≥450)		+?							1.5
Conchostraca	+(≥180)		?						+	2
Ostracoda	(about 8000)		++	+		+(patchy)	±(some)		++	6.5
Mystacocarida	++(10)				±(patchy)	++(interstitial)	±(some)			5
Copepoda	(9000)		?	+	±(some)	+(some parasites)	±(some)	±(some parasites, relative)		3.5
Branchiura	+(150)		+	+		+(ectoparasites)	+			5
Tantulocarida	++(5)		++		++	++(parasites)	+	±(relative)		9.5
Ascothoracida	+(60)		?	+	+	+(parasites)	+		++	7
Acrothoracica	+(50)			+	+	+ (borers)			+	5
Thoracica	(≥800)		+		±(some)	+(intertidal, commensal)			++	4.5
Rhizocephala	+(230)		?		±(local)	++(endoparasites)	+	+(relative)		5.5
Leptostraca	++(14)		+		+				+	5
Stomatopoda	+(400)		+	+	±(some)	±(some)	+(active)	±(some)	+	6.5
Anaspidacea	++(≥15)	++	++		+	++(fresh water refugia)			±(allied to primitive groups)	9.5
Bathynellacea	+(100)		++		++?	++(ground water)			±(allied to primitive groups)	7.5

Mysida	(>745)		+		±(some)	+(some intertidal, interstitial, fresh water, caves)	±(some)	±(some long lived, low reproduction)	±(allied to primitive groups)	4
Lophogastrida	+(40)		?		+?		±(some)		+	3.5
Thermosbaenacea	++(11)	+	++		+	++(fresh and salt ponds, caves, hot springs, ground water)			+	9
Tanaidacea	(850)		+?				±(some)		+	2.5
Spelaeogriphacea	++(3)	+	++		++	++(fresh water caves)			+	10
Cumacea	(≥1000)		?		±(patchy)	±(patchy)	±(some)	±(some long lived)	+	3
Mictacea	++(3)		++		++	++(caves, deep sea)			+	9
Isopoda	(>4000)		+		±(some)	+(many, including terresrial, parasites)	±(some)		+	4
Amphipoda	(>6000)		+			±(some, including terrestrial, parasites)	±(some)			2
Euphausiacea	++(90)		±(some)				+			3.5
Amphionidacea	++(1)			++					++phylogenetically unique, 1 genus and species	6
Dendrobranchiata	+(450)			+			+(active)		+	4
Caridea	(>2000)		+	+	±(some)	±(some)	+(active)	±(some)		5.5
Procarididae	++(3)		++		+	++(brackish caves)			+	8
Stenopodidea	++(25)		+	+	++	++(commensal)	+(active)			9
Thalassinidea			+	+	±(some)	±(some, cryptic reef)	+(some, active)	+(many)		10
Astacida			+			+	+(active)	+	+	5
Palinura			±(some)	+	+		+(active)	+	+	5.5
Anomura				+		+(many)	±(some, active)			2.5
Brachyura			±(some)	+	±(some)	+(many)	+(many, active)	+	+?	6

produce long-lived pelagic larvae that confer protection against background extinction. The present analysis suggests that 12 of 39 groups (31%) are at possible risk due to their relatively large body sizes (if other life history characteristics do not outweigh the effect of body size).

One of the factors that should be included in prioritizing conservation efforts is the uniqueness of the lineage. Some crustacean groups represent a heritage that stems from half a billion years ago (the beginning of the Paleozoic). Among groups still alive today, the remipedes, ostracods, ascothoracidan barnacles, and thoracican barnacles are known from Cambrian fossils or their morphology suggests that they are derived from the most primitive lineages of crustaceans. Sixteen other groups are known from the Mid-Paleozoic (300–400 million years ago). These ancient groups are more likely than more recently derived groups to have totally novel characteristics, even in their biochemical and genetic constitution. As we enter a millenium in which biotechnology has only begun to offer human benefits through genetic manipulations, we can ill afford to lose these lineages. Some of these ancient groups are particularly vulnerable due to few constituent species, restricted geographic ranges, small populations, specialized habitat requirements, or other factors (especially the remipedes, cephalocarids, notostracans, ascothoracidan and acrothoracican barnacles, leptostracans, anaspidaceans, thermosbaenaceans, spelaeogriphaceans, mictaceans, procaridids, and especially the amphionidaceans). These groups merit special conservation attention.

Table 2 provides a tentative overall summation score for risk factors and intrinsic conservation value in each group. The overall mean risk/value score for the 39 groups was 5.9. Incorporating all of the different risk and value factors, eight groups (the remipedes, tantulocarids, anaspidaceans, spelaeogriphaceans, mictaceans, procaridids, stenopodideans, and thalassinideans) merit the closest conservation scrutiny overall (score of 9 or above). Of these, the remipedes are considered to be the highest priority. This represents only a beginning, however, at assessing which lineages are in most need of protection in order to ensure their survival in the next centuries. Identifying which lineages and where they are most at risk (as well as the most effective methods for conserving them) will be one of the most important tasks for at least the next decade.

See also: Aquaculture. Endangered Marine Invertebrates. Human Impact on Biodiversity, Overview. Invertebrates, Marine, Overview. Marine Ecosystems. Molluscs

References

Abele LG (ed.) (1982a) *The Biology of Crustacea*. Vol. 1: Systematics, the Fossil Record, and Biogeography. New York: Academic Press.

Abele LG (ed.) (1982b) *The Biology of Crustacea*. Vol. 2: Embryology, Morphology, and Genetics. New York: Academic Press.

Barnes RD (1980) *Invertebrate Zoology*, 4th edn. Philadelphia: Saunders.

Barnes RSK (ed.) (1998) *The Diversity of Living Organisms*. London: Blackwell Sci. Ltd.

Bowman TE and Abele LG (1982) Classification of the Recent Crustacea. In: Abele LG (ed.) *The Biology of Crustacea. 1. Systematics, the Fossil Record and Biogeography*, pp. 1–27. New York: Academic Press.

Brusca RC and Brusca GJ (1990) *Invertebrates*. Sunderland, MA: Sinauer Assoc.

Dominguez JH and Reaka ML (1988) Temporal activity patterns in reef-dwelling stomatopods: A test of alternative hypotheses, J. *Exp. Mar. Biol. Ecol.* 117: 47–69.

Feldmann RM and Manning RB (1992) Crisis in systematic biology in the "Age of Biodiversity." *J. Paleontol.* 66: 157–158.

Gaston KJ and May RM (1992) Taxonomy of taxonomists. *Nature* 356: 281–282.

Gore RH and Heck KL (1986) *Crustacean Biogeography*. Rotterdam: Balkema Press.

Grassle JF and Maciolek NJ (1992) Deep-sea species richness: Regional and local diversity estimates from quantitative bottom samples. *Amer. Nat.* 139: 313–341.

Jablonski D (1986) Background and mass extinctions: The alternation of macroevolutionary regimes. *Science* 231: 129–133.

Jablonski D (1991) Extinctions: A paleontological perspective. *Science* 253: 754–757.

Jablonski D (1993) The tropics as a source of evolutionary novelty through geological time. *Nature* 364: 142–144.

Jablonski D and Lutz RA (1983) Larval ecology of marine benthic invertebrates: Paleobiological implications. *Biol. Rev.* 58: 21–89.

Knowlton N (1993) Sibling species in the sea. *Ann. Rev. Ecol. Syst.* 24: 189–216.

Lovejoy TE (1980) Changes in biological diversity. In: Barney CO (ed.) *The Global 2000 Report to the President 2 (Tech. Rept.)*, pp. 327–332. Harmondsworth: Penguin Books.

May RM (1988) How many species are there on Earth? *Science* 241: 1441–1449.

May RM (1994) Biological diversity: Differences between the land and sea. *Phil. Trans. Roy. Soc. Lond.* B 343: 105–111.

Menge BA (1975) Brood or broadcast? the adaptive significance of different reproductive strategies in two intertidal sea stars. *Leptasterias hexactis and Pistaster ochraceus*. *Mar. Biol.* 31: 87–100.

Norse EA and McManus RE (1980) Ecology and living resources: Biological diversity. In: *Environmental Quality 1980: The Eleventh Annual Report of the Council on Environmental Quality*, pp. 31–80. Washington, D.C.: Council on Environmental Quality.

Pearse V, Pearse J, Buschbaum M, and Buschbaum R (1987) *Living Invertebrates*. Palo Alto, CA: Blackwell Sci. Pub.

Poore GCB and Wilson GDF (1993) Marine species richness. *Nature* 361: 579.

Raup DM (1986) Biological extinction in Earth history. *Science* 231: 1528–1533.

Ray GC (1991) Coastal zone biodiversity patterns. *Bioscience* 41: 490–498.

Reaka ML (1980) Geographic range, life history patterns, and body size in a guild of coral-dwelling mantis shrimps. *Evolution* 34: 1019–1030.

Reaka ML and Manning RB (1981) The behavior of stomatopod Crustacea, and its relationship to rates of evolution, *J. Crust. Biol.* 1: 309–327.

Reaka ML and Manning RB (1987) The significance of body size, dispersal potential, and habitat for rates of morphological evolution in stomatopod Crustacea, *Smithsonian* Contr. *Zool.* 448: 1–46.

Reaka-Kudla ML (1991) Processes regulating biodiversity in coral reef communities on ecological vs. evolutionary time scales. In: Dudley EC (ed.) *The Unity of Evolutionary Biology* 1, pp. 61–70. Portland, OR: Dioscorides Press.

Reaka-Kudla ML (1997) The global biodiversity of coral reefs: A comparison with rain forests. In: Reaka-Kudla ML, Wilson DE, and Wilson EO (eds.) *Biodiversity II: Understanding and Protecting Our Natural Resources*, pp. 83–108. Washington, D.C.: Joseph Henry/National Academy Press.

Reaka-Kudla ML, Wilson DE, and Wilson EO (1997) Santa Rosalia, the turning of the century, and a new age of exploration. In: Reaka-Kudla ML, Wilson DE, and Wilson EO (eds.) *Biodiversity II: Understanding and Protecting Our Natural Resources*, pp. 507–524. Washington, D.C.: Joseph Henry/National Academy Press.

Schram FR (ed.) (1983) *Crustacean Phylogeny*. Crustacean Issues I. Rotterdam: Balkema Press.

Schram FR (1986) *Crustacea*. Oxford: Oxford University Press.

Stanley SM (1973) An explanation for Cope's Rule. *Evolution* 27: 1–26.

Strathmann RR and Strathmann MF (1982) The relationship between adult size and brooding in marine invertebrates, *Amer. Nat.* 119: 91–101.

Systematics Agenda 2000 (1994) *Systematics Agenda 2000: Charting the Biosphere*. Tech. Rept. Consortium of Amer. Soc. Plant Taxonomists, Soc. Syst. Biol., Willi Hennig Soc., in cooperation with Assoc. Syst. Collections, New York.

Wilson EO (1985) Time to revive systematics. *Science* 230: 1227.

Wilson EO (1989) Threats to biodiversity. *Sci. Amer.* September: 108–116.

Wilson EO (1992) *The Diversity of Life*. Cambridge, MA: Belknapp Press.

Wilson EO and Peters FM (eds.) (1988) *BioDiversity*. Washington, D.C.: National Academy Press.

D

Darwin, Charles (Darwinism)

Michael T Ghiselin, California Academy of Sciences, San Francisco, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biota Fauna and flora; the organisms in a given place.
Coevolution Evolution where two species have a mutual influence on each other.
Genetic drift Change in gene frequencies due to random fluctuations in a finite population.
Heterostyly Having two or more kinds of flowers with stamens and pistils of different lengths.
Mimicry Defensive mechanism in which organisms of two species resemble each other, ordinarily because at least one of them is distasteful.
Morphology The study of form.
Phylogeny The history of a species or lineage.
Pleiotropy Phenomenon of a gene affecting more than one trait.
Sexual selection Kind of selection that Darwin distinguished from natural and artificial selection in which reproductive success depends on monopolizing mates.
Taxon Group of organisms in a formal system of classification.

Introduction

Although natural selection theory is generally considered as Darwin's most important and original contribution to knowledge, he also deserves the credit for providing compelling evidence that evolution is a fact and for making its fundamental implications clear. Indeed, Darwin began an intellectual revolution of the first magnitude, one that has affected the whole scope of modern life and culture. Indeed it is still going on. The emphasis here will be on those aspects of Darwin's contribution and of Darwinism in general that are particularly relevant to the study of biodiversity. Within biology it has been the sciences of systematics and ecology that have been most revolutionized by Darwinian thinking because it enables us to understand both how and why diversity originates and is maintained. Some of the larger philosophical and social issues will also be considered, especially as they relate to the central theme. These include the rejection of teleology, the replacement of typological thinking with population thinking, the recognition of the importance of historical contingency, and a competitive view of the natural economy.

Darwin's Life and Work

Childhood and Education

Charles Darwin was born on February 12, 1809, in Shrewsbury, England. His father was a wealthy physician and his mother was a member of the Wedgwood family, famous for the manufacture of pottery. She died when he was eight years old. Young Darwin was not very successful in school and at the age of sixteen he was sent to study medicine at Edinburgh University. There he was able to cultivate his interest in natural history, largely on an extra-curricular basis. A decision to abandon medicine in favor of preparation for a career as a clergyman led him to enroll in Cambridge University on October 15, 1827. During the years at Cambridge he continued his largely extracurricular studies of natural history and received much encouragement from both faculty and students. In those days there was nothing unusual about a clergyman also being a scientist, a university professor, or both. Darwin attended the botany lectures and field trips arranged by Professor John Stevens Henslow (1796–1861) and became an enthusiastic collector of beetles. He passed the examinations for his BA degree on January 22, 1831 but remained there to complete his residency requirements. He took up the study of geology and the following summer he obtained some field experience in the company of Professor Adam Sedgwick (1785–1873).

The *Beagle* Voyage

On September 1, 1831, Darwin accepted an invitation to join a surveying expedition on HMS *Beagle* as an unofficial naturalist and gentlemanly companion for the captain of the ship, Robert FitzRoy (1805–1865). The *Beagle* left for South America on December 27, 1831 and after sailing around the world finally got back to England on October 2, 1836. During the voyage, Darwin was able to observe and collect on a vast scale,

especially in South America, on the coasts of which most of the survey work was done. He also visited Australia, the Cape of Good Hope, and many islands along the way. At first his work emphasized the marine invertebrates that had interested him since his days at Edinburgh. But he took an increasing interest in geology. He read the first volume of the *Principles of Geology* by Charles Lyell (1797–1875) not long after the ship left port, and began to read the second volume the following November and the third in May of 1834. Darwin became an enthusiastic supporter of Lyell's uniformitarian methodology, and was also influenced by the discussions of biogeography and Lamarckism in the *Principles*. Darwin developed a theory of coral reef morphology while still working in South America. Later in the voyage he began to test it by means of observations on reefs. It was the coral reef theory, coupled with other geological work, that first established his high reputation as a scientist.

Observations during the voyage on the geographical distribution of animals and plants ultimately led Darwin to believe in evolution. The differences among geographically separated biotas and the peculiar character of insular ones were crucial in his thinking. It must be stressed, however, that these evolutionary insights did not happen instantaneously and that his commitment to evolution was a later development. Darwin began to take the possibility of evolution seriously only toward the end of the voyage, and the compelling evidence was not really in place until shortly after it was over, and some of his specimens had been examined by specialists. During his visit to the Galapagos Islands (from September 15 to October 20, 1835) he still assumed that species were immutable. He did, however, expect them to be somewhat adaptable, and to exist as local varieties. Some of the birds that he had supposed were just varieties turned out to be distinct species, implying that the inhabitants had diversified more than he had assumed was possible, and furthermore they had done it so locally. They were related by community of descent. This was just the beginning of many of his biogeographical insights. Facts such as the lack of frogs and other amphibians on islands far distant from land could easily be explained in terms of evolution, but otherwise remained a mystery. Contrary to the impression that one sometimes gets from text books, the most compelling evidence for evolution comes from biogeography, not the study of fossils.

Natural Selection

The fossils that Darwin collected in South America turned out to be recently extinct members of groups of animals that were still living in the same area. Again this suggested the possibility of local diversification. Around March, 1837, he began to search for an evolutionary mechanism. After much reading and theorizing, he discovered natural selection. This happened toward the end of September, 1838, when he read the *Essay on the Principle of Population* by Thomas Robert Malthus (1766–1834), who was among other things, a professor of economics. Darwin now understood how reproductive competition between organisms of the same species could produce evolutionary change.

The influence of Malthus on Darwin requires some elaboration here because it is so important for biodiversity studies. Malthus' ideas (or what have been represented as his ideas) have been controversial and have often entered into discussions on population planning and conservation policy. His basic point was that the growth of population tends to outstrip the growth of the food supply. His notion that the growth of population is exponential whereas that of the food supply is linear is well known, although it never was more than just a crude approximation. Malthus based his arguments on empirical evidence. After plagues and other catastrophes have reduced the numbers of human population there are fewer mouths to feed and the amount of food available per capita increases. Consequently people can and do support larger families and population grows rapidly. As time passes, the amount of food per capita decreases and population growth levels off, as people marry later and have smaller families. The standard of living tends to drop as well. A comparable situation could be observed where new land had been brought into cultivation, as in the United States, following which a high living standard was accompanied by rapid population growth.

Malthus saw a problem in this scenario with respect to the future economic prosperity of society: population growth would tend to depress the standard of living. There was a serious question in his mind, as well as those of his critics, whether his model of population growth need have those particular consequences. "Prudential constraints" (as he called sexual abstinence) would have some mitigating effect. The growth of technology has of course at least delayed the sort of famine that might be envisaged. However, the Malthusian model applies quite well to nonhuman populations. It contributed to Darwin's understanding of population dynamics, which he incorporated into his evolutionary and ecological theory. Darwin recognized that there was reproductive competition between individuals of the same species and because these individuals varied, those that were able to utilize resources more effectively would have more offspring. Natural selection follows when it is realized that the properties of those individuals that have more offspring are differentially transmitted by inheritance to the next generation. Extrapolating further, it made sense that natural selection would operate differently depending on whether or not population levels had been lowered by the effects of weather or other density-independent factors.

Given this kind of population dynamics, it became clear to Darwin that variation was far more important than had previously been understood. Natural selection would not work without it. This shift in emphasis is often discussed under the rubric of 'population thinking.' In contrast, the older 'typological' approach, which conceived of groups of organisms in terms of stereotypes, viewed the differences between individuals as a kind of departure from a norm, and these were screened out as unimportant. To the extent that variability had been recognized, it was assumed that something held it within definite limits. It was believed that species might vary but not so far as to give rise to new ones. At a deeper philosophical level, treating species are reproductive populations implied that they were very different from the 'natural kinds' of chemistry and other physical sciences. The new Darwinian thinking accepted that species were concrete, they had a

history, and they could evolve. Treating species as something more than abstractions was a major conceptual innovation.

The Long Delay

Following this shift in his thinking, Darwin immediately began elaborating a comprehensive theory of evolution together with its many implications. However, he was busy with other things and publication of his theory was long delayed. One factor in the delay was the decline of his health, which if not brought on by stress, was at least exacerbated by it. Ill health was a major consideration in his decision to leave London and live in the country. Darwin had married his cousin Emma Wedgwood on January 29, 1839, and the first of their many children was born on December 27th of that year. He moved to the village of Down, not far South of London, and resided there from September, 1842, until his death on April 19, 1882.

He had already published a semipopular account of his travels in 1839, usually called *The Voyage of the Beagle*; it was a great literary success. He also edited the scientific results of the voyage and published three books on the geology of the voyage, including one on coral reefs. He began to publish a few papers on zoology. One of these grew into a vast monograph on barnacles (Cirripedia). It was well over a thousand pages long, and consumed eight years of his working time.

The Origin of Species

At last, in 1854, the work on barnacles was completed and Darwin began to write a long book on evolution. It was to incorporate much original research as well as a detailed review of the literature. As it happened, Alfred Russel Wallace (1832–1913) was interested in the possibility of evolution too and spent many years in the field collecting animals and plants and studying their geographical distribution, first in South America (1848–1852) and then in the region that is now mostly part of Indonesia (1845–1862). They entered into correspondence and in a letter dated May 11,1857, Darwin informed Wallace that he had an evolutionary theory but did not let on what it was. Wallace may have gained valuable hints from his readings of Darwin's publications, especially the book on the voyage, which contains among other things a brief discussion on Malthus. Nonetheless, he is generally credited with an independent discovery of the principle of natural selection.

A manuscript by Wallace on that topic reached Darwin in June, 1858 and his colleagues arranged for a joint publication, which was read at the Linnean Society on July 1. It became available in print the following month but was not well understood at that time. Almost immediately Darwin began an abstract of the longer book that he had been writing. This "abstract" became *On the Origin of Species by means of natural selection, or the Preservation of favoured races in the struggle for life*. Published on November 24, 1859; it became the focus of a controversy such as the world has never seen before or since.

As a supporter, Darwin already had his friend and confidant, the botanist Joseph Dalton Hooker (1817–1911) and others soon joined them. There were of course public debates, including the one at Oxford on June 9, 1860. Although the exchange between Bishop Samuel Wilberforce (1805–1873) and Thomas Henry Huxley (1825–1895) is better known, Hooker was really more effective in support of Darwin's ideas. Darwin himself furthered his interests through behind the scenes negotiation and, more importantly, through research and publication. There were a total of six editions of the *Origin of Species*, the last of which was published in 1872.

Later Publications

Plans to complete and publish the big book were abandoned, but part of it, somewhat revised, appeared in 1868 as *The Variation of animals and plants under Domestication*. Concurrently Darwin had an active empirical research program under way, much of which was conducted on plants. His book of 1862, entitled *The various contrivances by which British and foreign orchids are fertilized by insects, and on the good effects of intercrossing*, dealt with the problems of adaptation. It was a seminal document of pollination ecology and the study of coevolution. Later works on floral biology, *The effects of cross and selffertilization in the vegetable kingdom* (1876) and *The different forms of flowers on plants of the same species* (1877) addressed mainly the effects of inbreeding and outbreeding.

Darwin was a first-rate plant physiologist, and his book *The movement and habits of climbing plants* had first appeared as a journal article in 1865; it was followed by *The power of movement in plants* (1880), which treated the underlying mechanisms of plant behavior. The book on climbing plants and *insectivorous plants* (1875) are early treatises in what later came to be called ethology. They deal with behavior from a comparative and evolutionary point of view. Treating humans, other animals, and even plants as part of an evolutionary continuum was a major contribution to comparative psychology, as well as a serious challenge to traditional philosophical ideas such as mind–body dualism. There was also a lot of behavior in *The descent of man, and selection in relation to sex* (1871) and its sequel, *The expression of the emotions in man and animals* (1872), and in the last of his books, *The formation of vegetable mold through the action of worms, with observations on their habits* (1881).

Darwin's, work on sexual selection and the evolution of social behavior as treated in, *The descent of man* is strikingly modern in its approach. In retrospect, his efforts to solve the problems of heredity seem much less successful, though they stimulated others to work on them. It is unfortunate that his theory of pangenesis was appended to the 1868 book, *Variation*, with the result that it is rarely read because that theory was superseded. Although pangenesis explained the inheritance of acquired characters, it was more a theory of development than of heredity. In fact the book presents a remarkably sophisticated treatment of the relationships between embryology and evolution.

Darwin's Contributions to the Study of Biodiversity

The Competitive Natural Economy

Darwin succeeded, where others had failed, in convincing the scientific community that evolution had in fact occurred and

that it was responsible for the patterns of diversity that may be observed in both fossil and living organisms. Biogeography, paleontology, and classification took on a new significance, becoming profoundly historical in nature. He was less successful at gaining immediate support for his more fundamental (or at least original) contribution, which was the basic mechanism for evolution and the new way of thinking that went along with it. Evolution by natural selection made it possible to reject all sorts of older ideas about the world, especially the notion that organisms and the 'natural economy' have had somehow been designed, or produced by a supernatural agency. Darwin had demonstrated that there was much less 'purpose' in the world than had been supposed that there was nothing inherently 'good' about the way things are and that there was no necessary progression from lower to higher levels of organization.

Darwin's theory was immediately recognized as having profound philosophical significance. From time immemorial it had seemed natural to interpret the living world and the universe in general as if they were the product of something like human intelligence. Such notions and the anthropomorphic way of thinking that underlies them, are called 'teleology.' The idea that somehow the world had been designed or was otherwise the product of supernatural action was of course fundamental of much religious thinking, even among those who had come to accept early versions of evolution. The remarkable adaptations of living organisms were taken as evidence for the existence of a supernatural order. Although many of the features of inanimate nature could be explained in terms of laws of nature and matter in motion, a more subtle purposefulness was widely attributed to the universe as a whole. Indeed, the great German philosopher Immanuel Kant (1724–1804) had proclaimed that there never could be a 'Newton of the grass-blades' able to explain adaptation in strictly naturalistic terms. But that is precisely what Darwin accomplished.

Darwin also provided a way of thinking about adaptation that allows us to avoid the naive and anthropomorphic thought processes that lead to the illusion of design. Rather than asking what things are for, we ask how they might have evolved and explain the facts in terms of what has happened and why. When this is done properly, it becomes clear that much ascription of adaptive significance has been mere guesswork and that there is far less 'optimality' in the world than is sometimes supposed even today. However, research has repeatedly revealed that supposedly 'nonadaptive' features turn out to be very important in the lives of organisms. Defensive metabolites are an excellent example.

Opponents of Darwinism of course include many persons who have been reluctant to abandon teleological interpretations of nature. It has not always been a matter of denying evolution or natural selection outright. Often evolution has been envisaged as a means of achieving some supernaturally ordained plan. And often the teleology is merely assumed or presupposed rather than being advocated explicitly. 'Holistic' notions that treat the world as if it were a super-organism have a long history in myth and pseudoscience that goes back to antiquity. They retain their popularity among religious persons and advocates of occult metaphysics. Because such older ways of thinking are still widespread, it is important to avoid giving the appearance of endorsing them. Functional explanations should be called 'teleonomic' (i.e., where the apparent purposefulness in organisms is derived from evolutionary adaptation) rather than 'teleological' and metaphors such as 'design' should be avoided.

The idea of progress has rarely been disassociated from teleological thinking. For that reason, among others, Darwin's ideas about progress have often been misinterpreted but rarely understood. They are often confused with quite different ideas, especially those of philosophers, social scientists, and non-Darwinian evolutionists. In pre-evolutionary biology it was widely accepted that organisms could be placed in a single series from lower to higher, connecting to nonliving matter at the bottom and to the various ranks of supernatural beings at the top. This great chain of being or *scala naturae* was later modified to make the order a historical or quasi-evolutionary one. Some authors treated the history of life as something like the development of an embryo and the succession of organisms would become increasingly more complicated because such change was in a sense built in to the ancestral germ. Others attributed the same basic pattern to laws of nature, laws that were supposedly preordained so as to have such an effect. In either case, it was Divine Providence on a geological scale.

Darwin forthrightly rejected that kind of progressionist thinking, especially as it had been presented by his predecessor Jean-Baptiste Lamarck (1744–1829) and his contemporaries Richard Owen (1804–1892) and Herbert Spencer (1820–1903). There was nothing like a cosmic escalator carrying species upward toward higher levels of organization. Nonetheless, Darwin believed that evolutionary progress is real and that his theory could explain it. He was well aware of the progress that had been made during his own lifetime in science and technology. Artificial selection was one technique by means of which domesticated breeds of plants and animals had been much improved. His study of the fossil record clearly indicated to him that analogous directional changes had taken place, even though Lyell and others had long denied that. Natural selection, he reasoned, would lead to the continued accumulation of adaptations that made their possessors increasingly able to compete with those around them. But he saw nothing truly inevitable about that kind of progressive change, for it depends on environmental and other circumstances that are not always realized.

Darwin replaced the older notion of a largely harmonious and cooperative natural economy with that of a fundamentally competitive one. The organisms within populations reproduce differentially as a consequence of how they utilize the resources in their environments and change occurs in the direction of more effective use of those resources. That accounts for adaptation and for such progressive and regressive changes as may in fact occur. But it greatly limits the kind of adaptation that can evolve, because it is organisms, and to some extent families, that compete as reproductive units. Darwin was profoundly aware of the point that his mechanism would not produce adaptations at the species level or at the level of the community or ecosystem. This point, however, was often neglected by his contemporaries and received due emphasis only in the latter half of the twentieth century.

Darwin's idea of what a species is lacked a clear conception of reproductively isolated populations with distinct gene pools.

Furthermore, he did not fully appreciate the difference between genotypic and phenotypic variability. Because his mechanism depended on differences between individual organisms, he needed to show that the necessary variability does in fact exist. He was able to show that species are indeed far more variable than his predecessors had believed. He also discovered that sexual reproduction and outcrossing are much more prevalent than even he had at first suspected. Therefore, his contributions to the study of population structure in nature were the foundation for a great deal of work along such lines.

Variation within and between the component populations of a species is one important form of biodiversity within species. Another is the sort of differences that may be observed between the sexes, and here again Darwin's contribution was fundamental. His theory of sexual selection helped to account for sexual dimorphism on the basis of males competing for the females, either through fighting with one another or through efforts to attract the females. He also discovered heterostyly in plants. This is a kind of polymorphism in which hermaphrodite flowers of the same species have different morphologies, furthering outcrossing and reducing inbreeding.

Darwin also was able to explain the global pattern of diversity among species. Of course the splitting up of species and their ability to undergo an indefinite amount of change allows diversification through geological time. However, that by itself does not tell us why or how the diversification takes place. To account for that Darwin invoked what he called a 'principle of divergence of character.' According to this principle, diversification allows the organisms to exploit a wider range of resources, so that groups of them can expand their numbers. As he put it, they can occupy more places in the natural economy, or as we would say, they can occupy more Eltonian niches. It is important not to view this diversity in a strictly negative way, as if it was merely a matter of 'avoiding competition.' Rather, new resources come to be exploited and the original ones may be exploited more effectively.

Darwin was also one of the first scientists to study competition experimentally. In *The Origin of Species* he described field experiments that document changes in diversity as a result of competition. He believed that competition is most severe between closely related organisms, and this tends to drive both intermediate forms to extinction and to bring about divergent evolution among the more modified ones. At lower taxonomic levels this means that higher biotic diversity should develop where the competition is most severe, notably in the tropics. At higher taxonomic levels, it means that there is adaptive radiation within groups such as families, orders, and classes. The lack of intermediate forms among extant taxa is the consequence of those forms that were close to the ancestral ones having been driven to extinction. On a global scale, whole biotas would tend gradually to diversify through time. Darwin specified the conditions under which such diversification would lead to the evolution of competitively effective organisms. These included a large area and a diverse topography, coupled with a long period of competitive interaction among the inhabitants. He realized that the biotas of separate biotic provinces have not undergone the same amount or kind of modification and that similar considerations apply to insular biotas. His insights were thus fundamental to our understanding of invasions, faunal interchanges, and many aspects of extinction.

Darwin contributed a great deal toward our understanding of the interdependencies among species by emphasizing how complicated such ecological relationships may be. He was particularly interested in pollination symbiosis and his work in that area initiated the study of coevolution among animals and plants. Although his research was focused on showing the importance of outcrossing, it also revealed the ecological significance of pollination.

Classification

Systematics is often defined as the study of biodiversity and Darwin made fundamental contributions to both the theory and the content of that science. Classification is part of everybody's language and thinking and is necessary for the organization of knowledge, whether it be the knowledge of a lay-person or a professional scientist. The objects of our experience can be referred to by more or less general terms, such as 'food,' 'pasta,' and 'spaghetti.' Because what is true of the more general groups is true of all the groups that fall under them; we have a powerful instrument for thinking about more than just single objects. If pasta is made of grain, and all spaghetti is pasta, then spaghetti is made of grain.

Scientific classification may be much more sophisticated than that of the lay-person. Scientists attempt to discover groups of things that share properties that are of theoretical importance. A good example of a scientific classification system is the periodic table of the elements, in which chlorine and bromine are both halogens and undergo similar chemical reactions. The groups in question may be worked out on an empirical basis, and only later understood in theoretical terms. This was approximately the case with Darwin's work on biological systematics. Human beings could be assigned to more general groups, such as mammalia, vertebrata, and animalia but it was not obvious why. Sometimes it was attributed to God having some plan in mind when He created organisms, and sometimes unknown laws of nature were thought to be responsible.

Darwin's solution to this puzzle was straight-forward and evolutionary. According to his biodiversity model, species have repeatedly undergone a kind of splitting, giving rise to separate lineages that evolve more and more differences as time since common ancestry becomes greater. The more general the group, the earlier the common ancestry. Once the theoretical reason for traditional classification was understood, a new approach could be developed in which the groups were explicitly recognized as lineages sharing a common ancestor. As Darwin put it: 'Our classifications will come to be, as far as they can be so made, genealogies.'

Putting that prognostication into practice was not easy and, generally speaking, classifications have been only rough approximations to genealogies. They also generally attempt to express the amount of difference that separates the various groups, and sometimes the criteria conflict. Be this as it may, Darwin was writing from his experience. On October 1, 1846, he began to study the anatomy of barnacles in order to write a short paper on an aberrant species that he had discovered. The project expanded to the point that he spent eight years revising and monographing the entire subclass, both living and fossil,

and treating all aspects of their biology in great detail. Although he did not explain what he was doing at the time, it is clear that his classification is based on his understanding of the phylogeny of the group. Subsequent workers have modified and enlarged his system, but the basic arrangement remains the same. Why Darwin delayed publishing his evolutionary theory in order to undertake a major research program in systematics has been somewhat of a puzzle to historians. One possible reason, however, is quite clear. He was interested in topics such as the evolution of organisms with separate sexes from ancestral hermaphrodites. The problems were solvable, but only if the sequence of changes could be documented, and that meant reconstructing the pattern of branching in the phylogenetic tree. Without good systematics, much of evolutionary biology would consist of untested hypotheses and conjectural history. This is as true today as it was for Darwin.

Because, as was already mentioned, Darwin did not have an adequate species concept, modern systematists have made some important changes in Darwin's barnacle classification. In particular, many of the groups that he treated as mere varieties are now classified as species. In many other respects, however, his approach was very advanced for his day. He worked with extensive series of specimens and went out of his way to document all kinds of variation.

Darwinism after Darwin

Alternatives to Darwinism

Darwin maintained that natural selection is the 'main but not exclusive' evolutionary mechanism. This seems about as good a criterion as any for distinguishing 'Darwinians' from a wider range of scientists who accepted evolution but either rejected natural selection or treated it as unimportant relative to other causes. Among these scientists, the followers of Jean-Baptiste Lamarck (1744–1829), who emphasized use and disuse, were particularly numerous. However, there were other alternatives. Étienne Geoffroy Saint-Hilaire (1772–1844), whose views are often confused with those of Lamarck, believed that the environment directly induces changes in structure. Indeed, Darwin was strongly influenced by Geoffroy and accepted the direct action of the environment as a minor cause of change. There were also various versions of 'orthogenesis' or evolution that is directed in a particular direction. Some versions of orthogenesis were based on the notion that evolution is like embryological development and change has been built in from the start. In other versions, laws of nature, similar to those that determine the structure of minerals, were invoked. Sometimes the appeal was to unknown causes, perhaps of a supernatural character. Many scientists were, or at least claimed to be, agnostic with respect to what causes evolution.

Early Darwinians

Given such a range of alternatives and the small amount of research that had been done, it makes sense that from the outset there were few Darwinians other than Darwin himself. Because he was largely responsible for getting evolutionary thinking in general accepted by the intellectual community, a lot of evolutionists who did not accept natural selection nonetheless considered themselves his followers. The degree to which Darwinism, in the sense that we use that term here, was a minority position has sometimes been exaggerated. We can identify quite a number of important contemporaries of Darwin who established successful research programs based on the study of natural selection. Foremost among these of course was Alfred Russel Wallace (1832–1913), its codiscoverer. Also very distinguished was Wallace's traveling companion, Henry Walter Bates (1825–1892), the discoverer of Batesian mimicry. Second only to Darwin in his mastery of the theory was Fritz Müller (1822–1897). He is best remembered for his discovery of Müllerian mimicry, but he also was the first to propound the idea that developmental stages may recapitulate evolutionary ones. Both Fritz Müller and his brother Hermann (1829–1883) conducted magisterial research on pollination symbiosis under Darwin's influence. It is worth emphasizing that these scientists were outstanding for their performance as naturalists in the field. The kind of research that they did has been fundamental to our understanding of biodiversity because it documents how natural selection takes place in real environments.

Darwin also had important followers whose work was more focused in the museum and the laboratory. He had a close circle of followers, botanist Joseph Dalton Hooker (1817–1911), zoologists George John Romanes (1848–1894), and John Lubbock (1834–1913). There was also August Weismann (1834–1914), whose ideas about the continuity of the germplasm made natural selection seem a much more plausible explanation for evolution than Lamarckism. Neo-Darwinism is rightly associated with the name of Weismann, whose basic position was that natural selection is not just the main but the exclusive evolutionary mechanism. Actually he admitted two minor ones that had been invoked by Darwin: sexual selection and pleiotropy.

Genetics and the Modern Synthesis

Beginning in 1900, Mendelism, developed from the revolutionary research of Gregor Johann Mendel (1822–1884), began to supply the understanding of heredity that Darwinism needed. However, it did not have the immediate effect of gaining adherents to Darwin's theory. Instead it gave rise to alternatives such as mutationism. Advocates of mutationism were more impressed by conspicuous mutants and erroneously concluded that their sudden appearance in a single generation could be extrapolated to the origin of new species and even higher taxa. However the study of genetics soon moved away from mutationism and gradually undermined some of the other alternatives, especially Lamarckism and orthogenesis.

The reconciliation of Darwinism with genetics depended in part on the elimination of such alternatives, in part on the growth of theoretical population genetics, and in part on the study of natural populations in the field. Studies of natural populations by systematists like Ernst Mayr (1904–2004) and geneticists like Theodosius Dobzhansky (1900–1975) were crucial to the 'evolutionary synthesis' that began in the 1930s

and matured by around 1950. A new and more sophisticated biological species concept emerged and the richness of genetic diversity within populations became much better understood. The result was a theory very much like Darwin's but without the minor elements of Lamarckism and Geoffroyism integrated with modern genetics.

This modernized Darwinism continued to be expanded and refined. Regarding biodiversity, there were two important developments, both of which involved a return to Darwin's original theory. First, there was Darwin's ecology, which was based on individualistic competition and which rejected the notion that organisms do things 'for the good of the species' or 'in order to benefit the ecosystem.' Ecologists had commonly treated species and ecosystems as if they were superorganisms and had attributed a kind of adaptation to them that could only have resulted from selection of populations. The rejection of 'group selection' in its unsophisticated and uncritical form was largely carried out in the 1960s and 1970s.

Second, there was a return to Darwin's idea that classification should be genealogical. Although both the extent and the manner to which this program should be carried out have been controversial, it is generally agreed that comparative biology is best conducted within the framework of phylogenetic relationships. Ill-defined 'similarity' has proved to be too subjective and not rigorous enough for modern biodiversity research. This development has partly been the result of new methodologies (especially cladistic analysis). It has also been strengthened by new kinds of evidence, such as isozyme studies and genomic sequencing that supplement the more traditional data.

Contemporary Developments

From time to time Darwinism and neo-Darwinism are challenged by what are purported to be genuine alternatives. Some of these challenges, such as efforts to resurrect Lamarckism or orthogenesis, have simply not measured up. Others, such as developmental constraints, are components of Darwinism that may not have received as much attention as they deserve. Still others, such as the notion of punctuated equilibria, with its occasional periods of rapid change interspersed with long interludes of stasis, are by no means incompatible with the Darwinian tradition.

Darwinism has never been so monolithic or devoid of major unanswered questions as to preclude a broad range of possibilities for new developments. At present it is by no means clear to what degree processes at various levels need to be invoked to provide a satisfactory account of evolution. The 'reductionist' view that would explain everything in terms of genes and essentially ignore organisms and species is obviously too simplistic. The models used in theoretical population genetics make all sorts of unrealistic assumptions and more realistic ones might give surprising results. There is still no satisfactory explanation for the prevalence and ecological patterns of sexuality. In spite of much recent progress, many of the traditional problems of phylogenetics remain highly debatable. Much work remains to be done in turning branching phylogenetic diagrams into explanatory historical narratives. The vast majority of extant species remain undescribed and the fossil record consists largely of 'roadkills' strung together with gaps. There are plenty of opportunities for important new discoveries within the Darwinian research tradition.

See also: Adaptation. Adaptive Radiation. Biodiversity, Origin of. Coevolution. Phylogeny. Speciation, Process of. Species, Concepts of. Systematics, Overview

References

Bowler PJ (1990) *Charles Darwin: The Man and his Influence*. UK: Oxford University Press.
Browne J (1995) *Charles Darwin Voyaging*. New York: Knopf.
Browne J (2002) *Charles Darwin: The Power of Place*. New York: Knopf.
Gayon J (1998) *Darwinism's Struggle for Survival: Heredity and the Hypothesis of Natural Selection*. Cambridge, UK: Cambridge University Press.
Ghiselin MT (2009) *Darwin: A Reader's Guide*. San Francisco: California Academy of Sciences.
Goldie P and Ghiselin MT (1997) *The Darwin CD ROM*. San Francisco: Lightbinders.
Herbert S (2005) *Charles Darwin, Geologist*. Ithaca: Cornell University Press.
Hodge MJS and Radick G (2003) *The Cambridge Companion to Darwin*. Cambridge, UK: Cambridge University Press.
Mayr E (1991) *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought*. Cambridge, MA: Harvard University Press.
Mayr E and Provine WB (1980) *The Evolutionary Synthesis: Perspectives on the Unification of Biology*. Cambridge, MA: Harvard University Press.
Tort P (1996) Dictionnaire du Darwwinisme et de l'Évolution, Presses Universitaires de France, Paris.

Defenses, Ecology of

Phyllis D Coley, University of Utah, Salt Lake City, UT, USA
John A Barone, Mississippi State University, Mississippi State, MS, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 11–21, © 2001, Elsevier Inc.

Glossary

Chemical defenses Compounds used by plants to deter or poison herbivores and pathogens.
Constitutive defenses Defenses that are manufactured and maintained by a plant, regardless of whether it has been attacked by an herbivore or pathogen.
Endophyte Fungus or other organisms residing or growing within plant tissues.
Herbivory Damage to plant tissues by herbivores or pathogens.
Induced defenses Plant defenses, including both chemical and physical defenses, that are produced, at least in their final form, only after the plant has been damaged by herbivores or pathogens.
Mutualism Interactions between organisms of different species that increase the fitness of both participants.
Secondary compounds A synonym for chemical defenses in plants; contrasts with chemical compounds used in primary metabolism, such as photosynthesis and cellular respiration.

Diversity and Function of Chemical Defenses

Plants contain a tremendous diversity of chemical compounds that have no function in any aspect of primary metabolism, such as in cellular respiration or photosynthesis. Therefore, these compounds are frequently called "secondary compounds" and were originally thought to be either waste products or storage molecules. In the 1950s, however, it was realized that these compounds were in fact defenses against herbivores and pathogens. There is an impressive variety of forms of these compounds, and we begin with a brief description of the more common types and how they act. For additional details, see Rosenthal and Berenbaum (1991) and Harborne (1988) (**Table 1**).

Phenolic Compounds

Perhaps the most common chemical defenses in plants are composed of phenols. Simple phenols consist of a benzene ring with a hydroxyl group attached, along with varying functional groups. More than 200 different simple phenols have been described, and they have been found to deter feeding of both invertebrate and vertebrate herbivores. Several studies have shown that plants with higher concentrations of phenolics were avoided by species of herbivorous birds (Harborne, 1988). For example, the Canada goose selects plant species that are low in phenolic content, and in the subarctic phenolic resins deter the feeding of ptarmigans and grouse on several plant species. In the latter case, the resins may be effective defenses because they have antimicrobial effects and may disrupt disgestion by microbes in the birds' caeca.

Compounds composed of multiple phenolic units are called tannins, most of which belong to either of two groups. Hydrolyzable tannins have relatively simple structures that are made up of phenolic acids, whereas condensed tannins (also called proanthocyanidins) have more complicated structures and are made from the condensation of hydroxyflavanol units. Condensed tannins are also more common across plant taxonomic groups, occurring in 54% of angiosperm and 74% of gymnosperm genera that have been examined. In contrast, hydrolyzable tannins are found in only 13% of angiosperm genera (all dicots) and are absent from gymnosperms. Although the actual mechanism of tannin action is controversial, tannins may function by binding proteins. As a herbivore chews a leaf, the tannins stored in the leaf vacuole are released and mix with the proteins, binding them and rendering them indigestible. Since proteins are a primary source of nitrogen for herbivores, this chemical reaction makes leaves less nutritious, although it is unclear whether this reaction is effective in the digestive tract of herbivores.

Table 1 Types of chemical defenses in plants

Class of compound	*Activity*	*Examples*
Alkaloids	Variety of mechanisms; often disrupt nervous system function	Caffeine, nicotine, morphine
Nonprotein amino acids	Disrupt proteins when incorporated in herbivores	Canavanine
Phenolics	May bind proteins, reducing the nitrogen available to herbivores	Hydrolyzable tannis and condensed tannins
Saponins	May reduce uptake of sterol from herbivore digestive system	Digitonin
Terpenes	Various; often toxic	Pyrethroids, hormone mimics
Toxic proteins	Variety of mechanisms; highly toxic compounds	Ricin

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00198-2

Nonetheless, evidence suggests that tannins do deter feeding and are even toxic to some herbivores.

Alkaloids

Though a structurally and biochemically diverse group of secondary compounds, all alkaloids have a heterocyclic ring that either contains a nitrogen atom or has one attached to it (Roberts and Wink, 1998). Some of the best known include atropine, the main poison in deadly nightshade (*Atropa belladonna*), quinine, caffeine, cocaine, morphine, and nicotine. It has been estimated that between 15 and 20% of all vascular plants contain alkaloids, and that they are found in one-third of angiosperm families, primarily the dicots. The distribution of alkaloids in relation to climate or geographic distribution remains in dispute, but in general it appears that alkaloids are a more common form of defense in tropical habitats than in the temperate zone. About 16% of temperate species in one survey contained alkaloids versus 35% of tropical species. The mechanisms by which alkaloids affect herbivores (including humans) are highly varied and specific and often effective at small doses, making them effective poisons (*see* Human Uses of Plant Defenses). Many that have been described affect the functioning of the autonomic nervous system by competing with neurotransmitters. An alkaloid found in the genus *Erythrina* (Fabaceae) blocks acetylcholine receptors in animals, whereas caffeine, an alkaloid found in *Coffea arabica* (Rubiaceae), causes the release of calcium from neural receptors. Both types of disruption can lead to the death of susceptible herbivores.

Terpenes

These are an extremely heterogeneous group of chemicals derived from a common metabolic pathway that produces polymers of isoprene units, a five-carbon compound. The isoprene units may then be arranged in a variety of ways, generating a diverse range of nonpolar compounds. Monoterpenes, which have 10 carbon atoms, are often volatile and include pyrethroids from the genus *Chrysanthemum* that are highly toxic to insects and used as commercial insecticides. These neurotoxins disrupt sodium channels in neurons, causing uncoordinated movement and paralysis. Sesquiterpenes, diterpenes, and triterpenes are larger polymers containing 15, 20, and 30 isoprene units, respectively. Terpenes are especially common in pine and fir trees, which secrete the chemicals in resins that reduce insect attack, although monoterpenes can act as feeding attractants for some species of beetles.

Terpenes can be combined with other molecules to form additional types of secondary compounds, for example, sesquiterpene lactones. In this case, a lactone ring (composed of three carbon atoms, with an oxygen and a carbonyl group) is attached to a sesquiterpene. These compounds are found predominantly in the family Asteraceae. Studies have shown that they are toxic to a variety of herbivorous insects and mammals, particularly livestock. Currently, the cause of the toxicity of sesquiterpene lactones is not known, although they cause tissue lesions in mammals and may disrupt hormone function in insects.

Another group of compounds that are chemically related to triterpenes have the remarkable feature that they can act as analogs to hormones found in animals. These analogs do not appear to have any function in plant metabolism, but their similarity to animal hormones is so striking that it is easy to believe that they evolved primarily as a defense against herbivores. Among these plant hormones are mimics of human female sex hormones, and these plant "phytoestrogens" are being used as alternative medicines.

Some plant species also produce analogs that mimic insect hormones. Two critical hormones in insects are juvenile hormone and molting hormone, both of which are necessary for the development of larvae or juveniles. Molting hormone mimics (or phytoecdysones) are common in ferns and gymnosperms (but rare in flowering plants) and, when consumed by insect larvae, disrupt hormonal control of development, in some cases causing death by preventing the insect from shedding its cuticle. Juvabione and other analogs of juvenile hormone are less common than mimics of molting hormone but appear to have similar, fatal effects in insects.

Three other important groups of terpenes are cucurbitacins (found in the family Cucurbitaceae), limonoids (found in Rutaceae, among others), and saponins, which are discussed in the following section.

Saponins

A type of triterpenoid glycoside, saponins are complicated molecules composed of a polycyclic structure (with either 27 or 30 carbon atoms) attached to a carbohydrate group. Consequently, these molecules have both hydrophobic and hydrophilic properties, making them water soluble. Taxonomic surveys suggest that saponins are widespread; one study found that they are present in 500 species from 80 plant families. Saponins have been shown to be toxic to many arthropods, including mites, beetles, and lepidopterans. In insects, saponins may reduce the uptake of sterol from the digestive system. Since sterol is necessary for insects to molt, saponins apparently interfere with insect development. There is little evidence, however, that saponins are toxic to most mammals.

Nonprotein Amino Acids

Many plant species produce amino acids that are not used in making proteins but instead are involved in defenses against natural enemies. More than 600 of these nonprotein amino acids have been identified from a wide range of plant species. For example, canavanine, a structural analog of the amino acid arginine, has been found in 1500 species in the legume family. Fed to sensitive insect herbivores, canavanine causes dramatic and often fatal defects in development. This effect occurs because canavanine becomes incorporated into newly synthesized proteins in place of arginine, thus altering their structure and function. A species that produces canavanine, the woody vine *Dioclea megacarpa*, has seeds that are 8% canavanine by dry weight. Not surprisingly, these seeds are largely free from attack except by one specialized beetle (Rosenthal and Berenbaum, 1991).

Toxic Proteins

Plants also make toxic defensive proteins that produce their effects through a variety of different mechanisms (Kiowa *et al.*, 1997). One such protein, ricin, is produced by the castor bean (*Ricinus communis*) and functions by inhibiting ribosomes in herbivores. By weight, it is among the most toxic compounds known. In general, there is considerable interest in identifying genes that code for toxic proteins in wild plants. If such genes could then be used to modify agricultural crops, they would be protected from herbivores without the use of chemical insecticides.

Physical and Structural Defenses

Spines

In addition to chemical defenses, plants also employ a wide range of physical or structural features that either reduce the accessibility of plant tissue to herbivores or deter herbivore feeding. The most apparent of these are spines or thorns on plants. There is little doubt that the primary function of these structures is as a defense against herbivorous animals, particularly vertebrates, although in some cases spines may help plants to thermoregulate or, in the case of some vines, to climb.

Trichomes

In some ways resembling small spines, trichomes are structures found on the leaves and stems of many plant species that inhibit attack by herbivores and pathogens, either physically or chemically. The shapes and sizes of trichomes vary considerably across species from slender hairs to stout spikes, and many are armed with recurved barbs. Furthermore, glandular trichomes of many species produce or store chemicals that act as deterrents to herbivores. There are countless examples from plants in both natural and agricultural settings of how trichomes can kill or deter herbivores (Levin, 1973). Acting as physical barriers, trichomes can reduce egg laying by herbivores on some varieties of wheat and increase plant resistance to leaf hoppers and mites in some varieties of cotton. The chemicals secreted by many trichomes include sticky gums that immobilize small insects, as found in some varieties of wild potato and tomato. On other species trichomes secrete chemicals such as alkaloids and terpenes as well as waxes, fatty acids, and alcohols. These chemicals are frequently toxic to herbivores, for example, nicotine (an alkaloid) secreted by tobacco plants (*Nicotiana* in the Solanaceae) that kills aphids on contact. Thus, trichomes not only act as physical defenses but also provide a way for plants to use chemical defenses against their natural enemies without first having to suffer tissue damage.

Toughness

Leaves with thick cell walls, consisting of lignin and fiber, are tough. In one assessment of both physical and chemical defenses of mature leaves in tropical forests, leaf toughness and fiber content were the two most important factors in reducing herbivory. Toughness is an effective defense against insect herbivores, particularly small or immature ones, because they have a difficult time cutting or chewing the leaves. Furthermore, cell walls are largely indigestible; therefore, tough leaves are nutritionally poor.

Biotic Defenses and the Third Trophic Level

In addition to using chemical and physical defenses against herbivores and pathogens, many plant species have evolved complex, mutualistic interactions with other groups of organisms that act as a type of defense against herbivores. In such interactions, the plants typically provide food, shelter, or both, whereas the other organisms defend the plant from its natural enemies.

Ants

The best studied defensive mutualisms in plants are those with ants (Huxley and Cutler, 1991). Given the incredible abundance of ants in most habitats and their propensity to forage on leaves and stems, it is not surprising that these mutualisms are so common. Ant – plant interactions vary from loose, facultative associations in which the plant offers rewards to any ants in its vicinity to more tightly coevolved relationships in which both partners display highly specialized traits. In facultative associations, a plant offers nectar, food bodies, and other rewards to lure ants that nest elsewhere to patrol its leaves and remove any herbivores they encounter. These associations are especially common in tropical forests.

There are also many obligate ant – plant mutualisms, in which one or both participants require the other to survive. The best known is the interaction between the tree *Acacia cornigera* (Fabaceae) and the ant *Pseudomyrmex ferruginea* (Pseudomyrmecinae) in Central America. The plants in this pair have numerous adaptations for playing host to the ants, including swollen thorns on which the ants have their colonies, enlarged nectaries on the leaves that provide the ants with sugars, and modified leaflet tips on which the ants can feed. The young queen ant establishes a colony by landing on a young tree and setting up a nest in one of the swollen thorns. After the colony is established, the worker ants patrol the tree constantly, removing any other insects on the tree and cutting away any other plants that touch it. In one study, when the ants were experimentally removed the plants suffered high rates of defoliation, leading to a reduction in growth and a doubling of mortality during the 11 months of the study (55 vs 28% for control plants with ants). Thus, there is little doubt that ants can play an important role in protecting plants in these tightly obligate mutualisms.

The clear defensive effect of ants on their host plants has been demonstrated for other species as well. Less easily measured is the cost of the ants as a defense: Presumably the benefits outweigh the costs, but what are they? Recent studies show that *Cecropia* trees produce different types of food rewards for their ants in relation to the nutrients available—in effect producing rewards that use the less limiting nutrients. Circumstantial

evidence also suggests that the ants are a costly defense since on islands on which native ants are rare or absent, few plant species have either extrafloral nectaries or food bodies.

Domatia, Mites, and Other Predators

The mutualistic relationship between ants and plants is apparently paralleled by similar but looser associations between plants and small arthropods. Many plant species have structures on their leaves called domatia that can serve as homes for mites and small, predaceous insects. Although there are a variety of forms, most domatia are pits, pockets, or tufts of hairs on the undersides of leaves. Domatia are widespread taxonomically among dicots, occurring in 28% of 290 plant families, and they are quite common, at least in some forests. Recent work has supported the hypothesis that domatia provide shelter to arthropods, which in turn attack herbivores on the plant. For example, when artificial domatia were placed on the leaves of cotton plants, the density of predaceous bugs increased, the number of herbivores decreased, and the number of fruit produced increased by 30%. Although preliminary, this work suggests that domatia are a part of loose mutualisms used by plants to defend themselves.

Endophytic Fungi

In addition to using animals as a means of defense against herbivores and pathogens, there is considerable evidence that plants can employ fungi in a similar manner. Most plants are infected asymptomatically by endophytic fungi—that is, fungi that live within the plant, usually between the plant cells. These endophytes are a potential defense since they produce chemical compounds not made by plants.

Although endophytes are common in an impressive diversity of plant species, most research has focused on their relationship with grasses. In fact, their defensive role in plants was only appreciated after it was discovered that cattle suffered toxic symptoms after eating tall fescue grass infected with a fungal endophyte. Since then, grasses infected with fungal endophytes have been shown to be toxic to other domestic and wild mammalian herbivores as well as herbivorous insects. At the community level, research has shown that pastures dominated by endophyte-infected grasses have lower populations of small mammals. Thus, it seems likely that plants can gain some protection against herbivores from fungal endophytes.

Predators and Parasitoids as Plant Defenses

Plant defenses have evolved within ecological communities in which predators and parasitoids that attack their herbivores are often common. Evidence suggests that some plant species may directly or indirectly elicit the help of these predators and parasitoids as a form of defense. Recent research has found that some plants release volatile compounds when damaged, and that these are used as location cues by parasitoids and predators (Turlings and Benrey, 1998). Frequently, these volatile chemicals are emitted only when damage is accompanied by oral secretions of herbivores, meaning that the signals are specifically released when a herbivore is present and not when the plants incur other types of damage. By facilitating the discovery of herbivores by their predators and parasites, plants may enjoy reduced herbivory.

Other defenses, such as tannins and toughness, do not kill herbivores outright but reduce their growth and prolong the time until pupation. As a result, herbivores may actually consume more leaf tissue than if the plants lacked these defenses. This seems paradoxical, but by slowing growth these defenses increase the chance that herbivores will be preyed on while still in the earlier instars (Benrey and Denno, 1997). Since the majority of leaf damage occurs in the final instar, this could greatly reduce herbivory. Thus, the effectiveness of some defenses may depend critically on help from parasitoids and predators.

Phenological Strategies

New, expanding leaves are generally more vulnerable to attack by herbivores and pathogens than mature leaves because they are tender and nutritious. In addition to the defenses already described, many plant species reduce new leaf damage by altering the timing of new leaf production. These phenological strategies are of two general types. In the first, plants produce new leaves during times of the year when herbivores and pathogens are rare. In temperate forests, the spring flush of leaves occurs during the short window of opportunity between the end of winter and the recovery of herbivore populations in the late spring. Similarly, studies conducted in tropical forests in India, Ghana, and Panama have shown that when plants produce new leaves during the dry season, when herbivores are rare, herbivory is reduced. Obviously, this seasonal escape strategy would not work in nonseasonal climates, and because of coincident changes in light and water availability with season it is difficult to demonstrate conclusively that the timing of leaf production is mainly an adaptation to reduce herbivore and pathogen damage.

The second phenological strategy to decrease damage to new leaves is for the plants of a given species to produce new leaves synchronously. By flushing simultaneously, plants may be able to overwhelm their specialist herbivores with an abundance of new leaves so that the chance that any particular new leaf will be discovered and eaten is decreased. One test of this strategy involved following new leaf production and herbivory in a moist tropical forest for a year. The results showed that highly synchronous species experienced lowered herbivory than those that were continuous in production. This phenological strategy may be more important in less seasonal forests in which a seasonal escape strategy may be less effective.

Plant Investments in Defenses

Assumptions

In their evaluations of plant defenses, ecologists typically make the following assumptions: (i) Herbivory and other tissue damage by natural enemies is bad for plants, reducing their fitness; (ii) defenses reduce damage by natural enemies;

and (iii) defenses cost the plants in terms of energy and nutrients.

Surprisingly, the first assumption about the negative effect of herbivory on plants has been controversial. Several studies have shown that some plant species, in certain circumstances, are able to "overcompensate" for tissue damage, achieving higher fitness than those that were never damaged. However, in general the evidence is quite clear that herbivory, for most plant species, reduces fitness. The billions of dollars spent annually throughout the world on pesticides to protect crops are a testament to the negative effects of herbivores and pathogens. Experimental studies have also demonstrated that damage reduces plant fitness. In a study by Marquis, for example, shrubs of the species *Piper arieianum* (Piperaceae) were artificially defoliated and their growth and seed production followed for 2 years. The results showed that individuals that had lost at least 30% of their leaf area suffered a 50% decrease in growth and produced approximately half the seeds of control plants. Clearly, damage affects plant fitness.

Costs and Benefits of Defenses

Plant defenses may be costly to plants for several reasons (Fritz and Sims, 1992). First, in order to invest in defenses, plants must divert resources from other aspects of growth and reproduction. For example, nitrogen that a plant invests in alkaloids as a defense cannot be used in making proteins necessary for greater photosynthesis. A second reason why defenses may be costly to plants is that some chemical defenses, although effective deterrents against natural enemies, may also be toxic to the plant, so the plant has to expend additional energy protecting itself from its own defenses. Tannins must be sequestered in the vacuole to avoid cell damage, and many terpenes and resins are restricted to specialized canals in plants.

Third, a defense that reduces plant damage from one herbivore or pathogen species may inadvertently make the plant more attractive to another. For example, Eisner and coworkers showed that the trichomes on the leaves and stems of *Mentzelia pumila* (Loasaceae) frequently ensnare a variety of insects, including many damaging herbivores. However, the trichomes do not affect a damaging aphid species, whereas they kill one of its predators. Thus, the trichome defense protects the plant from some herbivores but makes it more vulnerable to others.

Despite a variety of direct and indirect costs of defense, plants generally benefit from being defended from herbivores and pathogens. One instance was already discussed: When the ants were removed from ant-defended *Acacias*, herbivory and plant mortality increased. Other studies of chemical defenses have shown the benefits of chemical defenses as well as their costs. One such study was conducted by Berenbaum and associates, who examined whether wild parsnips benefited from being chemically defended in the presence of the herbivore parsnip webworms. They showed that plants with greater chemical defenses had higher fitness in the presence of the webworms, meaning that the defenses had a clear benefit. The chemicals were costly, however, because in the absence of herbivores in the greenhouse plants that were better defended chemically had lower seed production. Under natural conditions, however, the benefits of reduced herbivore damage outweighed the costs to these plants.

Induced Defenses

The chemical defenses described so far are usually constitutive—that is, they are produced and maintained regardless of whether herbivores or pathogens have damaged the plant. Many secondary compounds, however, are induced, with production (or at least the final stages of production) only occurring after the plant has been attacked (Karban and Baldwin, 1997). Usually, an aspect of damage, such as partially eaten cell walls, leads to a transduction process that causes the cell, tissue, or whole plant to begin synthesis of defensive compounds. This entire process can take place in less than 1 hr. In field and laboratory conditions, plants that have been induced suffer less herbivory and have higher fitness than controls. However, most species do not appear to have inducible defenses.

Plants should have induced defenses instead of constitutive ones if defenses are costly to plants, and energy and nutrients that are allocated for secondary compounds or other compounds cannot be used for growth or reproduction. Therefore, by producing defenses only when they are needed—during an attack by herbivores or pathogens—the plant is able to divert these resources to growth and reproduction. However, most species do not appear to have inducible defenses, suggesting that plants may frequently not be in the position to predict when it would be advantageous to induce defenses.

Theories of Plant Defense

Both across and within habitats, the amount of damage that plants suffer from herbivores and pathogens varies enormously. In a tropical forest, for instance, rates of herbivory to mature leaves can vary more than two orders of magnitude, from species that receive virtually no damage to those that lose more than 0.6% of their leaf area per day. This huge range implies that plants differ in both the types and amount they invest in defenses.

Much research has focused on trying to correlate defensive investment with other plant life history traits and habitat preferences. It is generally accepted that short-lived, fast-growing species, which typically are found in resource-rich habitats, are not as well defended as slower growing, longer lived woody species. However, the underlying reasons for these patterns are not understood, and numerous theories have been proposed. Two important elements appear in many of these theories: value and risk. First, selection should favor high investments in defense if the risk of herbivory is high. Second, effective defenses should be favored in tissues with a high value or replacement cost.

Apparency theory was the first influential attempt to provide a theoretical framework for plant defenses (Feeny, 1976). It emphasized differences in the risk of herbivory, arguing that defense investment should depend on how apparent plants are to herbivores. Species that are ephemeral, such as annuals, may not be readily apparent to herbivores, whereas long-lived

trees and shrubs should be readily found. Thus, unapparent plants should be able to largely escape specialist herbivores and, as a result, should invest less in defenses and use defenses that are toxic at low concentrations to generalist herbivores. These "qualitative" defenses include low-molecular-weight compounds such as alkaloids, cyanogenic glycosides, and monoterpenes. In contrast, apparent plants are certain to be discovered by host-specialist herbivores and should produce leaves that are generally less palatable. The theory predicts that they should invest in "quantitative" defenses, such as tannins and toughness, that are not poisonous per se but have dosage-dependent effects, making leaves harder to digest. One test of this prediction found that herbaceous plants had a greater abundance of toxic alkaloids than more apparent, woody plants. Several authors have also suggested that the risk of discovery by herbivores may depend on the frequency of other more palatable species in the community.

Other theories, although acknowledging that herbivore pressure is important, also suggest that the value of the plant tissue should influence the evolution of defense. For example, species adapted to resource-poor environments, such as those with low light or poor soils, have inherently slow growth rates and long leaf lifetimes. The resource availability hypothesis (Coley *et al.*, 1985) argues that slower growing species should invest more in defenses than fast-growing ones because the cost of replacing damaged leaves is higher when resources are limited. Moreover, in slow growers, the decrement of lost growth for a given investment in defenses is small because the species are already growing slowly. However, in a fast growing species, the investment in defenses leads to a higher percentage decrease in plant growth. This means that the relative or opportunity cost of plant defenses is greater for fast-growing species than for slow growers. Thus, for a given abundance of natural enemies, plant species adapted for growth in high resource environments should maximize their growth by investing less in defense and suffering greater herbivory. The optimal level of defense for species adapted to low resource environments should be much higher.

The resource availability theory also predicts the general types of defenses that plants should use. Species with long-lived leaves should invest in defensive compounds that, despite high initial costs due to their high concentrations, have low turnover rates. These defenses include fibers, which increase leaf toughness, and tannins, both of which would be quantitative defenses in the plant apparency theory. In fast-growing species with short-lived leaves, these defenses would not be profitable because the leaves would be dropped before the plants would have an opportunity to recoup their investment. Instead, the hypothesis predicts that species with short-lived leaves should invest in compounds that are effective at low concentrations even though they may have high turnover rates. The combination of low concentration (or low incremental cost) and high turnover rates means that these defenses are relatively less costly to plants that have short-lived leaves, especially because the components of these defenses can be withdrawn from leaves before they senesce. These defenses would be very costly for species with long-lived leaves since they would have to be continually regenerating these defenses over a long period of time.

Plasticity in Defenses

The theories of plant defenses presented in the last section attempt to explain why selection would favor different amounts and types of defenses in different species. An underlying assumption is that defense strategies will be optimized for the predominant conditions experienced by each species. However, there is variation around this optimum, depending on environmental perturbations. In general, if a particular resource is available in excess of normal growth requirements, the extra is shunted into making defenses (Bryant *et al.*, 1983). For example, if a plant is fertilized but retained in low light, it will shunt the extra nitrogen into nitrogen-containing defenses such as alkaloids. Of course, this occurs only in species that have the ability to make alkaloids or other nitrogen-based defenses. In contrast, if a plant is placed in a high light environment, the high rates of photosynthesis will lead to an excess of carbon in the form of starch. This extra carbon will be used to produce higher levels of carbon-based defenses such as tannins and terpenes. Note that these patterns are the opposite of the interspecific trends in which high resources select for lower defense (see Section V,D). Thus, shifts in the ratio of carbon and nitrogen available to an individual plant will cause phenotypic changes in defense allocation patterns that do not necessarily reflect the optimal trends seen across species.

Human Uses of Plant Defenses

In their daily lives, humans have long made use of plant defenses, especially plant secondary compounds. This use has been one focus of ethnobotany, which is the general study of the relationship between humans and plants. Here, we briefly review a few examples of how these plant chemicals have been employed by people. More detailed information can be found in Johns (1996).

Hunting and Fishing

Undoubtedly, one of the first uses of plant secondary compounds by humans was as poisons on the tips of arrows. As described by Neuwinger (as cited in Roberts and Wink, 1998), such poisons were formerly used with arrows throughout the world (except perhaps in Australia and New Zealand) and are still used in some areas, both in hunting and in warfare. Poisons for hunting must be highly toxic in small quantities and either easy to process during hunting trips or persistent while on arrow tips so they do not have to be replaced often. For these reasons, cardiac glycosides and alkaloids are the most common arrow poisons.

Cardiac glycosides, which are secondary compounds common in the families Asclepiadaceae and Apocynaceae, were mostly used as arrow poisons in Africa. They were often the primary active ingredient in concoctions of several different plant extracts, with the other plants added to enhance the effectiveness of the poison, to help it adhere better to the arrows, or to fulfill a magical purpose. In northern South America curare poisons were used in blowgun darts. These

were typically derived from plants in the genus *Strychnos* (Loganiaceae; also used in Africa) or from species in the Menispermaceae. Chemically, curare poisons all contain quaternary alkaloids, albeit with different structures. Once in the bloodstream, these block the transmission of neural impulses to the skeletal muscles by competing for receptors with the neurotransmitter acetylcholine. Progressive paralysis of the skeletal muscles follows, eventually leading to respiratory failure in the animal (Roberts and Wink, 1998).

In many traditional societies, poisons from plant secondary compounds have also been used in fishing. These poisons are normally put into small ponds or slow-moving streams, in which the concentration of poison will be high enough to kill or at least stun the fish. A variety of compounds have been used as fish poisons, including isoflavonoids, saponins, cardiac glycosides, alkaloids, tannins, and cyanogenic compounds. Some of the better known are flavonoids derived from the tropical legumes in the genera *Derris*, *Lonchocarpus*, *Mundulea*, and *Tephrosia*.

Medicines

As has been noted many times, the difference between a poison and a medicine is often only a matter of dosage, and this is certainly true of products from plant secondary compounds. One example is a curare poison derived from the vine *Chondrodendron tomentosum* that contains the alkaloid tubocurarine. Although it works like other curare poisons in paralyzing muscles, it was introduced to Western medicine in 1939 and is now used as a muscle relaxant in anesthesia.

It would be difficult to exaggerate the importance of plant secondary compounds in the history of medicine, both traditional and Western. It has been estimated that 25% of all prescriptions written in North America are from plant-derived drugs, and the majority of medicines in traditional cultures are derived from plants. However, only about one-half of 1% of the approximately 265,000 flowering plant species have been comprehensively evaluated for biological active compounds, the first step in determining whether a chemical can be effective as a drug. Of course, most plants do not contain medically useful compounds; therefore, an important task is to determine how to screen species in a more effective manner than simply collecting and testing plants at random. One approach, pioneered by ethnobotanists, has been to document the plants used by healers in traditional cultures. In a preliminary study, plants used by a healer in Belize produced four times as many positive results in laboratory assays than did species collected at random. A second approach is to collect plants based on ecological information. For example, as noted previously, an important defense of mature leaves in tropical forests in toughness, resulting from increased lignin and cellulose in the leaves. However, new, expanding leaves cannot be defended by toughness and so are more likely to use chemical defenses. For this reason, one way to increase the effectiveness of plant screenings for drugs may be to focus on young leaves instead of mature ones.

Regardless of the screening process, plant secondary compounds have been an important source of drugs. In particular, alkaloids have provided many noteworthy medicines in both traditional and Western societies (Schmeller and Wink as cited in Roberts and Wink, 1998). These include atropine, codeine, colchicine, ephedrine, morphine, reserpine, taxol, and theobromine. Codeine, for example, is derived from the fruit of *Paper somniferum* (Paperveraceae) and is a component of opium. Currently, more than 200 pharmaceuticals contain codeine, and it is used as a cough suppressant, a pain killer, and a sedative. In the West, ephedrine is used as a nasal decongestant in cold medicines and is a treatment for asthma in both Western and traditional societies. Its source is *Ephedra sinica* and *E. shunnungiania* (Ephedraceae), and it works in part by raising blood pressure and respiration while causing an opening of air passages. Taxol is a recently discovered diterpene alkaloid in the tree *Taxus brevifolia* (Taxaceae) that has therapeutic use as a treatment for breast and ovarian cancer. This compound disrupts microtubule assembly in cells, making it an effective weapon against rapidly dividing tumor cells.

Self-Medication in Animals

Mounting evidence suggests that humans are not the only animals that use plant chemical defenses as medicines. Naturalists have long observed that some animals occasionally eat or rub leaves or other plant parts on themselves in an apparent attempt to cure either diseases or kill off parasites (Clayton and Wolfe, 1993). Demonstrating that this behavior is truly self-medication requires showing that the behavior is deliberate, that the plant substance used kills parasites or disease-causing organisms, and that the self-medication leads to increased fitness in the animal. Since in many cases the behavior is rare, gathering evidence to meet these criteria is difficult, but some instances are suggestive. For example, kodiak bears chew the roots of the *Ligusticum* spp. and then rub their saliva through their fur. Since this species is also used by humans as a medicine, it is possible that the bears are also using the plant to cure or ward off infections. Other vertebrates, including chimpanzees, white-faced monkeys, and birds, also use leaves of particular plant species in unusual ways or only when they are sick, suggesting that they may have medicinal value. However, it remains difficult to demonstrate conclusively that animals are truly using plant defenses as medicines.

Spices in Foods

In addition to their uses in treatment of diseases, plant secondary compounds may also be used to prevent illness. Billing and Sherman (1998) argued that people use spices in food preparation to inhibit or kill pathogenic microorganisms. They marshaled several lines of evidence to support this "antimicrobial hypothesis." First, they noted that numerous studies have demonstrated that many spices, such as onions, garlic, pepper, and chili peppers, can inhibit the growth or even kill many species of bacteria. Second, by comparing recipes of traditional cuisines from 36 countries, they found that both the proportion of recipes calling for spices and the number of spices per recipe were positively correlated with the mean annual temperature of the country. In warmer climates food should spoil more quickly, they suggested, so spices may be an important way to preserve food. Third, in warmer

countries the spices that are commonly used in cooking more strongly inhibit bacterial growth, again reflecting the greater chance of food going bad. For example, despite its strongly antimicrobial properties, garlic is not part of any traditional dishes in Norway, but it is in 80% of the recipes from Indonesia. Billing and Sherman also noted that spices are used in small quantities and thus the alternative explanation that they are used for nutritional reasons seems unlikely. Although other factors may come into play, including the greater diversity of plants in equatorial countries and the relatively slow incorporation of newly available spices into traditional cuisines, the correlations observed by Billing and Sherman strongly support their hypothesis.

Summary and Conclusions

Ultimately, the theme of this article, like that of this book, is biodiversity. Plants have evolved a stunning diversity of defenses and strategies to reduce damage by herbivores and pathogens. The most notable of these are chemical defenses, which plants use either to deter feeding (such as tannins) or to poison their natural enemies (such as nonprotein amino acids.) Plant defenses also include physical defenses such as thorns, and many species have evolved elaborate mutualistic relationships with ants, mites, and fungi as an additional means of defense. All defenses have some cost, and a defense will only be favored by natural selection if its benefits in reduced damage and higher fitness outweigh the costs of producing and maintaining the defense. Why particular species invest what they do in defenses, and why they invest in some defenses and not others, remains a subject of controversy, but there are patterns across habitats and successional seres, making it clear that the evolution of plant defenses correlates with at least some abiotic factors. Finally, humans (and perhaps other animals) make use of plant defenses in their everyday lives. In food and with medicine, the lives of people have been greatly enhanced by the diverse wealth of plant defenses. Understanding how to make a fuller, more effective use of the diversity of plant defenses found in the natural world is an important, practical goal emerging from this type of research.

See also: Adaptation. Parasitism

References

Benrey B and Denno RF (1997) Slow-growth—high-mortality hypothesis: A test using the cabbage butterfly. *Ecology* 78: 87–99.

Billing J and Sherman PW (1998) Antimicrobial functions of spices: Why some like it hot. *Q. Rev. Biol.* 73: 3–49.

Bryant JP, Chapin FS III, and Klein DR (1983) Carbon/nutrient balance of boreal plants in relation to vertebrate herbivory. *Oikos* 40: 357–368.

Clayton DH and Wolfe ND (1993) Adaptive significance of self-medication. *Trends Ecol. Evol.* 8: 60–63.

Feeny P (1976) Plant apparency and chemical defense. *Recent Adv. Phytochem.* 10: 1–40.

Fritz RS and Sims EL (1992) *Plant Resistance to Herbivores and Pathogens: Ecology, Evolution and Genetics*. Chicago: Univ. of Chicago Press.

Harborne JB (1988) *Introduction to Ecological Chemistry*. San Diego: Academic Press.

Huxley CR and Cutler DF (1991) *Ant – Plant Interactions*. Oxford: Oxford Univ. Press.

Johns T (1996) *Origins of Human Diet and Medicine*. Tucson: Univ. of Arizona Press.

Karban R and Baldwin IT (1997) *Induced Responses to Herbivory*. Chicago: Univ. of Chicago Press.

Kiowa H, Bressan RA, and Hasegawa PM (1997) Regulation of protease inhibitors and plant defense. *Trends Plant Sci.* 2: 379–384.

Levin DA (1973) Role of trichomes in plant defense. *Q. Rev. Biol.* 48: 3–15.

Roberts MF and Wink M (1998) *Alkaloids: Biochemistry, Ecology and Medicinal Applications*. New York: Plenum.

Rosenthal GA and Berenbaum MR (1991) *Herbivores: Their Interactions with Secondary Plant Metabolites*. San Diego: Academic Press.

Turlings TCJ and Benrey B (1998) Effects of plant metabolites on the behavior and development of parasitic wasps. *Ecoscience* 5: 321–333.

Defining, Measuring, and Partitioning Species Diversity

Hanna Tuomisto, University of Turku, Turku, Finland

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alpha diversity The mean effective density of types (such as species) expressed per compositional unit in a dataset. See eqn [15].

Beta diversity The effective number of compositional units in a dataset. See eqn [17].

Compositional unit A virtual subunit that contains as many types (such as species) as the actual subunits do on average, but does not share types with the other virtual subunits. See **Figure 2**.

Diversity The effective number of types (such as species) into which the entities (such as individuals) that form the dataset of interest have been classified. Also known as "true diversity" and "Hill number." See **Figure 1**.

Effective species A virtual species that has the same proportional abundance as the actual species do on average. The number of effective species, or effective number of species, is species diversity. See **Figure 1**.

Evenness The effective number of types expressed as a proportion of the actual number of types in the dataset. See eqn [9].

Gamma diversity Total effective number of types (such as species) in a dataset. See eqn [10].

Generalized mean with exponent $q-1$ A general equation that allows deriving different kinds of mean (e.g., arithmetic and geometric) by varying the parameter q. See eqn [2].

Gini–Simpson index An index that quantifies the probability that two entities (such as individuals) drawn at random from the dataset represent different types (such as species). Often used as a diversity index, but not to be confused with diversity itself. See the explanation after eqn [6].

Proportional abundance The abundance (such as the number of individuals) of one type (such as species) expressed as a proportion of total abundance (number of individuals) of all types (species) in the dataset.

Richness The actual number of types (such as species) into which the entities (such as individuals) that form the dataset of interest have been classified.

Shannon entropy An index that quantifies the uncertainty regarding which type (such as species) is picked when one entity (such as individual) is drawn at random from the dataset of interest. Often used as a diversity index, but not to be confused with diversity itself. See eqn [4].

Simpson index An index that quantifies the probability that two entities (such as individuals) drawn at random from the dataset represent the same type (such as species). Often used as a diversity index, but not to be confused with diversity itself. See the explanation after eqn [6].

Species turnover A measure of the amount by which the species composition changes among subunits in the dataset.

Introduction

Confusion around the concept "diversity" has long hampered accurate communication among researchers. For example, Hurlbert (1971) lamented: "The term 'species diversity' has been defined in such various and disparate ways that it now conveys no information other than 'something to do with community structure'; species diversity has become a nonconcept." Hill (1973) argued for a logically coherent definition of diversity, which implies making a distinction between diversity itself and different kinds of diversity indices. Unfortunately, ecologists have continued to treat diversity indices as if they measured the same thing as diversity itself does, which makes much of the diversity literature very difficult to understand and interpret in a coherent way.

In recent years, awareness of the problems caused by inconsistent and often contradictory use of terminology has increased. Since Jost (2006, 2007) revived Hill's definition of diversity, it has become possible to appreciate that the concept of diversity is actually quite simple (Tuomisto, 2010a, b, c, 2011). The simplicity emerges from three main sources. Firstly, the concepts become less ambiguous when an explicit difference is made between diversity itself and diversity indices. An index of something is just a surrogate for the thing itself, and for an index to be useful, it needs to be clear what it is an index of. This necessitates an unambiguous definition of diversity itself. Secondly, Hill's definition of diversity is understandable in simple, everyday terms, which makes diversity a much less abstract concept than ecologists have often thought. Thirdly, separating the conceptual problem of defining diversity from the practical problem of deciding how to adequately quantify diversity for a community of interest allows a clearer view on the issue. As a result, it has become a feasible goal to reach a consistent terminology for diversity and related concepts.

Before getting into the details, it is worth noting that diversity can be quantified for any dataset for which observed entities have been classified into types. The entities are often individuals, but with organisms whose individuals are difficult to delimit (such as clonal plants or colonial animals), the entities may also be something else, such as ramets, colonies, or units of cover or biomass.

If species diversity is of interest, the types into which the entities are classified are species. In other cases, the entities may be classified into, for example, genera, families, haplotypes, or functional types. To calculate the diversity of the

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00378-6

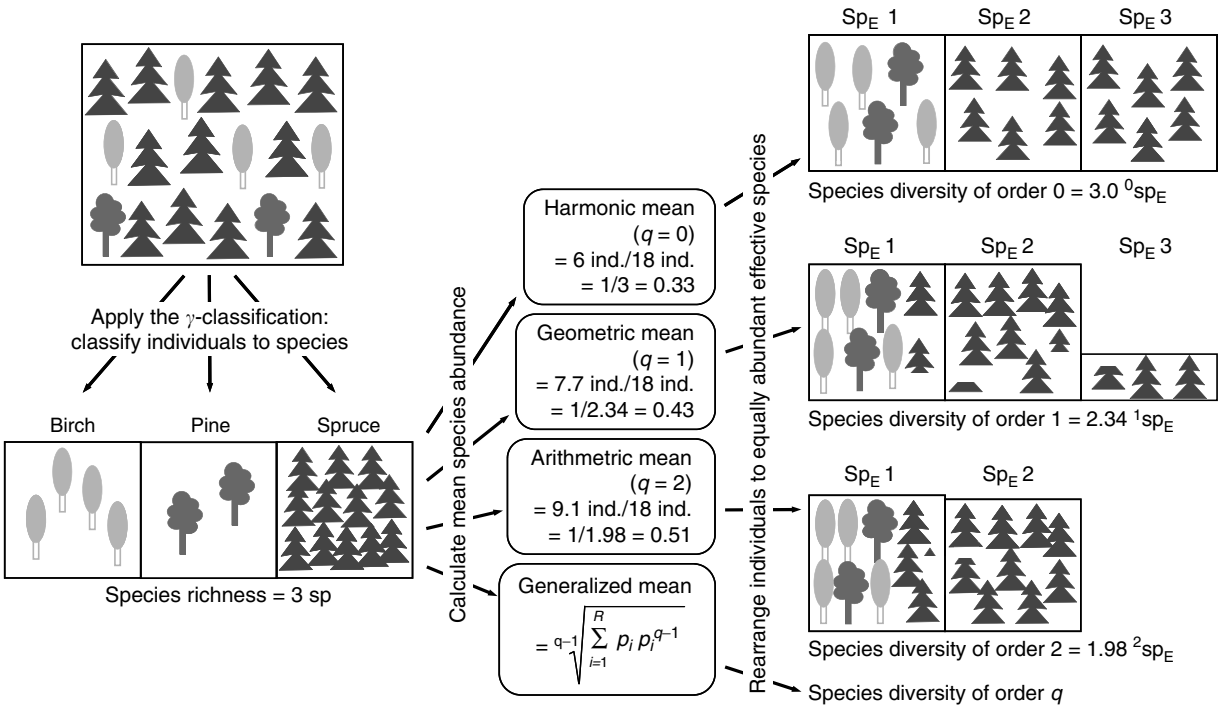


Figure 1 The calculation of species richness and species diversity in a small dataset consisting of a single tree plot. Species richness is the number of boxes (the set at left) needed to place all individuals into a box with an appropriate species name. Species diversity is the number of new boxes needed (the sets at right) to place all individuals into a box such that each box receives as many individuals (or as large a proportion of all individuals) as the named species have, on average (only a part of the last new box may be needed). The measure of "average" here is the weighted generalized mean with exponent $q-1$. This can be calculated by first taking the weighted mean of the absolute species abundances (6, 7.7, and 9.1 individuals for $q=0$, 1, and 2, respectively) and then dividing this by total abundance (18 individuals). Equally well, one can calculate the weighted mean of the proportional abundances directly (as shown in the equation of the generalized mean; R is the number of actual named species and p_i is the proportional abundance of the ith species). Species diversity ${}^qD$ equals the inverse of mean p_i, and it is the effective number of species (the number of equally abundant species that would give the observed mean species abundance). The measurement unit is hence actual species (sp) in the case of richness and effective species (sp_E) in the case of diversity. Reproduced from of Tuomisto H (2011) Commentary: Do we have a consistent terminology for species diversity? Yes, if we choose to use it. Oecologia 167: 903–911, with permission from Springer.

dataset, all that is needed is knowledge of the proportional abundance of each of the relevant types within the dataset. Species richness is related to species diversity, but is not the same thing: it does not take the proportional abundances into account. In this article, the letter D will be used for diversity and the letter R for richness. The traditional notation S refers specifically to species richness and is not used here to avoid restricting the discussion to this specific kind of richness. For simplicity, this article is nonetheless written in terms of species richness and species diversity, but the same logic applies to any types or categories of interest.

Defining Diversity

Species richness (R) of a dataset can be defined simply as the number of species names in the species list that is obtained after every individual (or other entity) that belongs to the dataset has been identified to species. This can be called the actual number of species. Species diversity (D) is a more complicated concept, because it takes into account not only how many species the dataset contains but also their relative abundances. This can be done in many different ways, which has led to a variety of definitions for the concept "diversity" (reviewed in Peet, 1974; Magurran, 2004; Jost, 2006, 2007; Gauthier *et al.*, 2010; Tuomisto, 2010a). In the interest of accurate scientific communication, "species diversity" is here defined as the effective number of species, following Hill (1973).

This definition of diversity D can be derived as follows. Let us first define the proportional abundance of species i in the dataset to equal $p_i=m_i/m$, where m_i is the absolute abundance of species i and m is the total abundance of all species in the dataset. For simplicity, abundance is here treated as the number of individuals, but the same principles also apply when other units of abundance are used (e.g., grams of biomass or square meters of cover). When all species are equally abundant, then $m=m_iR$, and the proportional abundance p_i of each species equals $m_i/(m_iR)$, which simplifies to the inverse of richness $1/R$. The mean proportional species abundance then also equals $1/R$. When all species are not equally abundant, some of them have a larger and others a smaller proportional abundance, but the mean of the proportional abundances is nevertheless a single value $\overline{p}_i$, which can also be expressed as $\overline{m}_i/(\overline{m}_iD)=1/D$. Species diversity D can hereby be defined as the number of equally abundant species that would be needed

to obtain the same mean proportional species abundance as that observed in the dataset of interest. This is the effective number of species (MacArthur, 1964, 1965; Hill, 1973; Routledge, 1979; **Figure 1**).

A mean can be calculated in several different ways, which has implications for the quantification of diversity. The different kinds of mean yield different values when applied to the same dataset, so the diversities based on them also obtain different values. Hill's (1973) definition of diversity can be written in the general form

$$ {}^{q}D = 1/{}^{q}\overline{p}_i \qquad [1] $$

Here ${}^{q}D$ is diversity of order q, and ${}^{q}\overline{p}_i$ is the corresponding mean of the proportional species abundances. The mean here is the weighted generalized mean with exponent $q-1$, which is calculated as

$$ {}^{q}\overline{p}_i = \sqrt[q-1]{\sum_{i=1}^{R} p_i p_i^{q-1}} \qquad [2] $$

When calculating the mean, each p_i value is nominally weighted (i.e., multiplied) by itself, but the exponent $q-1$ modifies the weighting. When $q=1$, the actual weights equal the nominal weights, which implies that each species contributes to the value of the mean in proportion to the amount of data (e.g., number of individuals) it contributed to the dataset. In other words, each individual has the same weight no matter which species it belongs to: each individual pulls the mean toward the p_i value of its own species by exactly the same weight irrespective of its species identity. Smaller values of q ($q<1$) deemphasize the abundance differences among species. This gives rare species more weight than implied by their p_i, and individuals of rare species hence pull the mean toward the p_i value of their own species with more weight than individuals of abundant species do. Conversely, larger values ($q>1$) exaggerate the abundance differences among species and hence give abundant species (and individuals belonging to them) disproportionate weight. Specific values of q correspond to familiar kinds of means (Hill, 1973):

$q \to -\infty$	minimum p_i value
$q=0$	harmonic mean $= 1/\sum_{i=1}^{R} p_i/p_i$
$q=1$	geometric mean $= \prod_{i=1}^{R} p_i^{p_i}$
$q=2$	arithmetic mean $= \sum_{i=1}^{R} p_i p_i$
$q \to \infty$	maximum p_i value

In practice, the parameter q determines where the mean proportional species abundance lands within the possible range between the minimum and the maximum p_i observed in the dataset (see **Figure 1** for a numerical example).

By combining eqns [1] and [2], the definition of diversity can be written:

$$ {}^{q}D = 1/\sqrt[q-1]{\sum_{i=1}^{R} p_i p_i^{q-1}} \qquad [3] $$

If all p_i values are identical in a dataset, the number of effective species (${}^{q}D$) in that dataset equals the number of actual species (R) no matter which value of q is chosen. If different species have different p_i values in a dataset, increasing q causes mean p_i to increase and hence diversity ${}^{q}D$ to decrease for that dataset. At $q=0$, the number of effective species (${}^{0}D$) equals the number of actual species (R) even when all species are not equally abundant. At $q=1$, diversity equals the exponential of the Shannon index, and at $q=2$, diversity equals the inverse of the Simpson index (these diversity indices will be discussed in more detail in the Diversity Indices section). Negative values of q cause the number of effective species to exceed the number of actual species (${}^{q}D>R$), and positive values of q cause the number of actual species to exceed the number of effective species ($R>{}^{q}D$).

In diversity studies, q is usually limited to nonnegative values. This is because increasing the weight given to the rarest species makes the mean proportional species abundance more dependent on the total size of the dataset. The rarest species in a dataset are often singletons (represented by a single individual), and their proportional abundance is $1/m$. As q approaches negative infinity, the mean of the proportional abundances approaches this minimum value, and the effective number of species therefore approaches m, that is, the total number of individuals in the dataset. It is difficult to think of practical applications where this property would be desirable in a diversity measure, and setting the restriction $q \geq 0$ ensures it does not cause problems of interpretation.

Obviously, ${}^{q}D$ can obtain very different values depending on the chosen value of the parameter q. However, the conceptual interpretation of ${}^{q}D$ does not change with q: it always quantifies the effective number of species (or other types of interest) in the dataset. Depending on the questions at hand, it might be appropriate to choose just a single kind of mean to calculate ${}^{q}D$, or several kinds of mean could be used in parallel (by choosing different values for the parameter q) to explore how the effective number of species varies when rare versus abundant species are emphasized in different ways (e.g., Ricotta and Avena, 2002; Kindt *et al.*, 2006; Tuomisto, in press).

The concept of diversity as an effective number of types was apparently first used in ecology by MacArthur (1964, 1965), but it was more thoroughly discussed by Hill (1973). Owing to Hill's influential paper on the topic, the diversity values have been known as "Hill numbers" (or "Hill's numbers") in the ecological literature. Although some researchers adopted Hill numbers relatively early (e.g., Routledge, 1977, 1979), they have generally been treated as just another diversity index (e.g., Magurran, 2004). However, Jost (2006, 2007, 2009) has argued that the effective number of types should be used as the only definition of diversity and proposed to call it "true diversity" to distinguish it from diversity indices that represent phenomena other than a count of effective species. This narrow definition of "diversity" is gaining in influence, because it considerably clarifies communication about diversity issues (e.g., Gregorius and Gillet, 2008; de Bello *et al.*, 2010; Gauthier *et al.*, 2010; Gregorius, 2010; Tuomisto, 2010a, b, c, 2011; Jurasinski and Koch, 2011; Moreno and Rodríguez, 2011).

It can be noted that the use of the word "true" in "true diversity" is analogous to its use in "true bugs" (Tuomisto,

2011). In both cases, the word is used as a synonym of "in the narrow sense." In the case of bugs, "true" indicates that the interest is in insects of the monophyletic order Heteroptera, rather than in the polyphyletic assemblage consisting of any insect-like animals and pathogenic microorganisms, which are also called "bugs" in everyday contexts. In the case of diversity, "true" indicates that the interest is in the effective number of types, rather than in the conceptually heterogeneous group of phenomena consisting of numbers of types, entropies, probabilities, and differences, which have also been called "diversity" in the ecological literature. Gregorius (2010) used the term "explicit diversity" instead of "true diversity," possibly because the latter has raised negative reactions as some researchers have misinterpreted "true" to mean "best for any purpose" (Hoffmann and Hoffmann, 2008; Anderson *et al.*, 2011; Gorelick, 2011).

The greater the largest p_i values in a given dataset, the larger the mean p_i and hence the smaller the number of equally abundant species that would be needed to obtain the same mean p_i value. These hypothetical, equally abundant species can be called effective species ($^{q}\mathrm{sp}_{\mathrm{E}}$), and they are the unit of measurement of species diversity ^{q}D (note the use of the superscript q to specify which mean is used). In contrast, species richness R is measured in units of actual species (sp; Tuomisto, 2010a, c). When the individuals (or other observed entities) are classified into types other than species (such as genera, families, functional types, haplotypes, or size classes), the measurement units of richness and diversity change accordingly (e.g., actual and effective genera, respectively).

Both richness and diversity follow the important replication principle: if each of the original species is replicated to form r different species, each of which has the same absolute abundance as the original, both richness and diversity of the dataset increase to r times their original values (Hill, 1973; Jost, 2006, 2007). The replication principle corresponds to an intuitive sense of how diversity is expected to behave. The traditional diversity indices (such as the Shannon and Gini–Simpson indices) do not follow the replication principle, which is one reason why their values are often misinterpreted (Jost, 2006, 2007, 2009).

Practical Issues in Quantifying Diversity

Often ecologists are interested in the species richness or diversity of a community, but are able to sample only a tiny fraction of the individuals belonging to that community. Since any one of the unseen individuals could represent a species not recorded in the sample, the species richness of the community is generally (and often considerably) underestimated by the recorded data. Therefore, richness and diversity values that are accurate for an existing dataset provide no more than estimates for the community of interest. How accurate such estimates are depends on how representative the dataset is of the community. In general, the more species a community contains, the more individuals need to be observed for the sample to become representative. Different kinds of sampling and data analysis methods can be used to improve the accuracy of the estimate, just as in other situations where results are extrapolated beyond the actual observations (e.g., Colwell and Coddington, 1994; Plotkin and Muller-Landau, 2002; Chao *et al.*, 2005; Beck and Schwanghart, 2010; Gauthier *et al.*, 2010).

Much of the discussion in the past on how to define diversity has actually focused on how to accurately estimate diversity for a community of interest from a sample of that community. For example, it has often been recommended that diversity is best measured using diversity indices that stabilize at small sample sizes (e.g., Lande, 1996; Magurran, 2004; Beck and Schwanghart, 2010). By this criterion, species richness is a very bad index of diversity, the Shannon index clearly better, and the Simpson index superior to both. However, the conceptual definition of a term should not depend on context-specific details such as sampling decisions, the ease of measurement, or the questions of interest in a particular study. Consider observing a mixed-species flock of birds. How the concept of "species diversity" is defined should not depend on how difficult it is to decide which individuals belong to the flock and which species each individual belongs to. Similarly, how the concept of "volume" is defined should not depend on how difficult it is to decide where the physical limits of the flock are. Of course, such decisions determine whether the measurement result is appropriate and sufficiently accurate for the intended purpose, but this is a separate issue that has no bearing on which phenomenon is represented by the variable to be measured (whether diversity or volume).

In this article, the focus is on conceptual definitions, so richness and diversity will be treated simply as properties of an existing dataset (see also Smith and Wilson, 1996; Gregorius and Gillet, 2008; Tuomisto, 2010a, b, c). Before diversity and richness can be quantified, a decision has to be made on the limits of the dataset, that is, which individuals (or other entities) belong to it. Every different set of individuals in effect corresponds to a different dataset and consequently may yield different richness and diversity values.

Each individual that is added to a dataset may represent a species that was not yet present in the dataset before. Therefore, drawing meaningful ecological conclusions from comparisons of richness and diversity values among datasets necessitates that the methods by which individuals have been selected for the datasets are comparable, which in practice implies that sampling efforts must have been standardized across the datasets in an appropriate way (Gotelli and Colwell, 2001; Gauthier *et al.*, 2010; Tuomisto, 2010b). For example, bird inventories might be based on making observations in a fixed observation point either for a predefined amount of time or until a predefined number of individuals has been recorded. Tree inventories could be based on recording all stems of a predefined minimum diameter in a plot of a predefined shape and surface area, or on recording a predefined number of stems using some plotless method. Inventory of marine benthic fauna could be based on animals captured by sieving cores of a specific volume through a mesh of a specific size.

If the communities of interest vary widely in species diversity, standardizing relative sampling effort (representativeness or coverage) may sometimes be more important than standardizing absolute sampling effort (e.g., number of individuals). A relevant measure of sampling effort in such cases would pay attention to species accumulation curves or the

proportion of singletons in the data (Good and Toulmin, 1956).

If all individuals within a fixed surface area have not actually been observed and included in the dataset, then surface area cannot be used as a measure of sampling effort. For example, the nominal sampling unit size in macroecological studies may be fixed (e.g., 50 km by 50 km grid cells), but the actual sampling effort invested in each sampling unit (e.g., time spent in the field or number of individuals observed) can vary by several orders of magnitude among grid cells. In such cases, the observed patterns in richness and diversity may depend more on the uneven distribution of field work effort than on any ecologically interesting phenomenon.

In addition to defining the limits of the dataset, it is important to define how abundance is measured. Many different ways are possible and justifiable. Possible measures include the number of individuals, the number of stems or ramets, biomass, basal area, volume, or the percentage of surface cover. Obviously, if the size of individuals varies (as it generally does), the number of individuals and biomass can be expected to give different diversity values even when calculated for the same dataset, and diversity values based on different abundance measures are therefore not comparable across datasets.

Both richness and diversity also depend on how the individuals (or other observed entities) are classified into types. A lumper and a splitter will rarely report the same species richness for a given dataset, and family richness can change dramatically depending on which family circumscriptions are used. Consequently, interpretation of results for richness and diversity requires careful specification of the classification that was used. Comparing richness or diversity values among datasets is meaningful only if all of them are based on the same classification. For example, if one dataset has identified tropical trees to the species level and another only to the genus level, comparing richness and diversity values between them yields meaningless and misleading ecological conclusions.

Diversity Indices

Shannon Index

Several indices of diversity have been popular in ecology (reviewed in Magurran, 2004; Gauthier *et al.*, 2010). Until recently, too little attention has been paid to the fundamental conceptual differences among those indices. Consequently, the term "diversity" has been used in several conceptually different ways in the ecological literature, as indices of diversity have been equated with diversity itself (Jost, 2006, 2007; Tuomisto, 2010a, c).

The most popular diversity index by far in the ecological literature has been the Shannon entropy H' (Shannon, 1948):

$$H' = -\sum_{i=1}^{R} p_i \log\ p_i = \log({}^1D) \qquad [4]$$

This measure is also known as the Shannon index, the Shannon–Weaver index, the Shannon–Wiener index, the Shannon–Weiner index, and (confusingly) the Shannon diversity. The Shannon entropy does not quantify species diversity (the effective number of species) but the uncertainty in the species identity of an individual that is picked at random from the dataset. The measurement unit of Shannon entropy depends on the base of the logarithm and can be, for example, bits, nats, or decits when using log bases 2, *e*, and 10, respectively. Taking the exponential of the Shannon entropy recovers true diversity; that is, ${}^1D = \exp(H')$.

The Shannon entropy is a special case for $q = 1$ of the Rényi entropy (Rényi, 1960; Tuomisto, 2010a)

$${}^qH' = \log({}^qD) = \log(1/{}^q\overline{p}_i) \qquad [5]$$

Here the value of q defines how the probability of choosing a particular individual is affected by the abundance of the species it belongs to. When $q = 1$, each individual has the same probability of being chosen, and hence the probability that the chosen individual represents species i equals p_i. When $q = 0$, each species has the same probability of being chosen irrespective of its proportional abundance. The probability that the chosen individual belongs to the most abundant species of the dataset increases monotonically with increasing q. With sufficiently large q, the chosen individual almost certainly belongs to the single most abundant species in the dataset, which has proportional abundance p_{imax}. The uncertainty in the species identity of the randomly chosen individual (Rényi entropy) hence decreases asymptotically toward $\log(1/p_{\text{imax}})$ as q increases toward infinity. This parallels the behavior of the mean proportional species abundance, which increases toward the abundance of the most abundant species as q increases (see explanation after eqn [2]).

Simpson Index and Gini–Simpson Index

Equation (3) is usually written in the form

$${}^qD = \left(\sum_{i=1}^{R} p_i^q\right)^{1/(1-q)} \qquad [6]$$

The term inside the parentheses (${}^qD^{1-q}$), known as the "basic sum," has often been used as a diversity index. When $q = 2$, the basic sum equals the Simpson index $\sum_{i=1}^{R} p_i^2$ (Simpson, 1949; Hill, 1973; Jost, 2006, 2007; Tuomisto, 2010a), which quantifies the probability that two individuals picked at random from the dataset (with replacement) represent the same species. Hurlbert (1971) called this the probability of interspecific encounter (PIE, which can also be calculated, in a slightly modified form, to account for not replacing the first individual before choosing the second). The Simpson index decreases as the species richness of a dataset increases, so in itself it is an impractical index of diversity. A popular transformation of the Simpson index that increases with increasing richness is the Gini–Simpson index, which is obtained by subtracting the Simpson index from unity. The Gini–Simpson index quantifies the probability that two individuals picked at random from the dataset (with replacement) represent different species.

Another transformation of the Simpson index that also increases with increasing richness is its inverse. Although less popular as a diversity index than the Gini–Simpson index, the inverse Simpson index actually equals ${}^2D$, that is, true diversity of order 2. In other words, the Simpson index itself represents

the weighted arithmetic mean of the proportional species abundances, rather than an effective number of species.

The Gini–Simpson index is a special case of an entropy that is known in ecology as the Tsallis or HCDT entropy (Jost, 2006; Mendes *et al.*, 2008). This entropy can be expressed as

$$ {}^{q}T = (1 - {}^{q}D^{1-q})/(q-1) \quad [7] $$

The HCDT entropy is a simple transformation of the basic sum, and its value is related to the probability that q individuals chosen at random from the dataset (with replacement) represent at least two different species.

From eqns [3] and [6] it can be seen that the basic sum is just an intermediate step in calculating mean proportional species abundance and its inverse, true diversity. Therefore, using the basic sum to derive an index of diversity is unnecessary; diversity itself is just as easy to calculate as the diversity index, and the risk of misinterpreting the results is diminished if the variable that is reported represents the actual phenomenon of interest rather than a surrogate (or transformation) thereof (Tuomisto, 2010c).

Partitioning Diversity

Multiplicative Partitioning: Richness × Evenness

Species diversity is traditionally thought to consist of two independent components, richness and evenness (e.g., Peet, 1974; Alatalo, 1981; Smith and Wilson, 1996; Magurran, 2004; Kindt *et al.*, 2006; Mendes *et al.*, 2008; Jost, 2010). One way to combine components that represent different phenomena is by using multiplication, which in this case leads to the following partitioning of diversity (Jost, 2010; alternative evenness concepts are reviewed in Tuomisto, in press):

$$ \text{diversity} = \text{richness} \times \text{evenness} \quad [8] $$

Since richness and diversity were already defined in the Defining Diversity section, evenness can be defined as

$$ {}^{q}E = {}^{q}D/R \quad [9] $$

Because richness is based on presence–absence data and hence does not depend on the proportional species abundances p_i, it is not sensitive to the value of q. In contrast, diversity and evenness decrease when the value of q is increased, unless all species (or other types of interest) are equally abundant in the dataset of interest. Consequently, the value of q must always be specified to allow accurate communication about diversity and evenness.

In practice, ${}^{q}E$ quantifies how many effective types there are for each actual type in the dataset of interest. Species richness is based on presence–absence (binary) data and is measured in units of actual species; species diversity is based on abundance (quantitative) data and is measured in units of effective species. Consequently, the measurement unit of evenness is ${}^{q}\text{sp}_\text{E}/\text{sp}$. The highest possible species evenness ${}^{q}E_{\max}$ equals $1\ {}^{q}\text{sp}_\text{E}/\text{sp}$. This is obtained when there are as many effective species as actual ones, that is, when either all species are equally abundant or $q=0$. When $q>0$, diversity ${}^{q}D$ decreases toward $1\ {}^{q}\text{sp}_\text{E}$ as the proportional abundance of the most abundant species increases toward unity. As a result, the smallest possible value of evenness ${}^{q}E_{\min}$ is $(1\ {}^{q}\text{sp}_\text{E})/R$.

For any given dataset, small values of q lead to higher evenness than large values of q. This follows directly from the fact that diversity decreases when q increases, but richness does not change (*see* Defining Diversity). Different kinds of species abundance distribution lead to different trajectories of diversity ${}^{q}D$, and hence of evenness ${}^{q}E$, when these are plotted as a function of q (Ricotta and Avena, 2002; Kindt *et al.*, 2006; Jost, 2010; Tuomisto, in press). Graphs showing ${}^{q}E$ and ${}^{q}D$ against q provide complementary insights for comparing datasets that differ in R.

Multiplicative Partitioning: Alpha × Beta

The General Principle

When Whittaker (1960) coined the terms alpha, beta, and gamma diversity, his intention was to understand the species diversity in a landscape (gamma diversity) as the combined result of two different phenomena, namely, the species diversity at a more local scale (alpha diversity) and the compositional heterogeneity among localities (beta diversity). Whittaker explored several approaches to quantifying compositional heterogeneity, and quite some confusion has resulted because he referred to all of them as "beta diversity." Other researchers have further expanded the circumscription of beta diversity, such that the term has been used to refer to more than 30 distinct phenomena in the ecological literature (e.g., Wilson and Shmida, 1984; Vellend, 2001; Koleff *et al.*, 2003; Jurasinski *et al.*, 2009; Tuomisto, 2010a, b; Anderson *et al.*, 2011).

Some of these "beta diversities" are not mathematically derived from alpha and gamma diversity in any way, and the values of many are uncorrelated with each other. Although most of the resulting measures are useful in themselves, each one of them quantifies a different phenomenon and is therefore relevant for addressing different ecological questions (conceptual differences among the measures are reviewed in Tuomisto, 2010a, b). Confusion among these concepts – because they are referred to by the same name – has caused many unwarranted comparisons among studies and incorrect inferences from data.

Whittaker (1960) did write that the simplest definition of beta diversity is $\beta=\gamma/\alpha$. This corresponds to the multiplicative partitioning of gamma diversity $\gamma=\alpha_\text{d}\beta_\text{Md}$. Subscript "M" is added to specify that the components come from a multiplicative partitioning, and subscript "d" to indicate that their measurement units correspond to the types whose diversity in the dataset is being quantified (notation as in Tuomisto, 2010a). This classical partitioning of diversity has been one of the most widely used definitions of beta diversity in ecological studies, because it yields alpha and beta diversities that are conceptually and numerically independent of each other (e.g., Routledge, 1977, 1979; Vellend, 2001; Jost, 2007; de Bello *et al.*, 2010; Tuomisto, 2010a, c, 2011).

Gamma diversity is conceptually the simplest one of the three diversities: it is the total species diversity observed in the dataset of interest. The dataset may or may not represent a landscape, but whether it does is irrelevant for the definition of the concepts. To quantify gamma diversity, only one

classification within the dataset of interest is needed, namely, the classification of its individuals into species. This classification was already used in the Defining Diversity section, and it can be called the gamma classification (Tuomisto, 2010a). Gamma diversity can hence be defined as the diversity in relation to the gamma classification and identified by adding the subscript γ to eqn [1]:

$$ {}^{q}D_{\gamma} = 1/{}^{q}\overline{p}_{i} \qquad [10] $$

The richness in relation to the gamma classification, that is, the total number of actual species in the dataset of interest, can be called gamma richness R_{γ}. The symbols α, β, and γ have been used to refer to both the diversity components and the corresponding richness components, and indeed it may be useful to continue using them in this broad sense, when the alpha, beta, and gamma components are discussed in general terms without specifying whether the interest is in diversity or in richness. However, the distinction between diversity and richness becomes relevant as soon as their values are to be quantified. At this point, more specific notation should be used, such as ${}^{q}D_{\alpha}$, ${}^{q}D_{\beta}$, and ${}^{q}D_{\gamma}$ for the diversity components, and R_{α}, R_{β}, and R_{γ} for the richness components.

Gamma diversity is the total number of effective species in a dataset, and gamma richness is the total number of actual species in the dataset. Both are independent of whether the dataset has been subdivided into smaller units, and if so, how this has been done. In contrast, the alpha and beta components are defined only when the dataset consists of subunits. These subunits may or may not be the same sampling units that were used during the field inventory. It is possible to subdivide a single dataset into subunits in many different ways, and each different subdivision corresponds to different alpha and beta values.

When quantifying alpha and beta diversity or richnesses, each individual is classified not only according to species but also according to subunit. The classification of individuals into subunits has been called the "omega classification" (Tuomisto, 2010a). The richness in relation to the omega classification (omega richness) is the actual number of subunits in the dataset. The diversity in relation to the omega classification (omega diversity) is the effective number of subunits in the dataset (the number of subunits the dataset would get divided into if each subunit received the same number of individuals as the actual subunits have, on average). Alpha diversity or richness is obtained by taking into account both the gamma and the omega classifications simultaneously, but in such a way that the primary interest is in the gamma classification. In other words, each individual is classified both into a species and into a subunit, and the interest is in quantifying the density of species (actual or effective) per subunit (Tuomisto, 2010a, c, 2011).

Exactly the same idea has been expressed by Gregorius (2010) using a different notation. He characterized individuals by two attributes, type (T) and subcollection affiliation (S). The classification of individuals into types corresponds to the gamma classification, and the classification of individuals into subcollections corresponds to the omega classification. Gregorius (2010) used the notation ν_T for gamma diversity (total species diversity in the dataset) and ν_S for omega diversity (total subunit diversity in the dataset). The earliest definition of gamma diversity in these terms may be that of Routledge (1979), who used the notation $N_{a\gamma}$ (where a has the same meaning as q in ${}^{q}D_{\gamma}$).

Alpha, Beta, and Gamma Richness

The richness components are quantified without taking into account species proportional abundances. Alpha richness is simply calculated as the unweighted arithmetic mean of the actual species density values in the subunits

$$ R_{\alpha} = \sum_{j=1}^{N} \frac{1}{N} R_{\alpha j} \qquad [11] $$

Here N is the total number of subunits and $R_{\alpha j}$ is the actual species density (species richness per subunit) in subunit j. The multiplicative partitioning leads to defining beta richness as

$$ R_{\beta} = R_{\gamma}/R_{\alpha} \qquad [12] $$

This can be interpreted as the number of actual compositional units (CU) in the dataset. The actual (as opposed to effective; see the next section) CU are obtained by rearranging the R_{γ} actual species of the dataset evenly into new hypothetical subunits such that each subunit receives R_{α} species. In other words, each compositional unit has the same species density as the original subunits do on average, but does not share species with any other compositional unit. Alpha richness equals the number of actual species per actual compositional unit, which gives it the measurement unit sp/CU (**Figure 2**; Tuomisto, 2010c, 2011).

Alpha, Beta, and Gamma Diversities

To quantify the diversity components, species proportional abundances need to be taken into account. The proportional abundance of species i that is conditional on the limits of subunit j is $p_{i|j} = m_{ij}/m_j$, where m_{ij} is the number of individuals belonging to species i in subunit j and m_j is the total number of individuals in subunit j. The weighted generalized mean of all these within-subunit proportional abundances is

$$ {}^{q}\overline{p}_{i|j} = \sqrt[q-1]{\sum_{j=1}^{N}\sum_{i=1}^{R} p_{ij} p_{i|j}^{q-1}} \qquad [13] $$

Here N is the number of subunits . The proportion of data that individuals of species i in subunit j contribute to the entire dataset is p_{ij} and this is used as the nominal weight when calculating mean within-subunit proportional abundance. The same mean $p_{i|j}$ can also be expressed as (see Tuomisto, 2010a for derivation)

$$ {}^{q}\overline{p}_{i|j} = \sqrt[q-1]{\sum_{j=1}^{N} w_j \sum_{i=1}^{R} p_{i|j} p_{i|j}^{q-1}} \qquad [14] $$

Here the subunit weight $w_j = m_j/m$ equals the proportion of the individuals in the entire dataset contributed by subunit j. Alpha diversity is obtained as the inverse of this mean:

$$ {}^{q}D_{\alpha} = 1/{}^{q}\overline{p}_{i|j} \qquad [15] $$

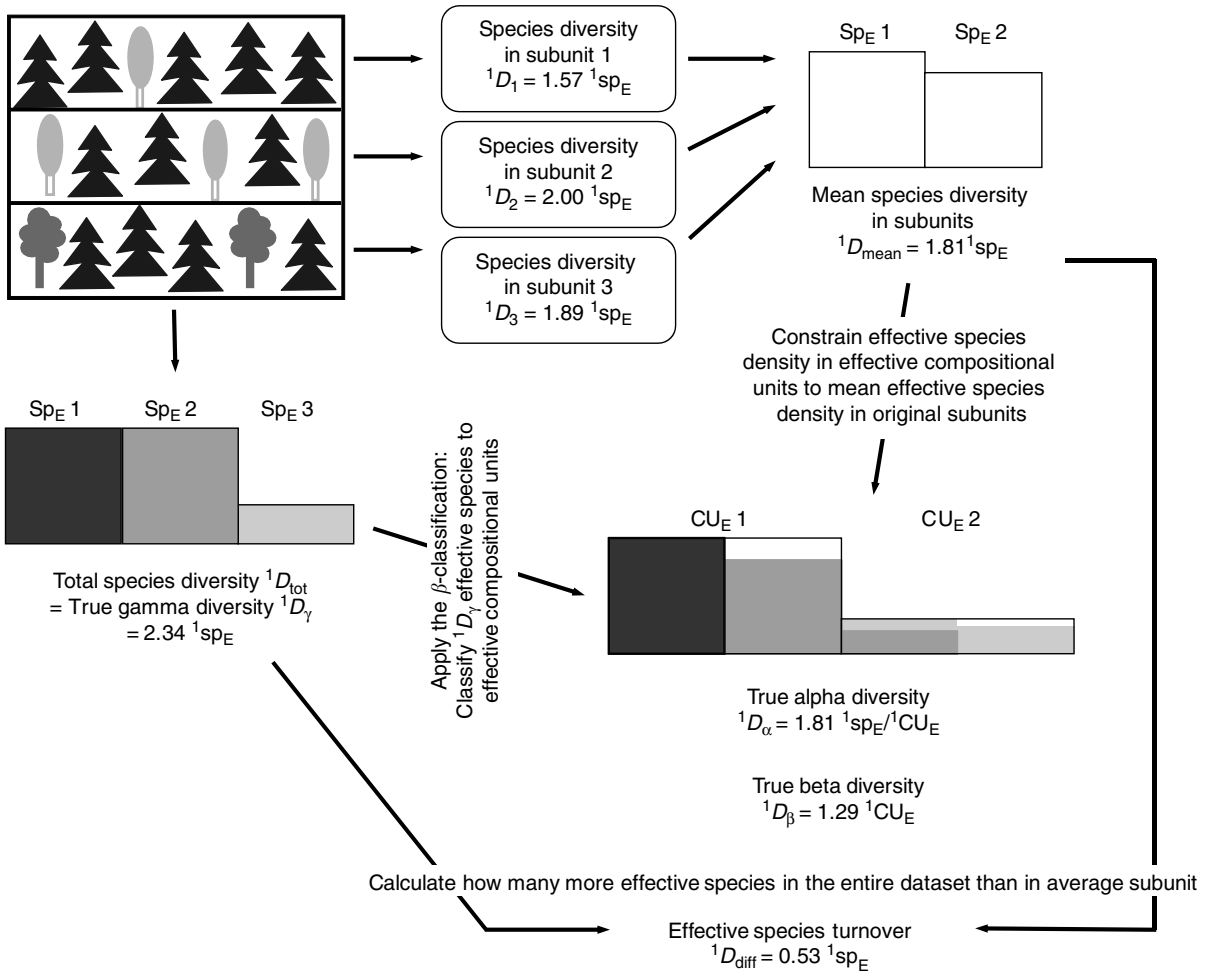


Figure 2 Two different ways of partitioning the total species diversity of a dataset (true gamma diversity) into two components, given three subunits. The dataset is the same as in **Figure 1**, but only true diversities of order 1 (which are based on the geometric mean of species proportional abundances) are shown for simplicity. Species diversity in each subunit j (^1D_j) is calculated in the same way as species diversity in the entire dataset ($^1D_\gamma = {}^1D_{tot}$; see **Figure 1**). Multiplicative partitioning gives two conceptually and mathematically independent components, namely the effective number of compositionally distinct units (true beta diversity; measurement unit ^1CU_E) and the mean effective species density in those units (true alpha diversity). Additive partitioning gives two components that represent the same concept as true gamma diversity does, namely species diversity. One of these components quantifies the mean within-subunit species diversity ($^1D_{mean}$), and the other ($^1D_{diff}$) quantifies the total number of effective species that differ (turn over) among all subunits. Reproduced from Figure 2 of Tuomisto H (2011) Commentary: Do we have a consistent terminology for species diversity? Yes, if we choose to use it. Oecologia 167: 903–911, with permission from Springer Science + Business Media.

This can also be expressed as (see Proof 2 in Tuomisto, 2010a for derivation)

$$^qD_\alpha = \sqrt[1-q]{\sum_{j=1}^{N} w_j (^qD_{\alpha j})^{1-q}} \qquad [16]$$

This formula expresses the weighted generalized mean with exponent $1-q$ of the mean effective species densities $^qD_{\alpha j}$ in all subunits j. In the notation of Routledge (1979), alpha diversity is $N_{a\alpha}$, and in the notation of Gregorius (2010), it is $v_{T|S}$.

The alpha diversity of Jost (2006, 2007) differs from the alpha diversity defined by eqn [16] in that Jost treated the gamma and the omega classification symmetrically, whereas eqn [16] considers the gamma classification of primary interest. The differences between the alpha diversity concepts of Routledge and Jost have been discussed by Tuomisto (2010a) and Gregorius (2010). Because Routledge's definition corresponds better with what seemed to be Whittaker's original idea behind diversity partitioning, it is followed throughout this article.

The multiplicative partitioning leads to defining beta diversity as

$$^qD_\beta = {}^qD_\gamma / {}^qD_\alpha \qquad [17]$$

This ratio can be interpreted as the number of effective compositional units (^qCU_E). An effective compositional unit is a hypothetical subunit that has the same number (and density) of effective species as the original subunits do on average, but

does not share effective species with the other effective compositional units. Alpha diversity equals the mean density of effective species as expressed per effective compositional unit, which gives it the measurement unit ${}^{q}\mathrm{sp_E}/{}^{q}\mathrm{CU_E}$ (**Figure 2**; Tuomisto, 2010a, c, 2011). In the notation of Routledge (1979), beta diversity is $N_{a\beta}$ and in the notation of Gregorius (2010), it is N_{eS}.

Properties of Beta Richness and Beta Diversity

In any dataset, there must be at least one compositional unit (both actual and effective), so $R_\beta \geq 1$ CU and ${}^{q}D_\beta \geq 1\ {}^{q}\mathrm{CU_E}$. Similarly, the total number of species (both actual and effective) in the dataset must be at least as large as the mean number of species per subunit, so $R_\gamma \geq R_\alpha$ and ${}^{q}D_\gamma \geq {}^{q}D_\alpha$. Some researchers have complained that the latter condition does not always hold (Gadagkar, 1989; Lande, 1996). However, this is an erroneous assertion based on averaging the effective species densities ${}^{q}D_{\alpha j}$ using the arithmetic mean, which corresponds to $q=0$, even if some other value of q is used in the calculations otherwise. Using the generalized mean with exponent $1-q$ (eqn [16]) correctly gives alpha diversity values that do not exceed gamma diversity for any value of q.

If all species occur in all subunits, beta richness necessarily equals 1 CU. The more the original subunits differ in species composition, the more compositional units are needed at a given species density to accumulate the given total number of species. Beta richness obtains its maximum value of N when none of the actual subunits share any species.

Because beta diversity depends both on species composition and on species abundances, its behavior is more complex than that of beta richness. When the proportional abundances of the species vary among subunits, ${}^{q}D_\beta$ may exceed $1\ {}^{q}\mathrm{CU_E}$ even if all species occur in all subunits. Similarly, when no species are shared among subunits, ${}^{q}D_\beta$ may be either smaller or larger than N. Beta diversity is constrained to the fixed range of $[1, N]$ only when either $q=1$ or the nominal weights w_j in eqn [16] are equal for all subunits j (Gregorius, 2010; Tuomisto, 2010a). Nominal subunit weights are equal when, for example, sampling effort has been standardized to the same number of individuals in all subunits, or when abundances are recorded as percentages that sum to 100% in each subunit.

Increasing the value of q makes ${}^{q}D_\beta$ more sensitive to the variation among subunits in the proportional abundances of species and less sensitive to variation in species composition. Therefore, changing within-subunit species proportional abundances without changing presence–absence patterns has no effect on ${}^{0}D_\beta$, and changing species composition without changing the proportional abundance of the single most abundant species has no effect on ${}^{\infty}D_\beta$ (Tuomisto, 2010a).

Beta richness and beta diversity can be thought of as the richness and diversity in relation to the beta classification. Here the entities of interest are species (actual or effective, respectively) that are classified into compositional units (actual or effective, respectively; compare with the classification of individuals into actual or effective species in **Figure 1**). Alpha richness and alpha diversity, in turn, correspond to the density of richness and diversity in relation to the gamma classification that is conditional on the limits set by the beta classification. This combined classification can be called the alpha classification.

Whittaker (1977) referred to α and γ as inventory diversity and to β as differentiation diversity, which has since become a common practice (e.g., Magurran, 2004; Jurasinski *et al.*, 2009). Some researchers have even argued that beta diversity should not be called diversity at all (e.g., Lande, 1996; Kiflawi and Spencer, 2004; Gregorius and Gillet, 2008). However, true beta diversity ${}^{q}D_\beta$ is the effective number of compositional units, which conforms with the general definition of diversity just as well as the effective number of species does.

Alpha, beta, and gamma diversities are conceptually similar in that each quantifies the effective number of types in a dataset according to a specific kind of classification of the entities that constitute that dataset. The three differ in two ways. Firstly, they classify different entities: individuals in alpha and gamma diversities versus effective species in beta diversity. Secondly, they quantify the relevant diversity at a different level: as a total diversity in the entire dataset in beta and gamma diversities versus a mean density of diversity as expressed per effective compositional unit. This conceptual similarity to ${}^{q}D_\alpha$ and especially ${}^{q}D_\gamma$ justifies singling ${}^{q}D_\beta$ out as the sole measure of true beta diversity and recommending that all other things that have been referred to as "beta diversity" in the past be called something else in the future (Jost, 2006, 2007; Tuomisto, 2010a, b, c, 2011).

Additive Partitioning: Alpha + Beta

Absolute Species Turnover

Lande (1996) proposed defining beta diversity as the difference rather than the ratio of gamma to alpha diversity, which leads to the equation

$$\beta_{\mathrm{At}} = \gamma - \alpha_{\mathrm{t}} \qquad [18]$$

This approach corresponds to partitioning gamma diversity into additive components $\gamma = \alpha_{\mathrm{t}} + \beta_{\mathrm{At}}$. The subscript "A" specifies that the beta component arises from additive partitioning, which is important because it differs from the beta component of the multiplicative partitioning β_{Md} both conceptually and in its numerical value. The subscript "t" specifies that the components relate to species turnover, which affects their measurement units: all components in the additive partitioning have the same measurement unit, namely species (actual species in the case of richness and effective species in the case of diversity).

Lande (1996) borrowed the additive approach from MacArthur (1964, 1965), who had used it with Shannon entropies. However, MacArthur applied the exponential function to the entropies before interpreting the results, and therefore he actually followed Whittaker's multiplicative approach with true diversities at $q=1$. In contrast, Lande used and interpreted the raw values of different diversity indices, and therefore the components obtained from his partitioning correspond to conceptually different phenomena for each different diversity index (see Tuomisto, 2010a for more detailed interpretations and discussion). The interpretations provided in this article are all based on the assumption that any diversity indices are converted to the corresponding

effective numbers of types before being subjected to additive partitioning.

The additive alpha component expresses how many species there are, on average, in one subunit, and the additive beta component expresses how many of the species in the dataset do not fit within the limits of the average subunit. In other words, both components quantify a count of species (i.e., the types based on the gamma classification), but they do so for different subsets of the data (an average subunit vs. the rest). In contrast, both multiplicative components (α_d and β_{Md}) give a measure that describes the entire dataset, but they differ in that each is based on a different classification (the alpha or the beta classification, respectively; Tuomisto, 2010a, c). In addition, α_d is conceptually a density measure (density of types per subunit), which differs from β_{Md}, β_A, α_t, and γ, all of which are simply counts of types.

If not all observed species are present in all subunits, there will be change in species composition, or species turnover, among subunits. Therefore, β_{At} can be interpreted as a measure of total species turnover in the dataset: if one were to start from one (actual or effective) compositional unit and walk through all the other compositional units, β_{At} would indicate how many (actual or effective) species are swapped for a new species along the way. Conceptually, every new species one encounters when moving to a new compositional unit equals half a species turned over, as does every old species that drops out. The additive beta component has recently become widely used in ecological studies (e.g., Veech *et al.*, 2002; Kiflawi and Spencer, 2004; Ricotta, 2008; de Bello *et al.*, 2010). Unfortunately, many studies refer to species turnover as "beta diversity," even though a count of species that change is both conceptually and numerically very different from a count of compositional units.

Equation (18) does not differentiate between presence–absence data (richness) and abundance data (diversity), and using the notation β for both species turnover and true beta diversity easily leads to confusing them. When the interest is in the turnover of actual species (actual species turnover), a more explicit notation is

$$R_{diff} = R_{tot} - R_{mean} \qquad [19]$$

Here the subscript "tot" refers to the total richness in the dataset, "mean" to the mean richness in the subunits, and "diff" to the difference between the total and the mean. The relevant mean here is the unweighted arithmetic mean. All components of the additive partitioning are based on the gamma classification, so all of them could have subscript γ. R_{tot} hence equals R_γ, but in the additive partitioning the components differ in which part of the dataset they refer to, rather than in which classification they are based on, so the subscript "γ" is omitted to simplify the notation. When the interest is in the turnover of effective species (effective species turnover), the equation can be written as

$${}^qD_{diff} = {}^qD_{tot} - {}^qD_{mean} \qquad [20]$$

The relevant mean here is the weighted generalized mean with exponent $1-q$ (calculated as in eqn [16]).

As with the multiplicative partitioning, the components of the additive partitioning behave in different ways depending on whether they are components of richness or of diversity. When all species are present in all subunits, R_{diff} invariably takes the value zero. In contrast, ${}^qD_{diff}$ does so only when either each species also has a constant abundance in all subunits, or when q is sufficiently large and the most abundant species is the same and has the same proportional abundance in all subunits.

The upper limit of absolute species turnover in a dataset depends both on the number of subunits N and on the mean species count within the subunits. If no subunits share any species, actual species turnover R_{diff} necessarily takes the value $(N-1)R_{mean}$. In the same situation, effective species turnover ${}^qD_{diff}$ takes the value $(N-1){}^qD_{mean}$ only if either all subunit weights are equal (i.e., all subunits contain the same abundance) or $q=1$ (*see* eqn [16]). Otherwise, the value that ${}^qD_{diff}$ takes when no subunits share any species cannot be predicted in advance, because it depends on dataset properties (Tuomisto, 2010a, c).

Whittaker's Species Turnover

In many cases, it is useful to quantify species turnover in relative terms, such that its value does not depend on the absolute number of species involved. Whittaker (1972) developed such a species turnover measure by subtracting unity from the value of the multiplicative beta component. The same measure, Whittaker's species turnover, can be obtained by dividing the additive beta component by the additive alpha component (notation from Tuomisto, 2010a; see also Kiflawi and Spencer, 2004, but note that β_M refers to both β_{Mt} and β_{Mt-1} in their text):

$$\beta_{Mt-1} = \beta_{At}/\alpha_t = (\gamma - \alpha_t)/\alpha_t = \gamma/\alpha_t - 1 = \beta_{Mt} - 1 \qquad [21]$$

Here β_{Mt} has the same numerical value as true beta diversity β_{Md} does, but the two differ in that β_{Mt} is unitless (since γ and α_t have the same measurement unit).

When the interest is in presence–absence data, Whittaker's actual species turnover can be expressed more explicitly as

$$R'_W = (R_{tot} - R_{mean})/R_{mean} = R_{tot}/R_{mean} - 1 \qquad [22]$$

The prime is added to specify that this is a relativized measure of richness; that is, one richness is expressed in multiples of another richness, which yields a unitless value. Subscript "W" refers to Whittaker's approach. When the interest is in abundance data, Whittaker's effective species turnover can be expressed as

$${}^qD'_W = ({}^qD_{tot} - {}^qD_{mean})/{}^qD_{mean} = {}^qD_{tot}/{}^qD_{mean} - 1 \qquad [23]$$

Whittaker's turnover measures express species turnover among the compositional units of the dataset in multiples of the within-unit species count (actual or effective). In other words, these measures quantify how many times the entire species composition turns over among the compositional units, rather than the number of species that turn over.

Whittaker's species turnover and absolute species turnover behave in a similar way in that both obtain their minimum and maximum values under the same conditions, but the numerical values of the actual species turnover measures differ by factor R_{mean} and those of effective species turnover differ by factor ${}^qD_{mean}$. As a result, the maximum value of Whittaker's actual species turnover is $N-1$, which is obtained when none of the N subunits share any species. The maximum value of

Whittaker's effective species turnover is also $N-1$, if nominal subunit weights are equal or $q=1$. In both cases, the values are independent of how many species are involved. This allows comparing compositional heterogeneity among datasets in such a way that the comparisons are not confounded by differences in species richness or diversity. Since the absolute species turnover measures (R_{diff} and ${}^{q}D_{diff}$) do depend on the absolute number of species (actual or effective, respectively), they are not monotonically related to the corresponding Whittaker's species turnover values, and each measure can hence lead to a different ranking of datasets.

Proportional Species Turnover

To obtain Whittaker's species turnover, absolute species turnover β_{At} is divided by the mean species count in the compositional units α_t. Equally well, β_{At} can be divided by the total species count in the dataset γ. Doing so leads to a new relative species turnover measure, *proportional species turnover* (Tuomisto, 2010a)

$$\beta_{Pt} = (\gamma - \alpha_t)/\gamma = 1 - \alpha_t/\gamma \qquad [24]$$

When the interest is in presence–absence data, proportional actual species turnover can be expressed more explicitly as

$$R'_P = (R_{tot} - R_{mean})/R_{tot} = 1 - R_{mean}/R_{tot} \qquad [25]$$

The prime is added to specify that this is a relativized measure of richness; that is, one richness is expressed as a proportion of another richness, which yields a unitless value. Subscript "P" refers to proportion. When the interest is in abundance data, proportional effective species turnover can be expressed as

$${}^{q}D'_P = ({}^{q}D_{tot} - {}^{q}D_{mean})/{}^{q}D_{tot} = 1 - {}^{q}D_{mean}/{}^{q}D_{tot} \qquad [26]$$

The proportional species turnover measures express what proportion of the species in the entire dataset is not contained in a single compositional unit (actual or effective). The term R_{mean}/R_{tot} also indicates the proportion of subunits in which the average actual species occurs (mean species frequency; Whittaker, 1972; Routledge, 1977).

Proportional species turnover obtains its maximum and minimum values in the same situations as Whittaker's and absolute species turnovers do. Its minimum value is zero, and its maximum value is $1-1/N$ (for effective species turnover, the latter is true only if all nominal subunit weights are equal or $q=1$). Just like Whittaker's species turnover, proportional species turnover is independent of how many species the dataset contains, and it can be used to compare the compositional heterogeneities of datasets that differ in the number of species. Proportional species turnover is monotonically (but curvilinearly) related with Whittaker's species turnover, so the two measures will always rank datasets in the same way (assuming, of course, that both are based either on richness data or on diversity data of the same order). Proportional species turnover has been derived several times (e.g., Roschewitz *et al.*, 2005; Ricotta, 2008; de Bello *et al.*, 2010). Unfortunately, it has usually been called beta diversity, even though it is both conceptually and numerically very different from true beta diversity (effective number of compositional units).

Resemblance Indices Related to Diversity

Several familiar similarity and dissimilarity indices can be derived as transformations of beta richness, beta diversity, or actual or effective species turnover in a dataset consisting of two subunits (Jost, 2006, 2007; Tuomisto, 2010a, b, c). For example, the Manhattan metric, as calculated using presence–absence data, is a simple transformation of absolute actual species turnover (Tuomisto, 2010a):

$$M = 2R_{diff} = 2(R_{tot} - R_{mean}) = b + c \qquad [27]$$

Here b is the number of actual species unique to the first subunit and c is the number of actual species unique to the second. The Manhattan metric is not bound to any fixed interval, but can increase indefinitely as the number of species in the dataset increases. In some applications, this can be a desirable property, but in most situations where compositional resemblance indices are used, it is not. Therefore, in most cases it is more appropriate to use indices whose values are constrained to a fixed interval. Typically, compositional similarity indices take the value zero when the two subunits that are being compared share no species, and the value unity when both subunits have identical species compositions.

The Sørensen index equals the one-complement of Whittaker's actual species turnover (see Tuomisto, 2010a for derivation)

$$C_S = 1 - R'_W = 2 - R_{tot}/R_{mean} = 2a/(2a + b + c) \qquad [28]$$

Here a is the number of species found in both subunits, and b and c the numbers of species found in one or the other subunit but not both subunits. The Sørensen index expresses the number of shared species as a proportion of the average number of species in the two subunits $R_{mean}=(2a+b+c)/2$.

The Sørensen index can easily be expanded to abundance data by replacing actual species turnover with effective species turnover:

$${}^{q}C_S = 1 - {}^{q}D'_W = 2 - {}^{q}D_{tot}/{}^{q}D_{mean} \qquad [29]$$

The addition of superscript q specifies that the calculations are based on diversity of order q. The classical (richness-based) Sørensen index is invariably constrained to the interval [0, 1], which is often a desirable property in a similarity index. With the diversity-based Sørensen index, the minimum value is fixed at zero only when the maximum value of ${}^{q}D'_W$ is fixed at unity, which happens when both subunits have the same nominal weight or $q=1$. Using ${}^{q}C_S$ as a similarity index therefore needs to be done with care to avoid misinterpreting the results. If different subunits have different total abundances, and it is for some reason not justified or desirable to rarefy them to the same number of individuals or convert them to within-subunit percentages, then only $q=1$ should be used when correct interpretation of the results depends on ${}^{q}C_S$ being restricted to the interval [0, 1].

The Jaccard index equals

$$C_J = 1 - 2R'_P = 2R_{mean}/R_{tot} - 1 = a/(a + b + c) \qquad [30]$$

The Jaccard index expresses the number of shared species as a proportion of the total number of species in the two subunits $R_{tot}=a+b+c$. Just as with the Sørensen index, the

Jaccard index can be adapted to abundance data:

$$ {}^{q}C_{\mathrm{J}} = 1 - 2^{q}D'_{\mathrm{p}} = 2^{q}D_{\mathrm{mean}}/{}^{q}D_{\mathrm{tot}} - 1 \qquad [31] $$

As with the diversity-based Sørensen index, the values of the diversity-based Jaccard index lie in the exact interval [0, 1] only if both subunits contain the same abundance or $q=1$.

Since both richness and diversity can be calculated for any dataset, irrespective of the number of subunits it contains, both the Sørensen index and the Jaccard index can be generalized to more than two subunits. The Jaccard index equals the one-complement of proportional species turnover between two subunits as ranged to the interval [0, 1] (obtained by dividing with the maximum value). Its values therefore naturally stay in this range even when $N>2$. The multiple-site Jaccard index for presence–absence data can be written as (Tuomisto, 2010a)

$$ C_{JN} = 1 - \frac{R'_{\mathrm{p}}}{1-1/N} = \frac{R_{\mathrm{mean}}/R_{\mathrm{tot}} - 1/N}{1-1/N} \qquad [32] $$

When $N=2$, the Sørensen index equals the one-complement of Whittaker's species turnover, but the value of the latter can exceed unity when $N>2$, so ranging is needed to keep its value within the interval [0, 1]. The multiple-site Sørensen index for presence–absence data then becomes (Diserud and Ødegaard, 2007; Tuomisto, 2010a)

$$ C_{SN} = 1 - \frac{R'_{\mathrm{W}}}{N-1} = \frac{N - R_{\mathrm{tot}}/R_{\mathrm{mean}}}{N-1} \qquad [33] $$

The one-complements of both C_{JN} and C_{SN} quantify how much the actual species turnover of the corresponding kind exceeds its minimum possible value, expressed as a proportion of the total possible range of values (given N). The one-complement of C_{SN} was proposed by Harrison *et al.* (1992) under the name "beta-1" (although their original equation gives values in percentages rather than proportions).

For effective species turnover, the maximum values are definable in terms of N only when all nominal subunit weights are equal or $q=1$, and ranging is therefore possible only when at least one of these conditions holds. Then the multiple-site Sørensen index for abundance data is (Tuomisto, 2010a)

$$ {}^{q}C_{SN} = 1 - \frac{{}^{q}D'_{\mathrm{W}}}{N-1} = \frac{N - {}^{q}D_{\mathrm{tot}}/{}^{q}D_{\mathrm{mean}}}{N-1} \qquad [34] $$

Similarly, the multiple-site Jaccard index for abundance data is (Jost, 2006; Tuomisto, 2010a)

$$ {}^{q}C_{JN} = 1 - \frac{{}^{q}D'_{\mathrm{p}}}{1-1/N} = \frac{{}^{q}D_{\mathrm{mean}}/{}^{q}D_{\mathrm{tot}} - 1/N}{1-1/N} \qquad [35] $$

The one-complements of both indices quantify how much the effective species turnover of the corresponding kind exceeds its minimum possible value, expressed as a proportion of the total possible range of values (given N).

Binary measures can be generalized to abundance data in different ways, and a particular nonlinear (but monotonic) transformation of diversity gives the following overlap index (Chao *et al.*, 2008):

$$ C_{qN} = \frac{({}^{q}D_{\mathrm{mean}}/{}^{q}D_{\mathrm{tot}})^{q-1} - (1/N)^{q-1}}{1-(1/N)^{q-1}} \qquad [36] $$

For $q<1$, the exponents become negative and hence cause all terms to be inverted. How this affects the calculation can be seen by rewriting the equation in the mathematically equivalent form that has positive exponents when $q<1$:

$$ C_{qN} = \frac{N^{1-q} - ({}^{q}D_{\mathrm{tot}}/{}^{q}D_{\mathrm{mean}})^{1-q}}{N^{1-q}-1} \qquad [37] $$

Note that the parameter q in C_{qN} is specified in a subscript rather than in a superscript. This distinction is important, because even though C_{qN} is expressed as a function of true diversities of order q, it is not linearly related with ${}^{q}D$ but with the basic sum ${}^{q}D^{1-q}$ (see the explanation in connection with eqn [6]). The C_{qN} index is therefore a measure of compositional overlap (Chao *et al.*, 2008), not of relative species turnover as ${}^{q}C_{SN}$ and ${}^{q}C_{JN}$ are. As a result, C_{qN} behaves in a different way than ${}^{q}C_{SN}$ and ${}^{q}C_{JN}$ do when the value of q is changed.

In the special case of $q=0$, the C_{qN} index converges on the ${}^{q}C_{SN}$ index (compare eqns [37] and [34]), and both thus converge on the classical Sørensen index when $N=2$ (see eqns [29] and [28]. In the special case of $q=2$, the C_{qN} index converges on the ${}^{q}C_{JN}$ index instead (compare eqns [36] and [35]), and both converge on the Morisita–Horn index when $N=2$. The Morisita–Horn index generalized to N subunits can be written as (see Jost, 2006 for derivation)

$$ C_{2N} = \frac{{}^{2}D_{\mathrm{mean}}/{}^{2}D_{\mathrm{tot}} - 1/N}{1-1/N} \qquad [38] $$

In the special case of $q=1$ and $N=2$, the C_{qN} index converges on the Horn index (Chao *et al.*, 2008). The Horn index generalized to N subunits can be written as (see Tuomisto, 2010a for derivation)

$$ C_{1N} = \frac{\log(N) - \log({}^{1}D_{\mathrm{tot}}/{}^{1}D_{\mathrm{mean}})}{\log(N)} \qquad [39] $$

The C_{qN} index provides a valid and useful measure of overall compositional overlap among all sampling units in a dataset of interest at all values of q. At some values of q, it can also be interpreted in terms of effective species turnover. At $q=0$, C_{qN} equals Whittaker's species turnover expressed as a proportion of the maximum value obtainable, given N. At $q=2$, C_{qN} equals proportional species turnover expressed on a relative scale between the minimum and the maximum possible, given N. However, C_{qN} is not primarily a species turnover measure, and at other values of q, interpreting the value of C_{qN} in terms of species turnover is misleading. For example, at $q=1$, C_{qN} quantifies the magnitude of relative Shannon entropy, not relative species turnover. Conversely, the value of ${}^{q}C_{SN}$ can be interpreted in terms of relative Whittaker's species turnover at any value of q, but in terms of compositional overlap, only at $q=0$; the value of ${}^{q}C_{JN}$ can be interpreted in terms of relative proportional species turnover at any value of q, but in terms of compositional overlap only at $q=2$.

To choose an appropriate index for a particular study, it is necessary to understand which aspect of the data each index quantifies. Ecological hypotheses invariably concern some

specific aspect of data behavior, and choosing an index that corresponds to the wrong aspect can therefore lead to incorrect interpretations of the results and unjustified conclusions on the tested hypotheses. The mathematical properties of the indices also determine what needs to be taken into account when comparing their values among datasets, and indeed whether the index values can be meaningfully compared at all. How an index is calculated is not a mere technical detail, because it defines the variable of interest and hence the ecological meaning of the results obtained.

See also: Biodiversity, Definition of. Complexity Versus Diversity. Diversity, Taxonomic Versus Functional. Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data. Species Diversity, Overview

References

Alatalo RV (1981) Problems in the measurement of evenness in ecology. *Oikos* 37: 199–204.

Anderson MJ, Crist TO, Chase JM, *et al.* (2011) Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters* 14: 19–28.

Beck J and Schwanghart W (2010) Comparing measures of species diversity from incomplete inventories: An update. *Methods in Ecology and Evolution* 1: 38–44.

de Bello F, Lavergne S, Meynard CN, Lepš J, and Thuiller W (2010) The partitioning of diversity: Showing Theseus a way out of the labyrinth. *Journal of Vegetation Science* 21: 992–1000.

Chao A, Chazdon RL, Colwell RK, and Shen T-J (2005) A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters* 8: 148–159.

Chao A, Jost L, Chiang SC, Jiang Y-H, and Chazdon RL (2008) A two-stage probabilistic approach to multiple-community similarity indices. *Biometrics* 64: 1178–1186.

Colwell RK and Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions: Biological Sciences* 345: 101–118.

Diserud O and Ødegaard F (2007) A multiple-site similarity measure. *Biology Letters* 3: 20–22.

Gadagkar R (1989) An undesirable property of Hill's diversity index N_2. *Oecologia* 80: 140–141.

Gauthier O, Sarrazin J, and Desbruyères D (2010) Measure and mis-measure of species diversity in deep-sea chemosynthetic communities. *Marine Ecology Progress Series* 402: 285–302.

Good IJ and Toulmin GH (1956) The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika* 43: 43–63.

Gorelick R (2011) Commentary: Do we have a consistent terminology for species diversity? The fallacy of true diversity. *Oecologia* 167: 893–902.

Gotelli NJ and Colwell RK (2001) Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4: 379–391.

Gregorius H-R (2010) Linking diversity and differentiation. *Diversity* 2: 370–394.

Gregorius H-R and Gillet EM (2008) Generalized Simpson-diversity. *Ecological Modelling* 211: 90–96.

Harrison S, Ross SJ, and Lawton JH (1992) Beta diversity on geographic gradients in Britain. *Journal of Animal Ecology* 61: 151–158.

Hill MO (1973) Diversity and evenness: A unifying notation and its consequences. *Ecology* 54: 427–432.

Hoffmann S and Hoffmann A (2008) Is there a "true" diversity? *Ecological Economics* 65: 213–215.

Hurlbert SH (1971) The nonconcept of species diversity: A critique and alternative parameters. *Ecology* 52: 577–586.

Jost L (2006) Entropy and diversity. *Oikos* 113: 363–375.

Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88: 2427–2439.

Jost L (2009) Mismeasuring biological diversity: Response to Hoffmann and Hoffmann (2008). *Ecological Economics* 68: 925–928.

Jost L (2010) The relation between evenness and diversity. *Diversity* 2: 207–232.

Jurasinski G and Koch M (2011) Commentary: Do we have a consistent terminology for species diversity? We are on the way. *Oecologia*.http://dx.doi.org/10.1007/s00442-011-2126-6.

Jurasinski G, Retzer V, and Beierkuhnlein C (2009) Inventory, differentiation, and proportional diversity: A consistent terminology for quantifying species diversity. *Oecologia* 159: 15–26.

Kiflawi M and Spencer M (2004) Confidence intervals and hypothesis testing for beta diversity. *Ecology* 85: 2895–2900.

Kindt R, van Damme P, and Simons AJ (2006) Tree diversity in western Kenya: Using profiles to characterise richness and evenness. *Biodiversity and Conservation* 15: 1253–1270.

Koleff P, Gaston KE, and Lennon JJ (2003) Measuring beta diversity for presence–absence data. *Journal of Animal Ecology* 72: 367–382.

Lande R (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5–13.

MacArthur RH (1964) Environmental factors affecting bird species diversity. *American Naturalist* 98: 387–397.

MacArthur RH (1965) Patterns of species diversity. *Biological Reviews* 40: 510–533.

Magurran AE (2004) *Measuring Biological Diversity*. Oxford: Blackwell Publishing.

Mendes RS, Evangelista LR, Thomaz SM, Agostinho AA, and Gomez LC (2008) A unified index to measure ecological diversity and species rarity. *Ecography* 31: 450–456.

Moreno C and Rodríguez P (2011) Commentary: Do we have a consistent terminology for species diversity? Back to basics and toward a unifying framework. *Oecologia* 167: 889–892.

Peet RK (1974) The measurement of species diversity. *Annual Review of Ecology and Systematics* 5: 285–307.

Plotkin JB and Muller-Landau HC (2002) Sampling the species composition of a landscape. *Ecology* 83: 3344–3356.

Rényi A (1960) On measures of entropy and information. In: *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 1960, pp. 547–561.

Ricotta C (2008) Computing additive β-diversity from presence and absence scores: A critique and alternative parameters. *Theoretical Population Biology* 73: 244–249.

Ricotta C and Avena CG (2002) On the information-theoretical meaning of Hill's parametric evenness. *Acta Biotheoretica* 50: 63–71.

Roschewitz I, Gabriel D, Tscharntke T, and Thies C (2005) The effects of landscape complexity on arable weed species diversity in organic and conventional farming. *Journal of Applied Ecology* 42: 873–882.

Routledge RD (1977) On Whittaker's components of diversity. *Ecology* 58: 1120–1127.

Routledge RD (1979) Diversity indices: Which ones are admissible? *Journal of Theoretical Biology* 76: 503–515.

Shannon CE (1948) A mathematical theory of communication. *The Bell System Technical Journal* 27: 379–423.

Simpson EH (1949) Measurement of diversity. *Nature* 163: 688.

Smith B and Wilson JB (1996) A consumer's guide to evenness indices. *Oikos* 76: 70–82.

Tuomisto H (2010a) A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33: 2–22.

Tuomisto H (2010b) A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33: 23–45.

Tuomisto H (2010c) A consistent terminology for quantifying species diversity? Yes, it does exist. *Oecologia* 164: 853–860.

Tuomisto H (2011) Commentary: Do we have a consistent terminology for species diversity? Yes, if we choose to use it. *Oecologia* 167: 903–911.

Veech JA, Summerville KS, Crist TO, and Gering JC (2002) The additive partitioning of species diversity: Recent revival of an old idea. *Oikos* 99: 3–9.

Vellend M (2001) Do commonly used indices of b-diversity measure species turnover? *Journal of Vegetation Science* 12: 545–552.

Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30: 279–338.

Whittaker RH (1972) Evolution and measurement of species diversity. *Taxon* 21: 213–251.

Whittaker RH (1977) Evolution of species diversity in land communities. *Evolutionary Biology* 10: 1–67.

Wilson MV and Shmida A (1984) Measuring beta diversity with presence–absence data. *Journal of Ecology* 72: 1055–1064.

Deforestation and Land Clearing

Jaboury Ghazoul, Department of Environmental Sciences, ETH Zurich, Switzerland

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Jaboury Ghazoul, Julian Evans, vol. 2, pp 23–36, © 2001, Elsevier Inc.

Glossary

Afforestation Planting of previously unforested land with trees.

Agroforestry Practice of combining agricultural crops or animal husbandry with the maintenance and cultivation of trees on the same patch of land.

Deforestation Complete or almost complete removal of trees to less than 10% of tree cover and conversion of forested land to other uses as a result of human activities.

Desertification Development of desert conditions as a result of human activity, frequently by overuse of trees and overgrazing, or climate change.

Forest Land with minimum tree height of 5 m, a tree canopy cover exceeding 10%, and minimum area of 0.5 ha

Forest certification Procedure whereby an independent third party inspects forest management and timber harvesting practices to assess compliance against a set of ecological, economic and social standards for sustainable forestry.

Forest degradation Damage to forest ecosystems through human activities that does not result in the total elimination of forest cover.

Natural forest Area of land that supports a minimum of 10% tree cover that has arisen as a result of natural processes of establishment and succession.

Plantation Area of land that supports planted forest, usually for commercial exploitation.

REDD Reduced Emissions from Deforestation and forest Degradation: one form of payment for environmental service where carbon storage value of forests threatened by degradation or clearance is financially recognized through payments to forest owners (usually, but not exclusively, nation states) to conserve the forest.

Secondary forest Area of previously forested land that was subsequently degraded or deforested through human or natural action but which now supports regenerating or mature natural forest.

Shifting cultivation Temporally or spatially cyclical agricultural system that involves the clearing of land followed by cultivation and fallow periods.

Introduction

Deforestation is the complete or almost complete removal of tree cover and conversion of the land to other uses. A clear definition of deforestation is, however, challenging on account of differing definitions for forests. The UN Food and Agriculture Organization (FAO) defines forest as land with minimum tree height of 5 m, a tree canopy cover exceeding 10%, and minimum area of 0.5 ha (FAO 2009). Other definitions use 2–5 m tree heights, 10–30% canopy area, and as little as 0.01–1 ha minimum area. This article uses the FAO definition, and therefore deforestation is taken to mean the semi-permanent depletion of tree crown cover to less than 10%. Deforestation differs from forest degradation in that the latter represents significant damage to forest ecosystems but where forest cover exceeding 10% is retained.

Historical Deforestation and Land Clearance

At the advent of agriculture, approximately 8000 years ago, forests are thought to have covered approximately 40% of the world's land area, or approximately 6000 million ha. Up to AD 1500 , the spread of agriculture across the globe resulted in the clearance of many forests, particularly those on the most accessible and fertile land. The areas most affected were the Middle East, the Mediterranean, South Asia, Europe, and the Far East. Despite the long history of human occupation of Africa, evidence for early human clearance of African forests is patchy and localized, although currently uninhabited and forested regions of Congo reveal evidence of early human occupation in the form of pottery shards and charcoal in the soil. Deforestation in the New World was, however, clearly underway long before the arrival of Europeans, including indigenous large-scale landscape transformation at sites across the Amazon basin from AD 250 to approximately AD 1600.

Deforestation in temperate forest regions accelerated rapidly from late eighteenth century driven. The industrial revolution, expanding populations and increasing affluence, put tremendous pressure on the remaining forests to supply fuel for industrial development and land for agriculture, and by the end of the nineteenth century woodlands and forests across Europe, Russia, and China had been cleared. In nineteenth century North America, settlers cleared around 142 million ha of temperate forests for agriculture and a further 8 million ha for industry and urban development. This trend was paralleled by other lands colonized by Europeans, notably New Zealand and Australia.

European conquest and settlement of the Neotropics decimated indigenous communities through disease resulting in the abandonment and secondary regrowth of formerly cultivated areas of Amazonia. Settlers from Europe, however, started to clear accessible forests of coastal Brazil and the Caribbean for sugar plantations in the seventeenth and eighteenth centuries. India was heavily exploited for its teak

and for its potential to grow cash crops by the British colonial administration in the early ninteenth century, resulting in extensive forest clearance. Deforestation in coastal regions of Southeast Asia was equally extensive and driven by improved transport infrastructure and peasant immigration particularly from 1850 to 1920. The onset of industrial timber extraction after 1920 dramatically accelerated Southeast Asian deforestation, as it has elsewhere in both temperate and tropical forests. Between 1850 and 1980, 15% of the world's forests and woodlands were cleared.

Contemporary Deforestation

Deforestation and land clearance in the twentieth century increased greatly, with the highest rates of clearance occurring since 1960. Most current deforestation occurs in the tropical regions, whereas in north temperate countries there has been a net annual 0.1% increase in forest cover from 1990 to 2005 due to reforestation and regeneration policies. In China, forest cover has increased even more dramatically with a positive annual change of 1.2% through 1990–2000 and 2.2% for 2000–2005. This overall increase in temperate forest cover, however, masks inconsistencies at national scales. Thus in Canada, although deforestation is declining, with the annual rate dropping from nearly 68,000 ha in 1990 to approximately 45,000 ha in 2008, large hydroelectric projects caused deforestation spikes in 1995 and 2006.

By 2008, global forest area shrunk to 397 million ha (Table 1), largely as a consequence of human exploitation in the latter half of the twentieth century. The total area of forest cleared globally between 1990 and 2005 is 125 million ha, approximately 3% of the 1990 forest extent. This represents a decline from the high rates of the 1980s, when annual global deforestation amounted to around 14.0 Mha, to the lower rates of 8.9 million ha annually in the 1990s (Table 1). Since 2000, global forest loss has declined further, and is estimated at 7.3 million ha annually. There are, however, large regional variations: almost all deforestation now occurs in the tropics. Indeed, since 1990 some 7.8% of tropical forest have been lost, with an equivalent amount lost in the 1980s.

Deforestation is now largely a tropical concern, and most deforestation is concentrated in relatively few tropical countries. The top 10 deforesting countries, which are all tropical or subtropical, account for 63% of total deforestation of all countries that show a net loss of forest area during 2000–2005 (Table 2). A large proportion of the world's biodiversity is found within tropical forests, and so tropical deforestation has considerable relevance to global biodiversity conservation (Table 3).

Table 1 Annual changes in forest area

	Forest area (2005)	*Annual forest change*			
		1990–2000		*2000–2005*	
	1000 ha	*1000 ha*	*%*	*1000 ha*	*%*
Central Africa	236,070	−910	−0.37	−673	−0.28
East Africa	77,109	−801	−0.94	−771	−0.97
North Africa	76,805	−526	−0.64	−544	−0.69
Southern Africa	171,116	−1,152	−0.63	−1,154	−0.66
West Africa	74,312	−985	−1.17	−899	−1.17
Total Africa	**635,412**	**−4,375**	**−0.64**	**−4,040**	**−0.62**
East Asia	244,862	1,751	0.81	3,840	1.65
West Asia	27,571	25	0.09	5	0.02
Central Asia	16,017	9	0.06	9	0.06
Temperate Asia	**288,450**	**1,785**	**0.62**	**3,854**	**1.34**
South Asia	79,239	213	0.27	−88	−0.11
Southeast Asia	203,887	−2,790	−1.20	−2,763	−1.30
Tropical Asia	**283,126**	**−2,577**	**−0.91**	**−2,851**	**−1.01**
Oceania	**206,254**	**−448**	**−0.21**	**−356**	**−0.17**
Eastern Europe	43,042	71	0.17	150	0.35
CIS countries	826,588	103	0.01	−73	−0.01
Western Europe	131,763	703	0.56	583	0.45
Total Europe	**1,001,393**	**877**	**0.09**	**660**	**0.07**
Caribbean	5,974	36	0.65	54	0.92
Central America	22,411	−380	−1.47	−285	−1.23
South America	851,540	−3,802	−0.44	−4,251	−0.5
North America	677,464	17	0	−101	−0.01
Tropical America	**911,142**	**−4,345**	**−0.48**	**−4,592**	**−0.50**
Temperate America	**646,247**	**216**	**0.03**	**9**	**0.00**
Total America	**1,557,389**	**−4,129**	**−0.27**	**−4,583**	**−0.29**
Tropical	**1,788,013**	**−10,066**	**−0.56**	**−10,142**	**−0.57**
Temperate	**2,184,011**	**1,199**	**0.05**	**2,826**	**0.13**
Global	**3,972,024**	**−8,867**	**−0.22**	**−7,316**	**−0.18**

Source: Reproduced from Food and Agriculture Organization (FAO) (2009) *State of the World's Forests*. Rome: FAO. Available online at: http://www.fao.org/forestry/49666/en/

Table 2 "Top 10" Deforesting countries in terms of total forest lost (2005)

Country	*1990–2000*		*2000–2005*	
	Annual loss (1000 ha)	*Rank*	*Annual loss (1000 ha)*	*Rank*
Brazil	−2,681	1	−3,103	1
Indonesia	−1,872	2	−1,871	2
Sudan	−589	3	−589	3
Myanmar	−466	5	−466	4
Zambia	−445	6	−445	5
Tanzania	−412	7	−412	6
Nigeria	−410	8	−410	7
D. Rep. Congo	−532	4	−319	8
Zimbabwe	−313	11	−313	9
Venezuela	−288	12	−288	10
Net Deforesting Countries	−13,047	–	−12,878	–
Total World	−8,868	–	−7,317	–

Source: Reproduced from Food and Agriculture Organization (FAO) (2009) *State of the World's Forests*. Rome: FAO. Available online at: http://www.fao.org/forestry/49666/en/

Table 3 Important deforesting countries and regions in terms of annual rate of loss (2005)

Country	*1990–2000*		*2000–2005*	
	% Annual loss	*Rank*	*% Annual loss*	*Rank*
Comoros	−4.00	1	−7.40	1
Burundi	−3.70	2	−5.20	2
Togo	−3.40	4	−4.50	3
Mauritania	−2.70	7	−3.40	4
Nigeria	−2.70	8	−3.30	5
Honduras	−3.00	5	−3.10	6
Afghanistan	−2.50	9	−3.10	7
Benin	−2.10	10	−2.50	8
Uganda	−1.90	13	−2.20	9
Philippines	−2.80	6	−2.10	10
Pakistan	−1.80	14	−2.10	11
Indonesia	−1.70	16	−2.00	12
Cambodia	−1.10	30	−2.00	13
Ghana	−2.00	12	−2.00	14
Wallis and Futuna Islands	−0.80	40	−2.00	15
Net Deforesting Countries	−0.44	–	−0.46	–
Total World	−0.22	–	−0.18	–

Source: Reproduced from Food and Agriculture Organization (FAO) (2009) *State of the World's Forests*. Rome: FAO. Available online at: http://www.fao.org/forestry/49666/en/

In temperate countries, and particularly Europe, there is no overall deforestation in terms of net area but rather a small increase in forest cover owing to policies of regeneration and afforestation. This, however, disguises a steady and continuing transition from natural or seminatural forest stands to plantation forests. Forest cover in Europe has consistently increased in recent years and a relatively high proportion of its land area is under forest cover. Much of this comprises planted forests or regrowth as most of the original forest cover was cleared centuries ago. Declining population, low land dependencies, concern for protection of the environment, and a well-developed policy and institutional framework all favor further expansion of forest area. The exception to this trend may be Eastern Europe and Russia, where the importance of forest resources as an accessible and tradeable commodity takes priority. This could be compounded by the projected impacts of increasing frequency and extent of fires (due to climate change) in the Russian Federation as well as in parts of Western Europe such as the Iberian Peninsula.

Causes of Deforestation

There is no single cause of deforestation but rather it is the result of the interaction of social, economic, political, and cultural forces with the environment. Several underlying socioeconomic causes create conditions that favor forest clearance by readily identifiable direct causes.

The causes of deforestation vary among regions. In Africa, FAO reports that the major direct cause of deforestation is clearance by farmers driven by increasing population pressures. In Latin America, settlement and infrastructure projects

in forested areas result in clearance of land for cattle ranching and permanent agriculture, often combined with financial incentives in the form of subsidies and favorable tax policies. In Asia, intensive timber harvesting and shifting cultivation as well as the expansion of large-scale agricultural projects and plantation crops, such as oil palm and rubber, and to a lesser extent transmigration projects, all contribute to deforestation. Forest land is often not suitable for sustainable agricultural development, and as soils become exhausted new areas of forest have to be cleared. For example, 80% of the Amazon basin is ill suited to sedentary farming. Desertification through unsustainable agricultural development has contributed to much deforestation in drier regions of Africa and Asia.

Underlying Causes of Deforestation

The underlying causes of deforestation are the factors that give rise to conditions in which forest clearance becomes a rational or necessary behavior. They may be local or national socioeconomic or political forces, or they may be external global forces such as the state of the global market economy. They are generally beyond the control of an individual but strongly influence the decisions individuals make regarding the management and use of forests and forest resources.

Population Growth and Poverty

Population growth is one of the most publicized underlying factors for tropical deforestation. Average global population growth over the past three decades has been 820 million per decade. Most of this increase occurs in developing countries in which deforestation is greatest. Increasing populations place pressure on forests by increasing demand for food, energy, water, wood, paper, and a variety of other products. The importance of population pressure as an underlying cause for deforestation can be misleading in that it is dependent on the ability of the land to support the population, the importance of forest for supplying goods and services to the people, and the level of control by national or local institutions.

An estimated 3.25 billion people (2004 figures) live in rural areas and are dependent on agriculture to meet their basic needs. The rural poor have very few economic options and are often forced to seek short-term solutions to their economic problems. These solutions include clearing forested land to grow subsistence crops. Opportunities for improving livelihoods by other means are limited due to the lack of rural capital, low capacity of subsistence farming to generate income, and the lack of infrastructure and education.

Development Policies and Tax Incentives

Debt repayments constitute a large proportion of the national budget of many tropical countries, and structural adjustment programs introduced as a result often favor the maximization of foreign exchange through direct and unsustainable exploitation of forest capital and by conversion of forests to agriculture for export crops. Large-scale extensive agricultural development, frequently at the expense of small farmers as well as forest cover, is further encouraged through the provision of state subsidies for agriculture and livestock expansion, reduction in income and corporate taxes, and tax breaks on imports of equipment for new industries. Expansion of agricultural crops for export or to satisfy national demands destroys forest directly but also causes the displacement of subsistence farmers who are forced to relocate and clear new and often marginal lands elsewhere.

The privatization of public resources, advocated by the World Bank and other bilateral donor agencies, favors management strategies that maximize the short-term economic gain for the new owners, whereas non-monetary forest services, such as soil conservation and watershed protection, are not valued highly in a market-driven environment. Government incentives and subsidies have allowed some otherwise uneconomical industries to prosper at the expense of forest cover, whereas development projects often fail to account fully for the value of forest capital lost.

A lack of understanding of the real value of forests goods and services results in poor policies. The institutional weakness of the national forest department or corruption within the government can lead to policy decisions that favor private interests at the expense of the benefits to society as a whole. In recent years there has been an improvement in the reformulation of forest policies of several tropical countries. Subsidies that promote cattle ranching have been withdrawn in Brazil, whereas Costa Rica is now beginning to account for the destruction of forest capital in its national economy.

Resettlement Programs

Government-sponsored resettlement programs, such as the transmigration programs of Indonesia and Brazil, encourage landless farmers or the urban poor to clear areas of forest for subsistence agriculture or for the cultivation of cash crops for export. These schemes relieve the pressure of urban crowding and allow governments to avoid the difficult issue of land reform by providing new agricultural land from forest. However, many forested lands are unsuitable for permanent agriculture and resettlement programs have invariably failed as a result. Transmigrants are thus forced to resort to shifting cultivation, leading to further forest degradation and deforestation.

Tenurial Policies

Eighty-four percent of the world's forests are publicly owned, although private ownership is on the rise, whereas much of the agricultural land is owned by large land owners or corporations and is therefore not accessible to the majority of the farming population. Small farmers, who may not have legal title to their land, are frequently displaced or forced to sell their land because of increasing debts. These farmers often move to the forest frontier to clear a new plot of land, a practice that continues because it is usually easier for governments to ignore deforestation than to deal with the difficult issues of land redistribution.

Without legal land title there is little incentive to invest in the land to make it more productive, and it becomes more economically logical to pursue short-term gain and to move on to clear new forest land once productivity declines. Because most tropical forest lands are owned by the state, clearance is often illegal and governments are unwilling to grant legal title to small farmers for land acquired in this way.

Market Demands

One-third of the world's forests are primarily used for production of wood and nonwood products. As populations grow and become more affluent, the demand for forest products increases, particularly for industrial timber and pulpwood for making paper. Some countries have significant forest product exports (Table 4), though the extent to which international markets contribute to deforestation varies greatly from country to country. In the top 10 deforesting countries, national demand for forest products accounts for a significant amount of the area deforested. Furthermore, there is no strong link between increasing international demand for forest products generally and deforestation, not least because about 80% of the world's pulp and paper is produced and processed in temperate regions. Also, much of the world's industrial-grade timber is softwood (i.e., coniferous), and tropical forests are overwhelmingly broad-leaved. Supplies of industrial lumber in tropical countries increasingly come from tropical plantations, notably pines, cypress, eucalyptus, and teak.

Deforestation can be driven by growth in both the domestic and export markets. Rising production of beef in Central and South America to feed a growing domestic market resulted in extensive deforestation by ranchers, farmers, and land speculators since the 1980s. Over the past two decades the growing international demand for palm oil led to the widespread establishment of oil palm plantations in Indonesia, with palm oil exports rising from approximately 50% of production in the mid 1980s to 75% in 2006. Although the importance of export markets should not be underestimated, national market forces can be as important as international trade in determining the rates and extent of tropical deforestation. Consequently, it is likely that international trade offers only limited scope for reducing deforestation rates in most tropical countries.

Clearance of forest for agricultural production also appears to be driven by increasing national demand for agricultural crops. Self-sufficiency in agricultural production has been a primary developmental goal of many of these countries and has led to policies that encourage conversion of forests to fields.

Undervaluation of Forests and Forest Products

Where logging has preceded and been the first stage in wholesale land clearance, it is often because the value placed on the timber is no more than the cost of extraction and marketing. In this sense, logging is not only unsustainable forest management but also becomes a mining activity in which timber value reflects the ease or otherwise of obtaining the raw material. The value does not reflect cost of replacement or cost of growing beyond what is often a nominal payment of royalty to the owner. If timber was valued to reflect its true cost of replacement, then growing trees to produce timber would become economically worthwhile and hence potentially a sustainable option. Although it is not worthwhile, deforestation, especially in the tropics, is likely to continue because clearance and conversion is perceived as more profitable.

The nontimber benefits and services that trees and forests provide are often more important than their timber products. Environmental benefits, including soil protection, shelter, microclimate amelioration, contribution to regional and global hydrological and carbon cycles, are rarely quantified and hardly ever incorporated in economic assessments. Only when massive downstream flooding is traced back to wanton deforestation in the catchment are such connections made.

Thus a major underlying cause of deforestation is the widespread failure to value sufficiently both forest products and the many environmental benefits forests bring. Regrettably, this is largely because simple and widely accepted approaches to such valuations in economic terms do not exist. For example, stumpage, that is the charges that governments demand from loggers for state-owned timber, often undervalues the resource, which encourages waste and makes other land uses more economically attractive.

Weak Government Institutions

Although almost all countries have explicit forest laws and policies designed to conserve forests, laws, regulations, and policies in producer countries are poor in most areas. Some countries, such as Brazil, have implemented a major overhaul

Table 4 The importance of exports to the forest-based economies of the top 10 deforesting countries (2005)[a]

Country	Timber products[b]			Paper and paperboard[c]			Fuelwood[d]			Total %
	Production	Export	%	Production	Export	%	Production	Exports	%	
Brazil	133,445	6,100	4.57	19,789	8,037	40.61	138,783	0	0.00	4.8
Indonesia	37,328	6,138	16.44	12,820	6,271	48.92	70,719	1	0.00	10.3
Sudan	2,226	0	0.00	3	0	0.00	17,901	0	0.00	0.0
Myanmar	5,905	1,804	30.55	85	0	0.00	38,286	0	0.00	4.1
Zambia	1,500	11	0.73	4	0	0.00	8,798	0	0.00	0.1
Tanzania	2,359	90	3.82	81	4	4.94	21,914	1	0.00	0.4
Nigeria	11,513	64	0.56	42	2	4.76	61,629	1	0.00	0.1
D. Rep. Congo	4,419	159	3.60	0	1	-	72,126	0	0.00	0.2
Zimbabwe	1,416	107	7.56	164	13	7.93	8,380	0	0.00	1.2
Venezuela	3,206	73	2.28	841	9	1.07	3,884	0	0.00	1.0

[a]Note all values in thousand cubic meters.

[b]Includes industrial roundwood, sawnwood, and wood-based panels.

[c]Includes pulp for paper.

[d]Does not include charcoal.

Source: Reproduced from Food and Agriculture Organization (FAO) (2009) *State of the World's Forests*. Rome: FAO. Available online at: http://www.fao.org/forestry/49666/en/

of policies and regulations in recent years, though ambiguity in forest legislation remains a common problem across tropical countries. Poor enforcement of forest legislation, through lack of political will, incoherence of the forest laws, or limited capacity, ensures that many of the laws that do exist are routinely flouted.

Extensive forests by their very geography are likely to be remote from towns and cities and hence far from the rule of law. It is therefore easy for illegal logging and clearance to continue unseen and unchecked. Furthermore, forest services are frequently the ''Cinderella'' organization of government, being viewed as inferior to agriculture and even wildlife and tourism. Few resources are attracted, and poorly paid staff often have difficulty in ensuring that sustainable management practices are implemented and in imposing their authority on perhaps large private sector interests. Quite apart from the risk of corruption that these circumstances afford, many staff once trained simply dislike the remoteness of forest management and supervision and prefer the white-collar work of the city office.

Direct Causes of Deforestation

Various forms of agricultural development have been the main drivers of deforestation (**Table 5**). Agricultural expansion into forests can take many forms, ranging from haphazard subsistence agriculture at the forest frontier to large-scale and well organized cash crop expansion. Other direct causes of forest destruction, such as fire, are often simply the tools by which land clearance for agriculture is achieved. Aside from agricultural, mining and unsustainable timber extraction account for a considerable area of forest clearance. All these direct drivers of deforestation are principally relevant in the tropics.

Shifting Cultivation

The contribution of small-scale shifting agriculture to tropical deforestation remains unresolved due to the widely variable agricultural practices that are encompassed by this term. Indeed, estimates of the number of shifting cultivators worldwide vary from as little as 40 million to as high as 500 million. Some types of small-scale agriculture undoubtedly cause deforestation, but the inherent stability and long-term viability of many shifting cultivation systems are unlikely to result in long-term forest clearance. The least destructive form of shifting cultivation is when land cultivated for two or three years is then left for a long fallow period. This long fallow shifting cultivation occurs only under conditions of very low population density. Expanding populations and land scarcity have led to more intensive farming systems and shorter fallow periods. Government policies have also created 'shifted' cultivators who are the typical slash-and-burn farmers of recent decades. Unlike the traditional farmers who have practiced shifting cultivation for decades, these shifted agriculturists have been forced by circumstances to cultivate habitats that are unfamiliar to them. Government resettlement and transmigration schemes attract migrants for whom forest cultivation is an unfamiliar means of generating a livelihood and income. In the Amazon, migrant cultivators are attracted to the forest frontier, where they clear and cultivate land for a few years. This land quickly becomes exhausted due to unsuitable soils or farming techniques and the land is sold to cattle ranchers and the ''farmers'' move on to clear more land.

Commercial Agriculture

Large-scale commercial agriculture is most frequently practiced by large corporations or state enterprises. These large operations can dispossess local land owners and farmers of the best and most fertile agricultural land, indirectly leading to deforestation in areas in which the farmers relocate. The establishment of oil palm plantations in valleys of Honduras in the 1970s, for example, displaced thousands of farmers who were forced to clear forests from steep slopes to establish new farms.

Commercial agriculture often leads to direct conversion of large tracts of forest to plantation estates and rice fields. This has been particularly prevalent in Indonesia and other regions of Southeast Asia where oil palm, coconut, or rubber plantations have been established on cleared forest land. In Indonesia, harvested oil palm plantations increased from approximately 300,000 ha in 1980 to 2 million ha in 2000 and to 5 million ha in 2008. Land clearance for agricultural development is often subsidized by governments, and because the owners of the agribusiness companies are politically well appointed there is little interest in forest protection. Areas for conversion are frequently burnt because this is the least expensive method of clearance. Climatic events such as the 1997–1998 El Niño in Southeast Asia and the 2005 Amazonian drought (caused by an increase in the North Atlantic sea surface temperatures) are used as opportunities to

Table 5 Percentage of total area change by individual change processes in the tropics (1990–2000)[a]

Driver	*% of total area change*		
	Africa	*Asia and Pacific*	*Americas*
Direct conversion to small-scale agriculture	59	13	13
Direct conversion to large-scale agriculture	12	29	47
Intensification of shifting cultivation areas	8	23	4
Expansion of shifting cultivation	4	9	2
Gains in forest area and canopy cover	8	6	6
Other	9	20	28

[a]Includes livestock and clearing for industrial tree plantations.

Source: Reproduced from Food and Agriculture Organization (FAO) (2001) *State of the World's Forests*. Rome: FAO. Available at: http://www.fao.org/forestry/49666/en/

burn forests, often illegally, though many such forests would already have been subjected to prior degradation.

Cattle Ranching and Livestock Grazing

Intensive clearing of forest land in South and Central America occurred as a result of the expansion of cattle ranching, which was economically attractive due to relatively low risk, little labor, well-established markets, and the availability of various government subsidies. Cattle ranching expanded greatly initially in response to the opening of large markets in North America but has been sustained by the development of the domestic markets for beef. In Colombia, the area of cleared forest lands used for grazing was 4.9 million ha at the end of 1800 and expanded to 23.8 million ha by 2000. In Central America the area of pasture is estimated to have increased from 3.9 million ha in 1955 to 13.4 million ha in 1995, largely at the expense of the region's tropical forests. Ranchers cleared forest land by purchasing it directly and employing laborers to clear the forest cover or by purchasing land or dispossessing it from slash-and-burn farmers that was then converted to grasslands. The shifting cultivators would move deeper into the forest to repeat the cycle. Thus, deforestation in Latin America due to ranching is also associated with slash-and-burn agriculture and land speculation.

Livestock grazing can be a serious cause of deforestation in Africa where livestock herds exceed the carrying capacity of the area. Such pressure is acute in the drier tropics such as the Sahel region of Africa and in the Middle East, where large flocks of sheep and goats are maintained. The history of deforestation around the Mediterranean is linked to grazing regimes, especially by goats, but past simplistic assumptions have given way to the recognition that climatic, sociological, and agrarian factors have also contributed to forest clearance.

Infrastructural Development and Industrial Projects

During the 1970s and early 1980s, development of the Amazon, largely illustrative of similar strategies throughout the tropics, was actively encouraged by the Brazilian government through the building of roads, tax incentives and subsidies, massive resettlement programs, and large-scale development programs. The Trans-Amazonian highway opened up millions of square kilometers of previously inaccessible forest to colonization and allowed further expansion of the cattle industry. Such roads improve access to poorly developed areas and therefore tend to increase the adjacent land value for nonforest uses and encourage land speculation and deforestation.

Logging roads in Asia also facilitate deforestation by allowing access to illegal loggers who might deforest an area that is otherwise merely degraded through selective logging.

Mining and oil exploration had widely publicized deforestation impacts. Large mines consume large amounts of wood for fuel, whereas oil companies clear forests to create lines for exploration, which are later used by colonizers following the exploration teams. Other large-scale development projects include hydroelectric power plants, whose reservoirs flood forests and transmission line paths are cut through the forest causing permanent loss of forest cover.

Crop and Tree Plantations

Much forest has been cleared for commercial plantation crops such as rubber, oil palm, and the crops of cocoa, coffee, and tea. Demand for land for agriculture and plantations continues to grow, fueled by the expanding human population and increasing affluence. To some extent, current climate mitigation policies have contributed to the problem by promoting the expansion of biofuels, which further increases demand for land. Increasing affluence has created a higher demand for meat and dairy products, which is often met by the conversion of forests to pasture or for the production of animal feedstocks.

Conversion of natural forest to tree plantations has also occurred, although currently such practice is deprecated and is less common. Indeed, a key principle of good forest stewardship is that forest plantations are only located on already cut-over, abandoned, or wasteland and in this way can actually help deflect pressure away from natural forest. Illustrative examples of significant conversion of natural forest formations to forest plantations include the following: conversion to exotic pines, *Gmelina*, and eucalypts at Jari in the drier east Amazon for pulp and paper production, conversion on the hills around Bulolo in Papua New Guinea – albeit with the species of *Araucaria* native to the area – for high-quality softwood veneer logs for making plywood, and clearance of miombo and savanna woodland for pine plantations in Zambia for industrial constructional material.

Fuelwood Collection and Charcoal Production

Fuelwood accounts for approximately 80% of all wood use in developing countries. Dependence on fuelwood is expected to decrease gradually with the introduction of electricity, kerosene, and propane, but to continue well into the twenty-first century. Fuelwood collection as an agent of deforestation is particularly marked near urban centers and in villages in which continuous collection results in the gradual degradation and eventual deforestation of accessible areas. This is critical in the dry tropics, along with domestic use of wood for other uses such as construction and fencing material.

Logging

Logging affects a substantial proportion of the world's tropical forests (**Table 6**). Most tropical logging consists of short-term exploitation of timber products with little or no concern for the future potential of the forest to provide additional resources and revenues. Indeed, estimates for developing countries in 2000 showed that only 123 million ha, or 6% of the total forest area, was covered by "a formal, nationally approved forest management plan covering a period of at least five years", in contrast to developed countries where the equivalent figure is 89% (FAO 2001). This is due to the insecurity of tenure and short concession periods, as well as poor governance, corruption and lack of capacity to enforce forest laws. The length of concessions is very often short term, sometimes less than 10 years, and in the absence of a long-term commitment, logging companies have little incentive to invest in long-term forest management. Concessions are also granted for timber with little regard for the other resources provided by the forest and the impact of the logging on local people.

Although the intensity of logging in the tropics is usually low, removal of only 10% of the timber trees can result in

Table 6 Approximate geographic extent of contemporary forest cover, deforestation, and selective logging by region in the humid tropical forest biome. Values are in km^2, with percentage of biome extent also given

Region	*Total biome extent*	*Area with 0–50% forest cover, 2005*	*Area with 50–100% forest cover, 2005*	*Forest area cleared 2000–2005*	*Selective logging (2000s)*
Africa	2,918,511	1,085,941 (37.2%)	1,832,569 (62.8%)	14,972 (0.5%)	561,153 (19.2%)
Asia/Oceania	7,191,529	5,234,293 (72.8%)	1,957,236 (27.2%)	93,955 (1.3%)	1,777,963 (27.2%)
Central America/ Caribbean	685,840	501,415 (73.1%)	184,425 (26.9%)	9,687 (1.4%)	36,097 (5.3%)
South America	8,826,966	3,194,632 (36.2%)	5,632,334 (63.8%)	156,001 (1.8%)	1,603,166 (18.2%)
Total	19,622,846	10,016,282 (51.0%)	9,606,564 (49.0%)	274,615 (1.4%)	3,978,379 (20.3%)

Source: Reproduced from Ghazoul J and Sheil D (2010) *Tropical Rain Forest Ecology, Diversity, and Conservation*. pp. 516. New York: Oxford University Press.

damage to 55% or more of the remaining trees. Even so, logging operations usually result in degradation of the forest rather than its complete elimination. Deforestation does occur along logging roads, where forest is cleared for several meters on either side of the road, allowing the sun to dry the road surface. Poorly designed roads can result in severe erosion and landslides as well as facilitate movement into the area by illegal loggers, slash-and-burn cultivators and land speculators.

Illegal logging remains a substantial problem across tropical forest regions, despite estimates that from 2000 to 2010 the extent of illegal logging fell by 50 to 75% in major tropical timber producer countries. Illegal harvesting is thought to account for 35–72 % of logging in the Brazilian Amazon, 22–35 % in Cameroon, 59–65 % in Ghana, 40–61 % in Indonesia, and 14–25 % in Malaysia (Lawson and MacFaul, 2010). Worldwide, more than 100 million cubic metres of timber are illegally harvested each year, leading to the degradation and possible eventual destruction of five million hectares of forest.

Fire

Serious losses in forest cover in Southeast Asia and South America have been reported as a result of forest fires in 1997 and 1998. The causes of these fires are new large-scale commercial agricultural projects (including plantations) and shifting cultivation. The fires were exacerbated by the dry, coarse, woody debris left after logging operations and the very dry climatic conditions caused by the El Niño phenomenon. The area of forest consumed by fires in 1997 and 1998 has been estimated at more than 20 million ha (Cochrane 2003). Extensive fires in tropical moist forests have been previously associated with the El Niño phenomenon, as in 1982, but the underlying causes are clearance of forest to establish plantations of oil palm, pulpwood, and rice, and, in South America, cattle pastures and shifting cultivation.

Consequences of Deforestation

Biodiversity

Perhaps the most severe consequence of the loss of tropical forest cover is the associated loss of biodiversity. It is not known how many species become extinct through the loss of tropical forest habitats, but estimates range up to a yearly loss of 50,000 species. Small forest fragments that often remain after clearance usually support only a small fraction of the original forest biodiversity.

Ecosystem Services

The loss or disruption of ecosystem services following deforestation is well documented, particularly for regulation of water flow through watersheds and maintenance of soil structure and fertility. Watersheds become susceptible to extremes of water flow following removal of tree cover, with which erosion of topsoil is associated. Thus, forest clearance often results in erratic water flow with periodic droughts punctuated by flood events, which put communities living downstream at risk. Loss of water purity and topsoil through erosion deprive rural communities of safe potable water and fertile soils. Tropical forest soils rapidly become marginal for farming following loss of vegetation cover due to compaction, loss of organic material, and leaching of soil nutrients. Sedimentation of coastal fisheries can deprive communities of their sources of food and income far from the sites of deforestation.

Clearance of forests disrupts these soil-conserving roles and subsequent rapid soil erosion is frequently observed, as the muddy red, yellow, and brown streams and rivers of so much of the tropics testify. Currently, over 300 million hectares of forests have been designated for soil and water conservation (FAO 2009).

Land degradation through loss or overuse of forest cover in dry forest zones can result in desertification. Firewood consumption is a significant contributing factor, but this is further compounded by overgrazing and overcultivation. The combination of these factors is resulting in widespread desertification of the Sahel in Africa, causing massive losses of topsoil and soil fertility and declines in productivity.

Global Climate Change

It is generally agreed that global climate is warming, with an increase in average global temperatures of approximately 0.3 °C each decade, due to the emissions of carbon dioxide

and other greenhouse gases from burning of fossil fuels mainly in industrialized countries. However, it is estimated that approximately 20% of total CO_2 emissions are a consequence of deforestation and forest fires as the carbon stocks in forests are liberated (IPCC 2007). The consequences of global warming are difficult to predict, but increased drought and desertification leading to crop failures and reduced productivity seem likely in many tropical regions. Higher temperature and reduced rainfall coupled with more frequent and severe El Niño events are predicted to result in increased frequency and intensity of fires in even moist low-land tropical forests. Geographic and altitudinal shifts in the distribution of vegetation types can also be expected.

Economic Losses

Annual loss of forest capital is estimated at $45 billion (World Commission on Forests and Sustainable Development, 1999). Destruction of forest represents opportunity costs through the loss of future revenues and employment that might be derived from the sustainable management of the forest for timber and nontimber products as well as potential income from ecotourism and other such enterprises. Economic losses can be much greater if the costs of lost ecosystem services are factored into the equation. If the nontimber benefits and services tropical forests provide are even partially costed, then the value per hectare of extant forest increases enormously. This is now being quantified. For example, the previously ignored carbon sequestration role of forests now has potentially very high financial values.

Social Consequences

Deforestation leads to the loss of traditional lifestyles associated with forested regions. Individual and collective rights to forest land have frequently been ignored by developers, ranchers, transmigrants, and governments. The loss of traditional lifestyles, customs, and religious beliefs of indigenous communities in the Amazon and Borneo has paralleled the loss and degradation of forests in these regions. Approximately 200 million people in the developing world live in or at the forest margins and wholly depend on forests for their livelihoods. Rarely are their concerns heeded or their rights even acknowledged.

Prospects for the future

Agriculture

The main cause of deforestation worldwide is the expansion of agriculture in its various forms. It is estimated that over 5 billion ha of new agricultural land will be needed to sustain global demands by 2050. This more than doubles the current agricultural land area. Although much of the new agricultural land will be derived from shrublands, pastures, and the intensification of subsistence plots, it is expected that a large proportion of the new agricultural land area will be derived from the conversion of logged and degraded forest as well as intact forest.

State-supported investment in deforested marginal lands, fair credit schemes, and educational development are all needed to improve the efficiency of land use. Greater productivity from and improved use of existing agricultural land will lessen the pressure for clearance of new forested lands and will promote private investment for currently occupied land, which in turn encourages sustainable use.

Protected area systems and REDD

Efforts to reduce the rate of tropical forest loss focus on protected area systems, improving logging practices, and increasing the emphasis on the valuation of ecosystem services that forests provide. Emerging policies concerned with reducing carbon emissions from deforestation activities seek to provide financial incentives to nations and land owners to maintain forest cover. One such policy instrument is Reducing Emissions from Deforestation and Forest Degradation (REDD), which has received wide political and scientific support, but is yet to be implemented globally beyond the voluntary pilot study stage.

Protected area systems offer some measure of protection from conversion, although designation does not necessarily protect forests from encroachment. Currently, around 11% of the world's forests are designated as protected areas. To be successful, protected area systems need to be supported by adequate funding and legislation and managed by strong institutional departments. Forest protection in the tropics must often be developed with the cooperation of the local communities that use forest resources as otherwise alienation of these communities undermines adequate protection.

Logging Practices

Improved logging practices that minimize damage to forests and, which reduce the probability of subsequent degradation leading to forest loss are well-known but rarely implemented due to the perceived high cost of doing so. National standards for logging operations, where they exist, are often flouted and enforcement is weak. There has been considerable progress in recent years in legality verification, although much of this has been to minimum legal standards, which often does not include an assessment of actual harvesting practices. Certification schemes have been widely introduced to encourage producers to adopt sustainable practices, which after verification, may be certified as such with the expectation that certified products have a competitive advantage and procure higher market prices. There are several difficulties associated with certification as a means to reduce forest degradation and deforestation, the principal one being the need to ensure that certified wood is kept separate from noncertified wood from the source through the chain of custody and to the retailer. The success of certification also depends on the market demand for certified forest products, which is currently greatest in Europe. The domestic markets of timber-producing tropical countries account for the majority of tropical timber consumption in these countries, in which the market demand for certified wood remains low.

Development Projects

Neo-liberal economic policies and trade liberalization provide a system framework that promotes development of forest areas, often by large-scale infrastructure, industrial or energy generation programs. Although these create economic opportunities and have undoubted national economic benefits, they also incur substantial costs of forest destruction. The development of infrastructure, mining, urbanization, and industrialization projects is often supported by bilateral and multilateral donors, and a careful redistribution of these financial flows could benefit forests and forest peoples.

Tree Plantations

Tree and forest plantations are not substitutes for natural forest, but appropriately sited they can alleviate deforestation pressures. Plantations on already degraded land can be an excellent source of industrial-grade timber, pulpwood for paper making, and fuel – solid firewood and charcoal. Thus, they provide an alternative source of such products. Globally, forest plantations cover 140 million ha of land, about half of which is located in the tropics and subtropics.

Small-scale plantings on farms and around villages helps satisfy domestic demand for building and fencing materials, fuel, and local constructional uses. The estimated 20–30 million ha of such "woodland" further diverts pressure away from natural forest.

Profitable plantations risk undermining the perceived value of natural forest and may accelerate conversion of these to plantations. If the primary cause behind deforestation is the conversion of land for agricultural development, it is unlikely that plantations will reduce deforestation. Tree plantations do, however, offer an opportunity to reduce the pressure on natural forests if the latter are exploited primarily for wood products.

Conclusion

Deforestation is largely a tropical issue. The quality of data on deforestation rates has improved but remains poor and is the source of contention and debate. The causes of deforestation are complex and multi-faceted, although socioeconomic factors and trade are foremost. Deforestation impacts the environment through loss of biodiversity and disruption of ecosystem processes and the economy by affecting ecosystem services and because of inefficient squandering of resources. Perhaps the most tragic consequence is the loss of traditional beliefs and customs and the displacement of forest-dependent communities. Solving the problem of deforestation requires a suite of strategies that include establishing an effective and global protected area network to preserve forest biodiversity and the genuine implementation of good practice guidelines in the use of forest resources. Ultimately, a reformulation of policy and a change in the attitudes of decision makers are needed to ensure that forests are correctly valued economically, socially, and biologically.

See also: Agriculture, Industrialized. Agriculture, Sustainable. Biodiversity-Rich Countries. Cattle, Sheep, and Goats, Ecological Role of. Desertification. Forest Ecology. Land-Use Patterns, Historic. Reforestation. Timber Industry

References

Food and Agriculture Organization (FAO) (2001) *State of the World's Forests*. Rome: FAO. Available online at: http://www.fao.org/forestry/49666/en/

Food and Agriculture Organization (FAO) (2009) *State of the World's Forests*. Rome: FAO. Available online at: http://www.fao.org/forestry/49666/en/

Ghazoul J and Sheil D (2010) *Tropical Rain Forest Ecology, Diversity, and Conservation*, pp. 516. New York: Oxford University Press.

Lawson S and MacFaul L (2010) Illegal Logging and Related Trade Indicators of the Global Response. Chatham House (The Royal Institute of International Affairs). Available online at: http://www.illegal-logging.info/uploads/CHillegalloggingpaperwebready1.pdf

Desert Ecosystems

Exequiel Ezcurra, University of California, Riverside, CA, USA
Eric Mellink, CICESE, Carretera Ensenada-Tijuana, Ensenada, BC, México

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by James A. MacMahon, vol. 2, pp. 37–59, © 2001, Elsevier Inc.

Glossary

Alluvial fan Fine sediments deposited by water that form a conical fan when laid down by streams or rivers.
Annual A plant, sometimes called an ephemeral, that survives year to year as a seed rather than as a "plant."
Convergence or convergent evolution Species develop similar behavioral, physiological, or morphological characteristics in response to living in similar environments in different places.
Desertification Process of changing a nondesert area into what appears to be a desert, regardless of the climate. This process often has a negative connotation relating it to overgrazing by domestic animals, or general human misuse; however, factors, not related to humans, may cause desertification.
Hadley cell A pattern of air circulation that causes deserts at about 30° N and 30° S of the equator. Sun heats the equator, air rises, pressure decreases, moist air moves north and south, cooling as it goes and falls onto the spinning earth, drying as it falls. The falling dry air increases pressure preventing the incursion of moist air.
Life-form A system of classification of the plants in communities based on the position of their perennating structures relative to the ground.
Mesic Moist, but not wet, environment.
Perennating structures The tissue that carries a plant into a new growing season. For example, buds of trees, shrubs, etc., bulbs of some plants (e.g., tulips) or seeds of annuals where the "plant" dies each year.
Physiognomy The general appearance of an area. For example, if you are in any desert in the world it has a certain recognizable "look," regardless of species composition.
Saltation Sand particles jumping across the surface due to wind movement.
Torpor A condition, usually in birds and mammals, when the body temperature falls a few degrees below ambient for a short period of time. Differs from hibernation in (a) the amount the body temperatures fall and (b) the length of the resting state.
Xeric A dry environment.

The Desert Biome

Seen from outer space, the earth is spotted by a series of conspicuous ochre areas that run in an east–west direction, parallel to the equator at 25–35° latitude in both the northern and southern hemispheres (**Figure 1**). The reddish-brown blotches are the midlatitude, vegetation-barren deserts that lie some 2000–4000 km away from the equatorial rainforests. In the northern hemisphere, the succession of midlatitude subtropical deserts is formed by (1) the Mojave, Sonoran, and Chihuahuan Deserts in North America, (2) the Sahara's immense swath in Northern Africa and the Somali-Ethiopian deserts in the Horn of Africa, and (3) the deserts of Asia, including the Arabian, Mesopotamian, Persian, and Thar deserts that cover all the way from the Middle East into Pakistan and India, as well as the Central Asian deserts in Uzbekistan, Turkmenistan, and the Taklimakan and Gobi deserts in China and Mongolia. In the southern hemisphere, the chain is formed by (1) the Atacama, Puna, and Monte Deserts in South America, (2) the Namib and the Karoo in southern Africa, and (3) the vast expanse of the Australian deserts (Allan *et al.*, 1993; McGinnies *et al.*, 1977; Pipes, 1998; Ricciuti, 1996).

There are many criteria to define a desert but perhaps the most important one is aridity – the lack of water, as the main factor limiting biological processes. Aridity can be measured through several approaches, one of the most common ones is through the "Aridity Index" (**Figure 2(a)**), the ratio between mean annual precipitation (*P*) and mean annual potential evapotranspiration (*PET*, the amount of water that would be lost from water-saturated soil by plant transpiration and direct evaporation from the ground, calculated as a function of temperature and insolation; Thornthwaite, 1948). Arid and hyperarid regions have a *P*-to-*PET* ratio of less than 0.20; that is, rainfall supplies less than 20% of the amount of water needed to support optimum plant growth (UNEP, 1997; FAO, 2004). Aridity is highest in the Saharan and Chilean-Peruvian deserts, followed by the Arabian, East African, Gobi, Australian, and South African Deserts, and it is generally lower in the Thar and North American deserts. Aridity index values in the different deserts in the world fall within the arid and hyperarid categories, which can be used as a good first approximation to the boundaries of the Desert Biome (**Table 1**).

A second approximation to map the Desert Biome is based on biological adaptation (**Figure 2(b)**), and considers all the global ecoregions that harbor desert vegetation (identified by the xerophilous life-forms and the general desert-adapted physiognomy of the dominant plants). An adequate source for this is the World Wildlife Fund's (WWF's) ecoregions of the world report (Olson *et al.*, 2001).

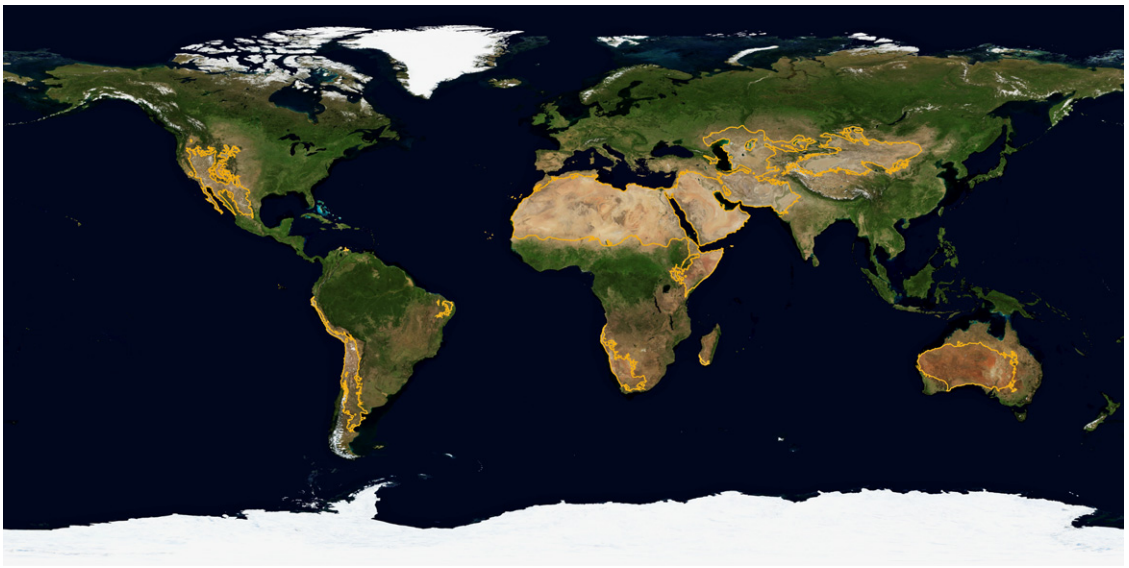


Figure 1 The Desert Biome, enclosed by a yellow line, is clearly discernible in this satellite image of the earth, both north and south of the equator, as a series of ochre-colored patches.

A third criterion to define deserts can be derived from land-cover data in Advanced Very High Resolution Radiometer (AVHRR)-satellite images of the world (**Figure 2(c)**). Using the Normalized Difference Vegetation Index (NDVI) land-cover values, the earth has been classified into different categories (GLOBIO, 2005; USGS, 2005). By this tool, a desert can be defined as an extensive region with little or no vegetation cover (the definition must exclude urban areas, as well as high altitude and polar regions, where low vegetation cover is not due to water limitations).

Each of these approaches may have their own sources of error and the three differ in their definition of what is a desert. However, the three alternative definitions coincide remarkably (**Figure 3**). The Desert Biome, in short, is formed by a set of geographic regions characterized by (1) extremely high aridity, (2) plants and animals showing well-defined evolutionary adaptations to survive in extremely dry environments, and (3) low plant cover and large proportions of bare soil.

Latitudinal Desert Belts

Deserts occur mostly in specific latitudes (25–35° N and 25–35° S of the equator) because of the general thermodynamics of our planet. Solar radiation hits the earth with highest intensity near the equator. Because the earth's axis is tilted 23.5° with respect to the plane of its orbit, during part of the year the zone of maximum solar interception shifts northward, toward the Tropic of Cancer, and during part of the year it moves southward, toward the Tropic of Capricorn. Thus, the warm tropics form a belt around the equator from latitude 23° N to latitude 23° S, where the tropical heat generates rising, unstable air. As it climbs, the air condenses the moisture evaporated from the warm tropical seas and forests, and produces the heavy downpours that characterize the wet tropics. As it moves away from the equator at high altitudes, the air cools again and eventually starts descending in the midlatitudes, some 3000 km away from the equator both north and south. This closed, low-latitude pattern of atmospheric circulation north and south of the equator is called the "Hadley cells." The air masses compress and heat in their descent and, having lost their moisture during their tropical ascent, they become extremely dry. Thus, by contrast with the equatorial forests, the midlatitude arid fringes that run alongside the tropical belt have a more stable atmosphere; these are the "horse" latitudes, where calm, dry air often dominates. Additionally, because of the stable atmosphere, not only winds are slack but also rainstorms seldom develop. For this reason most of the world's large deserts occur along the belt that separates the tropics from the temperate regions (Goudie and Wilkinson, 1977).

Inland Deserts

The sheer size of continents may be in some cases a direct source of aridity. Because most of the water in the atmosphere is ultimately derived from evaporation from the seas, an aridity gradient develops in large continents when the land closer to the sea receives a larger share of this ocean-derived water and, as air moves inland, it gets depleted of moisture and precipitation drops. Thus, regions lying deep within a continent may become deserts simply because air currents reaching them have already traversed vast land distances and lost most of the moisture they originally carried. Continentality is a major factor driving arid conditions in the Taklimakan and Gobi deserts of central East Asia, the central deserts of Australia, the Great Basin in North America, and the Monte Desert of South America (**Figure 4**). Continentality not only affects the precipitation regime of inland deserts, but also drives the temperature variability. Because of the higher thermal inertia of oceans compared to continental land masses, coastal deserts have a much lower temperature gradient throughout the year than deserts far from the sea (**Figure 5**).

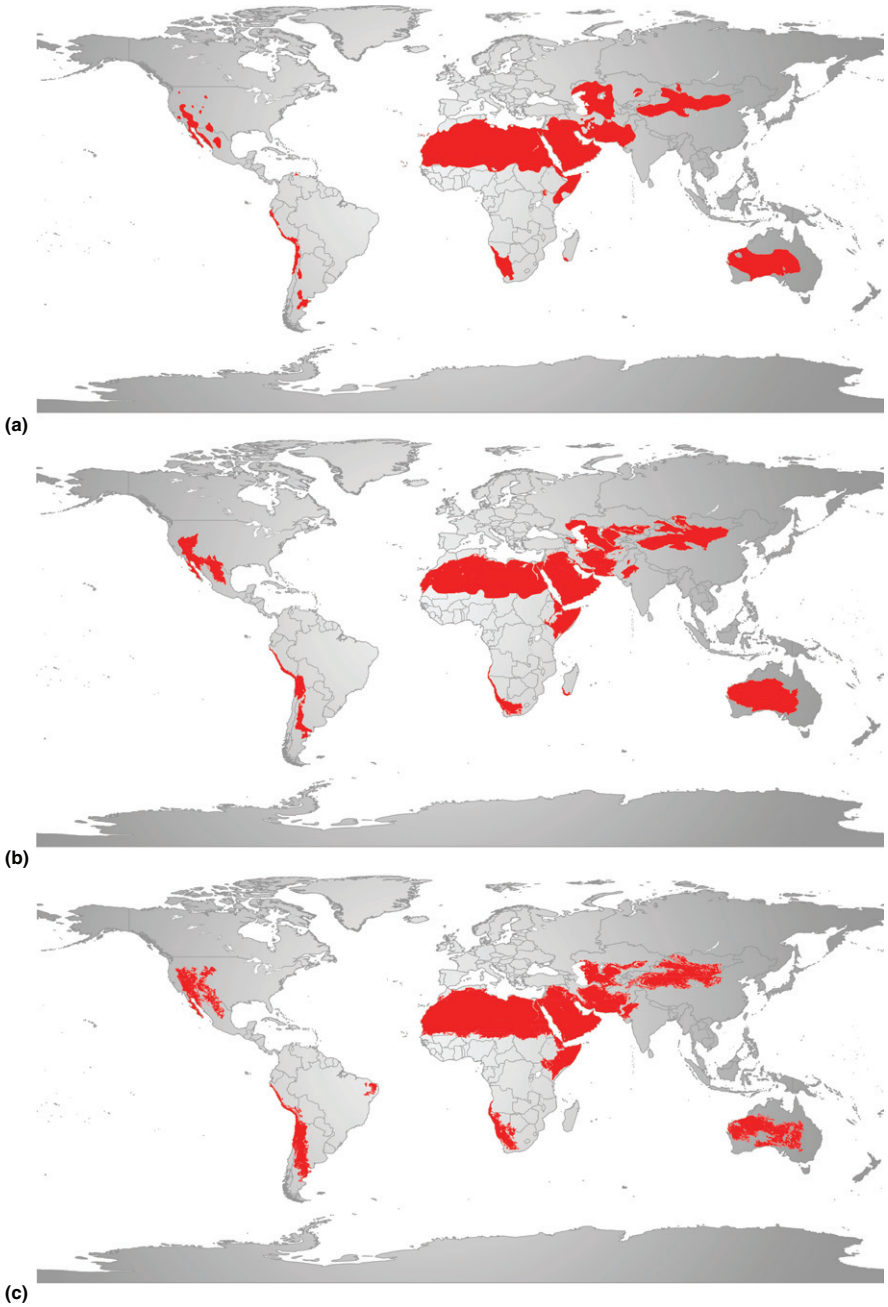


Figure 2 The Desert Biome defined by three alternative criteria: (a) Arid and hyperarid regions of the planet as defined in the World Atlas of Desertification (UNEP, 1997); (b) global ecoregions of the world (Olson *et al.*, 2001) that harbor desert vegetation; and (c) contiguous regions of extremely low vegetation cover derived from AVHRR satellite images of the world (USGS, 2005).

Coastal Deserts

The latitudinal explanation, however, does not account entirely for the global distribution of deserts. Around the mid-latitudinal belt, the western coastal fringes of continents are normally occupied by deserts, while the eastern sides are covered by forests. In North America, the western coast of the continent is occupied by the dry deserts of Baja California, while the east harbors the humid tropical ecosystems of Florida. In South America, the Atacama Desert in the west is matched by Brazil's luxuriant Atlantic Forest at the same latitude on the Atlantic coast. The Namibian Desert, on the western coast of southern Africa, is paralleled by lush forests on the eastern coast.

The reason for this has to do with the global circulation of ocean currents: gravitation from the sun and the moon pulls air and water on the earth's surface and tends to make them lag behind relative to the earth's rotational movement. The gravitational drag is greatest in the equator, where the centrifugal speed of the earth is fastest. Thus, as the earth turns,

ocean currents and winds flow in the equator from east to west, forming the equatorial currents and the easterly trade winds. As the west-bound surface waters move away from the continents, they pull temperate currents toward the equator along the eastern side of the large oceans, and drag cold, nutrient-rich waters to the surface inducing a cool, stable coastal atmosphere with little evaporation from the sea and very low rainfall other than morning fogs. In the coasts neighboring these oceanic upwellings, typical coastal fog deserts tend to develop, forming some of the driest ecosystems on earth. Thus, the large-scale circulation of the ocean is the main reason why coastal deserts are always found in the west coasts of continents, such as the Namib in Africa (**Figure 6**), Atacama in Chile, the Atlantic Coastal Desert of Morocco, or the deserts of Baja California (**Figure 7**).

Rainshadow Deserts

Topographic heterogeneity also contributes to the formation of deserts. Moisture-laden winds decompress and cool in their ascent as they climb continental mountain ranges, condensing fog and drizzle that feed montane cloud forests. Once the winds pass over the mountain divide, they start compressing and warming-up in their descent, but, having left behind their original moisture, they become hotter and dry. Thus, while the windward slopes of many mountain ranges are covered by cloud forests, the leeward part, known as the "rain shadow" of the mountains, is covered by arid scrub. The rain shadow effect is largely responsible for many arid lands that seem to defy the rule that deserts are found in the earth's midlatitudinal reaches. Mountain rain shadows are the cause of tropical drylands such as the Sechura Desert in Peru and Ecuador, the Caatinga scrub in equatorial Brazil, or the Tehuacán Valley desert in southern Mexico. Similarly, they are also responsible for some high-latitude cold drylands, such as the large deserts of Central Asia, the Great Basin, and Patagonia.

Table 1 The hyperarid and arid regions of the world – defined as those areas with an aridity index (*P*/*PET*) lower than 0.2 – cover in total some 36.2 million km^2, and occupy almost 20% of the terrestrial surface of the planet. Potential evapotranspiration (*PET*) is calculated from Thornthwaite's equations as a function of mean monthly temperatures and mean monthly number of daylight hours, while precipitation (*P*) is measured directly from weather stations

Classification	*Aridity Index (P/PET)*	*Area (km^2 × 10^6)*	*Area (%)*
Hyperarid	<0.05	10.0	7.5
Arid	0.05–0.20	16.2	12.1

Desert Landforms

The landforms of deserts, like those of high mountains and the polar areas, are much more visible than those of more vegetated landscapes. Bareness also allows much more active surface processes in all these areas, but in different combinations. Deserts suffer more wind erosion than any other environment, and, if slopes are steep, may also experience very fast water erosion during the short but intense rain pulses.

Desert landscapes come in two categories: (1) "shield" deserts and (2) "mountain-and-basin" deserts (Cooke *et al.*, 1992; Mabbutt, 1977). Shield deserts have developed on very ancient rock basements that have been folded, faulted, and hardened by heat and pressure over millions of years. The Sahara, the Arabian deserts, the southern African deserts, and the Australian deserts are in this group. In many places (as at Uluru in Australia), granite rocks formed originally deep within the earth have been uplifted by tectonic forces and

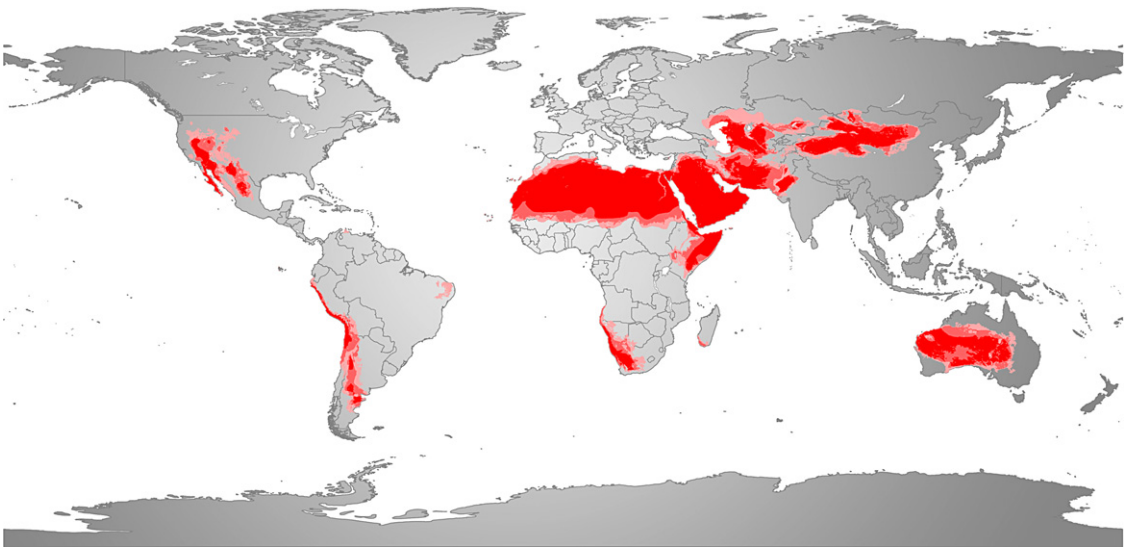


Figure 3 The Desert Biome defined by three criteria in combination; that is, aridity, ecoregional physiognomy, and soil cover (see **Figure 2**). The intensity of the red color indicates congruency between the three criteria: Areas in intense red correspond to regions where the three criteria coincide, areas in intermediate color highlight regions where two criteria coincide, and areas in pale red show areas where only one criterion operates.

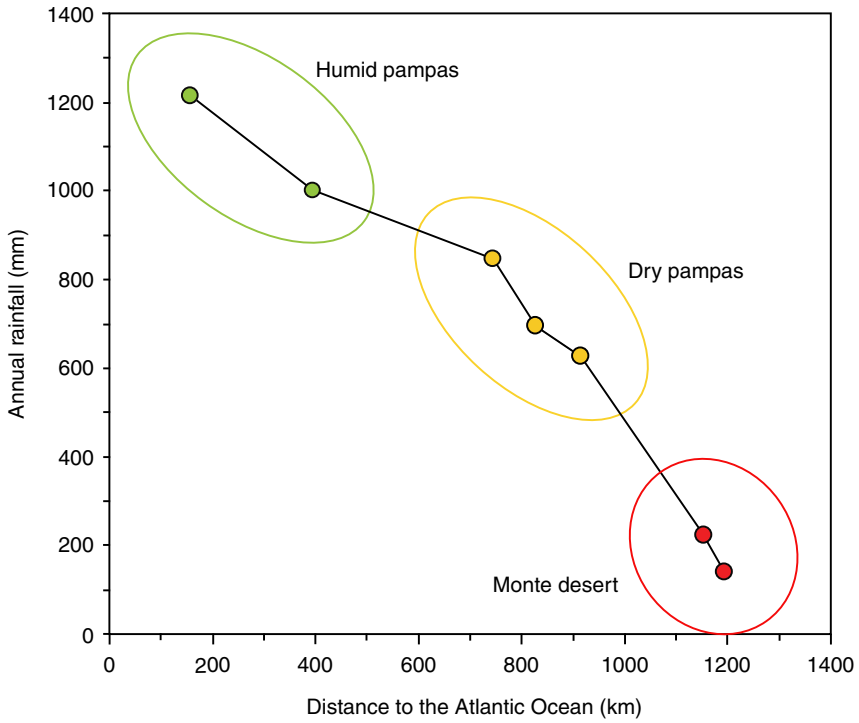


Figure 4 Continentality and aridity: Following a transect along the flat Argentine Pampas from the Atlantic coast to the foothills of the Andes, mean annual rainfall decreases as a function of the distance to the coast at a mean rate of approximately 1 mm of rainfall per km. Because the high Andean Cordillera stops any moisture coming from the Pacific, the continental distance to the Atlantic coast is the single best predictor of precipitation. The colors indicate local weather stations in the humid Pampas (green), the dry Pampas (yellow), and, finally, the Monte Desert (red) in the extreme west.

unearthed by erosion to form steep-sided hills. The basement has been folded into gently sloping swells and basins, and the basins have been filled over millions of years with sediments eroded from the swells. Shield deserts contain ample supplies of groundwater accumulated during past geologic ages, as in the northeastern and southern Sahara, and in Australia; and in some areas they are also rich in oil. Here and there, recent volcanic rocks have overflowed at the surface, as in Ahaggar and Tibesti mountains in the central Sahara. In their long lives, these landscapes have experienced many different climates, partly because they were moved by continental drift, and many features formed in the different climates survive. There are even ancient glacial features in parts of Arabia and the Sahara; there are many ancient river gorges, and ancient wet tropical soils like silcretes or laterites, in what is now a dry desert landscape.

Water is the main agent of erosion on the few hills of the shield deserts and cuts deep gullies on their edges. Elsewhere, low topographic gradients make water erosion ineffective, and the main erosive force weathering these plains is the wind. Great plumes of dust travel from the Sahara toward Europe, southwest Asia and even the Americas and the Caribbean, removing more sediment than do rivers in the same area. With the dust winnowed out by winds, sand is left behind, and most of it collects in dunes, which cover 20–30% of these desert shield landscapes. The median size of these "sand seas" is 123,000 km^2, but the larger cover more than 300,000 km^2.

In contrast, deserts in the much more recently folded and faulted rocks of the earth's active tectonic belts are mountain-and-basin deserts, in which up-faulted mountains alternate irregularly with down-faulted basins. The American deserts, both North and South, belong mostly in this category, as well as the deserts of central Asia (where some of the basins, however, cover many hundreds of thousands of square kilometers). Water erosion in the mountains (which, because of the rainshadow effect, catch more rain than the lower basins) cuts deep, steep-sided valleys and gorges, and takes the debris out to the basins, where the broadening and shallowing of the ephemeral washes (called *arroyos* in the Americas and *wadis* in Northern Africa and the Middle East) causes the coarser debris to be dropped in broad "alluvial fans" (**Figure 8**). Occasional extreme storms may carry huge boulders onto these as well. Further down-stream, the alluvial fans coalesce into a long slope of finer alluvium – the *bajada*. Sand may be winnowed out of the alluvial deposits and form dunes (and in central Asia, even a few sand seas). Only the finest debris (silt and clay) reaches the bottom of the basin, where it is deposited in ephemeral lakes or *playas*, in which the salts carried in the waters also accumulate in large blinding-white crusts.

Desert Soils

The most extensive soils of deserts are the stony, rocky soils of the uplands (*Leptosols*, in the World Reference Base for Soil

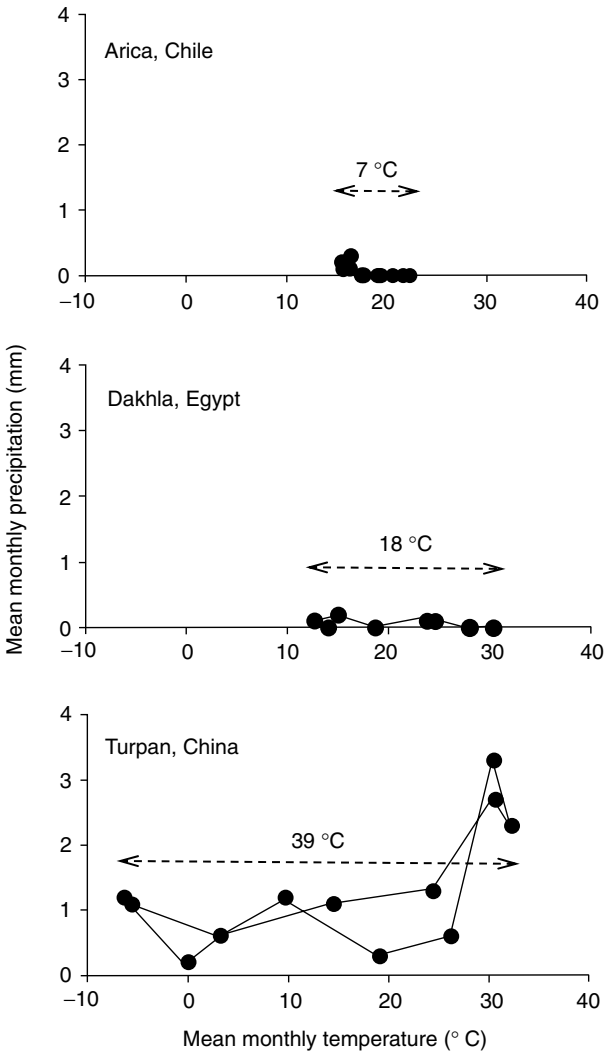


Figure 5 Continentality and temperature range in three hyperarid deserts: data points show mean temperature and mean precipitation for the 12 months of the year. The larger thermal inertia of water relative to dry soil dampens down temperature variations in coastal deserts. As a consequence, annual temperature fluctuations exhibit a range increase as the distance from the sea increases. Arica, in the Chilean Atacama Desert, lies in the Pacific coast and the annual temperature range is only 7 °C; Dakhla, in the Egyptian Sahara, is 520 km from the Red Sea and 590 km from the Mediterranean, and its annual range is 18 °C; while the highly continental area of Turpan, in the north of the Chinese Gobi Desert, is more than 2000 km away from the nearest ocean coast and its annual range is 39 °C.



Figure 6 The consequences of the upwelling of cold, nutrient-rich water from the Benguela current near the coast of Namibia are clearly visible in the greenish-white plume of sulfide emissions produced by decomposing phytoplankton accumulated after the winter productivity peak. To the right, the dunes of the Namib desert bear witness to the dramatic aridity of this coastal desert.

Resources classification), the sedimentary soils of *bajada* slopes and basins (*Regosols*), wind-blown sands (*Arenosols*), and the gravelly sands of arroyos and wadis (*Fluvisols*). These soils have undeveloped horizons and exhibit only a weakly distinct surface layer, hardly different from the underlying parent material. The most characteristic desert soils are those that exhibit significant accumulations of soluble salts, gypsum, calcium carbonate, or silica. These include *Solonchaks*, characterized by high levels of soluble salts which accumulate naturally in closed depressions; *Solonetz*, highly alkaline, sodium rich soils very slippery when wet; *Gypsisols*, or gypsum soils; *Calcisols*, characterized by accumulation of lime, or calcium carbonate, and *Durisols*, showing a hardpan cemented by dry silica.

Soil surfaces in deserts often have one of two unique characteristics – pavements and crusts. Superficial layer of stones, termed desert pavement, gibber, hammada, or reg, occurs in many deserts, caused by the removal of fine materials by wind or surface water, or by the upward migration of stones. The latter can be caused by a number of physical processes, but most likely causes are cycles of freezing and thawing or of wetting and drying. Often the stones of pavements develop a thin patina referred to as desert varnish.

In some deserts (e.g., southwestern US), 75% of the rock may be covered with desert varnish. The varnish is composed basically of particles of clay with iron and manganese oxides, and its color ranges from red to black depending on manganese content. There is considerable debate about the process of varnish formation. Some scientists believe only physicochemical processes are involved, while others promote a biological model involving lichens and bacteria. Both processes probably contribute to the variety of colors of varnishes. Surprisingly, desert varnish may form in as little as one or two decades.

The second unique soil surface in deserts is a thin crust-like layer formed by microscopic organisms. When viewed from

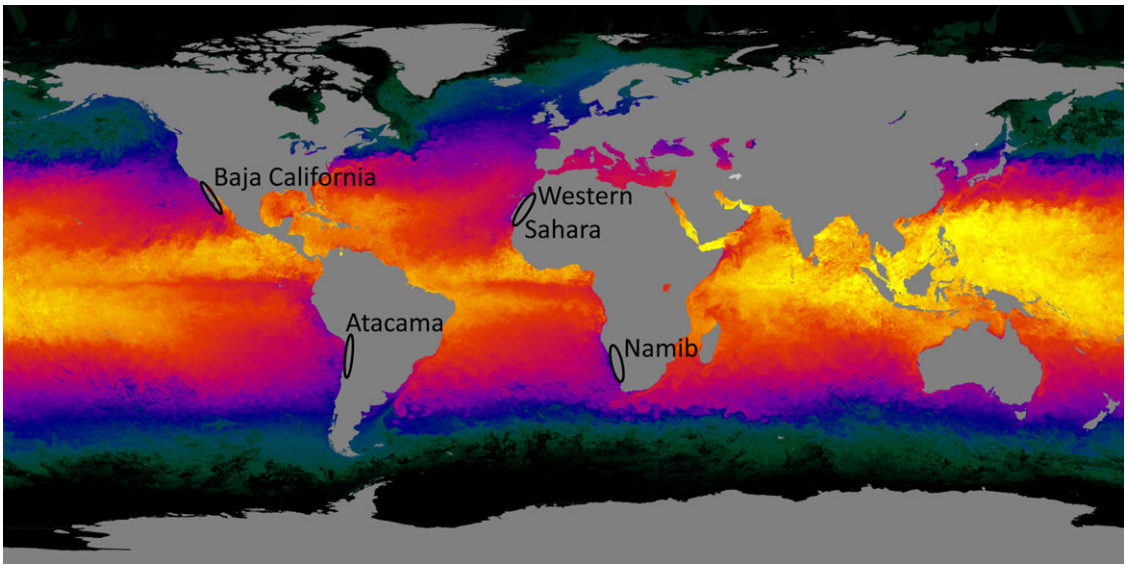


Figure 7 Sea surface temperatures in 2–9 June 2001, measured by NASA's Moderate Resolution Imaging Spectroradiometer (MODIS). Cold waters are black, dark green; and blue; purple, red, yellow, and white represent progressively warmer surface waters. Note the plumes of cold water along the Pacific coasts of North and South America (the California Current along Baja California, and the Humboldt Current along the Atacama Desert in Chile and Peru), as well as along the Atlantic coasts of Africa (the Canary Current along the Western Sahara, and Benguela Current along the Namib Desert), which give origin to the main coastal fog deserts of the world. Other coastal and fog deserts are also found around the Arabian Peninsula, in the Red Sea, the Persian Gulf, and the Oman coast, associated to the strong tidal currents in these narrow water bodies and straits.

the side, these crusts are full of vesicles caused by rainfall or perhaps the expansion of soil gases. Crusts that are formed in association with algae and lichens have increased infiltration rates, whereas others may be somewhat water repellant. Most biological crusts involve cyanobacteria, green algae, lichens, and xerophilous mosses. Because some of these organisms are nitrogen fixers, the disturbance of crusts from off-road vehicles, cattle, and human traffic may decrease nitrogen available to plants and impact soil fertility.

Desert Climates

Rainfall Pulses

In deserts, rainfall events trigger short periods of high resource abundance which, despite the overall scarcity of rain, can saturate the resource demand of many biological processes for a brief time. Furthermore, because rainstorms are also frequently very localized, deserts are extremely patchy environments in their resource availability, both in space and in time. Thus, although deserts are characterized often by their mean climatologic conditions (as in the case of the aridity index described previously) they are really driven by a succession of short pulses of abundant available water against a background of long periods of drought. Rainfall pulses are the driving force structuring desert ecosystems, and plants and animals have developed very specific adaptations to cope with ephemeral abundance, especially with regard to growth, population dynamics, and the cycling of organic matter and nutrients (Sher *et al.*, 2004). Within a desert, rainfall events may vary significantly from one pulse to the next. Some spells may occur in winter, others in summer; some events may bring very little precipitation, others may bring intense showers; and the period between pulses may also vary substantially.

Each desert organism's response threshold is often determined by its ability to make use of moisture pulses of different duration and infiltration depth. For example, brief and shallow pulses have an important effect over surface-dwelling organisms with fast response times and high tolerance for scarce resources, such as soil microorganisms. Short precipitation pulses are important to the survival of annual plants, but deep-rooted perennial plants respond to longer, more intense precipitation events. Thus, the diversity of pulses also promotes a diversity of responses in life-forms, migrations, or population cycles in different species. To a large extent, it is the heterogeneity of pulses that drives the surprisingly large biodiversity of desert ecosystems (Chesson *et al.*, 2004).

But, what commands pulses in deserts? Why are seasons, and even decades, so different from each other, and often so unpredictable? Pulse-type variations in desert environments are linked to global atmospheric and oceanic phenomena. Large-scale drivers of regional precipitation patterns include the position of the jet streams, the movement of polar-front boundaries, the intensity of the summer monsoon, El Niño events, and even longer term ocean cycles, such as the Pacific Decadal Oscillation (Loik *et al.*, 2004). Driven by these large-scale forces, the intensity of midlatitude continentality, ocean upwellings, and rain shadows – the major factors modulating the distribution of arid lands – is not constant but varies from 1 year to the next. As a result, the intensity, frequency, and separation of rain pulses at a local scale may vary substantially through time, often in a seemingly unpredictable fashion.



Figure 8 The basin-and-range topographic sequence is clearly visible at the Laguna Salada in the Mexican Sonoran Desert. The upland mountains discharge their runoff onto alluvial fans of the terraced foothills; down-stream, the water flows in the bajada coalescing into larger streams dotted by mesquite and other phreatophytes (plants that extract water from the deep aquifer), and finally the runoff and drainage water with all its accumulated salts arrives in the valley bottom playa, where it evaporates leaving the salts behind. Thus, the playa salt flat is fringed by a belt of halophytes, or salt tolerating plants. Along these desert topographic sequences, there is an inverse texture effect as compared to wetter regions: soil texture becomes finer toward the valley bottom where the clay is transported by water. The hill slopes are rocky, the foothills are stony, the bajadas are formed by sand and loam, and the playas by clay.

The influence of large-scale drivers on local desert conditions was noted many years ago by the fishermen and the farmers of the coastal desert of Peru, who realized that during some years the normally cold waters of the Pacific became warmer. In these years, they noted, the abundance of sardines decreased but abundant rainfall soaked the land and made the desert flourish. Because this phenomenon was normally observed around the month of December (a time of the year in which Christians commemorate the birth of Jesus, the child – *El Niño* in Spanish), they called the phenomenon "El Niño" (technically, "El Niño Southern Oscillation," or ENSO). During El Niño years, the trade winds and the west-bound equatorial currents slow down, and the upwelling of nutrient-rich waters in the coasts of the American Continent decreases or ceases. The sea becomes less productive while the coastal fog deserts of the Pacific Ocean become drenched in the abundant rainfall that originates from the now-warm sea waters (Holmgren *et al.*, 2001; Safriel *et al.*, 2006).

These pulses of abundance and scarcity of resources are a major force in the ecological organization of many deserts of the world. During pulses of bounty, the fragile seedlings of desert plants can germinate, establish, and prepare for long droughts by burying their roots deep into the desert soils. Ephemerals can replenish their seed banks (**Figure 9**); desert toads can reproduce in extraordinary numbers before entering again into their waterless torpor; and granivorous rodents, such as North American kangaroo rats, Australian hopping mice, and African jerboas, can stock up their underground caches. The desert becomes renewed, and ready to face again years or decades of extreme hardship.



Figure 9 During rainy years, the sand dunes of the Gran Desierto – the driest desert in North America – become lush with the growth of ephemeral plants (Patricio Robles-Gil).

Winter and Summer Rainfall Patterns

Two important weather systems bring precipitation to the world's deserts. The horizontal transport by wind of moist air from the sea onto the land (known as advective transport) during winter, when the land becomes colder than the sea, causes atmospheric condensation over the cold continents and generates winter rains. Because this particular pattern of summer droughts and winter rains dominates the coasts around the Mediterranean Sea, the areas of the world that show this type of seasonal variation are called "Mediterranean" regions. In summer, in contrast, a different weather system is the main driver of rainfall pulses in arid lands: As the continents become hot they generate low-pressure centers with rapidly ascending warm air (a phenomenon known as convective transport). As the air ascends the atmospheric mass cools rapidly and condenses large amounts of air moisture, which pour down in the form of summer thunderstorms. In many parts of the world this rainfall pattern is known as the "monsoon."

Most of the large deserts of the world lie between these two weather patterns: The large midlatitudinal deserts often share their boundary with winter-rain ecosystems in their higher latitudes, and with monsoon regions near their tropical reaches (**Figure 10**). The Sahara Desert, for example, is dominated by winter rains in its northern Mediterranean limit, and by summer rains in its more tropical Sahelian border. The Sonoran Desert, in North America, receives mostly winter rains in its northwestern reaches, near the Mojave, and is fed by the Mexican summer monsoon in its tropical southern reaches (Dimmit, 2000). In South America the Monte Desert receives mostly winter rains in its cold Patagonian austral limit

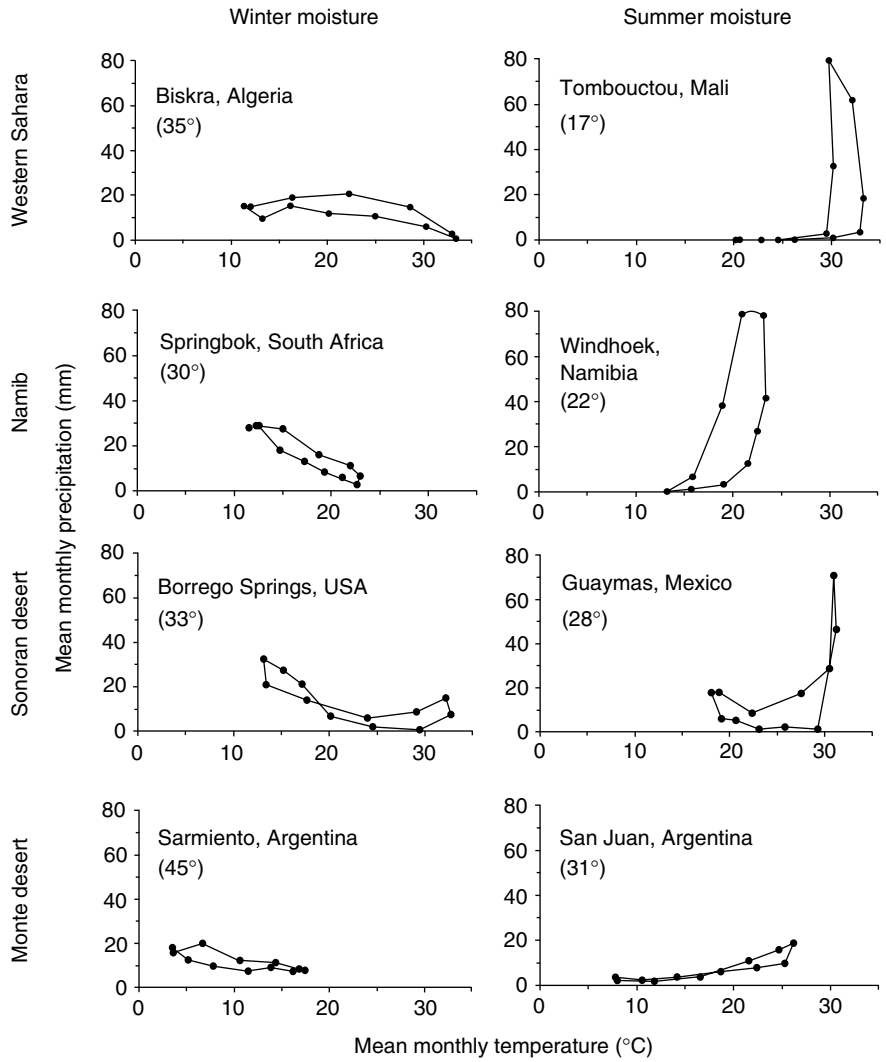


Figure 10 Rainfall seasonality at high and low-latitude extremes in four different deserts: data points show mean temperature and mean precipitation for each month of the year at weather stations north and south of the large midlatitudinal deserts of the world (Western Sahara, Namib, Sonoran Desert and Monte). All stations lying equator-ward of the desert show increased monsoon-type summer rains while all stations lying more distant to the equator, in the deserts' temperate edge, show Mediterranean-type winter rains.

but is dominated by summer thunderstorms near the Tropic of Capricorn (Ezcurra *et al.*, 1991).

Plants in winter-rain deserts have a rather narrow window of opportunity for growth. During winter, moisture accumulates in the soil but the weather is often too cold for plant growth. During summer, temperature may be adequate but the soil is parched and waterless. Only during spring, when temperatures start to rise and moisture is still present in the ground, can plants flourish. For perennial plants, retaining the leaves from the previous year gives the shrubs an early start and a competitive edge during the short spring. But in order to keep the foliage on until the next year the leaves have to survive the dry desert summer, when no rains occur and the soil is bone-dry. To avoid water loss during the hot season, many evergreen desert shrubs have developed small leaves, or leaves with a thick, leathery epidermis, and few, small stomata (the pores through which leaves breathe and fix carbon dioxide). Thus, evergreen shrubs with small and/or tough, leathery leaves are the dominant life-form in winter-rain deserts.

Tough leaves are not the only mechanism through which plants can use efficiently the short desert spring season. The winter-rain deserts of the world are also amazingly rich in short-lived spring ephemerals. These plants survive the scorching summer in the form of seeds, bulbs, or tubers, and quickly sprout during the narrow window of opportunity that the desert spring provides. In contrast with the evergreen shrubs, their leaves are soft and tender, and their stomata are large, allowing the plants to grow very fast when conditions are favorable. One of their most astonishing traits is the lack of any evident adaptations to aridity other than fast growth. Because desert ephemerals complete their life cycles during short periods of abundance, natural selection has favored fast growers and even faster reproducers.

In monsoon deserts, in contrast, rainfall pulses coincide with adequate temperatures for plant growth. Because moisture becomes available at a time in which plants can grow fast, leaf retention is rare and drought deciduousness is a dominant trait in shrubs and trees. In these deserts, perennial plants often show extensive networks of shallow roots, and many plants compete to extract water from the soil immediately after the rain has fallen. Giant fleshy trunks or cactoid succulent stems, adapted to accumulate water, are thus common in summer-rain deserts.

Evolution, History and Biogeography

The Evolution of Modern-Day Deserts

During the last 2 million years – the Pleistocene period – the earth underwent a series of alternate cycles of cooling and warming, induced by variations in the planet's orbit and in the inclination of its axis. During the colder periods – known as the "Ice Ages" – most of the high-latitude regions of the world became covered by massive glaciers, and temperate ecosystems such as cold grasslands and conifer forests shifted toward the equator. Because of the higher proportion of land mass in the northern hemisphere, this phenomenon was more conspicuous in North America and Eurasia, where a large proportion of the continental land mass became covered under the ice sheets.

From the late Pleistocene glacial period to our times significant changes occurred in the climate and ecosystems of the planet. During the Last Glacial Maximum (LGM) period, (25–17,000 years before present (yBP)), large ice sheets and lower concentration of atmospheric CO_2 had given rise to a colder climate and a reduced summer monsoon, resulting in a low global forest cover. The tropical belt narrowed and the deserts moved toward the equator, shrinking in the midlatitudes, where they were replaced by grasslands, semiarid scrubs, open woodlands, or cold steppes. The ancestors of the modern-day desert biota found refuge in what are now dry subtropical habitats, especially in places where arid conditions persisted under the rain shadow of large mountain ranges, or in areas that are now covered by dry tropical savannas, which were then more arid than at present. The last glaciation ended around 15,000 years ago, when the glaciers retreated, giving place to the warm interglacial period that followed: the Holocene, with our current global climate.

As a general rule, during the LGM the high-latitude border of the world's deserts became colder and wetter than they are at present, while the tropical fringe became drier than it is today. Thus, during the LGM large dunes developed in the southern Sahara-Sahel zone and in the Thar Desert of India, while the Mediterranean coast of Algeria and Morocco became wetter and colder. Similarly, during the late Pleistocene the Chihuahuan Desert of North America harbored woodlands of piñon pines, junipers, and oaks, while the tropical dry woodlands of central Mexico became drier, evolving the rich cactus flora that characterizes the region today.

During the warm early- to mid-Holocene (8000–5000 years BP), the global climate that resulted from glacial retreat brought an increase in the intensity of the monsoon throughout the subtropical arid lands. Lake Chad became a freshwater inland lake bigger than today's Caspian Sea, in an area that has become a complete desert. Tropical forests and dry woodlands extended around the equator expanding north and south, while deserts moved into the midlatitudes. During that period, the southern African Sahara and the Sahel were much wetter than today, with extensive vegetation cover, thriving animal communities, and numerous human settlements.

Sometime between 6000 and 5000 years BP, there was again a transition to more arid conditions. Mesic vegetation communities disappeared rapidly, lake levels declined dramatically, and highly mobile pastoralist cultures started to dominate the region replacing sedentary lacustrine and riparian traditions. The Liwa region of the United Arab Emirates, for example, experienced phases of sand deposition that lead to the formation of a megadune (up to 160 m high). In North America, the late Holocene transition toward more arid conditions brought the arrival of Mojave, Chihuahuan and Sonoran desert scrub elements from the south, such as the agaves, cacti, ocotillos (*Fouquieria*), and creosote bushes that characterize these areas today.

An explanation for these climatic variations is that changes in incoming solar radiation, associated with slow, gradual shifts in the Earth's orbit (i.e., changes in the tilt angle, eccentricity of the orbit, and precession of the equinoxes), enhanced the strength of the summer monsoon rains at the beginning of the Holocene. These rains, in turn, increased the extent of vegetation cover and wetlands with two major concomitant effects – a reduction in surface albedo (reflectance) and an increased ability to recycle water back to the atmosphere through evapotranspiration. Both effects helped fuel the monsoons with additional energy and moisture, increasing the summer rains. In Africa, the climate–vegetation system maintained a "green Sahara" climatic regime through the middle Holocene, when a sudden transition occurred to a "desert Sahara," the regime that we know at present. The aridization trend of the mid-Holocene fed back into the deserts themselves by decreasing vegetation cover, reducing local inputs of moisture into the atmosphere, and further increasing the dry conditions.

The Formation of Sky-Islands

When the ice sheets started to retreat, some 20,000 years ago, most of the temperate flora and fauna slowly shifted back into higher latitudes and the Desert Biome gradually expanded across the midlatitudes to its current extent. A subset of the temperate biota, however, stayed behind in the rugged and cool mountain ranges that emerge from the desert plains. Establishing higher-up with each passing generation as the climate warmed, the ice-age organisms were able to persist in the cool mountain environments, where conditions are similar to the ones they had enjoyed in the lower plains during the Ice Ages. As they ascended into the isolated desert mountains, the communities of the desert "sky-islands" became separated from other mountains by harsh desert plains. Like prehistoric castaways, the Ice Age species now survive high-up in the cool refuges of the desert mountains; a biological memory of past evolutionary history subsisting high-up in the mountains like ghosts of climates past. And, because they have been

reproducing in isolation for 15,000–20,000 years, many of their populations have developed unique genetic traits and have evolved into new species or subspecies. Thus, in a similar fashion to evolution in remote oceanic islands, the biota of the desert sky-islands is composed by a large number of endemic taxa and has immense value for biological conservation (Axelrod, 1950).

As the effect of the Ice Ages was more severe in the more land-covered northern hemisphere than in the more oceanic southern half of the globe, most of these Pleistocene montane relicts are found north of the equator. In North America, where the desert relief is highly folded, mountainous sky-islands dapple the central part of the Sonoran and the Chihuahuan Deserts, and of the Great Basin. All these ranges contain endemic pines, oaks, madrones, and chaparral species, remnants of the "Madro-Tertiary" flora, a unique temperate ecosystem that covered much of the now-dry North American deserts during the last 6 million years.

In Africa, similar relict mountains emerge from the harsh Saharan plains: near the Mediterranean coast, the Atlas Mountains in northern Morocco shelter rich pine and oak forests. Further south, the Ahaggar and Tassili-n-Ajjer ranges of south-eastern Algeria and the Aïr massif in northern Niger harbor a number of endemic and rare Mediterranean species such as the tarout, the wild olive, and the Saharan myrtle. To the east, the Tibesti mountains in southern Lybia hold some Mediterranean as well as some tropical relicts. These Saharan mountains also provide prime habitat for migratory birds and a key refuge for threatened wildlife.

In the Somali Peninsula, and across the Gulf of Aden in the southern part of the Arabian Peninsula, high mountain ranges shelter similar temperate relicts: along the northern Somalian coast, in the tip of the Horn of Africa, the Somali Montane Woodlands thrive along the coastal ranges fed by moisture brought in from the sea. A biogeographic refuge and center of endemism, these mountain habitats harbor many endemic species of both plants and animals. The higher part of the ranges contains an evergreen scrub quite similar to the Mediterranean maquis, with species such as kadi, mountain ox, Ethiopian pistachio, and remnants of ancient juniper forests.

Across the Aden Straits, and overlooking the coast of the Red Sea, the high Asir mountains of Yemen and Saudi Arabia give shelter to an even richer relict montane ecosystem with juniper, bramble, and zaytoon olive. Fed by the coastal fogs of the Red Sea, the region shelters more than 170 endemic plants, and is a critical refuge for a large number of endangered animals such as the Arabian leopard and the Hamadryas baboon. Finally, in the eastern part of the Arabian peninsula, near the coast of the Gulf of Oman, the Al Hajar Mountain range forms a spectacular wall of mountains that rise almost 3000 m from the surrounding deserts, providing an important refuge for endemic and relict species of plants of Mediterranean and Indo-Iranian origin, often growing in the vicinity of tiyu (a carob relative) and other trees of African ancestry. The endemic Arabian tahr, a type of wild goat, is still frequent in these precipitous slopes.

These mountains, and many others, play an immensely important role in the maintenance of desert diversity and in the conservation of desert biota. Sharp ecotones exist between arid plains and island-like, desert-surrounded mountains. In these slopes, altitudinal change reflects evolutionary time and in their intense transitions it is possible to explore the evolutionary history of deserts during the Pleistocene (**Figure 11**).



Figure 11 The Maderas del Carmen range in the Chihuahuan Desert is a prime example of a montane Pleistocene relict, or desert "sky island". The lowlands are covered by creosote bush (*Larrea tridentata*) and prickly pears (*Opuntia* sp.), two hardy desert perennials, but high-up in the mountains the ancient pine forests that once covered a large part of the region still survive (Jaime Rojo).

Distribution and Characteristics of World Deserts

True deserts are unique, highly adapted natural ecosystems. The desert biome includes arid and hyperarid drylands, both providing ecosystem services and supporting human populations that may degrade the land resources in much the same ways as in other ecosystems. The deserts of the world occur in six of the world's large biogeographic domains, or eco-realms. A detailed description of each desert ecoregion within these realms can be found at National Geographic's Wild World website (www.nationalgeographic.com/wildworld/terrestrial.html) and at the World Wildlife Fund's terrestrial ecoregions website (http://www.worldwildlife.org/science/ecoregions/biomes.cfm). A short narrative description of the world's main deserts is given in the sections that follow.

Afrotropic Realm: Deserts of East and Southern Africa

The Afrotropic deserts comprise four large lowland desert ecoregions (Ethiopian xeric grassland and shrublands, Nama Karoo, Somali *Acacia-Commiphora* bushlands and thickets, and Southwestern Arabian foothills savanna); eight ecoregions occupying coastal desert fringes (Arabian Peninsula coastal fog desert, Eritrean coastal desert, Gulf of Oman desert and semidesert, Madagascar spiny thickets, Namib desert, Northern Namib's Skeleton Coast, Red Sea coastal desert, and Succulent Karoo), and four montane Pleistocene relict ecoregions, or desert "sky-islands" (Al Hajar montane woodlands, Ethiopian montane forests, Somali montane xeric woodlands, and Southwestern Arabian montane woodlands). They cover in total 2.7 million km^2, of which some 10% is under

environmental protection. Their mean population density is 21 persons km^{-2}, and their mean human footprint index (20; see Sanderson *et al.*, 2002) is moderate on average, but relatively high in the Horn of Africa and Madagascar.

Among the plants that are unique to these deserts are *Welwitschia mirabilis* of the Namib; a great variety of woody legumes and succulent-stemmed species such as baobabs (*Adansonia*), commiphoras or myrrh-trees (*Commiphora*), bottle-trees (*Pachypodium*), and kokerbooms (*Aloe dichotoma*). The Succulent Karoo, which shelters many of these plants, is the world's only plant hotspot (Myers *et al.*, 2000) that is entirely found within the desert biome and is entirely arid. The Madagascar thorny thickets represent a unique, very diverse assemblage of plants and animals, most of them not found anywhere else – such as the local baobabs and the octopus tree (*Didierea madagascariensis*). Much of the Namib is protected but some significant areas are at risk because of prospecting and mining of diamonds and copper. In contrast, very little of the Madagascar thorny thickets are protected, and more than 90% of its original habitat has disappeared through extraction of wood for firewood and charcoal, grazing and clearance for farming. Browsing by livestock and firewood collection by humans also threaten the deserts around the Horn of African and in the southern Arabian Peninsula, from the Red Sea to the Gulf of Oman.

Australasian Realm: Deserts of Australia

The deserts in the Australasian realm comprise 10 flat, lowland ecoregions in WWF's classification (Carnarvon xeric shrublands, Central Ranges xeric scrub, Gibson desert, Great Sandy-Tanami desert, Great Victoria desert, Nullarbor Plains xeric shrublands, Pilbara shrublands, Simpson desert, Tirari-Stuart stony desert, Western Australian Mulga shrublands), covering some 3.6 million km^2, of which around 9% is under some sort of environmental protection scheme. Hardly inhabited at all, their mean population density is less than 1 person km^{-2}. These deserts show, by far, the lowest human footprint of the global biome.

All the arid ecoregions identified in Australia receive only sporadic rain and carry sparse vegetation but each one presents its own peculiarities. The Gibson Desert, the "great undulating desert of gravel," is an area of red gravel soils with mulga (*Acacia aneura*) accompanied by other drought-resistant shrubs and hardy spinifex (*Triodia*) grass. The Central Ranges xeric scrub has low myall and desert oak woodlands growing over spinifex on the sand plains, whereas, throughout the ranges, a profusion of endemic ferns, reeds, and rushes grows in gorges where there is abundant water. The Great Sandy-Tanami Desert is an extensive area of reddish sand plains out of which rises the great Uluru (Ayers' Rock); the plains are covered with clumps of spinifex and scattered saltbush (*Atriplex* spp.) with sparse, spiny acacias and tall desert oaks (*Allocasuarina decaisneana*). The vegetation in the Great Victoria Desert is restricted to the hardiest of species, such as spinifex. The dunes and sand plains of the Simpson Desert carry sparse shrubs and spinifex with cane grass (*Eragrostis australasica)* on the deep sands along the crests of dunes, and coolibah (*Eucalyptus* sp.) snaking across the northeastern corner. The Tirari Desert is characterized by vast sand dunes that harbor an abundant and diverse fauna, an array of unique creatures that have adapted to the high temperatures and soil aridity (marsupials, lizards, frogs, and even small parrots); important fossil deposits have been discovered in the area. Over the Tirari-Sturt Stony Desert sparse saltbush sustains sheep grazing, while its southern portion is woodland with native cypress, black oak, and eucalyptus. Both the Gibson and the Great Sandy-Tanami Deserts are extensive rangelands, mostly uninhabited but supporting important mining activities at the Tanami. Threats include overgrazing, feral, and exotic animals, in addition to the localized and sometimes severe disturbances caused by tourism and mining.

Indo-Malay Realm: Deserts of India and Pakistan

The Indo-Malay region harbors only two hot lowland deserts – the Indus Valley and the Thar – covering 0.26 million square kilometers, of which some 20% receives some level of legal environmental protection. In terms of population, these are the densest and most heavily used deserts in the world, with a mean of 151 persons km^{-2}; while their mean human footprint (33) is the highest in the desert biome. These deserts receive 10 and 80 mm of annual rainfall respectively, enough to support trees, shrubs and grasses, in particular *Halostachys caspica* on salty soils and *Aristida pennata* on sandy ones. The rural population depends mostly on raising sheep and goats. About 63% of the Indus Valley and 17% of the Thar Desert are protected (the latter in 11 different spots).

Nearctic Realm: Deserts of North America

The Nearctic deserts are formed by five lowland deserts (Chihuahuan desert, Sonoran desert, Mojave desert, Great Basin shrub steppe, and Meseta Central matorral); two coastal deserts (Baja California desert and Gulf of California xeric scrub), and four montane relict sky-islands (Western Madrean Archipelago, Eastern Madrean Archipelago, Great Basin montane forests, and Sierra de Juárez and San Pedro Mártir pine-oak forests). They cover 1.7 million km^2, of which 19% is under some level of legal protection. Because of the growth of large urban conglomerates such as Phoenix in the US, their mean population density is high (44 persons km^{-2}) and their mean human footprint (21) is the second highest of the world's deserts.

These deserts present a combination of sagebrush (*Artemisia tridentata*) and creosotebush (*Larrea tridentata*) shrublands, halophytes (such as *Atriplex confertifolia)*, and open grasslands in the moister parts. There are numerous cactus species except in the Great Basin shrub steppe, where freezing temperatures limit cactus growth. The giant saguaro cactus is emblematic of the Sonoran Desert; it grows very slowly but may reach a height of 15 m over its long lifespan of more than 200 years. The Baja California desert is home to over 3500 species of endemic plants, almost a quarter of which are endemic, ranging from a diversity of shapes and sizes of cacti, to thick-stemmed trees and shrubs in the rocky mountain soils. Likewise, around 1000 of the 3200 plant species in the Chihuahuan Desert are endemic, while creosote bush, tarbush,

viscid acacia, yucca, and cactus are characteristic. The Joshua tree (*Yucca brevifolia*) is probably the most recognizable plant of the Mojave; growing with a wide variety of cacti, creosote bush, white bur-sage (*Ambrosia dumosa*), jojoba (*Simmondsia chinensis*), and small trees such as paloverde (*Parkinsonia microphylla*) and ironwood (*Olneya tesota*). Grazing, hunting and salt extraction (in the Baja California desert) are significant anthropic activities in these ecoregions, which also contain several large and fast-growing cities: Las Vegas, Reno, and Salt Lake City in the Great Basin, Los Angeles near the Mojave, Phoenix and Tucson in the Sonoran Desert.

Neotropic Realm: Deserts of South America

The Neotropic deserts comprise two lowland deserts (Argentine Monte and High Monte), a high-altitude inland desert (Central Andean Dry Puna), and two coastal deserts (Atacama and Sechura deserts). They cover 1.1 million km^2, of which only 6% receives legal protection. Their mean population density is 18 persons km^{-2}, and their mean human footprint (16) is lower than in their North American counterparts, with most pressure concentrating in the Sechura Desert along the Peruvian coast. These deserts form a long arid continuum that cover South America's "arid diagonal", starting from the Pacific Ocean just south of the equator in Peru, and running in a southeast direction to the Atlantic coast of northern Patagonia, at latitude 43° S.

The barren landscape of the Atacama Desert is represent one of the driest deserts on earth. The almost complete absence of vegetation in its interior is due to the lack of precipitation and the high mineral content of the soils. Rare rainfall events cause ephemeral plants to germinate and burst with flowers for a short period of time. The Monte, east of the Andes, is a fold desert with sandy plains, plateaus, and rocky foothills with an open, low thorn-scrub harboring a characteristic endemic flora of zygophylls (family Zygophyllaceae). There are also edaphic communities of many species such as *Prosopis* thickets in ravines, shrublands of broom rape (*Baccharis*) and saltbush (*Atriplex*) on clay soils, and *Allenrolfea vaginata* and *Suaeda divaricata* in salty soils. The dry Central Andean Puna carries tall tussocks of bunchgrass and other high-altitude grasses, shrubs like *Parasthrephia lepydophilla* and *Baccharis*, and a unique flora of high-altitude cushion plants. The main land use in the Puna is grazing with llamas, alpacas, goats, and sheep. In ancient, pre-Hispanic times, these deserts were an important part of the Inca Empire; a network of roads stretched through them forming the *Qhapaq ñan*, or *Camino del Inca* (the Inca Road).

Palearctic Realm: Deserts of Central Asia, the Middle East, and Northern Africa

By far the largest set of deserts in the world is found in the Palearctic realm. The region includes 21 lowland desert ecoregions that cover an immense corridor of deserts stretching from the Atlantic coasts of Morocco, to the Mediterranean coasts of the Sahara, to the Gobi in Central Asia, including the deserts of Northern Africa, Arabia, Azerbaijan, Taklimakan, Central Persia, and the Caspian lowlands (the full list of ecoregions includes Alashan Plateau semidesert, Arabian Desert and East Sahero-Arabian xeric shrublands, Azerbaijan shrub desert and steppe, Badghyz and Karabil semidesert, Caspian lowland desert, Central Asian northern desert, Central Asian southern desert, Central Persian desert basins, Eastern Gobi desert steppe, Gobi Lakes Valley desert steppe, Great Lakes Basin desert steppe, Junggar Basin semidesert, Mesopotamian shrub desert, North Saharan steppe and woodlands, Qaidam Basin semidesert, Red Sea Nubo-Sindian tropical desert and semidesert, Registan-North Pakistan sandy desert, Sahara desert, South Iran Nubo-Sindian desert and semidesert, South Saharan steppe and woodlands, and Taklimakan desert). The Palaearctic realm also harbors three coastal deserts (Atlantic coastal desert, Persian Gulf desert, and Red Sea coastal desert), as well as five montane sky-island relict ecosystems (Tarim Basin deciduous forests and steppe, Kuh Rud and Eastern Iran montane woodlands, Afghan Mountains semidesert, Tibesti-Jebel Uweinat montane xeric woodlands, and West Saharan montane xeric woodlands). The Palaearctic deserts cover a remarkable 16 million km^2, totaling 63% of all the deserts on the planet. Of this area, 9% is legally protected. Their population density is 16 persons km^{-2}, and their mean human footprint (15) is the second lowest in the planet, possibly because of the sheer inaccessibility and the extreme aridity of many of its large ecoregions.

The great shield desert of the Sahara has immense dune fields covering 33% of the total area, plains of metasediments (sediments or sedimentary rocks that have been subject to metamorphism), and sky-islands, mostly on volcanic rocks in the Tibesti, Ahaggar, and Aïr ranges. The Red Sea rift separates the Sahara from the Arabian Peninsula, which dips toward the Persian Gulf. The deserts of the Middle East and Central Asia are all rain-shadow deserts, mostly encircled by young fold mountains. The Sahara alone occupies 4.6 million km^2, or 10% of the African continent. It includes an undisturbed hyperarid central area of sand and rock, but with small areas of permanent vegetation. Vegetation tends to be much more diversified in the Western Sahara, with xerophytes and ephemeral plants in the open desert plains, and halophytes in the moister areas. Currently the Sahara Desert is not well protected legally, but its sheer inaccessibility plays a major role in conserving its ecosystems. The nomadic population depends on pastoral activities, hunting, and trade. Most settled people along the desert fringes do not venture into its interior. The high mountains shelter wild ancestors of many trees that have been domesticated for their fruits and nuts, such as pistachio and wild olive. Considering the hyperarid conditions, the fauna of the Sahara is relatively rich, with 70 mammal species, 20 of them large; 90 species of resident birds, and around 100 species of reptiles.

In contrast with the Sahara and Arabian deserts, the deserts of Central Asia exhibit fold mountain slopes cut by steep valleys with mostly ephemeral streams, fringing alluvial fans, and enclosed basins, some of which contain lakes (Caspian and Aral Seas, Lop Nor) or playas. The Turpan Depression in Western China reaches 154 m below sea level. The predominant vegetation types are all formed by typical desert species of the Old World: grasses such as *Panicum*, *Poa*, and *Stipa*; sedges (*Carex*); desert bulbs such as wild tulips (*Tulipa* spp.); halophytes such as *Salicornia* and *Atriplex*; and shrubs

such as *Artemisia*, *Euphorbia*, and *Caragana*. The most common vegetation found in the ecoregion is the desert sagebrush and other *Artemisia* species. The range of the flora goes from sagebrush to psammophilous (dune-adapted) plants, and includes salt-tolerant chenopod communities (Chenopodiaceae) in Afghanistan. The salt pans have almost no vegetation. In some parts, the deserts of Central Asia still support small populations of valuable animals like wild Bactrian camels (*Camelus ferus*) and Asian wild asses (*Equus hemionus*) that have been largely extirpated in the wild.

The Mesopotamian shrub desert is transitional between the deserts to the south and the steppes to the north. The flora includes umbrella-thorn acacia (*Acacia tortillis*), shrubby rock-rose species (*Cistus* spp.), and many dwarf shrubs. Reeds and rushes grow in the wetland areas, while poplar (*Populus euphratica*) and tamarisk (*Tamarix*) grow along river channels.

The whole ecorealm has evolved under continual grazing pressure and most plants are adapted to grazing pressure. However, overgrazing is a constant threat. Illegal hunting is a serious threat in the South Iran Nubo-Sindian desert; while egg collection and nest disturbance affect nesting migratory waterfowl. Significant portions of the Azerbaijan shrub desert and the Central Asian Northern desert are farmed under irrigation, bringing on water and soil pollution by the use of fertilizers and pesticides. The region as a whole contains most of the world's oil reserves and large gas reserves, increasingly exploited and with associated urban areas and infrastructure development. The vital ground waters are also under heavy development pressure.

Biological Adaptations to Aridity

At its origins, life evolved in water, and water is the most crucial element for the survival of all living organisms. Thus, it is no surprise that some of the most remarkable adaptations for survival are found in desert species, that is, in the environments where water is most scarce (Louw and Seely, 1982). The short pulses of abundance that contrast sharply with the background condition of aridity and scarcity are the major force that has driven evolution, natural selection, and adaptation in desert biota. Plants and animals are adapted to these seasonal strokes. Natural selection and evolution have molded in very precise ways the life-forms of desert organisms to their harsh and unpredictable environment. Furthermore, because most deserts of the world have developed in relative isolation from each other, many of their constituent species have evolved from different ancestors (Morton, 1979). Thus, deserts are prime ecosystems to study and understand the phenomenon of convergent evolution – the development of similar growth forms and adaptations derived from different ancestors (**Table 2**, **Figure 12**).

Adaptations of Plants to Aridity

Most desert species have found remarkable ways to survive drought (Robichaux, 1999). Desert succulents, such as cacti or rock plants (*Lithops*) for example, survive dry spells by accumulating moisture in their fleshy tissues. They have an extensive system of shallow roots that allows them to capture soil water only a few hours after it has rained. Their photosynthesis is modified to exchange gases and fix CO_2 during the night, when evaporative demand is low, and to accumulate the fixed carbon in the form of malic acid, which is later used by the plant as a building block of more complex organic molecules (this photosynthetic pathway is called "Crassulacean Acid Metabolism" or CAM). Additionally, many cacti and other stem-succulent plants of hot deserts present columnar growth, with leafless, vertically erect, green trunks that maximize light interception during the early and late hours of the

Table 2 Convergent evolution of animal guilds in four hot deserts of the world. In spite of their different phylogenetic origins and their different biogeographic histories, many of these animals look strikingly similar in appearance

Guild	*Deserts*			
	North American	*Australian*	*North African*	*South African*
Bipedal rodent granivores[a]	*Dipodomys*	*Notomys*	*Jaculus*	
Quadrupedal rodent granivores[a]	*Chaetodipus*		*Taterillus*	*Gerbillus*
	Perognathus		*Pachyuromys*	*Tatera*
			Gerbillus	*Gerbillurus*
			Sekeetamys	
Small mammalian insectivores[a]	*Onychomys*	*Antechinus*	*Crocidura*	*Crocidura*
	Notiosorex	*Sminthopsis*		*Elephantulus*
		Antechinomys		*Macroscelides*
Mid-sized herbivores[a]	*Lepus*	*Onychogalea*	*Lepus*	*Pedetes*
	Sylvilagus			
Mammalian carnivore[b]	*Vulpes*		*Fennecus*	
Ant eating reptile[b]	*Phrynosoma*	*Moloch*		
Horned snakes[b]	*Crotalus*		*Cerastes*	*Bitis*
Open space, long-legged lizard[c]	*Callisaurus*	*Amphibolurus*		
Medium-sized, lizard-eating lizard[c]	*Crotaphytus*	*Varanus*		

[a]Mares MA (1980) Convergent evolution among desert rodents: A global perspective. *Bulletin of the Carnegie Museum of Natural History* 16: 1–51.
[b]Schmidt-Nielsen K (1964) *Desert Animals: Physiological Problems of Heat and Water.* Oxford: Clarendon Press.
[c]Pianka ER (1986) *Ecology and Natural History of Desert Lizards: Analysis of the Ecological Niche and Community Structure.* Princeton, NJ: Princeton University Press.



Figure 12 The Australian horny devils (*Moloch horridus*, shown in this picture taken in Alice Springs, Australia) are surprisingly similar to the North American horned lizards (*Phrynosoma coronatum*, in Baja California) in both their feeding habits and their ecology. Their similarity, however, is due to convergent evolution in distant deserts, as they belong to different families and are taxonomically unrelated (Patricio Robles-Gil).

day, but avoid the midday sun, when excessive heat may damage, or even kill, the plant tissues. Thus, erect columnar plants may avoid drought by (1) accumulating water, (2) exchanging gases at night, and (3) morphologically avoiding midday exposure to solar radiation (Zavala *et al.*, 1998). Instead of having developed succulent stems adapted for water storage, other CAM plants such as aloes and agaves have developed ground-level rosettes of succulent leaves. Because of the funnel-shape of their leaf-whorls, these plants seem to be adapted to collect dew from morning fogs in desert mountains, hill slopes, and ocean coasts, accumulating the water thus harvested in their fleshy leaves and central bud (Martorell and Ezcurra, 2002).

Woody desert trees, such as acacias, cannot store much water in their trunks but many of them evade drought by shedding their leaves as the dry season sets-in, entering into a sort of drought-induced latency. Many of these desert species also have long taproots that explore deep underground water layers. Other trees have evolved convergently a mixture of these strategies: they can store water in gigantic trunks and have a smooth bark that can do some cactus-like photosynthesis during dry periods; but, when it rains they produce abundant green leaves and shift their metabolism toward that



Figure 13 The fleshy stems of the kokerbooms (*Aloe dichotoma*) – one of Africa's many giant, fleshy-stemmed trees – are true landmarks in the otherwise barren landscape of the Namib Desert (Patricia Rojo).

of normal-leaved plants. This group is formed by trees with famously "bizarre" trunks, such as the African baobab (*Adansonia*), the Baja-Californian Boojum-tree (*Fouquieria columnaris*) and elephant-trees (*Bursera* and *Pachycormus*), and the South African commiphoras (*Commiphora*), bottle-trees (*Pachypodium*), kokerbooms (*A. dichotoma*, **Figure 13**) and botterbooms (*Tylecodon*).

A third group of plants, the "true xerophytes" or true desert plants, have simply adapted their morphology and their metabolism to survive extremely long droughts. These species have remarkably low osmotic potentials in their tissue, which means that they can still extract moisture from the soil while most other plants cannot do so. True xerophytes, such as the creosote bush (*Larrea*), are mostly shrubs with small, leathery leaves that are protected from excessive evaporation by a dense cover of hairs or a thick varnish of epidermal resin. Their adaptive advantage lies in their capacity to extract a fraction of soil water that is not available to other life-forms. However, because their leaves are so small and protected from transpiration, their gas-exchange metabolism is very inefficient during rain pulses when moisture is abundant. In consequence, these species are extremely slow growers, but extremely efficient water users and very hardy.

Finally, one of the most effective drought-survival adaptations for many species is the evolution of an ephemeral life cycle. A short life and the capacity to leave behind resistant forms of propagation is perhaps one of the most important evolutionary responses in most deserts, found not only in plants but also in many invertebrates. Desert ephemerals are amazingly rapid growers capable of reproducing at a remarkably high rate during good seasons, leaving behind myriad resistance forms that persist during adverse periods. Their population numbers simply track environmental bonanzas; their way to evade critical periods is to die-off, leaving behind immense numbers of propagules (seeds or bulbs in the case of plants, eggs in the case of insects) that will restart the life cycle when conditions ameliorate. These opportunist species play an immensely important role in the ecological web of deserts: myriad organisms, like ants, rodents, and birds, survive the dry spells by harvesting and consuming the seeds left behind by

the short-lived ephemeral plants. Granivory (the consumption of seeds) and not herbivory (the consumption of leaves) is at the base of the food chain in most deserts, as those few plants that maintain leaves during dry spells usually endow them with toxic compounds or protect them with spines. The onset of rainy periods brings to the desert a reproduction frenzy of desert ephemerals, and a subsequent seed-pulse that drives the entire food web for years.

Thus, the survival strategies of desert plants are classifiable along a gradient ranging between two extreme categories: (1) adaptation for quick use of ephemerally abundant resources, or (2) adaptation for the efficient use of poor but more permanent resources (Shmida, 1985). The first category, typically exhibited by desert ephemerals, represents a "maximum variance" behavior that consists essentially in tracking environmental variation, while the second category, exhibited by true xerophytes and cacti, is a "minimum variance" behavior that consists in adapting to the worst possible conditions. Drought-deciduous perennials and grasses represent a compromise between the two extremes. Attributes necessary for the quick use of water include rapid growth (often at the cost of low water-use efficiency) and abundant seed production. Attributes for survival with little water include high water-use efficiency, slow growth, and passive cooling. Drought deciduousness, as an intermediate strategy, requires the capacity to shed leaves and to quickly recover them when moisture conditions improve.

The survival strategies of desert plants present some of the most striking cases in nature of evolutionary convergence: Plants from widely different families and from divergent evolutionary origins have developed in the different deserts life-forms so similar that it is sometimes difficult to tell them apart. Such is the case of the succulent cactoid growth form, evolved in Africa in the families Euphorbiaceae, Asclepediaceae, and Aizoaceae, and in the Americas in the family Cactaceae. Similarly, bottle-trees in Africa evolved in the families Apocynaceae, Aloeaceae, and Crassulaceae, while in the New World they belong to the Fouquieriaceae, Anacardiaceae, and Burseraceae. Superficially, however, they are almost undistinguishable.

Adaptations of Animals to Aridity

Adaptation of animals to desert conditions can be divided into morphological and anatomical, physiological or behavioral. Among the first, in mammals, desert fur coats are short, hard and compact, but at the same time well ventilated, to allow sweat to evaporate directly from the skin. Birds can fluff or compact their feathers to regulate heat exchange. In the ostrich, a desert dweller, the uncovered head, throat, legs and abdomen allow for radiation and convection cooling, while the feathers on the back protect the larger part of the body from direct solar radiation (**Figure 14**). Bipedalism, a common trait in small desert mammals such as kangaroo rats, allows for fast travel in open spaces and also keeps the body separated from the extreme temperatures of the ground surface. Indeed, bipedal desert rodents use open microhabitats much more frequently than their quadrupedal relatives, who restrict their activities to sheltered habitats.



Figure 14 Despite the grueling heat and extreme drought, a family of ostriches (*Struthio camelus*) survives and thrives in the Namib Desert (Patricio Robles-Gil).

Sand-dwellers have evolved several traits that allow them to survive in dunes, including fleshy foot-pads in camels, scaly fingers in certain lizards and digital membranes in some geckoes. Additionally, camels have long dense lashes that protect their eyes and they can close their nostrils to protect them from wind-blown dust. Many snakes have upwardly turned nostrils that allow them to burrow rapidly in loose sand; others are flat and can bury laterally. Many reptiles show also adaptations that protect their eyes, nose, and ears from sand and dust, and many insects have especially adapted legs that allow them to bury themselves rapidly and to walk efficiently on hot sand.

The most basic physiological problem of desert animals is to maintain their water balance by maximizing water intake and/or minimizing water loss. In deserts, free-standing water is scarce, found only in isolated oases and reservoirs. Camels and wild asses, for example, are able to drink large quantities of water in a very short time, which causes a dramatic dilution of the bloodstream sufficient to cause death in other animals, with no ill effects. In coastal deserts, animals obtain water by licking fog-drenched rocks. Desert amphibians can absorb water through the skin from humid underground dens by accumulating urea in their blood and raising its osmotic pressure.

Most herbivores, like eland and oryx, obtain water from the foliage of the shrubs that compose their diet, often feeding at night when the plants are turgent. Some succulent plants have high salt contents, toxic compounds, or spines to deter their consumption. Herbivores, however, have found their way around these obstacles: some reptiles and birds have developed efficient salt-excreting glands, and many mammals have kidneys that can cope with salty water. White-throated packrats (*Neotoma albigula*), which feed almost exclusively on juicy cacti, have metabolic adaptations to prevent poisoning from the oxalates contained in these plants.

Animals lose water through urine, feces, respiration, and transpiration. Desert rodents have kidneys that are capable of producing highly concentrated urine, with an electrolyte concentration many times higher than that of blood plasma. Reptiles, birds and insects excrete uric acid, which requires less

water, and sometimes complement the excretion process with specialized excretion from salt glands. Amphibians produce little urine, and can store large amounts of urea within their bodies, drastically reducing water loss. In droughts, some rodents can produce dry feces, efficiently reabsorbing liquids in the rectum.

Metabolism produces CO_2 and water as byproducts of respiration. In most animals, this metabolic water is exhaled through the lungs, but many desert animals, including invertebrates, reptiles, and mammals, possess physiological and anatomical adaptations to reduce respiratory water loss, including modifications in the morphology of the nasal passages and the capacity to reabsorb water along the respiratory tract. One of the most extreme examples of this is given by the kangaroo rats (*Dipodomys*), which can survive on a diet of perfectly dry seeds.

In addition to the mechanisms that reduce water loss, many desert animals are extremely tolerant to dehydration, a condition that causes a fatal increase in blood viscosity in nondesert dwellers. To achieve this, camels, for example, are able to lose water selectively from tissues other than blood. In contrast, desert amphibians are tolerant to increased fluid viscosity, and some reptiles can excrete excess electrolytes through urine and salt glands, avoiding the thickening of the blood as they dehydrate. A problem related to dehydration is that of temperature regulation. In smaller animals the high surface-to-body ratio makes sweating a dangerous enterprise and panting is the most common method of cooling. Even larger animals that usually sweat, like the oryx, begin to pant when their body temperature exceeds 41 °C.

Nocturnal hypothermia, exhibited by some large mammals like the eland, allows them to reduce their metabolic rate and to exhale air with less humidity during the night. Diurnal hyperthermia allows animals to reach body temperatures that would be normally lethal for nondesert vertebrates and to save on water needed to prevent overheating. Camels and elands, for example, can reach body temperatures of 44 °C with no harmful consequences, and saving as much as 5–10 l of water during extremely hot days. Hyperthermal species have a special disposition of veins and arteries that allows their brains to remain at a temperature lower than that of the overheated body.

Like ephemeral plants, many smaller desert animals can also evade drought by entering into a dormant phase: Desert butterflies and grasshoppers thrive in huge numbers when conditions are good and survive dry spells in the form of eggs or pupae. Spade-foot toads (*Scaphiopus*) spend most of their lifetime buried in dry mud and become active only after rains refill their ephemeral pools. Many other organisms go into some form of torpor during dry periods.

Behavioral adaptations are many. Numerous birds and most large mammals, like pronghorn antelopes or wild sheep, can evade critical spells by migrating along the desert plains or up into the mountains. Smaller animals cannot migrate such long distances, but regulate their environment by seeking out cool or shady places. In addition to flying to other habitats during the dry season, birds can reduce heat loads by soaring. Many rodents, invertebrates, and snakes avoid heat by spending the day in caves and burrows, and procuring food during the night. Even diurnal animals may reduce their activities by resting in the shade during the hotter hours of the day. Fossoriality, a lifestyle based in burrows, is the norm for small animals in all deserts, as it allows them to stay away from the grueling heat during the hotter part of the day and it also provides them with a warm refuge during the cold desert nights. Additionally, humidity inside burrows (approximately 30–50%) allows desert animals to preserve water. When the normal mechanisms to prevent increases of body temperature above acceptable limits fail, many small rodents and some desert tortoises (*Testudo*) salivate to wet the chin and throat and allow evaporative cooling. Such mechanism has a high cost in water and is used only as an emergency measure to prevent death.

At dawn, the dry desert ground may approach freezing temperatures and at midday it may heat up into an 80 °C inferno. A few inches above the ground, variations in air temperature are much less pronounced, and, just a few inches below the surface, temperatures are almost constant between day and night. For this reason, thermoregulating is a particularly challenging problem for small surface-dwellers and especially for reptiles, which cannot regulate metabolically their body temperature. Most desert reptiles have developed peculiar ways of travelling over hot sandy surfaces. Side-winding, a form of lateral movement in which only a small part of the body is in contact with the surface, is employed by many sand snakes (**Figure 15**). Many lizards and some ground birds avoid overheating by running rapidly over the hot desert surface while maintaining their bodies well separated from the ground (Safriel, 1990). Some lizards assume an erect, bipedal position when running, while others regulate their contact with the hot desert pavement by doing "push-ups" with their forelegs. By changing from quadrupedal to bipedal travel, several species move more efficiently as they direct their running force downward rather than laterally (MacMahon, 2001).

Many large mammals that cannot avoid being in the sun during a large part of the day orient their bodies so as to reduce the incidence of the sun's rays. By standing upright, guarding ground squirrels reduce solar incidence upon their bodies. The African ground squirrel *Xerus inauris* even orients



Figure 15 The characteristic crawling patter of the sidewinder (*Crotalus cerastes*) leaves a tell-tale trail in the dunes of the Gran Desierto. Its perfect matching to the color of the sand protects it from predators such as kestrels and falcons (Patricio Robles-Gil).



Figure 16 Using its tail like a parasol, the African ground squirrel (*Xerus inauris*) protects itself from the sun in the Namib Desert (Patricio Robles-Gil).

toward the sun and shades itself with its tail when foraging (**Figure 16**). The jackrabbit *Lepus californicus* warms its body in the early morning by exposing its large, highly vascularized ears perpendicular to the sun's rays, using them as a form of solar collector. Similarly, it cools at midday by keeping in the shade and putting the ears parallel to the incoming solar radiation, thus minimizing exposure while keeping the same radiative surface.

Desert Ecosystem Processes

Primary Productivity and Soil Fertility

Although deserts cover around 22% of the world's terrestrial area, only about 1% of global primary productivity occurs in them. There is a strong correlation between annual net primary productivity (NPP) and precipitation; NPP in extreme deserts ranges from 0 to 10 g m^{-2} $year^{-1}$, and, in shrubby deserts, from 10 to 250 g m^{-2} $year^{-1}$. These low values contrast with those of temperate grasslands (200–1500 g m^{-2} $year^{-1}$) and forests (600–1500 g m^{-2} $year^{-1}$). In North America, the Mojave Desert is, on average, the lowest producing area with 25 g m^{-2} $year^{-1}$, while average values in the Chihuahuan Desert are close to 120 g^{-2} $year^{-1}$.

Nitrogen (N), phosphorus (P), and potassium (K) are important nutrients for plant growth and, in arid areas, these nutrients are usually associated positively with shrubs. For example, in a creosotebush community there were 0.253 mol of N per kg of soil beneath the canopy and 0.148 and 0.10 at the canopy edge and interspace, respectively. There may be several reasons for this association, the most general one relating to the concept of "islands of fertility." Once established, a shrub sheds leaves (at a rate of 19–53 g m^{-2} $year^{-1}$) and portions of its belowground roots die adding organic matter to the soil. Additionally, the roots may extend into the barren area between shrubs and extract nutrients from those areas, concentrating them under shrubs. Rodents often make burrows at the bases of shrubs and the addition of their nesting material and digging activities can increase soil nutrients.

Nitrogen is generally scarce in deserts. It enters the system through fixation by organisms in soil crusts, through fixation by the nodules on the roots of woody legumes such as mesquite and acacias, and through airfall deposition. In the past it was thought that low quantities of nitrogen limited primary productivity. More recently, it has been suggested that phosphorus may be more limiting because it is relatively unavailable in the alkaline soils of deserts. When nutrients exist in desert soils, they may not be available because decomposition occurs more slowly in the absence of water, and because the microfauna and microflora that usually affect decomposition are at much lower densities than in forest or grassland systems. None of these decomposers can continue activity when soils are very dry. In the Chihuahuan Desert, termites are extremely important in decomposition, but this seems less so in drier areas, where biological decomposition of organic matter occurs mostly under the protective shade of shrub canopies and grass clumps. In the open bare space, ultraviolet radiation, the same factor that hinders biological decomposition, is the main cause of organic matter degradation. It is solar UV radiation, and not humus-forming microorganisms, that finally weathers and decomposes organic matter in the bare soil of extreme deserts (Montaña *et al.*, 1988).

Interactions Between Species

The harsh conditions of desert ecosystems has promoted the evolution of a complex set of relations among desert organisms, a surprising number of which are positive interactions (Cloudsley-Thompson, 1996). Desert shrubs in general and woody legumes in particular, create microhabitats that are critical for the survival of other species. Small animals seek the shade of desert trees and shrubs, birds find refuge and nesting sites in their canopies, and many small plants recruit their juveniles under the nitrogen-rich canopy of desert woody legumes such as acacias, carobs, and mesquites. Because of their CAM metabolism, desert succulents such as agaves, aloes, and cacti are poor thermoregulators as young seedlings, and cannot survive the harsh ground-level midday temperatures. For this reason, they can successfully germinate and establish only under the protective shade of shrubby "nurse plants" that act as true cornerstone species in desert conservation (**Figure 17**). If the desert trees and shrubs are cut, all the accompanying biota soon disappears.

Additionally, many desert plants have very specific requirements in terms of their pollinators and seed dispersers



Figure 17 Giant saguaro cacti (*Carnegiea gigantea*) in the Pinacate mountains of the Sonoran Desert establish and grow under the protective shade of a paloverde (*Parkinsonia microphylla*), a legume species that generates true islands of fertility in the harsh desert environment (Patricio Robles-Gil).



Figure 18 A Gambel's quail (*Callipepla gambelii*) forages on the ripe, red, juicy fruits of the saguaro. After digesting the sweet pulp it drops the seeds, often under the protective shade of nurse plants, where the cactus will be able to germinate and establish (Patricio Robles-Gil).

(**Figure 18**). Although some desert ephemerals are truly unspecific in their requirements and produce thousands of seeds with only wind-pollination, the slow-growing desert perennials are frequently highly specialized in their reproductive habits, and depend strictly on coevolved animals to help them out in their sexual and reproductive processes. Many African cactoid plants (euphorbs and asclepias) produce foul-smelling flowers that attract carrion insects as pollinators. New World giant cacti and agaves produce sugar-rich nocturnal flowers that engage the pollinating services of nectar-feeding bats, while the sweet pulp of their fruits lures birds to disperse the seeds miles away. The red tubular flowers of many desert shrubs attract hummingbirds and giant sphinx-moths.

In almost all deserts of the world, three types of granivorous organisms are influential: harvester ants, small mammals, and birds. These organisms disperse or consume seeds, and, in the case of ants and burrowing small mammals, they turn over soil at very high rates, change the physical conditions of the soil through digging activities, incorporate organic matter, and remove vegetation from the sites of their nests. Some bipedal rodents, such as kangaroo rats, dig small surface caches to store seeds, from where new plants may germinate after a pulse of rain. Interestingly, the relative importance of these three groups differs around the world: In North America and Israel, small mammals are most important and remove the greatest number of seeds, followed by ants, and then by birds. In the Karoo, ants take more seeds than small mammals, as they also do in the Monte of Argentina. Birds may be quite important seasonally. For some species of plants, nearly 100% of the seeds are removed by granivores while other species lose a very small portion. This differential feeding can clearly alter the composition of the plant community. Additionally, for species like the harvester ants that carry seeds back to their nests and belowground, there are dramatic changes in soil properties where these nests exist. Clearly, soil disturbance and granivory, with its concomitant on seed dispersal, have an important role in structuring desert ecosystems.

Humans in the Desert

"Coevolution" is a term evolutionary theorists use to describe a sort of ecological "intimacy," when the evolution of one organism is shaped by and in turn shapes the evolution of another (Ehrlich *et al.*, 1988). It is interesting to consider the evolutionary trajectories of desert biota and of human beings in that light. Relative to the time-scales of the geologic and atmospheric processes that created desert conditions, the advent of humans is a very recent event. Nonetheless, the effect of humans on deserts – and of deserts on humans – is pronounced.

Desert landscapes and desert biota have had profound effects on human cultural evolution. Humans display

remarkable behavioral and cultural adaptations to the aridity and unpredictability of deserts, and traditions derived there have influenced human and biological communities far beyond the desert edge; that the three "religions of the book" had their origins in these environments well-illustrates that fact (Safriel *et al.*, 2006). Plants and animals from these harsh landscapes have also played an important role in the evolution of modern human societies. Dryland biota provided much of the "raw material" for species that could and did become domesticated, which helped usher in the dawn of pastoral and agricultural societies. The early domestication of ungulates (cattle, sheep, and goats) began in the drylands of the Near East, on the edge of the Arabian Deserts, some 9000 years ago (Davis, 2005), and the domestication of llamas and alpacas took place in the Andean Puna of South America some 6000 years ago just north of South America's "arid diagonal" formed by the Atacama, Dry Puna, and Monte deserts (**Table 3**). In many regions of the world, dryland annuals have been at the base of the plant domestication process and drylands have been the cradle of agricultural societies. The first records of cultivated wheat and barley (two dryland ephemerals) come from the Fertile Crescent of the Near East some 7000–9000 years ago. In the American Continent the first agricultural records come from the Tehuacán Valley in southern Mexico, a hot tropical dryland where corn and squash (two annual, drought-tolerant fast growers) were first domesticated some 6000 years ago. Not too long after that, gatherers in the Andean Puna started domesticating two other dryland ephemerals: the quinoa (*Chenopodium*, a fast-growing annual) and the potato (*Solanum*, a tuber ephemeral).

That biota of drylands would be a source of such innovation is not surprising, given the life history of those plants and animals. Arid-land herbivores, in particular desert ungulates, are extremely hardy. They can use water very efficiently, they can withstand long periods without drinking, and when forage is plentiful they can convert plant material quickly into animal protein, with very high efficiency. Furthermore, many of them are migratory and move naturally in herds following a leader, looking for new foraging grounds, and socially protecting themselves from predators. For all these evolutionary reasons, ungulates native to drylands were ideal candidates for domestication: hardy animals, efficient foragers, and amenable to shepherding, as social aggregation is a natural behavior for them. Some of the same factors that made wild goats, mountain sheep, or guanacos evolutionarily adapted to desert environments are what drove early hunter-gatherers to start breeding their offspring and selecting them for desirable domestic attributes. As with desert ungulates, the same traits that have made some desert annuals apt to survive and thrive on ephemeral water pulses are what make them so apt for agriculture: fast growth, short life cycle, and the capacity to direct most of their metabolic budget toward the abundant production of seeds. Because dryland ephemerals grow so fast and produce so much seed in just a few weeks, they grow at an amazingly fast rate when planted at the desert's edge and make ideal grain plants, especially cereals and pulses.

The effect of humans on the ecology and the evolutionary trajectory of deserts can be similarly pronounced. The following ecological "anachronism" provides an illustrative example: some desert plants have seed dispersal mechanisms that reflect the existence of seed dispersers that are no longer present. Trees like the mesquites (*Prosopis*), for example, have pods with nutritious sweep pulp and extremely tough seeds, which need intense scarification in order to germinate. Similarly, the tough seeds of the prickly pears (*Platyopuntia*) germinate successfully only when chewed and digested for a long time. During the Pleistocene period, this abrasion was provided by the digestive system of large ungulates, such as gomphotheres or giant ground sloths. At the end of the last glaciation some 15,000 years ago, however, much of that Pleistocene megafauna went extinct – a fate that humans likely contributed to (Martin, 1973; Martin and Klein, 1989; Alroy, 2001; Brook and Bowman, 2004). Loss of that fauna resulted in the loss of seed dispersal and regeneration mechanisms for a number of plant species. Desert plant species with anachronistic seed dispersal have merely survived for the last millennia through vegetative growth and accidental abrasion of seeds in the deserts' sand and gravel, in the absence of their effective seed dispersers. Not surprisingly, when humans reintroduced ungulates – cattle – into the New World some five centuries ago, the population of many of these plant species rebounded to large numbers.

Humans continue to affect desert ecology, at times fundamentally. Being areas of such low productivity, deserts can be easily degraded – even irreparably – by the increasing intensity of human land and resource use. Desert soils, which are of

Table 3 Domestic mammals evolved from wild dryland ancestors

Domestic form		*Wild progenitor*		*First domestication*		*Distribution of wild progenitor*
Common Name	*Scientific name*	*Common name*	*Scientific name*	*Date*	*Place*	
Perissodactyla						
Donkey	*Equus asinus*	African ass	*Equus africanus*	4000 BC	Egypt	North Africa, possibly West Asia
Artiodactyla						
Llama	*Lama lama*	Guanaco	*Lama guanicoe*	4000 BC	Peruvian Andes	South American drylands
Alpaca	*Lama pacos*	Guanaco	*Lama guanicoe*	4000 BC	Peruvian Andes	South American drylands
Dromedary	*Camelus dromedarius*	Dromedary	*Camelus* sp.	3000 BC	West Asia	Asia, possibly North Africa
Bactrian camel	*Camelus bactrianus*	Bactrian camel	*Camelus ferus*	3000 BC	Central Asia	Central Asia
Goat	*Capra hircus*	Wild goat	*Capra aegagrus*	7000–8000 BC	West Asia	West Asia
Sheep	*Ovis aries*	Mouflon	*Ovis orientalis*	7000–8000 BC	West Asia	West Asia

generally limited profundity and high fragility, are highly susceptible to compaction, erosion, and salinization when exploited for agricultural, industrial, or recreational purposes. Invasive non-native plants, whether introduced intentionally (such as in the case of the planting of grasses for livestock forage, which have the effect of introducing a grass-fire cycle to an ecosystem that has no natural fire regime) or not (such as the case of the invasion of *Tamarix ramosissima* in Nearctic deserts, which can substantially alter desert hydrological regimes), can have cascading effects on ecosystems function and native species viability in deserts. Human industry in and beyond deserts alters not only desert weather patterns via anthropogenic-induced climate change, but also desert nutrient cycling via atmospheric deposition. Paradoxically, fertilization of deserts through increased deposition of nutrients like nitrogen can favor the invasive dispersal of non-native species and reduce native diversity. Whether through direct or indirect pathways, humans clearly have a hand in determining the future course of desert evolution.

Concluding Remarks

To the untrained eye, deserts look barren, especially during dry periods. However, because of their evolution in relative geographic isolation, most deserts of the world are rich in rare and endemic species, and are hence highly vulnerable to biological extinction and environmental degradation. In spite of their remarkable convergence in adaptation, deserts are different in their origin and their evolutionary history. Their incredible variation of the world's deserts in rainfall patterns, continentality, temperature regime, and evolutionary history have all contributed not only to their biological uniqueness, but also to their wondrous wealth of life-forms and adaptations, from some of the shortest-lived ephemeral plants, to some of the longest lived giant cacti; from seed-eating rodents that do not need water to survive and depend on their burrows to regulate their metabolism almost as if the burrow was an extended part of their body, to amazing pollinators like nectar-feeding bats that migrate thousands of miles following the flowering seasons (Davis, 1998). This adaptive diversity – what Darwin, strongly influenced by deserts himself, called "forms most beautiful and most wonderful" – is what makes deserts unique. In the hot deserts we may find giant cacti and trees with mammoth fleshy stems coexisting with some of the toughest hardwoods; ground-creeping succulents side by side with fog-harvesting rosettes, incredibly fast-growing annuals together with the hardiest drought-resistant perennials; shrubs of enticing odors with some of the nastiest, spiniest plants. Very few parts of the earth contain a richer collection of natural adaptations (Ezcurra *et al.*, 2006; Ezcurra and Mellink, 2005).

The fragmented evolutionary history of the deserts of the world has been the driving force of their biological rarity, of adaptation to local conditions, of specialization to isolated environments. After millions of years in isolation, the forces of evolution and fragmentation have yielded unique life-forms in each desert, strangely shaped desert plants and extraordinary animals. The world's deserts are biological and cultural islands, lands of fantasy and adventure, habitats of surprising, often bizarre growth forms, and territories of immense natural beauty.

See also: Africa, Ecosystems of. African and Asian Savannas. Agriculture, Traditional. Asia, Ecosystems of. Biodiversity of the Succulent Karoo. Biodiversity-Rich Countries. Cattle, Sheep, and Goats, Ecological Role of. Central America, Ecosystems of. Desertification. Domestication of Crop Plants. El Niño and Biodiversity. Grazing, Effects of. Hunter-Gatherer Societies, Ecological Impact of. Indigenous Peoples and Biodiversity. Land-Use Issues. Mediterranean-Climate Ecosystems. Near East Ecosystems, Animal Diversity. Near East Ecosystems, Plant Diversity. Neotropical Seasonally Dry Forests. North America, Patterns of Biodiversity in. Paleoecology. Plant Invasions. Plant–Animal Interactions. Pollinators, Role of. South American Natural Ecosystems, Status of. Terrestrial Ecosystems

References

Allan JA, Warren A, Tolba M, and Allan T (1993) *Deserts: The Encroaching Wilderness (A World Conservation Atlas).* Oxford: Oxford University Press.

Alroy J (2001) A multispecies overkill simulation of the end-Pleistocene megafaunal mass extinction. *Science* 292: 1893–1896.

Axelrod DI (1950) Evolution of desert vegetation in western North America. *Publications of the Carnegie Institute of Washington* 590: 215–306.

Brook BW and Bowman DMJS (2004) The uncertain blitzkrieg of Pleistocene megafauna. *Journal of Biogeography* 31: 517–523.

Chesson P, Gebauer RLE, Schwinning S, *et al.* (2004) Resource pulses, species interactions, and diversity maintenance in arid and semi-arid environments. *Oecologia* 141: 236–253.

Cloudsley-Thompson JL (1996) *Biotic Interactions in Arid Lands.* Berlin, Heidelberg: Springer.

Cooke RU, Warren A, and Goudie AS (1992) *Desert Geomorphology.* London: University College Press.

Davis SJM (2005) Why domesticate food animals? Some zoo-archaeological evidence from the Levant. *Journal of Archaeological Science* 32(9): 1408–1416.

Davis W (1998) *Shadows in the Sun: Travels to Landscapes of Spirit and Desire.* Washington, DC: Island Press.

Dimmit MA (2000) Biomes and communities of the Sonoran Desert Region. In: Phillips SJ and Comus PW (eds.) *A Natural History of the Sonoran Desert*, pp. 3–18. Tucson: Arizona-Sonora Desert Museum and University of California Press.

Ehrlich PR, Dobkin DS, and Wheye D (1988) *The Birder's Handbook.* New York: Simon and Schuster.

Ezcurra E and Mellink E (2005) Hot deserts. In: *Encyclopedia of Life Sciences. Online Encyclopedia* Chichester, UK: John Wiley and Sons.

Ezcurra E, Mellink E, Wehncke E, *et al.* (2006) Natural History and Evolution of the World's Deserts. In: Ezcurra E (ed.) *Global Deserts Outlook*, pp. 1–26. Nairobi, Kenya: United Nations Environment Programme (UNEP).

Ezcurra E, Montaña C, and Arizaga S (1991) Architecture, light interception, and distribution of Larrea species in the Monte Desert, Argentina. *Ecology* 72(1): 23–34.

FAO (2002) *Pastoralism in the New Millennium.* FAO Animal Production and Health Papers No. 150. Rome: Food and Agriculture Organization of the United Nations.

FAO (2004) *Carbon Sequestration in Dryland Soils.* World Soils Resources Reports, No. 102. Rome: Food and Agriculture Organization of the United Nations.

GLOBIO (2005) *Global Methodology for Mapping Human Impacts on the Biosphere.* Norway: UNEP/GRID-Arendal. http://www.globio.info/region/world/

Goudie A and Wilkinson J (1977) *The Warm Desert Environment.* Cambridge: Cambridge University Press.

Holmgren M, Scheffer M, Ezcurra E, Gutiérrez JR, and Mohren GMJ (2001) El Niño effects on the dynamics of terrestrial ecosystems. *Trends in Ecology and Evolution* 16(2): 59–112.

Loik ME, Breshears DD, Lauenroth WK, and Belnap J (2004) A multi-scale perspective of water pulses in dryland ecosystems: Climatology and ecohydrology of the western USA. *Oecologia* 141: 269–281.
Louw GN and Seely MK (1982) *Ecology of Desert Organisms*. London: Longman.
Mabbutt JA (1977) *Desert Landforms*. Cambridge, Massachusetts: MIT Press.
MacMahon JA (2001) Desert Ecosystems. In: Levin S (ed.) *Encyclopedia of Biodiversity*, vol. 2. pp. 37–59. San Diego, CA: Academic Press.
Mares MA (1980) Convergent evolution among desert rodents: A global perspective. *Bulletin of the Carnegie Museum of Natural History* 16: 1–51.
Martin PS (1973) The Discovery of America. *Science* 179(4077): 969–974.
Martin PS and Klein RG (eds.) (1989) *Quaternary Extinctions: A Prehistoric Revolution*. Tucson, Arizona: University of Arizona Press. 892pp.
Martorell C and Ezcurra E (2002) Rosette scrub occurrence and fog availability in arid mountains of Mexico. *Journal of Vegetation Science* 13: 651–662.
McGinnies WG, Goldman BJ, and Paylore P (eds.) (1977) *Deserts of the World*. Tucson: University of Arizona Press.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
Montaña C, Ezcurra E, Carrillo A, and Delhoume JP (1988) The decomposition of litter in grasslands of northern Mexico: A comparison between arid and non-arid environments. *Journal of Arid Environments* 13: 551–556.
Morton RR (1979) Diversity of desert-dwelling mammals: A comparison of Australia and North America. *Journal of Mammalogy* 60: 253–264.
Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience* 51(11): 933–938. http://www.worldwildlife.org/science/ecoreigons/terrestrial.(fm).
Pianka ER (1986) *Ecology and Natural History of Desert Lizards: Analysis of the Ecological Niche and Community Structure*. Princeton, NJ: Princeton University Press.
Pipes R (1998) *Hot Deserts (World Habitats)*. New York: Raintree.
Ricciuti ER (1996) *Desert (Biomes of the World)*. New York: Benchmark Books.
Robichaux RH (ed.) (1999) *Ecology of Sonoran Desert Plants and Plant Communities*. Tucson: University of Arizona Press.
Safriel U (1990) Winter foraging behaviour of the Dune Lark in the Namib Desert, and the effect of prolonged drought on behaviour and population size. *Ostrich* 61: 76–80.
Safriel U, Ezcurra E, Tegen I, *et al.* (2006) Deserts and the Planet: Linkages between Deserts and Non-Deserts. In: Ezcurra E (ed.) *Global Deserts Outlook*, pp. 49–72. Nairobi, Kenya: United Nations Environment Programme (UNEP).
Sanderson EW, Jaiteh M, Levy MA, Redford KH, Wannebo AV, and Woolmer G (2002) The human footprint and the last of the wild. *BioScience* 52(10): 891–904.
Schmidt-Nielsen K (1964) *Desert Animals: Physiological Problems of Heat and Water*. Oxford: Clarendon Press.
Sher AA, Goldberg DE, and Novoplansky A (2004) The effect of mean and variance in resource supply on survival of annuals from Mediterranean and desert environments. *Oecologia* 141: 353–362.
Shmida A (1985) Biogeography of Desert Flora. In: Evenari M, Noy Meir I, and Goodall D (eds.) *Ecosystems of the World*, vol. 12. pp. 23–75. Amsterdam: Elsevier.
Thornthwaite CW (1948) An approach toward a rational classification of climate. *Geographical Review* 38: 55–94.
UNEP (1997) *World Atlas of Desertification*, 2nd edn. Nairobi, Kenya: United Nations Environmental Programme.
USGS (2005) *Global Land Cover Characteristics Data Base. Earth Resources Observation and Science (EROS)*. Sioux Falls, South Dakota: United States Geological Survey. http://edcsns17.usgs.gov/glcc/globdoc2_0.html.
Zavala-Hurtado JA, Vite F, and Ezcurra E (1998) Stem tilting and pseudocephalium orientation in *Cephalocereus columna-trajani* (Cactaceae): A functional interpretation. *Ecology* 79(1): 340–348.

Desertification

James F Reynolds, Duke University, Durham, NC, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Combating desertification Activities specifically aimed at prevention and/or reduction of land degradation, rehabilitation of partly degraded land, and reclamation of desertified land.

Desertification Land degradation in drylands resulting from various factors, including climatic variations and human activities.

Dimensions of desertification The interactions and feedbacks of meteorological, ecological, and human components of land degradation.

Dryland Development Paradigm A conceptual framework to aid in understanding the complexity of desertification and dryland development by identifying and synthesizing those factors important to research, management, and policy communities.

Drylands Arid and semiarid croplands, pastures, rangelands, and subhumid woodlands where the index of aridity is less than 0.65; they cover more than two-fifths of the land surface of the globe, and are home to 35% of the human population.

Index of aridity Ratio of mean annual precipitation to mean annual potential evapotranspiration.

Land The terrestrial ecosystem that encompasses soils, vegetation, other biota, and the ecological, biogeochemical, and hydrological processes that operate therein.

Land degradation Reduction or loss of the biological and economic productivity and complexity of terrestrial ecosystems, including soils, vegetation, other biota, and the ecological, biogeochemical, and hydrological processes that operate therein.

Myths of desertification Controversy stemming from failure to consider all dimensions (meteorological, ecological, and human) of the problem of land degradation, and the alarmist tone connoted by the word "desertification," which incorrectly suggests the action of deserts "moving" across a landscape, engulfing fertile lands and leaving starving people in their wake.

Introduction

Land Degradation in Global Drylands

Water covers 70% of the Earth's surface, with the remaining 30% covered by land. In turn, more than 40% of this terrestrial land area is classified as drylands, which exist on all continents, but are most widespread in Africa and Asia (**Figure 1**). Drylands are home to 2 billion humans or 35% of the total population of the Earth, 90% of whom reside in developing countries. The fate of many of these people is ultimately dependent on the availability (and efficient utilization) of natural resources, for example, water, soils, plants, and wildlife. However, unsustainable use of water, poor agriculture practices, deforestation, and climate change have resulted in the degradation of extensive land areas of drylands.

It is estimated that approximately 6–12 million km^2 ($\sim$10–20%) of all drylands suffer from serious land degradation. This directly impacts the food security, health, material needs, and other aspects of the livelihoods of 250–300 million people (Millennium Ecosystem Assessment, 2005; Reynolds *et al.*, 2007a). Many more drylands are potentially at risk, especially as populations continue to grow. Dryland degradation currently costs developing countries approximately 4–8% of their gross annual domestic product (EMG, 2011).

Although present in all biomes, land degradation in drylands is generally referred to as "desertification." Desertification has occurred over large areas in Asia, the Mediterranean, Africa, Oceania, and the Americas, involving the loss of the structural and functional integrity of soils, vegetation, other biota, and those ecological, biogeochemical, and hydrological processes that operate therein. Typical consequences include the elimination or major reduction of the endemic plant and animal species, a decrease in the infiltration of precipitation into the soil, inefficient nutrient cycling, water erosion, and increases in soil salinity (**Figure 2(a–i)**). These changes have direct impacts on the human populations who rely on these ecosystems for their livelihoods.

The Creation of Deserts: Reality versus Myths

Deserts on the Move!

The topic of desertification has evoked much confusion and controversy over the years (Reynolds and Stafford Smith, 2002). Perhaps the number one culprit is the alarmist tone connoted by the word "desertification." The French forester Andre Aubréville is credited with coining the term desertification in 1949. Aubréville's usage was in the context of the indiscriminate felling and burning of forests and woodlands in subhumid Africa and the accelerated soil erosion that ensued: that is, what he and his colleagues saw as the "creation of deserts." It evokes an image of deserts literally moving across a landscape, engulfing fertile lands and leaving starving people in their wake. These popular images have their roots in a series of papers written in the 1930s by the British forester Stebbing, who used titles such as "The encroaching Sahara" and "The man-made desert in Africa."

In the late 1970s, desertification was first brought into the public spotlight when the United Nations (UN) announced that more than one-third of the terrestrial land surface of the

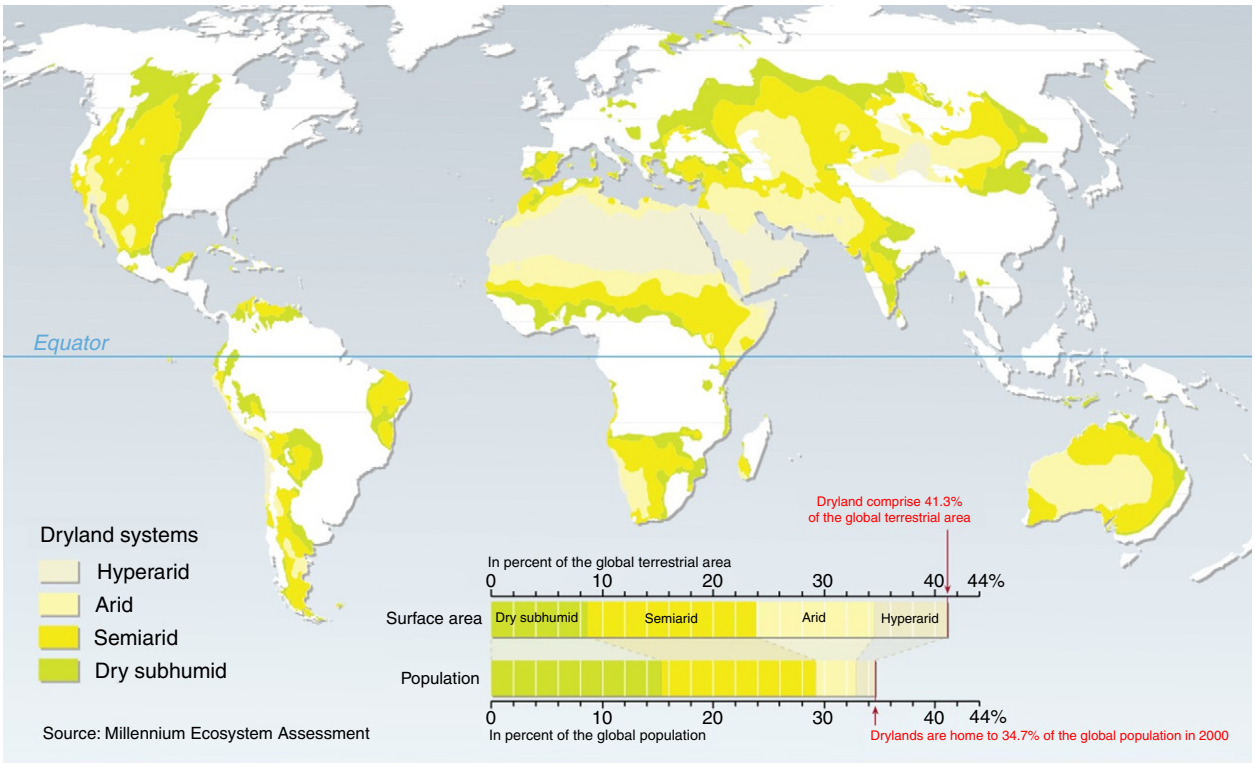


Figure 1 Drylands of the globe as defined by the aridity index, the long-term mean of the ratio of an area's mean annual precipitation to its mean annual potential evapotranspiration. This map is based on data from United Nations Environmental Programme Geo Data Portal. Global area based on Digital Chart of the World data (147,573,196.6 km^2). Data presented in the figure are from the Millennium Assessment core database for the year 2000. Data from Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-being: Desertification Synthesis*. Washington, DC: World Resources Institute.

globe was "at risk of desertifying.". Africa was identified as a particularly serious region of concern. This coincided with a series of major droughts in sub-Saharan Africa that, when combined with weak economies and civil strife in the region, resulted in widespread famine and the deaths of millions. The UN held a high profile Conference on Desertification in Nairobi in 1977 and adopted a Plan of Action to Combat Desertification (PACD) "to curb this onslaught."

The droughts and famines of the 1970s served to reinforce the impression that massive deserts were being formed in drylands. The word "desertification" was applied indiscriminately, even for climatic regions that were natural deserts. To varying degrees, this myth persists to this day, especially in the general news media and also in some policy and scientific circles. Provocative news articles with titles such as "Droughts, Deserts and Death," "Sahara Jumps Mediterranean into Europe," and "Greedy Sahara Devours Land Along Its Border" certainly evokes emotions but not necessarily reality (Reynolds and Stafford Smith, 2002).

The root problem is that land degradation is not a simple phenomenon. One key misunderstanding stems from the fact that drylands are fundamentally fragile ecosystems and, hence, highly susceptible to disturbance. A decrease in vegetation cover due to a short-term response to drought, for example, those captured by satellite images in the Sahara, and soil degradation (a longer-term or permanent response resulting from chronic disturbances) are distinct and not necessarily related phenomena. The former is reversible when the drought is over, but severe land degradation may not be so. Satellite images of the dynamics of plant cover on the margins of the Sahara desert have clearly demonstrated that its "boundary" is quite dynamic (**Figure 3**): There are periods of "expansion" and "retreat," which coincide with the severity and duration of droughts. Another example of short- versus long-term dynamics is given in the section Dimensions of Desertification.

A Stakeholders Perspective

To a large extent, the degree to which land is considered degraded in a particular area depends on the priorities of local users. Clearing native woodland in order to grow crops may increase soil erosion and negatively impact the economic business of firewood collection, but increase food production. With respect to land degradation, changes in the characteristics of soils, water, and vegetation do not have a linear relationship to the productivity potential of the land. Reynolds and Stafford Smith (2002) provide an example to illustrate how different stakeholders may view land degradation through the lens of their particular needs and priorities.

Imagine a cattle ranch in central Mexico, where a herd of cattle is grazing in a pasture that has a large number of erosion gullies (e.g., **Figure 1(b)**). Whereas one might deduce that these gullies are the result of overgrazing, which removes plants and exposes the soil to water and wind erosion



Figure 2 Images of land degradation in drylands. (a) Plowed field in Salar de Uni region of southern Bolivia, leaving soil highly susceptible to wind and water erosion. (b) Erosion gullies being formed in cultivated landscape north of Port Elisabeth, South Africa. (c). Cattle and sheep disturbance around a watering hole in southern New Mexico. (d) Terracing of hillsides on Loess Plateau (China) to curb soil erosion. (e) Mixed-use cultivation and goat grazing in village of La Amapola, Mexico. (f) Woody plant encroachment in native grasslands of central Argentina and dust storm from resulting exposure of sensitive soil surface. (g) Revegetation efforts to restore plant cover near La Sepultura (Chiapas), Mexico. (h) Dune formation resulting in loss of plant cover due to overgrazing in Negev Desert, Israel. (i) Wood gathering in Ivory Coast, Africa. All photos by author except (f) (courtesy of R. Distel), and (i) (courtesy of UNESCO-MAB).

(which is usually true), alternative stake-holder views of these gullies are possible (**Figure 4**):

1. the erosion gullies are the result of natural phenomena (wind and water);
2. in some landscapes gullies, whether natural or induced by overgrazing (and depending on the severity and length of time since forming), may have no effect on what matters in terms of human values (such as meat production);
3. although the erosion gullies may not connote a loss in meat production on this ranch *per se,* they may be creating salinity problems downstream from the ranch or threaten the survival of endemic plant species; and
4. even if the gullies are the direct result of overgrazing, there is room for debate as to whether the root cause is overstocking by the rancher, a product of the land tenure system of this area, climate-driven, or an indication of broader institutional problems (or possibly any combination of these).

An ecologist will view erosion gullies as representing a loss of ecosystem structure and function, for example, the ability of the soil to retain water, nutrient cycling, long-term soil stability, and reduced forage production. Although true, this may resonate with a local rancher only if it has a demonstrable impact on his beef production. Of course, broader concerns for appearances or genuine environmentalism may also arise, but the rancher would unlikely invest in gully control on these grounds alone unless, for example, it was linked to an environmental accreditation scheme for the ranch. Gullies could also matter to a local tourist operator who finds that ecotourists are put off by the appearance of environmental damage. Simply put, different stakeholders may see this problem differently.

The Dryland Syndrome

Reynolds *et al.* (2007a) posited that drylands tend to have a common nexus of environmental, ecological, and social–economic characteristics. They termed this as "dryland syndrome." Some of the main features of the dryland syndrome are presented below.

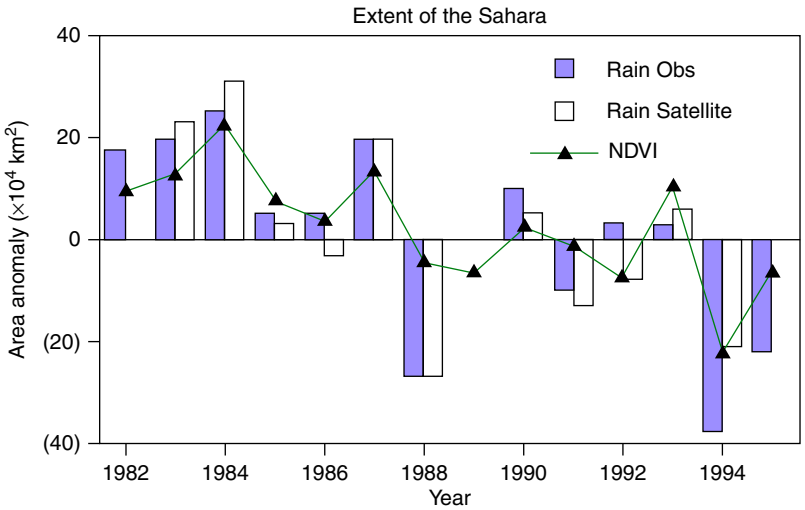


Figure 3 Extent of the Sahara Desert, shown in terms of annual deviations from average size defined by a 200-mm rainfall isohyet extending to the north of 25° N. Calculations based on rainfall data collected at 141 locations throughout West Africa (Rain Obs), from Meteosat satellite data (Rain Satellite) and from the NDVI (a simple ratio between the red and near-infrared reflectance bands obtained from the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellite). The NDVI correlates well with percent vegetation cover and biomass and thus the interannual fluctuations of the desert boundary, as assessed from NDVI, generally mimics that of rainfall. This work by Sharon Nicholson and coworkers at Florida State University shows considerable agreement between estimates based on rainfall and those based on NDVI and demonstrates that there is no "march" of the desert over West Africa. Plotted from data in Nicholson SE, Tucker CJ, and Ba MB (1998) Desertification, drought, and surface vegetation: An example from the West African Sahel. *Bulletin of the American Meteorological Society* 79: 815–829.

Environmental/Ecological Characteristics

Drylands occur on all continents (**Figure 1**). Annual precipitation is quite low, often consisting of high-intensity events that occur over short periods of time. Rainfall is also highly variable throughout the year. Most, if not all, of the annual precipitation is lost from the system via evapotranspiration (evaporated from the soil or transpired by plants), rather than via drainage to ground water or surface runoff. This is due to the high evaporative demand of the atmosphere resulting from high air temperatures, low humidity, and abundant solar radiation, that is, the potential for evapotranspiration is high. It is estimated that only 8% of the world's renewable freshwater supply is available to drylands, which, on a *per capita* basis, is 30% less than the minimum needed for human well-being and sustainable development (Safriel *et al.*, 2005).

Quantitatively, drylands are determined by the ratio of mean annual precipitation (P) to mean annual potential evapotranspiration, which is known as the index of aridity (IA). If the IA of a region is less than 0.65, it is considered a dryland. As shown in **Figure 1**, drylands are classified into four types: hyperarid (IA < 0.05; examples include the Atacama, Arabian, and Sahara deserts), arid (0.05–0.20), semiarid (0.20–0.50), and dry subhumid (0.50–0.65).

The greatest concentrations of drylands are in Asia and Africa, which combined account for almost 60% of all drylands. The remaining drylands, especially when expressed on a global basis, appear relatively small (e.g., 3% in Europe and 10% in South America). However, these numbers can be deceptive: in Australia, drylands represent only 15% of the world's total but cover nearly 90% of the land surface area of the continent.

Drylands are classified as rangelands, croplands, or urban (**Figure 5**). Drylands tend to have low amounts of soil organic matter and low soil fertility. In spite of this, and the xeric nature of drylands, more than 44% of all cultivated land in the world is found within drylands (EMG, 2011). Most of the cropping systems are concentrated on the more mesic range of the gradient (in dry subhumid regions). Another notable feature is that endemic dryland plant species constitute 30% of all plants cultivated in the world (**Figure 6**).

Social–Economic Characteristics

Drylands are home to more than 2 billion people, or approximately a third of the total human population. Approximately 2% live in the hyperarid regions, 4% in arid, 14% in semiarid, and 15% in dry subhumid (**Figure 2**). Importantly, 90% of these 2 billion people live in developing countries, including 40% of the human populations of both Africa and Asia. The latter statistic is significant in that the UN World Population Prospects (2010 Revision) estimates that there will be an increase of approximately 2.3 billion in the global population by 2050, with the majority of growth attributable to "high-fertility countries" located in Africa and Asia.

Although urban systems constitute only a tiny fraction of drylands (**Figure 5**), they contain nearly 45% of its total population. In fact, eight of the world's 20 megacities are located in dryland regions, for example, Cairo, Mexico City, and New Delhi. However, the remaining 1 billion people live in rural areas and thus rely directly on the land for their livelihoods. Their fate is ultimately dependent on the availability of ecosystem goods and services via rainfed or irrigated farming, harvesting native plant products, hunting, and pastoralism. Potential climate changes, limited resources, and increasing

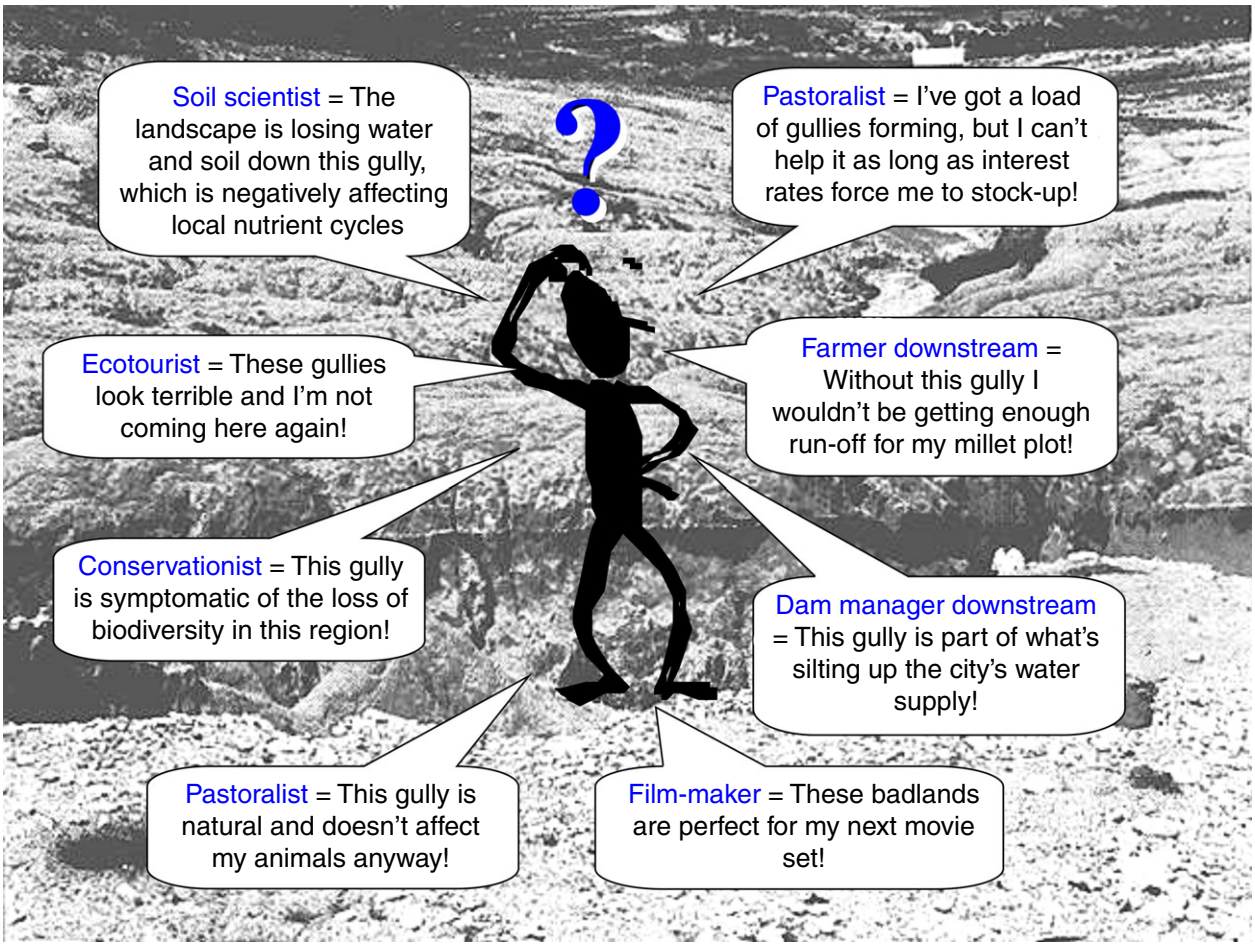


Figure 4 Vastly different perceptions by state-holders arising from concerns over erosion gullies, illustrating how land degradation is a concept that resides in the "eye of the beholder."

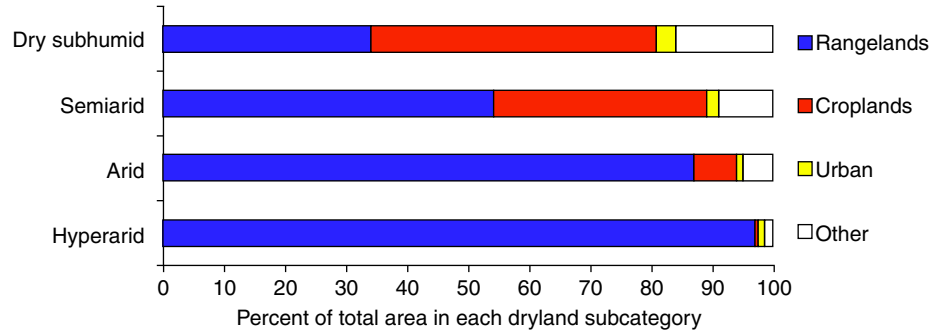


Figure 5 Land use in drylands. Reproduced with permission from United Nations Environmental Programme (UNEP) (ed.) (1997) United Nations Environmental Programme. In: Middleton NJ and Thomas DSG (Coordinating Editors) *World Atlas of Desertification*, 2nd edn. New York: Edward Arnold, and Safriel U, Adeel Z, Niemeijer D, *et al.* (2005) Dryland systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystems and Human Well-being: Current State and Trends*. The Millennium Ecosystem Assessment Series, pp. 623–662. Washington, DC: Island Press.

population size increases the pressure on limited dryland resources, which can in turn lead to rapid land degradation threatening people's livelihoods.

As compared to more mesic systems, the distribution of these rural populations tend to be sparse and more mobile, and there is a scarcity of human capital (e.g., labor). They are also apt to be physically isolated, that is, (1) located far from markets (and hence higher transaction costs) and (2) bureaucratic decisions are made by distant policymakers. As a result, dryland populations tend to lag behind populations in



Figure 6 Quinoa (*Chenopodium quinoa*) field in the Salar de Uyuni region of southern Bolivia. Quinoa has long been a staple in the diets of indigenous people of the Altiplano in the central Andes. This endemic plant played an important role in pre-Columbian Andean civilizations because of its very high protein content. Demand for quinoa in organic stores of North America and Europe in recent decades has resulted in the rapid expansion of the cultivation of this plant throughout the Altiplano. Although an economic boom, this has resulted in various forms of land degradation (due to overcultivation, the introduction of tractors, etc.; see **Figure 2(a)**) and has impacted the livelihoods of local populations in many ways (e.g., changes in diet, community structure). Ironically, the high foreign demand for quinoa means it is now less accessible for Bolivians because prices have been pushed-up.

other parts of the world with regard to many economic and health indices, such as having higher infant mortality rates, severe shortages of potable water, and lower *per capita* income (Reynolds *et al.*, 2007a,b).

Combating Desertification

Formation of the United Nations Convention to Combat Desertification (UNCCD)

Despite the work of the 1977 UNC PACD, as well as numerous other international efforts, by 1991 it was concluded that the PACD was not working and the problem of land degradation in the world was actually intensifying, rather than abating. As a result, desertification remained a major issue for the UN when they convened for the Earth Summit in Rio de Janeiro 15 years later.

At the 1992 Rio Summit, the loss of biodiversity, climate change, and desertification were singled-out as the greatest threats to humankind and sustainable development. For each, legally-binding international agreements were forged, known respectively as the Convention on Biological Diversity, the United Nations Framework Convention on Climate Change, and the UNCCD. The UNCCD (established in 1994) deals with land degradation in arid, semiarid, and dry subhumid areas of the globe, collectively known as drylands, where some of the most vulnerable ecosystems and peoples live. In fact, the Millennium Ecosystem Assessment (2005) concluded that "desertification is potentially the most threatening ecosystem change impacting livelihoods of the poor."

Defining Land Degradation and Desertification

The UNCCD (1994) formally defined desertification as the "… reduction or loss of the biological and economic productivity and complexity of terrestrial ecosystems, including soils, vegetation, other biota, and the ecological, biogeochemical and hydrological processes that operate therein… resulting from various factors including climatic variations and human activities." Land degradation was defined as a "reduction or loss, in arid, semiarid, and dry subhumid areas, of the biological or economic productivity and complexity of rainfed cropland, irrigated cropland, or range, pasture, forest, and woodlands resulting from land uses or from a process or combination of processes, including processes arising from human activities and habitation patterns, such as: (1) soil erosion caused by wind and/or water; (2) deterioration of the physical, chemical, and biological or economic properties of soil; and (3) long-term loss of natural vegetation." Although hundreds of definitions for desertification and land degradation have been proposed (review in Reynolds and Stafford Smith, 2002), we believe the UNCCD's definition to be most straightforward and unambiguous.

Combating Desertification

As its title suggests (i.e., "to combat desertification"), the Convention adopted a proactive approach. What are the essential elements involved in "combating" desertification? The UNCCD adopted the following definition: "combating desertification includes activities that are part of the integrated development of land … for sustainable development. These are activities specifically aimed at: (1) prevention and/or reduction of land degradation; (2) rehabilitation of partly degraded land; and (3) reclamation of desertified land."

Although there are language and contextual differences, many of the goals and objectives of the UNCCD are not unlike those contained in the 1977 PACD. Hence, it is crucial to understand why the PACD failed. Three major reasons are

Table 1 Desertification involves ecological, meteorological, and human dimensions. These do not operate in isolation from each other but, rather, are closely coupled via feedbacks, synergies and interactions. Some key questions usually associated with each dimension are given here

Dimension	*Key questions (examples)*
Meteorological	Will changes in surface energy and water balance caused by changes in land cover significantly affect vegetation? Can changes in near surface climate caused by land degradation be translated into a recognizable contribution to global-scale changes, that is, a desertification "warming signal?" Do bio-geophysical feedback-based climate models contain sufficient mechanism to accurately predict droughts? What is the relationship between the loss of vegetation, albedo, surface roughness, and drought? Will global climate change exacerbate the high natural variability of precipitation and temperature in drylands?
Ecological	Are short-term ecosystem dynamics (e.g., decreases in plant cover) indicative of desertification? Do long-term changes in dryland ecosystems alter the resource base of the entire system such that it moves beyond a threshold whereby degradation accelerates and becomes irreversible? Do multiple thresholds exist? How do the key physical drivers influence other factors such as animal disease carriers, for example, how do temperature and drought affect insect behavior? When is animal production a function of rainfall (plant production) and when is it decoupled? If global climate change further exacerbates the already high natural variability of precipitation in dryland ecosystems, will this lead to permanent degradation of their productive potential, particularly since there is a lack of "buffering" by large reserves of organic matter in the soils or in woody vegetation? Can "early warning" functional indices (based on soils, vegetation, biota, and ecological, biogeochemical, and hydrological processes) be developed to indicate major physical restructuring of a system symptomatic of land degradation? What impact do exotic species have on land degradation processes? What is the relationship of bush encroachment to desertification processes?
Human	How do human populations in the various natural, seminatural, and intensively-managed dryland ecosystems of the globe, affect those ecosystem goods and services considered vital to the sustainability of human populations? What are the key socioeconomic drivers of land use that lead to desertification? For example, is there a relationship between how land is used (accountability) and ownership (which tends to be low in poor countries and high in rich ones)? Are regional and national programs to combat desertification based on economic integration and sustainability of local people who are most directly affected? How can local people be integrated into the decision-making process? Is the "stake-holder" concept a feasible one at all scales of interest? Is it possible to reconcile the different viewpoints of different stakeholders in relation to desertification issues? What are the respective roles of government, local communities, and land users in maintaining sustainability of drylands? Do ecologically-sound land-use practices lie at the level of recognizing the rights and environmental knowledge of local communities? What resources are needed to accomplish this? What is the response of human activities to different stages of land degradation? Are there adaptive adjustments? What is the role of technology in providing new opportunities for ecological sustainability in drylands of the globe?

most often cited: (1) many processes and issues leading to land degradation, for example, overcultivation of the land, deforestation, overgrazing by livestock, and poverty, have their roots in socioeconomic and political realms, and thus are not amenable to technical or science "solutions" as proposed in the PACD; (2) local populations were not included in drawing up solutions; and (3) there was a failure to integrate the PACD into other existing development programs. These problems reflect the failure to recognize the simultaneous importance of meteorological, ecological, and human dimensions of land degradation (**Table 1**).

Recognizing these shortcomings, the Earth Summit supported a new, integrated approach to the desertification problem, emphasizing programs and strategies to promote sustainable development at the community level. Consequently, the UNCCD places considerable emphasis on the social dimensions of the problem and the role of local peoples and nongovernmental organizations in solving local problems. There is a conspicuous emphasis on the human dimensions (a "grass-root" effort). There is now a clear recognition that regardless of emphasis, dryland degradation – given its ecological, meteorological, and human dimensions (**Table 1**) – requires contributions from a multitude of disciplines to understand and identify useful resolutions (Reynolds *et al.*, 2007a).

10-Year Strategy

In 2007, a new 10-year strategy of the UNCCD (2008–2018) was implemented with the goal "… to forge a global partnership to reverse and prevent desertification/land degradation and to mitigate the effects of drought in affected areas in order to support poverty reduction and environmental sustainability." The new strategy is composed of three general strategic objectives. A list of these objectives and their anticipated benefits are given in **Table 2**.

At the 2011 UNCCD conference in the Republic of Korea, UN Secretary-General Ban Ki-moon reaffirmed the importance of global desertification to the world community: "If we

Table 2 As the UNCCD entered its second decade, it adopted a new 10-year strategic plan (2008–2018) to enhance its implementation (EMG, 2011)

Strategic objectives	*Expected benefits*
To improve the living conditions of affected populations	● People living in areas affected by desertification/land degradation and drought to have an improved and more diversified livelihood base and to benefit from income generated from sustainable land management. ● Affected populations' socioeconomic and environmental vulnerability to climate change, climate variability, and drought is reduced.
To improve the condition of affected ecosystems	● Land productivity and other ecosystem goods and services in affected areas are enhanced in a sustainable manner contributing to improved livelihoods. ● The vulnerability of affected ecosystems to climate change, climate variability, and drought is reduced.
To generate global benefits through effective implementation of the UNCCD	● Sustainable land management and combating desertification/land degradation contribute to the conservation and sustainable use of biodiversity and the mitigation of climate change.
To mobilize resources to support implementation of the Convention through building effective partnerships between national and international actors	● Increased financial, technical, and technological resources are made available to affected developing country Parties, and where appropriate Central and Eastern European countries, to implement the Convention. ● Enabling policy environments are improved for UNCCD implementation at all levels.

protect, restore and manage land and soils we can tackle many challenges simultaneously, such as poverty, food and energy insecurity, biodiversity loss, climate change, forced migration, and geopolitical instability."

Drivers of Desertification

Proximal and Underlying Factors

Desertification is driven by factors that vary from region to region. Geist and Lambin (2004) conducted an analysis of over 130 existing case studies (from throughout the world) in an attempt to elucidate patterns. They identified four major categories of proximal causal agents (i.e., drivers) (**Figure 7**): (1) increased aridity; (2) agricultural impacts, including livestock production and crop production; (3) wood extraction and other economic plant removal; and (4) infrastructure extension, which could be separated into irrigation, roads, settlements, and extractive industry (e.g., mining, oil, and gas). Only 10% of the case studies were driven by a single cause (with approximately 5% due to increased aridity and 5% to agricultural impacts). Approximately 30% of the case studies were attributable to a combination of two causes (primarily increased aridity and agricultural impacts), whereas the remaining cases were combinations of three or all four proximal causal factors.

Certain drivers are more or less important in a given region. The drivers evolve from underlying forces associated with specific socioeconomic and biophysical factors characteristic of a particular region. For example, Geist and Lambin found that two underlying drivers, climate and technological factors (either new technologies or deficiencies in technology) were the key drivers of desertification in the majority (54%) of case studies in southern Europe. In Africa, they found that climate, alone or acting in concert with population demography, was a key driver in the 38% of the case studies whereas in the USA, 50% of the case studies were attributable to a combination of climate and technology drivers or these two factors interacting with economic ones. Desertification processes in Asia, Latin America, and Australia could only be attributed to more complex interactions among four or more underlying forces.

Climate Change

One of the conclusions of the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) is that many dryland areas of the globe will suffer decreases in water resources as a result of climate change. Both Africa and Asia are expected to experience major changes, for example, increasing air temperature and changes in precipitation (total amount and season distribution). Under a range of climate change scenarios, the IPCC posits that arid and semiarid land in Africa will increase by 5–8% by 2080; overall, lower precipitation will lead to lower agricultural production in many African countries. In Asia, especially in the central, south, east, and southeast, freshwater availability is projected to decrease in the next 40 years.

Global warming-induced droughts will likely exacerbate land degradation. People will likely have to endure greater and more frequent weather extremes (droughts, flooding). However, it is important to note that humans have the capacity to adapt. For example, in the Sahel of West Africa there was a decline in average rainfall of more than 30% between the 1960s and the 1990s, yet a rapidly growing human population ultimately survived, demonstrating the capacity to adapt to climatic uncertainty in the future (EMG, 2011).

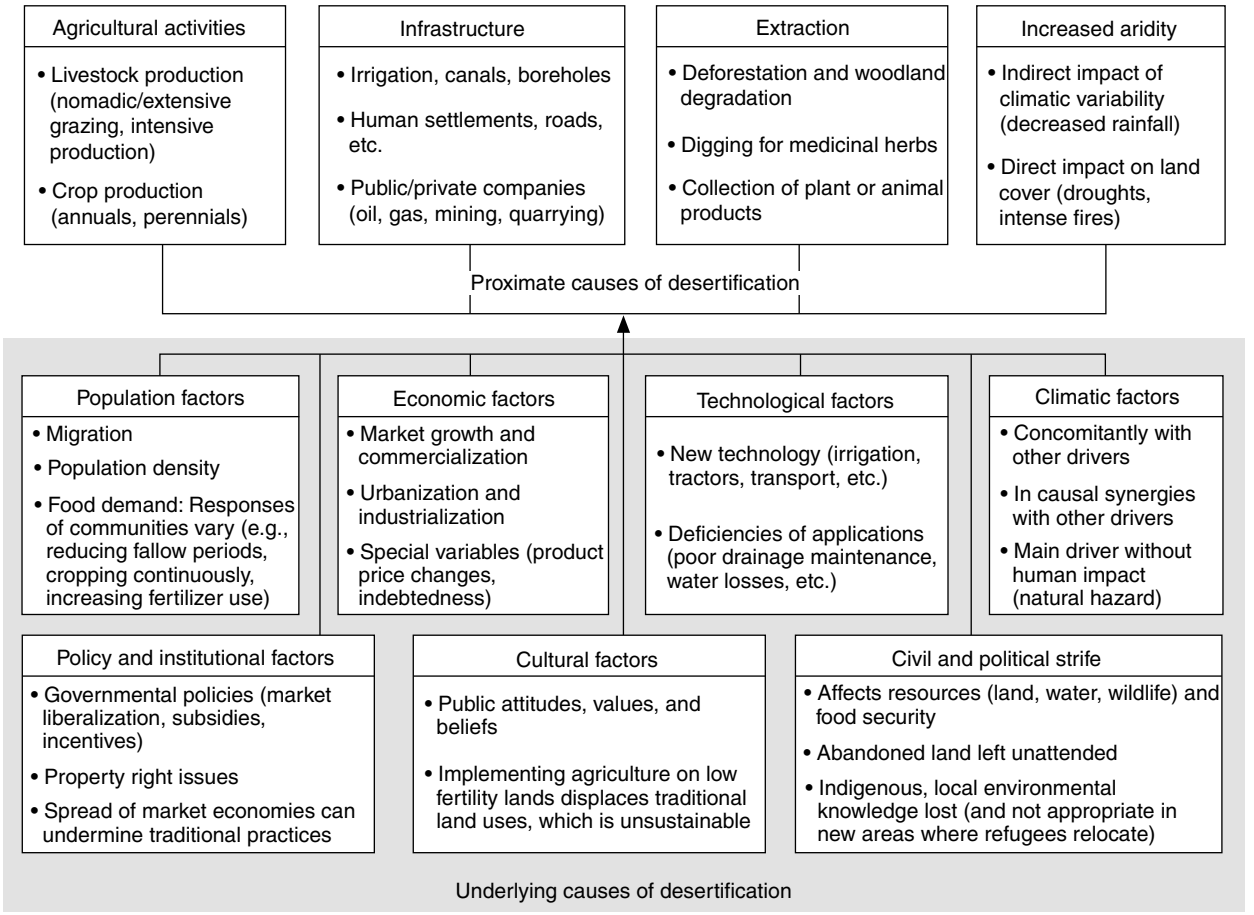


Figure 7 In a global survey Geist and Lambin found repeating causal patterns for drivers of land degradation, which resolved into four major proximate causes explained by six major underlying drivers. Modified from Geist HJ and Lambin EF (2004) Dynamic causal patterns of desertification. *Bioscience* 54: 817–829, with permission from University of California press.

Drylands may also play an important role in mitigation of climate change via the potential for sequestering of carbon in soils. Whereas on a per unit area basis drylands may have a low sequestration rate, their large expanse (~40% of the land area of the globe, **Figure 1**) makes them very important.

Dimensions of Desertification

As alluded to above, we view desertification in terms of its (1) meteorological dimensions, (2) ecological dimensions, and (3) human dimensions (**Table 1**). Each of these is by itself quite complex and difficult to predict. Importantly, the three are highly interdependent. A failure to recognize the simultaneous role of – and feedbacks between – these three dimensions has led to many of the controversies and misconceptions alluded to above. In this article, we explore the concept of desertification based on these critical dimensions.

Ecological Dimensions

Short- versus Long-Term Ecosystem Dynamics

Nearly all drylands are characterized by extreme year-to-year precipitation fluctuations. Hence, it is often difficult to distinguish between short-term variability and long-term changes in ecosystem appearance, as well as between temporary and permanent changes. Short-term variability in precipitation tends to affect the range and frequency of resource pulses, whereas long-term change alters the resource base: that is, the entire system moves beyond some threshold. Once this threshold has been crossed, it is unlikely that the system can be returned to its predisturbed state, without expensive and often extensive remediation efforts. All of this depends on the interactions of numerous climatic, edaphic, and biological factors in combination with the economic feasibility of rehabilitation.

We have stressed the inherent difficulties in teasing out underlying causal factors that give rise to short- versus long-term ecosystem dynamics in drylands (e.g., **Figure 3**). A case study from the Serengeti in Tanzania is an excellent illustration of this. In the 1970s, there was much concern that the Serengeti was degrading and turning into a "desert." The large bush and acacia forests that characterized the Serengeti were declining at an alarming rate. Elephants – who fed on these trees – were largely considered responsible and many believed that a culling program was the only way to save these "pristine" woodlands. In fact, these woodlands were not pristine but instead a relatively recent feature of the Serengeti

reflecting a sequence of complex interactions among ecological factors (e.g., the lifespan of the acacias, the number of wildebeest), human dimensions (cattle grazing, fire setting), and meteorological factors (precipitation patterns) (Figure 8).

Woodlands in the Serengeti started appearing after an outbreak of rinderpest disease in the 1890s, which devastated wildebeest herds and local cattle populations, and which led to mass human starvation. This decline in human populations meant a decrease in fires, which were set to create new pasture for cattle – and fewer fires meant higher establishment of acacia seedlings and the eventual establishment of a forest in the midst of grassland. By the 1920s and 1930s, the practice of setting fires was becoming re-established as the human population recovered; however, the new fires were more frequent and hotter since conditions existed that favored grass growth: low herbivory (wildebeest herds were still low) and a relatively wet climate in the 1960s. These fires resulted in high mortality of acacias seedlings and, by the 1960s, the oldest trees (which have a lifespan of approximately 60–70 years) started to die off. Yet, concerns about the Serengeti turning into a "desert" were abated by the late 1990s as the Serengeti was experiencing a "return of the forest." Heavy poaching of elephants, the recovery of the wildebeest populations, and fewer fires set by humans all favored high recruitment of tree seedlings.

This example illustrates several key elements of the conceptual model of desertification: (1) that short-term dynamics may not necessarily be indicative of long-term phenomena (i.e., system thresholds are extremely difficult to establish without long-term studies); (2) natural ecosystems are complex and difficult to predict; and (3) humans have major impacts on natural systems in often unknown and unpredictable ways. This serves to reinforce the view that land degradation is a complex phenomenon that frequently evades simple characterizations.

Ecological Degradation

Drylands typically have a number of distinguishing ecological characteristics that contribute to their susceptibility to disturbance and, ultimately, to desertification. The natural vegetation is composed of various mixtures of grasslands, shrublands, and savannas, with trees either scattered or concentrated along watercourses. Since vegetation cover is usually relatively sparse, much of the soil is exposed directly to rain, overland flow, sunlight, and wind. Many dryland soils are sensitive to disturbances because they contain small amounts of organic matter and have low aggregate strength. Both tillage and grazing by domesticated animals can have profound effects – in a very short period of time – on these soils, including lowering their permeability to water (thus decreasing infiltration), disturbing their surface integrity (thus, increasing susceptibility to erosion and sedimentation), and decreasing their quality (decreased nutrient status) for plant growth.

Although land degradation is often equated with soil degradation, it is a more general phenomenon that involves whole ecosystems. The UNCCD defines land degradation as the reduction or loss of the biological or economic productivity and complexity. This reduction in the complexity of the land is the major physical restructuring of a system, that is, symptomatic of land degradation. In drylands, this includes erosion and sedimentation by both water and wind, often resulting in a redistribution of topsoil, compacting of the soil, loss of soil silt fraction, dune formation, and gully formation. There may be shifts in natural fire cycles with a disruption of biogeochemical cycling, including the redistribution of essential nutrients,

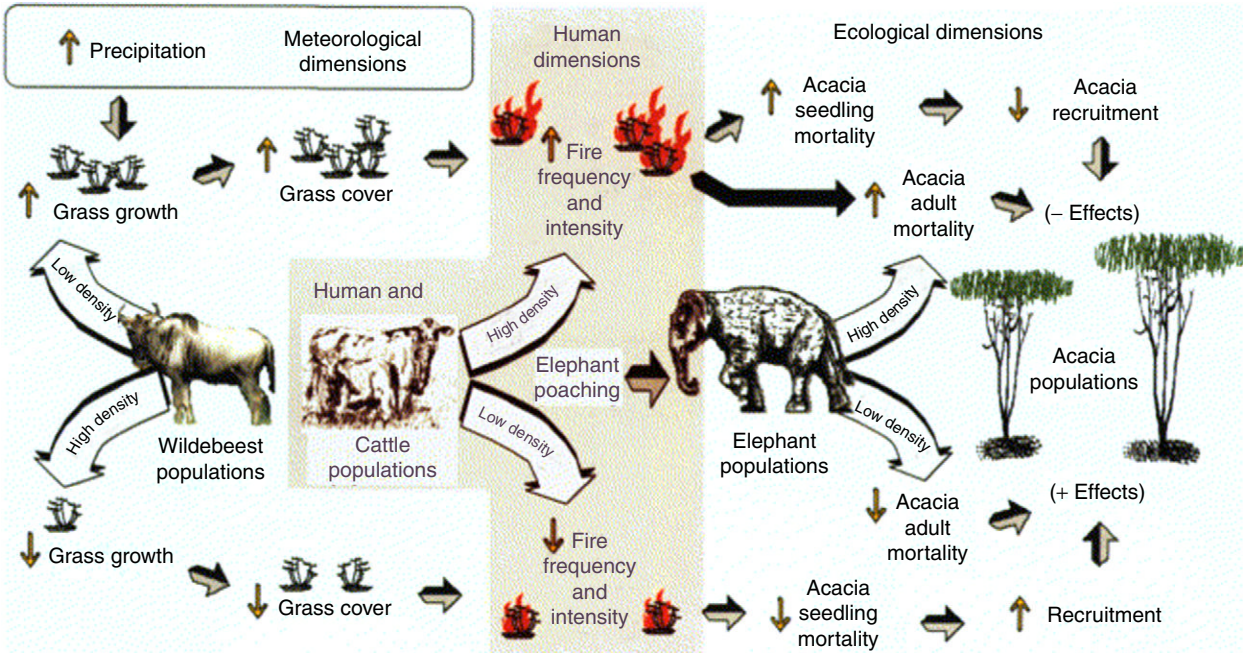


Figure 8 Dynamics of acacia forests in the Serengeti grasslands are a function of complex interactions between meteorological, ecological, and human dimensions. Past failures to recognize the interdependencies of these dimensions has led to controversies and misconceptions about the concept, for example, short- vs. long-term changes.

decreased efficiency of nutrient cycling, and increased nutrient losses from the system. Both native perennial plants (cover and biomass) and many associated microbial and animal populations are reduced, whereas exotic species increase in dominance. All of this results in a decrease in the biological and, hence, economic potential of the land.

''Greening'' of Drylands

In recent years, numerous studies have reported a "greening-up" of drylands over large regions of the globe. These data are akin to those presented in **Figure 3** relating normalized difference vegetation index (NDVI) to precipitation. For example, using data from 1983–2003, Helldén and Tottrup (2008) found a strong greening trend in the Sahel (from west to east), moderate greening in the Mediterranean and East Asia, but insignificant greening in South Africa and South America. They suggest that the very strong greening found in the Sahel is potentially a return to pre-Sahelian-drought rainfall conditions, whereas climatic changes may be responsible for enhanced plant growth in the Mediterranean and East Asia. These trends appear at odds with the UNCCD view that desertification is increasing worldwide in response to climate change, population growth, unsustainable land-use practices, and urbanization. However, due to numerous time and spatial scale issues, technical concerns, and other unknowns, Helldén and Tottrup conclude that "at the moment there is no clear understanding of this dryland '"greening," although it is recognized that an explanation must include both biophysical as well as human factors."

Meteorological Dimensions

Substantial improvements have been made in our understanding of the causes of interannual variability in dryland climates, including the natural causes of droughts. For example, it is accepted that variations in annual precipitation levels are related to natural variations within the global-scale climate systems. However, great uncertainties remain, and the relationship between desertification and climate resembles the proverbial "chicken and egg" problem. To what extent does climate induce desertification? How do changes in land surface properties due to land degradation, in turn, affect climate? Since the two processes often work together, it is virtually impossible to separate the impact of drought from that of desertification.

Drylands are particularly vulnerable to climate variability, of which precipitation is the most important component. For example, a slight shift in seasonal precipitation and/or frequency of extreme rain events could potentially lead to the overexploitation of the meager resources of drylands and contribute to the further degradation of the resource base on which human populations are so dependent. Preliminary studies with general circulation models (GCMs) in the 1980s projected that a doubling of atmospheric carbon dioxide (due to the rapidly expanding human population and associated activities) would result in lower precipitation, as well as shifts in the timing and frequency of rains, in the interior of large continents. Recent GCM studies also predict increases in rainfall intensity and longer dry periods in many dryland regions of the globe. In the long run, global climate change may further exacerbate the already high natural variability of drylands, leading to permanent degradation of their productive potentials, particularly since there is a lack of "buffering" by large reserves of organic matter in the soils or in woody vegetation.

The array of impacts of climate on land and the implications of degraded land surface for the climate system are varied and complex. Human activities impact surface characteristics (and, hence, its albedo) and the atmospheric composition of dryland regions, including the breakdown of soil structure, reductions in soil moisture, increased surface runoff, changes in species composition and cover, increase dust in the atmosphere, and increases in aerosol and trace gas emissions from burning. In response to such human impacts, dryland climate is greatly influenced mainly via changes in energy balance. Changes in albedo affect the amount of solar radiation absorbed by the surface and changes in soil moisture levels affect the portion of energy used in evaporation and transpiration processes. Changes in surface roughness influence wind speeds and turbulence, which have a bearing on evapotranspiration. Finally, changes in atmospheric composition may affect atmospheric temperature profiles, and thus influence capacity to generate precipitation over land.

Human Dimensions

Rapidly changing social and economic conditions – along with the potential for climate change – pose serious challenges to many dryland regions of the world. Globally, there are differences in socioeconomic factors (e.g., human population growth rates) and biogeography (e.g., natural vegetation), which play a large role in the type of major human activities in any given area. Moreover, there are differences in how human interference is affecting biodiversity and ecosystem functioning in poor and richer countries. Taking these differences into account, there are roughly six major classes of human interference in drylands:

1. the loss of habitat;
2. the fragmentation of crucial habitat;
3. overexploitation (mainly overgrazing by domestic animals);
4. the spread of exotic organisms (aliens);
5. air, soil, and water pollution; and
6. climate change.

Although it is somewhat difficult to generalize about the vulnerability of terrestrial ecosystems, some problems virtually exist everywhere (e.g., habitat loss, habitat fragmentation, and the direct and indirect effects of exotic species), whereas pollution and climate change tend to be more significant threats in richer, temperate zone nations. Key ecosystem "goods and services" (e.g., food, construction materials, water purification, flood control, climate regulations, soil maintenance, carbon sequestration, nutrient recycling, wildlife habitat, erosion control, tourism/recreation, etc.) are being seriously affected (**Table 3**; *see* Ecosystem Services).

Given that all three of the developing regions of the world – Africa, Asia, and Latin America – have similar percentages of land degradation, and the highest levels of population growth, desertification is a growing problem with major ecological,

Table 3 Four broad categories of ecosystem services used by the Millennium Ecosystem Assessment (MEA), including a general assessment of their economic value and application to drylands

Service	*Brief explanation*	*Examples*	*Economic value*	*Application to desertification*
Provisioning (=“goods”)	Material items for direct or near-direct human use. Globally, often at the expense of the other types of services, there has been a trend to increase production of most provisioning services, largely due to the expansion and intensification of agriculture.	Crops, forage, livestock, wild-caught fish or animals, freshwater, timber, fuel, fiber.	Provisioning services are relatively easy to quantify and value, since formal or informal markets often exist, and where they do not, substitute values can be estimated.	Desertification is often a direct consequence of overuse of provisioning services, driven by complex external and internal factors such as population pressures and poverty. Examples of declining provisioning services in drylands include wild plant and animal products, fuelwood, genetic resources, clean air, and freshwater.
Regulating	Keep the flow of other ecosystem services within a desired range, and prevent ecosystem “disservices” such as natural disasters from getting out of hand. The MEA suggests that globally 70% of regulating services have been degraded. At a global scale, climate change can be seen as a breakdown of climate regulating services.	Pollination, climate regulation, flood attenuation, air and water quality maintenance, pest regulation.	Drylands are inherently variable so valuing these services is difficult (the value of regulating services is often delivered via other services). Nevertheless, an increase in certainty and a decrease in risk have economic value. The breakdown of regulating services can have dire social and environmental consequences.	Vegetation cover buffers both the hydrological and production systems. The weakening of regulating services can be slow and continuous, or irregular and only apparent during extreme events. For instance, catastrophic floods may be a combination of reduced infiltration, reduced slowing of overland flow due to a combination of soil compacting and the loss of ground cover coupled with an intense storm event.
Cultural	Cultural services make the world a place in which people want to live. The MEA suggests there has been a rapid decline in many cultural services over the past century.	Recreation, esthetic, artistic appreciation, educational, scientific benefits, spiritual inspiration, sense of place.	People of all cultures value nature. Many of these services are easily quantified and valued, because well-established markets exist, such as for ecotourism. Others require more indirect approaches, such as hedonistic pricing, for example, the increase in value of a property when it has a nice view.	Empirical evidence suggests cultural services are readily traded for provisioning services in times of need, or simply because their economic value is not appreciated. For example, some dryland ranchers use excessive stocking rates to maximize short-term profit with the knowledge that it will lead to destructive feedbacks (erosion, loss of plant cover) to the very land they value.
Supporting	Underlying ecosystem processes that allow provisioning, regulating and cultural services to be provided. Often constitute the “slow variables.”	Net primary production, habitat for biodiversity, soil formation and nutrients cycling.	Always delivered via provisioning, regulating and cultural services, and should be valued through those other services to avoid double counting.	In desertification it is typically the reduction of supporting services that ultimately leads to the persistent decrease in the ability of the system to provide provisioning and regulatory services.

Source: Modified with permission from Safriel U, Adeel Z, Niemeijer D, *et al*. (2005) Dryland systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystems and Human Well-Being: Current State and Trends*. The Millennium Ecosystem Assessment Series, pp. 623–662. Washington, DC: Island Press.

economic, and social implications. However, to understand the human dimensions of desertification, it is important to compare and contrast different regions of the world to each other, seeking out generalities, particularly in richer versus poorer countries. As new economic policies are adopted, many pastoral and dryland farmers are marginalized and must move to urban areas; socioeconomic conditions are rapidly changing, for example, the rise of tourism, intensification of high-tech agriculture, and the shifting of populations to urban environments. Often, such changes result in the abandonment of land for traditional agriculture and the rise in demand for water for urban expansion, tourism, and irrigation, resulting in increased land- and resource-use conflicts.

The human dimensions of desertification represent some of the most complicated and perplexing issues in identifying the causes of and solutions to land degradation. For example, in large areas of Africa, civil strife and government policies are key factors influencing resources (land, water, and wildlife), land degradation, and food security. Civil strife displaces people, often to other marginal lands and in great numbers; their abandoned land is left unattended, indigenous management systems are lost, and they are forced to use methods of ranching and farming that are often not appropriate in these new areas. In some countries, government policies have encouraged planned human settlements in the wetter margins of arid and semiarid lands or nearer to water supplies. This has led to land-use conflicts between agriculture, livestock, wildlife, and human settlements, as a result of the intrusion of agriculture into lands traditionally used for domestic stock.

Ecosystem Services

Reduction in the provision of ecosystem services as a result of water scarcity, overintensive use of services, and climate change is a much greater threat in drylands than in nondryland systems (**Table 3**). As noted, a key feature of most dryland areas is that there is a very tight, direct coupling between human livelihoods and use of natural resources (for instance collecting firewood, obtaining water from wells, hunting game, and raising livestock). Hence, the protection of the supporting services is critical in drylands because they underpin all the other services, and degradation is expensive and difficult (sometimes effectively impossible) to reverse (**Table 3**).

In the section Combating Desertification, the author suggests the UNCCD's definition of land degradation and desertification ("the reduction or loss of the biological and economic productivity and complexity of terrestrial ecosystems … resulting from various factors including climatic variations and human activities") to be the most straightforward and unambiguous one available. The Millennium Ecosystem Assessment (2005) expressed this definition in terms of the balance between the demand and supply of ecosystem services that support the well-being of humans. Since human needs change over time, the viability of achieving and maintaining sustainable livelihoods depends on a suite of ecosystem services, rather any single one (see **Table 3**).

This is the objective of sustainable land management, which is defined as "the use of land resources, including soils, water, animals and plants, for the production of goods to meet changing human needs, while simultaneously ensuring the long-term productive potential of these resources and the maintenance of their environmental functions" (Liniger *et al.*, 2011). Nevertheless, not all ecosystem services are equally vulnerable to degradation and have differing importance. Within limits, trade-offs can be made between services of different types without incurring degradation but the nonlinear responses of ecosystem services to human appropriation must be considered as critical thresholds exist and, in some cases, once surpassed trade-offs cannot be safely made. For example, poorly managed cropland or the conversion of rangeland to cropland may affect multiple services (especially provisioning, regulating, and supporting) resulting in increased pressure on the remaining natural resources that have impacts beyond the local scale. Often, the expensive use of external inputs such as irrigation and fertilization may ensure a positive balance of services locally, but in the long term can lead to desertification through processes such as water-logging and salinization of the soil.

The Dryland Development Paradigm

With regard to addressing the causes and consequences of desertification, we can summarize the main points in this article as follows:

1. ecological, meteorological, and human (social–economic) variables interact to form coupled human (H)–environmental (E) systems;
2. change in these coupled H–E systems is omnipresent but is manifested in different ways, for example, slowly evolving (or chronic) and abrupt;
3. H–E systems are not in equilibrium, often exhibiting multiple ecological and social states;
4. cross-scale interactions between humans and the environment are scale-dependent (see **Figure 6**); and
5. local environmental knowledge (LEK) can play an important role in understanding both short- and long-term processes.

Based on the three dimensions of desertification (**Table 1**) and the aforementioned summary points, Reynolds *et al.* (2007a) proposed the Dryland Development Paradigm (hereinafter, the 'DDP'), which is a new synthetic framework for global desertification. The DDP consists of five principles (P1–P5, **Table 4**). This conceptual model has practical implications for an effective approach for understanding H–E systems and formulating models useful for decision-making. A conceptual model is shown in **Figure 9**.

As noted under dryland syndrome (*see* The Dryland Syndrome), there is a close dependency of human livelihoods (H) on the environment (E) in drylands. These dependencies are dynamic, always changing (P1, **Table 4**). The critical linkages between these subsystems are created by (1) human activities and decision-making on the one hand (H→E) and (2) environmental controls and feedbacks pertaining to the flow of ecosystem services on the other (E→H). As a result, ecosystem goods and services of importance to local

Table 4 Five principles of the DDP, for coupled H–E systems, their significance and implications for research, management and policy

Principle	*Why important in drylands*	*Key implications for research, management, and policy*
P1: H–E systems are coupled, dynamic, and coadapting, so that their structure, function, and interrelationships change over time.	The close dependency of most drylands livelihoods on the environment imposes a greater cost if the coupling becomes dysfunctional; variability caused by biophysical factors as well as markets and policy processes, which are generally beyond local control, means that tracking the evolving changes and their functionality is relatively harder and more important in drylands.	ki-1: Understanding dryland desertification and development issues always requires the simultaneous consideration of both human and ecological drivers, and the recognition that there is no static equilibrium "to aim for."
P2: A limited suite of "slow" variables are critical determinants of H–E system dynamics.	Identifying and monitoring the key slow H and E variables is particularly important in drylands because high variability in "fast" variables masks fundamental change indicated by slow variables.	ki-2: A limited suite of critical processes and variables at any scale makes a complex problem tractable.
P3: Thresholds in key slow variables define different states of H–E systems, often with different controlling processes; thresholds may change over time.	Thresholds particularly matter in drylands because the capacity to invest in recovering from the impacts of crossing undesirable thresholds is usually lower per unit (area of land, person, etc.); and, where outside agencies must be called on, the transaction costs of doing so to distant policy centers are usually higher.	ki-3: The costs of intervention rise nonlinearly with increasing land degradation or the degree of socioeconomic dysfunction; yet high variability means great uncertainty in detecting thresholds, implying that managers should invoke the precautionary principle.
P4: Coupled H–E systems are hierarchical, nested, and networked across multiple scales.	Drylands are often more distant from economic and policy centers, with weak linkages; additionally, regions with sparse populations may have qualitatively different hierarchical relationships between levels.	ki-4: H–E systems must be managed at the appropriate scale; cross-scale linkages are important in this, but are often remote and weak in drylands, requiring special institutional attention.
P5: The maintenance of a body of up-to-date LEK is key to functional coadaptation of H–E systems.	Support for LEK is critical in drylands because experiential learning is slower where monitoring feedback is harder to obtain (owing to more variable systems, larger management units, in sparsely populated areas) and, secondarily, where there is relatively less research.	ki-5: The development of appropriate hybrid scientific and LEK must be accelerated both for local management and regional policy.

Source: Reproduced from Reynolds JF, Stafford Smith DM, Lambin EF, *et al.* (2007a) Global desertification: Building a science for dryland development. *Science* 316: 847–851.

populations are changing and evolving over time. This close dependency of livelihoods on the environment may impose a high cost if the H–E linkages become unbalanced or dysfunctional (P1). Often this is due to variability caused by biophysical factors as well as markets and policy processes, which are generally beyond local control, means that tracking any changes that are occurring (which is always happening) and their overall importance to the stability of the H–E balance is crucial but very difficult.

The DDP emphasizes the importance of "slow" or controlling variables in both H and E domains (P2). For example, terrestrial primary production is dependent on both soil moisture and nutrients; however, nutrients in the soil change very slowly over time (hence, is a "slow" variable), whereas the residence time of soil water changes very rapidly (a "fast" variable). "Slow" variables are important, because they actually control changes of state, whereas "fast" variables usually reflect unimportant variability within states. For example, at any given time, the degree of soil fertility would be more important to a management decision than soil water content (the latter of which will change within hours).

As discussed, thresholds matter in drylands. However, the DDP posits that only thresholds of interest should be the "slow" variables. If a threshold is crossed, it can lead to major imbalances in the entire system. Furthermore, the capacity of H–E systems to recover from the impacts of crossing undesirable thresholds is often slow and if distant policy centers must be called on for assistance, the transaction costs of doing so are usually high (P3, **Table 4**).

The principles of the DDP introduced thus far – that is, "slow" and "fast" variables, thresholds and constant changes in the interactions, and thus balance of H–E interactions – all depend on the scale of interest (P4, **Table 4**), of which there are numerous ones. For example, the slow variables of interest regarding the welfare of a single household, an ejido (in Latin America, this is communal land shared by all people of a

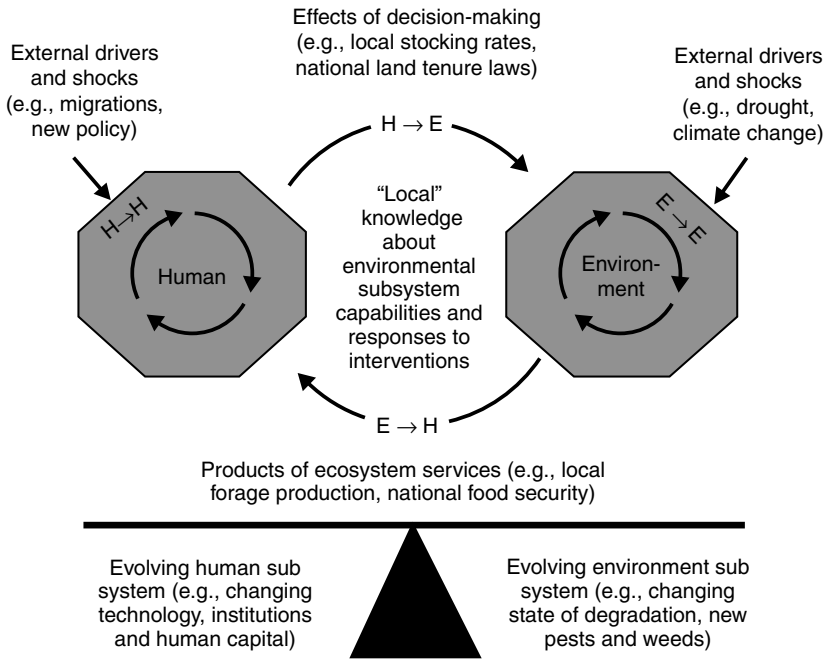


Figure 9 Schematic depicting interactions between human (H) and environmental (E) components of the land system. P1–P5 refer to DDP principles in **Table 2**. These systems coevolve as balanced networks of feedbacks and interactions between H–E components in spite of constantly changing external drivers. Decision-making (H→E) and ecosystem services (E→H, e.g., grazing land and clean air) are key linkages between components (moderated by an effective system of LSK, P5). Many studies have tended to focus on each subsystem (i.e., H→H and E→E) at the expense of the linkages between them. Reproduced from Verstraete MM, Scholes RJ, and Stafford Smith M (2009) Climate and desertification: Looking at an old problem through new lenses. *Frontiers in Ecology and the Environment* 7: 421–428, with permission from ESA.

community), or a nation may be quite different. In managing H–E systems, such cross-scale linkages are important in decision-making, but in drylands they are often remote and weak, requiring special institutional attention and monitoring support (*see* The Dryland Syndrome).

Lastly, the loop in **Figure 9** implies that links between the H and E subsystems are mediated in practice by human mental models and local and scientific knowledge (LSK) (P5, **Table 4**). Differing management and policy decisions are often dependent on LSK. "Local knowledge" *per se* is not always very local: It may be the knowledge of distant policymakers about how their institutions operate as much as local farmers' understanding of the effects of different levels of grazing.

Conclusions

Desertification (land degradation in drylands) is the loss of biological and economic productivity, and biodiversity in arid and semiarid croplands, pastures, rangelands, and subhumid woodlands of the world. It is due mainly to nonsustainable human activities, such as overcultivation, overgrazing, deforestation, and poor irrigation practices and is often triggered or exacerbated by climate variability, mainly drought. A new integrated assessment paradigm, the DDP, has been proposed as a tool to aid in tackling the complexity of this phenomenon. The DDP, which is based on the simultaneous roles and complex feedbacks among the meteorological, ecological, and human dimensions of this problem, provides a roadmap (not a recipe) for thinking about drylands management issues. Historically, the failure to recognize and include the interdependencies of these three dimensions has slowed progress by the UNCCD in tackling the enormous problem of dryland degradation, although recent developments are encouraging. International researchers – representing a wide range of disciplines such as ecology, atmospheric science, social sciences, policy, and integrated assessment – must work together to successfully address this pressing global problem.

References

Environment Management Group (EMG) (2011) *Global Drylands: A UN System-Wide Response*. Nairobi, Kenya: United Nations.

Geist HJ and Lambin EF (2004) Dynamic causal patterns of desertification. *Bioscience* 54: 817–829.

Helldén U and Tottrup C (2008) Regional desertification: A global synthesis. *Global and Planetary Change* 64: 169–176.

IPCC (2007) Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis*. Cambridge, UK and New York, USA: Cambridge University Press.

Liniger HP, Studer RM, Hauert C, and Gurtner M (2011) *Sustainable Land Management in Practice – Guidelines and Best Practices for Sub-Saharan Africa*. Rome, Italy: TerrAfrica, World Overview of Conservation Approaches and Technologies (WOCAT) and Food and Agriculture Organization (FAO) of the United Nations.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-being: Desertification Synthesis*. Washington, DC: World Resources Institute.

Morell V (1997) Counting creatures of the Serengeti, great and small. *Science* 278: 2058–2060.

Mortimore M, Anderson S, Cotula L, *et al.* (2009) *Dryland Opportunities: A New Paradigm for People, Ecosystems and Development,* 86 pp. Gland, Switzerland: IUCN; London, UK: IIED and Nairobi, Kenya: UNDP/DDC.

Reynolds J, Maestre F, Kemp P, Stafford Smith DM, and Lambin E (2007b) Natural and human dimensions of land degradation: Causes and consequences. In: Canadell JG, Pataki DE, and Pitelka LF (eds.) *Terrestrial Ecosystems in a Changing World (Global Change – The IGBP Series)*, pp. 247–255. Berlin, Heidelberg: Springer.

Reynolds JF and Stafford Smith DM (eds) (2002) Global desertification: Do humans cause deserts? *Dahlem Workshop Report 88*. Berlin: Dahlem University Press.

Reynolds JF, Stafford Smith DM, Lambin EF, *et al.* (2007a) Global desertification: Building a science for dryland development. *Science* 316: 847–851.

Safriel U, Adeel Z, Niemeijer D, *et al.* (2005) Dryland systems. In: Hassan R, Scholes R, and Ash N (eds.) *Ecosystems and Human Well-Being: Current State and Trends*. The Millennium Ecosystem Assessment Series, pp. 623–662. Washington, DC: Island Press.

Stafford Smith DM, Abel N, Walker B, and Chapin III FS (2009) Drylands: Coping with uncertainty, thresholds, and changes in state. In: Chapin, III FS, Kofinas GP, and Folke C (eds.) *Principles of Ecosystem Stewardship: Resilience-Based Natural Resource Management in a Changing World*, pp. 171–195. New York, NY: Springer.

United Nations Convention to Combat Desertification (UNCCD) (1994) United Nations Convention to Combat Desertification (Intergovernmental Negotiating Committee For a Convention to Combat Desertification), Elaboration of an International Convention to Combat Desertification in Countries Experiencing Serious Drought and/or Desertification, Particularly in Africa. *UN Doc.* A/AC.241/27, 33 I.L.M. 1328. United Nations, New York.

United Nations Environment Programme (UNEP) (1991) United Nations Environment Programme. *Status of Desertification and Implementation of the United Nations Plan of Action to Combat Desertification*. Nairobi, Kenya.

Verstraete MM, Scholes RJ, and Stafford Smith M (2009) Climate and desertification: Looking at an old problem through new lenses. *Frontiers in Ecology and the Environment* 7: 421–428.

World Meteorological Organization, *United Nations Environmental Programme*. London: Arnold Publishers.

Development by Design: Using a Revisionist History to Guide a Sustainable Future

Joseph M Kiesecker, Kei Sochi, and Michael Heiner, The Nature Conservancy, Colorado, CO, USA
Bruce McKenney, The Nature Conservancy, Virginia, VA, USA
Jeffrey Evans and Holly Copeland, The Nature Conservancy, Wyoming, WY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Avoidance Measures taken to prevent impacts from occurring in the first place—for instance, by changing or adjusting the development project's location or the scope, nature, and timing of its activities.

Biodiversity offsets Measurable conservation outcomes resulting from actions designed to compensate for significant residual adverse biodiversity impacts arising from project development after appropriate prevention and mitigation measures have been taken. The goal is to achieve no net loss and preferably a net gain of biodiversity on the ground with respect to species composition, habitat structure, and ecosystem function.

Environmental impact assessment (EIA) A formalized process, including public consultation, in which all relevant environmental consequences of a project are identified and assessed before authorization is given. The process of identifying, predicting, evaluating and mitigating the biophysical, social, and other relevant effects of development proposals prior to major decisions being taken and commitments made.

Footprint The area of land or water covered or affected by a project. This can include the direct physical coverage (i.e., the area on which the project physically stands) and the area indirectly affected by the project (i.e., the area affected by disturbances such as noise that directly emanate from the project).

Minimization Minimizing impacts by limiting the degree or magnitude of the development action and its implementation in a manner consistent with conservation goals.

Mitigation Measures that aim to reduce impacts to the point where they have no adverse effects.

Mitigation hierarchy Hierarchy that may include avoidance, minimization, restoration, or offset. In *avoidance*, to completely avoid impacts on certain components of biodiversity, measures are taken to avoid creating impacts from the outset, such as careful spatial or temporal placement of elements of infrastructure. In *minimization*, as far as is practically feasible, measures are taken to reduce the duration, intensity, and/or extent of impacts that cannot be completely avoided. In *restoration*, measures are taken to rehabilitate degraded ecosystems or restore cleared ecosystems after exposure to impacts that cannot be completely avoided and/or minimized. In *offset*, achieve no net loss or a net gain of biodiversity, measures are taken to compensate for any residual significant adverse impacts that cannot be avoided, minimized, and/or rehabilitated or restored. Offsets can take the form of positive management interventions such as restoration of degraded habitat, arrested degradation or averted risk, or protecting areas where there is imminent or projected loss of biodiversity.

No net loss Target for a development project in which the impacts on biodiversity caused by the project are balanced or outweighed by measures taken to avoid and minimize the project's impacts, to undertake on-site restoration, and finally to offset the residual impacts so that no loss remains. Where the gain exceeds the loss, the term *net gain* may be used instead of *no net loss*. No net loss (or net gain) of biodiversity is a policy goal in several countries, and is also the goal of voluntary biodiversity offsets.

Offset activities The set of activities identified to achieve no net loss or a net gain of biodiversity in the specific context of the development project concerned. They can involve a mixture of activities, including averting risk and undertaking positive management interventions. In *averting risk*, areas of biodiversity in which there is imminent or projected loss of that biodiversity are protected; agreements such as contracts or covenants with individuals are struck in which they forgo the right to convert habitat in the future in return for payment or other benefits received now. In *undertaking positive management interventions* to restore an area or stop degradation, the conservation status of an area of land is improved by restoring habitats or ecosystems and reintroducing native species. Where proven methods exist for successful reconstruction or creation of ecosystems these may be undertaken. In other instances, a project might reduce or remove current threats or pressures by, for instance, introducing alternative sustainable livelihoods or substitute materials.

Residual impact The remaining adverse impact on biodiversity after appropriate avoidance, minimization, and restoration actions have been taken according to the mitigation hierarchy.

Chapter Outline

- The world is changing more rapidly each day. Fast-paced development threatens some of our most cherished lands, waterways, and wildlife.
- Over the next two decades, there will be even more pressure to convert and develop our lands and waters to provide food, energy, transportation, and housing for a growing and hungry world.
- Achieving conservation goals will require a more integrated vision of development and conservation.
- Here we outline an innovative science-based approach called Development by Design (DbD) to help guide more sustainable development.
- DbD brings together conservation and mitigation planning to help reduce conflicts between development and conservation goals, improve the effectiveness of mitigation (to avoid, minimize, restore, and offset impacts), and deliver more sustainable outcomes for the development of natural resources.
- DbD seeks to achieve a net gain for both development and environment.

Introduction

The world is on the cusp of a sea change. World population is projected to grow to more than 9 billion by 2050 with a rapid pace of development in order to meet the growing demands for food, water, energy, minerals, and other resources. At the same time, most of the world's people don't care about biodiversity, and most don't think of nature conservation as essential to meet their basic needs. People want food and shelter, safety and security, and an opportunity to improve their livelihoods. And they see the use of natural resources as a way to further those objectives. Achieving conservation goals will require a more integrated vision of development and conservation. The conservation community needs to recognize that not everyone wants to set aside the world into a series of protected areas or parks for the sole purpose of wildlife conservation. The importance of economic development for improving human well-being means that society will likely choose to value development over conservation if we continue to present it as an either–or option.

In this chapter, we put forward a four-step framework called DbD that supports development *and* conservation goals by harnessing advances in conservation and mitigation planning. Then to illustrate how the DbD framework can advance truly sustainable outcomes, we apply it retroactively to the Wyoming Basins Ecoregion (WBE), a landscape rich in both energy resources and conservation values. We examine the additional conservation activities that would have been implemented if the framework had been applied, and discuss how taking these steps in the past could have reduced the need for current Endangered Species Act restrictions on development, putting the region in a better position to meet today's and perhaps tomorrow's energy development and conservation needs.

Over the next two decades, energy and mining companies will invest unprecedented sums—far more than $20 trillion—in projects around the world from Madagascar to Mongolia. Rapidly developing countries like China and Brazil are making trillion dollar investments in infrastructure. In most cases, the environmental mitigation process for addressing impacts of development projects is ad hoc, opaque, and insufficient, failing to deliver effective outcomes for biodiversity conservation (Environmental Defense, 1999; National Research Council, 2001; Environmental Law Institute, 2002; Federal Interagency Mitigation Workgroup, 2002; Brown, 2006; Gibbons, and Lindenmayer, 2007; Thorne *et al.*, 2009; McKenney and Kiesecker, 2010). Mitigation offers the opportunity to minimize the impacts of development through application of the mitigation hierarchy: avoid, minimize, restore, and offset (Council on Environmental Quality, 2000). Mitigation also represents the best global opportunity for mobilizing billions of dollars for conservation. When compensatory mitigation (also known as *biodiversity offsets*) is a normal part of project costs, the level of funding available for conservation dwarfs other funding sources by many multiples. Consider US wetland mitigation, where no net loss goals and the use of wetland offsets and banks generates $3 billion in funding annually (Environmental Law Institute, 2007). Compare this figure with federal appropriations of about $300 million annually for the US Land and Water Conservation Fund or $2.5 billion annually for the entire US National Park System. Likewise, in Wyoming $36 million was established as a mitigation fund for a single gas field and $24.5 million for another. Compare these figures with the $4 million available annually for wildlife conservation from the Wyoming Wildlife and Natural Resource Trust (Kiesecker *et al.*, 2009).

But in the US and around the world, there are many problems with how mitigation is applied. Traditionally, mitigation has been carried out on a project-by-project basis; specific measures are implemented to mitigate project impacts at a site, usually on or adjacent to the impact site. But with conservation goals defined at a landscape level, applying the mitigation hierarchy on a project-by-project basis, often at small spatial extents, underestimates the cumulative impacts of multiple current or future development projects and undermines the hierarchy's purpose and utility (**Figure 1**). In our opinion, the way mitigation is currently applied does not capture cumulative impacts associated with development. It does not provide a structured decision-making framework to determine when projects can proceed or should be avoided, and it does not harness the full potential of offsets (conservation actions applied away from the development site). There is a strong need to improve the mitigation process with better science and more transparency to address key challenges, including how best to ensure conformance with the mitigation hierarchy, identify the most environmentally preferable offsets within a landscape context, and determine appropriate mitigation replacement ratios. For these reasons, we think an overhaul of existing mitigation practices is in order to more effectively address expected development in the coming years. This is the ideal time for this kind of overhaul. Around the world, there are now 45 compensatory mitigation programs for biodiversity impacts, with another 27 programs in development (Madsen *et al.*, 2010). Major financial institutions are requiring infrastructure projects to meet performance standards that include the full sweep of mitigation, from avoiding

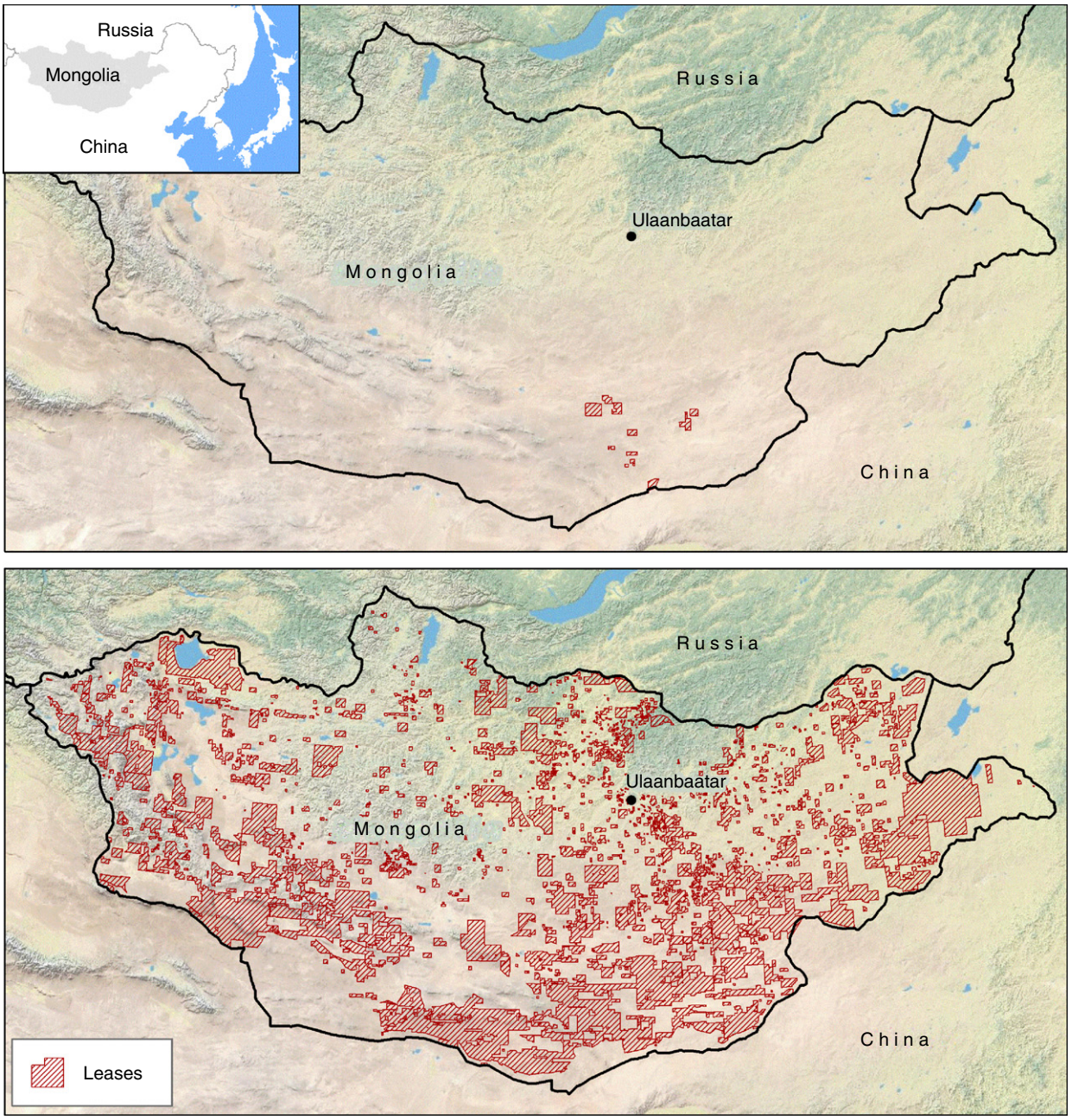


Figure 1 Map of mineral leases in Mongolia. Top panel shows leases of a single company currently moving from exploration to development. As it develops its leases, it will negotiate with the Mongolian government for infrastructure needs as well as actions necessary for mitigation. This company has a corporate policy of net gain for biodiversity associated with its development activities. Bottom panel shows all leases currently under either exploration or development. Attempts by any single leaseholder to develop necessary infrastructure or achieve positive mitigation outcomes will be difficult without consideration of cumulative impacts. Inset map shows the location of the study area.

impacts to compensatory actions (International Finance Corporation, 2012). And leading companies increasingly see this as a normal cost of doing business (Rio Tinto, 2004).

The Framework for Development by Design

The four-step DbD framework has been developed to address key deficiencies in mitigation and to mesh with existing regulatory approaches. It supports a more comprehensive approach to mitigation planning applied at two distinct spatial scales. First, it can be applied at a *landscape, watershed, or seascape level* to determine conservation priorities in the context of multiple objectives, to assess potential cumulative impacts, to identify possible conflicts between objectives, to assess the compatibility (or incompatibility) of impacts with conservation goals, and to evaluate opportunities for sustainable outcomes (steps 1 and 2). Second, it can be applied at a

project or site level to assess site-level impacts, to design a mitigation strategy involving offsets to address impacts, and to measure progress toward fully compensating for impacts (steps 3 and 4).

The key components of DbD are as follow.

1. *Setting landscape conservation priorities*: Where are the highest priority areas for conservation, ecosystem services, and other values in the region?
2. *Making mitigation decisions consistent with landscape planning*: How will projected development affect the region? Given expected impacts, what are the best options to support sustainability?
3. *Determine residual impacts and select optimal offset portfolio*: Where impacts occur, how can compensating mitigation (i.e., offsets) best deliver values ecologically and functionally equivalent to those lost, be located at an acceptable proximity from the impact site, and contribute to landscape conservation goals?
4. *Estimating offset contribution and progress toward goals*: To what extent will mitigation actions support sustainable outcomes? To what extent will offsets compensate for impacts to achieve a net gain or no net loss goal? Which offsets will provide the best conservation return on investment—highest conservation value at least cost and risk?

Step 1: Setting Landscape Conservation Priorities

Landscape-level conservation planning has a rich history (Pressey *et al.*, 1997; Ball and Possingham, 2000; Pressey and Bottrill, 2008), and our goal here is not an exhaustive review on the methods behind or importance of conservation planning but rather to illustrate how it can be utilized to direct mitigation. Conservation planning, particularly ecoregional assessments (e.g., Groves, 2003), provides the structure to ensure that mitigation is consistent with conservation goals, which often include the maintenance of large and resilient ecosystems to support both healthy wildlife habitats and human communities. Blending the mitigation hierarchy with landscape planning offers distinct advantages over the traditional project-by-project approach because it considers the cumulative impacts of both current and projected development, provides regional context to better guide which step of the mitigation hierarchy should be applied (i.e., avoidance vs. offsets), and offers increased flexibility for choosing offsets that can maximize conservation return by providing funding for the most threatened ecosystems or species.

Landscape-level conservation planning is the process of locating and configuring areas that can be managed to maintain viability of biodiversity and other natural features (Pressey and Bottrill, 2008). The resulting conservation plan is intended to clearly articulate a vision that incorporates the full range of ecosystem services and biological features, their distribution, and the minimum needs of each to persist in the long term (e.g., Tear *et al.*, 2005). Creation of the vision depends on the active involvement of host governments, experts of many disciplines, development organizations, and citizens in the region. A conservation portfolio of priority sites, the end product of conservation planning, is a selected set of areas that represents the full distribution and diversity of the biological features the plan attempts to conserve (e.g., Noss *et al.*, 2002). Often plans use an optimization approach automated with spatial analysis tools to meet the minimum viability needs of each biological target yet minimize the amount of area selected (Pressey *et al.*, 1997; Ball and Possingham, 2000). Optimization tools provide flexibility for reconfiguring conservation plans in the context of future development scenarios (Kiesecker *et al.*, 2010; Saenz *et al.*, 2010; Kiesecker *et al.*, 2011a; Heiner *et al.*, 2011) To date, conservation plans (i.e., ecoregional assessments) have been conducted in many places around the world, providing opportunities to harness these data to implement DbD (Dinerstein *et al.*, 2000; Groves, 2003). In places where conservation plans have yet to be completed, rapid assessments can be conducted in a manner that allows them to address the mitigation of future development in a time frame that is congruent with regulatory decision making (see Saenz *et al.*, 2010, and Heiner *et al.*, 2011, for examples).

Step 2: Making Mitigation Decisions Consistent with Landscape Planning

Development by Design entails analysis of conservation plans in the context of future development (i.e., Copeland *et al.*, 2009). Proposed projects, along with projected development activities (e.g., oil and gas, wind, solar, residential, agricultural conversion, and some types of mining development), can be mapped and assessed relative to existing conservation priorities or incorporated in the development of a new conservation plan (Kiesecker *et al.*, 2010; Saenz *et al.*, 2010; Heiner *et al.*, 2011). Whenever possible, development patterns should be designed within a framework that takes into account conservation goals for healthy ecosystems (see, e.g., Kiesecker *et al.*, 2011b). Under the DbD approach, overlap or conflicts between conservation priorities and development impacts can be addressed by a redrawing of the conservation portfolio to recapture habitat elsewhere in the study area to meet minimum viability needs of target species and ecological systems. However, if conservation goals cannot be met through redrawing of the conservation priorities, then development impacts would need to be minimized to the degree that biodiversity values are maintained or impacts must be avoided (Kiesecker *et al.*, 2010; **Figure 2**).

This approach offers important advantages over current approaches to mitigation. It provides an opportunity to avoid conflict between potential development and areas critical for biodiversity and ecosystem services, and it also provides the structure to guide decisions about which step in the mitigation hierarchy should be applied in response to proposed development (**Figure 2**). Whereas mitigation is often fraught with subjective and opaque decision making, applying science-based conservation planning through this framework provides a more transparent and credible basis for informing key issues: how compatible impacts are with conservation goals and what the opportunities are for moving beyond zero-sum development versus conservation conflicts. More broadly, by integrating conservation and development planning at a

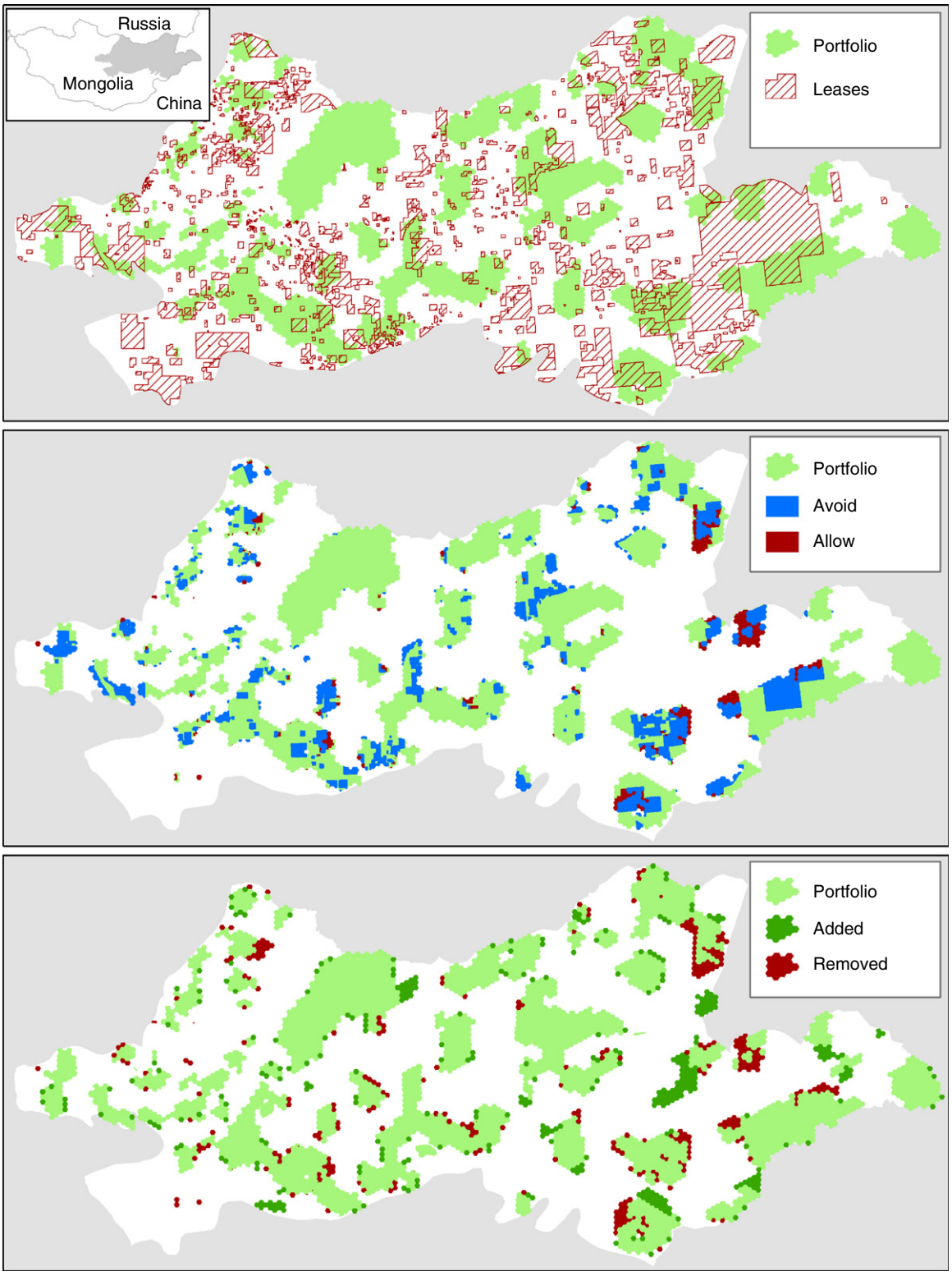


Figure 2 Landscape-level Development by Design approach in Mongolia. Top panel shows initial conservation portfolio and existing mineral and petroleum leases. Middle panel shows areas of potential conflict with mineral and petroleum development and mitigation recommendation (avoid or offset) for conflict areas. Bottom panel shows conservation portfolio redesigned to minimize conflict with mineral and petroleum development. Inset map shows the location of the study area within Mongolia. For details, see Heiner M, Davaa G, Kiesecker J, *et al.* (2011) *Identifying Conservation Priorities in the Face of Future Development: Applying Development by Design in the Grasslands of Mongolia.* Arlington, Virginia: The Nature Conservancy.

landscape scale, it supports application of mitigation at a more appropriate ecological scale. Projecting cumulative impacts at this scale moves mitigation beyond a piecemeal project-by-project approach to one that can support a dynamic vision for a sustainable landscape.

Step 3: Determine Residual Impacts and Select Optimal Offset Portfolio

Under the DbD framework, proposed development consistent with conservation goals of a landscape-level conservation plan would move forward, using best management practices to minimize impacts and harnessing offsets to mitigate the unavoidable impacts associated with the development. Offsets are increasingly used to achieve environmental benefits (Gibbons and Lindenmayer, 2007), providing a mechanism for maintaining or enhancing environmental values in situations where development is sought despite detrimental environmental impacts (McKenney and Kiesecker, 2010; ten Kate *et al.*, 2004). Offsets address residual environmental impacts of development after efforts have been undertaken to avoid, minimize, and restore them on site, with the overall aim of achieving a net neutral or positive outcome. Offsets offer substantial potential benefits for industry, government, and conservation groups alike. Benefits for industry include a higher likelihood that clearance is granted from regulators for new developments, greater societal support for development projects, and the opportunity to more effectively manage environmental risks. Offsets provide government regulators with the opportunity to encourage companies to make significant contributions to conservation and sustainability goals, particularly in situations where legislation does not require mandatory offsets. Conservation organizations can use offsets to move beyond piecemeal mitigation and thereby secure larger-scale, more effective conservation projects. Offsets can also be a mechanism ensuring that regional conservation goals are integrated into government and business planning.

Our objective with DbD is to ensure that offsets are ecologically equivalent to impacts, are consistent with goals of landscape conservation planning, and will persist at least as long as on-site impacts, resulting in net neutral or positive outcomes. The geographic location of offset sites can be accomplished as a separate stand-alone process (**Figure 3**) or as part of a landscape conservation plan (**Figure 4**). Comprehensive mitigation planning is in its infancy, and mitigation will probably occur on a project basis for some time (but see Thorne *et al.*, 2009). Whether conducted as a separate process or incorporated into landscape-level planning, effective offset design must capture impacts to representative biological targets and select offset sites in a manner that captures ecological equivalency and accounts for the landscape value and condition of potential sites (Ball, 2000; Ball and Possingham, 2000; Possingham *et al.*, 2000; Arponen *et al.*, 2007).

Step 4: Estimating Offset Contribution and Progress Toward Goals

Offset policies generally seek no-net-loss or net-gain outcomes for conservation, compensating for impacts with offsets (McKenney and Kiesecker, 2010). But how do we know when enough offset projects have been implemented to achieve no net loss? Under existing policies, offset benefits are often estimated using mitigation replacement ratios that establish the number of credit units that must be debited from an offset to compensate or replace one unit of loss at the project site. Policy guidance on replacement ratios tends to fall into three categories: predefined ratios, such as those based on the type of conservation action (e.g., 1 : 1 ratio for restoration, 5 : 1 for preservation); specified assessment methods that provide a specific approach for determining ratios; and subjective determinations based on the discretion of regulatory authorities after multiple considerations such as proposed conservation actions and risk factors (McKenney and Kiesecker, 2010). Each approach has weaknesses (Kusler, 2003). Although these approaches may simplify the offset accounting and implementation process, there is little reason to believe they deliver no-net-loss outcomes on a regular basis (King and Price, 2004). In sum, current accounting approaches are generally inadequate for ensuring no-net-loss outcomes. They are either too inflexible to address the ecological context for impacts and offsets, as is the case with predefined ratios and specified assessment methods, or too open to discretion and subjective judgment.

Under DbD, a portfolio of possible offsets is identified that represent the best opportunities for protection or restoration. The portfolio selection process ensures that, in comparison to the impact site, potential offsets provide equivalent or better ecological quality as defined by condition, function, connectivity, and contribution to landscape-level conservation. Such considerations are addressed in the rule-setting process for portfolio selection. After selection of an offset portfolio, the next question is, which offsets in that portfolio should be selected for implementation? The aim is to identify offsets that will deliver the greatest contribution toward achieving no net loss at the lowest cost and risk. Our accounting approach seeks to strike a balance by using ratio determinations based on a more structured and transparent approach. Among several considerations for measuring progress, three important factors to take into account include the following.

1. *Additionality*: an option's new contribution to conservation, which is additional to existing values. Offsets that preserve habitat also deliver conservation value when taking into account real-world conditions and threats; those offsets protect against an expected background rate of loss. For example, protecting a 1000 ha forest area that was experiencing an average deforestation rate of 1% per year delivers a new contribution to conservation of 10 ha per year (1% of 1000 ha).
2. *Probability of success*: the likelihood that an option will deliver expected conservation benefits. Success of restoration projects can vary greatly, depending on the ecosystem, restoration techniques, and other factors. In some cases, restoration approaches are known to be effective, but in other situations there may be great uncertainty due to a lack of experience.
3. *Time lag to conservation maturity*: how long it will take for an offset to deliver conservation value at a level similar to what was lost due to impacts. If an offset takes many years

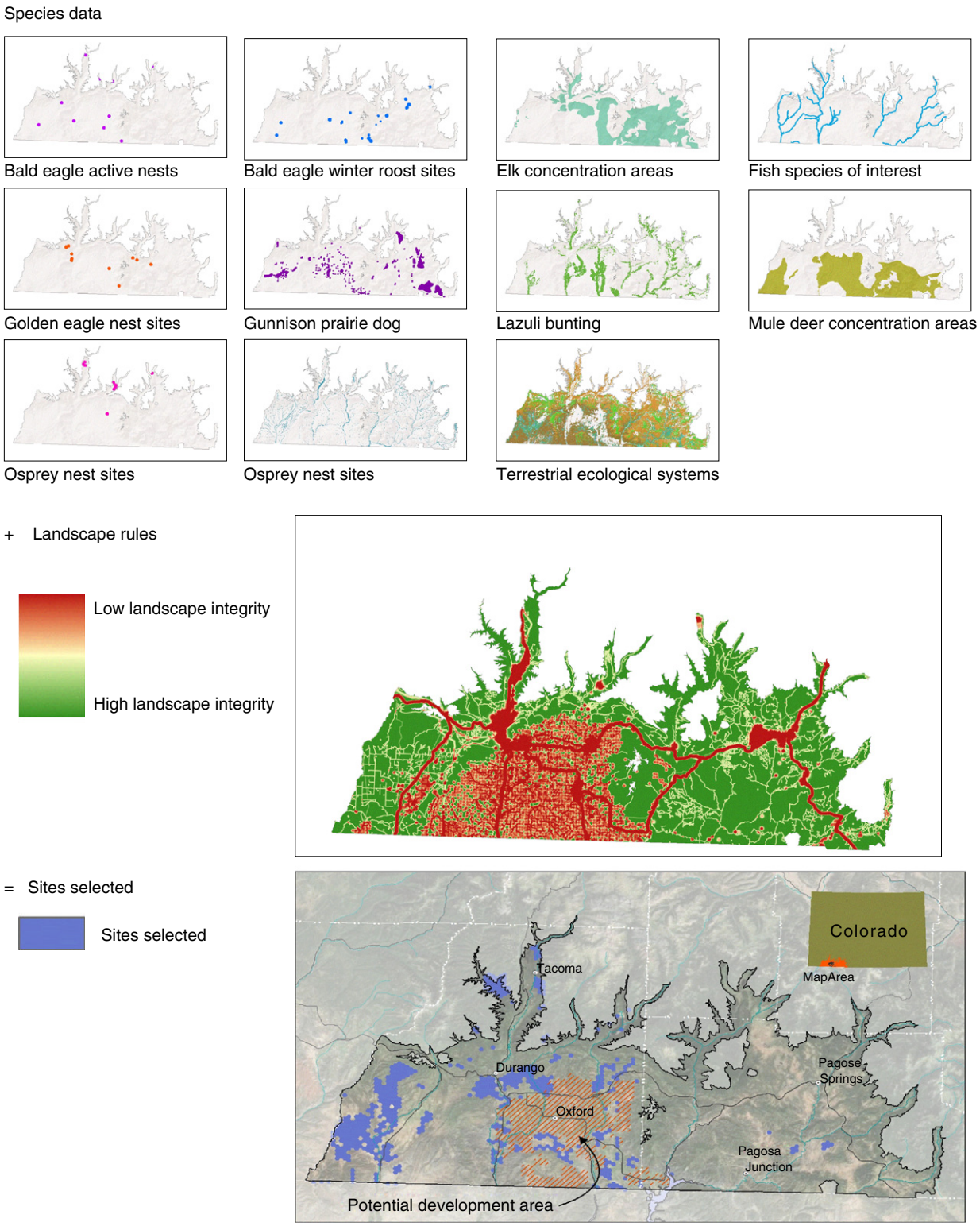
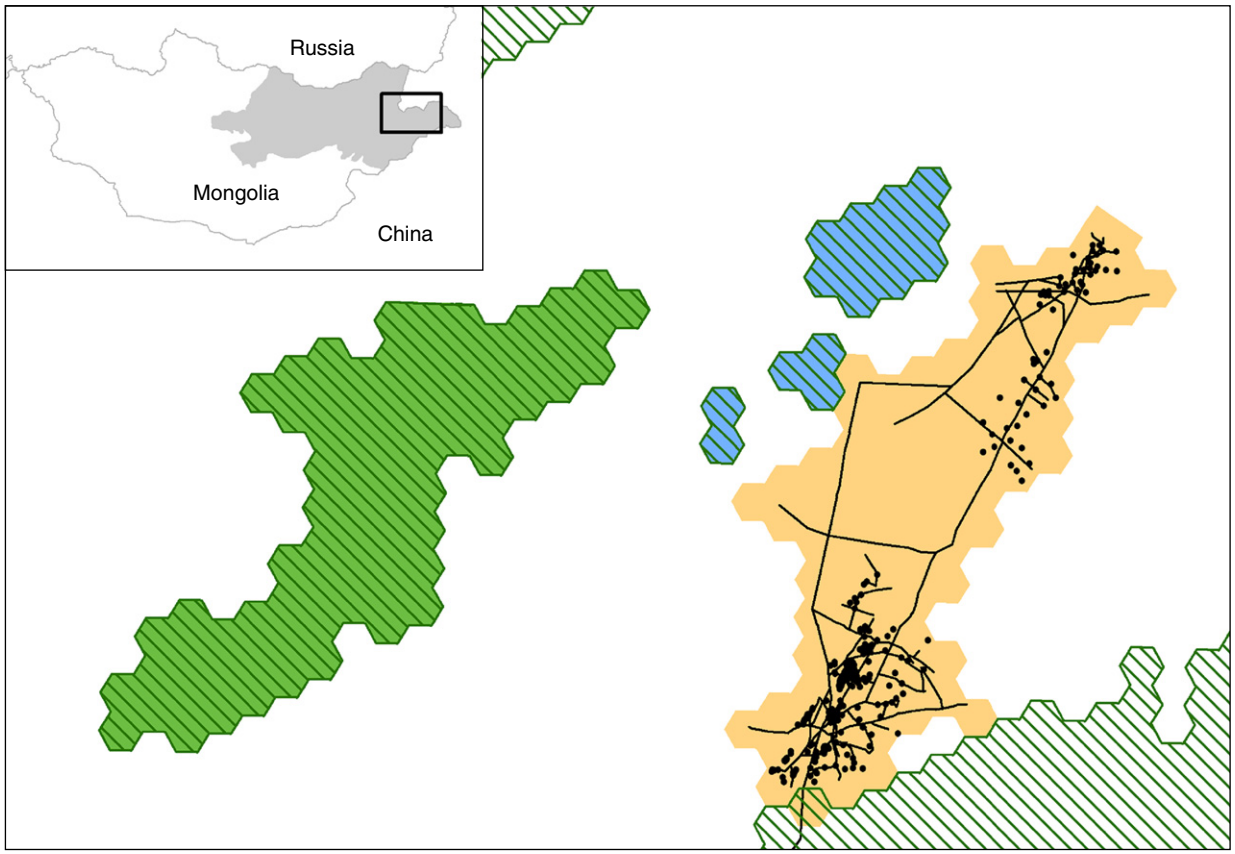


Figure 3 Use of the Marxan algorithm to select suitable offsets sites for oil and gas development in the San Juan Basin of Colorado. Spatial data layers were utilized for both assessing impacts as a result of development on the field and selecting suitable offset sites. Landscape rules: Intactness guided the selection of sites to areas of high habitat quality and low oil and gas development potential. Areas hatched represent the best fit of the Marxan algorithm based on these specific targets and specified rules. Inset map shows the location of the development area within the US state of Colorado. For details see Sochi K and Kiesecker JM (2011) *Energy by Design: Cooperative Mitigation Planning for the San Juan Basin. Appendix B: Selecting Offset Opportunities*. San Juan Basin Wildlife Mitigation Plan, Colorado Division of Wildlife, Denver, Colorado.



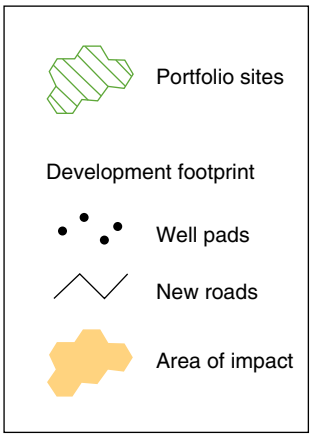
Composition of potential offset sites

Ecosystem type	*Area (km^2)*	
Small water bodies	383	0.3%
Dry steppe | low elev. flat	87,542	72.9%
Dry steppe | low elev. hills	13,212	11.0%
Dry steppe | valley bottom	4183	3.5%
Wet salty depressions	14,270	11.9%
Meadow steppe | low elev. flat	353	0.3%
Meadow steppe | low elev. hills	184	0.2%
Total	120,127	100.0%

Development footprint

Ecosystem type	*Area (km^2)*	
Small water bodies	810	0.1%
Dry steppe | low elev. flat	526,538	74.5%
Dry steppe | low elev. hills	83,689	11.8%
Dry steppe | valley bottom	27,142	3.8%
Wet salty depressions	68,196	9.7%
Total	706,375	100.0%

Ecosystem type	*Area (km^2)*	
Small water bodies	2358	0.4%
Dry steppe | low elev. flat	306,096	50.2%
Dry steppe | low elev. hills	182,521	29.9%
Dry steppe | valley bottom	23,795	3.9%
Wet salty depressions	19,267	3.2%
Meadow steppe | low elev. flat	6390	1.0%
Mod. dry steppe | low elev. flat	34,669	5.7%
Mod. dry steppe | low elev. hills	31,464	5.2%
Total	609,907	100.0%



before conservation benefits mature, the time lag represents a loss for biodiversity and should be accounted for in estimates of offset benefits. We propose accounting for this loss by estimating the time to maturity of a restoration action and applying a discount rate, a commonly used method for estimating the present value of future values. With an understanding of the conservation value of different mitigation actions, it is possible to compare these values to the cost of implementing the actions to estimate which actions will provide the highest conservation returns on investment.

We recognize that developing precise estimates for the accounting framework's key factors—additionality, probability of success, and time to conservation maturity—will be challenging in some contexts. But even approximate estimates that have been developed using science-based methods and presented in a transparent process would be a marked improvement over current practices for estimating compensatory mitigation. We firmly believe that offset accounting must incorporate these factors if offsets are to truly deliver on no-net-loss goals.

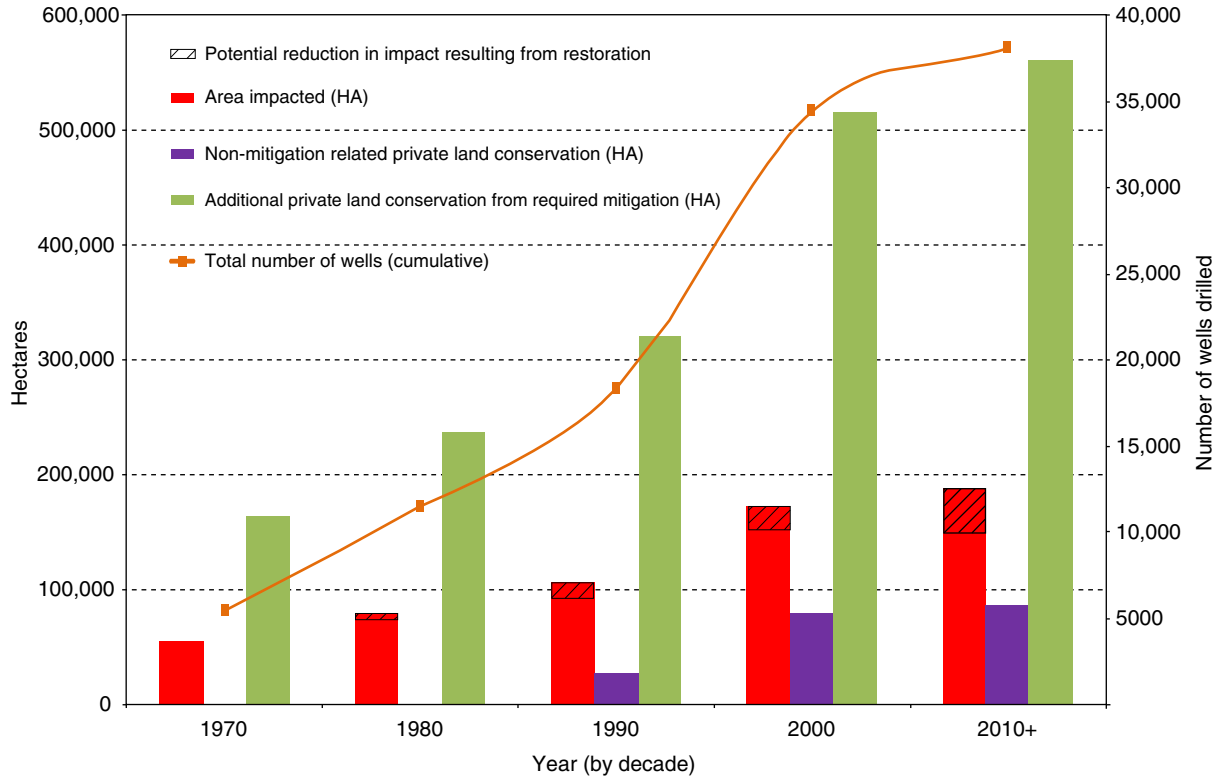
Applying Development by Design to Achieve Sustainable Landscapes: A Case Study of the WBE

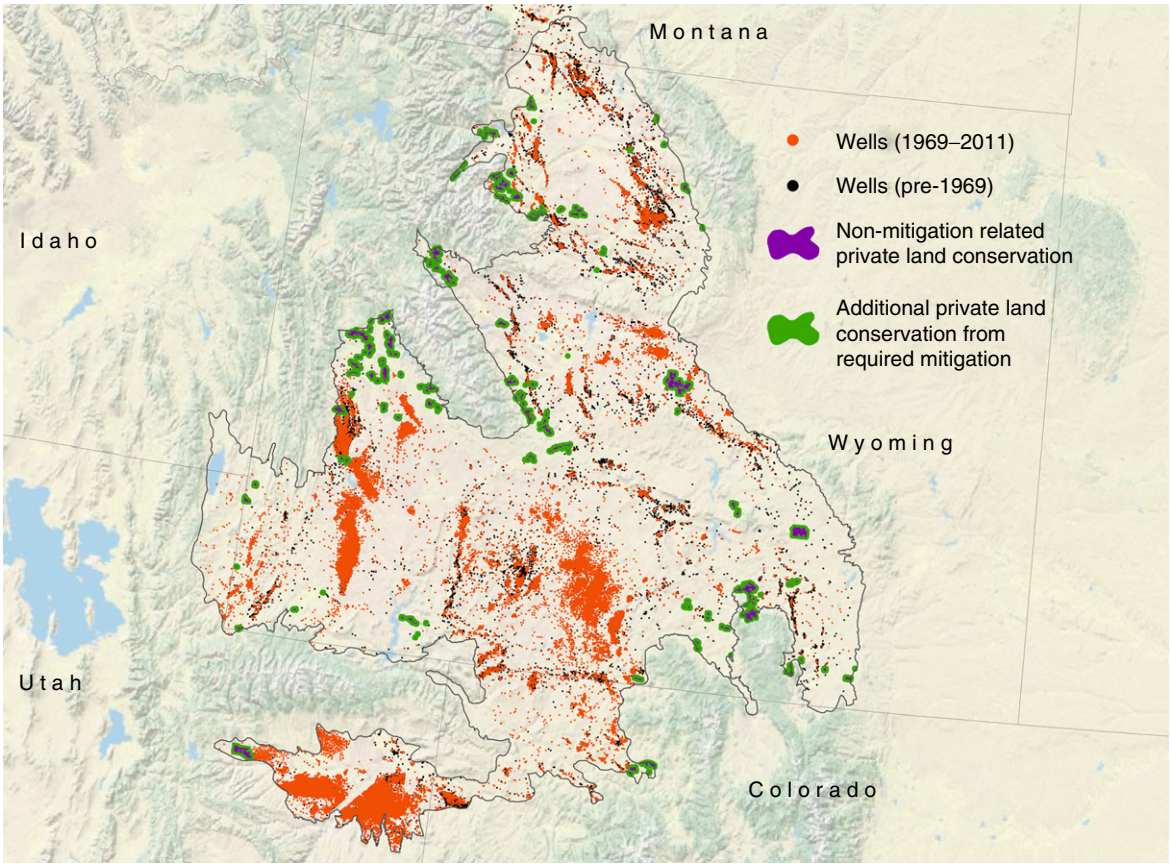
The WBE covers 13.3 million hectares of basin, plain, desert, and 'island' mountains in Colorado, Idaho, Montana, Utah, and Wyoming (Bailey, 1995). US energy demand, energy security issues, and concern over climate change are driving the exploitation of traditional and renewable energy sources abundant in the WBE (McDonald *et al.*, 2009; Copeland *et al.*, 2009). The WBE is home to some of the best onshore wind resources, as well as a significant amount of onshore natural gas deposits (Copeland *et al.*, 2009; Kiesecker *et al.*, 2011b). At the same time, the ecoregion provides critical habitat for migratory big game, songbirds, and raptors within the reaches of the Greater Yellowstone Ecosystem. Some of the world's largest herds of mule deer (*Odocoileus hemionus*) and pronghorn antelope (*Antilocarpa americana)* winter here, relying on the snow-free forage to get them through harsh winter weather. The area is also a stronghold for the greater sage grouse (*Centrocercus urophasianus*), an emblematic native game bird now currently listed as warranted but precluded under the Endangered Species Act. Because of the population declines of several species, including sage grouse, restrictions are increasingly being placed on energy development, curtailing exploitation of both gas and wind resources (Doherty *et al.*, 2010). The current situation appears to be suboptimal for both energy development and conservation. Could we have done better in balancing the competing demands for energy and conservation? In this case study, we apply the DbD framework retrospectively to the WBE to explore how the trajectory of energy development and conservation would have been different over the past four decades. We examine the lessons for mitigation planning and for promoting sustainable landscapes in the future.

To identify areas that would maintain long-term persistence of representative biodiversity, the Nature Conservancy—along with key state and federal land management and wildlife regulatory agencies, universities, and other conservation organizations—designed an ecoregional plan for the WBE (Freilich *et al.*, 2001). The portfolio of sites chosen during the ecoregional assessment conducted for the WBE (Freilich *et al.*, 2000) totals 3.5 million hectares or 27% of the total area in the ecoregion. Side by side with the WBE's wildlife resources are some of the western US' richest oil and gas deposits (US DOI, 2006; Copeland *et al.*, 2009), including some that intersect areas selected in the ecoregional assessment (Kiesecker *et al.*, 2010). To date, the WBE is home to more than 55,000 oil and gas wells (**Figure 5**). Conservation of the biological diversity in this ecoregion is in question, in part because the US government has authorized exploration and development in 4 million of the 8 million hectares (52%) of the federal mineral estate within the WBE (Doherty *et al.*, 2008). In fact, the number of producing wells in the ecoregion has nearly tripled from the 1980s to the present and is expected to increase over the next 30 years (Copeland *et al.*, 2007; Doherty *et al.*, 2011; **Figure 5**). At the same time, declines in populations of several species (i.e., sage grouse) have resulted in restrictions that have impacts on the future of traditional and renewable energy development alike.

Here we use past oil and gas development activities and the portfolio of sites selected in the ecoregional assessment to demonstrate how application of the mitigation hierarchy can be improved to deliver better outcomes for both conservation and development. We located the position of every well drilled in the WBE and made mitigation recommendations for each. Well data were acquired from state-level oil and gas regulatory agencies (Colorado: http://cogcc.state.co.us/; Montana: http://www.bogc.dnrc.mt.gov/onlinedata.asp; Utah: http://oilgas.ogm.utah.gov/Data_Center/DataCenter.cfm; Wyoming: http://wogcc.state.wy.us/urecordsMenu.cfmSkip='Y'&oops=#oops#). Where available, we used the spud date (the start of drilling on a new well) to indicate the year a well was put in. In the absence of a spud date, we used other information such as first production date to best approximate the start of activity on a well site. We were unable to establish well activity dates for less than 2% of the original well locations. Records where well status codes (e.g., abandoned location, unapproved permit to drill) suggested no physical activity ever occurred on a proposed well site were not included in the analysis. We focused on wells drilled from 1969 to the present since that is when the National Environmental Policy Act (NEPA) was established. NEPA is an environmental law that established a US national policy promoting the enhancement of the environment and the use of the mitigation hierarchy (Council on Environmental Quality, 2000). We calculated potential disturbance associated with each well and estimated the amount

Figure 4 Use of the ecoregional portfolio to select suitable offsets sites for oil and gas development in the Eastern Steppe of Mongolia. Ecoregional data provides the currency to assess impacts associated with development and compare which portfolio sites are best suited to serve as offset sites. Inset map shows the location of the study area within Mongolia. For details, see Heiner M, Davaa G, Kiesecker J, *et al.* (2011) *Identifying Conservation Priorities in the Face of Future Development: Applying Development by Design in the Grasslands of Mongolia.* Arlington, Virginia: The Nature Conservancy.


Potential reduction in impact resulting from restoration
Area impacted (HA)
Non-mitigation related private land conservation (HA)
Additional private land conservation from required mitigation (HA)
Total number of wells (cumulative)
Hectares
Number of wells drilled
600,000
500,000
400,000
300,000
200,000
100,000
0
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5000
1970
1980
1990
2000
2010+
Year (by decade)


Montana
Idaho
Wyoming
Utah
Colorado
Wells (1969–2011)
Wells (pre-1969)
Non-mitigation related private land conservation
Additional private land conservation from required mitigation

of 'required mitigation.' Estimates of the impacts associated with oil and gas development vary widely, due in part to the nature of the development and the species and systems potentially impacted. To generate an estimate of impacts associated with an 'average well,' we examined data available from oil and gas development in the WBE. Estimates of direct disturbance from several US Bureau of Land Management's environmental impact assessments suggest that ~2.5 ha (~6 acres) of disturbance is associated with each well's well pad, roads, and pipelines (US BLM, 2006, 2007). Studies that have examined indirect estimates of disturbance suggest that upward of ~24 ha (~60 acres) of disturbance is associated with each well (Sawyer, 2009; Doherty *et al.*, 2010; Copeland *et al.*, 2009). Although we recognize that impacts of development extend beyond the footprint of direct surface disturbance, we chose a conservative estimate of 4 ha (10 acres) of disturbance impact for each well and required that offsets for these impacts be implemented.

We simplified the analysis in other ways as well. We recognize that offsets do not typically deliver conservation value per unit area equal to the area that is lost as a result of development. Under the DbD framework, we would strive to require that offset benefits are estimated using the principles discussed in step 4 to ensure that offsets deliver additional conservation that can be appropriately valued. For simplicity in this analysis, we used a 3 : 1 replacement ratio established as part of the only oil and gas development project in the WBE that has required offsite mitigation (US DOI, 2006, Kiesecker *et al.*, 2009). Offset activities can span a variety of actions from restoring degraded habitats to protecting habitats from future degradation. In the WBE, one of the key threats to biodiversity is the conversion of large working ranches to small ranchettes or residential subdivisions (Gude *et al.*, 2007). Easements are an effective tool for preventing habitat conversion of private lands and the only form of mitigation we considered in this analysis (Kiesecker *et al.*, 2007; Pocewicz *et al.*, 2011). Although under the DbD framework we would consider a variety of offset actions, for simplicity we consider only private land conservation easements and compared the amount of private land conservation occurring from nonmitigation-related conservation to new private land conservation resulting from DbD related mitigation. We used the Protected Areas Database of the US (PAD-US v1.2) created by USGS GAP (http://gapanalysis.usgs.gov/data/padus-data/) to assess the quantity of protected private lands. We supplemented these data with locally available data (e.g., The Nature Conservancy) on additional known privately conserved lands in the project area. Approximately 10% of the available privately protected areas data had no date to indicate when a protection mechanism was placed on the land. In those cases, we redistributed acres proportionately across the three decades, starting in the 1980s when it is most probable that the private land conservation would have occurred. Consistent with the DbD framework, before estimating the amount of impacts associated with development and the amount of offsets needed to compensate for these impacts, we first sought to indicate which wells would have been incompatible with landscape conservation goals and should have been avoided. To decide which impacts should have been avoided, we followed the approach outlined in Kiesecker *et al.* (2010) and described in the preceding step 2.

Our analysis indicated that more than 55,000 oil and gas wells have been drilled in the WBE, with 41,072 of these drilled since 1969 and subject to our application of DbD. The pattern of development shows a marked increase in the number of wells over time, with a low of 5542 wells drilled from 1970 to 1979 and a high of 17,761 wells drilled from 2000 to 2009 (**Figure 5**). Strikingly, only 2886 (~7%) of all wells were drilled in areas where our DbD analysis would have recommended that impacts be avoided. Wells compatible with landscape-level conservation goals would have resulted in impacts totaling 154,533 ha (381,860 acres) based on a estimated footprint of 4 ha (10 acres) per well. If we required 3 ha of offset for every hectare of impact, we would have generated 463,599 ha (1,145,580 acres) of additional private land conservation. Compare this with the 87,412 ha (216,000 acres) of private land conservation conducted over the same period; we see application of DbD results have a more than fivefold increase in conservation (**Figure 5**). We recognize this analysis incorporates simplifying assumptions; a more robust approach would consider a range of other factors such as how restoration actions may be addressing energy development impacts and contributing to conservation values, as these impacts are unlikely to persist indefinitely. Although there are limited data on the efficacy of restoration, if we assume a 30% restoration rate per decade then impacts would be reduced to 108,173 ha (267,302 acres) (**Figure 5**). Nonetheless, we believe this retrospective application of the DbD framework shows how, if DbD had been implemented as part of normal practices associated with energy development over the past four decades, it could have contributed to significant benefits for sage grouse as well as other species and perhaps alleviated current restrictions placed on energy development. Consider that if the impacts of wells incompatible with conservation goals (7% of wells) had been avoided and offsets had been implemented for wells compatible with conservation goals (93% of wells), then the WBE would have 498,637 ha (1,232,160 acres) more in valuable protected wildlife habitat. This represents 7.5% of all core sage grouse habitat in the WBE and 13% of the private lands within sage grouse core areas (Doherty *et al.*, 2011). If these additional conservation efforts had been targeted at sage grouse and other species of conservation concern, then we might not be facing the current uncertainties regarding listing under the Endangered Species Act and the consequences that would have for future energy development.

Figure 5 Applying Development by Design retroactively in the Wyoming Basins Ecoregion. Map shows location of wells, private land conservation, and additional private land conservation resulting from "required offsets". The number of wells drilled per decade is shown from 1969 to the present. Impacts associated with development are based on an estimate of 4 ha footprint per well. Additional private land conservation resulting from "offsets" is compared to existing private land conservation. Estimates of additional private land conservation are based on a requirement of 3 ha offset for each hectare of impact. For simplicity, additional private land conservation resulting from "offsets" were clustered around existing private land conservation.

Conclusion

Offsets represent an opportunity to mobilize billions of dollars for conservation (Burgin, 2008; McKenney and Kiesecker, 2010). If we ensure that development is consistent with broader conservation goals, avoiding impacts where they are not, then development can be the fuel to achieve significant conservation outcomes. The results of our simple analysis suggest that when offsets are a normal part of project activities, as regular practice for fully internalizing project impacts, then the level of conservation can exceed traditional sources by many multiples. Although incorporation of offsets into normal practices could obviously have benefits for conservation, it could also benefit development activities. Given the extensive amount of development projected for the WBE, a requirement for development to attain no net loss through application of DbD could be the impetus to conserve biodiversity and proactively add the flexibility needed to develop resources. This will be important as the US moves to exploit more of its domestic energy resources, particularly renewable energy (McDonald *et al.*, 2009; Kiesecker *et al.*, 2011b). Moreover, global development patterns suggest that changes projected for the WBE are closer to the rule rather than the exception. Lessons learned here could have implications for designing development that are compatible with nature and human well-being, reducing conflicts for both biodiversity conservation and business.

Appendix

University Courses

- Applied Ecology
- Conservation Biology
- Environmental Impact Assessment

See also: Agriculture, Sustainable. Conservation and People. Ecological Footprint, Concept of. Economic Value of Biodiversity, Measurements of. Economics of the Regulating Services. Ecosystem Services. Energy Use, Human. Environmental Impact, Concept and Measurement of. Feeding the World and Protecting Biodiversity. Government Legislation and Regulations in the United States. Human Impact on Biodiversity, Overview. Identifying Conservation Priorities using a Return on Investment Analysis. Implications of Urbanization for Conservation and Biodiversity Protection. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production. Loss of Biodiversity, Overview. Market Economy and Biodiversity. Oil-Palm Plantations in the Context of Biodiversity Conservation. Restoration of Biodiversity, Overview. Sustainability and Biodiversity. The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity

References

Arponen A, Kondelin H, and Moilanen A (2007) Area-based refinement for selection of reserve sites with the benefit function approach. *Conservation Biology* 21: 527–533.

Bailey RG (1995) *Description of the Ecoregions of the United States*. USDA Forest Service Misc. Publ. No. 1391, 2nd edn. Washington, DC: US Department of Agriculture.

Ball IR (2000) *Mathematical Applications for Conservation Ecology: The Dynamics of Tree Hollows and the Design of Nature Reserves*. PhD Thesis, The University of Adelaide.

Ball IR and Possingham HP (2000) MARXAN (V1.8.2): *Marine Reserve Design Using Spatially Explicit Annealing, a Manual*.

Brown JW (2006) *Ecological: An Ecosystem Approach to Developing Infrastructure Projects*. Washington, DC: Office of Project Development and Environmental Review, Federal Highway Administration.

Burgin S (2008) BioBanking: An environmental scientist's view of the role of biodiversity banking offsets in conservation. *Biodiversity and Conservation* 17: 807–816.

Copeland H, Ward J, and Kiesecker JM (2007) Threat, cost, and biological value: Prioritizing conservation within Wyoming ecoregions. *Journal of Conservation Planning* 2007: 1–16.

Copeland HE, Doherty KE, Naugle DE, Pocewicz A, and Kiesecker JM (2009) Mapping oil and gas development potential in the U.S. Intermountain West and estimating impacts to species. *PLoS ONE* 4: e7400.http://dx.doi.org/10.1371/journal.pone.0007400.

Council on Environmental Quality (2000) Protection of the Environment (under the National Environment Policy Act) (40 CFR 1500–1517).

Dinerstein E, Powell G, Olson D, *et al.* (2000) A workbook for conducting biological assessments and developing biodiversity visions for ecoregion-based conservation. Conservation Science Program, November, 2000 WWF.

Doherty KE (2008) *Sage-grouse and Energy Development: Integrating Science with Conservation Planning to Reduce Impacts*. PhD Dissertation, University of Montana, Missoula, MT, USA.

Doherty KE, Naugle DE, and Evans JS (2010) A Currency for offsetting energy development impacts: horse-trading sage-grouse on the open market. *PLoS One* 5: e10339.http://dx.doi.org/10.1371/journal.pone.0010339.

Doherty KE, Naugle DE, Copeland HE, Pocewicz A, and Kiesecker JM (2011) Energy development and conservation tradeoffs: Systematic planning for greater sage-grouse in their eastern range. In: Knick ST and Connelly JW, (eds.) *Greater Sage-Grouse: Ecology and Conservation of a Landscape Species and its Habitats. Studies in Avian Biology* vol. 38. pp. 505–516. Berkeley, CA: University of California Press.

Environmental Defense (1999) Mitigation Banking as an Endangered Species Conservation Tool. Environmental Defense. November 1999.

Environmental Law Institute (2002) *Banks and Fees: The Status of Off-Site Wetland Mitigation in the United States*. Washington, DC: Environmental Law Institute.

Environmental Law Institute (2007) *Mitigation of Impacts to Fish and Wildlife Habitat: Estimating Costs and Identifying Opportunities*. Washington, DC: Environmental Law Institute.

Federal Interagency Mitigation Workgroup (2002) National Wetlands Mitigation Action Plan. 24 December 2002.

Freilich J, Budd B, Kohley T, and Hayden B (2001) *The Wyoming basins ecoregional plan*. Lander, Wyoming: TNC Wyoming Field Office.

Gibbons P and Lindenmayer DB (2007) Offsets for land clearing: No net loss or the tail wagging the dog? *Ecological Management & Restoration* 8: 26–31.

Groves CR (2003) *Drafting a Conservation Blueprint: A Practioner's Guide to Planning for Biodiversity,* 457 pp. Washington, DC: Island Press.

Gude PH, Hansen AJ, and Jones DA (2007) Biodiversity consequences of alternative future land use scenarios in Greater Yellowstone. *Ecological Applications* 17: 1004–1018.

Heiner M, Davaa G, Kiesecker J, *et al.* (2011) *Identifying Conservation Priorities in the Face of Future Development: Applying Development by Design in the Grasslands of Mongolia.* Arlington, Virginia: The Nature Conservancy.

International Finance Corporation (2012) *Performance Standard 6, Biodiversity Conservation and Sustainable Management of Living Natural Resources*. World Bank Group. Available at: http://www.ifc.org/ifcext/policyreview.nsf/AttachmentsByTitle/Updated_PS6_August1-2011/$FILE/Updated_PS6_August1-2011.pdf

Kiesecker JM, Comendant T, Grandmason T, *et al.* (2007) Conservation easements in context: A quantitative analysis of their use by the nature conservancy. *Frontiers Ecology and the Environment* 5: 125–130.

Kiesecker JM, Copeland H, Pocewicz A, *et al.* (2009) A framework for implementing biodiversity offsets: Selecting sites and determining scale. *Bioscience* 59: 77–84.

Kiesecker JM, Copeland H, Pocewicz A, and McKenney B (2010) Development by design: blending landscape level planning with the mitigation hierarchy. *Frontiers in Ecology and the Environment* 8: 261–266.

Kiesecker JM, Copeland HC, McKenney B, Pocewicz A, and Doherty K (2011a) Energy by design: Making mitigation work for conservation and development. In: Naugle DE (ed.) *Energy Development and Wildlife Conservation in Western North America*, pp. 159–182. Washington, DC: Island Press.

Kiesecker JM, Evans JS, Fargione J, *et al.* (2011b) Win-win for wind and wildlife: A vision to facilitate sustainable development. *PLoS ONE* 6: e17566.http://dx.doi.org/10.1371/journal.pone.0017566.

King DM and Price EW (2004) *Developing Defensible Wetland Mitigation Ratios: A Companion to "The Five-Step Wetland Mitigation Ratio Calculator"*. Solomons, MD: University of Maryland. Silver Spring, MD: National Oceanic and Atmospheric Administration.

Kusler J (2003) *Reconciling Wetland Assessment Techniques*. Berne, NY: Institute for Wetland Science and Public Policy, The Association of State Wetland Managers, Inc.

Madsen B, Carroll N, and Moore Brands K (2010) State of Biodiversity Markets Report: Offset and Compensation Programs Worldwide. Available at: http://www.ecosystemmarketplace.com/documents/acrobat/sbdmr.pdf

McDonald RI, Fargione J, Kiesecker J, Miller WM, and Powell J (2009) Energy sprawl or energy efficiency: Climate policy impacts on natural habitat for the United States of America. *PLoS One* 4: e6802.http://dx.doi.org/10.1371/journal.pone.0006802.

McKenney B and Kiesecker JM (2010) Policy development for biodiversity offsets: a review of offset frameworks. *Environmental Management* 45: 165–176.

National Research Council (2001) *Compensating for Wetland Losses 993 Under the Clean Water Act*, 348 pp. Washington, DC: National Academy Press.

Noss RF, Carroll C, Vance-Borland K, and Wuerthner G (2002) A multicriteria assessment of the irreplaceability and vulnerability of sites in the greater Yellowstone ecosystem. *Conservation Biology* 16: 895–908.

Pocewicz A, Kiesecker JM, Jones GP, Copeland HE, Daline J, and Mealor BA (2011) Effectiveness of conservation easements for reducing development and maintaining biodiversity in sagebrush ecosystems. *Biological Conservation* 144: 567–574.

Possingham HP, Ball IR, and Andelman S (2000) Mathematical methods for identifying representative reserve networks. In: Ferson S and Burgman M (eds.) *Quantitative Methods for Conservation Biology*, pp. 291–305. New York: Springer.

Pressey RL and Bottrill MC (2008) Opportunism, threats, and the evolution of systematic conservation planning. *Conservation Biology* 22: 1340–1345.

Pressey RL, Possingham HP, and Day JR (1997) Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. *Biological Conservation* 80: 207–219.

Rio Tinto (2004) *Rio Tinto's biodiversity strategy*. London, UK and Melbourne, Australia, 12 pp. http://www.riotinto.com/documents/ReportsPublications/RTBidoversitystrategyfinal.pdf

Saenz S, Walschburger T, Leon J, and Gonzalez J (2010) Propuesta metodologica para asignacion de compensaciones por perdida de biodiversidad. Convenio de Asociacion No.09 de 2008. Ministerio de Ambiente, Vivienda y Desarrollo Territorial, The Nature Conservancy, World Wildlife Fund, Conservacion Internacional, Colombia.

Sawyer H, Kauffman MJ, and Nielson RM (2009) Influence of well pad activity on the winter habitat selection patterns of mule deer. *Journal of Wildlife Management* 73: 1052–1061.

Sochi K and Kiesecker JM (2011) *Energy by Design: Cooperative Mitigation Planning for the San Juan Basin. Appendix B: Selecting Offset Opportunities*. San Juan Basin Wildlife Mitigation Plan, Colorado Division of Wildlife, Denver, Colorado.

Tear TH, Kareiva P, Angermeier PL, *et al.* (2005) How much is enough? The recurrent problem of setting quantitative objectives in conservation. *BioScience* 10: 835–849.

ten Kate K, Bishop J, and Bayon R (2004) *Biodiversity Offsets: Views, Experience, and the Business Case*. Gland, Switzerland and Cambridge, UK: IUCN, and London, UK: Insight Investment.

Thorne JH, Huber PR, Girvetz EH, Quinn J, and McCoy MC (2009) Integration of regional mitigation assessment and conservation planning. *Ecology and Society* 14: 47–63.

US BLM (United States Bureau of Land Management) (2006) *Record of decision for the Jonah Infill Drilling Project Environmental Impact Statement*. Wyoming: Department of the Interior, Bureau of Land Management Pinedale Field Office.

US BLM (Bureau of Land Management) (2007) *Hiawatha Regional Energy Development Project Environmental Impact Statement*. Department of the Interior, Bureau of Land Management Rock Springs Field Office Wyoming and Little Snake Field Office Colorado.

US DOI (United States Department of the Interior) (2006). *Energy Policy and Conservation Act of 2000 Phase II Inventory: Scientific Inventory of Onshore Federal Lands' Oil and Gas Resources and the Extent and Nature of Restrictions or Impediments to Their Development*. Report BLM/WO/GI-03/002 + 3100/REV06 Washington, DC.

Differentiation

Nicholas H Barton, IST Austria, Klosterneuburg, Austria

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptive landscape A graph of average fitness against the state of the population, represented by allele frequencies or trait means.
Balancing selection Selection which acts to maintain two or more alleles in a stable equilibrium.
Character displacement The displacement of characters caused by competition between species which live in the same place.
Coalescent process The merging of lineages traced backward in time from a sample of genes.
Fitness The total number of offspring produced over one generation.
Frequency-dependent selection Selection in which fitnesses vary with the frequency of different genotypes.
Gene flow The movement of genes from place to place.
Hybrid zone A narrow region in which genetically distinct populations meet, mate, and produce hybrids.
Neutral Having no effect on fitness.
Random genetic drift The random fluctuation in allele frequency caused by random variation in fitness between individuals.
Stabilizing selection Selection on a continuously varying trait toward an intermediate optimum.
Sympatric speciation The separation of a single population living in one place into two species.
Transposable element A DNA sequence which can move from one place in the genome to another.

Introduction

In evolutionary biology, the term differentiation usually refers to the accumulation of genetic differences between populations or species. However, the term can be applied more broadly, to the diversification of genes, organisms, and populations. The extraordinary conservation of the basic genetics and biochemistry across all living organisms implies that life evolved from one most recent common ancestor; the obstacles to the chemical evolution of the first replicator suggests that this ancestor itself evolved from a single origin – most likely, a short self-replicating RNA sequence. Thus, we must explain how past and present biodiversity has differentiated from this simple beginning. We can ask how the first replicating molecules diversified to produce the many kinds of gene in present-day organisms; how the first cells evolved into these diverse organisms; and how populations have come to contain so much genetic diversity. This article sets out to explain how elementary evolutionary and ecological processes lead to differentiation in both the narrow and broad senses.

If we regard species as simply groups of organisms with similar morphology, then the differentiation of populations becomes equivalent to the process of speciation. However, under the biological species concept, speciation consists of the evolution of genetic differences that prevent successful interbreeding; under the phylogenetic species concept, speciation consists of the separation of lineages. Thus, at least in the sexually reproducing populations to which these ideas of species apply, population differentiation is distinct from speciation.

In this article, the author introduces successively more complex processes involved in differentiation. In the simplest kind of evolution, organisms, and populations differentiate under mutation and random reproduction. If there is inherited variation in the rate of reproduction (i.e., in fitness), then natural selection will shape variation; populations in different environments will become differentiated in a straightforward way. This divergence may be opposed by the flow of genes from place to place, but will only be prevented over the smallest spatial scales. Just as diverse ecological niches are required if species are to coexist, so selection can maintain differentiation within a population only if genotypes exploit different limiting resources. Finally, the author considers how far these genetical and ecological processes can account for differentiation in the broadest sense, and over the largest scales.

Neutral Evolution

Even if all copies of a gene are equivalent, some will by chance leave more descendants than others, leading to random drift of gene frequencies. As we trace their descendant lineages forward in time, some will go extinct, while others will grow in number. Eventually, it is inevitable that one gene will become the ancestor of the entire population (**Figure 1**). Therefore, if we trace the ancestry of all present-day genes backward in time, the lineages will at some point coalesce in a single common ancestor. In an ideal population of N diploid individuals, the chance that any two lineages coalesce is just $1/2N$ per generation. A large sample of genes will take about $2N$ generations to coalesce into two ancestral lineages, and these two will then take, on average, another $2N$ generations to coalesce in one ancestor. Much attention has focused on identifying "mitochondrial Eve," the single female who (assuming strictly maternal inheritance) was the most recent common ancestor of all human mitochondria. The key issue here is when this ancestor lived: a recent Eve would imply a small effective size for the human population.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00031-9

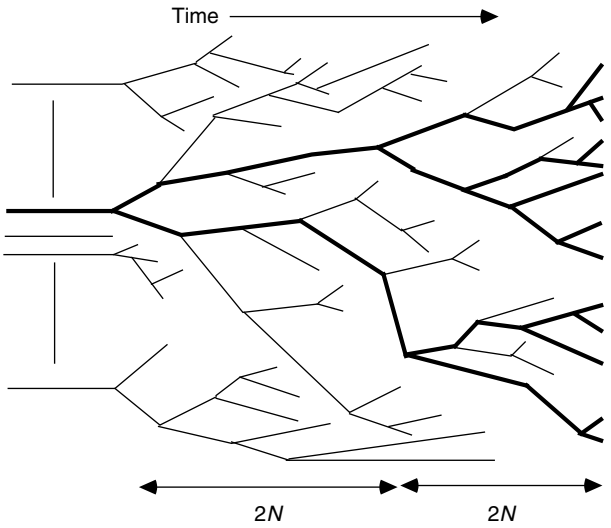


Figure 1 A sketch of the random reproduction of genes. Moving forward in time (left to right), genes leave random numbers of descendants. Eventually, only one family of descendants survives (in bold). Moving backward in time (right to left), the ancestry of the population takes on average $2N$ generations to coalesce to two lineages, and then the same time, on average, to reach the single common ancestor.

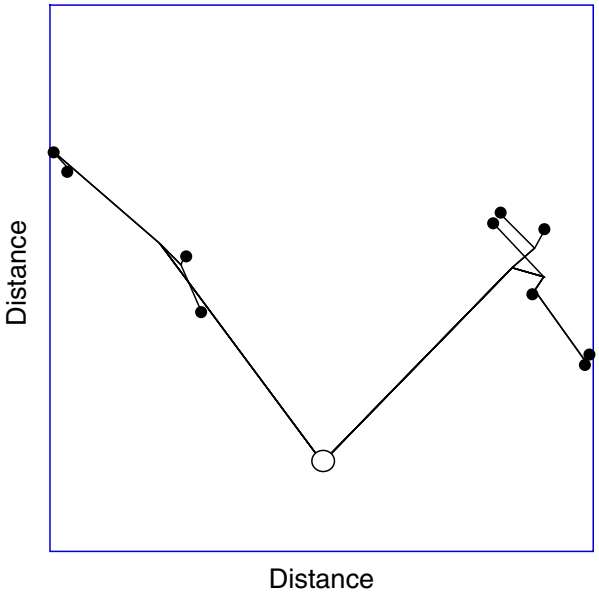


Figure 2 An example of the clustered pattern produced by random reproduction. Each dot represents an asexual organism, or a separate species. The position of each point represents its morphology, reduced to two dimensions; this evolves in a random walk, following a normal distribution. Each lineage goes extinct with constant probability, and splits at a rate which decreases with the total number of lineages. The open circle represents the ancestor.

This coalescent process determines the simplest form of differentiation of genes and populations. Suppose that the gene in question consists of a long stretch of DNA, which is always passed on as a unit. Mutations will occur at a rate μ, at random times and at random sites in the sequence; to a good approximation, mutations never occur at the same site twice (the infinite sites model). If we examine sequences which shared an ancestor a time T in the past, they will on average differ by $2\mu T$ mutations. If these gene sequences come from two well-separated species, then we can date the divergence of those species by the steady ticking of a molecular clock. Indeed, it was the striking observation that any given protein accumulates mutations at a steady rate that stimulated Kimura to develop the neutral theory of molecular evolution. Across a range of eukaryote lineages, sequences which have no effect on fitness mutate and diverge at a rate of $\sim 10^{-8}$ per base pair per year.

Variation within a population can be understood in a similar way. If two gene sequences are chosen at random, they will share an ancestor on average $T = 2N$ generations back; therefore, they will differ by, on average, $4N\mu$ mutations. This simple prediction is in qualitative agreement with observations on molecular genetic diversity within populations. However, variation in extremely large populations is not vastly greater than in much less abundant species: For example, about 5% of enzyme loci in humans are heterozygous, while about 25% are heterozygous in the bacterium *Escherischia coli*. There are two possible explanations for the weak relation between within-population diversity and census numbers. First, the effective population size may be much smaller than the actual census number, because of occasional bottlenecks and local extinctions, or because of selection on linked genes. For example, in *Drosophila melanogaster*, $\sim 1\%$ of noncoding base pairs are heterozygous, which, with $\mu = 10^{-8}$ per base per year, implies that a random pair of bases share a common ancestor $T \sim 0.01/2\,\mu = 5 \times 10^5$ years back. Assuming 4 generations per year, this implies an effective population number of $\sim 10^6$ flies – many fewer than the census number. Second, variation in coding and regulatory regions of the DNA may be maintained by selection, regardless of population size.

Neutral evolution does not lead to a homogeneous pattern of differentiation. Imagine that a process of random reproduction produces lineages as in **Figure 1**. (We might think of asexual organisms, or of species originating and going extinct.) If morphology changes by random increments, then the variance amongst a group reflects the age of that group. Now, random genealogies typically have a hierarchical structure, with most family groups being closely related compared with the age of the whole. Morphology is therefore expected to fall into well-separated clusters (**Figure 2**). Of course, the clustering that we see in the living world, which allows us to classify organisms into species, genera, and higher taxa, is not merely due to random reproduction. However, it is important to realize that what Darwin termed "descent with modification" tends naturally to produce clusters of distinct morphologies.

The recent availability of DNA-based markers, which for the most part can be presumed to have negligible effects on fitness, has transformed evolutionary biology. These provide a molecular clock against which to measure divergence times; an objective and quantitative means of reconstructing phylogeny; and a baseline level of differentiation within and between populations, against which to measure other processes. Having established this straightforward benchmark, we can consider how selection, gene flow, and ecological interactions influence differentiation.

Natural Selection

A difference in relative fitness of s (the selection coefficient) causes change over a time $\sim/f(1,s)$; it dominates over random genetic drift if $s > 1/N$, and over mutation if $s > \mu$. Thus, even slight fitness differences of $s > 10^{-5}$ will be the predominant force in a population of effective size larger than 10^5 individuals, and for genes consisting of (say) 10^3 base pairs, each with mutation rate 10^{-8} per generation. Such minute fitness differences cannot be measured directly. However, a multitude of studies show that selection on morphological traits and on individual genes can be strong ($s > 10\%$, say). Nevertheless, we will see that this does not imply that differentiation among populations is a simple consequence of selection alone.

In this section, the author considers the straightforward effects of selection in the simplest case, where each genotype (i.e., each set of genes) has a fixed effect on fitness. Variation within populations is then maintained by a balance between mutation generating deleterious variants, and selection eliminating them. (For the moment, we leave aside cases in which heterozygotes are fitter than either homozygote.) While mutation rates on individual bases are extremely low, the net mutation rate over the whole genome may be high, and so may sustain substantial genetic variation. For example, Eyre-Walker and Keightley (1999) used rates of DNA sequence divergence among primates to show that the human lineage has been subject to roughly one deleterious mutation per genome per generation. At the level of the whole organism, the variance of continuous traits such as growth rate or skeletal proportions increases by roughly 0.1% every generation as a result of new mutations; thus, much of the response to long-term artificial selection is due to mutation, rather than to standing variation. The genetic basis of such mutations is beginning to be understood. For example, Lai *et al.* (1994) found that $\sim$10% of the variance in bristle number in a natural population of *D. melanogaster* was due to the insertion of transposable elements around the *scabrous* gene; these insertions are clearly deleterious, since they were all found to be of relatively recent origin, and hence rapidly eliminated by selection.

There is a long-standing debate over the processes which maintain genetic variation. While this is as yet unresolved, it is at least plausible that most variation, in both discrete genes and continuous traits, is due to deleterious mutations. On this view, differentiation within populations is merely a fluctuation around some optimal state, while differentiation between populations is a direct consequence of changes in that optimum. The ability of artificial selection to rapidly change almost any aspect of an organism was Darwin's strongest argument for the efficacy of natural selection, and remains so today. Almost all traits show heritable variation, and therefore respond to selection; what is remarkable is that this response continues apparently without limit (at least, over the span of human experiments). For example, Weber (1992) selected on the angle between two small veins on *Drosophila* wings. Lines selected in opposite directions diverged steadily, and differed by 20 standard deviations after 15 generations. We now understand that at the genetic level, this divergence is due to the establishment of individual alleles, either present in the original base population, or generated by new mutations. Weber's lines, for example, differed by at least 11 genes on one chromosome, each of small effect.

There has been much dispute over whether adaptive differences between populations and species are due to the accumulation of minute differences, or to a few genes of large effect. Darwin emphasized the former view, on the grounds that complex and finely adjusted adaptations could only be built up by slight variations. In contrast, the geneticists who rediscovered Mendelian heredity emphasized the importance of major mutations. After a century of research, the answer is only beginning to become clear. On the one hand, sustained, steady responses to selection must be based on multiple genes of small effect, and genetic analysis of species differences usually shows that many genes are involved in maintaining their separation. On the other hand, differentiation in particular traits is often based on one or a few genes. Insecticide resistance usually evolves through mutations in particular target genes (e.g., in the acetyl cholinesterase enzyme that is the target of organophosphates). To take a natural example, the striking differences between warning color patterns of *Heliconius* butterflies are due to only a handful of genes (Counterman *et al.*, 2010). Much effort is now being devoted to identification of the quantitative trait loci responsible for human variation related to disease, for differences between varieties of domesticated plants and animals, and for adaptive differences between species. Such methods do often identify genes of major effect: indeed, the whole enterprise of genetic engineering rests on the possibility of obtaining useful results by manipulating a few key genes. However, there is a statistical bias toward finding genes of large effect, and recent large-scale studies of human height suggest that heritable variation depends on a very large number of alleles of small effect (Yang *et al.*, 2010).

Given that selection can rapidly adapt populations to changing conditions, and given that selection is often the strongest evolutionary force, one might suppose that differentiation would simply reflect the optimal genotype which is favored by the local environment. We could then say rather little about what patterns of differentiation to expect, since they would be the outcome of the particular interactions between organism and environment that determine reproductive success. However, matters are not so simple. There may be alternative combinations of genes which are locally optimal, and so populations may adapt to essentially the same conditions in different ways. Sewall Wright devised an influential way of thinking about this issue, the "adaptive landscape" (**Figure 3**). This is a surface which represents the average fitness of a population, plotted against a set of axes which describe the state of a population (e.g., gene frequencies or trait means). Under the simple assumptions of this section, natural selection pushes populations uphill on this landscape, toward a local "adaptive peak." In general, there will be many such peaks, and so populations will differentiate under selection even if they all experience the same environment.

Divergence may be observable at the phenotypic level, or may be cryptic. For example, suppose that a trait is determined as the sum of effects of many genes, each carrying either "+" or "−" alleles. If selection favors a single intermediate optimum, any combination of genes that achieves that optimum

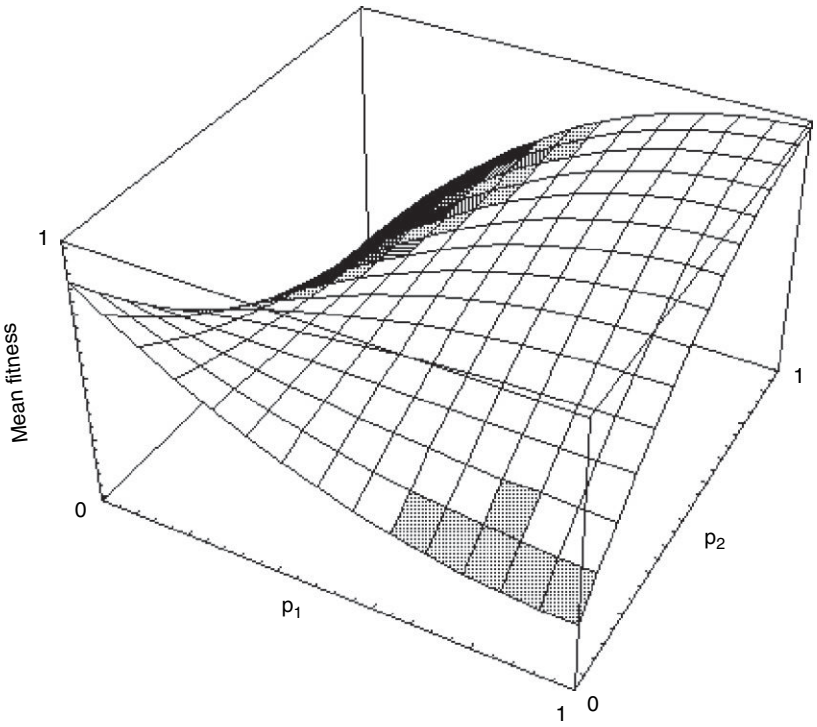


Figure 3 A simple "adaptive landscape" for two loci, each carrying two alleles. One allele at each locus is dominant, with frequency denoted by p_1 and p_2. The double dominant (P1P2) and double recessive (Q1Q2) genotypes are assumed to be most fit, with $W_{PP} = 1$, $W_{QQ} = 0.8$, $W_{PQ} = 0.4$, and $W_{QP} = 0.2$.

can be established ($++--$, $--++$, ...). Such cryptic divergence is revealed when crosses between populations produce progeny with reduced fitness and increased variance; over the largest scales, it is reflected by sibling species, which are morphologically indistinguishable and yet reproductively incompatible.

Divergence toward alternative "adaptive peaks" can occur in several ways. Random genetic drift in small populations may knock populations into the domain of attraction of different peaks. This process is especially likely if there are flat ridges on the adaptive landscape along which divergence is not opposed by natural selection. For example, if only the double recessive homozygote $P_1P_1P_2P_2$ is unfit, then populations initially containing Q_1 and Q_2 can fix either P_1 or P_2, with no effect on fitness. However, when crossed they have reduced fitness. Genetic analysis of hybrid inviability and sterility in *Drosophila* show that incompatibilities are largely recessive, which supports this model of quasi-neutral divergence. A different line of evidence comes from studies of functional RNA molecules (e.g., the transfer RNAs that translate the genetic code). Many changes in individual bases cause no appreciable change in three-dimensional structure, so that extensive divergence can occur without disruption of function.

Random processes can drive divergence even in large populations. If adaptation depends on new mutations, then different mutations may arise first in different places. For example, humans have evolved resistance to malaria by a variety of different mutations in the β-hemoglobin molecule: the S allele in Africa, the C allele in Asia, and various thalassemias in the Mediterranean. These alternative alleles tend to exclude each other, so that chance differences in first occurrence of mutations leads to permanent divergence. Finally, divergence may be driven by fluctuations in selection. Suppose that the "adaptive landscape" fluctuates as conditions change, but maintains overall distinct peaks. Then, populations may end up on different peaks, even if they experience the same mix of environments: the outcome depends on the particular sequence of environments which they encounter.

While the "adaptive landscape" is a helpful metaphor, it gives a misleading picture of the geometry of evolution. Evolution can proceed in very many dimensions, rather than just the few which can be visualized by the human mind. This dimensionality is nicely illustrated by the model experiments of Eigen and co-workers, in which RNA molecules ~ 100 bases long are replicated by a polymerase derived from the Qβ virus. Under given conditions, some definite sequence (and its mutational variants) is picked out by selection from amongst 4^{100}–10^{60} possibilities. This is possible because any sequence is connected to any other by no more than 100 mutational steps. At any stage, 300 variants are available from a single mutation, since each base can mutate to three alternatives. Provided that there is some sequence of steps which leads uphill toward the optimal sequence, selection will effectively search from amongst the astronomical numbers of possibilities. This multidimensional view emphasizes that the outcome depends on which mutations are available: while evolution could take any of a large number of directions at any stage, this number is very much smaller than the total space of possibilities, and sets a strong constraint.

Gene Flow

The most obvious impediment to differentiation is gene flow: that is, the mixing of genes from different populations caused by movement of the individuals that carry them. At least since the mid-nineteenth century, there has been controversy over whether such mixing prevents differentiation. Darwin held that populations could adapt to gradually varying conditions across their range and that this could lead to the origin of new species. In contrast, Wagner emphasized the importance of barriers to migration as prerequisites for speciation. This view has been strongly propounded in more recent years by Ernst Mayr. However, both population genetic theory, and observations of geographic divergence, strongly support the Darwinian view, at least if one considers adaptations which can be built up by individual alleles or by continuous changes in quantitative traits.

Gene flow has straightforward and well-understood effects. The rate of gene flow into a single population is measured by m, the fraction of genes derived by immigration in each generation. Mixing then occurs over a timescale $\sim 1/m$. In a broadly continuous habitat, gene flow can be well approximated as a diffusion: in the same way that gas molecules diffuse as a result of random collisions, so genes diffuse through their habitat as a result of the random movement of the organisms that carry them. The rate of diffusion is measured by σ^2, the variance of distance between parent and offspring along any particular axis. Over a time T, this diffusion causes genes to spread over a distance $\sim \sqrt{(2\pi\sigma^2 T)}$. For example, the annual grasshopper *Chorthippus parallelus* is divided into two subspecies *C. p. parallelus* and *Chorthippus p. erythropus*, which met in the Pyrenees after the past Ice Age, $\sim 10^4$ years ago. These grasshoppers move by about 30 m in a year, so that $\sigma^2 \sim 900\ \mathrm{m}^2$ per year. Enzymes which are fixed for different alleles in the two subspecies mix over a region ~ 15 km wide, which is of the same order as the distance predicted by simple diffusion, $\sqrt{(2\pi\sigma^2 T)} \sim 7.5$ km (Hewitt, in Harrison, 1993).

Whether populations can adapt to local conditions, despite mixing, depends essentially on the relative rates of gene flow and selection. An allele with selective advantage s in the heterozygote can be established in a single population despite gene flow, provided that the rate of immigration of individuals carrying the alternative allele is lower than the rate of selection ($m < s$). If genes diffuse across a continuous habitat, then an allele can become established provided that it is favored in a region larger than some critical distance proportional to the characteristic scale $\sigma/\sqrt{(s)}$. New mutations have a chance of fixation approaching $2s$ if they arise within the appropriate region, but the probability falls to zero outside this region. Similar considerations apply to adaptations based on continuously varying traits. Suppose that fitness falls away in a Gaussian function with variance V_s (i.e., stabilizing selection); the optimal trait value varies from place to place. The characteristic scale over which populations can respond to local conditions is now $\sigma\sqrt{(V_s/V_g)}$, where V_s is a measure of the strength of stabilizing selection, and V_g is the additive genetic variance.

Since the dispersal range σ is typically much smaller than the species' range, these theoretical considerations imply that weak selection can allow a species to adapt to diverse local conditions. This can be seen directly. For example, the grass *Agrostis tenuis* has evolved tolerance to heavy metals on mines only a few meters across. Similarly, narrow clines ~ 10 km wide, separate races of *Heliconius* butterflies which carry alternative warning color patterns (Mallet, in Harrison, 1993). Such examples leave little doubt that gene flow need not prevent divergence over very short scales.

If variation in selection can cause differentiation over local scales, why do species not fall apart into a myriad of separate forms? Mayr was led to argue that gene flow prevents differentiation of populations by his observation that species remain morphologically homogeneous over much of their contiguous range: it is this homogeneity that makes taxonomy practicable. How can we explain this apparent conflict with the potential power of selection to cause differentiation?

Much variation in protein or DNA sequence may have such small effects on fitness that selection is negligible relative to gene flow. Such neutral variation will differentiate as a result of random genetic drift. However, because drift is weak in even moderately abundant populations, weak gene flow homogenizes neutral variation. In a single population, the number of migrants, N_m, is crucial. The variance of allele frequency within the local population is $\bar{p}(1-\bar{p})/(1+4N_m)$, where $\bar{p}$ is the allele frequency in the source population; hence, if more than one migrant per generation enters, there will be little differentiation. This result carries over to populations spread over two dimensions, either in an array of subpopulations, or in a continuous habitat. (In the latter case, the crucial parameter is $\rho\sigma^2$, where ρ is the population density.) The effectiveness of gene flow in overcoming random drift has an important consequence: even if selection within local populations is ineffective relative to drift ($N_s \ll 1$), the cumulative effect of selection across the whole population of N_{tot} individuals dominates if $N_{tot}s > 1$. For many purposes, a population with $N_m > 1$ can be treated as a single gene pool.

Many surveys have been made of geographic variation in gene frequency, using both protein and DNA variation. The usual finding is that the variance in allele frequency between locations is small ($\rho/\bar{p}(1-\bar{p}) < 5\%$; this ratio is termed F_{ST}), implying that if this variance is at an equilibrium between random drift and gene flow, the number of migrants is large ($N_m > 5$). There are of course exceptions: for example, the salamander *Ensatina eschscholtzii* has exceptionally low dispersal, and correspondingly, $F_{ST} \sim 0.7$ (Wake and Yanev, 1986). However, the general conclusion is that gene flow is sufficient to homogenize differentiation generated by random drift.

Surveys of geographic variation often aim at estimating the number of migrants (N_m). However, even if large-scale differentiation is generated by random drift, its geographic pattern may be dominated by past history, rather than by present-day rates of gene flow (see Whitlock and McCauley, 1998). If a species is extinguished from a large region (as happened in temperate regions during the Ice Ages), and then reinvades, the pattern of variation reflects the pattern of recolonization. In European human populations, Cavalli-Sforza and coworkers have reconstructed the history of population movements from the detailed genetic data that are available: many loci show traces of the introduction of distinct alleles as agriculture spread from the Near East. When the genealogical

relationships amongst DNA sequences can be inferred, historical events can be seen more directly. For example, Avise (1994) has identified in many species a cryptic subdivision which runs across the south-western United States; this is a relict of a separation into different Pleistocene refugia. Such reconstructions should be based on several unlinked genetic loci, however: within a sexual species, different loci may have different histories.

Much molecular differentiation may have slight effects on fitness, allowing it to be used to reconstruct the history of gene movements. However, we cannot explain the apparent homogeneity of species simply by an absence of selection. There are many examples of geographic variation at enzyme loci which are maintained by selection for adaptation to different environments. Evidence of selection may come from direct measurement, from close correlation with environment, or from discordance in pattern between different loci. (For example, the geographic pattern at several enzymes in European humans differs significantly from the majority of genes, whose pattern reflects population history.) Morphological traits are usually strongly selected, and so their geographic patterns therefore do not match those predicted from putatively neutral markers. Because selection can cause rapid divergence, even in the face of gene flow, extinctions and recolonizations can only generate a transient homogeneity in selected traits. Thus, we expect to see morphological patterns which follow recent selection imposed by environmental features, and cut across patterns at neutral markers which reflect population history. For example, Schneider *et al.* (1999) found that in the skinks of north-eastern Australia, mitochondrial DNA and microsatellite markers tend to change along the Black Mountain Corridor, a biogeographic feature marking past range changes. In contrast, skeletal morphology changes in a quite different pattern, which parallels the transition from open to closed forest.

The evident power of selection to cause divergence remains hard to reconcile with the apparent similarity of species across large areas which Mayr and others have emphasized. The problem is close to that posed by examples of stasis in the fossil record, in which morphology remains similar over long times rather than large distances. One possibility is that in sexual species, different traits or genes can evolve independent spatial patterns, so that overall morphology is more stable than any one component, and varies more smoothly. In any case, the supposed homogeneity may be exaggerated. Sharp transitions ("hybrid zones") do often occur, prompting classification of the taxa on either side as species or subspecies. These may arise either when many traits respond to a sharp environmental discontinuity (as in the skink example, above), or when populations that have diverged to become partially incompatible are brought together by a range expansion. In the latter case, divergence remains permanently because the hybridizing populations are at two different "adaptive peaks."

Ecological Differentiation

A simple form of natural selection, in which each gene combination has a fixed fitness, can explain much of evolution: it accounts for differentiation between populations and species as a result of changing conditions, mutation, and random drift, and despite gene flow. However, at the very least absolute fitness (i.e., the number of offspring) must decrease with density if the population is to remain bounded. We expect the relative fitnesses of different genotypes to change with their abundance in the population: selection is in general frequency-dependent. If the reproductive success of an allele decreases as it becomes more common, then genetic diversity can be maintained. This is known as balancing selection.

Maintenance of genetic diversity by frequency-dependent selection arises from competition for different limiting resources, in just the same way as maintenance of species diversity. If the concept of a "limiting resource" is interpreted broadly, then all examples of balancing selection can be seen in this way. Evidence for a heterogeneity of resources, cited by Darwin and exploited by plant breeders, comes from the increased yield of a mixture of varieties compared with a pure stand. Here, the limiting resources may consist of different nutrients, though other mechanisms are possible. A classic example of frequency-dependent selection is Batesian mimicry, in which a palatable species avoids predation by mimicking a distasteful model species. As the mimic becomes common, predators fail to associate their pattern with unpalatability, and protection is lost. Batesian mimics such as swallowtail butterflies therefore evolve diverse mimetic patterns, each of which remains rare. The limiting resources here are the various different model species. In diploid species, variation can be maintained by heterozygote advantage. The classic example is the polymorphism for the β-hemoglobin S allele, where the heterozygote gains resistance to malaria, while the homozygote suffers from sickle-cell anemia. The relative fitnesses of diploid individuals are fixed, but from the gene's point of view, fitness decreases with frequency: a common allele is more likely to find itself in a homozygote. The limiting resources here are the two kinds of allele with which each gene can find itself paired in a diploid.

The existence of heterogeneous limiting resources is, in a broad sense, necessary for maintaining diverse species, and may be responsible for much genetic differentiation within species. However, we do not know how much genetic variation is maintained by balancing selection, as opposed to mutation and gene flow, and we do not know how far organisms respond to environmental heterogeneity by evolving diverse genotypes or diverse species. A trade-off between variation within and between species is revealed by character displacement, where coexisting species tend to diverge in morphology and behavior. The *Anolis* lizards of the Caribbean provide a classic example. Where one species lives on an island, it has a broad distribution of body size; where two species share an island, they take up nonoverlapping distributions, and so partition the available prey. A similar phenomenon can be seen where asexual clones compete with their sexually reproducing relatives. The freshwater fish *Poeciliopsis monacha* and *Poeciliopsis lucida* occasionally mate, and produce a parthenogenetically reproducing clone. Vrijenhoek (1994) found that in streams with many different asexual clones, there tend to be fewer of the sexual progenitor species. This suggests that the space of available ecological niches can be filled either by a set of distinct asexual clones, or by members of the genetically diverse sexual population.

If multiple forms, each exploiting a different resource, coexist in a sexual population, then matings between them may produce maladapted progeny. To make this point in another way, imagine that resource use depends on a set of continuously varying traits. If these traits are determined as the sum of effects of several genes, then sexual reproduction produces an approximately normal trait distribution. There is no reason why this should match the distribution of resources which are available. The problem can be avoided if the genetic system can somehow produce the appropriate distribution despite the random shuffling of genes that occurs with sexual reproduction. For example, swallowtail butterflies produce many distinct wing patterns, each mimicking a different distasteful model. These are determined by several genes, which are so tightly linked that they behave as a single genetic locus. If such genetic tricks do not evolve, then selection favors mating behaviors that causes like to mate with like. Ultimately, such selection could lead to complete reproductive isolation between the two forms. This splitting of a single population into two separate species in the absence of spatial separation is termed sympatric speciation. It is almost impossible to demonstrate directly, since any present pattern of divergence might have originated with some spatial isolation. However, there are several plausible cases. For example, several distinct morphs of cichlid fishes have evolved within isolated crater lakes (Schliewen *et al.*, 2001). A key question is how often ecological differentiation amongst genotypes evolves into differentiation between species in this way.

Darwin emphasized that natural selection is driven by competition for limited resources (the "struggle for existence"). Moreover, he argued that the lack of intermediate forms – the evident clustering of organisms that makes taxonomy possible – can be explained largely by the diversifying effects of competition. There is a "division of labor" amongst organisms, which (as with market economics) results from competition between individuals rather than any optimization of species' or community productivity. Similar processes account for the differentiation of the $\sim 10^5$ different genes in complex organisms. Often, one gene may acquire multiple functions. For example, arginosuccinate lyase in birds acts as a lens crystallin as well as having a catalytic function. Similarly, most of the genes responsible for the establishment of segmentation in early *Drosophila* development are also involved much later in specifying various organs. If by chance a multifunctional gene duplicates, then each copy may specialize to become more efficient at one of the two original functions. The sequencing of complete genomes has made clear the extent of gene duplication and differentiation: most yeast genes are members of closely related families. There are close analogies between the division of labor among different species within an ecological community, different genotypes within a population, and different genes within an organism.

Evolution Over Large Scales

The processes of selection, mutation and random drift that lead to differentiation are well understood, at least over the short timescales that are accessible to experiment and observation. How far do they account for the differentiation of genes and organisms over the whole span of evolution? First, consider the overall pattern. The first fossil evidence of complex multicellular organisms is found in the Ediacaran fauna, about 560 million years ago. An extraordinary diversity of animals with skeletons evolved soon after, at the beginning of the Cambrian, and included representatives of almost all present animal phyla. Molecular evidence suggests that the ancestral lineages of these phyla may trace back much further; nevertheless, there was a rapid diversification of body plans during this "Cambrian explosion." The subsequent fossil record shows an overall increase in diversity, as measured by numbers of species, but interspersed with occasional mass extinctions. Organismal complexity is hard to quantify, and trends in complexity are controversial: while the most complex organisms may become more complex as diversity increases, it is not clear that there is (or should be) any trend toward more elaborate organisms. Rates of morphological evolution can be quantified more readily, at least for hard parts. Morphology changes much more slowly over the long timespan of the geological record than in current populations (Gingerich, 1983). Only weak selection pressures need be invoked to explain observed rates, and indeed, it is surprising that morphology has often changed little over millions of years.

If we suppose that organisms differentiate primarily as a result of selection to fill different ecological niches (more precisely, to exploit different limiting resources), then we might imagine that these macroevolutionary patterns could be explained by diversification to fill a set number of preexisting niches. At equilibrium, there might be a balance between extinction and species formation, leaving many niches unfilled; nevertheless, diversity of species and their morphology would be expected to rise toward some definite value. This view is supported by statistical analysis of the fossil record, which suggests saturation of species numbers. Over shorter timescales, remote islands such as the Hawaiian archipelago are occupied by the descendants of a few colonizing species, which radiated to fill the roles normally occupied by other taxa. However, new organisms generate new niches. One host species is often plagued by several specialist parasites, which are restricted to that host by specific adaptations to evade its defenses. The evolution of novel features (multicellularity, bony skeletons, flight, etc.) are followed by a rapid diversification into novel ways of living. While the maintenance of diverse genotypes and species can be understood formally in terms of competition for distinct limiting resources, their number is not fixed, and seems inherently unpredictable.

The pattern of diversity that we see may not be entirely a matter of historical accident. Even though there is no difficulty in explaining the rapid evolution of any one gene or trait, there are fundamental limits on the numbers of genes, the rates at which new genes are substituted for old, and the number of different genotypes that can be maintained within a population. The total mutation rate cannot greatly exceed one per generation, since most mutations are deleterious, and each deleterious mutation requires roughly one selective death to be eliminated. The human mutation rate approaches this upper limit, which may limit the number of functional genes

that we carry. Each substitution of one allele by another requires that individuals carrying that allele leave more offspring than those that do not: Haldane (1957) showed that this implies a "cost of natural selection" which limits the rate at which new alleles are substituted. Again, this argument places a strong constraint on the number of adaptive changes that can have occurred during human evolution. Finally, the diversity that can be maintained within a finite population is limited: the maximum number of genes that vary independently can be no greater than the effective population size. These constraints give some hope that we can understand the overall features of organismal differentiation, even if the detailed pattern is a conglomeration of unique historical events.

See also: Speciation, Process of. Speciation, Theories of. Species, Concepts of. Species Diversity, Overview

References

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. New York: Chapman & Hall.

Cavalli-Sforza LL, Menozzi P, and Piazza A (1994) *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Counterman BA, Araujo-Perez F, Hines HM, *et al.* (2010) Genomic hotspots for adaptation: The population genetics of Mullerian mimicry in *Heliconius erato*. *PLoS Genetics* 6: e1000796.

Eigen M (1992) *Steps Towards Life*. Oxford: Oxford University Press.

Eyre-Walker A and Keightley PD (1999) High genomic deleterious mutation rates in hominids. *Nature* 397: 344–347.

Gingerich PD (1983) Rates of evolution: Effect of time and temporal scaling. *Science* 222: 159–161.

Haldane JBS (1957) The cost of natural selection. *Journal of Genetics* 55: 511–524.

Harrison RG (1993) *Hybrid Zones and the Evolutionary Process*. Oxford: Oxford University Press.

Hatfield T and Schluter D (1999) Ecological speciation in sticklebacks: Environment-dependent hybrid fitness. *Evolution* 53: 866–873.

Lai C, Lyman RF, Long AD, Langley CH, and Mackay TFC (1994) Naturally occurring variation in bristle number and DNA polymorphisms at the scabrous locus in *Drosophila melanogaster*. *Science* 266: 1697–1702.

Lai C and Mackay TFC (1993) Mapping and characterization of P-element induced mutations at quantitative trait loci in *Drosophila melanogaster*. *Genetical Research* 61: 177–194.

Schliewen UK, Rassmann K, Markmann M, Markert JA, Kocher T, and Tautz D (2001) Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. *Molecular Ecology* 10: 1471–1488.

Schneider CJ, Smith TB, Larison B, and Moritz C (1999) A test of alternative models of diversification in tropical rainforests: Ecological gradients vs rainforest refugia. *Proceedings of the National Academy of Sciences (USA)* 96: 13869–13873.

Vrijenhoek RC (1994) Unisexual fish: Model systems for studying ecology and evolution. *Annual Review of Ecology and Systematics* 25: 71–96.

Wake DB and Yanev KP (1986) Geographic variation in allozymes in a ring species, the Plethodontid salamander *Ensatina eschscholtzii* of western North America. *Evolution* 40: 702–715.

Weber KE (1992) How small are the smallest selectable domains of form? *Genetics* 130: 345–353.

Whitlock MC and McCauley DE (1998) Indirect measures of gene flow and migration: F_{ST} not equal to $1/(4\ N_m + 1)$. *Heredity* 82: 117–125.

Yang J, Benyamin B, McEvoy BP, *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 42: 565–569.

Dinosaurs, Extinction Theories for

J David Archibald, San Diego State University, San Diego, CA, USA
Norman MacLeod, Natural History Museum, London, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–9, © 2001, Elsevier Inc.

Glossary

Ammonites Shelled group of invertebrates related to octopus, squid, and nautilus.
Amniota Taxon of vertebrates including mammals and reptiles (including birds).
Archaic ungulates (or condylarths) Extinct group or grade of mammals possibly giving rise to hoofed mammals.
Chondrichthyeans Sharks and their relatives.
Ectotherm Animals that produce their heat from external sources such as the Sun.
Endotherm Animals that produce their heat internally through metabolic means.
K/T Abbreviation for Cretaceous/Tertiary, usually in reference to the K/T boundary.
Nonavian dinosaurs Dinosauria excluding birds.
Ornithischians Bird-hipped dinosaurs such as horned and duck-billed dinosaurs.
Palynomorphs Pollen and spores produced by plants.
Saurischians Reptile-hipped dinosaurs such as theropods.

Sixty-five and a half million years ago, a 10 km diameter rock from space slammed into the Earth, almost instantly annihilating over 70% of all species, including all nonavian dinosaurs. Textbooks, the press, and many scientists have accepted some version of this Cretaceous/Tertiary (K/T) mass extinction scenario for over 20 years. Yet, even as the popularity of this single-cause explanation increased, nagging questions continued to be voiced as to the possible biologic effects of such an impact event. This is particularly true with regard to the record of plants and animals living on land and in freshwater habitats. For example, 12 clades representing about 107 species of better-studied vertebrates known from very near the K/T boundary in western North America suffered 51% extinction. Of these 12, however, only five—chondrichthyeans, lizards, marsupials, ornithischians, and saurischians (without birds)—account for 75% of the extinctions. The obvious biological question is what did sharks, lizards, opossums, and nonavian dinosaurs have in common that made them collectively more susceptible to extinction, especially from the environmental consequences of a large bolide impact? Other statistically significant patterns emerge as well (**Figure 1**). Chances of survival were much lower if you were a large, terrestrial, amniotic endotherm rather than a small, freshwater, anamniotic ectotherm. None of the myriad of Dante's inferno events (e.g., sharp temperature increase, sharp temperature decrease, tsunamis, hurricanes, global wildfires, and acid rain) alleged or modeled to have occurred in the wake of this impact have been able to explain these curious extinction and survival patterns. In fact, several of these so-called inferno events have been discounted completely (e.g., acid rain, because most aquatic species survive), or are regarded as unlikely (e.g., sharp temperature decrease, because ectotherms do well) when tested against the known K/T fossil record.

Yet Another Dante's Inferno

A recently proposed postimpact environmental scenario may be the first that shows some broad agreement with the vertebrate fossil record (Robertson *et al.*, 2004). This scenario postulates that, following the K/T impact, ejecta reentering the atmosphere created an intense blanket of infrared radiation that covered the planet's entire surface. Such a pulse of intense thermal radiation would kill any unsheltered organisms. The radiation scenario predicts that aquatic organisms would have been sheltered from the searing thermal pulse by a layer of water. This might explain the survival of most aquatic species. Terrestrial forms, however, would have suffered greatly unless they found shelter underground or in some sort of natural cavity. At first blush, this mechanism appears to explain why typically larger dinosaurian species faired more poorly than did some mammals. It may not actually be large size itself that was being selected against. Dinosaurs grew from a small size after hatching from eggs and some may even have been small enough after leaving any parental care to have survived in some sort of cavity. But in such cases, there would have been enough survivors to form a viable postimpact population to ensure species survival. For the smaller mammals, it is given that much larger population sizes would have predominated. Indeed, their numbers in the fossil record indicate this. By virtue of these larger population sizes, coupled with small physical size, more individuals may have been able to shelter in some sort of cavity or underground, and then emerge to reassemble their population.

This is all well and good, but how does the infrared radiation scenario explain the fact that placental mammals faired much better than marsupial mammals in North America, while in South America both groups radiated in the Early Tertiary? Similarly, under this radiation scenario, why would the clearly aquatic sharks disappear completely? In the marine realm, this scenario becomes even more unsatisfactory as it fails to account for the distinction between victims as disparate as nannoplankton, planktonic foraminifera, ammonites, and marine reptiles and survivors as disparate as benthic foraminifera, sponges, corals, lophophorates, and echinoderms. Many of these groups exhibit

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00032-0

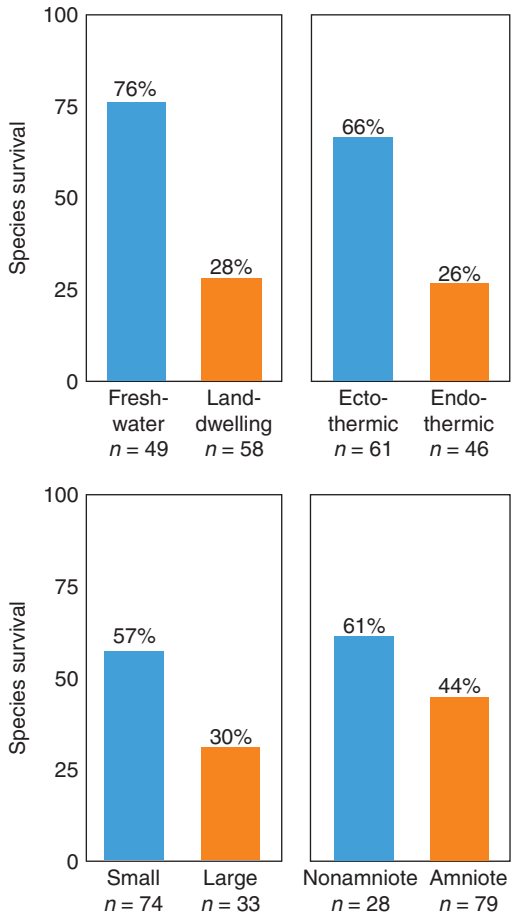


Figure 1 Differential patterns of vertebrate species survival at the K/T boundary (Archibald and Fastovsky, 2004). χ^2 tests of comparisons, except possibly the last, are significant (freshwater versus land-dwelling comparison and ectothermic versus endothermic comparison, $P \ll 0.005$, small versus large comparison, $0.010 < P < 0.025$, nonamniote versus amniote comparison $0.150 < P < 0.100$). Reproduced from Archibald JD and Fastovsky DE (2004) Dinosaur extinction. In: Weishampel DB, Dodson P, and Osmolska H (eds.) *The Dinosauria*, 2nd edn., pp. 672–684. Berkeley: California University Press.

patterns of species-richness turnover that are not consistent with a single-event scenario located at the very end of the Cretaceous. Indeed, even the timing of impact-related event scenarios has recently been rendered much more complicated by the discovery of evidence that the Chicxulub impact may have predated the K/T boundary by as much as 300 ky.

Back on land, there are still other unresolved questions. The radiation scenario predicts that heightened infrared flux would have caused global wildfires. Whether this is a necessary correlate remains to be seen, but others have argued that the physical evidence does not support the occurrence of a global wildfire at the K/T boundary. Thus, although it is a promising mechanism, this latest Dante's inferno scenario requires much further study before postimpact reentry infrared radiation can be accepted as either a partial, or the predominant, mass-extinction killing mechanism.

Multiple Causes

Other events that clearly occurred at the K/T boundary also fit the known pattern of vertebrate extinction in western North America, notably the loss of most epicontinental seas from the middle of the continent shortly before the K/T boundary. This was the greatest such loss in the past 250 My of Earth history. How it may have contributed to the K/T extinctions is not always straightforward, with one obvious exception. Shark species that frequented the rivers of western North America would have followed the regressing seas. Thus, their absence in the earliest Paleocene of this region is no great mystery. Different shark species made a brief reappearance when the seas once again lapped interior North American shores for a short time in the Early to Middle Paleocene. Aquatic species, except the fore-mentioned sharks, generally did well across the K/T boundary as the freshwater systems expanded in the wake of the sea's retreat.

But what of the marsupials? One possibility is that the earliest archaic ungulates (or condylarths) that definitely appear in North America in the earliest Paleocene (but probably were around by the latest Cretaceous) may have out-competed the marsupials in North America. The geographic origin for these archaic ungulates is uncertain. More and more evidence, however, points to Asia. As sea level fell, archaic ungulates presumably crossed the perennially appearing Bering Land Bridge. In South America, instead of replacing the marsupials, local ungulate lineages shifted more and more to herbivory while the marsupials remained omnivores or became more carnivorous.

Latest Cretaceous dinosaurs, which are only known in any numbers from low coastal plains, would have decreased in population size as the marine regression progressed. Such habitat loss is similar to that caused by human disturbance. It is doubtful that this alone would have been sufficient to cause the demise of dinosaurs, but we do know that with this sea level regression, global environmental change caused by increased volcanism from the Deccan Traps, and the terminal Cretaceous impact, the stresses on these creatures would have been great.

Long-term Trends in Dinosaur Diversity

The events discussed thus far occurred over hundreds of thousands of years in the case of the multiple-cause hypotheses (marine regression, habitat fragmentation, volcanism, and impact), or possibly over as short a span as hours–years in the case of an impact. The problem has been that to examine these competing views more thoroughly, we need a longer record stretching back millions, if not tens of millions, of years. Nowhere is this more keenly felt than in the question of what was happening to the dinosaurs during the close of the Cretaceous. The two components of this issue are: (i) What happened during the last 10 or so million years of the Cretaceous? and (ii) What happened during the last million years or less? Although among the rarest of almost any vertebrate species in the Late Cretaceous, dinosaurs have become the poster-children for this extinction. Not surprisingly then, this group has received most of the scientific attention as well. Because its stratigraphic record is far more complete through

this geological interval, once again, most work centers on western North America.

For at least the past 25 years, the general view was that the 75-million-year-old, Late Campanian beds in western North America held the richest known dinosaur fauna and that there was a substantial taxonomic decline by about 66 Ma, just before the terminal Cretaceous extinctions. There were a few who argued that this was the result of more poorly known latest Cretaceous dinosaur faunas, but the vast majority of studies found this to be a true decline in taxonomic diversity in the waning 10 My of the Cretaceous, something on the order of 40%.

A recent study has attempted to resuscitate this issue within the context of examining dinosaur diversity throughout the Mesozoic Era, in which the authors concluded that their data did not support a claim that dinosaur richness was decreasing toward extinction during the last 1 My before the K/T boundary (Fastovsky *et al.*, 2004). Separate from the issue of whether dinosaurs were decreasing toward extinction, the question at hand is how one can conclude dinosaurs were not declining in taxonomic diversity during the last 10 My of the Cretaceous. Even in such studies the results show a decline of genera going from the Late Campanian (about 75 Ma) into the Late Maastrichtian (about 65 Ma). These patterns of decline have been attempted to be explained by rarefaction analysis. Fortunately, a newly published compendium of dinosaur distribution has allowed tabulations of taxa for succeeding intervals of time.

Table 1 is a generic compilation from this new compendium. Comparing taxa at the generic level rather than at the species level can be problematic, but in this case most of the genera are either monotypic or only one species is known from a given fauna. Only genera identified without qualification are included in **Table 1**. Similarly, only localities were used for which unambiguous age ranges were provided. The compilation in **Table 1** is based on standard geological divisions of Early and Late Campanian, and Early and Late Maastrichtian.

Both the Early Campanian and Early Maastrichtian have a quite low diversity, which are almost certainly artifactual. This is supported by the fact that five genera (*Avisaurus*, *Leptoceratops*, *Pachycephalosaurus*, *Pentaceratop*, and *Troodon*) reported from the Late Campanian and Late Maastrichtian are not identified from the intervening Early Maastrichtian. The Late Campanian and Late Maastrichtian of North America are much better sampled intervals, with an obvious decline from 48 to 32 genera—a 33% drop. This is the case even though there are four more localities and 27 more repeated generic samplings for the North American Late Maastrichtian compared with the Late Campanian.

Although the drop in generic richness is clear, one could argue that such broad-brush geographic comparisons are not ecologically meaningful. To examine this concern, one can compare more clearly delineated faunas rather than make comparisons across the board (**Table 2**). For the Late Campanian, the richest dinosaurian fauna is from the Dinosaur Park Formation, Alberta, which includes 31 genera and 38 to 42 species. A recent reassessment of this fauna's age brackets it between about 76.5 and 74.2 Ma within Dinosaur Provincial Park, Alberta, with the actual span of time represented by this dinosaur fauna being somewhat shorter, by about 2 My. For the Late Maastrichtian, the richest fauna is from the Lance Formation, Wyoming, which contains 20 genera and 21 to 22 species. Although ages are not well known for the Lance Formation, the nearby biostratigraphically comparable Hell Creek Formation in Montana spans about the last 1.7 My of the Cretaceous. Comparing these figures shows that, although the Dinosaur Park and Lance Formations sample a similar magnitude of time, there is a 35% generic decline, and up to a 50% species decline between the Dinosaur Park (Late Campanian) and the Lance (Late Maastrichtian) faunas.

What is of particular ecological interest is that this decline is not especially high among the rarer carnivorous theropods, but rather among the more common hadrosaurid and ceratopsian genera. From six hadrosaurid and six ceratopsian genera in the Dinosaur Park fauna, we see a decline to only one hadrosaurid and four ceratopsian genera in the Lance fauna (**Table 2**). Comparable changes have been reported for the similarly aged Hell Creek Formation, but they are less well known for the Scollard Formation, Alberta. Thus for all the cases known, the percent generic and species declines are still over 30 and near 50, respectively, during the last 10 My of the Cretaceous.

Short-term Trends in Dinosaur Diversity

The question of what happened to vertebrates—and especially to dinosaurs—during the last million years or so of the Cretaceous is a less tractable problem. Unlike the broader question of patterns characteristic of the last 10 My of the Cretaceous, in tracking a much finer scale pattern we are now asking much more of the fossil record. The first attempt at quantification of this question counted the number of dinosaur teeth per metric ton of sediment recovered from screen washing samples collected in eastern Montana. Results showed a gradual decrease of dinosaurs leading up to the K/T boundary. As it turned out, all but the lowest sample were from the earliest Paleocene. Consequently, the dinosaur teeth were largely zombie taxa (i.e., taxa reworked from latest Cretaceous sediments). In 1992, another study examined surface-collected dinosaur material within the lower, middle, and upper third of the Hell Creek Formation, North Dakota. No changes in the dinosaur faunas were found throughout the section. Unfortunately, as this study was conducted at the relatively crude taxonomic level of family, a 40% loss of dinosaur species could have occurred without being detected.

In 2002, the results of a study conducted largely by an amateur paleontologist provided the first well-documented sampling of vertebrates through the entire thickness of the Hell Creek Formation in North and South Dakota. Despite a somewhat coarse-level sampling scheme, this study did not find any obvious changes in taxonomic diversity throughout the formation. These results have been used by some as evidence for abrupt dinosaur extinction. In fact, neither the stratigraphic nor the taxonomic levels of resolution employed in these studies are fine enough to address, let alone answer, this question. For example, this study reported that there are 2233 dinosaurian specimens representing 14 taxa. In the 3 m interval below the K/T boundary, there are only 26

Table 1 Dinosaur generic counts through the Campanian and Maastrichtian of North America[a,b] (Weishampel *et al.*, 2004)

E. Campanian *7 localities*		*L. Campanian* *14 localities*		*E. Maastrichtian* *3 localities*		*L. Maastrichtian* *18 localities*	
Hesperornis	4	*Achelousaurus*	1	*Albertosaurus*	2	*Alamosaurus*	7
Hadrosaurus	2	*Albertosaurus*	1	*Anchiceratops*	2	*Albertosaurus*	2
Ricardoestesia	1	*Anchiceratops*	1	*Arrhinoceratops*	1	*Ankylosaurus*	4
Troodon	1	*Apatornis*	1	*Aublysodon*	1	*Avisaurus*	1
		Aublysodon	2	*Caenagnathus*	1	*Bugenasaura*	2
4 genera	8	*Avaceratops*	1	*Chirostenotes*	1	*Caenagnathus*	1
		Avimimus	1	*Daspletosaurus*	1	*Chirostenotes*	2
		Avisaurus[b]	1	*Dromaeosaurus*	2	*Diceratops*	1
		Bambiraptor	1	*Edmontonia*	3	*Dromaeosaurus*	7
		Baptornis	2	*Edmontosaurus*	3	*Edmontonia*	6
		Brachyceratops	1	*Euoplocephalus*	1	*Edmontosaurus*	9
		Brachylophosaurus	3	*Hypacrosaurus*	1	*Leptoceratops*	3
		Centrosaurus	2	*Maiasaura*	1	*Montanoceratops*	1
		Chasmosaurus	1	*Montanoceratops*	2	*Nanotyrannus*	1
		Chirostenotes	2	*Ornithomimus*	1	*Ornithomimus*	5
		Coniornis	1	*Pachyrhinosaurus*	2	*Pachycephalosaurus*	4
		Corythosaurus	1	*Panoplosaurus*	1	*Pachyrhinosaurus*	1
		Daspletosaurus	3	*Parksosaurus*	1	*Palintropus*	1
		Dromaeosaurus	4	*Ricardoestesia*	1	*Parksosaurus*	1
		Dryptosaurus	1	*Saurolophus*	1	*Pentaceratops*	1
		Edmontonia	4	*Saurornitholestes*	1	*Potamornis*	1
		Einiosaurus	1	*Stegoceras*	1	*Ricardoestesia*	3
		Euoplocephalus	3	*Struthiomimus*	1	*Saurornitholestes*	6
		Gorgosaurus	2			*Sphaerotholus*	1
		Gravitholus	1	23 genera	32	*Stegoceras*	2
		Gryposaurus	2			*Struthiomimus*	1
		Hadrosaurus	1			*Stygimoloch*	4
		Hesperornis	2			*Thescelosaurus*	6
		Hypacrosaurus	2			*Torosaurus*	8
		Lambeosaurus	1			*Triceratops*	9
		Leptoceratops[b]	2			*Troodon*	6
		Maiasaura	1			*Tyrannosaurus*	12
		Monoclonius	2				
		Montanoceratops	1			32 genera	119
		Ornatotholus	1				
		Ornithomimus	3				
		Orodromeus	3				
		Pachycephalosaurus[b]	2				
		Panoplosaurus	1				
		Parasaurolophus	2				
		Pentaceratops[b]	1				
		Prosaurolophus	2				
		Ricardoestesia	4				
		Saurornitholestes	6				
		Stegoceras	4				
		Struthiomimus	1				
		Styracosaurus	1				
		Troodon[b]	5				
		48 genera	92				

[a]Number following each genus is the number of localities at which a particular genus occurs.

[b]Taxa occur in L. Campanian and L. Masstrichtian, but not intervening E. Maastrichtian.

Source: Reproduced from Weishampel DB, Barrett PM, Coria RA, *et al.* (2004) Dinosaur distribution. In: Weishampel DB, Dodson P, and Osmolska H (eds.) *The Dinosauria*, 2nd edn., pp. 517–606. Berkeley: California University Press.

dinosaurian specimens representing just three taxa. Within the upper 5 m of the Hell Creek only two dinosaurs (*Ricardoestesia* and *Tyrannosaurus*) were identified at the generic level. The next higher occurrences of dinosaur genera occur 8 m or lower within the Hell Creek Formation. What of the other 18 or so dinosaur genera that are known to have occurred in the Hell Creek Formation? Such a record is obviously much too poor to say anything about whether the local dinosaur extinction was instantaneous or gradual on a scale of tens to hundreds of thousands of years.

Table 2 Generic and species counts of dinosaurs for Dinosaur Park Formation (Late Campanian) and Lance Formation (Late Maastrichtian) of North America (Weishampel *et al.*, 2004)

Dinosaur park formation		*Lance formation*	
Genus	*No. of species*	*Genus*	*No. of species*
Theropoda		Theropoda	
Avimimus	1	*Albertosaurus*	1
Baptornis	1	*Dromaeosaurus*	1
Chirostenotes	2	*Palintropus*	1
Daspletosaurus	1	*Ornithomimus*	1
Dromaeosaurus	1–2	*Potamornis*	1
Gorgosaurus	1	*Ricardoestesia*	1
Ornithomimus	1	*Saurornitholestes*	1
Ricardoestesia	2	*Troodon*	1
Saurornitholestes	1	*Tyrannosaurus*	1
Struthiomimus	1		
Troodon	1		
Ankylosauria		Ankylosauria	
Edmontonia	2	*Ankylosaurus*	1
Euoplocephalus	1	*Edmontonia*	1
Panoplosaurus	1		
Euornithopoda		Euornithopoda	
Orodromeus	1	*Bugenasaura*	1
		Thescelosaurus	1
Hadrosauridae		Hadrosauridae	
Brachylophosaurus	1	*Edmontosaurus*	2
Corythosaurus	1		
Gryposaurus	2		
Lambeosaurus	2		
Parasaurolophus	1		
Prosaurolophus	1		
Pachycephalosauria		Pachycephalosauria	
Gravitholus	1	*Pachycephalosaurus*	1
Ornatotholus	1	*Stygimoloch*	1
Pachycephalosaurus	1		
Stegoceras	1		
Ceratopsia		Ceratopsia	
Anchiceratops	1	*Diceratops*	1
Centrosaurus	1–2	*Leptoceratops*	1
Chasmosaurus	3–4	*Torosaurus*	1
Leptoceratops	1	*Triceratops*	1–2
Monoclonius	1–2		
Styracosaurus	1	20 genera	21–22
31 genera	38–42		

Source: Reproduced from Weishampel DB, Barrett PM, Coria RA, *et al.* (2004) Dinosaur distribution. In: Weishampel DB, Dodson P, and Osmolska H (eds.) *The Dinosauria*, 2nd edn., pp. 517–606. Berkeley: California University Press.

Out with a Whimper, a Bang, or Both?

If questions of rate and taxonomic magnitude of terrestrial extinctions are to be addressed and possibly answered, a much finer scale of both stratigraphic and taxonomic sampling will be required. One study (**Figure 2**) that specifically addressed changes within the turtle community in the Hell Creek Formation during the last 1.7 My of the Cretaceous found that for turtles, as well as for mammals and plants, taxonomic richness was lower in the formation and dropped immediately before the K/T boundary. The maximum diversities in these three groups were correlated with the maximum latest Cretaceous warming and the drop in richness was correlated to a rapid drop in paleotemperature. In a detailed study of only the mammals from this same sequence, the observed change in climate was deemed insufficient to explain the extinction or

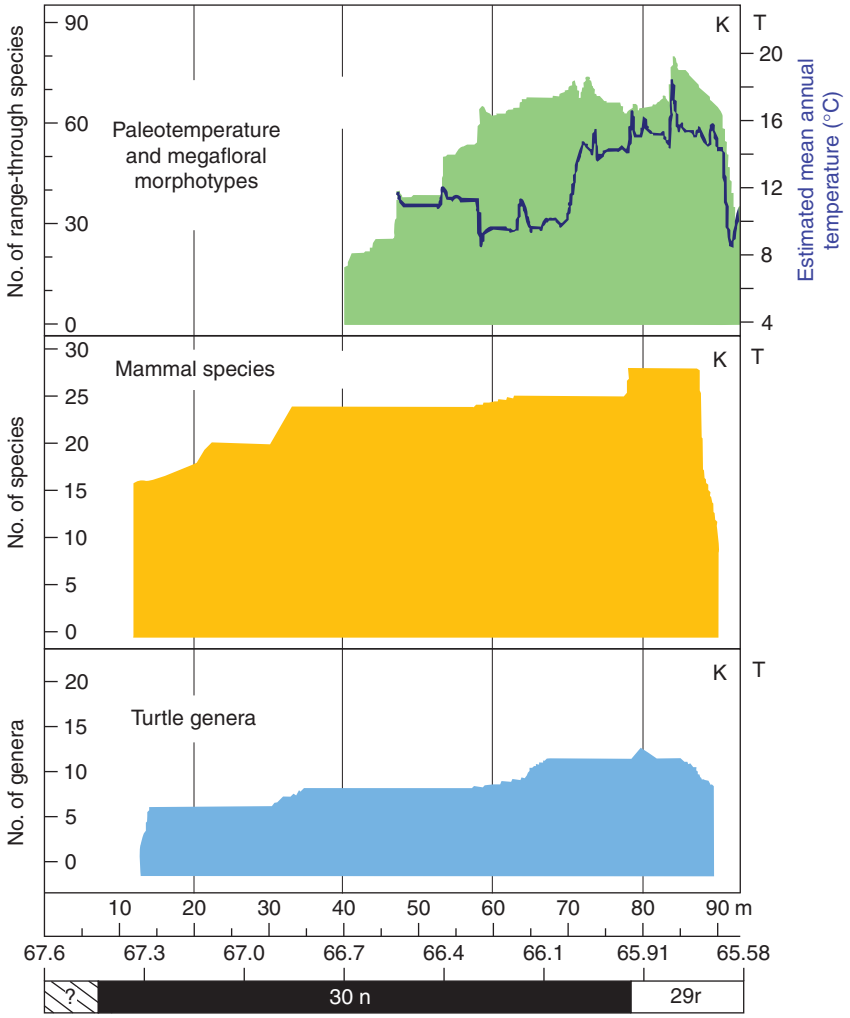


Figure 2 Compiled taxonomic richness data for plants, mammals, and turtles in the Hell Creek Formation. Independent of taxonomic level, taxonomic richness is lower in the early portion of the Late Maastrichtian and it drops immediately before the K/T boundary. The greatest taxonomic diversity is associated with maximum warmth, while the latest Maastrichtian drop is correlated with a rapid drop in paleotemperature. Reproduced from Hutchison JH, Holroyd PA, Gregory P, and Wilson GP (2004) The possible role of climate on successive turtle assemblages from the Upper Cretaceous Hell Creek Formation. *J. Vertebrate Paleontol.* 24(Suppl. 3): 73A.

pseudoextinction (disappearance resulting from speciation) of up to 27 species of mammal near or at the K/T boundary. A longer, gradual turnover pattern could be ruled out based on this analysis, but the time resolution of these data was not able to differentiate between stepwise and sudden extinction.

Until recently, the well-sampled megafloral (leaves and fruits rather than pollen) record had indicated as much as a 79% plant extinction at the K/T boundary in the northern western interior of the United States and even higher (84%) farther to the south in the United States. With much tighter stratigraphic control, however, this estimate has recently been downgraded to a minimum of 30% extinction based on palynomorphs that have higher stratigraphic but lower taxonomic resolution, compared with a maximum of 57% extinction based on megaflora that disappear within 5 m or less of the K/T boundary. This flora was interpreted then to have suffered between 30% and 57% extinction at the K/T boundary, while other floral extinctions were spread out within the 50 m of the upper part of the Hell Creek Formation. Certainly, there is a dynamic floral turnover throughout the Cretaceous part of the section, a sharp shift near the K/T boundary, and virtual stasis during the Paleocene; a pattern interpreted as being commensurate with "sudden ecosystem collapse, presumably caused by the Chicxulub impact." Once again, though, while the fact of the taxic richness decline is clear, the time interval over which this decline occurred, and its correlation to alternative sources of global environmental change, are not.

Conclusions

The newest generation of studies on the K/T and dinosaur extinctions shows that extinction levels for plants and animals, on land and in fresh water, are not as great as once thought. Instead of species extinction levels as high as 80%, newer

studies range from a possible low of 30% to a high of, at the most, 60%. Dinosaurs clearly declined in taxonomic richness during the last 10 My of the Cretaceous, at least in North America where they are best known. This is true whether broader geographic taxonomic counts or specific dinosaurian faunas are compared. But the dinosaur record remains too poor to determine what occurred in the last few tens or hundreds of thousands of years of the Cretaceous. The record is better for other terrestrial groups. Taxonomic richness for both plants, and vertebrates—at least turtles and mammals—tracks the climb to a warm maximum and then a sudden plunge in temperature just before the K/T boundary. These records appear sufficient to eliminate a long-term, gradual pattern of extinction leading up to the K/T boundary, but whether the extinctions were stepwise over hundreds of thousands of years, or sudden over as little as a few years cannot be determined, given the record currently available. Moreover, given the certain knowledge that a number of major environmental changes were operating in this latest Maastrichtian interval, singling out of any one cause is more a matter of speculation than fact.

See also: Extinction, Causes of. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of

References

Archibald JD and Fastovsky DE (2004) Dinosaur extinction. In: Weishampel DB, Dodson P, and Osmolska H (eds.) *The Dinosauria*, 2nd edn., pp. 672–684. Berkeley: California University Press.

Fastovsky DE, Huang Y, Hsu J, Martin-McNaughton J, Sheehan PM, and Weishampel DB (2004) The shape of Mesozoic dinosaur richness. *Geology* 32: 877–880.

Hutchison JH, Holroyd PA, Gregory P, and Wilson GP (2004) The possible role of climate on successive turtle assemblages from the Upper Cretaceous Hell Creek Formation. *J. Vertebrate Paleontol.* 24(Suppl. 3): 73A.

Robertson DS, Mckenna MC, Toon OB, Hope S, and Lillegraven JA (2004) Survival in the first hours of the Cenozoic. *Geol. Soc. Am. Bull.* 116: 760–768.

Weishampel DB, Barrett PM, Coria RA, *et al.* (2004) Dinosaur distribution. In: Weishampel DB, Dodson P, and Osmolska H (eds.) *The Dinosauria*, 2nd edn., pp. 517–606. Berkeley: California University Press.

Diseases, Conservation and

Barbara A Han and Sonia Altizer, University of Georgia, Athens, GA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Definitive host For parasites with complex lifecycles, this is the host in which the parasite reproduces sexually (also called the *primary host*).

Epidemic Characterizes a sudden increase in parasite prevalence or intensity beyond that which is normally present.

Helminth One of several classes of parasitic worms: nematodes, cestodes, trematodes (monogeneans and digeneans), and acanthocephalans.

Intensity The number of parasites per infected host; a related measure, abundance, refers to the average parasite load of the entire population (including uninfected individuals).

Intermediate host For parasites with complex lifecycles, earlier parts of their life cycle are completed within intermediate hosts (also called the *secondary host*).

Prevalence The proportion of hosts in a population that are infected or diseased.

Reservoir host A host organism that can carry and transmit a parasite or pathogen to other host species while suffering little to no negative effects of disease.

R_0 The basic reproductive ratio of a parasite or pathogen; for a microparasite, this describes the number of new infections generated by a single infected host entering an entirely susceptible population; for a macroparasite, R_0 is the number of adult offspring produced by a single adult parasite over its lifespan.

Vector An animal that transmits parasites among definitive hosts; for example, mosquitoes are vectors of malaria.

Virulence Disease-induced mortality rate, or the harm caused by parasites to individual hosts.

Types of Parasites and Causes of Disease

Parasites are organisms that live in or on and obtain resources from a host, usually to the host's detriment. The terms parasite, pathogen and disease are often used interchangeably; however, disease refers to the pathogenic condition of a host, including the signs and symptoms of infection, whereas parasites and pathogens are the disease-causing agents. Infectious diseases can be caused either by microparasites (such as viruses, bacteria, protozoa, and fungi) or by larger macroparasites, including worms and arthropods. The distinction between macro- and microparasites is particularly useful to ecologists and epidemiologists because these groups differ in the degree of within-host replication, their ability to generate a lasting adaptive immune response, and how they are quantified. They also differ in their impacts on host populations (**Table 1**).

Disease outbreaks require the presence of a susceptible host population, an infectious pathogen, and favorable environmental conditions; these three factors are often referred to as the disease triangle (Wobeser, 2007). It follows that disease prevalence or infection intensity can increase when changing environmental conditions favor pathogen transmission or host susceptibility. In many cases, these environmental modifications are caused by human activities that lead to host crowding, habitat degradation, and host stress. Other environmental changes can cause novel pathogen introductions or enhance opportunities for cross-species transmission through changes in the composition of ecological communities or shifts in the geographic distribution of a host, pathogen, or vector species.

Basic Epidemiological Principles

Since the pioneering work of Anderson and May in the late 1970s, ecologists have shown increasing interest in the spread and impacts of parasites at the population and community levels. A general understanding of host–pathogen population ecology can illuminate problems in conservation biology ranging from detecting disease threats for endangered species to predicting how human impacts on landscapes will affect pathogen invasion and persistence. Epidemiologists quantify and model the spread of infectious diseases over time and space to identify parameters that influence their prevalence and population-level effects. Models are used in epidemiology to detail how processes operating at the level of individuals (such as infection, recovery, and death) translate into population-level patterns such as changes in the numbers of susceptible and infected hosts. Epidemiological models differ from models of other antagonistic interactions (such as between predators and prey) because pathogens and parasites do not necessarily kill their hosts. Also, a single host can be infected by many parasites at once, and recovered hosts can develop a long-term immunity to reinfection by some pathogens. It is useful to address models designed for micro- versus macroparasites separately to identify basic principles that govern parasite spread and quantify their effects on host populations.

Microparasites

Mathematical models for microparasites divide the host population into categories that reflect their stage of exposure,

Table 1 General characteristics, examples, and ecological properties of micro- and macroparasites

	Microparasites	*Macroparasites*
Representative taxa	Viruses, bacteria, protozoa, fungi, microsporidians	Helminths (e.g., nematodes, cestodes, and acanthocephalans), arthropods (e.g., mites, ticks, and lice)
Size and reproduction	Small, unicellular, short generation times and rapid replication within individual hosts	Large, multicellular, longer generation times, usually no direct replication within individual hosts
Transmission of infectious stages	Transmission *via* direct contact (e.g., venereal and vertical), vectors, or contaminated air/soil/water	Complex life cycles and intermediate hosts are common; can also be transmitted by vectors, close contact, or ingestion of material contaminated by feces
Effects on adaptive host immunity	In many cases, temporary or lasting host immunity develops quickly and protects against reinfection	Some acquired immunity, but antigenic diversity of parasites usually too high for host to mount lasting adaptive immune response
Effects on host fitness	Disease can be acute or chronic, may have strong effects on host survival or fecundity	Depends on the number of parasites within the host (can affect mortality or fecundity, usually chronic infection with cumulative, sublethal effects)
Quantification in host populations	Prevalence, seroprevalence, incidence	Prevalence, intensity, degree of aggregation in individual hosts
Frequency of epidemics	Common	Rare

including susceptible (S), infected (I), and recovered/immune (R) classes (**Figure 1**). These compartment models track changes in the number of hosts within each category, but ignore the number of parasites within each host. SIR models are commonly used in modeling directly transmitted microparasites of vertebrates, and have been developed and analyzed extensively (e.g., Anderson and May, 1991). The model shown in **Figure 1** makes many assumptions that can be relaxed. For example, hosts are uninfected at birth, infection increases host mortality but does not affect host fecundity, and host populations are large enough that stochastic processes (random events that affect populations in unforeseeable ways) can be ignored. For pathogens for which hosts do not acquire immunity to reinfection (e.g., many plant and insect pathogens and vertebrate diseases such as tuberculosis and brucellosis), the recovered/immune class (R) is eliminated, and the equations simplify to a SI or SIS model.

The basic SIR model gives rise to several key principles that have consequences for the spread of infectious diseases in wild populations. Of fundamental importance is a value called R_0, the basic reproductive ratio of the pathogen. R_0 describes the initial rate of pathogen increase in a previously unexposed host population. This parameter is estimated by multiplying the expected number of new infections from a single infected host by the average duration of infectiousness. For the SIR model in **Figure 1**,

$$R_0 = \frac{\beta S}{\alpha + b + \nu} \quad (1)$$

R_0 must exceed 1.0 for a pathogen to invade and spread. The form of eqn (1) suggests that pathogens with relatively high transmission (β), low virulence (α), and low host recovery (ν) are most likely to establish in host populations. For pathogens that can successfully invade, a common dynamical outcome is an "epidemic curve" (**Figure 1**, orange line), whereby the number of infected hosts firsts increases, then decreases as the pathogen spreads through the population. At the same time, the numbers of susceptible hosts decline as they transition to the infected class, and the number of recovered hosts gradually increases. The general shape of the epidemic curve produced by simple compartment models is mirrored by many real-world epidemics including influenza in humans and phocine distemper in harbor seals (reviewed in Anderson and May, 1991; Hudson *et al.*, 2002).

The SIR model shown in **Figure 1** is useful for pathogens with density-dependent transmission, a process in which transmission (measured by the rate new susceptibles are infected per unit time) increases directly with host population density. Specifically, for density-dependent transmission there exists a threshold density, N_T, of hosts below which the pathogen cannot persist ($R_0 < 1$) in a host population. Assuming that the population is large and homogeneously mixed, this value is

$$N_T = \frac{\alpha + b + \nu}{\beta} \quad (2)$$

Pathogens that are highly virulent (high α) or have lower transmission (low β) are likely to require much higher host densities for establishment than those that are highly transmissible and relatively benign.

For some pathogens, transmission can remain relatively constant over a range of host densities, a process termed frequency-dependent transmission. A key result of frequency-dependent transmission is that there is no threshold density for pathogen invasion (so in theory, such pathogens can persist at arbitrarily low host densities). Pathogens spread by direct contact (such as touching), aerosol droplets (coughing or sneezing) or indirect contact (ingesting fecally contaminated food or water) are expected to show density-dependent transmission. By comparison, transmission of sexually transmitted diseases and some vector-borne diseases is thought to be frequency dependent. Importantly, field and experimental studies of pathogens affecting mice, voles, frogs, ladybird beetles, and wildflowers have shown that the transmission of most pathogens probably falls between these two extremes.

Analysis of simple microparasite models can generate important insights for considering pathogen risks to wild or captive populations. For example, models suggest that the population-level impacts of an infectious disease will depend

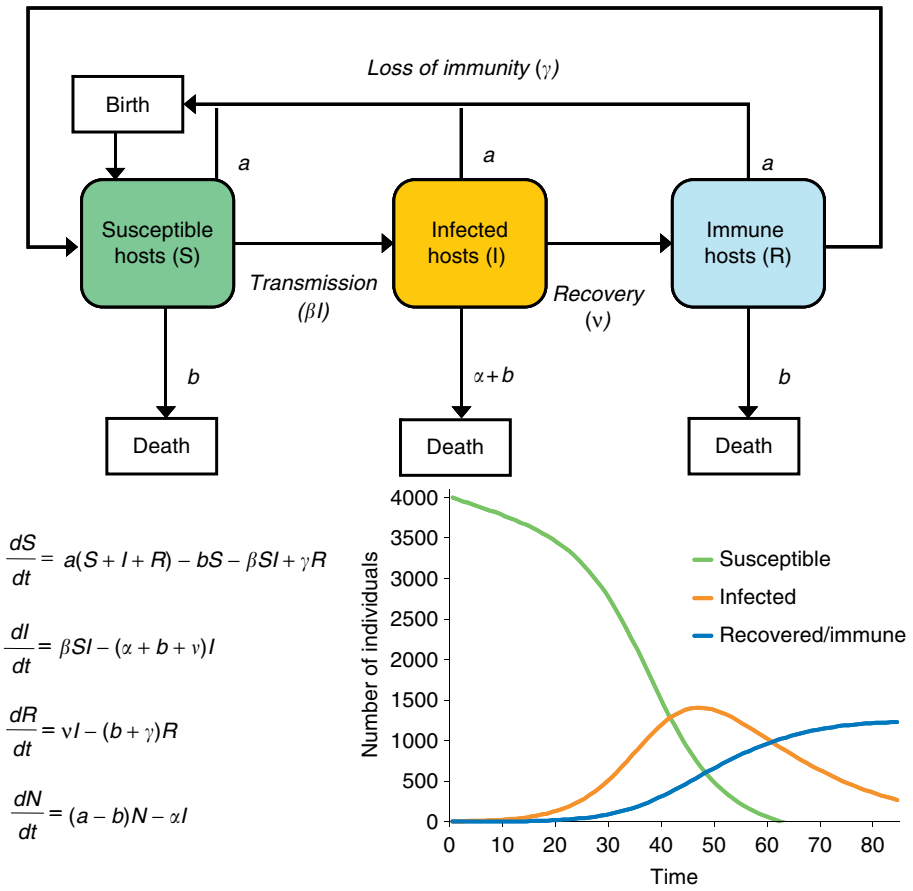


Figure 1 SIR compartment model for directly transmitted microparasitic disease. Diagram shows a population containing susceptible hosts (S), infected hosts that can transmit the parasite to others (I), and recovered or immune hosts that are no longer infected (R). Total host population size ($N = S + I + R$. Susceptible hosts arise from birth or immunity loss at per capita rates a and γ, respectively. Individuals leave the susceptible class through natural mortality (rate b) or by acquiring the parasite at rate β after encountering an infected host. Hosts leave the infected category through natural death or disease-induced mortality (rates b and α, respectively) or through recovery (rate ν) to an immune state. Differential equations that track changes in the numbers of susceptible, infected and recovered hosts are shown in the bottom left panel. Representative dynamics (changes in numbers over time) are graphed in the bottom right panel for the following parameter values: $a = 0.1$, $b = 0.05$, $\beta = 0.0006$, $\alpha = 0.3$, $\nu = 0.3$, $\gamma = 0.005$.

on several factors, including pathogen effects on individual host fitness. In general, for pathogens that lower host survival, those with intermediate virulence should have the largest negative impacts on host populations. If pathogens are too virulent, infected hosts will die before they cause many new infections, whereas nonvirulent pathogens can become highly prevalent but have minimal population-level impacts. For similar reasons, models also predict that sterilizing pathogens (those that reduce fecundity) can cause greater reductions in host population size than those that reduce survival (Anderson and May, 1991). Models can generate important predictions for pathogen control, including effects of vaccination or culling on the probability of pathogen eradication (*see* Intervention Methods).

Finally, the spread and impacts of pathogens will depend on a number of ecological and genetic heterogeneities. Unlike the simple homogeneous populations described in **Figure 1**, natural host populations tend to be stratified by age, sex, social rank, or clumped spatially due to fluctuating resources or habitat fragmentation. This heterogeneity can have large impacts on pathogen transmission and, as a result, on disease control efforts. Heterogeneity also arises when pathogens can be transmitted between several different host species. For these multihost pathogens, some host species are reservoirs or amplifiers (their presence increases transmission of the pathogen), whereas others may lower transmission or be unusually susceptible to pathogen-induced mortality. The identity and abundance of different host species has been shown to be important in the transmission of plant diseases such as sudden oak death and animal diseases such as brucellosis, Lyme disease, and chytridiomycosis (*see* Habitat Management).

Macroparasites

In contrast to microparasites, macroparasites typically cause persistent infections (**Table 1**), in large part because the host's immune response is often incomplete or short-lived.

Ecologists keep track of the number of adult macroparasites per host because the outcomes of macroparasite infections (e.g., the survival and fecundity of both hosts and macroparasites) depend strongly on infection intensity. The distribution of parasites among hosts almost always shows evidence for aggregation or clumping which means that most hosts have few or no parasites, and a few hosts have many.

Fundamental macroparasite models developed by Anderson and May (1991) have been modified by Dobson and Hudson (1992) and others to consider the presence of free-living infective stages, arrested parasite development, and parasites with complex life cycles (which can include multiple intermediate hosts and a definitive host). These models typically track the density of the entire host population (H), the abundance of adult parasites within hosts (P), and the number of free-living parasite stages (W) in the external environment (**Figure 2**). The model further assumes that parasites are aggregated within hosts according to the negative binomial distribution, where the degree of aggregation varies inversely with k. As indicated by the equations in **Figure 2**, the mortality of adult parasites is affected by within-host clustering, with parasite mortality increasing when k is small (and parasites are highly aggregated).

The basic reproductive ratio of macroparasites is the product of the mean number of new infections produced by a single adult parasite and the average life expectancy of adult and larval stages:

$$R_0 = \frac{\beta\lambda H}{(\mu + b + \alpha)(\gamma + \beta H)} \qquad (3)$$

As with microparasites, eqn (3) must exceed 1.0 for the parasite to establish when rare. Therefore, parasite invasion

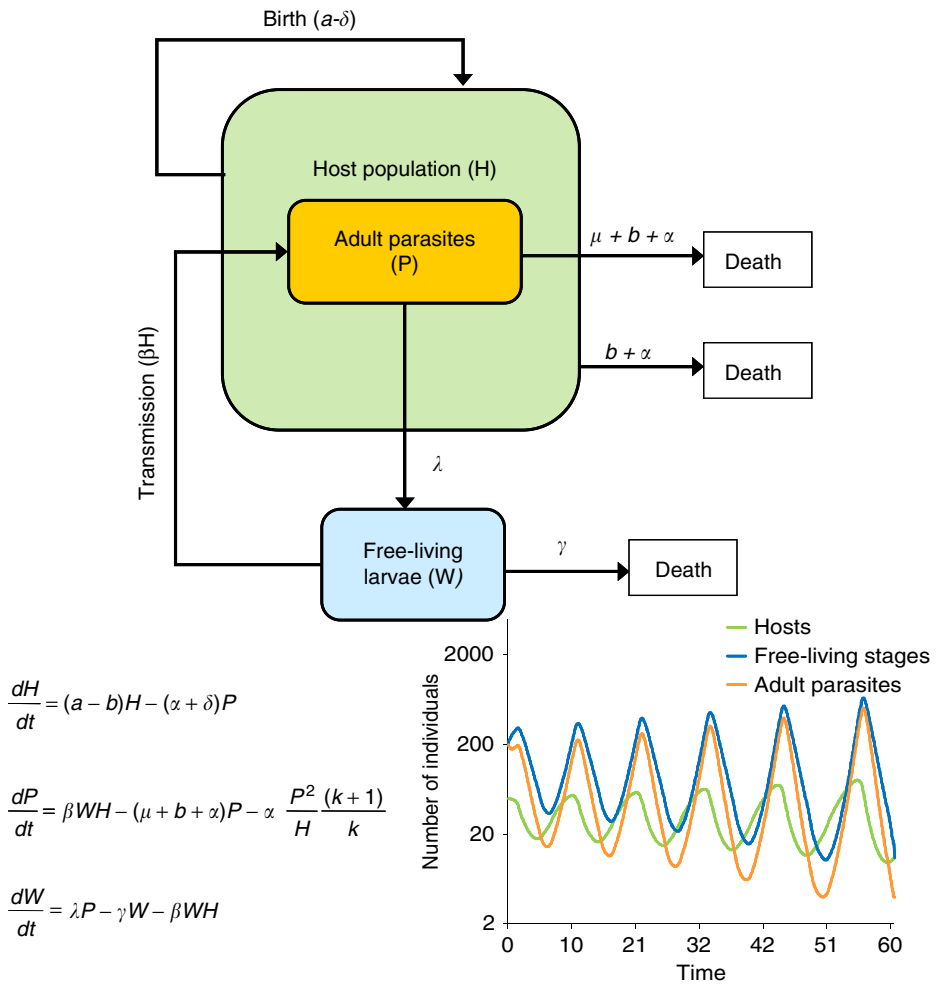


Figure 2 Box diagram of host and parasite life cycle for macroparasitic infections, showing a host population of size H, collectively harboring an adult parasite population of size P. As in the microparasite model (**Figure 1**), per capita host birth and death rates are denoted by a and b, respectively. δ and α are the per capita host sterility and mortality rates induced by each individual parasite and assume that overall host death rate increases linearly with parasite burden. Adult parasites give birth to free-living infective stages at rate λ and die at rates μ (background mortality), b (host mortality), and α (disease-induced mortality). Thus, the model assumes that when hosts die, so do their parasites. Free-living egg and larval stages die at rate γ, and transmission occurs when hosts eat these stages at rate β, thus giving rise to new adult infections. The aggregation parameter k corresponds to the negative binomial distribution, such that low k values indicate highly aggregated parasite distributions (in which most parasites are harbored by a few hosts). Representative dynamics (changes in numbers over time) are graphed in the bottom right panel for the following parameter values: $a = 0.9$, $b = 0.55$, $\beta = 0.1$, $\alpha = 0.03$, $\delta = 0.06$, $\lambda = 11$, $\mu = 0.4$, $\gamma = 1$, $k = 0.315$.

and persistence depend strongly on the rate of production of eggs or larval stages (λ), the rate at which parasites are consumed by hosts (β), and the survival of free-living infective stages ($1-\mu$). The macroparasite model shown in **Figure 2** is also associated with a threshold host population necessary to sustain infection. Because larval macroparasites often have long-lived resistant stages and adult worms can live for years within their hosts, many macroparasites can persist at lower host population densities than directly transmitted microparasites.

The effects of macroparasite virulence on host population thresholds and their ability to regulate host populations depend on the degree of parasite aggregation. Parasites of intermediate virulence (α) will depress host density more than those of low or high virulence, and parasite impacts will be maximized when aggregation is low (so that parasites are distributed across a greater number of host individuals). Importantly, highly aggregated parasite distributions tend to stabilize host–macroparasite interactions, whereas random or regular parasite distributions tend to destabilize them, leading to population cycles in host and parasite abundance (**Figure 2**). When parasites reduce host fecundity (i.e., $\delta > 0$) this can further destabilize the host–parasite interaction and increase the probability of parasite-induced host population cycles.

Field studies also support a role for macroparasites in wildlife population dynamics, although their effects are often more subtle than the dramatic population declines seen in response to some microparasitic diseases. Perhaps the best evidence comes from a handful of field experiments where researchers treated a fraction of animals or a subset of populations with antiparasitic drugs. This approach has been useful in demonstrating impacts of nematode parasites on host survival and population sizes of feral Soay sheep, white-footed mice, and red grouse (with several examples reviewed in Hudson *et al.*, 2002). In the case of red grouse, for example, cecal nematodes are only weakly aggregated among hosts and high parasite loads cause reduced fecundity in grouse. In addition, treating 20% or more of a local population to remove parasites was sufficient to halt periodic population crashes that occurred every 4–8 years (and *see* Alternative Interventions). This work suggests that macroparasites should not be overlooked as important causes of wildlife declines, either alone or together with other factors such as food limitation or harsh environmental conditions.

Case studies of macroparasite infections further emphasize the importance of sublethal effects of parasites for infection outcomes and host–parasite population dynamics. In the case of red grouse, for example, negative effects of cecal nematodes on host breeding success (and not survival) drive population cycles in abundance over time. Moreover, grouse that carry heavy infestations of cecal nematodes are more vulnerable to predation by red foxes and raptors. These effects of parasites on host fitness would be easy to miss, but their implications for host and parasite population dynamics are extremely important. In other examples, parasites have been observed to manipulate key behaviors of hosts as diverse as ants, amphipods, and fishes, making them behave in ways that increase their risk of being consumed by a predator and thereby improving the chances that the parasite will be transmitted to its definitive host (Moore, 2002). Another subtle effect of parasitism occurs when infection by one parasite species affects host susceptibility to other pathogens. In the case of African buffalo, Jolles *et al.* (2005) showed that infections by parasitic worms were negatively associated with the probability of bovine tuberculosis infection, as might be expected if coinfected animals experience sharp declines in body condition and greater mortality. The authors used a population dynamic model to show that high mortality of coinfected hosts (as might be caused by failure of the immune system to adequately control both parasite types) qualitatively captured the observed patterns of disease in free-ranging buffalo populations. Thus, the sublethal and cumulative effects of parasites can impact host population dynamics and community-level interactions in substantial and unexpected ways.

Infectious Diseases as Threats to Biological Diversity

Introduced Parasites and Species Declines

Exotic diseases and parasites are increasingly recognized as important factors driving population declines and geographic range contractions in many organisms (**Table 2**). Owing to their high rates of spread and potentially devastating effects on host populations, a handful of pathogens are now considered the greatest threats to the survival of some endangered species. Parasitic organisms have been shown to impact host populations in a variety of ways. Most directly, disease-induced mortality can reduce host population sizes below a threshold necessary for maintenance and growth. In very small populations, differential mortality between male and female hosts can sufficiently distort sex ratios or shift host life history to affect future reproduction. This is illustrated by a facial tumor disease infecting Tasmanian devils that spreads when infectious cancer cells are transmitted through aggressive encounters, especially among older males. Mortality caused by this disease has resulted in declines of up to 50% in Tasmanian devil populations. Moreover, as older, more dominant males succumb to this disease, researchers have observed a shift in breeding phenology such that devils are reaching sexual maturity sooner and are breeding at a younger age than before the epidemic. Tasmanian devils may eventually come to rely on single, early reproductive bouts as opposed to engaging in multiple reproductive efforts over a longer time period.

A recently discovered terrestrial fungus, *Geomyces destructans*, causes White Nose Syndrome (WNS) in multiple bat species of the eastern United States (**Figure 3(a)**). Following initial reports in New York state in the winter of 2006–2007, WNS spread rapidly, devastating populations of the little brown bat (*Myotis lucifugus*) in eastern North America. Once among the most abundant bat species in North America, local population declines of *M. lucifugus* have exceeded 75%, with bat mortality reaching 100% in some hibernation caves. Perhaps one of the best-studied examples of parasite-induced population declines involves a chytrid fungus, *Batrachochytrium dendrobatidis*, which causes amphibian chytridiomycosis in hundreds of frogs, toads, and salamander species worldwide. Multiple amphibian populations from North America, Central America, South America, Australia, and several island

Table 2 Selected disease outbreaks associated with declines in natural populations

Host species	*Disease (parasitic agent)*	*Location*	*Comments*
Plants			
American chestnut (*Castanea dentata*)	Chestnut blight (*Cryphonectria parasitica*)	Eastern North America	Caused massive extinction of dominant hardwood species
Flowering dogwood (*Cornus florida*)	Anthracnose bight (*Discula destructiva*)	North America	Fungus decimated dogwood populations throughout native range
Several native plant species	Fungal dieback (*Phytophthora cinnamoni*)	Western Australia	Responsible for large-scale diebacks and permanent plant community shifts
American elm (*Ulmus americana*)	Dutch elm disease (*Ceratocystis ulmi*)	North America	Pathogen introduced from Asia, spread by bark beetles
Multiple woody species including live oaks and tanoaks	Sudden oak death (*Phytophthora ramorum*)	Western North America	Introduced from Asia; wide host range
Invertebrates			
Elkhorn coral (*Acropora palmata*)	White pox disease (*Serratia marcescens*)	Caribbean Sea	Bacterial pathogen traced to human wastewater
Oysters (*Crassostrea virginica*)	Protozoan parasite (*Perkinsus marinus*)	Atlantic coast of North America	Warming ocean temperatures facilitated parasite spread along the coast
Fish			
Rainbow trout, salmon (*Salmo spp.*)	Whirling disease (*Myxobolus cerebralis*)	Montana	Introduced with stocked fish
Aral Sea sturgeon (*Acipenser nudiventis*)	Monogenean trematode (*Nitzschia sturionis*)	Aral Sea, former USSR	Introduced with stocked Caspian sturgeon
Amphibians			
Multiple species of frogs, toads, salamanders	Digenean trematode (*Riberoia ondatrae*)	North America	Causes limb malformations during metamorphosis; linked with eutrophication
Multiple species of frogs, toads, salamanders	Saprolegniasis *Saprolegnia* spp (oomycete)	North America	Lethal to amphibian embryos; linked with increasing UV-B radiation, decreased rainfall and stocking nonnative fishes
Reptiles			
Desert tortoise (*Gopherus agassizii*)	Upper respiratory tract syndrome (*Mycoplasma agassizii*)	Mojave Desert	Introduction through released pets
Birds			
Hawaiian honeycreepers	Avian malaria (*Plasmodium relictum*)	Hawaii	Implicated in the extinction of several Hawaiian forest birds
House finches (*Carpodacus mexicanus*)	Bacterial conjunctivitis (*Mycoplasma gallisepticum*)	North America	Likely transfer from domesticated poultry
Mammals			
Harbor seals (*Phoca vitulina*)	Phocine distemper virus (*Morbillivirus*)	North Sea	Outbreaks likely initiated by contact with Harp seals
Sea otter, some wild felids, red fox	Toxoplasmosis (*Toxoplasma gondii*)	Pacific Ocean and the Pacific coast of North America	Associated with ground surface run-off contaminated by fecal matter from infected felines
Gorillas, Chimpanzees	Ebola virus (*Filoviridae*)	Multiple countries in Central and sub-Saharan Africa	Associated with severe population declines of wild primates

countries have experienced severe declines, with up to 50% of species vanishing over a matter of months in some stream communities of Central America (**Figure 4**).

In addition to causing precipitous declines of initially abundant host populations, parasites can also threaten already-endangered species with extinction. For example, critically endangered black-footed ferrets in western N. America succumbed to an outbreak of canine distemper in the late 1980s, with only 18 individuals remaining for population recovery through vaccination and captive breeding efforts. It is important to keep in mind that such heavily endangered populations are not likely to sustain most parasites in the long run simply because remaining host populations are too small. This is particularly true for parasites that specialize on a single host species and for highly virulent parasites. However, parasites that infect multiple host species (i.e., generalist parasites) often pose the greatest conservation concern because they can persist in reservoir hosts and their transmission is not limited



Figure 3 Examples of infectious diseases that have caused population declines. (a) White nose syndrome caused by the fungus *Geomyces destructans* is illustrated by white spots on the noses and ear membranes of infected little brown bats (*Myotis lucifugus*); (b) Barley yellow dwarf virus causes yellowing of infected grasses compared to healthy grass; (c) Eurasian red squirrel (*Sciurus vulgaris*) severely infected by parapox virus introduced by gray squirrels into the UK; (d) Elkhorn coral (*Acropora palmata*) infected with white pox disease (white spots) caused by the human enteric bacterium *Serratia marcescens*.

by the small population sizes of many endangered species. As a case in point, rabies is a generalist pathogen that can be maintained in domesticated dog populations; in recent decades, rabies outbreaks have caused the near-extinction of local populations of endangered carnivores in parts of Africa, including African wild dogs and Ethiopian wolves. Additionally, theoretical studies suggest that parasites with frequency dependent transmission will cause severe problems for endangered species or species with small local population sizes (de Castro and Bolker, 2005) because they can continue to drive down host populations even after they have been drastically reduced in number.

Invasive Species and Pathogen Introductions

A number of introduced or exotic pathogens have caused catastrophic declines of plant and animal species. In epidemics where new pathogens are introduced into previously unexposed host populations, the disease often progresses rapidly through immunologically naive hosts and can cause mass mortalities. One of the best known examples is the introduction of the exotic malaria parasite *Plasmodium relictum* and its mosquito vector into the Hawaiian Islands. Although most nonnative birds on Hawaii showed a combination of resistance and tolerance to this infection, the parasite was

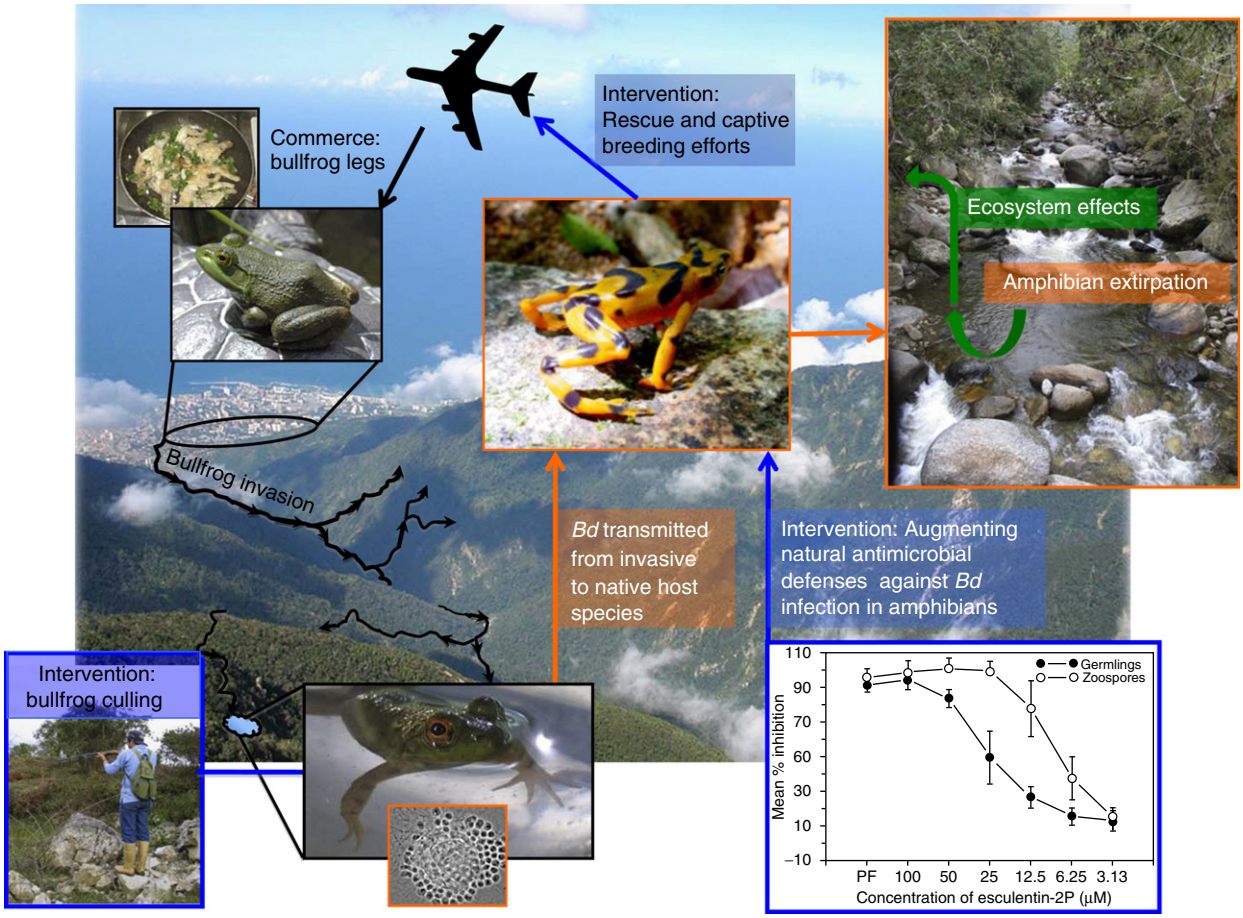


Figure 4 Amphibian chytridiomycosis can spread to native frog populations through contact with reservoir species such as the American bullfrog (*Lithobates catesbeianus*). Bullfrogs are disseminated globally to support the culinary industry for frog legs, and can be heavily infected with *Batrachochytrium dendrobatidis* (*Bd*). Bullfrogs that escape captivity can invade and contaminate new habitats with *Bd*. Chytridiomycosis epidemics can extirpate multiple amphibian species with dramatic effects on community structure and ecosystem function. Intervention efforts include continuous monitoring of amphibian populations, focused culling of established bullfrog populations, and alternative tactics such as augmenting antimicrobial peptides found on amphibian skin that can confer protection from *Bd*. The blue box (lower right) shows the mean percent of *Bd* growth inhibited by increasing concentrations of one antimicrobial peptide, esculentin. Amphibian species that are highly susceptible to lethal *Bd* infection have also been rescued to other locations for captive breeding and postepidemic reintroduction. Photo credits: bullfrog and cooked frog legs, Wikimedia Commons; infected *Atelopus zeteki* (center), Brenes R; inset *Bd* zoospores, Longcore JE; all other photos, Han BA).

extremely pathogenic to native bird species. Avian malaria epidemics throughout the Hawaiian Islands caused high mortality in the native birds, essentially clearing the lower island elevations of the native avifauna and contributing to the extinctions of several endemic bird species. In addition to species extirpations, introduced pathogens have facilitated drastic changes to community composition through invasion and displacement of native species. For example, Barley and Cereal Yellow Dwarf Viruses (**Figure 3(b)**) have enabled invasive annual grasses to outcompete and replace native perennial bunchgrasses in California. A parapox virus that is highly lethal to native red squirrels (**Figure 3(c)**) was introduced by invasive gray squirrels, and has accelerated the replacement of red squirrels by gray squirrels throughout much of the UK. Currently there is intensive research to determine whether invasive American bullfrogs are transmitting amphibian chytridiomycosis (**Figure 4**) to native amphibian populations in western North America, Central, and South America. Dissemination of the amphibian chytrid fungus *via* invasive reservoir species could be particularly devastating in areas such as the Venezuelan Andes, one of the world's biodiversity hotspots for endemic amphibians.

Negative Community and Ecosystem Consequences

The negative impacts of parasites and pathogens extend beyond their direct effects on hosts. Host population declines can generate secondary effects that ripple through natural communities, in severe cases triggering secondary extinctions termed "extinction cascades." Examples of such community-wide effects have been reported from both marine and terrestrial ecosystems. For example, the disappearance of long-spined sea urchins (*Diadema antillarum*) throughout the

Caribbean Sea by an unidentified pathogen caused a shift from coral- to algae-dominated reef communities (Harvell *et al.*, 1999). Similarly, the virtual disappearance of the American chestnut caused by the introduced fungal pathogen *Endothia parasitica* led to the extinction of eight Lepidopteran species that fed predominantly on this tree species.

Direct and indirect effects of infectious diseases can also damage the way an ecosystem functions. For example, the extirpation of adult amphibians by chytridiomycosis in Panama was linked to dramatic declines in amphibian larvae in stream communities. In the absence of tadpoles as filter feeders, grazers, and detritivores, studies report dramatic decreases in the quality and quantity of fine organic particulate matter suspended in the water column (i.e., seston, made up of periphyton, tadpole feces and bacteria). These changes are likely to affect communities downstream, which receive an influx of energy through the particulate matter generated by headwater communities. Ecosystem impacts can occur even when host mortality or species declines are not apparent. For example, experimental manipulations of intact grassland communities demonstrate that infection by foliar fungal pathogens affects net primary productivity and soil respiration by reducing root biomass, photosynthetic capacity, and leaf longevity in multiple grasses (*Andropogon*, *Poa*, and *Schizachytrium* spp.) (Mitchell, 2003).

Anthropogenic Drivers of Disease Emergence in Nature

In recent decades, a growing number of pathogen outbreaks in natural systems have been attributed to human activities. In particular, economic development influences disease emergence through changes in land use and human population growth. These activities can increase the risk of pathogen spread in natural populations through habitat destruction or fragmentation, pollution or other forms of habitat degradation, climate change, and global commerce.

Loss and Fragmentation of Natural Habitats

Humans destroy natural habitats in a plethora of ways, many of which have been summarized under the term habitat fragmentation – the process of reduction and isolation of a continuous natural habitat into smaller patches. Habitat fragmentation is of great conservation importance because it affects native communities directly by removing individuals and their resources, and indirectly through changes in community composition and species interactions. In terms of host–pathogen interactions, the reduction or subdivision of a host population can allow hosts in some patches to escape infection, especially if patches are isolated. Furthermore, if remaining host subpopulations are relatively small they may fall below the critical host threshold required for disease persistence (N_T). Consequently, one possible result of habitat fragmentation may be the decline or extinction of a parasite. Conversely, it is also possible that hosts moving between habitat fragments will continue to spread infection. Between these two contrasting scenarios, the benefits of increasing connectivity – namely the increased likelihood of species persistence and the potential for recolonization of locally extinct patches – appear to outweigh the potential costs of pathogen transmission between fragmented habitats on a landscape (McCallum and Dobson, 2002).

From a different perspective, habitat fragmentation can increase disease risks to native biota by providing opportunities for contact with domesticated crops, livestock, humans, and other native species crowded into adjacent habitats. For example, endangered red colobus monkeys living on the edge of forest fragments in Kibale National Park were exposed more frequently to multiple infections and had a higher prevalence and intensity of certain parasites, including harmful parasitic worms and protozoa, than monkeys living in the forest interior (Chapman *et al.*, 2006). The authors speculated that interactions with humans and domesticated species on the forest edge could cause greater exposure to some generalist parasites. Other work in this same region showed that forest fragmentation caused greater transmission of enteric bacteria among humans, livestock, and three wild primate species; this was partly associated with crop raiding behaviors of wild primates living in small fragments or on the edge of reserves.

Habitat Degradation and Pollution

Degradation of natural habitats takes many forms and has frequently been associated with pathogen outbreaks in wildlife and humans. Human activities such as logging, overexploitation of game and fisheries, erosion, and pollution can trigger pathogen epidemics. For example, direct injury to trees as a result of logging can increase disease in native plant communities. Many scientists suggest that the phocine distemper epidemic that spread through the seal populations in the North Sea can be attributed to the depletion of the fish stocks (by commercial fishing) in the North Atlantic ocean, leading to aberrant migrations of harp seals seeking food. In addition, wastewater run-off (which includes human sewage and agricultural byproducts like pesticides and fertilizers) into the ocean has been linked with toxoplasmosis in southern sea otters along the California coast, aspergillosis in Gorgonian sea fans in the Gulf of Mexico, and white pox in Elkhorn coral in the Caribbean sea (**Figure 3(d)**) among other diseases (reviewed in Harvell *et al.*, 1999). Across the eastern seaboard of the United States and in the Gulf of Mexico, eutrophication due to excess run-off from agricultural fertilizers has led to outbreaks of *Pfiesteria piscicida*, an aggressive protozoan that kills large numbers of fish. In freshwater ecosystems, the spread and duration of avian cholera outbreaks in waterfowl populations is greatly facilitated by pollution from nutrient-rich run-off into wetlands. This disease, caused by the introduced bacterium *Pasteurella multocida*, has become the second most important cause of waterfowl die-offs in North America.

Global Climate Change and Disease

Because the development and transmission of many parasites depend on environmental conditions, it is not surprising that

anthropogenic climate change is expected to influence the distribution, prevalence, and severity of pathogens in natural ecosystems. A number of recent examples illustrate that even small changes in climate can speed up parasite generation times, shift the geographic ranges of both hosts and vectors, and propagate complex trophic cascades leading indirectly to disease emergence. In the Arctic, warming temperatures appear to enable lungworm parasites to mature twice as quickly within their slug intermediate host. Infection severity in muskoxen (the definitive host) is postulated to increase as climate continues to warm. Coral reef ecosystems are among the most severely impacted by small increases in temperature, which can trigger coral bleaching (expulsion of symbiotic algae from the coral) and enhance the growth rates of a suite of opportunistic infections that have led to massive coral die-offs. Although such ocean warming events can occur naturally in association with El Niño-Southern Oscillation events, their frequency and severity have increased during the past 20 years. As a result of warming oceans, Caribbean yellow band disease has increased in prevalence and virulence to endanger all four coral species in the genus *Montastraea*, the most important reef building corals of this region (Harvell *et al.*, 2009). Because so many marine organisms depend on habitat provided by coral reefs, and because reefs can take many years to recover, the consequences of such massive die-offs are profound and long-lasting.

As climate warms, species range shifts will also bring new species in contact with each other and with new parasites. For example, the distribution of some animal diseases (such as the previously discussed *P. relictum*, and *B. dendrobatidis*) could lead to declines of many endangered bird and amphibian species surviving in higher elevations. Since global climate impacts all biological interactions, some ecologists have noted that the net effects of changing climate can increase disease in some regions (as exemplified above) and decrease disease in others. Predicting changing biological interactions as a result of changing climatic conditions is therefore an increasingly important goal for scientists and wildlife managers.

Worldwide Commerce and Travel

Worldwide commerce and animal trafficking are probably the most important causes of new disease outbreaks in wildlife populations. This traffic, which has dramatically increased in the past few decades, includes international trade of live animal and plant stock for commercial breeding purposes, for zoos and animal parks, for the pet trade and hunting, and for laboratory research. The situation is exacerbated by the staggering amounts of international transport of fruit, vegetables, and various animal parts which can facilitate the spread of pathogens. For example, most of the major pathogens that cause disease epidemics in wild birds in North America were probably introduced by humans. Both avian cholera (a highly virulent disease caused by the bacterium *P. multocida*) and duck plague (caused by a herpesvirus) appeared first in domesticated bird flocks and spread to wild waterfowl. Today these diseases are responsible for the deaths of tens of thousands of wild waterfowl each year in the United States alone. Similarly, avian malaria was introduced to Hawaii with imported birds, and brucellosis (caused by the bacterium *Brucella abortus*) was introduced to North America by domesticated livestock (and has since spread to wild deer, elk, and buffalo populations).

Disease spread into wildlife populations is further facilitated by the continual breakdown of barriers between farmed animals and wild animals. For example, aquaculture practices create crowded conditions where farmed Pacific salmon become heavily parasitized by sea lice (*Lepeophtheirus salmonis*). Wild juvenile salmon returning to open waters from spawning grounds travel migratory routes in close proximity to salmon farms, and become infected by sea lice on their way back out to the open sea. Similarly, there is concern that international demand for farmed bullfrogs is contributing to the spread of amphibian chytridiomycosis (**Figure 4**). As bullfrogs can carry very high pathogen loads with no clinical signs of infection, escaped bullfrogs that invade new habitats may transmit the chytrid fungus to native species that are less resistant to disease.

Sparse regulation of the growing exotic pet trade has exacerbated risks of novel pathogen introductions. Many exotic pets comprise a mix of captive reared and wild caught animals, which can facilitate the transmission of infectious diseases acquired from different environments. Currently there are no screening procedures enforced for exotic animals, especially for animals ordered over the internet. In 2007, the number of exotic animals imported into the US numbered over 37 million, reflecting a staggering diversity of reptiles, amphibians, rodents, primates, and various wild cat species.

Transport and commerce also pose a risk for the spread of invertebrate diseases due to the growing popularity of raising and releasing insects such as ants, ladybird beetles, bees, and butterflies for the purposes of biological control, pollination, education activities or for release at special events. Not surprisingly, the parasite communities of most insect species and disease risks for wild insect populations remain largely unknown. Honey bees are probably the best-studied insect species in terms of their infectious diseases, but disease threats have also caused concern for wild bumble bee species (*Bombus* spp.) in recent years. Since the 1990s, colonies of several native bumble bee species have been mass-produced and distributed for the pollination of greenhouse crops in North America. The movement of bumble bees from Europe to the US is thought to have introduced an exotic strain of a microsporidian parasite that is highly virulent to some North American bumble bees. This parasite appears to have caused the near-extinction of at least one bumble bee species (*Bombus affinis*). Furthermore, the prevalence of this and other pathogens was shown to be higher in commercially reared bumble bees than in nearby wild populations. Furthermore, infections in wild bees declined with increasing distance from industrial greenhouses, strongly implicating human rearing operations in the disease-induced declines of these pollinator species.

Positive Role for Pathogens in Biodiversity and Conservation

Although the risks they pose to threatened species are apparent, parasites and pathogens can also maintain biodiversity, in

part through their effects on the outcome of other species interactions. Indeed, recent advances in the field of disease ecology show that parasites can maintain diverse communities through their effects on competition, food web structure, and trophic interactions. Other work suggests that over longer timescales, parasites can promote evolutionary diversification among hosts. These interactions support the idea that parasites are integral components of ecological communities and can even mediate ecosystem processes. Although the conservation of parasites is almost never considered to be a priority, a key point highlighted by this section is that not all aspects of parasitism are necessarily negative; a strong argument can be made for conserving endemic parasites and pathogens.

Parasites, Species Interactions, and Biodiversity

One way that parasites can enhance biodiversity is by altering the outcome of competition between host species. In this case, generalist pathogens that are relatively benign to one species can lower the density of other hosts for which they are more pathogenic, potentially reversing the outcome of resource competition. Apparent competition is a related phenomenon whereby two or more hosts not directly competing for resources are affected by the same parasite to different degrees. In some cases, parasites shared between host species can threaten highly susceptible native species with extinction, as discussed *in* Infectious Diseases as Threats to Biological Diversity (e.g., red squirrel–gray squirrel parapoxvirus). In other cases, parasites can promote biodiversity by having disproportionate negative effects on otherwise dominant species, thus allowing many species to coexist at relatively low densities. Several case studies support the role of pathogens in determining plant and animal community structure and modifying the outcome of competition. For example, malaria parasites have been shown to facilitate coexistence between two *Anolis* lizard species on the Caribbean island of St Maarten; in parts of the island where lizard malaria is absent, only the dominant competitor was found, whereas the competitively inferior species could coexist at sites where malaria was present. Other studies found that pathogenic soil fungi can be more harmful to dominant tallgrass prairie species, facilitating the coexistence of less competitive species. Soil-borne diseases have also been implicated in the rate and direction of succession in plant communities and in determining patterns of seedling recruitment in tropical rain forests.

When parasites attack dominant herbivore species in a community, epidemics can cause major shifts in community composition through secondary effects on plant recruitment and abundance or on the density of predators and other natural enemies. One example involves the introduction of rinderpest in the Africa; this viral disease of cattle was introduced in the late 1800s and spread to native ungulate species including wildebeest and African buffalo, causing mortality rates of up to 80%. Starting in the 1950s, a vaccination campaign in cattle eventually eliminated the virus from wildlife, leading to a gradual rise in the numbers of wildebeest and other ungulate species. The growing numbers of herbivores affected species at other trophic levels: with the recovery of wildebeest populations, predator populations increased (including lions and hyenas) and the biomass of vegetation (especially grasses) declined. This example illustrates how the effects of pathogens on dominant or keystone host species can propagate through food webs and alter community structure.

From a different perspective, evolutionary adaptations of hosts and pathogens over long timescales could increase biodiversity through cospeciation (the joint divergence of two or more interacting lineages) and phylogenetic diversification. For example, one comparative study of wild primates and their pathogens showed that primate host species from more diverse lineages harbored a greater number of parasite species (including viruses, protozoa, and helminths) (Nunn *et al.*, 2004). This pattern could be caused by an evolutionary arms race between hosts and pathogens, whereby hosts mount greater resistance against infection, and parasites evolve higher virulence and transmissibility. Other explanations are possible, however, including the idea that parasites infecting primates from more diverse lineages experience greater opportunities for diversification themselves through cross-species transmission and host shifting. As a final possibility, coextinctions of hosts and parasites might drive these associations. Parasite lineages could lose transmission opportunities and go extinct themselves as their hosts decline in population size, or as their geographic ranges shrink.

Pathogens and Host Genetic Diversity

In many ways, the maintenance of genetic diversity within species parallels the coexistence of multiple species in ecological communities and should enable hosts to evolve in response to changing environments and ecological perturbations. In terms of host–pathogen interactions, parasites are likely to be powerful selective agents because they can spread rapidly and cause significant negative effects on host fitness. Host species that are continually exposed to a diverse array of parasites should therefore harbor a variety of resistance alleles or a repertoire of inducible defenses. A growing number of examples from wild populations, including parasitic worms infecting Soay sheep and snails, and fungal pathogens affecting plants show how parasites can favor allelic diversity or sexual recombination in their hosts (Altizer and Pedersen, 2008).

Given the strong selective pressures imposed by parasites and the benefits of host resistance traits, an important question is why aren't all individuals resistant to infectious diseases? Simple host–parasite models show that genetic variation underlying host resistance traits can be maintained by balancing selection or by costs of resistance traits for host survival or reproduction. Balancing selection simply refers to processes that can maintain genetic variation over time, including heterozygote advantage, frequency-dependent selection, and selection pressures that fluctuate over time or space. For example, under frequency-dependent selection, rare host genotypes are resistant to pathogen genotypes that attack more common host types; this advantage held by rare alleles can lead to cycles in both host and parasite allelic frequencies over time, resulting in the maintenance of host genetic variation in the long-term. The phenomenon of parasites tracking

common host genotypes has been demonstrated by studies of parasitic trematodes infecting freshwater snails in New Zealand and bacterial pathogens of *Daphnia* in European ponds, and is critical for arguments concerning the role of parasites in favoring sexual reproduction in their hosts.

The major histocompatibility locus (MHC) in vertebrates provides a notable example of extreme polymorphism and diversity of rare alleles maintained by balancing selection. Immune proteins coded by MHC genes can recognize and bind to pathogen proteins (antigens) inside infected host cells, and transport these antigens to cell outer membranes. Here, they are presented to other immune cells that destroy infected host cells and initiate antibody production. Specific MHC molecules preferentially bind to specific pathogen peptides, and hence different MHC alleles confer resistance to different pathogens. In natural populations, MHC class I and II genes show enormous variation, and even species with low overall genetic diversity can show high genetic variation at MHC genes. In one extreme example, Aguilar *et al*. (2004) demonstrated high levels of variation across five MHC loci in a population of the San Nicolas Island fox with essentially no genetic diversity across selectively neutral loci. These findings led authors to conclude that intense balancing selection had maintained MHC variation in the face of past bottlenecks.

In plants, gene-for-gene coevolution has also been demonstrated to promote a high diversity of resistance and virulence alleles. For example, long-term field studies of the interaction between wild flax and flax rust in natural populations in Australia indicate that many alleles can persist among metapopulations, and the distribution of genotypes can change rapidly during individual epidemics. Collectively, these examples emphasize that pathogens commonly exert selection pressures on their hosts, in many cases leading to genetic heterogeneity over space and time.

Parasites and Ecosystem Function

Parasites and pathogens have long been excluded from studies of food webs (maps of feeding interactions in an ecological community) and ecosystem function (flows of energy and nutrients), in large part due to their small size and assumptions that they contribute little biomass to ecosystems. However, several recent studies have challenged this view by showing that parasites can dramatically alter the connectivity and stability of ecological food webs, and can represent a surprising fraction of biomass in real-world ecosystems. For example, parasites are involved in more food web links than predators in some estuarine and salt marsh ecosystems. Moreover, adding parasites to food webs increased measures of food web connectance (the number of actual links between species relative to the total links possible) by up to 93%. Indeed, parasites can account for 3–13% of the biomass of their free-living plant and animal counterparts in some estuarine ecosystems. Some parasitic worms like trematodes had biomass levels comparable to those of fishes and small arthropods, and their collective biomass was greater than that of top predators (namely shorebirds) (Kuris *et al*., 2008). A similar study of plant fungal pathogens in a tallgrass prairie ecosystem showed that the estimated pathogen biomass was comparable to that of herbivores, and that pathogen effects on grass biomass appeared to be stronger than the effects of herbivory (Mitchell, 2003). Other researchers have pointed out that host death caused by pathogens can liberate carbon and nutrients from host cells and tissues, increasing their turnover in natural ecosystems. Not surprisingly, however, the effects of pathogens on biogeochemical cycling remains relatively unexplored. More generally, a growing number of studies point to parasites as key players controlling properties of food webs and biomass production, and suggest that parasites should be viewed as essential components of ecosystem function rather than as nuisances that interfere with management objectives.

Parasites Can Inform Conservation of Threatened Species

Pathogens can reveal valuable information about host population size and movement in ways that enhance conservation efforts. A handful of case studies show that the evolution of some pathogens can be rapid enough to reveal cryptic host geographic isolation and historical contact patterns. For example, an analysis of sequence variation in the feline immunodeficiency viruses of cougars in western N. America (Biek *et al*., 2006) showed that different pathogen lineages dominated in different parts of the cougars' range, indicating population structure in the host that could not be uncovered by analysis of molecular markers in the cougars themselves. Genetic analysis further revealed that the spatial occurrence of viral lineages is expanding, most likely due to increases in cougar population size. In another study, three species of whale lice revealed historical separation of populations of endangered right whales (Kaliszewska *et al*., 2005). Genetic analyses of parasite mitochondrial DNA showed that whale populations in the North Atlantic, North Pacific, and southern oceans diverged several million years ago, following the formation of the Isthmus of Panama. High genetic diversity among lice in the currently small populations of North Atlantic right whales also indicated that their host population sizes probably numbered in the tens of thousands before the modern era of commercial whaling. This work illustrates the usefulness of pathogen molecular markers for understanding historical population size, contemporary population movements and geographic structuring of their hosts.

Managing Diseases and Biodiversity in the Future

Parasites and pathogens that pose significant threats to conservation programs can be divided roughly into two categories. The overwhelming majority of epidemics begin with the introduction of a disease into a naïve population. A second category of concern are cases where environmental changes (such as habitat loss and pollution) shift host relationships with endemic parasites. Effective disease prevention must therefore address both of these areas by limiting the introduction of novel pathogens and mitigating large-scale environmental changes that facilitate the emergence of endemic disease or cross-species transmission with domesticated species and humans. Monitoring and disease screening programs, quarantine and vaccination regimes, and attention to

captive breeding programs comprise common approaches to preventing or managing disease outbreaks in wild populations.

Monitoring Populations for Infection

Most pathogens are discovered after epidemics have already spread through wild populations. A powerful tool in the management of diseases is thus to monitor threatened populations for signs of infection before disease-induced mortality occurs. Several tools exist to monitor host populations and track the spread of an epidemic, and the efficacy of a monitoring program will increase with the number of host animals included in the screening (Wobeser, 2007).

Parasite prevalence and intensity are often monitored by examining blood, tissue, and fecal samples in animals. The presence of blood parasites (e.g., trypanosomes, malaria, and filarial nematodes), anemia, elevated leukocyte levels, and pathogen-specific antibodies can all be detected from relatively small volumes of blood. Antibody assays (e.g., enzyme-linked immunosorbent assay) can show evidence of current and past infection. Molecular techniques based on polymerase chain reaction tests can reveal the presence and intensity of infection in host blood or tissue by identifying and quantifying pathogen genetic material. In addition, many microbial pathogens can be cultured directly from tissue showing signs of infection or from swabs taken from the mouth, ears, eyes, nose, genitals, anus, or skin. Scans of feces will also provide information on the diversity of intestinal parasites in an individual as well as fecal egg counts per host as a measure of infection intensity.

Hosts that are hunted or culled for other purposes are frequently examined for internal and external parasites. Parasitic arthropods (e.g., ticks, mites, fleas, and biting dipterans) are also monitored because they represent key vectors for many infectious diseases and can decrease the condition of their hosts by drawing on host resources. Combining infection-related data from animals that are opportunistically sampled over a long time period can be useful for examining temporal trends. Long-term monitoring over large spatial scales is usually time consuming and expensive, but such studies often provide important information on the conditions leading to elevated disease prevalence, and are invaluable to developing effective disease management and prevention strategies for wild populations.

Assessing Disease Threats

Screening programs will verify the presence of a particular pathogen in a wild population, but do not provide information about population-level consequences resulting from infection. Epidemiological models outlined in the section on Basic Epidemiological Principles indicate that this determination requires information on both the prevalence of a pathogen (y) and its effects on host fitness (α and δ). It is important to note that *post mortem* examinations are often performed to determine the effects of disease on individual hosts. However, the presence of a pathogen in dead animals does not necessarily indicate the cause of death, nor does it reveal subtle disease effects on host behavior or fecundity. Ideally, captive or wild hosts should be monitored throughout the course of infection to compare survival and fecundity between infected and uninfected hosts. Experimental manipulation of parasite loads in natural populations remains the most direct way of assessing the effects of pathogens on host populations.

Intervention Methods

Historically, diseases in wild populations have drawn the attention of wildlife managers only after an epidemic severely threatened the host population or when a pathogen threatens agricultural crops or livestock. In general, the types of management regimes used to limit disease spread vary depending on the type of pathogen, the threat it poses to host populations, and the availability of financial resources. Management regimes also depend on the existence of agricultural or veterinary information and tools, which often do not exist for free-living plant and animal populations. That being said, pathogen outbreaks in recent years have allowed researchers to employ some of the more common intervention methods available and test their efficacy in reducing pathogen spread and impacts.

Vaccination, Culling, and Quarantine

Several intervention measures center on decreasing the number of contacts between susceptible and infectious host individuals to prevent new infections. Viral infections and, less frequently, bacterial infections in vertebrate animals can be effectively controlled by vaccinations, which confer a period of immunity and essentially transfer susceptible hosts directly to the resistant/immune class (*see* Microparasites and **Figure 1**). However, the cost of vaccinations and lack of testing (for safety and efficacy) in most wildlife species may severely limit the success of population-wide disease control plans. Epidemiological models can be helpful in determining where to concentrate finite vaccination resources, and what proportion of the population to target to limit or eradicate disease. A frequent goal of vaccination is to treat enough hosts to prevent pathogen invasion, or to achieve local eradication. This critical vaccination threshold is $1–1/R_0$, which reduces R_0 below 1. Vaccinations have been successfully administered to slow the spread of rabies in European foxes and North American raccoons, and vaccination of domesticated reservoir hosts (cattle and domestic dogs) has reduced the threats of rinderpest and canine distemper outbreaks in wild ungulates and carnivores in east Africa. A key benefit from vaccination is that even when susceptible hosts are not vaccinated they are less likely to be infected because they are surrounded by immune individuals. It is also important to note that even if too few hosts are vaccinated to eradicate a pathogen, low-coverage vaccination (of a small fraction of hosts) can still prevent local host extinction, as recently illustrated by vaccination of Ethiopian wolf populations against a potentially devastating rabies virus outbreak.

Culling is the removal of host individuals by lethal means. Culling can be indiscriminate or targeted solely toward infected individuals and is most often used when transmission is

believed to be density-dependent. The goal is to reduce host densities below the threshold needed for parasite persistence, N_T. Selective culling targeted toward infected individuals is analogous to intensifying parasite-induced mortality (α; **Figure 1**). This strategy can effectively decrease disease prevalence and lower R_0. Selective culling has been implemented to counter the spread of certain tree diseases (e.g., Dutch elm disease). Although less frequently applied to vertebrate populations, routine culling has been applied to some populations of Cape buffalo in South African national parks when individuals are found with tuberculosis. In Venezuela, invasive American bullfrogs are routinely culled to slow the spread of amphibian chytridiomycosis to native species (**Figure 4**).

Quarantine involves the isolation and care of infected individuals from a population currently experiencing an outbreak, with the goal of decreasing contact rates between infectious and susceptible individuals. Although commonly employed to slow or stop human disease outbreaks, this approach is less often applied to wild populations. An exception occurs when host populations are dangerously low in numbers such that every member may be of great value and worth rehabilitating. This was the case during a canine distemper epidemic in Wyoming black-footed ferrets (*noted in* Introduced Parasites and Species Declines), and a similar situation may arise if future epidemics affect endangered African apes, Ethiopian wolves, or African wild dog populations.

Alternative Interventions

Administering focused treatment regimes of antiparasitic drugs could be the best option for infectious diseases in some populations, although little quantitative data exist on the success of these approaches. Drugs and other forms of chemotherapy are most frequently administered for bacterial, fungal, helminth, and ectoparasitic infections. This method of disease control least effectively addresses the ultimate cause of a disease and may be extremely costly for population-wide control measures (e.g., systemic fungicides to counter tree blights). Other examples of targeted intervention include the treatment of mites causing mange in cheetahs, Arctic foxes, mountain gorillas, and wombats, but the long-term efficacy of these treatments is still unknown. Recently, researchers discovered that the microbial community and a wide array of antimicrobial peptides covering the surface of amphibian skin can confer protection against chytridiomycosis in some frog species (**Figure 4**). The possibility of augmenting this unique form of innate immunity to help conserve amphibian populations is an active area of research.

Habitat Management

Management regimes that address the ultimate causes of disease outbreaks have the greatest potential for removing disease threats but are also the most difficult to implement. Indeed, the ultimate causes of marine invertebrate diseases, namely, pollution and ocean temperature changes, are so global and diffuse in origin as to be impossible to confront in any single species management plan. Managing terrestrial habitats may alleviate some disease-related conservation problems. Two such examples are the proposed creation of a bovid-free land zone around Yellowstone National Park to prevent contact between bison and cattle, and the removal of feral pigs from the Hawaiian Islands (because their activities increase mosquito-breeding areas and elevate the transmission of avian malaria).

An understanding of disease ecology is also pertinent to the design of habitat reserves. For example, how large should reserves be to prevent host crowding and limit disease transmission? Do corridors between reserves increase the threat of pathogen transfer among locations, or does host dispersal among habitats facilitate the spread of resistance genes or aid in parasite avoidance? Maintaining species richness and genetic diversity within reserves is also critical to limiting threats from disease. In particular, high species diversity may buffer natural communities from devastating epidemics, and habitats that are restored with genetically diverse stock may be less susceptible to pathogen invasion. In terms of host species diversity, this can be particularly important in slowing the spread of multihost pathogens. One hypothesis that has gained support from studies of both plant and animal communities is the dilution effect, which predicts a pattern whereby disease risk decreases with increasing host species diversity through one of several underlying mechanisms (Keesing *et al.*, 2006). This buffering effect of host diversity can be especially important for pathogens spread by vectors that feed on several host species, but for which only a few host species effectively amplify the pathogen. The best-studied examples are the Lyme disease bacterium and West Nile virus, two pathogens that are transmitted between dozens of animal species by ticks and mosquitoes, respectively. In both cases, the presence of a diverse mammal and bird community, respectively, reduces the fraction of vectors that feed on highly competent hosts, and lowers infection prevalence in vectors. Thus, habitat management practices that maintain diverse vertebrate communities (such as larger reserve sizes and corridors that connect existing reserves) could lower the risk of pathogen transmission to wildlife and humans.

Concerns for Captive Breeding Programs

Conservation efforts rely increasingly on captive breeding programs to augment and restore free-living populations. Because captive animals are particularly susceptible to infections, pathogens represent a significant concern of captive breeding managers. Captive animals may acquire novel infections from unrelated species kept in the same pen, foster parents, or from individuals of the same species or closely related species. For example, captive African elephants kept in mixed collections have been infected with a lethal herpesvirus that occurs without disease symptoms in their Asian elephant pen mates. Captive-bred hatchlings of the endangered Mauritius pink pigeon contracted and succumbed to a herpesvirus infection that their foster parents (domestic rock doves) were carrying without ill effects. Several human diseases, such as measles, tuberculosis, herpesvirus, and influenza are highly virulent for nonhuman primates, especially gorillas, chimpanzees and other apes, which poses a concern for zoo and sanctuary animals exposed to potentially infected humans (and for wild animals exposed to ecotourism groups).

Many animals in captive breeding programs are often held close together, a practice that poses two disease-related risks.

First, animals are likely to become stressed and hence more susceptible to infection (particularly those that are territorial or normally persist at low densities in the wild). Second, crowding in pens or cages can elevate host densities above the threshold necessary for virulent pathogens to invade, and will also increase transmission rates (e.g., hosts may reinfect themselves by ingesting the eggs of their own parasitic nematodes released into their pen). Vector-borne diseases or parasites with complex life cycles may be less of a concern to captive breeding programs because of the likely absence of intermediate hosts (or vectors) that are necessary for transmission.

Finally, additional complications exacerbating disease problems in zoos and captive breeding programs stem from inbreeding depression, or the genetic impoverishment of a captive colony due to loss of diversity and the expression of deleterious recessive alleles. This loss of genetic variability leads to homogeneous captive populations that can be very susceptible to a variety of pathogens, as implicated in the high mortality that captive cheetah populations experienced due to a feline infectious peritonitis virus. Hence, genetic and ecological problems can operate in synergy to reduce population size and diminish heterozygosity, leading populations toward increased disease susceptibility, and possible extinction. In conclusion, as captive breeding programs expand, disease-related problems are likely to become even more prevalent and will require serious precautions including screening and treatment, suitable housing and animal care, and separating potentially infected individuals (including some humans) from healthy captive populations.

Equal Rights for Parasites?

Parasites are an integral part of life on earth, with their biodiversity projected to be significantly greater than the species richness of free-living hosts, but biologists have uncovered only a tiny percentage of this diversity. Many micro- and macroparasites that live uniquely on threatened host species could go extinct long before their hosts, and this poses a distinct threat for parasite extinction. As with most taxa, we do not have accurate numbers of how many species of parasitic organisms might be affected by future extinctions. But in contrast to many other threatened species, the conservation of parasites is virtually never considered a primary goal of conservation strategies. This is most likely due to their small body size and cryptic nature, and because parasites tend to draw the most public attention during epidemics where large numbers of hosts die from infection.

Many parasites have fascinating life cycles and have evolved incredible strategies for dispersing their progeny to new hosts. Some parasites have become sources of pharmaceutically important products; for example, a tick salivary gland component (calreticulin) has been used in dogs to treat thrombosis and heart disease. Perhaps more importantly, as humans disturb natural balances and break transmission barriers among species, pathogen outbreaks among rare or threatened host species will continue to occur. Because pathogens are a major factor promoting both genetic and species diversity in natural communities, conservation strategies that result in the loss of native parasites might ultimately rob host populations of genetic diversity needed to respond to future epidemics. An important question that has yet to be answered is "do healthy populations have a diverse community of parasites?" If the answer is "yes," then maintaining endemic parasite populations, and hence the ability of wild populations to respond evolutionarily to parasite-mediated selection, could be one of the best long-term strategies for mitigating the risks of infectious diseases.

See also: Biodiversity and Human Health. Biodiversity, Evolution and. Captive Breeding and Reintroduction. Climate Change and Ecology, Synergism of. Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. El Niño and Biodiversity. Endangered Amphibians. Estuarine Ecosystems. Hemiparasitism. Implications of Urbanization for Conservation and Biodiversity Protection. *In Situ*, *Ex Situ* Conservation. Loss of Biodiversity, Overview. Microbial Biodiversity. Microorganisms (Microbes), Role of. Parasitism. Population Diversity, Overview. Species Interactions. Traditional Conservation Practices. Wildlife Management. Worms, Nematoda. Worms, Platyhelminthes

References

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, and Wayne RK (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America* 101: 3490–3494.

Altizer S and Pedersen A (2008) Host–pathogen evolution, biodiversity and disease risks for natural populations. In: Carroll SP and Fox CW (eds.) *Conservation Biology: Evolution in Action*, pp. 259–277. Oxford: Oxford University Press.

Anderson RM and May RM (1991) *Infectious Diseases of Humans*. Oxford: Oxford University Press.

Biek R, Drummond AJ, and Poss M (2006) A virus reveals population structure and recent demographic history of its carnivore host. *Science* 311: 538.

de Castro F and Bolker B (2005) Mechanisms of disease-induced extinction. *Ecology Letters* 8: 117–126.

Chapman CA, Speirs ML, Gillespie TR, Holland T, and Austad KM (2006) Life on the edge: Gastrointestinal parasites from the forest edge and interior primate groups. *American Journal of Primatology* 68: 397–409.

Dobson AP and Hudson PJ (1992) Regulation and stability of a free-living host–parasite system: Trichostrongylus tenuis in red grouse. II. Population models. *Journal of Animal Ecology* 61: 487–498.

Harvell CD, Kim K, Burkholder J, *et al.* (1999) Emerging marine diseases – climate links and anthropogenic factors. *Science* 285: 1505–1510.

Harvell D, Altizer S, Cattadori IM, Harrington L, and Weil E (2009) Climate change and wildlife diseases: When does the host matter the most? *Ecology* 90: 912–920.

Hudson PJ, Rizzoli AP, Grenfell BT, Heesterbeek JAP, and Dobson AP (eds.) (2002) *Ecology of Wildlife Diseases*. Oxford: Oxford University Press.

Jolles AE, Cooper DV, and Levin SA (2005) Hidden effects of chronic tuberculosis in African buffalo. *Ecology* 86: 2358–2364.

Kaliszewska ZA, Seger J, Rowntree VJ, *et al.* (2005) Population histories of right whales (Cetacea: Eubalaena) inferred from mitochondrial sequence diversities and divergences of their whale lice (Amphipoda: Cyamus). *Molecular Ecology* 14: 3439–3456.

Keesing F, Holt RD, and Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecology Letters* 9: 485–498.

Kuris AM, Hechinger RF, Shaw JC, *et al.* (2008) Ecosystem energetic implications of parasite and free-living biomass in three estuaries. *Nature* 454: 515–518.

McCallum H and Dobson A (2002) Disease, habitat fragmentation and conservation. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269: 2041–2049.

Mitchell CE (2003) Trophic control of grassland production and biomass by pathogens. *Ecology Letters* 6: 147–155.

Moore J (2002) *Parasites and the Behavior of Animals*. Oxford: Oxford University Press.

Nunn CL, Altizer S, Sechrest W, Jones KE, Barton RA, and Gittleman JL (2004) Parasites and the evolutionary diversification of primate clades. *American Naturalist* 164: 90–103.

Wobeser GA (2007) *Disease in wild animals: Investigation and management*. Berlin: Springer Verlag.

Dispersal Biogeography

Ran Nathan, The Hebrew University of Jerusalem, Jerusalem, Israel

© 2013 Elsevier Inc. All rights reserved.

Glossary

Dispersal The movement of disseminules away from their source, either their place of birth or their breeding site. Because of the variety of disciplines involved with the study of movements of organisms, the term is often misused and confused with a plethora of movement terms. The term dispersion refers to the fine-scale spatial distribution pattern of organisms, and migration is in some cases equivalent to dispersal, though it typically refers to a particular (directional and synchronized) type of movement. The term immigration is dispersal from the perspective of the recipient locality regardless of the source – the arrival of new individuals to an area previously not occupied by that species. It should be emphasized that dispersal *per se* does not deal with events that occur after a movement has been completed and the term "effective dispersal" should be used for events of dispersal followed by successful establishment. Similarly, the term colonization should be used for incidents of immigration that are followed by establishment of a viable population.

Dispersal barrier The species-specific physical and biological restrictions a propagule must overcome to accomplish successful colonization.

Dispersal biogeography Usually refers specifically to "a branch of historical biogeography that attempts to account for present-day distributions based on the assumption that they resulted from differences in the dispersal abilities of individual lineages" (Brown and Lomolino, 1998; p. 628). More generally, it is the study of the relationships between dispersal and the geographical distribution of organisms. The latter definition is adopted for this article.

Dispersal route The particular set of physical and biological conditions that allow organisms to cross dispersal barriers.

Disseminule The unit of dispersal; any part or stage in the life cycle of an organism that is used for dispersal.

Propagule The unit of colonization; a disseminule, or group of disseminules, which has the full potential to establish a new population.

Range expansion The spatiotemporal process in which successive colonization events increase the size of a species' geographical range. The terms invasion, migration (as used in the botanical literature), and spatial spread are roughly synonyms for range expansion. Invasion, in particular, is often used to describe colonization far outside the natural species range.

Introduction

In Chapters XII and XIII of *The Origin of Species*, Charles Darwin (1859) discussed how patterns of geographical distribution of species are related to his ideas of natural selection. He argued that "all the grand leading facts of geographical distribution are explicable on the theory of migration, together with subsequent modification and the multiplication of new forms." Darwin illustrated this argument through many examples, by explaining different biogeographic patterns, emphasizing the role of dispersal barriers (*see* Glossary) and the various means of long-distance dispersal (LDD) through which organisms are able to cross them. Although Carl Linnaeus (1707–1778) and Augustin-Pyramus de Candolle (1778–1841) had already expressed similar ideas about the importance of dispersal in determining species distributions, Darwin (1809–1882) and Alfred Wallace (1823–1913) are those most identified with the dispersalist school. This school asserts that disjunction and range expansion occur through rare LDD events across preexisting barriers. Joseph Hooker (1817–1911), one of Darwin's closest confidants, viewed LDD across barriers as extremely unlikely and argued that range expansion was possible only across ancient land bridges that once connected currently disjunct continents. He thus argued that biogeographic disjunction can be explained by vicariant (splitting) events in which new barriers have subdivided a previously continuous range. Evidence for Hooker's land bridges never emerged; however, the two basic opponent ideas of dispersal versus vicariant history were already brought onto the front stage of biogeography. In *The Secular Ark*, Browne (1983) describes the nineteenth-century discussion on this issue involving Darwin, Wallace, Hooker, and others.

The twentieth century has seen a further polarization of the vicariance/dispersal debate. Plant biogeographers such as Evgenii Wulff (1885–1941) and Stanley Cain (1902–1995) reiterated serious doubts about the ability of LDD to account for present plant disjunctions, in light of the extreme rarity and lack of direct empirical evidence of these events. Others, like Philip Darlington (1904–1983), advocated Darwin's dispersalist view, but without Darwin's reliance on specific centers of origin. In particular, the equilibrium theory of island biogeography (MacArthur and Wilson, 1967) ascribes a critically important role to immigration (*see* Glossary) in determining the number of species on islands. This landmark theory, which had a major impact on both ecology and biogeography, modeled the probability of immigration of a new species as roughly a negative exponential function of the number of species on an island, with the rate also decreasing with distance from the mainland.

At approximately the same time that Robert MacArthur and Edward Wilson published their theory, the important theory

of continental drift, or, more generally, of plate tectonics, first proposed early in the twentieth century, was finally accepted. This theory provided a vicariant mechanism to explain biogeographic disjunctions without LDD events. In the 1970s and the early 1980s, the concept of vicariance biogeography received much attention among biogeographers, as reflected in three books by Nelson and Rosen (1981), Nelson and Platnick (1981), and Wiley (1981). The main argument raised against dispersalists' explanations, following the similar criticism by Leon Croizat (1894–1982), was that LDD is a unique, highly stochastic and unpredictable phenomenon; hence, its relationships with species distributions are not testable. However, truth and testability remain distinct: Obstacles to proving the truth of an idea do not imply its falsehood. The explanations of biogeographic patterns by dispersalists were condemned as "story-telling" scenarios which "in explaining everything really explain nothing." However, solutions of complex biogeographic patterns by vicariant events alone, if taken to their limits, are also speculative. There are, indeed, examples of extreme cases in which disjunctions are exclusively explained by dispersal (e.g., colonization of remote newly created islands) or by vicariance (e.g., postglacial relicts). However, none of these extreme arguments provides a general explanation. In fact, as shall be seen, not only do both phenomena exist, but they also frequently co-occur.

The dispersal–vicariance debate reflects a more general controversy over the biogeographic relevance of dispersal: To what extent does dispersal influence species distribution? This debate concerns whether disjunct distributions have resulted from dispersal across a barrier or from a barrier splitting a continuous range. But dispersal and vicariant events are not the sole explanations for disjunctions, which may also result from differential extinction within a species range. The major problem in limiting dispersal biogeography to its currently narrow definition (*see* Glossary), is that such a definition reduces the entire phenomenon of dispersal to LDD across barriers, thereby ignoring the various biogeographic patterns to focus solely on disjunct distributions. As this discussion will show, there are many other distributional patterns that cannot be explained by either vicariant or extinction events but only by long- and short-distance dispersal events. Rich evidence gathered from diverse independent sources indicates not only that LDD exists, but also that it is the only possible explanation for the geographical distributions of many species, such as the Hawaiian biota (Carlquist, 1974). Furthermore, LDD is not necessarily as random or unpredictable as it is often considered. The relative importance of dispersal varies between species, biogeographic regions, and times. Thus, to assess the relative biogeographic importance of dispersal, one needs to explore the nature of dispersal and its relationships with the geographical distribution of species. The broader view of dispersal biogeography (*see* Glossary) propounded here aims to understand how dispersal influences, and is influenced by, the geographical distribution of organisms. This view has received enormous support since the publication of the first version of this article in 2001, giving rise to a balanced approach recognizing the importance of both dispersal and vicariance (de Queiroz, 2005; Crisp *et al.*, 2011).

Perspectives on Dispersal

The Ecological Perspective

As a general rule, most disseminules (*see* Glossary) travel relatively short distances from their parent source. The basic description of dispersal, called the dispersal curve, is a graph plotting the distribution of disseminules against the distance from their parent source (**Figure 1**). The dispersal curves, derived for a great number of organisms, show not only tremendous variability, but also a right-skewed leptokurtic distribution with overwhelming regularity. That is, they typically describe a rapid decline in disseminule density with distance from their parent source, with more disseminules both near the source and along the tail than a corresponding normal distribution (*see* Dispersal Models). The restricted spatial extent of dispersal has important implications for the dynamics of populations, metapopulations, and communities and for other topics addressed in ecological studies of dispersal (Nathan and Muller-Landau, 2000).

The ecological perspective of dispersal is concerned not only with small spatial scales, but also with relatively short temporal scales. Dispersal is often restricted to early stages during an organism's life cycle (seeds, eggs, larvae, and juvenile stages in animals); the term natal dispersal denotes dispersal of young animals whereas breeding dispersal indicates dispersal of reproductive adults. Disseminules of many plants, nematodes, rotifers, desert butterflies, copepods, and other invertebrates can persist for years – sometimes hundreds of years – as dormant seeds or eggs, providing the means for a temporal exploration for favorable establishment conditions. From another temporal perspective, dispersal is typically a seasonal phenomenon, and synchronization of dispersal in different individuals, populations, and species is fairly common in both animal and plant communities. It is important to keep in mind that dispersal does not always lead to successful establishment and reproduction in the next generation. In fact, the seed-to-adult survival probability is rarely high in any species and in many, if not most, species can be extremely low. Cases in which dispersal has been followed by successful establishment are often called "effective dispersal" and nearly all referrals to "dispersal" in the biogeographic perspective actually mean effective dispersal. However, the importance of distinguishing the (small) subsample of effective dispersal cases from all dispersal events is now widely recognized in dispersal ecology studies (*see* Ecological Opportunity: Postdispersal Establishment); the growing use of terms like "long-distance dispersal and establishment" (Crisp *et al.*, 2011) in dispersal biogeography literature implies an emerging agenda to explore the distinct role of transport processes and postdispersal establishment.

The Evolutionary and Population Genetics Perspectives

The great variety of dispersal-determining traits presents clear evidence that dispersal is of important adaptive value for many life forms on Earth. From the evolutionary perspective, the study of dispersal is concerned with the source and significance of this variation: The factors and processes that affect the advantage of dispersal. Typically, dispersal is advantageous for escaping the high mortality near the source, reducing

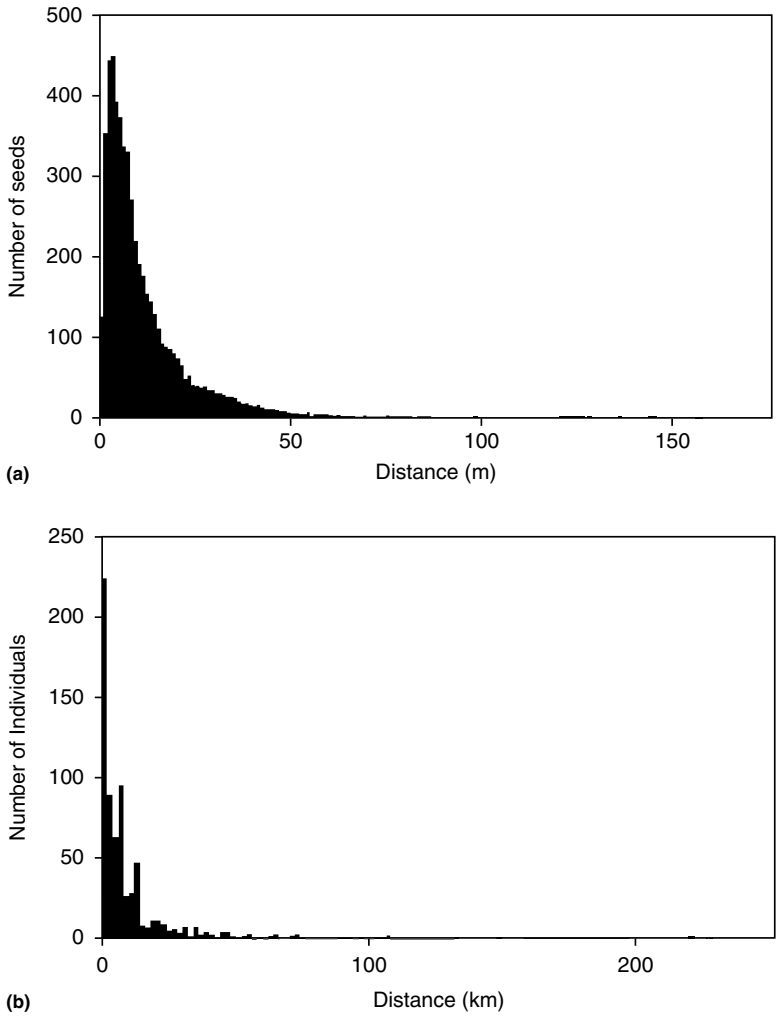


Figure 1 Dispersal diagrams representing the distribution of dispersal distances, for the winged seeds of Aleppo pine (*Pinus halepensis*) (a) and for first-year woodpigeons (*Columba palumbus*) (b). Dispersal of pine seeds was estimated by combining mechanistic simulation of wind dispersal with seed-trap data collected during six dispersal seasons in an isolated pine stand on Mt. Carmel, Israel, whereas bird natal dispersal was estimated with permission from ringing and recovery data gathered in Britain from 1909 to 1994. Note the three-order of magnitude difference between abscissas. Pine data reproduced from Nathan R, Safriel UN, and Noy-Meir I (2001) Field validation and sensitivity analysis of a mechanistic model for tree seed dispersal by wind. *Ecology* 82: 374–388, and the bird diagram was redrawn from Paradis E, Baillie SR, Sutherland WJ, and Gregory RD (1998) Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology* 67: 518–536.

competition with close relatives, minimizing the fitness cost of inbreeding depression by reducing the chances of mating with close relatives (in sexual species), and for locating new resources for establishment, growth, and reproduction. Theoretical models (see Johnson and Gaines, 1990, for a review) show that even in environments in which the physical conditions are stable in time and space, surprisingly high proportions of dispersing progeny may be adaptive. Since mutations beneficial to successful dispersal are typically rare, evolutionary effects are observable only at relatively long time scales (but *see* The Geography of Dispersal). However, there is no characteristic spatial scale for the evolutionary perspective of dispersal. Adaptive advantages of dispersal are expressed at both small and large spatial scales.

Dispersal provides the means for gene flow within and among populations; thus, studies of population genetics explore how dispersal, in conjunction with the breeding system and selection regime, affect the genetic structure of populations. In general, high rates of dispersal lead to mixing of genotypes and thus lower regional genetic diversity, whereas low rates of dispersal (isolation) typically result in low local genetic diversity, but high regional diversity. Although these effects may also occur at small spatial and temporal scales, they are notably more pronounced in large ones.

The Biogeographic Perspective

Dispersal becomes important at the biogeographic level at large temporal or spatial scales when and where propagules disperse outside the species range and successfully establish viable populations (*see* Components of Colonization Success). The identification of locations as "within" or "outside" the species range (and obviously the definition of a species range itself) is problematic since sampling of any species is typically

incomplete. Even if one assumes complete information on the exact location of each individual at any particular time, the issue of species boundaries is difficult because each species exhibits a profound variation in its local density and, in fact, is entirely absent from most locations. In biogeography, this distinction is less problematic when dealing with larger spatial and temporal scales.

Spatial spread can occur at different rates and can generate different spatiotemporal patterns, from slow gradual expansion to abrupt long-distance jumps. Both rate and pattern depend on the two basic components of spatial progression the movement itself (dispersal) and local population growth (reproduction) – and their interplay (*see* Modeling). The biogeographic literature provides numerous examples showing a continuum between these extremes, and it is also clear that gradual spread and long jumps can follow each other. However, there are basic differences between these two modes: Gradual range expansion is the outcome of multiple ordinary (i.e., short-distance) dispersal events over a relatively long time period, whereas jumps are rare individual LDD events that abruptly change the species' range.

Gradual Range Expansion

Gradual range expansion is the continuous spread of populations through multiple ordinary (short-distance) dispersal events, usually over quite long time scales. Very slow expansion has been termed secular dispersal (or secular migration); more rapid (though still gradual) expansion is referred to as diffusion (Pielou, 1979). Examples of gradual range expansion are the initial spread of mollusks, worms, crabs, and fishes from the Red Sea to the Mediterranean (*see* Removal of Barriers), and of muskrats in Europe (*see* The Spread of Alien Species).

Jump Range Expansion (Jump Dispersal)

Although most dispersal is limited to relatively short distances and therefore covered appropriately by the ecological scale (*see* The Evolutionary and Population Genetics Perspectives), rare LDD events have crucial implications for biogeography. There is enormous variation in the ability to disperse disseminules over large distances within and, more significantly, between species. Specific morphological characteristics are associated with LDD ability (**Box 1**). When such an event ends in establishment, the spatial spread is clearly discontinuous. The frequently used term for this type of spatial spread, jump dispersal, does not fully describe the process: Spatial spread is not only restricted to the movement phase (dispersal), but also necessitates establishment. Hence, the term jump dispersal should be replaced with jump range expansion. Note also that although the original term is appropriate for describing LDD over barriers, the term jump range expansion is also applicable to situations in which suitable habitats are as common in the gap as in the recipient area but are unoccupied (*see* Holocene Postglacial Spread *and* Dispersal Models).

Previous studies have suggested many methods of classifying the various means of dispersal. The most fundamental distinguishes between organisms that are actively dispersed (by their own power) and those that are passively dispersed (by a physical agent or by other organisms). Organisms unusually capable of active dispersal are said to be vagile, whereas those unusually capable of passive dispersal are called pagile. Another common classification, mostly used for plants, is based on the primary dispersal agent, using the suffix chory (to disperse) at different levels: For example, zoochory is dispersal by animals, omithochory by birds, and endozoochory is within an animal. Neither classification should be taken as an absolute: Many vagile organisms also disperse passively, and many pagile organisms can move actively over short distances. Seeds of many plants are actually dispersed by multiple agents, rather than by a single one (Nathan *et al.*, 2008b).

Relatively few animals have the capacity to travel long distances under their own power. Most of these are volant (flying) species such as birds (**Box 1(a)**), bats, and butterflies, as well as large mammals, reptiles, and fishes capable of swimming across the oceans. Although many migrate seasonally, this large-scale movement usually does not affect the geographical distribution of the breeding range. However, extra-limital sightings are well documented for many taxa, and occasionally may result in the establishment of a new colony outside the species range (*see* Recent Range Expansion). Nonvolant animals cannot vault dispersal barriers as do flying animals, rendering their active dispersal generally less effective. In plants, active dispersal is very rare and is restricted to the phenomenon of ballistic dispersal, in which seeds are explosively discharged from their base (**Box 1(b)**) and are dispersed at most tens of meters.

Passive dispersal is the main, or sole, means of movement for the vast majority of organisms. In plants, morphological appendages such as wings or hairs aid the passive wind dispersal of diaspores (**Box 1(c)**); barbs or spines adhere to mobile animals; fleshy fruits attract frugivores that may digest the fruit and defecate the seeds (**Box 1(d)**). Seed dispersal by volant vertebrates (especially birds and bats), either external (epizoochory) or internal (endozoochory), has traditionally been considered an efficient adaptation for LDD. However, plant seeds can be transported over long distances irrespective of their dispersal morphology; therefore, generalizations about LDD should not be based on seed morphology but on attributes of potential LDD vectors (Nathan *et al.*, 2008b).

Passive dispersal is also frequent among organisms other than plants. Many marine organisms have free-living juvenile stages in the form of cell, eggs, or larvae that are carried passively by ocean currents over large distances. Being concentrated near the water surface, these tiny planktonic propagules (**Box 1(e)**) are dispersed by the wind-driven oceanic circulation. Microbes, spores, light seeds, and tiny insects, spiders, and mites constitute the so-called aerial plankton known to be an extremely effective source of colonization of remote islands. Some spiders, for example, climb to a high point, release silk parachutes into the air, and are subsequently carried off by wind. Dispersal of larger disseminules by winds is usually very local, but strong updrafts and extreme storms can carry them over large distances. Another passive mechanism for dispersal of relatively large organisms is by floating islands or rafts (**Box 1(f)**), which may disperse entire trees, rodents, and reptiles, again, irrespective of their morphological adaptations for dispersal. However, the most influential dispersal agent at the biogeographic scale – at least during the past 10,000 years – is man. Anthropogenic dispersal has moved many species of almost any dispersal adaptation across the globe (Carlquist, 1974),

Box 1 Mechanisms and Adaptations for Dispersal


(a)


(b)


(c)


(e)


(f)


(d)



(g)

(a) Shy albatross (*Thalassarche cauta*) – a vagrant seen off the shore of California in 1999. Photo by Luke Cole. (b) Squirting cucumber (*Ecballium elaterium*). Photo by Pietro Pavone. (c) Seeds of milk thistle (*Silybum marianum*). Photo by Ran Nathan. (d) Fruits of thorny smilax (*Smilax aspera*). Photo by Avinoam Danin. (e) Veliger larva of neapolitan triton *Cymatium parthenopeum*, a teleplanic gastropod (Ranellidae). Photo by Rudolf S. Scheltema.. (f) Floating island of 21 m long, composed mainly of reeds, in Lake Malawi, Africa. Photo by Kenneth R. McKaye. (g) Brown treesnake (*Boiga irregularis*) coiled in an airplane's wheelwell in Guam. Photo by Gad Perry.

either intentionally or not. The arrival of the brown tree snake to Guam is considered to be unintentional (*see* Extinction); years later, a snake was found in an airplane's wheelwell just before leaving Guam for Japan (Box 1(g)).

The probability of LDD is inherently very low, and jump range expansion also requires the unlikely discovery of a suitable site and establishment of a viable population in a remote location. Jump range expansion is thus highly improbable (*see* Components of Colonization Success). This improbability might lead one to doubt the biogeographic importance of LDD (*see* Introduction). However, even extremely low (but nonzero) probabilities will eventually be realized given enough attempts. In fact, there are large numbers of examples of jump range expansion, including, but not restricted to, the colonization of newly formed volcanic islands (*see* Colonization of Virgin Habitats).

The low probability of jump range expansion implies that this phenomenon is highly stochastic and unpredictable; although this generalization is true to a large extent, there are several predictable mechanisms of jump range expansion. Nathan *et al*. (2008b) proposed six generalizations for the key mechanisms driving LDD of plant seeds and other passively dispersed organisms, stating that LDD is generally more common in open terrestrial landscapes, and is typically driven by large and migratory animals, extreme meteorological phenomena, ocean currents, and human transportation (*see* **Box 1**), each transporting a variety of seed morphologies.

Determinants of Dispersal and Colonization Success

Mechanisms and Adaptations for Dispersal

Most (if not all) organisms are able to move or be transported at some stage during their life cycle, exhibiting a diverse array of dispersal mechanisms (**Box 1**). Henry Ridley (1930), Sherwin Carlquist (1974), and Leendert van der Pijl (1982) provide many examples of dispersal-determining traits in animals and plants. Overall, organisms vary greatly in their dispersal ability, the capacity for dispersal, which is determined by a complex set of interacting factors. High dispersal ability can be achieved by either active or passive dispersal (**Box 1**), and it is often associated with relatively low competitive ability and short longevity.

Components of Colonization Success

Successful colonization necessitates that the physical presence of a propagule will be followed by establishment. Both dispersal and establishment require physiological, morphological, and behavioral preadaptations or adaptations enabling them to disperse far from the source, survive during dispersal, and become established in a remote location. In general, the larger the distance between any two locations, the more dissimilar the environmental conditions. Therefore, LDD typically takes propagules to environments that are very different from their usual habitats (Nathan, 2006). Thus, successful colonization in remote locations is generally extremely rare, because it is the product of the very low probabilities of these sequential processes. However, given enough time and multiple attempts, LDD may result in long-distance colonization.

Preadaptations

Species that often combine high dispersal ability with broad ecological tolerance are those adapted to exploit disturbed or newly created environments and those that inhabit ephemeral or uncertain environments. Such species are thus preadapted for successful colonization outside their species range. MacArthur and Wilson (1967) called these species r-strategists, characterized by small size, early maturity, rapid growth, high fecundity, high dispersal ability, and broad niche. In contrast, K-strategists are better adapted for local competition through the efficient use of limited resources. The dichotomous r/K scheme is obviously an oversimplification and the concept is rather controversial (see Williamson, 1996); yet empirical studies generally, but not always, do support the theoretical expectation that r-strategy preadaptations characterize effective colonizers (Brown and Lomolino, 1998; *see* Removal of Barriers).

Physical Access: Dispersal Barriers and Dispersal Routes

Physical access to a remote location depends not only on specific traits that enhance LDD (**Box 1**), but also on the environmental conditions in the area between the source and the deposition location. Crucial to the problem of physical access is that of dispersal barriers, which propagules must overcome during jump range expansion (*see* Jump Range Expansion (Jump Dispersal)). The manner of overcoming barriers defines the dispersal route (*see* Glossary), the other key term in the context of physical access; some routes may be traversed with ease, whereas others rely on chance. Examples of dispersal barriers and dispersal routes are given in the section Removal of Barriers and in other sections that follow.

It is important to emphasize that dispersal barriers are species-specific; that is, conditions that may prevent dispersal of one species may not affect the dispersal of another. They may even be specific for a particular stage in the species' life cycle. Dispersal barriers can be physiological, ecological, or behavioral, with organisms subjected to physiological stresses, ecological hazards, or behavioral difficulties in the environments they traverse.

Dispersal barriers become traversable via dispersal routes. George Simpson (1902–1984) made a distinction between corridors, filters, and sweepstakes routes, which roughly cover the entire range of probabilities of crossing a barrier from highly probable to highly improbable, respectively (Simpson, 1965). Corridor routes allow dispersal of many species from one region to another, conferring a considerable degree of similarity between the biotas of the two sides. Climatic changes and plate tectonics have repeatedly connected previously separated land or water masses, generating dispersal corridors such as those between Africa and Eurasia (*see* Spread of Early Hominins; **Figure 2**). The passage through filter routes is more selective than through corridors. Some species pass rather easily, whereas others do not. As a result, each biota represents a biased subset of the other. The Suez Canal between the Red Sea and the Mediterranean Sea (*see* Removal of Barriers) is an example of a filter route. Sweepstake routes in which crossings occur as a chance event

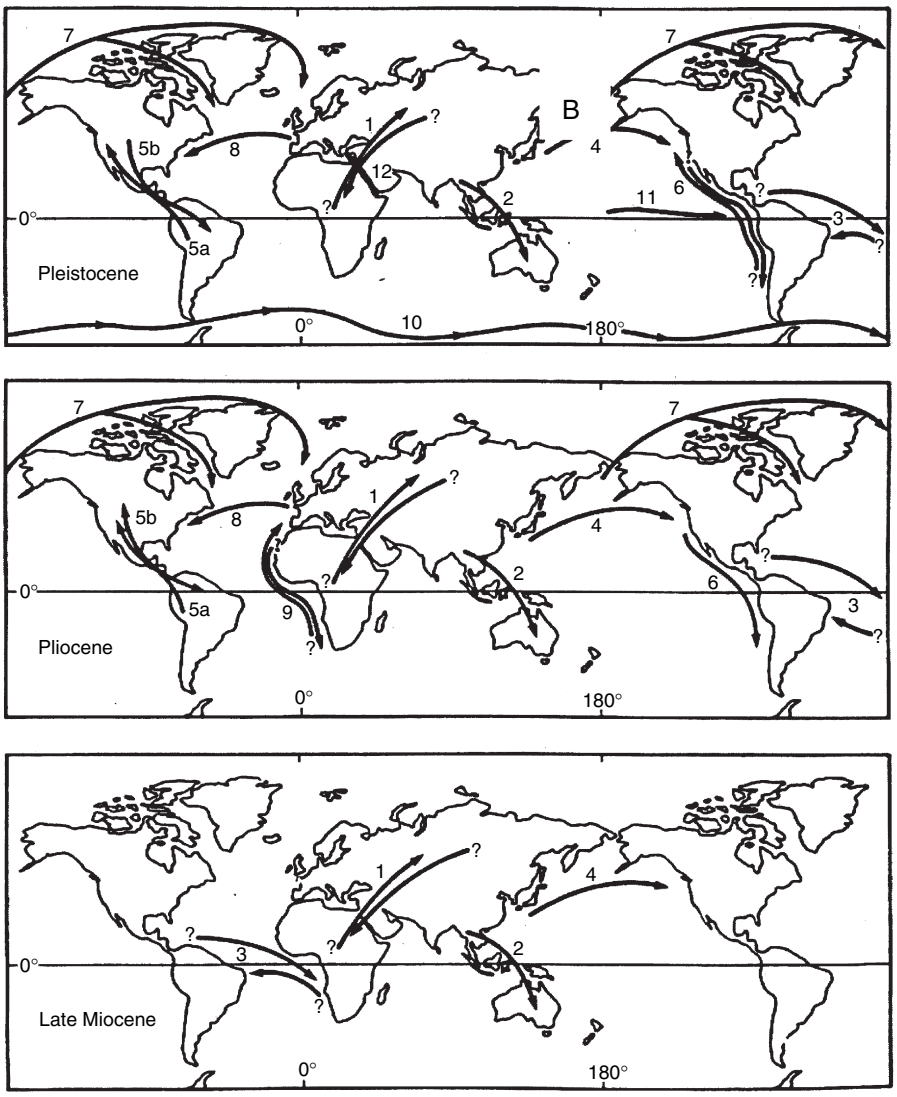


Figure 2 Major episodes of biotic interchange during the late Miocene, Pliocene, and Pleistocene-Recent. Roughly symmetrical interchanges are marked by two-headed arrows, asymmetrical ones by single-headed arrows pointing at the predominant direction, and uncertain direction by question marks. Reproduced from Vermeij GJ (1991) When biotas meet: understanding biotic interchange. *Science* 253: 1099–1104, with permission from American Association for the Advancement of Science.

at a very low frequency, such as by dispersal on floating islands (**Box 1**), represent the most improbable way to cross a barrier (*see* Colonization of Virgin Habitats).

Ecological Opportunity: Postdispersal Establishment

Given that physical access has been attained, nevertheless the probability of successful colonization in remote areas is still rather low. As indicated above (*see* Preadaptions), both the physical access and the establishment of a viable population require physiological and morphological preadaptations; often each characteristic is restricted to either dispersal or establishment, and plays little role in the other. The importance of postdispersal establishment for successful colonization is illustrated by the many examples in the section Methodological Approaches and Case Studies.

Establishment success is determined in part by the ecological conditions at the new site. For passively dispersed species, the interaction between the dispersal vector and the conditions in the particular arrival site matter, whereas actively dispersed propagules may exhibit habitat choice (also called habitat preference) due to their ability to reach preferred habitats among those locally available. Passively dispersed dormant propagules like seeds and invertebrate eggs may exhibit temporal habitat choice by becoming dormant until conditions favorable for establishment become available. Meeting the physical conditions for establishment may not be sufficient since colonizers may still be excluded by biotic interactions such as competition and predation. Thus establishment also requires traits specific to the recipient community, particularly in relation to the resident competitors and predators.

Successful colonization in remote areas requires that colonists be able to reproduce and establish a viable population. Given that LDD is likely to bring only one or very few

reproductive individuals close together in a remote area, the initial population size of a naturally colonizing species is typically very low. Very small populations are prone to extinction due to demographic, genetic, and environmental stochasticity – the influence of random events on population persistence – which is intensified as population size decreases. These risks may be reduced with repeated immigration to the site, by the so-called rescue effects (*see* Extinction). Hence, dispersal ability itself is a factor that determines not only arrival but also, in part, establishment probability. Small, newly established isolated populations exhibit reproductive difficulties due to Allee effects such as difficulty in finding a mate. Such effects are obviously graver for obligate sexual reproducers (higher animals and plants) than for asexual or self-compatible organisms that can reproduce without gametes from another individual. However, the severity of Allee effects might be reduced in some obligate sexual species by certain reproductive mechanisms, such as sperm storage and delayed implantation, which assure sufficient gametes of both types, as well as by social behavior such as flocking.

The Geography of Dispersal

The main issue addressed thus far is the effect of dispersal on the geographic distribution of organisms. The complementary issue of the effects of geography on dispersal is dealt with in the discussion of dispersal barriers and dispersal routes (*see* Physical Access: Dispersal Barriers and Dispersal Routes). This topic is concerned with the geographic variation in dispersal explained by geographic variation in the external factors affecting dispersal. It typically neglects geographic effects on the dispersal traits themselves. Theoretical models predict that dispersal is favored during colonization, but selected against when a population is strongly isolated. The degree of isolation is a relevant factor: (1) when isolation is low, colonization is easily achieved by dispersal. Thus, elevated levels of isolation initially select for greater dispersal abilities; (2) at higher levels of isolation colonization is unlikely, and dispersal is selected against. Although the second prediction has long been recognized and is supported by overwhelming evidence, there have been only a few recent tests of the first.

The prediction that very high degrees of isolation select against dispersal is evident in the loss of dispersal ability on oceanic islands, as in many flightless birds, land snails, insects, and plants (see Carlquist, 1974). Reduced dispersal is also evident in organisms in unique isolated habitats as diverse as subterranean caves, mountaintops, deep-sea trenches, thermal vents, and hot springs. The observation of this phenomenon raises an intriguing paradox, which Darwin (1859) explained through counteracting selective forces acting before and after colonization of a remote island. Cody and Overton (1996) demonstrated the strength of the selection for reduced dispersal potential in weedy, short-lived wind-dispersed plants of inshore islands in Canada. Over the course of just a few generations, island populations of two of three species with sufficiently large sample sizes showed marked reduction in dispersal potential relative to mainland populations or young island populations. Aphids, leafhoppers, crickets, and many other insects have the potential to develop into a normal or flightless adult, depending on the environmental conditions encountered during development.

Habitat fragmentation has substantial effects on dispersal at mesoscales by creating dispersal barriers and disrupting dispersal routes. Studies on British butterflies (Van Dyck and Matthysen, 1999) show that the relative investment in the thorax, which mainly contains flight muscles, is higher when patches are >40 km apart; similar selection appears to have occurred in a region where population expansion was most rapid. No differences were found in the relative investment in flight between sites <5 km apart. Since flight is associated with other functional needs of butterflies such as foraging and mate-location, further study is needed to test the relationship between morphological traits and dispersal ability.

Methodological Approaches and Case Studies

Observing Dispersal and Colonization in Action

Observations of Dispersal

Direct observation is potentially a powerful method of quantifying dispersal. Turchin (1998) provided an extensive review and explanation of the various methods used to track the movements of individuals and to estimate population redistribution in both animals and plants, including records of movement paths and mark-recapture methods. These methods are typically laborious and, from a biogeographic perspective, where large spatial or temporal scale are most important (*see* The Biogeographic Perspective), nearly impossible. The indirect approach, in which dispersal is estimated from its outcome (recruitment), is more feasible. The greater feasibility of indirect methods is a result of the larger size and more conspicuous nature of most mature organisms in relation to their propagules; moreover, it is usually easier to observe a pattern (recruitment) than a process (dispersal). As emphasized in the section Components of Colonization Success, dispersal is meaningless without subsequent recruitment; therefore, indirect evidence of dispersal is indeed important. Nevertheless, it does not tell us much about the dispersal process itself, since the distances dispersed, the number of dispersal events, and the mechanisms of unsuccessful recruitment are unknown. These are important parameters that should be incorporated into biogeographic models in order to enhance their generality and predictive ability.

Data on dispersal events are best obtained by following movements of individuals for which both source and end points are known. Unfortunately, for practical reasons, these kinds of data are relatively rare. In many cases, only the end location is known and there are multiple potential sources. Although our knowledge of LDD continues to increase, the paucity of direct observations of these dispersal events is one of the largest obstacles to the development of dispersal biogeography as a mature scientific discipline. One possible solution is the development of methods to track individual movements at both small and large spatial scales. Advances in developing molecular and remote sensing techniques promise significant progress in the near future and have already yielded new insights into dispersal processes. For example, using a variety of molecular tools, Vilà *et al*. (2003) showed that even

a single immigration event of an individual gray wolf (*Canis lupus*) can recover genetic diversity in a severely bottlenecked and geographically isolated population in Norway. However, the origin of this wolf has been debated, since the closest known population lies >900 km away in Finland, beyond the known dispersal distance of this species. Doubts have lessened given that a young GPS-tracked female wolf has been recorded as having traveled a total straight-line distance of >1000 km from Norway to the Finish–Russian border (Wabakken *et al.*, 2007). Note that in genetic studies, the term "direct evidence" is used for evidence of gene flow obtained by comparing the genotypes of recruits and their putative parents (and then applying methods such as paternity analysis to identify the most probable parent). That would be considered as indirect evidence under the preceding definition. However, for the study of LDD this approach is not very valuable. Indentifying putative parents becomes impractical as the distance from the offspring increases. In contrast, the so-called indirect genetic approach, which compares the genetic structure between populations, has provided compelling evidence for jump range expansions (*see* Holocene Postglacial Spread).

There are two types of methods for estimating the dispersal distances when there are multiple possible sources (reviewed in Nathan and Muller-Landau, 2000). The first makes *a priori* assumptions about the source location (of which the simplest assumes the nearest possible parent is the source), whereas the second uses inverse methods to fit a particular dispersal kernel (a functional form of the probability distribution of dispersal distances – *see* Dispersal Models). Both methods have been applied to study dispersal at the ecological scale. Since short-distance dispersal governs gradual range expansion, it is also important from the biogeographic perspective. However, using these methods to estimate LDD is problematic since the exact position along the dispersal kernel tail is typically very sensitive to small differences in probability, thus sampling errors are likely to obscure the actual probabilities. Furthermore, the number of endpoints is typically very small (in most cases there is only one), which prohibits inverse modeling techniques. With a better understanding of the process of LDD itself, mechanistic dispersal models help surmount this difficulty (*see* Dispersal Models).

Colonization of Virgin Habitats

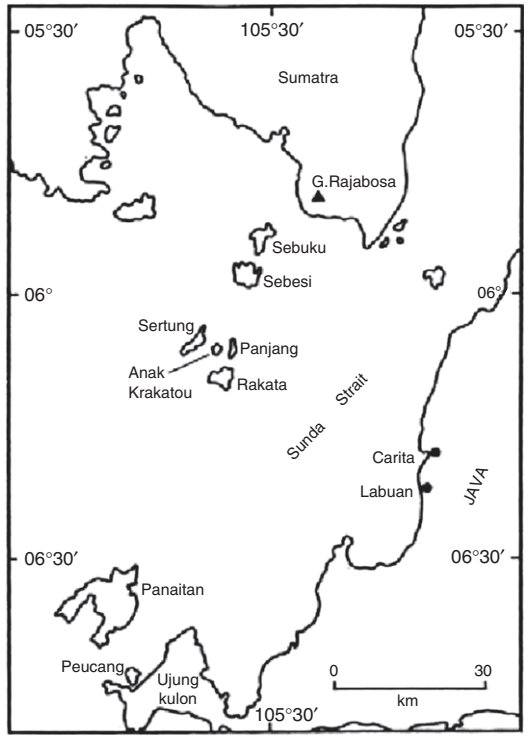
Volcanic eruptions that eliminate an entire local biota provide unique opportunities to study patterns of dispersal and colonization. Although there are some well-studied cases of recolonization in terrestrial systems after volcanic eruptions, such as the 1980 eruption of Mt. St. Helens (western North America), the best case studies are of oceanic islands. A well-known example is the island of Surtsey near Iceland that emerged from the sea in 1963. The best example, however, has its roots 80 years earlier. On 27 August 1883, the island of Krakatau, located between Sumatra and Java (**Figure 3(a)**), exploded with tremendous force, obliterating all traces of life. Less than 2 months after the eruption, a Dutch expedition found no evidence for any surviving life form. The first animal, a single tiny spider that had dispersed as aerial plankton (**Box 1**), and the first plant, were detected 9 and 13 months after the eruption, respectively. Later biological surveys documented the rapid recolonization of the island by various life forms, in a complex successional process, which began with fungi and lichens breaking down the volcanic rock to form soil, enabling seeds to germinate. The possible sources of immigrants were the mainland areas of Java and Sumatra, both approximately 40 km away, and two small islands, 12 and 20 km away (**Figure 3(a)**). More precise identification of the sources for colonization in Krakatau by the methods applied thus far is highly speculative. Furthermore, most species were discovered after they had already established populations, with no data on their actual arrival; obviously, there are very few data on arrivals not followed by recruitment. Yet, even with these restrictions, the main insights from this unique long-term classical case study are extremely important; Thornton's (1996) book tells the story of the reassembly of this island ecosystem.

The immigration data collected on Krakatau, albeit incomplete, provide important information on the dispersal process itself. One of the most important modes of dispersal was by sea, whether passively (e.g., fruits and seeds) or actively (e.g., the water monitor *Varanus salvator*). There was a striking example of rafting (**Box 1**) in a 20-m^2-area floating island with 3–4-m high palms observed in the archipelago in 1986. Dispersal by air, either passively (spores, light seeds, and other aerial plankton) or actively (butterflies, birds, and bats), was another important mode. Birds and bats served as dispersal agents for plant seeds over short (within the archipelago) and long (from the mainland to the archipelago) distances.

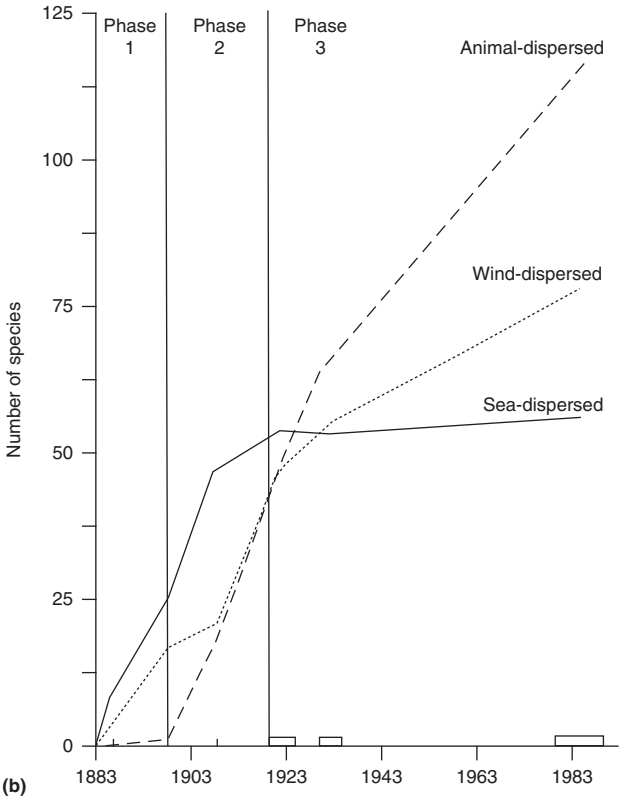
One of the most remarkable patterns observed in the Krakatau system is that of unequal colonization. This pattern was largely clarified for plants when dispersal agents were classified into three primary types (wind, sea, and animals) and examined in three time phases since the eruption (**Figure 3(b)**). Sea-dispersed species were the most important early colonizers but became the least important by the third phase; animal-dispersed species generally showed the opposing trend; the relative importance of wind-dispersed species remained fairly constant in all phases. The mode of dispersal also explains why several kinds of taxa were underrepresented. Species with large and heavy seeds dispersed either by wind (wings) or by bats and terrestrial mammals (**Figure 3(c)**). Moreover, some species reached, but did not colonize, the archipelago. Over the period of 8 years (1984–1992), 54 species of animals were recorded as new arrivals to Anak Krakatau (a small island in the center of the archipelago, which emerged from the sea in 1928; **Figure 3(a)**). Fourteen were new to the archipelago, whereas more than 20 did not establish. The absence of several sea-dispersed plant species is attributed to postdispersal conditions. The role of establishment conditions is also reflected in the decrease of the colonization rate with increasing species richness.

Removal of Barriers

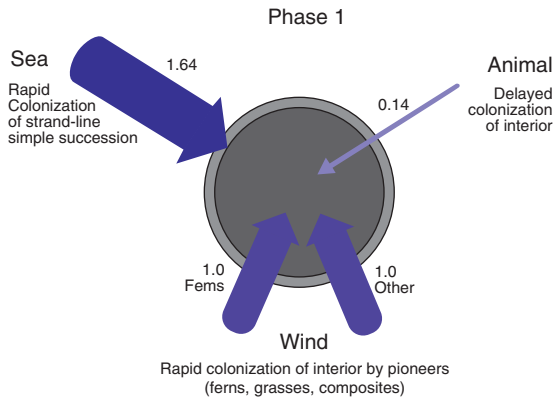
Dissolution of barriers is typically accompanied by species invasion in both directions. Although the term "biotic interchange" might be interpreted as a balanced exchange, the pattern is often highly asymmetric. Climatic changes and tectonic plate movements in geological times have repeatedly created and destroyed dispersal barriers, and both kinds of events have had important biogeographic and evolutionary consequences (**Figure 2**). Such events include the closure of

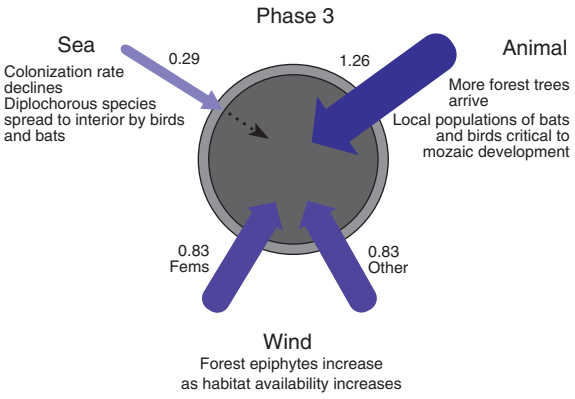

05°30′
105°30′
05°30′
Sumatra
G.Rajabosa
Sebuku
Sebesi
06°
06°
Sertung
Panjang
Anak
Krakatou
Rakata
Strait
Sunda
Carita
Labuan
JAVA
06°30′
06°30′
Panaitan
Peucang
Ujung
kulon
0
30
km
105°30′

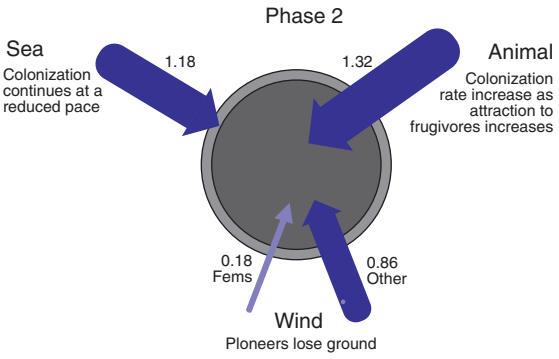
(a)

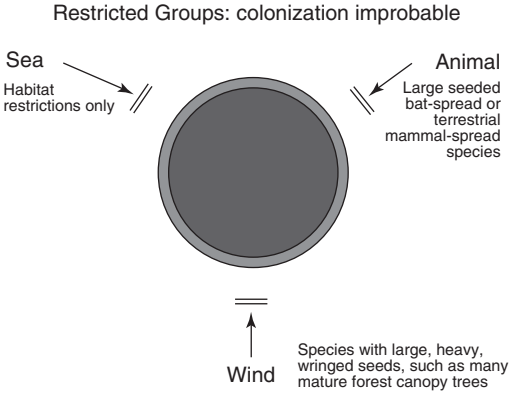

125
100
75
50
25
0
Number of species
Phase 1
Phase 2
Phase 3
Animal-dispersed
Wind-dispersed
Sea-dispersed
1883
1903
1923
1943
1963
1983

(b)


Phase 1
Sea
Rapid
Colonization
of strand-line
simple succession
1.64
Animal
Delayed
colonization
of interior
0.14
1.0
Fems
1.0
Other
Wind
Rapid colonization of interior by pioneers
(ferns, grasses, composites)


Phase 3
Sea
Colonization rate
declines
Diplochorous species
spread to interior by birds
and bats
0.29
1.26
Animal
More forest trees
arrive
Local populations of bats
and birds critical to
mozaic development
0.83
Fems
0.83
Other
Wind
Forest epiphytes increase
as habitat availability increases


Phase 2
Sea
Colonization
continues at a
reduced pace
1.18
1.32
Animal
Colonization
rate increase as
attraction to
frugivores increases
0.18
Fems
0.86
Other
Wind
Ploneers lose ground


Restricted Groups: colonization improbable
Sea
Habitat
restrictions only
Animal
Large seeded
bat-spread or
terrestrial
mammal-spread
species
Wind
Species with large, heavy,
wringed seeds, such as many
mature forest canopy trees

(c)

the Tethyan seaway and the formation of a wide land bridge between Africa and Asia during the Miocene Epoch 18 million years ago (Ma). Likewise, the reoccurring Bering land bridge and strait have intermittently served as a dispersal route between the terrestrial biota of northeastern Asia and northwestern North America, and between the marine biota of the north Pacific and Arctic–Atlantic basins. Another well-studied example is the current land bridge between South and North Americas, which was formed 3.5 Ma; it provided a dispersal route for a massive passage – the Great American Interchange – of terrestrial organisms, and a dispersal barrier for the marine biota of the Atlantic and Pacific oceans. These examples demonstrate how dispersal and vicariant events can take place simultaneously.

Since the fossil record of mammals is richer than that of other vertebrates, the passage of mammals during the Great American Interchange has been the most thoroughly documented. It is clearly asymmetric, with many more North American taxa passing through the filter and diversifying in South America than the converse. In other terrestrial vertebrates, however, the interchange was either roughly balanced (e.g., birds) or biased in the opposite direction – that is, predominantly from south to north (reptiles, amphibians, and even freshwater fishes). Indeed, the differential response to the creation of the land bridge shows large variation also at lower taxonomic levels; in general, most taxa in both continents actually did not expand their ranges to the other continent.

Of the several episodes of biotic interchange that have occurred recently, perhaps the one most intensively studied is that through the Suez Canal, which connects the Mediterranean Sea to the Red Sea. Before the canal was opened for navigation in 1869, only a very few euryhaline species were able to cross the hypersaline lakes that have occasionally connected the two seas (Por, 1978). The creation of the Suez Canal led to an almost unidirectional colonization named Lessepsian migration after Ferdinand de Lesseps, the designer and chief constructor of the canal. It involves the colonization of the Mediterranean by at least 120 mollusks, 70 crustaceans, and more than 80 fish species from the Red Sea. Two main explanations for the unidirectional colonization have been proposed. The first, that species from the Gulf of Suez, the saline arm of the Red Sea at the southern end of the canal, were better preadapted physiologically to cross the hypersaline canal and lakes than were species from the northern end of the Mediterranean. The second postulates that species from the more diverse biota of the Red Sea are superior in competition, defense, or reproduction to species of the impoverished biota of the eastern Mediterranean. A similar explanation has also been raised to explain the asymmetry in the passage of mammals during the Great American Interchange. Comparisons between species that colonize and closely related species that do not colonize show, in general, a tendency for r-selection attributes (*see* Preadaptations) among colonizers, especially higher dispersal ability and fast growth.

Natural Range Expansion

Recent Range Expansion

Well-documented major range expansions provide a unique opportunity to quantify and model the process of spatial spread at large scales. Charles Elton (1900–1991) and John G. Skellam (1914–1979) laid the foundations for the dynamic progress recently experienced in this field. Most range expansions at present are related to human activity (*see* The Spread of Alien Species), but some that are presumably natural range expansion are of particular interest to explore how species expand their range without human-aided transportation (see Hengeveld, 1989; Shigesada and Kawasaki, 1997; Williamson, 1996).

Range expansion can be very abrupt and may be followed by a rapid contraction, a phenomenon called boom-and-bust (Williamson, 1996). The most remarkable examples are three African locusts (*Schistocerca gregaria*, *Locusta migratoria*, and *Nomadacris septemfasciata*), which usually inhabit limited areas, but which reproduce rapidly, change their morphology and behavior, and disperse in huge swarms to forage over much larger areas under particularly favorable weather and food conditions. Populations of several rodent species such as lemmings (*Lemmus* spp.) also exhibit marked (cyclic) fluctuations, but due to their lower dispersal capacities, the resulting geographical impact is much more limited. Typically, these outbreaks do not result in expanding the breeding distribution of these species, as occasionally occurs after outbreaks of some bird species such as crossbills (*Loxia curvirostra*) and Pallas' sandgrouse (*Syrrhaptes paradoxus*).

The accidental occurrence of individuals far away from the breeding sites has been documented in many taxa, especially in volant animals. For example, in 1979, the author observed an accidental arrival to Elat (southern Israel) of a female long-tailed duck (*Clangula hyemalis*) whose southernmost breeding and wintering regions in Europe are in Sweden and northern France, respectively. Fifteen months later, at the very same salt pond, the author collected an exhausted shy albatross (**Box 1**), a rare seabird that breeds on islands near Tasmania and New Zealand. Some accidental vagrants may establish viable populations (jump range expansion), as did the purple gallinule (*Gallinula martinica*), a colorful waterfowl, in the Tristan da Cunha Islands at the Mid-Atlantic Ridge after years as a vagrant visitor. This tropical species has been reported as a rare vagrant as far north as Nova Scotia in Canada.

Events of boom-and-bust and of vagrant occurrence illustrate three important points: first, that species differ in their dispersal ability (*see* Mechanisms and Adaptations for

Figure 3 Colonization of plant species to the Krakatau archipelago between Sumatra and Java (a). The number of species colonizing the island of Rakata increased over the period 1883–1989. Species with different primary dispersal mechanisms (sea, animal, and wind) exhibit dissimilar colonization dynamics more than three phases of this period (b). A model of island colonization derived from these empirical data (c) incorporates the dispersal mechanisms and temporal phases with a distinction between strand-line (outer circle) and interior habitats, and emphasizes colonization constraints (the lower right figure). Arrow widths are proportional to the cumulative number of species in species per year (values are given by each arrow). Reproduced from Bush MB and Whittaker RJ (1991) Krakatau: colonization patterns and hierarchies. *Journal of Biogeography* 18: 341–356, with permission from Blackwell Science Ltd, and Whittaker RJ and Jones SH (1994) Structure in rebuilding insular ecosystems: An empirically derived model. *Oikos* 69: 524–530, with permission from Wiley.

Dispersal; **Box 1**); second, that the current breeding distributions of these species is not limited by dispersal *per se* (*see* Physical Access: Dispersal Barriers and Dispersal Routes); and third, that postdispersal factors are important determinants of colonization success (*see* Ecological Opportunity: Postdispersal Establishment). However, these events are episodic and, thus, difficult to study. In contrast, some persistent cases of recent broad-scale range expansions that have been documented in detail are important contributions to the study of dispersal and population spread. As in the previous case, the best examples come from highly mobile animals. One well-studied case is the collared dove (*Streptopelia decaocto*), which, since 1900, has expanded its range from Turkey through the Balkans to the rest of Europe. Another is the cattle egret (*Bubulcus ibis*), which expanded its range in all continents since the late 1800s. It naturally colonized South America from its breeding grounds in Africa by the mid-1930s (jump range expansion) and rapidly expanded throughout the New World. For a more detailed discussion of range expansion, *see* The Spread of Alien Species *and* Modeling.

Holocene Postglacial Spread

The considerable fluctuations in the global climate during the Pleistocene (the geological epoch between 1.8 Ma and 11 thousand years ago (Ka)) have led to several glacial–interglacial cycles. The last glacial period reached its maximum about 18 Ka, when ice sheets covered extensive parts of Eurasia and North America. Species responded in various independent ways, depending on their physiology, life history strategies, and dispersal abilities. Many species were able to track the shift of their most favorable climate and habitats, experiencing range contraction or range shift southward. From about 15 Ka, the climate warmed, the ice retreated, and species expanded their range northward from refugia. Valuable data on this range expansion during the Holocene (the last 11,000 years) come from insect exoskeletons, plant macrofossils, and pollen cores collected from many localities. Analyses of genetic diversity provide another important source of information about patterns of postglacial spread.

The postglacial expansion of trees in North America and Europe has been studied extensively based on fossil pollen cores from lake sediments across a wide geographical range, dated by radiocarbon (14C) analysis. The results are often expressed as isochrone maps in which contours join sites where similar pollen-stratigraphical events occurred at the same time, and as rates of range expansion during a given period (**Table 1**). These interpretations, however, should be made with care since pollen of many tree species are effectively dispersed over large distances; pollen data also contain a nonlocal component. Estimation of this component is possible through data on other plant remains (seeds, fruits, leaves, twigs, wood), the so-called plant macrofossils, that are less likely to be transported over large distances. The relatively few studies in which macrofossils have been well documented showed that the main conclusions based on pollen data alone were fairly robust. Yet, recent genetic and macrofossil studies cast doubts about the generality of these conclusions, revealing the existence of cryptic higher latitude refugia for a variety of plant and animal species.

The paleoecological pollen record shows an individualistic response of tree species, in the timing, direction, and speed of their spread. This variation can be partially explained by factors other than dispersal; for example, the existence of small cryptic high latitude refugia reasonably indicates adaptations to survive in cold environments. Yet, the paleoecological evidence also shows that species were able to cross dispersal barriers as large as the North American Great Lakes, the English Channel, and the Baltic Sea. In particular, it provides evidence that range expansion of trees can be very rapid with estimated average rates of spread of a few hundred meters per year (**Table 1**) and more. These estimated rates greatly exceed the rates expected from the observed dispersal rates of these species, collected at ecological spatial and temporal scales (*see* The Evolutionary and Population Genetics Perspectives). The discrepancy has been a long-standing puzzle for plant ecologists, termed Reid's paradox after Clement Reid who described it for British oaks. Even more puzzling than the rate of expansion of oaks – whose acorns may occasionally be dispersed over large distances by birds – is the rapid rate of expansion of woodland herbs that mostly propagate clonally and possess no obvious mechanism for LDD (Cain *et al.*, 1998). Indeed, the existence of cryptic high latitude refugia suggests that for those species postglacial spread rates were much slower than those inferred from the fossil pollen record.

The discussion of Reid's paradox has recently been revived because of the need to predict biological responses to human-induced climate warming. In fact, evidence on population expansions of a variety of plants and animals (Parmesan and Yohe, 2003) indicate that global climatic warming is already reflected in considerable poleward shift of species' distributions. Consequently, scientists have developed new mathematical models for seed dispersal that incorporate the long-distance component as well (*see* Modeling). The next step is development of frameworks capable of incorporating projected scenarios of dispersal, demographic and environmental changes into spread models to predict species spread in future environments. A recent example applied to predict the future spread rates of North American wind-dispersed tree species showed that spread rates on the order of > 50 m yr^{-1} are very unlikely under most conditions and for most species (Nathan *et al.*, 2011a).

Another advance in the study of postglacial range expansion follows the progress in molecular genetics that enables assessment of present genetic structure; that is, in turn, used to infer past colonization events (*see* Inference from Current Distributions). Godfrey Hewitt (1999) summarized the results of studies dealing with the postglacial colonization of the European biota (both animals and plants) from three main refugia at the eastern (Balkans), central (Italian), and western (Iberian) peninsulas of southern Europe. He proposed three main expansion patterns, which were termed, after their exemplars, "hedgehog," "grasshopper," and "bear" (**Figure 4(c)**): the hedgehog (*Erinaceus* spp.) expanded with three genomes advancing north from the three refugia; the grasshopper (*Chorthippus parallelus*) colonized mostly from a eastern refugium, with other genomes blocked at the Pyrenees and Alps; the bear (*Ursus arctos*) colonized from the east and the west, with its central refugium blocked at the Alps.

Table 1 Observed rates of range expansion, arranged in decreasing order within each taxonomic group

Species	*Period*	*Region*	*Average expansion rate (km y^{-1})*	*Reference*
Mammals				
European rabbit	1870–1900	Australia	54	Myres, 1970
Oryctolagus cuniculus				
Muskrat		Czechoslovakia		Andow *et al.*, 1990
Ondatra zibethicus	1905–1947	to east-southeast	25	
	1905–1948	to north	12	
	1932–1954	France Three localities	0.9–4.6	Andow *et al.*, 1990
	1923–1938	Finland Four localities	3.5–6.7	Andow *et al.*, 1990
Gray squirrel	1965–1981	Britain	7.7	Okubo *et al.*, 1989
Sciurus carolinensis				
Red deer	1861–1910	New Zealand	1–1.6	Shigesada and Kawasaki, 1997
Cervus elaphus	1910–1920		4.3	
California sea otter	1938–1972	US, California		Lubina and Levin, 1988
Enhydra lutris		to south	3.1	
		to north	1.4	
Himalayan tahr	1936–1966	New Zealand	0.7	Shigesada and Kawasaki, 1997
Hemitragus jemlahicus				
Birds				
Cattle egret	1950–1970	South America to North America	106	Van den Bosch *et al.*, 1992
Bubulcus ibis				
House sparrow	1864–1888	US	66	Okubo, 1988
Passer domesticus				
European Starling	1900–1915	US	11	Okubo, 1988
Sturnus vulgaris	1915–1950		51	
Collared dove	1928–1963	Europe	44	Van den Bosch *et al.*, 1992
Streptopelia decaocto				
House finch	1947–1962	US	3.5	Shigesada and Kawasaki, 1997
Carpodacus mexicanus	1962–1979		21	
Insects				
Africanized bee	1957–1992	South America to North America	300–500	Winston, 1992
Apis melifera scutellata				
Rice water weevil	1976–1986	Japan		Andow *et al.*, 1993
Lissorhoptrus oryzophilus		to northeast	28–470	
		to west	47–250	
Cabbage white butterfly	1860–1883	Canada, Quebec	15–170	Andow *et al.*, 1990
Pieris rapae				
Cereal leaf beetle	1962–1979	US, Michigan	27–90	Andow *et al.*, 1990
Oulema melanopus				
Gypsy moth	1900–1916	North America	9.5	Liebhold *et al.*, 1992
Lymantria dispar	1916–1965		2.8	
	1966–1990		21	
Vascular plants[a]				
Alder	8.5–5.5 Ka	British Isles	0.33	Birks, 1989
Alnus glutinosa	8.5–7 Ka		0.55	
	7–6 Ka		0.15	
Aspen	20 Ka–present	North America	0.26	Delcourt and Delcourt, 1987
Populus spp.	14–12 Ka		0.54	
	10–8 Ka		0.15	

(*Continued*)

Table 1 Continued

Species	*Period*	*Region*	*Average expansion rate (km y^{-1})*	*Reference*
Oak	9.5–6 Ka	British Isles	0.21	Birks, 1989
Quercus spp.	9.5–8.5 Ka		0.48	
	7–6 Ka		0.04	
	20 Ka–present	North America	0.13	Delcourt and Delcourt, 1987
	16–14 Ka		0.18	
	10–8 Ka		0.04	
Bacteria and viruses				
Bubonic plague (black death) *Yersinia pestis*	1347–1350	Europe	320–650	Noble, 1974
Rabies Lyssavirus	1939–1983	Europe	30–60	Shigesada and Kawasaki, 1997

[a]Postglacial expansion rates of trees during the late Quaternary, based on fossil pollen records. The first row in each data set shows the overall average rate during the entire period examined, the second row shows the period with the fastest rate, the third shows data from the same period (in North America and British Isles separately), for comparison.

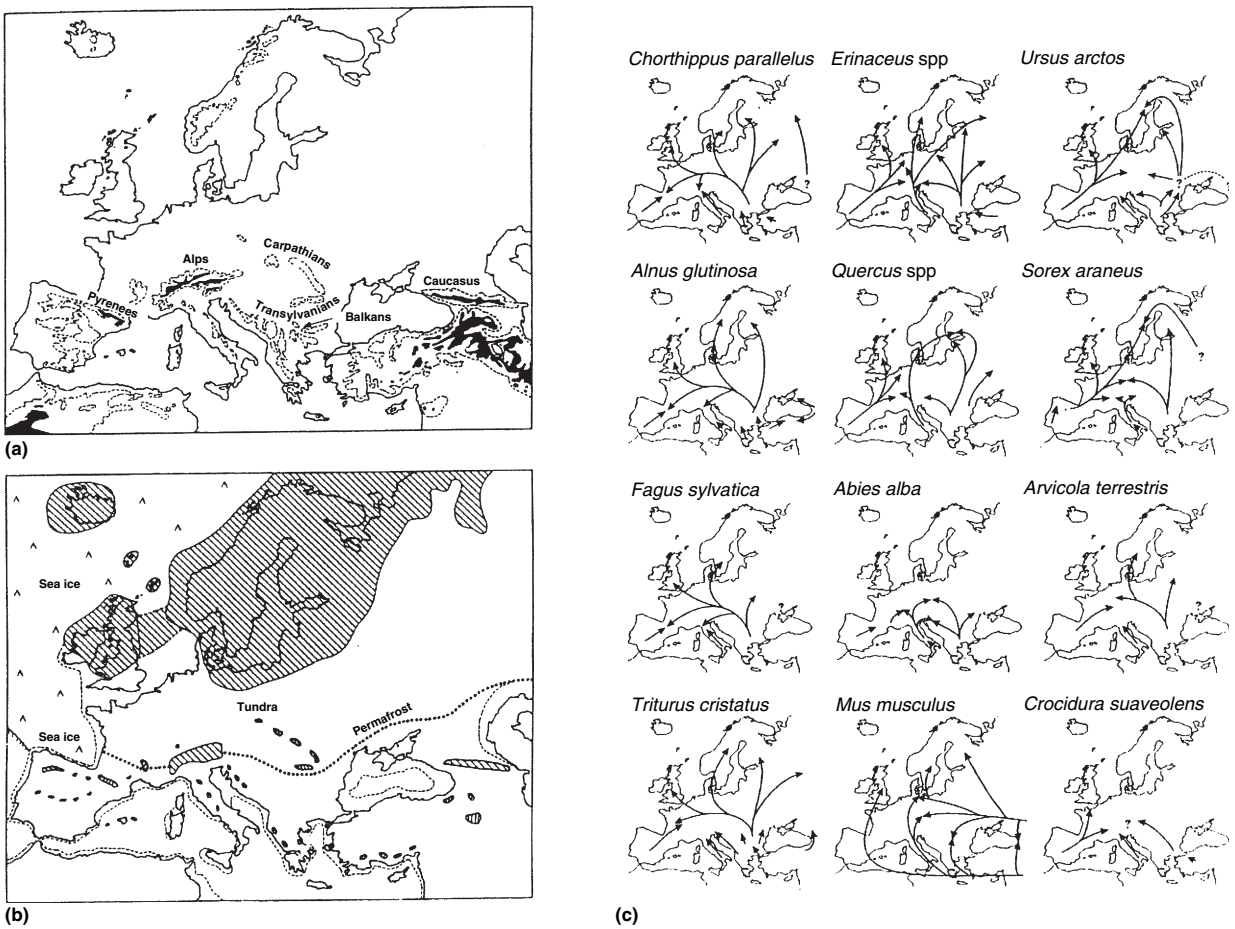


Figure 4 Postglacial recolonization of European biota. The physical geography of Europe (a) is characterized by mountain ranges running from the Pyrenees eastward through the Alps to Caucasus, marked as black regions (>2000 m) and dashed lines (>1000 m altitude). By the end of the Pleistocene ice age (b) extensive ice sheets (hatched) covered northern Europe, and a permafrost plain extended southward to the mountain ranges. The rapid postglacial expansion from the three southern peninsular refugia (c), as deduced from molecular data and fossil evidence, followed three main patterns as illustrated by four examples in each column (see text for details). Reproduced from Hewitt GM (1999) Postglacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87–112, with permission from Wiley.

Spread of Early Hominins

Modern man (*Homo sapiens sapiens*) is currently the species with the most widespread geographical distribution, and other globally distributed species are mostly human commensalists. The dispersal biogeography of humanity is a very intriguing subject; it is also an excellent case study that demonstrates many essential aspects of dispersal biogeography. The following discussion briefly summarizes the patterns of dispersal and range expansion of early hominins (note that there still is some argument over this topic).

The ancestral source location of early hominins is considered to be the plains of the eastern Rift Valley in Africa (now Ethiopia, Kenya, and Tanzania) where fossils of early, woodland-dwelling, hominins were found (*Ardipithecus kadabba*, 5.8–5.5 Ma, *Ardipithecus ramidus*, 4.5–4.3 Ma, *Australopithecus anamensis*, 4.2–3.9 Ma, and *Australopithecus afarensis*, 3.6–2.9 Ma). The cooling period during the Middle Pliocene (3.0–2.4 Ma) marks the emergence and dispersal of the genus *Homo*. The traditional hypothesis postulates that the earliest *Homo* species, omnivorous scavengers who lived in open habitats between 2.5 and 2.0 Ma, shifted from forests to open landscapes, like many other mammalian species, in response to expanding dry savannah. An alternative hypothesis emphasizes the behavioral and phenotypic plasticity of early hominins, enabling them to cope with environmental changes and a variety of habitats, rather than a particular environment. Two well-documented *Homo* species, *H. habilis* (since ~2.3 Ma) and *H. erectus* (since ~1.95 Ma), and possibly other *Homo* species such as *H. ergaster* coexisted in the African Rift Valley until ~1.4 Ma. The *H. erectus/ergaster* type (also called the erectine group) had a long torso and limbs, narrow hips, a large brain, reduced dentition, and elementary stone technology. The findings of much bone and stone refuse near lakes and streams suggest that early hominins congregated within catchment basins and utilized local resources until they were exhausted, and then moved to another catchment, and so on. All of these attributes served as morphological and behavioral preadaptations (*see* Preadaptations) for successful dispersal to and colonization of similar habitats outside Africa during the Pliocene (**Figure 5**). It has long been hypothesized that humans and certain animals (e.g., large bovids) emerged and dispersed together in response to environmental changes. Recent studies, however, emphasize the role of tool production and socialization, which enabled early hominin groups to succeed in a broad range of environmental conditions, as the key features leading to earliest dispersals out of Africa.

Among the earliest fossils of *Homo* found outside Africa, those found in Dmanisi (Georgia), dated ~1.8 Ma and defined as a distinct species (*H. georgicus*), are currently considered the oldest reliable record of *Homo* outside Africa. *Homo georgicus* was a short bipedal hominin with a small brain and, overall, a very primitive morphology resembling both *H. habilis* and erectines. The dating of early erectine skulls found in Mojokerto and Sangiran (Java, Indonesia) is still controversial with estimates ranging from 1.8 to 1.2 Ma for the former and 1.7 to 1.0 Ma for the latter. Skull remains found in Ubeidiya (Israel) are dated 1.4 Ma. The Dmanisi and Ubeidiya sites are rich with lithic artifacts and faunal remains assumed to be indicative of hominin presence; similar findings were also reported from Renzidong (China, 2.5–2.0 Ma) and Riwat (Pakistan, 2.4–1.9 Ma) and from Erq-el-Ahmar (Israel), Pabbi hills (Pakistan), and Longgupo (China), all dated 1.95–1.77 Ma. Thus, the first departure from Africa occurred



Figure 5 Dispersal routes that allowed passage from Africa across the Middle-East into South and East Asia during the late Pliocene and the main sites where human remains were discovered.

about 2 Ma, perhaps even earlier. During this period, there were wide land bridges between Africa and Eurasia at the Sinai Peninsula, and at the straits of Bab-el-Mandab between the Red Sea and the Indian Ocean. Probably both served as dispersal routes for *Homo*. Since this early departure from Africa 2 Ma, early hominins successively dispersed from Africa to Asia, and later to Europe, where the earliest *Homo* fossil found (in Atapuerca hills, Spain) is dated 1.2 Ma. Evidence shows that *Homo sapiens*, our own species, also originated in Africa from where it eventually colonized a much broader range than any of its extinct predecessors.

The Spread of Alien Species

Humans can facilitate the physical access of organisms from one region to another by breaking dispersal barriers or by providing dispersal routes. Species that have colonized non-native regions due to human activities are called alien species, and those that have been deliberately transferred are called introduced species. Many closely related terms, including adventive, exotic, and invasive species, have been used confusingly in the literature. Though examples are numerous, the spread of alien species represents merely a small, though growing, fraction of the total number of incidents in which humans have intentionally or accidentally facilitated physical access. Further, as a general rule – termed "propagule pressure" by Mark Williamson (1996), which is principally equivalent to the rescue effect (*see* Extinction) – single or very few propagules are unlikely to invade. Increasing the number of propagules (e.g., by repeated introduction attempts) increases the chance of an alien species to establish and spread.

The literature provides a wide array of case studies of various taxa across the entire globe, from zebra mussels and starlings in North America to rabbits and myxoma virus in Australia (**Table 1**). Here the classical case study of the muskrat (*Ondatra zibethicus*) has been chosen to illustrate the implications from such events to dispersal biogeography.

The muskrat is a large semiaquatic rodent native to North America (**Figure 6**). In 1905, five muskrats escaped from a fur-breeding farm near Prague in Czechoslovakia. Within 50 years, their descendants and others inhabited large parts of the



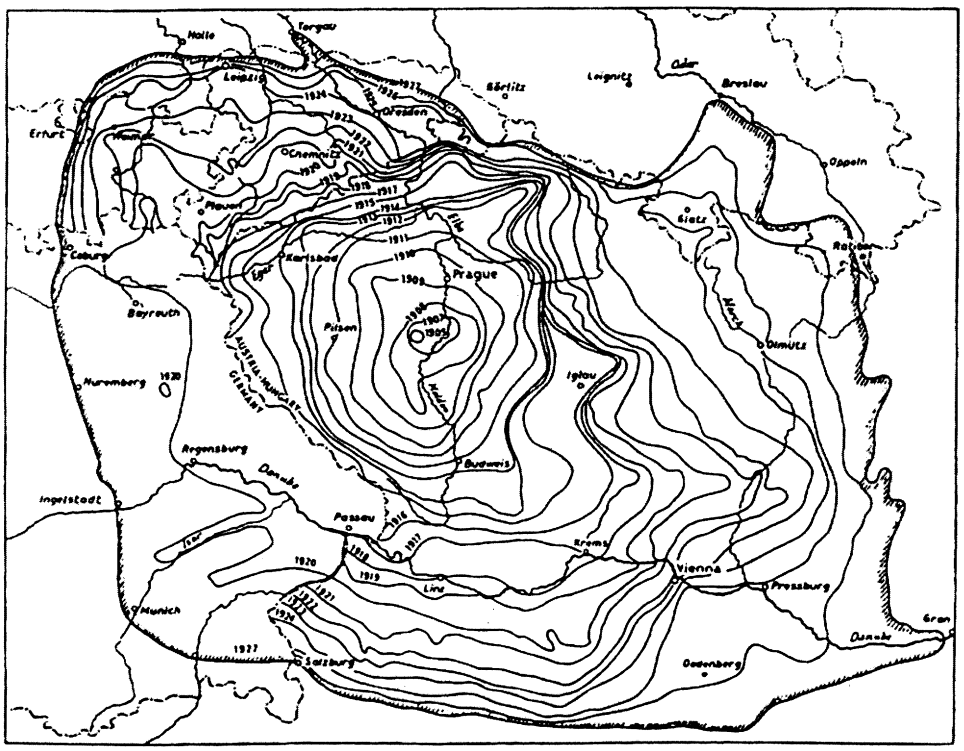


Figure 6 The spread of muskrats in central Europe, showing boundaries of apparent expansion for various years. Reproduced from Figures 4.3 and 4.4 in Williamson M (1996) *Biological Invasions*. Dordrecht, Boston, London: Kluwer Academic Publishers, with permission from Springer.

European continent (**Figure 6**), despite eradication programs initiated due to muskrat damage to roads, dikes, and agriculture. Expansion of the muskrat appeared to be rapid and of constant rate (*see* Reproduction–Dispersal Models for quantification of this case study). At finer scales, however, expansion was fairly variable in space and, to a lesser extent, in time (**Table 1**). Overall, the expansion was closely related to the geographical structure of the species' favorable wet habitats. River valleys, one kind of favorable habitat, provided dispersal routes, whereas mountain ranges acted as dispersal barriers, forming a starlike expansion pattern. Interestingly, favorable habitats do not necessarily increase, but may also decrease the rate of spread, as happened in regions with abundant wet habitats where muskrats settled and built up huge populations. Understanding the interplay between the processes of dispersal and of population growth is critical to the understanding of range expansion (*see* Modeling). Temporal changes should also be taken into account, since expansion rates were lower in dry years when the species' favorable habitats were more restricted.

Inference from Current Distributions

Early biogeographers have frequently compared species composition of different regions to evaluate their biogeographic affinities. This analysis, later scrutinized to include various quantitative methods, has revealed a common nonrandom pattern called nestedness, in which small biotas contain a nonrandom subset of the species in richer ones (Patterson and Atmar, 1986). Found in various taxonomic groups, nestedness is generally attributed to differential extinction and colonization, or a combination of both, though other factors may also be responsible. Differential extinction is arguably the prevalent mechanism of nestedness, especially in fragmented biotas. However, several biotas assembled by colonization also exhibited nested patterns. Some studies suggested that nestedness reflects differential dispersal ability coupled with geographical isolation. This suggestion can be tested by ranking the subsets (e.g., islands) according to their distance from the presumed source (e.g., mainland), and by examining the correlation between individual species contributions to the observed nestedness and species dispersal ability. Kadmon (1995) used this approach for woody plant species in seven man-made islands in a reservoir in eastern North America and confirmed that strong nestedness was attributable to geographic isolation and was correlated with dispersal ability. Wind-dispersed species showed no evidence of nestedness and species without particular adaptations for LDD showed the strongest nestedness. An equivalent pattern is seen in much broader spatial scales, where high dispersal abilities are associated with low levels of endemism, for example, in marine benthic invertebrates with teleplanic larvae such as the gastropod shown in **Box 1**. Similarly, the biotas of isolated islands such as Krakatau (*see* Colonization of Virgin Habitats), Galápagos, and Hawaii, are overrepresented by taxa having LDD mechanisms (but *see* The Geography of Dispersal), a characteristic termed disharmony to denote the lack of biotic elements that are conspicuously present on continents (Carlquist, 1974).

Another traditional approach to elucidate the role of dispersal in determining variation in geographical ranges among species is to test the relationship between some measures of species' dispersal ability and range size. Empirical tests of this relationship provided conflicting results. A key limitation of this approach lies in the difficulty of distinguishing cause and effect typical of correlation-based studies: Is the range size of a species small because of its low dispersal ability, or, as discussed *in* The Geography of Dispersal, is its dispersal ability low because of its small range?

During the 1970s and the early 1980s, the approach used in the study of historical biogeography shifted dramatically from explanations of the unique responses of individual groups to a quantitative and more rigorous analysis of general patterns. This progress was achieved mainly by the scholars of vicariance biogeography (*see* Introduction; Myres and Giller, 1988) who, based on Hennig's phylogenetic systematics and Croizat's panbiogeography, developed the area cladogram as a general technique for the analysis of biogeographic patterns. Area cladograms are branching diagrams representing history, similar to the traditional phylogenetic trees, but with the geographic areas replacing taxa at the terminal tips. Correspondence between phylogeny and independently derived geological history is taken to imply vicariance; dispersal is generally invoked to explain lack of correspondence. Congruence among area cladograms of different taxonomic groups from the same region more strongly implies a vicariant event and may be used to infer geological history in the absence of independent data.

In 1995, Juan Morrone and Jorge Crisci classified the historical biogeography methods into five groups: dispersalist, phylogenetic biogeography, panbiogeography, cladistic biogeography, and parsimony analysis of endemicity. The methods of cladistic biogeography are the most diverse and most widely used (Morrone and Crisci, 1995). These approaches construct an area cladogram assuming vicariance as a null hypothesis. An alternative interpretation is presented in the book edited by Wagner and Funk in 1995 where simple area cladograms were produced for 25 taxonomic groups, including insects, spiders, birds, and flowering plants, of the Hawaiian Islands. The cladistic analysis was undertaken using various attributes, including morphological and anatomical characters, banding patterns on chromosomes, and DNA sequences from chloroplast, mitochondria, and nuclear ribosomes. Area cladograms were compared with several hypothetical patterns that represent historical scenarios that putatively generated the observed patterns, based on the geological history of the islands. The most frequent pattern was of dispersal from older to younger islands; several patterns were interpreted either as recent colonization from the mainland or as vicariant events within the archipelago. A book on the historical biogeography of Southeast Asia, edited by Hall and Holloway in 1998, provides another exemplary integration of ecological and evolutionary processes that account for the impact of dispersal and vicariance along with a profound analysis of the complex geological history of this region. Comparing phylogenetic versus geological divergence times, a general approach recently gained credit by new dating techniques such as Bayesian molecular clocks, further revealing that dispersal explanations might be preferred even if

the species dispersal ability is assumed to be poor. For example, seeds of the woody plant genus *Nothofagus* (southern beech) have no apparent transoceanic LDD mechanism but the 35 extant species exhibit strong disjunct distribution in the southern hemisphere; the genus was therefore considered to be an iconic example for the Gondwanan-breakup history of vicariant events. Yet, using relaxed molecular clocks based on chloroplast DNA, Michael Knapp *et al.* (2005) found that divergence of several species currently endemic to Australia or New Zealand occurred after the two landmasses have split. The most likely explanation for the current distribution of these species is therefore transoceanic LDD during the Late Tertiary.

Ronquist (1997) developed the dispersal–vicariance analysis (DIVA), which reconstructs the ancestral distributions of the species in a given phylogeny without any prior assumption about the form of area relationships and allowing dispersal and extinction. The optimal ancestral distributions are those that minimize the number of dispersal and extinction events required to explain the pattern. The number of dispersal events is underestimated because similar distributions between species are interpreted as sharing vicariant history, though this is not necessarily the case. This method disentangles some important methodological and conceptual restrictions of previous methods and is currently the most widely used method in historical biogeography. Although DIVA, as its name implied, has been considered as a potential bridge between vicariance and dispersal biogeography, this method and similar event-based methods such as tree fitting, strongly bias against dispersal and extinction events in favor of vicariance and local speciation (within the original area). Nevertheless, analysis of 23 phylogenic datasets of plants from the southern hemisphere revealed that most (18) of them were consistent with the dispersal rather than the vicariance explanation (Sanmartín *et al.*, 2007). Interestingly, although most of the reconstructed dispersals from Australia to New Zealand fitted the predicted eastward transport by the western wind drift and the Antarctic circumpolar ocean current, the reconstructed dispersals from New Zealand to South America were more frequently in the opposite direction. In another example, Voelker (1999) reconstructed the ancestral distributions and the direction of dispersal events in the pipits (*Anthus* spp.), small songbirds taxonomically related to the wagtails that occur on all continents. The results show that 16 dispersal events are required to explain the biogeographic pattern of the 40 species analyzed. The reconstructed map of colonizations from the postulated ancestral area in central Asia (**Figure 7**) shows broad intercontinental movements, suggesting that dispersal played an important role in the distribution of this group.

The parsimony principle underlying DIVA and other event-based methods necessitates that the parameters representing processes such as dispersal will be fixed beforehand and not estimated directly from the data. To overcome this and other limitations, several maximum-likelihood methods (e.g., Ree and Smith, 2008) and Bayesian approaches (e.g., Hickerson and Meyer, 2008) have recently been developed and applied to elucidate the role of dispersal and vicariance in historical biogeography. In 2009, Juan Santos and colleagues, for example, used relaxed Bayesian clock analyses and a maximum-likelihood dispersal–extinction–cladogenesis method to reconstruct the biogeography of the poison frog clade

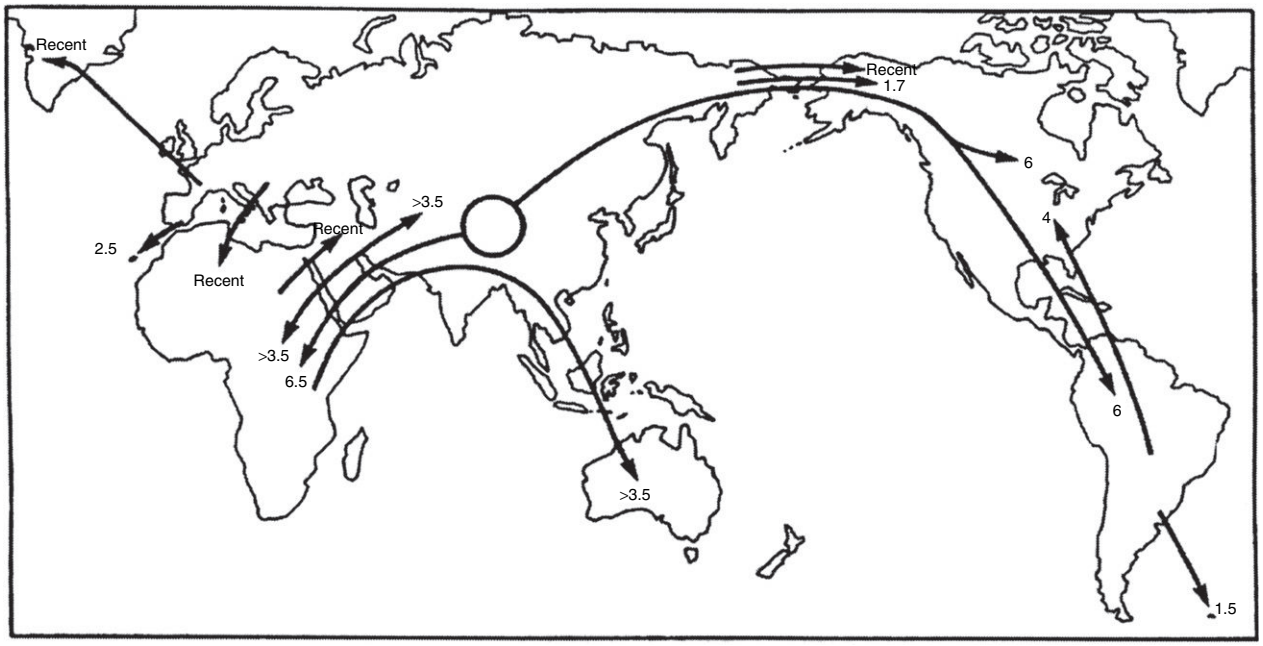


Figure 7 Major intercontinental and continent-to-island colonizations (arrows) by pipits (*Anthus*) form a postulated ancestral area in central Asia (large circle). The approximate time of each colonization event is given either in million years (numbers close to the arrowheads) or is defined as recent (single species that have dispersed onto another landmass). Reproduced from Voelker G (1999) Dispersal, vicariance, and clocks: Historical biogeography and speciation in a cosmopolitan passerine genus (*Anthus*: Motacillidae). *Evolution* 53: 1536–1552, with permission from Wiley.

(Dendrobatidae). The Amazonian center-of-origin hypothesis was rejected in favor of a complex set of vicariance and dispersal events, both out and into Amazonia, mostly from the Andes where significant *in situ* diversification took place.

Modeling

Mathematical models are useful tools for exploring the role of dispersal in the process of range expansion and for predicting expansion dynamics. A comprehensive study from various fields such as genetics, epidemiology, and biological control has led to the development of highly diverse mathematical approaches of varying complexity to modeling movement and spatial population dynamics. These models include dispersal models that focus on the movement alone and reproduction–dispersal models that incorporate dispersal with local population growth to model the spatial spread of populations. The rich literature on this subject was reviewed and summarized by Turchin (1998, p. 73), who concluded with the following prognosis: "the theory is well developed, and appears to be ready to incorporate whatever quantitative information about movement we can provide." Surprisingly, biogeographers have generally overlooked this progress, despite its obvious relevance and potential contribution to the field. A brief overview of some important dispersal and expansion models follows; details can be found in Mollison (1995), Shigesada and Kawasaki (1997), Turchin (1998), and Okubo and Levin (2001).

Dispersal Models

Dispersal models include phenomenological (empirical curve fitting), diffusion, individual-based, and mechanistic (transport) models. The basic objective of these models is to estimate the probability distribution of dispersal distances, which is often given in a standardized functional form called the dispersal kernel (Nathan and Muller-Landau, 2000). A classic dispersal model used in biogeography is the stepping-stone model of MacArthur and Wilson (1967), which assumes that a disseminule moves in a constant direction from one island to another and has a constant probability m per unit time of landing (cessation of movement). Mathematically, these assumptions lead to a widely used phenomenological model – the negative exponential function – in which the probability to disperse to distance x decreases with x at a rate equal to $\ln(m)$. Other widely used phenomenological dispersal models follow the Gaussian distribution and the inverse power law. The Gaussian models usually underestimate observed dispersal data, which almost universally exhibit a leptokurtic distribution (*see* The Evolutionary and Population Genetics Perspectives). The inverse power law has the problem of singularity (infinite density) at zero and hence its simple formulation cannot be used as a dispersal kernel.

Another set of dispersal models that is likely to prove beneficial to dispersal biogeography is that of the mechanistic models, which simulate the process by a set of mathematical expressions that mimic the relationships and joint effects of the principal operative factors. For example, mechanistic models of seed dispersal by wind typically incorporate factors such as the height of release, the seed free-fall velocity, and wind velocities (Nathan *et al.*, 2001). Although complicated to develop and parameterize, these models have the decisive advantages of actually predicting, rather than of merely fitting, changes in disseminule densities with distance, and of potential generalization to other systems. From the biogeographic perspective, special attention should be paid to the tail of the distribution, since LDD is disproportionately important in jump dispersal and range expansions. Mathematically, it has been shown that kernels differing significantly in their tails yield dramatically different rates of population spread, even if they have the same mean. Turchin (1998) summarized the relationships between dispersal models and the types of population spread, focusing on the tail of the dispersal kernel. If the tail of the dispersal kernel drops off exponentially or faster with distance (e.g., Gaussian tails that are predicted by simple diffusion), then we should observe linear spatial spread in which the range is expanded at a constant rate during each time step. If the tail drops off slower than the exponential tail, as in the so-called "fat-tailed" kernels, the type of spatial spread depends on whether the variance of the kernel is finite or infinite. In both cases, the rate of spatial spread accelerates (i.e., increases with successive time steps), but the former reaches an asymptote of a constant expansion rate, whereas the latter continues to increase. Thus, fat-tailed dispersal kernels lead to jump range expansion, whereas exponential and thinner tails lead to gradual range expansion (*see* The Biogeographic Perspective). Since data on LDD are rare, the assumptions of the characteristics of the distribution's tail cannot yet be tested; at the same time it is clear that extrapolating beyond the observed range is questionable.

The need to implement "fat-tailed" phenomenological and mechanistic dispersal models is now well established in dispersal research. Yet, the task of applying and testing models of LDD is inherently more complicated than for short-distance dispersal, despite recent technical developments with miniature remote sensing techniques and molecular markers (Nathan, 2006). Important attempts to model the mechanism of LDD has been made for the wind dispersal of aerial plankton (Kuparinen, 2006) and seeds (see Nathan *et al.*, 2011b for review), based on the knowledge of atmospheric dynamics. Similarly, knowledge of sea currents and the winds that drive them can be used to predict the trajectories of passively dispersed marine biota (Cowen *et al.*, 2006). Nevertheless, some enigmatic LDD phenomena, such as the dispersal of seeds and invertebrates by migratory birds and dispersal of various propagules by extreme weather events such as hurricanes remain out of reach for both modelers and field researchers (Nathan *et al.*, 2008b).

Reproduction–Dispersal Models

The dispersal models that have been discussed thus far are useful tools to understand the general relationships between dispersal and spatial spread. Reproduction–dispersal models, however, are more realistic since they consider the effects of birth and death during the process of range expansion. The simplest reproduction–dispersal models are reaction–diffusion models; and the simplest of those,

introduced to biologists by Skellam (1951), model dispersal as random diffusion and local population growth as exponential increase

$$\frac{dN}{dt} = rN + D\left[\frac{\partial^2 N}{\partial x^2}\right]$$

where N is the population size, t is time, and x is the distance from the source; the two parameters are D, the diffusion coefficient, and r, the intrinsic rate of population growth. Earlier, Fisher (1937) used a nonlinear variant of this equation in which the unbounded exponential growth (rN) is replaced by asymptotic logistic growth ($rN(1-N/K)$), where K is the equilibrium population size (or carrying capacity). As shown by Fisher, this model predicts, as does Skellam's model, a constant asymptotic front velocity equal to $2\sqrt{rD}$. This type of gradual range expansion does not incorporate the effect of LDD events. Tests of reproduction–dispersal models against observed expansion pattern, starting from Skellam's (1951) pioneering study, have frequently showed that the simple reaction–diffusion approach cannot provide a general description for range expansion. For example, Andow *et al.* (1993) found approximately linear rates of spread in two case studies and accelerating rates of spread in two others. Even for the cases that show linear rates, such as that of muskrats (*see* The Spread of Alien Species), there was significant variation in space and time. Expansion rates of muskrats were lower in the first 2–3 years, possibly because of Allee effects (*see* Ecological Opportunity: Postdispersal Establishment), and the average rates range from 0.9 to nearly 80 km yr^{-1}. Thus, reaction–diffusion models are best used as a primary null analysis of expansion patterns, followed by more complex approaches, including stratified diffusion, interacting particle systems, and spatial contact models. These models incorporate both short- and long-distance dispersal and thus are able to predict patterns of jump range expansion as well.

Biogeographic and Evolutionary Consequences of Dispersal

Dispersal, evolution (adaptation and speciation), and extinction are the fundamental processes in biogeography. Together, they determine the geographical distribution of organisms. Dispersal biogeography focuses on the effects of dispersal on evolution and especially on speciation, both within a lineage and for other lineages. The following section outlines some principal relationships; more details are given in related articles on speciation and extinction in this encyclopedia.

Speciation

Speciation is the process by which new independently evolving lineages (species) arise from a single ancestral lineage. In general, efficient dispersal between populations (patches) increases the rate of gene flow between them and hence tends to reduce genetic differentiation and to restrain speciation. In contrast, inefficient dispersal between existing populations leads to geographic isolation and tends to increase genetic differentiation and to promote speciation. This type of speciation, called allopatric speciation, is considered to be the most frequent one. Jump range expansion occurs with very low probability (*see* The Biogeographic Perspective) and is thus unlikely to be followed by additional dispersal; this isolation thereby causes rare dispersal events (in this context also called founder events) to be relevant to allopatric speciation. The adaptive radiation of Darwin's finches in the Galápagos Islands is one of the many examples of allopatric speciation that has resulted from dispersal events. However, as discussed in the Introduction, a dispersal barrier may divide a previously continuous distribution. Hence, allopatric speciation may also result from a vicariant event.

A second way in which dispersal affects speciation requires the opposite conditions. The breakdown of a dispersal barrier, and the following biotic interchange (*see* Removal of Barriers), brings organisms to dissimilar environments and thus opens new chances for diversification. The colonization of North American mammals during the Great American Interchange (*see* Removal of Barriers) was followed by extensive diversification in South America. In the same manner dispersal events cause speciation. They affect the genetic structure within the species range, sometimes to the point that the populations are evolving independently, and can be considered as distinct species. Theoretical models and molecular studies have shown clear effects of LDD on the spatial structure of the genetic variation within a species during the postglacial expansion of European and North American biota (*see* Holocene Postglacial Spread). As predicted by models, jump range expansion well ahead of the "normal" advancing wave generates large patches of low genetic variation, due to the so-called founder effects; consequently, genetic diversity decreases from south to north.

Extinction

Limited dispersal is associated not only with higher rates of speciation (*see* Speciation) but generally also with higher rates of extinction. This phenomenon has been observed at various spatial scales. For example, the fossil record shows that gastropods with planktonic larval dispersal, typically have high LDD capacity (**Box 1**), tend to have larger geographical ranges and to persist for longer periods of time than do those with nonplanktonic larvae, which are dispersed locally. A similar phenomenon is observed at regional scales, where populations with higher rates of immigration persist longer due to increased population size and higher levels of genetic variability, which counteracts the deleterious effects of inbreeding. This so-called rescue effect has been observed in a variety of organisms. The very similar notion of source–sink dynamics attributes a longer persistence time to populations in unfavorable sink habitats (where death rates exceed birth rates) due to a rescue effect from populations in favorable source habitats (where birth rates exceed death rates). Another relevant term "mass effect" emphasizes the contribution of rescue effects to the breeding stock. The regional framework of populations subjected to local extinction and recolonization is treated in metapopulation models in which dispersal plays a critical role.

Although dispersal tends to reduce the probability of extinction of the same species, dispersal, as immigration, may have the opposite effect on other species inhabiting the recipient site. Around 1950, the brown tree snake (*Boiga irregularis*) of Australia and Southeast Asia was accidentally introduced to the Pacific island of Guam. This fairly large (up to 3 m) snake, which eats bird eggs from nests either on the ground or in trees, caused the extinction of most of the native birds and the decline of almost all other native terrestrial vertebrates. Despite this and other spectacular examples of massive extinction caused by colonization of a given species, most colonization events appear to have minor consequences. The consequences of colonization of one species to the persistence of other species depend on the species' biology and on postdispersal processes, however, dispersal is the ultimate cause of such extinctions, since it determines whether the prerequisite species is present. Though the particular dispersal mechanisms should not matter to the severity of such incidents, it is noteworthy that in most documented cases dispersal was related to human activity.

Applied Dispersal Biogeography

Biogeography in general, and dispersal biogeography in particular, has an important role in managing the world's biodiversity. The processes currently threatening the world's biota typically act on large scales, and the movement (or lack thereof) of organisms is a key process. Dispersal is strongly involved with issues such as invasions, introductions, fragmentation, climatic changes, design of nature reserves, and control of gene flow; it is hard to imagine any issue of applied biogeography that is not related to the dispersal process. Thorough discussions of these practical aspects of biogeography can be found in Williamson (1996), Brown and Lomolino (1998), and Spellerberg and Sawyer (1999); this section outlines some major principles of applied dispersal biogeography.

Applied dispersal biogeography deals with all aspects of dispersal that are involved with the management of the Earth's biota and aims to predict the causes and the consequences of dispersal events. The main threats to global biodiversity involve excessive LDD of elements alien to ecosystems and insufficient dispersal of native species, for example, because of habitat fragmentation (Trakhtenbrot *et al.*, 2005). The mathematical models described above (*see* Modeling) are efficient tools on small scales. Although our ability to predict LDD continues to improve, much still remains to be done, especially the testing of these predictions. LDD has a substantial stochastic component, which will always remain somewhat unpredictable. However, since most LDD events with harmful consequences are related to human activity, our predictive ability of anthropogenic arrivals may be better than for natural arrivals. Even unintentional dispersal such as that of the brown tree snake to Guam (*see* Extinction; **Box 1**) is to some degree predictable and controllable. As emphasized earlier, the variation among organisms in dispersal ability is enormous (**Box 1**); responses to dispersal barriers and dispersal routes also vary considerably (*see* Physical Access: Dispersal Barriers and Dispersal Routes). Mathematical models, especially mechanistic ones (*see* Dispersal Models), can be used to predict dispersal of particular species using knowledge of the species-specific dispersal traits (**Box 1**) and the effect of environmental factors on dispersal (*see* Physical Access: Dispersal Barriers and Dispersal Routes *and* The Geography of Dispersal).

The invasion of muskrats to Europe (*see* The Spread of Alien Species) illustrates these points. The muskrat's arrival was related to human activity and the animal's escape was controllable; expansion rates are predictable through simple mechanistic models based on dispersal and reproduction (*see* Reproduction–Dispersal Models); the variation in expansion rates was related to habitat-related factors (*see* The Geography of Dispersal *and* The Spread of Alien Species) and was therefore also predictable. The large number of escapes and the initial slow rates of expansion demonstrate the importance of propagule pressure (*see* The Spread of Alien Species) and probably of Allee effects (*see* Ecological Opportunity: Postdispersal Establishment) as well. Attempts to control this expansion – mostly extensive trapping campaigns – have failed in continental Europe, but have succeeded in Britain and Ireland. Mark Williamson (1996) suggested that the relatively early response in the British Isles, before the muskrats spread too far and built overly large populations, along with presumably inferior climatic matching that would have led to lower reproduction rates, are possible reasons.

Synthesis

The importance of dispersal in determining the geographic distribution of organisms has been emphasized by Darwin and other dispersalists. These views have been seriously questioned by their opponents, who argue that vicariant events are the principal determinant of distributions. Clearly, both processes occur and their effects vary between species both in space and in time. The term "dispersal biogeography" is currently restricted to the arguments of extreme dispersalists concerning LDD and disjunct distributions. The broader and more fruitful usage of this term adopted here would be the study of the relationship between dispersal and the geographical distribution of organisms.

Dispersal is a widespread phenomenon occurring in nearly all organisms. There is a high variability of dispersal ability in organisms, due to differences in their physiological, morphological, and behavioral traits and responses to environmental conditions. Dispersal is a key biological process at the ecological, evolutionary, genetic, and biogeographic levels, which differ in spatial and temporal scales. LDD, disproportionately important at the biogeographic level, is typically rare and very difficult to measure; it contains a large stochastic component but in many cases is a result of a specific mechanism and hence can be predicted. If followed by successful establishment, LDD can lead to abrupt long-distance colonization jumps. Short-distance dispersal is overwhelmingly important at the ecological level; at the biogeographic level, multiple short-distance dispersal events over a relatively long time, if followed by establishment, can lead to gradual range expansion. Patterns of speciation and extinction

are strongly affected by dispersal, demonstrating the broad biogeographic consequences of this process, which itself is also affected by geography.

The field of biogeography has gone through a dramatic shift since the publication of the first version of this article in 2001. Molecular tools to quantify genetic variation continue to improve and to become increasingly more cost-effective, providing data of much higher quality and quantity compared to what were available just a few years ago. Tools to analyze multiscale spatial patterns such as various Geographical Information Systems (GIS) techniques have been rapidly developed. The field of phylogeography, established 25 years ago to integrate the closely related, yet traditionally isolated, disciplines of population genetics and phylogenetic biology (Avise, 2009), has recently experienced a boost of developments of new statistical frameworks enabling us to test key hypotheses on the processes responsible for the observed phylogeographic patterns (Knowles, 2009). New tools enable tracking of smaller organisms for longer duration and new theoretical frameworks can associate studies across diverse movement phenomena into an integrated movement ecology paradigm (Nathan *et al.*, 2008a). Thus, one can scrutinize the complex set of intrinsic properties of individuals and their interactions with the biotic and physical environment that determine their dispersal and movements and eventually the distribution of species at both local and large scales. It is therefore the time now to develop synergies among all these and other exciting recent developments, to test challenging hypotheses about the role of dispersal and other movement phenomena in the evolution and distribution of biota. For some promising integrative frameworks, see the recent reviews of Spear *et al.* (2010), Chan *et al.* (2011), and Crisp *et al.* (2011).

See also: Biogeography, Overview. Island Biogeography. Life History, Evolution of. Metapopulations. Migration. Plant Invasions. Speciation, Process of. Speciation, Theories of. Vicariance Biogeography

References

Andow DA, Kareiva PM, Levin SA, and Okubo A (1993) Spread of invading organisms: Patterns of spread. In: Kim KC and McPheron BA (eds.) *Evolution of Insect Pests: Patterns of Variation*, pp. 219–242. New York: Wiley.

Avise JC (2009) Phylogeography: Retrospect and prospect. *Journal of Biogeography* 36: 3–15.

Brown JH and Lomolino MV (1998) *Biogeography*. Sunderland, MA: Sinauer.

Browne J (1983) *The Secular Ark: Studies in the History of Biogeography*. New Haven: Yale University Press.

Cain ML, Damman H, and Muir A (1998) Seed dispersal and the Holocene migration of woodland herbs. *Ecological Monographs* 68: 325–347.

Carlquist S (1974) *Island Biology*. New York: Columbia University Press.

Chan LM, Brown JL, and Yoder AD (2011) Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution* 59: 523–537.

Cody ML and Overton JM (1996) Short-term evolution of reduced dispersal in island plant populations. *Journal of Ecology* 84: 53–61.

Cowen RK, Paris CB, and Srinivasan A (2006) Scaling of connectivity in marine populations. *Science* 311: 522–527.

Crisp MD, Trewick SA, and Cook LG (2011) Hypothesis testing in biogeography. *Trends in Ecology & Evolution* 26: 66–72.

Darwin C (1859) *The Origin of Species by Means of Natural Selection*. London: John Murray.

Fisher RA (1937) The wave of advance of advantageous genes. *Annals of Eugenics* 7: 355–369.

Hall R and Holloway JD (eds.) (1998) *Biogeography and Geological Evolution of SE Asia*. Leiden: Backhuys.

Hengeveld R (1989) *Dynamics of Biological Invasions*. London: Chapman & Hall.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87–112.

Hickerson MJ and Meyer CP (2008) Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical bayesian approach. *BMC Evolutionary Biology* 8.

Johnson ML and Gaines MS (1990) Evolution of dispersal: Theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* 21: 449–480.

Kadmon R (1995) Nested species subsets and geographic isolation: A case study. *Ecology* 76: 458–465.

Knapp M, Stockler K, Havell D, Delsuc F, Sebastiani F, and Lockhart PJ (2005) Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (southern beech). *PLoS Biology* 3: 38–43.

Knowles LL (2009) Statistical phylogeography. *Annual Review of Ecology Evolution and Systematics* 40: 593–612.

Kuparinen A (2006) Mechanistic models for wind dispersal. *Trends in Plant Science* 11: 296–301.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton: Princeton University Press.

Mollison D (1995) *Epidemic Models: Their Structure and Relation to Data*. Cambridge: Cambridge University Press.

Morrone JJ and Crisci JV (1995) Historical biogeography: Introduction to methods. *Annual Review of Ecology and Systematics* 26: 373–401.

Myres AA and Giller PS (eds.) (1988) *Analytical Biogeography: An Integrated Approach to the Study of Animal and Plant Distributions*. London: Chapman and Hall.

Nathan R (2006) Long-distance dispersal of plants. *Science* 313: 786–788.

Nathan R, Getz WM, Revilla E, *et al.* (2008a) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America* 105: 19052–19059.

Nathan R, Horvitz N, He Y, Kuparinen A, Schurr FM, and Katul GG (2011a) Spread of North-American wind-dispersed trees in future environments. *Ecology Letters* 14: 211–219.

Nathan R, Katul GG, Bohrer G, *et al.* (2011b) Mechanistic models of seed dispersal by wind. *Theoretical Ecology* 4: 113–132.

Nathan R and Muller-Landau HC (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology and Evolution* 15: 278–285.

Nathan R, Safriel UN, and Noy-Meir I (2001) Field validation and sensitivity analysis of a mechanistic model for tree seed dispersal by wind. *Ecology* 82: 374–388.

Nathan R, Schurr FM, Spiegel O, *et al.* (2008b) Mechanisms of long-distance seed dispersal. *Trends in Ecology and Evolution* 23: 638–647.

Nelson G and Platnick N (1981) *Systematics and Biogeography: Cladistics and Vicariance*. New York: Columbia University Press.

Nelson G and Rosen DE (eds.) (1981) *Vicariance Biogeography: A Critique*. New York: Columbia University Press.

Okubo A and Levin SA (eds.) (2001) *Diffusion and Ecological Problems: Modern Perspectives*. New York: Springer.

Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Patterson BD and Atmar W (1986) Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biological Journal of the Linnean Society* 28: 65–82.

Pielou EC (1979) Biogeography. New York: Wiley & Sons.

van der Pijl L (1982) *Principles of Dispersal in Higher Plants*. Berlin: Springer.

Por FD (1978) *Lessepsian Migration: The Influx of Red Sea Biota into the Mediterranean by Way of the Suez Canal*. Berlin: Springer.

de Queiroz A (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution* 20: 68–73.

Ree RH and Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14.

Ridley HN (1930) *The Dispersal of Plants Throughout the World*. Ashford: Reeve.

Ronquist F (1997) Dispersal–vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology* 46: 195–203.

Sanmartín I, Wanntorp L, and Winkworth RC (2007) West wind drift revisited: Testing for directional dispersal in the southern hemisphere using event-based tree fitting. *Journal of Biogeography* 34: 398–416.
Santos JC, Coloma LA, Summers K, Caldwell JP, Ree R, and Cannatella DC (2009) Amazonian amphibian diversity is primarily derived from late Miocene Andean lineages. *PLoS Biology* 7: 448–461.
Shigesada N and Kawasaki K (1997) *Biological Invasions: Theory and Practice*. New York: Oxford University Press.
Simpson GG (1965) *The Geography of Evolution*. Philadelphia: Chilton.
Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* 38: 196–218.
Spear SF, Balkenhol N, Fortin MJ, McRae BH, and Scribner K (2010) Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology* 19: 3576–3591.
Spellerberg IF and Sawyer JWD (1999) *An Introduction to Applied Biogeography*. Cambridge: Cambridge University Press.
Thornton I (1996) *Krakatau: The Destruction and Reassembly of an Island Ecosystem*. Cambridge, MA: Harvard University Press.
Trakhtenbrot A, Nathan R, Perry G, and Richardson DM (2005) The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions* 11: 173–181.
Turchin P (1998) *Quantitative Analysis of Movement*. Sunderland: Sinauer.
Van Dyck H and Matthysen E (1999) Habitat fragmentation and insect flight: A changing 'design' in a changing landscape? *Trends in Ecology & Evolution* 14: 172–174.
Vilà C, Sundqvist AK, Flagstad O, *et al*. (2003) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London Series B* 270: 91–97.
Voelker G (1999) Dispersal, vicariance, and clocks: Historical biogeography and speciation in a cosmopolitan passerine genus (*Anthus*: Motacillidae). *Evolution* 53: 1536–1552.
Wabakken P, Sand H, Kojola I, *et al*. (2007) Multistage, long-range natal dispersal by a global positioning system-collared Scandinavian wolf. *Journal of Wildlife Management* 71: 1631–1634.
Wagner WL and Funk VA (eds.) (1995) *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago*. Washington, DC: Smithsonian Institution Press.
Wiley EO (1981) *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. New York: Wiley & Sons.
Williamson M (1996) *Biological Invasions*. London: Chapman & Hall.

Disturbance, Mechanisms of

Frank W Davis and Max A Moritz, University of California, Santa Barbara, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Disturbance Relatively discrete event in time that disrupts ecosystem, community, or population structure and changes resources, substrate availability, or the physical environment.
Disturbance regime The collective spatial, temporal, physical, and ecological characteristics of a disturbance process operating in an area.
Resilience Measure of a system's ability to recover to its original state when subjected to disturbance.
Resistance A measure of system's ability to remain unchanged when subjected to a disturbance.

Disturbance Ecology

Ecological communities are subjected to sharp environmental perturbations such as the passage of a fire, a storm wave, an avalanche, or a large animal, that suddenly reduce standing biomass. Such events also alter ecosystem properties such as energy balance, nutrient fluxes, substrate texture, and chemistry. Generally referred to as "disturbances," these events are distinguished as a special form of environmental variability because they are relatively discrete in both time and space and cause unusual mortality or tissue loss in affected populations. Disturbances promote environmental heterogeneity and free up limiting resources, such as space, light, and nutrients, thereby triggering successional processes of community recovery.

Widespread agents of disturbance in terrestrial ecosystems include fire, wind, extreme temperature, desiccation, gravity (as a force on water, ice, rocks, and soil), and organisms. With the exception of fire, one can find analogs in aquatic systems, in which the main agents are heat, solutes, currents, desiccation, waves, ice, sediments, and organisms. Disturbance agents operate by exerting mechanical force, altering physico–chemical conditions, or through biological consumption and disease. Practically, any physical mass can mechanically disturb ecosystems at some scale provided it has sufficient velocity to dislodge organisms or kill tissues. In contrast, disturbance by fire involves chemical combustion of biomass. Disturbance of terrestrial communities by prolonged inundation is largely a chemical disturbance resulting from oxygen depletion. Freshets in estuarine and marine environments are also chemical disturbances. These mechanical and physico–chemical disturbances contrast with biological disturbances such as herbivory or predation, which involve tissue removal and digestion by individual mobile consumers.

During the first half of the twentieth century, ecologists paid more attention to postdisturbance processes of species recovery and community succession than to specific disturbance mechanisms. Disturbance was generally treated as a temporary setback to communities that otherwise would tend to develop toward a relatively steady state or "climax community" whose structure and composition were determined by climate and other physical factors and were regulated by endogenous biological interactions. Recently, disturbance has been recognized as intrinsic to and ongoing in virtually all ecological systems. Closer attention has been paid to disturbance processes and their role in promoting characteristic scales of spatial and temporal environmental heterogeneity and in regulating ecosystem processes, population dynamics, species interactions, and species diversity (Paine and Levin, 1981; Sousa, 1984; Pickett and White, 1985; Turner, 2010). Efforts to improve the understanding of disturbance mechanisms have accelerated due to concern over the effects of global environmental change on disturbance regimes and associated ecosystem dynamics.

In practice, it may not be easy to distinguish disturbances from other environmental variation. Most agents of disturbance operate over a continuum, and environmental perturbation is sudden and severe only relative to some set of reference conditions and from the perspective of the affected organisms. The mound of soil produced by a burrowing gopher is a significant disturbance to the underlying herbaceous plants and soil animals but likely to be of little consequence to a large tree, a few meters away. A windstorm that topples trees in an open savanna may have no immediate effect on small herbaceous plants located a short distance outside of the canopy. Thus, it is important to bear in mind that disturbance is a relativistic concept and it can span a very broad range of spatial and temporal scales. Not surprisingly, the term disturbance has been applied somewhat indiscriminately in ecology. The most general definition of a disturbance – any process that causes a sudden decrease in standing live biomass and frees up ecological resources (Sousa, 1984) – is perhaps the most unambiguous. Another widely used definition, provided by White and Pickett (1985, p. 7) is "any relatively discrete event in time that disrupts ecosystem, community, or population structure and changes resource, substrate availability, or the physical environment".

A disturbance differs from a stress in that the latter is a more chronic condition inhibiting the growth or normal functioning of an organism (e.g., a lack of key nutrients or physical abrasion). A disturbance is termed a catastrophe if it causes extraordinary ecological impact.

Variables commonly used to describe a single disturbance event include timing, extent, and magnitude, where magnitude encompasses both intensity (e.g., energy per area per time) and severity (biological impact). These and other stochastic variables, such as event frequency or recurrence interval

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00034-4

between events, have statistical properties that serve to define a disturbance regime. More broadly, a disturbance regime is the collective spatial, temporal, physical, and ecological characteristics of a disturbance process operating in an area. Predictability, which can be defined as the inverse of the variance in disturbance frequency, size, and magnitude (Christensen, 1988), is also an important consideration. In general, predictability increases as the spatio–temporal scale of analysis is expanded from local (the typical size of a disturbance event) to landscape or regional domains (the entire area over which the disturbance regime is manifested).

Because the magnitude of disturbance is defined relative to its ecological impact, it is practically tautological that disturbance regimes are dominated by events of relatively low magnitude and high frequency, whereas higher magnitude events are increasingly rare. However, the impact of a disturbance may not increase linearly with size, frequency, or duration. Romme *et al*. (1998) distinguished three classes of disturbance response: (1) threshold response, (2) scale-independent response, and (3) continuous response. Individuals and communities manifest threshold responses when there are discrete limits in their ability to resist a large disturbance (e.g., the wind speed at which a tree is uprooted). Disturbances can have their greatest ecological impact when one or more events follow close on the heels of another, preventing or disrupting normal community recovery (Paine *et al*., 1998).

Some mechanisms of disturbance, such as earthquakes or storm waves, are exogenous to the biological communities being impacted, whereas others such as treefall or fire could be considered endogenous. In the former, there is little or no feedback between the state of the ecosystem and the likelihood of a disturbance event so that the disturbance regime depends mainly on location and environmental context. In the latter, the likelihood of a disturbance depends on the state of the ecosystem as well as location. Although the categories of endogenous and exogenous disturbances are somewhat artificial, it is useful to examine the relative strength of coupling between disturbance processes and the biota. In many cases, disturbance processes and their effects are tightly coupled to the biological properties of individual organisms and communities. This coupling may promote the formation of specific scales of ecological pattern and reinforce certain ecological and evolutionary processes (Levin, 1992).

In thinking about patterns and processes, it is also useful to distinguish spatially propagating from nonpropagating disturbances (Reiners and Driese, 2003). Disturbances such as fire and flood spread from neighboring areas, and the spatial pattern of "susceptible" areas or organisms may have a constraining effect on disturbance dynamics, thereby linking spread to previous disturbance events.

Disturbance and Biodiversity

The "intermediate disturbance hypothesis" (IDH; Connell, 1978) predicts that maximum levels of biodiversity should be observed under some intermediate disturbance frequency because few species are able to tolerate very intense disturbance regimes, and few are able to compete successfully in habitats that experience little or no disturbance. The IDH also implies that maximum diversity should be found at some intermediate span of time since the last disturbance. The IDH has been expanded to incorporate intermediate levels of disturbance intensity and extent, and it has been tested and supported in a wide variety of ecosystems.

The best experimental examples of the IDH come from the study of sessile species competing for space or some space-associated resource. Counter-examples come largely from the study of mobile consumers (e.g., freshwater invertebrates), for which rapid immigration may override local disturbance effects. Defining "intermediate" in the context of specific organisms and choosing the scale to measure diversity are also important issues in assessing the validity of the IDH, about which there is ongoing debate. In general, the IDH applies to small-scale disturbances and to plants and sessile filter feeders. The relationship between disturbance and diversity is more complicated at larger scales and may not apply when interactions among multiple trophic levels are considered.

At spatial scales, much larger than the characteristic size of single disturbance events, disturbance regimes generate disturbance mosaics that maintain beta diversity in landscapes or regions by promoting coexistence of dispersal-limited competing species or prey species and by maintaining environmental heterogeneity and multiple serai stages (Pickett and White, 1985).

If they recur with sufficiently high frequency (e.g., on average at least once per generation), ecological disturbances can be a strong selective force operating on species' morphology, physiology, and behavior. Not surprisingly, many organisms are adapted to or depend on specific kinds of disturbances and disturbance regimes. Grime (1979) proposed that herbaceous plant life histories can be ordered along three fundamental axes: stress tolerance, competition, and disturbance. Somewhat analogously, animal species are often called *r*-strategists or *k*-strategists depending on whether they have high intrinsic rates of reproduction and tend to be favored by disturbances or whether they have lower reproductive rates but exert competitive dominance in the absence of disturbance. Because local abundance of many species is increased or maintained by disturbance, the long absence of specific kinds of disturbances (e.g., fire or floods) may have large negative impacts on biodiversity.

The life history strategies of some organisms may promote specific disturbance regimes (e.g., some fire-adapted shrub species possess canopy structure and foliar chemistry that promote fire spread). Organisms that recover quickly after a disturbance are said to be resilient to that disturbance, as opposed to those that show little response to disturbance and are considered resistant. These concepts are also applied to ecological communities, and the relationship between community diversity and community stability and resilience has long preoccupied the ecologists (Holling, 1973). The subject has received renewed attention due to increasing concern over human-caused species extinctions and community impoverishment associated with habitat fragmentation. To date, the hypothesis that community stability to disturbance increases with community richness has met with mixed results in modeling and empirical studies, in part owing to differences in spatial and temporal scale and in how stability is measured.

Mechanical Disturbances

Most aquatic and terrestrial communities are constantly subjected to mechanical forces that reduce standing biomass at some scale. Avalanches, landslides, and debris flows remove soil, shear off and uproot plants, and bury plants and animals. High winds, snow, and ice uproot trees and break off branches and leaves that then bury other plants and animals. Ice in the near shore zone plucks and scours littoral communities. Floodwaters topple riparian vegetation and scour streambeds. Waves dislodge encrusting intertidal organisms and break down branching corals. Large animals break twigs and compact the soil as they move. Humans clear or clip vegetation and plow soils. Burrowing animals excavate plants and animals as they dig. In all of these cases, biomass is dislocated but not immediately chemically transformed or consumed. Most important, space occupied by living tissues is evacuated as those tissues are moved someplace else.

Thus, mechanical disturbances leave conspicuous openings and persistent biological legacies, such as large woody debris embedded in landslides, downed trees next to treefall gaps, floating wracks of detached marine plants and animals, or rubble piles of coral. Materials are often transported to a location where they act to create another disturbance. In turn, the legacy of biological debris from mechanical disturbances not only affects the rate and pattern of community recovery but also may influence the timing and location of subsequent disturbance events.

The spatial distribution of mechanical forces in the environment is highly nonrandom. Features such as substrate, topography, and standing biomass create persistent biophysical pathways in which specific mechanical disturbances are concentrated, thereby compounding existing physical environmental heterogeneity. For example, the location of avalanche paths is strongly associated with topography, geology, and lithologic structure. The impacts of severe wind storms on forest communities varies systematically with topography and soil characteristics as well as vegetation composition and structure (which depend, in part, on disturbance history). Wave energy is concentrated on jutting shorelines. Animal burrowing is concentrated in specific soils and sediments. In other words, many mechanical disturbance processes can have somewhat predictable spatial distributions.

Most mechanical disturbances operate indiscriminately (i.e., they do not target specific organisms, as would selective herbivory, for example) and their impact depends on the ability of organisms to withstand the mechanical force. Thus, mechanical disturbances often winnow biological communities. For example, snow avalanches apply a bending stress on trees in their path that is a function of the snow's density and the avalanche's mass and drag. Susceptibility of woody species to breakage (vs. bending) increases with tree size, leading to complex interactions between avalanche recurrence frequency and forest structure and composition in avalanche paths. Similarly, masses of attached barnacles and mussels in rocky intertidal environments become increasingly susceptible to detachment by waves as organism size and density increases. The same mechanical agent may create different kinds of disturbances depending on the size of the organism. In forests, for instance, woody stems may be most susceptible to breakage by wind when they are very small or very large, whereas intermediate-sized stems may be the most susceptible to uprooting by wind (Everham and Brokaw, 1996). When impacts of mechanical disturbances depend on the size or density of organisms, this produces feedbacks that influence population dynamics and create pronounced spatial patchiness at scales well below that of the disturbing process (Levin, 1992).

Some mechanistic models have been developed to better understand and predict the interactions between physical forces and biological communities. A good example is the family of models relating slope or stream bank stability to plant root systems. Soils have low tensile strength but high compressive strength; conversely, roots have high tensile strength but are weak under compression. Thus the susceptibility of soil masses to shearing forces and associated landslides or bank collapses is strongly related to the composition, density, and size distribution of roots at different soil depths (Pollen-Bankhead *et al.*, 2009).

Physico–Chemical Disturbances

Physico–chemical disturbances such as extreme cold, prolonged inundation of terrestrial organisms, hypoxic episodes in aquatic systems, freshets in coastal waters, and releases of toxic compounds (e.g., biological exotoxins such as those produced by the dinoflagellate *Pfiesteria piscicida*), reduce biomass by imposing lethal physiological stress. Physico–chemical disturbances do not involve mechanical force (except perhaps in the case of extreme cold when tissue damage is due to the intracellular ice formation), and most physical and chemical disturbances leave dead tissues and organisms *in situ*. This contrasts with mechanical disturbances, in which space and resources are freed up through displacement of biological material. There are few studies comparing recovery processes under these two situations, but one might expect differences simply due to the presence or absence of residual detritus that could affect light, nutrient levels, and substrate quality. For example, management experiments in conifer forests of the western US comparing effects of mechanical thinning to controlled burning or both thinning and burning have produced similar forest structures but very different understory plant communities (Wayman and North, 2007).

Like mechanical disturbances, physico–chemical disturbances have organism-specific impacts because of species-, age-, and size-specific physiological tolerances to environmental extremes. Thus, like mechanical disturbances, physico–chemical disturbances operate as an environmental filter of biological assemblages.

As mentioned previously, much of the emphasis in disturbance studies has been on sessile organisms such as plants and benthic invertebrates. Some largescale physico–chemical disturbances may be somewhat distinctive in that they can kill large numbers of highly mobile organisms. For example, episodes of anoxia can cause massive fish kills in estuaries and lakes. Severe cold weather can selectively kill large numbers of birds. In these instances, the suddenness and magnitude of the events prevent animals from escaping, a phenomenon, which can also be observed during extreme wildfires.

Fire is obviously a special case of physico–chemical disturbance because biomass is both transformed and removed in the combustion process. Fire is also unquestionably one of the most pervasive and obvious disturbances in terrestrial ecosystems, and an enormous scientific literature has been devoted to the subject of fire events and fire regimes. It is discussed here to illustrate in more detail the concepts of disturbance mechanisms and disturbance regimes.

The occurrence of fire requires an ignition, fuel combustion, and spread, each of which in turn depends on many physical and biological factors, notably climate, weather, fuel structure and chemistry, and topography. The most important effects of fire are partial or complete combustion of aboveground vegetation, mineralization and deposition of plant tissues as ash and charcoal, extreme heating of the local atmosphere and topsoil, and killing of selected plants, seeds, spores, and animals both above and below ground.

Figure 1 presents a simple conceptualization of factors that influence fire dynamics at three distinctive spatial and temporal scales in Mediterranean-climate shrublands. The fire fundamentals triangle (**Figure 1a**), captures essential elements of the actual combustion process, which occurs at minute scales and depends on oxygen, heat, and fuel. The timing, size, and severity of a wildfire (or controlled fire) event, which may last hours to weeks and extend over hectares to thousands of square kilometers, depend on local weather, topography, and fuel bed. Finally, the fire regime of a landscape depends on complex interactions among climate, vegetation, and ignition factors operating over decades to centuries.

Fire creates strong biological patterns over a broad range of spatial scales. At the scale of individual plants, biomass combustion, soil heating, and nutrient removal can vary widely, imposing very fine-grained variation in mortality and postfire conditions that is manifested in equally fine-grained patterns of biological recovery. A coarser pattern can be created at the scale of entire population (vegetation stands). Fire severity and vegetation removal can vary with topography, soil, vegetation, and wind conditions, and distance from source populations for seeds and animal colonists can influence community recovery. At still broader scales, landscapes are composed of fire mosaics, a patchwork of stands with differing fire histories and biotic communities. This landscape mosaic is one of the factors aiding in the long-term persistence of species with different fire tolerances. The spatial patterns from individual plant distributions to landscape mosaics all affect the location and severity of future fires and their ecological impacts.

Fires are obvious, discrete events, and research on wildfire history has provided many useful models and methods for characterizing temporal and spatial properties of disturbance regimes (Moritz *et al.*, 2009). Temporal properties such as time-since-fire (survivorship) and fire interval (mortality) are described using probability distribution functions, notably the Weibell and negative exponential distributions. Fire mosaics and their dynamics have been described using spatial Markov chain models, percolation theory, and spatial autocorrelation statistics. These quantitative characterizations are useful for describing changes in fire regime over time, for comparing different areas, and for estimating wildfire risk.

Disturbance by Herbivores and Predators

In many ways, consumers (and pathogens) have the same effects as other agents of disturbance in that biomass is removed and new opportunities for recruitment are created. There are some distinctive features, however, of disturbances by herbivores and carnivores. Most obviously, they operate at the relatively fine scale of individual consumers, and they leave a very different biological legacy than mechanical and physico–chemical disturbances because biomass is concentrated and converted into heat, new tissues, and discrete waste products. The waste products are frequently deposited away from the

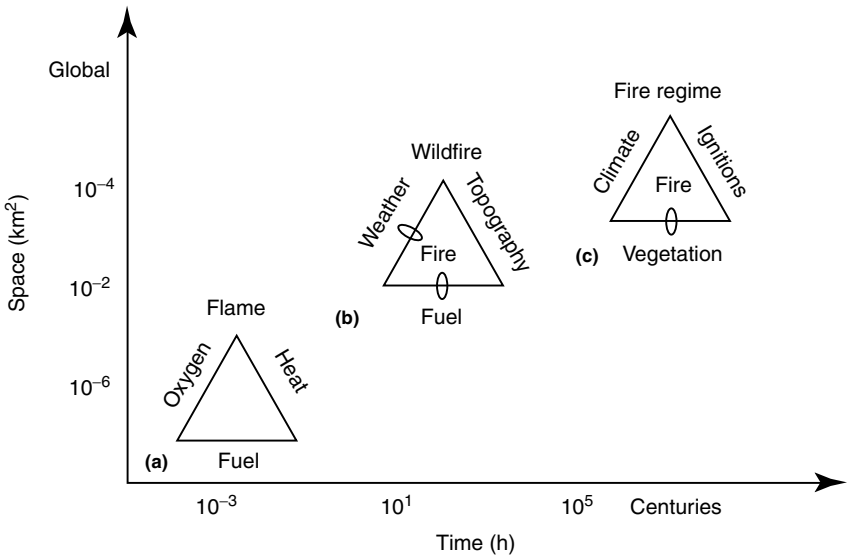


Figure 1 Hierarchical model of fire controls at successively coarser scales of space and time. Connections between fire and individual controls indicate potential feedback loops (reproduced from Moritz MA (ed.) (1999) *Controls on Disturbance Regime Dynamics: Fire in the Los Padres National Forest*. Ph.D Dissertation, Department of Geography. Santa Barbara: University of California).

disturbed area. For example, frass from defoliating caterpillars in tree canopies falls to the soil surface. Shade-seeking ungulates in African savannas deposit a disproportionate amount of urine and feces under trees. Elk migrations lead to a net transport of nitrogen from summer ranges to winter ranges.

Foraging is also a far more selective form of disturbance than other disturbance mechanisms. Organisms (including humans) preferentially occupy specific sites and microsites and focus on specific plant resources, creating patchiness at many scales. Most plants are unpalatable to most herbivores, and selective herbivory by dominant grazers can exert strong, directional effects on community composition, structure, and ecosystem processes. These effects may be amplified by strong feedbacks among consumption, food quantity, and food quality. For example, moose in boreal ecosystems selectively browse nitrogen-rich deciduous species and avoid coniferous species. Over time, this promotes nitrogen sequestration in conifer litter and accelerates succession of deciduous to coniferous plant communities (Hobbs, 1996). Disease spread from infected to susceptible hosts and the development of resistance in infected individuals is somewhat analogous in terms of strong spatio–temporal feedbacks between the disturbance process and state of the system.

Many consumers act as mechanical, chemical, and biological agents of disturbance. Perhaps the most sensational example is that of beavers, which selectively fell trees, dam streams, and inundate floodplains. Similarly, burrowing animals such as pocket gophers can operate as ''physical ecosystem engineers'' creating pervasive changes in the abiotic and biotic environment by redistributing soil and nutrients, altering soil hydrology and canopy light regimes, and altering vegetation composition and dynamics.

Interactions among Disturbance Mechanisms

Disturbances can interact in complex, nonlinear ways across space and time. One disturbance may promote another, such as the occurrence of dry-season fire in Mediterranean-climate shrublands promotes flooding and debris flows during the ensuing wet season. Conversely, one disturbance mechanism may reduce the likelihood of another. For example, grazing reduces the likelihood of fires in grasslands and the likelihood of crown fires in some forest ecosystems. Nearshore communities that are heavily impacted by wave disturbance may also experience lowered rates of predation. It is not difficult to imagine many such interactions operating across a wide range of spatial and temporal scales.

Disturbances often play an important role in invasion and spread of exotic species. In Hawaii, for instance, invasion of the nitrogen-fixing shrub, *Myrica faya*, was facilitated when native vegetation was thinned by heavy ash deposition from a volcanic eruption. There is a growing list of examples in which disturbance promotes the spread of an invasive species and then that species initiates ecosystem changes that introduce new disturbance processes. For example, in the western US cattle grazing in shrub steppe promoted spread of the exotic grass, *Bromus tectorum*, and the presence of a dry grass layer then increased wildfire size and frequency in that system, in which fires were historically uncommon.

As mentioned in *Disturbance Ecology*, compounded disturbances, or disturbances of communities already stressed by abiotic or biotic forces, can have large and persistent effects on ecological communities if they occur in such rapid succession that they disrupt normal recovery processes. Repeated burning over a series of two or more years is a common means of converting shrublands to grasslands in semi-arid ecosystems. In a particularly dramatic example, Hughes (1994) describes the decline of corals in Jamaica owing to the combined effects of overfishing, two major hurricanes, and disease that caused mass mortality in sea urchins. These impacts appear to have collectively precipitated a massive phase shift from coral-dominated to algal-dominated reefs.

Humans as Agents of Disturbance

For many millennia, humans have been manipulating ecosystem processes with varying impacts on ecosystem function and species distribution. Human activities that create disturbances at spatio–temporal scales to which organisms are adapted generally do not effect long-term changes in community composition or structure. For example, shifting agriculture in the humid tropical forests, like hurricanes, creates small, temporary forest clearings and results in landscape mosaics within which most species can persist. In contrast, commercial forest cutting and cropland and pasture development are large and persistent disturbances that can produce regional extinction of species and require much longer periods for forest recovery (Attiwill, 1994).

Another widespread human impact has been to regulate or arrest disturbances such as fires, floods, or animal activities that create habitats on which some species depend. Even when human ecosystem manipulations are, by design, modeled after a particular disturbance, it is difficult to incorporate all disturbance-related changes in the environment. For example, forest harvesting can occur at the spatio–temporal scale of a natural fire regime, but the postfire nutrient pulse and legacy of dead snags are no longer factors in the recovering ecosystem.

As humans fragment terrestrial habitats through land-use conversion and aquatic systems through impoundments, diversions, dredging, trawling, and construction, permanent fragmentation of habitats is superposed on dynamic disturbance mosaics, causing changes in disturbance regimes. Such fragmentation can impede disturbance spread (e.g., the spread of fire or disease), alter disturbance regimes in remnant patches, and create edge environments with novel disturbance regimes. Local human activities can have large cumulative effects. For example, the building of roads alters the natural cycles of erosion and landslide initiation, whereas agricultural land use often leads to the concentrated nutrient inputs and changes storm flows in receiving water bodies. In fact, most human activities have the capacity for altering historical disturbance regimes in some way.

At the broadest scale, humans are changing global climate patterns, and nearly all mechanisms of disturbance and disturbance effects are ultimately tied to climate processes. Changes in disturbance regimes during global climate change are inextricably tied to other ecological responses to changing climate, affecting whether species can persist in their current

ranges or can invade new areas of more suitable environmental conditions. The tight coupling between biotic processes and structures and most disturbance patterns and processes – a recurrent theme in this article – creates complex feedbacks that could amplify or moderate the effects of global climate change on disturbance regimes, species, and ecosystems.

See also: Diseases, Conservation and. Ecosystem Function, Principles of. Fires, Ecological Effects of. Grazing, Effects of. Herbaceous Vegetation, Species Richness in. Human Impacts on Ecosystems: An Overview

References

Attiwill PM (1994) The disturbance of forest ecosystems: The ecological basis for conservative management. *Forest Ecology and Management* 63: 247–300.

Christensen NL (1988) Succession and natural disturbance: Paradigms, problems, and preservation of natural ecosystems. In: Agee JK and Johnson DR (eds.) *Ecosystem Management for Parks and Wilderness*. Seattle: University of Washington Press.

Connell JH (1978) Diversity in tropical rainforests and coral reefs. *Science* 199: 1302–1310.

Everham EM and Brokaw NLV (1996) Forest damage and recovery from catastrophic wind. *Botanical Review* 62(2): 113–185.

Grime JP (1979) *Plant Strategies and Vegetation Processes*. New York: Wiley.

Hobbs NT (1996) Modification of ecosystems by ungulates. *Journal of Wildlife Management* 60: 695.

Holling CS (1973) Resilience and stability of ecological systems. *Annual Review of Ecology* 4: 1–23.

Hughes TP (1994) Catastrophes, phase shifts, and large scale degradation of a caribbean coral reef. *Science* 265: 1547–1551.

Levin SA (1992) The problem of pattern and scale in ecology. *Ecology* 73: 1943.

Moritz MA (ed.) (1999) *Controls on Disturbance Regime Dynamics: Fire in the Los Padres National Forest*. Ph.D Dissertation, Department of Geography. Santa Barbara: University of California.

Moritz MA, Moody TJ, Miles LJ, Smith MM, and deValpine P (2009) The fire frequency analysis branch of the pyrostatistics tree: Sampling decisions and censoring in fire interval data. *Environmental and Ecological Statistics* 16: 271–289.

Paine RT and Levin SA (1981) Intertidal landscapes: Disturbance and the dynamics of pattern. *Ecological Monographs* 51: 145–178.

Paine RT, Tegner MJ, and Johnson EA (1998) Compounded perturbations yield ecological surprises. *Ecosystems* 1: 535–545.

Pickett STA and White PS (eds.) (1985) *The Ecology of Natural Disturbance and Patch Dynamics*. New York: Academic Press.

Pollen-Bankhead N and Simon A (2009) Enhanced application of root reinforcement algorithms for bank stability modeling. *Earth Surface Processes and Landforms* 34: 471–480.

Reiners WA and Driese K (2003) Transport of energy, information, and material through the biosphere. *Annual Review of Environment and Resources* 28: 107–135.

Romme WH, Everham EH, Frelich LE, Moritz MA, and Sparks RE (1998) Are large, infrequent disturbances qualitatively different from small, frequent disturbances? *Ecosystems* 1: 524–534.

Sousa WP (1984) The role of disturbance in natural communities. *Annual Review of Ecology and Systematics* 15: 353–391.

Wayman RB and North M (2007) Initial response of a mixed conifer understory plant community to burning and thinning restoration treatments. *Forest Ecology and Management* 239: 32–44.

White PS and Pickett STA (1985) In: Pickett STA and White PS (eds.) *The Ecology of Natural Disturbance and Patch Dynamics*. New York: Academic Press.

Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems

Robert Keane, Missoula Fire Sciences Laboratory, Missoula, MT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Disturbance regime Cumulative effects of disturbance events over space and time.
Ecosystem management Conservation of major ecological services and restoration of natural resources while meeting society's needs for current and future generations.
Ecosystem process Any exchange of matter or energy within an ecosystem.
Focal landscape Area being evaluated in an HRV analysis.
Historical range of variation (HRV) Expression of the full range of landscape characteristics that occurred in the past.
Mosaic Spatial patch distribution of vegetation communities.
Scale Temporal or spatial extent before some quantity of interest changes.
Simulation buffer Area surrounding the context landscape that allows for the spread of spatial processes into the context landscape.
Simulation landscape Area being simulated to produce an HRV time series that usually includes both the context landscape and the simulation buffer.
Succession Process of plant assemblages replacing another; vegetation development.

Disturbance Regimes

Picture a tranquil landscape with undulating topography, idyllic streams, scenic glades, and verdant vegetation. Left to its own devices, this landscape would gradually become dominated by late successional communities that would slowly shift in response to climate changes over long time periods. This scene often forms the foundation and reference for most land management across the globe. However, this peaceful panorama rarely happens in nature because gradual successional change rarely drives landscape dynamics. Abrupt change is usually the rule, with vegetation development suddenly truncated by a set of ecological processes more dynamic than succession: disturbance. A wide variety of insect, disease, animal, fire, weather, and even human disturbances can interact with current and antecedent vegetation and climate to perturb the landscape and create a shifting mosaic of diverse seral vegetation communities and stand structures that in turn affect those very disturbances that created them. This complex interaction of vegetation, climate, and disturbance results in unique landscape behaviors that foster a wide range of landscape characteristics, which ensures high levels of biodiversity. The impacts of disturbances on landscape pattern, structure, and function drive most ecosystem processes, and it is disturbances that ultimately set the bounds of management for most landscapes of the world. In this chapter, disturbance regimes are discussed in terms of how they affect landscape dynamics and how historical disturbance regimes can form the range and variation of possible landscape conditions that can be used as a reference for managing today's landscapes.

Background

"Disturbance regime" is a general term that describes the temporal and spatial characteristics of a disturbance agent and the impact of that agent on the landscape. More specifically, a disturbance regime is the cumulative effects of multiple disturbance events over space and time. Any description of a disturbance regime must encompass an area that is large enough that the full range of disturbance sizes are manifest and with time measurement that is long enough that the full range of disturbance characteristics are captured. It is important to recognize that disturbance regimes are fundamentally different from individual disturbance events; for ecological restoration to succeed, for example, disturbance regimes should be emulated, not individual disturbance events, to fully capture the range and variation of disturbance effects. Biodiversity is intricately linked to disturbance regimes in that disturbances create shifting mosaics of diverse plant communities and habitats across a landscape (Watt, 1947), and the spatial and temporal fluctuations of these communities ensure the conservation of biodiversity (Naveh, 1994). Biodiversity is highest when disturbance is neither too rare nor too frequent on the landscape (Grime, 1973).

In this chapter, disturbance regimes can be generally described by 11 characteristics (**Table 1**) (Simard, 1996; Agee, 1993, 1998; Skinner and Chang, 1996). The disturbance "agent" is the entity that causes the disturbance, such as wind, fire, and beetles. Sometimes disturbance agents have a "source" that triggers the agent. Lightning can be a source for wildland fire, and heavy snow loads may be the source for avalanches. The disturbance agent occurs at a particular frequency that is often described over a period of time, depending on scale and objective. Point-level measures, such as disturbance return interval and occurrence probability, describe the number of disturbance events experienced over time at one point on the landscape (Skinner and Chang, 1996; Baker and Ehle, 2001). Spatial measures of disturbance rotation and disturbance cycle estimate the number of years it takes to disturb an area the size of the landscape (Johnson and Gutsell, 1994; van Wagner, 1978; Reed *et al.*, 1998). The frequency distribution of disturbance sizes on a landscape or region, for example, will depend primarily on the size and number of the largest events and landscape complexity (Yarie, 1981; Strauss *et al.*, 1989).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00389-0

Table 1 The 11 terms used to describe disturbance regimes in this chapter from Agee (1993), Simard (1996), and Skinner and Chang (1996)

Disturbance Characteristic	*Description*	*Example*
Agent	Factor causing the disturbance	Mountain pine beetle is the agent that kills trees
Source, cause	Origin of the agent	Lighting is a source for wildland fire
Frequency	How often the disturbance occurs or its return time	Years since last fire or beetle outbreak (scale dependent)
Intensity	A description of the magnitude of the disturbance agent	Mountain pine beetle population levels; wildland fire heat output
Severity	The level of impact of the disturbance on the environment	Percent mountain pine beetle tree mortality; fuel consumption in wildland fires
Size	Spatial extent of the disturbance	Mountain pine beetles can kill trees in small patches or across entire landscapes
Pattern	Patch size distribution of disturbance effects; spatial heterogeneity of disturbance effects	Fire can burn large regions, but weather and fuels can influence fire intensity and therefore the patchwork of tree mortality
Seasonality	Time of year of that disturbance occurs	Species phenology can influence wildland fires effects; spring burns can be more damaging to growing plants than fall burns on dormant plants
Duration	Length of time of that disturbances occur	Mountain pine beetle outbreaks usually last for 3–8 years; fires can burn for a day or for an entire summer
Interactions	Disturbance interact with each other, climate, vegetation, and other landscape characteristics	Mountain pine beetles can create fuel complexes that facilitate or exclude wildland fire
Variability	The spatial and temporal variability of the above factors	Highly variable weather and mountain pine beetle mortality can cause highly variable burn conditionsm resulting in patchy burns of small to large sizes

Disturbance "intensity" is the level of the disturbance agent as it occurs on the landscape. Insect and disease intensities are often described by population levels. Wildland fire intensity is described by its heat output. Windthrow intensity can be described by wind speed. "Severity" is different from intensity in that it reflects the impact of that disturbance and its characteristics on the biophysical environment. The primary and direct effects of most disturbance agents are biotic damage and mortality, but some physical disturbances such as fire can act on abiotic factors such as soil fertility, necromass consumption, and atmospheric emissions. Most disturbance regimes are described by their cumulative severities because it is the most important factor that directly impacts land management.

The "sizes" (area) and "patterns" (spatial variability) of disturbance events also shape disturbance regimes and influence biodiversity. Distributions of disturbance sizes reflect critical features of a disturbance regime, and many think that disturbance size distributions support the theory that they are self-organized processes (Malamud, 1998; Ricotta *et al.*, 1999). Moreover, patterns of disturbance severity and intensity often dictate landscape heterogeneity, which influences a wide variety of landscape characteristics such as wildlife habitat, hydrology, biodiversity, and other disturbances (Turner, 1987; Knight, 1987; Gustafson and Gardner, 1996). "Pattern" refers to the size, shape, and spatial location of the perturbed patches. Seasonality is the time of year of the disturbance event because plant and animal phenology can produce differential spatial effects, and disturbance patterns are often linked with the "duration" of the disturbance agent on the landscape, with durations ranging from seconds (wind) and minutes (avalanches) to days (fire) and years (insects).

The last two terms are the most important and tend to cross over all other disturbance descriptive terminology. It is the "variability" of disturbance characteristics such as severity, frequency, and size, coupled with the *interaction* of these characteristics with other disturbance characteristics such as previous patterns, duration, and seasonality, and climate that make disturbance regimes some of the most complex processes governing landscape dynamics and controlling biodiversity. It is this great complexity that confounds the description and classification of disturbance regimes into the simple abstractions often used in land management (Ryan and Noste, 1985). The highly variable spatial and temporal feedbacks and interactions of landscape patterns with disturbance characteristics and climate dynamics also make disturbance regimes so difficult to understand. It is far more illustrative to present the concept of disturbance regimes with an example from one of the world's most ubiquitous disturbances – wildland fire (Bowman *et al.*, 2009).

Disturbance Regime Example: Wildland Fire

Wildland fire regimes generally result from the cumulative interaction of fire, vegetation, climate, humans, and topography over time (Crutzen and Goldammer, 1993), though there are many other factors that influence disturbance regimes (e.g., other disturbances, weather, and fuels) (**Figure 1**). These interactions are spatially and temporally correlated. Future fires are influenced in space by the pattern of previously burned stands, fire-prone topographic features (e.g., midslopes, riparian bottoms), and areas with high fuel accumulations (e.g., older stands). They are influenced in time by the timing and severity of past climate (e.g., drought, wind), the rate of vegetation development (e.g., succession), and the frequency of other disturbance events (e.g., previous fires, insect outbreaks). A change in any of these factors will ultimately result in a change in the fire regime, and because all factors are constantly changing, fire regimes are inherently dynamic. For

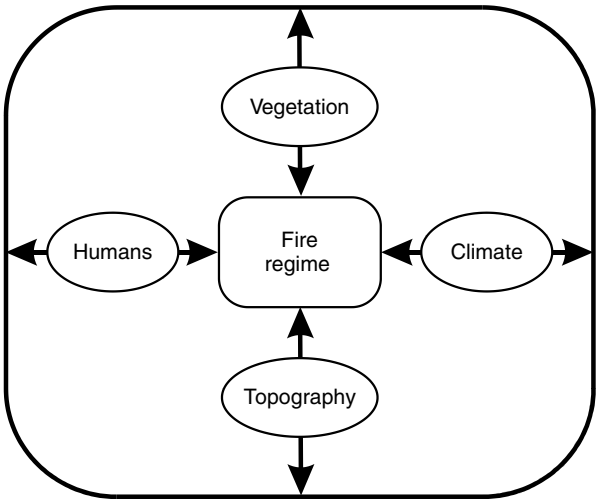


Figure 1 Four of the most important factors affecting fire regimes and their interactions.

example, climate change can affect fire regimes by altering fire ignition patterns (i.e., lightning), vegetation characteristics, and fuels (Flannigan and Van Wagner, 1991; Balling *et al.*, 1992; Swetnam, 1993). And exotic invasions such as white pine blister rust (*Cronarium ribicola*), cheatgrass (*Bromus tectorum*), and spotted knapweed (*Centaurea maculosa*), have changed fire regimes in many semiarid ecosystems of the US (Whisenant, 1990; Knapp, 1997).

Fire regimes are created by the interaction of bottom-up and top-down controls (Heyerdahl *et al.*, 2001). Bottom-up controls such as vegetation, fuels, topography, and patch distributions dictate fire spread, intensity, and severity at fine scales, and it is the bottom-up controls that can often be manipulated by land management. Coarse-scale, top-down controls are mostly climate and weather, which often dictate fire frequency, duration, and synchrony (Swetnam, 1990). Climate and weather trend signals are often embedded in a complex set of atmospheric teleconnections (simultaneous variations in climate observed over distant areas) interacting with land surface weather patterns. Drought-induced wildfires have been associated with global circulation anomalies such as the El Niño-Southern Oscillation (Swetnam and Betancourt, 1998; Heyerdahl *et al.*, 2002) and more recently the Pacific Decadal Oscillation (Hessl *et al.*, 2004). Coarse-scale, top-down climate interacts with fine scale landscape characteristics to influence the timing, size, frequency, and severity of fire events. As a result, a fire regime is actually a spatial disturbance gradient that does not follow discrete mapping units, so it should not be viewed as an attribute or characteristic of an ecosystem or cover type. Attempts to predict fire regimes solely from fuels (Olson, 1981), vegetation (Frost, 1998), or topography (Keane *et al.*, 2004b) have only partially succeeded because they did not recognize the spatial pervasiveness of fire on the landscape and the multiscale interactions of all factors that control fire dynamics (Morgan *et al.*, 2001).

The role of ignition in the formation of fire regimes is relatively misunderstood. Although fires can start from several sources such as spontaneous combustion, volcanic eruptions, and sunlight magnification, two sources start most wildland fires – lightning and humans. Ignition patterns are quite different between lightning and humans, with lightning strikes being mostly randomly distributed across landscapes over long temporal scales (Fuquay, 1980; Van Wagtendonk, 1991). Historical and contemporary peoples, however, have tended to light more fires along major transportation routes and settlements. These ignition patterns can create different fire regimes, depending on the geographical location. Lightning strikes in moist, productive ecosystems rarely start a fire because the fuel is too wet most years and the lightning occurs in the seasons when fuel is the wettest (Barrows *et al.*, 1977; Fuquay, 1980). Humans, however, can start fires when the fuel is driest and can control the number and types of ignition.

Fire regimes are most often described in terms of severity and frequency (Agee, 1998; Heinselman, 1981) (**Table 1**) because these two factors are usually the most important to land management and also because these two characteristics are probably the most important factors in creating fire regimes. In general, species with fire-adapted survival traits such as thick bark, high crowns, and sprouting tend to dominate landscapes with frequent fires, whereas other fire-adapted traits such as serotiny, soil seed banks, and far-ranging dispersal become important as fires become less frequent (Grime, 1979; Connell and Slayter, 1977). Although the size, pattern, and shifting mosaic of fire severity can be quite complex, severity types are often grouped into three general categories: (1) nonlethal surface fires, (2) mixed-severity fires, and (3) stand-replacement fires. Nonlethal surface fires are usually frequent and burn surface fuels at low intensities, causing low mortality ($<10\%$). Stand-replacement burns are usually rarer and reduce or kill the majority of the dominant vegetation, often trees and shrubs (90%) (Brown, 1995) as both lethal surface fires and active crown fires (Agee, 1993). Mixed-severity burns contain elements of both nonlethal surface and stand-replacement fires mixed in time and space (Arno *et al.*, 2000; Perry *et al.*, 2011). Passive crown fires, patchy stand-replacement fires, and low-intensity underburns are common in mixed-severity burns (DeBano *et al.*, 1998). Typically, mixed-severity fires are used to describe an area of patchy burn patterns created during one fire event. However, mixed-severity fire regimes can also be used to describe mixed-severity fires over time (e.g., nonlethal surface fire followed by stand-replacement fire) (Shinneman and Baker, 1997). There are other fire regime types, including ground fires (i.e., fire burning extensive organic layers), but they are not as prevalent as these three types (Agee, 1993).

Fire frequency and severity can be finely to broadly quantified using a number of field methods (Swetnam *et al.*, 1999; Humphries and Bourgeron, 2001; Heyerdahl *et al.*, 2001). Fire scar dates can be measured from trees, snags, stumps, and downed logs, but scarred trees are rarely distributed across large regions and diverse ecosystems at the densities needed to adequately describe fire regimes and corresponding landscape characteristics (Heyerdahl, 1997). Landscapes with long fire return intervals dominated by stand-replacement fire, for example, contain few fire-scarred trees. Charcoal samples from lake and ocean sediments and soil profiles provide important sources of historical fire data, but the temporal resolution of

the data is often inadequate for quantifying the annual variation in fire regimes (Whitlock and Millspaugh, 1996). Burn-boundary maps or fire atlases are another source for quantifying fire regimes, but these maps usually span a short temporal scale and rarely describe fire severity (Rollins *et al.*, 2001).

Perhaps the best way to illustrate the impact of a fire regime on the landscape is to describe a specific example from the high mountains of western North America (see **Box 1**). Whitebark pine (*Pinus albicaulis*) forests are declining across most of their range because of the combined effects of the modification of three disturbance factors (Arno, 1986). Furthermore, predicted changes in climate brought about by global warming could further exacerbate the decline by increasing the frequency and duration of beetle epidemics, blister rust infections, and severe wildfires (Logan and Powell, 2001; Running, 2006). The loss of whitebark pine could have serious consequences for the biodiversity of upper subalpine ecosystems because it is considered a keystone species (Tomback *et al.*, 2001). This "stone" pine produces large, wingless seeds that are an important food source for more than 110 animal species (Hutchins, 1994). In the Yellowstone ecosystem, the endangered grizzly bear (*Ursus arctos horribilis)* depends on whitebark pine seeds as a major food source, which it raids from red squirrel (*Tamiasciurus hudsonicus*) middens.

One major goal of ecosystem management, especially in the case of whitebark pine restoration, is to emulate historical disturbance regimes on contemporary landscapes to promote those processes that were considered necessary for healthy ecosystems. However, as is evident from the previous discussions, the complexity, interactions, and variability of disturbance regimes makes implementing historical disturbance regimes difficult if not impractical. An easier and more straightforward method of assessing ecosystem condition and evaluating ecosystem health was needed. The variability of historical disturbance regimes on landscape dynamics provided a foundation for a novel method of assessing ecosystem condition.

Historical Range and Variation

The Background

To effectively implement ecosystem management, land managers found it necessary to obtain a reference or benchmark to represent the conditions that described fully functional ecosystems (Cissel *et al.*, 1994; Laughlin *et al.*, 2004). Contemporary conditions could be evaluated against this reference to determine status, trend, and magnitude of change and also to design treatments that provide society with its sustainable and valuable resources while also returning declining ecosystems to a more natural or sustainable condition (Hessburg *et al.*, 1999b; Swetnam *et al.*, 1999). Ecologists and land managers were beginning to recognize that landscapes were not static but constantly changing, so it was critical that these reference conditions represented the dynamic character of ecosystems and landscapes as they vary over time and across space (Watt, 1947; Swanson *et al.*, 1994). Describing and quantifying ecological health has always been difficult because ecosystems are highly complex, with immense biotic and disturbance variability and diverse processes interacting across multiple space and time scales from genes to species to landscapes, and from seconds to days and centuries. One central concern with implementing ecosystem management was identifying appropriate reference conditions that could be used to describe ecosystem health, prioritize those areas in decline for possible treatment, and design feasible treatments for restoring their health (Aplet *et al.*, 2000).

Box 1 Whitebark pine ecology, management, and HRV

Whitebark pine is a long-lived, seral tree of moderate shade tolerance that can live well beyond 400 years (one tree is more than 1300 years), and on many sites it is eventually replaced, in the absence of fire, by the shade-tolerant subalpine fir (*Abies lasiocarpa*), with minor amounts of spruce (*Picea engelmannii*) and mountain hemlock (*Tsuga mertensiana*) in the mesic parts of its range (Arno and Hoff, 1990; Keane, 2001). Clark's nutcracker (*Nucifraga columbiana*) plays a critical role in the dispersal of whitebark pine's heavy, wingless seed (Tomback, 1998; Lorenz *et al.*, 2008) (**Photo 1**). The bird harvests seed from purple cones during late summer and early fall and carries as many as 100 of them in a sublingual pouch to sites as far as 10–20 km away, where it buries as many as 15 seeds in caches 2–3 cm below the soil surface. Seeds that remain unclaimed eventually germinate and grow into whitebark pine trees (Tomback, 2005). Nutcrackers often cache in open areas with a high degree of ground pattern in high-mountain settings that are often created by wildland fire (Morgan and Bunting, 1989).



Photo 1 Clark's nutcracker harvesting seeds from a whitebark pine (Photo by D. Tomback).

Three types of fires describe the diverse fire regime that occurs in whitebark pine forests (Morgan and Bunting, 1989; Murray *et al.*, 1998). Some high elevation whitebark pine stands experience nonlethal surface fires because sparse fuel loadings and widely spaced trees foster low-intensity fires (Keane *et al.*, 1994). The more common, mixed-severity fire regime is characterized by fires of mixed severities in space and time, creating complex mosaics of tree survival and mortality on the landscape (Murray, 2008; Arno and Weaver, 1990). Burned patches are important caching habitat for Clark's nutcracker because it prefers to cache in nearby burned patches, probably because of the heavy seeds (Norment, 1991). Many whitebark pine forests in Montana, northern Idaho, and the Cascades originated from large stand-replacement fires that occurred at long time intervals (greater than 250 years) where nutcrackers cache whitebark pine great distances (Keane *et al.*, 1994; Murray, 2008). The great variability in historical fire severity, frequency, pattern, and size has ensured the continued dominance of whitebark pine on upper subalpine landscapes in western North America.

The relatively new concept of historical range and variability (HRV) was introduced in the 1990s to bring understanding of past spatial and temporal variability into ecosystem management (Cissel *et al.*, 1994). HRV provided land-use planning and ecosystem management a critical spatial and temporal foundation to plan and implement possible treatments to improve ecosystem health and integrity (Landres *et al.*, 1999). Why not let recent history be a yardstick to compare ecological status and change by assuming that recent historical variation represents the broad envelope of conditions that supports landscape resilience and its self-organizing capacity (Hessburg *et al.*, 1999b)? Managers initially used "target" conditions developed from historical evidence to craft treatment prescriptions and prioritize areas (Harrod *et al.*, 1999). However, these target conditions tended to be subjective and somewhat arbitrary because they represented only one possible situation from a wide range of conditions that could be created from historical disturbance and vegetation dynamics (Keane *et al.*, 2002). This single objective, target-based approach was then supplanted by a more-comprehensive theory of HRV that incorporated the full variation and range of conditions that occurred across multiple scales of time and space into a metric of ecosystem health.

The idea of using historical conditions as reference for land management is not new (Egan and Howell, 2001). Since the 1990s planners have been using target stand and landscape conditions that resemble historical analogs to guide landscape management, and research has provided various examples (Christensen *et al.*, 1996; Fulé *et al.*, 1997; Harrod *et al.*, 1999). However, the inclusion of temporal variability of ecosystem elements into land management has only recently been employed. Landres *et al.* (1999) presented some of the theoretical underpinnings behind HRV and extensive reviews, and other background material on HRV and associated terminology can also be found (Egan and Howell, 2001; Swanson *et al.*, 1994; Kaufmann *et al.*, 1994; Morgan *et al.*, 1994; Foster *et al.*, 1996; Millar, 1997; Aplet and Keeton, 1999; Hessburg *et al.*, 1999a; Perera *et al.*, 2004; Veblen, 2003). This section was taken from material in Keane *et al.* (2009). The major advancement of HRV over the historical target approach is that the full range of historical ecological characteristics is used as the critical criterion in the evaluation and management of ecosystems (Swanson *et al.*, 1994). It is this variability that ensures continued health, self-organization, and resilience of biodiversity, ecosystems, and landscapes across spatiotemporal scales (Holling, 1992). Understanding the causes and consequences of this variability is key to managing landscapes that sustain ecosystems and the services they offer to society.

The Concept

The theory behind HRV is that the broad historical envelope of possible ecosystem conditions such as burned area, vegetation cover type area, and patch size distribution can provide a representative time series of reference conditions to guide land management (Aplet and Keeton, 1999) (see **Figure 2** as an example). This theory assumes the following: (1) Ecosystems are dynamic, not static, and their responses to changing processes are represented by past variability (Veblen, 2003); (2)

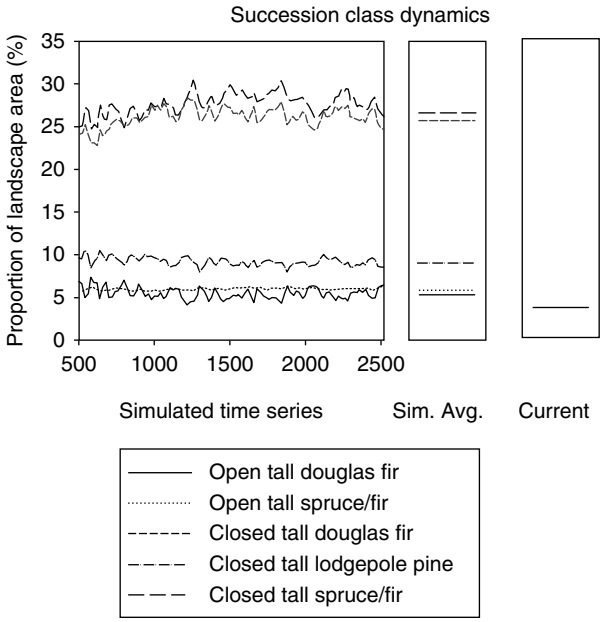


Figure 2 An example of a simulated time series for southwestern Montana, USA, mountainous landscape showing the five most dominant succession communities for simplicity. Also shown are the simulated average and the composition of the current landscape with only one succession class present (Keane *et al.*, 2008).

ecosystems are complex and have a range of conditions within which they are self-sustaining, and beyond this range they transition to disequilibrium (Egan and Howell, 2001; Wu *et al.*, 2006); (3) historical conditions can serve as a proxy for ecosystem health (Swetnam *et al.*, 1999); (4) the time and space domains that define the HRV are sufficient to quantify observed variation (Turner *et al.*, 1993); and (5) the ecological characteristics being assessed for the ecosystem or landscapes match the management objective (Keane *et al.*, 2002). In this chapter, we will focus the discussion on historical variations of landscape, not stand dynamics, and specifically landscape composition (i.e., aerial extant of vegetation communities) and structure (i.e., patch distributions).

Any quantification of HRV requires an explicit specification of the spatial and temporal context. The spatial context is needed to ensure that the variation of the selected landscape attribute is described across the most appropriate area relative to the spatial dynamics of ecosystem or landscape processes. The variability of the area occupied by a vegetation type over time, for example, generally decreases as the spatial context increases until it reaches an asymptote, which can be used to approximate optimal landscape size (Fortin and Dale, 2005; Karau and Keane, 2007). The optimal size of evaluation area will depend on (1) the ecosystem attribute evaluated, (2) the dynamics of major disturbance regimes, and (3) the relevant management issues (Tang and Gustafson, 1997). Fine woody fuel loadings, for example, would vary across smaller scales than coarse woody debris loads (Tinker and Knight, 2001).

The time scale over which HRV is evaluated must also be specified to properly interpret the underlying biophysical processes that influenced historical ecosystem dynamics,

especially climate (Millar and Woolfenden, 1999). The HRV of landscape composition evaluated from 1300 to 1600 AD might be entirely different if evaluated from 1600 to 1900 AD because of the vast differences in climates and land use between those periods (Mock and Bartlein, 1995). Temporal scale and resolution is usually dictated by the temporal depth of the historical evidence used to define and describe HRV, but they can also be selected to match specific management objectives. These time and space scale constraints are both a benefit and limitation of the HRV concept.

One advantage of the HRV approach is that it can use many elements to describe ecosystems, stands, or landscapes at any scale (Egan and Howell, 2001). The HRV of tree basal area, for example, can be assessed at the stand, landscape, and regional spatial scales; similarly, the HRV for landscape composition and patch structure can be computed for a watershed, National Forest, or an entire region. This multiscaled, multicharacteristic approach allows HRV attributes to be matched to the specific land management objectives at their most appropriate scale. For instance, fuel managers might decide to evaluate the HRV of coarse woody fuels and severe fire behavior at a watershed level (Hessburg *et al.*, 2007) to manage landscapes for continued ecological integrity and high biodiversity. Similarly, each HRV element can be prioritized or weighed based on its importance to the land management objective. This forms the critical linkage to adaptive land management in which iterative HRV analyses can be used to balance trade-offs in landscape integrity of ecosystems with other social issues and economic values. Perhaps the biggest advantage of HRV is that it forced land managers to recognize the dynamic character of landscapes in crafting management plans (Keane *et al.* 2009).

Its Application

HRV has been used in many land management projects. The departure of current conditions from historical variations have been used to prioritize and select areas for possible restoration treatments (Reynolds and Hessburg, 2005; Hessburg *et al.*, 2007) or to identify areas to conserve biological diversity (Aplet and Keeton, 1999). US fire management agencies have used fire regime condition class (FRCC), which is based on HRV of fire and vegetation dynamics, to rate and prioritize lands for fuel treatments (Hann and Bunnell, 2001; Schmidt *et al.*, 2002; Hann, 2004) (http://www.frcc.gov). The HRV of patch size and contagion was demonstrated to design the size of treatment area and landscape composition to select the appropriate management treatments to mimic patch characteristics (Keane *et al.*, 2002). Keane et al. (1996) used the HRV approach to design coarse-scale restoration of the previously discussed whitebark pine ecosystem in the Pacific Northwest.

Reference conditions for HRV have been described for many ecosystems across the western US and Canada. Veblen and Donnegan (2005) synthesized available knowledge on forest conditions and ecosystem disturbance for national forest lands in Colorado, USA. The ecological and economic implications of forest policies designed to emulate historical fire regimes were investigated by Thompson *et al.* (2006) using a simulation approach. Historical vegetation and disturbance dynamics for southern Utah were summarized in the Hood and Miller (2007) report. Wong *et al.* (2003) compiled an extensive reference of historical disturbance regimes for the entire province of British Columbia, Canada. Several studies have detailed historical variations in upland vegetation for two national forests in Wyoming (Dillon *et al.*, 2005; Meyer *et al.*, 2005). Although these studies are good qualitative references for understanding and interpreting historical conditions, they do not provide the quantitative detail needed to implement the described reference conditions directly into management applications.

Quantification of HRV demands temporally deep and spatially explicit historical data. Data sets that represent long-term empirical landscape dynamics are rarely available, inconsistent, and difficult to obtain (Humphries and Bourgeron, 2001; Barrett *et al.*, 2006). Historical reconstructions of landscape characteristics can be made from many sources if they exist for a particular landscape (see Egan and Howell, 2001, for a summary). Historical vegetation conditions can be reconstructed or described from (1) pollen deposits in lake or ocean sediments; (2) plant macrofossil assemblages deposited in middens, sediments, soils, and other sites; (3) dendrochronological stand reconstructions and fire scar histories; (4) land survey records; and (5) repeat photography (Gruell *et al.*, 1982; Arno *et al.*, 1995; Humphries and Bourgeron, 2001; Friedman *et al.*, 2001; Montes *et al.*, 2005; Schulte and Mladenoff, 2005). Unfortunately, these data have either a confined or unknown spatial domain because they were collected on a very small portion of the landscape, or they pertain to a broad, undefined area (middens, lake sediments) and lack spatial specificity with respect to patterns. Moreover, some ecosystems on a landscape have little evidence of past conditions with which to quantify HRV, and any available data are usually limited in temporal extent. In general, those methods that describe HRV at fine time scales, such as tree fire scar dating, are constrained to multicentenary time scales, whereas those methods that cover long time spans (millennia), such as pollen and charcoal analyses, have a resolution that may be too coarse for management of spatial patterns of structure and composition (Swetnam *et al.*, 1999).

For most landscape-level HRV quantification, there are three main sources of spatial data to quantify historical conditions (Keane *et al.*, 2006; Humphries and Bourgeron, 2001). The best sources are spatial chronosequences or digital maps of historical landscape characteristics over many time periods. Unfortunately, temporally deep and spatially explicit time series of historical conditions are missing for many US landscapes because aerial photography and satellite imagery are rare or were nonexistent before 1930 AD and comprehensive maps of forest vegetation are scarce, inconsistent, and limited in coverage prior to 1900. Tinker *et al.* (2003) quantified HRV in landscape structure using digital maps of current and past landscapes in the Greater Yellowstone Area from aerial photos and stand age interpretation.

Another HRV data source is to substitute space for time and collect spatial data across similar landscapes, from one or more times, across a large geographic region (Hessburg *et al.*, 1999a, 1999b). This assumes landscapes of similar biophysical environments with similar disturbance and climatic

regimes can provide a representative cross section of temporal variation of landscape dynamics. In effect, differences in space are equivalent to differences in time, and inferences may be drawn regarding variation in spatial pattern that might occur at a single location over time. However, subtle differences in landform, relief, soils, and climate make each landscape unique, and grouping landscapes may tend to overestimate range and variability of landscape characteristics (Keane *et al.*, 2002). Landscapes may be similar in terms of the processes that govern vegetation such as climate, disturbance, and species succession, but topography, soils, land use, and wind direction also influence vegetation development and fire growth (Knight, 1987).

A third approach for quantifying HRV involves using computer models to simulate historical dynamics to produce a time series of simulated data to compute HRV statistics and metrics. This approach relies on the accurate simulation of succession and disturbance processes in space and time (Keane *et al.*, 2002a). Many spatially explicit ecosystem simulation models are available for quantifying HRV patch dynamics (Gardner *et al.*, 1999; Mladenoff and Baker, 1999; Humphries and Baron, 2001; Keane *et al.*, 2004a), but most are (1) computationally intensive, (2) difficult to parameterize and initialize, and (3) overly complex, thereby making them difficult to use, especially for large regions, long time periods, and inexperienced staffs. However, those landscape models designed specifically for management planning may oversimplify vegetation development and disturbance. Even the most complex landscape models rarely simulate spatial interactions between climate, fire dynamics, and vegetation development at the scales needed to quantify HRV because of the lack of critical research in those areas and the immense amount of computer resources required for such an effort. Even with its shortcomings, simulation modeling is the most common method of creating HRV time series, and many studies have used simulation modeling to quantify HRV time series for landscapes and ecosystems using a wide variety of models. Nonspatial models such as VDDT (Beukema and Kurtz, 1995) were used to estimate landscape composition in a wide variety of areas from the Pacific Northwest to the northern Rocky Mountains (Hann *et al.*, 1997; Barbour *et al.*, 2007; Kurz *et al.*, 1999). The LADS model (Wimberly *et al.*, 2000) was used for the Oregon Coast Range to determine the appropriate level of old growth forests to quantify HRV in landscape structure (Nonaka and Spies, 2005), and to simulate the effect of forest polices (Thompson *et al.*, 2006). McGarigal *et al.* (2003) quantified historical forest composition and structures of Colorado landscapes. Keane *et al.* (2002) simulated historical landscape patch dynamics using the LANDSUM model for northern Rocky Mountain USA landscapes. The LANDFIRE prototype project quantified historical time series for landscapes across the US using the LANDSUM model (Keane *et al.*, 2007) (**Figure 3**).

There is a common misconception that long-term simulation model HRV outputs are inappropriate because the simulation of fire and landscape dynamics occurred while unrealistically holding climate and fire regimes constant (Keane *et al.*, 2006). This would be true if the objective of the modeling were to replicate historical fire events. However, the primary purpose of HRV modeling efforts is to describe variation in historical landscape dynamics, not to replicate them. Simulation modeling allows the quantification of the entire range of landscape conditions by simulating the static historical fire regime for long time periods (e.g., thousands of years) to ensure all possible fire ignitions and burn patterns are represented in the HRV time series. In contrast, HRV time series from empirical historical records will tend to underestimate variation of landscape conditions because there are a limited number of fire events in the historical record. Model input parameters represent the actual temporal context, whereas the simulation time represents the length of time needed to adequately capture the range and variation of historical conditions. Because the temporal domain of model parameters often represents only four or five centuries, it may seem that only 500 years of simulation are needed. However, the parameters quantified from sampled fire events that occurred during this time represent only one unique sequence of the fire ignitions and growth that created the unique landscape compositions observed today. If these events happened at different times or in different areas, an entirely different set of landscape conditions would have resulted.

Its Limitations

Although easily understood, the concept of HRV can be quite difficult to implement due to scale, data, and analysis limitations (Wong and Iverson, 2004). Inappropriate temporal and spatial scales to evaluate landscapes will introduce bias and increased variability into the computation of HRV statistics because the scales of climate, vegetation, and disturbance interactions are inherently different across landscapes (Morgan *et al.*, 1994). Karau and Keane (2007) found that simulated HRV chronosequences of landscapes smaller than 100 km^2 had increased variability in landscape composition due to the truncated spatial dynamics of simulated disturbance processes as they ran into the landscape boundary. This is why HRV approaches are inappropriate when applied on small areas such as stands. A Douglas-fir stand that was historically dominated by ponderosa pine, for example, may be appear to be outside HRV, but if it is within a 100-km^2 landscape composed primarily of ponderosa pine, it will certainly be within HRV. However, when evaluation landscapes are large (>500 km^2), it is often difficult to detect significant changes caused by ecosystem restoration or fuel treatments implemented on small areas (Keane *et al.* 2006). There is an optimal landscape extent for HRV simulation, but this optimum depends on subtle differences in topography, climate, and vegetation across large regions, and it also changes with spatial resolution.

There are few statistical techniques to compare HRV time series data to current landscape composition and structure (**Figure 2**). Many have used departure statistics to describe the dissimilarity of current conditions from historical variations to prioritize and select areas for possible restoration treatments (Reynolds and Hessburg, 2005; Hessburg *et al.*, 2007) or areas to conserve biological diversity (Aplet and Keeton, 1999). These departure methods use the similarity metrics from landscape and community ecology to estimate departure from

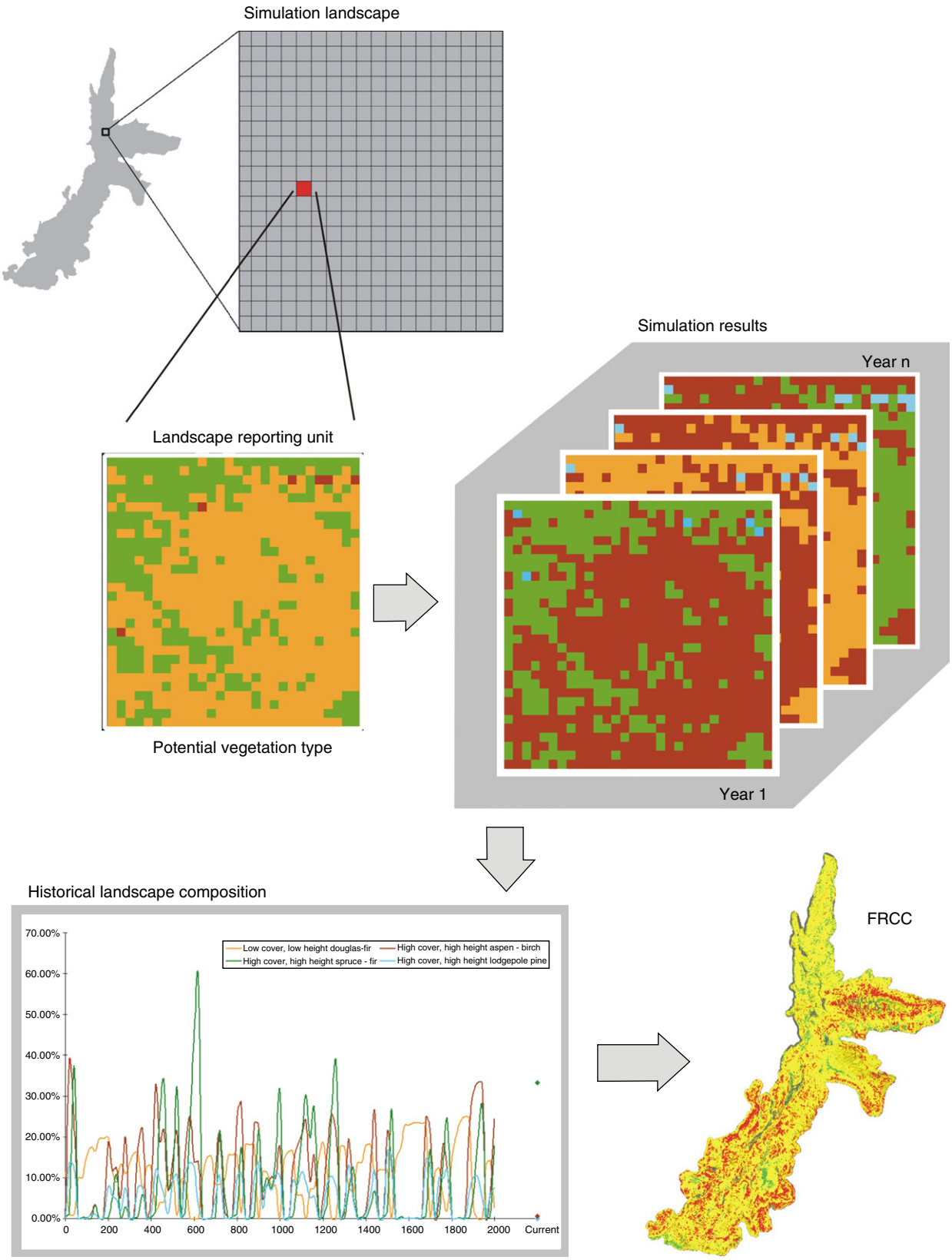


Figure 3 The computation of departure from HRV used in the LANDFIRE project (Keane *et al.* 2007). A simulated historical landscape composition time series is created using the LANDSUMv4 model for a landscape reporting unit (focal landscape). This time series is then compared to current conditions using an ecological similarity analysis to compute fire regime condition class (FRCC). FRCC is a three-category ordinal index reflecting the degree of departure from historical conditions, with red indicating areas where the landscapes are the most departed.

HRV, and most have limitations when used in HRV applications (Keane *et al.*, 2011) (**Figure 3**). For example, the Sorenson's index is sensitive to the number of classes used to describe the landscape (Keane *et al.*, 2008, 2011). Furthermore, some departure indices are insensitive to subtle changes in landscape composition when the same categories appear in all time sequences. Better multivariate statistical techniques with hypothesis testing are needed for more-credible HRV analyses.

Many limitations are associated with the use of the simulation approach to quantify HRV (Keane et al., 2011). It is impossible to build a model that includes all those landscape and ecosystem processes that directly affect those variables selected to represent the HRV time series. Information and data concerning the important processes and their linkages used to construct models is often inadequate. And, as model complexity increases, so does instability, computational requirements, input parameters, and reliability; there is a trade-off between model complexity and applicability. Simple, empirical modeling approaches are easier to understand and yield the most accurate answers, but they are limited in scope, data intensive, and often incapable of directly simulating complex interactions. Mechanistic modeling approaches include greater detail in simulating ecological processes, making them more robust and comprehensive; these models, however, can be inaccurate, difficult to use, and somewhat unstable. There is also a lack of sufficient expertise and input parameter data to parameterize and execute most models, and it is difficult to test and validate models because there are few spatial historical time series that are temporally deep and in the right context for comparison with model results (Keane and Finney, 2003). As a result, the simulated variation also includes undesirable and unquantifiable sources such as unintended stochasticity, model flaws, inadequate parameterization, and oversimplifications. No single model will satisfy the varied HRV demands of management, so compromises in simulation design must always be made. The best model might not be the most useful because (1) few people know how to run the model, (2) there may be insufficient computer resources (software, hardware requirements) to run the model, and (3) there may be insufficient field data for input parameter approximation to run the model. Most HRV simulation projects are designed around the people responsible for their completion.

One major problem in defining landscapes for HRV quantifications is that the landscape edges create artificial boundaries across which spatial processes such as seed dispersal and wildland fire spread cannot traverse. Areas near the edge of the landscape, for example, have a limited number of surrounding pixels from which a seed can fall or a fire can spread into them (**Figure 4**). Spatial processes such as fires cannot immigrate into the simulation landscapes, resulting in decreased occurrence near landscape edges. This problem is exacerbated by biophysical processes such as fire, seed dispersal, and insect migrations that are controlled by directional vectors such as wind that differentially act on parts of a landscape. Areas downwind, for example, have a higher probability of burning than those upwind (Keane *et al.*, 2002a). The best way to mitigate the edge effects is to surround the simulation landscape with a buffer (**Figure 4**). Each landscape is unique, so buffer width may differ for each setting. HRV simulations should be inspected to determine if the buffer is large enough to minimize edge effects within the context landscape, keeping in mind that simulation time increases exponentially as landscapes get larger.

The shape of the focal landscape is also an important factor in the simulation of HRV landscape dynamics. Fire frequencies in long, thin landscapes, for example, tend to be underestimated because simulated fires rarely reach their full size because they reach the landscape boundary first. Even with a large buffer, simulated fires spread quickly across a thin context landscape and burn only a fraction of its full size. Keane *et al.* (2002a) found fire frequencies were approximately 20–40% less in long landscapes with high edge. Many like to use watersheds to define simulation landscapes, but watersheds are often long and linear with high edge, making them somewhat undesirable for HRV simulation. The best shapes are squares, rectangles or circles that are large enough to contain both the buffer and context landscapes. The directional orientation of the simulation landscape is also important. Long, thin landscapes that are oriented perpendicular to the wind direction will have far less burned area over time than the same landscapes oriented parallel to the wind (Keane *et al.*, 2002a). The orientation is especially important if elongated landscapes are positioned at right angles to the wind direction; they will tend to have significantly less fire and narrow HRV time series.

Perhaps the most important limitation of the HRV application is that it may no longer be a viable concept for managing lands in the future because of expected climate warming and increasing contemporary human activities across the landscape (Millar *et al.*, 2007). Tomorrow's climates might change so rapidly and dramatically that they will no longer be similar to historical climates that created past conditions, and the continued spread of exotic plants, diseases, and other organisms by human transport will permanently alter ecosystems. Climate warming is expected to trigger major changes in disturbance processes, plant and animal species dynamics, and hydrological responses (Botkin *et al.*, 2007; Schneider *et al.*, 2007) that may create new plant communities, change biodiversity assemblages, and alter landscapes where they will be quite different from historical analogs (Notaro *et al.*, 2007). Whitebark pine, for example, is predicted to significantly decline under new forecast climates, so does it make sense to restore the species on landscapes where they are predicted to be reduced? Therefore, it may seem obvious that using historical references may no longer be reasonable in this rapidly changing world. However, a critical evaluation of possible alternatives may indicate that HRV, with all its faults and limitations, might be the most viable approach for the near term because it has the least amount of uncertainty.

One other alternative to HRV is to forecast the future variations of landscapes under changing climates using highly complex spatial, empirical, and mechanistic models, and this option is fraught with compounding uncertainties. The range of predictions for future climate from the major general circulation models may actually be greater than the variability of climate over the past two or three centuries (Stainforth *et al.*, 2005). This uncertainty increases when we factor in society's

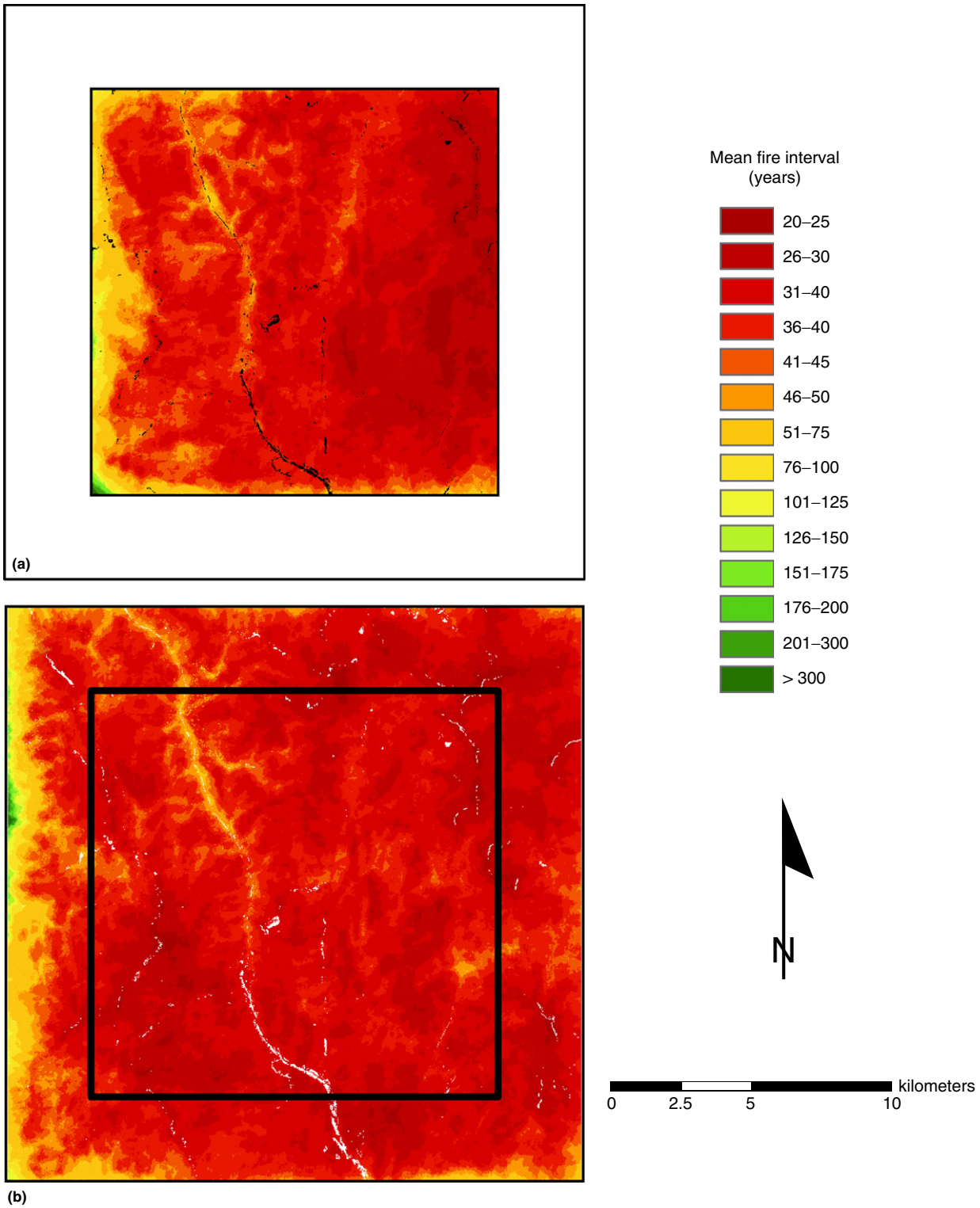


Figure 4 Illustration of the edge effect when simulating fire regimes across landscapes and mitigation of this effect through the inclusion of a buffer in the simulation landscapes.

projected responses to climate change through technological advances, behavioral adaptations, and population growth (Schneider *et al.*, 2007). Moreover, the variability of climate extremes, not the gradual change of average climate, will drive most ecosystem response to the climate-mitigated disturbance and plant dynamics, and these extreme events are the most difficult to predict (Easterling *et al.*, 2000, Smith, 2011). Uncertainty will also increase as the climate predictions are

extrapolated to the finer scales and longer time periods that are needed to quantify future range and variation (FRV) for landscapes. And this uncertainty will surely increase as we try to predict how ecosystems will respond to this simulated climate change (Araujo *et al.*, 2005). Mechanistic and empirical ecological models that simulate climate, vegetation, and disturbance dynamics across landscapes are often missing detailed representations and interactions of disturbance, hydrology, land use, and biological processes that will catalyze most climate interactions (Notaro *et al.*, 2007). Moreover, little is known about the interactions of climate with critical plant and animal life-cycle processes, especially reproduction and mortality (Keane *et al.*, 2001; Gworek *et al.*, 2007; Ibanez *et al.*, 2007; Lambrecht *et al.*, 2007), yet these process could be the most important in determining species response to climate change (Price *et al.*, 2001). Given large cumulative uncertainties involved in predicting future climates and subsequent ecosystem responses, it may be that HRV time series may have significantly lower uncertainty than any simulated predictions for the future. Recall that large variations in climates of the past several centuries are already reflected in the parameters used to simulate HRV time series. So before throwing the HRV baby out with the climate change bathwater, it may be more prudent to wait until simulation technology has improved enough to create credible FRV landscape pattern and composition time series for regional climate forecasts based on extensive model validation and testing, and this may take decades. In the meantime, it is doubtful that the use of HRV to guide management efforts will result in inappropriate activities considering the large genetic variation in most species (Rehfeldt *et al.*, 1999; Davis *et al.*, 2005).

In conclusion, the concept of HRV has a valid place in land management, at least for the near future. Landscape models can be used to simulate fire regimes and their interaction with climate and vegetation to create HRV time series that can be used as reference conditions to assess, plan, evaluate, design, and implement ecosystem-restoration treatments. HRV should be used to guide land management and not as a target on which to evaluate success or failure. There are few measures of ecosystem health that match the scale, scope, flexibility, and robustness of HRV analysis.

Appendix

List of Courses

1. Landscape Ecology and Management
2. Ecosystem Management

See also: Biodiversity and Ecosystem Services. Computer Systems and Models, Use of. Conservation Biology, Discipline of. Ecosystem Services. Fires, Ecological Effects of. Impact of Ecological Restoration on Ecosystem Services. Landscape Legacies. Landscape Modeling. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Restoration of Biodiversity, Overview. Wildlife Management

References

Agee JK (1993) Fire regimes and approaches for determining fire history. In: Hardy CC and Arno SF (eds.) *The Use of Fire in Forest Restoration, 1995*, pp. 12–13. Seattle, Washington: Intermountain Research Station, Forest Service, USDA.

Agee JK (1998) The landscape ecology of western forest fire regimes. *Northwest Science* 72: 24–34.

Aplet G and Keeton WS (1999) Application of historical range of variability concepts to biodiversity conservation. In: Baydack RK, Campa H, and Haufler JB (eds.) *Practical Approaches to the Conservation of Biological Diversity*. New York: Island Press.

Aplet G, Thomson J, and Wilbert M (2000) Indicators of wildness: Using attributes of the land to assess the context of wilderness. In: Mccool SF, Cole DN, Borrie WT, and O'loughlin J (eds.) *Wilderness Science in a Time of Change*, pp. 89–98. Missoula, MT: USDA Forest Service Rocky Mountain Research Station. RMRS-P-15, Vol 2.

Araujo MB, Whittaker RJ, Ladle RJ, and Erhard M (2005) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography Letters* 14: 529–538.

Arno SF (1986) Whitebark pine cone crops – A diminishing source of wildlife foods? *Western Journal of Applied Forestry* 1(3): 92–94.

Arno SF and Hoff RJ (1990) *Pinus albicaulis* Engelm. Whitebark Pine. *Silvics of North America. Vol. I. Conifers. Agr. Handbook*.

Arno SF, Parsons DJ, and Keane RE (2000) Mixed-severity fire regimes in the northern Rocky Mountains: Consequences of fire exclusion and options for the future. *Wilderness Science in a Time of Change Conference, Vol. 5: Wilderness Ecosystems, Threat, and Management, Missoula, Montana, May 23–27, 1999*. Fort Collins, CO: US Department of Agriculture Forest Service Rocky Mountain Research Station.

Arno SF, Scott JH, and Hartwell MG (1995) *Age-Class Structure of Old Growth Ponderosa Pine/Douglas-Fir Stands and its Relationship to Fire History*. Intermountain Research Station, Forest Service, USDA.

Arno SF and Weaver T (1990) Whitebark pine community types and their patterns on the landscape. In: Schmidt WC and McDonald KJ (eds.) *Proceedings – Symposium on Whitebark Pine Ecosystems: Ecology and Management of a High Mountain Resource*. Bozeman, MT, USA: USDA Forest Service. (Compilers).

Baker WL and Ehle D (2001) Uncertainty in surface-fire history: The case of ponderosa pine forests in the western United States. *Canadian Journal of Forest Research* 31: 1205–1226.

Balling RC, Meyer GA, and Wells SG (1992) Climate change in Yellowstone National Park: Is the drought-related risk of wildfires increasing? *Climate Change* 22: 35–45.

Barbour RJ, Hemstrom MA, and Hayers JL (2007) The Interior Landscape Analysis System: A step toward understanding integrated landscape analysis. *Landscape and Urban Planning* 80: 333–344.

Barrett SW, Demeo T, Jones JL, Zeiler JD, and Hutter LC (2006) Assessing ecological departure from reference conditions with the Fire Regime Condition Class (FRCC) mapping tool. In: Andrews PL and Butler BW (eds.) *Fuels Management – How to Measure Success*, pp. 575–585. Portland, OR, USA: USDA Forest Service Rocky Mountain Research Station.

Barrows JS, Sandberg David V, and Hart JD (1977) *Lightning Fires in Northern Rocky Mountain Forests*. Missoula, MT, USA: USDA Forest Service Intermountain Fire Sciences Laboratory.

Beukema SJ and Kurtz WA (1995) *Vegetation Dynamics Development Tool User's Guide*. Vancouver, BC, Canada: ESSA Technologies Ltd.

Botkin DB, Saxe H, Araujo MB, *et al.* (2007) Forecasting the effects of global warming on biodiversity. *Bioscience* 57: 227–236.

Bowman DMJS, Balch JK, Artaxo P, *et al.* (2009) Fire in the Earth system. *Science* 324: 481–484.

Brown JK (1995) Fire regimes and their relevance to ecosystem management. In: *Proceedings of the Society of American Foresters 1994 Annual Meeting*, pp. 171–178. Bethesda, MD.

Christensen NL, Bartuska AM, brown JH, *et al.* (1996) The report of the ecological society of America committee on the scientific basis for ecosystem management. *Ecological Applications* 6: 665–691.

Cissel JH, Swanson FJ, Mckee WA, and Burditt AL (1994) Using the past to plan the future in the Pacific Northwest. *Journal of Forestry* 92: 30–31. 46.

Connell JH and Slayter RO (1977) Mechanisms of succession and their role in community stabilization and organization. *American Naturalist* 111: 1119–1144.

Crutzen PJ and Goldammer JG (1993) *Fire in the Environment: The Ecological, Atmospheric and Climatic Importance of Vegetation Fires*. New York, USA: John Wiley and Sons.

Davis MB, Shaw RG, and Etterson JR (2005) Evolutionary responses to changing climates. *Ecology* 86: 1704–1714.
DeBano LF, Neary DG, and Ffolliott PF (1998) *Fire's Effect on Ecosystems*. New York, USA: John Wiley and Sons.
Dillon GK, Knight DH, and Meyer CB (2005) *Historic Range of Variability for Upland Vegetation in the Medicine Bow National Forest, Wyoming*. Fort Collins, CO, USA: USDA Forest Service Rocky Mountain Research Station.
Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, and Mearns LO (2000) Climate extremes: Observations, modeling, and impacts. *Science* 289: 2068–2074.
Egan D and Howell EA (eds.) (2001) *The Historical Ecology Handbook*. Washington, DC, USA: Island Press.
Flannigan MD and Van Wagner CE (1991) Climate change and wildfire in Canada. *Canadian Journal of Forest Research* 21: 66–72.
Fortin MJ and Dale MR (2005) *Spatial Analysis: A Guide for Ecologists*. Cambridge, UK: Cambridge University Press.
Foster DR, Orwig DA, and Mclachlan JS (1996) Ecological and conservation insights from reconstructive studies of temperate old-growth forests. *Trends in Ecology and Evolution* 11: 419–424.
Friedman SK, Reich PB, and Frelich LE (2001) Multiple scale composition and spatial distribution patterns of the north-eastern Minnesota presettlement forest. *Journal of Ecology* 89: 538–554.
Frost CC (1998 1998) Presettlement fire frequency regimes of the United States: A first approximation. *Proceedings from the Tall Timbers Fire Ecology Conference*. Tallahassee, FL: Tall Timbers Research Station.*Proceedings from the Tall Timbers Fire Ecology Conference*. Tallahassee, FL: Tall Timbers Research Station. pp. 70–81.
Fulé PZ, Covington WW, and Moore MM (1997) Determining reference conditions for ecosystem management of southwestern ponderosa pine forests. *Ecological Applications* 7: 895–908.
Fuquay DM (1980) Lightning that ignites forest fires. *Proceedings of the Sixth Conference on Fire and Meteorology*, April 22–24, Seattle, WA, USA. Society of American Foresters, Washington DC, USA, pp. 109–112.
Gardner RH, Romme William H, and Turner MG (1999) Predicting forest fire effects at landscape scales. In: Mladenoff DJ and Baker WL (eds.) *Spatial Modeling of Forest Landscape Change: Approaches and Applications*. Cambridge, UK: Cambridge University Press.
Grime JP (1979) *Plant Strategies and Vegetation Processes*. New York, USA: Wiley.
Grime JP (1973) Competitive exclusion in herbaceous vegetation. *Nature* 242(5396): 344–347.
Gruell GE, Schmidt WC, Arno SF, and Reich WJ (1982) *Seventy Years of Vegetative Change in a Managed Ponderosa Pine Forest in Western Montana – Implications for Resource Management*. Ogden, UT: USDA Forest Service, Intermountain Forest and Range Experiment Station.
Gustafson EJ and Gardner RH (1996) The effect of landscape heterogeneity on the probability of patch colonization. *Ecology* 77: 94–107.
Gworek JR, Vander Wall SB, and Brussard PF (2007) Changes in biotic interactions and climate determine recruitment of Jeffrey pine along an elevation gradient. *Forest Ecology and Management* 239: 57–68.
Hann WJ (2004) Mapping fire regime condition class: A method for watershed and project scale analysis. In: Engstrom RT, Galley KEM, and De Groot WJ (eds.) *22nd Tall Timbers Fire Ecology Conference: Fire in Temperate, Boreal, and Montane Ecosystems, 2004*, pp. 22–44. Tall Timbers Research Station
Hann WJ and Bunnell DL (2001) Fire and land management planning and implementation across multiple scales. *International Journal of Wildland Fire* 10: 389–403.
Hann WJ, Jones JL, Karl MG, *et al.* (1997) *An Assessment of Ecosystem Components in the Interior Columbia Basin and Portions of the Klamath and Great Basins. Vol. II – Landscape Dynamics of the BASIN*. USDA Forest Service Pacific Northwest Research Station
Harrod RJ, Mcrae BH, and Hartl WE (1999) Historical stand reconstruction in ponderosa pine forests to guide silvicultural prescriptions. *Forest Ecology and Management* 114: 433–446.
Heinselman ML (1981) Fire intensity and frequency as factors in the distribution and structure of northern ecosystems. In: Mooney HA, Bonnicksen TM, Christensen NL, Lotan JE, and Reiners WA (eds.) *Proceedings of the Conference Fire Regimes and Ecosystem Properties, 1981*, pp. 7–55. USDA Forest Service General Technical Report WO-26.
Hessburg PF, Reynolds KM, Keane RE, James KM, and Salter RB (2007) Evaluating wildland fire danger and prioritizing vegetation and fuels treatments. *Forest Ecology and Management* 247: 1–17.
Hessburg PF, Smith BG, and Salter RB (1999a) A method for detecting ecologically significant change in forest spatial patterns. *Ecological Applications* 9: 1252–1272.
Hessburg PF, Smith BG, and Salter RB (1999b) Detecting change in forest spatial patterns from reference conditions. *Ecological Applications* 9: 1232–1252.
Hessl AE, Mckenzie D, and Schellhaas R (2004) Drought and pacific decadal oscillation linked to fire occurrence in the inland pacific northwest. *Ecological Applications* 14: 425–442.
Heyerdahl E (1997) *Spatial and Temporal Variation in Historical Fire Regimes of the Blue Mountains, Oregon and Washington: The Influence of Climate*. PhD Dissertation, University of Washington.
Heyerdahl EK, Brubaker LB, and Agee JK (2001) Spatial controls of historical fire regimes: A multistage example from the interior west, USA. *Ecology* 82: 660–678.
Heyerdahl EK, Brubaker LB, and Agee JK (2002) Annual and decadal climate forcing of historical fire regimes in the interior Pacific Northwest, USA. *The Holocene* 12: 597–604.
Holling CS (1992) Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological Monographs* 62: 447–502.
Hood SM and Miller M (2007) *Fire Ecology and Management of the Major Ecosystems of Southern Utah*. Fort Collins, CO, USA: USDA Forest Service Rocky Mountain Research Station.
Humphries HC and Baron J (2001) Ecosystem structure and function modeling. In: Jensen ME and Bourgeron PS (eds.) *A Guidebook for Integrated Ecological Assessments*. New York, USA: Springer-Verlag.
Humphries HC and Bourgeron PS (2001) Methods for determining historical range of variability. In: Jensen ME and Bourgeron PS (eds.) *A Guidebook for Integrated Ecological Assessments*. New York, USA: Springer-Verlag.
Hutchins HE (1994) Role of various animals in dispersal and establishment of whitebark pine in the Rocky Mountains, USA. In: Schmidt WC and Holtmeier FK (eds.) *Proceedings: International Workshop on Subalpine Stone Pines and their Environment: The Status of Our Knowledge. St. Moritz, Switzerland*. Ogden, UT, USA: US Department of Agriculture, Forest Service, Intermountain Research Station. (Compilers).
Ibanez I, Clark JS, Ladeau S, and Lambers JHR (2007) Exploiting temporal variability to understand tree recruitment response to climate change. *Ecological Monographs* 77: 163–177.
Johnson EA and Gutsell SL (1994) Fire frequency models, methods and interpretations. *Advances in Ecological Research* 25: 239–287.
Karau EC and Keane RE (2007) Determining landscape extent for succession and disturbance simulation modeling. *Landscape Ecology* 22: 993–1006.
Kaufmann MR, Graham RT, Boyce DAJ, *et al.* (1994) *An Ecological Basis for Ecosystem Management*. Fort Collins, CO, USA: USDA Forest Service, Rocky Mountain Forest and Range Experiment Station.
Keane RE (2001) Successional dynamics : Modeling an anthropogenic threat. In: Tomback D, Arno S, and Keane R (eds.) *Whitebark Pine Communities: Ecology and Restoration*. Washington DC, USA: Island Press.
Keane RE, Cary G, Davies ID, *et al.* (2004a) A classification of landscape fire succession models: Spatially explicit models of fire and vegetation dynamic. *Ecological Modelling* 256: 3–27.
Keane RE and Finney MA (2003) The simulation of landscape fire, climate, and ecosystem dynamics. In: Veblen TT, Baker WL, Montenegro G, and Swetnam TW (eds.) *Fire and Global Change in Temperate Ecosystems of the Western Americas*. New York, USA: Springer-Verlag.
Keane RE, Garner J, Teske C, Stewart C, and Hessburg P (2001b) Range and variation in landscape patch dynamics: Implications for ecosystem management. *Proceedings from the 1999 National Silviculture Workshop, 2002a Kalispell*. MT, USA: USDA Forest Service Rocky Mountain Research Station. pp. 19–26.
Keane RE, Holsinger L, Parsons R, and Gray K (2008) Climate change effects on historical range of variability of two large landscapes in western Montana, USA. *Forest Ecology and Management* 254: 274–289.
Keane RE, Holsinger L, and Pratt S (2006) *Simulating Historical Landscape Dynamics Using the Landscape Fire Succession Model Landsum Version 4.0*. Fort Collins, CO, USA: USDA Forest Service Rocky Mountain Research Station.
Keane RE, Austin Michael, Dalman R, *et al.* (2001) Tree mortality in gap models: Application to climate change. *Climatic Change* 51: 509–540.
Keane RE, Morgan P, and Menakis JP (1994) Landscape assessment of the decline of whitebark pine (*Pinus albicaulis*) in the Bob Marshall Wilderness Complex, Montana, USA. *Northwest Science* 68: 213–229.
Keane RE, Parsons R, and Hessburg P (2002) Estimating historical range and variation of landscape patch dynamics: Limitations of the simulation approach. *Ecological Modelling* 151: 29–49.

Keane RE, Parsons R, and Rollins MG (2004b) Predicting fire regimes across multiple scales. In: Perera A and Buse L (eds.) *Emulating Natural Disturbances: Concepts and Techniques*. Cambridge, UK: Cambridge University Press.

Keane RE, Rollins MG, and Zhu Z (2007) Using simulated historical time series to prioritize fuel treatments on landscapes across the United States: The LANDFIRE prototype project. *Ecological Modelling* 204: 485–502.

Keane RE, James P Menakis and Hann WJ (1996) Coarse scale restoration planning and design in Interior Columbia River Basin Ecosystems – An example using Whitebark Pine (*Pinus albicaulis*) forests. *The Use of Fire in Forest Restoration – A General Session at the Annual Meeting of the Society of Ecosystem Restoration "Taking a broader view", USDA Forest Service General Technical Report INT-GTR-341*, Seattle, WA, USA. pp. 14–20.

Keane RE, Hessburg PF, Landres PB, and Swanson FJ (2009) A review of the use of historical range and variation (HRV) in landscape management. *Forest Ecology and Management* 258: 1025–1037.

Keane RE, Holsinger L, and Parsons RA (2011) *Evaluating Indices that Measure Departure of Current Landscape Composition from Historical Conditions*, 19 pp. USA: US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Knapp PA (1997) Spatial characteristics of regional wildfire frequencies in Intermountain West grass-dominated communities. *Professional Geographer* 49: 39–51.

Knight DH (1987) Parasites, lightning, and the vegetation mosaic in wilderness landscapes. In: Turner MG (ed.) *Landscape Heterogeneity and Disturbance*. New York: Springer-Verlag.

Kurz WA, Beukema SJ, Merzenich J, Arbaugh M, and Schilling S (1999) Long-range modeling of stochastic disturbances and management treatments using VDDT and TELSA. *Society of American Foresters 1999 National Convention, 1999*. Portland, OR, USA: Society of American Foresters. pp. 349–355.

Lambrecht SC, Loik ME, Inouye DW, and Harte J (2007) Reproductive and physiological responses to simulated climate warming for four subalpine species. *New Phytologist* 173: 121–134.

Landres PB, Morgan Penelope, and Swanson FJ (1999) Overview and use of natural variability concepts in managing ecological systems. *Ecological Applications* 9: 1179–1188.

Laughlin DC, Bakker JD, Stoddard MT, *et al.* (2004) Toward reference conditions: Wildfire effects on flora in an old growth ponderosa pine forest. *Forest Ecology and Management* 199: 137–152.

Logan JA and Powell JA (2001) Ghost forests, global warming, and the mountain pine beetle (Coleoptera: Scolytidae). *American Entomologist* 47: 160–173.

Lorenz TJ, Aubry C, and Shoal R (2008) *A review of the literature on seed fate in whitebark pine and the life history traits of Clark's Nutcracker and pine squirrels*. Portland, OR: USDA Forest Service Pacific Northwest Research Station.

Malamud BD, Morein Gleb, and Turcotte Donald L (1998) Forest fires: An example of self-organized critical behavior. *Science* 281: 1840–1842.

McGarigal K, Romme WH, Goodwin D, and Haugsjaa E (2003) *Simulating the Dynamics in Landscape Structure and Wildlife Habitat in Rocky Mountain Landscapes: The Rocky Mountain Landscape Simulator (Rmlands) and Associated Models*. Amherst, MA: Department of Natural Resources Conservation, University of Massachusetts.

Meyer CB, Knight DH, and Dillon GK (2005) *Historic Range of Variability for Upland Vegetation in the Bighorn National Forest, Wyoming*. Fort Collins, CO, USA: USDA Forest Service Rocky Mountain Research Station.

Millar CI (1997) Comments on historical variation and desired future conditions as tools for terrestrial landscape analysis. In: Sommarstrom S (ed.) *Sixth Biennial Watershed Management Conference*, pp. 105–131. University of California at Davis

Millar CI, Stephenson NL, and Stephens SL (2007) Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* 17: 2145–2151.

Millar CI and Woolfenden WB (1999) The role of climate change in interpreting historical variability. *Ecological Applications* 9: 1207–1216.

Mladenoff DJ and Baker WL (1999) *Spatial Modeling of Forest Landscape Change*. Cambridge, UK: Cambridge University Press.

Mock CJ and Bartlein PJ (1995) Spatial variability of late-quaternary paleoclimates in the western United States. *Quaternary Research* 44: 425–433.

Montes F, Sanchez M, Del Rio M, and Canellas I (2005) Using historical management records to characterize the effects of management on the structural diversity of forests. *Forest Ecology and Management* 207: 279–293.

Morgan P, Aplet GH, Haufler JB, Humphries HC, Moore MM, and Wilson WD (1994) Historical range of variability: A useful tool for evaluating ecosystem change. *Journal of Sustainable Forestry* 2: 87–111.

Morgan P and Bunting SC (1989) Whitebark pine: Fire ecology and management. *Women in Natural Resources* 11: 52.

Morgan P, Hardy CC, Swetnam TW, Rollins MG, and Long DG (2001) Mapping fire regimes across time and space: Understanding coarse and fine-scale fire patterns. *International Journal of Wildland Fire* 10: 329–342.

Murray MP (2008) Fires in the high Cascades: New findings for managing whitebark pine. *Fire Management Today* 68: 26–29.

Murray MP, Bunting SC, and Morgan P (1998) Fire history of an isolated subalpine mountain range of the intermountain region, United States. *Journal of Biogeography* 25: 1071–1080.

Naveh Z (1994) From biodiversity to ecodiversity: A landscape-ecology approach to conservation and restoration. *Restoration Ecology* 2: 180–189.

Nonaka E and Spies TA (2005) Historical range of variability in landscape structure: A simulation study in Oregon, USA. *Ecological Applications* 15: 1727–1746.

Norment CJ (1991) Bird use of forest patches in the subalpine forest-alpine tundra ecotone of the Beartooth Mountains, Wyoming. *Northwest Science* 65: 1–10.

Notaro M, Vavrus S, and Liu Z (2007) Global vegetation and climate change due to future increases in CO_2 as projected by a fully coupled model with dynamic vegetation. *Journal of Climate* 20: 70–88.

Olson J (1981) Carbon balance in relation to fire regimes. In: Mooney HA, Bonnicksen TM, Christensen NL, Lotan JE, and Reiners WA (eds.) *Proceedings of the Conference Fire Regimes and Ecosystem Properties*, pp. 327–378. USDA Forest Service (TECHNICAL COORDINATORS).

Perera AH, Buse LJ, and Weber MG (eds.) (2004) *Emulating Natural Forest Landscape Disturbances: Concepts and Applications*. New York, USA: Columbia University Press.

Perry DA, Hessburg PF, Skinner CN, *et al.* (2011) The ecology of mixed severity fire regimes in Washington, Oregon, and Northern California. *Forest Ecology and Management* 262: 703–717.

Price DT, Zimmerman NE, Van Der Meer PJ, *et al.* (2001) Regeneration in gap models: Priority issues for studying forest responses to climate change. *Climatic Change*

Reed WJ, Larsen CPS, Johnson EA, and Macdonald GM (1998) Estimation of temporal variations in historical fire frequency from time-since-fire map data. *Forest Science* 44: 465–475.

Rehfeldt GE, Ying CC, Spittlehouse DL, and Hamilton DA (1999) Genetic responses to climate in *Pinus contorta*: Niche breadth, climate change, and reforestation. *Ecological Monographs* 69: 375–407.

Reynolds KM and Hessburg PF (2005) Decision support for integrated landscape evaluation and restoration planning. *Forest Ecology and Management* 207: 263–278.

Ricotta C, Avena G, and Marchetti M (1999) The flaming sandpile: Self-organized criticality and wildfires. *Ecological Modelling* 119: 73–77.

Rollins MG, Swetnam TW, and Morgan P (2001) Evaluating a century of fire patterns in two Rocky Mountain wilderness areas using digital fire atlases. *Canadian Journal of Forest Research* 31: 2107–2133.

Running SW (2006) Is global warming causing more, larger wildfires. *Science* 313: 927–928.

Ryan KC and Noste NV (1985) Evaluating prescribed fires. In: Lotan JE, Kilgore BM, Fischer WC, and Mutch RW (eds.) *Wilderness Fire Symposium, Missoula, MT*, pp. 230–237. Ogden UT, USA: Intermountain Research Station. USDA Forest Service General Technical Report INT-182.

Schmidt KM, Menakis JP, Hardy CC, Hann WJ, and Bunnell DL (2002) *Development of Coarse-Scale Spatial Data for Wildland Fire and Fuel Management*. Fort Collins, CO, USA: USDA Forest Service Rocky Mountain Research Station.

Schneider SH, Semenov S, Patwardhan A, *et al.* (2007) Assessing key vulnerabilities and the risk from climate change. In: Parry ML, Canziani OF, Palutikof PJ, Van Der Linden P, and Hanson CE (eds.) *Climate Change 2007: Impacts, adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

Schulte LA and Mladenoff DJ (2005) Severe wind and fire regimes in northern forests: Historical variability at the regional scale. *Ecology* 86: 431–445.

Shinneman DJ and Baker WL (1997) Nonequilibrium dynamics between catastrophic disturbances and old-growth forests in ponderosa pine landscapes of the Black Hills. *Conservation Biology* 11: 1276–1288.

Simard AJ (1996) Fire severity, changing scales, and how things hang together. *International Journal of Wildland Fire* 1: 23–34.

Skinner CN and Chang C-R (1996) *Fire Regimes, Past and Present*. Davis, CA: University of California at Davis.

Smith MD (2011) The ecological role of climate extremes: Current understanding and future prospects. *Journal of Ecology* 99: 651–655.

Stainforth DA, Aina T, Christensen C, *et al.* (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* 433: 403–406.

Strauss D, Bednar L, and Mees R (1989) Do one percent of forest fires cause ninety-nine percent of the damage? *Forest Science* 35: 319–328.

Swanson FJ, Jones JA, Wallin DO, and Cissel JH (1994) Natural variability – implications for ecosystem management. Vol. II: Ecosystem management principles and applications. In: Jensen ME and Bourgeron PS (eds.) *Eastside Forest Ecosystem Health Assessment*, pp. 80–94. USDA Forest Service Pacific Northwest Research Station

Swetnam TW (1990) Fire history and climate in the southwestern United States. *Effects of fire Management of Southwestern Natural Resources: proceedings of the Symposium*. Fort Collins, CO, USA: USDA Forest Service, Rocky Mountain Forest and Range Experiment Station. p. 293.

Swetnam TW (1993) Fire history and climate change in giant sequoia groves. *Science* 262: 1993.

Swetnam TW, Allen CD, and Betancourt JL (1999) Applied historical ecology: Using the past to manage for the future. *Ecological Applications* 9: 1189–1206.

Swetnam TW and Betancourt JL (1998) Mesoscale disturbance and ecological response to decadal climatic variability in the American Southwest. *Journal of Climate* 11: 3128–3147.

Tang SM and Gustafson EJ (1997) Perception of scale in forest management planning: Challenges and implications. *Landscape and Urban Planning* 39: 1–9.

Thompson JR, Johnson KN, Lennette M, Spies TA, and Bettinger P (2006) Historical disturbance regimes as a reference for forest policy in a multiowner province: A simulation experiment. *Canadian Journal of Forest Research* 36: 401–417.

Tinker DB and Knight DH (2001) Temporal and spatial dynamics of coarse woody debris in harvested and unharvested lodgepole pine forests. *Ecological Modelling* 141: 125–149.

Tinker DB, Romme WH, and Despain DG (2003) Historic range of variability in landscape structure in subalpine forests of the greater Yellowstone area, USA. *Landscape Ecology* 18: 427–439.

Tomback D, Arno Stephen F, and Keane RE (2001) *Whitebark Pine Communities: Ecology And Restoration*. Washington DC, USA: Island Press.

Turner M (1987) *Landscape Heterogeneity and Disturbance*. NY: Springer Verlag.

Turner MG, Romme WH, Gardner RH, O'neill RV, and Kratz TK (1993) A revised concept of landscape equilibrium: Disturbance and stability on scaled landscapes. *Landscape Ecology* 8: 213–227.

van Wagner CE (1978) Age-class distribution and the forest fire cycle. *Canadian Journal of Forest Research* 8: 220–227.

Van Wagtendonk JW (1991) Spatial analysis of lightning strikes in Yosemite National Park. *Proceedings of the 11th Conference on Fire and Forest Meteorology: April 16-19, 1991, Missoula, Montana*. Bethesda Md: Society of American Foresters c1991.

Veblen TT (2003) Historic range of variability of mountain forest ecosystems: Concepts and applications. *The Forestry Chronicle* 79: 223–226.

Veblen TT and Donnegan J (2005) Historical Range of Variability for Forest Vegetation of the National Forests of the Colorado Front Range. *USDA Forest Service Rocky Mountain Research Station, Final Report USDA FOREST SERVICE AGREEMENT No. 1102-0001-99-033*. USA: Fort Collins, CO.

Watt AS (1947) Pattern and process in the plant community. *Journal of Ecology* 35: 1–22.

Whisenant SG (1990) Changing fire frequencies on Idaho's Snake River Plains: Ecological and management implications. In: Mcarther ED (ed.) *Proceedings on Cheatgrass Invasion, Shrub Dieoff, and Other Aspects of Shrub Biology*. USDA Forest ServiceIn: Mcarther ED (ed.) *Proceedings on Cheatgrass Invasion, Shrub Dieoff, and Other Aspects of Shrub Biology*. USDA Forest Service General Technical Report INT-276, 4-10.

Whitlock C and Millspaugh SH (1996) Testing the assumptions of fire-history studies: An examination of modern charcoal accumulation in Yellowstone National Park, USA. *The Holocene* 6: 7–15.

Wimberly MC, Spies TA, Long CJ, and Whitlock C (2000) Simulating historical variability in the amount of old forest in the Oregon Coast Range. *Conservation Biology* 14: 167–180.

Wong CM, Dorner B and Sandmann H (2003) Estimating historical variability of natural disturbances in British Columbia. Victoria, BC Canada: BC Ministry of Forests, Forest Science Program, Ministry of Sustainable Resource Management, Resource Planning Branch.

Wong CM and Iverson K (2004) Range of natural variability: Applying the concept to forest management in central British Columbia. *BC Journal of Ecosystems and Management Extension Note* 4: 1–56.

Wu J, Jones B, Li H, and Loucks OL (2006) *Scaling and Uncertainty Analysis in Ecology*. Dordrecht, the Netherlands: Springer.

Yarie J (1981) Forest fire cycles and life tables: A case study from interior Alaska. *Canadian Journal of Forest Research* 11: 554–562.

Diversity and Conservation of Neotropical Mammals

Ricardo A Ojeda, Institute for Aridland Research (IADIZA-CONICET), Mendoza, Argentina

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptive radiation The evolution of ecological and phenotypic diversity within a rapidly multiplying lineage.
Ecoregion A large unit of land or water containing a distinct assemblage of natural communities and species.
Endemic Any taxon (e.g., species, genus, family, order) confined to a particular region.
GABI The Great American Biotic Interchange was the major exchange of land mammals between South and North America after the formation of the Isthmus of Panama, about 2.7 million years ago.
Rapoport effect A pattern where the size of a species' geographic range increases toward higher latitudes.

Introduction

The characteristics and biogeography of Neotropical mammals have been the focus of copious literature. The interest in the Neotropical region lies in its long isolation during the Cenozoic Era (Age of Mammals), particularly in South America, and the extensive radiations that led to a remarkable fauna within a context of heterogeneous topography and extraordinary diversity of ecoregions. The foundations of the knowledge of Neotropical mammals were laid in the voyages of Christopher Columbus in 1492, naturalist expeditions up into the nineteenth century, as well as by pioneers of South American mammal paleontology (Hershkovitz, 1987).

Neotropical Land Mammals: History of Place and Lineages

Among the most prominent contributions to Neotropical mammal biogeography are the studies of Simpson (1950, 1980), Hershkovitz (1958, 1972), Cabrera and Yepes (1960), Patterson and Pascual (1972), Reig (1981), Webb and Marshall (1982) and Mares and Genoways (1982), among others.

The evolutionary history of Neotropical mammals (that is, the history of lineages) should be framed within the dynamic history of place (i.e., changes in geography, geology, climate, and environmental characteristics; Lomolino *et al.*, 2010). Thus, major paleogeographic events were the Division of Pangea into Gondwana and Laurasia, the breakup of Gondwana (into South America and Africa) during the upper Cretaceous, the isolation of South America as an "island continent" for more than 100 million years, throughout much of the Tertiary, the emergence of the Panamanian land bridge, and the rise of the Andean Cordillera.

Chronologically, the immigration of recent land mammals into the Neotropical Region can be divided into three distinct faunal strata (Simpson, 1980; Webb and Marshall, 1981) reflecting older and younger faunal contingents. These strata are the result of a fragmentary fossil record, and represent only a general heuristic framework synthesizing the history of the Neotropical mammal fauna **(Figure 1)**.

Prior to the first stratum, a protohorofauna (i.e., a Mesozoic Gondwanian fauna made up of mammal-like reptiles and true mammals from which part of the following fauna might have taken their origin; Reig, 1981) was present or had evolved in South America from Western Gondwanian or Laurasian groups; some Neotropical lineages of Stratum 1 may have originated from them (Reig, 1981). Stratum 1 is made up of the ancient autochthonous fauna of the early Cenozoic, and includes the living armadillos, giant anteater, marsupials, and various diverse and now-extinct groups of autochtonous ungulates. Stratum 2 represents the allocthonous lineages present in early Oligocene that includes the caviomorph rodents and primates (the origin of both groups are contentious, but they are probably African). The Stratum 3 corresponds to the young taxa, associated with the emergence of the Panamanian land bridge, and consequently, the large-scale experiment of faunal interchange known as the Great America Biotic Interchange (GABI), during late Pliocene, at about 3 million years ago. Among the mammal invaders from the North American stock are the procyonids (coati and racoon), vespertilionid and molossid bats, skunks, peccaries, cricetid rodents, mastodons, camelids, tapirs, bears, horses, deer, dogs, weasels, cats, leporids and squirrels (Marshall, 1988).

There is still debate regarding these phases, specifically the time and place of entry of some lineages. An example is the long-standing debate regarding the arrival and diversification of the cricetid field mice in South America, particularly the diverse subfamily sigmodontine with more than 300 species. Although there is a general consensus on North America as the place of origin of this lineage, there is controversy on whether their explosive diversification took place in North or South America. Some authors suggest that it occurred in Central America, and that the subfamily was already taxonomically diverse when they arrived in South America, whereas others suggest that the subfamily underwent a large scale adaptive radiation once in South America. The centers of diversity are now in South America, although many species are distributed into Central America and a few in the southern USA (Marshall, 1988; Steppan, 1996).

The Great American Interchange

The most impressive change in the composition of the South American land mammal fauna began in the late Miocene (Webb and Marshall, 1982). During this period, representation of genera from Stratum 1 (autochtonous ungulates and large herbivorous xenarthrans) declined from 70% to less than 20%.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00353-1

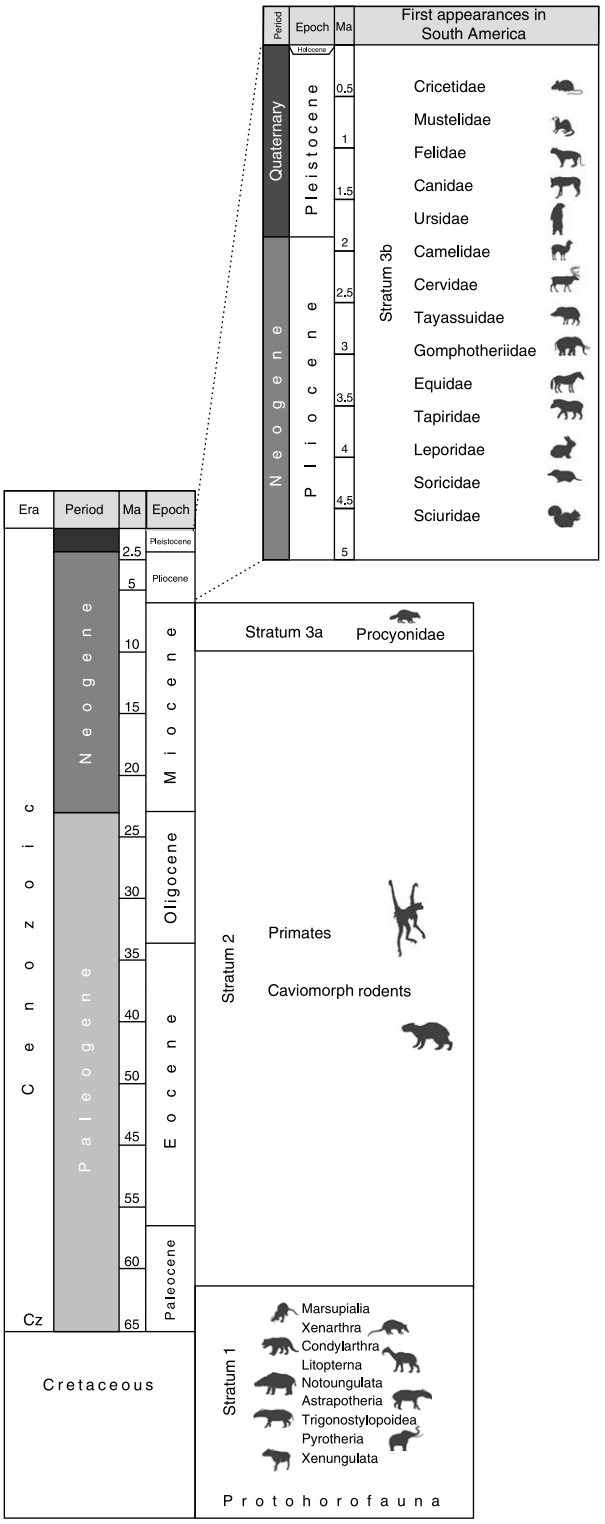


Figure 1 Cenozoic South American land mammal ages and faunal strata.

Concurrently, caviomorph rodents (making up much of Stratum 2) experienced remarkable adaptive radiation into the newly vacant herbivorous niches, ecologically equivalent to mammal groups in other continents (Mares and Ojeda, 1982). However, the most remarkable episode involved the increase in the diversity of Stratum 3 fauna from 0% to more than 50% of the total genera. At least half of the recent land mammal genera of South America belong to this stratum.

The geological events that took place about 3 million years ago (e.g., tectonic changes and a decrease in sea level) gave rise to the Isthmus of Panama, triggering the GABI. The arrival of the species from Stratum 3 to South America, and a reciprocal faunal interchange with North America, is partly linked to this event. However, some of the mammals of Stratum 3 entered South America before the emergence of the land bridge. During the late Miocene and early Pliocene, island archipelagos may have facilitated waif dispersal (i.e., dispersal across water barriers by island-hopping) across what is now Central America. At the beginning of the Cenozoic, Central America was probably the source area of ancient lineages of marsupials, xenarthrans, and ungulates that dispersed to South America as waif immigrants across the water barrier (Patterson and Pascual, 1972). A successful group that probably arrived as island hoppers in the late Miocene are the sigmodontine rodents (Hershkovitz, 1972; Reig, 1981). At the same time (7–9 million years ago), South American ground sloths entered North America. A summary of the reciprocal land mammal interchange between North and South American is portrayed in **Figure 2**.

A mass extinction, particularly megaherbivores, affected the Neotropical faunal assemblages at the end of the Pleistocene (Martin, 1967; Marshall, 1988). This coincided with the entrance of humans from Asia into North America across the Beringian land bridge at the end of the last glaciation and continued to South America. Earliest human settlements in South America may date back 14,220 years (Dillehay *et al.*, 2008). In the case of North America, there were about 34 extinct genera, whereas in South America the loss of megafauna (i.e., more than 44 kg) reached 52 genera including mastodons, ground sloths, glyptodonts, and the diverse group of native ungulates (Barnosky and Lindsey, 2010).

The megafaunal extinction during the Quaternary period (Late Pleistocene–Holocene) is hypothesized to be the result of synergistic interactions of environmental modifications and hunting, associated with climate change and the dispersal of early humans into the Americas. Although there has been improvement in the dating of extinctions and paleoclimatic reconstruction, the association of human artifacts and megafaunal extinctions are still a matter of active debate (Barnosky and Lindsey, 2010).

The Zoogeographic Region and their Mammals

Biogeographical Limits and Transition Zones

Hershkovitz (1958) defined a zoogeographic region as

> A zoogeographic area, whether Realm, Region, or subdivision thereof, regarded as a faunal district of the earth at a stated time in geologic history. The fauna of a zoogeographic area includes all animals inhabiting that area at the same time irrespective of their place of origin. A zoogeographic boundary between faunal areas is the barrier that prevents a natural and continuous faunal flow from one area to the other. A zoogeographic transition zone is an area on either or both sides of the arbitrarily drawn boundary line between contiguous and contemporaneous faunal areas. Its entire fauna is derived from neighboring areas (1958, p. 584).

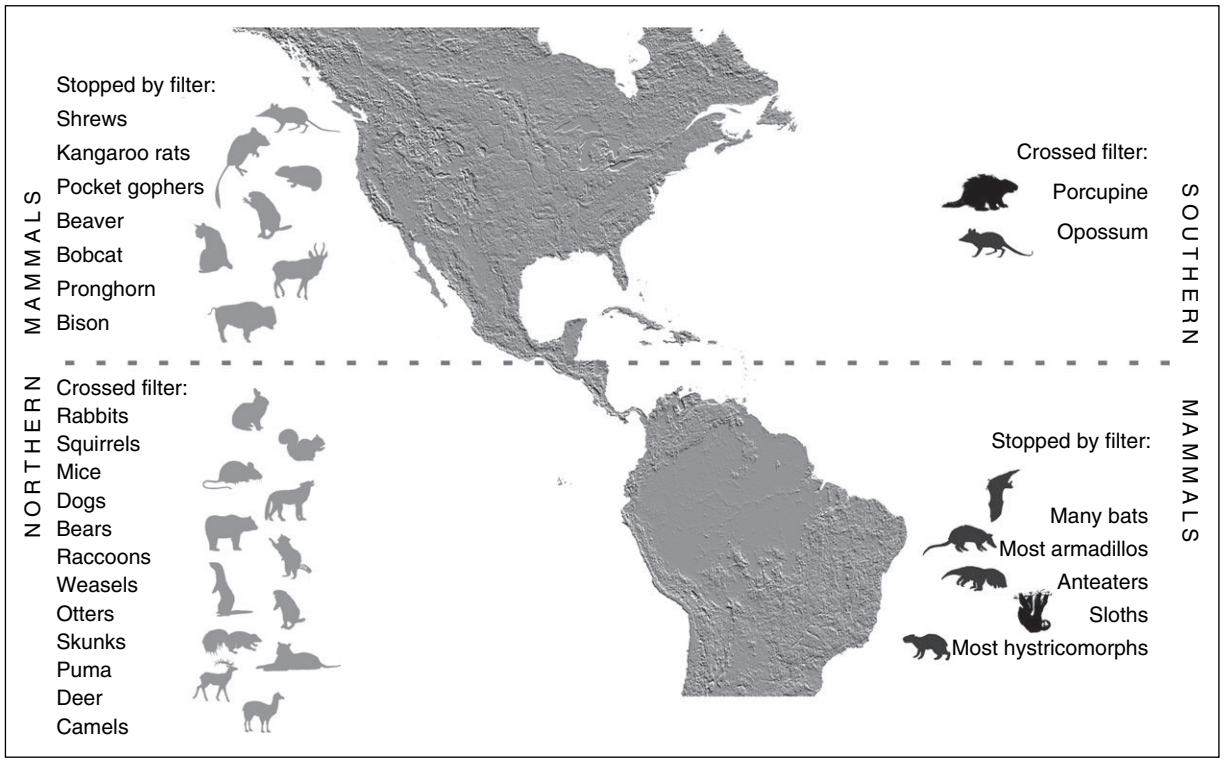


Figure 2 The Great American Biotic Interchange and the Central America filter effect. The expansion and retreat of glaciers affected the ecosystem dynamics in Middle America and acted as a filter for the dispersal of certain mammal groups. During the advance of glaciers there was an expansion of the savanna habitats which favored reciprocal dispersal of certain types of fauna to pass through the Isthmus of Panama, whereas during glacier retreats there was an advance of tropical habitats, a contraction of the savanna corridor and consequently a restriction to the dispersal of the savanna faunal groups.

The Neotropical Region was defined by Sclater and Wallace as one of the six major zoogeographical divisions of earth based on their vertebrate assemblages (Sclater, 1858; Wallace, 1876). The region extends from Central America to the southern tip of South America, including continental and oceanic islands, the Bahamas, West Indies, Galapagos, and Malvinas (Falklands). It is subdivided into hierarchically lower divisions known as the West Indian, Brazilian, and Patagonian Subregions.

The delimitation between the Neartic and the West Indian Subregion is an imaginary line across the Gulf of Mexico and the strait of Florida. The West Indian Subregion consists of a chain of islands including the Bahamas, Cuba, and the Grenadines. The Brazilian Subregion extends from Central America and northern South America up to northeastern Argentina. Major topographic features include the Andes extending along the western portion of South America and the Guianian and Brazilian highlands in the northeast and southeast, respectively. Included in its diverse vegetation are rainforests, savannas, mountain meadows, deciduous and evergreen broad-leave forests, and coastal mangroves. The Patagonian Subregion includes Uruguay, most of Argentina (including Malvinas islands), Chile, the highlands of Bolivia and Peru, and the paramos of Ecuador up to the equator. Major topographic and vegetation features include the Pacific coastal desert, the Andes, the vast Patagonian steppe desert, *Nothofagus* beech forests, arid, and semiarid scrublands, savanna and open thornscrub forests, and grassland habitats (Hershkovitz, 1974).

There have been various different biogeographical approaches and methodologies to set the boundaries between the Neartic and Neotropical zoogeographic regions. Some of these proposed limits based on the southern and northern geographic range edges of mammals are synthesized in the inset of **Figure 3**. Disagreement among these regions are due to different methodologies and taxa analyzed. Thus, the boundary between the Neartic and Neotropical Regions based on the distributional patterns for different taxonomic levels (i.e., families, genera, and species) shows that the southern boundaries of Neartic taxa are located in northern Mexico up to the Mexican Plateau (Escalante *et al.*, 2010). Another quantitative analysis of the distribution of bats in Mexico and Central America (i.e., cells of presence/absence occupancy) found a similar pattern for the southern limits of the Neartic region, with minor differences regarding the extent of the transitional zone (Ortega and Arita, 1998).

The boundaries between biogeographic divisions (referred to as transition zones especially on land), are difficult to determine and are unrealistically portrayed as sharp lines. Instead, these boundaries represent dynamic areas of biotic overlap, active belts of intense faunistic interactions and large-scale population dynamics (Ortega and Arita, 1998; Escalante *et al.*, 2010; Morrone, 2009; Lomolino *et al.*, 2010).

The transition zone between the Brazilian and Patagonian Subregions runs diagonally between southern Ecuador, Peru, Bolivia, and southeastern Brazil, and into Argentina. In

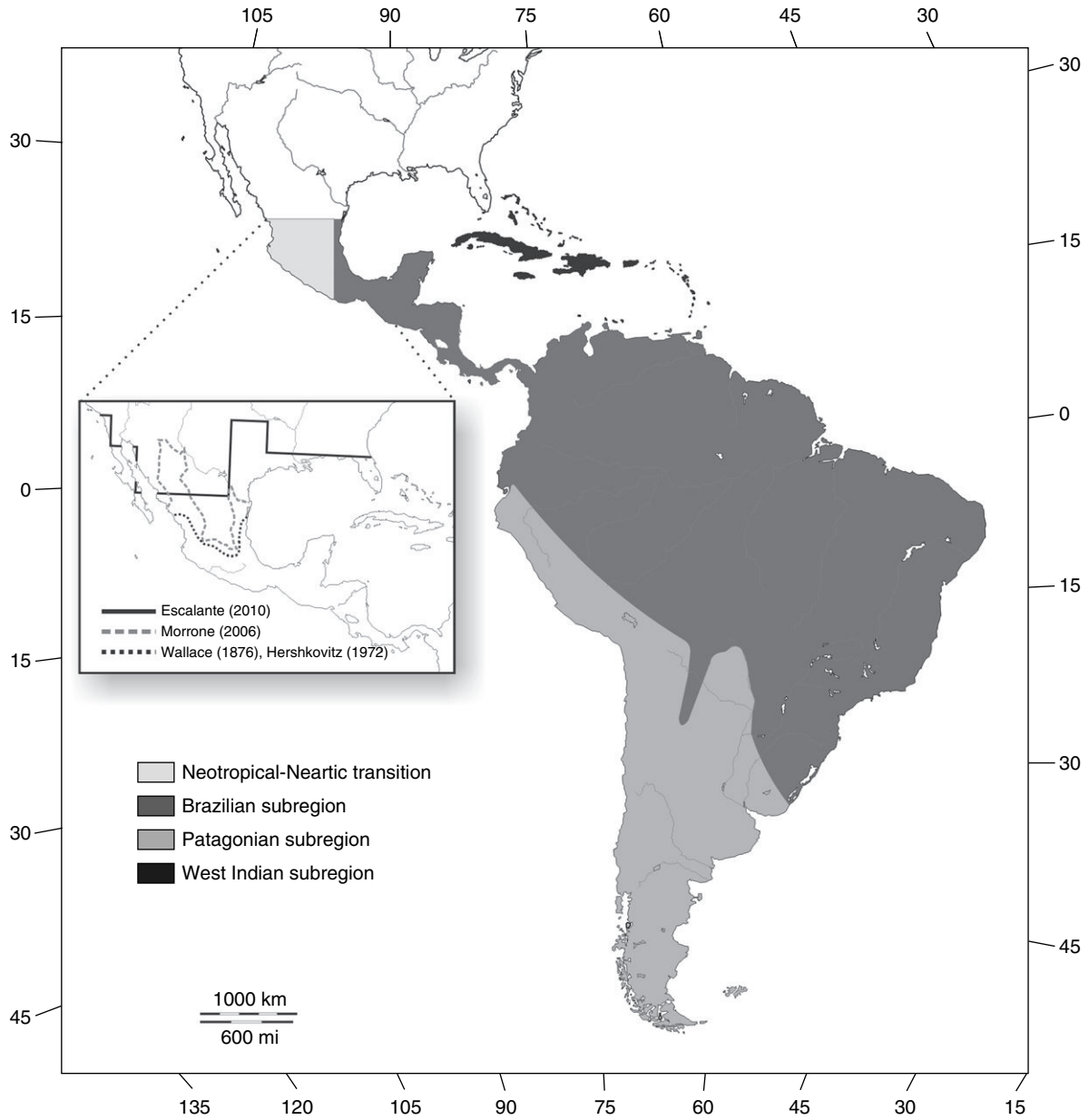


Figure 3 The Neotropical Region and subregions (adapted from Hershkovitz, 1972). The insert map depicts several delimitations of the Neartic–Neotropical transition zone that were proposed by different authors.

northwestern Argentina, the Brazilian Subregion extends like a peninsula (see **Figure 3**) including the southernmost montane cloud forests which stretch to northern South America along the eastern flank of the Andes (Tropical Andes hotspot). Mammal distributions in this transition zone result in a broader transitional area which is roughly congruous with previous boundaries established by Sclater and Wallace as well as the "subtropical line" of Ruggiero *et al.* (1998).

Diversity and Classification of Mammal Groups

The geological history and biotic diversity of the Neotropical Region are the focus of active research, with active discovery of new taxa at various levels. Extant Neotropical mammals represent one quarter of the worlds mammal fauna. Eight of the 20 richest countries in mammals are in the Neotropical Region (Brazil, Mexico, Peru, Colombia, Argentina, Ecuador, Bolivia, and Venezuela, in descending order). However, tabulations of mammal species are imprecise, with new species (particularly marsupials, insectivores, primates, bats, and rodents) being discovered and described rapidly (Patterson, 1994, 2000). New mammal discoveries are particularly significant in South America, especially in areas of high endemism such as tropical biomes (e.g., Central America, eastern tropical Andes, the Amazon basin, and Atlantic forests), and temperate semiarid habitats.

On average, 223 new mammal species are described every 10 years. Of the 62% newly-described New World mammal species in recent years, 93% were from the Neotropical region,

mostly from the tropical and semitropical regions of Mexico and Brazil, the Andes of Colombia, Peru and Ecuador, as well as the temperate areas of Chile and Argentina (Ceballos and Ehrlich, 2009; Reeder *et al.*, 2007).

Of the about 1550 mammal species recently estimated for the Neotropical Region (Solari *et al.*, 2012), 1421 species are from land habitats, and 946 of these (66%) are endemic.

The Neotropical Region has the highest diversity of mammals in the New World, with 15 orders of terrestrial and marine mammals three of which are endemic. The marsupials belonging to the orders Microbiotheria (e.g., monito del monte) and Paucituberculata (i.e., rat and shrew opossums) are mostly restricted to the Patagonian Subregion, whereas the order Pilosa (giant and lesser anteater, and two and three-toed sloths) is restricted to the Brazilian Subregion.

With respect to families, the Neotropical Region has the largest number of endemic families of terrestrial mammals, 29/56. The endemicity ratio (endemic families/number of families) of other biogeographic regions is: Paleartic: 1/36; Neartic: 2/31; Afrotropic: 20/58; Indo-Malayan: 8/49, and Australasian: 21/35 (Vaughan *et al.*, 2011).

A General Account of the Neotropical Mammal Fauna

This section provides an overview of the Neotropical mammals. **Figure 4** depicts some representatives of these orders and endemic families.

Didelphimorphia – This is an ancient lineage of small to medium-sized marsupials in the Neotropical Region. They occupy a wide diversity of habitats and macroniches (arboreal, scansorial, and semiaquatic; insectivores, frugivores, omnivores, and carnivores). Some genera are *Didelphis*, *Caluromys*, *Glironia*, *Lutreolina*, *Phylander*, *Thylamys*, and *Monodelphis*. All species are endemic to the Neotropical Region, except for the common opossum, *Didelphis virginiana*, which extends its geographic range into the Neartic Region up to Canada.

Paucituberculata – This endemic order of small insectivorous marsupials is confined to South America and is represented by a single family containing the genus *Caenolestes*, *Lestoros*, and *Rynchолestes*. Their distribution extends across Ecuador, Peru, and Bolivia (*Caenolestes*, *Lestoros*), and southern Chile (*Rhyncholestes*).

Microbiotheria – This is a monospecific order with a single family and species, endemic to South America, the small marsupial *Dromiciops australis*. (“monito del monte”). It feeds on larvae, insects, and fruits, and is restricted to the Valdivian temperate rainforest of *Nothofagus* in Argentina and Chile Phylogenetically. it is more closely related to Australian marsupials than to South American ones.

Sirenia – This group is represented by the family Trichechidae, and composed of three species which are endemic to the Amazon and Orinocco basin of the Brazilian Subregion. They are strictly aquatic and herbivorous, and include the manatee (*Trichechus*) and Amazon river dolphins (*Sotalia*, *Inia*).

Cingulata – This order, along with Pilosa and the marsupials, are among the emblematic “old inhabitants” of the Neotropical region. The order Cingulata includes a diverse suite of armadillos with varying body sizes, ranging from the pink fairy armadillo (100 g) to the giant armadillo (30 kg). Representative genera are: *Dasypus*, *Euphractus*, *Cabassous*, *Priodontes*, *Tolypeutes*, *Chaetophractus*, *Chlamyphorus*, *Calyptophractus* and *Zaedyus*. Armadillos do not occur in the Neartic Region (North America) with the single exception of *Dasypus novemcinctus*, which extends into southeastern United States. They are omnivores that feed on insects, larvae, eggs and fruits.

Pilosa (=Edentata) – This order includes the tree sloths of the genus *Bradypus* and *Choloepus*, and the anteaters (*Tamandua*, *Myrmecophaga*, and *Cyclopes)*. Tree sloths are arboreal folivores, feeding on leaves and buds; they are mostly tropical in their distribution, and occur from Central America to near the Tropic of Capricorn in northern Argentina. Anteaters are terrestrial and semiarboreal and feed on ants and termites.

Primates – Primates are represented by two endemic and one cosmopolitan (Hominidae) families which includes humans, *Homo sapiens*. Neotropical monkeys are mostly tropical forest species and are small to medium size. Among them, the pigmy marmoset, *Cebuella pygmaea* with a body weight of 70 g, is the smallest of the living primates. Neotropical monkeys are arboreal and with a diet of insects, leaves, and fruits. Whether they originated in Africa or Asia is still a matter of debate. Representative genera, many of them highly endangered, include *Cebuella*, *Saguinus*, *Callithrix*, *Leontopithecus*, *Aotus*, *Saimiri*, *Ateles*, *Alouatta*, *Cebus*, among others.

Rodentia – Rodents are represented by four major groups, or suborders: Sciuromorpha, Castorimorpha, Hystricomorpha, and Myomorpha. Of these, the Castorimorpha and Sciuromorpha are of northern origin, and they reach their southern distributional limits in distribution in northern South America and near the Tropic of Capricorn, respectively. Representative genera of Sciuromorpha are squirrels, *Sciurus*, whereas the Castorimorpha is represented by subterranean pocket gophers (*Geomys*, *Orthogeomys*) and spiny pocket mice (*Liomys*, *Heteromys*). The other two suborders, Hystricomorpha and Myomorpha, include the endemic and highly diversified caviomorphs (the South American hystricomorphs) and sigmodontine rodents, respectively, which underwent explosive adaptive radiation in the Neotropics, mainly in South America. These two groups account for more than 550 species, ranging from 12 g to 50 kg (e.g., the largest living rodent, capybara), and occupy a full array of habitats and modes of life. Caviomorph rodents comprise 11 endemic families and some genera are *Galea*, *Dolichotis*, *Abrocoma*, *Ctenomys*, *Octodon*, *Tympanoctomys*, *Spalacopus*, *Proechimys*, *Hydrochaeris*, *Agouti*, *Dasyprocta*, *Cuniculus*, and *Coendu*, among the most conspicuous. Representative genera of sigmodontine rodents are: *Isthmomys*, *Icthyomys*, *Oryzomys*, *Thomasomys*, *Wiedomys*, *Calomys*, *Sigmodon*, *Akodon*, *Eligmodontia*, *Phyllotis*, *Salinomys*, *Euneomys*, and *Abrothrix*.

Lagomorpha – This order is represented by the family Leporidae (rabbits and hares) and includes the North American cottontail rabbits. In the Neotropical Region it is represented by the genus *Sylvilagus*. Some of these herbivorous mammals have distributions from Central America to northern Argentina and southern Brazil. Other members of this family are represented by the European rabbit (*Oryctolagus)* and European hare (*Lepus*), introduced by humans in the Neotropical region.

Soricomorpha – This order includes the insectivorous shrews which are distributed across Central America, the West Indian Subregion, and barely into the subtropical Andes in northern South America (genus *Cryptotis*). The endemic family

Solenodontidae includes two endangered species of the genus *Solenodon* distributed in Cuba and the Dominican Republic.

Chiroptera – Bats, the second most speciose mammals in the Neotropical Region. Six out of nine families are endemic here; their feeding niches are highly diversified and include sanguinivores, frugivores, insectivores, carnivores, nectarivores, pollinivores, and piscivores. Representative genera of different families are *Noctilio*, *Glossophaga*, *Brachyphylla*, *Carollia*, *Tonatia*, *Desmodus*, *Vampyrops*, *Artibeus*, *Mormoops*, *Peropteryx*, *Thyroptera*, *Natalus*, *Molossus*, *Lasiurus*, and *Eumops*.

Carnivora – This is a diversified group of mammals occupying terrestrial, arboreal, aquatic and marine habitats. They are represented by eight families which are widely distributed in the Neotropical Region and include the large and medium-size spotted cats, weasels, skunks, river otters, elephant seal, sea lions, foxes, raccoons, coatis, and the spectacled bear. Representative genera are *Panthera*, *Felis*, *Galictis*, *Conepatus*, *Lutra*, *Lontra*, *Nasua*, *Potos*, *Tremarctos*, *Hydrurga*, *Mirounga*,*Arctocephalus*, and *Otaria*.

Perissodactyla – This ungulate group possess an enlarged middle digit on the anterior and posterior limbs, hence the name of odd-toed ungulates. The order is represented by three species of tapirs in the genus *Tapirus*, among the largest terrestrial mammals in the Neotropical Region. They occur from Central Mexico to northern Argentina and feed on grasses and fruits. They are mostly distributed in tropical and gallery forests, forested foothills, tropical Andean paramos, and riverine environments.

Artiodactyla – These are the even-toed ungulates. In the Neotropical Region this order is represented by three families. Representative genera are *Tayassu*, *Lama*, *Vicugna*, *Pudu*, *Mazama* and *Ozotoceros*, and include the peccaries, guanacos, vicuñas, pampas deer and brocket deer, and pygmy pudu deer. Their diet is mostly herbivorous (deer and camels) and omnivorous (peccaries) and they are distributed from southern North America (collared peccaries; genus *Tayassu*) to the southern tip of South America (guanacos).

Cetacea – A group of strictly marine mammals that is distributed worldwide and represented by whales, dolphins, and porpoises; they feed on fish and small crustaceans. Representative genera are *Pontoporia*, *Delphinus*, *Lagenorhynchus*, *Orcinus*, *Tursiops*, *Phocoena*, *Mesoplodon*, *Physeter*, *Balaenoptera Megaptera*, and *Balaena*.

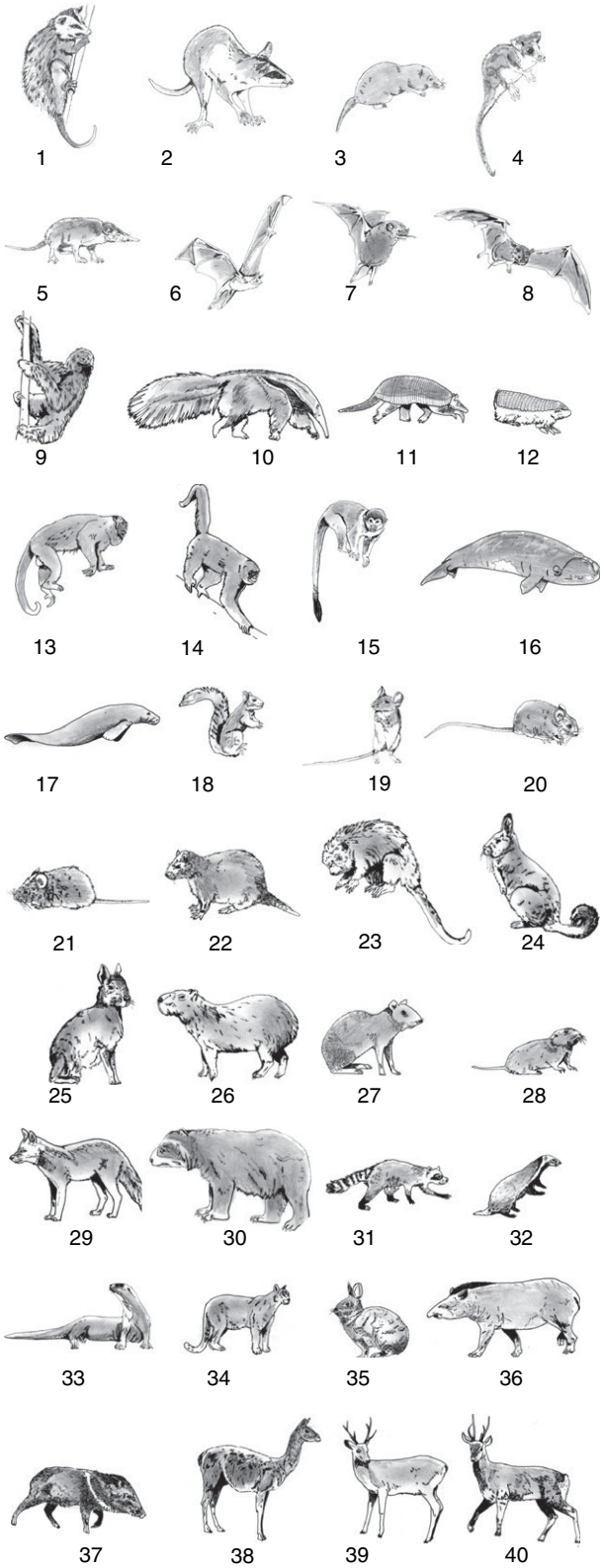


Figure 4 Representative orders and species of Neotropical mammals. Didelphimorphia: (1) *Didelphis albiventris*; (2) *Metachirus nudicaudatus*; Paucituberculata: (3) *Rhyncholestes raphanurus*; Microbiotheria: (4) *Dromiciops australis*; Soricomorpha: (5) *Solenodon cubanus*; Chiroptera: (6) *Noctilio leporinus*; (7) *Glossophaga soricina*; (8) *Eumops auripendulus*; Pilosa: (9) *Bradypus torquatus*; Cingulata: (10) *Myrmecophaga tridactyla*; (11) *Priodontes maximus*; (12) *Chlamyphorus truncatus*; Primates: (13) *Alouatta palliate*; (14) *Lagothrix flavicauda*; (15) *Saimiri sciureus*; Cetacea: (16) *Eubalaena australis*; Sirenia: (17) *Trichechus manatus*; Rodentia: (18) *Sciurus ignitus*; (19) *Eligmodontia typus*; (20) *Oryzomys longicaudatus*; (21) *Akodon molinae*; (22) *Capromys pilorides*; (23) *Sphigurus mexicanus*; (24) *Lagidium viscacia*; (25) *Dolichotis patagonum*; (26) *Hydrochoerus hydrochaeris*; (27) *Dasyprocta punctata*; (28) *Ctenomys talarum*; Carnivora: (29) *Dusicyon culpaeus*; (30) *Tremarctos ornatus*; (31) *Procyon cancrivorus*; (32) *Lyncodon patagonicus*; (33) *Pteronura brasiliensis*; (34) *Puma concolor*; Lagomorpha: (35) Sylvilagus brasiliensis; Perissodactyla: (36) *Tapirus terrestris*; Artiodactyla: (37) *Tayassu tajacu*; (38) *Lama guanicoe*; (39) *Ozotoceros bezoarticus*; (40) *Hippocamelus antisensis*.

Biogeographical–Macroecological Patterns

Patterns of Species Richness

The distribution and diversity of Neotropical mammals have been analyzed for many groups utilizing different biogeographic approaches, from the continental scale (e.g., Mesoamerica and South America) to regional biomes and countries. The inverse relationship between richness of Neotropical land mammals and latitude is similar to that found for many other taxa of plants and animals. One of the first to quantify and explain this large-scale pattern of decreasing species richness from low equatorial latitudes to high temperate latitudes for the mammals of the Neartic region (North America) was Simpson (1964). This pattern has been analyzed and corroborated for the New World mammals and several taxa such as marsupials, bats, caviomorph rodents and carnivores, among others. The highest concentration of terrestrial (nonvolant) mammals in South America, is found along the eastern Andes of Bolivia, Peru and Ecuador, the Amazon Basin, and the Atlantic Rainforests of Brazil (Mares and Ojeda, 1982; Kaufmann, 1995; Ruggiero, 1994; Tognelli and Kelt, 2004 (**Figures 5** and **6**).

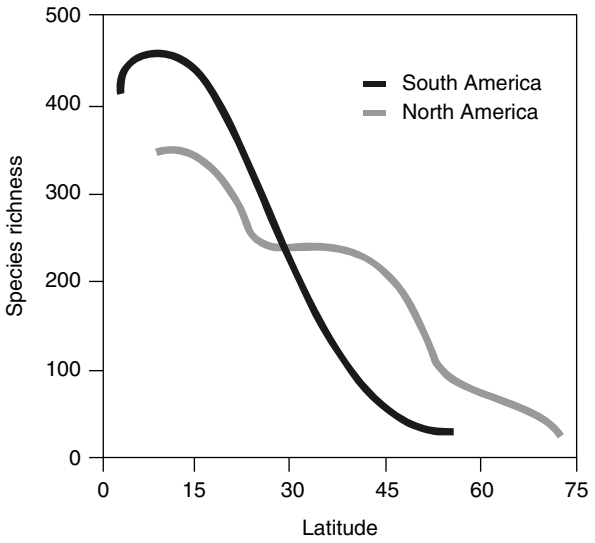


Figure 5 Patterns of latitudinal diversity in North and South American mammals.

Although researchers agree on the generality of this latitudinal trend, there is considerable debate over its causes (i.e., historical, ecological, climatic stochastic and integrative hypotheses, Chown *et al.*, 2004; Lomolino *et al.*, 2010). For example, the large area of tropical South America has been suggested to explain diversity of biotas at both continental and island scales (Mares and Ojeda, 1982; Rosenzweig, 1995; Rosenzweig and Sandlin, 1997). However, the latitudinal diversity gradient for New World marsupials and bats could be explained by a null model (i.e., the pattern in species richness could be a consequence of stochastic mechanisms independent of any particular factor or environmental gradient), up to some extent (Willig and Lyons, 1998). However, further analysis of the joint effects of latitude and area has shown that latitude is the dominant factor which explains greater than 50% of the variation in mammal species richness. This relationship is independent of area, which contradicts the explanations involving area as one of the major drivers in species diversity (Kaufmann, 1995; Lyons and Willig, 1999).

Regardless of this debate, global trends of latitudinal diversity in mammals are driven by a combination of factors, including energy availability, topographic complexity, and area, among others. As a result, land mammal species richness peaks toward the equator, in conjunction with primary productivity, whereas richness of marine mammals peaks at approximately 40° N and S, corresponding to belts of high oceanic productivity (Schipper *et al.*, 2008). In the particular case of South America, energy input converted into larger primary productivity and then translated into greater resources available to consumers, explains a large part of the latitudinal diversity gradient (Tognelli and Kelt, 2004). Other variables such as habitat heterogeneity, topographic variability, climate, and number of ecosystems, are less important in determining mammalian species richness at this continental scale of analysis, although they are relevant in explaining mammal diversity at regional and local scales. Furthermore, despite corroboration of this relationship, some of the reported differences between results on large scale latitudinal gradients are likely due to the utilization of different databases, methodology, scale of analyses, number and type of taxa involved, and statistical models (Kaufmann and Willig, 1998; Tognelli and Kelt, 2004).

In general, ongoing research regarding the explanation of latitudinal gradients in species diversity is shifting from the need to find a unique driving factor to a more integrative approach of hypotheses and theories combining several factors (Lomolino *et al.*, 2010).

Detailed reviews and summaries of current knowledge and compelling hypotheses for a latitudinal gradient in species richness have been provided by Rosenzweig (1995), and Lomolino *et al.* (2010).

Areographic Patterns and Ecogeographic Rules

Sizes and shapes of geographic ranges are influenced by the topography of the region. Thus, South American mammal geographic ranges become more asymmetrical (relative to a circle) toward the west as a result of the topographical complexity of the Andes. The most widespread species are located in eastern Brazil, whereas the most restricted species are associated with the Andes. The major contrast in South American mammal range sizes therefore occurs in an east – west direction, along the boundaries between the Guayana-Brazilian and Andean-Patagonian Subregions, in accordance with the classical zoogeographical delimitations (Ruggiero *et al.*, 1998).

The increase in body sizes from the equator toward the Poles (toward high, colder latitudes) is a generalized pattern known as Bergmann's Rule (Lomolino *et al.*, 2010). However, there is marked contrast between patterns among the mammal faunas of the Neartic (North America) versus the Neotropical Region. Although in the Neartic Region body size is negatively correlated with temperature, increasing toward high latitudes, in the Neotropical Region there is a positive correlation with temperature such that mammals are larger in warmer tropical latitudes and smaller in the mountainous Andean regions.

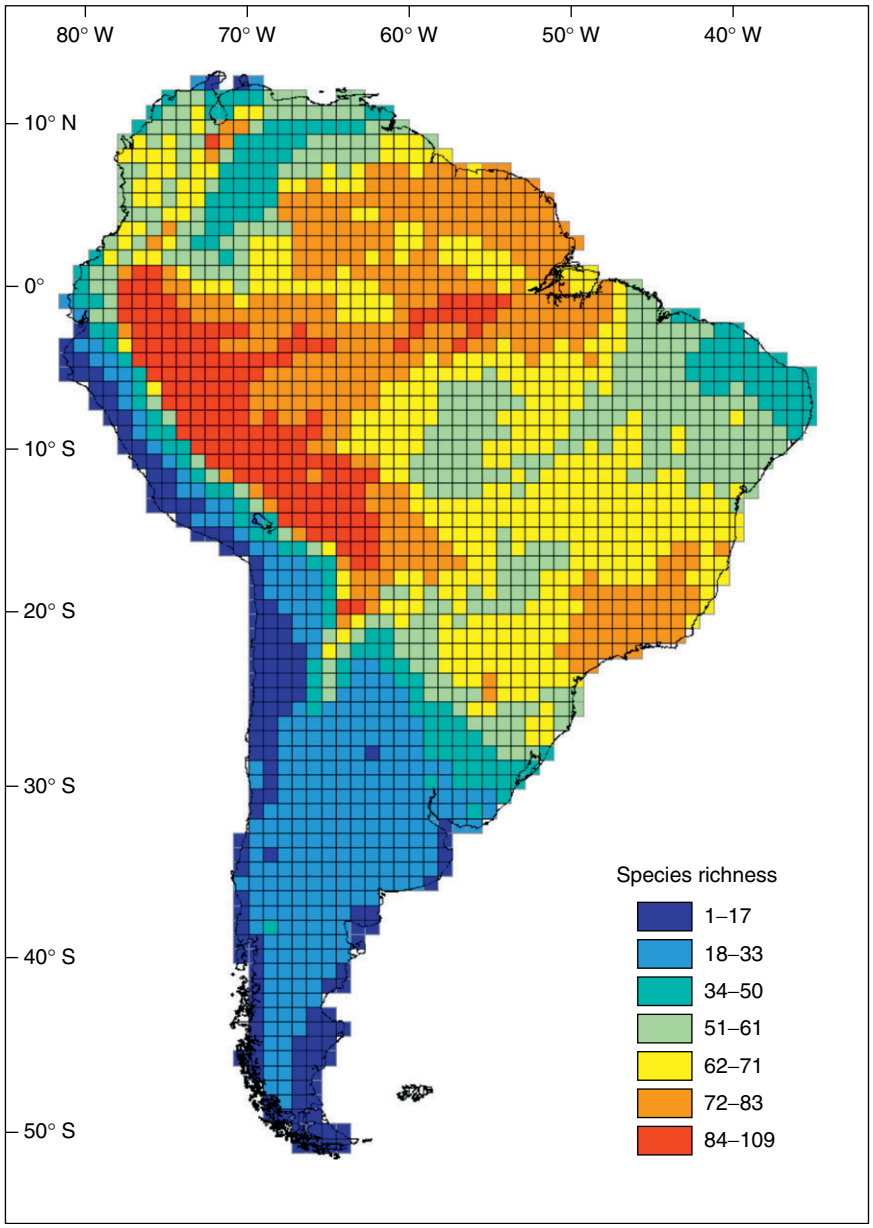


Figure 6 Patterns of diversity of South American nonvolant mammals.

A suggested explanation for this is that reduced habitat in mountains produced an area effect resulting in fewer large-sized mammals (Rodriguez *et al.*, 2008; Ruggiero, 1994).

Another ecogeographic rule known as the Rapoport Rule (Stevens, 1989), states that the latitudinal ranges of species tend to increase toward higher latitudes. An explanation for this trend is that greater climate seasonality and variability at high latitudes exerts a selective effect promoting greater flexibility and larger geographic ranges of species. In South America, Rapoport's rule is consistent with observed patterns only for carnivores and primates; support is less clear for caviomorph rodents, marsupials, and artiodactyls. Anteaters, tree sloths, and armadillos show the opposite trend, with their geographic ranges decreasing toward higher latitudes Ruggiero (1994). Overall, contrary to the Neartic Region, there is a lack of evidence regarding increases in geographic ranges toward higher latitudes in South America. This could be associated with the fact that there is less temperature variation in the southern hemisphere, compared to the northern hemisphere (Chown *et al.*, 2004).

Mammal Conservation and Major Threats

As a result of the Neotropic's high mammal diversity and endemicity, their conservation is a key issue in the region. Roughly, of the 76 world mammal species that have become extinct in the past 500 years (http://roboconsumer.wordpress.com/2007/09/23/extinct-mammals/) 34 (45%) occurred in the Neotropical Region.

One of the first extinct canids was the Malvinas (or Falkland Islands) fox (*Dusycyon australis*), an event already predicted by Darwin (1860) during the voyage of the Beagle

> "…As far as I am aware, there is no other instance in any part of the world, of so small a mass of broken land, distant from a continent, possessing so large an aboriginal quadruped peculiar to itself. Their numbers have rapidly decreased; they are already banished from that half of the island which lies to the eastward of the neck of land between St. Salvador Bay and Berkeley Sound. Within a very few years after these islands shall have become regularly settled, in all probability this fox will be classed with the dodo, as an animal which has perished from the face of the earth." (Charles Darwin, The Voyage of the Beagle, 1860, entry for March 16th, 1834)

Most of these extinctions were the result of a combination of factors, including ecological attributes (i.e., habitat or diet specialists), coupled with restricted insular distributions (e.g., Haiti, Dominican Republic, Cuba, Puerto Rico, and Malvinas), and anthropogenic factors affecting mammals worldwide, such as overhunting, habitat destruction, and the introduction of alien species (e.g., rats, cats, and dogs) (Schipper *et al.*, 2008). In the long term, global climate change is becoming a matter of increasing concern (MacLean and Wilson, 2011).

About 52% of mammals worldwide whose population trends are known, appear to be contracting their geographic ranges and declining their populations (Schipper *et al.*, 2008). At the scale of countries, a detailed overview of the diversity and major conservation problems facing the mammal fauna of the Neotropical Region, such as deforestation, hunting, wildlife trade, and introduction of exotics, was provided by Ceballos and Simonetti (2002). Habitat loss and fragmentation from extensive deforestation is common throughout the Neotropics. In South America, most losses of mammals are associated with cattle grazing and intensive agricultural activities, particularly in the Pampas region (Argentina), Mata Atlantica (Brazil), and coastal areas of Ecuador and Peru. These disturbances are particularly relevant in mammals of temperate South America because of its characteristic high number of endemisms and small geographic ranges (Ceballos and Ehrlich, 2002; Schipper *et al.*, 2008; Lamoreux and Lacher, 2010).

Subsistence hunting, and particularly faunal commercialization, are important threats to mammal populations in the Neotropical Region (Ojeda and Mares, 1982; Iriarte and Jaksic, 1986; Robinson and Redford, 1991) **(Table 1)**. An example of the commercial scale of mammal exploitation is provided by official statistics from the trade of wildlife products (e.g., skins, hides, fur, and meat) from Argentina which represented more than 30 million dollars during 1976–1979 (Ojeda and Mares, 1982). Between 1972 and 1979, Argentina exported 22 million mammals. Despite international agreements on the conservation of marine mammals, the dominant threats are harvesting and accidental mortality, (Schipper *et al.*, 2008).

Biological invasions are another threat to Neotropical biodiversity. An appropriate comment regarding exotic mammals was again made by Charles Darwin (1860) during his crossing of the Pampas and Patagonian landscapes

> "few countries [than Argentina] have undergone more remarkable changes since the year 1535, when the first colonist of La Plata landed with seventy-two horses. The countless herds of horses, cattle, and sheep, have altered the whole aspect of the vegetation, but they have almost banished the guanaco, deer, and ostrich. Numberless other changes must likewise have taken place; the wild pig in some parts probably replaces the pecari; packs of wild dogs may be heard howling on the wooded banks of the less frequented stream; and the common cat, altered into a large and fierce animal, inhabits rocky hills…". (Charles Darwin, The Voyage of the Beagle, 1860, entry for September 19th, 1833)

Table 1 Mammal harvesting in the Neotropical Region

Subsistence hunting[a]	*Commercialization*[b]	
Species	*Species*	*1972–1978*
Common Opossum	Nutria	11,011,288
Nine – banded armadillo	Gray Fox	5,789,011
Giant Armadillo	Opposums	1,541,717
Giant Anteater	Skunks	1,243,129
Lesser Anteater	Plains Vizcacha	820,177
Three-toed sloth	Geoffroy' s Cat	481,333
Two-toed Sloth	Guanaco	443,655
Night Monkeys	Peccaries	312,115
Titi Monkeys	Red Fox	101,251
Capuchin Monkeys	Capybara	89,656
Howler Monkeys	Pampa's Cat	82,195
Spider Monkeys	Puma	2571
Brazilian Cottontail	Total	21,918,098
Squirrels		
Agouti		
Capybara		
Common Tapir		
White-lipped Peccary		
Collared Peccary		
Brocket Deer		
Coatí		
Kinkajou		

[a]Most common mammals for subsistence hunters (adapted from Redford and Robinson, 1991).

[b]Mammal species and number of individuals (skins) exported legally from Argentina during 1972–1978 (data from Ojeda and Mares, 1982).

The most impacted by invasions are the West Indian and Patagonian Subregions. Exotic mammals in South America represent about 20% of world mammal introductions, but the highest density of exotic mammals is found in the temperate Patagonian Subregion. Most introductions occurred in the eighteenth and nineteenth centuries, with more successful establishment in temperate ecosystems, between 34° and 55° S. They were facilitated by accidental introductions from ships of early explorers, or deliberate introductions in order to establish sport hunting or provide food and fur. Exotic Neotropical mammals include a variety of taxa (e.g., rodents, deer, wild boar, and mustelids), ecological groups (terrestrial and semiaquatic herbivores and carnivores, omnivores). Among the traits that were suggested for the success of invasive species are a high reproductive rate, large size of native geographic range, vacant niches, climatic matching, and so on (Novillo and Ojeda, 2009). Most species are of Eurasian origin and occupy similar ecoregions as in their native areas, although some have expanded their range of habitats. In Cabo de Hornos (55° S), for example, exotic species richness is higher

than that of natives. However, the understanding of the impacts of invasive mammals in the temperate region remains largely anecdotal for many of them (Jaksic *et al.*, 2002; Novillo and Ojeda, 2009).

Approaches and Strategies for Conservation in a Heterogeneous Region

The Neotropical Region includes a diverse and complex mosaic of macrohabitats. **Figure 7** depicts large-scale species richness (=number of species) and level of endemism in South America. Despite the well-known diversity of the Amazon rainforest, the pattern in **Figure 7** highlights the importance of mammal diversity and endemism in non-amazonian dryland biomes (Mares, 1992). One of these is the semiarid thornscrub forests of the Gran Chaco, which covers more than 1 million km^2 across Paraguay, Bolivia, Argentina, and Brazil. It possesses a rich fauna of medium- and large-sized mammals equivalent to some of the richest tropical rainforests (Redford *et al.*, 1990). Other important dryland biomes are Mediterranean scrublands, Cerrado, Patagonian steppe, high Andean Altiplano, and Monte desert. Some of these, such as the Mediterranean shrublands and the Chaco, are poorly protected in existing reserves. Further, the plains of the Gran Chaco are currently undergoing a rapid and extensive process of transformation driven by conversion to soybean production (Zak *et al.*, 2008).

Conservation of Neotropical mammals has been approached at various scales and political levels. Scales range from entire assemblages to populations and from major hotspots (e.g., Mesoamerica) to ecoregions or areas of particular interest (e.g., transitional zones between Neotropical subregions). Political approaches have been at the national level

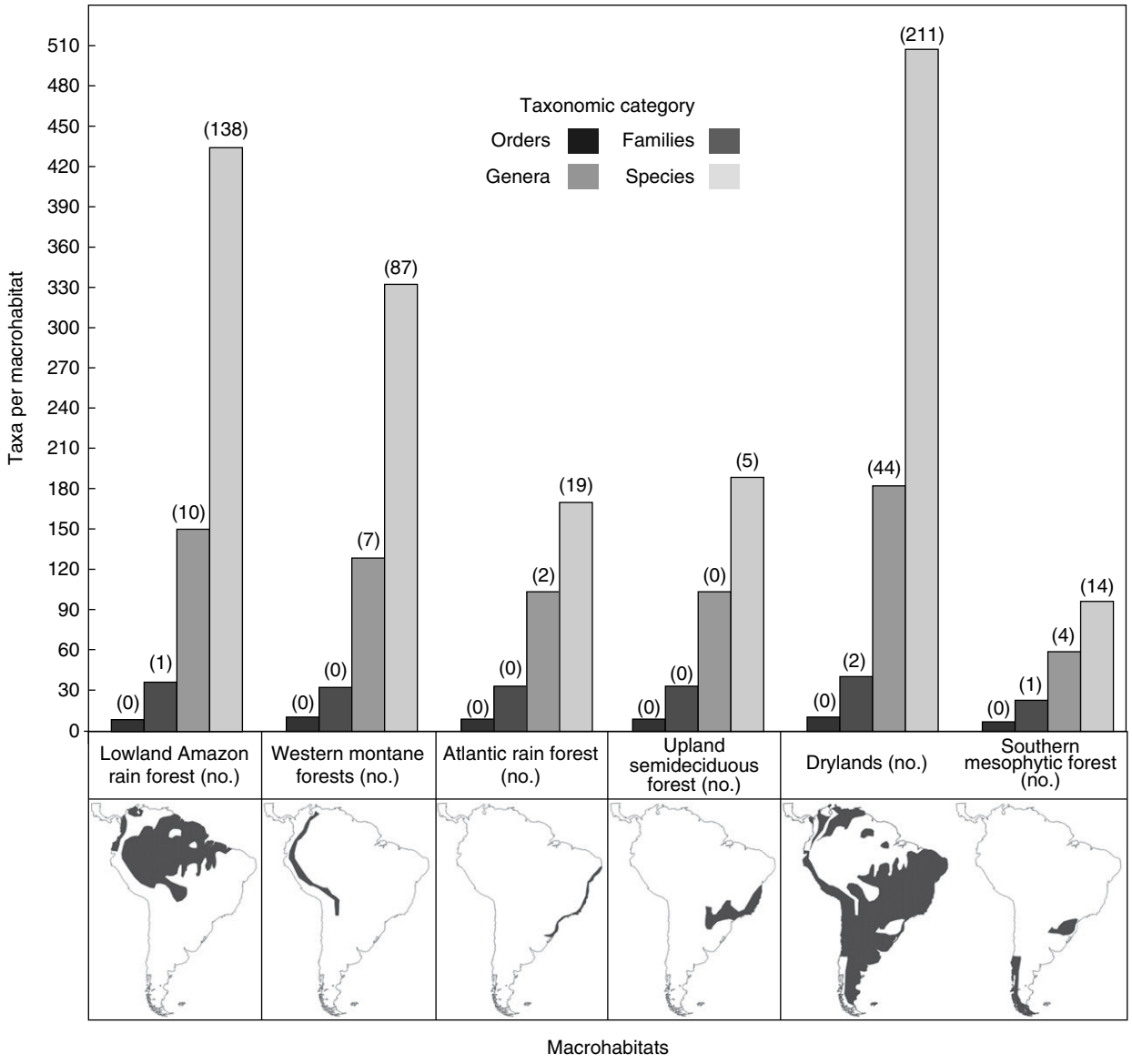


Figure 7 Mammal richness (orders, families, genera, and species) and endemisms in major South American macrohabitats. These macrohabitats differ in size (i.e., area) and the number of species they support is directly related to their area.

down to smaller subdivisions (e.g., provinces). Biodiversity hotspots, i.e., areas which combine extraordinary levels of endemism with low proportions of remaining original habitat, are important in planning and prioritizing mammal conservation at a global scale. The Neotropical Region includes seven hotspots in Central and South America. Among them are the tropical Andes, one of the most diverse regions on Earth. Other hotspots include the Caribbean, which retain 11.3% of their primary vegetation, and the Brazilian Atlantic Forest, with less than 10% of the original habitat. The other hotspots are the Cerrado (the second largest biome in Brazil), the Valdivian temperate rainforests of Chile, and the Tumbes-Chocó-Magdalena wetlands (stretching from Panama to Peru). (Myers *et al.*, 2000; http://www.biodiversityhotspots.org).

Despite their potential value, most hotspots are incongruent when considering several biodiversity attributes such as species richness, endemism, and extinction threats. Consequently, there is a need for complementary approaches that incorporate conservation priorities and establishment of reserves and networks.

A useful alternative approach to hotspots is the use of ecoregions. Ecoregion divisions are based on classical biogeography and provide a finer resolution in assessing biodiversity. They can be used as a tool for conservation of species, habitats, and ecological processes of the ecoregion under study (http://www.worldwildlife.org/science/ecoregions/delineation.html) (Olson *et al.*, 2001).

At a country scale, reserve networks (i.e., areas representing the total number of species, endemic species, endangered species, and species with restricted distributions) provide contrasts for maintaining regional biodiversity. For example, in megadiverse Mexico, 82% of mammal species are represented in the reserve network, which covers a small portion (3.8%) of the country (Ceballos, 2007). In contrast, in temperate Chile, only a small fraction of the ranges of endemic threatened species are under protection of the reserve system (Cofré and Marquet, 1999; Tognelli *et al.*, 2008).

Prioritizing conservation in areas with high species richness and large numbers of endemic and endangered species is difficult due to the limited geographic overlap in these attributes for any given area. An example of an heterogeneous region is the tropical–temperate transitional belt of the southern cone of South America, where faunal differences in a variety of biomes requires the addition of many complementary reserves to existing ones in order to protect 90% of the total number of mammal species (Ojeda *et al.*, 2003). Overall, regional biogeographic and macroecological analyses of mammal composition, population abundance, degree of occupancy, endemicity, and endangerment across the diversity of ecoregions can provide a baseline for additional studies and the establishment of global conservation strategies such as a network of protected areas maximizing complementarity (i.e., protected areas are chosen with the goal of efficiently achieving representation of all species within a particular criteria, such as endemic species, rare species, and so on; Ceballos, 2007).

Prospects

Prospects for understanding the biogeography and conservation of Neotropical mammals are promising. This is not only because of the growing development and quality of georeferenced databases of mammalian distributions and geographic ranges (e.g., http://www.natureserve.org/infonatura/) which help address the Wallacean shortfall (i.e., lack of biogeographic data and its variability), but also because of progress in delineating taxonomic units (Linnean shortfall) and systematic relationships through the advance of genetic and molecular techniques (Lomolino, 2004). Further advances in software applications and sophisticated modeling techniques in spatial analysis of distributions (e.g., GLM, generalized linear models; niche models, as BIOCLIM, GARP, and so on) hold promise of further development for biogeography, conservation, and the macroecology of Neotropical mammals. Together, all of these developments are strengthening the scientific basis of biodiversity conservation and management policies at several geographic scales and political units. However, there is also a need for interdisciplinary *dialog* in order to integrate and build a stronger, more rigourous biogeographical approach, instead of isolated research programs (Brown, 2004). This includes the consolidation of the theoretical foundations of this interdisciplinary science, and making research in this area more useful and relevant to the many problems of rapidly changing human landscapes at a global scale.

The extraordinary biotic diversity and heterogeneous matrix of the Neotropics, one of the most diverse biogeographic regions of the world, is presently experiencing unprecedented threats to major habitats due to a wide array of rapidly expanding human pressures, particularly those associated with commodity production (e.g., impacts of habitat conversion for soybean farming in the Amazon Basin, Cerrado, Chaco, and Pantanal).

A concerted effort among different segments of society is necessary to conserve the enormous mammalian diversity of the region. Unless there is a restructuring process, within the political domain, to equilibrate the disparity among social, environmental, and economic dimensions, the biodiversity losses in the Neotropical Region could escalate rapidly in the near future in the face of rapidly accelerating threats.

Acknowledgments

The author appreciates the invitation extended by Sandra Díaz to participate in the new edition of the Encyclopedia. He appreciates the thorough reviews of Peter Meserve and Marcelo Tognelli, as well as the suggestions on style and editorial improvements of Maria Periago, Mike Wisdom, and Agustina Ojeda. The collaboration of Benjamin Bender with the figures and original illustrations was decisive for the completion of the chapter. He welcomes the support of his institution, the Instituto Argentino de Investigaciones de Zonas Aridas (IADIZA) and the funding agencies CONICET and SECYT to his research program on mammals.

Appendix

Courses

- Mammalogy
- Biogeography
- Conservation Biology

See also: Central America, Ecosystems of. Ecosystems of South America. Endangered Mammals. Hotspots. Latitudinal Gradients of Biodiversity. Loss of Biodiversity, Overview. Mammals, Biodiversity of. Mammals, Conservation Efforts for. Mammals (Late Quaternary), Extinctions of. Mammals (Pre-Quaternary), Extinctions of. Marine Mammals, Extinctions of. Megaherbivores. South American Natural Ecosystems, Status of

References

Barnosky AD and Lindsey EL (2010) Timing of Quaternary megafaunal extinction in South America in relation to human arrival and climate change. *Quaternary International* 217: 10–12.

Brown JH (2004) Concluding remarks. In: Lomolino MV and Heaney LR (eds.) *Frontiers of biogeography: New Directions in the Geography of Nature*, pp. 361–368. Massachussetts: Sinauer Press.

Cabrera A and Yepes J (1960) Mamíferos Sud Americanos 2nd ed., 2 vols. Ediar, Buenos Aires.

Ceballos G (2007) Conservation priorities for mammals in megadiverse Mexico: The efficiency of reserve networks. *Ecological Applications* 17: 569–578.

Ceballos G and Ehrlich P (2009) Discoveries of new mammal species and their implications for conservation and ecosystem services. *Proceedings of the National Academy of Science* 106: 3841–3846.

Ceballos G and Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science* 296: 904–907.

Ceballos G and Simonetti J (2002) *Diversidad y conservación de los mamíferos Neotropicales*. Mexico: CONABIO.

Chown SL, Sinclair BJ, Leinaas HP, and Gaston KJ (2004) Hemispheric asymmetries in biodiversity – a serious matter for ecology. *PLoS Biology* 2: 1701–1707.

Cofré H and Marquet PA (1999) Conservation status, rarity, and geographic priorities for conservation of Chilean mammals: An assessment. *Biological Conservation* 88: 53–68.

Darwin C (1860) *The Voyage of the Beagle*. New Jersey: Doubleday.

Dillehay TD, Ramírez C, Pino M, Collins MB, Rossen J, and Pino-Navarro JD (2008) Monte Verde: Seaweed, food, medicine, and the peopling of South America. *Science* 320: 784–786.

Escalante Tania, Rodriguez-Tapia G, Szumik C, Morrone JJ, and Rivas M (2010) Delimitation of the Nearctic region according to mammalian distributional patterns. *Journal of Mammalogy* 91: 1381–1388.

Hershkovitz P (1958) A geographic classification of Neotropical mammals. Chicago natural history museum. *Fieldiana Zoology* 36: 581–620.

Hershkovitz P (1972) The recent mammals of the Neotropical Region: A zoogeographic and ecological review. In: Keast A, Erk FC, and Glass B (eds.) *Evolution, Mammals, and Southern Continents*, pp. 311–431. Albany: State University of New York Press.

Hershkovitz P (1987) A history of the recent mammalogy of the Neotropical Region from 1492 to 1850. In: Patterson BD and Timm RM (eds.) *Studies in Neotropical Mammalogy. Essays in Honor of Philip Hershkovitz* 39. pp. 11–98. Field Museum of Natural History, Fieldiana Zoology.

Iriarte JA and Jaksic F (1986) The fur trade in Chile: An overview of seventy-five year of export data (1910–1984) *Biological Conservation* 38: 243–253.

Jaksic FM, Iriarte JA, Jimenez J, and Martınez DR (2002) Invaders without frontiers: Cross-border invasions of exotic mammals. *Biological Invasions* 4: 157–173.

Kaufman DM (1995) Diversity of New World mammals: Universality of the latitudinal gradients of species and bauplans. *Journal of Mammalogy* 76: 322–334.

Kaufmann DM and Willig MR (1998) Latitudinal patterns of mammalian species richness in the New World: the effects of sampling method and faunal group. *Journal of Biogeography* 25: 795–805.

Lamoreux JF and Lacher Jr. TE (2010) Mammalian endemism, range size and conservation status in the southern temperate zone. *Diversity and Distributions* 16: 922–931.

Lomolino MV (2004) Conservation biogeography. In: Lomolino MV and Heaney LR (eds.) *Frontiers of Biogeography: New Directions in the Geography of Nature*, pp. 361–368. Massachussetts: Sinauer Press.

Lomolino MV, Riddle BR, Whittaker RJ, and Brown JH (eds.) (2010) *Biogeography*. Sinauer Press.

Lyons KS and Willig MR (1999) A hemispheric assessment of scale dependence in latitudinal gradients of species richness. *Ecology* 80: 2483–2491.

MacLean IMD and Wilson RJ (2011) Recent ecological responses to climate change support predictions of high extinction risk. *Proceedings of the National Academy of Science* 108: 12337–12342.

Mares MA (1992) Neotropical mammals and the myth of Amazonian biodiversity. *Science* 255: 976–979.

Mares MA and Genoways H (1982) *Mammalian Biology in South America*. Special Publication Series, Pymatuning Laboratory of Ecology, University of Pittsburgh, PA.

Mares MA and Ojeda RA (1982) Patterns of diversity and adaptation in South American hystricognath rodents. In: Mares MA and Genoways HH (eds.) *Mammalian Biology in South America*, pp. 393–432. Special Publication Series, Pymatuning Laboratory of Ecology, University of Pittsburgh, PA.

Marshall L (1988) Land mammals and the Great American interchange. *American Scientist* 76: 380–388.

Martin PS (1967) Pleistocene overkill. In: Martin and Wright (eds.) *Pleistocene Extinctions: The Search for a Cause*, pp. 75–120. New Haven, Connecticut: Yale university Press.

Morrone JJ (2009) *Evolutionary Biogeography. An Integrative Approach with Case Studies*. New York: Columbia University Press.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Novillo A and Ojeda RA (2009) The exotic mammals of Argentina. *Biological Invasions* 10: 1333–1344.

Ojeda RA and Mares MA (1982) Conservation of South American mammals: Argentina as a paradigm. In: Mares MA and Genoways HH (eds.) *Mammalian Biology in South America*, pp. 505–521 Spec. Publ.Series No 6, Pymatuning Lab. Ecology, Linnesville, PA.

Ojeda RA, Stadler J, and Brandl R (2003) Diversity of mammals in the tropical–temperate Neotropics: Identifying hotspots on a regional scale. *Biodiversity & Conservation* 12: 1431–1444.

Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the World: A new map of life on Earth. *BioScience* 51: 933–938.

Ortega J and Arita H (1998) Neotropical–Neartic limits in Middle America as determined by distribution of bats. *Journal of Mammalogy* 79.

Patterson B and Pascual R (1972) The fossil mammal fauna of South America. In: Keast A, Erk FC, and Glass B (eds.) *Evolution, Mammals and Southern Continents*, pp. 247–309. Albany: State University of New York Press.

Patterson BD (2000) Patterns and trends in the discovery of new Neotropical mammals. *Diversity and Distributions* 6: 145–161.

Redford KH and Robinson JG (1991) *Neotropical Wildlife Use and Conservation*. Chicago: The University of Chicago Press.

Redford KH, Taber A, and Simonetti J (1990) There is more to biodiversity than the tropical rainforests. *Conservation Biology* 4: 328–330.

Reeder DM, Helgen KM and Wilson DE (2007) Global trends and biases in new mammal species discoveries. OCCASIONAL PAPERS, Museum of Texas Tech University 269, pp. 1–35.

Reig OA (1981) Teoría del origen y desarrollo de la fauna de America del Sur. *Monografie Naturae, Museo Municipal de Ciencias Naturales Lorenzo Scaglia*, Número 1. Mar del Plata, Argentina.

Rodríguez MA, Olalla-Tárraga MÁ, and Hawkins BA (2008) Bergmann's rule and the geography of mammal body size in the Western Hemisphere. *Global Ecology and Biogeography* 17: 274–280.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge University Press.

Rosenzweig ML and Sandlin EA (1997) Species diversities and latitudes: listening to area's signal. *Oikos* 80: 172–176.

Ruggiero A (1994) Latitudinal correlates of the sizes of mammalian geographical ranges in South America. *Journal of Biogeography* 21: 545–559.

Ruggiero A, Lawton J, and Blackburn T (1998) The geographic ranges of mammalian species in South America: Spatial patterns in environmental resistance and anisotropy. *Journal of Biogeography* 25: 1093–1103.

Schipper J, Chanson JS, Chiozza F, *et al.* (2008) The status of the world's land and marine mammals: Diversity, threat, and knowledge. *Science* 322: 225–230.

Sclater PL (1858) On the general geographical distribution of the members of the class Aves. *Journal of the Linnean Society, Zoology* 2: 130–145.

Simpson GG (1950) History of the Fauna of Latin America. *American Scientist* 38(1950): 361–389.

Simpson GG (1964) Species density of North American recent mammals. *Systematic Zoology* 13: 57–73.

Simpson GG (1980) *Splendid Isolation*. New Haven: Yale University Press.

Solari S, Velazco PM, and Patterson BD (2012) Hierarchical organization of Neotropical mammal diversity and its historical basis. In: Patterson BD and

Costa S (eds.) *Bones, Clones, and Biomes: The History and Geography of Recent Neotropical Mammals*. Chicago: University of Chicago Press.

Steppan, Scott J (1996) Sigmodontinae. Neotropical mice and rats. Version 01 January 1996. http://tolweb.org/Sigmodontinae/16548/1996.01.01 in The Tree of Life Web Project, http://tolweb.org/

Stevens GC (1989) The latitudinal gradients in geographical range: How so many species coexist in the tropics. *American Naturalist* 133: 240–256.

Tognelli MF and Kelt DA (2004) Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. *Ecography* 27: 427–436.

Tognelli MF, Ramirez de Arellano PI, and Marquet PA (2008) How well do the existing and proposed reserve networks represent vertebrate species in Chile? *Diversity and Distributions* 14: 148–158.

Vaughan TA, Ryan JM, and Czaplewsky NJ (2011) Mammalogy. Fifth Ed. Jones and Bartlett Publishers.

Wallace AR (1876) *The Geographical Distribution of Animals*. New York: Harper and Brothers.

Webb SD and Marshall LG (1982) Historical biogeography of recent South American land mammals. In: Mares MA and Genoways HH (eds.) *Mammalian Biology in South America*, pp. 39–52. Special Publication Series, Pymatuning Laboratory of Ecology, University of Pittsburgh, PA.

Willig MR and Lyons SK (1998) An analytical model of latitudinal gradients of species richness with an empirical test for Marsupials and Bats in the New World. *Oikos* 80: 93–98.

Zak MR, Cabido M, Caceres D, and Diaz S (2008) What drives accelerated land cover change in Central Argentina? Synergistic consequences of climatic, socioeconomic,and technological factors. *Environmental Management (2008)* 42: 181–189.

Diversity, Community/Regional Level

Howard V Cornell, University of Delaware, Newark, DE, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 166–177, © 2001, Elsevier Inc.

Glossary

Between-habitat diversity Degree of dissimilarity in species composition between two habitats. The greater the number of species shared by two habitats, the lower the between-habitat diversity.

Metacommunity An assemblage of species in a patchy landscape, each species comprising a metapopulation in that its distribution shifts among habitat patches because the distribution is determined by colonization and extinction dynamics.

Metapopulation Group of subpopulations distributed among habitat patches in a landscape, each subpopulation being subject to periodic extinction and subsequent recolonization from other occupied patches.

Richness Simplest measure of diversity, represented by a count of the number of species in some ecologically meaningful unit (community, metacommunity, region) without regard to variation in the number of individuals per species.

Saturation Upper limit to species richness within a community, set by species interactions, and independent of the pool of colonists to which that community is accessible.

Species pool Set of species that are able to colonize a habitat of interest.

Type I community An assemblage of species where the richness is a constant proportion of the number of species occurring in a larger geographic unit in which the community is embedded.

Type II community An assemblage of species where the richness is independent of the number of species in a larger geographic unit in which the community is embedded.

Two Perspectives on the Control of Community/Regional Diversity

Local Control of Species Richness

The Volterra–Gause Perspective

One of the most obvious and fundamental facts of ecology is that all species do not occur everywhere. Instead, each species has a unique distribution determined by its ecological context and evolutionary history. Nevertheless, species distributions overlap to varying degrees so that some places support more species than others do. Such variations in species richness have been a central issue in ecological and biogeographical studies at least since the turn of the century. Until recently, explanations focused mainly on processes operating within ecological communities over relatively short periods of time. This perspective arose out of a desire to add scientific rigor to ecological studies by focusing on mechanisms and ignoring history. A mechanistic approach made it possible to model ecological systems with simple equations and to test these models and their assumptions with carefully collected data.

One of the earliest truly influential models in community ecology was developed by Vito Volterra (and independently by Alfred Lotka) in the mid-1920s. This model is a set of simple differential equations describing competition between two species in terms of reductions in per capita population growth rates by interactions among individuals. When the model is analyzed at equilibrium, the outcome of competition (coexistence or competitive exclusion) is easily predicted from the intensity of interactions within species relative to those among species. If populations limit their own growth more than the growth of other populations, the two species will coexist. If they limit other populations more than themselves, one will exclude the other. The model and its assumptions were later tested by G. F. Gause in laboratory systems comprising several species of protozoa. The tests lent strong support to the model, and as a result the so-called Volterra–Gause perspective has become a metaphor for the way in which ecologists think about control of species richness.

According to this perspective, species richness in a biogeographical region may indeed be determined in part by history. However, history has little relevance to the upper limits on local richness within communities. Instead local richness is limited primarily by species interactions (predation, parasitism, herbivory, competition, mutualism) in conjunction with abiotic conditions (water, nutrient, and energy availability, chemistry, temperature, disturbance). In other words, the outcome of interactions is determined in large part by the abiotic conditions in which they occur. For example, lower concentrations of a particular nutrient or lower water availability might favor one plant species in competition with another, but higher values of these factors might result in a reversal of competitive dominance. The number of species in the biogeographical region provides a species pool from which the local communities are drawn, and all species have the potential for colonizing all habitats in the region. Since habitats vary in their abiotic conditions and resident assemblages, only those species capable of coexisting in a given milieu will establish successfully. Differences in species richness thus result from differences in the abiotic template and species composition among habitats. In a recent series of publications from the early 1990s, James Drake has shown that the colonization order of species into a habitat can also have some effect, and communities can exhibit multiple stable points representing different levels of species richness.

A crucial element of this perspective that derives directly from the Volterra model is that communities exist at a dynamic equilibrium. For example, in the case of competition, population levels of the resident species are limited by resource availability and physical conditions. Thus, if sufficient time elapses and there are no perturbing disturbances to the community, the number of species will saturate at a level determined by the number of limiting resources and physical factors. At saturation, species are sufficiently similar in their resource use that finer subdivision of resources becomes increasingly difficult and new colonizations into the community must be accompanied by extinctions of current residents. If habitats throughout the region are all identical and saturated with species, then the total number of species in the region is limited to the total number coexisting within each identical habitat. That is, community and regional richness are identical and regional richness is limited by the same factors that limit local richness.

Since all habitats are not identical, species within the region have the option of partitioning habitats as well as resources and thus the same set of species will not occur everywhere in the region. Nevertheless, local processes are still limiting both local and regional richness, but limits on regional richness now result from interplay between species interactions and spatial habitat variation. The greater the number of habitats, the greater the between-habitat diversity and the greater the number of species in the region relative to those within a single community. Historical processes such as massive extinctions might reduce the number of species in the region for a time, and subsequent evolution, speciation, and dispersal might generate a new biota that differs from the previous one. However, the essential feature of this viewpoint is that given sufficient time, the imprint of history on species richness will fade, communities will fill up and become saturated, and processes operating locally will also limit richness regionally.

Even strict adherence to the Volterra–Gause viewpoint still allows for other factors to affect local and regional richness. All else being equal, larger localities and regions will support more species than smaller ones as predicted by the species-area relationship. In addition, some species within a region might be ecologically equivalent and thus might be able to substitute for one another within a given community type. Chance might determine what species occupies what particular community, in which case ecologically equivalent species could increase regional richness without affecting local richness. Finally, spatial interspersion of habitat types might allow species adapted to one habitat type to persist in an adjacent, less favorable habitat type by means of propagule flow. This so-called source–sink effect proposed by Ron Pulliam in 1988 would increase local richness in the less favorable habitat above that which would be self-maintaining in the absence of immigration. Although these factors may have important impacts on species richness, they do not alter the premise that the primary direction of control on species richness is from the local to the regional level.

An ecological community at equilibrium and not subject to vagaries of history evokes a powerful image that continues to influence modern ideas about species diversity. Robert MacArthur's (1972) book represents the pinnacle of this viewpoint, although he recognized that the equilibrium assumption and local control of richness did not apply to all communities. There are clear recent demonstrations that species interactions have strong influences on community structure and might exclude species from the community (see Lawton, 1999, for a brief review). In short, the Volterra–Gause perspective has been a useful guide to a limited understanding of the mechanisms by which particular communities are structured. However, this understanding has led to few if any general patterns or rules that apply to most systems (Lawton, 1999).

Species–Energy Theory

Another view of local control of species richness comes from the ecosystem perspective, and has been termed "species–energy theory." The theory was developed mainly by David Wright, David Currie, and Brian Maurer (1993). Species-energy theory posits that local and, ultimately, regional richness are limited by available energy within the local environment. Available energy limits local richness by placing ceilings on the population densities of individual species. As energy decreases, densities decrease, random extinctions increase, and local richness is reduced. Available energy can be scaled up to whole regions by multiplying the energy per unit area by the total area of the region. Thus, population densities and species richness over the whole region are also controlled by energy availability.

Although the theory in this simplified version posits local control of local and regional richness, it differs from the Volterra–Gause viewpoint in that species interactions are of secondary importance in limiting richness. Energy can have direct and independent effects on population levels to increase richness rather than mediating interactions among species. Interactions can be included in the model, but they are not necessary for limits on richness to occur. Species–energy theory is very similar to island biogeography theory, which proposes that species richness is a balance between colonization and extinction rates in a local habitat. Smaller habitats have higher per species extinction rates and thus lower species richness as predicted by the species–area relationship. The main difference between the two theories is that species–energy theory substitutes energy availability for area as the predictor of species richness.

A recent literature review by David Wright and others demonstrated that several proxy variables for energy availability (precipitation, potential or actual evapotranspiration, productivity) are indeed strong predictors of species richness for various taxonomic groups (**Table 1**). However, it is not clear whether energy is the direct cause of richness variation, or whether it simply covaries with other possible causes such as area effects, habitat heterogeneity, species interactions, and historical effects. It is also not clear whether the assumptions of the theory are always met in particular studies. For example, it is assumed that energy limitation should operate at the local scale where individuals of each species seek sufficient energy to persist in the community. However, it is at the scale of the local community that the correlation between energy and richness is weakest and most variable; the correlation is strongest at larger spatial scales (104–108 km^2). It is also assumed that average population size and species richness

should increase together with available energy, but this is not always the case. In particular, David Tilman and Pacala have pointed out that plant systems often show an inverse relationship between per species population sizes and energy as predicted by the $-3/2$ power law, yet species richness commonly increases with energy in the expected way. Despite these limitations, energy can be a useful variable for estimating richness at larger scales because covariation is high regardless of the underlying causes, and because energy data (or proxies) are available for most geographic areas.

At more local scales, species interactions among resident populations may come into play, altering any potential increases in richness with energy availability. Such interactions might partially explain the more variable relationship between energy and richness within communities. Increased energy can lead to increased biomass of the total community, which in turn can lead to increases or decreases in average population density depending on the intensity of interindividual interactions and the size plasticity of the taxa involved. But it is the changes in the relative allocation of energy among species that are critical for predicting changes in richness. If allocation changes as energy availability increases, for example, caused by shifts in competitive superiority or ability to avoid predation, then the population densities of some species may increase while other species may be driven extinct. Richness can thus increase or decrease with increased energy depending on the circumstances. It is at this point that the species–energy theory converges with the Volterra–Gause viewpoint since the environment is now mediating interactions among species.

Table 1 Energy-related and other factors which correlate with species richness

Factor cited	*Correlations with species richness*	
	Significant	*Not significant*[a]
Energy-related factors		
Mean annual heat and/or humidity variables	29	0
Productivity (sometimes est. as standing crop)	9	0
Nutrient, food availability	3	0
Seasonality	4	0
Other factors		
Habitat complexity, subdivision, microrelief	9	3
Disturbance	5	5 (2)
Environmental chemistry (pH, cations)	3	1 (1)
Isolation, peninsula effects	2	1 (1)
Diversity of food, prey	6	1
Regional richness	2	0
Time, historical factors	1	2 (2)
Competition, predation	0	1
Other	3	1

[a]Numbers in parentheses indicate the number of correlations that were omitted from further analysis because no correlation coefficient was reported.

Note: Factors investigated as the principle determinant of species richness in 82 correlations from 53 studies. Reproduced from Wright, *et al.* in Ricklefs RE and Schluter D (eds.) (1993) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. Chicago: University of Chicago Press.

The Diversity–Productivity Relationship

The idea that richness can change with energy availability in unpredictable ways has had a long history under the guise of the diversity–productivity relationship. Productivity is a measure of energy availability to the community and has been shown to affect local richness in contradictory ways; richness either increases, decreases, or shows no relationship with productivity depending on the assemblage. Recently it has been suggested that these contradictory relationships may just be incomplete segments of an overall hump-shaped, unimodal relationship over a broader range of productivity. In other words, richness peaks at intermediate productivity levels and declines toward the extremes. Different taxa can show richness peaks at different levels of productivity, resulting in unimodal curves that are displaced along the productivity spectrum. No ecological pattern is universal and, accordingly, contradictory relationships between richness and productivity need not always imply a truncated hump. Indeed, evidence for this possibility is currently limited at best (Waide *et al.*, 1999). Monotonic relationships may be just as frequent as unimodal ones. Nevertheless, the unimodal curve is a useful synthetic construct and it is sufficiently frequent to invite explanation. Moreover, the same explanations for each arm of the hump may apply equally well to monotonically increasing or decreasing relationships over a particular range of productivity.

It is not clear why richness shows this strange relationship with productivity, and the pattern remains one of the most important unsolved mysteries in community ecology. The hypothesis that productivity differences are directly responsible for the unimodal relationship assumes that productivity is uniformly distributed among sites. In other words, low-, medium-, and high-productivity environments occur with equal frequency. If medium-productivity sites are more frequent than low- or high-productivity sites, then the unimodal curve might just be a reflection of simple species–area effects. There is some evidence to suggest that this is not the case (Rosenzweig, 1995), but the possibility has not been extensively tested. At the very least, diversity-productivity correlations should be corrected for unequal area effects before a causative link between these variables is seriously considered.

Unimodal curves (as well as monotonically increasing or decreasing ones) can occur at different spatial scales (local, landscape, regional, continental), but their frequency changes as spatial scale increases. Monotonically increasing patterns are more common at larger scales and unimodal patterns are more common at local scales (**Figure 1**). Moreover, the mechanisms responsible for patterns at each scale probably differ (Waide *et al.*, 1999). Since the within-community, among-community, and local scales are where ecological processes become important, theories that consider changes in species interactions and local environmental conditions along productivity gradients are the most relevant.

The decline to the left of the hump may derive from increased environmental stress or the direct effects of energy reduction on population levels as proposed by species–energy theory. Other mechanisms, such as increases in disturbance rate, poor recovery from disturbance due to slower population growth rates, declines

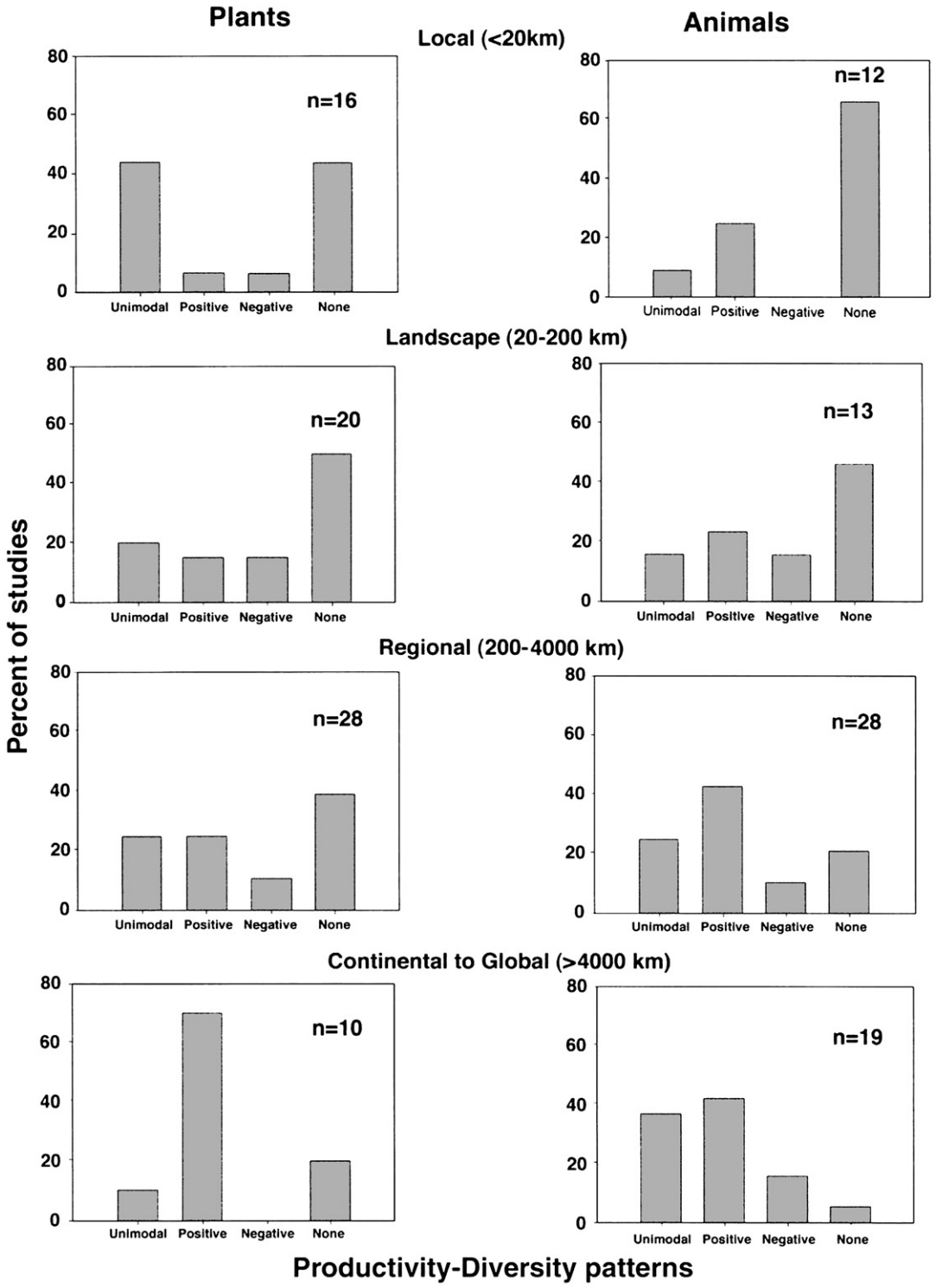


Figure 1 Breakdown of published plant and animal studies showing a unimodal, positive, negative, or no relationship between species diversity and productivity at four spatial scales. The number of studies in each histogram (*n*) is shown. Reproduced from Waide RB, Willig MR, Steiner CF, Mittelbach GG, Gough L, and Dodson SI (1999) The relationship between productivity and species richness. *Ann. Rev. Ecol Systematics* 30: 257–300.

in territoriality, and environmental heterogeneity have also been suggested to play a role. The decline to the right of the hump may involve some form of interaction mediation among species, such as increased intensity of competition and predation or increased population dynamic instability within food webs at high productivities. Such declines would then provide strong support for the idea that species interactions limit species richness in accordance with the Volterra–Gause viewpoint. However,

it is important to point out that richness declines in highly productive areas may also result from regional effects. For example if productive habitats are relatively rare, they are likely to be surrounded by larger, less productive areas and thus isolated from a large source pool of species adapted to productive conditions. As previously mentioned, area effects may also contribute to low richness. Regional effects must therefore always be taken into consideration when analyzing diversity–productivity relationships.

Tests that distinguish among these various theories are difficult to perform and have thus far been limited. Comparative approaches, although they require less disruption of the community, might not be able to distinguish among alternative theories unless natural "experiments" in which selected mechanisms are not operating can be found. Experimentation has better control of confounding variables, but care must be taken to provide enough time for all processes that contribute to changes in richness to ensure that they be allowed to respond to the manipulation. For example, in a 1999 review, Laura Gough and co-workers found that fertilization experiments consistently resulted in increases in inorganic nutrient concentrations, which in turn intensified species interactions in herbaceous plant communities. Fertilization increased productivity and reduced species richness across a broad range of community type, due primarily to increased dominance by one or a few species. These experiments certainly demonstrate the potential of interaction mediation to limit community richness. However, the relevance of such studies to richness declines in natural communities at highly productive sites is questionable. Fertilization reduces diversity even at poorly productive sites, where theory predicts that diversity should increase.

This anomalous result may be a consequence of the short-term nature and small spatial scale of such experiments. Species adapted to the higher productivities of the fertilized sites might not be able to reach them if they are isolated in space or time. More appropriate experiments should therefore manipulate the species pool to simultaneously test for the effects of productivity and the dispersal effects on richness. If pool limitation is the cause of the anomalous results, then experimentally boosted recruitment to low-productivity sites by an appropriate species pool should generate higher species richness with higher fertilization rates. Naturally, the species selected for the boosted recruitment should be well-adapted to the increased nutrient levels at the fertilized sites. High-productivity sites, on the other hand, should remain depressed in diversity owing to interspecific interactions. Some combination of comparative and experimental approaches will probably be required to solve the mystery of the diversity–productivity relationship. Regardless of the methods used, it is important that it be solved, since it represents one of the clearest possibilities that local richness is locally controlled. It also provides a single framework that subsumes both species–energy theory and Volterra–Gause processes.

Theory Doesn't Always Predict Local Control

Without question, the Volterra–Gause viewpoint has stimulated much research on the factors that control species richness. Nevertheless, the fundamental premise that local processes control local and ultimately regional richness may be incorrect for many systems. Even the restrictive Volterra–Gause models do not always predict limits on species richness due to local mechanisms. After appropriate modification to incorporate some realistic features of environments and species, such as spatial heterogeneity, environmental fluctuations, nonlinear population growth responses to environmental change, and trade-offs among species in their ability to compete in different environments, the models often predict that there are no upper limits on local richness (Tilman and Pacala, in Ricklefs and Schluter, 1993). Moreover, the Volterra–Gause assumption that population levels are at equilibrium, may not commonly be met in natural communities. In a classic example, periodic disturbance can drive populations below equilibrium, where the outcomes of interspecific interactions are never fully realized. More species can thus invade than would be allowed by equilibrium population levels.

If real communities have any of these features, it is not likely that local richness will reach saturation as predicted by the simplest models. Evidence from real communities suggests that lack of saturation is indeed common. Lack of saturation raises the possibility that factors external to the community may ultimately determine local richness. As a consequence, the factors that determine regional richness must also be sought elsewhere. In the Volterra–Gause view, regional richness is limited by a combination of the number of species within saturated habitats and the number of habitats within the region. If communities within habitats are not saturated, control on regional richness cannot be exerted at the within-habitat scale. Moreover, the number of habitats within a region is not fixed and may be more an effect of the number and kinds of species in the region than a cause of richness limitation at the between-habitat scale (Rosenzweig, 1995). In other words, the more species there are in the region, the more finely habitats can be partitioned and the more distinct habitat types that can be identified. How, then, do differences in richness among regions arise? And what relevance do these differences have for local richness? An alternative to the Volterra–Gause viewpoint that addresses these questions is thus needed.

Regional Control of Richness

Regional Processes

The alternative view rests on an idea that actually predates the Volterra–Gause view, namely, that regional richness is set to a large degree by historical biogeographical factors. The importance of history is implicit in the traditional definition of a biogeographical region that relies in part on the presence of a distinctive native biota. The presence of large groups of indigenous species implies that the region has undergone a period of isolation and independent evolution. Thus, the numbers and kinds of species that occur across the region are determined to some degree by its unique historical development. These different histories might result in widely divergent species richnesses among some components of the biota, which in turn may be reflected in local communities. Under this scenario, local richness is not limited by species interactions as predicted by simple Volterra–Gause models, but

becomes some function of the number of species in the region that are able to colonize and persist in a given habitat. The richness of the community embedded within the region thus retains some of the region's historical signature.

The key prediction arising out of this scenario that distinguishes it from the Volterra–Gause model is that localities in richer regions can support more species than those in poorer ones even though the localities are environmentally similar. This is simply because species co-occurrences within localities are determined to some extent by the same factors that determine regional co-occurrence. In their 1991 book *Phylogeny, Ecology, and Behavior*, Brooks and McLennan pointed out that regional co-occurrence can come about via two processes: association by descent and association by colonization. In association by descent, two species can co-occur because their ancestors co-occurred. In other words, the community was assembled from ancestors in the distant past, and the species present today evolved *in situ* from members of the ancestral community, most likely by geographic isolation (**Figure 2**). The community is thus rich or poor because the region in which it is embedded was historically rich or poor. In association by colonization, two species can co-occur because one evolved elsewhere and migrated into the region. Communities are open to colonization because (1) interactions are not limiting richness, (2) the colonizing species uses niche space that is not occupied by the *in situ* residents, or (3) after contact with the residents, evolution of niche partitioning occurs. In association by colonization, the community is rich or poor because its region has high or low barriers to colonization from elsewhere. It is likely that real communities are mixtures of colonists and *in situ* descendants, and Brooks and McLennan have pioneered the development of phylogenetic methods to distinguish between these two groups.

The influence of higher colonization rates into a region on regional and thus local richness is obvious, but the role of history is less so. Why are some regions historically richer than others? The answer to this question must involve region-specific differences in historical speciation rates, extinction rates, or both. Some processes might act to increase the speciation rate relative to the extinction rate. For example, some taxa may have developed key innovations that allowed them to enter a new adaptive zone and undergo an adaptive radiation. The region may have been simultaneously dissected by colonization barriers, disrupting gene flow among many populations and resulting in a burst of speciation.

Latham and Ricklefs present a possible example in the evolution of frost tolerance and proliferation into the temperate zone by forest trees. Asian temperate forests have three times more tree species than North American forests and six times more than European forests. Because Asia may be particularly diverse topographically and has no geographic barrier between the tropical and temperate zones, such proliferation may have occurred more frequently than in North America and Europe, where such barriers do exist (Ricklefs and Schluter, 1993). Alternatively, some taxa may undergo non-adaptive bursts of speciation resulting from inherent biological attributes, including genetic peculiarities (e.g., chromosomal anomalies or polyploidy), mate recognition and selection, or other factors. A dramatic example is the anomalous diversity of the Hawaiian *Drosophila*. The proliferation of this genus in Hawaii far exceeds that in any other

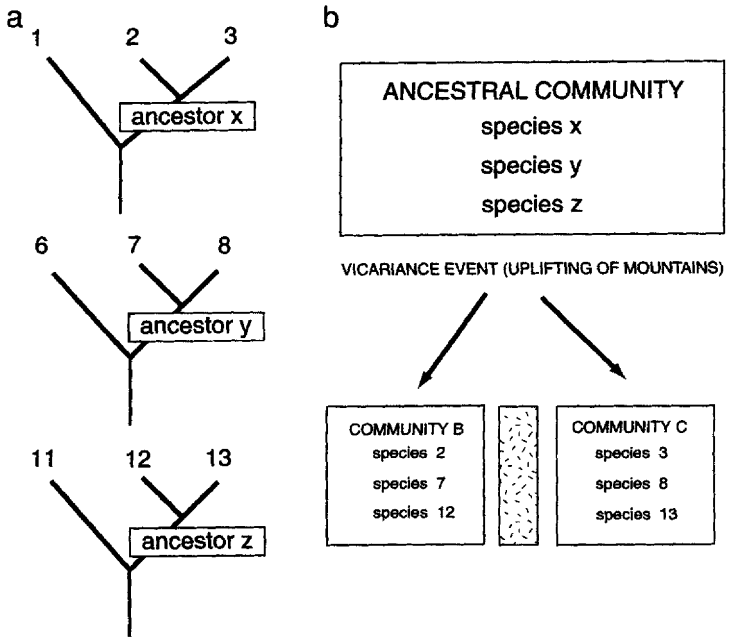


Figure 2 Phylogenetic and biogeographical patterns that suggest association by descent. Species sets (2, 7, 12) and (3, 6, 13) comprise communities that were found on opposite sides of a geographic barrier. The phylogeny (A) indicates that species sets (2, 3), (7, 8), and (12, 13) each had a common ancestor. These two patterns suggest that both communities descended from a common ancestral community that exited before the barrier was formed. (B) The species in each extant community are thus associated because their ancestors were associated. Reproduced from Brooks and McLennan, in Ricklefs RE and Schluter D (eds.) (1993) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. Chicago: University of Chicago Press.

region of equivalent area and has been attributed to a strong tendency for sexual selection in the Hawaiian taxa (see Cornell and Lawton, 1992). Adaptive and nonadaptive radiations might be distinguishable by differences in the relationship between speciation rate and richness. In adaptive radiation, speciation rate should at first exceed extinction rates and then come into balance with extinction as regional richness increases and the adaptive zone fills up. Nonadaptive radiations do not require ecological diversification, and thus speciation and extinction rates should be unaffected by regional richness. Adaptive and nonadaptive radiations are not mutually exclusive and both could be operating in the same clade.

Other processes might act to increase extinction rates relative to speciation rates in some regions. Increased probability of extinction is believed to be linked to low average population size, which in turn is correlated with geographic distribution. Thus any event that reduces a species' range, such as habitat destruction or displacement by an invader, is likely to decrease its time to extinction. Returning to temperate forest trees, glaciation in conjunction with mountain barriers in southern Europe might have squeezed the geographic ranges of many temperate tree species below a critical threshold, driving them to extinction. The absence of mountain barriers in Asia and North America may have allowed temperate species to survive in southern refuges, accounting in part for the higher tree species richness in those regions.

The relative importance of colonization, speciation, and extinction rates in setting regional richness will vary depending on the taxon and the history of the region. Since regional and local richness are linked, an understanding of community assembly will require a new focus on these macroevolutionary processes. Fortunately, new phylogenetic methods involving sister group comparisons and other innovations have been developed by Mitter, Farrell, and others, which can distinguish among various macroevolutionary causes for regional and local richness variation. Such methods have given history and biogeography the consideration they deserve in some modern investigations of community structure (e.g., McPeek and Brown, 2000), but more work needs to be done. Community structure is undoubtedly influenced by contemporaneous processes as well, and distinguishing historical from present-day effects such as energy, local environment, and species interactions will be one of the biggest future challenges in this area.

Metacommunity Processes

Once species become part of the biota of a particular region, they can disperse among habitats, and as a result communities can be quite dynamic in space and time. The species within the region can thus be viewed as a species pool for a given local habitat. The species pool will be considerably larger than the number of species in any habitat for at least three reasons: (1) some species in the region will not be able to disperse to all sites because of colonization barriers or isolation-by-distance; (2) all habitats in the region are not favorable for all species; and (3) species can be driven locally extinct by species interactions, demographic and environmental stochasticity, disturbances, or other unfavorable changes in local conditions. It is of some interest to explore situations where species cut off by highly effective colonization barriers are not considered part of the pool for habitats from which they are isolated. The dynamic links between local and regional richness should thus be examined at a spatial scale where all species can reach all habitats over ecological time. Moreover, we should focus on a single habitat type, since looking at more than one habitat will confound regional effects with local environmental variation. This can generate artifactual relationships between local and regional richness (Srivastava, 1999). The metacommunity within the landscape, in its simplest form, fulfills both of these criteria and thus represents an appropriate theoretical context in which to examine these links. The term metacommunity was first used by Gilpin and Hanski in 1991. The metacommunity can be thought of as a pool of species in a fragmented landscape of environmentally identical sites, each species behaving as a metapopulation in that its distribution among sites in the landscape is determined by colonization and extinction. In this scenario, regional processes still determine the potential species pool, but now the effects of the pool on local richness are filtered through the metacommunity.

The simplest metacommunity model derives from Richard Levins' original 1969 equation for a single-species metapopulation:

$$\frac{dp_i}{dt} = c_i p_i(1 - p_i) - m_i p_i$$

Tilman and Lehman used this equation to construct a metacommunity model by writing independent equations for as many species as occur in the metacommunity (Tilman and Kareiva, 1997). The value of p_i is the proportion of local sites in the landscape occupied by the ith species, and c_i and m_i are the site-specific colonization and extinction rates, respectively. Thus, when $dp_i/dt = 0$, each p_i is set by a balance between colonization and extinction rates per site. The model assumes that all sites are identical. This assumption permits the examination of links between local richness and pool size without the confounding effects of local environmental variation. If p is defined as the mean of p_i's for all species, then p becomes the average proportion of local sites occupied by a species. Alternatively, it is the average proportion of the metacommunity found at a particular site. Thus, local richness = p(landscape richness).

What this expression tells us is that if p is a constant then local richness is a linear function of landscape richness, a pattern that has been called Type I or proportional sampling (**Figure 3**). Proportional sampling predicts that matched habitats embedded in richer landscapes will always have richer communities, but that local richness will be less than landscape richness by a constant proportion. The slope of the relationship, p, will be determined by the size of the pool balanced by the rate of species extinction at individual sites. In this simple model, species-specific colonization and extinction rates are assumed to be independent of pool size and the number of species at a site. Thus the model predicts the relationship between local and landscape richness in the absence of species interactions. Under these conditions, communities will be unsaturated and open to colonization from the species pool.

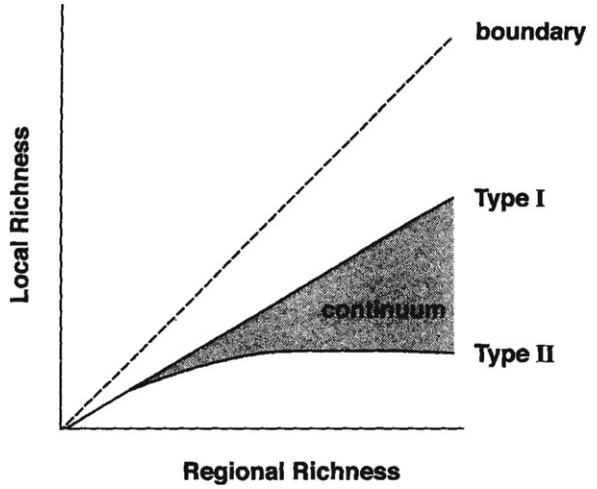


Figure 3 The shape of the relationship between local and regional richness in Type I and Type II communities. The linear relationship in Type I communities suggests that they are strongly influenced by the size of the regional species pool. Type II communities are asymptotic, suggesting that they are saturated and that further increases in the regional pool will not affect local richness. Real communities can also occupy the gray zone. Reproduced from Cornell HV and Lawton JH (1992) Species interactions, local and regional processes, and the limits to richness in ecological communities. *J. Animal Ecol.* 61: 1.

At least one recent study by Hugueny and Cornell (2000) has shown that the proportional sampling pattern in a real species assemblage can be predicted from colonization and extinction rates. Metacommunity models can be made more realistic by adding rescue effects, by adding provisions for variable population sizes among patches, by making habitats differentially suitable to different species, and by incorporating source-sink effects and habitat specialism. Such modifications may alter the details of the Type I curve, but they will not alter the conclusion that pool effects strongly influence community richness. Thus, for the purposes of conceptual distinction between saturated and unsaturated communities, the Levins model suffices.

The Levins model not only predicts the relationship between landscape and local richness; it also shows that the landscape filter constrains the species pool. Most obviously, landscapes that are smaller will have smaller species pools as predicted by the species–area relationship. More subtly, landscapes of a given area that are more highly fragmented will have smaller species pools because they suffer higher landscape-scale extinction rates. Fragmentation also reduces the average proportion of the pool (p) found at a local site by lowering per site colonization rates and increasing per site extinction. Thus any differences in landscape characteristics must be considered when estimating the species pool and evaluating the proportional sampling relationship.

If competitive interactions are added to the model, and it is assumed that competition leads to extinction rather than character displacement and speciation, then increased competition can result in increased extinction rates and/or decreased colonization rates per species. As the species pool increases, competition could become sufficiently intense that colonization by a new species into the community would either fail or result in the rapid extinction of other species. Under this condition, further increases in the pool do not result in increases in local richness and the community becomes saturated with species. At saturation, the relationship between local and landscape richness levels off and becomes independent of the rate of colonization and the size of the species pool. This leveling relationship has been designated Type II (Cornell and Lawton, 1992). Species interactions are thus limiting local richness consistent with the Volterra–Gause perspective. Other dynamic models developed by Caswell and Cohen (in Ricklefs and Schluter, 1993) and Morton and Law (1997), which allow for dispersal, species interactions, and local extinction, also predict leveling off and saturation.

However, the spatial character of many dynamic models allows for realistic complexities that ameliorate the impact of species interactions on local richness. For example, if competition occurs only among close neighbors, and inferior competitors are better dispersers, then a potentially unlimited number of species can coexist locally (Lehman and Tilman, in Tilman and Kareiva, 1997). Other spatial models incorporating disturbance, limited dispersal, spatially local competition, and habitat heterogeneity also predict coexistence at high levels of richness in the face of competitive interactions (Cornell and Karlson, and Lehman and Tilman, in Tilman and Kareiva, 1997). Thus Type I relationships can still occur in the presence of intense competition.

It is important to point out that just because metacommunity and other models predict proportional sampling or saturation under certain conditions, it does not mean that the Type I or Type II relationships need be explained exclusively by mechanisms operating at the landscape scale. The models simply represent theoretically plausible ways in which the pattern can be generated. Factors operating at other spatial scales can also produce Type I and Type II relationships. For example, suppose the scale of the analysis is expanded to the entire region. One reason to do this is that regions are easier to define than landscapes and thus more data on regional richness have been collected. Type I and Type II relationships are thus more conventionally represented as plots between local and regional richness (see **Figure 3**). Under these circumstances, Type I relationships can result simply from dispersal limitation. That is, when regions of different species richness are compared, a constant proportion of the regional biota is unable to reach the habitat type of interest. Type II relationships, on the other hand, can be generated by historically stable communities in which all niche space is occupied and into which invasion is limited. In such situations, the community need not be particularly dynamic, need undergo little turnover in ecological time, and its species composition may bear the strong stamp of its historical development. Nevertheless, a Type I relationship would still be an indication that hard limits on local richness have not been reached even in the richest regions, whereas a Type II relationship would suggest saturation. The Type I and Type II relationships should thus be viewed as empirical patterns with multiple causes, but also as evidence that local richness is or is not limited by local factors.

Testing for Saturated Communities

The Method

The preceding theoretical exploration suggests that one can test for saturation by examining the relationship between local

richness and some reasonable estimator of the species pool. Regional richness has been the most commonly used estimator for the reasons stated earlier. To test for saturation, matched habitats are compared in different geographical regions of similar size that support different numbers of species. Habitats in richer regions are presumably exposed to richer species pools. Choosing regions of similar size should reduce any differential effects of colonization barriers that might generate differential dispersal limitation among regions (see Cornell and Lawton, 1992; Caley and Schluter, 1997). If the communities are unsaturated, local richness should increase steadily with increases in regional richness. The shape of the relationship can be linear (Type I) or curvilinear, but it cannot level off. Curvilinearity would suggest some resistance to invasion and thus a tendency toward saturation in richer regions, but that the limit to local richness has not been reached. If communities in the richest regions are saturated, then the curve will reach an asymptote and local richness becomes independent of regional richness (Type II).

The method assumes that matched habitats and species in different regions are similar in ways that are critical to species coexistence. This means that if there is the potential for saturation at some level of local richness determined by local environmental conditions and the characteristics of individual species, that level will be the same from region to region. However, even if matched habitats are environmentally identical, some communities might exhibit multiple stable points and saturate at different levels, for example, due to differences in colonization order as shown by Drake. If the community is near saturation, multiple stable points might increase the variance around the putative asymptote but should not affect the average tendency for local richness to level off.

The Tests

Two recent surveys by Cornell and Karlson (1997) and Srivastava (1999) have evaluated the frequency of Type I and Type II relationships in a variety of species assemblages. The surveys were combined by Lawton (1999) to yield 38 papers. Some of the data were not analyzed strictly according to the protocol presented here, and in these cases interpretation of the pattern was conservative and based on common sense. By conservative estimate, there were at least 23 examples of Type I relationships (wood-boring beetles, tiger beetles, fig wasps, lizards, corals, plants, helminth parasites on introduced fishes, 6 examples of fish, 4 examples of plant-feeding insects, 3 examples of parasitoids, 2 examples of birds, and a variety of vertebrate taxa (birds, fish, mammals, reptiles) analyzed separately and in combination). By contrast, there were 15 examples of curvilinear relationships, and not all of these were Type II (saturating). The examples are classified as saturating (strong), weakly curved (weak), or uncertain (u) as follows: marine crustaceans (u), deep-sea gastropods (weak), tiger beetles (u), 2 examples of helminth parasites on fish (strong), helminth parasites on amphibians (strong), 2 examples of fish (u), 2 examples of birds (weak), 2 examples of mammals (strong), and 3 examples of higher plants (1 strong, 2 weak).

Caveats

Several methodological concerns about the saturation test have been raised by various authors. The most relevant of these to the present discussion are that various artifacts such as nonhomogeneous variance structures, insufficient sampling, and overestimates of the true species pool in richer regions sometimes exaggerate the curvilinearity of the relationship between local and regional richness. To the extent that such artifacts are present in some studies, true Type I relationships might appear curvilinear, and curvilinear but unsaturated relationships might appear to be Type II. The count of unsaturated patterns in the survey is thus a conservative and robust result, at least with respect to the presence of such artifacts.

Saturated or Unsaturated?

Given these caveats, at least it can be said that unsaturated patterns are common and widespread in ecological communities. Indeed, unsaturated patterns have been reported about twice as frequently as saturated ones. The survey thus confirms that local communities are often open to regional influences and have not reached any upper limit set by species interactions as predicted by the Volterra–Gause model. Those communities that have reached an asymptote might be saturated but have to be examined closely to rule out artifacts that might be responsible for the Type II shape. A Type II curve is not sufficient in and of itself to conclude that species interactions are limiting local richness. For example, helminth parasite communities on fish and amphibians appear to be saturated, but at least one expert opinion is that they are probably not (see Lawton, 1999).

The survey also supports the theoretical position that strong interspecific interactions are not always sufficient to generate a curvilinear or Type II relationship. Taxa that are known to experience strong interspecific interactions such as fish and birds show Type I or weakly curvilinear relationships that are characteristic of unsaturated communities. Although it is not certain that these particular assemblages were experiencing the strong interactions that are supposedly characteristic of these taxa, the results at least suggest that evidence for species interactions by itself might not permit generalizations about the effects of these interactions on broad patterns of community structure.

Experimental Support

Additional evidence from experimental introductions supports the survey results that communities are often unsaturated. Experimental manipulations of the species pool are only just beginning, but they are particularly useful because there is better control of confounding variables. In separate studies by Tilman (1997) and Robinson, Quinn, and Stanton (1995), seeds of native species were added to experimental plots containing different numbers of plant species and establishment was monitored. In the Robinson *et al.* study, California poppy was the test colonizer, and in the Tilman study, anywhere from 0 to 54 species were added to plots. Neither study showed evidence of saturation; seeds became

established and persisted in many plots. Unsaturation was dramatic in the Tilman study, as the number of species in the experimental plots continued to increase over the full range of species additions and the increases persisted for 4 years (Figure 4). In other words, an asymptote characteristic of the Type II relationship was never reached. However, plots that were richer in resident species had lower establishment rates, suggesting some resistance to invasion.

The results of these experiments must be interpreted with caution, as they represent short-term responses to pulsed increases in the species pool. There are almost certainly time lags in the response of the community to species additions, and long-term monitoring will be required to confirm that the "invaders" have become established without driving "residents" extinct over ecological time. Nevertheless, pool manipulations offer a powerful probe for saturation and need to be expanded to other taxa and habitats to test the generality of these results.

Matters of Spatial Scale

A central issue in the exploration of saturation in communities is the spatial scale at which local richness should be measured. Ideally, a locality should be ecologically defined. That is, a scale should be chosen within which all species of interest are interspersed and can potentially interact and within which the local environment is homogeneous. Since different species interact at different scales, and environments that are homogeneous for one species may be heterogeneous for others, the ideal locality will probably have indistinct boundaries. It may circumscribe the number of herbivore species associated with a host plant species, the number of insectivorous birds found in a patch of woods, or the number of flowering plant species associated with a serpentine outcrop. Local samples must be standardized in some way so that they are comparable among regions. Typically, the number of individuals or the unit area assessed is uniform, or local assemblages are sampled until the species/sampling curve levels off.

If regional richness is sufficiently high to engender saturation, it is not patently clear at which scale it should occur. Communities are hierarchically structured assemblages of organisms that may be saturated at one scale and not other scales because the intensity of dynamic coupling among species varies with the unit of study. Most comparisons of local and regional richness offer few clues about the relationship between hierarchical structure and saturation because local

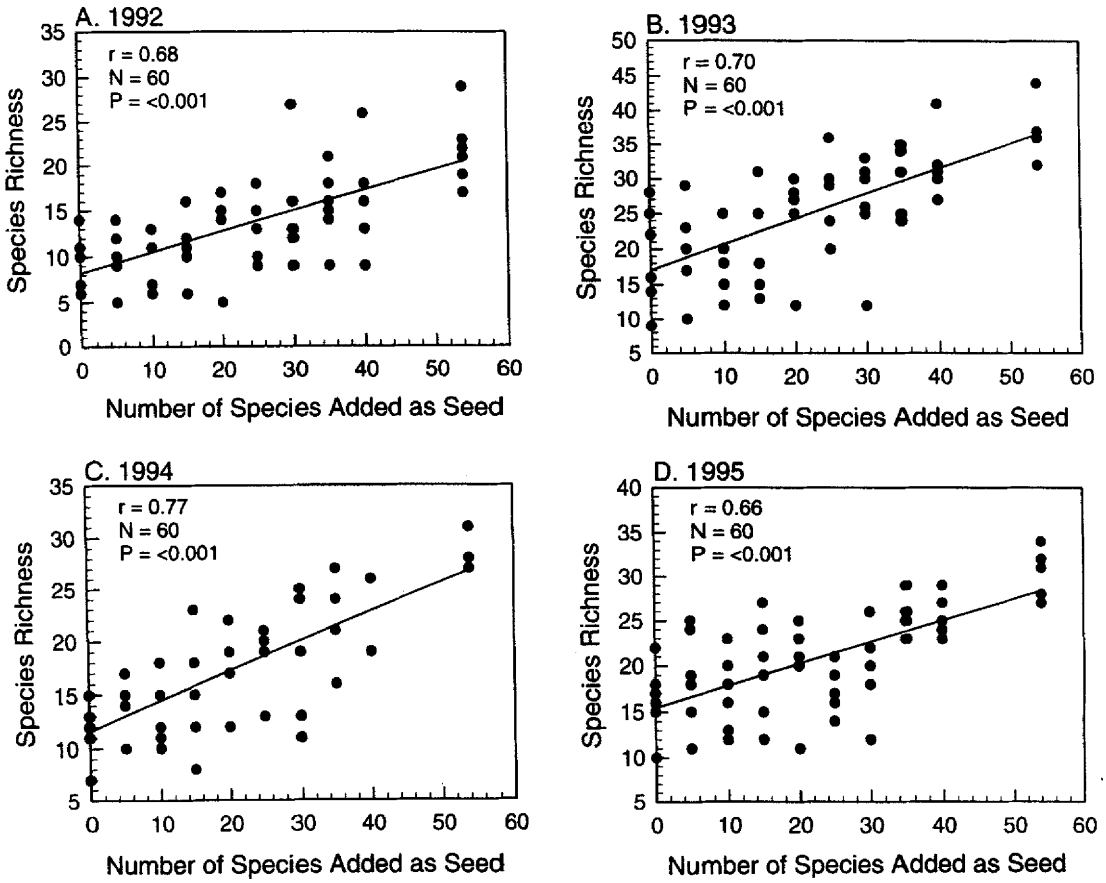


Figure 4 Species richness of vascular plants in 0.25-m^2 prairie plots (1-m^2 plots in 1995) after the experimental addition of seeds in 1992. Seed additions ranged from 0 to 54 species and the plots were followed for 4 years (1992–1995). The 0 species addition (data on the Y-axis) shows changes in species richness in plots not subject to manipulation. Species richness increased steadily over the entire range of species additions and the increases persisted for at least 4 years. The results suggest that these plant communities are highly recruitment limited and not saturated with species. Reproduced from Tilman D (1997) Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* 78: 81.

richness is evaluated at only one scale. Clearly, a more methodological exploration of hierarchical structure that involves multiscale sampling is required. A few studies have made limited attempts at a multi-scale analysis. Caley and Schluter (1997) measured local richness from distributional data of 10 different taxa at two local scales (1% and 10% of regional area) in various regions. All taxa fell along a Type I trajectory at both scales (**Figure 5**). In a 1996 paper Kennedy and Guégan showed that parasite communities reached a hard upper limit of four species in individual eels whereas parasite richness in the eel population as a whole continued to increase steadily with regional richness. Finally, data on coral assemblages examined by Karlson and Cornell (2002) showed that local richness was less sensitive to regional richness in 1-m^2 quadrats than in 10-m line transects. In the latter example, the 1-m^2 quadrat is so small relative to the size of some coral colonies that local richness, particularly in very rich regions, could be limited simply by the number of individuals that can be packed into this small sample. In all three examples, only two local scales were examined, but for two of these examples, this was sufficient to show differential responses of local richness to changes in regional richness. Clearly, multiscale analyses will permit stronger inferences regarding saturation, or lack thereof, in communities. They will also be useful in defining the spatial arena within which local processes, which might in some cases limit local richness, should be sought.

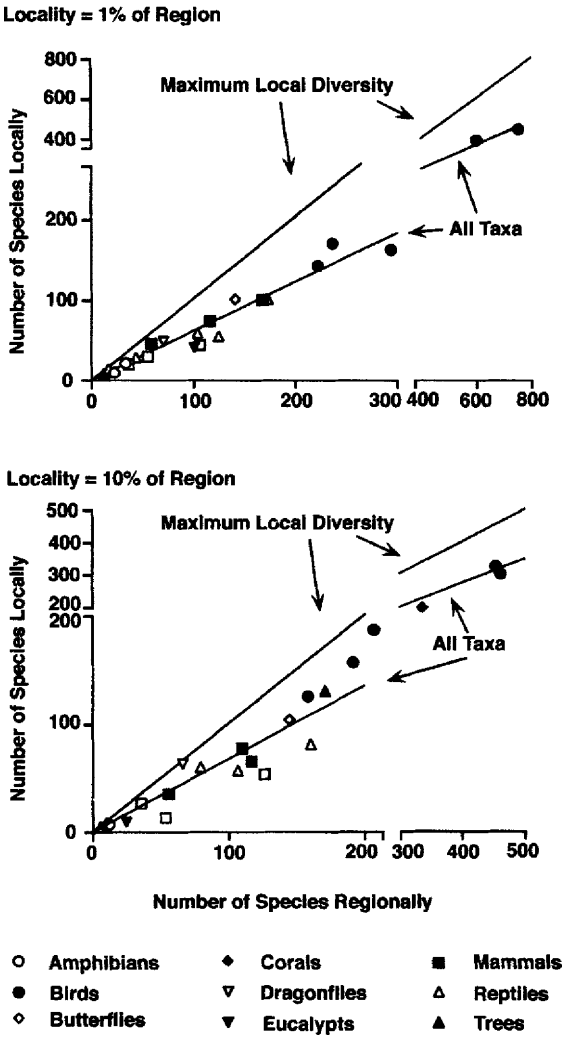


Figure 5 Local versus regional richness relationships at two local scales (1% and 10% of regional area) for 23 different taxa that have been combined for this analysis. The relationship is Type I at both scales, with the main difference being that the slope of the curve is roughly 10% lower at the smaller scale, consistent with the species–area relationship. The size of the locality has no effect on the tendency for saturation over this range of spatial scales. Reproduced from Caley MJ and Schluter D (1997) The relationship between local and regional diversity. *Ecology* 78: 70.

Synthesis of Community and Regional Perspectives

A Thought Experiment

Since processes operating at scales larger than the local habitat (e.g., landscape or region) almost surely affect local richness, research emphasis must shift to these larger-scale processes. Only then can patterns of community assembly be completely understood. However, this shift does not imply that local processes should be ignored. Local- and larger-scale processes may both affect local richness and their relative influence must be determined. The advantages of this multi-scale perspective can be illustrated with a thought experiment. Suppose we have two climatically similar regions with different regional richness due to historical factors. The Asian and European temperate forests discussed previously provide a good example. Suppose further that physically identical habitats of two types (e.g., sandy versus loamy soils) were transplanted into these regions and were allowed to achieve a steady state that showed no influence of initial colonization history. Once the steady state was reached, by how much would per locality tree species richness vary among habitat types and regions? The transplanted habitats would make it possible to partition the contribution to per locality richness among processes operating at local and regional spatial scales. The design would also make it possible to evaluate interactions between regional and local factors and to examine mediation of species interactions by changes in the local environment.

Manipulations of Productivity and the Pool

Although the preceding experiment is possible in principle, it is likely to be difficult to carry out. Also, it is difficult to be sure that climatic conditions that may affect local processes are truly identical among regions. An alternative to habitat transplantation is a two-factorial manipulation of the species pool and local environmental conditions in the same region. For example, suppose some high-productivity lakes in a particular region had unusually low fish species richness. Is the low richness due to interaction mediation or pool limitation? Manipulations could be designed in which the entire fish species pool could be added to lakes of different productivity. In addition, productivity itself could be manipulated by adding fertilizer to particular lakes. Care would have to be taken that the fish were adapted to live at the prevailing productivity levels. A full two-factorial design would make it possible

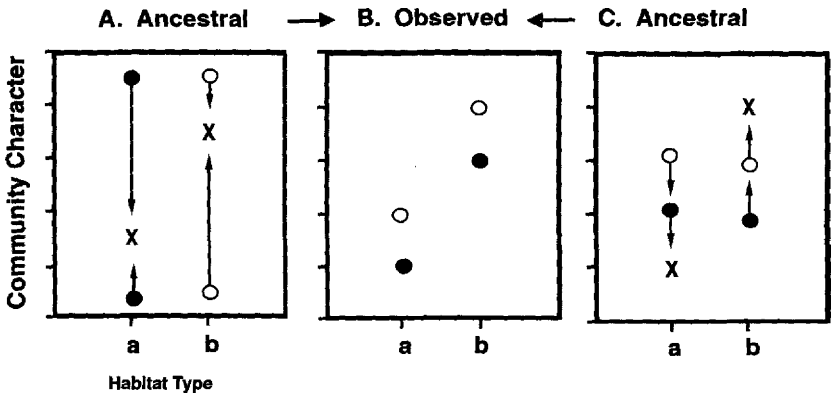


Figure 6 Hypothetical changes in communities over time showing that differences in the community due to habitat characteristics can be used to estimate community convergence. The community character in this case is species richness. (A) X indicates the richness value toward which the community will converge because of habitat characteristics. Open circles represent some unusually rich ancestral state of the communities in both habitats because, for example, they are located in a rich region and have been subject to a massive colonization event. Solid circles represent an alternative, depauperate ancestral state, because, for example, the communities are located in a species-poor region and have been subject to a disturbance. (B) Over time, communities will converge on X and the variance components due to habitat and region can be used to estimate the relative contribution of habitat and regional effects on local richness. (C) A slightly different case, but the inferences are the same. Reproduced from Schluter D (1986) Tests for similarity and convergence of finch communities. *Ecology* 67: 1073.

to sort out pool effects, productivity effects, and pool X productivity interactions on local richness (*see* Two Perspectives on the Control of Community/Regional Diversity). If regional control of community structure is important, then pool effects or pool X productivity interactions should have important effects on local richness.

A Comparative Approach

In situations where experiments are intractable or when it is desirable to examine longer-term effects, the problem can be addressed using a statistical comparative approach. For example, Schluter and Ricklefs (in Ricklefs and Schluter, 1993) have suggested using multi factorial methods to partition the variance in local richness into components that can be explained by habitat differences within regions and species richness differences among regions. To the extent that local environments constrain local richness, species richness in similar habitats from different regions should converge relative to some hypothetical ancestral state of the community (Figure 6). Detecting convergence is thus equivalent to measuring the variance component due to habitat differences. Important regional effects on local richness will be detectable in the variance component due to differences in regional richness. Schluter and Ricklefs calculated variance components for 36 different studies and found strong convergence in many cases, consistent with significant local effects on local richness. Yet there were strong regional differences in richness within similar habitats as well. Using regression methods to study coral reef communities, Karlson and Cornell (1998) similarly found both local and regional effects to be important.

Conclusions

The results from experimental and statistical analyses confirm that both local and regional processes have important effects on local richness. Thus, not only must experimental work on local communities be continued, but investigations need to be expanded to encompass larger-scale processes as well. The Volterra–Gause image of an ecological system at equilibrium and independent of historical processes has undeniable appeal, and has engendered much productive research on community assembly. Nevertheless, integration of historical, biogeographical, and metacommunity processes with local interactions and environmental variation is essential for a complete understanding of richness variation in species assemblages.

See also: Biogeography, Overview. Diversity, Molecular Level. Diversity, Organism Level. Energy Flow and Ecosystems. Landscape Diversity. Temperate Forests

References

Caley MJ and Schluter D (1997) The relationship between local and regional diversity. *Ecology* 78: 70.

Cornell HV and Lawton JH (1992) Species interactions, local and regional processes, and the limits to richness in ecological communities. *J. Animal Ecol.* 61: 1.

Hugueny B and Cornell HV (2000) Predicting the relationship between local and regional species richness from a patch occupancy dynamics model. *J. Animal Ecol.* 69: 194.

Karlson RH and Cornell HV (1998) Scale-dependent variation in local vs. regional effects on coral species richness. *Ecol. Monogr* 68: 259.

Karlson RH and Cornell HV (2002) Species richness of coral assemblages: Detecting regional influences at local spatial scales. *Ecology* 83(2): 452–463.

Lawton JH (1999) Are there general laws in ecology? *Oikos* 84: 177.

MacArthur RH (1972) *Geographical Ecology: Patterns in the Distribution of Species*. New York: Harper and Row.

McPeek MA and Brown JM (2000) Building a regional species pool: Diversification of the *Enallagma* damselflies in eastern North American waters. *Ecology* 81: 904.

Morton RD and Lau R (1997) Regional species pools and the assembly of local ecological communities. *J. Theoretical Biol.* 187: 321–331.

Ricklefs RE and Schluter D (eds.) (1993) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. Chicago: University of Chicago Press.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, United Kingdom: Cambridge University Press.

Srivastava DS (1999) Using local–regional richness plots to test for species saturation: Pitfalls and potentials. *J. Animal Ecol.* 68: 1.

Tilman D and Kareiva P (eds.) (1997) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*. Princeton, NJ: Princeton University Press.

Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea

Matthew Church, University of Hawaii at Manoa, Honolulu, HI, USA
Daniela Böttjer, University of Hawaii, Honolulu, HI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biological carbon pump Describes the biologically mediated transfer of carbon from the euphotic zone to the ocean's interior, maintaining a vertical gradient in inorganic carbon concentrations. Major processes involved in the pump include photosynthetic consumption of inorganic carbon and production of organic matter in the euphotic zone, gravitational settling of particulate organic matter, and subsequent remineralization of organic matter back to inorganic carbon at depth in the ocean.

Chemoorganoheterotrophy A form of metabolism that relies on oxidation of organic compounds for energy, reducing equivalents, and carbon.

Diazotroph Free-living or symbiotic bacteria or archaea that are capable of reducing N_2 to a biologically utilizable form of nitrogen (N), ammonia (NH_3).

Dinitrogen (N_2) fixation Dinitrogen (N_2) is the most abundant form of nitrogen (N) in the atmosphere. It is unavailable to most microorganisms because it consists of a stable $N \equiv N$ triple bond, and breaking this molecule requires substantial energy. Only selected microorganisms, termed "diazotrophs," have evolved the capacity to fix N_2 that is the conversion of N_2 to ammonia (NH_3), a biologically more available source of nitrogen (N) to many microorganisms. N_2 fixation is catalyzed by the enzyme nitrogenase and is an energy-expensive process; 16 ATP molecules and eight reducing equivalents are required to reduce N_2 to 2 NH_3.

Euphotic zone Is the uppermost or "well-lit" layer of the ocean in which there is sufficient light energy available to support net photosynthetic growth. The euphotic zone extends to a depth where light intensity falls from 0.1% to 1% of that at the surface and in the open sea generally varies from <50 m to ~200 m.

Heterocyst The site for N_2 fixation in some species of filamentous cyanobacteria. A heterocyst is a differentiated cell that may be present terminally or intercalary. It lacks photosystem II, but photosystem I is present, providing a means to generate additional energy (ATP) required to fuel N_2 fixation.

New versus regenerated production Depending on the source of nitrogen (N), primary production in the sea can be divided into "new" and "generated" production. Production maintained by assimilation of new sources of N to the euphotic zone, such as nitrate (NO_3^-) or N_2, is referred to as "new production," whereas production derived from recycling of nutrients within the euphotic zone such as via the uptake of ammonia (NH_4^+) or dissolved organic N is termed "regenerated production."

Nitrogenase A metalloenzyme complex that catalyzes the process of N_2 fixation. The nitrogenase enzyme consists of two highly conserved protein subunits: An iron protein (Fe) known as "dinitrogen reductase" and a molybdenum–iron protein (MoFe) known as "dinitrogenase." Nitrogenase is sensitive to the presence of oxygen (O_2), becoming irreversibly inhibited at high O_2 concentrations.

Photolithoautotrophy A form of metabolism that depends on light as an energy source, inorganic compounds (e.g., H_2O, H_2S) for reducing equivalents, and CO_2 as a source of carbon.

Photoorganoheterotrophy A form of metabolism that utilizes light as a source of energy and organic compounds as sources of both reducing power and carbon.

Introduction

Earth's largest ecosystems are found in the sea, and these massive habitats comprise vast reservoirs of planktonic biodiversity. The combined metabolic activities of diverse planktonic communities make the oceans globally significant bioelemental reactors. Numerous selective forces collectively shape extant microbial diversity, and the emergence of specific traits offer glimpses into successful strategies for adapting to life in the sea. The wealth of biodiversity found in contemporary marine ecosystems reflects a long history of evolutionary adaptations, and some of the best examples of this selective adaptation can be found among microorganisms catalyzing the nitrogen (N) cycle. N is an essential element for life required for biosysnthesis of nucleotides, amino acids, proteins, and nucleic acids, but bioavailable forms of N are often found in vanishingly low concentrations (nanomolar) in the euphotic zone of the open sea. In these habitats, biological demand for N often outpaces its supply, and the availability of N restricts planktonic growth and limits the capacity of these ecosystems to support biomass. Ironically, both the atmosphere and oceans hold a nearly inexhaustible reservoir of N in the form of dinitrogen (N_2). However, only a select group of microorganisms termed "diazotrophs" are capable of utilizing this pool of N as a nutrient resource. By converting N_2 to ammonia (NH_3), these microorganisms constitute a key microbial functional guild: the metabolic activities of these organisms supply new N to otherwise N-depleted ecosystems, making them fundamental to plankton ecology and ocean biogeochemistry.

Microorganisms catalyze nearly all of the major processes in the N cycle. The ability of various N-containing compounds

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00403-2

to accept or donate electrons situates N as an important participant in numerous biologically catalyzed oxidation–reduction reactions. The N cycle includes processes that utilize N as a nutrient resource in support of cellular biomass production (assimilatory processes) and those that rely on N compounds as either electron donors or acceptors (dissimilatory processes). All of these processes collectively interact to govern the sizes of N inventories and the movement of N from one pool to another in the sea. Among the assimilatory N cycle processes, only N_2 fixation is capable of introducing new bioavailable N to an ecosystem. Globally, biological N_2 fixation is estimated to supply ~250 Tg N per year, with ~50% of that input occurring in the open ocean (Gruber and Galloway, 2008). Introduction of new N via N_2 fixation is countered by various removal processes that result in production of the gaseous products N_2 or N_2O; these removal processes include denitrification (dissimilatory reduction of nitrate and anaerobic ammonia oxidation) and to a lesser extent nitrification (via production of nitrous oxide).

Marine N_2 fixers appear most abundant in tropical and subtropical ecosystems where upper ocean temperatures typically vary between ~18 °C and 30 °C and concentrations of fixed N are low relative to other bioessential nutrients. In the nutrient-poor regions of the upper ocean, activities of N_2-fixing microorganisms provide an important source of N fueling plankton production. N_2-fixing microorganisms are both responsive to and modifiers of their environment. In addition, the process of N_2 fixation is highly regulated at the genetic level and can respond rapidly to fluctuations in the environment. Intriguingly, diazotrophs are frequently found at very low population abundances relative to other microorganisms growing in these waters (Figure 1). Nonetheless, through their ability to control the availability of an essential limiting resource, diazotrophs constitute a keystone functional role in various regions of the ocean. Introduction of fixed N to an otherwise N-starved ecosystem results in a cascade of responses, altering food web ecology, biomass productivity and standing stocks, and ultimately the carbon (C) sequestration potential of the ecosystem. The fate of N introduced by N_2-fixing microorganisms and its consequences on ocean biogeochemistry have been active areas of research for decades, but much less is known about the ecology and diversity of the organisms that underpin the process. The vast majority of ocean habitats are remote and largely inaccessible, resulting in chronic undersampling of marine diazotroph habitats.

The marine N_2-fixing microorganisms we understand best are those most easily recognized under a microscope and amenable to isolation and cultivation. Much of our understanding of marine diazotroph physiology has developed based on laboratory studies of cyanobacteria belonging to the genera *Nostoc*, *Trichodesmium*, *Crocosphaera*, and *Nodularia*. However, the diazotroph guild is populated by diverse groups of microorganisms with equally diverse metabolic capabilities. Many of these organisms have no cultivated representatives from which to build our understanding of their physiological behavior, metabolic flexibility, or sensitivity to perturbation. Like many major advances in microbial ecology over the past decade, the study of marine diazotrophs has been greatly facilitated through fusion of cultivation-independent molecular biology techniques with biogeochemical rate measurements.

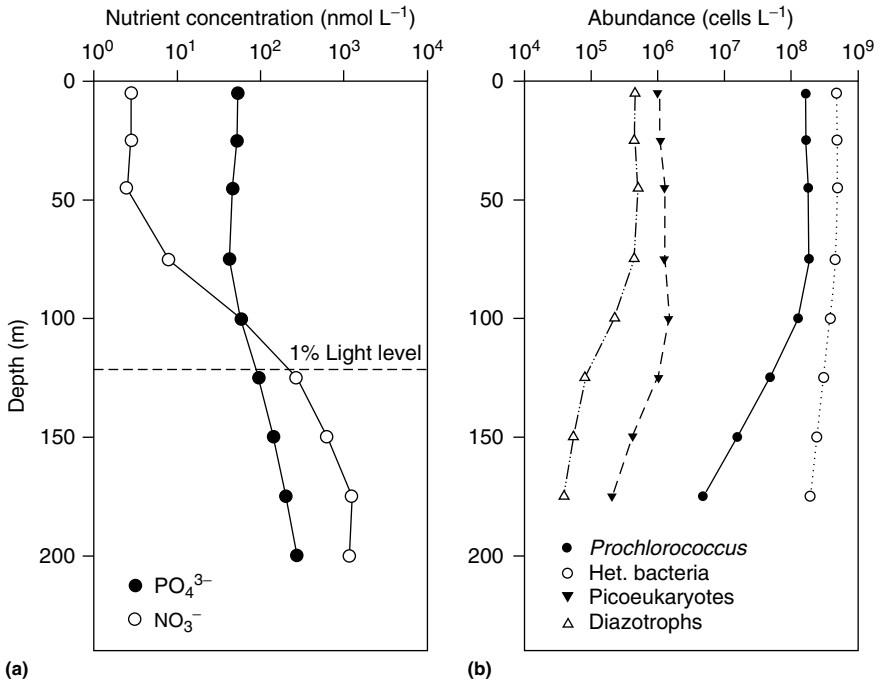


Figure 1 Habitat characteristics of the euphotic zone in the oligotrophic North Pacific Ocean. Vertical distributions of inorganic nutrients (NO_3^- and PO_4^{3-}, (a)) and abundances of prominent groups of planktonic microorganisms (b); also depicted is the depth to which 1% of the incident irradiance penetrates (dashed line in (a)). Diazotroph abundances estimated based on quantitative PCR-based determinations of *nifH* gene copies. Het. Bacteria refers to presumed heterotrophic picoplankton.

Perhaps nowhere has the marriage of these disparate tool kits yielded more unique insights than to our understanding of the microbes driving the N cycle. Such work has begun to transform ecological studies of marine microorganisms, allowing mapping of organisms onto processes, and providing new information on rates of biogeochemical transformations mediated by specific microorganisms (Zehr and Kudela, 2011). In this chapter, we highlight the importance of biological N_2 fixation as a key determinant of biogeochemical cycling and plankton biology, with a focus on describing some of the recent advances on the diversity and ecology of N_2-fixing microorganisms in the open sea.

Biodiversity of N_2-Fixing Microorganisms in the Sea

Marine diazotrophs are morphologically, genetically, and metabolically diverse. They have been reported to occur in marine habitats ranging from sea ice to warm tropical seas, and to occur as both free-living organisms or associated with other organisms through various forms of symbiotic relationships. In the oceans, these organisms appear most abundant throughout the low-nutrient, high-light regions of the subtropical ocean gyres that include ~40% of the surface area of the planet. Much of our understanding of marine diazotroph diversity has derived from advances along several avenues of research that include application of cultivation-independent molecular techniques, biogeochemical rate measurements at the level of single cells to whole communities, and cultivation-dependent cell isolations and manipulative laboratory-based model diazotroph systems.

Nitrogenase and *nifH*

The capacity for N_2 fixation is widely distributed among members of bacteria and archaea; to date, there are no reports of N_2 fixation by eukaryotes. The maintenance of diverse microbial lineages performing N_2 fixation may reflect the importance of this process to ocean ecosystems. In an environment where fluctuations in habitat conditions are prevalent and comprise important selective forces, functional redundancy may ensure that critical ecosystems services are retained in the face of variations in population structure.

N_2-fixing microorganisms utilize the enzyme complex nitrogenase to catalyze the reduction of N_2 to NH_3; the reaction can be described as

$$N_2 + 8\,H^+ + 16ATP + 8e^- \rightarrow 2NH_3 + H_2 + 16ADP + 16PO_4^{3-}$$

The conventional nitrogenase enzyme consists of two transiently interacting protein clusters: a tetraheterodimer molybdenium–iron containing protein called "dinitrogenase" and a homodimer iron–sulfur protein called "dinitrogenase reductase." Some organisms may also possess alternative nitrogenases in which molybdemium in dinitrogenase is substituted by vanadium, or, in other cases, this protein contains iron only. Dinitrogenase appears to be the substrate-binding protein, whereas the role of dinitrogenase reductase is to facilitate the transfer of electrons needed to reduce dinitrogenase. Nitrogenase is rapidly inactivated in the presence of oxygen (O_2), and microorganisms that catalyze the reaction have evolved various strategies for establishing anoxic conditions necessary to protect the enzyme from O_2. The high activation energy required to break the triple bond in N_2 and the subsequent reduction of N_2 to NH_3 are energy- and reductant-demanding processes. Sixteen moles of ATP and eight electrons are required for every 2 mol of NH_3 produced by the reaction. The evolution of hydrogen (H_2) by the reaction represents a potential reductant sink; however, many N_2-fixing microorganisms utilize hydrogenases to internally recycle the H_2. The key role of iron (Fe) as a co-factor in this enzyme system, together with low supply of Fe to large regions of the sea, has led to the suggestion that Fe might be a key element limiting N_2 fixation in many ocean ecosystems.

The nitrogenase enzyme complex is encoded by a suite of genes in the *nif* operon; dinitrogenase is encoded by *nifD* and *nifK* genes, whereas dinitrogenase reductase is encoded by the *nifH* gene. The process of N_2 fixation is highly genetically regulated through transcriptional and translational modification of the *nifHDK* gene products. Such strict regulation has been hypothesized to minimize energy losses incurred if the gene product was translated to a functional protein under conditions not conducive to N_2 fixation – for example, in the presence of O_2 (Zehr and Paerl, 2008). The evolutionarily conserved nature of the *nifH* gene has been capitalized on for its utility as a molecular target for studying N_2-fixing microorganisms. Analyses of the phylogeny of *nifH* genes reveal four deeply branching gene lineages (**Figure 2(a)**; Zehr *et al.*, 2003). One of these lineages (termed Cluster IV) derives from *nif*-related gene families, whereas the remaining three clusters represent *nifH* genes from widely diverse groups of prokaryotes. Among the Cluster I *nifH* genes are representatives of both the conventional molybdenum-containing nitrogenases and selected alternative vanadium-containing nitrogenases. Cluster II contains microorganisms encoding the alternative nitrogenases (those containing Fe only) as well as selected groups of N_2-fixing archaea. Cluster III includes numerous and diverse anaerobic *nifH*-containing microorganisms. The vast majority of cultivated *nifH*-containing marine microorganisms are found among Cluster I; this cluster includes the cyanobacteria as well as a large proportion of N_2-fixing proteobacteria (including the α-, β-, γ-, and ε-proteobacteria).

Despite the value of metagenomic techniques to identify physiological and metabolic capabilities among environmental microorganisms, these approaches have provided only limited information on marine diazotrophs. In large part this stems from the low population abundances of these organisms, making them highly underrepresented in these large-scale sequencing efforts (Johnston *et al.*, 2005). However, *nif* genes have been retrieved (albeit in very low numbers) from selected metagenomic sequence libraries from ocean plankton and analyses of these sequences have provided new insights into evolutionary strategies and selective adaptations by marine N_2 fixers. Metagenomic sequence analyses from naturally occurring populations of *Crocosphaera* in both the Pacific and Atlantic Oceans led to the conclusion that these diazotrophs demonstrate remarkable gene conservation (Zehr *et al.*, 2007), a finding that contrasts studies on more-abundant non–N_2 fixing oceanic cyanobacteria (e.g., *Prochlorococcus*). These results suggest oceanic cyanobacterial diazotrophs may

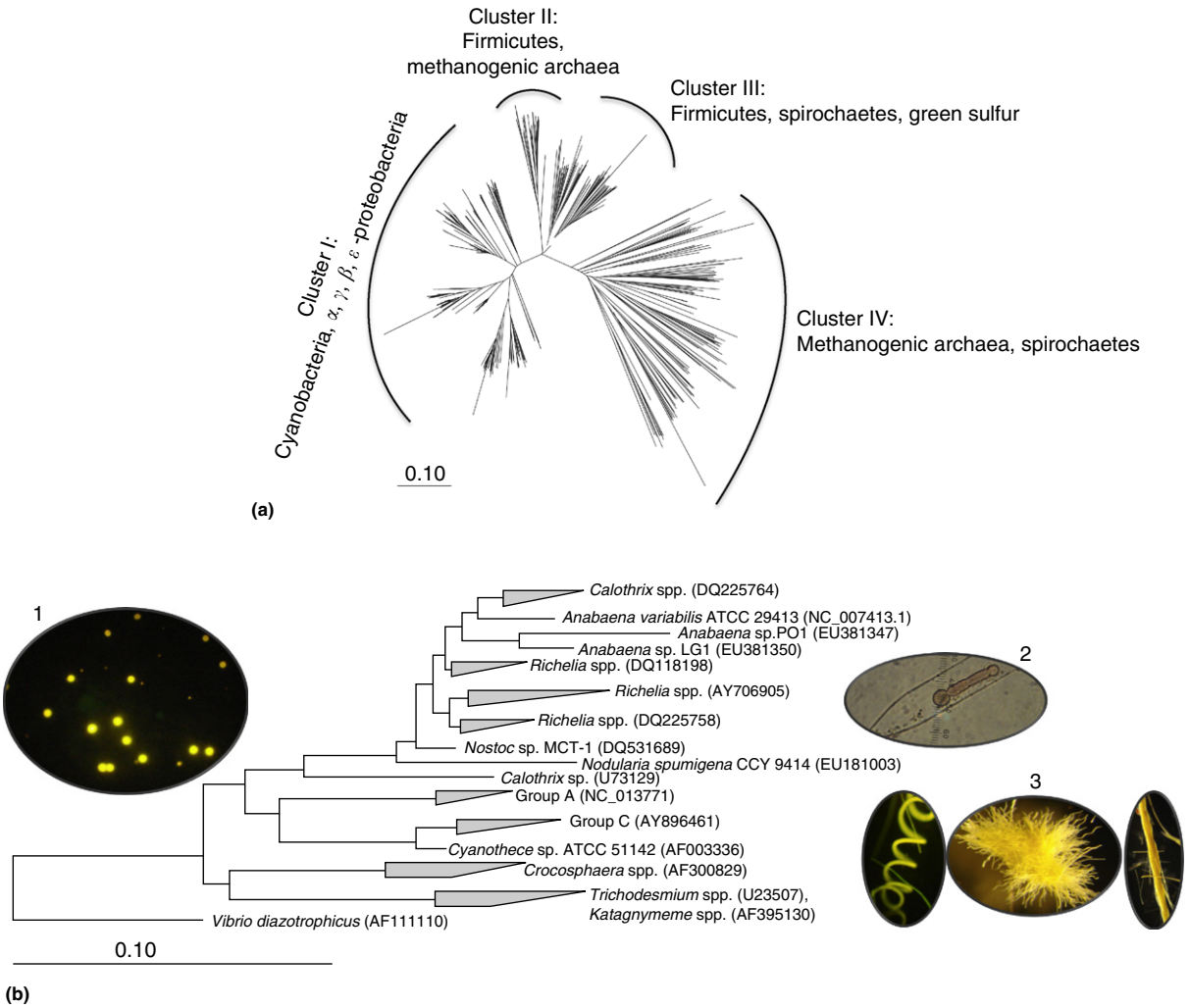


Figure 2 Phylogenetic relationships among nitrogenase (*nifH*) genes (a). Representative groups of bacteria and archaea from each of the deeply branching clusters (I–IV) of genes are shown. Phylogenetic relationships and microscopy images of several major groups of *nifH*-containing cyanobacteria common to the subtropical oceans (b). Images include *Crocosphaera* (1), endosymbiotic heterocystous *Richelia* sp. shown inside the diatom *Rhizosolenia* sp. (2), and three morphologies of organisms belonging to the genus *Trichodesmium* (3). Microscopy images courtesy of Angelique White (Oregon State University).

undergo different selective pressures than their more abundant relatives and rely on different evolutionary strategies for survival.

Many of the advances in our understanding of the diversity of N_2-fixing microorganisms have derived from polymerase chain reaction (PCR) amplification of *nifH* genes from naturally occurring plankton communities. This approach has yielded sequence information, enabled new approaches to studying *nifH* diversity, and provided tools to quantify *nifH* gene abundances and gene transcripts in the sea. In large part as a result of PCR amplification, cloning, and sequencing efforts, there are currently more than 36,000 publicly available *nifH* gene sequences. This wealth of sequence information provides insight into numerous aspects of diazotroph diversity, including understanding how *nifH* genes have become so widely distributed among different groups of microorganisms and tailoring of molecular-based approaches to studies of specific groups of *nifH*-containing microorganisms.

Phenotypic and Genetic Diversity of Marine N_2 Fixers

Marine cyanobacteria are morphological diverse, including both unicellular and colonial organisms ranging in size across more than three orders of magnitude ($\sim$1 μm to $>$1 mm, **Table 1**). The ability to enact selective pressures as part of cultivation strategies (namely, N-free media and light) has yielded isolated representatives from many of the dominant lineages of oceanic N_2-fixing cyanobacteria. Arguably the best-studied marine cyanobacteria are members of the genus *Trichodesmium*; these non-heterocystous cyanobacteria exhibit a wide range of cell morphologies, growing as either single filaments (termed "trichomes") or colonial aggregates of filaments (Capone and Carpenter, 1982). On average, each trichome contains $\sim$100 individual cells, with each colony comprised of 100–200 trichomes (Letelier and Karl, 1996). *Trichodesmium* abundances vary widely (**Table 1**), with abundances occasionally very high ($>10^5$ cells L^{-1}) during large

Table 1 Characteristic features of various marine diazotrophs commonly found in tropical and subtropical ocean ecosystems

Organism	*Cultivated representatives*	*Size range*	*Abundance (cells L^{-1})*	*Notable physiological features*
Filamentous cyanobacteria				
Trichodesmium spp.	*T. contortium, T. erythraeum, T. Hildebrandt, T. T. tenue, T. thiebautii K. spiralis, K. pelagica*	~10 μm–1 mm	ND–>10^5	Non-heterocystous, photolithoautotroph, bloom-forming, daytime N_2 fixer, buoyancy control
Richelia spp.	*R. intracellularis*	~10–40 μm	ND–>10^6	Heterocyst-forming, photolithoautotroph, bloom-forming, endosymbiont of diatoms *Hemiaulus* spp. and *Rhizosolenia* spp., daytime N_2 fixer
Calothrix spp.	*C. rhizosolenia*	~15–30 μm	ND–10^3	Heterocyst-forming, photolithoautotroph, bloom-forming, epibiont of diatoms *Chaetoceros* sp., daytime N_2 fixer
Unicellular cyanobacteria				
Group A	Uncultured	<1 μm[a]	10^3–10^5[b]	Photoorganoheterotroph, presumed daytime N_2 fixer, lacks photosystem II, possible symbiont
Crocosphaera spp.	*Crocosphaera watsonii* (WH0001, WH0002, WH0003, WH0004, WH0005, WH0006, WH0401, WH0402, WH8501, WH8504)	2–8 μm	10^3–10^6	Photolithoautotroph, free-living, nighttime N_2 fixer
Group C	TW3	4–6 μm	ND–>10^5[b]	Photolithoautotroph, free-living, nighttime N_2 fixer
Non-cyanobacteria	*Vibrio diazotrophicus, Klebsiella, Pseudomonas*	<1 μm	~10^2–10^3[b]	Chemoorganoheterotrophs, obligate and facultative anaerobes

[a]Size estimated based on flow-cytometric cell sorting and size-fractionated QPCR results.
[b]Abundances estimated based on QPCR determinations of *nifH* gene copies, assuming one gene copy equals one cell.

blooms. Intensive cultivation efforts have yielded numerous laboratory isolates of *Trichodesmium*, including *T. thiebautti*, *T. erythraeum*, *T. tenue*, *T. hildebrandt*, and *T. contortium*. The genome of *T. erythraeum* also has been sequenced (7.8 Mega bases; Mb). In addition, organisms previously identified as belonging to the genus *Katagyneme* (including *Katagyneme spiralis* and *Katagyneme pelagic*) are genetically so similar to *Trichodesmium* that it appears these organisms should in fact be considered members of this genus (Lundgren *et al.*, 2005), providing additional evidence of the phenotypic plasticity exhibited by *Trichodesmium*.

Another prominent assemblage of marine N_2 fixers includes various heterocyst-forming cyanobacteria. These filamentous cyanobacteria form 5–10 cell chains of vegetative (photosynthetically active) cells terminally bound by nonvegetative (lacking photosystem II) cells called "heterocysts." The heterocyst forms the active site of N_2 fixation; the thick glycoprotein cell wall acts as a diffusion barrier for O_2, thereby protecting nitrogenase. Compared to freshwater or brackish waters, oceanic heterocystous cyanobacteria are relatively rare, but these organisms have been known as important contributors to open ocean N cycling for many years (Venrick, 1974; Carpenter *et al.*, 1999). Several distinct groups of oceanic heterocytous diazotrophs have been described, including *Richelia intracellularis* and *Calothrix rhizosoleniae*, and the abundances of these organisms is highly variable, occasionally exceeding 10^6 cells L^{-1} during blooms (**Table 1**). Notably, these cyanobacteria are frequently observed in symbiotic relationships with various genera of oceanic diatoms. Among the best studies of these interactions are those involving *R. intracellularis* and diatoms belonging to the genera *Rhizosolenia* and *Hemiaulus*. *Richelia* forms intracellular symbiotic partnerships with these diatoms, providing fixed N to the hosts. However, it is unknown whether the cyanobionts benefit from these interactions, but such benefits might include exchange of nutritional or energy-demanding substrates, buoyancy regulation by the host, or protection from predation (Carpenter, 2002). Unlike the endosymbiotic relationships between *Richelia* and *Rhizosolenia* or *Hemiaulus*, *Calothrix* are often observed attached as an epiphyte to the spaces between chains of diatoms belonging to the genus *Chaetoceros*. *C. rhizosoleniae* has been isolated in culture without the host diatom, indicating the symbiosis is not obligate (Foster *et al.*, 2010). Moreover, consistent with these results, the genome of the isolate is relatively large (6.6 Mb) compared to obligate symbionts. Genetic analyses, based on various functional genes and ribosomal RNA genes, suggest that different groups of diazotroph symbionts form specific associations with the

diatom hosts. Such symbiotic interactions do not appear strictly limited to heterocystous cyanobacteria; the oceanic diatom *Climacodium frauenfeldianum* also hosts presumed N_2-fixing unicellular cyanobacteria (Carpenter and Janson, 2000). How these symbioses develop and are maintained over generations remains largely unknown. Moreover, it remains unknown to what extent these symbioses have shaped the co-evolution of the diazotrophs and their hosts.

Unicellular organisms appear to comprise some of the most abundant and geographically dispersed groups of the marine N_2-fixing cyanobacteria (Moisander *et al.*, 2010). Research into the diversity and ecology of these microorganisms has been greatly facilitated by both cultivation-dependent approaches and application of various genetic tools. Initial surveys of oceanic *nifH* genes revealed the presence of sequence types that clustered phylogenetically among unicellular N_2 fixers. Subsequent studies revealed that the *nifH* sequence of one of these phylotypes was identical to *Crocosphaera*, a unicellular cyanobacterium initially isolated in the mid-1980s but which had been largely overlooked in its potential contributions to marine N_2 fixation (Zehr *et al.*, 2001). Members of the genus *Crocosphaera* are readily identified based on both their spherical morphology, size (2–8 μm in diameter), and brightly fluorescent characteristics of the pigment phycoerythrin. *Crocosphaera* abundances are often reported between $\sim 10^3$ and 10^4 cells L^{-1}, with occasional blooms exceeding 10^6 cells L^{-1} (**Table 1**). Several isolates of *Crocosphaera* now exist, and the genome (6.2 Mb) of *Crocosphaera watsonii* WH8501 has been sequenced. The strains exhibit some notable differences in morphology and physiologies, despite low apparent genetic variability. For example, examination of 10 strains of *Crocosphaera* indicated that, despite minimal differences in rRNA genes and internal transcribed spacer regions, the strains varied in size, temperature optima for growth (ranging between 22 and 36 °C), and production of extracellular organic material (Webb *et al.*, 2009).

Subsequent studies in the Atlantic Ocean revealed the presence of *nifH* genes closely related to *Cyanothece* ATCC51142 (Langlois *et al.*, 2005) belonging to a N_2-fixing cyanobacteria termed "group C" (Foster *et al.*, 2007). The *nifH* gene abundances of these organisms were subsequently examined in the waters of the tropical Atlantic, suggesting that on occasion this organism could be relatively abundant ($\sim 10^5$ cells L^{-1}). Cultivation efforts in the western Pacific yielded heavily pigmented, 4–6-μm diameter cells whose *nifH* and 16 S rRNA genes cluster closely among group C and *Cyanothece*, respectively. These isolated aerobic diazotrophs (strain name TW3) actively fix N_2 at night (**Table 1**) and grow most rapidly at water temperatures $\sim$30 °C, suggesting these organisms may be largely restricted to tropical waters (Taniuchi *et al.*, 2011).

A third group of unicellular N_2-fixing microorganisms, termed "group A," was first identified based on *nifH* gene sequences. The *nifH* gene sequences of group A are also most closely related to the marine unicellular cyanobacterium *Cyanothece* sp. strain ATCC 51142. However, to date, no representatives of group A microorganisms have been isolated in culture. Information on these microorganisms mostly derives from PCR- and reverse-transcriptase PCR-dependent (including both 16 S rRNA and *nifH* genes) methodologies to examine their distributions and patterns of *nifH* gene expression. Quantitative PCR (QPCR) based surveys revealed that the group A phylotype is often the most abundant ($\sim 10^5$ *nifH* gene copies L^{-1}) and widely distributed of the N_2-fixing cyanobacteria. A major breakthrough in our understanding of these organisms came from targeted flow cytometric cell sorting and subsequent metagenomic analyses of the sorted cell populations (Zehr *et al.*, 2008). This approach provided several important pieces of information about these uncultivated organisms: (1) they are small ($\sim$1 μm in diameter), (2) possess a highly reduced genome (1.44 Mb), and (3) lack a number of fundamental biosynthetic pathways (Zehr *et al.*, 2008; Tripp *et al.*, 2010). Notably absent from the genome of group A were genes required for inorganic C fixation, photosystem II, genes necessary for synthesis of purines and several amino acids; and nearly all the major genes required for a functional tricarboxylic acid cycle (Tripp *et al.*, 2010). The lack of these basic biosynthetic pathways suggests group A are likely symbionts, providing fixed N to hosts in exchange for the supply of fixed C substrates required for basic cellular biosynthesis.

The prevalence of different forms of diazotroph symbioses in the marine environment hints that mutualism is an important mode of existence in the open sea. Such observations should not be surprising – the majority of N_2 fixed in terrestrial ecosystems appears controlled by diazotroph symbioses. Symbiotic interactions involving many of the most common marine diazotrophs suggests we may need to revisit the prevailing notion that competition is the primary factor shaping adaptation in the nutrient-depleted upper ocean. The struggle to acquire nutrients in persistently oligotrophic ocean habitats also appears to select for cooperative interactions as an evolutionary strategy, although perhaps at the expense of rapid growth.

In addition to the cyanobacteria, diverse non-cyanobacterial diazotrophs have been reported from marine habitats, including representatives of the α-, β-, δ-, and γ-proteobacteria, *Chlorobi*, *Actinobacteria*, *Firmicutes*, *Spirochaetes*, and archaea. In the well-oxygenated waters of the open sea, many of the N_2 fixers identified through isolation or via genetic techniques appear related to facultative or obligate anaerobes although studies on the distributions and activities of these organisms in the open sea are lacking. Based on the limited surveys available for evaluating these non-cyanobacterial N_2 fixers, it appears these organisms are often found at low population abundances in the upper ocean ($< 10^4$ *nifH* gene copies L^{-1}). Nonetheless, the application of PCR-dependent methodologies has begun to provide important clues into the diversity of non-cyanobacterial N_2-fixing microorganisms, suggesting much greater diversity of these organisms than currently represented in culture collections. Amplification and sequencing of *nifH* genes from environmental samples collected from both oligotrophic and eutrophic ocean waters indicates that diverse bacterial lineages dwell in the sea, including those phylogenetically grouping among the Clusters I, III, and IV *nifH* sequences. Although there is relatively little known about the metabolic capabilities of these organisms or their potential contribution to N_2 fixation, they appear widespread (Riemann *et al.*, 2010). Several studies have found expression of *nifH* genes by various groups of non-cyanobacterial diazotrophs, suggesting an active role for these microorganisms in marine

N_2 fixation. Organisms phylogenetically affiliated with the Cluster I *nifH* phylotypes are the most common non-cyanobacterial diazotrophs, and they include members of the α-, β-, γ- and δ-proteobacteria. Among those, several *nifH* phylotypes clustering among γ-proteobacteria seem to be widespread, having been reported in clone libraries not only from the Pacific and Atlantic Oceans but also from the Arabian and Baltic Seas (Langlois *et al.*, 2005). In addition, *nifH* genes similar to those belonging to obligate anaerobes (most closely related to phototrophic green-sulfur bacteria and spirochetes) contained within the Cluster III *nifH* sequences have been retrieved from *nifH* gene clone libraries in the North Pacific and Atlantic Oceans. These *nifH* gene sequences appear most abundant ($\sim 10^3$–10^4 *nifH* copies L^{-1}) in the dimly waters of the lower euphotic zone. Nevertheless, our current understanding of the ecological processes that control these presumed N_2-fixing microorganisms is woefully lacking and ripe for future progress.

Metabolic and Physiological Diversity

Oceanic diazotrophs include aerobes and anaerobes, phototrophs and chemotrophs, autotrophs and heterotrophs, in addition to various hybrid modes of metabolism. The process of N_2 fixation has a high demand for energy and reducing power and exhibits extreme sensitivity to O_2. The biosynthesis of the metalloenzyme nitrogenase places unique nutritional demands on these cells, including requirements for specific trace metals often found in vanishingly low concentrations in the sea. Furthermore, in many organisms the presence of ammonium (NH_4^+) suppresses N_2 fixation due to transcriptional and post-translational regulation of nitrogenase. The fairly stringent demands imposed by a diazotrophic lifestyle greatly narrow the window of available niche spaces able to support the growth of these organisms. Some of the most notable examples of physiological diversification among the marine N_2 fixers are found in the way they acquire energy, harvest nutrients, and protect nitrogenase from O_2, suggesting these selective forces factor prominently into shaping the evolutionary history of these organisms.

Given the relatively high energetic demands of N_2 fixation, it should not be surprising to find that many of the most common ocean diazotrophs utilize sunlight as a relatively plentiful source of energy in the sea. Such organisms include obligate and facultative photolithoautotrophs and photoorganoheterotrophs. Oxygenic photolithoautrophic N_2-fixing cyanobacteria have received considerable attention, with much of this focused on how these microorganisms protect nitrogenase from O_2 evolved during photosynthesis (Gallon, 1992). Most of these organisms rely on one of two strategies: (1) spatial separation of photosynthesis and N_2 fixation, or (2) temporal separation of these processes (Paerl, 1990). In the case of heterocystous cyanobacteria, nitrogenase is confined to the heterocyst; these specialized nonvegetative cells lack photosystem II but retain photosystem I, thereby enabling cyclic photophosphorylation to provide energy to fuel N_2 fixation during daylight hours. In addition, the heterocysts retain the catabolic machinery necessary for aerobic consumption of carbohydrate supplied by vegetative cells. Hence, by spatially separating photosynthesis from N_2 fixation, heterocystous cyanobacteria are able to actively fix N_2 during the daytime, and at night these organisms can continue to fix N_2 at a reduced rate through catabolism of photosynthetically fixed organic matter (Stal and Zehr, 2008). It is unclear whether this strategy for N_2 fixation might influence the establishment or maintenance of symbiotic interactions prevalent among oceanic heterocystous cyanobacteria and other photolithoautotrophs (e.g., diatoms).

Similar to heterocystous cyanobacteria, the non-heterocystous *Trichodesmium* also fixes N_2 during the day coincident with photosynthetic production of O_2. The mechanisms underlying the ability of *Trichodesmium* to fix N_2 at the same time when it is also actively producing O_2 through photolithoautrophic growth is the subject of continuing debate. Both photosynthesis and N_2 fixation in *Trichodesmium* are regulated at the level of gene transcription through a tightly coordinated circadian rhythm; moreover, *Trichodesmium* appears to adjust its respiratory rate to control intracellular O_2 concentrations. Two different hypotheses have emerged to explain how *Trichodesmium* fixes N_2 during the day: (1) nitrogenase is spatially segregated into differentiated nonphotosynthetic cells similar to heterocysts, or (2) cells within *Trichodesmium* filaments cycle rapidly between N_2 fixation and photosynthesis, thereby temporally separating N_2 fixation and photosynthesis (Stal and Zehr, 2008). To date, it remains unclear which of these strategies is utilized in support of N_2 fixation.

Free-living, non-heterocystous photolithoautotrophs face additional challenges in protecting nitrogenase from O_2; this is particularly acute for unicellular N_2-fixing microorganisms where spatial isolation of N_2 fixation and photosynthesis is not possible. Various groups of N_2-fixing cyanobacteria have evolved the capacity to fix N_2 at night in the absence of photosynthetic O_2 production; marine N_2 fixers among the genera *Crocosphaera*, *Cyanothece*, *Lyngbya*, and *Gloethece* all rely on temporal separation of N_2 fixation and photosynthesis. These organisms fuel the energetic demands of N_2 fixation via catabolism of photosynthetically fixed organic matter, a process that also consumes O_2.

Although photolithoautotrophic cyanobacteria have received considerable attention for their role in ocean N_2 fixation, recent discoveries suggest that some of the most abundant N_2-fixing cyanobacteria rely on a rudimentary photoorganoheterotrophic metabolism. Reconstruction of the group A cyanobacterial genome revealed that these organisms lack photosystem II but retain photosystem I, suggesting these organisms harvest sunlight for energy but do not evolve O_2 during photosynthesis (Zehr *et al.*, 2008). This discovery provides insight into the seemingly enigmatic observations that nitrogenase gene expression by group A peaks during the day (Church *et al.*, 2005), contrasting the strategy of temporal separation of N_2 fixation and photosynthesis observed in other unicellular cyanobacteria. The absence of photosystem II and the presence of photosystem I suggests N_2 fixation by group A is coupled to energy harvested through cyclic photophosphorlyation without concomitant evolution of O_2. The source of reductant needed for reducing N_2 remains unknown but could include recycling of H_2 or catabolism of organic matter.

Non-cyanobacterial marine N_2 fixers include both aerobes and facultative and obligate anaerobes, and these organisms rely on diverse modes of acquiring energy and reductant such as chemoorganoheterotrophy, methanotrophy, and methanogensis. In the open sea, various strains of chemoorganoheterotrophic γ-proteobacteria have been isolated that fix N_2 optimally under microaerobic or anaerobic conditions, including several strains belonging to the genera *Vibrio*, *Klebsiella*, and *Pseudomonas* (Zehr and Paerl, 2008). Although O_2 concentrations in the upper ocean are generally at or even slightly above saturation relative to the atmosphere, low O_2 niches for these organisms might exist within decomposing particles or within the guts of marine invertebrates. Other strategies employed by chemoheterotrophs for shielding nitrogenase from O_2 include building gas-diffusion barriers around the cell. Several groups of aerobic chemoheterotrophs, including *Azotobacter*, produce copious quantities of extracellular polysaccharides that, when combined with relatively high rates of respiration, appear to help minimize oxidative deactivation of nitrogenase (Fenchel *et al.*, 1998). Although little is known about the contribution of chemoorganoheterotrophs to oceanic N_2 fixation, the energetic demands of these organisms would likely require high rates of organic matter supply and thus likely limit their contributions to more-eutrophic but N-limited marine ecosystems (e.g., the Baltic Sea).

The Role of N_2 Fixation in Marine Ecosystems

Marine ecosystems are fundamental components of the global C cycle and N_2-fixing microorganisms factor prominently into the processes controlling ocean C inventories and fluxes. The biological C pump describes the collective set of processes that partly maintain the large vertical gradients observed in oceanic inorganic C concentrations. Net photosynthetic production of organic matter in the well-lit regions of the upper ocean consumes carbon dioxide (CO_2) and fuels the metabolic demands of the ocean's food web; however, due to gravitational settling of particles or convective mixing of dissolved organic matter, a fraction of this photosynthetically derived organic matter escapes immediate consumption in the upper ocean and is transported downward to the cold, unlit waters of the deep sea. During the downward transfer of this material, microorganisms actively decompose the sinking material, remineralizing organic matter back to inorganic constituents, including CO_2 (**Figure 3**).

The magnitude of C removal via the biological C pump depends in large part on the supply of growth requiring nutrients to the upper ocean (Eppley and Peterson, 1979). Although nutrient recycling in the upper ocean often supports a major fraction of biological productivity, the steady loss of material from the upper ocean via the biological pump needs to be supported by external nutrient subsidy. Dugdale and Goering (1967) formalized the terms "new" and "regenerated" production to describe the sources of nutrients (in their model based on N) supporting plankton production in the upper ocean. Regenerated production includes that fraction of production sustained by nutrient recycling in the euphotic zone, whereas new production describes plankton production supported by nutrients supplied from outside the euphotic zone. If the upper ocean is assumed to be in steady state over some reasonable time and space scales, then input of new nutrients should be largely balanced by removal of these nutrient subsidies in the form of organic matter. Thus, in N-limited regions of the ocean, the supply of new N should be quantitatively related to organic matter export to the deep sea. Sources of new nutrients to the upper ocean include nitrate introduced largely by physical processes (turbulent mixing, advection, or upwelling), atmospheric deposition of N-containing compounds, and N_2 fixation. As a result, the supply of bioavailable N to the upper ocean via N_2 fixation forms a central component of the biological C pump.

Quantification of N_2 Fixation

Over the past several decades, considerable efforts have been made in quantifying the importance of N_2 fixation to marine ecosystems (Capone, 2001). Such efforts have included direct biological rate measurements or geochemical approaches to derive rates of N_2 fixation based on anomalous stoichiometric behaviors of inorganic C or nutrient pools, or stable isotopic signatures of planktonic N. Prior to routine use of the stable isotopes in biological oceanography, most biological efforts to quantify rates of N_2 fixation were based on the reduction of acetylene to ethylene. Nitrogenase demonstrates low substrate specificity to N_2, allowing the enzyme to reduce several other compounds that share molecular similarity to N_2, including acetylene. Hence, biological reduction of acetylene to ethylene has been used as a surrogate measurement for N_2 fixation. The method suffers from several uncertainties, perhaps most notably the selection of appropriate acetylene to N_2 conversion factors. The technique typically relies on gas chromatograph separation of ethylene followed by flame ionization or thermal conductivity detection of the gas; for natural samples where the abundance of N_2 fixers is relatively low, the methodology has historically demanded preconcentration of N_2-fixing microorganisms to increase the analytical sensitivity of this method. This preconcentration is often done by handpicking colonies or filaments of visible N_2-fixing cyanobacteria, with the resulting rate of acetylene reduction per colony or filament extrapolated based on enumeration of the abundances of the organisms in the environment. Although this approach provides useful information, in particular on physiological variability of ocean diazotrophs (Capone *et al.*, 1990), growing realization that many of the abundant and potentially important marine diazotrophs may not be easily distinguishable based on phenotypic characteristics (e.g., unicellular N_2 fixers) presents impediments to utilizing these derived rate measurements to obtain estimates of rates of N_2 fixation in the sea.

The development and application of ^{15}N stable isotope-based approaches marked a major advance in quantifying rates of N_2 fixation. Tracing the assimilation of ^{15}N-labeled N_2 gas into plankton biomass provides a direct means of determining rates of N_2 fixation; the sensitivity of the method has proven useful even in ecosystems where rates of N_2 fixation and abundances of N_2 fixers are low. However, the methodology also faces several limitations. In particular, dilution of the $^{15}N_2$ substrate by the large natural pool of dissolved N_2

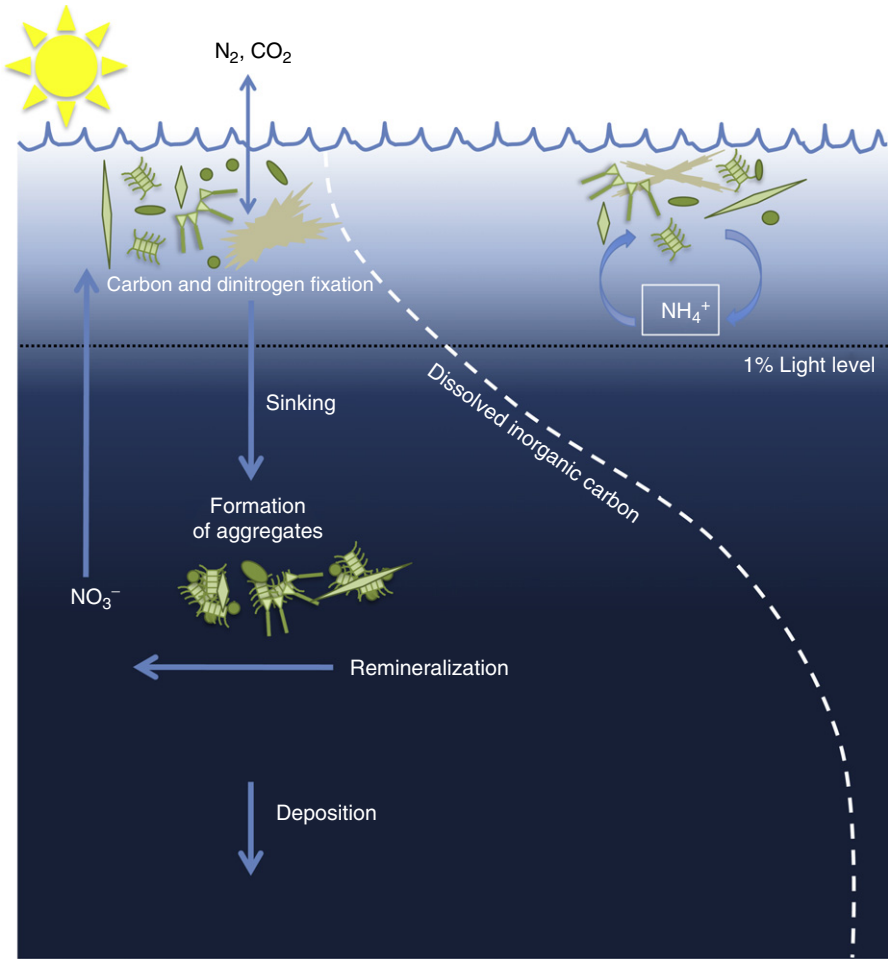


Figure 3 Depiction of the role of new production in fueling the biological carbon pump. Vertical gradient in dissolved inorganic carbon (dashed line). New production and carbon export supported by NO_3^- and N_2 shown on left side of figure, whereas productivity fueled by regenerated nutrients (via recycling of NH_4^+) is depicted on right side of figure.

(>600 μM) limits the sensitivity of the assay. Moreover, as commonly employed, the method measures the accumulation of ^{15}N into particulate material (cell biomass); and to the extent that release of ^{15}N labeled dissolved material (amino acids or NH_4^+) constitutes an important pathway for fixed N, the method would underestimate the rate of fixation. This is one potential advantage to using the acetylene reduction method: it may provide an estimate of N_2 fixation closer to gross rather than the net estimate likely obtained using the $^{15}N_2$ technique. Moreover, a recent study has uncovered a potentially serious flaw with the standard $^{15}N_2$ methodology: when $^{15}N_2$ is injected as a gas into seawater, the equilibration time for the $^{15}N_2$ with the seawater appears relatively slow, potentially resulting in underestimation of the rate of N_2 fixation (Mohr *et al.*, 2010). These authors modified the conventional method by dissolving the $^{15}N_2$ gas into the seawater prior to the incubation, then adding the equilibrated mixture to seawater samples; the resulting rates of fixation by *Crocosphaera* were ~60% greater than rates estimated by the conventional "bubble injection" technique.

The just described methods have provided considerable insight into processes underlying variability in N_2 fixation in different ecosystems and also suggest that the process can be highly variable in space and time. However, the value of extrapolating such measurements to obtain estimates of ecosystem- or global-scale rates of N_2 fixation remains unclear. Several geochemical approaches based on anomalous nutrient and inorganic C distributions provide additional constraints on basin- to global-scale N_2 fixation. Time-series measurements of inorganic C drawdown at the Bermuda Atlantic Time-series Study (BATS) in the Sargasso Sea revealed an intriguing seasonal pattern (Michaels *et al.*, 1994): wintertime convective mixing delivers nutrients and inorganic carbon to the euphotic zone; these nutrients are rapidly consumed in support of biological production and subsequent vertical export during the early- to mid-spring, resulting in net removal of inorganic C. However, throughout the warm summer and fall months when upper-ocean nutrient concentrations are at their annual minimum, inorganic C inventories continue to disappear, presumably a manifestation of an active biological C pump. Mixing, upwelling, and advection of nutrients do not appear sufficient to support new production during these periods (Michaels *et al.*, 1994). Similar observations emerged from the Hawaii Ocean Time-series (HOT) program in the

North Pacific (Karl *et al.*, 2002). These observations highlighted a potentially important role for N_2-fixing microorganisms in the open sea: by supplying a source of N to the upper ocean during the thermally stratified summer months, these organisms may provide nutrient subsidy to the oligotrophic oceans that ultimately fuels the biological C pump. Subsequent extrapolation of these observations to other regions of the ocean (based on satellite-derived sea surface temperatures) provided basin- to global-scale estimates of N_2 fixation–supported new production (Lee *et al.*, 2002).

Additional efforts to constrain basin-scale rates of N_2 fixation hinged on observed stoichiometric patterns in inorganic and organic nutrient pools. The congruence in elemental stoichiometries of deep ocean nutrient pools and planktonic biomass has been known for many years. Globally, the molar ratio of N to phosphorus (P) found in seawater nutrient pools – namely, nitrate (NO_3^-) and phosphate (PO_4^{3-}) – converges near 16 : 1, similar to the average N : P ratio of these elements found in marine plankton (Redfield *et al.*, 1958). However, examples of anomalous behaviors from this canonical ratio abound, and such examples have provided insight into the magnitude of N inputs and losses in ocean ecosystems. In the subtropical North Atlantic, inorganic N : P ratios of waters in the upper thermocline can exceed 50 : 1; whereas in the subtropical North Pacific, the N : P ratios of upper ocean and thermocline nutrient pools are nearly always $<1:1$ (**Figure 4**). Two important N-cycle processes are known to effectively decouple the cycling of N and P in the sea: N_2

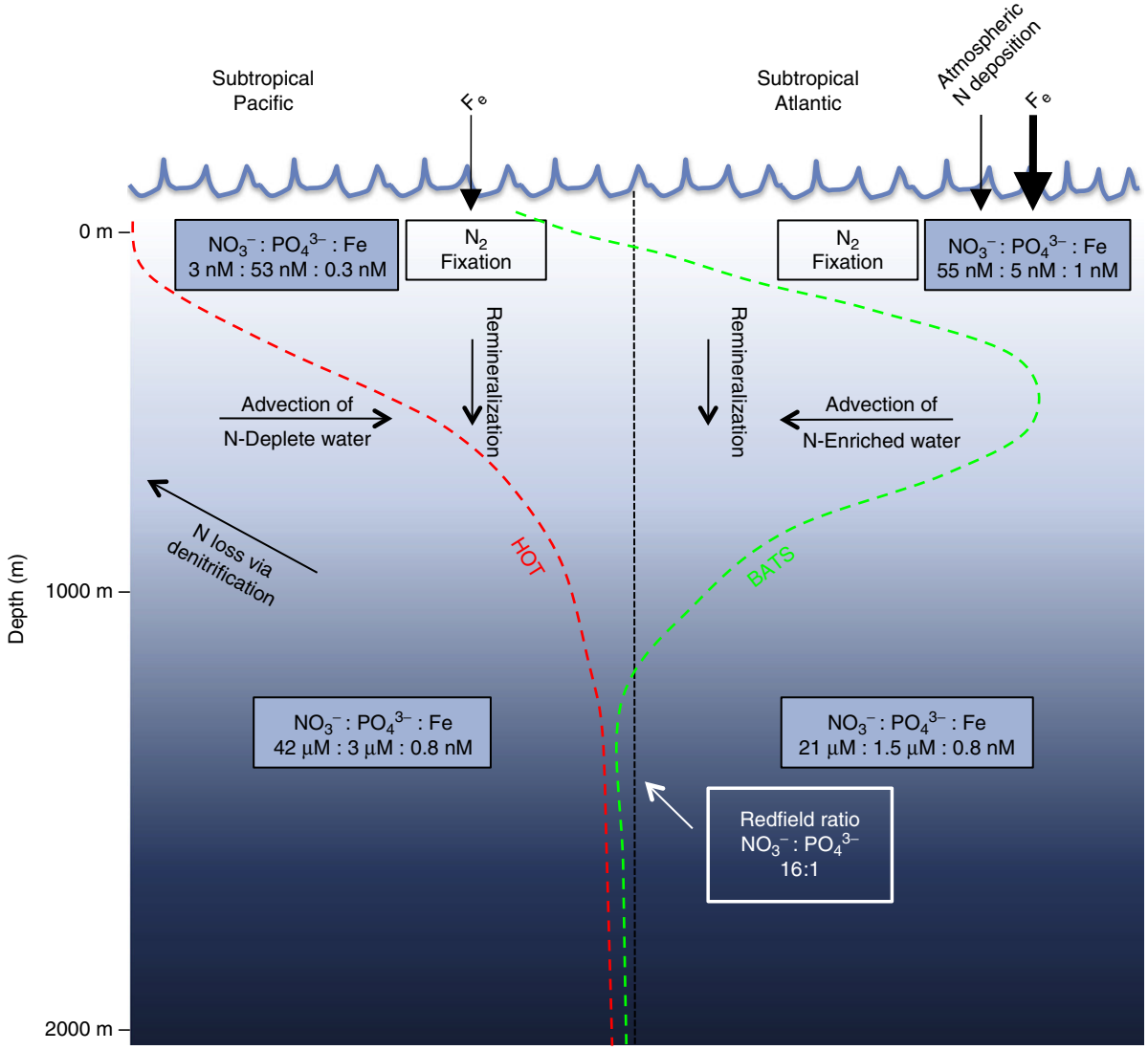


Figure 4 Processes contributing to vertical relationships in inorganic nutrient (NO_3^- and PO_4^{3-}) N : P ratios. N : P ratios for the subtropical Pacific depicted by dashed orange line, whereas N : P ratio for subtropical Atlantic depicted by dashed green line; Redfield ratio (16N : 1P) shown with black dashed line. Low Fe deposition to the waters of subtropical North Pacific is hypothesized to limit N_2 fixation, which together with advection of denitrified waters to this region results in thermocline N : P ratios $\ll$ Redfield ratio. In the subtropical Atlantic, enhanced Fe-enriched dust deposition is hypothesized to stimulate N_2 fixation and thereby consume P, resulting in thermocline N : P ratios of the remineralized nutrient pools $\gg$ Redfield ratio.

fixation and denitrification. N_2 fixation provides a source of N to the ocean in a manner often decoupled from introduction of P, whereas denitrification removes fixed N independent from removal of P. Diazotrophs appear to demonstrate considerable flexibility in their biomass N : P ratios, thereby potentially introducing an N-rich source of organic matter that could fuel net remineralization in the deep ocean. Moreover, through direct release of dissolved N substrates (either NH_4^+ or amino acids), N_2-fixing microorganisms have the potential to significantly alter seawater nutrient pools. Thus, in ecosystems where N_2 fixation constitutes an important supply term for fixed N, the remineralized inorganic nutrient pools and dissolved organic matter pools should reflect excess input of N relative to P. Such anomalous behavior in N : P nutrient stoichiometry has been used to constrain basin- to global-scale rates of N_2 fixation and denitrification (Michaels *et al.*, 1996; Gruber and Sarmiento, 1997; Deutsch *et al.*, 2007). Seawater nutrient N : P ratios indicate excess N input (presumably due to N_2 fixation) in the Atlantic, whereas net N loss (presumably due to denitrification) predominates in the Pacific. Note that these patterns do not necessarily imply that rates of N_2 fixation are greater in the Atlantic than the Pacific; high rates of N input by N_2 fixation in the Pacific could be obscured by greater denitrification occurring in this region (Gruber and Sarmiento, 1997).

There have been a number of challenges to the hypotheses that N_2 fixation supports the observed drawdown in inorganic C observed at HOT and BATS and that contemporaneous rates of N_2 fixation are sufficient to produce the anomalous N : P ratios observed in the thermocline of the North Atlantic. Isotopic analyses of the ^{15}N signatures of NO_3^- and total N (largely dissolved organic N) suggest N_2 fixation may be only a small source ($\sim$2%) of new N input to the euphotic zone in these systems (Knapp *et al.*, 2005). These studies highlighted a number of considerations when constraining rates of N_2 fixation-based on patterns of anomalous nutrient stoichiometry. Foremost among these considerations is that we currently have poor estimates of the importance of processes other than N_2 fixation in contributing to anomalous nutrient N : P ratios; in particular, the potential importance of atmospheric deposition and non-Redfieldian nutrient remineralization. In addition, these studies point out that deviations in seawater nutrient N : P ratios from the canonical Redfield ratio may not provide insight into local-scale processes, but rather reflect a host of spatially integrated processes that ultimately result in N-enrichment of thermocline waters; such processes could include advection of water from regions that support high rates of N_2 fixation, or transport of water from regions where N is supplied to upper ocean via atmospheric deposition (Knapp *et al.*, 2005; Hansell *et al.*, 2007).

N_2 Fixation Fueled Linkages Among Nutrient Cycles

Marine N_2-fixing microorganisms appear well adapted to growth in warm, stratified, N-depleted habitats (Karl *et al.*, 2002). But if N_2-fixing microorganisms provide a source of fixed N to such ecosystems, why do so many regions of the tropical and subtropical ocean appear so tightly controlled by the supply and availability of N? What limits N supply by N_2 fixation in these ecosystems? Why aren't N_2-fixing microorganisms more abundant in these N-limited ecosystems? The persistent dearth of nutrients found in large regions of the open sea suggests that bioessential nutrients likely limit the capacity of these ecosystems to support new biomass. Plankton demand for nutrients has long been known to exceed supply in many regions of the sea, and given that low nutrient supply defines oligotrophic habitats, there is considerable interest in identifying how N_2-fixing microorganisms obtain nutrients needed for growth. Much of the focus has been directed toward understanding how N_2-fixing microorganisms meet their nutritional demands for P or Fe or vitamins. Over geological time scales the availability of Fe and P appear important regulators of ocean C sequestration via the biological C pump, in part through the role these elements play in controlling N_2 fixation (Falkowski, 1997). Both Fe and P are introduced to the oceans through weathering of rock and soils, with P delivery to the oceans mostly via riverine runoff and Fe supplied by wind-borne deposition of dust. Both Fe and P are essential components of biomass: P is central to synthesis of nucleotides and lipids, whereas Fe is central to numerous protein systems including those required for electron transport, photosynthesis, and nutrient assimilation.

Understanding how the introduction of bioavailable N to the planktonic food web by the activities of N_2-fixing microorganisms impacts the availability and turnover of other nutrient elements in the ocean continues to be a subject of intense study. Such new N input should increase planktonic demands for P, and in fact examples of secular-scale changes in P inventories have been attributed to changes in N_2 fixation (Karl *et al.*, 1997). The eventual drawdown of P due to N_2 fixation establishes various feedbacks, including imposing P stress or even limitation on diazotroph growth. N_2-fixing microorganisms possess high affinity P uptake systems that permit them to harvest very low concentrations of PO_4^{3-}. Nonetheless, kinetic studies on *Trichodesmium* suggest both growth and P-uptake limitation can be limited at PO_4^{3-} concentrations <0.2 μM (Fu *et al.*, 2005), which occurs in most of the regions of the sea where these organisms thrive. To cope with such P-stress, various groups of marine diazotrophs appear to acquire P through utilization of organic compounds, including phosphomonoesters or polyphosphates (Karl and Björkman, 2002). Although *Crocosphaera* and *Trichodesmium* encode alkaline phosphatases and polyphophatases for hydrolyzing phosphate esters and polyphosphates, respectively, *Trichodesmium* appears to have further diversified its mode of acquiring P through utilization of phosphonate compounds. Phosphonates are a unique subset of C–P bonded organic substrates that comprise a relatively large fraction of organic P found in seawater (Dyhrman *et al.*, 2006). Intriguingly, laboratory and field-based studies indicate that diazotroph growth on selected alkylphosphonates (methyl- and ethylphosphonate) results in production of the corresponding alkanes (e.g., methane and ethane) thereby linking the P scavenging metabolism of these organisms to production of climate-reactive gases (Karl *et al.*, 2008). Marine N_2 fixers also demonstrate considerable flexibility in their cellular P quotas, with cellular N : P ratios ranging from 16 : 1 to 90 : 1 (White *et al.*, 2006). Part of this flexibility likely derives from

the ability to limit their cellular demand for P through substitution of sulfur in membrane lipids (Van Mooy *et al.*, 2009) or through reduction in P-rich ribosomes (Arrigo, 2005). These adaptations to low P conditions further highlight the complexity of factors that shape diazotroph biodiversity in the sea.

The large Fe-demands of nitrogenase (enzyme contains up to 50 mol Fe per mole of the enzyme complex) appears to place relatively large Fe requirements on N_2-fixing microorganisms. *Crocosphaera* appears to partly deal with this high demand for Fe through intense intracellular Fe recycling (Saito *et al.*, 2011); however, net diazotroph growth in the oceans still demands a supply of Fe. Such elevated Fe requirements have focused attention on whether low delivery of Fe to the oceans might limit contemporary rates of N_2 fixation. Stimulation of N_2 fixation through eolian Fe deposition has been proposed as a mechanism underlying basin-scale differences in $NO_3^-:PO_4^{3-}$ ratios observed between the Atlantic and Pacific (**Figure 4**); whereas $NO_3^-:PO_4^{3-}$ ratios in the main thermocline of the Atlantic exceed the Redfield ratio (16 : 1), in the Pacific ratios of these nutrient pools are consistently below this canonical ratio (**Figure 4**). Delivery of Fe-rich Saharan dust to the central Atlantic has been hypothesized to support euphotic zone N_2 fixation (Michaels *et al.*, 1996; Moore *et al.*, 2006), a process that would fuel formation of N-enriched organic matter that ultimately results in N-enriched remineralized inorganic nutrients below the euphotic zone. Several direct field tests of this hypothesis indicated both Fe and P can limit N_2 fixation in these waters, consistent with the view that Fe delivery controls N_2 fixation which in turn regulates P inventories (Mills *et al.*, 2004; Sañudo-Wilhelmy *et al.*, 2001). In contrast, eolian Fe input to the relatively remote waters of the Pacific appears lower and possibly limits N_2 fixation, thereby creating the relative surplus in PO_4^{3-} observed throughout the euphotic zone waters of the Pacific (Wu *et al.*, 2000; Cavender-Bares *et al.*, 2001).

Several recent intriguing developments in the study of factors controlling rates of N_2 fixation in the global ocean have explored coupling between different processes in the N cycle and how these processes influence the relative distributions of nutrient elements necessary to support diazotrophs. In particular, ocean models hint that physical transport of denitrified waters with low $NO_3^-:PO_4^{3-}$ ratios could effectively couple the processes of N_2 fixation and denitrification at the global scale (Deutsch *et al.*, 2007). Denitrification removes fixed N and leaves residual P relative to the Redfield ratio. The process occurs in low-O_2 waters, including regions in the Arabian Sea and along the eastern boundaries of the Pacific and Atlantic Oceans. Reintroduction of these relatively P-enriched waters to the euphotic zone has been hypothesized to promote N_2 fixation. Together, these hypothesized interactions between N_2 fixers and Fe supply or coupling with denitrification reinforce the key role of N_2-fixing microorganisms in linking elemental cycles in the sea and highlight possible climate-sensitive feedbacks regulating the ocean's N cycle. Future perturbations to Earth's climate, either via the hydrological cycle or warming, could alter the magnitude and spatial distribution of dust inputs and denitrification, impacting N_2 fixation, nutrient distributions, and the biological C pump.

Patterns and Progress in Diazotroph Ecology

A hallmark of N_2-fixing microorganisms is their ability to modify their environment through introduction of fixed N, thereby provoking a suite of subsequent changes in overall ecosystem dynamics. As previously described, the inevitable drawdown of other bioessential nutrient pools (including P and Fe) modifies the stoichiometry of seawater nutrient pools that the diazotrophs depend on for growth. Through their growth, N_2-fixing microorganisms appear to establish self-imposed feedbacks that ultimately influence their competitive success and shape ecosystem biodiversity. There has been considerable effort directed toward defining emergent patterns in the distributions and activities of marine N_2 fixers, and this has provided a wealth of information on spatial and temporal dynamics underlying diazotroph ecology. It has become increasingly clear that despite apparent functional redundancy among members of the marine diazotroph guild, the genetic and metabolic diversity of N_2-fixing microorganisms has direct bearing on ocean ecology and biogeochemical cycling. In the following sections, we highlight a small subset of intriguing findings that point to the importance of oceanographic processes in structuring diazotroph diversity and the role of this diversity in driving biogeochemical cycles.

Spatiotemporal Variability

With the recognition that N_2 fixers are keystone functional members of the planktonic food web, interest in defining the biogeography of diazotrophs, with an eye to identifying factors that shape the resulting patterns, has increased. Although the generalized distributions of microscopically identifiable N_2-fixing microorganisms has been known for many years, the use of molecular-based approaches has provided new tools to examine the distributions and activities of different groups of N_2 fixers (including many only known based on gene sequences) and in some cases expanded the geographical boundaries of habitats known to support N_2-fixing microorganisms. Although these studies have confirmed that N_2-fixing microorganisms appear best suited to the warm, nutrient-poor regions of open sea, they have also highlighted several important biogeographical distinctions in the types of organisms that predominate in these regions. For example, the uncultivated unicellular group A appears better adapted to lower temperatures and greater nutrient concentrations than other marine N_2 fixers, extending its geographic range in the oceans to include the extremes of the subtropical gyres (Moisander *et al.*, 2010; Sohm *et al.*, 2011). Although we recognize that abundances and activities of N_2-fixing microorganisms are highly variable in space and time, despite years of research, in most cases the processes underlying this variability remain unknown. Moreover, most studies on distributions and diversity of marine N_2-fixing microorganisms have focused on cyanobacteria and our understanding of spatial and temporal variability associated with non-cyanobacterial N_2 fixers remains weak. Those few studies that have explored distributions and diversity of these non-cyanobacterial N_2 fixers suggest such diazotrophs may exhibit a larger geographic range than the cyanobacteria (Farnelid *et al.*, 2011).

At the basin scale, several prominent biogeographical features emerge from studies on marine diazotroph diversity and distributions: (1) *Trichodesmium* appears to be a dominant N_2 fixer in the tropical and subtropical waters of the Atlantic Ocean; (2) unicellular N_2 fixers often dominate diazotroph abundances in the subtropical Pacific; (3) despite the presence of diverse N_2-fixing microorganisms in the Mediterranean Sea, the P-depleted waters of this ecosystem appear to support very low rates of N_2 fixation; and (4) diazotroph dynamics in the southern hemisphere remain grossly undersampled, limiting our understanding of N_2 fixation in large areas of the Indian Ocean, and both the South Pacific and Atlantic Oceans. Although it is tantalizing to infer that temperature is the primary selective force on these larger-scale biogeographic patterns, changes in temperature often coincide with alterations in physical forcing of the upper ocean, which can have indirect consequences on light availability and nutrient supply. Although both the physiology and distributions of various N_2-fixing microorganisms demonstrate temperature dependent patterns, diazotroph abundances in the warm, nutrient-enriched waters of the equatorial Pacific are low (**Figure 5**), suggesting temperature is not the primary controller of N_2 fixers in this ecosystem. Such meridional patterns could stem from a variety of processes, including limitation of N_2 fixers by low Fe supply, or the inability of diazotrophs to effectively compete for growth limiting nutrients in a habitat where NO_3^- is supplied to the upper ocean via physical processes (in this case upwelling). Studies with laboratory isolates of *Trichodesmium* and *Crocosphaera* indicate that N_2 fixation is repressed in the presence of NH_4^+ at concentrations above ~0.2 μM, but that rates of N_2 fixation by these organisms can be less sensitive to changes in NO_3^- concentrations (Ohki *et al.*, 1991; Dekaezemacker and Bonnet, 2011). However, NO_3^- supply to the N-starved upper ocean would almost certainly instill intense competition for growth limiting nutrients among naturally occurring plankton assemblages. N_2 fixers generally appear to grow more slowly than non-diazotrophic phytoplankton, presumably as a result of the relatively high energy demands imposed by N_2 fixation, and this could limit their ability to effectively compete for resources in environments where NO_3^- supply is not limiting.

Our expanding knowledge of broad-scale patterns related to diazotroph population structure has proven fruitful for formulating new hypotheses on the ecological and physiological dynamics underlying these patterns (LaRoche and Bretibarth, 2005). However, exploration of notable exceptions to these patterns has also proved important to developing mechanistic understanding into processes controlling diazotroph diversity in the sea. The oceans are temporally and spatially heterogeneous habitats and this heterogeneity factors

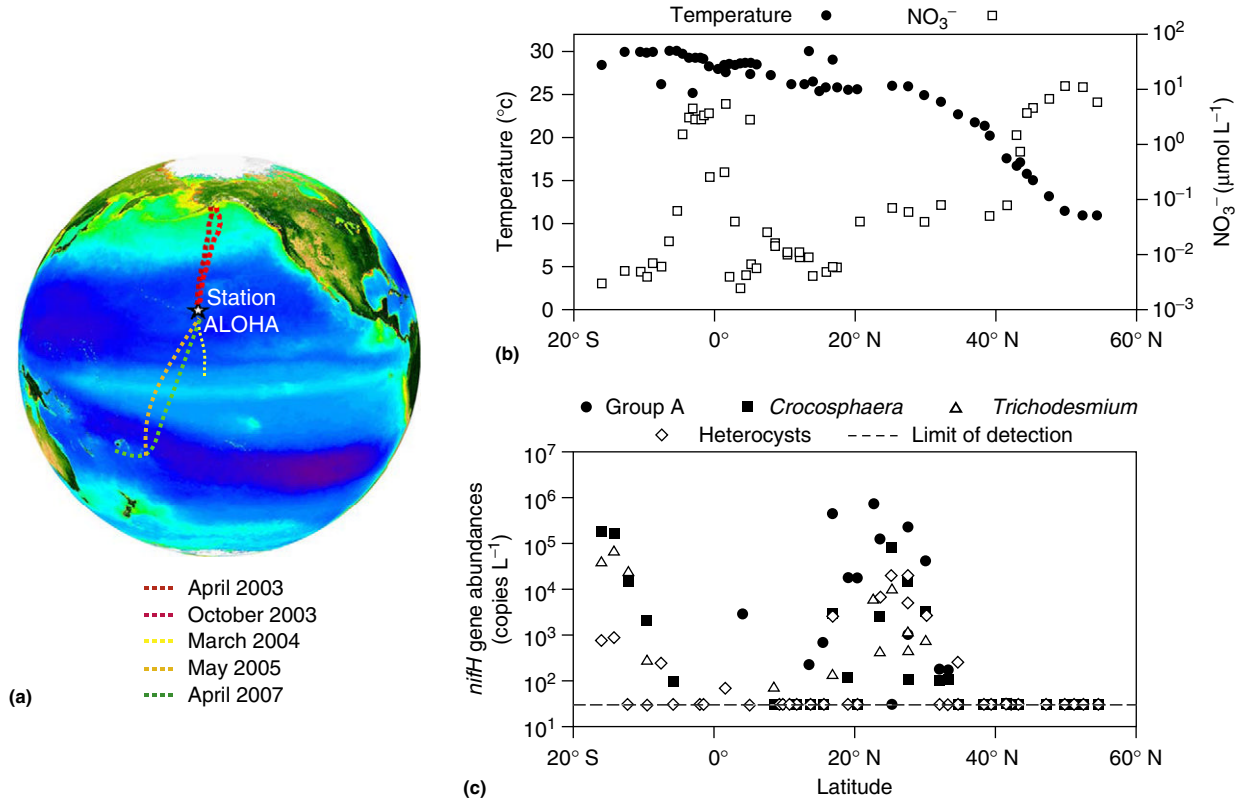


Figure 5 Sea-viewing wide field-of-view sensor (SeaWiFS) satellite-composite image depicting distributions of surface ocean chlorophyll in the Pacific Ocean, with cruise tracks from five research expeditions shown (a). Panels (b) and (c) depict meridional distributions of near-surface ocean nitrate (NO_3^-), temperature, and cyanobacterial *nifH* gene abundances obtained from the five cruises shown in (a). Dashed line (c) indicates approximate detection limit of quantitative PCR-based *nifH* gene determinations. Heterocysts includes sum of *Richelia* and *Calothrix* phylotypes.

prominently into shaping diazotroph dynamics. Oceanic fronts, riverine plumes, and mesoscale eddies all superimpose physical and biogeochemical variability onto otherwise relatively quiescent regions of the sea, and such features appear to influence diazotroph dynamics. For example, the diazotroph habitat expands and contracts with meridional changes in the major frontal features that define the extremes of the subtropical gyres (Church *et al.*, 2008). In the North Pacific, interannual- to seasonal-scale changes in the strength of westerly and trade wind forcing alters the latitudinal positions of fronts along both the northern and southern boundaries of the oligotrophic subtropical gyre, expanding or contracting the areal extent of this habitat by $>10^6$ km^2. In addition, riverine input to marginal zones along the tropical and subtropical oceans is an important factor in control on N_2 fixer dynamics in these ecosystems. Studies in both the Amazon and Mekong river plumes indicate that riverine discharge of terrestrial material establishes a spatially extensive succession of planktonic organisms that includes a prominent role for N_2 fixers (Foster *et al.*, 2007; Subramaniam *et al.*, 2008; Voss *et al.*, 2006). The silica, Fe, and P enriched waters associated with such river plumes appears to promote the growth of N_2 fixers, specifically diatom-associated heterocystous cyanobacteria. Hence, time-varying changes in the intensity and spatial dynamics of the plume discharge directly contribute to variability in diazotroph dynamics in such regions.

Mesoscale eddies have long been recognized as important controls on physical and biogeochemical variability in the oceans. Eddies are ubiquitous features in the sea; these isolated vortices are typically 10–100 s of kilometers in diameter, and most persist for weeks to months. Eddies develop due to instabilities in flow along frontal boundaries or through wind-driven forcing at topographic features. Depending on the direction of their rotation, eddies alter nutrient availability in the upper ocean waters. Most of the attention on how mesoscale eddies impact marine plankton has focused on cyclonic eddies, whose rotation results in uplift of nutrient-enriched waters into the euphotic zone. However in both the Atlantic and Pacific, the activities and abundances of N_2-fixing organisms appear enhanced by downwelling anticyclonic eddies; the growth of *Trichodesmium*, *Crocosphaera*, and diatom-associated heterocystous N_2 fixers appear stimulated by such physical forcing (Davis and McGillicuddy, 2006; Church *et al.*, 2009). Nevertheless, the mechanisms that promote N_2 fixation by specific organisms associated with these features remain unknown.

We currently have considerably more information on spatial variability in N_2-fixing microorganisms than we do on time-varying changes in these organisms. However, deconvolving spatial and temporal variability in a dynamic fluid environment is often difficult. Much of our understanding of temporal variability in marine diazotrophs derives from time series sampling at a few locations in the world's oceans. Several studies based at HOT and BATS examined time-variability in the dynamics of N_2-fixing microorganisms, specifically *Trichodesmium* and heterocystous cyanobacteria (Letelier and Karl, 1996; Orcutt *et al.*, 2001). Among other findings, these studies highlighted episodic blooms as important features of the ecology and biogeochemistry of marine N_2 fixers. Although diazotrophs are generally found in low abundances in the open sea, *Trichodesmium*, *Calothrix*, and *Richelia* are notable for occasionally forming large blooms and possibly dominating plankton biomass during these events. In some cases, these blooms are often large enough in magnitude as to be detectable by Earth-orbiting satellites (Westberry and Siegel, 2006; White *et al.*, 2007). Most often these blooms are observed during the summer months, when the upper ocean is stratified, and nutrient supply to the upper ocean is presumably at its annual minimum. The net accumulation of biomass during such bloom events implies that on some time scale, processes controlling the growth and loss of these organisms have become decoupled. The expansion of biomass during blooms events could derive from non–steady state behavior in either bottom-up or top-down controls on diazotroph populations. Although top-down processes are undoubtedly important controls on population ecology, in many regions of the open sea limited nutrient supply would appear to leave little room for expansion of diazotroph population sizes. To date it remains unclear what processes result in the initiation of net growth of N_2-fixing microorganisms or how those processes change to eventually terminate these blooms. However, the impact of these blooms on ocean biogeochemistry is apparent. At HOT, a recurrent summertime increase in the flux of organic carbon to the deep sea accounts for >30% of the annual carbon export in this ecosystem (Dore *et al.*, 2008). Microscopic analyses of sediment trap samples from this site revealed that this temporally focused pulse of material to the deep sea is driven by sinking diatoms, many of which contain endosymbiotic *Richelia* (Scharek *et al.*, 1999). Although the processes that control this sedimentation phenomenon remain unknown, the observations highlight an important role for time-varying changes in diazotroph population structure on ocean carbon sequestration.

Time series sampling of *nifH* gene abundances and upper ocean N_2 fixation (based on assimilation of $^{15}N_2$) at HOT has provided insight into seasonal to event scale fluctuations in the diazotroph assemblage dynamics (**Figure 6**). During most of the year, *nifH* gene abundances are dominated by unicellular cyanobacteria, specifically group A. However, as the upper ocean warms and becomes more stratified between the late spring and fall, filamentous N_2 fixers (including *Trichodesium*, *Richelia*, and *Calothrix*) become increasingly abundant. These seasonal-scale changes in diazotroph community structure appear to directly influence rates of N_2 fixation. More than 70% of the annual N_2 fixation appears controlled by microorganisms <10 μm in diameter (presumably reflecting the activities of group A and *Crocosphaera*); however, during those months when filamentous organisms become more dominant, rates of N_2 fixation by larger (>10 μm in diameter) organisms increase. Thus, larger filamentous organisms appear best adapted to periods when upper ocean nutrient supply is low, but energy is plentiful, whereas unicellular microorganisms appear to flourish during the cooler, more nutrient-enriched winter and spring months. These seasonal-scale shifts in the size structure of marine diazotrophs appear contradictory to conceptual models which predict larger celled plankton flourish under periods of elevated nutrient supply, whereas picoplankton appear better suited to periods when turbulence is low and nutrient supply limiting.

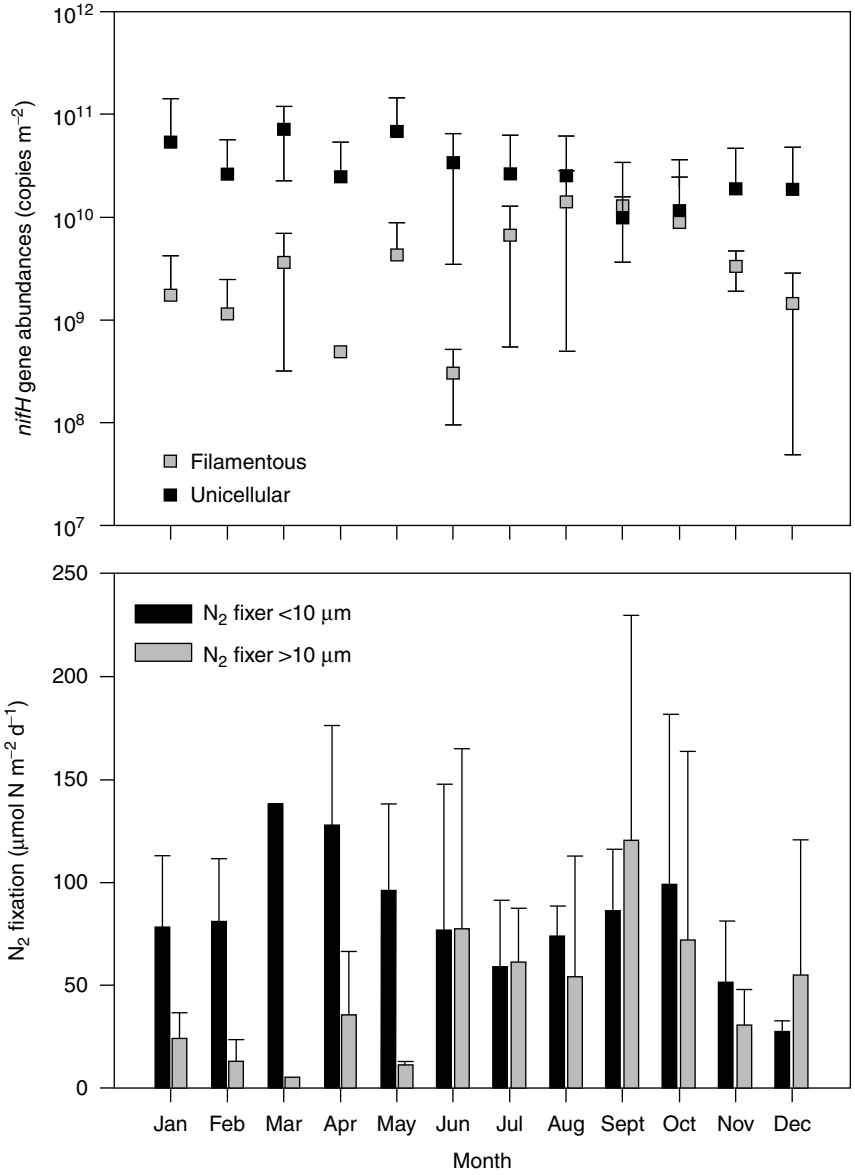


Figure 6 Monthly determinations of depth-integrated (0–100 m) *nifH* gene inventories (a) and rates of N_2 fixation (b) by two size classes (<10 μm and >10 μm) of microorganisms in the subtropical North Pacific Ocean. Filamentous *nifH* groups includes the sum of *Trichodesmium*, *Richelia*, and *Calothrix*; whereas unicellular phylotypes include sum of *Crocosphaera* and group A phylotypes.

Diazotrophy in the Changing Sea

The organization of microbial assemblages underlies many of the emergent properties of marine ecosystem functions, including the pathways and efficiency through which energy gets dispersed and material gets retained within the food web. Fluctuations in environmental forcing – that is, nutrient supply, solar energy input, redox conditions – all directly influence the organization and structure of planktonic communities. Hence, understanding the sensitivities and resilience of contemporary ecosystem functions to changes in microbial diversity is central to our predictive knowledge on how these systems may change in the future. There are clear examples of progressive changes occurring in ocean habitats, many attributable to anthropogenic activities, but our knowledge of how such changes are shaping plankton biodiversity, and more specifically diazotroph diversity, remain largely rudimentary. Both measurements and models hint that climate-driven changes may favor expansion of warm, low nutrient marine ecosystems that comprise key habitats for N_2-fixing microorganisms, likely increasing the ecological and biogeochemical footprint of diazotrophy in the sea. However, such expanded habitats could differ significantly from the oligotrophic gyres we recognize in contemporary oceans, with lower pH, elevated temperatures, and dampened seasonality in nutrient entrainment. If the apparent genomic conservation demonstrated by members of *Trichodesmium* and *Crocosphaera* (Zehr *et al.*, 2007) turns out to be a distinguishing feature

among other groups of marine diazotrophs, this characteristic may reduce the genetic flexibility available to these organisms and make these populations potentially less resilient to future changes in environmental selective pressures.

Some of the most robust examples of progressive ocean change are found in decadal-scale records of the seawater carbonate system. Ocean CO_2 inventories are increasing in step with atmospheric CO_2, driving concomitant decreases in seawater pH. These changes in the seawater carbonate system influence life in the sea in a myriad of ways, ranging from direct physiological changes to alteration in the speciation of bioessential metals, and thus are likely to exert increasingly important selective pressures on marine plankton, including the diazotrophs. Laboratory studies have documented how N_2-fixing microorganisms acclimate to abrupt experimental manipulations in key habitat properties (e.g., light, nutrients, temperature, CO_2). Such studies have revealed sensitivities in N_2 fixers to changes in CO_2; however, the responses by different organisms can be varied and complex. Growth and N_2 fixation by different strains of *Trichodesmium* and *Crocosphaera* increase under elevated CO_2 conditions (Hutchins *et al.*, 2009; Levitan *et al.*, 2007), whereas other diazotrophs such as the bloom-forming heterocystous cyanobacterium *Nodularia* demonstrate a negative response to increasing CO_2 (Czerny *et al.*, 2009).

Designing and conducting experiments that adequately simulate adaptive responses by the microorganisms to slow and progressive ecosystem change remains challenging. It is currently unclear to what extent information obtained from short-term acclimation experiments will be useful in helping to predict future changes to ocean diazotrophs. Although many of the projected changes to the ocean habitat would appear favorable to growth of specific N_2-fixing microorganisms, such changes could also trigger new ecosystem stress points that feedback onto the N_2 fixers. For example, possible stimulation of diazotroph growth by increased CO_2 availability might further stress demand for limiting nutrients in the open sea, potentially offsetting any physiological benefit associated with increased CO_2 availability. We currently have limited information on the physiological tolerances of naturally occurring ocean diazotrophs or how simultaneous changes in multiple climate-sensitive factors might interact to influence the resulting physiologies of these organisms.

Conclusions

Diazotrophs have long been recognized as functionally important members of the ocean plankton. Increasing interest in the role these organisms play in ocean C cycling has pushed understanding those processes controlling their growth, rates of N_2 fixation, distributions, diversity, and abundances to the forefront of biological oceanography. Despite often comprising a low proportion of plankton biomass, N_2-fixing organisms are metabolically diverse and thus spatiotemporal shifts in their population structure can impart distinct biogeochemical signatures on marine ecosystems. Moreover, these diverse groups of microorganisms appear sensitive to environmental forcing, suggesting these organisms could be early responders to future changes in the sea. There are many unanswered questions regarding their ecology and biogeochemistry, but it is clear that these organisms provide several unique ecosystem services, making them essential components of contemporary marine systems. Several key themes related to marine N_2-fixing microorganisms that we have sought to emphasize here include:

1. Marine diazotrophs exhibit a wide range of metabolic diversity, including organisms that are photolithoautotrophs, photoorganoheterotrophs, and chemoorganoheterotrophs, and those both obligate and facultative aerobes and anaerobes.
2. Introduction of fixed N to the open sea by N_2 fixation has numerous consequences on ecosystem dynamics, including regulating material export to the deep sea, altering the elemental stoichiometry of ocean nutrient pools, and influencing production of reduced gases.
3. Input of new N via N_2 fixation has required reassessment of the processes and mechanisms that supply other bioessential nutrients (e.g., P and Fe) to ocean ecosystems.
4. Marine diazotrophs demonstrate complex and diverse ecological interactions, including prominent roles for symbiotic interactions, many of which remain poorly understood.
5. Our knowledge of the biogeochemical importance and ecological dynamics of non-cyanobacterial marine diazotrophs remains woefully scant.
6. The diversity and activities of marine diazotrophs appear susceptible to projected future changes to the ocean's climate and C cycle.

Acknowledgments

We acknowledge David Karl and the Center for Microbial Oceanography: Research and Education (C-MORE) for the opportunity to contribute this work. In addition, the US National Science Foundation supported the authors through grants OCE 08–50827 and 09–26766 (to MJC).

See also: Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Services. Marine Symbioses: Metazoans and Microbes. Microbial Biogeography. Microorganisms (Microbes), Role of. Ocean Ecosystems. Plankton, Status and Role of

References

Arrigo KR (2005) Marine microorganisms and global nutrient cycles. *Nature* 437: 349–355.

Capone DG (2001) Marine nitrogen fixation: What's the fuss? *Current Opinion in Microbiology* 4: 341–348.

Capone DG and Carpenter EJ (1982) Nitrogen fixation in the marine environment. *Science* 217: 1140–1142.

Capone DG, O'Neil JM, Zehr JP, and Carpenter EJ (1990) Basis for diel variation in nitrogenase activity in the marine planktonic cyanobacterium *Trichodesmium thiebautii*. *Applied and Environmental Microbiology* 56: 3532–3536.

Carpenter EJ (2002) Marine cyanobacterial symbioses. Biology and Environment. *Proceedings of the Royal Irish Academy* 102B: 15–18.

Carpenter EJ and Janson S (2000) Intracellular cyanobacterial symbionts in the marine diatom *Climacodium frauenfeldianum*. *Journal of Phycology* 36: 540–544.

Carpenter EJ, Montoya JP, Burns J, *et al.* (1999) Extensive bloom of a N_2-fixing diatom/cyanobacterial association in the tropical Atlantic Ocean. *Marine Ecology Progress Series* 185: 273–283.

Cavender-Bares KK, Karl DM, and Chisholm SW (2001) Nutrient gradients in the western North Atlantic Ocean: relationship to microbial community structure and comparison to patterns in the Pacific Ocean. *Deep-Sea Research I* 48: 2373–2395.

Church MJ, Björkman KM, Karl DM, Saito MA, and Zehr JP (2008) Regional distributions of nitrogen-fixing bacteria in the Pacific Ocean. *Limnology and Oceanography* 53: 63–77.

Church MJ, Mahaffey C, Letelier RM, *et al.* (2009) Physical forcing of nitrogen fixation and diazotroph community structure in the North Pacific subtropical gyre. *Global Biogeochemical Cycles* 23: GB2020.http://dx.doi.org/10.1029/2008GB003418.

Church MJ, Short CM, Jenkins BD, Karl DM, and Zehr JP (2005) Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. *Applied and Environmental Microbiology* 71: 5362–5370.

Czerny J, Barcelos e Ramos J, and Riebesell U (2009) Influence of elevated CO_2 concentrations on cell division and nitrogen fixation rates in the bloom-forming cyanobacterium *Nodularia spumigena*. *Biogeosciences* 6: 1865–1875.

Davis CS and McGillicuddy DJ (2006) Transatlantic abundance of the N_2 fixing colonial cyanobacterium *Trichodesmium*. *Science* 312: 1517–1520.

Dekaezemacker J and Bonnet S (2011) Sensitivity of N_2 fixation to combined nitrogen forms (NO_3^- and NH_4^+) in two strains of the marine diazotroph *Crocosphaera watsonii* (Cyanobacteria). *Marine Ecology Progress Series* 438: 33–46.

Deutsch C, Sarmiento JL, Sigman DM, Gruber N, and Dunn JP (2007) Spatial coupling of nitrogen inputs and losses in the ocean. *Nature* 445: 163–167.

Dore J, Letelier R, Church M, Lukas R, and Karl D (2008) Summer phytoplankton blooms in the oligotrophic North Pacific Subtropical Gyre: Historical perspective and recent observations. *Progress in Oceanography* 76: 2–38.

Dugdale RC and Goehring JJ (1967) Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnology and Oceanography* 12: 196–206.

Dyhrman ST, Chappell PD, Haley ST, *et al.* (2006) Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* 439: 68–71.

Eppley RW and Peterson BJ (1979) Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282: 677–680.

Falkowski PG (1997) Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* 387: 272–276.

Farnelid H, Andersson AF, Bertilsson S, *et al.* (2011) Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. *PLoS ONE* 6: e19223.http://dx.doi.org/10.1371/journal.pone.0019223.

Fenchel T, Blackburn TH, and King GM (1998) *Bacterial Biogeochemistry*. London: Academic Press.

Foster RA, Goebel NL, and Zehr JP (2010) Isolation of *Calothrix rhizosoleniae* (Cyanobacteria) strain SC01 from *Chaetoceros* (Bacillariophyta) spp. diatoms of the subtropical North Pacific Ocean. *Journal of Phycology* 46: 1028–1037.

Foster RA, Subramaniam A, Mahaffey C, *et al.* (2007) Influence of the Amazon River plume on distributions of free-living and symbiotic cyanobacteria in the western tropical North Atlantic. *Limnology and Oceanography* 52: 517–532.

Fu F, Zhang Y, Bell PRF, and Hutchins D (2005) Phosphate uptake and growth kinetics of *Trichodesmium* (Cyanobacteria) isolates from the North Atlantic Ocean and the Great Barrier Reef, Australia. *European Journal of Phycology* 41: 62–73.

Gallon J (1992) Reconciling the incompatible: N_2 fixation and O_2. *New Phytologist* 122: 571–609.

Gruber N and Galloway JN (2008) An earth-system perspective of the global nitrogen cycle. *Nature* 451: 293–296.

Gruber N and Sarmiento JL (1997) Global patterns of marine nitrogen fixation and denitrification. *Global Biogeochemical Cycles* 11: 235–266.

Hansell DA, Olson DB, Dentener F, and Zamora LM (2007) Assessment of excess nitrate development in the subtropical North Atlantic. *Marine Chemistry* 106: 562–579.

Hutchins DA, Mulholland MR, and Fu F (2009) Nutrient cycles and marine microbes in a CO_2-enriched ocean. *Oceanography* 22: 128–145.

Johnston AWB, Li Y, and Ogilvie L (2005) Metagenomic nitrogen fixation – feast or famine? *Trends in Microbiology* 13: 416–420.

Karl D, Letelier R, Tupas L, *et al.* (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* 388: 533–538.

Karl D, Michaels A, Bergman B, *et al.* (2002) Dinitrogen fixation in the world's ocean. *Biogeochemistry* 57/58: 47–98.

Karl DM and Björkman KM (2002) Dynamics of DOP. In: Hansell DA and Carlson CA (eds.) *Biogeochemistry of Marine Dissolved Organic Matter*, pp. 249–366. San Diego: Academic Press.

Karl DM, Bidigare RR, Church MJ, *et al.* (2008) The nitrogen cycle in the North Pacific trades biome: an evolving paradigm. In: Capone DG, Bronk DA, Mullholland MR, and Carpenter EJ (eds.) *Nitrogen in the Marine Environment*, 2nd ed, pp. 705–769. Burlington: Elsevier.

Knapp AN, Sigman DM, and Lipischultz F (2005) N isotopic composition of dissolved organic nitrogen and nitrate at the Bermuda Atlantic Time-series Study site. *Global Biogeochemical Cycles* 19: GB1018.http://dx.doi.org/10.1029/2004GB002320.

Langlois RJ, LaRoche J, and Raab PA (2005) Diazotrophic diversity and distribution in the tropical and subtropical Atlantic Ocean. *Applied and Environmental Microbiology* 71: 7910–7919.

LaRoche J and Breitbarth E (2005) Importance of the diazotrophs as a source of new nitrogen in the ocean. *Journal of Sea Research* 53: 67–91.

Lee K, Karl DM, Wanninkhof R, and Zhang JZ (2002) Global estimates of net carbon production in the nitrate-depleted tropical and subtropical oceans. *Geophysical Research Letters* 29.http://dx.doi.org/10.1029/2001GL014198.

Letelier RM and Karl DM (1996) Role of *Trichodesmium* spp. in the productivity of the subtropical North Pacific Ocean. *Marine Ecology Progress Series* 133: 263–273.

Levitan O, Rosenberg G, Setlik I, *et al.* (2007) Elevated CO_2 enhances nitrogen fixation and growth in the marine cyanobacterium *Trichodesmium*. *Global Change Biology* 13: 531–538.

Lundgren P, Janson S, Singer A, Jonasson S, and Bergman B (2005) Unveiling of novel radiations within the *Trichodesmium* cluster by hetR gene sequence analysis. *Applied and Environmental Microbiology* 71: 190–196.

Michaels AF, Knapp A, Dow R, *et al.* (1994) Seasonal patterns of ocean biogeochemistry at the U.S. JGOFS Bermuda Atlantic Time-series Study site. *Deep-Sea Research II* 41: 1013–1038.

Michaels AF, Olson D, Sarmiento JL, *et al.* (1996) Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic. *Biogeochemistry* 35: 181–226.

Mills MM, Ridmae C, Davie M, La Roche J, and Geider RJ (2004) Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* 429: 292–294.

Mohr W, Großkopf T, Wallace DWR, and LaRoche J (2010) Methodological underestimation of oceanic nitrogen fixation rates. *PLoS ONE* 5: e12583. http://dx.doi.org/10.1371/journal.pone.0012583.

Moisander P, Beinart RA, Hewson I, *et al.* (2010) Unicellular cyanobacterial distributions broaden the oceanic N_2 fixation domain. *Science* 19: 1512–1514.

Moore JK, Doney SC, Lindsay K, Mahowald N, and Michaels AF (2006) Nitrogen fixation amplifies the ocean biogeochemical response to decadal timescale variations in mineral dust deposition. *Tellus* 58B: 560–572.

Ohki K, Zehr JP, Falkowski PG, and Fujita Y (1991) Regulation of nitrogen-fixation by different nitrogen sources in the marine non-heterocystous cyanobacterium *Trichodesmium* sp. NIBB 1067. *Archives of Microbiology* 156: 335–337.

Orcutt KM, Lipschultz F, Gundersen K, *et al.* (2001) A seasonal study of the significance of N_2 fixation by *Trichodesmium* spp. at the Bermuda Atlantic Time-series Study (BATS) site. *Deep Sea Research II* 48: 1583–1608.

Paerl H (1990) Physiological ecology and regulation of N_2 fixation in natural waters. *Advances in Microbial Ecology* 11: 305–344.

Redfield A (1958) The biological control of chemical factors in the environment. *American Scientist* 46: 205–221.

Riemann L, Farnelid H, and Steward GF (2010) Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters. *Aquatic Microbial Ecology* 61: 225–237.

Saito MA, Bertrand EM, Dutkiewicz S, *et al.* (2011) Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph *Crocosphaera watsonii*. *Proceedings of the National Academy of Science* http://dx.doi.org/10.1073/pnas.10069431.

Sañudo-Wilhelmy SA, Kustka AB, Gobler CJ, *et al.* (2001) Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean. *Nature* 411: 66–69.

Scharek R, Tupas LM, and Karl DM (1999) Diatom fluxes to the deep sea in the oligotrophic North Pacific Gyre at station ALOHA. *Marine Ecology Progress Series* 182: 55–67.

Sohm JA, Webb EA, and Capone DG (2011) Emerging patterns of marine nitrogen fixation. *Nature Reviews Microbiology* 9: 499–508.

Stal LJ and Zehr JP (2008) Cyanobacterial nitrogen fixation in the ocean. Diversity, regulation and ecology. In: Herrero A and Flores E (eds.) *Cyanobacteria:*

Molecular Biology, Genomics and Evolution, pp. 423–446. Norfolk, UK: Caister Academic Press.

Subramaniam A, Yager PL, Carpenter EJ, *et al.* (2008) Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean. *Proceedings of the National Academy of Science* http://dx.doi.org/10.1073/pnas.0710279105.

Taniuchi Y, Lee Chen Y-L, Chen H-Y, Tsai M-L, and Ohki K (2011) Isolation and characterization of the unicellular diazotrophic cyanobacterium Group C TW3 from the tropical western Pacific Ocean. *Environmental Microbiology* http://dx.doi.org/10.1111/j.1462-2920.2011.02606.x.

Tripp HJ, Bench SR, Turk KA, *et al.* (2010) Metabolic streamlining in an open ocean nitrogen-fixing cyanobacterium. *Nature* 464: 90–94.

Van Mooy BAS, Fredricks HF, Pedler BE, *et al.* (2009) Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* 458: 69–72.

Venrick EL (1974) The distribution and significance of *Richelia intracellularis* Schmidt in the North Pacific Central Gyre. *Limnology and Oceanography* 19: 437–445.

Voss M, Bombar D, Loick N, and Dippner JW (2006) Riverine influence on nitrogen fixation in the upwelling region off Vietnam, South China Sea. *Geophysical Research Letters* 33: L07604.http://dx.doi.org/10.1029/2005GL025569.

Webb EA, Ehrenreich IM, Brown SL, Valois FW, and Waterbury JB (2009) Phenotypic and genotypic characterization of multiple strains of the diazotrophic cyanobacterium, *Crocosphaera watsonii*, isolated from the open ocean. *Environmental Microbiology* 11: 338–348.

Westberry TK and Siegel DA (2006) Spatial and temporal distribution of *Trichodesmium* blooms in the world's oceans. *Global Biogeochemical Cycles* 20: GB4016.http://dx.doi.org/10.1029/2005GB002673.

White AE, Spitz YH, Karl DM, and Letelier RM (2006) Flexible elemental stoichiometry in *Trichodesmium* spp. and its ecological implications. *Limnology and Oceanography* 51: 1777–1790.

White AE, Spitz YH, and Letelier RM (2007) What factors are driving summer phytoplankton blooms in the North Pacific Subtropical Gyre? *Journal of Geophysical Research* 112: C12006.http://dx.doi.org/10.1029/2007JC004129.

Wu J, Sunda W, Boyle EA, and Karl DM (2000) Phosphate depletion in the western north Atlantic Ocean. *Science* 289: 759–762.

Zehr JP and Kudela RM (2011) Nitrogen cycle of the open ocean: From genes to ecosystems. *Annual Review of Marine Science* 3: 197–225.

Zehr JP and Paerl HW (2008) Biological nitrogen fixation in the marine environment. In: Kirchman DL (ed.) *Microbial Ecology of the Oceans*, 2nd ed, pp. 387–426. New York: Wiley-Liss, Inc.

Zehr JP, Bench SR, Carter BJ, *et al.* (2008) Globally distributed uncultivated oceanic N_2-fixing cyanobacterial lack oxygenic photosystem II. *Science* 322: 1110–1112.

Zehr JP, Bench SR, Mondragon EA, McCarren J, and DeLong EF (2007) Low genomic diversity in tropical oceanic N_2 fixing cyanobacteria. *Proceedings of the National Academy of Sciences* 104: 17807–17812.

Zehr JP, Jenkins BD, Short SM, and Steward GF (2003) Nitrogenase gene diversity and microbial community structure: A cross-system comparison. *Environmental Microbiology* 5: 539–554.

Zehr JP, Waterbury JB, Turner PJ, *et al.* (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. *Nature* 412: 635–638.

Diversity, Molecular Level

Carlos A Machado, University of Maryland, College Park, MD, USA
Marcos A Antezana, Universidade de Lisboa, Lisboa, Portugal

© 2013 Elsevier Inc. All rights reserved.

Glossary

Background selection Selection that takes place at loci other than that of interest. Such events can strongly increase the intensity of genetic drift as well as hinder the action of selection at a locus of interest. These effects are normally modeled by reducing the effective population size applicable to the locus of interest.

Divergence A fixed genetic difference between two species or populations, or the process of evolving such a difference.

Effective population size The size that a random mating population must have to experience the same intensity of genetic drift as a population of interest. The size can vary across genes due to changes in background selection across the genome.

Functional constraint A limit to the kinds of mutations that can appear in a gene or protein without compromising function, often also called "selective" constraints.

Genetic drift Random changes in gene frequency that occur when a finite number of zygotes or propagules is formed by random sampling gametes or genotypes from the previous generation to create the next generation.

Mutation A change in the nucleotide sequence of a DNA or RNA molecule.

Nearly neutral mutation A mutation whose population genetic dynamics are influenced by both genetic drift and selection.

Neutral mutation A mutation whose population genetic dynamics are influenced only by genetic drift.

Nonsynonymous mutation A nucleotide change in a protein coding region that changes the encoded amino acid.

Polymorphism A genetic difference between two or more individuals of the same species; to exclude new mutations from the definition, the underlying mutations are required to segregate above an arbitrarily set minimal frequency (1% or 5%).

Replacement Divergence due to a change of an amino acid in a protein, or the process to evolve such a change.

Selection Differential reproduction within a set of genes, individuals, groups, etc. caused by differential performance in fitness determining tasks. The appropriate unit of selection (gene, individual, group) is determined by the level at which the causation of the fitness differences must be described (interactions or performances at the level of genes, individuals, groups, etc.).

Substitution Divergence due to a change of base in a DNA sequence, or the process of evolving such a change.

Synonymous mutation A nucleotide change in a protein coding region that does not change the encoded amino acid.

Understanding the population genetic mechanisms that generate molecular diversity is fundamental to understanding how biodiversity is maintained and generated. This article deals with molecular diversity at the DNA and protein levels, and concentrates on diversity that is genetically heritable. We will discuss molecular diversity patterning and causation primarily at the level of single genes and genomic regions. Therefore, important aspects of molecular diversity that concern higher level genomic structuring and evolution will not be mentioned (e.g., genome size evolution, gene duplication, and gene family evolution). The article will deal first with DNA and protein diversity within species (molecular polymorphism) and then with diversity generated by genetic differentiation between species (molecular divergence). Such order is justified because visualizing diversity in this way helps understanding the evolutionary and population genetic forces and processes that generate it.

The information presented here is almost exclusively drawn from knowledge gathered by molecular population geneticists and molecular evolutionary biologists. Molecular population genetics studies the form and extent of molecular polymorphism, and tries to identify the forces that pattern it. Molecular evolutionary research aims at characterizing the degree and forms of molecular divergence, at clarifying how such differentiation takes place, and – by taking advantage of the former – at reconstructing the history of DNA sequences and proteins and the genealogical relationships among extant and sometimes extinct organisms.

Introduction

Phenotypic differences between organisms are due to genetic and environmental differences, and the former are the result of genetic divergence or polymorphism. Realizing that polymorphism is a common and striking phenomenon in the living world is not only important when trying to make sense of molecular diversity. Indeed, the transition from typological (essentialistic) to variational–populational thinking was a major contribution of the Darwinian revolution to biology.

Measuring genetic diversity within and between species is a long-standing interest of biologists. However, this only became possible in a systematic way in the 1960s when novel techniques from molecular biology (protein sequencing and enzyme electrophoresis) began to be applied to the study of evolutionary and population genetic questions. These techniques opened the possibility to study variation at the molecular level (amino acid sequence divergence and allozyme variation)

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00282-3

and gave birth to the fields of molecular evolution and molecular population genetics. The availability of protein sequences from various taxa triggered sophisticated studies of the forces driving protein evolution and led to the first reconstructions of historical relationships among organisms using protein sequence data (molecular phylogenetics). However, protein sequencing could not be used for quantitatively oriented studies of polymorphism because of its laboriousness. Electrophoretic studies of variation were carried out in hundreds of taxa uncovering the presence of large amounts of genetic variation in natural populations, although levels of molecular variation were mostly underestimated due to methodological limitations.

The advent of DNA sequencing in the 1980s improved resolution and made it possible to describe exhaustively the genetic variation segregating at single loci. Before DNA studies this was not possible, and in fact investigators developed and tested models that predicted patterns for electrophoretic variation segregating at many loci, because only such data could be produced in statistically analyzable counts. The shift to DNA studies allowed researchers to quantify conclusively not only amino acid polymorphism but also coding region variation that does not change the encoded amino acids (synonymous variation) and variation in noncoding regions, both of which were found to occur in copious amounts. This increased drastically the amount of data obtainable from a single locus and thus the number and the statistical quality of contrasts among classes of variation subject to arguably different functional constraints. New methods were developed to describe the processes shaping DNA level variation within and between populations and new, often more powerful, modeling approaches were developed and/or given a more solid empirical grounding (e.g., coalescence and maximum likelihood estimation of genetic distances). More recently, developments in DNA sequence technology have allowed fast and cheap genome sequencing, allowing researchers to obtain large-scale, and more precise, estimates of polymorphism and divergence for a wide variety of species. Unfortunately, despite all the technical and conceptual advances in quantifying and interpreting patterns of molecular variation, we still know very little about the genetic basis of adaptation and about the relationship between genetic and phenotypic differences among organisms.

The Generation and Maintenance of Genetic Diversity

There are three main sources of genetic variation in natural populations: mutation, recombination, and migration. Mutation is the only mechanism that can generate variation in a genetically homogeneous population, since recombination and migration only shuffle and shuttle around preexisting variation, respectively. Mutation is therefore the original source of all genetic variation.

To understand the evolutionary and population genetic forces patterning genetic variation, it is useful to consider variation at two main levels: within and between species. The first level constitutes the polymorphism component of genetic diversity, which can be also considered at finer levels if gene flow is structured within the species under study ("geographic" structure). Within sexual populations this type of diversity more or less freely combined each generation to form individual genotypes. The second level constitutes the divergence component of genetic diversity, which can also be considered at multiple levels depending on how closely related the species under study are. This type of diversity consists of fixed differences between species and is only exchanged under special circumstances (horizontal gene transfer, introgression).

The patterns of genetic variation within species arise from the combined action of two main set of processes: (1) a molecular-level origination process in which a mutation in a DNA sequence arises and (2) a population-level "fate" process that determines whether the mutation is quickly eliminated from the population or enters the pool of intraspecific variation becoming a polymorphism. The fate process also includes the possibility that the mutation becomes fixed in the population and thus becomes part of the divergence component of diversity. If a new mutation is passed on to the next generation of individuals, its final fate depends on the stochastic error produced when creating a finite number of zygotes or propagules by sampling gametes or genotypes from the previous generation (random genetic drift), on demographic forces like changes in population size and gene flow, as well as on the deterministic action of natural selection if the mutation is deleterious or advantageous.

Diversity as Polymorphism

What is the quantity (and quality) of genetic variation found in natural populations? What are the evolutionary forces responsible for the observed patterns of variation in natural populations? What is the meaning of that variation? What is the relationship between adaptive changes, phenotypic changes, and extant genetic variation? These are some of the questions asked by evolutionary geneticists. This section will give examples and summarize the current answers to the first two questions, but will not elaborate on the last two given the inconclusiveness of the current knowledge on the subject.

The Mutation Process: Molecular Biological Considerations

The mutation process is known to be affected by physiological and life history characteristics of the organism that affect the rate and accuracy of DNA replication (generation time, number of divisions before gametogenesis, sex, age) and by factors that increase damage to DNA (metabolic rate, age, transcription rate, frequency, and degree of packing and unpacking of chromosomes for expression). Exogenous agents that can damage DNA include UV light, ionizing radiation from X-rays, heat, alkylating agents, and hydroxyl radicals originated by the interaction of ionizing radiation with water. Endogenous agents include oxidant by-products of normal metabolism. Changes generated by endogenous agents cause the highest fraction of DNA damage. In human cells, oxidants damage between 10^4 and 10^5 nucleotide sites every day. Such damage would be extremely deleterious were it not for the presence of enzymes involved in finding and repairing it.

Of the five major types of DNA mutations – substitutions, deletions, insertions, inversions, and duplications – nucleotide substitutions and single base pair deletions or insertions (all called point mutations) account for 85% of the mutation events. With the exception of macromolecular events like chromosomal rearrangements, most mutations are the result of errors made by enzymes involved in DNA replication and repair. Replication errors are often due to the incorporation of noncomplementary nucleotides by the DNA polymerase and incorrect subsequent proofreading by the same enzyme. The latter process, however, is normally extremely effective in lowering the error rate by a millionfold relative to the intrinsic misincorporation rate of the polymerase.

Table 1 Direct estimates of the genome-wide mutation rate in five eukaryotes using whole genome sequencing

Organism	*Mutation rate per base pair/ generation (genome size)*
Saccharomyces cerevisiae[a]	0.33×10^{-9} (13 Mb)
Caenorhabditis elegans[b]	2.7×10^{-9} (100 Mb)
Drosophila melanogaster[c]	3.5×10^{-9} (130 Mb)
Arabidopsis thaliana[d]	7.0×10^{-9} (120 Mb)
Homo sapiens[e]	11.8×10^{-9} (3200 Mb)

[a]Lynch M, Sung W, Morris K, *et al.* (2008) A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Science USA* 105: 9272–9277.
[b]Denver DR, Dolan PC, Wilhelm LJ, *et al.* (2009) A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proceedings of the National Academy of Sciene USA* 106: 16310–16314.
[c]Keightley PD, Trivedi U, Thomas M, Oliver F, Kumar S, and Blakster ML (2009) Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Research* 19: 1195–1201.
[d]Ossowski S, Schneeberger K, Lucas-Lledo JI, *et al.* (2010) The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* 327: 92–94.
[e]Conrad DF, Keebler JE, DePristo MA, *et al.* (2011) Variation in genome-wide mutation rates within and between human families. *Nature Genetics* 43: 712–714.

Estimating the rate at which spontaneous mutations arise is challenging but constitutes a very important aspect of evolutionary studies. In the past, genome-wide mutation rates were extrapolated from indirect measures at single or few loci using experimental (e.g., mutation accumulation lines and reporter constructs) or evolutionary approaches (e.g., using patterns of sequence divergence among species). Due to the recent development of high-throughput methods to sequence complete genomes, precise estimates of genome-wide mutation rates have been obtained for five eukaryotes (**Table 1**). A recent analysis that includes a wide variety of organisms suggests that mutation rate is inversely correlated with genome size in viruses and prokaryotes, although this correlation is extremely weak. However, mutation rate is strongly and positively correlated with genome size in unicellular and multicellular eukaryotes (**Figure 1(b)**). Those differences in the fidelity of DNA replication across organisms with different genome size may be explained by the interplay of natural selection and population size, given the negative correlation between inferred effective population sizes and genome size (**Figure 1(a)**): selection to lower the mutation rate is more efficient in organisms with large effective population sizes (which have smaller genome sizes), whereas in organisms with smaller population size (and higher genome size) genetic drift becomes a limiting factor on how much selection can drive mutation rates to lower levels.

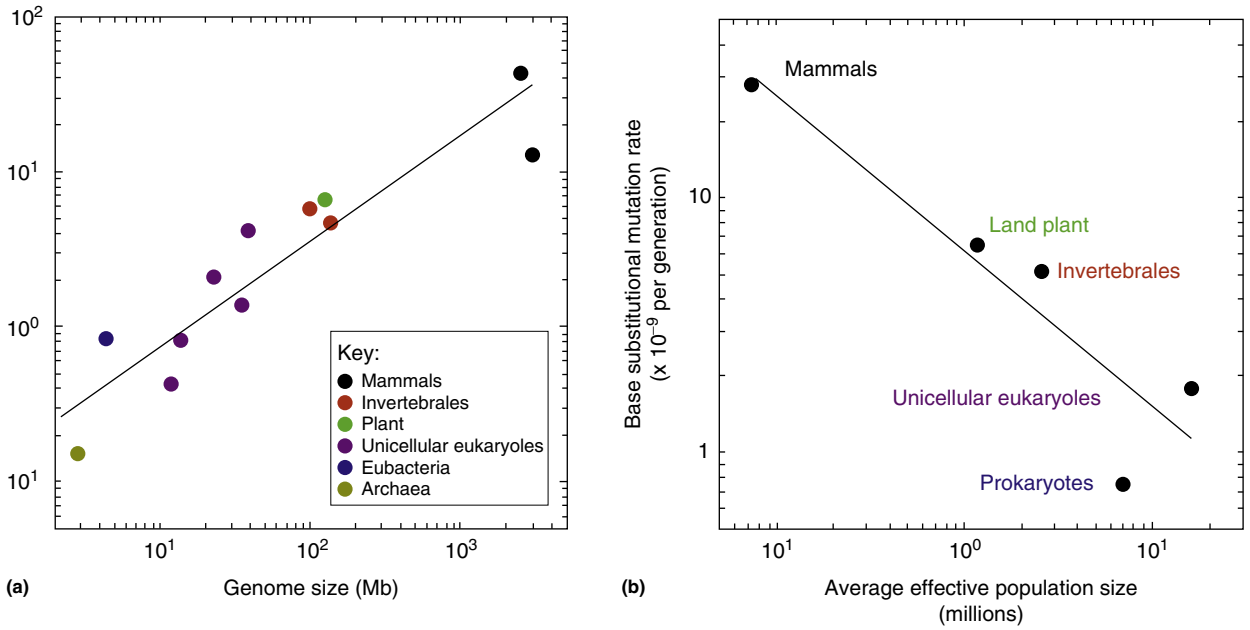


Figure 1 Scaling of the substitution rate per nucleotide site per generation with: (a) genome size or (b) inferred effective population size. Each data point represents the average estimate for a separate taxon. Data for most bacteria are based on just one or a few reporter constructs, whereas those for most multicellular taxa are based on very large datasets, including whole genome sequences. Reprinted from Lynch M (2010) Evolution of the mutation rate. *Trends in Genetics* 26: 345–352.

The Mutation Process: Cytological and Ontogenetic Considerations

In viruses, prokaryotes, and unicellular eukaryotes – with the notable exception of protists like *Paramecium* that have a macro- and a micronucleus – all mutations that are not strongly deleterious are heritable and become part of the polymorphism component of diversity. In metazoans most mutations arise in somatic cells and are therefore not heritable, although they may very well have an impact on the fitness of the individual carrying them. Mutations that arise in the cell lineages leading to propagules or gametes are however heritable. This applies in general to all multicellular organisms. In organisms with a germline (a terminal differentiation of specialized gamete-producing cells), differences in the timing of germline sequestration and in the number of cell divisions before gametogenesis can be studied carefully, and have been shown to have a major impact on the number of new mutations found in gametes. A good example is the observation that most of the evolution at the molecular level in different groups of animals (flies, fish, birds, and mammals) is "male driven." In mammals, in X- and Y-linked homologous genes the male-to-female ratio of nucleotide substitution is close to 6 in humans and 2 in rodents, which corresponds closely to the ratio of male-to-female germline divisions by the time of sexual maturity. This means that in mammals, humans included, most mutations participating in molecular divergence arise in the male germline.

Factors Influencing the Fate of Mutations

Patterns of molecular diversity other than those created by a continuous flow of short-lived mutations, which remain unique to the individuals in which they arise, can be very striking. For instance, some of the loci encoding the antigen-presenting major histocompatibility complex (MHC) in humans show extremely high levels of amino acid and synonymous polymorphism as well as many alleles segregating at intermediate frequencies. However, loci in the human Y chromosome show little polymorphism and mainly low-frequency mutations.

In general, patterns of polymorphism are determined by mutation pressure, genetic drift, and natural selection. The effectiveness of the action of the last two forces depends on the effective population size and on the fitness effect of the mutation involved. In particular, genetic drift increases in strength with decreasing population size and this in turn hinders the deterministic action of selection.

The patterns of variation of selectively neutral mutations are almost completely determined by genetic drift, population size, and the rate of mutation to neutral variants. For selected mutations both drift and selection can affect the patterns. In the case of strongly deleterious mutations, the observed pattern is one of mutations unique to each individual and none segregating at high frequencies as such mutations remain in the population for only very few generations. When a steady state is reached between the number of new deleterious alleles being produced by mutation and those being removed by selection, one speaks of a mutation–selection balance. Strongly favorable mutations should not show any observable polymorphism patterns for such mutations are normally fixed soon after being introduced in the population (if they are not immediately lost to drift) and practically do not participate in polymorphism. However, favorable mutations can be maintained at intermediate frequencies, for example, when the heterozygote has higher fitness. In this case one speaks of a balanced polymorphism (as in the MHC previously mentioned). For weakly selected mutations, genetic drift and selection are important, as we shall see below (*see* Nearly Neutral Theory). Note that the above considerations are mainly applicable to recombining populations because in asexual populations and in cases with very weak selection and low recombination intensity, situations can arise that require considering patterns and forces at the level of the whole genome or of long stretches of the genome.

Models of Molecular Evolution

The main models that make predictions about the patterns of polymorphism and divergence are the neutral theory proposed in the late 1960s by Motoo Kimura and the nearly neutral theory proposed in the early 1970s by Tomoko Ohta. Those two models emphasize the importance of genetic drift and mutation on molecular evolution. Predictions from those models have been compared to empirical patterns and to predicted patterns from complex models that involve the action of natural selection. All models agree that the vast majority of newly arisen mutations are deleterious, but they differ greatly in their assumptions about the relative proportions of neutral and beneficial mutations that participate in polymorphism and divergence.

Neutral Theory

Kimura's neutral theory (Kimura, 1983) postulates that both the fixation of genetic differences between populations and the patterning of genetic variation within populations are the results of random changes in the frequency of neutral alleles driven by genetic drift. This model assumes that most new mutations are either deleterious or neutral and that advantageous mutations arise quite infrequently and are quickly fixed or lost to drift so that they are not important for polymorphism. Most of the theory's predictions for polymorphism patterns assume mutation–drift equilibrium, i.e., that the rate of allele loss to drift is balanced by the production of new alleles by mutation. The predictions therefore apply only when the population size and the mutation rate have been constant for a long period of time.

The theory predicts that the amount of polymorphism within a population is proportional to the effective population size (N) and the neutral mutation rate (μ). Nucleotide polymorphism is normally expressed in terms of the parameter θ ($=4N\mu$), which can be interpreted as the average number of differences per nucleotide site between two alleles randomly chosen from the population. This divergence is the result of nucleotide mutations that have accumulated since the two alleles last shared a common ancestor. Theory has shown that the expected time back to this ancestor for an autosomal locus in a diploid is $2N$ generations, so that when each allele has been evolving independently for $2N$ generations, the expected

number of nucleotide differences between them is $2 \times 2N$ times the neutral mutation rate μ, i.e., $4N\mu$.

The neutral theory explains the fact that different genes and regions of the genome show different levels of polymorphism by assuming that the neutral mutation rate varies among such regions and genes because they are under different functional constraints. Alternatively, levels of polymorphism can also vary across genes because the effective population size can be reduced by selection against deleterious mutations at linked and unlinked sites ("background selection"). This effect can compel neutral polymorphism to change drastically across regions that differ in recombination intensity and deleterious mutation production.

Nearly Neutral Theory

Ohta's nearly neutral theory (Otah, 1992), considered by many to be the "dominant paradigm for molecular evolution", was proposed in the early 1970s in order to account for two observations that could not be explained by the neutral model: the generation time effect for DNA divergence and the lower-than-expected levels of protein variation in natural populations. The model deviates from the neutral theory mainly because it assumes (1) that most observed molecular variation is slightly deleterious, (2) that mutation rates are correlated with the generation time (g), and (3) that the rate of molecular divergence for slightly deleterious alleles is inversely proportional to N. The nearly neutral theory is biologically realistic because it proposes plausible correlations of population genetic parameters like g and N with the origination and fate processes of molecular evolution, and because it avoids the unrealistic assumption of complete neutrality for most variation participating in polymorphism and divergence.

Under this model, slightly deleterious or weakly selected alleles, defined as those with selective coefficients close to the neutral boundary ($s \approx 1/N$), are considered major contributors to polymorphism and divergence in natural populations. Whereas the neutral theory proposes that the fate of neutral variants depends only on random drift, the slightly deleterious model builds on the fact that the fate of slightly deleterious variants depends on the joint action of both drift and selection. Which of these two processes is more important depends on the effective size of the population, because the efficacy of selection in removing slightly deleterious variants and fixing advantageous ones is higher in larger populations. In smaller populations, slightly deleterious mutations can increase in frequency and become fixed by genetic drift, but in large populations selection prevails and eliminates deleterious mutations quickly. The nearly neutral model predicts therefore that if most of the observed variation in natural populations is slightly deleterious, polymorphism patterns should be dominated by rare variants.

This model explains the lack of correlation between genetic variation and population size (especially in large populations) with the argument that when divergence and variation involve mainly slightly deleterious alleles, the amount of diversity in a population should be proportional to N but nevertheless reach a plateau when N is large enough for selection to impede slightly deleterious variants increasing in frequency. One main problem with this interpretation is that in organisms with very large population sizes one would not expect any variation if most polymorphism is due to slightly deleterious variants (at least no amino acid polymorphism), which is clearly not the case.

Selection Models

Models where natural selection is the main evolutionary mechanism shaping patterns of polymorphism have originated from two contrasting philosophies. In one approach, selection is invoked to explain patterns of variability not consistent with neutrality (genetic hitchhiking and background selection models). That approach assumes that neutral explanations may still be appropriate null models of molecular evolution. In the other approach, complex selection models have been developed to explain patterns of polymorphism that fit (or do not) neutral expectations. That approach is centered on the premise that neutral models are incorrect explanations of the molecular evolutionary process (Gillespie's models).

Levels of neutral polymorphism can be reduced by the indirect effect of positive or negative selection on neighboring (linked) genomic regions. The extent of that effect is dependent on the level of recombination in the genomic region, on the strength of selection, and, in one of the models, on the mutation rate to deleterious alleles. In the "genetic hitchhiking" model, the fixation of a beneficial allele by directional selection (a selective "sweep") leads to reduction in linked neutral variation, as any neutral variants present in the chromosome where the beneficial allele first arises will also increase in frequency very rapidly as the beneficial allele rises in frequency toward fixation (thus "hitchhiking" with the selected allele). The width of the region affected by the selective sweep is predicted to be $\sim 0.1s/r$, that is, proportional to the selection coefficient (s) of the beneficial allele that is sweeping to fixation and inversely proportional to the recombination rate (r) per base pair (Gillespie, 2004). For instance, in genomic regions with average recombination rates in *Drosophila melanogaster*, a beneficial allele with a 1% advantage ($s = 0.01$) could affect patterns of neutral variability in a region up to 100 kb around the selected site (Sella *et al.*, 2009). In the "background selection" model, selection against recurrent deleterious mutations reduces levels of neutral polymorphism at linked sites. The extent of this effect increases with the deleterious mutation rate and the strength of purifying selection, but decreases with the recombination rate. The strongest effects on reduction of variability are expected for deleterious mutations with intermediate selection coefficients, as strongly or weakly deleterious alleles will be eliminated too quickly or too slowly, respectively, to significantly reduce levels of linked neutral variation.

Selection models developed in the 1980s by John Gillespie seem to account for most of the empirical observations about molecular polymorphism and divergence (Gillespie, 1991). The models assume that balancing and episodic selection dominate the patterning of divergence and polymorphism at the molecular level. They predict that fluctuations of the environment can lead to the maintenance of polymorphism by balancing selection and, under the right circumstances (when the selective advantage of an allele is greater than the variance of the fitness effects of the alleles at that locus), to the fixation of favorable alleles. The models assume fitness effects that are

underdominant, overdominant, or that change in a random environment (Stochastic Additive Scale-Concave Fitness Function (SAS-CFF) model) or in a temporally fluctuating environment (Takahata, Iishi, Matsuda (TIM) model). The addition of environmental variability is biologically realistic since the existence of fluctuations in the fitness value of segregating alleles due to changes in the environment over ecological time scales is quite plausible. These models predict patterns of polymorphism similar to those predicted by the neutral models, which makes hypothesis testing very problematic.

Statistical Tests on Molecular Polymorphism

Of all the models of molecular evolution, only the neutral theory makes simple quantitative predictions about expected levels and patterning of genetic polymorphism. That fact makes the neutral model a very popular null hypothesis against which to test the importance of other evolutionary forces in shaping patterns of polymorphism. There are two basic families of statistical tests for neutrality: (1) models that test neutrality using polymorphism data from a single locus (below, this section) and (2) models that test neutrality using polymorphism and divergence data (*see* Combined Predictions for Divergence and Polymorphism).

The data required for the single locus tests are homologous DNA sequences collected from individuals of a single species. This family of tests compares three estimates of the neutral parameter θ, which equals $4N\mu$ for an autosomal locus in a diploid. The three estimates are or are straightforwardly calculated from (1) π, the average number of pairwise differences between the sampled alleles, which should be equal to θ, (2) S, the number of polymorphic sites in the sample, and (3) η, the number of mutations present in only one of the sampled alleles (singletons). Under neutrality the value of these estimates should be the same. The tests therefore ascertain whether departures from zero of the differences between any two of these estimates are statistically significant for the sample at hand.

Tajima, for instance, proposed to compare the values of π and θ_S (θ estimated by S) using his statistic D which is the difference between π and θ_S normalized by dividing the statistic by a factor proportional to the variance of the difference. Negative values of D are expected in the presence of purifying selection or following a purging of variation from the region by selection driving the quick fixation of a new mutation or a rapid increase in frequency of an allele previously present at a lower frequency (selective "sweep"). Positive values of D are expected with balancing selection and population subdivision.

Although the patterns predicted by the neutral theory are always used as null hypothesis, the patterns produced by the two main kinds of selection regimes (directional selection and overdominance) have also been described thoroughly, and their matching with the data has sometimes been striking. For instance, selective sweeps, the result of directional selection, tend to reduce neutral variation in regions linked to the selected site. This "hitchhiking" effect on neutral variation is more pronounced in zones with low recombination. Balancing selection increases the levels of variation over the neutral expectation in the sites adjacent to the balanced site by making the times to the common ancestor at these linked sites more similar to that of the balanced site (i.e., to the time since the establishment of the polymorphism) than to the neutral time at unlinked sites. If the polymorphism is older than $2N$ generations, one should observe an increase in neutral variation in the region of the balanced site relative to the neutral expectation.

Empirical Observations on Levels of Polymorphism

Observations about patterns of variation can be very informative when asking questions about the nature and sometimes the magnitude of the evolutionary forces generating the patterns, such as: Why do some species have more variation than others? Is it due to differences in the rate of origination of mutations? Is it due to differences in the size of their populations? Is it due to recent evolutionary, demographic, and ecological events? Is it due to the species' distributions over geographic space? Is it due to their breeding system? Is it due to the genomic context in which the mutations occur (e.g., the intensity of recombination and of deleterious mutation in and around the studied locus)?

The advent of methods for high-throughput genome sequencing has facilitated obtaining unbiased genome-wide estimates of nucleotide polymorphism for several species, although the most complete datasets still come from model species with fully assembled genomes like fruit flies and humans. Levels of nucleotide polymorphism vary widely across species (**Table 2**). Those differences are likely the result of differences in historical effective population sizes, which should influence levels of standing genetic variation as

Table 2 Levels of neutral nucleotide polymorphism in autosomal loci from different species of animals, plants, or bacteria

Organism	*Nucleotide heterozygosity per silent site*	*Estimated effective population size (N_e)*[a]
Homo sapiens	0.0011[b]	9300[d]
Pan troglodytes	0.0013[c]	21,300[c]
Gorilla gorilla	0.0016[c]	25,200[c]
Mus musculus castaneus	0.0079[d]	580,000[d]
Drosophila melanogaster	0.0232[e]	1,150,000[g]
Escherichia coli	0.0200[f]	25,000,000[g]

[a]Estimated N_e value depends on inferred neutral mutation rate.
[b]Livingston RJ, von Niederhausern A, Jegga AG, *et al.* (2004) Pattern of sequence variation across 213 environmental response genes. *Genome Research* 14: 1821–1831.
[c]Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, and Li WH (2004) Nucleotide diversity in gorillas. *Genetics* 166: 1375–1383.
[d]Halligan DL, Oliver F, Eyre-Walker A, Harr B, and Keightley PD (2010) Evidence for pervasive adaptive protein evolution in wild mice. *PLoS Genetics* 6: e1000825.
[e]Shapiro JA, Huang W, Zhang C, *et al.* (2007) Adaptive genic evolution in the *Drosophila* genomes. *Proceedings of the National Academy of Science USA* 104: 2271–2276.
[f]Charlesworth and Eyre-Walker (2006) The rate of adaptive evolution in enteric bacteria. *Molecular Biology and Evolution* 23: 1348–1356.
[g]Charlesworth (2009) Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* 10: 195–205.

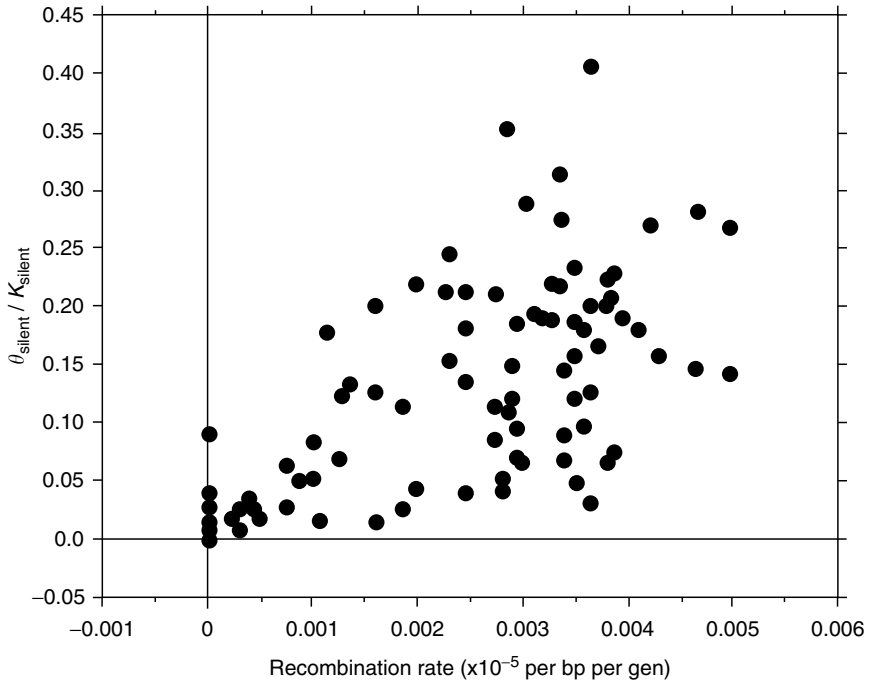


Figure 2 The relationship between silent site nucleotide polymorphism and local recombination rate in *D. melanogaster*. Silent site diversity (Θ_{silent}) is standardized by the silent mutation rate (K_{silent}). Reprinted from Presgraves D (2005) Recombination enhances protein adaptation in *Drosophila melanogaster*. *Current Biology* 15: 1651–1656.

predicted by neutral models (i.e., at mutation–drift equilibrium nucleotide polymorphism should equal $4N\mu$). However, within a species the level of polymorphism can vary extensively across genomic regions, showing, in species where those comparisons can be conducted like *D. melanogaster*, a positive correlation between levels of variability at neutral sites and the rate of recombination (**Figure 2**). That correlation is consistent with patterns expected under the genetic hitchhiking and background selection models. Although both of those models can account for the low levels of polymorphism observed in regions of low recombination, only the hitchhiking model (especially with recurring selective sweeps) can account for the overrepresentation of low-frequency alleles in those regions.

An important conclusion from estimates of polymorphism is that the full haploid genomes of two randomly sampled individuals from any of the species studied to date should show a large number of nucleotide differences: $\sim 5 \times 10^5$ differences in *D. melanogaster* or $\sim 3 \times 10^6$ differences in humans. Thus, the level of standing genetic variation on which natural selection can act is substantial.

Diversity as Phylogenetic Divergence

How do genes, proteins, and genomes differentiate, and what limits how different they could become? Are there heterogeneities in the rates of divergence among different proteins, genes, and genomic regions? Do molecules from different organisms evolve at different rates? What are the mechanisms of nucleotide and amino acid substitution? Can we infer the genealogical relationships among different organisms from molecular data? Can we use molecular divergence data to date evolutionary events? These are some of the questions addressed by researchers in the field of molecular evolution. This section will give examples and summarize the current answers to the first three questions, i.e., those most relevant to understanding the evolutionary dynamics of molecular diversity.

The Substitution Process

The fixation of new mutations gives rise to the divergence component of genetic diversity. This process describes how a new allele replaces the previously most common allele in the population, thereby becoming the common ancestor of all alleles at that locus. The dynamics of the substitution process are determined (1) by the rate of origination of new mutations in the population and (2) by the probability of fixation of such mutations. The main evolutionary forces affecting this process are mutation, genetic drift, and selection, with their relative importance differing for neutral, weakly selected, and strongly selected mutations.

Neutral autosomal mutants are produced in a diploid population at a rate of $2N\mu$ per generation, i.e., the number of new mutations produced in each generation is equal to the product of the number of gene copies in the population that can mutate ($2N$) times the mutation rate to neutral alleles per gene per generation μ. The same result applies for selected mutations if we let μ become the rate of mutation to strongly selected or weakly selected alleles.

The probability of fixation of a mutation depends on the selective advantage or disadvantage of the mutation, on the population size, and on the initial frequency of the mutation. For a newly arisen autosomal mutation the initial frequency is $1/2N$ in a diploid population, which for the neutral case

equals the probability of fixation of the mutation (since any of the $2N$ alleles present at any given moment can become the common ancestor of all alleles in the population some time later). For a nonrecessive mutation with selective coefficient s close to $1/N$, the probability of fixation is about $2s/(1-e^{-4Ns})$. If the value of s is positive (advantageous mutant) and the population size is large such that $Ns \gg 1$, the fixation probability is twice the selective advantage of the mutant ($\approx 2s$). Even in this case, however, genetic drift is the main determinant of the fate of new mutations (e.g., 98% of mutations with 1% selective advantage are lost to drift). For a deleterious mutant (negative s) the fixation probability is lower than in the neutral case (but nonzero, stressing the importance of genetic drift) and is practically zero for strongly deleterious mutations.

The rate of substitution equals the average number of new mutant alleles entering the population each generation, times the probability that one of them will fix. For a neutral allele this equals $(1/2N) \times (2N\mu) = \mu$, i.e., the rate of substitution of neutral alleles equals the mutation rate (Kimura, 1983). For selected mutations the rate of substitution equals $4Ns\mu$, and is therefore dependent on the mutation rate of favorable selected mutations, the population size and the selection coefficient of the mutant. It is easy to see that for the neutral rate of substitution to remain constant across different populations it suffices that the neutral mutation rate does not change. For selected mutations, however, the substitution rate could vary much more since its constancy requires that three parameters do not change or that they vary in ways that balance changes in the others.

Models of Molecular Evolution

From their conception most theoretical models of molecular evolution dealt with both divergence and polymorphism using an explicit population genetic perspective.

Neutral Theory

Kimura showed that genetic divergence at neutral loci should increase over time due to the fixation of neutral mutations by random drift at a rate that equals the neutral mutation rate μ (*see* The Substitution Process) (Kimura, 1983). This predicted constancy in the rate of substitution matched empirical observations showing constant rates of amino acid substitutions per year for different proteins in a wide range of vertebrate lineages that had led E. Zuckerkandl and L. Pauling to propose the molecular clock hypothesis in the early 1960s. Although the match is often viewed as positive evidence for Kimura's neutral theory, some selection models can also produce clock-like divergence (*see* Selection Models) The observed differences in rates of evolution among different genes and the heterogeneity of divergence rates in different regions of a gene are explained by the theory as being the result of differences in the fraction of all mutants that are neutral for each given gene or region (i.e., different functional constraints).

Nearly Neutral Theory

Ohta's model assumes that most molecular diversity is slightly deleterious, and predicts that the rate of production and fixation of neutral mutations is negatively correlated with organismal generation time (generation time effect), and that the rate of divergence for slightly deleterious changes is inversely proportional to the population size (Ohta, 1992). The negative correlation between generation time and mutation rate is expected when mutations occur mainly during DNA replication and the number of germinal cell replications per generation is the same across different organisms. The germinal cells of short-lived organisms should indeed experience more DNA replications per year than those of long-lived organisms so that more neutral mutations can be produced in short-lived organisms. Such an assumption corresponded well with observations from divergence studies from the early 1970s based on DNA hybridization that showed a strong generation time effect for noncoding DNA evolution. The generation time effect can be obscured in comparisons among distantly related organisms when they differ in developmental mode and germline ontogeny.

An important contradiction this theory tried to address was that protein evolution appears to be clock-like whereas noncoding DNA evolution shows a generation time effect. The neutral model explained the former by asserting that most amino acid replacements are neutral. The nearly neutral model asserted that protein divergence is mainly due to the fixation of slightly deleterious alleles, so it had to resort to postulating that generation time is inversely proportional to population size in order to account for the constancy of amino acid replacement rates. It was argued that small organisms have higher mutation rates per year due to their shorter generation times but that due to their larger population sizes, they fix slightly deleterious amino acid changes less readily. However, organisms with longer generation times have lower mutation rates but smaller population sizes so that they fix slightly deleterious mutations more easily. Therefore, the fact that across different taxonomic levels generation time correlates negatively with population size lets the rate of amino acid divergence appear clock like.

A major problem with this theory concerns divergence: slightly deleterious variants cannot accumulate indefinitely because they would decrease the fitness of the population thereby giving new favorable mutations a major role in the process. Additionally, in large populations, no divergence should occur because in them selection should be efficient enough to eliminate most slightly deleterious alleles.

Selection Models

Divergence data prompted Gillespie to propose his SAS-CFF and TIM selection models (*see* Diversity as Polymorphism: Selection Models) (Gillespie, 1991). He showed that there is excessive variation in amino acid divergence rates in different mammalian proteins using the index of dispersion $R(t)$, which is the ratio of the variance in substitution rate to the mean substitution rate. The index ranged between 2.5 and 7.5 although the neutral theory predicted it to be around 1.0 (as expected for a simple Poisson process). These high values of $R(t)$ led Gillespie to propose that the process of divergence is characterized by bursts of substitution during which alleles previously segregating at high frequencies become rapidly fixed by selection, followed by quiet periods with little substitution and long-lived intermediate frequency polymorphisms. This still allows for a molecular clock if the periods of bursting and

stasis succeed each other more or less regularly, leading to an "episodic molecular clock." Nevertheless, neither these models nor the neutral and nearly neutral models can completely account for the excessive variance of the rate of amino acid replacement in mammals. Improvements of these models should include increasing the time frame over which the fitness effects of mutations change or fluctuate. However, the complexity of the models has already reached the point that they can fit almost any observation, which compromises their value for hypothesis testing. At this point one will have to assess the plausibility of their assumptions rather than the accuracy of their predictions.

Statistical Tests on Sequence Divergence

The comparison of the rate of substitution of nonsynonymous (replacement) and synonymous (silent) mutations (dN/dS) in homologous sequences from different species have been used to infer the occurrence of natural selection during protein evolution. In a protein-coding sequence that is evolving neutrally (and under no selective constraints) the rate of substitution at nonsynonymous or synonymous sites should be the same; thus, one expects $dN/dS = 1$. Alternatively, if purifying selection is removing a fraction of nonsynonymous mutations while not removing or removing a lower fraction of synonymous mutations then one expects $dN/dS < 1$. The vast majority of protein coding genes that have been studied show evidence of purifying selection ($dN/dS < 1$). This is expected because amino acid changes often disrupt the function of the protein whereas synonymous changes are less likely to have comparable consequences. Evidence of adaptive substitutions in a protein-coding gene are inferred when $dN/dS > 1$, that is, when the rate of amino acid substitution is greater than the rate of silent substitution. Although this approach is widely used to identify genes evolving under positive selection, it has relatively low power given that evolutionary rates are compared across the whole length of the gene. For that reason, statistical methods that allow inferring selection at particular sites of the protein or along particular lineages in a phylogeny have been developed.

Empirical Observations on Divergence Levels

One of the main patterns of divergence that have been described by protein and DNA sequencing studies is that rates of amino acid replacement in different kinds of proteins and in different organisms are extremely variable (**Table 3**). Histone H4, a protein involved in the packaging of DNA in chromosomes, is a striking case of an extremely conserved protein. Comparisons of H4 histones from plants and mammals show that only two amino acid replacements, over the 100 sites compared, have taken place since the two groups last shared a common ancestor about one billion years ago. However, viral envelope proteins, which are under constant challenge by the host immune system, change at extremely high rates – between a million- to ten-thousand-times faster than the fastest mammalian protein (γ-Interferon).

Table 3 also shows that rates of substitution at synonymous sites are higher than rates at nonsynonymous sites ($dN/dS < 1$).

Table 3 Rates of amino acid and silent substitution (in events per site per year)

Organism	*Protein*	*Nonsynonymous rate*	*Synonymous rate*
Human–Rodent	Histone 4	0.00×10^{-9}	3.94×10^{-9}
	Insulin	0.20×10^{-9}	3.03×10^{-9}
	Hemoglobin α chain	0.56×10^{-9}	4.38×10^{-9}
	Interleukin-1	1.50×10^{-9}	3.27×10^{-9}
	Ig kappa chain	2.03×10^{-9}	5.56×10^{-9}
	γ-Interferon	3.06×10^{-9}	5.50×10^{-9}
Drosophila[a]	*Adh*	0.90×10^{-9}	9.50×10^{-9}
	Esterase 6	3.03×10^{-9}	21.52×10^{-9}
Plants			
Grasses	*Adh*1	0.32×10^{-9}	7.00×10^{-9}
Grasses	*Adh*2	0.89×10^{-9}	5.99×10^{-9}
Palms	*Adh*A	0.41×10^{-9}	2.61×10^{-9}
DNA viruses			
Hepatitis B	P	1.45×10^{-5}	4.57×10^{-5}
RNA viruses			
Influenza A	Hemagglutinin	3.59×10^{-3}	13.10×10^{-3}
HIV-1	*gag*	1.70×10^{-3}	9.70×10^{-3}

[a]Estimates are from comparisons between the obscura and the melanogaster groups.

Source: Reproduced from Li W-H (1997) *Molecular Evolution*. Sunderland, Massachusetts: Sinauer Associates.

However, in several proteins, amino acid replacements occur more frequently than synonymous substitutions ($dN/dS > 1$), for example, in the antigen recognition site of the Human Leukocyte Antigen (HLA) loci in humans, in fertilization proteins of abalone, in HIV proteins, in proteins involved in male reproduction (*Acp26Aa*), or associated with hybrid male sterility (*OdsH*) in *Drosophila*. In the latter locus, for instance, more amino acid changes have occurred in less than a million years than during the previous 700 million years. The rapid evolution of this gene is likely due to a change in function.

Scans for genes evolving under natural selection using the dN/dS approach have been conducted in species pairs for which annotated genome sequences are available. Humans have been a major focus of such scans given the interest in finding evidence of adaptation in genes that can provide any insights into what "make us human". For instance, a comparison of 8079 homologous genes from humans and chimpanzee identified 733 genes with a dN/dS ratio greater than 1. Genes involved in immune response, olfaction, and spermatogenesis were among the most overrepresented genes showing evidence of adaptation.

Another major pattern evidenced by the data is that different taxa can evolve at quite different rates (lineage effects). For instance, a survey of substitution rates in 54 mammalian genes shows that rodents tend to have higher rates of synonymous divergence than primates, but their nonsynonymous rates are not significantly different. Similarly, sharks show a sevenfold reduction in rates of silent mitochondrial DNA evolution with respect to primates. Finally, endosymbiotic bacteria associated with aphids show higher rates of amino acid replacement than closely related free-living strains.

Combined Predictions for Divergence and Polymorphism

Two tests based on the mutual compatibility of divergence and polymorphism patterns have been developed and used with much success to test the neutral theory. When they reject neutrality, outside information can help in further identifying the forces that might have caused the rejection. The Hudson, Kreitman, and Aguadé (HKA) and the McDonald–Kreitman (MK) tests focus on the positive correlation between the amounts of polymorphism and the amounts of divergence expected under neutrality (given the linear dependence of both patterns on the mutation rate).

The HKA test compares the correlation between polymorphism and divergence at two or more loci. Its rationale is that if loci have different neutral mutation rates, their rates of divergence and levels of polymorphism should be equally affected. Selection is therefore inferred when the ratio of polymorphism to divergence varies too much over loci, a conclusion that requires sophisticated corrections to deal with the fact that, due to stochastic reasons, different loci can have different levels of polymorphism if they are not totally linked.

The (MK) test uses a similar contrast between polymorphism and divergence but concentrates on "loci" that are interspersed with each other over the sequence, thereby avoiding the need for the above corrections and drastically increasing the statistical power. This test is mostly applied to contrast synonymous (silent) and nonsynonymous (replacement) variation between species under the assumption that these two kinds of variation are homogeneously distributed along the coding region studied. The MK test compares the ratio of nonsynonymous to synonymous divergence (dN/dS) to the ratio of nonsynonymous to synonymous polymorphism (pN/pS). Under neutrality the two ratios should be the same, but adaptive protein evolution should make dN/dS greater than pN/pS. Furthermore, the difference between the dN/dS and pN/pS ratios is now typically used to estimate the fraction of amino acid substitutions (α) that have fixed by positive selection: $\alpha = 1 - [(pN/pS)/(dN/dS)]$.

Rejections of neutrality with the HKA or MK tests can be due to departures at the divergence or at the polymorphism level. For instance, an excess of amino acid polymorphism relative to amino acid divergence should be observed if variation is slightly deleterious. However, this "excess" could very well be due to a deficiency in silent polymorphism and/or an excess in silent divergence. If the ratio of amino acid replacement to silent substitution is significantly greater than the ratio of replacement to silent polymorphism, one could infer that natural selection has been driving the fixation of amino acid mutants beyond the neutral expectation, but again this could be due to a deficiency in amino acid polymorphism and/or in silent divergence. This is why outside information must be used to identify the causes of a rejection of neutrality by these tests.

Empirical Observations for Combined Divergence and Polymorphism Levels

Application of the MK tests to large-scale datasets has found evidence of adaptive amino acid substitutions in close to 50% of the species surveyed (**Table 4**). Furthermore, in species where evidence of selection has been inferred, α, the fraction of amino acid substitutions that have fixed by positive selection, is quite high (often between 40% and 60%) (**Figure 3**). The range of estimated α values is quite similar (0–60%) regardless of what taxonomic comparison is considered (e.g., unicellular vs multicellular organisms, plants vs animals, and invertebrates vs vertebrates) and regardless of the inferred size of the populations being compared. In plants, only ~30% of the surveyed species show evidence of selection, but in every case the species with significant evidence of adaptive substitution are those with the largest effective population sizes. However, in bacteria, where every species is expected to have a very large population size, evidence of selection has only been observed in close to 40% of the species. Two of the most widely studied groups, *Drosophila* and humans, show contrasting patterns: although α is quite high (>30%) for all six species of *Drosophila* that have been surveyed, the evidence is equivocal in humans, where most studies show no evidence of adaptive substitution whereas few others suggest that α falls between 10% and 35% although the evidence is weak.

Table 4 Reported evidence of positive selection on large-scale datasets using the MK test

Group of organisms	*Number of species*	*Evidence of positive selection (%)*
Vertebrates	3	2 (67%)
Insects[a]	6	6 (100%)
Plants	13	4 (31%)
Fungi	2	0 (0%)
Bacteria	14	6 (43%)
All	38	18 (47%)

[a]All insect species are *Drosophila*.

Source: Reproduced from Fay JC (2011) Weighing the evidence for adaptation at the molecular level. *Trends in Genetics* 27: 343–349.

Outlook and Future Questions

Despite the suggestion that molecular adaptation is common in some organisms (e.g., more common in *Drosophila* than in plants or humans), the neutral and nearly neutral models may still remain useful for formulating null hypotheses. What is important is that more genome-wide polymorphism and divergence datasets are collected from a more diverse group of organisms. The use of next-generation sequencing methods will facilitate that task.

One important issue, however, that affects most models discussed here and that is yet to be satisfactorily confronted is the fact that most of the neutrality- and selection-based modeling of polymorphism and divergence is based on single-site theoretical approaches. In models with selection, single-site results, which assume that no further concomitant mutant is segregating, have been the standard, and the few efforts that have dealt with many concomitant mutations have assumed multiplicative fitness effects of mutations, i.e., that fitness is $(1-s)^i$ when one carries i mutations. The latter assumption is made out of convenience and lack of clear alternatives but is not justified *a priori* nor has it been *a posteriori*. It is indeed not clear

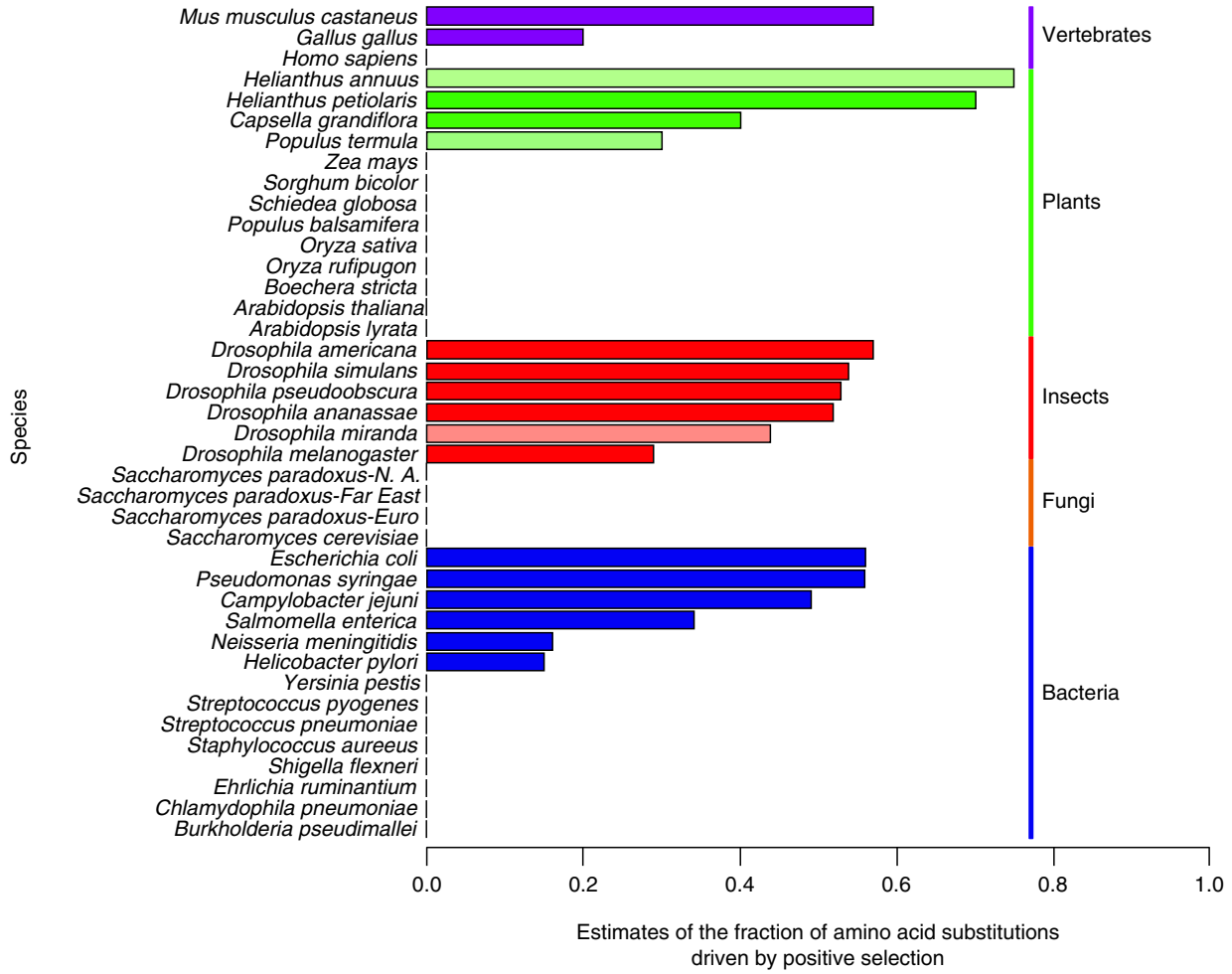


Figure 3 Current estimates of the fraction of nonsynonymous substitutions driven by positive selection (α) across different taxonomic groups. Two subpopulations of *Saccharomyces paradoxus* are analyzed separately because they show strong evidence of genetic divergence and reproductive isolation. Reprinted from Fay JC (2011) Weighing the evidence for adaptation at the molecular level. *Trends in Genetics* 27: 343–349.

how selection acts on say synonymous or insertion/deletion variation and one cannot exclude that it acts at the level of the total individual genotype via a possibly complex epistatic relationship between the number of mutations and fitness. Similarly, it is very likely that the assumption that silent variation is weakly selected ($s \approx 1/N$) is incorrect as in fact too many species with too disparate values of N nevertheless manage to harbor intermediate levels of variation for synonymous changes, and to avoid fixing the major synonymous codon or having codon usages solely determined by mutation pressures. All of this could not be possible under the slightly deleterious model if N changes by more than a factor of 3 up or down from the value $1/s$. This point is also stressed by the fact that genes on the 4th chromosome of *Drosophila* show some clear signs of biased usage of synonymous codons despite the fact that the extremely low levels of polymorphism in this chromosome indicate that its local N should be reduced well beyond 10-fold, making the maintenance of such bias not possible. Exciting new patterns of polymorphism and divergence will have to be mined in order to address this and many other important questions, an endeavor that will fundamentally bring the evolutionary and population genetic understanding of molecular diversity well beyond its present state.

The patterns of molecular diversity presented above have received attention because they facilitate the study and testing of hypotheses about evolutionary and population genetic dynamics. They are thus not necessarily representative of the phenomenon of molecular diversity in a more comprehensive organismal context. This explains, for instance, the absence of results on the evolutionary dynamics of molecular diversity with developmental consequences (albeit tentative efforts in this direction have been made during the last decade). Biologists, however, are well aware that this kind of diversity may be extremely important in evolution, but the present lack of knowledge about its behavior and nature makes it unsuitable for evolutionary and population genetic research that focuses on hypothesis testing. The narrow focus on variation suitable for hypothesis testing explains also this article's lack of an extensive presentation of some important molecular biological and biochemical constraints on molecular diversity

patterns. For instance, the dinucleotide motif TA is extraordinarily under-represented in all coding regions of *Drosophila*, yeast, and vertebrates (as is CG in vertebrate coding regions), and research on the forces underlying such motif preferences shows that they constrain molecular diversity in major ways and are rigidly conserved phylogenetically in some taxonomic groups (and thus under strong selection), but that they can also change markedly between groups or over time across chromosomes within groups.

Because these forces exert their effects on DNA without connection to gene homologies and chromosomal synteny maps, they are unapproachable using the methods presented above. It is an important upcoming task to widen our knowledge of such major and fundamental genome-architectural constraints and of how they affect molecular diversity and its dynamics in the natural-historical and organismic evolutionary sense.

The Value of Molecular Diversity and Conservation Biology

Diversity on earth is endangered and so is molecular diversity both at the polymorphism and divergence levels. Much has been written about the value of the divergence component of diversity (e.g., species richness), and also about the need to preserve the biological information contained in adaptations evolved over millions of years in species and locally adapted gene pools all over the world. Conservation biologists are aware that the polymorphism component of genetic diversity may be important for the survival of a species over evolutionary and even ecological time, and therefore they deem polymorphism worth preserving too. However, that they do it and the way in which they do it is important not only for evolutionary research but also for the other biological sciences. In fact the polymorphism patterns of many species with large effective population sizes have been shaped during millions of generations by the action of very small selection intensities. Such patterns may take too long to recreate by expanding a population and allowing selection to act for sufficiently long enough time and/or they may be the result of selective effects too weak to study with standard laboratory techniques. The information in these patterns can indirectly guide researchers to subtle but important molecular biological, biochemical, cytological, and physiological effects of the selected mutations involved but will go lost if such populations disappear or their numbers are reduced drastically. For that reason, conservation biologists should also strive to protect large "unendangered" populations.

Appendix

List of Courses

1. Undergraduate Courses on Evolution (e.g., Evolution, Principles of Evolution, Evolutionary Biology)
2. Evolutionary Genetics
3. Genetics
4. Population Genetics

See also: Conservation Genetics. Ecological Genetics. Evolution, Theory of. Genes, Description of. Genetic Diversity. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms. Population Genetics. Recombination

References

Gillespie JH (1991) *The Causes of Molecular Evolution*. New York: Oxford University Press.

Gillespie JH (2004) *Population Genetics: A Concise Guide*, 2nd edn. Baltimore: Johns Hopkins University Press.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. New York: Cambridge University Press.

Ohta T (1992) The nearly neutral theory of molecular evolution. *Annual Review of Ecological Systems* 23: 263–286.

Sella G, Petrov DA, Przeworski M, and Andolfatto P (2009) Pervasive natural selection in the *Drosophila* genome. *PLoS Genetics* 5: e1000495.

Diversity, Organism Level

Daniel R Brooks and Deborah A McLennan, University of Toronto, Toronto, Canada

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Daniel R. Brooks, volume 2, pp 191–203, © 2001, Elsevier Inc.

Glossary

Entropy An expression or measurement of the energy available for use by a system, including living systems; often described as the extent to which the system tends toward a state of disorder or randomness.

Extremophile An organism that requires, or grows optimally in, extreme environmental conditions; e.g., extremes of temperature, pressure, or acidity.

Heterotroph Literally, feeder on others; an organism dependent on organic material from an external source to provide carbon for growth.

Organism The smallest entity of life that can function as a whole, distinct from others of the same type; a single living being.

Photoautotroph An organism that uses inorganic material as a source of carbon for growth, and light as an energy source; e.g., plants.

The Nature of the Organism

The Basic Units of Life

Prior to modern evolutionary thinking in the nineteenth century, organisms were a primary focus of observation and explanation in biology. Darwin extended this tradition by reserving a central role for organisms and organismal diversity in his theory of evolution. He believed that evolutionary change resulted from the interaction of two factors, which he called "the nature of the organism" and the "nature of the conditions." Of these two, Darwin (1872) proposed that the nature of the organism seems to be much more important; for nearly similar variations sometimes arise under, as far as we can judge, dissimilar conditions; and, on the other hand, dissimilar variations arise under conditions which appear to be nearly uniform, (p. 32).

Darwin thought that organisms were historically and developmentally cohesive wholes, and therefore it was in the nature of the organism to produce offspring that were all highly similar (but not identical) to each other and to their parents and other ancestors. He also postulated that reproduction produced variation without regard for environmental conditions and therefore it was in the nature of the organism to produce these offspring in numbers far exceeding the resources available for their support. This produced *Darwin's Necessary Misfit* (Brooks and Hoberg, 2007; Brooks, 2010, 2011), an emergent property of which is natural selection. Reproductive overproduction produced variant offspring, many of which perished. Those that survived all had positive Darwinian fitness, but some were "fitter" (functionally superior) to others in the environment in which they were produced, and produced more descendants than their less fit relatives. Those merely adequate relatives, however, played a decisive role in distinguishing Darwinian selectionism from Lamarckian adaptationism. Whenever an environment changes, the most fit organisms in the old environment might not survive in the new, whereas some of the merely adequate in the old environment might have the adaptations necessary to survive in the new one. Darwin's perspective contrasted sharply with Lamarck's proposal that adaptation was an immediate and directed response by organisms to their surroundings. Lamarck also believed that the nature of the organism was important in the production of diversity, but only because all organisms have the same ability to change according to their needs. Therefore, whereas Darwin postulated that the nature of the organism included autonomous, self-regulating properties, Lamarck believed that the nature of the organism was to be directly and completely connected to the environment.

The distinction between Lamarckian adaptationism and Darwinian selectionism became increasingly blurred in the second half of the twentieth century as biologists focused more attention on parts of organisms and less on organisms as wholes. Whatever the reasons for this reductionist movement, losing the perspective on whole organisms led to a loss of Darwin's panoramic view of biological diversity. In the last quarter of the twentieth century there were many efforts to reemphasize the nature of the organism in evolutionary biology: treatises by Eldredge (1986), Brooks and Wiley (1988) and Maynard Smith and Szathmary (1995) are particularly attuned to Darwin's original assertion about the nature of the organism and the nature of the conditions (Brooks, 2011).

Organisms as Energy Flow Systems

Lotka (1913, 1925) recognized that biological systems persist in space and time by transforming energy from one state to another in ways that generate and maintain organized structure. Maurer and Brooks (1991) proposed two classes of such energy transformations. Heat-generating transformations involve a net loss of energy from the system, usually in the form of heat. Conservative transformations involve changing free energy into stored states (e.g., structure). Because all conservative transformations in biological systems are coupled with heat-generating transformations, there is a heavy energetic cost to maintaining structure. Lotka (1913) suggested

that the interplay between flow and partitioning of energy in biological systems acts to slow the rate at which energy stored by conservative transformations is degraded by heat-generating transformations.

Nonequilibrium thermodynamics allows us to generalize this perspective. Living systems are nonequilibrium thermodynamic systems; they exchange matter and energy irreversibly with their surroundings and they maintain themselves in far from equilibrium conditions. The basic features of nonequilibrium systems can be summarized heuristically as

$$dS = d_eS + d_iS, \quad \mathrm{d}_iS > 0$$

Total entropy changes (dS) can be subdivided into two components: d_eS, which measures exchanges between the system and its surroundings (changes in the surroundings), and d_iS, which measures production by irreversible processes internal to the system (changes within the system). This heuristic equation describes the "cost of living" for organisms because all organisms must take in high-grade energy and matter and dissipate lower-grade energy to their surroundings in order to survive. Energy degraded in the uptake of raw materials from the surroundings into the system is dissipated into the surroundings (d_eS). These exchanges are accompanied by a great deal of waste; hence, d_eS is very large compared with d_iS. Nonetheless, it is in the manifestations of internal production that we perceive organismal diversity.

Internal production (d_iS) includes (i) dissipation from the system, called the external dissipation function (γ_α, or heat-generating transformations). Heat-generating processes occur when energy and entropy flow in opposite directions, with entropy production tending to move the system toward disordered states and (ii) dissipation within the system, called the bound dissipation function (γ_μ or conservative transformations). Conservative transformations are characterized by energy and entropy flowing in the same direction, with entropy production being retained within the system and tending to move the system towards more structured states. In biological systems, γ_μ can be further subdivided into allocations for accumulating biomass (γ_μ^b) and allocations for accumulating genealogical information (γ_μ^i). Heuristically (Brooks and Wiley, 1988),

$$d_iS = \gamma_\alpha + \gamma_\mu^b + \gamma_\mu^i$$

Organisms thus have a dualistic nature. As open thermodynamic systems, they must simultaneously interact with their surroundings and perform critical functions internally. They maintain themselves in a viable state by exchanging matter and energy irreversibly with their surroundings, taking in relatively high-grade energy and using it to perform useful work within themselves. This requires sensing of, and causal engagement with, the surroundings mediated by a physical distinction between the organism and its surroundings. That is, there must be an "inside" and an "outside" of the organism. For all organisms, cell membranes are simultaneously physical barriers between the inside and outside of the organism and highly selective mechanisms for modulating the exchange of matter and energy between the organism and its surroundings. For multicellular organisms, this barrier is a complex of cell membranes.

Organisms as Information Systems

Discovering the genetic code made it possible to begin thinking of organisms as information systems. Information systems consist of a source of signals, a channel through which the signals are transmitted, and a receiver and converter to translate the portion of the signals that made it through the channel into information. Early discussions considered the genetic system as the source, reproduction and ontogeny the channel, and the environment the receiver and converter. Genetic possibilities become phenotypic signals as a result of reproduction and ontogeny and become meaningful biological information as a result of causal interactions between the phenotype and the environment. The environment cannot be a receiver in a Darwinian world because it does not measure or interpret; its only causal interaction with biological information is elimination of some of it. If biological information is a material part of biological systems, however, it is possible for biological systems to be their own sources and receivers (Brooks and McLennan, 1990; Brooks, 2001, 2002, 2011). Although it is true that biological systems are localized in space, they are also localized in time. In other words, the receiver can be a "time." The source is a genetic system at time t_0, the channel is reproduction and ontogeny, and the receiver is the same genetic system at any given time $t_{1\ldots n}$; thus, the receiver is temporally distinct from the source. If an information source precedes its receiver in time, it can produce the system that acts as receiver, and that system can then become a source. Information theorists use this perspective in designing self-correcting computer programs, which can enhance their own abilities to store and transmit information efficiently. The same holds true for biological systems: DNA has significant self-repair capabilities.

Biological systems are physical information systems, a type of nonequilibrium thermodynamic system, open to exchanges of matter and energy but maintaining a closed information system internally which functions to reproduce the system – to perpetuate lineages through time. They impose themselves and their functions on their surroundings and thus are self-stabilizing and self-organizing. They produce organized complexity cheaply (d_iS is small compared to d_eS, and the portion of d_iS allocated for the information system is small, in part because a small number of chemical templates are used to generate many organisms), variably (because even chemical templates are subject to the statistical mechanical vagaries of the second law of thermodynamics), and functionally (because organisms must exchange matter and energy with their surroundings in order to maintain themselves), but without regard for details of the surroundings (because the information system is embodied in relatively autonomous internal chemical production, d_iS, of the system). As the source and receiver of organized information, they can be the embodiment of the organizing principles for that information. Biological systems thus transmit information *through*, not *to*, their surroundings. This supports Darwin's view that the nature of the organism creates the necessary conditions for selection processes.

Treating biological systems as physical information systems provides a causal basis for the origin of selection. Selection processes originate as a result of the necessity that biological

systems obtain matter and energy from their surroundings coupled with the relative autonomy of their information systems, which permits production of organisms regardless of the details of their surroundings. Without the constraints provided by this autonomy, there would be no selection; simultaneously, however, constraints provide systems with properties that limit the ways in which and the extent to which the system will respond to selection.

Putting it all Together

Biological systems, beginning with organisms, are functional wholes with respect to the way they engage their surroundings and with respect to their internal organization. A major component of internal organization is functional integration through the interdependence of parts, and this is most evident in the dynamics of ontogeny and in physiological processes. Evolutionary changes in biological systems do not occur all at the same time; thus, when such changes occur, only part of the system changes. All changes, whether point mutations on a chromosome or alteration of part of a complex mating ritual, must integrate with the rest of the system, which has not changed, in such a way that viable organisms result. The functional necessity of developmental integration creates stability domains within bioinformational phase space. Orderliness and organization in biological systems result from the interaction of selection processes with three aspects of the nature of the organism: (i) historical uniqueness, (ii) cohesive properties, and (iii) hierarchical organization (Brooks, 2001, 2002, 2011).

Historical uniqueness manifests itself in historical contingency and temporal irreversibility. Important organismal processes such as reproduction, development, aging, and death are inherently irreversible. Spontaneous irreversible behavior of this sort always involves growth and increasing complexity, and physical manifestations of at least some of the systems' history.

Cohesive properties range from membrane-bound nucleic acids to cell–cell adhesion and recognition, sexual reproduction and specific mate recognition systems, and common history of inheritance. Cohesion is especially important to evolutionary explanations; it is the "glue" of functional integration and hierarchical organization that are so characteristic of biological systems. Many biological processes that demonstrate irreversible behavior manifest such changes as a result of interactions among cohesive factors.

Hierarchical organization plays an important role in organismal diversity. Hierarchies provide stability, reinforce boundaries between organisms and their surroundings, allow increasing amounts of complexity without losing organizational coherence, and provide a way in which causation and control can be tied together. Eldredge (1985, 1986) emphasized two major forms of hierarchically organized biological structure. The ecological hierarchy is manifested by patterns of energy flow in ecosystems, and the genealogical hierarchy is manifested by patterns of ancestral relationships among organisms and species. Environmental and genealogical phenomena are good starting points for investigating hierarchical interactions because they are intimately connected in biology. Prebiotic environmental conditions established the boundary conditions within which life could originate. Conversely, genealogical processes that characterize life are autonomous enough from environmental conditions to be capable of overrunning available resources and of changing the environmental conditions substantially. The longer life exists on this planet, the more it shapes the environment of the planet. Organisms are truly *Children of Time* (Brooks, 2010, 2011).

Diversity in the Real World: A Précis of Current Organismal Diversity

Organismal diversity is evolved diversity. No matter how distantly related two organisms are, they still share at least some features in common, such as cell membranes made of lipid bilayers and a common genetic code, linking them to a single phylogeny of life on this planet. Highlights of organismal diversity can thus be best recounted in a précis of organismal diversification – phylogeny. However, phylogeny is not a linear historical timeline. It is hierarchically branching and reticulating, producing organisms that are historically unique mosaics of ancient and recent traits, both functional and structural. Therefore, it is not possible to give even a précis of organismal diversity without twists and turns. The major transitions in evolution (Maynard Smith and Szathmary, 1995) provide excellent reference points. Maynard Smith and Szathmary suggested such transitions occur whenever life functions were compartmentalized and modularized in such a way that the reproductive flow of biological information was enhanced. They may enhance the survivorship of organisms in particular environments, thus increasing the likelihood that they will reproduce, or may enhance the mechanisms by which information is stored and transmitted, regardless of the particular environment in which it evolves. Maynard Smith and Szathmary highlighted what they considered to be the nine most important evolutionary transitions, all of which were examples of enhanced efficiency of information storage and transmission.

The oldest lineages of organisms on the planet are various prokaryotes, including bacteria and archaeans. These organisms are single celled, lack organelles or membrane-bound genomes, and exhibit only asexual or simple parasexual reproduction. Their cells form only rods, spheres, or spirals, and yet they have inhabited this planet continuously for almost 3 billion years and represent by far the largest number of differentiated lineages on the planet today, including the largest number of unnamed taxa. Representatives of the prokaryotes include chemoautotrophs (thought to have been the lifestyle adopted by the first living things), chemoheterotrophs, photoheterotrophs, and photoautotrophs. Photoautotrophic prokaryotes, primarily Cyanobacteria, were the first to demonstrate that organisms could change the environment of the planet, substantially affecting standing biodiversity. They did this by releasing large amounts of molecular oxygen into a reducing atmosphere, changing it in such a way that most prokaryotes became extinct or restricted to marginal habitats. This release of oxygen (the Great Oxidation Event), as well as associated nitrogen fixation, by prokaryotic photoautotrophs began about 2.4 billion years ago,

creating the conditions necessary for other life forms to evolve, so it is safe to say that prokaryote evolution played a major role in setting the stage within which all subsequent biological diversification occurred. Much later, an endosymbiotic association between a eukaryote and a cyanobacterium would lead to the evolution of the chloroplast (the modified cyanobacterium). Prokaryotes today include the so-called Archaea, sometimes called *extremophiles*, which may represent a collection of distantly related remnants of the diversity affected so much by the evolution of those photoautotrophs. Among the extremophiles are members of the halophilic *Halobacterium*, which grow best at pH 11.5, and the acidophilic *Sulfolobus*, growing best at pH 2 or 3 but able to survive at pH 0.9. There are methanogens, producing most of the planet's methane, including a host of symbionts living in the rumens of cattle and their relatives where they produce digestible carbohydrates for their hosts as well as methane. Another methanogen, *Methanopyrus*, which lives near ocean-bottom thermal vents, grows optimally at 98 °C but is able to survive at 110 °C and "dies of cold" at 84 °C. *Sulfolobus* is also a hyperthermophile, living optimally at 70–75 °C and dying at temperatures lower than 55 °C. *Thermus aquaticus*, isolated from hot springs in Yellowstone National Park, was the original source of high-temperature DNA polymerase, the enzyme used to catalyze the polymerase chain reaction (PCR) technology that has become indispensable to molecular biology. Most other prokaryotes are bacteria, which include some of mankind's most valued and most loathed organisms. Bacteria living in our digestive tracts are essential for proper nutrition, and *Escherichia coli* is the workhorse of molecular biology and biotechnology, including the pharmaceutical industry. On the other hand, diseases such as syphilis, anthrax, diphtheria, tuberculosis, smallpox, tetanus, cholera, and plague and a variety of agents of food poisoning, including *Salmonella* and botulism, are all caused by bacteria. These same disease-causing prokaryotes have also shown a tremendous degree of adaptability in the face of substantial human efforts to eradicate them. Strangely, there are no known pathogenic archaeans.

From within this great prokaryotic diversity evolved the first eukaryotes. Most eukaryote species, including the oldest surviving lineages, comprise single-celled organisms differing from prokaryotes by well-organized compartmentalization of functions within the cell. Cell energetics are handled by organelles called mitochondria and chloroplasts, which are thought to have originated as endosymbiotic prokaryotes (the mitochondrion being a modified α-proetobacterium symbiont). Digestion and excretion are handled by specialized vesicles within the cell that facilitate the movement of matter and energy in and out of the cell. Synthesis of amino acid building blocks for structural components takes place in specialized organelles called ribosomes. Finally, the storage, production, and transmission of information are compartmentalized inside a membrane-bound nucleus. Powered locomotion, first evolved by prokaryotes and often involving "appendages" (whip-like flagella, small, thin proteinaceous micro-filaments) to propel, swim, glide, twitch, and swarm through water or across surfaces, becomes highly effective in unicellular eukaryotes as a result of increasingly sophisticated cytoskeletal integration.

Most unicellular eukaryotes were once classified together in a nonmonophyletic collection of lineages called the Protista. We now know that "protists" are distributed across all of the major eukaryotic kingdoms, including the two kingdoms containing true multicellular organisms, the Unikonts (metazoans, fungi and related unicellular species such as amebas and choanozoans) and the Plantae (red algae, green algae, and land plants). Within this evolutionarily heterogeneous and diverse collection of lineages are the origins of the first sophisticated sexuality as well as alternation of sexual and asexual generations (numerous examples are found in ciliates, apicomplexans such as malaria, and various algae), the first sophisticated modifications of the cell membrane in ways that modulate organism–environment interactions (diatoms produce complex and beautifully sculpted exoskeletons), and the beginnings of true multicellularity, manifested by colonial aggregations of unicells that exhibit some degree of functional differentiation, such as the fruiting bodies of slime molds.

Single celled eukaryotes are major underpinnings of the heterotrophic portions of the biosphere. They are saprobes and detritivores, predators, and a variety of symbionts, ranging from commensal ciliates living in the tests of sea urchins and the rectums of frogs to mutualist flagellates living in termite guts, providing digestible carbohydrates for the insect hosts in exchange for their own food and a safe haven, and parasites of global socioeconomic significance, including the causal agents of amebic dysentery, malaria, giardiasis, African sleeping sickness, Leishmaniasis, and Chagas' disease. Most show enhanced reproductive output and complex life cycles, but *Giardia* and *Entameba* show some secondary losses of structures.

The photoautotrophic unicellular eukaryotes comprise four distinct lineages: dinoflagellates, red algae, diatoms, and green algae. The dinoflagellates, which owe their photosynthetic ability to an ancient endosymbiotic relationship with a red algae, are best known for causing *red tide*. Massive blooms of certain dinoflagellates produce large amounts of waste products that act as nerve toxins capable of killing large numbers of fish or of being concentrated in shellfish and killing those who eat them. The red algae include the well-known edible genera *Palmaria* (dulce) and *Porphyra* (nori). Green algae often form symbiotic relationships with other eukaryotes, including fungi, sponges, cnidarians, and various unicellar species. One species, *Oophilia amblystomatis*, lives in the jelly surrounding spotted salamander embryos, where it thrives on the nitrogen waste from the developing salamander and produces oxygen in return. More astonishingly, this alga has been found growing inside the embryo's cells where it is surrounded by mitochondria. This implies that the salamander may be using the oxygen and carbohydrates produced by the photosynthesizing algae. Diatoms are ubiquitous phytoplankton that represent the primary source of photosynthesis (and oxygen production) in the ocean, although many are freshwater inhabitants. They are characterized by having bipartite siliceous shells (tests), exhibiting species-specific forms and beautiful surface sculpting. Because they are siliceous, these tests are highly resistant to decomposition. As a result, fossil deposits of diatoms are excellent indicators of past environmental conditions. Large deposits produce "diatomaceous earth," which has many industrial uses including

insulation, filtration, and metal polishing. Diatomaceous earth has also been proposed as a "nonpolluting" insecticide. Spreading diatomaceous earth on fields may clog the tracheae of pest insects but will also clog the tracheae of beneficial insects and may also affect the breathing structures of other animals, including those humans who apply it or work in fields on which it has been applied.

True multicellular lineages appear to have evolved at least four times, in the brown algae and plants, which are photoautotrophs, and the fungi and animals, which are heterotrophs. According to Maynard Smith and Szathmary, multicellularity represents an advance over unicellularity because in multicellular organisms the partitioning of functions can occur at the level of entire cells, tissues, organs, or organ systems rather than having to take place within each cell. Complex reproductive systems can evolve because multicellular organisms can produce gametes without having to become gametes. Given the large numbers of unicellular eukaryote lineages and the small number of multicellular eukaryote lineages, we conclude that it is difficult to achieve a multicellular state, despite the evolutionary payoffs. Perhaps this is because it is difficult to coordinate a population of cells. There are no true multicellular prokaryotes (bacteria form strings, strands, and mats of individual organisms), perhaps because prokaryotes show no functional partitioning at the cellular level. If so, true multicellularity could only have arisen after unicellular eukaryotes evolved. In a complementary vein, there are no known examples of secondary loss of the multicellular state.

The evolution of true multicellularity followed four different paths. Brown algae, up to 60 m long in the case of *Macrocystis*, can be found on ocean shores with pounding surf, in relatively quiet near-shore meadows, or floating in the open ocean. The Sargasso Sea, for example, is an enormous mass of floating *Sargassum*. Although relatively species poor, brown algae can be abundant. They exhibit differentiated tissues and even organs, including holdfasts and filaments, which are analogs of roots and leaves that provide homes to a diverse array of animal life. Some have gas-filled bladders which may contain up to 5% carbon monoxide – enough to kill a human. Cellular cohesion in brown algae is provided by alginic acid in the cell walls. This is a gummy polymer of sugar acids that cements cells and filaments together and provides the glue for the holdfasts. Alginic acid is an important commercial emulsifier in ice cream, cosmetics, and other products. Larger pieces of brown algae, especially some kelp species, are considered delicious and healthy food sources for humans.

The second group of multicellular photoautotrophs are the plants, sister group of the Chlorophyta; a group containing most, but not all, of the green algae, in which aggregates of single-celled organisms showing some degree of functional differentiation has evolved more than once (e.g., *Volvox* and *Micrasterias* (sea lettuce)). Unlike the brown algae, which originated in the sea and have remained there, plants are the descendants of an aquatic lineage that invaded and diversified on land. Cellular cohesion in true plants is provided by plasmodesmata, cytoplasmic strands connecting adjacent cells. Many plants exhibit a high degree of adaptability or permissiveness with respect to genome-level evolutionary modifications. Polyploidization is common among many groups, including both autopolyploidization and allopolyploidization as a result of interspecific hybridization.

The land represented an excellent habitat for photoautotrophs because air transmits more sunlight than does water, in which light of sufficient quality for photosynthesis does not penetrate deeper than approximately 50 m. However, being on land also posed problems, one of which was structural: air provides little support compared to water. In the absence of inherent structural support, plants could only sprawl on the land – they could never rise above it. Cell walls made rigid by the carbohydrate cellulose provided the necessary structural support. Subsequent evolution of vascular tissue in a moss-like ancestor, then a sophisticated tubular transport system (tracheids) in the ancestor of the sister group of the mosses (those plants called, not surprisingly, Tracheophytes) enabled water and nutrients to be transported against gravity to considerable heights. Rigid cell walls and the tracheid system provided opportunities for tremendous vertical growth, represented by the convergent evolution of tree forms in many groups (including giants up to 110 m tall) and a tremendous flow rate of water and nutrients. A single maple tree only 15 m tall may support more than 175,000 leaves with a total surface area of more than 675 m^2, transporting up to 220 l of water per hour from the soil throughout the tree and eventually into the atmosphere through evaporation from the leaves.

Terrestrial life that is truly independent of aquatic environments must be associated with suitable reproductive modes. When ancestral plants moved onto land, they brought with them the large stationary eggs characteristic of their closest relatives. In all land plants, when these eggs are fertilized, the resulting embryos are sporophytes contained within a protective covering produced by the parent. Those ancestral plants also brought with them flagellated sperm, maintaining a dependency on aqueous surroundings for reproduction. This limited plants to living in moist environments, places in which thin layers of water or water droplets allowed sperm to swim to the egg. Only when nonaquatic male gametes (pollen) evolved in early angiosperms did plants become completely terrestrial. The evolution of pollen, however, created a new problem; since the sperm are no longer motile, the plant depended on other means of transportation, such as wind or pollinators like birds and insects, to achieve fertilization.

Once firmly established on land, major events in plant diversification were associated with increasing compartmentalization and modularization of the organism, all integrated and coordinated chemically and without a centralized nervous system. They elaborated functionally differentiated nonreproductive body parts, especially roots and leaves, that became specialized habitat for other terrestrial dwellers, creating novel opportunities for diversification of those species and coevolutionary feedback that affected their own evolution as well. Their root systems helped plants modify the soil in ways that enhanced the lives of the plants themselves but also affected the lives of other types of organisms living in the soil. Some legumes (e.g., clover, soybeans, alfalfa, and peanuts) have a symbiotic relationship with bacteria living in root nodules. The bacteria convert atmospheric nitrogen to ammonia, a more useable form of nitrogen. This process helps fertilize the soil and is the reason why many farmers rotate

crops, planting clover or buckwheat every so often to rejuvenate fields naturally. Some plants, on the other hand (e.g., black walnut), produce toxins in their roots that diffuse in the soil, inhibiting the growth of other plants. Still others, including common plants such as marigolds, secrete substances from their roots that deter plant-consuming soil nematodes and thus can be planted as a nontoxic nematicide border in gardens.

Aboveground vertical growth in tracheophytes led to a proliferation of habits – plants, shrubs, bushes, vines, and trees. This vertical structuring allowed complex communities to form, within which the various plants created their own ecological interactions by shading each other. Plants differ in the relative amount of biomass allocated to below- and aboveground structure. This led to complex ecological interactions among root systems. In the tropics, for example, most plant biomass is aboveground. Tropical ecosystems have nutrient-poor soils, so most of the productivity is tied up in the aboveground biomass, which is why deforestation can be so devastating to the ecosystem. In temperate areas in which there is more humus in the soil, a much larger proportion of plant biomass is belowground. In fact, a given plant may participate simultaneously in one set of aboveground and another set of belowground ecological interactions, just with other plants, depending on what species the plant belongs to and where it happens to be growing.

Plant organismal diversity is also associated with major modifications of the reproductive system, each of which has many variations. Primary among these was the evolution of seeds, which are embryos protected by a covering of nutrient tissue provided by the parent. These multicellular protective coats, containing cellular contributions from as many as three generations, are a source of protection and nutrition for the embryos that may allow them to withstand substantial periods of environmental harshness and still be capable of germinating. The earliest seed plants produced seeds exposed directly to the environment, called gymnosperms (naked seeds: e.g., conifers). Angiosperms (hidden seeds: e.g., flowering plants) evolved when the ovaries, containing the seeds, expanded into fruits, hiding the seeds and providing more protection. The evolution of double fertilization at more or less the same point in time, producing triploid endosperm, provided even more nutrition for the embryo. In essence one grain of pollen fertilizes an egg, while another grain fuses with two maternal nuclei, forming a nutritive tissue that is genetically distinct from the developing embryo. Coconut milk, coconut meat, corn, and wheat flour are examples of extensive endosperm.

Angiosperm diversification produced some amazing organisms. Some plants, like Venus Flytraps and Sundews, are predators. The pitcher plant, *Nepenthes rajah*, produces large urn shaped structures that can hold up to 2.5 l of digestive fluids. Small animals, mainly insects but occasionally frogs, lizards, birds, and even mammals, which fall into the fluid cannot escape, drown and are eventually digested, providing nutrition to the plant. The pitcher, in turn, is home to numerous species, including many predatory arthropods that may feed on the trapped organisms, and various mosquito larvae that may not be able to develop anywhere else. Other plants are parasitic. Indian pipe, which has lost its chlorophyll, is parasitic on fungi, which are themselves associated with tree roots. The fungi exchanges minerals for carbohydrates with the tree, and the Indian Pipe taps into that exchange, in effect "stealing" nutrition and giving nothing back. Dodder seedlings grow preferentially toward volatile chemicals released by suitable host plants. Once contact is made, the dodder sends projections into the vascular system of the host, loses its roots and become completely parasitic. The seeds of some plants need to be burned by forest fires in order to germinate, others must pass through the intestinal tracts of herbivorous vertebrates to germinate and/or be dispersed. Perhaps the most famous example of this latter interaction is "civet coffee," the most expensive type of coffee in the world. A civet eats coffee fruits, digests the sweet pulpy, outer layers then passes the seeds, which have been slightly modified by the digestive process, when it defecates to mark its territory. The final, and perhaps most recognizable, major innovation in land plants was the evolution of flowers, which are specialized and self-contained reproductive modules. Flowers represented an enormous potential for reproductive cohesion and diversification.

The evolution of plants was followed closely by the evolution of herbivores, especially among the nematodes and insects. This in turn was followed by the evolution of a variety of antiherbivore adaptations. Defenses against herbivory include both diminution and elaboration of the part of the plant being eaten, the evolution of deterrents such as thorns, and storage of toxic metabolic by-products in tissues that are the targets of herbivory such as leaves and stems. Likewise, the diversification of flowering parts is also associated strongly with the diversification of particular groups of animal pollinators, again most notably insects, although nectivorous birds (e.g., hummingbirds) and mammals (e.g., bats) also play a role.

Plants exist in many seemingly inhospitable habitats, particularly desert and boreal regions, where the evolution of resistant seeds as well as specialized reproductive, root, and stem and leaf systems allow them to survive and thrive. Cacti and other succulents represent convergent evolution of water-retaining and evaporation-resistant body forms that can survive in desert conditions. Many plants are able to flourish in nearly marine salinity conditions, although none is truly marine. Mangroves, for example, are able to excrete excess salt from the water they absorb in the tropical coastal estuaries in which they live. Plants also thrive in seasonally fluctuating environments. This is helped by the evolution of resistant seeds but is also enhanced by the ability to synchronize growth and reproduction with permissive times of the year. For example, the ability to shed leaves deciduously apparently arose as an adaptive response to seasonally dry climates but today functions in seasonally cold climates as well.

Tapping an almost unlimited energy source (sunlight), coupled with their sophisticated means of compartmentalization, has made plants extremely important to humans and to the biosphere. Terrestrial plants improve air quality by fixing carbon and producing oxygen and improve soil quality by fixing nitrogen. Plants that have adopted life in freshwater habitats are often quite sensitive to changes in water quality and are thus useful as indicators of pollution. Those having extensive aquatic root systems may also be excellent biofiltration systems; water hyacinths, for example, can actually

improve water quality. The cultivation of plants was a major event in human evolution, which ironically provided the conditions for the growth of human population and technology that now threaten our existence by threatening the biosphere. Different plants allocate substantial amounts of nutrients into root systems, especially tuberous ones, into stems and leaves, and into reproductive structures such as nuts, seeds, and fruits. Virtually any part of a plant may be highly edible for humans or their domestic animals. Simultaneously, substances functioning to deter herbivory or competition from other plants may also serve as sources of useful materials for humans, ranging from spices to pharmaceuticals. The American yew tree sequesters a compound called taxol, which has significant effects on certain forms of cancer; a close relative, the European yew, produces a similar compound called taxotene, which also has anticancer properties.

Fungi represent one of the two major groups of multicellular heterotrophs. Among living groups, fungi are apparently the sister group of choanoflagellates and animals. Cellular cohesion in fungi is provided mostly by default: The vegetative body of a fungus, called a mycelium, comprises many hyphae, which are filamentous multicellular extrusions maintaining cohesion among the cells by forming only incomplete cellular divisions, or septa, between them.

To many people, the presence of fungi indicates that something has gone wrong – one's basement is too damp or one's bread is too old. To others, fungi are something quite different: they are the major recyclers and decomposers of the biosphere, found most commonly in association with the refuse of the biosphere, including that of our own civilization. This is because most fungi are very adaptable saprobes, highly tolerant of great ranges in temperature and osmotic, especially hyperosmotic, conditions. A considerable number of fungi are parasitic and a few are even predators, attacking small animals such as soil nematodes. Fungi are very flexible reproductively, reproducing asexually or sexually with mating types not distinguishable as male or female and exhibiting an enormous diversity of variations on the general eukaryote theme of alternation of generations. They are the primary source of cellulose and lignin breakdown in the environment. All of these features make them excellent agents for bio-remediation. They can become established in highly degraded habitats and actually improve them, creating conditions that can support yet more life. Tropical forest soils, for example, lack much humus but have enormous mats of fungal hyphae forming mycorrhyzal mutualisms on and in plant root systems, which maintain vital nutrients and structural cohesion in the small amount of soil that is present. Fossil evidence suggests that mycorrhyzal associations are at least 300 million years old.

Of the numerous fungi phyla, only three are well known to most people. The Chytridiomycota, or water molds, may be the sister group of the rest of the fungi. They have recently made the news because they are implicated in massive die-offs of amphibians living above 1000 m elevation throughout the world. Many chytrids are known to be parasitic, with some exhibiting complex life cycles involving two different hosts. A few have been found that are predators of nematodes. Members of *Arthrobotrys*, *Dactylaria*, and *Dactyella* produce three-celled loops that entrap nematodes as they wend their way through the soil; once trapped, the nematode cuticle is invaded by hyphae and ingested. The Ascomycota or sac fungi contain some of the most valuable and dangerous species to human health and cultural development. The mold that tells you your bread is too old is likely *Penicillium*, the original source of the world's first broad-spectrum antibiotic, penicillin. Two other members of the genus, *Penicillium camembertii* and *Penicillium roqueforti*, are required to produce fine cheeses. Some members of *Aspergillus* are essential to the production of soy sauce and sake, whereas others can cover edible products such as peanuts with carcinogenic aflatoxins. *Saccharomyces cerevisiae* (Baker's yeast) is used to make bread, wine and beer, while morels and truffles are some of the most expensive fungi on the culinary circuit. *Pneumocystis jiroveci*, which lives in most people's lungs, causes severe pneumonia in people with compromised immune systems (e.g., HIV/AIDS or cancer patients). The sister group to the sac fungi, the Basidiomycota or club fungi, includes most of the edible fungi, the mushrooms and chanterelles, as well as fungi that are sources of legal (ergotamine) and illegal pharmaceuticals (e.g., psilocybins such as magic mushrooms) and some of history's most deadly toxins (e.g., Death cap, Destroying Angel, Deadly Webcaps). This group also includes rusts and smuts, which can cause serious diseases in agricultural plants. Many club fungi produce microscopic spores that are forcibly ejected into the air (they are called ballistospores). Being light allows the spores to be transported on the slightest of air currents, but being small means that friction (drag) will quickly slow them down, dramatically decreasing their dispersal potential. At least one species (*Sclerotinia sclerotorium*) has solved the drag problem by releasing puff after puff of spores in a synchronized wave. The ballistic release creates its own airflow, moving the spores outwards, and even allowing them to stream around obstacles. This cooperative venture among spores carries them up to 20 times farther than a single spore ejected on its own.

On their own, fungi are ubiquitous and numerous throughout the world. In addition, some fungi form mutualistic associations with green algae or cyanobacteria. These associations, called lichens, are perhaps the champion extremophiles of the multicellular world. Extremes of heat and cold and dry and wet do not seem to deter lichens, who often flaunt their hardiness by living on rocks, slowly but inexorably breaking them down into dust. They can survive long periods of desiccation, and were even shown to return unharmed after 10 days of exposure to the harsh radiation, temperatures, and vacuum in outer space.

The second group of multicellular heterotrophic eukaryotes are the Metazoa, or true animals. Animals maintain cellular cohesion by means of extracellular matrices of cell-adhesion proteins. Regulated cell growth, programmed cell death, cell–cell/cell–matrix adhesion, developmental signaling, recognition of self and cell type specialization, originated in the common ancestor of the metazoans. These major novelties were driven, in part, by rampant gene duplications; nearly 75% of animal-specific gene families arose by duplication in the metazoan ancestor greater than 600 myra. Unlike plants, which invaded the land early and diversified there, most animal diversification occurred in the ocean, and various groups have invaded freshwater and terrestrial habitats convergently. Even today, most metazoan organismal diversity is

in the sea. A great deal is known about metazoans, probably because many of them are large enough to be seen with the naked eye and active enough to engage visually oriented organisms such as ourselves. Most metazoans, including their sister group the Choanoflagellates, however, are microscopic. Highly complex and sophisticated adult organisms, such as some rotifers, may be as small as 50 μm.

The array of metazoan body plans provides clear evidence of modularization and compartmentalization. Metazoan bodies range from aggregates of specialized cells in sponges and placozoans to defined body forms composed of specialized tissues in cnidarians and ctenophores, with organ-level organization beginning with platyhelminths, and complex organ systems in the majority of the other metazoans. Body cavities emerged very early in metazoan evolution as effective ways of physically compartmentalizing organ-level modules and laid the evolutionary groundwork for specialized compartments within body cavities, called segments, metamers, and tagmata. Metazoan bodies may be asymmetrical as in sponges, radially symmetrical as in cnidarians, bilaterally symmetrical as in the majority of metazoans, or even biradially symmetrical as in ctenophores and echinoderms. Developmental patterns among metazoans tend to be highly conserved within major groups, although secondary loss of structures is not uncommon; the platyhelminthes appear to be descended from an ancestor that lost its body cavity, and a large group of parasitic platyhelminths, dominated by the true tapeworms, has secondarily lost its digestive tract as well.

Many metazoans are conspicuous because they tend to move around a lot. Modes of locomotion are diverse, but most are variations on swimming, creeping, or crawling. Sophisticated locomotion in metazoans is highly correlated with enhanced cephalization and neural integration Cephalopod mollusks have water-jet propulsion provided by specialized siphons. Complex appendages for swimming, walking, and running, evolved in the arthropods and the chordates. Indeed a few chordates like snakes and some lizards are capable of sophisticated movements without limbs. It is within these two groups that organisms capable of true powered flight emerged.

All metazoans are heterotrophs, but they exhibit the entire range of heterotrophic functions – they are saprobes, detritivores, filter and suspension feeders, predators, herbivores, commensals, mutualists, parasitoids, and parasites. Herbivory is uncommon among marine metazoans but is a major attribute of the structure of terrestrial ecosystems. Modern herbivore diversity is concentrated in the insects, the amniote vertebrates, and the nematodes. Interestingly, no metazoans ever evolved the ability to digest cellulose so many of them dine on plant cells and tissues that contain nutrients in a form other than cellulose. Animals that actually gain nutritional benefit from cellulose are hosts to a variety of prokaryotic and eukaryotic microbes capable of breaking cellulose down into less complex carbohydrates, which can be absorbed by their host. Although this form of herbivory seems to be very successful when it evolves, it does not appear to have evolved often, presumably because it is difficult to acquire the proper symbiotic microbes, provide a suitable home for them in the intestine, and provide a means of transmitting them to offspring. For example, hatchling iguanas must consume fecal matter from adults in order to acquire the necessary microfauna to digest cellulose.

Parasitism has evolved in virtually every group of metazoans. The phylum Acanthocephala is entirely parasitic, as are the vast majority of platyhelminths, including the well-known digeneans (trematodes or flukes) and eucestodes (tapeworms), and many nematodes and arthropods. There are very few parasitic species of chordates, the most famous being lampreys that attach themselves to fishes, rasp through their skin, and feed on the blood and fluids oozing out of the wound. In some deepsea anglerfishes, a male begins life as a very small, free living creature that spends his time trying to locate the much larger female. Once he finds her, he attaches, digests her skin until he fuses with her bloodstream, at which point he begins to atrophy, losing gut, brain, eyes, etc. until he is nothing but a bag of sperm, nourished, and transported by his mate.

Reproductive modes are diversified in metazoans. Most metazoan organisms are either males or females, although hermaphroditism is not uncommon in members of some of the oldest lineages. Gonads may be transient specialized cells or tissues or permanent organs and organ systems. Reflecting their ancient origins and diversification in the ocean, fertilization is plesiomorphically external in most groups. Internal fertilization, in which the female provides the aqueous medium within which the sperm can swim to the ova, has evolved multiple times. The ancient eukaryotic theme of alternation of generations is carried forward into metazoan evolution, although it drops out in several lineages convergently. In many cases, the asexual and sexual portions of the life cycle are ecologically partitioned; in most cnidarians these stages are characterized by distinct medusa and polyp forms with distinct ecologies. However, even for metazoans that do not exhibit alternation of generations, it is common for the larvae and adults to have very different life-styles. Most larval frogs, for example, are herbivores, whereas virtually all adults are insectivores. Among ecdysozoans (nematodes, rotifers, tardigrades, kynorhynchs, gastro-trichs, acanthocephalans, and arthropods), haplodiploidy is a common theme, setting the stage for extreme sexual size dimorphisms and asymmetrical sex ratios biased heavily toward females. Many metazoans are parthenogenetic, forming unisexual female lineages. In some cases, these lineages are entirely self-sustaining. For example, in some gynogenetic lizards, females court one another, with one partner displaying typical "male" behaviors, and the other typical "female" behaviors, and vice versa. This mutual courtship stimulates ovulation. In others, such as the gynogenetic poeciliid fish called the Amazon molly, sperm from males of a parental species are required to trigger embryogenesis, even though no genetic information from the sperm is incorporated into the embryo. A few metazoans, primarily tropical reef fishes, begin life as one sex, then change to the other sex depending on the social conditions around them. The production of two different sexes itself is not a simple process. Sexual differentiation is controlled by genes on sex chromosomes in many metazoans, although the particular chromosomes are not the same in all groups. In other animals, most turtles, crocodylians and many lizards, there are no sex chromosomes; offspring gender is determined by incubation temperature. Small changes in

temperature (e.g., through global climate change) can have a dramatic effect on sex ratio in these animals, even eliminating one gender altogether.

Another major theme of metazoan organismal diversity is temporal synchronization of reproductive cycles. Perhaps the most extreme case is that of the tropical Pacific polychaete called the Palolo worm, in which the entire species reproduces in a single night each year. Frogs living in deserts escape desiccation and heat by aestivating in burrows most of the year. They are awoken by the sound of rain and emerge en masse to breed during the brief, torrential rains. A more familiar example is that of anadromous salmonid fishes, in which adults return to their natal streams after spending 3–5 years in the ocean, reproduce, and die. Once they begin the movement into freshwater, salmon stop feeding and begin breaking down their own tissues for energy, all of which is used for getting to the spawning grounds and reproducing. In herbivorous insects reproductive synchronization is tied to cycles of growth and reproduction in the different plants they eat.

One of the most arresting features of animals is complex behavior, especially with respect to reproduction. This can be manifested in sophisticated interactions among males, among females, between males and females, in parental care, or even in complex combinations of all possibilities. Complex mating rituals have been documented in tardigrades, in which males are known to apply ritualized strokes to females to encourage them to reproduce. Mollusks (especially cephalopods), arthropods (especially insects), and chordates (especially vertebrates) are the best examples of sophisticated behaviors among metazoans. All three groups have evolved complex control and sensory integration involving neurochemical systems and strong cephalization.

Parental care is not restricted to mammals, nor is it primarily the domain of females. Some leeches are known to carry their offspring on their abdomens for a period of time after birth. In many species of fish, the male alone cares for the eggs and fry. Amphibians, particularly frogs, have evolved more ways to care for offspring than any other group. Some species carry their eggs and tadpoles on their backs, others swallow the fertilized eggs and the tadpoles develop in either the mother's stomach or the father's vocal sac (depending on species) and climb out through their parent's mouth when they have metamorphosed into froglets. Most caecilians, relatives of frogs, are viviparous. In some species the mother secretes a rich assortment of lipids and proteins into the top layer of her skin, which her offspring then peel off and eat using specialized teeth.

One outcome of integrated and sophisticated behaviors is sociality. Maynard Smith and Szathmary (1995) designated the evolution of social systems as one of the nine major transitions in biological evolution. Social systems permit functional compartmentalization at the level of different organisms within a social group. This level of modularization and compartmentalization is best seen in the social insects, notably bees, ants, and termites, whose colonies function very much like super-organisms. In bees and their relatives, colonies consist of many sexually inactivated females (workers) and a smaller number of reproductively active males called drones, whose only function is to mate with the only reproductively active female, the queen. Ants, also social hymenopterans, differ from bees and their relatives by having males that die shortly after reproducing with a queen and by exhibiting a tremendous range of morphologically distinct, functionally stereotyped workers and soldiers. Ant societies fascinated Darwin, especially their propensities to wage war, make slaves, use tools, construct complex domiciles, and engage in agriculture, farming fungi and milking aphids for nutrients. Termites are the third major group of social insects. Their colonies comprise a single queen and king who do all the reproducing. Sexual inactivation of offspring is maintained from one generation to another by pheromones spread through trophallaxis, which is the mutual feeding of nymphs. Workers and soldiers produce pheromonal secretions whose proportions in the colony act to maintain suitable proportions of each type of individual in succeeding generations. One of the strangest examples of social caste systems has just been recorded in digeneans, parasitic flatworms with a complicated life-cycle that includes the repeated clonal production of infective larvae inside a snail. There are two types of infective larvae: large, slow moving, reproductively active primary morphs and small (about 64 times smaller), thin, very active secondary morphs. The secondary morphs do not reproduce. Instead, they are specialized for recognizing and attacking larvae from different flatworm species. These "soldiers" cluster in areas of the snail that are the most likely to be penetrated by other larvae. When an invader is detected, the soldier attacks, grasping the invader with its large muscular pharynx, either swallowing it whole, or lacerating it and sucking it dry.

Some vertebrates, including many primates, also form social systems. These tend to be much less stereotyped than those of the insects, with more flexibility in roles in which individuals perform multiple tasks and all are reproductively competent. Members of vertebrate societies never lose their individuality. Actual reproductive patterns are constrained by behavioral dominance hierarchies, both male and female, that can be quite complex and can be based in large part on kinship. None, other than humans, come close to the frenetic and compartmentalized social activity seen in the colonies of social insects.

The predominant metazoan groups are the chordates, nematodes, arthropods, mollusks, annelids, and echinoderms. The most organismally diverse group is the Ecdysozoa, which includes the nematodes and the arthropods. Within the arthropods, the insects alone represent more than 1 million named species, with many more remaining undescribed. Nematodes may be as species rich as insects, but only 15,000 nematodes species have been named. Feats of relative strength and tolerance of extreme environmental conditions are legendary among the insects; among the ecdysozoa, however, tardigrades may be the most amazing extremophiles of all, being capable of entering a state of cryptobiosis. Under desiccating conditions, tardigrades can experience a loss of total water content from 85% of their body weight to 3%. Furthermore, while in this state, tardigrades can withstand, sometimes for years, extreme temperatures, including cold to near absolute zero, ionizing radiation, almost total lack of oxygen, and up to 1000 atm of pressure. When moisture levels return to normal, normal life functions resume.

Although they represent a relatively small percentage of the total species on this planet, metazoans have a major impact on

the biosphere. Early in the evolutionary history of metazoans, reef-building sponges and corals evolved, creating biogenic alterations in ocean current patterns and novel forms of habitat for additional life forms. Benthic metazoans make use of, and help maintain, the ocean bottom while annelids, nematodes, rotifers, and a variety of arthropods perform the same function for terrestrial soil systems. Many metazoans serve as habitat for other metazoans: in the life-cycle of an individual tapeworm, a copepod, a mollusk, a ray-finned fish, and an elasmobranch may all serve as habitat. The many metazoans that are filter or suspension feeders may act as bioaccumulators. In some cases, their activities may enhance local environmental quality, but this may be a mixed blessing. Oysters and other edible mollusks, for example, are excellent at extracting bacteria and other microbes from the water, but if those microbes include fecal coliform bacteria from human sewage the value of the mollusks as food is severely compromised. The array of complex ecological interactions between plants and insects has permitted the classification of insects as "beneficial," meaning that they do not compete with us for plant biomass, "pest," meaning that they do compete with us for plant biomass, or "vector" when they transmit diseases to us or our livestock. We use parasitoid insects as agents of biocontrol by infecting pests and vectors with them.

Summary

Organismal diversity is not just amazing, it is important. The extent to which parts of today's planetary array of organisms can be reduced without creating a cascade of extinctions is unknown but hotly debated. As well, virtually every species has some value to humans, directly or indirectly. Organisms provide us with food, shelter, raw materials, things of beauty, and pharmaceuticals; help us with our work; and maintain the ability of our air, soil, and water to sustain life, including our own. They are valuable to us indirectly as the building blocks of the biosphere within which we originally evolved and which we still require in order to survive.

There is also clear evidence that all organisms are intimately tied together in the structure of the biosphere because they are all simultaneously parts of larger genealogical and ecological wholes. The significance of the duality of organismal diversity is most apparent in the recognition that new types of organisms are derived from preexisting organisms, while at the same time almost all organisms make extensive use of the biodiversity that predated their origins. Newly evolved organisms always have an impact on preexisting ones. Because it is in the nature of the organism to be relatively autonomous from its surroundings, these interactions are not necessarily positive, often taking the form of "conflicts of interest." The evolutionary resolution of these conflicts of interest has produced an increasingly complex biosphere. As Maynard Smith and Szthmary (1995) observed, each major transition in evolution has been associated with the emergence of organisms, and by extension the entire biosphere, with enhanced abilities to produce, maintain, and transmit information cohesively and also associated with the emergence of novel forms of selection resulting from the evolution of those new organisms. In this way, each newly evolved form of organism becomes intimately involved with both local and global ecology, maintaining the biosphere as a relatively isolated system with its own windows of viability (Ulanowicz, 1997).

See also: Adaptation. Archaea, Origin of. Biodiversity, Origin of. Darwin, Charles (Darwinism). Diversity, Community/Regional Level. Diversity, Molecular Level. Eukaryotes, Origin of. Thermophiles, Origin of

References

Brooks DR (2001) Evolution in the information age: Rediscovering the nature of the organism, Semiotics, Evolution. *Energy and Development* 1: 1–26. http://www.library.utoronto.ca/see

Brooks DR (2002) Taking evolutionary transitions seriously. *Semiotics, Evolution, Energy and Development* 2: 6–24. http://www.library.utoronto.ca/see

Brooks DR (2010) Sagas of the children of time: the importance of phylogenetic teaching in Biology. *Evolution: Education and Outreach* 3: 495–498.

Brooks DR (2010) The mastodon in the room: How Darwinian is neo-Darwinism? *Studies in History and Philosophy of Biological and Biomedical Sciences*. http://dx.doi.org/10.1016/j.shpsc.2010.11.003.

Brooks DR and Hoberg EP (2007) Darwin's necessary misfit and the sloshing bucket: The evolutionary biology of emerging infectious diseases. *Evolution: Education and Outreach* 1: 2–9.

Brooks DR and McLennan DA (1990) Searching for a general theory of biological evolution. *Journal of Ideas* 1: 35–46.

Brooks DR and Wiley EO (1988) *Evolution as Entropy: Toward a Unified Theory of Biology*, 2nd edn. Chicago: University of Chicago Press.

Darwin C (1872) *The Origin of Species*, 6th edn. London: John Murray.

Eldredge N (1985) *Unfinished Synthesis*. New York: Columbia University Press.

Eldredge N (1986) Information, economics and evolution. *Annual Review of Ecology and Systematics* 17: 351–369.

Lotka AJ (1913) Evolution from the standpoint of physics, the principle of the persistence of stable forms. *Scientific American Supplement* 75: 345–346. 354, 379.

Lotka AJ (1925) *Elements of Physical Biology*. Baltimore, MD: Williams & Wilkins.

Maurer BA and Brooks DR (1991) Energy flow and entropy production in biological systems. *Journal of Ideas* 2: 48–53.

Maynard Smith J and Szathmary E (1995) *The Major Transitions in Evolution*. Oxford: Freeman Spektrum.

Ulanowicz RE (1997) *Ecology, The Ascendent Perspective*. New York: Columbia University Press.

Diversity, Taxonomic versus Functional

John C Moore, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Functional diversity A component of biological diversity that enumerates the different types of processes in a community that are important to its structure and dynamic stability.
Functional group A group of species that utilize similar resources; synonymous with *guild*.
Guild A group of species that utilize similar resources (usually food).
Species evenness A measure of the relative abundance of species in an area.
Species richness The number of species in an area.
Species turnover The change in the composition of species in an area due to the extinction of some species and the replacement by a new species through colonization.
Taxonomic diversity The number and the relative abundance of species in a community.

Species, Communities, and Ecosystems

Linnaeus (1735) and Hutchinson (1959) provided a convenient starting point to address biodiversity in terms of taxonomic and functional diversity. Linnaeus formalized the two-kingdom scheme of Plantae and Animalia and made the distinction between sessile photosynthetic plants and motile food-ingesting animals. Predating an understanding of photosynthesis or modern evolutionary theory, this work was descriptive and simply attempted to categorize and count organisms rather than posit mechanisms. Function was at the basis of the classification. In his address to the Society of American Naturalists, Hutchinson, like Linnaeus, made a deliberate distinction between primary producers and heterotrophs by asking "Why are there so many kinds of animals?" The unique feature of this address was that it studied biodiversity by invoking principles from ecology and evolution, formalizing niche theory, and including energetics. More important, it marks an early attempt to link species diversity to the functioning of an ecosystem.

The choice of taxonomic versus functional diversity illustrates a broader issue in science. In both cases, traditional taxonomies based on individual traits or ones based on the functional attributes of a species in a community or ecosystem are human constructions. Ehrlich and Holm (1962) summarized the dilemma nicely:

There seems to be no theoretical reason why there must be complete congruence among estimates of relationships based on characters from different developmental stages or on characters from different organs systems of the same stage…. In a taxonomy based on ecological requirements, whales will be more closely related to sharks than to bears. Such a relationship is no more or less "true" than the classical one: it is merely based on different attributes.

At some point, the question comes down to the type of information that is used, the way it is used, and why it is used. The transition from phylogenies based on taxonomies of morphological characteristics to ones based on taxonomies of genetic information have revealed ecological and functional coherence that Ehrlich and Holm (1962) envisioned for higher taxonomic rankings (phyla, orders, and families), if not for species.

Different Approaches to Studying Biodiversity

There are different subdisciplines within ecology: autoecology (individual and population ecology), community ecology, and ecosystem ecology. The focus and approach of each subdiscipline and the type of information each uses have influenced the degree to which taxonomic or functional diversity have been used (**Table 1**).

Individual and Population Approaches

Autecology is the study of individual organisms. The approach originally focused on the adaptiveness of an organism's physiology to the environment but has since been expanded to include the study of the distribution and dynamics of populations. In terms of biodiversity, autecology has embraced taxonomic diversity.

The Community Approach

The community approach defines diversity in terms of the number of species and relative abundances of species. This information is used to study the effects of the number and types of interactions and the consequences that the interactions have on the distributions and abundances of the species and the dynamic stability of the community. Community ecology tends to focus more on taxonomic diversity than on functional diversity. An ideal description of a community would include all species and species interactions. In practice, descriptions of communities include a mixture of individual species and aggregations of species. The aggregations have been based on functional attributes defined by similarities in resource utilization among species and on taxonomic affinities (e.g., "mammals").

The Ecosystem Approach

The ecosystem approach defines diversity in terms of processes and function within the community. The focus is not on species *per se* but on the products of species interactions, usually different forms of carbon or nitrogen. Underlying this approach is the belief that species interact in ways that reinforce stabilizing feedbacks that operate to maintain a persistent steady state. The feedbacks operate through the availability of nutrients in one pool that are critical to another pool.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00036-8

Table 1 Information used by different subdisciplines of ecology

	Subdiscipline		
	Autecology	*Community ecology*	*Ecosystem ecology*
Focus	Individuals and populations Species	Populations Species and guilds	Forms of matter Functional groups
Units	Individuals	Individuals	Biomass
Biodiversity	Taxonomic	Taxonomic and functional	Functional

A typical description of an ecosystem does not include many species (except for the dominant species) but rather aggregates of species that are similar in the ways in which they process matter relative to a key process. For example, the decomposition of pine needles on a forest floor is a complex process that is influenced by the nutrient quality of the pine needles, the temperature and moisture of the forest floor, and the organisms that are present. For the microbes alone, the process might involve more than 100 species of saprobic fungi and an equal number of strains of saprobic bacteria. Depending on the level of resolution desired, the microbes could be aggregated into a single pool of *decomposers* or separated into pools of *fungi* and *bacteria*. The justification for the seemingly casual treatment of the individual species is that the focus is not on how many species are present but on the general composition of the pool and the rate at which the pool processes matter.

Species, Function, and the Ecological Niche

Taxonomic and functional diversity use similar information but in different ways. The species is central to taxonomic diversity. Species are arranged into phylogenetic relationships that are based on how they acquire energy, their cell types, the level of cellular organization, their embryology and life history, their physiology, and their genetic makeup (RNA or DNA or both). Different classification schemes have been proposed that use all or some of these criteria with different weights given to any one of the criteria. Ecosystem processes are central to functional diversity. Functional diversity has been based on ecological characteristics. Species are aggregated into functional units based on how they use resources. Various schemes have been proposed to aggregate species into functional groupings. Species are arranged into functional relationships based on how they acquire energy, where they acquire energy, the rates at which they process energy (physiology), and when they process energy (life history).

Linking taxonomic and functional diversity would require integrating information that describes the makeup of an organism (its taxonomic characteristics) along with the rates at which it processes and transforms matter (its functional characteristics). Niche theory offers a way to link taxonomic and functional diversity.

The Ecological Niche

The ecological niche of a species has been defined by the resources it requires and by its role or function in the environment. The concept of the ecological niche when introduced by

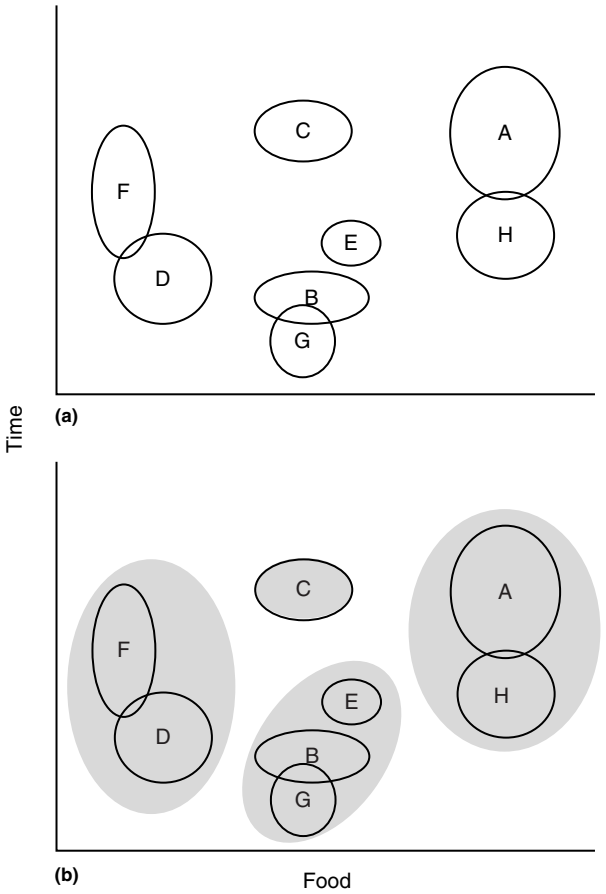


Figure 1 Projection of eight species (A–H) onto two of the principal niche axes (food and time). (a) The ovals surrounding each species represent the region of the food and time space that each species occupies. Ovals that overlap one another represent regions in which the species share (compete for) resources. (b) Ecologists may aggregate species that occupy similar niche spaces into functional groups (species within the shaded ovals).

Grinnell (1917) focused on the habitat requirements of a species for it to survive and reproduce. Hutchinson (1959) expanded on this definition of niche by disaggregating the habitat into the multiple resources it embodied. The physical space a species occupies, the temperature and moisture conditions of the space, and the seasonality in abiotic and biotic conditions that the space experiences, along with the food requirements and the interactions that a species engages in with other species are considered. Here the niche of a species can be viewed graphically as an n-dimensional hypervolume in which each resource represents an independent axis (**Figure 1**).

The space that each species occupies within this hypervolume is defined by its basic resource requirements (i.e., the fundamental niche) and is winnowed down by antagonistic interactions with other organisms (i.e., realized niche). Schoener (1974) presented an analysis of the resource requirements for several species from different taxa, revealing that the principal niche axes are food, habitat use, and time.

Elton (1927) defined the ecological niche in terms of the role of a species in the environment. This definition is subjective but not at odds with resource-based definitions offered by Grinnell (1917) and Hutchinson (1959); rather, it is conditioned on them. The role, or in keeping with this article, function of a species is a property that emerges from the adaptations of species in meeting their resource requirements to survive and reproduce. As for information, both the requirements-based and functional-based definitions of niche are grounded in the genetics of the species, providing a means to link species diversity and functional diversity.

Functional Groupings of Species

Both the community and the ecosystem perspectives have grouped species into larger units when studying interactions among species or processes mediated by species. The schemes used have ranged from those of convenience, which are usually based on a broad taxonomic category, to ones that have adopted strict criteria that may include organisms from different taxonomic groupings. For this discussion, the groupings that have been based on the resource requirements of a species and its position in niche space relative to other species makes the most sense, given that function in an ecosystem context is directly related to how species utilize resources. Hence, a functional group is composed of species that utilize similar resources (i.e., the resource-based niche definition) and perform similar roles (i.e., the functional-based niche definition) in the environment (**Figure 1**). There is no standard number of resources (axes in niche space) used to establish functional groups, but many are based on the principal resources of food, habitat, and time. This has resulted in the groupings being given different names. For example, a *guild* is defined as a group of species that feed on similar food in a similar way. Soil ecologists define a *functional group* as a group of species that feed on similar food in a similar way, occupy the same habitat, and possess similar life histories.

Estimating Diversity

Procedures to estimate species diversity originated from the fields of information theory and statistics. If presented a code with a set number of symbols (s) of known proportions (p), it is desirable to estimate its information content. Ecologists simply substituted the species (or aggregations of species) for the symbols. Species diversity of a community could be estimated in the same way as the information content of a code. Two aspects of species diversity are the number of species (species richness) and the relative abundance of species in a community (species evenness).

Species Richness

Species richness (S) is the number of species within a defined region. The species richness of a region is obtained through sampling or via a census. Because "region" is defined by the observer, species richness has been further categorized into three components to account for changes in spatial scale.

Alpha diversity, sometimes referred to as *point diversity*, is the species richness that occurs within a given area within a region that is smaller than the entire distribution of the species. *Beta diversity* is the rate at which species richness increases as one moves in a straight line across a region from one habitat to another habitat. In other words, it is the rate of change in species richness that occurs with a change in spatial scale. *Gamma diversity* is the species richness within an entire region. As the area being surveyed approaches that of the entire region, alpha diversity approaches gamma diversity and beta diversity approaches zero.

Species Evenness

Species evenness takes into account the number of species and the relative abundance of species in a community. Several indices have been proposed. Two of the commonly used measures of evenness are the Shannon index (H) and the Simpson index (D).

The Shannon index (H) is a measure of the information content of a community rather than of the particular species that is present. The index is as follows:

$$H = -\sum_{i=1}^{S} p_i \log p_i$$

where s is the number of species in the community (species richness) and p_i is the relative proportion of species i. Hence, if two communities (A and B) with the same number of species (not necessarily the same species) present were compared, and the distributions of p_i were the same, then $H_A = H_B$. If either the number of species or the proportions differed, however, then $H_A \neq H_B$. In the latter case, the community with the greater number of species whose relative abundances were equal would possess the higher diversity.

The Simpson index (D) measures the dominance of a multispecies community and can be thought of as the probability that two individuals selected from a community will be of the same species. The Simpson index was originally proposed as follows:

$$D = -\log \sum_{i=1}^{s} p_i^2$$

where S is the species richness of a community and p_i is the relative proportion of species i. The index can be modified to $1-D$ to give it the property of increasing as diversity increases (the dominance of a few species decreases).

Linking Species Richness and Species Evenness

Species richness and species evenness are special cases of dichotomous-type rarity measures. For dichotomous-type rarity measures for a pool of s species, the rarity of species i, whose

function is denoted $R(p_i)$, depends only on its relative abundance (p_i). The rarity measures (Δ) take the following form:

$$\Delta = \sum_{i=1}^{s} p_i R(p_i)$$

The relationships between species richness (S), the Shannon index (H), and the Simpson index (D) are presented in Table 2. Notice that the indices share three important properties: (1) Δ is maximized when $p_i = 1/S$ for all i species; (2) given that $p_i = 1/S$, the community with the higher species richness has the greater diversity; and (3) $\Delta = 0$ if the community possesses a single species ($S = 1$).

Biodiversity and Systems Theory

One of the enduring tenets of ecology in the early part of the 20th century was that a diverse community was a stable community. The linkage between biodiversity and stability has an intuitive appeal. The more species there are in a community, the more likely they would be tightly networked, and hence the more stable the community would be. Interestingly, function plays an important role in this argument in that in a highly diverse community, there would be alternative species or assemblages of species to perform key tasks to sustain the cycling of matter and the current diversity. These ideas have been challenged and modified. Ecologists agree that biodiversity is important to the function and stability of an ecosystem. The current debates concerns the nature of the relationship between biodiversity, function, and stability.

Although useful, the functions used to estimate biodiversity that were presented previously are descriptive measures and imply neither the mechanisms that shape biodiversity nor the importance of biodiversity to maintaining the structure and dynamic stability of a community. If the diversity of one community were higher than that of another, does this increase its likelihood of persisting in time or recovering from a disturbance? The answer depends on the species involved. For example, the colonization of an exotic species increases the species richness of a community but does not necessarily increase its stability – to the contrary, it often destabilizes the system. The addition of a predator, however, can stabilize an ecosystem by tempering the oscillations in the population densities of its prey through time. Biodiversity is a property of complex dynamic biological systems. The steady state and stability are two concepts important to the study of complex dynamic systems.

Table 2 Rarity functions and forms of diversity indices

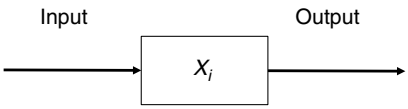
	Rarity function	*Diversity index*
Species richness (S)	$R(p_i) = (1/p_i) - 1$	$\Delta = S - 1$
Shannon index (H)	$R(p_i) = -\log p_i$	$\Delta = -\sum p_i \log p_i$
Simpson index ($1 - D$)	$R(p_i) = 1 - p_i$	$\Delta = \sum p_i(1 - p_i)$

The Steady State

A steady state is an equilibrial condition in a dynamic process in which the rate of input of a state variable is equal to the rate of output of that variable (Figure 2). The dynamic process can involve any variable that changes in time. For taxonomic and functional diversity, the variables of interest are the number of species or functional groups, the densities and biomass of each species or functional group, and the mass of key elements within different pools within the ecosystem.

The Number of Species

The number of species in a community is at steady state when the rate of colonization equals the rate of extinction (Figure 3). The classic model of this process, the theory of island biogeography, assumes that the rate of colonization is affected by



Number of species	S	Colonization	S	Extinction
Species composition	[Si]	Colonization	[Si]	Extinction
Population size	N	Births Immigration	N	Deaths Emigration
Community structure				
	[Ni]	Births Immigration	[Ni]	Deaths Emigration
	[Si]	Colonization	[Si]	Extinction
Ecosystem process Carbon flux	C	Photosynthesis	C	Respiration

Figure 2 Diagram and list of the processes (inputs and outputs) that contribute to the dynamics of a single state variable (X) or several state variables [X_i].

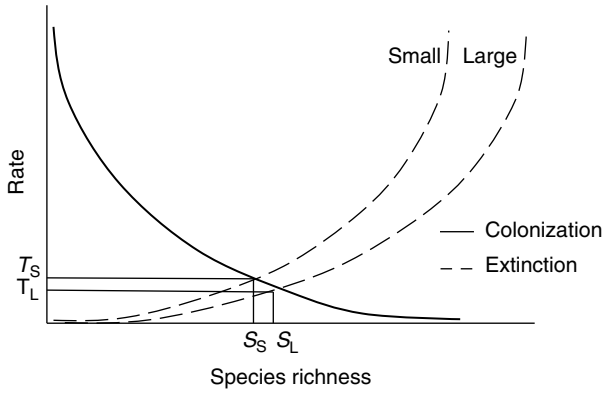


Figure 3 The effect of the rates of colonization and extinction on the species richness of large and small islands (S_L and S_S, respectively). T_L and T_S represent the points at which the rates of colonization and extinction are equal. Adapted from MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*, *Monographs in Population Biology*, vol. 1. Princeton, NJ: Princeton University Press.

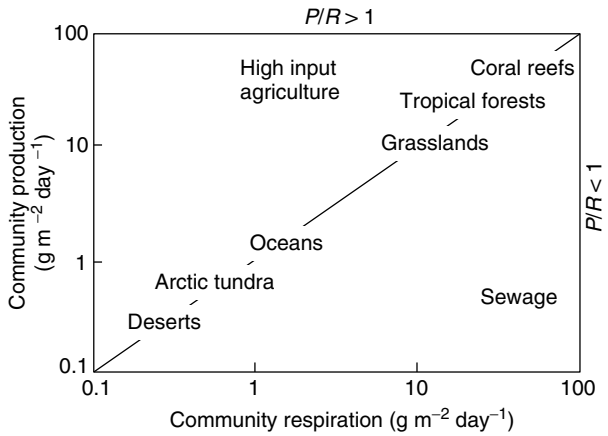


Figure 4 The relationship of various ecosystems in terms of their community production (grams per square meter per day), *P*, and community respiration (grams per square meter per day), *R*. The dashed line represents the steady state in terms of *P* and *R* ($P/R=1$). Note the position of the disturbed ecosystems.

how far the habitat being colonized is from the source of species and by the size of the habitat being colonized. Colonization rates are higher for large habitats that are close to the source of species than for smaller distant habitats. The rate of extinction is affected by the size of the habitat. The rates of extinction for smaller habitats are higher than those for larger habitats.

Species Composition

In the model presented previously, although the number of species is constant at steady state, the species composition of the community does not remain constant. The model simply assumes that once the rates of colonization and extinction are equal, then for every new species that enters a habitat, one of the current resident species goes extinct. This point is important when trying to estimate the diversity of a habitat by sampling the habitat over time. If a habitat were sampled long enough, the estimate of the number of species within the habitat would equal the number of species in the habitat serving as the source, even though the source has a higher number of species at its steady state than the habitat being colonized.

Population Densities

The population density of an individual species within a habitat is at steady state when the number of births and immigration equal the number of deaths and emigration. When modeling this process, ecologists typically assume that the habitat is closed to immigration and emigration and focus entirely on the rates of birth and death. More sophisticated approaches that involve multiple habitats will include the immigration and emigration of species among habitats.

Community Structure

A community is at steady state when the species composition of the community and the densities of the species remain constant. This implies that the species richness and evenness of the community do not change. This condition is rarely met since the species composition and the densities of the species fluctuate through time.

Ecosystem Processes

An ecosystem process is at steady state when the rate of input of matter or energy is equal to the rate of output. The input of matter and energy occurs via photosynthesis and from sources external to the ecosystem (e.g., leaves entering a stream). The export of matter and energy occurs through heterotrophic respiration and the physical movement of matter from the ecosystem (e.g., export of leaves from forest to a stream). The production of organic matter from inorganic matter (photosynthesis) and the transformation of organic matter to other organic forms and inorganic matter (heterotrophic respiration, hereafter referred to as *respiration*) are ecosystem processes that approach a steady state and illustrate the interplay between matter and energy. When the rate of photosynthesis equals the rate of respiration, the system is said to be at steady state (**Figure 4**). In other words, at steady state the ratio of production to respiration for matter and for energy is 1 ($P/R=1$).

Stability

A population, community, or ecosystem is stable if it can persist through time and can recover from minor disturbances. A disturbance is any perturbation to the system that causes any of the species or functional groups to deviate from their steady state. There are many facets to stability and many operational definitions as well. Aspects of stability include local asymptotic stability and short-term transient dynamics.

Local Asymptotic Stability

An ecosystem possesses local asymptotic stability if over time it returns to its original steady state following a minor disturbance. The minor disturbance is relative to the sizes of the populations involved and occurs in the neighborhood of the steady state, hence the reference to "local." This concept of stability has been widely used by biologists because it is simple and exact in its interpretation. It has been criticized for the same reasons that it has been accepted. For example, high species diversity has been shown to negatively impact the local

stability of model ecosystems. This result was the basis for the hypothesis that dynamic stability constrains biodiversity. However, competing models of ecosystems with different sets of assumptions and criteria for stability, although they do not disagree that dynamic stability affects biodiversity, yield different results with regard to the nature of the relationship.

The concepts of local stability and the theory of island biogeography may seem at odds with one another, even though they rely on similar mathematics and assumptions and are modeling the same systems. An obvious question is, how can an ecosystem be locally stable if the species composition of the ecosystem is changing? The discrepancy can be clarified by revisiting the processes that are being modeled. When modeling populations within a community, local stability is in reference to the densities or biomass of the populations. The dynamics of the populations are governed by the difference between the rates of birth and death. When modeling the species diversity within the same community, local stability is in reference to the number of species that are present. The dynamics of the number of species is governed by the rates of colonization and extinction. To carry this a step further, the local stability of the carbon or nitrogen cycles within an ecosystem is in reference to the mass of the different forms of carbon or nitrogen. The dynamics of the different forms is governed by the functional attributes of the species present and the rates at which the different forms are acted on.

Transient Dynamic Behavior

Transient dynamics represent the short-term dynamic behavior that asymptotically stable systems exhibit following a minor disturbance. The nature of and degree of the transient response can influence the stability and diversity of the system. Transient behavior includes the amplification and decay of a perturbation, represented by an amplification envelope, $\rho(t)$, encompassing the magnitude and extent of the deviation and its temporal components (**Figure 5**). Following a disturbance, the perturbation to the system, now portrayed as deviations from equilibrium, amplify to a maximum ρ_{max} at time T_{max}, with a rate defined by its reactivity during the short-term transient phase.

Eventually, the perturbation decays to the long-term asymptotic phase at a time defined by its return time at a rate defined by its resilience. The principle of resilience can be applied to any state variable present in an ecosystem: population densities, the number of species, or the mass of nutrients. Resilience is a function of the rates of inputs of the state variables, the rate at which the state variables turn over, and the number of state variables. The return time of a system is a measure of its resilience and is estimated as the time between the point when the disturbance altered the ecosystem and the point at which the ecosystem recovers. More resilient ecosystems have shorter return times than less-resilient ones.

The persistence of a system and hence its diversity may be compromised if the system possesses a highly reactive transient phase or if the frequency of disturbance is too great given the length of return time because a system may not have time to recover from any one incident. Under these circumstances, the perturbations may grow initially and the populations may oscillate wherein the densities of individual species may approach zero, making them vulnerable to future disturbances, or some may dip below a lower threshold where the likelihood of successful reproduction is compromised. Frequent disturbances of reactive systems may also push the system and its dynamic state from one state and attractor to another. Again, although there is uncertainty of the effects of this type of transition on the system, a consequence is the loss of species or the introduction of new species.

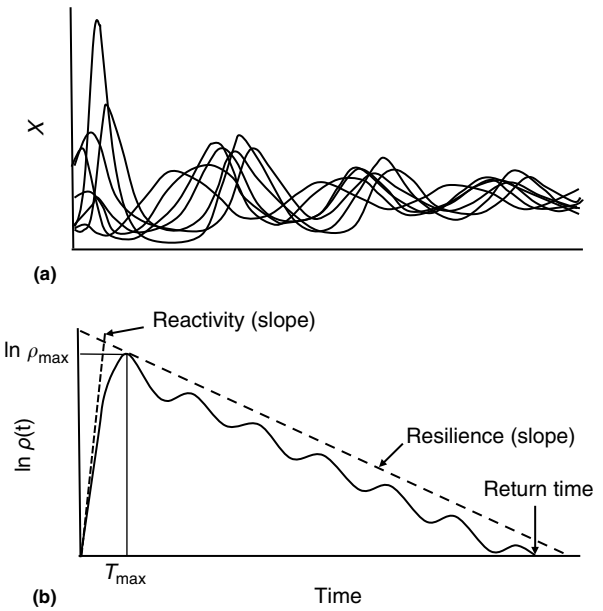


Figure 5 An overview of transient dynamic behavior. (a) The transient dynamics of a multispecies system that has been disturbed as it returns to equilibrium. (b) A summary of the components of the transient behavior of a reactive system in terms of the amplification envelope $\rho(t)$ through time. The deviations from equilibrium amplify to a maximum ρ_{max} at time T_{max} with a rate defined by its reactivity (slope of —) during the short-term transient phase. The disturbance eventually decays at a rate defined by its resilience (slope of - - - -) to the long-term asymptotic phase at a time defined by its return-time. Adapted from Chen XIN and Cohen JE (2001) Global stability, local stability, and permanence in model food webs. *Journal of Theoretical Biology* 212: 223–235.

Modeling Species Interactions

A simple mathematical exercise illustrates how the number of interacting units (species) within a system and the number of interactions between these units (species interactions) are related to the stability of the system. The exercise constructed simple systems of four, seven, and 10 units assumed to be at a steady state. For a system of a given size, connections are made between the units in a way that ensures that all units would be connected to the system, creating several variations of the system in terms of the connections. Next, the strength of the connections is allowed to vary, and the systems are disturbed by moving the units away from their steady state. Two results emerge from this exercise. First, as the number of interacting units increase within a system the likelihood of the system returning to a steady state after a small disturbance (stability) decreases. Second, for a given number of interacting units, as the number of connections between the units increases, the stability of the system declines.

There is a clear connection between the simple mathematical exercise previously presented for estimating diversity

and the diversity and stability of ecological communities. The species is substituted for the generic units modeled previously, and the interactions are defined in terms of predation, competition, or mutualism. Like the generic modeling exercise, high biodiversity does not necessarily lead to stability. Additional study demonstrates that systems composed of species aggregated into blocks of highly interactive species that are connected by weaker interactions are more stable than systems with the same number and types of species and interactions arranged at random. Hence, the way in which the community is constructed affects its diversity and stability.

The types of interactions among species affect the stability of a community as well. Theory and observation agree in general terms on this point, if not on specifics. For example, mutualistic interactions, whereby the actions of each species benefit the growth and dynamics of one another, and omnivory, in which species derive energy from different trophic levels, have been shown to be destabilizing in simple models, but both are prolific in natural communities. All the dominant terrestrial ecosystems rely on a mutualism between the vascular plants and microbes, and nearly every described ecosystem possesses numerous examples of omnivory.

Modeling Species and Processes

Process-oriented models focus on the transformation of elements among various inorganic to organic states through ecosystems. Carbon, nitrogen, and phosphorus are the more common choices of elements to be modeled given their importance to plant and animal growth. Models of species interactions have been connected to models of processes. The general approach to linking these models is to use the structure and formulation adopted when modeling species with the currency used in models of processes. First, species are aggregated into functional groups using one of the schemes described previously, and the interactions among the functional groupings are established. Second, the biomass (expressed in terms of carbon, nitrogen, or phosphorus) of each functional group is estimated. Third, a series of differential equations is developed that establishes which functional groups interact with one another and the nature of the interactions. Parameters within the models include birth and death rates, feeding rates, and the efficiencies by which the functional groups convert consumed matter into new biomass. The models are used to estimate the flow of elements among the functional groups, the flow of elements out of the system, the stability of the ecosystem, and the resilience of the ecosystem.

The following results have emerged from these models: First, the models are sensitive to the changes in the parameters associated with the physiology and life histories of the species within the functional groups. For example, microbes possess high reproductive rates and high rates of energy conversion, and they process matter at a higher rate compared to a functional group dominated by arthropods or nematodes, which possess lower reproductive rates and rates of energy conversion. Second, functional groups are arranged into interactive blocks that form pathways of material flow that originate from a specific source and end with a top predator. The blocked arrangement of functional groups, the distribution of biomass within the ecosystem, and the distribution of the flow of elements through the ecosystem influence the stability and resilience of the system.

Linking Taxonomic and Functional Diversity

The linkage between taxonomic diversity and functional diversity involves many facets. The morphology of an organism determines the type of habitat in which it resides and its ability to colonize new habitats. The physiology of a species influences its adaptiveness to a habitat and the rates and efficiency with which it transforms matter. The life history of an organism influences the rate at which it processes matter (births, deaths, and consumption). Linking taxonomic diversity to functional diversity requires finding measurable attributes that all organisms possess that have meaning to all the ecological subdiscplines, but it also requires flexibility in terms of how the information is used. Three case studies are presented in the following sections. The case studies illustrate how, when studying biodiversity, aspects of both taxonomic and functional diversity are used interchangeably.

Diversity of C3 and C4 Grasses

The first case study involves a comparison of the distribution of grass species that possess either the C3 or C4 photosynthetic pathways in North America. The study illustrates how the information of autecology can be used to find mechanisms to explain patterns in diversity. The study demonstrates how the functional aspects of an organism's physiology influence a species adaptation to a habitat and ultimately the diversity of species with similar physiologies.

Plants that possess C3 and C4 photosynthetic pathways differ in terms of their morphology and biochemistry as they relate to the dark reaction of photosynthesis. For C3 plants, photosynthesis occurs entirely within the outer mesophyll cells, in which CO_2 enters the Calvin–Benson cycle through the plants' stomata. The enzyme (RuBP carboxylase or Rubisco) that incorporates CO_2 into the Calvin–Benson cycle also has an affinity toward O_2. If O_2 were incorporated instead of CO_2, the plant would experience photorespiration – energy is expended, no carbon is fixed, and the plant loses organic carbon. For C4 plants, there are two types of photosynthetic cells: the outer mesophyll cells and the inner bundle sheath cells. In C4 plants, CO_2 enters the mesophyll through the stomata and is incorporated into a four-carbon compound before being exported to the bundle sheath cells in which it enters the Calvin–Benson cycle. The additional steps that occur within the C4 plants prior to the Calvin–Benson cycle cost the plant additional energy and organic carbon. However, the enzyme (PEP carboxylase) that adds the CO_2 to the four-carbon molecule in the mesophyll cells is more efficient than Rubisco and less likely to initiate the loss of organic carbon to photorespiration. This trade-off places plants with the C4 pathway at a disadvantage under cool and moist conditions compared to those with the C3 pathway, but it affords them an advantage in warmer and dry conditions.

Within North America, the northern latitudes are dominated by species of C3 grasses, whereas the southern latitudes are dominated by species of C4 grasses. This pattern of

Table 3 Functional groups of arthropods by number and relative proportion (%) on islands before and 1 year after fumigation

	Number and relative proportion (%)							
	Herbivore	*Scavenger*	*Detritivore*	*Wood borer*	*Ant*	*Predator*	*Parasite*	*Total*
Before	55 (34)	7 (4)	13 (8)	8 (5)	32 (20)	36 (22)	12 (7)	163
After	55 (40)	5 (4)	8 (6)	6 (4)	23 (17)	31 (23)	9 (6)	137

Source: Adapted from Heatwole H and Levins R (1972) Trophic structure, stability and faunal change during colonization. *Ecology* 53: 531–534.

diversity has been attributed to the adaptiveness of each pathway to patterns of temperature and light intensity. The lower latitudes possess higher temperature and higher light intensity than the upper latitudes given the orientation of Earth to the sun. Grasses that possess the C3 pathway have been shown to be more productive than those with the C4 pathway under conditions of cool temperature and low light intensity. At higher temperatures and light intensity, C3 plants suffer greater water loss through their stomata and increased photorespiration than do C4 plants.

Test of the Theory of Island Biogeography

This case study illustrates how information from the community perspective has been used to study biodiversity. The experimental manipulations of several small mangrove islands off the Florida coast by Simberloff and Wilson (1969) and the reanalysis of their work by Heatwole and Levins (1972) offer not only the best validation of the theory of island biogeography but also some insight into how taxonomic diversity and functional diversity are related.

The arthropods on four small mangrove islands were surveyed. The islands were then fumigated with methyl bromide to eliminate the arthropods. For 2 years following the disturbance, the islands were resurveyed for arthropods to capture the recolonization process. The results were consistent with those predicted by the model proposed for the theory of island biogeography (**Figure 3**). The colonization rates were higher and the steady states in species richness were achieved sooner on islands closer to the mainland than on the more distant island. Once a steady state in species richness was achieved, the species composition changed due to the predicted turnover in species.

The study provides an explicit link between the recovery process, species diversity, and functional diversity when the colonization of functional guilds – groups of species that share common food and foraging behaviors – is compared to that of the individual species. One year following the fumigation, the taxonomic diversity of the islands was lower than it was prior to the fumigation (**Table 3**). However, the proportion of species within each of the guilds was the same as it was prior to fumigation. These results suggest that the functional diversity of the islands recovered before the species diversity. In other words, the taxonomic diversity of the communities was in some way dependent on functional characteristics of the community.

Invasion of an Exotic Species

The final case study illustrates how elements of taxonomic and functional diversity and information from autecology, community ecology, and ecosystem ecology have been used to study biodiversity. Vitousek *et al.* (1987) studied the natural history of the invasive plant species *Myrica faya* and the effects that it has had on the nitrogen cycle and plant community development of recent lava flows in Hawaii. Since its introduction to Hawaii from the Canary Islands by Portuguese immigrants at the end of the 19th century, *M. faya* had colonized more than 34,365 ha. *Myrica faya* forms a symbiotic association with the soil nitrogen-fixing actinomycete, *Frankia*. The plant possesses two characteristics that separate it from the native nitrogen-fixing plants on the islands. Unlike the native plant nitrogen fixers (*Acacia koa* and *Sophora chrysophylla*), the seeds of *M. faya* are small, have an edible fleshy fruit, and pass through the digestive systems of birds intact. Moreover, *M. faya* is a more efficient nitrogen fixer than the native species. When combined, these factors make *M. faya* a superior colonizer of recent lava flows, pastures, and forests.

Once established, the *M. faya* and its soil symbiont, *Frankia*, alter the nitrogen status of the soils. For example, soil nitrogen inputs within open canopies of *M. faya* were four times more than those of similar sites without the exotic plant (23.5 and 5.5 kg per ha per year, respectively). The added soil nitrogen was shown to facilitate invasions by other exotic species (e.g., strawberry guava, *Psidium catleianum*) to the detriment of native plant species. The functional attributes of a single species and its symbiont affected the biodiversity of a community by altering the course of community development.

Ecological Coherence of High Taxonomic Ranks

The acquisition and disposition of matter and energy are key processes that all organisms possess that directly link them to ecologically relevant functions. The case studies previously presented share this common feature. The acquisition of matter and energy are characteristics of the interactions in question, and they occur or emerge at taxomomic ranks higher than species. This should not be too surprising given that the modern phylogenies developed from the times of Linnaeus (1735) through Whittaker (1969) relied on energy acquisition and morphology (which is related in many ways to energy acquisition). Recent treatment of prokaryotes that are based on gene sequences have entertained this idea of including environmental information into the calculus of forming taxonomic clades. Assessments of bacterial and archaeal phylogenies when viewed through this lens confirm that there appears to be greater ecological coherence of organisms at high taxonomic ranks. Reducing ecological diversity to functional groups of species that share these attributes may in fact

be the way to organize information captured within species to make the connection between functional and species diversity.

Natural Selection

Species evolve and communities and ecosystems change, but do communities and ecosystems evolve? Does natural selection operate on communities and ecosystems even though they do not possess genes? Natural selection needs to be defined before these questions can be answered. Natural selection is the differential perpetuation of genes in successive generations caused by different degrees of adaptiveness to the environment.

One argument is that ecosystems do evolve because the composition of species within a community changes and natural selection operates on the gene pools of each species. In addition to the changes in the genetic makeup of the species, there are changes in key processes that feed back to reinforce stable configurations or that initiate change. A criticism of this analogy is that natural selection at the ecosystem level for the type of coordination among species that leads to a stable ecosystem is not the same as selection at the organismal level on genes. Selection on genes operates through differential reproductive success of individuals, whereas this may not be true at the ecosystem level since the boundaries of an ecosystem are not so clearly defined.

An alternative view posits that ecosystems do not evolve but that the species composition of the stable ecosystem is a result of the selection operating on individuals and the coevolution of species interacting with one another. The outcome of selection on individuals can influence function within a community, which in turn feeds back as a form of selection. Unique to this argument is that the adaptiveness of a species includes the degree to which its dynamics and function contribute to and are compatible with the stability of the ecosystem as a whole.

Appendix

List of Courses

1. Biogeography
2. Ecosystem Ecology
3. Introduction to Ecology

See also: Guilds. Habitat and Niche, Concept of. Island Biogeography. Measurement and Analysis of Biodiversity. Plant Invasions. Species Diversity, Overview. Stability, Concept of. Taxonomy, Methods of

References

Ehrlich PR and Holm RW (1962) Patterns and populations: Basic problems of population biology transcend artificial disciplinary boundaries. *Science* 137: 652–657.

Elton C (1927) *Animal Ecology*. New York: McMillan.

Gardner MR and Ashby WR (1970) Connectance of large, dynamical (cybernetic) systems: Critical values for stability. *Nature* 228: 784–785.

Grinnell J (1917) The niche-relationships of the California Thrasher. *Auk* 34: 427–433.

Hastings A (2010) Timescales, dynamics, and ecological understanding. *Ecology* 91: 3471–3480.

Heatwole H and Levins R (1972) Trophic structure, stability and faunal change during colonization. *Ecology* 53: 531–534.

Hutchinson GE (1959) Homage to Santa Rosalia, or why are there so many kinds of animals? *The American Naturalist* 93: 145–159.

Linnaeus C (1735) *Systema naturae sive regna tria naturae systematice proposita per Classes, Ordines, Genera, and Species*. Leiden: Haak.

May RM (1973) *Stability and Complexity in Model Ecosystems, Monographs in Population Biology*, vol. 6. Princeton, NJ: Princeton University Press.

Patil GP and Taillie C (1979) An overview of diversity. In: Grassle JF, Patil GP, Smith W, and Taillie C (eds.) *Ecological Diversity in Theory and Practice*, pp. 3–27. Fairfield, MD: International Cooperative Publishing House.

Philippot L, Andersson SGE, Battin TJ, *et al.* (2010) The ecological coherence of high bacterial taxonomic ranks. *Nature Reviews Microbiology* 8: 523–529.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.

Schoener TW (1974) Resource partitioning in ecological communities. *Science* 185: 27–39.

Simberloff DS and Wilson EO (1969) Experimental zoogeography of islands: The colonization of empty islands. *Ecology* 50: 267–278.

Vitousek PM, Walker LR, Whiteaker LR, Mueller-Dombois D, and Matson PA (1987) Biological invasion by *Myrica faya* alters ecosystem development in Hawaii. *Science* 238: 802–804.

Whittaker RH (1969) New concepts of kingdoms or organisms. Evolutionary relations are better represented by new classifications than by the traditional two kingdoms. *Science* 163: 150–160.

Domestication of Crop Plants

Daniel Zohary, The Hebrew University, Jerusalem, Israel

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 217–227, © 2001, Elsevier Inc.

Glossary

Cultivated plants Plants that are grown for their produce.

Self pollination The pollination of a flower by pollen from the same flower or from another flower on the same plant.

Vegetative reproduction (vegetative propagation, vegetative multiplication, vegetative cloning) Is a form of asexual reproduction in plants. It is a process by which new individuals arise without production of seeds or spores.

Introduction

The original focus of farming was on growing food crops; but this was soon extended to include plants essential for other utilities such as fiber sources, stimulants, medicinal aids, dye crops, forage, and timber production. In addition, the past two or three centuries witnessed another major development, namely a massive introduction of ornamental plant species into cultivation. (Today the number of ornamental species exceeds the number of food crops.) Most recently, plants with medicinal properties are getting new attention. The potentials of scores of such plant species (traditionally collected from the wild) have been evaluated. To ensure a steady supply, and for improving quality and yields, many of them are presently introduced into cultivation. A parallel trend is apparent in forestry. Here, too, dozens of timber producing trees are being quickly domesticated. To date, thousands of plant species, native to the various phytogeographic regions of the World, have been introduced into cultivation. In the majority of these cases, the domestic derivatives have already been drastically altered by this move. So much so that the cultivated forms can no more survive in the wild and have become dependent on humans for their existence. Grain crops, for example, lost the wild type adaptation to disperse their seeds and depend on the harvesting, threshing, and sowing by the farmer. Vegetatively propagated tuber plants (under domestication) are commonly sterile. So are many seedless varieties of fruit crops (e.g., bananas). Moreover, as a result of (a) the ecological shift from the original wild habitats into the greatly different and varied anthropogenic environments and (b) repeated cycles of deliberate selection by humans, many domesticates have greatly diverged under domestication. In most major food crops (e.g., bread wheat, rice, maize, apple, or grapevine) or in many popular ornamentals (e.g., roses, rhododendrons, narcissi, tulips, irises), hundreds or even thousands of distinct cultivated varieties ("cultivars") or ornamental forms are now recognized, fitting the wide range of conditions under which these plants are being maintained and the different and contrasting wishes of the growers who constantly try to improve their plants by selection. All in all, domestication provides us with most impressive examples of rapid evolution in plants. Vast arrays of morphological, anatomical, physiological, and chemical changes have evolved in crops and ornamentals, and this in a very short span of evolutionary time. Moreover, selection under domestication molded the gene pools of the crops (the "cultivated gene pools") to become the backbone for the production of food and other essential utilities for humankind. Today, our very existence depends on this genetic diversity. Thus not only do the crops became dependent on humans; also humans became dependent on crops. Obviously, civilizations and domesticated plants have coevolved; and this coevolution continues today.

The Traditional Agricultural Systems of the World

In classical times (about 2500–2000 years ago), and to a large extent until Columbus's time, agriculture was structured quite differently from what we find it today. At present, each country grows a mixture of crops introduced from the various parts of the world. But this fusion is very recent. Traditionally, farming was practiced in largely separated, independent regions, each occupying extensive territories and each possessing a characteristic and largely unique ensemble of native crops. When the Europeans discovered America, they found that the flourishing agricultural civilizations in the New World were based on growing crops that were totally different from what they were familiar with in Eurasia and in Africa. The two hemispheres are indeed very rich in cereals, pulses, tubers, fruits, vegetables, spices, and other types of cultivated plants. Yet the list of domesticated plants in each of them is unique. With only one or two exceptions, these two landmasses originally had no crops in common. This strongly indicates that domestication and agriculture in each of them evolved independently.

Also within each of the two hemispheres, one can detect similar patterns. Each landmass seems to contain several historical agricultural regions, previously separated by difficult-to-cross geographic and cultural barriers. Thus the crops of the traditional agriculture system in Southwest Asia, the Mediterranean basin, and temperate Europe were very different from those grown in East Asia. Moreover, the crop composition in each one of these two agricultural systems is very different from the crop assemblage in Africa south of the Sahara. Similarly in the New World, numerous crops grown in South America were not known in Mexico and middle America, and vice versa.

The main food crops in the traditional agricultural regions of the world are given in Appendix 1. A comparison between the various regional lists of crops reveals that each agricultural system usually contains dominant cereals, companion pulses, oil plants, fruit crops, vegetables, spices, tubers and corms, and stimulants. In other words, the composition of crop *kinds* repeats itself. Only the plant species are different in each traditional region (for example, wheats and barley, lentil, and pea in Southwest Asia, the Mediterranean basin, and Europe; rice and soybean in East Asia; maize and *Phaseolus* beans in America; and sorghum and other millets and cowpea in Africa south of the Sahara).

The Emergence of Agriculture

The shift from hunting and gathering to farming started relatively very late in the history of *Homo sapiens*. The archaeological and botanical evidence already assembled on this development is not even. Some parts of the world (the Mediterranean basin, the Nile Valley, temperate Europe, Southwest Asia, North America) have already been extensively studied. Other parts of the world (East Asia, the Indian subcontinent, South America) are much less explored but at least provide us with some critical evidence on the beginning of agriculture in them. In still others (Africa south of the Sahara, most of the tropical parts of the world) archaeobotanical information is still very poor or almost nonexisting. Even so, the available information clearly shows that farming was independently initiated in several geographically and culturally separated "nuclear areas," both in the Old World and in the New World, between 10,000 and 5,000 years before present (B.P). uncalibrated radiocarbon time.[1] In each such area cultivation seems to have been independently invented, and native wild plant species—largely unique to each area—were introduced and tested in cultivation. Some evolved to be the "founder crops" that initiated farming in these independent places. Once this new way of life had been successfully established, these domesticates also formed the basis for a territorial expansion of the newly formed agricultural populations and for further development in food production.

The First Old World Territory

The oldest and the most extensively studied nuclear area is the "fertile crescent" belt of the Near East. In this rather small geographic territory, a string of early Neolithic farming villages started to appear some 9500 years ago and further established themselves in the subsequent 1500 years. Plant remains retrieved from these early archaeological sites show that eight grain plants growing wild in this Near Eastern "arc" were domesticated at these times. Most common in these preceramic Neolithic Near Eastern sites were remains of three cereals: emmer wheat, einkorn wheat, and barley. They were accompanied by pulses (lentil, pea, chickpea, bitter vetch) and by flax. Soon later, by 8500–8000 B.P., signs of domestication of sheep, goat, cattle, and swine appeared in the sites as well, and the Near Eastern "package" for food production was formed.

Once the package was assembled, and the early Neolithic farming villages were soundly established, farmers started to expand. They did so explosively. By 8000 B.P. this type of agriculture was already introduced into Greece, and by 7000 B.P. it had reached the Danubian basin, the Nile Valley, the Caucasus, and Turkmenistan, and soon afterward it was introduced in the Indus basin. By 6500 to 6000 B.P., grain crop agriculture (and cattle and caprine rearing) was already established all over major parts of temperate Europe—from Ukraine in the east to northern France in the west. More or less at the same time, this new technology also appeared in the middle and western parts of the Mediterranean basin as far as south Spain. All over these vast territories of Southwest Asia, the Nile Valley, the Mediterranean basin, and temperate Europe, agriculture was started by the introduction of the same Near Eastern crops (wheats and barley, and their companion pulses and flax). Only later were additional plants, native to other parts of this huge region, added (e.g., poppy *Papaver sommiferum*) in the west Mediterranean basin.

Some 3000 years after the start of Neolithic grain agriculture, fruit crop horticulture (based on the invention of vegetative propagation) appeared in this region. As with grain crops, the earliest convincing signs of fruit tree cultivation were found in the Near East. Olive, fig, grapevine, and date palm have been cultivated here at least since Chalcolithic times (5500 B.P.). Horticulture, too, spread quickly. By the late Bronze Age, olive, grapevine, and fig (as well as pomegranate and almond) were already used as principal elements of food production in the Aegean belt; and date palm groves flourished in the warm fringes of the Near East and the eastern Mediterranean basin (including Egypt) and extended eastward as far as the Indus Valley.

From the Bronze Age onward there are sound indications of cultivation of vegetables. Melon, watermelon, onion, garlic, leek, and lettuce were the first vegetable crops grown in Mesopotamia or in Egypt. Definite signs of their cultivation appear by 4500 to 3500 B.P. By 2800 to 2000 B.P., the list of Mediterranean and Southwestern Asian vegetable crops had grown considerably; and beet, turnip, cabbage, radish, carrot, parsnip, celery, parsley, and asparagus had also entered cultivation. More or less at the same period (2400–2000 B.P.), a second group of native fruit trees (those in which cultivation depends on grafting) was also added. Most conspicuous among them are apple, pear, plum, and cherries. Contrary to the earlier crops that were almost all taken into cultivation in the Near East "arc," many of the vegetables and the late fruit trees were probably picked up for domestication, not in the Near Eastern nuclear area but in other parts of the already vast Mediterranean and Southwest Asian system of agriculture. Thus, starting in the Neolithic and ending in classical times, an impressive assemblage of native crops were domesticated and diffused all over Southwest Asia, the Mediterranean basin, and temperate Europe. Most of them remained economically important until today.

[1]Uncalibrated radiocarbon ^{14}C years are used throughout this chapter and expressed as years before present (B.P.). Note that dendrological (tree rings) calibration adds to these values—more to the older dates, less to the younger ones. For early ^{14}C dates (9000–5000 B.P.), the correction adds some 700 to 1000 years. By 3000 B.P. radiocarbon time, the differences are negligible.

The Second Old World Territory

A second Asian nuclear area—or more exactly two closely situated, but independent initiations of farming—have been discovered in East Asia. The first was uncovered in the loess soil belt in Honan province and adjacent areas, in the middle part of the Yellow River Valley in North China. The second in the Hupei basin in Central China, along the middle Yangtze River, and also in the wetlands of its delta. The evidence available on the start and on the subsequent development of agriculture in China is still fragmentary compared to that available on the Near East. Yet very ancient farming villages have been discovered both in the northern and in the central parts of this large country. They represent the oldest undisputed signs of plant domestication found yet in East and in South East Asia.

In the Yellow River Valley the oldest farming sites date from 8000 to 7000 B.P. and they contain very large amounts of foxtail millet, indicating that in temperate North China this cereal had been the principal founder crop of the local agriculture. Also in later times foxtail millet continues to be a major food crop in China. More or less in the same time, farming villages seem to have appeared in the middle of the Yangtze River Valley. The earliest sites discovered there date even somewhat older (8500–7000 B.P.) than those in north China. Significantly, in these sites the dominant crop is rice, which later emerges as the most important crop of the vast agricultural economies of East and Southeast Asia. Massive remains of this cereal were unearthed in the Hupei basin, and they continue to appear in the Yangtze River Valley and in later archaeological contexts. These are the oldest remains of rice found yet. But because they occur on the fringe of the wild rice distribution area, some workers suspect that even older rice sites will be discovered in the future further south.

The Third Old World Territory

A third territory in the Old World suspected to have had an independent initiation of agriculture is the Savanna belt and its forest margin in Africa south of the Sahara. The traditional agriculture in this region (which stretches from Senegal and Guinea to Sudan) is based, almost entirely, on a rich assemblage of native crops such as sorghum, pearl millet, African rice, fonio, cow-pea, and bambara groundnut (Appendix 1). The highlands of Ethiopia border the Savanna belt in the east and also harbor an impressive list of local crops, such as teff, finger millet, coffee, and noog. But here the African plants grow side by side with wheat, barley, and many other Near Eastern elements that reached Ethiopia in the past. All in all, African agriculture is diverse and adapted to a wide range of situations. Apart from the Near Eastern elements in Ethiopia, its rich crop assemblage seems almost fully indigenous. This is borne by the fact that the wild relatives of the crops are restricted in their distribution to this part of the World. Most of them could have been introduced into cultivation only somewhere in this geographic belt. All this seems to suggest the independent initiation of cultivation. Whether this is indeed the case is yet hard to say. Archaeologically, Africa south of the Sahara is practically unexplored. The few data available on cultivation of pearl millet or other crops are no more than 3000 years old.

The American Territory

There is little doubt that farming was independently invented in the Americas. As already noted, in pre-Columbus times the Eastern hemisphere and the Western hemisphere shared almost no crop in common. On basis of this evidence it is hard to imagine agricultural contacts between the two landmasses before that time. Also in America itself archaeological research has revealed definite signs of early cultivation, not in a single area but in several distantly located places.

Relatively early rich finds (the age of which has been recently corrected by direct ^{14}C accelerator mass spectrometry tests to 4700 B.P. onward) were discovered in several caves and rock shelters in central and SW Mexico. They contain numerous remains of what are clearly primitive cultivated maize forms, as well as common bean and squash. Because the wild progenitor race of maize—the principal crop in this triad—is restricted only to this general area, it seems reasonable to regard southwestern Mexico as a nuclear area where maize (and its two accompanying crops) were taken into cultivation. This crop combination also remained the main source for food production in the Aztec days.

Indications of early, possibly independent initiation of cultivation come from South America. Chemical comparisons have shown that the common bean and lima bean were independently domesticated both in middle and in South America. Finds in the central and south Andes show that potato and Quinua cultivation in these elevated areas might be 5000 years old, although dating is yet problematic. Finally, sound evidence has been assembled that independent domestication of plants took place in eastern North America. Here goosefoot *Chenopodium berlandieri*, sunflower, and marsh elders were cultivated at least 4000 to 3500 years ago. This is some 2000 years before maize or any other outsider crop reached this area.

Reproductive Systems in Plants Under Domestication

Cultivated plants differ markedly from wild plants in the way by which they are reproduced and maintained. In the various families of the flowering plants, cross-pollination is the principle genetic system. Most species build out-breeding populations. Other reproductive systems such as self-pollination and vegetative propagation (including apomixis) occur as well and are even common in some genera and families. Yet compared to sexual reproduction by cross-pollination, their overall weight is small. In contrast, cross-pollination is relatively rare in plants under domestication. Most crops are maintained by one of the following two systems: (a) self-pollination and (b) vegetative propagation. In contrast to cross-pollination, both these reproductive systems are effective in bringing about immediate "fixation" of desired genotypes.

Self-Pollination

Self-pollination, or more exactly almost full self-pollination, is the principle mating system found in grain crops and in many vegetables. The majority of the 50–60 main grain crops of the world are predominantly self-pollinated. Only a few (such as maize, rye, pearl millet, buckwheat, or scarlet runner bean) are cross-pollinated. Now that the wild progenitors of the majority of the grain crops are already satisfactorily identified, we know that the wild ancestors of the self-pollinated crops are also self-pollinated. In other words, self-pollination in grain crops did not evolve under domestication. It is rather a "pre-adaptation" of the wild ancestor, which considerably enhances its chance to perform successfully in cultivation. One major advantage of self- over cross-pollination is the fact that self-pollination isolates the crop reproductively from its wild progenitor. It enables the farmer to grow a desired genotype in the same area in which the wild relatives abound without endangering the identity of the cultivar by genetic swamping. A second advantage of self-pollination lies in the genetic structure maintained within the crop. Self-pollination results in splitting the crop's gene-pool into independent homozygous lines. Variation is thus structured in the form of numerous true breeding cultivars. Because they are automatically "fixed" by the pollination system, they can be easily maintained by the farmer, even if they are planted together. In contrast, the preservation of varietal identity in cross-pollinated plants is much more problematic. It requires repeated selection towards the desired norms, constant care to avoid the mixing of types and the prevention of contamination from undesirable plants. It is therefore not surprising that early successes in plant domestication frequently involved selfers. In fact, all the eight Neolithic Near Eastern founder crops have this system of reproduction.

Vegetative Propagation

Vegetative propagation is the second widely adapted means to fix and maintain desired genotypes under domestication. This way of handling prevails in the fruit trees, in tuber and corm crops, and in numerous ornamentals. Here, domestication means first of all changing the reproductive biology of the plants involved by shifting from sexual reproduction (in the wild) to vegetative propagation (under cultivation). As a rule, cultivated varieties of these plants have been maintained as clones by cuttings, rooting of twigs, suckers, or by the more sophisticated technique of grafting (and recently also by meristem tissue culture propagation). This is in sharp contrast with their wild progenitors, which reproduce from seeds. Their wild growing populations are usually variable, maintain themselves through sexual reproduction, and are distinctly cross-pollinated. Consequently, seedlings raised from any mother plant segregates widely in numerous traits, including the size, shape, and palatability of the fruits or the tubers.

In the hands of the grower, vegetative propagation has been a powerful device to prevent genetic segregation and to "fix" desired genotypes. By discarding sexual reproduction and inventing clonal propagation, the farmer was able to select, in a single act, exceptional individuals with desirable traits from among a large numbers of variable, inferior plants. Moreover, the farmer could duplicate the chosen types to obtain genetically identical saplings. This is no small achievement. Because the plants are cross-pollinated and widely heterozygous, most progeny obtained from them (even from "superior" plants) are economically worthless. The change from seed planting to vegetative propagation has been a practical solution to assure a dependable supply of desired genotypes. In most fruit trees—and in many tuber crops and ornamentals—it made domestication possible. The history of the Mediterranean fruit trees illustrates this clearly. Olive, grapevine, fig, and date palm that can be relatively simply cloned (by twig rooting or by suckers) were the first to enter cultivation (already by 5500 B.P.). Other native fruit trees that do not lend themselves to easy rooting (e.g., apple, pear, plum, cherry) were domesticated much later (about 2400–2000 B.P.). Their incorporation into horticulture became possible only after the invention of a new cloning tool, namely scion grafting.

Conscious Versus Unconscious Selection

Two main types of selection operate (and complement each other) in plants under domestication:

1. Conscious selection deliberately applied by the growers for traits of interest to them.
2. Unconscious selection brought about by the fact that the plants concerned were picked up and isolated from their original wild environment and placed in a new (and usually very different) human-made environment. This shift in the ecology led automatically to drastic changes in selection pressures. In response to the introduction of the plants into the anthropogenic environment, numerous adaptations vital for survival in the wild lost their fitness. New sets of traits were automatically selected for to fit the new conditions, resulting in the buildup of characteristic domestic adaptations and the appearance of "domestication syndromes"—each fitting the specific agricultural system provided by the domesticators.

Although the pressures and the effects of conscious selection in shaping domesticated plants are familiar to both plant breeders and crop plant evolutionists, unconscious selection has been less frequently evaluated. Because this type of selection shapes so many of the principal traits that characterize crops, several developments caused by the shift in the ecology are discussed in some detail in the following sections.

Methods of Maintenance of Crops and their Impact

As already noted, two main methods of maintenance are employed by the cultivators to maintain their crops: (a) planting of seeds and (b) vegetative propagation. The choice between these two agronomic practices is also the choice between two contrasting patterns of selection and modes of evolution under domestication.

With very few exceptions of apomixis (such as nucellar seed in citrus fruits and mango) planting of seeds can be equated with sexual reproduction. Cultivated plants maintained by seeds (the bulk of the grain crops, numerous vegetables and forest trees, some ornamentals) therefore undergo a recombination-and-selection cycle every planting. In other words, such crops had, under domestication, hundreds or even thousands of generations of selection. They have been continually molded either as (a) clusters of inbred lines (in predominantly self-pollinated crops) or as (b) distinct cultivated races (in cross-pollinated plants). In numerous sexually reproducing crops, the results of such repeated selection are indeed striking. Under domestication these crops have diverged considerably from their wild progenitors. They are now distinguished from them by numerous morphological, developmental, physiological, and chemical traits.

Crops maintained by vegetative propagation (most fruit trees and tuber crops, few vegetables, many ornamentals) have had an entirely different history of selection. Cultivars in these crops are not true races but only clonal replications of exceptional individuals, which are as a rule highly heterozygous. They had been picked up by the cultivator from among numerous sexually produced variable individuals and "fixed" by cloning. In terms of selection, the development of a vegetatively propagated cultivar is largely a single-step operation. With the exception of rare somatic mutations, selection is completed the moment the clone is picked up. In traditional fruit tree horticulture, the turnover of clones has been low. Appreciated genotypes were maintained for long periods of time. Thus, clonal crops underwent in cultivation only a few recombination-and-selection cycles. In sharp contrast to sexually reproducing cultivated plants, their cultivars do not represent true breeding races but only clones, which, as a rule, are highly heterozygous and segregate widely when progeny tested. Significantly, the majority of such segregating progeny are not only economically inferior or even worthless, but they often regress toward the means found in wild populations, showing striking resemblance to the wild forms.

The Purpose for which the Plant is Being Grown

In different crops, different parts of the plant body are being used. Some are raised for their vegetative parts (tubers, leaves, stems, etc.). In others the reproductive parts (inflorescences, flowers, fruits, seeds) constitute the used products. Also the choice of the desired plant parts introduces automatically contrasting selection pressures, particularly in regard to the reproductive system of the crops involved.

When crops are *grown for their seeds* (or at least when they are reproduced by seed), they are kept (like their wild relatives) under constant stabilizing selection that keeps their reproductive system intact. Grain crops provide us with the most rigid cases of such normalizing selection. Yields in these crops depend decisively upon the streamlined development of flowers and fruits, normal chromosome pairing in meiosis, and full fertility. Deviants are weeded out automatically and the reproductive system is kept in balance. It is no wonder that among cultivated plants, grain crops are the most conservative in this regard. They are characterized by strictly balanced chromosome systems and show very little chromosome "problems" under domestication. With very few exceptions (such as the formation of hexaploid bread wheat), the chromosome complements in cultivars of grain crops are identical to those found in their wild progenitors.

Drastic changes in seed fertility (as well as in the chromosome system) can be tolerated when the plants are *grown for their fruits* but are *maintained by vegetative propagation*. Crops in this group (the bulk of the fruit trees) do not depend on seed fertility for their maintenance. Yet they have to keep the basic reproductive growth functions to assure the development of fruits. Moreover, growers usually prefer types in which the size and the number of stones or pips had been reduced, or seedless fleshy fruits. Several solutions for reducing seed fertility without harming fruit set evolve automatically under cultivation. They include the establishment of polyploid clones, some of them meiotically unbalanced (e.g., triploidy in some pears and in bananas), or the incorporation of mutations inducing parthenocarpy—that is, the induction of fruit development without fertilization and without seed set (e.g., bananas, fig, some pears).

Crops *grown for their vegetative parts* and *maintained by vegetative propagation* exhibit the most drastic disruption of their reproductive system and the most variable and bizarre chromosomal situations among cultivated plants. Because such crops are clonally propagated, the conscious selection exerted on them by the growers to increase the output of the desired vegetative part is rarely counterbalanced by normalizing selection to retain sexual reproductive functions. Tropical root and tuber crops provide us with outstanding examples for this development under domestication. Cultivated clones of cassava, yam, sweet potato, or garlic frequently show drastic reduction in flowering. In some, flowering ceases altogether, or almost altogether. When flowers develop they are frequently sterile. Also chromosomally many of these crops are highly polymorphic and frequently contain clones with different levels of polyploidy or variable aneuploid chromosome numbers. Clones may contain 3x, 5x, or even higher meiotically unbalanced chromosome complements. Thus in the yams, *Dioscorea alata* shows all chromosome levels between 3x to 8x; whereas in *D. esculenta*, 4x, 6x, 9x, and 10x forms are known. Sugarcanes confront us with even more complex chromosome pictures. Cultivated clones in this crop are all highly polyploid and frequently aneuploid. Modern cultivars range from $2n = 100$ to $2n = 125$ chromosomes. Older cultivars vary from $2n = 80$ to $2n = 124$.

The Impact of Sowing and Reaping

Traditional grain agriculture is based on *sowing* the seeds of the crop in a tilled field, *reaping* the reproductive parts of the mature plants, and *threshing* out the grains. The introduction of grain plants into such a farming practice initiated, automatically, selection toward the following changes, setting them apart from their wild counterparts.

1. Most conspicuous is the unconscious selection for mutants in which the mature seed is retained on the mother plant

(i.e., for the breakdown of the wild mode of seed dissemination). In cereals this implies a shift from shattering spikes or panicles (in wild forms) to nonshattering ones (in plants under domestication). In pulses it usually means the prevention of the pod from dehiscing and from shedding the seeds. In wild wheat, wild barley, or wild maize, the seed dispersal unit comprises a single internode of the ear. Full disarticulation of the ear, at each segment, is thus an essential element for wild type dispersal, and plants in wild populations are constantly selected for quick shattering of their mature spikes. In contrast, the introduction of planting and harvesting brings about automatic, unconscious selection in exactly the opposite direction. Under the new system a sizable proportion of seed produced by brittle plants will shatter and would not be included in the harvest, whereas all grains produced by nonshattering mutants "wait on the stalks" to be reaped by the grower. Under cultivation, nonshattering individuals have therefore a much better chance to contribute their seed to the subsequent sowing. To summarize, nonbrittle mutants that were totally disadvantaged under wild conditions became highly successful under the new system. Thus, when wild cereals are introduced into the system of sowing and reaping, one should expect selection for nonbrittle forms whether or not the cultivator is aware of this trait. Furthermore, the incorporation of nonbrittle mutants makes the crops fully dependent on humans, as nonshattering plants lose their seed dispersal ability and can no longer survive under wild conditions. For this reason, the presence of nonshattering remains of grain crops in archaeological contexts provide critical indication that these plants were under cultivation. Finally, both theoretical considerations and experimental evidence seem to suggest that at least in wheats and barley, the establishment of nonshattering mutants under the system of sowing and reaping was very likely a fast process. The shift could have been accomplished in the course of several scores of generations.

2. A second major outcome of introducing wild grain plants into this regime of cultivation is the breakdown of the wild mode of their seed germination. Most wild plants, especially annuals, depend for their survival on regulation of their germination in space and time. A common and vital adaptation is a delay of germination and its spread over two or more years. Again, under the system of sowing and reaping such inhibition of germination is automatically selected against. Most grain crops have lost their wild-type germination inhibition patterns. Their seeds germinate fully whenever the farmer plants them.
3. Several other traits characterize cultivated grain crops and are part of the domestication syndrome under sowing and reaping conditions. Because of a rather dense stand in the cultivated field there is a stress to develop forms with erect habit and to reduce the growth of side tillers or branches. Because of the mode of harvest there is a selection for synchronous ripening. In cases of biennial or perennial wild stocks, there is an automatic selection for the shift to annuality. Because the seeds stay protected in the granary, thick shells are selected against. Finally because tilling increases soil fertility there is also a pressure to increase the number of fertile flowers in the reproductive parts of the crops.

Appendix 1: Crops of Various Regions

This appendix lists the native crops of the various traditional agricultural regions of the world (*see* The Traditional Agricultural Systems of the World). For illustrations and short notes on the various food crops, consult Vaughan and Geissler (1997). The important crops of the world and their evolution are surveyed in Smartt and Simmonds (1995).

Southwest Asia, Mediterranean Basin, and Temperate Europe

Cereals
Emmer and durum-type wheats *Triticum turgidum*
Einkorn wheat *Triticum monococcum*
Bread wheat *Triticum aestivum*
Barley *Hordeum vulgare*
Rye *Secale cereale*
Common oat *Avena sativa*

Pulses
Lentil *Lens culinaris*
Pea *Pisum sativum*
Chickpea *Cicer arientinum*
Faba bean (broad bean) *Vicia faba*

Oil or Fiber
Flax *Linum usitatissimum*
Oil seed turnip *Brassica rapa*

Fruits and Nuts
Olive *Olea europaea*
Grapevine *Vitis vinifera*
Fig *Ficus carica*
Date palm *Phoenix dactylifera*
Apple *Malus pumila*
Pear *Pyrus communis*
European plum *Prunus domestica*
Sweet cherry *Prunus avium*
Almond *Amygdalus communis*
Walnut *Juglans regia*

Vegetables and Spices
Melon *Cucumis melo*
Watermelon *Citrullus lanatus*
Onion *Allium cepa*
Garlic *Allium sativum*
Leek *Allium porrum*
Lettuce *Lactuca sativa*
Beet *Beta vulgaris*
Turnip *Brassica rapa*
Cabbage *Brassica oleracea*
Carrot *Daucus carota*
Celery *Apium graveolens*
Parsley *Petroselinum sativum*
Asparagus *Asparagus officinalis*
Cumin *Cuminum cyminum*

Coriander *Coriandrum sativum*
Fenugreek *Trigonella foenum-graecum*

East Asian Agriculture

Cereals and Cereal-Like Grain Crops
Rice *Oryza sativa*
Foxtail millet *Setaria italica*
Broomcorn millet *Panicum miliaceum*
Buckwheat *Fagopyrum esculentum*

Pulses
Soybean *Glycyne max*

Oil or Fiber
Hemp *Cannabis sativa*

Vegetables and Spices
Pe Tsai (Peking cabbage) *Brassica pekinensis*
Pak-choi (Chinese cabbage) *Brassica chinensis*
Rakkuyu *Allium chinense*
Chinese chives *Allium tuberosum*
Ginger *Zingiber officinale*
Chinese pickling melon varieties *Cucumis melo*

Fruits and Nuts
Chinese chestnut *Castanea henryi*
Chinese white pear *Pyrus bretschneideri*
Chinese sand pear *Pyrus pyrifolia*
Loquat *Eriobotrya japonica*
Oriental persimmon *Diospyros Kaki*
Litchi *Litchi chinensis*
Apricot *Prunus armeniaca*
Peach *Prunus persica*

Corms and Tubers
Chinese varieties of turnip *Brassica rapa*
Chinese yam *Dioscorea esculenta*
Lotus *Nelumbium speciosum*

Stimulants
Tea *Camellia sinensis*

Southeast Asian Agriculture (Including the Indian Subcontinent)

Cereals
Rice *Oryza sativa*
Job's tears *Coix lachryma-jobi*

Pulses
Sword bean *Canavalia gladiata*
Mung bean *Vigna radiata*
Black gram *Vigna mungo*
Rice bean *Vigna calcarata*

Oil or Fiber
Sesame *Sesamum indicum*
Tree cotton *Gossypium arboreum*
Coconut *Cocus mucifera*
Jute *Corchorus capsularis*

Fruits and Nuts
Banana: cultivated derivatives of *Musa accuminata* and *M. balbisiana*
Citrus fruit crops (several species)
Carambula *Averrhoa carambula*
Jackfruit *Artocarpus integrifolia*
Durian *Durio zibethinus*
Mangosteen *Garcinia mangostana*
Rambutan *Nephelium lapaceum*

Vegetables and Spices
Cucumber *Cucumis sativus*
Egg plant *Solanum melongena*
Black pepper *Piper nigrum*
Cardamom *Elettaria cardamomum*
Tumeric *Cucurma longa*
Nutmeg *Myristica fragrans*
Cloves *Syzygium aromaticum*
Ginger *Zingiber officinale*

Corms and Tubers
Taro *Colocasia esculenta*
Greater yam *Dioscorea alata*
East Indian arrowroot *Tacca leontopetaloides*

Sugar
Sugar-cane *Saccharum officinarum*

Sub-Saharan African Agriculture

Cereals
Sorghum *Sorghum bicolor*
Pearlmillet *Pennisetum glaucum*
Finger millet *Eleusine coracana*
Fonio *Digitaria exilis*
Teff *Eragrostis tef*
African rice *Oryza glaberrina*

Pulses
Cowpea *Vigna unguiculata*
Bambara groundnut Voandzeia subterranea

Oil or Fiber
Oil palm *Elaeis guineensis*
Noog *Guizotia abyssinica*
Old Word cotton *Gossypium herbaceum*

Vegetables and Spices
Gherkin *Cucumis anguria*

Corms and Tubers
Yam *Dioscorea cayenensis*
Enset *Ensete ventricosa*

Stimulants
Coffee *Coffea arabica*

American Agriculture

Cereals and Pseudo-Cereal Crops
Maize *Zea mays*
Amaranth *Amaranthus cruentus*
Amaranth *Amaranthus caudatus*
Quinua *Chenopodium quinoa*
Goosefoot, *Chenopodium berlandieri*
Marsh elder *Iva annua*

Pulses

Common bean *Phaseolus vulgare*
Lima bean *Phaseolus lunatus*
Scarlet runner bean *Phaseolus coccineus*
Tepary bean *Phaseolus acutifolius*
Jack bean *Canavalia ensiformis*
Peanut (groundnut) *Arachis hypogea*

Oil or Fiber
Sunflower *Helianthus annuus*
Island cotton *Gossypium hirsutum*
Upland cotton *Gossypium barbadense*

Fruits and Nuts
Avocado *Persea americana*
Prickly pears *Opuntia ficus-indica*
Sapodilla *Manilkara zapota*
Guava *Psidium guajava*
Passion fruit *Passiflora edulis*
Papyia Carica papaya
Pineapple *Ananas comosus*
Cherimoya *Anona cherimola*

Vegetables and Spices
Tomato *Lycopersicum esculentum*
Peppers *Capsicum annuum* and *C. frutescens*
Squashes and pumpkins *Cucurbita pepo*, *C. maxima*, *C. moschata*, *C. mixta* and *C. ficifolia*
Vanilla *Vanilla fragrans*

Corms and Tubers
Potato *Solanum tuberosum*
Sweet potato Ipomoea batatas
Oca *Oxalis tuberosa*
Ysanu *Tropaeolum tuberosum*
Ulluco *Ullucus tuberosus*
Cassava (manioc) *Manihot esculenta*
Cush-cush Yam *Dioscorea trifida*

Stimulants
Cacao *Theobrama cacao*
Tobacco *Nicotiana tabacum*
Coca *Erythroxylon coca*

See also: Agriculture, Traditional. Biodiversity in Plant Breeding. Breeding of Animals. Edible Plants. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Plant Sources of Drugs and Chemicals

References

Harlan JR (1992) *Crops and Man*, 2nd ed. Madison, WI: American Society of Agronomy.
Smartt J and Simmonds NW (1995) *Evolution of Crop Plants*, 2nd ed. Harlow, UK: Longman.
Smith BD (1995) *The Emergence of Agriculture*. New York: Scientific American Library.
Smith NJH, Williams JT, Plucknett DL, and Talbot JP (1992) *Tropical Forests and their Crops*. Ithaca, NY: Cornell University Press.
Vaugham JG and Geissler CA (1997) *The New Oxford Book of Food Plants: A guide to the Fruit, Vegetables, Herbs and Spices of the World*. Oxford: Oxford University Press.
Zohary D and Hopf M (2000) *Domestication of Plants in the Old World: The Origin and Spread of Cultivated Plants in West Asia, Europe and the Nile Valley*, 3rd ed. Oxford: Oxford University Press.

Dormancy and Diapause

Nelson G Hairston Jr, Cornell University, Ithaca, NY, USA
Jennifer A Fox, Georgetown University, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Nelson G. Hairston, Jr., volume 2, pp 79–84, © 2001, Elsevier Inc.

Glossary

Bet-hedging strategy A trait of an organism, living in a variable environment, that leads to low variation in fitness. In general, such a trait provides an organism greater net fitness over a range of environmental conditions than would a trait specialized for any single environment. A bet-hedging trait is expected to evolve when the environment in which a species lives fluctuates over a fixed range of conditions that is sufficiently broad for fitness to vary significantly, and when precisely which state the environment will take in the immediate future is unpredictable.

Diapause A state of dormancy in some animals that is induced by a "token" environmental cue, such as day length. The token cue serves as a reliable indicator of a coming onset of harsh environmental conditions, but is not by itself harsh.

Dormancy Any state of reduced metabolic activity of an organism. Typically, dormant organisms have associated characteristics such as cessation of development, absence of reproduction, and enhanced resistance to harsh environmental conditions. Some disciplines have distinct meanings for dormancy that add further constraints to its meaning.

Egg Bank An accumulation of long-lived diapausing eggs (i.e., those eggs that persist in diapause for longer than a single growing season) of aquatic invertebrates in the sediments of marine or freshwater habitats.

Seed Bank An accumulation of dormant plant seeds that persist in dormancy for longer than a single growing season.

Storage Effect A general mechanism for the maintenance of biodiversity within a single habitat, based on differences between competing species in their responses to environmental conditions. A resistant life-history stage can allow coexistence by the storage effect if each species reproduces successfully under the conditions favorable for that species, and can survive through unfavorable periods (e.g., when a competing species dominates) in the resistant stage. Often the resistant stage has prolonged dormancy. This mechanism can also serve to promote the maintenance of genetic diversity within a single population.

Temporal Dispersal The emergence of individuals from dormancy over a range of years (or other time interval), when those individuals entered dormancy in a single year. Often the years (or other time intervals) have different environmental qualities for growth and reproduction.

Temporal Migration The avoidance of harsh environmental conditions in an environment by an individual organism that enters dormancy before conditions become harsh and emerges from dormancy when favorable conditions return.

What are Dormancy and Diapause?

Dormancy is a very general term that encompasses a wide variety of rather different physiological states. The names applied to each of these states depend on the scientific traditions accompanying particular taxon-based disciplines. Only a flavor of this terminological diversity can be covered here, but it is important to recognize that the biological similarities (homologies) among the various physiological conditions called "dormancy" in different kinds of organisms are often questionable. In its most general sense, a dormant organism is simply one with a reduced metabolic rate. For bacteria, where life cycle and cell cycle are synonymous, dormancy is defined as a temporary loss of the ability to reproduce (Henis, 1987), though many single-celled organisms also produce hardened cases at the time of dormancy (Whitton BA, Henis Y *et al.*, and Bradbury PC in Henis, 1987). At the other end of the phylogenetic tree, vertebrate dormancy is most clearly expressed in "adaptive hypothermia" which includes both the daily torpor exhibited by a variety of small mammals and birds, and the longer seasonal hibernation or estivation seen in many taxa (Bartholomew, 1972). These latter states are often induced in part by "token" day-length cues that initiate physiological responses (e.g., fat storage), in advance of the onset of harsh conditions. In higher plants, dormancy is divided into "seasonal" (induced by external day-length cues that foreshadow a change in season) and "opportunistic" (imposed by direct exposure to harsh conditions) categories (Harper, 1977). Similarly, in arthropods, Tauber *et al.* (1986) distinguished between "aseasonal quiescence" ("a reversible state of suppressed metabolism" rather like the "opportunistic dormancy" of plants) and "dormancy" ("a seasonally recurring period – during which growth, development, and reproduction are suppressed"). By their definition, dormancy is then further divided as "diapause-mediated dormancy" which is anticipatory, being induced by token cues (similar to "seasonal dormancy" of plants), and "nondiapause dormancy" which is essentially seasonal quiescence. The common theme here is that organisms from bacteria to vertebrates exhibit dormancy as an adaptation for survival in temporally varying

environments. Over a remarkable range of organismal complexity and phylogenetic origin, species have evolved broadly similar mechanisms for avoiding harsh conditions either by direct physiological responses to the imposition of harsh conditions, or through distinct anticipatory responses to seasonally predictable changes. These mechanisms are generally lumped under the broad umbrella of "dormancy."

Depending on the organism and the environment in which it resides, the duration of dormancy or diapause can be brief or can last for extraordinary periods of time. For microbial cysts, dormant seeds, and diapausing eggs, there are reports of remarkable abilities to survive for centuries in dormancy: There are examples of bacteria becoming active after 200 years, of plant seeds germinating after thousands of years, and of crustacean eggs hatching after 300 years (for a review, see Hairston *et al.*, 1996). The diapausing eggs of *Artemia* (fairy shrimp from salt pans along San Francisco Bay) can survive more than 4 years of complete anoxia with no measurable amounts of stored carbohydrates used during this period, and no detectable metabolic rate, even down to 1/50,000th of the aerobic respiration rate (Clegg and Jackson in Brendonck *et al.*, (1998)). Thus, at least some organisms can apparently survive as "living dead" – that is, in a state without detectable metabolism but still capable of becoming active when exposed to favorable environmental conditions.

Dormancy can occur at many points in an organism's life cycle, from embryo to juvenile and adult stages, depending on the species. Indeed, for some organisms, dormancy may occur at several different life history stages in a single population. Perennial plants may express both bud and seed dormancies (Harper, 1977). Some invertebrates, such as cyclopoid copepods, can enter diapause at several different stages ranging from immatures to adults (Elgmork K in Alekseev and Fryer, 1996) and even as fertilized females (Alekseev and Fryer, 1996). Some fishes can estivate as adults, whereas others, such as the annual killifishes make diapausing eggs capable of surviving several years in desiccation. Some mammals can exhibit both adult hibernation and delayed implantation, a kind of embryonic diapause. In general, short-term dormancy (from a few weeks to a few years) is characteristic of species that are dormant as immature individuals or adults, whereas long-term dormancy (years to centuries) is only found in species that possess embryonic (i.e., seed or egg) dormancy. However, many seeds and eggs actually remain dormant for only brief time periods.

Dormancy and Diapause as Adaptive Traits

It is axiomatic that if individuals that possess dormancy within a population survive a harsh period whereas others that lack dormancy do not survive, then genotypes expressing dormancy will be favored by natural selection. It is less obvious, however, precisely where the tradeoff lies if the harsh environment is not fatal to nondormant individuals. Dormancy represents a reproductive delay (i.e., a lengthening of generation time), and therefore a reduction in potential rate of population increase. Cohen (1970), modeling insect diapause, showed that all individuals in a population should enter dormancy when the arithmetic mean of potential reproduction in the active stage drops below the harmonic mean of survival in the dormant stage. The "means" in Cohen's analysis represent long-term probabilities of successful reproduction or survival. Likewise, Cohen (1966), this time taking plant seed dormancy as his inspiration, showed that the optimal fraction of seeds germinating in any given year depends on the probability that an individual will contribute to long-term fitness through reproduction as a growing plant compared with the contribution it would make by simply surviving another season as a dormant seed. Cohen's important contribution in these two papers was to show that dormancy is not only a response to the certain extermination of active individuals in a highly seasonal environment, but also that it is expected to evolve as a response to environmental uncertainty. It is because not all harsh periods are fatal, and not all growing seasons are bountiful that variation in the expression of dormancy exists both within and among species.

Variation in Dormancy and Diapause

There is often variation within a group of species living in a single environment either in whether or not any given species makes dormant eggs or seeds, or in the time of year that these dormant stages are produced (see Harper, 1977 for examples in plants, and Tauber *et al.*, 1986 for examples in insect diapause). This may be expected if different species react to the environment in different ways. One species may perceive a change as highly unfavorable to which another is relatively immune. There are many examples of populations in which only a fraction of individuals enter dormancy (Venable DL in Leck *et al.*, 1989; Tauber *et al.*, 1986). In these cases, it seems likely that a kind of "bet-hedging" strategy has evolved in response to variation through time, typically among years. In the models reviewed, whether dormancy is favored or not depends on the relative expectations of growth and reproduction in the active stage and survival in the dormant stage. If these expectations vary substantially over time, then the most successful phenotype may be one in which some individuals remain active (in case conditions remain favorable) and others enter dormancy (in case conditions become too harsh). For the same reason, there may be variation in the time of year that individuals within a population enter dormancy or diapause: In some years, the seasonal onset of harsh conditions comes early and in other years, the onset comes later. An example of this variation is a population of freshwater crustaceans (the calanoid copepod, *Diaptomus sanguineus*, living in a small lake in Rhode Island) in which one subpopulation switches from making eggs that hatch immediately to making diapausing eggs significantly earlier in the season than the second subpopulation (Ellner SP *et al.*, in Brendonck *et al.*, (1998).

Time Travelers: Dormancy and Diapause as "Migration from the Past"

Many organisms migrate seasonally away from environments that become uninhabitable at one time of year (usually winter). These include both vertebrates (birds, mammals, and fish)

and invertebrates (e.g., some butterfly and lobster species). There is a large group of organisms, however, that are unable to undertake long-distance movement, and so must survive the seasonally harsh environments where they live. Microbes, fungi, plants, and a great many invertebrates have this life history constraint, as do some vertebrates that hibernate rather than migrating. In a very real sense, the organisms that cannot migrate spatially have evolved dormancy as a kind of temporal migration from one favorable season to the next. Migration is a directed movement between spatially separated habitats. Dormancy is especially analogous to migration when an organism not only begins dormancy before conditions become harsh, but also emerges only when conditions become favorable again. For plant seeds, this is known as ''predictive germination'' (Venable DL in Leck *et al.*, 1989), but many other types of dormancy termination have a predictive character, such as spring arousal from hibernation in mammals or the hatching of crustacean diapausing eggs only when their temporary pond habitat refills with water.

Also analogous to migration are spatial and temporal dispersal in unpredictably varying environments. Spatial dispersal contrasts with migration in that it is a nondirectional movement of individuals away from a source population. It is often interpreted as an adaptation for survival when habitat patches become uninhabitable asynchronously. An individual that produces many offspring, each of which disperses to a new habitat, plays a kind of probabilistic game with many young landing in unfavorable habitats where they die, but with at least some landing in sites where they can successfully grow and reproduce. For organisms lacking effective spatial dispersal in an environment that is unpredictable through time, prolonged dormancy is an effective alternative mechanism for continuing success. An individual that produces many dormant seeds, cysts, or eggs, each of which is capable of surviving long periods of time (i.e., over what would be multiple generations or seasons in the active life history stage) before germinating or hatching, exhibits a kind of temporal dispersal. This is true in particular if the dormant propagules also distribute their emergence from dormancy over multiple seasons. Then, as in spatial dispersal, the dormant offspring awaken in a range of different conditions. Many will emerge at times when the habitat is harsh, but some will emerge and have successful growth and reproduction, producing new long-lived dormant propagules. Seeds with prolonged dormancy are said to belong to a ''seed bank'' (Leck *et al.*, 1989), and eggs in prolonged diapause are, by analogy, said to belong to an ''egg bank'' (Hairston *et al.*, 1996).

Spatial dispersal and prolonged dormancy are predicted in theory to be alternative life-history adaptations (Hairston NG and McPeek M and Kalisz S in Brendonck *et al.*, 1998). Organisms that have excellent spatial dispersal capabilities should not experience strong selection for long-term survival in dormancy, and *vice versa*: Organisms with effective prolonged dormancy should not also evolve mechanisms for spatial dispersal (Cohen D and Levin SA, 1987). Consistent with this theory, there are significant negative relationships between spatial dispersal ability and seed dormancy in plants (Venable DL in Leck *et al.* 1989; Rees, 1993). Similar considerations may explain why winged insects, although often exhibiting single-season diapause (Tauber *et al.*, 1986), very rarely possess diapause that extends over longer periods.

For a variety of species that make dormant propagules (i.e., either dormant seeds or diapausing eggs), there is evidence that these stages can facilitate both spatial and temporal dispersal. The resistant stages that permit dormant organisms to survive stressful periods *in situ* also make it possible for them to withstand unfavorable conditions during transport, whether by physical forces such as wind or water currents, or as hitch-hikers attached to more mobile animals. There is evidence, however, that tradeoffs still exist in which function is most important for a given organism: Plant species with elaborations attached to their seeds that promote wind dispersal have shorter-term dormancy than those that lack such structures (Rees, 1993).

Limitations inherent in spatial dispersal for invertebrates living in lakes and other inland pools very likely explain why diapause and dormancy are so much more prevalent among taxa living in this habitat than in related animals found in marine environments where ocean currents provide for passive dispersal. In a survey of 167 species of crustaceans, Hairston and Cáceres (in Alekseev and Fryer, 1996) found that greater than 55% of those living in inland water bodies possessed a long-lived diapausing stage, whereas less than 10% of marine species possessed this trait. More broadly, prolonged dormancy occurs much more frequently in invertebrate phyla with species that occur in nonmarine habitats than in invertebrate phyla that are exclusively marine. Indeed, no exclusively marine phylum, and only one of 28 exclusively marine classes is known to exhibit prolonged dormancy (Cáceres, 1997a). Several studies have suggested that possession of prolonged dormancy or diapause may facilitate invasion of inland waters by aquatic invertebrates (Hairston NG and Cáceres CE in Alekseev and Fryer, 1996; Cáceres, 1997a; Hairston and Bohonak, 1998). Furthermore, those taxa that have persisted the longest in freshwater habitats, over geological time, are the ones with prolonged diapause (Alekseev and Starobogatov YI in Alekseev and Fryer, 1996).

The Storage Effect, Prolonged Dormancy and Diapause, and the Maintenance of Biodiversity

Long-lived dormant stages that spread their germination or hatching over an extended period of time can play a major role in maintaining both the coexistence of species within a community, and of genotypes within a population. This is especially true when the environment varies through time with some species (or genotypes) favored under some environmental conditions and others favored under other conditions. As an example, consider two competing plant species for which the environment varies through time so that each species has years in which it does well and others in which it does poorly. Also, suppose that in those years when one species does well, the other does poorly. In the absence of seeds with prolonged dormancy, one species would be expected to be competitively superior on average, and thus to eliminate the other over time. However, if both species have long-lived seeds, then the one that is on average the poorer performer can persist by producing seeds when conditions are favorable and

survive in dormancy through years of poor growth and reproduction. This scenario can be extended to communities of many species: If for each species there is an environmental condition in which it does better than its competitors, and if each has some seeds that germinate in each year, then all species can take advantage of years favorable to their own growth and reproduction while not suffering serious fitness loss in years that are unfavorable. Theoretical studies have shown that many species can coexist by this scenario, which is one example of a more general mechanism called the "storage effect" (Chesson, 1994). Its importance in natural communities has been demonstrated for two species of competing freshwater crustaceans in the genus *Daphnia* that make long-lived diapausing eggs (Cáceres, 1997b), and for desert annual plants with long-lived seeds (e.g., Philippi TE in Brendonck *et al.*, 1998).

A similar process can maintain biodiversity of genotypes within a single species if individuals with different genotypes perform best in years of differing environmental quality (Hairston *et al.*, 1996). The major difference between the theory for species competing within a community and that for genotypes competing within a population is evolution. In a community of competing species, one needs only to ask if those species present can coexist. With evolution, however, it is theoretically possible for natural selection to produce a single genotype that has a greater long-term fitness than any other genotype. The question now becomes whether other genotypes with lower long-term fitnesses can coexist in the presence of this most-fit type. The answer is "no" if dormancy is absent or only short term, but "yes" if dormant stages are sufficiently long lived and if natural selection fluctuates sufficiently to favor different genotypes at different times (Hairston *et al.*, 1996). This has been shown to be the case for a population of freshwater crustacean copepods, *D. sanguineus*, with long-lived diapausing eggs, that experience fluctuating natural selection due to year-to-year changes in fish predation pressure (Ellner *et al.*, 1999), and for the freshwater cladoceran *Daphnia magna*, which experiences antagonistic coevolution with the bacterial parasite *Pasteuria ramosa* DeCaestecker *et al.*, (2007).

The characteristics of organisms and their environments that make it likely that dormancy plays an important role in maintaining biodiversity are (1) an environment that fluctuates through conditions that favor a diversity of types of organisms and (2) dormancy or diapause of sufficient duration to span the time it takes for favorable conditions to recur for each organism type. A literature review on these topics by Hairston *et al.* (1996) shows that both of these conditions are quite common in nature for a wide variety of taxa and habitats.

Through its influence on genetic diversity, dormancy or diapause can influence the evolutionary potential of populations. As genotypes from different generations accumulate in the dormant propagule bank, fluctuating or variable selection as described above (in this section) can lead to greater variance of traits compared with the active population, whereas directional selection can cause the mean genotype of the dormant stages to be quite different from that in the active population (Levin, 1990; Hairston *et al.*, 1996). Depending on the dynamics of hatching and viability, the greater genetic variation of this pool can lead to more rapid evolution (Levin, 1990) or to retarded rates as "old" genotypes continue to reappear (Templeton and Levin, 1979). For example, Hairston and Destasio (1988) demonstrated that the dormant egg bank of the copepod *D. sanguineus* slowed down the rate of change in the timing of diapause egg production.

Summary

Dormancy is a term that covers a variety of physiological states in a wide range of kinds of organisms. In each of its forms, dormancy plays a role in the ability of species to live where they do, and is thus important in explaining what organisms are found in which environments. The presence of prolonged dormancy as a character in a taxonomic group of organisms can be critical to the ability of that group to invade and colonize new habitats. This is particularly true for habitats that vary greatly through time and for groups with restricted abilities to disperse spatially. Finally, prolonged dormancy, when combined with a temporally varying environment, can be a significant factor in maintaining both genetic and community biodiversity within a habitat.

See also: Biodiversity Generation, Overview. Desert Ecosystems. Endangered plants. Gene Banks. Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity. Lake and Pond Ecosystems. Life History, Evolution of. Microbial Biodiversity. Migration. Plant Phenology Changes and Climate Change. Population Dynamics. Species Assemblages, Macroecology, and Global Change. Species Coexistence

References

Alekseev VR and Fryer G (eds.) (1996) Diapause in the crustacea. *Developments in Hydrobiology*. Boston: Kluwer. Vol. 114.

Bartholomew GA (1972) Energy metabolism. In: Gordon MS (ed.) *Animal Physiology: Principles and Adaptations*, 2nd edn. New York: Macmillan.

Brendonck L, De Meester L, and Hairston Jr. NG (eds.) (1998) *Evolutionary and ecological aspects of crustacean diapause. Advances in Limnology*, vol. 52. Stuttgart: E. Schweizerbart'sche.

Cáceres CE (1997a) Dormancy in invertebrates. *Invertebrate Biology* 116: 371–383.

Cáceres CE (1997b) Temporal variation, dormancy, and coexistence: A field test of the storage effect. *Proceedings of the National Academy of Sciences of the United States of America* 94: 9171–9175.

Chesson PL (1994) Multispecies competition in variable environments. *Theoretical Population Biology* 45: 227–276.

Cohen D (1966) Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology* 12: 119–129.

Cohen D (1970) A theoretical model for the optimal timing of diapause. *American Naturalist*. 104: 389–400.

Cohen D and Levin SA (1987) The interaction between dispersal and dormancy strategies in varying and heterogeneous environments. In: Teramoto E and Yamaguti M (eds.) *Mathematical Topics in Population Biology, Morphogenesis and Neurosciences. Lecture Notes in Biomathematics 71*, pp. 110–122. Heidelberg: Springer-Verlag.

DeCaestecker E, Gaba S, Raeymaekers JAM, *et al.* (2007) Host-prasite "Red Queen" dynamics archived in pond sediment. *Nature (London)* 450: 870–874.

Ellner SP, Hairston Jr. NG, Kearns CM, and Babaï D (1999) The roles of fluctuating selection and long-term diapause in microevolution of diapause timing in a freshwater copepod. *Evolution* 53: 111–122.

Hairston Jr. NG and Bohonak AJ (1998) Copepod reproductive strategies: Life-history theory, phylogenetic pattern and invasion of inland waters. *Journal of Marine Systems* 15: 23–34.

Hairston Jr. NG and De Stasio Jr. BT (1988) Rate of evolution slowed by a dormant propagule pool. *Nature (London)* 336: 239–242.

Hairston Jr. NG, Ellner S, and Kearns CM (1996) Overlapping generations: The storage effect and the maintenance of biotic diversity. In: Rhodes OE, Chesser RK, and Smith MH (eds.) *Population Dynamics in Ecological Space and Time*, pp. 109–145. Chicago: University of Chicago Press.

Harper JL (1977) *Population Biology of Plants*. New York: Academic.

Henis Y (ed.) (1987) *Survival and Dormancy of Microorganisms*. New York: Wiley Interscience.

Leck MA, Parker VT, and Simpson RL (eds.) (1989) *Ecology of Soil Seed Banks*. New York: Academic.

Levin DA (1990) The seed bank as a source of genetic novelty in plants. *American Naturalist* 135: 563–572.

Rees M (1993) Tradeoffs among dispersal strategies in British plants. *Nature (London)* 366: 150–152.

Tauber MJ, Tauber CA, and Masaki S (1986) *Seasonal Adaptations of Insects*. Oxford, UK: Oxford University Press.

Templeton AR and Levin DA (1979) Evolutionary consequences of seed pools. *American Naturalist* 114: 232–249.

Dynamic Global Vegetation Models

Iain Colin Prentice, Macquarie University, North Ryde, Australia
Sharon A Cowling, University of Toronto, Ontario, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acclimation Processes by which plants adapt to changes in environment over time scales of days to weeks and longer.

Aerosol A suspension of liquid droplets or solid particles in air.

Arrhenius equation The standard relationship between the rate of a chemical reaction and temperature.

Big leaf An approximation to calculate the total rate of photosynthesis by a vegetation canopy, which is formally equivalent to treating the whole canopy as a single leaf.

Biome A broad vegetation type characterized by the dominance of one or more particular PFTs and with a particular structure and appearance, for example, boreal evergreen forest, savanna, desert.

Breathing of the biosphere The land takes up CO_2 during spring and summer, when photosynthesis exceeds respiration, and gives out CO_2 during autumn and winter, when the converse is true. This seasonal signal appears in atmospheric CO_2 measurements. The largest signal globally is from the mid- to high latitudes of the Northern Hemisphere.

CAM Crassulacean acid metabolism, a specialized photosynthetic pathway adopted especially by some desert plants, in which stomata close during the daytime (restricting water loss) and open during the nighttime allowing CO_2 to be fixed by PEP carboxylase (see C_4 plants).

Carbon cycle The circulation of carbon between atmospheric carbon dioxide and other inorganic and organic forms of carbon on land and in the ocean.

CASA The Carnegie-Ames-Stanford Approach model.

CTEM The Canadian Terrestrial Ecosystem Model, which is the land-surface component embedded in the Canadian Climate Model.

C3 plants Plants with the common photosynthetic pathway, including almost all trees and almost all plants of cool climates.

C4 plants Plants with a preliminary CO_2 concentration mechanism involving the enzyme phosphoenolpyruvate carboxylase (PEP carboxylase). Many C_4 plants achieve this with the help of a distinct leaf anatomy, although 'single cell' C_4 photosynthesis is known and may be evolutionarily ancient.

Cenozoic The past 65.5 million years, characterized by cooling and falling CO_2 concentrations.

Chilling A period with temperatures below about 5 °C, required by many temperate plants before they will respond quickly to spring warming.

CLM The Community Land Model of the National Center for Atmospheric Research, USA.

CO_2 fertilization The enhancement of NPP by increased atmospheric concentrations of carbon dioxide.

Colimitation The situation where the rate of supply of two or more essential resources are in balance, so that neither is in excess or deficit. Colimitation is economically optimal in situations where the rates of supply depend on investment.

Complexity (in models) A complex model is one that represents many different processes. Although there has been a tendency for models, including DGVMs, to become more and more complex, there are well-recognized problems with this, especially if the number of parameters increases. A rapidly growing school of thought in biology favors the use of models that are just complex enough for their purpose.

Cretaceous An epoch within the Mesozoic from 145.5 to 65.5 million years ago, characterized by the diversification of angiosperms.

CTEM The Canadian Terrestrial Ecosystem Model, which is coupled to the climate model of the Canadian Centre for Climate Modeling and Analysis.

Decomposition The breakdown of litter and soil organic matter by microbial respiration, leading to the release of carbon dioxide and inorganic forms of reactive nitrogen. In the global carbon cycle, NPP and decomposition are approximately balanced. The small difference is the net biosphere uptake (or release) of carbon.

DGVM Dynamic Global Vegetation Model.

Diffuse radiation Solar radiation from the sky or clouds, which is approximately the same in all directions. Diffuse radiation penetrates dense vegetation canopies more effectively than direct radiation.

Direct radiation Solar radiation from the solar beam, coming from one direction.

Disturbance Removal of biomass, for example, by fire, grazing or extreme winds.

D–O events Dansgaard–Oeschger events, large and rapid climate warming events that took place during the ice ages. The origins of D–O events are not fully established, but their effects are manifest around the world: they are strongest in the North Atlantic and surrounding land regions but also detectable in the tropics and southern hemisphere. The response of vegetation to D–O events is recorded by pollen analysis.

Down-regulation Acclimation that involves a reduction in the maximum rate of a physiological process.

Eccentricity The deviation of the Earth's orbit from circular. Eccentricity varies with periods around 100 thousand years, 400 thousand years and 1 and 2.4 million years.

Eocene A warm epoch within the Cenozoic, from 55.8 to 33.9 million years ago.

ESM Earth System model.

Evapotranspiration The sum of transpiration and evaporation from leaves and bare ground.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00412-3

FACE Free air carbon dioxide enrichment, a computer-controlled system by which carbon dioxide is pumped into an experimental vegetation plot – taking account of varying wind speeds and direction – in such a way as to maintain elevated CO_2 concentration in the air above the plot. FACE gives clearer results than techniques involving chambers, because it eliminates the effect of enclosing the plants in chambers.

fAPAR Fraction of absorbed photosynthetically active radiation: that is, the fraction of solar radiation in the 'photosynthetically active' range of wavelengths that is absorbed by plants. fAPAR is measured from space and provides information about the 'greenness' of vegetation, which is related to total photosynthesis.

GPP Gross primary production, the rate at which CO_2 is incorporated into organic compounds by photosynthesis.

Holocene The past 11,700 years, corresponding to the present interglacial period of the Quaternary.

HYBRID A DGVM that incorporates an individual-based representation of vegetation dynamics, developed by Andrew Friend at various locations.

IBIS The Integrated Biosphere Simulator DGVM, developed at the University of Wisconsin-Madison.

Insolation Incoming solar radiation at the top of the atmosphere.

IPCC The Intergovernmental Panel on Climate Change, a United Nations body which conducts periodic comprehensive assessments of the science of climate change and its implications, with the involvement of a large number of scientists from all over the world.

JULES The Joint UK Land Ecosystem Simulator, a community land-surface model embedded in the Met Office Hadley Centre climate model.

Leaf economics spectrum The tendency of plant leaves to fall along a continuum of variation from thin leaves with fast turnover to thick leaves with slow turnover.

LGM Last glacial maximum, around 21 thousand years ago.

Liebig's law (or the 'law of the minimum') The incorrect hypothesis that growth is always limited by the rate of supply of a single, 'most limiting' resource.

LPJ The Lund–Potsdam Jena DGVM.

LPJmL The 'managed Land' LPJ DGVM, developed at the Potsdam Institute for Climate Impacts Research, which includes process-based representations of crops and grazing lands.

LPJ-GUESS The LPJ-Generic EcoSystem Simulator, a variant of LPJ with more explicit vegetation dynamics, developed at Lund University.

LPX The Land-surface Processes and eXchanges DGVM, a variant of LPJ with a process-based fire model, developed by Colin Prentice's team at various locations.

LSM Land-surface model (a component of a climate model).

MC1 The MAPPS/CENTURY 1 DGVM, developed at the US Forest Service. MAPPS is a previously developed biogeography model and CENTURY is a widely used biogeochemistry model.

Megafauna Large animals.

Mesic Environments with moderate to adequate water supply.

Mesozoic The period from 251 to 65.5 million years ago.

Mid-Holocene Around 6 thousand years ago, characterized by higher than present insolation in northern summer, warmer conditions than present in northern latitudes, and expanded northern hemisphere monsoons.

Miocene An epoch within the Cenozoic, from 23 to 5.3 million years ago.

Nitrogen cycle The circulation of nitrogen between atmospheric nitrogen gas and reactive inorganic and organic forms of carbon on land and in the ocean. The nitrogen cycle and carbon cycle are linked because of stoichiometric requirements, i.e., tissues of organisms have to have at least a certain proportion of nitrogen (the most abundant element after carbon, hydrogen and oxygen) relative to carbon in order to function.

Nitrogen deposition The deposition of reactive nitrogen that has been generated elsewhere, for example, in combustion, after atmospheric transport.

Nitrogen fixation The conversion of atmospheric nitrogen gas into reactive forms that can be used by organisms. The largest natural nitrogen fixation process is biological fixation, carried out by bacteria. Some nitrogen is fixed during lightning. The Haber–Bosch process (used in the industrial production of nitrogen fertilizer) is an additional major source of reactive nitrogen.

NDVI Normalized difference vegetation index, a simple remotely sensed measure related to fAPAR. Remotely sensed NDVI data are available from the early 1980s onward.

NPP Net primary production, the rate of production of new plant material. NPP is equivalent to the rate of GPP minus the rate of plant respiration, which is returned to the atmosphere as carbon dioxide.

Obliquity The tilt of the Earth's axis, which influences total annual insolation at different latitudes. Obliquity varies on an approximately 40 thousand year cycle.

Oligocene An epoch during the Cenozoic, from 33.9 to 23.0 million years ago.

Optimality The idea that natural selection has eliminated all but the most effective modes of behavior by organisms that can be achieved given the biochemical and biophysical constraints. Optimality is a powerful generator of testable hypotheses for modeling biological systems.

ORCHIDEE ORganizing Carbon and Hydrology In Dynamic EcosystEms, the land-surface model embedded in the French (Institut Pierre Simon Laplace) climate and Earth System model.

Parameter A prescribed quantity in a model that does not have a precisely known or measurable value.

Perihelion The time of year when the Earth is closest to the Sun, which is currently in early January.

PFT Plant functional type. A group of plants assumed to have similar properties and responses to environmental changes.

ppm Parts per million, used to express molar mixing ratios of gases in the atmosphere.

Phenology The timing of seasonal events in the life of plants, such as spring budburst in climates with cold winters.
Photoperiod The length of the daylight period. Photoperiod is sensed by green tissues, and is involved in phenological events such as leaf abscision in autumn.
Plant functional traits Directly observable or measurable properties of plants that are related to how they work.
Pliocene An epoch within the Cenozoic, the most recent prior to the Quaternary, from 5.3 to 2.6 million years ago. The Pliocene is characterized by largely modern flora and fauna, and is the most recent epoch to have experienced extended periods globally warmer than present.
Precession The drift of the timing of perihelion through the year, which influences the distribution of insolation with season and latitude. Precession has two major periods of 19 and 23 thousand years.
Quaternary The past 2.6 million years. The Quaternary has been characterized by the alternation of ice ages and interglacials with fluctuating, low CO_2 concentrations that were lowest during peak ice-age conditions and highest during interglacial conditions (warm periods).
Recycling (of water) Transpiration and evaporation from the land surface followed by reprecipitation elsewhere within the continent.
Resprouting Refers to woody plants that are not killed by fires, but sprout new leaves, either from the base or from epicormic shoots (new shoots originating beneath the bark).
Rubisco (or RuBisCO) Ribulose-1,5-bisphosphate carboxylase oxygenase, the central enzyme of photosynthesis. *Rubisco capacity* is the maximum rate of carboxylation (incorporation of carbon dioxide into carbohydrates), determined principally by the amount of Rubisco and other proteins needed for photosynthesis. Because maintenance of these proteins is expensive, the Rubisco capacity is actively regulated by the plant.
Runoff Water that is left over after evapotranspiration. Runoff is the source of freshwater in streams and rivers.
SDGVM The Sheffield DGVM, developed at the University of Sheffield.
Seeding Refers to plants whose seeds germinate abundantly after fires, thus quickly re-establishing a population from seeds.
Soil moisture deficit The difference between the water content of soil and the water content of the same soil at field capacity, i.e., after saturation followed by drainage.
Stoichiometry The ratios of different elements in a material. For example, a plant tissue requires not only carbon, but also nitrogen, phosphorus, and other elements in certain 'stoichiometric' proportions.
Stomatal conductance The conductance for water vapor or carbon dioxide across the leaf surface, which depends on the degree to which the stomates are open or closed. Conductance is the inverse of resistance, so the rate of exchange of carbon dioxide across the leaf surface is equal to the product of stomatal conductance and the difference in the concentration of carbon dioxide outside and inside the leaf.
Stomates (or stomata) Pores in the leaf surface, surrounded by guard cells that regulate the openness of the pores.
Trace gases Gases present in the atmosphere in low mixing ratios (typically less than a few ppm). Some trace gases, including methane, nitrous oxide, and ozone, are greenhouse gases. Others are reactive in the atmosphere and influence the greenhouse effect indirectly. For example, carbon monoxide – produced in fires – competes with methane for the available oxidizing capacity of the atmosphere, and so an increase in carbon monoxide emission leads indirectly to a higher methane concentration.
Transpiration The process by which vascular plants take up water through the roots, transport it through the xylem, and evaporate it from the surface of the leaves.
TRY The global data base of plant functional traits, maintained by the Max Planck Institute for Biogeochemistry in Jena, Germany and governed by an international steering committee.
Vapor pressure deficit The difference between the water vapor content of the air and the saturated water vapor content of the air at its current temperature, the driving force of transpiration.
VOC Volatile organic compound. Many VOCs are emitted by plant leaves, of which the most abundant is the unsaturated hydrocarbon isoprene. VOCs have many functions in plants, including protection against heat stress damage and damage by insect herbivores.
Wetland A terrestrial ecosystem with a water table that is permanently close to the surface.
Xeric Environments with low water supply.
Xylem Water-conducting tissue in the stem or transport roots of vascular plants.

Article Outline

This article is a critical reflection on a class of numerical models that represents large-scale processes in the terrestrial biosphere. DGVMs serve multiple functions. They provide a means of systematizing knowledge about plants and ecosystems that is functional rather than descriptive, and universal rather than particular to a biome or region. DGVM development has stimulated additional research revealing general, empirical relationships among plant functional traits, and between traits and environment. DGVMs also allow predictions of the impacts of climate change on ecosystems, and biogeochemical feedbacks to climate change. But despite impressive achievements in modeling emergent phenomena, such as the 'breathing of the biosphere', predictions of climate change impacts and feedbacks are still highly uncertain, varying greatly among models. The authors argue that the tools and datasets are at hand to reduce such uncertainties. The authors pay particular attention to the role that can be played by modeling the nature and function of past

ecosystems, including those in 'deep time' when the dominant plants were functionally different from today's, as well as those in more recent geological history. The latter provide case studies for the responses of the modern biota to a range of environmental changes, including large and rapid climate changes. The former should provide fundamental insights on Earth system evolution.

Why make Global Models of Ecosystems?

During the summers of 1988 and 1989, Allen M. (Al) Solomon invited a group of scientists to engage in extended brainstorming and collaboration at the International Institute for Applied Systems Analysis in Laxenburg, near Vienna. His agenda was deceptively simple. It was to start development of a new class of multifunctional models, which would seamlessly represent the multiple processes involved in predicting how ecosystem structure and function respond to global environmental change. For many participants, Al's initiative resets their research priorities for the next decade. Dynamic Global Vegetation Models were the outcome.

The idea of DGVMs was taken up by the Global Change and Terrestrial Ecosystems project of the International Geosphere–Biosphere Programme, which announced an intercomparison project for DGVMs before any had actually been constructed. The ensuing scramble to make a working model eventually led to the much-cited article by Cramer *et al.* (2001), who presented historical simulations and future projections with DGVMs that had been developed by six independent groups.

By that time, several climate modeling groups had seen the potential of DGVMs to provide an interactive lower boundary for the atmosphere, and a dynamically changing carbon cycle in which the carbon balance of ecosystems would feed back on the atmospheric concentration of carbon dioxide. After groundbreaking publications on this new field of 'coupled climate–carbon cycle modeling' by Cox *et al.* (2000) and Friedlingstein *et al.* (2001), applications of DGVMs to carbon cycle questions took center stage. Indeed, Cramer *et al.* (2001) already focused on the carbon cycle implications of projected future changes in CO_2 uptake by ecosystems, including both negative and positive feedbacks due to CO_2 fertilization and accelerated decomposition, respectively. The paper's findings on the role of terrestrial ecosystems in the carbon cycle were showcased in the Third Assessment Report of the Intergovernmental Panel on Climate Change (Prentice *et al.*, 2001). Carbon cycle applications have dominated the DGVM literature ever since.

Several other applications of DGVMs were envisaged from the outset and remain important goals. These include predictions of how different aspects of the terrestrial biosphere, such as vegetation structure and habitats, river flows (an indirect consequence of water use by vegetation), and trace gas emissions, respond to environmental change, and how the terrestrial biosphere has influenced and been influenced by climatic and environmental changes in the geological past. These fields of application are summarized later in this article.

More fundamentally, behind the development of DGVMs lay a more radical, unwritten agenda: to realign ecology's research priorities with the requirements of global change research. In the late 1980s, ecology almost completely lacked a focus on (or funding for) either the study of global-scale phenomena, such as the relation between vegetation structure and macroclimate (then often derided for being nonexperimental science), or even more crucially, a search for empirical generalizations that would apply globally, across different biomes and regions. The prevailing wisdom at the time held that models should be developed (if at all!) for one biome at a time, emphasizing processes important in a particular biome – neglecting the difficulty that climate change was already known (from paleoecology) to be capable of converting one biome into another, so without a generic model, there would be no way to predict the dynamics of such transitions. The development of DGVMs has made this position untenable, and has helped to establish 'global ecology' or 'macroecology' as a high-profile area of science.

Comparative information on plant traits across species and biomes was scarce at that time, and modelers were obliged to rely on a very limited information base dominated by a few species of commercial importance. This situation too has changed drastically. The past decade has seen the emergence of important generalizations based on quantitative trait data from thousands of species, such as the leaf economics spectrum (Wright *et al.*, 2004), while TRY data base has brought together more than 2 million pieces of quantitative and qualitative information on over 70,000 species of vascular plants (Kattge *et al.*, 2011). This effort has been stimulated, and frequently justified to funding agencies, precisely by the need to provide better information for DGVMs.

This article is a critical reflection on the state of DGVMs. The authors suggest that they have only partially fulfilled their original promise, while providing a number of pointers to how their usefulness could be enhanced; not least, by taking on board the implications of the 'data revolution' alluded to above. The reader is also referred to Arora (2002), Woodward and Lomas (2004), Prentice *et al.* (2007), and Tang and Bartlein (2008). The authors note that applications to questions of biophysics (land–atmosphere interactions) and biogeochemistry (carbon and to a lesser extent nitrogen cycling) have tended to dominate over applications to questions of community composition and biodiversity. Yet the framework of DGVMs would allow far more to be done in terms of predicting climate change impacts on habitats and species, while the newly data-rich world of plant functional ecology should allow greater realism and reliability to be achieved in models.

Types of Ecosystem Model and their Relation to DGVMs

The building blocks of DGVMs and how they were fitted together have been reviewed by Prentice *et al.* (2007). Briefly, DGVMs represent a merger of four previously distinct modeling strands, developed by separate research communities, *biogeography* models originated from vegetation-oriented climate classifications such as that of Köppen (1918). Biogeography models predict the broad-scale distribution of biomes, assumed to be static and in equilibrium with climate. Woodward (1987) and Prentice *et al.* (1992) showed how a combination of independent limiting

values of bioclimatic variables (indices of winter cold, summer warmth, and moisture availability) could predict biome distributions in the different continents with surprising realism, thus injecting a more explicit process basis into biogeography modeling. *Vegetation dynamics* models emerged from regionally specific, individual-based models that predicted forest succession patterns from basic allometric and growth traits of trees (Botkin *et al.*, 1972; Shugart, 1984). Although these models in general did not preserve mass or energy balance, they were a step forward from earlier attempts to model vegetation dynamics purely from observed transitions between discrete vegetation states. *Biogeochemistry models* (e.g., Running and Coughlan, 1988; Running and Hunt, 1993; Raich *et al.*, 1991; Melillo *et al.*, 1993; Parton *et al.*, 1993; Del Grosso *et al.*, 2001) by contrast focused on mass balance as opposed to vegetation structure and composition, and described the coupled cycling of carbon, water, and (usually) nitrogen through ecosystems of a given type on multiannual time scales. *Soil–vegetation–atmosphere models* focused on fast cycling processes (carbon uptake and release, evapotranspiration, and water balance, usually with subdaily time steps), assuming the structure of the ecosystem and the masses of different plant compartments to be fixed. This last category of models also provided the foundations for *land-surface models*, built into atmospheric general circulation models (climate models) (Seller *et al.*, 1997). In DGVMs, the different levels of process represented by the four model categories are linked through partially synchronous coupling so that (for example) plant water and carbon balances are updated at daily or subdaily time steps whereas vegetation dynamics is modeled on annual time steps (Prentice *et al.*, 2007).

These modeling strands have continued to develop independently or in various combinations. For example, some contemporary models for vegetation dynamics use explicit approximations to allow the modeling of individual-based competition without actually simulating individuals (e.g., Medvigy *et al.*, 2009). Models for biogeography, appropriate to the representation of 'snapshots' of paleoecological time, now incorporate elements of biogeochemistry modeling so as to allow the representation of competitive interactions between trees and grasses, evergreen versus deciduous trees, C_4 versus C_3 herbaceous plants, and effects of changes in atmospheric CO_2 concentration on these interactions (Haxeltine and Prentice, 1996a). But DGVMs are a well-established line of development in their own right, and in principle they can be used for all the purposes suitable to their constituent submodels.

DGVMs also form a component of *coupled climate–carbon cycle models* (Cox *et al.*, 2000; Friedlingstein *et al.*, 2001, 2006), and of comprehensive *Earth system models*. Comprehensive ESMs aim to include a range of land–atmosphere and ocean–atmosphere interactions in a fully coupled manner, for example, trace gas emissions from ecosystems, which affect atmospheric composition, which in turn affects climate and ecosystems (Scholze *et al.*, 2012). The first generation of ESMs has been developed and run by climate modeling groups for the Fifth Assessment Report of the IPCC, for publication in 2013–2014. At the time of writing, these ESMs are in an early stage of development. Nevertheless, there are good reasons to aspire to coupled modeling of the biological, physical, and chemical aspects of the global environment. DGVMs now form an indispensable part of the Earth system modeling toolkit.

There are now many more DGVMs than the original six compared in Cramer *et al.* (2001). Some of the more extensively used and developed DGVMs include CASA (Potter *et al.*, 1993; deFries *et al.*, 1999), CLM (Levis *et al.*, 2004), CTEM (Arora and Boer, 2006; Christian *et al.*, 2010), HYBRID (Friend and White, 2000; Friend and Kiang, 2005; Friend, 2010), IBIS (Foley *et al.*, 1996; Kucharik *et al.*, 2000), JULES (Alton *et al.*, 2007, Clark *et al.*, 2011), LPJ and LPJmL (Sitch *et al.*, 2003; Gerten *et al.*, 2004; Bondeau *et al.*, 2007), LPJ-GUESS (Smith *et al.*, 2001), LPX (Prentice *et al.*, 2011a), MC1 (Lenihan *et al.*, 1998), ORCHIDEE (Krinner *et al.*, 2005), and SDGVM (Woodward *et al.*, 1998; Woodward and Lomas, 2004). This is not a comprehensive list. New models are being developed all the time for either coupled or stand-alone applications or both. Sitch *et al.* (2008) presented an intercomparison of five DGVMs. Murray *et al.* (2012) provide a longer list of DGVMs that have been used in hydrological applications.

Principles and Processes in DGVMs

From Empirical to Process-Based Modeling: A Well-Traveled (But Bumpy) Road

The representation of fundamental leaf-level processes in DGVMs and their precursor models evolved from early attempts based on empirical dependencies of plant growth on climate variables (solar radiation, temperature, and precipitation) toward more explicitly process-based representations, which generally have superior power to predict the behavior of plants in changing biotic and abiotic environments. Plant growth, for example, is now usually quantified in DGVMs as a balance of photosynthesis, predicted either using the standard model of Farquhar *et al.* (1980) or a common simplification thereof, and respiration associated with the growth and maintenance of different tissues. The Farquhar model includes a number of coefficients that are almost constant among different plants at any given temperature, and with a known temperature dependence (e.g., the affinities of Rubisco for CO_2 and oxygen). The model describes photosynthesis by C_3 plants but it has been modified to apply to C_4 plants as well. The rationale of applying such a general model is that a few basic characteristics (e.g., photosynthetic pathway) and measurements (e.g., Rubisco capacity) on plants then provide enough information to predict photosynthesis under a vast range of environmental conditions.

But there are still difficulties, because Farquhar's model does not tell us, for example, how one might expect Rubisco capacity to vary seasonally or between different climates. Many DGVMs simply apply a constant value for each type of plant, but this approach ignores acclimation processes. Similarly, plant respiration in most DGVMs is assumed to vary with temperature following the Arrhenius equation, which is well-established for short-term measurements, or a simple exponential relationship, which is an adequate approximation to the Arrhenius equation over a moderate range of temperatures. But recent research has shown that plant respiration acclimates to temperature so as to almost entirely negate this temperature dependence on time scales of weeks or longer, i.e., the main time scales at which we would like to apply

DGVMs (Atkin and Tjoelker, 2003; Atkin *et al.*, 2008; Zaragoza-Castells *et al.*, 2008). Most current DGVMs do not allow for this acclimation and therefore must produce incorrect longer term responses of plant respiration to temperature. Some models however have assumed that plant respiration constitutes a fixed fraction of photosynthesis; this may be oversimplified, but at least it is consistent with the evidence for respiratory acclimation (Gifford, 2003).

Adoption of process-based models is not a panacea. The way in which models represent the longer term acclimation of basic physiological processes as discussed in the previous paragraph is just one area where even DGVMs that appear to use the same, well-tested process model can get very different (and potentially wrong) results. The following subsections provide a (nonexhaustive) list of key processes for which many different approximations are in use, some better justified than others, and where more critical evaluation and comparison with observations would be desirable.

Carbon–Water Coupling

A basic tenet of ecophysiology, made explicit in DGVMs, is that plants lose water during transpiration when stomata open to take up CO_2 from the atmosphere during photosynthesis. Thus, there is a close coupling between carbon gain and water loss, both being regulated by stomatal 'openness.' There is no established mechanistic model for stomatal behavior. But there is a powerful and well-supported theory (Cowan and Farquhar, 1977), based on the idea that stomata behave in such a way as to optimize the marginal carbon gain per unit water loss (see also Schymanski *et al.*, 2009). Some DGVMs explicitly simulate stomatal conductance (the measure of stomatal openness) based on some variant of this principle (e.g., Friend, 1991). Others use long-standing empirical formulations, such as the Ball-Berry model (Ball *et al.*, 1987) which relates stomatal conductance to photosynthetic rate, ambient CO_2 concentration, and relative humidity.

The analysis by Medlyn *et al.* (2011) has shown that the difference between the Ball-Berry formulation and the optimization approach, mathematically speaking, is slight. Both lead to a strong inverse dependence of stomatal conductance on atmospheric vapor pressure deficit (dryness). But important open questions remain to afflict the modeling of plant responses to drought. For example, the Cowan–Farquhar theory relies on an 'exchange rate' between water and carbon, which is expected to vary among plant types, and with soil moisture. Much less is known about these variations than about the rapid response of stomatal conductance to atmospheric dryness. There is also evidence that photosynthetic rates decline with soil moisture more steeply than would be expected from the behavior of stomatal conductance alone. Some models incorporate an inhibition of photosynthesis by soil moisture deficit, others do not.

Responses to Atmospheric CO_2 Concentration

The short-term response of photosynthesis to increase the ambient concentration of CO_2 is well established. At present-day ambient CO_2 or (especially) lower ambient CO_2, there is a positive response of photosynthesis in all C_3 plants. This response is predicted accurately by the Farquhar model, and is included in most DGVMs. In addition, stomata tend to close somewhat in response to increasing ambient CO_2. This behavior is consistent with the maintenance of a constant ratio of leaf-internal to ambient CO_2. Models that treat this ratio as conservative will predict it, others may not do so.

There is evidence from many plant species for an acclimation response of photosynthesis, such that Rubisco capacity declines with increasing CO_2 (Medlyn *et al.*, 1999). But the decline in Rubisco capacity is generally accompanied by an increase in photosynthesis. This behavior can be predicted from an optimality hypothesis formulated by Dewar (1996) and Haxeltine and Prentice (1996b), namely that Rubisco capacity optimizes the balance of photosynthesis and respiration at the leaf level, over a period of days to weeks. To the authors' knowledge, only the Lund–Potsdam–Jena (LPJ) model (Sitch *et al.*, 2003) and its descendants include this criterion. Furthermore, the experimental literature commonly attributes this acclimation (or 'down regulation') to 'dilution,' nitrogen deficiency, or other limitations on growth. The jury is out.

In any case, most models assume Rubisco capacity to be fixed for a given type of plant. The particular consequences of allowing photosynthetic acclimation in a model have rarely been investigated. Arora *et al.* (2009) used the Canadian coupled climate–carbon cycle model to test the effect of photosynthetic acclimation over the period 1850–2000 AD. In simulations without photosynthetic acclimation, atmospheric CO_2 at 2000 was 15–20 ppm lower than observations. Inclusion of a realistic level of acclimation corrected this discrepancy. But it is not automatic that reducing carboxylation capacity would reduce CO_2 uptake. Too high a value would incur respiratory costs in excess of achievable gains.

Photosynthetic responses do not necessarily predict responses of plant growth. Growth may be constrained by low temperatures, and/or insufficient availability of essential nutrients to convert fixed carbon into plant tissues, whose stoichiometry can only vary within certain limits. Knowledge of the growth responses of field-grown plants and *in situ* ecosystems has increased greatly through the implementation over the past decade or so of free air carbon dioxide enrichment (FACE) experiments (Hendry *et al.*, 1993; Hendry and Kimball, 1994; Kimball *et al.*, 1995; Nowak *et al.*, 2004; Ainsworth and Long, 2005; Albert *et al.*, 2011; Norby and Zak, 2011). Norby *et al.* (2005) showed that a 200 ppm increase in ambient CO_2 concentration in forest FACE experiments induces an increase in NPP consistently close to 23% across a wide range of absolute NPP values. Hickler *et al.* (2008) showed that the LPJ-GUESS model produces approximately this level of NPP enhancement in temperate forests, but noted that the model also predicts much smaller enhancements in boreal forests and much larger enhancements in tropical forests, where as yet there are no FACE experiments to test the prediction. There is also considerable uncertainty about the sustainability of the growth response to CO_2, and the sources of the additional nitrogen that would be required to sustain such an increase (e.g., Finzi *et al.*, 2007; Norby *et al.*, 2010; Hofmockel *et al.*, 2011).

Carbon–nitrogen Coupling

The topic of CO_2 'fertilization' of plant growth leads naturally to the consideration of nutrient limitations on growth. Hungate *et al.* (2003) criticized the predictions of future global land CO_2 uptake made by the DGVMs in Cramer *et al.* (2001), on the grounds that the models did not include limitations on growth due to limited nitrogen availability. In fact, two of the models (SDGVM and Hybrid) did include nitrogen limitations, and these two showed the smallest growth responses. But they still exceeded Hungate *et al.*'s (conservative) range of estimates for the maximum possible response.

The interactions of the terrestrial carbon and nitrogen cycles are not well understood. There is evidence for a reduction in the CO_2 fertilization effect on forest growth under more nitrogen-deficient conditions (e.g., Oren *et al.*, 2001; Dieleman *et al.*, 2010; Norby *et al.*, 2010). It is often asserted that nitrogen supply limits the response of plant growth CO_2. However, plenty of counter-examples exist. Nowak *et al.* (2004) analyzed results from FACE experiments in a variety of ecosystems, including several where nitrogen status had been manipulated as well as CO_2. They found that there was often a very large response of growth to the combination of nitrogen addition and elevated CO_2. However, they were not able to show that the relative enhancement of growth by elevated CO_2 was different between plots with normal and increased nitrogen nutrition. Lee *et al.* (2011) showed a mean enhancement of NPP in grasslands by about 10% over a decade – less than some models would predict, but the enhancement was independent of the nitrogen status of the ecosystems.

The inclusion of interactive nitrogen cycling has recently become a major focus of DGVM development (Ostle *et al.*, 2009). Motivations include attempting to more realistically simulate the effects of increasing CO_2 on growth, the feedbacks from soil organic matter decomposition to climate, and the effects of anthropogenic nitrogen deposition on growth (e.g., Gu *et al.*, 2010). New models or versions of models with interactive carbon and nitrogen cycling have been presented by Thornton *et al.* (2007), Sokolov *et al.* (2008), Xu-Ri and Prentice (2008), Jain *et al.* (2009), Gerber *et al.* (2010), Zaehle and Friend (2010), Zaehle *et al.* (2010), and Esser *et al.* (2011). These models have yielded a startling diversity of results. There is unfortunately no consensus on how (and what) observations could be used to discriminate between different models. Results of field experiments may provide the most valuable benchmarks. Zaehle and Friend (2010), for example, showed how modeled responses to soil warming could be improved by the inclusion of an interactive nitrogen cycle.

Outstanding problems for modeling include the poorly constrained nature of estimates of total global nitrogen fixation, the ultimate source of most reactive nitrogen to ecosystems, and denitrification, the major pathway of gaseous nitrogen loss; failure to consider the recycling of nitrogen from sedimentary rocks, a previously overlooked natural nitrogen source (Morford *et al.*, 2011); and failure to consider the ways in which plants can allocate carbon in order to acquire nitrogen. This last consideration forms the basis for a promising new approach that seeks to optimize the marginal gain in growth (due to increased nitrogen acquisition) per unit of carbon invested, and allows plants to switch accordingly between alternative pathways of nitrogen uptake (Fisher *et al.*, 2010).

Carbon Allocation and Multiple Resource Limitations

One important ecophysiological process resulting in major differences in DGVM simulations of carbon gain in plants is how fixed carbon is allocated between plant organs (leaves, woody tissues, and fine roots) (Malhi *et al.*, 2011). This too is relevant to the question of carbon–nitrogen cycle coupling, because all mechanisms that would increase the uptake of nitrogen from soils involve increased allocation below ground and therefore a trade-off between nutrient and light capture. Below-ground carbon allocation is important not only for its effects on nutrient acquisition and productivity, but also as a major carbon source for soil organic matter (Johnson *et al.*, 2011) and as a regulator of soil decomposition rates (Metcalfe *et al.*, 2011). There are abundant (albeit unsystematically collected) experimental data on plant allocation responses to a variety of environmental manipulations, forming the basis for the analysis by Poorter *et al.* (2012). This analysis provides a wealth of quantitative information, which should now be used to improve the representation of carbon allocation in DGVMs.

DGVMs have adopted various *ad hoc* representations of how different environmental constraints in combination control plant growth. Some seem to have been influenced by 'Liebig's law,' the hypothesis that growth is controlled by a single, 'most limiting' factor. This hypothesis is generally false for complex organisms such as higher plants – as shown, for example, by the fact that plants can show positive growth responses to increased supplies of water, nitrogen, and CO_2 in a single experiment. A general theoretical approach is much needed to replace this misleading hypothesis. The authors suggest that the key is recognizing that the acquisition of any resource depends on the allocation of carbon. So when a particular resource is in short supply, the optimal response may involve allocating more carbon to acquisition of that resource (e.g., Rastetter *et al.*, 2001; Rastetter, 2011). This principle is in direct contradiction to 'Liebig's law,' but it corresponds to a well-established tenet in microeconomics and has considerable predictive power for plants and ecosystems. For example, it predicts colimitation of photosynthesis by water and nitrogen (Wright *et al.*, 2003) and colimitation of plant growth by nitrogen and phosphorus, now recognized to be widespread (e.g., Domingues *et al.*, 2010; Harpole *et al.*, 2011).

Interactions specific to particular resources need to be considered as well, and can sometimes be counter intuitive. For example, it has been observed that nitrogen fertilization can lead to a large enough decrease in below-ground carbon allocation that trees' drought resistance is impaired (e.g., Graciano *et al.*, 2005). The prevalence of symbiotic nitrogen fixation in tropical ecosystems, where phosphorus is often scarce, has been explained by the high nitrogen content of extracellular phosphatases that can help plants to acquire phosphorus from otherwise inaccessible chemical forms. This last suggestion forms part of the basis for the biogeochemical model of Wang *et al.* (2010b).

Scaling from Leaf to Canopy

DGVMS face the challenge of representing, in a simple but sufficient way, the properties of canopies that are almost intractably heterogeneous – both horizontally and vertically. The first and simplest approach involves a 'big leaf' (sometimes derogatorily called 'green slime') approximation. The 'big leaf' approximation has been justified mathematically on the grounds that it holds, in theory, when the vertical profiles of light availability and nitrogen content of foliage are similar (Sellers *et al.*, 1992; Friend, 2001). Both decline with depth in the canopy, but in reality the gradient of leaf nitrogen per unit area is less steep than the gradient of light attenuation (e.g., Lloyd *et al.*, 1995, 2010). It has also been shown that a 'two big-leaf' approximation, which divides the canopy into a sunlit and shaded fraction, is more exact (e.g., de Pury and Farquhar,1999; Zeng *et al.*, 2003). The scaling of leaf-level processes to the canopy remains an active issue (Ren and Henderson-Sellers, 2006; Popp *et al.*, 2009; Chen *et al.*, 2011; Zhu *et al.*, 2011). Some DGVMs now include an explicit vertical profile of canopy properties (Mercado *et al.*, 2007), which facilitates modeling of some processes such as the different canopy penetration by diffuse and direct radiation, with implications for total canopy photosynthesis (Mercado *et al.*, 2009).

Plant Functional Types

The characterization of vegetation cover as being composed of a quantitative mixture of plant functional types, as opposed to classifying vegetation into qualitatively distinct biomes, was a major step forward for process-based biogeography modeling (Prentice *et al.*, 1992) and a prerequisite for dynamic modeling. The ecophysiological foundations for bioclimatic distinctions among PFTs have been reviewed by Harrison *et al.* (2010a). The PFT concept also applies to distinctions, for example, between early and late successional plants, or plants with different responses to fire or grazing pressure. There is an extensive literature on PFTs (e.g., Woodward and Cramer, 1996; Díaz and Cabido, 1997; Wright and Westoby, 2002; Wright *et al.*, 2005; Lavorel *et al.*, 2011; Fyllas *et al.*, 2012). The fundamental idea behind PFTs is that there are inevitable trade-offs between different plant traits, such that not all combinations of traits are possible; therefore, plants can be classified into a manageable number of PFTs. This idea also leads to the notion of a low-dimensional continuum of plant trait combinations (Wright *et al.*, 2004; Prentice *et al.*, 2011b). The evidence supports continuous variation for many plant traits, and also shows the coexistence of many different trait values within any given community.

Current DGVMs, which classify all plants into a small number of PFTs (typically between five and 15), therefore probably underestimate the ability of communities to adapt to environmental change through selection for species with different trait values within a model-defined PFT. They certainly underestimate the functional diversity of communities. It is likely that future DGVMs will include methods to better represent continuous trait variations within a limited number of discrete (or almost-discrete) classes, such as C_3 versus C_4 and CAM photosynthetic pathways; the tree habit versus the shrub, forb, and graminoid habits; deciduous versus evergreen phenological types; and gymnosperm versus angiosperm wood anatomies.

Phenology

Phenology describes the seasonal timing of events in the life of an organism. In the DGVM context, the key phenological events that are modeled are the timing of budburst, the length of time required to the full 'leaf-on' state, and the timing of leaf senescence and/or abscision. Although these events take place in all green plants, most DGVMs consider them only for deciduous woody plants in seasonal climates, and perennial herbaceous plants with seasonal display of above-ground biomass. The key importance of modeling these particular cases lies in the control they exert on the seasonal pattern of 'greenness,' i.e., the fraction of absorbed photosynthetically active radiation, which is a principal control on the rate of photosynthesis at the ecosystem level.

fAPAR is routinely measured by satellite instruments that record the relative reflectance of the ground surface in different spectral bands. The prototype of this measurement is the normalized difference vegetation index, which is based on a comparison of spectral reflectances in just two frequency bands: the red (where chlorophyll absorbs most strongly) and the near infrared, where chlorophyll does not absorb. The global NDVI record goes back to the early 1980s. Owing to the availability of satellite data, a subclass of DGVMs has emerged in which phenology and leaf area are not explicitly modeled, but instead prescribed from data (e.g., the CASA model, DeFries *et al.*, 1999). However, as phenology is highly responsive to climate, it is important for many purposes also to be able to predict phenology and leaf area (Arora and Boer, 2005a). Most DGVMs have this capability. It is best developed for high latitude climates with a short growing season, where 'leaf-on' can be predicted reliably with a simple growing degree day index (Harrison *et al.*, 2010a) and 'leaf-off' predicted, albeit less reliably, with a threshold temperature. Lucht *et al.* (2002) showed that the high latitude 'greening' (increasing leaf area and duration) trend observed in NDVI data by Myneni *et al.* (1997) could be reproduced using the LPJ model, as could the associated advancement of leaf-on in spring, and the temporary setback to both trends that was observed later after the eruption of Mount Pinatubo in 1991. However, phenology in other regions is more complex than in high latitudes. Woody plants have evolved mechanisms to prevent a too-rapid response to warm intervals in winter so that the accumulated warmth required for budburst depends inversely on the length of a preceding period of chilling (Murray *et al.*, 1989). Botta *et al.* (2000) combined NDVI and climate observations to fit different models for the phenology of different PFTs, including a chilling-dependent model for temperate trees. The resulting scheme has been included in the ORCHIDEE DGVM. But this is certainly not the last word. Experimental studies have shown a subtle interplay of photoperiodic and climatic responses (Chuine, 2000), which has yet to be incorporated in a DGVM.

The simulation of phenology in tropical biomes has caught researchers' attention in part because simulations indicate that

some evergreen tropical rainforest may be replaced by more seasonal vegetation (seasonal forest or savanna) in a warmer world. But there is a dearth of experimental information on the phenology of tropical plants, compared with temperate and boreal regions (Bradley *et al.*, 2011). A preliminary study of the Amazon Basin yielded two distinct responses with solar radiation being the cue for phenology in upland (more mesic) forests, and precipitation being the regulator in savanna ecosystems.

Fire

Fire is the prinicipal cause of vegetation 'disturbance' (removal of biomass) worldwide and endemic to all biomes except the wettest and the most sparse (driest or coldest) (Bowman *et al.*, 2009; Krawchuk *et al.*, 2009; Harrison *et al.*, 2010b). There is a vast literature on the modeling of fire at the landscape scale, but there have been relatively few attempts to model fire in a generic, global way. Nevertheless, several DGVMs do now include representations of fire. The basic principle of fire modeling is that it must be represented as a feedback process. Vegetation state (availability and dryness of fuels) controls the probability of burning and the rate, intensity, and duration of fire spread. Fire in turn, depending on its extent and properties, impacts the vegetation state by removing fuel, killing plants (or simply temporarily removing their leaves in the case of perennial herbaceous plants and 'resprouting' woody PFTs), and allowing the establishment of new plants in the case of 'seeding' PFTs.

The development of fire modeling during the early development of PFTs was handicapped by a lack of a consistent global historical dataset that could be used to test and refine models. This situation has now changed, with the advent of global remotely sensed data products (Giglio *et al.*, 2010; van der Werf *et al.*, 2010). At least two distinct strands of process-based, global fire modeling have emerged. One draws on the work of Rothermel (1972), who developed the first physically based approach to fire modeling in the context of operational fire management. Global modeling can use Rothermel's equations directly although it requires many quantities, which are measured in operational work, to be estimated somehow from the modeled vegetation properties and prescribed traits of PFTs. This approach was pioneered by the group of Ron Neilson at the US Department of Agriculture Forest Service (Lenihan *et al.*, 1998) and adopted by Thonicke *et al.* (2010) and Prentice *et al.* (2011a) in the LPJ family of models. An alternative, less complex process-based formulation, tailored more specifically to the requirements of DGVMs, was developed by Arora and Boer (2005b) for the CTEM.

It is likely that the representation of fire in DGVMs will evolve considerably during the coming years. One major complication needing to be tackled is the role of human activities. Although many human-set fires can be considered to 'pre-empt' fires that would otherwise have started naturally (especially in highly fire-prone ecosystems such as savannas), this is a great simplification of a complex reality. On the one hand, models that disregard human ignitions not surprisingly fail to reproduce 'deforestation fires' linked to El Niño events in Amazonia and southeast Asia. On the other hand, contemporary human land use in many regions has a powerful suppressive effect, which current models underestimate (Prentice *et al.*, 2011a).

Wetlands and Drylands

Vegetation growing in hydrologically extreme environments (very wet or very dry) has been the last to be considered seriously in DGVM development (Ivanov *et al.*, 2008; Meron, 2011). Some DGVMs now have explicit representation of wetlands, with distinct PFTs to represent plants adapted to waterlogged conditions (e.g., Wania *et al.*, 2009a, b). Drylands have not been treated in as much detail. But with drylands expanding in some regions (Naito and Cairns, 2011), in part, because of changes in land use (i.e., Knapp *et al.*, 2008; Bollasina and Nigam, 2011), and with the potential for further expansion in some regions due to climate change (MacDonald, 2010; Tietjen *et al.*, 2010; Gutzler and Robbins, 2011), the modeling of drylands is assuming greater importance.

Fire is a crucial intermediary in some seasonally dry environments (e.g., mediterranean-type and savanna ecosystems: Bond *et al.*, 2003, 2005), and is modeled explicitly in several DGVMs. In the drylands, by contrast, fire is rare or absent because of the lack of a continuous fuel bed. Vegetation in desert environments depends on several processes that are not represented at all in most DGVMs. Facilitation or the 'nursery effect' describes the ability of pioneering plants to provide shade, shelter, and soil moisture to less drought-tolerant plants (Prentice and Werger, 1985; Brooker *et al.*, 2008). Once established, these newly established plants provide further shade, shelter, and moisture to additional plants, resulting in the grouping of vegetation in clusters across arid landscapes. Clustering is also promoted by microtopography, leading to a separation of 'run-off' and 'run-on' zones, the latter able to support (denser) vegetation (Vasquez-Mendez *et al.*, 2010; Aerts *et al.*, 2006). Hydraulic redistribution is the passive release of water from deep tap roots into upper soil horizons when hydraulic conductivity decreases at night (Katul and Siqueira, 2010). It is most pronounced in areas with a strong gradient of soil water potential within the rooting zone. Water is generally distributed from wet to dry regions within the soil profile. Plants with deep roots are therefore able to provide soil moisture to neighboring shallow-rooted plants. Both facilitation and hydraulic redistribution can result in a positive feedback on plant establishment.

Surprisingly, simulations including hydraulic redistribution have been initially conducted not for arid regions but rather for wet environments like the Amazon Basin that experience only periodic drought. Recent simulations incorporating the effect indicate that the overall consequences for vegetation performance depends primarily on the duration of drought (Wang, 2011; Wang *et al.*, 2011). When drought periods are short, hydraulic redistribution increases transpiration and NPP; however, when drought is prolonged, hydraulic redistribution ultimately reduces NPP because it accelerates the depletion of deep soil water storage. Simulations of the Amazonian forest incorporating the effects of deep roots in extracting water from soils near the groundwater table show that deep root uptake of water was limited, due to

the high concentration (70–90%) of clay in the soils and a high soil-to-root resistance to water uptake (Markewitz *et al.*, 2010). The lack of representation of facilitation, run-off/run-on, and hydraulic redistribution is a likely contributory explanation for the relatively poor ability of current DGVMs to represent the vegetation in arid and semiarid regions such as the southwestern USA, southern Africa, and interior Australia.

Another limitation of DGVMs' ability to simulate dryland vegetation is the very short list of PFTs used by current models. The diversity of plant adaptations to life in the desert is an often remarked-on feature of these environments, but today's DGVMs do not account for characteristic traits of dryland plants such as succulence, leaflessness, traits that avoid excessive absorption of solar radiation such as reflective hairs, waxy epidermis, and vertical leaf insertion, and the specialized form of photosynthesis known as Crassulacean Acid Metabolism.

Fields of Application of DGVMs

Carbon Cycle

The story of modern carbon cycle science begins with the first precise measurement of annually and seasonally varying CO_2 concentrations by Keeling (1958), who was thus able to confirm a hypothesis (that CO_2 concentration was rising) which had first been put forward more than half a century earlier. Since the 1980s, several overlapping networks of remote atmospheric measurement stations (for CO_2, other trace gases and valuable tracers of the carbon cycle such as the stable carbon isotope composition of atmospheric CO_2) have been maintained. The resulting global network of measurement stations contains far more information than could be extracted from Keeling's original two, at Mauna Loa (Hawaii) and the South Pole. These records already showed that in addition to the year-on-year increase of CO_2 concentration, there is a seasonal cycle, which is much stronger at Mauna Loa than at the South Pole and which shows opposite phase in the two hemispheres. Keeling quickly realized that this cycle is caused by the 'breathing' (seasonal CO_2 uptake and release) of the terrestrial biosphere. The great value of the more extensive CO_2 observation network that exists today is that it provides more highly resolved information about the latitudinal variations in this seasonal cycle, and even about longitudinal patterns (albeit these are smoothed by rapid mixing), which give information about the spatial patterns of sources and sinks of CO_2.

This information is valuable as a test of DGVMs, since they explicitly predict a spatial and seasonal pattern of CO_2 uptake and release. To use this information, however, it is necessary to involve an atmospheric transport model to translate the signals at grid cells to a combined signal at a remote place in the atmosphere. A simple recipe for this comparison, presented by Heimann *et al.* (1998), provides a powerful (but under-used) benchmark for DGVMs and coupled climate–carbon cycle models (Prentice *et al.*, 2000a, Cadule *et al.* 2010). The ability to reproduce this seasonal signal at different latitudes is a major achievement of DGVMs.

The rate of increase in CO_2 concentration varies greatly from year to year. These variations bear the imprint of the El Niño-Southern Oscillation (Bacastow, 1976), and of large volcanic eruptions such as Mount Pinatubo that have significant transient effects on climate (Mercado *et al.*, 2009). The interannual variability of the CO_2 growth rate has been shown to be dominated by variations in the fraction of anthropogenic CO_2 that is taken up by the land (McGuire *et al.*, 2001; Denman *et al.*, 2007). Thereby it provides a further test and demonstration of the capabilities of DGVMs (Cadule *et al.* 2010).

The longer term mean rate of uptake of CO_2 by the land, which can be separated from the ocean component by independent measurements of O_2 concentration or the stable isotope composition of atmospheric CO_2 (Keeling and Shertz, 1992; Keeling *et al.*, 1996; Battle *et al.*, 2000; Prentice *et al.*, 2001), is a further quantity that can be predicted by DGVMs. It is modeled as a direct consequence of the continuing CO_2 rise and its fertilizing effect, combined with the residence time of carbon in the biosphere (Friedlingstein *et al.*, 1995; Kicklighter *et al.*, 1999; McGuire *et al.*, 2001; Prentice *et al.*, 2001). The residence time is long enough to ensure that the size of the carbon pool does not keep up with the rate of increase in growth. This historical rate of carbon uptake by the land is potentially an important constraint on models' future projections of carbon uptake (Arora *et al.*, 2009). In Friedlingstein *et al.* (2006), the models with the lowest and highest uptake rate of atmospheric CO_2 in the future, respectively, under- and overpredicted the rate of CO_2 uptake during recent decades, suggesting that this very simple carbon budget criterion could provide one way to identify poorly performing models.

The IPCC Fourth Assessment Report (Denman *et al.*, 2007) highlighted the uncertainty in the modeled terrestrial carbon cycle feedbacks, represented by these coupled climate–carbon cycles. It is a matter of urgency that the uncertainty be reduced, which will be possible through systematic comparison of model outputs with the most relevant observations – i.e., atmospheric carbon cycle measurements.

Hydrology

Hydrology is a relatively new field of application for DGVMs (see e.g., Rost *et al.*, 2008; Murray *et al.*, 2011). But *global hydrological models* have been developed in parallel with DGVMs, primarily for the assessment of large-scale regional changes in freshwater availability in a changing climate, and contain many of the same processes (e.g., Arnell, 1999a, b). The potential advantage of using a DGVM for water resources assessment is that a DGVM explicitly represents changes in vegetation cover, leaf area, stomatal conductance, and transpiration resulting from changes in climate and atmospheric CO_2 concentration, in a biologically and physically consistent way.

However, additional work needs to be done to bring DGVMs 'up to speed' for hydrological applications. Foremost, it is essential to use hydrological observations (streamflow measurements) as a benchmark for the terrestrial runoff simulated by DGVMs (Murray *et al.*, 2011). This will bring a dual benefit, as runoff is highly sensitive to the representation of vegetation (being a relatively small fraction, about 1/3 of

precipitation globally and much less in the drylands) and is therefore powerful as a test for the model's performance in simulating vegetation function. Furthermore, hydrological modelers have developed methods to improve runoff simulation by taking account of heterogeneity in the hydrological properties of the land surface. This approach has been used in LSMs (Liang *et al.*, 1994; Liang and Xie, 2001). In one case where the concept of variable infiltration capacity was applied to the LPJ DGVM, it was shown to greatly improve the model's simulation of runoff (Li and Ishidaira, 2011). And just as atmospheric transport modeling is the key to bringing DGVM predictions into comparison with atmospheric measurements, river routing modeling will be the key to bringing DGVM predictions into comparison with streamflow measurements.

Trace Gases and Aerosols

The terrestrial biosphere is the predominant source of the powerful greenhouse gases methane (CH_4) and nitrous oxide (N_2O); the precursors of ozone (O_3), which (when present in the surface atmosphere) is a potent greenhouse gas as well as a toxic pollutant, and reactive gases that indirectly influence greenhouse gas concentrations through atmospheric chemistry, including carbon monoxide (CO), nitric oxide (NO) and volatile organic compounds – predominantly hydrocarbons (isoprene, C_5H_8 and monoterpenes, $C_{10}H_{16}$) and methanol (CH_3OH). These natural atmospheric constituents are produced variously by direct emission from leaves, soil microbial transformations of organic matter derived from plants, and in fires. Global modeling of the terrestrial sources (and in some cases, sinks) of these compounds is a natural extension of DGVMs. Arneth *et al.* (2010a) reviewed the state of the art in this field.

Modeling methane emissions has proved to be a complex challenge as it required the explicit simulation of wetlands, including PFTs adapted to low-oxygen environments, and the microbial processes specific to wetlands and waterlogged mineral soils (Gedney *et al.*, 2004; Wania *et al.*, 2009a, b, 2010; Spahni *et al.*, 2011; Huntingford *et al.*, 2011). For Earth system modeling, it is important to model not only processes within wetlands and wet soils, but also the hydrological conditions that allow permanent wetlands to be sustained (Ringeval *et al.*, 2010). This is a new frontier for DGVMs.

In addition to trace gases, terrestrial ecosystems provide the main source for some types of atmospheric aerosol that influence climate. Airborne dust has many different, both 'good' and 'bad' consequences, including providing a natural source of iron to fertilize marine ecosystems (Harrison *et al.*, 2001). Dust is emitted from unvegetated parts of the land surface. Its emissions are thus controlled by vegetation. Dust modeling depends critically on modeling vegetation density and phenology (Tegen *et al.*, 2002) as well as geomorphic factors that control the emission of dust from nonvegetated surfaces (Bullard *et al.*, 2011). VOC emissions also contribute to secondary organic aerosol, a complex and poorly understood atmospheric constituent whose physical importance has only recently been recognized. Earth system models are just beginning to integrate these many aspects of the interaction of the atmosphere and biosphere, all of which depend on the representation of plant and ecosystem processes by the embedded DGVMs.

Biophysical Land–Atmosphere Interactions

Responses of vegetation to changes in CO_2 and climate bring about changes in the physical properties of the land surface, which in turn influence climate. These feedback effects become explicit when DGVMs are coupled with climate models (e.g., Foley *et al.*, 1998; Betts, 2001; Betts *et al.*, 2007b; Cowling *et al.*, 2007a; Roy, 2009; Lunt *et al.*, 2010; Quillet *et al.*, 2010; Delire and de Noblet-Ducoudré, 2011). The three most important land-surface variables that influence local climate, and change according to vegetation properties, are the surface albedo; surface energy partitioning; and surface roughness. Surface albedo (the reflectivity of the Earth's surface) varies according to vegetation type, with deep green vegetation (such as evergreen forests) absorbing more solar radiation than light-colored vegetation such as grassland or desert. Typical surface albedos are for coniferous forest 0.05–0.15, deciduous forest 0.15–0.20, grassland 0.16–0.26, cropland 0.18–0.25, and bare ground 0.05–0.40 depending on soil type and wetness (Bonan, 2002). Snow cover dramatically increases surface albedo in treeless vegetation, but less so in forests (especially boreal evergreen forests) where dark plant surfaces remain above the snowpack (snow masking). Surface energy partitioning refers to the amount of latent heat produced relative to sensible heat. Latent heat is the energy required to change the phase of water (denoted Q_E); sensible heat is the energy available to warm the surface (denoted Q_H). The partitioning of surface energy is expressed by the Bowen ratio (Q_H/Q_E). The Bowen ratio is influenced by the rate of transpiration, which in turn depends on leaf area index (the ratio of leaf surface area to underlying ground area) and stomatal conductance (per unit leaf area). Surface roughness is related to the theoretical height at which wind speed above the canopy is zero, defining the maximum height of atmospheric mixing. In general, surface roughness for grass is smaller than for forests. Surface roughness ultimately influences surface energy partitioning as a wind gradient (air mixing) is required to maintain the gradient that drives transpiration (low vapor pressure of water outside the leaf, relative to inside). All three properties can be altered by vegetation dynamics and thus can potentially induce feedback responses in a climate model.

Uncertainties in the behavior of the physically coupled vegetation–atmosphere system, however, are at least as large as those for the carbon cycle. The strength of the coupling between land-surface properties and the hydrological cycle in the atmosphere differs greatly among climate models (Seneviratne *et al.*, 2006; Denman *et al.*, 2007), so the simulated climate effects of a given vegetation change can be very different depending on the climate model. Climate–vegetation feedbacks in a future climate have been indicated to cause severe shrinkage of the Amazon Basin rainforests (Betts *et al.*, 2004; Cox *et al.*, 2000, 2004; Harris *et al.*, 2008). But a change in the recycling of water in tropical rainforests can in principle either increase or decrease the total areal coverage of vegetation, depending on its cause. Recent simulations in general have shown less Amazon dieback (Huntingford *et al.*, 2008;

Rammig *et al.*, 2010; Salazar and Nobre, 2010) than earlier ones. The general uncertainty regarding biophysical feedbacks also afflicts attempts to quantify the consequences of past and present land use (Betts *et al.*, 2007a), which have shown major discrepancies among models (Pitman *et al.*, 2009).

One question that has received considerable attention, but incomplete resolution, is the influence of 'physiological forcing' of CO_2 on climate (Boucher *et al.*, 2009; Andrews *et al.*, 2011). Physiological forcing refers to changes in land-surface properties brought about by the response of plants to changes in CO_2 concentration, which can be contrasted with 'physical forcing' of climate due to the greenhouse effect of CO_2. The physical forcing is certainly larger, but modeling suggests that physiological forcing is not negligible. Early sensitivity experiments, where land-surface properties in a climate model were varied, suggested that reduced stomatal conductance due to increased CO_2 concentration led to reduced transpiration and increased surface warming (Sellers *et al.*, 1997; Bounoua *et al.*, 1999). This is not the whole story, however, because physiological effects of CO_2 (increased NPP and reduced stomatal conductance) can also lead to an increase in LAI (Cowling and Field, 2003), with compensatory effects on transpiration and climate (Betts *et al.*, 1997). In general, if LAI increases then the Bowen ratio decreases, due to an increase in rate of transpiration and production of latent heat, opposite to the effect of reduced stomatal conductance (Reyers *et al.*, 2011).

Bounoua *et al.* (2010) continued the theme of using sensitivity experiments to assess the physiological forcing of CO_2. This study attracted media attention because in one scenario, greatly enhanced evapotranspiration led to a substantial reduction in the overall global warming due to CO_2. However, this scenario assumed that the down-regulation of Rubisco capacity would lead to a large increase in LAI. There is no evidence for such an effect.

Many additional vegetation–climate interactions are likely to arise as a result of the physical forcing of climate. For example, a longer snow-free season will create a positive albedo feedback in high latitudes. Forest expansion in high latitudes, as simulated by DGVMs (Cramer *et al.*, 2001), will create additional positive feedback due to snow masking. And if DGVM-simulated vegetation dynamics are allowed to proceed to equilibrium with a new, high CO_2 concentration, one set of climate model simulations has suggested that the 'greening' effect promoted by reduced stomatal conductance may lead to additional warming, due to albedo feedback (O'ishi *et al.*, 2009).

The balance of conflicting effects of rising CO_2 is thus still not clear in general, and is likely to differ among regions. Yet this balance has wider implications, for example, for hydrology (Betts *et al.*, 2007b). Gedney *et al.* (2006) attributed an increase during recent decades in runoff relative to precipitation to the effect of CO_2 on stomatal conductance, but this conclusion has been disputed by Piao *et al.* (2007) and several others. Resolution of this issue will not come through more sensitivity experiments, but it will require careful comparative evaluation of the various responses of vegetation to CO_2 in DGVMs. Vegetation–climate feedbacks are a fertile field for the application of DGVMs, but large uncertainties are likely to persist until ways are found to constrain the strength of the coupling between the land surface and the atmosphere. In addition, there may be interactions between processes at the local ecosystem versus the landscape scale that could bring about threshold behavior in the coupled system, but which are not yet represented in DGVMs (e.g., Brovkin *et al.*, 2009; Dekker *et al.*, 2010; Rietkerk *et al.*, 2011).

Vegetation Structure, Composition, and Biodiversity

There have been very few analyses of the global impacts of projected twenty-first century environmental changes on vegetation distribution, structure, and composition using DGVMs (but see: Neilson *et al.*, 1998; Scholze *et al.*, 2006; Gonzalez *et al.*, 2010). This is surprising, given that this was one of the principal applications originally envisaged for DGVMs. It may be that the lack of regional detail implied in the use of a limited set of PFTs, and other deficiencies noted above, has inhibited such effort. If so, then it will be important to develop more realistic representations of the diversity of plants. The authors suggest that this will be possible, given the dramatic recent improvements in the availability of trait data on wild plants. It will require a bold increase in the diversity of plants represented within DGVMs, whether as discrete PFTs, or via plasticity in trait values, or (most likely) a combination of the two.

Meanwhile, the dominant type of model used for regional assessments of climate change impacts on biodiversity has been quite different, namely *species envelope models* developed for one species at a time based on empirical relationships between geographic distributions and climate variables (e.g., Thomas *et al.*, 2004; Thuiller *et al.*, 2005). The list of problems and deficiencies with this approach is rather long (Pearson and Dawson 2003; Dawson *et al.*, 2011). Nevertheless, there is valuable climatic information in species range limits. Alternative approaches have been developed whereby 'bioclimate variables' for predictive purposes are specified *a priori* from process considerations, then species' limiting values fitted empirically (e.g., Sykes *et al.*, 1996). Better still, independent observations and experimental evidence can be used to predict species' limits, so the predictions can be tested using the observed limits (e.g., Chuine and Beaubien, 2001; Morin *et al.*, 2007).

One limitation of conventional species envelope models is that they provide no way to represent the potential modification of range limits by CO_2 concentration. Keenan *et al.* (2011) showed that this makes an important difference. Predictions of the effect of climate change on tree species distributions in the Mediterranean region were found to depend critically on whether or not raised CO_2 concentration was allowed to compensate for decreased rainfall. There is a potential to merge elements of process-based range limit modeling with representations of physiological processes used in DGVMs. This has been done for regional applications of the LPJ-GUESS DGVM (Hickler *et al.*, 2012) and provides (among other benefits) a natural way to build in CO_2 effects on species' potential ranges.

A quite different approach to modeling has attempted to represent explicitly the controls on the number of species that can be 'packed' into a given environment (Kleidon and

Mooney, 2000; Kleidon *et al.*, 2009; Reu *et al.*, 2011). This approach has been successful in reproducing the broad global patterns of plant species diversity, and deserves consideration as an element of future DGVMs.

Interpreting the Past: Toward a Deeper Understanding of the Natural World

Vegetation, Climate, and Carbon Cycle Feedbacks During the Quaternary Period

The major controls on the climate variations of approximately the past million years are relatively well understood. There is a well-developed theory for how variations in the astronomical parameters of the Earth (orbital eccentricity, precession and axial tilt) combine with variations in ice-sheet extent, which are heavily lagged responses to orbital variations, and natural variations in the concentrations of the greenhouse gases CO_2, methane, and nitrous oxide, which are directly measured in the air bubbles trapped in ice cores (the record is now 850,000 years long): see the review by Harrison and Bartlein (2012). The continuous time sequence of climate changes in different latitude bands and regions has been modeled over the last deglaciation, i.e., the period since the last glacial maximum, 21 thousand years ago, using a fast climate model (Timm and Timmermann, 2007). Many standard simulations with climate models have also been made for the iconic periods, the LGM and the mid-Holocene (six thousand years ago), in successive phases of the Palaeoclimate Modelling Intercomparison Project. These two periods have also been a major focus for the synthesis and analysis of palaeodata (e.g., Prentice *et al.*, 2000b; Bartlein *et al.*, 2011).

The main factor that changed between the mid-Holocene and recent preindustrial time was precession. Perihelion in the early to mid-Holocene occurred in boreal summer, leading to increased summer and decreased winter insolation in the mid-latitudes of the Northern Hemisphere. This in turn led to greater land–sea contrast and expanded monsoons in the northern tropical continents, exemplified by the 'green Sahara' as manifested in pollen and plant data and relics of former lakes, swamps, megafauna, and human habitation. Much has been written, based on simulations without dynamic vegetation, about the requirement for biophysical feedback (specifically, the reduced surface albedo of a wetter and more vegetated landscape) to explain the full extent of the monsoon expansion (e.g., de Noblet-Ducoudré *et al.*, 2000). However, the more recent application of DGVMs within climate models has changed the picture: the feedbacks appear to be less strong in the available models, and insufficient to explain the data (Harrison and Bartlein, 2012; Wolff *et al.*, 2012). Much depends on the specification of surface albedo for different vegetation types, for example, when surface albedo was modeled based on leaf area index in combination with the state of the soil beneath the forest canopy (increased organic matter, litter, and standing dead biomass), Vamborg *et al.* (2011) could better simulate the increase in Saharan precipitation required to account for the data. Another controversial area concerns the abruptness, or otherwise, of the reinstatement of the Sahara desert, after the mid-Holocene.

The field evidence is contradictory. Strong positive feedbacks could in principle allow an abrupt transition from a 'green' to a 'brown' state, and this type of dynamics has been simulated with an Earth System model that includes a simple DGVM (Claussen *et al.*, 1999). But analysis with the LPJ DGVM has suggested that an abrupt transition, in response to gradually declining precipitation, may be a consequence of vegetation dynamics alone (Liu *et al.*, 2006, 2007).

The LGM is interesting for the carbon cycle because of the low CO_2 concentration (180–190 ppm, compared with 260–280 ppm during the preindustrial Holocene); long-standing evidence from stable carbon isotopes (reviewed by Prentice and Harrison, 2009) that the terrestrial biosphere stored less carbon then, compared with the Holocene; and new evidence from stable oxygen isotopes that terrestrial GPP was even more strongly reduced in proportional terms (Ciais *et al.*, 2011). The mere presence of large continental ice sheets in North America and northern Europe cannot account for reduced carbon storage at the LGM because a comparable area of continental shelf was exposed, mainly in the tropics. A variety of modeling studies has led to the conclusion that the greatly reduced global forest cover during the LGM can be explained only when physiological effects of CO_2 are considered (Cowling, 1999; Cowling and Sykes, 1999; Bond *et al.*, 2003; Harrison and Prentice, 2003), and that the reduced carbon storage cannot be explained solely by the colder and drier climate (Prentice and Harrison, 2009). In a study using the LPX DGVM and results from four climate models, Prentice *et al.* (2011c) showed that realistic reductions in forest cover, GPP, and carbon storage could be simulated for the LGM. Ciais *et al.* (2011) further showed that the reduction in carbon storage may have been partly compensated by increased storage of inert carbon in permafrost soils.

Paleoecological records contain far more information than these limited 'snapshot' studies have used so far. Now that transient climate model simulations over multimillennial periods have become feasible, there is a potential to apply DGVMs through periods of current focus for climate modeling such as the last millennium, the last deglaciation, and the abrupt warming episodes (Dansgaard–Oeschger events) that punctuated the last glacial period. Large-scale analyses of data from these periods have yielded surprising insights, such as the sensitive dependence of biomass burning on small variations in temperature during recent centuries, the very large and rapid decline in biomass burning since the mid-nineteenth century (Marlon *et al.*, 2008; Wang *et al.*, 2010a), and the rapid responses of many global indicators including greenhouse gas concentrations, dust concentrations, and fire regimes to D–O events (Arneth *et al.*, 2010b; Hopcroft *et al.*, 2011). There is also evidence of remarkably rapid responses of vegetation composition to D–O events (Harrison and Sanchez-Goñi, 2010). Reproducing such phenomena will be an important challenge for DGVMs.

Vegetation–Climate Interactions in Deep Geological Time

Studies of the Quaternary period, especially the period from the LGM onwards, have two great advantages: a wealth of accurately dated information, and the fact that very little

macroevolution or natural extinction has taken place on this time scale (Dawson *et al.*, 2011) so that observed biotic transitions can be interpreted based on knowledge of the present-day biota. But there is increasing interest in the climate and vegetation of more distant geological periods, especially periods when global temperatures and CO_2 levels were higher than today.

Climate models can legitimately be used for any geological epoch for which information on paleogeography and atmospheric composition can be supplied. But the land surface is a sticking point for deep time studies. Land-surface models of all kinds, including DGVMs, incorporate information derived from extant plants. We do not yet have the tools to analyze ecosystems composed of ancient plants (Shellito and Sloan, 2006a, b). A number of studies have therefore simply used models designed for modern vegetation. One of the first explicit simulations of ancient ecosystems was an assessment of the recovery time of terrestrial productivity following the meteorite impact that ended the Cretaceous (Lomax *et al.*, 2001). This study accepted the modern biogeography of five PFTs (evergreen and deciduous needle-leaved and broad-leaved trees and a herbaceous understory), so it did not account for the fact that angiosperm trees – with their higher rates of xylem water flow – were far less important in the Cretaceous than they have become since. Donnadieu *et al.* (2009) used the LPJ DGVM, without modification, to simulate vegetation–climate feedbacks in the Mesozoic; as a result they simulate grasslands (including C_4 grasses, with much lower stomatal conductance than C_3 plants) that had not yet appeared on Earth. A study of vegetation–climate feedbacks in the Paleozoic (Horton *et al.*, 2010) removed the grass PFTs from a biogeography model, but nevertheless assumed that the Paleozoic woody plants were appropriately represented by modern trees, despite major differences in leaf and wood anatomy. These are not isolated examples of studies whose findings are rendered moot by the use of models that have been developed for modern plants.

We know from the fossil record that ecosystems of past geological periods were unlike modern ones in functionally important ways (e.g., Hermann *et al.*, 2011). Angiosperms did not expand greatly until the beginning of the Cenozoic (Magallon and Castillo, 2009), but during the Cenozoic angiosperms largely out-competed the gymnosperms that were widespread forest dominant in the late Cretaceous (Coiffard and Gomez, 2010). The evergreen needle-leaved gymnosperm trees of today are cold-adapted and generally confined to northern temperate and boreal latitudes, whereas those of the Cretaceous were warm-adapted and found at all latitudes (Escapa *et al.*, 2011; Hermann *et al.*, 2011; Leslie, 2011). Tree ferns and ground-covering surface ferns were plentiful in the Cretaceous, whereas today they are a very minor component of vegetation globally (DiMichelle and Phillips, 2002). These situations raise intriguing questions about the extent to which the predominant plant taxa in different epochs were particularly well adapted to environmental conditions and, conversely, the extent to which past climates and environments were influenced by the biophysics of the plants of their time. To what extent did ancient trees recycle water inward from coastal regions? Did the properties of ancient PFTs contribute to the higher atmospheric humidity levels reconstructed for the Mesozoic? Did falling CO_2 levels, with the attendant need for more rapid water conduction, drive the expansion of the angiosperms during the Cenozoic? Such questions cannot be addressed with contemporary plant models.

Indicative answers to 'what if' questions about ancient ecosystems can sometimes be provided nevertheless by sensitivity analyses using currently available models. Herold *et al.* (2011), wishing to determine whether the Australian monsoon climate could be simulated for the early to mid-Miocene, ran simulations with grassland covering the Australian continent; forest covering the continent; and more realistic cases incorporating broad-leaved forests, shrublands, and mixtures of these vegetation types. Lunt *et al.* (2007) simulated the distribution of C_4 grasses in the late-Oligocene and found that the model predicted greater C_4 distributions relative to the preindustrial (curiously, a result opposite to what the paleodata show). Global simulations without C_4 grass were carried out to assess the implications of C_4 grass evolution for regional energy and water exchanges during the Miocene (Cowling *et al.*, 2007b). Cowling *et al.* (2009) also conducted hypothetical global simulations in which all plants used C_4 photosynthesis, to determine the consequences for the global distribution of sensible and latent energy.

Some recent simulations of climate and vegetation during the late Miocene have paid more specific attention to representing past vegetation types. Boreal forest was simulated to extend beyond 80° N and temperate forest above 60° N, consistent with the fossil record (Pound *et al.*, 2011). Pound *et al.* (2011) concluded that tropical latitudes were characterized by altered distributions of tree types found there today, whereas northern latitudes contained tree types not found there today. François *et al.* (2011) incorporated nonextant warmth-requiring tree PFTs in their modeling of the late Miocene in Europe. Several simulations of Pliocene climate and vegetation have been conducted with DGVMs embedded within GCMs (Jost *et al.*, 2009; Salzmann *et al.*, 2008; Haywood *et al.*, 2002; Haywood and Valdes, 2006), representing a flora more similar to that of the Quaternary, so fewer problems arise with the application of current models.

Confusion can be caused if words we use today to describe vegetation are applied to ancient vegetation. For example, terms like 'xeric adapted' have been used to apply to Cretaceous plants. But most of the extant desert plant lineages (e.g., agaves, ice plants, and cacti) evolved more recently, between the Eocene and the Miocene (Arakaki *et al.*, 2011). Furthermore, large variations in atmospheric CO_2 concentration have occurred during the period since the origin of land plants. These have been linked to large variations in the density and morphology of stomates (e.g., Franks and Beerling, 2009) implying that plant water relations have been very different indeed at past high CO_2 levels. Thus, it cannot be assumed that morphological features relating, for example, to water stress have the same quantitative significance in the deep past as they do today. To understand the function of ancient plants and vegetation, it will be necessary to analyze and model plant adaptations from a process-based, biophysical perspective.

In conclusion, the authors propose that the modeling of ancient ecosystems will become a highly rewarding field of investigation if serious attention is paid to the representation of ancient plants in land-surface models and DGVMs. The key

point is that modeling needs to be based on biophysical principles, and on measurable traits of ancient plants. It should *not* be based on the observed distributions and habitat preferences of extant plants. Many relevant morphological traits fossilize well, and there is a growing literature based on the compilation of measurements such as leaf size and dissection, petiole diameter, bole diameter, root length and diameter, stomatal density and aperture, leaf anatomy and the diameters of xylem conducting elements (which govern the hydraulic properties of stems) in fossil plants. All of these traits have a direct relevance to plant biophysics and could be used in a 'next generation' DGVM, if such a model were to be formulated in terms of the observed characteristics of plants – whether modern or ancient.

Prospects

DGVMs are a necessary tool for modeling and understanding vegetation responses to environmental changes, past, present, and future. They are also needed to link different kinds of Earth System observations (such as climate, the carbon cycle, land-surface properties, and atmospheric composition), and they are an indispensable part of the toolkit required for modeling the interactions among different Earth System components. They are still a relatively new class of model, however, and they need a good deal of further development. There is a potential to reduce the current (large) discrepancies among models by paying greater attention to systematic benchmarking of models against large-scale observations (including remotely sensed data, atmospheric measurements and hydrological measurements). The utility of DGVMs could be expanded by making better use of the 'data revolution' in comparative plant ecology. Some uncertainties in modeling reflect uncertainties in the underlying science (for example, carbon–nitrogen cycle coupling). The authors suggest that modeling has a proactive role to play in such cases, by analyzing the consequences of alternative hypotheses rather than simply assuming a particular interpretation of the evidence.

The authors propose that the next generation of DGVMs will be not necessarily be more 'complex' than the present one, but future models will take better account of the observed diversity and morphology of plants and their varied responses to environmental triggers, including (especially) aridity and fire. DGVMs will continue to be applied to past vegetation and climate, with emphasis on transient changes in plant community composition and fire in recent climates, and the interplay of plant morphological and functional evolution with climate, atmospheric composition and the global hydrological cycle in ancient climates. This last objective will require closer attention to the biophysical representation of ancient plants.

Appendix

List of Courses

1. Climatology
2. Biogeography
3. Global Change
4. Global Ecology
5. Paleobiology

References

Aerts R, Maes WH, November E, *et al.* (2006) Surface runoff and seed trapping efficiency of shrubs in a regenerating semiarid woodland in northern Ethiopia. *Catena* 65: 61–70.

Ainsworth EA and Long SP (2005) What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist* 165: 351–372.

Albert KR, Mikkelsen TN, Michelsen A, Ro-Poulsen H, and van der Linden L (2011) Interactive effects of drought, elevated CO_2 and warming on photosynthetic capacity and photosystem performance in temperate heath plants. *Journal of Plant Physiology* 168: 1550–1561.

Alton P, Mercado L, and North P (2007) A sensitivity analysis of the land-surface scheme JULES conducted for three forest biomes: Biophysical parameters, model processes, and meteorological driving data. *Global Biogeochemical Cycles* 20. GB1008.

Andrews T, Doutriaux-Boucher M, Boucher O, and Forester PM (2011) A regional and global analysis of carbon dioxide physiological forcing and its impact on climate. *Climate Dynamics* 36: 783–792.

Arakaki M, Christin PA, Nyffeler R, *et al.* (2011) Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences of the United States of America* 108: 8379–8384.

Arnell NW (1999a) A simple water balance model for the simulation of streamflow over a large geographic domain. *Journal of Hydrology* 217: 314–335.

Arnell NW (1999b) Climate change and global water resources. *Global Environmental Change* 9: S31–S49.

Arneth A, Harrison SP, Zaehle SP, *et al.* (2010b) Terrestrial biogeochemical feedbacks in the climate system. *Nature Geoscience* 3: 525–532.

Arneth A, Sitch S, Bondeau A, *et al.* (2010a) From biota to chemistry and climate: Towards a comprehensive description of trace gas exchange between the biosphere and atmosphere. *Biogeosciences* 7: 121–149.

Arora V (2002) Modeling vegetation as a dynamic component in soil–vegetation–atmosphere transfer schemes and hydrological models. *Reviews of Geophysics* 40: 1006.

Arora VK and Boer GJ (2005a) Parameterization of leaf phenology for the terrestrial ecosystem component of climate models. *Global Change Biology* 11: 39–59.

Arora VK and Boer GJ (2005b) Fire as an interactive component of dynamic vegetation models. *Journal of Geophysical Research – Biogeosciences* 110. G02008.

Arora VK and Boer GJ (2006) Simulating competition and coexistence between plant functional types in a dynamic vegetation model. *Earth Interactions* 10: 10.

Arora VK, Boer GJ, Christian JR, *et al.* (2009) The effect of terrestrial photosynthesis down regulation on the twentieth-century carbon budget simulated with the CCCma Earth System Model. *Journal of Climate* 22: 6066–6088.

Atkin OK, Atkinson LJ, Fisher R, *et al.* (2008) Using temperature-dependent changes in leaf scaling relationships to quantitatively account for thermal acclimation of respiration in a coupled global climate–vegetation model. *Global Change Biology* 14: 2709–2726.

Atkin OK and Tjoelker MG (2003) Thermal acclimation and the dynamic response of plant respiration to temperature. *Trends in Plant Science* 8: 343–351.

Bacastow RB (1976) Modulation of atmospheric carbon cycle by the southern oscillation. *Nature* 261: 116–118.

Ball J, Woodrow I, and Berry J (1987) A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. *Progress in Photosynthesis Research*, vol. IV, pp. 221–224. Dordrecht: Martinus Nijhoff.

Bartlein PJ, Harrison SP, Brewer S, *et al.* (2011) Pollen-based continental climate reconstructions at 6 and 21 ka: A global synthesis. *Climate Dynamics* 37: 775–802.

Battle M, Bender ML, Tans PP, *et al.* (2000) Global carbon sinks and their variability inferred from atmospheric O_2 and $\delta^{13}C$. *Science* 287: 2467–2470.

Betts RA (2001) Biogeophysical impacts of land use on present-day climate: Near-surface temperature change and radiative forcing. *Atmospheric Science Letters* 2: 39–51.

Betts RA, Boucher O, Collins M, *et al.* (2007b) Projected increase in continental runoff due to plant responses to increasing carbon dioxide. *Nature* 448: 1037–1041.

Betts RA, Cox PM, Collins M, *et al.* (2004) The role of ecosystem-atmosphere interactions in simulated Amazonian precipitation decrease and forest dieback under global climate warming. *Theoretical and Applied Climatology* 78: 157–175.

Betts RA, Cox PM, Lee SE, and Woodward FI (1997) Contrasting physiological and structural vegetation feedbacks in climate change simulations. *Nature* 387: 796–799.

Betts RA, Falloon PD, Goldewijk KK, and Ramankutty N (2007a) Biogeophysical effects of land use on climate: Model simulations of radiative forcing and large-scale temperature change. *Agricultural and Forest Meteorology* 142: 216–233.

Bollasina M and Nigam S (2011) Modeling of regional hydroclimate change over the Indian subcontinent: Impact of the expanding Thar Desert. *Journal of Climate* 24: 3089–3106.

Bonan G (2002) *Ecological Climatology: Concepts and Applications*. Cambridge: Cambridge University Press.

Bond WJ, Midgley GF, and Woodward FI (2003) The importance of low CO_2 and fire in promoting the spread of grasslands and savannas. *Global Change Biology* 9: 973–982.

Bond WJ, Woodward FI, and Midgley GF (2005) The global distribution of ecosystems in a world without fire. *New Phytologist* 165: 525–538.

Bondeau A, Smith P, Zaehle S, *et al.* (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology* 13: 679–706.

Botkin DB, Janak JF, and Wallis JR (1972) Some ecological consequences of a computer model of forest growth. *Journal of Ecology* 60: 849–872.

Botta A, Viovy N, Ciais P, Friedlingstein P, and Monfray P (2000) A global prognostic scheme of leaf onset using satellite data. *Global Change Biology* 6: 709–725.

Boucher O, Jones A, and Betts RA (2009) Climate response to the physiological impact of carbon dioxide on plants in the Met Office unified Model HadCM3. *Climate Dynamics* 32: 237–249.

Bounoua L, Collatz GJ, Sellers PJ, *et al.* (1999) Interactions between vegetation and climate: Radiative and physiological effects of doubled atmospheric CO_2. *Journal of Climate* 12: 310–324.

Bounoua L, Hall FG, Sellers PJ, *et al.* (2010) Quantifying the negative feedback of vegetation to greenhouse warming: A modeling approach. *Geophysical Research Letters* 37. L23701.

Bowman DMJS, Balch JK, Artaxo P, *et al.* (2009) Fire in the Earth system. *Science* 324: 481–484.

Bradley AV, Gerard FF, Barbier N, *et al.* (2011) Relationship between phenology, radiation and precipitation in the Amazon region. *Global Change Biology* 17: 2245–2260.

Brooker RW, Maestre FT, Callaway RM, *et al.* (2008) Facilitation in plant communities: The past, the present, and the future. *Journal of Ecology* 96: 18–34.

Brovkin V, Raddatz T, Reick CH, Claussen M, and Gayler V (2009) Global biogeophysical interactions between forest and climate. *Geophysical Research Letters* 36. L07405.

Bullard JE, Harrison SP, Baddock M, *et al.* (2011) Preferential dust sources: A geomorphological classification designed for use in global dust-cycle models. *Journal of Geophysical Research* 116. F04034.

Cadule P, Friedlingstein P, Bopp L, *et al.* (2010) Benchmarking coupled climate-carbon models against long-term atmospheric CO_2 measurements. *Global Biogeochemical Cycles* 24. Doi: 10.1029/2009GB003556.

Chen HS, Dickinson RE, Dai YJ, and Zhou LM (2011) Sensitivity of simulated terrestrial carbon assimilation and canopy transpiration to different stomatal conductance and carbon assimilation schemes. *Climate Dynamics* 36: 1037–1054.

Christian JR, Arora VK, Boer GJ, *et al.* (2010) The global carbon cycle in the CCCma earth system model CanESM1: Preindustrial control simulation. *Journal of Geophysical Research – Biogeosciences* 115. G03014.

Chuine I (2000) A unified model for the budburst of trees. *Journal of Theoretical Biology* 207: 337–347.

Chuine I and Beaubien E (2001) Phenology is a major determinant of temperate tree range. *Ecology Letters* 4: 500–510.

Ciais P, Tagliabue A, Cuntz M, *et al.* (2011) Large inert carbon pool in terrestrial ecosystems during the last glacial maximum. *Nature Geoscience* 5: 74–79.

Clark DB, Mercado LM, Sitch S, *et al.* (2011) The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development* 4: 701–722.

Claussen M, Kubatzki C, Brovkin V, *et al.* (1999) Simulation of an abrupt change in Saharan vegetation in the mid-Holocene. *Geophysical Research Letters* 26: 2037–2040.

Coiffard C and Gomez B (2010) The rise to dominance of the angiosperm kingdom: Dispersal, habitat widening and evolution during the Late Cretaceous of Europe. *Lethaia* 43: 164–169.

Cowan LR and Farquhar GD (1977) Stomatal function in relation to leaf metabolism and environment. In: Jenning DH (ed.) *Integration of Activity in the Higher Plant*, pp. 471–505. Cambridge: Cambridge University Press.

Cowling SA (1999) Simulated effects of low atmospheric CO_2 on structure and composition of North American vegetation at the Last Glacial Maximum. *Global Ecology and Biogeography* 8: 81–93.

Cowling SA and Field CB (2003) Environmental controls of leaf area production: Implications for vegetation and land surface modelling. *Global Biogeochemical Cycles* 17: 1007.

Cowling SA, Jones CD, and Cox PM (2007b) Consequences of the evolution of C4 photosynthesis for surface energy and water exchange. *Journal of Geophysical Research – Biogeosciences* 112: G01020.

Cowling SA, Jones CD, and Cox PM (2009) Greening the terrestrial biosphere: Simulated feedbacks on atmospheric heat and energy circulation. *Climate Dynamics* 32: 287–299.

Cowling SA, Shin Y, Pinto E, and Jones CD (2007a) Water recycling by Amazonian vegetation: Coupled versus uncoupled vegetation–climate interactions. *Philosophical Transactions of the Royal Society B – Biological Sciences* 36: 1865–1871.

Cowling SA and Sykes MT (1999) Physiological significance of low atmospheric CO_2 for plant–climate interactions. *Quaternary Research* 52: 237–242.

Cox PM, Betts RA, Collins M, *et al.* (2004) Amazonian forest dieback under climate–carbon cycle projections for the 21st century. *Theoretical and Applied Climatology* 78: 137–156.

Cox PM, Betts RA, Jones CD, Spall SA, and Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* 408: 184–187.

Cramer W, Bondeau A, Woodward FI, *et al.* (2001) Global responses of terrestrial ecosystems to changes in CO_2 and climate. *Global Change Biology* 7: 357–373.

Dawson TP, Jackson ST, House JI, Prentice IC, and Mace GM (2011) Beyond predictions: Biodiversity conservation in a changing climate. *Science* 332: 53–58.

DeFries RS, Field CB, Fung I, Collatz GJ, and Bounoua L (1999) Combining satellite data and biogeochemical models to estimate global effects of human-induced land cover change on carbon emissions and primary productivity. *Global Biogeochemical Cycles* 13: 803–815.

Dekker SC, de Boer HJ, Brovkin V, *et al.* (2010) Biogeophysical feedbacks trigger shifts in the modelled vegetation–atmosphere system at multiple scales. *Biogeosciences* 7: 1237–1245.

Del Grosso SJ, Parton WJ, Mosier AR, *et al.* (2001) Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT model. In: Schaffer M, Ma L, and Hansen S (eds.) *Modeling Carbon and Nitrogen Dynamics for Soil Management*, pp. 303–332. Boca Raton, Florida: CRC Press.

Delire C and de Noblet-Ducoudré N (2011) Vegetation dynamics enhancing long-term climate variability confirmed by two models. *Journal of Climate* 24: 2238–2257.

Denman KL, Brasseur G, Chidthaisong A, *et al.* (2007) Couplings between changes in the climate system and biogeochemistry. In: Solomon S, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis*. Cambridge: Cambridge University Press.

Dewar RC (1996) The correlation between plant growth and intercepted radiation: An interpretation in terms of optimal plant nitrogen content. *Annals of Botany* 78: 125–136.

Díaz S and Cabido M (1997) Plant functional types and ecosystem function in relation to global change. *Journal of Vegetation Science* 8: 463–474.

Dieleman WIJ, Luyssaert S, Rey A, *et al.* (2010) Soil N modulates soil C cycling in CO_2-fumigated tree stands: A meta-analysis. *Plant Cell and Environment* 33: 2001–2011.

DiMichelle WA and Phillips TL (2002) The ecology of Paleozoic ferns. *Review of Palaeobotany and Palynology* 119: 143–159.

Domingues TF, Meir P, Feldpausch TR, *et al.* (2010) Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant Cell and Environment* 33: 959–980.

Donnadieu Y, Goddéris Y, and Bouttes N (2009) Exploring the climatic impact of the continental vegetation on the Mezosoic atmospheric CO_2 and climate history. *Climate of the Past* 5: 85–96.

Escapa IH, Taylor EL, Cuneo R, *et al.* (2011) Triassic floras of Antarctica: Plant diversity and distribution in high paleolatitude communities. *Palaios* 26: 522–544.

Esser G, Kattge J, and Sakalli A (2011) Feedback of carbon and nitrogen cycles enhances carbon sequestration in the terrestrial biosphere. *Global Change Biology* 17: 819–842.

Farquhar GD, Caemmerer SV, and Berry JA (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta* 149: 78–90.

Finzi AC, Norby RJ, Calfapietra C, *et al.* (2007) Increases in nitrogen uptake rather than nitrogen-use efficiency support higher rates of temperate forest productivity under elevated CO_2. *Proceedings of the National Academy of Sciences of the United States of America* 104: 14014–14019.

Fisher JB, Sitch S, Malhi Y, Fisher RA, Huntingford C, and Tan S-Y (2010) Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation. *Global Biogeochemical Cycles* 24. GB1014.

Foley JA, Levis S, Prentice IC, Pollard D, and Thompson SL (1998) Coupling dynamic models of climate and vegetation. *Global Change Biology* 4: 561–579.

Foley JA, Prentice IC, Ramankutty N, *et al.* (1996) An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles* 10: 603–628.

François L, Utescher T, Favre E, *et al.* (2011) Modelling late-Miocene vegetation in Europe: Results of the CARAIB model and comparison with palaeovegetation data. *Palaeogeography, Palaeoclimatology, Palaeoecology* 304: 359–378.

Franks PJ and Beerling DJ (2009) Maximum leaf conductance driven by atmospheric CO_2 effects on stomatal size and density over geologic time. *Proceedings of the National Academy of Science of the United States of America* 106: 10343–10347.

Friedlingstein P, Bopp L, Ciais P, *et al.* (2001) Positive feedback of the carbon cycle on future climate change. *Geophysical Research Letters* 28: 1543–1546.

Friedlingstein P, Cox P, Betts R, *et al.* (2006) Climate–carbon cycle feedback analysis: Results from the C^4MIP model intercomparison. *Journal of Climate* 19: 3337–3353.

Friedlingstein P, Fung I, Holland E, *et al.* (1995) On the contribution of CO_2 fertilization to the missing biospheric sink. *Global Biogeochemical Cycles* 9: 541–556.

Friend AD (1991) Use of a model of photosynthesis and leaf microenvironment to predict optimal stomatal conductance and leaf nitrogen partitioning. *Plant Cell and Environment* 14: 895–905.

Friend AD (2001) Modelling canopy CO_2 fluxes: Are 'big-leaf' simplifications justified? *Global Ecology and Biogeography* 10: 603–619.

Friend AD (2010) Terrestrial plant production and climate change. *Journal of Experimental Botany* 61: 1293–1309.

Friend AD and Kiang NY (2005) Land-surface model development for the GISS GCM: Effects of improved canopy physiology on simulated climate. *Journal of Climate* 18: 2883–2902.

Friend AD and White A (2000) Evaluation and analysis of a dynamic terrestrial ecosystem model under preindustrial conditions at the global scale. *Global Biogeochemical Cycles* 14: 1173–1190.

Fyllas NM, Quesada CA, and Lloyd J (2012) Deriving plant functional types for Amazonian forests for use in vegetation dynamics models. *Perspectives in Plant Ecology, Evolution and Systematics*.

Gedney N, Cox PM, Betts RA, *et al.* (2006) Detection of a direct carbon dioxide effect in continental river runoff records. *Nature* 439: 835–838.

Gedney N, Cox PM, and Huntingford C (2004) Climate feedback from wetland methane emissions. *Geophysical Research Letters* 31. L20503.

Gerber S, Hedin LO, Oppenheimer M, Pacala SW, and Shevliakova E (2010) Nitrogen cycling and feedbacks in a global dynamic land model. *Global Biogeochemical Cycles* 24. GB1001.

Gerten D, Schaphoff S, Haberlandt U, Lucht W, and Sitch S (2004) Terrestrial vegetation and water balance – hydrological evaluation of a dynamic global vegetation model. *Journal of Hydrology* 286: 249–270.

Gifford R (2003) Plant respiration in productivity models: Conceptualisation, representation and issues for global terrestrial carbon-cycle research. *Functional Plant Biology* 30: 171–186.

Giglio L, Randerson JT, van der Werf GR, *et al.* (2010) Assessing variability and long-term trends in burned area by merging multiple satellite fire products. *Biogeosciences* 7: 1171–1186.

Gonzalez P, Neilson RP, Lenihan JM, and Drapek RJ (2010) Global patterns in the vulnerability of ecosystems to vegetation shifts due to climate change. *Global Ecology and Biogeography* 19: 755–768.

Graciano C, Guiamet JJ, and Goya JF (2005) Impact of nitrogen and phosphorus fertilization on drought responses in *Eucalyptus grandis* seedlings. *Forest Ecology and Management* 212: 40–49.

Gu FX, Zhang YD, Tao B, *et al.* (2010) Modeling the effects of nitrogen deposition on carbon budget in two temperate forests. *Ecological Complexity* 7: 139–148.

Gutzler DS and Robbins TO (2011) Climate variability and projected change in the western United States: Regional downscaling and drought statistics. *Climate Dynamics* 37: 835–849.

Harpole WS, Ngai JT, Cleland EE, *et al.* (2011) Nutrient co-limitation of primary producer communities. *Ecology Letters* 14: 852–862.

Harris PP, Huntingford C, and Cox PM (2008) Amazon Basin climate under global warming: The role of the sea surface temperature. *Philosophical Transactions of the Royal Society B – Biological Sciences* 363: 1753–1759.

Harrison SP and Bartlein PJ (2012) Records from the past, lessons for the future: What the palaeorecord implies about mechanisms of global change. In: Henderson-Sellers A and McGuffie K (eds.) *The Future of the World's Climate*, pp. 403–436. Amsterdam: Elsevier.

Harrison SP, Kohfeld KE, Roelandt C, and Claquin T (2001) The role of dust in climate changes today, at the last glacial maximum and in the future. *Earth Science Reviews* 54: 43–80.

Harrison SP, Marlon JL, and Bartlein PJ (2010b) Fire in the Earth system. In: Dodson J (ed.) *Changing Climates, Earth Systems and Society*, pp. 21–48. Dordrecht: Springer.

Harrison SP and Prentice IC (2003) Climate and CO_2 controls on global vegetation distribution at the last glacial maximum: Analysis based on palaeovegetation data, biome modeling and palaeoclimate simulations. *Global Change Biology* 9: 983–1004.

Harrison SP, Prentice IC, Barboni D, *et al.* (2010a) Ecophysiological and bioclimatic foundations for a global plant functional classification. *Journal of Vegetation Science* 21: 300–317.

Harrison SP and Sanchez-Goñi MF (2010) Global patterns of vegetation response to millennial-scale variability and rapid climate change during the last glacial period. *Quaternary Science Reviews* 29: 2957–2980.

Haxeltine A and Prentice IC (1996a) BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability and competition among plant functional types. *Global Biogeochemical Cycles* 10: 693–709.

Haxeltine A and Prentice IC (1996b) A general model for the light use efficiency of primary production. *Functional Ecology* 10: 551–561.

Haywood AM and Valdes PJ (2006) Vegetation cover in a warmer world simulated using a dynamic global vegetation model for the mid-Pliocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* 237: 412–427.

Haywood AM, Valdes PJ, Francis JE, and Sellwood BW (2002) Global middle Pliocene biome reconstruction: A data-model synthesis. *Geochemistry, Geophysics, Geosystems* 3: 1072.

Heimann M, Esser G, Haxeltine A, *et al.* (1998) Evaluation of terrestrial carbon cycle models through simulations of the seasonal cycle of atmospheric CO_2: First results of a model intercomparison study. *Global Biogeochemical Cycles* 12: 1–24.

Hendry GR and Kimball BA (1994) The FACE program. *Agricultural and Forest Meteorology* 70: 3–14.

Hendry GR, Lewin KF, and Nagy J (1993) Free air carbon dioxide enrichment: Development, progress, results. *Vegetation* 104: 17–31.

Hermann E, Hochuli PA, Bucher H, *et al.* (2011) Terrestrial ecosystems on North Gondwana following the end-Permian mass extinction. *Gondwana Research* 20: 630–637.

Herold N, Huber M, Greenwood DR, Muller RD, and Seton M (2011) Early to middle Miocene monsoon climate in Australia. *Geology* 39: 3–6.

Hickler T, Smith B, Prentice IC, *et al.* (2008) CO_2 fertilization in temperate FACE experiments not representative of boreal and tropical forests. *Global Change Biology* 14: 1531–1542.

Hickler T, Vohland K, Feehan J, *et al.* (2012) Projecting tree species-based climate-driven changes in European potential natural vegetation with a generalized dynamic vegetation model. *Global Ecology and Biogeography* 21: 50–63.

Hofmockel KS, Gallet-Budynek A, McCarthy HR, *et al.* (2011) Sources of increased N uptake in forest trees growing under elevated CO_2: Results of a large-scale ^{15}N study. *Global Change Biology* 17: 3338–3350.

Hopcroft PO, Valdes PJ, and Beerling DJ (2011) Simulating idealized Dansgaard-Oeschger events and their potential impacts on the global methane cycle. *Quaternary Science Reviews* 30: 3258–3268.

Horton DE, Poulsen CJ, and Pollard D (2010) Influence of high-latitude vegetation feedbacks on late Palaeozoic glacial cycles. *Nature Geoscience* 3: 572–577.

Hungate BA, Dukes JS, Shaw MR, Luo Y, and Field CB (2003) Nitrogen and climate change. *Science* 302: 1512–1513.

Huntingford C, Cox PM, Mercado LM, *et al.* (2011) Highly contrasting effects of different climate forcing agents on terrestrial ecosystem services. *Philosophical*

Transactions of the Royal Society A – Mathematical Physical and Engineering Sciences 369: 2026–2037.
Huntingford C, Fisher RA, Mercado LM, *et al.* (2008) Towards quantifying uncertainty in predictions of Amazon 'dieback'. *Philosophical Transactions of the Royal Society B – Biological Sciences* 363: 1857–1864.
Ivanov VY, Bras RL, and Vivoni ER (2008) Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks. *Water Resources Research* 44. W03429.
Jain A, Yang X, Kheshgi H, *et al.* (2009) Nitrogen attenuation of terrestrial carbon cycle response to global environmental factors. *Global Biogeochemical Cycles* 23. GB4028.
Johnson KD, Scatena FN, and Silver WL (2011) Atypical soil carbon distribution across a tropical steepland forest catena. *Catena* 87: 391–397.
Jost A, Fauquette S, Kageyama M, *et al.* (2009) High resolution climate and vegetation simulations of the late-Pliocene, a model-data comparison over western Europe and the Mediterranean region. *Climate of the Past* 5: 585–606.
Kattge J, Díaz S, Lavorel S, *et al.* (2011) TRY: A global database of plant traits. *Global Change Biology* 17: 2905–2935.
Katul GG and Siqueira MB (2010) Biotic and abiotic factors act in coordination to amplify hydraulic redistribution and lift. *New Phytologist* 187: 3–7.
Keeling CD (1958) The concentration and isotopic abundances of atmospheric carbon dioxide in rural areas. *Geochimica et Cosmochimica Acta* 13: 322–334.
Keeling RF, Piper SC, and Heimann M (1996) Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration. *Nature* 381: 218–221.
Keeling RF and Shertz SR (1992) Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle. *Nature* 358: 723–727.
Keenan T, Serra JM, Lloret F, Ninyerola M, and Sabate S (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO_2 matters! *Global Change Biology* 17: 565–579.
Kicklighter DW, Bruno M, Dönges S, *et al.* (1999) A first order analysis of the potential of CO_2 fertilization to affect the global carbon budget: A comparison of four terrestrial biosphere models. *Tellus* 51B: 343–366.
Kimball BA, Pinter PJ, Wall GW, *et al.* (1995) Carbon dioxide effects on crop energy balance: Testing ecosystems with a free-air CO_2 enrichment (FACE) experiment. *Agronomy Journal* 97: 446–457.
Kleidon A, Adams J, Pavlick R, and Reu B (2009) Simulated geographic variations of plant species richness, evenness and abundance using climatic constraints on plant functional diversity. *Environmental Research Letters* 4: 014007.
Kleidon A and Mooney HA (2000) A global distribution of biodiversity inferred from climatic constraints: Results from a process-based modelling study. *Global Change Biology* 6: 507–523.
Knapp AK, Briggs JM, Collins SL, *et al.* (2008) Shrub encroachment in North American grasslands: Shifts in growth form dominance rapidly alters control of ecosystem carbon inputs. *Global Change Biology* 14: 615–623.
Köppen W (1918) Klassifikation der Klimate nach Temperatur, Niederschlag und Jahresablauf (Classification of climates according to temperature, precipitation and seasonal cycle). *Petermanns Geographische Mitteilungen* 64: 193–203 and 243–248.
Krawchuk M, Moritz M, Parisien M-A, Van Dorn J, and Hayhoe K (2009) Global pyrogeography: The current and future distribution of wildfire. *PLoS ONE* 4: e5102.
Krinner G, Viovy N, de Noblet-Ducoudré N, *et al.* (2005) A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system. *Global Biogeochemical Cycles* 19. GB1015.
Kucharik CJ, Foley JA, Delire C, *et al.* (2000) Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance and vegetation structure. *Global Biogeochemical Cycles* 14: 795–825.
Lavorel S, Grigulis K, Lamarque P, *et al.* (2011) Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology* 99: 135–147.
Lee TD, Barrott SH, and Reich PB (2011) Photosynthetic responses of 13 grassland species across 11 years of free-air CO_2 enrichment is modest, consistent and independent of N supply. *Global Change Biology* 17: 2893–2904.
Lenihan JM, Daly C, Bachelet D, and Neilson RP (1998) Simulating broad-scale fire severity in a dynamic global vegetation model. *Northwest Science* 72: 91–103.
Leslie AB (2011) Shifting functional roles and the evolution of conifer pollen-producing and seed-producing cones. *Paleobiology* 37(4): 587–602.
Levis S, Bonan GB, Vertenstein M, and Oleson KW (2004) The community land model's dynamic global vegetation model (CLM-DGVM) technical description and user guide. *NCAR Technical Note*, May 2004.
Li Q and Ishidaira H (2011) Development of a biosphere hydrological model considering vegetation dynamics and its evaluation at basin scale under climate change. *Journal of Hydrology* 412–413: 3–13.
Liang X, Lettenmaier DP, Wood E, and Burges SJ (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research* 99: 14415–14428.
Liang X and Xie Z (2001) A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. *Advances in Water Resources* 24: 1173–1193.
Liu Z, Wang Y, Gallimore R, Notaro M, and Prentice IC (2006) On the cause of abrupt vegetation collapse in North Africa during the Holocene: Climate variability vs. vegetation feedback. *Geophysical Research Letters* 33. L22709.
Liu Z, Wang Y, Gallimore R, *et al.* (2007) Simulating the transient evolution and abrupt change of northern Africa atmosphere–ocean–terrestrial ecosystem in the Holocene. *Quaternary Science Reviews* 26: 1818–1837.
Lloyd J, Grace J, Miranda AC, *et al.* (1995) A simple calibrated model of Amazon rainforest productivity based on leaf biochemical properties. *Plant Cell and Environment* 18: 1129–1145.
Lloyd J, Patiño S, Paiva RQ, *et al.* (2010) Optimisation of photosynthetic carbon gain and within-canopy gradients of associated foliar traits for Amazon forest trees. *Biogeosciences* 7: 1833–1859.
Lomax B, Beerling DJ, Upchurch Jr. G, and Otto-Bliesner B (2001) Rapid (10-yr) recovery of terrestrial productivity in a simulation study of the terminal Cretaceous impact event. *Earth and Planetary Science Letters* 192: 137–144.
Lucht W, Prentice IC, Myneni RB, *et al.* (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science* 296: 1687–1689.
Lunt DJ, Haywood AM, Schmidt GA, *et al.* (2010) Earth system sensitivity inferred from Pliocene modelling and data. *Nature Geoscience* 3: 60–65.
Lunt DJ, Ross I, Hopley PJ, and Valdes PJ (2007) Modelling late-Oligocene C4 grasses and climate. *Palaeogeography, Palaeoclimatology, Palaeoecology* 251: 239–253.
MacDonald GM (2010) Water, climate change, and sustainability in the southwest. *Proceedings of the National Academy of Sciences of the United States of America* 107: 21256–21262.
Magallon S and Castillo A (2009) Angiosperm diversification through time. *American Journal of Botany* 96: 349–365.
Malhi Y, Doughty C, and Galbraith D (2011) The allocation of ecosystem net primary productivity in tropical forests. *Philosophical Transactions of the Royal Society B – Biological Sciences* 366: 3225–3245.
Markewitz D, Devine S, Davidson EA, Brando P, and Nepstad DC (2010) Soil moisture depletion under simulated drought in the Amazon: Impacts on deep root uptake. *New Phytologist* 187: 592–607.
Marlon JR, Bartlein PJ, Carcaillet C, *et al.* (2008) Climate and human influences on biomass burning over the past two millennia. *Nature Geoscience* 1: 697–702.
McGuire AD, Sitch S, Clein JS, *et al.* (2001) Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO_2, climate and land-use effects with four process-based ecosystem models. *Global Biogeochemical Cycles* 15: 183–206.
Medlyn BE, Badeck F-W, de Pury DGG, *et al.* (1999) Effects of elevated [CO_2] on photosynthesis in European forest species: A meta-analysis of model parameters. *Plant, Cell and Environment* 22: 1475–1496.
Medlyn BE, Duursma RA, Eamus D, *et al.* (2011) Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology* 17: 2134–2144.
Medvigy D, Wofsy SC, Munger JW, Hollinger DY, and Moorcroft PR (2009) Mechanistic scaling of ecosystem function and dynamics in space and time: The Ecosystem Demography model version 2. *Journal of Geophysical Research – Biogeosciences* 114. G01002.
Melillo JM, McGuire AD, Kicklighter DW, *et al.* (1993) Global change and terrestrial net primary production. *Nature* 363: 234–240.
Mercado LM, Bellouin N, Sitch S, *et al.* (2009) Impact of changes in diffuse radiation on the global land carbon sink. *Nature* 458: 1014–1017.
Mercado LM, Huntingford C, Gash JHC, Cox PM, and Jogireddy V (2007) Improving the representation of radiation interception and photosynthesis for climate model applications. *Tellus Series B – Chemical and Physical Meteorology* 59: 553–565.
Meron E (2011) Modeling dryland landscapes. *Mathematical Modelling of Natural Phenomena* 6: 163–187.
Metcalfe DB, Fisher RA, and Wardle DA (2011) Plant communities as drivers of soil respiration: Pathways, mechanisms, and significance for global change. *Biogeosciences* 8: 2047–2061.

Morford SL, Houlton BZ, and Dahlgren RA (2011) Increased forest ecosystem carbon and nitrogen storage from nitrogen rich bedrock. *Nature* 477: 78–81.

Morin X, Augspurger C, and Chuine I (2007) Process-based modeling of tree species' distributions: What limits temperate tree species' range boundaries? *Ecology* 88: 2280–2291.

Murray MB, Cannell MGR, and Smith RI (1989) Date of budburst of fifteen tree species in Britain following climatic warming. *Journal of Applied Ecology* 26: 693–700.

Murray SJ, Foster PN, and Prentice IC (2011) Evaluation of global continental hydrology as simulated by the Land-surface Processes and exchanges Dynamic Global Vegetation Model. *Hydrology and Earth System Science* 15: 91–105.

Murray SJ, Watson M, and Prentice IC (2012) The use of Dynamic Global Vegetation Models for simulating hydrology and the potential integration of satellite observations. *Progress in Physical Geography.*

Myneni RC, Keeling CD, Tucker CJ, Asrar G, and Nemani RR (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* 386: 698–702.

Naito AT and Cairns DM (2011) Patterns and processes of global shrub expansion. *Progress in Physical Geography* 35: 423–442.

Neilson RP, Prentice IC, and Smith B (1998) Simulated changes in vegetation distribution under global warming. In: Watson RT, Zinyowera MC, Moss RH, and Dokken DJ (eds.) *The Regional Impacts of Climate Change: An Assessment of Vulnerability*, pp. 439–456. Cambridge: Cambridge University Press.

de Noblet-Ducoudré N, Claussen M, and Prentice IC (2000) Mid-Holocene greening of the Sahara: First results of the GAIM 6000 yr BP experiment with two asynchronously coupled atmosphere/biome models. *Climate Dynamics* 16: 643–659.

Norby RJ, DeLucia EH, Gielen B, *et al.* (2005) Forest response to elevated CO_2 is conserved across a broad range of productivity. *Proceedings of the National Academy of Science of the United States of America* 102: 18052–18056.

Norby RJ, Warren JM, Iversen CM, Medlyn BE, and McMurtrie RE (2010) CO_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences of the United States of America* 107: 19368–19373.

Norby RJ and Zak DR (2011) Ecological lessons from free-air CO_2 enrichment (FACE) experiments. *Annual Review of Ecology, Evolution and Systematics* 42: 181–203.

Nowak RS, Ellsworth DS, and Smith SD (2004) Functional responses of plants to elevated atmospheric CO_2 – do photosynthetic and productivity data from FACE experiments support early predictions? *New Phytologist* 162: 253–280.

O'ishi R, Abe-Ouchi A, Sitch S, and Prentice IC (2009) Vegetation dynamics and plant CO_2 responses as positive feedbacks in a greenhouse world. *Geophysical Research Letters* 36. L11706.

Oren R, Ellsworth DS, and Johnsen KH (2001) Soil fertility limits carbon sequestration by forest ecosystems in a CO_2-enriched atmosphere. *Nature* 411: 469–472.

Ostle NJ, Smith P, Fisher R, *et al.* (2009) Integrating plant–soil interactions into global carbon cycle models. *Journal of Ecology* 97: 851–863.

Parton WJ, Ojima DS, Schimel DS, and Kittel TGF (1993) Modelling the effects of climatic and CO_2 changes on grassland storage of soil C. *Water, Air and Soil Pollution* 70: 643–657.

Pearson RG and Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography* 12: 361–371.

Piao S, Friedlingstein P, Ciais P, *et al.* (2007) Changes in climate and land use have a larger direct impact than rising CO_2 on global river runoff trends. *Proceedings of the National Academy of the United States of America* 104: 15242–15247.

Pitman AJ, de Noblet-Ducoudré N, Cruz FT, *et al.* (2009) Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study. *Geophysical Research Letters* 36. L14814.

Poorter H, Niklas KJ, Reich PB, Oleksyn J, Poot P, and Mommer L (2012) Biomass allocation to leaves, stems and roots: Meta-analyses of interspecific variation and environmental control. *New Phytologist* 193: 30–50.

Popp A, Vogel M, Blaum N, and Jeltsch F (2009) Scaling up ecohydrological processes: Role of surface water flow in water-limited landscapes. *Journal of Geophysical Research – Biogeosciences* 114.

Potter CS, Randerson JT, Field CB, *et al.* (1993) Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles* 7: 811–841.

Pound MJ, Haywood AM, Salzmann U, *et al.* (2011) A Tortonian (late-Miocene, 11.61-7.25 Ma) global vegetation reconstruction. *Palaeogeography, Palaeoclimatology, Palaeoecology* 300: 29–45.

Prentice IC, Bondeau A, Cramer W, *et al.* (2007) Dynamic global vegetation modelling: Quantifying terrestrial ecosystem responses to large-scale environmental change. In: Canadell J, Pitelka L, and Pataki D (eds.) *Terrestrial Ecosystems in a Changing World*, pp. 175–192. Berlin: Springer.

Prentice IC, Cramer W, Harrison SP, *et al.* (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19: 117–134.

Prentice IC, Farquhar GD, Fasham MJR, *et al.* (2001) The carbon cycle and atmospheric carbon dioxide. In: Houghton JT, Ding Y, Griggs DJ, *et al.* (eds.) *Climate Change 2001: The Scientific Basis*. Cambridge: Cambridge University Press.

Prentice IC and Harrison SP (2009) Ecosystem effects of CO_2 concentration: Evidence from past climates. *Climate of the Past* 5: 297–307.

Prentice IC, Harrison SP, and Bartlein PJ (2011c) Global vegetation and terrestrial carbon cycle changes after the last ice age. *New Phytologist* 189: 988–998.

Prentice IC, Heimann M, and Sitch S (2000a) The carbon balance of the terrestrial biosphere: Ecosystem models and atmospheric observations. *Ecological Applications* 10: 1553–1573.

Prentice IC and Jolly D and BIOME 6000 participants (2000b) Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *Journal of Biogeography* 27: 507–519.

Prentice IC, Kelley DI, Harrison SP, *et al.* (2011a) Modeling fire and the terrestrial carbon balance. *Global Biogeochemical Cycles* GB3005.

Prentice IC, Meng TT, Wang H, *et al.* (2011b) Evidence of a universal scaling relationship for leaf CO_2 drawdown along an aridity gradient. *New Phytologist* 190: 169–180.

Prentice IC and Werger MJA (1985) Clump spacing in a desert dwarf shrub community. *Vegetatio* 63: 133–139.

de Pury DGG and Farquhar GD (1999) A commentary on the use of a sun/shade model to scale from the leaf to a canopy. *Agricultural and Forest Meteorology* 95: 257–260.

Quillet A, Peng CH, and Garneau M (2010) Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: Recent developments, limitations and future challenges. *Environmental Reviews* 18: 333–353.

Raich JW, Rastetter EB, Melillo JM, *et al.* (1991) Potential net primary productivity in South America: Application of a global model. *Ecological Applications* 1: 399–429.

Rammig A, Jupp T, Thonicke K, *et al.* (2010) Estimating the risk of Amazonian forest dieback. *New Phytologist* 187: 694–706.

Rastetter EB (2011) Modeling coupled biogeochemical cycles. *Frontiers in Ecology and the Environment* 9: 68–73.

Rastetter EB, Vitousek PM, Field C, *et al.* (2001) Resource optimization and symbiotic nitrogen fixation. *Ecosystems* 4: 369–388.

Ren DD and Henderson-Sellers A (2006) An analytical hydrological model for the study of scaling issues in land surface modeling. *Earth Interactions* 10: 20.

Reu B, Zaehle S, Proulx R, *et al.* (2011) The role of plant functional trade-offs for biodiversity changes and biome shifts under scenarios of global climatic change. *Biogeosciences* 8: 1255–1266.

Reyers M, Kruger A, Werner C, *et al.* (2011) The simulation of the opposing fluxes of latent heat and CO_2 over various land-use types: Coupling a gas exchange model to a mesoscale atmospheric model. *Boundary-Layer Meteorology* 139: 121–141.

Rietkerk M, Brovkin V, van Bodegom PM, *et al.* (2011) Local ecosystem feedbacks and critical transitions in the climate. *Ecological Complexity* 8: 223–228.

Ringeval B, de Noblet-Ducoudré N, Ciais P, *et al.* (2010) An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales. *Global Biogeochemical Cycles* 24. GB2003.

Rost S, Gerten D, Bondeau A, *et al.* (2008) Agricultural green and blue water consumption and its influence on the global water system. *Water Resources Research* 44. W09405.

Rothermel RC (1972) A mathematical model for predicting fire spread in wildland fuels. *United States Department of Agriculture Forest Service General Technical Report INT-115*, Intermountain Forest and Range Experiment Station, Ogden, Utah.

Roy SB (2009) Mesoscale vegetation–atmosphere feedbacks in Amazonia. *Journal of Geophysical Research – Atmospheres* 114: D20111.

Running SW and Coughlan JC (1988) A general model of forest ecosystem processes for regional applications. I. Hydrologic balance, canopy gas exchange and primary production processes. *Ecological Modelling* 42: 125–154.

Running SW and Hunt RE (1993) Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global-scale models. In: Ehleringer JR and Field CB (eds.) *Scaling Physiologic Processes: Leaf to Globe*, pp. 141–158. San Diego: Academic Press.

Salazar LF and Nobre CA (2010) Climate change and thresholds of biome shifts in Amazonia. *Geophysical Research Letters* 37. L17706.
Salzmann U, Haywood AM, Lunt DJ, Valdes PJ, and Hill DJ (2008) A new global biome reconstruction and data-model comparison for the middle-Pliocene. *Global Ecology and Biogeography* 17: 432–447.
Scholze M, Allen I, Collins W, *et al.* (2012) Earth System models: A tool to understand changes in the Earth System. In: Cornell SE, Prentice IC, House JI, and Downy CJ (eds.) *Understanding the Earth System: Global Change Science for Applications*. Cambridge: Cambridge University Press.
Scholze M, Knorr W, Arnell NW, and Prentice IC (2006) A climate change risk analysis for world ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 103: 13116–13120.
Schymanski SJ, Sivapalan M, Roderick ML, Hutley LB, and Beringer J (2009) An optimality-based model of the dynamic feedbacks between natural vegetation and the water balance. *Water Resources Research* 45. W01412.
Seller PJ, Bounoua L, Collatz GJ, *et al.* (1996) Comparison of radiative and physiological effects of doubled atmospheric CO_2 on climate. *Science* 271: 1402–1406.
Sellers PJ, Berry JA, Collatz GJ, Field CB, and Hall FG (1992) Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of the Environment* 42: 187–216.
Sellers PJ, Dickinson RE, Randall DA, *et al.* (1997) Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* 275: 502–509.
Seneviratne SI, Koster RD, Guo Z, *et al.* (2006) Soil moisture memory in AGCM simulations: Analysis of global land–atmosphere coupling experiment (GLACE) data. *Journal of Hydrometeorology* 7: 1090–1112.
Shellito CJ and Sloan LC (2006a) Reconstructing a lost Eocene paradise: Part I. Simulating the change in global floral distribution at the initial Eocene thermal maximum. *Global and Planetary Change* 50: 1–17.
Shellito CJ and Sloan LC (2006b) Reconstructing a lost Eocene paradise: Part II. On the utility of dynamic global vegetation models in pre-Quaternary climate studies. *Global and Planetary Change* 50: 18–32.
Shugart HH (1984) *A Theory of Forest Dynamics: The Ecological Implications of Forest Succession Models*. New York: Springer.
Sitch S, Huntingford C, Gedney N, *et al.* (2008) Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using 5 Dynamic Global Vegetation Models (DGVMs). *Global Change Biology* 14: 1–25.
Sitch S, Smith B, Prentice IC, *et al.* (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology* 9: 161–185.
Smith B, Prentice IC, and Sykes MT (2001) Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Global Ecology and Biogeography* 10: 621–637.
Sokolov AP, Kicklighter DW, Melillo JM, *et al.* (2008) Consequences of considering carbon–nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle. *Journal of Climate* 21: 3776–3796.
Spahni R, Wania R, Neef L, *et al.* (2011) Constraining global methane emissions and uptake by ecosystems. *Biogeosciences* 8: 1643–1665.
Sykes MT, Prentice IC, and Cramer W (1996) A bioclimatic model for the potential distributions of north European tree species under present and future climates. *Journal of Biogeography* 23: 203–233.
Tang GP and Bartlein PJ (2008) Simulating the climatic effects on vegetation: Approaches, issues and challenges. *Progress in Physical Geography* 32: 543–556.
Tegen I, Harrison SP, Kohfeld K, *et al.* (2002) The impact of vegetation and preferential source areas on global dust aerosol: Results from a model study. *Journal of Geophysical Research* 107: D21.
Thomas CD, Cameron A, Green RE, *et al.* (2004) Extinction risk from climate change. *Nature* 427: 145–148.
Thonicke K, Spessa A, Prentice IC, *et al.* (2010) The influence of vegetation, fire spread and fire behaviour on global biomass burning and trace gas emissions: Results from a process-based model. *Biogeosciences* 7: 1999–2011.
Thornton PE, Lamarque J-F, Rosenbloom NA, and Mahowald NM (2007) Influence of carbon–nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. *Global Biogeochemical Cycles* 21. GB4018.
Thuiller W, Lavorel S, Araújo MB, Sykes MT, and Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America* 102: 8245–8250.
Tietjen B, Jeltsch F, Zehe R, *et al.* (2010) Effects of climate change on the coupled dynamics of water and vegetation in drylands. *Ecohydrology* 3: 226–237.
Timm O and Timmermann A (2007) Simulation of the last 21,000 years using accelerated transient boundary conditions. *Journal of Climate* 20: 4377–4401.
Vamborg FSE, Brovkin V, and Claussen M (2011) The effect of a dynamic background albedo scheme on Sahel/Sahara precipitation during the mid-Holocene. *Climate of the Past* 7: 117–131.
Vasquez-Mendez R, Ventura-Ramos E, Oleschko K, *et al.* (2010) Soil erosion and runoff in different vegetation patches from semiarid Central Mexico. *Catena* 80: 162–169.
Wang G (2011) Assessing the potential hydrological impacts of hydraulic redistribution in Amazonia using a numerical modeling approach. *Water Resources Research* 47. W02528.
Wang G, Alo C, Mei R, and Sun S (2011) Droughts, hydraulic redistribution, and their impact on vegetation composition in the Amazon forest. *Plant Ecology* 212: 663–673.
Wang J, Chappellaz J, Park K, and Mak JE (2010a) Large variations in southern hemisphere biomass burning during the last 650 years. *Science* 330: 1663–1666.
Wang YP, Law RM, and Pak B (2010b) A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. *Biogeosciences* 7: 2261–2282.
Wania R, Ross I, and Prentice IC (2009a) Integrating peatlands and permafrost into a dynamic global vegetation model: 1. Evaluation and sensitivity of physical land surface processes. *Global Biogeochemical Cycles* 23. GB3015.
Wania R, Ross I, and Prentice IC (2009b) Integrating peatlands and permafrost into a dynamic global vegetation model: 2. Evaluation and sensitivity of vegetation and carbon cycle processes. *Global Biogeochemical Cycles* 23. GB3014.
Wania R, Ross I, and Prentice IC (2010) Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3. *Geoscientific Model Development* 3: 1–59.
van der Werf GR, Randerson JT, Giglio L, *et al.* (2010) Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics* 10: 11707–11735.
Wolff EW, Harrison SP, Knutti R, *et al.* (2012) How has climate responded to natural perturbations? In: Cornell SE, Prentice IC, House JI, and Downy CJ (eds.) *Understanding the Earth System: Global Change Science for Applications*. Cambridge: Cambridge University Press.
Woodward FI (1987) *Climate and Plant Distribution*. Cambridge: Cambridge University Press.
Woodward FI and Cramer W (1996) Plant functional types and climatic change: Introduction. *Journal of Vegetation Science* 7: 306–308.
Woodward FI and Lomas MR (2004) Vegetation dynamics – simulating responses to climatic change. *Biological Reviews* 79: 643–670.
Woodward FI, Lomas MR, and Betts RA (1998) Vegetation–climate feedbacks in a greenhouse world. *Philosophical Transactions of the Royal Society of London B* 353: 29–39.
Wright IJ, Reich PB, Cornelissen JHC, *et al.* (2005) Assessing the generality of global leaf trait relationships. *New Phytologist* 166: 485–496.
Wright IJ, Reich PB, and Westoby M (2003) Least-cost input mixtures of water and nitrogen for photosynthesis. *American Naturalist* 161: 98–111.
Wright IJ, Reich PB, Westoby M, *et al.* (2004) The worldwide leaf economics spectrum. *Nature* 428: 821–827.
Wright IJ and Westoby M (2002) Leaves at low versus high rainfall: Coordination of structure, lifespan and physiology. *New Phytologist* 155: 403–416.
Xu-Ri and Prentice IC (2008) Terrestrial nitrogen cycle simulation with a dynamic global vegetation model. *Global Change Biology* 14: 1745–1764.
Zaehle S and Friend A (2010) Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biochemical Cycles* 24. GB1005.
Zaehle S, Friend AD, Friedlingstein P, *et al.* (2010) Carbon and nitrogen cycle dynamics in the O-CN land surface model: 2. The role of the nitrogen cycle in the historical terrestrial C balance. *Global Biogeochemical Cycles* 24. GB1006.
Zaragoza-Castells J, Valladares F, Sánchez-Gómez D, *et al.* (2008) Climate-dependent variations in leaf respiration in a dry-land, low productivity Mediterranean forest: The importance of acclimation in both high-light and shaded habitats. *Functional Ecology* 22: 172–184.
Zeng XM, Zhao M, Yu RC, *et al.* (2003) Application of a "Big-Tree" model to regional climate modeling: A sensitivity study. *Theoretical and Applied Climatology* 76: 203–218.
Zhu QA, Jiang H, Peng CH, *et al.* (2011) Evaluating the effects of future climate change and elevated CO_2 on the water use efficiency in terrestrial ecosystems of China. *Ecological Modelling* 222: 2414–2429.

E

Ecogenomic Sensors

Christopher A Scholin, Monterey Bay Aquarium Research Institute, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Autonomous sensor Device that allows for automatic detection of chemical, physical, or biological properties absent human intervention.

Cabled observatory Cable on the seafloor that provides electrical power and high-speed communications between sensors in the ocean and the Internet (*see also* Ocean observatory).

DNA probe array Collection of short, single-stranded, uniquely addressable segments of deoxyribonucleic acid (DNA) attached to a solid surface that are used to detect the presence or expression of specific genes via the principles of nucleic acid hybridization.

Functional gene Recognizable, heritable segment of DNA that encodes for proteins involved with cell metabolism.

Genomics Discipline of genetics that focuses on the study of the complete genetic content of species, often by determining the entire DNA sequence of organisms and expression and function of their genes in relation to various stimuli or metabolic perturbations.

In situ Latin phrase meaning "in position"; here, its use is synonymous with "in water."

Ocean observatory Collection of airborne and in-water sensors for characterizing oceanic chemical, physical, and biological parameters that may vary over time. Sensors comprising an ocean observatory are mounted on various platforms (e.g., satellite, aircraft, moorings, autonomous underwater vehicles, benthic cable) that typically provide communications capabilities for transmitting sensor outputs and for receiving instructions from a remote location. An assortment of sensors and platforms spread over a distance may also be referred to as a *distributed ocean sensor array.*

Protein probe array As per DNA probe array, only with proteins that serve as sensing elements.

Real-time Instantaneously; without latency. In the ocean sciences, this phrase is often used with respect to the amount of time between sensor function and data transmission.

Remote operations Controlling and communicating with autonomous systems from afar.

History of the Ecogenomic Sensor Concept

Ocean scientists had conceived of ecogenomic sensors many years before the phrase was coined and put into common practice (Scholin, 2010). The origins of that name are closely tied to an image put forth in 2005 by Hunter Hadaway, a graphic artist (Figure 1). Prior to that time, the term *ecogenomics* was being used to describe the merger of high-throughput DNA sequencing with environmental science interests. Among the many attempts to capture this idea graphically was a figure promoted by Anna Palmisano, who once combined images of a satellite, a ship, and a DNA microarray to illustrate how genomic approaches for studying marine microbes could be nested within a variety of other ocean-observing capabilities. Virginia Armburst encouraged Hadaway to meld Palmisano's ideas with the notion of an ocean bottom cabled observatory and satellite imagery. From those inputs emerged Hadaway's iconic image carrying with it the term *ecogenomic sensor.* The concept of an autonomous ocean sensor network had been successfully fused with the burgeoning field of ecogenomics. Those who had been working independently for years at that very interface could at once point to a common image that captured the essence of their collective vision: a field-deployable device that enables the remote detection of targeted organisms, genes, and gene products using molecular probe technologies.

Implicit in Figure 1 was that a larger ocean observing system provided an environmental context through which the outputs of ecogenomic sensors would be viewed. In keeping with the vanguard cabled-observatory concept proposed at that time, nearly unlimited power and high-bandwidth communications networks would allow ecogenomic sensors to transmit results of analyses instantly and also allow researchers to redirect activity of those devices based on inputs conveyed via the larger, distributed sensor array. Ecogenomic sensors would thus become an extension of ship- and shore-based laboratories, opening a new window for observing biological processes in the sea. It was a grand vision.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00408-1

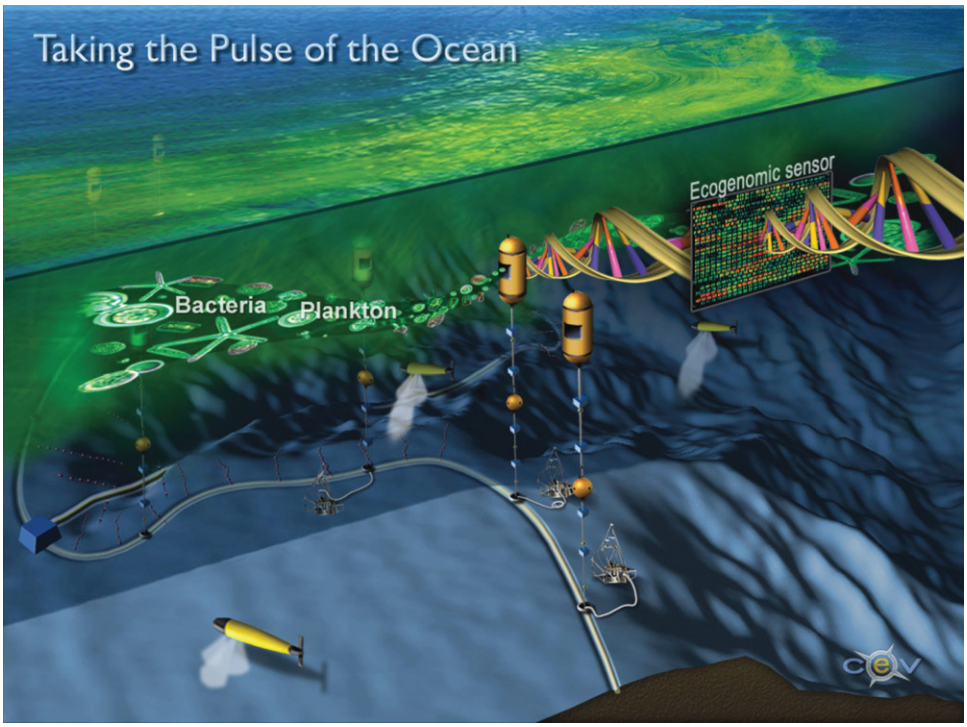


Figure 1 Artist's rendition of ecogenomic sensors embedded within an array of airborne and in-water sensors. This image was cast in the broader vision of the US Ocean Observatories Initiative (OOI).

Pieces of that puzzle were in fact in various stages of development in 2005, with some parts operational and others experimental, but the complete system as depicted did not yet exist.

The initial motivation for developing ecogenomic sensors took root many years prior to the release of Hadaway's figure and emerged from much more humble beginnings. The widespread use of molecular biological methods as a means to reveal the diversity, relationships, and functions of marine organisms had been a fixture in the ocean sciences for decades. Traditionally, such investigations often required transporting discrete field samples to a laboratory for subsequent processing and data generation. That requirement stood in sharp contrast to characterizations of many basic chemical, physical, and optical properties of waters (e.g., temperature, salinity, chlorophyll fluorescence) that can be obtained near instantaneously while at sea. The disparity between the effort and time required for molecular analytical analyses versus real-time, bulk ocean characterizations was a long-standing limitation that drove the conceptualization of ecogenomic sensors. Previous engineering developments had made real-time assessments of many fundamental ocean properties possible. By extension, that progression applied equally to inventing new machines that would use molecular biological techniques to conduct measurements *in situ*; initially, it was a relatively modest and unencumbered proposition.

The push to develop ecogenomic sensors stemmed from studies of microorganisms such as viruses, bacterioplankton, phytoplankton, small zooplankton, and larval stages of larger animals. Under many circumstances, there is a need to rapidly identify and quantify particular species that may be rare and that cannot be distinguished easily. Applications related to water-quality monitoring, aquaculture and detection of invasive species exemplify such situations. DNA probe technology offers one way for enhancing organism detection, but the reliance on returning samples to a laboratory often nullifies many of the advantages that approach provides. Similarly, the idea of assessing the genomic capacity and activity of a community or population of organisms in real-time is desirable, particularly as a means to understand how organisms respond to environmental perturbations over daily, seasonal, annual and interannual time scales. In all cases, an ability to assess organism presence, abundance and activity *in situ*, instantaneously, can enhance decision making opportunities with respect to resource management and field research operations (e.g., Bowler *et al.*, 2009). Most importantly, a remote sensory presence allows for continued observations absent a human presence. It was from that confluence of innovations in genomics and ocean observatories that the first attempts to develop and deploy ecogenomic sensors grew.

Technical Challenges

The application of molecular analytical techniques in the aquatic sciences traditionally relies on laboratory personnel collecting discrete samples and processing them by applying a series of procedures that utilize various pieces of equipment housed in a laboratory. Automating and integrating that sequence of steps – from sample acquisition through

conditioning/fractionation to biochemical analyses – is a major challenge to realizing the ecogenomic sensor vision. Current, accepted laboratory methods that encompass that full sample processing cycle were never constrained by a need for hands-free operation for extended periods of time under ambient field conditions, let alone whilst submerged below the sea surface. Therefore, the development of ecogenomic sensors requires a philosophical departure from traditional wisdom derived from decades of conventional shipboard and laboratory operations.

A primary technical difficulty associated with such sensors is the need to acquire relatively large volumes of native water (often liters (l's) since target organisms may be rare), concentrate various sized particles, and then apply to the material retained a variety of chemical solutions on milli- and microfluidic (ml, μl) scales or lower. For example, using traditional laboratory methods, collection of particulate matter from 1 to 2 l of seawater might typically yield 50–100 μl of extracted nucleic acid, from which one to several microliters of extract (or less) may be used in a subsequent polymerase chain reaction (PCR) reaction (e.g., Preston *et al.*, 2011 and references therein). In that simple case, automation of the procedure dictates fluid handling capabilities that scale over at least six orders of magnitude. In addition, the procedure demands accessing "live" sample water that may contain a diverse assemblage of organisms and particulate matter, as well as highly refined, sterile and sensitive biochemical reagents. Raw sample water and other reagents must be partitioned and the integrity of each not compromised by cross contamination with others.

Most traditional molecular analytical methods also rely on expendable supplies and chemicals maintained and utilized under strict temperature regimes. Translating this approach to an autonomous machine creates a need for storing sample processing materials and ensuring that they remain stable and in good working order under conditions that are far from that of a shore based laboratory. Any design is further complicated when the details of different detection methodologies and intended uses or deployment locations are brought under scrutiny. Consequently, there is no single, overarching electromechanical design that satisfies the generic definition of an ecogenomic sensor (Scholin, 2010).

Those who have worked to automate sample processing for environmental science applications are by no means alone in their quest. Commercial ventures have pursued similar ideas for decades, giving rise to present-day companies that offer highly sophisticated, automated laboratory and field portable instruments (e.g., Cepheid (California, USA), Enigma Diagnostics (Salisbury, UK), Seattle Sensor Systems (Washington, USA)). However, such platforms were primarily designed around specific biomedical diagnostic or bio-threat detection applications, and the input volume of raw sample is extremely limited. Many research laboratories have also worked on developing portable instrument systems, or components thereof, that incorporate nucleic acid, antibody and/or receptor-based assays for use with environmental samples. Like the commercially available products, those systems are not necessarily suited for underwater use and generally require some type of laboratory infrastructure or personnel to effect a complete cycle of sample collection, processing and data transmission (e.g., Bruckner-Lea *et al.*, 2002; Belgrader *et al.*, 2003; Chandler and Jarrell, 2004; Fukuba *et al.*, 2004; Hindson *et al.*, 2004, 2005, 2008; Straub *et al.*, 2005; Casper *et al.*, 2007; Chinowsky *et al.*, 2007; LaGier *et al.*, 2007; Regan *et al.*, 2008; Bahi *et al.*, 2011).

In a brief review of optical and molecular analytical ocean instrumentation published in 2007, Paul *et al.* (2007) summarized the progress that had been made in realizing the ecogenomic sensor vision. At that time, there were only two systems that had succeeded in integrating the steps of sample collection and molecular analytical analyses into remotely operable, subsurface capable, ocean-deployable instruments: the Autonomous Microbial Genosensor (AMG) and the Environmental Sample Processor (ESP). Each of these devices met the conceptual definition an ecogenomic sensor, but neither was as highly evolved as that implied in **Figure 1**. Despite their limitations, the AMG and ESP provided a glimpse into research efforts that were aimed at fielding

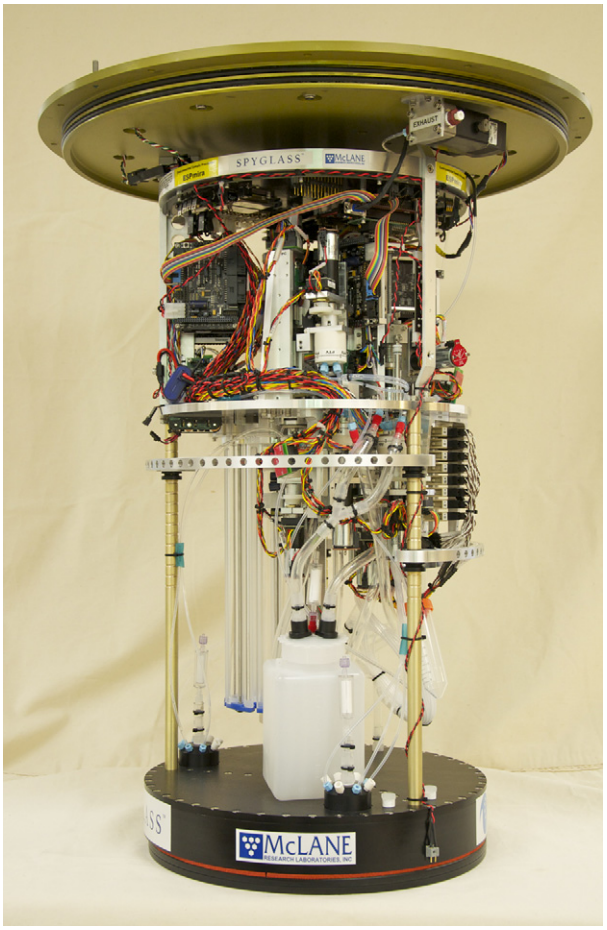


Figure 2 Photo of the second-generation (2G) ESP. The instrument is shown fitted (on top) with a surface-mooring pressure housing endcap (see **Figure 5**). The bulk of the instrument's electromechanical systems are on the top half of the device; visible at back in the mid-lower left side is a portion of the rotating carousel that holds pucks as shown in **Figure 3**. The black base is hollow and used to capture waste fluids. Reagents used for sample processing (not shown) are attached to the frame. The ESP is made available through Spyglass Biosecurity in partnership with McLANE Research Laboratories. The ESP is ~0.5-m diameter and 1-m tall. The instrument is protected under US Patent No. 6187530. Photo by © McLANE Research Laboratories, Inc.

automated molecular detection systems in an ocean setting. They also illustrated the maturity gap between molecular analytical instruments designed for ocean use versus the plethora of commercial systems aimed at the biomedical, food safety and military markets (Scholin, 2010). As time moved on, the ESP was shown to support a diverse set of detection chemistries and it was proven to be capable of operating in coastal, deep-sea and open ocean environments. In that way, the ESP came to exemplify an emergent ecogenomic sensor and served to highlight the potential for further development and application of that class of instrument.

The ESP

System Overview

Work on the ESP began in the mid-1990s. It was originally built as an experimental system to help define both the technological and operational elements that underlie the ecogenomic sensor concept. Accounts detailing the early fabrication and use of the ESP are summarized elsewhere (Scholin *et al.*, 2008; Roman *et al.*, 2007). It is a robotic device that was conceived around three, major functional elements: external sampling modules, a core sample processor ("core ESP"; **Figure 2**) and analytical modules (as detailed in Scholin *et al.*, 2009). These three components allow for a cascade of autonomous operations from acquiring a coarse-screened, multi-L sample from surface to deep ocean waters, to the detection and quantification of specific nucleic acid sequences or other biotic materials on a microliter scale.

Within the core ESP, "pucks" (**Figure 3**) are used to compartmentalize various steps associated with sample acquisition and processing. Pucks are stored in a rotating carousel. Robotic mechanisms serve to move pucks into and out of different sample processing stations, with each station being dedicated to specific functions like sample concentration, chemical processing or probe array imaging. Pucks can be configured to contain a variety of filters, or chemically reactive media depending on the specific application required. By programming a sequence of actions that coordinate moving pucks through the various stations in concert with actuating valves for collecting a sample and pumping reagents, the ESP functions autonomously to detect a variety of organisms and substances they produce (Roman *et al.*, 2007).

Sample Processing Capabilities

Initially, the ESP's design was heavily influenced by methods geared toward detection of groups of organisms and specific species based on ribosomal ribonucleic acid (rRNA) sequences. Such "signature sequences" can be detected within intact cells or through a variety of cell-free methods wherein target molecules are liberated into solution (e.g., Scholin *et al.*, 1996). By detecting and quantifying rRNA molecules one can ascertain the presence of certain species and under certain circumstances estimate their abundance as well (e.g., Metfies *et al.*, 2006 and references therein).

The sandwich hybridization assay (SHA) is a cell-free method that allows for rapid detection of rRNA (Van Ness and Chen, 1991; Marin and Scholin, 2010). When applied in a DNA probe array format, the technique provides a means to sense the presence of a variety of organisms simultaneously



Figure 3 Pucks used in the ESP to compartmentalize collection and processing of samples. All pucks conform to the same overall size and shape (~30 mm diameter × 17 mm tall) so that robotic systems used to move them are standardized against a constant form factor. Shown from left to right are pucks for sample collection and homogenization (for use with conventional membrane filters), large and small format probe arrays (arrays printed on 25-mm or 12-mm diameter membranes, respectively), sample preservation (for use with conventional membrane filters), and high-volume sample collection and homogenizing (for use with sintered wafer filters). Scale at the bottom is in inches. Photo by © McLANE Research Laboratories, Inc.

in a single test (e.g., Jones *et al.*, 2007; Greenfield *et al.*, 2008; Preston *et al.*, 2009; **Figure 4(a)**). Reagents for SHA are stored as liquids that are held in flexible bags or other containers attached to the core instrument; these chemicals are stable for many months over a wide temperature range. Similar to the SHA is a sample processing procedure that employs protein probe arrays (Doucette *et al.*, 2009). Protein probe arrays make it possible to use antibody-based techniques to quantify the presence of target molecules other than rRNA. Historically, the DNA and protein probe arrays as used in the ESP have been used in tandem to detect harmful algae and toxins they produce. The presence and abundance of molecules detected by both types of assays is conveyed via a chemiluminescent reaction. A camera captures an image of the light emitted from a puck (**Figure 3**), thus recording the intensity of reactions associated with each sensing element on a probe array. The entire process, from collection of a live sample to broadcast of an imaged array, takes approximately 2 h (Scholin *et al.*, 2009; http://www.mbari.org/ESP/espworks.htm).

The fluid handling system of the core ESP is capable of manipulating volumes down to roughly 100 µl. Although this is appropriate for demonstrating the use of the probe arrays, for example, that level of resolution precludes application of other desirable detection technologies, such as quantitative PCR (qPCR), that for a number of reasons demand precise fluid handling at the scale of at least 1 µl (Scholin, 2010). To meet the latter need, a separate fluid handling system was added. This ancillary system is referred to as the "microfluidic block," or MFB; it bolts to the core instrument and connects via single electrical and fluidic interfaces. The core ESP functions to concentrate and condition samples prior to passage to the MFB. The MFB then acts as a take-off point for distributing very small volumes of sample extracts and reagents to one or

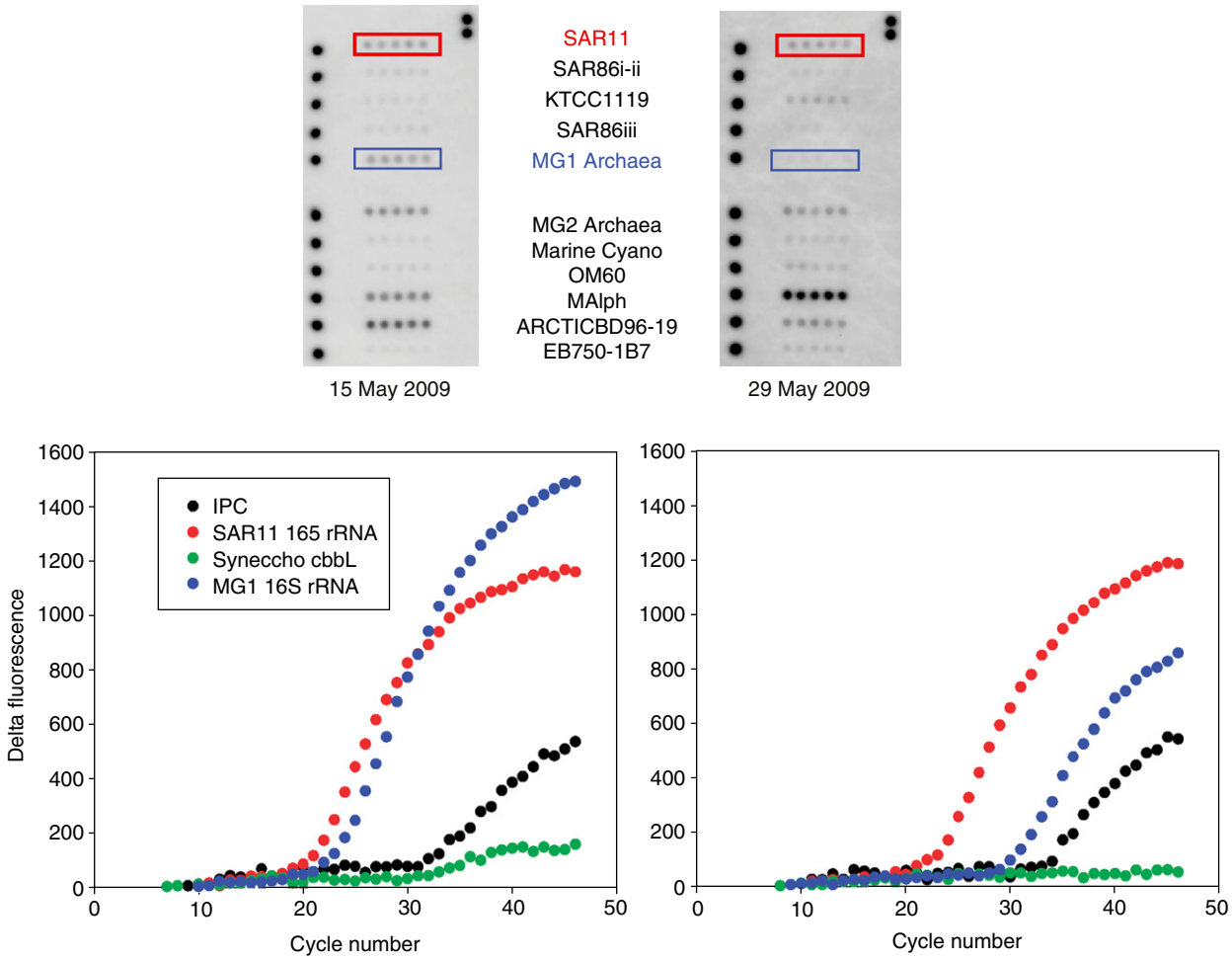


Figure 4 Examples showing results of DNA probe array (top) and quantitative PCR assays (bottom) obtained from two different periods during a field deployment of an ESP fitted with a MFB/PCR module. Probes targeting specific ribosomal RNA sequences (rRNA) indicative of various bacterioplankton are arranged horizontally; positive controls for verifying array orientation are arranged vertically (actual array size is ~10-mm tall by 5-mm wide, roughly half of the full large-array format used in the puck); darker spots indicate more intense reaction (more abundant target). Changes in the microbial community structure are apparent between the two sample periods. Highlighted on the arrays are rRNA probes indicative of SAR 11 (red) and marine crenarchaea (blue). The qPCR plots show reactions for the corresponding rRNA gene sequences along with a functional gene (RuBisCo) from *Synechococcus* (green) and an internal positive control (IPC; black). From this raw data one can determine that the marine crenarchaea were more abundant on May 15 than on May 29, 2009 (etc.). See Preston *et al.* (2011) for more details.

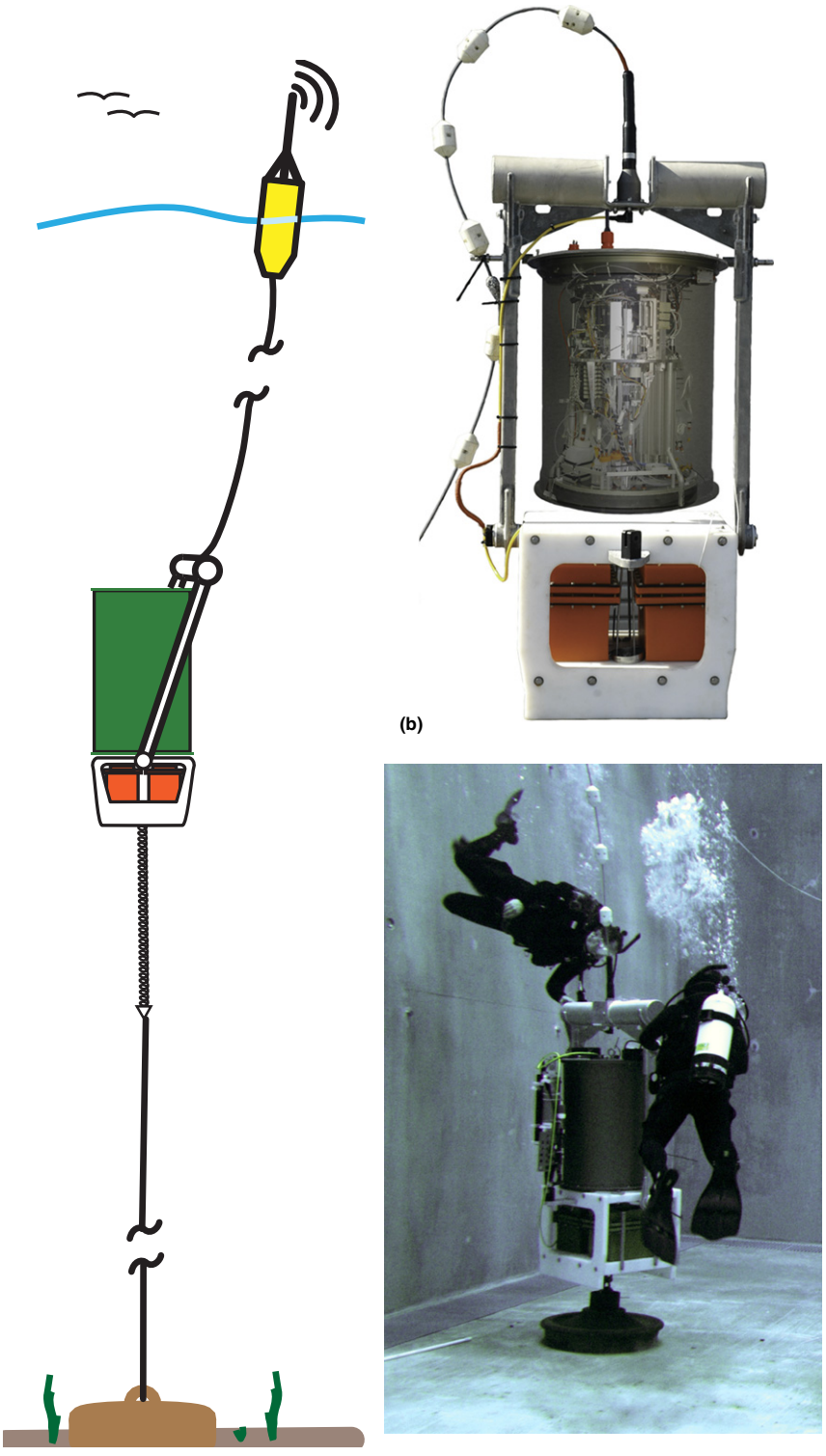
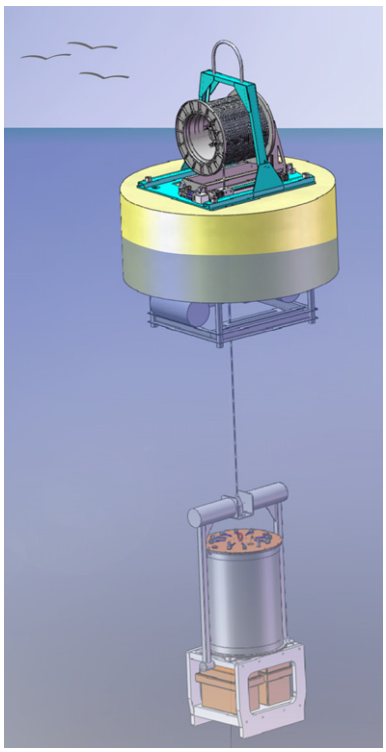
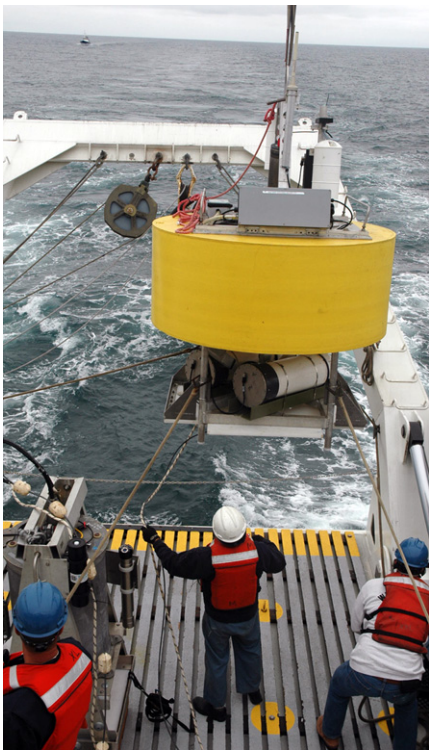


Figure 5 ESP surface mooring. The ESP is moored to the ocean bottom as illustrated in (a). The ESP is mounted on a subsurface platform (b) that carries batteries (in the orange boxes) and other ocean sensors; a electromechanical cable allows for communications and connection with the surface buoy. The assembly shown in (b) has a cutaway view of the pressure housing, revealing the ESP inside. Divers inspect the subsurface assembly during a test tank trial (c). Photo by Todd Walsh © 2006 MBARI.



(a)

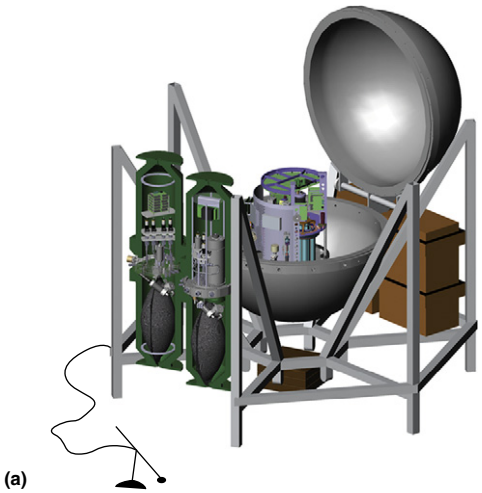


(b)



(c)

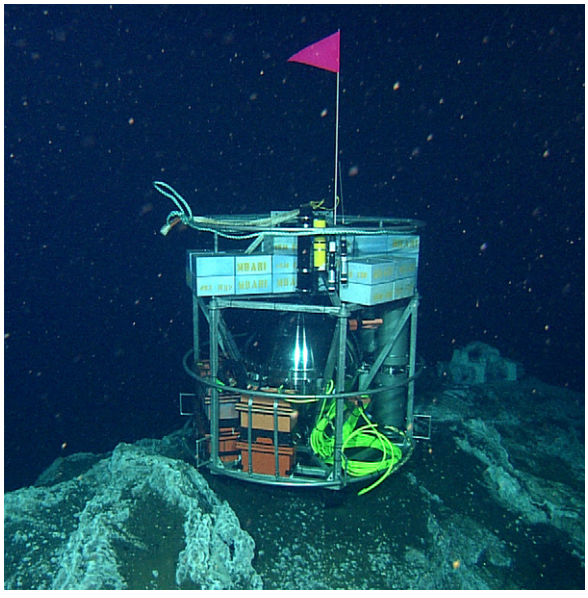
Figure 6 ESP open-ocean drifter. A surface float holds the ESP at a fixed depth from the surface via a electromechanical cable (a). As with the surface mooring, the ESP is mounted on a subsurface platform that carries other ocean sensors. Photo showing the deployment of the drifter surface float (b) and the freely drifting assembly with the R/V Western Flyer in the background (c). Photo by Philip Sammet © 2010 MBARI.



(a)



(b)



(c)

more molecular detection systems (or analytical modules as per Scholin *et al.*, 2009).

A DNA extraction system and flow-through, two-color (i.e., two optical channels), real-time PCR machine was recently incorporated into the MFB (Preston *et al.*, 2011; Robidart *et al.*, 2011). The MFB/PCR module permits application of a variety of polymerase enzyme master mixes, primer/probe combinations and assay controls. Presently, up to six qPCR reactions can be applied to a given sample at any one time. Reagents used with those assays are stored as liquids and have proven to be stable at room temperature for up to five months (the longest test conducted to date). The SHA and qPCR techniques can be applied to the same sample so that both an rRNA community profile and select gene abundances can be ascertained (e.g., **Figure 4**). As before, it takes approximately 2 h between collecting a live sample to recording an array image; results from each qPCR reaction are available roughly every 2 h thereafter.

Complementing the real-time data provided by the probe arrays and qPCR is the core ESP's ability to preserve discrete samples for subsequent laboratory-based analyses. Preserving samples *in situ* has a number of applications. That capability was originally motivated by a desire to validate trends in species abundance observed during field deployments (e.g., Greenfield *et al.*, 2006). It also provides additional sample materials that can be used to support analyses that are not yet appropriate for use with an instrument like the ESP (e.g., nucleic acid sequencing (Ottesen *et al.*, 2011), proteomics research (Saito *et al.*, 2011)).

Operational Uses

The ESP has been deployed on a variety of platforms including coastal moorings (**Figure 5**), an open ocean drifter (**Figure 6**), and a 4000 m-rated benthic lander for delivering the instrument to the seafloor (**Figure 7**). When deployed, the ESP is bundled with sensors for assessing basic ocean properties like temperature, salinity, chlorophyll (etc.) so that samples collected, and the assay results obtained, can be referenced against prevailing environmental conditions. The ESP can be operated in a standalone mode or as part of a larger ocean-observing system (e.g., Ryan *et al.*, 2011). Two-way communications with the machine are achieved using a radio or acoustic modem, satellite link or direct (wired) network connection.

Figure 7 Free-fall vehicle for deploying the ESP to 4000-m depth. As illustrated in a cutaway view (a), the ESP is mounted in spherical pressure housing that is attached to a larger structure for free-fall deployment to the ocean bottom. An external sampling module (two tall cylinders at left) allows for inhaling up to 10 l of water at a time prior to making it available to the core ESP; a hose emanating from the sampling module directs sample intake from precise locations on the seafloor. The author stands with the D-ESP prior to its deployment (b); the lander holds a variety of other sensors, batteries, and data telemetry systems. The D-ESP as seen by remotely operated vehicle atop a carbonate mound at 890-m depth off Southern California (c); the mound vents methane gas, and the surrounding waters are enriched with methane-oxidizing bacteria. Copyright by © 2010 MBARI.

To date, most applications of the ESP have focused on surface waters with foci being bacterioplankton, harmful algal blooms (HABs) and invertebrates (*see* Sample Processing Capabilities). Individual instruments have been used to detect one or all of those species during coastal mooring deployments (**Figure 5**) each of which lasts roughly 30 days. Most of the work has concentrated on feasibility studies and methods demonstrations. Recently, progress has been made in utilizing the ESP to reveal novel marine microbial community dynamics in coastal waters (Preston *et al.*, 2011; Robidart *et al.*, 2011) and patchiness of species responsible for HABs (Ryan *et al.*, 2011). Other studies have focused on testing specific hypotheses regarding HAB initiation and evaluating different strategies for detecting small invertebrates. Academic and government agencies are presently evaluating the utility of the ESP to carry out water-quality assessments associated with sewage effluents.

Oceanic microbial community structure and function is generally referenced against environmental conditions that are assessed at the instant samples are collected. However, the biological community collected at any one time is derived from past as well as current conditions. Consequently, there is

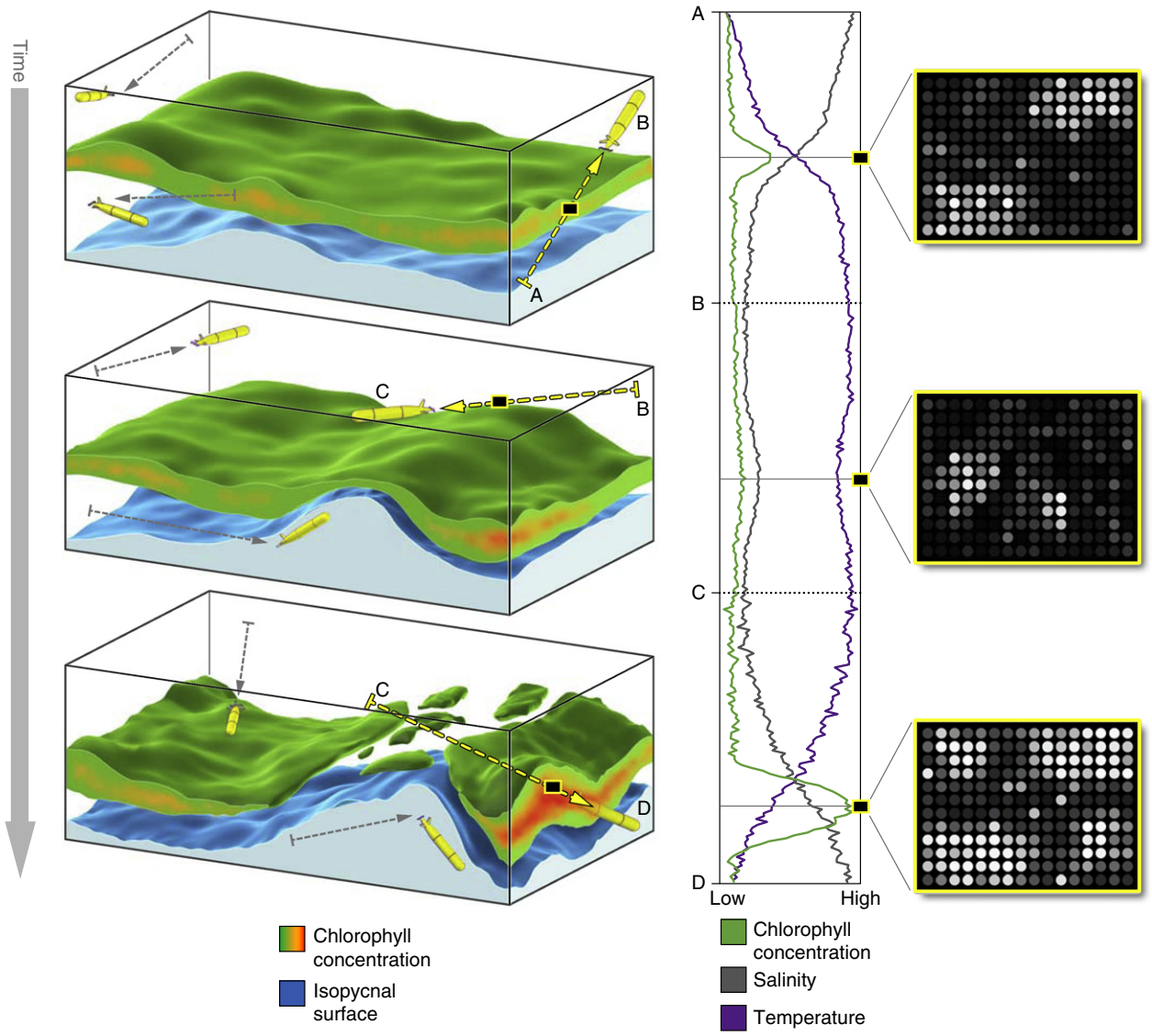


Figure 8 Hypothetical illustration showing a fleet of autonomous underwater vehicles (AUVs; far left panels) characterizing properties of the water column (middle panel) that vary over both space (denoted in steps A–D) and time (progressing from top to bottom). In this illustration, a layer of phytoplankton is broken apart and redistributed as a result of an internal wave (propagating density discontinuity). Sensor outputs obtained from one of the AUVs as it traverses through the water column are shown in the middle panel. For each time/space step (A–B, B–C, C–D as seen in the far left panels), the vehicle moves through the water column recording basic ocean properties like temperature, salinity and chlorophyll. A discrete sample collected during each transit (black box along the AUV track) is processed using an ecogenomic sensor and DNA probe arrays (on far right) to reveal microorganisms' gene content and expression. Brightness of each probe spot indicates the abundance of a particular gene and/or its corresponding messenger RNA. By coordinating the operation of the AUVs and other ocean observing equipment, chemical/physical perturbations and the biological community's response are thus recorded in volumetric fashion.

an inherent difficulty in establishing detailed cause-and-effect relationships since the history of the community or population in question is generally not known. For this reason, a platform that allows the ESP to drift freely in the open ocean was created (**Figure 6**). So far, the drifting ESP has been deployed only for short periods (up to 10 days as of 2011), but it runs continuously throughout that period to yield very-high-resolution time-series data. Such deployments have recently taken place along the central California coast (Ottesen *et al.*, unpublished) and in waters north of Hawaii (Robidart *et al.*, unpublished). A long-term goal is to translate this experimental system into one that is capable of actively identifying, tracking, and sampling coherent microbial populations so that changes in their community structure and gene expression can be assessed with respect to both space and time domains.

The advent of cabled observatories as depicted in **Figure 1** is opening a new door for operating novel sensor systems on the seafloor. With that are new opportunities for assessing microbial community structure and function in particular locations and on temporal scales not achievable previously. These developments spurred on the idea for a platform that would allow the core ESP to operate in the deep sea (**Figure 7**), in particular at sites dominated by hydrothermal vents, hydrocarbon seeps, and other features of intense microbiological interest. A special pressure housing and external sampling module that isolates the core ESP from such extreme environments was thus devised, allowing the instrument to function in a deep-ocean setting as it does in surface waters. It is expected that the deep-sea ESP will operate for periods up to 1 year between servicing intervals when powered by an ocean bottom cable node.

Conclusions and Future Prospects

Ecogenomic sensors are fast becoming a reality. With the expansion of coastal and global ocean sensing systems, opportunities for developing and fielding ecogenomic sensors should increase. For the first time in the history of the ocean sciences, it is realistic to conceive of carrying out interactive experiments remotely *in situ* from a molecular biological perspective along the lines of that conceptualized in **Figure 1**. Recent progress made in bringing the ESP to operational status has helped lend credibility to that vision, but much more remains to be done if ecogenomic sensors are to become commonplace in the environmental sciences.

A primary need is to develop a new generation of these devices that are much smaller and mechanically less complex than the current ESP system. Once that type of ecogenomic sensor exists, it is likely that the push will be to integrate such sensors with mobile platforms such as autonomous underwater vehicles (AUVs). If that can be accomplished, then researchers will have an ability to collect and analyze discrete samples in response to changing environmental conditions without the current restrictions imposed by traditional shipboard operations, much in the same way that profiling floats equipped with chemical sensors have revolutionized our ability to observe the interior of the global ocean (e.g., Johnson *et al.*, 2010). No longer is it a pipe dream to imagine fleets of mobile systems working cooperatively to identify specific communities of microorganisms and follow them over time (cf. Ryan *et al.*, 2010; Zhang *et al.*, 2010, 2011), revealing time and space relationships that ultimately unite environmental forcing and organism function and activity (Zehr *et al.*, 2011; **Figure 8**).

From an engineering perspective, future development of ecogenomic sensors should be driven by a clear set of science or monitoring requirements rather than what may be possible to achieve technologically (Scholin, 2010). For example, ecogenomic sensors may be needed on moorings, cabled networks, profiling systems, near the surface, or at great depth. They may be readily accessible or in very remote locations, highly configurable, or very limited in their functionality. Accordingly, operations and maintenance requirements as well as system costs will vary greatly. Proxy measurements and modeling techniques will be needed to a greater or lesser extent to interpolate between a discrete and limited set of sensors collecting molecular analytical data. Although some common themes will almost certainly emerge, it is highly doubtful that a single design will meet all the ecogenomic sensing needs articulated by researchers and resource managers.

Appendix

List of Courses

1. C-MORE Summer Course on Microbial Oceanography
2. Stanford Summer Course on Marine Microbiology (Hopkins Marine Station)
3. Marine Biological Laboratory Summer Course on Microbiology
4. Oceans and Human Health Course at the University of Miami

Acknowledgments

This work was made possible by grants from the David and Lucile Packard Foundation, National Science Foundation (NSF), NASA Astrobiology, the Gordon and Betty Moore Foundation, and the Keck Foundation. The author gratefully acknowledges partners comprising the Center for Microbial Oceanography: Research and Education (C-MORE) for their help in shaping the development and application of ecogenomic sensors; E. Virginia (Ginger) Armbrust for providing the history behind **Figure 1**; Christina Preston, Julie Robidart, John Ryan, David Fierstein, Annette Gough, Shannon Boedecker, Todd Walsh, and James Birch for their help with figures and photos; and Patricia Duran for assistance with manuscript preparation.

See also: Biodiversity of Harmful Marine Algae. Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea. Eutrophication and Oligotrophication. Marine and Aquatic Communities, Stress from Eutrophication. Marine Ecosystems. Marine Viruses. Microbial Biogeography. Proteorhodopsins: Widespread Microbial Light-Driven Proton Pumps

References

Bahi MM, Tsaloglou MN, Mowlem M, and Morgan H (2011) Electroporation and lysis of marine microalga *Karenia brevis* for RNA extraction and amplification. *Journal of the Royal Society Interface* 8(57): 601–608. http://dx.doi.org/10.1098/rsif.2010.0445.

Belgrader P, Elkin CJ, Brown SB, *et al.* (2003) A reusable flow-through polymerase chain reaction instrument for the continuous monitoring of infectious biological agents. *Analytical Chemistry* 75: 3446–3450.

Bowler C, Karl DM, and Colwell RR (2009) Microbial oceanography in a sea of opportunity. *Nature* 459: 180–184.

Bruckner-Lea CJ, Tsukuda T, Dockendorff B, *et al.* (2002) Renewable microcolumns for automated DNA purification and flow-through amplification: From sediment samples through polymerase chain reaction. *Analytica Chimica Acta* 469: 129–140.

Casper ET, Patterson SS, Bhaunushali P, *et al.* (2007) A handheld NASBA analyzer for the field detection and quantification of *Karenia brevis*. *Harmful Algae* 6: 112–118.

Chandler DP and Jarrell AE (2004) Automated purification and suspension array detection of 16s rRNA from soil and sediment extracts by using tunable surface microparticles. *Applied and Environmental Microbiology* 70: 2621–2631.

Chinowsky TM, Soelberg SD, Baker P, *et al.* (2007) Portable 24-analyte surface plasmon resonance instruments for rapid, versatile biodetection. *Biosensors and Bioelectronics* 22: 2268–2275.

Doucette GJ, Mikulski CM, Jones KL, *et al.* (2009) Remote, subsurface detection of the algal toxin domoic acid onboard the Environmental Sample Processor: Assay development and initial field trials. *Harmful Algae* 8: 880–888.

Fukuba T, Yamamoto T, Niganuma T, and Fuji T (2004) Microfabricated flow-through device for DNA amplification – Towards in situ gene analysis. *Chemical Engineering Journal* 101: 151–156.

Greenfield D, Marin III R, Doucette GJ, *et al.* (2008) Field applications of the second-generation Environmental Sample Processor (ESP) for remote detection of harmful algae: 2006–2007. *Limnology and Oceanography: Methods* 6: 667–679.

Greenfield DI, Marin III R, Jensen S, Massion E, *et al.* (2006) Application of the Environmental Sample Processor (ESP) methodology for quantifying *Pseudo-nitzschia australis* using ribosomal RNA-targeted probes in sandwich and fluorescent in situ hybridization. *Limnology and Oceanography: Methods* 4: 426–435.

Hindson BJ, Brown SB, Marshall GD, *et al.* (2004) Development of an automated sample preparation module for environmental monitoring of biowarfare agents. *Analytical Chemistry* 76: 3492–3497.

Hindson BJ, Gutierrez DM, Ness KD, *et al.* (2008) Development of an automated DNA purification module using a micro-fabricated pillar chip. *Analyst* 133: 248–255.

Hindson BJ, McBride MT, Makarewicz AJ, *et al.* (2005) Autonomous detection of aerosolized biological agents by multiplexed immunoassay with polymerase chain reaction confirmation. *Analytical Chemistry* 77: 284–289.

Johnson KS, Riser C, and Karl DM (2010) Nitrate supply from deep to near-surface waters of the North Pacific subtropical gyre. *Nature* 465: 1062–1065.

Jones WJ, Preston C, Marin III R, Scholin C, and Vrijenhoek R (2007) A robotic molecular method for in situ detection of marine invertebrate larvae. *Molecular Ecology Resources* 8: 540–550.

LaGier MJ, Fell JW, and Goodwin KD (2007) Electrochemical detection of harmful algae and other microbial contaminants in coastal waters using hand-held biosensors. *Marine Pollution Bulletin* 54: 757–770.

Marin III R and Scholin C (2010) Sandwich hybridization. In: Karlson B, Cusack C, and Bresnan E (eds.) *Microscopic and molecular methods for quantitative phytoplankton analysis*. Paris: UNESCO. Ch 12, IOC Manuals and Guides, no. 55. (IOC/2010/MG/55).

Metfies K, Töbe K, Scholin C, and Medlin LK (2006) Laboratory and field applications of ribosomal RNA probes to aid the detection and monitoring of harmful algae. In: Granéli E and Turner JT (eds.) *Ecology of Harmful Algae*, pp. 311–325. Berlin, Heidelberg, New York: Springer Verlag.

Ottesen EA, Marin III R, Preston CM, *et al.* (2011) Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *The ISME Journal* online publication, June 30, 2011. http://dx.doi.org/10.1038/ismej.2011.70.

Paul JH, Scholin C, van Den Engh G, and Perry MJ (2007) In situ instrumentation. *Oceanography* 20: 58–66.

Preston C, Marin III R, Jenson S, *et al.* (2009) Near real-time, autonomous detection of marine bacterioplankton on a coastal mooring in Monterey Bay, California, using rRNA-targeted DNA probes. *Environmental Microbiology* 11: 1168–1180.

Preston CM, Harris A, Ryan JP, *et al.* (2011) Underwater application of quantitative PCR on an ocean mooring. *PLoS ONE* 6(8): e22522. http://dx.doi.org/10.1371/journal.pone.0022522.

Regan JF, Makarewicz AJ, Hindson BJ, *et al.* (2008) Environmental monitoring for biological threat agents using the autonomous pathogen detection system with multiplexed polymerase chain reaction. *Analytical Chemistry* 80: 7422–7429.

Robidart JC, Preston CM, Paerl RW, *et al.* (2011) Seasonal dynamics of Synechococcus and Thaumarchaeal microbial populations in Monterey Bay resolved in real time with remote in situ instrumentation. *The ISME Journal* online publication, October 6, 2011. http://dx.doi.org/10.1038/ismej.2011.127.

Roman B, Scholin C, Jensen S, *et al.* (2007) Controlling a robotic marine water sampler with the ruby scripting language. *Journal of American Laboratory Automation* 12: 56–61.

Ryan J, Greenfield D, Marin III R, *et al.* (2011) Harmful phytoplankton ecology studies using an autonomous molecular analytical and ocean observing network. *Limnology and Oceanography* 56: 1255–1272.

Ryan JP, Johnson SB, Sherman A, *et al.* (2010) Mobile autonomous process sampling within coastal ocean observing systems. *Limnology and Oceanography: Methods* 8: 394–402.

Saito MA, Bulygin VV, Moran DM, *et al.* (2011) Examination of microbial proteome preservation techniques applicable to autonomous environmental sample collection. *Frontiers in Aquatic Microbiology* 2: article 215 (10.3389/fmicb.2011.00215).

Scholin C, Doucette G, Jensen S, *et al.* (2009) Remote detection of marine microbes, small invertebrates, harmful algae, and biotoxins using the Environmental Sample Processor (ESP). *Oceanography* 22: 158–167.

Scholin CA (2010) What are "Ecogenomic sensors?" – A review and thoughts for the future. *Ocean Science Discussion* 6: 51–60.

Scholin CA, Buck KR, Britschgi T, Cangelosi J, and Chavez FP (1996) Identification of *Pseudo-nitzschia australis* (Bacillariophyceae) using rRNA-targeted probes in whole cell and sandwich hybridization formats. *Phycologia* 35: 190–197.

Scholin CA, Doucette GJ, and Cembella AD (2008) Prospects for developing automated systems for in situ detection of harmful algae and their toxins. In: Babin M, Roesler CS, and Cullen JJ (eds.) *Real-Time Coastal Observing Systems for Ecosystem Dynamics and Harmful Algal Blooms*, pp. 413–462. Paris, France: UNESCO Publishing.

Straub TM, Dockendorff BP, QuiÒonez-DÌaz MD, *et al.* (2005) Automated methods for multiplexed pathogen detection. *Journal of Microbiological Methods* 62: 303–316.

Van Ness J and Chen L (1991) The use of oligodeoxynucleotide probes in chaotrope-based hybridization solutions. *Nucleic Acids Research* 19: 5143–5151.

Zehr JP, Robidart J, and Scholin C (2011) Marine microorganisms, biogeochemical cycles, and global climate change. *Microbe* 6: 169–175.

Zhang Y, McEwen RS, Ryan JP, *et al.* (2011) A peak-capture algorithm used on an autonomous underwater vehicle in the 2010 Gulf of Mexico oil spill response scientific survey. *Journal of Field Robotics* 28: 484–496. http://dx.doi.org/10.1002/rob.20399.

Zhang Y, McEwen RS, Ryan JP, and Bellingham JG (2010) Design and tests of an adaptive triggering method for capturing peak samples in a thin phytoplankton layer by an autonomous underwater vehicle. *IEEE Journal of Oceanic Engineering* 35: 785–796.

Ecological Footprint, Concept of

William E Rees, University of British Columbia, Vancouver, BC, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Carrying capacity Usually defined as the average maximum number of individuals of a given species that can occupy a particular habitat without permanently impairing the productive capacity of that habitat.

Competitive exclusion The displacement of one species from its habitat or ecological niche by another. When humans appropriate other species' ecological space, it often leads to the local or even the global extinction of the nonhuman organism.

Ecological deficit An ecological deficit exists when the load imposed by a given human population on its own territory or habitat (e.g., region, country) exceeds the productive capacity of that habitat. Under these circumstances, if it wishes to avoid permanent damage to its local ecosystems, the population must use some biophysical goods and services imported from elsewhere (or, alternatively, lower its material standards).

Human load The total human load imposed on the environment by a specified population is the product of population size times average per capita resource consumption and waste production. The concept of load recognizes that human carrying capacity is a function not only of population size but also of aggregate material and energy throughput. Thus, the human carrying capacity of a defined habitat is its maximum sustainability supportable load.

Overshoot A population is in overshoot when it exceeds available carrying capacity. A population in overshoot may permanently impair the long-term productive potential of its habitat, reducing future carrying capacity. It may survive temporarily but will eventually crash as it depletes vital natural capital (resource) stocks.

Patch disturbance The measurable habitat and ecosystem modification caused by large animals, including humans, as they forage for food or other resources. Patch disturbance is most pronounced near the den site, temporary camp, or other central place within the overall home range of the individual or group.

Sustainability gap The global ecological deficit – that is, the difference between any excessive human load on the ecosphere and the long-term carrying (or load-bearing) capacity of the planet.

Ecological footprint analysis (EFA) is a quantitative tool that represents the ecological load imposed on the earth by humans in spatial terms. Thus, the ecological footprint (EF) of a defined population is the total area of land and water ecosystems required to produce the resources that the population consumes, and to assimilate the wastes that the population generates, wherever on earth the land/water are located. Ecofootprinting is used to assess the ecosystem area effectively appropriated in support of any specified human population or economic activity. We can then compare this with available productive area (i.e., load-bearing capacity). The size and nature of the human EF is relevant to biodiversity conservation because energy and material resources extracted from nature to serve human purposes are irreversibly unavailable to other species. The larger the human footprint, the less the nonhuman biodiversity. This article describes the contemporary context for ecofootprint analysis and the theory on which it is based. It then discusses the implications of the analysis for sustainable development in general and for the future of biodiversity in particular.

The Conceptual Framework: Why Do We Need Ecofootprint Analysis?

Human Activity and Global Ecological Change

There is agreement among scientists that humanity has assumed a major role in changing the face of the earth. Up to one-half of the land has been directly transformed by human action; more than half of the planet's accessible fresh water is already being used by humans; atmospheric carbon dioxide has increased from 280 to 393 parts per million (40%) during the industrial era and the resultant climate change is increasingly being accepted as a fact; more atmospheric nitrogen is fixed and injected into terrestrial ecosystems by humans than by all natural terrestrial processes combined; two-thirds of the world's major fisheries are fully or overexploited; and biodiversity losses are accelerating (Hansen *et al.*, 2010; Lubchenco, 1998; NOAA, 2011; Vitousek *et al.*, 1997).

What is in dispute is whether any of this really matters to socioeconomic sustainability. The more pessimistic analysts see human domination of the earth and the deterioration of nature as ultimately fatal to civilization itself. Various mechanisms are invoked in coming to this conclusion. For example, because of our culture's dependence on liquid petroleum and the present lack of affordable substitutes, energy analyst Richard Duncan (1993) argues that the total life expectancy of western technoindustrial society is less than 100 years – counting from the 1930s! (Significantly, the International Energy Agency conceded in 2010 that global production of conventional crude oil peaked in 2006, and is now in decline (EIA, 2010).) Hern attacks industrial culture's obsession with growth and compares humanity with a self-destructive cancer on the planet. The continuous expansion of the human enterprise apparently "exhibits all four major characteristics of a malignant process: rapid uncontrolled growth; invasion and

destruction of adjacent tissues (ecosystems, in this case); metastasis (colonization and urbanization, in this case); and dedifferentiation (loss of distinctiveness in individual components)'' (Hern, 1997). According to geographers Smith and Sauer-Thompson, environmental decay, coupled with geopolitical stresses induced by resource scarcity, will inevitably cause industrial society to ''self-destruct, producing massive ecological damage, social chaos, and mega-death'' (Smith and Sauer-Thompson, 1998). From these perspectives, the human enterprise has already overshot global carrying capacity, is eroding future options, and will inevitably crash – there is virtually nothing we can do to prevent the collapse of global civilization.

At the other extreme, we can take comfort in the views of those scientists and economists who believe that human ingenuity will prevail. Their logic is simple. Stimulated by incipient scarcity, rising prices, and the opportunity for profit, scientists and entrepreneurs will find technological substitutes for exhaustible resources and natural processes. As energy economist Gordon explains, ''the tendency to technical progress is viewed as the most critical economic law involved.'' Moreover, because human ingenuity has historically been remarkable at increasing living standards and warding off the pressures of resource depletion, ''the immediate need for avoiding depletion is nil'' (Gordon, 1994).

Perhaps the most extreme expression of this belief is revealed in the ebullient optimism of the late Julian Simon, who wrote: ''Technology exists now to produce in virtually inexhaustible quantities just about all the products made by nature'' and ''We have in our hands now … the technology to feed, clothe, and supply energy to an ever-growing population for the next seven billion years'' (Simon, 1995). Inspired by Simon, Danish statistician Bjørn Lomborg, in *The Skeptical Environmentalist*, dismissed concerns about population growth, potential resource shortages, and pollution as being unsupported by the data (Lomborg, 2001). Some conventional economists even argue the benefits of continuous population growth on grounds that ''additional people can create more resources than they use up, thanks to technological improvements'' (Block, 1990). Clearly, from the techno-optimists' perspective, there are no practical limits at all on human carrying capacity – we do not need to do anything to sustain global civilization but stay our present, increasingly market-driven, global development course.

With such diametrically opposing assessments of the human prospect, both starting from present circumstances and held with equal conviction, it is little wonder that ordinary citizens and policy makers alike are confused about global ecological change and what – if anything – to do about it. This article assesses the overall sustainability conundrum using EF analysis and links the general problem to the specific issue of biodiversity conservation. First, however, we examine the origins of the ongoing perceptual conflict.

Evolving Economic Paradigms

Many of the conflicting views about sustainable development can be traced to differing fundamental beliefs, values, and assumptions about the nature of reality, particularly humankind–environment relationships. The fact is that human beings socially construct their perceptions of reality (Berger and Luckmann, 1966). Different societies and interest groups therefore possess quite different conceptual lenses (worldviews, narratives, preanalytic visions, and paradigms) through which their members interpret new data, assertions, or circumstances. This is no trivial observation – group members may vigorously defend existing perceptions against contrary data, even though, as social constructs, their perceptions may be only tenuously connected to reality.

Remarkably, people are often unaware that they possess a particular worldview – each person acquires his or her culturally specific narrative unconsciously, simply by growing up in a particular socio-cultural milieu. We may therefore be unconscious of the subtle ways in which the prevailing worldview shapes our approach to critical issues or that there may be more viable alternatives. When we think that ''the world is flat'' was once accepted as self-evident truth, it raises the unsettling possibility that much of our present cultural worldview consists largely of shared illusions.

The dominant social paradigm in western (and increasingly global) technoindustrial capitalist culture can be characterized as the expansionist or the cornucopian worldview. As we have already seen, its strictest adherents believe that humankind has achieved mastery over the natural world and that technology, abetted by market processes, will be able to compensate for the depletion of natural resources and the loss of life-support services.

The economics of this expansionist vision is the neoclassical (neoliberal) economics that has come to dominate geopolitics in the past 50 years. The starting point for neoliberal analysis is the circular flow of exchange value in which money recycles repeatedly between producer firms (as payment for goods and services) and consumer households (as wages, salaries, and returns on investment). This model portrays the economy as a self-renewing, self-feeding perpetual motion machine, as an isolated system with no inlets or outlets to anything outside itself; the economy floats free of the natural world, its productivity and growth unconstrained by biophysical limits. Because conventional economics treats the economy as separate from, and essentially independent of, the environment, development policy has historically treated humans solely as economic agents.

While still dominant in policy circles, the neoliberal perspective as described is increasingly being challenged. The accelerating pace of global change, including an alarming accumulation of market externalities, has convinced many mainstream economists of the need to reconsider the implications of unlimited growth, to acknowledge the complex dynamics of de facto world economy–ecosystems relationships, and to address the failures of market prices to reflect real environmental and social costs (e.g., Arrow *et al.*, 1996).

The most 'out-of-the-box' critique of the mainstream is leveled by those environmentalists and ecological economists who profess an ecological or a steady-state worldview. The latter implies that energy/material flows through the economy should be maintained at a nongrowing optimum, safely within the means of nature. These analysts see the present economy not as a separate system, but rather as an inextricably integrated, completely contained, and wholly dependent

growing subsystem of a dynamic but nongrowing ecosphere (e.g., Daly, 1991). From this perspective, the economy (i.e., the human enterprise) is potentially parasitic on nature. Hence, people must be considered not only as economic agents but also as potentially destructive ecological agents.

Indeed, the ecological perspective actually sees the so-called environmental crisis as a human ecobehavioral crisis. This is a critical distinction. The former term externalizes the problem, effectively blaming it on a defective environment, which then needs to be fixed. By contrast, the latter term traces the problem to its source – the nature and behavior of people themselves – and suggests that it is the latter that need fixing. As we shall see, this is the starting point for EFA.

Human Ecology/Economy and the Second Law

If humans are ecological entities, then human activity is governed by natural laws. In particular, economic production requires continuous, irreversible energy and material transformations, and these transformations are ultimately governed by the second law of thermodynamics. This fundamental law is ignored by conventional economic models.

In its simplest form, the second law states that with every spontaneous change, an isolated system will tend toward equilibrium; alternately, the entropy of isolated systems always increases. Available energy dissipates, concentrations disperse, and gradients disappear. An isolated system thus becomes increasingly randomized in an inexorable slide toward thermodynamic equilibrium. This is a state of maximum entropy in which there is no structure and nothing can happen.

The second law was originally formulated for simple isolated systems close to equilibrium. However, even complex open systems are subject to the forces of entropic decay. Any differentiated far-from-equilibrium system has a natural tendency to erode and unravel.

But not all complex systems disintegrate in this way. Many biophysical systems, from individual fetuses to the entire ecosphere, actually gain in organizational complexity and mass over time (i.e., they increase their distance from equilibrium). This seemingly paradoxical behavior can readily be reconciled with the second law. Biophysical systems exist in loose, nested hierarchies, each component system being contained by the next level up and itself comprising a chain of linked subsystems at lower levels. Living systems can therefore import available energy and material (negentropy) from their host environments and use it to maintain their internal integrity against entropic decay and to grow (Kay and Regier, 2000). They also export the resultant waste (entropy) back into their hosts. In effect, modern formulations of the second law posit that all highly ordered systems grow and develop (increase their internal order) at the cost of destroying order at higher levels in the systems hierarchy (Schneider and Kay, 1995). Because such systems maintain themselves by continuously degrading and dissipating available energy and matter, they are called "dissipative structures" (Prigogine, 1997).

Now, the human enterprise is clearly one such highly ordered, complex, dissipative structure. Both our biological and our industrial metabolisms require enormous inputs of high-grade energy and material resources from the rest of the ecosphere for their maintenance and growth. Indeed, from the ecological perspective, humans are strictly secondary producers – all production by the human enterprise, from the increase in population to the accumulation of manufactured capital, requires the consumption of a much larger quantity of energy and material first produced by nature (we can sometimes increase the amount of incoming solar energy captured by photosynthesis through irrigation or by building terraces for agriculture).

Moreover, recall that within the global hierarchy, the economy is a dependent subsystem of the ecosphere. In effect, the expanding human enterprise is positioned to consume the ecosphere from within. Beyond certain limits, continuous population and material economic growth is, therefore, inherently self-destructive. This reality suggests a necessary second law condition for sustainability: material consumption and waste production by the economy must be no greater than the resource production and waste assimilation capacity of the ecosphere.

Homo Sapiens: The Archetypal Patch Disturbance Species

Evidence to support the ecological view has accumulated throughout history. Even preagricultural humans had significant effects on local ecosystems' structure and function and on the biodiversity of their habitats. This was the inevitable consequence of the second law combined with two additional facts of human biology: human beings are large animals with correspondingly large individual energy and material requirements (i.e., we are huge consumers) and humans are social beings who live in extended groups. Consequently, whenever human hunter-gatherers invaded a previously stable ecosystem, their material demands would significantly alter established energy and material pathways, benefiting some nonhuman species and harming others. People invariably perturb or disturb the systems of which they are a part.

Perhaps the most dramatic evidence is the permanent systemic changes that occur when humans first invade and settle a new habitat. Consider biodiversity loss. The recent paleo ecological, anthropological, and archeological literature tells a convincing story of the extinctions of large mammals and birds that accompanied first contact and settlement of their habitats by human beings (Diamond, 1992; Flannery, 1994; Ponting, 1991). It seems that everywhere on earth that paleontologists have studied and that humans first reached within the past 50 ky, human arrival approximately coincided with massive prehistoric extinctions. In North America, South America, and Australia, about 72%, 80%, and 86%, respectively, of large mammal genera ultimately became extinct after human arrival. Scientists estimate that with only Stone Age technology, the Polynesians exterminated more than 2000 bird species, about 15% of the world total.

All this is to emphasize that humans are, by nature, a patch-disturbance species, a distinction we share with other large mammals ranging from beavers to elephants. Large animals, due to their size, longevity, and food and habitat requirements, tend to have substantial physical and systemic

impacts on the ecosystems that sustain them. A patch-disturbance species may thus be defined as any organism that, usually by foraging, degrades a small central place greatly and disturbs a much larger area away from the central core to a lesser extent.

There is, of course, a major difference between human patch disturbance and that of other species. Because human knowledge and technology are uniquely cumulative, human patch disturbance has been intensifying since the Neolithic. It received a major boost with agriculture and became the dominant force in the ecosphere with the use of fossil fuels and the industrial revolution. (Cheap, plentiful fossil fuels have enabled humans to accelerate the exploitation of everything else.) Today, human patch disturbance is evident on a global scale in the form of such persistent negative trends as greenhouse gas accumulation, increasing climatic variability, ozone depletion, tropical deforestation, desertification, ocean acidification, and accelerating biodiversity loss. Ultimately, the erosion of systems integrity may lead to the irreversible loss of basic life support functions.

Whatever joy one might take from the human symphony, the fact that material economic growth now threatens critical biophysical systems sounds a discordant note. Global change implies that economic activity is pressing against the limits of human carrying capacity at the expense of other species – and ultimately of people themselves. Regrettably, the concept of human carrying capacity has long been rejected by mainstream analysts, but for reasons explained below, it is the organizing principle of ecological footprinting.

Revisiting Human Carrying Capacity: The Roots of Ecofootprint Analysis

Ecologists who study nonhuman species generally define carrying capacity in terms of the average number per unit area of deer, elk, or other species that a particular habitat can support indefinitely (i.e., without permanent deterioration). However, because of such factors as climatic variability and community succession, instantaneous carrying capacity is constantly changing. Carrying capacity is therefore not a particularly rigorous concept, even when applied to nonhuman species with simple stable demands on their ecosystems. It is all the more problematic when applied to humans, whose demands on the environment are anything but simple and stable.

Indeed, many economists and other cornucopians argue that interregional trade alone is virtually enough to cancel concerns about human carrying capacity. Any region or country that can trade services or surpluses of resource A for needed supplies of resource B need not be limited in numbers or economic growth by the theoretical carrying capacity of its home territory. Furthermore, if trade fails, we can always rely on human ingenuity and technology to increase carrying capacity. These are seemingly powerful arguments, and they explain how such densely populated countries as the Netherlands and the UK can remain economically vigorous despite having long exceeded their domestic carrying capacities.

But let us examine the assumptions here. The idea that trade acts to increase local carrying capacity treats each trading region as an isolated open system. Looked at in the aggregate, the situation is cloudier. The world as a whole is a materially closed system. Resources imported to and consumed in region A are no longer available for consumption in the exporting region B and vice versa. Hence, the exchange may result in a one-time increase in the population of each region taken separately, but it also increases global consumption and total pollution, and may deplete local stocks of natural capital, thereby limiting future options. Thus, although the total human load increases, there is no unambiguous increase in total load-bearing capacity.

Indeed, under some circumstances, unfettered trade can lead to a permanent loss of carrying capacity. Global trade exposes pockets of scarce resources everywhere to the largest possible market. This subjects them to ever-greater exploitation pressure, often to the point of depletion or collapse. (Such is the history of trade in fisheries products, e.g.) In short, instead of increasing load-bearing capacity, trade just shuffles it around, increasing local populations, resource depletion, and ecosystems degradation. This ensures that all countries, their economies happily expanding through trade, hit the (now shrinking) limits to growth simultaneously. People everywhere will suffer the consequences of significant climate change, for example.

What about human ingenuity? The general argument here is that cumulative knowledge and technological skills enable humans to squeeze more out of the environment than nature-in-the-raw can provide. This is achieved mainly by continuously increasing resource productivity and by creating substitutes for any resources that are eventually depleted. In these ways, we are able to increase the capacity of any given habitat to support humans and their activities without practical limit.

Let us take these points one at a time. The so-called green revolution, for example, is frequently cited as a miracle of human technological prowess. High-input fossil-fuel-dependent production agriculture has, in fact, greatly increased the short-term productivity (and carrying capacity) of arable land. This seemingly banished the threat of Malthusian famine and certainly facilitated the modern population explosion. However, this miracle was achieved largely by supplementing renewable sun, soil, and rain with nonrenewable fossil fuel, fertilizer, pesticides, and depletable groundwater at a great cost of soil degradation, biodiversity loss, widespread pollution, and falling water tables. The net effect is a swollen human population, likely in overshoot, and increasingly dependent for survival on nonrenewable artificial inputs, even as pollution mounts and natural long-term load-bearing capacity is steadily eroded. The peaking of conventional oil production, diminishing returns to fertilizer, growing water shortages, and globally rising food prices all suggest that we may be entering an era of global food insecurity.

Another vaunted potential fix is our demonstrated ability to greatly increase resource productivity, to do more with less. In theory, more efficient technologies should enable a constant energy and material supply to support a given population at a higher material standard, or a higher population at the same material standard, thereby seeming to increase carrying capacity.

Unfortunately, real-world practice has so far betrayed theoretical promise. The steady gains in material efficiency in the post World War II period have been accompanied by both an exploding population and increasing consumption. In a study of material flows in a selection of the world's most technologically advanced and efficient countries, the World Resources Institute (WRI) showed that, although technology and economic restructuring have substantially weakened the link between economic growth and resource throughput since 1975, there has also been a gradual increase in per capita natural resource use. The study therefore concludes that actual dematerialization has not been achieved (WRI, 2000, p. 19). In fact, efficiency gains may actually work against conservation by lowering prices and raising incomes. This gives people more money to spend on cheaper goods and services so that total throughput – human load – keeps expanding. Also, note that the WRI analysis says nothing about the explosive increase in per capita consumption in developing countries, where most future population and material growth is expected to occur.

Finally, overconfidence in resource substitution is also problematic. It is true that human ingenuity has been markedly successful at finding replacements for simple material inputs to production. For example, optical fiber is as much a product of mind as of nature and has displaced much of the demand for copper wire. However, as ecological economist Daly (1991) has emphasized, there are many more circumstances under which technology and nature are complements, not substitutes. More fish boats cannot for long substitute for fewer fish; more saws and carpenters are no substitute for less lumber. Indeed, the primary products of natural capital (e.g., lumber, steel, aluminum) are actually prerequisites for manufactured capital (e.g., a fish boat) since the latter is made from the former. More generally, we should emphasize that the growing need for technology-based substitutes (e.g., fertilizer) is a warning signal that natural load-bearing factors (e.g., productive soil) are being permanently eroded.

In summary, confidence that expanding trade and technical wizardry have infinitely expanded human carrying capacity and banished the Malthusian specter forever is premature. Carrying capacity is still very much applicable to humans, and given present trends, it is likely to become a major preoccupation of global development policy in the coming decades. However, as this paper shows, we need new approaches to the analysis of carrying capacity, methods that account for uniquely human cultural attributes, including trade and technology.

The Logic and Structure of Ecological Footprint Analysis

Basic Concepts

EFA is one such analytic method. Ecofootprinting starts from the premise that, conscious of it or not, modern human beings are integral components of the ecosystems that support them and are therefore still very much dependent on ecosystem integrity and the land (Rees, 1996; Wackernagel and Rees, 1996).

The method also explicitly recognizes (1) that whether one consumes locally produced products or trade goods, the land connection remains intact, however far removed from the point of consumption some of that land may be, and (2) that no matter how sophisticated our technology, the production/consumption process requires some land- and water-based ecosystems services. Ecofootprint analysis thus incorporates the trade and technology factors simply by inverting the standard carrying capacity ratio: rather than asking what population can be supported by a given area, ecofootprinting estimates how much productive area is needed to support a given population, regardless of the location of the land or the state of technology.

Ecofootprinting builds on traditional trophic ecology by constructing what is, in effect, an elaborate foodweb for the study population. This requires quantifying the material and energy flows supporting the population using production and trade data obtained from national statistical agencies and such international data sources as the Food and Agriculture Organization's Corporate Statistical Database (FAOSTAT). Of course, the human foodweb differs significantly from those of other species. In addition to the material and energy required to satisfy our biological metabolism, a human foodweb must also account for the needs of humanity's industrial metabolism.

Ecofootprinting is further based on the fact that many material and energy flows (resource consumption and waste production) can be converted into a corresponding area of productive land and water ecosystems. Summed up, the total area represents the biocapacity needed to produce the biophysical goods and services used by the study population. Thus, the EF of a specified population is the area of productive land and water ecosystems that the population requires, on a continuous basis, to produce the resources it consumes and to assimilate the wastes it produces, wherever on earth the relevant land/water may be located. A complete ecofootprint analysis would therefore include the population's domestic ecosystem requirements, plus the area it appropriates through net commodity trade, plus its demands on the global common pool for free land- and water-based services (e.g., the carbon sink function).

The area of a given population's ecofootprint depends on four factors: the size of the population, the people's average material standard of living, the productivity of the land/water base, and the technological efficiency of resource harvesting, processing, and use. Regardless of how these factors interact, a population's ecofootprint represents much of its global demand on biocapacity, including ecosystems located half a planet away. We can therefore think of the EF as the globally extended habitat patch ecologically occupied by the study population.

It is critical to recognize that ecofootprints represent ecologically exclusive areas. The productive capacity used by one human population is not available for use by another. Although two or more human groups may share in the output of the same exporting region, the total ecosystem area appropriated is the sum of the areas required by the individual populations. In the final analysis, all human populations are in competition for the available load-bearing capacity of the earth.

The Method in Brief

EF estimates are conceptually simple but data rich. Population EFs are generally based on the final demand for goods and materially intensive services by the referent population. Thus, the first step in calculating the EF of a study population is to compile the total annualized consumption of each significant commodity or consumer good used by that population.

For accuracy, consumption data should be trade-corrected. Thus, the population's consumption of wheat can be represented as follows:

$$\text{domestic consumption}_{\text{wheat}} = \text{domestic production}_{\text{wheat}} + \text{imports}_{\text{wheat}} - \text{exports}_{\text{wheat}}$$

The second step is to convert the consumption of each item into the land area required to produce that item by dividing the total consumption by land productivity or yield. This actually gives us the EF of the individual item. In general: $a_i = c_i/\gamma_i$ where: a_i is the ecofootprint of item i in hectares, c_i is the total consumption of item i in kilograms, and γ_i is the yield of item i in kilograms per hectare. Thus, for wheat:

$$a_{\text{wheat}} = c_{\text{wheat}}/\gamma_{\text{wheat}} = \text{kg}_{\text{wheat}}/(\text{kg}_{\text{wheat}} \times \text{ha}_{\text{wheat}}^{-1})$$

Next, the aggregate EF of the population, (F_p), is determined by summing the footprints for the n individual items:

$$F_p = \sum_{i=1}^{n} a_i$$

Finally, the per capita EF, f_c, is obtained by dividing the total population footprint by population size, N:

$$f_c = F_p/N$$

For some wastes (such as carbon dioxide emissions) or nutrients (such as phosphates and nitrates), it is also possible to calculate the land (or aquatic) ecosystem area required for sustainable assimilation and recycling. In these cases, the assimilation rate per hectare and year is substituted for γ (yield) noted earlier. Most published ecofootprint estimates include an estimate of the carbon sink that would be necessary to assimilate the population's carbon dioxide emissions (but this does not imply an adequate supply of assimilative ecosystems).

Ecofootprint analysts avoid double-counting whenever possible. For example, some consumer products such as leather goods are the byproduct of another industry (such as beef production). In such cases, one would generally count only the primary land requirements (the grazing and grain lands required for feeding cattle). For some purposes, of course, it may be of interest to apportion the estimated ecofootprint among the various products associated with the primary use.

In general, we err on the side of caution in making EF estimates. For example, if there is dispute over, or several estimates of, land productivity, we use the higher estimate (this reduces footprint size). Most EF and biocapacity calculations are therefore likely to under rather than overestimate. **Box 1** describes a common elaboration of basic EFA that involves conversion of primary ecofootprint estimates into their equivalent in hectares of global average productivity (gha). This step simplifies data management and facilitates international comparisons.

Human Ecological Footprints: Implications for Global Sustainability

So, how big are our EFs? Most citizens of industrialized high-income countries feel so distanced from nature that it has

Box 1

In practice, population ecofootprints and biocapacities are usually estimated in terms of global hectares (gha of hectares of world average productivity). In converting ecosystem areas to gha, analysts use yield factors to reflect national differences in land productivity and equivalence factors to account for differences among ecosystems types. For example, if Country A's cropland is twice as productive as world average cropland, and world average cropland is twice as productive world average landtypes, then 1 ha of A's cropland is the equivalent of 4 gha. Conversely, if country A has an estimated per capita cropland ecofootprint of 2 gha, this is equivalent to just 0.5 actual ha of A's domestic biocapacity.

Converting national consumption data to gha simplifies EF estimates because we do not have to identify the sources of trade goods or locations of waste sinks, or determine the productivity and assimilative capacities of the corresponding production/assimilation areas. As important, using a common base yield facilitates comparison among countries and comparisons of individual countries with the global total EF. Using gha is also increasingly realistic for many countries that are heavily dependent on trade flows of commodities from various sources. (Note that the world population ecofootprint estimate in gha would be the same as the global estimate based on actual yields and ecosystem types.)

For some kinds of analyses, of course, it might be necessary or useful to base the EF calculation on actual land/water yields where sources are known and data are available (e.g., see Kissinger and Rees, 2010). Also, to compare a country's or a region's estimated EF in gha with its domestic biocapacity (i.e., to determine that country's ecological deficit or surplus), we must know the actual productivity of various domestic land categories (cropland, pasture land, forests, carbon sinks, etc.).

For example, assume that we want to know whether a given country is potentially food self-sufficient. If that country's cropland is twice as productive as world average cropland, we would multiply the area of its cropland by a yield factor of two before comparing it with (i.e., subtracting it from) its estimated cropland footprint in gha. Note, however, that if the higher productivity of domestic land is the result of intensive fertilizer and pesticide use, any apparent reduction in the ecological deficit due to higher cropland productivity will be at least partially canceled out by the increase in the ecological footprint attributable to the use of agricultural chemicals (fossil fuels are an important feed-stock in the manufacture of these inputs).

At present, the Global Footprint Network estimates both the ecological footprint (demand on ecosystems) and the biocapacity (domestic supply or ability to meet demand) of more than 200 countries and territories. The results, updated annually, use greater than 5000 data points for each country, each year, derived primarily from national statistical accounts and internationally recognized sources such as the United Nations' Food and Agriculture Organization, the International Energy Agency, the UN Statistics Division, and the Intergovernmental Panel on Climate Change. Details of the current methods and results are available from WWF (2008), WWF (2010), and on-line from the Global Footprint Network by following the links at http://www.footprintnetwork.org Lubchenco, 1998.

simply never occurred to them to ask, "How much of the earth's surface is dedicated to supporting just me in the style to which I have become accustomed?" They are generally shocked to learn that average residents of the United Arab Emirates, North America, Australia, Europe, Japan, and other rich countries require the biophysical output of 4–10 gha (10–25 acres) of biophysically productive land and water each to support their consumer lifestyles (Wackernagel *et al.*, 1999; WWF, 2008, 2010). Such large ecofootprints are also sported by members of wealthy urban elites throughout the developing world. Consider these data in light of the fact that there are only 1.8 gha of productive land and water ecosystems per capita (biocapacity) on Earth.

The Functional Footprints of Cities

These findings should alter our perceptions about many things. To begin, they should change how we think about cities and urban land. Over 40 years ago, American ecologist Eugene Odum wrote, "Great cities are planned and grow without any regard for the fact that they are parasites on the countryside which must somehow supply food, water, air, and degrade huge quantities of waste" (Odum, 1971). Ecofootprinting enables us to quantify the extent of this urban parasitism. For example, in 2006, Canada's largest city, Toronto, had a population of approximately 2,503,281 living in an area of 630 km^2. Assuming they are typical Canadians, Toronto's citizens had an average EF of about 7.4 gha (18 acres). Thus, the EF of Toronto was 185,243 km^2 or about 294 times the size of its political area. Clearly, most of the city's supportive ecosystems are located at a great distance from the people they sustain; indeed, they are scattered all over the planet.

This situation is characteristic of high-income cities. In a particularly comprehensive analysis, Folke *et al.* (1997) estimated that the 29 largest cities of Baltic Europe appropriate for their resource consumption and waste assimilation an area of forest, agricultural, marine, and wetland ecosystems 565–1130 times larger than the areas of the cities themselves. A study for the International Institute for Economy and Development in London shows that the ecological demands of that city alone appropriate an area scattered around the world equivalent to the entire biocapacity of the UK.

Such findings have important implications for both urban development and rural sustainability in the twenty-first century. Some analysts interpret the global migration of people from rural areas to the city as implying that modern humans are abandoning the countryside and becoming less dependent on the land. This is an illusion. The reality is that productive croplands, pasture lands, and forests everywhere are being used more intensely than ever to sustain the world's burgeoning urban populations. The human occupants of cities may think of the latter as their principal habitat, but cities per se represent only a tiny fraction of the total increasingly urban-centered human ecosystem.

There is, of course, a certain mutualism between the city and the countryside. Cities need the resources of rural areas and rural areas benefit from urban markets and technology transfers from cities. However, while rural areas could survive without cities, the dependence of cities on rural environments is absolute. In short, there can be no urban sustainability without rural sustainability. In this light, we might even want to reconsider what we think of as urban land – in a whole-systems ecological sense, the great plains of North America, one of the world's major breadbaskets, are an essential component of the increasingly urban global human ecosystem.

According to United Nations' projections, there will be 27 cities with populations exceeding 10 million by 2015 (up from only 1 in 1950). Forty-four more cities will have populations of 5–10 million by the same year. This should trigger at least a cautionary alarm. The world's megacities – all cities for that matter – are dependent on a vastly larger area of productive lands outside their boundaries and political control. But just how secure can any urban population be if the lands represented by its ecofootprint are under threat from climate change or the population's access to essential biocapacity is denied by resource scarcity (think peak oil) or geopolitical turmoil?

Ecofootprints, Global Development, and Social Equity

It is not just cities that overshoot the productive capacity of ecosystems within their political domain. As a result of continuous population growth, rising material demand, globalization and expanding trade, most high-income countries have an ecofootprint several times larger than their domestic biocapacities. In effect, they are running massive ecological deficits with the rest of the world. The EFs of wealthy consumers in particular now roam all over the planet.

Ominously, ecofootprint studies suggest that humanity is now running a global ecological deficit of up to 50% and is accumulating a nonrepayable ecological debt. The global average ecofootprint is 2.7 gha compared with only 1.8 gha per capita of available biocapacity (WWF, 2010). (We can refer to this 1.8 gha as the per capita equitable or fair Earth-share, each person's equal allocation of global biocapacity.) Thus, even at current average levels of economic production and consumption, the human load already exceeds the long-term carrying capacity of the earth – the world's population is living, in part, by depleting essential natural capital and overcharging vital waste sinks. The direct empirical evidence (e.g., climate change, ocean dead zones, tropical deforestation, fisheries collapses) is the stuff of daily headlines.

This means that contrary to the assumptions of prevailing international development models, the so-called first world material lifestyles are not sustainably extendible to the entire world population using prevailing technology. Wealthy consumers, who, on average, require the goods and life support services of 4–10 gha of productive ecosystems to support their consumer life-styles, use greater than two- to five- times their equitable share of global biocapacity. Meanwhile, the great majority of humankind is understandably not satisfied with its present material lot and is determined to improve it. (Despite the industrial revolution in China, in 2007, the EF of the average Chinese was only about 2.2 gha and Bangladeshis were getting by on only 0.6 gha.) The problem is that to support all seven billion members of the human family at the average North American material standard would require over three additional Earth-like planets (and this does not account

for the needs of the two billion or more people to come by the middle of the century).

If we are at or beyond the human load-bearing capacity of the planet, then the only way additional population or material growth can be sustained, without ravaging biodiversity and ultimately destroying the ecological basis of human life, is through massive cutbacks in resource consumption by the presently wealthy. North Americans, for example, should be planning to reduce their EFs by about 80%. This is necessary to vacate the ecological space necessary for justifiable growth in low-income countries. In theory, such cuts could be achieved while actually increasing life quality through a tightly managed new efficiency revolution in rich countries (doing much more with less) combined with an absolute reduction in material demand (i.e., a shift to simpler but more satisfying lifestyles).

Other carrying capacity studies have reached similar conclusions. Using data on current waste discharge rates and the assimilative capacity of biophysical systems, the Sustainable Europe Campaign has estimated that sustainability requires an approximately 50% reduction in the material throughput of the global economy in the coming decades. However, equity considerations demand that developed countries reduce their consumption of various nonrenewable resources by between 88% and 94% (Carley and Spapens, 1999).

These might seem to be extreme conclusions and impossible goals. However, the United Nations Environment Program's *Global Environment Outlook 2000* report also argues the need for a 10-fold reduction in resource consumption by high-income countries. Even the Geneva-based world Business Council for Sustainable Development acknowledged as early as 1993 that "industrialized world reductions in material throughput, energy use, and environmental degradation of greater than 90% will be required by 2040 to meet the needs of a growing world population fairly within the planet's ecological means" (BCSD, 1993).

If so-called postindustrial society fails to meet this challenge, the anticipated doubling and doubling again of human material demand in the next half-century will be disastrous for sustainability. It will certainly negate most present conservation efforts. As explained next, the expanding human EF inevitably reduces the ecological space available for other species.

Human Ecological Footprints and the Biodiversity Crisis

Human beings are uniquely successful among large mammals in mastering the variety of environments on earth and in expanding both numerically and spatially over the globe. Our capacity to extend our aggregate ecofootprint continuously is attributable to several species-specific qualities, of which perhaps three stand out.

First, humans have a remarkably variable diet – we have wide-ranging omnivorous tastes and if we cannot consume something directly (such as grass), we domesticate an animal that will and then consume the animal or its products such as milk. This means that humans may well possess the broadest food niche of any vertebrate on the planet. Second, humans are as behaviorally adaptable as they are catholic in their diets. Together, these two factors make virtually any terrestrial ecosystem on earth accessible to *Homo sapiens*. No other species has managed successfully to colonize and dominate virtually all major terrestrial ecosystem types, from grasslands and forests to deserts and tundra, on all the world's significant land masses. Finally, we are creatures of language, culture, and cumulative learning. Technological advances have enabled humans relentlessly to increase the intensity of resource exploitation and to extend their ecofootprint into virtually all remaining productive habitats on the planet. For example, thanks to electronic fish-finding devices and ruthlessly effective fishing gear, humans have also become the functionally dominant marine mammal. By the mid-1990s, the world's fishers were appropriating the equivalent of 25–35% of the product of net photosynthesis from the 10% of the oceans that produces 96% of the catchable fish.

We previously noted that human ascendance has historically been achieved at the expense of other species. People's material demands, culminating in today's global consumer society, have steadily appropriated an increasingly disproportionate share of the planet's finite productive and sink capacities. The recent human dominance of ocean ecosystems provides several examples of the resulting biodiversity loss.

Certain modern fish-harvesting techniques have devastating impacts on the physical and biological structure of marine habitats, thereby reducing the productivity, diversity, and abundance of both target and nontarget species. Various surveys reveal that fishers now drag heavy trawls over all the oceans' continental shelves at least once every two years and some areas are hit several times in a season. In heavily fished regions of the North Sea, trawlers make up to seven passes annually with nets and chains weighing up to five tons. The total area affected by trawls is 150 times larger than the area of forest clear cut globally each year.

Many fish-stock depletions and the destruction of benthic communities are the direct result of the increasing intensity of commercial fishing. However, other systemic problems stem from the fact that expanding the human marine ecofootprint inevitably diverts an ever greater proportion of the finite energy budget of the sea away from other marine predators, including various fish, birds, and mammals.

A particularly interesting, if still somewhat speculative example of the resultant domino effect, begins with the depletion by human predation of fin and sperm whales in the North Pacific Ocean and southern Bering Sea during the post World War II period through the 1980s (Estes *et al.*, 2004). The weight of evidence suggests that the decline of whale stocks forced mammal-eating orcas (killer whales) to shift a significant part of their diet to pinnipeds – Steller's sea lions and harbor seals. The populations of these mammals in the Aleutian Archipelago declined drastically during the 1970s and 1980s and had been reduced to 10–20% of their former abundance by the 1990s. With the decline of sea lions and seals, the orcas turned to eating sea otters – smaller, less fatty animals – whose populations also declined by 80% during the 1990s.

But the systemic shifts and biodiversity losses do not end there. Sea otters are a keystone species in the shallow marine kelp forests that constitute their preferred habitat. Among the

otters' major prey are sea urchins, which, in turn, graze on kelp. As the otter population declined, there was an explosion of sea urchins, leading to the destruction of the kelp beds by overgrazing. So it is that human overharvesting of whales seems to have precipitated the collapse of one of the most productive inshore-marine ecosystems in the North Pacific, as so-called sea urchin barrens spread along the Aleutian Island chain and the adjacent coast of Alaska. (Dramatic as this example might seem, it is only the final chapter of a longer book. Direct human overharvesting of sea otters for their pelts, beginning in the nineteenth century, had long since ravaged the more southern kelp forests that previously characterized much of the Pacific coast of North America from California north.)

Some popular accounts of this phenomenon have emphasized the role of orcas in switching prey from pinnipeds to sea otters. Certainly, this was an important link in the chain of events, but the most interesting structural change leading to the collapse of the near-shore ecosystem was the human displacement of orcas from their role as the top carnivores in the offshore ecosystem. As human predation on whales forced orcas from their historic place in the foodweb, they were forced to switch to seals and, finally, sea otters (a much inferior food source).

This example underscores how, conscious of it or not, *Homo sapiens* has become the ecologically dominant marine mammal. But the main point to take away is that uncontrolled increases in human appropriations of bioproductivity inevitably result in the rerouting of energy and material flows through the affected ecosystems – marine or terrestrial – with consequent and potentially permanent changes in ecosystem structure and function.

We now recognize that several mechanisms are at work as humans extend their ecofootprints. People:

1. Passively displace other species from their food niches or appropriate their habitats. (Human predation on whales restructured the ecosystems of the northwest Alaska coast; agriculture replaced bison and many other species on the Great Plains of North America; clearing for crops and grazing has destroyed thousands of species in former tropical forests.)
2. Actively eliminate nonhuman competitors – other species that compete with us for our food. (We shoot wolves that hunt ungulates and seals that eat commercially valuable fish; we poison insects that would devour our crops.)
3. Deplete both self-producing and nonrenewable natural capital stocks. (Humans overexploit many wild prey populations from rhinos and tigers to fish; destroy whole ecosystems such as forests; and deplete vital assets such as groundwater, soils, and fossil fuels.)

These processes are all consumption related. The first two are forms of competitive exclusion. Technological man is simply more effective than other organisms at appropriating nature's bounty for our own use. Because energy and material flows used by people are irreversibly unavailable for other species, the latter decline, even to extinction, at least locally.

The third mechanism, stock depletion, is the product of many things, including blind ignorance, material greed, confidence in technological substitution, sheer desperation, and the relentless working of the so-called common property problem on an overcrowded planet. Sometimes, it is even the result of willful disregard on the part of those who give no moral standing to other creatures or who simply do not care about the state or the fate of the world.

Economic globalization, the sanctioning of greed, the rise of consumerism, and the spread of energy-intensive technologies have intensified all these processes. Overharvesting and habitat destruction are driving what some analysts refer to as "the sixth extinction." According to Biologist Wilson, the earth is now experiencing the greatest extinction episode since the natural catastrophes at the end of the Paleozoic and Mesozoic eras. The current species extirpation rate is 100–1000 times prehuman levels inferred from the fossil and paleontological record. A quarter of bird species have already been extinguished by people and fully a quarter of the 4400 mammal species living today are on a path of decline that, if not reversed, is also likely to end in extinction.

The contemporary human ecofootprint is also characterized by massive increases in waste production. Not surprisingly, the resultant pollution is adding to the toll on biodiversity. A poignant example is the phenomenon of "coral bleaching," which has been in the news since the mid-1980s. This effect is produced when the animal component of coral organisms – corals are a mutualistic union of animal polyps and plant (algae) cells – expel their algal cohabitants as a result of thermal or other stress. The loss of algae leaves the coral pale, washed out (bleached), and unable to grow or reproduce. In the exceptionally warm spring and summer of 1998, coral bleaching associated with a 1 °C increase in ocean temperature extended across the tropics, including, for the first time, the Indian Ocean. Reports indicate that 70–90% of corals in the Indian Ocean basin were affected. By 2004, approximately 20% of the world's coral reefs had been destroyed by bleaching and other human impacts and showed no sign of recovery.

Some scientists have referred to the bleaching and die-back of corals as an unprecedented ecological disaster. Coral reefs are among the most diverse and productive ecosystems on the planet. Numerous fisheries, particularly locally important food fisheries, and millions of people are sustained, in part, by the productivity of corals. What economic and social costs would be incurred if the decline of corals and associated reef habitats continues unabated? How will we measure the ultimate consequences of the sheer loss of biodiversity? Most importantly, what does the apparent spread of the problem indicate about the state of the ecosphere in general and the marine environment in particular?

These are important questions, but the main point is that to the extent that coral bleaching is caused by human-induced global warming, it is the indirect result of greenhouse gas accumulation in the atmosphere. (The earth's carbon sinks are overflowing.) The attendant biodiversity loss is therefore the distal product of the industrial world's addictive dependence on fossil fuels and our ever-expanding industrial metabolism. Regrettably, because of lag effects and the sheer inertia of the climate system, the world is committed to additional climate change/global warming for decades to come, even if significant mitigative actions were to be taken today. Certainly, coral bleaching will not be reversed by any

political action feasible under the prevailing global development model.

All these examples underscore the inverse relationship between growth of the human ecofootprint and biodiversity (the competitive exclusion principle at work). Unfortunately, at present, ecofootprint analysis does not directly track biodiversity losses. However, the negative correlation is most directly associated with the conversion of wildlands to agriculture, intensive (including plantation) forestry, and various forms of habitat fragmentation. Many of these conversions are measurable. It may therefore eventually be possible, using the historical record and principles of biogeography, to link some index of biodiversity loss to the growth of population ecofootprints.

Whether or not this relationship can be rigorously defined, the basic point stands. To the extent that further human population and material growth requires the clearing of forests and the intensification of agriculture, and is accompanied by urban sprawl and pollution, it necessarily means less ecological space to sustain nonhuman species and ecosystems. Biodiversity loss is thus virtually certain to accelerate with anticipated growth of the human enterprise.

Assessing the Ecological Footprint Approach

Conceptual and Methodological Strengths

Ecofootprint analysis brings several methodological strengths to bear in assessing the state of the world. First, it is compatible with various concepts developed by other analysts to address human ecological problems. For example, the EF concept

- incorporates and extends George Borgstrom's 1960s notion of "ghost acreage," referring to the extraterritorial food lands required to support densely populated regions and countries.
- corresponds closely to Paul Ehrlich's and John Holdren's 1970s definition of human impact on the environment: I = PAT, where I is impact, P is population, A is affluence, and T is technology. The population EF (F_p) corresponds to impact (I) in the Ehrlich–Holdren formulation and is a function of population size (P) and consumption (C). However, because consumption is a function of income (A) and the state of technology (T), I, like F_p, is also a function of population and consumption. Thus, F_p is an area-based analog of I.
- provides one measure of human load defined by Catton (1980) as a function of population size and average individual impact (consumption again). The larger the footprint, the greater the load. (Catton defined carrying capacity as the "maximum persistently sustainable load.")
- is conceptually related to the embodied energy (emergy) analyses of ecologist Howard Odum and recognizes the importance of the second law of thermodynamics to human affairs as stressed by ecological economists Nicholas Georgescu-Roegen and Herman Daly. Indeed, the area represented by the EF can be conceived as the photosynthetic surface (solar collector) needed to replace the free energy (essergy or negentropy) continuously being dissipated by humans and their industrial metabolism.

However, for all its technical compatibility, perhaps the major strength of ecofootprint analysis is conceptual simplicity and intuitive appeal. First, the "footprint" metaphor seems particularly effective in communicating the idea that we each have an impact on the Earth for which we are responsible through consumption choices. Second, ecofootprinting consolidates real data on a variety of energy and material flows into a single concrete variable, land area. Land is a particularly powerful indicator because it too is readily understood by ordinary people.

Because the supply of land is finite at any scale, ecofootprint analysis provides a measure of human demand (load) for comparison with supply (load bearing capacity) at all geographic levels from local to global. (The ability to compare supply and demand is a major strength of EF analysis shared by few other sustainability indicators.) Such comparisons suggest several secondary indices of (un)sustainability that can be used to stimulate discussion of sustainability options or in establishing measurable policy targets. For example, if we are running a local or a regional ecological deficit, the question becomes: is this a major problem and, if so, what must be done to reduce it? In an era of increasingly ecological and potential geopolitical uncertainty, is it wise for any region or nation to become significantly dependent on distant sources of supply for critical resources? Does a growing ecological deficit suggest that a region or a nation should develop policies to become more self-reliant, by, for example, investing in conservation or restoration of vital natural capital such as arable soils? What does a growing deficit imply for population policy?

Note that national ecodeficits may ultimately prove to be more important than fiscal deficits and yet they are totally ignored in most national accounting systems. Moreover, what about humanity's total ecological deficit that constitutes the global sustainability gap?

Methodological Limitations and Other Criticisms

Although it has considerable conceptual and operational strength, ecofootprint analysis has attracted its share of brickbats and objections. Some criticisms stem from real weaknesses of the method or from misunderstandings of its scope and potential. Others, however, seem more to reflect the critic's discomfort with the findings and implications of EFA than they do flaws in the method. Some of the methodological limitations and common critiques are as follows:

The Question of Scope

Some critics have argued that the term EF is misleading because the method does not capture the full range of ecologically significant impacts on the ecosphere.

It is quite true that ecofootprint calculations are not all-encompassing. They do not account for many pollution effects such as ozone depletion or endocrine (hormone) mimicry, to cite just two examples. Ecofootprinting was designed and is best suited for those human load factors that are readily convertible into a land or water ecosystem equivalent. In fact,

estimates do not even tell the whole land and water story. Various minor consumption categories are yet to be included and analysts are only beginning to examine the spatial implications of waste discharges other than carbon dioxide and essential nutrients.

To be sure, it is unlikely that ecofootprint analysis will ever cope satisfactorily with various toxic waste discharges into the ambient environment. The biological effects of chronic low-level chemical contamination (e.g., carcinogens), for example, do not readily translate into functional land area. However, this may not be as problematic as it seems. Environmental and ecological economists already agree that society should have zero tolerance for highly toxic chemical wastes and radioactive substances for which the ecosphere has no measurable assimilative capacity. Such substances should simply be banned or phased out (as was the case for ozone-depleting CFCs and chlorinated pesticides in much of the world).

The fact that ecofootprint analysis does not incorporate all human ecological impacts merely confirms that no single index can represent every aspect of any phenomenon. It certainly does not invalidate the findings of existing analyses. Rather, it suggests that, despite the alarming magnitude of many national ecofootprints, present calculations almost certainly underestimate actual ecosystem appropriations and that methodological improvements and extensions may well result in even larger ecofootprints. As it is, ecofootprinting provides unambiguous results (e.g., it shows that the world is in overshoot, quantifies the increasingly unsustainable entanglement of nations, reveals the extent of ecological inequity, and challenges material growth as the best way to address poverty) and suggests clear directions for policy if nations and the world are serious about achieving equitable sustainability. Increasing the accuracy and scope of the analyses may add to our sense of urgency but will not likely shift the general direction of needed policy initiatives.

Ecofootprinting is too Simplistic

Some commentators have argued that the ecofootprint concept oversimplifies both nature and society and has little predictive value. It merely generates static estimates of a single aggregate variable.

Again, it is true that footprint analysis is not dynamic modeling and cannot directly project the effects of technological change or social adaptation. However, rigorous modeling and prediction were never the intent. Ecofootprinting provides a simple ecological camera – each analysis is a snapshot of current demands on nature, a portrait of how things stand right now under prevailing technology and social values.

This in itself is important information. Current snapshots indicate that many countries are in a potentially risky ecodeficit, that humanity has already significantly exceeded global carrying capacity, and that some people contribute significantly more to ecological overshoot than do others. This is sufficient to suggest approximately how much we must reduce our consumption, improve our technology, redistribute wealth, or change our behavior to achieve sustainability.

Of course, once an initial baseline has been established, subsequent ecofootprint assessments can contribute to time-series studies – repeated snapshots over years or decades – to monitor progress toward, for example, reducing national ecological deficits or closing the global sustainability gap. (After all, even a motion picture is a series of snapshots.) As shown by the Global Footprint Network's annual updates of the ecofootprints of most of the world's nations, such sequential analyses are capable of demonstrating the actual effects of new technologies, significant economic policy initiatives, changes in income or population, and shifts in consumer behavior or cultural values. We can also assess the potential impacts of alternative technologies, policy changes, or settlement patterns on population ecofootprints using simulation studies. If we know how some technological innovation or policy initiative is likely to affect the consumption of critical commodities, the standard calculations can be used to address such questions as: how would a shift from fossil fuels to biomass fuels or solar energy affect a population's ecofootprint? Would a general shift to vegetarianism or organic foods reduce our food footprint?

To summarize this point, EF analysis is not intended to provide a dynamic window on the future but rather a snapshot in time. As such, it can both help to assess current reality and to test alternative what-if scenarios on the road to sustainability.

On a related matter, some critics have stressed that while both the ecosphere and the economy are complex dynamic systems, ecofootprint analysis ignores the wide-ranging implications of complexity theory. This comment, while true, misses two points. First, ecofootprinting was not intended to provide a comprehensive picture of socioecological systems. It was created to quantify just one key indicator of humanity's engagement with the rest of nature – our dependence on ecosystems services – heretofore largely ignored in sustainability policy debates. To the extent that this index is reliable and has policy relevance to sustainability, it can be used in conjunction with economic, social, or other indicators in a comprehensive approach to the issues at hand.

Second, although science has certainly gained important insights from complex systems theory – the behavior of dynamical, self-organizing, adaptive systems often confounds standard forecasts and derails simplistic approaches to ecosystems management – the fact remains that simple models are adequate for many phenomena. Under these circumstances, Occam's razor still applies: there is no need to derive a complex explanation where a simple one suffices.

For example, one can hardly imagine a better example of a complex human-dominated system than the world's great cities. Cities are a marvel of biophysical, economic, political, social, and cultural subsystems that somehow come together to function as a complex, dynamic, integrated, self-organizing whole. Nevertheless, if any city as currently conceived were enclosed in an impermeable glass dome whose base coincided with its political or geographic boundaries, it would rapidly disintegrate.

The explanation is simple – the city would simultaneously be starved of resources and forced to suffocate on its own wastes. This is because most of the ecosystem functions and biophysical processes supporting the city lie outside our hypothetical bell jar. Ecofootprint analysis measures the city's (region's, country's, etc.) dependence on nature's services and estimates the de facto ecosystem area needed to produce them. (Remarkably, most studies of cities-as-systems still ignore this

critical biophysical dimension.) By accounting for the city's full metabolic demand, ecofootprint analysis shows that the built-up core – the nominal city – is only a tiny component of the total system. Cities are dissipative structures totally dependent on the biocapacity of a vast, increasingly global hinterland hundreds of times the size of the city itself. The consumptive core and the productive hinterland together constitute the complete urban ecosystem. There is certainly room to debate the methods used to make such EF estimates. However, we do not need to know anything about city governance or the complex relationships among its physical and socioeconomic subsystems to predict the collapse of our city-in-a-jar, to quantify its dependence on distant ecosystems, or to assess its vulnerability to global change.

Is the Carbon (Energy) Footprint a Real Land Use?

For many high-income industrialized countries, the fossil energy or carbon footprint may make up half of the total ecofootprint. Modern society makes heavy demands on available carbon sinks. However, because science cannot account fully for the global carbon budget and there is dispute over the relative roles of terrestrial and marine ecosystems in carbon sequestration, some critics see energy footprint calculations as unreliable or hypothetical. Others do not regard carbon assimilation as a real land use. Still others, mostly techno-optimists, suggest that existing ecofootprint estimates exaggerate the impact of prevailing energy use while not taking renewable energy into account.

The first criticisms reflect our cultural bias that land use requires physical occupancy or the consumption of some tangible product of the land. People are unaccustomed to thinking of the provision of other (particularly nonmarket) biophysical services of nature, no matter how essential to sustainability, as effective uses of landscape. From the perspective of thermodynamic law and mass balance, it is this cultural bias that needs adjustment, not ecofootprint theory. All energy and material use is throughput, implying the generation of wastes, and to the extent that the assimilation and recycling of those wastes (including carbon dioxide) requires an exclusive dedication of land or water, it is a legitimate component of ecofootprint calculations.

As it happens, carbon dioxide is the greatest waste output by weight of industrial economies. Also, although the details of the carbon budget may be in dispute, there is no question that: (1) carbon dioxide levels are increasing in the atmosphere and that (2) these increases represent about half of current carbon emissions from fossil fuel combustion. This implies that available land and water carbon sinks are insufficient to sequester all anthropogenic carbon dioxide at current rates of emission. In ecofootprint terms, humanity's demand for natural sequestration services exceeds available supply. We are running a global carbon sink deficit in exactly the same sense that many countries run a food-land deficit.

All of which is to say that the notion of carbon sink lands is not hypothetical even if the real thing is in short supply. Energy footprint calculations therefore estimate the area of dedicated carbon-sink forests that would be necessary to assimilate a study population's carbon emissions, assuming average estimated carbon sequestration rates of the world's growing forests. It is precisely such thinking that has led various electric utilities to plant dedicated carbon sink forests to offset the emissions from their fossil fuel-burning generating plants. Similarly, some countries have contemplated dedicated carbon sink forests as an alternative to reducing their carbon emissions or as a way to generate credits in regional or future global carbon trading schemes. (Carbon trading in various forms has become commonplace.)

Having established the legitimacy of carbon sink estimates, we should note that a dedicated carbon sink forest is an exclusive land use insofar as it cannot be harvested for any purpose such as pulp and paper production that would quickly release sequestered carbon back into the atmosphere.

Finally, we should recognize that any alternative to fossil fuels will also generate an energy ecofootprint and that the effect of any wholesale shift to renewable energy technologies will show up in subsequent ecofootprint estimates. Techno-optimists should note that the new energy footprints will not necessarily be smaller than today's carbon footprint. For example, consider the ecofootprint of corn-based ethanol production, which has become a major industry in the US. The ethanol footprint includes not only a fuel-crop growing area larger than the carbon sink area required to sequester the carbon emissions from an equivalent amount of fossil fuel but also a de facto carbon sink to assimilate the carbon emitted by the various fossil-fuel based technologies (mechanized cropping, diesel irrigation, transportation, biomass distillation, etc.) used to make the ethanol. Various studies suggest that corn ethanol production is a net energy sink or only a marginally positive energy source – it takes as much or more fossil energy to produce the ethanol as is contained in the product – hence, the carbon footprint of a gallon of corn ethanol is more or less the same as that of the fossil fuel the ethanol was supposed to displace. In total, therefore, aggregate energy ecofootprint of the so-called renewable corn ethanol is much larger than the carbon ecofootprint from burning an equivalent quantity of fossil energy. So much for sustainability. The main point is that all energy sources have an ecofootprint, and when subjected to full life-cycle analysis, at least some renewable energy sources may not perform any better than fossil fuels in ecofootprint terms.

Ecofootprinting is Antitrade and Antigrowth

Some critics, particularly neoliberal economists, suggest that ecofootprint analysis is biased against cities, trade, and globalization. Others argue that the method is neo-Malthusian because it raises concerns about carrying capacity and suggests that there are limits to growth. In fact, however, ecofootprint analysis is merely a number-crunching tool whose output is subject to alternative interpretations and policy responses. The numerical results do not, in themselves, reflect an antitrade, antiurban, or antigrowth bias.

That said, there can be little question that many sustainability analysts are alarmed by the findings of ecofootprint studies. The results show unambiguously the biophysical dependence of densely populated, high-consuming regions/countries on other countries and the global commons; they reveal that many countries are running significant ecological deficits and that the world as-a-whole is in overshoot. Given that we live in an era of accelerating global change and uncertain geopolitics, it seems reasonable to conclude that trade

dependence carries increasing risks and that greater regional self-reliance might be a good thing. Similarly, is rampant overshoot not sufficient grounds to question global society's so far unwavering commitment to continuous economic growth?

The problem is that all such reasoning is anathema to unreformed economists and other expansionists who favor unrestrained growth facilitated by trade and for whom carrying capacity is a nonissue. Many critics of ecofootprinting thus seem to reject the method, not on theoretical or methodological grounds, but rather because of the questions it raises and because logical answers to these questions confront their own disciplinary or ideological biases. By reopening the debate on carrying capacity and exposing the downside of globalization and trade, ecofootprint analysis challenges the core developmental economic ideology of our time.

Epilog

Whatever its weaknesses, the EF concept is strong enough to have captured both the scientific and the popular imaginations. It is now being applied as an assessment tool by nongovernmental organizations and government agencies in towns, cities, and countries around the world. A small industry of consulting firms has sprung up to develop and commercialize various products based on the ecofootprint concept.

Ecofootprinting seems to concretize the human ecological crisis in a way other energy and material flows studies do not. The ability to compare the demand for biocapacity with supply is an important factor. Knowing our EF not only makes us conscious of our personal contribution to global change but also makes us responsible for doing something about it. Pluralistic democracies can work only if the public are well informed about key policy issues and have a keen sense of their capacity to make a difference in improving the human prospect and those of other species. To the extent that EF analysis helps to generate a more ecologically literate and politically active populace, it will have achieved its most important objective.

See also: Carrying Capacity, Concept of. Energy Flow and Ecosystems. Energy Use, Human. Human Impact on Biodiversity, Overview. Loss of Biodiversity, Overview. Sustainability and Biodiversity. Urban–Suburban Biodiversity

References

Arrow K, Bolin B, Costanza R, *et al.* (1996) Economic growth, carrying capacity and the environment. *Ecological Applications* 6(1): 13–15.

Berger PL and Luckmann T (1966) *The Social Construction of Reality.* Garden City, NY: Doubleday.

BCSD (1993) Getting eco-efficient *Report of the BCSD First Antwerp Eco-Efficiency Workshop*. Geneva: Business Council for Sustainable Development.

Block W (1990) Environmental problems, private property rights solutions. In: Block W (ed.) *Economics and the Environment: A Reconciliation*, pp. 281–332. Vancouver, BC: The Fraser Institute.

Carley M and Spapens P (1999) *Sharing the World: Sustainable Living and Global Equity in the 21st Century.* London: Earthscan Publications.

Catton WR (1980) *Overshoot: The Ecological Basis of Revolutionary Change.* Urbana and Chicago: University of Illinois Press.

Daly H (1991) *Steady-State Economics*. Washington, DC: Island Press.

Diamond J (1992) *The Third Chimpanzee*. New York: HarperCollins.

Duncan RC (1993) The life-expectancy of industrial civilization: The decline to global equilibrium. *Population and Environment* 14: 325–357.

Estes JA, Danner EM, Doak DF, *et al.* (2004) Complex trophic interactions in kelp forest ecosystems. *Bulletin of Marine Science* 74(3): 621–638.

Flannery TF (1994) *The Future Eaters: An Ecological History of the Australasian Lands and Peoples*. Chatsworth, NSW: Reed Books.

Folke C, Jansson A, Larsson J, and Costanza R (1997) Ecosystem appropriation by cities. *Ambio* 26: 167–172.

Gordon RL (1994) Energy, exhaustion, environmentalism, and etatism. *The Energy Journal* 15: 1–16.

Hansen J, Ruedy R, Sato M, and Lo K (2010) Global surface temperature change. *Reviews of Geophysics* 48: 29, RG4004.http://dx.doi.org/10.1029/2010RG000345.

Hern WM (1997) Is human culture oncogenic for uncontrolled population growth and ecological destruction? *Human Evolution* 12: 97–105.

IEA (2010) *World Energy Outlook 2010 (Executive Summary).* Paris: International Energy Agency.

Kay JJ and Regier H (2000) Uncertainty, complexity, and ecological integrity. In: Crabbé P, Holland A, Ryszkowski L, and Westra L (eds.) *Implementing Ecological Integrity: Restoring Regional and Global Environment and Human Health, NATO Science Series IV: Earth and Environmental Sciences*, vol 1. pp. 121–156. Dortrecht: Kluwer Academic Publishers.

Kissinger M and Rees WE (2010) Importing terrestrial biocapacity: The US case and global implications. *Land Use Policy* 27: 589–599.

Lomborg B (2001) *The Skeptical Environmentalist: Measuring the Real State of the World*. Cambridge, UK: Cambridge University Press.

Lubchenco J (1998) Entering the century of the environment: A new social contract for science. *Science* 297: 491–497.

NOAA (2011) *Trends in Atmospheric Carbon Dioxide*. Earth System Research Laboratory, National Oceanic and Atmospheric Administration. http://www.esrl.noaa.gov/gmd/ccgg/trends/ (accessed 20 April 2011).

Odum EP (1971) *Fundamentals of Ecology*, 3rd edn. Philadelphia: Saunders.

Ponting C (1991) *A Green History of the World*. London: Sinclair-Stevenson.

Prigogine I (1997) *The End of Certainty: Time, Chaos and the New Laws of Nature*. New York: The Free Press.

Rees WE (1996) Revisiting carrying capacity: Area-based indicators of sustainability. *Population and Environment* 17: 195–215.

Schneider ED and Kay JJ (1995) Order from disorder: The thermodynamics of complexity in biology. In: Murphy MP and O'Neill LAJ (eds.) *What is Life: The Next Fifty Years – Reflections on the Future of Biology.* Cambridge: Cambridge University Press.

Simon J (1995) The state of humanity: Steadily improving. *Cato Policy Report 17.* Washington, DC: The Cato Institute, p. 5.

Smith JW and Sauer-Thompson G (1998) Civilization's wake: Ecology, economics and the roots of environmental destruction and neglect. *Population and Environment* 19: 541–575.

Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of earth's ecosystems. *Science* 277: 494–499.

Wackernagel M, Onisto L, Bello P, *et al.* (1999) National natural capital accounting with the ecological footprint concept. *Ecological Economics* 29: 375–390.

Wackernagel M and Rees WE (1996) *Our Ecological Footprint: Reducing Human Impact on the Earth*. Gabriola Island, BC/Philadelphia, PA: New Society.

WRI (2000) *The Weight of Nations: Material Outflows from Industrial Economies*. Washington, D.C: World Resources Institute.

WWF (2008) *Living Planet Report 2008*. Gland, Switzerland: World Wide Fund for Nature.

WWF (2010) *Living Planet Report 2010*. Gland, Switzerland: World Wide Fund for Nature

Ecological Genetics

Beate Nürnberger, The University of Edinburgh, Edinburgh, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Directional selection Natural selection that favors phenotypes that differ from the current mean phenotype of a population in one direction (e.g., those that are larger than the current mean).

Disruptive selection Natural selection that favors phenotypes that deviate in either direction from the current population mean.

Dominance Phenotypic effect of a particular heterozygous combination of two alleles at a single locus that deviates from the mean of the two homozygous phenotypes.

Epistasis Phenotypic effects of particular combinations of alleles at two or more loci that differ from the sum of additive effects of these alleles (i.e., the analog of dominance for more than one locus).

Gene flow The influx of genetic variants into a local gene pool as consequence of immigration of individuals or gametes (i.e., pollen).

Genetic drift Stochastic change in allele frequencies at a given locus from generation t to generation $t+1$ due to the chance events that affect the number of offspring produced per parent in generation t. The effect of genetic drift on allele frequencies and, by extension, on population mean phenotypes increases with decreasing population size.

Genetic linkage map For a given taxon, a map of genetic marker loci that are arrayed based on the recombination rates among them. Recombination rates are estimated from crossing-over events in controlled crosses. Given a sufficiently high marker density, the resulting linkage groups correspond to chromosomes.

Polymorphism The co-existence of two or more distinct variants in a population either in the form of alleles for a given locus or as distinct phenotypes.

Stabilizing selection Natural selection against phenotypes that deviate from the current population mean.

Introduction

Historical Sketch

Ecological genetics is a multidisciplinary endeavor with the aim to understand phenotypic adaptation in response to specific selective agents in a particular ecological setting and in terms of the underlying genetic changes. While this definition has remained largely constant over the past 50 years since the publication of the first treatise on the topic (Ford, 1964), there have been marked shifts in the kinds of traits under study and in the time scales of adaptation that have been considered. Given the available methodology at the time, Ford (1964) presented mainly studies of strongly selected visible polymorphisms with a simple genetic basis, such as banding patterns in snails, cryptic coloration in moths, mimicry in butterflies, and heterostyly in primroses. He concluded that a wealth of evidence "forces us completely to readjust our thoughts on evolution and to recognize that a population may adapt itself very rapidly to changed conditions" (p. 296). This was in contrast to the prevailing view that evolutionary change and ecological interactions took place on different time scales and that species could therefore be viewed as fixed entities from an ecological perspective. However, this view was to change gradually over the next decade (Brussard, 1978).

Antonovics (1976) eloquently advocated a "new ecological genetics" that abandoned the distinction between ecological and evolutionary time scales and viewed biological communities "as consisting of evolving components" (p. 239). The scope for empirical tests of this assertion was greatly enlarged with the introduction of quantitative genetics to the study of evolution in natural populations (Lande, 1979, 1982; Lande and Arnold, 1983). Continuously varying phenotypes including life history traits such as rates of growth and development and schedules of birth and death became amenable to evolutionary analysis. Finally, Endler's (1986) comprehensive review of empirical studies proved that directly observable adaptive responses to natural selection in the wild were not uncommon.

Since then, empirical studies of contemporary adaptive evolution have been published at an increasing rate and with ever finer resolution of the contributing processes. Thanks to the recent availability of abundant, highly variable genetic markers, phenotypic change can be analyzed in the context of a known genetic population structure that is inferred from the distribution of neutral marker alleles. Moreover, marker-based linkage maps can be used to study the genomic distribution of genetic effects ($\approx$loci) that underlie a continuously varying trait. The identification of such effects can serve as a first step toward pinpointing the molecular basis of a given adaptive response. Moreover, genetic markers allow for the construction of pedigrees in intensively studied natural populations. Such pedigrees are prerequisites for the analysis of phenotypic traits with the so-called animal model (Kruuk, 2004, Kruuk *et al.*, 2008). Over the past decade, the application of this model to studies of natural populations has transformed the analysis of microevolutionary processes. It represents a powerful statistical tool with which to dissect the complexities of phenotypic trait variation that lie at the very heart of ecological genetics. Additive genetic variation, the key component of inheritance, can be estimated alongside maternal effects, effects of sex, age, spatial heterogeneity, or temporal trends (see Wilson *et al.*, 2010 for a concise, nontechnical introduction).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00038-1

Scope and Outline of the Article

This article takes a microevolutionary perspective on adaptive phenotypic evolution, i.e., it concentrates on processes that are observable on a maximum time interval of roughly 100 years (Thompson, 1998) and mostly within the time frame of a researcher's career. Whenever possible examples are drawn from long-term field studies, because they offer priceless insight into the temporal dynamics of ongoing evolutionary change and allow for a dissection of multiple cooccurring processes. Given the focus on adaptive change brought about by either directional or disruptive selection, any unqualified use of the term selection refers implicitly to these two modes. However, it is worth keeping in mind that stabilizing selection also plays an important role in natural populations. Furthermore, the article focusses on microevolutionary change within populations. It can often be the starting point of regional differentiation and may ultimately contribute to the formation of new species. These processes are considered in other articles of this encyclopedia. Finally, this article is about concepts and insights from empirical studies rather than about methodological issues such as the estimation of allele frequencies, gene flow rates, or heritabilities. These are discussed in detail in several text books (Falconer and Mackay, 1986; Lynch and Walsh, 1998; Conner and Hartl, 2004; Lowe *et al.*, 2004).

The first three sections lay out the building blocks of any microevolutionary analysis: the amount of heritable genetic variation in the target species, the strength of natural selection and the rate and pattern of any adaptive response. Next, genetic correlations are discussed as sources of trade-offs that can limit phenotypic evolution. Strategies of and limits to microevolution in heterogeneous environments are the themes of the sections on phenotypic plasticity and local adaptation, respectively. The final two sections present insights from the genetic dissection of phenotypic traits and from analyses of the wider ecological implications of contemporary adaptive evolution.

Genetic Variation for Phenotypic Traits in Natural Populations

Every phenotypic trait results from a complex interplay of many gene products that influence its mature expression over the course of ontogeny. A mutation at any one of the underlying loci can in principle alter a trait. Yet only those genetic loci that are polymorphic at any one time contribute to the trait's heritable variation. In the context of any (micro-) evolutionary analysis, the focus is thus on the genetic basis of phenotypic variation of a given trait, not on the full complement of genes that are involved in the trait's expression. This implies that the inheritance of a trait either as a distinct polymorphism or as a continuous spectrum of phenotypes is partly a function of the number and effects of genetic variants that segregate in a population at any one time. This genetic architecture may change if, for example, an allele of large effect is fixed by selection.

If discrete phenotypic variation is based on polymorphisms at one or two loci, then crossing experiments can in principle map trait values to alleles and the heritable variance in the trait follows directly from the allele frequencies in the population. More typically, variation in a trait is based on many polymorphic loci and also reflects substantial environmental influences. Quantitative genetics provides the necessary framework in which to analyze such traits (see **Box 1** for a brief summary of the basic model). Over the past 30 years, quantitative genetic studies have demonstrated the existence of substantial additive and thus heritable genetic variation for

Box 1 Quantitative genetics

Quantitative genetics represents a statistical analysis of the resemblance among relatives to uncover the causal components of phenotypic variation and to predict phenotypic responses to natural or artificial selection either for single or for multiple, correlated traits. At its core is a model that assumes the existence of a large number of polymorphic loci with alleles of small and mostly additive effect. The particular set of alleles carried by an individual causes its phenotype P to deviate from the population mean phenotype μ by a certain amount. This deviation is called the individual's genotypic value G and has a population mean of zero. Any nongenetic influences on the phenotype are modeled similarly as the environmental deviation E, which is assumed to be normally distributed with mean zero. Consequently, an individual's phenotypic value P is defined as

$$P = \mu + G + E$$

The summation over many small genetic effects and the assumptions about E ensure that the resulting distributions of P, G, and E are normally distributed. This framework thus allows for an analysis of phenotypic variation in terms of a small number of parameters (i.e., means and variances) without any knowledge of the underlying genetic detail. Empirical support for this model comes partly from the common observation of normally distributed phenotypes in nature and from the immense utility of quantitative genetics for plant and animal breeders.

Clearly, gene action is not necessarily additive. At a single locus, one allele may be partly or entirely dominant over another and there may be epistatic interactions among alleles at two or more loci. Under these conditions, the effect of an allele depends on the genetic background in which it occurs. Nonadditive gene action is incorporated into the model by partitioning the genotypic value into an additive component A, a dominance deviation D and a deviation due to epistatic interactions, I:

$$G = A + D + I$$

The additive effect of an allele is defined accordingly as the average deviation from the population mean phenotype caused by a single copy of that allele. Just as the overall mean phenotype depends on the allele frequencies at all underlying loci, the average effect of an allele depends on the frequency distribution of genetic backgrounds in the population. It is thus affected by nonadditive gene action and may vary from one population to another. For a given phenotype, the sum of all additive allelic effects is A.

Because genotypes are broken up and reshuffled in the process of sexual reproduction, dominance and epistatic effects on parental phenotypes are not passed on to the offspring. The laws of inheritance thus revolve around the additive genetic variation. The additive component A of an individual's phenotype is defined as its breeding value. It predicts the mean phenotype of its offspring and is therefore in principle measurable. Finally, the proportion of the total phenotypic variation V_P that is due to additive genetic variance V_A (or, equivalently, to the variance in breeding values) is defined as the trait's heritability h^2:

$$h^2 = V_A / V_P$$

Falconer and Mackay (1996) and Lynch and Walsh (1998) present detailed treatments and extensions of this model.

Box 2 Estimating the strength of selection on quantitative traits

In its original univariate form, the so-called breeder's equation predicts the response of a trait to selection given estimates of the trait's heritability h^2 (cf. **Box 1**) and the strength and direction of selection on that trait:

$$R = h^2 S$$

where the selection differential S is defined as the difference between the mean phenotype before selection, μ, and that after selection within one generation. Equivalently, S represents the phenotypic covariance between the trait and relative fitness, where relative fitness is distributed with mean 1. The selection response R is the difference between μ and the mean phenotype in the following generation. This model can predict R provided that selection acts directly on the trait under consideration (z_1). It will fail, however, if the shift in the mean of z_1 in the parent generation results from a merely phenotypic covariance with a second trait z_2 that is in fact under selection. In this situation, there can be no selection response in z_1 because there is no additive genetic covariance between both traits that would reflect a (partially) shared genetic basis.

The example points to the more fundamental issue that individual traits do not evolve in isolation. Particularly, additive genetic covariation in a number of traits implies that selection on any one trait should produce a correlated response in the others. To determine which traits actually cause variation in fitness and to predict the responses to direct and indirect selection, a broader modeling framework is needed.

To this end, Lande and Arnold (1983) derived a multivariate analog of the breeder's equation. Recall that the heritability is the quotient of additive genetic variance over phenotypic variance, V_A/V_P. Each of these terms is now replaced by a matrix, which contains for the set of measured traits the appropriate variances (diagonal elements) and pairwise covariances (off-diagonal elements). Of particular interest is the additive genetic matrix **G**, which contains the variances and covariances of breeding values (see **Box 1**). It encapsulates not just the per-trait amount of genetic variation on which selection can act but also the interdependencies among traits that are due to a shared genetic basis and that may either enhance or constrain the selection response of any one trait. The matrix **P** is the phenotypic analog of **G**. The multivariate breeder's equation is thus

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S}$$

where **R** and **S** are vectors of selection responses and differentials, respectively. This equation can be rewritten as

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$$

where $\boldsymbol{\beta}$ is a vector of selection gradients which express the strength of direct selection on each trait. Selection gradients can be estimated as partial regression coefficients when relative fitness (or a component thereof) is regressed on the set of measured traits. They quantify the strength of directional selection on a given trait when all other traits are held constant. Quadratic terms to estimate either stabilizing or disruptive selection can be added to the regression model. Note however, that these are not easily incorporated into the breeder's equation.

While the consideration of multiple traits should generate much improved estimates of selection strengths and responses, the breeder's equation typically fails as a predictive tool when applied to natural populations. Violations of the model assumptions (Lande and Arnold, 1983) are partly responsible for this. Phenotypes, breeding values and environmental deviations are assumed to follow a multivariate normal distribution (or can be transformed into such). Highly correlated traits mean that the inverse matrix $\mathbf{P}^{-1}$ cannot be found, selection gradients cannot be estimated and the dimensionality of the model needs to be reduced (e.g., through the use of principal components). Moreover, estimates of selection strengths and responses will still be in error to the extent that additional pertinent traits are missing from the analysis. Finally, the variance–covariance matrices apply to the base population before selection acts. Different environmental conditions in the offspring generation may alter these matrices to the extent that **R** cannot be correctly predicted.

Endler (1986) and Mitchell-Olds and Shaw (1987) discussed a range of statistical issues that follow from these assumptions. Moreover, since "missing traits" can lead to erroneous inference of selection on one or more of the measured traits, these authors argued that the results of the regression analysis should be validated in the context of a longer term research program that includes detailed field observation and, ideally, experimental manipulation. Without this context, estimates of selection gradients are in Endler's words (1986, p. 201) akin to "refined alchemy." The problem of causation and the common discrepancy between predicted adaptive change and observed stasis have recently prompted a more fundamental critique of the current use of the breeder's equation in observational studies of natural populations. Morrissey et al. (2010) specified the condition under which the breeder's equation will arrive at correct inferences of causal factors and selection responses and provided examples of commonly encountered situations in nature that violate this condition. The authors advocated an alternative approach to predict selection responses based on the Robertson–Price identity (Robertson 1966; Price, 1970).

practically any trait that has been studied (Mousseau and Roff, 1987; Roff, 1997). However, there are a few notable exceptions (Blows and Hoffmann, 2005): low heritabilities were found, for example, for mass-specific metabolic rates in mammals and there are a small number of cases in which artificial selection experiments did not produce a phenotypic response.

Estimates of heritability (**Box 1**) are typically used to predict the response of a trait to selection. According to the so-called breeder's equation, this response is equal to the product of heritability and selection strength (**Box 2**). When different types of traits are compared, morphological characters have the highest and life history traits the lowest heritabilities (means: 0.461 and 0.262, respectively, Mousseau and Roff, 1987). This discrepancy between morphology and traits closely correlated with fitness is widely observed not just under controlled laboratory conditions but also in wild, unmanipulated populations (e.g., Merilä and Sheldon, 2000). It agrees with the expectation that persistent selection toward an optimum should erode additive genetic variance in fitness and in traits associated with it. However, this interpretation is not supported by more detailed analyses. Instead, the lower heritability in fitness-related traits is primarily due to greatly increased residual variation (V_R), which subsumes non-additive and environmental effects (Houle, 1992; Campbell, 1997; Merilä and Sheldon, 2000; Coltman *et al.*, 2005; Teplitsky *et al.*, 2009). The level of additive genetic variation may either show no correlation with the fitness association of a trait (Campbell, 1997; Kruuk *et al.*, 2000; McCleery *et al.*,

2004; Coltman *et al.*, 2005) or even increase with it (Houle, 1992; Merilä and Sheldon, 2000, but see Teplitsky *et al.*, 2009 for a counter example).

These patterns may derive from the particular nature of fitness itself, a composite trait that is expressed late in life. Compared to ordinary traits, it should reflect the cumulative environmental variation over an individual's lifespan (increased V_E) as well as genetic variation at a larger number of loci. It might thus represent a large "mutational target" (Houle *et al.*, 1996, increased V_A). Moreover, low genetic variation in total fitness does not rule out substantial additive genetic variation in fitness components, such as survival or fecundity, if there are negative genetic correlations among them (Charlesworth, 1987, *see* Genetic Correlations).

The heritability of a given trait can vary as a function of environmental conditions. In comparisons between benign and stressful conditions, studies of size-related traits in *Drosophila* initially suggested that harsh conditions caused an increase in heritability (Hoffmann and Parsons, 1991) but analogous studies in birds showed the opposite trend (Hoffmann and Merilä, 1999). More recent detailed analyses of genetic variance components across a range of animal taxa have corroborated the pattern seen in birds (Charmantier and Garant, 2005): overall, stressful conditions reduce the heritability of a trait. The effect is stronger in size related as opposed to fitness-associated traits and can be due to a reduction of V_A, an increase in V_E, or a combination of both. To explain these effects, it has been suggested that harsh conditions (1) prevent individuals from reaching their genetic potential and (2) create increased local variation in habitat quality (Hoffmann and Merilä, 1999; Charmantier and Garant, 2005). More generally, these studies illustrate that the ecological context plays an important role in the phenotypic manifestation of genetic variation (Charmantier and Garant, 2005; Ellegren and Sheldon, 2008).

The Strength of Natural Selection: Point Estimates

Selection on Discrete Polymorphisms: An Example

In the case of discrete polymorphisms, the strength of directional selection on a given phenotype can be measured from its rate of increase or decrease in frequency. Because genotypes can be inferred from phenotypes in this case, there is no need to distinguish between the strength of selection and the selection response, i.e., the shift in mean phenotypes from one generation to the next. For example, the melanic morph of the peppered moth (*Biston betularia*) reached local frequencies close to one in Britain only 50 years after it was first observed, which suggests a 30% higher fitness relative to the light-colored wildtype (Haldane, 1924). The fate of the dark morph which is tightly linked to levels of air pollution has now turned and its rate of decline implies a selective disadvantage of similar but somewhat lower magnitude (Cook and Turner, 2008). This classic case of "industrial melanism" continues to be a textbook example of evolution in action (Majerus, 2009).

A Review of Selection Gradients

In continuously varying traits, the strength of directional selection is estimated from a linear regression of a measure of fitness, such as survival or fecundity, on individual phenotypes. The slope of this regression estimates the selection differential or, if several traits are analyzed simultaneously, the selection gradient for a given trait (see **Box 2**).

An influential review of published selection gradients (Kingsolver *et al.*, 2001) found that their absolute values followed an exponential distribution: most estimates were small, but moderate to large values were also observed. The change in relative fitness per phenotypic standard deviation (SD) had an overall mean of 0.22 and a median of 0.16. About 13% of the estimates were greater than 0.5. Note that the median estimate implies substantial fitness variance, because a trait may vary over several SDs within a population (Hereford *et al.*, 2004). Moreover, assuming a heritability of 0.5, sustained directional selection of this strength would cause the trait mean to shift by 3 SDs within 50 generations (Kingsolver and Pfennig, 2007).

The review also showed that selection was on average stronger on morphology than on life history or phenological traits. However, this need not reflect a true biological pattern other than that morphology is easier to estimate accurately. In studies that spanned months or years (as opposed to days), larger estimates were observed when fitness was measured in terms of fecundity or mating success as opposed to survival (Hoekstra *et al.*, 2001). Finally Kingsolver *et al.* (2001) also estimated the strength of stabilizing versus disruptive selection and found that these modes were detected with equal frequency. This observation challenges the commonly held view that stabilizing selection should be the rule and disruptive selection the exception.

Potential Biases in Estimated Selection Strengths

Kingsolver *et al.* (2001) discuss a number of reasons why the estimates of linear selection gradients may be biased upward. They are probably not a random sample of selection strengths in nature, because researchers focus on study systems in which they suspect selection to be operating. The majority of studies had small sample sizes which may have introduced biases (Hersch and Philipps, 2004, but see Knapczyk and Conner, 2007). Finally, almost all studies measured fitness components (survival, mating success, or fecundity) only. Their less than perfect correlation with total fitness necessitates a downward adjustment of the selection estimates (Hereford *et al.*, 2004).

However, there are also reasons to believe that natural selection may actually be stronger than suggested by the review. Hereford *et al.* (2004) showed that standardizing selection gradients by the phenotypic SD (as in Kingsolver *et al.*, 2001) confounds the strength of selection with the ability of the population to respond to it. The authors advocated instead standardization by the mean phenotype, which expresses the change in relative fitness for a proportional change in the trait. This approach is applicable to traits with a non-arbitrary minimum trait value (i.e., true ratio traits, e.g., body size with a minimum of zero, but not "flowering date"). For the appropriate subset of the Kingsolver *et al.* (2001) data, standardization by the mean yielded a median linear selection gradient of 0.31, which implied selection that is 31% as strong as selection on fitness itself (Hereford *et al.*, 2004).

Furthermore, Blows and Brooks (2003) detected strong nonlinear directional selection on certain trait combinations in the dataset of the review.

One final and important source of bias in estimating selection strengths comes from environmental covariances between fitness and trait, because they can imply selection where in reality there is none. Price *et al.* (1988) discussed the hypothetical example of a bird population in which individuals that breed earlier in the spring produced larger clutches. This apparent selection on breeding date may result from purely environmental variation in overall "condition" such that some birds have sufficient resources to produce large clutches early in the year while others cannot but breed later and lay fewer eggs. The problem can in principle be remedied by estimating selection based on breeding values (approximated as offspring means per family) rather than individual observations (Rausher, 1992). To the extent that family members are randomized over local environmental conditions, this approach eliminates the environmental covariance (but see Postma, 2006 and Hadfield, 2008 for an important caveat). Based on studies of three annual plants, Stinchcombe *et al.* (2002) found that environmental covariances produced significant and substantial biases in 25% of the estimated selection gradients. Path analysis represents an alternative approach to disentangling the causal factors underlying (apparent) selection on a given trait (Scheiner *et al.*, 2002, see also Kruuk *et al.*, 2003).

Thus, there is at present no straightforward conclusion about the strength of selection in unmanipulated natural populations in a single generation. While strong selection per trait and generation may occur only occasionally, it appears that mean phenotypes quite often do not coincide with the current, local optimum.

Patterns of Selection and of Adaptive Responses

Examples of Rapid Evolution

Cases of abrupt changes in ecological conditions have provided the most clear-cut examples that heritable genetic variation and directional natural selection can bring about observable and adaptive phenotypic change. When Trinidadian guppies (*Poecilia reticulata*) were translocated into sites with reduced predation pressure (Endler, 1980), two such replicate populations evolved within 12 and 18 generations, respectively, in the expected direction of earlier sexual maturity and smaller size (Reznick *et al.*, 1997). Recent anthropogenic habitat change has caused classic examples of rapid phenotypic evolution, such as industrial melanism (*see* Selection on Discrete Polymorphisms: An Example), heavy metal tolerance (Antonovics *et al.*, 1971) and crop mimicry in plants (Barrett, 1983). Similarly, species introductions due to human transport have provided numerous examples of rapid evolution of both new arrivals (Lee, 2002) and residents (Strauss *et al.*, 2006), including novel host use, life history adaptations, and altered defenses against predators (Thompson, 1998).

One of the fastest responses to natural selection is the gradual establishment of a latitudinal cline in body size in North American populations of *Drosophila subobscura* within 20 years after their introduction from Europe (Huey *et al.*, 2000). At the end of this period, female wing length clines in Europe and North America were statistically indistinguishable but, intriguingly, based on changes in different parts of the wing. Overall, recent phenotypic changes in response to anthropogenic disturbances were larger on average than those associated with "natural" environmental perturbations, but this increase was partly due to phenotypic plasticity (*see* Phenotypic Plasticity, Hendry *et al.*, 2008).

Fluctuating Selection

When one considers the rates of phenotypic change more generally, a discrepancy emerges: extrapolation of the responses to contemporary selection predicts much faster long-term phenotypic change than is actually observed in the fossil record (Gingerich, 1983; Hendry and Kinnison, 1999). Selection that varies in strength and direction over time represents one possible way out of the dilemma (Thompson, 1998; Bell, 2010). This hypothesis of fluctuating selection is *a priori* appealing given ubiquitous environmental heterogeneity (Connell and Sousa, 1983). Theoretical support comes from Estes and Arnold's (2007) demonstration that macroevolutionary patterns of phenotypic change and divergence can be produced by a microevolutionary model in which the population mean phenotype tracks the bounded meanderings of the local optimum around an overall mean. Finally, a recent survey reported extensive temporal variation in selection strengths and directions (Siepielski *et al.*, 2009). However, Morrissey and Hadfield (2012) found that sampling error was a major source of the observed variation and arrived at a more cautious assessment based on their reanalysis of the data. There is consensus, however, that better knowledge of the causative agents of natural selection would greatly help to clarify the role of fluctuating selection, especially *vis à vis* the alternative hypotheses of drift and immigration (Bell, 2010). In general, it has proven much easier to demonstrate that particular environmental conditions tend to select for certain phenotypic attributes than to identify the particular habitat features that exert the selection pressure.

Empirical demonstrations of fluctuating selection necessarily rely on long-term studies with sufficient temporal resolution (see Bell, 2010 for a list of examples). Females of the freshwater copepod *Onychodiaptomus sanguineous* switch in the spring from producing eggs that hatch immediately to those that undergo diapause (Hairston and Dillon, 1990). The switch date is highly heritable ($h^2 = 0.48$) and is linked to the seasonal increase in fish activity and thus predation: Diapausing eggs are safe from fish, but females that switch too early forego one or more rounds of reproduction. The optimal switch date in a given year also depends on fish density. Hairston and Dillon (1990) documented several correlated cycles of fish density and mean switch date over a 10-year period in a small North American lake. Interestingly, the response to selection was dampened by the influx of individuals from the pool of diapausing eggs (Ellner *et al.*, 1999), which differed in their mean switch date from the current optimum.

The best available time series of evolving phenotypes, selection differentials and pertinent ecological parameters

documents the evolution of body size, beak size, and beak shape in two species of Darwin's finches (*Geospiza fortis* and *Geospiza scandens*) on the Galápagos island of Daphne Major (Grant and Grant, 1989, 2002). Data from the first 30 years of the study produced evidence of both abrupt and more gradual increases and decreases in trait means as well as periods of stasis. These patterns differed between the two species (Grant and Grant, 2002). The main causative agent of selection was variation in food supply that was driven in turn by fluctuations in rainfall levels ranging from severe droughts to the unusually heavy rains of an El Niño event. Significant selection differentials for body size were observed about every 3 years in *Ge. fortis* and runs of significant unidirectional selection in any measured trait lasted no longer than 3 years. The distribution of selection gradients was comparable to the published large-scale survey (Kingsolver *et al.*, 2001): most of them were small ($\beta < 0.15$ phenotypic SDs) and only a few were large ($\beta > 0.5$ SD). Over the 30-year period, however, they contributed to a marked net change in all trait means with one exception (beak size in *Ge. fortis*).

Lack of an Evolutionary Response

The study of Darwin's finches is remarkable also, because the phenotypic responses to natural selection could be accurately predicted from estimated heritabilities, selection gradients, and genetic correlations among the measured traits (Grant and Grant, 1995, 2002). Overall, this is the exception rather than the rule in quantitative genetic studies of natural selection in the wild. Predictions of this quality are usually restricted to studies under controlled laboratory conditions (Lynch and Walsh, 1998). Particularly, there are numerous examples from unmanipulated natural populations in which substantial heritabilities and selection gradients failed to produce a change in the trait mean (Merilä *et al.*, 2001). Careful dissection of some of these cases has illustrated how the interplay between organismal features and environmental variables affects the efficacy of natural selection.

A contrast between apparent selection and observed phenotypic change was discovered in an analysis of red deer (*Cervus elaphus*) antler size based on a 30-year study on the Scottish island of Rum. By applying the animal model, Kruuk *et al.* (2002) showed that antler size was substantially heritable ($h^2 = 0.33$) and also significantly correlated with lifetime breeding success ($\beta = 0.445$ phenotypic SD). From these figures, antler size was predicted to increase over the course of the study, but instead it showed a decrease that was driven by increasing population density. After correction for the latter effect, there was no evidence for a genetic correlation between antler size and fitness and no indication of any evolutionary change in this trait. Instead, the data strongly suggested that the phenotypic correlation between antler size and fitness came about via a third, unmeasured, condition-dependent trait. Here, as in the hypothetical example of egg-laying dates in birds (*see* Potential Biases in Estimated Selection Strengths), environmental covariances between fitness and a given trait can generate the spurious inference of natural selection in action (see also Kruuk *et al.*, 2003; Morrissey *et al.*, 2010).

This example also illustrates how contrasting trends of phenotypic versus genetic change can obscure ongoing evolutionary processes. Classic examples are phenotypic clines of increasing body size and development time with altitude that mask trends in genotypic values in the opposite direction. For example, mountain populations of green frogs (*Rana clamitans*) face a much shorter larval growing season than their conspecifics in the lowlands. This forces them to overwinter once or even twice before metamorphosis. The resulting intense selection for rapid development is revealed under constant laboratory conditions, where tadpoles from upland populations develop significantly faster than those from the lowlands and especially so at low temperatures (Berven *et al.*, 1979). Such counter-gradient selection (for a review see Conover and Schultz, 1995) can result in phenotypic homogeneity, when environmental and genetic influences balance each other. Similarly, phenotypic stasis within a given population over time might be caused by an ongoing evolutionary response to directional selection that balances an environmental trend in the opposite direction.

Temporal variation in heritability contributed to an unexpectedly low selection response in an intensely studied population of Soay sheep (*Ovis aries*) on one of the St. Kilda islands (Wilson *et al.*, 2006). Birth weight in these sheep is both heritable and strongly positively associated with the survival of lambs during their first summer. Nevertheless, it showed no increase over the 20-year study period. Overall, variation in density and climatic conditions caused substantial annual variation in the total first-summer survival. This variation affected both the heritability and strength of selection on birthweight: strong selection and low heritability coincided in low survival years, whereas weak selection and increased heritability were observed in good years. The resulting negative correlation between the two independent variables in the breeder's equation reduced the maximum possible response to selection.

There is an interesting twist in this story. Total heritability in birthweight has two components: one based on the additive genetic variance of the trait in the offspring and a second one due to genetic maternal effects. Maternal effects comprise all means by which a mother's phenotype influences the phenotype of her offspring over and above the additive effects of the transmitted genes. Maternal effects used to be thought of mainly as components of environmental variance, as in the case of larger offspring from females that are in good condition. However, there can also be genetic maternal effects due to heritable phenotypic variation among females that contributes to the variation in some measure of offspring phenotype. Any component of maternal investment and parental care might in principle display such effects, and they are increasingly recognized as an important variance component in natural populations (for a review see Räsänen and Kruuk, 2007). In the case of the Soay sheep, temporal variation in total heritability of birthweight could indeed be attributed to maternal genetic effects. Under good environmental conditions, there was relatively more heritable variation in the mothers' ability to promote the *in utero* growth of their offspring.

Taken together, the examples in this section illustrate that point estimates of heritabilities and selection strengths

provide only a rough, first approximation of the scope for microevolutionary change in natural populations. Temporal variation in either of these variables can considerably alter the course of ongoing adaptive phenotypic change, and insufficient temporal resolution in a given dataset may miss some or all of the dynamics. Furthermore, a purely phenotypic correlation between a trait and fitness gives the erroneous impression that the trait plays a causative role in fitness variation. Finally, spatial or temporal trends in environmental effects can override or balance genetic change in the opposite direction. An additional impediment to adaptive phenotypic change is considered in the next section.

Genetic Correlations

Pleiotropy

While there is an effectively infinite number of traits that ecologists and evolutionary biologists might choose to measure on a given organism, the number of developmental pathways that lead to their manifestation is clearly finite. This suggests widespread genetic interdependencies among traits due to a partially or entirely shared genetic basis. Genetic variation that affects the expression of any one trait is likely to affect others as well, a condition known as pleiotropy. Comparisons of estimated mutation rates per trait versus per genome and considerations of total fitness variance within populations similarly argue for widespread pleiotropy (Barton, 1990). Pleiotropy causes genetic correlations among traits with a shared genetic basis which are measurable as correlations between breeding values. As a consequence, direct selection on one trait will cause indirect selection in any correlated trait. Selection gradients (*see* A Review of Selection gradients and **Box 2**) were devised to estimate the strength of direct selection on a given trait in such cases. Note that a genetic correlation can be the result not just of pleiotropy but also of tight genetic linkage between loci that affect different traits.

Genetic correlations can impede the response to joint selection on several correlated traits (antagonistic pleiotropy). Of particular interest is the case of a genetic trade-off in which the fitness gain due to selection on one trait is cancelled by an associated fitness loss in a second, correlated trait. Antagonistic pleiotropy and genetic trade-offs feature prominently in the context of ecological specialization (see e.g., Hawthorne and Via, 2001), life history theory (*see* Life History Trade-Offs) and sexually antagonistic selection (e.g., Foerster *et al.*, 2007).

Genetic Correlations: A Case Study

Limited genetic variation for certain trait combinations is a recurring theme in the study of flower shapes. In the wild radish (*Raphanus raphanistrum*), the lengths of filaments and of the corolla are highly correlated such that anthers protrude from the corolla by a certain amount (the so-called anther exsertion). The strength of the genetic correlation is illustrated in a plot of filament length against corolla length: individual breeding values cluster along a straight line with positive slope (Conner, 2003). In other words, most genetic variation is arrayed along that one axis (=flower size), whereas little genetic variation exists along the perpendicular axis that represents variation in anther exsertion. Indeed, the latter trait is under stabilizing selection in the field, because plants with intermediate trait values achieve the highest seed-siring success (Morgan and Conner, 2001). The current correlation pattern may thus reflect a shared developmental pathway and/or past selection (Conner, 2006). However, the genetic correlation between filament and corolla length ($r=0.85$) did not pose an absolute constraint ($r=1$ or $r=-1$) and anther exsertion did respond to five generations of artificial selection (Conner 2003).

Life History Trade-Offs

Genetic trade-offs are key components of life history theory. Classic examples are egg size versus egg number and survival versus reproduction. The most clear-cut evidence for their existence comes from artificial selection (Fry, 2003) and manipulative experiments (Zera and Harshman, 2001). *Drosophila melanogaster* that have been artificially selected for early reproduction have a shorter lifespan. When egg production was experimentally suppressed in females at the end of the selection experiment, the mortality difference between early and late reproducing selection lines disappeared, which established the direct causal link between early investment into eggs and increased mortality later in life (Sgro and Partridge, 1999). Experimental manipulations of the side-blotched lizard (*Uta stansburiana*, Sinervo, 1999) illustrated how hormones mediate the trade-off between egg size and egg number in this species. Such trade-offs are expected to leave their mark on additive genetic correlations among traits within populations. In the case of the lizards, one might expect that V_A for egg number should entail measurable negative genetic and phenotypic correlations between the number and size of eggs. However, numerous attempts to demonstrate trade-offs in the form of negative genetic correlations across a wide range of taxa and traits have produced surprisingly few clear-cut examples.

Several hypotheses have been advanced to explain this puzzling pattern. While an individual necessarily faces a trade-off when allocating a given resource pool to competing organismal functions, no such trade-off need be apparent at the population level. This will be the case when individuals differ not only in their allocation rules but also in the amount of resources at their disposal. Those that have acquired a larger resource pool can allocate more to each fitness enhancing trait, which results overall in a positive correlation between these traits (van Noordwijk and de Jong, 1986). Using a standard modeling framework for quantitative traits, Houle (1991) showed that a simple dichotomy between acquisition and allocation loci can produce either negative or positive genetic correlations depending on the relative number of loci in each category. Trade-offs may also be masked when there is a hierarchy of allocation decisions (Worley *et al.*, 2003). Allocation to somatic growth versus reproduction will often precede decisions regarding, for example, size versus number of offspring or, in hermaphrodites, investment into male versus female function. In such a hierarchy, positive genetic

correlations between the traits that trade-off later in life can arise if the population mean moves toward the overall optimal allocation scheme and if there is relatively more genetic variance for the earlier allocation decision.

Inferring Genetic Constraints from the G Matrix

Pairwise genetic correlations will also be misleading when additional genetic correlations with unmeasured traits play a role. Such "missing traits" pose a problem not just in the prediction of selection responses (**Box 2**) but also in the inference of constraints from the additive genetic variance–covariance matrix (the **G** matrix, **Box 2**). In general, strength and pattern of such constraints depend on the structure of the **G** matrix as a whole and cannot be inferred from particular matrix elements (for a review see Walsh and Blows, 2009). In fact, some of the trade-offs between component traits are expected to result in positive pairwise covariances (Charlesworth, 1990). Similarly, negative genetic correlations larger than -1 do not translate into "weak" constraints that would merely delay rather than prevent an adaptive response. Pertinent information instead comes from the eigenvalues and eigenvectors of the **G** matrix. An eigenvalue of zero signifies an absolute genetic constraint in the direction of the associated eigenvector. But even a matrix with only nonzero but highly variable eigenvalues imposes substantial constraints on adaptive evolution.

This is true in particular, if selection acts simultaneously on several correlated traits, a situation that should be commonplace in nature. In that case, only relatively few mutations at the underlying pleiotropic loci are expected to meet the manifold demands of multitrait selection (directional and/or stabilizing) and so increase net fitness (cf. review by Walsh and Blows, 2009). The contrast between successful artificial selection on single traits and the frequently observed lack of a selection response in natural populations may be due, at least in part, to the more complex selection regimes in nature. While we presently lack the necessary data to test this hypothesis, it seems safe to conclude that the sole focus on abundant heritabilities and rare absolute pairwise trade-offs results in an overly optimistic view of adaptive microevolutionary change.

Phenotypic Plasticity

An Adaptive Strategy in Heterogeneous Environments

Phenotypic plasticity refers to the observation that a given genotype expresses different phenotypes in different ecological settings. At its most basic level, the concept applies to any differences in trait means between environments. For example, plants may show stunted growth in habitats other than the one to which they are adapted. However, when trait expression in different habitats has been shaped by natural selection, phenotypic plasticity represents an ecological strategy that adapts organisms to heterogeneous environments and forms an integral part of species interactions (Agrawal, 2000). Such adaptive plasticity can range from repeated ontogenetic adjustments in behavior, physiology, or life history to the expression of distinct, irreversible morphologies (polyphenism) such as fully winged and flightless forms in insects (Harrison, 1980; Roff, 1986), the shade avoidance growth form of plants (Schmitt *et al.*, 2003) carnivorous and omnivorous tadpole morphs of spadefood toads (Pfennig *et al.*, 2007) inducible defensive morphs of *Daphnia* water fleas (Tollrian and Dodson, 1999) and the seasonally alternating morphs of *Bicyclus anynana* butterflies (Lyytinen *et al.*, 2004).

Theoretical Predictions and Empirical Tests

If phenotypic optima differ across space and time, then a genotype that can reliably express any local optimum at no extra cost is clearly at an advantage and should outcompete all others. Clearly no such organism exists, because it would have an infinite ecological niche. Numerous theoretical studies have specified conditions under which a costly plastic strategy is favored over that of a nonplastic generalist or a collection of locally adapted specialists (reviewed in Berrigan and Scheiner, 2004). Key considerations include the temporal and/or spatial patterns of environmental change, the predictability of such change at the time when developmental "decisions" are made by the organism and the relative frequencies and carrying capacities of each habitat type. The following empirical examples address (1) the assumption that there is a cost to plasticity and (2) the prediction that more variable environments should select for greater plasticity.

To date, many thorough attempts to demonstrate various kinds of plasticity costs (see DeWitt *et al.*, 1998; van Kleunen and Fischer, 2005; Auld *et al.*, 2010 for reviews and classification schemes) have resulted in numerous suggestive trends but only a few clear-cut examples (Auld *et al.*, 2010). In Swedish populations of the common frog, *Rana temporaria*, Merilä *et al.* (2004) detected a cost of accelerated tadpole development in response to simulated pond drying: genotypes that exhibited a stronger plastic response also metamorphosed at a smaller size when reared in controls at a constant water level. This result lends support to the hypothesis that the maintenance of a plastic response system *per se* can be costly. Other examples of demonstrated plasticity costs come from the shade response (Dorn *et al.*, 2000) and flowering time plasticity (Stinchcombe *et al.*, 2004) in *Arabidopsis thaliana*. Despite the limited number of such examples, costs may nevertheless be critical in determining whether phenotypic plasticity in a given trait evolves at all. By the time an established plastic strategy is subjected to scientific analysis, the initial costs may have been absorbed by further adaptive changes (Sultan and Spencer, 2002).

A classic study by Cook and Johnson (1968) tested the hypothesis that greater levels of plasticity are expected in populations from more variable environments. The semiaquatic lesser spearwort, *Ranunculus flammula*, exhibits so-called heterophylly: partially submerged individuals of this species produce different leaf shapes above versus below the water line. Under controlled conditions, the level of individual plasticity (measured as the difference in shape between the leaf types) increased with the level of environmental heterogeneity both among and within populations. Similarly, two populations of orange jewelweed (*Impatiens capensis*) within 1 km of each other differed in their plastic response to density

in a way that was predicted by estimated selection gradients for plasticity (Donohue *et al.*, 2000; Donohue *et al.*, 2001).

Natural Selection of the Norm of Reaction

A compact graphical presentation of phenotypic plasticity is provided by the norm of reaction (reviewed in Via *et al.*, 1995). It is equal to the mean phenotypes that a given genotype expresses along an environmental gradient (**Figure 1**). Significant differences among the slopes of reaction norms for a collection of genotypes (cf. **Figure 1(b)**) are referred to as genotype × environment interactions (G × E). They indicate genetic variation for plasticity and are frequently observed in natural populations (Pigliucci, 2005). How efficiently this variation can be molded by natural selection depends partly on the patterns of environmental heterogeneity. In a so-called fine-grained environment, in which a range of environmental conditions is typically encountered by an individual during its lifetime, plasticity is a feature of that individual. It can be estimated from repeated measurements and analyzed like any other quantitative trait (Nussey *et al.*, 2007).

In contrast, a coarse-grained environment changes more slowly such that an individual experiences only one of several environmental states. Consequently, plasticity does not exist at the level of the individual but only at the level of its genotype. Estimation of reaction norms then requires clonal replicates of a genotype or, as a second-best option, full-sib families to be distributed along an environmental gradient. Selection experiments to increase or decrease this type of plasticity (i.e., to alter the slope of the reaction norm) have met with limited, if any, success (Scheiner, 2002), whereas selection on the overall mean of the reaction norm (i.e., its elevation) have generally been successful. These results may be due to the fact that selection above the level of the individual is less efficient (Williams, 1992) or point to genetic constraints (Via and Lande, 1985), such that a significant response is less likely to materialize over the course of a selection experiment.

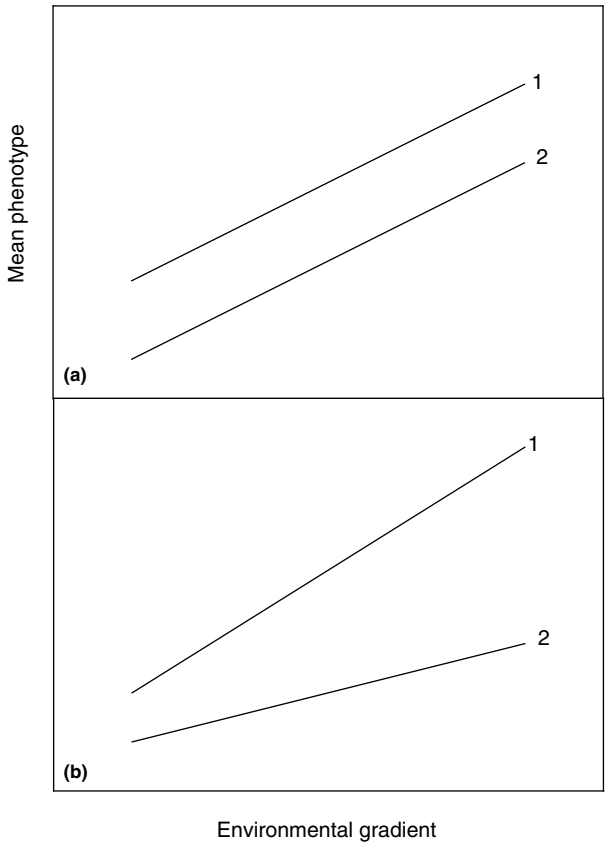


Figure 1 Norms of reaction. Shown are the norms of reaction for a given trait and for two genotypes, 1 and 2, along an environmental gradient. (a) The genotypes differ in the overall trait mean but are equally plastic, as indicated by the identical slopes of their reaction norms. (b) Genotype 1 not only expresses larger trait values overall but also shows greater plasticity than genotype 2.

A striking example of how two populations of the same species can differ in their level of adaptive plasticity comes from two long-term studies of great tits (*Parus major*). In response to increasing spring temperatures over the past few decades, the mean egg-laying date as well as the peak in food abundance (moth caterpillars) have advanced considerably both in a British and a Dutch population. In the British population, the interval between the two events has been maintained such that peak food abundance continues to coincide with the greatest food demand by the offspring and the population is thriving (Charmantier *et al.*, 2008). In the Dutch population, however, laying date and food peak have advanced at different rates. The increasing mismatch between both has resulted in a steady decline of reproductive success over the past 30 years and in measurable selection for increased plasticity. But while there is significant heritable variation in plasticity of laying date ($h^2 = 0.32$), there is as yet no evidence for a selection response (Nussey *et al.*, 2005). While the established reaction norm in the British population happens to fit the temperature trend over the past 30 years, there are apparently constraints on achieving a similar fit in the Dutch population despite available genetic variation.

Phenotypic Plasticity versus Microevolutionary Change

In any analysis of phenotypic change following an environmental perturbation or a colonization event, phenotypic plasticity is the obvious alternate hypothesis to adaptive microevolutionary change and needs to be evaluated. This holds even for traits that are not generally thought of as being responsive to environmental conditions. The experimental introductions of *Anolis sagrei* lizards to 14 small islands in the Bahamas from a nearby source population are a case in point (Losos *et al.*, 1997). The shorter vegetation on the islands was predicted to select for relatively shorter hindlimbs, because these allow for greater agility on narrower perches. Within 14 years after the introduction, exactly this pattern was observed, which is also manifest more broadly in intra and interspecific comparisons in the spectacular radiation of Caribbean *Anolis* species. Phenotypic plasticity most likely played a role in the island introductions, because in a separate study the lizards were shown to vary in their relative hindlimb length depending on the kinds of perches on which they had been raised (Losos *et al.*, 2000). Note that genetic drift may

also be responsible for phenotypic changes especially when population numbers have been severely reduced following an introduction, a range of expansion or an environmental perturbation. However in that case, one does not expect a correlation between ecological parameters and the direction of phenotypic change.

The role of plasticity in adapting organisms to novel habitats has been viewed from two different perspectives: either as a source of phenotypic novelty that may be subsequently fixed by the genetic system ("genes as followers," West-Eberhard, 2003) or as an evolved, genetically codified strategy that is subject to modification by natural selection ("genes as leaders," de Jong, 2005). In the latter case, plasticity is expected to aid the colonization of a novel environment if conditions there resemble those that have been previously experienced and have thus shaped the plastic response (Ghalambor *et al.*, 2007). Price *et al.* (2003) suggested that an intermediate level of plasticity is optimal, because it allows the organisms to persist in a new habitat while still offering scope for genetic adaptation. Tests of these hypotheses would allow for predictions of the success of invasive species and the persistence of established populations at a time of global change.

Local Adaptation versus Gene Flow

"Site-Specific" Adaptation is Common

Environmental heterogeneity that is coarse grained and mostly spatial rather than temporal favors genetically based local adaptation over costly phenotypic plasticity (Kawecki and Ebert, 2004). For a given habitat, local adaptation is defined as a higher fitness of resident as opposed to immigrant genotypes. A recent survey of reciprocal translocation experiments showed that this condition was fulfilled in 65% of all cases (Hereford, 2009). In 29% of the studies, however, immigrants had a higher fitness than residents, which suggests that there are impediments to the spread of locally beneficial alleles that have arisen elsewhere. Hereford (2009) also found evidence for trade-offs in 49% of all cases such that a given genotype had a fitness advantage in its own but a fitness disadvantage in the alternate habitat. This analysis underscores the frequent occurrence of "site-specific" adaptation, but does not take the distance between the experimental sites into account. Therefore, it does not address the tension between selection and gene flow that limits local adaptation on small spatial scales. The following is an overview of the theory on this issue.

Population Genetic Models

In the case of strictly spatial heterogeneity, the critical conditions for local adaptation can be illustrated with a simple deterministic model of two populations of constant and equal sizes in which alternate alleles at a single locus are selectively favored (see Lenormand, 2002 for a more detailed presentation of the theory). Local adaptation in this model is equivalent to the persistence of both alleles at equilibrium. Without gene flow, only the local fitnesses determine the equilibrium conditions and each allele will be fixed in the population in which it is favored. At the other end of the spectrum, imagine that in every generation all offspring are thoroughly mixed and then each randomly assigned to one or the other population. In this case, only the mean fitnesses across both populations matter. Both alleles will then persist, if their mean fitnesses are exactly identical (Levene, 1953). For even slight differences in mean fitness, local adaptation is impossible unless genetic exchange is limited: the gene flow rate m (i.e., the proportion of each population that is replaced by immigrants per generation) must lie below a threshold that is determined by the strength of local selection in each population (Bulmer, 1972).

For example, if one allele has a local fitness advantage in its "own" population of 5% but a mean fitness disadvantage of 2.5%, then m must be smaller than 0.05 in order for it to be maintained at equilibrium. For $m \geq 0.05$, mean fitness "wins" over local fitness and the allele will be lost. The maximum permissible m becomes even smaller as the mean fitness disadvantage increases. Conversely for a given set of mean fitnesses, local adaptation can only be sustained in the presence of more gene flow if local selection is proportionately stronger. Finally, if the conditions for local adaptation are met, then the regular influx of the other, maladapted allele causes the population mean in each population to deviate from the local optimum. This deviation is called the migration load.

This model can be extended to populations of unequal size. Mean fitness is then weighted by population size, because it reflects the average selective regime experienced by individuals. Consequently, the allele that is favored in the smaller population has a reduced mean fitness and the conditions for its persistence become more stringent. Local adaptation to environmental conditions that are rarely encountered, for example, at the edge of a species range, is thus harder to achieve.

Another set of models takes the spatial scale of dispersal and of selective regimes explicitly into account. In these models, individuals are continuously distributed in space. Dispersal is quantified in terms of the mean distance that an individual travels between birth and reproduction (σ). Assuming a sharp ecotone, that separates two different, infinitely large habitats, and two differently adapted alleles at a single locus (as above), one expects at equilibrium a clinal change of allele frequencies across the ecotone and fixation of the locally favored allele on either side. The width of the cline is proportional to the characteristic length $\sigma/\sqrt{s}$ (Slatkin, 1973), where s is the selective advantage of the locally favored allele. Weak selection and wide-ranging dispersal thus widen the area in which the maladapted allele persists and imposes a migration load. The characteristic length also determines whether a locally adapted allele can be maintained in an "environmental pocket" that is embedded in a larger region in which the alternate allele is favored. For all but the strongest selection pressures, the minimum size of the pocket must extend over several to many dispersal distances (Slatkin, 1973).

The analysis of quantitative traits at an ecotone is more complicated in part because the mixing of differently adapted

populations generates statistical associations among alleles at different loci (so-called linkage disequilibria) which violate the assumptions of the quantitative genetic model. However, based on a number of approximations, it could be shown that the trait mean should vary on a scale of $\sigma/\sqrt{2sV_A}$ (Slatkin, 1978, Barton, 1999) and so should be able to track the environment quite closely. Spichtig and Kawecki (2004) simulated clines in polygenic traits at an ecotone and found that local adaptation was favored by strong selection for the locally adapted phenotype and by moderate selection against intermediate phenotypes. If intermediates were selected too strongly, then a locally adapted population could not become established, whereas weak selection allowed generalist intermediates to outcompete the locally adapted specialists.

Empirical Studies

There are a number of clear illustrations how gene flow can reduce the degree of local adaptation. For example, the optimal egg-laying dates of blue tits (*Parus caeruleus*) along the Mediterranean coast differ between deciduous and evergreen forest by about a month. On the French mainland, frequent gene flow from the predominant deciduous forest into smaller patches of evergreen forest caused a mismatch between laying date and bud burst in the latter (Dias and Blondel, 1996). On the island of Corsica however, gene flow rates are demonstrably much lower such that blue tit populations in deciduous forest (the minority habitat there) were well adapted and strongly genetically differentiated from nearby populations in evergreen forest (Blondel *et al.*, 1999). Negative correlations between estimated gene flow rates and adaptive phenotypic divergence among nearby populations were found in streamside salamanders (*Ambystoma barbouri*, Storfer and Sih, 1998) and threespine sticklebacks (*Gasterosteus aculeatus*, Hendry and Taylor, 2004). A long-term study of local selection on a suite of behavioral traits in the spider *Agelenopsis aperta* documented in detail how selection acted on immigrants and their offspring. It also included theoretical modeling to define the phenotypic optima and to predict how quickly these would be reached if gene flow ceased (Riechert, 1993; Riechert *et al.*, 2001).

Gene Flow Can Promote Local Adaptation

Gene flow need not be an antagonist to adaptation. Most importantly, it spreads universally favored mutations across a species range (for a review see Morjan and Rieseberg, 2004). It can also aid local adaptation by supplying new alleles to populations with limited genetic variance. This can facilitate adaptation at the edge of a species range (see Geber, 2011 and associated papers) or after a drastic reduction in population size. The coevolution between parasites and hosts represents another evolutionary process in constant need of new genetic variation.

The theoretical prediction that gene flow can facilitate local adaptation in such systems (Gandon, 2002) has been experimentally tested with bacterial hosts and bacteriophage "parasitoids." Experimental evolution in microbes allows for the study of large, replicated populations over many generations and for the direct comparisons of genotypes across time (i.e., from frozen stocks). In the present context, it also enables the experimenter to control a suite of key parameters that would be difficult to measure let alone manipulate in nature. Compared to a set-up without gene flow, regular genetic exchange between replicate microcosms of bacteria and phage increased both the overall occurrence and the spatio-temporal variability of local adaptation in the phage (Forde *et al.*, 2004). The latter observation might explain why snapshots of unmanipulated host–parasite systems in nature give overall inconsistent results with regard to local adaptation. In another study (Morgan *et al.*, 2005), host and parasitoid gene flow rates were manipulated independently (either host or parasitoid or no gene flow). As before, parasitoid gene flow led to increased local adaptation in parasitoids. In addition, there was evidence at the end of the experiment that parasitoid gene flow had increased parasitoid adaptation to nonlocal hosts, which could reflect the spread of universally favored alleles. Interestingly, host gene flow had no effect on the level of local adaptation in the host.

Genetic Basis of Ecological Variants

Quantitative Trait Locus (QTL) Mapping

Methodological developments over the past 20 years have provided insights into the genetic basis of quantitative trait variation that had so far been hidden inside the "black box" of quantitative genetics. The development of abundant genetic markers, linkage maps, and statistical methodology has led to new, sophisticated implementations of an old idea (Sax, 1923): the approximate genomic locations and effect sizes of genes underlying a trait can be identified through statistical associations between trait values and marker genotypes (Mackay, 2001; Mauricio, 2001). The typical design involves a cross between individuals from genetically and (usually) phenotypically divergent selection lines or populations for which a linkage map of genetic markers exist. The largely heterozygous F1 generation is then further crossed to generate sets of recombinant offspring (F2 or backcross) in which significant cosegregation of a trait with alleles at a particular marker locus identifies a "quantitative trait locus" (QTL) that is closely linked to the marker. In reality, a QTL may represent more than one linked locus influencing the trait. More recently, additional statistical methods have been developed to map QTL in outbred populations including pedigreed populations in the wild (Slate *et al.*, 2010).

Test of the Quantitative Genetic Model

Quite often, such studies have detected a few loci of large effect and a relatively large number of small effects for a given trait. In addition, there are numerous examples of dominance and of epistatic interactions (Roff, 2007). At face value, these data contradict the assumptions on which the quantitative genetics model is built (cf. **Box 1**). However, several considerations narrow the apparent gap between theory and data. Hill *et al.* (2008) emphasized that significant nonadditive allelic interactions in test crosses do not necessarily imply substantial nonadditive genetic

variance at the population level, because the latter depends on allele frequencies. For loci underlying QTL, these are expected to have U-shaped distributions (i.e., most alleles have frequencies either near 0 or near 1) in the case of neutrality and under various selective regimes. This implies that most nonadditive allelic effects would mainly contribute additive genetic variance to the trait. Furthermore, if the commonly observed epistatic QTL effects are largely contingent on the particular subset of alleles that is present in test crosses, then some large additive effects of "major loci" may similarly disappear if the alleles in question were randomized across genetic backgrounds in the population (Roff, 2007). In addition, the small sample sizes used especially in early QTL studies ($N \approx 200$) led to a systematic underestimation of QTL numbers and overestimation of their effects (Beavis, 1998). Finally, traits based on structural loci of large effect may still have normal genotypic and phenotypic distributions if regulatory genes generate a continuous distribution of expression levels (Roelofs *et al.*, 2006). Thus, while loci of large effect clearly do contribute to continuous trait variation in natural populations and, indeed, are expected to play an important role in the adaptive process (*see* The Genetics of Adaption), currently available QTL data most probably overestimate their frequency and, on balance, do not amount to a refutation of the quantitative genetics model.

The Genetics of Adaptation

The distribution of allelic effects underlying a trait depends in part on the strength of directional selection acting on it. This is dramatically illustrated by the adaptive responses to strong anthropogenic selection, for example, via insecticides, warfarin or heavy metal contamination of mine sites. Resistence in these cases is overwhelmingly based on no more than three loci and mostly just one locus (Macnair, 1993; ffrench-Constant *et al.*, 2004). In fact, a gradual process based on genes of small effect is not an option here, given that a) the new optimum lies decidedly outside the initial distribution of phenotypes and b) survival requires a considerable level of resistance from the outset (Macnair, 1991). For example, the establishment of grasses on contaminated mine sites depended on the presence of the tolerance allele at low frequency in the surrounding population. None of the species that grew nearby but lacked such pre-existing alleles have succeeded in colonizing the mine sites (Bradshaw, 1991). More generally, theory predicts that adaptation to a new phenotypic optimum consists of a sequence of allelic substitutions of diminishing effect size. The typical finding in QTL studies of few "large" and many "small" loci agrees qualitatively with this prediction, as does the observation in microbial studies that alleles of large effect are fixed early in the adaptive process (for a review see Orr, 2005).

Insights from the Genes Underlying Adaptation

There is usually a long way to go from the identification of a QTL to the discovery of the gene(s) that actually produce the phenotypic effect. However, once identified, they offer new insights that can go beyond the scope of ecological genetics and so provide a much broader context to the field study. The following examples illustrate, respectively, a molecular mechanism of heterozygote advantage, contrasting pleiotropy per locus versus per trait and the phylogenetic origin of an adaptation.

In *Colias* butterflies, phosphoglucose isomerase (*Pgi*) is a key enzyme involved in the energy supply to flight muscles. In both *Colias eurytheme* and *Colias meadii*, two common alleles persist at intermediate frequencies due to heterozygote advantage in terms of enzyme kinetics, flight ability, and fitness (including male mating success and female survival; Watt *et al.*, 1983). The underlying mechanism of this effect has been explored at the level of enzyme structure in *Co. eurytheme* (Wheat *et al.*, 2006). In this dimeric enzyme, heterozygote advantage appears to originate from the heterodimer's flexible conformation that allows enyzme performance to alternate between maximizing substrate binding strength and processivity. Under this hypothesis, the heterozygote can resolve a trade-off that limits the performance of either homozygote.

Following the last ice age, many independent colonizations of Canadian freshwater lakes by marine threespine sticklebacks (*Ga. aculeatus*) have prompted parallel changes in a whole suite of traits. While QTL studies have found most of these to be polygenic, one trait, the number of pelvic plates, is clearly associated with a major locus, ectodysplasin (*Eda*, Albert *et al.*, 2008). The high-armored morph of marine populations is thought to protect against fish predation. In lakes, the low-armored allele has become established in most cases. Laboratory and artificial pond experiments that mimicked the freshwater habitat, led to a significant increase in its frequency and it conferred a growth advantage to very young fish, which should reduce mortality due to invertebrate predation (Barrett *et al.*, 2008). However, an initial, substantial frequency decline of the low-armored allele in the ponds (and only there) as well as apparent heterozygote disadvantage in a natural polymorphic lake population indicate that selection at this locus is considerably more complicated (Schluter *et al.*, 2010). Similarly complex antagonistic selection due to pleiotropy or closely linked genes was also discovered for a flower-color locus in the morning glory (*Ipomoea purpurea*, Coberly and Rausher, 2008) and for a coat-color locus in a wild sheep population (Gratten *et al.*, 2008).

Finally, phylogenetic analyses of genes linked to several QTL have provided important new insights into the rapid evolution of a new host race in the apple maggot fly (*Rhagoletis pomonella*) in North America. In this classic example of sympatric differentiation, development of the larvae on apple (the new host) as opposed to hawthorn requires a different timing of winter diapause. The genes underlying this trait map only to three chromosomal regions all of which are located in inversions (Feder *et al.*, 2003). This genetic architecture greatly facilitates sympatric differentiation, because the inverted regions do not recombine with the "wildtype" homologs. Moreover, molecular phylogenies of genes within these inversions suggest that the flies' response to disruptive selection since the introduction of apples about 150 years ago made use of an existing diapause syndrome that is 1.5 million years old (Feder *et al.*, 2003).

Eco-Evolutionary Dynamics

From Populations to Communities and Ecosystems

Phenotypic change due to microevolutionary processes such as adaptation to a new resource, more efficient exploitation of an old one or the acquisition of resistance to a pathogen is likely to effect the abundance of the evolving species. Changes in density commonly alter the optimal strategies of growth and reproduction, which might trigger further adaptive responses. This "quintessential link" between selection response and numerical abundance can generate an "ecogenetic feedback loop" in the target species (Kokko and Lopéz-Sepulcre, 2007). It may also have profound effects on other species with which it interacts. There may be "ripple effects" (Thompson, 1998) throughout a biological community that could range in principle from the exclusion of a competitor to changes in food web structure or in patterns of nutrient cycling. Alternatively, concurrent subtle microevolutionary changes in more than one interacting species may not immediately alter the structure of a community but affect its stability (Jones *et al.*, 2009). Thompson (1998, p. 331) argued that "to dismiss or ignore these rapid, fluctuating evolutionary processes as trivial evolutionary meanderings… is to miss the real interface of evolution and ecology in community dynamics." The increase in successful demonstrations of adaptive microevolution in single species has led to a recent surge of interest in these community effects (e.g., Neuhauser *et al.*, 2003; Hairston *et al.*, 2005; Fussmann *et al.*, 2007; Kokko and Lopéz-Sepulcre, 2007; Pelletier *et al.*, 2009 and associated papers).

The ecogenetic feedback loop focuses on single species: it is closed when adaptive change in a population leads to a change in population density which, in turn, causes variation in reproductive success among individuals based on some aspect of their phenotype (Kokko and Lopéz-Sepulcre, 2007). Eco-evolutionary dynamics are more broadly defined (Pelletier *et al.*, 2009; Post and Palkovacs, 2009). They comprise the unidirectional links from evolution to ecology and vice versa as well as bidirectional interactions. The link from evolution to ecology may consist of demographic change in one or more species, adaptive responses in interacting species, changes in community structure or stability or changes in ecosystem processes. The link in the opposite direction can affect not just mean phenotypes but also the levels of heritable genetic variation and the spatial distribution of genetic variation across landscapes of fragmented and potentially heterogeneous habitats (see Zheng *et al.*, 2009 for an example). Strong links are expected between strongly interacting species, whereas effects are expected to diminish as one moves from pairwise interactions to whole communities and to ecosystems (Pelletier *et al.*, 2009). Since antagonistic pairwise species interactions, for example, between hosts and pathogens can be both strong and persistent, they have long featured prominently in eco-evolutionary studies (e.g., Dwyer *et al.*, 1990; Thrall and Burdon, 2003).

Demonstrating a complete feedback poses a formidable challenge especially in natural populations (Kokko and Lopéz-Sepulcre, 2007). A large amount of high-resolution data is required to establish the causal links between concurrent changes. Once a system has settled at a new equilibrium, it will often be impossible to reconstruct these in hindsight. In addition, the strength of the interaction may vary locally. For example, a third species that weakens the link between two primary contestants may occur in some places but not in others. According to the geographic mosaic theory of coevolution (Thompson, 2005), spatially variable conditions across sites mean that reciprocal evolutionary change will occur in some sites ("coevolutionary hot spots") but not others and may subsequently spread to other populations.

The analysis of eco-evolutionary dynamics typically requires a combination of observational data, experimental studies, and the analysis of theoretical models. A rich theoretical literature has accumulated over the past 40 years (reviewed in Fussmann *et al.*, 2007; Kokko and Lopéz-Sepulcre, 2007) that provides inspiration for the analysis of any particular eco-evolutionary study system. Theoretical models feature in all of the following three examples.

Case Studies

Populations of *Daphnia dentifera*, a cyclical parthenogen, are subject to episodic outbreaks of infection with a yeast parasite (*Metschnikowia*), which severely reduces fitness and ultimately kills the host. Duffy and Sivars-Becker (2007) found that the susceptibility to infection in *Da. dentifera* clones from populations with a recent parasite outbreak was lower on average and also less variable compared to lake populations that had not experienced a recent epidemic. The authors therefore included clonal selection by the parasite into a model of this species interaction and incorporated field estimates of all key parameters. The model generated correct predictions of outbreak duration and host population size during an outbreak only in the case of a genetically variable host population. Particularly, outbreaks came to a definite end in nature and in this version of the model, whereas they persisted endemically in models of genetically uniform hosts. Thus, eco-evolutionary dynamics may be able to resolve a commonly observed discrepancy between theory and data that has puzzled epidemiologists for some time.

The Soay sheep population on St. Kilda undergoes frequent fluctuations in abundance. Population crashes have been caused by a combination of high densities and adverse weather conditions (mild winters and heavy spring rains), but different age groups of each sex were differently affected by these conditions. For a given weather regime and total population size, the population growth trajectory will therefore be affected by the current age structure (Coulson *et al.* 2001). Moreover, for several sex/age subgroups within the population, the patterns of survival and reproduction were nonrandom with regard to body size (Pelletier *et al.*, 2007). The strength of natural selection on body size was considerably greater in crash years as opposed to noncrash years. Pelletier and coworkers traced the effect of natural selection on body size via survival during and reproduction following a crash all the way to the level of additive genetic variance for that trait.

Finally, in a chemostat experiment of a copepod predator (*Brachionus calyciflorus*) and its algal prey (*Chlorella vulgaris*),

evolutionary processes had to be invoked to explain the observed population cycles (Yoshida *et al.*, 2003). The results from initial experiments had been incompatible with standard models: the observed predator and prey cycles had periods that were two long and population maxima that were too strongly offset relative to each other. However, agreement with the theory was achieved when only a single algal clone was used in the chemostat. Separate selection experiments established the existence of genetic trade-off among *Chlorella* clones between reduced nutritious value (a defense against the predator) and competitive ability, which were favored in the presence and in the absence of *Brachionus*, respectively. When clonal diversity reflecting this trade-off was incorporated into the model, it correctly predicted the length and phase relationship of the observed cycles. Moreover, when the relative frequencies of two well-characterized algal clones were directly monitored, the chemostat dynamics were indeed attributable to clonal selection as predicted by the trade-off (Meyer *et al.*, 2006).

Conceptual Developments

The recent empirical advances in the study of eco-evolutionary dynamics have been paralleled by a number of conceptual developments. Hairston *et al.* (2005) have argued that the rate of evolutionary change per time is not a good metric with which to determine whether evolution has operated "rapidly," i.e., on an ecological time scale. This is partly because published rates tend to vary inversely with the length of the measurement interval and because they are uninformative unless they are viewed in the context of a concurrent ecological process. Hairston *et al.* (2005, p. 1117) propose that "evolution is rapid… only if the heritable phenotypic change occurs sufficiently quickly to alter the trajectory of an ecological process while it is still in progress." They further propose a method to partition the rate of change in an ecological variable of interest into components due to evolutionary and ecological causes. The relative contribution from evolutionary change can thus be objectively determined.

The joint consideration of evolution and demography departs from traditional evolutionary analyses which model variation in fitness among genotypes or phenotypes as their relative per-capita contributions to the next generation such that the mean fitness is one (Saccheri and Hanski, 2006). In other words, the population size is assumed to be constant and also large enough so that stochastic effects can be ignored. In an eco-evolutionary context, population size and its change over time are products of the absolute rates of survival and reproduction of individuals (Antonovics, 1976). To explore how individual variation affects population growth in this context, Coulson *et al.* (2006) suggested a theoretical framework that captures the entire process from the action of individual genes to the rate of population increase (or decrease) through a hierarchy of three "maps": genotype → phenotype, phenotype → demography, demography → population growth. The mechanistic or statistical mapping rules determine how sensitive variation at one level is to perturbations at the level below. The framework is intended for theoretical exploration and, in particular, as a tool to identify levels of detail that can be justifiably suppressed in a given situation.

Outlook

The interface between ecology and genetics has long been considered to be a source of important insight for both "parent disciplines." But for a long time, the field was represented by a collection of intriguing case studies with diverse conceptual affiliations. Emboldened by the success of earlier pioneering studies and equipped with new tools for data collection and analysis, researchers have since brought ecological genetics into focus as an integrated discipline. Patterns have emerged from the wealth of accumulated information and these have led to new questions. A selection based on this article is summarized here.

Why do we often not see a selection response when it is predicted from the breeder's equation? What are the relative contributions of fluctuating selection, counter-gradient selection, environmental trait/fitness covariance and genetic correlations (Merilä *et al.*, 2001) to the solution of this puzzle? Answers to these questions should help to clarify how natural selection in the wild differs from selection experiments in the lab and how organismal integration affects the true amount of genetic variation on which selection can act. Such answers should also shed light on two important evolutionary issues: the maintenance of genetic variation and the question to what extent microevolutionary processes determine macroevolutionary patterns. Finally, we need to understand to what extent microevolutionary potential and phenotypic plasticity reduce the risk of extinction due to global change.

What are the critical factors that determine whether a given species' ongoing adaptive change triggers eco-evolutionary dynamics? It is worth considering that two of the three case studies (cf. Eco-Evolutionary Dynamics) involved adaptation by clonal selection in strong pairwise species interactions, i.e., under conditions that are most favorable for eco-evolutionary dynamics. In the Soay sheep, the selection strengths varied significantly with density but were small in absolute value. The potential implications of constantly evolving species for community organization and stability are profound, but at present it is not known what proportion of microevolutionary change simply dissipates without any knock-on effects.

Methodological advances will help to generate the high-resolution data sets with which to address these questions, especially as dense linkage maps become available even for nonmodel species. These maps allow, for example, for direct estimation of relatedness without a pedigree. Adaptation can be monitored at the molecular level by mapping QTLs for a trait and then following allele frequency changes at linked marker loci in nature. Moreover, breeding values can be estimated from such marker data so that microevolution can be distinguished from purely phenotypic change that is driven by environmental trends (Kruuk *et al.*, 2008).

Finally, progress depends as ever on painstaking and long-term field research. Even in the age of high-throughput DNA sequencing and genotyping, the importance of monitoring carefully selected environmental variables in the "messy" real

world and of knowing the natural history of the study organism is undiminished (Rausher, 2005, Conner, 2010).

Appendix

List of Courses

1. General Ecology
2. Conservation Genetics
3. Evolution

See also: Adaptation. Coevolution. Conservation Genetics. Differentiation. Evolution, Theory of. Inbreeding and Outbreeding. Insecticide Resistance. Landscape Genetics. Life History, Evolution of. Phenological Shifts in Animals Under Contemporary Climate Change. Phenotype, a Historical Perspective. Population Genetics. Speciation, Process of

References

Agrawal AA (2000) Plasticity in the interactions and evolution of species. *Science* 294: 321–326.

Albert AYK, Sawaya S, Vines TH, *et al.* (2008) The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size. *Evolution* 62: 76–85.

Antonovics J (1976) The input from population genetics: "The new ecological genetics". *Systematic Biology* 1: 233–245.

Antonovics J, Bradshaw AD, and Turner RG (1971) Heavy metal tolerance in plants. *Advances in Ecological Research* 7: 1–85.

Auld JR, Agrawal AA, and Relyea RA (2010) Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society of London B* 22: 503–511.

Barrett RDH, Rogers SM, and Schluter D (2008) Natural selection on a major armor gene in threespine stickleback. *Science* 322: 255–257.

Barrett SCH (1983) Crop mimicry in weeds. *Economic Botany* 37: 255–282.

Barton NH (1990) Pleiotropic models of quantitative variation. *Genetics* 124: 773–782.

Barton NH (1999) Clines in polygenic traits. *Genetical Research* 74: 223–236.

Beavis WD (1998) QTL analyses: Power, precision, and accuracy. In: Patterson AH (ed.) *Molecular Dissection of Complex Traits*, pp. 145–162. New York: CRC Press.

Bell G (2010) Fluctuating selection: The perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society London B* 365: 87–97.

Berrigan D and Scheiner SM (2004) Modeling the evolution of phenotypic plasticity. In: DeWitt TJ and Scheiner SM (eds.) *Phenotypic Plasticity – Functional and Conceptual Issues*, pp. 82–97.

Berven KA, Gill DE, and Smith-Gill SJ (1979) Countergradient selection in the green frog, *Rana clamitans*. *Evolution* 33: 609–623.

Blondel J, Dias PC, Perret P, Maistre M, and Lambrechts MM (1999) Selection-based biodiversity at a small scale in a low-dispersing insular bird. *Science* 285: 1399–1402.

Blows MW and Brooks R (2003) Measuring nonlinear selection. *American Naturalist* 162: 815–820.

Blows MW and Hoffmann AA (2005) A reassessment of genetic limits to evolutionary change. *Ecology* 86: 1371–1384.

Bradshaw AD (1991) Genostasis and the limits to evolution. *Philosophical Transactions of the Royal Society London B* 333: 289–305.

Brussard PF (1978) Preface. In: Brussard PF (ed.) *Ecological Genetics: The Interface*, pp. iii–iv. New York: Springer Verlag.

Bulmer MG (1972) Mulitple niche polymorphism. *American Naturalist* 106: 254–257.

Campbell DR (1997) Genetic and environmental variation in life-history traits of a monocarpic perennial: A decade-long field experiment. *Evolution* 51: 373–382.

Charlesworth B (1987) The heritability of fitness. In: Bradbury JW and Andersson MB (eds.) *Sexual Selection: Testing the Alternatives*, pp. 21–40. Chichester: John Wiley & Sons, Limited.

Charlesworth B (1990) Optimization models, quantitative genetics, and mutation. *Evolution* 44: 520–538.

Charmantier A and Garant D (2005) Environmental quality and evolutionary potential: Lessons from wild populations. *Proceedings of the Royal Society of London B* 272: 1415–1425.

Charmantier A, McCleery RH, Cole LR, Perrins C, Kruuk LEB, and Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* 320: 800–803.

Coberly LC and Rausher MD (2008) Pleiotropic effects of an allele producing white flowers in *Ipomoea purpurea*. *Evolution* 62: 1076–1085.

Coltman DW, Pilkington JG, Smith JA, and Pemberton JM (2005) Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution* 53: 1259–1267.

Connell JH and Sousa WP (1983) On the evidence needed to judge ecological stability or persistence. *American Naturalist* 121: 789–824.

Conner JK (2003) Artificial selection: A powerful tool for ecologists. *Ecology* 84: 1650–1660.

Conner JK (2006) Ecological genetics of floral evolution. In: Harder LD and Barrett SCH (eds.) *Ecology and Evolution of Flowers*, pp. 260–277. New York: Oxford University Press.

Conner JK (2010) Natural selection in plants 151 years after *The Origin*: Introduction. *International Journal of Plant Science* 171: 927–929.

Conner JK and Hartl DL (2004) *A Primer of Ecological Genetics*. Sunderland, MA: Sinauer Associates.

Conover DO and Schultz ET (1995) Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology and Evolution* 10: 248–252.

Cook LM and Turner JRG (2008) Decline of melanism in two British moths: Spatial, temporal and inter-specific variation. *Heredity* 101: 483–489.

Cook SA and Johnson MP (1968) Adaptation to heterogeneous environments. I. Variation in heterophylly in *Ranunculus flammula* L. *Evolution* 22: 496–516.

Coulson T, Benton TG, Lundberg P, Dall SRX, and Kendall BE (2006) Putting evolutionary biology back in the ecological theatre: A demographic framework mapping genes to communities. *Evolutionary Ecology Research* 8: 1155–1171.

Coulson T, Catchpole EA, Albon SD, *et al.* (2001) Age, sex, density, winter weather, and population crashes in Soay sheep. *Science* 292: 1528–1531.

DeWitt TJ, Sih A, and Wilson DS (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution* 13: 77–81.

Dias PC and Blondel J (1996) Local specialization and maladaptation in the Mediterranean blue tit (*Parus caeruleus*). *Oecologia* 107: 79–86.

Donohue K, Messiqua D, Pyle EH, Heschel MS, and Schmitt J (2000) Evidence of adaptive divergence in plasticity: Density- and site-dependent selection on shade-avoidance responses in *Impatiens capensis*. *Evolution* 54: 1956–1968.

Donohue K, Pyle EH, Messiqua D, Heschel MS, and Schmitt J (2001) Adaptive divergence in plasticity in natural populations of *Impatiens capensis* and its consequences for performance in novel habitats. *Evolution* 55: 692–702.

Dorn LA, Pyle EH, and Schmitt J (2000) Plasticity to light cues and resources in *Arabidopsis thaliana*: Testing for adaptive value and costs. *Evolution* 54: 1982–1994.

Duffy MA and Sivars-Becker L (2007) Rapid evolution and ecological host-parasite dynamics. *Ecology Letters* 10: 44–53.

Dwyer G, Levin SA, and Buttel L (1990) A simulation model of the population dynamics and evolution of myxomatosis. *Ecological Monographs* 60: 423–447.

Ellegren H and Sheldon BC (2008) The genetic basis of fitness differences in natural populations. *Nature* 452: 169–175.

Ellner SP, Hairston Jr. NG, Kearns CM, and Babai D (1999) The roles of fluctuating selection and long-term diapause in microevolution of diapause timing in a freshwater copepod. *Evolution* 53: 111–122.

Endler JA (1980) Natural selection on color patterns in *Poecilia reticulata*. *Evolution* 34: 76–91.

Endler JA (1986) *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.

Estes S and Arnold SJ (2007) Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence on all timescales. *American Naturalist* 169: 227–244.

Falconer DS and Mackay TFC (1996) *Introduction to Quantitative Genetics*. Harlow: Longman.

Feder JL, Berlocher SH, Roethele JB, *et al.* (2003) Allopatric genetic origins for sympatric host shifts and race formation in *Rhagoletis*. *Proceedings of the*

National Academy of Sciences of the United States of America 100: 10314–10319.
Feder JL, Roethele JB, Filchak KE, Niedbalski J, and Romero-Severson J (2003) Evidence for an inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* 163: 939–953.
ffrench-Constant RH, Daborn PJ, and Le Goff G (2004) The genetics and genomics of insecticide resistance. *Trends in Genetics* 20: 163–170.
Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH, and Kruuk LEB (2007) Sexually antagonistic genetic variation for fitness in red deer. *Nature* 447: 1107–1111.
Ford EB (1964) *Ecological Genetics*. London: Methuen & Co Ltd.
Forde SE, Thompson JN, and Bohannan BJ (2004) Adaptation varies through space and time in a coevolving host–parasitoid system. *Nature* 431: 841–844.
Fry JD (2003) Detecting ecological trade-offs using selection experiments. *Ecology* 84: 1672–1678.
Fussmann GF, Loreau M, and Abrams PE (2007) Eco-evolutionary dynamics of communities and ecosystems. *Functional Ecology* 21: 465–477.
Gandon S (2002) Local adaptation and the geometry of host–parasite coevolution. *Ecology Letters* 5: 246–256.
Geber MA (2011) Ecological and evolutionary limits to species geographic ranges. *American Naturalist* 178: S1–S5.
Ghalambor CK, Mckay JK, Carroll SP, and Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21: 394–407.
Gingerich PD (1983) Rates of evolution: Effects of time and temporal scaling. *Science* 222: 159–161.
Grant PR and Grant BR (1989) *Evolutionary Dynamics of a Natural Population: The Large Cactus Finch of the Galápagos*. Chicago: The University of Chicago Press.
Grant PR and Grant BR (1995) Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49: 241–251.
Grant PR and Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296: 707–711.
Gratten J, Wilson AJ, McRae AF, *et al.* (2008) A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science* 319: 318–320.
Hadfield JD (2008) Estimating evolutionary parameters when viability selection is operating. *Proceedings of the Royal Society of London B* 275: 723–734.
Hairston Jr. NG and Dillon TA (1990) Fluctuating selection and response in a population of freshwater copepods. *Evolution* 44: 1796–1805.
Hairston Jr. NG, Ellner SP, Geber MA, Yoshida T, and Fox JA (2005) Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters* 8: 1114–1127.
Haldane JBS (1924) A mathematical theory of natural and artificial selection. *Transactions of the Cambridge Philosophical Society* 23: 19–40.
Harrison RG (1980) Dispersal polymorphisms in insects. *Annual Review of Ecology and Systematics* 11: 95–118.
Hawthorne DJ and Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412: 904–907.
Hendry AP, Farrugia TJ, and Kinnison MT (2008) Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology* 17: 20–29.
Hendry AP and Kinnison MT (1999) The pace of modern life: Measuring rates of contemporary microevolution. *Evolution* 53: 1637–1653.
Hendry AP and Taylor EB (2004) How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake–stream stickleback pairs. *Evolution* 58: 2319–2331.
Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist* 173: 579–588.
Hereford J, Hansen TF, and Houle D (2004) Comparing strengths of directional selection: How strong is strong? *Evolution* 58: 2143–2153.
Hersch EI and Phillips PC (2004) Power and potential bias in field studies of natural selection. *Evolution* 58: 479–485.
Hill WG, Goodard ME, and Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. *PloS Genetics* 4: e1000008.
Hoekstra HE, Hoekstra JM, Berrigan D, *et al.* (2001) Strength and tempo of directional selection in the wild. *Proceedings of the National Academy of Sciences of the United States of America* 98: 9157–9160.
Hoffmann AA and Merilä J (1999) Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology and Evolution* 14: 96–101.
Hoffmann AA and Parsons PA (1991) *Evolutionary Genetics and Environmental Stress*. Oxford: Oxford University Press.
Houle D (1991) Genetic covariances of fitness correlates: What genetic correlations are made of and why it matters. *Evolution* 45: 630–648.
Houle D (1992) Comparing evolvability and variability of quantitative traits. *Genetics* 130.
Houle D, Morikawa B, and Lynch M (1996) Comparing mutational variabilities. *Genetics* 143: 1467–1483.
Huey RB, Gilchrist GW, Carlson ML, Berrigan D, and Serra L (2000) Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287: 308–309.
de Jong G (2005) Evolution of phenotypic plasticity: Patterns of plasticity and the emergence of ecotypes. *New Phytologist* 166: 101–118.
Jones LE, Becks L, Ellner SP, Hairston Jr. NG, Yoshida T, and Fussmann GF (2009) Rapid contemporary evolution and clonal food web dynamics. *Philosophical Transactions of the Royal Society London B* 364: 1579–1591.
Kawecki TJ and Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters* 7: 1225–1241.
Kingsolver J, Hoekstra HE, Hoekstra JM, *et al.* (2001) The strength of phenotypic selection in natural populations. *American Naturalist* 157: 245–261.
Kingsolver J and Pfennig DW (2007) Patterns and power of phenotypic selection in nature. *BioScience* 57: 561–572.
Knapczyk FN and Conner JK (2007) Estimates of the average strength of natural selection are not inflated by sampling error or publication bias. *American Naturalist* 170: 501–508.
Kokko H and Lopéz-Sepulcre A (2007) The ecogenetic link between demography and evolution: Can we bridge the gap between theory and data? *Ecology Letters* 10: 773–782.
Kruuk LEB (2004) Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society London B* 359: 873–890.
Kruuk LEB, Clutton-Brock TH, Slate J, Pemberton JM, Brotherstone S, and McGuiness FE (2000) Heritability of fitness in a wild mammal population. *Proceedings of the National Academy of Sciences of the United States of America* 97: 698–703.
Kruuk LEB, Merilä J, and Sheldon BC (2003) When environmental variation short-circuits natural selection. *Trends in Ecology and Evolution* 18: 207–209.
Kruuk LEB, Slate J, Pemberton JM, Brotherstone S, Guinness FE, and Clutton-Brock TH (2002) Antler size in red deer: Heritability and selection but no evolution. *Evolution* 56: 1683–1695.
Kruuk LEB, Slate J, and Wilson AJ (2008) New answers for old questions: The evolutionary quantitative genetics of wild animal populations. *Annual Review of Ecology, Evolution, and Systematics* 39: 525–548.
Lande R (1979) Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. *Evolution* 33: 402–416.
Lande R (1982) A quantitative genetic theory of life history evolution. *Ecology* 63: 607–615.
Lande R and Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* 17: 386–391.
Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution* 183–189.
Levene H (1953) Genetic equilibrium when more than one ecological niche is available. *American Naturalist* 87: 331–333.
Losos JB, Creer DA, Glossip D, *et al.* (2000) Evolutionary implications of phenotypic plasticity in the hindlimb of the lizard *Anolis sagrei*. *Evolution* 54: 301–305.
Losos JB, Wahrheit KI, and Schoener TW (1997) Adaptive differentiation following experimental island colonizations in *Anolis* lizards. *Nature* 73.
Lowe A, Harris S, and Ashton P (2004) *Ecological Genetics: Design, Analysis, and Application*. Oxford: Blackwell Science Ltd.
Lynch M and Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer.
Lyytinen A, Brakefield PM, Lindström L, and Mappes J (2004) Does predation maintain eyespot plasticity in *Bicyclus anaynana*? *Proceedings of the Royal Society of London B* 271: 179–283.
Mackay TFC (2001) Quantitative trait loci in *Drosophila*. *Nature Reviews Genetics* 2: 11–20.
Macnair MR (1991) Why the evolution of resistance to anthropogenic toxins normally involves major gene changes: The limits to natural selection. *Genetica* 84: 213–219.
Macnair MR (1993) Tansley review no. 49: The genetics of metal tolerance in vascular plants. *New Phytologist* 124: 541–559.
Majerus MEN (2009) Industrial melanism in the peppered moth, *Biston betularia*: An excellent teaching example of Darwinian evolution in action. *Evolution: Education and Outreach* 2: 63–74.

Mauricio R (2001) Mapping quantitative trait loci in plants: Uses and caveats for evolutionary biology. *Nature Reviews Genetics* 2: 370–381.

McCleery RH, Pettifor RA, Armbruster P, Meyer K, Sheldon BC, and Perrins CM (2004) Components of variance underlying fitness in a natural population of the Great Tit *Parus major*. *American Naturalist* 164: E62–E72.

Merilä J, Laurila A, and Lindgren B (2004) Variation in the degree and costs of adaptive phenotypic plasticity among *Rana temporaria* populations. *Journal of evolutionary Biology* 17: 1132–1140.

Merilä J and Sheldon BC (2000) Lifetime reproductive success and heritability in nature. *American Naturalist* 155: 301–310.

Merilä J, Sheldon BC, and Kruuk LEB (2001) Explaining stasis: Microevolutionary studies in natural populations. *Genetica* 112 – 113: 199–222.

Meyer JR, Ellner SP, Hairston Jr. NG, Jones LE, and Yoshida T (2006) Prey evolution on the time scale of predator–prey dynamics revealed by allele-specific quantitative PCR. *Proceedings of the National Academy of Sciences of the United States of America* 103: 10690–10695.

Mitchell-Olds T and Shaw RG (1987) Regression analysis of natural selection: Statistical inference and biological interpretation. *Evolution* 41: 1149–1161.

Morgan AD, Gandon S, and Buckling A (2005) The effect of migration on local adaptation in a coevolving host–parasite system. *Nature* 437: 253–256.

Morgan MT and Conner JK (2001) Using genetic markers to directly estimate male selection gradients. *Evolution* 55: 272–281.

Morjan CL and Rieseberg LH (2004) How species evolve collectively: Implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* 13: 1341–1356.

Morrissey MB and Hadfield JD (2012) Directional selection in temporally replicated studies is remarkably consistent. *Evolution* 66: 435–442.

Morrissey MB, Kruuk LEB, and Wilson AJ (2010) The danger of applying the breeder's equation in observational studies of natural populations. *Journal of evolutionary Biology* 23: 2277–2288.

Mousseau TA and Roff DA (1987) Natural selection and the heritability of fitness components. *Heredity* 59: 181–197.

Neuhauser C, Andow DA, Heimpel GE, May G, Shaw RG, and Wagenius S (2003) Community genetics: Expanding the synthesis of ecology and genetics. *Ecology* 84: 545–558.

Nussey DH, Postma E, Gienapp P, and Visser ME (2005) Selection on heritable plasticity in a wild bird population. *Science* 310: 304–306.

Nussey DH, Wilson AJ, and Brommer JE (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology* 20: 831–844.

Orr HA (2005) The genetic theory of adaptation: A brief history. *Nature Reviews Genetics* 6: 119–127.

Pelletier F, Clutton-Brock TH, Pemberton JM, Tuljapurkar S, and Coulson T (2007) The evolutionary demography of ecological change: Linking trait variation and population growth. *Science* 315: 1571–1574.

Pelletier F, Garant D, and Hendry AP (2009) Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society London B* 364: 1483–1489.

Pfennig DW, Rice AM, and Martin RA (2007) Field and experimental evidence for competition's role in phenotypic divergence. *Evolution* 61: 257–271.

Pigliucci M (2005) Evolution of phenotypic plasticity: Where are we going now? *Trends in Ecology and Evolution* 20: 481–486.

Post DM and Palcovacs EP (2009) Eco-evolutionary feedbacks in community and ecosystem ecology: Interactions between the ecological theatre and the evolutionary play. *Philosophical Transactions of the Royal Society London B* 364: 1629–1640.

Postma E (2006) Implications of the difference between true and predicted breeding values for the study of natural selection and micro-evolution. *Journal of Evolutionary Biology* 19: 309–320.

Price G (1970) Selection and covariance. *Nature* 227: 520–521.

Price T, Kirkpatrick M, and Arnold SJ (1988) Directional selection and the evolution of breeding date in birds. *Science* 240: 798–799.

Price TD, Qvarnström A, and Irwin DE (2003) The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London B* 270: 1433–1440.

Räsänen K and Kruuk LEB (2007) Maternal effects and evolution at ecological time-scales. *Functional Ecology* 21: 408–421.

Rausher MD (1992) The measurement of selection on quantitative traits: Biases due to environmental covariances between traits and fitness. *Evolution* 46: 616–626.

Rausher MD (2005) Plant evolutionary ecology. *New Phytologist* 165: 2–5.

Reznick D, Shaw FH, Rodd FH, and Shaw RG (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* 275: 1934–1937.

Riechert SE (1993) Investigation of potential gene flow limitation of behavioral adaptation in an aridlands spider. *Behavioral Ecology and Sociobiology* 32: 355–363.

Riechert SE, Singer FD, and Jones TC (2001) High gene flow levels lead to gamete wastage in a desert spider system. *Genetica* 112-113: 297–319.

Robertson A (1966) A mathematical model of the culling process in dairy cattle. *Animal Production* 8: 95–108.

Roelofs D, Overhein L, de Boer ME, Janssens TKS, and van Straalen NM (2006) Additive genetic variation of transcriptional regulation: Metallothionein expression in the soil insect *Orchesella cincta*. *Heredity* 96: 85–92.

Roff DA (1986) The evolution of wing dimorphism in insects. *Evolution* 42: 717–728.

Roff DA (1997) *Evolutionary Quantitative Genetics*. New York: Chapman & Hall.

Roff DA (2007) A centennial celebration for quantitative genetics. *Evolution* 61: 1017–1032.

Saccheri I and Hanski I (2006) Natural selection and population dynamics. *Trends in Ecology and Evolution* 21: 341–347.

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8: 552–560.

Scheiner SM (2002) Selection experiments and the study of phenotypic plasticity. *Journal of Evolutionary Biology* 15: 889–898.

Schluter D, Marchinko KB, Barrett RDH, and ROGERS SM (2010) Natural selection and the genetics of adaptation in threespine stickleback. *Philosophical Transactions of the Royal Society London B* 2010: 2479–2486.

Schmitt J, Stinchcombe JR, Heschel M, and Huber H (2003) The adaptive evolution of plasticity: Phytochrome-mediated shade avoidance responses. *Integrative and Comparative Biology* 43: 459–469.

Sgro CM and Partridge L (1999) A delayed wave of death from reproduction in *Drosophila*. *Science* 286: 2521–2524.

Siepielski AM, DiBattista JD, and Carlson SM (2009) It's about time: The temporal dynamics of phenotypic selection among fitness components in the wild. *Ecology Letters* 12: 1–16.

Sinervo B (1999) Mechanistic analysis of natural selection and a refinement of Lack's and Williams's principles. *American Naturalist* 154: S26–S42.

Slate J, Santure AW, Feulner PGD, *et al.* (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics* 26: 275–284.

Slatkin M (1973) Gene flow and selection in a cline. *Genetics* 75: 733–756.

Slatkin M (1978) Spatial patterns in the distribution of polygenic characters. *Journal of Theoretical Biology* 70: 213–228.

Spichtig M and Kawecki TJ (2004) The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *American Naturalist* 164: 70–84.

Stinchcombe JR, Dorn LA, and Schmitt J (2004) Flowering time plasticity in *Arabidopsis thaliana*: A reanalysis of Westerman & Lawrence (1970). *Journal of evolutionary Biology* 17: 197–207.

Stinchcombe JR, Rutter MT, Burdick DS, Tiffin P, Rausher MD, and MAURICIO R (2002) Testing for environmentally induced bias in phenotypic estimates of natural selection: Theory and practice. *American Naturalist* 160: 511–523.

Storfer A and Sih A (1998) Gene flow and ineffective antipredator behavior in a stream-breeding salamander. *Evolution* 52: 558–565.

Strauss SY, Lau JA, and Carroll SP (2006) Evolutionary responses of natives to introduced species: What do interactions tell us about natural communities. *Ecology Letters* 9: 357–374.

Sultan SE and Spencer HG (2002) Metapopulation structure favors plasticity over local adaptation. *American Naturalist* 160: 271–283.

Teplitsky C, Mills JA, Yarrall JW, and Merila J (2009) Heritability of fitness components in a wild bird population. *Evolution* 63: 716–726.

Thompson JN (1998) Rapid evolution as an ecological process. *Trends in Ecology and Evolution* 13: 329–332.

Thompson JN (2005) *The Geographic Mosaic of Coevolution*. Chicago, IL: University of Chicago Press.

Thrall PH and Burdon JJ (2003) Evolution of virulence in a plant host–pathogen metapopulation. *Science* 299: 1735–1737.

Tollrian R and Dodson SI (1999) Inducible defenses in *Cladocera*: Constraints, costs and multipredator environments. In: Tollrian R and Harvell CD (eds.) *The Ecology and Evolution of Inducible Defenses*, pp. 177–202. Princeton: Princeton University Press.

van Kleunen M and Fischer M (2005) Constraints on the evolution of adaptive phenotypic plasticity in plants. *New Phytologist* 166: 49–60.

Van Noordwijk AJ and de Jong G (1986) Acquisition and allocation of resources: Their influence on variation in life history traits. *American Naturalist* 128: 137–142.

Via S, Gomulkiewicz R, de Jong G, Scheiner SM, Schlichting CD, and Tienderen PHv (1995) Adaptive phenotypic plasticity: Consensus and controversy. *Trends in Ecology and Evolution* 10: 212–217.

Via S and Lande R (1985) Genotype–environment interaction and the evolution of phenotypic plasticity. *Evolution* 39: 505–522.

Walsh B and Blows MW (2009) Abundant genetic variation + strong selection = multivariate genetic constraints: A geometric view of adaptation. *Annual Review of Ecology, Evolution, and Systematics* 40: 41–59.

Watt WB, Cassin RC, and Swan MS (1983) Adaptation at specific loci. III. Field behaviour and survivorship differences among *Colias* PGI genotypes are predictable from *in vitro* biochemistry. *Genetics* 103: 725–729.

West-Eberhard MJ (2003) *Developmental Plasticity and Evolution*. New York: Oxford University Press.

Wheat CW, Watt WB, Pollock DD, and Schulte PM (2006) From DNA to fitness differences: sequences and structures of adaptive variants of *Colias* phosphoglucose isomerase (PGI). *Molecular Biology and Evolution* 23: 499–512.

Williams GC (1992) *Natural Selection: Domains, Levels and Challenges*. Oxford: Oxford University Press.

Wilson AJ, Pemberton JM, Pilkington JG, *et al.* (2006) Environmental coupling of selection and heritability limits evolution. *PLoS Biology* 4: e216.

Wilson AJ, Réale D, Clements MN, *et al.* (2010) An ecologist's guide to the animal model. *Journal of Animal Ecology* 79: 13–26.

Worley A, Houle D, and Barrett SCH (2003) Consequences of hierarchical allocation for the evolution of life-history. *American Naturalist* 161: 153–167.

Yoshida T, Jones LE, Ellner SP, Fussmann GF, and Hairston Jr. NG (2003) Rapid evolution drives ecological dynamics in a predator–prey system. *Nature* 424: 303–306.

Zera AJ and Harshman LG (2001) The physiology of life history trade-offs in animals. *Annual Review of Ecology and Systematics* 32: 95–126.

Zheng C, Ovaskainen O, and Hanski I (2009) Modelling single nucleotide effects on *phosphoglucose isomerase* on dispersal in the Glanville fritillary butterfly: Coupling of ecological and evolutionary dynamics. *Philosophical Transactions of the Royal Society London B* 364: 1519–1532.

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF
Simon A Levin
Princeton University, Princeton, NJ, USA

VOLUME 3



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xv

Contributors to Volume 3 xxi

Foreword to the First Edition xxv

Preface to the First Edition xxvii

Preface to the Second Edition xxix

How to use of the Encyclopedia xxxi

VOLUME 3

Ecology, Concept and Theories in 1
Peter Kareiva and Michelle Marvier

Ecology of Agriculture 9
Alison G Power

Economic Control of Invasive Species 16
David Finnoff, Jason F Shogren, Richard D Horan, Shana M McDermott, and Charles Sims

Economic Growth and the Environment 25
Karl-Göran Mäler

Economics of Agrobiodiversity 31
Unai Pascual, Louise E Jackson, and Adam G Drucker

Economics of the Regulating Services 45
Edward B Barbier

Economic Value of Biodiversity, Measurements of 55
Robert Mendelsohn and Seth Binder

Ecosystem, Concept of 59
Eugene P Odum

Ecosystem Function Measurement, Aquatic and Marine Communities 64
John T Lehman

Ecosystem Function Measurement, Terrestrial Communities 72
Sandra Díaz

Ecosystem Function, Principles of 90
Ross A Virginia and Diana H Wall

Ecosystem Services 96
Heather Tallis, Anne Guerry, and Gretchen C Daily

Ecosystems of South America 105
Luis Anibal Solorzano Cardenas

Ecotoxicology 118
JM Lynch, A Wiseman, and FAAM De Leij

Edible Plants 127
Eduardo H Rapoport and Barbara S Drausal

Education and Biodiversity 133
Shirley M Malcom

Effects of Ecotones on Biodiversity 142
Salit Kark

Elevational Trends in Biodiversity 149
John-Arvid Grytnes and Christy M McCain

El Niño and Biodiversity 155
Emily Meuser, Arne Ø Mooers, and Daniel FR Cleary

Endangered Amphibians 164
Tim Halliday

Endangered Ecosystems 169
Raymond C Nias

Endangered Freshwater Invertebrates 176
David L Strayer

Endangered Mammals 188
Peter Zahler and Tatjana Rosen

Endangered Marine Invertebrates 199
James T Carlton

Endangered Plants 205
Thomas J Stohlgren and Sunil Kumar

Endangered Reptiles 216
Tim Halliday

Endangered Terrestrial Invertebrates 219
Piotr Naskrecki

Endemism 228
RM Cowling

Energy Flow and Ecosystems 237
Alan P Covich

Energy Use, Human 250
Patrick Gonzalez

Environmental Ethics 267
Philip J Cafaro and Richard B Primack

Environmental Impact, Concept and Measurement of 278
Ellen W Chu and James R Karr

Estuarine Ecosystems 297
G Carleton Ray and Jerry McCormick-Ray

Ethical Issues in Biodiversity Protection 309
Philip J Cafaro and Richard B Primack

Ethnobiology and Ethnoecology 319
Gary J Martin

Eukaryotes, Origin of 329
B DeRennaux

Europe, Ecosystems of 333
Ladislav Mucina

Eutrophication and Oligotrophication 347
JoAnn M Burkholder and Patricia M Glibert

Evaluation of Ecosystem Service Policies from Biophysical and Social Perspectives: The Case of China 372
Jianguo Liu, Zhiyun Ouyang, Wu Yang, Weihua Xu, and Shuxin Li

Evolution in Response to Climate Change 385
Julie R Etterson and Ruth G Shaw

Evolution, Theory of 392
Gregory C Mayer and Catherine L Craig

Extinction, Causes of 401
Richard B Primack and Rachel A Morrison

Extinction in the Fossil Record 413
Jeffrey S Levinton

F

Feeding the World and Protecting Biodiversity 426
Paul C West, Reinette Biggs, Bruce A McKenney, and Chad Monfreda

Fires, Ecological Effects of 435
William John Bond

Fish Conservation 443
Carl Safina and Alan Duckworth

Fishes, Biodiversity of 456
Gene S Helfman

Fish Stocks 477
Rainer Froese and Daniel Pauly

Flies, Gnats, and Mosquitoes 488
Brian V Brown

Food Webs 497
Gary R Huxel and Gary A Polis

Forest Canopies, Animal Diversity 511
Terry L Erwin

Forest Canopies, Plant Diversity 516
Nalini M Nadkarni, Mark C Merwin, and Jurgen Nieder

Forest Ecology 528
Timothy J Fahey

Fossil Record 537
Sean R Connolly and Matthew A Kosnik

Framework for Assessment and Monitoring of Biodiversity 545
Francisco Dallmeier, Robert C Szaro, Alfonso Alonso, James Comiskey, and Ann Henderson

Freshwater Ecosystems 560
Robert G Wetzel

Freshwater Ecosystems, Human Impact on 570
Kaj Sand-Jensen

Functional Diversity 587
David Tilman

Functional Diversity Measures 597
Norman WH Mason and David Mouillot

Functional Groups 609
Robert S Steneck

Fungi 624
Thomas J Volk

G

Gene Banks 641
Simon H Linington and Hugh W Pritchard

Genes, Description of 654
Michael F Antolin and William C Black IV

Genetic Diversity 662
Eviatar Nevo

Geologic Time, History of Biodiversity in 678
James W Valentine

Global Declines of Amphibians 691
Kellie Whittaker, Michelle S Koo, David B Wake, and Vance T Vredenburg

Globalization Effects on Common Plant Species 700
Thomas J Stohlgren, Petr Pyšek, John Kartesz, Misako Nishino, Aníbal Pauchard, Marten Winter, Joan Pino, David M Richardson, John Wilson, Brad R Murray, Megan L Phillips, Laura Celesti-Grapow, and Jim Graham

Global Species Richness 707
Kevin J Gaston

Government Legislation and Regulations in the United States 712
Kathryn A Saterson

Grasshoppers and their Relatives 722
Piotr Naskrecki

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 3

Alfonso Alonso
Smithsonian Conservation Biology Institute, Washington, DC, USA

Michael F Antolin
Colorado State University, Fort Collins, CO, USA

Edward B Barbier
University of Wyoming, Laramie, WY, USA

Reinette Biggs
Stockholm University, Stockholm, Sweden

Seth Binder
University of Minnesota, St Paul, MN, USA, and Yale University, New Haven, CT, USA

William C Black IV
Colorado State University, Fort Collins, CO, USA

William John Bond
University of Cape Town, Rondebosch, South Africa

Brian V Brown
Natural History Museum of Los Angeles County, Los Angeles, CA, USA

JoAnn M Burkholder
North Carolina State University, Raleigh, NC, USA

Philip J Cafaro
Colorado State University, Fort Collins, CO, USA

Luis Anibal Solorzano Cardenas
Gordon and Betty Moore Foundation, CA, USA

G Carleton Ray
University of Virginia, Charlottesville, VA, USA

James T Carlton
Williams College and Mystic Seaport, Mystic, CT, USA

Laura Celesti-Grapow
Università La Sapienza, Rome, Italy

Ellen W Chu
University of Washington, Seattle, WA, USA

Daniel FR Cleary
University of Aveiro, Aveiro, Portugal

James Comiskey
National Park Service, Fredericksburg, VA, USA

Sean R Connolly
James Cook University, Townsville, QLD, Australia

Alan P Covich
University of Georgia, Athens, GA, USA

RM Cowling
University of Cape Town, Rondebosch, South Africa

Catherine L Craig
Harvard University, Cambridge, MA, USA

Gretchen C Daily
Stanford University, Stanford, CA, USA

Francisco Dallmeier
Smithsonian Conservation Biology Institute, Washington, DC, USA

FAAM De Leij
University of Surrey, Guildford, UK

B DeRennaux
University of Tennessee, Knoxville, TN, USA

Sandra Díaz
Universidad Nacional de Córdoba – CONICET, Córdoba, Argentina

Barbara S Drausal
Universidad Nacional del Comahue, Bariloche, Argentina

Adam G Drucker
Bioversity International, Rome, Italy

Alan Duckworth
Blue Ocean Institute, Cold Spring Harbor, NY, USA

Terry L Erwin
Smithsonian Institution, Washington, DC, USA

Julie R Etterson
University of Minnesota, Duluth, MN, USA

Timothy J Fahey
Cornell University, Ithaca, NY, USA

David Finnoff
University of Wyoming, Laramie, WY, USA

Rainer Froese
IFM-GEOMAR, Kiel, Germany

Kevin J Gaston
University of Exeter, Exeter, UK, and University of Sheffield, Sheffield, UK

Patricia M Glibert
University of Maryland Center for Environmental Science, Cambridge, MD, USA

Patrick Gonzalez
University of California, Berkeley, CA, USA

Jim Graham
Colorado State University, Fort Collins, CO, USA

John-Arvid Grytnes
University of Bergen, Bergen, Norway

Anne Guerry
Stanford University, Stanford, CA, USA

Tim Halliday
Open University, Milton Keynes, UK

Gene S Helfman
University of Georgia, Athens, GA, USA

Ann Henderson
National Geographic Channel, Washington, DC, USA

Richard D Horan
Michigan State University, East Lansing, MI, USA

Gary R Huxel
University of Arkansas, Fayetteville, AR, USA

Louise E Jackson
University of California, Davis, CA, USA

Peter Kareiva
National Marine Fisheries Service, Seattle, WA, USA

Salit Kark
The Hebrew University of Jerusalem, Jerusalem, Israel, and University of Queensland, Brisbane, QLD, Australia

James R Karr
University of Washington, Seattle, WA, USA

John Kartesz
Biota of North America Program, Chapel Hill, NC, USA

Michelle S Koo
University of California, Berkeley, CA, USA

Matthew A Kosnik
Macquarie University, North Ryde, NSW, Australia

Sunil Kumar
Colorado State University, Fort Collins, CO, USA

John T Lehman
University of Michigan, Ann Arbor, MI, USA

Jeffrey S Levinton
State University of New York, Stony Brook, NY, USA

Shuxin Li
Michigan State University, East Lansing, MI, USA

Simon H Linington
Royal Botanic Gardens, London, UK

Jianguo Liu
Michigan State University, East Lansing, MI, USA

JM Lynch
University of Surrey, Guildford, UK

Shirley M Malcom
American Association for the Advancement of Science (AAAS), Washington, DC, USA

Karl-Göran Mäler
Beijer International Institute of Ecological Economics, Stockholm, Sweden

Gary J Martin
The Global Diversity Foundation, Marrakech, Morocco

Michelle Marvier
Santa Clara University, Santa Clara, CA, USA

Norman WH Mason
Landcare Research, Hamilton, New Zealand

Gregory C Mayer
University of Wisconsin-Parkside, Kenosha, WI, USA

Christy M McCain
University of Colorado at Boulder, Boulder, CO, USA

Jerry McCormick-Ray
University of Virginia, Charlottesville, VA, USA

Shana M McDermott
Dartmouth College, Hanover, NH, USA

Bruce A McKenney
The Nature Conservancy, Charlottesville, VA, USA

Robert Mendelsohn
Yale University, New Haven, CT, USA

Mark C Merwin
The Evergreen State College, Olympia, WA, USA

Emily Meuser
Simon Fraser University, Burnaby, BC, Canada

Chad Monfreda
Arizona State University, Tempe, AZ, USA

Arne Ø Mooers
Simon Fraser University, Burnaby, BC, Canada

Rachel A Morrison
Institution of Oceanography, La Jolla, CA, USA

David Mouillot
James Cook University, Townsville, Queensland, Australia

Ladislav Mucina
Curtin University, Perth, Australia

Brad R Murray
University of Technology Sydney, Sydney, NSW, Australia

Nalini M Nadkarni
The Evergreen State College, Olympia, WA, USA

Piotr Naskrecki
Harvard University, Cambridge, MA, USA

Eviatar Nevo
University of Haifa, Haifa, Israel

Raymond C Nias
TierraMar Consulting, Sutherland, NSW, Australia

Jurgen Nieder
Botanische Institüt der Universität Bonn, Bonn, Germany

Misako Nishino
Biota of North America Program, Chapel Hill, NC, USA

Eugene P Odum
University of Georgia, Athens, GA, USA

Zhiyun Ouyang
Chinese Academy of Sciences, Beijing, China

Unai Pascual
University of Cambridge, Cambridge, UK; Basque Centre for Climate Change (BC3), Bilbao, Spain, and IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Aníbal Pauchard
Universidad de Concepción and Institute for Ecology and Biodiversity (IEB), Concepción, Chile

Daniel Pauly
University of British Columbia, Vancouver, BC, Canada

Megan L Phillips
University of Technology Sydney, Sydney, NSW, Australia

Joan Pino
University of Barcelona, Barcelona, Spain

Gary A Polis[†]
University of California, Davis, CA, USA

Alison G Power
Cornell University, Ithaca, NY, USA

Richard B Primack
Boston University, Boston, MA, USA

Hugh W Pritchard
Royal Botanic Gardens, London, UK

Petr Pyšek
Academy of Sciences and Charles University, Prague, Czech Republic

Eduardo H Rapoport
Universidad Nacional del Comahue, Bariloche, Argentina

David M Richardson
Stellenbosch University, Matieland, South Africa

Tatjana Rosen
Wildlife Conservation Society, Bronx, NY, USA

Carl Safina
Blue Ocean Institute, Cold Spring Harbor, NY, USA

Kaj Sand-Jensen
University of Copenhagen, Copenhagen, Denmark

Kathryn A Saterson
U.S. Environmental Protection Agency, NC, USA

Ruth G Shaw
University of Minnesota, St Paul, MN, USA

Jason F Shogren
University of Wyoming, Laramie, WY, USA

Charles Sims
Utah State University, Logan, UT, USA

Robert S Steneck
University of Maine, Orono, ME, USA

Thomas J Stohlgren
U.S. Geological Survey, Fort Collins, CO, USA

David L Strayer
Cary Institute of Ecosystem Studies, Millbrook, NY, USA

Robert C Szaro
Smithsonian Conservation Biology Institute, Washington, DC, USA

[†]Deceased.

Heather Tallis
Stanford University, Stanford, CA, USA

David Tilman
University of Minnesota, St Paul, MN, USA

James W Valentine
University of California, Berkeley, CA, USA

Ross A Virginia
Dartmouth College, Hanover, NH, USA

Thomas J Volk
University of Wisconsin-La Crosse, La Crosse, WI, USA

Vance T Vredenburg
San Francisco State University, San Francisco, CA, USA

David B Wake
University of California, Berkeley, CA, USA

Diana H Wall
Colorado State University, Fort Collins, CO, USA

Paul C West
University of Minnesota, St. Paul, MN, USA

Robert G Wetzel
University of Alabama, Tuscaloosa, AL, USA

Kellie Whittaker
University of California, Berkeley, CA, USA

John Wilson
South African National Biodiversity Institute (SANBI), Claremont, South Africa

Marten Winter
Helmholtz Center for Environmental Research - UFZ, Leipzig, Germany

A Wiseman
University of Surrey, Guildford, UK

Weihua Xu
Chinese Academy of Sciences, Beijing, China

Wu Yang
Michigan State University, East Lansing, MI, USA

Peter Zahler
Wildlife Conservation Society, Bronx, NY, USA

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

E

Ecology, Concept and Theories in

Peter Kareiva, National Marine Fisheries Service, Seattle, WA, USA
Michelle Marvier, Santa Clara University, Santa Clara, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 259–268, © 2001, Elsevier Inc.

Glossary

Coevolution Process of reciprocal genetic changes in interacting species that result from the mutual selection pressures that strongly interacting species can exert on one another.

Keystone species Species whose presence fundamentally alters the structure or function of an ecological community or ecosystem.

Metapopulation Collection of subpopulations connected by dispersal. Subpopulations within a metapopulation may undergo local extinction and subsequent recolonization.

Threshold theorem of epidemiology If the density of susceptible individuals in a population falls below some critical threshold, then a disease may not be able to maintain itself in the population.

Historical Development of Ecology's Central Concepts and Theories

Ecological theory has three major origins: (1) attempts to manage fisheries, pests, or wildlife, (2) explorations of patterns in nature—especially spatial pattern and the apparent order of species assemblages, and (3) investigations of "the balance of nature" (and how that balance can be disturbed). To some extent, ecological theories developed separately in the sub-disciplines of plant and animal ecology. For example, quantitative plant ecologists tended to focus on patterns of plant dispersion and relationships between individual plants and their environments, whereas animal ecologists focused more on population dynamics. But the most important early ecological models came from neither plant nor animal ecologists; instead these pioneering models were largely contributed by scientists from other fields, such as physics, mathematics, and human demography.

The acceptance of mathematical approaches into mainstream ecology has been slow (and some might argue that mathematical theory is still not accepted by many ecologists). The reluctance of ecologists to embrace mathematical models reflects the different ways that empiricists and theoreticians view the world. In particular, whereas field ecologists are typically concerned with the uniqueness of their study systems and pay attention to the role of history and "details," mathematical ecologists often neglect the role of history and gloss over details to arrive at more general results. Fortunately, there has been increasing reconciliation between the theoretical and empirical approaches to ecology, largely because modern ecological theory has matured enough to allow examination of how details such as age structure, spatial heterogeneity, or environmental variation alter the generalities derived from more simple models. In addition, recent ecological theory relies less on "equilibrium analyses," which tended to downplay historical effects.

Whereas many branches of natural science are grounded in some core body of theory and first principles, this is not the case for ecology. Instead, ecological theory is remarkably varied and discordant, with natural selection as the only fundamental "law." Moreover, although one could argue that little in ecology makes sense without the perspective of natural selection, there is a great deal of useful ecological theory that never makes any contact with evolutionary principles. Thus, even though all ecologists accept evolutionary theory, different subfields of ecology do not share a common view of the most important theoretical questions (such as ecosystem analysis versus population biology). However, in spite of this notable lack of a single unified and ascendant ecological theory, current attention to conservation and resource management has stimulated the practical integration of models that range from evolving gene pools to ecosystem nutrient cycling.

Population Dynamics

Births and deaths are the bottom line of ecology—this is because variation in birth and death rates determines the distribution, density, and dynamics of populations. However, most population models avoid tabulating particular sequences

of births and deaths and instead keep track of average rates of change. In most cases, models of population change are not precise forecasting tools, but rather are conceptual frameworks with which to understand when a population is expected to stay constant as opposed to fluctuate wildly. Models are also valuable for asking how external factors such as harvesting and habitat destruction might alter a population's long-term fate. Two fundamental features of population dynamics are inescapable:

1. Because of resource limitation, per capita rates of change for populations decline with increasing density.
2. If there is any time lag in the response of demographic rates to declining per capita resources, the population may overshoot their resources, then crash to low levels, then rebound, only to repeat the process again.

Formal mathematical representations of these principles have led to elegant models of nonlinear population dynamics. "Nonlinear" refers to the fact that the rate of change in populations varies nonlinearly with changes in population density. As a consequence, under a wide variety of circumstances populations are prone to fluctuate, sometimes so erratically that a graph of numbers through time would appear random, and such that any small changes in initial population size become amplified through time (an attribute called chaotic dynamics). This theoretical result is now widely appreciated by ecologists, but its illumination in the 1970s represented a major advance in the field. First, it emphasizes that populations may fluctuate without any environmental or climatic variation. Second, the phenomenon of chaos illustrates how complicated and unpredictable population dynamics can arise in the simplest systems, if nonlinear feedbacks are strong enough to cause the overshooting of equilibria followed by pronounced population crashes. Prior to the development of a theory of chaotic and nonlinear dynamics there was a tendency to always look to the external environment as the cause of population fluctuations, and to think that complicated patterns of fluctuation must have complicated explanations.

Applications of population theory have largely focused on three types of species: rare species that are of conservation concern, species that reach extraordinarily high population densities ("outbreaks") that cause severe damage to crops or forests, and finally species that are harvested by humans, especially many fish species.

To assess threats to rare species, ecologists have developed a suite of analytical tools collectively known as population viability analysis (or PVA). PVAs are commonly used to estimate the probability of extinction under a range of alternative management strategies or to quantify the minimum number of individuals that will ensure survival of the species. PVA models may incorporate the effects of environmental and demographic stochasticity, catastrophes, and spatial structure and movement (*see* Spatial Patterns). One approach that has been quite successful entails detailed demographic analyses that keep track of the birth and death rates for groups of similar individuals. When the groups are defined by age of the organisms, a life table can be developed. However, demographic matrices are flexible enough to incorporate groups based on size, stage, or sex, as well as age. The principal advantage of demographic matrices is that one can ask how sensitive the growth rate for the entire population is to changes in the survival or fecundity of a particular class of individuals. This approach is being used increasingly to compare the efficacy of alternative management scenarios. Analyses of this nature have revealed that adult survival is generally a very sensitive demographic parameter for long-lived species, whereas fecundity is often the most sensitive parameter for shorter-lived species. These findings have shaped a number of conservation decisions—for example, a demographic matrix applied to loggerhead sea turtles made it clear that even substantial improvements in the survival of baby hatchling turtles would provide little benefit to the turtle population, whereas elevating the survival of adults could dramatically help this threatened long-lived species.

For outbreaking or pest species, population ecologists have asked what factors lead to population outbreaks and whether certain life-history traits are shared among those species with a propensity to outbreak. The goal of this research is to determine how best to control these species and prevent future outbreaks from occurring. One quantitative focus concerning outbreaking species has been on the role of time lags. Time lags can occur when, for example, a population does not respond immediately to changes in its own density, perhaps because the individuals store up energy reserves. Alternatively, a population may lag in its response to one of its natural enemies, such as a predator or a disease. Any such time lag, especially when combined with a high potential for population growth, can lead to dramatic oscillations in population abundance.

Finally, for harvested species there has been a focus on determining the maximum sustainable yield (MSY) that can be harvested from a population without placing it at risk of extinction. Clearly most natural populations do not grow without bound, but rather are regulated around some equilibrium density, called the carrying capacity of the environment. This assumption of density-dependent population growth (a reduction in the per capita rate of population growth as population density increases) is central to the calculation of MSY. In particular, theory predicts that the largest number of individuals can be sustainably harvested when a population is maintained at approximately half its carrying capacity. Although density dependence is a central tenet of population ecology and particularly of fisheries management, actually detecting density dependence from a time series of population abundance estimates remains difficult, and discussions about the appropriate methodology have been contentious. A large part of the difficulty in detecting pattern is the random variation inherent in estimates of population size. While part of the variation is due to observer error, a large fraction is often due to random fluctuations in the environment that change either the population growth rate or the equilibrium population size from one year to the next. It is straightforward to investigate in theory the interplay of density dependence and these external sources of variation. The bigger challenge lies in developing management models for real-world data sets in which the apportionment of variation among observer error and different effects of environmental variation is unknown. Fisheries biologists have made the

greatest progress in this area, but the theory is still incomplete; it remains especially challenging to connect models of population dynamics to economic models that integrate the biological and economic aspects of resource management.

Theories of Species Interactions

The traditional core of mathematical ecological theory involves dynamical models of pairwise interactions between a predator and its prey, or between two competing species. Both of these types of interaction lead easily to the extinction of species—either because a predator imprudently consumes its entire food base or because a superior competitor usurps so much of the resources that other species cannot make a living. Not surprisingly then, the focus of species-interaction theory has been to ask what processes mitigate the tendency toward extinction in these antagonistic interactions.

Although the elimination of an inferior competitor seems inherent to competitive interactions, in nature we regularly find many species coexisting. Early on, theoretical ecologists suggested that coexistence generally occurs only when the populations of the two species are limited by different resources. In other words, coexisting species must occupy distinct ecological niches. The concept of the ecological niche motivated numerous attempts to quantify just how different two species must be in order to coexist. Measures of various traits related to resource use, such as beak size, body size, and habitat preferences, were often invoked as evidence for ''niche partitioning'' (the use of different resources by potentially competing species). Careful attention to such data indicated that this evidence was often not as clear-cut as initially assumed, and research was devoted to alternative explanations for the coexistence of competitors. A second generation of competition models made it clear that competing species may readily coexist without any niche partitioning, as long as some form of disturbance interrupts the exclusion of the inferior competitor. Such disturbance can take the form of predation, or abiotic factors such as storms or droughts. Implicit in disturbance-mediated coexistence is the concept of a trade-off between the ability to be a superior competitor and the ability to recover from a disturbance.

As with competition models, early models of predators and their prey suggested that stable coexistence would require certain special conditions. In the absence of these conditions, predators tend to drive their prey to extinction, and then, having exhausted their food supply, go extinct themselves. One answer to the dilemma of predator–prey coexistence was provided by a classic series of laboratory experiments using predatory mites and their mite prey that live on oranges. These experiments demonstrated that, in a simple ''landscape'' (a tray of oranges), the mite predator was able to quickly find and consume all of its prey, leading to the extinction of both species. However, more complicated laboratory microcosms that impeded the predator's movement could prolong the coexistence of the predatory and prey mites by as much as an order of magnitude—slower predator movement allowed the prey to stay ''one orange ahead'' of predators and enjoy a temporary refuge from attack. Subsequent mathematical models have made it clear that this is a general principle: predator-prey interactions can be stabilized by any attribute that affords prey some refuge from attack. The refuge can take the form of patches of habitat, or age or size classes of individuals that enjoy low attack rates. With spatial refugia, there is still a tendency for predator–prey systems to fluctuate locally, but different patches will fluctuate out of phase so that the entire collection of patches never goes through a phase in which prey are everywhere absent.

In the last three decades predator–prey theory has been modified to model plant–herbivore and host–disease (or parasite) interactions. These models are similar to classic predator–prey theory in structure and in terms of the possible dynamics associated with the interactions. However, there are also important differences. First, the effects of herbivory are not as obvious as those of predation—individual plants are usually not killed outright by herbivory, and some plant species may actually perform better after being partially consumed. Also, studies of herbivory have typically had more of an evolutionary emphasis in focusing on how herbivores act as selective pressures on plant traits, especially plant structural and chemical defenses.

Host–disease interactions also have been a rich subject for mathematical modeling, with much of the initial theory originating from work on the epidemiology of human diseases. Epidemiological models reveal that disease spread depends on the transmission rate, virulence, latency, and infectious period of the disease, as well as the density of susceptible hosts. More specifically, a disease will be unable to persist if the density of susceptible hosts falls below a threshold value. Immunization programs represent an application of this result, sometimes called the ''threshold theorem of epidemiology.'' Immunization can lead to the elimination of a disease, because if a sufficiently large percentage of the population is vaccinated, an infected individual will, on average, fail to transmit the disease to a susceptible host before the infected host either recovers or dies. Epidemiological models further reveal that diseases can potentially maintain host populations at low levels or cause pronounced cycles in the host population.

Although much theoretical work has focused on interactions within species pairs, species do not, of course, occur only in pairs in nature. Most natural communities include scores or hundreds of different species, simultaneously interacting with one another to various degrees. Recently, increased attention has been paid to relatively complex interactions occurring among multiple species. The food web or, more generally, interaction web has proven to be a useful way to visualize the interactions among large groups of species with vertical links representing trophic interactions and cross-links representing competitive or mutualistic interactions (**Figure 1**). Enormous effort has been focused on understanding the relationship between the patterns of connections among species and the traits of the community itself. However, interactions vary in strength, and it is clearly as important to measure the relative strength of an interaction as to merely document its presence.

The study of multiple species interacting simultaneously has led to the recognition that indirect effects can be important determinants of community structure, and several distinct types of indirect effects have been distinguished. For example, apparent competition occurs when two species appear to be

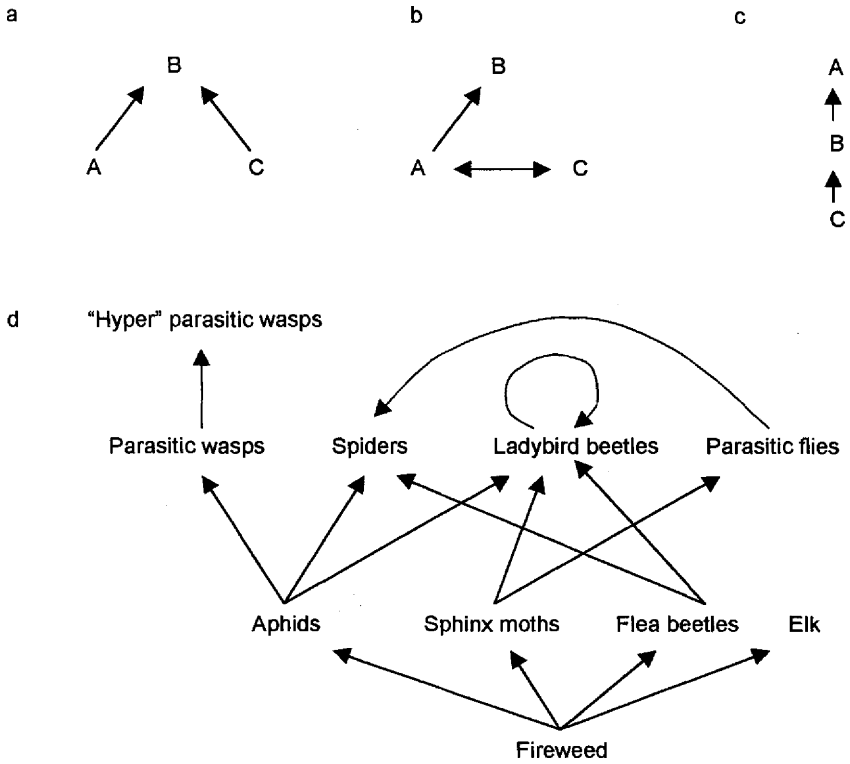


Figure 1 Interaction webs demonstrating three types of indirect interaction: (a) apparent competition; (b) predator-mediated competition; and (c) trophic cascade, as well as (d) a more realistic, species-rich interaction web depicting the major interactors associated with fireweed at Mt. St. Helens, in Washington.

competing because, as the population of one species grows, the abundance of the other declines. However, these changes in abundance are not caused by direct competition, but rather by a third species that consumes each prey species in proportion to its relative abundance (see **Figure 1a**). A second type of indirect interaction occurs when a consumer species indirectly affects a species via its competitor. For example, a consumer that limits a competitively dominant species may indirectly benefit a competitively inferior species (see **Figure 1b**). A third type of indirect effect thought to be common and important is a trophic cascade. An example of a trophic cascade is when a carnivore species limits the abundance of herbivores, resulting in luxuriant plant growth (see **Figure 1c**). The presence of a trophic cascade is demonstrated when removal or exclusion of the carnivore results in a dramatic increase in the herbivore population and the subsequent decimation of plant biomass. Field studies are increasingly documenting the ubiquity and importance of indirect interactions. However, the difficulties encountered when studying the simplified food webs depicted in **Figure 1a–c** are greatly exacerbated when trying to understand the interactions occurring in more realistic, species-rich food webs (see **Figure 1d**). The significance of this is that simple ecological experiments that focus on just one or two species may fail to reveal the important determinants of community structure.

In sum, theories of species interaction have largely focused on the problem of coexistence or, as it has also been called, the paradox of diversity. These theories address how so many species manage to coexist in nature, whereas simple lab experiments and mathematical models so often predict extinction. A major shift in thinking about this problem was the recognition that natural populations are subjected to environmental variability and other disturbances that can prevent the process of competitive exclusion. In addition, the complexities of multispecies interactions, including indirect interactions, can promote the coexistence of species.

Although the earliest models of species interactions seem abstract and disconnected from reality, ecology has a rich tradition of using experiments to test models. For example, the Russian biologist G. F. Gause pioneered the testing of species-interaction models using laboratory cultures of protozoans and yeast in which he could follow population dynamics on the time-scale of weeks, and thus could rapidly see whether or not a species could persist in his experimental systems. When examining competition between *Paramecium* species, Gause often observed that some species could drive others to extinction, a process termed competitive exclusion (and a result predicted in theory by the differential equations that depict continuous reproduction of two protozoans in competition for the same resource). Currently, laboratory microcosms involving protozoan communities are receiving fresh attention—nowadays these experiments are aimed at testing models of multispecies interactions and interaction webs.

Theories of Biodiversity

Biodiversity is the variety of all living things, and most commonly refers to the number of different species in a particular

area. Ecologists have documented striking patterns in the number of species occurring in different regions of the planet. For example, for most groups of species, there is a strong gradient of declining diversity from the equator to the poles, and the vast majority of species are therefore concentrated in tropical and subtropical regions.

Ecologists have proposed a variety of mechanisms that may be responsible for generating global patterns of biodiversity. For instance, areas with more varied landscapes and more topographic relief tend to support more diverse species assemblages than more uniform areas. Thus, spatial heterogeneity is likely to provide more ecological niches for species to occupy and therefore allow more species to coexist. As discussed earlier, classic ecological theory predicts that competitive interactions among species should lead to competitive exclusion and therefore reduced biodiversity. Several of the proposed mechanisms for global diversity patterns focus on forces that counter the process of competitive exclusion. Accordingly, it has been hypothesized that tropical regions harbor many species because their benign climate favors diseases and predators—these predators and diseases increase species diversity by attacking competitively superior species, thereby allowing weak competitors an escape from competitive exclusion. Disturbance is thought to similarly preclude competitive exclusion, although disturbance that is too intense or that occurs too frequently may result in reduced biodiversity. The idea that diversity should be reduced when disturbance is either too rare or too frequent has been termed the intermediate disturbance hypothesis.

Another robust pattern of biodiversity is the observation that islands generally have fewer species than mainlands, and that larger islands tend to have more species than smaller islands. R. H. MacArthur and E. O. Wilson proposed the equilibrium theory of island biogeography to account for this pattern. According to this theory, there is a balance between colonizations of new species on islands and the subsequent extinctions of established species. On larger islands, colonizations are relatively more frequent and extinctions are less frequent, resulting in a higher equilibrium number of species on larger islands.

Human activities such as the direct harvesting of species, introductions of alien species, habitat destruction, and various forms of habitat degradation have greatly accelerated the loss of biodiversity. Consequently, current extinction rates are estimated be 100 to 1000 times higher than pre-human extinction rates. This extinction crisis has spurred a great deal of scientific interest in the biological *and* ecological functions of biodiversity. Organisms are responsible for a variety of ecosystem functions, including maintenance of the gaseous composition of the atmosphere, regulation of the global climate, generation and maintenance of soils, and recycling of nutrients and waste products. However, it is not obvious whether these biological services require many species or only a handful of species. Much current research is focused on understanding how much (if any) diversity we can afford to lose yet still maintain the necessary ecosystem functions (*see* Ecosystem Theories).

Some species play obviously important roles in ecosystems—the addition or deletion of these "keystone" species leads to dramatic changes in ecosystem functions such as productivity or nutrient uptake. However, most species probably do not exert such important effects on ecosystems. In other words, it may be possible to lose a number of species from an ecosystem with little overall impact on ecosystem function. This could be the case if several species that perform approximately the same function are present in the original ecosystem. The situation where multiple species play a similar role has been termed "species redundancy." If species redundancy is a common phenomenon, ecosystem function should be largely independent of species diversity, so long as major functional types are represented. Thus, when one species is lost from an ecosystem, some other species with a similar function may become abundant and compensate for the lost species, leaving the ecosystem as a whole relatively unaffected. Indeed, ecosystem processes often do remain stable despite large fluctuations in the abundance of the various species involved. The term species redundancy may seem to imply that all species are not necessary for an ecosystem to function properly. However, species redundancy may in fact be an essential feature for the long-term health of ecosystems, and the importance of any particular species may only become evident during the occurrence of rare, drastic events.

Spatial Patterns

Spatial patterns and habitat patchiness or heterogeneity have spawned a host of concepts that try to both describe these spatial patterns and explain their impacts and sources. The two most important general results of spatial theory in ecology are:

1. The fact that species interact in spatially extensive and patchy worlds creates opportunities for persistence and coexistence of species that would otherwise not exist.
2. In uniform environments, the interplay of dispersal and nonlinear population dynamics can produce complex spatial patterning in population densities, and this patterning can itself promote coexistence in predator–prey or competitive interactions.

More generally, spatial theory emphasizes that it may be impossible to understand ecological processes and interactions if the focus and scale of a study is too small. This is an important caution, given the fact that many experimental ecology studies are conducted at scales of less than 1 m^2.

One of the most important practical applications of spatial theory stems from the use of metapopulation models in conservation biology. Metapopulation theory describes a species as a collection of populations (or patches) in which there is turnover due to local extinction and subsequent recolonization by dispersing individuals, and in which the fate of a species can only be understood by tracking the fate of a collection of populations or patches, as opposed to a single local population or patch. Metapopulation models are of considerable importance in conservation applications because they identify critical thresholds for habitat destruction or for the creation of barriers to plant and animal dispersal, thresholds that, once crossed, imply doom for a species (because colonization becomes too infrequent to counterbalance local extinctions). It is obvious to anyone who has looked out an airplane window that one of humankind's major impacts

on the world is to destroy habitats and to create ever more fragmented landscapes dissected by highways, agricultural fields, and commercial or residential developments that surely hinder the dispersal of species. If this trend proceeds too far, metapopulation theory predicts the slow but inexorable loss of species. Spatially detailed models of human-modified landscapes are increasingly being used to identify strategies for land use and human development that will minimize the undesirable effects of habitat loss and habitat fragmentation. The two most common strategies involve clustering of habitats and the protection of dispersal corridors.

At a more theoretical level, spatially explicit models of species interactions indicate that the interactions themselves can result in pattern formation. For example, high rates of predation combined with limited prey dispersal can result in a dense patch of prey surrounded by a zone that is relatively devoid of prey due to intense predation. This "ring" of intense predation can limit the spread of prey populations. It is important to realize that these spatial patterns can emerge without any underlying environmental heterogeneity. Spatial patterning can allow more species to coexist because the species occupy different portions of the pattern. However, the opportunity for spatial segregation of species into different characteristic portions of spatial patterns requires sufficiently large habitats—if habitats are carved into patches that are too small, patterns cannot be generated, and opportunities for coexistence will be lost. Thus, theories of pattern formation reveal yet another mechanism by which biodiversity can be maintained, and conversely a route by which biodiversity can be degraded.

Scaling Laws

Because ecological processes operate across a huge range of scales, there have been numerous attempts in ecology to identify scaling laws—laws that allow us to study phenomena on one scale and predict those same phenomena at different scales. Just as important as identifying rules for extrapolating across scales is the idea that there are critical spatial scales at which scaling laws break down and processes fundamentally change in character.

An especially good example of successful extrapolation across scales concerns the prediction of the rate at which exotic species invade new areas and expand their ranges. Simple invasion models extrapolate short-term observations of dispersal (over the scale of one year or less and involving the movements of single individuals) to the broad expanse of an entire species range after decades and over thousands of kilometers. For organisms as diverse as muskrats and cabbage butterflies, invasion models have successfully used small-scale and short-term data to predict long-term rates of range expansion.

To quantitatively examine the legitimacy of extrapolations across scales, it is often useful to graph the variance of a measured variable as a function of the spatial scale at which it was measured. For example, population density could be measured in anything from 1-m^2 quadrats to hectare plots, and the variance in density (per unit area) among plots is expected to decline with increasing plot size. The rate of this decline is a measure of spatial correlation. However, marked changes in the slope relating variance in population density to measurement scale indicate some fundamental shift in biological processes—such as a shift from patterns generated by the dispersal range of a plant's seeds to patterns generated by large-scale climatic heterogeneity. Such nonlinear relationships between variance and scale indicate circumstances where it may not be appropriate to extrapolate across scales.

Most ecologists continue to conduct research at one and only one scale. Although theories regarding scaling laws are still in their infancy, they plainly draw attention to the fact that any extrapolation across scales is tenuous, and can only be justified if the processes have been studied or measured at more than one scale.

Ecosystem Theories

Ecosystem theories focus on patterns of energy flow, nutrient cycling, and the factors that control or limit the rates of these processes and the stability or predictability of these processes through time. To assess the flow of energy and nutrients, it is useful to recognize that organisms fill different trophic roles: autotrophs (usually plants) acquire energy, usually from the sun, and nutrients from nonliving materials, heterotrophs derive energy and nutrients by consuming living matter, and decomposers transform once-living materials back to their inorganic forms. Both nutrients and energy flow from autotrophs to heterotrophs and decomposers, and the connections among these groups, constitute a food chain or food web.

The trophic interactions among species affect the rates of ecosystem functions, including biogeochemical processes, the flow of nutrients, water, and atmospheric gases, and the processing of energy through ecosystems. Species are the key working parts of ecosystems, and therefore biodiversity must be related in some way to ecosystem function. Certainly ecosystems would not function if all species were lost, but it is unclear just how many species are necessary for an ecosystem to function properly. The current extinction crisis has prompted ecologists to assess the relationship between biodiversity and various ecosystem properties. Evidence regarding the importance of biodiversity for ecosystem function largely comes from comparing ecosystems that naturally differ in the number of species present. More recently, ecologists have undertaken more controlled manipulative experiments in which the number of species has been directly varied. Several studies have demonstrated that various measures of ecosystem function, such as production of biomass and nutrient uptake, increase as the number of species present increases. However, some studies also report no effect or even negative relationships between biodiversity and ecosystem processes.

Although some evidence suggests that biodiversity increases or improves the overall functioning of ecosystems, the underlying mechanisms remain unclear. For example, a positive relationship between species diversity and productivity could result simply because including more species increases the chance of including particularly productive, or fast-growing, species. Alternatively, a diverse group of species may use the available resources more efficiently, because each species has a slightly different strategy, resulting in higher overall growth.

A second commonly purported benefit of biodiversity is that more diverse ecosystems may be more stable compared to species-poor ecosystems. Stability can be defined at the community level as fewer invasions and fewer extinctions, meaning that a more stable community will contain a more stable composition of species. However, stability can also be defined at the population level as reduced fluctuations in population size, meaning that a more stable population will contain a more constant number of individuals.

The idea that biodiversity confers stability to ecosystems has enjoyed a long and controversial history. Ecologists initially hypothesized that diverse ecosystems were more stable than those with fewer species because, it was reasoned, with more species in an ecosystem, there are more paths through which energy and nutrients can flow. Consequently, in diverse ecosystems each species should be less affected by changes in the abundance of other species, leading to higher overall stability. However, mathematical models exploring the relationship between biodiversity and population stability indicated that higher species diversity generally reduces the stability of population dynamics of individual species. The conflict between these modeling results and the original intuitions of ecologists remained unresolved for many years.

Recent experimental manipulations of species diversity have helped to resolve this long-standing controversy. In particular, several manipulations of plant and microbe diversity have been performed to examine the stability of various measures of ecosystem function over time. Generally, such studies have shown that, although the abundance of individual species fluctuates more dramatically in high-diversity ecosystems, the total abundance or productivity of all species combined is actually more stable. These empirical results are consistent with both views of diversity and stability. First, the fact that individual species fluctuate more in diverse systems supports the results of mathematical models regarding diversity and stability. On the other hand, the observation that aggregate attributes such as productivity become more stable with increased diversity supports the intuition that multiple pathways for energy (high diversity) foster stability.

Although the relationship between biodiversity and ecosystem stability is becoming increasingly clear, the mechanisms generating this pattern are not. In particular, diverse groups of species may be more stable because complementary species compensate for one another. Alternatively, variation in aggregate measures such as total productivity may increase with richness due to averaging of random fluctuations in the growth of each species. On the basis of simple probability theory, it is expected that as more independently varying species are summed together, the more stable will be the sum of their abundances. The strength of this averaging effect depends on correlations among the species' fluctuations, but a positive relationship between biodiversity and the stability of aggregate measures of ecosystem function should almost always be expected, simply due to averaging.

Evolutionary Theory in Ecology

Evolutionary theory has entered ecology in two major areas: the evolution of species attributes and how they might vary in a predictable manner, and the coevolution of interacting species, such as the evolution of virulence and resistance in host–pathogen interactions.

There is often a remarkable matching between the traits of organisms and the local conditions in which they live. Ecologists and population geneticists alike have worked to identify the factors contributing to or preventing the adaptation of populations to various aspects of their local environments, including food sources, predators, and microclimatic conditions. Reciprocal transplant and common garden experiments have been the major tools used to test for local adaptation. In particular, there is interest in the balance between selection for locally adapted traits and the immigration of individuals with nonadapted traits into the population. This balance between differentiating selection and the homogenizing effects of migration can create a difficult conundrum for conservation biologists—for a rare species that exists in several small populations, it might be better to pool all the individuals together to bolster the size and viability of a single population. Alternatively, the subpopulations might be so genetically distinct and locally adapted to their environments that pooling the individuals might actually harm the species. Thus, one of the more difficult challenges in conservation is to balance the short-term viability of a species with preservation of its long-term evolutionary potential.

As a species evolves to become locally adapted to other species in its environment, it can exert selection pressure on the co-occurring species, leading to changes in the traits of these species with which it interacts. The reciprocal evolution of two species in response to one another is called coevolution. The term coevolution was originally coined to describe the presence of toxic compounds found in plants and the ability of butterfly species to detoxify these compounds. Indeed, many of the best-known examples of coevolution involve reciprocal changes between adversaries, a sort of ever-escalating "arms race" as might occur between a predator and its prey or between a parasite and its host.

Plant breeders try to directly manipulate the revolutionary process that occurs between crop plants and the herbivores and pathogens that attack these plants. Selective breeding is used to maximize the defenses naturally present in crop species and, more recently, genetic engineering has allowed breeders to co-opt defenses of very distantly related organisms and introduce completely novel defense traits to crops. Unfortunately, when highly defended crops are planted over the enormous spatial scales of modern farms, herbivore and pathogen species tend to quickly evolve resistance to the plant defenses. Thus, an important application of evolutionary ecology involves determining how we can best deploy plant defensive traits such that evolution on the part of plant pathogens and herbivores does not make the plant defenses obsolete. Theory suggests that alternating plantings of crops with distinct defensive traits may reduce the ability of herbivores to evolve resistance. In addition, incorporating combinations of defensive traits into single individuals should substantially reduce the likelihood that herbivores will become resistant. Thus, theory in evolutionary ecology has contributed concrete strategies that could greatly enhance the "longevity" of plant defenses, with enormous economic consequences.

Synthesis

Of the numerous specific concepts and theories discussed in this article, some general ideas emerge that relate pattern to process and dynamics to structure. First, complicated spatial or temporal patterns in population densities do NOT require environmental explanations or complicated underlying dynamics—rather, we should expect patterning, cycling, and complexity simply because of the feedbacks inherent in ecological systems. This is of more than theoretical interest, because it implies that any management of ecological systems must heed the implications of nonlinear dynamics. In other words, we must anticipate that rates of change in population growth or in number of species may change dramatically and suddenly, not just because of random fluctuations in the environment, but also as a result of deterministic processes. Second, several ecological models, including epidemiological and meta-population models, point to critical thresholds across which systems change qualitatively (e.g., a disease can no longer persist, or a population dwindles toward extinction because colonizations fail to balance local extinctions). Third, opportunities for coexistence of species arise whenever the world in which organisms live becomes less uniform—spatial heterogeneity creates refugia, disturbances interrupt the process of competitive exclusion, and so on. In contrast, human actions tend to enforce regularity and control on the natural world, and ecological theory suggests that this homogenization of nature is contributing to the erosion of our planet's biodiversity. The consequence of this reduction in biodiversity remains an open theoretical question, with the greatest challenges to ecological theory being the development of models that contend with rich webs of species interactions in complex ecosystems. The future of ecological theory lies in blending human activities with natural processes, so that we can better understand how to encourage a sustainable biosphere while accommodating human interests and demands.

See also: Coevolution. Complexity Versus Diversity. Keystone Species. Population Dynamics. Population Viability Analysis. Resource Partitioning. Species Interactions

References

Gotelli NJ (1995) *A Primer of Ecology*. Sunderland, Massachusetts: Sinauer Associates.

Huston MA (1994) *Biological Diversity: The Coexistence of Species on Changing Landscapes*. Cambridge, United Kingdom: Cambridge University Press.

Krebs CJ (1994) *Ecology: The Experimental Analysis of Distribution and Abundance*, 4th ed. Menlo Park, California: Addison-Wesley.

Levin SA (1999) *Fragile Domain*. Reading, Massachusetts: Perseus Books.

Ricklefs RE and Schluter D (1993) *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. Chicago: University of Chicago Press.

Roughgarden J, May R, and Levin SA (1989) *Perspectives in Theoretical Ecology*. Princeton, New Jersey: Princeton University Press.

Tilman D and Kareiva P (1997) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*. Princeton, New Jersey: Princeton University Press.

Ecology of Agriculture

Alison G Power, Cornell University, Ithaca, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Facilitation Species interactions in which at least one of the participants benefits and neither partner is harmed.
Monoculture Growing a single crop in a field.
Natural enemy A predator, parasite, parasitoid or pathogen of another animal, especially an insect.
Niche complementarity Species that differ in at least one aspect of their resource use, and thereby experience less intense competitive interactions, allowing them to coexist.
Nitrogen Fixation The process by which gaseous nitrogen (N_2) in the atmosphere is converted by biological or chemical means to ammonia and other forms usable by plants and animals.
Polyculture Growing two or more crops in the same field at the same time.
Resilience The rate at which an ecosystem returns to original conditions after a perturbation.
Resistance The degree to which an ecosystem changes in response to a perturbation.

Introduction

The popularity of agriculture as a focus of ecological study has waxed and waned during the past century. Agricultural systems have been touted as model systems for testing ecological theory because of their relatively well-characterized components, simple food webs, and manipulability. In the past, they have often served as effective model systems for plant ecology, plant–insect interactions, and predator–prey theory. In recent years, the growth of sustainable agriculture has resulted in increasing interest in understanding ecological processes in agriculture so that they can be manipulated more effectively for enhancing agricultural productivity and reducing negative environmental impacts of agricultural activities. A variety of types of ecological theory have been applied to agricultural systems, and thus the subdiscipline of "agroecology" encompasses a very broad range of approaches, including physiological ecology, population biology, community ecology, ecosystem ecology, and landscape ecology. All these approaches offer significant insights into the ecology of agricultural systems and provide useful guidelines for improving the functioning of these systems. Because the role of diversity in ecological systems has been a consistent element in much agroecological research, this article focuses on diversity in agricultural systems, with particular attention to areas of agroecology that appear to have the most potential to offer agricultural and environmental enhancement.

Planned and Unplanned Diversity in Agroecosystems

The biodiversity in agroecosystems can be classified as either "planned" or unplanned diversity. Planned diversity includes the spatial and temporal arrangement of domesticated plants and animals that farmers purposely include in the system. It may also include beneficial organisms that are deliberately added to the agroecosystem, such as biological control agents or plant-associated nitrogen-fixing bacteria. Unplanned diversity (or "associated diversity") includes all the other associated organisms that persist in the system after it has been converted to agriculture or colonize it from the surrounding landscape. This component of diversity is likely to include a variety of weedy plants, herbivores, predators, parasites, and microorganisms that make up the majority of species in any ecosystem, even a simplified one such as an agroecosystem. It has become increasingly clear that the two components of diversity are significantly linked: As planned diversity increases, the diversity of the associated biota also increases. Although the evidence for the plant-based food web is strongest, this relationship also appears to exist for the detritus-based food web.

Given this relationship between planned and unplanned diversity, how do both types of diversity influence agricultural productivity and other ecological processes that occur in and around agroecosystems? The impact of planned diversity is relatively well understood, but much remains to be elucidated about the role of unplanned diversity in agroecosystem function.

Impact of Planned Diversity on Ecological Processes

Agroecosystems vary dramatically in their complexity and degree of planned diversity, which may include genetic diversity, species diversity, structural diversity, and functional diversity (**Figure 1**). The input-intensive monoculture systems that dominate commercial agriculture worldwide have low planned diversity and relatively simplified community assemblages. At the other extreme, traditional agricultural systems and home gardens of the tropics often have high planned diversity and complex assemblages of associated biota. This range of planned diversity in existing systems has provided rich opportunities for exploring the impact of planned diversity on a suite of ecological processes, including primary production, pest regulation, decomposition, and nutrient cycling.

Planned Diversity, Productivity, and Stability

The idea that diversity should impart stability to ecological systems, including agricultural systems, dates back at least to

Agroecosystem	Genetic	Species	Structural	Functional
Herbaceous monoculture				
Herbaceous varietal mixture				
Herbaceous polyculture				
Plantation crops				
Agroforestry				
Tropical home gardens				

Figure 1 Degree of different types of planned diversity found in some typical agroecosystems. Intensity of shading indicates degree of diversity. Reproduced from Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005b) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* 8: 857–874, with permission from Wiley.

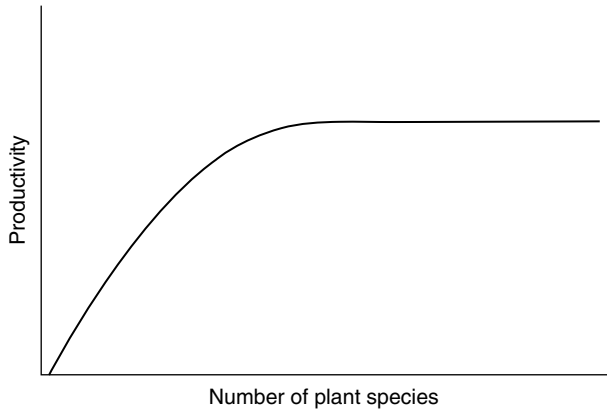


Figure 2 Predicted relationship between productivity and plant species richness. Reproduced from Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005b) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* 8: 857–874, with permission from Wiley.

the 1950s and has stimulated much ecological theory, some empirical research, and extensive debate. Until the 1990s, the best evidence for the idea that diverse systems should be more stable than simple ones came from agricultural ecosystems rather than natural ecosystems. Many agricultural studies have demonstrated increased productivity (i.e., significant yield increases) in diverse cropping systems compared to monocultures. However, the majority of these studies address systems of relatively low plant species richness, comparing monocultures to polycultures of two or three species. We expect that the relationship between productivity and plant species richness should resemble the asymptotic curve in **Figure 2**, as it appears to do so for natural grasslands (Tilman *et al.*, 1996) and long-term experimental grasslands (Tilman *et al.*, 2006), but we have little data to indicate where the asymptote might be. Researchers have speculated that the increase in productivity with increasing species richness in agroecosystems may level out at approximately 10 plant species, although this number is likely to vary according to crop species identity, type of agricultural system, and environmental conditions.

The tendency of diverse cropping systems to have higher productivity than would be expected on the basis of the productivity of their component crops grown in monoculture is often called "overyielding." Overyielding may result from niche complementarity or facilitation resulting from a variety of mechanisms, such as more efficient use of resources (light, water, and nutrients) or reduced pest damage. There have been numerous experimental studies examining these mechanisms (Malezieux *et al.*, 2009; Letourneau *et al.*, 2011). When interspecific competition for a limiting factor is less than intraspecific competition for that factor, overyielding is predicted. Facilitation occurs when one crop modifies the environment in a way that benefits a second crop, for example, by lowering the population of a critical herbivore or releasing nutrients that can be taken up by the second crop. Facilitation may result in overyielding even where direct competition between crops is substantial (Vandermeer, 1989). Note that not all crop combinations exhibit niche complementarity or facilitation, and not all combinations overyield. For obvious reasons, however, most of the classic polycultures used by farmers over centuries, such as grass–legume mixtures, tend to overyield.

There is evidence that diverse cropping systems exhibit greater yield stability and higher productivity, suggesting that resistance to environmental perturbation may be higher in diverse systems. Yield stability has been measured in at least three ways: by calculating coefficients of variation of yield, by computing regressions of yield against an environmental index, and by estimating the probability of crop failure. Polycultures exhibit greater yield stability according to all criteria. Polycultures tend to have lower coefficients of variation than crops grown as separate monocultures; the response of polycultures to environmental change tends to be as stable or more stable than the most stable component crop grown in monoculture; and polycultures tend to have a much lower probability of crop failure than the component crops grown in monoculture. The probability of crop failure is an estimate of risk and lower probabilities result from both the higher yields of polycultures and the putative yield stability. Overall, studies indicate that diverse cropping systems are

more stable than monocultures, both agronomically and economically.

Several mechanisms may lead to greater yield stability in diverse systems. When one crop performs poorly because of drought or pest epidemic, for example, the other crop(s) can compensate, using the space and resources made available. Such compensation is obviously not possible if the crops are grown separately. If the yield advantages of polycultures are greater under stress conditions, then yield stability is higher. This polyculture advantage has been demonstrated for crops under nutrient stress and drought stress. Moreover, where polycultures lead to reduced pest attack, as they often do (Letourneau *et al.*, 2011), then greater yield stability may result from the dampening of pest outbreaks and disease epidemics.

These processes may operate simultaneously in diverse systems and are consistent with the idea that species richness buffers productivity under conditions of environmental variability and that diversity imparts resistance to perturbation. The ability of polycultures to compensate for losses might also be considered to represent resilience. While systems with greater resistance will be less impacted by a perturbation, systems with greater resilience return rapidly and reliably to original conditions.

Ecological research on the role of diversity in natural grassland systems around the world (Hooper *et al.*, 2005; Tilman *et al.*, 2006) indicates that diverse natural communities may be more productive than simple systems, at least partially as a result of increased nitrogen use efficiency in more diverse systems. These studies also suggest that more diverse plant communities are more resistant to disturbance and more resilient in the face of environmental perturbations such as drought. For example, the productivity of diverse grassland communities appears to decline less during a drought and to return more quickly to predrought levels than is the case for species-poor communities. Research in grasslands thus tends to support the conclusions from agricultural research that diverse plant communities have higher productivity and stability.

Planned Diversity and Pest Regulation

The planned diversity of the agroecosystem has important effects on herbivorous insects and the microbial community attacking crops. Traditional agricultural systems often include substantial planned genetic and species diversity, and genetically diverse grain crops are used in many parts of the world to control pathogens. In contrast, the low planned diversity of most commercial monocultural systems often results in large crop losses from a pest complex that is less diverse but more abundant than that in more diverse systems. The trend for higher pest densities in monocultures compared to diverse cropping systems is especially strong for specialist insect herbivores with a narrow host range. As planned diversity increases, population densities of these specialist herbivores decrease (Andow, 1991). In a recent meta-analysis of experiments on plant diversification in agricultural systems, Letourneau *et al.* (2011) found that herbivores and crop damage were strongly suppressed and natural enemy populations were enhanced in diversified systems compared to monocultures.

In a classic paper, Root (1973) offered two hypotheses to explain higher densities of herbivorous insects in simple, monocultural systems: the resource concentration hypothesis and the enemies hypothesis. The first hypothesis predicts that herbivores, particularly specialists, in pure, dense host plant stands will be more likely to find their hosts and more likely to survive and reproduce. In contrast, herbivores in less dense or more host plant-diverse stands should be less likely to find their hosts and more likely to lose them. Although the details of herbivore–host interactions vary considerably, subsequent experimental tests of this hypothesis have generally supported it with respect to the effects of diversity (Letourneau *et al.*, 2011), especially for specialist herbivores. Host-finding behavior and insect movement, both colonization and emigration, appear to play important roles in the response of herbivorous insects to agroecosystems. Densities of specialists may be lower in diverse systems because (1) they have difficulty locating hosts due to interference with olfactory or visual cues or (2) they leave hosts more often due to lower plant quality and then have difficulty relocating them. These behaviors are significantly affected by the chemical, nutritional, and structural diversity that accompanies planned plant species diversity.

The enemies' hypothesis predicts that diverse systems should have higher densities of herbivore natural enemies (predators and parasites) because they provide more resources for these natural enemies, such as alternate prey or hosts, nectar, pollen, and refugia. This hypothesis has also largely been supported in experimental studies (Letourneau *et al.*, 2011). Compared to monocultures, diverse systems are likely to have higher predation rates, higher parasitism rates, and higher ratios of natural enemies to herbivores, all of which may contribute to lower pest densities. Natural enemies comprise one component of the unplanned diversity that increases as planned diversity increases in agroecosystems.

Microbial pathogens also respond to the planned diversity of agroecosystems, but their response is more variable than that of herbivorous insects. The majority of plant viruses are transmitted by insects, and these tend to be found at lower incidence in diverse systems due to the effects of plant species diversity on their insect vectors. The impacts of host-specific pathogens are likely to be lower in diverse systems due to density-dependent population growth. However, crop diversification can modify the microclimatic conditions that play an important role in the development and severity of many fungal and bacterial diseases. Pathogen growth and reproduction may be either encouraged or inhibited in more diverse cropping systems, depending on the particular requirements of the organism. The effects of diversity depend on a variety of dispersal processes, infection efficiency, and the rate of disease progress. Interestingly, some experiments in grasslands have shown that low productivity in low-diversity communities can largely be attributed to specialized soil-borne pathogens (Schnitzer *et al.*, 2011).

The genetic diversity of crops can dramatically reduce pathogen impacts on crop productivity. Mixtures of genotypes of a single species, such as multiline cultivars and varietal mixtures, have been used effectively to retard the spread and

evolution of fungal pathogens in small grains and other crops. There is evidence that genetic mixtures may also have lower densities of insect herbivores and a lower incidence of plant viruses. Typically, these mixtures include both resistant and susceptible crop genotypes, although they may be mixtures of several different resistant genotypes. The reduction in pathogen spread is greater than would be expected on the basis of the proportion of resistant genotypes in the mixtures, and therefore it appears to be due to the effects of diversity *per se* on the ability of pathogens to disperse (Mundt, 2002).

High planned diversity that includes genetic diversity, species diversity, and structural diversity thus has a strong influence on populations of herbivorous insects and plant pathogens. The intentional manipulation of planned diversity for the purpose of pest control is common in traditional agricultural systems, and it offers real potential for increasing the productivity of contemporary agricultural systems without the negative environmental impacts of pesticides.

Planned Diversity and Soil Processes

Planned diversity may also have significant impacts on the soil community. Plant pathogens and their antagonists in the soil are well known to respond to crop diversity. Long-term, continuous monoculture can result in dramatic shifts in the competitive balance among microbial species and increase the aggressiveness of plant pathogens. Conversely, populations of plant parasitic fungi, bacteria, and nematodes may decline when monocultures of their host plants are replaced by diverse cropping systems. Rhizosphere microbial communities, including a variety of bacteria and fungi, respond to plant species composition of the cropping system, plant phenology, plant nutrient status, and even plant genotype. Plant mutualists such as mycorrhizal fungi or nitrogen-fixing bacteria may be strongly determined by crop composition and phenology. Some ecological processes such as nitrogen cycling may be substantially controlled by crop diversity and plant composition since the inclusion of legumes can increase rates of biological nitrogen fixation (Drinkwater and Snapp, 2007). Owing to the effects of litter diversity, decomposition rates may also be quite responsive to planned diversity.

Despite increasing awareness of the effects of planned crop diversity on the soil community, our ability to predict the impacts of manipulating planned diversity is limited because our understanding of the complex interactions among soil organisms and between plants and soil microbes is still rather weak. Greater attention to the ecological dynamics of these soil communities under different cropping systems will aid in determining the extent to which we can manage these communities to enhance agricultural productivity using planned diversity.

Unplanned Diversity and Ecological Services

The unplanned diversity that accompanies planned diversity in agricultural systems can provide many ecological services to agriculture. This means that the conservation of biodiversity can offer significant benefits to agriculture. Uncultivated species, including wild relatives of crops that occur in and around the agroecosystem, are an important source of germplasm for developing new crops and cultivars. Increasing planned crop diversity can augment the resources available to pollinators and to natural enemies such as parasitic wasps, resulting in higher populations of these beneficial organisms. Increasing planned diversity may also foster beneficial soil organisms and the conservation of functional processes, such as decomposition and nutrient cycling.

Overall, unplanned diversity within the agroecosystem affects plant and soil processes that can in turn affect crop productivity. For example, many traditional cacao and coffee systems in the tropics resemble forest ecosystems because they are well shaded and humid, with a diversity of shade trees, a diverse herbaceous understory, and a thick leaf litter layer. These traditional agroecosystems have high planned species diversity, high genetic diversity, and high structural diversity. Studies have shown that shaded cacao and coffee plantations support higher levels of biodiversity than full-sun plantations grown without shade trees, that is, monocultures (Perfecto *et al.*, 1996; Tscharntke *et al.*, 2011). The shaded systems may be traditional "rustic" systems in which the understory of native forests is cleared and the crop is planted under the natural forest overstory. They may also be improved systems in which a few carefully chosen tree species are planted interspersed among the crop. In either case, these shaded systems have high planned diversity accompanied by high unplanned diversity.

Shaded coffee and cacao have been demonstrated to be good habitats for migrant and resident forest birds, as well as bats, and these vertebrates can be important consumers of pest insects. They also contain relatively high diversities of insects, lizards, and other fauna, some of which contribute to pest regulation. Pollination also appears to be enhanced in shaded systems. Since many of the commonly used shade trees are leguminous, the trees reduce the amount of nitrogen inputs required. In addition, the shade provides significant weed control and, at least in cacao, reduces the incidence of several important diseases. The high levels of unplanned diversity thus contribute significantly to agricultural productivity, and recent studies have shown that there is no necessary trade-off between biodiversity and yield (Clough *et al.*, 2011). Moreover, by serving as habitat for a relatively high diversity of vertebrate and invertebrate animals, such agroecosystems may enhance the conservation value of an agricultural landscape.

Unplanned Diversity and Pest Regulation

One of the ecological services provided by the unplanned diversity in agroecosystems is the regulation of herbivorous insects and microbial pathogens by competitors, predators, and parasites. As discussed previously (*see* Planned Diversity and Pest Regulation), high planned diversity often results in higher densities of predators and parasites. Even in agroecosystems of low planned diversity, the diversity of natural enemies may be quite high as long as pesticides are not used. For example, paddy rice monocultures managed without pesticides can have surprisingly high diversities of herbivorous insects, predators, and parasitoids compared to similar

monocultures in which pesticides are used. Recent pest management programs in Southeast Asian paddy rice have taken deliberate advantage of this diversity and drastically reduced pesticide inputs without sacrificing yields. In traditionally managed rice fields, the predators are likely to include fish and amphibians, which contribute to pest regulation and also provide additional nutritional resources for farm families.

Unplanned Diversity and Soil Processes

Despite some significant advances in our understanding of soil processes, we know relatively little about the soil biota and their impact on agricultural productivity. In natural ecosystems, decomposition and soil nutrient cycling are regulated by a diverse community of invertebrates and microorganisms, such as termites, earthworms, nematodes, fungi, and bacteria. The composition, abundance, and activity levels of the soil biota in agricultural systems are markedly different from those in surrounding natural ecosystems. For example, the diversity and abundance of soil insects and earthworms in tropical agroecosystems are typically significantly reduced compared to those of a wide range of undisturbed tropical ecosystems. In cases in which abundance remains high in agricultural systems, the soil communities are often dominated by a single or small number of species highly adapted to the modified environment.

The changes in the soil community under agriculture result from a variety of perturbations to the soil environment. The initial conversion from undisturbed ecosystems typically involves the removal and/or burning of plant biomass followed by tillage – activities which have drastic impacts on soil structure and soil chemistry. The physical changes at the soil surface amplify diurnal and seasonal fluctuations in temperature and moisture. In addition, organic inputs to the soil are significantly reduced as a result of plant biomass removal, and the chemical composition of organic inputs is altered. These extreme modifications of the soil environment can result in the elimination of some soil organisms and, at a minimum, are likely to change the competitive balance among species. To the extent that agricultural systems minimize these perturbations (e.g., by reducing tillage or burning), the impacts on the soil community may be less severe.

Despite these well-known effects of agricultural conversion on the soil community, the link between loss of soil biodiversity and various ecological processes has not been extensively studied. Decomposition and nutrient mineralization, for example, are controlled by the activities of a diverse community of organisms. It is not clear that the loss of some species will result in significant changes to these functions, but the degree of functional redundancy among different species is still controversial. However, some processes such as nitrogen fixation are carried out by very specific organisms, whose loss might substantially affect nitrogen cycling. Given the importance of soil biota for decomposition and nutrient cycling, it is essential that the link between this component of unplanned diversity and ecological processes be explored more fully.

Agriculture and Diversity in the Landscape

Agroecosystems typically sit in a landscape of multiple land uses. Human-modified landscapes fall along a continuum from relatively homogeneous landscapes consisting of a small number of cropping systems to heterogeneous mosaics of diverse cropping systems within a matrix of natural ecosystems. Agricultural activities can have important effects on natural ecosystems through the movement of organisms, pesticides, nutrients, and soil. At the same time, the landscape context can strongly influence ecological processes that occur within agroecosystems, such as pollination and pest regulation. The spatial structure of the landscape, including the size, shape, and connectedness of different ecosystem types can have dramatic effects on these ecological processes.

Consequences of Agricultural Activities

Ecological concepts have often been utilized to address the consequences of agricultural perturbations for natural ecosystems in agricultural landscapes. Agricultural intensification can diminish ecosystem services such as the provision of clean water due to increased nutrients, agrochemicals, and dissolved salts entering groundwater and surface water. Significant sedimentation of waterways and dams results from the loss of riparian vegetation that often accompanies intensification. The impacts of nutrient inputs and pesticides for downstream aquatic and terrestrial ecosystems have been well studied. Much of the pesticide applied to crops fails to reach the target pests and instead moves into adjacent ecosystems, where it can have significant impacts on the diversity and abundance of nontarget species. Studies have shown that pesticides have strong effects on trophic dynamics and a range of ecosystem processes in complex natural ecosystems, both aquatic and terrestrial. Similarly, movement of nitrogen from agroecosystems to natural ecosystems through leaching or deposition may lead to significant impacts at several trophic levels. Increased nitrogen loading may reduce plant species richness while increasing the biomass of dominant species. In response to these changes, both pathogens and herbivorous insects are likely to exhibit lower species richness but higher population abundance, just as they do in agricultural systems. Changes in pathogen and herbivore populations may in turn influence the productivity and fitness of their host plants. Even in the absence of changes in herbivore population pressure, research has shown that the effects of herbivory on plant growth and reproduction may be more severe at high levels of nutrient availability. As with the complex effects of pesticides, these impacts of inadvertent fertilization are likely to impact trophic dynamics and ecosystem processes.

Spillover of Organisms from Agroecosystems

In addition to the movement of soil and agricultural inputs, a variety of organisms move from agroecosystems to natural ecosystems in agricultural landscapes. There is abundant evidence that pathogens move from domestic animals to wild animals (and humans), along with more limited evidence for movement of plant pathogens from crops to wild plants

(Power and Mitchell, 2004). Despite human efforts to breed crops resistant to pests and pathogens, many crops are excellent hosts for a variety of parasites that can spill over to wild hosts. Herbivorous insect pests build up on high quality, nutrient-dense crop hosts and then move to wild host plants, where increased herbivory may have significant effects on plant demography (Tscharntke *et al.*, 2005a). The reproductive success of a number of rare or endangered plant species has been significantly reduced due to herbivores moving from crops (Blitzer *et al.*, 2012). Movement out of crop fields is particularly likely at mowing, harvest, and other times when the crop resource is disturbed, degraded, or diminished. Similar movements have been seen for natural enemies of crop pests, which seek alternative prey in natural habitats when crop pests are no longer available. As a result, enemies that increase in agricultural systems due to high pest densities can suppress the densities of herbivorous insects in natural ecosystems, thereby changing the dynamics of plant–insect interactions.

Finally, there is some evidence for spillover of pollinators from agroecosystems to natural habitats (Blitzer *et al.*, 2012). Many pollinators forage in crop fields, but move to natural habitats to complete their lifecycle. Studies suggest that spillover of bees from agroecosystems may sometimes benefit native plant species in natural habitats by compensating for the decline in native pollinators due to habitat fragmentation and other changes. If pollinators are limiting, however, then crops may compete with native plants for pollination services. In general, organisms that build up in agroecosystems clearly have the potential to modify species interactions and the structure of natural plant communities, but the consequences of movement of organisms from managed to natural ecosystems have not been extensively studied.

Landscape Influences on Ecological Processes in Agriculture

The spatial structure of the landscape can shape many ecological processes in agroecosystems, including water and nutrient uptake by crops, weed dispersal, pest control, and pollination. Globally, the intensification of agriculture has caused a simplification of landscape structure through the expansion of agricultural land, increase in field size, loss of field margin vegetation, and elimination of natural habitat. This simplified landscape structure can affect the movement of organisms across the agricultural landscape, often limiting the movement of wildlife and organisms beneficial to agriculture if natural habitat corridors are not available (Tscharntke *et al.*, 2005b; Kremen *et al.*, 2007). In contrast, wildlife, pests, natural enemies, and pollinators are more likely to move among agricultural and natural habitats in heterogeneous landscapes with smaller, intermixed patches. In some cases, natural habitats provide beneficial organisms with refugia and resources that are unavailable in agroecosystems, thus contributing valuable ecosystem services to agriculture.

Meta-analyses of published studies on the effects of landscape structure on natural enemies and pests in agriculture indicate that landscape complexity often enhances natural enemy populations and biological control of insect pests. Relatively few studies have quantified the direct effects on pest

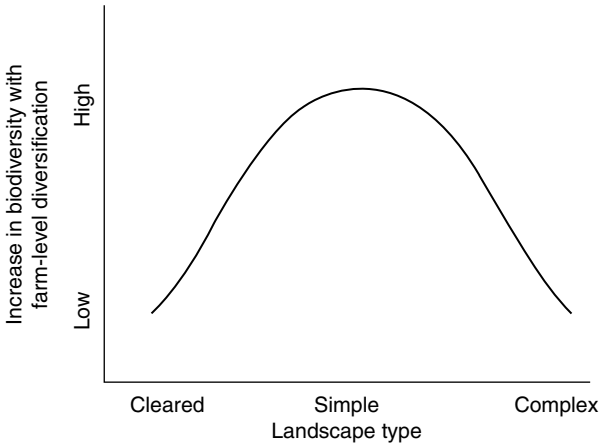


Figure 3 Predicted relationship between biodiversity and local, farm-level diversification in landscapes of varying diversity. Reproduced from Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005b) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* 8: 857–874.

populations, but there is some evidence that landscape simplification leads to higher population growth rates of pests and greater plant damage (Chaplin-Kramer *et al.*, 2011). One unresolved issue is whether pests and natural enemies respond to landscape complexity at different spatial scales. Some studies have suggested that enemies may respond at smaller scales than pests, and that parasitoids respond at smaller scales than other enemies. Meta-analyses do not strongly support the overall difference between pests and enemies, but they indicate that specialist natural enemies such as parasitoids may respond to landscape structure at smaller spatial scales than generalist natural enemies such as predators (Chaplin-Kramer *et al.*, 2011). As a rule, natural enemies that are capable of dispersing long distances are more likely than weak dispersers to survive in simplified agricultural landscapes with few or distant patches of natural habitat. As described above, crop diversification and management of seminatural habitat, such as hedgerows or field borders, within a single farm can also lead to higher natural enemy populations. However, farm-level diversification is more likely to influence pests and natural enemies if the wider landscape is structurally simple, than if it is already heterogeneous with many patches of natural habitat (**Figure 3**). Within complex landscapes, pest control services may not be significantly affected by additional farm-level diversity (Tscharntke *et al.*, 2005b; O'Rourke *et al.*, 2011). Landscapes that completely lack natural habitat, however, may have such a limited species pool of natural enemies that farm-level diversification may not result in notable increases in pest control services.

Like natural enemies, pollinators tend to be more abundant in complex landscapes with patches of natural habitat. Pollination by bees or other animals significantly improves productivity for more than 70% of globally important crops. As honeybee populations have declined due to disease, pollination by wild bees has become an essential ecosystem service provided by natural habitats. Agricultural intensification threatens wild bee communities and may destabilize

pollination processes at the landscape level. Both pollinator visitation to crops and the species richness of pollinators decline with increasing distance from natural habitats, and the decline is steeper in tropical landscapes compared to temperate landscapes (Ricketts *et al.*, 2008). This implies that a heterogeneous landscape with intermixed patches of agriculture and natural vegetation would be most successful in delivering pollination services to agriculture.

Future Directions

Although the study of agricultural ecosystems has provided many insights into the role of diversity in agroecosystem function and the ecological impacts of agriculture, the subdiscipline of "agroecology" is still in its infancy. Two areas that urgently warrant more attention from ecologists are (1) the influence of planned and unplanned agroecosystem diversity on soil processes such as decomposition and nutrient cycling and (2) the functional consequences of landscape complexity for community structure in agroecosystems and natural habitats. A better understanding of the ecological processes operating in both areas could result in significant benefits to both agriculture and environmental protection.

See also: Agriculture, Industrialized. Agriculture, Sustainable. Biodiversity and Pest Control Services. Parasitism. Plant Conservation. Plant–Soil Interactions

References

Andow DA (1991) Vegetational diversity and arthropod population response. *Annual Review of Entomology* 36: 561–586.

Blitzer EJ, Dornmann CF, Holzschuh A, Klein A-M, Rand TA, and Tscharntke T (2012) Spillover of functionally important organisms between managed and natural habitats. *Agriculture, Ecosystems, and Environment* 146: 34–43.

Chaplin-Kramer R, O'Rourke ME, Blitzer EJ, and Kremen C (2011) A meta-analysis of crop pest and natural enemy response to landscape complexity. *Ecology Letters* 14: 922–932.

Clough Y, Barkmann J, Juhrbandt J, *et al.* (2011) Combining high biodiversity with high yields in tropical agroforests. *Proceedings of the National Academy of Sciences of the United States of America* 108: 8311–8316.

Drinkwater LE and Snapp SS (2007) Nutrients in agriculture: Rethinking the management paradigm. *Advances in Agronomy* 92: 163–186.

Hooper DU, Chapin FS, Ewel JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75: 3–35.

Kremen C, Williams NM, Aizen MA, *et al.* (2007) Pollination and other ecosystem services produced by mobile organisms: A conceptual framework for the effects of land-use change. *Ecology Letters* 10: 299–314.

Letourneau DK, Armbrecht I, Rivera BS, *et al.* (2011) Does plant diversity benefit agroecosystems? A synthetic review. *Ecological Applications* 21: 9–21.

Malezieux E, Crozat Y, Dupraz C, *et al.* (2009) Mixing plant species in cropping systems: Concepts, tools and models. A review. *Agronomy for Sustainable Development* 29: 43–62.

Mundt CC (2002) Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytopathology* 40: 381–410.

O'Rourke ME, Rienzo-Stack K, and Power AG (2011) A multi-scale, landscape approach to predicting insect populations in agro-ecosystems. *Ecological Applications* 21: 1782–1791.

Power AG and Mitchell CE (2004) Pathogen spillover in disease epidemics. *American Naturalist* 164: S79–S89.

Ricketts TH, Regetz J, Steffan-Dewenter I, *et al.* (2008) Landscape effects on crop pollination services: Are there general patterns? *Ecology Letters* 11: 499–515.

Root RB (1973) Organization of a plant–arthropod association in simple and diverse habitats: The fauna of collards (*Brassica oleraceae*). *Ecological Monographs* 43: 95–124.

Schnitzer SA, Klironomos JN, HilleRisLambers J, *et al.* (2011) Soil microbes drive the classic plant diversity-productivity pattern. *Ecology* 92: 296–303.

Tilman D, Reich PB, and Knops JMH (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441: 629–632.

Tilman D, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379: 718–720.

Tscharntke T, Clough Y, Bhagwat SA, *et al.* (2011) Multifunctional shade-tree management in tropical agroforestry landscapes – a review. *Journal of Applied Ecology* 48: 619–629.

Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005b) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters* 8: 857–874.

Tscharntke T, Rand TA, and Bianchi F (2005a) The landscape context of trophic interactions: Insect spillover across the crop–noncrop interface. *Annales Zoologici Fennici* 42: 421–432.

Vandermeer J (1989) *The Ecology of Intercropping*. Cambridge: Cambridge University Press.

Economic Control of Invasive Species

David Finnoff and Jason F Shogren, University of Wyoming, Laramie, WY, USA
Richard D Horan, Michigan State University, East Lansing, MI, USA
Shana M McDermott, Dartmouth College, Hanover, NH, USA
Charles Sims, Utah State University, Logan, UT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptation Actions that reduce the consequences of invasive species damage.

Comparative statics Comparing economic outcomes both before and after a change in an underlying exogenous parameter.

Control Controlling invasive species population or spread to reduce the damage an invasive species causes after becoming established in a new habitat.

Endogenous risk The idea that private citizens and public servants have some control over the set of probabilities and outcomes that define the relevant states of the world.

Feedback loop An ecological response to a human action having environmental consequences, followed by a human behavioral response to ecological changes having economic consequences.

Joint determination Ecological and economic systems are both affected by, and affect, each other.

Optimal control theory A technique used to study problems involving optimal decisions in a multiperiod framework.

Prevention Efforts to reduce the probability of an invasive species being introduced and subsequently becoming established in a new habitat.

Introduction

Invasive species destroy key natural resources used to produce market and nonmarket goods and services causing damages that outweigh any benefits they may generate (Elton, 1958; Kareiva, 1996; Lodge *et al.*, 2009). Their expanding global numbers play a primary role in global environmental change (ESA, 2004; Mack *et al.*, 2000; Sala *et al.*, 2000). In response, invasive species policies have expanded in scope and magnitude. For instance, in the USA, Executive Order 13112 points out the need to "minimize the economic, ecological, and human health impacts that invasive species cause" (also see Miller and Fabian, 2004; National Invasive Species Counsel (2001); CENR, 1999; FICMNEW, 1998). Internationally, countries can propose trade restrictions to prevent introductions of invasive species through the World Trade Organization's (WTO) Agreement on Sanitary and Phytosanitary (SPS) measures (see Peterson and Orden, 2008; Calvin and Krissoff, 1998; Josling *et al.*, 2004; Smith, 2003).

For these public policies to be cost-effective, they should be informed by both economic and ecological principles (Finnoff *et al.*, 2009a, b, 2010a, b). Invasive species risks are jointly determined by economic and ecological interactions (Crocker and Tschirhart, 1993). Economic activity introduces many invasive species who find the new ecosystem suitable; they then spread through natural migratory patterns or hitch-hiking on commerce and recreation. Both economics and ecology and the feedback loops between the two systems play a key role (feedback loops here are defined as ecological responses from human actions having environmental consequences, followed by human behavioral responses to ecological changes having economic consequences). Effective public policy also requires detailed knowledge about the portfolio of risk reduction technologies – prevention, control, and adaptation (see Shogren, 2000). (These guiding principles have generated an explosion of bioeconomic research, as shown in special issues of *Ecological Economics*, Volume 52, 2005; *Agricultural and Resource Economics Review*, Volume 35, 2006; and *Resource and Energy Economics*, Volume 32, 2010, a book edited by Keller *et al.* (2009) and a book edited by Perrings *et al.* (2010).) This article defines "prevention" as investments to reduce the probability of an invasive species being introduced and subsequently becoming established in a new habitat; "control" is defined as actions to curtail invasive species population or spread to reduce the damage an invasive species causes after becoming established in a new habitat; "adaptation" is defined as actions to reduce the consequences of invasive species damage.

This article combines the idea of *endogenous risk* into an optimal control setting to provide a framework to address the challenge on how to best manage the prevention and control of invasive species (Kamien and Schwartz, 1991). Endogenous risk captures the idea that private citizens and public servants have some control over the set of probabilities and outcomes that define the relevant states of the world (Ehrlich and Becker, 1972; Shogren and Crocker, 1991). The dynamic endogenous risk approach used here captures the following: (1) the biological and economic circumstances of invasions, (2) the interdependency of economic and ecological systems, and (3) accounts for inherent uncertainty associated with the bioeconomic parameters. This framework provides a means to identify the feedback loops between systems and to evaluate the trade-offs inherent in designing and implementing cost-effective public policies directed at the threats of invasive species. The discussion begins by presenting the problem of invasive species in the dynamic endogenous risk optimal control framework.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00185-4

Optimal Control of Invasive Species

When considering how to manage renewable natural resources cost-effectively over time, economists typically use dynamic optimization, the most common variant being optimal control theory (for a primer in this setting, see Kamien and Schwartz, 1991; Lewis *et al.*, 2009). Economists use optimization techniques either to maximize net benefits (benefits from a management plan net its costs) or to minimize costs of the management plan. The optimal control method provides the derivation of an optimal jointly determined evolution of state variables (e.g., invader abundances) and control variables (prevention, control, and adaptation). (Examples include Barbier (2001), Eiswerth and Johnson (2002), Olson and Roy (2002, 2008), Barbier and Shogren (2004), Knowler and Barbier (2005), Burnett *et al.* (2006, 2008), Kim *et al.* (2006), Potapov *et al.* (2007), and Finnoff *et al.* (2010b).

In the standard setting, the invasive species problem is cast as a renewable resource problem in which a single control is chosen to maximize the negative of a discounted stream of damages and control costs (Eiswerth and Johnson, 2002; Finnoff *et al.*, 2010b). The result is an optimal path for control that is monotonic in invasion size, yet typically exhibits features such as multiple equilibria that make the evaluation of optimal policies difficult (see Olson and Roy, 2008, for a discussion). The standard optimal control setting is extended to one that incorporates endogenous risk in the spirit of Shogren (2000), as in Horan and Fenichel (2007). The dynamic endogenous risk framework incorporates more of the key risk reduction strategies that are available to managers.

Consider a benevolent resource manager charged with managing the risks of invasion. Following Jerde and Lewis (2007), **Figure 1** details the invasion process as a series of stages.Individual members of the invasive species are initially transported from a population in a source region. Some fraction survives, providing a dispersal pool of individuals. A fraction of the dispersal pool is then introduced into the region of interest from which some fraction can further the spread of invaders and damages. Depending on the species and circumstances of the invasion process, the relative importance of each of these stages varies.

At a basic level there are three public policies (or management strategies) available to the resource manager in response to the risk. The manager can employ strategies that lower the probability of transport and introduction, a catch-all strategy termed *prevention*. Following introduction, the manager can use strategies that *control* the invasion to stop or slow the spread and subsequent damage from the invader. Finally, the manager can *adapt* to the consequences of the invader and employ strategies that alleviate the realized damages. Although prevention and control work to inhibit the spread and damage of the invader, adaptation does not influence the growth and spread of the ecological process of a specific invasive species (note that adaptation might affect the ecological process of some future invader and also the incentives for control).

This article divides the management problem into two separate, though interrelated, time intervals. Specifically, it splits the ecological invasion process in **Figure 1** in half, so that invasion occurs at time τ. This partition also separates management into two stages: prevention occurs prior to time τ, whereas control and adaptation occur after time τ. The time of invasion τ is a random variable that can be delayed by the manager's investments in prevention. There are no damages during preinvasion (i.e., when $t<\tau$), although costly investments in prevention help avoid or delay future damages. Following an invasion ($t\geq\tau$), realized damages and social welfare depend on the manager's choices of control and adaptation in response to the invasion, as typically seen in the standard model of controlling an established invasion (which ignores prevention).

To include all of these features in a single model, the standard model of controlling an established invasion is merged with a rigorous treatment of the uncertainty underlying the timing of invasion and its prevention. This extension follows the model developed by William Reed (1987) for the optimal preventative maintenance schedule of a revenue-earning asset subject to random catastrophic failure. The method states the stochasticity of the problem in such a manner that deterministic methods of optimal control can be used to find a solution (see Reed and Heras, 1992, for applications toward natural resources and Horan and Fenichel, 2007, for an epidemiological application). Kim *et al.* (2006) present a variant of the method directed at invasive species, but they do not focus on differentiation in timing of prevention and control. Burnett *et al.* (2006, 2008) employ alternative frameworks that include prevention and control in which prevention works to prevent repeated introductions.

Here Reed's method is used to illustrate what are believed to be the key features in the economics of invasive species that can be addressed with the use of optimal control theory. Consider a natural asset or resource that generates a flow of social benefits in each period prior to invasion ($t<\tau$) given by $B(t)$. If an invader is able to successfully establish and cause damage (at time $t=\tau$ and beyond), then the benefits are

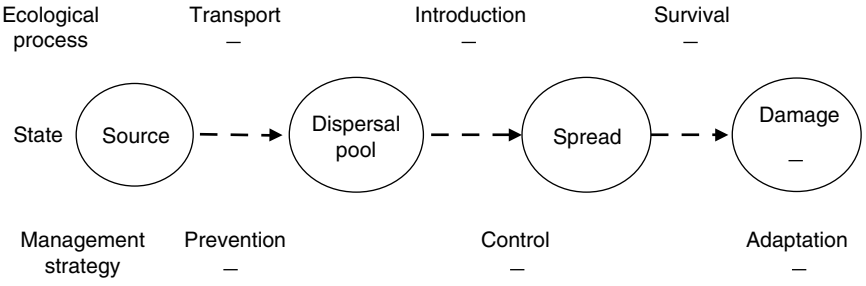


Figure 1 Flow diagram of the invasion process.

diminished by damages $D(X(t))$, where the invader stock is given by $X(t)$, where $X(t)=0$ before the invasion at time $t=\tau$ and $X(\tau)>0$. Greater levels of X lead to greater damages so that $(\partial D/\partial X)>0$. Expenditures on prevention $n(t)$ reduce net benefits before the invasion, V, so that

$$V(B,n,t) = B(t) - n(t) \quad [1]$$

These expenditures also reduce the probability of invasion and so the realization of damages. Following Reed (1987), the probability the invader is introduced at any time t, given that it has not been introduced up to that point in time, is given by the effective hazard function

$$\psi(n(t),b(t)) = \lim_{\Delta t\to 0}\{Pr(\textit{Invasion in}(t,t+\Delta t)|\textit{no invasion at } t)/\Delta t\} \quad [2]$$

where $b(t)$ is the background hazard or probability of invasion in the absence of prevention. The greater the background probability of invasion, the greater the hazard $\psi_b>0$ and the more the prevention reduces the hazard $\psi_n<0$ (where subscripts indicate partial derivatives in what follows). The probability of no invasion having occurred up to time t is given by the survivor function

$$S_p(t) = e^{-\int_0^t \psi(n(t),b(t))\,dv} = e^{-\gamma(t)} \quad [3]$$

where the change of variables allows one to define $\dot{\gamma}=\psi(n,b)$, $\gamma(0)=0$ so that $S_p(0)=1$.

Following an invasion at time τ, the invader stock grows following a density-dependent growth function $F(X(t))$. Control expenditures $h(t)$ reduce or reverse the growth in the invader following an invasion through a "kill function" $K(h(t))$ so that

$$\dot{X} = F(X(t)) - K(h(t)),\ X(\tau)=X_\tau,\ t\in\{\tau,\infty\} \quad [4]$$

More money spent on control can be expected to remove more of the invader, so $K_h>0$. Adaptation expenditures $a(t)$ simply lower realized damages while allowing the invader stock to remain so that $D(X(t), a(t))$ with partial derivatives $D_X>0$ and $D_a<0$. Let the flow of social net benefits from time τ onward be given by VX so that

$$VX(B,X,a,h,t) = B(t) - D(X(t),a(t)) - h(t) - a(t) \quad [5]$$

The expected net present value of net benefits earned over an infinite horizon is given by

$$J = E_\tau\left\{\int_0^\tau e^{-rt}V(n,t)\,dt + \int_\tau^\infty e^{-rt}VX(X,a,h,t)\,dt\right\} \quad [6]$$

where the expectations operator reflects the uncertainty of invasion time τ and r is the discount rate. If one defines the present value of an optimal program of *ex post* management from τ through time to be given by $JX^*(X(\tau)) = \max_{h,a}\int_\tau^\infty e^{-rt}VX(X,a,h,t)\,dt$ (where the star indicates the function has been optimized by the optimal choices of $h^*(t)$ and $a^*(t)$ through time), it can be seen that JX^* depends on *ex ante* management through prevention before τ through both the point in time that damages begin to accrue and the initial stock of the invader in τ. $JX^*(X(\tau))$ is the solution to the standard renewable resource model of an optimally controlled invasion as developed in Eiswerth and Johnson (2002), Brown *et al.* (2002), Lichtenberg and Lynch (2006), Burnett *et al.* (2006, 2008), Potapov *et al.* (2007), and Finnoff *et al.* (2010b). First the *ex ante* management schemes are discussed, and then the insight provided by optimal control for *ex post* management is explored.

Ex ante Management: Prevention

Rewriting eqn [6] allows the objective function for *ex ante* prevention to be written as

$$\max_s E_\tau\left\{\int_0^\tau e^{-rt}V(n,t)\,dt + e^{-r\tau}JX^*(X(\tau))\right\} \quad [7]$$

Following the steps laid out in the Appendix A, eqn [7] can be rewritten in a manner similar to that provided by Reed and Heras (1992)

$$J = \max_n \int_0^\infty [V(n,t) + \psi(n,b)JX^*(X(t))]e^{-rt-\gamma(t)}\,dt \quad [8]$$

subject to $\dot{\gamma}=\psi(n,b)$, $\gamma(0)=0$, and the equation of motion in eqn [4]. The method results in a problem of deterministic optimization (Reed, 1987) that can be solved applying Pontryagin's maximum principle (see Kamien and Schwartz, 1991). The beauty of the method is that it incorporates the endogenous risk of invasion directly into the associated conditional current value Hamiltonian

$$\overline{H} = V(n) + \psi(n,b)JX^* + \rho\psi(n,b) \quad [9]$$

where ρ is the costate variable for γ. This deterministic formulation allows the application of the maximum principle (Pontryagin *et al.*, 1962). The associated necessary conditions require prevention to be chosen along the optimal path to maximize the conditional current value Hamiltonian

$$\frac{\partial\overline{H}}{\partial n} = V'(n) + \psi_n[JX^* + \rho] = 0 \quad [10]$$

given the evolution of γ follows

$$\dot{\gamma} = \psi(n,b),\ \gamma(0)=0 \quad [11]$$

and the evolution of ρ is

$$\dot{\rho} = \rho(r + \psi(n,b)) + V(n) + \psi(n,b)JX^* \quad [12]$$

which has solution (Horan and Fenichel, 2007)

$$\rho(t) = -\int_t^\infty [V(n) + \psi(n,b)JX^*]e^{-r(s-t)-\gamma(s-t)}\,ds \quad [13]$$

The implication (Reed and Heras, 1992) is that $-\rho(t)$ is the expected present value of net benefits from the current time onward, or the *ex ante* value of an optimally managed system facing the threat of invasion. This value depends on the state of the system before $V(n)$ and after an invasion $JX^*(X,a^*,h^*)$ (i.e., the severity) and the probability of invasion $\psi(n,b)$, where each in turn depends on both the ecological baseline risk of invasion and the key aspects of manager behavior – prevention, control, and adaptation. The problem is now one defined by endogenous risk.

Analytical solutions for the optimal time path of prevention are lacking although numerical solutions are available, as shown in Reed (1987). This article can gleam some insight by evaluating eqn [10]. Given the interpretation of $-\rho(t)$, the implication from the maximum principle becomes apparent – prevention should be employed up until the level that balances the current marginal cost of prevention to its expected marginal benefits or the marginal welfare savings expected from a reduced risk of invasion (Horan and Fenichel, 2007). Such a prevention level is optimal unless (1) there is no probability of invasion, (2) unless prevention has no effect on the probability of invasion ($\psi = \psi_n = 0$), (3) if the invasion has minor consequences or is benign ($JX^* + \rho \to 0$), or (4) if the invasion were to improve societal welfare so that $JX^* + \rho > 0$.

From the system of eqns [10]–[12], it is possible to derive a differential equation governing the evolution of prevention. The procedure starts from time differentiating eqn [10] and substituting in the definitions provided by the initial structure of the problem and the findings from the maximum principle. To reduce the complexity, it is assumed that the present value of an optimal program of *ex post* management JX^* is time invariant from an *ex ante* perspective. The procedure allows one to determine the "key equation" governing the optimal time path of prevention (Reed, 1987)

$$\dot{n} = \frac{\psi_n(\psi + r)(V'(n) + \psi_n JX^*) + [V'(n)\psi_{nb}]\dot{b} - [V(n) + \psi JX^*]\psi_n^2}{V''(n)\psi_n - V'(n)\psi_{nn}} \qquad [14]$$

Again there are no analytical solutions to eqn [14], making generalizations problematic. Numerical solutions are possible and provide some insight into the optimal prevention schedule. For example, if the analysis is restricted to consideration of time invariant levels of prevention (or $n(t) = \overline{n}$ and $\dot{n} = 0$) and if it is assumed that the background risk is constant over time so that $\dot{b} = 0$, then as it is known that $V''(n) = 0$, $\overline{n}$ can be found as the solution to

$$\psi_n(\psi + r)(V'(n) + \psi_n JX^*) - [V(n) + \psi JX^*]\psi_n^2 = 0 \qquad [15]$$

To provide some illustrative insight, this expression was solved for a hypothetical collection of functions and parameters selected to be consistent with an invasion and with one another. Taking an exponential form for $\psi = e^{-c\beta(b)n}$, letting $r = 0.03$ and $B = \$1000$, restricting JX^* to being less than the value of an uninvaded state so $JX^* < B/r$, and assuming a linear form for the influence of the baseline risk of invasion $\beta(b) = a - db$ and $c = 1$, $a = 0.1$, and $d = 0.075$. For this case, a comparative static analysis of changes in fundamental parameters on the time invariant level of prevention can be performed. **Table 1** presents the description of parameters.

These comparative static effects conform to intuition. Note, however, it is uncertain whether these would hold after relaxing the restrictive assumptions. There have been few comprehensive analyses of prevention in this *ex ante* setting, although the results of the stochastic dynamic programming model in Finnoff *et al.* (2007) corroborate the results provided in **Table 1**. Kim *et al.* (2006) present a similar model with no differentiation in the timing of prevention and control (all decisions are made before an invasion has occurred and this uncertainty has been resolved). Although they do not focus on the timing of invasion as is done in this article, they incorporate preventative measures before and after the discovery of an invader and perform a comparative dynamic analysis on the time paths of prevention and control. The authors are able to use their model to demonstrate that if the expected damages from the invader are large enough (low JX^* in the authors' model), it is efficient to expend a larger share of resources on prevention before discovery. However, this result can be reversed if the marginal costs of this early prevention activity are sufficiently high.

Burnett *et al.* (2006, 2008) examine prevention but again in a setting that places prevention and control decisions at the same time period. Unlike the model developed here, Burnett *et al.* specify prevention by reducing repeated introduction, or reducing a flow of invaders that would otherwise create or increase an established stock. Burnett *et al.* (2008) show that the optimal prevention schedule depends on whether or not the species has established in the initial conditions of the model. If none have established, the optimal path of prevention gradually increases over a period of time to a fixed level that should be maintained into the future. If the population is already established at the start of the problem, it is then optimal to prevent at a constant level into the future (where this prevention reduces the magnitude of population growth and damage).

Ex post Management: Control

Control of an established and spreading invader has been the most well studied aspect of the problem using optimal control theory. The invasive species problem transforms a basic renewable resource problem by (typically) choosing control to maximize the negative of a discounted stream of damages and control costs. The models capture the interaction of state variables (e.g., invader abundances) and control variables

Table 1 Comparative static effects on prevention

Parameter	*Description*	*Effect on fixed level of prevention ($\bar{n}$)*
r	Discount rate	Increases in r lead to less use of prevention
b	Baseline risk of invasion	Increases in b lead to more use of prevention
B	Social benefits at risk from invasion	Greater levels of B lead to more use of prevention
c	Effectiveness of prevention	If prevention is not very effective (low c), an increase leads to more prevention. If prevention is already highly effective (high c), an increase leads to less prevention
JX^*	Present value of an optimal program of *ex post* management from time of invasion into perpetuity	Higher damages (or lower JX^*) lead to more use of prevention

(policy choices, adaptation, eradication, and control) over time and in equilibrium (steady states).

There is a long history in economics of using optimal control theory to answer related natural resource policy questions as seen in the use of the method to inform policies aimed at agricultural pests (Lichtenberg and Zilberman, 1986; Archer and Shogren, 1996; Levy and Caputo, 2008). For policies concerned with invasive species, the focus has been on the control of an invasion and whether eradication of the invader is optimal. This article illustrates the use of optimal control for managing an established invader by continuing the derivation of the value of an optimally managed system following an invasion, JX, under an assumption that the invasion occurs immediately (or $\tau = 0$). The objective functional can be written as

$$JX(X) = \max_{h,a} \int_0^\infty e^{-rt} VX(X,a,h,t)\, dt \qquad [16]$$

where the strategies are control h and adaptation a, as defined in eqn [5], and the optimization is subject to the growth of the invader, as given in eqn [4]. Optimal policies for maximizing JX can be obtained by applying the maximum principle to the current value Hamiltonian (omitting time notation)

$$\mathcal{H} = B - D(X,a) - h - a + \mu(F(X) - K(h)) \qquad [17]$$

where μ is the costate variable or the shadow price of the invader and signals to society the value (loss) of allowing an additional unit of the species to persist into the future. The invader as defined here results only in costs, making $\mu \leq 0$. Following the maximum principle, the necessary conditions require control and adaptation to be chosen along the optimal path to maximize the current value Hamiltonian

$$\frac{\partial \mathcal{H}}{\partial h} = -1 - \mu K_h \leq 0, \quad h \geq 0 \qquad [18]$$

$$\frac{\partial \mathcal{H}}{\partial a} = -D_a - 1 \leq 0, \quad a \geq 0 \qquad [19]$$

given the evolution of X follows

$$\dot{X} = F(X) - K(h), X(0) = X_0 \qquad [20]$$

and the costate equation requires that μ satisfy

$$\dot{\mu} = \mu(r - F_X) + D_X \qquad [21]$$

Equations [18] and [19] place restrictions on the optimal use of expenditures on control and adaptation. If control is to be used ($h > 0$), then a dollar spent on control must result in at least a dollar's "worth" of invader removals, where the marginal productivity of control expenditures K_h is valued by the shadow price μ. Similarly, if it is optimal to adapt ($a > 0$), then each dollar spent on adaptation must lower damages by at least a dollar.

Equations [18]–[21] define a system of differential equations governing the optimal evolution of the system. The solution to these systems is tricky as there is the possibility of multiple equilibria and complex transition dynamics. Furthermore, as the shadow price of the invader is negative, the problem lacks the structure typically relied upon to determine the optimality of alternative paths (Olsen and Roy, 2008; Finnoff *et al.*, 2010b). The point is that standard sufficiency conditions can no longer be used to determine the optimality of one path and equilibrium combination in relation to another. To determine which path and equilibrium is optimal requires a comparison of the value of the entire program of management, that is, a solution-by-solution comparison of JX^* (for an elegant alternative, see Tahvonen and Salo (1996)). In most cases, analytical solutions cannot be derived and the comparisons have to be made numerically. Although the numerical solutions are dependent on the parameters and functional forms employed and so are case specific, this article has found two basic cases relating to whether or not there are repeated introductions.

The first case considers the control of an established invader that does not face the threat of repeated introductions as seen in Eiswerth and Johnson (2002) and Olsen and Roy (2008). This would be appropriate for those invasions with rare yet potentially disastrous introduction events. In this case, eqns [18]–[21] can be reduced to a pair of differential equations with a solution typically shown by the construction of a phase diagram in X and h space (ignoring adaptation) and illustrated in **Figure 2**. Phase diagrams are used to illustrate the optimal solution to the dynamic problem. They depict the key equilibrium points and nonequilibrium (or dynamic) directions of movement for the state and control variables. They are particularly useful in these problems as they show both the optimal paths to equilibrium points (here unique saddle paths) and the nonoptimal paths (how any pair of state and control variables would move following the small directional arrows).

The horizontal axis of the figure is the invader abundance X (with units here scaled between 0 and 1) and vertical axis expenditures on control h. The solutions to the system of eqns [18]–[21] are shown using the isoclines for X and h (trace out the combinations of X and h that keep each of the variables invariant over time) and sketches of optimal paths. The intersections of the isoclines (and the origin when the invader and control are both zero) provide the candidate equilibria for the system.

In **Figure 2** there are three possible equilibrium points, a highly invaded steady state (X_H,h_H), a low invaded steady state (X_L,h_L), and an eradication steady state (0,0). The small arrows provide directions of movement in the figure, demonstrating the conditional stability of both the eradication and the highly invaded equilibrium points and the instability of the low invaded equilibrium point. Which (of the stable) equilibrium points are optimal depends (for a given set of circumstances) on the initial conditions in relation to a Skiba point in the neighborhood of the unstable equilibrium point (X_L,h_L). If the invasion is detected before the invader population has risen to X_L, it will be optimal to control to eradication. If the invasion is detected in its moderate or latter stages (i.e., an abundance above X_L), it is optimal to control the invader to maintain a heavily invaded equilibrium. In relation to the model of prevention discussed earlier (i.e., eqn 7), the optimized value function JX^* for early detection and eradication is always higher than moderate and late detection that leads to maintenance at X_H (although which equilibrium is optimal depends on the initial conditions).

The exact location of the equilibrium points and optimal path(s) depends on the specifics of a problem. A key ecological variable is the rate of spread or growth of the invader. In comparison to the situation depicted in **Figure 2**, a species

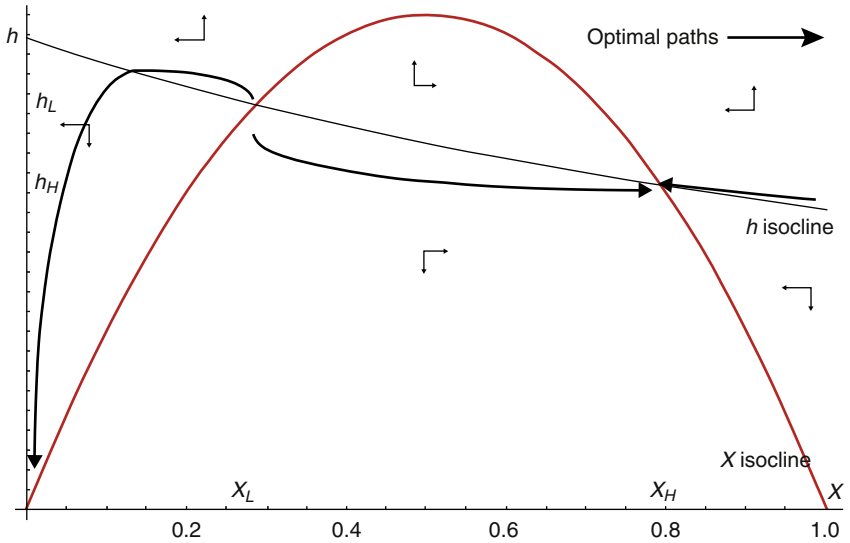


Figure 2 Phase diagram for control of an established invader without reintroduction.

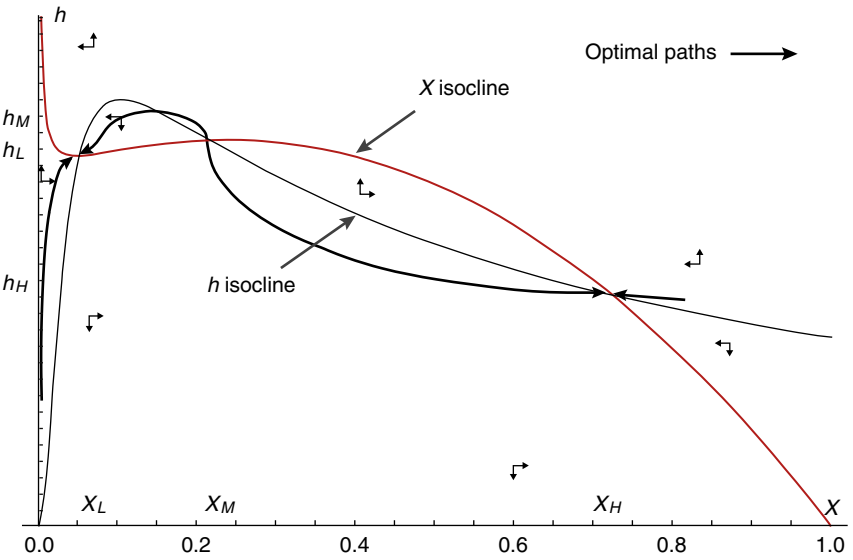


Figure 3 Phase diagram for control of an established invader with reintroduction.

with a lower rate of spread (or intrinsic growth) would have a lower X isocline and a higher h isocline. This makes eradication optimal over a wider range of initial conditions. If the spread rate or growth rate of the species is higher, the opposite occurs – eradication is now optimal over a narrower window of initial conditions. Key economic variables are the economic damage of the invasion, the discount rate, and the marginal cost of control. Greater damages (lower discount rate and lower marginal costs of control) raise the h isocline, leading to an increase in the range of initial conditions over which eradication is optimal and lowering the highly invaded steady state. Lower damages (higher discount rate and higher costs of control) in turn lower the h isocline, reducing the interval of initial conditions for eradication and increasing the highly invaded steady state.

The second case considers situations of repeated introductions, as considered by Burnett *et al.* (2008) and Finnoff *et al.* (2011). This is appropriate for invasions with numerous entry pathways. Now the phase space and isoclines undergo modification due to the inclusion of repeated introductions. The main implication is the isocline for X no longer intersects the origin. This indicates that through repeated introductions and a high positive amount of grow at low invader abundances, the invader will soon be established. **Figure 3** presents a sketch of one potential configuration of the phase space with all labeling consistent with **Figure 2**.

As **Figure 3** illustrates, a completely eradicated equilibrium does not exist for invasions with repeated introductions (i.e., perfect prevention is impossible). There are three equilibrium points, consisting of a slightly invaded equilibrium (X_L,h_L), a

moderately invaded equilibrium (X_M,h_M), and a highly invaded equilibrium (X_H,h_H). The slightly and highly invaded equilibria are conditionally stable (saddle points); the moderate equilibrium is again unstable. Repeated introductions imply that a one-time eradication program is insufficient; even if the invader can be held in check at a low abundance, it will have to be consistently maintained over time through a persistent high application of control.

The optimality of the slightly invaded steady state in relation to the highly invaded steady state again depends on the relationship between the initial conditions and the moderately invaded equilibrium point. Invasions detected early enough so their initial conditions are less than the moderately invaded equilibrium point should be optimally directed to the slightly invaded equilibrium (with high JX^*). Invasions detected later in their progression should be managed to the highly invaded steady state (with low JX^*). Variations in key parameters shift the isoclines in a similar fashion, as documented in Finnoff *et al.* (2010b). Across their set of key parameters, Finnoff *et al.* demonstrate that the control of the invasion process to the slightly invaded steady state is optimal for early detection and relatively high damages, low rates of density-dependent spread, and low rates of repeated introductions. Invasions that should be controlled to the highly invaded steady state are those detected in their latter stages with low relative damages, high rates of repeated introductions, and high levels of uncertainty in species location.

Concluding Comments

This article has reviewed the insight that optimal control theory can provide for the bioeconomics of invasive species policy. The key message is that a bioeconomics framework can provide useful insight into cost-effective invasive species management. People invest scarce resources to prevent and control the risks invasive species pose to ecological and economic systems. Just how people invest and how they *ought to* depend on what they know about the economic and biological underpinnings of the problem and how the systems respond to one another. Here the authors have used the idea of endogenous risk in an optimal control setting to frame the question of how to best manage the prevention and control of invasive species. The approach is inclusive and accounts for uncertainty, both biological and economic circumstances of invasions, and interdependent systems.

The dynamic endogenous risk framework provides a means to evaluate the trade-offs at the foundation of designing and implementing policies directed at invasive species. Creating scenarios that capture realistic trade-offs still requires more information on the economic and ecological systems, and the critical feedback loops between the two systems. More modeling and data collection is needed. As better insight is gained into understanding the alternative risk reduction strategies, it is more likely to understand the degree to which prevention has been underfunded (Leung *et al.*, 2002).

The authors appreciate the limits of their approach. The optimal control framework is well suited for control actions that can be continuously adjusted such as pesticide/herbicide application or biomass removal. The existing optimal control literature has found that more aggressive control actions of this type are optimal early in the invasion process when the costs of such actions are relatively low (e.g., Olson and Roy, 2002). But for the control of invasive species at the state and federal levels, control actions are delayed with less aggressive actions initially implemented. These broad-scale state- and federal-level policies require significant economic commitments that are typically difficult to adjust over time. Examples include decisions to "blacklist" potentially invasive species, the construction of an electric barrier in the Chicago Ship and Sanitary Canal to prevent Asian carp spread into the Great Lakes, and the mandatory ballast water-exchange policy for ocean-going shipping. For broad-scale control decisions, alternative approaches such as an optimal stopping framework could be considered (Neher, 1990).

The goal is to understand better the consequences of invasion and the policies that can alleviate these consequences. And although more is known now than a decade ago about the bioeconomics of invasive species, more work is needed on the *economic welfare* consequences of invasions and of associated public policies. This will require more realistic economic parameter assumptions and accounts of ecological thresholds in the authors' models.

Appendix A

For the derivation of J: The following is based on Rick Horan's implementation of the Reed and Heras method. Let the random variable denoting the time of invasion be given by Z, for some time horizon T. If $Z=T$, no invasion occurs. As the time of invasion is a random variable, so is the present value of rewards

$$\begin{cases} \int_0^Z V(n,t)\mathrm{e}^{-rt}\,\mathrm{d}t + \mathrm{e}^{-rZ}JX^*(X(Z)) & \text{if } Z<T \\ \int_0^Z V(n,t)\mathrm{e}^{-rt}\,\mathrm{d}t + \mathrm{e}^{-rT}JX^*(X(T)) & \text{if } Z=T \end{cases} \quad [A1]$$

The cumulative distribution function (cdf) of Z is $(1-S_p(z))$ for $z<T$ and 1 for $z\geq T$. This allows one to write the expected present value of rewards as

$$\begin{aligned} J = &\int_0^T \int_0^Z V(n,t)\mathrm{e}^{-rt}\,\mathrm{d}t\,d\big(-S_p(Z)\big) \\ &+ \int_0^T \mathrm{e}^{-rz}JX^*(X(z))d\big(-s_p(z)\big) \\ &+S_p(T)\int_0^Z V(n,t)\mathrm{e}^{-rt}\,\mathrm{d}t + S_p(T)\mathrm{e}^{-rt}\psi(n(T),b(T)) \end{aligned} \quad [A2]$$

Let the first term in eqn [A2] be given by J^1

$$J^1 = \int_0^T \int_0^T V(n,t)\mathrm{e}^{-rt}\,\mathrm{d}t\,d\big(-S_p(z)\big) \quad [A3]$$

As $d\big(-S_p(z)\big) = d\left(-\mathrm{e}^{-\int_0^t \psi(n(t),b(t))\,\mathrm{d}v}\right) = S_p(z)\psi(n(z),b(z))\,\mathrm{d}z$, eqn [A3] can be simplified to

$$J^1 = \int_0^T \int_0^Z V(n,t)E^{-rt}S_p(z)\psi(n(z),b(z))\,\mathrm{d}t\,\mathrm{d}z \quad [A4]$$

The integration can be simplified by switching the order of the integrals in eqn [A4]. Noting the limits of z for a given t

will be $t \leq z \leq T$ and time ranges from $0 \leq t \leq T$ finds

$$J^1 = \int_0^T \int_t^T V(n,t)\mathrm{e}^{-rt} S_p(z)\psi(n(z),b(z))\mathrm{d}z\,\mathrm{d}t$$

$$= \int_0^T V(n,t)E^{-rt}S_p(t)\,\mathrm{d}t - S_p(T)\int_0^T V(n,t)\mathrm{e}^{-rt}\mathrm{d}t \qquad [A5]$$

Let the second term in eqn [A2] be given by J^2, make the same substitution for the cdf as in the earlier equation, and change the dummy variable of integration

$$J^2 = \int_0^T \mathrm{e}^{-rt} JX^*(X(t))S_p(t)\psi(n(t),b(t))\,\mathrm{d}t \qquad [A6]$$

Substituting eqns [A5] and [A6] into eqn [A2] finds

$$J = \int_0^T [V(n,t) + \psi(n(t),b(t))JX^*(X(t))]\mathrm{e}^{-rt}S_p(t)\,\mathrm{d}t$$

$$+S_p(T)\mathrm{e}^{-rT}\psi(n(T),b(T)) \qquad [A7]$$

To push the time horizon T to infinity, take the limit as $T \to \infty$, and as it is known that $S_p(t) = \mathrm{e}^{-\gamma(t)}$, eqn [A7] simplifies to

$$J = \int_0^\infty [V(n,t) + \psi(n(t),b(t))JX^*((X(t))]\mathrm{e}^{-rt-\gamma(t)}\,\mathrm{d}t \qquad [A8]$$

where the change of variables allows $\dot{\gamma} = \psi(n,b)$, $\gamma(0) = 0$ so that $S_p(0) = 1$.

For the derivation of ρ:

From the current value Hamiltonian

$$H = \mathrm{e}^{-\gamma}(V(n) + \psi(n,b)JX^*) + \mu\psi(n,b)$$

the first-order necessary condition of the maximum principle can be derived

$$\frac{\partial H}{\partial n} = \mathrm{e}^{-\gamma}(V'(n) + \psi_n JX^* + \mu\psi_n \leq 0$$

and also the equation of motion for the current value costate variable (shadow price)

$$\dot{\mu} = r\mu - \frac{\partial H}{\partial \gamma} = r\mu - \frac{\partial(\mathrm{e}^{-\gamma}(V(n) + \psi(n,b)JX^*) + \mu\psi(n,b)}{\partial \gamma}$$

$$= rm + \mathrm{e}^{-\gamma}(V(n) + \psi(n,b)JX^*)$$

If one defines $\rho = e^{\gamma}\mu$, then the conditional current value Hamiltonian is $\overline{H} = He^{\gamma}$. One can then time differentiate ρ to find $\dot{\rho} = \dot{\mu}e^{\gamma} + \dot{\gamma}e^{\gamma}\mu = e^{\gamma}(\dot{\mu} + \dot{\gamma}\mu) = e^{\gamma}(r\mu + e^{-\gamma}(V(n)) + \psi(n,b)JX^*) + \psi(n,b)\mu)$, which, given the definition of $\rho = e^{\gamma}\mu$, can be written as

$$\dot{\rho} = (r\rho + V(n) + \psi(n,b)JX^* + \rho\psi(n,b))$$

$$= \rho(r + \psi(n,b)) + V(n) + \psi(n,b)JX^*$$

To determine the key equation, the authors time differentiate the interior solution for the first-order condition, rearrange terms, and resubstitute from the maximum principle

$$\dot{n} = \frac{\psi_n(\psi + r)(V'(n) + \psi_n JX^*) + [V'(n)\psi_{nb}]\dot{b} - [V(n) + \psi JX^*]\psi_n^2}{V''(n)\psi_n - V'(n)\psi_{nn}}$$

Appendix B

List of Courses

1. Resource Economics
2. Environmental Economics
3. Environmental Management

Acknowledgment

Thanks to the financial support of the Program of Research on the Economics of Invasive Species Management (PREISM), ERS – USD, NSF (DEB 02-13698) and NOAA CSCOR grant number NA09NOS4780192.

See also: Agricultural Invasions. Plant Invasions

References

Archer D and Shogren J (1996) Endogenous risk in weed control management. *Agricultural Economics* 14: 103–122.

Barbier E (2001) A note on the economics of biological invasions. *Ecological Economics* 39: 197–202.

Barbier E and Shogren J (2004) Growth with endogenous risk of biological invasion. *Economic Inquiry* 42: 587–601.

Brown C, Lynch L, and Zilberman D (2002) The economics of controlling insect-transmitted plant diseases. *American Journal of Agricultural Economics* 84: 279–291.

Burnett K, Kaiser B, Pitafi BA, and Roumasset JA (2006) Containment of invasive species: Illustrations from Hawaii. *Agricultural and Resource Economics Review* 35: 63–77.

Burnett KM, D'Evelyn S, Kaiser BA, Nantamanasikarn P, and Roumasset JA (2008) Beyond the lamppost: Optimal prevention and control of the brown tree snake in Hawaii. *Ecological Economics* 67: 66–74.

Calvin L and Krissoff B (1998) Technical barriers to trade: A case study of phytosanitary barriers and US–Japanese apple trade. *Journal of Agricultural and Resource Economics* 23: 351–366.

Committee on Environment and Natural Resources (CENR) 1999 *Ecological Risk Assessment in the Federal Government*. National Science and Technology Council (NSTC). Executive Office of the President of the United States.

Crocker T and Tschirhart J (1993) Ecosystems, externalities, and economics. *Environmental and Resource Economics* 2: 551–567.

Ehrlich I and Becker G (1972) Market insurance, self-insurance, and self-protection. *Journal of Political Economy* 80: 623–648.

Eiswerth ME and Johnson WS (2002) Managing nonindigenous invasive species: Insights from dynamic analysis. *Environment and Resource Economics* 23: 319–342.

Elton C (1958) *The Ecology of Invasions by Animals and Plants*. London: Methuen.

ESA (2004) *INVASION!* Ecological Society of America. http://www.esa.org/education/edupdfs/invasion.pdf

Federal Interagency Committee for the Management of Noxious and Exotic Weeds (FICMNEW) (1998) *Invasive Plants: Changing the Landscape of America*. Washington, DC: Fact Book.

Finnoff D, Lewis MA, and Potapov AB (2009a) Second best policies in invasive species management: When are they "good enough"? In: Perrings C, Mooney H, and Williamson M (eds.) *Bioinvasions and Globalization: Ecology, Economics, Management, and Policy*. Oxford: Oxford University Press.

Finnoff D, Lewis MA, and Potapov AB (2010b) Control and the management of a spreading invader. *Resource and Energy Economics* 32: 534–550.

Finnoff D, McIntosh C, Shogren JF, Sims C, and Warziniack T (2010a) Invasive species and endogenous risk. *Annual Review of Resource Economy* 2: 77–100.

Finnoff D, Settle C, Shogren JF, and Tschirhart J (2009b) Integrating economics and biology for invasive species management. In: Keller RP, Lodge DM, Lewis

MA, and Shogren J, (eds.) *Bioeconomics of Invasive Species: Integrating Ecology, Economics, Policy, and Management,* vol. 2. pp. 25–43. New York: Oxford University Press.

Finnoff D, Shogren JF, Leung B, and Lodge DM (2007) Take a risk – preferring prevention over control of biological invaders. *Ecological Economics* 62: 216–222.

Horan RD and Fenichel EP (2007) Economics and ecology of managing emerging infectious animal diseases. *American Journal of Agricultural Economics* 89: 1232–1238.

Jerde CJ and Lewis MA (2007) Waiting for invasions: A framework for the arrival of non-indigenous species. *The American Naturalist* 170: 1–9.

Josling T, Roberts D, Orden D (2004) *Food Regulation and Trade: toward a Safe and Open Global System*. Peterson Institute.

Kamien MI and Schwartz NL (1991) *Dynamic Optimization: the Calculus of Variations and Optimal Control in Economics and Management*. Amsterdam: North-Holland.

Kareiva P (1996) Developing a prediction ecology for non-indigenous species and ecological invasions. *Ecology* 77: 1651–1652.

Keller R, Lodge D, Lewis MA, and Shogren J (2009) *Bioeconomics of Invasive Species: Integrating Ecology, Economics, Policy, and Management*, 298 pp. New York: Oxford University Press.

Kim CS, Lubowski RN, Lewandrowski J, and Eiswerth ME (2006) Prevention or control: Optimal government policies for invasive species management. *Agricultural and Resource Economics Review* 35: 29–40.

Knowler D and Barbier EB (2005) Managing the Black Sea anchovy fishery with nutrient enrichment and a biological invader. *Marine Resource Economics* 20: 263–285.

Leung B, Lodge DM, Finnoff D, Shogren JF, Lewis MA, and Lamberti G (2002) An ounce of prevention or a pound of cure: Bioeconomic risk analysis of invasive species. *Proceedings: Biological Sciences* 269: 2407–2413.

Levy A and Caputo MR (2008) Optimal control of locusts in subsistence farming areas. *European Journal of Operational Research* 191: 503–515.

Lewis MA, Potapov AB, and Finnoff DC (2009) Modeling integrated decision-making responses to invasive species. In: Keller RP, Lodge D, Lewis MA, and Shogren J (eds.) *Bioeconomics of Invasive Species: Integrating Ecology, Economics, Policy, and Management*, pp. 180–204. New York: Oxford University Press.

Lichtenberg E and Lynch L (2006) Exotic pests and trade: When is pest free status certification worthwhile? *Agricultural and Resource Economic Review* 35: 52–62.

Lichtenberg E and Zilberman D (1986) The econometrics of damage control: Why specification matters. *American Journal of Agricultural Economics* 68: 261–273.

Lodge D, Lewis MA, Shogren J, and Keller RP (2009) Introduction to biological invasions: Biological, economic, and social perspectives. In: Keller RP, Lodge DM, Lewis MA, and Shogren J (eds.) *Bioeconomics of Invasive Species: Integrating Ecology, Economics, Policy, and Management,* 1: pp. 1–24. New York: Oxford University Press.

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, and Bazzaz FA (2000) Biotic invasions: Causes epidemiology, global consequences, and control. *Ecological Applications* 10: 689–710.

Miller M and Fabian R (2004) *Harmful Invasive Species: Legal Responses*. Washington, DC: Environmental Law Institute.

National Invasive Species Counsel (2001) *Management Plan: Meeting the Invasive Species Challenge*. January 18. Washington, DC: National Invasive Species Counsel.

Neher P (1990) *Natural Resource Economics: Conservation and Exploitation*. New York: Cambridge University Press.

Olson L and Roy S (2002) The economics of controlling a stochastic biological invasion. *American Journal of Agricultural Economics* 84: 1311–1316.

Olson L and Roy S (2008) Controlling a biological invasion: A non-classical dynamic economic model. *Economic Theory* 36: 453–469.

Perrings C, Mooney HA, and Williamson MH (eds.) (2010) *Bioinvasions and Globalization: Ecology, Economics, Management, and Policy*. Oxford: Oxford University Press.

Peterson E and Orden D (2008) Avocado pests and avocado trade. *American Journal of Agricultural Economics* 90: 321–355.

Pontryagin LS, Boltyanskii VG, Gamkrelize RV, and Mishchenko EF (1962) *The Mathematical Theory of Optimal Processes*. Wiley.

Potapov AB, Finnoff DC, and Lewis MA (2007) Optimal control of biological invasions in lake networks. *Natural Resource Modeling* 20: 351–379.

Reed WJ (1987) Optimal preventative maintenance, protection, and replacement of a revenue-earning asset. *Applied Mathematics and Computation* 24: 241–261.

Reed WJ and Heras HE (1992) The conservation and exploitation of vulnerable resources. *Bulletin of Mathematical Biology* 54: 185–207.

Sala OE, Chapin III FS, Armesto JJ, *et al.* (2000) Biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.

Shogren JF (2000) Risk reductions strategies against the explosive invader. In: Perrings C, Williamson M, and Dalmazzone S (eds.) *The Economics of Biological Invasions*, pp. 56–69. Northhampton, MA: Edward Elgar.

Shogren J and Crocker T (1991) Risk, self-protection, and *ex ante* valuation. *Journal of Environmental Economics and Management* 20: 1–15.

Smith J (2003) *Exotic Pests and Diseases: Biology and Economics for Biosecurity*. Iowa: Ames.

Tahvonen O and Salo S (1996) Nonconvexities in optimal pollution accumulation. *Journal of Environmental Economics and Management* 31(2): 160–177.

Economic Growth and the Environment

Karl-Göran Mäler, Beijer International Institute of Ecological Economics, Stockholm, Sweden

© 2013 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp. 277–284, © 2001, Elsevier Inc., with revisions made by the Editor.

Introduction

The title of this article has many meanings. Economic analysis of environment and growth can mean many different things. In this article, the focus is on the following issues:

- What is economic growth?
- Will economic growth deteriorate or improve the environment?
- What is the role of environmental resources in economic growth?

The first question is typically answered by growth in gross domestic product (GDP) or GDP per capita. However, seldom is it explained why this is an interesting definition of growth. Of course, behind this notion of growth is the notion that GDP or gross national product (GNP) is a measure of well-being in society. It is therefore necessary to discuss what we mean by well-being and possible indices for such a concept. This will, of course, immediately bring us to a discussion of green accounting. There is a huge literature on this subject, but in this article the author will not try to survey those contributions but instead examine one particular formulation of the conceptual basis of green accounting in order to determine the difficulties in creating such an index (The discussion of these issues is based on Dasgupta and Mäler (2000) and Mäler (1999)).

The second question has been discussed in terms of so-called Kuznets curves – that is, empirically estimated relations between environmental quality and GDP per capita. The general finding for many pollutants is that a country with a very low income does not have much pollution but when the scale of the economy grows, for example, as measured by GDP per capita, emissions of these pollutants will increase. However, when the income per capita is high enough, the economy will reach a turning point and pollution will decrease with further increases in per capita income. This has been taken by the World Bank to mean that growth implies a win–win strategy. By growing, a country may in the long run obtain both increased material well-being and a better environment. Of course, this kind of analysis has been criticized, from conceptual points of views and also from econometric and empirical points of view. Once again, the literature is huge, and no attempt to summarize this literature will be done. However, some central issues will be discussed.

Finally, the role of environmental resources as inputs for the growth process has very seldom been studied. When Solow (1957) studied economic growth in the US, he found that most of the growth could not be explained by the two traditional factors of production (capital and labor), and he attributed the "unexplained" growth or the residual to technical progress. Later investigations by Dale Jorgenson, Zwi Grilliches, and others (Jorgenson (1995a, 1995b) provides masterful surveys of these issues) reduced the residual considerably (e.g., by using better measurements of capital and by incorporating human capital), but it still remains quite high. This means that economists basically cannot account for all the productivity growth Western economies experienced during the past 100 years. To solve this problem, the theory of endogenous growth was developed during the 1980s and the beginning of 1990s. The main idea is to incorporate the externalities that research and development activities generate because the benefits from these cannot be captured completely by the companies that carry out these activities. These externalities will therefore contribute to growth without being accounted for in the usual way that we analyze productivity growth. However, there are other externalities that also should be accounted for, such as networking externalities or social capital (Dasgupta, 1999). In this context, there is one particular externality that is of interest – environmental or resource externality. Because environmental resources or natural resources in general often lack well-defined property rights, they are not transacted on markets and therefore not accounted for in the standard national accounts (SNAs.) This omission of natural capital in growth accounting studies may explain at least part of the residual.

Net National Product (NNP)

Aggregate measures such as GNP or GDP were not constructed to measure welfare or well being. They were aimed at measuring total monetary income in an economy and the total supply of goods produced in man-made facilities to provide economic planners with a database suitable for economic policies. However, despite warnings from many economists, GNP and GDP became used as indices of welfare. During approximately the past 25 years, economists have tried to define welfare and construct indices measuring welfare, and there is now a substantial literature on the subject. Instead of reviewing the literature, the author will discuss some issues not very much discussed. Furthermore, most of the literature is based on the assumption that the economy is following an optimal growth path, which is rarely a realistic assumption. Instead, we will assume that we are able to make forecasts arbitrarily into the future (The basic reason behind the assumption of an optimal growth path in much of the literature is that predictions become very simple. It will be shown that nothing essential will change when we switch to arbitary allocation mechanisms).

The issues that will be discussed in this section are

- The notion of social well being
- Social well-being and population growth
- The search for a linear index – NNP as a marginal social cost–benefit criterion for elementary projects
- The search for a linear index – NNP as a measure of sustainable development
- Sustainable development and technical progress.

Social Well-Being

To discuss social well-being, it is useful to imagine a social planner who is in charge of planning an economy from now (date 0) to infinity. Koopmans (1960,1972a,b) has convincingly shown that the rational social planner should evaluate different policies with regard to how they affect the present value of the stream of future utilities, where the utility function is strictly concave and is a function of all the factors that influence the current well being. In other words, the social welfare function is written

$$W = \int_0^{\infty} e^{-\delta t} u(c(\tau))\mathrm{d}\tau \qquad [1]$$

where $c(\tau)$ is the rate of consumption at moment τ. In principle, $c(\tau)$ is a vector, the components of which represent all the factors that affect well-being at time τ. However, it will not change any conclusions if we restrict ourselves to the interpretation that c is the aggregate consumption good.

One of Koopmans' conditions is that $\delta > 0$; that is, the discount rate must be strictly positive. The discount rate is necessary in order to be able to define a social well ordering over all possible future paths of utilities. However, this is the rate for discounting utilities. The rate for discounting consumption will in general be different and depends on the growth rate of consumption. The formula is

$$r = \delta + \eta g \qquad [2]$$

where r is the consumption discount rate, δ is the utility discount rate, ν is the elasticity of marginal utility of consumption, and g is the growth rate of consumption. It is clear that if δ is small and the growth rate is sufficiently negative, the social planner will want to discount future costs and benefits measured with consumption as a numeraire (i.e., the way we usually measure them in cost–benefit analysis) with a negative interest rate.

The strict concavity of the utility function guarantees an attitude in favor of equalization between generations. Holding everything else unchanged, a transfer from a rich generation to a poor generation will increase social welfare.

Mathematical Preludes

No in-depth mathematics will be used in this section and an elementary knowledge of calculus should be enough for the derivations made here. However, reading this section is not necessary for an understanding of later sections.

At time t, the economy has inherited assets from the past. Denote the vector of these assets K_t. Assume that the social planner can predict the future flow of consumption, given K_t. To do this, the social planner must know how these assets are going to be used (now and in the future) for consumption and accumulation of some of the assets and perhaps depletion of others (in particular, stocks of nonrenewable resources). Such predictions can be summarized by an allocation mechanism, which will be called α:

$$c(\tau) = \alpha(\tau, t, K_t), \forall \tau > t \qquad [3]$$

Thus, the mapping α is the information the social planner uses for mapping forecasts.

Note that α is not only a function of the initial stocks but also of calendar time t. The reason for this is that there may be autonomous changing over time. If there is exogenous technical progress, for example, future predicted consumption will of course depend on this and α must therefore depend on calendar time. Other reasons for autonomous changes in the allocation mechanism may be trends in trade and forecasted changes in government policies. The author will discuss these issues later.

Given the allocation mechanism, the social welfare beginning at time t will be a function of the capital stocks at t, the allocation mechanism, and the calendar time. Thus,

$$W = V(t, \alpha, K_t) \qquad [4]$$

V gives the social welfare achievable from the initial stocks K_t with the allocation mechanism α.

A marginal change in one of the stocks at time t will change the welfare, and this change is basically the marginal value of that stock; that is,

$$p_i(t) = \frac{\partial V(t, \alpha, K_t)}{\partial K_i} \qquad [5]$$

where K_{it} is the stock of asset i at time t. In most of the literature on green accounting, the accounting prices are those that support an optimal growth path. The prices defined here do not. However, they reflect the true values of the initial stocks at time t.

It is practical to define the Hamiltonian H at time t as

$$H_t = u(c(t)) \sum_{i=1}^{N} P_{it} \frac{\mathrm{d}K_{it}}{\mathrm{d}t} \qquad [6]$$

One can now show that (see Mäler (1999) for a proof)

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \delta p_i - \frac{\partial H_t}{\partial K_i}, i = 1, 2, \ldots, N \qquad [7]$$

Although these equations are formally identical to the equations governing the costate variables in Pontryagin's maximum principle, they do not reflect any optimization except that the prices are the true measures of the marginal social values of the initial stocks.

Population and Welfare

So far, the author has neglected population changes in the discussion of social welfare. However, we know that humans have a finite life and new generations will be born. How do we take this into account when we try to formulate a welfare function for a society?

The simplest way is perhaps to assume that population changes exogenously in a predictable manner. Then, at each moment in time we can represent the welfare of the cohort by the utility of an individual with average per capita consumption multiplied by the number of individuals. However, this is seriously flawed in that fertility behavior is not exogenous. We now know that human reproduction behavior is governed by many social and economic factors. There are many studies showing that reproduction depends on the need for support during retirement, on the need for labor, on the need for risk diversification, and of course because children

can be regarded as consumption goods – that is, they give immediate welfare to the parents. Dasgupta (1993, 1997) provides discussions of these issues and references to empirical material. Thus, we should treat fertility as endogenous, something that depends on the economic system in all its characteristics. However, this is exactly what the allocation mechanism introduced in the previous section represents. Thus, the population at any particular future point of time is determined by α.

It is now easy to assume that the welfare function can be written

$$\int_0^\infty e^{-\delta t} L(\alpha(0,K_0))u(c(t))\mathrm{d}t \qquad [8]$$

However, this formulation does not take into account the fundamental difference between a potential human being and an actual human. What is the value of having one more individual to that individual? There is no reasonable answer to this question (see Dasgupta (1997) for an extremely interesting discussion of these issues). So far, we do not have an ethical base for assessing welfare to endogenous population changes.

A Marginal Cost–Benefit Rule

Define NNP with utility as the numeraire as the linearized Hamiltonian

$$\mathrm{NNP} = u^i(c)c + \sum_{i=1}^{N} p_{it}\frac{\mathrm{d}K_{it}}{\mathrm{d}t} \qquad [9]$$

If, instead, we choose consumption at time 0 as the numeraire, NNP is defined by

$$\mathrm{NNP} = c + \sum_{i=1}^{N} \tilde{p}_{it}\frac{\mathrm{d}K_{it}}{\mathrm{d}t} \qquad [10]$$

where $\tilde{p}_{it}$ is $p_{it}/u'(c(0))$.

Note that NNP is a linear index in all quantities, which makes it an interesting measure. A nonlinear measure such as the Hamiltonian is hopeless in terms of empirical applications. We cannot estimate all the consumers' surpluses needed to get an empirical estimate of the Hamiltonian.

Second, note that c is basically a proxy for all variables that affect current utility. Therefore, all these variables must be included in a generalized consumption set. Among these variables is the disutility of labor, which leads us to the conclusion that the wage bill, as long as the wage rate reflects the disutility of work, should be deducted. Immediate environmental damages should also be deducted (Mäler (1991) discusses these issues in detail.). In all these cases the damages should be valued at a rate that reflects the marginal disutility of the environmental damage with consumption as the numeraire.

Third, note that the investment part of the NNP must include net changes in all assets that will affect future well being. This means of course that the current narrow asset boundary in SNAs must be substantially extended.

Also note that the asset prices p_i must correctly reflect the social scarcity values captured by eqn [7]. These prices will not always coincide with market prices. However, surprisingly often, market prices will capture the social scarcity values of assets that are transacted on markets. However, for all the ecological assets, the values must be estimated by modeling the ecological systems and their connectedness. There is a growing literature on these procedures (Freeman (1992) provides a good introduction).

It should be noted that the NNP measure so defined deviates radically from the United Nations recommendations. The NNP measure we have defined has one important property, namely, that if there is a small perturbation (so that price changes are small) of the allocation mechanism during a small time interval, and if this small perturbation increases the social well-being, then it will also increase NNP and vice versa (For a proof, see Mäler (1999) or Dasgupta and Mäler (2000)). This is what makes NNP interesting – a linear marginal cost–benefit criterion with which we can evaluate such perturbations (which henceforth will be called policy reforms).

Here, we consider discreet time, and the following example will indicate the use of the theorem stated previously. Assume that there is one allocation mechanism α and we consider a perturbation or policy reform $\Delta\alpha$ that for the first period prescribes the same allocation as α did for the second period. We compute the change in NNP when we make this policy reform and this change is positive. This can be interpreted as saying that the second period gives higher social welfare than the first period. By following the allocation mechanism α, welfare increases over time. However, note that this depends on the assumption that the prices do not change much.

This theorem can now easily be generalized to situations in which we take atemporal equity into account, and even population changes can be accommodated. The trick is to find the appropriate shadow prices.

The Hamiltonian as Constant-Equivalent Utility

In the previous two sections it was shown that NNP can be used as an index for conducting social cost–benefit analysis of policy reforms. However, the theoretical literature on green NNP has been directed toward a quite different end. It has argued that NNP measures "constant-equivalent consumption."

It can be shown that the present value of a constant stream of utility equal to the Hamiltonian at time 0 is the maximum feasible constant utility stream that an economy can afford. This was first shown by Weitzman in the case in which the utility function is linear. In this case, NNP and the Hamiltonian coincide, and NNP gives the maximum constant-equivalent consumption.

However, a linear utility function is ethically flawed: It is insensitive to distributional issues. Furthermore, a large body of evidence concerning household saving behavior is at odds with linear utility functions.

Social Well-Being and the Concept of Sustainability

The World Commission (1987) defined "sustainable development" as an economic program in which, loosely speaking, the well-being of future generations is not jeopardized.

There are many possible interpretations of this. Consider the following:

a. An economic development is sustainable if $dU_t/dt > 0$, where $U_0 > \lim U_t$ as t tends to 0.
b. An economic development is sustainable if $dU_t/dt > 0$.
c. An economic development is sustainable if $dV/dt > 0$, where $V_t = \int_0^\infty e^{-\delta t} U(C_t) dt$

It is clear that (a.) lacks ethical foundation. For example, it may be desirable to reduce U in the short run in order to accumulate assets so that the flow of U is still higher in the future. In this sense, (b.) offers greater flexibility in ethical reasoning: It permits initial sacrifices in the current standard of living, U (a burden assumed by the generation engaged in the reasoning), but requires that no future generation should have to experience a decline in its standard of living.

In contrast to (b.), the focus of (c.) as a notion of sustainable development is social well-being, V. The criterion permits the first generation to make initial sacrifices in V (relative to the past) but requires that social well-being should never decline in the future. Note that although (b.) implies (c.), (c.) does not imply (b.). In other words, (c.) is more general. In what follows, we adopt (c.) as our notion of sustainable development and develop criteria for judging whether a given economic program represents sustainable development.

One can now easily show that

$$\frac{dV}{dt} = \frac{\partial V}{\partial t} + \sum_i^N p_i \frac{dk_i}{dt}$$

The first term on the right-hand side is the autonomous change in V, reflecting exogenous technical progress or exogenous changes in terms of trade. This term has been discussed by the authors in the last section. For now, we assume that there are no autonomous changes. The previous equation simply indicates that the rate of change in welfare is equal to the value of the net investment in the economy. Thus, an economic program represents sustainable development if, and only if, the program net investment in the economy's capital assets is always nonnegative.

The result has intuitive appeal. It says that social welfare is higher today than it was yesterday if the economy is wealthier today. Here, an economy's "wealth" is interpreted as the accounting value of all its capital assets, and wealth comparisons are made at constant prices. In a famous article, Samuelson (1961) argued in connection with national income accounting that welfare comparisons should deal with "wealth-like" entities. Proposition 4 formalizes this insight.

Note, however, that what we have obtained is an equivalence result: the theorem cannot on its own tell us if sustainable development is feasible. Whether the economy is capable of growing wealthier indefinitely depends on, among other things, the extent to which different assets are substitutable in production.

This idea of examining changes in wealth dates back to Samuelson (1961). The connection to sustainable development was first suggested by Pearce and Atkinson (1993), although they did not derive the theorem rigorously and they did not interpret it as an equivalence result. The idea was then investigated by the World Bank, which published many attempts to measure the value of net investment or genuine savings (see Expanding the Measure of Wealth (1997) and Kunte and Clemens (1998) for further details.)

Note that we cannot in general use NNP to determine whether the economy is on a sustainable path. Only if the prices are stationary is an increasing NNP equivalent to saying that the economy is on a sustainable path.

In conclusion, NNP can be used for cost–benefit analysis in situations in which prices do not change very much. In particular, in such a situation it may give some guidance as to whether social welfare is increasing or not. Note that there is no conflict between growth in the NNP and the concern for the environment. These concerns have been incorporated into the linear index. If the economy is on a path in which environment is degraded, this will be captured by the NNP, and it can be expected that NNP will decrease if the degradation is sufficiently large.

NNP is not useful for determining whether the economy is on a sustainable path because the relative prices will change along the path. Here, concepts related to wealth become more interesting. In particular, the value of the net investments in a time period gives a correct criterion for judging whether the economy is on a sustainable path. However, it is only a necessary condition. The sufficient criterion is that in each future period, the value of net investments must be positive. *A priori*, we do not know whether this is feasible.

Environmental Kuznets Curves

Since the early 1990s, economists and econometricians have been industrious in estimating empirical relations between GDP and environmental quality indicators. Typically, they have used cross-sectional data – country data for a particular year. For some quality indicators, they have found a relation that looks like an inverted U (**Figure 1**).

The interpretation is simple. Very poor countries with low industrialization are not damaging the environment much and as a result the environmental quality indicator is good (in this case, a low sulfur emission). When the country starts the growth process, the environmental damages will increase. When income per capita increases, the demand for a cleaner environment will increase, environmental legislation will be enacted, and the environment will improve. Because the curve is similar (but inverse) to the curves Kuznets (Economic Growth and Income Inequality, 1955) estimated for the relationship between income inequality and GDP per capita, they

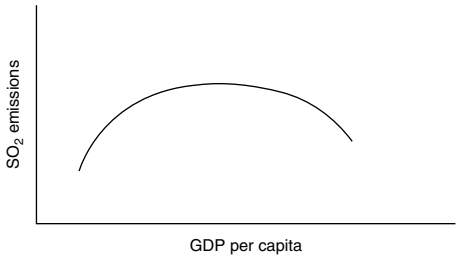


Figure 1 An estimated empirical relation between GDP and environmental quality indicators.

have been known as the environmental Kuznets curves. (For an overview and examples of environmental Kuznets curves, see *Environment and Development*, 2 (4), October 1997.)

The breakthrough for this kind of analysis came with the World Development Report (1992), which was devoted to the environment. The existence of these curves was taken to be an indication of possible win–win situations. By stimulating growth, one would also in the long run stimulate environmental regulations. It is clear, however, that the analysis is quite complex.

First, environmental Kuznets curves have only been established for a few environmental indicators. Second, at least earlier, the specifications of the relation between the quality measure and GDP per capita were quadratics, which almost certainly forced the curves to look like an inverse U. Third, the studies did not include environmental spillovers between countries. Fourth, they did not examine possible relations between indices of freedom and the environment. It has been shown that whenever people have had the right to voice concern about environmental degradation, the environment has improved, irrespective of the per capita income. Locally, in many cultures people have had a direct influence on, for example, local water pollution, which has improved even in very poor countries. It is therefore possible that the environmental Kuznets curves measure not only the effect of income growth but also growth in freedom.

However, most important is that the environmental Kuznets curves – if they exist in any meaningful way – are not optimal in any sense of the concept. Typically, a country would do much better to develop institutional and regulatory reforms for better environmental management than just trust that economic growth will automatically solve environmental problems. In this sense, the belief in environmental Kuznets curves and the associated belief in win–win situations are very dangerous.

Technological Change and Growth Accounting

Productivity growth as discussed previously is productivity growth in GDP in which the flow of services that are not transacted on markets are excluded (with of course some important exceptions, e.g., government output). Here, we are interested in productivity growth in NNP as defined previously. There is no reason why productivity growth in GDP would correspond to a similar growth of NNP (In fact, Repetto *et al.* (1989) showed that by including depreciation of oil reserves, deforestation, and soil erosion in Indonesia, the growth rate was reduced to approximately half the official growth rate in GDP). Furthermore, if, as Jorgenson (1995a, ch. 1) pointed out, changes in energy prices affected the observed productivity growth rate substantially, it should be clear that the utilization of other natural resources should also affect the growth rate of the total production. Thus, we should analyze the following issues:

Growth in NNP
The role of natural capital in the growth of NNP

The following calculations are very speculative, but the author hopes that someday someone will find it worthwhile to study the empirical consequences from the analysis. The simplest possible model to demonstrate the main ideas has been used (which probably will be seen to be perhaps too obvious).

Let there be one aggregate good x which is produced with man-made capital K_x and with one input representing natural capital services, denoted γ. Thus,

$$x = f(K_x, \gamma) \qquad [11]$$

Assume that the production function is homogeneous to degree 1.

The total output is used for consumption c, for investiment in man-made capital $I = K_x$, and for investment in man-made capital for exploiting the natural resources $I_\gamma = K_\gamma$. Let the production function for γ be

$$\gamma = g(K_\gamma) \qquad [12]$$

Thus, γ is an intermediary good in the system and x will correspond to GDP:

$$x = c + K_x + K_\gamma = c + K \qquad [13]$$

Assume that capital is freely mobile between the two uses so that we can reasonably talk about the total capital stock $K = K_x + K_\gamma$. Then we can define the productivity growth in GDP as

$$\theta = \frac{1}{x}\frac{\mathrm{d}x}{\mathrm{d}t} - \frac{1}{K}\frac{\mathrm{d}K}{\mathrm{d}t} \qquad [14]$$

Now, let the dynamics of the natural or ecological capital K_E be

$$\frac{\mathrm{d}K_E}{\mathrm{d}t} = \phi(K_E) - \gamma \qquad [15]$$

Assume that there is an institutional failure in the economy so that there are no assigned property rights to natural capital and therefore no market transactions in K_E. This is of course the reason why investments in natural capital are not included in GDP.

NNP according to the definition in equation is

$$NNP = c + p_E\frac{\mathrm{d}K_E}{\mathrm{d}t} + \frac{\mathrm{d}K}{\mathrm{d}t} = x + p_E\frac{\mathrm{d}K_E}{\mathrm{d}t} \qquad [16]$$

It is clear from eqn [16] that a positive productivity growth for GDP can be offset by depreciation of the stock of natural capital so that the productivity growth of NNP is 0 or even negative. In fact, if we assume that $\dot{K} = 0$, we have

$$\frac{1}{NNP}\frac{\mathrm{d}NNP}{\mathrm{d}t} - \frac{1}{K_E}\frac{\mathrm{d}K_E}{\mathrm{d}t} = \frac{GDP}{NNP}\theta + \frac{\dot{K}_E}{K_E}\left\{\frac{p_E K_E}{NNP}\frac{\ddot{K}_E}{\dot{K}_E} - 1\right\} \qquad [17]$$

The last bracket in eqn [17] will surely be negative because of the overexploitation of the resource as a result of the institutional failure. Thus, this bracket will compensate for the positive productivity growth of GDP. In this case, because we know that the production functions are stationary, we know for sure that the productivity growth rate of NNP (if the production functions have constant returns to scale in K_x, K_γ, and K_E) must be 0. We can therefore conclude that even if measured GDP growth rates show a positive productivity

growth, this does not necessarily imply that NNP will show growth irrespective of factor inputs.

This result also follows from the observation that conventional SNA regards the inputs of resources as intermediary goods (if there are costs in producing them and it neglects them completely if they are supplied free of any costs to the production processes), and the estimation of an aggregate production function with only labor and capital as inputs, necessarily will be biased because resources are the basis for material production. Thus, the total factor productivity of GDP should be measured as

$$\frac{1}{x}\frac{\mathrm{d}x}{\mathrm{d}t} - w_K \frac{1}{K}\frac{\mathrm{d}K}{\mathrm{d}t} - w_v \frac{1}{v}\frac{\mathrm{d}v}{\mathrm{d}t} \qquad [18]$$

where w_K is the value share of man-made capital and w_v is the value share of the resource input. This will necessarily be zero if the production function is linearly homogeneous and stationary. Thus, it may be that the neglect of resource inputs in the production functions is an important explanatory variable of the measured positive factor productivity growth. However, the only way to find out whether this is important is to carry out new empirical measurements. In conclusion, it does not seem reasonable to assume autonomous productivity growth in models designed to create a new accounting system for welfare measurements. Therefore, the argument for a nonstationary value function is not necessarily true.

Appendix

List of Courses

1. Environmental Management Systems Institute, University of Florida, USA
2. California Polytechnic's Orfalea College of Business, San Luis Obispo, USA
3. St Andrews Centre for Social and Environmental Accounting Research, University of St. Andrews, Scotland
4. School of Business and Informational Management, The Australian National University, Australia
5. Osaka City University, Japan
6. Chuo Graduate School of International Accounting, Kobe University, Japan

See also: Biodiversity as a Commodity. Economic Value of Biodiversity, Measurements of. Market Economy and Biodiversity

References

Dasgupta P (1993) The population problem. *Journal of Economic Literature* 1879–1902.
Dasgupta P (1997) The population problem. *Journal of Ecological Econo*mics.
Dasgupta P (1999) Social capital. Economic growth and income inequality (1955). *American Economic Review* 45: 1–28.
Dasgupta P and Mäler K-G (2000) Net national product, wealth and social well-being. *Environment and Development Economics* 5(1 and 2): 69–94.
Freeman III A (1992) *The Measurement of Environmental and Resource Values*. Washington, DC: Resources for the Future.
Jorgenson D (1995a) *Productivity, vol. 1*. Cambridge: MIT Press.
Jorgenson DW (1995b) *Productivity, vol. 2*. Cambridge: MIT Press.
Koopmans T (1960) Stationary ordinal utility and impatience. *Econometrica* 28: 287–309.
Koopmans T (1972a) Representation of preference orderings with independent components of consumption. In: McGuire CB and Radner R (eds.) *Decision and Organization*. Amsterdam: North-Holland.
Koopmans T (1972b) Representation of preferences ordering over time. In: McGuire CB and Radners R (eds.) *Decision and Organization*. Amsterdam: North-Holland.
Kunte AKHJD and Clemens M (1998) Estimating national wealth: Methodology and results. *Technical Report*. Washington, DC: World Bank, Environment Department.
Mäler KG (1991) National accounting and environmental resources. *Environmental and Resource Economics* 1–5.
Mäler K-G (1999) National income, welfare measurements and local prices.
Pearce D and Atkinson G (1993) Capital theory and the measurement of sustainable development: An indicator of weak sustainability. *Ecological Economics* 8(2): 103–108.
Repetto R, Magrath WWMBC, and Rossini F (1989) *Wasting Assets: Natural Resources and the National Income Accounts*. Washington, DC: World Resources Institute.
Samuelson P (1961) The evaluation of social income: Capital formation and wealth. In: Lutz F and Hague D (eds.) *The Theory of Capital*. London: Macmillan.
Solow R (1957) Technological progress and productivity change. *Review of Economics and Statistics* 39: 312–320.
World Bank (1997) *Expanding the Measure of Wealth*. Washington, DC: World Bank.
World Development Report (1992) New York: Oxford University Press.

Relevant Websites

http://data.worldbank.org/data-catalog/environmental-accounting
http://unstats.un.org/unsd/envaccounting/seea.asp
System of Environmental-Economic Accounting (SEEA).

Economics of Agrobiodiversity

Unai Pascual, University of Cambridge, Cambridge, UK; Basque Centre for Climate Change (BC3), Bilbao, Spain, and IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
Louise E Jackson, University of California, Davis, CA, USA
Adam G Drucker, Bioversity International, Rome, Italy

© 2013 Elsevier Inc. All rights reserved.

Glossary

Agrobiodiversity The variety and variability of living organisms that contribute to food and agriculture in the broadest sense, and that are associated with cultivating crops and rearing animals within ecological complexes. It is further expanded in some contexts to include all the organisms present in an agricultural landscape. Examples consist of crops and animal breeds, their wild relatives, and the species that interact with and support these species.
Direct use value The economic value derived from direct use or interaction with a biological resource or resource system.
***Ex situ* conservation** A conservation method that entails the removal of germplasm resources (seed, pollen, sperm, and individual organisms) from their original habitat or natural environment with the aim to keeping components of biodiversity alive outside their original habitat or natural environment. It includes gene bank storage (seed and field), *in vitro* storage, pollen storage, and DNA storage. Seed gene banks are the most common storage practice. In the case of plant breeding it generally involves the collection, classification, evaluation, and utilization of agrobiodiversity.
***In situ* conservation** A conservation method that attempts to preserve the genetic integrity of gene resources by conserving them within the evolutionary dynamic ecosystems of the original habitat and seeks to maintain and recover viable populations of species in their natural surroundings. It is also known as on-farm conservation, which encourages farmers to continue selection and management of local crop populations.
Indirect use value The economic value associated with the services supported by a resource as opposed to the actual use of the resource itself.
Landrace A local variety of a domesticated animal or plant species that has developed largely by natural processes, by adaptation to the natural and cultural local environment. The term traditional variety is sometimes applied to plant landraces.
Market value chain It describes the sequence of activities from producing raw material and transforming the same into products that can be purchased by final consumers. It can be used as a strategy to create incentives to famers to conserve landraces with traits that are valued by consumers. It is based on interlinked value-adding activities that convert landraces as inputs into marketable outputs which, in turn, add to competitive advantage.
Natural insurance value It is associated with the value of biodiversity in terms of reducing risk to its users and can be more or less substitutable for financial insurance.
Option value The potential benefits of a resource not derived from actual use and expresses the preference for the conservation of such resource against some probability that it can be used at some later date.
Payments for agrobiodiversity conservation services A variant of the term payments for ecosystem services and used in the context of agrobiodiversity.
Plant genetic resources for food and agriculture Any genetic material of plant origin of actual or potential value for food and agriculture. Its diversity embodies the diversity among cultivated species or interspecific diversity and the diversity within species, i.e., the variation of populations belonging to one species, or infraspecific diversity.
Plant species A group of organisms capable of interbreeding freely with each other, but not with members of other species. In taxonomic classification, a subdivision of a genus; a group of closely related individuals descended from the same stock.
Plant variety A subdivision of a plant species. An agricultural variety is a group of similar plants that by structural features and performance can be identified from other varieties within the same species. Synonymous with cultivar.
Underutilized varieties It is also known as neglected varieties and refer to those species with under-exploited potential for contributing to food security, health (nutritional/medicinal), income generation, and environmental services.

Introduction

The economics of agrobiodiversity deal with the causes and implications of changes in biological diversity in agricultural landscapes. Agrobiodiversity is a key asset that sustains food production systems worldwide and provides cultural as well as aesthetic values for society (Jackson *et al.*, 2007a). One type of reasonably well-understood agrobiodiversity is the domesticated or "planned diversity," which is associated with the genetic diversity of cultivars and breeds. This is the basis of local as well as global food supply and nutrition in both subsistence-based societies and technologically advanced societies. However, today, only about 30 plant species out of the global agricultural biodiversity are used to meet 95% of the world's food energy needs (FAO, 1998) and it is estimated that the total global food supply is provided on average by just 100

species (Will, 2008). It is also estimated that more than 1.4 billion people live in farm households that critically depend on local plant genetic resources (Brush, 2000). In fact, smallholder farmers in the developing world continue to maintain the majority of such remaining agrobiodiversity (Wale *et al.*, 2010). The value of this agrobiodiversity extends far beyond the livelihoods of smallholder farmers, however, since it forms the basis of all plant breeding programs for stress tolerance and pest resistance (e.g., Reynolds *et al.*, 2007; FAO, 2010; Varshney *et al.*, 2010), and of much current work to exploit genetic resources via molecular methods (Varshney *et al.*, 2009). The narrow agricultural portfolio of present day's agriculture raises the important question of how effectively major crops alone can contribute toward food security, poverty reduction, and the continued delivery of valuable ecosystem services in agricultural landscapes **(Figure 1)**.

Within such planned diversity, we focus on the diversity of plant genetic resources relevant for food and agriculture. Plant genetic resources for food and agriculture (PGRFA) diversity have two main dimensions: (1) the diversity among cultivated species or interspecific diversity; and (2) the diversity within species, i.e., the variation of populations belonging to one species (intraspecific diversity). Both types depend on the underlying genetic diversity present within and among cultivated species, and the evolutionary processes of human selection within agroecosystems (Bellon, 2009). Since the Neolithic revolution 10,000 years ago, humans have managed the diversity of crops and livestock through selection, experimentation, innovation, and exchange. Modern crops are the result of thousands of years of such evolutionary processes (Brush, 2000). Most of the crop genetic diversity used historically (and easily available to humans) is contained in landraces of both major and minor crops. Landraces are well-adapted to local environmental conditions and have the capacity to tolerate biotic and abiotic stress, resulting in high yield stability partly because they are typically genetically diverse (Villa *et al.*, 2006). They often produce an intermediate yield level under a low-input agricultural system that depends on a set of farmers' practices of seed selection, field management, as well as a local knowledge base (Bellon, 2009).

Local varieties or landraces can be conceptualized as "open, dynamic and decentralized genetic systems since they are the crop populations that farmers manage, and which result from farmers' seed selection practices, the flows of seed among them, and farmers' production and utilization strategies" (Bellon, 2009: p. 54). Clonal propagation of genetic material (such as for agroforestry trees) often requires grafting and budding practices, which somewhat constrains farmer selection and utilization, and thus may be more successful via participatory domestication and community partnerships (Leakey, 2012).

A variety of scientific disciplinary fields have contributed to our understanding of the functions of planned agrobiodiversity (Smale *et al.*, 2001). Examples from the natural sciences include the agricultural sciences, and their pursuit of a better understanding of how to manage crop breeding for enhancing agricultural productivity and food security (e.g., Cooper *et al.*, 2001; Leakey, 2012), and to promote sustainable agriculture more generally through its ecological functions (e.g., Swift *et al.*, 2004; Jarvis *et al.*, 2011). The cultural role of agrobiodiversity is often stressed by ecological and agricultural



Figure 1 Regions of diversity of major cultivated plants. Reproduced from FAO (1998) *The State of the World's Plant Genetic Resources for Food and Agriculture*. Rome, Italy: Food and Agriculture Organization of the UN, and adapted by IDRI, http://web.idrc.ca/en/ev-30549–201–1-DO_TOPIC.html

anthropology (e.g., Brush, 2004; Veleto and Skarbø, 2009) and human ecology (e.g., Dumaresq *et al.*, 2010; Stromberg *et al.*, 2010) and questions about its governance is addressed by political science (e.g., Andersen, 2008; Sperling *et al.*, 2008; Moore). Economics focuses on the causes and consequences of biodiversity loss in general, and agrobiodiversity loss in particular society (Perrings, 2001). Ecological economics as transdisciplinary field is well located to pursue an integrated effort to understand the complexity behind crop agrobiodiversity benefits and its role in supporting human well-being. For other types of agrobiodiversity (e.g., wild plants, soil biota, or arthropods) at the community and landscape level, knowledge is more scattered because inventories and functions are so much more difficult to assess (Jackson *et al.*, 2009).

Despite the importance of agrobiodiversity in contributing to the sustainability of agricultural practices, many plant and animal genetic resources are increasingly being lost from agricultural landscapes worldwide (Perrings *et al.*, 2006; Hajjar *et al.*, 2008). Additionally, as a direct result of growing commercialization and industrialization of farming systems (e.g., via the "Green Revolution") agroecosystems are increasingly characterized by high levels of intensification with low levels of agrobiodiversity (Thrupp, 2000; Jackson *et al.*, 2007b; Tscharntke *et al.*, 2012). It is estimated that humans have made use of about 7000 plant species (FAO, 1998). According to a comprehensive FAO assessment of the state of the world's animal and plant genetic resources, out of 7600 livestock breeds, 20% are classified as being at risk, and 9% are already extinct. Most of these are located in developing countries where the risk of loss is highest, and between 2000 and 2006 alone 62 breeds became extinct (FAO, 2007a, 2009). Illustrative examples are those of Mexico and China where since the 1930s, 90% of all local maize varieties and 80% of all local wheat varieties have been lost, respectively (FAO, 1997). The 2010 FAO follow-up report explicitly noted that the threats to diversity are intensifying, partly due to climate change (Ceccarelli1 *et al.*, 2010; FAO, 2010).

Off-farm, *ex situ* gene bank facilities store growing collections of germplasm of local crop varieties and wild relatives (FAO, 2010). These *ex situ* facilities provide a relatively low-cost method of safeguarding genetic resource material in the state it was collected, avoiding loss or degeneration (Hawkes, 1983; Plucknett *et al.*, 1987; Brush, 2004), and are considered key to the protection of genetic diversity in the face of climate change (Ceccarelli1 *et al.*, 2010; FAO, 2010). Accessions in gene banks are used mainly for crop breeding programs by scientists and farmers (Smale and Day-Rubenstein, 2002; Hodgkin *et al.*, 2003). The total number of accessions conserved *ex situ* in the world is 7.4 million, from which it is estimated that less than 30% are distinct, i.e., between 1.9 and 2.2 million (FAO, 2010). But even if the world's major crops are found in *ex situ* facilities, many minor crops and crop wild relatives that are a key asset for the livelihoods of the rural poor worldwide are underrepresented in such gene-banks (Padulosi *et al.*, 2002). Moreover, agrobiodiversity conservation is inherently a dynamic process requiring ecological relationships such as gene flow between different populations and species, coevolutionary adaptation and selection to predation and disease, and systems of agricultural knowledge and practice associated with genetic diversity, which cannot be achieved *ex situ* (Altieri and Merrick, 1987; Wood and Lenne, 1997; Brush, 2004).

On-farm or *in situ* agrobiodiversity conservation is based on the maintenance of crop diversity by farmers themselves. It involves management of local crop populations including the maintenance of their evolutionary processes of natural and human selection based on traditional knowledge, seed saving, and exchange (Bellon *et al.*, 2009). *In situ* and *ex situ* agrobiodiversity conservation can be seen as complementary strategies. The former maintains a stock of genetic diversity in continuous evolution and provides commercial and crop breeder programs with valuable genetic material for innovation (Brush, 2000).

Based on knowledge of the role of agrobiodiversity, stemming from research conducted in the last couple of decades, increasing emphasis is being put on biodiversity-based paradigms for sustainable agriculture (Collins and Qualset, 1999; MEA, 2005; Letourneau *et al.*, 2011), both as potential solutions to the many problems associated with intensive agricultural practices and to generate greater resilience to food systems (Jackson *et al.*, 2007a,b). More specifically, agrobiodiversity is viewed as a means of survival for the world's rural poor (Smale, 2006) and as a mechanism for buffering against output losses due to pests and diseases in traditional as well as commercialized agricultural systems (Heisey *et al.*, 1997). It also serves as a key input into locally sustainable cropping (Bellon *et al.*, 1997) indigenous technology systems. More globally it serves as a repository of valuable traits that can help satisfy the evolving tastes and preferences of consumers (Evenson *et al.*, 1998) and allow future genetic improvements on which the global supply of food products depends (Koo *et al.*, 2004).

The challenge is threefold: (1) to better understand the drivers of agrobiodiversity change; (2) to enhance the knowledge of the intertwined ecological and social functions of agrobiodiversity by determining its contribution to ecosystem services and their economic value; and (3) to evaluate alternative options for the sustainable use and conservation of agrobiodiversity on-farm, where the latter implies "the combined set of actions (and policies) by which a sample, or the whole, of a plant/animal population is subjected to processes of genetic and/or environmental manipulation with the aim of sustaining, utilizing, restoring, enhancing and understanding (characterizing) the quality and/or quantity of the genetic resources and its products" (Drucker and Smale, 2007: p. 623).

Economics is well suited to studying the drivers of agrobiodiversity change, to linking agrobiodiversity with ecosystem service values, and to devising instruments for its sustainable conservation and utilization in agricultural landscapes. In what follows, the economics of agrobiodiversity is reviewed where the latter refers to planned (domesticated) plant and animal genetic diversity and its conservation refers to the *in situ* or on-farm approach. The next section describes the problem of agrobiodiversity loss, its economic drivers, and the role of Agrobiodiversity in supporting economic development. The section The Value of Agrobiodiversity addresses the different types of economic value associated with agrobiodiversity and ways to estimate them in order to

demonstrate its social worth as natural capital. The section Economic Instruments for Agrobiodiversity Conservation focuses on the economic instruments that can be used to capture and share the social value by farmers generated through their on-farm conservation effort.

The Economic Problem of Agrobiodiversity Loss

As with the management of biodiversity in general, changes in agrobiodiversity is seen in economics as an investment/disinvestment decision (Swanson, 1998) made in the context of evolving sets of preferences, "value systems," moral strictures, information, technological possibilities, and social, cultural, and institutional conditions (Pascual and Perrings, 2007). From an economic perspective, agrobiodiversity loss becomes a problem wherever it is socially inefficient (given social distributional priorities).

Many of the services provided by on-farm agrobiodiversity are not traded in markets, leading them to have no price in markets or, at the very best, to be undervalued. Furthermore, since, agrobiodiversity not only provides private on-farm benefits, but also quasi-public benefits such as improved pollination or control of pests or diseases that reduce production risks on neighboring farms, a market failure problem arises. Economic theory predicts that the inability by farmers to capture the value of such benefits leads to the underprovision of such services, in other words to a socially suboptimal level of agrobiodiversity conservation at national, regional, and global levels (Drucker and Smale, 2007; Baumgartner, 2007; Pascual and Perrings, 2007). Policy interventions are therefore necessary to support the sustainable use and management of agrobiodiversity in order to fulfill the needs of society at large.

The field of agrobiodiversity economics recognizes that besides supply-side factors that may hamper access to resources required to maintain socially desired agrobiodiversity levels (e.g., access to seeds by farmers), the bulk of the problem behind agrobiodiversity loss is demand-based. Farmers' individual decisions regarding the desired level of planned agrobiodiversity in their farms usually depend on conditions in the relevant agricultural markets (Smale *et al.*, 2001), and these often do not recognize the full suite of values of agrobiodiversity.

Although agrobiodiversity management decisions are taken at the farm-level, with the opportunity costs of conservation privately born by individual farmers, a significant part of the benefits from agrobiodiversity conservation accrue at the landscape (community), regional, and even global levels (O'Farrell and Anderson, 2010). Since the social benefits of higher levels of crop genetic diversity being conserved are not transmitted to those farmers who invest in their conservation, the latter have little or no private incentive to conserve agrobiodiversity (Perrings, 2001). The most profitable decision is to grow only a few crop varieties frequently, and not to invest in conservation of the varieties that are less "favored" by the market. In this sense, farmers who conserve and utilize on-farm or *in situ* crop genetic diversity are essentially conserving an impure public good that has significant global public spillover benefits and thus they can be seen as net-subsidizers of modern agriculture and food consumers worldwide (Pascual and Perrings, 2007).

The case of underutilized crop species is an example of such market failure in agrobiodiversity. Following Gruère *et al.* (2009: p. 62), the notion of "underutilized" species can be applied to those landraces that simultaneously share the following three characteristics (1) they are locally abundant but globally rare, (2) local users have practical knowledge of them but scientific information about the species is scant, and (3) current use of the plant is limited relative to its economic potential. Minor millets (finger millet, kodo, barnyard, and little millet) in India are one example of underutilized species. Many other underutilized plant species are collected rather than cultivated (Padulosi *et al.*, 2002). The market failure associated with underutilized species stem from the fact that their real and potential economic value that surpasses their current (observed) value and this induces the socially suboptimal provision of these species.

Besides the market failure associated with agrobiodiversity, there are important intervention or policy failures at the local, regional, and global scales that further distort farmers' decision toward agrobiodiversity conservation. For instance, the bulk of subsidized inputs associated with intensive agriculture, free extension services, or market price support related to high yielding or improved plant genetic resources are sources of such policy failure (Drucker and Rodriguez, 2009). Additionally, as Pascual and Perrings (2007) note, other important institutional failures arise from the decisions of powerful intergovernmental organizations (World Bank, International Monetary Fund, and United Nations Development Programme) and international agreements (the General Agreement on Tariffs and Trade, the Sanitary and Phytosanitary Agreement, and the International Plant protection Convention). In some cases, they affect agrobiodiversity by limiting the choice of management strategy or technology used by farmers. In others, they work by encouraging the diffusion of new technologies or by dispersing new crop varieties, biocontrol agents, pests, and pathogens. As in the case of direct subsidies, these indirect influences on farmers' decisions change the private returns on farm investments, often in ways that discourage agrobiodiversity conservation. Hence, from an economic viewpoint, the problem of agrobiodiversity loss can be interpreted as a typical market failure. It stems from the mixed public-good nature of agrobiodiversity conservation and the policy or institutional failures that make a biodiversity-based agriculture uncompetitive (e.g., through agricultural production subsidies, tax breaks, price controls, or incomplete property rights) and lack of appropriate governance structures that support its conservation (Pascual and Perrings, 2007).

Figure 2 describes a heuristic model adapted from Drucker and Rodriguez (2009) and Narloch *et al.* (2011a), and serves to illustrate the market failure problem associated with the socially suboptimal level of agrobiodiversity conservation and the displacement of local landraces or underutilized species by commercially improved, high yielding varieties from modern breeding programs (e.g., Brush and Perales, 2007).

Given local agroecosystem characteristics and local economic context, including prices of different crop varieties in the market or the implicit value attached to them by farmers, a

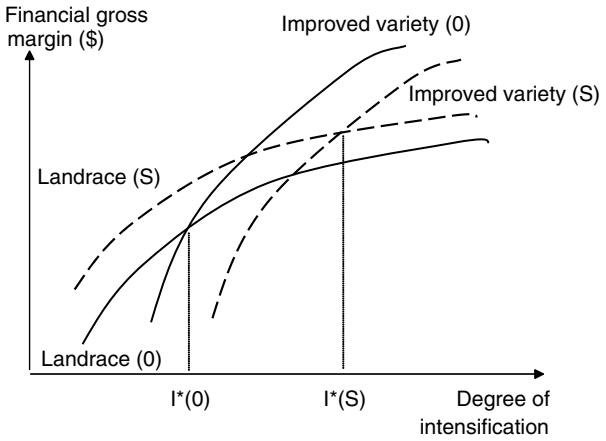


Figure 2 Replacement of a landrace by improved crop variety. Adapted from Drucker AG and Rodriguez LC (2009) Development, intensification and the conservation and sustainable use of farm animal genetic resources. In: Kontoleon A, Pasqual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 7, pp. 92–110. Abingdon, UK: Routledge, with permission from T&F.

local crop variety (or landrace) might be expected to outperform a higher yielding or improved variety at low levels of input intensification of the agricultural system, denoted by I^*. Here the term intensity is used in a general sense and includes also, *inter alia*, factors related to access to markets and extension services. The curves represent the profitability functions observed by farmers under full information. After $I^*(0)$ is reached, farmers face increasing incentives to replace the locally adapted landrace with the improved crop variety. The magnitude of such incentives is determined by the vertical distance between the two curves and represents the opportunity cost associated with the conservation and utilization of the local landrace.

As Drucker and Rodriguez (2009) argue, there are a number of reasons to suggest that the current replacement point, $I^*(0)$, is located to the left of the socially optimal replacement point, or $I^*(S)$, implying a suboptimal conservation of agrobiodiversity on-farm. The difference between $I^*(0)$ and $I^*(S)$ reflects a market failure. The distance increases in two cases. The first is where the perceived financial performance of the improved crop variety is artificially supported, for instance based on policies that support the free availability of improved seeds, capital subsidies for inputs such as fertilizer or pesticides, free or subsidized support services, or subsidized market prices associated with the improved variety instead of with the landrace. The second is where the possible public benefits of conservation, such as those associated with the provision of multiple ecosystem services and the potential nonuse values associated with local agrobiodiversity, including maintaining socio-cultural traditions, local identities, and traditional knowledge values, are ignored.

Figure 2 assumes that to the left of $I^*(0)$, farmers have financial incentives not to replace local crop varieties, and thus to conserve agrobiodiversity since this is what provides a higher economic return. Beyond $I^*(S)$, displacement of landraces by improved varieties would be economically justified. The dilemma occurs between these two points, that is, when the displacement takes place between $I^*(0)$ and $I^*(S)$ as it is associated with a suboptimal loss of agrobiodiversity from a social perspective, while this might be privately optimal from the farmer's point of view. In this case, the additional loss of public and nonmarket values due to the erosion of agrobiodiversity outweighs the benefits of introducing the higher yielding variety.

As Drucker and Rodriguez (2009) and Narloch *et al.* (2011a) contend, from this heuristic model the solution to the social dilemma seems straightforward. Society should first demonstrate the public values of agrobiodiversity and design economic instruments so that farmers can capture the true economic value of on-farm agrobiodiversity. All distorting policies that bias against landraces artificially ought to be dismantled. These include, for example, subsidies to improved varieties that shift the curve for improved varieties downward and to the right (dotted line). In addition, by putting in place mechanisms to permit the "capture" of the total economic values associated with the use of landraces, the profitability curve for local varieties would shift upward to the left (dotted line). The next section addresses the value of agrobiodiversity and how this can be demonstrated. Policy mechanisms to capture and share the values are presented in the subsequent section.

The Value of Agrobiodiversity

Valuing biodiversity has been the focus of much research and debate in economics (e.g., Pearce and Moran, 1994; Nunes and van den Bergh, 2001; Pascual *et al.*, 2010). There is a growing theoretical and empirical literature, largely based on case studies, which forms one of the most fertile grounds of environmental economics. The economic value of biodiversity is an anthropocentric dimension of value and as such it can be used to complement, but not to substitute for, other legitimate ethical or scientific reasoning and arguments relating to biodiversity conservation (Turner and Daily, 2008). In essence, the economic value of biodiversity is one way to capture the intensity of people's preferences for changes in the quantity or quality of biodiversity. There is a growing realization among conservation biologists that demonstrating the value of biodiversity influences policy makers to increase conservation (Pascual *et al.*, 2010; ten Brink, 2011).

During the past decade, a small but growing literature on valuing crop and livestock agrobiodiversity emerged (e.g., Smale, 2006). The ongoing efforts to demonstrate the economic value of agrobiodiversity is generally seen as a first step in justifying its conservation through the design and implementation of policy instruments that can capture and share the benefits of conservation (Jackson *et al.*, 2007b; Pascual and Perrings, 2007). Demonstration refers to the identification and measurement of agrobiodiversity. Capture refers to the appropriation of biodiversity values through incentives for its conservation by those who ultimately bear the costs of conservation.

Agrobiodiversity is largely not priced by the market. Its value reflects a socially evolving context, and there are differences in the way that social groups calculate its worth (Pascual and Perrings, 2007; Pascual *et al.*, 2010). Price information on

agricultural genetic resources is sparse, mainly because the attributes or traits of such resources are not easily identifiable even by those managing them, such as farmers. Compared to wild biodiversity, proportionately more of the economic value of (domesticated) agrobiodiversity is related to use rather than nonuse values, although the latter can be important in certain contexts associated with farmers' cultural identity and the maintenance of traditional knowledge (Cicia *et al.*, 2003; Smale, 2006; Smale and Drucker, 2009).

As pointed out by the Millennium Ecosystem Assessment (MEA) "…where agriculture already dominates landscapes, the maintenance of biodiversity within these areas is an important component of total biodiversity conservation efforts, and, if managed appropriately, can also contribute to agricultural productivity and sustainability through the ecosystem services that biodiversity provides (such as through pest control, pollination, soil fertility, protection of water courses against soil erosion, and the removal of excessive nutrients)" (MEA, 2005: p. 13). These regulating and supporting ecosystem services, largely associated with resilience at the landscape level, are quite distinct from those typically associated with forest ecosystems and the carbon sequestration services they provide. An additional distinction is that, unlike most wild biodiversity, agrobiodiversity has mixed-good properties. As discussed by Smale (2006), it has an impure public good characteristic as it shares properties of a private good (rival in use and excludability possibility) as well as public good (low rivalry and low excludability). While seed is clearly rival and the cost of exclusion is low, the genetic information embodied in the seed is non rival and the cost of controlling its use can be costly. Similarly, seeds as planting material can be considered a private good consumed as a production input, with no two farmers planting the same unit of seed, but the genetic resources embodied in the diversity of seeds is a public good for those two farmers. This public good aspect also extends to the functionality of the underlying seed system that allows evolutionary processes to be maintained, as well as the underlying traditional knowledge related to the management of use of the genetic resource in question. Furthermore, biodiversity has an option value for future generations in terms of beneficial traits for breeding new varieties adapted to changing environmental conditions.

Private versus Public Value of Agrobiodiversity

Underutilized or neglected species of minor or major crops provide important plant genetic diversity worldwide and they are associated with all three types of values. Gruère *et al.* (2009) provide a conceptual framework to guide empirical analysis of changes in use of underutilized species and policy options to conserve them. The value of underutilized plant species involves two components: a private and public value component. The public component includes the value of improved nutrition to both current and future generations, the mitigation of private production risks, and the nonuse values associated with tradition and culture (Gruère *et al.*, 2009). They propose a classification of underutilized plant species based on their private and public value. By definition, the observed value of any underutilized plant species is inferior to its potential value **(Figure 3)**.

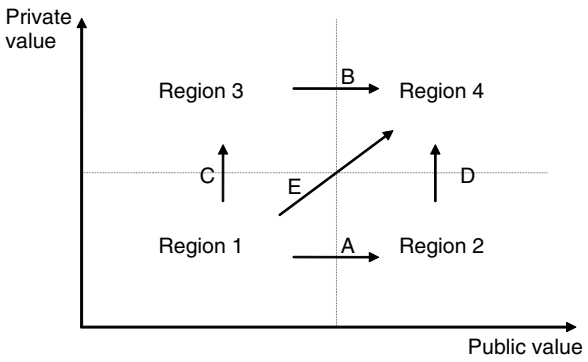


Figure 3 Characterization of underutilized plant species according to private and public values. Reproduced from Gruère G, Giuliani A, and Smale M (2009) Marketing underutilized plant species for the poor: A conceptual framework. In: Kontoleon A, Pascual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 5, pp. 62–81. Abingdon, UK: Routledge, with permission from T&F.

A given underutilized plant species is represented by an outward vector linking its observed value to its potential (current or expressed) value, with the origin of the vector being the observed value. The current or observed value of the species can be located in Region 1, 2, or 3 but not in region 4 as by definition species that have both high private and public values cannot be identified as being underutilized (Padulosi *et al.*, 2002). Gruère *et al.* (2009) identify five types of directional vectors, each representing a different type of underutilized species. The horizontal vectors A and B represent species that are neglected by policy-makers despite their known contribution as public goods. Vectors C and D both have a vertical direction representing species for which their current or observed private value is inferior to their private potential value. In this case, public intervention could be geared to help markets function better. Finally Vector E is associated with a species that currently has a lower private and public value than what could be potentially realized. This categorization allows us to identify alternative types of intervention related to the need for enhancement of private good and/or the capture of public good values to support the maintenance of agrobiodiversity.

While many underutilized plant species are locally valued as revealed by their use based on traditional knowledge, the potential public value often arises through new scientific evidence related to their intrinsic properties, or due to changing consumer preferences in turn linked to evolving cultural trends or fashion. For example quinoa (*Chenopodium quinoa*), typically grown in the Andean altiplano of Bolivia and Peru, has become fashionable in urban areas of developing and developed countries (Gruère *et al.*, 2009).

The Total Economic Value of Agrobiodiversity

The value of agrobiodiversity is generally argued to comprise: (1) the private benefits to farmers via the consumption and production values they derive from producing crops; (2) private, as well as locally and regionally public benefits to farmers, and to consumers, when the choices make farming more resilient to biotic and abiotic stresses; and (3) global

public benefits to future farmers, plant breeders, and consumers, when the choices they make protect against genetic erosion (Lipper and Cooper, 2009, p. 30). These types of value are associated with a direct (private) use value, a local public indirect use value, and global public option/bequest values, respectively. These are described in the next section.

Direct Use Value

Current direct use values are most obvious when referring to plant genetic resources and refer to the harvest and use of crop varieties in noncommercial or commercial/industrial processes. They derive from the utility gained by an individual from the consumption of a good or service or from the consumption by others of a good or service associated with agrobiodiversity (Smale, 2006). They reflect the production and consumption benefits of agrobiodiversity. For example, the financial rate of return to investment in plant breeding programs in the development of world agriculture has been demonstrated to be significant, and to generate economic returns that far outweigh the costs of investment (Alston *et al.*, 2000; Evenson and Gollin, 2003). The use of genetic resources for plant breeding that lowers food prices in less developed countries is especially relevant, as these consumers generally spend a greater share of their income on food than consumers in more developed countries (Drucker and Smale, 2007). Studies to date attest to the significant net marginal benefits of exploiting individual genetic accessions in commercial agricultural use versus conserving it in *ex situ* facilities such as large national and international gene banks (Gollin *et al.*, 2000, Koo *et al.*, 2004, Drucker and Smale, 2007).

For on-farm agrobiodiversity, Bellon (2009) argues that most studies of the benefits associated with maintaining crop diversity on-farm link them to the provision of local private goods and services to farmers (see e.g., Smale, 2006), to the provision of desirable varieties to other farmers in a particular community or region (e.g., Badstue *et al.*, 2006; Stromberg *et al.*, 2010), and to the management of tolerance to stress, pests and diseases (a quasipublic good) (e.g., Jarvis *et al.*, 2011).

The economic valuation of agrobiodiversity includes both studies of farmers' revealed preferences and surveys to elicit preference over a range of crops, farmers' socio-economic and agroecological environments. Case studies support the notion that farmers value various dimensions of agrobiodiversity. From these studies, the data indicate that often farmers in less productive and more remote regions value agrobiodiversity the most (Smale, 2006).

Understanding farmers' choices regarding crop diversity (Van Dusen and Taylor, 2005; Birol *et al.*, 2006, 2008) and varietal traits affecting productivity, environmental adaptability, and yield stability (Gemessa and Zander, 2010; Pant *et al.*, 2010) is key to developing incentives for agrobiodiversity conservation. Stated preference methods to estimate farmers' willingness to pay for different agrobiodiversity attributes show that farmers' preferences for variety attributes are shaped by economic (resource constraints, markets, and risk) as well as by noneconomic factors, including cultural norms and attitudes. In addition, farmers' demand for varieties also depends on consumers' preferences for valuable attributes embodied in landraces such as taste, aroma, etc. Through attribute preference analysis based on hedonic pricing models, consumers' willingness to pay for different traits of landraces have been estimated for crops such as eggplants in India (Krishna *et al.*, 2010) and rice in Nepal (Pant *et al.*, 2010).

Other studies have focused mainly on the productivity effects of agrobiodiversity to derive their economic value. When on-farm biodiversity supports the productivity of crops by enhancing yields or their stability, or by substituting for the use of purchased inputs such as pesticides, agrobiodiversity has an instrumental or "use-value" for farmers. In addition, a direct consumptive value refers to the benefits that consumers, including farmers, obtain from being able to access a greater variety of foods to support food security through a more diverse diet, taste preferences, cooking characteristics, and storage (FAO, 2009). Based on the MEA framework, direct use values are associated with provisioning services of agrobiodiversity. More specifically, production benefits refer to supporting mean productivity as well as stability of yields especially in areas prone to significant environmental (e.g., pest, disease, and competition), climatic and/or economic risks.

Besides theoretical models linking agrobiodiversity and productivity (e.g., Omer *et al.*, 2010), other economic approaches linking genetic diversity to crop productivity have econometrically estimated mean-variance production functions to test the relationship of crop biodiversity to mean yield productivity and variability. As Drucker and Smale (2007) point out findings are rather location-, time-period-, and cropping-system-specific. In their groundbreaking studies Smale *et al.* (1998) on wheat production in the Punjab of Pakistan and Widawsky and Rozzelle (1998) on rice production in China used the Just and Pope-type stochastic production function approach. While Smale *et al.* (1998) found that in rainfed districts, crop genetic diversity was linked positively to mean yield levels and negatively to yield variance, Widawsky and Rozzelle (1998) found that diversity was instead negatively related to both mean and variance of yields. In an extended approach to account for risk aversion to crop failure (downside risk exposure) in rainfed durum wheat farming in Sicily (Italy), Di Falco and Chavas (2006) have shown that under low pesticide use, crop genetic diversity can increase farm productivity and reduce yield variance. Additionally, when longer-term effects can be factored in, spatial genetic diversity appears to be linked with greater productivity in more intensive cereal-based agroecosystems (Omer *et al.*, 2007; Di Falco and Chavas, 2008).

Indirect Use Value

The indirect use value of agrobiodiversity refers to the environmental services rendered by agrobiodiversity in providing stability of yields, thus reducing the risk of crop failure. It has been noted that risk-averse farmers use crop diversity in order to hedge their income risk (Birol *et al.*, 2006, Di Falco and Perrings, 2003; Di Falco and Chavas, 2009). This is akin to considering investment in agrobiodiversity as a form of "self insurance" by risk-averse farmers who act *ex ante* in order to reduce a potential income loss, such as reliance on landraces by the most vulnerable farmers in a drought-prone region (Cavatassi *et al.*, 2011). In addition to the private decisions by individual risk-averse farmers faced with risk of crop failure, the natural insurance mechanism of

agrobiodiversity is also associated with a broader notion of social-ecological resilience delivering important public benefits at the local and regional level (Pascual and Perrings, 2007). The public insurance value of agrobiodiversity stems from the notion that a more biodiverse agricultural system allows the system to absorb higher disturbance levels, be it climatic, environmental, or economic, and still remain within the same state while being able to self-organize and have the capacity for learning and adaptation (Carpenter *et al.*, 2001).

In this vein, the public value of agrobiodiversity is also associated with the capacity of the agricultural system to adapt to economic or environmental external changes. For instance, multicrop and tree biodiversity in agroecosystems can enhance the ability of farmers to remain economically viable by diversifying revenues affected by volatile international market prices for inputs, crops, or tree products (Lin, 2011). In other words, in intensive agricultural systems with low levels of agrobiodiversity, and high dependence on the use of nonrenewable inputs, the cost of eroding agrobiodiversity is that the system gets locked into a narrow range of highly productive and stable (under low perturbations) agricultural technologies. But this reduces the capacity of the system to absorb greater environmental or economic shocks, such as sudden and unexpected commodity price changes or climatic extreme events, especially where other supporting mechanisms (government subsidies and extension services) may not be sufficient. By eliminating options toward productive diversification, a reduction in agrobiodiversity may also lock farmers into obsolete agricultural technologies (Perrings, 1998).

While there are some economic studies that explore the private insurance value of agrobiodiversity (Cavatassi *et al.*, 2011), the public insurance value of crop diversity in regulating ecosystem services at the landscape level (e.g., pollination, water retention) is not yet fully supported empirically. However, a view is emerging that links the broader notion of agrobiodiversity at the landscape level (beyond crop diversity) with the regulating ecosystem services, and so identifies positive spill-over effects of agrobiodiversity on local on-farm productivity and stability, especially under global perturbations (Chapin *et al.*, 2000; Folke *et al.*, 2004; Tscharntke *et al.*, 2005; Jackson *et al.*, 2007a,b). Often this is associated with species in neighboring ecosystems that provide pollination and biological control services against pests and invasive species that, in turn, can imply significant cost savings to farmers (Ricketts *et al.*, 2004; Perrings, 2005; Tscharntke *et al.*, 2012). In addition, genetic diversity may contribute to an array of ecosystem services by facilitating evolutionary processes *in situ* that select for better-adapted crop varieties (Faith *et al.*, 2010).

The private indirect use value for farmers is based on the private (natural) insurance value that agrobiodiversity offers. Reliance on agrobiodiversity to manage risk can be affected by agricultural policies such as subsidized crop yield insurance or direct financial assistance (Di Falco and Perrings, 2005) as well as the existence of financial insurance institutions. Market insurance products can serve as a substitute for self-insurance through agrobiodiversity investments by risk-averse farmers. The insurance issue is addressed in a model developed by Quaas and Baumgärtner (2008) and Baumgärtner and Quaas (2009). In this model it is shown that the price of financial insurance has an impact on the level of on-farm agrobiodiversity managed for risk management purposes. The model also shows that as financial insurance becomes cheaper, it can optimally drive out agrobiodiversity. In addition, in the case of agrobiodiversity having a quasipublic good characteristic, i.e., when it has positive spill-over effects in turn leading to market failure, the trade-off between improved access to financial insurance and lower agrobiodiversity conservation has ambiguous welfare effects and these mainly depend on agroecosystem properties (Quaas and Baumgärtner, 2008). They show that when the private insurance value of agrobiodiversity is low compared to its public insurance value, then improved access to financial insurance can have a negative impact on total welfare at the local or regional scale. The public insurance value stems from the local and regional stability and wider societal benefits that are not directly associated with harvest or consumption at the plot or farm level.

Agrobiodiversity through its insurance mechanism may not generally be enough to justify conservation by an individual farmer when there is easy access to improved artificial capital inputs, for example, fertilizers, improved seeds, etc. However where financial insurance markets are thin or nonexistent, or access to intensive agricultural technologies is restricted, there is some evidence that agrobiodiversity can mitigate production and consumption vulnerability in marginal environments. This evidence is still sparse, even if it is well known that farmers manage risk through crop or varietal choices in their *ex ante* farm production decisions (Smale *et al.*, 1998; Di Falco and Chavas, 2006). For example, Di Falco and Chavas (2008, 2009) have showed that in areas prone to production failure due to lack of rainfall, such as in the South of Italy and in the Tigray region of Ethiopia, respectively, crop agrobiodiversity is used to buffer negative productivity effects of declining rainfall levels.

Global Option Value

The option value of agrobiodiversity is the value of the option to use it under future uncertain technological, market, or environmental conditions (Smale, 2006). (While agrobiodiversity may also have existence and bequest values, these are likely to be relatively unimportant both compared to the other types of value generated by agrobiodiversity, as well as to that of wild biodiversity in general. We therefore do not further consider such nonuse values in this article.) This type of value is not necessarily factored in by individual farmers in their decentralized decisions and instead may be associated with the maintenance of crop evolutionary processes (Brush, 2000, chapter 1). As Bellon (2009) argues, by conserving landraces we keep the option for the generation of new and improved crop varieties. Agrobiodiversity is the basis for scientific plant breeding and farmers to adapt crops to heterogeneous and changing environments in order to increase agricultural productivity as well as stability of production by providing, for example, environmental stress and resistance to pests and diseases. There is ample evidence that continuing releases of improved varieties, based on conserved genetic resources, have generated economic returns that far outweigh the costs of investment (Qualset and Shands, 2005). Further, the global option value of agrobiodiversity is considered to be very

significant especially in the face of climate change as it can provide natural insurance for an unpredictable future (Ceccarelli1 *et al.*, 2010; FAO, 2010; Pascual *et al.*, 2011). But while some case studies exist at the local level that link farmers' *ex ante* agrobiodiversity choices to climatic risk exposure (e.g., Di Falco and Chavas, 2009; Cavatassi *et al.*, 2011), the global option value of both *ex situ* and *in situ* agrobiodiversity in terms of uncertainty to global economic and climatic changes represents a significant research gap.

Economic Instruments for Agrobiodiversity Conservation

A significant policy challenge is that most of the globally valuable agro-biodiverse landscapes are located in poor and often marginal regions of the world exposed to rapid social and economic change (Drucker and Smale, 2007). Further, market developments have major impacts on agrobiodiversity by affecting farmers' choice of crops and varieties (Lipper *et al.*, 2009) and by reducing the probability that farmers will cultivate traditional varieties that add to both intra (or varietal)- and intercrop diversity (van Dusen and Taylor, 2005; Smale, 2006). Hence in the context of rapid economic change in developing countries, incentive mechanisms for *in situ* agrobiodiversity conservation are being devised. Noneconomic policy instruments attempt to influence farmers' decisions toward conservation through moral suasion, regulation, and planning (Pascual and Perrings, 2007). The supply of, and demand for, genetic resources by farmers depends both on relative prices and on regulations governing plant breeding and the seed sector (Lipper and Cooper, 2009). While most regulations related to land use in agricultural landscapes affect agrobiodiversity indirectly, various national and international regulatory frameworks exist to support agrobiodiversity, such as the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA).

The ITPGRFA entered into force in 2004 as a legally binding international agreement that promotes the conservation and sustainable use of PGRFA and the fair and equitable sharing of the benefits arising out of their use, in line with the Convention on Biological Diversity (FAO, 2009). The ITPGRFA has established a multilateral system of access and benefit sharing to facilitate access to plant genetic resources of the most important crops for food security. It involves 126 countries plus the European Union. The Convention on Biological Diversity (CBD), which views *in situ* conservation as the preferred approach, has called for the development of positive incentives (COP 10 Decision X/2) for conservation as part of its Strategic Plan for 2011–2020 (see http://www.cbd.int/decision/cop/?id=12268). In addition, a handful of countries worldwide are experimenting with new ways of protecting and rewarding traditional knowledge through the realization of Farmers' Rights by national governments, as explicitly called for Article 9 of the ITPGRFA (FAO, 2009). It is therefore apparent that innovative policy instruments are being seen as necessary to facilitate and enhance incentives for farmers to utilize and conserve agrobiodiversity.

As is the case for biodiversity in general, a partial solution may lie within the market system itself without explicit need for government intervention, i.e., market formation for agrobiodiversity conservation (Pascual and Perrings, 2007; Perrings *et al.*, 2009). The idea is that since conserving on-farm agrobiodiversity can yield goods and services valued beyond the farm, in certain circumstances, these could be traded in the market and doing so could generate income for farmers, making on-farm conservation financially viable for them. The challenge is thus to identify the specific components of agrobiodiversity that can be shown to provide benefits for different consumers at different scales. There are two main types of market instruments for agrobiodiversity: (1) those that stimulate market value chains for landrace-derived commodities, including through the use of labeling and certification schemes, and (2) voluntary direct payments to farmers. Both can be seen as complementary. While emergent economic theories and applications are being developed about such instruments, to date they remain to be mainstreamed into policy making.

Market Value Chain Development

Developing market value chains for crop genetic resources (CGR) offers an opportunity to link farmers' incomes and the diversity of crops and varieties that they manage. The idea is that adding value to the products derived from traditional varieties means that consumer demand for the crop can increase its market price and the income received by farmers. Promoting such value chains for agrobiodiversity is thus seen as one sustainable way to maintain the cultivation of targeted landraces and to inform which types of landraces could succeed along the value chain (Will, 2008; Gruère *et al.*, 2009; Wale *et al.*, 2010).

A value chain approach regarding plant genetic resources targeted for conservation can be promoted through eco-labeling, certification, or origin schemes, which highlight the specific attributes of agro-biodiversity related commodities (Hermann and Bernet, 2009). While there is ample experience with eco-labels in developed countries (Basu *et al.*, 2003) to support products derived from traditional plant genetic resources, the potential marketability of traditional crop varieties through labeling and certification in developing countries is also gaining attention (Kruijssen *et al.*, 2009; Gautam and Paant, 2010). For example, Hellin *et al.* (2010) points to a number of successful experiences in developing countries, such as with native potatoes and quinoa in the Andes, maize landraces in Mexico, capers in Syria, and minor millets in India. However, there are few studies that estimate both the implicit consumer willingness to pay for landraces and farmer willingness to accept compensation, especially in contexts where higher yielding varieties exist. An exception is Krishna *et al.* (2010), who have demonstrated, for the case of eggplant landraces in India, that the implicit price premium urban consumers are willing to pay for landraces instead of modern (hybrid) eggplant varieties can compensate for the lower productivity of the former.

However, although it has been shown that market value chain development can facilitate the conservation through sustainable use of "neglected and underutilized" or traditional crop species/varieties by enhancing the values of such CGR

(amongst others see: Hoeschle-Zeledon *et al.*, 2009; Padulosi *et al.*, 2009; Rojas *et al.*, 2009, respectively, for Italian hulled wheat, Indian minor millets, and Andean grain examples), the potential to use such an approach as part of a strategic national agrobiodiversity conservation program remains to be established. Indeed, Narloch *et al.* (2011a) noted that "relying solely on market development might be a dangerous strategy for the conservation of a diverse genetic resource pool, especially as market conditions can change rapidly and generally consumers and agribusiness tend to favor a narrow suite of crop species/varieties."

In fact, in order for such market value chain development interventions to be able to effectively contribute to such a strategic national program, a number of challenges need to be considered. These include:

1. The market potential of priority crop genetic resources for conservation: Not every crop species/variety within a priority conservation portfolio (i.e., the strategically selected set of crop species/varieties considered to be priorities for conservation given their contribution to overall diversity and their current threat levels and conservation costs) will necessarily have improved marketing potential. In fact, in the absence of a strategic approach, market value chain development interventions would instead tend to focus on high market potential traditional crop species that may either be unthreatened or only locally threatened.
2. Initial investment costs and long-term success rates: Not every niche product development project is successful and initial investment requirements may be high. Generating appropriate product volumes for market may also be far greater than those required to achieve modest conservation goals.
3. Displacement/leakage: Market development may be so successful that it displaces other priority crop species/varieties. Examples of such displacement include minor millets in India (Padulosi *et al.*, 2009), Andean grains in Latin America (Rojas *et al.*, 2009), potatoes in the Andes, and capers in Syria (Hellin *et al.*, 2010). In such cases, farmer-selection to conform to characteristics demanded by the market has reduced the suite of other characteristics associated with local varieties.

Given this, it is worth exploring the potential for recently developed payment for agrobiodiversity conservation services (PACS) incentive schemes to overcome some of these challenges. Such schemes, which are discussed further below, have been shown to be a potentially effective instrument for promoting the cost-effective maintenance of threatened crop varieties while accounting for social equity considerations.

Voluntary Direct Payments

The creation of markets for biodiversity can also involve direct payments (Ferraro, 2001; Ferraro and Simpson, 2002). These include payment for ecosystem services (PES), which build on compensation flows from the beneficiaries of an ecosystem service to the providers of that service (Engel *et al.*, 2008; Muradian *et al.*, 2010). PES has been hailed as a promising solution for conservation dilemmas in many different contexts, but mainly in the contexts of forest conservation in developing countries (Wunder *et al.*, 2008) following the so-called '*Provider Gets Principle*' (Hodge, 2000).

In developed countries, such as in EU, there are payments to farmers to change agricultural practices toward more biodiversity- and ecosystem-friendly practices (Kleijn and Sutherland, 2003). These resemble PES, though the buyer in this case is not a direct beneficiary but a government institution. There are few examples of such types of PES in agricultural landscapes in developing countries. These include paying farmers in silvopastoral systems to manage diverse ecosystems services, to introduce natural vegetation contour strips in farmers' plots, and to integrate short-term improved fallow systems and shade-grown coffee cultivation (FAO, 2007b; Lipper *et al.*, 2009).

To date, however, only a couple of PES-like schemes have been applied to conserve plant genetic diversity worldwide. While not presented as a PES scheme, a Global Environmental Facility project in Ethiopia involved paying farmers for conserving traditional sorghum varieties based on an incremental cost approach related to the yield gap between traditional and improved varieties (Wale, 2008). Recently a novel approach, known as PPACS, which is a subcategory of agriculture-related PES, has focused specifically on socially valuable and threatened local plant genetic resources (Narloch *et al.*, 2011a, b). Such schemes may be viewed as an economic instrument to tackle the market, policy, and global appropriation failures associated with the public good characteristics of agrobiodiversity conservation services through the use of (monetary or in-kind) reward mechanisms in order to increase the private benefits of agrobiodiversity conservation to farmers.

Given that the marginal commercial value of large scale on-farm agrobiodiversity conservation for the private sector may not be an option, it may be expected that government agencies and NGOs at a local, regional, national, or even international level may be required to take the role of the beneficiary payer in PACS (Narloch *et al.*, 2011a). Furthermore, the implementation of a PACS approach requires a prioritization of the targeted threatened genetic resources, as well as the definition of conservation goals (trade-offs may exist between environmental effectiveness, economic efficiency, and social equity – see Narloch *et al.*, 2011b) and the levels of conservation required to achieve the provision of given ecosystem services derived from agrobiodiversity. Such levels may be associated with safe minimum populations, in turn related to land areas, farmer/community numbers and their spatial distribution, the functionality of local seed systems, and the maintenance of traditional knowledge and culture.

To date, only one such PACS scheme for plant genetic resources has been designed and implemented in the context of on-farm agrobiodiversity conservation. This PACS scheme was implemented in the Bolivian and Peruvian Andean altiplano based on a reverse tender approach at the community level. It includes in-kind payments with the objective of conserving endangered quinoa landraces (Narloch *et al.*, 2011b **(Figure 4)**.

However, PACS or other incentive schemes based around direct payments also have their limitations, particularly with regard to issues related to the sustainability of their funding. It may therefore be that the two types of market creation



Figure 4 Andean farmers conserving quinoa landraces through a competitive tender PACS approach.

instruments (i.e., market chain and PACS approaches) should be seen as complements rather than as substitutes.

In terms of addressing the market development challenges (1–3) highlighted above, it might be expected that a complementary PACS scheme could ensure that market development efforts are focused on priority threatened species/varieties, and the protection of safe minimum populations (Drucker, 2006) of specific threatened crop genetic resources. Furthermore, a PACS approach could facilitate adequate evaluation through farmer experimentation of specific crop genetic resources that are no longer widely used. The overall goal, of course, is that through time-limited PACS support, threatened crop genetic resources would have a potential market. Others would continue to depend on alternative sources of sustainable financing or *ex situ* conservation.

Under such a complementary approach, market value chain development could be used to support the financial sustainability of PACS, while PACS could support the environmental sustainability of the market chain. However, the framework for undertaking a comparative analysis and exploring potential complementarities remains to be developed and a number of key economic and underlying questions need to be answered. These include:

1. the degree to which crop genetic resources within a typical priority conservation portfolio have market potential;
2. the initial investment costs and long-term success rates of agrobiodiversity-based market value chain development (and the volumes required for such market development);
3. how such volumes and costs compare to that required to ensure that an *in situ* population of the crop species/variety in question is considered to be not at risk, as well as how such a safe minimum population can be defined and how can it be linked with the actual provision of specific ecosystem services;
4. how a safe minimum population is spatially distributed (geographically but also in terms of farmer and community numbers) so as to reduce risk of catastrophic loss, as well as help maintain evolutionary processes, seed system functionality, and indigenous knowledge.

Appendix

List of Courses

1. Environmental Economics
2. Ecological Economics
3. Agricultural Economics
4. Economic and Rural Development
5. Economic Development and the Environment

See also: Agrobiodiversity. Economic Value of Biodiversity, Measurements of. Economics of the Regulating Services. Market Economy and Biodiversity. Poverty and Biodiversity. The Value of Biodiversity

References

Alston JM, Chan-Kang C, Marra MC, Pardey PG, and Wyatt TJ (2000) A meta-analysis of rates of return to agricultural R and D: Ex Pede Herculem? *IFPRI Research Report No 113*. Washington, DC: International Food Policy Research Institute.

Altieri MA (1999) The ecological role of biodiversity in agroecosystems. *Agriculture, Ecosystems & Environment* 74: 19–31.

Altieri MA and Merrick L (1987) In-situ conservation of crop genetic resources through maintenance of traditional farming systems. *Economic Botany* 41(1): 86–96.

Andersen R (2008) *Governing Agrobiodiversity: Plant Genetics and Developing Countries*. Hampshire: Ashgate Publishing.

Badstue LB, Bellon MR, Berthaud J, *et al.* (2006) Examining the role of collective action in an informal seed system: A case study from the central valleys of Oaxaca, Mexico. *Human Ecology* 34: 249–273.

Basu AK, Chau NH, and Grote U (2003) Eco-labeling and stages of development. *Review of Development Economics* 7(2): 228–247.

Baumgärtner S (2007) The insurance value of biodiversity in the provision of ecosystem services. *Natural Resource Modelling* 20(1): 87–127.

Baumgärtner S and Quaas M (2009) Agro-biodiversity as natural insurance and the development of financial insurance markets. In: Kontoleon A, Pascual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 16, pp. 293–317. Abingdon, UK: Routledge.

Bellon MR (2009) Do we need crop landraces for the future? Realizing the global option value of in situ conservation. In: Kontoleon A, Pascual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 4, pp. 51–61. Abingdon, UK: Routledge.

Bellon MR, Pham JL, and Jackson MT (1997) Genetic conservation: A role for rice farmers. In: Maxted N, Ford-Lloyd B, and Hawkes JG (eds.) *Plant Genetic Conservation: The In Situ Approach*, pp. 263–289. London/New York: Chapman and Hall.

Birol E, Rayn-Villalba E, and Smale M (2008) Farmer preferences for *Milpa* diversity and genetically modified maize in Mexico: A latent class approach. *Environment and Development Economics* 14(4): 521–540.

Birol E, Smale M, and Gyovai Á (2006) Using a choice experiment to estimate farmers' valuation of agrobiodiversity on Hungarian small farms. *Environmental and Resource Economics* 34(4): 439–469.

ten Brink (2011) *The Economics of Ecosystems and Biodiversity in National and International Policy Making*. London: Earthscan.

Brush S (2000) *Genes in the Field. On-Farm Conservation of Crop Diversity*. Boca Raton, FL, USA: Lewis Publishers.

Brush SB (2004) *Farmers' Bounty: Locating Crop Diversity in the Contemporary World*. New Haven, CN, USA: Yale University Press.

Brush SB and Perales HR (2007) A maize landscape: Ethnicity and agro-biodiversity in Chiapas Mexico. *Agriculture, Ecosystems and Environment* 121: 211–221.

Carpenter SR, Walker BH, Anderies JM, and Abel N (2001) From metaphor to measurement: resilience of what to what? *Ecosystems* 4: 765–781.

Cavatassi R, Lipper L, and Narloch U (2011) Modern variety adoption and risk management in drought prone areas: Insights from the sorghum farmers of eastern Ethiopia. *Agricultural Economics* 42: 279–292.

Ceccarelli S, Grando S, Maatougui M, *et al.* (2010) Plant breeding and climate changes. *The Journal of Agricultural Science* 148: 627–637.

Chapin FS, Zavaleta ES, Eviner VT, *et al.* (2000) Consequences of changing biodiversity. *Nature* 405: 234–242.

Cicia G, D'Ercole E, and Marino D (2003) Costs and benefits of preserving farm animal genetic resources from extinction: CVM and bio-economic model for valuing a conservation program for the Italian Pentro horse. *Ecological Economics* 45(3): 445–459.

Collins WW and Qualset CO (1999) *Biodiversity in Agroecosystems*, 334 pp. Boca Raton, FL, USA: CRC Press, Inc.

Cooper HD, Spillane C, and Hodgkin T (2001) *Broadening the Genetic Base of Crop Production*. London: CABI Publishing.

Di Falco S and Chavas J-P (2006) Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture. *European Review of Agricultural Economics* 33(3): 289–314.

Di Falco S and Chavas J-P (2008) Rainfall shocks, resilience and the dynamic effects of crop biodiversity on the productivity of the agroecosystem. *Land Economics* 84(*1*): 83–96.

Di Falco S and Chavas J-P (2009) On crop biodiversity, risk exposure and food security in the highlands of Ethiopia. *American Journal of Agricultural Economics* 91(3): 599–611.

Di Falco S and Perrings C (2003) Crop genetic diversity, productivity and stability of agroecosystems: A theoretical and empirical investigation. *Scottish Journal of Political Economy* 50(2): 207–216.

Di Falco S and Perrings C (2005) Crop biodiversity, risk management and the implications of agricultural assistance. *Ecological Economics* 55(4): 459–466.

Drucker A (2006) An application of the use of safe minimum standards in the conservation of livestock biodiversity. *Environment and Development Economics* 11(1): 77–94.

Drucker AG and Rodriguez LC (2009) Development, intensification and the conservation and sustainable use of farm animal genetic resources. In: Kontoleon A, Pasqual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 7, pp. 92–110. Abingdon, UK: Routledge.

Dumaresq D, Carpenter D, and Lockie S (2010) The human ecology of agrobiodiversity. In: Lockie S and Carpeter D (eds.) *Agriculture, Biodiversity and Markets: Livelihoods and Agroecology in Comparative Perspective*, pp. 33–46. London: Earthscan.

Edwards C and Abivardi C (1998) The value of biodiversity: Where ecology and economy blend. *Biological Conservation* 83: 239–246.

Engel S, Pagiola S, and Wunder S (2008) Designing payments for environmental services in theory and practice: An overview of the issue. *Ecological Economics* 65(4): 663–674.

Evenson RE and Gollin D (eds.) (2003) *Crop Variety Improvement and its Effect on Productivity: The Impact of International Agricultural Research*. Wallingford, UK: CABI Publishing.

Evenson RE, Gollin D, and Santaniello V (eds.) (1998) *Agricultural Values of Plant Genetic Resources*. Wallingford, UK: CAB International.

Faith DP, Magallón S, Hendry AP, Conti E, Yahara T, and Donoghue MJ (2010) Ecosystem services: An evolutionary perspective on the links between biodiversity and human well-being. *Current Opinion in Environmental Sustainability* 2: 66–74.

FAO (1998) *The State of the World's Plant Genetic Resources for Food and Agriculture*. Rome, Italy: Food and Agriculture Organization of the UN.

FAO (2007a) *The State of the World's Animal Genetic Resources for Food and Agriculture*. Rome, Italy: Food and Agriculture Organization of the UN.

FAO (2007b) *State of Food and Agriculture: Paying Farmers for Environmental Services*. Rome, Italy: Food and Agriculture Organization of the UN.

FAO (2010) *Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture*. Rome, Italy: Food and Agriculture Organization of the UN.

Ferraro P and Simpson R (2002) The cost-effectiveness of conservation payments. *Land Economics* 78: 339–353.

Ferraro PJ (2001) Global habitat protection: Limitations of development interventions and a role for conservation performance payments. *Conservation Biology* 15(4): 923: 990–1000.

Folke C, Carpenter S, Walker B, *et al.* (2004) Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology and Systematics* 35: 557–581.

Gautam JC and Paant KP (2010) Commerzialization and market linkages fro promoting use of local rice varieties. A Nepalese case study. In: Wale E, Drucker AG, and Zander KK (eds.) *The Economics of Managing Crop Diversity On-Farm: Case Studies from the Genetic Resource Policy Initiative*, pp. 111–124. London, UK: Earthscan.

Gemessa SA and Zander KK (2010) Economic analysis of Ethiopian farmers' preferences for crop variety attributes: A choice experiment approach. In: Wale E, Drucker AG, and Zander KK (eds.) *The Economics of Managing Crop Diversity On-Farm: Case studies from the Genetic Resources Policy Initiative*, pp. 25–44. London: Earthscan.

Gepts P (2006) Plant genetic resources conservation and utilization: The accomplishments and future of a societal insurance policy. *Crop Science* 46: 2278–2292.

Gollin D, Smale M, and Skovmand B (2000) Searching an *ex situ* collection of wheat genetic resources. *American Journal of Agricultural Economics* 82: 812–827.

Gruère G, Giuliani A, and Smale M (2009) Marketing underutilized plant species for the poor: A conceptual framework. In: Kontoleon A, Pascual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 5, pp. 62–81. Abingdon, UK: Routledge.

Hajjar R, Jarvis DI, and Gemmill-Herren B (2008) The utility of crop genetic diversity in maintaining ecosystem services. *Agriculture, Ecosystems and Environment* 123: 261–270.

Hawkes JR (1983) *The Diversity of Crop Plants*. Cambridge MA: Harvard University Press.

Heisey PW, Smale M, Byerlee D, and Souza E (1997) Wheat rusts and the costs of genetic diversity in the Punjab of Pakistan. *American Journal of Agricultural Economics* 79: 726–737.

Hellin J, Higman S, and Keleman A (2010) Value chain coordination for agrobiodiversity conservation. In: Lockie S and Carpeter D (eds.) *Agriculture, Biodiversity and Markets: Livelihoods and Agroecology in Comparative Perspective*, pp. 213–228. London: Earthscan.

Hermann M and Bernet T (2009) The transition of maca from neglect to market prominence: Lessons for improving use strategies and market chains of minor crops. Agricultural Biodiversity and Livelihoods Discussion Papers 1. Rome, Italy: Bioversity International. Available at www.bioversityinternational.org

Hodge I (2000) Agri-environmental relationships and the choice of policy mechanism. *World Economy* 23: 257–273.

Hodgkin T, Rana R, Tuxill J, *et al.* (2007) Seed systems and crop genetic diversity in agroecosystems. In: Jarvis DI, Padoch C, and Cooper HD (eds.) *Managing Biodiversity in Agricultural Systems*, pp. 77–116. New York: Columbia University Press.

Hodgkin T, Rao R, Cibrian-Jaramillo A, and Gaiji S (2003) The use of *ex-situ* conserved plant genetic resources. *Plant Genetic Resources* 1(1): 19–29.

Hoeschle-Zeledon I, Padulosi S, Giuliani A, and Al-Haj Ibrahim U (2009) Making the most of wild and relict species – Experiences and lessons. *Bocconea* 23: 129–143. ISSN 1120-4060.

Jackson LE, Pascual U, and Hodgkin T (2007a) Utilizing and conserving agrobiodiversity in agricultural landscapes. *Agriculture, Ecosystems and Environment* 121: 196–210.

Jackson LE, Rosenstock T, Thomas M, Wright J, and Symstad A (2009) Managed ecosystems: Biodiversity and ecosystem functions in landscapes modified by human use. In: Naeem S, Bunker D, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity and Human Impacts*, pp. 178–194. Oxford, UK: Oxford University Press.

Jarvis DI, Brown AHD, Imbruce I, *et al.* (2007) Managing crop disease in traditional agroecosystms: Benefits and hazards of genetic diversity. In: Jarvis DI, Padoch C, and Cooper HD (eds.) *Managing Biodiversity in Agricultural Systems*, pp. 292–319. New York: Columbia University Press.

Jarvis DI, Hodgkin T, Sthapit BR, Fadda C, and Lopez-Noriega I (2011) A heuristic framework for identifying multiple ways of supporting the conservation and use of traditional crop varieties within the agricultural production system. *Critical Reviews in Plant Sciences* 30: 125–176.

Kleijn D and Sutherland WJ (2003) How effective are European agri-environment schemes in conserving and promoting biodiversity? *Journal of Applied Ecology* 40(6): 947–969.

Kontoleon A, Pascual U, and Smale M (2009) *Agrobiodiversity Conservation and Economic Development*. Abingdon, UK: Routledge.

Kontoleon A, Pascual U, and Swanson T (2007) *Biodiversity Economics*. Cambridge, UK: Cambridge University Press.

Koo B, Pardey PG, and Wright BD (2004) *Saving Seeds: The Economics of Conserving Crop Genetic Resources Ex Situ in the Future Harvest Centres of the CGIAR*. Wallingford, UK: CABI Publishing.

Krishna V, Pascual U, and Zilberman D (2010) Assessing the potential of labelling schemes for in-situ landrace conservation: An example from India. *Environment and Development Economics* 15: 127–151.

Kruijssen F, Keizer M, and Giuliani A (2009) Collective action for smale-scale producers of agricultural biodiversity products. *Food Policy* 34: 46–52.

Leakey RBB (2012) Participatory domestication of indigenous fruit and nut trees: new crops for sustainable agriculture in developing countries. In: Gepts P, Famula RR, Bettinger RL, *et al.* (eds.) *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability*, pp. 479–501. Cambridge: Cambridge University Press.

Letourneau DK, Armbrecht I, Rivera BS, *et al.* (2011) Does plant diversity benefit agroecosystems? A synthetic review. *Ecol. Appl.* 21: 9–21.

Lin BB (2011) Resilience in agriculture through crop diversification: Adaptive management for environmental change. *BioScience* 61: 183–193.

Lipper L and Cooper D (2009) Managing agricultural biodiversity for sustainable use: Balancing the benefits in the field. In: Kontoleon A, Pascual U, and Smale M (eds.) *Agrobiodiversity Conservation and Economic Development*, ch. 2, pp. 27–39. Abingdon, UK: Routledge.

Lipper L, Sakuyama T, Stringer R, and Zilberman D (eds.) (2009) *Payment for Environmental Services in Agricultural Landscapes: Economic Policies and Poverty Reduction in Developing Countries*. Rome, Italy: Food and Agriculture Organization of the UN. Springer.

MEA (Millennium Ecosystem Assessment) (2005) *Ecosystems and Human Well-being: Biodiversity Synthesis*. Washington, DC, USA: World Resources Institute.

Moore G (2010) Multilateral and national regulatory regimes for agrobiodiverity. In: Lockie S and Carpeter D (eds.) *Agriculture, Biodiversity and Markets: Livelihoods and Agroecology in Comparative Perspective*, pp. 47–60. London: Earthscan.

Muradian R, Corbera E, Pascual U, Kosoy N, and May PH (2010) Reconciling theory and practice: An alternative conceptual framework for understanding payments for environmental services. *Ecological Economics* 69(6): 1202–1208.

Narloch U, Drucker A, and Pascual U (2011a) Payments for agrobiodiversity conservation services for sustained on-farm utilization of plant and animal genetic resources. *Ecological Economics* 70(11): 1837–1845.

Narloch U, Pascual U, and Drucker AG (2011b) Cost-effectiveness targeting under multiple conservation goals and equity considerations in the Andes. *Environmental Conservation* 38(4): 417–425.

Nunes PALD and van den Bergh CJM (2001) Economic valuation of biodiversity: Sense or nonsense? *Ecological Economics* 39(2): 203–222.

O'Farrell PJ and Anderson PML (2010) Sustainable multifunctional landscapes: A review to implementation. *Current Opinion in Environmental Sustainability* 2: 59–65.

Omer A, Pascual U, and Russell NP (2007) Biodiversity conservation and productivity in intensive agricultural systems. *Journal of Agricultural Economics* 58: 308–329.

Omer A, Pascual U, and Russell NP (2010) A theoretical model of agrobiodiversity as a supporting service for sustainable agricultural intensification. *Ecological Economics* 69(10): 1926–1933.

Padulosi S, Bhag Mal S, Bala Ravi J, *et al.* (2009) Food security and climate change: Role of plant genetic resources of minor millets. *Indian Journal of Plant Genetic Resources* 22(1): 1–16.

Padulosi S, Hodgkin T, Williams JT, and Haq N (2002) Underutilized crops: Challenges and opportunities in the 21th century. In: Engels JMM, Rao VR, Brown AHD, and Jackson MT (eds.) *Managing Plant Genetic Diversity*, pp. 323–338. Wallingford, UK: CABI Publishing.

Pant KP, Gautam JC, and Wale E (2010) Valuation of rice diversity in Nepal: A trait-based approach. In: Wale E, Drucker AG, and Zander KK (eds.) *The Economics of Managing Crop Diversity On-Farm: Case studies from the Genetic Resources Policy Initiative*, pp. 45–64. London: Earthscan.

Pascual U, Muradian R, Brander L, *et al.* (2010) The economics of valuing ecosystem services and biodiversity. In: Kumar P (ed.) *The Economics of Ecosystems and Biodiversity Ecological and Economic Foundations*, ch. 5, pp. 183–256. London: Earthscan.

Pascual U, Narloch U, Nordhagen S, and Drucker A (2011) The economics of agrobiodiversity conservation for food security under climate change. *Economía Agraria y Recursos Naturales* 11(1): 191–220.

Pascual U and Perrings C (2007) Developing incentives and economic mechanisms for in-situ biodiversity conservation in agricultural landscapes. *Agriculture, Ecosystems and Environment* 121: 256–268.

Pearce DW and Moran D (1994) *The Economic Value of Biodiversity.* London: Earthscan.

Perrings C (1998) Resilience in the dynamics of economy-environment systems. *Environmental and Resource Economics* 11(3–4): 503–520.

Perrings C (2001) The economics of biodiversity loss and agricultural development in low income countries. In: Lee DR and Barrett CB (eds.) *Tradeoffs or Synergies? Agricultural Intensification, Economic Development and the Environment*, pp. 57–72. Wallingford, UK: CAB International.

Perrings C (2005) Mitigation and adaptation strategies for the control of biological invasions. *Ecological Economics* 52: 315–325.

Perrings C, Baumgärtner S, Brock WA, *et al.* (2009) The economics of biodiversity and ecosystem services. In: Naeem S, Bunker D, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*, ch. 17. Oxford, UK: Oxford University Press.

Perrings C, Jackson L, Bawa K, *et al.* (2006) Biodiversity in agricultural landscapes: Saving natural capital without losing interest. *Conservation Biology* 20(2): 263–264.

Plucknett DL, Smith NHJ, Williams JT, and Anishetty NM (1987) *Gene Banks and the World's Food*. Princeton: Princeton University Press.

Quaas MF and Baumgärtner S (2008) Natural vs. financial insurance in the management of public-good ecosystems. *Ecological Economics* 65(2): 397–406.

Qualset CO and Shands HL (2005) *Safeguarding the Future of US Agriculture: The Need to Conserve Threatened Collections of Crop Diversity Worldwide*. Davis, CA, USA: University of California, Division of Agriculture and Natural Resources, Genetic Resources Conservation Program.

Reynolds M, Dreccer M, and Trethowan R (2007) Drought-adaptive traits derived from wheat wild relatives and landraces. *Journal of Experimental Botany* 58(2): 177–186.

Ricketts TH, Daily GC, Ehrlich PR, and Michener CD (2004) Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences of the United States of America* 101: 12579–12582.

Rojas W, Valdivia R, Padulosi S, *et al.* (2009) From neglect to limelight: Issues, methods and approaches in enhancing sustainable conservation and use of Andean grains in Peru and Bolivia. *Journal of Agriculture and Rural Development in the Tropics* (supplement 92): 1–32.

Smale M (2006) *Valuing Crop Biodiversity: On-Farm Genetic Resources and Economic Change*. Wallingford, UK: CABI Publishing.

Smale M, Bellon RM, and AguirreGomez JA (2001) Maize diversity, variety attributes, and farmers' choices in Southeastern Guanajuato, Mexico. *Economic Development and Cultural Change* 50: 201–225.

Smale M and Day-Rubenstein K (2002) The demand for crop genetic resources: International use of the US national plant germplasm system. *World Development* 30: 1639–1655.

Smale M and Drucker AG (2007) Agricultural development and the diversity of crop and livestock genetic resources: A review of the economics literature. In: Kontoleon A, Pascual U, and Swanson T (eds.) *Biodiversity Economics: Principles, Methods and Applications*, pp. 623–648. Cambridge, UK: Cambridge University Press.

Smale M, Hartell J, Heisey PW, and Senauer B (1998) The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan. *American Journal of Agricultural Economics* 80: 482–493.

Sperling L, Cooper HD, and Remington T (2008) Moving towards more effective seed aid. *Journal of Development Studies* 44: 586–612.

Stromberg P, Pascual U, and Bellon M (2010) Seed systems and farmers' seed choices: The case of maize in the Peruvian Amazon. *Human Ecology* 38: 539–553.

Swanson T (ed.) (1998) *The Economics and Ecology of Biodiversity Decline: The Forces Driving Global Change*. Cambridge: Cambridge University Press.

Swift MJ, Izac AMN, and van Noordwijk M (2004) Biodiversity and ecosystem services in agricultural landscapes – are we asking the right questions? *Agriculture, Ecosystems, and Environment* 104: 113–124.

Thrupp LA (2000) Linking agricultural biodiversity and food security: The valuable role of agrobiodiversity for sustainable agriculture. *International Affairs* 76(2): 283–297.

Tscharntke T, Clough Y, Jackson L, *et al.* (2012) Global Food Security, Biodiversity Conservation and the Future of Agricultural Intensification. *Biological Conservation* 151(1): 53–59.

Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, and Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity ecosystem service management. *Ecology Letters* 8: 857–874.

Turner RK and Daily GD (2008) The ecosystem service framework and natural capital conservation. *Environmental and Resource Economics* 39: 25–35.

Van Dusen ME and Taylor JE (2005) Missing markets and crop diversity: Evidence from Mexico. *Environment and Development Economics* 10(4): 513–531.

Varshney RK, Glaszmann J-C, Leung H, and Ribaut J-M (2010) More genomic resources for less-studied crops. *Trends in Biotechnology* 28L: 452–460.

Varshney RK, Hoisington DA, Nayak SN, and Graner A (2009) Molecular plant breeding: Methodology and achievements. In: Somers DJ, Landridge P, Gustafson JP, *et al.* (eds.) *Methods in Molecular Biology, Plant Genomics*, Vol. 513, pp. 283–305.

Veleto JR and Skarbø K (2009) Sowing the seeds: Anthropological contributions to agrobiodiversity studies. *Culture and Agriculture* 31(2): 73–87.

Villa TCC, Maxted N, Scholten M, and Ford-Loyd B (2006) Defining and identifying crop landraces. *Plant Genetic Resources* 3(3): 373–384.

Wale E (2008) A study on financial opportunity costs of growing local varieties of sorghum in Ethiopia: Implications for farm conservation policy. *Ecological Economics* 64: 603–610.

Wale E, Drucker AG, and Zander KK (2010) *The Economics of Managing Crop Diversity On-Farm: Case studies from the Genetic Resources Policy Initiative*. London: Earthscan.

Widawsky D and Rozelle SD (1998) Varietal diversity and yield variability in Chinese rice production. In: Smale M (ed.) *Farmers, Gene Banks, and Crop Breeding: Economic Analyses of Diversity in Wheat, Maize and Rice*, pp. 159–172. Dordrecht: Kluwer Academic Press and CIMMYT.

Will M (2008) *Promoting Value Chains of Neglected and Underutilized Species for Pro-Poor Growth and Biodiversity Conservation*. Rome, Italy: Global Facilitation Unit for Underutilized Species.

Wood D and Lenné JM (1997) The conservation of agrobiodiversity on-farm: Questioning the emerging paradigm. *Biodiversity and Conservation* 6: 109–129.

Wunder S, Engel S, and Pagiola S (2008) Taking stock: a comparative analysis of payments for environmental services programs in developed and developing countries. *Ecological Economics* 65(4): 834–852.

Economics of the Regulating Services

Edward B Barbier, University of Wyoming, Laramie, WY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Consumer surplus A measure of the welfare that people gain from the consumption of goods and services, usually measured as the difference between the price that they are actually willing and able to pay for a good or service and the actual market price.

Direct use values Benefits gained through consumptive and nonconsumptive uses that involve some form of direct physical interaction with environmental goods and services, such as recreational activities, resource harvesting, drinking clean water, breathing unpolluted air, and so forth.

Economic valuation The various methods used by economists to determine the willingness to pay of individuals for the benefits derived from ecosystems, many of which are nonmarketed.

Ecosystem functions Biogeochemical cycling, including primary production (photosynthesis), nutrient and water cycling, and materials decomposition, and the flow, storage and transformation of materials, and energy through an ecosystem, including through food web processes such as pollination, predation, and parasitism.

Ecosystem services The benefits (i.e., goods and services) people obtain from ecosystems.

Ecosystem structure The biotic (living) and abiotic (nonliving) components of an ecosystem, and the interactions between them.

Indirect use values Ecosystem services or benefits whose values can only be measured indirectly, since they are derived from supporting and protecting activities that have directly measurable values (e.g., coastal habitat as breeding and nursery grounds for off-shore fisheries; coral reefs providing protection against storm surge; upper watershed forests controlling soil erosion and runoff).

Nonuse values Values ascribed to the natural environment that are independent of any human interaction with it; for example, an individual valuing the pure "existence" of a natural habitat or ecosystem or wanting to "bequest" it to future generations.

Producer surplus A measure of the welfare that producers gain from selling goods and services, usually measured as the difference between the actual market price and the minimum amount that they would be willing to accept for the good or service.

Regulating services The benefits obtained from the regulation of ecosystem processes, including flood protection, climate regulation, human disease regulation, water purification, air quality maintenance, pollination, and pest control.

Introduction

The Millennium Ecosystem Assessment defines ecosystem services as "the benefits people obtain from ecosystems" (MA 2005). Among the myriad types of benefits that arise through normal ecosystem functioning, a subset can be defined as *regulating services*. These are the benefits obtained from the regulation of ecosystem processes, and include flood protection, climate regulation, human disease regulation, water purification, air quality maintenance, pollination, and pest control.

As indicated in **Table 1**, many important regulating services are declining globally due to human-induced change to ecosystems. The main reason for this loss is because we have failed to value these benefits adequately. Too often, economic development decisions are taken that harm ecosystems and the services they provide, because the value of important ecosystem services is either ignored or unknown. Because they are the most difficult to value, important regulating services are usually the main ecosystem benefits that are not considered in decisions to convert or exploit natural habitats.

However, recent interest in regulating services has grown because of concerns of the global impacts of climate change, including effects on future freshwater supplies, the spread of disease and pests, degradation of forest and wetland habitats, and biological invasions. The 2004 Indian Ocean Tsunami and the 2005 Hurricanes Katrina and Rita have also increased interest in the storm protection and food control benefits of coastal ecosystems.

Although the ecological analysis of key regulatory ecosystem functions has been well established, the economic valuation of the benefits provided by these functions is relatively new. As a report from The US National Research Council (2005, p. 2) has emphasized, "the fundamental challenge of valuing ecosystem services lies in providing an explicit description and adequate assessment of the links between the structure and functions of natural systems, the benefits (i.e., goods and services) derived by humanity, and their subsequent values". Moreover, it has been increasingly recognized by economists and ecologists that the greatest "challenge" they face is in valuing the ecosystem services provided by ecosystem regulatory functions.

The following reviews current efforts to develop the economic analysis of regulating services. As these services arise through the regulatory functioning of ecosystems, the review begins by briefly outlining this critical ecological production process (The linkages between biodiversity and ecosystem function are also addressed in two articles of this encyclopedia.) A further problem is that regulating services are not marketed. In addition, for many regulating services, there is

Table 1 Global status of key ecosystem services

Condition globally has been enhanced	*Condition globally has been degraded*	*Condition globally is mixed*
Crops	Capture fisheries	Timber
Livestock	Wild foods	Cotton, hemp, silk, and other fiber crops
Aquaculture	Wood fuel	Water regulation
Global climate regulation	Genetic resources	Disease regulation
	Biochemicals, natural medicines, and pharmaceuticals	Recreation and ecotourism
	Fresh water	
	Air quality regulation	
	Regional and local climate regulation	
	Erosion regulation	
	Water purification and waste treatment	
	Pest regulation	
	Pollination	
	Natural hazard regulation	
	Spiritual and religious values	
	Esthetic values	

Notes: Enhancement is defined as either increased production of or change in the ecosystem good or service that leads to greater benefits for people.
Degradation is defined as if current use exceeds sustainable levels, or a reduction in the benefits obtained from the good or service due to either some human-induced change or use exceeding its limits.
Mixed status implies that the condition of the good or the service globally has experienced enhancement in some regions but degradation in others.

inadequate knowledge to link changes in ecosystem structure and function to the production of these valuable benefits. These measurement issues are important limitations to the valuation of regulating services, and thus much of the review discusses how economists, ecologists, and other natural scientists are collaborating on overcoming such challenges. To illustrate the significance of valuing regulating services to development versus conservation decisions, a case study of mangrove conversion to shrimp farming in Thailand will be discussed in detail. Finally, there is growing awareness that the regulatory functions underlying key services are often influenced by the spatial extent, or scale, of ecosystems. As a consequence, the value of regulating services will vary across an ecological landscape. The significance of spatially varying regulating services for conservation versus development decisions is shown with the example of storm protection provided by mangroves in the Thailand case study.

Ecosystems, Biodiversity, and Regulating Services

Within its prescribed area or location, an ecosystem comprises its abiotic (nonliving) environment and the biotic (living) groupings of plant and animal species, or communities. The biotic and abiotic components, and the interactions between them, are often referred to as the *ecosystem structure*.

Two important *ecosystem functions* are carried out in every ecosystem: biogeochemical cycling and flow of energy. Important processes of biogeochemical cycling include primary production (photosynthesis), nutrient and water cycling, and materials decomposition. The flow, storage, and transformation of materials and energy through the system are also influenced by processes that link organisms with each other, such as the food web, which is made up of interlocking food chains. These food chains are often characterized by other important functions, such as pollination, predation, and parasitism.

Increasingly, evidence suggests that biodiversity plays a pivotal role in ecosystem functioning (Duffy, 2009; Duffy *et al.*, 2007; Naeem *et al.*, 2009). The consequence, as described by Levin and Lubchenco (2008, p. 30), is that diversity and heterogeneity is also critical to the adaptability of an ecosystem, and "lowered diversity exposes ecosystems to catastrophic change".

The structure and functions of an ecosystem provide valuable goods and services to humans. For example, some of the living organisms found in an ecosystem might be harvested or hunted for food, collected for raw materials, or simply valued because they are esthetically pleasing. Some of the ecosystem functions, such as nutrient and water cycling, can also benefit humans through purifying water, controlling floods, recharging aquifers, reducing pollution, or simply by providing more pleasing environments for recreation. These various benefits provided by an ecosystem via its structure and functions are now referred to as *ecosystem services* (Millennium Ecosystem Assessment, 2005).

However, it is important to recognize that the structure and functions of an ecosystem are not synonymous with its services or economic benefits. Ecosystem structure and functions describe the components of an ecosystem and its biophysical relationship regardless of whether or not humans benefit from them. Only if they contribute to human well-being do these components and relationships generate an ''ecosystem service''. The role of economic valuation is to determine how changes in ecosystem goods and services affect an individual's well-being, and then to determine how much the individual is either willing to pay for changes that have a positive welfare impact, or conversely, how much the individual is willing to accept as compensation to avoid a negative effect. In other words, the *ecological production* of ecosystem services is linked to the *economic valuation* of these services, but these two relationships are distinct (Barbier, 2007, 2011; Barbier *et al.*, 2009).

Figure 1 provides a visual summary of the links between the structure and function of ecosystems, the ecological

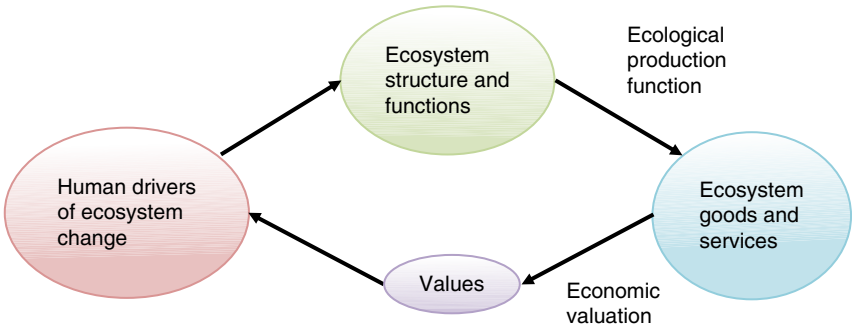


Figure 1 Links between the structure and function of ecosystems, the ecological production of ecosystem goods and services, and the economic valuation of the latter benefits.

Table 2 Some regulating services of ecosystems

Ecosystem regulatory functions	*Ecosystem processes and components*	*Ecosystem services (benefits)*
Gas regulation	Role of ecosystems in biogeochemical processes	Ultraviolet-B protection Maintenance of air quality Influence of climate
Climate regulation	Influence of land cover and biologically mediated processes	Maintenance of temperature, precipitation
Disturbance prevention	Influence of system structure on dampening environmental disturbance	Storm protection Flood mitigation
Water regulation	Role of land cover in regulating runoff, river discharge and infiltration	Drainage and natural irrigation Flood mitigation Groundwater recharge
Soil retention	Role of vegetation root matrix and soil biota in soil structure	Maintenance of arable land Prevention of damage from erosion and siltation
Soil formation	Weathering of rock and organic matter accumulation	Maintenance of productivity on arable land
Nutrient regulation	Role of biota in storage and recycling of nutrients	Maintenance of productive ecosystems
Waste treatment	Removal or breakdown of nutrients and compounds	Pollution control and detoxification

Source: Adapted from Table 3-3 in National Research Council (2005) *Valuing Ecosystem Services: Toward Better Environmental Decision Making*. Washington, DC: The National Academies Press.

production of ecosystem goods and services, and the economic valuation of the latter benefits. Human drivers of ecosystem change affect important ecosystem processes and functions and their controlling components. Assessing this change is crucial yet difficult. However, even more challenging is determining how the ecological production of ecosystem goods and services arises from the structure and functions of an ecosystem. Probably the greatest remaining difficulty in the successful valuation of ecosystem goods and services is to integrate studies of the ecological production function (PF) with nonmarket valuation of the resulting benefits.

Thus, mapping ecosystem structure and function to a flow of ecosystem services that can then be valued is no easy task. This is especially the case for regulating services. As Table 2 indicates, for these services, we are beginning to understand how certain ecosystem regulatory functions, moderated by key ecosystem processes and components, yield a range of ecosystem services. Many of these benefits, such as storm and flood protection, climate regulation, and water purification, are highly valuable. As human economic activities exploit, degrade, and convert ecosystems, many such important regulating services will be affected. But tracing through the resulting changes in ecosystem structure and function to the impacts on various regulating services, and finally, to the effects on human welfare requires overcoming a number of measurement challenges.

Measurement challenges

The most significant problem is that regulating services are not marketed. These services arise from ecosystem processes and functions that benefit human beings largely without any additional input from them. This is especially the case for many important services, such as coastal protection, nutrient cycling, erosion control, water purification, and carbon sequestration. In recent years, substantial progress has been made by economists working with ecologists and other natural scientists in applying environmental valuation methodologies to assess the welfare contribution of some of these regulating services. The various nonmarket valuation methods used for ecosystem services are essentially the standard techniques that are available to economists. Later on in this article, we discuss the issues that arise in applying these methods in

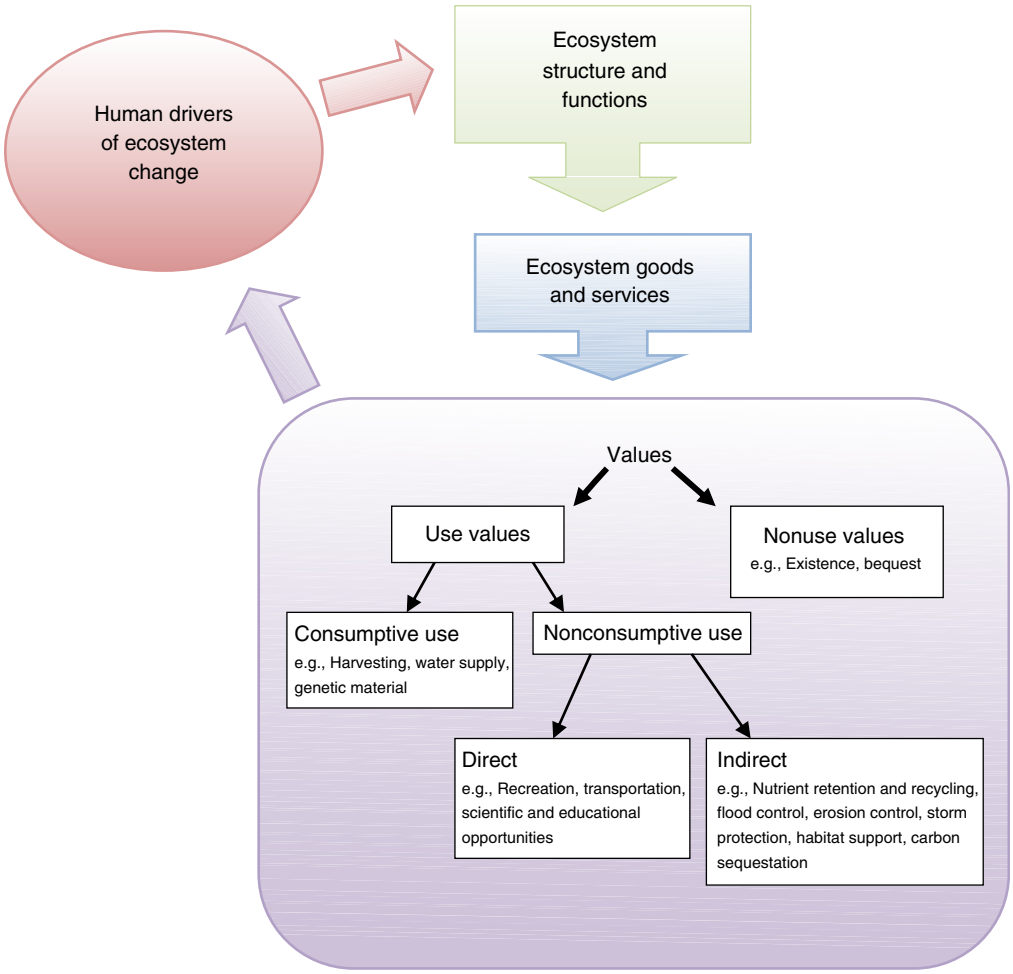


Figure 2 More detailed version of Figure 1 that emphasizes the economic valuation component.

regulating services. Nevertheless, what makes applying these methods to estimate the value of a nonmarketed regulating service especially difficult is that it requires three important, and inter-related, steps.

The first step involves determining how best to characterize the change in ecosystem structure, functions, and processes that gives rise to the change in the ecosystem regulating service. For example, the change could be in the spatial area or quality of a particular type of ecosystem, such as a mangrove forest, marsh vegetation, or watershed extent. It could also be a change in a key population, such as fish or main predator. Alternatively, the change could be due to variation in the flow of water, energy, or nutrients through the system, such as the variability in tidal surges due to coastal storm events or the influx of organic waste from pollution upstream from estuarine and coastal ecosystems.

The second step requires tracing how the changes in ecosystem structure, functions, and processes influence the quantities and qualities of ecosystem regulating service flows to people. Underlying each ecosystem service is a range of important energy flow, biogeochemical, and biotic processes and functions (Naeem *et al.*, 2009). For example, water purification by seagrass beds is linked to the ecological processes of nutrient uptake and suspended particle deposition. However, the key ecological process and functions that generate an ecosystem service are in turn controlled by certain abiotic and biotic components that are unique to each ecosystem's structure. The various controlling components that may affect nutrient uptake and particle deposition by seagrass ecosystems include seagrass species and density, nutrient load, water residence time, hydrodynamic conditions, and light availability. Only when these first two steps are completed is it possible to conduct the final step, which involves using an existing economic valuation method to assess the impact on human well-being that results from the change in ecosystem goods and services.

Despite the progress made in improving the application of environmental valuation methodologies to nonmarket ecosystem regulating services, a number of important challenges still arise in applying these methods. To aid our subsequent discussion of valuation issues, it is useful to look at a more detailed version of **Figure 1** that emphasizes the economic valuation component of the latter diagram (see **Figure 2**).

Valuing Nonmarket Regulating Services

As indicated in **Figure 2**, there are a number of different ways in which humans benefit from, or value, ecosystem goods, and

Table 3 Various nonmarket valuation methods applied to ecosystem services

Valuation method	*Types of value estimated*	*Common types of applications*	*Ecosystem services valued*
Travel cost	Direct use	Recreation	Maintenance of beneficial species, productive ecosystems, and biodiversity
Averting behavior	Direct use	Environmental impacts on human health	Pollution control and detoxification
Hedonic price	Direct and indirect use	Environmental impacts on residential property and human morbidity and mortality	Storm protection; flood mitigation; maintenance of air quality
Production function	Indirect use	Commercial and recreational fishing; agricultural systems; control of invasive species; watershed protection; damage costs avoided	Maintenance of beneficial species; maintenance of arable land and agricultural productivity; prevention of damage from erosion and siltation; groundwater recharge; drainage and natural irrigation; storm protection; flood mitigation
Replacement cost	Indirect use	Damage costs avoided; freshwater supply	Drainage and natural irrigation; storm protection; flood mitigation
Stated preference	Use and nonuse	Recreation; environmental impacts on human health and residential property; damage costs avoided; existence and bequest values of preserving ecosystems	All of the above

Source: Adapted from Table 3-3 in National Research Council (2005) *Valuing Ecosystem Services: Toward Better Environmental Decision Making*. Washington, DC: The National Academies Press.

services. The first distinction is between the *use values* as opposed to *nonuse values* arising from these goods and services. Typically, use values involve some human "interaction" with the environment whereas nonuse values do not, as they represent an individual valuing the pure "existence" of a natural habitat or ecosystem or wanting to "bequest" it to future generations. Direct use values refer to both consumptive and nonconsumptive uses that involve some form of direct physical interaction with environmental goods and services, such as recreational activities, resource harvesting, drinking clean water, breathing unpolluted air, and so forth. Indirect use values refer to those ecosystem services whose values can only be measured indirectly, since they are derived from supporting and protecting activities that have directly measurable values. For example, for wetlands, the indirect use values associated with ecosystems services include coastal protection, erosion control, flood protection, water purification, carbon sequestration, maintenance of temperature and precipitation, and habitat support for fishing, hunting, and foraging activities outside the wetlands.

It is clear that many regulating services, such as those listed in **Table 2**, are mainly associated with indirect use values. These are some of the most difficult values to estimate. **Table 3** indicates the various nonmarket methods that can be used for valuing ecosystem goods and services. As shown in the table, the methods used are essentially the standard nonmarket valuation techniques that are available to economists. However, the application of nonmarket valuation to ecosystem goods and services is not without difficulties. Although it is beyond the scope of this review to discuss in detail the advantages and the shortcomings of the different methods and their application, we can summarize some of the key issues (for further elaboration on these issues, see Barbier, 2007; Barbier *et al*., 2009; Bateman *et al*., 2011; EPA, 2009; Hanley and Barbier, 2009; Mendelsohn and Olmstead, 2009; National Research Council, 2005).

Firstly, the application of some of the valuation methods listed in **Table 3** is often limited to specific types of ecological goods and services. For example, the travel cost method is used principally for those environmental values that enhance individuals' enjoyment of recreation and tourism; averting behavior models are best applied to the health effects arising from environmental pollution. Thus, these valuation methods are best suited to estimating direct use values. These methods are unlikely to be used to value regulating services, unless of course such services have a direct impact on human welfare, such as through reducing pollution that affects human health or recreational activities. Similarly, hedonic wage and property models are used primarily for assessing work-related environmental hazards and environmental impacts on property values, respectively. These hedonic methods can only be applied to valuing regulating services if they influence work-related hazards or property values.

In contrast, stated-preference methods, which include contingent valuation methods and choice modeling, have the potential to be used widely in valuing ecosystem goods and services. These valuation methods share the common approach of surveying individuals who benefit from an ecological service or a range of services, in the hope that analysis of these responses will provide an accurate measure of the individuals' willingness to pay for the service or services. In addition, stated-preference methods can go beyond estimating the value to individuals of single and even multiple benefits of ecosystems and in some cases elicit nonuse values that individuals attach to ensuring that a preserved and well-functioning system will be around for future generations to enjoy. For example, a study of mangrove-dependent coastal communities in Micronesia demonstrated through the use of contingent valuation techniques that the communities place some value on the existence and ecosystem functions of mangroves over and above the value of mangroves' marketable products (Naylor and Drew, 1998). Similarly, choice modeling has the potential to elicit the relative values that individuals place on different ecosystem services. A study of wetland restoration in southern Sweden revealed through

choice experiments that individuals' willingness to pay for the restoration increased if the result enhanced overall biodiversity but decreased if the restored wetlands were used mainly for the introduction of Swedish crayfish for recreational fishing (Carlsson *et al.*, 2003).

However, as emphasized by the National Research Council (2005), to implement a stated-preference study, two key conditions are necessary:

1. the information must be available to describe the change in an ecosystem in terms of the goods and services that people care about, in order to place a value on those goods and services and
2. the ecosystem change must be explained in the survey instrument in a manner that people will understand and not reject the valuation scenario.

For many of the regulatory services listed in **Table 2**, one or both of these conditions may not hold. For instance, it has proven very difficult to describe accurately through the hypothetical scenarios required by stated-preference surveys how changes in ecosystem processes and components affect ecosystem regulatory functions and thus the specific benefits arising from these functions that individuals value. If there is considerable scientific uncertainty surrounding these linkages, then not only is it difficult to construct such hypothetical scenarios but also any responses elicited from individuals from stated-preference surveys are likely to yield inaccurate measures of their willingness to pay for ecological services (Bateman *et al.*, 2009). Valuation workshop methods may, however, help in terms of conveying information about complex ecological goods and investigating the effects on people's values of scientific uncertainty about linkages within the system (Christie *et al.*, 2006).

In contrast to stated-preference methods, the advantage of PF approaches is that they depend on only the first condition, and not both conditions. That is, for those ecological regulatory functions where there is sufficient scientific knowledge of how these functions link to specific ecological services that support or protect economic activities, then it may be possible to use the PF approach to value these services. The basic modeling approach underlying PF methods, also called "valuing the environment as input", is similar to determining the additional value of a change in the supply of any factor input (Barbier, 2007; Freeman, 2003, Chapter 9; McConnell and Bockstael, 2005). If changes in the structure and functions of ecosystems affect the marketed production activities of an economy, then the effects of these changes will be transmitted to individuals through the price system via changes in the costs and prices of the final good and services. This means that any resulting improvements in the resource base or environmental quality as a result of enhanced regulating services lower costs and prices and increase the quantities of marketed goods, leading to increases in consumers' and perhaps producers' surpluses (Freeman, 2003, Chapter 9).

An adaptation of the PF methodology is required in the case where ecological regulatory functions have a protective value, through various ecological services such as storm protection, flood mitigation, prevention of erosion and siltation, pollution control, and maintenance of beneficial species (Barbier, 2007; McConnell and Bockstael, 2005). In such cases, the environment may be thought of producing a nonmarketed service, such as "protection" of economic activity, property, and even human lives, which benefits individuals through limiting damages. Applying PF approaches requires modeling the "production" of this protection service and estimating its value as an environmental input in terms of the expected damages avoided by individuals. However, PF methods have their own measurement issues and limitations when they are used to value ecosystem goods and services.

For instance, applying the PF method raises questions about how changes in the ecological service should be measured, whether market distortions in the final goods market are significant, and whether current changes in ecological services may affect future productivity through biological "stock effects". A common approach in the literature is to assume that an estimate of ecosystem area may be included in the "PF" of marketed output as a proxy for the ecological service input. For example, this is the standard approach adopted in coastal habitat-fishery PF models, as allowing wetland area to be a determinant of fish catch is thought by economists and ecologists to proxy some element of the productivity contribution of this important habitat function (Barbier, 2007; Freeman, 2003, Chapter 9; McConnell and Bockstael, 2005). In addition, market conditions and regulatory policies for the marketed output will influence the values imputed to the environmental input. For instance, in the previous example of coastal wetlands supporting an off-shore fishery, the fishery may be subject to open access conditions. Under these conditions, profits in the fishery would be dissipated, and price would be equated to average and not marginal costs. As a consequence, producer values are zero and only consumer values determine the value of increased wetland area. Finally, a further measurement issue arises in the case where the ecological service supports a natural resource system, such as a fishery, forestry, or a wildlife population, which is then harvested or exploited through economic activity. In such cases, the key issue is whether or not the effects on the natural resource stock or biological population of changes in the ecological service are sufficiently large that these stock effects need to be modeled explicitly. In the PF valuation literature, approaches that ignore stock effects are referred to as "static models" of environmental change on a natural resource production system, whereas approaches that take into account the intertemporal stock effects of the environmental change are referred to as "dynamic models" (Barbier, 2007; Freeman, 2003, Chapter 9).

Finally, measurement issues, data availability, and other limitations can prevent the application of standard nonmarket valuation methods to many ecosystem regulating services. In circumstances where an ecological service is unique to a specific ecosystem and is difficult to value, then economists have sometimes resorted to using the cost of replacing the service or treating the damages arising from the loss of the service as a valuation approach. For example, a number of studies that have attempted to value the storm prevention and flood mitigation services of the "natural" storm barrier function of mangroves, salt marsh, and estuarine ecosystems have used the replacement cost method by simply estimating the costs of replacing these systems by constructing physical barriers to perform the same services.

The replacement cost method can provide a reliable valuation estimation for an ecological service, but only if the following conditions are met: (1) the alternative considered provides the same services; (2) the alternative compared for cost comparison should be the least-cost alternative; and (3) there should be substantial evidence that the service would be demanded by society if it were provided by that least-cost alternative. Unfortunately, very few replacement cost studies of regulating services meet all three conditions (Barbier, 2007; Freeman, 2003; Hanley and Barbier, 2009; McConnell and Bockstael, 2005). One study that does conform to these criteria estimates the value of using wetlands for abatement of agricultural nitrogen load on the Baltic Sea coast of Sweden (Byström, 2000). The replacement value of the wetlands is defined and estimated as the difference between two cost-effective reductions of agricultural nitrogen pollution: one that uses wetlands for nitrogen abatement and one that does not. The results show that the use of wetlands as nitrogen sinks can reduce by 30% the total costs of abating nitrogen pollution from agriculture in Sweden.

However, economists consider that the replacement cost approach should be used with caution, especially as it has a tendency to overestimate values. A comparison of using an expected damage function approach and replacement cost method of estimating the welfare impacts of a loss of the storm protection service due to mangrove deforestation in Thailand confirms that the latter method tends to produce extremely high estimates – almost 4 times greater than even the largest upper-bound estimate of the expected damage function approach (Barbier, 2007). But the expected damage function has its own limitations, especially when households are risk averse, and under such circumstances, can be a poor proxy for the ex ante willingness to pay to reduce or avoid the risk from storm damages (Freeman, 2003).

Another measurement issue for economic valuation is assessing the multiple benefits that arise from large-scale habitats and ecosystems. One approach is simply to account for as many of the important ecosystem values necessary for determining conservation versus development tradeoffs. This approach was used to estimate the benefits arising from a wide range of ecosystem services provided by the Peconic Estuary in Long Island, New York (Johnston *et al.*, 2002). The tidal mud flats, salt marshes, and seagrass (eelgrass) beds of the estuary support the shellfish and demersal fisheries. In addition, bird watching and waterfowl hunting are popular activities. Incorporating PF methods, the authors simulate the biological and food web interactions of the ecosystems to assess the marginal value per acre in terms of gains in commercial value for fish and shellfish, bird watching, and waterfowl hunting. The aggregate annual benefits are estimated to be $67 per acre for intertidal mud flats, $338 for salt marsh, and $1065 for seagrass across the estuary system. Using these estimates, the authors calculate the asset value per acre of protecting existing habits to be $12,412 per acre for seagrass, $4291 for salt marsh, and $786 for mud flats; in comparison, the asset value of restored habitats is $9996 per acre for seagrass, $3454 for marsh, and $626 for mud flats.

A second approach is to develop integrated ecological-economic modeling to determine the complex ecological production underlying multiple ecosystem services, including regulating services (Barbier, 2011; Nelson *et al.*, 2009; Sanchirico and Springborn, 2011; Tschirhart, 2009). Such an approach has proven useful when one or more wetland habitats contribute to a range of goods and services. This has been especially appropriate for mapping the life-cycle of fish through coral reef–mangrove–seagrass systems to determine the contribution of each habitat to the biological growth and productivity of marine fisheries and to the protection of coastal properties from storm damages (Sanchirico and Springborn, 2011).

Integrated ecological-economic modeling has been combined with economic valuation methods to analyze land-use options in the Willamette Basin, which covers nearly 30,000 km^2 in central Oregon (Nelson *et al.*, 2009). The analysis shows how three different land-use scenarios for the Willamette Basin affect hydrological services (water quality and storm peak mitigation), soil conservation, carbon sequestration, biodiversity conservation, and the value of several marketed commodities (agricultural crop products, timber harvest, and rural residential housing). Both the baseline and development scenarios generate higher net present value (NPV) over 1990–2050 in terms of rural residential housing, timber production, and agricultural crops ($15.26 billion and $15.27 billion, respectively) than the conservation scenario ($14.78 billion). However, by 2050, the latter scenario clearly conserves more biodiversity, and thus provides more ecosystem services, than the other two landscape scenarios. The NPV of carbon sequestration from 1990 to 2050 is especially large under the conservation scenario ($1.6 billion) compared with the baseline and development scenarios ($0.9 billion and $0.8 billion, respectively). In fact, if these carbon sequestration values are added to the NPV of marketed commodities, then this scenario generates the most monetary value over 1990–2050. The overall conclusion of the study is that if markets for carbon sequestration could be established in the region, payments for sequestered carbon may make it more profitable for landowners to choose conservation over development.

An Example from Thailand

Overcoming measurement issues is critical as the most important role for valuing nonmarket regulatory services is to inform conservation versus development choices concerning ecosystem management. Such valuation is essential for public policy, because often, the *private benefits* arising from such choices are very different from the *public benefits*. Valuation, then, can show this difference explicitly, thus aiding the correct policy decision with regard to developing or conserving ecosystems. In some cases, valuing nonmarket regulatory services can be critical to this decision, as the failure to value these services properly or to ignore them can bias the policy choice significantly. An example of economic valuation applied to mangrove land use in Thailand shows the importance of this role (Barbier, 2007).

Since 1961, Thailand has lost from 1500 to 2000 km^2 of coastal mangroves or about 50–60% of the original area. Over 1975–1996, 50–65% of Thailand's mangroves were lost to shrimp farm conversion alone. Although declining in recent

years, conversion of mangroves to shrimp farm ponds and other commercial coastal developments continues to be a major threat to Thailand's remaining mangrove areas. Thus, the choice between conserving versus developing mangroves is an important land-use policy decision in Thailand as well as many other tropical countries.

As illustrated in **Figure 3**, when this decision is made on the basis of the private benefits of commercial shrimp farm areas compared with retaining mangroves, it looks as if the aquaculture option should be preferred. For a typical shrimp farm, the present value of commercial profits is around \$9600 ha^{-1}), whereas the value to local communities of exploiting mangroves directly for a variety of products, such as fuelwood, timber, raw materials, honey and resins, and crabs and shellfish, is only about \$580 ha^{-1}. But this comparison of private benefits is misleading. Many of the conventional inputs used in shrimp pond operations are subsidized, below border-equivalent prices, thus increasing artificially the private returns to shrimp farming. A further problem is that intensive shrimp farming usually lasts for only a short period of time, usually 5 years or even less. After this period, there tend to be problems of drastic yield decline and disease; shrimp farmers then usually abandon their ponds and find a new location. Once the generous subsidies to shrimp farming are accounted for, the actual value of the private benefits of aquaculture, when discounted over the 5-year period of normal operation, amounts to \$1220 ha^{-1}.

However, during their operation, intensive shrimp farming generates substantial water pollution, as the ponds are often flushed with water that is then discharged into the surrounding environment. The cost of this pollution is estimated to be around \$1000 ha^{-1} in present value terms. There is also the problem of the highly degraded state of abandoned shrimp ponds after the 5-year period of their productive life. Across Thailand, those areas with abandoned shrimp ponds degenerate rapidly into wasteland, since the soil becomes very acidic, compacted, and too poor in quality to be used for any other productive use, such as agriculture. Rehabilitation of the abandoned shrimp farm site would require treating and detoxifying the soil, replanting mangrove forests, and maintaining and protecting mangrove seedlings for several years. These restoration costs are considerable, and are estimated to be \$9318 ha^{-1}. As indicated in **Figure 3**, when pollution and restoration costs are included, the *net public costs* of shrimp farming are about \$9100 ha^{-1}.

But there are also additional *public benefits* from mangroves beyond the direct private benefits to local communities from using the resources. Mangroves serve as nursery and breeding habitats for many species of fish that are important to offshore fisheries. Mangroves also play an important role as natural "storm barriers" to periodic coastal storm events, such as wind storms, tsunamis, storm surges, and typhoons. The present value of mangroves as habitat in support of fisheries is \$987 ha^{-1}, and that for storm protection is \$10,821 ha^{-1}. After adding these two public benefits of mangroves, their value increases to nearly \$12,400 ha^{-1}.

The Thailand mangrove case study illustrates how basing a land-use decision solely on comparing the *private benefits* of

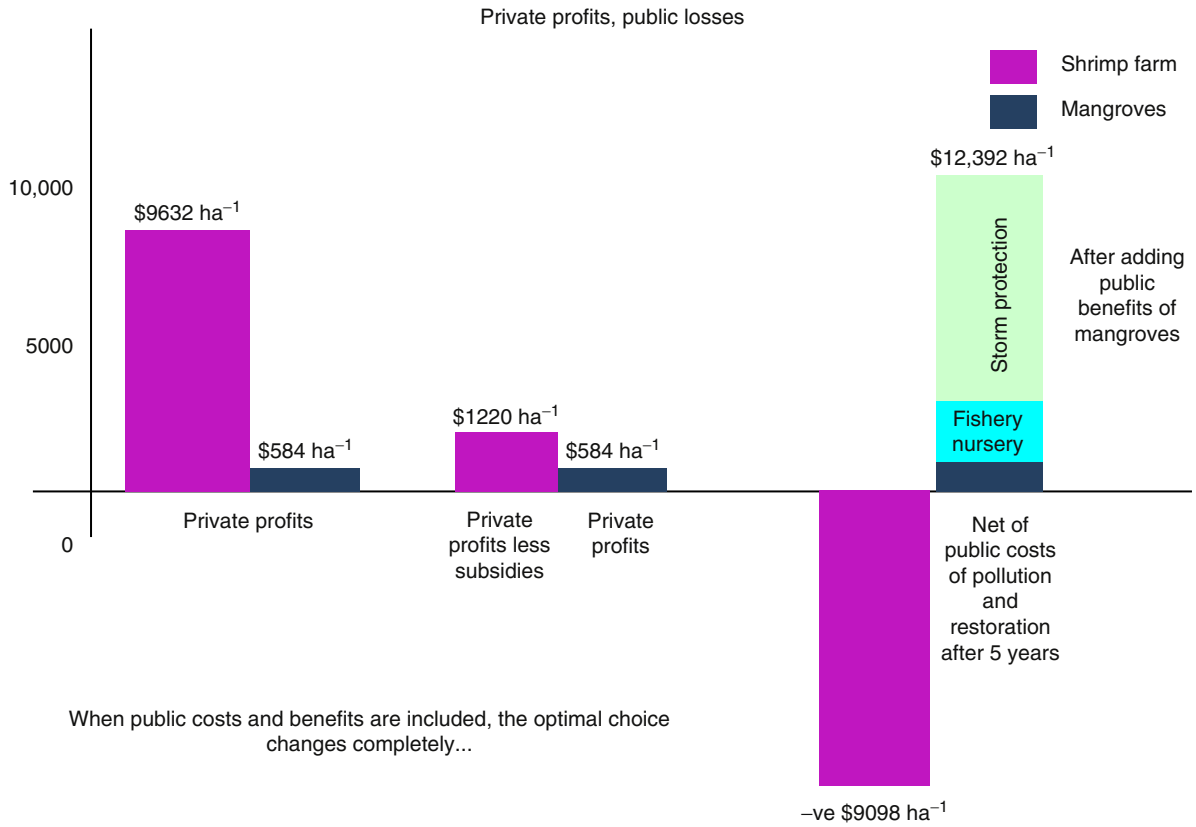


Figure 3 Shrimp farming versus mangrove conservation, Thailand.

conservation versus development options is misleading. The irreversible conversion of mangroves for aquaculture results in the loss of ecological services that generate significantly large economic benefits. In fact, the largest of these public benefits is from a regulating service, the storm protection benefits of mangroves. Any loss of this and other ecosystem services should be taken into account in land-use decisions that lead to the widespread conversion of mangroves, but typically are ignored in private sector calculations. The high restoration costs also reflect the fact that "reversing" mangrove conversion is difficult, and should not always be considered *ex post*. Instead, before the decision to allow shrimp farming to take place, the restoration costs could be treated as one measure of the public costs of converting mangroves irreversibly, and this cost should be deducted from the estimation of the net returns to shrimp aquaculture. Other public costs that occur during the operation of the shrimp farm, such as pollution, should also be included. As the example shows, shrimp farming actually can lead to a large net public cost, whereas conserving the mangroves instead yields a substantial public benefit.

Spatially Variable Ecosystem Services

A common assumption is often made that the "value" of an ecosystem regulating service changes "linearly" with critical habitat variables, such as size (e.g., area). One reason for invoking such an assumption is that little data exist for examining the marginal losses associated with changes in nonlinear ecological functions, making it difficult to value accurately the changes in ecosystem services in response to incremental changes in habitat characteristics (e.g., area). Thus, a "point estimate" for the value of an ecosystem service, in terms of benefits per hectare (ha), is simply multiplied by the total land area of an ecosystem to obtain the value of the service provided by the entire system. If, however, the relationships between ecosystem structure and function are nonlinear, then assuming that the value of an ecosystem service varies linearly with respect to changes in habitat or ecosystem area will mislead management decisions.

The potential problem can be illustrated by revisiting the case study of Thailand mangrove use (*see* An Example from Thailand *and* Figure 3) (Barbier *et al.*, 2008). As indicated previously, in this case study, the present value of mangroves for storm protection is estimated at \$10,821 ha^{-1}. The present value of mangroves as habitat in support of fisheries is much lower, at \$987 ha^{-1}, whereas the value to local communities of exploiting mangroves directly for a variety of products, such as fuelwood, timber, raw materials, honey and resins, and crabs and shellfish, is only about \$580 ha^{-1}.

But what if the per hectare values in **Figure 3** were used as the basis for a land-use decision to convert a mangrove ecosystem, such as a forest extending 1000 m seaward along a 10 km coastline? The decision as to how much of this forest to convert to shrimp aquaculture may depend critically on whether or not all the mangroves in the 10 km^2 ecosystem are equally beneficial in terms of coastal storm protection.

For example, suppose it is assumed that the annual per ha values for the various ecosystem benefits are "uniform" across the entire 10 km^2 land area. Consequently, these benefits are essentially "linear" with respect to the area of mangroves, that is, to derive the total benefits of an area of mangrove, we simply multiply the "point estimate", or per unit value, of each benefit times the total number of hectares of a mangrove ecosystem. As a result of this assumption, it does not matter if we are deciding to convert one hectare of mangroves or an entire landscape of 10 km^2. The outcome would be the same. As shown in **Figure 3**, since the total of ecosystem benefits (\$12,392 ha^{-1}) exceeds even the private profits of shrimp farming (\$9632 ha^{-1}), then the decision would always be to preserve mangroves.

However, the assumption that the storm protection service is uniform across a mangrove landscape is unlikely to be correct. Not all mangroves along a coastline are equally effective in storm protection. It follows that the storm protection value is unlikely to be uniform across all mangroves either. The reason is that the storm protection "service" provided by mangroves depends on their critical ecological function in terms of "attenuation" of storm waves. That is, the ecological damages arising from tropical storms arise mostly from the large wave surges associated with these storms. Ecological and hydrological field studies show that, where mangroves are effective as "natural barriers", against storms that generate waves less than 6 m in height, the wave height of a storm decreases quadratically for each 100 m that a mangrove forest extends out to sea. As a study shows, if the nonlinear wave attenuation function for mangroves is used to revise the estimate of storm protection service value for the Thailand case study, then a different land-use outcome emerges (Barbier *et al.*, 2008). Because small losses in mangroves will not cause the economic benefits of storm buffering by mangroves to fall precipitously, the correct land-use decision now is to allow up to 2 km^2 of mangroves to be converted to shrimp aquaculture and to preserve the remainder of the ecosystem.

The spatial variability of the ecological functions underlying ecosystem regulating services can therefore have an important bearing on the valuation of these services and land-use decisions. A number of ecological functions appear to have this property. For instance, wave attenuation appears to be nonlinear not just for mangroves but for a wide variety of interface coastal habitat, including salt marshes, seagrass beds, nearshore coral reefs, and sand dunes (Barbier *et al.*, 2011). The assumption that underlying ecological functions are spatially uniform is likely to be inaccurate for many other ecosystem regulating services that depend on habitat size, a result that could have important implications for both the valuation of these services across landscapes and the land-use decisions based on such valuations (Barbier, 2011; Brauman *et al.*, 2007; Polasky and Segerson, 2009).

Conclusion

The rapid decline of ecosystems globally has focused attention on the resulting loss in important regulating services (see **Table 1**). However, our current ecological and economic knowledge of how ecosystems produce these benefits and what values they generate are hampering efforts to improve ecosystem management.

In recent years, considerable progress has been made by economists, ecologists, and other natural scientists in determining the values of some key regulating services. However, there are still many nonmarket services that are insufficiently studied. The difficulty of determining how changes in ecosystem structure, functions, and processes influence the quantities and qualities of ecosystem regulating service flows has prevented further progress in the application of standard nonmarket valuation methods. Overcoming these measurement issues is critical if valuing nonmarket regulating services is to inform policy choices concerning ecosystem conservation versus development.

Integrated ecological-economic modeling has been combined with economic valuation methods to analyze a wide range of ecosystem goods and services, including regulating services. Such methods hold promise for future analysis. But such modeling is not a substitute for understanding better how regulating services arise through the structure and functioning of ecosystems. For example, a common assumption is often made that the value of an ecosystem regulating service is uniform across an ecological landscape. But as the case study from Thailand shows, if wave attenuation by mangroves declines from the seaward edge of a mangrove ecosystem to locations further inland, then the value of the regulating service for the entire landscape will also vary. This can have implications for both the decision as to how much mangroves to convert to alternative economic activities or where to locate such development.

Clearly, more progress is needed in determining how the structure and functioning of ecosystems influence the value of regulating services. In the past, because of the complexity of this task, such services were either improperly valued or simply ignored. What is clear from the limited studies conducted to date is that these services can often prove to be highly beneficial, and to disregard their value can distort important ecosystem conservation and development decisions. As the global decline of ecosystem regulatory services suggests, the failure to value these services properly may be the biggest obstacle to sound management of our remaining ecosystems.

See also: Ecosystem Function Measurement, Terrestrial Communities. Functional Diversity Measures. Priority Setting for Biodiversity and Ecosystem Services

References

Barbier EB (2007) Valuing ecosystem services as productive inputs. *Economic Policy* 22: 177–229.

Barbier EB (2011) *Capitalizing on Nature: Ecosystems as Natural Assets*. Cambridge: Cambridge University Press.

Barbier EB, Baumgärtner S, Chopra K, *et al.* (2009) The valuation of ecosystem services. In: Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing*, ch. 3.18, pp. 248–262. Oxford: Oxford University Press.

Barbier EB, Hacker SD, Kennedy C, Koch EW, Stier AC, and Silliman BR (2011) The value of estuarine and coastal ecosystem services. *Ecological Monographs* 81: 169–183.

Barbier EB, Koch EW, Silliman BR, *et al.* (2008) Coastal ecosystem-based management with nonlinear ecological functions and values. *Science* 319: 321–323.

Bateman IJ, Day BH, Jones AP, and Jude S (2009) Reducing gain-loss asymmetry: A virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management* 58: 106–118.

Bateman IJ, Mace GM, Fezzi C, Atkinson G, and Turner RK (2011) Economic analysis for ecosystem service assessments. *Environmental and Resource Economics* 48: 177–218.

Brauman KA, Daily GC, Duarte TK, and Mooney HA (2007) The nature and value of ecosystem services: An overview highlighting hydrologic services. *Annual Review of Environment and Resources* 32: 67–98.

Byström O (2000) The replacement value of wetlands in Sweden. *Environmental and Resource Economics* 16: 347–362.

Carlsson F, Frykblom P, and Lilijenstolpe C (2003) Valuing wetland attributes: An application of choice experiments. *Ecological Economics* 47: 95–103.

Christie M, Hanley N, Warren J, Murphy K, Wright R, and Hyde T (2006) Valuing the diversity of biodiversity. *Ecological Economics* 58: 304–317.

Duffy JE (2009) Why biodiversity is important to the functioning of real-world ecosystems. *Frontiers in Ecology and Environment* 7: 437–444.

Duffy JE, Cardinale BJ, France KE, McIntyre PB, Thébault E, and Loreau M (2007) The functional role of biodiversity in ecosystems: Incorporating trophic complexity. *Ecology Letters* 10: 522–538.

Environmental Protection Agency (2009) Valuing the protection of ecological systems and services. *A Report of the EPA Science Advisory Board*. Washington, DC: Environmental Protection Agency.

Freeman AMIII (2003) The Measurement of Environmental and Resource Values: Theory and Methods, 2nd edn. Washington, DC: Resources for the Future.

Hanley N and Barbier EB (2009) *Pricing Nature: Cost–Benefit Analysis and Environmental Policy*. London: Edward Elgar.

Johnston RJ, Grigalunas TA, Opaluch JJ, Mazzotta M, and Diamantedes J (2002) Valuing estuarine resource services using economic and ecological models: The Peconic estuary system. *Coastal Management* 30: 47–65.

Levin SA and Lubchenco J (2008) Resilience, robustness, and marine ecosystem-based management. *BioScience* 58: 27–32.

McConnell KE and Bockstael NE (2005) Valuing the environment as a factor of production. In: Mäler K-G and Vincent JR (eds.) *Handbook of Environmental Economics*, vol. 2, ch. 14, pp. 621–669. Amsterdam: Elsevier.

Mendelsohn R and Olmstead S (2009) The economic valuation of environmental amenities and disamenities: Methods and applications. *Annual Review of Environment and Resources* 34: 325–347.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press.

Naeem S, Bunker DE, Hector A, Loreau M, and Perrings C (eds.) (2009) *Biodiversity, Ecosystem Functioning, and Human Wellbeing*. Oxford: Oxford University Press.

National Research Council (2005) *Valuing Ecosystem Services: Toward Better Environmental Decision Making*. Washington, DC: The National Academies Press.

Naylor R and Drew M (1998) Valuing mangrove resources in Kosrae, Micronesia. *Environment and Development Economics* 3: 471–490.

Nelson E, Mendoza G, Regetz J, *et al.* (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and Environment* 7: 4–11.

Polasky S and Segerson K (2009) Integrating ecology and economics in the study of ecosystem services: Some lessons learned. *Annual Review of Resource Economics* 1: 409–434.

Sanchirico JN and Springborn M (2011) How to get there from here: Ecological and economic dynamics of ecosystem service provisions. *Environmental and Resource Economics* 48: 243–267.

Tschirhart J (2009) Integrated ecological-economic models. *Annual Review of Resource Economics* 1: 381–407.

Economic Value of Biodiversity, Measurements of

Robert Mendelsohn, Yale University, New Haven, CT, USA
Seth Binder, University of Minnesota, St Paul, MN, USA, and Yale University, New Haven, CT, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Robert Mendelsohn, volume 2, pp. 285–290, © 2001, Elsevier Inc.

Glossary

Consumer surplus Value to people of large amounts of goods or services minus the cost of creating them.
Marginal value Additional value of adding or subtracting one unit of a good or service.
Productive benefits Values that are bought and sold in markets such as honey and timber or which enhance goods and services sold in markets such as real estate.
Valuation Practice of measuring the economic values that people have for a good or service.
Willingness to pay Highest payment a person would make in return for a good or service.

Introduction

There has been a great deal of confusion and controversy surrounding the valuation of biodiversity. Part of the confusion lies in understanding the nature of economic value and the relationship between an individual's values and society's values. Part of the confusion lies with the different definitions of the term "biodiversity" and part of it lies with the different methods employed to value biodiversity.

Economic value is a fundamentally anthropocentric concept representing the trade-off, that the people are willing to make amongst valuable outcomes. How much income would a particular person give up to get something of value such as an additional deer or an additional pair of bald eagles nearby? Economic values reflect the benefits each person feels they receive from a specific change in the environment. Economics does not judge whether these individual benefits should be high or low. Economics simply measures the values people have.

Biodiversity benefits many people, not just a single person. The social value of biodiversity is the sum of the individual values of all the people who benefit from that biodiversity. Economic studies reveal that individual values vary a great deal across the population, especially in a heterogeneous society. There is not a single individual value for biodiversity but rather a distribution of individual values. However, the social value, the sum of all these individual values, is a single number. This makes social values controversial. Someone with high personal values for biodiversity may feel that the social value is not high enough. In contrast, someone with low personal values for biodiversity may feel that the social value is far too high. Social values are not intended to reflect a single person's values. They are supposed to reflect the values of the entire society.

Environmental protection cannot provide the exact amount of biodiversity that each person wants because only one level of biodiversity can be provided in each location. People who place a high individual value on biodiversity are often frustrated that not enough is being done, while at the same time people who place a low value on biodiversity, often feel too much is being done to protect biodiversity. This problem is not unique to biodiversity, it is a common problem of all shared or public goods. But it is one of the reasons why biodiversity values are so controversial.

Social values for biodiversity help resource managers determine what trade-offs society wants to make across the many services that are possible. For example, society must choose between having more food (farms), more timber products (managed forests), or having more wildlife (natural lands). Even within natural lands, society must choose between allocating land for species that prefer complete forest cover, a mix of forest cover and openings, or open spaces. Valuation guides how to choose across the many different possible outcomes. Valuation is a key component of resource management and it is intended to guide actual management decisions.

Note that valuation extends beyond just goods and services traded in markets. It also includes nonmarket goods and services as well. Although the prices of market goods are readily observable in most circumstances, the marginal value of nonmarket goods and services are generally more difficult to measure. In fact, some nonmarket services may not be possible to measure at all, such as nonuse values.

The term 'biodiversity' is defined differently depending on the policy issue at hand. Sometimes it is meant to reflect species richness or the number of species in a particular location. Sometimes it is meant to capture the existence of an animal that might be eliminated entirely (endangered species). Sometimes biodiversity is defined as an entire ecosystem. Biodiversity sometimes simply reflects the abundance of a particular animal even if that animal is common (starlings, rabbits, or deer). The multiple meanings of biodiversity are a source of confusion and cloud what is meant by the value of biodiversity. In this essay, the authors consequently try to be as specific as possible about what is being valued and try to avoid valuing 'biodiversity' *per se*.

It is important to understand that valuation is context-specific. There is no single value for a maple tree, a wetland, or a coral reef. The value of a particular ecosystem, species, or individual depends on the circumstances surrounding it, including its relative scarcity and the tastes, number and income level of the people placing a value on it. As natural areas or populations become scarce, each remaining area and animal

becomes more valuable. A single maple tree is going to be quite common and therefore of low value in the middle of a maple forest, but it may be very special and of high value in a city park where it is the only one. An acre of wetland may be highly valued in the middle of a dense metropolitan area where it provides scarce habitat and open space but it may have very low value in an immense, remote river delta. A Minnesota resident may value a whooping crane near their home more than a different, though possibly rarer, crane in distant Russia. A small natural park will have more value in a dense urban setting than in a rural area simply because more people are nearby to value it. Finally, richer people can afford to give up more income (have higher values) for a specific type of biodiversity than poorer people solely because they have more money.

The fact that biodiversity values are context-specific suggests to researchers that they should not be seeking a single number to measure the value of biodiversity. Valuation studies should focus on measuring specific values for biodiversity in specific circumstances as precisely as they can. This allows policymakers to evaluate the consequences of concrete management decisions.

Valuation of Productive Benefits

The marginal values of biodiversity traded in markets are readily observable as market prices. For example, crops grown on farmland, animals raised on pasture, and trees harvested for timber are all traded in markets and have readily observable prices (marginal values). The nonmarginal values of large changes in these goods can be estimated from market data as well, using consumer surplus. Consumer surplus is the area underneath the demand function minus the cost. It captures the fact that prices increase (decrease) if there is a large decrease (increase) in the quantity of the goods supplied.

Although market prices are generally widely available, there are some circumstances where they might be difficult to observe. For example, the prices of many nontimber forest products (NTFP) collected in remote locations may not be easily observed. Careful market studies of the collection of NTFPs from specific forestland can help address these shortcomings.

Many of the market values for biodiversity are actually inputs into production. Although crops, timber, and livestock are readily recognized as goods bought by consumers, the value paid at a store is much higher than the price paid to farmers and forest owners at their farm gate. The price in the store often involves processing, spoilage, transportation, and presentation. Although these costs are reflected in the final price to consumers and must be subtracted to get the price paid to the resource owner. A more obscure use of biodiversity is an input that occurs with research and development (R&D). Biodiversity is a source of breeding material and new pharmaceuticals (Simpson *et al.*, 1996). However, the process of finding new species and drugs is very expensive and involves many other inputs, so that the contribution of biodiversity is generally low. Markets place an especially low value on R&D values because any sample of a plant or animal contains all the valuable information of that species. It is also true that many species share common characteristics that are of value to R&D so that each single species has a small probability that it alone can contribute a valuable gene or compound.

The marginal value of biodiversity services that are not traded in markets are not as readily observed. For example, Yosemite National Park, a local city park, a flock of geese, old-growth (primary) forests, or the remaining spotted owls are not traded in markets. Without market trading, there is no price for these goods and services. However, that does not imply they have no value. People may value these nonmarket goods and services because they contribute to their quality of life, even if they do not contribute to the economy.

There are many methods to measure the nonmarket values of biodiversity (Bockstael *et al.*, 2000; Freeman, 2003; Haab and McConnell, 2003; Mendelsohn and Olmstead, 2009). But it is critical that users of estimated values of nonmarket goods and services understand the strengths and weaknesses of these methods as some methods are far more reliable and accurate than others.

Sometimes, the value of biodiversity can be inferred from the costs its presence avoids. For example, certain watersheds in the Catskill mountains are dedicated to providing the water supply of New York city (Chichilnisky and Heal, 1998). Buying more land for watershed protection reduces the cost of water treatment plants. The reduced water treatment cost for New York city is estimated to be worth $6–8 billion whereas the cost of setting aside the additional land is estimated to be just $1–1.5 billion (Chichilnisky and Heal, 1998).

Some species or resources have indirect values. They are not used or appreciated directly by people but they are an important component of an ecosystem that generates final ecosystem goods and services. For example, people might not actively use fungus, bacteria, and worms but they provide a valuable ecosystem function by breaking down decaying plant material. These indirect values can be estimated using a production function approach. Inputs to the ecosystem production function can be valued from their contribution to output. Of course, one cannot add these input values to output values because that would be double counting. But one can argue that without these inputs, the outputs would be reduced. Society should protect these inputs in order to protect the output.

In practice, the production function approach is sometimes misused. Studies should measure the value of output in terms of the net value that the output generates. Unfortunately, many studies rely on the gross revenues (price times quantity) of outputs and forget to subtract other costs needed to obtain the output. The most grievous example of this concerns fisheries. An efficient fishery requires substantial resources to catch the fish (boats and fishermen). These costs must be subtracted from the gross revenue of the fish caught to measure the net rent of the fishery. In the frequent case of a common property fishery, there is an excess amount of fishing effort applied in each fishery. Quite often, the cost of harvesting the fish is equal to gross revenue. There is no net revenue in this case. If the fishery were lost, the lost fish would be equal in value to the regained inputs that now could be used for some other purpose. The fishery has zero net value. Any resource required to promote that fishery such as mangroves has zero value as well.

Another critical element of valuation is whether one is valuing a small change or a very large one. Small changes tend not to affect marginal values (prices) but large changes will. For example, if one were to damage a single fishery, the global price of fish would not be affected. The loss can be measured in terms of the current price of fish. However, if one were to damage all the world's fisheries, fish would become scarce and the price or marginal value of fish would change. Large reductions (increases) in supply can lead to large increases (reductions) in price or marginal value. With large scale changes, it is necessary to take these price changes into account using consumer surplus. However, when valuing small changes, it is just as important not to include price changes (consumer surplus).

Valuation of Recreational and Aesthetic Benefits

There are also several methods for valuing the recreational benefits of biodiversity. For example, the travel cost method can measure the recreational value of specific sites. Travel cost methods also have been used to measure the recreational value of special biodiversity 'hotspots', such as the Galapagos Islands, Costa Rica's Monteverde Cloud Forest, and the Great Barrier Reef.

However, much less progress has been made measuring the value of specific characteristics of ecosystems. Resource managers of conserved sites thus have little insight into how to manage their lands to provide the greatest biodiversity benefits. Should natural lands be left untended or would these lands provide more value if protected from fire? Should land managers cut trees to create openings for wildlife or leave entirely intact forest canopies? Should the populations of certain wildlife be actively reduced or increased to favor other species?

There are several techniques that value site characteristics: discrete-choice, random utility travel cost models; hedonic travel cost methods; and stated preference methods, including choice experiments and contingent ranking (Bockstael *et al.*, 2000; Brown and Mendelsohn, 1984; Mendelsohn and Olmstead, 2009). For example, hedonic travel cost methods have been used to value old growth, forest species along backcountry trails, deer, and sport fishing populations of trout. Discrete-choice random utility models have valued the quality and number of fish species, the abundance of pheasants in wetlands, deer on agricultural lands, and the 'environmental diversity' of forested areas. Choice experiments have been used to value changes in the quality of coral reef ecosystems in the Caribbean (Hanemann, 1984).

Hedonic property studies can also capture the value of biodiversity when it provides local recreational or aesthetic benefits for which proximity is important. For example, waterfront, nearby parks, and invasive aquatic species all affect the value of nearby property. By examining how property values are affected, one can value these attributes.

Another controversial technique used in the literature is to value ecological services by the cost of substitute services. For example, a tree in New York city provides shade and therefore some cooling. The substitute technique is controversial because many studies do not use an appropriate substitute. For example, one study assumed that air conditioning is a close substitute for a tree. The tree is worth the reduced air conditioning cost. However, a more appropriate choice is a parasol. The tree is worth the cost of a parasol to people underneath the tree.

One lesson with respect to all of these valuation methods is that it is important not to double count the same value. Recreation values can be measured by several of these techniques. The recreational values also reflect aesthetic values and they are related to preferences for genetic diversity. Using multiple valuation methods to capture these values helps to confirm that the values are robust. However, adding them together may be counting the same thing twice.

Valuation of Nonuse Benefits

The most controversial set of values for biodiversity concerns nonuse values. By definition, nonuse values involve no contact between the person with the value and the environment being valued. They are strictly conceptual values. Whether the site or species continues to exist or not is really not known or directly experienced by the person. Some people argue that such 'existence values' should be included as one component of biodiversity values (Portney, 1994). Unfortunately, it is not possible to reliably measure nonuse values (Diamond and Hausman, 1994). There is no behavior to examine to see what people are actually doing to appreciate the resource since people have no contact with the resource. Stated preference methods for existence values are subjected to substantial errors and biases since people often have no experience with, or knowledge of the site or species in question. Respondents' stated values cannot grasp the scale or scope of the resource. For example, people respond with the same value to protect 1000, 10,000, or 100,000 birds in a remote flyway.

Conceptually, it is not clear that existence values can be very large. If one had to pay for the existence of each unit of biodiversity on earth, what would be the aggregate payment? Divide that payment into specific amounts for specific resources and the existence value for each resource becomes quite small. Further, people will also pay for the existence of jobs, managed or developed sites, and prosperity. The net existence value for conservation may not be as large as some people would think.

The policy and management implications of existence value are also unclear. High existence values imply that society should avoid total destruction of a species, but they may not have any relevance for most management decisions that do not permanently eliminate species.

Other Valuation Approaches

Valuation studies are expensive and so some researchers have attempted to interpolate between careful site-specific valuation studies to determine overarching values for species not included and for areas not measured. Such 'meta-analyses' are only as good as their assumptions. Interpolation can be made across a number of sites that are alike. Housing appraisers value a house not for sale by looking at the market values of

similar homes that have been sold in the market. However, if care is not taken in selecting representative sites, the studies can be biased. For example, the values of ordinary sites cannot be inferred by looking at the estimated values of the most popular destinations. Since researchers tend to value the most popular, most valuable, and most accessible sites first, there is a real danger that meta-analysis will overestimate the value of the remote and rarely visited places that have been overlooked.

Ecological models measure the productivity of ecosystems using Net Primary Production (NPP). Sites with higher NPP tend to grow faster. For example, a timber forest with high NPP also has high growth rates for timber. Farms with higher NPP produce more crops. Managed systems with higher NPP are generally more valuable. However, unmanaged natural ecosystems with higher NPP do not necessarily have more value. People rarely value natural areas by how quickly trees, plants, and animals grow. They are more interested in the state of the forest. So NPP can lead to unrealistic rankings of ecosystem values. For example, NPP rankings suggest that marine systems on continental shelves are worth more than terrestrial ecosystems, and wetlands are by far the most valuable terrestrial ecosystem. However, natural systems produce many plants and animals that have little value and may even be a nuisance. So measuring NPP without making any distinction as to whether a plant or animal is desired has little significance.

The valuation of biodiversity has come a long way over the past few decades. There are now more methods to value ecosystems than ever before and there are more places that have been valued. However, the valuation of biodiversity is not at all complete. There remains a pressing need for more valuation studies of ecosystem characteristics so resource managers can begin to manage conserved areas for their highest and best use.

See also: Biodiversity as a Commodity. Ecosystem Services. Property Rights and Biodiversity. The Value of Biodiversity

References

Bockstael NE, Freeman III AM, Kopp RJ, Portney PR, and Smith KV (2000) On measuring economic values for nature. *Environmental Science and Technology* 34: 1384–1389.

Brown Jr. G and Mendelsohn R (1984) The hedonic travel cost method. *The Review of Economics and Statistics* 66: 427–433.

Chichilnisky G and Heal G (1998) Economic returns from the biosphere. *Nature* 391: 629–630.

Diamond PA and Hausman JA (1994) Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives* 8: 45–64.

Freeman AM (2003) *The Measurement of Environmental and Resource Values: Theory and Methods*, 2nd edn. Washington, DC: Resources for the Future.

Haab TC and McConnell KE (2003) *Valuing Environmental and Natural Resources: The Econometrics of Nonmarket Valuation*. Northampton, MA: Edward Elgar Publishing, Inc.

Hanemann WM (1994) Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8: 19–43.

Mendelsohn R and Olmstead S (2009) The economic valuation of environmental amenities and disamenities: Methods and applications. *Annual Review of Environment and Resources* 34: 325–347.

Portney P (1994) The contingent valuation debate: Why economists should care. *Journal of Economic Perspectives* 8: 3–17.

Simpson RD, Sedjo RA, and Reid JW (1996) Valuing biodiversity for use in pharmaceutical research. *Journal of Political Economy* 104: 163–185.

Ecosystem, Concept of

Eugene P Odum, University of Georgia, Athens, GA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 305–310, © 2001, Elsevier Inc.

Glossary

Autotroph Literally, a self-feeder; an organism that is able to utilize inorganic carbon (carbon dioxide) as the sole carbon source for growth; for example, green plants and certain bacteria.

Black box Entity that can be examined at the system level without specifying its internal contents.

Heterotrophy Literally, a feeder on others; an organism that is dependent on organic material from an external source to provide carbon for growth; for example, vertebrates.

Industrialized agriculture Modern form of agriculture that differs from traditional agriculture in the use of elaborate and expensive machinery, the control of pests with toxic chemicals rather than biocontrols, fertilization by synthetic rather than organic products, excessive consumption of water, and farm ownership and management by corporations rather than individuals.

Input environment Collective term for all energy and materials moving into a given system.

Mega-city Modern city with a large, expanding population, characterized by high consumption levels of energy, water, and food from sources outside the city.

Output environment Collective term for all energy and materials moving out of a given system.

Techno-ecosystem Technology-based ecosystem in the contemporary world that is fundamentally distinct from natural ecosystems in the use of energy sources other than sunlight (fossil fuels, nuclear power), an urbanized concentration of human population, and the generation of substantial amounts of air and water pollutants and waste materials.

Introduction

The ecosystem is the first unit in the molecule to ecosphere hierarchy (as shown in **Figure 1**) that is complete, that is, it has all the components, biological and physical, necessary for survival. Accordingly, it is the basic unit around which to organize both theory and practice in ecology. Furthermore, as the shortcomings of the "piecemeal" short-term technologic and economic approaches to dealing with complex problems become ever more evident with each passing year, management at this level, that is, *ecosystem management*, emerges as the challenge for the future.

Since ecosystems are functionally open systems, consideration of both inputs and outputs is an important part of the concept, as shown in **Figure 2**. A diversity of species and genetic forms, together with a variety of functions and niches, are essential properties of natural ecosystems. Ecosystem diversity provides redundancy in times of environmental uncertainty.

The Ecosystem Concept

The term "ecosystem" was first proposed in 1935 by the British ecologist A. G. Tansley, but of course the concept is by no means so recent. Allusions to the idea of the unity of organisms and the environment (as well as the oneness of humans and nature) can be found as far back in written history as one might care to look. Not until the late 1800s did formal statements begin to appear, interestingly enough, in a parallel manner in the American, European, and Russian ecological literature. Thus, in 1877 Karl Mobius wrote (in German) about the community of organisms in an oyster reef as a "biocoenosis," and in 1887 S. A. Forbes, an American, wrote his classic essay on the lake as a "microcosm." The pioneering Russian V. V. Dokuchaev and his chief disciple, G. F. Morozov (who specialized in forest ecology), emphasized the concept of the "biocoenosis," a term later expanded by Russian ecologists to "geobiocoenosis" (Sukachev, 1959).

In addition to biologists, physical scientists and social scientists began to consider the idea that both nature and human societies function as systems. In 1925, the physical chemist A. J. Lotka wrote in a book entitled *Elements of Physical Biology* that the organic and inorganic worlds function as a single system to such an extent that it is impossible to understand either part without understanding the whole. It is significant that a biologist (Tansley) and a physical scientist (Lotka) independently and at about the same time came up with the idea of the ecological system. Because Tansley coined the word "ecosystem" and it caught on, he gets most of the credit, which should be shared with Lotka.

In the 1930s, social scientists developed the holistic concept of regionalism, especially Howard W. Odum, who used social indicators to compare the southern region of the United States with other regions (Odum, 1936; Odum and Moore, 1938). More recently, Machlis *et al.* (1997) and Force and Maddie (1997) have promoted the idea of the human ecosystem that combines biological ecology and social theories as a basis for practical ecosystem management. Accordingly, the concept of the ecosystem now brings together organisms, the physical environment, and humans.

As shown in **Figure 2**, a graphic model of an ecosystem can consist of a box that we can label the system, which represents the area we are interested in, and two large funnels that we can

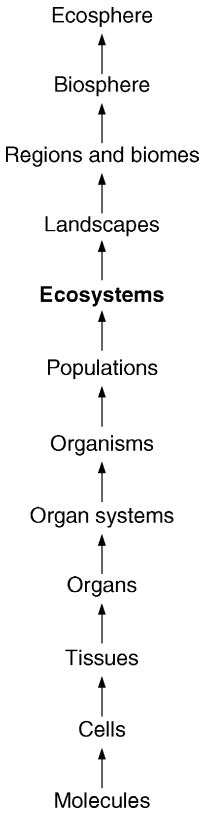


Figure 1 Hierarchical organization of living systems. The ecosystem is the first level that is complete.

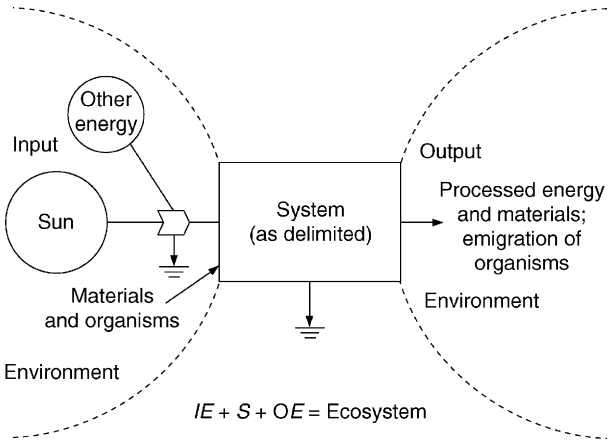


Figure 2 This ecosystem model emphasizes the external environment, which must be considered an integral part of the ecosystem concept. See **Figure 3** for a model emphasizing internal structures and processes.

label *input environment* and *output environment*. The boundary for the system can be arbitrary (whatever is convenient or of interest), delineating an area such as a block of forest or a section of beach, or it can be natural, such as the shore of a lake where the whole lake is to be the system, or ridges as boundaries of a watershed.

Energy is a necessary input. The sun is the ultimate energy source for the biosphere and directly supports most natural ecosystems within the biosphere. But there are other energy sources that may be important for many ecosystems, for example, wind, rain, water flow, or fuel (the major source for urban-industrial society). Energy also flows out of the system in the form of heat and in other transformed or processed forms such as organic matter (e.g., food and waste products) and pollutants. Water, air, and nutrients necessary for life, along with all kinds of other materials, constantly enter and leave the ecosystem. And, of course, organisms and their propagules (seeds and other reproductive stages) enter (immigrate) or leave (emigrate).

In **Figure 2** the system part of the ecosystem is shown as a "black box," which is defined by modelers as a unit whose general role or function can be evaluated without specifying its internal contents. **Figure 3** is a graphic model of the solar-powered natural ecosystem showing internal system components and functions. The interactions of the three basic components, namely, (1) the community, (2) the flow of energy, and (3) the cycling of materials, are diagrammed in this simplified compartment model. Energy flow is one-way; some of the incoming solar energy is transformed and upgraded in quality (i.e., converted into organic matter, a more upgraded form of energy than sunlight) by the community, but most of it is degraded and passes through and out of the system as low-quality heat energy (heat sink). Energy can be stored, then "fed back," or exported, as shown in the diagram, but it cannot be reused. In contrast with energy, materials, including the nutrients necessary for life (carbon, nitrogen, phosphorus, and so on), and water can be used over and over again. The efficiency of recycling and the magnitude of imports and exports of nutrients vary widely with the type of ecosystem.

Each "box" in **Figure 3** is given a distinctive shape that indicates its general function. Circles are renewable energy sources, bullets are autotrophs, hexagons are heterotrophs, and the round-bottomed shapes are storages (in this case of nutrients and organics). The community is depicted as a "food web" of autotrophs and heterotrophs linked together with appropriate energy flows, nutrient cycles, and storages.

Both graphic models (**Figures 2** and **3**) emphasize that a conceptually complete ecosystem includes inputs and outputs along with the system as delimited, that is, an Ecosystem = IE + S + OE (input environment + system + output environment). This scheme solves the problem of where to draw lines around an entity that one wishes to consider, because it does not matter very much how the box portion of the ecosystem is delimited. Often, natural boundaries, such as a lakeshore or forest edge, or political ones, such as city limits, make convenient boundaries, but limits can just as well be arbitrary so long as they can be accurately designated in a geometric sense. The box is not all there is to the ecosystem, because if the box were an impervious container, its living contents (lake or city) would die.

It is important to emphasize that it is the *diversity of ecosystem functions* including microbial recycling, inputs and outputs as well as habitats and human land uses that need to be maintained, not just the diversity of species or biodiversity in the narrow sense.

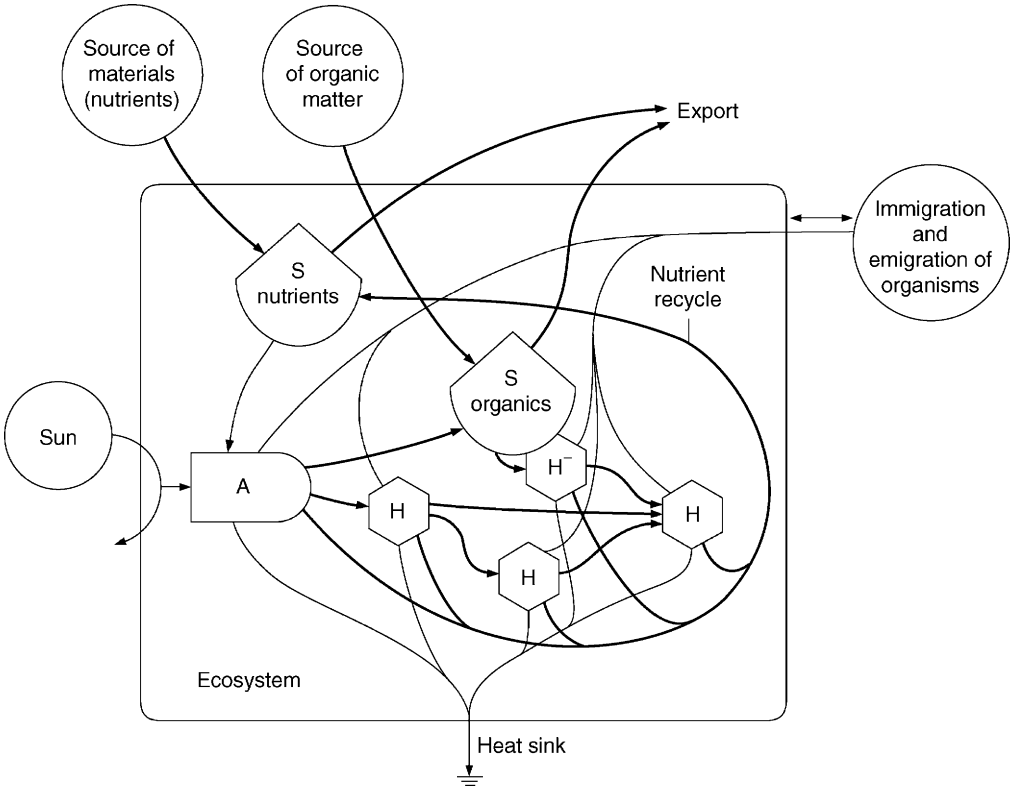


Figure 3 A functional diagram of a natural ecosystem, with emphasis on internal dynamics involving energy flow, material cycles, and storage (S), as well as food webs of autotrophs (A) and heterotrophs (H).

The Human-Dominated Techno-Ecosystem

Current urban-industrial society not only impacts natural life-support ecosystems in negative and sometimes positive ways, but has created entirely new arrangements that we can call *techno-ecosystems* that are competitive with and parasitic on natural ecosystems. These human-made systems involve new, powerful energy sources, technology, money, and fuel-powered cities that have little or no parallel in nature. It is imperative that techno-ecosystems interface with natural life-support ecosystems in a more positive or mutualistic manner than is now the case if our rapidly growing urban-industrial society is to survive in a finite world.

Before the industrial revolution, humans were a part of—rather than apart from—natural ecosystems. In the ecosystem model of **Figure 3**, humans functioned as top predators and omnivores (the terminal H box in the food web). Early agriculture, as is the case with traditional or preindustrial agriculture as still widely practiced in many parts of the world, was compatible with natural systems and often enriched the landscape in addition to providing food. But with the increasing use of fossil fuels and atomic fission—energy sources many times more powerful than sunlight—together with the mushrooming growth of cities and increasing use of money-based market economics as the basis for decision making, the model of **Figure 3** is no longer adequate. We need to create a new model for this techno-ecosystem, a term suggested by pioneer landscape ecologist Zev Naveh (1982).

Figure 4 is our graphic model for these new (in terms of human history) fuel-powered systems. It includes the four components listed earlier: powerful energy sources, technology, money, and cities. The model shows the inputs of the new fuel energy sources and natural resources, and the increasing outputs of air, water, and solid waste pollution that are very much larger and more toxic than anything that comes out of natural ecosystems. In **Figure 5** we add to the techno-ecosystem model some natural ecosystems that provide life-supporting goods and services (breathing, drinking, and eating!) and that maintain homeorhetic (i.e., pulsing) global balances in the atmosphere, soils, freshwater, and oceans. Note that money circulates as a two-way flow between society and human-made systems, but not natural systems, thereby creating a vast market failure when society fails to pay for ecosystem services.

A modern city,[1] of course, is the major component of the fabricated techno-ecosystem. It is a very energetic hot spot that requires a large area of low-energy natural and seminatural countryside to maintain it. Current cities clean and recycle no air and or water (to the point of being redrinkable), grow little or no food, and generate a huge waste stream that impacts wide areas of downstream rural landscapes and oceans. The city does export money that pays for some natural resources,

[1]The term "city" is used synonymously with the geographers' term "standard metropolitan district (SMD)," which includes industrial areas and residential suburbs that often extend far beyond official city limits.

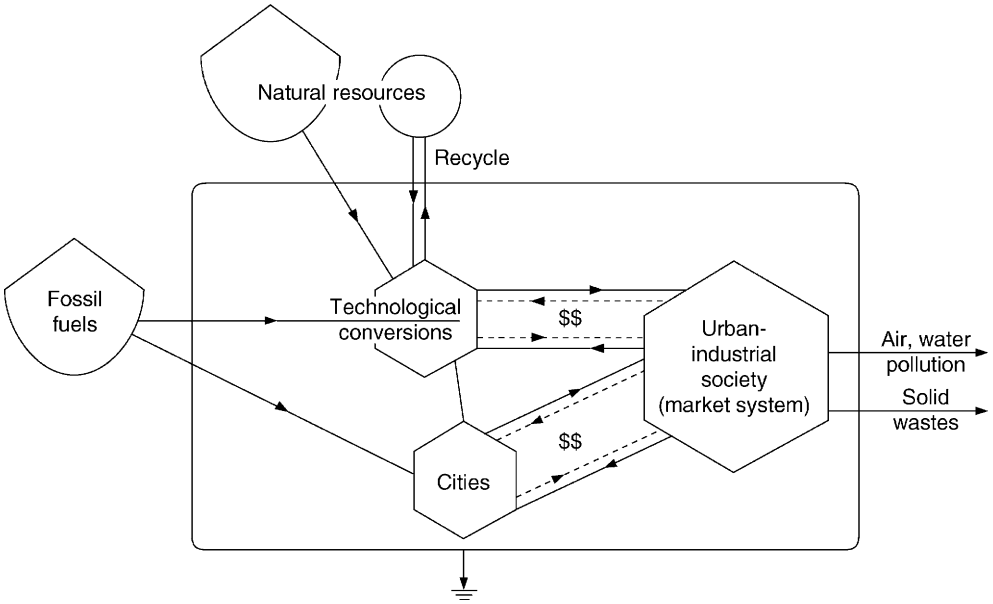


Figure 4 A human-dominated techno-ecosystem.

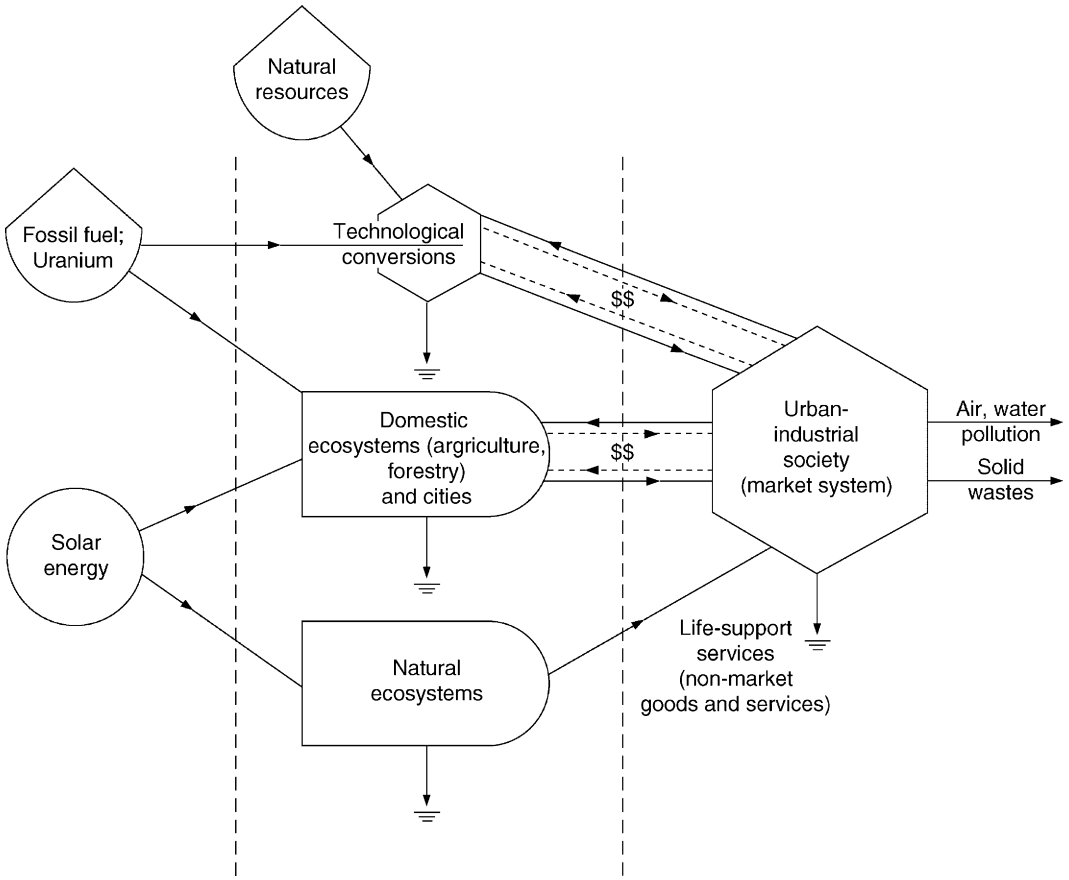


Figure 5 A human-dominated techno-ecosystem and natural ecosystem.

and the city provides many desirable cultural institutions, such as museums and symphonies, that are not available in rural areas.

In summary, cities are essentially parasites on the low-energy countryside. To call a city a parasite is not to belittle it, but to be realistic. In undisturbed nature, parasites and hosts tend to coevolve for coexistence; otherwise, if the parasite takes too much from its host, both die if the parasite has only one host. Currently humans have only one habitable host—the earth.

Especially threatening to the global life-support ecosystems is the explosive growth of *mega-cities* in the less-developed nations, caused in part by the increasing dominance of another techno-ecosystem, that of *industrialized agriculture,* with its often excessive consumption of water and use of toxic and enriching chemicals. These systems produce more food products per unit of space, but in turn are prodigious polluters and by their economic might drive small farmers out of business worldwide, forcing them into cities that are unable to assimilate them. This current situation illustrates what engineer and former president of MIT Paul Gray (1989) has written: "A paradox of our time is the mixed blessing of almost every technological development." In other words, technology has its destructive as well as beneficial side. To bring the natural and technical ecosystems into a mutualist relationship will be society's greatest challenge in the twenty-first century.

See also: Ecosystem Services. Energy Flow and Ecosystems. Human Impacts on Ecosystems: An Overview

References

Force JE and Maddie GE (1997) The human ecosystem. *Society and Natural Resources* 10: 369–382.

Gray PE (1989) The paradox of technological development. In: *Technology and the Environment*, pp. 192–205. Washington, DC: National Academy Press.

Lotka AJ (1925) *Elements of Physical Biology.* Baltimore: Williams and Wilkins.

Machlis GE, Force JE, and Birch WR (1997) The human ecosystem as an organizing concept in ecosystem management. *Society and Natural Resources* 10: 347–367.

Naveh Z (1982) Landscape ecology as an emerging branch of human ecosystem science. *Adv. Ecol. Res.* 12: 189–237.

Odum HW (1936) *Southern Regions of the United States.* Chapel Hill: University of North Carolina Press.

Odum HW and Moore HE (1938) *American Regionalism.* New York: Henry Holt.

Sukachev VN (1959) The correlation between the concepts "forest ecosystem" and "forest biogeocoenose" and their importance for classification of forests. *IX International Botanical Congress* 2: 387.

Tansley AG (1935) The use and abuse of vegetational concept terms. *Ecology* 16: 284–307.

Ecosystem Function Measurement, Aquatic and Marine Communities

John T Lehman, University of Michigan, Ann Arbor, MI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aerobic In the presence of oxygen.
Allochthonous Imported from outside the ecosystem.
Chemoautotrophy Use of energy-yielding chemical reactions as an energy source for synthesis of organic matter from inorganic precursors.
Compensation depth Depth where photosynthesis and respiration are in balance.
Euphotic zone Water depth with sufficient light for photosynthesis.
Photoautotrophy Use of light as an energy source for synthesis of organic matter from inorganic precursors.
Production Newly formed biomass of a population or trophic level, including the organic matter eliminated during the period of observation.
Trophic level Position in a food chain, defined by the number of energy transfer steps to that level.
Trophogenic region Region where net production of organic matter occurs by photoautotrophy or chemoautotrophy.
Tropholytic region Region where respiration and decomposition of organic matter proceed in the absence of primary production.

Conceptual Framework

Biological mass flux and energy transfer obey the first law of thermodynamics. Investigations that follow this theme require basic accounting for inputs and outputs balanced on carbon mass, nitrogen mass, or the chemically bound potential energy present in organic matter. The different bases for accounting are interlinked by the common stoichiometry of all protoplasm. As Alfred Redfield and his colleagues began to point out in the 1930s, there are consistent ratios among carbon, nitrogen, and phosphorus in living matter drawn from lakes and oceans. The ratios are not as rigid as the elemental composition of a crystalline mineral, but they are reliable within limits. They stem from the fact that organisms exist as biochemical aggregates of proteins, lipids, carbohydrates, and nucleic acids. The basic biochemical building blocks are universal, so the scope for differences among species is limited. If a class of organisms dominates the organic nitrogen pool within an ecosystem, it likely dominates in terms of organic carbon or calories as well. Thus the different measurement bases converge to common results.

The energy and mass-based approaches offer consistency of accounting, but they are inherently abiotic. They are indifferent to the forces that govern self-propagating organisms with heritable traits, for which the law of natural selection is as important as the laws of thermodynamics. Survival, persistence, and production of viable offspring are properties that define successful species, whether or not they dominate ecosystem biomass and material flux. Another theme in measuring ecosystem function focuses on the properties of genetic entities rather than on the properties of conservation laws. This second theme focuses on processes that add or subtract individuals or genomes within an ecosystem, processes such as birth, death, or migration, as well as the explicit match between individual age and the passage of time. Viewed through a lens that traces heritable lines of descent, mass and energy transformations are incidental consequences. Transformations of energy and material are the byproducts of life struggles whose object is indefinite persistence.

In 1961, Hutchinson coined the phrase "paradox of the plankton" to define the challenge to explanation presented by the species diversity of plankton communities. Nets towed from several hundred meters depth to the surface in the Pacific Ocean north of Hawaii, for example, routinely collect more than 300 species of zooplankton, and the richness of phytoplankton species in the surface waters is equally great. The oceanic pelagic region is the oldest continual habitat on the planet and its denizens are the product of ceaseless natural experiments. Near coastal regions and in lakes and rivers, the permanent plankton are joined by the larvae of benthic organisms. To Hutchinson, the contemporaneous existence of so many distinct genetic lines in habitats that lack obvious structural complexity begged important questions about coexistence, interactions, competition, and resource use. Faced with a bewildering array of potential species interactions, investigations of the marine pelagic ecosystems historically have emphasized mass and energy flow rather than individual species dynamics. Compounding the problem of overwhelming diversity is the fact that marine plankton are notoriously difficult to census accurately, owing to the physical movements of water masses, to such an extent that some workers believe that time series data that are essential to population studies are nearly impossible to gather.

In an effort to span the breach between measurements of organic composition in units of carbon or nitrogen and the realities of the ways matter is packaged as individual organisms, there has been great interest in finding rules of general validity and broad predictive strength that blend the approaches. Special attention has been paid to relationships based on individual body size or biomass. These size-based or allometric models of ecosystem processes offer the attraction of predicting metabolic rates and trophic interactions from simple measurements. Extensive tabulations have become available for physiological processes such as respiration, motility, body growth, and feeding rates versus individual size. This activity has spurred the drive toward models of ecosystem dynamics that treat the size structure of organisms present in a system more so than the phyletic composition of the communities.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00154-4

Biological Comparisons Between Freshwater and Marine Ecosystems

Freshwater ecosystems exhibit reduced species richness compared with marine systems. Many important invertebrate groups (e.g., Echinodermata, Ctenophora, Chaetognatha) have failed to colonize freshwater habitats, although in some cases their roles have been assumed by successful radiation of aquatic insects, particularly the Diptera, in lakes and streams. Much of the difference may be owed to the greater depth, antiquity, and continuity of oceanic plankton environments. The role of age alone is problematic, however. The great ancient lakes of the planet such as Baikal, Tanganyika, and Malawi, with basin ages measured in hundreds of thousands to a few millions of years, exhibit endemic species radiation of some groups but not all. The zooplankton faunas of the ancient lakes are notably undiversified, despite rich endemism among some fish, mollusks, or amphipods.

Not only are there differences in overall species richness, but there are some differences in latitudinal trends as well. In general, the diversity of marine plankton is greater at low latitudes than at high latitudes. The latitudinal trend in species richness for freshwater plankton is the opposite. Tropical lakes have abbreviated zooplankton faunas compared with temperate sites. The tropical lakes are depauperate of large-bodied Cladocera and copepods, but they are poor in small-bodied pelagic rotifers as well.

Primary Production

A basic unit of measurement, and also a fundamental basis for comparisons among aquatic ecosystems, is the rate of primary biological production. Several methods are used (Vollenweider, 1974; Lampert and Sommer, 2007), but they rely on the overall chemical reaction of carbon fixation in photosynthesis:

$$12H_2A + 6CO_2 \xrightarrow{\text{ln}} C_6H_{12}O_6 + 12A + 6H_2O \qquad [1]$$

where A represents an element that serves as an electron donor for the photochemical oxidation–reduction reaction. In photoautotrophic reactions of algae and bacteria in lakes and oceans, oxygen, sulfur, or reduced organic compounds are used as the electron donors. Methods for measuring primary production differ depending on whether the focus is on the transformation of inorganic carbon into organic matter, the resulting release of oxidized product, or changes in the internal cellular redox system or photochemical state. The different methods are not equivalent, and the stoichiometry of carbon fixed to oxidation product released is not strictly 1 to 2 as indicated in eqn [1]. Equation [1] is a simplification of the true synthetic reactions involved with cellular growth. Synthesis products are not only hexose sugars, but include all sorts of carbohydrates, as well as proteins, lipids, and nucleic acids. The photosynthetic quotient (PQ) of oxidized product released to carbon fixed varies according to the dominant synthesis products.

In practical application, measurement of primary production relies on one of two approaches. Either measurements are made on subsets of the natural community, which are enclosed and isolated for an experimental time duration, or the measurements are made directly on the natural community without isolation. Both approaches have advantages and potential complications.

Oxygen-Based Methods

To avoid the complexity of diffusion or advection of dissolved substances and gases in or out of a water parcel, most measurements of primary production are conducted in enclosures. During the 1920s, oceanographers introduced one such method that is still in wide use: the light and dark bottle oxygen method. The method relies on making three measurements of oxygen concentration: (1) the initial concentration of dissolved oxygen in a water sample at the start of the experiment, (2) the final concentration of oxygen in a water sample that was enclosed in a transparent bottle called the Light bottle, and (3) the final concentration of oxygen in a water sample enclosed in an opaque bottle called the Dark bottle. From these three experimental measurements it is possible to deduce rates of respiration (R), gross primary production (GPP), and net primary production (NPP):

$$R = \left([O_2]_{\text{Initial}} - [O_2]_{\text{Dark}}\right)/\Delta t \qquad [2]$$

$$\text{GPP} = \left([O_2]_{\text{Light}} - [O_2]_{\text{Dark}}\right)/\Delta t \qquad [3]$$

$$\text{NPP} = \left([O_2]_{\text{Light}} - [O_2]_{\text{Dark}}\right)/\Delta t \qquad [4]$$

where Δt is the duration of the incubation. The experimental bottles are either suspended at specified depths in the water column at the sampling site for duration of the experimental incubation, or they are placed in light and temperature conditions that simulate submarine conditions. The duration of experimental incubation Δt is invariably a compromise between the desire to obtain measurable changes in oxygen concentration and the desire to minimize artifacts resulting from prolonged exposure, such as growth of microorganisms on bottle surfaces or development of supersaturated oxygen concentrations inside the bottles. Typical incubation durations range from 4 to 24 h.

Carbon-Based Methods

Methods that record fixation of inorganic carbon into organic matter rely either on measuring the removal of inorganic carbon or the appearance of new carbon in the organic matter. The pool of inorganic carbon available for photosynthesis in aqueous solution includes not only dissolved aqueous carbon dioxide, but also the reaction products of carbon dioxide with water molecules. Carbon dioxide reacts with water to form carbonic acid:

$$CO_2 + H_2O = H_2CO_3 \qquad [5]$$

Carbonic acid is a weak acid that dissociates to form its conjugate bases bicarbonate and carbonate:

$$H_2CO_3 = HCO_3^- + H^+ \qquad [6]$$

$$HCO_3^- = CO_3^{2-} + H^+ \qquad [7]$$

At equilibrium conditions the concentration ratio of dissolved aqueous carbon dioxide to carbonic acid is about 600 to 1. Proportions of carbon dioxide to bicarbonate to carbonate vary widely, but at the environmental pH of the ocean and many lakes, bicarbonate concentrations exceed all other molecular carbon species. The pool of dissolved inorganic carbon (DIC) available for photosynthesis is defined as

$$\mathrm{DIC} = \mathrm{CO_2} + \mathrm{HCO_3}^- + \mathrm{CO_3}^{2-} \quad [8]$$

The three dominant chemical species that comprise DIC exist in known stoichiometric relationships to each other as functions of pH, temperature, and ionic strength of the aqueous solution. There are several ways to measure or calculate DIC. An aqueous sample may be acidified, transforming all DIC to carbon dioxide, and the resulting gas can be measured. Alternatively, pH and CO_2 may be measured and then the other species are calculated. Or the concentrations of the conjugate bases can be measured by acid titration, and from their total plus measured pH, the DIC is calculated.

C-14 Method

An alternative to measuring oxygen by light and dark bottle method is a method that uses radiocarbon as ^{14}C-bicarbonate as a tracer. A small but known amount of radiocarbon is introduced to a water sample and the bottle is incubated for a time period in the same manner as with the oxygen method. At termination of the incubation period, the contents of the experimental bottles are filtered. Radiocarbon retained on the filters is considered to represent inorganic carbon fixed into particulate organic matter during the experiment. Any fixed carbon that leaked into the water from algal cells can also be assessed by acidifying the filtrate under vacuum to drive residual inorganic carbon out of solution. The radioactivity of residual organic release products can then be measured. The beta particles emitted from disintegrating C-14 nuclei are detected and counted by various methods, including gas or planchet Geiger–Muller (ionization) detector, liquid scintillation, or track autoradiography.

Calculation of total carbon fixation by the C-14 method requires knowledge of the specific activity of the inorganic carbon pool and of the isotopic fractionation differences between C-12 and C-14. Specific activity is the ratio of added radioisotope activity to the mass of stable isotope that the radioisotope is intended to trace. In this method the radiocarbon traces the pool of DIC. Uptake of inorganic carbon by photosynthetic organisms involves the kinetic process of diffusion and active transport across cell membranes. In such processes there are differences in the mobility of isotopes of the same element owing to differences in molecular mass. Empirical study has demonstrated that the ratio of C-12 uptake to C-14 uptake is approximately 1.06. Accordingly, the biological fixation of inorganic carbon is calculated as

$$\begin{aligned}\text{DIC fixed} &= 1.06 \times (\text{DIC in the sampled}) \\ &\times (\text{Radiocarbon fixed}) \\ &/(\text{Total inorganic radiocarbon added})\end{aligned} \quad [9]$$

Both light and dark bottles are used in estimation of primary production by the C-14 method, but interpretation of results from the dark bottles are quite different from interpretations for the oxygen method. In the case of C-14, the dark bottle does not measure respiration, but rather it records the rate of nonphotosynthetic carboxylation reactions by algae and bacteria. In simultaneous comparisons between C-14 and oxygen methods, it appears that C-14 measures something intermediate between net and GPP. The reason for this is that some but not all of the carbon fixed during the experiment is preferentially respired during the incubation period.

Fluorescence Methods

Indirect methods for assessing primary production rely on the fact that photosystem pigments alternate between a ground state in which they are able to absorb excitation photons of characteristic wavelengths and an energized state in which they either transfer the excitation energy to a chemical reaction or fluoresce a photon at longer wavelength and return to their ground state. The proportion of pigments in one state or the other varies with the rate of primary production expressed per unit mass of pigment. The usual reference pigment is chlorophyll *a*, the antenna pigment present in all aerobic photoautotrophs.

Application of indirect approaches requires calibration against direct measurements of primary production. Their attraction is their potential to estimate primary production in natural communities without isolation.

Interpretation and Analysis of Primary Production Results

Rates of primary production vary with depth according to light intensity, temperature, algal biomass, and physiological state of the algae. The photosynthetically active spectrum includes light from 400 to 700 nm. Near surface, light intensities are often sufficient to saturate the antenna pigments with photons. At high light, rates of photosynthesis are limited by the rates of biochemical dark reactions, specifically the carboxylation reactions, rather than by interception of light photons. Because carboxylation is enzyme dependent, the maximum rates of carbon fixation are temperature dependent and they vary with nutritional status as well.

Light intensities diminish with depth. At low light, the rate of primary production is limited by the rate at which photons are intercepted, and carbon fixation varies directly with light intensity. Temperature variations are less important at low light because in photochemical reactions limited by interception of photons, the role of temperature-dependent enzyme kinetics is minimized.

Respiration rates vary with temperature. Near the surface, as a daily average, GPP often exceeds respiration and NPP is positive. At greater depth, however, GPP declines and a depth is reached at which NPP equals zero. This depth is termed the photosynthetic compensation depth. For purposes of predicting the physiological state and growth response of the algae, it is important to recognize that organisms in aqueous suspension rarely maintain fixed vertical positions in the water. What is most important is the integral difference

between the rates of GPP and of respiration during 24 h:

$$\int_0^{24} \int_0^{z_{mix}} \text{GPP}\,(z,t) - R\,(z,t)\,\text{d}z\,\text{d}t \qquad [10]$$

where z_{mix} is the maximum depth of water involved in vertical mixing during 24 h. When the integral defined by eqn [10] is positive, net growth will occur. Seasonal variations in photoperiod, water transparency, temperature, and species composition influence the components of eqn [(10)], and so the critical mixing depth for net positive algal growth varies likewise.

Secondary Production

The principle of mass and energy balance used in primary production studies can be extended to heterotrophs (Downing and Rigler, 1984). Among herbivores, predators, and detritivores, material and energy are considered commodities that continually turn over through biochemical means and by replacement of individuals within a population. Secondary production is the total growth increments over a given time interval as experienced by all individuals within a population that were alive at the beginning of the interval, whether or not they survive the complete interval. Unlike primary production, there is no distinction between gross and net production; secondary production (P) is inherently a net balance of several factors:

$$P = I - \text{Eg} - \text{Ex} - R - M - G \qquad [11]$$

where I is the ingestion, Eg is the egestion of unassimilated food, Ex is the excretion, R is the respiration, M is the molts or exuvia, and G is the gametes and reproductive products such as egg yolk and spermatophore capsules.

The efficiency of energy or material transfer during secondary production is expressed by several quotients. Assimilation efficiency is the ratio of assimilation ($I - \text{Eg}$) to ingestion (I). Net growth efficiency is the ratio of growth (P) to assimilation. Gross growth efficiency is the ratio of growth (P) to ingestion (I). Ecological efficiency is sometimes defined for predator–prey trophic interactions. Ecological efficiency is the ratio of predator production to prey production.

Bioenergetic Analysis

Bioenergetic analysis places its frame of reference on individual organisms of specified size, weight, or age. When eqn [11] is applied at the individual level, the result (P) is incremental growth. The units of accounting can be total body mass, content of carbon, nitrogen, or phosphorus, or calories. Analysis is facilitated if each of the terms on the right side of eqn [11] can be expressed with reference to body mass. For example, ingestion rate may be expressed as a function of food types and availability, temperature, and body mass of the consumer. Respiration may be expressed as a function of temperature, activity level, and body mass. In this way, the incremental growth of individual organisms is linked with organism size and a set of environmental conditions. The approach can then be extended to the population level by summing calculated growth increments across the empirical size distribution of the population.

Biomass Accrual

The standing stock biomass of a population is the sum of the weights of all individuals in the population. Changes in biomass occur owing to changes either in number (N), or mean weight (W):

$$\Delta\,\text{Biomass} = \Delta(N \times W) = N\Delta W + W\Delta N \qquad [12]$$

Equation [12] suggests a means to calculate secondary production for populations consisting of distinct cohorts. Presence of distinct cohorts is often the case among long-lived, seasonally reproducing populations. In such cases, a survivorship chart or life table for each cohort (decreasing N vs. time) defines ΔN. A parallel chart or table reporting the average individual weights of cohort members versus age or time defines ΔW.

When cohorts cannot be distinguished, as happens with continually reproducing populations, secondary production can be calculated by summing the growth increments (Δw_i) of individual life history stages:

$$P = N_1\Delta w_1/T_1 + N_2\Delta w_2/T_2 + N_3\Delta w_3/T_3 + \cdots + N_k\Delta w_k + T_k \qquad [13]$$

where the subscripts denote individual stages (e.g., instars) or size classes and where T_i represents the mean duration of each stage. Abundances at each stage (N_i) are considered to be the mean stage-specific abundances during the period of interest.

Birth and Death Rate Analysis

Numerical changes in populations can be represented as a simple function of population size (N):

$$\text{d}N/\text{d}t = r \times N \qquad [14]$$

where r is the net intrinsic population growth rate (time^{-1}) resulting from additions and removals:

$$r = \text{Birth} - \text{Death} + \text{Immigration} - \text{Emigration} \qquad [15]$$

In cases where changes owing to migration are not an issue, net population growth amounts to the balance between birthrate (b) and death rate (d):

$$r = b - d \qquad [16]$$

Over short time intervals an assumption is made that birth and death rates are stationary, so that the simple integral solution to eqn [14] is

$$N(r + \Delta t) = N(t)\text{e}^{r\Delta t} \qquad [17]$$

and net intrinsic population growth rate r can be calculated from population sizes at the beginning and end of any time interval Δt:

$$r = \ln[N(t + \Delta t)/N(t)/\Delta t \qquad [18]$$

To decompose the net rate r into its individual components, an independent estimate is needed for either birth or death rate.

Egg Ratio Method

The egg ratio method, pioneered by Edmondson (1974), calculates birthrate from demographic properties of populations in nature. The method is particularly successful in application to parthenogenic populations, especially to species in which females produce and brood clutches of eggs, such as among rotifers and Cladocera.

The egg ratio, E, is defined as the ratio of eggs and embryos to the sum of all postembryonic stages (juveniles and adults) in the population. From egg ratio and duration of development for the embryonic stage (D), the best estimate for population birthrate (b) is

$$b = \ln(E + 1)/D \quad [19]$$

Derivation of eqn [19] requires assumption of a stable age structure for the population, and further that all embryonic and postembryonic stages experience a common death rate (d), such that d can be calculated from r and b by difference:

$$d = b - r \quad [20]$$

These assumptions are most nearly applicable to shortlived species with abbreviated juvenile stages. Accuracy of the birthrate estimate is improved if the egg ratio is based on a census of egg stages that are in final phases of development and if the development time is correspondingly restricted in range. Most of the populations to which this analysis is applied are ectotherms, so egg development times are strong reciprocal functions of water temperature.

Instar Analysis

For populations that exhibit distinct morphological stages, such as size classes or the developmental instars of arthropods, it is possible to estimate stage-specific mortality. This method requires knowledge of the durations (T_i) of each stage including their variation with time and water temperature. The numerical change in each stage is defined as

$$\mathrm{d}N_i/\mathrm{d}t = N_{i-1}/T_{i-1} - N_i/T_i - \mathrm{d}_i N_i \quad [21]$$

The first term on the right side of eqn [21] represents recruitment of new individuals into the stage or age class from the preceding stage. The second term is recruitment out of the stage, and the third term is mortality within the stage. This approach assumes a uniform age distribution within each stage. It further assumes that the calculation time step is small compared to the stage durations T_i.

Instar analysis may be combined with birthrate estimation by the egg ratio method to estimate recruitment into the youngest instar category. Addition of mass or energy information about each instar or stage permits further calculation of secondary production according to eqn [13].

Functional Type Analysis

Owing to the extraordinary phylogenetic diversity of many marine plankton communities, and a desire to develop models of entire ecosystems that can help evaluate global change, an approach has emerged that simplifies the representation of community members according to their functions, such as primary producer, herbivore, carnivore, or decomposer. These approaches diverge from those that focus on genetical entities, but they bridge the biotic world with the processes discussed in section "Biogeochemistry and Nutrient Cycling." They incorporate explicit treatment of elemental stoichiometries of the organisms and allometric variation of physiological processes with organism size or individual biomass. Application to field conditions sometimes involves the use of automated image analysis to categorize organisms and to develop size spectra of community members.

Biogeochemistry and Nutrient Cycling

Transformations of matter and energy in aquatic ecosystems are dependent on, and in turn influence, the availability of all the elemental constituents of protoplasm including both major components like carbon and nitrogen as well as trace constituents such as iron and manganese. Even elements and compounds that are not essential constituents of biological material are affected by ecosystem processes. For example, redox reactions and pH changes caused by biological reactions can influence the solubility and reactivity of metals such as mercury and cadmium, which can become toxic.

Redfield is credited with the conceptual theory of "biochemical circulation" within ocean ecosystems, and the theory is equally applicable to fresh waters. According to this view, processes of synthesis and decomposition may be separated in space and time, but they remain linked. Organic matter is generated at the ocean surface where radiant energy from sunlight permits positive net primary production. Carbon is assimilated from inorganic sources into organic form, together with nitrogen, phosphorus, silicon, iron, and all the other constituents of protoplasm. Over time, a portion of this synthesized organic matter will be transported actively or passively to deeper depths in perpetual darkness where photosynthesis is impossible. Once there, the allochthonous organic matter is used in secondary production, but its ultimate fate is to be recycled to inorganic mineral form or to be buried in sediments for long periods of geological time. Organic matter that is not transferred to the depths is recycled *in situ*, owing to the production inefficiencies of multiple secondary production pathways including herbivory, predation, and detritivory.

Physical processes of upwelling and thermohaline circulation permit some mineral nutrients dissolved in deep water to be returned to the surface for another round of synthesis reactions. Over geological time scales even the elements buried in the ocean sediments can become available again through weathering of the ancient deposits. This alternation between the inorganic state of an element and a state in which it is associated with organic processes at both biological

and geological time scales is the basis for the term biogeochemistry.

For elements such as carbon and nitrogen, organic transformations entail changes in oxidation state and allied chemical properties. For others such as phosphorus and silicon, transformations affect not oxidation state but rather solubility.

Elemental nutrients that are essential for the production of organic matter can be traced through ecosystem transformations owing to the principle of conservation of matter. Quantitative illumination of the pathways has led to production of descriptive "cycles" for each element. Different aquatic ecosystems exhibit characteristic magnitudes of different pathways. Even water itself is involved in a hydrological cycle, in which water alternates among liquid, gas, and solid phases, and among ocean, atmosphere, and terrestrial environments. Water evaporates from the sea surface to the atmosphere, from where some of it falls as rain on the continents.

Surface and groundwater on the continents react with carbon dioxide from the air and from respiration of soil biota. The resulting solution thus includes carbonic acid. In some regions the carbonic acid is supplemented by strong acids introduced by anthropogenic activities. For example, combustion of fossil fuels introduces oxides of nitrogen and sulfur into the atmosphere, because the elements were components of the ancient protoplasm that underwent diagenesis to form the fuel sources. After reacting with atmospheric water vapor, the oxides enter aqueous solution as nitric and sulfuric acids. The resulting mix of weak and strong acid is the primary agent of chemical weathering, a process that dissolves the rocks of the continents. Chemical weathering accounts for most of the ions in river water.

Elements that enter the ocean by runoff from the continents or as wind-blown dust are subject to further reactions and removal processes. Biological processes are prominent among these. Microbiological processes reduce sulfate to sulfide, followed by loss of hydrogen sulfide gas or chemical precipitation of metal sulfides. Inorganic carbon is transformed into particulate organic matter or is precipitated from solution with calcium as the skeletons of marine organisms. Silica is precipitated likewise, as cell wall covering or skeletal material for a variety of plants and animals.

Inputs of new material and removal processes are in close balance, for the bulk chemical composition of the ocean appears to be in steady state. The relationship between material flux and pool sizes in biogeochemical cycles is expressed through the concept of residence time, defined as

$$\text{Residence time (yr)} = \frac{\text{Amount of a given element in the ocean}}{\text{Rate of addition or removal per year}} \quad [22]$$

Elements that exhibit great biological reactivity, such as the nutrient elements N, P, and Si, have short residence times and are usually present at low concentrations. Elements with lower significance to biological reactions such as Na and Cl have long residence times, and correspondingly account for most of the dissolved salt in sea water.

New and Regenerated Production

In the late 1960s, Dugdale and Goering introduced a conceptual model of biological production in the surface ocean euphotic zone. The model represented the euphotic zone as a compartment or box with inputs and outputs of nitrogen. The inputs were delivered via upwelling, nitrogen-fixation of dissolved diazo-nitrogen gas, and river runoff from the continents. Outputs were the sinking material consisting of dead plankton and fecal pellets. In addition to the inputs and outputs, they described internal recycling processes within the euphotic zone, through which nitrogen was regenerated by zooplankton and bacteria.

To maintain steady state, the outputs from the euphotic zone must equal the inputs, irrespective of the magnitudes of regeneration. Dugdale and Goering reasoned that primary biological production that was sustained by the inputs, or "new" nitrogen, should be called new production, and that the primary production sustained by regenerated nitrogen should be called regenerated production. "New" nitrogen typically arrives in the form of nitrate, a form that is thermodynamically favored at the redox potential of surface sea water, and is the form to which organic matter is decomposed under aerobic conditions. "Regenerated" nitrogen is typically ammonium, which is at the same reduced oxidation state as the primary amino nitrogen that comprises proteins. Most aquatic animals are ammonotelic in their excretion metabolism. Thus, the forms of nitrogen used to support primary production of organic matter are a clue to the processes that dominate production.

The concepts were further refined by Eppley and Peterson into a modern ecosystem paradigm of biological oceanography. New production is sustained by imported nutrients. It is considered to be equal to the sum of sinking flux of organic matter plus transfers of organic matter to higher trophic levels, which may be exported from the ecosystem by, for example, fisheries yields. By empirical evidence, ecosystems dominated by "new" production exhibit higher primary production overall and are more capable of sustaining exploitation such as harvest of fisheries stocks. Ecosystems dominated by "regenerated" production cannot endure prolonged exploitation of their sustaining elements.

The concept of new production and its linkage with export production has further implications for global biogeochemical cycles and specifically for carbon dioxide levels in the global atmosphere. Organic matter that is exported from the surface ocean to the deep sea represents a sink for DIC from the surface ocean. Deficits of DIC are replenished in part by dissolution of carbon dioxide from the atmosphere. As a consequence of this linkage, the fraction of global oceanic primary production that is new production represents an upper limit to the amount of carbon that can be exported and sequestered in the deep ocean.

Techniques for measuring the rates of nutrient regeneration differ depending on whether the measurements are conducted in trophogenic regions or in tropholytic regions. In tropholytic regions, primary production is minimal and thus regeneration can be assessed by measuring the accumulation of mineral nutrients directly. In trophogenic regions, synthesis reactions dominate over regeneration, and regenerated nutrients are

quickly reassimilated. Techniques for measuring nutrient recycling under such conditions rely on assessment of uptake and regeneration simultaneously, such as by isotope dilution. In isotope dilution experiments, a deliberate tracer such as ^{15}N-ammonium is introduced to a closed system in which ^{14}N is the predominant natural isotope. Over time, the ratios of ^{15}N to ^{14}N in the ammonium pool become diluted as regenerated ^{14}N-ammonium is added, even while uptake of both isotope forms proceeds.

Stable Isotopes

Stable isotopes of natural elements are ubiquitous. They have often been collected and used as deliberate tracers, such as in studies of new and regenerated production, but they can be measured and used at natural occurrence as well (Lajtha and Michener, 1994; Griffiths, 1998). Owing to the proportions in which different nuclear species survived the first instants of the universe, and the proportions that have been retained in the planetary composition of the Earth, many common elements exhibit more than one stable form. Hydrogen, carbon, nitrogen, oxygen, and sulfur, for example, all have two or more nuclear configurations that differ in numbers of neutrons and hence differ in atomic mass. Oxygen-16, O-17, and O-18 are stable (not radioactive), but they exist in different proportions in the environment. The different forms can be distinguished by mass spectrometry. The preferred form of analysis is by isotope ratio mass spectrometer, which compares the stable isotope ratios in a sample with ratios in a standard.

Measurement of stable isotope ratios in natural materials has been a technique used by geochemists for decades to help interpret complex processes. The relative proportions of the stable isotopes vary from one medium or substance to another. Mechanisms that can change the proportions of two stable isotopes of the same element are called fractionation processes. Fractionations result from differing mobilities and reactivities by the same chemical element owing to differences in atomic mass. Both kinetic and thermodynamic factors affect the magnitudes of fractionation. Five general rules apply:

1. Elements with high atomic mass generally exhibit smaller fractionation among their isotopes than do lightweight elements.
2. If multiple isotopes of an element exist, fractionation is greatest between isotopes that have the largest atomic mass difference.
3. In unidirectional reactions, the lighter isotope becomes enriched in the endpoint.
4. In reactions that reach equilibrium, the lighter isotope is enriched in the reactant compound or phase that has the weaker bond strength.
5. As temperature increases, fractionation decreases asymptotically to zero.

Changes in the fractionation ratios with temperature have been the basis for using stable isotopes in geo-chemical studies as "paleothermometers."

Atmosphere, terrestrial, and marine environments are heavily dominated by single isotope forms of common elements, and the alternate isotopes are usually very rare. For example, in the atmosphere, O-16, O-17, and O-18 co-occur in overall ratios by atoms as 1 : 0.0004 : 0.002. Similarly, C-12 and C-13 occur in atmospheric carbon dioxide in the mean ratio 1 : 0.011. Given such disparities of magnitude, isotope ratio differences between samples and standard are calculated in δ notation with units of parts per thousand, or per mille (‰). For example,

$$\delta - {}^{18}\mathrm{O} = \left[\left({}^{18}\mathrm{O}/{}^{16}\mathrm{O}\right)_{\mathrm{sample}}/ \left({}^{18}\mathrm{O}/{}^{16}\mathrm{O}\right)_{\mathrm{standard}} - 1\right] \times 1000 \quad [23]$$

Isotopes are useful in ecosystem studies because physical reactions and enzyme reactions cause reproducible discrimination between isotopes. Reaction differences lead to differences in the composition of biological material, skeletal products, and inorganic nutrient pools.

Stable isotopes have been used to infer and trace the origins and pathways of materials in aquatic ecosystems. They are useful when different sources are isotopically unique and when the sources change in predictable ways. Successful applications require that the magnitudes of all isotope fractionations that result from physical, chemical, and biological processes be known.

Isotopes of carbon and nitrogen have been used to trace food-web relations in freshwater and marine environments. Phytoplankton discriminate against C-13 during photosynthetic carbon fixation, just as they discriminate against C-14. The magnitude for discrimination against C-13 is less than for C-14, or about -20 to -30‰ with respect to the source DIC. Organic matter that originates from terrestrial sources sometimes has very different carbon isotope ratios than material that is of aquatic origin. Terrestrial plants fractionate atmospheric CO_2 differentially during photosynthesis according to whether they use C-3, C-4, or CAM pathways for carbon fixation. Similarly, organic matter generated microbiologically from methane as a substrate typically exhibits unique carbon isotope ratios. Stable isotopes may thus be used to identify chemoautotrophic pathways as well as photoautotrophic ones.

Carbon isotopic ratios in animals resemble those of their diets within about 1‰. Slight enrichment of C-13 in an animal versus its diet is presumed to be caused by preferential loss of C-12 during respiration or by preferential assimilation of C-13 from the food. Synthesis of lipid rather than protein can also influence the isotope ratios because there are reported fractionation differences between biochemical constituents. As long as the isotope ratios in different potential food sources are well defined and consistently different, the relatively conservative levels of fractionation for carbon by trophic level can indicate whether an animal is eating its food from one source, another source, or a definable mixture of the two.

Nitrogen isotope ratios also reflect the composition of an animal's diet, but the animal is usually enriched with N-15 by about $+3$‰ compared with its food. This increased retention of N-15 over N-14 seems to be favored by the thermodynamics associated with transamination reactions, and by kinetic factors leading to preferential loss of N-14 as excreted

ammonium and urea. Reproducible differences in δ–^{15}N with trophic level provide a diagnostic measure of relative trophic position among organisms that depend on the same primary source of organic matter. Carnivores are isotopically heavier than herbivores, which in turn are isotopically heavier than the algae; omnivores are isotopically intermediate between herbivores and carnivores.

Interpretation of nitrogen isotope dynamics can be complicated by temporal and spatial variations in isotope fractionation and nutrient source materials. In laboratory experiments when nutrients are supplied in great excess of consumptive demand, isotope fractionations associated with assimilation of nitrate and of ammonium can be large. However, if a nutrient element becomes limiting and all of it is consumed, there can be no isotopic fractionation. In nature, seasonal alternation between times of nutrient excess and nutrient limitation is common, with the result that isotope ratios can vary greatly in the source material and in the primary producers. Such variations provide useful information about nutrient dynamics, but simultaneously introduce complication to interpretations of trophic relations particularly for long-lived organisms.

See also: Biogeochemical Cycles. Marine Ecosystems

References

Downing JA and Rigler FH (1984) *A Manual on Methods for the Assessment of Secondary Productivity in Fresh Waters*, 2nd edn. Oxford: Blackwell.

Edmondson WT (1974) Secondary production. *Mitteilungen International Vereinigung Limnologie* 20: 229–272.

Griffiths H (1998) *Stable Isotopes*. Oxford: Bios Scientific Publishers.

Lajtha K and Michener RH (eds.) (1994) *Stable Isotopes in Ecology and Environmental Science*. Oxford: Blackwell.

Lampert W and Sommer U (2007) *Limnoecology: The Ecology of Lakes and Streams*, 2nd edn. Oxford: Oxford University Press.

Vollenweider RA (1974) *A Manual on Methods for Measuring Primary Production in Aquatic Environments*, 2nd edn. London: Blackwell.

Ecosystem Function Measurement, Terrestrial Communities

Sandra Díaz, Universidad Nacional de Córdoba – CONICET, Córdoba, Argentina

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity Number, abundance, composition, spatial distribution, and interactions of genotypes, populations, species functional types and traits, and landscape units in a given system.

Ecosystem Conceptual view of an assemblage of organisms and of physical and chemical components in their immediate environment, and the flow of materials and energy between them.

Ecosystem functioning Flow of energy and materials through the arrangement of biotic and abiotic components of an ecosystem.

Ecosystem stability Capacity of an ecosystem to persist in the same state. It has two components. Ecosystem resistance is the ability to stay in the same state in the face of perturbation. Ecosystem resilience is the ability to return to its former state following a perturbation.

Functional diversity Value, range, distribution and relative abundance of the functional traits of the organisms in a given ecosystem.

Resource dynamics Inputs, outputs, and internal cycling of key resources, such as carbon, water, and mineral nutrients, in an ecosystem.

Trophic transfer The amount of biomass and/or energy which is transferred from the primary producers to the herbivore-based food chain rather than directly from plants to detritivores.

Why Measure Ecosystem Functioning?

Ecosystem functioning involves processes such as primary production, trophic transfer from plants to animals, nutrient cycling, water dynamics, and heat transfer. This term refers to functioning in equilibrium, namely the amount (how much), the rate (how fast), and sometimes the seasonal variations of these processes. It also includes the responses of ecosystem processes to perturbation, resilience, and resistance. Ecosystem processes provide essential benefits to humankind, also called ecosystem services, such as food, fiber, fodder, fuel, water provision and detoxification, amelioration of weather, soil formation, retention of carbon that otherwise would be released into the atmosphere contributing to climate change, and medicinal, recreational, and cultural resources. As the impacts of changes in the quantity and quality of benefits provided by terrestrial ecosystems to humankind come to the focus of public attention, it is increasingly obvious that their value in the market, if any, greatly underestimates their true value to society. Therefore, measurement of an ecosystem's structural and functional properties such as how much animal biomass an ecosystem can produce or how much water, carbon, and nutrients can be retained *in situ* are key inputs for an evaluation of ecosystem services to humankind. Concepts, methods, and data retrieval on ecosystem functioning have progressed enormously in the past few decades, and today an impressive body of quantitative information on flows and stocks of materials and energy through a wide spectrum of ecosystems is available. Much less is known, however, of the effects of perturbations on these processes, in particular of thresholds beyond which the ecosystem can no longer recover its former basic properties. Physiological processes underlying biomass production, water balance, and nutrient cycling are well understood. However, there are still many theoretical and methodological difficulties and information gaps in the understanding – and thus the prediction – of processes at coarser scales of time (decades or centuries) and space (ecosystems, landscapes, and biomes). To determine how the biosphere will respond to the changes in climate, atmospheric composition, and land use projected for the future, a better understanding of ecosystem properties at these scales is vital.

Ecosystem Resource Dynamics

Three main processes that are linked to the acquisition and loss of chief resources for terrestrial ecosystems (light, carbon, water, and mineral nutrients) will be analyzed here. These processes are biomass production (primary productivity and trophic transfer), nutrient (especially nitrogen) cycling, and water dynamics. There is a strong association between water flow, carbon assimilation, and nutrient uptake by a plant. Given that carbon dioxide (CO_2) input and water output occur through open stomata, and nutrient and water uptake occur through root hairs and fine roots, they are inextricably linked. Nutrient availability controls the increase in plant mass, and incorporation of mineral nutrients into biomass is not possible without carbon assimilation. The derived processes at the ecosystem level, namely biomass production, water balance, and nutrient cycling, are therefore closely linked and basically regulated by net radiation, temperature and precipitation, soil composition and structure, and the kinds of plants that are present. Any consideration of carbon, nutrient, or water balance in isolation is probably artificial. The examples presented here are described in different sections for the sake of clarity, but they highlight the strong interdependence of these processes.

Approaches to Analysis and Measurement of Ecosystem Functioning

Whole-Ecosystem Approaches and Approaches Based on Community Structure and Composition

The measurement of ecosystem functioning focuses on the sizes of major pools of resources, such as water, carbon, and

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00042-3

mineral nutrients, and on the rates of flows connecting them. In the 1970s compartmental models become an important tool in 'systems ecology' (illustrations of this view can be found in the classic textbook on ecology by E.P. Odum, 1971). In these models, boxes represent pools or stocks (usually major trophic levels), and arrows signify the flows between them. This still represents a first and very useful approximation. However, pools are often treated as 'black boxes' in which only the general size, and not the internal composition, is important. This often makes this approach too coarse for the management of real situations and of little predictive value for novel situations. This is because in many the whole pool is not altered, but instead subtle intrapool changes occur until a threshold is reached or substances are chemically altered or sequestered because organisms that are members of the same pool (e.g., primary producers, detritivores, and primary or secondary consumers) often differ in the ways they affect, and respond to other biotic or abiotic compartments in the ecosystem. Hence other approaches have increasingly focused on the internal composition of pools (what taxa or functional groups are present within the boxes and in what abundance) and how changes in that composition can alter flow rates between pools (see for e.g., the highly innovative textbook by Chapin *et al.*, 2011).

Measurement of Short-Term Resource Dynamics and Long-Term, Large-Scale Ecosystem Processes

Some methods of ecosystem analysis emphasize the measurement of short-term process rates (e.g., photosynthesis, nutrient uptake, and evapotranspiration), whereas other methods emphasize the measurement of the size of pools accumulated as a result of these processes, with or without consideration of their internal components (e.g., biomass production, nitrogen stocks in different compartments, and water use efficiency integrated over a whole season) (**Table 1**). The choice depends mostly on the objectives pursued and on the scale of observation selected. However, as a general rule, methods based on short-term flows provide more mechanistic understanding, but they do not provide a clear picture of what happens at the time and space scales most relevant to ecosystem functioning (meters to thousands of kilometers and months to decades or centuries). Measurements of flow rates (*F* measurements in **Table 1**) tend to be very precise but highly variable. They give a snapshot of ecosystem functioning, but they do not necessarily reflect processes over a growing season or longer. For example, short-term (hourly or daily) variation in gas exchange per leaf area is not directly reflected in annual biomass production. Plant growth and vegetation productivity cannot be directly equated with leaf photosynthetic rates. The productivity of plant communities is determined more by the amount of photosynthetic tissue (which is in turn controlled by carbon allocation) than by photosynthetic rate. There is often only a slight relationship between these two parameters. Gas exchange measures the instantaneous plant performance, whereas growth necessarily integrates along time and reflects plant allocation patterns. At the community level, additional regulations occur; therefore great caution is necessary in drawing ecosystem-level conclusions on the basis of leaf-level ecophysiological processes. For example, a 10- to 40-fold difference between species in maximum rates of photosynthesis at the leaf level is reduced to only a factor of two to four at the primary production level between different forest types.

Therefore, for responses at the whole-ecosystem level, methods that focus on changes of pools (*P* measurements in **Table 1**) are usually more meaningful and convenient. This point is illustrated by **Figure 1** and by some examples of the use of different methods for assessing biomass production at different scales (*see* Biomass Production). An example is provided in **Figure 2**. A long-term study of the responses of tundra-dominant plants to different treatments showed that processes that are readily integrated at annual time steps (shoot growth, mortality, and allocation) were more useful than instantaneous physiological measurements in predicting decadal vegetation changes. Strong treatment effects on photosynthetic rate (**Figure 2(a)**) and nutrient uptake were poorly related to longer term changes in production and nutrient concentration (**Figure 2(b)** and **2(c)**). This was probably due to the operation of buffering mechanisms (e.g., allocation, nutrient relations, altered phenology, interspecific interactions, and positive and negative feedback mechanisms) that compensate for immediate physiological responses to the environment.

Some methods are increasingly used that are intermediate or bridge the gap across scales and therefore partially defy the framework given above. A prominent example is micrometeorological methods that measure the net water and carbon exchange between the canopy and the atmosphere (e.g., the eddy covariance and eddy accumulation methods). These are flow-based methods, which typically work at time scales of hours and spatial scales of a few square kilometers, and are used to integrate leaf- and plot-level fluxes with scales relevant to global climate modelers (Sala *et al.*, 2000).

Biomass Production

Net primary production (NPP) is strictly defined as the difference between the energy fixed by autotrophs and their respiration, and it is most commonly equated to increments in biomass per unit of land surface and time. Because the increment in biomass over a given time depends on the rate at which new biomass is produced and also on the initial amount of carbon-assimilating photosynthetic tissue, stands with a large standing biomass often show higher NPP than stands with lower biomass. Therefore, another useful concept is that of relative productivity rate, or the time needed by a vegetation stand to produce its standing biomass. For example, the estimated relative productivity rate for a dry tropical forest can be many years, whereas in an annual grassland it is less than 1 year.

The fate of assimilated carbon – that is, whether it is allocated to increase the pools of aboveground or belowground biomass, root exudates, litter, soil organic matter, grazers, symbionts, or parasites – varies strongly between ecosystems, depending on prevailing climatic conditions, disturbance regimes, and allocation patterns of dominant plant functional groups (**Figure 3**). It should be noted that NPP is not the same as Net Ecosystem Productivity or NEP. As stressed by

Table 1 Some examples of ecosystem-level measurements with emphasis on short-term flow rates (F) and on longer term pool sizes (P)[a]

Parameter measured	*Usual units*	*Techniques*	*Emphasis*
Biomass production			
Photosynthesis	μmol CO_2 cm^{-2} s^{-1} or mg CO_2 m^{-2} h^{-1}	Gas exchange systems (CO_2, O_2), or simultaneous H_2O and CO_2 exchange	F
Standing biomass	g m^{-2} or t ha^{-1}	Harvest at the end of growing season	P
Productivity	g m^{-2} $year^{-1}$ or t ha^{-1} $year^{-1}$	Sequential harvest; equations including climatic and soil factors; satellite imagery	P
Cover or frequency	% ground area (cover); % total interception (frequency)	Estimation through nondestructive measurement, which can be converted into units of biomass using allometric equations developed on similar neighboring individuals that are destructively harvested	P
Trophic transfer			
Consumer biomass	g or kJ m^{-2}; t ha^{-1}	Harvest (with or without replacement); equations including NPP	P
CO_2 efflux from respiration (estimates microfaunal biomass)	μmol CO_2 m^{-2} or mm^{-3} s^{-1}	Gas exchange measurements	F
Secondary productivity	g or kJ m^{-2} $year^{-1}$; t ha^{-1} $year^{-1}$; No. animals m^{-2} $year^{-1}$	Sequential harvest (with or without replacement); animal countings; equations including NPP	P
Consumption	g or kJ m^{-2} $year^{-1}$; t ha^{-1} $year^{-1}$; % plant tissue; % leaf surface	Plant harvest in herbivore-free and grazed situations; estimation of proportion of plant tissue removed; equations including NPP	P
Nutrient cycling			
Decomposition through biomass loss	% initial dw loss	Sequential weighing of litter samples	P
Decomposition through CO_2 efflux from soil or jars	μmol $CO_2 m^{-2}$ s^{-1} or μg CO_2g dw substrate^{-1} h^{-1}	Gas exchange system (soil-only control samples or isotope labeling sometimes used in order to distinguish from root respiration)	F
Nutrient status	Mg nutrients g dw^{-1} or % nutrient dw	Harvest followed by chemical analysis; estimation through remote sensing	P
Nutrient use efficiency in litter (amount of biomass produced per unit of nutrient expended and lost; it reflects initial leaf status and efficiency of nutrient resorption)	kg dw mol N^{-1} in litter	Harvest followed by chemical analysis	P
Nutrient uptake	Mg nutrient g dw^{-1} $year^{-1}$	Change in nutrient pool in live biomass over time interval; estimated through sequential harvest followed by chemical analysis of samples	P
Nutrient uptake potential	μg radioisotope g^{-1} h^{-1}; pg radioisotope mg^{-1} min^{-1}	Soil labeling with radioisotopes such as ^{32}P, ^{33}P, and ^{42}K followed by monitoring of isotope appearance in plant tissue	F
N transformations in the ecosystem (e.g., 'tightness' or 'openness' of nitrogen cycle)	$\delta^{15}N$ ($^{\circ}/_{00}$)	Ratio of stable isotopes ^{15}N to ^{14}N in plant material compared with standard (air) using mass spectrometer	P
Nitrogen in soil available to plants	$NO_3^- + NH_4^-$ (μg g^{-1}; %)	Soil sampling followed by chemical analysis; ion exchange resin procedure	P
Mineralization rate	μg N g substrate dw^{-1} day^{-1}	Tracing of ^{15}N released from control samples and from samples treated with chloroform fumigation to kill microbial cells; ion exchange resin procedures	F
Nutrient retention or loss	mg nutrient L^{-1} leachate; μg nutrient g^{-1} soil or plant sample; ^{15}N recovered (mg; %)	Measured as nutrient concentration in leachates (the higher the concentration in leachates, the lower the retention by the ecosystem) or as ^{15}N recovered in plants and soil organic matter	P
Water balance			
Water content	mg H_2O g dw^{-1} or % H_2O dw	Harvest followed by measurement of fresh and dry weights	P

(*Continued*)

Table 1 Continued

Parameter measured	*Usual units*	*Techniques*	*Emphasis*
Water use efficiency (C assimilated per unit water transpired, integrated over a long period, e.g., a whole season or longer)	μmol CO_2 mmol H_2O^{-1}; $\delta^{13}C^{\circ}/_{00}$	Ratio of stable isotopes ^{13}C to ^{12}C in plant material compared with standard (PeeDee Belemnite)	P
Evapotranspiration	mg H_2O m^{-2} min^{-1}	H_2O exchange measurements	F
Water potential (the more negative the water potential, the more negative the balance between absorption and transpiration)	Mpa	Pressure chamber measurements (pressure needed to expel a drop of sap out of the xylem)	F

[a]dw, dry weight.

Source: More details can be found in Pearcy *et al.* (1989), Ehleringer and Field (1993), Sala *et al.* (2000), Berg and McClaugherty (2000), and Hoover *et al.* (2008).

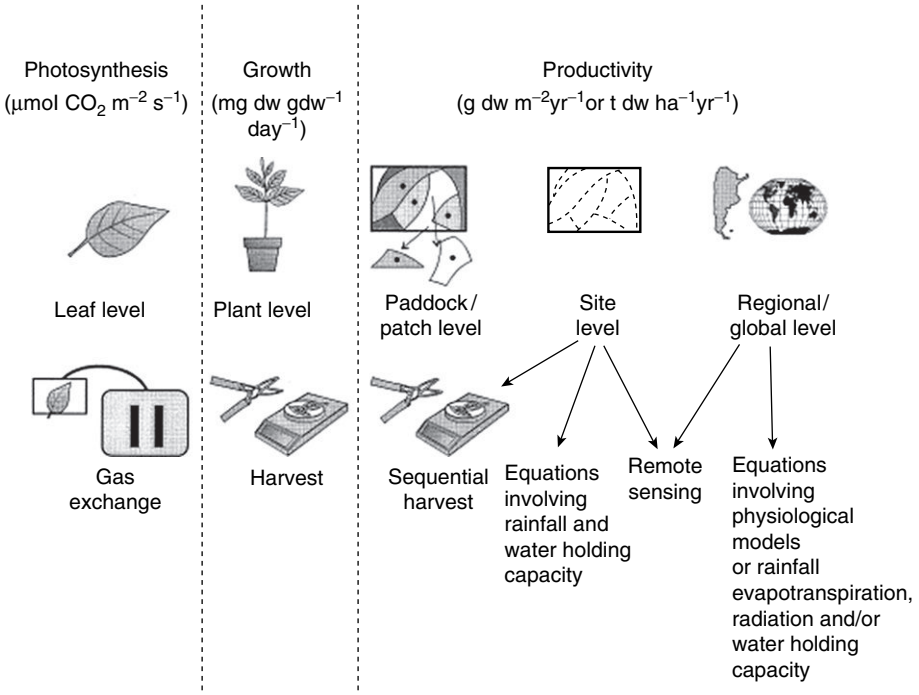


Figure 1 Carbon assimilation processes at different scales and some methods for measuring them.

Chapin *et al.* (2005), NEP is the net biomass accumulation by a whole ecosystem and depends not only on NPP, but also on carbon losses due to the respiration of animals and microbes, leaching, erosion, exportation by animals, and in some cases volatilization due to fires.

At the regional scale, NPP can be largely accounted for by climatic factors, rainfall, and temperature. For example, precipitation, potential evapotranspiration, and radiation are enough to account for the net aboveground primary production (NAPP) of North American forests, deserts, and grasslands. In regions of the US with up to 1400 mm of annual rainfall, annual precipitation is enough to account for 90% of the variability in NAPP of grasslands (**Figure 4(a)**). At higher precipitation, NAPP depends more on other factors, and equations based on annual rainfall lose part of their predictive power. At the site level, variability in production seems to be accounted for by annual precipitation and soil water-holding capacity (whc; **Figure 4(b)**). Soil whc can have a positive or negative effect depending on the precipitation value. In dry regions, major losses of soil water occur via bare soil evaporation. However, where sandy soils occur, bare soil evaporation is lower than in loamy soils because water penetrates deeper into the soil. For the same reason, surface runoff is also lower in sandy soils than in loamy soils. In more humid regions, substantial water losses occur via deep percolation, which is reduced in soils with high whc. This is known as the inverse texture hypothesis, proposed by Noy-Meir in 1973.

At finer scales of analysis (e.g., paddocks and vegetation patches), more variables are needed to account for NAPP. Species composition and land-use regime become important factors, although drivers at a coarser scale are still in operation and constrain responses (e.g., irrespective of management or species composition, annual precipitation will set an upper boundary to NAPP). For instance, the NAPP of Argentine

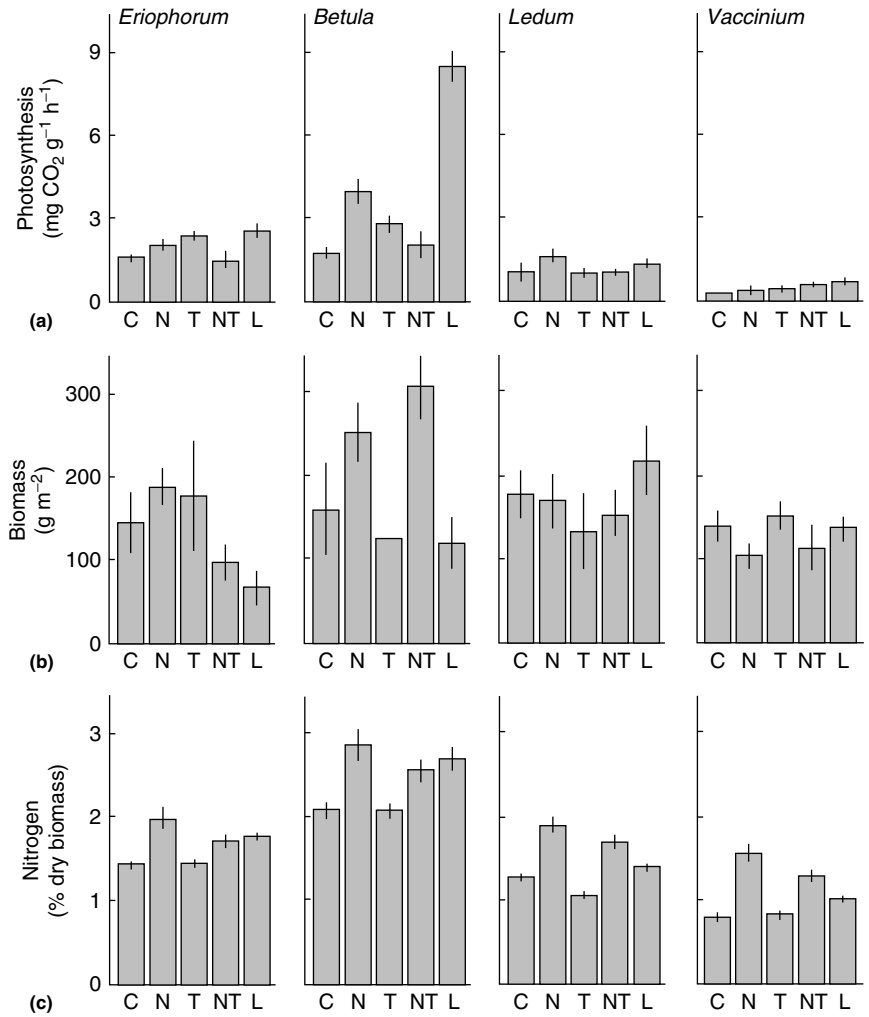


Figure 2 Effects of 3-year environmental manipulations on tussock tundra species at different scales: (a) photosynthetic rate at 20 °C and in full sun, (b) total peak-season biomass (excluding roots), and (c) leaf nitrogen content. Treatments are control (C), nutrient addition (N), temperature increase (T), nutrient × temperature (NT), and light attenuation (L). Reproduced from Chapin III FS and Shaver GR (1996) Physiological and growth responses of arctic plants to a field experiment simulating climate change. *Ecology* 77: 822–840.

natural grasslands has been shown to decrease between 50% and more than 300% under moderate to heavy grazing, depending on regional climatic conditions. Species composition is crucial at this level; for example, NAPP tends to be higher in legume-dominated pastures than in grass-dominated ones because legume growth is much less limited by soil nitrogen availability due to their capacity for symbiotic nitrogen fixing.

Biomass production from local to global scales can also be estimated by indices obtained from remote sensing. The normalized difference vegetation index (NDVI) is tightly related to the fraction of the photosynthetically active radiation absorbed by the vegetation. This spectral index is derived from the reflectance in the red and infrared bands measured by different sensors (e.g., MODIS, AVHRR, and LandSat TM). Combined with information about incident radiation and an energy conversion factor, NDVI can be used to calculate productivity. The source of images to be used will depend on the spatial and temporal scales and resolution required for a particular study.

Trophic Transfer

Trophic transfer is defined here as the amount of biomass or energy, which is transferred from the primary producers to the herbivore-based food chain rather than directly from plants to detritivores. In this article, only the transfer from plants to herbivores will be analyzed. Two concepts are particularly relevant: consumption, or the amount of plant biomass consumed by herbivores, and net secondary productivity (NSP), or the amount of biomass/energy at the herbivore level that is available to carnivores. Both depend on the amount of available biomass (NPP), the quality of plant biomass, the kind of metabolism of the herbivores, and the linkages between herbivores and plants.

Effects of Nutrient Availability, Plant Production, and Plant Quality on Herbivore Performance

Plant nutrient quality can directly affect animal populations. Plants growing in low-nutrient sites tend to have elevated

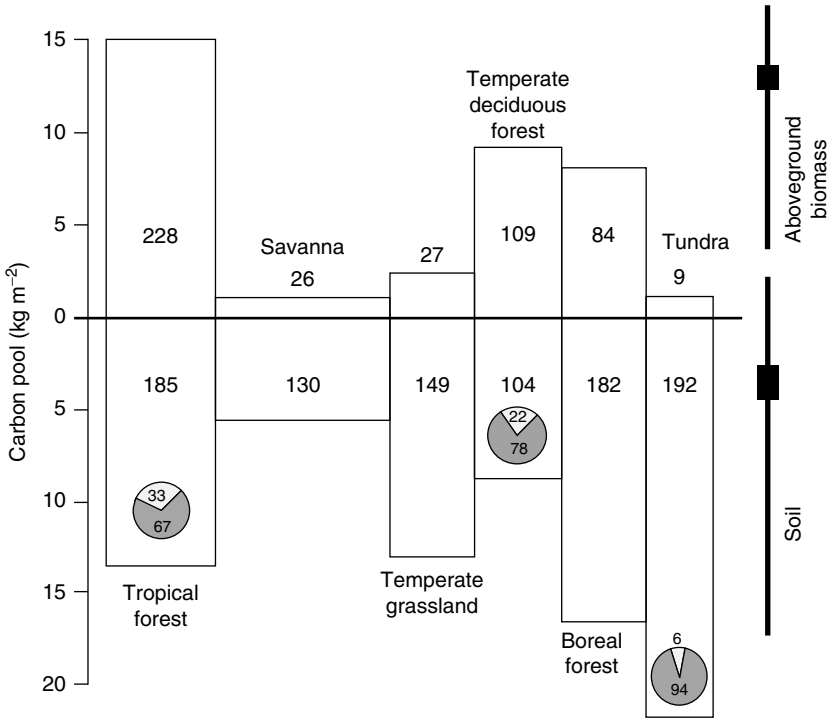


Figure 3 Carbon pools in major ecosystem types. Soil stocks include biomass, soil organic mass, and litter. Pie diagrams indicate percentage of soil carbon in belowground biomass (gray) and in soil organic mass (white). Modified from Anderson JM (1991) The effects of climate change on decomposition processes in grassland and coniferous forests. *Ecological Applications* 3: 326–347, and Larcher W (1995) *Physiological Plant Pathology*. Berlin: Springer.

concentrations of carbon-based secondary compounds that deter consumption. In addition, low foliar concentrations of nutrients can result in either decreased or increased consumption by herbivores. In the second case, herbivores (most commonly insects) consume higher quantities of biomass in order to meet nutritional requirements, with or without reduced fitness as the final result. This effect, sometimes called the 'nutrient dilution effect,' is particularly true in the case of generalist herbivores, whose opportunities for coevolutionary adjustment to the chemistry of a particular plant species are slight.

Secondary chemicals in foliage seem to be more important as regulators of the types of herbivores prevailing in different ecosystems and their dietary habits than as regulators of consumption at the ecosystem level. There is evidence that herbivore biomass, consumption, and productivity are closely correlated with plant productivity across a wide range of ecosystems, such as deserts, savannas, agricultural grasslands, tropical forests, and salt marshes (**Figure 5**). Respiratory costs per unit production at the level of herbivores are higher for homeotherms (such as ungulates) than for heterotherms (such as insects). In forests, in which most herbivores are heterotherms, most production is allocated to wood, whereas in grasslands, in which homeotherms predominate, a much higher production is allocated to green tissue. This explains why net foliage production (that of leaves only) predicts consumption considerably better than net aboveground biomass.

Herbivore biomass and consumption increase as a power of NAPP, whereas net secondary production increases linearly (**Figure 5**). This indicates that highly productive ecosystems sustain a larger level of herbivory per unit of NAPP than unproductive systems. This larger level of herbivory, however, is accompanied by a lower secondary production per unit of consumption. Because the foliage unconsumed by herbivores will flow into decomposer food webs, the relationship shown in **Figure 5** between consumption by herbivores and NAPP indicates that the relative importance of the direct flow to detritus decreases as ecosystem productivity increases. However, the relative amount of litter that is actually consumed is small, even in the most productive systems.

Effects of Herbivores on Nutrient Cycling and Primary Productivity

Animal activity can affect nutrient cycling directly and indirectly. The amount of NAPP consumed by herbivores can vary from less than 10% in tropical rain forests to more than 50% in meadows (Larcher, 1995). Herbivores can short-circuit the decomposer pathway since urine and feces are much easier to decompose than plant litter. Many of them, such as ungulates, rodents, and ants, can also redistribute nutrients and create patches, and they can promote secondary compound production in plants, which may further deter grazers and may retard decomposition. Mammals, ants, and termites can play important roles in spatial distribution of nutrients at different scales, with impacts on the dynamics of the whole food web.

In many systems primary productivity increases with light grazing, then decreases, and finally decreases more or less

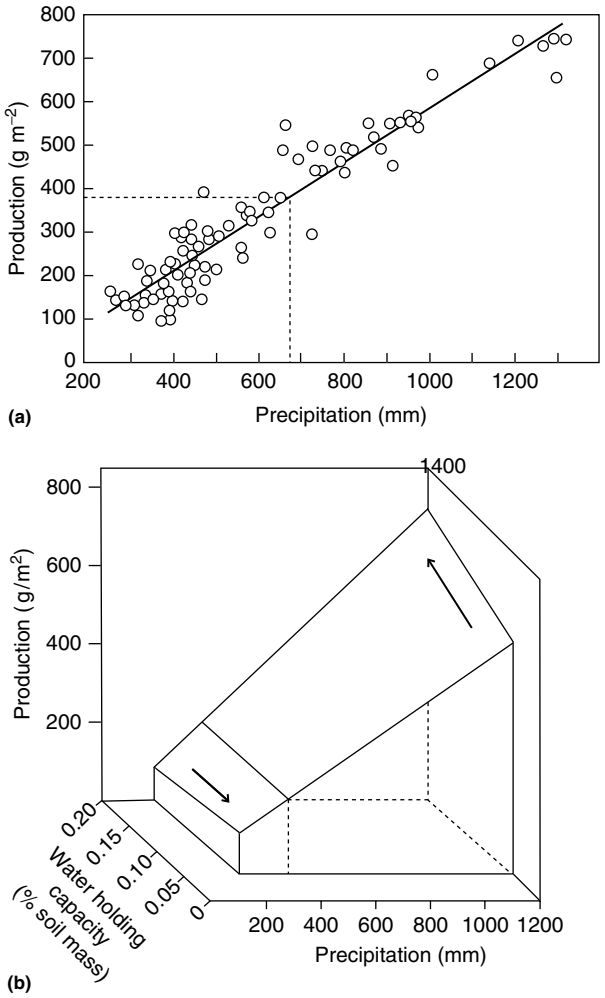


Figure 4 Regional- and site-level controls over net aboveground primary production (NAPP) of US grasslands, (a) Annual precipitation (APPT) is the main factor at the regional level, with NAPP = 0.6 (APPT − 56) (r^2 = 0.90), where 0.6 represents the average water use efficiency of the community, and 56 mm year^{-1} is the 'ineffective precipitation' (precipitation volume which is not enough to result in production). Addition of temperature and potential evapotranspiration did not improve the model, (b) Annual precipitation and soil water-holding capacity (whc) are the main factors at the site level, with NAPP = 32 + 0.45 APPT − 352 whc + 0.95 whc APPT; r^2 = 0.67). Reproduced from Sala OE, Parton WJ, Joyce LA, and Lawenroth WK (1988) Primary production of the central grassland region of the United States. *Ecology* 69: 40–45.

sharply as grazing becomes severe. This was at the heart of the herbivore optimization curve hypothesis proposed by S. J. McNaughton in the 1970s. Herbivore optimization models are still under debate, and to determine whether total (aboveground and belowground) biomass actually increases under herbivory is operationally difficult. However, supporting evidence has been found in a wide range of terrestrial ecosystems, including not only grasslands and wild and domestic ungulates but also forbs and geese, trees and mosses, and crops and birds. It is generally accepted that the highest production rates of plants occur when grazing occurs but it is not too high. Factors like water and nutrient availability, grazing history and timing, and links with other trophic levels modulate this general response in particular systems.

Bottom-Up and Top-Down Controls of Food Chains

Bottom-up or resource control in food chains emphasizes the importance of resource availability (water and soil nutrients) for primary producers and the subsequent energy and nutrient flow through a series of trophic levels. Organisms at each trophic level are food-limited. The rationale of the opposite, or top-down, view is that organisms at the top of the food chain (e.g., top carnivores) are food-limited but those at lower trophic levels (herbivores) are alternatively predator- and food-limited. According to the Oksanen–Frettwell theory (Oksanen *et al.*, 1981) the importance of top-down control increases with primary productivity. This idea has found empirical support in naturally occurring ecosystems and has also been supported by smaller-scale manipulative experiments. For example, for invertebrate herbivory and predation on limestone grasslands through a series of experiments summarized in **Figure 6** involving pesticide treatments, transplant of turves, and the use of bioassays (lettuce discs and blowfly maggots in order to evaluate the degree of herbivory and predation, respectively).

The literature contains clear examples of both bottom-up and top-down controls from various ecosystems around the world (see, for e.g., the classical book by White (1993) on bottom-up control and the recent review by Estes *et al.* (2011) on top-down control of some major ecosystem processes). Current thinking tends to see both processes as combined rather than mutually exclusive factors, operating in ecosystems, and switching in relative importance according to spatial and temporal variability in environmental drivers.

Nutrient Cycling

Nutrient Capture, Retention, and Release by Plants

A key, albeit controversial, concept in relation to nutrient cycling is that of resource use efficiency, or the relationship between a limiting resource (light, nitrogen, and water) and a biological process (photosynthesis and primary production). Nutrient use efficiency was originally defined by Vitousek (1982) as the total NPP (above- plus belowground) per unit nutrient absorbed annually. In practice, it has usually been measured as the ratio of dry mass to nutrient content in litter (**Table 1**), which is a good index of the nutrient economy in a stand as a whole and is based on information reasonably easy to obtain.

Major Controls Over Nutrient Cycling

The main factors underlying variations in nutrient cycling in different ecosystems are climatic factors, soil fertility, time from major disturbance events, and species composition. Nitrogen content is a strong control of productivity, carbon exchange, and composition on many ecosystems. In some cases, other macronutrients also exert strong controls (e.g., phosphorus and calcium in some tropical rain forests and grasslands, respectively).

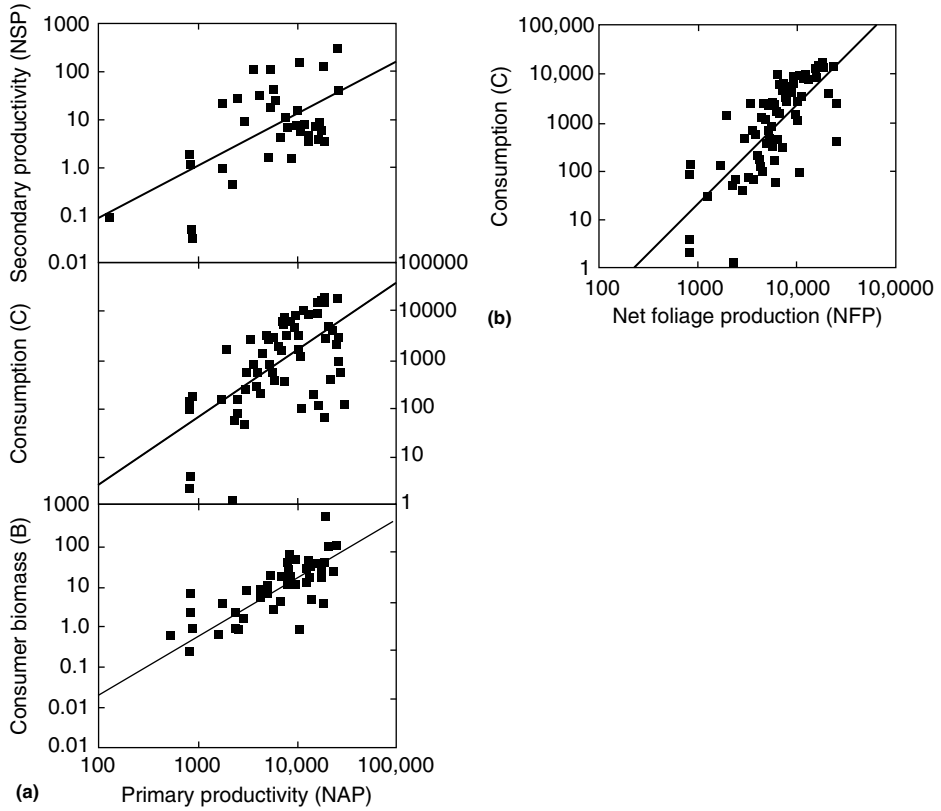


Figure 5 (a) Net aboveground primary productivity (NAP) can predict net secondary productivity (NSP), consumption (C), and consumer biomass (B). Log NSP = 1.10(log NAP) − 327, $r^2 = 0.364$; log C = 1.35(log NAP) − 2.32, $r^2 = 0.367$; log B = 1.52(log NAP) − 4.79, $r^2 = 0.367$. (b) Net foliage primary productivity (NFP) is a better predictor of consumption. Log C = 2.04(log NFP) − 4.80, $r^2 = 0.594$. Units are kJ m^{-2} $year^{-1}$ except for biomass, which is kJ m^{-2}. All relationships are significant at $p < 0.0001$ (modified from McNaughton SJ, Oesterheld M, Frank DA, and Williams KJ (1989) Ecosystem-level patterns of primary productivity and herbivory in terrestrial habitats. *Nature* 341: 142–144).

Open Versus Tight Nutrient Cycles

A concept strongly linked with nutrient use is the degree of 'tightness' or 'openness' of nutrient cycling in different ecosystems. This refers to the relative importance of within-system nutrient cycling versus external inputs and outputs. Tropical forests are considered systems with tighter nutrient cycling than temperate forests because within-system nutrient recycling is more important than influx into or effluxes out of the system.

The Effects of Rainfall

Concentrations of available nutrients in the soil are relatively (although not absolutely) high in semiarid sites and decrease with increasing precipitation. Soil organic carbon tends to follow the opposite pattern, increasing with increasing rainfall. Whereas total soil nutrient content may also increase, carbon : nutrient ratios in soil increase with higher mean annual precipitation. This suggests that, as rainfall increases, rate of carbon accumulation in soils is higher than that of total nutrient accumulation due to differences in mineralization.

Changes in nutrient cycling 'tightness' with climatic factors are well illustrated by an analysis of soil and foliar nutrients in a rainfall gradient in Hawaii (**Figure 7**). As rainfall increases with altitude, there is a shift from relatively high nutrient availability to relatively high carbon gain by producers, indicated by a decrease of leaf mass per area and leaf nitrogen concentration and an increase in lignin concentration with altitude (**Figure 7(b)**). A progressively depleted ^{15}N signature in both soils and vegetation in the wetter sites (**Figure 9(a)**) suggests that N cycling shifts from more open at the drier sites (larger turnover) to tighter (smaller losses) as precipitation increases.

The Effects of Soil Type, Land Use, and Vegetation Structure

A comparison among different types of Amazonian rain forest ecosystem, differing in soil properties and topographic positions (**Table 2**) illustrates how patterns of nutrient allocation depend on soil chemical properties and flooding regimes. In turn, soil properties influence nutrient supply, and flooding regimes affect nutrient uptake ability. Mixed and *guaco* forests, located in higher topographical positions, show relatively high nitrogen contents in both soil and vegetation. In the tall *caatinga* forest, the proportion of total nitrogen in living biomass is much higher than in the mixed forest and *guaco* forests, and decomposition is slower, associated with waterlogging and nitrogen limitation. Nitrogen circulates in larger amounts in mixed and *guaco* forest, which show more 'open' nitrogen cycling than tall *caatinga* forest.

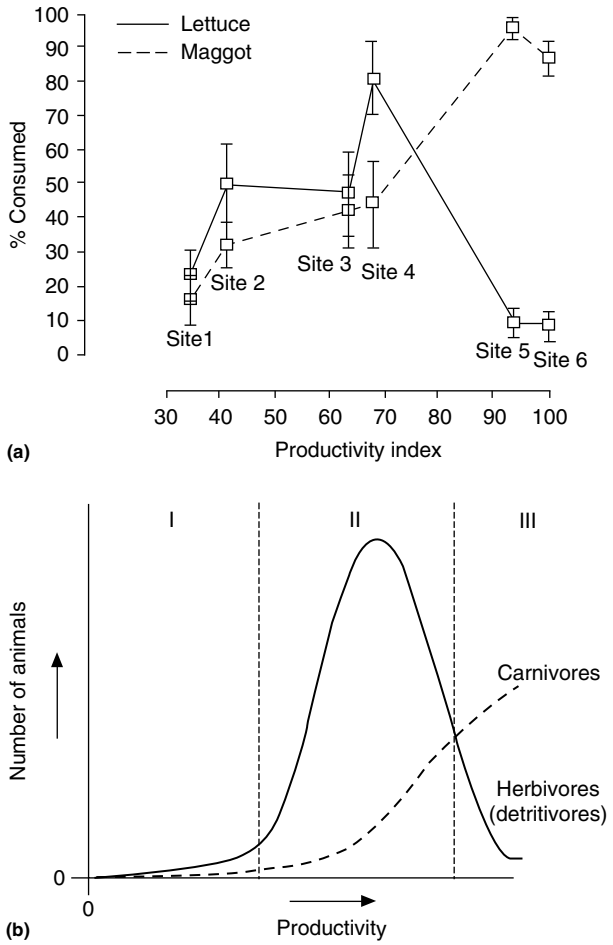


Figure 6 (a) Estimation of herbivore and carnivore activity, measured as a percentage of lettuce discs and maggots consumed, respectively, and (b) illustration of the Oksanen–Fretwell theory of interaction between trophic dynamics and primary productivity in British limestone grasslands with low (I), intermediate (II), and high (III) productivity. At very low productivity, the vegetation did not experience a detectable amount of herbivory. Vegetation of intermediate productivity supported a high level of herbivory and responded strongly to the removal of herbivores. Productivity in this situation is believed to be insufficient to sustain a high intensity of 'top-down' control. In the case of highly productive vegetation, carnivory was strong and the intensity of the top-down control by carnivores on plant mass removal by herbivores was maximum. Reproduced from Fraser LH and Grime JP (1997) Primary productivity and trophic dynamics investigated in a north Derbyshire, UK, dale. *Oikos* 80: 499–508.

Changes in vegetation structure due to land use can lead to different nutrient relations in sites under similar climatic and original soil conditions. A comparison between a tropical dry deciduous forest and savannas derived from the same forest and now maintained by grazing and fire in northern India (**Table 3**) shows that nutrient cycling is faster in the vegetation, litter, and soil of the savanna. The forest shows a higher nutrient use efficiency: Whereas the biomass and nutrient content are much higher, the annual net production and nutrient uptake are similar to those of the savanna. In the savanna there is smaller permanent nutrient storage, and significant nutrient leakage from the system, reflected in lower soil content, indicating more open nutrient cycling.

Decomposition: From Nutrient Organic to Inorganic Forms Available to Plants

The process of decomposition, or disintegration of plant and animal residues by the soil detrital food web, is a key step in nutrient cycling. This is because it transforms nutrients already present in the system in organic form into the inorganic forms available for new plant growth. Most plants absorb nutrients in inorganic form, although some plants can incorporate organic forms of nitrogen with the intervention of mycorrhizal symbionts or through highly specialized roots. However, these cases seem mostly restricted to very nutrient-poor systems. In most ecosystems, the majority of primary production is not consumed by herbivores but is passed directly to detritus. A smaller fraction of primary production is incorporated into herbivores and carnivores, which becomes detritus when these organisms die. The ultimate end product of organic matter breakdown is inorganic forms of carbon (CO_2) and nutrients (nitrates, ammonium, and phosphates). Decomposition depends on soil environment (water potential, temperature, and aeration), microsite characteristics (slope, texture, drainage, aspect, and cover type), substrate quality, and composition of decomposer community (size and specific composition of animal and microbe assemblages and synergistic or antagonistic relationships among them).

The control of decomposition by macroclimatic parameters is very strong. Whereas mean NPP increases by a factor of approximately 20 from tundra to tropical rain forests, mean residence times of dead organic matter decrease by a factor of approximately 200. This is primarily because soil temperatures limit decomposition more than air temperatures limit production.

At a more local scale, litter quality plays an extremely important role in determining decomposition rates. The decomposability of litter from different plant species tends to be strongly correlated with the palatability of living leaves. This suggests that the same compounds that determine palatability aboveground control litter decomposition by soil microbiota. The most widely used indexes to describe litter quality are the carbon : nitrogen and lignin : nitrogen ratios, with higher ratios being associated with lower decomposition rates (**Table 4**).

A worldwide review of litter decomposition studies by Cornwell *et al.* (2008) found that litter decomposition rate can vary more than 18 times among species. It is also found that large differences in decomposition rate among coexisting species can be found in most major biomes of the world.

Mineralization and Immobilization of Nutrients in the Soil

The release of organically bound nutrients into the inorganic forms available to most plants is called mineralization. Because decomposer soil organisms require these nutrients to be incorporated into their bodies, decomposition can only

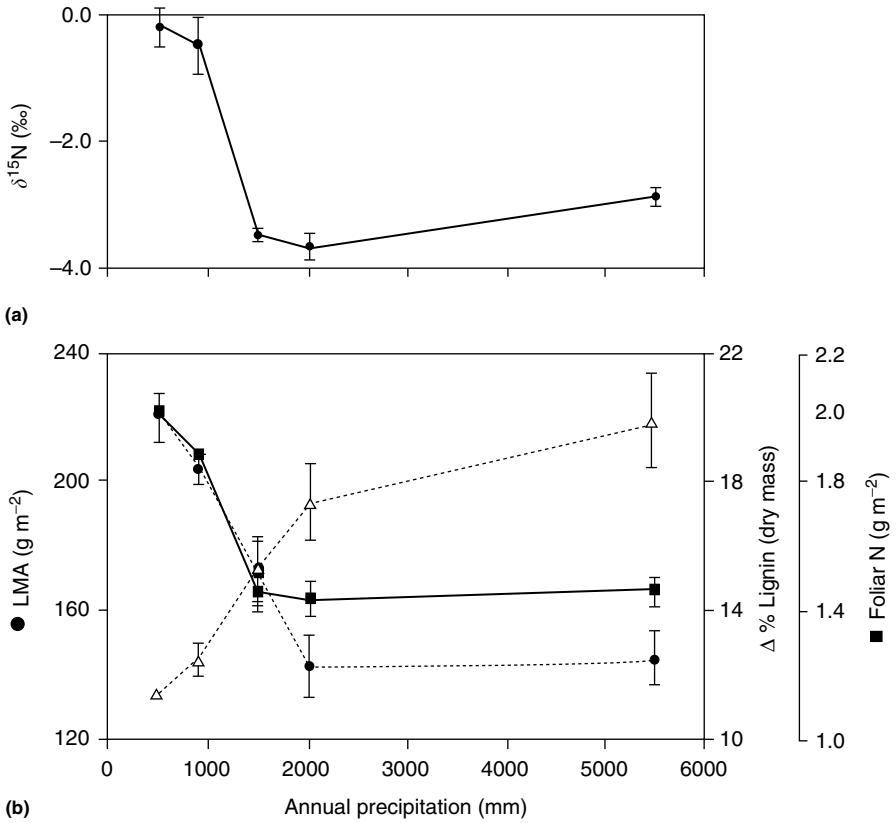


Figure 7 Influence of precipitation on nutrient cycling. Nitrogen cycle becomes increasingly closed with increased precipitation in Hawaii, (a) δ ^{15}N average values of leaves of seven species, (b) Leaf mass per area (LMA), lignin concentration, and foliar N concentration of the dominant *Metrosideros polymorpha* Gaud. Modified from Austin AT and Vitousek PM (1998) Nutrient dynamics on a precipitation gradient in Hawai'i. *Oecologia* 113: 519–529.

Table 2 Patterns in nutrient flows and stocks in three different Amazonian forests under similar macroclimate

	Mixed	Guaco	*Tall* caatinga
Soil	Concretional oxisol	Yellow ultisols	Tropaquods
Geomorphological features	Rolling hills, laterite cap, never flooded	Clayey hillsides, never flooded, sometimes water saturated	Sandy valley fills, flood-prone
Total biomass (t ha^{-1})	310	465	400
Leaf fall (t ha^{-1} year^{-1})	6.5	7	5.2
Disappearance constant (k; year^{-1})	3.68–0.34	–	0.93–0.62
Biomass: nitrogen ratio (g: g)	~110	~120	~260
Litter N content (kg ha^{-1})	137	69.5	132
Soil N content (at 10 cm depth) (kg ha^{-1})	1474	2490	716

Source: Data from Medina E and Cuevas E (1989) Patterns of nutrient accumulation and release in Amazonian forests of the upper Rio Negro Basin. In: Proctor J (ed.) *Mineral Nutrients in Tropical Forests and Savanna Ecosystems*, pp. 217–240. Oxford: Blackwell.

proceed if there is an adequate nutrient supply to the decomposers. Otherwise, they can act as a net sink for available nutrients in the system rather than as a net source, and decomposition is then said to be nutrient limited. The retention of nutrients by decomposers of nutrient-deficient substrates is termed 'immobilization.' Mineralization and immobilization tend to occur simultaneously in most systems; but one of them usually predominates over the other, with important implications in the whole functioning of the system. For example, the decrease in soil fertility immediately after the addition of cereal straw to the soil is due to immobilization, with carbon : nitrogen ratios of 100 to 150 : 1. In arable soils, a carbon : nitrogen ratio of 20 : 1 is considered the threshold between net immobilization and mineralization.

Table 3 Nutrient cycling in two tropical dry deciduous vegetation types under the same climatic and soil conditions, with different land-use history in Northern India

	Dry deciduous forest			*Savanna*		
Production (t ha^{-1} $year^{-1}$)		15			11	
Total biomass (t ha^{-1})		95			67	
	N	P	K	N	P	K
Nutrient uptake per unit energy captured (mg 1000 $kcal^{-1}$)	3000	225	1720	3290	320	2740
Vegetation nutrient content (kg ha^{-1})	680	53	451	87	8	74
Litter nutrient content (kg ha^{-1})	37	3	11	14	1	11
Soil nutrient content (0–30 cm depth) (kg ha^{-1})	2906	126	377	2386	134	160
Total litter fall (kg ha^{-1} $year^{-1}$)	80	6	38	82	8	77
Total nutrient release (leaf + root decomposition; kg ha^{-1} $year^{-1}$)	104	8	58	124	11	107
Nutrient retention in vegetation (kg ha^{-1} $year^{-1}$)	62	4	38	20	2	17

Source: Data from Singh KP (1989) Mineral nutrient dynamics in tropical dry deciduous forest and savanna ecosystems in India. In: Proctor J (ed.) *Mineral Nutrients in Tropical Forests and Savanna Ecosystems*, pp. 153–168. Oxford: Blackwell.

Table 4 Relationships between leaf litter quality (C : N) and decomposition rate (% dry mass loss) in a wide variety of plant species

References	*Plant material*	*r*	*P*	*n*
1	British shrubs and trees	−0.780	<0.01	12
2	Mediterranean shrubs and trees	−0.720	<0.05	8
3	South American temperate and subtropical graminoids, forbs, succulents, shrubs, and trees	−0.520	<0.001	52

r, Spearman's correlation coefficient; *p*, significance level; *n*, number of species involved.

Source: Reproduced from Cornelissen JHC (1996) An experimental comparison of leaf decomposition rates in a wide range of temperate plant species and types. *Journal of Ecology* 84: 573–582; Gallardo A and Merino J (1993) Leaf decomposition in two mediterranean ecosystems of southwest Spain: Influence of substrate quality. *Ecology* 74: 152–161, and Pérez-Harguindeguy N, Díaz S, Cornelissen JHC, Vendramini F, Cabido M, and Castellanos A (2000) Chemistry and toughness predict leaf litter decomposition rates over a wide spectrum of functional types and taxa in central Argentina. *Plant and Soil* 218: 21–30.

Litter Decomposition and Composition of Detrital Food Webs

The composition of the decomposer communities, which include macrofauna (earthworms and arthropods that carry out initial comminution, mixing, and dispersion of litter and microbial propagules), mesofauna (springtails, mites, and enchytraeid worms), and microbiota (fungi, bacteria, protozoa, nematodes, and actinomycetes), affects the local decomposition rate. Several experiments manipulating densities of bacteria, fungi, and their predators (e.g., protozoa and nematodes) as well as macrofauna (e.g., millipedes) have shown that the composition of the detrital food web is a strong factor determining the balance between net mineralization and immobilization. Belowground food webs are often connected with aboveground food webs in complex positive and negative feedbacks, with consequences at the ecosystem-level processes such as carbon sequestration and soil nutrient supply. For example, aboveground grazing by vertebrates on plants can increase microbial activity and mineralization. Further examples of the ecosystem-level consequences of the interactions between aboveground and belowground food webs can be found in the synthesis works by Wardle *et al.* (2004) and Bardgett and Wardle (2010).

Water Dynamics

Water flow through the plant compartment of the soil–plant–atmosphere continuum is regulated at the plant–air interface, where the transition of liquid water to vapor occurs and the steepest gradient of water potential exists. The shoots are exposed to the high vapor pressure deficit of the air and a flow of water through the plant is set in motion. In this way, and depending on xylem structure, the plant component of the system bridges the steep water potential gradient between soil and air.

Vapor loss from land and vegetation is governed by leaf area, by stomatal and aerodynamic conductances of plant canopies, and by the contribution of evaporation from soils, and it affects numerous terrestrial processes ranging from the biogeochemical cycling of elements to the development of regional climate. The aerodynamic conditions above and in the canopy, determined by the density and the architecture of a vegetation stand, strongly influence transpiration. These can also affect climate, and therefore water availability, at the regional level. For example, deforestation, by reducing roughness and increasing albedo, tends to result in higher temperatures and decreased precipitation.

Much of the terrestrial evaporative water loss passes through the stomata pores of plant leaves. Given certain meteorological conditions, the strongest biotic determinants of transpiration from vegetation are stomatal opening (usually expressed as diffusive conductance of leaves for water vapor) and the amount of leaf area per land area (leaf area index; LAI). Maximum canopy conductance increases with LAI up to a point. Below an LAI of approximately three, wet soil evaporation contributes significantly to total ecosystem evaporation. Examples of vegetation types with LAI below three are

most deserts and tundra ecosystems, some temperate grasslands, and some semiarid shrubs dominated by members of the genera *Eucalyptus* or *Acacia*.

Maximum diffusive conductance of canopy plus soil and maximum photosynthetic rate tend to be correlated, with the correlation being better for herbaceous plants; woody plants tend to assimilate less carbon per unit of transpirational water loss. Evapotranspiration and carbon assimilation are so coupled that water use efficiency at the ecosystem level (how much biomass can be built up per unit of water transpired) can be approximately estimated as NPP/precipitation.

Because of the link between water loss and CO_2 absorption at the leaf level, and because it is a very important factor in soil processes, water availability can also indirectly control ecosystems by affecting cycling of carbon and nitrogen (*see* Open Versus Tight Nutrient Cycles *and* The Effects of Rainfall). There is evidence suggesting that the nutritional status of a canopy type determines the capacity for exchanging gases with the atmosphere. Maximum evaporation from vegetation plus soil and carbon assimilation rates have been shown to be determined by plant nutrition across very different ecosystem types. However, these patterns apply for maximum rates, which do not always reflect seasonal carbon or water balances. For example, coniferous forests are less sensitive than grasslands to soil drought. They appear to close their stomata at lower soil water content than do herbaceous plants. In addition, woody vegetations are generally deeper rooted and thus access a larger soil and water volume than do herbaceous species. Apparently, the main variable that determines seasonal carbon balance is the total length of the growing season, which depends mainly on functional rooting depth of the plants. The constraints during the growing season, rather than those during the unfavorable season, seem to determine the success of different plant life-forms in different regions of the world.

Ecosystem Functioning Under Disturbance: Resistance and Resilience

Ecosystem stability is often divided into two components: resistance and resilience. Ecosystem resistance is the ability of a system to avoid change – the capacity to stay in the same state in the face of perturbation (e.g., fire, unusual frost or drought, plowing, eutrophication). Ecosystem resilience is the rate at which a system returns to its former state after being displaced from it by a perturbation. Not all aspects of ecosystem functioning are equally resistant or resilient. The most commonly measured ones (which are not necessarily the most important or sensitive ones) are species composition and biomass. Usually, a standard or ad hoc index is constructed that relates vegetation structure or composition before the perturbation to that after the perturbation. A higher dissimilarity means the ecosystem has a lower resistance. The longer amount of time needed to achieve maximum similarity between the predisturbance and postdisturbance situation, the lower the resilience. Some authors have used indicators other than community composition, such as nutrient loss rate, to estimate stability; the higher the relative nutrient loss rate

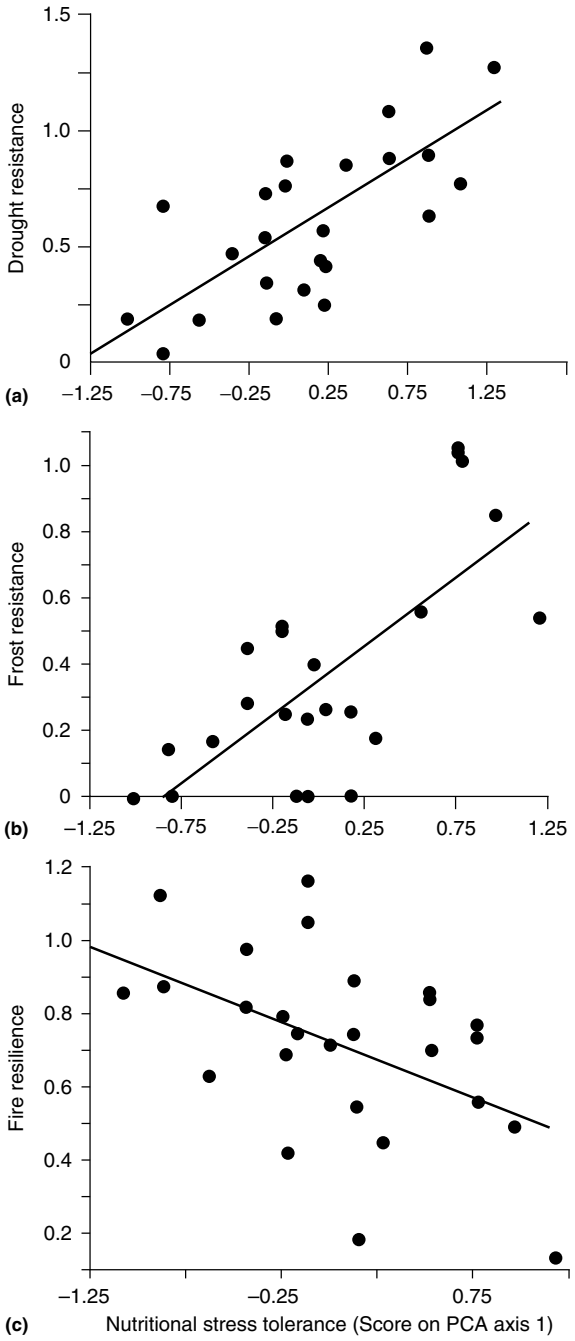


Figure 8 Relationships between estimates of (a) drought resistance, (b) frost resistance, (c) fire resilience and nutrient stress tolerance (expressed as scores on a PCA axis) for 26 British herbaceous species. Experimental treatments of fire, drought, and frosts were applied to turves of natural herbaceous vegetation. The resistance of each species was measured as the ratio of the biomass of that species in the treated turves to its biomass in the control turves immediately after the application of the treatments. The resilience (capacity for recovery) of each species was measured as the ratio of the biomass of that species in the treated turves to its biomass in the control turves 8 weeks and 1 year after applying the treatments. Modified from MacGillivray CW, Grime JP, and the ISP Team (1995) Testing predictions of the resistance and resilience of vegetation subjected to extreme events. *Functional Ecology* 9: 640–649.

following disturbance, the lower the resistance, and the longer the time to restore 'normal' nutrient loss rate, the lower the resilience.

Evidence has accumulated for the idea that both components of ecosystem stability are primarily determined by key traits of the dominant plant species. Highly productive communities, dominated by fast-growing plants, tend to have high resilience and low resistance, with the opposite being true for communities dominated by slow-growing plants. Productivity and seed production (especially persistent seeds) favor resilience, whereas preferential allocation to storage and defense favors resistance. These ideas were formalized by Lepš *et al.* (1982) on the basis of studies of old-field successional communities. Recently, new experimental support has been provided for the idea that vegetation resistance and resilience in the face of extreme events are a function of the nutrient stress tolerance of the component species. **Figure 8** illustrates that resistance to extreme events increases, and resilience decreases with increasing nutrient stress tolerance in herbaceous communities.

The size and the turnover rate of the detrital compartment also have implications for ecosystem stability. Ecosystems with high carbon : nutrient ratios in the soil are effective at immobilizing large amounts of nutrients. This immobilization can substantially control the nutrient losses as long as a sufficient amount of plant residue is left on the site to sustain the microorganisms that immobilize the nutrients in their biomass. Therefore, soil organic mass is a major determinant of ecosystem resistance. However, systems with tight nutrient cycles have increased return time to equilibrium (lower resilience). When substantial amounts of biomass are removed from these systems, causing high nutrient loss, recovery may be very slow because there is little throughflow of nutrients coming from outside the system compared with the nutrient lost. As a general rule, higher nutrient mean residence time in an ecosystem increases its resistance and decreases its resilience.

Ecosystem Functioning and Functional Diversity

Taxonomic and Functional Diversity

The most common way of assessing biodiversity is still the measurement of the number and relative abundance of species per area. Taxonomic diversity *per se* is an important parameter for conservation, but it provides little information on the magnitude and rate of ecosystem processes. However, as discussed in the section Approaches to Analysis and Measurement of Ecosystem Functioning, measurements of pools and flow rates at the whole ecosystem level provide little information about a system's conservation value or about its likely response in the face of a perturbation. An intermediate approach is the one based on plant functional types. Functional types are sets of plants sharing similar responses to environmental conditions and similar effects on major ecosystem processes. This approach is as old as ecology itself, but it has gained renewed interest in the past few years because it bridges the gap between individual species and whole-ecosystem functioning, explains particular values of pools and flow rates, and improves predictions of how ecosystems can be modified by the introduction of new abiotic (e.g., fertilization) or biotic factors (e.g., invasion by alien species). The practical value of distinguishing discrete 'types' is obvious. However, it is important to bear in mind that most real plants represent transitions between, rather than typical examples of, different functional types. In fact, functional ecology has now moved from an emphasis on discrete types toward a focus on these recurrent suites of associated traits, also called functional syndromes. This concept reflects better the continuous nature of the trait variation that occurs in natural ecosystems. However, functional types are still indispensable for many purposes that require clear limits between target- and nontarget organisms.

Ecosystem Functioning and Biodiversity

Although the role of species richness *per se* remains controversial, and the evidence is sparse, the importance of the role of functional types, particularly the dominants, is incontestable. Until the early 1980s, studies on the relationship between diversity and ecosystem functioning usually emphasized the impacts of ecosystem processes on biodiversity. More recently the question has been reversed, and there is

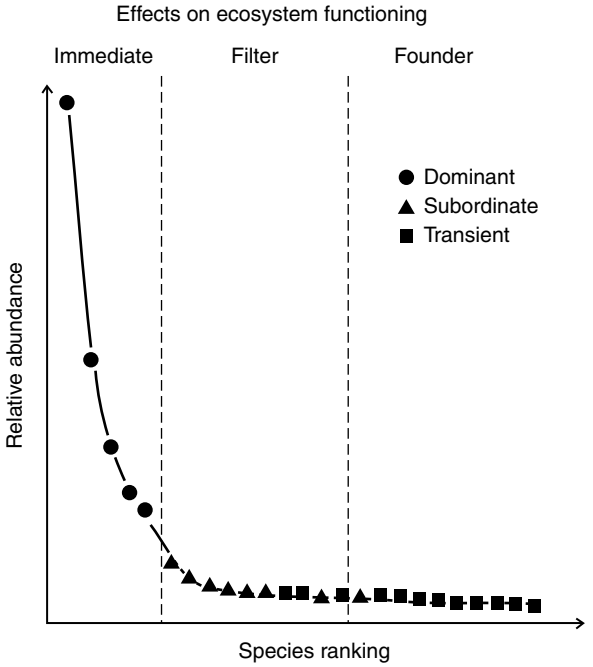


Figure 9 Proposed ecosystem effects of species in different positions along an idealized dominance–diversity curve. Transients are those species unable to regenerate and persist *in situ;* their sources are soil seed bank and seed rain from the surrounding landscape. Immediate effect, control over major processes of resource dynamics; filter effect, positive or negative influence on the regeneration of dominants following major perturbations; founder effect, reservoir of potential colonizing dominant and subordinate species in the event of ecosystem reassembly. Adapted from Grime JP (1998) Benefits of plant diversity to ecosystems: Immediate, filter and founder effects. *Journal of Ecology* 86: 901–910.

much more emphasis on how diversity influences ecosystem functioning. Biodiversity has moved from the γ to the x axis. Dominant species or functional types, which account for most of the standing biomass and energy flow, tend to be the most important (and sometimes the only) components of biodiversity accounting for ecosystem functioning at the local level. This, and the fact that they tend to represent a small proportion of the local species assemblage, provides strong

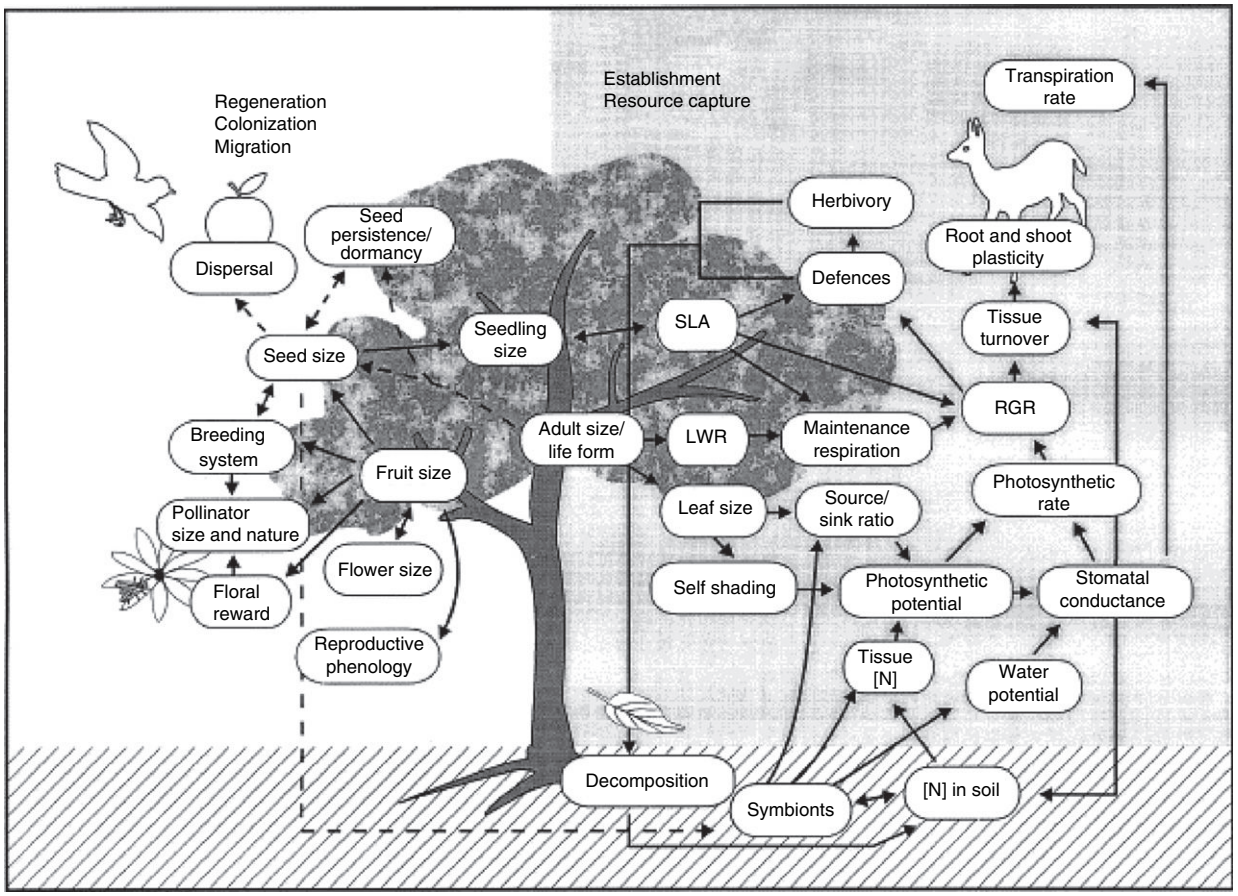


Figure 10 Some of the most common associations among vegetative traits (right), among regeneration traits (left), between vegetative and regenerative traits, and between traits and major ecosystem processes. SLA, specific leaf area; LWR, leaf weight ratio; RGR, relative growth rate.

Table 5 Examples of individual plant traits which strongly influence processes of the community/ecosystems in which they are dominant

Individual trait	*Ecosystem/community process*
Relative growth rate	Productivity, resilience, trophic transfer
Leaf turnover rate	Nutrient cycling, detritivore diversity and biomass
Nutrient content	Nutrient cycling, trophic transfer, detritivore biomass
Life span	Resistance
Canopy structure	Aerodynamic conductance, interception, water relations, runoff, roughness/albedo, temperature buffering, soil stability, consumer biodiversity
Secondary growth	Carbon sequestration, trophic transfer, nutrient cycling
Ramification	Structural complexity, consumer biodiversity, resistance (particularly drought), temperature buffering
Root architecture	Water uptake, soil stability
Reserve organs	Resilience
Pollination mode	Expansion over landscape
Persistent seed bank	Resilience
Seed number	Expansion over landscape
Dispersal mode	Expansion over landscape

Source: Modified from Díaz S, Cabido M, and Casanoves F (1999) Functional implications of trait–environment linkages in plant communities. In: Weiher E and Keddy PA (eds.) *Ecological Assembly Rules*, pp. 338–362. Cambridge, UK: Cambridge University Press.

evidence that further studies on the role of biodiversity on ecosystem functioning should focus on them. However, the likely (but far less known) role of subordinates and even rare species in maintaining long-term ecosystem functioning, especially in the face of disturbances, has also been stressed (Figure 9).

Links Between Plant Functional Traits and Ecosystem Functioning

The existence of suites of vegetative traits consistently associated across taxa and ecosystems (e.g., plants that 'go for it' and plants that 'sit and wait') has been repeatedly documented and is increasingly accepted. Central to the plant functional type approach is the idea that life history, allocation, and phenological, physiological, architectural, and reproductive traits of dominant plant species appear to be associated in a limited number of combinations and strongly determine vegetation structure and ecosystem functioning (Figure 10). By being primary components of ecosystem-level fluxes of matter and energy, dominant plant functional types are indirect determinants of the biogeochemical cycles of carbon, water, and nutrients. Individual examples of strong links between plant traits and ecosystem processes have accumulated in the past few decades (Table 5).

Major ecosystem processes appear more strongly and directly linked to recurrent suites of vegetative traits than to regeneration traits (Figure 10). Plant regeneration traits appear to be under the control of selective forces different from those that operate on resource dynamics during the established phase. *In situ* plant regeneration by seed is an important aspect of resilience and also strongly influences migration across the landscape in the face of climatic changes. Seed production, germination, dispersion, and establishment are therefore key aspects in determining ecosystem functioning in the face of major changes of climate and land use, although their role has proved much more difficult to document than that of adult-phase traits.

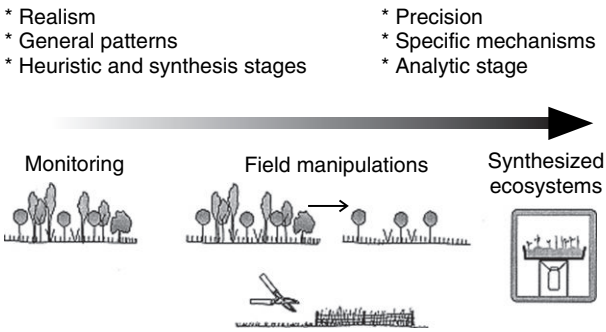


Figure 12 General characteristics of different approaches to the study of ecosystems.

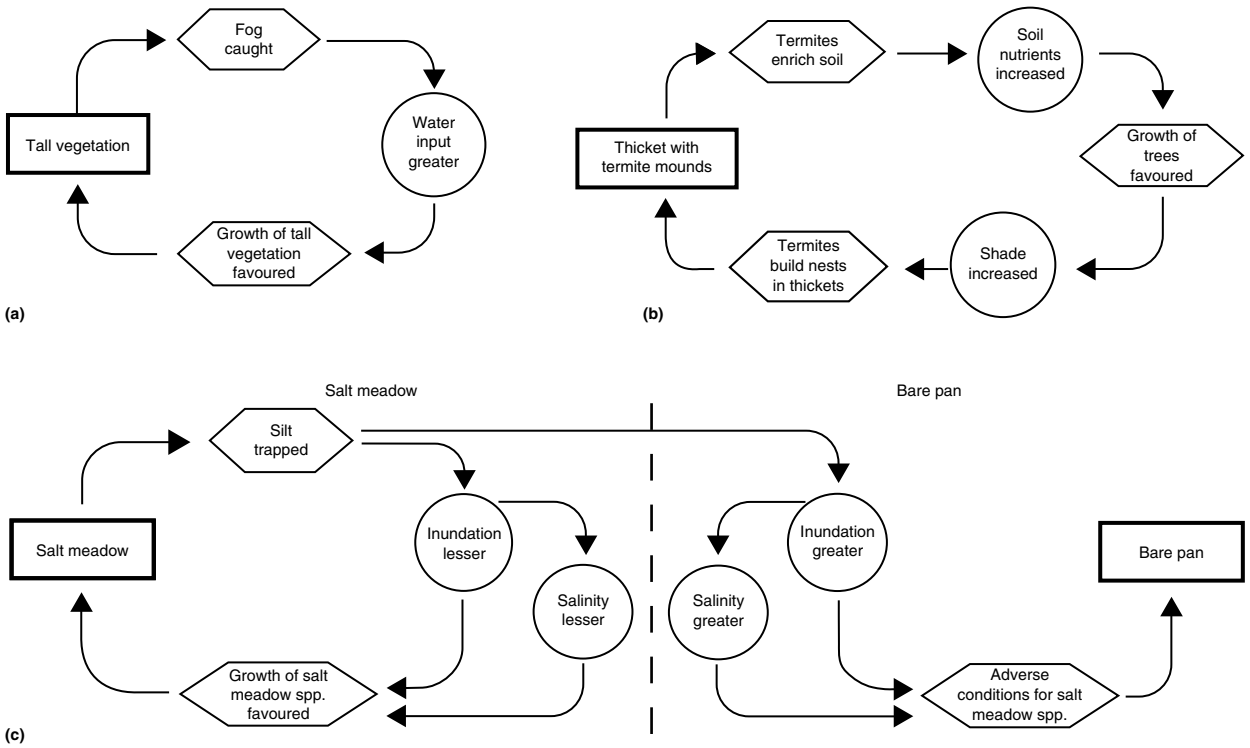


Figure 11 Switches of (a) fog precipitation on a hillside in the montane tropics, (b) termite mounds in tropical savanna, and (c) water (sediment entrapment)/salt with salt pans on a salt marsh. Reproduced from Wilson JB and Agnew ADQ (1992) Positive-feedback switches in plant communities. *Advances in Ecological Research* 23: 263–335.

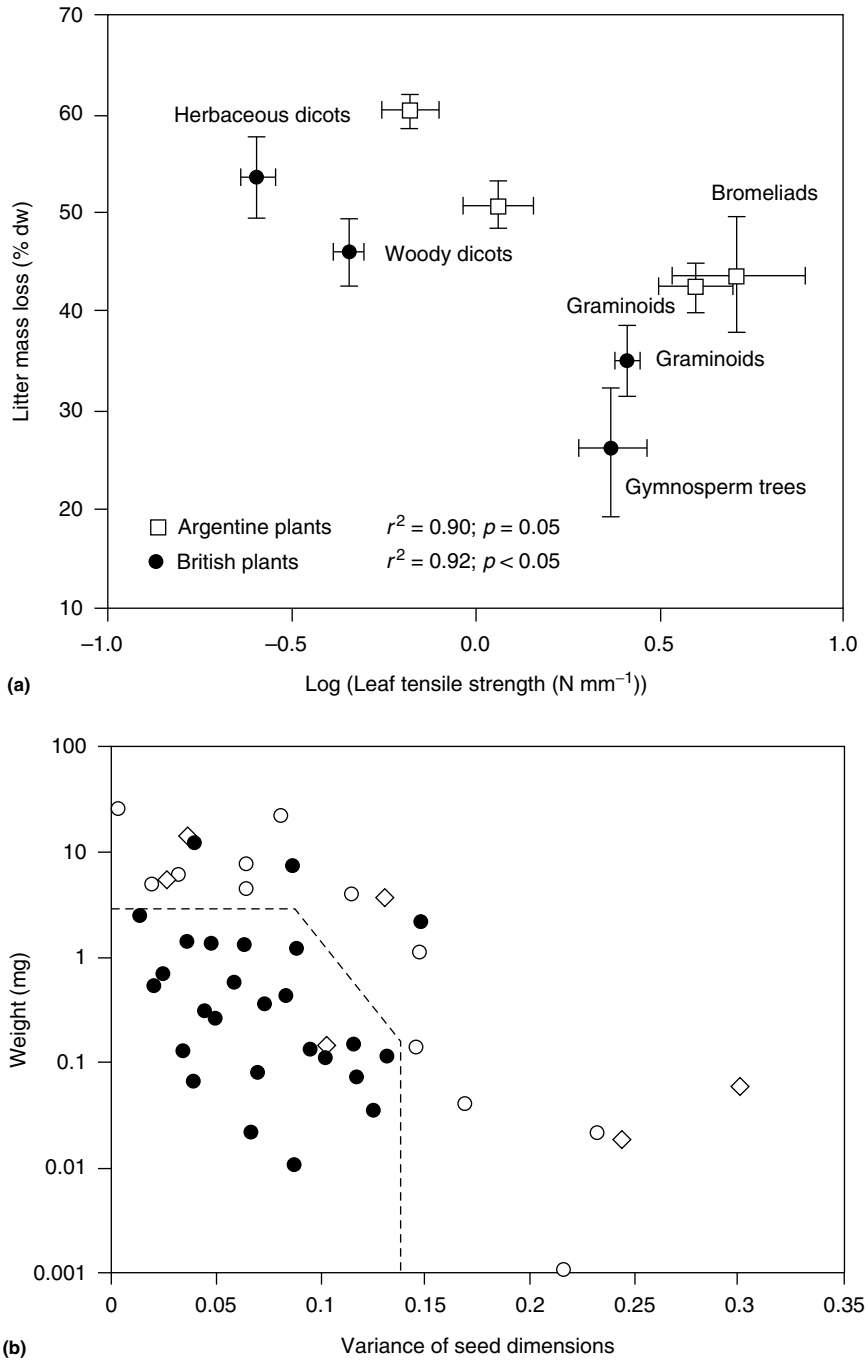


Figure 13 Examples of calibration of easily measured against more time-consuming vegetative and regeneration traits, (a) Percentage leaf litter mass loss of different plant groups as a function of tensile strength of fresh leaves (reproduced from Cornelissen JHC, Pérez-Harguindeguy N, Díaz S, *et al.* (1999) Leaf structure and defence control litter decomposition rate across species, life forms and continents. *New Phytology* 143: 191–200). (b) Relationship between seed weight and shape (expressed as variance of seed dimensions) in 44 British species: ●, species with seeds that persist in the soil for at least 5 years; ○, species with seeds that persist in the soil for <0.5 years; ◇, species whose seed bank type could not be determined. The dashed line encloses the region within which all seeds examined are long lived (reproduced from Thompson K, Band SR, and Hodgson JG (1993) Seed size and shape predict persistence in soil. *Functional Ecology* 7: 236–241).

Positive-Feedback Switches

The effect of the biotic component on ecosystem dynamics often takes the form of 'switches' or positive-feedback processes in which members of a community modify their environment, making it more suitable for themselves. This term was coined in 1992 by J.B. Wilson and A.D.Q. Agnew (Wilson and Agnew, 1992), who provided an extensive list of examples, some of which are shown in **Figure 11**.

Measurement, Analysis, and Prediction of Ecosystem Functioning: Major Protocols and Obstacles

Monitoring, Field Manipulations, and Synthesized Ecosystems: A Gradient of Questions and Methods

The understanding of how ecosystems function can be achieved by a whole gradient of protocols, from simply documenting what is happening in real systems to synthesizing model ecosystems from scratch. These extreme approaches represent a trade-off between realism and precision and between the documentation of general patterns and that of the specific mechanisms underlying them. The monitoring of real ecosystems often needs to be performed by 'soft' approaches. It is important at the initial heuristic stage and when testing whether results found in simpler experiments can be reasonably applied to the real world (synthesis stage). Experiments on synthesized ecosystems, however, are ideal for the stage of testing specific hypotheses on mechanisms that may account for the patterns observed (analytic stage). Field manipulations lay in the middle of this gradient (**Figure 12**).

Monitoring Real Ecosystems: Approaches Based on Functional Traits

Processes in natural ecosystems can be recorded by applying the usual methods of measuring pools and flows described previously (*see* Whole-Ecosystem Approaches and Approaches Based on Community Structure and Composition *and* Measurement of Short-Term Resource Dynamics and Long-Term, Large-Scale Ecosystem Processes). Alternatively, ecosystem functioning can be inferred from the presence and abundance of plant traits which are easily measured but at the same time have clear implications at the ecosystem level. Some examples of traits considered good indirect indicators of ecosystem processes are given in **Table 5**. Trait-based approaches have the advantage that they need a minimum investment in financial and technological resources and can be utilized to characterize extensive areas or high numbers of systems in a short period of time. The result is usually a comparative estimation of ecosystem functioning. This in turn needs to be calibrated against direct measurements of the plant's more direct contribution to ecosystem processes of interest, performed in a considerably smaller subset of samples. For example, leaf tensile strength correlates well with decomposition rate, which is a trait known to strongly influence nutrient cycling at the ecosystem level. The easily measured seed mass and shape correlate well with seed persistence in the soil, a trait directly involved in vegetation resilience that is considerably more difficult to measure; (**Figure 13**).

Experimental Manipulations Involving 'Natural' Versus Synthesized Ecosystems

The major disadvantage of field experiments is that the degree of control of independent and external variables is low compared to manipulating them. The amount of 'noise' (unwanted variance) is usually high, and enough replicates in order to reduce it are usually unavailable or involve prohibitive costs. However, synthesized ecosystems, such as microcosms, never come close to the realism of experiments involving manipulations of real ecosystems, they are always a badly simplified version of nature, and they have severe size limitations (e.g., for ecosystems based on woody vegetation). Considering the realism–precision trade-off illustrated in **Figure 12**, and taking into account that some processes scale up poorly from single systems to complex ecosystems (*see* Whole-Ecosystem Approaches and Approaches Based on Community Structure *and* Composition and Measurement of Short-Term Resource Dynamics and Long-Term, Large-Scale Ecosystem Processes), field experiments are indispensable. However, microcosms work has produced highly relevant insights into ecosystem functioning, including relationships between trophic levels, community roles of symbionts, feedbacks involved in soil fertility and climate manipulation; indeed some of the most controversial and inspiring experiments on the role of biodiversity in determining ecosystem functioning have involved microcosms. Field manipulations, such as the application of treatments to real vegetation plots, or the selective removal of certain components of the community, have some limitations. These include the difficulty of separating the effects of removal of one component from the effect of the disturbance caused during the removal process and the difficulty of minimizing spatial heterogeneity that usually masks treatment effects. However, they are an excellent intermediate step between laboratory experiments and the long-term monitoring of undisturbed systems. Additionally, they have produced important evidence of the difficulties and potentialities of scaling up from individual physiology to ecosystem functioning (see **Figure 2**), and they will likely continue to provide fruitful insight in ecosystem ecology.

Appendix

List of Courses

1. General Ecology
2. Ecosystem Ecology
3. Terrestrial Ecology
4. Community Ecology
5. Functional Diversity
6. Global Change Ecology
7. Biogeochemistry

Acknowledgments

I am grateful to Diana Abal-Solís for drawing the figures and to N. Pérez-Harguindeguy, L. Enrico, D. E. Gurvich, A. M. Cingolani, M. R. Zak and to two anonymous reviewers for fruitful comments.

See also: African and Asian Savannas. Alpine Ecosystems. Arctic Terrestrial Ecosystems. Biodiversity and Ecosystem Services. Boreal Forest Ecosystems. Carbon Cycle. Desert Ecosystems. Ecosystem, Concept of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Services. Energy Flow and Ecosystems. Forest Ecology. Functional Diversity. Functional Diversity Measures.

Grassland Ecosystems. Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity. Mediterranean-Climate Ecosystems. Nitrogen, Nitrogen Cycle. Remote Sensing and Image Processing. Soil Biota, Soil Systems, and Processes. Southern (Austral) Ecosystems. Stability, Concept of. Temperate Forests. Terrestrial Ecosystems. Tropical Forest Ecosystems. Valuing Ecosystem Services

References

Bardgett RD and Wardle DA (2010) *Aboveground–Belowground Linkages*. Oxford: Oxford Series in Ecology and Evolution.

Begon M, Harper JL, and Townsend C (1990) *Ecology: Individuals, Populations, and Communities*. Oxford: Blackwell.

Berg B and McClaugherty C (2008) *Plant Litter Decomposition, Humus Formation, Carbon Sequestration*. Berlin, Heidelberg: Springer.

Chapin III FS, Matson PA, Vitousek PM, and Chapin MC (2011) *Principles of Terrestrial Ecosystem Ecology*. New York: Springer.

Chapin III FS, Woodwell GM, Randerson JT, Rastetter EB, Lpvett GM, and Baldocci DD (2005) Reconciling carbon-cycle concepts, terminology, and methods. *Ecosystems* 9: 1041–1050.

Cornelissen JHC (1996) An experimental comparison of leaf decomposition rates in a wide range of temperate plant species and types. *Journal of Ecology* 84: 573–582.

Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, *et al.* (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany* 51: 335–380.

Cornwell WK, Cornelissen JHC, Amatangelo K, Dorrepaal E, Eviner VT, *et al.* (2008) Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters* 11: 1065–1071.

DeAngelis DL (1992) *Dynamics of Nutrient Cycling and Food Webs*. London: Chapman & Hall.

Díaz S, Lavorel S, de Bello F, Quetier F, Grigulis K, and Robson M (2007) Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences USA* 104: 20684–20689.

Ehleringer JR and Field CB (1993) *Scaling Physiological Processes – Leaf to Globe*. San Diego: Academic Press.

Estes JA, Terborgh J, Brashares JS, Power ME, Berger J, *et al.* (2011) Trophic downgrading of planet Earth. *Science* 333: 301–306.

Gallardo A and Merino J (1993) Leaf decomposition in two mediterranean ecosystems of southwest Spain: Influence of substrate quality. *Ecology* 74: 152–161.

Grime JP (2001) *Plant Strategies, Vegetation Processes and Ecosystem Properties*. Chichester: Wiley.

Herms DA and Mattson WJ (1992) The dilemma of plants: To grow or defend. *Quarterly Review of Biology* 3: 283–335.

Hoover CM (2008) *Field Measurements for Forest Carbon Monitoring*. New York: Springer.

Lambers H, Chapin III FS, and Pons T (1999) *Plant Physiological Ecology*. Berlin: Springer.

Lambers H, Poorter H, and Van Vuuren MMI (1998) *Inherent Variation in Plant Growth – Physiological Mechanisms and Ecological Consequences*. Leiden: Backhuys.

Larcher W (1995) *Physiological Plant Ecology*. Berlin: Springer.

Lepš J, Osbornová-Kosinová J, and Rejmánek M (1982) Community stability, complexity and species life-history strategies. *Vegetatio* 50: 53–63.

Loreau M, Neem S, and Inchausti P (2002) *Biodiversity and Ecosystem Functioning*. Oxford: Oxford University Press.

McNaughton SJ, Oesterheld M, Frank DA, and Williams KJ (1989) Ecosystem-level patterns of primary productivity and herbivory in terrestrial habitats. *Nature* 341: 142–144.

Medina E and Cuevas E (1989) Patterns of nutrient accumulation and release in Amazonian forests of the upper Rio Negro Basin. In: Proctor J (ed.) *Mineral Nutrients in Tropical Forests and Savanna Ecosystems*, pp. 217–240. Oxford: Blackwell.

Mooney HA, Medina E, and Schindler DW (1991) *Ecosystem Experiments, SCOPE Series No 45*. Chichester, UK: Wiley.

Noy-Meir I (1973) Desert ecosystems I. Environment and producers. *Annual Review of Ecological Systematics* 4(1973): 25–52.

Oksanen L, Fretwell SD, Aruda J, and Niemela P (1981) Exploitation ecosystems in gradients of primary productivity. *The American Naturalist* 118: 240–261.

Pearcy RW, Ehleringer J, Mooney HA, and Rundel PW (1989) *Plant Physiological Ecology – Field Methods and Instrumentation*. London: Chapman & Hall.

Pérez-Harguindeguy N, Díaz S, Cornelissen JHC, Vendramini F, Cabido M, and Castellanos A (2000) Chemistry and toughness predict leaf litter decomposition rates over a wide spectrum of functional types and taxa in central Argentina. *Plant and Soil* 218: 21–30.

Sala OE, Jackson RB, Mooney HA, and Howarth RW (2000) *Methods in Ecosystem Science*. New York: Springer.

Schlesinger WH (1997) *Biogeochemistry – An Analysis of Global Change*. San Diego: Academic Press.

Schulze E-D and Caldwell MM (1995) *Ecophysiology of Photosynthesis*. Berlin: Springer.

Singh KP (1989) Mineral nutrient dynamics in tropical dry deciduous forest and savanna ecosystems in India. In: Proctor J (ed.) *Mineral Nutrients in Tropical Forests and Savanna Ecosystems*, pp. 153–168. Oxford: Blackwell.

Sterner RW and Elser JJ (2002) *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton: Princeton University Press.

Vitousek PM (1982) Nutrient cycling and nutrient use efficiency. *The American Naturalist* 109: 153–172.

Vitousek PM (2004) *Nutrient Cycling and Limitation*. Princeton: Princeton University Press.

Vitousek PM and Sanford Jr. RL (1986) Nutrient cycling in moist tropical forest. *Annual Review of Ecology and Systematics* 17: 137–167.

Wardle DA, Bardgett RD, Klironomos JN, Setala H, van der Putten WH, and Wall DH (2004) Ecological linkages between aboveground and belowground biota. *Science* 304: 1629–1633.

White TCR (1993) *The Inadequate Environment*. Berlin, Heilderberg, New York: Springer-Verlag.

Wilson JB and Agnew ADQ (1992) Positive-feedback switches in plant communities. *Advances in Ecological Research* 23: 263–335.

Ecosystem Function, Principles of

Ross A Virginia, Dartmouth College, Hanover, NH, USA
Diana H Wall, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ecosystem All the individuals, species, and populations in a spatially defined area and the interactions among them and with the abiotic environment.
Ecosystem functioning The sum total of processes such as the cycling of matter, energy, and nutrients operating at the ecosystem level.
Functional group A group of species that perform similar roles in an ecosystem process.
Nutrient cycle (or biogeochemical cycle) The repeated pathway of mineral elements, such as carbon, nitrogen, phosphorus, and water from the environment through organisms and back into the environment.
Succession The predictable change in species that occupy an area over time caused by a change in biotic or abiotic factors benefiting some species but at the expense of others.

Development of the Ecosystem Concept

The concept of the ecosystem as a functioning unit in the natural world is a relatively recent one. The term ecosystem was coined by the British ecologist Tansley in 1935 and has since then become a common word in science and with the public. An ecosystem encompasses all the organisms of a given area and their relationships with one another and the physical or abiotic environment. The ecosystem contains the linkages and dynamic interactions between life and the environment, many of which are essential to society. A focus on the ecosystem as the unit of study represents a shift from studying the ecology and behavior of individual organisms and species (natural history) to the study of processes and how they influence or are influenced by organisms and their interactions with the environment.

Dividing the complexity of nature into convenient units of study is required for scientific investigation but can present problems. Ecological systems can be organized in a hierarchy of increasing levels of organization and complexity: individual, population, species, community, ecosystem, landscape, and biome. The size (scale) of an ecosystem is defined by the purposes of the study. Ecosystems may have distinct boundaries as in the case of a lake or a watershed. More often, the boundaries of one ecosystem (a forest) may grade gradually into another (a meadow) across an intermediate area called an ecotone. The ecotone is often a zone of higher diversity because it may be a suitable habitat for species from each of the adjoining ecosystems. At one extreme of scale, the earth is sometimes treated as an ecosystem. At the other extreme, the complex symbiotic community of organisms inhabiting the gut of a termite has all the functional properties of an ecosystem. The definition and delineation of an ecosystem has practical importance because ecosystems are increasingly seen as a functional unit for resource and conservation management purposes. It has become evident that the management of lands for sustained levels of ecosystem services and natural resources requires an understanding of how ecosystems function, how they respond to disturbance, and how the role of biodiversity is regulating their function and stability.

Ecosystem Functioning and Ecosystem Services

Society depends on the functioning of ecosystems for many essential ecosystem services on which we place economic and esthetic values (Daily, 1997; Millennium Ecosystem Assessment, 2005). Ecosystem functioning results from the collective activities of organisms, their life processes (production, consumption, and excretion) and the effects of these activities on the condition of the environment. These functions (services when they provide utility to humans) include production of food, fuel, and fiber, the cycling and purification of water, and the maintenance of organisms that have a role in ecosystem functioning or that provide products for human use (**Table 1**). Humans are rapidly changing the earth's ecosystems and their services by altering land use or by harvesting biological resources (forest cutting and fisheries) (Vitousek *et al.*, 1997a). Approximately 40% of the earth's primary production is diverted to human use. One consequence of these economic activities is an abrupt increase in the rate of change in biological diversity leading to species extinction, replacement of high-biodiversity ecosystems with less diverse managed

Table 1 Examples of the biological and physical processes or interactions that contribute to important ecosystem functions

Process	*Ecosystem function*
Photosynthesis	Primary production
Plant nutrient uptake	
Microbial respiration	Decomposition
Soil/sediment food web dynamics	
Nitrification	Nitrogen cycling
Denitrification	
Nitrogen fixation	
Plant transpiration	Hydrologic cycle
Root activity	
Mineral weathering	Soil formation
Soil bioturbation	
Vegetation succession	
Predator–prey interactions	Biological control

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00041-1

systems, and invasions of natural ecosystems by exotic species (Millennium Ecosystem Assessment, 2005). This pattern of ecosystem change has raised serious concern that the functioning and stability of our global ecosystem are threatened by the loss of biodiversity.

What Do Ecosystem Scientists Study?

Ecosystems share certain characteristics and functions that allow scientists to study ecosystem types (e.g., deciduous forest, temperate grassland, arctic tundra, coral reef, and deep-ocean hydrothermal vents) that vary greatly in structure, biodiversity, and spatial extent. For example, all ecosystems require inputs of energy (usually solar) and a supply of the mineral elements (nutrients) essential for life. These inputs support many ecological processes operating at multiple scales. For example, sunlight, carbon dioxide, and water are inputs for the process of photosynthesis, which can be measured and studied at the scale of individual cells, a leaf, the plant canopy, or an entire ecosystem. Photosynthesis acting with other processes such as mineral uptake by roots combine to create an ecosystem function – primary productivity.

Scientists can discover basic principles about the behavior of ecosystems by studying the functions that very different ecosystems, such as the polar desert of Antarctica and the rangelands of the southwestern US, share in common (Virginia and Wall, 1999). The movement of energy and materials within and between ecosystems and the role of organisms in mediating these processes are the parameters used by ecosystem scientists to compare the functioning of ecosystems and their responses to disturbance. Some of the important processes and functions central to the integrity and sustained activity of an ecosystem are summarized in **Table 2**. Ecosystem scientists study the rate at which ecosystems remove carbon from the atmosphere by photosynthesis, store it in the soil as organic matter, and then return the stored carbon to the atmosphere during decomposition. They study how nitrogen is cycled through ecosystems to sustain continued plant productivity. Our knowledge of how carbon and nitrogen move in the ecosystem helps us to understand when an ecosystem has been seriously altered by humans, for example, by adding nitrogen in the form of air pollution (acid rain) and fertilizers.

Many basic principles provide insight into the functioning of ecosystems and their response to human use and disturbance. Here, we will consider some of the essential functions of ecosystems and examine the principles that govern their operation, with an emphasis on the role of organisms (biodiversity) in determining ecosystem functioning.

Table 2 Examples of ecosystem services that would be affected by a decline in ecosystem function

Pest control
Insect pollination
Fisheries
Climate regulation
Soil retention
Flood control
Soil formation and maintenance of soil fertility
Cycling of matter
Composition of the atmosphere
Maintenance of genetic diversity

Source: Reproduced from Daily GC (ed.) (1997) *Nature's Services. Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.

Important Ecosystem Functions

Ecosystem Productivity

A central process of most ecosystems is photosynthesis, the capture of solar radiation, and its conversion to stored chemical forms (biomass). Plants require sunlight, water, and essential nutrients for the process of photosynthesis. Photosynthesis is coupled with other plant processes that result in plant growth, for example, the accumulation of biomass. Primary productivity, the change in plant biomass per unit area and time, is an important index of ecosystem function. Primary productivity (often referred to as ecosystem productivity) has been related to plant species diversity as well as the diversity of organisms (soil biota) that influence the availability of limiting resources. Humans depend on ecosystem productivity as the basis of our agriculture, forestry, and fisheries. Thus, factors that alter ecosystem productivity (for example, climate change and biodiversity loss) affect us directly.

Ecosystems with high rates of primary productivity have favorable amounts of resources required for plant growth and optimal climate. These systems also tend to have higher diversity (**Table 3**). The highest rates of terrestrial ecosystem productivity are seen in the tropics, where temperature and moisture are favorable for plant growth throughout the year. In contrast, water-limited hot and cold deserts have much lower productivity, averaging less than 10% of that of tropical systems.

Limits to Ecosystem Productivity

A basic principle invoked to explain variation among ecosystems in their productivity is Liebig's Law of Minimum. Justus Liebig formulated this concept during pioneering studies of the mineral nutrition of plants in the early 1800s. He found that addition of a single "limiting element" to a soil would increase plant growth. Once this element was in sufficient supply, another mineral element would have to be supplied in increased amounts to stimulate additional increase in plant growth. From these observations, he proposed that a limiting factor was responsible for limiting the growth or reproduction of an organism or population. This factor

Table 3 Typical values for the net primary productivity of major ecosystems[a]

Ecosystem type	*Net primary production (g C/m^2 year)*	*Relative species diversity*
Tropical rain forest	900	Highest
Temperate forest	540	Intermediate
Grassland	315	Intermediate
Desert	32	Low
Extreme desert	1.5	Lowest

[a]Ecosystem productivity and biodiversity are often positively related.

might be a chemical factor (a growth-stimulating nutrient such as nitrogen), a physical factor such as moisture, or a biological factor such as the presence of a competing species. Thus, any change in a limiting factor is expected to have large effects on ecosystem functioning.

There are many examples in which a change in a limiting factor alters ecosystem function. The large increase in the amount of nitrogen cycling in the environment from fertilizers and fossil fuel should have significant effects on rates of ecosystem functions since nitrogen frequently is the primary limiting element for plant growth in terrestrial ecosystems. Humans have doubled the rate of nitrogen inputs to ecosystems with increases in carbon storage and declines in biodiversity (Vitousek *et al.*, 1997b). In fact, the forests of the northeastern US may have reached "saturation" in their ability to absorb and retain anthropogenic inputs of nitrogen. The limits to production in ecosystems may also be related to the balance (ratio) of essential elements for growth (C:N:P). This theory of ecological stoichiometry predicts that there are multielement constraints on ecological interactions such as energy flow in food webs, competition, herbivory, and on ecological processes such as primary production (Sterner and Elser, 2002).

Are plant species diversity and primary production related? Ecologists are accumulating evidence from experiments in controlled growth facilities and in the field that ecosystem primary productivity increases with increasing plant species diversity. The theoretical basis for the expectation that productivity and diversity should be related derives from an understanding of how limiting resources (water and nutrients) are distributed in ecosystems and an appreciation for the diversity of physiological or "functional" traits that organisms have evolved to capture and utilize these resources for growth. Differences between plant species in rooting depth, phenology (seasonality of growth), photosynthetic rates, and other physiological traits allow multispecies communities to more fully utilize the available resources.

The ability of diverse plant communities to obtain higher productivity than low-diversity systems is demonstrated in traditional (low-input) agriculture in which polycultures (multiple-species plantings) often have higher yields than single-species plantings (monocultures) (Gliessman, 1998). For example, corn (*Zea mays* L.) yields at comparable densities are higher when corn is grown in the presence of nitrogen-fixing beans (*Vicia* spp.). The bean crop forms a symbiotic association with bacteria that "fix" atmospheric nitrogen (N_2) to an inorganic form (ammonia, NH_3) useable by plants. The nitrogen fixed by the bean crop improves the overall supply of this limiting element in the soil and increases the growth of the interplanted corn. The functioning provided by the diverse corn–bean–nitrogen-fixing bacteria association is often replaced in intensive agriculture by applying inorganic nitrogen fertilizers. With external inputs (fertilizers) the corn monoculture can produce higher yields than the polyculture. Substituting an industrial source of nitrogen for a biological source has environmental costs resulting from the production and combustion of fossil fuels used to produce fertilizers and pesticides (Mäder *et al.*, 2002). In addition, over application of fertilizers and pesticides are a major source of water pollution in surface and groundwaters.

There are similar examples of diversity influencing productivity in natural ecosystems. In a California grassland ecosystem, Hooper and Vitousek (1997) manipulated the number of plant functional groups in a community (early versus late-season forbs, perennial grasses, and nitrogen-fixing plants) in combinations of one to four groups in a given plot. They found that the number of plant functional groups was not the main factor that determined productivity. Rather, certain functional characteristics of individual species within functional groups contributed more to ecosystem productivity than overall diversity of the plot. This study points to the complexity of trying to simply relate species diversity to function. As a general principle, ecologists recognize that some species play particularly important roles in regulating important ecosystem functions such as productivity and nutrient cycling.

Keystone Species

Certain species, termed keystone species, have a disproportionate influence (relative to their biomass) on ecosystem functioning. The loss of a keystone species will produce a cascade of effects on the diversity and function of the remainder of the ecosystem (Menge *et al.*, 1994). Consequently, since keystone species can control ecosystem diversity and associated ecosystem functions, they and the habitats they live in often receive high priority in conservation management plans. There are many well-documented studies of keystone species and how they interact with ecosystem functioning, for example, the North Pacific sea otter preys on sea urchins, which consume kelp. In the absence of the keystone predator, sea urchin populations increase and create areas devoid of kelp and, consequently, the myriad of fish and other species that depend on the kelp forest (**Figure 1**). This is an example of a food web – the

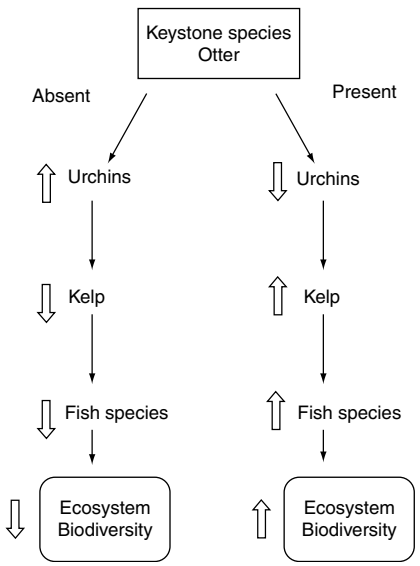


Figure 1 The influence of a keystone species on the biodiversity of an entire ecosystem can be large. Arrows indicate an increase or decrease in population size or species diversity in response to the presence or absence of the keystone species. The removal of the Pacific sea otter from California coastal ecosystems leads to the loss of the kelp community and many fish species.

representation of trophic (feeding) relationships between species in an ecosystem.

There are many examples of keystone species in terrestrial ecosystems. A large change in the number of African elephants has dramatic effects on the diversity and structure of the vegetation types (savanna woodlands and forests) they consume, altering ecosystem productivity, soil nutrient cycles, and plant community diversity. The much smaller tsetse fly shares the elephant's habitat and also has the attributes of a keystone species. The tsetse fly is the vector for the human disease sleeping sickness (African trypanosomiasis). This biting fly also influences the behavior of large herbivores that tend to avoid heavily infested areas. Consequently, herbivore-related impacts on plant communities and associated ecosystem functions are altered in tsetse-occupied ecosystems. This small insect may control the biodiversity of large tracts of Africa through another mechanism. Diverse native ecosystems have been "protected" from agricultural development and species loss because humans avoid regions where the tsetse and therefore sleeping sickness are endemic.

Nutrient Cycling

The sustained functioning of any ecosystem requires a minimum number of species to develop the intricate relationships between producers, consumers, and decomposers that regulate the flow of energy and nutrients. The productivity of all ecosystems is dependent on the cycling of essential elements. The movement and biological transformations of organic matter and nutrients are mediated by biota, especially those found in soil and sediments (Wall and Virginia, 1997). Therefore, changes in the biodiversity of ecosystems can alter biogeochemical processes.

Succession

Scientists study the process of ecological succession (ecosystem change with time, often in response to disturbance) in part to untangle relationships between biodiversity and function. Although not all ecosystems follow a predictable pathway as they develop in time, examples of succession highlight the linkage between organisms, diversity, and ecosystem function. They include the recovery of a forest after harvest or following damage by a hurricane, the reestablishment of grassland following fire, and the old-field succession of natural vegetation reclaiming abandoned agricultural land. During succession, ecosystems change in generally predictable ways as they accumulate species, increase in biomass, and gain structural complexity. Odum (1969) proposed a model of ecological succession (development) that relates ecosystem diversity, structure, and functioning as ecosystems redevelop and "mature" following disturbance (**Table 4**). Odum's model related the stability (constancy) of function and the conservation of nutrients to increasing diversity – themes that are at the center of biodiversity and ecosystem research today.

The relationships represented in Odum's (1969) model between ecosystem function and diversity are elucidated in the Hubbard Brook watershed experiment (Likens and Bormann, 1995). One of the first long-term ecosystem studies, the Hubbard Brook project began in 1963 in the White Mountains of New Hampshire. The study was designed to understand the process of forest recovery following harvest with a focus on ecosystem functions related to production, nutrient cycling, and nutrient loss. Measurements of the mature intact deciduous forest showed that less than 0.1% of the nitrogen contained in living forest biomass and dead organic matter in the soil and litter left the site in stream flow. A nutrient cycle in which outputs are low and internal recycling of nutrients is high (the loop from soil to vegetation and back to soil) is called a closed nutrient cycle.

Table 4 A model of ecological succession showing relative changes in energy flow, nutrient cycling, and diversity over time

Ecosystem trait	*Ecosystem status*	
	Developing	*Mature*
Energetics		
Net primary production	High	Low
Food chains	Linear	Weblike
Communities		
Species diversity	Low	High
Nutrient cycling		
Mineral cycles	Open	Closed
Nutrient conservation	Poor	Good
System dynamics		
Stability	Poor	Good

After the unperturbed patterns of growth and nutrient cycling were known, an entire Hubbard Brook watershed was clear-cut, followed by a dramatic change in ecosystem functioning. Stream flow increased by approximately 40% because water used by plants had been nearly eliminated by the forest harvest. The previously "closed" nutrient cycle of this forest became "open." After clear-cut, the concentrations of nitrogen (nitrate, NO_3^-) in the stream water draining the watershed increased approximately 60-fold. Concentrations of elements that are important to the biology of the ecosystem leaked into the streams and were exported from the ecosystem. Elements not essential for plant growth or required in very small amounts (e.g., sodium) were not lost to the same degree, indicating their cycling was not regulated by biotic activity of the forest. Odum (1969) predicted that nutrient losses would decline with increasing plant biomass and function. After the Hubbard Brook forest was allowed to regrow (undergo succession), nutrients resumed being absorbed by plants and nutrient losses to streams declined to near baseline levels. The Hubbard Brook ecosystem experiment informed forest management practices by providing a better understanding of how forest removal and regrowth affect the retention of soil nutrients and therefore the long-term productivity and diversity of the ecosystem.

Ecosystem Stability

Ecosystems are dynamic. They experience change in species composition and function in response to variations in climate and an array of disturbances. Fire, flood, drought, frost, and biological events such as the outbreak of pathogens and pests can "stress" ecosystems and alter their condition. Ecosystems vary widely in their responses to disturbance. The ability of an

ecosystem to withstand stress without a loss of function (resistance) or to recover rapidly from disturbance (resilience) is an important ecosystem trait. Some ecosystems appear very stable (high resistance and resilience) and their functioning is little affected by variations in factors external to the system. Ecosystems with high resilience are buffered against perturbation. Many ecosystems, however, show large decreases in productivity and biodiversity when disturbed. These ecosystems are "fragile" and have low resistance.

The relationship between ecosystem stability and diversity has been the subject of many field studies and theoretical tests using mathematical modeling. Ecologists have hypothesized that ecosystems with high biodiversity are more resistant (will experience less change) in response to a given level of disturbance and will also exhibit resilience – a high rate of recovery to predisturbance functioning (Folke *et al.*, 1996).

Does diversity influence the stability of ecosystem functioning? There is experimental evidence that it can do so (Chapin *et al.*, 1998). Several mechanisms have been proposed and tested to varying degrees to examine this relationship. Higher species diversity means that the trophic structure (feeding relationships among species) of the ecosystem is more complex, providing alternate pathways for energy flow within and between trophic levels (producers, consumers, and decomposers). Alternative pathways for energy transfers within the ecosystem could increase resistance to disturbance (species loss). Naeem and Li (1997) tested the hypothesis that redundancy (multiple species with similar functions in a food web) would stabilize ecosystem functioning by creating experimental microcosms with a varying number of species in each functional group. The simple systems contained producers (algae), decomposers (bacteria), and a primary and secondary consumer trophic level (protists) – the trophic structure of a typical aquatic ecosystem. Nutrient levels, light, and the number of species per trophic level were manipulated, and the biomass and density of the producers and decomposers were measured as indicators of ecosystem functioning. As the number of species in a trophic level increased, the biomass and density of replicate communities were more consistent. Thus, the communities with more species were more predictable in function (biomass production) and had higher reliability, i.e., the probability that an ecosystem will provide a given level of performance over a specified period of time. In contrast, the functioning of low-diversity ecosystems may be highly susceptible to decreases in biodiversity or species abundance. In the McMurdo Dry Valley polar desert ecosystems of Antarctica, a population decline in the dominant soil nematode *Scottnema lindsayae* was predicted to cause a 30% decline in soil carbon cycling function (Barrett *et al.*, 2008).

Higher species diversity may ensure functioning by reducing the risk of invasion by species that have the capacity to alter the structure or function of the ecosystem. An example is the higher resistance of species-rich natural systems to pest outbreaks compared to low-diversity agricultural ecosystems growing under the same environmental conditions. The spatial arrangement of individuals in an ecosystem can affect their risk to disease, predation, or consumption. In higher diversity systems the mean distance between individuals of the same species is on average greater than that of low-diversity systems. The wider spacing of individuals acts to slow the movement of pathogenic organisms, which should limit the occurrence of pest outbreaks that alter the performance of the ecosystem. These and other observations lead to the general expectation that diversity increases the resistance of ecosystems to disturbance.

The benefits of biodiversity to ecosystem functioning should be multiple since the processes of production and nutrient cycling are coupled by the biological interactions of organisms. The response of a Minnesota grassland to a severe drought (disturbance) illustrates this principle (Tilman *et al.*, 1996). In 1987 and 1988, a drought decreased productivity of the grassland. The species diversity of experimental plots prior to the drought explained the degree of productivity loss. Diverse plots experienced about a 50% decline in productivity, whereas productivity in the least diverse plots declined by more than 90%. The greater resistance of the higher diversity plots resulted from compensatory increases in productivity shown by drought-resistant species. The more diverse plots also had lower concentrations of nitrate in the rooting zone, indicating a more efficient use of this limiting resource.

This experiment demonstrates that species diversity has an effect on productivity and nutrient cycling and that declining species diversity influences these functions. However, we lack an understanding of the mechanisms producing these patterns of ecosystem response to disturbance and biodiversity change. Increasing diversity may increase the chance that a single drought-adapted and productive species will be present in the community, ensuring relatively high productivity. Alternatively, higher diversity may provide for a more efficient utilization of limiting resources, as suggested by the lower soil nitrate in more diverse plots. Before the basic relationships between biodiversity and ecosystem functioning can be more fully formalized, we need more detailed information on the critical levels (thresholds) of diversity associated with specific ecosystem functions and how environmental conditions operating over time alter their relationship (Folke *et al.*, 1996). The ability of societies to adapt to uncertainty from rapid environmental change will require resilience-based ecosystem management strategies that sustain essential ecosystem functions and provide for the delivery of critical ecosystem services (Chapin *et al.*, 2009).

Conclusions

Humans have become major agents of environmental change and influence the biodiversity and structure of ecosystems in many ways. Air pollution, clearing of natural systems for agriculture, forestry and urban development, the spread of exotic species, changes in the composition of the atmosphere, and other anthropogenic influences are altering ecosystem functioning. By changing ecosystem biodiversity and altering the processes that biota mediate, we significantly decrease the ability of ecosystems to provide services and resources for our use. The management of ecosystems for sustained levels of services and the restoration of damaged ecosystems will require greater knowledge about the role that species play in ecosystem functions related to production and nutrient cycling. Although we cannot know with certainty the roles of

most species in ecosystems, it is prudent to assume that the entire biodiversity is essential to ecosystem function and stability, and should be valued and protected.

See also: Ecosystem, Concept of. Ecosystem Services. Energy Flow and Ecosystems. Keystone Species. Nitrogen, Nitrogen Cycle

References

Barrett JE, Virginia RA, Wall DH, and Adams BJ (2008) Decline in a dominant invertebrate species contributes to altered carbon cycling in a low-diversity soil ecosystem. *Global Change Biology* 14: 1734–1744.

Chapin III FS, Kofinas GP, and Folke C (eds.) (2009) *Resilience-Based Natural Resource Management in a Changing World*. New York, NY: Springer.

Chapin III FS, Sala OE, Burke IC, *et al.* (1998) Ecosystem consequences of changing biodiversity. *BioScience* 48: 45–52.

Daily GC (ed.) (1997) *Nature's Services. Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.

Folke C, Hollings CS, and Perrings C (1996) Biological diversity, ecosystems, and the human scale. *Ecological Applications* 6: 1018–1024.

Gleissman SR (1998) *Agroecology: Ecological Processes in Sustainable Agriculture*. Chelsea, MI: Sleeping Bear Press.

Hooper DU and Vitousek PM (1997) The effects of plant composition and diversity on ecosystem processes. *Science* 277: 1302–1305.

Likens GE and Bormann FH (1995) *Biogeochemistry of a Forested Ecosystem*, 2nd edn. New York, NY: Springer-Verlag.

Mäder P, Fließach A, Dubois D, Gunst L, Fried P, and Niggli U (2002) Soil fertility and biodiversity in organic farming. *Science* 296: 1694–1697.

Menge BA, Berlow EL, Blanchette CA, Navarrete SA, and Yamada SB (1994) The keystone species concept: Variation in interaction strength in a rocky intertidal habitat. *Ecological Monographs* 64: 249–286.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Biodiversity Synthesis*. Washington, DC: World Resources Institute.

Naeem S and Li S (1997) Biodiversity enhances ecosystem reliability. *Nature* 390: 507–509.

Odum EP (1969) The strategy of ecosystem development. *Science* 164: 262–270.

Sterner RW and Elser JJ (2002) *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton, NJ: Princeton University Press.

Tilman D, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379: 718–720.

Virginia RA and Wall DH (1999) How soils structure communities in the Antarctic Dry Valleys. *BioScience* 49: 973–983.

Vitousek PM, Aber JD, Howarth RW, *et al.* (1997b) Human alteration of the global nitrogen cycle: Sources and consequences. *Ecological Applications* 7: 737–750.

Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997a) Human domination of the earth's ecosystems. *Science* 277: 494–499.

Wall DH and Virginia RA (1997) The world beneath our feet: Soil biodiversity and ecosystem functioning. In: Raven P and William TA (eds.) *Nature and Human Society: The Quest for a Sustainable World*, pp. 225–241. Washington, DC: National Academy of Sciences Press.

Ecosystem Services

Heather Tallis, Anne Guerry, and Gretchen C Daily, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Gretchen Daily, Shamik Dasgupta, volume 2, pp 353–362, © 2001, Elsevier Inc.

Glossary

Ecosystem services The wide array of benefits that ecosystems, and their biodiversity, confer on humanity.

Marginal value The economic value of the next incremental unit of something. In this context, marginal values are those associated with managing the next small unit of an ecosystem in a particular way (e.g., preserving, rather than clearing, the next unit of forest). They can also be the partial contribution of natural capital to a final good that is produced with other inputs. For example, the marginal value of irrigation water for crop production is the value of the incremental crop yield that can be attributed to irrigation, rather than to labor, fertilizer inputs, etc.

Natural capital Here we focus on living, renewable forms of natural capital, which constitute a stock – of an ecosystem and the biota that makes it up – that generates a flow of ecosystem services. For example, a forest constitutes a stock that generates a flow of timber, carbon sequestration, water quality, biodiversity, serenity, and other benefits, depending upon how it is managed. (Fossil fuels and other minerals constitute non-living natural capital, which is generally non-renewable on time scales of interest to society.)

Introduction

Ecosystem services are essential to sustaining and fulfilling human life, and yet their supply is seriously threatened by the intensification of human impacts on the environment. This article provides an overview of issues concerning the identification, biophysical and economic characterization, and safeguarding of ecosystem services.

Over the past decade, efforts to value and protect ecosystem services have been promoted by many as the last best hope for making conservation mainstream – attractive and commonplace worldwide. In theory, if institutions recognize the values of Nature, then we can greatly enhance investments in conservation and foster human well-being at the same time. In practice, scientific and policy communities have not yet developed the scientific basis or the policy and finance mechanisms for integrating natural capital into resource and land-use decisions on a large scale. Here we describe the history, progress, and key frontiers for advancing the science and practice of accounting for natural capital in the decisions of individuals, communities, corporations, and governments.

What are Ecosystem Services?

Definition and Classification

Ecosystem services are defined simply as the benefits that people obtain from ecosystems (MA, 2005). They sustain and fulfill human life and flow from many conditions and processes of ecosystems, and the species making them up (Daily, 1997). The processes and features generating ecosystem services (ES) are so tightly interconnected that any classification is inherently somewhat arbitrary. The most widely used classification was developed through the Millennium Ecosystem Assessment (MA) and identifies four classes of ES based on their types of benefits to society:

1. *provisioning services* including the production of goods such as food, water, timber, and fiber;
2. *regulating services* that stabilize climate, moderate risk of flooding and disease, and protect or enhance water quality;
3. *cultural services* that provide recreational, esthetic, educational, community, and spiritual opportunities; and
4. *supporting services* that underlie provision of the other three classes of benefits, including soil formation, photosynthesis, nutrient cycling, and the preservation of options (**Figure 1**; MA, 2005).

The classification of ES is still a topic of debate and several other classification approaches have been suggested (e.g., De Groot *et al.*, 2002).

The ecosystem services framework is system- and scale-neutral, applying equally to terrestrial, freshwater and marine ecosystems and their processes, anywhere on the spectrum from relatively pristine to heavily managed conditions. Indeed, all ecosystems provide, to differing degrees, a set of ES. Human conversion of ecosystems from one type to another is often motivated by a desire for a different set of ES, though consideration of the services of the two systems and their tradeoffs is often incomplete.

Ecosystem Services Across Systems

In the terrestrial realm, croplands, natural and managed forests, grasslands, and wetlands can provide a wide range of ES. In each of these systems, for example, vegetation can protect and enhance soils, preventing their loss through erosion and

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00284-7

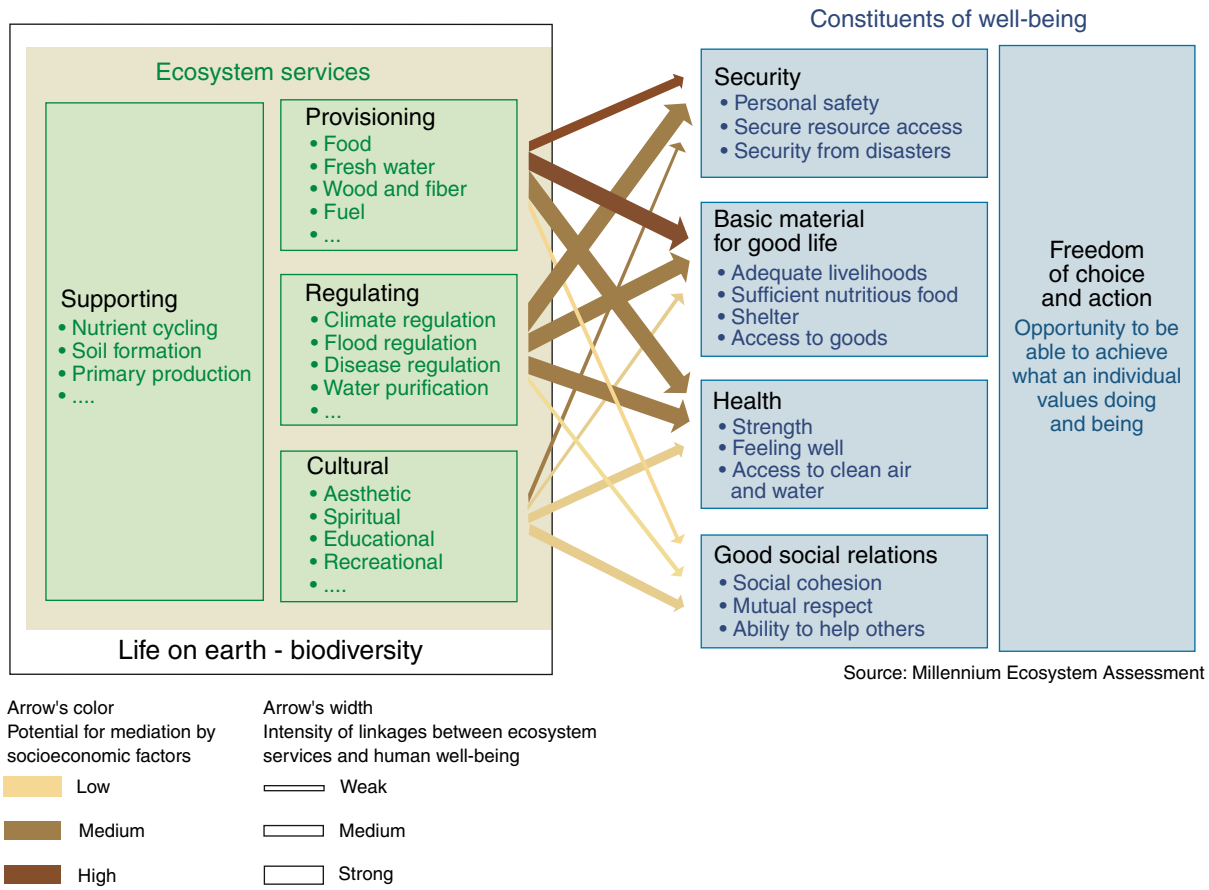


Figure 1 Ecosystem service categories and their linkages to human well-being as described in the MA (Millennium Ecosystem Assessment) (2005) *Ecosystems and Human Well-being: The Assessment Series (Four Volumes and Summary)*. Washington, DC: Island Press, with permission from World Resources Institute.

improving fertility by retaining moisture and storing and recycling nutrients. Vegetation and soils together regulate the quantity, quality, and timing of water flows, thus moderating floods and droughts and providing cleaner, more reliable supplies.

Forests stand out as important in regulating water and carbon cycles in their strong influence on local, regional and global climate and because of the multiple, interacting threats to their future. They also provide natural products for subsistence use or sale including timber, firewood, mushrooms, fruits and seeds, medicinal plants, rubber, cork, and bushmeat. Forest and woodland habitats harbor species that provide pollination and pest control to commercial or subsistence crops. Grassland and other dryland systems play these same critical roles in addition to supporting vast livestock populations (MA, 2005). Wetlands occupy a small fraction of Earth's surface, but dominate the landscape where they are concentrated and provide a wide array of water quality, flood mitigation, coastal protection, and biogeochemical services (MA, 2005). Each of these systems, however natural or managed, can provide habitat for biodiversity and opportunities for recreational activities, spiritual experiences, and creative, cultural expression.

Freshwater ecosystems provide a suite of highly visible and widely appreciated ES. The freshwater regulated by terrestrial systems and the atmosphere is used for drinking, hydropower production, irrigation, household activities (washing etc.), industrial purposes (cooling, manufacturing, etc.), and cultural experiences. People also gain large revenues and nutrition from freshwater fisheries and aquaculture. Less appreciated is the value of sediment transport and deposition in rivers that supply river reaches and downstream beaches with important sand and gravel resources. Wetlands and other aquatic vegetation can regulate flood waters and cycle nutrients, improving water quality. Finally, freshwater systems serve as pathways for human transportation and recreational or cultural activities.

Marine ecosystems also provide all four classes of ecosystem services described in the MA. Marine fisheries and aquaculture provide nutrition, feed for animals, livelihoods and important recreational and cultural opportunities. Harvests of other species for food additives, cosmetics, and pharmaceuticals also support health, nutrition, and livelihoods. Marine biogenic habitats (such as coral reefs, oyster reefs, and kelp forests) regulate natural hazards including storm surges, and may play a critical role in helping coastal communities adapt to sea level rise. Marine systems also transform, detoxify, and sequester wastes. In addition, oceans are the center of the global water cycle; they hold 96.5% of the earth's water and are a primary driver of the atmosphere's

temperature, moisture content, and stability. Oceans are also key players in the global cycles of carbon, nitrogen, oxygen, phosphorus, sulfur, and other major elements and are responsible for approximately 40% of global net primary productivity. Finally, coastal communities reap many benefits from coastal tourism (one of the world's most profitable industries), and numerous coastal communities define their very identities in relation to the sea and all it brings.

The Ecosystem Service Supply Chain

All of these services flow to people along a supply chain from biophysical systems to people (Tallis *et al.*, 2012). All services are generated by some function or element of a natural or managed system (**Figure 2**). The full suite of these elements or functions can best be considered in three discrete steps: Supply, service, and value. For example, consider protection from coastal storm surges. Many different types of coastal elements (e.g., coral reefs, mangroves, oyster beds, barrier islands) confer protection from storm surges by attenuating waves. The full set of locations of these coastal elements represents the supply of protection from storms. People do not receive storm surge protection from all of these locations, however, because some are far from human infrastructure and settlement. Both the distribution of human infrastructure and settlement, together with the location and condition of supply, are required to give a clear picture of how much "service" is actually delivered at a given time. Finally, the service delivered to human communities is often valued differently, depending on the context. For example, coastal protection services provided by nearshore habitats to easily accessible, popular, public beaches might be seen as more valuable than those to more remote sites.

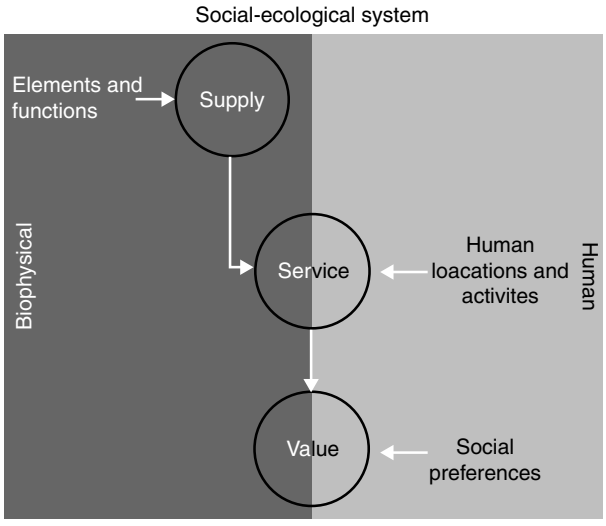


Figure 2 Three measurement points for ecosystem services (Reproduced from Tallis H, Lester SE, Ruckelshaus M, *et al.* (2012) New metrics for managing and sustaining the ocean's bounty. *Marine Policy* 36: 303–306.). Supply metrics deal only with the biophysical system underpinning the service of interest. Service metrics include critical information linking supply to beneficiaries. Value metrics weight the level of service based on people's preferences or social policy goals.

History of the Concept of Ecosystem Services

The nature and value of ecosystem services have been illuminated primarily through their disruption and loss. For instance, deforestation has demonstrated the critical role of forests in the hydrological cycle – in particular, in mitigating floods, droughts, the erosive forces of wind and rain, and the silting of dams and irrigation canals. Release of toxic substances, whether accidental or deliberate, has revealed the nature and value of physical and chemical processes, governed in part by a diversity of microorganisms, that disperse and break down hazardous materials. Thinning of the stratospheric ozone layer sharpened awareness of the value of its service in screening out harmful ultraviolet radiation. And the loss of coastal wetlands has brought into relief their importance in regulating coastal hazards such as hurricanes and tsunamis.

Initial Development of the Ecosystem Services Concept

A cognizance of ecosystem services, expressed in terms of their loss, dates back at least to Plato and probably much earlier:

> What now remains of the formerly rich land is like the skeleton of a sick man with all the fat and soft earth having wasted away and only the bare framework remaining. Formerly, many of the mountains were arable. The plains that were full of rich soil are now marshes. Hills that were once covered with forests and produced abundant pasture now produce only food for bees. Once the land was enriched by yearly rains, which were not lost, as they are now, by flowing from the bare land into the sea. The soil was deep, it absorbed and kept the water..., and the water that soaked into the hills fed springs and running streams everywhere. Now the abandoned shrines at spots where formerly there were springs attest that our description of the land is true. (– Plato)

Mooney and Ehrlich (1997) trace modern concern for ecosystem services to George Perkins Marsh, a lawyer, politician, and scholar. Indeed, his 1864 book *Man and Nature* describes a wide array of services, again, often expressed in terms of their loss. Remarking on the terrain of the former Roman Empire, he notes that it "is either deserted by civilized man and surrendered to hopeless desolation, or at least greatly reduced in both productiveness and population" (p. 9). He goes on to describe the reduction of hydrological services: "Vast forests have disappeared from mountain spurs and ridges, the vegetable earth … [is] washed away; meadows, once fertilized by irrigation, are waste and unproductive, because … the springs that fed them dried up; rivers famous in history and song have shrunk to humble brooklets" (p. 9). He also made connections between deforestation and climate: "With the disappearance of the forest, all is changed. At one season, the earth parts with its warmth by radiation to an open sky – receives, at another, an immoderate heat from the unobstructed rays of the sun. Hence the climate becomes excessive, and the soil is alternately parched by the fervors of summer, and seared by the rigors of winter. Bleak winds sweep unresisted over its

surface, drift away the snow that sheltered it from the frost, and dry up its scanty moisture" (p. 186). Finally, he even wrote of decomposition services: "The carnivorous, and often the herbivorous insects render an important service to man by consuming dead and decaying animal and vegetable matter, the decomposition of which would otherwise fill the air with effluvia noxious to health" (p. 95).

Following World War II, other eloquent writers on the environment emerged, including Fairfield Osborn (*Our Plundered Planet*, 1948), William Vogt (*Road to Survival*, 1948), and Aldo Leopold (*A Sand County Almanac and Sketches from Here and There*, 1949). Each discusses ecosystem services without using the term explicitly. In *The Population Bomb* (1968), Paul Ehrlich describes anthropogenic disruption of ecosystems and the societal consequences of doing so, addressing the need to maintain important aspects of ecosystem functioning. Along these lines, the *Study of Critical Environmental Problems* (1970) presents a list of key "environmental services" that would decline with a decline in "ecosystem function." This list was expanded on by Holdren and Ehrlich (1974). Meanwhile, in the 1960s and 1970s, economists set out to measure "the value of services that natural areas provide," with efforts focused on agricultural production and environmental amenities.

By the early 1980s, efforts were initiated to investigate two questions: The extent to which ecosystem function (and the delivery of services) depends on biodiversity, and the extent to which technological substitutes could replace ecosystem services. The first question is addressed in Biodiversity and Ecosystem Services in this volume. The second question was tackled by Ehrlich and Mooney (1983). Work on these topics proliferated and, in 1997, a collective effort was made to synthesize the wealth of scientific information that had accumulated on the functioning of ecosystem services, with a preliminary exploration of their economic value, and of key issues meriting further work (Daily, 1997).

Recent Advances

Four major advances of the last decade have revitalized research on ecosystem services and brought them into the public eye. First, the MA represented a visionary and seminal step in global science – it was the first comprehensive global assessment of the status and trends of all of the world's major ecosystem services. It was requested by United Nations Secretary General Kofi Annan in 2000 and carried out between 2001 and 2005 with contributions from more than 1360 experts worldwide. The key finding of this assessment was that two-thirds of the world's ecosystem services were declining (MA, 2005). This captured the attention of world leaders and emphasized the connections between human decisions and the natural environment that feedback to the human condition via changes in the flow of ecosystem services.

Work following the MA clarified this chain of connections (**Figure 3**) (Daily *et al.*, 2009). Human decisions shape individuals' actions relating to the use of land, water, oceans, and other elements of natural capital. These actions often alter the state or functioning of ecosystems, which in turn provide altered flows of benefits (goods or services) to people. People express different values (monetary, cultural) associated with these altered streams of benefits and it is the expression of these values that leads to changes in institutions that guide decisions. The following three recent advances all concern the connections in this flow.

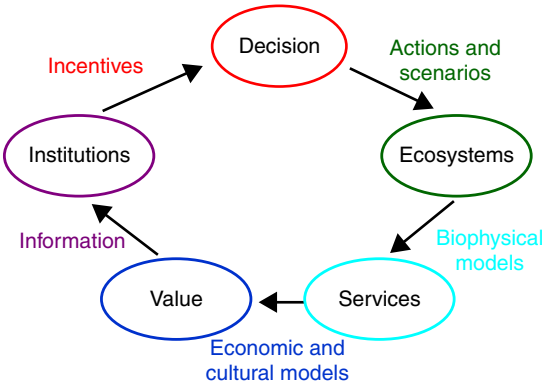


Figure 3 A framework showing key elements for integrating ES into decision-making (reproduced from Daily GC, Polasky S, Goldstein J, *et al.* (2009) Ecosystem services in decision-making: Time to deliver. *Frontiers in Ecology in Ecology and the Environment* 7: 21–28). One could link any two ovals, in any direction, in different decision contexts.

We have recently seen large advances in our understanding of the links between ecosystem functions and processes and the provision of ecosystem services (**Figure 3**). For some ecosystem services, we now better understand the key ecological system components that drive provision and we can now measure and model, with uncertainty, the impacts of land use and resource management decisions on a wider variety of ecosystem processes and associated services. Ecological science has also advanced spatially explicit modeling, which is essential for mapping ecosystem services and their flows to people (e.g., Chan *et al.*, 2006). Finally, we are starting to see patterns in how multiple ecosystem services and biodiversity change in relation to each other. Recent work has started characterizing bundles of ecosystem services, and exploring their synergies and tradeoffs (e.g., Bennett *et al.*, 2009).

Further, economic valuation methods have been applied to the spatial provision of ecosystem services to estimate the monetary value of benefits and, in some cases, the distribution of those benefits to various segments of society (NRC, 2005). In addition, qualitative and quantitative methods from other social sciences have been applied to gain better understanding of the social and cultural importance of ecosystem services (e.g., MA, 2005).

Lastly, experiments in payments for ecosystem services in ecosystem-based management and in regional planning have begun, giving us opportunities to learn about how science can play a role in altering institutions, and how institutions alter decisions and the resulting flow of ecosystem services. The following section describes some of these efforts in more detail.

Incorporating Ecosystem Services into Decisions

The urgent challenge today is to move from theory to practical implementation of ES tools and approaches in resource decisions taken by individuals, communities, corporations, and

governments. The framework in **Figure 3** connects the science of quantifying services with valuation and policy to devise payment schemes and management actions that take account of ecosystem services. This connection is expressed in the real world in a variety of ways across scales from local to global.

Over the past decade, a great number and diversity of efforts to implement the ES framework have emerged worldwide. Individually, most of these efforts are small and idiosyncratic. But collectively, they represent a powerful shift in the focus of conservation organizations and governments (primarily) toward a more inclusive, integrated, and effective set of strategies. Taken together, these efforts span the globe and target a full suite of ecosystem services, including principally forest-generated services of carbon sequestration, water supply, flood control, biodiversity conservation, and enhancement of scenic beauty (and associated recreation/tourism values).

Many local or regional ES efforts focus on a single service that stands out as sufficiently important, from economic and political perspectives, to overcome the activation energy required to protect it. Under the institutional umbrella created for the focal service, it is possible that other services may be at least partially protected. Beginning in the late 1990s, large-scale investment in natural capital for water flow regulation in China – and for a broad suite of ES in Costa Rica – set pioneering examples that are now being adapted elsewhere and scaled up.

The brief summaries offered next illustrate several models of success at different scales and in different kinds of social-ecological systems. In each case, there is an acute or looming crisis, innovative leadership, and pursuit of dual goals: Improving both human and ES condition.

Local Scale: Water Funds

New York City made one of the first and most famous investments in ecosystem service provision in recent history, in the mid-1990s. The city invested *ca.* US$1.5 billion in a variety of watershed protection activities to improve drinking water quality for 10 million users rather than spending the estimated US$6–8 billion needed (excluding annual operating and maintenance costs) for building a new filtration plant. This seminal example is widely cited as evidence of the business case for investing in natural capital instead of built capital. Yet the effort remains very much an experiment in the science and policy of investing in natural capital, and one on which there is international focus.

Globally, watersheds are now emerging as the target of a range of creative policy and finance mechanisms that link beneficiaries to suppliers through a payment system. In these ''water funds,'' water users voluntarily pay into a pool that is collectively managed by contributors and invested in watershed management improvements. The Nature Conservancy (TNC) has now established more than 10 water funds in Latin America, has plans to create 22 more by 2015 and is exploring the possibility of establishing the first funds in Africa.

One of the established water funds, Agua por la Vida y la Sostenibilidad, demonstrates the diversity of water users that are becoming engaged in these funds and the kinds of watershed management changes these funds motivate. Formally established in the Cauca Valley, Colombia in 2009, this water fund is supported by the region's sugar cane grower's association (PROCAÑA), the sugar producers' association (ASOCAÑA), 11 local watershed management groups, TNC and a Colombian peace and justice nongovernment organization. Each member of the water fund voluntarily pays a self-determined amount into the fund that is then jointly managed by the members to improve landscape management in 11 watersheds covering over 3900 km^2.

Members in this fund have currently committed to contributing US$10 million over 5 years to be invested in five kinds of management changes: Protection of native vegetation, restoration of denuded lands, enrichment of degraded forests, fencing of rangelands, and implementation of silvopastoral practices. The fund is starting a monitoring program that will ensure that these investments lead to measurable improvements in water quality for approximately 1 million water users downstream and significant improvements in terrestrial and freshwater biodiversity.

Local Scale: Coastal and Marine Spatial Planning

People have historically thought of oceans as relatively featureless expanses that defy the drawing of lines on maps. However, recent political and scientific advances have highlighted the need for a comprehensive approach to planning marine and coastal uses and the need for practical tools to make this more comprehensive approach a reality on the ground and in the water. In a marine spatial plan, a wide range of uses of the marine environment are put on one map. But an understanding of how such plans are likely to yield changes in the delivery of the broad range of services people receive from the system has, until recently, remained elusive.

Along the west coast of Vancouver Island Canada, multiple, often competing interests are struggling to define the future character of the place. Existing extractive, industrial, and commercial uses; traditional First Nations subsistence and ceremonial uses; recreation and tourism; and emerging ocean uses such as the extraction of wave energy are all in the mix. The West Coast Aquatic Management Board (WCA) is charged with creating a marine spatial plan for the region. WCA is a public–private partnership with participation from four levels of government (Federal, Provincial, local, and First Nations), and diverse stakeholders. Ultimately, WCA's vision is to manage resources for the benefit of current and future generations of people and nonhuman species and communities.

Some key pillars of the partnership's strategy are to: use a precautionary, ecosystem-based approach to protect, maintain, and restore marine and coastal resources; respect and protect First Nations' food, social and ceremonial requirements and treaty obligations; integrate expertise and knowledge from First Nations, local, scientific, and other sources; ensure broad participation in the planning process; and foster initiatives that maintain or enhance opportunities for coastal communities to benefit from local resources, while achieving sustainable social, cultural, and economic benefits for the region. WCA has partnered with the Natural Capital Project to

explore how alternative spatial plans might affect a wide range of ES and to provide information about tradeoffs among ES.

Some key considerations for WCA and their stakeholders include balancing important industrial and commercial activities (such as shipping, mining, logging, aquaculture, and fisheries), increased development of tourism and recreation, renewable energy generation, and a strong cultural desire for sustaining the remote, wild feeling of the place. WCA is exploring the suitability of alternative regions for these different activities. For example, maps of coastal vulnerability to erosion and flooding from storm surge are helping to direct coastal development permits to low-risk areas. Similar maps of the value of captured wave energy are being overlaid with existing ocean uses (e.g., fishing and recreational activities) to highlight regions of high wave energy value, where wave energy generation facilities might be constructed while having minimal impacts on other activities. Examinations of tradeoffs among aquaculture (finfish, shellfish), wild salmon fisheries, recreation (e.g., kayaking, whale watching, and diving), coastal development (on the coast, as well as floathomes), and habitat and water quality are underway.

The general framework of ES and ES modeling, in particular, are helping to articulate connections between human activities that are often considered in isolation, to align diverse stakeholders around common goals, and to make implicit decisions explicit. ES modeling results have informed early iterations of the marine spatial plan and will inform the creation of the final plan in 2012.

National Scale: Land-Use Planning and Human Development in China

The ecosystem service investments being made in China today are impressive in their goals, scale, duration, and innovation. Following massive droughts and flooding in 1997–1998, China implemented several national forestry and conservation initiatives, into which investments exceeded 700 billion yuan (*ca.* USD100 billion) over 2000–2010 (Liu *et al.*, 2008). The larger and older of these initiatives are being rigorously evaluated to determine their biophysical and socioeconomic impacts, to improve their design and efficacy.

These initiatives have dual goals: To secure critical natural capital through targeted investments across landscapes and regions, and to alleviate poverty through targeted wealth transfers from coastal provinces to inland regions where many ES originate. The Chinese government aims to reduce the loss of soil, improve water retention, reduce desertification, and generally protect biodiversity and ecosystems in the west of the country for flood control, hydropower production efficiency, irrigation supply, more productive agriculture, and ecotourism. In addition, it wants to change the economic structure in rural areas to increase local household income while simultaneously making local households' patterns of land utilization and agricultural production more sustainable (Liu *et al.*, 2008).

The initiatives include two national PES programs, the Natural Forest Conservation Program (NFCP) and the Sloping Land Conversion Program (SLCP), established in 1998 and 1999, respectively. Implementation was tested in a few provinces, and then rapidly scaled to the whole country. Evaluation of the programs shows significant achievement of the biophysical goals, with remarkably rapid land conversion in the desired directions. For example, by the end of 2006, the SLCP had converted *ca.* 9 million hectares of cropland into forest/grassland and had afforested *ca.* 12 million ha of barren land. Village level field measurements have shown not only that the payments for ES have altered land-use patterns, but in turn soil erosion has been decreased in some areas by as much as 68% (Cao *et al.*, 2009).

Overall social impacts of the programs are mixed. In some places, payment levels and types are leading to improvements in economic measures of well-being, whereas in others payments were not sufficient to compensate for loss of income from shifting livelihoods (Liu *et al.*, 2008). In addition, in some places where participation in the SLCP has significant positive impacts on household income, it has not yet transferred labor toward nonfarming activities as the government wished. Payments are now being adjusted to improve success in achieving goals of poverty alleviation and growth of new economic sectors in rural areas.

China is also now in the process of establishing a new network of Ecosystem Function Conservation Areas (EFCAs), specifically for ES provision. Their exact delineation is now being determined through quantitative ecosystem service mapping and valuation. They are expected to span *ca.* 25% of the country.

The current and potential future impacts of ES investments in China are enormous, certainly within the country – and also globally, in the form of enhanced carbon sequestration and reduced dust export, and perhaps most importantly in lessons on making the investments needed in natural capital and human well-being everywhere.

International Scale: Global Policy and Research Efforts

As described above, the MA was the first major effort to establish ES in the international policy arena. Activities stemming from that effort are now aimed at bringing countries together in making tangible commitments to safeguard ES (e.g., 2020 targets for the Convention on Biodiversity) and to assess national and international progress toward those commitments (e.g., through Group on Earth Observations Biodiversity Observation Network (GEO BON) and the Programme on Ecosystem Change and Society (PECS), which synthesize knowledge for the International Platform for Biodiversity and Ecosystem Services (IPBES), formally established in 2010). Several new international research efforts aim to feed into these international processes, including the Natural Capital Project, The Resilience Alliance, and the Stockholm Resilience Centre. Other entities are focused on establishing and tracking ES markets, as one mechanism for bringing larger attention to ES benefits to society (e.g., The Katoomba Group and The Ecosystem Marketplace, both initiated by Forest Trends). As one example of many burgeoning international efforts, we describe in greater detail the Natural Capital Project.

The Natural Capital Project (NatCap) (www.naturalcapitalproject.org) is an international partnership working to align economic forces with conservation, by developing tools that

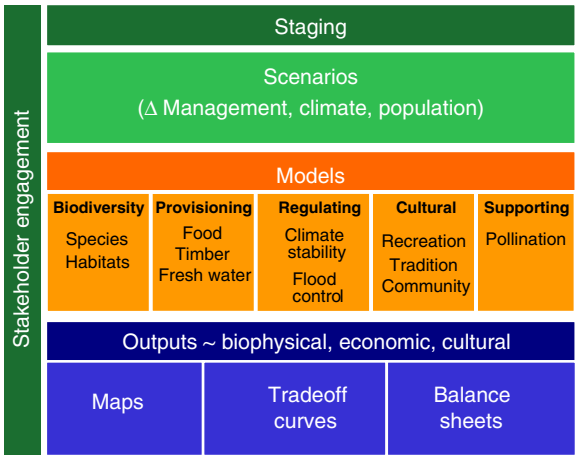


Figure 4 An iterative process for integrating ecosystem services into decisions. The process begins with stakeholder engagement around impending decisions, with a focus on realistic, alternative scenarios for the future. The modeling is shaped by stakeholders, and typically focused on the services and scenarios deemed most important. Outputs are displayed in accordance with stakeholder preferences, in the form of maps, tradeoff curves, and/or balance sheets. These can be expressed in biophysical (e.g., tons of carbon), economic (e.g., dollars), or cultural (e.g., visitor-days) terms. Reproduced from Daily GC, Polasky S, Goldstein J, *et al.* (2009) Ecosystem services in decision-making: Time to deliver. *Frontiers in Ecology and the Environment* 7: 21–28.

make incorporating natural capital into decisions easy and replicable; by demonstrating the power of these tools in important, contrasting places; and by engaging leaders globally. NatCap is developing InVEST, a family of tools for Integrated Valuation of Ecosystem Services and Tradeoffs.

InVEST helps decision makers visualize the impacts of potential policies by modeling and mapping the delivery, distribution, and economic value of ES under alternative scenarios. The outputs identify tradeoffs and compatibilities between environmental, economic, and social benefits. InVEST is designed for use as part of an active decision-making process (**Figure 4**) and can be applied at local, regional, or global scales. The first phase of the approach involves working with stakeholders to identify critical management decisions and to develop scenarios that project how the provision of services might change in response to those decisions as well as to changing climate, population, etc. Based on these scenarios, a modular set of models quantifies and maps ES. The outputs of these models provide decision makers with information about costs, benefits, tradeoffs, and synergies of alternative investments in ES provision.

NatCap is using InVEST in major natural resource decisions in diverse contexts around the world, including in the three examples given above (water funds, coastal and marine spatial planning, and land-use planning and human development in China). The aim is to demonstrate the power of these approaches and to learn how to replicate and scale up models of success. The Project is engaged in a suite of international efforts, including GEO BON and IPBES, to offer a common, unifying platform for regional and national efforts that are spawned by these initiatives.

Outstanding Challenges

With the rapid rate of development of ES mapping, from the biophysical and economic modeling through to policy application in diverse socioeconomic contexts, it is likely that great advances will be made in coming years. What we report here is only a beginning. There are key arenas in which further learning is crucial to understand what drives variation in the provision of ES, how they percolate through various arms of society and how social reaction leads to sustainable or unsustainable change in ES provision.

Relating Ecosystem Services and Human Health

The relationships between biophysical attributes of ecosystems and human health are complex. Destruction of natural ecosystems can at times improve aspects of public health. Draining swamps, for example, can reduce habitat for the mosquito vector that transmits the parasite that causes malaria. However, destruction of other systems can have sharp negative consequences for human health. There is emerging evidence that loss of tropical forests, for example, leads to an increase in transmission of malaria (Patz and Olson, 2006). Similarly, fragmentation of, and biodiversity loss from, eastern North American forests is associated with an increase in lyme disease (Keesing *et al.*, 2010). Natural and managed ecosystems provide many services that sustain human health, through provision of human nutrition (especially of protein and micronutrients); purification and regulation of drinking water; regulation of air quality; regulation of vector-borne disease; and psychological benefits. There is a great need for research illuminating the links between biodiversity, ecosystem conditions and processes, and human health.

Tradeoffs and Synergies

The relationship between ES and biodiversity and among different ES varies with socio-ecological context. In some cases, clear tradeoffs and synergies among services have been defined in specific contexts, but there is still much to be learned about what determines the nature of these relationships. Advancing this knowledge is essential because policies addressing management change can only be successful if management controls ES relationships. If policies are established to align multiple ecosystem services, but biophysical conditions in the system lead to innate tradeoffs among services, management changes are bound to fail in delivering the desired improvements to social benefits.

Distributional Effects

Much of the science of mapping ES has focused on identifying where ES are generated and where they are delivered. However, less work has focused on identifying to whom ES actually flow. This connection is essential if policies addressing ES delivery are to be equitable and either improve the well-being of the poor or avoid unintended distributional consequences. Past work in this arena has focused on overlaying maps of ES provision with an array of poverty indicators. Missing from this spatial analysis

is information on access to and ability to control the delivery of ES. In many cases (e.g., for services such as clean drinking water, hydropower production, agriculture, water for irrigation, wave power generation), the actual delivery of services to specific people is affected by the location of infrastructure or institutions regulating access to resources. New science is needed that allows the ready mapping of these connections and the prediction of how they will change under future conditions.

Dynamic Effects – and Shocks and Uncertainty

Dynamic changes, such as in climate and in the nitrogen cycle – as well as changes arising through economic development and evolving human preferences over time – are very important. The possibility of feedbacks within ecosystems, and between ES and human behavior, is a key area for further development. Feedback effects can give rise to thresholds and rapid changes in systems that can fundamentally alter system outcomes. The ability to incorporate shocks and the possibility of surprises is another area where further development is needed. Fires, droughts, and disease all can have major influences on ecosystems and affect the services produced. Changes in economic conditions or fads in human behavior can similarly cause major changes in systems (e.g., financial crises). The occurrence of each of these and other potential disturbances is difficult to predict but virtually certain to come about. Understanding their likely impacts on ecological and social systems will help us prepare for them.

Valuation in Monetary and Non-Monetary Terms for Decision-Making

Monetary valuation of ES is not nearly as prevalent as sometimes assumed. More typically, real-world applications of the ES framework rely on biophysical values to inform policy design, such as measures of water quality or flood risk.

Value is not always easily characterized or fully captured in monetary terms, so it is important to characterize value in multiple dimensions, including health, livelihood support, cultural significance, etc. (e.g., Dasgupta, 2001). This will help ensure that valuation and broader decision-making approaches are inclusive of the range of benefits and people concerned. Interdisciplinary efforts are presently underway to create a conceptual framework that is useful both in theory and in practice for a broad suite of cultural ES.

Institutional Design

However ES are measured, there is a need for political and social science research to design institutions and mechanisms that better capture externalities. Efforts such as national accounts are blossoming now, but it is unclear how they will evolve and how successful governments will be at incorporating natural capital into national measures of wealth. There is great work to be done in determining the merits and limitations of alternative policy and finance mechanisms, in different economic, governance, and other social contexts (e.g., Ostrom, 2005). There is also great work to be done in developing institutions that achieve representation and participation by stakeholders as part of adaptive governance systems (e.g., Cowling *et al.*, 2008).

Conclusions

Ecosystem services have had a relatively long history through indirect recognition of the importance of nature for the persistence of the human endeavor. There are scientific challenges for biogeochemists, hydrologists, ecologists, economists, political scientists, epidemiologists, anthropologists, and other social scientists to understand how human actions affect ecosystems, the provision of ES and the value of those services. At least as demanding are the social and political questions associated with incorporating this understanding into decision-making. There is also a need to design effective and enduring institutions to manage, monitor, and provide incentives that reflect the social values of ecosystem services. Information is becoming more readily available for individuals, corporate managers, and government officials who make decisions affecting ecosystems and their services to consider a more complete set of costs and benefits associated with their choices. We are likely to see continuing growth in our scientific ability to measure and predict changes in ES, our ability to design policies and institutions that accurately represent these changes and in turn, the ability of the environment to continue providing the many benefits society needs to prosper.

See also: Biodiversity and Ecosystem Services. Biodiversity and Pest Control Services. Biodiversity in Plant Breeding. Dormancy and Diapause. Evaluation of Ecosystem Service Policies from Biophysical and Social Perspectives: The Case of China. Modeling Marine Ecosystem Services. Priority Setting for Biodiversity and Ecosystem Services. Valuing Ecosystem Services

References

Bennett E, Peterson G, and Gordon L (2009) Understanding relationships among multiple ecosystem services. *Ecology Letters* 12: 1394–1404.

Cao S, Zhong B, Yue H, Zeng H, and Zeng J (2009) Development and testing of a sustainable environmental restoration policy on eradciating the poverty trap in China's Changing County. *Proceedings of the National Academy of Sciences* 106: 10712–10716.

Chan K, Shaw R, Cameron D, *et al.* (2006) Conservation planning for ecosystem services. *Public Library of Science, Biology* 4: 2138–2152.

Cowling R, Egoh B, Knight AT, *et al.* (2008) An operational model for mainstreaming ecosystem services for implementation. *Proceedings of the National Academy of Sciences* 105: 9483–9488.

Daily GC (ed.) (1997) *Nature's Services*. Washington, DC: Island Press.

Daily GC, Polasky S, Goldstein J, *et al.* (2009) Ecosystem services in decision-making: Time to deliver. *Frontiers in Ecology and the Environment* 7: 21–28.

Dasgupta P (2001) *Human Well-Being and the Natural Environment*. Oxford: Oxford Univ. Press.

De Groot D, Wilson MA, and Boumans RMJ (2002) A typology for the classification, description and valuation of ecosystem functions, goods and services. *Ecological Economics* 41: 393–408.

Ehrlich P and Mooney H (1983) Extinction, substitution and ecosystem services. *BioScience* 33: 248–254.

Holdren J and Ehrlich P (1974) Human population and the global environment. *American Scientist* 62: 282–292.

Kareiva PK, Tallis H, Ricketts TH, Daily GC, and Polasky S (eds.) (2011) *Natural Capital: Theory & Practice of Mapping Ecosystem Services*. Oxford: Oxford University Press.

Keesing F, Belden LK, Daszak P, *et al.* (2010) Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468: 647–652.

Kremen C (2005) Managing for ecosystem services: What do we need to know about their ecology? *Ecology Letters* 8: 468–479.

Liu J, Li S, Ouyang Z, *et al.* (2008) Ecological and socioeconomic effects of China's policies for ecosystem services. *Proceedings of National Academy of Sciences* 105: 9489–9494.

MA (Millennium Ecosystem Assessment) (2005) *Ecosystems and Human Well-being: The Assessment Series (Four Volumes and Summary)*. Washington, DC: Island Press.

Mooney HA and Ehrlich PR (1997) Ecosystem services: A fragmentary history. In: Daily GC (ed.) *Nature's Services*, pp. 11–19. Washington, DC: Island Press.

NRC (National Research Council) (2005) *Valuing Ecosystem Services: Toward Better Environmental Decision-Making*. Washington, DC: National Academies Press.

Ostrom E (2005) *Understanding Institutional Diversity*. Princeton: Princeton University Press.

Patz JA and Olson SH (2006) Malaria risk and temperature: Influences from global climate change and local land use practices. *Proceedings of the National Academy of Sciences* 103: 5635–5636.

Tallis H, Lester SE, and Ruckelshaus M (2012) New metrics for managing and sustaining the Ocean's bounty. *Marine Policy* 36: 303–306.

Ecosystems of South America

Luis Anibal Solorzano Cardenas, Gordon and Betty Moore Foundation, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Caatinga Collective name assigned to the semiarid ecosystem of eastern South America. Some authors use this term as a general name for a type of thin Amazonian forests. In this article, however, the term is reserved for the xeric caatinga region of northeastern Brazil.

Cerrado Name assigned by phytogeographers to the tropical savannas of the central Brazilian shield. Cerrado vegetation is further divided into four categories: (1) campo limpio (i.e., clean field), (2) campo sujo (i.e., dirty field or grasslands with scattered shrubs), (3) campo cerrado (i.e., closed fields or grasslands with numerous trees and shrubs), and (4) cerradao (i.e., when the vegetation is dominated by a closed canopy of trees).

Chaco Region located between the Paraná basin, the central Brazilian shield, and the Andes. The ecosystems present in the Chaco region include temperate grasslands, savannas, and arid and semiarid environments.

Llanos Region of tropical South America, in the east side of the Andean range of Colombia and Venezuela. The llanos landscapes occupy lower elevations (2–300 m), and the landforms vary from flat to rolling terrain. The dominant vegetation is savanna with areas of dry forest, gallery forest, and morichales.

Mangal/mangle Name assigned to the vegetation of mangrove ecosystems.

Pampa(s) Region of temperate South America, west of the Andes in Argentina. The Pampa region contains several ecosystem types, including temperate grasslands, dry forest, and xeric shrub lands.

Pantanal Largest wetland of South America drained by the Paraguay River and its tributaries.

Paramo Ecosystem of the high Neotropical mountains, found in South America above tree line in the northern Andean mountains of Colombia, Venezuela, Ecuador, Peru, and Bolivia. The dominant vegetation is similar to an alpine meadow or grassland; cacti and giant-rosette plants are conspicuous.

Puna Ecosystem of the high Andean mountains found above tree line from Bolivia to southern Peru. The Puna comprises of four distinct ecological regions: wet, dry, thorny, and desert Puna. The dominant vegetation types vary from grassland/shrub lands to desert.

Varzea Floodplain of the whitewater rivers of the Amazon basin.

Outline

South America contains a diversity of ecosystems from tropical, subtropical, alpine, and temperate environments (Archibold, 1995). The geology, climate, and biogeographical history of the continent have important consequences in shaping the contemporary geographic distribution of ecosystem types and their structural attributes and functioning (Beard, 1955). The stability of the old Precambrian shields and the recent orogenic evolution of the Andean range have been the principal factors that determine geomorphologic, edaphic, and continental climatic patterns that correlate with the geographic distribution of major ecosystem types. The lifting Andean range causes altitudinal zoning and different environments that have significantly increased the diversity of ecosystems that occur in the continent. Most of the area of South America occupies tropical latitudes, hence tropical ecosystems dominate the landscapes. The main tropical ecosystems include tropical rain forest, dry forest, cloud forest, savannas, shrub lands, xeric formations, and high Andean ecosystems (Archibold, 1995). Ecosystems of the southern temperate regions include temperate grasslands, the Mediterranean Matorral, and temperate forest and deserts. There are extensive wetland ecosystems across the continent occupying areas along the floodplains of major rivers, lakes, estuarine areas, and seasonally flooded savannas. Mangrove ecosystems occur along the tropical coastlines of both the Atlantic and Pacific oceans.

Physical Environment

South America occupies an area of approximately 18 million square kilometers and extends from about 11° N to 56° S, and from about 35° W to 81° W. The continent has been connected to Central and North America by the land bridge of the Panamanian Isthmus for approximately 3.5 million years and is bounded on the east by the Pacific Ocean, on the north by the Caribbean Sea, and on the west by the Atlantic Ocean.

Geology and Geomorphology

The geological history of South America cannot be fully covered in this summary as it is very complex and reflects many tectonic and orogenic processes operating over millions of years. However, the structural characteristics of South American ecosystems – in particular the development, properties, and geography of their soils – are significantly associated with the geophysical environment. Likewise, the major landforms

of the continent create distinctive regional climatic patterns, which ultimately affect the functioning of ecosystems. A general overview of the geology and physiography of the continent will help to explain the distribution and general properties of its major ecosystems.

The basic geomorphic features of South America are (1) the high Andean Mountain chain on the western border of the continent spreading from Venezuela to Chile, (2) the low plains of the piedmont that occupy the eastern side of the Andes, (3) the very large sedimentary lowlands of the Amazon valley, and (4) the continental shields outcropping on the east (**Figure 1**). These major geotectonic categories of the continent have important differences in the origin, thickness, and stage of weathering and composition of the underlying materials, which are of essential importance to determine the type, depth, and physical composition of soils. Geomorphologic processes such as soil formation and erosion have produced characteristic surface features, which identify discrete areas developed under different climatic conditions (Harrington, 1973).

The central and eastern parts of the continent are made of three cratonic areas: the Guyana Shield, the central Brazilian Shield, and the Coastal Brazilian Shield (**Figure 1**). These cratonic areas have not been geologically disturbed since the Silurian ($\sim 408.5 \times 10^6$ years ago). Large plateaus and rolling highlands occupy the most eastern part of the continent covering most of the Guyana and Brazilian shields. From remote geological times, in the east side the stable Precambrian shields of the continent have held positions above sea level and have been affected by simple deformations. Landscapes in these areas, therefore, have evolved continuously under conditions of direct exposure to the atmosphere, resulting in the widespread development of old erosion surfaces. The shields consist primarily of crystalline basement complexes from the Precambrian period and the dominant rocks are granites, genisses, and mica schists. The waste mantle can be very thick

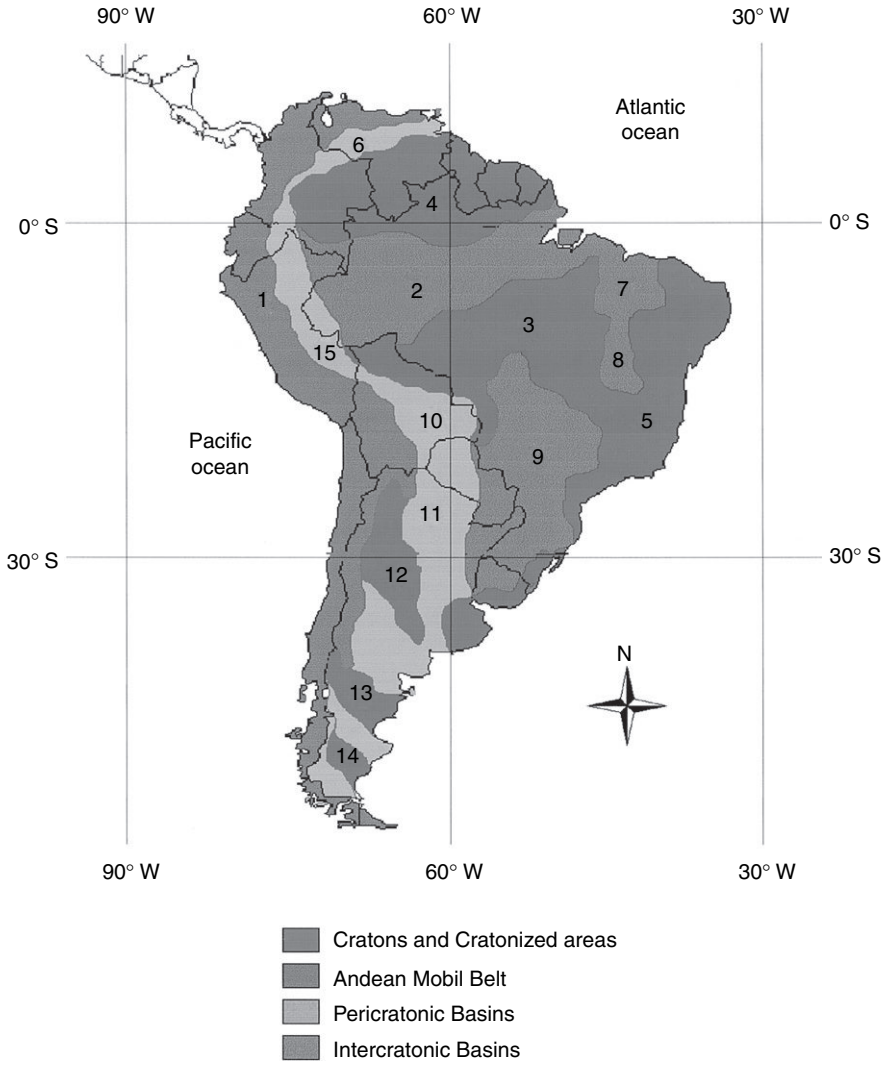


Figure 1 Major geotectonic components of South America. 1 = Andean Belt; 2 = Amazon Basin; 3 = Central Brazilian Shield; 4 = Guyana Shield; 5 = Coastal Brazilian Shield; 6 = Llanos; 7 = Parnaiba Basin; 8 = Sao Francisco Basin; 9 = Paraná Basin; 10 = Chaco; 11 = Pampas; 12 = Pampean Massif; 13 = Patagonian Massif; 14 = Deseado Massif; 15 = Lowlands of Iquitos, Acre, and El Beni.

on the old erosion surfaces of the shields. The materials originated from these rocks are at an advanced stage of *in situ* weathering and are millions of years old. A shallower layer of waste material covers the landscapes of residual relief. The valley bottoms and lowlands are covered with drift material that can reach considerable depths.

Along the Andean belt (**Figure 1**) there are Precambrian rocks masked by different younger deposits. Since the Paleozoic, the ocean has invaded pre-Andean areas several times, and extensive marine deposits mixed with metamorphic and intrusive rocks occur all along the range from Venezuela to Chile. One important consequence of marine transgressions (i.e., ingressions and regressions) is the inland deposition of rich marine sediments. During marine regressions, erosion increased the loss of previously deposited layers. In the Early Devonian ($408–362 \times 10^6$ years ago), the largest of all marine transgressions covered the Amazon and Parnaiba basins (**Figure 1**), which remained submerged until the middle Devonian. In the upper Paleozoic ($290–245 \times 10^6$ years ago) only the Parnaiba basin received mixed marine and continental sediments. In the Early Mesozoic ($245\sim \times 10^6$ years ago), eolian and fluviatile sediments were deposited in the Parnaiba and Amazon basins. The Triassic sediments ($208\sim \times 10^6$ years ago) came from continental areas and were deposited in an arid environment. During the Jurassic ($208–145 \times \sim 10^6$ years ago), the products of major volcanic activity in central and western Argentina were deposited in the Paraná, Sao Francisco, Parnaiba, and Amazon basins (**Figure 1**). During the Late Cretaceous, marine depositions also occurred over large areas of eastern South America (Harrington, 1973).

Most of the current physical features of South America – topography, coastline, and river systems – developed toward the end of the Tertiary period during the Pliocene ($2–8 \times \sim 10^6$ years ago) when the major uplift of the Andes took place. During the Tertiary, several tectonic movements accompanied by acid intrusions and volcanic activity lifted the western part of the continent forming what is now Andean Cordillera (**Figure 1**). During the uplifting, continuous sedimentation occurred in the inter-Andean valleys and along the eastern piedmont, forming deposits that were later involved in Andean orogenic movements that resulted in complex folded geological sequences. During the Miocene–Pliocene, most of the northern Andean ranges were strongly lifted while the southern Pampas were broken into faulted blocks. These movements were accompanied by intense volcanic activity in Colombia, Ecuador, Peru, Bolivia, Argentina, and Chile, which also cause the uplifting of the Brazilian and Guianan highlands in the eastern part of the continent. These orogenic and volcanic activities continue today, particularly in Colombia, Ecuador, and Chile, and the areas affected have very fertile soils. At present, the Andean mountains have fluctuating altitudes from 3000 to 5000 m with higher summits between Ecuador and Central Chile and the highest peak of the Western Hemisphere – Cerro Aconcagua – reaching an altitude of 7005 m in Argentina. Owing to the position of the Andes, most of the hydrological networks of the continent drain westward to the Atlantic Ocean *via* three main river systems: the Amazon River, the Paraná–Paraguay Rivers, and the Orinoco River.

A wide belt of low sedimentary plains extends along the eastern piedmont of the Andes, from the Llanos of Colombia and Venezuela to the Argentinean Pampas, and eastward along the valley of the Amazon River (**Figure 1**). The majority of these plains reach altitudes from 200 to 500 m. These plains are composed of Tertiary and Quaternary sediments deposited from the South American and Brazilian shields and more recently from the Andes (Harrington, 1973). The intercratonic lowlands of the Amazon basin are made of heavy sedimentary clays originated from erosion of uplifted deposits washed from the Andes (**Figure 1**). Those deep clay sediments were deposited on the flat bottom of an ancient inland sea that covered most ofthe basin during the Late Tertiary–Early Pleistocene when the ocean level was higher.

Climate

Climate is an important factor that determines to some degree the type of soils and vegetation of any given region and, therefore, influences the structure and functioning of ecosystems. **Figure 2** illustrates the major South American climate types according to the classification of Köppen. The shape of South America essentially determines its large-scale climatic patterns, which in turn influence the geographical distribution of ecosystems. South America has the shape of an acute triangle with most of its area situated in tropical latitudes and its southern portion extending well into high latitudes. Although the southern portion of South America reaches more than 20° farther South than the tip of Africa, the continent is so narrow at the southern end that the landmass of this region rather constitutes a continental peninsula. Because most of South America occupies tropical regions and the continental area decreases at higher latitudes, the property of climatic continentality is absent. Accordingly, the more important factors that determine the diversity of climatic regimes observed across South America are the major global atmospheric circulation patterns, the proximity to oceans and their currents, the Andean relief, and other coastal and inland topographic variations (Colinvaux, 1989).

In general, tropical temperatures dominate the northern part of the continent, declining smoothly toward the south. Major variations in temperature are due to the Andean Mountains and ocean currents. The presence of the Andes causes climatic diversity over the continent due to altitudinal variations that change wind circulation patterns and to adiabatic responses of air masses along the altitudinal gradients of the mountain chains. South American environments can range from hot deserts and warm rainy climates in the lowlands to cold deserts, temperate, and iced polar climates in the Andean highlands (**Figure 2**). Two ocean currents are particularly important to the regional temperatures: the cold Humboldt Current that brings lower temperatures from the south up north along the West Coast and the warm Brazilian current that brings warmer temperatures down south along the eastern part of the continent (Edit, 1968).

Owing to the lack of continentality, the distinct climates of South America are often characterized by differences in precipitation. Rainfall across most of the northern and central parts of the continent depends on the intertropical convergence (ITC). The ITC occupies its most northerly position from June to September, when it is located between 7° and 9° N. Its advance and retreat migration does not take place

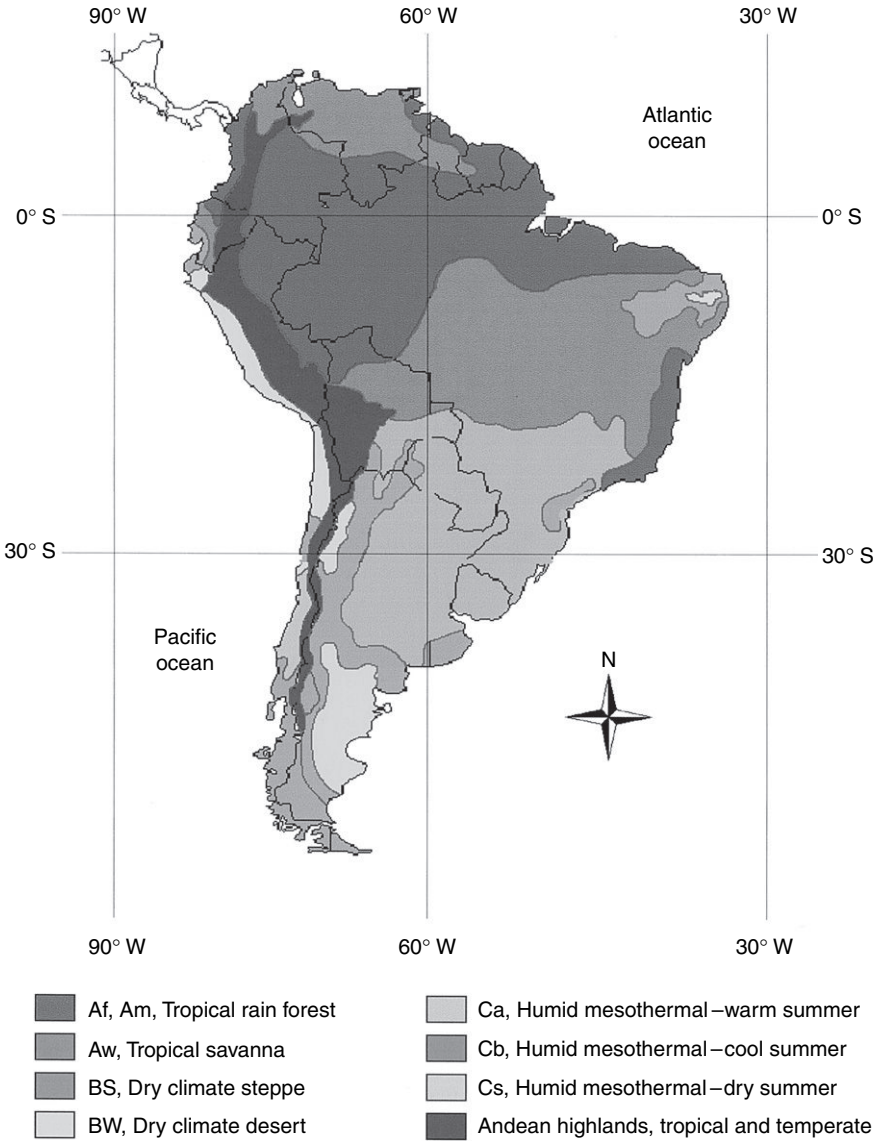


Figure 2 Broad climatic regions of South America defined according to the Köppen classification system.

parallel to the equatorial line. When the ITC drifts south, a prevalent High Pressure Cell in the Atlantic Ocean restrains its movement in the eastern part of the continent causing an arch of precipitation that reaches southeast Brazil in February–March, while dry conditions persist in the northeast. This Atlantic high cell also causes higher precipitation in most of the East Coast north of Patagonia.

The Andean chain also has a significant effect on continental patterns of precipitation. Owing to adiabatic cooling, the Andean windward slopes remove moisture from the air masses moving from the Pacific causing a large rain shadow that extends from east of the Andes from Bolivia to Tierra del Fuego (**Figure 2**). Conversely, in the windward slopes of the Pacific coasts of Colombia and Chile, the Andes cause exceptional orographic rainfall levels, which in the Choco region of Colombia can surpass 10,000 mm per year (**Figure 2**). A similar effect is observed in the eastern Andean slopes of central South America where along the western edge of the Amazon basin, moist winds rise and deliver higher precipitation levels than those registered in lower elevations of the basin. In the north of Venezuela, the northern Andean ranges block the easterly trade winds and create a tropical savanna climate that extends from northern Colombia to Surinam (**Figure 2**). In the western coast of the continent, between 27° and 10° S, the edge of the western Pacific high-pressure cell combined with the cold Humboldt current to produce an arid zone called, in Chile, the Atacama Desert (**Figure 2**). The structural complexity of the Andean relief includes inter-Andean cordilleras and valleys (e.g., the Eastern, Central, and Western Colombian cordilleras) and high plateaus (e.g., the Altiplanos of Peru and Bolivia) that further create many local climates according to variations in wind exposure and altitude (Edit, 1968).

Other extra-Andean topographic features that produce orographic climatic modifications of rain shadows or rainfalls are the highlands of the Brazilian shield that block onshore winds between Porto Alegre and Bahia; the uplands of

northeast Brazil that cause a dry rain shadow in the interior northeast region; the Guyana Highlands that block the trade winds and cause a dry winter effect in their leeward slopes; and the Sierras de Cordoba and Patagonian Plateau in Argentina, which affect air temperature and can collect moisture from humid air masses crossing Patagonia (**Figure 2**). There are also dry inter-Andean valleys localized along its range.

Biogeography

Early History and Associations

Biogeographers include central and northern South America in the Neotropical Kingdom, which also contains the Antilles and most of tropical Central America. Until approximately the Early Cretaceous, South America, Africa, Australia, and Antarctica were joined together composing the supercontinent Gondwanaland, thus allowing for a continuous interchange of their ancient biotas. Africa and South America share many plant families like Annonaceae, Myristicaceae, Cecropiaceae, Sterculiaceae, and Bombacaceae. Some plant genera are very diverse in one continent and barely represented on the other, for instance, Mayaca (about eight species in America and one in Africa), Duvernoya (~35 species in Africa and three in America), and Hyptis (~400 species in America and two in Africa). The southern cone of South America is included within the Holantartic Kingdom, which includes the temperate mesic-adapted floras of southern South America, New Zealand, and southwestern Australia. The genera of trees Araucaria, Podocarpus, and Nothofagus are typical although not exclusive to the southern cone. Fossil records have also shown strong ancient faunal associations between South America, Africa, Antarctica, and Australia. For instance, some of the existing species of fishes of South America, as well as their parasites, show affinities with African groups. There are also close phylogenetic relations among groups of South American, African, and New Zealand invertebrates including, among others, arachnids and mollusks.

After the separation of South America from the rest of Gondwanaland in the Early Cretaceous, the biota of the continent was isolated for millions of years. Because of their isolation, South American flora and fauna are extraordinarily rich in endemic groups and species. It is estimated that about 7% of the total world's superior plants are endemic species from Brazil, 6.8% from Colombia, 3.2% from Venezuela, 2.1% from Peru, and 2% from Ecuador. Brazil is the country with more mammal species reported in the world (>520 of which more than 130 are endemic), Colombia the fourth (>455), and Peru the ninth (>345 of which 46 are endemic). Endemic faunal families include the armadillos Dasypodidae, the anteaters Myrmecophagidae, and the monkeys Cebidae. Particularly important are the marsupials that in South America are represented by two genera, two families, and 87 species. There are no fewer than 3000 bird species with two orders and 30 endemic families. The humming birds, which are endemic of the New World, have more than 250 species in South America and populate ecosystems from the Amazonian lowlands to the high Andean Mountains. The diversity of endemic reptiles is exceptional; there are 520 endemic species reported for Colombia, 468 in Brazil, 374 in Ecuador, 298 in Peru, and 293 in Venezuela, including species of snakes, turtles, and iguanas –which other than in South America only occur in the Fiji and Tonga Islands of the Pacific. Brazil is also considered as the country with highest diversity of freshwater fish species in the world with >3000 species reported, Colombia the second (>1500), Venezuela the fourth (>1200), and Peru the seventh (>850). The diversity of invertebrates is also outstanding; for example, there are more than 350 endemic species of butterflies reported for Peru, 300 for Colombia, 200 for Brazil, and 200 for Bolivia.

Late Tertiary

Although it is likely that some exchange occurred between North and South America during the Eocene and Miocene, the land bridge of the Panama Isthmus connected permanently South, Central, and North America about 3.5 million years ago. The completion of the Panamanian land bridge in the Early Pliocene was perhaps one of the most influential biogeographical events that has occurred in the continent and it dramatically changed the ancient biota of South America. The significant exchange of South and North American biotas is called by biogeographers as "the great American Interchange". Among the many invaders that entered South America through the isthmus are rattlesnakes, rodents, tapirs, deer, peccaries, felids (including jaguars and pumas), canids, and humans. About 40% of the mammal fauna of South America are species that arrived since the Pliocene. Among the South American forms that migrated to North America are the ground sloth, which disappeared about 10,000 years ago, and opossums and armadillos, which are still expanding their range. The dynamic interchange of plant and animal forms between South and North America continues at present.

Quaternary

During the past 2 million years, climatic fluctuations of dry glacial and wet interglacial periods affected the geographic ranges of both temperate and tropical areas. Palynological sequences indicate that at least during the Pleistocene, major climatic changes affected speciation and biogeographic patterns in South America. The modern geographical distributions of many endemic taxa resulted from Quaternary speciation and redistribution of species overlaid with previously existing range patterns. The explicit mechanisms that generated current biogeographical patterns are not fully understood, in part due to their inherently historical nature that is difficult to reconstruct (Bush, 1994, 2003).

Paleoecological studies initially assumed that the richness and complexity of plant communities were a consequence of continuous favorable growing conditions persisting over long time periods. Many areas have been identified as centers of endemism and proposed as tropical forest refugia occurring during the period spanning 18,000 to 13,000 years ago (Haffer, 1969). According to this hypothesis, during the last four glaciating cycles of the Pleistocene, tropical forest contracted during dry glacial periods causing most of the Amazon Basin, as well as other forested areas, to be occupied by savanna vegetation. Continuous but isolated islands of forest

(refugia) might have contained populations of plant and animal species that underwent speciation while they remained isolated. During mesic interglacial periods similar to currently existing conditions, the forest expanded again and new geographic patterns appeared including the signal of speciation processes that occurred within the refugia themselves. Paleoecological data have also shown that in the Andes, the tree line has migrated down and up-slope by as much as 1500 m, matching the frequency of each northern hemispheric ice age. The estimated cooling experienced during the glacial maxima at the Sabana de Bogota is 7–9 °C. A consequence of the lowering of the tree line would be the connection of some of the high Andean Paramo ecosystems, which are currently isolated on top of mountains.

An alternative explanation for the distribution of centers of endemism proposes changes in temperature rather than precipitation as the driving force of vegetation rearrangements during glacial periods. It has been proposed that the areas of refugia were not isolated islands of stability but rather of maximal disturbance due to cooler glacial temperatures, reduced atmospheric CO_2, and moderate reductions of precipitation. Far from being static places of refuge, under this hypothesis the regions of endemism would be the dynamic edges between the forest below and the cool-tolerant vegetation above (Colinvaux, 1989).

The uplifting of the Andes during the Quaternary has also contributed to the isolation or dispersal of some taxa. While the mountain range has served as a dispersal corridor from north to south, it also represents a dispersal barrier (like the Atacama Desert) to some plants and animals. For instance, the floristic individuality of the High Andean deserts is attributed to the short history of these environments that appeared after the recent and fast uplifting of the cordillera. The age of these ecosystems might have not allowed enough evolutionary time for many species to adapt and disperse under those new conditions. There is evidence also that during interglacial dry periods the southern Andes cone experienced a cold and arid climate that served as an effective dispersal obstacle for rodents and plant species from the northern arid ecosystems.

Not only long-term events like Pleistocene glaciations or orogenic processes such as the uplifting of the Andes have had significant effects on the ecosystems of South America, but short-term climatic episodes like El Niño cause important also impacts as well. For example, changes in water temperatures of the western Pacific during El Niño years trigger profound modifications in rainfall patterns across all of South America. Major large-scale ecosystem disturbances have been associated with El Niño: fires in savannas and forests of Amazonia; droughts in the llanos of northern South America, in the Chaco region and in northeastern Brazil; and floods in Peru, Ecuador, and Colombia. There is evidence that these perturbations affect the dynamics of populations and ecosystems, which over evolutionary time could modify the composition and geographic patterns of their biota (Hoorn *et al.*, 2010).

Major Ecosystems of South America

Although it is possible to make some generalizations about the nature of the biomes of the continent and their major ecosystem types, it is important to bear in mind that across several spatial scales they are in fact complex mosaics of microhabitats (Beard, 1955). The categories of ecosystem types applied in this summary are very broad and based on the classification system developed by FAO-Unesco (1971), satellite-based maps and the ecoregions study of the World Bank/World Wildlife Fund (Dinerstein *et al.*, 1995). In this review, very general categories are used and ecosystems are geographically delimited based on broad vegetation formations (Beard, 1955). Only some key features of these major ecosystem types are briefly discussed.

Tropical Forest

Tropical forests are among the most complex ecological communities of the world and offer a variety of habitats that support an immense diversity of plants and animals (Leigh, 1975; Terborgh, 1992). South America contains more than 40% of the tropical forest of the world.

Distribution and Structure

Tropical Rain Forest

Tropical wet forests occupy the drainage of the Amazon and southern Orinoco basin, the piedmont of the Andes in northern and western South America, the Choco-Darien region between the Andes and the Pacific Ocean in western Colombia, and the southwest and Atlantic coast of Brazil (**Figure 3**). Among other types of tropical rain forests are Babassu palm forests, which occupy the southeastern border of Amazonia, and the liana forests, which are found in southern Amazonia between the rivers Xingu, Tapajos, and Tocantins. A liana forest is an open forest with well-spaced trees, rich in woody climbers, and often completely entwined by lianas. It occurs on all soil types, but there are no detailed studies of this vegetation, nor explanation of the factors that cause its wide distribution.

Tropical Semideciduous and Dry Forests

These ecosystems occupy part of the eastern edge of the Andes in Bolivia and Brazil, the inter-Andean valleys of Colombia, the northern edge of the Venezuelan Llanos, and the western regions of Ecuador and Peru (**Figure 3**). Most of the South American regions originally covered by dry forests are today populated areas, and much of the primary vegetation has been cleared. Tropical semideciduous forests usually occur on alluvial lowlands, submontane, and mountain terrains up to 1000 m. Between the Amazonian rain forest and the *Cerrado* (discussed later) of central Brazil, there is a so-called *transition forest*, which comprises a large number of plant communities. Transition forests are generally less species-rich than the moister forest in terms of plant and vertebrate species; however, plant physiognomy diversity seems to be greater in dry than in wet forests. Both transition dry forest and savannas can be found under similar climatic conditions (Beard, 1955). In eastern South America, transition forests run along most of the boundary between Amazonia and the Guyana and Brazilian shields, and they vary in width from a few kilometers to more than 100 km in some areas. The transition forest is taller and with a more closed canopy than the closed Cerrado, *cerradao*

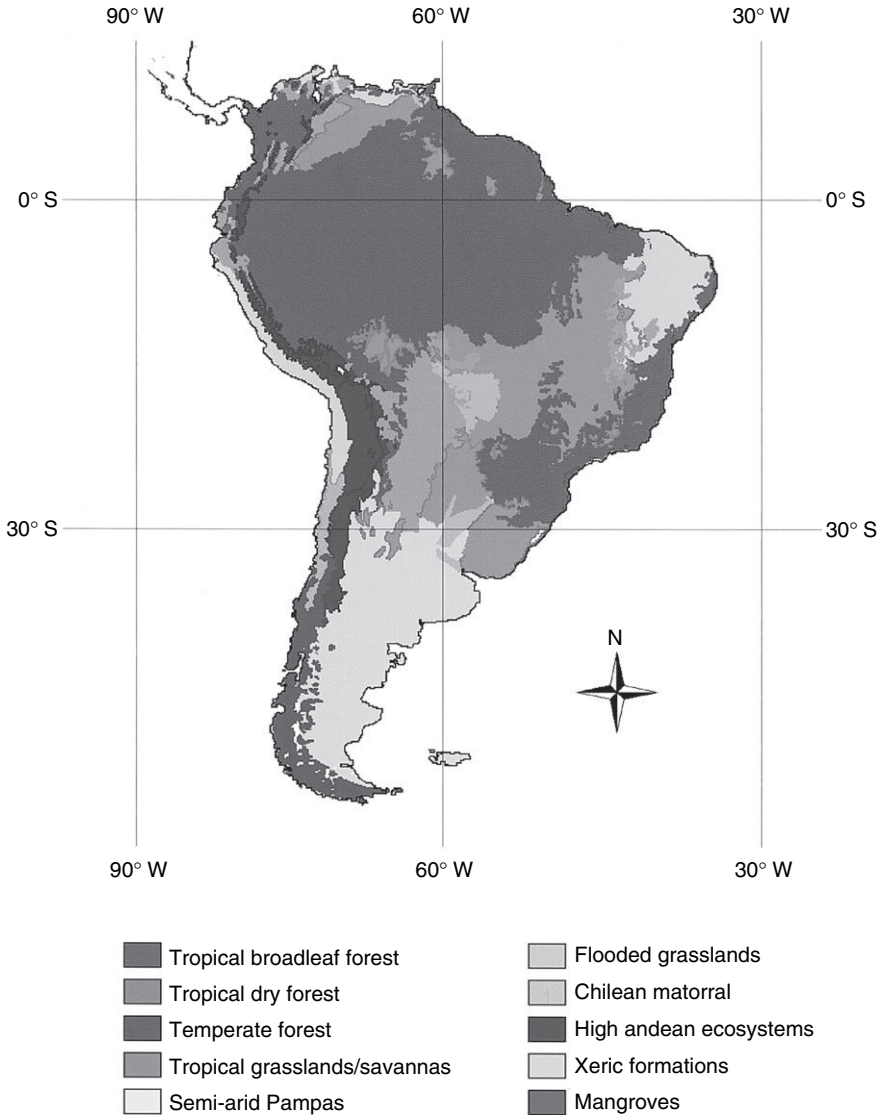


Figure 3 Simplified map of the major ecosystem types of South America geographically defined on the basis of dominant vegetation.

(discussed later), but is lower and with a more open canopy than the typical Amazonian rain forests.

Mountain Cloud Forests

These ecosystems occur in the humid tropical slopes of the Andes from Venezuela to northwest Argentina and in the Santa Marta Massif in Colombia. They occupy defined altitudinal belts with the lower belts intermixing with lowland forests and the upper belts limiting with the high Andean grasslands at the tree line about 3000 m above sea level. Fog is persistent in these ecosystems and the vegetation is luxuriant with an extraordinary variety and abundance of epiphytic forms and mosses.

General Patterns

Structural variations between tropical forests are mainly governed by differences in the amount and seasonality of precipitation and the temperature regimes. In general, the proportion of deciduous woody components increases along a gradient of rainfall, as the amount of annual precipitation decreases below 2000 mm. Although exact boundaries between ecosystems are somewhat arbitrary – because vegetation is not always so sharply delimited – the transition from arid types to rain forest seems to be primarily determined by continental climatic patterns, especially rainfall intensity and seasonality. When nonevergreen trees are present in wet tropical forests, they never shed all their foliage at the same time. A large number of woody species, of which most have evergreen foliage, characterize these ecosystems and the most frequent tree species rarely represent more than 15% of all the species present.

In general, tropical forests have a vertical structure with several strata where the tallest trees form a discontinuous layer of emergent individuals, which in South America can reach heights of 30–40 m. A denser canopy grows below and consists of two or more layers of shorter trees, small palms, tall

herbs, ferns, and shrubs. Lowland forest trees are generally shallow rooted and many of them develop buttresses at the base of their trunks, an adaptation thought to provide mechanical support. The epiphytes and vines are conspicuous in tropical forests, where the epiphytic Orchidaceae and Bromeliaceae are usually prominent. The barks of lowland tropical rain forest trees are usually thin and smooth, and this morphological characteristic makes the anchorage of epiphytic forms difficult. In the Andean cloud forests, the heights of trees decrease at increased altitudes and their shapes become irregular (e.g., elfin forest). In cloud forests, the bark of trees and the forest floor are covered with mosses, lichens, and ferns, and the diversity of fungi, algae, bryophytes, and seedless vascular plants is remarkable.

Functional Aspects

Tropical Rain Forests

A mean monthly temperature of $>26\,^{\circ}C$ and mean annual rainfall usually exceeding 1800 mm characterizes the Tropical Rain Forest (*Af*) climate that coincides with the geographic distribution of tropical rain forest (**Figures 2** and **3**). Some areas of tropical rain forests are the result of orographic precipitation. Isolated mountains (e.g., Sierra Nevada de Santa Marta in Colombia) or coastal escarpments (e.g., the Atlantic Brazilian Shield and the lower slopes of the Andes in the Pacific coast of Colombia and Ecuador) act as barriers to airflow, forcing moist air to ascend upslopes where adiabatic cooling generates clouds and precipitation (Terborgh, 1992). These geophysical conditions conduce to forested landscapes with a physiognomy similar to the Amazon wet evergreen forest. In areas of lowland tropical forest (e.g., Amazonia), minimum changes in temperature between summer and winter and continuous heating coupled with the high evapotranspiration rates of vegetation create a predictable pattern of daily convectional rain that falls during the middle afternoon. It is estimated that up to 80% of the precipitation falling in Amazonia is retained within the hydrological cycle of the basin.

The geological material underlying the hydrographic network also influences the development of forest ecosystems. In the Amazon lowlands, large areas of forest are subject to inundation (varzea) and these ecosystems differ from those that occupy high terrain (terra firme). Rivers running from the Andes carry large amounts of silt and have a light brown color (whitewater). These large loads of silt are deposited in areas of low relief and form broad floodplains of high fertility. During the flooding season, these areas remain under several meters of water. In contrast, rivers running through the old Precambrian Shield are poor in nutrients and minerals and their waters are clear. The Shields have been exposed to millions of years of leaching and in order to conserve nutrients, plants protect their foliage with tannins. The tannins leached into the rivers give a tainted coloration to the water (blackwater). The Rio Negro and other blackwater rivers also have floodplains, but they are different because of the lack of sediments and because plant communities are adapted to acidic conditions (igapo).

Because of the high temperatures and abundant rainfall of tropical forest ecosystems, there is intense chemical action on bedrock, and soils and all soluble components continuously leach, producing characteristic Ultisol and Oxisol soils (Terborgh, 1992). These soils are rich in iron, manganese, and aluminum, which stay behind after all other soil soluble constituents have been leached. Large amounts of these minerals arrange in stratified layers forming laterites.

Field studies have shown that tropical forests exploit a larger volume of soil than previously thought. Soil depth is therefore quite important in water balance variability because it allows evergreen forest to keep evapotranspiring during dry periods by absorbing water from soil depths of more than 8 m. Although the most rapid changes in water availability take place at the soil surface, where it is depleted by plants, the process of hydraulic lift can take water from deep soil layers and discharge it into the dryer upper layers. During the day, the water potential (ψ) gradient causes water to move from the ground into the roots, shoots, and through stomata to the atmosphere. During the night when stomata are closed, ψ in the plant equals ψ in the deep soil and a water potential gradient can move water from the deeper moist soil and plant interior into the dryer upper soil layers.

Decomposition and nutrient dynamics proceed rapidly under the warm, moist conditions of tropical forest ecosystems. High temperatures also favor an intense bacterial activity on upper soil layers and therefore there is no accumulation of humus. Many nutrients are stored in the forest biomass where they are kept from leaching. Nutrients are partitioned in different compartments of the forest and, on a weight basis, leaves contain the highest concentrations. Most of the phosphorous is stored in the leafy biomass and potassium and calcium in stem tissues. Nutrient losses in intact forests are generally low because of the high concentration of roots in the upper soil layers. The presence of mycorrhizal associations in roots enhances their nutrient uptake. Epiphytes also reduce the loss of nutrients washed away from standing biomass. Over long successional periods, the forest biomass is progressively stored in the woody components and nutrient accumulation in leaves decreases.

Invertebrates can account for up to 60% of the animal biomass of tropical forest ecosystems. About 19% of the woody fraction of the biomass is consumed by termites, beetles, and the larvae of insects. Litter is the main source of food for the animal community and leaves, flowers, and fruits are consumed mainly by insects, birds, bats, and small mammals. Chemical defenses are common among tropical forest plants and many species have developed mechanical defenses by incorporating silica, lignin, and fibers in their tissues.

Tropical Semideciduous and Dry Forests

Along the gradients from tropical rain forest to dry forest, the structural and functional attributes of the ecosystems change. Tropical deciduous forests correlate with climates of two well-determined seasons: a rainy period followed by a dry one. More than 50% of the tree species shed their foliage during the dry period. The predominant characteristic of the habitat of these forests is the prolonged seasonal drought, causing desiccation of the topsoil and lowering atmospheric humidity. The length of the dry season determines the degree of divergence in physiognomy and structure of the seasonal forest (Gentry, 1995). From wet evergreen to dry deciduous forests, prolonged drought correlates with a reduction in

physiognomy and floristic composition and the lowering of canopy height. At the same time, the degree of deciduousness of the trees of the overstory increases. The phenology of these ecosystems is determined by a seasonal climate with a period of intense summer rainfalls followed by a dry winter. In the southern dry deciduous forests, the winter season also brings changes in mean temperature down to 15 °C. Although between 20% and 50% of the woody species found in these ecosystems are deciduous, these forests are still dominated by Amazonian genera widely distributed across South America, such as *Parapiptadenia*, *Peltophorum*, *Cariniana*, *Lecythis*, *Tabebuia*, *Astronium*, and others of lesser physiognomic importance (Murphy and Lugo, 1986).

Although tropical rain forests contain the highest diversity of species measured in terrestrial ecosystems, life-form diversity is higher in dry forest probably because of habitat heterogeneity. Woody plants in tropical rain forest tend to converge to a small number of lifeforms. In drier seasonal environments, the proportion of deciduous trees and shrubs increases, the presence of epiphytes decreases, and vines become more frequent. The presence of epiphytes is associated with high air humidity and the presence of dew (Gentry, 1995). Along seasonality, gradients succulent plants conspicuously increase, including those with Crassulacean acid metabolism (CAM). Succulent-stemmed plants have very stable water relations and are drought resistant. Evergreen woody plants dominate at both extremes of the gradient, and although they all belong to the C3 photosynthesis type, there are differences in their leaf structure and drought resistance, which allow them to function in both mesic and arid environments (Murphy and Lugo, 1986).

Savannas, Grasslands, and Shrub Lands

Distribution and Structure

Savannas and grasslands occupy nearly 25% of South America in both tropical and subtropical regions.

Tropical Savannas

These ecosystems are characteristic of the warm lowland tropics and occur in areas with a strongly seasonal rainfall regime and a dry period lasting from 4 to 7 or 8 months. Although a herbaceous cover consisting mostly of bunch grasses and sedges is dominant in savannas, the term "savanna" embraces a variety of vegetation types found in tropical latitudes. An important characteristic of tropical savannas is a clear seasonality in their phenology and a period of lower activity associated with times of water stress. According to Sarmiento (1984), tropical savannas can be divided into four functionally distinct types: (1) semiseasonal savannas, with weak water stress conditions; (2) seasonal savannas, with a distinct rainfall seasonality and common fires during the dry season; (3) hyperseasonal savannas, where in addition to a marked dry season there is also a period of water flooding; and (4) "esteros," which are areas without a clear dry season, but an excess of soil water for most of the year.

The largest extension of savanna in South America occupies the old Precambrian Brazilian shield and has been dubbed by phytogeographers "Cerrados." Savanna ecosystems are also found in the Orinoco Llanos of Colombia and Venezuela, in Suriname, Guyana, in some Amazonian regions of Brazil, and in Paraguay, Bolivia, Argentina, and the northern Chaco region (**Figure 3**). The vegetation of the Guyanan savannas is more related floristically to the llanos than the cerrado.

Perennial grasses, herbs, and shrubs are the dominant life forms in savanna ecosystems. They are well adapted to the seasonal climate, soils poor in nutrients, and the fire disturbance regime that dominate in these environments. At least five phenological groups of grasses have been recognized in savanna ecosystems: perennial with a seasonal semidormant period, annual ephemeral with a short cycle, annual with long cycle, perennial with a seasonal dormant period, and of continuous flowering and growth (Coupland, 1992). Tree species found in savannas have adapted to these conditions by increased allocation to underground biomass. Trees in woody savannas can reach 25–30 m in height, but their diameter rarely exceeds 50 cm. Some savanna tree species are described as subterranean trees because their roots can penetrate to depths of 18 m. Other species like *Curatella americana* have shallow roots that can grow more than 20 m horizontally and secondary roots that can reach depths of 6 m. Other conspicuous components of savanna landscapes are the gallery forest along rivers and smaller streams and the seasonally inundated palm communities (*morichales*) dominated by the palm *Mauritia flexuosa*.

In the Brazilian Cerrado, there is a series of plant communities from open grasslands to dense woodlands and more or less recognizable stages in this continuum receive vernacular names. Dry grasslands without shrubs or trees are called "campo limpio" (i.e., clean field); grasslands with scattered shrubs are called "campo sujo" (i.e., dirty field); grasslands with numerous trees and shrubs are called "campo cerrado" (i.e., closed field); when the vegetation is dominated by a closed canopy of trees it is called "cerradao" (i.e., the vegetation is closed). The latter is a woodland composed of trees often 8–12 m or even taller, with ground vegetation reduced because of the shade. Many factors probably determine which of these forms of Cerrado vegetation occurs in a given locality, among them topography and soil texture seem to have the main causal effect. Cerrados have a markedly seasonal climate and posses a large characteristic flora of fire-resistant plants, including about 800 species of trees and large shrubs, and many times that number of herbs and subshrubs. The vast majority of these species are endemic to the Cerrado, an ancient vegetation formation dating back perhaps 50 million years (Coupland, 1992).

Temperate Grasslands

These ecosystems occupy the eastern part of southern South America in Argentina, Uruguay, Chile, and southern Brazil. They expand across the pampas of Rio de la Plata in Argentina, the Chaco region, the semiarid region west of the humid pampa (*Monte*), and reach the edges of desert and semidesert areas of the Patagonian region (**Figure 3**). The dominant landforms are large plains with a few table-shaped outcroppings no more than 500 m above the plains. The dominant climate across the large grassland region is characterized by a mean temperature of 10–20 °C with an annual precipitation ranging from 400 to 1600 mm. The dominant soil types

are young and include Mollisols, Alfisols, Vertisols, and Entisols.

Vegetation in the Pampean temperate grasslands is composed of about 1000 species and varies from grass-dominated communities in the east, to xeric woodlands and semidesert communities in the west. In the Argentinean Pampas, plant communities are dominated by species of *Agrostis*, *Bouteloa*, *Elyonurus*, *Festuca*, *Panicum*, *Paspalum*, and *Stipa*. In Chile and northward in the Andes, the ecosystems of grasslands are dominated by species *of Andropogon*, *Bromus*, *Calamagrostis*, *Poa*, *Sporobolus*, *Eragrostis*, and *Distichlis*. It is difficult to determine the proportion of Patagonia that can be classified as grasslands, because the sparse vegetation is mostly dwarf shrubs and grasses do not form a continuous cover, therefore many areas are treated rather as deserts.

High Andean Grasslands

These ecosystems are found above the tree line (3000–4500 m) of the Andean range in Colombia, Venezuela, Ecuador (*Paramo*), Peru, Bolivia, and Chile (*Puna*). The actual lower limit of paramo and puna ecosystems depends on exposure, the climatic regime, and in particular precipitation. The dominant life-forms of the Andean grasslands are giant-rosette plants (known locally as "*frailejon*"), tussocks of grasses or sedges, acaulescent rosette plants, cushion plants, cacti, and sclerophyllous shrubs. The giant-rosette plants can reach heights of several meters and have developed a number of morphological and anatomical adaptations including the retention of dead leaves that cover the shoot and provide insulation. At higher elevations, the vegetation cover becomes sparser and shrubs and herbs replace bunch grasses.

Functional Aspects

Tropical Savannas

Relatively homogeneous temperature and a well-defined wet season characterize the Tropical Savanna (*Aw*) tropical savanna climate or dry–wet tropical climate (**Figure 2**). The alternation of wet and dry periods coincides with the presence of tropical savanna vegetation, characterized by open spaces covered by grasses and spaced trees. Soils are mainly latosols of low fertility, formed by similar processes but not as deep as those found in the *Af* climate.

Savanna ecosystems are characterized by seasonal growth cycles that follow the annual precipitation cycles. However, the phenology patterns of savanna plants are very asynchronous resulting in different patterns of productivity, reproduction, and decomposition. Most grasses usually start growing at the beginning of the rainy season and reach reproductive stages early in the wet season or a few months later. Some vegetative growth continues throughout the year but at the end of the dry season, all aboveground biomass is dead. In perennial grasses, all the aboveground biomass dies by the end of the dry season and in the following year is replaced by vegetative growth of subterranean perennial organs. In evergreen species, growth of leaf and shoot biomass and reproduction generally occur during the dry season. Leaf development stops at the beginning of the wet season, and before the dry season begins, leaves are shaded. In most evergreen species, however, growth of stems and other woody parts only occurs during the wet season.

Usually there are layers of laterites under savanna soils that originated during the Cretaceous and Tertiary periods. Continuous weathering has produced soils poor in nutrients and with high acidity and concentration of aluminum that immobilizes phosphorous as iron–aluminum phosphate. In the soil environment, marked annual differences occur particularly in the upper soil layers where moisture levels can exceed field capacity during the wet season and diminish to wilting point during the dry period.

The stock of nutrients in savannas varies over the year as they move from different compartments of the ecosystem. The largest fraction of nutrients is stored in the aboveground biomass. Savanna trees tend to accumulate fewer nutrients than similar dry forest species and some nutrients reallocate from leaves to other organs before shedding occurs. Some nutrients accumulated in grasses return to the soil after fires but large fractions of nitrogen, phosphorous, and potassium are lost to the atmosphere when biomass burns. Decomposition of leaf litter by microorganisms occurs during the wet season and termites consume most of the woody material during the dry period.

The flora of the Brazilian cerrado and other South American savannas has typical features of pyrophytic vegetation. Trees are low, of contorted form, with thick, corky, fire-resistant barks; sclerophylly is common, and many leaves have thick cuticles, silicified tissues, and are often of considerable longevity. The underground organs of perennial grass species are protected from fire and their seeds as well as those of annual species are adapted to fire; in fact, the germination of many grass species is stimulated by fire events (Eiten, 1972). Because savanna vegetation is adapted to resist burning, occasional natural fires must have been an environmental factor throughout its history. Frequent burning favors the herbaceous, as opposed to the woody, component of the vegetation and when this occurs, savanna woodlands are transformed into grasslands. Conversely, protection from fire often has the opposite effect and a type of dense savanna woodland with considerable shade and only a very sparse grass layer is established. Some authors have correlated increasing production of woody vegetation with an increasing soil fertility gradient. Soil moisture is another factor that controls many of the most obvious differences in vegetation in the savanna landscape (Eiten, 1972). In general, woody vegetation occurs only on soils well drained throughout the year. Swampy gallery forest is found where the water table is permanently high.

Temperate Grasslands

There are different hypotheses that attempt to explain the physiognomic origin of the temperate grasslands. An initial hypothesis proposed that pre-Columbian people burned an original scrub forest. Another hypothesis points to an adaptive dominance of grasses due to a climatic effect that maintains a negative water balance during part of the year. Although there are significant regional differences, edaphic factors such as the fine texture of soils are proposed as major causes for the dominance of grasses. A lack of air spaces in the soil combined with summer dry periods would favor the competitive ability of grasses over trees. A variety of root growth habits is observed among both grasses and shrubs and the stratification of root

biomass along the soil profile minimizes competition for water and nutrients. The dominant species of grasses are perennial, live for 3–4 years, and are short and leafy.

The soils of the southern grasslands usually present bad drainage conditions, relatively acidic pH, and the humus layer can be 30–40 cm deep. Deeper soil layers are less organic, more alkaline, and composed mostly of clay or calcareous sediments. From west to east, soil depth and organic matter content decrease because in drier areas productivity is lower. Soils are relatively poor in nitrogen and phosphorous, which frequently limit grassland productivity. Nitrogen accumulates mainly in organic forms; it is quickly absorbed and stored in leaves and live root biomass. Fungi are the most important decomposers in grassland ecosystems, and nematodes are the principal soil invertebrates. In the flooded Pampean grasslands, phenological studies have shown that there is not a well-defined growing season. These grasslands are subject to periodic flooding and drought almost every year and floristically different plant communities occur according to various degrees of perturbation (Coupland, 1992).

South American grasslands have been greatly modified by humans through burning, grazing livestock, and agriculture. Before these intensive activities took place, there were a vast number of ungulates like the Guanaco (*Lama guanicoe*) and the Pampas deer (*Ozotoceros bezoarticus*). Birds are still abundant with more than 390 species described, of which the most conspicuous is the Ñandú (*Rhea americana*). Reptiles and amphibians are also abundant with poisonous snakes (e.g., *Bohtrops alternata*), caimans (e.g., *Caiman latirostris*), lizards (e.g., *Teius teyou*), toads (e.g., *Bufo refus*), and frogs (e.g., *Hyla pulchella*).

High Andean Grasslands

The most important functional attributes of these ecosystems are associated with the extraordinary amplitude of daily fluctuations in temperature and humidity, which are more pronounced than the seasonal ones. In these environments, high daytime temperatures follow nocturnal frost causing significant physiological stress in both plants and animals. Plants have evolved mechanisms that allow them to grow during the whole year. The frailejon giant-rosettes, for instance, at night protect their buds in the center of the rosette by adjusting the position of their leaves. In addition, these plants have large intercellular air spaces and excrete mucilage that contributes to their heat-storage capacity. The discontinuous areas of paramo ecosystems are isolated by montane wet (cloud) forests that grow at lower altitudes and as a result there is a large number of endemic species. Nutrient dynamics in high Andean ecosystems are not fully documented. In general, low temperatures and moist acid soils slow decomposition rates, and a thick layer of organic humus usually accumulates in sites where the relief allows deposition.

Xeric Formations

In South America there is a variety of xerophytic ecosystems, all marked by an intensive and prolonged period of seasonal drought during which the vegetation suffers water deficit.

Distribution and Structure

A mean annual precipitation of 400–1200 mm and more extreme daily changes of temperature characterize the Dry Climate Steppe (*BS)* climate. Soils are brown and yellowish, reflecting a higher content of humus. The *BS* climate coincides with the geographic distribution of semiarid environments and generally represents transitions between desert environments of the type Dry Climate Desert (*BW)* and savanna environments of the type *Aw* (**Figures 2** and **3**). In South America, xeric ecosystems are represented by the xeric scrubs of La Guajira peninsula in Colombia and Venezuela, the Paraguaná xeric scrub of Venezuela, the dry forest of west Ecuador, the Sechura and Atacama deserts in Chile, the Gran Chaco region in southern Bolivia, west Paraguay and north Argentina, and the Caatinga xeric scrub of northeastern Brazil (**Figure 3**). The dry highland plateaus of the Puna region and the thorn woodlands of Patagonia are usually included within the dry formations of the continent. There is also a xeric formation called "agreste," which forms a narrow strip between the Caatinga and the seasonal forest of eastern Brazil. Although not a strictly xeric ecosystem, the Matorral of Chile is also usually included as a semiarid type of vegetation (**Figure 3**).

Deciduous forest trees that severely restrict their transpiration during the middle of the day – even during the rainy period – characterize the thorny woodlands of Colombia, Venezuela, and northeastern Brazil. Significant areas of the Guajira region consist of sand, moving dunes. Only a few trees are evergreen, but along the periodically dry channels of streams, riparian vegetation retain their leaves and reflect the presence of groundwater. The mean canopy height of the woody components is 5 m, sometimes reaching 7 m. The most important genera in these plant communities are *Acacia*, *Bulnesia*, *Bursera*, *Caesalpina*, *Capparis*, *Croton*, *Jacquina*, *Opuntia*, and *Prosopis*. The caatinga ecosystems in the larger "Sertao" depression of northeastern Brazil are dominated by the genera *Canavillesia*, *Chorisia*, and *Schinopsis*. Communities on the dry tablelands are dominated by the genera *Spondias*, *Commiphora*, *Cnidoscolus*, *Aspidosperma*, and *Mimosa*. The driest parts of the Sertao are dominated by the genera *Mimosa*, *Auxema*, *Combretum*, and *Aristida*. The Agreste formation between the Caatinga and seasonal forest of Brazil has an open canopy of xerophytic trees (mostly Leguminoseae and Myrtaceae) with a very scanty understory of palms (e.g., *Copernicia cerifera*) and cactus.

The ecosystems of the Gran Chaco region are structurally diverse, reflecting the varied environmental changes observed across the Chaco range. Precipitation increases from the center of the region, both to the east and west. The landforms are flat in the west and north and during the rainy season, flooding is common. The thorn woodlands of Patagonia occupy the west margin of the Pampa region, where annual precipitation is lower (300–500 mm) and the woody components of the vegetation (Acacias and palms) are characteristically undersized.

Owing to the effect of the Andes and climatic circulation patterns (*see* Climate), an arid zone occupies the zone from the coasts of Peru and Chile across the mountains into the Patagonian steppes (**Figure 3**). The Pacific coastal area is a desolated desert where rain can be absent for years. The

Atacama Desert is the driest climatic region of South America and occupies the eastern Pacific coast of the continent from about 5° to 26° S (**Figures 2** and **3**). In many areas there is no fog to provide moisture for vegetation and the cover is dominated by lichens and algae that depend on water condensed as dew. In those areas where fog and clouds develop, bringing a few millimeters of annual precipitation, there are some short-cycled annual plants from the genera *Aristida*, *Loasa*, *Malesherbia*, *Nolana*, and *Tropelum*. On the west-facing slopes, between 100 and 150 m, moisture from fog carried by the Humboldt Current allows the presence of some vegetation. Some of the common species in these ecosystems are *Prosopis juliflora, Distichilis spicata, Capparis angulata, Sporobolus virginicus, Acacia macracanta*, and the giant cactus *Neoraimondia gigantea*. In the high Andean slopes, the lowland forests and the grassland Puna limit the dry ecosystems of the Andean Peruvian valleys. In the dominant genera there are *Acacia*, *Bombax*, *Bursera*, *Caesalpinia*, *Cercidicum*, *Prosopis*, *Puya*, and several cacti.

The Mediterranean Matorral occupies the coastal west-facing slopes of the Andes in central Chile (**Figure 3**). Plant communities in these ecosystems are composed of evergreen shrub species that reach 3–5 m in height and have characteristic sclerophyllous leaves (e.g., *Lithaea caustica* and *Fluorencia thurifera*), while succulents and herbs cover the ground. Some Acacias and Nothofagus are present in communities of higher areas and those found further inland. Despite its latitudinal position, the flora of the Chilean Matorral is very rich, with more than 2000 species reported.

Functional Aspects

There are only a few studies on the physiology and phenology of South American xeric ecosystems. Most of these ecosystems are characterized by at most 3–5 months of rainy season with not more than 400–800 mm per year. In tropical lowland areas, temperatures are 23–27 °C year round. Relative humidity is stable around 50% and the potential evapotranspiration is usually 1500–2000 mm per year. Plant communities are affected by water deficit, which varies greatly from region to region, and so do the physiognomic and physiologic responses of the vegetation. Some annual plants germinate, grow, and reproduce only during periods of available water (e.g., most herbaceous). Other plants evade stress conditions by restricting growth to periods of available water, while the shedding of leaves is a common adaptation to cope with the long dry periods. Since leaves are only present during the rainy season, many species have not developed mechanisms to control evapotranspiration. In addition, many species from these ecosystems produce leaves with chemicals that inhibit the germination of other plants. CAM is common and succulent perennials use this type of photosynthesis to endure the scarcity of water.

Annual precipitation of about 300–900 mm and a dry summer climate characterize Mediterranean ecosystems. The dry, hot weather favors the occurrence of fires, which is the main disturbance in these ecosystems. Topographical conditions are also important in Mediterranean ecosystems, where erosion and continuous rearrangement of surface materials are common. In Chile, the duration and intensity of the drought period increases southward and Matorral communities become dominated by deciduous evergreen shrubs and succulent species. The climatic conditions allow all shrub species to use C3 photosynthesis to fix carbon and there are not species with C4 metabolism. As already pointed out, the availability of soil resources is important in determining the investment that plants make on belowground biomass. In the Chilean Matorral it has been estimated that shrubs have an average root to shoot ratio of 1.5 and the roots spread in a radius at least two times the area of their crowns. Most of the fine absorbing roots, however, are found under the crown where they absorb nutrients from decomposing leaves and debris. Water availability appears to be the most important controlling factor of productivity in Matorral ecosystems. Annual productivity in these ecosystems range from 2 to 6.5 kg m^{-2}.

Coastal Ecosystems and Wetlands

Mangroves

In South America, mangroves occur along fragments of the Atlantic coastlines of Colombia, Venezuela, Guyana, Surinam, French Guyana, and Brazil, and the Pacific coastlines of Colombia, Ecuador, and Peru (**Figure 3**). Mangroves develop well in areas where the ocean temperature remains above 24–27 °C. Because of the effect of ocean currents (*see* Climate), the southern limits of mangroves in South America occur at about 25° S in the Atlantic coast of Brazil and 4° S in the Pacific coasts of Ecuador and Peru.

These woody communities can grow up to 30 m. The mangle is best developed in coastal areas where high rates of sedimentation are enhanced by the mangroves themselves. Local topography, runoff channels, and sediment stability determine distinct zonal patterns. The vegetation of mangrove ecosystems (mangal) is adapted to salinity and is dominated by *Rhizophora mangle, Avicenia* sp. and in high ground by *Laguncularia racemosa*. Plant communities directly influenced by seawater but occupying high grounds are characterized by genera typical of seashore environments: *Remirea, Salicorna, Canavalia, Acicarpa*, and *Vignia*.

Mangrove areas are sometimes associated with river deltas and coastal swamps that have habitats in which the soils are inundated for at least part of the year. In the mouth of the Amazon, however, mangroves do not grow well. According to the quality of the water, it is possible to distinguish freshwater swamps and brackish water swamps. To the first type belong the freshwater swamp forests of the Atlantic coast between the Orinoco delta and the mouth of the Parnaiba River. Most of the other mangrove swamp forests belong to the second type. Other conspicuous coastal ecosystems are the Restingas of the Brazilian coast. These ecosystems are characterized by sand dunes where low scrubby vegetation grows sparsely. Their flora is rich in endemic species.

Wetlands

In South America, there are many dispersed wetland ecosystems. Some savanna areas of the Llanos region in the Orinoco watershed and the Rio Branco in Brazil have bad drainage and are seasonally inundated. In Ecuador, west of the Daule River, there is also a region of wetlands. The complex of

grassland and woodland savanna vegetation occupies flat plains with tropical seasonal climate. The different types of vegetation are closely associated with the duration of the period of inundation. Areas with a long inundation period maintain communities dominated by the genera *Euterpe* and *Mauritia*. The genera *Typha*, *Cyperus*, and *Juncus* dominate swampy areas. In well-drained floodplains, the dominant genera are *Panicum*, *Paspalum*, and *Thalia*.

In western and eastern Amazonia there are also some areas of flooded grasslands but larger wetland ecosystems usually occupy the floodplain of large rivers (e.g., varzea and igapo in Amazonia) and the structure of the plant communities depends on the level and duration of the flooding period (*see* Tropical Forest: Functional Aspects). The larger wetlands of the continent are in the Gran Pantanal region in Brazil, Bolivia and Paraguay, and the Paraná flooded savannas in Argentina (**Figure 3**). The Pantanal occupies the upper watershed of the Paraguay basin and is one of the world's largest wetlands, which is flooded in more than 80% of its area during the rainy season (from May to December). The geomorphology is essentially flat and small high-ground islands separate hundreds of small lagoons. Cerrado tree species and palms occupy the high-ground islands of the Pantanal. The region accommodates large seasonal migrations of wildlife including aquatic birds and caimans.

Temperate Forests

In South America, temperate forests have a limited distribution and are restricted to relatively small areas. These ecosystems occur in the southern Andean areas of Chile and Argentina (**Figure 3**). Evergreen temperate forest occupies areas of frost-free oceanic climate. These ecosystems are dominated by tall trees, some reaching heights up to 50 m, and composed of species such as *Eucryphia cordifloria* and *Laurelia philippiana* and some conifers like *Araucaria araucana*.

In the western humid slopes of the Andes, temperate rain forests are divided into the Valdivian and Magellanic forests. The Valdivian rain forest has an upper limit of 500 m and occurs from 40° S to 47–49° S where it becomes the Magellan forest. Lush vegetation rich in evergreen species, abundant epiphytes, and lianas characterizes Valdivian ecosystems. The Valdivian forests can support an extraordinary biomass, in part because the climate is very humid and the mean annual temperature ranges from 10 to 12 °C. In areas with excessive humidity, coniferous trees (*Fitzroya patagonica*) – whose stands can reach heights of 50–60 m – replace common species of the rain forests. Coniferous forests also grow well at higher altitudes up to the tree line, which in the southern cone occur at about 1600–1900 m above sea level. Between the Valdivian forests and the sclerophyllous vegetation of central Chile there are limited extensions of deciduous forests composed of *Nothofagus obliqua*, which are found up to 1200 m above sea level.

The Magellanic forests occupy the western slopes of the Cordillera Patagonica in southern Chile. In this region, the coastline is very rocky, curved by fjords, and has an abrupt slope. Abundant rainfall, low evapotranspiration, and strong winds characterize the climate. The Magellanic forests are poor in species. As precipitation decreases rapidly to the east, in the leeward slopes, deciduous forests occur and are composed almost of pure stands of *Nothofagus* with an understory of small evergreen trees. These forests grow to heights of about 30 m, but they become shorter as altitude increases. The tree line falls from 1200 to 1400 m in the north to about 600 m in the region of Ushuaia.

See also: Alpine Ecosystems. Amazon Ecosystems. Biodiversity-Rich Countries. Climate, Effects of. Desert Ecosystems. Endangered Ecosystems. Grassland Ecosystems. Mediterranean-Climate Ecosystems. Neotropical Seasonally Dry Forests. Rainforest Ecosystems, Animal Diversity. Rainforest Loss and Change. South American Natural Ecosystems, Status of. Wetlands Ecosystems

References

Archibold OW (1995) *Ecology of World Vegetation*. London: Chapman & Hall.
Beard SJ (1955) The classification of tropical American vegetation types. *Ecology* 36(1): 89–100.
Bush MB (1994) Amazonian speciation: A necessarily complex model. *Journal of Biogeography* 21: 5–17.
Bush M (2003) *Ecology of a Changing Planet*, 3rd edn. Upper Saddle River, NJ: Prentice Hall.
Colinvaux PA (1989) Ice-age Amazon revisited. *Nature* 340: 188–189.
Coupland RT (1992) Overview of South American Grasslands. In: Coupland RT (ed.) *Natural Grasslands: Introduction and Western Hemisphere*, pp. 363–365. Amsterdam: Elsevier.
Dinerstein E, Olson DM, Graham D, Webster A, Primm S, and Bookbinder M (1995) *A Conservation Assessment of the Terrestrial Ecoregions of Latin America and the Caribbean*. Washington, DC: The World Bank & World Wildlife Fund.
Edit RC (1968) The climatology of South America. In: Fittaku E, Illes J, Klinge G, Schwabe G, and Sioli H (eds.) *Biogeography and Ecology in South America*. The Hague: Junk.
Eiten G (1972) The cerrado vegetation of Brazil. *The Botanical Review* 38(2): 201–341.
FAO-Unesco (1971) *Soil map of the World, 1:5.000.000*. Paris: Food and Agriculture Organization of the United Nations and the United Nations Educational, Scientific and Cultural Organization.
Gentry AH (1995) Diversity and floristic composition of Neotropical dry forests. In: Bullock SH, Mooney HA, and Medina E (eds.) *Seasonally Dry Tropical Forests*, pp. 146–194. New York: Cambridge University Press.
Haffer J (1969) Speciation in Amazonian forest birds. *Science* 165(3889): 131–137.
Harrington H (1973) Geology of South America. In: Fairbridge RW (ed.) *The Encyclopedia of Earth Sciences*, vol. VIII, pp. 456–465. Stroudsburg: Dowden, Hutchinson & Ross.
Hoorn C and Wesselingh FP (eds.) (2010) *Amazonia Landscape and Species Evolution a Look Into the Past*. West Sussex, UK: Blackwell Publishing.
Leigh EG (1975) *Structure and Climate in Tropical Rain Forest. Annual Review of Ecology and Systematics* 6: 67–86.
Lovejoy TL and Hannah L (eds.) (2005) *Climate Change and Biodiversity*. New Haven: University Press.
Murphy PG and Lugo AE (1986) Ecology of tropical dry forest. *Annual Review of Ecology and Systematics* 17: 67–88.
Sarmiento G (1984) *The Ecology of Tropical Savannas*. Cambridge: Harvard University Press.
TerborghJ (1992) Diversity and the Tropical Rain Forest. New York.
Whitmore TC and Prance GT (1987) *Biogeography and Quaternary History in Tropical America*. Cambridge: Cambridge University Press.

Ecotoxicology

JM Lynch, A Wiseman, and FAAM De Leij, University of Surrey, Guildford, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 363–373, © 2001, Elsevier Inc.

Glossary

Bio-accumulation The accumulation of toxicants or other chemicals by successive organisms in the food chain.
Biological integration Organization of life functions in distinct, self-regulating units (cells, tissues, organs, organisms, populations, communities, and ecosystems).
Biosphere The region of the earth where life exists, including parts of the lithosphere, hydrosphere, and atmosphere.
Ecosystem A self-regulating assembly of communities of animals, plants, and microorganisms interacting with one another and with their nonliving environment.
Endocrine disruptor Chemical that interferes with the chemical communication (mainly performed by hormones) within an organism.
Remediation Removal of toxicants from a contaminated environment using chemical, physical or biological means.
Resolution The magnitude of a biological response due to a chemical insult.
Specificity The ability to relate a chemical effect to a distinct biological function or organism.
Xenobiotics Toxic substances not naturally produced within organisms.

Introduction

The manifest principles of ecotoxicology involve the application of the principles of toxicology to the environment: focused on human activity leading to the release of molecules such as polynuclear aromatic hydrocarbons (PAH), polychlorinated biphenyls (PCBs), xenoestrogens (XEs), chlorofluorocarbons (CFCs), pesticides, heavy metals, radionucleotides, greenhouse gasses (CO_2, CH_4), sulfur dioxide (SO_2) and oxides of nitrogen (NO_X) into the environment.

Biomonitoring is undertaken sometimes as the "prelude to remediation." Remediation aims to remove potential toxic substances from a contaminated site, thus restoring ecosystem function as far as is reasonably possible following the removal of the ecotoxicity by biological, physical, or chemical means.

Molecular toxicology is the branch of toxicology (the study of the effects of manifestly poisonous substances on individuals) that adopts the biochemical approach to the understanding of the detrimental threats to life, often recognized by a range of morbid conditions in a variety of fauna and flora. Such pathology can be understood at a molecular level in terms of biomolecular damage undergone by such macromolecules as DNA (deoxyribonucleic acid: the genetic determinant mainly residing in the nucleus of the living cell), RNA (ribo-nucleic acid: responsible for transfer of the genetic message inherent in DNA structure sequence to the cytoplasm of the cell), and diverse proteins constructed from a choice of sequence of residues of 20 amino acids (obtained mainly from dietary protein hydrolysis during digestion).

It is the increased understanding of chemistry and its associated biochemistry that allows a meaningful prediction of toxicity of chemicals in all forms of life (ranging from bacteria and higher fungi to plants and animals including *Homo sapiens*). Although no meaningful distinction can be made between natural and manmade chemicals released into the environment, detoxifying enzymes such as the ubiquitous oxygen-requiring enzyme family, cytochromes P-450, present in most forms of life, are useful biomarkers for constitutive or acquired tolerance to chemicals such as polycyclic aromatic hydrocarbons (PAHs), polycyclic halogenated biphenyls (PCBs), and xenoestrogens (XE). In the course of millions of years of biological evolution of species, survivors now display a range of such biochemical defenses against toxicants including the more recent pollutants released through industrial activity in processing and manufacturing industries.

Particular concern has been displayed recently for the contamination of the environment with endocrine disruptors (environmental hormones). These mimic the biochemistry of natural estrogens (such as 17β-estradiol) and these have been associated with feminization of males, leading to displayed hermaphroditism in a few cases, in polluted environments. Some of these endocrine disruptors are derived from the release of industrial chemicals and are therefore referred to as xenoestrogens (foreign-compound estrogens). This is using the accepted nomenclature of xenobiotics as "foreign" or "manmade" chemicals, as contrasted with natural chemicals of which the environment, the diet, and all life is composed.

Several important concepts have been incorporated into the principles of ecotoxicology. For example, "hazardous chemicals" have the potential for causing harm to a variety of organisms, but the assessment of the risk to the biosphere posed by these chemicals must be based on consideration of the exposure (dose) to individuals, populations and communities as well as direct toxicity to the different species involved. Environmental exposure is not only determined by the quantity of a chemical released but depends also on the characteristics of a particular environment and the distribution of the biota in that environment. Soil that contains large quantities of clay minerals and humic substances might lead to immobilization of hazardous chemicals through absorption, while chemicals released in aquatic environments might be dispersed rapidly. In both cases, toxicity to the biota

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00043-5

living in these environments might be limited as exposure (bioavailability) is small. Also the distribution of different species in the environment is not random, but populations of a particular species are concentrated at particular sites (niche). For example, in soil some species are found in the litter that lays on the soil surface (epigeic and hemigeic species), while others frequent the deeper mineral soil layers (eudaphic species). It is clear that epigeic and hemigeic species are more likely to be affected by toxic chemicals that enter the soil environment from above than eudaphic species. Besides distribution, exposure will be determined by feeding patterns. Organisms that feed on prey that is likely to have accumulated the substance will be exposed to larger quantities of the toxicant than those that feed on substances that are relatively uncontaminated. Especially recalcitrant, low water soluble chemicals such as PCBs and organochlorines (DDT, for example) can be biomagnified in the environment to lethal concentrations higher up the food chain. Furthermore, toxicity of each hazardous chemical should be related to the metabolism of each major species present, as affected by the toxicant. Therefore, signs of toxicity need to be sought by a battery of methods before meaningful conclusions can be drawn on the environmental risk of a potential chemical released in the environment; environmental pollutants at a specified dose may cause changes in finely balanced ecosystems due to differential toxicity to different species and differential exposure of different species that inhabit the environment.

Much of ecotoxicology relates to the toxic effects of chemicals in natural ecosystems, but it must be remembered that man is an integral part of these ecosystems and effects on the other biotic components of the ecosystems of which he is part can affect him directly or indirectly. For example, in relation to bioaccumulation of toxic chemicals, substances like PCBs, chlorinated organopesticides, and heavy metals, are likely to accumulate (and cause toxic effects) in the human body because of the simple fact that man is high up in the food chain. Indirectly, man is dependent on the biosphere for the maintenance of the climate, purification of water and air, and the provision of a sustained food supply. Disruption of biological activity that impairs the capacity of the living world to fulfill these functions will ultimately affect humankind's well-being and survival. In this respect, therefore, human toxicology and ecotoxicology are of necessity closely interlinked.

Application of Different Scientific Disciplines to Ecotoxicology

Chemistry

Many different xenotoxicants of many chemical structures are now ubiquitous in aquatic, terrestrial, and atmospheric environments. Most, however, are present at extremely low concentration. For example, there may be 60,000 molecular-species of "hormone-disrupters" such as the xenoestrogens that may be detected (albeit by remarkably sensitive detector systems that can respond to even a few hundred molecules for a positive identification). Nevertheless, most of these xenoestrogens are extremely weak molecular-mimics of human estrogens, and many of these can, paradoxically, stop the estrogenic response of particular target tissues.

Proposed chemical approaches to ecotoxicology are often analytical in nature and seek to quantify the effect of each chemical species that can be identified and assayed so as to define the concentration-dependence of any hazardous substance and its consequent environmental or human risk. Quick tests for biohazards are essential, such as that achieved in the famous Ames Test for mutagenicity (and related carcinogenicity) in selected bacteria, for example, by the polyaromatic hydrocarbon, benzo(a)pyrene. Such mutagenicity (novel growth-behavior) due to genetic (DNA) damage may manifest as cancer in mammalian organs (because of loss of growth-control in particular tissues, due to subtle alterations in regulation of cellular growth by particular genes).

Biochemistry

Biochemistry studies the chemical processes that take place in living organisms. The main pertinent principle of biochemistry that can be applied to ecotoxicology is the modern concept of enzyme (biocatalytic proteins) regulation by molecular-intermediates of tissue metabolism in cells. Many xenobiotics (foreign chemicals unusual in nature) can interfere with the finely balanced biochemical reactions of living cells by perturbation of the web of molecular interactions necessary for life.

Additional problems are due to the molecular damage to DNA (deoxyribonucleic acid) and its related RNA (ribonucleic acid) along with cellular membranes (phospholipids that are readily destroyed by reactive oxygen species (ROS), generated by reactions of these chemicals with atmospheric molecular-oxygen). The ecotoxicity of ROS may, however, be ameliorated by antioxidants in the environment (and by dietary antioxidants such as vitamins C, E, A, and D as well as phytoestrogens in the diet).

Microbiology

The activity of microorganisms is at the basis for most functions on which life on this planet depends. Due to their almost limitless metabolic capacities, microbial communities are responsible for the transformation and recycling of organic and inorganic molecules in the environment. This activity results in the maintenance of nutrient cycles, the maintenance of soil fertility, and the detoxification of toxic substances in the environment. Furthermore, many microorganisms are intimately associated with every conceivable higher life form on this earth. They perform functions without which higher organisms could not function. Examples of this kind are organelles that have a microbial origin, including mitochondria, which are responsible for the energy generation in eukaryotic organisms (organisms with a true nucleus), and chlorophyll that allows algae and higher plants to convert carbon dioxide into sugars via photosynthesis. Other symbiotic interactions between microorganisms and plants that are of crucial importance for plant growth and nutrition are dinitrogen fixing microorganisms and mycorrhizae, which have close

associations with the roots of many plant species. In fact plant roots have adapted specially to accommodate for these microorganisms. Furthermore, microorganisms associated with the gut of animals provide essential vitamins that animals cannot synthesize themselves. On the negative side, microorganisms are responsible for causing a wide variety of diseases in plants and animals. Although this is negative from the viewpoint of the organism (or population) that is affected, disease causing organisms are important in ensuring that excessive population growth of particular populations is curtailed, thus ensuring the maintenance of an environment that contains a rich diversity of species. The interdependence between microorganisms and higher life forms means that toxic effects that affect either group of organisms can have important consequences for the other.

Biotechnology

Biotechnology is an applied science that aims to harness different life forms for the benefit of man. Agriculture, antibiotic production, and bioremediation are examples of such applications. Recent advances in molecular biology have given biotechnologists new tools to change the nature of life by genetic modification. This technology has resulted in the creation of organisms that would very unlikely have arisen in nature via normal mating and exchange of genetic material. It is now possible for humans to combine the genes of widely different organisms into one organism, giving this recombinant organism the means to express novel characteristics. Such "novel characteristics" include resistance to diseases and pests, resistance to pesticides, or the means to produce products that are of commercial or medical value. Whereas these organisms are of potential value to society, it is necessary to evaluate their environmental impact before they can be safely released into the environment. Therefore, with the advance of biological science comes the responsibility for biotechnologists to ensure that the integrity of the environment is maintained by ensuring that the biological functions that allow life on this planet to thrive are not impaired. Interestingly, recombinant technology has also opened the way to construct organisms that can be used to monitor the environmental impact of toxicants more sensitively. An important interface of ecotoxicology with biotechnology arises in the use of such genetically modified organisms as biomonitoring tools. However, the detection (and assay) of ecotoxicants can be achieved by a variety of methods. An interesting example is the detection of carbon monoxide. This can be achieved using a caged canary (that responds sensitively to carbon monoxide in coal mines) or using an enzyme-based detector electrode that functions as a biosensor for this gas. A detailed account of biotechnological techniques for biomonitoring and bioremediation in relation to ecotoxicology can be found in the book by Lynch and Wiseman (1998).

Ecology

Ecology is the study of the interactions between organisms and their environment that determine the distribution and abundance of organisms. These interactions can be studied at different levels of biological integration starting from individuals, which are part of populations, which are part of communities, which are part of ecosystems, which are part of the earth's biosphere. To gain a living from nature, humans have to understand the interactions that determine the abundance and distributions of organisms on which human life depends. This is not only important for the harvest of species from natural ecosystems by hunting, fishing, and or gathering of plant products, but also where the aim is to modify the natural environment to yield predominantly products that are of economical value as is the case in agriculture and forestry. Not only do we need to understand how we can most efficiently use the environment, but also increasingly we need to be aware of the impact of human activities on ecosystem function in general. Pollution of the environment by human activities (manufacturing, mining, waste disposal, transport, energy use, pesticide use, etc.) can have a large impact on ecosystem function via effects at all levels of biological integration. Especially since the scale of human activities has increased so dramatically over the past century there is a real danger that human activity interferes significantly with the basic biological processes of sustained life, including our own. Therefore, ecological knowledge applied to ecotoxicology aims to evaluate the consequences of human activity to ecosystem function, especially in relation to pollution caused by human activity.

Toxicology

Important principles of toxicology that relate to ecotoxicology include the concepts of extent of exposure, persistence, and distribution of chemicals in the environment. Subsequently predictions can be made on the toxicity of such chemicals to individual organisms or populations. The starting point of such analysis is often the chemical structure of each toxin, which will allow some prediction of the behavior of the chemical in the environment to be made. One of the best examples of this kind is the environmental impact assessment of the insecticide DDT. DDT is almost completely insoluble in water but readily soluble in fat. Furthermore, DDT and its degradation product DDE are highly persistent. Its low solubility means that it is easily dispersed in aquatic environments (in the case of DDT it is found at low concentrations all over the globe) and only accumulates in places with a high fat content (i.e., living organisms). Once it has entered living organisms, it persists and accumulates in the fat tissues of organisms that are higher up the food chain resulting in toxicity. For this reason, populations of both fish eating birds and bird eating raptors such as peregrines and hawks were badly affected by DDT, even in cases where they inhabited pristine environments.

Another aspect of toxicology is risk analysis. This quantitative topic is most easily studied when death rates within a population can be quoted. Whereas the rates can be determined experimentally for animals, plants, and microorganisms, the determination of the effects on the human population is inevitably more difficult to determine, and one of the few situations where effects are well known is the mutagenic effects of radionucleotides, resulting from

contamination caused by accidents at nuclear power installations or nuclear warfare.

Entry, Movement, and Fate of Pollutants in Ecosystems

Organic and metallic (including radionucleide) toxicants that enter the environment are regarded as pollutants of air, soil, and water. Their movement and transport is through air and water depending on their volatility and solubility, respectively. Soil and the detritus of sediments provide important solid supports for adsorption to regulate the movement and flow of the pollutants through terrestrial ecosystems, as well as influencing the localization and persistence of the pollutants in the environment. In this respect clay and humus particles that are electrostatically charged have an important role to play. Some organisms will take up the pollutant and concentrate them in their cells, a phenomenon known as bioaccumulation or biomagnification. This is particularly serious where the toxicant enters the food chain leading to progressive accumulation of toxic molecules higher up the chain, including humans. This is often the case with molecules that are relatively biologically inert such as heavy metals, PCBs, and organochlorine insecticides. Accumulation of such molecules in different species can lead to toxicity expressed as reduced growth, reduced fecundity, changes in behavior, susceptibility to diseases, or increased mortality. However, on the positive side, where there is storage of the pollutant in the cell, as is normally the situation with (heavy) metals, it is possible that the pollutant might be harvested and therefore removed from the environment. This especially applies to plants that hyper accumulate metals in their tissues. The application of this process for the cleanup of contaminated land is known as phytoremediation. Also, many pollutants are biologically degradable and can therefore be detoxified. Both plants and mammals have a range of enzymes that are involved in detoxification of molecules that are potentially harmful to them (cytochromes P-450, for example), whereas the almost limitless metabolic capacity of a wide variety of microorganisms allows degradation of pollutants, especially hydrocarbons, into nontoxic molecules such as water and carbon dioxide. The stimulation of microbial degradation and metabolism of pollutants is known as bioremediation and is currently used to clean up contaminated land while microbial degradation of pollutants present in sewage and water is used in a variety of water purification systems. The fate of pollutants in the environment is therefore not only dependent on the molecular characteristics of the pollutant but also on a range of biotic and abiotic factors.

Biomonitoring

The ultimate goal of biomonitoring is to use biological effects resulting from chemical exposure for making predictions and deductions about the quality of the environment for life in general and human life in particular. Biomonitoring activities can be conducted on various levels of biological integration, over a period of time. Several overlapping terms are commonly used in this process. It is worth noting that some of the tools are, strictly speaking, alternatives for chemical/physical methods to measure bioavailability of chemicals in the environment. In other words, they do not provide information on ecological effects per se.

- The term *bioreporter* is used for a molecular tool, often genetically modified cells, to transform the presence of a chemical into an easy measurable signal (e.g., luminescence).
- The measurement of biochemical and physiological variables in individuals or in their excretion products, providing information on exposure or damage, is indicated by the term *biomarker*.
- A *biosensor* or *bioprobe* is a physical device that allows the detection of a chemical as an electrical signal derived from a biocatalyst, such as an enzyme or an antibody.
- A *bioassay* is a toxicological test system in which the activity of a chemical is measured as an adverse effect on a population of a specific test species.
- A *bioindicator* is used to detect the effects of chemicals at the individual, population, community, or even ecosystem level.

Important in this respect is that when we move up to the higher levels of biological integration (populations, communities, ecosystems), effects are increasingly difficult to link to a specific chemical, even though the effects measured are of more importance. Furthermore, it takes an increasingly long time for higher levels of integration to respond to a chemical insult. For example, DDT took years to bioaccumulate up the food chain resulting in negative effects at the population/community level. On the other hand, chemicals that are apparently nontoxic at the lower levels of biological integration (carbon dioxide, chlorofluorocarbons (CFCs), sulfur dioxide, etc.) can have profound effects on all levels of biological integration. These chemicals might not directly affect biological processes but affect the chemical balance of the earth's biosphere (air, water, and soil). For example, greenhouse gases (carbon dioxide, methane, and water vapor) trap the sun's long wave radiation augmenting global warming, which could lead to profound changes of the world's ecosystems. CFCs contribute to the destruction of the earth's protective ozone layer, leading to an increase in harmful UV radiation that reaches the earth, resulting in increased lethal mutation rates and cancer. Sulfur dioxide and NO_X emissions lead to acid rain and acidification of soil and water, which in turn leads to the release of toxic quantities of metal ions in soil and water that kill both vegetation and animal life. An interesting example of this kind is the acidification of lakes in Scandinavian countries as a result of acid rain. The resulting lower pH of the lake water means that the water can contain a higher concentration of aluminum ions. However, the pH around the gills of fish is slightly alkaline, resulting in the precipitation of aluminum on the gills. This results in the impairment of gill function and suffocation of the fish. Even though Scandinavian lakes affected in this way look perfectly clean, they are devoid of fish.

A variety of specific tests have been employed to investigate the effects of chemicals on the environment. Generally the

Table 1 Some tests of chemicals that affect soil function

Retention capacity of soil and endangering of groundwater	*Living space for plant production*	*Living space for soil communities*
Growth (algae)	Dehydrogenase (*Bacillus cereus*)	Biomass (microorganisms)
Immobilization (*Daphnia*)	Nodulation (*Rhizobium*)	Enzyme activity (microorganisms)
Light emission (*Photobacterium phosphoreum*)	Shoot or root growth (various plants)	Nitrification (bacteria)
Mortality (fish, nematodes)	Yield (various crops)	Mortality (earthworms)
Mutagenicity (Ames, Umu)		Reproduction (collembola, earthworms)
		Respiration (microorganisms)

biotic effects used for biomonitoring deploy biota from the environment and expose them to different concentrations of the test chemical. Mortality of fishes, for example, has been used to assess toxicity of chemicals that are used in the soil environment (**Table 1**). In the soil environments a variety of effects can be investigated as to whether the retention capacity of soil would affect the toxicity of a chemical leading to leaching into groundwater or the living space for plant production and soil communities. In this respect, the assays relate to the effects on populations, gene products (enzymes and metabolites), or impact (plant bioassays) (**Figure 1**). However, the exciting development of recent years is the recognition that these effects are all controlled by genes and another method of biomonitoring is to use gene (DNA) probes (**Figure 2**). This option is very much in its early stages and the concept is to link a gene, which responds to a specific chemical or general effect, to a promoter DNA sequence and a reporter DNA sequence that either encodes for an enzymatic effect or any other specific signal. It is therefore possible to use genes, which respond to mutagenic substances with a light signal that is proportional to its mutagenic potential in the environment. Whereas this approach offers much promise, the reality is that it is complementary to the other methods currently deployed and should be used in conjunction with them.

A particular useful activity, which has been monitored in aquatic environments, is cytochrome P450. (**Table 2**). This is the terminal component of the mixed-function oxygenase, which catalyses expoxidation, hydroxylation, dealkylation, and desulfurisation, all critical in detoxification of toxic molecules in invertebrates and vertebrates (including humans). To date, more than 750 cytochrome P450 genes have been identified (Nelson, 1998). The cytochrome P450/A1 (CYPAI) isoform is a biotransformation enzyme involved in a wide range of xenobiotic metabolisms (chemicals foreign to the natural environment) including metabolism of polyaromatic hydrocarbons (PAHs). Usually CYPIAI has been used as a biomarker in fish and mollusks. The linking of this enzymatic function with its genetic control offers a great deal of potential for the biomonitoring of toxicology of both marine and freshwater ecosystems.

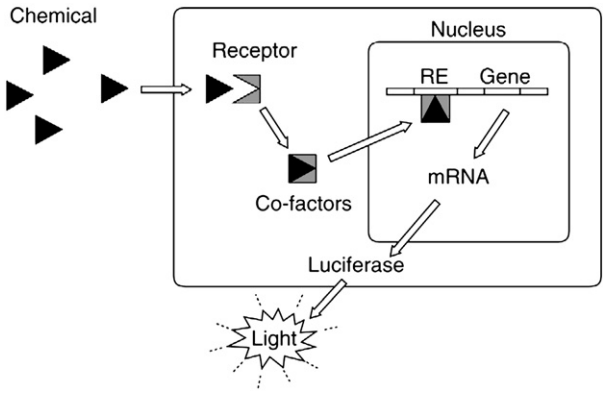


Figure 1 Receptor/reporter gene assays; RE: responsive element.

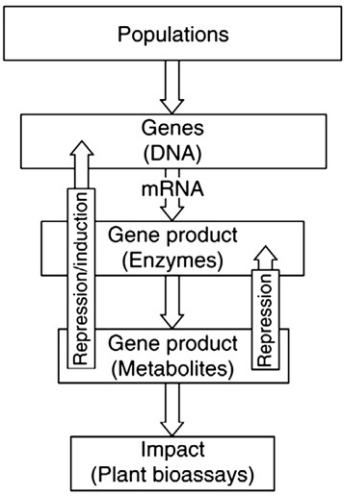


Figure 2 Levels of perturbation measurement.

Biological Processes

The question can be raised why it is necessary to analyze complicated biological processes if the chemical in question can be measured directly using chemical or physical methodologies. There are at least three situations where biomonitoring provides answers that cannot be arrived at using direct chemical measurements. The first use for biomonitoring is in situations where one wants to trace the chemical history of an environment. For example, the pH history of lakes can be deduced from the percentage acidophilous and alkalophilous species within the diatom communities preserved in sediment cores (Renberg and Hellberg, 1982). Similarly, pollen records preserved in soil provide information on past vegetation from which the past environmental conditions can be deduced. The second use of

Table 2 Hepatic cytochrome P450 1A (CYPIA) as a biomarker of organic pollution in fish

Fish species
Atlantic tomcod (*Microgadus tomcod*), cod (*Gadus morbua*), Largemouth bass (*Micropterus salmoides*), Rainbow trout (*Oncorbynchus mykiss*), Whitefish (*Coregornus lavaratus*)
Point sources of pollution
Chemical plants, incineration plants, industrial complexes, landfill sites, oil industry, oil spills, natural oil seeps, pulp mills, sewage
Diffuse areas of pollution
Bays, bights, estuaries, fjords, gulfs, harbors, lakes, offshore areas, river systems.

biomonitoring is a situation where the activity of a chemical is short-lived but the effects are persistent. For example, pesticides that disperse rapidly or are broken down rapidly cannot be measured easily, but their biological effects might persist for long periods. The insecticide deltamethrin is active only for a few hours after spraying, while effects on sensitive ground spiders and beetle fauna remain visible for weeks after spraying (Everts, *et al.*, 1989; Jagers op Akkerhuis, (1993). The third reason, and probably the most important one, for using biomonitoring is that effects of chemicals are dependent on the interactions that take place within the environment itself after their release. These interactions can be of a physical/chemical nature where the chemical is immobilized on to soil particles and is therefore not bioavailable. For example, the toxicity of metals, such as copper and zinc, in soil is determined by their bioavailability in soil. Both copper and zinc ions bind to soil minerals (mainly clay particles) and humic substances and as such are unavailable to the biota present in soil. Soils contain large quantities of these potential binding sites and therefore toxic effects of heavy metals are often buffered, even at relatively high concentrations. Because of this soil-buffering capacity, it usually takes some time of sustained metal inputs before metal ions become bioavailable at concentrations that are toxic. Both the mineral and organic fractions of soil interact with metal ions, providing negatively charged surfaces on to which cations can bind. Such adsorption displaces the cation that was previously counter-balancing the colloid's negative charge; hence the ability of colloids to adsorb ions in this way is known as the soil's cation exchange capacity (CEC). Heavy metals may also specifically adsorb on to hydrated oxides of aluminium, iron, and manganese by the formation of partly covalent bonds and can be chelated by solid-phase humic substances or by other organic ligands. The adsorption of metal ions is also affected by the soil pH. Acidification decreases the CEC of humic substances when carboxyl and ammonium groups will become protonated, reducing their capacity to bind metal ions. The increase in hydrogen ions will also increase the competition for negative binding sites, leading to the displacement of metal ions. Therefore, the size distribution of mineral particles, the amount of organic matter and the soil pH will to a large degree determine the metal buffering capacity of the soil (Gupta, 1991).

Besides physical/chemical interactions, chemicals interact with the biota present in the soil. For example, sustained use of specific pesticides leads to the development of microbial communities that degrade the pesticide, reducing their half-life time and therefore their environmental impact. On the other hand, chemicals such as PCBs and organochlorides are likely to accumulate up the food chain resulting in high concentrations of residues distant from the point of release, thus increasing their environmental impact. As biological systems are not just collections of independent organisms but are strongly dependent on each other for their survival, it follows that chemical effects on one species can have important consequences for species that depend on the affected species. An example of this type of indirect effect is the decline of the gray partridge (*Perdix perdix*) in West Sussex, England. It could be shown that 48% of chick mortality was explained by the density of preferred insects. Given that reduced availability of insects is the key to the partridge decline, it was postulated that this might be due to increased pesticide usage, especially since there was a strong correlation between the partridge decline and increased herbicide use. The effect of the herbicides was shown to be indirect (they are relatively nontoxic to partridges), removing the host plants on which the insect larvae fed that were the food source of the partridge chicks, resulting in less insects and therefore less food for the partridge chicks. In conclusion, because biological effects are dependent on the behavior (movement, persistence, bioavailability) of the chemical in the environment, the sensitivity of biota and the role these biota play in ecosystem function, biomonitoring is often the only available way for monitoring chemical effects.

Early Warning Systems

Life as a whole can be seen as an hierarchical systems with distinct levels of biological integration. Going from a low level of organization to a high level of organization we can recognize cells that make up tissues, that make up organs, that are part of an organism; these organisms are arranged in populations that form part of communities, that are ultimately part of ecosystems. In general, it can be said that biological systems are at each level of organization buffered to resist change in overall performance by adjusting the components that make up that level of integration. The result of these buffering effects is that effects at lower levels of integration are dampened so that they do not affect higher levels of integration. Only when the buffering capacity at a certain level of integration is exceeded will the next level of integration be affected. This implies that toxic effects can be measured at the lower levels of biological integration before they result in effects at higher levels of integration. Because the aim of biomonitoring is "to provide a means by which (impending) changes in ecosystem function can be detected," compensatory effects at lower levels of integration can provide an early warning system of adverse chemical effects. A good example of such an early warning system is the thinning of bird eggshells as a result of DDE (one of the metabolites of DDT). DDE affects Ca deposition in eggshells (physiological change). However, only when egg shell thinning is in excess of 16 to 18% will this lead to a decrease in breeding success of the birds affected. Eggshell thinning is therefore an early warning system for population

change. Not all bird populations are likely to be affected as DDE accumulates up the food chain. Only birds that are at the top of the food chain (raptors and fish-eating birds) will be affected in the first instance. However it can be assumed that if DDT use had continued, the decline in raptors and fish-eating birds would eventually have resulted in community changes.

What is important in this context is to recognize which factors determine the buffering capacity at each level of biological integration. These factors can be used profitably in biomonitoring. It is, however, important to recognize that if we are interested in impending changes at the community level, it is probably most sensible to look for changes that are occurring at the level below (i.e. the population level). Similarly, impending population effects are most profitably investigated using effects on individuals. However, changes at the organism level will be too sensitive for making reliable predictions at community level and above. This simple fact is often overlooked in biomonitoring where current molecular technology has allowed precise investigation at the lowest levels of integration. These techniques allow the development of sensing systems that pick up changes that occur at the molecular and physiological level. Whereas these approaches are valuable for medical purposes in which we are interested in warning systems that help prevent problems at the individual level, they are inappropriate for monitoring ecosystem changes, in that they are far too sensitive. In choosing the most appropriate monitoring tool, one has to evaluate its relevance, reliability, robustness, responsiveness, and reproducibility. These factors, known as the 5 Rs, will differ according to level of biological integration and the composition of biological units that make up that level. It is unlikely that an environmental monitoring system will be "a freeze dried, talking bug on a stick" (van Straalen, 1998). As interactions that govern processes at each level of biological integration become apparent, new and more appropriate biomonitoring tools will become available. The simple fact that only a fraction of the species that make up "biological life" are classified shows that we have still a long way to go.

Specificity and Resolution

Within each level of biological integration the biological units that make up the level of integration display different degrees of specificity and sensitivity to different forms of (chemical) stress. With the term *specificity* it is meant that a biological effect can be related to a specific stimulus. For example, inhibition of the enzyme ALA dehydrogenase is only induced by lead, while induction of an immune response might be triggered by a wide variety of causes. The term *resolution* is used to indicate the ability of a bioindicator to respond to small environmental changes. For example, a specialized predator at the top of a food chain is a more sensitive indicator of bioaccumulation of persistent pesticides than omnivores with a varied diet. Although it seems that sensitive indicators with a great resolving power hold the greatest promise for biomonitoring, these properties are likely to result in overcaution and false alarms. In general, it can be said that indicators at lower levels of biological integration are more sensitive than characteristics that relate to higher levels of biological integration. Therefore, indicator species are more sensitive than diversity indices, for example. What makes a "good" bioindicator will depend on the information that is required. Looking at a "soil ecosystem" for example, the biological components that make up the system consist of plants, vertebrates, invertebrates (insects, earthworms, nematodes, mites and springtails), protozoa, bacteria, and fungi, all of which perform crucial roles in the maintenance of soil ecosystem function. When the system is exposed to a potentially toxic chemical the ability of each population to cope with this exposure will depend on genotypic plasticity within the population, phenotypic plasticity, mutation rate, metabolic capacity, and level of exposure. Genotypic plasticity will be high in microorganisms that multiply fast and can acquire resistance genes via mutations or by acquisition of plasmids carrying resistance genes. As a result, adaptation of microbial communities to chemical stress is likely to be high. Furthermore, because of the almost limitless metabolic diversity displayed by microorganisms, populations that contribute to the detoxification of a toxic chemical might increase rapidly (comparable with the activation of the immune system as a response to foreign molecules/organisms that enter a higher organism). Other organisms (plants, vertebrates, and invertebrates) that have a large phenotypic plasticity might respond to a chemical insult by avoiding those areas that are most contaminated. Furthermore, spatial distribution of the biota in the soil environment and their mobility will both determine exposure and therefore toxicity. Only when there is a clear understanding of "normal" distribution and diversity of the biota in a given ecosystem is it possible to relate biological patterns to the effects of a toxicant in a meaningful way.

Because of the biological complexity of ecosystems, efforts have been made to classify organisms not according to species but according to similarities in life strategy. Biological life can be roughly classified by the way organisms spend resources on their offspring. Organisms that invest in large numbers of offspring (giving each limited resources) are termed r-strategists, while those species that invest in few offspring (each of which is provided with ample resources) are termed K-strategists. Looking at the spectrum of life, there is a general trend that large organisms are K-strategists, while small ones are r-strategists. Although this is true to a large extent, within a certain body class size there are organisms that have adopted r- or K-type survival strategies. For example, mice are typical r-strategists (they have many large litters each year), while bats produce every two years a single offspring, making them typical K-strategists. Not surprising, K-strategists are characteristic for stable, climax environments while r-strategists are characteristic for unstable pioneer environments. In terms of "stress," K-strategists are adapted to cope with biological stress (competition, predation, etc.), while r-strategists are adapted to cope with physical/chemical stress but are not good at coping with biological stress. In general, it can therefore be stated that r-strategists are insensitive to chemical stress, while K-strategists are sensitive. When an ecosystem is challenged by toxins (pesticides, for example), the r-strategists (the pest organisms) are the first to develop pesticide resistance, while the K-strategists (nonpest species) will disappear. The reason for this might be found in the fact that large numbers of offspring provide large genetic variability leading to rapid selection of

the best-adapted individuals. Chemical/physical stress on an ecosystem will therefore lead to a shift in the balance between r- and K-strategists in favor of r-strategists. This principle has been used successfully to monitor environmental pollution and disturbance using colonizer/persister (c-p) indices applied to nematodes (Bongers, 1990) or perturbations of rhizophere microbial communities (De Leij *et al.*, 1993).

Another useful way of looking at communities and ecosystems is by classifying the organisms that inhabit that ecosystem by their "trophic function." Using this approach, organisms can be classified (aboveground) as primary producers (plants and algae), herbivores, predators, and top predators, while organisms that feed on a variety of food sources are classified as omnivores. In general, food chains have no more than four trophic levels because less and less energy is available higher up the food chain. Furthermore, dependence of predators and top predators on exclusive feeding sources make them vulnerable to perturbations lower down the food chain. Using the concept of trophic levels, it can be argued that specialized feeders are more sensitive to environmental stress (when this stress affects their food source) than organisms that are less specialized. Not surprisingly, when recalcitrant chemicals enter the food chain it is the specialized feeders (for example, peregrines and sparrow hawks that feed exclusively on birds) that are the first to be affected by toxins. Beside the fact that they are exposed to relatively high levels of a pollutant, these organisms are often very skillful specialized hunters. Any impairment of these skills due to a toxin will result in starvation. Therefore, these species can be sensitive bioindicators of toxins, even at sublethal concentrations.

Another important ecological concept that can be used to simplify ecosystems in a meaningful way is the concept of keystone species. The idea is that even though there is dependency between all species that are involved in a certain food web, some species are more important than others. It is important in this context to make a distinction between dominant species that derive their importance to the ecosystem in terms of biomass or energy flow and keystone species. Keystone species are often relatively rare but have a major impact because of their key regulatory function in the system as a whole. This is because these species are involved in more links or interact with parts of the ecosystem that maintain living conditions for a wide variety of other species. In other words, the species diversity of an ecosystem is determined in large part by a few keystone species on which many others depend. A detrimental effect on these species will result in the collapse of the whole ecosystem (or at least a large part thereof). Several different categories of keystone species have been recognized, such as keystone predators, keystone prey, and keystone habitat modifiers. One of the best-studied examples of the importance of keystone predators is the starfish (*Pisaster*) and predatory whelks (*Nucella*) in pools along the rocky inter tidal zone. Removal of these predators led to the disappearance of 80% of the prey species and the nearly complete dominance of the prey community by mussels (*Mytilus*). Sea otters (*Enhydra lutis)* have also been labeled as keystone species as they limit the density of sea urchins, which, in turn, eat kelp and other macro algae that provide the habitat for a large variety of species. Without sea otters the sea urchin population explodes resulting in the destruction of the kelp forests leading to the loss of habitat for those species that depend for their survival on the kelp forests for reproduction, shelter, and food. The way keystone species are affected by toxins is an important consideration in ecotoxicology, because they represent the Achilles' heel of an ecosystem. Not surprisingly, keystone species such as earthworms and honeybees are often used in toxicity testing because they are important for ecosystem function in general.

Clearly a negative effect on ecosystem function is often due to alterations of complex interactions within the ecosystem itself. In most cases, toxic effects are buffered due to the compensating capacity of an ecosystem. However, occasionally toxic effects can affect species that are responsible for essential functions in that ecosystem. Further insight in the way that different species contribute to ecosystem function is an essential requirement for the development of the most appropriate biomonitoring tools for assessing the ecotoxicological effects of chemicals.

Conclusion

Currently there is growing concern internationally about environmental toxicants affecting biodiversity and human health. Biodiversity has come to the fore by the United Nations meetings, the first of which was in Rio de Janeiro in 1992. The medical profession is increasingly concerned about a range of conditions that might be induced by chemicals in the environment. This is resulting in moves to increase legislation. The study of ecotoxicology seems therefore certain to grow in importance, as well as there being a stimulus to prevent problems occurring in the first place via the development of clean technology and stimulation of sustainable land use. On the other hand, ecotoxicology can help to identify existing pollution problems, which in many cases can be cleaned up using a variety of remediation strategies.

See also: Air Pollution. Ecological Footprint, Concept of. Environmental Impact, Concept and Measurement of. Greenhouse Effect. Keystone Species. Soil Biota, Soil Systems, and Processes

References

Alloway B (1995) *Heavy Metals in Soil*. Glasgow: Blackie Academic and Professional.

Bongers T (1990) The maturity index: an ecological measure of environmental disturbance based on nematode species composition. *Oecologia* 83: 14–19.

Carson R (1965) *Silent Spring*. Harmondsworth, England: Penguin Books.

De Leij FAAM, Whipps JM, and Lynch JM (1993) The use of colony development for the characterisation of bacterial communities in soil and on roots. *Microbiol. Ecology* 27: 81–97.

Everts JW, Aukema B, Hengeveld R, and Koeman JH (1989) Side effects of pesticides on ground dwelling predatory arthropods in arable ecosystems. *Environmental Pollution* 59: 203–225.

Gupta SK (1991) Metabolisable metal in antropogenic contaminated soils and its ecological significance. In: Vernet JP (ed.) *Impact of Heavy Metals in the Environment*, pp. 299–310. Amsterdam: Elsevier.

Jagers op Akkerhuis GAJM and Van der Voet H (1992) A dose-effect relationship for the effect of deltamethrin on a linyphiid spider population in winter wheat. *Archives of Environmental Contamination Toxicology* 22: 114–121.

Lovelock JE (1979) *Gaia: A New Look at Life on Earth*. Oxford: Oxford University Press.

Lynch JM and Wiseman A (eds.) (1998) *Environmental Biomonitoring: The Biotechnology Ecotoxicology Interface*. Cambridge: Cambridge University Press.

McCarthy JF and Shugart LR (1990) *Biomarkers of Environmental Contamination*. Boca Roton: Lewis Publications.

Nelson DR (1998) Metazoan cytochrome P450. *Comparative Biochemistry and Physiology* (121 Part C) 15–22.

Potts GR (1986) *The Partridge*. London: Collins.

Renberg I and Hellberg T (1982) The pH history of lakes in Southwestern Sweden, as calculated from the subfossil diatom flora of sediments. *Ambio* 11: 30–33.

Van Straalen NM (1998) Evaluation of bioindicator systems derived from soil arthropod communities. *Applied Soil Ecology* 9: 429–437.

Walker CH, Hopkin SP, Silby RM, and Peakall DB (1996) *Principles of Ecotoxicology*. London: Taylor and Francis.

Zelikoff JT, Schepers J, and Lynch JM (eds.) (1997) *Ecotoxicology: Responses, Biomarkers and Risk Assessment*. Fair Haven, NJ: SOS Publications.

Edible Plants

Eduardo H Rapoport and Barbara S Drausal, Universidad Nacional del Comahue, Bariloche, Argentina

© 2013 Elsevier Inc. All rights reserved.

Glossary

Cultivar Cultivated variety or genetic strain of a domesticated food plant.

Domesticate Plant that has been selected by humans and adapted for use as a food crop, nutrient, fiber, or for other purposes.

Ethnobotany Study of the variety, natural history, and characteristics of the plants used by human cultures.

Food Habits

Animals, whether terrestrial or marine, have limitations with respect to the variety of species they consume. They depend on what nature offers in the place where each animal lives. Since the geographic ranges of species are relatively widespread, the individuals of a given animal species may vary their food resources in different sites of the range, according to their opportunities. For this reason, the variety of plants ingested by a given species is always greater than the variety of plants ingested by an individual or by an entire population.

Domestic as well as wild herbivores may eat a wide variety of food items, but they have clear preferences for particular plant species. Some ungulates are capable of selecting the most nutritious individual plants among those of the same species. They select forage mainly by smell, and secondarily by taste. Volatile substances in the plants may either inhibit or attract foragers, and are largely responsible for their palatability. According to Klein (1970), the nutritive value of plants available to wild ruminants depends on the stage of maturity of the vegetation, with highest nutritive quality coinciding with the initiation of growth, as well as with soil type and climate. Rapid growth in plants is correlated with high nutritive quality. Similar conclusions were reported by Gardarsson and Moss (1970) in a study of food selection by the Icelandic ptarmigan (*Lagopus mutus*). This bird consumed leaves and flowers of 8, 11, 10, and 11 species in summer, autumn, winter, and spring, respectively.

Similarly, the European hare (*Lepus europaeus*) in northwestern Patagonia makes use of 17–21 species of plants in a given season out of a total of 28 species that it consumes over the year. **Table 1** provides a short sample of the variety of food consumed by different herbivorous and omnivorous mammals. The case of the Soay sheep was included to show how a domestic herbivore may restrict its diet in a species-poor environment. Of course, the range of foods ingested by other animals may vary widely. Monophagous insects restrict themselves to only one plant species, soil amoebae (*Acanthamoeba*) normally ingest five species of microscopic algae (Heal and Felton, 1970), and polyphagous, pest arthropods may feed on more than 300 species of crops and wild plants. The number of food plants eaten by humans is not far from the figures shown in **Table 1**.

Domestic and wild ungulates may show a copious range of food species but they focus on a few, preferred plants. A number of less palatable species are used only in time of food scarcity. Studies performed in western Argentina by Kufner and Monge showed that the rodent *Lagostotmus maximus* increases the variety of its food sources in degraded habitats and during droughts.

Standards of Consumption

On an individual basis, people use a small number of plant items per day, perhaps between 10 and 20 species or products.

Table 1 The number of food plants consumed annually by different mammal species

Mammal species	*Common name*	*Number of plant species*	*Source*
Alouatta fusca	Brown howler monkey, SE Brazil	52	Galetti *et al.* (1994)
Ateles spp.	Spider monkey, Panama	14	Milton (1981)
Bos Taurus	Cattle, NW Patagonia	23	Relva (1998)
Copra hircus	Goat, Mendoza, Argentina	76	Dalmasso *et al.* (1995)
Cebus paella	Capuchin monkey, SE Brazil	73	Galetti and Pedroni (1994)
Cervus elaphus	European red deer, NW Patagonia	34	Relva (1998)
Chiropotes satanas	Monkey, Venezuela	29	Kinsey and Norconk (1993)
Ctenomys mendocinus	Tuco-tuco, Argentina	28	Madoeri (1993)
Lagidium viscacia	Vizcacha de la sierra, Argentina	21	Galende and Grigera (1998)
Lama guanicoe	Guanaco, Mendoza, Argentina	47	Candia and Dalmasso (1995)
Lepus europaeus	European hare, NW Patagonia	28	Galende and Grigera (1998)
Ovis aries	Soay sheep, St. Kilka Island, UK	~12	Gwynne and Boyd (1970)
Pithecia pithecia	Monkey, Venezuela	25	Kinsey and Norconk (1993)

A normal diet includes common vegetables, fruits, seeds (in the form of flour or oil), roots, sugar, beverages (beer, wine, colas), condiments, teas, and herbals. The number of plant species normally used by an individual over a year, however, is about 100, although this is limited by the number of edible plant items commonly offered by popular markets and supermarkets. Taking into account all the variety of greens, vegetables, fruits, grains, nuts, and condiments in an exceptionally well-provisioned supermarket, the figure (including different varieties and brands) may rise to 600, according to Duke (1992).

The stomach contents of two mummies that were found well-preserved in Danish bogs provided interesting information about the gastronomic habits of people during the Iron Age. Their last meals contained 66 different plant taxa (Godwin, 1960; King, 1966), many of which are nowadays considered as cosmopolitan weeds. Before the invention of agriculture, in the Paleolithic Age, humans were hunter-gatherers, and probably had a better knowledge of the variety of edible wild plants than modern people. This knowledge, however, has slowly been lost since the Neolithic and, in present times, is still lost after one or two generations of acculturation in aboriginal communities (Plotkin, 1993). The process of civilization goes hand in hand with the loss of knowledge, as well as with the abandonment of traditional crop varieties and the habit of gathering wild plants. But at the same time, new cultivars, coming from distant countries, are constantly increasing the variety of foods. A recent case is that of the kiwi fruit (*Actinidia deliciosa*), which originated in China and then was renamed, cultivated, and popularized in New Zealand. The roots of the ahipa (*Pachyrhizus ahipa*) and jicama (*P. erosus*), of Central American origin, have become increasingly popular in the US and Southeast Asia, especially the latter species, which has a similar texture and flavor to bamboo shoots and is used by Asian food restaurants in Western countries (National Research Council, 1989; Brucher, 1989). Rice (*Oryza sativa*), of Asian origin, is at present the most popular staple in warm countries of South America, whereas the South American potato (*Solanum tuberosum*) has become mandatory in European cuisine.

The Diversity of Food Plants

No one has compiled a complete record of edible plants for the entire world. The Food and Agriculture Organization (FAO), part of the United Nations, publishes an annual report of the production of the commercially most important foods. This list includes about a hundred species of plants. In *The Oxford Book of Food Plants* (Nicholson *et al.*, 1969), the number increases to 389 species distributed among 81 plant families. These are both locally and widely known cultivated plants. Duke (1992) estimated that North American Indians ate 1112 plant species. This figure is set at 1886 species according to Moerman (1998). More than 3000 edible species are carefully listed and commented on in the voluminous book *Cornucopia*, compiled by Facciola (1990), but in its preface Noel D. Vietmeyer suggests that there are about 20,000 edible species across the world. Probably, the most complete inventory is Kunkel's (1984) book, which lists approximately 12,560 species from 3100 genera belonging to about 400 families of flowering plants and ferns. This list, however, is being constantly enriched by the contribution of many ethnobotanical studies.

The proportion of edible plant species in slightly disturbed communities is variable. In the Sonoran Desert it is about 15%. Ona Indians from Tierra del Fuego made use of at least 6% of this island flora, whereas the Chácobo Indians in the Bolivian Amazon use 21% of their surrounding flora. Medium to highly disturbed communities may contain similar or higher proportions of edible species for human consumption. For example, in western Uruguay the proportion is 17%, in southwestern Córdoba province (Argentina) it is 19%, in the outskirts of Havana (Cuba) it is 33%, in swidden (slash-and-burn) fields of northern India it is 43%, and in experimental fields in Saskatchewan (Canada) it may reach 61% of all wild plants. Yet studies like these do not necessarily reveal the actual possibilities offered by nature, but rather the knowledge of informants or the perspicacity, experience, and fieldwork time employed by the investigators. According to the estimate that about 10% of any flora represents food resources, then 10%, or 27,000, of the 270,000 species of plants already recognized by world botanists should be edible.

Since historical times, because of written testimony, Europe has conserved people's knowledge of gastronomic matters. From the botanical point of view, the UK is probably the best-known country in the world. If we compare the floristic list compiled by Martin in 1976 with Kunkel's list of food plants, and discard the exotic species, hybrids, and other sub-specific taxa, as well as plants used only during famine times, we can verify that out of 1503 species considered, 350 are edible. In other words, 23% of the British flora is edible. Thus, we have two estimates of the possible richness of edible vascular plants – 10% and 23% – and they represent between 27,000 and 62,000 species, respectively, based on the 270,000 known at present. Because the description of the world flora has not been completed yet, the final list of comestibles will probably increase in the future.

By comparison, less than 2% of the Central American flora is eatable, based on the list prepared by Duke (1992). Possible explanations for this remarkable difference are: (1) greater taxonomic ignorance or less exploration of the natural resources – because the flora of Central America is much richer than the flora of Britain, humans may have concentrated on fewer, more abundant, and profitable plants, and disregarded the less useful ones; and (2) widespread loss of cultural heritage and environmental knowledge following the conquest and colonization by European countries.

The Most Prolific Taxa

A first, rough estimation at higher taxonomic ranks indicates that the proportions of edible species are quite similar to the proportions of common (edible and nonedible species) species present in the plant kingdom. The right-hand column in **Table 2** is based on a random sample of 1790 food plants appearing in Kunkel (1984).

If the property of being eatable or palatable were randomly distributed among the different taxa, then it would be

Table 2 A comparison between common and edible Species in higher taxonomic groups (figures represent percentages of their respective totals)

Taxonomic group	*Common species (n = 270,000)*	*Edible species (n = 1790)*
Pterydophyta	3.9	1.9
Gymnospermae	0.3	1.0
Dicotiledoneae	69.9	75.5
Monocotiledoneae	25.9	21.6

predictable that the most numerous families of plants would contain a higher number of edible species (**Table 3**). This relation seems to be valid since many of the most prolific families are also among the most productive in edible plants. This is the case with the Compositae (Asteraceae), Leguminosae (Fabaceae), Gramineae (Poaceae), Euphorbiaceae, and Rubiaceae, which occupy the top positions in the ranking of both common species and edible species. The most numerous of all plant families, the Orchidaceae, however, has few edibles. Among the most suggestive cases is the number of genera in the Cactaceae family, which appears in the fourth place for edibles but is the 31st in the rank of common plants.

There is no clear relation among the ratios of number of species per genus that allows us to differentiate edible from common species. Among families with the lowest ratios are the Asclepiadaceae (8.0), Cruciferae (8.6), and Rutaceae (10.0). Families that are richer in species per genus are the Begoniaceae (255.0), Aizoaceae (208.3), and Eriocaulaceae (92.3), yet they show no evidence of having experienced a process of selection, that is, of proclivity or rejection by humans, that would have led them to speciate toward palatability or distastefulness.

Of the 389 more frequently cultivated species considered by Nicholson *et al.* (1969), the ranking goes as follows: Rosaceae (13.3% of total species), Leguminosae (8.5%), Gramineae (6.4%), Compositae (5.9%), Umbelliferae (5.4%), and Cruciferae (5.1%). They are followed by Palmae, Cucurbitaceae, Rutaceae, Alliaceae, Chenopodiaceae, Dioscoreaceae, Annonaceae, Ericaceae, Grossulariaceae, and 66 less prolific families. Kunkel (1984) states that the Rosaceae is the richest family among food plants. The analysis of a sample of 6222 items from his list confirms this assessment. Rosaceae appears at the top of the list, comprising 5.8% of the cases, insofar as Leguminosae are split into Fabaceae (fourth place), Mimosaceae (sixth place), and Caesalpinaceae (82nd place). On the contrary, if the latter three families are considered as a unit, the Leguminosae stand in first place, comprising 6.7% of the sample. The rank of the first 30 families is shown in **Table 4**.

Table 3 A ranking of the abundance of genera per plant family

Common species	*Edible species*
Compositae (Asteraceae)	Compositae
Orchidaceae	Leguminosae
Leguminosae (Fabaceae, Mimosaceae)	Gramineae
Gramineae (Poaceae, Bambusaceae)	Cactaceae
Rubiaceae	Umbelliferae
Cruciferae (Brassicaceae)	Palmae
Umbelliferae (Apiaceae)	Rubiaceae
Euphorbiaceae	Labiatae
Liliaceae	Euphorbiaceae
Asclepiadaceae	Cruciferae
Acanthaceae	Rosaceae
Labiatae (Lamiaceae)	Myrtaceae
Palmae (Arecaceae)	Araceae
Scrophulariaceae	Apocynaceae
Rutaceae	Moraceae

Table 4 A ranking of the abundance of edible species in the 30 most prolific families[a]

1. Rosaceae
2. Compositae (Asteraceae)
3. Dioscoreaceae
4. Fabaceae
5. Liliaceae
6. Mimosaceae
7. Moraceae
8. Ebenaceae
9. Rubiaceae
10. Myrtaceae
11. Solanaceae
12. Cactaceae
13. Gramineae (Poaceae)
14. Fagaceae
15. Euphorbiaceae
16. Cruciferae (Brassicaceae)
17. Polygonaceae
18. Palmae (Arecaceae)
19. Ericaceae
20. Rutaceae
21. Umbelliferae (Apiaceae)
22. Sapotaceae
23. Guttiferae (Clusiaceae)
24. Asclepiadaceae
25. Caesalpiniaceae
26. Annonaceae
27. Zingiberaceae
28. Vitaceae
29. Araceae
30. Cucurbitaceae

[a]The richest family of common edible and nonedible species, the Orchidaceae, appears in 45th place.

Edible Parts

Some plant genera are extremely abundant in edible species and may show particular tendencies toward a given kind of food (**Table 5**). For example, all of the 205 species of *Rubus* appearing in Kunkel's list provide edible fruits. Among them, there are three species whose leaves are also used as tea. Similarly, the 80 or more species of *Prunus* provide edible fruits, as also occurs with *Rosa* spp. and *Ribes* spp. The majority of *Piper* species are used as black or white pepper or as a spice for curries. *Rumex* provides 44 species with leaves used as vegetables and three species with edible roots. Of the 100 edible species of *Solanum*, 59 are used only for their fruits, 20 species only for their tubers, 14 for both fruits and leaves, six species only for their leaves, and a single species exclusively for its seeds.

Table 5 The most prolific genera of food plants

Genus	*Family*	*Number of food species*	*Total number of species*	*% food species*
1. *Ruibus*	Rosaceae	205	2500–3000	7–8
2. *Ficus*	Moraceae	137	700	20
3. *Dioscorea*	Dioscoreaceae	110	600	18
4. *Solarium*	Solanaceae	100	1400–1700	6–7
5. *Acacia*	Leguminosae	80	800	10
6. *Eugenia*	Myrtaceae	79	800	10
7. *Diospyros*	Ebenaceae	69	200	35
8. *Garcinia*	Guttiferae	68	400	17
9. *Quercus*	Fagaceae	67	470–1000	7–14
10. *Vaccinium*	Ericaceae	66	300–400	17–22
11. *Passiflora*	Passifloraceae	58	500	12
12. *Opuntia*	Cactaceae	52	250	21

Table 6 An estimate of food usage (as percentages of regional totals) from different sources

	World flora (Facciola, 1990)	*World flora (Kunkel, 1984)*[a]	*Panama Isthmus (Duke, 1990)*	*United Kingdom (Martin, 1976)*	*Andes (NRC, 1989)*	*Cambodia (Ito, 1969)*	*Botswana (Campbell, 1986)*
Leaves[b]	21.8	28.9	20.6	46.8	7.6	22.8	16.2
Fruits	19.6	30.5	37.4	9.5	61.4	29.9	33.3
Seeds	13.9	13.2	17.5	10.1	10.6	17.0	8.1
Condiments, flavorings	11.5	6.2	3.2	6.9	6.1	6.5	0.0
Tea, herbals	9.5	2.9	1.6	7.6	0.0	3.1	4.0
Beverages	7.0	0.8	6.0	0.0	2.3	1.7	4.0
Flowers, capers	6.7	4.0	4.1	4.8	0.0	9.2	1.0
Roots[c]	5.6	8.7	6.3	12.6	12.1	7.5	26.3
Saps[d]	4.2	3.9	2.2	0.6	0.0	2.0	7.1
Barks	0.2	0.9	1.0	1.1	0.0	0.3	0.0
Number of species	3000	1790	182	350	98	186	77
Number of items (multiple uses)	6311	2442	336	476	132	294	99

[a]A more detailed analysis of food usages appears in **Table 8**.
[b]Including stems, sprouts, and meristems.
[c]Including bulbs and rhizomes.
[d]Including gums, latex, sugars, and masticatories.

An analysis of the data compiled by Duke (1972) for Central America indicates that 85% of the plant species are used for a single purpose, 10% have two uses, 4% have three uses, and 1% have four uses (leaves, flowers, fruits, and seeds). A similar analysis of the 350 edible species of native British flora yields slightly different proportions: 78%, 19%, $<3\%$, and $<1\%$ for one, two, three, and four purposes, respectively. One of the exceptional cases of variability of uses is that of hops (*Humulus lupulus*). Its leaves, roots, flowers, and bark may be used as food and as condiment. The pumpkin (*Cucurbita moschata*) provides fruits, seeds, flowers, young leaves, and shoots for human consumption.

It is clear from **Table 6** that some kinds of food are scarcely used, such as bark, flowers, sap, and liquorice. In contrast, leaves (including stems, sprouts, seedlings, and shoots) and fruits seem to be the most preferred food.

Geographic Patterns of Food Plants

On the basis of a sample of 1790 species from Kunkel's list of food plants, it appears that edible herbs are more numerous than trees, shrubs, and vines. Vines include all the climbing, creeping, and epiphytic plants. **Table 7** shows that in decreasing order of species richness, the Indomalayan regions appear first, followed by the Neotropical, Palearctic, Ethiopian, Nearctic, and Australian (or Australasian) regions. By means of a chi-square test, at a significance level $P<.05$, the Nearctic and Australian regions show greater, and the Palearctic lower, proportions of food trees than expected. The proportion of edible shrubs does not vary significantly among regions, although at a $P<.1$ the Neotropics seem to have a higher proportion than the rest of the world. The Palearctic region shows a higher and the Neotropics a lower proportion of herbs, whereas the latter region is richer in vines ($P<.05$). In contrast, the Northern Hemisphere (Palearctic and Nearctic regions) shows a significantly lower proportion of vines than the Southern Hemisphere.

According to biogeographic regions (**Table 8**), the Palearctic shows a significantly greater proportion of species that provide edible leaves, stems, and sprouts than the other regions. The Australian region is characterized by a greater

Table 7 The distribution of food plants by growth form in different geographic regions[a]

Growth form	*Total number of spp.*	*Percentages (from totals)*	*Nearctic (North America)*	*Neotropical (C. & South America)*	*Indomalayan*	*Australian*	*Ethiopian (Africa S. of the Sahara)*	*Palearctic (Eurasia and North Africa)*
Trees	504	25.9	14.3	24.9	33.9	45.1	35.4	9.6
Shrubs	452	23.2	31.1	31.6	20.0	17.7	16.5	20.0
Herbs	787	40.4	51.2	23.7	33.9	32.4	38.8	66.5
Vines	203	10.4	3.4	19.8	12.1	4.9	9.3	3.9
Totals	1946	~100.0	12.2	22.3	25.4	5.2	16.6	18.2

[a]The sample includes 1790 species. Because a number of species are shared between two or more regions, the total number of items classified increases to 1946. The third column, as well as the bottom row, are percentages of the total number of species considered. The remaining columns are percentages calculated from their respective regional subtotals.

Table 8 A sample of 1790 species used in 2442 ways, classified by types of food in different regions (figures are percentages of regional totals)

	Nearctic	*Neotropical*	*Indomalayan*	*Australian*	*Ethiopian*	*Palearctic*
Leaves[a]	24.3	14.3	30.0	25.0	27.3	48.0
Fruits	28.3	52.5	31.9	14.1	27.3	14.4
Seeds	12.8	11.7	13.7	28.9	13.3	10.0
Condiments, flavorings	3.0	3.9	8.2	2.3	6.4	8.7
Tea, herbals	4.9	2.3	2.2	4.7	2.2	3.6
Beverages	1.0	2.8	0.3	0.0	0.2	0.0
Flowers, capers	3.3	2.7	4.8	1.6	5.4	4.2
Roots[b]	15.8	7.6	5.8	8.6	9.1	9.1
Saps[c]	4.6	2.0	2.0	13.3	8.1	1.7
Barks	2.0	0.4	1.1	1.6	0.7	0.2
Number of items	304	487	643	128	407	471

[a]Including stems, sprouts or shoots, and meristems.
[b]Including bulbs and rhizomes.
[c]Including gums, latex, sugars, and masticatories.

proportion of edible seeds, and the Neotropics by its abundance of fruit species. Of course, these differences were derived from the sample analyzed by us. For instance, the appearance of zero values for beverages from the Australian and Palearctic regions does not mean that there are no species of this kind in their floras; they were simply not registered in our sample of 1790 species. On the contrary, the high proportion of edible fruits in the Neotropical region is repeatedly verified in ethnobotanical studies performed in different countries of Central and South America. These contrasts may be attributed to differential characteristics of seed dispersal evolution within their plant communities, as well as to prevailing cultural trends in the use of natural resources. The possibility of ethnobotanists showing biased attention toward particular kinds of food should not be discarded.

Are Edible Weeds Weeds?

The habit of gathering wild food plants has not been totally lost. Spencer maintained that "any plant is a weed if it insists upon growing where the husbandman wants another plant to grow. It is a plant out of place in the eye of man; in the nice eye of nature it is very much in place." Many of our dearest crops, however, have originated from weeds. In the course of time, humans learned how to profit from them. Oats (*Avena sativa*), foxtail millet (*Setaria italica* derived from *S. viridis*), chicory (*Cichorium intybus*), pak choi (*Brassica rapa*), spinach beet (*Beta vulgaris*), and many more species appeared originally as invaders in ancient cultivated fields. Furthermore, several cosmopolitan weeds, such as dandelion (*Taraxacum officinale*), garden rocket (*Eruca vesicaria*), common purslane (*Portulaca oleracea*), and shepherd's purse (*Capsella bursapastoris*), are nowadays cultivated for the specialty food market.

After analyzing a set of 22,521 species of plants inhabiting natural ecosystems of North and South America, we concluded that 11.3% were edible. In another sample of 1264 species from seminatural communities in the temperate Neotropical region, edibles were 20.3% of the total. But if only weeds are considered (2455 widely spread species), the proportion mounts to 35.6%. This demonstrates that the degree of environmental disturbance correlates with the proportion of weeds, which in turn increases the proportion of edible food resources. This should not be surprising given that 17 of the world's 18 most aggressive weeds provide parts for human consumption. Because weeds are so numerous (more than 10,000 species cataloged) and so abundant (averaging 1.3 and 2.1 ton ha^{-1} in a temperate and a tropical area of northwestern Patagonia and eastern Mexico, respectively; Díaz-Betancourt *et al.*, 1999; Rapoport *et al.*, 1998), they stand

ambiguously as both enemies and potential benefactors of humanity.

The prospects for the future of food plant diversity appear to be auspicious. Further research should be conducted to manage and profit from these varied and abundant natural resources. At present, more than 15,000 species of food plants are recorded and this figure is constantly growing. Yet most of the world is fed with about 20 crops. As has been pointed out by Facciola (1990), 8000 cultivars of apples have been developed by humans, but only a handful are available in supermarkets.

See also: Agriculture, Traditional. Domestication of Crop Plants. Plant Biodiversity, Overview

References

Brucher H (1989) *Useful Plants of Neotropical Origin and Their Wild Relatives*. New York: Springer-Verlag.

Campbell A (1986) The use of wild food plants, and drought in Botswana. *Journal of Arid Environment* 11: 81–91.

Díaz-Betancourt M, Ghermandi L, Ladio A, López-Moreno IR, Raffaele E, and Rapoport EH (1999) Weeds as a source for human consumption. A comparison between tropical and temperate Latin America. *Revista de Biologia Tropical* 47: 329–338.

Duke JA (1972) *Isthmian Ethnobotanical Dictionary*. Maryland: Fulton.

Duke JA (1992) *Handbook of Edible Weeds*. Boca Raton, FL: CRC Press.

Facciola S (1990) *Cornucopia. A Sourcebook of Edible Plants*. Vista, CA: Kampong Publications. (A second, revised edition was published in 1998.)

Gardarsson A and Moss R (1970) Selection of food by Icelandic Ptarmingan in relation to its availability and nutritive value. In: Watson A (ed.) *Animal Population in Relation to Their Food Resources*, pp. 47–71. Oxford: Blackwell.

Godwin H (1960) The history of weed in Britain. In: Harper JL (ed.) *The Biology of Weeds*, pp. 1–10. Oxford: Blackwell.

Heal OW and Felton MJ (1970) Soil amoeba: Their food and their reaction to microflora exudates. In: Watson A (ed.) *Animal Population in Relation to Their Food Resources*, pp. 145–162. Oxford: Blackwell.

Ito S (1969) Economical plants and their distribution in Cambodia. *Journal of Agricultural Science Tokyo University of Agriculture* 13: 163–218.

King LW (1966) *Weeds of the World. Biology and Control*. London: Leonard Hill.

Klein DK (1970) Food selection by North American deer and their response to overutilization of preferred plant species. In: Watson A (ed.) *Animal Population in Relation to Their Food Resources*, pp. 25–46. Oxford: Blackwell.

Kunkel G (1984) *Plants for Human Consumption*. Koenigsten, Germany: Koeltz Scientific Books.

Martin WK (1976) *The Concise British Flora in Colour*. London: Ebury Press and Michael Joseph.

Moerman DE (1998) *Native American Ethnobotany*. Portland, Oregon: Timber Press.

National Research Council (1989) *Lost Crops of the Incas*. Washington, DC: National Academy Press.

Nicholson BE, Harrison SG, Masefield GB, and Wallis M (1969) *The Oxford Book of Food Plants*. London: Oxford University Press. (A second, updated edition appeared as Vaughan JG and Geissler CA. (1997) *The New Oxford Book of Food Plants: A Guide to the Fruit, Vegetables, Herbs, and Spices of the World*.)

Plotkin MJ (1993) *Tales of a Shaman's Apprentice*. New York: Penguin Books.

Rapoport EH, Ladio A, Raffaele E, Ghermandi L, and Sanz EH (1998) Malezas comestibles. *Ciencia Hoy* 9(49): 30–43.

Education and Biodiversity

Shirley M Malcom, American Association for the Advancement of Science (AAAS), Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Formal education Education that takes place in a school environment.

Informal education Education that takes place in nonschool settings, such as museums, zoos, parks, or through the media.

Out of school experiences Things that individuals experience that may support or reinforce learning. These might include participation in youth-serving organizations or hobbies such as bird watching.

Science literate Referring to a person who has fundamental knowledge and understanding of, as well as comfort with, the concepts, ideas, and nature of science.

Systematics Area of biology related to the classification of living things based on their relationships and evolutionary history.

Introduction

The Convention on Biological Diversity articulates a case for nations of the world to come together to undertake activities to improve conservation of biodiversity and sustainable use of biological resources. As of November 2010 more than 190 nations had ratified the convention and 170 had developed national biodiversity strategies and action plans. In addition to calls for better management, more research and study, protection of traditional knowledge, and international and regional cooperation, there was also recognition of the role of education, public participation, public information, and the development of a cadre of professionals to support the goals of the convention.

A 1998 report to President Clinton from the President's Committee of Advisors on Science and Technology (PCAST) makes recommendations to strengthen "the understanding and management of biological resources" (PCAST, 1998). Among the recommendations in *Teaming with Life: Investing in Science to Understand and Use America's Living Capital,* are calls for increased opportunities for formal and informal education centered on biodiversity and ecosystems, for interactions between scientists and students, and for continuing professional education for K–12 teachers.

Biodiversity: Connecting with the Tapestry of Life (Alonso *et al.*, 2001), prepared by the Smithsonian Institution's Program on Monitoring and Assessment of Biodiversity, under the Conservation and Research Center of the National Zoological Park and President Clinton's PCAST, sounded a similar theme as to the importance of biodiversity and the role of education and public awareness in its protection.

This article will outline aspects of formal and informal education focused on biodiversity. Specifically, it will outline learning goals around biodiversity for K–12 education, recommended school-based experiences that can lead to an attainment of these goals, and evidence as to the extent to which these understandings are being achieved by students. The article will review school-based environmental education initiatives and describe how these compare to and differ from education about biodiversity. The article will then continue with a discussion of informal learning opportunities in "places of science" as well as those available through youth-serving organizations, tourism, and field experiences. The section on general education will conclude with information on public interest in and awareness of biodiversity.

A brief section will discuss biodiversity and tertiary education: Biodiversity themes and courses as part of liberal education as well as issues in the education of professionals and specialists who work in biodiversity research, management, and conservation.

Learning About Biodiversity: K–12 Education

How do students learn about biodiversity? What specific concepts must they learn and what ideas must they acquire to support that understanding? What class work, materials, curriculum, set of courses, and experiences would provide an adequate background so that they come to an understanding of this concept?

In 1985 the American Association for the Advancement of Science (AAAS) began Project 2061, a long-term initiative to reform K–12 education in natural and social sciences, mathematics, and technology. *Science for All Americans*, (Project 2061, American Association for the Advancement of Science, 1989) was conceived as a statement of learning goals for science, mathematics, and technology education, defining what all students should know and be able to do by the time they conclude secondary education. *Science for All Americans* includes learning goals related to biodiversity (see **Box 1**). *Benchmarks for Science Literacy (Benchmarks (Project 2061, American Association for the Advancement of Science, 1993))* and the *National Science Education Standards (Standards)*, published by the National Research Council in 1996, describe the ideas that students must grasp at different ages (grades or developmental levels) in order to understand science concepts, including the overall concept of biodiversity.

Grades K–2

For young children in early primary education, *Benchmarks* recommends that students learn the following:

- Some animals and plants are alike in the way they look and in the things they do, and others are very different from one another.

Box 1 Biodiversity: What would a science literate adult understand?

Diversity of Life

Millions of different types of individual organisms inhabit the earth at any one time – some very similar to each other, some very different. Biologists classify organisms into a hierarchy of groups and subgroups on the basis of similarities and differences in their structure and behavior. One of the most general distinctions among organisms is between plants, which get their energy directly from sunlight, and animals, which consume the energy-rich foods initially synthesized by plants. But not all organisms are clearly one or the other. For example, there are single-celled organisms without organized nuclei (bacteria) that are classified as a distinct group.

Animals and plants have a great variety of body plans, with different overall structures and arrangements of internal parts to perform the basic operations of making or finding food, deriving energy and materials from it, synthesizing new materials, and reproducing. When scientists classify organisms, they consider details of anatomy to be more relevant than behavior or general appearance. For example, because of such features as milk-producing glands and brain structure, whales and bats are classified as being more nearly alike than are whales and fish or bats and birds. At different degrees of relatedness, dogs are classified with fish as having backbones, with cows as having hair, and with cats as being meat eaters.

For sexually reproducing organisms, a species comprises all organisms that can mate with one another to produce fertile offspring. The definition of species is not precise, however; at the boundaries it may be difficult to decide on the exact classification of a particular organism. Indeed, classification systems are not part of nature. Rather, they are frameworks created by biologists for describing the vast diversity of organisms, suggesting relationships among living things, and framing research questions.

The variety of the earth's life forms is apparent not only from the study of anatomical and behavioral similarities and differences among organisms but also from the study of similarities and differences among their molecules. The most complex molecules built up in living organisms are chains of smaller molecules. The various kinds of small molecules are much the same in all life forms, but the specific sequences of components that make up the very complex molecules are characteristic of a given species. For example, DNA molecules are long chains linking just four kinds of smaller molecules, whose precise sequence encodes genetic information. The closeness or remoteness of the relationship between organisms can be inferred from the extent to which their DNA sequences are similar. The relatedness of organisms inferred from similarity in their molecular structure closely matches the classification based on anatomical similarities.

The preservation of a diversity of species is important to human beings. We depend on two food webs to obtain the energy and materials necessary for life. One starts with microscopic ocean plants and seaweed and includes animals that feed on them and animals that feed on those animals. The other one begins with land plants and includes animals that feed on them, and so forth. The elaborate interdependencies among species stabilize these food webs. Minor disruptions in a particular location tend to lead to changes that eventually restore the system. But large disturbances of living populations or their environments may result in irreversible changes in the food webs. Maintaining diversity increases the likelihood that some varieties will have characteristics suitable to survival under changed conditions. As noted in *Science for All Americans*, Project 2061, from the American Association for the Advancement of Science: PCAST, 1998.

- Plants and animals have features that help them live in different environments.
- Stories sometimes give plants and animals attributes they really do not have.

Recommended learning activities include providing students with the opportunity to observe a variety of plants and animals in the classroom; on the school grounds; in the community; at home in parks, streams, and gardens; and at the zoo. The observations will prompt students to pursue questions about how the organisms live, where they are found, or how they interact with other organisms.

Grades 3–5

For children in later primary education, *Benchmarks* recommends that students learn the following:

- A great variety of kinds of living things can be sorted into groups in many ways using various features to decide which things belong to which group.
- Features used for grouping depend on the purpose of the grouping.

Recommended learning activities include providing students with the opportunity to learn about an increasing variety of living organisms and offering them a chance to invent schemes for classification. Students would be encouraged to develop different classification schemes (without being introduced to the Linnean classification system) and shown how their usefulness varies depending on the purpose of the classification. The purpose of the work would be to help students develop a deeper understanding about the relatedness of organisms.

Grades 6–8

For children in upper primary and lower secondary education *Benchmarks* argues that science should provide students with opportunities to enrich their growing knowledge of the diversity of life on the planet and to begin to connect that knowledge to what they are learning in geography. That is, whenever students study a particular region in the world, they should learn about the plants and animals found there and how they are like or unlike those found elsewhere.

Food patterns of development, and external and internal structures, would all be used to illustrate interrelationships, interdependence, similarities, and differences. Students would be introduced to the features that biologists use in classification systems and would be taught why these classifications are made.

Benchmarks suggests that students in this group should know the following:

- One of the most general distinctions among organisms is between plants, which use sunlight to make their own food, and animals, which consume energy-rich foods. Some kinds of organisms, many of them microscopic, cannot be neatly classified as either plants or animals.
- Animals and plants have a great variety of body plans and internal structures that contribute to their being able to make or find food and reproduce.

- Similarities among organisms are found in internal anatomical features, which can be used to infer the degree of relatedness among organisms. In classifying organisms, biologists consider details of internal and external structures to be more important than behavior or general appearance.
- For sexually reproducing organisms, a species comprises all organisms that can mate with one another to produce fertile offspring.
- All organisms, including the human species, are part of and depend on two main interconnected global food webs. One includes microscopic ocean plants, the animals that feed on them, and finally the animals that feed on those animals. The other web includes land plants, the animals that feed on them, and so forth. The cycles continue indefinitely because organisms decompose after death to return food material to the environment.

Grades 9–12

For students at the secondary level, curricular objectives lead to understanding diversity within and among species by looking at same and different features at a molecular level. Students would learn the following:

- The variation of organisms within a species increases the likelihood that at least some members of the species will survive under changed environmental conditions, and a great diversity of species increases the chance that at least some living things will survive in the face of large changes in the environment.
- The degrees of kinship between organisms or species can be estimated from the similarity of their DNA sequences, which often closely matches their classification based on anatomical similarities. Understanding built up over this period of study would lead students to comprehend the diversity of ecosystems, diversity of species, and the genetic diversity within species.

While *Benchmarks* sets out a recommended sequence of learning goals to help students come to an understanding of biodiversity as a complex idea, it is not clear that most students have access to the education, ideas, concepts, and learning experiences needed to achieve such understandings. Therefore it would be necessary to explore what students are taught or expected to learn over time during their schooling.

Biodiversity and School Science

Formal education in science is an important contributor to students' fundamental understandings about science. While self-directed study – books, articles, the Internet, museum visits, and field experiences – augment science learning for many students, the quality of the curriculum, textbooks, and other instructional materials, the preparation of teachers, the school-mediated experiences provided to students both inside and outside of the classroom all interact to shape what students take away from school science. Understandings of biodiversity would be based on accumulated experiences and knowledge. These would include the early school focus on natural history and nature study and development of an intuitive understanding of biological diversity and the relationships among living organisms (National Research Council, 1990). Students' out-of-school experiences, where such are available, would reinforce school learning. The curriculum focus shifts in lower and upper secondary levels (grades 6–12) to more formal, taxonomic instruction.

This pattern of topic coverage for life sciences concepts is present in the curriculum of other countries around the world and was prevalent among the majority of the 50 countries that participated in the Trends in International Mathematics and Science Study (TIMSS). Data were collected in 2007 and results presented in 2008. In addition, the 2011 framework for the upcoming assessment contains similar topical coverage of life sciences concepts; over 60 countries are slated to participate.

In the US, according to statistics provided by the U.S. Department of Education (2007), biology is the most frequently taken high school science course, with over 92% of 2005 graduates of public and private high schools reporting having completed such a class. Analysis of content core is not available to determine the extent to which the ideas critical to a student's understanding of biodiversity are actually taught. It is also not clear if earlier foundational ideas are provided to students as a part of instruction at primary and lower secondary levels. One trend that has emerged since the late 1990s is the increase in the percentage of students taking Advanced Placement/Honors biology courses *in lieu* of so-called general biology.

In the US there has been considerable discussion about the adequacy, pacing, and structure of curriculum and, especially, of textbooks in middle and high school biology. Researchers have criticized their encyclopedic nature, with too many ideas covered too superficially, and too much focus on vocabulary at the expense of big ideas. In addition, there has been a movement to avoid instruction of controversial topics such as evolution, a central idea in biodiversity. A 1999 study that evaluated the science textbooks used in middle grades against three concepts from *Benchmarks* and *Standards*, one each from the earth, physical, and life sciences, led the AAAS Project 2061 to conclude that all nine titles examined were inadequate to help students achieve understanding of the ideas that they were attempting to explain.

Efforts to actually measure what students know about science are undertaken through the regularly administered National Assessment of Educational Progress and recently have been internationally benchmarked through TIMSS. At lower and upper secondary levels, the results suggest that US students gain little real understanding of the big ideas of science, including life sciences, in spite of the fact that they are most likely to have had formal classes in this area than in other areas of science. It has largely been in the study of biology that the basic concepts of biodiversity have been advanced; diversity within species, across species and ecosystems; relatedness of species; interdependence among living things; threats to this diversity due to population pressures; loss of habitat; changes in species over time; natural extinctions; and human induced extinctions. While biodiversity and biodiversity education are more recent concepts, nature study, conservation, and environmental education have older roots, in a more

general emphasis on environmental improvement and appreciation. The following section will explore some of the history of environmental education and consider how it might relate to biodiversity education.

Environmental Education

In 1977 the world's first intergovernmental conference on environmental education was organized by the United Nations Education, Science, and Cultural Organization (UNESCO) in cooperation with the United Nations Environment Programme and held in Tbilisi, Georgia (Soviet Union). The Tbilisi Declaration, adopted at the conference, stressed the importance of environmental education in the preservation and improvement of the world's environment. Agenda 21, Chapter 36, outlines issues related to formal education, public awareness, and training to promote sustainable development and emphasizes environmental and development education as an essential aspect of all education. There were also calls for linking to environmental education in recommendations from the World Conference on Education for All (held in Jomtien, Thailand in 1990) that urged a move toward universal access to basic education for girls and boys.

The recommendations include calls for better informal education, promotion of environmentally sound leisure and tourism activities, programs to involve young people and children, as well as respect for and support of efforts to promote dissemination of traditional and socially learned knowledge through mechanisms based in local cultures.

Environmental education was called "nature study" when it got its start in the 1920s with Junior Audubon Clubs teaching children to appreciate nature. According to Karen Schmidt in a December 13, 1996, article in *Science*, the movement was transformed into conservation education in the 1930s when the Dust Bowl environmental tragedy led to incorporation of ideas into some schools about the management of natural resources.

With the initiation of Earth Day in 1970 and passage of the National Environmental Education Act of 1970, teachers received supplemental training in environmental education. Many states enacted their own environmental education laws, and schools began incorporating these topics into science classes. On the 20th anniversary of Earth Day, President Bush signed the National Environmental Education Act of 1990, which created an Office of Environmental Education at the Environmental Protection Agency and supported curriculum development and teacher training in the states.

In the US concern has been expressed about the place of environmental education, as currently configured into the curriculum, and especially about the quality of programs sometimes offered in lower secondary schools. This has included concerns about an advocacy orientation in instruction, about the need for and balance in materials used and instruction provided. The North American Association for Environmental Education (NAAEE), a professional group that includes college faculty and K–12 educators among its members, developed Guidelines for Excellence (North American Association for Environmental Education, 2004) to help guide teachers and others in the selection of quality, balanced materials. NAAEE membership includes individuals throughout North America as well as in 55 countries around the world.

Teaming with Life made a strong recommendation that environmental education should have a stronger base in science, using scientifically grounded curricula. Innovative programs such as Global Learning and Observation to Benefit the Environment (GLOBE) (see **Box 2**) that depend on student–scientist partnerships and collection of real data may point the way to science education based around environmental and biodiversity concerns.

Learning in the Informal Education Sector

Overview

A wide range of informal education experiences are available for adults and children to extend their knowledge about biodiversity. These include organizations that incorporate biodiversity education and exhibition within their missions, such

Box 2 Global Learning and Observation to Benefit the Environment (GLOBE)

GLOBE, a hands-on, school-based international environmental science program, was introduced by US Vice President Al Gore in April 1994 and began operation on Earth Day in April 1995. In 1999 there were more than 6000 participating schools in more than 70 countries. GLOBE brings together students, teachers, and scientists from around the world to enhance environmental awareness of individuals worldwide, increase scientific understanding of the earth, and support improved student achievement in science and mathematics.

Students make environmental observations, or take environmental measurements near their school site, report their observations *via* the Internet, receive and use GLOBE images created from the combined data, and study the environment by relating their observations to larger environmental topics. GLOBE educational materials were developed by environmental educators and curriculum development specialists working with scientists. Materials are used in schools under the guidance of teachers who have received training using GLOBE materials. Teachers include GLOBE activities as appropriate within their local curriculum. GLOBE materials are translated into the six United Nations' languages (Arabic, Chinese, English, French, Russian, and Spanish) and are also available in over 45 other languages. GLOBE international partners sign bilateral agreements with the US and manage participation of schools in their countries. In the US, GLOBE is an interagency program with management infrastructure funded by the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration (NASA), and the National Science Foundation, supported by, the U.S. Department of State and implemented through a cooperative agreement between NASA and the University Corporation for Atmospheric Research. In 2010 there were 111 participating countries, 54 000 GLOBE-trained teachers, and 23 000 schools in the program. Over 1.5 million students have participated, taking over 21 million measurements.

Observations range from basic weather parameters (temperature, atmospheric pressure, and precipitation) to measurements such as water chemistry, biodiversity, and biomass assessment. The focus to date has been more heavily directed toward physical systems measurements. Potential is great to increase the biodiversity and life systems aspects of the project.

as the following: Zoos; Botanical gardens; Aquariums; Museums; and National parks.

Depending on their size, these places of science might also incorporate research, collections, conservation, or other functions important to biodiversity. Signage and docent-led and audio tours provide additional information to visitors. Classes, lectures, and workshops (including those for teachers) are also often provided to extend the learning experience. Increasingly, materials and web sites incorporate aspects of a visit, bringing resources to audiences at a distance to make some part of the visitor experience remotely available.

Informal education also includes more intensive immersion experiences such as where it is incorporated into visits to natural preserves such as parks (including those within national park systems) and forests, or that provided through ecotourism.

Interpretive programs using volunteers, staff, and written and video materials provide enhanced learning experiences by bringing the science, the issues, and the concerns into sharp focus as a part of the overall environmental experience. Programs such as Earthwatch have biodiversity-focused visits that involve the participant in the research as data collector. Other informal learning opportunities are available through television and IMAX programs, web sites, and books.

Youth-serving groups provide a broad range of activities and experiences that can support education around concepts of biodiversity (see **Box 3**). Many may involve long-term projects of environmental monitoring, animal and plant breeding, habitat restoration, and other activities undertaken individually or in groups.

Exciting new online resources have been developed that serve both the formal and informal education communities around biodiversity education.

Zoos and Other Places of Science

Animal parks were established by and for rulers. Maier and Page, in their volume *Zoo: The Modern Ark* (Maier *et al.*, 1990), describe how animals were kept by royalty for entertainment and as a show of wealth. The third dynasty ruler of the Sumerian city of Ur had a park that dated around 2300 BC. A millennium later as civilization spread in the Near East and Asia, rulers and pharaohs exchanged "exotic" animals for their zoos. Emperor Wu Wang of the Chou dynasty laid out a zoological garden called the Park of Intelligence. Animal collections were found around the globe in early civilizations such as in Egypt some 3500 years ago.

Alexander the Great, perhaps influenced by Aristotle's private menagerie, installed what was perhaps the first public zoo in Alexandria, Egypt.

With the coming of the Dark Ages of Europe, monasteries became the keepers of menageries and game parks. When Cortes arrived in the Aztec capital of Tenochtitlan, he found a large zoo behind Emperor Montezuma's palace. Zoos in India were established by Akbar toward the end of the 16th century. He, like the Aztecs, employed people specially trained to care for and medically tend to animals.

The zoo at Vienna was reinvigorated by Maria Theresa and her husband as the Imperial Menagerie at Schönbrun for the

Box 3 Selected youth serving groups providing informal science/environmental education

American Camping Association (ACA)
- Founded in 1910
- ACA provides an accrediting mechanism for camps. Of the more than 2400 accredited camps listed, 284 provided nature/environmental study as part of the camp experience.

Boy Scouts of America (BSA)
- Enrollment: 5.6 million
- Program: Incorporated in 1910 BSA provides programs for boys that include outdoor skills, nature study, and conservation activities through an elaborated badge structure and group activities. BSA is a charter member of the World Scout Conference.
- World Scout Conference includes 145 member associations representing more than 25 million scouts.

Boys and Girls Clubs of America
- Enrollment: 3 million youth; served in 1006 local organizations
- Program: Activities include outdoor and environment education (The Ultimate Journey).

Camp Fire Boys and Girls
- Enrollment: 750,000 young people (birth–age 21)
- Program: Camping and environmental education programs offering children an appreciation and commitment to the natural environment.

Girl Scouts of the USA (GSUSA)
- Enrollment: 3.5 million
- Program: Activities include out-of-doors, nature study. Badge structure that includes environmental issues. GSUSA is a member of the World Association of Girl Guides and Girl Scouts.
- World Association of Girl Guides and Girl Scouts undertakes world projects including building world citizenship. Environment is one of the themes of this program.

Girls Incorporated
- Enrollment: 900,000 girls (ages 6–18) at more than 1000 sites nationwide
- Program: Activities include experiences in mathematics and science education through Operation SMART

National 4-H Clubs
- Enrollment: 6.5 million in more than 90,000 clubs
- Program: Part of the U.S. Department of Agriculture's Cooperative Extension Service established in 1914. Project areas include agricultural and natural sciences, and technology; Cornell Nest Box network.

convenience and entertainment of the nobility. The zoo remains today as likely the oldest in continuous operation, dating from the 1750s.

Democratization of Europe and establishment of urban centers that accompanied industrialization led to the modern zoo as a repository of exotic specimens of life that were to be studied as a way of understanding the flora and fauna of the world. Public monies (rather than private patrons or royal largesse) were available to begin systematic scholarly study. Maier and Page date the modern zoo to 1826 when the Zoological Society of London founded the zoological gardens at Regent's Park for the purpose of understanding the natural history of the animals inhabiting the reaches of the British Empire.

Since zoos as public institutions had to raise funds and attract money (independent of their research and conservation goals), they had to become popular attractions. Zoo organizers also had to learn to manage space and figure out and meet animals' requirements, such as for social interaction. Zoos' role in conservation became educational as they raised visitor awareness about endangered species and loss of habitat. Where larger zoos also developed significant breeding herds, they established breeding farms. In San Diego, for example, this "wild animal park" has become an additional attraction.

The National Zoological Park (National Zoo), associated with the Smithsonian Institution, established a "biopark," Amazonia, to emphasize the relationships among soil, plant, invertebrate, and other animal forms and the need to preserve the habitats of the world. Zoos, aquariums, and game parks are being seen as tools to affect public attitudes regarding the variety of life on earth.

As these places of science intentionally blend education and entertainment they are increasingly adding materials from museum collections and incorporating interactive exhibits from science–technology centers to reinforce conservation messages, concern about loss of species numbers and diversity, and loss of habitat.

Botanical Gardens

According to BGCI, the global network devoted to the conservation of threatened plant species, more than 200 million persons visited botanical gardens around the world. In addition to visits and guided tours, gardens offered continuing education for adults, workshops and hands-on experiences for children and families, and professional education courses and seminars for K–12 teachers. The New York Botanical Garden and Missouri Botanical Garden are examples of two gardens offering graduate studies programs, usually in collaboration with universities in their area.

Museums

Through collections, education programs, exhibitions, and graduate-level research, museums have been very active in promoting biodiversity in both the formal and informal sectors. The American Museum of Natural History in New York (AMNH) provides an interesting example of an institution with current involvement in all these areas:

- Exhibition. The 11,000-square-foot Hall of Biodiversity is a permanent exhibit of AMNH which opened in 1998. The exhibit uses collections, interactive technologies, and an immersive environmental replica of a portion of the rain forest of the Central African Republic – complete with sound, smell, movement, and running water – to provide a unique visitor experience.
- Graduate and continuing education. The Center for Biodiversity and Conservation collaborates within and outside the museum in the development of courses and programs. AMNH is home to the oldest and largest doctoral and postdoctoral training program of any scientific museum in the world, collaborating with Yale, Columbia, Cornell, and City University of New York.
- Education. The National Center for Science Literacy, Education and Technology, supported by the National Aeronautics and Space Administration (NASA) and a number of private foundations, has developed a number of projects related to the theme of biodiversity, including *Biodiversity Counts: A Student Inventory Project*, a program for middle school students across the US to inventory plant and animal life in their communities and to share their findings through publications and on-line field journals.

Biodiversity Experiences and Resources

Overview

For most adults, biodiversity education will take place in the informal sector as they read books; visit zoos, museums, and national parks; listen to lectures; and watch programs on the increasing number of science- and nature-based cable channels, public television, or the increased coverage of science on the news or news magazines. Others will visit the World Wide Web, where an increasing number of excellent sites developed by universities, museums, federal agencies, and nonprofit organizations provide high-quality information. Several examples of resources for biodiversity education available to the adult public follow.

Earthwatch

Earthwatch Institute is an international nonprofit organization founded in 1971 that supports scientific field research worldwide. Volunteers participate in actual field research, assisting scientists in gathering data. Since its beginning it has deployed over 85,000 volunteers into the field to work with 1000 scientists. Since 1995 Earthwatch's Capacity Development Program has trained over 1000 scientists including around species identification and other conservation issues. The Earthwatch web site lists active projects that volunteers can join in a number of topical areas including endangered species.

National Biological Information Infrastructure (NBII)

The NBII attempts to organize the disparate sources of information available through agencies, departments, museums, and other organizations, providing a source of links to sites. A "Biodiversity, Systematics and Collections" section connects to other web sites, many of which have education or "for kids" sections. Many federal agencies such as the U.S. Geological

Survey, the Environmental Protection Agency, NASA, and others have relevant sites.

Nonprofit Environmental and Biodiversity Groups

A number of organizations produce materials to support education about environmental and biodiversity issues. These include groups such as the World Wildlife Fund, the Sierra Club, and the Audubon Society. These groups develop a wide variety of public information and educational materials.

While most mainstream advocacy groups are conscious of the need to "get the science right" and to present balanced viewpoints, concerns are sometimes expressed about school use of materials that emanate from an advocacy position. Guidelines have been developed by NAAEE to assist educators in assessing the scientific accuracy of such materials.

Online Resources

Resources available online have added a new dimension to biodiversity education, most notably the Encyclopedia of Life (EOL). Initiated as a vision of biologist E.O. Wilson and funded by the MacArthur and Alfred P. Sloan Foundations, EOL is a free online collaborative reference tool that proposes to document every one of the 1.9 million known living species. As so-called Cornerstone Institutions, it brings together the resources of several of the world's leading natural history museums, botanical gardens, and libraries. The encyclopedia will continue to grow as more global partnerships are developed. The EOL went live on 27 February 2008 with 30,000 entries.

Available to support biological science educators, BiosciEdNet (BEN) is a collaboration of biological science organizations and other groups, managed by the AAAS, that aim to provide quality bioscience materials to support teaching and learning. Nearly 1688 reviewed materials covering 77 topics are aggregated under this portal, including almost 800 teaching resources for biodiversity. BEN is one of the National Science Digital Library sites supported by the National Science Foundation.

Public Awareness

How much does the public understand about environmental issues in general and biodiversity in particular, and what are the attitudes toward these issues? Surveys from a number of sources indicate that there is strong public interest in and support for issues related to the environment. The National Science Board's *Science and Engineering Indicators* (2010) suggested strong interest and "informedness" of the public around environment and health topics, especially when compared with other science and technology areas. The 1997 National Environmental Report card, an attitudinal and knowledge survey of American adults conducted by the National Environmental Education and Training Foundation and Roper Starch Worldwide, concluded that there was "an alarming lack of knowledge about some of our most critical environmental problems." With regard to biodiversity, however, 73% of adults surveyed correctly responded about the direct relationship between species loss and habitat destruction. The 1996 Biodiversity Poll (Belden & Russonello and Roper Starch Worldwide, 1997, 1998, 1999), conducted by the public opinion research firms Belden & Russonello and R/S/M, revealed the following about the environment and biodiversity:

- People care about the environment, but it isn't in the top tier of public concerns.
- Of environmental concerns, the public considers the most serious problems to be toxic waste, destruction of the rain forest, loss of places in nature, and air and water quality.
- Extinction is a concern, but it is not high on the list.
- People understand that nature is connected and interdependent, but most people do not recognize or use the word biodiversity. Only two in 10 said they had heard about the "loss of biological diversity."
- The public understands that species are declining and that human activity is largely responsible. But the public does not understand much about specific reasons or about the seriousness of the rate of loss.
- Public support for biodiversity conservation (once biodiversity is explained) is wide – 87%. But this support is shallow.
- Countervailing pressures (values) can peel support away from biodiversity protection. These include concerns about jobs, individual property rights, comfort and convenience, and preservation of unattractive species. However, 51% of Americans agree that the world would suffer if such unattractive species (e.g., mosquitoes) are eliminated.

The reasons Americans think biodiversity should be conserved included personal and family issues (79%), responsibility to future generations (71%), and spiritual concerns of stewardship (67%) (i.e., Nature is God's work).

Another survey of biological scientists, science educators, and the general public conducted by the American Museum of Natural History in April 1998 revealed that most scientists believe we are in the middle of a mass extinction largely caused by human activity (American Museum of Natural History, 1998). While 70% of the scientists surveyed rated loss of biodiversity as major and urgent, the general public was generally unaware of species loss and the threats this posed. Even science teachers, who are aware of the biodiversity crisis, did not believe that there was mass extinction, and only 38% of teachers rated themselves as being very familiar with the concept of biodiversity. Both 1996 and 1998 polls revealed a large gap between scientists' perception and the public's awareness and concern about biodiversity, this in spite of the public's perceived attentiveness and informedness about environmental issues in general.

The gap may relate to the following:

- A general lack of attention to biodiversity and its consequences by the media.
- The way that the public message about biodiversity is conveyed.
- The biological understandings that members of the public bring to the discussion.

Interestingly, the spottiness of adult knowledge and concepts (some individual ideas understood, but not the overall

concepts nor the consequences that flow from them) tracks with observations made about the K–12 student understanding of life sciences ideas. For most members of the public, high school is the last time for a formal course in the life sciences. After that, they generally rely on the informal sector for additional information and updates to their knowledge.

The 1999 National Environmental Report Card revealed that 84% of Americans believe that loss of animal and plant species is very serious or somewhat serious. Despite this support, however, specific concern for biodiversity within the larger frame of environmental issues appeared to have declined over the ensuing decade, as concerns for air and water quality have grown.

The United Nations declared 2010 as the International Year of Biodiversity; and later declared 2011–2020 as the United Nations Decade on Biodiversity. Such enhanced visibility will likely increase the number of events and level of attention directed to the public over this period of time.

Colleges and Universities

According to *Science and Engineering Indicators,* persons taking college-level science courses are more likely to be informed about, supportive of, and interested in science and technology topics. College level courses in biology influence knowledge and attitudes of the public. *Beyond Biology 101*, produced by the Howard Hughes Medical Institute, describes efforts to transform college-level courses in biology for both majors and nonmajors, moving away from vocabulary-driven courses to ones that are more integrative and that include meaningful laboratory and field experiences. These include an innovative program in human biology developed and in place since the 1960s at Stanford University. The program, taught by faculty from biology, education, anthropology, psychology, and other disciplines, focuses on the relationship between human biology and human behavior including human interactions with environment. It is a major course of interdisciplinary study and provides introductions to integrative biology for nonmajors.

The program for majors involves students in a range of activities designed to get them to think like scientists. Biodiversity and human impacts on the environment are explicit foci of instruction, and in the late 1980s and 1990s the college program was translated into a middle grades life sciences project (HumBio). Biodiversity concepts and activities are explicitly included among the curriculum materials in the ecology theme.

While most four-year institutions offer majors in biology or related subspecialties (including ecology), biodiversity was found as an area of major concentration at only a few US colleges and universities. A 2011 report published by the American Association for the Advancement of Science, *Vision and Change in Undergraduate Biology Education: A Call to Action*, summarized the consensus of conversations and discussions of life sciences educators as to what constitutes critical themes and competencies for introductory courses in biology; important among these was evolution as the driver for the diversity of life.

A search of the Society of Conservation Biology site, led to the international directory of conservation programs which lists 519 schools and 588 academic programs. Some 47 countries are listed in the directory as having academic programs.

The United Nations University, Institute for Natural Resources in Africa (UNU/INRA) was established in 1986 to "support the building of African capacity by strengthening national institutions to promote sustainable use of the continent's natural resources for development." Biodiversity and medicinal plants are included as one of the four thematic areas in which activities are focused.

Education and training are key areas of interest to UNU/INRA. In cooperation with other agencies of the UN, the program develops curricula and contributes to training in areas such as ecological economics, natural resource economics and environmental accounting, germplasm and biodiversity conservation, wildlife management, and taxonomy. Gender and Natural Resources is a major cross-cutting theme in the work of the institute.

Systematics Research and Training

While the demand for expertise in conservation, biodiversity, and systematics has been increasing, concern is being expressed about the human resources, especially in developing countries, available to manage and inform natural resources utilization around the globe. A 1995 workshop on Priorities in Systematics Research and Training organized by the United Kingdom Systematics Forum and held at the Linnean Society of London raised issues about the adequacy of support for systematic biology, the declining interest in systematics among students, and the decline in the teaching of systematics in many universities.

Fueled by the growing interest in biodiversity an increasing number of university departments worldwide provide training in systematics and taxonomy, often in partnership with research-oriented museums and botanical gardens. New emphases, such as work in molecular systematics, may exist in cellular, molecular, microbiology, or biochemistry programs. BIO NET-INTERNATIONAL is a global network of people and institutions that develop biosystematics capacity in developing countries. Training at all levels (in service, short courses, distance courses, and joint graduate programs) is a major focus of the network's activities.

Conclusion

A combination of school-based learning and out of school experiences combine to provide young people with knowledge of and attitudes about biodiversity that they then take into adulthood. For those who pursue higher education, college-level courses are available in some institutions that integrate biodiversity education into larger biological, environmental, or human impacts courses. Other adults must depend on the informal education sector, with experiences provided by a variety of different institutions and media. Biodiversity education may also, in some cultures, rely on community

transmission of locally held knowledge of plants and animals of a region. Whatever the process for developing understanding, education and public awareness have been seen as crucial precursors to building support for biodiversity.

See also: Biodiversity, Definition of. Conservation Efforts, Contemporary. Government Legislation and Regulations in the United States. Historical Awareness of Biodiversity. Human Impact on Biodiversity, Overview

References

Alonso A, Dallmeier F, Granek E, and Raven P (2001) *Biodiversity: Connecting with the Tapestry of Life*. Washington, DC: Smithsonian Institution/Monitoring and Assessment of Biodiversity Program and President's Committee of Advisors on Science and Technology (PCAST).

American Association for the Advancement of Science (2011) *Vision and Change in Undergraduate Biology Education: A Call to Action*. Washington, DC: AAAS.

American Museum of Natural History (1998) *Biodiversity in the Next Millennium*. New York, NY: Louis Harris and Associates, Inc.

Belden & Russonello and R/S/M (1996) *Human Values and Nature's Future: American's Attitudes on Biological Diversity, An Analysis of Findings from a National Survey*. Washington, DC: Belden & Russonello Research and Communications.

Committee on High School Biology Education, National Research Council (1990) *Fulfilling the Promise: Biology Education in the Nation's Schools*. Washington, DC: National Academy Press.

Maier F, Page J, and Durrell G (1990) *Zoo: The Modern Ark*. New York: Facts on File.

National Environmental Education and Training Foundation/Roper Starch Worldwide (1997, 1998, 1999) *National Report Card on Environmental Knowledge, Attitudes and Behaviors*. Washington, DC: National Environmental Education and Training Foundation.

National Research Council (1996) *National Science Education Standards*. Washington, DC: National Academy Press.

National Science Board (2010) *Science and Engineering Indicators*. Virginia: National Science Foundation.

North American Association for Environmental Education (2004) *Environmental Education Materials: Guidelines for Excellence*. Washington, DC: NAEE.

President's Committee of Advisors on Science and Technology (PCAST) (1998) Panel on Biodiversity and Ecosystems. *Teaming with Life: Investing in Science to Understand and Use America's Living Capital*. Washington, DC: PCAST.

Project 2061 (1993) Benchmarks for science literacy. *American Association for the Advancement of Science*. New York: Oxford University Press.

Project 2061 (1989) Science for all Americans. *American Association for the Advancement of Science*. New York: Oxford University Press.

U.S. Department of Education. Institute of Education Sciences (2007) *High school and beyond longitudinal study of 1980 sophomores*, high school transcript study; and 1987, 1990, 1994, 1998, 2000, and 2005 high school transcript study (HSTS). Washington, DC: National Center for Education Statistics.

Effects of Ecotones on Biodiversity

Salit Kark, The Hebrew University of Jerusalem, Jerusalem, Israel, and University of Queensland, Brisbane, QLD, Australia

Published by Elsevier Inc.

Glossary

Beta diversity Also called species turnover, beta diversity refers to the change in species as one moves across habitats, communities, or ecosystems.

Divergence-with-gene-flow model of speciation A model explaining the process of species formation (speciation) in which new species diverge in the face of gene flow, the movement of genes within a group that results from mating with immigrant individuals.

Ecotone A sharp transition zone between two or more different ecological communities or regions.

Ecotone effect The pattern of increased species richness (number of species) and abundance in ecotones and the occurrence of unique ecotonal species.

Edge effect The effect of the juxtaposition of contrasting environments on an ecosystem.

Geographic information systems (GIS) A computer-based system for creating and managing spatial data and associated attributes. It enables the capture, storage, retrieval, analysis, and display of spatial (location-based) data.

Remote sensing The science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation (e.g., *via* a satellite image).

Concepts and Terminology

Ecotones are areas where ecological communities, ecosystems, or biotic regions coincide. They often occur in areas of steep environmental transition, along environmental gradients. In these transitional regions, the environment rapidly shifts from one type to another based on abiotic (e.g., climatic) or biotic (e.g., community structure) factors (Holland *et al.*, 1991; Kent *et al.*, 1997). The origin of the word "ecotone" is in the Greek roots "oikos" (home) and "tonus" (tension). Many different definitions and terms have been used in the literature to describe ecotones and areas of ecological transition. These include boundary regions, borders, meeting zones, transitional zones, tension zones, zones of intermingling, and zones of transgression (Kent *et al.*, 1997). The definition often relates ecotones to more homogenous areas found on both sides of the transition or to the landscape as a whole. Ecotones can occur in both terrestrial and aquatic systems, and cover several spatial scales, ranging from large spatial-scale ecotones, where biomes meet to local-scale transitions, such as mountain treelines (Gosz, 1993; see **Figure 1**). These areas are sometimes considered to be dynamic zones of interaction between communities, which are unstable over time (Kent *et al.*, 1997). As suggested by Odum (1953), they do not simply represent a boundary or an edge; the concept of an ecotone assumes the existence of active interaction between two or more ecosystems with properties that do not exist in either of the adjacent ecosystems. Ecotonal regions show a diversity of boundary types that range from natural boundaries (e.g., altitudinal and latitudinal transitions) to human-generated ecotones, sometimes termed anthropogenic ecotones (e.g., forest clear-cut edges or urban ecotones) (Lloyd *et al.*, 2000), as shown in **Figure 1**.

Early work on the Effects of Ecotones on Biodiversity

Since the late 19th century, there has been interest in boundary regions and edges of ecological systems. A conceptual ecological framework for the study of ecotones was given by Odum in 1953. Odum discussed the ecotone effect, which he characterized as increased richness and abundance in ecotones and the occurrence of unique ecotonal species. Until the 1970s, there was considerable interest in ecotones within the scientific community (reviewed by Risser, 1995). This interest subsided when a focus on more homogenous and well-defined ecosystems and communities (e.g., tropical rainforests and tundra) became common. A revival of research in the field focusing on ecotones and their effects on biodiversity was seen in the late 1980s and 1990s, with the development of new research areas, especially those of landscape ecology and of conservation biology (Holland *et al.*, 1991). Studies on ecotones in the 1980s often focused on material flow (e.g., water and nutrients) across communities and on ecosystem processes in these boundary regions (Hansen and di Castri, 1992). Much of the work focused on wetlands and on riparian zones, where land–water interfaces occur (reviewed in Zalewski *et al.*, 2001). Later work in the 1990s more directly examined the effect of ecotones on biological diversity, and especially the relationship between ecotones and processes leading to morphological divergence, patterns of genetic and phenotypic diversity, species richness, rarity, and their

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00234-3

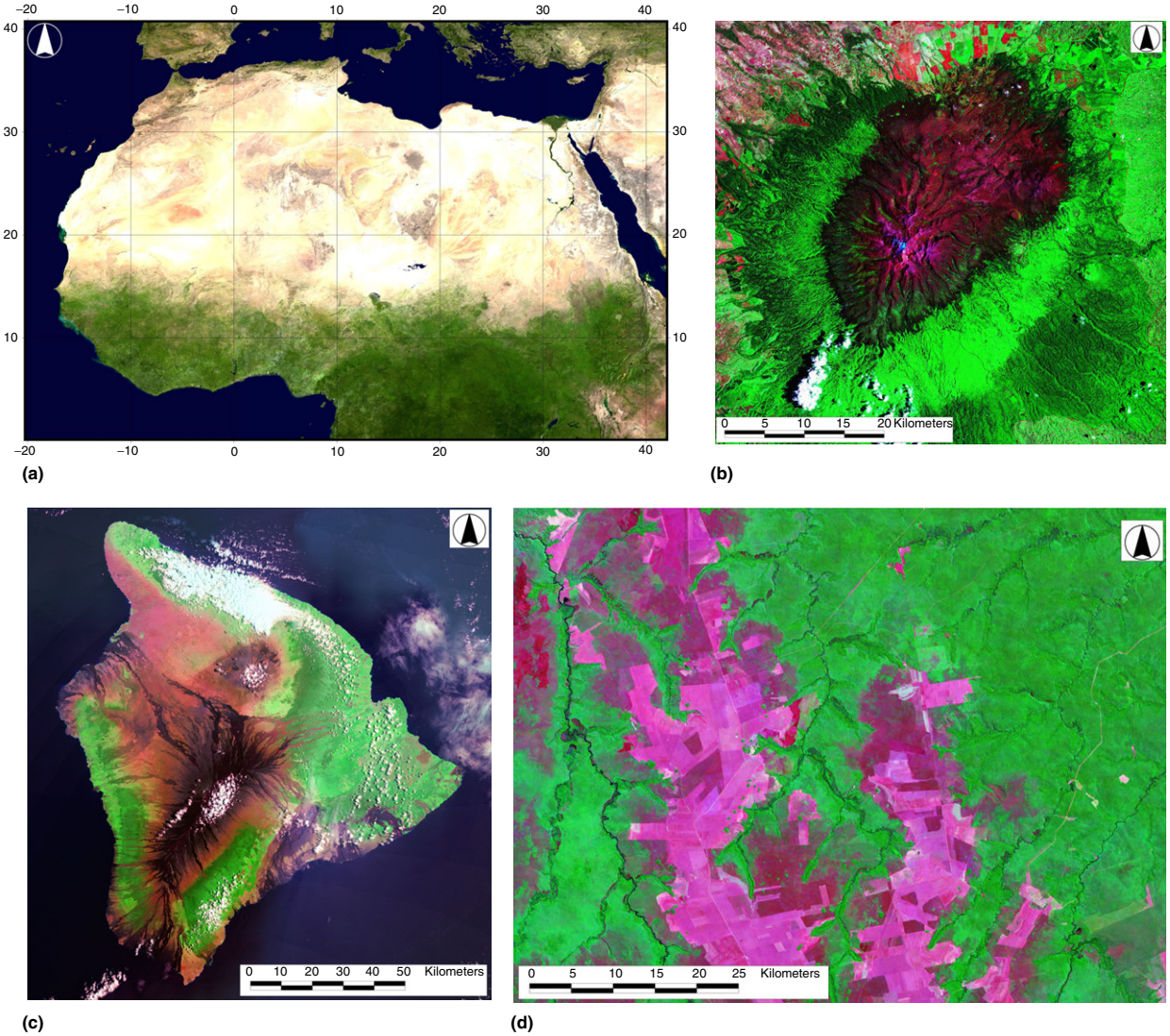


Figure 1 Satellite images showing a variety of natural and human-generated ecotones at several spatial scales: (a) Ecotones in the Sahel region of Africa (see text). (b) Altitudinal ecotones between vegetation belts in Mt. Kenya. (c) Natural and human-made ecotones on Hawaii's big island. Ecotones vary over different slopes and elevations. Note also the sharper ecotones in areas where lava has flowed. (d) Human-related ecotones resulting from deforestation (in pink) in the Amazon Basin, Brazil. On the far left and far right, in dark green, natural riparian ecotones can be detected along the rivers. Figure 1(a) reproduced from Stöckli R, Vermote E, Saleous N, Simmon R, and Herring D (2005) *The Blue Marble Next Generation – A True Color Earth Dataset Including Seasonal Dynamics from MODIS*. Greenbelt, MD, USA: NASA Earth Observatory. Downloaded from: http://visibleearth.nasa.gov/. Figures 1(b)–(d): False color composite of Landsat bands 7, 4, and 2. Green shades represent vegetation, magenta and brown represent bare soil, black represent water bodies, recent lava flows of shadows, cyan shades represent snow, and white represents clouds. Reproduced from MDA Federal (2004) Landsat GeoCover (1990/TM) Edition Mosaics; Tiles N-37-00 and S-37-00 (Mt. Kenya), N-14-15_20 (Hawaii), and S-21-10 (Amazon). Sioux Falls, South Dakota: US Geological Survey. Downloaded from: https://zulu.ssc.nasa.gov/mrsid/ – Applied Sciences Directorate (accessed November 2006).

conservation implications (reviewed in Kark and van Rensburg, 2006).

The Measurement of Ecotones

Because ecotones can rarely be delimited by a fine line, their measurement and mapping is not simple. A wide range of research approaches and tools have been used to detect and quantify ecotones. These include, among others, simulation modeling, geographic information systems (GIS), remote sensing, and statistical tools that enable quantification and analysis of ecotones of different types and over several spatial scales. Diverse approaches for the quantification of the steepness of gradients exist (Kark and van Rensburg, 2006). Methods for measuring and characterizing ecotones depend on the data available (e.g., quantitative or qualitative, grid- or transect-based data), with one of the simplest approaches, proposed by Womble in 1951, being the quantification of the magnitude of the first and second derivatives (rates of change in a given variable or several variables along a spatial gradient) (Womble, 1951). These approaches often examine the values

of variables in an area (e.g., a 1×1 km grid square) relative to its neighboring regions. The basic idea is to detect areas of sharp transition (which are referred to as boundaries or ecotonal regions) by finding the areas with the highest rate of change in the value of a given variable or several variables between adjacent squares (pixels). Specific software for the detection of boundary regions and analysis are now available (e.g., BoundarySeer: http://www.terraseer.com/products_ boundaryseer.php), enabling more widespread use of advanced statistical tools for the study of ecotones (Fortin *et al.*, 2000, Colwell, 2006).

In recent years, new approaches to quantify changes in diversity across gradients and boundary regions have been developed and are being applied. Among these is a range of new beta-diversity estimates of species turnover in space (Koleff *et al.*, 2003, Mena and Vázquez, 2005). These have been developed in the past decades, since Wilson and Shmida's (1984) review on beta-diversity estimates. Beta diversity and species turnover are often used when studying gradients, and although they do not focus necessarily on ecotonal areas, they can be applied to the study of ecotones.

One of the most promising directions in ecotone and boundary measurement is the use of tools developed in other areas of science. These include fields such as physics, remote sensing, and image analysis, where substantial advancements in boundary detection and gradient quantification methods have been made. Remote-sensing tools, for example, can use data occurring over several orders of magnitude, from satellite-derived data currently available at a resolution of 0.5 m to 100 km, to electronic microscope data. Further application of these tools at multiple spatial resolutions will provide a better understanding of ecotones. Recent work has shown that remote-sensing tools can be effectively used to detect ecotones and to predict species richness and rarity (e.g., range size rarity) in ecotones, especially in mountains where environmental transitions are sharp (e.g., Levin *et al.*, 2007, Levanoni *et al.*, 2011).

Patterns of Biodiversity in Ecotones

Studies on patterns of biodiversity in ecotonal areas have led to a range of results. Recent work is providing increasing evidence that boundary regions between ecological communities can be highly diverse at both the within-species and community levels. Ecotones have been shown to hold especially high biological diversity over several spatial scales, at both the community level (when examining species richness, i.e., the number of species in an area, e.g., Shmida and Wilson, 1985) and at the within-species level (morphological and genetic diversity) as reviewed by Kark and van Rensburg (2006). Other studies, however, have shown conflicting results, making it difficult to generalize without carefully examining each case, community, and region. Early on, Odum (1953) pointed at high species richness and abundance in ecotones and suggested that the ecotonal community commonly contains many species that are characteristic of, and sometimes restricted to, the ecotone. In a recent continental-scale study of New World birds, Kark *et al.* (2007) examined the relationship between passeriform richness and rarity of ~2300 passeriform species in 4889 one-degree New World grid cells and the distance of the cells to boundaries between adjacent plant-based ecoregions. They found that areas nearer to boundaries between ecoregions had more bird species, and also scored more highly in terms of species rarity. The findings of their work suggest that transitional environments harbor many rare species, in addition to high richness. At the community level, there is also some evidence for high species richness in ecotonal areas in marine systems. For example, van Rensburg *et al.*, (2009) showed at the sub-continental scale in South Africa that species richness and range size rarity at a spatial resolution of quarter degree are generally negatively correlated with distance to transition areas between vegetation communities for both birds and frogs. Areas with more range-restricted species were located significantly closer to transition areas between vegetation communities than expected by chance (van Rensburg *et al.*, 2009). Similarly, in the Gulf of Aden, Kemp (2000) found high-reef fish diversity in an ecotone harboring a unique mixing of the three distinct faunas of Oman, the Red Sea, and the Indian Ocean.

Processes Shaping Biodiversity in Ecotones

Ecotones and Evolutionary Processes

The process by which new species form is called speciation. This process is of major interest to evolutionary biologists who define three major types of speciation: allopatric, parapatric, and sympatric. These models are based on the degree of geographical subdivision between populations that lead to the formation of new species. Allopatric speciation happens in geographical isolation, and has been for many years considered the major form of speciation. Parapatric speciation occurs in adjacent populations with gene flow among them, often along clines. Sympatric speciation occurs when populations are geographically congruent, and are found in the same area. The study of ecotones has led to a better understanding of the potential importance of parapatric and sympatric speciation as mechanisms for speciation. Ecotones have been proposed to be centers of evolutionary novelty that maintain evolutionary process, and especially regions where parapatric (or sympatric) speciation processes may take place (Schilthuizen, 2000). As such, ecotones have been suggested as natural laboratories where evolutionary processes and barriers to gene flow can be examined (Schilthuizen, 2000). A review by Moritz *et al.* (2000) summarized the major models of evolutionary processes that promote diversification of rainforest faunas. They include the Gradient Model (either parapatric or sympatric), which suggests that adaptive divergence caused by selection forces occurs across sharp environmental gradients, leading to speciation even in the face of gene flow across ecotones. This means that speciation does not require isolation in cases where selection is strong enough to separate populations. This process is expected to occur especially where very different environments meet in the ecotone, for example, at the border between a dry and wet rainforest. In such regions, sharp boundaries may mean that even when gene flow continues, strong selection pressures can lead to divergence.

Support for the gradient model comes from recent research examining divergence using molecular genetic, phenotypic,

and experimental approaches (Smith *et al.*, 1997; Schneider *et al.*, 1999; Moritz *et al.*, 2000; Schilthuizen, 2000). Smith *et al.* (1997), studying the little greenbul (*Andropadus virens*), a passeriform bird in the rainforest–savannah ecotone region of Cameroon, found especially high morphological divergence in the ecotone. They proposed that when the ecotone is large enough, natural selection processes could be strong enough to generate morphological differences similar to those seen in reproductively isolated species even when high rates of gene flow occur. Their data support the divergence-with-gene-flow model of speciation (Smith *et al.*, 2000), leading them to propose that ecotones may be integral to the production and maintenance of high biodiversity in tropical rainforests. Increasing evidence suggests that ecotone regions may serve as centers for rainforest speciation. Quantification of morphological and geographic distances in olive sunbird (*Nectarinia olivacea*) populations in West African forests and ecotones revealed similar divergence patterns (Smith *et al.*, 2000). Smith studied morphological divergence in another Central African species, the black-bellied seedcracker (*Pyrenestes ostrinus*). Seedcrackers show polymorphism in bill size. A megabilled morph was found in the ecotone that specializes on a very hard-seeded sedge found only in ecotonal areas (Moritz *et al.*, 2000). This ecotonal megamorph was maintained in the population, despite high levels of gene flow with rainforest populations that had only smaller-billed morphs, owing to its selective advantage for feeding on the hard seeds. Morphological divergence between habitats across an ecotone was also found in leaf-litter skinks (*Carlia rubigularis*) in the wet tropical rainforest of Australia (Schneider *et al.*, 1999). Adult skinks occurring across sharp ecotones from open (wet sclerophyll) forests to adjacent rainforests showed large morphological and life history differences over short distances despite moderate to high levels of mitochondrial gene flow (Schneider *et al.*, 1999). Populations occurring across the ecotone had larger size differences than populations located dozens to hundreds of kilometers away that were geographically isolated millions of years ago, suggesting that in this case, and perhaps in many others that await research, speciation with gene flow may have great importance, even compared with allopatric speciation.

If boundary regions harbor unique and endemic species and alleles, this may provide support for the notion that these regions also serve as centers of speciation. If this is the case, ecotonal regions are expected to contain a preponderance of recently derived species that are yet to expand their ranges (neoendemics). Fjeldså and Rahbek (1998) suggested that more recently evolved species are concentrated in transitional ecotones surrounding the main central African rainforest. This reasoning is consistent with the finding that terrestrial ecotones sustain high morphological divergence, providing evidence that current speciation processes may indeed be taking place in these regions (Moritz *et al.*, 2000; Schilthuizen, 2000; Smith *et al.*, 1997).

Ecotones, Gradients, Species Ranges, and Variability

What might lead to higher species richness in ecotonal regions compared with adjacent areas? Evolutionary processes, as discussed above, may be one explanation. Ecological factors may also have importance in shaping this pattern. Several authors, such as Gosz (1993) and Risser (1995) have suggested that transitional areas not only share the two types of environments of the habitats that coincide in the ecotone, but also have a unique ecotonal environment. Indeed, Odum (1953) proposed that transition zones often support a unique community with characteristics additional to those of the communities that adjoin the ecotone, although also commenting this is by no means a universal phenomenon. Studies testing these predictions show mixed results, some pointing toward the occurrence of ecotonal species, while others not finding evidence for species unique and highly abundant in ecotones. The inconsistency among studies is complicated by the fact that different species, systems, scales, and regions were used in different studies or due to methodological factors, such as differences in sampling and analysis approaches. In addition, ecotones tend to shift in space and time over several spatial scales (Gosz, 1993; Kent *et al.*, 1997), as a response to climatic variation, other environmental changes (Crumley, 1993; Neilson, 1993; Kent *et al.*, 1997), and human activity (Gehrig Fasel *et al.*, 2007). They show high spatial and temporal heterogeneity, which may serve as important factors contributing to their high genetic and species diversity (Risser, 1995). For example, multiple ecotones can be defined within and around the African Sahel (**Figure 1a**), depending on the scale of interest and on the definitions used (Agnew and Chappell, 2000). The different transitions (e.g., that between the Sahel and the desert to the north) experience shifts in time and space, showing high spatiotemporal variability.

Another, simpler, process shaping this pattern is that ecotones, comprising meeting areas between adjoining communities, include a combination of species from two or more community types (Risser, 1995). Ecotonal areas often comprise the edge of the range for species on both sides, where many peripheral populations occur (Shmida and Wilson, 1985; Kark and van Rensburg, 2006). An important question is whether populations occurring in ecotonal areas are viable populations that exist over time within ecotones, or rather persist temporarily due to the constant flow of individuals from other parts of their range into the ecotone areas, and are not self-sustainable over time, and will disappear if this flow is stopped. Shmida and Wilson (1985) proposed that the high number of species in transitional areas could be due to a process they called the mass effect, which is the flow of individuals from favorable to unfavorable areas. For species that reach the margins of their range at the ecotone, this effect may result in some individuals of a given species establishing in ecotonal areas where they cannot maintain viable populations, existing in sinks adjacent to larger source populations (Shmida and Wilson, 1985). This may lead to increased species richness in ecotonal areas, which is maintained by constant immigration of individuals from more favorable environments. Some evidence for the existence of mass effects can be found in the literature, yet these effects seem to be rather weak, and it is currently unclear whether they can indeed act as a major factor generating high diversity in boundary regions (Kunin, 1998).

The mass effect, however, does not predict the occurrence of unique or endemic ecotonal species or genotypes. If some

species or genotypes are characteristic of an ecotone or occur at the ecotone in higher abundances than in the neighboring habitats, as proposed above, this could suggest that some ecotone populations are ecologically viable. Several recent studies have found peak genetic and morphological diversity within species in ecotone regions, with populations in these regions harboring unique and rare alleles not found elsewhere (Kark and van Rensburg, 2006). For example, a study examining allozyme (protein level) diversity within chukar partridge (*Alectoris chukar*) populations across a rainfall gradient from northern to southern Israel found that the highest levels of diversity occur in the sharp ecotone area between the Mediterranean region and the desert, in the northern Negev area. Populations in this area not only showed higher genetic diversity based on 32 loci (proteins) examined, but also had unique and rare alleles that did not occur elsewhere across the range (Kark *et al.*, 1999). The same species also had peak levels of morphological diversity in the ecotone region based on 35 traits and 23 ratios between traits (Kark *et al.*, 2002). More studies on genetic and species uniqueness in ecotonal areas are needed to enable generalizations.

Implications for Conservation

While an understanding of the effects of ecotones on biodiversity is important for evolutionary and ecological pure science purposes, it also has many implications for conservation of biodiversity and practical decision making. There has been an interesting discussion in the literature in recent years regarding whether transitional areas are valuable for conservation (e.g., see Smith *et al.*, 2001 versus Brooks, 2001). Some scientists have argued that because ecotones hold marginal populations at the edge of the range of many species, where abundances may be lower and populations more prone to local extinctions than other parts of the range, these ecotonal areas have low value for conservation as they will not persist over time, for example, if the surrounding environments become fragmented (e.g., Gaston *et al.*, 2001). Substantial conservation attention has been given in recent years to the understanding and mapping of biodiversity patterns and the underlying processes, and toward predicting the effects of global change. Ecotone and boundary regions, where change, shifts, and variability occur naturally in both space and time, could serve as useful models for understanding, monitoring, and predicting the response of individuals, populations, and communities to changing environments (Crumley, 1993; Neilson, 1993; Allen and Breshears, 1998). In addition, while some scientists suggest that ecotone populations are more likely to be negatively affected by climate change (Gaston *et al.*, 2001), other researchers have argued that ecotonal areas sustain populations that are adapted to changing, fluctuating, and unstable environments and, as such, these populations may be better able to persist in the face of predicted change (see further discussion in Kark and van Rensburg, 2006). This is especially relevant since climate shifts are expected, according to some studies, to be rapid and extreme in boundary regions between ecosystems (Allen and Breshears, 1998). By examining changes in ecotone locations over time, these areas may potentially serve as "early warning" indicators of global changes (Crumley, 1993; Neilson, 1993). However, the response depends on the spatial and temporal scales examined and may be a useful indicator mainly at global spatial scales and rather coarse timescales. Therefore, this area deserves further attention owing to the complexity of the factors affecting the location of ecotones in space and time.

Much research in the past decade has focused on prioritizing conservation efforts and determining what areas are most important and valuable for conservation. Two main approaches have been suggested. The first approach includes a search for biodiversity hotspots, or areas with especially high species richness, endemism, and rarity. A second approach has been to select areas that are complementary, and hold biodiversity not present in other areas. Ecotones may provide a unique opportunity to conserve both high species richness and high complementarity. Owing to their relatively small size, it may prove a cost effective strategy to further conserve ecotone regions and to explicitly include them in future systematic conservation planning, given they potentially provide a high return on investment having small area and high biodiversity. Undoubtedly, if these areas have the potential to maintain and to generate species richness as well as unique and novel species and forms, they deserve far greater research attention than they are currently receiving (Smith *et al.*, 2001). Nevertheless, conservation plans for ecotones should not be considered independent of their surrounding environments. Ecotonal regions are important for our understanding of evolutionary processes (such as speciation, divergence with gene flow, and adaptation) and ecological processes shaping biodiversity (e.g., response of populations to fluctuating environments). They may enable us to better predict the responses of populations to environmental change and to further identify previously neglected biodiversity hotspots valuable for biodiversity research and conservation.

Human-related Ecotones

While early literature discussing ecotones largely dealt with natural ecotones that are generated by environmental factors such as soils, geology, and sharp climatic gradients, recent research is increasingly including human-related boundaries. Human activity is generating boundaries that did not exist before, changing their steepness and shifting their location. These boundaries include a diversity of ecotone types, such as forest clear-cut edges, margins between built-up and natural landscapes, and human-generated features, such as lakes and plantations. Research on these human-generated ecotones and their effects on biodiversity is related to a study area that is sometimes termed "countryside biogeography," examining biodiversity in human-dominated landscapes. Another recently developing research area that is relevant is that of urban ecology. Sharp, human-caused transitions may result from human activities such as urbanization, land-use changes, agriculture, grazing, or burning (see **Figure 1**). These boundaries occur at multiple spatial scales, ranging from local ecotones between agricultural plots, urban areas, and roads, and their neighboring native habitat, to large-scale ecotones such as shifting desert borders owing to desertification processes and river divergence (**Figure 1**). These ecotones may be either static and fixed in space or dynamic and shifting in location over space and time.

A substantial amount of work in both natural and human-related landscapes has focused on what has been called the "edge effect." This is the effect of the juxtaposition of contrasting environments on an ecosystem. It refers to how the local environment changes along some type of boundary, or edge and how biodiversity is affected by such edges. This idea is attributed by animal ecologists to Aldo Leopold and his 1933 book Game Management. It encompasses a wide range of both biotic and abiotic trends associated with boundaries between adjacent habitat types, natural or anthropogenic. Much of the reference to edge effects in the recent landscape ecology literature has been related to human-caused boundaries, and especially to boundaries between forest fragments and neighboring patches of habitat that have been cleared. Again, there are no clear-cut conclusions as to the effect of human-dominated boundaries on biodiversity. The response largely depends on the type of edge and its history as well the species in focus. Its conservation implications are therefore complex and deserve further scientific attention. A vast amount of research has been done on the effect of forest edges, especially the effect of human-made fragments and their edges on biodiversity. A review of edges (Murcia, 1995) suggested there are many discrepancies in the literature, and a better understanding and search for general patterns requires a much more mechanistic approach to examine the processes underlying such patterns. In this context, there has been much effort to understand the effect of forest edges on predation, brood parasitism, and the breeding success of birds. Paton (1994) found that in the majority of studies, nest success varied near edges with an increase in both depredation and brood parasitism rates. The most conclusive studies suggest that edge effects in birds usually occur within 50 m of an edge (Paton, 1994). Since these reviews, multiple studies on dozens of species and regions have been conducted. While results show mixed patterns, small and intermediate-sized fragments of natural habitat (e.g., forest) tend to show stronger edge effects than large fragments.

Conclusions

Research on the effect of ecotones on biodiversity suggests that these areas of sharp transition are, at least in some cases, centers of high species richness, genetic and phenotypic diversity. Ecotones can sustain unique forms or species that are less abundant or do not occur elsewhere. In addition, studies suggest that ecotones are areas where some populations are diverging to new species in the face of gene flow (across the ecotone). However, these patterns are not repeated in all systems and we cannot yet say how general they are. Thus, the importance of ecotones to the generation and conservation of biodiversity, especially in the face of global climate and other changes, is an area of active research interest and potential conservation investment.

Acknowledgments

My thanks to Noam Levin for his help in generating the figure. This work was supported by the Israel Science Foundation (grant no. 740/04 to SK).

See also: Biodiversity, Definition of. Herbaceous Vegetation, Species Richness in. Remote Sensing and Image Processing

References

Agnew CT and Chappell A (2000) Drought in the Sahel. *Geological Journal* 48: 299–311.

Allen CD and Breshears DD (1998) Drought-induced shift of a forest-woodland ecotone: Rapid landscape response to climate variation. *Proceedings of the National Academy of Sciences* 95: 14839–14842.

Brooks TM (2001) Prioritizing hotspots, representing transitions. *Trends in Ecology and Evolution* 16: 673.

Colwell RK (2006) Estimates: Statistical estimation of species richness and shared species from samples. User's guide and application.

Crumley CL (1993) Analyzing historic ecotonal shifts. *Ecological Applications* 3: 377–384.

Fjeldså J and Rahbek C (1998) Continent-wide conservation priorities and diversification processes. In: Mace G, Balmford A, and Ginsberg J (eds.) *Conservation in a Changing World*, pp. 107–138. Cambridge: Cambridge University Press.

Fortin M-J, Olson RJ, Ferson S, *et al.* (2000) Issues related to the detection of boundaries. *Landscape Ecology* 15: 453–466.

Gaston KJ, Rodrigues ASL, van Rensburg BJ, Koleff P, and Chown SL (2001) Complementary representation and zones of ecological transition. *Ecology Letters* 4: 4–9.

Gehrig Fasel J, Guisan A, and Zimmermann N (2007) Tree line shifts in the Swiss Alps: Climate change or land abandonment? *Journal of Vegetation Science* 18: 571–582.

Gosz JR (1993) Ecotone hierarchies. *Ecological Applications* 3: 369–376.

Hansen A and Di Castri F (1992) *Landscape Boundaries: Consequences for Biotic Diversity and Ecological Flows*. New York: Springer.

Holland M, Risser P, and Naiman R (1991) *Ecotones*. New York: Chapman & Hall.

Kark S, Alkon PU, Safriel UN, and Randi E (1999) Conservation priorities for chukar partridge in Israel based on genetic diversity across an ecological gradient. *Conservation Biology* 13: 542–552.

Kark S, Allnutt TF, Levin N, Manne LL, and Williams PH (2007) The role of transitional areas as avian biodiversity centres. *Global Ecology and Biogeography* 16: 187–196.

Kark S, Mukerji T, Safriel UN, Noy-Meir I, Nissani R, and Darvasi A (2002) Peak phenotypic diversity in an ecotone unveiled in the chukar partridge by a novel estimator in a dependent sample (EDS). *Journal of Animal Ecology* 71: 1015–1029.

Kark S and van Rensburg BJ (2006) Ecotones: Marginal or central areas of transition? *Israel Journal of Ecology and Evolution* 52: 29–53.

Kemp JM (2000) Zoogeography of the coral reef fishes of the north-eastern Gulf of Aden, with eight new records of reef fishes from Arabia. *Fauna Arabia* 18: 293–321.

Kent M, Gill WJ, Weaver RE, and Armitage RP (1997) Landscape and plant community boundaries in biogeography. *Progress in Physical Geography* 21: 315–353.

Koleff P, Gaston KJ, and Lennon JJ (2003) Measuring beta diversity for presence–absence data. *Journal of Animal Ecology* 72: 367–382.

Kunin WE (1998) Biodiversity at the edge: A test of the importance of spatial "mass effects" in the Rothamsted Park Grass experiments. *Proceedings of the National Academy of Sciences* 95: 207–212.

Levanoni O, Levin N, Turbé A, Pe'er G, and Kark S (2011) Can we predict butterfly diversity along an elevation gradient from space? *Ecography.* 34: 372–383.

Levin N, Shmida A, Levanoni O, Tamari H, and Kark S (2007) Predicting mountain plant richness and rarity from space using satellite derived vegetation indices. *Diversity and Distributions* 13: 692–703.

Lloyd KM, McQueen AAM, Lee BJ, Wilson RCB, Walker S, and Wilson JB (2000) Evidence on ecotone concepts from switch, environmental and anthropogenic ecotones. *Journal of Vegetation Science* 11: 903–910.

MDA Federal (2004) Landsat Geocover (1990/TM) Edition Mosaics; Tiles N-37-00 and S-37-00 (Mt. Kenya), N-14-15_20 (Hawaii), and S-21–10 (Amazon). Sioux Falls, South Dakota: US Geological Survey.

Mena J and Vázquez Domínguez E (2005) Species turnover on elevational gradients in small rodents. *Global Ecology and Biogeography* 14: 539–547.

Moritz C, Patton JL, Schneider CJ, and Smith AR (2000) Diversification of rainforest faunas: An integrated molecular approach. *Annual Review of Ecology, Evolution, and Systematics* 31: 533–563.

Murcia C (1995) Edge effects in fragmented forests: Implications for conservation. *Trends in Ecology and Evolution* 10: 58–62.

Neilson RP (1993) Transient ecotone response to climate change: Some conceptual and modelling approaches. *Ecological Applications* 3: 385–395.

Odum EP (1953) *Fundamentals of Ecology*. Philadelphia: W. B. Saunders Company.

Paton PWC (1994) The effect of edge on avian nest success: How strong is the evidence? *Conservation Biology* 8: 17–26.

van Rensburg BJ, Levin N, and Kark S (2009) Spatial congruence between ecotones and range-restricted species: Implications for conservation biogeography at the sub-continental scale. *Diversity and Distributions* 15: 379–389.

Risser PG (1995) The status of the science examining ecotones. *BioScience* 45: 318–325.

Schilthuizen M (2000) Ecotone: Speciation-prone. *Trends in Ecology and Evolution* 15: 130–131.

Schneider CJ, Smith TB, Larison B, and Moritz C (1999) A test of alternative models of diversification in tropical rainforests: Ecological gradients vs. rainforest refugia. *Proceedings of the National Academy of Sciences* 96: 13869–13873.

Shmida A and Wilson MV (1985) Biological determinants of species diversity. *Journal of Biogeography* 12: 1–20.

Smith TB, Kark S, Schneider CJ, Wayne RK, and Moritz C (2001) Biodiversity hotspots and beyond: The need for conserving environmental transitions. *Trends in Ecology and Evolution* 16: 431.

Smith TB, Wayne RK, Girman DJ, and Bruford MW (1997) A role for ecotones in generating rainforest biodiversity. *Science* 276: 1855–1857.

Smith TB, Wayne RK, Girman DJ, and Bruford MW (2000) Evaluating the divergence-with-gene-flow model in natural populations: The importance of ecotones in rainforest speciation. In: Bermingham E, Dick CW, and Moritz C (eds.) *Rainforests Past and Future*, pp. 148–165. Chicago: University of Chicago Press.

Stöckli R, Vermote E, Saleous N, Simmon R, and Herring D (2005) *The Blue Marble Next Generation—A True Color Earth Dataset Including Seasonal Dynamics from MODIS*. Goddard Space Flight Center, Greenbelt, MD, USA: NASA Earth Observatory.

Wilson MV and Shmida A (1984) Measuring beta diversity with presence–absence data. *Journal of Ecology* 72: 1055–1064.

Womble WH (1951) Differential systematics. *Science* 114: 315–322.

Zalewski M, Schiemer F, and Thorpe J (2001) Fish and land-inland water ecotones: Overview and synthesis. *Ecohydrology Hydrobiology* 1: 261–266.

Elevational Trends in Biodiversity

John-Arvid Grytnes, University of Bergen, Bergen, Norway
Christy M McCain, University of Colorado at Boulder, Boulder, CO, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alpha diversity Number of species within a restricted area, usually a small sample from a homogenous habitat.

Beta diversity Change in species composition from one area to another.

Ecotone An abrupt spatial transition from one community to another often caused by change in environmental conditions or in the disappearance of a dominant species like the forest limit.

Gamma diversity Total number of species within a large area or landscape encompassing a number of samples from different communities.

Mid-domain effect A hypothesis based on random distribution of species ranges within a bounded domain resulting in maximum species richness in the middle of the domain.

Source–sink dynamics Dynamics between populations that are net exporters of individuals and populations that are net importers of individuals.

When walking along an elevational gradient we see many features change as we walk upward passing through various life zones. Most dramatically it may change from a dense tropical forest that is packed with visible biological life to the snow capped tops with no sight of life. When comparing these extremes it is obvious that species richness is lower at the extreme mountain tops than in the lowlands. The striking changes in climate, plant communities, and faunal assemblages with increasing elevation have led to the formulation of many of the most widely accepted ecological concepts, including niche theory, life zones, community assembly, and insular biogeography. The interest in elevational patterns, particularly in species richness, has fluctuated over the last 100 years, but has greatly increased over the last decade as a result of the increasing interest in broad-scale patterns of biodiversity. Rigorous quantification of elevational species richness patterns from different parts of the world may prove to be the key to a much-needed understanding of broad-scale diversity mechanisms.

History of Elevational Studies

It has probably always been acknowledged, and it was clearly noted by the nineteenth century naturalists, that diversity of plants and animals was higher in the lowlands than at the mountain tops. Willdenow, von Humboldt, Darwin, and Wallace made detailed observations along altitude and noted that the stature and diversity behaved in the same way along altitude and latitude (Lomolino, 2001). In fact, Linne and Willdenow went further and actually explained the latitudinal distribution of species to be a result of elevational distributions during the Great Flood (Lomolino, 2001).

Among the first rigorous quantitative descriptions of elevational species richness patterns were works by Grinnell and Storer (1924), Whittaker (1952, 1960), and Whittaker and Niering (1965). Grinnell, known for coining the niche concept, detailed species distributions, abundance patterns, and range limits of vertebrates along an elevational transect in the Yosemite region of the Sierra Nevada, CA, USA (1924). His study emphasized that the "amplitude of the general environment – the number and extent of distinct ecologic niches it compasses – determines the richness of the fauna, both as regards number of species, and the number of the individuals to the unit of area representing each species". While Grinnell used elevational distribution patterns to understand niches and richness patterns, Whittaker undertook elevational gradient studies in the Great Smoky Mountains (insects; 1952), Siskiyou Mountains (plants; 1960), and Santa Catalina Mountains (plants; Whittaker and Niering, 1965) to tease apart important potential climatic drivers of species richness and abundance (e.g., water availability, soil type, temperature), and to determine the cohesiveness of communities.

Whittaker's elevational studies tested community cohesion by contrasting the support for the two main hypotheses: Gleasonian continuum hypothesis versus Clementsian discontinuity hypothesis. The discontinuity hypothesis proposes that groups of species have a similar distribution along an environmental gradient with more or less a clearly defined transition to another group of species, i.e., that the community actually behaved as a superorganism with respect to the environment. The continuum hypothesis proposes that the individual species are distributed independently from each other and that centers and boundaries of species populations are scattered along the environmental gradient. With numerous studies, using elevational gradients as their test system, Whittaker and colleagues found that species tended to have independent range boundaries providing evidence against the discontinuity hypothesis and supporting the continuum hypothesis. These studies had an enormous impact on vegetation ecology, sparked gradient analyses, and ignited the use and development of ordination techniques and related statistical methods in ecology.

While discussing the cohesiveness of communities, Whittaker and colleagues detailed both monotonically decreasing species richness with increasing elevation and humped pattern with maximum species richness at intermediate elevations (e.g., **Figure 1(a)**). Grinnell and Storer (1924) determined

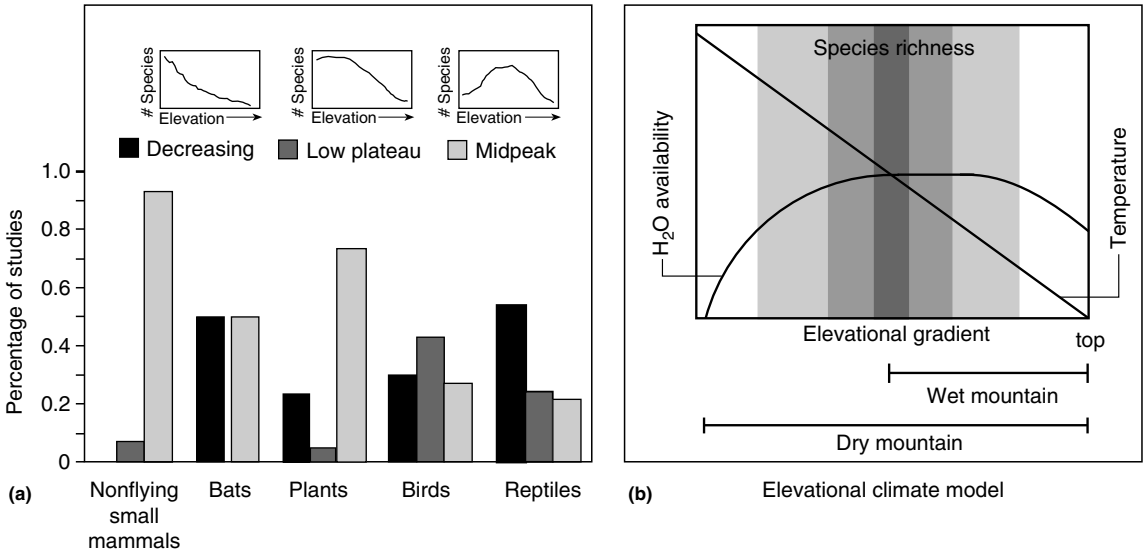


Figure 1 (a) The percentage of each of the three main elevational richness patterns demonstrated in the robust, informative montane gradients across the globe: nonflying small mammals (McCain, 2005); bats (McCain, 2007a, b); plants (from Rahbek, 2005, Figure 3f; plant percentages rescaled to include only these three diversity patterns); birds (McCain, 2009); and reptiles (McCain, 2010). (b) Generalized climatic model for elevational species richness, incorporating a linearly decreasing temperature gradient and a unimodal water availability gradient. Species richness is depicted in gray tones with darker tone indicating more species. The placements of generalized wet and dry montane gradients are shown below the *x*-axis (e.g., McCain, 2007a, b). Adapted from McCain CM (2007a) Could temperature and water availability drive elevational diversity? A global case study for bats. *Global Ecology and Biogeography* 16: 1–13.

that each group of vertebrates (bats, rodents, breeding birds, amphibians, and reptiles) exhibited a humped richness pattern with the highest species richness between 1000 and 1300 m. One important conclusion is that richness patterns found in Whittaker's studies differed among species groups (e.g., trees, bushes, and herbs; flies, beetles, and grasshoppers), whereas different vertebrate groups in Grinnell's studies showed consistent diversity trends with elevation. The combination of studying patterns in species richness and gradient analyses naturally leads to a way of quantifying the degree of change in composition of communities along a gradient, which Whittaker (1960) termed beta diversity.

After Whittaker, most attention to broad-scale patterns in diversity was focused on latitudinal gradients. This continued until the seminal work of John Terborgh on bird communities along an elevational transect in the Peruvian Andes (e.g., Terborgh, 1977, 1985). Terborgh's widely cited bird study detailed strongly decreasing richness with increasing elevation and showed support for the importance of habitat complexity, competition, and ecotones in bird diversity patterns. Thereafter, decreasing diversity along elevational clines became the accepted and assumed pattern for all taxonomic groups, that is, the elevational pattern mirrored the latitudinal pattern of diversity (e.g., Brown and Lomolino, 1998). As time passed the humped elevational patterns of Grinnell and Whittaker were forgotten, although their contributions on niches, community cohesiveness, and gradient analysis were not. Rahbek (1995) began to question the pervasiveness of the decreasing diversity pattern, which he claimed was based on very few studies that had serious sampling biases and called for more rigorous studies on elevational patterns. He presented a preliminary overview of elevational diversity patterns, and found more support for humped richness patterns with the highest richness at intermediate elevations than decreasing patterns. After Rahbek's study in 1995 the interest in elevational patterns in species richness has increased enormously. Today, elevational gradient studies are often motivated by increasing our knowledge about broad-scale patterns in diversity and the search for global drivers of biodiversity is so urgently needed for conservation efforts worldwide. As noted by Grinnell and Whittaker, and repeated recently by Lomolino (2001) and Nogués-Bravo *et al.* (2008), the elevational gradient is ideal for doing studies on broad-scale distribution of species and species richness.

Observed Patterns and Ways of Quantifying Patterns

Two types of data are used to quantify the elevational pattern in species richness. Alpha diversity studies use local field sampling of plots along a transect usually on one mountain slope preferably with equal sampling effort at each elevational band. Alternatively, gamma diversity studies use regional data from previously collected specimens and field records for an entire mountain or mountainous region. Such regional data are often readily available and summarized in flora and faunas from mountainous areas around the world.

Unlike other broad-scale gradients, field sampling is feasible for elevational transects and together with appropriate statistical methods can give a robust richness estimate. When using museum or historical data the observed pattern may be very sensitive to the intensity of sampling. One common way to account for sampling intensity is to assume that each species is present at all elevations between the upper and lower

elevations of observation (interpolation). Interpolation can create an artificial hump in species richness along the elevational gradient if sampling is incomplete (Grytnes and Romdal, 2008). If more detailed specimen information is available, uneven sampling may be accounted for by using rarefaction or extrapolation (Grytnes and Romdal, 2008), without such information error simulations can test for pattern robustness (e.g., McCain, 2007a).

Rahbek (2005) conducted an overview of elevational richness patterns from the literature. The majority of these studies treated plants and he demonstrated that almost 50% of these studies found a humped pattern, and around 25% had a monotonic decrease with elevation. The fraction of hump-shaped patterns increased to around 70% after excluding studies that did not consider the whole gradient (**Figure 1(a)**: plants). Rahbek's study also demonstrated the importance of scale, for example, a hump-shaped pattern is more common if a single transect is studied (i.e., alpha diversity), than if the pattern is studied on a whole mountain range (i.e., gamma diversity).

McCain has performed a series of global meta-analyses on elevational richness patterns among taxonomic groups from published studies. These and previous analyses clearly show that the observed pattern depends on the organism studied and the local climatic conditions (**Figure 1(a)**; McCain, 2005, 2007a, 2009, 2010; Rahbek, 2005). Nonflying small mammals (rodents, shrews, marsupials) almost ubiquitously demonstrate unimodal richness patterns with the highest richness at intermediate elevations (robust, informative gradients (RIG) = 54; McCain, 2005). Bats demonstrate two global patterns: half of the studies found decreasing species richness with increasing elevation and the other half found unimodal richness patterns (RIG = 12; McCain, 2007a). As stated above, Rahbek found that plants tend to show mostly unimodal richness patterns (RIG = 21; from Rahbek, 2005, Figure 3f). Birds show more variation in their elevational richness patterns: 30% are decreasing, 43% have high diversity across most of the lower portion of the gradient then decrease (low plateau in diversity; e.g., **Figure 1(a)**), and the final 27% have unimodal richness (RIG = 95; McCain, 2009). Finally, for reptiles 54% show a decreasing pattern, 25% have a low plateau, and 21% have a unimodal richness pattern (RIG = 24; McCain, 2010; **Figure 1(a)**).

Mountain regions often host a large fraction of endemic species (e.g., Orme *et al.*, 2005). Considering that isolation is an important factor for speciation; it is no surprise that fraction of endemic species tends to increase with altitude resulting in a peak in species richness at intermediate elevations above the peak in total species richness. For vascular plants in the highest mountains of the world the fraction of endemics increases linearly from the lowlands to the highest point where plants are found (around 6000 m above sea level) (Vetaas and Grytnes, 2002). This results in a peak in endemic richness around 4000 m, whereas the total number of species peaks around 1500 m. Studies on avian mountain endemics demonstrated their greatest diversity at intermediate elevations between 1500 and 3000 m, although somewhat lower on shorter mountains, even though overall diversity decreases monotonically with elevation (e.g., Stotz *et al.*, 1996). Such contrasting patterns in total species richness and endemic species richness are most likely commonplace along elevational gradients, particularly for highly diverse groups.

Discussion of Possible Causes

Elevational gradients are invaluable for discerning between diversity hypotheses. The small spatial scale, the thousands of independent replicates on mountains across the globe of various heights and in various climates, the high variability in richness patterns among taxonomic groups, and the predictable trends in abiotic factors with elevation allow globally distributed elevational gradients to be used as natural experiments, allowing for rigorous testing of hypotheses. The causes commonly mentioned for elevational patterns in species richness are very similar to the causes used to explain other broad-scale factors in species richness. These can be grouped into four categories: climatic hypotheses based on current abiotic conditions, spatial hypotheses of area and spatial constraint, historical hypotheses invoking processes occurring across evolutionary time scales, and biotic hypotheses (e.g., community overlap (ecotones), source–sink dynamics, and habitat heterogeneity). Below the authors describe some of the most commonly asserted hypotheses and assess their current level of support.

Climatic Hypotheses

Climatic variables like temperature, rainfall, and productivity are probably the most commonly cited causes for broad-scale patterns in species richness and elevational patterns are no exception. Temperature has a simple relationship with altitude as it decreases monotonically by 0.3–0.6 °C per 100 m elevational gain. Rainfall often follows a more complex relationship with altitude and maximum rainfall is often found at intermediate elevations, but is also known to increase with elevation or be high across a broad band of low to intermediate elevations. In tropical areas the zone of maximum humidity often corresponds to the cloud zone and horizontal precipitation from low-lying clouds can significantly increase the water availability at those elevations.

Climate may affect elevational species richness patterns in several ways. First, climatic tolerances of the studied species may put restrictions on how many species that can survive at different elevations. This will have different effects on different species groups, and may be a result of their evolutionary history and niche conservatism (Wiens *et al.*, 2010). Some species groups (e.g., epiphytic plants, salamanders) are dependent on high and constant moisture, whereas others may be restricted by a certain winter temperature. As a result, different species groups will show different elevational richness patterns, exactly what was demonstrated by Whittaker's early studies and confirmed by McCain's meta-analyses. Second, species richness may depend on productivity through the number of individuals that are found in an area. Higher productivity leads to higher number of individuals, which in turn leads to higher species richness. Primary productivity is dependent on temperature and precipitation. Because rainfall in many cases increases with elevation or has a humped relationship with

altitude, highest productivity may be found in the middle of the elevational gradient in many cases. In arid or seasonally dry areas where precipitation is low at the lowest elevations, productivity may decrease with temperature because higher temperature leads to higher evaporation. This may explain differences in patterns between mountains with wet versus dry local climates so that species richness on wet mountains is monotonically decreasing with increasing elevation whereas on dry mountains richness peaks at mid-elevations (**Figure 1(b)**). Evaluation of this model shows very good support for bats, as all bat diversity patterns studied from dry based mountains show peaks in richness at intermediate elevation, whereas all but one study on wet based mountains show strongly decreasing bat richness with increasing elevation (McCain, 2007a). Brown and Lomolino (1998) also concluded that dry montane environments may show mid-elevational peaks in diversity across multiple taxonomic groups due to the higher water availability at intermediate elevations.

Spatial Hypotheses

Area

The classical species–area relationship is often asserted to explain elevational species richness patterns, predicting more species in elevational bands that cover more area. Most area on mountains occurs at lower elevations (Körner, 2000) but in some areas, particularly mountainous regions, steep valleys at low elevations cover less area and lead up to a large plateau at intermediate elevations and in such cases area is more extensive at mid-elevations. This variation predicts different elevational richness patterns between mountains if area determines richness and may serve as an exceptional test system for evaluating the importance of area on broad-scale patterns.

So far, very few studies have investigated the effect of area on elevational patterns in species richness. McCain (2007b) evaluated the species–area relationship across 34 globally distributed mountains. Area influences montane richness patterns to a surprisingly low degree; overall only 38% of the studies showed strong responses to area (i.e., had a significant impact on the elevational pattern in species richness). In these cases correcting for area generally resulted in changing linearly decreasing patterns to mid-elevational richness peaks, but the area effect does not appear to be consistent enough among studies to be the main driver of richness patterns (McCain, 2007b). Another important point is that the high impact of human influence in the lowlands may reduce the available area for species at lower elevations along many elevational gradients (Nogués-Bravo *et al.*, 2008)

Mid-Domain Effect (MDE)

MDE is a relatively new hypothesis for explaining broad-scale patterns in species diversity. The hypothesis predicts a humped species richness pattern when species ranges are randomly distributed within a geometrically constrained domain (i.e., base and top of a mountain). A terrestrial species range cannot extend over the top of the mountain or below the base at sea level or the lowest regional elevation. Most of the elevational studies published lately discuss the potential of MDE as an explanatory factor. The conclusion from these results is that MDE predictions can sometimes be highly correlated with the observed pattern (e.g., Kluge *et al.*, 2006), but in the majority of the studies the fits to the model are low (Dunn *et al.*, 2007). MDE may play a role in concert with other factors such as area and climate, but are probably not the main driver of elevational richness patterns.

Biotic Hypotheses

Ecotone and Source–Sink

Along elevational gradients the distance between very different climatic zones and hence also between different communities and biomes is short. This implies that there will be short distances between optimal and suboptimal areas for many species along the gradient. This may result in a net flow of seeds or propagules from optimal to suboptimal areas. Even though a population of a species usually would not survive in a suboptimal area over time, the extra propagules received from populations in the nearby optimal areas (usually called source populations) may result in persistence of the populations in the suboptimal areas (sink populations) (Pulliam, 1988; mass effect is another term used to describe this process, Shmida and Wilson, 1985).

Source–sink dynamics will generally inflate species richness along the whole elevational gradient as new species are added locally by sink populations. This may affect elevational species richness patterns in two ways. An elevational pattern in species richness may be created by source–sink dynamics only if some areas receive more sink species than other areas. This may be the case around ecotones where more sink species may be found than in surrounding areas resulting in peaked species richness around ecotones. Studies that have looked for the ecotones specifically have difficulties in actually detecting an abrupt change in species composition or increased species richness around the assumed ecotones (Terborgh, 1985). Alternatively, source–sink dynamics may create an elevational pattern if the lower and upper part of the elevational gradient receives less sink species than the mid-elevational parts. Mid-elevational areas will receive sink populations from source populations both above and below, whereas the upper and lower part of the gradient will only receive sink species from one direction. This will create a humped pattern, with maximum species richness in the middle of the elevational gradient. Using sterile species as indications of sink species studies on plants have shown that the mass effect may be important for shaping the elevational species richness pattern (Grytnes *et al.*, 2008; Kessler, 2009).

Habitat Heterogeneity

Heterogeneity will certainly have a large effect on species diversity. It is, however, difficult to say anything general about how heterogeneity will vary with elevation. The relevant type of heterogeneity will depend very much on the species group studied and on the scale of study. For bird species that forage in forest trees or for epiphytic plants the important heterogeneity will certainly depend on height of canopy and number of strata that can be defined in the forest (e.g., MacArthur and MacArthur, 1961; Terborgh, 1977). This will generally decrease

with altitude, or be related to productivity. An opposite example can be found by looking at moisture variation at relatively small scale. At high elevations water runs in small canals creating a high heterogeneity of moisture. At lower elevations the water gathers in larger and larger waterways and at small to intermediate scales the moisture heterogeneity will increase with elevation. Owing to the difficulty in defining the taxon-appropriate habitat heterogeneity and obtaining robust field measurements, habitat heterogeneity has not been tested rigorously beyond the elevational studies of Terborgh (1977).

Historical Hypotheses

Ultimately, species richness patterns result from differences in speciation, extinction, and immigration rates. Speciation rates, extinction rates, and clade age are thought to be correlated with latitude causing the highest richness near the equator through influence from climate and/or area (Brown and Lomolino, 1998). Examining elevational gradients eliminates this problem by examining diversity patterns within a single region of potentially uniform clade ages and latitudinally similar speciation and extinction rates. The predictions of evolutionary rates have been less developed along elevational gradients, but in addition to a positive relationship between climate and evolutionary rates that may cause high evolutionary rate at low elevation; the increasing isolation toward higher elevation will work in opposite direction. Elevational trends in evolutionary rates will certainly have little effect on mountains whose biota mainly comprised organisms that have colonized from a larger regional pool (e.g., areas covered by ice during the last ice age) compared to large mountain ranges like the Andes or Himalayas that also have endemic species generated through various gradients of speciation, extinction, and climatically fluctuating range shifts. The consistency of taxonomically and thus ecologically linked climate trends on elevational gradients (e.g., small mammals, bats, birds; **Figure 1**) lends support not only to current climate drivers but also to past climatic affinities. It may be that the signal of highest richness on mountains – in wet and warm conditions – is because these were the conditions under which most taxonomic groups and species originated. This lends support to importance of niche conservatism in shaping the elevational gradients (Kozak and Wiens, 2010) and diversity gradients (e.g., Wiens *et al.*, 2010), which posits that modern large-scale species richness patterns result from that fact that most modern groups and species originated when the majority of the earth was experiencing tropical-like conditions and these strong affinities still exist in current climatic regimes.

Conclusions

The unique opportunities for evaluating hypotheses for broad-scale patterns along elevational gradients in species diversity have not been fully utilized to date. Comparing independent transects and searching for similarities and differences in patterns between transects and among taxonomic groups in different climates (tropics, deserts, temperate regions), biogeographic regions (islands, continents) or between transects on mountains of different size or aspects (dry versus wet slopes) will certainly improve our understanding of mechanisms underlying broad-scale patterns in diversity. We need to develop more predictions specifically for comparisons of multiple elevational transects (as those shown in **Figure 1(b)**), gather robust climatic data at small spatial scales, and optimally design new elevational studies to test predictions of diversity theory. Elevational trends in other aspects of species beyond diversity may help define other important aspects of global ecological patterns. For example, studies of elevational trends in range size coupled with physiological tolerances will improve our understanding of role of climatic tolerances driving broad-scale diversity patterns (e.g., Janzen, 1967: why mountains are higher in the tropics) and the potential impacts of climate change. Similarly, understanding how biotic interactions (e.g., disease, food resources, competition) influence species ranges and abundance patterns is fundamental to understanding niche dynamics, as well as the strength of coupled responses to climate warming. For these reasons and because climatic shifts may be more rapid along the smaller scaled montane gradients, elevational gradients may be ideal for long-term monitoring of species responses to climate change.

See also: Diversity, Community/Regional Level. Latitudinal Gradients of Biodiversity. Scale, Concept and Effects of. Species Diversity, Overview. Species–Area Relationships

References

Brown JH and Lomolino MV (1998) *Biogeography*, 2nd edn. Sunderland: Sinauer Associates Inc.

Dunn RR, McCain CM, and Sanders NJ (2007) When does diversity fit null model predictions? Scale and range size mediate the mid-domain effect. *Global Ecology and Biogeography* 16: 305–312.

Grinnell J and Storer TI (1924) *Animal Life in the Yosemite*. Berkeley, CA: University of California Press.

Grytnes JA, Heegaard E, and Romdal TS (2008) Can the mass effect explain the mid-altitudinal peak in vascular plant species richness? *Basic and Applied Ecology* 9: 373–382.

Grytnes JA and Romdal TS (2008) Using museum collections to estimate diversity patterns along geographical gradients. *Folia Geobotanica* 43: 357–359.

Janzen DH (1967) Why mountain passes are higher in the tropics. *American Naturalist* 101: 233–249.

Kessler M (2009) The impact of population processes on patterns of species richness: Lessons from elevational gradients. *Basic and Applied Ecology* 10: 295–299.

Kluge J, Kessler M, and Dunn RR (2006) What drives elevational patterns of diversity? A test of geometric constraints, climate and species pool effects for pteridophytes on an elevational gradient in Costa Rica. *Global Ecology and Biogeography* 15: 358–371.

Körner C (2000) Why are there global gradients in species richness? Mountains might hold the answer. *Trends in Ecology & Evolution* 15: 513–514.

Kozak KH and Wiens JJ (2010) Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *American Naturalist* 176: 40–54.

Lomolino MV (2001) Elevational gradients of species-density: Historical and prospective views. *Global Ecology & Biogeography* 10: 3–13.

MacArthur RH and MacArthur JW (1961) On bird species diversity. *Ecology* 42: 594–598.

McCain CM (2005) Elevational gradients in diversity of small mammals. *Ecology* 86: 366–372.
McCain CM (2007a) Could temperature and water availability drive elevational diversity? A global case study for bats. *Global Ecology and Biogeography* 16: 1–13.
McCain CM (2007b) Area and mammalian elevational diversity. *Ecology* 88: 76–86.
McCain CM (2009) Global analysis of bird elevational diversity. *Global Ecology and Biogeography* 18: 346–360.
McCain CM (2010) Global analysis of reptile elevational diversity. *Global Ecology and Biogeography* 19: 541–553.
Nogués-Bravo D, Araújo MB, Romdal TS, and Rahbek C (2008) Scale effects and human impact on the elevational species richness gradients. *Nature* 453: 216–220.
Orme CDL, Davies RG, Burgess M, *et al.* (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature* 436: 1016–1019.
Pulliam HR (1988) Sources, sinks, and population regulation. *American Naturalist* 132: 652–661.
Rahbek C (1995) The elevational gradient of species richness: A uniform pattern? *Ecography* 18: 200–205.
Rahbek C (2005) The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters* 8: 224–239.
Shmida A and Wilson MW (1985) Biological determinants of species diversity. *Journal of Biogeography* 12: 1–20.
Stotz DF, Fitzpatrick JW, Parker TA III, and Moskovits DK (1996) *Neotropical Birds: Ecology and Conservation*. Chicago: University of Chicago Press.
Terborgh J (1977) Bird species diversity on an Andean elevational gradient. *Ecology* 58: 1007–1019.
Terborgh J (1985) The role of ecotones in the distribution of Andean birds. *Ecology* 66: 1237–1246.
Vetaas OR and Grytnes JA (2002) Distribution of vascular plant species richness and endemic richness along the Himalayan elevational gradient in Nepal. *Global Ecology and Biogeography* 11: 291–301.
Whittaker RH (1952) A study of summer foliage insect communities in the Great Smoky Mountains. *Ecological Monographs* 22: 1–44.
Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30: 279–338.
Whittaker RH and Niering WA (1965) Vegetation of the Santa Catalina Mountains, Arizona: A gradient analysis of the south slope. *Ecology* 46: 429–452.
Wiens JJ, Ackerly DD, Allen AP, *et al.* (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters* 13: 1310–1324.

El Niño and Biodiversity

Emily Meuser and Arne Ø Mooers, Simon Fraser University, Burnaby, BC, Canada
Daniel FR Cleary, University of Aveiro, Aveiro, Portugal

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bottom-up versus top-down Description of the forces governing overall community structure: availability of nutrients is a bottom-up force and predators are a top-down force. Both may regulate, for example, the amount of standing plant biomass, or whether a lake is clear or murky.
Density dependence A common ecological phenomenon whereby the effect of an entity (e.g., a species' rate of population growth) is a nonlinear function of its density. So, for instance, for the same input of energy, a species may increase in population rapidly at low and intermediate densities, but because of physical interference while foraging, more slowly at higher densities.
Primary productivity The *in situ* transformation of chemical or (more commonly) solar energy into biomass, producing the primary energy source for the rest of the community.
Selection regime A situation in which a particular set of evolved traits is favored. Here, the newly established population evolves in a different direction from the ancestral population.
Teleconnections Atmospheric interactions between widely separated regions that have been identified through statistical correlations. For example, El Niño involves large-scale changes in climatic conditions over the Pacific that are linked to increased winter rainfall in the southwestern United States through a teleconnection.
Trophic inputs Energy entering a community *via* pathways other than local primary production, and often at higher levels in the food chain. For example, nutrients from salmon carcasses are a periodic trophic input to the forest community adjacent to salmon streams.
Upwelling An ocean phenomenon whereby cooler, more nutrient-rich water rises up from the bottom of the ocean to the surface. This occurs where currents of contrasting temperatures meet. The extra nutrient influx allows for specific and rich upwelling ecosystems in these areas (e.g., off the coasts of Chile and Peru, and southwest Africa).

Introduction to El Niño-Southern Oscillation

The El Niño-Southern Oscillation (ENSO) is a climatic and oceanographic phenomenon centered in the Tropical Pacific. Extreme oscillations in ENSO are termed "El Niño" and "La Niña" events, depending on phase (see **Figure 1**) with El Niño generally being the more extreme deviation. Such ENSO events, which have historically occurred every 2–7 years on average, can have dramatic, cascading, and long-lasting effects on terrestrial and marine ecosystems spanning numerous regions worldwide.

Some recent models suggest that predicted global climate change will result in more frequent and more extreme El Niño events (Tudhope *et al.*, 2001), though there are still uncertainties in the exact nature of the predicted changes. Evidence from corals suggests that the amplitude of modern ENSO events is significantly greater now than during any other period in the past 130,000 years (Tudhope *et al.*, 2001), while evidence from Galapagos diatoms indicates that the most recent 50 years have had the warmest sea-surface temperature (SST) of any 50-year period in the past 1200 years for the eastern equatorial Pacific Ocean (Conroy *et al.*, 2009). However, the complexities of modeling the interactive factors contributing to the onset, frequency, and severity of ENSO events currently preclude predictions about the magnitude and direction of future changes to this oscillation (Collins *et al.*, 2010). In recent years, the 1997/1998 record-breaking El Niño followed only 15 years after the previous record El Niño event of 1982/1983. The 1997/1998 El Niño developed so quickly that each month set a new record for warm SST based on records dating back to the middle of the nineteenth century (McPhaden, 1999).

Extreme ENSO events change weather patterns, increasing precipitation in normally dry areas and reducing precipitation in normally wetter areas, as well as increasing SST and reducing nutrient upwelling in the eastern Pacific Ocean. ENSO events have been shown to influence ecosystem functioning (Curran *et al.*, 1999); changes to ENSO may therefore affect the long-term biodiversity of large parts of the planet, especially when it is appreciated that climatic changes associated with both El Niño and La Niña occur in some of the most biodiverse areas of the planet (Holmgren *et al.*, 2006; see **Figure 2(a)** and 2(b)). Here, we survey some of the best-documented effects of ENSO events on biodiversity. We have chosen to organize these effects by the direct ENSO-induced change (fire and drought, increased precipitation, increased SST and changes in ocean currents) rather than the more common biome-based approach.

Expected Effects

If the climatic effects are large enough (e.g., drought-induced fire, *see* Observed Effects: Fire and Drought), ENSO events can be characterized as cyclical disturbance events. If moderate in intensity and infrequent, such disturbances could theoretically increase local diversity over that expected under steady-state conditions (the Intermediate Disturbance Hypothesis; Connell, 1978). So, past ENSO cycles might have encouraged

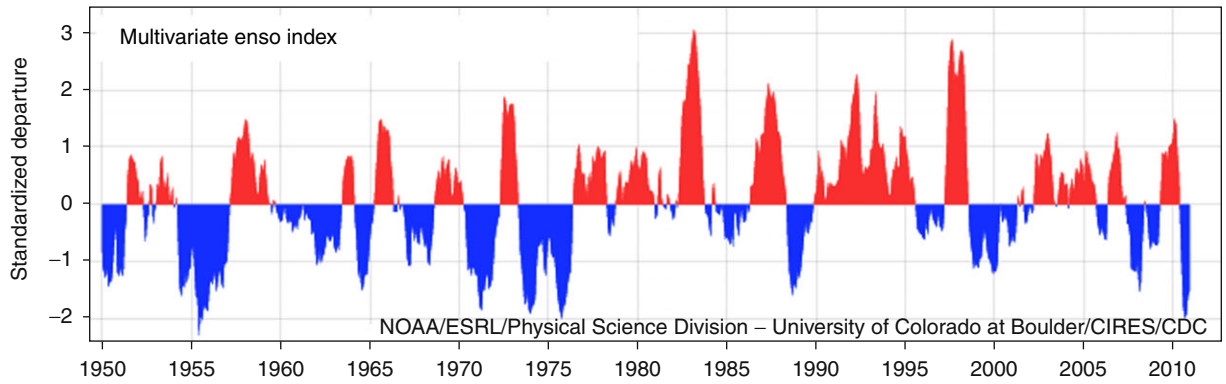


Figure 1 Time series of the Multivariate ENSO Index (MEI) from 1950–2010. The MEI is based on six observed variables over the Tropical Pacific: sea level pressure, surface zonal and meridional wind components, sea-surface temperature, surface air temperature, and cloudiness. Positive values of MEI (red) indicate warm, El Niño conditions, while negative values of MEI (blue) indicate cold, La Niña conditions. For more information about the MEI, see Wolter and Timlin (1993, 1998).

increased local species richness, though retrospective studies are difficult.

Long-lived organisms (e.g., trees) may be adapted to an evolutionary history of repeated, within-lifetime disturbance. Regardless, rapid changes in ENSO intensity can increase environmental variation beyond a species' threshold, causing a die off. If disturbances occur at too high a rate relative to recruitment, successive decimations can produce severe population bottlenecks, increasing the probability of both deterministic and stochastic extinction.

In contrast, short-lived organisms that can evolve within ENSO cycles may be constantly adapting to changing conditions. Microcosm experiments suggest that intermediate temporal variation in environments (e.g., oscillations between one benign and one 'harsh' environment, such as El Niño-induced drought) may increase genetic variation, though with less of an effect than spatial heterogeneity (Kassen, 2002). However, a sharp increase in amplitude and frequency of temporal disturbance can erode genetic variation, with evolution toward either generalist phenotypes that do well under both regimes, or, if the harsh environment is severe enough, specialist phenotypes adapted to the harsh, though rarer, environment (e.g., adapted to El Niño-induced drought years). ENSO events are also likely to have species-specific effects on short-term evolutionary dynamics. For example, on the Galapagos Islands trait evolution of finches changed rapidly in response to the strong El Niño evens of both 1976/1977 and 1982/1983; for instance, drought killed more females than males of one species, leading to interspecific hybridization and subsequent population genetic changes (Grant and Grant, 2002).

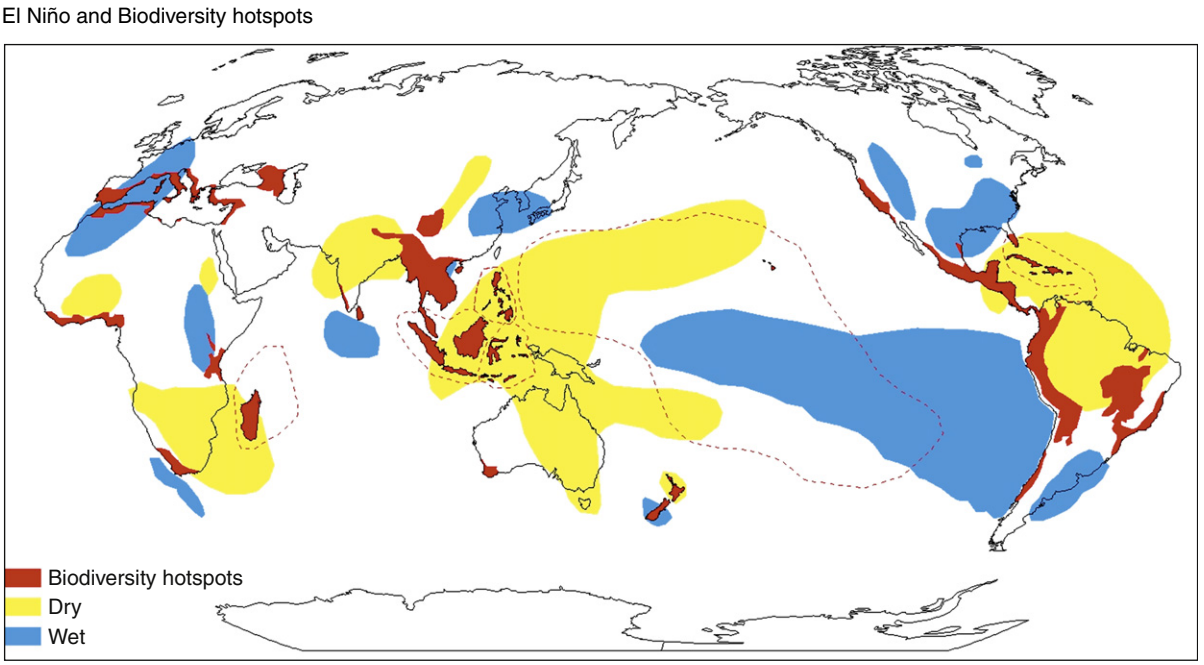
Through changes in current movement and increased SST (*see* Observed Effects: Increased Sea-surface Temperature), regular ENSO events also increase the potential for long-range dispersal. Such migration and gene exchange often decreases geographic genetic structure. However, rarer and more extreme events may cause entirely new habitats to be colonized, especially by species with plankton. This could promote speciation by creating separated populations in novel selection regimes. As well, depending on how they affect the complexity of the new habitats, newly introduced species can either positively or negatively influence species richness in existing communities (Crooks, 2002).

Observed Effects

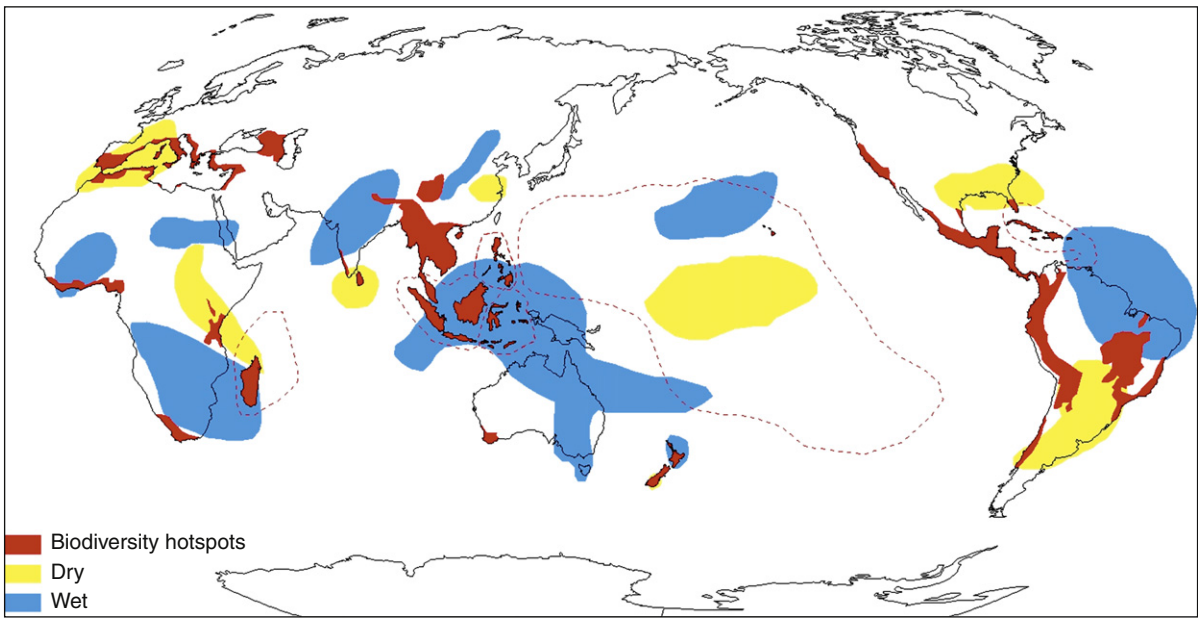
Fire and Drought

Although fires are not new to wet tropical areas, their rate of occurrence has generally been on a scale of hundreds to thousands of years so that many forests and tree species are ill-adapted to frequent large-scale fire events (Cochrane, 2003). The observed and projected increase in ENSO-induced wildfires means that forest fires are now considered to be a major threat to tropical biodiversity (Laurance, 2003).

The Brazilian Amazon and Western Indonesia, particularly eastern Borneo, are two of the areas worst affected by ENSO-induced forest fires; they also harbor among the most diverse floras and faunas in the world. El Niño-associated fires were described for East Kalimantan (Indonesian Borneo) as early as 1914, but until recently these were restricted in scale (Brown, 1998). The first recorded major burning event in East Kalimantan was during the 1982/1983 ENSO event, which caused a conservative 5.5×10^9 US$ of damage, not including ecological costs (Siegert *et al.*, 2001). The 1997/1998 El Niño-induced fires in East Kalimantan were even larger, costing 1.4 billion US$ for smoke-related damage alone. These fires mainly occurred in recently accessed logging concessions (Siegert *et al.*, 2001). In the Amazon, 20,000 km^2 of forest burned during the extreme 1997/1998 El Niño; satellites revealed almost 45,000 individual fires, most of which were sparked by humans (Laurance, 1998). An additional 1.5 million km^2 of forest became susceptible to burning, but did not burn because of insufficient ignition sources (Nepstad *et al.*, 2001). However, frontier expansion is bringing people and thus ignition sources closer to forest tracts. Although the effects are poorly known in the Amazon, an estimated 50% of the remaining forests have already been affected by fires, which have caused more deforestation than intentional clearing in recent years (Cochrane *et al.*, 1999).



(a)



(b)

Figure 2 ENSO and biodiversity. A Robinson projection of the world with areas most greatly affected by the two phases of ENSO, El-Niño (a) and La Niña (b), highlighted (Allan *et al.*, 1996) Superimposed are the dark red terrestrial areas designated as world biodiversity "hotspots" (Myers *et al.*, 2000). The overlap is fairly strong, at least in part because both ENSO effects and species are concentrated in the tropics.

Deforestation fires can smolder for up to a week (**Figure 3**), and can destroy soil-stored seed, reduce the number of resprouting plants, and lead to a longer-term decline in soil fertility due to combustion of organic materials. Repeated burning can deplete the seed supply and cause additional mortality to upper-canopy trees. Typical low-intensity ground fires kill nearly 40% of trees less than 10 cm in diameter at breast height (dbh) and most vines and ground forbs. Mortality is highly species-specific (Slik, 2004). With multiple burns, large fire-contacted trees have no survival advantage over smaller trees (Laurance, 2003). Owing to the lack of a seed bank for climax species in the tropics, their regeneration



Figure 3 Lowland tropical Dipterocarp rainforest, Wanariset, Kalimantan, during the 1997–1998 El Niño fires.

following burn events depends on sapling survival, resprouting and the seed rain from surviving trees. Sapling survival, however, can be very low following burning. In burned forest in East Kalimantan sapling density was only 2.5% of that found in adjacent unburned forest (Cleary *et al.*, 2006b). Resprouting is also highly species-dependent, in part due to changed light conditions (Slik, 2004); whereas a few rainforest species (e.g., *Fordia splendidissima*) are able to resprout profusely after burning, many others fail to resprout at all. Regeneration will therefore, to a large extent, depend on surviving trees and areas of unburned forest that tend to be located in low-lying topography such as forest flood plains (Slik, 2004). Multiple fire events reduce the area of remnant unburned forest because fires tend to become more probable along previously disturbed edges. Under a frequent fire regime, low-diversity stable grasslands can develop. Under a less frequent burn regime, novel pioneer-dominated forests (e.g., dominated by typical pioneer genera such as *Macaranga*, *Mallotus*, or *Trema*) may replace the normal dipterocarp forests, with unknown effects on forest biodiversity.

Primary forests are less susceptible to forest fires than logged and already burned forests because they are more humid (Brown, 1998; Cochrane, 2003) and contain less dead wood (Nepstad *et al.*, 2001). Therefore, during 1982/1983 and 1997/1998, logged forests suffered greater tree mortality than unlogged forests, and fire intensity was directly related to the intensity of logging (Siegert *et al.*, 2001).

Animal responses to fire are still poorly understood. Small mammals, reptiles, and amphibians, for example, were reported to increase in abundance and species richness following forest fires in South America while birds and insects declined (Cochrane, 2003; Barlow *et al.*, 2002; Fredericksen and Fredericksen, 2002). Animals negatively affected by forest fires in Amazonia included slow species with poor climbing ability, cavity-breeding birds, army ants, understory wasp colonies, and leaf-litter invertebrates. Affected forests were also conspicuously silent, generally lacking in the usual vocal activity of monkeys and birds. The birds most affected were uncommon species, habitat specialists, and species that avoided disturbed areas (Laurance, 2003). All of the insectivorous guilds were severely affected by fire, of which the dead-leaf-gleaners and ant-followers were worst affected.

In Borneo, butterfly species richness declined dramatically following forest fires; specialist and restricted range butterfly species recorded before the 1997/1998 ENSO event were also much less likely to return to the regenerating forest than were generalist and wide-ranging species (Charrette *et al.*, 2006). In addition to this, the allelic richness of populations of one species, *Drupadia theda*, that survived the fires, was substantially lower in burned forest and a small unburned isolate than in a large unburned isolate (Cleary *et al.*, 2006a).

In general, while large frugivores and other vertebrates were adversely affected by single burn events in the Amazon, most primary forest specialists across all guilds were extirpated from twice-burnt forest (Barlow and Peres, 2006); we have found the same extreme effects of multiple fires for Bornean butterflies and dragonflies (Cleary *et al.*, 2004).

Even without direct burning, increasing El Niño events could affect tropical forest biodiversity. In Panama, there is a strong relationship between El Niño events, which increase fruiting in rainforest trees, and subsequent famine-induced population changes among frugivorous mammals when the following year is dry (Wright *et al.*, 1999). ENSO-induced drought can also increase tree mortality, particularly in previously disturbed forest (Harrison, 2000; Slik, 2004). In Northern Borneo, El Niño-related drought in a small forest fragment (< 6500 ha) led to mass local extinction of fig wasps, thereby disrupting the important mutualism between fig wasps and their keystone fig hosts (Harrison, 2000). Also in Borneo, we found that butterfly species richness was not only substantially lower in burned forest, but also in small (< 5000 ha) and large (> 100,000 ha) unburned isolates surrounded by burned forest than in continuous forest or forest sampled before the 1997/1998 El Niño (Cleary, 2003).

Increased Precipitation

Water as a Limiting Factor

ENSO-induced increases in precipitation act as resource pulses for primary producers in water-limited ecosystems. The expected effect of such pulses is an immediate increase in ephemeral plant density, with a delayed increase in perennial plant abundance, and herbivore and predator abundance. Over the short term, local biodiversity can change dramatically, as in the Flowering Desert of Copiapó, Chile. For example, normally arid islands in the Gulf of California experience an increase in plant cover from 0–4% in non-El Niño years to 54–89% during El Niño events (Polis *et al.*, 1997). While annual plants account for the majority of ground cover increase, perennial plants also respond to increased water availability by increasing growth, and flower and fruit set production. Seed banks in arid and semiarid regions have been observed to experience a 5–10 fold increase during ENSO-induced wet periods. The species composition of these plant communities also changes during wet years, with normally uncommon or absent species becoming dominant. Thus, the species richness apparent during the common, dry years is not representative of the overall diversity contained in the seed bank (Gutierrez *et al.*, 2000).

As well as such short-term changes in plant communities, ENSO-associated increases in precipitation can also aid in the

recruitment of shrubs and trees in dry areas. Years with extreme, wet ENSO events may act as a window of opportunity for establishment of seedlings, which may then help to create a less-hostile environment for future generations of seedlings (Holmgren *et al.*, 2001). In semiarid ecosystems that have alternate stable states of vegetation (i.e., barren land, degraded savannah, shrubland, or dry forest), pulses of precipitation can play a role in moving the ecosystem from one stable state to another (Holmgren *et al.*, 2006). Interestingly, the process that leads to this type of phase shift may involve the periodic loss of top-down control of plant populations by herbivores. Pulses of precipitation allow plants to grow more rapidly than the numeric response of their herbivores, and this can result in a higher plant biomass equilibrium state, particularly when plants become less palatable with greater size (Scheffer *et al.*, 2008). However, accumulation of plant biomass during wet years can also lead to greater fuel loads, facilitating the ignition and spread of wildfires during contrasting dry years (Block and Richter, 2000).

The effects of increased precipitation can be even more complex. For instance, the dry ecosystems of the Gulf of California islands are influenced both by El Niño-induced influxes of primary productivity and by trophic inputs from the adjacent, highly productive marine ecosystem. Seabirds deposit guano on the islands, which enhances primary productivity increases during El Niño years, and carrion from seabird feeding (i.e., remains of marine organisms) act as a food resource for scavengers more generally (Sanchez-Pinero and Polis, 2000). Omnivorous rodents on these islands obtain marine food resources during dry years, but shift to seed consumption during wet years. So, on these islands, the omnivorous rodent guild experienced a fourfold population increase in response to one El Niño-induced pulse, and a subsequent crash when dry conditions returned. In contrast, granivorous rodents on the same islands experienced a more moderate population increase during the El Niño (1.6 times), probably mediated by food competition with the omnivorous species, but maintained a more stable population after dry conditions returned (Stapp and Polis, 2003). The importance of these direct (food resource) and indirect (nutrients *via* guano) spatial trophic subsidies illustrates the tight linkage between the marine and terrestrial realms in the California Gulf island system. This linkage tends to weaken during El Niño events (Stapp and Polis, 2003); thus, changes in ENSO could have effects on food web dynamics and biodiversity of these islands.

In areas where water is a major limiting factor, the periodic bursts of precipitation associated with ENSO events can exert significant bottom-up effects on multiple trophic levels. The expected effect of rapidly increased primary productivity is a rise in herbivore populations, which in turn positively affect both predator and pathogen populations. For instance, the 1992/1993 El Niño led to increased deer mouse density in New Mexico, US, allowing for a hantavirus outbreak (Holmgren *et al.*, 2006). Carnivorous bird populations respond to the variations in food resources by tracking these changes temporally (Jaksic *et al.*, 1992), and, at the community level, by forming looser guilds with divergence in diet (Jaksic *et al.*, 1993). However, top-down cascades can also be important, and tend to follow bottom-up effects with a delay. During an El Niño event in the Chihuahuan Desert, Arizona, some herbivore species that were previously abundant disappeared, one novel species colonized the area, and other species experienced little change in their populations, while overall diversity remained stable (Brown *et al.*, 1997). This indicates that the net effect on populations is dependent on community-specific interactions within and among trophic levels. In addition, the state of the community prior to the change in precipitation is important in determining the effect of this change on the community. For example, although Galapagos finches generally experience a population boost during wet ENSO events, this boost can be mitigated by density-dependent effects if population levels had previously been high (Grant *et al.*, 2000).

Flooding

Floods cause catastrophic ecosystem change, with immediate abiotic and biotic impacts that may influence community regeneration in the long run. Erosion, changes in salinity, increased sediment in water, oxygen deprivation and habitat disruption can all have effects on biodiversity. For example, seabirds on the Galapagos lost their nests in flooding events during the 1982–1983 El Niño causing reproductive failure for the duration of this wet episode (Duffy, in Glynn, 1990).

The most important ecological effect of flooding is likely its influence on the structure of communities by providing periodic connectivity to ecosystems that are usually isolated. Floods can also maintain wetlands through periodic inundation, provide pulses of nutrients, sediment and species to areas that normally lack water connectivity, and can cause rapid habitat conversion. Rapid water flow through riverbeds can have a scouring effect, redistributing and reducing the abundance of zooplankton (de Paggi and Paggi, 2008). In some cases, resource inputs from flooding can maintain an ecosystem in a nonequilibrium state, or can speed up the process of succession (Ward *et al.*, 2003). Though not otherwise well studied with respect to ENSO, flooding can cause differential mortality of species due to different degrees of flood resistant adaptations. For example, during the 1997/1998 El Niño event, mangrove forests in a Tanzanian wetland experienced species-specific mortality and regeneration due to extensive flooding (Erftemeijer and Hamerlynck, 2005). Whether this short-term variation in survival translates into long-term changes in biodiversity is presently unknown.

Increased Sea-surface Temperature

Increases in SST have both direct and indirect consequences for biological systems. ENSO-related positive SST anomalies correspond with changes in the speed and direction of ocean currents, reductions in nutrient upwelling, and changes in the local environments of marine species.

Ocean Current Shifts

Dispersal of organisms by ocean currents is an important process in establishing and maintaining biodiversity in aquatic ecosystems. For example, the Hawaiian biota comprises a subset of Polynesian biota, and this similarity can be

attributed to movement of organisms on currents associated with El Niño events (Richmond, in Glynn, 1990).

El Niño-induced changes in ocean currents may lead to long distance dispersal of marine organisms from west to east. Genetic similarity between shallow water populations separated by 5400 km of ocean in the eastern Pacific (known as the Eastern Pacific Barrier, EPB) indicates that gene flow occurs between these populations (Lessios *et al.*, 1998). During non-El Niño years, parcels of water carried by the North Equatorial Counter-Current take 100–155 days to cross the EPB; in El Niño years this is shortened to 50–81 days. Given the known maximum time that larvae can stay in the plankton and successfully settle, the number of species able to colonize a new region may more than double under El Niño conditions (Richmond, in Glynn, 1990).

Nutrient Level Shifts

During non-El Niño years, the eastern equatorial Pacific Ocean comprises a large upwelling ecosystem, where the relatively shallow thermocline allows nutrient-rich cool water to rise along the west coast of South America. This tongue of cool water extends as far west as the International Date Line and along the coast of Central and South America between $\sim 10^{\circ}$ N and $\sim 20^{\circ}$ S. During ENSO events, upwelling decreases, and the area of cool, productive water is reduced to a pocket $\sim 10\%$ of its usual size along the South American coast. ENSO-associated reductions in upwelling also occur in the near-coastal southeast Atlantic Ocean through teleconnections. As nutrient upwelling is the first-order process governing ocean primary productivity, the effects of reduced upwelling have significant impacts on both marine and terrestrial biota. These impacts can be seen along the equator in the eastern Pacific, where primary productivity can fall to as low as 6% of normal during El Niño years, while in the remaining productive region primary productivity falls to 20–50% of normal (Barber and Kogelshatz, in Glynn, 1990). Such El Niño-induced changes impact heavily on, for example, fish and zooplankton-grazer species. The well-known short and longer-term fluctuations of anchoveta and sardine in the Pacific (Chavez *et al.*, 2003) highlight how species-specific traits are important in determining response to climate fluctuations.

Declines in small fish and other food source populations have bottom-up effects on higher trophic levels. For example, Christmas Island's great frigate bird population fell from 20,000 to fewer than 100 over the course of six months following the extreme 1982/1983 ENSO. While nest flooding and heavy rains may have caused some nestling death, adult abandonment of the island along with their young has been attributed to the disappearance of fish and squid as food resources. Also, multi-decadal records of species diversity and abundance in the Galapagos suggest that many marine species have yet to recover from the effects of the 1982/1983 El Niño event.

In upwelling ecosystems, water temperature and availability of nutrients are tightly linked. Therefore, seabirds, fish, and marine mammals within these ecosystems have evolved behavioral adaptations that enable these species to use thermal cues as indicators of areas of abundant food. Short-term responses of species to changes in productivity may then be mediated through responses to anomalous SST during ENSO events. This results in a concentration of organisms residing in those areas where upwelling still occurs during El Niño events.

Coral Bleaching

Coral reefs are among the most diverse and productive ecosystems known and are one of the first to be severely affected by global climate change. The marked and rapid response of corals to global climate change is largely a result of the pronounced thermal sensitivity of most coral species (Graham *et al.*, 2006). Corals' physiological tolerance ranges between 18 °C and 30 °C, and many reefs exist at close to their upper temperature threshold. Against the baseline of rising SST associated with global climate change, the most catastrophic effects of climate change on coral reef communities are predicted to be through increasing the magnitude and frequency of extreme, periodic climate oscillations such as ENSO events (Reaser *et al.*, 2000). Increases in SST and solar irradiance have been heavily implicated in the widespread bleaching (loss of symbiotic zooxanthellae) and mortality of reef-building corals during ENSO events, and these effects are compounded when reefs already have reduced resilience due to other, often human-induced disturbance (e.g., cyanide poisoning, nutrient enrichment; Zhu *et al.*, 2004).

Over the course of a severe episode, corals may lose 60% to $>90\%$ of their symbiotic algae, and the remaining algae may lose 50–80% of their photosynthetic pigments (Glynn, 1996; **Figure 4**). Depending on the severity and duration of the bleaching event, corals may regain their obligate symbionts with the return of favorable conditions. Alternatively, individuals or entire assemblages may die. Coral mortality corresponds with an often-dramatic loss of coral species and reef cover of live corals. For example, two out of 12 coral species were virtually eliminated from Panama during the 1982/1983 ENSO (Glynn and Feingold, 1992). Also, coral mortality leads to loss of reef structural complexity due to reef disintegration. In addition to the initial loss of diversity, species such as the crown of thorns starfish are able to exploit remaining corals more effectively, at least in part due to

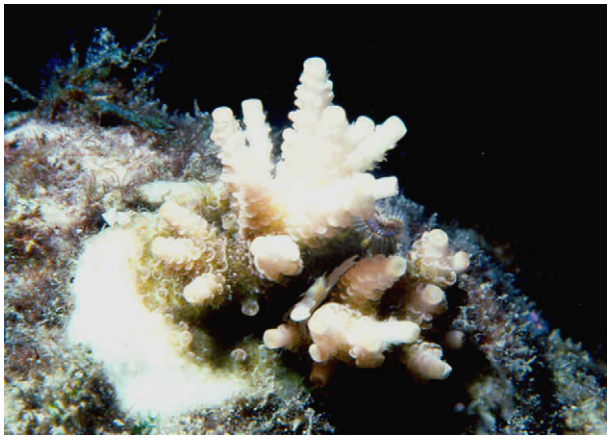


Figure 4 Example of El Niño-induced coral bleaching. Photo of a juvenile staghorn coral (Acropora sp.) at Pulau Pari, Thousand Islands, Indonesia, taken during the 1983/1984 El Niño event. Bleaching has started at the tips and at the periphery of the coral base (see Hoeksema 1991).

the corals' crustacean guards becoming less aggressive with the deterioration of their hosts (Glynn, in Glynn, 1990). Bioerosion and larval recruitment failure on damaged reefs also contribute to changes in species composition. On death, many corals are colonized by algae, which prevents coral regeneration and transforms these areas into non-reef building communities (McManus and Polsenberg, 2004). As coral reefs play host to some of the world's greatest biodiversity, loss of these habitats could have profound effects on that diversity. For example, in the Galapagos Islands direct or interactive El Niño effects, often mediated by coral reef deterioration, were recognized as the major threatening process for 29 of 45 species recognized as being at risk by the IUCN (Edgar *et al.*, 2010).

During the 1982/1983 El Niño, tropical eastern Pacific coral populations experienced up to 70–90% mortality; in the Galapagos Islands, mortality of most reefs was ~95%. Decimation of these populations has particularly profound effects in this region, where isolation from other reefs that could otherwise aid in repopulation prolongs recovery time. In coral reefs, the 1997/1998 bleaching event was the most severe and widespread ever recorded (Nyström *et al.*, 2000; Wilkinson, 2004). Some scientists fear that a threshold may already have been reached where reefs will not be able to cope with the increasing intensity and frequency of future bleaching events associated with ENSO-induced disturbances (Nyström *et al.*, 2000). McWilliams *et al.* (2005) reported exponential increases in the extent and intensity of bleaching with increasing SST anomalies in the Caribbean; a rise of only 0.1 °C resulted in 35% and 42% increases in the geographic extent and intensity of coral bleaching, respectively. Maxima of both are predicted to occur at SST anomalies of less than +1 °C; this would coincide with the most conservative projection for warming in the Caribbean by the end of the century. Bleaching will, therefore, probably be a chronic stress factor for Caribbean coral reefs in the future. The loss of live corals in turn has had a number of cascading effects on other taxa. Long-term censuses of fish communities in Australia spanning bleaching events have shown species loss and distinct phase shifts in composition from pre- to postbleaching with no evidence of regeneration (Bellwood *et al.*, 2006; Garpe *et al.*, 2006). However, and interestingly, some evidence suggests that strong directional selection by frequent ENSO-induced thermal stress may lead to coral reef communities that experience less severe bleaching during subsequent ENSO events. While coral reef communities may become more resistant to thermal stress, the differential survivorship of corals and algal symbionts that leads to this resistance will likely also lead to large changes in species composition of coral reefs (Thompson and van Woesik, 2009).

Feedback and Conclusions

Interestingly, many of the ecological changes caused by extreme El Niño events may be self-reinforcing (see **Figure 5**).

Predicted changes in the frequency and severity of ENSO events have the potential to greatly alter biodiversity and ecosystem functioning (Harrison, 2000; Holmgren *et al.*, 2001; Holmgren *et al.*, 2006). In particular, severe ENSO-induced droughts can cause devastating forest fires with profound effects on ecosystem dynamics (Holmgren *et al.*, 2001). Although fires have always been present in Southeast Asia and in much of the world, population growth, habitat fragmentation, logging, changes in land-use, ENSO events, and the ubiquity of human ignition sources have increased the probability of catastrophic fires and are responsible for the catastrophic fires in 1982/1983 and 1997/1998 (Laurance, 1998; Laurance, 2003). In addition, both logging and previous fires increase the probability of future fires (Nepstad *et al.*, 1999), which are often exacerbated by build up of fuel from dead wood. Secondary fires are, therefore, much worse than initial fires (Cochrane *et al.*, 1999), leading to positive feedback.

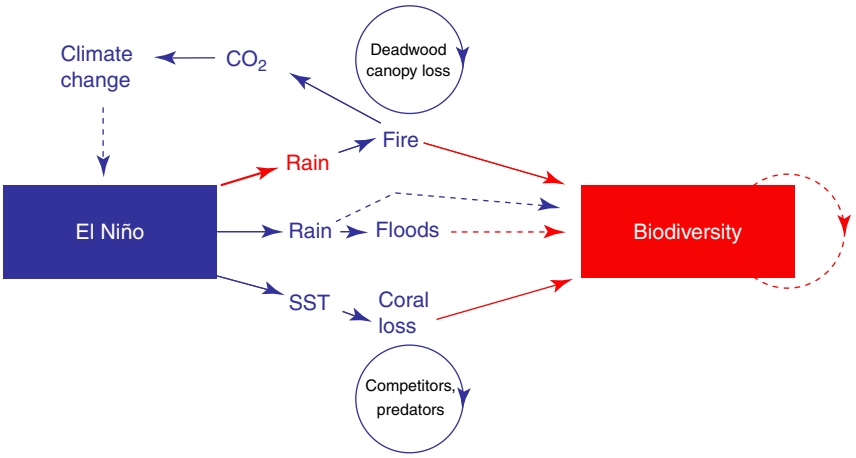


Figure 5 El-Niño feedbacks on biodiversity. Blue arrows indicate increases, and red lines indicate decreases. Solid arrows are well-established connections, and dotted arrows are more speculative. So, El Niño causes decreased rainfall, which causes fires. The fires increase the severity of future fires, and these lead to a loss in biodiversity. Such fires also contribute CO_2 to the atmosphere, and if global climate change is linked to increased severity or frequency of ENSO-events, the feedback is positive. Likewise, increased sea-surface temperature (SST) is implicated in coral bleaching, which decreases local biodiversity as the community shifts. Because of trophic cascades and loss of stability, losses of biodiversity may also be reinforcing.

Ironically, these fires may also be contributing to greater global warming, which may, in turn, feed back into the ENSO cycle (Fedorov and Philander, 2000). Because of their scale of occurrence, fires now have the potential to more than double carbon emission in dry years (Malhi and Grace, 2000) and fires may have released the carbon equivalent of *ca.* 41% of world fossil fuel use in 1997/1998 (Cochrane, 2003). It is estimated that this, combined with other changes, has switched the tropics from a net carbon sink (2.0 Pg C $year^{-1}$) to a net source (0.4 Pg C $year^{-1}$) (Malhi and Grace, 2000). The release of large amounts of aerosols to the atmosphere can, furthermore, affect atmospheric stability and cloud formation and thereby reduce rainfall over large areas (Cochrane, 2003), thus compounding the risk of fire in these regions.

Conversely, increased precipitation in normally arid or semiarid regions can drive the restoration of degraded scrubland (Holmgren *et al.*, 2006). In areas that have been subjected to overexploitation, episodes of relatively extreme moisture in conjunction with decreased grazing pressure can lead the ecosystem to move to an alternative stable state of vegetation. In this manner, ENSO events may be partnered with land management strategies, with a favorable outcome for biodiversity in these regions.

ENSO is one climatic cycle among many that occur over various temporal and spatial scales. It is a natural phenomenon, and has been functioning for at least the past 130,000 years (Tudhope *et al.*, 2001). However, human activity may be causing ENSO events to become more frequent and extreme. As well, anthropogenic factors such as overharvesting, pollution, fragmentation, and edge effects not only worsen the direct effects of ENSO but also compromise the ability of ecosystems to maintain or re-establish levels of biodiversity following disturbance. There is little doubt that ENSO-induced disturbance, at least that mediated by human activity, has strong effects on biodiversity at many spatial and temporal scales, and the evidence suggests that the bulk of these effects are undesirable.

Acknowledgments

We thank M. Holmgren, I. Côté and an anonymous reviewer for constructive comments, B. Hoeksema for **Figure 4**, D. Austin for creating **Figure 1**, and the Dutch and Canadian National Science Foundations and the Institute for Advanced Study, Berlin for support.

See also: Biodiversity in Logged and Managed Forests. Climate Change and Ecology, Synergism of. Climate Change and Extinctions. Climate Change and Wild Species. Climate, Effects of. Deforestation and Land Clearing. Dispersal Biogeography. Fires, Ecological Effects of. Hotspots. Impact of Past Global Warming on Biodiversity. Mediterranean-Climate Ecosystems. Rainforest Ecosystems, Animal Diversity. Rainforest Ecosystems, Plant Diversity. Rainforest Loss and Change

References

Allan R, Lindesay J, and Parker D (1996) *El Niño Southern Oscillation and Climatic Variability.* Collingwood: CSIRO.

Barlow J, Haugaasen T, and Peres CA (2002) Effects of ground fires on understorey bird assemblages in Amazonian forests. *Biological Conservation* 105: 157–169.

Barlow J and Peres CA (2006) Effects of single and recurrent wildfires on fruit production and large vertebrate abundance in a central Amazonian forest. *Biodiversity and Conservation* 15: 985–1012.

Bellwood DR, Hoey AS, Ackerman JL, and Depczynski M (2006) Coral bleaching, reef fish community phase shifts and the resilience of coral reefs. *Global Change Biology* 12: 1587–1594.

Block M and Richter M (2000) Impacts of heavy rainfalls in El Niño 1997/98 on the vegetation of Sechura Desert in Northern Peru. *Phytocoenologia* 30: 491–517.

Brown JH, Valone TJ, and Curtin CG (1997) Reorganization of an arid ecosystem in response to recent climate change. *Proceedings of the National Academy of Sciences of the United States of America* 94: 9729–9733.

Brown N (1998) Out of control: Fires and forestry in Indonesia. *Trends in Ecology & Evolution* 13: 41–51.

Charrette NA, Cleary DFR, and Mooers AØ (2006) Range-restricted, specialist Bornean butterflies are less likely to recover from ENSO-induced disturbance. *Ecology* 87: 2330–2337.

Chavez FP, Ryan J, Lluch-Cota SE, and Niquen M (2003) From anchovies to sardines and back: Multidecadal change in the Pacific Ocean. *Science* 299: 217–221.

Cleary DFR (2003) An examination of scale of assessment, logging and ENSO-induced fires on butterfly diversity in Borneo. *Oecologia* 135: 313–321.

Cleary DFR, Fauvelot C, Genner MJ, Menken SBJ, and Mooers AØ (2006a) Parallel responses of species and genetic diversity to El Niño Southern Oscillation-induced environmental destruction. *Ecology Letters* 9: 301–307.

Cleary DFR, Mooers AØ, Eichhorn KAO, van Tol J, de Jong R, and Menken SBJ (2004) Diversity and community composition of butterflies and odonates in an ENSO-induced fire affected habitat mosaic: A case study from East Kalimantan, Indonesia. *Oikos* 105: 426–446.

Cleary DFR, Priadjati A, Suryokusumo BK, and Menken SBJ (2006b) Butterfly, sapling, seedling and tree diversity in a fire-affected Bornean rainforest. *Austral Ecology* 31: 46–57.

Cochrane MA (2003) Fire science for rainforests. *Nature* 421: 913–919.

Cochrane MA, Alencar A, Schulze MD, *et al.* (1999) Positive feedbacks in the fire dynamic of closed canopy tropical forests. *Science* 284: 1832–1835.

Collins M, Ans S-I, and Cai W (2010) The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geoscience* 3: 391–397.

Conroy JL, Restrepo A, Overpeck JT, *et al.* (2009) Unprecedented recent warming of surface temperatures in the eastern tropical Pacific Ocean. *Nature Geoscience* 2: 46–50.

Connell JH (1978) Diversity in tropical rain forests and coral reefs - high diversity of trees and corals is maintained only in a non-equilibrium state. *Science* 199: 1302–1310.

Crooks JA (2002) Characterizing ecosystem-level consequences of biological invasions: The role of ecosystem engineers. *Oikos* 97: 153–166.

Curran LM, Caniago I, Paoli GD, *et al.* (1999) Impact of El Niño and logging on canopy tree recruitment in Borneo. *Science* 286: 2184–2188.

Edgar GJ, Banks SA, Brandt M, *et al.* (2010) El Niño, grazers and fisheries interact to greatly elevate extinction risk for Galapagos marine species. *Global Change Biology* 16: 2876–2890.

Erftemeijer PLA and Hamerlynck O (2005) Die-back of the mangrove Heritiera littoralis dryland, in the Rufiji Delta (Tanzania) following El Niño floods. *Journal of Coastal Research* 42: 228–235.

Fedorov AV and Philander SG (2000) Is El Niño changing? *Science* 288: 1997–2002.

Fredericksen NJ and Fredericksen TS (2002) Terrestrial wildlife responses to logging and fire in a Bolivian tropical humid forest. *Biodiversity and Conservation* 11: 27–38.

Garpe KC, Yahya SAS, Lindahl U, and Ohman MC (2006) Long-term effects of the 1998 coral bleaching event on reef fish assemblages. *Marine Ecology Progress Series* 315: 237–247.

Glynn PW (ed.) (1990) *Global ecological consequences of the 1982-83 El Niño-Southern Oscillation.* Amsterdam: Elsevier.

Glynn PW (1996) Coral reef bleaching: Facts, hypotheses and implications. *Global Change Biology* 2: 495–509.

Glynn PW and Feingold JS (1992) Hydrocoral species not extinct. *Science* 257: 1845.
Graham NAJ, Wilson SK, Jennings S, *et al.* (2006) Dynamic fragility of oceanic coral reef ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 103: 8425–8429.
Grant PR and Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296: 707–711.
Grant PR, Grant BR, Keller LF, and Petren K (2000) Effects of El Niño events on Darwin's finch productivity. *Ecology* 81: 2442–2457.
Gutierrez JR, Arancio G, and Jaksic FM (2000) Variation in vegetation and seed bank in a Chilean semi-arid community affected by ENSO 1997. *Journal of Vegetation Science* 11: 641–648.
Harrison RD (2000) Repercussions of El Niño: Drought causes extinction and the breakdown of mutualism in Borneo. *Proceedings of the Royal Society of London Series B-Biological Sciences* 267: 911–915.
Hoeksema BW (1991) Control of bleaching in mushroom coral populations (Scleractinia: Fungiidae) in the Java Sea: Stress tolerance and interference by life history strategy. *Marine Ecology Progress Series* 74: 225–237.
Holmgren M, Scheffer M, Ezcurra E, Gutierrez JR, and Mohren GMJ (2001) El Niño effects on the dynamics of terrestrial ecosystems. *Trends in Ecology & Evolution* 16: 89–94.
Holmgren M, Stapp P, Dickman CR, *et al.* (2006) Extreme climatic events shape arid and semiarid ecosystems. *Frontiers in Ecology and the Environment* 4: 87–95.
Jaksic FM, Feinsinger P, and Jimenez JE (1993) A long-term study on the dynamics of guild structure among predatory vertebrates at a semi-arid neotropical site. *Oikos* 67: 87–96.
Jaksic FM, Jiménez JE, Castro SA, and Feinsinger P (1992) Numerical and functional response of predators to a long-term decline in mammalian prey at a semi-arid neotropical site. *Oecologia* 89: 90–101.
Kassen R (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology* 15: 173–190.
Laurance WF (1998) A crisis in the making: Responses of Amazonian forests to land use and climate change. *Trends in Ecology & Evolution* 13: 411–415.
Laurance WF (2003) Slow burn: The insidious effects of surface fires on tropical forests. *Trends in Ecology & Evolution* 18: 209–212.
Lessios HA, Kessing BD, and Robertson DR (1998) Massive gene flow across the world's most potent marine biogeographic barrier. *Proceedings of the Royal Society of London Series B-Biological Sciences* 265: 583–588.
Malhi Y and Grace J (2000) Tropical forests and atmospheric carbon dioxide. *Trends in Ecology & Evolution* 15: 332–337.
McManus JW and Polsenberg JF (2004) Coral-algal phase shifts on coral reefs: Ecological and environmental aspects. *Progress in Oceanography* 60: 263–279.
McPhaden MJ (1999) Climate oscillations – Genesis and evolution of the 1997–98 El Niño. *Science* 283: 950–954.
McWilliams JP, Cote IM, Gill JA, Sutherland WJ, and Watkinson AR (2005) Accelerating impacts of temperature-induced coral bleaching in the Caribbean. *Ecology* 86: 2055–2060.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
Nepstad D, Carvalho G, Barros AC, *et al.* (2001) Road paving, fire regime feedbacks, and the future of Amazon forests. *Forest Ecology and Management* 154: 395–407.
Nepstad DC, Verissimo A, Alencar A, *et al.* (1999) Large-scale impoverishment of Amazonian forests by logging and fire. *Nature* 398: 505–508.
Nyström M, Folke C, and Moberg F (2000) Coral reef disturbance and resilience in a human-dominated environment. *Trends in Ecology & Evolution* 15: 413–417.
de Paggi SBJ and Paggi JC (2008) Hydrological connectivity as a shaping force in the zooplankton community of two lakes in the Paraná River floodplain. *International Review of Hydrobiology* 93: 659–678.
Polis GA, Hurd SD, Jackson CT, and Pinero FS (1997) El Niño effects on the dynamics and control of an island ecosystem in the Gulf of California. *Ecology* 78: 1884–1897.
Reaser JK, Pomerance R, and Thomas PO (2000) Coral bleaching and global climate change: Scientific findings and policy recommendations. *Conservation Biology* 14: 1500–1511.
Sanchez-Pinero F and Polis GA (2000) Bottom-up dynamics of allochthonous input: Direct and indirect effects of seabirds on islands. *Ecology* 81: 3117–3132.
Scheffer M, van Nes EH, Holmgren M, and Hughes T (2008) Pulse-driven loss of top-down control: The critical-rate hypothesis. *Ecosystems* 11: 226–237.
Siegert F, Ruecker G, Hinrichs A, and Hoffmann AA (2001) Increased damage from fires in logged forests during droughts caused by El Niño, *Nature* 414: 437–440.
Slik JWF (2004) El Niño droughts and their effects on tree species composition and diversity in tropical rain forests. *Oecologia* 141: 114–120.
Stapp P and Polis GA (2003) Influence of pulsed resources and marine subsidies on insular rodent populations. *Oikos* 102: 111–123.
Thompson DM and van Woesik R (2009) Corals escape bleaching in regions that recently and historically experienced frequent thermal stress. *Proceedings of the Royal Society B Biological Sciences* 276: 2893–2901.
Tudhope AW, Chilcott CP, McCulloch MT, *et al.* (2001) Variability in the El Niño - Southern oscillation through a glacial-interglacial cycle. *Science* 291: 1511–1517.
Ward KM, Callaway JC, and Zedler JB (2003) Episodic colonization of an intertidal mudflat by native cordgrass (Spartina foliosa) at Tijuana estuary. *Estuaries* 26: 116–130.
Wilkinson C (2004) *Status of Coral Reefs of the World*. Townsville: Australian Institute of Marine Science.
Wolter K and Timlin MS (1993) Monitoring ENSO in COADS with a seasonally adjusted principal component index. *Proceedings of the 17th Climate Diagnostics Workshop*. Norman, OK: NOAA/NMC/CAC, NSSL, Oklahoma Climate Survey, CIMMS and the School of Meteorology, University of Oklahoma, pp. 52–57.
Wolter K and Timlin MS (1998) Measuring the strength of ENSO events - How does 1997/98 rank? *Weather* 53: 315–324.
Wright SJ, Carrasco C, Calderon O, and Paton S (1999) The El Niño Southern Oscillation, variable fruit production, and famine in a tropical forest. *Ecology* 80: 1632–1647.
Zhu BH, Wang GC, Huang B, and Tseng CK (2004) Effects of temperature, hypoxia, ammonia and nitrate on the bleaching among three coral species. *Chinese Science Bulletin* 49: 1923–1928.

Endangered Amphibians

Tim Halliday, Open University, Milton Keynes, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Amphibian Member of a class of vertebrates (the Amphibia), comprising frogs and toads (order Anura), newts and salamanders (order Caudata), and caecilians (order Gymnophiona), which typically return to water to breed and pass through an aquatic larval stage with gills. Amphibians have a moist skin without scales, which is permeable to water and gases.

Amphibians

Amphibians are a group of vertebrate animals that are descended from fishes and which, like fishes, are ectothermic, meaning that the heat that they require to maintain physiological processes is derived externally, directly, or indirectly from the sun. Their dependence on external heat sources limits amphibians to tropical and temperate regions of the Earth and they are especially numerous in the tropics; 80% of the world's amphibian species live in tropical habitats.

Amphibians have a complex life cycle, in which eggs, typically laid in water, hatch into larvae (called tadpoles in frogs and toads), which grow rapidly before undergoing a complex metamorphosis into the adult form. Adult amphibians have a thin, permeable skin that severely limits their ability to retain water within the body. Furthermore, the eggs of amphibians are bounded by only a thin membrane and can only survive if they are kept wet and amphibian larvae are typically aquatic. Consequently, amphibians are restricted, with very few exceptions, to those habitats in which water is available for all or most of the year.

Amphibians are not as well-known, in terms of documentation of species, as either the birds or the mammals. They have been studied less intensively than birds and mammals, are rather secretive in their habits, and are most numerous in parts of the world, notably tropical forests, that have not been fully explored. Consequently, new species of amphibians are being discovered and described at a much higher rate than is the case for birds or mammals. The number of recognized species of amphibians is approximately 6800, but this figure is continuously being revised upward as new species are described.

There are three major groups of amphibians: the frogs and toads (approximately 6000 species), the salamanders and newts (approximately 600 species), and the caecilians (approximately 200 species).

The Conservation Status of Amphibians

Amphibians are threatened worldwide. A recent assessment of the entire group found that nearly one-third (32%) of the world's amphibian species are threatened with extinction, representing 1856 species. Amphibians have existed on Earth for a very long time, about 300 million years, yet over the past few decades nearly 168 species are believed to have gone extinct and at least 2469 (43%) species are declining in population size, indicating that the number of threatened and extinct species will probably continue to rise. **Table 1** presents the results of the Global Amphibian Assessment, a major international effort to determine the conservation status of most of the world's amphibian species.

Threats to Amphibians

Many environmental factors have been identified that are detrimental to the continued survival of amphibian species. Most detailed studies of amphibian population declines have identified more than one factor as a cause of decline and, in many instances, multiple factors act synergistically (Blaustein and Kiesecker, 2002).

Habitat Destruction

The single most important factor that adversely affects amphibians is habitat destruction and modification. Many habitats that are essential for eggs, larvae, or adult stages are shrinking or disappearing at an accelerating rate as a result of human population growth and economic development. Tropical forests provide some of the most species-rich habitats in the world and are particularly vulnerable to destruction by humans, with land being turned over to agriculture or to provide residential areas for people. In 1991, the United Nations Food and Agriculture Organization reported that the

Table 1 Number of amphibian species in each of the IUCN categories

Conservation status	*Number of species*	*% of species*
Data-deficient (DD)	1294	22.5
Least concern (LC)	2203	38.4
Not threatened (NT)	359	6.3
Vulnerable (VU)	668	11.6
Endangered (EN)	761	13.3
Critically endangered (CR)	427	7.4
Extinct (EX)	35	0.6
Total	5743	

The three categories VU, EN, and CR are grouped together as Threatened (T); this group consists of 1856 species, representing 32% of all known amphibian species. *Source*: Reproduced from Stuart SN, Chanson JS, and Cox NA (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00278-1

world's tropical forests were being destroyed at a rate that was 50% faster than a decade previously. Tropical forests support a high diversity of amphibian species, many of them still to be described. At the current rate of deforestation, within 30 years all extensive areas of tropical forest, together with their amphibian fauna, will have disappeared.

Amphibians are particularly dependent on freshwater habitats, especially for breeding. The World Wildlife Fund publishes regular reports which suggest that freshwater habitats are the most seriously threatened on Earth. Between 1970 and 2007, the global diversity of freshwater species decreased by 35%, that of tropical freshwater species by almost 70%, a faster rate of decline than has been detected in any other component of the global ecosystem (Pollard, 2010). Although there is legal protection for some of the world's larger lakes and rivers, largely because of their importance in providing fish to support the human population, the smaller streams, ponds, and swamps that are essential habitat for amphibians are generally not protected. The draining of wetlands to make way for agriculture and housing has had a particularly serious negative effect on many amphibian populations.

In much of western Europe, traditional methods of agriculture previously provided good habitat for amphibians in the form of small woodlands, hedgerows, and numerous ponds created to provide water for livestock. In the past 50 years, however, agricultural practices have changed and all these landscape features have been destroyed over very large areas. In parts of Britain, for example, the number of ponds suitable for amphibians to breed in has declined by 90% in the past 50 years.

Habitat destruction in a given area is not always total; often, small pockets of forest, heathland, or wetland are set aside for conservation purposes. This results in the fragmentation of previous areas of habitat and there is increasing evidence that habitat fragmentation is a serious threat to the continued survival of species that it was assumed were afforded some degree of protection. Fragmentation of habitat leads to the isolation of small populations of amphibian species. Inbreeding in such populations reduces their genetic diversity, and their isolation prevents interchange of individuals and thus of genetic diversity, and their isolation prevents interchange of individuals and thus of genetic variation with other populations. As a result, isolated populations tend to decline slowly and eventually die out, even though they are protected. Many amphibians have only limited powers of dispersal, and even a road built through an area of otherwise suitable habitat will reduce dispersal.

Climate Change

There is increasing evidence that the earth's climate is undergoing major changes as a result of human activities such as the destruction of forests. Most notably, average temperatures are steadily increasing in many parts of the world and there are major changes in rainfall patterns. The long-term effects of such changes on amphibians are largely a matter of speculation, but there is evidence that they have had an impact on some species. For many amphibians, rainfall patterns are critical because they determine whether, and for how long, the small streams and ponds in which they breed contain water.

The apparent extinction of the golden toad (*Bufo periglenes*), along with several other frog species, in the montane forest of Costa Rica appears to be due, at least in part, to climate change (Pounds *et al.*, 1999). A critical factor in the habitat of this species is low-lying cloud, which provides the water that maintains water flow in the small streams in which many of the native frogs breed. The extent of such cloud has decreased in recent years, leading to extensive reductions in the amount of available water and thus a general reduction in stream habitat. In Britain, winters are becoming increasingly less severe and several native amphibians are now breeding earlier in the spring than they were 20 years ago. There is also evidence that common toads (*Bufo bufo*) in Britain are in poorer body condition, and less fecund, following milder winters (Reading, 2007).

Ultraviolet Radiation

The steady erosion of the ozone layer in the earth's stratosphere has led to an increase in the amount of ultraviolet (UV) radiation reaching the Earth's surface. Such radiation, especially UV-B, is harmful to living organisms and there is accumulating evidence that it may be a factor in the decline of some amphibian species, especially those living at high altitudes, where incident levels of UV-B are highest. Increased UV-B causes the genetic material DNA to mutate, leading to the abnormal development and eventual death of embryos. Frog and toad species that lay their eggs close to the surface of water, which filters out UV-B, may be particularly susceptible to increased UV-B. Although there is considerable experimental evidence that ambient levels of UV-B radiation are harmful to the early life stages of amphibians (Blaustein *et al.*, 1995) it is not clear to what extent elevated UV-B has been a factor in the decline of amphibian populations in nature (Alford and Richards, 1999).

Amphibian species vary in their susceptibility to elevated UV-B. Some species, such as the Pacific tree frog (*Hyla regilla*), produce high levels of the enzyme photolyase, which repairs DNA damage of the kind caused by UV-B. This species is one of those that has not declined in the Pacific Northwest of North America, unlike species with lower photolyase levels, such as the Cascades frog (*Rana cascadae*) and the western toad (*Bufo boreas*) (Blaustein *et al.*, 1994). If elevated UV-B does affect natural populations of amphibians, it is likely that it acts synergistically with other adverse environmental factors, such as acidification and pathogens. There is experimental evidence that both these factors have a more detrimental effect on the survival of amphibian embryos when they are combined with elevated UV-B.

Chemical Pollution

Environmental pollution has caused the decline or extinction of some local populations of amphibians, and it is likely that it also has a widespread harmful effect. For example, declines of amphibians in Yosemite National Park in California appear to be due to chemical pollution that has drifted on the wind

from agricultural areas many miles away. Amphibians may provide sensitive biological indicators of pollution because their highly permeable skin rapidly absorbs toxic substances. Examples of pollutants include fertilizers, herbicides, pesticides, heavy metals, and poisoning resulting from logging and mining operations.

A major form of pollution that can affect very large areas is atmospheric acid deposition or acid rain. Much of the rain that falls in regions downwind from major industrial areas, such as in the eastern United States, Scandinavia, and western Europe, is markedly acidic, with a pH of approximately 4.5; unpolluted rain has a pH of approximately 5.6. Acid rain lowers the pH of natural water bodies below the tolerance level of many species of amphibians. The toxic effects of low pH on amphibian development are well-documented. Acidic conditions reduce the mobility of sperm and may cause them to disintegrate, with a consequent reduction in the fertilization success of eggs. Eggs that are fertilized may develop abnormally; if they hatch, they produce deformed tadpoles that soon die. Acid precipitation has been implicated in the declines of tiger salamanders (*Ambystoma tigrinum*) in the Rocky Mountains of Colorado and of the natterjack toad (*Bufo calamita*) in lowland heaths in Britain.

Some chemical compounds of human origin are readily absorbed by animals and interfere with their endocrine systems; these are known as endocrine disrupters and include estrogen mimics, which are compounds that disrupt the reproductive development of both sexes (Stebbins and Cohen, 1995). Males may become feminized to varying degrees, for example, suffering lowered sperm counts and developing ovarian tissue in their testes.

Many pesticides, such as DDT, are not only poisonous to a wide range of animals but also, at low concentrations, have endocrine-disrupting effects. Populations of amphibians in Africa and the USA have been adversely affected by the widely used herbicide atrazine because of its feminizing effects on males (Hayes *et al.*, 2010). Many of these compounds are highly volatile and so are transported very widely in the environment from their point of origin. The long-term effects of widespread chemical pollution of this kind are poorly understood but are very alarming not just for amphibians and reptiles but also for animals of all kinds, including humans.

Another kind of pollution which may affect amphibians in agricultural habitats results from nitrates derived from fertilizers. Recent research has shown that frog tadpoles reared in water containing nitrates at levels low enough to be considered safe for human consumption suffer physical abnormalities, paralysis, and death (Marco and Blaustein, 1999). Nitrate fertilizers are widely used throughout the world and may constitute a serious threat to many amphibian populations.

Disease

Diseases are a natural cause of morbidity and mortality among animals, but very little is known about their role as a determinant of population size among amphibians. In the past few years, however, major outbreaks of disease among amphibians in many parts of the world have occurred, and some of these have had a major impact on the populations of many species. During the 1990s, mass mortalities occurred among populations of the common frog (*Rana temporaria*) in southeast Britain. These were caused by ranaviruses similar to pathogens that have caused mass mortalities among tiger salamanders (*A. tigrinum*) at several localities in the United States and Canada in recent years (Daszak *et al.*, 1999).

More recently, a disease called chytridiomycosis has become widespread throughout the world, and has wiped out many wild populations of amphibians (Daszak *et al.*, 1999). The disease is caused by a chytrid fungus, called *Batrachochytrium dendrobatidis*, an organism that invades the skin of amphibians. These sudden, major outbreaks of the same, previously unknown disease in widely distant parts of the world raises several questions that are urgently being addressed: Is the chytrid fungus a new organism that has only just evolved? Is it a long-established organism that has recently found its way to these localities and, if so, how? Current evidence suggests that amphibian chytrid is new to at least some parts of the world and that it has been spread, largely by human activities, from Africa.

Commercial Exploitation

Many people in many parts of the world have long eaten amphibians because they are a good and readily available source of protein. Until recently, such exploitation rarely had a serious impact on natural populations because it was localized, seasonal, and of low intensity. Unfortunately, modern commercialization of amphibians for the world's luxury food market is generally done with little regard for the long-term protection of natural populations. Most of the frogs that are killed for human consumption are not a vital component of the diet of local people in the developing world but are a luxury item in the diet of people in developed countries.

The scale of trade in frog legs has long been substantial, but it is beginning to be controlled and reduced. In 1976, 2.5 million kg of frog legs was imported into the United States, mostly from India and Japan. In France, the annual consumption is estimated to be 3000 or 4000 tons, imported mostly from Bangladesh and Indonesia. Until recently, 200 million pairs of legs were exported annually from Asia to the United States, Europe, and Australia; however, since 1987, India has banned this trade because declines in natural frog populations had caused dramatic increases in the densities of insect pests.

The emergence of the disease chytridiomycosis as a serious threat to amphibians has added an additional, urgent reason to curb the international trade in frogs for food. There is strong evidence that chytridiomycosis has been introduced into Europe, on several occasions, on American bullfrogs imported for food.

Products derived from amphibians are constituents of many traditional medicines and are used in many other ways. The skin secretions of poison-dart frogs have been used in hunting by Central and South American Indians, and those of many toad species are used in many parts of the world as hallucinogens. Whether any of these forms of exploitation has had a harmful effect on natural populations is doubtful, but

the future exploitation of amphibians and reptiles could either seriously threaten their continued existence or ensure their survival, depending on how it is managed. There is a lengthening list of compounds, found in the skin of frogs, that have considerable potential as medicines. For example, poison-dart frogs of the genus *Epipedobates* produce a unique alkaloid called epibatidine which is a much more effective painkiller than morphine and that appears not to be addictive.

Introduced Species

As the human population has expanded and people have colonized new parts of the world, they have taken with them, deliberately or accidentally, a variety of organisms that have been harmful to indigenous wildlife. The eggs and larvae of amphibians make easy prey for fishes and other freshwater predators, and many amphibian species are dependent on water bodies that are free of fish to breed successfully. In many parts of the world, exotic fishes have been introduced by humans to the detriment of native amphibians. In California, trout have been introduced, to provide sport fishing, to high-altitude lakes and have caused the decline of several amphibian species; in many parts of the world, mosquito fish (*Gambusia affinis*) have been introduced to control mosquitoes and other insect pests, with equally disastrous results for native amphibians.

Some amphibians have been implicated in the declines of other species where they have been introduced to parts of the world in which they did not previously occur. The most notable example is the marine toad (*Bufo marinus*), a native of Central America, which was deliberately introduced to Hawaii, Australia, and many other places to control insect pests in sugarcane plantations. In Australia, where it is known as the cane toad, it has spread relentlessly outward from the Queensland coast to the detriment of much of Australia's native fauna, including many frog species. The American bullfrog (*Rana catesbeiana*), a native of eastern North America, has been introduced to many parts of the world to be farmed to supply the trade in frog legs, and it has had a harmful effect on many native frog species both because it is a greatly superior competitor and because it is a predator of smaller frogs.

Declining Amphibian Populations

In the Monteverde cloud forest reserve in Costa Rica, the formation of large mating aggegations of the spectacularly colorful golden toad (*Bu. periglenes*) used to be an annual event. In 1989, however, this species failed to appear and it has not been seen since (Pounds and Crump, 1994). The golden toad has become the icon of the declining amphibian phenomenon, but it is only one of many species throughout the world that have disappeared because of this deeply disturbing process. The Monteverde reserve was established to protect biodiversity and yet, since 1990, 40% of its native frog and toad species have disappeared (Pounds *et al.*, 1999). Similar catastrophic declines of amphibian populations in supposedly pristine, protected habitats have occurred in Queensland, Australia, the Atlantic forests of Brazil, and the Pacific Northwest of North America. In a section of California's Sierra Nevada that includes Yosemite National Park, five of seven amphibian species have seriously declined in recent years. In protected remnants of tropical rain forest in eastern Queensland, 14 species of stream-dwelling frogs have drastically declined or totally disappeared. These events carry a disturbing message for conservationists: As one scientist said at a recent workshop on amphibian declines, "locking up nature just isn't working."

Although it is clear that the majority of the world's amphibians that have declined or become extinct in the past 50 years have done so as a result of habitat change or destruction, these declines in protected areas suggest that some other kind of process is adversely affecting amphibian populations. Because of their dual life history, spent partly in water and partly on land, and because none of their life stages (egg, larva, or adult) have the kind of protective covering possessed by animals such as insects, reptiles, birds, and mammals, it has been argued that amphibians are especially sensitive to environmental insults such as chemical pollution (Stebbins and Cohen, 1995). According to this argument, amphibian declines may be the prelude to an environmental catastrophe that could affect many forms of life. Like the coal miner's canary, frogs may be providing an early warning to all biodiversity.

It is becoming increasingly clear that this kind of argument is not applicable to all amphibians. Indeed, a feature of all the declines that have occurred recently in supposedly protected habitats, such as Monteverde, is that although some species have declined or vanished, others have been quite unaffected. It is not clear, however, what character or characters differentiate amphibian species that have declined in pristine areas from those that have not. Amphibians are a very diverse group of animals in terms of their habits, their life histories, and their physiology. For example, although some species do have a highly permeable skin and are extremely sensitive to pollutants, for others the skin is highly effective as a protective covering. The American bullfrog (*R. catesbeiana*), for example, seems to be remarkably resistant to the effects of pesticides.

Possible Link with Deformities

Since a group of children found some severely deformed frogs in Minnesota in 1995, there has been much media attention devoted to this phenomenon in the United States, often making a link with amphibian declines there and elsewhere. These deformities reflect abnormalities in development and include missing eyes, digits, or entire limbs as well as extra limbs, sometimes growing from unexpected parts of the body. There are several reasons for being cautious about a possible link between deformities and the global decline phenomenon, however. First, such deformities are not new but rather have been reported periodically for more than 200 years, especially in the mid-northern United States. Second, although reports have become more frequent recently, there is no reason to believe that deformed frogs have become more common; there are simply more people looking for them. Third, developmental abnormalities are caused by a variety of natural causes, including parasites and injury following attack by

predators, as well as by anthropogenic factors, such as chemical pollution and increased UV-B radiation.

Deformed amphibians are a cause for concern because they are often symptomatic of local environmental degradation. Before we can conclude anything of wider significance from them, however, we have to remember that they are both naturally occurring and typical of amphibians and that they do not represent a new phenomenon. Only additional research into the role of various pollutants and of UV-B into the development of deformities can reveal to what extent they are another facet of processes that are adversely affecting amphibians on a global scale.

See also: Agriculture, Nutrient Management, and Water Quality. Amphibians, Biodiversity of. Climate Change and Extinctions. Diseases, Conservation and. Ecotoxicology. Freshwater Ecosystems, Human Impact on. Global Declines of Amphibians. Herbicides. Human Impact on Biodiversity, Overview. Introduced Species, Impacts and Distribution of. Loss of Biodiversity, Overview. Mass Extinctions, Concept of. Nitrogen Deposition and Terrestrial Biodiversity. Pesticides, Uses and Effects of. Ultraviolet Radiation

References

Alford RA and Richards SJ (1999) Global amphibian declines: A problem in applied ecology. *Annual Review of Ecology and Systematics* 30: 133–165.

Blaustein AR, Edmund B, Kiesecker JM, Beatty JJ, and Hokit DG (1995) Ambient ultraviolet radiation causes mortality in salamander eggs. *Ecological Applications* 5: 740–743.

Blaustein AR, Hoffman PD, Hokit DG, Kiesecker JM, Walls SD, and Hays JB (1994) UV repair and resistance to solar UV-B in amphibian eggs: A link to population declines? *Proceedings of the National Academy of Sciences of the USA* 91: 1791–1795.

Blaustein AR and Kiesecker JM (2002) Complexity in conservation: Lessons from the global decline of amphibian populations. *Ecology Letters* 5: 597–608.

Daszak P, Berger L, Cunningham AA, Hyatt AD, Green DE, and Speare R (1999) Emerging infectious diseases and amphibian population declines. *Emerging Infectious Diseases* 5: 735–748.

Halliday TR (2008) Why amphibians are important. *Intlernational Zoo Yearbook* 42: 7–14.

Hayes TB, Khoury V, Narayan A, *et al.* (2010) Atrazine induces complete feminization and chemical castration in male African clawed frogs (*Xenopus laevis*). *Proceedings of the National Academy of Sciences of the USA* 107: 4612–4617.

Marco A and Blaustein AR (1999) Sensitivity to nitrate and nitrite in pond-breeding amphibians from the Pacific Northwest, USA. *Environmental Toxicology and Chemistry* 18: 2836–2839.

Pollard D (2010) *Living Planet Report 2010. Biodiversity, Biocapacity and Development.* Gland, Switzerland: World Wildlife Fund International.

Pounds JA and Crump ML (1994) Amphibian declines and climate disturbance: The case of the golden toad and the harlequin frog. *Conservation Biology* 8: 72–85.

Pounds JA, Fogden MPL, and Campbell JH (1999) Biological response to climate change on a tropical mountain. *Nature* 398: 611–615.

Reading CJ (2007) Linking global warming to amphibian declines through its effects on female body condition and survivorship. *Oecologia* 151: 125–131.

Stebbins RC and Cohen NW (1995) *A Natural History of Amphibians.* Princeton, NJ: Princeton University Press.

Stuart SN, Chanson JS, and Cox NA (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786.

Wake DB and Vredenburg VT (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the USA* 105: 11466–11473.

Endangered Ecosystems

Raymond C Nias, TierraMar Consulting, Sutherland, NSW, Australia

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Raymond C. Nias and John R. Mooney, pp 1–15, © 2001, Elsevier Inc.

Glossary

Biodiversity hotspots A region that contains at least 1500 species of vascular plants (>0.5% of the world's total) as endemics and has lost at least 70% of its original habitat.
Conservation status A classification of species and ecosystems according to their extinction risk; such as vulnerable, endangered, and critically endangered.
Ecoregion A relatively large area of land or water that contains characteristic, geographically distinct assemblages of natural communities, and species.
Ecosystem composition The species or functional groups of organisms found within a given ecosystem.
Ecosystem function The biological and chemical processes that occur within ecosystems such as primary production and nutrient cycling.
Ecosystem structure The different trophic (feeding) levels within an ecosystem and the interactions between them.
Extinction risk The likelihood of extinction of a species with specific demographic characteristics and distribution under various threatening processes over time.

The article outlines recent progress in the development of systematic approaches to the identification of the world's ecosystems and their conservation status. A synopsis is given of how from the mid-1990s through to the present conservation biologists and practitioners have been devising more explicit classifications systems for ecosystem types that has allowed for the identification and mapping of terrestrial, freshwater, and marine ecosystems from the global to local scale. Summaries are provided of these studies as well as recent progress in developing a new and more systematic approach to the classification of extinction-risk for ecosystems. In addition to overall trends derived from global assessments some examples are provided of globally significant ecosystems considered to be endangered.

Defining Ecosystems and their Conservation Status

In contrast to the situation with threatened species, the definition of an ecosystem as a "unit of conservation concern" has been problematic. Difficulties with the identification of distinct ecosystems and what constitutes the extinction of an ecosystem have hampered attempts to develop objective criteria for categorization of ecosystems and their conservation status. The definition of an ecosystem as "a dynamic complex of plant, animal, and microorganism communities and their nonliving environment interacting as a functional unit" (Convention on Biological Diversity, http://www.cbd.int/convention/text/) for example, does not define any specific dimensions or boundaries. At the smallest scale there are ecosystems of invertebrates and microbes that exist within a few liters of water or a few kilograms of soil. In fact each ecosystem is nested within a larger ecosystem up to the scale of the biosphere which can be considered as a single planetary ecosystem. Apart from the question of scale, ecosystems are also dynamic, with many natural changes both in areal extent and composition occurring over time. In aquatic ecosystems, for example, the area covered will fluctuate with changing water levels, floods, and tides and entire ecosystems such as ephemeral wetlands may come and go in the space of a few months. Terrestrial ecosystems are typically defined by their major vegetation type such as moist broad-leaved (tropical) forest or lowland temperate grasslands, but may also be characterized by unique physical features such as particular geological substrates (e.g., lime-stone-based karst ecosystems). In the marine environment ecosystems may be defined geographically and also by particular attributes of water temperature or chemistry, or by the characteristics of the substrate (e.g., rocky, sandy, or muddy bottoms). Specific plant and animal communities such as seagrass beds, kelp forests, and coral reefs may also form part of the definition of a particular marine ecosystem.

For individual species, the identification of conservation status as measured by the degree of extinction risk has been widely accepted for decades. Species are classified into various categories (such as not at risk, of least concern, vulnerable, endangered, and critically endangered) based on well-established criteria that relate to overall population size and trends (Mace *et al.*, 2008). In relation to species the idea of biological extinction is relatively easy to define but for ecosystems made up of multiple species this is far more difficult. In a pioneering study of North American ecosystems Noss *et al.* (1995) described ecosystem loss as taking two forms: quantitative loss and qualitative loss. Quantitative loss is the decline in areal extent of a discrete ecosystem type (e.g., the conversion of a native prairie grassland into a corn field) while qualitative loss is a change or degradation in the structure, function, or composition of an ecosystem (e.g., the change in species composition of a steppe ecosystem due to overgrazing, or the removal of key predator species in a marine ecosystem though overfishing). Problems arise when the concept of extinction risk is applied to ecosystems and the extent to which the loss of any species from an ecosystem impacts on the overall ecosystem structure and function is still the subject of much uncertainty and research (Duffy, 2003; Thébault *et al.*, 2007). Despite these obvious difficulties there have been

arguments made to codify the conservation status of ecosystems from at least the mid-1990s (e.g., Noss, 1996) to the present (Rodríguez *et al*., 2007, 2011; Nicholson *et al*., 2009).

Assessments of Endangered Ecosystems

Some of the first systematic reviews of the conservation status of particular ecosystems were carried out in the early 1990s including tropical rivers in Asia (Dudgeon, 1992) and North American ecosystems (Noss *et al*., 1995). The ongoing development of systematic approaches for mapping ecosystems at large scales – much of which is based on the concept of geographically and biologically distinct areas or "ecoregions" (Dinerstein *et al*., 1995) – has lead to the identification and mapping of global terrestrial (Olson *et al*., 2001), freshwater (Abell *et al*., 2008), and marine (Spalding *et al*., 2007) ecoregions. Although "ecoregions" and "ecosystems" may not necessarily be the same thing, systematic methods to describe spatial patterns of biodiversity distribution have greatly facilitated global overviews as well as continental-scale assessments of terrestrial and freshwater biodiversity (e.g., Burgess *et al*., 2004; Thieme *et al*., 2005, among others). Significant milestones in the application of systematic approaches to ecosystem conservation include the identification of threatened biodiversity hotspots (Myers *et al*., 2000) priority ecoregions (Olson and Dinerstein, 1998), globally significant wilderness areas (Mittermeier *et al*., 2003), and key biodiversity areas (Eken *et al*., 2004) and the synthesis of these approaches (Brooks *et al*., 2006). Finer-scale approaches can also assist in identifying alternative strategies for prioritizing conservation efforts within a particular ecosystem; and this has been well demonstrated in places as diverse as the Cape Floristic Region of South Africa (Cowling and Pressey, 2003), the Atlantic forests of Brazil (Paese *et al*., 2010) and the Great Barrier Reef of Australia (Fernandes *et al*., 2005). There are also some long-established monitoring programs that track trends in conservation status of ecosystems over time (e.g., coral reefs; Wilkinson, 2008). Recently the International Union for Conservation of Nature (IUCN) has begun work on a classification scheme for threatened terrestrial ecosystems that intends to build on the successful application of the Red List of Threatened Species and protocols in use by some government conservation agencies (Rodríguez *et al*., 2011). This system proposes the use of a naming system analogous to the present IUCN species-specific system (Least Concern, Vulnerable, Endangered and Critically Endangered) using four criteria: (A) reduction of land cover and continuing threat, (B) rapid rate of land cover change, (C) increased fragmentation, and (D) highly restricted geographical distribution. However a number of challenges remain to be solved and it remains unclear to what extent important changes in species composition (such as the loss of a key predator species for example) can be incorporated into this approach.

The Millennium Ecosystem Assessment (MA)

The Millennium Ecosystem Assessment (MA) (2005) was called for by the United Nations Secretary-General in 2000 and was initiated in 2001 through the United Nations Environment Programme (UNEP). The purpose of the MA was to "assess the consequences of ecosystem change for human well-being and to establish the scientific basis for actions needed to enhance the conservation and sustainable use of ecosystems and their contributions to human well-being." The report includes global overviews of marine, coastal, inland water, forest, dryland, island, mountain, and polar ecosystems as well as 18 subglobal assessments. The conclusions of the MA are stark in their portrayal of ecosystem loss; stating that "virtually all of Earth's ecosystems have now been dramatically transformed through human actions." Over half of the 14 biomes that the MA assessed have experienced a 20–50% conversion to human use, with temperate and Mediterranean forests and temperate grasslands being the most affected. Marine ecosystems have also been badly affected with 20% of coral reefs and 35% of mangrove ecosystems being lost in the past few decades.

General Trends in Conservation Status of Major Ecosystem Types

Terrestrial Ecosystems

Forest ecosystems currently occupy approximately 31% of the Earth's land surface and are estimated to contain more than half of all terrestrial animal and plant species, the great majority of them in the tropics. In most of the world the major threat to forest ecosystems is conversion to agricultural land. According to the third United Nations Global Biodiversity Outlook (http://gbo3.cbd.int/home.aspx) nearly 130,000 km^2 of forest were converted to other uses or lost through natural causes each year from 2000 to 2010, compared to nearly 160,000 km^2 per year in the 1990s. Many tropical moist forest ecosystems are considered to be endangered, particularly in western Africa, the lowland forests of Southeast Asia, the Atlantic coast of Brazil, and forests on islands such as Madagascar, Indonesia, and the Philippines. Although rapid development of these regions for agriculture and large-scale commercial logging has been primarily responsible for this extensive forest loss, large-scale fire now poses a major problem in some of the previously contiguous forests of Indonesia and Malaysia, especially during periods of prolonged drought. Similar problems are faced the seasonally dry forests, especially the monsoonal forests and conifer forests such as those found in Central America and Madagascar. Owing to their restricted distribution the monsoonal forests and conifer forests in the tropical and subtropical regions of the world are among the most endangered of forest ecosystems. The biologically rich dry forests of Madagascar for example are now much reduced but still some of the richest forest ecosystems in the world in terms of their natural diversity. Most of the forest has been cleared for slash-and-burn agriculture, pasture, firewood, or construction materials and secondary grasslands now cover most of the region. The remaining fragments of forest are steadily being degraded and lost as a result of uncontrolled burning of surrounding savannas. Most temperate forest ecosystems of the world have also been modified significantly to increase their commercial timber value at the

Table 1 Examples of some ecosystems and ecoregions commonly referred to as endangered

Region	*Terrestrial ecosystems*		*Freshwater ecosystems*		*Marine and coastal ecosystems*	
Africa	Tropical moist forests	Madagascar forests, thicket, and scrublands	Lakes	Great Lakes (Kivu, Edward, George, Victoria)	Oceans and seas	Canary Current (Eastern Atlantic)
		West African moist forests		Lake Chad	Coastal	Gulf of Guinea mangroves
		Eastern arc and East African coastal forests		Cameroonian Crater lakes		East African mangroves
		Albertine rift highland forests (central Africa)	Rivers and streams	Upper Guinea rivers and streams		Madagascar mangroves
	Savannas and woodlands	Sudanian savannas		Cape rivers and streams (South Africa)		
		Ethiopian highlands	Deltas	Niger river delta		
	Shrublands	Fynbos shrublands of the Cape of South Africa				
		Drakensberg Montane shrublands and woodlands				
	Deserts	Namib and Karoo deserts and shrublands				
Asia and Middle East	Tropical moist forests	Moist forests of Sumatra	Lakes	Aral Sea (central Asia)	Oceans and seas	Yellow Sea
		Southwestern Ghats moist forests (India)		Lake Biwa (Japan)		Sulu-Sulawesi Sea
		Peninsular Malaysia lowland and Montane forests		Yunnan lakes and streams (China)		Arabian Sea
	Temperate forests	Central China Temperate forests	Rivers and streams	Western Ghats rivers and streams (India)	Coastal	Greater Sundas mangroves
		Eastern Himalayan broad-leaved and conifer forests		Xi Jiang (Pearl) river and streams (China)		Sundarbans mangroves
	Dry forests	Nusa Tenggara dry forests (Indonesia)		Philippines freshwater ecoregion		Mangrove ecosystems (Thailand)
		Indochina dry forests	Swamps and marshes	Peat swamps of lowland Borneo		Nansei Shoto Coral Reefs (Japan)
		Chota-Nagpur dry forests (India)		Mesopotamian Marshes (Iraq)		
	Savannas and grasslands	Terai-Duar savannas and grasslands (north Indian subcontinent)	Deltas	Indus river delta		
		Rann of Kutch flooded grasslands (India, Pakistan)				
	Deserts	Socotra Island Desert (Yemen)				

(*Continued*)

Table 1 Continued

Region	*Terrestrial ecosystems*		*Freshwater ecosystems*		*Marine and coastal ecosystems*	
Australasia and Pacific	Forests	New Caledonia forests	Wetlands	New Caledonia rivers and streams		
		New Zealand subtropical forests		Freshwater Lakes of Southwest Australia		
		Hawaiian moist forests		Alpine Sphagnum Bogs and associated Fens (Australia)		
	Woodlands, grasslands, and shrublands	Southwest Australia woodlands and heaths				
		Buloke and Grassy White Box woodlands (Australia)				
		Lowland temperate grasslands of Australia				
Europe and Russia	Forests	Caucasus–Anatolian–Hyrcanian temperate forests	Rivers and streams	Volga River delta	Seas	Mediterranean sea
		Southern European Montane forests		Danube river delta		Baltic Sea
	Woodlands and scrubs	Mediterranean forests, woodlands, and scrub		Balkan rivers and streams		
				Anatolian rivers and streams		
South America and Caribbean	Tropical moist forests	Brazilian Atlantic forests	Lakes	High Andean Lakes	Oceans and seas	Gulf of Mexico
		Northern Andean Montane forests	Rivers and streams	Upper Paraná River		Northeast Atlantic Shelf
		Río Negro-Juruá moist forests		Greater Antillean freshwater ecoregion	Coastal and islands	Galápagos Islands
	Temperate forests	Valdivian temperate rain forests (Chile)				Coral reefs of the Caribbean
	Dry forests	Tumbesian and North Inter-Andean Valleys dry forests				Mangrove ecosystems of Ecuador
		Chiquitano dry forests				
	Savannas	Pantanal flooded savannas				
	Shrublands and grasslands	Chilean Matorral				
		Patagonian steppe				

North America	Forests	Klamath–Siskiyou coniferous forests	Lakes	Mexican highland lakes	Seas	Grand Banks
		Sierra Madre Oriental & Occidental Pine-Oak forests		Great Lakes (USA)		Gulf of California
		Mesoamerican Pine-Oak forests	Rivers and streams	South-eastern rivers and streams and Florida Everglades		
		Sierra Nevada conifer forests		Pacific Northwest rivers and streams		
		Southeastern USA conifer and broad-leaved forests		Colorado River and delta		
	Woodlands	California Chaparral and woodlands		Mississippi Piedmont rivers and streams		
	Grasslands	Northern Prairie		Chihuahuan freshwater ecoregion		
Polar					Seas	Barents-Kara Seas

expense of natural complexity and structure and they are more likely to contain uniform, relatively even-aged stands, with little dead timber, and with a significantly reduced complement of animal species. The deliberate introduction of non-native timber species has also altered the composition of temperate forest ecosystems across many parts of the world.

Despite their often uniform appearance, many woodland, heath, and grassland ecosystems are complex and varied, with many regional variations and subtle differences resulting from soil type and climate. They are usually present where soil nutrients, water availability, or climate are not conducive to the growth of forests. Temperate and semiarid scrubs, heaths, and grasslands in particular are well suited to grazing by domestic animals and for conversion to cultivation, especially cereal crops, and as a consequence are among the most modified of the world's major ecosystems. The loss and degradation of semiarid ecosystems is also a key driver for the process of desertification in many parts of the world.

Freshwater Ecosystems

Freshwater ecosystems including rivers and their floodplains, lakes, and wetlands have undergone more dramatic changes than any other type of ecosystem due to a combination of human activities including drainage for agriculture, abstraction of water for irrigation, industrial and household use, the input of nutrients and other pollutants, introduction of alien species, and the damming of rivers. Even if the physical structure of wetlands remains intact, they are often vulnerable to the effects of chemical pollution. Such pollution can significantly modify the ecosystem as key organisms are lost as a result. The use of lake and river water for industrial purposes often affects water quantity (when abstracting water) and water quality (when reintroducing it as wastewater, sometimes polluted, or of higher temperature). Many industries, such as pulp and paper production and mining, wash large quantities of particulate matter into lakes and rivers. Pollution and sedimentation threaten many inland waters and other wetlands. In some cases, the source of this pollution may be distant, with contaminants being transported through the air in the form of industrial emissions. The most significant impact on global freshwater ecosystems however is the removal of water for human use, particularly for irrigation.

Marine and Coastal Ecosystems

Coastal marine ecosystems are among the most productive ecosystems on earth but are also highly threatened by a range of human activities and in particular, pollution from land and ocean-based sources, coastal development, and overfishing. These impacts are being compounded by sea-level rise creating a "coastal squeeze." Coastal ecosystems and near-shore marine areas are among the most threatened of all ecosystem types with some 35% of mangrove ecosystems having been lost or converted, 20% of coral reefs having been destroyed, and up to 20% of coastal wetlands lost in some places of the world on an annual basis (MA, 2005). Recent assessments have shown a high proportion of mangrove and seagrass species – which are the basis of these ecosystem types, are at risk of extinction globally (Hughes *et al.*, 2009; Polidoro *et al.*, 2010). The loss of these coastal ecosystems also represents a significant loss of natural resources of great importance to many human societies and economies and their importance in storing carbon dioxide is also gaining increased attention (MA, 2005).

The world's oceans contain some of the least explored areas of the world, and only a tiny percentage of the deep seabed has been subject to biological investigation. Nevertheless, studies have revealed a wealth of diverse ecosystem types in the deep sea and some such as seamount ecosystems and cold water coral reef ecosystems are threatened by high-impact fishing methods, such as bottom trawling. Like other ecosystems, marine ecosystems also show the effects of changes in species abundance caused by direct human consumption. The impact of overfishing on marine ecosystems has been the subject of considerable debate in the past 10 years (e.g., Coll *et al.*, 2008). The full impacts of global warming on marine ecosystems are only just beginning to be understood although the predicted impacts may well be catastrophic in many cases (Hoegh-Guldberg and Bruno, 2010).

Although coral reefs occupy a small fraction ($<0.2\%$) of the total area of the world's oceans, they are among the most biologically diverse of all ecosystems. It is estimated that about 20% of the world's coral reefs have been destroyed and an additional 20% is badly degraded or under imminent risk of collapse (MA, 2005). The destruction of coral reefs is caused by many human activities, ranging from coastal development and destructive fishing practices to over-exploitation of resources, marine pollution, and runoff from agricultural activities and deforestation. A number of reefs of particular biological interest are also under the most serious threat, including almost all the reefs of the Philippines and coral communities in coastal Indonesia, Tanzania, the Comoros, and the Lesser Antilles in the Caribbean. Of all the world's ecosystems, coral reefs are among the most vulnerable to the effects of climate change. Coral bleaching events due to high water temperatures events often lead to mass mortality although some recovery may occur given sufficient time between bleaching episodes (Wilkinson, 2008) (**Table 1**).

Conclusions

Ecosystems are dynamic and range in scale from the biosphere to the microscopic. As a result the exact delineation of an ecosystem and its description as being "endangered" has been a largely subjective matter for many years. There has been progress in the past decade in the identification and description of the world's ecosystems and perhaps the most important advance has been the identification and global mapping of terrestrial, freshwater, and marine ecoregions. Recent work on the development of a systematic set of protocols for defining the conservation status of ecosystems may provide the basis for even more rigorous assessments of the true state of biodiversity to be determined at the global, continental, and regional scale. Overall the conservation trends for ecosystems and biological diversity in general, are extremely negative (Butchart *et al.*, 2010).

See also: Africa, Ecosystems of. Alpine Ecosystems. Amazon Ecosystems. Antarctic Ecosystems. Arctic Terrestrial Ecosystems. Asia, Ecosystems of. Australia, Biodiversity of Ecosystems. Boreal Forest Ecosystems. Central America, Ecosystems of. Coastal Beach Ecosystems. Conservation Biology, Discipline of. Conservation Efforts, Contemporary. Ecosystem, Concept of. Ecosystems of South America. Estuarine Ecosystems. Europe, Ecosystems of. Freshwater Ecosystems. Grassland Ecosystems. High-Temperature Ecosystems. Human Impact on Biodiversity, Overview. Intertidal Ecosystems. Lake and Pond Ecosystems. Loss of Biodiversity, Overview. Mangrove Ecosystems. Marine Ecosystems. Mediterranean-Climate Ecosystems. Ocean Ecosystems. Rainforest Loss and Change. River Ecosystems. South American Natural Ecosystems, Status of. Southern (Austral) Ecosystems. Subterranean Ecosystems. Temperate Forests. Terrestrial Ecosystems. Tropical Forest Ecosystems. Wetlands Ecosystems

References

Abell R, Thieme ML, Revenga C, *et al.* (2008) Freshwater ecoregions of the world: A new map of biogeographic units for freshwater conservation. *BioScience* 58(5): 403–414.

Brooks TM, Mittermeier RA, de Fonseca GAB, *et al.* (2006) Global biodiversity conservation priorities. *Science* 313: 58–61.

Burgess N, D'Amico Hales J, Underwood E, *et al.* (2004) *Terrestrial Ecoregions of Africa and Madagascar: A Conservation Assessment*. Washington, DC: Island Press.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328(5982): 1164–1168.

Coll M, Libralato S, and Tudela S (2008) Ecosystem overfishing in the ocean. *PLoS ONE* 3(12): e3881.

Cowling RM and Pressey RL (2003) Introduction to systematic conservation planning in the Cape Floristic Region. *Biological Conservation* 112: 1–13.

Dinerstein E, Olson DM, Graham DJ, *et al.* (1995) *A Conservation Assessment of the Terrestrial Ecoregions of Latin America and the Carribean*. Washington, DC: World Bank.

Dudgeon D (1992) Endangered ecosystems: A review of the conservation status of tropical Asian rivers. *Hydrobiologia* 248: 167–191.

Duffy JE (2003) Biodiversity loss, trophic skew and ecosystem functioning. *Ecology Letters* 6: 680–687.

Eken G, Bennun L, Brooks TM, *et al.* (2004) Key biodiversity areas as site conservation targets. *BioScience* 54(12): 1110–1118.

Fernandes L, Day J, Lewis A, *et al.* (2005) Establishing representative no-take areas in the Great Barrier Reef: Large-scale implementation of theory on Marine Protected Areas. *Conservation Biology* 19(6): 1733–1744.

Hoegh-Guldberg O and Bruno JF (2010) The impact of climate change on the World's marine ecosystems. *Science* 18(328): 1523–1528.

Hughes RA, Williams SL, Duarte CM, *et al.* (2009) Associations of concern: Declining seagrasses and threatened dependent species. *Frontiers in Ecology and the Environment* 7: 242–246.

Mace GM, Collar NJ, Gaston KJ, *et al.* (2008) Quantification of extinction risk: IUCN's system for classifying Threatened Species. *Conservation Biology* 22: 1424–1442.

Millennium Ecosystem Assessment (MA) (2005) *Ecosystems and Human Well-being: Synthesis*. Washington, DC: Island Press.

Mittermeier RA, Mittermeier CG, Brooks TM, *et al.* (2003) Wilderness and biodiversity conservation. *Proceedings of the National Academy of Sciences* 100(18): 10309–10313.

Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Nicholson E, Keith DA, and Wilcove DS (2009) Assessing the threat status of ecological communities. *Conservation Biology* 23: 259–274.

Noss RF (1996) Ecosystems as Conservation Targets. *Trends in Ecology and Evolution* 11: 351.

Noss RF, LaRoe ET III, and Scott JM (1995) Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. *Biological Report 28*. Washington, DC: National Biological Service.

Olson DM and Dinerstein E (1998) The Global 200: A representation approach to conserving the Earth's most biologically valuable ecoregions. *Conservation Biology* 12(3): 502–515.

Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience* 51: 933–938.

Paese A, Paglia A, Pinto LP, *et al.* (2010) Fine-scale sites of global conservation importance in the Atlantic forest of Brazil. *Biodiversity and Conservation* 19: 3445–3458.

Polidoro BA, Carpenter KE, Collins L, *et al.* (2010) The loss of species: Mangrove extinction risk and geographic areas of global concern. *PLoS ONE* 5(4): e10095.

Rodríguez JP, Balch JK, and Rodríguez-Clark KM (2007) Assessing extinction risk in the absence of species-level data: Quantitative criteria for terrestrial ecosystems. *Biodiversity and Conservation* 16: 183–209.

Rodríguez JP, Rodríguez-Clark KM, Baillie JEM, *et al.* (2011) Establishing IUCN red list criteria for threatened ecosystems. *Conservation Biology* 25: 21–29.

Spalding MD, Fox HE, Allen GR, *et al.* (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* 57(7): 573–583.

Thébault E, Huber V, and Loreau M (2007) Cascading extinctions and ecosystem functioning: Contrasting effects of diversity depending on food web structure. *Oikos* 116: 163–173.

Thieme ML, Abell R, Stiassny MLJ, *et al.* (2005) *Freshwater Ecoregions of Africa and Madagascar: A Conservation Assessment*. Washington, DC: Island Press.

Wilkinson C (2008) *Status of Coral Reefs of the World*. Townsville, Australia: Global Coral Reef Monitoring Network and Reef and Rainforest Research Centre.

Endangered Freshwater Invertebrates

David L Strayer, Cary Institute of Ecosystem Studies, Millbrook, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aquifer A geological formation that contains and allows movement of ground water.
Endangered species A species that is at substantial risk of extinction as a result of human activities.
Endemic species A species that occurs only over a limited geographical range.
Eutrophication The process of increasing the productivity of an ecosystem by enriching it with nutrients.
Ground water Water that occurs in saturated soils and geological formations beneath the Earth's surface.
Nonnative species A species that has been moved and established outside of its native range as a result of human activities; also called alien species, exotic species, introduced species, nonindigenous species.

The Earth's fresh waters contain an impressive diversity of invertebrate life. More than 100,000 species of have already been described, and tens of thousands of species remain to be discovered and described by scientists. This diversity is not spread evenly over the surface of the globe, but concentrated in local "hot spots," usually ancient lakes, streams, or ground waters. These "hot spots" often contain dozens to hundreds of species of freshwater invertebrates that are found nowhere else in the world. Because fresh waters are such an important resource for people, and have been used intensively for water supply, power, irrigation, fisheries, navigation, waste disposal, and as sites for cities, environmental conditions in many of the world's fresh waters have been altered greatly from their original states. Especially where "hot spots" of diversity coincide with areas of intensive human development, many freshwater invertebrates have disappeared from their native habitats. Some invertebrate species already have become extinct, and thousands more are in danger of disappearing from the Earth. Careful management of fresh waters, especially in "hot spots" of high biological diversity, is needed to prevent catastrophic extinctions of freshwater invertebrates in the next century.

Introduction

Freshwater Invertebrates of the World

Over 100,000 species of freshwater invertebrates have been described, representing about 750 families and 15 phyla (Balian *et al.*, 2008). The world's freshwater invertebrate fauna is actually much larger than this; probably 10,000–100,000 species await discovery, and new genera and families are discovered regularly. We are particularly ignorant about what lives in ground waters, what lives outside of Europe and parts of North America, and small, soft-bodied invertebrates. The most widespread and species-rich groups of freshwater invertebrates are the insects, crustaceans, mollusks, mites, nematodes, and rotifers, but many other groups are common as well. Invertebrates live in nearly all kinds of freshwater habitats; lakes, rivers, brooks, ephemeral ponds, wetlands, caves, alluvial ground waters, and even hot springs each contain a rich and characteristic invertebrate community. Among important freshwater habitats, perhaps only the deepest ground waters usually lack invertebrates. A typical lake or stream contains a few hundred species of invertebrates representing several dozen families and eight to 12 phyla. By comparison with the better-known vertebrates, freshwater invertebrates possess a wide range of biological traits. Life spans range from days to more than a century. Many invertebrates reproduce sexually, while others reproduce asexually by budding or parthenogenesis, or change their sexuality or mode of reproduction depending on environmental conditions. Some invertebrates produce eggs or other reproductive bodies that remain viable for years to centuries. Freshwater invertebrates include herbivores, bacteriovores, fungivores, predators, and parasites, and exhibit a wide range of specialized morphologies and behaviors to aid in food-gathering. Some even use symbiotic algae to photosynthesize!

Table 1 NatureServe's system for ranking global conservation status of species

GX	Presumed extinct: not located despite intensive searches and virtually no likelihood of rediscovery
GH	Possibly extinct: known only from historical occurrences but still some hope of rediscovery
G1	Critically imperiled: at very high risk of extinction due to extreme rarity (often five or fewer populations) very steep declines, or other factors
G2	Imperiled: at high risk of extinction or elimination due to very restricted range, very few populations, steep declines, or other factors
G3	Vulnerable: at moderate risk of extinction due to a restricted range, relatively few populations, recent and widespread declines, or other factors
G4	Apparently secure: uncommon but not rare; some cause for long-term concern due to declines or other factors
G5	Secure: common; widespread and abundant

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00270-7

What is ''Endangered''?

Various terms such as ''endangered,'' ''threatened,'' ''imperiled,'' and ''at risk,'' have been used to describe species that are in danger of extinction through human activities. Conservation organizations and governments typically have tried to develop a graded series of carefully defined terms, running from species only remotely threatened with extinction to those on the verge of extinction. **Table 1** shows an example of such a system. Here, ''endangered'' is used loosely to mean a species or population that is at substantial risk of becoming extinct over the next few decades as a result of human activities.

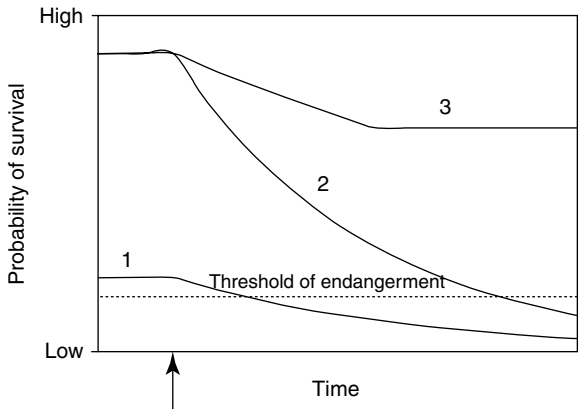


Figure 1 Diagram illustrating general causes of species endangerment. The three solid lines represent different species, and the dashed line shows the probability of long-term species survival below which we regard a species as endangered. The heavy arrow shows the point at which human impacts began. Species 1 and 2 are now endangered, but for different reasons. Species 1 became endangered because it was vulnerable to extinction before human impacts, and was somewhat sensitive to human activities. Species 2 and 3 were not very vulnerable to extinction before human impacts. Species 2 was either highly sensitive to human activities or lived in an area where human activities were intensive, while species 3 was either not very sensitive to human activities or lived in an area where human activities were weak.

Causes of Endangerment

The specific causes of endangerment of freshwater invertebrates are highly varied from case to case. It would be impractical (and probably not very illuminating) to discuss all the known cases of endangerment of freshwater invertebrates. Further, because information about the world's freshwater invertebrates is still so incomplete, a catalog of known cases of endangerment would be so biased as to be misleading. Instead, the problem of endangerment is discussed here in a more general way, and selected case studies are used to illustrate major points.

In a general way, endangerment is a product of three factors: the pre-existing vulnerability of a species, the pressure of human activities, and the sensitivity of the species to specific human activities (**Figure 1**). Thus, a species may become endangered if it already was vulnerable to extinction before human involvement, if human activities heavily affect most of the regions or habitats that it occupies, or if it is especially sensitive to a particular human activity. Conversely, a species is likely to avoid endangerment only if it evades all three of these conditions. Each of these factors is discussed below in more detail.

Vulnerability of Freshwater Invertebrates

Small Ranges

Many species of freshwater invertebrates had small ranges even before human intervention. Species with small ranges are called narrowly endemic species. For example, half of the 281 North American pearly mussel species were found only in one to three states, even before human intervention. Such species may have had elevated probabilities of extinction through natural catastrophes, and certainly are especially vulnerable to human activities (**Figure 2**). Small natural ranges often arise though a small number of understandable processes. Because these processes are focused in certain regions and on species with characteristic biological traits, narrowly endemic species

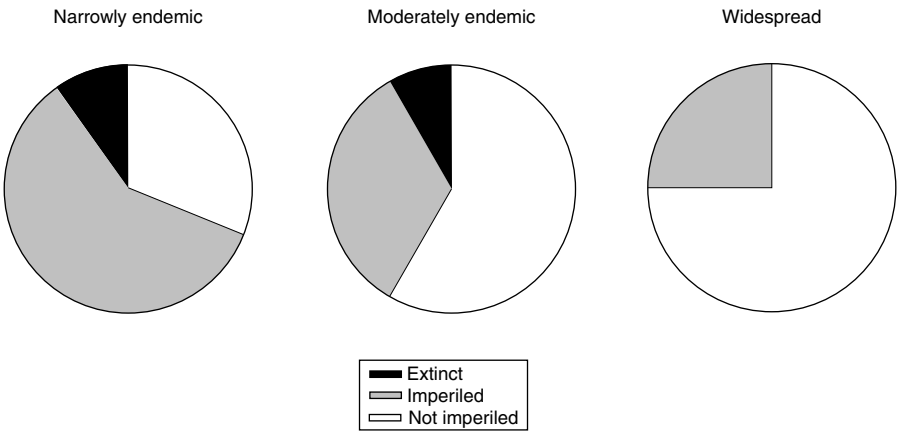


Figure 2 Current conservation status of North American pearly mussels (Unionoida) as a function of their native range sizes. Narrowly endemic species were found in one to three states and provinces, moderately endemic species in four to six states and provinces, and widespread species in more than six states and provinces.

often are clustered together into small regions and concentrated in certain taxonomic or ecological groups.

Causes of Small Ranges

A primary cause of small ranges in freshwater invertebrates is the limited dispersal abilities of these animals. All freshwater habitats are islands in a sea of terrestrial habitats, and are more or less isolated from other similar habitats. Although streams are connected together into drainage networks, the streams of one drainage network are isolated from those in other drainage networks. The dispersal abilities of freshwater invertebrates, and thus the perceived isolation of freshwater habitats, vary widely. For animals like dragonflies, whose long-lived aerial adults are strong fliers, or ectoproct bryozoans, whose tough resting stages ("statoblasts") are readily dispersed by migratory waterfowl, the separation of fresh waters probably does not present an important barrier to dispersal or gene flow. For other animals, such as fragile groundwater crustaceans that are poor swimmers, avoid the light, and lack tough dispersal stages, adjacent streams or aquifers may be nearly as remote as distant continents, and even small barriers may prevent migration and gene flow. For instance, the present-day distribution of microparasellid isopods nearly follows the pattern of marine beaches from over 20 million years ago (**Figure 3**), where these species presumably arose and from which they subsequently apparently have been unable to disperse.

The isolation of freshwater habitats may produce small ranges in two ways. First, endemic species with small ranges may evolve in place following infrequent crossing of dispersal barriers, resulting in a group of more or less closely related species whose ranges are separated by barriers to dispersal. Second, a formerly widespread species may be eliminated from most of its former range, for instance by a changing climate or the arrival of a competitor, stranded in small refuges, and be unable subsequently to disperse out of the refuges. This second mechanism may become especially important for the freshwater biota if humans cause large changes in regional or global climate, especially because habitat alterations and pollution have eliminated many of the natural dispersal corridors between freshwater habitats.

Finally, a species may have a small range because it requires an unusual habitat, which is itself rare. For example, the thermosbaenacean crustacean *Thermosbaena mirabilis* was described from ancient Roman warm baths and is known from only a few thermal springs in Tunisia. Its small range presumably owes to its unusual habitat requirements as well as its limited dispersal abilities.

"Hot Spots" of High Endemism

Because the processes of speciation, extinction, and dispersal do not occur uniformly over the earth, species richness and endemism vary greatly across the world's fresh waters. Some bodies of water contain more than 1000 invertebrate species, many of them unique to that single body of water. At the other extreme, some bodies of water support fewer than 100 invertebrate species, all of them widely distributed. We might expect sites of high richness and endemism to be habitats of great age, habitats where dispersal is limited, either by geographic isolation or by characteristics of the habitat, or habitats that harbor animal groups that are prone to speciate. Thus, many ancient lakes (Baikal in Siberia, Tanganyika, and Malawi in Africa) and river systems (the Tennessee in the United States, the Mekong in southeast Asia) that have not been

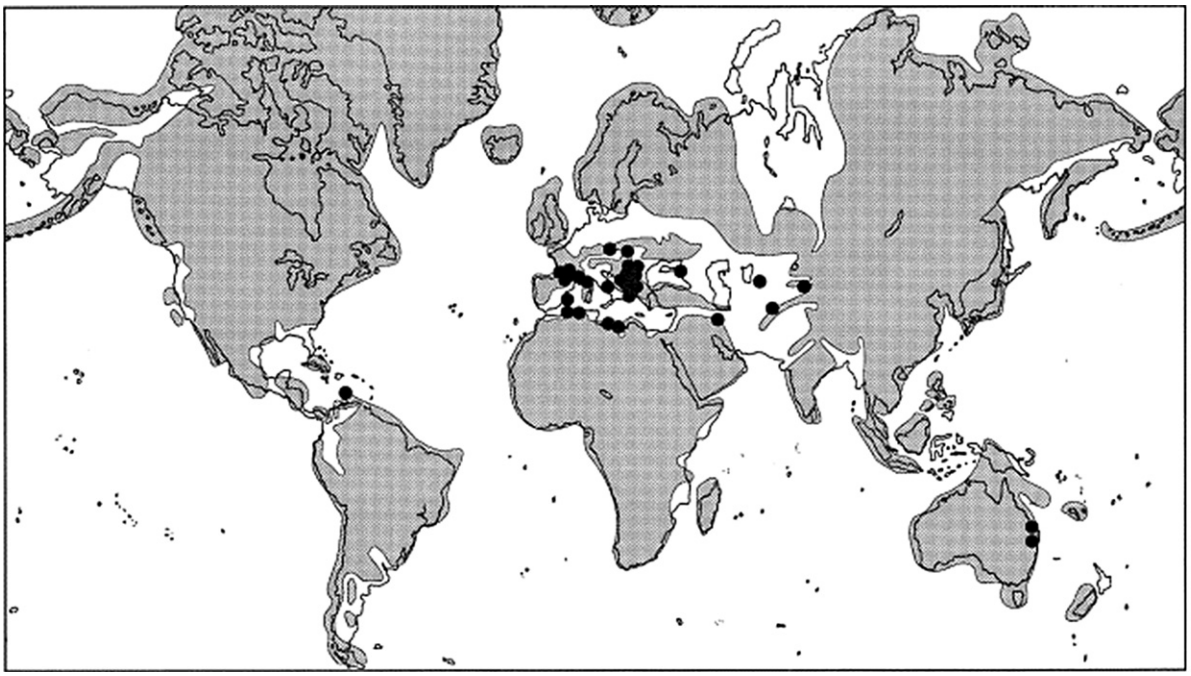


Figure 3 Distribution of freshwater microparasellid isopods and Oligocene shorelines (24–37 million years ago). Stippled areas were land during the Oligocene and closed circles show places where freshwater microparasellids have been found.

recently disturbed by glaciation, marine submergence, or desiccation support unique assemblages of invertebrates (**Boxes 1** and 2). Many aquifers seem to contain a high proportion of endemic species, probably because both the characteristics of aquifers (slow water flow, tortuous passageways within aquifers, and barriers between aquifers) and the characteristics of their inhabitants (e.g., fragile bodies, strong thigmotaxis) discourage long-range dispersal. Conversely, glacial lakes and temporary ponds rarely support locally endemic species. Although we know many hot spots of freshwater invertebrate diversity, it probably is not yet possible to produce a reliable global map that shows all major hot spots, as has been recently done for freshwater fishes (Abell *et al.*, 2008).

Species with High Endemism

Groups of animals vary in their tendency to form new species with small ranges. As already suggested, narrow dispersal probably allows the development of local species, while broad dispersal probably provides so much gene flow across populations that speciation is unlikely. Other traits that have been suggested to encourage local speciation include a requirement for outcrossing (as opposed to selfing hermaphroditism or parthenogenesis), production of large young, live-bearing (as opposed to egg-laying), and narrow habitat requirements. The importance of each of these (and other) factors is unclear, but it is clear that groups of freshwater invertebrates do differ widely in their degree of endemism. **Figure 4** shows two examples. Dragonflies, most of which are strong fliers and easily cross drainage divides, are much less likely to have small ranges than pearly mussels, which do not readily cross drainage divides. Even the same group of animals may have dramatically different degrees of endemism, depending on the habitat occupied. Thus, groundwater cyclopoid copepods have very much smaller ranges than their relatives in surface waters. This difference presumably arises because groundwater animals have distinctive behaviors and especially because dispersal between aquifers is more difficult than dispersal between lakes or streams.

Traits of Narrowly Endemic Species

Narrowly endemic species may possess traits (other than small range size) that influence their vulnerability to endangerment. For example, the limited dispersal abilities and specialized habitat requirements of many narrowly endemic species may make them particularly sensitive to and slow to recover from catastrophes, whether natural or human-caused. Further, some environments that contain endemic species may encourage the development of traits that influence species vulnerability. Thus, in the food-poor groundwater environment, many animals have sparse populations, delayed maturity, and low reproductive rates, all of which probably add to their sensitivity to human impacts.

Sparse Populations

Species may be vulnerable to endangerment because their populations are sparse. Because population densities of freshwater invertebrates are much less well known than their geographic ranges, relatively little is known about the

Box 1 Lake Baikal, Russia

Lake Baikal, part of a rift system in southeastern Siberia, is the oldest (>25 million years old) and deepest (>1600 m deep) lake in the world, and covers 31,500 km^2. It is the only great rift lake that is oxygenated to its bottom, allowing colonization of the entire lake by a wide range of invertebrates. So far, more than 1400 species of invertebrates have been found in the lake, about 60% of which live nowhere else in the world (Kozhov, 1963; Kozhova and Izmest'eva, 1998). Many genera and several families of invertebrates (including the Lubomirskiidae [Porifera], Baicalarciidae [Turbellaria], Baicaliidae [Gastropoda], Benedictiidae [Gastropoda]) are endemic to Baikal. Probably the most remarkable group of endemic species in Baikal is the huge flock (seven families, 72 genera, and >350 species) of endemic gammaroid amphipod crustaceans (**Figure 7**). These amphipods, which constitute about half of all known gammaroid species in the world, have diversified into a wide range of morphologies and behaviors, and occupy a wide range of ecological niches, including planktonic and benthic herbivores, detritivores, and predators, and semiparasites on sponges. New species of invertebrates are found regularly in Baikal, so it is clear that the true diversity in the lake is even higher than these figures imply. A distinctive feature of the Baikal fauna is that, while nearshore areas contain a mixture of endemic and widespread species, the open water and abyssal sediments are inhabited chiefly by Baikalian endemics.

In many ancient lakes (e.g., Victoria in Africa, Biwa in Japan, Lanao in the Philippines), pollution, habitat destruction, overfishing, and introductions of nonnative species have extinguished many endemic species (Martens *et al.*, 1994; Rossiter and Kawanabe, 2000). Baikal has been protected by its remoteness and vast size, and its fauna has so far been relatively unaffected by human activities. Nevertheless, industrial and domestic waste and siltation arising from deforestation of the catchment have polluted nearshore areas. Although lakewide water quality seems not to have suffered yet, pollution is a concern in Baikal because the long residence time of water in the lake means that contaminants entering Baikal may remain in the lake for a very long time. In addition, nonnative species (e.g., including several fish and the aquatic plant *Elodea canadensis*) have the potential to affect the nearshore fauna. Baikal is a remarkable example of an ancient lake with a richly endemic invertebrate fauna which may yet be preserved through careful management.

References

Kozhov MM (1963) *Lake Baikal and its Life*. The Hague: W. Junk.
Kozhova OM and Izmest'eva LR (eds.) (1998) *Lake Baikal. Evolution and Biodiversity*. Leiden: Backhuys.
Martens K, Goddeeris B, and Coulter G (eds.) (1994) Speciation in ancient lakes. *Advances in Limnology*, vol. 44. Stuttgart: E. Schweitzerbart'sche Verlagsbuchhandlung.
Rossiter A and Kawanabe H (eds.) (2000) Ancient lakes: Biodiversity, ecology and evolution. *Advances in Ecological Research*, vol. 31.

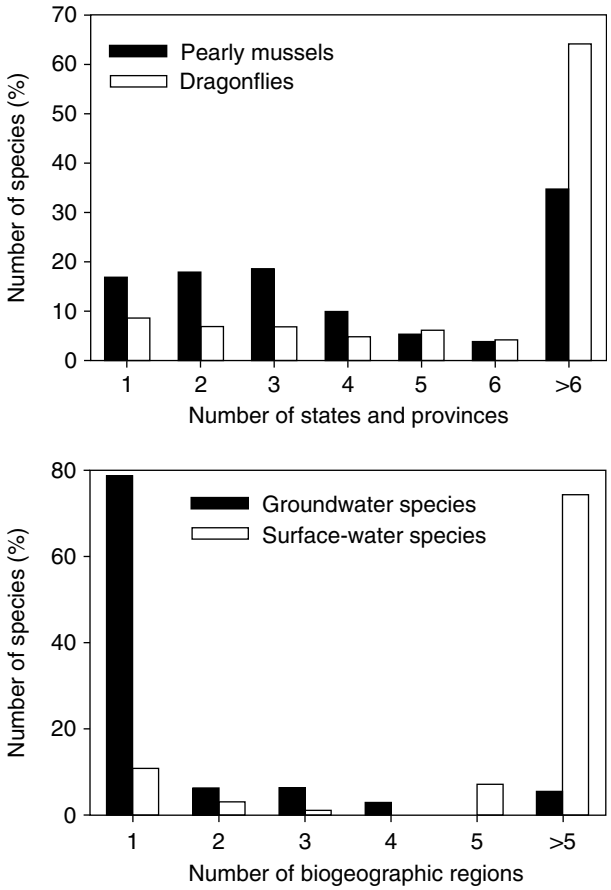


Figure 4 Different kinds of animals have different characteristic range sizes. The upper panel shows the number of states or provinces occupied by species of North American dragonflies and pearly mussels. The lower panel shows the number of biogeographic regions occupied by groundwater and surface-water cyclopoid copepods in Europe.

occurrence of sparse populations of freshwater invertebrates. Likewise, relatively little is known about the causes of population sparseness, or about how population density per se affects the probability of species extinction. Generally, large-bodied animals have lower population densities than small-bodied animals, and predators have lower population densities than their prey, although many exceptions exist to these generalizations. Further, population densities often are lower near the edge of a species range than in its center, and may be lower in unproductive habitats than in more productive habitats. Thus, we might expect to find sparse populations especially in large-bodied invertebrates and in unproductive habitats like deep lakes and ground waters.

Pressure of Human Activities

Human activities endanger freshwater invertebrates in many different ways (Dudgeon *et al.*, 2006). Here, five broad classes of important activities that are especially important in altering fresh waters and endangering their inhabitants are simply recognized.

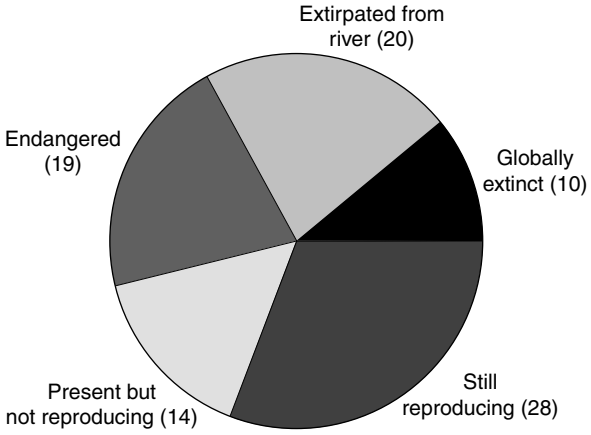


Figure 5 Status of the 91 species of freshwater mussel species that formerly occurred in the Tennessee River (USA), which now consists largely of a series of reservoirs. Reproduced from Benz GW and Collins DE (eds.) (1997) *Aquatic Fauna in Peril: The Southeastern Perspective*. Decatur, GA: Southeast Aquatic Research Institute, Lenz Design and Communications. Special Publication 1.

Habitat Destruction and Degradation

Humans have massively altered the physical characteristics of many fresh waters, usually without considering consequent effects on the biota. These physical alterations probably are the largest cause of endangerment of freshwater invertebrates. Dams have been especially damaging (e.g., **Figure 5**). Above dams, running-water habitats are converted into artificial pools usually unsuited to the native invertebrate species. Below the dam, the water temperature and flows of water and sediment often are so altered that downstream reaches support a highly artificial biota as well. Finally, the dam itself is a barrier that blocks normal migrations and movements of the riverine biota. Thus, even a single dam may endanger the riverine biota for hundreds of kilometers, and many river systems are now dismembered by dozens or hundreds of dams. Over the long term, fragmentation of river systems by dams and other barriers may prevent local populations from recovering from local catastrophes or normal population fluctuations, and thereby increase extinction rates.

In many rivers and streams, floodplains and other shallow-water marginal habitats have been destroyed by dredging, channelization, or filling, or separated from the main channel by dikes and levees. Further, humans often simplify shoreline habitats in both lakes and streams by straightening or filling shorelines and removing "debris" along and in the water. Because marginal habitats often are important for the feeding and spawning of the freshwater biota, the loss or isolation of these habitats may have grave consequences for the native biota. When water is removed from a river for irrigation or held up in a reservoir for hydroelectric generation or flood control, downstream reaches may dry up or lose critically important floods. Several of the world's major rivers (e.g., the Colorado, the Nile) no longer flow to the sea during dry periods. Likewise, the drawdown of many of the world's aquifers from overuse of ground water presumably has major effects on the groundwater biota, although these effects have scarcely even been studied. Finally, mining of underwater

deposits (for gravel or gold, for instance) may have devastating effects both at the site of mining, as well as far downstream through sediment transport and far upstream though head-cutting of the stream bed.

Pollution

Water pollution is another widespread activity with severe effects of freshwater invertebrates. Rivers and lakes often have been used for waste disposal. These wastes include sewage and other organic matter, the decomposition of which may reduce concentrations of dissolved oxygen to levels too low to support most species of invertebrates. Other wastes include substances (e.g., mercury used in gold mining, acid precipitation from power plants and automobiles) that are directly toxic to freshwater invertebrates. Particularly in industrialized regions, long reaches of streams and rivers have been nearly sterilized of invertebrates as a result of severe, chronic pollution (e.g., Figure 6). In addition to pollution caused by deliberate waste disposal, pollution may arise from a wide range of human activities in the watershed. Thus, conversion of forests or native grasslands to agricultural fields or cities typically greatly increases loadings of sediments, nutrients, and toxins that are washed in from the altered watershed. These "nonpoint-source" pollutants are more difficult to track down and control than point loadings of pollutants from factories, yet may have equally serious effects on freshwater ecosystems. While pollution has come under partial control in many developed parts of the world, residual pollution from past releases and inadvertent spills still damage the freshwater biota. Spectacular recent examples include a large spill of pesticides into the River Rhine following a fire at the Sandoz chemical plant, which killed fish and invertebrates for hundreds of kilometers, and the overturning of a truck that spilled a rubber accelerant into the Clinch River, Virginia, which killed most aquatic animals in a 10-km reach, including hundreds of endangered mussels, the largest "take" of endangered species in the United States since the Endangered Species Act was passed in 1973. Of course, in many less developed parts of the world, water pollution is still poorly controlled.

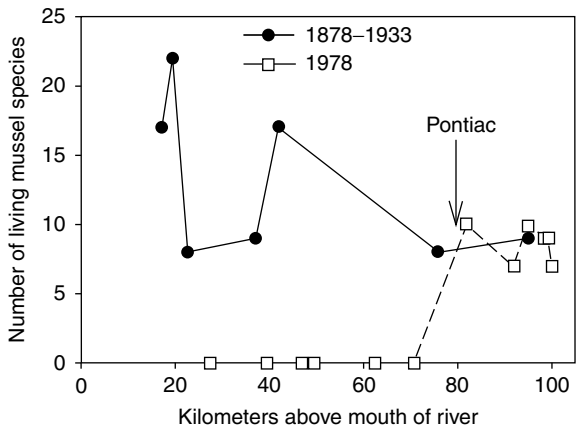


Figure 6 Destruction of the freshwater mussel fauna in the Clinton River, Michigan (USA), in the mid-twentieth century by pollution from the city of Pontiac.

Direct Harvest

Some freshwater invertebrates are harvested for human use, which may contribute to their endangerment. Freshwater mussels have been fished for their shells, pearls, and meat since prehistoric times. Especially in the nineteenth and twentieth centuries, harvest rates became so high that many populations were locally depleted in Europe and North America. For example, over 13 million kg of shells from living unionids were taken from streams and rivers in the state of Illinois in a single year. Harvests of Australian crayfish from the wild reached hundreds of tons year^{-1}, and have contributed to the endangerment of some species. Other invertebrates are collected for bait or the pet trade, which may contribute to local depletion of populations. Invertebrates sometimes are protected by harvest regulations (e.g., closed seasons, size regulations, bag limits), but such regulations may be inadequately conceived and poorly enforced. Fortunately, for economic reasons, harvests usually (but not always) concentrate on common species rather than rare ones.

Nonnative Species

Humans often move species outside of their native ranges. Such introductions may be deliberate, such as the stocking of trout throughout much of the temperate world, or inadvertent, such as the widespread movement of species in ships' ballast water. Whatever the cause, these nonnative species often have strong ecological impacts, and sometimes are responsible for the endangerment of freshwater invertebrates. A spectacular example is the loss of native unionid mussel populations throughout much of northeastern North America as a result of competition with the introduced zebra mussel (*Dreissena polymorpha*). It is projected that the zebra mussel will be the final blow that will drive several species of unionids into global extinction. Nonnative species may also serve as important predators of freshwater invertebrates, as in the case of the brown trout (*Salmo trutta*), which when introduced to Tasmania apparently preyed on and reduced the range of the unusual and endemic anaspidacean crustacean *Anaspides tasmaniae*. Because the effects of nonnative species tend to be cumulative and difficult to reverse, this is a difficult and growing problem in invertebrate conservation.

Global Climate Change

Humans have changed the chemistry of the earth's atmosphere so much that large changes in global climate are expected in the twenty-first century. At this point, it is difficult to make precise predictions about how these changes will affect specific bodies of fresh water. In many bodies of water the changes may be varied and large, involving such diverse characteristics as temperature, hydrology, water level, rising sea level, stratification, the nature and severity of disturbances, increases in damaging ultraviolet light, water chemistry, riparian vegetation, and food quality. Further, in regions where fresh water becomes scarcer while human demands for water continue to grow, human destruction and degradation of freshwater habitats probably will become more severe. Even though we cannot yet specify the details of global climate

change, it is almost certain that this change will endanger or extinguish many freshwater invertebrate species. A rapid, large change in climate will make habitats unsuitable for some of their native species. To survive, such a species will have to disperse to a body of water with suitable ecological conditions. As we have seen, though, the dispersal rates of many freshwater invertebrates are slow, almost surely too slow to keep up with the pace of climate change that current models predict. Further, human modifications to waterways (e.g., impoundments) probably have made long-distance dispersal more difficult.

Table 2 Number of species of freshwater invertebrates included on the 2010 IUCN red list of extinct, endangered, or vulnerable animals, by taxonomic group and continent

Turbellaria (flatworms)	1
Gastropoda (snails, limpets)	634
Bivalvia (clams, mussels)	143
Amphipoda (scuds)	71
Syncarida (anaspidaceans, bathynellaceans)	4
Cladocera (water fleas)	8
Anostraca (fairy shrimps)	25
Conchostraca (clam shrimps)	4
Notostraca (tadpole shrimps)	1
Copepoda (copepods)	74
Decapoda (crayfish, crabs, prawns)	359
Isopoda (sow bugs)	35
Ostracoda (seed shrimps)	13
Other crustacea	5
Ephemeroptera (mayflies)	3
Plecoptera (stoneflies)	3
Odonata (dragonflies, damselflies)	255
Coleoptera (beetles)	23
Trichoptera (caddisflies)	4
North and Central America	542
South America	61
Africa	296
Europe	299
Asia	268
Australia	196
Oceanic islands	12
Total	1665

Number and Distribution of Endangered Freshwater Invertebrates

How many of the world's freshwater invertebrates are endangered, and where do they live? The most comprehensive list of endangered animals available is the IUCN red list, which includes 1665 species of freshwater invertebrates (Table 2). Although an enormous amount of work by experts went into compiling the IUCN list, it is clearly incomplete. The list is dominated by large, conspicuous, and attractive animals (mollusks, decapods, dragonflies, and damselflies). There is no reason to believe that smaller and less attractive animals are less endangered, but there is simply insufficient information on the status and trends of their populations to identify many endangered species. Likewise, most IUCN-listed species are from North and Central America, Australia, or Europe, which probably reflects the geographical distribution of conservation biologists as much as the actual distribution of endangered freshwater invertebrates.

Another way to assess global endangerment of freshwater invertebrates is to examine where damaging human activities coincide with areas of high endemism (Table 3). Areas where freshwater invertebrates are especially likely to be endangered include river systems throughout much of the unglaciated world, ground waters and springs in arid and semiarid regions, and many industrialized areas.

Except in or near areas covered by Pleistocene ice, recently emerged from the sea, or desiccated, river systems often support species that are endemic to that drainage basin. Most river systems have been very highly modified through impoundment and other physical modifications, water withdrawals, and pollution. Thus, we could project that most old river systems probably contain endemic invertebrates, and that many of these invertebrates probably are endangered as a result of human activities (Box 2). Only species that are good dispersers and thus live in many drainage basins or are habitat

Table 3 Expected global patterns of endangerment of freshwater invertebrates as a function of human activities

Activity	*Geographic distribution*	*Groups of animals affected*
Impoundment	Global, especially North America, China, India, arid regions	Many, especially migratory species or those that depend on flooding or turbid water
Physical alterations (diking, channelization, shoreline modification)	Global, especially highly developed regions	Many, especially those that depend on marginal habitats (shallows, floodplains)
Water withdrawal	Global in arid and semiarid regions	Many, perhaps especially species of ground waters and springs
Toxic pollution	Global, especially industrialized regions	Most species
Eutrophication	Global, especially densely populated or farmed regions	Species of lake profundal sediments or plant beds
Harvest	Locally important throughout the world	Bivalves, decapods, sometimes others
Nonnative species	Global in surface waters, perhaps rare in ground waters	Many
Climate changes	Global, especially in high latitudes	Many, perhaps especially species that disperse poorly

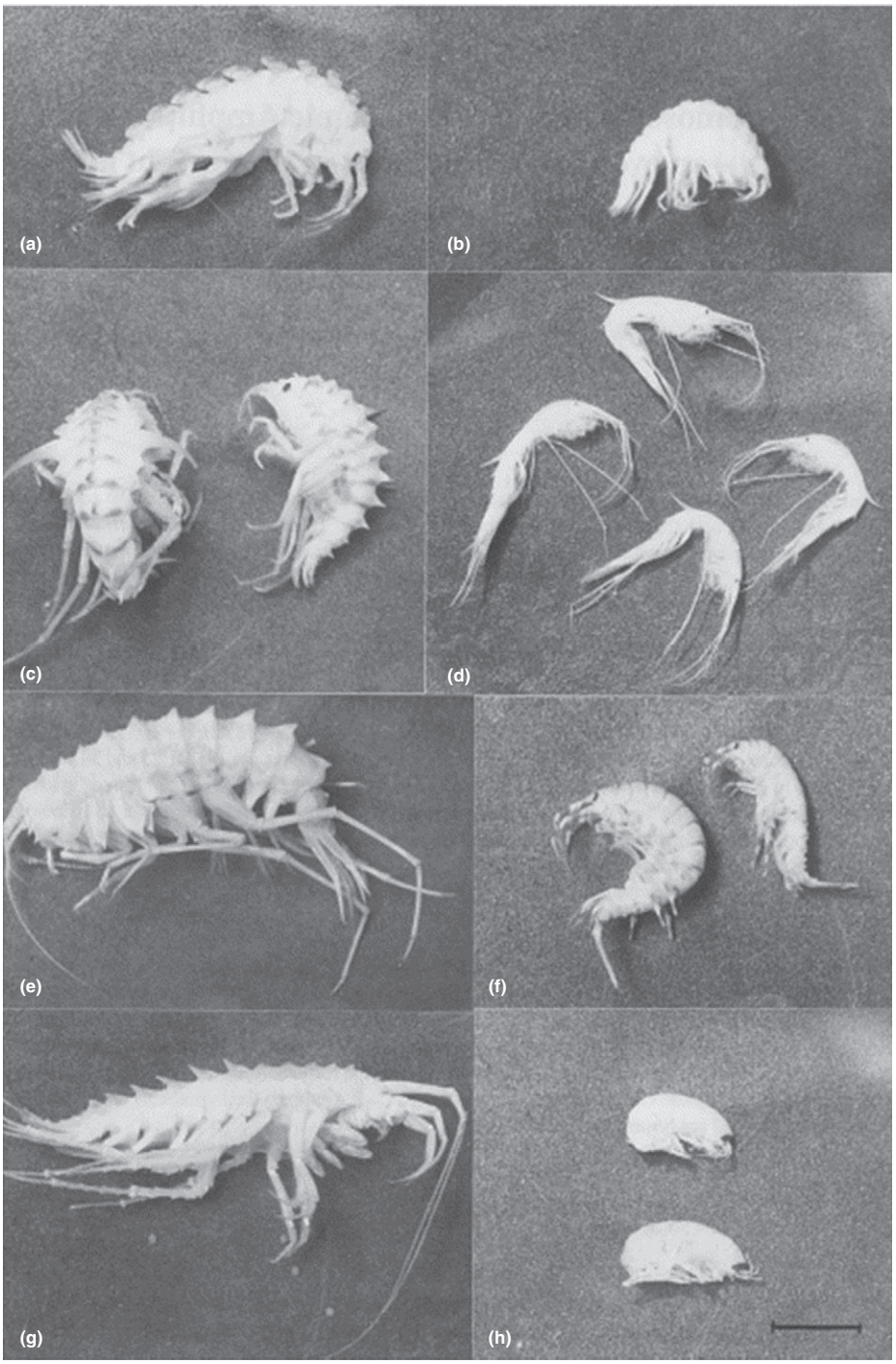


Figure 7 Endemic gammarid amphipods from Lake Baikal, showing some of the wide diversity of body forms. Body lengths are 1.5–6 mm. Reproduced from Salemaa H and Kamaltynov R (1994) The chromosome number of endemic Amphipoda and Isopoda – an evolutionary paradox in the ancient lakes Ohrid and Baikal. *Archiv für Hydrobiologie*, *Ergebnisse Limnologie* 44: 247–256.

generalists and thus occur in nonriverine environments are likely to escape endangerment. Human impacts on rivers are almost global, and even river systems that have not yet been heavily impounded and modified (e.g., some basins in southeast Asia and central Africa) are facing impoundment and other large modifications in the coming decades.

A second area where we might expect to find many endangered invertebrates are ground waters and springs in arid and semiarid regions (e.g., northern Africa, the American Great Plains and Southwest, Australia). These environments typically support highly endemic invertebrate faunas that are probably very sensitive to human impacts. Throughout much

Box 2 Rivers in the American Southeast

The American Southeast (extending roughly from the Ohio River south to the Coastal Plain of Alabama and Georgia, plus the highlands of Arkansas, Missouri, and Oklahoma) contains ancient river systems with an extraordinarily rich biota (Lydeard and Mayden, 1995; Benz and Collins, 1997; Williams *et al.*, 2008). This region was not covered by Pleistocene glaciers, nor was it covered by the sea or desiccated for hundreds of millions of years, so river systems like the Tennessee, Cumberland, and Alabama (and their associated aquifers) are very old. The rivers and ground waters of the Southeast are examples of biologically rich ecosystems that have suffered badly from human activities.

The freshwater invertebrate fauna of the Southeast contains hundreds of species of mollusks, crustaceans, insects, mites, and other animals (**Figure 8**) that are found nowhere else in the world. Dozens of genera are known only from this region. Many of these species have small ranges within the Southeast, and may have occurred in only one stream. Despite a long history of scientific study in the Southeast, new species and genera of freshwater invertebrates are discovered regularly in the region.

As is the case for many river systems, the streams and rivers of the Southeast have been deeply affected by human activities. Impoundments have been especially damaging to the invertebrate fauna. All of the large rivers in the region have been extensively impounded for flood control, hydroelectric power, and navigation, to the point that some of the large rivers have been converted into a continuous series of reservoirs. These reservoir systems differ from natural rivers in their hydrology, temperature, chemistry, sediments, and so on, and often are unsuitable for the native riverine biota. Other physical alterations of stream channels, such as channelization, dredging, diking, and instream gravel mining also have severely damaged the freshwater biota in parts of this region. Further, as in most developed countries, southeastern river systems have been badly polluted by toxins, nutrients, and sediments from industries, farms, and cities. Coal is mined in parts of the Southeast, which brings acid mine drainage and fine sediments into streams. Finally, nonnative species, particularly the zebra mussel and biotic exchanges through the Tennessee–Tombigbee Waterway, may affect the southeastern freshwater fauna.

As a result of these massive changes to southeastern rivers, much of the freshwater invertebrate fauna is extinct or imperiled. Among the mollusks, the only group for which reasonably complete data are available, about 60 species and four genera from the Southeast are now extinct. Literally hundreds of additional mollusk species, representing over half of the native fauna, are threatened or endangered. Hundreds of southeastern crayfish and aquatic insects are likewise rare or endangered, and additional species of small, poorly known animals like copepods, isopods, amphipods, and oligochaetes are doubtlessly extinct or at risk of extinction.

The southeastern fauna is now receiving some protection from the US Endangered Species Act and parallel state laws. Nevertheless, unless the continuing damaging effects of human activities like impoundments are reversed or remediated, it is difficult to be optimistic about the long-term prospects for the southeastern freshwater biota.

References

Benz GW and Collins DE (eds.) (1997) *Aquatic Fauna in Peril: The Southeastern Perspective*. Decatur, GA: Southeast Aquatic Research Institute, Lenz Design and Communications. Special Publication 1.

Lydeard C and Mayden RL (1995) A diverse and endangered aquatic ecosystem of the southeast United States. *Conservation Biology* 9: 800–805.

Williams JD, Bogan AE, and Garner JT (2008) *Freshwater Mussels of Alabama and the Mobile Basin in Georgia, Mississippi and Tennessee*. Tuscaloosa: University of Alabama Press.

of the world, humans living in arid and semiarid regions are pumping water out of aquifers faster than it can be replenished, resulting in large, rapid drops in the water table, which in turn dries up springs and aquifers. We know that desert spring communities are endangered (**Box 3**), and it is possible that aridland aquifers are experiencing large but unseen losses in biodiversity.

In addition to these current threats to freshwater invertebrates, we can expect increasing problems in any regions of rapid human population growth or economic development (e.g., China and southeast Asia), from the wide range of impacts that typically accompany human populations. Further, any rapid changes in climate probably will cause endangerment and extinction of freshwater invertebrates. These climate changes are projected to be most severe in mid- to high latitudes. Because of glaciation, endemism of freshwater invertebrates is higher in midlatitudes than at high latitudes, so impacts of climate change may be most severe at midlatitudes or even in the tropics, especially where the freshwater fauna and environments already have been damaged by human activities. Again, it is the poorly dispersing species that probably will be most severely affected.

There is still no precise answer as to the question of how many freshwater invertebrates actually are endangered (or extinct) globally. The IUCN estimates likely to be an accurate reflection of actual endangerments only for a few conspicuous and well studied invertebrates (e.g., unionoid mussels, odonates). A reasonable guess might be that 10,000–20,000 of the world's freshwater invertebrate species are extinct or endangered as a result of human activities.

Protection of Endangered Freshwater Invertebrates

Although some freshwater invertebrates are protected by international, national, and local regulations, this protection often is inadequate, for several reasons (New, 1995). First, lists of protected species usually underlist invertebrates and include only the largest, most conspicuous invertebrates. For instance, in the United States, only 120 domestic species of freshwater invertebrates are protected under the Endangered Species Act, all but 29 of them mollusks. Thus, many species of endangered freshwater invertebrates, especially small and inconspicuous animals, are not being protected by existing regulations. Second, simply being listed as a protected species may not provide enough help to endangered invertebrates. In many countries, resources for managing endangered species are insufficient, and attention naturally goes to the larger,

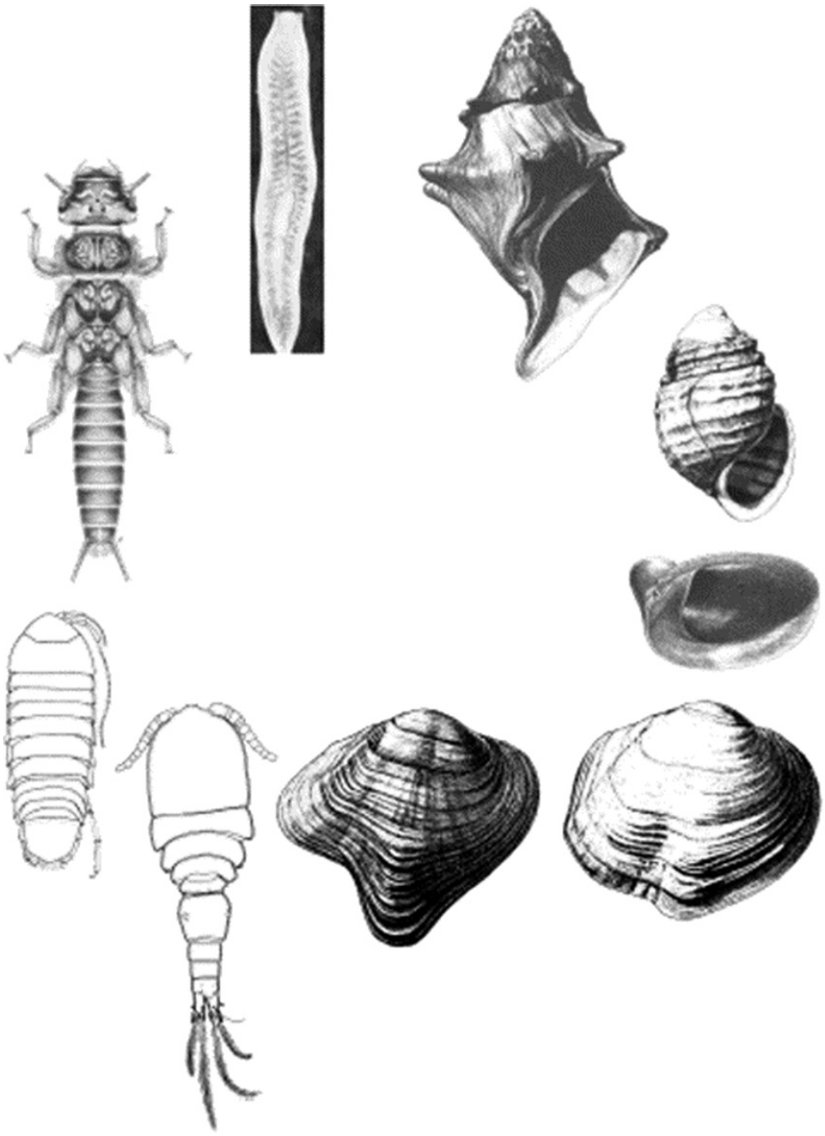


Figure 8 Some endemic freshwater invertebrates from the American Southeast. Clockwise from upper left (with length in parentheses): the flatworm *Sphalloplana holsingeri* (14 mm), the snails *Io fluviatilis* (73 mm), *Gyrotoma alabamensis** (25 mm), and *Amphigyra alabamensis** (2 mm wide), female (50 mm wide) and male (38 mm wide) of the pearly mussel *Epioblasma lewisii**, the copepod *Rheocyclops carolinianus* (0.4 mm), the isopod *Antrolana lira* (20 mm), and the stonefly *Beloneuria georgiana* (22 mm). Species marked* are thought to be extinct.

more charismatic species. In a recent year in the United States, half of all money spent on endangered species was directed to just seven species, all of them vertebrates. As a result, so little money and attention may be spent on invertebrates that plans for recovery of invertebrate species often are general and not pursued aggressively. Third, current approaches to species protection may be inadequate to protect species over the long term. Often, legal protection focuses on trying to prevent further losses from the remaining small populations of an endangered species, without adequate attention to removing the threats that endangered the species in the first place. Consequently, legal protection may slow the rate at which an invertebrate species approaches extinction without reversing its downward trajectory.

How might we more effectively reduce extinction rates of freshwater invertebrates? (Abell *et al.*, 2006; Strayer and Dudgeon, 2010) First, we need to devise ways to protect species without formally listing them. Human activities will endanger or extinguish many species of freshwater invertebrates before we ever gather enough information on their status to satisfy requirements for legal listing, and before we have enough information on their biology to develop effective species-specific plans to protect them. One way to do this is to take advantage of the fact that many species of endemic freshwater invertebrates co-occur in hot spots by protecting such hot spots from the most damaging of human activities (e.g., dams, excessive water withdrawals, toxic pollution). Such a program of hot-spot protection will require better identification, recognition, and protection of hot spots. Nonetheless, it may require less research and provide more effective protection to the world's freshwater invertebrates than existing species-based programs. Of course, not all endangered species

Box 3 Mound springs of the Great Artesian Basin in Australia

Much of the arid interior of eastern Australia is underlain by a large aquifer called the Great Artesian Basin. Fresh- or brackish-water springs occur along the margins of this aquifer in Queensland, New South Wales, and South Australia. These springs range in size from small, moist seeps to large (>100 L s^{-1}), flowing springs, and some of them have built up large (>10 m high) hills of sand and mineral deposits, and so are locally called "mound springs." Although perhaps connected to one another in the past, when the Australian climate was wetter, the springs are now separated from one another by a few meters to many kilometers of desert, and are not connected by streams or rivers.

Like many springs in arid regions, the springs of the Great Artesian Basin support animals that live nowhere else in the world (Ponder, 1986, 1995; Ponder *et al.*, 1989; Ponder and Clark, 1990; Knott and Jasinka, 1998; Perez *et al.*, 2005). Many of these species are found in only one or a few neighboring springs. Only the fish and the snails of these Australian springs have received much study. About 25 species and three genera of snails have so far been found to be endemic to the springs (**Figure 9**). All of the endemic snails belong to the Hydrobiidae, a widespread family that has produced flocks of endemic species in springs, caves, and ground waters in the Balkans, the arid Southwest of the United States and Mexico, and elsewhere. Although these snails may be very abundant in the Australian springs ($>1{,}000{,}000$ m^{-2}), some species are restricted to one or a few springs, and all are highly vulnerable to human impacts. The endemic fauna is thought to have originated by speciation in the more or less isolated springs, perhaps after a more widespread fauna was stranded in the springs by an increasingly arid Australian climate in the Pleistocene.

The chief threat to the spring fauna is from development of wells in the Great Artesian Basin. Because this is an arid region, there is great demand for water for humans, livestock, and mining. When new wells are brought into production, the groundwater level drops, causing springs to dry up. Additional threats include trampling of springs by livestock, which has badly degraded many springs, conversion of springs into pools by excavation or damming, and introduction of non–native species. Over the past three decades, the springs of the Great Artesian Basin have come to be recognized as important habitats for conservation, and steps are being taken to limit at least local impacts from grazing and habitat alterations. Nevertheless, many springs, especially in New South Wales, have dried up as a result of groundwater extraction, and many have been badly altered by livestock or people. It seems likely that at least some of the unique invertebrates of the Great Artesian Basin have gone extinct, and the remaining fauna are at risk of loss.

In arid regions around the world, extreme isolation of aquatic habitats has promoted speciation and development of endemic invertebrate faunas. As in the Great Artesian Basin, water in arid regions is a critically important resource that has been exploited heavily by people. Consequently, freshwater invertebrates of arid regions around the world are endangered by forces similar to those at work in the Great Artesian Basin.

References

Knott B and Jasinska EJ (1998) Mound springs of Australia. In: Botosaneanu L (ed.) *Studies in Crenobiology: The Biology of Springs and Springbrooks*, pp. 23–38. Leiden: Backhuys.

Perez KE, Ponder WF, Colgan DJ, Clark SA, and Lydeard C (2005) Molecular phylogeny and biogeography of spring-associated hydrobiid snails of the Great Artesian Basin, Australia. *Molecular Phylogenetics and Evolution* 34: 545–556.

Ponder WF (1986) Mound springs of the Great Artesian Basin. In: De Deckker P and Williams WD (eds.) *Limnology in Australia*, pp. 403–420. Melbourne: CSIRO.

Ponder WF (1995) Mound spring snails of the Australian Great Artesian Basin. In: Kay EA (ed.) *The Conservation Biology of Molluscs.*, Occasional Paper of the IUCN Species Survival Commission 9, pp. 13–18. Gland: IUCN.

Ponder WF and Clark GA (1990) A radiation of hydrobiid snails in threatened artesian springs in western Queensland. *Records of the Australian Museum* 42: 301–363.

Ponder WF, Herschler R, and Jenkins B (1989) An endemic radiation of hydrobiid snails from artesian springs in northern South Australia: Their taxonomy, physiology, distribution and anatomy. *Malacologia* 31: 1–140.

occur in hot spots, so species-based research and protection will necessarily have to accompany any program of hot-spot protection. Further, we could be more aggressive about identifying and removing threats that endanger species rather than just trying to protect the few populations that have somehow escaped threats. This will require creative thinking about how to preserve or restore essential features of habitat without making unrealistic demands on humans.

A second class of possible solutions could be focused on alleviating the dispersal limitations that are so acute for many freshwater invertebrates by actively establishing new populations of endangered species. This class of solutions is motivated by two main concerns. First, simple protection of existing populations of endangered species may fail to assure long-term survival because natural or human-caused catastrophes (e.g., chemical spills) or normal population fluctuations ultimately may take many isolated populations into extinction. Second, it seems likely that global climate change may be faster than the abilities of some freshwater invertebrates to disperse into suitable habitats. To preserve species under these conditions it may be necessary to deliberately establish populations of endangered species in new locations where suitable habitat exists. This strategy, which is called "assisted migration," has received a lot of attention from conservationists in recent years. However, this strategy has serious problems when applied to freshwater invertebrates because we know so little about these animals (the typical species of endangered freshwater invertebrate in the United States has been mentioned in a total of just one scientific publication!). First, we currently cannot reliably identify "suitable habitat" for most freshwater invertebrates. Second, we do not have good protocols for reintroductions for most species. Third, species introduced outside their native ranges may have unpredictable and undesirable effects on ecosystems and species. Finally, many biologists feel that it is unethical to introduce species outside of their known historical ranges (of course, for many invertebrate species, the known historical range is much smaller than the actual and unknowable historical range). Despite these problems, it may be necessary to confront the problem of species reestablishments, especially if climate change in the twenty-first century is substantial. Clearly, we will need much better information on how fast species are able to disperse in response to a changing climate (to identify which species, if any, will perish without intervention), practical information on how to establish populations of freshwater

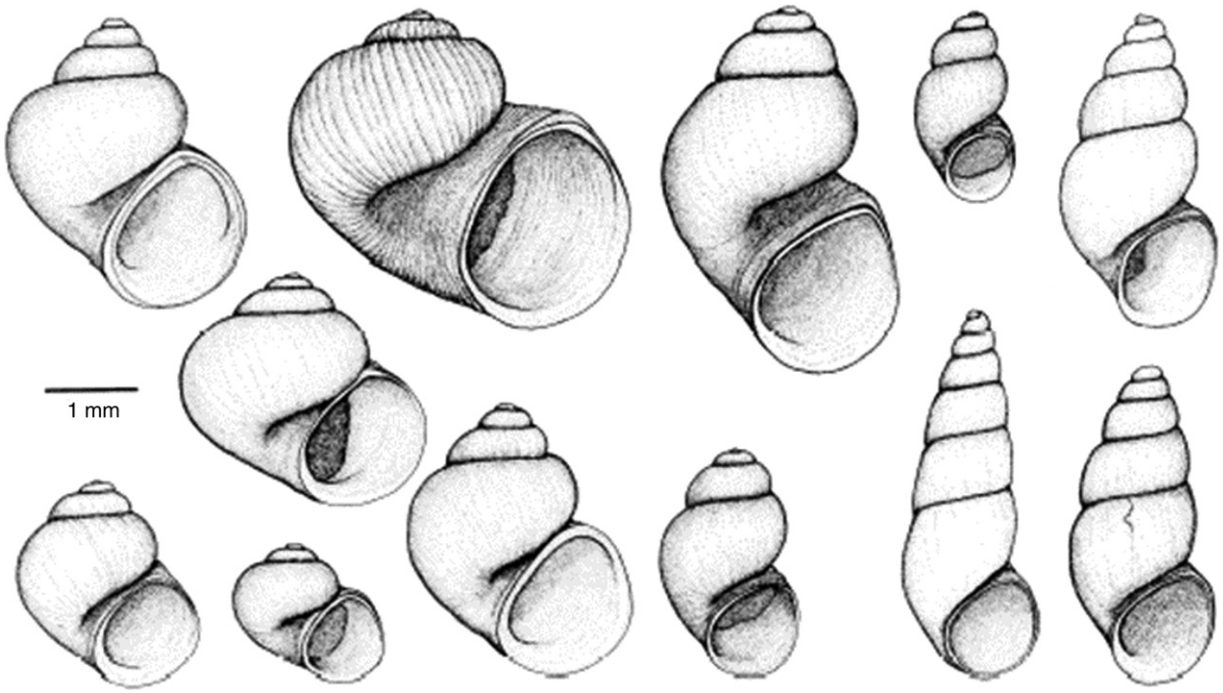


Figure 9 Endemic hydrobiid snails (*Jardinella* spp.) from springs in Australia's Great Artesian Basin. Shells are 1.5–3.5 mm high. Reproduced from Ponder WF and Clark GA (1990) A radiation of hydrobiid snails in threatened artesian springs in western Queensland. *Records of the Australian Museum* 42: 301–363.

invertebrates, and reliable models to predict whether the species we move will have undesirable effects in their new homes.

See also: Endangered Marine Invertebrates. Endangered Terrestrial Invertebrates. Freshwater Ecosystems. Freshwater Ecosystems, Human Impact on. Human Impact on Biodiversity, Overview. Invertebrates, Freshwater, Overview

References

Abell R, Allan JD, and Lehner B (2006) Unlocking the potential of protected areas for freshwater. *Biological Conservation* 134: 48–63.

Abell RA, Thieme ML, Revenga C, *et al.* (2008) Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience* 58: 403–414.

Balian EV, Lévêque C, Segers H, and Martens K (eds.) (2008) Freshwater animal diversity assessment. *Hydrobiologia*, vol. 595, pp 1–637. Dordrecht, The Netherlands: Springer.

Dudgeon D, Arthington AH, Gessner MO, *et al.* (2006) Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews* 81: 163–182.

New TR (1995) *An Introduction to Invertebrate Conservation Biology.* Oxford: Oxford University Press.

Strayer DL and Dudgeon D (2010) Freshwater biodiversity conservation: Recent progress and future challenges. *Journal of the North American Benthological Society* 29: 344–358.

Endangered Mammals

Peter Zahler and Tatjana Rosen, Wildlife Conservation Society, Bronx, NY, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Peter Zahler, volume 2, pp 441–454, © 2001, Elsevier Inc.

Glossary

Anthropogenic Caused by humans.
Depauperate Impoverished, as in a region with low taxonomic diversity.
Endemic Native and restricted to a specific geographic region.
Fossorial Adapted for living or digging underground.
Stochastic event An event that occurs by chance; a random event.
Taxon Any taxonomic group, at any level, that is considered distinct enough from other groups to be considered a separate unit; plural taxa.
Therapsids An order of reptiles that existed during the Late Paleozoic and Early Mesozoic and from which mammals are believed to have evolved.

Mammalian Diversity

Mammals can be found from the arctic to the tropics, on every continent, on many of the smaller islands, and in all of the oceans. There are wholly aquatic mammals (Cetacea), semi-aquatic mammals, terrestrial mammals, arboreal mammals, fossorial mammals, and even aerial mammals (Chiroptera). Mammals fall into three main groups: The monotremes, who lay shell-covered eggs and have a number of physiological structures in common with reptiles, including a single urogenital opening; the marsupials, noted for a suite of physiological traits, including an incomplete placenta, a female bifid reproductive tract, and usually an abdominal pouch for the almost-embryonic young; and the eutherians, who have a chorioallantoic placenta and give birth to relatively precocial young. Mammals may live for more than 100 years, as suspected in some cetaceans, or as short as 1 year, as in the "annual" males of the marsupial *Antechinus*. Mammalian social systems vary from solitary individuals through a wide range of social organizations to the termite-like sociality of the naked mole rat (*Heterocephalus glaber*). Mammals can have individual home ranges as small as a few square meters to systems where groups of tens of thousands migrate together (e.g., wildebeest and caribou) and migrations may cover thousands of kilometers (e.g., gray whales). Mammals have evolved to feed on virtually anything that might qualify as edible, from invertebrates to other mammals, and to fungi, grasses, leaves, bark, and even conifer needles. The variety of niches utilized by mammals and their behavioral and physiological adaptations are enormous – it is difficult to think of a niche not already occupied by a mammal, from aquatic grazer (the Sirenia) to aerial piscivore (*Noctilio* and two other genera of bats).

Not only can mammals be found in most places on earth, but also many species can have dramatic effects on their ecosystems. Classic examples include the alteration of drainage systems and vegetation patterns caused by beaver dams; the alteration of bush and forest to grassland caused by elephant foraging; the alteration of forest pattern and structure caused by sciurid seed predation, hoarding, and burial; and the alteration of grassland pattern and structure caused by grazing by both large ungulates such as American bison (*Bison bison*) and a suite of African ungulates, and small rodents such as prairie dogs (*Cynomys*) and viscachas (*Lagostomus maximus*). Many of these effects have been shown to increase biodiversity of both plants and animals within the region of modification.

Despite the high visibility and appeal of mammals and their significance to ecosystems, we astonishingly still know little about most species. This is especially true for the smaller, more cryptic, and nocturnal groups, such as the rodents, insectivores, and bats, which make up the bulk of mammalian species diversity, but it is also true for larger, more conspicuous species. Of all known species of mammals, 408 (10%) have been discovered since 1993 (Ceballos and Ehrlich, 2009), although a significant number of those have been 'found' through laboratory genetic investigations. New species include small creatures such as rodents and bats, as well as larger species such as monkeys (55 new species), the lesser mouse–deer or Kanchil (*Tragulus kanchil*) (the smallest known deer), and *Prionodiaptomus incarum*, a living representative of Diatomidae, a family considered extinct for 11 million years (My). For many of these and other species, there is not even basic information on population size, geographic range, or even whether the species is still extant. For example, in the case of the Asian elephant (*Elephas maximus*), Blake and Hedges (2004) and Hedges (2006) suggest that the global population 'estimate' of approximately 40,000 to 50,000 Asian elephants is no more than a guess and what is known about the status of Asian elephants is only the location and relative abundance of some (probably most) populations; in some cases it is uncertain whether populations are still extant.

A Brief History

Early Mammalian History

The history of mammals has been traced back approximately 200–250 million years ago (Ma) to offshoots of the therapsid reptiles in the Triassic. According to limited fossil evidence, the

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00218-5

first major radiation of mammals occurred in the Late Triassic, approximately 200 Ma, and subsequently in the Jurassic. Although the diversity of mammals was increasing at this time, with marsupials and placentals in evidence by the Early Cretaceous, these early mammals tended to have small and fairly uniform body sizes. A number of hypotheses have been generated to explain both the increase in mammalian diversity and the small body size. The breakup and then reattachment of the continents allowed for isolation, speciation, and then faunal exchange, and the development of angiosperm plants also undoubtedly resulted in coevolutionary adaptations and speciation by a wide range of mammals. Small body size perhaps was influenced by the dominance of reptiles, which may have forced mammals into a secretive, nocturnal lifestyle where a small body size would be an advantage. An explosive mammalian radiation occurred in the Paleocene, again partly from repeated separations, isolation, and reintroductions as continents continued their movements, and perhaps further spurred by the loss of a wide range of large reptilian competitors during the great extinction event 65 Ma.

The Pleistocene Extinctions

Between 30,000 and 10,000 years ago, a number of sudden and major die-offs occurred almost entirely among large terrestrial mammals. Between 15,000 and 10,000 years ago in North America, a megafaunal extinction event claimed up to 34 genera of large mammals, including mastodons, mammoths, ground sloths, camels, horses, and various large predators such as saber-toothed cats and dire wolves. A similar extinction event involving 52 genera of mammals occurred in South America between 12,000 and 8000 years ago. In Australia, 14 genera of large mammals disappeared between 50,000 to 32,000 years ago. Debate continues on the causes for these extinction events. The hypothesis that they are climate change-driven is overshadowed by evidence that megafaunal extinctions have coincided, nearly without exception, with human arrival. This suggests that the answer lies in the realm of a variety of human-caused situations, possibly exacerbated by climate change. They include rapid overharvesting, biological invasions, habitat transformation, and disease.

Recent Mammalian Extinctions

Many island mammal fauna suffered extinction events similar to those that occurred on the continents during the Pleistocene, although most island extinctions occurred in more recent times. On islands in the Mediterranean, 13 endemic genera vanished, including small goats and even dwarf elephants; most disappeared approximately 4000 years ago. A dwarf form of the mammoth disappeared from Wrangell Island approximately 3700 years ago. In Madagascar, eight genera of lemurs, two genera of pygmy hippos, and what is now considered a distinct order, the Bibymalagasia (comprising two species that are variously aligned with the aardvarks and the ungulates), became extinct approximately 1000 years ago. In each of these cases human activity, either hunting, habitat alteration, or a combination, is thought to be the main cause. Island extinction events have continued into modern times. In the Caribbean Islands, five insectivores, 12 rodents, a raccoon, and a seal have all become extinct since 1500 – almost a quarter of all recorded mammal extinctions during that period (**Figure 1**).

Although the total number of mammalian extinctions in recent times does not appear to be large, the current rate of extinction, approximately 0.01% per year, is anywhere from 100 to 1000 times greater than what would be predicted from

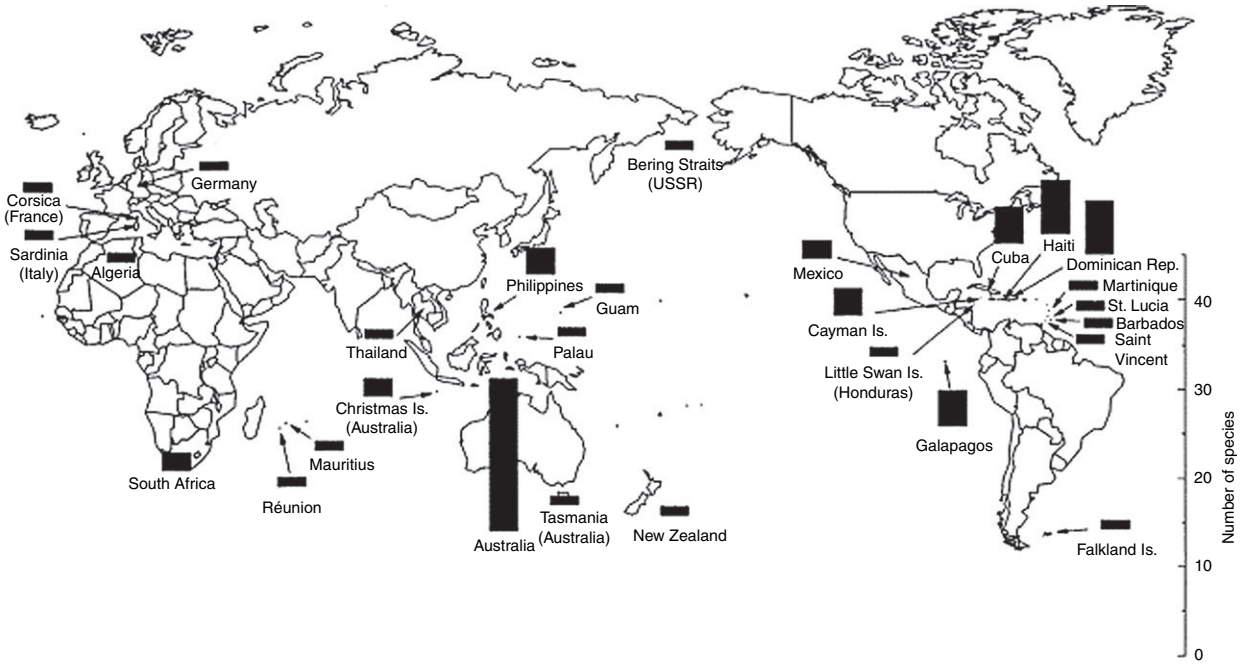


Figure 1 Known mammal extinctions since 1500.

the fossil record, and this rate will very likely increase in the near future. The status of 5488 species of mammals was reviewed in the 2010 IUCN Red List of Threatened Animals. Of those reviewed, approximately 25% were considered at risk of extinction. This percentage is undoubtedly an under-estimate, as the status of many mammal species is still unknown.

Causes and Threats

Many of the dangers that threaten other taxa also threaten mammals. These dangers include habitat loss, exploitation, disease, and exotic introductions, but there are some differences in the manner and level of these threats for mammals as compared to other taxa. For example, human persecution and harvesting pressure is probably greater for mammals than for most other taxa, with the probable exception of fish. True mammalian specialists are rare compared to taxa such as insects, where many species are entirely dependent on only one species of plant and thus are extremely vulnerable to habitat changes that alter community structure. However, many mammals are relatively large, and their habitat needs in terms of area are greater than for most other taxa, putting them at high risk of habitat loss and fragmentation.

A number of historical mammalian extinctions have occurred due to human exploitation or introductions rather than habitat loss. Of 18 mammal species that have become extinct since 1600 and whose cause of extinction is known, eight became extinct due to direct human persecution, eight became extinct as the result of introduced predators and competitors, and only two became extinct because of habitat destruction. At present, many mammalian species are now under a growing threat from habitat changes in addition to direct human persecution. According to the Third edition of the Global Biodiversity Outlook (2010), mammals have suffered the steepest increase in risk of extinction in South and South-East Asia due to the combined impact of hunting and loss of habitat.

Mammalian Physiology and Its Relationship to Threats

Mammals constitute a class of organisms that span an enormous size range of approximately eight orders of magnitude, from 2 g in some bats and shrews to 190,000 kg in the blue whale. However, the majority of mammals tend toward large sizes compared to other taxa, and they tend to have a correspondingly long maturation and slow reproductive rate. This puts them at risk from exploitation and other threats, as breeding rates may not be able to keep up with losses in a population. An extreme example is the blue whale (*Balaenoptera musculus*), whose population has been depleted by at least 70%, and possibly as much as 90%, in less than a century. Despite a halt in whaling that targeted this species in 1966 (identified as the key driver of the decline of the species), the blue whale remains endangered. It has been proposed that its subspecies, the Antarctic blue whale (*B. m. intermedia*), be separately listed as Critically Endangered given the population decline over the same period of more than 97%. Large body size often means large habitat needs, which also puts mammals at risk. Estimates for self-sustaining populations of carnivores such as brown bears (*Ursus arctos*) and African wild dogs (*Lycaon pictus*) range from 1000 to more than 10,000 km^2 of suitable habitat, which means that many protected areas are too small to sustain long-term populations of these species. However, for some species it is not simply a question of large habitat. For tigers (*Panthera tigris*) the actual available prey density is a key determinant of tiger numbers – areas such as India that contain high densities of deer, boar, and other prey have tiger densities roughly 10 times that of the Russian Far East, where prey densities are similarly lower (Miquelle *et al.*, 2010; Karanth *et al.*, 2004).

Habitat Loss, Degradation, and Encroachment

Habitat loss and fragmentation is the single greatest threat to biodiversity worldwide, and this certainly holds true for mammals today. Conversion of habitats by humans into other land uses can fragment and separate mammal populations and increase the likelihood of local population extinctions and eventual species extinction. Rapid deforestation of tropical areas is a growing threat to a number of mammalian species, including many large, wide-ranging, or specialist species of primates, cats, and forest ungulates, as well as numerous small species with restricted ranges such as rodents, insectivores, and marsupials. Most of these species cannot adapt to a highly fragmented or altered landscape, and the few that do adapt may come into conflict with humans by feeding on crops or livestock.

The example of the giant panda (*Ailuropoda melanoleuca*) shows some of the complexities related to fragmentation and habitat loss. Pandas feed primarily on bamboo that may live for decades but then tends to flower, seed, and die en masse within certain areas. When this happens pandas must switch to other bamboo species, often having to move to new locations to find these alternative food sources. According to studies following the latest major bamboo die-off in the early 1980s, pandas were still able to survive by finding patches that had not flowered, moving to new locations or switching to other less-favored species of bamboo. The increase in human population within the panda's range in China has now limited most populations of pandas to very small islands of habitat. Widely separated and very small populations of pandas may not be viable over the long term, even without the problems faced from the fluctuations in their food source.

Because mammals are often relatively poor dispersers, the creation of corridors linking habitats has been suggested as a way to help some species, especially large or wide-ranging (including nomadic or migratory) ones. But for many mammals the necessary size and structure of corridors is unknown, and few management plans have yet to put this idea into practice.

Genetic Loss

It is generally assumed that a population's long-term survival is at least partly dependent on sufficient genetic variation for individual fitness and population adaptability. Loss of genetic diversity and reduced fitness from inbreeding depression and the chance fixation of detrimental alleles has been presumed to reduce adaptive potential and increase the probability of

extinction. However, this has been notoriously difficult to quantify and has not yet been proven to have directly caused a decline in a mammal's population in the wild. Other events, such as disease or predation, often complicate analysis, especially as lowered fitness may increase an individual's susceptibility to these factors. There are some examples for which reduced genetic fitness may well be a cause for declines in mammal populations. The genetic bottleneck that occurred in cheetahs (*Acinonyx jubatus*) in the Pleistocene have been associated with increased fluctuating asymmetry in metric skull measurement, relatively low levels of normal spermatozoa in males, immunologically accepted reciprocal skin graphs between unrelated individuals, dental anomalies and palatal erosion, and an increased susceptibility to infectious disease agents. However, since then, cheetah populations have been reconstituting genetic variation, and current levels of cheetah microsatellite variation approach those of several other outbred populations of felids (Marker *et al.*, 2008). Isolation and reduction in numbers leading to inbreeding of Florida panthers (*Puma concolor coryi*) is considered to be the cause of low sperm counts, malformed spermatozoa, and cryptorchidism (undescended testicles), as well as an increase in heart defects within this population. The genetic restoration effort in the 1990s that introduced animals from Texas has shown signs of improving the genetic fitness of the Florida population.

Another potential genetic threat to endangered mammals is hybridization. Documented cases of hybridizations involving endangered species include Ethiopian wolves (*Canis simensis*) with domestic dogs (*Canis familiaris*), and Asian wild asses (*Equus hemionus*) with domestic horses (*Equus caballus*). Hybridization with domesticated forms or relatives is also seen as a threat to wild yaks (*Bos grunniens*), Bactrian camels (*Camelus bactrianus*), Mongolian wild horses (*Equus ferus przewalskii*), and European wild cats (*Felis silvestris*), among others.

Another genetic issue is the loss of native species and their replacement by common generalist species characteristic of human landscapes. This can result in a local numerical increase in species diversity but a worldwide loss of genetic diversity. One example is the introduction of the American grey squirrel (*Sciurus carolinensis*) into the British Isles and Italy and its replacement of the native red squirrel (*Sciurus vulgaris*).

Livestock and Disease

Domestic livestock and their attendant parasites and diseases have had dramatic effects on some wild mammal populations. Land clearing and overgrazing by domestic cattle, sheep, and goats have altered and fragmented landscapes, increased desertification, and reduced forage for other herbivores. The African rinderpest epidemics of the late 1800s, introduced by cattle, devastated a wide range of wild ungulate populations across the continent. More recently, African horse sickness decimated Indian wild asses, canine distemper killed one-quarter of the lions (*Panthera leo*) in the Serengeti National Park, and all eight packs of wild dogs (*L. pictus*) involved in a study in the Serengeti region disappeared in 1991 with rabies as the suspected or confirmed cause. Disease can spread in the opposite direction, with equally problematic results for wildlife; for example, bison have been shot when they wandered outside of protected areas in the Yellowstone National Park in an attempt to control the spread of brucellosis to cattle. Another example is that of the proposed culling of Mongolian gazelles (*Procapra gutturosa*) in Mongolia to limit the spread of foot and mouth disease outbreaks.

Another factor related to livestock and disease has been the erection of fences to separate wildlife and domestic animals. The results have sometimes been dramatic, such as the case in Botswana, where enormous fences were built in an attempt to control the transmission of disease. These fences blocked the natural migration routes of wild ungulates during the dry season and resulted in the deaths of tens of thousands of red hartebeest (*Alcelaphus bucelaphus caama*), blue wildebeest (*Connochaetes taurinus*), and other antelopes.

Exploitation

Exploitation of mammals has occurred for many different reasons. Subsistence hunting is a growing pressure due to the explosion of the human population, especially in developing countries where much of the populations' dietary protein may come from wildlife. Along with the increase in the human population has come an increase in the number of accurate firearms, including automatic weapons, which are now available to subsistence hunters. Subsistence hunting especially threatens ungulates, such as various deer and tapir species in Asia and South America – more than one third of all deer species and three of the four tapir species are now considered endangered or critically endangered. Subsistence hunting also threatens aquatic mammals whose ranges have already been limited from other causes or who, like sirenians and river dolphins, can primarily be found in coastal or riverine areas. Steller's sea cow (*Hydrodamalis gigas*), the largest and only cold-water sirenian, was hunted to extinction by hungry sailors and other visitors to the north Pacific within 27 years of the first scientific description of this monotypic genus.

Commercial meat hunting has also threatened some species of mammals. The saiga antelope (*Saiga tatarica*), a migratory species of the Central Asian Steppes that once numbered in the millions, was hunted almost to extinction – a nearly 97% decline – in only a few decades. The main threat appears to have been uncontrolled illegal hunting for horns (male horns are exported for the traditional Chinese medicine trade), although habitat degradation and destruction of traditional migratory routes, and hunting for meat in the wake of the break-up of the former USSR in the 1990s, all led to the catastrophic fall in numbers. Recent studies also show that selective hunting of young males and subsequent distortion of sex ratios may have contributed to reproductive collapse.

Until recently the fur trade was an important part of many countries' economies and cultures. In North America, this trade led to the depletion of many fur-bearing species by the early to middle 1800s, including beaver, marten, fisher, otter, and a number of species of seals. Fur hunting for sea otter (*Enhydra lutris*) caused the population of this species to drop from an estimated 300,000 to less than 2000 by 1911, and hunting for skins was the main cause of depletion in numbers for the now-extinct Caribbean monk seal (*Monachus tropicalis*).

More recently, the market for fur has resulted in heavy losses in most wild cat species, including large cats such as tigers and jaguars and the smaller spotted cats of South America and Asia. However, it should be noted that other mammal species, especially wide-ranging, abundant, and fast-breeding species such as the muskrat, raccoon, and coyote, have maintained their numbers despite heavy trapping efforts that continue to this day.

A second result of the fur trade was the introduction of fur-bearing animals outside of their ranges, often with unintended results. The accidental release of American mink (*Mustela vison*) into Europe now threatens the European mink (*Mustela lutreola*), as the American species breeds earlier and appears to be outcompeting its European counterpart.

Commercial hunting is not limited to fur. The hunting of whales for oil as well as baleen led to severe depletions of populations of many species in both hemispheres, with some populations driven to extinction and others lowered to below 5% of their original estimated numbers. Although bans have been in place for most species of whales for some time, many populations have yet to show signs of recovery.

Exploitation for body parts used in traditional medicine has also led to the decline in a number of mammal species. Almost every part of a tiger (*P. tigris*) is used for medicinal purposes, including the feet, fat, bones, blood, testes, penis, bile, whiskers, claws, and tail. The Asiatic black bear (*Ursus thibetanus*) is under great pressure from poaching for its gall-bladder, which is used in traditional remedies, and the poaching of rhinos for their horns, which are used both for knife handles and traditional medicines, is the major cause of their precipitous decline. Although habitat loss, trophy hunting, and persecution played an important role in the initial decrease in tigers, bears, and rhinos, poaching for the medicinal market is now the greatest threat to their continued survival.

Persecution

Persecution usually takes the form of purposeful attempts to control or extirpate a species. Examples include persecution of potential livestock predators such as wolves and tigers, potential livestock disease threats such as antelope species in Africa and bison in the Yellowstone National Park, potential livestock competitors such as prairie dogs in the United States, and, in the case of the American bison in the 1800s, an effort to control and eradicate the indigenous human culture that once depended on this large ungulate. In some cases, control efforts have driven nontarget species to endangered status, such as poison control efforts aimed at coyotes that eradicated the swift fox (*Vulpes velox*) throughout much of its range, and the prairie dog eradication efforts that eliminated the rodents more than 98% of their range and drove the black-footed ferret, a predatory specialist on prairie dogs, to the edge of extinction.

Interference

Interference can occur in a number of ways. As mentioned, fences have interfered with migratory ungulates in Africa and elsewhere. Dams have negatively affected Ganges and Indus river dolphins (*Platanista gangetica and Platanista minor*) by fragmenting populations, increasing siltation, and reducing fish prey – and this may have also contributed to the apparent recent extinction of the baiji or Chinese river dolphin (*Lipotes vexillifer*). For the critically endangered North Atlantic right whale (*Eubalaena glacialis*) injuries from boat collisions and entanglement in fishing and lobster gear have accounted for as much as a third of all mortalities in a year, whereas for manatees in Florida boat collisions have accounted for up to a quarter of all mortality in a year. The population of Hector's dolphin (*Cephalorhynchus hectori*), a small coastal species from New Zealand, continues to decline as death from gillnets entanglement exceeds the already low reproductive rate of the species. Spinner dolphin (*Stenella longirostris*) populations were significantly reduced due to drowning by tuna seiners who targeted the species because of their association with tuna schools. In Florida, the greatest cause of mortality for the endangered subspecies of key deer (*Odocoileus virginianus clavium*) and Florida panther is from being struck by cars. Even tourism may pose a threat. For example, tourist vehicle chases in desert environments have led to death from exhaustion of threatened African wild asses (*Equus asinus*) and addax (*Addax nasomaculatus*) in Northern Africa. The expansion of the whale watching industry has led to concerns that the increasing number of boats, motor noise, and chases may negatively affect whale behavior, including disturbing migratory patterns, breeding efforts, and even separating mothers and young.

War is an extreme example of interference, and it can have an equally dramatic effect on already-rare mammal populations. Although there are arguments that historical tribal warfare may have created source buffer zones for large mammal populations (Martin and Szuter, 1999), today's warfare results in numerous negative effects including habitat destruction from defoliation and bombing efforts and the increase in killing of wildlife due to improved hunting efficiency from the availability of modern firearms. This combination is thought to have had an effect on most wild mammals in Vietnam, including the endangered Douc langur (*Pygathrix nemaeus*). Soldiers and guards may also negatively affect wildlife. For example, the disappearance of most of the remaining population of Marco Polo sheep (*Ovis ammon polii*) along the China–Pakistan border is attributed to target practice and trophy hunting by armed border forces. War may also cause local increases in subsistence hunting and habitat destruction by desperate indigenous and displaced people, especially in and around formerly protected areas. This appears to be the case in Ethiopia and Somalia, where rare antelope and wild ass are heavily hunted, and in Western Africa, where mountain gorillas (*Gorilla gorilla beringei*) have been under periodic pressure from human refugees seeking an escape from fighting as well as food and firewood in protected forests.

Pollution

Most documented cases of pollution affecting wildlife relate to taxa other than mammals, such as DDT and peregrine falcons or acidification and brook trout. There are cases where

pollution has had an effect on populations of threatened or endangered mammals. For example, bats in southwestern North America suffered from DDT spraying, and a number of European bats have been affected by wood preservatives used in buildings that serve as roosts. Most documented mammal-pollution relationships, however, have involved aquatic systems. Perhaps the most obvious example was the death of thousands of sea otters from the effects of the Exxon *Valdez* oil spill in Alaska. Less obvious cases include the St. Lawrence population of beluga whales (*Delphinapterus leucas*) that have been shown to have high levels of organochlorines that have been implicated in cancers, ulcers, infections, and decreased fertility compared to open ocean belugas. Decreased pup production has been noted in seals fed with fish from polluted waters in Europe, and mercury released from mining is thought to pose a serious threat to Amazon river dolphins (*Inia*) and manatees (*Trichechus inunguis*).

Pollution may also have a more subtle effect by increasing a populations' vulnerability to disease by depressing the immune system. Seals have suffered outbreaks of disease in the Mediterranean, northeast Atlantic, and in Lake Baikal. Immune depression has been shown to exist in harbor seals (*Phoca vitulina*) from the Baltic. Pollution is thought to have led to infections from a calcivirus in California sea lions (*Zalophus californianus*) that caused abortion and premature parturition. Dall porpoises (*Phocoenoides dalli*) have been found to have an inverse correlation between serum testosterone level and DDE concentrations in their blubber. Noise pollution, in the form of underwater noise, can affect marine mammals behaviorally and physiologically. In 2000, two minke whales, 15 Blainville's beaked whales (*Mesoplodon densirostris*), and one Atlantic spotted dolphin (*Stenella frontalis*) were stranded in the Bahamas due to noise pollution caused by a US navy exercise, and in 2005,34 short-finned pilot whales (*Globicephala macrorhynchus*), one minke whale (*Balaenoptera acutorostrata*), and two dwarf sperm whales (*Kogia sima)* died in a similar incident in North Carolina.

Introductions

Introductions of exotics, both purposeful and accidental, have had very strong impacts on some native species of mammals. These impacts may involve predation, competition, and the concurrent introduction of diseases or parasites. Most documented introduction problems have involved mammals as the pest species and have taken place on islands. Cases include the introduction of goats, sheep, and other herbivores to islands where endemic plants lack defenses against herbivory, and the introduction of omnivores such as rats, pigs, and macaques or carnivores such as mongooses and stoats to islands where endemic animals lack defenses against predators. The introduction of rabbits (*Oryctolagus cuniculus*) to Laysan appears to have driven three species of birds to extinction from grazing that destroyed avian food sources, both plants and insects dependent on the plants. The results from the explosive population growth of rabbits introduced into Australia are well documented, although predation by nonnative red foxes probably has had a greater negative effect on indigenous small marsupials. Both pet and feral cats and dogs also pose a threat to wildlife. Feral cats in Australia are estimated to kill more than 400 million native mammals, birds, and reptiles each year, whereas in Argentina, predation by dogs is the biggest threat to the endangered pink fairy armadillo (*Chlamyphorus truncatus*). Introduced aoudad (*Ammotragus lervia)* and feral burros in the Southwest United States are considered to be potential competitors and disease vectors for the desert bighorn sheep (*Ovis canadensis nelsoni*). Introduced deer species have had major effects on vegetation in New Zealand, and in South America they may be serious competitors of the endangered Andean huemel (*Hippocamelus bisculus*).

Multiple Threats

For many species, a series or combination of the factors listed has led to precipitous population declines and endangerment. For the tiger, trophy hunting, eradication programs, market exploitation for fur and traditional medicines, and loss of habitat all played a role in the extinction of three races and a decline from perhaps more than 100,000 tigers worldwide in 1920 to as few as 3200 today.

Even for mammals that have recovered from previous threats, new factors may arise to threaten the species. The saiga antelope recovered from commercial hunting once a ban and then regulations were put in place, rebounding from a low of only a thousand in the early part of the last century to more than 2 million in just 50 years. However, poaching for medicinal products, overgrazing, desertification, and migratory barriers now threaten the species, and the population has dropped to around 60,000 animals in five widely separated regions.

Locations for Endangered Mammals

Geographic Analysis of Endangered Mammals

As with most other taxa, mammalian species diversity tends to increase with decreasing latitude. The top eleven countries for mammalian diversity are Indonesia (670), Brazil (648), China (551), Mexico (523), Peru (467), Colombia (442), United States (440), Democratic Republic of Congo (430), India (412), Kenya (376), and Argentina (374). Only the USA is not found entirely or primarily within the tropical zone. The number of threatened mammals is shown in **Table 1**, with Indonesia leading the list with 183 and Mexico following with 100. Most of the listed countries also have some of the highest human densities in the world. However, some of these countries, such as Australia, have high numbers of threatened species yet relatively low human densities. In most of these cases a second factor, endemism, can explain much of the data.

Endemism: The Case of Australian and Island Mammals

Oceanic islands tend to have depauperate mammalian faunas, primarily because terrestrial mammals make poor oceanic dispersers. Many islands also do not have enough habitats to maintain viable populations of terrestrial mammal species.

Table 1 Countries with the most species of threatened (CR, EN, VU) mammals

	Number of threatened mammals	*Total number of mammals in country*	*Percent threatened*	*Total number of threatened endemics*
Indonesia	183	670	27	115
Mexico	100	523	20	80
India	96	412	23	31
Brazil	82	648	13	55
China	74	551	14	17
Malaysia	70	336	21	6
Madagascar	62	228	29	57
Thailand	57	311	19	1
Australia	55	349	22	50
Vietnam	54	359	12	19
Peru	53	467	15	27
Colombia	51	442	19	3
Lao People's Democratic Republic	46	215	21	3
Myanmar	45	294	15	0
Ecuador	43	372	12	14
Cameroon	41	335	12	14
Papua New Guinea	41	269	15	24
Philippines	39	209	19	27
Cambodia	37	164	23	0
United States	37	440	9	

Source: Reproduced from IUCN red list of threatened animals (IUCN Gland, Switzerland, and Cambridge, UK, 2010).

However, on islands that terrestrial mammals were able to colonize, and that were large enough to sustain populations, isolation often led to speciation. Some examples of island endemism are truly spectacular, and often these locations also contain some of the most threatened species in the world. Indonesia, an archipelago consisting of more than 13,000 islands, has the greatest species richness in the world and also has the greatest number of endemic mammal species with 115. Indonesia also has the greatest number of threatened mammals with 183 (27% of its mammalian fauna). In Australia, out of a total of 349 species of mammals, 247 species are considered endemic. Twenty-one species are recorded as extinct since 1500 and, of the remaining species, 55 (22%) are considered threatened with extinction. Madagascar is the fourth largest island in the world and has been separated from mainland Africa since the Cretaceous. Because of its long isolation, all but one of the terrestrial species are endemic forms (the exception is a pig that may have been introduced). There are five families and 15 genera of living endemic primates (the lemurs), an endemic family of insectivores (the Tenrecidae), three endemic subfamilies of the carnivore family Viverridae, at least eight endemic genera of rodents, and a family of bats (the Myzopodidae). Of the endemic species, 57 (31%) are currently threatened, many from the recent loss of 80% of the island's original forest cover. Papua New Guinea has 269 species of mammals, of which 73 are endemic; 41 species are also considered threatened. In the Philippines, more than half (113 out of 207) of the mammal species are endemic, and 27 of these are considered threatened. As mentioned, many of these islands have already suffered extinction spasms: Of a total of 102 mammal species recorded as extinct since 1500,65 are from islands, of which 21 occur in Australia.

Which Mammal Taxa are Most Threatened?

IUCN Red Data Book

Lists such as the Red Data Books are useful in drawing attention to the plight of certain species and in developing conservation and management efforts, including the creation of protected areas. The 2010 IUCN Red List of Threatened Animals attempted to assess the conservation status of every mammal species in the world. Each species was placed within one of eight categories. The first three are categories of threat: critically endangered, endangered, and vulnerable. The next three are lower risk categories: conservation dependent, near threatened, and least concern. The last two categories are data deficient and not evaluated. As well as these categories, two more exist: extinct and extinct in the wild. Out of a total of 5488 species of mammals whose status was reviewed by IUCN, 22% were considered to be globally threatened or extinct. 15% have insufficient data to determine their threat status (**Figure 2**). This compares with 12.5% of birds globally threatened, with less than 1% data deficient. A possible explanation for why 22% of mammals are under direct threat as compared to only 12.5% of birds is that mammals are not as efficient dispersers. Birds in general are more likely to be found over a wide area, and as habitats become fragmented or disappear some populations are likely to survive, and individual birds can also move to new locations more easily than mammals.

The development of programs such as national Red Data Books that cover individual countries or the Natural Heritage Data Centers that cover North and most of Central America has led to a new spatial scaling of endangerment. These local lists often list many more species and subspecies of mammals

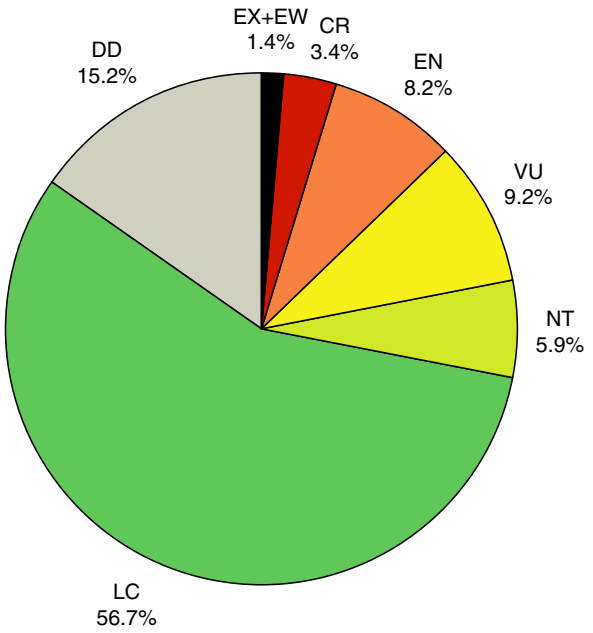


Figure 2 IUCN red list of threatened mammals (2010) showing the categories: data deficient, extinct and extinct in the wild, critically endangered, endangered, vulnerable, near threatened, least concern.

than appear in the IUCN Red Data Book. This is both from inclusion of species that are locally but not globally endangered and endemic subspecies that have yet to receive international attention. The inclusion in a national or regional list of an otherwise-common species whose range barely reaches the borders of a country, may at first seem questionable. However, there is the likelihood that it is a genetically distinct subpopulation adapted to living in what would otherwise be considered marginal habitat for the species.

Endangered Mammals by Taxa

An analysis of endangered mammals by taxa, using the IUCN Red List categories and criteria (2010), shows that among the larger orders of mammals primates are the most endangered, with 49% of the species under threat (critically endangered, endangered, or vulnerable). All primate species are either on Appendix I or II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Primate threats are many and varied, but the principal threat is habitat destruction. Because most primates have arboreal habits, tropical forest loss and fragmentation have had strong negative effects. Captures for the pet trade, zoos, and medical research had a major impact on a number of species until recent legislation limited wild captures. Other species have been affected by subsistence hunting or accidental take from snares set for other species.

Only four other orders have a higher percentage of threatened or extinct species: Sirenia (manatees and dugongs) (100%), Perissodactyla (odd-toed hoofed mammals, such as horses, tapirs, and rhinos) (81%), Monotremata (like platypus and echidna) (60%), and Proboscidea (elephants) (50%); however, these four orders have only five, sixteen, five, and two species, respectively. Two other orders with more than 100 species and a significantly higher than average level of threat are the Cetartiodactyla (whales, dolphins, and even-toed ungulates) (35.9%) and Diprotodontia (marsupials) (35.6%). Although much attention has been given to threatened members of the order Carnivora, "only" approximately one-quarter (26.7%) of the species are threatened. Rodents have approximately 17.5% of species listed as threatened (although comparatively little conservation research has focused on the status of rodent species), but because of the relatively large number of rodent species this order has the highest number of species under threat with a total of 358.

Conservation

Conservation Legislation

An enormous number of treaties and laws govern protection for wildlife or their habitat. These measures include prohibitions or controls on taking, collection, possession, and trade of specific species, control of exotic species, and the protection of habitat, either directly through the creation of protected areas or through indirect measures such as tax incentives, permits, and zoning.

Three levels of conservation legislation exist. The first is multilateral treaties, involving conservation of habitat as in the Convention of Wetlands of International Importance known as RAMSAR, international trade as in CITES, conservation, sustainable use and access and benefit sharing of biodiversity as in the United Nations Convention on Biological Diversity (CBD), and the conservation of terrestrial, avian, and marine migratory species as in the Convention on Migratory Species of Wild Animals (CMS). The second is regional treaties, such as the Bern Convention on the Conservation of European Wildlife and Natural Habitats. The last involves bilateral agreements, such as the 2000 USA and Russia agreement for the long-term conservation of shared polar bear populations in Alaska, in the USA, and Chukotka, in Russia.

Among the multilateral treaties, CITES is especially important for conservation of endangered mammals. CITES was established in 1973 to deal with the enormous level of international trade in wildlife, with estimated earnings of more than US$20 billion. CITES has three appendices, with Appendix I listing species threatened with extinction, Appendix II listing species that may become threatened with extinction unless trade is regulated, and Appendix III listing species protected under national law. Unfortunately, a treaty is only as good as the on-ground enforcement of its provisions, and while CITES has gone a long way toward stemming the tide of illegal trade in wildlife, black market commercial exploitation continues to threaten many species throughout the world.

Species Survival Commissions (SSC)

The International Union for the Conservation of Nature (IUCN) is an intergovernmental organization that counts as members government and nongovernment organizations and volunteer scientists dedicated to the conservation of

biodiversity, primarily through monitoring and advisory roles. IUCN now has more than 100 specialist groups within the Species Survival Commission SSC, each focused on a particular taxa. More than 30 of the specialist groups work on mammals, ranging from multiple orders such as the Afrotheria Specialist Group to individual species such as the Polar Bear Specialist Group. SSC specialist groups consist of volunteer scientists and other experts who provide information on their taxa and even lobby governments for conservation and research efforts. One method of providing this information is through Status Surveys and Conservation Action Plans that describe trends, threats, and conservation options, which most specialist groups have now published for their particular taxa.

Protected Areas

With human population increasing and unmodified habitat rapidly dwindling, habitat protection and maintenance is certainly the single most important conservation method for the preservation of biodiversity, including mammals. Large mammals often live at low densities and over large individual areas. A few protected meadows with the appropriate plant species may be enough habitat to maintain a population of butterflies, but an enormous tract of wilderness is necessary to sustain a viable population of tigers or rhinos. Sadly, many large mammals, such as rhinos and tigers, are now entirely or mostly found within park boundaries, and it is uncertain whether populations are large enough to maintain themselves in the face of stochastic events and anthropogenic changes. Mammals within protected areas often come into conflict with people sharing the land or living on the edges of the protected areas, and poaching may continue to cause declines in mammal populations within protected areas for a variety of social or economic reasons. With climate change, there also exists the possibility that protected areas currently designed around maintaining particular types of habitat will have those habitats altered due to changes in temperature or precipitation, reducing their efficacy for certain species. A new focus on conservation of threatened mammals on the edges of protected areas faces daunting challenges from multiple use, social, cultural, and economic growth issues. However, targeting conservation efforts on the edges of protected areas by engaging local communities may be the only hope for some mammals, as the money and political will to continue creating protected areas is limited.

Captive Breeding and Reintroduction

Captive breeding has become a well-accepted way of managing and increasing populations of critically endangered species. New techniques have led to breakthroughs in captive breeding with some species, including the use of extra-specific surrogate mothers and embryo manipulation, including transfer, cryo-preservation, and microsurgical division. However, captive breeding has not worked in every case. For example, the critically endangered Sumatran rhino (*Dicerorhinus sumatrensis*) has been bred in captivity only after many years of unsuccessful attempts, and the loss of 32 of the original 40 captive animals.

A number of captive breeding programs have led to reintroduction efforts for mammals that were extinct in the wild, including wisent (*Bison bonasus*) in Europe, Arabian oryx (*Oryx leucoryx*) in Oman, and Pére David's deer (*Elaphurus davidianus*) in China. Other reintroduction programs have attempted to supplement decreasing wild populations; for example, one third of wild golden lion tamarins (*Leontopiihecus rosalia*) are now from captive-bred stock. Perhaps the best-known case of captive breeding and reintroduction of a mammal involves the black-footed ferret (*Mustela nigripes*). This small carnivore was once common across the North American plains, where it specialized in hunting large, colonial ground squirrels called prairie dogs (*Cynomys*). Prairie dogs were considered to be pests and competitors with cattle, so an intensive eradication program eventually reduced historic prairie dog range by 98%. The black-footed ferret subsequently declined and was thought to be extinct in the wild in the 1970s, although a small number of ferrets still existed in captivity. This captive colony suffered from physical problems, perhaps related to inbreeding as well as disease, and the colony died out in 1979, leading to fears that the species had truly gone extinct. However, in 1981 a new colony of ferrets was discovered in Wyoming. The first six ferrets to be captured for breeding died of canine distemper. The remaining wild ferrets were captured and successful captive breeding resulted in their numbers reaching approximately 300, with some released back to the wild. However, the program has not been without criticism: Politically motivated decisions and arguments between state, federal, and private organizations have caused numerous problems and may have even jeopardized the success of the project at times. Finding suitable prairie dog colonies (both in terms of finding colonies of adequate size and the political difficulty of maintaining large numbers of what many consider to be a "pest" species) for continued reintroduction is yet another roadblock to ferret recovery.

Another successful and controversial example has been the reintroduction of gray wolves (*Canis lupus*) into the Rocky Mountain region of the USA in 1995 on Forest Service lands in Central Idaho and in 1996 in the Yellowstone National Park. The reintroduced wolf population, numbering more than 1500 in the Northern Rockies of Montana, Wyoming, and Idaho, has had an obvious effect on the ecosystem, causing a decrease in some species (i.e., coyotes, through competition and killing) and an increase in other species (due, at least in part, to an increase in carrion from wolf kills, but also from ripple effects related to changes in elk habitat use and distribution). This reintroduction has not been without controversy, partly stemming from wolf predation on livestock. In 2009, during a single attack on a ranch in Montana, wolves killed 122 sheep, surpassing the number of sheep killed by wolves in the entire state in 2008.

Other predator reintroduction attempts have not fared so well, however. Attempts to reintroduce lynx (*Lynx canadensis*) in the Adirondacks of New York have failed. The decision by French authorities not to augment the population of brown bears reintroduced in the Pyrenees Mountains of France amidst a hostile social context is casting doubt on the long-term viability of the bear population. The success of many reintroductions has also been threatened by legal challenges from individuals or organizations opposed to land-use

restrictions or worried about potential direct conflicts with humans or livestock.

Reintroductions do not always involve captive breeding. A number of African antelope and Eurasian Caprinae have been successfully translocated from still-viable populations into areas where they had been eradicated. One issue regarding this technique is whether there was a distinct local genotype. This is a serious problem if translocations are being considered as a method to bolster an existing but declining or no-longer-viable population, in which case mixing genotypes may result in the swamping of possible adaptive differences. However, in the case of the Florida panther, the controversial program to mix Texas cougars (*Puma concolor stanleyana*) with the Florida panthers has apparently boosted survival of the species: A study shows that hybrids between the two subspecies are three times more likely to survive until adulthood than purebreds, and adult female hybrids survive longer than adult female purebreds (Pimm *et al.*, 2006). There have also been introductions into areas that were not part of the original range of a species, with occasional unintended results of competition with indigenous species and habitat destruction. In one example, 11 Bahamian hutias (*Geocapromys ingrahami*), an endangered rodent, were introduced to a nearby island in a successful attempt to establish a second population of the species. However, it appeared that the hutia had not lived on this second island at least within historic time, and after 16 years the population had expanded to the point where seven plant species had vanished from the island due to overgrazing.

Reintroductions are a complex undertaking, involving a solid understanding of the biology of the species, from feeding and breeding to individual and population genetics. The animals must be trained to survive in the wild, from finding food to socializing, and in the case of arboreal primates, even learning to move through the trees properly. The location for reintroduction should be part of the original range, be large enough to sustain a viable population, contain the correct habitat specifications, and no longer have whatever pressures resulted in the original loss of the species in the wild. They must be free of disease before being released. Introduced animals must be closely monitored, and there must be strong political and local support and extended funding for the project. Thus captive breeding and reintroduction may only be appropriate in specific cases where this kind of information and support is available.

Economic Incentives

Because protected areas only cover less than 13% of the earth's surface, and captive breeding and reintroductions are only feasible for certain well-studied and well-funded species, other methods will be necessary to ensure the survival of a number of mammal species. One method is economic utilization, which applies the fact that mammals have always been heavily exploited. It also has the important potential advantage of paying for itself. Because much of the present extinction crisis of mammals has an economic incentive, this idea has won wide if cautious acceptance, although controversy still exists on issues such as lack of data on population numbers and dynamics, whether overuse is inevitable, and concerns over animal welfare.

Sports and trophy hunting has been linked with historic declines in a number of mammal species, often in conjunction with other factors such as habitat loss. However, sports and trophy hunting is being considered as a management tool today. For some species, especially big game animals, it is suggested that conservation can "pay for itself" when fees from hunters are redirected back into management. This can be most effective for polygynous species where males are not a limiting resource and are also considered desirable as trophies. Licensing and management for white-tailed deer in North America has resulted in a population explosion of that species to the point where it is now necessary to actively control deer as pests in many locations. Trophy hunting fees may also be directed back to local communities to help replace losses from subsistence hunting or nuisance animals and encourage conservation awareness and behavior. Wild goat trophy hunting programs in the mountains of Northern Pakistan have resulted in locally effective conservation efforts on behalf of populations of ibex (*Capra ibex*) and markhor (*Capra falconeri*) – with trophy markhor going for up to US$80,000 an individual, and laws stating that 80% of fees must go back to local communities. In Zimbabwe, the Communal Areas Management Programme for Indigenous Resources (CAMPFIRE) has had success with trophy hunting as an economic incentive for local people to manage their wildlife in a sustainable manner. However, sports hunting as a management tool is only possible when populations of the target species are large enough to support regular culling, monitoring of the target species is constant, regulations are capable of being altered if necessary, and where the political situation enables funds generated from hunting fees to be returned either directly to continued conservation efforts or to local people to encourage their support. From the biological side, potential problems associated with the removal of dominant males include skewed sex and age ratios, lowered genetic variability, altered behavior of other group members, and even the destruction of group cohesion in some targeted species, such as big-horn sheep (*Ovis canadensis*) and elephants (*Loxodonta africana*).

Tourism is another method of raising funds so that the organism under protection pays for itself. This has worked well for a few charismatic species, such as mountain gorillas, but the vast majority of threatened mammal species are not capable of generating a level of public interest that will result in tourism revenues. However, it has been argued that certain charismatic species may act as "umbrellas" to create protection that will cover other species.

Umbrella Species: The Case for and Against Charismatic Megafauna

The umbrella principle of conservation involves the protection of a large or wide-ranging species in the hopes that this will result in the protection of a number of other, smaller species. This concept has numerous advantages: It is easier to mobilize public interest toward large and charismatic species, it is easier to get funding for conservation work, and efforts to conserve large species are likely to result in conservation of

large numbers of smaller species, due to the large habitat needs of the large species overlapping or including those of smaller ones.

The tiger is often used as an example of a species where conservation success will likely mean protection for a number of other species. However, the 38 reserves set up for tigers in India, and four more approved to be set up, have not even adequately protected tigers, and poaching and encroachment threatens the ability of these reserves to protect many other species found within them. Another problem associated with the concept of mammalian umbrella species is that areas of species richness or endemism for other taxa, such as invertebrates, may not be within the designated core area of protection for a large mammal. A third issue of concern is that focused conservation management for a single species, such as antipoaching efforts for tigers, may not satisfactorily protect other species whose main threats may involve other factors such as fire, fragmentation, or the spread of exotic species.

Present and Future Trends

Although mammals get a disproportionate amount of attention and funding for conservation efforts, the future for many species still looks grim. As of 2010, one-fifth of all mammal species were considered to be globally threatened or extinct. For a number of these mammals, a long-term commitment will be needed scientifically, financially, and politically if they are to survive. However, species-based conservation efforts for mammals can have enormous costs. The conservation of the northern subspecies of the white rhino (*Ceratotherium simum cottoni*) was estimated to run approximately US$10,000 a year per rhino, which was not enough to save the population given that the last four rhinos left in the wild have apparently been poached to extinction. Another example is the cost over 7 years for reintroduction of the golden lion tamarin estimated at more than US$1.5 million. The cost of creating and maintaining a protected area that will preserve populations of many mammal species (as well as other taxa) is only a fraction of that needed for species-based management.

Unfortunately, for many endangered mammals there is a desperate need for immediate and focused management efforts. There also is a tremendous need for more protected areas and better protection within the protected areas themselves, and a need for directing conservation efforts toward nonprotected areas by mainstreaming wildlife conservation concerns into land-use planning. For this to succeed, efforts must also be directed toward engaging and educating local people about conservation issues; minimizing wild mammal–people conflicts by deploying conflict mitigation measures and promoting the design of creative insurance schemes and policy changes to stimulate incentives for people to coexist with wildlife; and managing the threat of disease transmission. In addition, illegal wildlife trade, one of the key culprits for mammalian decline, must be curtailed through better enforcement and cooperation and by eliminating the demand for endangered mammal products. In all cases, research and monitoring of wildlife populations and threats is critical, especially in situations where sustainable harvesting is being attempted.

Appendix

List of Courses

1. Conservation Biology
2. Evolutionary Biology and Diversity
3. Mammalogy
4. Vertebrate Biology

See also: Endangered Amphibians. Endangered Reptiles. Mammals, Biodiversity of. Mammals, Conservation Efforts for. Marine Mammals, Extinctions of. Modern Examples of Extinctions

References

Blake S and Hedges S (2004) Sinking the flagship: The case of forest elephants in Asia and Africa. *Conservation Biology* 18: 1191–1202.

Ceballos G and Ehrlich PR (2009) Discoveries of new mammal species and their implications for conservation and ecosystem services. *Proceedings from the National Academy of Sciences* 106: 3841–3846.

Hedges S (2006) Conservation. In: Fowler ME and Mikota SK (eds.) *Biology, Medicine and Surgery of Elephants*, pp. 475–490. Oxford, UK: Blackwell Publishing.

Karanth KU, Nichols JD, Kumar NS, Link WA, and Hines JE (2004) Tigers and their prey: Predicting carnivore densities from prey abundance. *Proceedings of the National Academy of Science* 101: 4854–4858.

Marker LL, Pearks AJ, Wilkerson RJ, *et al.* (2008) Molecular genetic insights on Cheetah (*Acinonyx jubatus*) ecology and conservation in Namibia. *Journal of Heredity* 99(1): 2–13.

Martin PS and Szuter CR (1999) War zones and game sinks in Lewis and Clark's West. *Conservation Biology* 13(1): 36–45.

Miquelle DG, Goodrich JM, Smirnov EN, *et al.* (2010) The Amur Tiger: A case study of living on the edge. In: MacDonald DW and Loveridge A (eds.) *Biology and Conservation of Wild Felids*, pp. 325–339. Oxford, UK: Oxford University Press.

Pimm SL, Dollar L, and Bass Jr OL (2006) The genetic rescue of the Florida panther. *Animal Conservation* http://dx.doi.org/10.1111/j.1469-1795.2005.00010.x.

Endangered Marine Invertebrates

James T Carlton, Williams College and Mystic Seaport, Mystic, CT, USA

Published by Elsevier Inc.

Glossary

Ecotone The transitional region between two habitats or environments.
Keystone A species with a critical regulatory role (often disproportionate to its abundance) in a community.
Planktotrophic Feeding on plankton by invertebrate larvae to survive and grow.
Stenotopic Occupying a highly specialized habitat or tolerating a narrow range of environmental conditions.
Supralittoral The portion of beaches, rocky shores, and other marine habitats that typically lies above the high tide.

We do not know how many ocean species have gone extinct in historical time. Thus, not surprisingly, we have only a very rudimentary concept of how many species are endangered, threatened, or vulnerable, as defined below (*see* Endangerment – Definition *and* Tables 2, 3, 4, and 5). This said, thousands of marine invertebrates alone may be at risk of extinction due to a plethora of human-induced pressures, including physical habitat destruction, water quality, and climate change.

Endangerment – Definition

One set of internationally recognized definitions (IUCN, 2011) treats *endangered* and *vulnerable* taxa under the broad category of *threatened* species. *Critically Endangered* species face an "*extremely high* risk of extinction in the wild". *Endangered* species are at a "*very high* risk of extinction in the wild". *Vulnerable* species are "considered to be facing a *high* risk of extinction in the wild" (IUCN, 2011). To attempt to apply quantitative boundaries to these otherwise seemingly hazy categories, a series of criteria have been erected based on the observed or projected changes in population size, the geographic range of a species, estimated population sizes, and an overall assessment of the probability of extinction (**Table 1**).

Data required to populate these fields include temporal data (e.g., rate of population decline and population fluctuations), spatial data (extent of occurrence and number of populations), and population data (size and number of adults).

In the USA, the federal Endangered Species Act (ESA) defines endangered as a species "in danger of extinction throughout all or a significant portion of its range" and threatened as "likely to become an endangered species within the foreseeable future throughout all or a significant portion of its range" (Environmental Protection Agency, 1973).

Although International Union for Conservation of Nature (IUCN) (and other) definitions focus on global extinction, there are other scales of endangerment as well. These include:

- Local endangerment, when a species faces extinction within a region, as in within a bay or estuary, or from a habitat.
- Regional endangerment, when a species faces extinction in a broad region, such as a biogeographical province.
- Functional endangerment, when a species that was performing a keystone or engineering role (as a competitor, predator, parasite, pathogen, or disturbing agent) in a community becomes so rare that it no longer operates to influence the system structure.
- Commercial endangerment, when a species which is the focus of a fishery becomes so rare that it is no longer economically feasible to hunt it.

It may be noted that certain species that are trade protected under the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES, 2011) may, by virtue of the convention name, be categorized as endangered species. Species that are the "most endangered" are listed in CITES Appendix I ("all species threatened with extinction, which are or may be affected by trade"); species that "are not necessarily now threatened with extinction, but may become so unless the trade in specimens of such species is subject to strict regulation to avoid utilization incompatible with their survival" are treated in CITES Appendix II. Marine invertebrates are treated only under Appendix II and include the mussel *Lithophaga lithophaga*, giant clams in the family Tridacnidae, the queen conch *Strombus gigas*, black corals (Antipatharia), the blue coral *Heliopora coerulea* (Helioporacea), stony corals (Scleractinia), organ-pipe corals (Tubiporidae), fire corals (Milleporidae), and lace corals (Stylasteridae). Although these taxa may thus be referred to as endangered under CITES, many (such as numerous stony coral species) might not be so listed under IUCN or ESA definitions.

Causes of Endangerment and Examples of Endangered Species

Numerous studies in the 1990s and 2000s (e.g., Pew Oceans Commission, 2003; U.S. Commission on Ocean Policy, 2004) defined five general drivers of human-mediated pressures on marine biodiversity: fisheries activities (both overextraction and the mechanical means by which the target species are collected), habitat destruction, changes in water quality, introduction of nonnative species, and climate change.

Table 1 IUCN criteria for critically endangered, endangered, and vulnerable species

Criteria	*Critically Endangered (CR)*	*Endangered (EN)*	*Vulnerable (VU)*
(A) Reduction in population size			
Based on:			
(a) direct observation			
(b) an index of abundance appropriate to the taxon			
(c) a decline in area of occupancy, extent of occurrence, and the quality of habitat			
(d) actual or potential levels of exploitation			
(e) the effects of introduced taxa, hybridization, pathogens, pollutants, competitors, or parasites			
Past			
1. Observed, estimated, inferred, or suspected population reduction (OEISPR), where the causes of reduction are clearly reversible, understood, and have ceased;	>90% over the last 10 years or 3 generations[a]	>70% over the last 10 years or 3 generations[a]	>50% over the last 10 years or 3 generations[a]
2. OEISPR, where the reduction or its causes may not have ceased, or may not be understood, or may not be reversible;	>80% over the last 10 years or 3 generations[a]	>50% over the last 10 years or 3 generations[a]	>30% over the last 10 years or 3 generations[a]
Future			
3. Population reduction projected or suspected to be met;	>80% within the next 10 years or 3 generations[a] (up to 100 years maximum)	>50% within the next 10 years or 3 generations[a] (up to 100 years maximum)	>30% within the next 10 years or 3 generations[a] (up to 100 years maximum)
Past and future			
4. OEISPR where the time period must include both the past and future, and where the reduction or its causes may not have ceased, or may not be understood, or may not be reversible	>80% over the last 10 years or 3 generations[a] (up to 100 years maximum in the future)	>50% over the last 10 years or 3 generations[a] (up to 100 years maximum in the future)	>30% over the last 10 years or 3 generations[a] (up to 100 years maximum in the future)
(B) Geographic range			
Based on at least 2 of the following 3 criteria:			
(a) Severely fragmented or known to exist at only:	1 location	Maximum of 5 locations	Maximum of 10 locations
(b) Continuing decline, observed, inferred, or projected, relative to extent of occurrence; area of occupancy; area, extent, and quality of habitat; number of locations or subpopulations; or number of mature individuals			
(c) Extreme fluctuations in extent of occurrence, area of occupancy, number of locations or subpopulations, or number of mature populations			
In the form of B1, B2, or both:			
B1 Extent of occurrence	<100 km^2	<5000 km^2	<20,000 km^2
B2 – Area of occupancy	<10 km^2	<500 km^2	2000 km^2
(C) Population size relatively small			
Based on:			
Number of mature individuals	<250	<2500	<10,000
And either:			
1. Estimated continuing decline	at least 25% within 3 years or 1 generation[b]	at least 20% within 5 years or 2 generations[b]	at least 10% within 10 years or 3 generations[b]
Or:			
2. Continuing decline, observed, projected, or inferred, in numbers of mature individuals and at least one of the following two:			
(a) No subpopulations estimated to contain:	>50 mature individuals, or at least 90% of mature	>250 mature individuals, or at least 95% of mature	>1000 mature individuals, or all mature individuals are in one subpopulation

(*Continued*)

Table 1 Continued

Criteria	*Critically Endangered (CR)*	*Endangered (EN)*	*Vulnerable (VU)*
	individuals in one subpopulation	individuals in one subpopulation	
(b) Extreme fluctuations in number of mature individuals			
(D) Population size very small or restricted			
Number of mature individuals	<50	<250	<1000
Population with a very restricted area of occupancy or number of locations, such that it is prone to human activities or stochastic events within a very short period of time in an uncertain future, and is thus capable of becoming CR or even EN in a very short time period			Area of occupancy: <20 km^2 Number of locations: <5
(E) Probability of extinction			
Probability of extinction in the wild	At least 50% within 10 years or 3 generations[b]	At least 20% within 20 years or 5 generations[b]	At least 10% within 100 years

[a]Whichever is longer.
[b]Whichever is longer, and up to a maximum of 100 years in the future.
Source: Extracted and reorganized from IUCN (International Union for Conservation of Nature) (2011) *The IUCN Red List of Threatened Species*. http://www.iucnredlist.org/ (accessed November 2011).

Table 2 Marine invertebrates classified as threatened or endangered under the U.S. Endangered Species Act

Species	*Category (date listed)*	*Range*	*Population status*	*Threats*
Haliotis cracherodii (Black Abalone)	Endangered (2009)	Northern California to Mexico	"Locally extinct" in most areas south of Point Conception CA	Abalone diseases, overextraction (historical), climate change, habitat loss
Haliotis sorenseni (White Abalone)	Endangered (2001)	Southern California to Mexico	99% reduction in density since 1970	Overextraction (historical); population reproduction failure
Acropora cervicornis (Staghorn coral)	Threatened (2006)	Florida to Venezuela, including Caribbean	Declines up to 98% throughout range	Coral diseases and climate change, including increased hurricane activity?
Acropora palmata (Elkhorn Coral)	Threatened (2006)	Florida to Venezuela, including Caribbean	In many areas, 90–95% reduction in abundance since 1980	Coral diseases and climate change, including increased hurricane activity?

Source: Reproduced from Environmental Protection Agency (1973) *Endangered Species Act, 16 U.S.C. § 1531 et seq.* http://www.epa.gov/lawsregs/laws/esa.html (accessed November 2011).

The state of our knowledge on how these phenomena and processes impact on an individual basis, let alone their interactions, on any given species, is such that it is often difficult to determine a single major cause of decline for species widely recognized as threatened (Tables 2 and 3). Human-mediated impacts that are the primary cause of endangerment include habitat destruction and fisheries impacts. That said, whether the modern-day surge of critical diseases that have impacted some of the most important reef corals of the Caribbean and the Florida Keys, an abalone disease in California, or stronger and more frequent hurricanes in the Western Atlantic are linked to human-mediated climate change remains unknown (and thus are often discussed without reference to their possible anthropogenic roots), thus potentially obfuscating the overall role of human perturbation. Nevertheless, the possibility that climate change (including coastal warming and ocean acidification) and introduced species may play critical roles in the future in species' endangerment appears clear.

Table 2 presents only four of the threatened and endangered marine invertebrates currently listed under the ESA by the US National Marine Fisheries Service (NMFS) (NMFS, 2011) and the US Fish and Wildlife Service. Table 3 presents the 11 species considered "critically endangered" on the IUCN "Red List" (three of these overlap with the ESA list; one US species, the "critically endangered" coral *Porites pukoensis*, is not so recognized as even a species of concern by NMFS). In all cases, threats are either referenced generically, unweighted multiple threats are cited, or no threats are noted at all, further underscoring challenges in clearly understanding (and predicting) the demise of even those few species that are on higher profile radars.

IUCN lists other marine invertebrates as either endangered or vulnerable: one of these is the North American Pacific coast abalone *Haliotis kamtschatkana*, listed as "endangered," but, like *Porites pukoensis*, not so recognized under US law (and, indeed, not recognized under any category, including "candidate" or "species of concern"). Not appearing on IUCN lists, or any general lists of endangered species, is the Galapagos seastar *Heliaster solaris*, last seen in 1977. A commensal pea crab, *Enigmatheres canfieldi*, likely collected intertidally from a

Table 3 Marine invertebrates listed as critically endangered by the IUCN

Species	*Location*	*Habitat*
Polychaeta		
Mesonerilla prospera	Bermuda	Marine caves
Cnidaria		
Stony corals		
Acropora cervicornis	Western Atlantic	Nearshore reefs (see **Table 2**)
Acropora palmata	Western Atlantic	Nearshore reefs (see **Table 2**)
Rhizopsammia wellingtoni	Galapagos Islands	Nearshore reefs (not seen since 2000; possibly extinct)
Tubastraea floreana	Galapagos Islands	Nearshore reefs (not seen since 2004)
Porites pukoensis	Hawaii (Molokai Island)	Nearshore reefs (known from less than 50 colonies)
Siderastrea glynni	Panama (Uraba Bay)	Nearshore reefs
Fire coral		
Millepora boschmai	Panama	Nearshore reefs (possibly extinct, unless extant in the Indo-West Pacific)
Mollusca: Gastropoda		
Haliotis cracherodii	Eastern Pacific	Rocky shores (see **Table 2**)
Siphonaria compressa	South Africa	Eelgrass specialist in a few lagoons
Crustacea: Amphipoda		
Ingolfiella longipes	Bermuda	Brackish cave

Source: Reproduced from IUCN (International Union for Conservation of Nature) (2011) *The IUCN Red List of Threatened Species*. http://www.iucnredlist.org/ (accessed November 2011).

Table 4 Marine invertebrate extinctions in historical time

Species	*Former location*	*Last known living*	*Habitat/extinction cause*
Mollusca: Gastropoda			
Lottia alveus (eelgrass limpet)	Northwest Atlantic: Labrador to New York	1929	Eelgrass (*Zostera marina*) blades; loss of habitat due to demise of eelgrass, due in turn to epidemic of eelgrass disease (natural resurgence or introduced?)
Lottia edmitchelli (shore limpet)	Southern California	1861	Rocky intertidal; habitat loss
Littoraria flammea (periwinkle)	China	<1840s	Mangrove and other intertidal habitats; probable habitat loss
Cerithidea fuscata (horn snail)	Southern California	1935	Estuarine mudflats; habitat loss

Source: Reproduced from Carlton JT, Geller JB, Reaka-Kudla ML, and Norse EA (1999) Historical extinction in the sea. *Annual Review of Ecology and Systematics* 30: 515–538.

common rocky shore limpet in Monterey Bay sometime before 1873, has not been seen in over 135 years, but is absent from virtually all discussions of marine endangered species (the genus name, newly proposed in 2009, refers to the mystery surrounding the species). The Mediterranean limpet *Patella ferruginea* is described as "the most endangered marine invertebrate species" on the European Council Directive on the Conservation of Natural Habitat of Wild Fauna and Flora but is absent from IUCN lists. These examples of differences in listings – and omissions or exclusions from different international and national registries – further underscore the challenges in properly assessing the status of many marine species and of appropriately assessing their standing independent of the political, economic, or other pressures.

Habitats and Species Susceptible to Endangerment

An examination of Tables 2 and 3 and consideration of those few marine invertebrates known to have gone extinct (Table 4) provides some predictive power for those habitats and species that are at greatest risk (Table 5). Species with highly restricted distributions under heavy fisheries pressure or in habitats strikingly susceptible to human pressures would appear to be low-hanging fruit for investigation of their historical and modern population status. Species in habitats that could be eliminated *in toto* have likely suffered extensively even if their demise has gone unnoticed. A marine habitat that has been extensively altered and extirpated over the past several hundred years globally is the supralittoral zone (variously known also as the wrack, drift, or strand line). This ecotonal world supports – or supported – a large number of species of crustaceans, other arthropods, nemerteans, polychaete worms, gastropods, and other taxa found only in this unique habitat, but they are now replaced in many regions by carparks, seawalls, container ship terminals, housing developments, coastal roads, and so on. Finally, an unknown number of marine invertebrate symbionts, commensals, and parasites associated with endangered and vulnerable marine vertebrates, other invertebrates, and marine plants are without doubt themselves threatened.

Table 5 Marine invertebrates at risk of endangerment and extinction

Examples of risk groups	*Examples of species*
Species with restricted distributions	
Species restricted to lagoons, estuaries, isolated marshes and mangroves, remnant seagrass beds, and other limited or now-restricted coastal habitats	Seaslug *Phyllaplysia smaragda* Seaslug *Stiliger vossi* Limpet *Siphonaria compressa* Amphipod *Transorchestia enigmatica* Many other species in coastal habitats, including estuarine seagrass stenotopic endemics
Species with limited distributions under heavy extraction pressure or under pressure from related extraction activities (such as bottom trawling)	White abalone *Haliotis sorenseni*, certain benthic invertebrates in areas of heavy trawling, endemic coral reef species
Short-range endemics with non-planktotrophic development, especially those under extraction pressure, fisheries activities, or at risk from coastal development and pollution	Many invertebrates species in numerous phyla, including short-range cone snails (*Conus* spp.); endemic coral reef species (all invertebrates); species restricted to islands, offshore banks, and sea mounts
Species restricted to extinguishable habitat	
Invertebrates found only in supralittoral (maritime, strand, and wrack-line) habitats	Talitrid amphipods, oniscid isopods, staphylinid beetles, other beach insects, marine centipedes, pseudoscorpions, nemerteans, polychaetes
Species associated with other endangered species	
Commensals; symbionts; and parasites of endangered marine vertebrates, invertebrates, and plants	Species associated with ● Mammal: *Phocoena sinus* (Gulf of California vaquita) ● Fish: *Syngnathus affinis* (Texas pipefish) ● Fish: *Raja laevis* (barndoor skate) ● Bird: *Oceanodroma macrodactyla* (Guadalupe storm petrel)[a] ● Other endangered marine vertebrates ● Endangered marine invertebrates ● Endangered marine plants

[a]If not already extinct.

Challenges with Determining the Number of Endangered Marine Invertebrates

Already mentioned is that disparate lists of threatened, endangered, and vulnerable lists of marine invertebrates (and all other marine species) exist in different countries and through different international organizations. Further impeding a clear view of the scale of endangerment and extinction is (despite the completion of a recent Census of Marine Life) a very limited picture of the historical baseline of nonedible marine invertebrate populations (accompanied by a striking lack of exploration of historical resources, such as earlier literature and museum collections), and the modern-day (the last 75 or so years) demise of marine invertebrate taxonomy, systematics, biogeography, and natural history. Perhaps underlying much of our general lack of understanding of the scale of endangerment and extinction among marine invertebrates is the enduring assumption that if common species have been on a steady decline, or had entirely disappeared, then "someone would have noticed." Numerous accounts of the overlooked demise of many marine species (and, concomitantly, commonly overlooking additions to marine communities by way of newly invading species) argue otherwise.

Recovery Strategies: Extraction Management, Culture, and Transplantation

The often late recognition that overextracted populations are in severe decline, or are almost extirpated, is typically accompanied by government regulations that close or restrict fisheries long after the target species has passed into functional or commercial extinction. Thus, the endangered abalones of the California coast are now under strict regulation, and IUCN trade rules attempt to prohibit the international movement of many species of corals. Additional approaches to prevent species demise now also include growing attempts to culture species – such as corals – in the laboratory (with subsequent outplanting back in the wild), and proposals to transplant species to new regions (also known as "managed relocation" and "assisted migration"). The latter strategy, although not new, has stirred concerns about the possibility that such transplants could inadvertently lead to species becoming potential nuisance invaders. Thus, the marsh snail *Assiminea parasitologica*, considered a rare, threatened species (due to habitat loss) in Japan, was accidentally introduced in ships' ballast water in the 1990s or early 2000s to the Oregon coast, where it now occurs by the billions, and threatens to replace a native marsh snail, *Assiminea californica*.

Appendix

List of Courses

1. Marine Ecology
2. Marine Biology
3. Environmental Studies
4. Conservation Biology
5. Invasion Biology

See also: Census of Marine Life. Climate Change and Extinctions. Extinction, Causes of. Government Legislation and Regulations in the United States. Invertebrates, Marine, Overview. Marine Conservation in a Changing Climate. Translocation as a Conservation Strategy

References

Carlton JT, Geller JB, Reaka-Kudla ML, and Norse EA (1999) Historical extinction in the sea. *Annual Review of Ecology and Systematics* 30: 515–538.

CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). *Appendices I, II, and III* (2011). http://www.cites.org/eng/app/index.php (accessed November 2011).

Dulvy NK, Pinnegar JK, and Reynolds JD (2008) Holocene extinctions in the sea. In: Turvey ST (ed.) *Holocene Extinctions*, pp. 129–150. Oxford University Press.

Edgar GJ, Samson CR, and Barrett NS (2005) Species extinction in the marine environment: Tasmania as a regional example of overlooked losses in biodiversity. *Conservation Biology* 19: 1294–1300.

Environmental Protection Agency (1973) *Endangered Species Act, 16 U.S.C. § 1531 et seq*. http://www.epa.gov/lawsregs/laws/esa.html (accessed November 2011).

IUCN (International Union for Conservation of Nature) (2011) *The IUCN Red List of Threatened Species*. http://www.iucnredlist.org/ (accessed November 2011).

National Marine Fisheries Service (2011) *Species under the Endangered Species Act (ESA)*. http://www.nmfs.noaa.gov/pr/species/esa/ (accessed November 2011).

Pew Oceans Commission (2003) *America's living oceans: Charting a course for sea change. A Report to the Nation,* 143 pp. Arlington, Virginia: Pew Oceans Commission.

Roberts CM and Hawkins JP (1999) Extinction risk in the sea. *Trends in Ecology and Evolution* 14: 241–246.

U.S. Commission on Ocean Policy (2004) *An Ocean Blueprint for the 21st Century. Final Report*. Washington, DC. http://www.oceancommission.gov/

Endangered Plants

Thomas J Stohlgren, U.S. Geological Survey, Fort Collins, CO, USA
Sunil Kumar, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Thomas J. Stohlgren, volume 2, pp 465–477, © 2001, Elsevier Inc.

Glossary

Critical habitat Habitat of a threatened or endangered species that is itself threatened by destruction, disturbance, modification, or human activity, potentially resulting in a reduction in the numbers, distribution, or reasonable expansion or recovery of that species.
Endangered species Those species in danger of extinction throughout all or a significant portion of their range.
Endemism Condition in which a species' distribution is restricted to a given geographic region.
Rare species Species with small populations globally that are not presently listed as endangered or vulnerable, but are at risk because of their small population size.
Threatened species Those species that are likely to become endangered in the foreseeable future throughout all or a significant portion of their range.

Introduction

Extinction has always played an important role in nature: 95–99% of all species that ever existed are now extinct. It is the current rapid rate of extinction that has many plant ecologists worried. A recent global analysis by Royal Botanic Gardens, Kew together with the Natural History Museum, London and IUCN showed that over one fifth world's plant species are threatened with extinction (IUCN, 2010). Worldwide, as many as 654 species of plants have gone extinct since A.D. 1600 (Heywood, 1995). In this time period, about 110 plant species may have gone extinct in Hawaii according to records of the U.S. Fish and Wildlife Service and The Nature Conservancy. Estimates of plant extinctions in the continental United States range from below 25 (IUCN, 1998) to above 90 (Davis *et al.*, 1986; Flather *et al.*, 1994). The discrepancy in estimates is due to incomplete systematic surveys, limited monitoring activities, and poor information on viable seed banks. Estimates of extinctions in other countries include 71 plant species in Australia, 53 in South Africa, 47 in Mauritius, and 23 in Cuba. Three endemic species of ebony (*Diospyros* spp.) went extinct in the Mascarene Islands before they were described. Nuttall's mudwort (*Micranthemum micranthemoides*) had been recorded in Delaware, Maryland, New Jersey, New York, Pennsylvania, Virginia, and the District of Columbia, but, despite extensive searches it has not been seen since 1941. The single individual of *Dicliptera dodsonii*, which clings to life in western Ecuador (Gentry, 1986), may be the next victim of accelerated extinction.

Estimates of future plant species extinctions vary widely. Raven *et al.* (1986) estimated that about 40,000 tropical plant species may go extinct in the wild within the next several decades. The New York Botanical Garden suggests that about 700 of the more than 2000 species of threatened and endangered plant species in the United States may be extinct in the next 10 years. The Center for Plant Conservation reported that 680 U.S. plant species were critically endangered, with approximately 253 species estimated to become extinct in 5 years and 427 species to become extinct in 10 years (based on unpublished data in 1988). Today, more than 20 years later, there is little evidence of these mass extinctions, but the concern may still be well founded in several areas.

Hawaii provides a good example of endangered plants. Up to 47% of the Hawaiian flora may meet the criteria for official listing by the U.S. Fish and Wildlife Service, with as many as 16% immediately threatened with extinction (see the Missouri Botanical Gardenweb-site listed in the References). The genus *Hibiscadelphus* in Hawaii includes only six species (with two extinct species) with a total of 14 live individuals, and one species with only one individual (Cody, 1986). The Hawaii Endangered Plant Task Force, which includes many federal, state, and non-government partners, now tracks 597 threatened, endangered, and rare plant species.

There is a growing backlog of candidates for listing as endangered species. Between July 1976 and August 1992, about 21 species per year were added to the U.S. endangered species list. In the second half of that 16-year period, 33 species per year were added to the list (Flather *et al.*, 1994). More than 200 plant species are being petitioned for future listing in the United States.

It has been said that "all species are rare somewhere" (Orians, 1997), because most plant species have larger populations in well-suited habitats and smaller populations in marginal or suboptimal habitats. Many threatened vascular plants in Finland were found to be rare largely because the species were in marginal rather than suitable habitat (Cropper, 1993). Along with the increasing threats of rapid environmental change, habitat loss, contaminants, and invasive exotic species, we may face an uphill battle in protecting our treasured botanical resources in these suboptimal habitats.

Maintaining biodiversity requires a considerable understanding of rarity and the processes and stresses that promote the endangerment of certain plant species. This article reviews the kinds of rarity, patterns of endangerment, causes of endangerment, consequences of rarity, and legal mandates to protect endangered plant species. It then provides selected examples of endangered plants, an example of a technique used for detection and mapping of endangered plants, and

concludes with the management and social implications of protecting endangered plants.

Kinds of Rarity

There are many published definitions of rarity. In an extreme example, DuMond (1973) stated nine criteria for rarity, including species that are: (1) found out of its expected context; (2) particularly subject to extinction or severe reduction in total population size by human activities that have already caused a significant population reduction; (3) found only in a very specific habitat of limited occurrence; (4) thought to be a relict of a no-longer extant vegetation association; (5) an indicator of a unique extant vegetation association; (6) recognized as an example of a wide, disjunction pattern; (7) at its natural distribution limits within the area in question; (8) known to be introduced and has become naturalized only on a very small scale; or (9) does not consistently occur as a member of any particular natural plant community. Gaston (1997) described rarity in another way, noting that rare species can be delimited on the basis of one, two, or at most a few of the following variables: abundance, range size, habitat specificity (habitat occupancy), temporal persistence (e.g., taxon age), threat (probability of, or time to, extinction), gene flow, genetic diversity, endemism, and taxonomic distinctness. Rare species in the IUCN Red Book (1998) are described in terms of population size rather than an assessment of extinction risk, which is reserved for threatened and endangered status. The Nature Conservancy and associated network of Natural Heritage Programs rate species endangerment based on five levels of global rarity (i.e., G1, G2, G3, etc.), national rarity (N1, N2, etc.), and state rarity (S1, S2, etc.).

The kinds of rarity that concern conservation biologists relate specifically to the potential vulnerability to extinction. Commonly recognized kinds of rarity (Rabinowitz, 1981; Cropper, 1993) include species with:

- Small populations, large geographic range, and occurrences in several habitats. For example, American chaffseed (*Schwalbea americana*) had a historical distribution from Mississippi to Massachusetts. It resides now in 20 relatively small populations in five states, with the greatest populations in South Carolina. It can survive in several vegetation types as long as there is enough light in the understory. Fire suppression, which allows for succession and canopy closure, threatens several *Schwalbea* populations. Another example is *Hypochoeris maculata*, which has small populations over a wide range in Great Britain.
- Large populations, large geographic range, but occurrences in specific habitats. For example, several grassland species of *Calochortus* (lilies) in California have large populations following wildfires. They were thought to be far more widespread prior to a century of fire suppression and encroachment of shrubs into grassland habitats. Sparse populations in a large geographic range but in specific habitats. For example, *Psilotum nudum* grows only on rocky outcrops and always in sparse populations, but it occurs in Australia, New Zealand, and Easter Island (Cropper, 1993).
- Small geographic range, but locally abundant in specific habitats. For example, the 48 endemic species of Bignoniaceae in Amazonia have a total home range of only a few thousand square kilometers, but some of these species are locally abundant in restricted habitats (Gentry, 1986). A specific example in Scotland is *Primula scotica*, which has a tiny range with a few large populations.
- Small geographic range, and sparse in specific habitats. This may be the most critical type of rarity, for small populations restricted to small and specific habitats are highly vulnerable to extinction. For example, the scouring of a new stone quarry in South Africa ended life in the wild for *Moraea loubseri*, a small, sparse-population iris (Stermer, 1995).
- Small populations of concern in any region despite populations elsewhere. For example, many floristically poor northern European countries typically have 200 or more plant species listed as threatened or endangered. However, many of these species invaded broadly since the last Ice Age, and they are buffered from extinction with widely scattered populations in several countries (Davis *et al.*, 1986).
- Small populations of new species. For example, a newly discovered species of *Calyptranthes* from El Yunque, Puerto Rico, has a population of four trees, and *Auerodendron pauciflorum* from Quebradillas and *Calyptranthes luquillensis* from the Luquillo Mountains have only five individuals each (Cody, 1986).

Generally, species may be ranked for protection based on overall rarity, magnitude of potential threats to populations, and the immediacy of the threats. Thus, a monotypic genus with high-magnitude and imminent threats might receive a higher priority ranking than a subspecies with moderate or nonimminent potential threats.

Endangerment Patterns

Slightly over 90% of threatened plants are single-country endemics. Species with restricted ranges face the greatest threat of extinction. It follows that endemic plant species on islands appear to be the most highly endangered. About 95% of the plant species on the Canary Islands are endemic, and 50% are considered endangered (Cody, 1986). *Senecio heritieri*, for example, is restricted to a small area of rocky slopes on the south coast of the island of Tenerife, one of the Canary Islands. On Crete, two-thirds of the 155 endemic species are endangered (Cody, 1986). Compare these numbers to those of the entire United States, where there are about 20,000 plant species of which 2050 species are rare and threatened.

About half the plant species in Mediterranean climate areas (parts of California, South Africa, Australia, and the Mediterranean basin) are narrow endemics that dominate the threatened and endangered species lists in their various countries (**Figure 1**) (Davis *et al.*, 1986). California, for example, contains 669 endemic species of the 2050 species on the U.S. rare and threatened list (Davis *et al.*, 1986). Chile has 50 forest tree species; 47 of these species are endemic and 38 species are listed as endangered, vulnerable, or rare.

Over half of the rare and endangered plants of the continental United States grow within the borders of 12 western

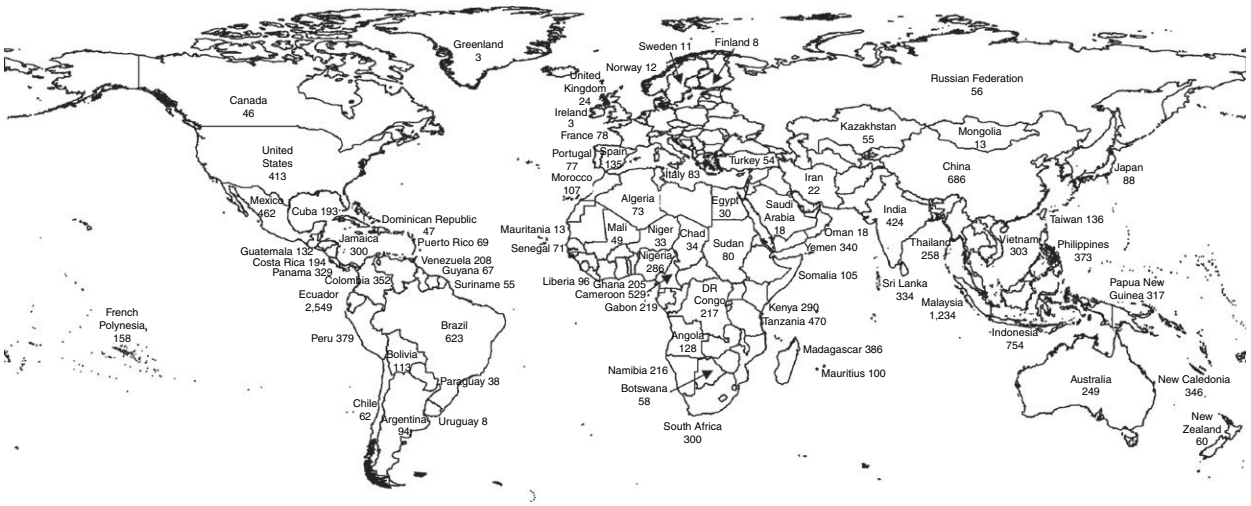


Figure 1 Distribution of rare plants in the world. Data from IUCN (2010) *IUCN Red List of Threatened Species*. Version 2010.3. http://www.iucnredlist.org. Downloaded on 05 October 2010.



Figure 2 Distribution of rare plants in the continental United States. The ''top 10'' continental states for rare plants include six western states, and California clearly dominates. Reproduced from Flather CH, Knowles MS, and Kendall IA (1998) Threatened and endangered species geography. *BioScience* 48: 365–376.

states (Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Texas, Utah, Oregon, Washington, and Wyoming; **Figure 2**) (Flather *et al.*, 1994). However, among U.S. states Hawaii is the hot spot of endangerment: about 950 native plant species of Hawaii are extinct, endangered, or threatened (Raven *et al.*, 1986).

Endangerment patterns vary by habitat type. Flather *et al.* (1994) reported that 44% of threatened and endangered plants in the continental United States are found in rangelands. Twenty-five percent of endangered plant species were associated with barren land or rocky habitats. However, wetlands, which comprise only 5% of the land base, contain 15% of the listed species.

Endangerment patterns also vary by region and land-use characteristics. In northern California, southern Nevada/Sonoran Basin, and the eastern Gulf Coast, >50% of the listed species are associated with wetland habitats. U.S. Department of Defense lands, which comprise only 3.4% of federally administered lands, contain 26% of the listed species. The high number of rare plants in the U.S. Southwest when combined

with rapid development and land-use change, leads to special problems for the western states that wish to conserve their floral heritage.

Worldwide, Australia and New Zealand have attracted much attention. Official listings show 1931 endangered plant species in Australia and 22 species in New Zealand (Stermer, 1995). Australia has already lost 117 plant species to extinction in modern times (Davis *et al.*, 1986), and has 3329 plant species that are considered rare or threatened (6.2% of the flora). An estimated 993 species are at risk of extinction in the next 50 years (Cropper, 1993).

Some plant families and genera are more prone to extinction than others. Twenty vascular plant families have at least 50% of their species threatened (e.g., Brunelliaceae, 91.9%; Zamiaceae, 88.9%; Araucariaceae, 78.9%; Taxaceae, 75%; Limnanthaceae, 72.7%). In Australia, species-rich genera such as *Acacia*, *Eucalyptus*, and *Grevillia* contain a proportionally greater number of rare species than do species-poor genera. Gymnosperms, containing relatively ancient species, may be less adapted to today's rapidly changing environment (IUCN, 1998). Cacti in the United States are particularly prone to extinction: 72 of 268 native species are very rare (Benson, 1982). In California, shrubs and subshrubs of the species-rich genera *Arctostaphylos*, *Ceanothus*, *Eriogonum*, and *Ribes* have a higher proportion of threatened and endangered plants relative to their proportion in the flora (Cody, 1986).

Curiously, some rare plant habitats may have human origins. Some moorlands, or blanket bogs, in western Europe were created around 7700 years ago by human activities of burning and grazing. These heathlands and grasslands today support many endangered plant species.

Causes of Endangerment

There are several natural causes of rarity that lead to endangerment. Ancient taxa, like cycads, are thought to be prone to extinction as they cling to increasingly shrinking relict habitats. New taxa, resulting from recent speciation, are thought to be susceptible to extinction because they have not had time to spread in distribution to reduce risk. Rare species generally lack an ability to rapidly colonize areas and they are often poor competitors. Many rare species are edaphically restricted to specific soils or geology. However, based on recent rates of extinction, natural causes of rarity (individually or combined) are no match for human-related causes of endangerment.

Species usually become endangered because of multiple, human-related causes. The overwhelming cause of plant species endangerment is habitat loss, directly linked to agriculture, forestry, and urbanization. Remaining small populations have a higher risk of extirpation (local extinction) or extinction. Coinciding with habitat loss is reduced habitat quality related to invasive species, grazing, and other land-use changes (Flather *et al.*, 1994). There are many examples of habitat loss to choose from. In the wetter forested Sierra Madres of Mexico, Guatemalan fir (*Abies guatemalensis*) or Pinabete trees are considered endangered from over-exploitation of old-growth forests, land-use change, and livestock grazing (Burton, 1991). *Persea theobromifola*, once an important timber species in western Ecuador, has been reduced to fewer than 12 trees at Rio Palenque (Gentry, 1986). In the United States Burton (1991) reported that only 1% of the 1,036,000 km^2 (400,000 mi^2) of tallgrass prairie remains, now existing in isolated patches and small nature reserves. Over half of all continental U.S. wetland and aquatic habitats have been destroyed. Fire suppression and succession, invasive exotic plants, and habitat fragmentation will make it increasingly difficult for rare species to persist.

Causes of rarity vary by region. Though habitat loss was generalized as an overall problem, agricultural development was specifically noted in the southern Appalachian region, whereas urban development and forest clearing were specifically noted in Florida. In the Gulf Coast areas, shoreline modification and development were specifically noted in endangerment patterns (Flather *et al.*, 1994).

Commercial exploitation is also a problem. Cacti in the southwestern United States are a prime example. Legislation regulating the collection of cacti in Arizona and California date from 1929. Yet, in a single year, over half a million small cacti were illegally collected for sale in the United States, Japan, and Europe (Burton, 1991). One of the two colonies of the 2.5-cm-diameter Nellie Cory cactus (*Coryphantha minima*) was completely eliminated by thieves in the 1960s. The rarest species demand the highest prices from collectors. Rare orchids are plucked one by one by various "collectors." Some threatened orchids in India, such as *Paphiopedilum druri*, *Dendrobium pauciflorum*, *Dendrobium nobile*, and *Diplomeris hirsuta*, are used for medicinal purposes. As such, they are facing increasing pressure from growing and aging human populations.

Natural environmental change can both cause and maintain rarity. Cropper (1993) reported that some rare species, such as *Leptorhychos gatesii* in Australia, are observed only following natural fire. A riparian herb, *Collomia rawsoniana*, required periodic high and low streamflows to persist. Likewise, disruption of the natural hydrology of Lake Okeechobee, the largest lake in Florida, has led to the endangered listing of *Cucurbita okeechobeensis*. Land management decisions can obviously upset required disturbance patterns. Fire suppression activities and flood control will likely increase the rarity of such species.

The introduction of exotic grazers such as goats, pigs, rabbits, and sheep often leads to devastated native floras. For example, 47 of the 49 native plant species on St. Helena Island in the south Atlantic Ocean are now rare or threatened. The island was once dominated by beautiful forests of St. Helena ebony (*Trochetia melanoxylon*). Goats were introduced on the island in 1513, and the goat population skyrocketed by 1588. Goats ate ebony seedlings and humans used mature trees as fuel, so by 1810 the forests were destroyed and the ebony was extinct. About 260 naturalized exotic plant species were also introduced to the island (Davis *et al.*, 1986).

Introduced rabbits have greatly threatened *Acacia carnie* in arid western New South Wales, Australia (Cropper, 1993), and nearly decimated *Dudleya traskiae* (ironically called the Santa Barbara live-forever) on Santa Barbara Island off California (Benseler, 1987). Hawaii's beautiful Haleakala silversword (*Argyroxiphium sandwicense* ssp. *macrocephalum*), with a flower stalk up to 2 m tall, almost went extinct in the 1920s due to vandalism and grazing by introduced cattle and goats. Like many endangered plant species, several additional threats must be held at bay. A major effort in Hawaii is under way to

protect the habitat of the Haleakala silversword from invasive exotic plants (*Verbascum thapsus*, mullein; and *Pennisetum setaceum*, fountain grass) and the Argentine ant (*Iridomyrmex humilis*). The Argentine ant poses a significant threat to native pollinators of the silversword.

Competition from invasive exotic plant species may be a major contributor in the future to native plant endangerment. Large, nearly pure stands of exotic purple loosestrife (*Lythrum salicaria*) have directly influenced the endangered small spike-rush (*Eleocharis parvula*) in New York and Long's bulrush (*Scirpus longii*) in Massachusetts.

Consequences of Rarity

The most obvious consequence of rarity is extinction. Charles Darwin, like many naturalists of his time and since, recognized that rarity often preceded extinction. Small populations of sessile organisms are vulnerable to catastrophes. Landslides, fire, flooding, hurricanes, and other disturbances can simply wipe out populations. Small populations of vascular plants are also vulnerable to breeding problems from higher variability in breeding success caused by inbreeding. Inbreeding has been documented in several localized endemics such as *Limnanthes bakeri*, a vernal pool species in Mendocino County, California. In contrast, a common congener *Limnanthes douglassii* reproduces almost exclusively by cross-fertilization. *Stephanomeria malheurensis*, a plant species confined to one small hilltop in Burns, Oregon, is auto-fertile, whereas wide-spread congeners are not. Reproductive failure is not uncommon in sparse species. Lower genetic variation in small populations may also make them more vulnerable to rapid environmental change.

Not all small populations march rapidly toward extinction. Many rare plant populations can persist for centuries and millennia. Sparse populations can often avoid pathogens and herbivory. *Pinus ponderosa*, which may have been restricted to small refugia in Arizona and New Mexico at the end of the last Ice Age, has become the most widespread pine in the western United States; it now occurs from Mexico to Canada, and from California to Nebraska. Other small populations may speciate (form new species). For example, two species of *Ranunculus* in alpine areas of the North Island of New Zealand (*Ranunculus verticillata* and *Ranunculus insignis*) have given rise to *Ranunculus nivicola*.

Legal mandates to Protect Endangered Plants

Policy and Legal Mandates Throughout the World

The International Plant Protection Convention held in Rome in 1951 set forth recommendations for the protection and promotion of plant life throughout the world. Since then, the International Union for Conservation of Nature and Natural Resources (IUCN; now known as the World Conservation Union) has taken center stage in the protection of the world's flora. The IUCN Plant Red Book strongly defines "endangered species" as a species in danger of extinction and whose survival is unlikely if the causal factors (e.g., over-exploitation, extensive habitat destruction) continue operating, "including taxa whose numbers have been reduced to a critical level or whose habitats have been so drastically reduced that they are deemed in immediate danger of extinction" (IUCN, 1998). It also identifies and tracks "vulnerable species" as those "believed likely to move into the endangered category in the near future if the causal factors continue operating." Finally, it tracks "rare species"—"taxa with small world populations that are not presently endangered or vulnerable, but are at risk." These IUCN classifications are determined by scientists and government officials around the world and classified species are not necessarily afforded legal protection after designation (Stermer, 1995).

Several countries have policies or legislation that protect endangered plant species. Following the Convention on International Trade in Endangered Species (CITES) of Wild Flora in 1973, 175 countries have agreed not to trade certain threatened species. Many countries augment these agreements with additional legislation. For example, the Mauritius National Plant Protection Legislation (the Plants Act of 1976) and the Forest and Reserves Act (1983) legally protects endangered species and habitats in the territories of Mauritius. Enforcement of endangered species laws and policies in many countries is generally considered to be well-intentioned but weak.

The United States Endangered Species Act

The intent of the U.S. Endangered Species Act of 1973 (16 USC1531-1543) is to prevent further decline and help restore endangered and threatened species and the habitats on which such species depend and to "provide a means whereby the ecosystems upon which endangered species and threatened species depend may be conserved" (Greenwalt and Gehringer, 1975). Thus, the Act recognizes the inseparable link between protecting a species, its habitat, and the surrounding ecosystem. The Act also provides broad-ranging protection for all species threatened with extinction in the "foreseeable future."

The Act makes the "taking" of endangered species anywhere within the United States a federal offense, requires federal agencies to use their existing authorities to conserve listed species, prohibits federal agencies from taking actions that may jeopardize a species' existence, provides a formal structure for listing endangered species, and provides a means for citizens to bring suit against any federal agency for failure to meet its obligations under the Act (Flather *et al.*, 1994).

Ayensu and DeFilipps (1978) noted that a species may be rare at the edge of its range, but not endangered or threatened as a whole. In determining national endangered, threatened, and extinction status, the total range and abundance of the species must be considered. However, states may further protect a species threatened with extirpation. Still, the cost of protecting individual species and habitats against multiple stresses is high, and the reality is that enforcement on public and private lands is generally weak.

Selected Examples of Endangered Plants

Coleus Forskohlii (Willdenow)

Coleus forskohlii, a 40-cm-tall, rare herb, is found in the Yunnan Province of China, Bhutan, India, Nepal, Sri Lanka, and Africa at

about 2300 m on steep slopes (**Figure 3**). Compounds extracted from the roots have long been a Hindu Ayurvedic traditional medicine. Rampant collection has increased the rarity of this species. A German pharmaceutical company holds at least six U.S. patents for use of the plant as treatments for high blood pressure, cardiovascular disease, colic, respiratory problems, insomnia, painful urination, and convulsions. This is sometimes referred to as "biopiracy," when the intellectual property of indigenous peoples is appropriated and used by foreign companies to develop and patent commercial products.

Fitzroya Cupressoides

Fitzroya cupressoides is an ancient tree species in southern Chile (**Figure 4**). Only 5% of the world's temperate forests are in the Southern Hemisphere, and one-third of the threatened temperate forests occur in Chile. The highest biodiversity in any temperate forest is also found in Chile. These forests are remarkably productive, with some of the world's largest concentrations of biomass. Old-growth *F. cupressoides* trees often reach 4–5 m in diameter and may live for 4000 years. One-third of Chile's forests were burned or cleared by 1955. Some *Fitzroya* populations are protected in nature reserves, but most of the forests containing this species are privately owned.

Sarracenia Rubra ssp.

Sarracenia Rubra ssp. Alabamensis

The Alabama canebrake pitcher-plant is a carnivorous plant with maroon flowers on 0.6-m stalks originating from rhizomes (underground roots; **Figure 5**). It is found in sandy and gravelly bogs, seeps, springs, and swamps, and flowers from late April to early June. *Sarracenia rubra* ssp. *alabamensis* is restricted to only 12 localized sites in a three-county area in



Coleus forskohlii root

The root of *Coleus forskohlii* is the only source of the biologically active compound Forskohlin. This ancient medicinal herb has been re-discovered by modern pharmaceutical researchers who have patented uses of the plant. *Coleus forskohlii* grows in the Yunnan Province of China, and in Bhutan, India, Nepal, Sri Lanka, and Africa.



Figure 3 *Coleus forskohlii.*



Figure 4 *Fitzroya cupressoides*. (Photograph courtesy of Dr. T T Veblen.)

central Alabama. Four of the 12 sites have 70–300 plants each, and half the populations have 2–20 plants each. Much of the original habitat (16 other sites) has been modified or destroyed by agriculture and construction of farm ponds in boggy areas. Fire exclusion, gravel mining, and invasive plants pose additional threats. Several populations have also been lost or degraded by plant collectors.

Encephalartos Longifolius

South African cycads are ancient gymnosperms and the most primitive living seed-bearing plants on Earth (**Figure 6**). They flourished 50–60 million years ago and provided forage for dinosaurs before then. In South Africa, all 40 cycad species are endangered, and some species are extinct in the wild. The thick-trunked plants with rigid spiked leaves grow very slowly; some species take 100 years to grow 1 m. Habitat loss and competition from invasive exotic plant species are often cited as causes of rarity, but illegal collecting is also a major problem. South Africa has some of the world's strictest laws controlling cycad theft, but the thefts continue. Some cycads now have implanted microchips for identification and tracking purposes in an effort to curb poaching.

Pterostylis Truncata

Brittle greenwood is a ground-dwelling orchid that emerges in the fall with large, squat flowers (**Figure 7**). It is a clonal plant that can regenerate vegetatively, and its flowers can be pollinated by flies or fungus gnats. *Pterostylis truncata* is found in only three locations in south-central New South Wales, Australia. One of the populations is threatened by introduced European rabbits (*Oryctolagus cuniculus*), feral goats (*Capra hircus*), and eastern gray kangaroos (*Macropus giganteus*). An additional recent threat comes from the highly invasive weed *Chrysanthemoides monifilera*. The second population is on private land, where it is potentially threatened by kangaroos and rabbits. The third population occurs in a forest reserve; here it is threatened from trampling by orchid enthusiasts, weed invasions, and bird predation.

Species-Environmental Matching Models

A new tool for rare and threatened plant species detection and conservation is the use of species-environmental matching models (Phillips *et al.*, 2006). These models use species

The Alabama canebrake pitcher-plant *(Sarracenia rubra ssp. alabamensis)* is an endangered, carnivorous plant. It is restricted to only 12 sites in central Alabama. Many factors have contributed to the listing of this plant as endangered, including habitat destruction, fire suppression, and the introduction of non-native plants.

Sarracenia rubra ssp. alabamensis



Figure 5 *Sarracenia rubra* ssp. alabamensis.

locations, with dozens of geographic information themes (e.g., climate factors, soils, topography, land use, vegetation) to model habitat suitability of a species. Kumar and Stohlgren (2009) showed the utility of the models for a threatened and endangered tree *Canacomyrica monticola* in New Caledonia (**Figure 8**). Only 11 locations were known (De Lang *et al.*, 2010). Maximum entropy model or Maxent (Phillips *et al.*, 2006), one of the species-environmental matching models, had a 91% success rate. Model results were successfully used to find *C. monticola*'s new populations at 12th location in southern New Caledonia (**Figure 8**). This approach is quite promising for predicting suitable habitat for threatened and endangered species and locating new populations or artificial introduction sites.

Implications

Reducing the rate of habitat loss around the world will be difficult. Between 1980 and 1990, species-rich tropical forests were cleared at the rate of 6.3 million ha year^{-1} (15.4 million acres year^{-1}), or 0.8% of the forest per year (Heywood, 1995). Deforestation causes habitat loss, habitat fragmentation, and edge effects at the boundaries, so the effects are greater than indicated by the actual deforested area. Large areas will be lost to urbanization and agriculture. Many dry grasslands in Germany have been converted to range and arable land in recent decades, making it more difficult to protect the nation's 164 threatened plant species (Heywood, 1995). Thus, saving critical habitats throughout the world is seen as increasingly important and extremely urgent.

A species-by-species approach to rare plant conservation is expensive and difficult. However, the unstated assumption of the habitat preservation approach is that species are inseparably linked to habitats, which in turn are stable and predictable. Yet because habitats may be neither stable nor predictable (Flather *et al.*, 1994), both species and habitats must be monitored to protect endangered plants and animals. In Australia, management techniques to protect critical habitats and species include: burning or slashing overgrown vegetation, removal of weeds, removal of grazing animals, hand pollination of selected species, propagation and seed storage, reintroduction into restored habitats, and quarantine to reduce the threat of introduced pathogens (Cropper, 1993).



Encephalartos species with male cones.

Encephalartos is the second largest genus in the Cycad genera. All *Encephalartos* are endemic to the African continent, with most in xeric habitats. Natural populations of *Encephalartos* suffer intense pressure from plant collectors.



Figure 6 *Encephalartos longifolius*. (Photograph courtesy of Dr. D A Steingraeber.)



Pterostylis truncata

Brittle greenhood (*Pterostylis truncata*) is a clonal orchid, found in New South Wales and Victoria, Australia. Causes of this orchid's decline include overgrazing by feral goats and European rabbits and the introduction of a highly invasive weed into its habitat.

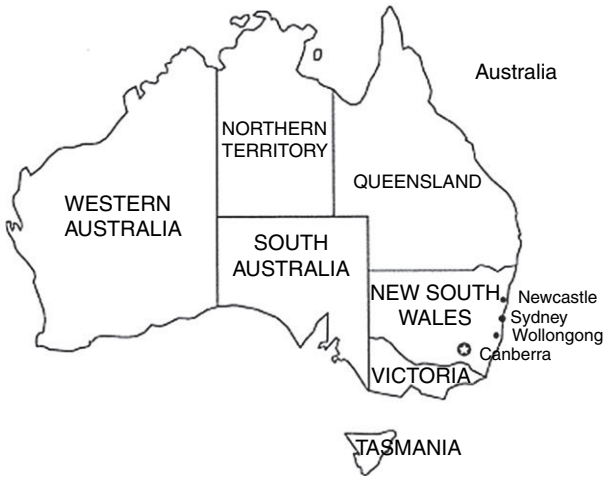


Figure 7 *Pterostylis truncata*.

Less than 10% of named plant species have been analyzed for medicinal or nutritional properties (Stermer, 1995). Nonetheless, 25–40% of all drug prescriptions in the United States contain plant ingredients (Durant and Saito, 1985), and many of these ingredients cannot be synthetically made. About 80% of people in developing countries use traditional medicines. As more and more plant species go extinct, so may our chances to find the next heart medicine (e.g., *Digitalis*; foxglove) or treatment for childhood leukemia (from *Catharanthus roseus*; the rosy periwinkle).

To improve the effectiveness of biodiversity conservation, increased emphasis is needed on systematic surveys and monitoring. Many species may be classified as rare owing to poor surveys, as evidenced by the rate of species discoveries.



Canacomyrica monticola

Model accuracy = 91% (P = 0.003)
Jackknife test (leave-one-out)

Important predictors
Precipitation seasonality (51%)
Temperature seasonality (26%)

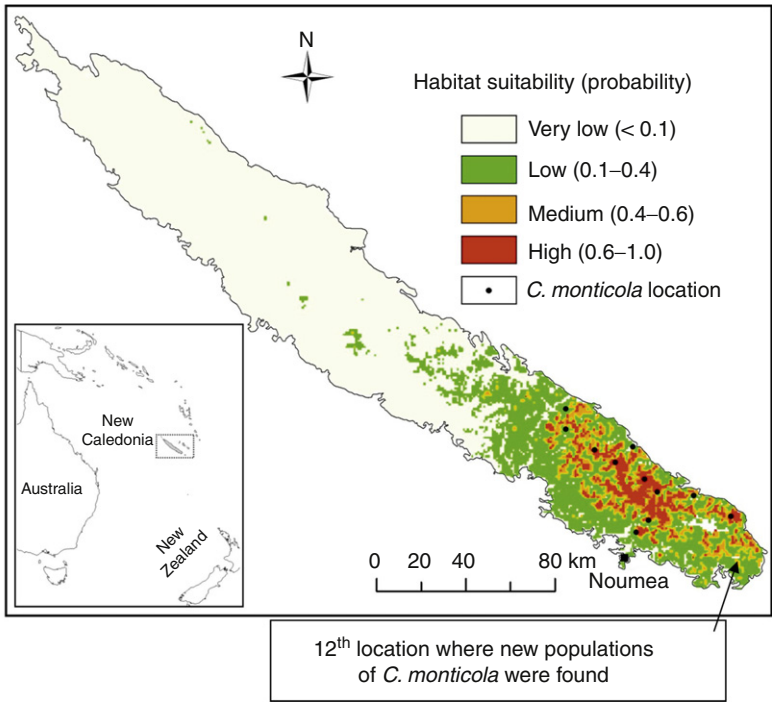


Figure 8 Potential suitable habitat for *Canacomyrica monticola* on Grande Terra, the main island of New Caledonia.

Since 1970, for example, one botanist has discovered 55 new plant species in Utah, and more than 80 species have been named since 1970 (Durant and Saito, 1985). Botanists searching the well-studied Rocky Mountain National Park in Colorado added more than 100 species to the Park's plant checklist between 1987 and 1992 (Stohlgren *et al.*, 1997). On average, one or two native plant species are added to the flora of New York each year. The Nature Conservancy and Natural Heritage Programs have found that plant surveys often show that many plant species are more common than previously believed. For example, several populations of rare orchids have been found in Australia after intensive searches. Systematic surveys of plants and rare habitats are badly needed.

The preservation of intact ecosystems may be the most promising way to protect endangered plants. It is now widely understood that maintaining natural disturbance regimes, such as fire and flooding, is important for many rare species (Cropper, 1993), and that preserving the habitat around rare plant locations is vitally important. However, subtle assaults to endangered species may include the use of fertilizers, herbicides, and insecticides that may reduce plant pollinators, invasive plant species that may out-compete the endangered plants, and introduced foreign diseases, insects, or domesticated/feral animals that may further endanger plants. Where habitat loss and degradation have been significant, habitat restoration efforts and species re-introductions are needed.

Cultivation, the last resort for endangered plants, may become necessary in some cases to preserve genetic variation. Franklinia (*Franklinia alatamaha*) is a familiar example. Franklinia is a small tree that was restricted naturally to the Altamaha River basin in southeastern Georgia. Now extirpated in the wild, the species survives as an ornamental throughout the eastern United States. In Australia, 1053 of 3329 rare and threatened plant species are found only in botanical gardens, and 515 plant species are held in only one collection (Cropper, 1993). Several botanical gardens in the United States have active programs to study, collect, and grow rare and endangered plants from all over the world. However, duplicate and more complete collections are needed in most countries, and such a system of collections should not be viewed as a substitute for preserving native habitats and ecosystems.

Some plants have appeared to bounce back from extinction. Mountain golden heather (*Hudsonia montana*) was described by Thomas Nuttall in 1818 in North Carolina, and was thought to be extinct in the 1960s. About 2000 individuals were found in the Blue Ridge Mountains around 1990 (Burton, 1991). Running buffalo clover (*Trifolium stoloniferum*) was also believed to be extinct, but it was rediscovered in 1983 in West Virginia and later discovered in four other states. Also in 1983, *Lomatium peckianum* was rediscovered in Oregon after "disappearing" for over 50 years. In 1985, one species of clover (*Trifolium microcephalum*) that had not been seen since the Lewis and Clark expedition in 1805–1806 was rediscovered. Such stories are not uncommon. Eleven plant species in New York have been "rediscovered" in the past 10 years. However, expecting miraculous rediscoveries for many long-lost species is probably unrealistic.

This article has focused on endangered vascular plants. However, many nonvascular plants are similarly threatened. For example, >50% of the mushrooms in Europe are listed as endangered or threatened in at least one country. Species that were once common, such as *Hydnum repandum*, have been extirpated in some countries. Over-collection of the lichen *Gymnoderma lineare* in the Great Smoky Mountains National Park has led to its recommendation for listing. Air pollution and acid deposition are known to adversely affect many

nonvascular plant species. The assessment of endangered plants cannot be done without assessing the interacting species (e.g., pollinators and symbiotic fungi) and the ecological processes that affect complexes of rare and common species.

Understanding the causes and consequences of rarity also requires a comprehensive knowledge of biology, evolutionary and recent history, and species demography. Often, the causes of population decline may be elusive. For example, *Torreya taxifolia*, the Florida torreya, is a narrowly restricted endemic conifer that has suffered catastrophic declines since the 1950s. Pathologists and ecologists have studied the problem relentlessly and found no obvious cause for the decline. There are now fewer than 1500 trees in the wild, with no solution in sight. In many cases, careful field and laboratory experiments may be necessary to isolate the causes of rarity.

It is equally important to increase public awareness about the ecology of rarity and the economic and social consequences of losing our endangered plants. In the end, protecting endangered species and biodiversity is a question of ethics and values. Endangered plants are best viewed as valuable resources and inherently valuable species with which we share the Earth.

Acknowledgments

April Owen and Lisa Schell assisted with the literature review and graphics. Curtis Flather provided information and Figure 2. Geneva Chong, April Owen, and Lisa Schell provided helpful comments on an earlier version of the manuscript. To all I am grateful.

See also: Endangered Ecosystems. Endemism. Extinction in the Fossil Record. Plant Biodiversity, Overview. Plant Invasions

References

Ayensu ES and DeFilipps RA (1978) *Endangered and Threatened Plants of the United States*. Washington, DC: Smithsonian Institution/World Wildlife Fund.

Benseler RW (1987) Conservation ethics, animals, and rare plant protection. In: Elias TS (ed.) *Conservation and Management of Rare and Endangered Plants*, pp. 623–626. Sacramento: California Native Plant Society.

Benson and Lyman D (1982) *The Cacti of the United States and Canada*. Stanford, CA: Stanford University Press.

Burton JA (ed.) (1991) *The Atlas of Endangered Species*. New York: Macmillan Publishing.

Cody ML (1986) Diversity, rarity, and conservation in Mediterranean-climate regions. In: Soulé ME (ed.) *Conservation Biology, the Science of Scarcity and Diversity*, pp. 122–152. Sunderland, MA: Sinauer Associates.

Cropper SC (1993) *Management of Endangered Plants*. East Melbourne, Australia: CSIRO Publications.

Davis SD, Droop SDM, Gregerson P, *et al.* (1986) *Plants in Danger: What Do We Know? International Union for Conservation of Nature and Natural Resources*. Switzerland: Gland.

De Lange P, Heenan P, Norton D, Rolfe J, and Sawyer J (2010) *Threatened Plants of New Zealand*, pp. 472. Canterbury University Press.

DuMond DM (1973) A guide for the selection of rare, unique and endangered plants. *Castanea: The Journal of the Southern Appalachian Botanical Club* 38(4): 387–395.

Durant M and Saito M (1985) The hazardous life of our rarest plants. *Audubon* 87(4): 50–61.

Flather CH, Joyce LA and Bloomgarden CA (1994) Species Endangerment Patterns in the United States, Fort Collins, Colorado General Technical Report RM-241, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station.

Gaston KJ (1997) What is rarity? In: Kunin WE and Gaston KJ (eds.) *The Biology of Rarity: Causes and Consequences of Rare-Common Differences*, pp. 31–47. London: Chapman & Hall.

Gentry AH (1986) Endemism in tropical versus temperate plant communities. In: Soulé ME (ed.) *Conservation Biology, the Science of Scarcity and Diversity*, pp. 153–181. Sunderland, MA: Sinauer Associates.

Greenwalt LA and Gehringer JW (1975) Endangered and threatened species: Notice on critical habitat areas. *Federal Register* 40(78): 17764–17765.

Heywood VH (ed.) (1995) *Global Biodiversity Assessment, United Nations Environment Program*. Cambridge, UK: Cambridge University Press.

IUCN (1998) *1997 IUCN Red List of Threatened Plants*. Walter KS and Gillett HJ (eds.) Morges, Switzerland: World Conservation Union (IUCN).

IUCN (2004) *IUCN Red List of Threatened Species: A global species assessment*. Baillie EM, Hilton-Taylor C, and Stuart SN (eds.) Gland, Switzerland and Cambridge, UK: IUCN.

IUCN (2010) *IUCN Red List of Threatened Species*. Version 2010.3. http://www.iucnredlist.org. Downloaded on 05 October 2010.

Kumar S and Stohlgren TJ (2009) Maxent modeling for predicting suitable habitat for threatened and endangered tree Canacomyrica monticola in New Caledonia. *Journal of Ecology and Natural Environment* 1: 94–98.

Missouri Botanical Garden. Online, June 2000, http://www.mobot.org

Orians GH (1997) Evolved consequences of rarity. In: Kunin WE and Gaston KJ (eds.) *The Biology of Rarity: Causes and Consequences of Rare-Common Differences*, pp. 190–208. London: Chapman & Hall.

Phillips SJ, Anderson RP, and Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259.

Rabinowitz D (1981) Seven forms of rarity. In: Synge H (ed.) *The Biological Aspects of Rare Plant Conservation*, pp. 205–217. New York: John Wiley & Sons.

Raven PH, Evert RF, and Eichhorn SE (1986) *Biology of Plants*. New York: Worth Publishers.

Stermer D (1995) *Vanishing Flora: Endangered Plants around the World*. New York: Harry N. Abrams.

Stohlgren TJ, Chong GW, Kalkhan MA, and Schell LD (1997) Rapid assessment of plant diversity patterns: A methodology for landscapes. *Ecology Monitoring and Assessment* 48: 25–43.

U.S. Endangered Species Act. 1973. Title 16 – Conservation. Ch. 35 – Endangered Species.

Endangered Reptiles

Tim Halliday, Open University, Milton Keynes, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Reptile Member of a class of vertebrates (the Reptilia), comprising turtles and tortoises (order Testudinata), lizards, snakes, and worm-lizards (order Squamata), the tuatara (order Rhynchocephalia), and crocodiles and alligators (order Crocodylia), which typically lay eggs with a leathery, impermeable shell. Reptiles have a dry, horny skin with scales, plates, or scutes.

Reptiles

Reptiles are ectothermic, meaning that the heat that they require to maintain physiological processes is derived externally, directly or indirectly from the sun. Reptiles are the ancestors of birds and the mammals, both of which are endothermic, meaning that body heat is primarily generated within their bodies. Their dependence on the external heat sources limits reptiles to tropical and temperate regions of the earth and they are especially numerous in the tropics. Reptiles are found in all continents of the world, except Antartica, living in a wide variety of habitats, including the world's driest deserts. Many are aquatic, living in both fresh water and the sea.

The skin of reptiles is covered in impermeable scales and they are able to retain water very effectively. Their eggs are covered by a hard shell, providing a sealed environment for the developing embryo, which does not depend on external sources of water. These two features have enabled reptiles successfully to colonize very dry habitats, including deserts. While tortoises are typically slow and ponderous in their movements, many reptiles are very agile and capable of moving very fast, often over large distances. Marine turtles, in particular, routinely undertake enormous journeys across the world's oceans.

Reptiles are divided into six major groups: The worm-lizards (*c.* 181 species), the lizards (*c.* 5461 species), the snakes (*c.* 3315 species), the turtles and tortoises (*c.* 317 species), the alligators and crocodiles (*c.* 23 species), and the tuataras (two species).

The Conservation Status of Reptiles

A comprehensive review of the conservation status of reptiles, like that of amphibians, has yet to be made, but reptile species are monitored by the World Conservation Union (IUCN) and by the Convention on International Trade in Endangered Species (CITES) (see **Table 1**). The latter is important because commercial exploitation is a particularly serious threat to many reptile species.

Threats to Reptiles

Many reptile species have become endangered by human activities that affect them either directly or indirectly. Direct impacts include the widespread persecution of snakes, whether or not they are venomous, killing of turtles and some lizards for food, and collection of tortoises, lizards, and snakes for the global pet trade. Indirect effects include habitat destruction, climate change, and pollution.

Habitat Destruction

Those reptiles that spend all or part of their lives in fresh water, such as alligators, crocodiles, turtles, and water snakes, are threatened in many parts of the world by the draining of wetlands. Some regions of the USA retain less than 20% of their original wetlands and numerous aquatic and semiaquatic reptiles have declined as a result. In Australia, destruction of

Table 1 Conservation status of reptiles, according to the Convention on International Trade in Endangered Species (CITES) and the World Conservation Union (IUCN)

Taxon	*CITES*[a]				*IUCN*		
	Approximate number of species	*Appendix III*	*Appendix II*	*Appendix I*	*Vulnerable*	*Endangered*	*Extinct*
All reptiles	7150	19	383	70	153	100	20
Turtles and tortoises	260	6	49	25	58	38	6
Crocodilians	22	0	8	16	3	7	0
Tuataras	2	0	0	2	1	0	0
Lizards and worm-lizards	5066	0	238	16	66	30	11
Snakes	1800	13	88	11	25	25	3

[a]Appendix I species are threatened with extinction and are, or may be affected by trade; Appendix II species are not currently threatened but are likely to become so if trade is not restricted; Appendix III species are listed to prevent exploitation.

Source: Data from FWS (2000).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00279-3

woodland and scrub to create agricultural land has been the major cause of population declines among snakes. Declines among reptiles as a result of deforestation have been particularly marked on Madagascar.

Roads present a particular risk to reptiles, not only because they have to cross them as they move about, but also because they are attractive places to bask in the sun. Large numbers of snakes, tortoises, and lizards are killed on roads. In some turtles, females are at greater risk than males, apparently because they move about more than males.

Animals of all kinds that live on small oceanic islands are especially vulnerable to extinction, and reptiles are no exception. Island species are often very vulnerable to introduced predators, such as rats, cats, dogs, and mongooses, having evolved in an environment in which such predators were previously absent. Small islands are also highly susceptible to habitat destruction and excessive hunting. Among endangered reptiles that live on islands are giant tortoises such as those found on Aldabra and the Galapagos.

Some reptiles are vulnerable to exploitation by humans and to natural predation because they gather to breed in large numbers. A marine turtle, Kemp's ridley *(Lepidochelys kempi*), was reduced to near extinction in the 1970s by the commercial exploitation of breeding females and eggs, gathered at their coastal breeding sites in Mexico and Texas. This species is now internationally protected and its numbers are recovering.

Migratory species, such as some sea turtles, are very vulnerable both because their conservation requires that they be protected in several, often distant, parts of the world and because they have to make long journeys across possibly hazardous environments. The green turtle (*Chelonia mydas*), for example, is protected in Australia but is exploited in Indonesia. Leatherback turtles (*Dermochelys coriacea*) nesting in Gabon in central Africa undertake huge journeys across the Atlantic, to South America or to southern Africa. Satellite-tracking has revealed that one female leatherback traveled 7563 km between a feeding area off South America to her nesting site in Gabon (Witt *et al.*, 2011). Such long journeys take turtles through many areas where they are in danger of being caught by fishermen. As a result, their effective conservation would require active cooperation by at least 11 different nations.

Commercial Exploitation

Throughout human history, reptiles have been used by people for a variety of purposes. Apart from being a food source, tortoise shell is used for decorative purposes. Snake, lizard, and crocodile skins are used to make shoes, belts, and handbags, and the flesh of many reptiles is believed to have medicinal or aphrodisiac properties.

In some parts of the world, there is also a substantial trade in reptile meat. Between 1979 and 1987, the hunting of alligators in Louisiana yielded 45,000 kg of alligator meat each year. Many turtles are heavily exploited as a food source, both by indigenous people and the markets in developed countries to supply luxury food. In Asia, the trade in turtles for food and medicines has reached crisis proportions, with 67 of 90 species categorized as threatened and 18 as critically endangered.

The green iguana (*Iguana iguana*) has declined severely as a result of being hunted for food. Eaten by humans in Central America for centuries, destruction of its habitat and the expansion of the human population have led to it being hunted much more intensely in the past 30 years. In the late 1960s, as many as 150,000 iguanas were eaten each year in Nicaragua alone. Not only is iguana flesh regarded as a great delicacy but also their fat and eggs are used for a variety of medicinal purposes, including a cure for impotence.

The adverse effects of the overexploitation of natural populations can, in some cases, be countered by sustainable farming. There is now a thriving trade in American alligators (*Alligator mississippiensis*) with more than 30 farms in Florida providing more than 136,000 kg of alligator meat and 15,000 skins each year.

The highly durable skin of reptiles makes it an excellent alternative to leather, and reptile skins have long been used for making footwear, purses, belts, and many other accessories. Many species of pythons, boas, varanid lizards, and crocodiles are declining because they are hunted for their skins. In 1981, the USA imported 304,189 pairs of shoes made from boa skin and 176,204 pairs of python skin shoes. In the absence of commercial breeding programs for these species, this trade was entirely based on the exploitation of wild populations. Today, trade in reptile skins is increasingly regulated and sustained by breeding animals in captivity, but there remains a substantial illegal market in skins derived from natural populations.

Products derived from reptiles and amphibians are constituents of many traditional medicines and are used in many other ways. Broths made from snakes and tortoises are believed to combat many diseases; the fat of monitor lizards (*Varanus* spp.) is used to combat skin infections; sea turtle eggs are used as aphrodisiacs; turtle oil is used to make perfumes and lubricants.

Introduced Species

As the human population has expanded and people have colonized new parts of the world, they have taken with them, deliberately or accidentally, a variety of organisms that have been harmful to indigenous wildlife. This is a particularly serious threat to species living on small islands where, in the absence of indigenous predators, they have lost their defense against predators. Populations of many reptile species that are endemic to islands have been devastated by the introduction of alien rats, cats, dogs, and other animals that feed on them or their eggs, and introduced pigs and goats have destroyed the ground cover that supports their food supply and which they need to escape from predators.

The tuataras (*Sphenodon*) have become extinct on the two major islands and on many of the smaller islands of New Zealand and are now confined to a very few small islands that remain free of alien immigrants. Intensive efforts are now being made to clear other islands of predators and prevent further releases of alien species. In addition, tuatara eggs are being incubated in captivity and the resulting lizards are being released across the various protected islands in such a way as to maintain genetic diversity in these small populations (Keall *et al.*, 2010).

The rarest lizard in the world is the Jamaican iguana (*Cyclura collei*). Believed for many years to be extinct, as the result of predation of eggs and young by introduced mongooses, a tiny population was rediscovered in 1990. A population of 100 to 200 animals is now being sustained by protection of its habitat, exclusion of mongooses, and captive breeding.

Some reptiles are themselves invasive species. Brown tree snakes (*Boiga irregularis*), accidentally introduced onto the island of Guam from their native Australia and Papua New Guinea, have seriously reduced or extirpated not only populations of endemic birds but also several native reptiles.

Climate Change

For reptiles, temperature is crucial because many species rely on basking in the sun to raise their body temperature to a level at which they can carry out the essential activities such as feeding and reproduction. Consequently, future climate change is likely to pose a serious threat to many reptile species. It has been predicted that this threat is most serious for tropical species, even though temperature increases are predicted to be less marked at tropical than at temperate latitudes (Deutsch *et al.*, 2008). This is because many tropical reptiles are currently living very close to the maximum temperature that they can tolerate. In contrast, many temperate reptiles are currently living in climates that are cooler than what is optimal for them and they may benefit from future climate change. However, because reptile diversity is much higher in the tropics, more species are likely to be adversely affected by climate change than will benefit from it.

Climate change is likely to affect reptiles in a more subtle way, because sex-determination in many reptiles is strongly influenced by the temperature at which their eggs develop. The details of this process vary across different groups of reptiles. In many turtles, low temperatures produce male offspring, high temperatures females; the reverse is true in some lizards. In crocodiles and alligators, low and high incubation temperatures produce females, while intermediate temperatures produce males. These effects have been intensively studied in the laboratory, but there is evidence that they may affect wild populations. Studies of painted turtles (*Chrysemys picta*) have revealed that small increases in air temperature during July reduce the proportion of offspring that are male (Janzen, 1994). Extrapolating from these results, it is predicted that an increase in mean temperature of 4 °C would effectively eliminate the male offspring altogether.

Pollution

Pollution in the form of solid plastic waste is harmful to some sea turtles that may eat it. Green turtles (*Chelonia mydas*) eat plastic bags while feeding on aquatic plant food and leatherback turtles (*Dermochelys coriacea*) mistake plastic bags for their jellyfish prey. Half of the sea turtles examined in some localities have plastic debris in their intestines. This may interfere with their digestion, respiration, and buoyancy, and some plastics are toxic.

Some chemical compounds of human origin are readily absorbed by animals and interfere with their endocrine systems; these are known as endocrine disrupters and include estrogen mimics, which are compounds that disrupt the reproductive development of both sexes. Males may become feminized to varying degrees, for example, suffering lowered sperm counts. Reptiles are especially susceptible to such effects because of their sex-determination mechanisms which, in many species, are very dependent on environmental factors. Of particular concern are the polychlorinated biphenyls (PCBs), industrial chemicals such as those used in fire retardants and adhesives, which persist for a very long time and accumulate in the environment. Some PCBs have a molecular structure so similar to estrogen that they mimic its effects when they enter an animal's body. PCBs can turn male red-eared slider turtles (*Trachemys scripta*) into females, drastically reducing the reproductive success of populations.

Many pesticides, such as DDT, are not only poisonous to a wide range of animals but also, at low concentrations, have endocrine-disrupting effects. Populations of alligators in Florida have been adversely affected by such compounds because of their feminizing effects on males. Many of these compounds, including PCBs, are highly volatile and so are transported widely in the environment from their point of origin. The long-term effects of widespread chemical pollution of this kind are poorly understood but are very alarming not just for reptiles but also for animals of all kinds, including humans.

See also: Captive Breeding and Reintroduction. Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. Climate, Effects of. Human Impact on Biodiversity, Overview. Introduced Species, Impacts and Distribution of. Marine Protected Areas: Static Boundaries in a Changing World. Migration. Reptiles, Biodiversity of

References

Deutsch CA, Tewksbury JT, Huey RB, *et al.* (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences* 105: 6668–6672.

Janzen FJ (1994) Climate change a temperature-dependent sex determination in reptiles. *Proceedings of the National Academy of Sciences* 91: 7487–7490.

Keall SN, Nelson NJ, and Daugherty CH (2010) Securing the future of threatened tuatara populations with artificial incubation. *Herpetological Conservation and Biology* 5: 555–562.

Pough FH, Andrews RM, Cadle JE, Crump ML, Savitsky AH, and Wells KD (1998) *Herpetology.* Englewood Cliffs, NJ: Prentice Hall.

Witt MJ, Bonguno EA, Broderick AC, *et al.* (2011) Tracking leatherback turtles from the world's largest rookery: Assessing threats across the South Atlantic. *Proceedings of the Royal Society of Britain* 278.

Endangered Terrestrial Invertebrates

Piotr Naskrecki, Harvard University, Cambridge, MA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Mark Deyrup, volume 2, pp 487–495, © 2001 Elsevier Inc.

Glossary

Centinelan extinction Disappearance of species before their formal recognition by science.
Hyperparasitoid A parasitoid of a parasitoid (e.g., a parasitoic wasp larva living in a parasitoid fly larva that feeds on a beetle).
Parasitoid An organism that spends a significant portion of its life history attached to or within a single host organism; unlike a true parasite, it ultimately sterilizes or kills the host.
Taxonomic impediment Insufficient amount of taxonomic expertise and tools to identify and describe the world's biodiversity.
Troglobitic Cave-dwelling organisms, adapted to life in complete darkness.

Overview of Terrestrial Invertebrates

Diversity and Species Numbers

To talk about terrestrial invertebrates is to talk about the majority of life. Although the numbers of species of living organisms can only be estimated to the nearest orders of magnitude, recent calculations based on the rate of description of higher taxonomic categories predict the existence of 8.7 million currently living species ($\pm$1.3 million.) Of these, approximately 7.7 million are terrestrial animals, of which only approximately 35,000 are vertebrates (Mora *et al.*, 2011). Thus, by even the most conservative estimates, terrestrial invertebrates constitute at least 80% of all living species, distributed in all inhabitable terrestrial environments. Consequently, the endangerment of any terrestrial habitat is often equivalent to the endangerment of many of its invertebrate inhabitants, the great majority of which remain unknown to science. The relatively small proportion of terrestrial invertebrates that have been formally named (fewer than 2 million species) is in most cases not studied sufficiently enough to make species-level assessment of their endangerment or even provide estimates of their population sizes and distribution. (*Note:* The term "endangered" is used here to indicate any potential threat to the long-term survival of a species, and not as a category within the International Union for Conservation of Nature (IUCN) Red List ranking system; thus, for the purpose of this article, species listed as near threatened, vulnerable, endangered, etc., by IUCN are considered to be endangered.)

The overwhelming majority of terrestrial invertebrates are insects (including springtails and relatives), and between 1.4 and 1.8 million species of insects have been named. Other groups of terrestrial invertebrates include arachnids (spiders, mites, and relatives) with approximately 100,000 named species; myriapods (centipedes and millipedes) with about 13,000 named species; crustaceans, with fewer than 5000 terrestrial species; velvet worms, with about 200 species; tardigrades, with about 1200; annelids (earthworms), with about 6000 species; land planarians (Geoplanidae), with fewer than 1000 known species; soil nematodes, with at least 15,000 species; and land mollusks (snails and slugs), with approximately 25,000 species.

Taxonomic Impediment

One of the most serious obstacles in effective invertebrate conservation is the existence of the taxonomic impediment, understood as the lack of taxonomic expertise to identify and describe their diversity. Without the ability to recognize species, it is virtually impossible to implement effective measures for their targeted, species- or population-level protection. The taxonomic impediment encompasses not only the shortage of biologists capable of identifying invertebrate animals and recognizing those in need of protection but also the nearly complete lack of tools, such as field guides or keys for identification that could be used by conservation practitioners, such as managers of nature reserves or technical personnel in national parks to identify and monitor population of endangered invertebrates. The impediment is particularly severe in areas with the highest level of invertebrate biodiversity, such as exceptionally species-rich tropical areas of South America and Africa, as well as taxa with presumably little economic importance (e.g., earwigs and centipedes), or those that because of their small size require specialized equipment to identify (e.g., soil mites and nematodes). The only terrestrial invertebrate groups where taxonomic impediment has been largely ameliorated are a handful of groups of particularly attractive butterflies and large beetles, which have a long history of being collected and studied by both entomologists and amateurs. Recently, similar efforts have been underway to develop tools for quick and reliable identification of dragonflies and damselflies, and a few additional groups of terrestrial invertebrates, albeit those efforts concentrate primarily on faunas of the more developed, but less biodiverse countries.

The invention and development of Internet-based technologies within the last two decades have made publishing

and distribution of taxonomic identification tools easier and less expensive, but efforts are still hampered by the insufficient numbers of taxonomists capable of their development. Several international initiatives, such as Global Biodiversity Information Facility (GBIF) or Encyclopedia of Life (EOL), have been involved in the efforts to accelerate digital distribution of tools to reduce the taxonomic impediment. The development of rapid and inexpensive genetic sequencing technologies has lead lately to a more widespread use of genetic barcodes, small fragments of (usually mitochondrial) deoxyribonucleic acid (DNA) to identify species. Large, comprehensive databases of genetic barcodes and automated, widely available sequencing will likely soon lessen the taxonomic impediment in invertebrate conservation.

There is evidence that areas too small to support endemic species of vertebrates (e.g., the Antioch Dunes of California) have rich faunas of endemic invertebrates, which are at risk of extinction due to relatively small alteration of those habitats. The current general paucity of expertise in invertebrate taxonomy and identification as well as insufficient efforts to sample and document terrestrial invertebrates in threatened habitats may lead to the negative phenomenon known as Centinelan extinction, understood as the extinction of species before their very existence is recognized and documented. While, by the very definition of the process, it is difficult to prove Centinelan extinction, the existing knowledge of distribution ranges and dispersal abilities of many invertebrate taxa support the notion that the disappearance or anthropogenic alteration of natural habitats will lead to extirpation of species restricted to those habitats; in the case of Antioch Dunes the extinction of endemic invertebrates, such as the Shield katydid (*Neduba extincta*) following the habitat alteration, has already been demonstrated.

Life Histories

A compounding factor in efforts to monitor and protect terrestrial invertebrates is the presence of polymorphism and complex life cycles in many of their species, where immature forms may have very different appearance and habitat requirements than the adult forms. For example, insects that undergo complete metamorphosis as a rule have very different microhabitat and food requirements than the adults, such as caterpillars of butterflies that often require very specific food plants, different from those that adults use to obtain their nutrients (Van Swaay *et al.*, 2010). Many invertebrates that are terrestrial as adults spend their entire larval development in the water (e.g., dragonflies, mayflies, or some midges). An extreme example of the disjunction between the adult and larval lifestyle is the coconut crab (*Birgus latro*), the largest living terrestrial invertebrate, whose early developmental stages include oceanic, pelagic larvae. This means that any attempts in species- or habitat-level conservation of such organisms must take into consideration all requirements and environments needed by them.

None of the terrestrial invertebrates lives in isolation from other members of its biological communities, and many form obligatory, symbiotic relationships with other species. A good example of such association is that between blue butterflies (Lycaenidae) and ants. European Scarce Large Blue (*Phengaris teleius*) is a species whose caterpillars need to be picked up by worker ants of the genus *Myrmica* and carried off to the ants' nest, where they feed on ant grubs and eventually pupate. In addition, young caterpillars of this species need to feed on the shrub *Sanguisorba officinalis* for the first 2 or 3 weeks of their life. Because of such complex life history, this butterfly is vulnerable to any changes in the environment that affect either the host plants or host ants, and in large parts of its range, this species declines because of either intensification (e.g., drainage, fertilization, and use of pesticides) or abandonment (where its habitat gets invaded by scrubs and later forest) of its wet meadow habitat. Larvae of the moth genus *Ceratophaga* can only survive in the keratin, such as that of horns of antelopes or shells of tortoises, and thus are dependent on the existence and availability of their vertebrate hosts.

A special case of complex, symbiotic life histories in terrestrial invertebrates is parasitic relationships between insects and their hosts. This includes not only such well-known (and negatively perceived) cases as lice and fleas on vertebrate hosts but also parasitoids and hyperparasitoids that develop in other insects. Braconid wasps, ichneumonid wasps, or meloid beetles all require other insect species to complete their life cycles, and the decline or loss of the host species invariably leads to decline or loss of the parasitic ones.

Threats to Terrestrial Invertebrates

Habitat Loss and Fragmentation

The principal threat to the survival of terrestrial invertebrate species is the loss of their natural habitats. There is evidence that the average distribution range sizes of invertebrate taxa are smaller than the average range sizes of vertebrates, and their dispersal abilities are also lower than those of vertebrates. In addition, species associated with small, isolated or island-like habitats (e.g., peatlands of Europe, subalpine ecosystems, and small oceanic islands) appear to have even lower dispersal abilities, and consequently are more prone to be affected by changes to those habitats, and pushed below the minimal viable size of their population (Bönsel and Sonneck, 2011; Kisdi, 2002). The highly developed parts of the world, such as Europe or the coastal regions of the USA, have already lost the majority of their natural habitats due to centuries of agricultural practices and urbanization, and those habitat that are still relatively intact are usually highly fragmented. Fragmentation is considered a major factor that impacts the survival ability of many terrestrial invertebrate species that require a certain, critical minimum area of continuous habitat or habitat-dependent resources to survive. For example, the American burying beetle (*Nicrophorus americanus*) has disappeared from over 90% of its former range within the past 150 years as a result of anthropogenic habitat loss and fragmentation of relatively continuous stands of deciduous forests across the presettlement range of this species; these fragmented habitats are now currently too small to sustain the fauna of small mammals, on carcasses of which the beetles feed, extensively enough to supports its survival. Flightless

species are particularly vulnerable to the effect of fragmentation: brachypterous grasshopper *Prionotropis hystrix rhodanica*, endemic to the Crau Plain area of southern France, is declining rapidly due to increasing fragmentation of the habitat that supports "coussou" type of vegetation, following the intensification of industrial development in the area (Blight *et al.*, 2011). In Japan, the traditional "Satoyama" ecosystem, consisting mostly of secondary woodlands and grasslands, supports a high proportion of the country's endangered species of butterflies. Within the past 40 years, most of the "Satoyama" habitats have been industrialized, causing an over 80% decline in the populations of these insects (Nakamura, 2011).

It is not known what precise effect on local terrestrial invertebrate faunas has been caused by the extensive deforestation and loss of natural habitats in many tropical areas of the world, but it can be safely assumed that it has already contributed to the extinction of many thousands, if not millions of species. For example, Madagascar, an island reputed for its exceptionally high levels of endemism, which in some cases may encompass 100% of species in a given higher taxon, has lost over 90% of its original forest coverage within the past 150 years. This has undoubtedly caused Centinelan extinction on an unimaginable scale, and the remainder of the Malagasy invertebrate fauna is under continuous threat from further losses of natural habitats. The negative effect of deforestation in areas with lower levels of invertebrate endemism, such as the extensive Amazon and Congo Basins, is probably proportionately lower, but the increasing shrinkage and fragmentation of individual forest habitats are likely to accelerate the loss of even relatively widespread terrestrial invertebrate species.

The effect of anthropogenic change is felt particularly strongly by insular faunas of invertebrates: the majority of threatened terrestrial mollusks in Europe occur on the Macronesian Islands (the Canary Islands, the Azores, and Madeira), where a massive pressure due to the increasing urbanization of these islands and the ongoing expansion of the tourist infrastructure exists; it has been shown that even small-scaled disturbances, such as a single road, can destroy the habitat of narrow-range endemic snails, with little chance of recovery from neighboring areas. The orthopteroid fauna (grasshoppers and relatives) of the Seychelles, because of its isolation and inability to be replenished by immigration from other areas, has suffered greatly from habitat disturbance caused by human activity and is now restricted to ever shrinking patches of native vegetation; some species have not been seen for the past 100 years and should be presumed extinct. The effect of roads, especially wide, paved roads, has been demonstrated to increase fragmentation of habitats and depress biodiversity even among relatively vagile invertebrates like beetles (Dunn and Danoff-Burg, 2007).

Cave habitats exhibit many characteristics shared with oceanic islands in terms of their isolation and levels of endemism. Some troglobitic species are dependent for their survival on the availability of external resources, for example, the Table Mountain cave cricket (*Speleiacris tabulae*), depends on bat guano and hence on survival of the bat population. Cave habitats, because of their isolation and relatively static nature, are more sensitive to low-level disturbance than surface habitats, and even seemingly small anthropogenic impact, such as the increase in the compactness of the soil on the cave floor caused by human visitors, may have negative impact on the troglobitic fauna.

Climate Change

As ectotherms, animals that cannot easily regulate their body temperature, invertebrates are predicted to be affected more severely from climate change than warm-blooded animals. Global climate change will favor generalist, highly mobile species, and there is already ample evidence of species of butterflies and mosquitoes from lower latitudes moving toward higher latitudes and thus into areas where mean annual temperatures used to be too low for their survival. At the same time, cold, island-like habitats near mountain peaks that often shelter endemic invertebrate taxa are shrinking, and populations of cold-loving species that live there are showing signs of decline. The glacial and alpine habitats of the Sierra Nevada mountain range in the western USA have been steadily contracting or entirely disappearing, and several species of endemic ice crawlers (*Grylloblatta*) are presumed extinct due to the loss of these microhabitats.

The general unpredictability of the change in both faunistic and floristic composition of specific habitats that is likely to accompany the global climate change makes planning strategies for invertebrate conservation very difficult. Single site reserve selection or even designation of large protected areas may not be effective in counterbalancing the synergistic effects of the climate change on habitat fragmentation and other anthropogenic disturbance. It is certain that long-term maintenance of the status quo of species composition within any discrete geographic area is unattainable, and future invertebrate faunas will be different, and certainly less species-rich and more homogenous on larger spatial scales.

Pest Control

Invertebrate conservation and pest management are often viewed as directly antagonistic. Invertebrates, and such groups of insects as grasshoppers in particular, are perceived as something to be controlled and curbed, perhaps entirely exterminated. This is despite the fact that less than 10% of all grasshopper species have the potential to become "pests", and some "pest" species have been demonstrated to be highly beneficial to agricultural practices and increase the productivity of pastures (Samways and Lockwood, 1998). Yet the general perception of invertebrates as pests or disease vectors makes it easy to enact control measures against them but often exceedingly difficult to introduce measures toward their conservation.

Over 5 million tons of chemical pesticide are used annually worldwide to control the actual or perceived threat of pest invertebrates. Highly toxic, broad-spectrum insecticides and acaricides (poisons targeting mites) may have a particularly detrimental effect on invertebrate populations, as they do not discriminate between the target and nontarget species, and collateral damage of hundreds, perhaps thousands of species of nontarget invertebrates may result from such control

measures. But although these treatments have immediate and lethal effect on local invertebrate populations, they do not appear to have caused widespread extinction of species. However, they may have negative effects on populations of threatened invertebrate species in habitats adjacent to agricultural areas, and sudden increase in their mortality due to the synergistic effect of the pesticides and habitat fragmentation/alteration may cause them to drop below sustainable population levels. One well-documented case of extinction due to a combination of intense eradication efforts and the overall anthropogenic landscape changes is that of the Rocky Mountain grasshopper (*Melanoplus spretus*), which was considered one of the most important pests in North America in the nineteenth century but became extinct by the early 1900s. A similar case may be unfolding in the Mediterranean region of Europe, where the Moroccan locust (*Dociostaurus maroccanus*), once one of the most dangerous agricultural pests, is becoming rare following the conversion of grasslands into croplands and effective pest-control methods. Cases like this raise the question of whether pest species should be protected before they become extinct or allowed to disappear completely.

Introduction of genetically modified (GM) organisms into agricultural practices, particularly plants containing transgenes (genes transferred from other organisms) has caused concern among conservation biologists about their impact on populations of other organisms. The most contentious case was initially that of B.t. (*Bacillus thuringiensis*) gene transferred into corn, which enabled the corn to produce its own pesticides against insects such as the European corn borer. There is evidence that exposure of Monarch butterfly (*Danaus plexippus*) caterpillars to pollen from GM corn increases their mortality, but the overall effect on the population of this species of B.t. corn appears to be relatively low. The use of GM plants may in fact lead to a decrease in the use of broad-spectrum insecticides and have a long-term positive effect on the populations of nonpest species of invertebrates (although the introduction of the "Roundup Ready" GM variety of corn has caused an increase in the use of herbicides, and consequently a decline in insect populations associated with plants perceived as "weeds", including the Monarch butterflies associated with milkweed).

Invasive Species

The threat from alien invasive species is widely perceived to be one of the major causes of species decline and local extinction rates among animals, including terrestrial invertebrates. In South Africa, invasive alien trees have caused exclusion of sun-loving dragonflies from some areas due to the increase in the amount of shade in the habitat, and a similar effect has been noted on the local diversity of acridid grasshoppers. But the overall impact of invasive alien plants on local terrestrial invertebrate communities is poorly known, and in some cases may even lead to an increase in the alpha diversity of insects. A far greater risk is associated with invasion by alien animals, both vertebrate and invertebrate. Introduction of turkeys, hens, and rabbits to Iberian Mediterranean Islands has had a major negative effect on the numbers of tenebrionid beetles, whereas in Hawaii, the introduction of nonnative poecilid fish has severely impacted populations of endemic *Megalagrion* damselflies. Invasive rats, pigs, and possums in New Zealand have caused local extinctions of several species of land snails, including kauri snail *Paryphanta* spp. and the giant *Powelliphanta* spp. Rats and introduced alien hedgehogs have caused massive decline and local extinctions of giant, cricket-like wetas (*Deinacrida* spp. and related genera) (Watts *et al.*, 2008), and it is presumed that rats, feral pigs, and other alien mammals have lead to the disappearance of an unknown number of land invertebrates within the past 200 years in New Zealand.

Invasive insects, and ants in particular, are a major cause of decline or extinction of land invertebrates (New, 2010). The crazy ant (*Anoplolepis gracilipes*) is one of the worst offenders, and this species has been documented to cause exclusion of local ant species in the Solomon Islands, a population crash in the terrestrial red crab (*Gecarcoidea natalis*) on Christmas Island, and local extirpation of native ants (*Odontomachus simillimus*) and ghost crabs (*Ocypode* spp.) in the Seychelles. The Argentine ant (*Linepithema humile*) has displaced native ant species in an ecologically sensitive area in Spain and has been associated with local extinctions of native ants in California, whereas the African big-headed ant (*Pheidole megacephala*) is known to have had a major adverse effect on both local insect and spider populations in those parts of the world where it was inadvertently introduced.

Displacement of native species has also been caused by other groups of insects: the number of alien species of ladybugs (Coccinellidae) in the USA and Canada has increased steadily over the last century while the average proportion of native individuals has remained fairly constant until 1987, followed by a rapid decrease in the past 25 years; consequently, formerly common native species, *Adalia bipunctata* and *Coccinella novemnotata* have become rare in their former ranges.

Other invasive invertebrates include New Zealand predaceous land planarian (*Arthurdendyus vegrandis*), which has been shown to cause strong decline among native earthworms in the British Isles. The European hammock spider (*Linyphia triangularis*) has been introduced to NE, USA, and its large size, competitive ability, and aggressive nature has caused native linyphiid spiders to become scarce or virtually absent in areas it invades. The millipede *Scolopendra morsitans* is believed to be at least partly responsible for the decline of the endemic St. Helena giant earwig (*Labidura herculeana*), giant ground beetle (*Calosoma helenae*) on Saint Helena, and probably other endemic invertebrates on Ascension.

One of the most dramatic examples of the impact of alien species on local terrestrial invertebrates was the deliberate introduction of the predatory snail *Euglandina rosea* to Hawaii, done with the intention to control another snail, a pestiferous African invasive *Lissachatina fulica*. What followed instead was rapid extinction of the majority of land snails in the endemic to Hawaii family Amastridae. A similar devastation of the native snail fauna by *E. rosea* took place in the Pacific Society Islands, where the large majority of the 61 described endemic Society Islands partulid tree snails (four *Samoana* and 57 *Partula* species) are now extinct in the wild (Lee *et al.*, 2009).

Overharvesting and Overcollecting

Few terrestrial invertebrates are commercially harvested, and thus the impact on their populations from harvesting is much lower than that on marine invertebrates. Coconut crab (*B. latro*) has become rare or locally extinct due to overharvesting on many islands of the Solomon Islands, and harvesting quota are now in place in other parts of its range, including Guam and Vanuatu (**Figure 1**). *Cornu aspersum*, *Helix pomatia*, and *Iberus* spp. are terrestrial snails frequently harvested for human consumptions throughout Europe. It is difficult to get reliable consumption numbers, but estimates for France, the main consumer, range between 20 and 40,000 t of snail meat being sold per year. While the first two species are not threatened, many species of *Iberus* are highly localized, and their overcollecting is bound to have negative impact on their populations (Cuttelod *et al.*, 2011).

Several species of katydids (Tettigoniidae) and crickets (Gryllidae) are collected and traded as singing or fighting insects in China. There is evidence that this trade is causing a reduction in the ranges and populations of some species as the demand for these insects occasionally reaches 100,000 specimens *per day* during the fighting season (Jin and Yen, 1998).

Collecting invertebrates for live pet trade has had some impact on populations of tarantulas and other arachnids. Malaysian trapdoor spiders (*Liphistius* spp.) have been locally overcollected and are now legally protected. Giant African scorpions (*Padinus* spp.) are also collected in massive numbers for the pet trade, and their exportation from the countries of origin is subject to strict quota.

International trade in a number of species of Birdwing butterflies (*Ornithoptera* and related genera) and South African Stag beetles (*Colophon* spp.) is now restricted, although most of the specimens of butterflies in the insect collector trade come from commercial captive breeding facilities. Overcollecting has lead to at least one case of local species extinction: the British Large Copper Butterfly (*Lycaena dispar dispar*) appears to have been collected out of existence by 1848 (although this species is still widespread in other parts of Europe). Overall, however, there is little evidence of a negative impact of collectors of insect and other terrestrial invertebrates on their survival, especially if compared to sweeping losses of species caused by deforestation and other anthropogenic impacts. Rather, insect collecting is often seen as a gateway activity to a greater awareness of environmental issues, and a valuable educational tool for the next generations of naturalists and conservationists (Naskrecki, 2005).



Figure 1 Coconut crab (*Birgus latro*) is the largest living terrestrial invertebrate, reaching the weight of up to 4.1 kg and leg span of up to a meter. These animals have been overharvested on many Pacific and Indian Ocean islands and disappeared from across some of their former range.

Status Assessment of Endangered Invertebrates

Despite the overwhelming majority of invertebrates in virtually all terrestrial habitats, or rather because of it, their diversity is not reflected proportionately in formal conservation efforts and documents. The assessment of invertebrate species endangerment is conducted by a mix of local, national, and international bodies and organizations, some of which have legal, executive power (e.g., federal list of endangered species in the USA), whereas others only provide recommendations for protection (e.g., IUCN Red Lists.) These lists use various, nonstandardized criteria for the assessment of the threat status of the invertebrates, and sometimes provide conflicting recommendations. Local lists (e.g., state lists in the USA, country-based Red Lists in Europe) frequently include species of invertebrates that appear threatened locally because their local occurrences are on the periphery of their range but are not threatened on the global or even country scale.

As of 2012, the IUCN Red List of Threatened species, the most prominent body that assesses the conservation status of species, includes 6113 species of invertebrates that can be classified as terrestrial, i.e., the species lives in a terrestrial environment during at least one phase of its developmental cycle. This includes 3844 species of insects, 2180 species of snails, 33 species of arachnids, 31 species of millipedes, 11 species of velvet worms, seven species of crustaceans, six species of earthworms, and one species of centipedes; overall the list includes less than 0.1% of the estimated number of terrestrial invertebrate species. In contrast, 34,779 species of vertebrates, or nearly 100% of known species, have been evaluated by the IUCN.

In addition to this global database of endangered species, local Red Lists are maintained by individual countries or regions (e.g., European Red List of Saproxylic Beetles, Carpathian List of Endangered Species) (Nieto and Alexander, 2010; Witkowski *et al.*, 2003), but the coordination between country and global Red Lists is not perfect. For example, the Brazilian Red List of terrestrial invertebrates includes 130 species, many of which do not appear on the global Red List. This is despite the fact that most of these species are endemic to Brazil and thus endangered on a global scale (Lewinsohn *et al.*, 2005).

Comprehensive assessment of even a small fraction of the global invertebrate diversity is not feasible, and for this reason, a simplified methodology has been introduced to get an overview of their status. The Red List Index (sampled approach) (SRLI) has been developed in order to determine the threat status and trends of lesser-known and less "charismatic" species groups, such as several insect orders, and is based on a comprehensive assessment of a small subsample of members of such groups (e.g., only 10% of all known species of

Papilionoidea butterflies are being evaluated); the results are then used to interpolate the general trends within populations of all its members.

There appears to be, however, a number of biases in conservation assessment efforts for terrestrial invertebrate taxa, which may give a false picture of species distribution and abundance (Cardoso *et al.*, 2001). For example, species tend to be monitored more in places where high species richness is to be expected. The selection of taxa being monitored or evaluated is strongly skewed toward "charismatic" and easily identifiable groups, such as butterflies, but these groups are typically species-poor and may not be representative of invertebrates as a whole. There is also a tendency to overestimate the overall species' rarity based on regional occurrence, which often falls on the margins of the species' range.

The Convention on International Trade in Endangered Species (CITES) is a multilateral treaty designed to ensure that international trade in specimens of wild animals and plants does not threaten the survival of the species in the wild, and it accords varying degrees of protection to more than 33,000 species of animals and plants. It currently lists 348 species of terrestrial invertebrates, albeit there is little correspondence between the CITES list, which has executive power to control the international trade in these organisms, and the IUCN Red List, which ranks species according to relatively rigorous criteria of current or potential threat. For example, of the 23 species of spiders listed by CITES, only one appears on the IUCN Red List (categorized as Lower Risk); conversely, of the 25 species of spiders listed as threatened by the IUCN, none is listed by CITES.

Approaches to Terrestrial Invertebrate Conservation

Coarse and Fine Filters in Invertebrate Conservation

The concept of conserving entire plant and animal communities in reserves was originally viewed as an efficient coarse filter approach to conserving biodiversity that would protect 85–90% of all species. This was based on the assumption that carefully selected areas, often containing so-called "umbrella species", will preserve enough habitat heterogeneity to sustain a large selection of unrelated taxa. More recently, the coarse filter has evolved to a concept of conserving species diversity by providing adequate representation (distribution and abundance) of ecological land units based on an understanding of their natural disturbance regimes. This coarse filter approach does not necessarily prescribe reserves but rather recognizes ecological processes and provides for a dynamic distribution of ecological units across the landscape over time (Samways, 2005). The reality of a matrix of agricultural landscapes and high fragmentation of the remaining natural habitats across most of Europe and many other developed areas makes it necessary to develop conservation strategies that focus on management and restoration of what remains. In the UK, landscape-scale conservation projects are now the main delivery mechanism to conserve its threatened Lepidoptera. It is achieved by providing advice to landowners and assistance with obtaining agri-environment or woodland grant schemes, and by habitat management under the guidance of the UK's Butterfly Conservation organization (Ellis *et al.*, 2011).

The complementary fine filter approach focuses on conserving individual rare or specialized species that slip through the coarse filter and are not necessarily protected in reserves. The Red List assessments are designed to pinpoint such species, and its fine filter assessments are conducted to evaluate whether sufficient amount and distribution of habitat is provided under the coarse filter strategy.

The creation in 1978 of the Tripuí Reserve in Minas Gerais province of Brazil to protect the population of a velvet worm (*Peripatus acacioi*) is an example of a conservation action based on a fine filter assessment of the status of a single species.

Surrogates in Terrestrial Invertebrate Conservation

Plants have long been used as environmental surrogates for invertebrate diversity, particularly that of insects. Because they are easier to identify and map than invertebrates, and there is an unquestionable dependence of most insects on their existence, species-rich plant communities can be used to designate protected areas with expectedly high levels of invertebrate diversity. It has been shown that species richness of even those invertebrates that do not directly rely on any particular plant species, such as dragonflies, is positively correlated with plant diversity. Plants have also been demonstrated to be much better predictors of invertebrate diversity than vertebrates. At the same time, there exist confounding issues, and spatial distribution of insects does not necessarily overlap completely with that of plants; for example, although there is a strong relationship between the distribution of butterflies and leafhoppers, and that of tallgrass prairie plants, no such correlation was found between other prairie plants and other insects. This suggests that focus on "hotspots" of plant diversity may not include many insect species at finer scales.

Plant surrogacy approaches usually emphasize indigenous vegetation, but in the absence of that, even exotic plants may provide suitable habitats for native invertebrate faunas: in the Azores exotic forests have provided alternative habitats suitable for native saproxylic beetles, whereas in South Africa pine plantations turned out to provide sufficient habitats for velvet worms (*Opisthopatus cinctipes*) that used to inhabit indigenous, but now largely gone coastal forests (Hamer *et al.*, 1997). Also in South Africa, organic vineyards have the potential to make an important contribution to arthropod conservation in the Cape Floristic Region at the field scale, although at landscape scale, the preservation of natural fragments in the vineyard landscape is the most effective measure to increase biodiversity.

Animals, and birds in particular, have been used as surrogates for invertebrate distribution and richness. At some spatial scale, birds and butterflies show similar trends in species richness and diversity. In Uganda, the selection of potential forest reserve sites based on either birds or butterflies was as effective at representing other taxa as the selection based on the richness of plants and a wide range of animal taxa. In the Azores species richness of four groups of endemic arthropods was a good predictor of overall species richness (Borges *et al.*, 2010), but in other parts of the world vertebrate or

single-taxon invertebrate distribution does not correlate strongly with that of invertebrates in general and may give a misleading picture of the overall biodiversity of an area.

''Umbrella'' Species

"Umbrella" or "flagship" species are species whose conservation is assumed to correlate positively with protection of other, unrelated species, or communities. This concept has been widely applied in the designation of nature reserves, despite the fact that little proof exists for the effectiveness of this approach in protecting broader biodiversity. There is some evidence that it may be effective in fine filter, species-level approaches to conservation of individual invertebrate communities, for example, European beetle *Osmoderma eremita* serves as a suitable umbrella species for other beetles associated with tree hollows.

Recovery Programs

Terrestrial invertebrate recovery can be conducted on both coarse- and fine-filter scales. Coarse filter approaches focus on restoration of habitats, such as the effort to restore stone cover and reduce grazing in French pseudo-steppe, which resulted in a great increase in the abundance and species richness of beetle species. Conversely, restoration of grazing in areas where plant succession eliminates savannah-like, open habitats often results in the increase of the species richness of native pollinators and other invertebrates.

At the species level, translocation of New Zealand wetas (*Deinacrida* spp.) to areas where invasive rats were eliminated has proven quite successful in restoring their populations (Stringer and Chappell, 2008). Captive breeding programs have been used to restore populations of locally declining or extinct species, although such programs can only be used for species with relatively simple life histories (Samways *et al.*, 2010). The Lord Howe Island stick insect (*Dryococelus australis*) had been thought to be extinct until a small population was found on a tiny Ball's Pyramid Islet (New 2012); a small fraction of the population has been collected and is now successfully breeding in captivity at Melbourne Zoo, although the captive animals have not yet been reintroduced into the wild. In the UK, two species of the Orthoptera, field cricket (*Gryllus campestris*) and wart biter katydid (*Decticus verrucivorus*), are captive bred and have been reintroduced to their former, suitable habitats. (**Figure 2**).



Figure 2 European wart biter katydid (*Decticus verrucivorus*) virtually disappeared across its former range in Britain and has been subject to an experimental captive breeding program. Although the program was successful and the katydids were subsequently reintroduced to their natural habitats, captive breeding of this species is costly and labor intensive.

Endangered Terrestrial Invertebrates Listed by the Global IUCN Red List and CITES

Of the approximately 12 currently recognized classes of invertebrates that include terrestrial or partially terrestrial species, eight classes include species whose level of endangerment has been evaluated by the IUCN Red List. This includes species assigned to the following threat categories: extinct – 261 species, extinct in the wild – 12 species, critically endangered – 299 species, endangered – 347 species, vulnerable – 836 species, near threatened and least concern – 3036 species, and data deficient – 1359 species.

Earthworms (Oligochaeta)

The IUCN Red List includes eight species of earthworms, of which four are listed as vulnerable and one as extinct; no species of earthworms are listed by CITES. Earthworms are threatened primarily by habitat loss and agricultural development; the population of the Australian Giant Gippsland Earthworm (*Megascolides australis*) has apparently declined following the switch from cattle ranching to crop-based agriculture. The Giant Palouse Earthworm (*Driloleirus americanus*) of NW USA is considered vulnerable due to habitat loss but also because of the competition from nonnative, introduced species of earthworms. In Europe and the USA, the invasive New Zealand planarian (*Artioposthia triangulata*) that feeds on earthworms has apparently caused a decline of native species.

Velvet Worms (Onychophora)

Eleven species of these distant relatives of arthropods are considered threatened on the global scale by the IUCN Red List, although additional species appear on local, country, or regional lists. Velvet worms are threatened by several factors. In Tasmania, government-mandated planned fires have lead to local extirpation of several endemic species and may eventually contribute to their extinction. Overcollecting of these animals to meet the demands for teaching specimens for invertebrate zoology classes at Australian universities has apparently caused local declines, whereas in South Africa urbanization of native habitats has caused decline, and perhaps extinction of species. Conversely, in Brazil, a local nature reserve was created to protect the endemic species *P. acacioi* (**Figure 3**).



Figure 3 Velvet worms (Onychophora), like this individual from Daintree Rainforest of Queensland, Australia, include many narrowly endemic species with specific microhabitat requirements. Several species of these animals are probably already extinct, whereas many others are known only from single, small populations.

Spiders, Daddy Longlegs, and Pseudoscorpions (Arachnida)

Thirty-three species of arachnids are listed, 30 of which are spiders. Many have been placed on the list because they are narrow range endemics (e.g., species restricted to single caves), rather than because of a demonstrable decline in their population sizes. Several tarantula species, such as the Indian *Haploclastus kayi*, are restricted to very few locations in a rapidly shrinking natural habitat. Interestingly, there exists the overlap of only one species, Red-knee tarantula (*Brachypelma smithi*), between the IUCN Red List and CITES; the latter includes 22 species of tarantulas and three species of scorpions, all threatened by collecting for the exotic pet trade.

Centipedes (Chilopoda)

Only one species, Serpent Island Centipede (*Scolopendra abnormis*), is listed by the IUCN Red List, and no species are on the CITES list. This species is endemic to Mauritius and considered vulnerable.

Millipedes (Diplopoda)

The IUCN Red List includes 31 species of millipedes, all placed in the southern African genus *Doratogonus*. They are threatened mostly by the habitat loss caused by the extraction of wood by the local community for fuel and building.

Crabs, Pillbugs, and Woodlice (Crustacea)

Seven species of terrestrial crustaceans are listed by the IUCN Red List. The largest terrestrial invertebrate, Coconut crab (*B. latro*) from the Pacific and Indian tropical islands, is listed as Data Deficient, although there is evidence of its decline and local extirpation on some islands. The remaining six species are woodlice (Isopoda) with narrowly restricted ranges.

Insects (Insecta)

Insects constitute the majority of terrestrial invertebrates listed by the IUCN Red List, with 3844 species. Most of them are dragonflies and damselflies (2657 species), threatened by the disappearance of suitable, unpolluted aquatic habitats required for the development of their larval stages. Butterflies and moths are represented by 720 species; most are threatened by habitat loss and fragmentation, climate change, increased frequency and intensity of fires, and tourism development; some species are threatened by collecting for trade in preserved butterflies. Beetles are represented by 220 species, mostly those with narrow, endemic ranges and few known populations; some species, such as the American burying beetle (*N. americanus*) or violet click beetle (*Limoniscus violaceus*), have extensive ranges but are known only from small, widely scattered relict populations. The presence of the Pygmy Hog Sucking Louse (*Haematopinus oliveri*) from India on the list illustrates the point that the decline of the host organism's populations; in this case, the Critically Endangered Pygmy Hog (*Porcula salvania*) leads to the decline of their symbionts.

CITES lists 271 insect species and subspecies, primarily narrow endemics or species threatened by overcollecting.

Land Snails (Gastropoda)

This is the second best represented group on the IUCN Red List, with 2180 evaluated species. This includes virtually all endemic snails from Hawaii and the Society Islands (many of which are now considered extinct), decimated by the introduction of alien, predaceous snail *E. rosea*. A large proportion of the snails listed here are also European endemics, often restricted to small Mediterranean islands and threatened by urbanization and tourism industry. The CITES list includes 51 species of Pacific island land snails.

See also: Endangered Freshwater Invertebrates. Endangered Marine Invertebrates. Invertebrates, Terrestrial, Overview. Terrestrial Ecosystems

References

Blight O, Fadda S, Orgeas J, Ponel P, Buisson E, and Dutoit T (2011) Using stone cover patches and grazing exclusion to restore ground-active beetle communities in a degraded pseudo-steppe. *Journal of Insect Conservation* 15: 561–572.

Bönsel AB and Sonneck A-G (2011) Habitat use and dispersal characteristic by *Stethophyma grossum*: The role of habitat isolation and stable habitat conditions towards low dispersal. *Journal of Insect Conservation* 15: 455–463.

Borges PAV, Serrano AR, and Quartau JA (2010) Ranking the Azorean Natural Forest Reserves for conservation using their endemic arthropods. *Journal of Insect Conservation* 4: 129–147.

Cardoso P, Borges PAV, Trianti KA, Ferrández MA, and Martín JL (2011) Adapting the IUCN Red List criteria for invertebrates. *Biological Conservation* 144: 2432–2440.

Cuttelod A, Seddon M, and Neubert E (2011) *European Red List of Non-marine Molluscs*. Luxembourg: Publications Office of the European Union.

Dunn RR and Danoff-Burg JA (2007) Road size and carrion beetle assemblages in a New York forest. *Journal of Insect Conservation* 11: 325–332.

Ellis S, Wainwright D, Berney F, Bulman C, and Bourn N (2011) Landscape-scale conservation in practice: Lessons from northern England, UK. *Journal of Insect Conservation* 15: 69–81.

Hamer ML, Samways MJ, and Ruhherg H (1997) A review of the Onychophora of South Africa, with discussion of their conservation. *Annals of Natal Musuem* 38: 283–312.
Jin X-B and Yen AL (1998) Conservation and the cricket culture in China. *Journal of Insect Conservation* 2: 211–216.
Kisdi E (2002) Dispersal: Risk spreading versus local adaptation. *The American Naturalist* 159: 579–596.
Lee T, Burch JB, Coote T, *et al.* (2009) Moorean tree snail survival revisited: A multi-island genealogical perspective. *BMC Evolutionary Biology* 9: 204.
Lewinsohn TM, Lucci Freitas AV, and Prado PI (2005) Conservation of terrestrial invertebrates and their habitats in Brazil. *Conservation Biology* 19: 640–645.
Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B (2011) How many species are there on Earth and in the Ocean? *PLoS Biology* 9(8): e1001127.http://dx.doi.org/10.1371/journal.pbio.1001127.
Nakamura Y (2011) Conservation of butterflies in Japan: Status, actions and strategy. *Journal of Insect Conservation* 15: 5–22.
Naskrecki P (2005) *The Smaller Majority*, 288 pp. Harvard University Press.
New TR (2010) *Insect Conservation and Islands*, 258 pp. Springer.
New TR (2012) *Insect Conservation: Past, Present and Prospects*, 437 pp. Springer.
Nieto A and Alexander KNA (2010) *European Red List of Saproxylic Beetles*. Luxembourg: Publications Office of the European Union.
Samways MJ (2005) *Insect Diversity Conservation*, 356 pp. Cambridge University Press.
Samways MJ and Lockwood JA (1998) Orthoptera conservation: Pests and paradoxes. *Journal of Insect Conservation* 2: 143–149.
Samways MJ, McGeoch MA, and New TR (2010) *Insect Conservation: A Handbook of Approaches and Methods (Techniques in Ecology & Conservation)*, 432 pp. Oxford University Press.
Stringer IAN and Chappell R (2008) Possible rescue from extinction: Transfer of a rare New Zealand tusked weta to islands in the Mercury group. *Journal of Insect Conservation* 12: 371–382.
Van Swaay C, Cuttelod A, Collins S, *et al.* (2010) *European Red List of Butterfies*. Luxembourg: Publications Office of the European Union.
Watts C, Stringer I, Sherley G, Gibbs G, and Green C (2008) History of weta (Orthoptera: Anostostomatidae) translocation in New Zealand: Lessons learned, islands as sanctuaries and the future. *Journal of Insect Conservation* 12: 359–370.
Witkowski ZJ, Król W, and Solarz W (eds.) (2003) *Carpathian List of Endangered Species*. Vienna-Krakow: WWF and Institute of Nature Conservation, Polish Academy of Sciences.

Endemism

RM Cowling, University of Cape Town, Rondebosch, South Africa

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 497–507, © 2001, Elsevier Inc.

Glossary

Biotope Region that is distinguished by particular environmental conditions (climate, soil, altitude, etc.) and therefore a characteristic assemblage of organisms.

Stenotopic Referring to taxa with restricted habitat requirements (i.e., confined to a single biotope) and hence restricted distributions.

Endemic taxa are those restricted to a specified geographical area. Therefore, the concept is a relative one; the patterns, correlates, and causes of endemism will vary according to the size and location of the geographical area, as well as the taxonomy and phylogenetic relatedness of the assemblage under consideration. At a global scale, all taxa are endemic and there is relatively little to say on the topic. Most research has focused on species that are endemic to relatively small areas. In this sense, endemism is best viewed as a form of rarity, that is, range-restricted rarity. This article presents biogeographical, evolutionary, ecological, and conservation perspectives on endemism and discusses generalizations regarding the patterns, correlates, and causes of species-level endemism in relatively small areas.

Categories

Endemics may be categorized according to their spatial distribution, inferred evolutionary age, affinities, and local abundance.

Spatial Distribution

Endemics are loosely and commonly categorized in four contexts of spatial distribution: site or restricted area; biotope; biogeographical region; and political area.

Evolutionary Age and Affinity

Categorization of endemics according to evolutionary age and affinity is summarized in **Box 1**. These schemes have been widely used by botanists but rarely by zoologists. Problems associated with the schemes are that (1) age is regarded as a categorical rather than continuous variable; (2) the establishment of relationships among taxa lacks rigor; and (3) many palaeoendemics are diploid.

Phylogenetic methods, which consider the distribution of characters among taxa in a cladistic context, provide a rigorous categorization of endemics in terms of relative age and propinquity of descent. In this context, low-ranking taxa correspond to neoendemics and high-ranking taxa to palaeoendemics (cf. **Box 1**). An absolute estimate of the age of endemics can be given when congruent phylogenetic relationships correlate with identifiable historical events.

Local Abundance

The classical, biogeographical perspective on endemism has tended not to consider the local abundance of species. However, in the more recently developed ecological and conservation perspectives, in which endemism is conceived as a category of rarity, population abundance is invariably explicitly considered. Thus, geographical range size as a categorical variable (wide/narrow) has been used as one of the

Box 1 Categorization of endemics according to evolutionary age and affinities

A. Engler's scheme, published in 1882.

1. Neoendemics: comprising clusters of closely related species and subspecies that have evolved relatively recently.
2. Palaeoendemics: comprising phylogenetically high-ranking taxa, usually monotypic sections, subgenera, or genera that may be regarded as evolutionary relicts.

R. C. Favarger and J. Constandriopoulos's scheme, published in 1961. This scheme uses cytological data to provide a more rigorous basis for assessing the age and affinities of endemics.

1. Palaeoendemics: ancient isolated taxa with a high ploidy level, whose diploid ancestors are extinct or unknown.
2. Schizoendemics: vicariant species of equal ploidy level, resulting from either gradual or rapid divergence.
3. Patroendemics: restricted diploid species that have spawned younger, widespread polyploid species.
4. Apoendemics: polyploid endemics that are derived from widespread species of a lower ploidy level.

Schizo-, patro-, and apoendemics are further subdivisions of Engler's neoendemics.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00044-7

factors in defining seven forms of rarity recognized for plants. Endemics (narrow range) may belong to any four categories of rarity according to biotope specificity (broad/restricted) and local population size (somewhere large/everywhere small).

Perspectives

The concept of endemism has a long history in biology, dating back to A. P. De Candolle's treatise, published in 1820. Most research on the topic has been in the field of descriptive biogeography, where distribution patterns of taxa have been used to define centers of endemism at various spatial scales. This approach provides a static perspective of endemism.

Over the past few decades, historical biogeographers have evaluated areas of endemism for monophyletic lineages in a phylogenetic context. This approach provides a dynamic perspective of endemism, especially when endemic taxa show congruent phylogenetic relationships that can be correlated with historical events.

Evolutionary biologists, studying both fossil and extant biotas, have explored the role of range restriction as a cause and consequence of speciation. Recently, several statistical techniques have been employed—collectively termed the comparative method—to exploit the phylogenetic relationships among species to extract independent information on the evolutionary correlates of endemism. These techniques acknowledge that related species may have similar range sizes, that is, range size cannot be assumed to be independent among species. However, in at least some cases, variance in range sizes seems to be partitioned mostly at the species level.

Community ecologists have conceptualized endemism as one of several forms of rarity, namely, range-restricted rarity, and have explored its role as an explanatory variable for taxon-specific ecological traits, such as local population size, body size, reproductive fitness, and dispersal distance. Increasingly, they are using comparative methods to correct for phylogenetic relatedness among biotas. However, for every cause–effect relationship documented, there are numerous exceptions.

Conservation biologists view range-restricted rarity as an attribute that predisposes a taxon to extinction. They seek to understand the abiotic and biotic correlates of this form of rarity as a basis for management guidelines that will reduce rates of extinction. A distinction is often, although not always, made between naturally rare species that may have some adaptation to rarity and those that have previously been widespread and are now restricted. Conservation planners often use patterns of endemism to identify reserve systems that are representative of a region's biodiversity. Many reserve selection algorithms have been formulated to select sites that have a unique or endemic complement of species.

Measurement

In quantifying patterns of endemism, the units of measurement (spatial scale and taxonomic entity), the mode of reporting of the data (percentages or counts), and a number of biases all influence the interpretation of the results. Of great importance is the relative nature of endemism: evaluation is always dependent on the spatial context and biological assemblage under consideration. This section provides a clarification of the problems and approaches associated with the measurement of endemism.

Units of Measurement

A variety of methods have been used to measure the range sizes of taxa. A useful distinction can be made between measures that attempt to estimate the extent of occurrence of a taxon—the distance between the outermost limits of a species' occurrence—and the area of occupancy—the area over which the species is actually found. The latter measure is particularly relevant for ecological studies that seek correlations between range size and environmental tolerances, as well as for conservation planning research; extent of occurrence is widely used in biogeographic studies.

Measures of endemism invariably seek to identify a subset of taxa within an assemblage that can be classified as having a lower than average range size. Within the biotas of larger-scale regions—biogeographic areas or countries—many researchers have recognized "local" endemics as a distinct category. However, the range size, or extent of occurrence, for defining this category is often arbitrarily set, varying between 50,000 km^2 (for Neotropical birds and plants, as well as for birds globally) to 2000 km^2 or less (for plants in the Cape Floristic Region). Endemics with extremely small range sizes—< 5 km^2—are regarded as point endemics. An approach that is increasingly being used is to evaluate endemism as a continuous variable, calculated as the sum of the inverse range sizes of all taxa in each quadrat (cell grid or map unit).

From both the biogeographical and ecological perspectives, patterns of endemism are best studied in relation to ecologically homogeneous, biogeographical regions. However, conservation planners often use political regions or property boundaries when evaluating endemism, since these may be the most effective decision-making unit for the preservation of endemics.

The taxonomic or phylogenetic scales employed also influence patterns of endemism. Centers of endemism identified on the basis of patterns among low-ranking taxa (sub-species or closely related species) often differ from those where the units are high-ranking members of the same lineage. Similarly, the spatial scale for defining endemism will vary among different taxa of the same rank.

Percentage versus Counts

Endemism may be expressed as a percentage of all extant taxa present, or as the absolute number of endemics in an area. Depicting plant endemics in biogeographic regions as percentages or counts, and using area and latitude as explanatory variables, results in different patterns with different significance (Figure 1). Some species-poor areas, such as oceanic islands and arid regions, although low in actual numbers of endemics, may support a high percentage of endemic taxa. Other areas, such as Madagascar, the Cape Floristic Region, and parts of the Neotropics combine high richness and high endemism for some taxa. Ideally, both measures

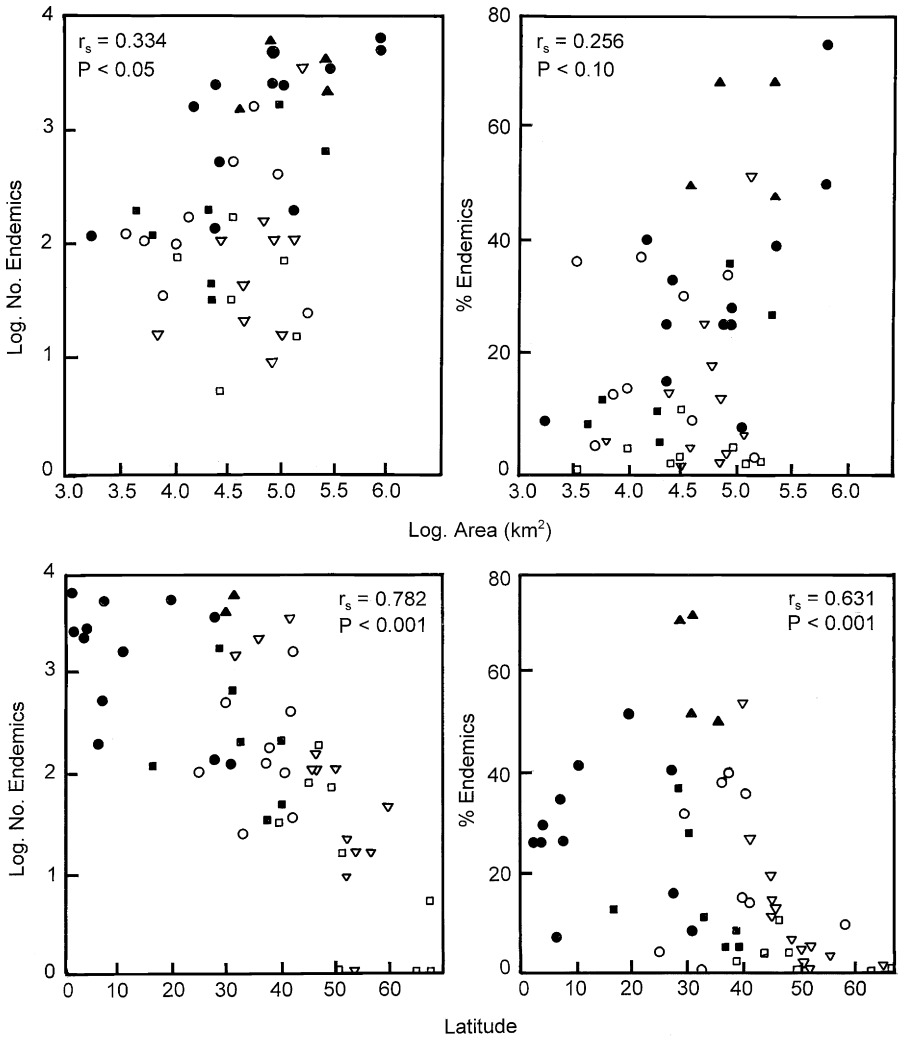


Figure 1 Relationships (Spearman's rank correlation) between two measures of endemism (percentages and counts) and area and latitude for plants in 52 biogeographical units in tropical and subtropical forests and savanna (●), temperate forest and woodland (○), Mediterranean-climate shrubland and woodland (▲), warm desert and steppe (■), cold desert and steppe (□), and boreal forest and tundra (▽) on continental landmasses across the globe. Reprinted from Cowling RM and Samways MJ (1995) Endemism and biodiversity. In: Heywood VH (ed.) *Global Biodiversity Assessment*, pp. 174–191. Cambridge, United Kingdom: Cambridge University Press.

of endemism should be considered when explaining patterns, but seldom are.

Biases

Endemism is influenced by taxonomic interpretation, sampling error, and human perceptions of rarity. Of particular importance is the fact that limited geographical exploration, as well as variation in the application of taxonomic concepts, introduces biases in the identification of endemics and the significance of their status. Pseudoendemics are widespread species incorrectly classified as endemics, whereas non-apparent endemics are endemic species that are incorrectly classified as widespread. The fact that widespread species are usually more thoroughly researched than those with smaller range sizes introduces biases in studies that explore the correlates of range size.

Patterns

There are very clear global and regional patterns of endemism for a wide range of taxa: endemics are not randomly distributed across the globe. However, these patterns are constrained by poor taxonomic knowledge and distributional data in key areas (e.g., the tropics) and for some taxa (e.g., most invertebrate groups).

Latitudinal Gradients

The incidence of endemism for whole assemblages in biogeographic zones increases with decreasing latitude (see **Figure 1**). Range sizes, as measured by latitudinal extent, increase for a wide range of organisms above a latitude of approximately 40°–50° N, but the same patterns are not evident in the Southern Hemisphere. There are many patterns that are not consistent with the generalization—termed Rapoport's Rule—that range sizes of taxa

decrease with decreasing latitude, as a consequence of greater ecological specialization in less seasonal environments. For example, very high endemism for terrestrial taxa is recorded in the mid-latitudes of the Southern Hemisphere, particularly in and adjacent to Mediterranean-climate regions. Marine teleost fishes have smaller range sizes at higher than at lower latitudes, and endemism for marine algae peaks in mid-latitude areas. These patterns are probably more a product of speciation and extinction processes than contemporary ecological conditions. Thus, widespread glaciation during the Pleistocene at high latitudes in the Northern Hemisphere resulted in the extinction of less tolerant terrestrial taxa. In mid-latitude Mediterranean-climate regions that escaped glaciation, rates of speciation, at least for plants, have overwhelmed extinction rates, resulting in an accumulation of habitat-specialist, range-restricted species.

Centers

Many centers of endemism—areas of higher than average concentrations of range-restricted taxa—have been recognized globally and regionally, principally for higher plants and large-bodied terrestrial vertebrate faunas. Generally, many groups of organisms show a concentration of centers at lower latitudes (**Figure 2**). Following from the previous section, it is no surprise that the high-latitude areas of the Northern Hemisphere support few centers. However, this is not always the case in the Southern Hemisphere, where large numbers of range-restricted taxa occur in middle- to high-latitude land-masses that were never glaciated during the Pleistocene.

Congruence

Overlapping or congruent areas of endemism for different taxa have been used extensively by biogeographers to reconstruct historical events. Patterns of congruence of endemism are also important for identifying reserve systems that maximize the preservation of different biotas.

Although strong patterns of congruence have been recognized for some taxa at the global scale—for example, swallowtail butterflies and tiger beetles, amphibians, birds, and mammals—higher plant centers often do not coincide with faunal centers. Nonetheless, on the basis of congruent patterns of endemic species diversity for mammals, reptiles, amphibians, and higher plants, it has been possible to identify 17 "megadiversity countries" (**Table 1**), that is, political units of high conservation value.

At a finer scale, patterns are highly variable among different taxa and in different regions, and no generalizations have emerged. This lack of strong congruence underlines the fact that endemism is an expression of many different causes, both ecological and historical.

Correlates and Causes

Range size or degree of endemism shows some clear relationships with a wide array of abiotic and biotic factors. These correlations are very useful in conservation biology since they may be used to identify factors that predispose endemic species to extinction. However, correlates may be either a cause or a consequence of endemism. To identify the causes of endemism in an evolutionary context, comparative methods that exploit phylogenetic relationships must be employed.

The causes of endemism are complex and numerous, and include intolerance of widespread habitats, niche specialization, isolation in marginal habitats owing to climate change, phylogenetic predisposition to narrow habitat selection, competition from alien species, and recent speciation of

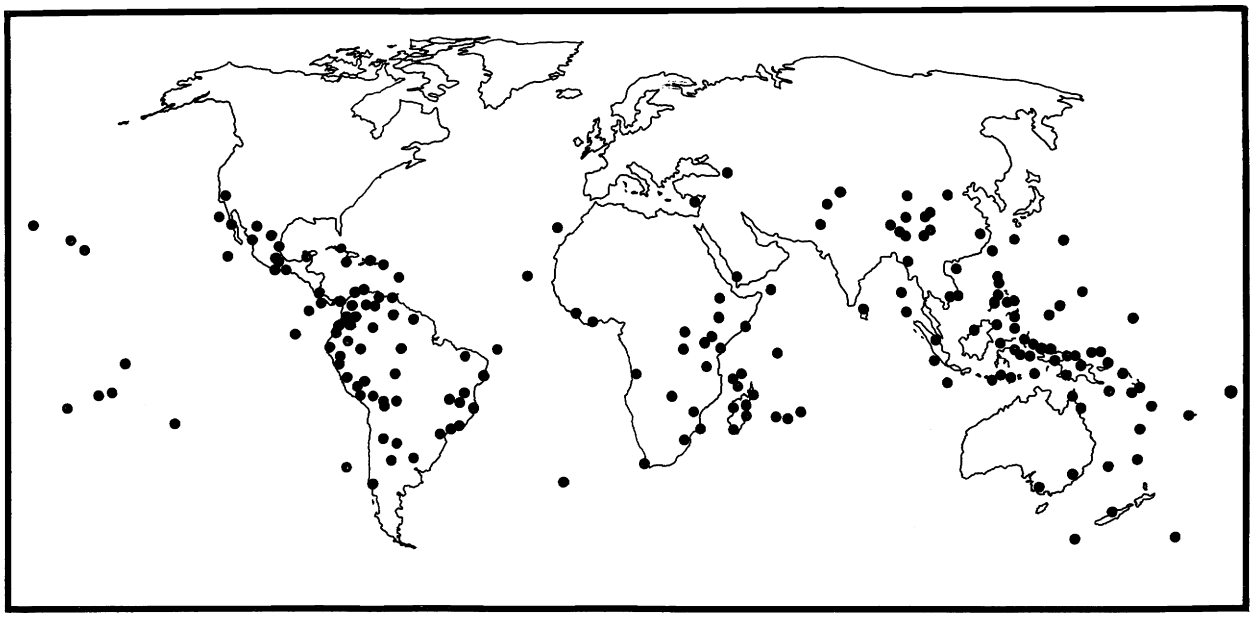


Figure 2 The distribution of Endemic Bird Areas of the globe, as recognized by Birdlife International. These centers are identified on the basis of the distributions of 2609 bird species that have had in historical times a global breeding range of less than 50,000 km^2. Reprinted with permission from Bibby CJ, Crosby MJ, Heath MF, *et al.* (1992) *Putting Biodiversity on the Map: Global Priorities for Conservation.* Cambridge: ICBP.

Table 1 Vertebrate and higher plant endemism in the world's 17 megadiversity countries

Country	*Area ($km^2 \times 10^3$)*	*Mammals*	*Birds*	*Reptiles*	*Amphibians*	*Plants*
Brazil	8512.0	131 (4)[a]	>191 (3)	172 (5)	294 (2)	ca. 17,500
Indonesia	1916.6	201 (2)	397 (1)	150 (6)	100 (11)	ca. 16,000
Colombia	1141.7	28	>142 (5)	97 (11)	367 (1)	ca. 16,000
Mexico	1972.5	140 (3)	125 (6)	368 (2)	169 (6)	ca. 12,500
Australia	7686.8	210 (1)	355 (2)	616 (1)	169 (5)	14,458
Madagascar	587.0	77 (8)	103 (8)	274 (3)	176 (3)	ca. 9200
China	9561.0	77 (7)	99 (9)	133 (7)	175 (4)	ca. 10,000
Philippines	300.8	116 (5)	183 (4)	131 (8)	44	ca. 5000
India	3287.8	44 (12)	52 (12)	187 (4)	110 (10)	ca. 7500
Peru	1285.2	46 (11)	109 (7)	98 (10)	>89 (12)	5356
PNG Papua New Guinea	475.4	57 (9)	85 (10)	79	134 (8)	ca. 13,000
Ecuador	283.6	21	37	114 (9)	138 (7)	ca. 4500
United States	9372.1	101 (6)	71 (11)	90	126 (9)	4036
Venezuela	912.1	11	45	57	76	ca. 6000
Malaysia	329.7	27	11	68	57	ca. 7250
South Africa	1221.0	27	7	76	36	16,500
Democratic Republic of Congo	2344.0	28	23	33	53	3200

[a]Figures in parentheses are rankings for the number of endemic species among the top 12 countries.

isolates in marginal habitats. Therefore, historical processes, contemporary ecological factors, and inherent biological properties of lineages are involved. In many cases, historical factors may be overriding, resulting in a poor relationship between measures of endemism and explanatory variables reflecting the contemporary environment.

Establishing correlates is a useful step in explaining patterns and causes of endemism. Most pertinent studies have addressed the following question: When compared to more widespread taxa, are endemics, however defined, a random subset of the biota with regard to abiotic and biotic factors? Developing these profiles, however, has been complicated by different definitions of endemism, multiple interactions between different traits, and a failure to consider phylogenetic relatedness. This section provides a brief review of the abiotic and biotic interspecific correlates of narrow range size, and concludes with an assessment of the role of endemism in speciation.

Regional Species Richness

There is often a positive relationship between the incidence of endemism and regional-scale richness. This results from the importance of high habitat-related and geographical compositional turnover (beta and gamma diversity, respectively) in producing regional richness. Habitat specialists (or stenotopic species) and geographical vicariants often have narrow range sizes. However, there are also many cases where patterns of endemism and diversity are largely noncoincident. Examples include plants in the Neotropics, birds in the Andes, dragonflies and terrestrial vertebrates in southern Africa, and the biotas of many oceanic islands and deserts.

Area

As a generalization, proportionate and absolute measures of endemism increase with increasing area (see **Figure 1**), irrespective of the taxonomic level. However, the relationship between number of endemic species (counts) and area is not as tight as that for the more widely studied species–area relationship. This results from the lack of congruence between endemism and richness in many areas (e.g., arid lands and oceanic islands).

Abiotic Environmental Factors

Levels of endemism may vary in a predictable way along gradients of rainfall, temperature, productivity, and habitat heterogeneity. Models that accurately predict levels of endemism on the basis of easily measurable environmental variables have been used for the rapid identification of endemic-rich areas.

For higher plants, levels of endemism increase with increasing productivity, with increasing elevation (reflecting increased habitat heterogeneity and isolation in high-altitude areas), and with higher rainfall in low-and middle-latitude areas, although many exceptions to these patterns exist. In the Mediterranean-climate regions of the Cape and southwestern Australia, there is a negative relationship between local endemism and soil fertility. In the California Floristic Province, palaeoendemics are clustered in the wettest and driest areas, whereas neoendemics occur in transitional rainfall areas where rates of speciation are highest. Similar patterns exist for Afrotropical birds and Neotropical butterflies. For a wide range of marine taxa, endemism is more pronounced in exposed and variable nearshore environments than in the more stable distant-shore habitats.

Biotope

Geographically isolated areas and biotopes, such as certain islands, mountain peaks, ancient lakes, caves, thermal vents, hot springs, vernal pools, the abyssal zone, and chemically imbalanced substrata, support a disproportionately high number of stenotopic endemics.

Most studies have focused on endemism on islands, mountains, and unusual substrata. Generally, larger continental islands such as Madagascar, New Caledonia, and New Zealand support the greatest number and proportion of endemic taxa, especially of higher plants. Elevational range explains the incidence of plant endemism on the Canary Islands and bird endemism in Indian Ocean archipelagos, suggesting the importance of topographical diversity. Continental islands are typically rich in palaeoendemics, whereas some taxa have undergone extensive and unusual adaptive radiation on oceanic islands such as the Canaries and Hawaii.

Mountains are also often rich in endemics, in both tropical and temperate regions, but not in recently glaciated, high-latitude areas of the Northern Hemisphere. Many desert inselbergs (granitic outcrops) act as mesic refugia that support endemics; this is particularly pronounced for plants in middle Asia. As on islands, endemism on mountains results from both historical (e.g., isolation) and ecological (e.g., heterogeneity) factors.

The restriction of endemic plant species to nutritionally imbalanced substrata, especially when these occur in an islandlike configuration, is widespread in Mediterranean-climate and humid tropical regions. These sites provide both a strong selective force for the evolution of neoendemics and a refuge from competition for palaeoendemics. The restriction of animal taxa to unusual substrata has not been studied in any detail, but is likely to be a response to habitat effects on vegetation structure rather than nutritional peculiarities per se.

Biology

Very few studies have addressed the relationships between restricted range size and biological factors such as body size, growth form, life-history traits, population size, and genetic architecture. Of these studies, few have considered multiple trait interactions or phylogenetic relatedness.

As a generalization, there is a positive, albeit weak relationship between a species' range size and its local population abundance for a wide range of taxa (**Figure 3**). However, not all endemics have low local abundances; indeed, many narrow plant endemics are extremely abundant locally. There are a number of hypotheses to explain the positive relationship between range size and local abundance. These are based principally on artefacts (e.g., sampling effects), resource use, metapopulation dynamics, and spatially independent rates of

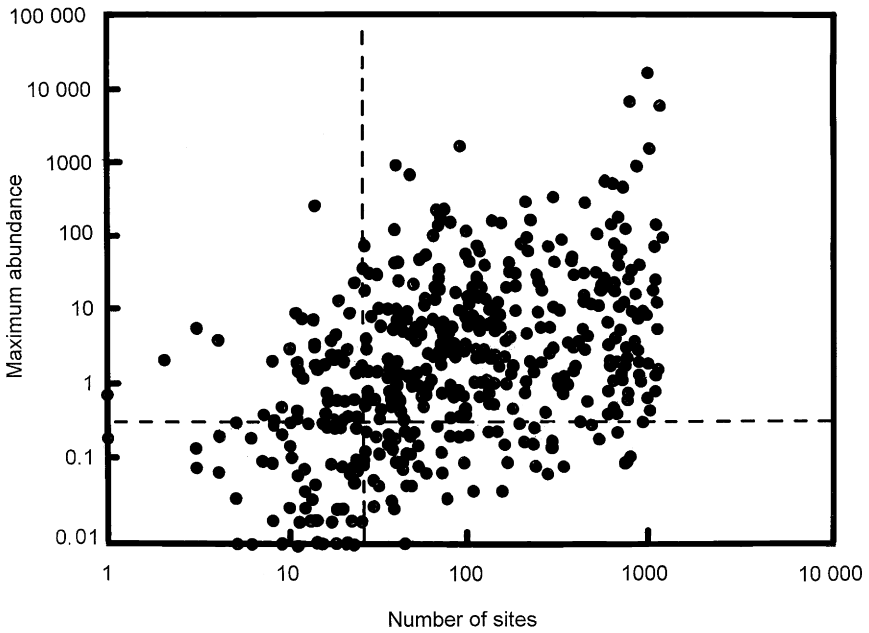


Figure 3 Relationship between the maximum local abundance value for North American birds and the number of sites at which each was recorded. Reprinted from Gaston KJ (1994) *Rarity*. London: Chapman & Hall, with permission from Springer.

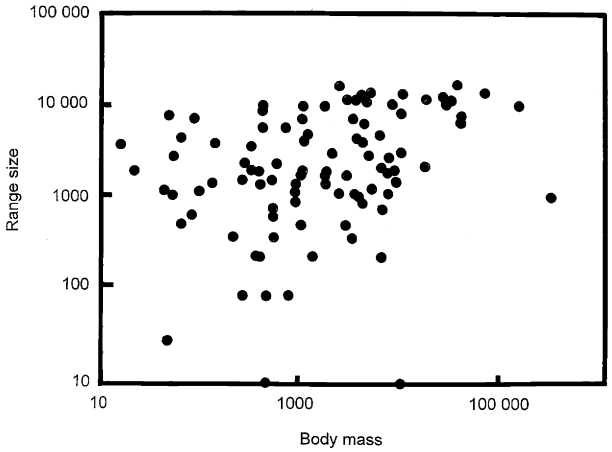


Figure 4 Relationship between the geographic range size ($km^2 \times 1000$) and the body mass (g) of Neotropical forest mammals. Reprinted from Gaston KJ (1994) *Rarity*. London: Chapman & Hall, with permission from Springer.

population growth. This pattern and its causes are currently attracting considerable attention.

There is a broad positive relationship between geographical range size and body size for animal species (**Figure 4**). There is also a trend, both within regional floras and specific taxa, for low-stature growth forms to be overrepresented among plant endemics. This is especially true of the South African Mediterranean-climate region, where low shrubs (**Figure 5**) dominate the endemic flora. Among Neotropical forest plants, endemics tend to be herbs, shrubs, or epiphytes rather than forest trees, whereas in the rain forests of Sri Lanka, endemics are overrepresented among long-lived, late-successional trees.

Gigantism is a common feature among some plant groups endemic to alpine habitats at low latitudes. Among animals, gigantism, dwarfism, and flightlessness are widespread among island endemics, as well as among some continental endemics associated with insular biotopes.

The reproductive correlates of endemism have been more extensively studied than other biological attributes. There are a number of pertinent generalizations, although exceptions exist for all of them. Range-restricted species differ from common ones in that they:

- tend to be self-compatible or rely on asexual reproduction;
- tend not to be wind-pollinated or have other inefficient forms of pollen transfer;
- invest less in reproduction;
- have poorer dispersal abilities;
- have shorter generation times.

The last two attributes are shown in **Figure 5**, where short-distance ant dispersal and fire sensitivity (rapid generation time) are overrepresented among plant endemics in a mountain region of the Cape Floristic Region in South Africa. However, these and other reproductive traits, such as seed size, seed number, and reproductive investment, all interact in complex ways. Furthermore, these traits are not independently distributed among species.

Many studies indicate that plant and animal endemics have lower levels of genetic variation in comparison with

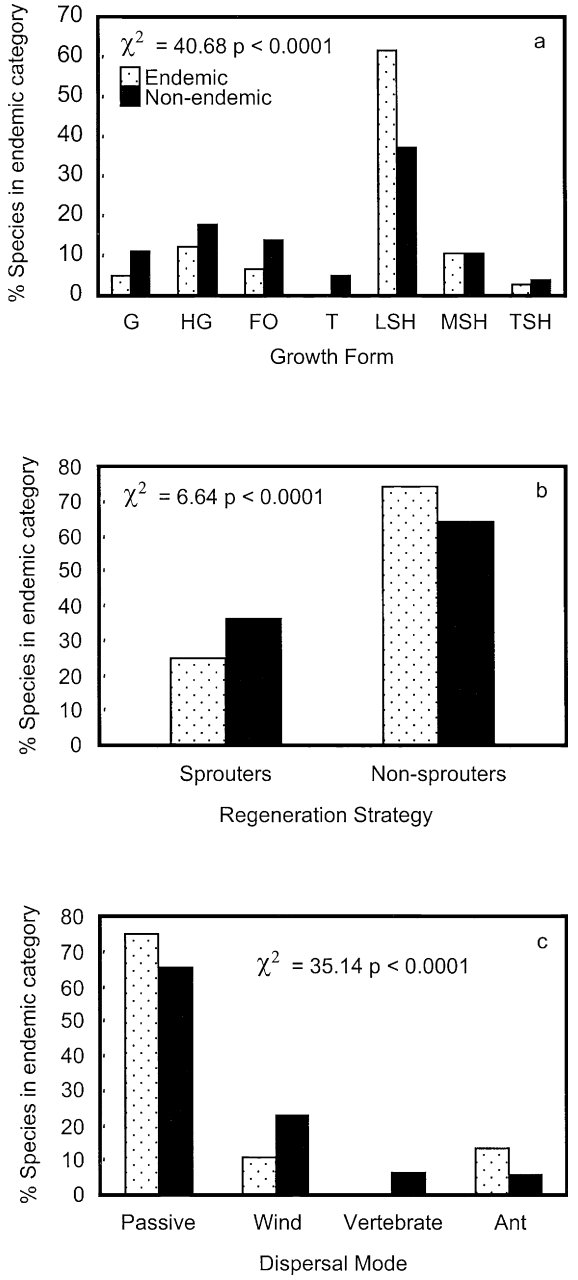


Figure 5 Percentage of endemic and nonendemic species in the Langeberg mountain flora (Cape Floristic Region, South Africa) in (a) seven growth form classes (G = geophyte, HG = graminoid, FO = forb, T = tree, LSH = low shrub, MSH = mid high shrub, TSH = tall shrub); (b) two postfire regeneration classes; and (c) four dispersal mode classes. Chi-square analyses were performed on untransformed data. Reprinted from McDonald DJ and Cowling RM (1995) Towards a profile of an endemic mountain fynbos flora: Implications for conservation. *Biol. Conserv.* 72: 1–12.

widespread congeners. This may be due to several factors, including adaptations to narrow ecological conditions, small population size, and self-incompatibility in plants. However, there are also cases of little difference in genetic diversity between closely related endemic and widespread plant species.

Taxonomy and Phylogeny

Many biotas that are endemic to biogeographic regions are not a random phylogenetic assemblage. Some plant families are significantly overrepresented among the endemic floras in many parts of the globe, especially in Mediterranean-climate regions. The same is true of certain dragonfly families in southern Africa. Among plants, Cyperaceae and Poaceae are underrepresented as endemics in many floras throughout the world. In many cases, these patterns can be attributed to taxon-specific biological attributes that predispose a lineage to endemism. Thus, the existence of discernible phylogenetic correlates of endemism implies that range size may be an evolutionarily stable character of a lineage. Hence there is a need for the comparative approach to assess the role of phylogenetic relatedness in explaining patterns. However, it is important to establish the taxonomic level at which these relationships are manifested. For example, for several data sets, the majority of variation in range size is explained at the level of species within genera.

Endemism and Speciation

At face value, the relationship between range size and speciation appears to be quite simple: a reduction in range size will always accompany a speciation event, and a species nearing extinction—in an advanced stage of the taxon cycle—will occupy a limited range size. The deeper issue of the extent to which range size is a cause or consequence of speciation is a question of considerable interest.

There has been a long-standing and as yet unresolved debate regarding the causal relationship between range size and speciation. The arguments assume positive relationships between range size, population size, and dispersal ability. One viewpoint suggests that owing to extensive gene flow and reduced extinction rates, widespread taxa should have lower rates of speciation than range-restricted taxa. An alternative hypothesis is that owing to greater genetic variability and a higher frequency of founder effects, species that comprise large and well-dispersed populations that occupy large range sizes are prone to vicariant speciation.

Many studies of fossil and extant lineages suggest that turnover (speciation and extinction) is associated with relatively low local population abundance, poor dispersal, and narrow range size. Clearly, at extremely low values for these variables, extinction rates will overwhelm rates of speciation. Elevated speciation and extinction rates are also associated with increased specialization, reduced body size, and increased generation times; all of these are correlates of narrow endemism. Thus, endemism and its correlates are responsible not only for enhanced rates of speciation, but also rapid rates of extinction. In E. S. Vrba's parlance, these processes are flip sides of the same coin.

The alternative view, that speciation is associated with large, centrally located and wide-ranging populations, and that peripheral isolates are relictual taxa, also has support. Ultimately, aspects of this debate will be resolved by studies that assess range size and its correlates in a phylogenetic context.

Conservation

Because of their restricted geographical range size, high habitat specificity, and generally low population abundance, endemics are more vulnerable to extinction than are widespread and common species, as a result of both deterministic (habitat transformation) and stochastic (small population effects) factors. Therefore, considerable attention has been given to the conservation of local endemics. Attempts have been made to use the correlates of local endemism to devise management plans that will reduce anthropogenic extinctions.

Recent advances in systematic conservation planning have identified priorities for conservation on the basis of complementarity of biotas (representation), but also for the retention of biodiversity in the face of threatening processes. This approach involves the assessment of the irreplaceability of an area—a measure of the likelihood that the area will be needed to achieve a conservation goal—and its vulnerability to biodiversity loss as a result of current or impending threatening processes. Endemic-rich areas inevitably emerge as priorities since they combine high irreplaceability, owing to their unique biota, and high vulnerability, since endemics are prone to extinction. However, some endemics, particularly plants, may be preadapted to persist in small populations and could be effectively preserved in small, fragmented areas.

Conclusions

There are few generalizations regarding geographical patterns and correlations of endemism. This is understandable, given that definitions of endemism are mostly study-specific, and that endemism is partly a consequence of regional-specific historical events acting on phylogenetically distinct biotas. Furthermore, species with similar range size often have different local abundances that are likely to be manifested in very different biological attributes. Finally, within-region analyses invariably lump together palaeoendemics and neoendemics, groups with different origins and phylogenetic relationships, and often, different biologies. The recent trend to correct for phylogenetic relatedness holds much promise for understanding the ecological and evolutionary correlates of endemism.

The most active fields of research currently are studies on the correlates of range size, particularly local population abundance, body size, and reproductive traits; the role of endemism in reserve selection, especially as a measure of irreplaceability and surrogate measure of vulnerability; and historical reconstructions using congruent areas of endemism in phylogenetic studies. Much less classical biogeographic research is being carried out on the identification of centers of endemism, despite the fact that reliable distribution data are lacking for many areas and taxonomic groups. This lack of data has serious consequences for the identification of endemic-rich areas for conservation purposes.

See also: Biodiversity-Rich Countries. Biogeography, Overview. Diversity, Community/Regional Level. Extinction, Causes of. Island Biogeography

References

Anderson S (1994) Area and endemism. *Quart. Rev. Biol.* 69: 451–471.

Bibby CJ, Crosby MJ, Heath MF, *et al.* (1992) *Putting Biodiversity on the Map: Global Priorities for Conservation*. Cambridge: ICBP.

Cowling RM and Samways MJ (1995) Endemism and biodiversity. In: Heywood VH (ed.) *Global Biodiversity Assessment*, pp. 174–191. Cambridge, United Kingdom: Cambridge University Press.

Gaston KJ (1994) *Rarity*. London: Chapman & Hall.

Kruckeberg AR and Rabinowitz D (1985) Biological aspects of endemism in higher plants. *Annu. Rev. Ecol. Syst.* 16: 447–479.

Kunin WE and Gaston KJ (eds.) (1997) *The Biology of Rarity. Causes and Consequences of Rare–Common Differences*. London: Chapman & Hall.

Major J (1998) Endemism: A botanical perspective. In: Myers AA and Giller PS (eds.) *Analytical Biogeography. An Integrated Approach to the Study of Animal and Plant Distributions*, pp. 117–146. New York: Chapman & Hall.

McDonald DJ and Cowling RM (1995) Towards a profile of an endemic mountain fynbos flora: Implications for conservation. *Biol. Conserv.* 72: 1–12.

Ricklefs RE and Schluter D (1993) *Species Diversity in Ecological Communities. Historical and Geographical Perspectives*. Chicago: University of Chicago Press.

Energy Flow and Ecosystems

Alan P Covich, University of Georgia, Athens, GA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Chemoautotrophs Species that use inorganic compounds as a source of carbon and energy, and function as primary producers.

Decomposition The biotic breakdown of dead organic matter (detritus) by bacteria and fungi that releases carbon dioxide and nutrients for recycling.

Ecosystems Composed of species assemblages (producers and consumers) that interact with each other and their associated abiotic environment within well-defined natural or conceptual boundaries.

Food chains Composed of species that are connected by the flow of energy and material from producers to consumers.

Food webs "Flow maps" that depict connections among multiple food chains.

Functional groups Aggregations of species that perform similar ecosystem processes, such as grazers, suspension or filter feeders, leaf shredders, predators, and decomposers.

Photoautotrophs These species, such as green plants and some bacteria, use solar energy and inorganic compounds to synthesize organic matter as primary producers.

Primary productivity The rate of synthesis of organic matter by plants (biomass per unit area of habitat per unit time or, in some cases, biomass per unit volume per unit time).

Secondary productivity The rate of assimilation and growth by animals (biomass per unit area of habitat per unit time or, in some cases, biomass per unit volume per unit time).

Trophic cascades These occur when changes in the presence or absence (or shifts in abundance) of a top predator alter the production at several lower trophic levels; primary and secondary production at lower levels are alternately constrained or unconstrained by the feeding activities of consumers at upper levels.

Trophic levels Groups of individuals classified as primary producers or primary or secondary consumers within food webs; individuals feeding both as primary and secondary consumers are omnivores. A single species may be represented on more than one trophic level.

Introduction

Ecosystems are thermodynamically open, hierarchically organized communities of producers, consumers, and decomposers together with the abiotic factors that influence species growth, reproduction, and dispersal. These abiotic factors include the flow of energy and the circulation of materials together with the geological, hydrological, and atmospheric forces that influence habitat quality, species distributions, and species abundances. Energy flows through many species, and the way in which this flow affects the persistence of ecosystems is influenced by land-use changes, precipitation, soil erosion, and other physical constraints such as geomorphology.

Energy flow through ecosystems is essential for nutrients to cycle through food webs. These food webs are often subwebs of more complex species assemblages and may be only partial descriptions of more complex hierarchies of energy flows. The connections among species' interactions in natural food webs typically result in some important feedback loops and recycling of nutrients and materials within the conceptually defined boundaries of an ecosystem. Many species of producers are generalized in using solar energy (photoautotrophs) and form the basis of most food webs. Certain species live in deep, ocean thermal vents and use inorganic chemical compounds as energy sources (chemoautotrophs). The connections among producers and consumers are often complex and change with species age and size during an individual's life history. These food webs are diagrams that can function as "flow maps" to document which species interact with other species, either directly or indirectly, as energy flows through the community and determines the movement of nutrients and other materials. This flow of energy is associated with the potential bioaccumulation of toxic chemicals that can be concentrated in predators who consume large amounts of prey that, in turn, have fed on large amounts of plants.

Different species have important functional values, such as for organic matter production (plant and animal growth) or organic matter breakdown (decomposition), oxygen production, nitrogen fixation, and nutrient cycling. These species have intrinsic values as the unique end products of evolution, and native species are likely to have adapted specific ways to respond to local or regional environmental changes or other physical disturbances. Conceptually, the loss of even a single native species, or the introduction of a nonnative species, could alter how the other remaining native species continue to perform different ecosystem functions. Disruptions of ecosystem processes are known to have occurred after certain well-adapted, native species were lost through local extinction following intense (pulse) or prolonged (press) disturbances. However, predicting species that are essential to ecosystem functions has generally remained difficult because information is lacking on many species interactions as well as on life history, adaptations to different disturbances, and dispersal abilities among key species. Ecosystem studies can provide a broad perspective regarding species relationships and recycling of essential nutrients. Species' shifts in patterns

of abundance (or local extinctions) following natural and anthropogenic disturbances can sometimes illustrate how certain key species regulate nutrient cycling and other ecosystem functions. Field testing of many concepts related to understanding the importance of key species in ecosystem functioning is just beginning. Results of these long-term ecosystem studies across a range of spatial and temporal scales can provide guidelines for the stewardship of biodiversity.

Ecosystem Boundaries: Inputs, Outputs, and Transformations of Energy

An ecosystem approach can be used to address many different questions spanning scales from the global biosphere to small ponds or patches of habitat. As the questions change, so do the boundaries and the complexities of species interactions within and among different compartments or across trophic levels. Ecosystem boundaries are often defined to include natural species assemblages and to analyze inputs and outputs of energy and materials for cross-site comparisons of efficiencies in energy transfers and studies of changing conditions. These analyses often take the form of mathematical models such as computer simulations or individual-based models of species and their specific functions within the biotic assemblage and environmental conditions under study.

One ecosystem may export nutrients and organic matter (stored energy) to other ecosystems so that cross-site linkages often become important. For example, in studies of nutrient cycling in terrestrial ecosystems, the definition of boundaries would likely have some compartments of organic matter production by living plants and their relationships with herbivores and carnivores. Macro- and micronutrient inputs would likely be derived from the atmosphere through dry deposition of particulates – nitrogen gas being taken up by some nitrogen-fixing species of microbes. Other sources of nutrients, especially phosphorus, would come from weathering and erosion of soil and bedrock deposits, with movement among other compartments derived from actions of wind and water. This combination of interactions illustrates the importance of defining clear boundaries for subsystems within the complete ecosystem so that measurements of movements (fluxes) among compartments can be measured accurately. Generally, ecosystems and their boundaries are abstractions that can only be useful and insightful when combined with sufficient knowledge regarding the natural history and general ecology of communities and their physical environment. There is wide recognition that a combination of direct field observation, experimentation, and modeling is essential when conducting ecosystem studies.

Currently, fundamental questions dealing with relationships between energy flow and the species-specific roles of organisms are attracting increased attention. For example, can results of controlled small-scaled (fine-grained) experimental studies of productivity be used to predict responses of other natural assemblages at larger scales (coarse grained) of ecosystem dynamics? Does energy flow through an ecosystem increase, decrease, or remain the same if one species goes locally extinct but the abundances of other "similar species" change rapidly to compensate for the lost species? Under what environmental conditions do species substitute for one another and compensate functionally for the lost species? These and many other questions are beginning to be answered, but studies related to biodiversity and the persistence of species assemblages remain incomplete. A series of symposia during the 1990s dealt with the relationships between biodiversity and ecosystem functions (**Figure 1**) and stimulated many new ideas (Lawton, 1997; Palmer *et al.*, 1997; Naeem, 1998).

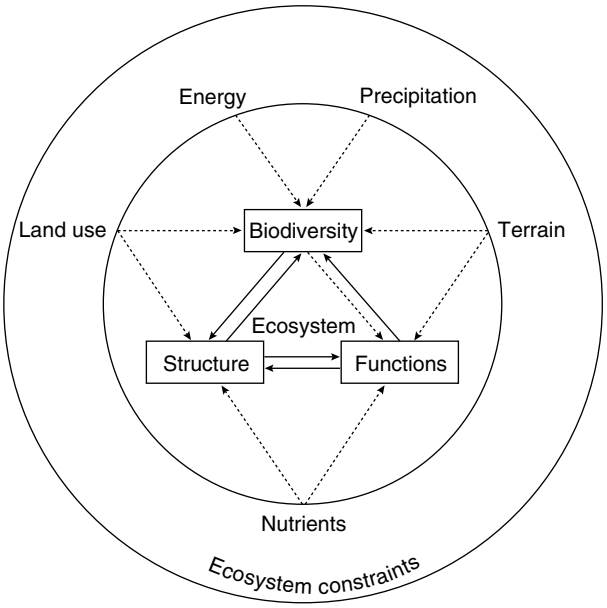


Figure 1 Interactions among various factors constrain how ecosystems function. The climatic controls over inputs of solar energy and precipitation are two "forcing functions" which strongly influence how ecosystems operate. Land-use changes are also important in determining how water and nutrients move through the ecosystem and influence biodiversity. Nutrients moving among biotic and abiotic components are a type of "transfer function" that depends on abiotic and biotic factors (Modified from Schulze and Mooney, 1994).

A Historical Perspective: The Ecosystem Concept

How energy moves from one group of species to another has been an active area of study at least since Charles Darwin and Alfred Russel Wallace first wrote in the 1850s about the interconnections among species. They were intrigued with the general proportions of population abundances that were thought to exist among different groups of large predators and their prey, and they emphasized competition and predation as important factors for regulating species interactions. These early observers lacked a conceptual approach to what was later viewed as ecosystem-level dynamics. Darwin's studies of earthworms and their roles in soil development and his views on the roles of multiple species in the "tangled bank" metaphor stimulated others to consider how these many interactions could be viewed holistically. In 1887, Stephen Alfred Forbes described material cycles within lakes and used a table of predator–prey data to examine which fish species

consumed similar or different prey species in Illinois's rivers and lakes. Forbes drew his conceptual boundaries and predation matrix to coincide with the shoreline and emphasized putting the many pieces together in a type of system homeostasis in his new metaphor of the lake as a microcosm. This holistic view was taken up in a different way by E.A. Birge in 1915 with his work on heat budgets of lakes in Wisconsin. Birge measured inflows and outflows of energy, and in so doing he set the stage for viewing ecosystems in terms of their general physical attributes rather than numerous component parts. This "black box" approach allowed for observations at larger scales without full analysis of the controlling variables within the process under study.

The British ecologist, Charles Elton, recognized that food chains existed in the context of energy flow and nitrogen cycling from his early work with V.S. Summerhayes on Spitzbergen and Bear Islands. The sum of these chains formed a food web which Elton called "food cycles" and included in a diagram combining nutrient cycling and energy flows across landscapes scales. Elton built on earlier concepts of Victor Shelford and introduced the idea of a "pyramid of numbers" and a "pyramid of energy" that reflected general patterns of community organization. More individual plants and plant material (biomass) can usually be harvested in a plot than herbivores or the relatively rare carnivores. (Biomass is the weight of all individuals of a species or trophic level found in a square meter of habitat that can be converted into calories of energy.) These concepts of pyramids of biomass and energy led others to formalize ways to analyze these relationships. However, it was not until 1935 that Sir Arthur George Tansley first defined ecosystems as the "basic units of nature." Since then, ecologists have recognized and debated the importance of ecosystems as objects of study (Golley, 1993; Weathers *et al.*, 2011).

The Trophic Dynamic Concept

A major breakthrough came when Raymond Lindeman and G. Evelyn Hutchinson first proposed the concept of trophic levels (Lindeman, 1942; Hutchinson, 1948). Lindeman emphasized that aggregations of individuals rely on similar sources of energy and that this grouping of species provided a conceptual structure and a means for quantification and testing predictions. Lindeman built his concept on Elton's pyramid of numbers by emphasizing the importance of biomass as well as the linkages of "food chains" and "food cycles" as Elton had done. In addition, Lindeman used concepts from studies of biogeochemistry by V.I. Vernadsky. Vernadsky stressed integrating abiotic and biotic relationships among components of the biosphere, atmosphere, hydrosphere, and geosphere (Hagen, 1992). These new ideas on ecological efficiency of energy transfers within the biosphere were being developed by Lindeman's mentor, G.E. Hutchinson at Yale University, in his long-term studies of Linsley Pond, Connecticut. Hutchinson introduced the use of P-32 in studies of phosphorus cycling with his early work on the biogeochemistry of Lindsey Pond in Connecticut. These isotopic studies made it possible to determine the rates of nutrient uptake and recycling by grazing zooplankton. Lindeman's trophic dynamic concept recognized that groups of species both "co-act" and "re-act" in their relationships with each other and with their abiotic environment.

Lindeman also emphasized that these patterns of productivity change as the biotic and abiotic community develops over time through succession. Based on his own detailed studies at Cedar Bog Lake, Minnesota (and other studies by Chauncey Juday on Lake Mendota, Wisconsin, who was actively studying lake productivity with E.A. Birge), Lindeman proposed that lake productivity increased during succession as nutrients accumulated in the basin during eutrophication. This natural aging process of a lake ecosystem was thought to reach a climax stage and undergo senescence before developing into a bog forest ecosystem. As with the trophic dynamic concept, some of Lindeman's thinking on patterns of change in productivity over time was influenced by earlier studies by G.E. Hutchinson and his colleagues on Linsley Pond, Connecticut. Lindeman (1942) noted that their work suggested that "these generalized changes in the rate of production may be expressed as a sigmoid curve showing a rough resemblance to the growth curve of an organism or of a homogenous population" (p. 409). Such analogies between ontogeny of individuals, populations, communities, and ecosystems were commonly used during that time period. However, as more accurate data on rates of sedimentation, nutrient cycling, and productivity became available as a result of the improved development of isotopic dating, the generality did not hold up and multiple patterns of Lake Eutrophication were later documented. Hutchinson and one of his students, E.S. Deevey, helped develop the use of radioactive isotopes (C-14 and Pb-210) for determining rates of deposition of sediments from cores of lake sediments. This early work led to the field of paleolimnology that has contributed much to the recent studies of climate change, land-use shifts, and the effects of people on ecosystems. Prior to radiocarbon dating techniques it was not possible to be certain about patterns from one region to another being caused by certain events. With the use of C-14 to cross-correlate major changes in the sedimentary records, these "floating" chronologies became strong evidence for tracking rates of changes across regional landscapes.

As a result of studies in many different types and locations of lakes, the patterns of eutrophication varied and reflected land-use changes as well as natural processes of accumulation of carbon and nutrients over time. Some lakes such as man-made reservoirs were observed to begin with relatively high concentrations of dissolved nutrients derived from leaching of inundated soils and microbial breakdown of organic matter in the flooded basin. These high nutrient concentrations often diminished over time and resulted in lower production. This pattern is the opposite of what was observed in changes from low to high production in many natural lakes (from oligotrophy to eutrophy). More recently, the rates of eutrophication have been documented to respond to foodweb dynamics in trophic cascades. Removal or addition of top predators (such as large-mouth bass) altered the prey populations of small species of consumers that, in turn, changed the number of grazers and the rate of algal production. Whole-lake experiments conducted in Wisconsin led to a new understanding of how phosphorus could limit algal growth and how the entire food web could affect rates of primary production.

As a result of these studies, it was recognized that recreational fishing can be managed by harvesting many of the larger fish and thus changing rates of algal growth and removing nutrients in both reservoirs and natural lakes. As discussed later, this biomanipulation links fisheries management and lake management to control of nutrient cycling and plant growth. Rainfall, runoff, and transport of sediments from the watershed (as well as inputs from wastewater treatment plants) can add various amounts of nutrients over time in lakes and wetlands. Consequently, the rate of eutrophication varies from one watershed to another depending on land uses, slopes, and climate.

The concept of trophic dynamics and its focus on transfers of energy between trophic levels was not widely accepted until a new postwar influx of investigators began detailed studies of energetic and nutrient cycling. Several questions have persisted for decades: How can energy flow regulate the number of trophic levels within an ecosystem? How does one group of consumers regulate the numbers of individuals and energy flow in other trophic levels? Which nutrients limit plant growth and how are consumers involved in recycling nutrients? These questions and many others were rapidly taken up by ecologists such as G. Evelyn Hutchinson, Eugene Odum, and Howard Odum and their students in the 1940s and into the 1970s.

After considerable debate and continued developments of the concept, many ecologists today rely heavily on models and field experiments using modifications of Lindeman's approach. Some still have concerns about fundamental issues regarding how trophic levels are defined relative to the complexities of natural food webs. As a result, there are various definitions of what constitutes a trophic level. Some prefer to use a more general term, trophic position, to indicate relative feeding relationships but do not measure transfer efficiencies across trophic levels. Because most descriptive field-based studies of food webs really are studies of subwebs, and therefore are incomplete, a thorough test of predictions regarding food chain length or long-term stability relationships derived from trophic models is usually not feasible except in very simple food webs or in laboratory studies.

Conceptually, the predictions of how energy flow regulates trophic dynamics are relatively straightforward. First, some of the initial energy entering the first trophic level is lost by reflectance from the plants, lost as heat, expended in metabolism and evapotranspiration, or lost because of a less than complete coverage of foliage (leaf area) or algal volume. Thus, some warming occurs by energy absorption by the physical habitat (e.g., soil, rock, or water). Then, from constraints imposed by the second law of thermodynamics, energy is lost at each step in the flow of energy from the first trophic level to successively higher levels within food webs. The use of solar energy or chemical energy by primary producers and the consumption of plants and animals at higher trophic levels are relatively inefficient because some energy is lost to metabolism and as heat at each transfer across trophic levels. The total energy flow through the plant trophic level is termed gross primary productivity (GPP). Once the energetic costs of respiration are subtracted from GPP, the remaining energy is called net primary productivity (NPP). This NPP is usually a very small portion of the available solar energy that entered the first trophic level. NPP is the only amount available for transfer to upper trophic levels. In some ecosystems, there is an "energy subsidy" provided by inputs of organic matter from another ecosystem. For example, leaf litter entering a stream, lake, or open cave can be an essential source of stored energy for use by detritivores in a different ecosystem than the forest ecosystem in which it was produced. In all energy transfers between trophic levels, the assimilation of energy is variable but generally of low efficiency. Typically, efficiencies (output:input ratios) are less than 10%, but higher values are known for some food webs. The consistently low values that were first measured in ecological studies in the 1940s and 1950s led to the hypothesis that the number of trophic levels within any food web was determined primarily by the amount of incoming energy and the efficiency of energy transfers.

Food Webs and Trophic Levels

Empirical studies demonstrate that most food webs contain fewer than four trophic levels. However, the number of trophic levels is not a consistent measure because of the complexity of feeding relationships over time and space, the mobility of consumer species, and the movement of food resources across ecosystem boundaries. Many species vary in how they obtain their energy and how efficient they are at different stages of their life histories and under different conditions. Among consumer species, many rapidly growing juveniles or reproductive adults require high-quality, nutrient-rich foods. These same individuals typically feed on lower protein foods when they become nonreproductive adults. Numerous species are omnivores and feed on plants and animals from different trophic levels. Because of these complexities, there has not been complete agreement on how to operationally define trophic levels.

The transformation of inorganic elements into organic matter requires energy to be converted into biomass by species of algae, green plants, and a few types of bacteria. These micro- and macroautotrophs are often represented by many species. The degree of similarity (niche overlap) in their abilities to produce and to store organic matter is important in predicting the consequences of any losses of species. Many species have evolved into persistent assemblages that store carbon and nutrients such as nitrogen and phosphorus. Energy stored in the form of plant-produced organic matter is later passed on directly to grazing species and then indirectly to predators within food webs.

Efficiency of energy transfer from one trophic level to the next higher level is of fundamental importance in understanding conceptually how different ecosystems function. Measures of efficiency, however, are only a part of the explanation for why some ecosystems have longer food chain lengths than others. Relatively "inefficient" food webs with few trophic levels appear to be adapted to certain types of frequent disturbances. Ecologists realize that a single explanation or mechanism is unlikely to account for all the various complexities that exist in determining how ecosystems are organized in terms of energy flow. However, comparisons among well-studied ecosystems and their numbers of trophic levels (food chain length) can provide a useful basis for predicting vulnerability of food webs to major disturbances and movement of toxins such as mercury and other heavy metals.

Analysis of guilds and that of functional groups are different approaches used to study shifts in feeding behavior and to complement trophic-level analysis. Guilds are defined as assemblages of species or individual age classes that share a common source of energy at any given time (e.g., nectar-feeding birds and bees and insect-feeding birds and spiders). Functional groups are defined as being similar in their mode of feeding (e.g., filter feeders, shredders, and pursuit predators), but individuals may use a variety of different sources as resource availability shifts. Thus, individuals of a single species may be distributed over several trophic levels and belong to different guilds and functional groups during each individual's life span and reproductive period. As nutritional requirements change and the availabilities of different types of food resources also change, individuals can often adapt to find different available sources of energy. Analyses of similarities in these adaptations and the degree of overlapping functionality are used to understand the degree to which producer and consumer species are interdependent. Numerous complex linkages (e.g., herbivory, predation, decomposition, parasitism, and mutualism) imply that few species are likely to be complete substitutes for other species. As discussed later, some species may interact positively, negatively, or neutrally in association with other species. Such complex relationships among species, especially under changing environmental conditions, complicate field experiments and make predictive models difficult to test fully.

Controls of Energy Flow in Food Webs

A small increase in species richness can have a large effect on how energy flows through food webs. The main features are the number of linkages among species and, especially, the type and strength of those linkages (Paine, 1969). For example, a simple trophic structure would be a linear series of three species (A–C) in a food chain with one species in each trophic level (**Figure 2**). Thus, a single species of plant is consumed by a single species of herbivore, which is consumed by a single predator species. Although analysis is relatively definitive in these types of communities with few species, this simple food chain structure may preclude consideration of some questions of general concern, such as resiliency of the assemblage following a disturbance and species loss.

If one more plant species is added (D) to a simple community, then this slightly more species-rich food web provides some important additional dynamics in terms of alternative pathways for energy to flow. Moreover, the herbivore (B) can switch from one food resource to another and this additional complexity increases generality and realism incrementally. With two herbivore species (B and C) the food web is more complex and the predator (A) has a choice of prey resources. Even with the same number of species, much more realism is added by considering the predator (A) to be an omnivore, and even more is added if the predator and one of the herbivores (C) are also cannibalistic. These simple diagrams show how quickly the types and numbers of linkages (connectance) within food webs can change the dynamics of energy flow even with only a few species. This last example is typical of some low-diversity stream food webs on isolated tropical islands currently under study and discussed later.

Ecosystem Analysis

General rules regarding the relationships between energy flow and the control of food web complexity are currently incomplete. In some habitats a complex relationship apparently does exist among the total annual amount (and seasonal distribution) of energy, the nutrient inputs to ecosystems, and the number of different species in a habitat. In other habitats there is no evidence for a cause-and-effect relationship among the rate of energy flow, species growth and productivity, and the number of species in an ecosystem. Other likely variables include evolutionary time and biogeographical distributions as well as the frequency and intensity of disturbances. The particular species composition of a food web may also alter productivity. Empirical evidence for predicting the importance of species-specific relationships is increasing, but methods for establishing which species regulate ecosystem functions remain controversial. Currently, only a few studies have focused on the species-specific roles to determine which species have unique roles and how these roles shift as environmental conditions change.

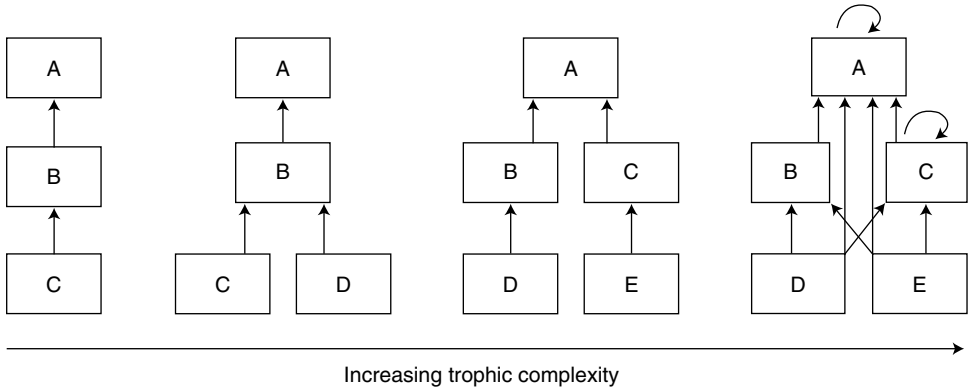


Figure 2 Energy flow through simple food chains and webs. Small increases in species number lead to high trophic complexity as connectance (arrows) and cannibalism (looped arrows) increase. The number and strength of connections among species are more important in regulating the flow of energy through food webs than simply the number of different species in food webs.

Although recent experiments have examined some fundamental relationships, we do not have a full understanding of the effects of varied energy inputs on the richness of consumer species that naturally coexist in ecosystems. For example, in most tropical forests and coral reefs there is generally a high number of species, but the cause and effect of this species diversity are open to different interpretations other than the potential importance of relatively high and continuous inputs of energy.

At the earth's surface, the annual, seasonal, elevational, and latitudinal distribution of solar energy provides varied inputs of energy to deserts, grasslands, forests, wetlands, lakes, rivers, and oceans. Generally, energy always flows through ecosystems but does so at different rates under different global geographical locations and local conditions of slope and aspect. Depending on the latitude, ecosystems generally receive a seasonally pulsed or a continuous annual supply of solar energy for primary producers. The rate of energy flow and associated biological productivity are dependent not only on the availability of energy but also on water, on combinations of different macronutrients (nitrogen and phosphorus) and micronutrients (trace elements such as iron, manganese, and silica), and on the presence of an assemblage of interactive plant and animal species. Some natural food webs in extreme environments, such as hot springs, saline lakes, caves, or certain deep-sea thermal vents, have relatively simple linear food chains and have species adapted for specific habitats. For example, dark, closed caves that are deep underground only receive indirect sources of detrital energy from sunlit surfaces aboveground and are typically characterized by a relatively low number of endemic species not found on the surface. Deep-sea thermal vents in the oceans rely solely on chemical energy derived from microbial breakdown of gases such as hydrogen sulfide and are characterized by sulfur bacteria and unique consumer species. These simple ecosystems continue to provide an opportunity to test some fundamental concepts regarding food webs and energy flow relationships.

Multiple Energy Pathways

It has been evident since Lindeman's work that energy travels along different pathways and includes microbial species and macro-species in various interconnected relationships. Experimental assemblages are now being widely used to provide some insights into which mechanisms control ecosystem dynamics. There is evidence for biotic control mechanisms (interspecific competition for resources, predation, parasitism, and mutualism), abiotic controls (nutrient limitation and frequent and/or intense disturbances), and combinations of controls in different ecosystems. These studies have also provided important insights regarding two main energy pathways. The distinction between direct, solar-driven, photosynthetically based food webs and indirect transfers of stored energy in the form of detritus (that can be wind driven or washed into habitats) has sorted energy flows into two main classes. The earliest work on ecosystems recognized this bimodal classification and it remains an important organizing framework in linking aspects of terrestrial and aquatic ecology.

The importance of organic detritus as a means for storage of energy was recognized by studies of Jerry S. Olson in the 1960s at Oak Ridge National Laboratory on the carbon cycle. Analysis of time lags requires an understanding of how rapidly organic detritus accumulates and then breaks down to recycle carbon, nitrogen, phosphorus, and other materials. These insights are critical in current discussions regarding carbon dioxide accumulation in the atmosphere as fossil fuels are burned (i.e., coal, oil, and natural gas taken from storage that accumulated over geological timescales and are now being rapidly cycled back into the atmosphere following combustion). Debates regarding global warming, the greenhouse effect, and where the "missing" carbon is in the present-day ecosystem require a thorough understanding of the entire biosphere and the carbon cycle as it relates to other nutrient cycles.

Studies of deep-sea vents and hot mineral springs illustrate another distinct class of ecosystems. This food web is not solar driven but depends on chemical energy sources used by chemosynthetic microbes. Geologic sources of hydrogen sulfide and other gases provide examples of chemical energy pathways that may well have been the first modes of ecosystem formation by the earliest microbial species on Earth before the evolution of photosynthetic species. Various lines of evidence, such as the banded iron formations in pre-Cambrian rock strata, indicate that the earliest atmosphere lacked oxygen, suggesting that chemoautotrophs dominated the first phases of evolution. Once oxygen-producing photoautotrophs evolved and dominated the oceans and lakes (and later developed terrestrial forms of green plants), their high levels of primary productivity resulted in an accumulation of oxygen in the atmosphere and a decrease in carbon dioxide (possibly through carbon uptake and storage by plants and deposition of sedimentary limestones). This early shift into a photoautotrophically based ecosystem apparently put the chemoautotrophs at a competitive disadvantage in an oxygen-rich environment. These remnants of the earliest species of microbes coexist with specialized species of tube worms, shrimps, and crabs that now dominate deep-sea thermal (DDT) vents where these chemoautotrophs use inorganic chemical compounds as energy sources.

Lake Ecosystems

To illustrate the flow of energy through ecosystems it is useful to consider some examples derived from lake studies. These convenient habitats have been used for comparative ecosystem studies because distinct boundaries provide clear definitions of inputs and outputs (**Figure 3**). The main boundaries include any inflowing and outflowing rivers as well as the lake surface-atmosphere and the sediment–water interfaces and also shorelines and topographic ridges (that delimit drainage basins). Water temperatures, nutrient inflows and outflows, and mixing and transport processes all influence species distributions and abundances in generally predictable ways.

Solar energy transformed into organic matter through the process of photosynthesis is the main source of energy for most ecosystems, especially in large lakes. Different sizes and types of plants in lakes vary greatly in their rates of productivity. For example, in shallow-water ecosystems solar energy can be used by microphotoautotrophs (attached algae or suspended phytoplankton) and macrophotoautotrophs (pond weeds such as cattails and water lilies). The ratio of the biomass of organisms

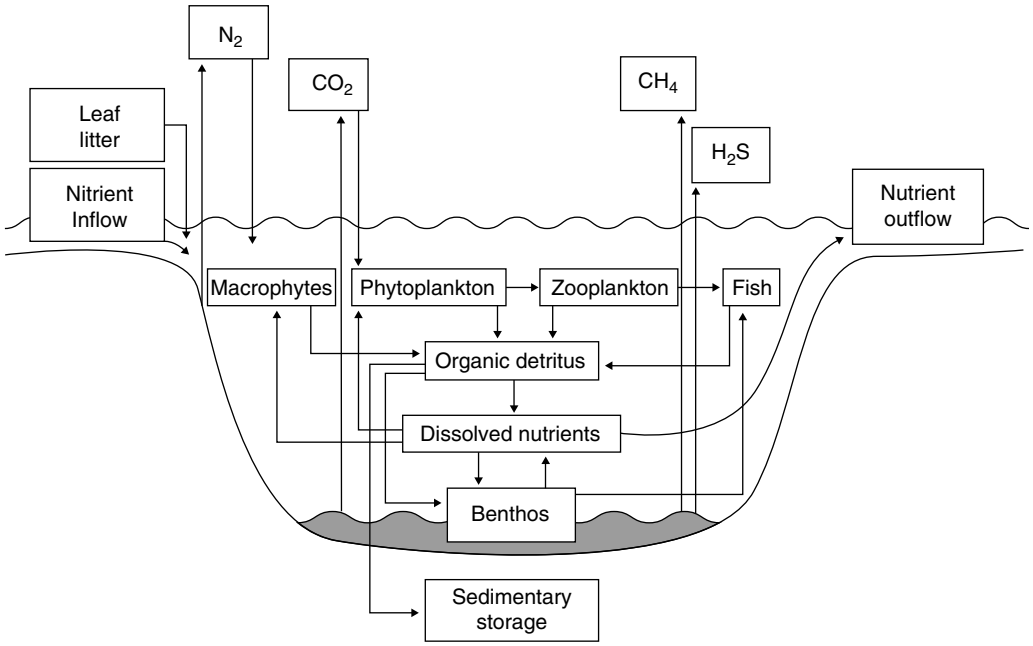


Figure 3 A nutrient-rich lake ecosystem and the flow of energy and materials through major compartments. The inputs of dissolved nutrients and particulates (leaf litter and suspended organics) move through the lake ecosystem at a rate determined primarily by the amount of solar energy entering from the surface of the lake, the stored energy in the form of organic matter inputs (for bacteria, fungi, and detritivores), and through-flow of water. Productive lakes can be either sources or sinks of nutrients relative to downstream river and lake ecosystems, depending on their uptake and storage of energy and nutrients. Shallow lakes have a higher surface to volume ratio than deep lakes and are usually more productive because a larger proportion of their volume receives solar energy inputs. Winds mix nutrients from bottom waters and often circulate limited concentrations of dissolved phosphorus that flux from the deeper sediments. Energy transformations from solar input to green plants to herbivores and carnivores are essential for nutrients to move through the food web and be recycled (reproduced from Covich MA, Palmer, and Crowl TA (1999) The role of benthic invertebrate species in freshwater ecosystems. *BioScience* 49: 119, with permission from American Institute of Biological Sciences).

relative to their rate of production is termed "biomass turnover time." Biomass is generally measured by multiplying the number of individuals in a population by each individual's weight. Turnover time is related to how rapidly organisms increase in biomass during their life span. Rapid turnover is associated with high rates of productivity by small, fast-growing individuals of plants and animals. Small species with rapid turnover usually exploit resources at relatively fine spatial scales.

Importance of Depth and Spatial Heterogeneity

Different wavelengths of light energy penetrate into waters of different depths. Shade-tolerant species of plants living in deep water (or in deep shade of the canopy trees in terrestrial ecosystems) have distinct physiological adaptations that allow them to obtain sufficient energy to grow and reproduce even at low light intensities. The range of wavelengths that is used by plants is light visible to the human eye. This range of visible light (between approximately 400 and 700 nm) is termed photosynthetically active radiation. Short wavelength (ultraviolet light) and long wavelength (infrared light) are not used in photosynthesis but are important in regulating floating and emergent aquatic plants (or terrestrial plants in other ecosystems) because ultraviolet radiation degrades organic molecules and infrared increases leaf temperatures. Daytime warming from the sun is essential for some species to survive, especially insects and cold-blooded vertebrates (poikilotherms). Other transfers of solar energy are essential to various Lake Ecosystem processes (e.g., sediment and nutrient transport by wind and water). Work done by solar-generated winds mixes the lake and thereby influences nutrient cycling. Similarly, wind-driven currents disperse planktonic larvae and aerate deeper waters. Winds also increase salinity through evaporation.

Consumer species are limited by the availability of plant-derived organic materials (or microbially derived organic materials from chemoautotrophs). The well-mixed, brightly lit, open water (pelagic zone) is dominated by small species of zooplankton that are well adapted to feed on suspended phytoplankton. Microbial breakdown of dead organic matter (detritus or seston) and grazing on algae by numerous invertebrate and vertebrate species function to recycle nutrients in the open waters. In other distinct habitats such as the shallow littoral zone near the shoreline, a vegetated zone is dominated by invertebrate species that feed on larger, rooted plants that have a slower turnover rate compared to the small, suspended algae. Some species of fish live in the littoral zone while young and move in and out of the pelagic open waters as they grow larger. As these predators increase in size, they switch from feeding on invertebrates and small fishes to consuming larger fish and crayfish prey.

Internal and External Nutrient Cycling

Ecosystems are generally "open" with regard to external additions of nutrients from the atmosphere and from catchment

basins. Deserts, grasslands, forests, and wetlands as well as most lakes and rivers continue to receive new supplies of nutrients from precipitation (rain and snow), from dust and eroded soils that are carried to the ecosystem by wind and water, and from human-derived fertilizers, sewage effluent, air pollutants, and other agricultural and industrial wastes. Decomposers recycle essential elements for continued productivity by other species. As a result of rapid decomposition, some ecosystems can be relatively "closed" in that nutrients may be rapidly recycled internally so that these nutrients remain within the ecosystem's boundaries and are not transported or lost to other ecosystems. Nutrients can recycle internally from temporary storage in living tissue (biomass), detritus, or storage in soils and sediments.

Different species of plants and consumers within the food web and specific environmental conditions (such as warm temperatures, low concentrations of nutrients, and high pH) influence rates of internal cycling. Land-use practices also have a large effect on vegetation and on how nutrients recycle within a catchment or move into streams and lakes. For example, by measuring the concentrations of nutrients in a stream draining forested catchments, a team of ecologists (led by Gene Likens and Herbert Bormann) identified different pathways for nutrients such as gaseous nitrogen and erosional phosphorus in Hubbard Brook, New Hampshire, from a series of long-term studies on the forest and its components (streams, lakes, and catchments). The research at Hubbard Brook demonstrated the importance of large-scale experiments (such as removal of forest vegetation) to determine how various components are related. The increased outputs of some major nutrients (nitrogen, phosphorus, and calcium) following the experimental manipulation of the forest cover provided information about the role of trees in taking up and storing water and nutrients. By tracking the movement of phosphorus out of the forest and into the stream, Judy Meyer concluded that pulses of stream flow were the primary mechanism governing movement of this relatively scarce element. Phosphorus can be stored by mosses and other plants in their biomass. Later studies led others to recognize the generality of Meyer's observations in 1979 and to identify the important role that aquatic mosses play in taking up and storing phosphorus in stream ecosystems as well as providing microhabitats for many types of benthic invertebrates. These studies at Hubbard Brook and elsewhere later became important in understanding how calcium biogeochemistry and acid deposition altered the forests, soils, and stream chemistry. Comparative watershed studies were conducted by researchers at several sites, such as the Coweeta Hydrologic Laboratory in North Carolina, the H.J. Andrews Experimental Forest in Oregon, and the Kuparuk River at the Toolik Lake Long Term Ecological Research site in the Alaskan tundra to track changes in mineral cycling as food webs responded to experimental manipulations of nutrient inputs.

Loops, Spirals, and Chains

Different species can be complementary in their roles as they interact in obtaining their energy. For example, benthic species transfer nutrients deposited in the sediments and stored in organic detritus back into overlying waters and thus help move the stored energy up food chains and to other associated food webs. Burrowing and mixing of sediments by benthic species enhance productivity of freshwater ecosystems and lead to interconnected food webs. The rate of movement from one form to another can be facilitated by how species use resources in different ways. Size differences in the resources and the species can be extremely important.

Microbial species play especially effective roles in internal nutrient cycling by using organic detritus as an energy source. Their small size and fast turnover (biomass: productivity ratios) of carbon and nitrogen make them highly important. Bacteria and fungi rapidly break down dead organic matter before it accumulates within terrestrial soil and aquatic sediment (benthic) ecosystems. In many soil, stream, and estuarine ecosystems the amount of energy and nutrients cycled by microbes is relatively larger than that of the photosynthetic pathway. The soil and sediment biota have similar groups of specialized species that shred organic detritus (from leaves, fine roots, dead algae, and other small invertebrate consumers). Fine particulate organic matter (FPOM) is composed of small fragments of detritus and aggregates of dissolved organic matter (DOM) that form from breakdown products and from cell exudates and leachates in aquatic ecosystems (such as groundwaters, streams, lakes, and estuaries). The release of DOM and inorganic nutrients is accelerated by some species of microbes and is a source of energy and nutrients for other microbial species, such as bacteria, fungi, and ciliated protozoans. DOM is also a food resource for larger invertebrate consumers and provides a parallel set of pathways for energy flow.

Invertebrate detritivores break down dead organic materials to obtain their energy. Many detritivores are dependent on microbes to condition the detritus before it can be consumed by invertebrates or vertebrates. Nutrient cycling allows for continued energy flow through the ecosystem. The continued input and flow of energy through the food web is likewise required for nutrient uptake and recycling. Numerous species of bacteria and fungi form a "microbial loop" that provides for high rates of energy flow by breaking down organic matter. Excretion of wastes by consumers provides one source of nutrients for species in the microbial loop. The microbial loop is especially important in marine and freshwater pelagic (open-water) ecosystems. For example, a crater lake (e.g., Crater Lake, Oregon) with a limited surface area and positioned in a volcanic depression usually has a very small drainage basin compared to large lakes with extensive drainage runoff. A lake with a small surface area will also receive very little atmospheric inputs of nutrients. Primary production remains low and mostly dependent on rapid, internal recycling of nutrients from fine, suspended detritus (dead plankton) as a result of breakdown by microbial species. Nutrients may only slowly accumulate in the sediments of the deep Crater Lake and not be mixed by currents back into the upper layers of well-lit waters (photic zone). Whenever nutrients are low (as in midsummer in some more productive lakes following the spring bloom of algae) the role of microbes and invertebrate species in cycling nutrients is important to maintaining energy flow in the ecosystem.

Similarly, nutrient concentrations are relatively low in the open ocean, far away from continental sources of nutrient

runoff and from coastal upwellings of deeper, nutrient-rich waters. Microbes attach to dead plankton and use this organic detritus as a source of energy. In the process of breaking down this fine detritus, the microbes release nutrients for further growth by phytoplankton (floating algae). Nutrients are also released in the photic zone by leakages from algal cells and by excretion of herbivores. The small size of the detritus particles slows their rate of sinking through the well-lighted photic zone and allows the microbes to release nutrients where phytoplanktonic photosynthesis is not light limited. Larger pieces of detritus fall more rapidly and accumulate in deeper, darker waters where short-term seasonal storage occurs. The nutrients in these deep, dark waters are not available for continued photosynthesis until this entire layer of nutrient-rich water is mixed vertically by wind- and gravity-driven currents back into the photic zone and taken out of temporary, seasonal storage. In very deep lakes, the wind energy is insufficient to mix the entire volume and nutrients accumulate over many years in these deep waters.

There are also some important horizontal linkages in that nutrients and organisms (especially differently sized fishes and wind-driven currents carrying drifting zooplankton) move from the shallow, nearshore waters (littoral zones) of lakes to the open waters (pelagic zones). In these surface waters, nutrients are also recycled by the grazing zooplankton (feeding on phytoplanktonic algae) and by consumption of zooplankton and phytoplankton by fishes. These nutrients are returned to the well-illuminated surface waters, where they are again available for continued algal growth.

This series of vertical and horizontal transformations of organic detritus and primary production of algae in the pelagic zones has some similar analogs in streams, in which the dynamics occur mostly horizontally, but to some extent vertically, along the network of stream channels. The combined vertical and horizontal currents in stream channels form a spiral in the downstream flow of water that carries nutrients and organisms to various distances. Dissolved nutrients move in and out of solution as they are briefly taken out of solution by adsorption on sediments and by active uptake by microbes and attached algae (growing along the bottom of the channel), and are later consumed and released back into the water by cell leakage, excretion from grazers, and predatory fishes. This spiral pattern of nutrients being transported downstream while moving in and out of sediments and organisms and then taken up again by other organisms or adsorbed onto sediment particles farther downstream was first examined by ecologists at the Oak Ridge National Laboratory using isotopes to trace pathways and is now widely studied, for example, with stable isotopes of nitrogen (N15). The distance that a particular nutrient element such as nitrogen moves downstream is termed the "spiral length." If the nutrient is taken from solution and tightly held in different species' biomass, then the spiral length is relatively short and the role of microbes and other organisms along the channel is relatively important compared to that for streams where biotic interactions are less significant and during other periods when the spiral length is longer.

Use of functional groups such as shredders and filter feeders (or scrapers, burrowers, and predators) allows for analysis of groups of different species in terms of particular attributes (e.g., how each species processes different sizes of detritus). In both vertically and horizontally structured interactions, the processing rate of organic detritus by microbes is accelerated by the presence of those invertebrate species that convert large organic fragments into finer fragments. If one species (shredders) uses larger sizes of suspended particles and breaks them down into smaller particles as a result of its feeding, then the downstream supply of FPOM is increased for use by other species (filter feeders). However, if the species that shreds detritus is lost, the filter-feeding species may not function effectively. Different species may form "processing chains" that require particular combinations of linked-species to complete certain ecosystem processes efficiently. These processing chains are one way that increased species diversity can increase efficiency of processing detrital resources, especially if detritus is available at low levels of abundance or if the detritus quality is low. Similar linkages are known to occur among burrowing organisms in soils and sediments where strong interactions are generally important in understanding how different ecosystems function.

Biodiversity Effects on Energy Flow

Productivity and related ecosystem processes such as nutrient cycling and decomposition are generally known to be influenced by particular species. Species attributes, including length of life span, rates of dispersal, and tolerance of frequent and intense disturbances under various environmental conditions, have important effects on ecosystem processes. Key processes (such as nitrogen fixation) performed by only a limited number of species or particular modes of feeding characteristic of only a few species are recognized as important characteristics associated with biodiversity. The importance of a single species in biogeochemical cycling was demonstrated by Peter Vitousek and his students in studies of the invasion of some habitats in Hawaii by the nitrogen-fixing plant *Myricafaya* and by determining the long-term consequences for nitrogen cycling and impacts on the ecosystem. These landscape changes in the biomass production and species composition are now being studied over large scales by use of remote sensing technology (MODIS satellites and airborne sensors such as LiDAR).

David Tilman and others conducted several experiments to test the hypothesis that primary productivity in prairie grasslands is related to the number of species of plants grown in the same plot. These experiments and others have the advantage of considerable replication and careful controls but were conducted at relatively small scales. Although the relationship between species richness and energy flow in ecosystems has attracted considerable study, the scale of these experiments and the composition of the species assemblages studied have resulted in different interpretations of general relationships. Results of field tests suggest that the number of species per se is apparently not as important as the particular attributes of different species that relate to their efficiency of nutrient uptake and retention, as well as growth. The larger the number of species included in a study, the more likely some species will be included that are well adapted for the conditions and function effectively. Thus, although ecologists have conducted

field tests to determine how different species alter ecosystem processes, there is no complete consensus on how energy flow by itself influences species richness or vice versa. Part of this lack of consensus is a result of using different scales and methods in field studies of many different types of ecosystems (e.g., boreal and tropical forests, grasslands, deserts, deep and shallow lakes, and large and small rivers). Theoretical and conceptual developments are being actively developed and are stimulating additional field testing of these relationships between biodiversity and productivity.

Top-Down and Bottom-Up Regulation of Energy Flow

In the 1960s, some of the first experiments designed to identify the importance of different species and key attributes included the removal of predators. Because top predators are relatively few in numbers, their removal can be operationally feasible over some relatively large areas. Robert Paine removed predatory starfish from the intertidal zone and observed a shift in the relative and absolute abundances of some of the molluscan prey species. His results demonstrated that competing species can be held in check by selective predation, especially if the predators consume more individuals of the more numerous prey species. Work by James Estes in the 1970s on the effects of declines in abundances of sea otters and increased abundances of sea urchins and other prey (as well as the shifts of algal regrowth in kelp beds) along the west coast of North America also demonstrated that predators had large effects on entire food webs. These "keystone species" were viewed as important regulators of energy flow in natural food webs in that they had a disproportionate effect despite their relatively small numbers or biomass. The declines in top, apex predators (overharvest, habitat loss, or disease) can lead to a cascade of effects throughout the food web.

Recognizing the importance of keystone species led some ecologists to use this idea for managing certain ecosystems. Researchers introduced the concept of bio-manipulation to study different combinations and abundances of algal species and grazing species in order to regulate nutrient cycling in aquatic ecosystems. In a series of lake studies by several ecologists (especially Joseph Shapiro, J. Hrbacek, John Brooks, and Stanley Dodson), they controlled the presence or absence of top fish predators. From these studies it became apparent that consumers in upper trophic levels could regulate populations of prey, and this regulation in turn had consequences for lower trophic levels as well as nutrient cycling. Manipulating consumer species as a means of removing algae or altering the uptake and storage of nutrients became an area of active study in order to manage lakes and improve water quality. These types of food web studies were also recognized as a means to monitor and to understand movements of toxic compounds such as DDT and mercury in lake ecosystems. Even very low concentrations of toxins in the open water can build up in sediments and bioaccumulate over time in upper trophic levels.

Studies of introduced predators by Robert Paine and Tom Zaret further emphasized that nonnative predatory species disrupted food webs. However, some ecologists argued that native predators rarely altered prey populations or shifted species composition. Steve Carpenter and James Kitchell demonstrated experimentally that these species changes did occur in open-water food webs of temperate lakes. Certain fishes could selectively remove large-sized zooplankton grazers (as Brooks and Dodson had earlier demonstrated), resulting in increased biomass of phytoplankton and decreased levels of dissolved nutrients in a trophic cascade. In other lakes, additions of certain predatory fish species had major effects on phytoplankton growth and nutrient cycling through the predators' effects on grazers. Carpenter and Kitchell performed a series of whole-lake experiments to show that fish predation altered zooplankton species composition and size spectra. Shifts in sizes and types of zooplankton herbivores (and other invertebrate predatory species) in turn altered the phytoplankton community and its productivity. These trophic cascades have mostly been observed in aquatic ecosystems (Carpenter *et al.*, 2010). Recent studies, however, demonstrate top-down effects through four trophic levels in tropical forests and coastal marine ecosystems.

Many studies have demonstrated that bottom-up control by nutrients also influences the rate of energy flow in food webs. For many years, additions of nutrients were known to increase phytoplankton biomass and to alter species compositions of algae. Zooplankton abundances and species richness, in turn, often decline when the quality of their algal food resources is reduced. Thus, the distinct pathways of energy flow are altered by the amounts and proportions of nutrients available to algae and bacteria. High phosphorus concentrations lead to phytoplankton communities dominated by blue–green algae (cyanobacteria) because some species of blue-greens can fix nitrogen (incorporate it into their cells) from nitrogen gas in the atmosphere. Their capacity to use high amounts of both phosphorus and nitrogen allows them to outcompete the species of green algae and diatoms that cannot use atmospheric sources of nitrogen. Blue-greens can also take up and store nutrients internally beyond their metabolic needs for nutrients and deprive other species of recycled nutrients. Many species of blue-greens have distinct features (slime sheaths, toxins, and large sizes of colonies) that make them relatively unpalatable to many grazers. Some species of blue-greens produce toxic substances. By avoiding grazers, blue–green algal species can grow rapidly in the upper waters of lakes and ponds, thereby shading out the competing species. Blue–green algae often form long, chain-like colonies of cells that are less readily filtered by certain species of large zooplankton. Thus, zooplankton production often declines in waters dominated by blue-greens, and with fewer large zooplankton available as prey the production by fish predators also declines.

In many ecosystems a combination of top-down and bottom-up control can be expected. The roles of specific species in determining how nutrients are used and how rapidly energy flows through different food webs are being studied. The generality of these complex relationships and the predictability of patterns of food web responses to additions of nutrients or the additions or deletions of predators are not completely established. There are some traits that appear to increase the likelihood of certain patterns, and these are under intensive study. For example, as plants and animals die and sink into deep, density-stratified lakes the breakdown of organic matter releases nutrients that can accumulate and

dissolved oxygen can be depleted. Deoxygenation sets up chemical conditions in the sediments that release previously bound nutrients that are added to bottom waters. During mid- or late summer nutrients will be scarce in uppermost brightly lit waters (photic zone) in temperature-density stratified lakes because essential nutrients have become increasingly concentrated (during the spring and early summer growing season) in the lowest layers of water below the photic zone. With these accumulated nutrients stored seasonally in deep waters in which light is limited and nutrients cannot be taken up by plants, the roles of grazers and predators are less likely to control nutrient dynamics. Biomanipulation of consumer species may be more effective in shallow, frequently wind-mixed lakes and ponds.

Stressful Environments as Testing Grounds: Ecosystem Services and Biodiversity

In some ecosystems the number of producer and consumer species is relatively low because they are adapted to severe environments. Similar situations can occur on isolated islands or in deep caves or hot springs. Only a few species can tolerate the physiological stresses associated with very high or low temperatures or highly variable or extreme salinities, acidities, or nutrient concentrations. Such sites are useful for conducting field experiments because the low number of species can be studied in detail and manipulated over relatively brief periods.

Because of their distance from mainland sources of colonizing species, headwater streams on tropical islands contain a few abundant species of freshwater decapods (shrimp and crabs) and only a few species of other detritivores. Recent studies in streams on Caribbean islands illustrate how species differ in their effectiveness as leaf shredders. Furthermore, these studies show that different species form processing chains and interact to transform suspended organic detritus into benthic biomass that is retained within headwater food webs rather than being washed downstream or accumulating in deep pools.

Whole-pool experiments in the Luquillo Experimental Forest on the island of Puerto Rico were used to study rates of leaf litter processing. Leaves of *Cecropia schrevenriana* were placed into pools that were cleared of other detritus. To start the experiments, either *Xiphocaris elongata* or *Atya lanipes* was placed into the treatment pools in which other detritivorous shrimp had been removed. Controls were pools with *Cecropia* but with no shrimp. Downstream concentrations of suspended fine, medium, and coarse particulate organic matter, dissolved organic carbon, total dissolved nitrogen, nitrate, and sulfate were measured during 23 days.

Results of the experiments illustrate how differently these two species function as detritivores. Both species of shrimp accelerated leaf breakdown rates relative to controls where only microbial decomposition occurred. *Xiphocaris* shredded *Cecropia* leaves much faster than *Atya*. *Xiphocaris* rapidly shredded the large, intact *Cecropia* leaves and converted them into fine suspended particulates. *Xiphocaris* increased the rate of downstream export of particulate organic matter and concentrations of total dissolved nitrogen and dissolved organic carbon relative to controls. These differences in processing rates also affected downstream distributions of suspended particulate organics and nutrients. *Atya* increased rates of leaf breakdown less than did *Xiphocaris* and apparently filtered out fine organic particulates resulting in less downstream export. *Atya* are especially well adapted to shred conditioned leaves and to scrape microbes from leaf surfaces at low flows. *Atya* are also well adapted with highly modified cheliped fans to filter suspended organic particulates from the water when flow rates are higher than 20 cm s^{-1}. Although both of these shrimp are detritivores, they are not complete substitutes for each other. Their co-location and relative abundances affect rates of detrital processing when they form detrital processing chains.

Why Species Matter

Recent studies have demonstrated that some ecosystem processes apparently do change as the number of species increases. One of the most cited examples is the relationship between increased numbers of species in mixtures of annual plants that are grown in experimental plots. The total area of the plot covered by plant growth (an indirect measure of primary productivity) increased as the number of species increased from 1 to 24. The mechanisms for this relationship are not clear, but progress is being made in interpreting the effects of species composition on these and similar replicated experiments dealing with primary productivity.

Linking species to different ecosystem processes (such as productivity and decomposition) highlights the importance of how organisms interact. In some cases, these interactions may facilitate how rapidly processes occur. For example, decomposition of organic matter prevents buildup of organic detritus, which could lead to increased or reduced growth of macrophytes depending on the amounts and types of litter. In other cases the role of different benthic species in breaking down litter can prevent deoxygenation of lakes and streams and maintain supplies of clean water.

Future Studies

Given the rapid and accelerated loss of species and the irreversibility of global extinction, it is imperative that experimental and conceptual studies do more to examine multiple levels of ecological organization, from populations and communities to ecosystems and landscapes, before more species are lost. Unfortunately, we lack sufficient information about how the loss of different species can disrupt natural ecosystem services. These losses represent a type of "warning light" that should draw more attention to analyzing the consequences of losing native species and introducing nonnative species that may disrupt ecosystem dynamics. The increased dispersal of non-native, invasive species increasingly threatens the stability and sustainability of ecosystems. New technology for monitoring these changes in species distributions and their consequences are developing as part of the National Ecological Observatory Network (NEON). Airborne sensors are beginning to monitor the rates of changes in vegetation across landscapes in response to climate changes, outbreaks of pathogens, and expansions of invasive species.

Improved techniques using stable isotopes, enhanced computer models, and many other developments have expanded the means of addressing ecosystem questions. However, only recently have ecologists evaluated the importance of species-specific attributes in ecosystem processes. Part of the reason for this slowness is that complete food webs are complex and dynamic. Furthermore, the specific functions of a single species are difficult to isolate from the functions of other species in most natural communities. Probably the most important factor has been the development of different scientific perspectives by population ecologists, community ecologists, and ecosystem ecologists. The recent emphasis on integration across subdisciplines has created a new perspective on the importance of different species' roles in performing different ecosystems "services" or essential functions for society, such as nutrient cycling and productivity.

Appendix

List of Courses

University of Georgia

1. ANTH 3090 – Evolution of Human Ecosystems
2. ANTH 8420 – Human Ecosystem Evolution
3. CRSS(HORT) 4400/6400 – Agro-Ecology
4. ECOL 2100 - Global Climate Change: Past, Present, and Future
5. ECOL 3500 Introduction to Ecology
6. ECOL 3520 – Ecological Applications
7. ECOL 4060–6060 Ecosystem Ecology
8. ECOL 4160 – Ecology of North America
9. ECOL(FISH)(WASR) 4310/6310–4310L/6310L – Limnology
10. ECOL(ANTH) 8110 – Tropical Ecological and Cultural Systems
11. ECOL(WILD) 8322 – Concepts and Approaches in Ecosystem Ecology
12. ECOL 8500 – Theoretical Ecology
13. ECOL 8580–8580L – Theory of Systems Ecology
14. ECOL 8600– Nuclear Tracers in Ecology
15. ECOL(CRSS) 8660–8660L – Soil Biology and Ecology
16. EHSC(ECOL)(FISH)(WASR) 8610 – Aquatic Toxicology
17. ENVE 3220 – Energy Analysis II
18. ENVE 4230/6230 – Ecosystem Energetics
19. ENVE 8110 – Ecological Energetics
20. FANR 2888E – Forest Ecosystem Services
21. FANR 3200 – Ecology of Natural Resources
22. FANR 4460/6460–4460L/6460L – Modeling the Effects of Climate Change on Animal Communities
23. FISH(EHSC) 4600/6600 – Ecotoxicology
24. FISH(EHSC)(ECOL)(ENTO)(VPHY)(PHRM) 8350 – Fundamentals of Ecotoxicology
25. FORS 5610/7610 – Prescribed Fire in the Forest Ecosystem
26. FORS 8100 – Advanced Forest Ecology
27. GEOG 3110 – Climatology
28. GEOG 4810/6810 – Conservation Ecology and Resource Management
29. LAND 4360 – Applied Landscape Ecology
30. LAND 6350 – Ecological Landscape Restoration
31. MARS 1020–1020L – Biology of the Marine Environment
32. MARS(MIBO) 4620/6620–4620L/6620L – Microbial Ecology
33. PBIO(ECOL) 4750/6750 – Tropical Ecology and Conservation
34. PBIO(FORS)(ECOL) 8770 – Communities and Ecosystems
35. PBIO(CRSS)(FORS)(ECOL) 8850–8850L – Terrestrial Biogeochemical Cycling
36. WASR 8300 – System Identification for the Environmental Scientist
37. WILD 4000/6000–4000L/6000L – Management of Wildlife Habitat

See also: Carbon Cycle. Economic Control of Invasive Species. Ecosystem, Concept of. Ecosystem Services. Energy Use, Human. Food Webs. Lake and Pond Ecosystems. Nitrogen, Nitrogen Cycle. Trophic Levels

References

Anderson T, Carstensen J, Hernandez-Garcia E, and Duarte C (2009) Ecological thresholds and regime shifts: Approaches to identification. *Trends in Ecology and Evolution* 24: 49–57.

Carpenter SR, Brock WA, Cole JJ, Kitchell JF, and Pace ML (2008) Leading indicators of trophic cascades. *Ecology Letters* 11: 128–138.

Carpenter SR, Cole JJ, Kitchell JF, and Pace ML (2010) Trophic cascades in lakes: Lessons and prospects. In: Terborgh J and Estes J (eds.) *Trophic Cascades*, pp. 55–69. Washington DC: Island Press.

Chapin FS, Matson PA, and Vitousek PM (2011) *Principles of Terrestrial Ecosystem Ecology*, 2nd edn. New York: Springer.

Coleman DC (2011) *Big Ecology. The Emergence of Ecosystem Science.* Berkeley, CA: University of California Press.

Covich AP (1993) Water and ecosystems. In: Gleick PH (ed.) *Water in Crisis*, pp. 40–55. Oxford: Oxford Univ Press.

Covich AP, Austen MC, Barlocher F, *et al.* (2004) The role of biodiversity in the functioning of freshwater and marine benthic ecosystems. *BioScience* 54(2004): 767–775.

Covich AP, Crowl TA, Hein CL, Townsend MJ, and McDowell WH (2009) Importance of geomorphic barriers to predator–prey interactions in river networks. *Freshwater Biology* 54.

Covich AP, Palmer MA, and Crowl TA (1999) The role of benthic invertebrate species in freshwater ecosystems. *BioScience* 49: 119.

Crowl TA, McDowell WH, Covich AP, and Johnson SL (2001) Species specific effects of freshwater shrimp on detrital processing and localized nutrient dynamics in a montane tropical rain forest stream. *Ecology* 82: 775–783.

Elser JJ, Fagen WF, Kerkhoff AJ, and Enquist BJ (2010) Biological stoichiometry of plant production: Metabolism, scaling, and ecosystem response to global change. *New Phytologists* 186: 593–608.

Golley FB (1993) *A History of the Ecosystem Concept in Ecology.* New Haven, CT: Yale University Press.

Hagen JB (1992) *An Entangled Bank: The Origins of Ecosystem Ecology.* New Brunswick: Rutgers University Press.

Heal GM, Barbier EB, Boyle KJ, *et al.* (2005) *Valuing Ecosystem Services: Toward Better Environmental Decision-Making.* Washington, DC: National Research Council.

Hobbs RJ, Higgs E, and Harris JA (2009) Novel ecosystems: Implications for conservation and restoration. *Trends in Ecology and Evolution* 24: 599–605.

Hutchinson GE (1944) Nitrogen in the biogeochemistry of the atmosphere. *American Scientist* 32: 178–195.

Hutchinson GE (1948) Circular causal systems in ecology. *Annals of the New York Academy of Sciences* 50: 221.

Hutchinson GE and Bowen VT (1950) Limnological studies in Connecticut IX. A quantitative radiochemical study of the phosphorus cycle in Lindsey Pond. *Ecology* 31: 194–203.

Lawton JH (1997) The role of species in ecosystems: Aspects of ecological complexity and biological diversity. In: Abe T, Levin SA, and Higashi M (eds.) *Biodiversity: An Ecological Perspective*, pp. 215–228. New York: Springer-Verlag.

Lindeman RL (1942) The trophic-dynamic aspect of ecology. *Ecology* 23: 399.

Naeem S (1998) Species redundancy and ecosystem reliability. *Conservation Biology* 12: 39.

Paine RT (1969) A note on trophic complexity and community stability. *American Naturalist* 103: 91.

Palmer M, Covich AP, Findlay BJ, *et al.* (1997) Biodiversity and ecosystem processes in freshwater sediments. *Ambio* 26: 571.

Parker JD, Burkepile DE, Collins DO, Kubanek J, and Hay ME (2007) Stream mosses as chemically defended refugia for freshwater invertebrates. *Oikos* 116: 302–312.

Post DM and Palkovacs EP (2009) Eco-evolutionary feedbacks in community and ecosystem ecology: Interactions between the ecological theatre and the evolutionary play. *Philosophical Transactions of the Royal Society B* 364(1523): 1629–1640.

Post DM and Palkovacs EP (2009) Eco-evolutionary feedbacks in community and ecosystem ecology: Interactions between the ecological theatre and the evolutionary play. *Philosophical Transactions of the Royal Society B* 364: 1629–1640.

Scheffer M (2009) Critical transitions in nature and society. *Princeton Studies in Complexity*. Princeton, New Jersey: Princeton University Press.

Terborgh J and Estes JA (2010) *Trophic Cascades: Predators, Prey and the Changing Dynamics of Nature*. Washington, DC: Island Press.

Tilman D, Lehman CL, and Bristow CE (1998) Diversity–stability relationships: Statistical inevitability or ecological consequences? *The American Naturalist* 151: 277.

Waide RB, Willig MR, Steiner CF, *et al.* (1999) The relationship between productivity and species richness. *Annual Review of Ecology and Systematics* 30: 257.

Weathers KC, Strayer DL, and Likens DE (eds.) (2011) *Fundamentals of Ecosystem Science*. Elsevier.

Energy Use, Human

Patrick Gonzalez, University of California, Berkeley, CA, USA

Published by Elsevier Inc.

Glossary

Energy The capacity to perform work. Potential energy is this capacity stored as position (e.g., in a gravitational or electromagnetic field) or as structure (e.g., chemical or nuclear bonds). Kinetic energy is this capacity as manifested by the motion of matter. The joule (J) is the common SI unit of energy, where 1 J equals the amount of energy required to increase the temperature of 1 g of water by 1 K. Other units include kilocalories (kcal), kilowatt-hours (kWh), and British thermal units (BTU).

Energy efficiency A measure of the performance of an energy system. First Law efficiency, the most commonly used measure, equals the ratio of desired energy output to the energy input. Second Law efficiency equals the ratio of the heat or work usefully transferred by a system to the maximum possible heat or work usefully transferable by any system using the same energy input.

Energy, industrial Forms of energy generally transformed in bulk at centralized facilities by means of complex technology. The major forms of industrial energy are oil, coal, natural gas, nuclear, and hydroelectric. In addition to hydroelectric, industrial energy also includes other technologically complex renewable energy systems, including solar photovoltaics, wind turbines, geothermal systems, and biofuels.

Energy, nonrenewable Forms of energy whose transformation consumes the energy source. The major forms include oil, coal, natural gas, and nuclear.

Energy, renewable Forms of energy whose transformation does not consume the ultimate source of the energy, harnessing instead solar radiation, wind, the motion of water, and subsurface water and geologic formations. The major forms of renewable energy are solar, wind, geothermal, hydropower, and biomass. Renewable energy systems that depend on complex technology are forms of industrial energy. The simpler renewable systems are forms of traditional energy.

Energy, traditional Forms of energy generally dispersed in nature, renewable, and utilized in small quantities by rural populations. The principal forms of traditional energy are firewood, charcoal, crop residues, dung, and small water and windmills.

Fossil fuels Forms of stored energy produced by the action of pressure and temperature on organic matter buried over geologic time. The major types of fossil fuels are oil, natural gas, and coal.

Law of Thermodynamics, First Physical principle that energy is neither created nor destroyed, only converted between different forms. Energy is therefore conserved. In thermodynamic terms, the change in energy of a system equals the difference of the heat absorbed by the system and the work performed by the system on its surroundings.

Law of Thermodynamics, Second Physical principle that any system will tend to change toward a condition of increasing disorder and randomness. In thermodynamic terms, entropy must increase for spontaneous change to occur in an isolated system.

Power The rate of energy transformation over time. The watt (W) is the common SI unit of power, where 1 W equals the power expended by the transformation of 1 J s^{-1}.

Patterns and Scale of Human Energy Use

We use energy to meet subsistence needs and to fulfill non-essential wants. In a subsistence society, a farmer burns wood to cook the day's meals. In an industrial society, people drive a car to go see a movie. Yet the forms of energy involved in these activities – wood, gasoline, electricity – merely comprise the means to end-uses – cooking, driving, operating a movie theater – that ultimately provide desired services – food, transportation, entertainment.

As used by humans, energy falls into two broad categories: industrial and traditional. Industrial energy includes those forms of energy generally transformed in bulk at centralized facilities by means of complex technology. In general, these forms fuel the technologies developed since the Industrial Revolution. The major forms of industrial energy are oil, coal, natural gas, nuclear, and hydroelectric. Industrial energy also includes other technologically complex renewable energy systems, including solar photovoltaics, wind turbines, geothermal systems, and biofuels. Cogeneration systems that produce electricity and industrial heat from natural gas and hybrid gasoline-electric vehicles combine two forms of energy transformation to increase overall efficiency.

Traditional energy includes those forms generally dispersed in nature and utilized in small quantities by rural people. The principal forms are firewood, charcoal, crop residues, dung, and small water and windmills. Because traditional energy sources occur widely and because their transformation does not rely on complex technology, they comprise the most important sources for rural people in the less industrialized regions of the world. A rural household generally harvests traditional energy sources for its domestic needs. Because no commercial transaction occurs in these situations and because most governments do not regulate the use of traditional sources, official statistics do not closely track traditional energy use.

Traditional energy is a form of renewable energy, which includes forms of energy whose transformation does not

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00242-2

consume the ultimate source of the energy. Renewable energy harnesses solar radiation, wind, water movement, and cold and hot reservoirs in water and geologic strata. The major forms of renewable energy are solar, wind, geothermal, hydropower, and biomass. Conversely, nonrenewable energy systems consume the very source of the energy, primarily oil, coal, natural gas, and nuclear fuel.

In addition to traditional energy and industrial hydroelectric, renewable energy sources include technologically complex systems: solar photovoltaics, wind turbines, geothermal systems, and biofuels. These systems require complex machinery associated with industrial energy, yet depend only upon nondestructive methods of harnessing natural energy sources.

World energy use has increased substantially in the past 150 years (**Figure 1**). In the twentieth century, energy use increased by a factor of 12. During this time, the world witnessed an explosion in the use of fossil fuels, while renewable energy use increased slightly.

In 2008, the world used energy at a rate of 16.2 trillion watts (16.2 TW) (**Table 1**). As a comparison, this rate of energy use is equivalent to the power drawn continuously by 1.1 trillion compact fluorescent light bulbs rated at 15 W. As another comparison, world energy use in 2008 required the equivalent of the continuous output of 16,000 standard 1 GW nuclear plants.

The world uses renewable energy sources for 12% of its energy (**Table 2**). The rate of hydroelectric energy use (**Tables 1** and **2**) is equal to the hydroelectricity output. Unlike fossil fuel combustion systems, hydroelectric systems do not waste energy as heat. If an equivalent amount of fossil fuel were used to generate the 0.4 TW generated by hydroelectric plants, then the world rate of energy use would total 16.8 TW. Thus, hydroelectric displaces 0.6 TW of fuel use and the impacts from that amount of industrial energy transformation.

The world depends on industrial energy for 90% of its energy use. Fossil fuels and nuclear are the primary industrial energy sources. Traditional energy comprises approximately one-tenth of world energy use. Countries in Africa, Asia, and Latin America account for the majority of world traditional energy use, with firewood and charcoal as the main sources.

While the United States (U.S.) is home to 4.5% of the world's population, it uses 19% of the world's energy (IEA, 2010). The U.S. fraction of global energy use has decreased steadily in the twenty-first century because of large increases in other countries, especially the People's Republic (P.R.) of China, where energy use has grown to the amount used in the U.S. (IEA, 2010).

One measure of energy efficiency is energy use per person. In 2008, average world energy use was 2400 W person^{-1} (IEA, 2010). Average U.S. energy use (10,000 W person^{-1}) exceeds energy use per person in many other industrial countries, for example, France (5500 W person^{-1}). On average, each person in the U.S. uses six times the amount of energy as each person in Brazil (1700 W person^{-1}) and 15 times the amount of energy of each person in India (700 W person^{-1}) (**Figure 2a**).

Another measure of energy efficiency is energy intensity, the amount of energy used per unit of economic production. In 2008, average world energy intensity was 0.25 W \$$^{-1}$ of gross domestic product (GDP), adjusted for purchasing power parity (IEA, 2010). U.S. energy intensity matched the world average. The U.S. has steadily increased its energy efficiency for decades, although it still is less efficient than many other industrial countries, for example, the United Kingdom (0.15 W \$$^{-1}$) (**Figure 2b**).

In 2008, the world used 37% of energy for industry, 36% for the commercial and residential sectors, and 27% for transportation (IEA, 2010). In 2009, U.S. energy end-uses were: industry 30%, commercial 19%, residential 22%, and transportation 29% (U.S. Department of Energy, 2010). Globally, a third of energy goes to electricity generation, mainly from coal, hydroelectric, and nuclear. Power plants release two-thirds of that as waste heat (see the next section). In the U.S., personal cars use half of all transportation energy.

The traditional energy sources of firewood and charcoal mainly serve the end-uses of cooking and heating. Cooking 1 J

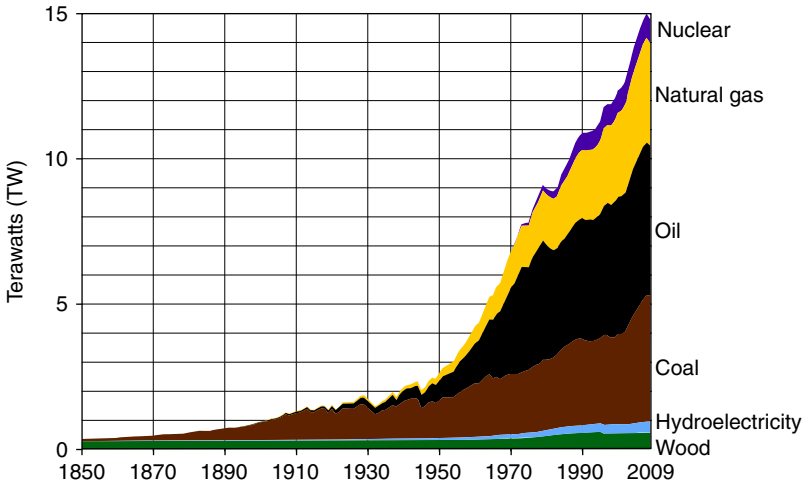


Figure 1 World energy use 1850–2009 (data from WEC and IIASA, 1995; BP, 2010; Food and Agriculture Organization (FAO) FAOSTAT <http://faostat.fao.org>; International Energy Agency (IEA) Statistics <http://www.iea.org/stats>). Because of a difference in FAO and IEA estimates of wood and biomass use, total 2008 energy use using FAO data is lower (15 TW) than total energy use using IEA data (16.2 TW).

Table 1 Energy use 2008

	Oil	*Natural gas*	*Coal*	*Nuclear*	*Hydroelectric*	*Renewable technologies*	*Wood and other biomass*	*Total*	*Fraction of total*
					(TW)				*(%)*
Africa	0.2	0.1	0.1	<0.1	<0.1	<0.01	0.5	1.0	6
Asia	2.0	1.1	2.7	0.2	0.1	0.01	0.6	6.8	42
Australia and New Zealand	0.1	<0.1	0.1	0	<0.1	<0.01	<0.1	0.2	1
Europe	1.0	0.7	0.5	0.4	0.1	0.02	0.1	2.8	17
Latin America	0.5	0.2	<0.1	<0.1	0.1	<0.01	0.2	1.1	7
Russia	0.2	0.5	0.1	0.1	<0.1	<0.01	<0.1	0.9	6
USA and Canada	1.3	0.9	0.8	0.3	0.1	0.01	<0.1	3.5	21
World	5.3	3.6	4.4	0.9	0.4	0.04	1.6	16.2	100
Fraction of total (%)	32	22	27	6	2	<1	10	100	

Source: Data from BP, 2010; IEA, 2010, Food and Agriculture Organization FAOSTAT http://faostat.fao.org; International Energy Agency Statistics http://www.iea.org/stats

Table 2 Renewable energy generation capacity 2009

Source	*GW*
Wood and other biomass	1600
Hydroelectric, large (>10 MW)	310
Biomass-to-energy (heat)	270
Solar thermal heat	180
Wind electric	160
Hydroelectric, small (<10 MW)	60
Geothermal heating and cooling	60
Biomass-to-energy (electricity)	54
Solar photovoltaics	21
Geothermal electricity	11

Source: Data from International Energy Agency (IEA) (2010) *Key World Energy Statistics* 2010. Paris: IEA and Renewable Energy Policy Network for the 21st Century (REN21) (2010) *Renewables 2010 Global Status Report*. Paris: REN21.

of food requires approximately 2 J of firewood or 8 J of wood converted to charcoal. Consequently, rural people use 1–2 kg wood person^{-1} day^{-1}, a rate of energy use of 250–500 W person^{-1} (FAO, 2009). Only 20–40 W person^{-1} enters the cooked food and warmed people because open fires diffuse the rest as waste heat (Gonzalez, 2001).

In urban areas of nonindustrial countries, people often rely on charcoal for energy. Even though the conversion of wood to charcoal releases waste heat, the end product has a higher energy density than firewood, making charcoal easier to store and transport. Urban people use 100–150 kg charcoal person^{-1} year^{-1}, requiring 800–1200 kg wood person^{-1} year^{-1}. The ultimate energy end-use is 30–45 W person^{-1}.

Implications of the Laws of Thermodynamics

The First Law of Thermodynamics states that energy is not created or destroyed, only converted between different forms. This is the principle of conservation of energy (**Table 3**). The First Law means that any energy that a process does not convert into useful forms will go into useless forms. Humans dump this waste energy into the environment.

The Second Law of Thermodynamics states that any system will tend to change toward a condition of increasing disorder and randomness. This is the principle of increasing entropy (**Table 3**). The Second Law means that no energy transformation can convert 100% of one energy form completely into a useful form. The process will always waste energy in forms that are unrecoverable due to the disorderliness or randomness of the waste energy forms. Therefore, the fewer energy transformations that a system performs, the fewer chances it creates for Second Law energy losses.

For example, the objective of an automobile's internal combustion engine is to convert the chemical energy in the covalent bonds of hydrocarbons in gasoline to heat energy in an expanding fuel–air mixture in the piston to kinetic energy in the drive shaft to kinetic energy in the main axle. No matter how efficient an engine and automobile can be, combustion processes will always waste energy as heat in the friction of engine parts, sound in the banging of vehicle components, heat in the friction of tires on the road, kinetic energy of the wind displaced by the vehicle, and countless other unrecoverable losses.

As another example, the objective of a coal-fired electric power plant is to transform the chemical energy in the covalent bonds of hydrocarbons in coal to heat energy in the boiler to heat energy in steam to kinetic energy in the turbine fan to electromagnetic energy in the generator coil. Along the way, the conversion processes lose energy as light and sound of the boiler fire, vibration of turbine parts, heat of power plant components, and, most of all, waste heat carried by the power plant cooling water.

Theoretically, the maximum efficiency across a heat gradient is the Carnot efficiency:

$$\eta = \frac{\text{Temperature of heat sink}}{\text{Temperature of transformation process}}$$

with temperatures in Kelvin (K).

For a coal-fired power plant, materials limit boiler temperatures to 1000–1200 K. At an ambient environmental temperature of 293 K (20 °C), the maximum efficiency will be 70–75%. Typically, coal plants only achieve efficiencies of 30–35%, releasing two-thirds of the total as waste heat.

Biodiversity Impacts of Industrial energy

Oil

The major impacts of oil on biodiversity come from a fuel cycle and end-uses that range over vast areas. Exploration, drilling, crude oil transport, refining, and combustion in vehicles alter terrestrial and marine habitat and pollute air, land, and water.

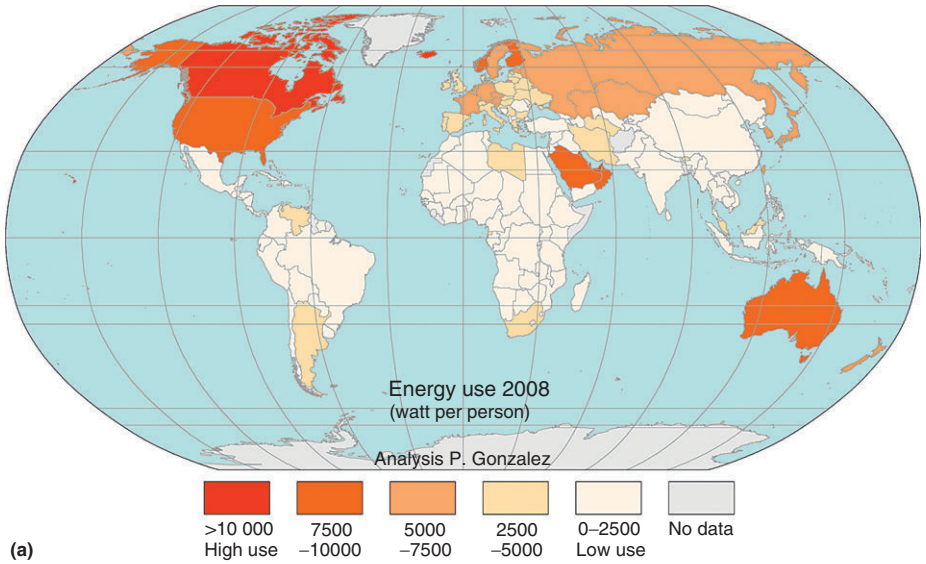
Petroleum, or oil, consists of a complex mixture of hydrocarbons formed over geologic time from organic matter compressed under anoxic conditions. The most important chemical constituents are alkanes (e.g., octane and methane) and aromatic hydrocarbons (e.g., benzene and toluene).

Oil deposits originated in aquatic vegetation deposited in inland seas and coastal basins during the Cretaceous period 100 million years ago. In the early stages of formation, bacteria initiated the anoxic reduction of the organic matter. Over time, pressure and temperature replaced microbial activity as the primary agent of transformation. Eventually, these forces drove off most of the water, oxygen, and nitrogen in the condensate, leaving oil compounds enriched in carbon and hydrogen. Dispersed between sediment granules, the oil eventually migrated to low-pressure geologic traps at depths of 1–7 km. On average, the stoichiometric composition of crude oil is $CH_{1.5}$, with a small amount of sulfur.

Petroleum exploration involves geologic surveys over extensive areas, often with low human population densities and relatively undisturbed natural communities. Exploration brings vehicle traffic and construction activity, generating local disturbances, yet the most serious impacts occur with seismic detection. This involves controlled detonations along lines or at points so that seismometers can extrapolate the layout of subsurface formations. These activities destroy vegetation, disturb ground-nesting birds and other wildlife, and fragment habitat. If these activities disturb animal behavior during breeding periods, the impact can last over many growth cycles.

Highest energy use	W person^{-1}
Qatar	26 000
Iceland	21 900
Trinidad and Tobago	19 300
United Arab Emirates	17 300
Bahrain	16 000
Netherlands Antilles	14 800
Kuwait	12 800
Brunei Darussalam	12 100
Luxembourg	11 200
Canada	10 600

Lowest energy use	W person^{-1}
Comoros	76
Malawi	72
Uganda	47
Burkina Faso	46
Central African Republic	45
Niger	41
Rwanda	34
Mali	32
Burundi	27
Chad	12



Lowest efficiency	W $^{-1}$
Iraq	1.45
Trinidad and Tobago	1.22
Uzbekistan	1.09
Netherlands Antilles	0.96
Qatar	0.94
Nigeria	0.88
Tanzania	0.84
Zambia	0.77
Bahrain	0.72
Kazakhstan	0.72

Highest efficiency	W $^{-1}$
Central African Republic	0.04
Burundi	0.04
Lesotho	0.04
Niger	0.04
Uganda	0.03
Burkina Faso	0.03
Guinea	0.02
Rwanda	0.02
Mali	0.02
Chad	0.01

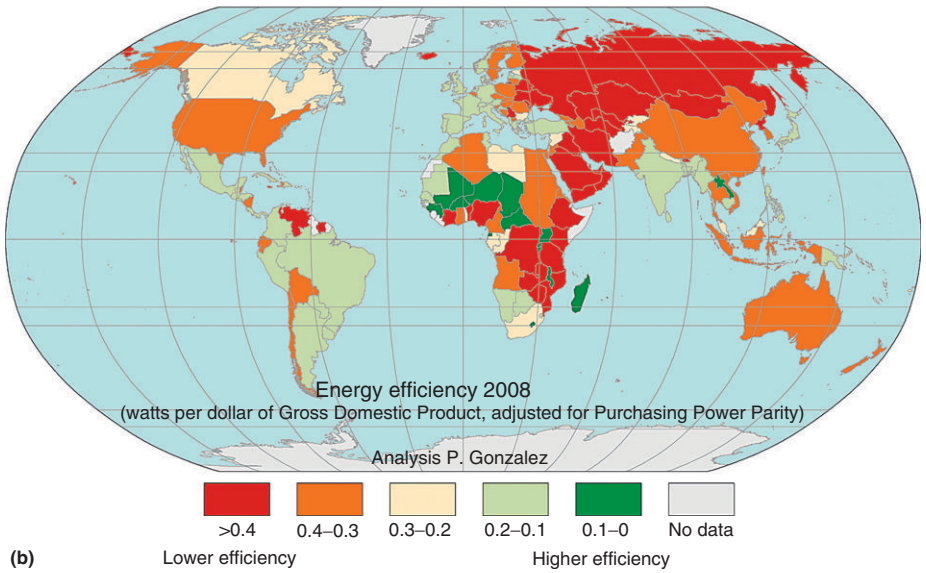


Figure 2 Energy efficiency by country (data from IEA, 2010; U.S. Department of Energy International Energy Statistics <http://www.eia.doe.gov>). (a) Energy use per person in 2008 and countries with the highest and lowest energy use per person. (b) Energy intensity in 2008 (watts per dollar (year 2000) of purchasing power parity (PPP)-adjusted gross domestic product (GDP)) and the most and least energy efficient countries.

Table 3 Formulations of the First and Second Laws of Thermodynamics

	First Law	*Second Law*
Universal	The total energy in the universe is constant	All physical processes proceed such that the entropy of the universe increases
Concise	Energy is conserved	Entropy increases
Colloquial	You can't get something for nothing	You can't even break even

Based on original formulations by John P. Holdren.

Since the drilling of the world's first commercial oil well in Titusville, Pennsylvania in 1859, oil fields have grown to cover vast areas, especially in the Middle East, North America, Russia, Mexico, Venezuela, and Nigeria. Construction of oil well pads, buildings, roads, electricity and water lines, and other infrastructure destroys vegetation and natural habitat. Roads and other linear infrastructure fragment a landscape that extends outside core areas of operation. For example, while the core infrastructure at the Prudhoe Bay field, Alaska (opened for drilling in 1968) covers 70 km^2, the 2000-km network of roads, short-distance pipelines, and electric lines cover 2600 km^2 of Arctic tundra. Proposed oil exploitation in the

Arctic National Wildlife Refuge in Alaska would further fragment the Arctic ecosystem.

Hundreds of thousands of kilometers of pipelines carry oil and refined products around the world, fragmenting and polluting natural areas. For example, the 1000 km pipeline for the Chad–Cameroon petroleum project sliced through intact tropical forests and attracted new migrants who cleared land for agriculture. The 700-km Camisea natural gas pipeline in Peru cut across Amazon rainforest, fragmented habitat, spilled gas liquids, and eroded soil into rivers.

To mitigate environmental impacts of the 1300-km Trans-Alaska pipeline and adjacent access road (operations started in 1977), oil companies integrated a set of environmental protection features. To prevent thawing of permafrost, brackets elevate 700 km of pipeline to heights of 3 m. Pipes at the bracket legs dissipate heat generated by the friction of oil passing through the pipe. Elevated sections provide passages for caribou (*Rangifer tarandus*). Bridges carry the pipeline over 800 streams. Zigzags along the pipeline translate the longitudinal movement of pipes expanding under heat to lateral movement, reducing the risk of leakage.

Routine drilling operations produce air and water pollution. Serious water pollution comes from drilling muds, a mixture of water and low-molecular-weight oil that drillers pump down to the drill bit to carry away rock cuttings, keep the bit from overheating, and protect the drill shaft from surrounding rock. Used muds contain bits of metal and trace metals mobilized from drilled rock.

All stages of oil production, from drilling to end-use, spill oil into surface and ground waters. Oil spills into oceans and freshwater bodies average 1.3 million tons $year^{-1}$ (NRC, 2003). Approximately 15% of spilled oil comes directly from oil wells, pipelines, and supertankers, 40% comes from nonpoint urban runoff, and the remaining comes from natural seeps. During oil production, spills occur at well blowouts when equipment fails to contain naturally high fluid pressures in oil-bearing strata. Spills also occur along the long lengths of pipeline from the wellhead to tank farms to supertanker ports to refineries to gas stations. Pipes, valves, and tanks leak from fatigue and from human error. The most catastrophic accidental oil spills in history have occurred from offshore oil wells and supertankers (**Table 4**).

Oil floats on top of water. Gravity and wind will spread a floating slick out to a thickness of 0.5–10 μm. Patches 0.1–5 mm thick can cover just 10% of the total slick area, yet contain 90% of the total slick volume. Some oil dissolves and emulsifies into the water column, forming emulsions containing 80% H_2O. Oil will generally not sink to depths below 20 m. Recovery teams can deploy an array of countermeasures that includes booms, skimmers, sorbents, pumps, burning, and surfactants for chemical dispersion.

Exposure to sunlight initiates photolysis of hydrocarbons into lower molecular-weight compounds. Heterotrophic bacteria will also oxidize hydrocarbons to smaller compounds, CO_2, and water. The lightest hydrocarbons and aromatic compounds (e.g., benzene) volatilize. Loss of the lighter fraction leaves the remaining slick more viscous over time. This thick oil forms tar balls and pancake-like forms. Some oil will sink into sediments; other oil will coat beaches; a fraction of oil will remain suspended in the water column for years.

Oil kills birds by coating, matting, and waterlogging their feathers. The water repellency, buoyancy, and insulating properties of plumage derive from an orderly arrangement of feather barbules and barbicelles. Contact with oil disrupts these arrangements. Soaked birds can die of hypothermia and drowning. Survivors remain at risk for chronic exposure to toxic organic compounds through ingestion, inhalation of fumes, or absorption. Moreover, eggs are highly sensitive to contact with oil.

Many of the aromatic hydrocarbons in petroleum, including benzene, toluene, xylene, and phenols, are lethal to animals on contact and carcinogenic under chronic exposure. Moreover, polycyclic aromatic hydrocarbons bond to lipophilic sites, an affinity that concentrates these compounds in body tissues as they move up the food chain. Even when not deadly, sublethal disruption of physiology or behavior reduces resistance to infection and induces stress.

Oil obliterates the insulating properties of marine mammal pelage, leaving animals to die of hypothermia. In addition, oil can clog the nostrils of seals, causing them to suffocate. Whales, insulated by layers of oily blubber, not hair, resist the effects of oil. Marine mammals may experience chronic problems after a spill because oil can accumulate in bile and fatty tissues.

Important fisheries often lie over high-yield offshore oil fields on the continental shelves. Oil spills kill fish directly and chronic effects reduce fitness years after initial exposure. Oil at the air–water interface acts as a physical barrier that impedes gas exchange. Under a thick slick, fish larvae can suffocate.

Table 4 Most voluminous oil spills since 1960

Rank	*Incident*	*Volume (million liters)*	*Date*
1	Persian Gulf War	900	26 January 1991
2	Deepwater Horizon unit, Macondo well, Gulf of Mexico	700	20 April 2010
3	Ixtoc I well, Gulf of Mexico	530	3 June 1979
4	Fergana Valley well, Uzbekistan	330	2 March 1992
5	*Atlantic Empress* tanker, off Trinidad and Tobago	320	19 July 1979
6	Nowruz platform no. 3, Persian Gulf	300	4 February 1983
7	*Castillo de Bellver* tanker, off coast of South Africa	300	6 August 1983
8	*ABT Summer* tanker, off coast of Angola	290	28 May 1991
9	*Amoco Cadiz* tanker, off coast of France	260	16 March 1978
10	*Odyssey* tanker, off coast of Canada	160	10 November 1988

Source: Etkin, DS (1999) Historical overview of oil spills from all sources (1960–1998). *International Oil Spill Conference*, paper no. 169; Crone TJ and Tolstoy M (2010) Magnitude of the 2010 Gulf of Mexico oil leak. Science 330: 634, and International Tanker Owners Pollution Federation (ITOPF) (2010) *Oil Tanker Spill Statistics: 2009*. London: ITOPF.

Fish eggs, which often float at the sea surface, and fish larvae, which are often distributed in the upper water column, both occur in the areas of high oil concentration. Hydrocarbons damage eggs on contact. Oil concentrations will quickly exceed the LC_{50} for fish larvae of 1–10 ppm. The early-life stages of intertidal-spawning fish are especially susceptible.

Over time, tidal action coats the shore of the intertidal zone in a band of oil. This oil ring smothers intertidal invertebrates, crustaceans, mussels, barnacles, limpets, and algae. Oil will asphyxiate filter feeders.

Oil spills also damage phytoplankton and other marine plants. Because oil absorbs photosynthetically active radiation, a coat of oil will hinder plant growth and increase plant tissue temperature. Aromatic hydrocarbons may disrupt the orderly arrangement of grana in chloroplasts. Blue-green algae often bloom at oil spills, producing anoxic conditions.

The largest accidental oil spill in the world began with the 20 April 2010 blowout of the Deepwater Horizon mobile offshore drilling unit and the Macondo oil well in the Gulf of Mexico, off the coast of Louisiana, USA. A drilling pipe extended from the Deepwater Horizon on the ocean surface 1500 m to the Macondo wellhead on the ocean floor and 5600 m more to the oil reservoir. After the failure of a blowout preventer device at the wellhead, oil and natural gas shot up to the Deepwater Horizon and ignited in a massive explosion that killed 11 people and sunk the unit 2 days later. For 84 days, the well spilled 4.4 $\pm$ 0.9 million barrels of oil (700 $\pm$ 140 million liters) into the Gulf of Mexico (Crone and Tolstoy, 2010). Direct observations and modeling estimated the fate of the oil as: recovered 16–17%; burned 5–6%; skimmed 2–4%; chemically dispersed 10–29%; naturally dispersed 12–13%; evaporated or dissolved 20–25%; and washed up on coast, sunk due to sedimentation, exported from area in ocean currents 11–30% (Federal Interagency Solutions Group, 2010).

To reduce the amount of oil that could reach beaches, crews released 7 million liters of chemical dispersants, a record amount, at the wellhead and on the surface of the Gulf. As a consequence, much of the oil remained in the water column, increasing contact to and ingestion of oil by marine animals. U.S. Fish and Wildlife Service counts indicated that laughing gulls (*Larus atricilla*), northern gannets (*Morus bassanus*), brown pelicans (*Pelecanus occidentalis*), and royal terns (*Sterna maxima*) had the highest numbers of oiled birds, of over 100 species recorded. Also, leatherback sea turtles (*Dermochelys coriacea*), an endangered species, were oiled and biologists relocated nests of Kemp's ridley sea turtle (*Lepidochelys kempii*), another endangered species, to reduce the chance of exposure to spilled oil. The Gulf provides important marine habitat for sperm whales (*Physeter catodon*), an endangered species, bottlenose dolphins (*Tursiops truncatus*), and other marine mammals. Oil also coated areas of coastal wetlands, damaging marsh plants. Continued contact with oil could cause chronic harm to coastal plant and animal species for years. On the Mexican side of the Gulf of Mexico, oil spilled from the Ixtoc I remains under rocks in some areas decades after the spill. In France, local marine species and oyster farms needed a decade to recover from the Amoco Cadiz oil spill.

The largest oil spill in continental U.S. waters occurred on 24 March 1989, when the supertanker *Exxon Valdez* ran aground in Prince William Sound, Alaska, and poured out 42 million liters of crude oil. The spill caused acute damage to birds, marine mammals, and intertidal communities and caused chronic damage to fish species and intertidal and subtidal communities.

The spill occurred in early spring, just before the young of many wildlife species emerged to rejuvenate marine animal populations. Pacific herring (*Clupea pallasi*) were spawning inshore. Millions of pink salmon (*Oncorhynchus gorbuscha*) fry were soon to be washed from gravel spawning beds into the spring plankton bloom offshore. Harbor seal (*Phoca vitulina*) and sea otter (*Enhydra lutris*) pups were testing the frigid waters. Seabirds were beginning to converge on breeding colonies in the gulf. The *Exxon Valdez* spill killed more than a quarter of a million birds of over 90 species, including marbled murrelets (*Brachyramphus marmoratus*), bald eagles (*Haliaeetus leucocephalus*), puffins (*Frattercula* spp.), and loons (*Gavla* spp.). Exposure to toxics from the *Exxon Valdez* spill caused chronic problems in fish, with elevated egg, larvae, and adult mortality, larval deformities, and poor adult growth rates.

The high energy density and flexibility of liquid fuels refined from oil render them convenient for powering motor vehicles. Onshore refining of oil into other products generates liquid and solid waste that pollutes natural water bodies. Refining employs catalytic cracking of carbon–carbon bonds of long-chain alkanes to produce lower molecular-weight hydrocarbons. Refineries attempt to recover every possibly useful organic compound, from light products, such as methane, benzene, and kerosene, to medium weight products, like gasoline and diesel fuel, to heavy tars and asphalt. Unless treated, wastewater from these processes and from sulfur recovery will pollute surface waters.

Most constituents of petroleum and refined oil products volatilize easily, so each step of the petroleum fuel cycle generates air pollution. Methane, benzene, toluene, and other compounds evaporate from crude oil exposed to air. The major emissions from oil refineries are CH_4, CO, CO_2, H_2S, NO_x, and SO_2.

This section has concentrated on the impacts from the core stages of the petroleum fuel cycle: exploration, extraction, transportation, and refining. Nevertheless, the entire life cycle, from manufacturing to disposal, of oil infrastructure generates air and water pollution and alters natural habitats. Combustion of refined oil products for automobiles, heating, and other end-uses generates the most serious emission of the fuel cycle, carbon dioxide, the principal greenhouse gas. Moreover, armed conflicts caused, in part, by efforts to control access to oil fields and refineries, directly harm human wellbeing and ecosystem health.

Natural Gas

Natural gas is a mixture of hydrocarbons that exists in a gaseous state at standard temperature and pressure. Methane (CH_4) is the main constituent, but the presence of higher molecular-weight alkanes, including ethane, propane, and butane, changes the average stoichiometric composition for natural gas, with water vapor removed to 0.79 $CH_{3.62}$. Formed by the same processes that formed oil, natural gas is often found at the top of oil deposits. The most voluminous natural gas reservoirs occur in Cretaceous strata.

Land use changes caused by exploration and extraction of natural gas produce the same biodiversity impacts described for oil. In particular, exploration grids, roads, and other linear infrastructure fragment landscapes that extend far outside core areas of operation. Accelerated natural gas leasing in the intermountain region of the western U.S. since 2001 has created networks so dense that no undisturbed ecosystems remain in some areas, threatening grassland habitats and species such as the greater sage-grouse (*Centrocercus urophasianus*) and the pronghorn (*Antilocapra americana*).

In the nineteenth century, companies had not yet erected natural gas pipelines or processing facilities. Moreover, industry had not yet developed extensive technology that used natural gas. Because companies found natural gas uneconomical to exploit, they burned it off to reduce the risk of fire and explosion. Consequently, the entire history of natural gas production has flared the equivalent of 8 years worth of U.S. energy use (Boden *et al.*, 2010). In 2008, natural gas flaring globally wasted an amount of energy equivalent to a quarter of U.S. natural gas use, or all of the natural gas use of Latin America, or twice the natural gas use of Africa (Elvidge *et al.*, 2009). Currently, Russia and Nigeria flare the most natural gas.

Gas companies generally pump natural gas straight from the well to a processing plant, avoiding the need for storage facilities at the wellhead and reducing the potential for leakage. Gas companies generally divide natural gas into three fractions: natural gas liquids (NGL), liquified petroleum gas (LPG), and liquified natural gas (LNG). NGL consists of the higher molecular-weight fraction of natural gas that often settles out by gravity. Processing of natural gas from oil wells produces LPG. Finally, pressurization of natural gas produces LNG, a product that is expensive because of the special containers required for transport.

Hundreds of thousands of kilometers of natural gas pipelines destroy and fragment natural ecosystems and create avenues for people to settle and clear forests. Construction of specialized natural gas terminals can destroy or alter coastal ecosystems.

The major end-uses of natural gas, cooking and heating, burn the fuel directly with no further transformation. Electricity generation from natural gas employs a gas turbine, which directly uses the hot gas products of combustion to turn the turbine fan, eliminating the intermediate step of steam generation used in oil and coal-fired plants. Cogeneration plants increase the energy efficiency of gas turbine systems by utilizing the waste heat of gas turbines for space heating or industrial processes.

Extraction and combustion of natural gas pollute much less than extraction and combustion of oil. Because it exists in a gaseous state for much of the fuel cycle, natural gas exploitation does not produce significant amounts of water pollution. Yet, methane is a greenhouse gas with a global warming potential of 21, indicating an impact on global warming 21 times more intense than carbon dioxide. The combustion of methane also produces carbon dioxide and contributes to climate change.

Coal

Coal consists of hard carbonaceous material formed by compression and transformation of terrestrial plant matter rich in cellulose, buried at the bottom of ancient freshwater swamps and bogs. The richest coal-bearing strata date from the Cretaceous period 100–200 million years ago and the Permian period 250 million years ago. Similar to the process of petroleum formation, deposited plant matter undergoes incomplete decay in anoxic conditions.

In geologic time, the pressure of overlying rock and the heat generated therein drive off oxygen and hydrogen, leaving thick seams of reduced carbonaceous rock containing much more organic than mineral matter. The average stoichiometric equation of coal is $0.75CH_{0.8}$, but elemental sulfur contaminates most coal deposits. The four major types of coal, in order of decreasing carbon content and increasing sulfur, are: anthracite, bituminous, subbituminous, and lignite. Bituminous coal is the most abundant type worldwide. Peat, the partially oxidized, moist, organic soil that forms in marshes and bogs, is an early precursor of coal. In certain regions, people burn peat for heating, cooking, and lighting.

The coal fuel cycle extends from extraction at the mine to combustion at a power plant to distribution across the electric grid to end-uses in lighting, heating, and all the other uses of electricity. Coal mines sprawl over vast land areas, including areas for excavation, dumps for extracted rock, and the support infrastructure of buildings, roads, and rail.

Deep and open pit mines remove huge amounts of rock, termed overburden, lying over the coal. Land over deep mines will sink, a process termed subsidence, physically changing a landscape. Underground coal fires in abandoned mines and refuse banks not only exacerbate subsidence, but they also release CO_2 and other air pollutants.

Miners dump large amounts of unwanted extracted rock, termed mine tailings, in abandoned parts of active mines or on the surface. Pyrite (FeS_2) usually comprises a significant fraction of the tailings. The reaction of water and pyrite produces sulfuric acid (H_2SO_4). In addition to being poisonous to plant and animal life, sulfuric acid mobilizes other toxic substances. The leaching of acids, trace metals, dissolved solids, and toxic organics produces a liquid known as acid mine drainage that can devastate surface waters. Selenium and cadmium often occur in high concentrations in tailings, so acid mine drainage can initiate the bioaccumulation of these trace metals in the surviving sections of the food chain.

Surface mining consumes vast tracts of land. Heavy machinery removes the upper layer of a landscape to expose relatively shallow coal seams, completely destroying the mined area. Although coal companies generally fill the overburden back into the mined area and replant it, strip-mined land can never completely recover its original characteristics. In addition, disturbance creates opportunities for invasive species to expand where perennial native plant species may have dominated. Rodents and other animals that adapt readily to human disturbance also take advantage of reclaimed areas.

Coal mines often impound surface streams to satisfy the large water needs of mine operations. These needs include water cannon drilling, transport by slurry, fugitive dust spraying, coal washing, and size sorting.

Mines crush and screen coal for uniform size, then wash and dry the coal for open-air storage. Fugitive emissions from these processes consist of particulates that coat any exposed surface, blocking photosynthetically active radiation from

plants, contaminating food and water sources for animals, and acidifying affected soil. Leaching of toxic substances from coal storage piles adds to the pollution of surface waters. Rail transport provides the most cost-effective means of moving the bulky commodity of coal, but fugitive emissions from unit trains increase particulate concentrations in rail corridors. To save money on rail transport, many utilities build electric power plants next to the mine and transmit the electricity by wire. In certain regions, this shifts the pollutant load from urban to rural areas.

Most coal goes to generate electricity. A conventional power plant burns coal in a boiler to boil water that circulates through a closed loop of pipes. The steam from the boiler enters a turbine to turn huge fans that power an electric generator that converts kinetic energy to electric energy. The generator exploits a physical principal in which the movement of a conductor across a magnetic field creates electric current in the conductor. In a coal-fired electric generator, the conductor consists of stationary coils of wire surrounding a magnet on a shaft rotated by the turbine fan. Much of the steam that moves through the fan transfers its heat energy to the kinetic energy of the fan, causing the steam to condense back to water. A condenser will then allow heat to transfer from any steam that continues past the turbine to an external supply of cold water. The water in the internal loop from the condenser returns back to the boiler to enter the steam cycle again.

Coal combustion releases CO, CO_2, SO_2, NO_x, particulates, fly ash, arsenic, cadmium, chromium, mercury, and selenium. Approximately 40% of anthropogenic CO_2 emissions come from burning coal, while coal burning produces approximately 80% of human SO_2 emissions. Consequently, greenhouse gases and acid precipitation may be the agents of coal's most extensive environmental impacts.

Slag remaining from burned coal contains high amounts of trace metals, especially cadmium and mercury. In addition, sludge from flue gas desulfurization units, the pollution control devices known as scrubbers, contains trace metals and toxic organics. Disposal of sludge can pollute soil and water.

Internal steam turbine water is the working fluid circulating from the boiler to the turbine to the condenser and back to the boiler. Cooling water is the medium that draws heat from the internal steam turbine water. In most conventional coal-fired power plants, the internal steam turbine water remains separate from power plant cooling water. A typical condenser consists of copper coils, carrying cooling water, that pass through larger structures carrying the internal steam turbine water. Heat passes from the steam turbine water through the walls of the copper coils into the cooling water.

A 1 GW coal-fired power plant typically uses 4 million m^3 water per day for all operations, mostly for cooling. Consequently, electric utilities often build plants next to natural water bodies. Power plants use freshwater because of the corrosive effects of salt water. Water withdrawals alter the hydrology of a watershed and change water levels, surface area of mudflats, surface area of wetlands, and other important habitat characteristics that can strand hydrophilic plant species such as willows (*Salix* spp.) and harm fish and shorebird populations. Impingement on intake screens kills significant numbers of fish and other aquatic species. Organisms that get through the screens can get entrained in the condenser, causing even greater mortality. The stress that any surviving organisms undergo reduces fitness considerably.

Industrial electric power plants (coal, oil, nuclear) generate three-quarters of the waste heat dumped into U.S. surface waters and into the atmosphere above the U.S. Once-through systems dump waste heat directly into local waters. Cooling towers systems dump waste heat into the atmosphere, condensing steam from the air. Cooling ponds provide a buffer for releasing some of the heat from cooling water into the atmosphere, reducing the temperature of cooling water before it enters surface waters.

Thermal discharges into freshwater and coastal zones cause many negative impacts on aquatic species:

1. Direct lethality to fish and crustaceans at water temperatures $\geq 35\ ^{\circ}C$.
2. Decrease in dissolved oxygen.
3. Increase in metabolic rates and nutrition needs for fish and changes in nutrition requirements for other taxa.
4. Displacement of diatoms by green and blue-green algae.
5. Inhibition of vertical migration by zooplankton.
6. Thermal plumes blockage of migratory fish movement.
7. Avoidance of warm areas by migratory waterfowl.
8. Early emergence of aquatic insect adult life stages into inhospitable environmental conditions.
9. Copper contamination from condenser coils.

Long-range transmission of electricity occurs across high-voltage lines strung on metal towers up to 60 m tall. The network of high-voltage electricity lines (230, 345, 500, 765 kV) in the U.S. stretches across 250,000 km and occupies more than 13,000 km^2 of land (U.S. Department of Energy, 2010). Clear-cutting of forest to create transmission corridors 30–60 m wide destroys vegetation. Maintenance periodically disturbs the corridors, favoring ruderal plant species and animals that adapt readily to human disturbance. Herbicides used for periodic clearing can kill pollinating insects and birds. Transmission line corridors fragment habitat and increase the area of habitat susceptible to edge effects while providing avenues for the dispersion of invasive weeds. Cleared areas can also block migrating land animals.

Short-range electricity transmission occurs across low-voltage lines strung on wood, metal, or concrete poles generally 5 m tall. Harvesting wood poles can produce all the potential biodiversity impacts of commercial logging, timber plantations, and milling. In many countries, utilities treat the wood with creosote to guard against insects and weather. Creosote, a by-product of crude oil refining, contains significant amounts of toxic organics that can leach and contaminate surface waters.

The materials and energy used to build the massive infrastructure of the coal fuel cycle produce wide-ranging environmental impacts. Because most coal goes to electricity generation, the end-uses of coal produce the environmental impacts associated with air conditioning, commercial machinery, residential appliances, and other electric devices. Owing to massive economic growth, P.R. China switched from a net exporter to a net importer of coal in 2009. Australia, Canada, Columbia, Indonesia, South Africa, and the U.S. export substantial amounts of coal to P.R. China. Environmental

impacts of coal mining and processing occur in the exporting countries while environmental impacts of coal burning occur in P.R. China and around the entire world due to the emission of greenhouse gases.

Nuclear Fission

Nuclear fission is the splitting of high-molecular-weight elements to release energy held by protons and neutrons in the nucleus of the atom. Uranium and plutonium are the elements that provide the most effective yield from fission. A fission reaction produces energy in the form of light, heat, motion of the fission products, and radiation. Radiation consists of kinetic energy of small molecules and atomic particles and electromagnetic energy of photons traveling at certain frequencies. When radiation passes through living tissue, the particles or photons impart their energy to atoms and molecules in the tissue, disrupting molecular and atomic structures.

Fission products emit radiation until they reach a stable atomic state. While the half-life of strontium-90 is 29 years and the half-life of cesium-137 is 30 years, plutonium-239 decays with a half-life of 25,000 years, and a quantity of iodine-129 will decay to half of its mass only after 17 million years. The similarity of the atomic structure of strontium to calcium increases the uptake of strontium by animals and its incorporation into bones.

Nuclear fission plants require highly processed uranium fuel. Uranium rests in geologic strata in the minerals uraninite and pitchblende. The isotope uranium-238 accounts for over 99% of the uranium in nature, but nuclear fission fuel requires the uranium-235 isotope. A standard 1 GW nuclear fission plant requires 150,000 Mt uranium-containing ore to fabricate enough fuel for 1 year. Milling, roasting, and acid leaching of the ore produces 150 Mt uranium oxide (U_3O_8) in a granular form called yellowcake as well as substantial amounts of ore tailings and chemical effluents. Fluorination of the yellowcake produces 188 Mt of uranium hexafluoride (UF_6).

Processors use one of three methods – gaseous diffusion, gas centrifuge separation, or liquid thermal diffusion – to divide UF_6 into separate fractions, one of which is enriched in a higher concentration of uranium-235 than found in nature. Nuclear fission for electricity generation requires enrichment to 2–3% uranium-235. Continuation of the process produces material enriched to 97–99% uranium for nuclear warheads.

The original 150,000 Mt of ore for a standard 1 GW nuclear fission plant has thus yielded 31 Mt UF_6 enriched in uranium-235. Fuel fabrication then produces 30 Mt of uranium dioxide (UO_2) pellets for use in the nuclear reactor core.

Mining and milling of uranium ore creates most of the same environmental problems described for deep coal mines and coal processing. Uranium conversion, enrichment, and fuel fabrication require toxic chemicals, including fluorine gas, which is lethal on contact to animals, damages vegetation, and forms toxic by-products.

In December 2008, 438 nuclear fission plants with a combined rated capacity of 372 GW were operating in 31 countries (IAEA, 2009). Among these, the U.S. operates 104 nuclear fission plants with a combined rated capacity of 101 GW.

Nuclear plants generate electricity in a steam cycle similar to the system in coal plants, except that nuclear fission is the source of heat for the boiler. Higher operating temperatures require more cooling water than a coal-fired plant of the same electricity generation capacity. A 1-GW nuclear fission plant requires 6 million m^3 of cooling water each day, so the effects of water intake and thermal discharge described in the section on coal are all greater with nuclear plants.

Because nuclear plants involve fossil fuel combustion only in construction and in support vehicles, they produce few direct air emissions. Nuclear plants, however, produce wastes that can remain radioactive for millions of years. Low-level wastes include reactor containment water, worker clothing, exposed tools, and plant fixtures irradiated for limited periods of time. High-level wastes consist of spent fuel and the fuel rods in which they are encased. The U.S. has not constructed a permanent repository for high-level wastes, which nuclear plants continue to store on site.

The greatest single release of radiation from a nuclear fission power plant came from the Chernobyl Unit 4 accident on 26 April 1986 in an area of the Republic of Ukraine that was in the former Soviet Union. Operator error combined with design aspects of the RBMK graphite-moderated reactor generated a nearly instantaneous catastrophic increase of thermal power and a steam explosion. The explosion destroyed the reactor and released 3% of the reactor fuel and up to 60% of the volatile products in the reactor core, mainly iodine-131, cesium-134, and cesium-137. The accident deposited radioactive fallout over the entire Northern Hemisphere.

Twenty-eight people died from acute radiation doses, while thyroid cancer has been found in at least 4000 people due to ingestion of iodine-131 (Chernobyl Forum, 2005). The Soviet government evacuated 116,000 people from a zone of 30 km radius and constructed a cement sarcophagus to contain the remains of the reactor. Increased risks of thyroid cancer and leukemia are possible for the 5 million people who lived in the contaminated parts of Belarus, Russia, and Ukraine.

Lethal radiation killed many conifers and small mammals within 10 km of the accident in the first few weeks. Radioactivity remains in trees and soil in the form of cesium-137. Grass, mushrooms, and berries continue to incorporate the isotope, perpetuating a source of exposure for people and wildlife that feed on contaminated plants. Atmospheric fallout from Chernobyl may also cause genetic abnormalities in the long term.

Hydroelectric

Hydroelectric systems harness the potential energy held by an elevated mass. The potential energy of water will convert to increased kinetic energy of the water when it runs to a lower elevation. A dam concentrates the difference in elevation, termed hydraulic head, in a spillway equipped with a turbine and an electric generator. The electricity immediately enters the electric grid. In this manner, a hydroelectric plant will generate electricity with few direct air emissions and little thermal discharge. The principal effects of hydroelectric plants come from the considerable alteration of topography and flow of a watercourse and the partial inundation of its watershed.

Besides the forced removal of people and inundation of homes, hydroelectric plants also cause significant ecological changes.

Over 45,000 large dams block rivers around the world, creating reservoirs that inundate up to 1 million km^2 (World Commission on Dams, 2000). In 2010, P.R. China completed filling the reservoir of the Three Gorges Dam, the hydroelectric plant with the highest generation capacity in the world (18.2 GW). The project has inundated 630 km^2 along 600 km of the world's third longest river, the Yangtze, and displaced 1.2 million people. In 1965, the Akosombo Dam on the Volta River in Ghana created Lake Volta, at 8500 km^2 the largest impoundment in the world.

Inundation of formerly dry land destroys vegetation and decreases the area of terrestrial habitat. Inundated forests represent ecosystem services forfeited and biomass wasted. A dam blocks nutrient-rich sediment that a river system would have deposited in floodplains, wetlands, and the outlet delta. Sediment fills the reservoir, impairs electricity generation, and traps rich organic carbon and other nutrients at the bottom of the reservoir. At the outlet delta, dams alter bay and estuary topography, reduce the area of mudflats and wetlands, and decrease the upwelling of nutrient-rich water. The Aswan High Dam in Egypt blocks 98% of the 120 million tons of sediment that the Nile River had carried each year, formerly depositing 10 million tons on the floodplain and delta. Consequently, soil depth has thinned and agricultural production has declined in the Nile Valley. Blockage of sediment and freshwater by the Akosombo Dam in Ghana has caused the decline of clams in the Volta estuary and barracuda (*Sphyraena barracuda*) in the Gulf of Guinea. The reduction of estuaries and mudflats at a dammed river's outlet delta renders the coast more susceptible to tidal erosion.

Utilities start and stop the flow of water based on electricity use and operational objectives. One operational objective, smoothing out extremes caused by natural flood regimes, changes the meandering and other channeling processes of a river. After the Glen Canyon Dam ended spring floods in the near downstream section of the Colorado River, sandbar erosion increased because the river did not flow fast or deep enough to move the amount of silt required for extensive sandbar formation. The resulting disappearance of some riparian tree species has led to the decline of southwestern willow flycatcher (*Empidonax traillii*) and other birds. To mitigate the problem, the U.S. Department of the Interior began to stage controlled floods.

For some dammed rivers, the flow of water unburdened by silt can deepen the riverbed. The depletion of riverbed gravel can harm any species of fish, insect, mollusk, or crustacean that requires a gravel stream bottom to spawn. Many insect, amphibian, and fish species also use gravel areas for habitat or protection.

The depth of a reservoir will often keep water at a temperature lower than in the native river. For example, the Glen Canyon Dam changed the water temperature in the near downstream section of the Colorado River from a range of 0–27 °C to a relatively constant 8 °C. This has contributed to the extinction of Colorado squawfish (*Ptychocheitus lucius*), roundtail chub (*Gila robusta*), and bonytail chub (*Gila elegans*) and the endangerment of five other fish species. Whereas release water is clear, reservoir water often becomes slightly turbid and eutrophic.

Impacts of dams on anadromous fish relate to the migratory behavior and timing of the life cycles of these unique species. Dams erect a hazard for the downstream migration of young fish and block the upstream migration of adults. Moreover, salinity and temperature adaptations in anadromous fish occur on a precise schedule, making long delays lethal. Also, disoriented and fatigued fish more easily fall prey to predation. Despite the deployment of extraordinary means to facilitate fish migration, including fish ladders, elevators, and trap-and-haul trucking, dams have eliminated anadromous fish from many rivers. Runs of Atlantic salmon (*Salmo salar*) and American shad (*Alosa sapidissima*) have disappeared from many rivers in the northeastern U.S. In the Columbia River Basin in the northwestern U.S., overfishing, pesticide runoff, and hydroelectric plants have endangered populations of Snake River sockeye salmon (*Oncorhynchus nerka*) and Snake River chinook salmon (*Oncorhynchus tshawytscha*).

Exotic fish species introduced into reservoirs for sport fishing often outcompete native species. In the 1350 km^2 reservoir on the Brazil–Paraguay border behind the Itaipu Dam, the hydroelectric plant with the second highest generation capacity in the world (12.6 GW), a nonnative species, curvina (*Plagioscyon squamosissimus*), has become the second most numerous fish species.

Polychlorinated biphenyls (PCBs) from electric transformers and oil leaking from machinery comprise the worst direct industrial pollution from dams. These substances and toxic organic chemicals accumulate in sediments and magnify up the food chain. Impingement of aquatic organisms on intake screens and entrainment through turbines kills many individuals and causes stress and injury in survivors.

Materials and energy used to construct the massive infrastructure of the hydroelectric energy cycle produce wide-ranging environmental impacts. In addition, end-uses of hydroelectricity will produce the environmental impacts associated with air conditioning, lighting, commercial production machinery, residential appliances, and other electric devices. Because smelting aluminum from bauxite ore requires large amounts of electricity, aluminum smelting comprises an end-use closely tied to the hydroelectric option. Air emissions from smelters include CO, CO_2, particulates, NO_x, and trace metals. Major water pollutants include trace metals and sulfates.

Renewable Energy Technologies

Renewable energy includes those forms of energy whose transformation does not necessarily consume the ultimate source of the energy, harnessing instead solar radiation, wind, the motion of water, or geologic heat. This section covers renewable energy technologies, including solar heating, solar thermal electric, solar photovoltaics, electric wind turbines, geothermal, biomass-to-energy (electricity and heat), and biofuels. These are renewable energy systems that depend on complex technology, so they are forms of industrial energy. The section on traditional energy covers the simpler forms of renewable energy – firewood, charcoal, and small water and windmills.

Forms of renewable energy generally require specific site characteristics, exhibit high temporal variability in availability, and capture diffuse flows. Except for biomass-to-energy (electricity and heat), renewable energy technologies do not involve combustion, so they do not directly produce much air pollution. The principal environmental impacts come from fabrication, installation, and maintenance of renewable energy devices.

Solar energy systems fall into two categories: passive and active. Passive systems consist of architectural forms that more effectively follow the diurnal and seasonal patterns of sunlight for the efficient heating and cooling of a building. Passive systems use the natural phenomena of radiation and convection. Active systems use moving devices to achieve heat transfer. Solar heating systems include passive solar building designs and simple active systems that use pipes or other collectors to heat water for residential or commercial use. Solar heating systems are nearly environmentally benign.

Complex active solar systems include solar thermal electric and photovoltaic systems. Solar thermal electric systems, also known as concentrating solar power systems, use arrays of reflective parabolic or trough collectors to focus sunlight on a water boiler to power a turbine to generate electricity. Photovoltaics are solid-state devices in which photons stimulate the emission of electrons and semiconductor materials channel the electrons for collection. They are flat rectangular chips placed into long flat arrays. Photovoltaics directly convert sunlight to electricity with no moving parts, except for motors that move the arrays to track the sun. These systems use water to wash photovoltaic surfaces.

The two types of complex active solar systems require vast areas of land for the solar collectors or the photovoltaic arrays. In 2010, the U.S. Government permitted 3.6 GW of solar thermal electric and photovoltaic plants in California and Nevada and advanced plans for large solar plants throughout the southwestern U.S. Many of these projects will occupy and possibly damage natural ecosystems and habitat for threatened species such as desert tortoise (*Gopherus agassizii*). Because solar is most economically feasible in sunny areas, solar sites are generally arid and water is scare. Water withdrawals for turbines and for washing collectors and arrays contribute to water flow reductions that can harm aquatic ecosystems.

Concentrating solar power systems are industrial scale plants that alter extensive areas of natural land far from energy users. In contrast, decentralized photovoltaic arrays above parking lots in urban areas have emerged in California as a more environmentally sound and economically feasible energy solution. Numerous colleges and schools, mainly in the San Francisco Bay Area, have erected solar arrays above parking lots. These projects use land that has already been urbanized and is close to the energy users, reducing transmission infrastructure and electricity losses. In some cases, solar arrays can provide for all school electricity uses. These systems demonstrate the solar potential of urban areas, especially parking lots and building roofs.

Fabrication of photovoltaic cells produces toxic environmental impacts. Mining the quartz base material of a photovoltaic cell produces many of the same impacts on aquatic and terrestrial biodiversity described for coal mining. Production of metallurgical grade silicon requires the refining of the quartz to 99% purity at 3000 °C in an electric arc furnace. Production of semiconductor-grade silicon takes place in a fluidized bed reaction of the silicon with hydrochloric acid. The process produces semiconductor-grade polycrystalline silicon by electrical heating of the semiconductor-grade silicon at 1000 °C and vapor deposition onto a silicone substrate. Remelting the polycrystalline silicon produces a form that can grow into crystals. Precision saws cut these crystals into wafers 0.5 mm thick. Robotic devices wire and encapsulate the cells in glass 3 mm thick.

Trace metals are used to dope the semiconductor material for the principal types of photovoltaic cells: cadmium telluride, copper indium diselenide, gallium arsenide, and gallium indium phosphide. Trace metals in these compounds, chlorinated organic solvents, and phosgene gas all produce hazardous air, water, and solid wastes that can be lethal on contact or carcinogenic in small doses.

Biomass-to-energy plants are facilities that convert waste wood from lumber mills, wood from dedicated plantations, agricultural waste, or municipal waste to electricity either through direct combustion of biomass in low-pressure boilers to power electric steam turbines or through gasification of organic matter into methane to power natural gas electric turbines. Biomass-to-energy plants often employ cogeneration to provide process heat for an adjacent industrial facility. The section on coal details the negative effects of electric turbines and condensers that are also components of biomass-to-energy systems.

Utilities in Brazil generate electricity and cogenerate heat from the organic wastes (bagasse) from sugar cane and orange juice processing. Landfills around the world use networks of collection pipes to capture methane released by decomposing municipal waste and send the methane to natural gas electric turbines. Most biomass-to-energy plants use material that otherwise would be wasted and discarded in landfills. The biomass would otherwise oxidize and release its energy as waste heat. The productive reuse of waste products in biomass-to-energy production therefore reduces the environmental impact of the disposal of biomass waste.

Experimental schemes, however, that grow wood or agricultural crops specifically to burn for electricity generation, can produce negative environmental impacts if the schemes use conventional fossil-fuel-powered machinery and energy-intense fertilizers. Indeed, energy-intensive methods could expend more energy in the cultivation and processing of energy crops than the amount of energy generated by burning the biomass to generate electricity. The principal species used by such schemes include short rotation trees such as poplars and aspen (*Populus* spp.), grown at densities of 1600–5000 trees ha^{-1}, and switchgrass (*Panicum virgatum*). An energy crop can damage natural ecosystems if it is grown in monoculture or if vegetation is clear-cut to prepare for the crop. If previous land uses were less environmentally sound than the energy crop, then the crop could mitigate the impacts of the previous land use.

Conversion of biomass into alcohol fuels, known as biofuels, also requires the dedicated growing of energy crops, primarily corn, palm trees (for palm oil), soybeans, and sugar cane. Fermentation of cellulose and other complex carbohydrates produces ethanol, which certain engines burn

straight or mixed with gasoline. Otto Cycle engines burn neat ethanol, a mixture of 96% ethanol and 4% water. Modified conventional automobile engines can burn gasohol, a mixture of 78% gasoline and 22% ethanol. The U.S. produces ethanol mainly from corn. Brazil produces enough ethanol from sugar cane to provide for 10% of the country's energy use.

Biofuel production that uses conventional fossil-fuel-powered machinery and energy-intense fertilizers can expend more energy in the cultivation and processing of energy crops than the amount of energy contained in the final product. Moreover, combustion of fossil fuels and energy use for inorganic fertilizer production generate the negative impacts identified in previous sections. Because the combustion of ethanol mainly produces CO_2 and water and much smaller amounts of hydrocarbons and NO_x than gasoline combustion, ethanol produced by nonfossil fuel-intense processes can mitigate the most harmful direct effects of oil production and use.

Conversion of the principal areas of rainforest, peatlands, and grasslands to produce biofuels in Brazil, Southeast Asia, and the U.S. could potentially release 20–400 times more CO_2 than the displaced fossil fuels in the next half-century (Fargione *et al.*, 2008). In contrast, biofuel from waste biomass or biomass grown on abandoned agricultural lands planted with perennials can produce less CO_2 than displaced fossil fuels.

Since the 1980s, electric utilities have used wind turbines to power electric generators. Early electric wind turbines were typically 10 m tall with rotors diameters of 5–10 m and electricity generation ratings of 100–300 kW per turbine. Wind turbines are now typically 50–100 m tall with rotor diameters of 30–100 m and electricity generation ratings up to 2.5 MW per turbine. Offshore wind turbines are generally larger, to take advantage of steady winds and economies of scale.

Wind farms of hundreds of turbines occupy considerable land and offshore areas. The first large-scale wind farms operated in the Altamont Pass in the San Francisco Bay Area, California, the Tehachapi Pass in Southern California, in the Netherlands, and Denmark. They have now become widespread throughout the world.

Wind farms fragment terrestrial habitats and access road networks cause soil erosion. Wind turbines kill substantial numbers of birds and bats. Rotors produce substantial noise. Offshore wind turbines can act as artificial reefs, potentially changing local marine food webs. Electromagnetic fields from turbines and underwater cables can disrupt marine mammal navigation abilities.

Wind turbines could potentially produce all of the world's electricity. Wind turbines operating at only 20% of their rated capacity (due to diurnal and seasonal wind cycles, some turbines only turn 20% of the time) on available nonforest, ice-free, and nonurban areas could potentially generate electricity at a rate of 100 TW for the world and 8 TW for the contiguous U.S. (Lu *et al.*, 2009).

Geothermal devices capture heat or cold stored underground and directly use the heat or cold for different purposes on the surface or process the heat through turbines to produce electricity. The U.S., P.R. China, and Sweden led the world in installed capacity of geothermal direct use systems in 2010 while the U.S., Philippines, and Indonesia led in installed capacity of geothermal electric systems (REN21, 2010).

Geothermal direct heat systems pipe water from underground hot geologic formations to the surface to heat greenhouses, buildings, and fish farm ponds or to heat streets and sidewalks from below to melt snow and ice on the surface. Geothermal heating and cooling systems pump water or another working fluid through a closed loop of pipes and a heat exchanger in contact with shallow underground areas or water bodies that maintain a relatively constant range of temperatures. In temperate zone residential geothermal systems, underground soil and rock are cooler than the surface in summer and warmer than the surface in summer. The heat exchanger cools the house in summer and transfers heat from the ground to the house in winter.

In the largest geothermal cold storage system in the world, the 2 GW capacity Alderney 5 project in Halifax, Nova Scotia, pipes take cold harbor water for 7 months and stores it in a field of 200 m deep boreholes that cool bedrock under a ferry terminal parking lot. A heat exchanger uses the stored water to cool a commercial and industrial complex in summer.

Geothermal electric systems capture the heat of hot geologic formations from a system of pipes sunk down into hot strata. These systems capture deep hot water or inject water so that it boils on contact with hot rocks and process the heat through turbines to produce electricity. These systems mobilize trace metals contained in certain geologic formations and release H_2S gas from geothermal deposits.

Biodiversity Impacts of Traditional Energy

The most important sources of traditional energy are firewood and charcoal, which is produced from firewood. Local people harvest firewood by coppicing shrubs and cutting branches from mature trees. Rural people in Africa, Asia, and Latin America generally avoid felling whole trees for domestic firewood because of the time and labor required to cut the tree and split the logs.

In semiarid areas of Africa, women prefer the straight, moderately sized branches that coppiced shrubs produce. Rural people go out in the dry season and coppice (cut at the base) small shrubs (family *Combretaceae*). Women carry branches back to the village and let them dry. Just before the first rains, men and women cut a store of firewood for the rainy season. This serves, first, to avoid cutting wood that is wet and difficult to burn and, second, to complete a time-consuming and strenuous chore before the exhausting and rushed rainy season.

Coppiced shrubs resprout in the rainy season and regrow branches in a year. When shrubs become scarce, women pull branches from adult trees, sometimes using long-handled hooks. This harms the growth potential of a tree by removing shoot apical meristem tissues and can yield thorny branches that are difficult to handle. When branches are depleted, women in semiarid areas of Africa fall back on noxious, dead stalks of spurge (family *Euphorbiaciae*). The last resort is animal dung.

While women carry firewood for rural use, rural people also load beasts of burden and carts to transport wood for sale in urban areas. In this way, a town or city can generate impacts far beyond its borders.

Wood contains energy at a density of approximately 15 GJ t^{-1}, one-third the energy density of oil. The relatively low-energy density of wood renders its transport onerous relative to the energy gained. Conversion of firewood to charcoal creates a product with double the energy per unit mass, but emits up to two-thirds of the energy contained in the original wood as waste heat. Charcoal makers cut down live and dead trees, targeting sturdy tree trunks. They pile the wood, cover it with soil to form a kiln 1–4 m in height and 1–3 m in diameter, and ignite a slow burn. Over 3–6 days, the wood converts to charcoal by partially anaerobic pyrolysis.

If wood harvesting for firewood or charcoal exceeds the natural regeneration of shrubs and trees, then wood harvesting will reduce vegetative cover. Reduction of vegetative cover and conversion of forested or wooded land to savanna or grassland comprise the principal potential impacts of traditional energy on natural ecosystems. People's preferences for certain species for firewood can also create a risk of overharvesting preferred species.

In some areas of Africa, Asia, and Latin America, international development agencies have funded the massive plantation of nonnative tree species for firewood production. Plantations that replace native forest or woodland eliminate natural ecosystems. In general, the biological and structural diversity of natural forests makes an area more resilient to fire, wind, and insect disturbances.

Natural regeneration of local species can restore native forest cover in ecosystems changed by overharvesting. Natural regeneration is a traditional practice in which farmers and herders protect and promote the growth of young native trees. Traditionally, local people protect small trees that have germinated naturally or resprouted from roots, prune them to promote growth of the apical meristem and, if necessary, set a stake to straighten the small tree. In Africa, natural regeneration has expanded *Acacia albida* from an original restricted range along rivers in Southern Africa to an extensive range that reaches across the continent north to the Sahel. Natural regeneration requires no external inputs. It concerns species well known and appreciated by rural people. It focuses on young trees that have demonstrated their hardiness by surviving with no human caretaker, no watering, and no special treatment. Natural regeneration not only augments the supply of wood, poles, fruit, medicine, and other forest products, it puts trees where farmers and herders need them – in fields to maintain soil fertility and in pastures to provide forage.

Photosynthetic activity converts only a fraction of total available solar radiation to wood. Nevertheless, the inefficiency of human tools for the conversion of wood to heat and light renders human end-uses even more wasteful. **Table 5** shows this energy chain from sunshine to wood end-use in the West African Sahel.

Improved cook stoves with higher fuel efficiencies can help conserve vegetative cover in rural areas that depend on firewood. In many areas, women customarily cook with a kettle over an open fire. International development agencies have helped to develop stoves such as, in Senegal, the *ban ak suuf*, a horseshoe-shaped hearth constructed from local clay that provides an enclosed combustion space that more effectively channels heat to the cooking vessel. The *lorena* in Guatemala is another improved efficiency earthen stove. The *jiko* in Kenya and *sakkanal* in Senegal are enclosed metal or ceramic charcoal stoves that more effectively contain heat than traditional open charcoal burners.

Table 5 Energy chain from sunshine to wood end-use in the West African Sahel

Stage	*W ha^{-1}*
Insolation at ground	2,400,000
Net primary productivity	1700
Total wood production	120
Human wood energy use	210
Imported fossil fuels	93
Food consumption	53
Human wood energy end-use	13

Source: Gonzalez P (2001) Desertification and a shift of forest species in the West African Sahel. *Climate Research* 17: 217–228.

For centuries, society has channeled water to mill grain and captured wind to move sailing ships and pump water. Small water and windmills comprise a form of traditional energy that generally requires only local materials and expertise for building and maintenance. The simplest water mills are run-of-the-river systems in which a water wheel is placed in the current of a perennial stream or river. A water wheel typically turns a circular stone for the grinding of grain. More advanced systems use water channels, pipes, Pelton wheels, or other devices to increase rotation speeds enough to turn a turbine to produce electricity. Small hydropower is most common in mountainous countries like Perú. The simplest windmills have a fan with wood or metal blades erected on a tower 5–15 m tall. The fan axis rotates on a horizontal axis, lifting a vertical rod connected to a plunger in a pipe well dug down to the water table. The movement of the plunger lifts water to the surface. This type of windmill is most common in flat grassland regions. Windmills in the Netherlands pumped water from extensive inundated areas that the Dutch enclosed with dikes, dried, and used for agriculture.

Future Energy Paths

Human energy use directly alters biodiversity through changes in land use and through industrial pollution. Indirectly, human energy use is altering biodiversity through the emission of greenhouse gases that cause global climate change and through other broad impacts on the natural function of ecosystems. Not only does the direct processing of energy generate environmental impacts, but the end-uses that convenient energy forms make possible produce impacts at all scales: individual species, local sites, landscapes, continents, and the world.

Table 6 summarizes the major environmental impacts of human energy use on biodiversity. **Table 7** summarizes the extensive land requirements and CO_2 production of energy sources. Land use change for energy use destroys and fragments natural ecosystems. Globally, climate change caused by emissions of CO_2 and other greenhouse gases constitutes the most severe impact of fossil fuels, but nonfossil fuel sources also produce air and water pollution. No energy transformation system operates without negative environmental impacts, yet renewable sources generally restrict harmful effects to the

Table 6 Major sources of biodiversity impacts from human energy use

Impact	*Oil*	*Natural gas*	*Coal*	*Nuclear fission*	*Hydroelectric*	*Renewable technologies*	*Wood*
Terrestrial habitat destruction and fragmentation	Exploration, access roads, pipelines	Exploration, access roads, pipelines	Mining, electricity transmission lines	Mining, electricity transmission lines	Inundation of vast land areas, changes to river hydrology	Land requirement for solar collectors, wind turbines, crops for biofuels	Destruction or fragmentation of habitat from unsustainable harvesting
Water pollution and damage to aquatic species	Oil spills, drilling muds, soil eroded from pipeline corridors and roads	Natural gas liquids spills, soil runoff from pipeline corridors and roads	Acid mine drainage, water removal for mining and plant cooling, entrainment, impingement, thermal pollution	Acid mine drainage, water removal for mining and plant cooling, entrainment, impingement, thermal pollution	Complete alteration of habitat, barriers to migration, entrainment, impingement, thermal changes, PCBs from transformers	Toxics from photovoltaic production, water removal for cleaning solar collectors and irrigating biofuel crops, pesticide and fertilizer from production of biofuels	Soil runoff from overharvested areas and roads
Air pollution	CO_2, toxic organic compounds from refining	CO_2, flaring, volatilization of CH_4	CO_2, SO_2	Radiation, toxic halogenated compounds in fuel processing		CO_2 emissions from biofuels farm machinery	CO_2 from burning
Solid waste	Oil spills, sludge	Natural gas liquids spills, sludge	Mine tailings, sludge	Radioactive waste	PCBs from transformers	Toxics from photovoltaic production	
Major end-uses	Automobiles	Cooking, heating	Electricity	Electricity	Electricity, smelters	Electricity	Cooking, heating

capital formation stage and do not produce much pollution from operations.

Environmental impact is equivalent to the multiplicative effect of population, affluence, and technology (Holdren and Ehlich, 1974):

$$\text{Environmental impact} = \text{Population} \times \frac{\text{Resource use}}{\text{Person}} \times \frac{\text{Environmental impact}}{\text{Resource use}}$$

Because the environmental impact of human energy use is proportional to the rate of energy use, and energy use is proportional to economic production, the IPAT (impact = population × affluence × technology) equation for energy becomes:

$$\text{Environmental impact} \propto \text{Human energy use} = \text{Population} \times \frac{\text{Economic production}}{\text{Person}} \times \frac{\text{Energy use}}{\text{Economic production}}$$

Economic production per person, often expressed as dollar of GDP per person, indicates a society's level of material affluence, although GDP does not price ecosystem services such as clean water and natural habitat at their full value and values activities with negative impacts, such as military spending, as positive. Energy use per unit of economic production, expressed as watts per dollar of GDP, indicates a society's level of technological efficiency.

This relationship highlights the leverage of energy conservation and efficiency in reducing the environmental impact of energy use. Improvements in energy efficiency reduced the energy intensity of economic activity in the U.S. by nearly one-third between 1975 and 1995 (U.S. Department of Energy, 2010). Energy conservation and currently available efficiency improvements could substantially reduce industrial energy use and greenhouse gas emissions. Energy efficient vehicles and appliances, installation of insulation and other residential weatherization features, and simple changes in driving behavior such as slower acceleration and reduced highway speeds could reduce U.S. greenhouse gas emissions 7% (Dietz *et al.*, 2009). In addition to those conservation and efficiency actions, walking or using public transit instead of cars, substitution of renewable energy sources for nonrenewable sources, and other currently available improvements could meet global energy uses and stabilize greenhouse gas emissions by the mid-twenty-first century (Pacala and Socolow, 2004).

If the world does not adopt such improvements, the historical path of industrialization has only left costly and environmentally disruptive energy sources. The earliest exploitation of fossil fuels depleted the most convenient oil and gas deposits. This phenomenon explains why fossil fuel production over time follows the bell-shaped Hubbert Curve. Deposits are now more remote, deeper underground and underwater, and dispersed. Furthermore, low-cost energy has shaped the expectations of people around the world for inexpensive on-demand energy services. Societies subsidize the provision of convenient energy through infrastructure support to energy industries, tax breaks to oil drillers, preferential treatment to automobile companies, and other schemes.

Depletion of nonrenewable resources (**Table 8**) and other environmental and social constraints hobble most future energy options: oil and gas reserves will only last decades; coal burning releases the principal cause of global warming, CO_2;

Table 7 Land requirements and carbon dioxide emissions for electric generation

	Land required ($ha\ MW^{-1}$)	*CO_2 emissions* ($t\ GW^{-1}\ h^{-1}$)
Coal	0.8–8.0	900
Oil	0.3–0.8	850
Natural gas turbine	0.3–0.8	560
Solar photovoltaics	3–7	50
Wood and other biomass	150–300	40
Solar thermal electric	1–4	40
Nuclear	0.8–1.0	30
Hydroelectric	2–1000	20
Geothermal	0.1–0.3	14
Wind electric	0.4–1.7	11

Source: Data from US OTA, 1995; Cho A (2010) Energy's tricky tradeoffs. Science 239: 786–787.

Table 8 Remaining world energy reserves 2000

Nonrenewable energy stocks	*Reserve (TW year)*	*Time remaining at 2008 rates (years)*	*Carbon emissions (Gt)*
Uranium	80	80	20
Oil	220	40	1600
Natural gas	230	60	1100
Coal	1100	250	8600

Renewable power flows	*Potential (TW)*	*Time remaining*	*Carbon emissions ($Gt\ year^{-1}$)*
Hydroelectric	1.6	Lifetime of sun	0.3
Geothermal	4	Lifetime of sun	0.5
Wood and other biomass	9	Lifetime of sun	3
Wind electric	20	Lifetime of sun	2
Solar electric	50–1600	Lifetime of sun	20–600

Source: Data from Cho A (2010) Energy's tricky tradeoffs. Science 239: 786–787; UNDP, 2000.

biomass energy requires vast amounts of land; a small number of exploitable sites limits the potential for hydroelectric and wind power; and health and safety problems prevent expansion of nuclear energy.

Although engineers have placed enormous effort into the development of technologies such as hydrogen cars, fuel cells, and nuclear fusion, exotic devices will not resolve our energy problems if they depend on nonrenewable fuels. Instead, energy efficiency, conservation, and renewable energy offer a sustainable future energy path for the world. For this path, the goal is not the acquisition of energy stocks and devices, but the provision of services and end-uses. In effect, we don't require light bulbs, we need illumination.

See also: Air Pollution. Economic Growth and the Environment. Greenhouse Effect. Pollution, Overview

References

Boden TA, Marland G, and Andres RJ (2010) *Global, Regional, and National Fossil-Fuel CO_2 Emissions*. Oak Ridge, TN: U.S. Department of Energy, doi:10.3334/CDIAC/00001_V2010

BP (2010) *BP Statistical Review of World Energy 2010*. London: BP.

Chernobyl Forum (2005) *Chernobyl's Legacy: Health, Environmental and Socio-economic Impacts and Recommendations to the Governments of Belarus, the Russian Federation, and Ukraine*. Vienna: International Atomic Energy Agency.

Cho A (2010) Energy's tricky tradeoffs. *Science* 239: 786–787.

Crone TJ and Tolstoy M (2010) Magnitude of the 2010 Gulf of Mexico oil leak. *Science* 330: 634.

Dietz T, Gardner GT, Gilligan J, Stern PC, and Vandenbergh MP (2009) Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *Proceedings of the National Academy of Sciences of the USA* 106: 18452–18456.

Elvidge CD, Ziskin D, Baugh KE, *et al*. (2009) A fifteen year record of global natural gas flaring derived from satellite data. *Energies* 2: 595–622.

Etkin, DS (1999). Historical overview of oil spills from all sources (1960–1998). *International Oil Spill Conference*, paper no. 169.

Fargione J, Hill J, Tilman D, Polasky S, and Hawthorne P (2008) Land clearing and the biofuel carbon debt. *Science* 319: 1235–1238.

Federal Interagency Solutions Group (2010) *Oil Budget Calculator – Deepwater Horizon*. Washington, DC: National Oceanic and Atmospheric Administration.

Food and Agriculture Organization (FAO) of the United Nations (2009) *State of the World's Forests 2009*. Rome: FAO.

Gonzalez P (2001) Desertification and a shift of forest species in the West African Sahel. *Climate Research* 17: 217–228.

Holdren JP and Ehlrich PR (1974) Human population and the global environment. *American Scientist* 62: 282–292.

International Atomic Energy Agency (IAEA) (2009) *Nuclear Power Reactors in the World*. Vienna: IAEA.

International Energy Agency (IEA) (2010) *Key World Energy Statistics 2010*. Paris: IEA.

International Tanker Owners Pollution Federation (ITOPF) (2010) *Oil Tanker Spill Statistics: 2009*. London: ITOPF.

Lu X, McElroy MB, and Kiviluoma J (2009) Global potential for wind-generated electricity. *Proceedings of the National Academy of Sciences of the USA* 106: 10933–10938.

National Research Council (NRC) (2003) *Oil in the Sea III: Inputs, Fates, and Effects*. Washington, DC: National Academy of Sciences of the USA.

Pacala S and Socolow R (2004) Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. *Science* 305: 968–972.

Renewable Energy Policy Network for the 21st Century (REN21) (2010) *Renewables 2010 Global Status Report*. Paris: REN21.

United Nations Development Programme (UNDP) (2000) *World Energy Assessment*. New York: UNDP.

U.S. Department of Energy (2010) *Annual Energy Review 2009*. Washington, DC: U.S. Department of Energy.

U.S. Office of Technology Assessment (OTA) (1995) *Renewing Our Energy Future*. Washington, DC: U.S. Government Printing Office.

World Commission on Dams (2000) *Dams and Development: A New Framework for Decision-Making*. London: Earthscan.

World Energy Council (WEC) and International Institute for Applied Systems Analysis (IIASA) (1995) *Global Energy Perspectives to 2050 and Beyond*. London: WEC.

Environmental Ethics

Philip J Cafaro, Colorado State University, Fort Collins, CO, USA
Richard B Primack, Boston University, Boston, MA, USA

Published by Elsevier Inc.

Glossary

Anthropocentrism Position that only human beings have moral worth or intrinsic value.
Biocentrism Position that all living beings have moral worth or intrinsic value.
Deep ecology Activist philosophy that advocates radical personal and political change to protect wild nature.
Environmental ethics Philosophical discipline that specifies proper human relationships to the natural world.
Ethical holism Position that complex aggregates, such as species, ecosystems, or human groups have intrinsic value.
Humanism Environmental philosophy devoted to maximizing human wealth, flourishing, power, and numbers.
Instrumental value Value of something relative to human interests or desires.
Intrinsic value Value of something independent of its value to people.
Rights Justified claims that people respect a being's existence and important interests, and act accordingly.
Stewardship God-given responsibility to protect natural resources and preserve biological diversity.

Duties to Protect Biodiversity: Intrinsic Value Arguments

From Economics to Ethics

Biodiversity has high economic value, both direct and indirect. Direct economic values are provided by wild game, seafood, fuelwood and timber products, and indigenous and high-technology medicines. Indirect economic values accrue from biodiversity's roles in waste disposal, climate regulation, building soil and purifying water, recreation and tourism, and much more. Both directly and indirectly biodiversity safeguards or adds to human wealth, often justifying its protection in purely economic terms.

Thus, economic arguments by themselves provide a basis for valuing and protecting species, especially when we expand our concept of economic value from short-term, next-quarter profits to include longer term and indirect benefits. But economic arguments can also provide grounds for extinguishing species or for saving one species rather than another (Czech, 2002). Halting profitable developments or making costly attempts to preserve species may not have any obvious economic justification, while in some circumstances, economic justification could exist for destroying an endangered species, particularly organisms that cause disease, attack crop plants, or simply get in our way. Still, many people believe that such destruction is morally wrong even if it is economically profitable.

To say that an object has economic value is to say that it is useful to human beings or that they desire to possess it. In other words, it has an instrumental value: one can use it for his or her purposes. However, we recognize that at least some entities have value regardless of whether anyone else values or uses them. These entities have an intrinsic value: a value that is grounded not in their usefulness to others, but in what they are themselves.

Human beings, we usually think, have intrinsic value, and therefore possess certain rights that no one can legitimately infringe. Conversely, we have certain duties toward other people that specify how we should treat them in various situations. We cannot help but look at the people we interact with each day in terms of their usefulness to us, but if we look at them solely in these terms we disrespect their (and our) humanity and if we treat them solely as a means to our own ends we are likely to behave immorally.

Many conservationists argue that similar duties restrict the morally acceptable treatment of wild nature (Leopold, 1987). They consider it wrong to destroy a rare woodland or cause a species to go extinct, even if this action is in an individual's or corporation's self-interest. Not only do we not have any right to destroy species, we also have a moral responsibility to actively protect species from extinction as the result of our activities. Opponents counter that this position is illegitimate because only human beings have intrinsic value. Unless our actions affect other people, directly or indirectly, any treatment of the natural world is morally acceptable.

Extensionist Arguments for Intrinsic Value

Proponents provide both extensionist and nonextensionist arguments for the intrinsic value of wild nature (Agar, 2001). Extensionist arguments ask what qualities give intrinsic value to human beings, and then assert that some other beings possess these same qualities. Therefore, they conclude, we should recognize the intrinsic value of these other beings and extend proper treatment to them.

One common justification for valuing human beings is our ability to reason. But some of the so-called higher animals also seem to possess the rudiments of reason. Chimpanzees and gorillas have been taught sign language involving several hundred words; wolves have an elaborate social life and the ability to coordinate long hunts; and dolphins, whales, and other cetaceans send complex signals that we are just beginning to understand. Many argue that because of these factors, we should not hunt these animals for food or sport, use them in research, or in general treat them solely as means to our own ends.

Other philosophers believe that sentience – the ability to perceive the surrounding world and feel pleasure and pain – demands moral consideration. The extensionist argument here states that because pain is bad we should avoid inflicting pain unnecessarily on others who can feel it – even if those others have fur, feathers, scales, or numerous legs. Immoral actions against human beings are wrong, at base, because they cause unnecessary suffering; actions that cause other animals unnecessary suffering are likewise immoral. Millions of people around the world act on such beliefs by abstaining from eating animals or using products that can only be procured by killing or harming them. Once they have visited a modern concentrated animal feeding operation (CAFO), even most meat-eaters recognize that these animal factories are immoral, because they inflict gross, unnecessary suffering on the animals. This suggests a widespread belief that where there is sentience there exists some moral responsibility.

One possible problem with the extensionist arguments just discussed, from the point of view of conservationists, is that they only encompass certain organisms (Rolston, 1989). Basing intrinsic value on some rudimentary form of reasoning or complex mental experience would appear to rule out most animals; for example, almost all invertebrate species. Basing moral consideration on the ability to feel pain includes a wider class of animals, but again many simpler animals and all the other kingdoms of life are ruled out.

However, extensionist arguments can be extended further by recognizing that all organisms have a drive to stay alive and reproduce. In the same way that we acknowledge the rights of people to live, have children, and satisfy basic needs, we may extend these rights to individuals of all species. In his autobiography, Albert Schweitzer maintained that "a man is ethical only when life, as such, is sacred to him, that of plants and animals as that of his fellow men, and when he devotes himself helpfully to all life that is in need of help."

This position may lead to a very demanding code of conduct! Believers in strict biocentric equality assert that it is always wrong to kill individuals of any species (because of their intrinsic value) unless we need to do so to survive. Others believe that using nature to provide necessities and some measure of comfort is morally acceptable, but not to provide luxury goods. On this view, cutting down trees for firewood or to build a modest house is morally acceptable, particularly if this is done sustainably, whereas harvesting mahogany trees from rain forests to make expensive furniture for wealthy individuals half a world away is morally unacceptable. Whatever the particular judgments made, recognition of an intrinsically valuable organic world leads to distinctions between essential and inessential human uses and to a more limited, respectful use of natural resources.

Nonextensionist Arguments for Intrinsic Value

In addition to extensionist arguments, which point out similarities between wild nature and intrinsically valuable humans, there are arguments that find value in nature without referring to such similarities. Some believe it is a mistake to value other beings only for the ways in which they resemble humans. Robert Elliot in *The Monist* (1992) suggested that natural organisms may have the following properties that give them intrinsic value: "diversity, stability, complexity, beauty, harmony, creativity, organization, intricacy, elegance, and richness." These are qualities of natural organisms that we can appreciate – and that may call forth responses of personal restraint and active protection.

All species represent unique biological solutions to the problem of survival. They have solved the challenges placed before them by their environments and thrived. Some people see a value in this creativity and time-tested uniqueness. Others appreciate the beauty and elegance of the forms created by these natural processes. Still others value the complexity and ingenious structures that science and close observation have revealed. People who know and value this uniqueness, beauty, and complexity feel a special horror at its permanent disappearance. After all, if an individual dies it may be replaced by another individual more or less the same, but take away the last passenger pigeon or giant moa and their like will never return. Nonextensionist arguments thus support preservation of species, as well as protection of individual organisms (*see* Ethical Holism).

While philosophers debate whether it makes sense to assign rights to species, conservation activists are working to make such rights a legal reality. Arguably, the US Endangered Species Act, passed in 1973, created a *de facto* species right to continued existence, free from untimely human-caused extinction. In 2008, the people of Ecuador went further, explicitly adding a new chapter on the rights of nature to their national constitution. The additions, proposed by a constitutional convention and ratified by popular vote, included the following:

Article 1. Nature or *Pachamama*, where life is reproduced and exists, has the right to exist, persist, maintain itself, and regenerate its own vital cycles, structure, functions, and its evolutionary processes.
Article 3. The State will motivate natural and juridical persons as well as collectives to protect nature; it will promote respect towards all the elements that form an ecosystem.
Article 4. The State will apply precaution and restriction measures in all the activities that can lead to the extinction of species, the destruction of the ecosystems or the permanent alteration of natural cycles.

While individual human beings remain our paradigm rights-holders, societies have wide latitude to create novel rights for larger wholes. For example, the United Nations Convention on the Prevention and Punishment of the Crime of Genocide, adopted in 1948, established a right against efforts to extinguish racial or ethnic groups. Such human group rights are no different in principle from nonhuman species rights. The key question is whether the creation of such rights furthers justice and the common good.

Human maturity leads naturally to self-restraint and a respect for others. Many conservationists agree with Arne Naess (1989), who writes that the further maturation of the human species will involve an "identification with all life forms" and "the acknowledgment of the intrinsic value of these forms" in an expanding circle of moral obligations. Moving outward from oneself, the circle would include duties to our family and relatives, our local community, our country, all humanity,

mammals (Save the Whales!), all animals (Save the Snail Darter!), all species (Save the Yellow Lady Slipper Orchid!), ecosystems (Save the Rain Forest!), and ultimately the whole Earth (**Figure 1**). Such an expansion of ethical concern involves new limitations on acceptable actions, but also new opportunities for personal growth and flourishing.

Anthropocentric Denials of Intrinsic Value

Skeptics believe that even though some people do value nonhuman organisms and species, we are not obligated to do so, because only human beings have intrinsic value or possess genuine rights (Ferry, 1995). Humans have a value beyond all other beings because only we are fully conscious and rational. Unless our actions affect other people, directly or indirectly, any treatment of the natural world is morally acceptable.

Such a viewpoint is anthropocentric – locating intrinsic value solely in humans – and to many it seems the most obvious common sense, while departures from it seem irrational or overly sentimental. The appellation "tree hugger" well expresses this view, suggesting inappropriate sentiments toward trees, leading to inappropriate actions. It also suggests a callous disregard for the interests of people, like those who cut trees for a living, who we really should care about. Consider the bumper sticker "Hug a Logger, Not a Tree." People who value nature counter that anthropocentrism is selfish and that speciesism – the privileged treatment of one species over another – is no more justified than racism or sexism.

Another argument against nature's intrinsic value is that recognizing it leads to absurdity. Even people who appreciate wild nature often resist granting it intrinsic value, since this seems to demand so much from us. If nature is wonderful and complex, as science and our own experiences tell us it is, how can we go on using it? But we must do so to survive. Besides, the world is already filled with rules limiting our actions; adding even more would be burdensome. Finally, since our modern lifestyles depend to a large extent on an ecologically destructive economic system, many despair of protecting the world and give up trying to live in an environmentally responsible manner.

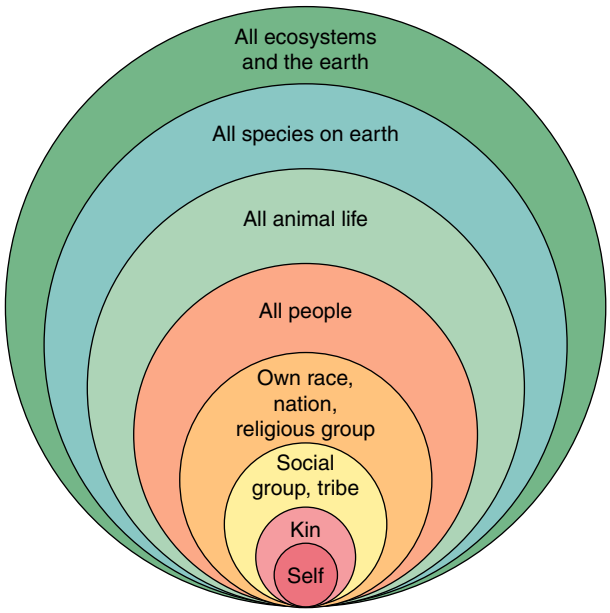


Figure 1 Human development. An ethical sequence in which the individual extends concern outward beyond the self to progressively more inclusive levels. Courtesy of Reed Noss.

These are legitimate concerns. Still, effective action to protect biological diversity is both possible and desirable. First, it is possible to use natural resources in a respectful and limited way: It is necessary to use nature, but not all uses of nature are necessary. Second, while no one likes more rules, growing up and living moral lives involves recognizing our duty to others. If nature does in fact have intrinsic value, we should respect that value – whether doing so is convenient or not. Opponents of this view counter that even though some people do value certain qualities in wild nature, they are not morally required to do so. And they repeat that humans have a value beyond all other species' value, because only we are fully conscious, rational, and free beings.

Although the debate between anthropocentrists and biocentrists has tended to incorporate all the uncertainties attending ethical justification in general, some clarification has still resulted. Charges of irrationalism to the contrary, it is possible to love and value nonhuman nature and to act on this view. The challenge for biocentrists is to fashion good lives that limit their negative environmental impacts and help preserve and celebrate nature (Mills, 2003). Anthropocentrism also remains a rationally defensible position, which may be consistently acted on. Anyone who values humanity based on qualities that we share with other species, however, cannot consistently deny intrinsic value to those other species. Furthermore, anthropocentrists who value humanity primarily for our ability to reason may consider the many arguments in favor of lessening pollution and preserving wild nature that appeal to our rational self-interest (*see* Biodiversity and Human Flourishing). Anthropocentrists are more likely than biocentrists to accept some amount of pollution in rivers or the extinction of certain species, but they also acknowledge human needs for clean air and drinking water, the enjoyment we get from fishing, swimming, and canoeing, and the value of biodiversity to science, art, and business. When it comes to particular environmental policies, anthropocentrists and biocentrists may find considerable common ground (Norton, 2003).

Ethical Holism

The preservation of biodiversity seems to demand that the needs of endangered species take precedence over the needs of individual organisms. For example, the US National Park Service killed hundreds of introduced rabbits on Santa Barbara Island off the California coast to protect a few plants of the endangered Santa Barbara live-forever (*Dudleya traskiae*) (**Figure 2**). In this case, one endangered species was judged to be more valuable than hundreds of individual animals of a common species (Rolston, 1994). Similarly, conservation biologists would not find it acceptable to destroy the last remnant of a rare biological community even if every species living there could be maintained in captivity; the ecological



Figure 2 Beasts vs the biosphere. Government agencies judged the continued existence of the endangered plant Santa Barbara live-forever (*Dudleya traskiae*; the tall plant in the photo) to be more valuable than the common rabbits on its island home. The rabbits, which fed on the plant's fleshy leaves (shown at the bottom right), were killed to stop their destruction of this fragile plant species. Reproduce with permission from National Park Service photograph.

interactions and evolutionary processes of the community would be lost if the species only lived in captivity. These examples illustrate that most conservationists are holists, finding value in larger groupings, such as species and biological communities. They are thus sometimes willing to sacrifice the interests of individuals to preserve species and communities.

Many writers, especially animal welfare advocates, have difficulty assigning rights to species. Peter Singer and Tom Regan, two prominent philosophical animal rights advocates, argue that species are not conscious entities and so do not have interests. In their view, to sacrifice the genuine interests of an individual animal to the imagined interests of a species is mistaken. Many animal welfare advocates also reject conservationists' special concern for native species over exotics: an animal is an animal, with a greater or lesser ability to suffer, and this should determine our behavior toward it, rather than its point of origin.

On both biological and ethical grounds, however, most conservation biologists argue that species, rather than individual organisms, are the appropriate targets of conservation efforts. All individuals eventually die; it is the species that continues,

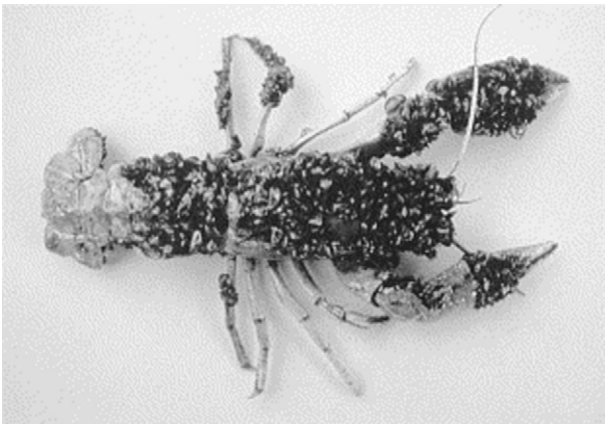


Figure 3 Unwelcome guests. The Zebra mussel (*Dreissena polymorpha*), a native of the Caspian Sea, was accidentally introduced into the Great Lakes and associated rivers in 1988. Since then the species has formed dense populations over a wide and ever-increasing area, outcompeting and choking out native species. In this case, thumb-sized Zebra mussels have almost totally encrusted a crayfish shell. Photograph courtesy of Ontario Ministry of Natural Resources from the Great Lakes Sea Grant Exotic Species Library.

evolves, and sometimes forms new species. In a sense, individuals are temporary representatives of species. Whether or not we allow them rights, species carry high value as the repositories of the accumulated experience and history of millions of previous life forms through their continuous, evolutionary adaptation to changing environments. The premature extinction of a species due to human activities destroys this natural process and obliterates its history. It can be regarded as a "superkilling," because it kills future generations of the species and eliminates whole lines from the processes of evolution and speciation.

Furthermore, conservationists typically argue that species should be prevented from spreading beyond their natural ranges as a result of direct or indirect human activity. For example, the zebra mussel (*Dreissena polymorpha*) is native to the Caspian Sea, but it has recently become an aggressive invader of North American aquatic habitats (**Figure 3**). Arguably this species should be destroyed whenever possible in North America, for two reasons. First, exotics often displace native species, sometimes contributing to their extinction. To prevent this loss of biological diversity we destroy individual exotics, on the assumption that the exotic species will continue to thrive in its native habitats. Second, and more controversially, many conservationists deny intrinsic value to individual plants and animals that have been spread beyond their natural range by people. Part of what gives a species value is its unique evolutionary history and its ecological roles in native habitats, both of which are tied to particular places and biological communities. When these species instead invade new natural areas and destabilize or radically change their species composition, these justifications of their value no longer hold. Species contribute to biodiversity in their native habitats but often diminish biodiversity when they become common in exotic new locales. If our overarching conservation goal is to protect life's full flourishing and diversity, exotic species are immensely harmful and should be reined in where practical.

Duties to Protect Biodiversity: Obligations to People and to God

Duties to Humans

That human beings have direct moral duties to other species remains controversial; that we have direct moral duties to other human beings is not (although great disagreement remains regarding their scope). Our duties to other human beings may support environmental protection and biodiversity preservation, based on their instrumental value to intrinsically valuable people (Moore and Nelson, 2010).

We have responsibilities toward our neighbors and fellow citizens. Because of this, we must arguably minimize damage to the natural environment, since such damage harms not only other species but people as well. Increasingly, connections are being made between environmental pollution and high levels of human disease. For example, massive environmental pollution in the former Soviet Union greatly increased cancer, birth defects, lead poisoning, and lung disease among the population. In the United States, evidence is growing that increased levels of endocrine-disrupting chemicals are harming people as well as wildlife. Much pollution and environmental degradation is unnecessary and could be minimized with better planning. Often pollution occurs because corporate leaders are unwilling to spend money to prevent it, despite the resulting ill effects on the health, wealth, or happiness of other people. Citizens must recognize the biological and social costs of environmental damage and force corporations to be good neighbors and governments to enact and enforce strong environmental laws for the common good.

We also have responsibilities toward future generations. Economic decision making tends to focus on the short term and it is this economic system that is driving environmental degradation (Speth, 2009). In response, many ethicists have emphasized the importance of intergenerational justice (Gardiner *et al.*, 2010). Humanity's unprecedented numbers and technological power mean that one generation can now radically remake the Earth that the next will inhabit, for better or worse. If in our daily living we degrade the natural resources of the Earth and cause species to become extinct, future generations will pay the price in terms of a lower quality of life.

This truth gives us arguments not just for keeping ecosystems unpolluted to safeguard human health, but also for preserving biodiversity for future human use, enjoyment, and development. If species and wild places are lost, children will be deprived of one of their most exciting experiences growing up – the wonder of seeing "new" animals and plants in the wild. These concerns are a powerful motivating force for the members of organizations such as the Sierra Club and the Nature Conservancy, who see themselves as land stewards preserving biodiversity for future generations. Over one hundred and fifty years ago, reflecting on the depauperate ecological landscape around him in Concord, Massachusetts, Henry Thoreau wrote in his journal:

> "When I consider that the nobler animals have been exterminated here, I cannot but feel as if I lived in a tamed, and, as it were, emasculated country … I take infinite pains to know the phenomena of the spring, thinking that I have here the entire poem, and then, to my chagrin, I hear that it is but an imperfect copy that I possess and have read, that my ancestors have torn out many of the first leaves and grandest passages, and mutilated it in many places." (March 23, 1856).

Like Thoreau, many of us believe that our descendants will "wish to know an entire heaven and an entire Earth." They have a right to know and explore wild nature. We should not take that away from them.

Holmes Rolston in the *Encyclopedia of Environmental Biology* (1995) predicted: "It is safe to say that in the decades ahead, the quality of life will decline in proportion to the loss of biotic diversity, though it is often thought that one must sacrifice biotic diversity to improve human life." Of course, this contention is debatable. Many argue that creating jobs or increasing wealth is more important to future generations than the preservation of biodiversity (Lomborg, 2010). Debate on this issue, while inconclusive, is essential, because it forces us to specify the actual, long-term benefits of development projects and the sorts of societies we want to create for our children. Given our unprecedented power and the tendency to put personal interests above the common good, even inconclusive debate may be a force for creating a better future for all.

At a minimum, our duties to posterity seem to require us to live sustainably – that is, to limit our consumption so as not to degrade essential life-support systems or deplete natural resources that future generations will need. The world's governments have formally recognized this through treaties governing air pollution, whale hunting, disposal of wastes in the ocean, and more. One important treaty, the Montreal Protocol of 1987, reduced the permitted use of ozone-depleting chlorofluorocarbons (CFCs), leading to a gradual phaseout of production. The grave dangers these chemicals posed to future generations, with their potential to thin the Earth's protective ozone layer and cause millions of new cases of deadly skin cancer, clearly outweighed the interests of producers and users of CFCs, especially since many of these uses were not essential to human survival (e.g., luxury items such as hairsprays and air conditioners) and alternatives were available or quickly developed for most uses. Similar appeals to the good of future generations have so far proved less successful in convincing the world community to reduce fossil fuel use to slow global climate change. This is partly due to well-financed lobbying by oil and coal corporations, whose leaders have placed short-term profits above the health and well-being of future generations. Partly though it is because the risks of increased fossil fuel use, though grave, are somewhat uncertain, while the inconvenience and economic costs of decreasing fossil fuel use are far higher than they were for cutting back on CFCs. Nevertheless, it seems immoral for this current, wealthy generation of human beings to avoid serious action to combat global warming, when our inaction risks the lives and livelihoods of fellow human beings at home and abroad (Arnold, 2011).

Some people worry that preserving nature requires taking resources and opportunities away from human beings. Sometimes it does. But protecting biological diversity also can be linked to better health and greater opportunities for people (Cafaro and Sandler, 2010) and some of the most exciting developments in conservation biology attempt to do just that. Helping poor people establish sustainable plots of cash crops

and achieve a degree of economic independence may reduce the need to overharvest wild species. Working with indigenous people to secure legal title to their land may give them the means to protect the biological communities in which they live (although they do not always choose to do so in practice). Working to benefit people is often not just compatible with effective conservation, but necessary to its success.

Conservationists need to look for "win win" scenarios where the interests of people and nature can both be furthered, whether we are working in underdeveloped or overdeveloped nations. However, in some situations, protecting biological diversity will be incompatible with meeting immediate human needs or promoting short-term human interests. In such situations, hard choices will need to be made and these choices should not automatically favor people over other species, or current over future generations. Extinction is forever and extinguishing other species is rarely if ever necessary to preserve human lives. Our duties to future human generations arguably include a strong duty to preserve other species from extinction.

Religious Stewardship

Some preindustrial cultures successfully coexisted with a rich local flora and fauna for hundreds of years, in part because their religious and ethical beliefs encouraged the thoughtful use of resources. People in these societies respected wild animals and plants, even as they harvested them or borrowed their habitat for human purposes. For example, the Cherokee Indians of the southeastern United States spoke special prayers to the deer that they killed, telling the deer spirit that they indeed needed the meat, that they would not waste it, and that they would bury the bones with due solemnity. Traditional peoples often treated rivers, mountains and other ecosystems as sacred places to be approached with reverence and an appreciation for what they were, rather than with haste, focused on what human beings could make of them.

Many modern religious believers abhor the destruction of species, because they are God's creation. Within the Jewish and Christian traditions, human responsibility for protecting animal species is explicitly described in the Bible as part of a covenant with God. The Book of Genesis describes the creation of the Earth's biological diversity as a divine act, after which "God saw that it was good" and "blessed them." In the story of Noah's Ark, God commanded Noah to save two of all species – not just the ones human beings found useful. God provided detailed instructions for building the ark – an early species rescue project – saying "Keep them alive with you," and after the flood subsided, the animals were released to repopulate the Earth. The prophet Mohammed, founder of Islam, continued this theme of human responsibility, saying "The world is green and beautiful and God has appointed you as His stewards over it. He sees how you acquit yourselves."

Belief in the value of God's creation supports a stewardship argument for preserving biodiversity: Human beings have been given responsibility for God's creation and must preserve, not destroy, what they have been given.

Many other religious traditions also support the preservation of nature. For example, Hinduism locates divinity in certain animals, and recognizes a basic kinship between humans and other beings (including the transmigration of souls from one species to another). A primary ethical concept in Hinduism, Jainism, and Buddhism is *ahimsa*: nonviolence, or compassion for all life. To live by this ideal, many believers become vegetarians and live materially simple lives.

Of course, some religions articulate views that put human beings at the center of creation, supporting a domineering attitude toward nature. But such interpretations are contested. Since many people base their ethical values on a religious faith, the development of religious arguments in support of conservation might be effective in motivating people to conserve biodiversity (Wirzba, 2003).

In September 1986, an interfaith ceremony was held in the Basilica of St. Francis, in Assisi, Italy. It included "Declarations on Nature" by representatives of the five participating religions: Buddhism, Christianity, Hinduism, Islam, and Judaism. For the first time in history, leaders of these faiths came together to declare that their religions mandate the conservation of nature. Excerpts from the five declarations follow.

The Buddhist Declaration on Nature

Venerable Lungrig Namgyal Rinpoche, Abbot, Gyuto Tantric University

> The simple underlying reason why beings other than humans need to be taken into account is that, like human beings, they too are sensitive to happiness and suffering …. Many have held up usefulness to human beings as the sole criterion for the evaluation of an animal's life. On closer examination, one discovers that this mode of evaluation of another's life and right to existence has also been largely responsible for human indifference as well as cruelty to animals.

We regard our survival as an undeniable right. As coinhabitants of this planet, other species too have this right for survival. And since human beings as well as other nonhuman sentient beings depend on the environment as the ultimate source of life and well-being, let us share the conviction that the conservation of the environment, the rectification of the imbalance caused by our negligence in the past, be implemented with courage and determination.

"All sentient beings, all breathing things, creatures without exception, may no evil befall them." – Gradual sayings of the Buddha.

The Christian Declaration on Nature

Father Lanfranco Serrini, Minister General, Order of Friars Minor (Franciscans)

> To praise the Lord for his creation is to confess that God the Father made all things visible and invisible; it is to thank him for the many gifts he bestows on all his children …. By reason of its created origin, each creature according to its species and all together in the harmonious unity of the universe manifest God's infinite truth and beauty, love and goodness, wisdom and majesty, glory and power.
>
> Man's dominion cannot be understood as license to abuse, spoil, squander, or destroy what God has made to manifest his glory. That dominion cannot be anything other than a stewardship in symbiosis with all creatures …. Every human act of

irresponsibility toward creatures is an abomination. According to its gravity, it is an offense against that divine wisdom which sustains and gives purpose to the interdependent harmony of the universe.

The Hindu Declaration on Nature

Dr Karan Singh, President, Hindu Virat Samaj

The Hindu viewpoint on nature is permeated by a reverence for life, and an awareness that the great forces of nature – the earth, the sky, the air, the water and fire – as well as various orders of life including plants and trees, forests and animals, are all bound to each other within the great rhythms of nature. The divine is not exterior to creation, but expresses itself through natural phenomena. The *Mahabharata* says that 'even if there is only one tree full of flowers and fruits in a village, that place becomes worthy of worship and respect'.

Let us declare our determination to halt the present slide toward destruction, to rediscover the ancient tradition of reverence for all life and, even at this late hour, to reverse the suicidal course on which we have embarked. Let us recall the ancient Hindu dictum: "The Earth is our mother, and we are all her children."

The Muslim Declaration on Nature

Dr Abdullah Omar Nasseef, Secretary General, Muslim World League

The essence of Islamic teaching is that the entire universe is God's creation. Allah makes the waters flow on the earth, upholds the heavens, makes the rain fall and keeps the boundaries between day and night …. It is God who created the plants and the animals in their pairs and gave them the means to multiply.

For the Muslim mankind's role on earth is that of a *khalifa*, viceregent or trustee of God. We are God's stewards and agents on Earth. We are not masters of this Earth; it does not belong to us to do what we wish. It belongs to God and He has entrusted us with its safekeeping …. The *khalifa* is answerable for his/her actions, for the way in which he/she uses or abuses the trust of God.

The Jewish Declaration on Nature

Rabbi Arthur Hertzberg, Vice President, World Jewish Congress

The encounter of God and man in nature is conceived in Judaism as a seamless web with man as the leader and custodian of the natural world … Now, when the whole world is in peril, when the environment is in danger of being poisoned and various species, both plant and animal, are becoming extinct, it is our Jewish responsibility to put the defense of the whole of nature at the very center of our concern.

We have a responsibility to life, to defend it everywhere, not only against our own sins but also against those of others. We are all passengers together in this same fragile and glorious world. Let us safeguard our rowboat – and let us row together.

"… and the rainbow shall be in the cloud; and I will look upon it, that I may remember the everlasting covenant between God and every living creature … that is upon the Earth." – Genesis 9:16

Biodiversity and Human Flourishing

Economic arguments stress that preserving biological diversity is in our material self-interest. Ethical arguments based on the intrinsic value of wild nature and our duties to others stress that we should act altruistically toward nature regardless of our material self-interest. A second ethical framework appeals to our enlightened self-interest, arguing that preserving biodiversity and developing our knowledge of it will make us better and happier people (Sandler, 2007). Protecting biological diversity is in our enlightened self-interest, for the following reasons.

Esthetic and Recreational Enjoyment. Nearly everyone enjoys wildlife and landscapes esthetically, and joy increases the quality of our lives. Nature-related activities are important in childhood development. The beauty of a field of wildflowers in Glacier National Park or a migrating warbler on a spring morning in a city park enriches people's lives (Thoreau, 1971). And for many people, truly experiencing nature means experiencing it in a natural setting – simply reading about species or seeing them in museums, gardens, or zoos does not suffice (**Figure 4**). Recreational activities such as hiking, canoeing, and mountain climbing are physically, intellectually, and emotionally satisfying. People spend tens of billions of dollars annually in these pursuits, proof enough that they value them highly.

Artistic Expression and Philosophical Insight. Throughout history, writers, artists, and musicians of all cultures have drawn inspiration from wild nature. Nature provides countless forms and symbols for painters and sculptors to render and interpret (**Figure 5**). Poets have often found their greatest inspiration in either wild nature or the pastoral countryside. Preserving biological diversity preserves possibilities for all artists and for everyone who appreciates their works. Philosophers too go to nature to find insights into human existence and our place in the wider universe. A loss of biodiversity limits such experiences and diminishes our intellectual resources. It is hard to see how, in a world without mockingbirds, Americans could appreciate Walt Whitman's "Out of the Cradle Endlessly Rocking," one of their greatest poet's greatest poems. And it is sad to think what other poems, songs, stories, or pictures will remain unsung, untold, or unmade, in a world without butterflies or broad meadows filled with wildflowers.



Figure 4 Reaching out. Most people find interacting with other species to be an educational and uplifting experience. Here, people greet a minke whale that is being rescued after it became entangled in a trawler's gill net; the float behind the whale was attached to the net to keep the whale at the surface so that it could breathe. Later, rescuers were able to release the whale from the netting. Reproduced with permission from Scott Kraus, New England Aquarium.



Figure 5 Celebrating nature. Rare wildflowers and butterflies are the inspiration for botanical sculptor Patrick O'Hara. In his studio in western Ireland, O'Hara molds, sculpts, and paints delicate porcelain scenes from nature that inspire an appreciation of biodiversity in a worldwide audience. Photograph courtesy of O'Hara (2006).

Scientific Knowledge. Science and our growing knowledge of nature are among humanity's greatest achievements. This knowledge is facilitated by the preservation of wild nature. Wild areas allow the study of natural ecological processes and interactions. Wild species preserve the record of evolution. Young people are inspired to become scientists by personal contacts with wild nature, and those who do not become professional scientists can apply their basic knowledge of science to understanding their own local forests, streams, and fields.

Three of the central mysteries in the world of science are how life originated, how the diversity of life interacts to form complex ecosystems, and how humans evolved. Biologists working on these problems are coming ever closer to the answers. However, when species become extinct and ecosystems are damaged, important clues are lost, and the mysteries become harder to solve. For example, if *Homo sapiens'* closest living relatives – chimpanzees, bonobos, gorillas, and orangutans – disappear from the wild, we will lose important clues regarding human physical and social evolution.

Historical Understanding. Knowing nature, both scientifically and through personal experience, is key to an understanding of human history. In walking the landscapes our ancestors walked, we gain insight into how they experienced the world at a slower pace and without mechanized aids. We often forget just how recently humankind has moved to ultrafast transportation, fully illuminated cities that shut out the night, and other aspects of modern life. We need to preserve natural areas in order to develop our historical imaginations and preserve ties to our past (Leopold, 1987).

Religious Inspiration. Many religions have traditions of "wandering in the wilderness" in order to commune with God or with spirits, or to purify themselves of the temptations and evils associated with life within human communities. Moses, Isaiah, St. John the Baptist, St. Francis of Assisi, and Jesus all sought out the solitude of wilderness to obtain spiritual strength and receive the guidance of God. Generations of Sioux, Ute, Cheyenne, and other Native American vision seekers have done the same. Being in nature allows us to clear and focus our minds and sometimes, experience the transcendent. When we are surrounded by the artifacts of civilization, our minds usually stay fully focused on human purposes and our everyday lives. Religion probably would not disappear from an environment totally tamed by humans, but it might become diluted for many.

Protecting our Life-Support Systems and Economies. It cannot be repeated too often that biological diversity preserves our basic life-support systems of food production, water supply, oxygen replenishment, waste disposal, soil production, and more. People will be healthier and happier in clean, intact environments. We depend on this and should value it. In addition to providing life-support, biodiversity allows us to create tremendous economic wealth, directly and indirectly. An article published in *Nature* in 1997 by Robert Costanza *et al*. "The Value of the World's Ecosystem Services and Natural Capital," estimated that the world's ecosystems provided $32 trillion per year of value to people, substantially higher than the $18 trillion per year of goods and services those people produced themselves. Human society could not exist without what the natural world provides us with for free, nor could we afford to pay for substitutes, even if they existed.

For all these reasons, degrading ecosystems and destroying species is almost always contrary to people's real interests. When it appears otherwise, that is usually because we are taking a short-term, selfish, or overly materialistic view of what those interests are (Norton, 2003). While the preservation of biodiversity sets limits to some human activities, it is a prerequisite for our continued enjoyment of others. There are good reasons to believe that preserving and exploring biodiversity can make us better, happier people (Cafaro and Sandler, 2010). Many environmental ethicists are convinced that a better understanding of our true self-interest would lead to greater efforts to protect biological diversity.

Environmental Philosophies

Deep Ecology

Recognition of the many instrumental and intrinsic values of biological diversity leads naturally to new limits on human action. This can make it seem like conservation is simply a never-ending list of "thou shalt nots." But this is only part of the story, and many environmentalists come to a different conclusion. According to Arne Naess (1989): "The crisis of life conditions on Earth could help us choose a new path with new criteria for progress, efficiency, and rational action… The ideological change is mainly that of appreciating life quality rather than adhering to a high standard of living."

One well-developed environmental philosophy that embraces this new approach to living in nature is known as deep ecology (Naess, 2008). It builds on the basic premise of biocentric equality: that all creatures have an equal right to live and thrive in their own ways. Proponents oppose what they see as the dominant worldview, which places human concerns above all others and views human happiness in materialistic terms (**Table 1**).

Deep ecologists may hold various foundational beliefs. Some are Christians, Jews, or Buddhists, while others are agnostics or atheists; philosophically, they may be utilitarians, communitarians, or rights-theorists, or they may have no well-worked-out theory of ethics. What they have in common is a belief in wild nature's great value and a strong commitment to act to preserve it. This belief and commitment are captured in the eight-point Deep Ecology Platform, formulated by Arne Naess and George Sessions:

1. The well-being and flourishing of human and nonhuman life on Earth have value in themselves (synonyms: inherent worth, intrinsic value, inherent value). These values are independent of the usefulness of the nonhuman world for human purposes.
2. Richness and diversity of life forms contribute to the realization of these values and are also values in themselves.
3. Humans have no right to reduce this richness and diversity except to satisfy vital needs.
4. Present human interference with the nonhuman world is excessive, and the situation is rapidly worsening.
5. The flourishing of human life and cultures is compatible with a substantial decrease of the human population. The flourishing of nonhuman life requires such a decrease.
6. Policies must therefore be changed. The changes in policies affect basic economic, technological, and ideological structures. The resulting state of affairs will be deeply different from the present.
7. The ideological change is mainly that of appreciating life quality (dwelling in situations of inherent worth) rather than adhering to an increasingly higher standard of living. There will be a profound awareness of the difference between big and great.
8. Those who subscribe to the foregoing points have an obligation directly or indirectly to participate in the attempt to implement the necessary changes (Naess, 1989).

Table 1 The dominant anthropocentric philosophy contrasted with deep ecology

Dominant worldview	*Deep ecology*
Humans dominant over nature	Humans living in harmony with nature
Natural environment and species are resources for humans	All nature has intrinsic worth, regardless of human needs
A growing human population with a rising standard of living	A stable human population living a materially simple life
Earth's resources are unlimited	Earth's resources are limited and must be used carefully
Ever higher technology brings progress and solutions	Appropriate technology must be used with respect for the Earth
Emphasis on material progress	Emphasis on spiritual and ethical progress
Strong central government	Local control, organized according to watersheds, bioregions, or other natural units

Deep ecologists and other biocentric philosophers see acceptance of the intrinsic value of nature less as a limitation than as an opportunity to live better lives. Over the past 200 years, the Industrial Revolution, with its accompanying technological advances and social changes, has generated tremendous material wealth and improved the lives of millions of people. But a law of diminishing returns seems to apply: for many people throughout the world, heaping up further wealth makes little sense, while the continued loss of biodiversity and taming of natural landscapes will not improve their lives (McKibben, 2007). What is being lost is unique and increasingly more precious as wealth increases and opportunities to experience wild nature diminish. These radical environmental philosophies thus ask us to imagine creating new sorts of economies that provide a sufficiency of goods and services for a limited number of people, not ever-increasing amounts of goods and services for endlessly more people.

Some biocentric philosophers also ask us to imagine different kinds of people living and working within these economies – if only because we will have to be somewhat different in order to accept living within limits. The kind of selfish, shortsighted behavior that is destroying biodiversity today will not change because of a few good arguments, or even solely because of new laws and institutions – although new laws and institutional change are certainly needed. Put bluntly, the claim is that in order to preserve biodiversity, we are going to have to become *better people* than we are right now.

Virtue ethics is the area of philosophy focused on ideals of human excellence. Virtue ethicists work to specify those qualities of human character ("the virtues") that make people *good* people: more likely to treat others well, and better able to succeed in important human endeavors. Over the past decade, environmental virtue ethicists have worked to specify those character traits that we will have to exemplify as citizens of truly sustainable societies (Sandler, 2007). They include:

- justice, in our treatment of other species and one another;
- generosity, involving a willingness to go beyond the demands of justice in sharing resources;
- temperance, shown in an ability to limit our cravings for more wealth and possessions and to act with self-control;
- prudence, in order to think long-term and follow the dictates of reason regarding grave environmental threats;
- attentiveness, to the beauty and complexity of nature;
- gratitude, in appreciating what we have and taking the time to give thanks for it, rather than always hankering after more;
- wisdom, in recognizing the real value of this world we live in, knowing how to live well within it, and acting on that knowledge (Cafaro and Sandler, 2010).

The claim, a new one in the history of philosophy, is that just as no one can be a virtuous person if they do not treat

other people justly, so no one can be a virtuous person if they do not live sustainably themselves and further efforts to create sustainable societies. It would be nice if we could just rearrange our institutions a bit, perhaps build in a few more safeguards, and preserve biodiversity and the essential ecosystem services that we depend on, without having to change ourselves. But that seems unlikely. From here onwards, in the busy, crowded world we have created, human excellence, well-being and sustainability are necessarily intertwined.

Some readers may see such talk as utopian, or even a counsel of despair. If it is hard to imagine radical economic change, it is even harder to imagine radical changes in human nature. Furthermore, in a democratic age, we are not used to demanding that our fellow citizens become better people. However, these changes do not have to happen all at once, or solely as a matter of individual willpower. Instead, we need to begin reforming our institutions, so that they help create better people over time: Schools that educate new citizens who are just, temperate, and generous, in addition to knowledgeable about nature and environmental issues; economies that create intelligent consumers whose purchasing decisions are thoughtful and restrained, rather than intemperate. Who is to say that such changes are impossible? The career of *Homo sapiens* may be just beginning! Perhaps human beings can be much more than we are now. Perhaps we will have to be, in order to preserve biodiversity.

Humanism: Our Default Philosophy

Like deep ecologists, many conventional ethical philosophers, have begun to address environmental issues. Most philosophers who have considered these matters (including deontologists, utilitarians, social contract theorists and theologically-grounded ethicists) have found strong reasons to support environmental protection (Schmidtz and Willott, 2012). As previously noted, biocentrists and anthropocentrists may agree on a wide range of practical measures to protect the environment, despite large philosophical differences.

Still, some philosophers and many members of the general public remain unconvinced of any moral imperative to protect biological diversity. Their position may be characterized as humanism, a philosophy committed to the following propositions (Ferry, 1995; Lomborg, 2010):

1. Biological diversity exists for humans, has no value apart from humans, and need not exist apart from humans.
2. The continued transformation of wild nature into natural resources adds value to nature, since nature possesses value only in human use. Indeed, the ever more thorough transformation of wild nature allows greater numbers of people to lead longer, happier, and better lives.
3. The creation of just societies filled with flourishing individuals is the highest achievement of which humans are capable. We should judge ourselves based on our technological, scientific, artistic, and social progress – not on whether we preserve nature.

Like other environmental philosophies, humanism may be developed in a variety of ways. A humanist's ideal society may be more or less egalitarian, more or less wealthy, more or less racially and culturally diverse. But humanists share a belief in the centrality of humans. They generally applaud increased human numbers, wealth, and technological power, and the development of new arts and activities that flourish in highly artificial environments. For these reasons they see little to lament in the loss of biodiversity.

A main motivation for embracing humanism is its rejection of limits to growth. Limiting both population growth (Grant, 2006) and economic growth (Czech, 2002) is essential to preserving biodiversity. But even many proponents of biodiversity protection shrink from defending explicit limits to growth. Two spectacular recent examples of this failure were the 2005 Millennium Ecosystem Assessment (MEA) and the 2007 4th Assessment Report of the Intergovernmental Panel on Climate Change. In laying out future scenarios for detailed consideration, the authors of the MEA, who included numerous biologists and ecologists, failed to include any which explicitly limited economic growth. This meant that all the scenarios they considered included massive biodiversity loss. The IPCC's 4th Assessment Report identified the primary causes of climate change as unremitting economic and demographic growth, and warned that under "business as usual," the world courted potentially disastrous levels of global warming. Yet their detailed discussion of measures to limit (or "mitigate") climate change, which ran to many hundreds of pages, failed to consider measures which would directly limit such growth. Instead, they focused nearly exclusively on technological improvements to accommodate more growth – an obviously self-defeating approach.

When even nature's defenders refuse to question the necessity and goodness of continued growth, we see the powerful hold this ideal has on the contemporary mind (Moore and Nelson, 2010). This refusal to probe growth appears to motivate many climate change deniers, for how could something so good result in such disastrously bad consequences? It lies behind the willingness to consider radical geoengineering schemes, such as seeding the oceans with iron or pumping aerosols into the stratosphere – rather than embracing simpler, cheaper, and more obvious solutions to climate change which involve economic and demographic restraint. Of course, as with biodiversity preservation, there are ample human-focused reasons for opposing climate geoengineering, based on its uncertain impacts on future generations of people. But we should also be impressed by its breathtaking anthropocentrism, which typically goes unremarked in the scientific literature and popular press. In effect, such proposals accept that Earth, which over eons has been home to an amazing and ever-increasing diversity of hundreds of millions of species, will henceforth be managed for the benefit of a single species. Us.

In the end, more people and more human economic activity equal less biodiversity: more of us equals less of them. So as a practical matter, whether or not our responses to climate change, species extinction, ocean acidification, and other global environmental challenges accept or reject limits to growth, will answer the question of whether we are biocentrists or anthropocentrists, deep ecologists or confident planetary managers. As Aldo Leopold (1987) noted, we write our environmental philosophies on the land itself (and in the air and under the seas). Our answers will be clear, in the

presences and absences, the sounds and silences which remain after we are gone.

Humanism may be called humanity's default mode, since current trends are moving us more and more in this direction, whether it is desirable or not. We are creating a world with much less biological diversity and this artificial world has come to seem normal, indeed natural, to many people. It is not, however, inevitable. Human beings can reconnect to nature and curb the activities which threaten it. Environmental ethics reminds us that we have many reasons to value and protect Earth's remaining biological diversity. The next hundred years should tell us a lot about whether self-proclaimed *Homo sapiens* really is wise, or merely clever.

See also: Ethical Issues in Biodiversity Protection. Religious Traditions and Biodiversity. Social and Cultural Factors. Stewardship, Concept of. The Value of Biodiversity. Traditional Conservation Practices

References

Agar N (2001) *Life's Intrinsic Value: Science, Ethics, and Nature*. New York: Columbia University Press.

Arnold D (ed.) (2011) *The Ethics of Climate Change*. Cambridge: Cambridge University Press.

Cafaro P and Sandler R (eds.) (2010) *Virtue Ethics and the Environment*. Dordrecht, Netherlands: Springer-Verlag Press.

Czech B (2002) *Shoveling Fuel for a Runaway Train: Errant Economists, Shameful Spenders, and a Plan to Stop them All*. Berkeley, CA: University of California Press.

Ferry L (1995) *The New Ecological Order*. Chicago: University of Chicago Press.

Gardiner S, Caney S, Jamieson D, and Shue H (eds.) (2010) *Climate Ethics: Essential Readings*. New York: Oxford University Press.

Grant L (ed.) (2006) *The Case for Fewer People*. Santa Ana, CA: Seven Locks Press.

Leopold A (1987) *A Sand County Almanac: And Sketches Here and There*. New York: Oxford University Press.

Lomborg B (2010) *Smart Solutions to Climate Change: Comparing Costs and Benefits*. Cambridge: Cambridge University Press.

McKibben B (2007) *Deep Economy: The Wealth of Communities and the Durable Future*. New York: Henry Holt.

Mills S (2003) *Epicurean Simplicity*. Washington, DC: Island Press.

Moore K and Nelson M (eds.) (2010) *Moral Ground: Ethical Action for a Planet in Peril*. San Antonio, TX: Trinity University Press.

Naess A (1989) *Ecology, Community, and Lifestyle: Outline of an Ecosophy*. Cambridge: Cambridge University Press.

Naess A (2008) In: Drengson A and Devall B (eds.) *The Ecology of Wisdom: Writings of Arne Naess*. Berkeley, CA: Counterpoint Press.

Norton B (2003) *Searching for Sustainability: Interdisciplinary Essays in the Philosophy of Conservation Biology*. New York: Cambridge University Press.

Rolston III H (1989) *Philosophy Gone Wild: Environmental Ethics*. Buffalo: Prometheus Books.

Rolston III H (1994) *Conserving Natural Value*. New York: Columbia University Press.

Sandler R (2007) *Character and Environment: A Virtue-Oriented Approach to Environmental Ethics*. New York: Columbia University Press.

Schmidtz D and Willott E (eds.) (2012) *Environmental Ethics: What Really Matters, What Really Works*. New York: Oxford University Press.

Speth G (2009) *The Bridge at the Edge of the World: Capitalism, the Environment, and Crossing from Crisis to Sustainability*. New Haven: Yale University Press.

Wirzba N (2003) *The Paradise of God: Renewing Religion in an Ecological Age*. Oxford: Oxford University Press.

Environmental Impact, Concept and Measurement of

Ellen W Chu and James R Karr, University of Washington, Seattle, WA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biological integrity Wholeness of a living system, including the capacity to sustain the full range of organisms and processes having evolved in a region.
Biosphere The thin layer of life at or near the surface of the Earth.
Biota Living things, in particular, the flora and fauna of a region.
Biotic impoverishment Systematic reduction in Earth's capacity to support life.
Ecosystem engineers Organisms that physically shape their environment, including organisms that create or modify the environments of other organisms.
Environment Surroundings; the complex of physical, chemical, and biotic factors acting upon a living system and influencing its form, function, and survival.
Health A flourishing condition, well-being; capacity for self-renewal.
Impact A forceful contact; a major effect of one thing on another.

One Species' Impact

All organisms change their environment as they live, grow, and reproduce. Over millennia, organisms evolve to contend with changes in their environment. Those that do not adapt go extinct. Those that survive are molded by evolution and biogeography as the environment changes. Even unusual or seemingly catastrophic events, like tidal waves from earthquakes, are an integral part of the ecological contexts to which organisms adapt over long time spans.

Some organisms, like beavers and elephants, change their surroundings so dramatically that they have been called ecosystem engineers. Beaver dams alter the flow of rivers, increase dissolved oxygen in downstream waters, create wetlands, and modify streamside zones. African elephants convert wooded savanna to open grassland by toppling trees as they browse. Change brought about by living things, including ecosystem engineers, has been slow and incremental in evolutionary terms. Ecosystem engineers evolve along with other inhabitants of their environments, developing ways to coexist; like other environmental components, ecosystem engineers, and their effects are part of an evolving ecological context.

In contrast, human effects since the Industrial Revolution – including many that may be invisible to a casual observer – outpace the capacity of living systems to respond. In evolutionary terms, these effects are recent and outside the experience of most organisms. Over the past two centuries – barely more than two human lifetimes – humans have profoundly altered living and nonliving systems everywhere. For the first time in the Earth's history, the environmental impact of one species, *Homo sapiens*, is the principal agent of global change. Understanding the environmental impacts of human actions is one of modern science's greatest challenges. Understanding the consequences of those impacts, and managing them to protect the well-being of human society and other life on Earth, is humanity's greatest challenge.

Human Environmental Impact Through Time

The human evolutionary line began in Africa about 7 million years ago (Ma). It took some 5 or 6 million years (My) for protohumans to spread from Africa to Asia and then to Europe. These early humans, like other primates, made their living by seeking food and shelter from their environment, gathering plant foods, and hunting easy-to-kill prey. Sometimes, they also experienced threats *from* their environment, including accidents, droughts, vector-borne diseases, and attacks from predators. At this stage, with relatively low population densities and limited technologies, humans were not ecosystem engineers.

By some 50,000 years ago, however, humans had learned to use fire; developed complex tools, weapons, and language; and created art. On local scales, these modern humans were very much ecosystem engineers. Sometimes their enhanced abilities to make a living outstripped their local environment's capacity to provide that living, and they caused local ecological disruptions. On several continents, for example, humans hunted large mammals to the point where many, such as the marsupial lion of Australia, went extinct. As humans became more efficient at exploiting their local environments, they spread farther. By 13,000 years ago, modern humans had spread to all continents and many islands across the globe. Then, about 10,000 years ago, people began to domesticate plants and animals. Instead of searching for food, they began to produce food.

Food production changed the course of human and environmental history. Domestication of plants and animals enabled people to adopt a sedentary lifestyle. As detailed by geographer and ecologist Diamond (1997, 2002), populations grew as agriculture developed, because larger sedentary populations both demanded and enabled more food production. Local ecological disruptions became more numerous and widespread and more intense. With animal domestication, contagious diseases of pets and livestock adapted to new, human hosts. Diseases spread more quickly in crowded conditions; inadequate sanitation compounded the effects. From agriculture, civilization followed and with it, cities, writing, advanced technology, and political empires. In just 10,000 years, these developments led to nearly 7 billion people on Earth, industrial societies, and a global economy founded on complicated technologies and fossil fuels. Humans have emerged as ecosystem engineers on a global scale. The ecological disruptions we cause are no longer just

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00253-7

local or regional but global, and we have become the principal threat *to* the environment.

Yet despite today's advanced technologies, humans are as dependent on their environments as other organisms are. History, not just ecology, has been very clear on this point. From the Old Kingdom of Egypt more than 4000 years ago to the culture that created the huge stone monoliths on Easter Island between 1000 and AD 1550 to the 1930s Dust Bowl of North America, civilizations or ways of life have prospered and failed by using and (mostly unwittingly) abusing natural resources.

In Old Kingdom Egypt, the resource was the valley of the Nile, richly fertilized with sediment at each flooding of the river, laced with canals and side streams, blessed with a luxuriant delta. Agriculture flourished and populations swelled, until unusually severe droughts brought on the civilization's collapse. On Easter Island, the resource was trees, which gave Polynesians colonizing the island the means to build shelter, canoes for fishing the open waters around the island, and log rollers for moving the ceremonial stone monuments for which the island is famous. Deforestation not only eliminated the humans' source of wood, but also further deprived the already poor soil of nutrients and made it impossible to sustain the agriculture that had sustained the island's civilization. On the dry Great Plains of North America, settlers were convinced that rain would follow the plow, and so they plowed homestead after homestead, only to watch their homesteads' soils literally blow away in the wind.

In these cases and many others, human civilizations damaged their environments, and their actions also worsened the effects on their civilizations of climatic or other natural cycles. Yet in each case, a human culture was operating precisely the way it had evolved to operate: The culture of Old Kingdom Egypt enabled its people to prosper on the Nile's natural bounty, but prolonged, unprecedented drought brought starvation and political disorder. Easter Islanders thrived and populated the island until its resources were exhausted. Dust Bowl farmers lived out their culture's view of dominating and exploiting the land for all it was worth. The inevitable outcome for all three cases was a catastrophe for the immediate environment and the people it supported – not only because the people were unprepared to cope with dramatic natural changes in their environments but because their own actions magnified the disastrous effects of those changes.

Quoting an apt bit of cynical graffiti, historical philosopher Wright (2004, p. 107) sums up what he calls the hazards of human progress this way: "Each time history repeats itself, the price goes up." Indeed, as the second decade of the 21st century begins, humans are ecosystem engineers on a planetary scale, and our global civilization threatens the life-sustaining capacity of all of Earth's environmental "spheres":

- *Geosphere (lithosphere): Earth's crust and upper mantle, containing nonrenewable fossil fuels, minerals, and nutrients that plants require.* The activities of plants, animals, and microorganisms weather mineral soils and rocks, create organic soils, and alter erosion and sedimentation rates. Humans mine minerals, metals, and gems; extract fossil fuels including coal, oil, and natural gas; and increase erosion and sedimentation by removing or altering natural plant cover through agriculture, logging, and urbanization.
- *Atmosphere: the thin envelope of gases encircling the planet.* Living systems modify the atmosphere, its temperature, and the amount of water it contains by continually generating oxygen and consuming carbon dioxide through photosynthesis and affecting the amount and forms of other gases. Humans release toxic chemicals into the air and alter the climate by raising the atmospheric concentration of greenhouse gases, such as carbon dioxide and methane, through the burning of fossil fuels in motor vehicles, ships, airplanes, and power plants.
- *Hydrosphere: Earth's liquid surface and underground water; its polar ice caps, oceanic icebergs, and terrestrial permafrost; and its atmospheric water vapor.* Living systems alter the water cycle by modifying the Earth's temperature and the amount of water plants send into the atmosphere through a process called evapotranspiration. Humans build dams, irrigation canals, drinking-water delivery systems, and wastewater treatment plants. They use water to generate electricity; they mine groundwater from dwindling underground aquifers for farming as well as drinking; they alter the flows of surface waters for everything from transportation to the mining of gold; they drain wetlands to gain land area and abate waterborne diseases. Modern human interference in global climate is likely to disrupt the entire planetary water cycle.
- *Biosphere: the totality of Earth's living systems, that part of the Earth inhabited by living organisms.* Life on Earth emerged 3.9 billion years ago and has sustained itself through changes in form, diversity, and detail since then. No planet yet discovered supports complex life as we know it on Earth. As predators, humans have decimated or eliminated wild animal populations worldwide. As domesticators of animals and plants, humans have massively reshaped landscapes by cutting forests, burning and plowing grasslands, building cities, desertifying vast areas, and overharvesting fish and shellfish. Human actions have precipitated a spasm of extinctions that today rivals five previous mass extinctions caused by astronomical or geological forces, each of which eliminated more than 70% of species then existing.

Humans themselves may be thought of as a sphere within the greater biosphere: the *ethnosphere*, or the sum total of all thoughts and intuitions, myths and beliefs, ideas and inspirations brought into being by the human imagination since the dawn of consciousness. In the words of anthropologist Davis (2009, p. 2), who coined and defined the term in 2002, "The ethnosphere is humanity's greatest legacy. It is the product of our dreams, the embodiment of our hopes, the symbol of all we are and all that we, as a wildly inquisitive and astonishingly adaptive species, have created." But, Davis notes, just as the biosphere is being severely eroded, so too is the ethnosphere, but at a much faster pace.

Today, the scientific consensus is that *H. sapiens* – a single species – rivals astronomical and geological forces in its impact on life on Earth.

Biotic Impoverishment

The first step in dealing with the present impact of human activity is to correctly identify the nature of humanity's

relationship with the environment and how human actions affect that relationship. Many people still see the environment as something people must overcome, or they regard environmental needs as something that ought to be balanced against human needs (for example, jobs vs. the environment). Most people still regard the environment as a provider of commodities or a receptacle for waste.

When asked to name humanity's primary environmental challenges, people typically think of running out of nonrenewable raw materials and energy or about water and air pollution. Environmental research and development institutions focus on ways technology can help solve each problem, such as fuel cells to provide clean, potentially renewable energy or scrubbers to curb smokestack pollution. Even when people worry about biodiversity loss, they are concerned primarily with stopping the extinction of species, rather than with understanding the underlying losses leading up to species extinctions or the broader biological crisis that extinctions signal.

These perspectives miss a crucial point: the reason pollution, energy use, extinction, and dozens of other human impacts are important is their larger impact on the biosphere. Ecosystems, particularly their living components, have always provided the capital to fuel human economies. When populations were small, humans making a living from nature's wealth caused no more disruption than other species. But with nearly 7 billion people occupying or using resources from every place on Earth, humans are overwhelming the ability of other life-forms to make a living and depleting the planet's natural wealth. At this point in the planet's history, one species is compromising Earth's ability to support the living systems that evolved here over millions of years. The systematic reduction in Earth's capacity to support life – which Woodwell (1990) termed 'biotic impoverishment' – is thus the most important human-caused environmental impact. At best, the ethics of this impact are questionable; at worst, it is jeopardizing our own survival.

The connection between biotic impoverishment and extinction is intuitively obvious. By overharvesting fish, overcutting forests, overgrazing grasslands, or paving over land for cities, we are clearly killing other organisms outright or eliminating their habitats, thereby driving species to extinction and impoverishing the diversity of life. But biotic impoverishment takes many forms besides extinction. It encompasses three categories of human impacts on the biosphere: (1) indirect depletion of living systems through alterations in physical and chemical systems, (2) direct depletion of nonhuman life, and (3) direct degradation of human life (**Table 1**). Identifying and understanding the biological significance of our actions – their effects on living systems, including our own social and economic systems – is the key to developing effective ways to manage our impacts.

Table 1 The many faces of biotic impoverishment

Indirect depletion of living systems through alterations in physical and chemical systems

1. Degradation of water (redirected flows, depletion of surface and ground water, wetland drainage, organic enrichment, destruction, and alteration of aquatic biota)
2. Soil depletion (destruction of soil structure, erosion, salinization, desertification, acidification, nutrient leaching, destruction, and alteration of soil biota)
3. Chemical contamination (land, air, and water pollution from pesticides, herbicides, heavy metals, and toxic synthetic chemicals; atmospheric ozone depletion; ocean acidification; fish kills; extinctions; biotic homogenization, and biodiversity loss; bioaccumulation; hormone disruption; immunological deficiencies, reproductive and developmental anomalies; respiratory diseases; intergenerational effects)
4. Altered biogeochemical cycles (alteration of the water cycle; nutrient enrichment; acid rain; fossil fuel combustion; particulate pollution; degradation of land and water biota; outbreaks of pests, pathogens, and red tides)
5. Global climate change (rising greenhouse gas concentrations, altered precipitation and airflow patterns, rising temperatures, effects on individual and community health, shifts among and within global ecosystems)

Direct depletion of nonhuman life

1. Overharvest of renewable resources such as fish and timber (depleted populations, extinctions, altered food webs)
2. Habitat fragmentation and destruction (extinctions, biotic homogenization, emerging and reemerging pests and pathogens, loss of landscape mosaics and connectivity)
3. Biotic homogenization (extinctions and invasions)
4. Genetic engineering (homogenization of crops, antibiotic resistance, potential extinctions and invasions if genes escape, other unknown ecological effects)

Direct degradation of human life

1. Emerging and reemerging diseases (occupational hazards, asthma and other respiratory ills, pandemics, Ebola, AIDS, hantavirus, tuberculosis, Lyme disease, West Nile fever, antibiotic resistance, diseases of overnutrition, higher death rates)
2. Loss of cultural diversity (genocide, ethnic cleansing, loss of cultural and linguistic diversity, loss of knowledge)
3. Reduced quality of life (malnutrition and starvation, failure to thrive, poverty, environmental refugees)
4. Environmental injustice (environmental discrimination and racism; economic exploitation; growing gaps between rich and poor individuals, segments of society, and nations; gender inequities; trampling of the environmental and economic rights of future generations)
5. Political instability (civil violence, especially under intransigent regimes; resource wars; international terrorism; increased number of environmental refugees)
6. Cumulative effects (environmental surprises, increased frequency of catastrophic natural events, boom-and-bust cycles, interactions between disease and biodiversity, collapse of civilizations)

Indirect Biotic Depletion

Humans affect virtually all the physical and chemical systems life depends on: water, soils, air, and the biogeochemical cycles linking them. Some human-driven physical and chemical changes have no repercussions on the biota; others do, becoming agents of biotic impoverishment.

Degradation of Water

Humans probably spend more energy, money, and time trying to control the movement and availability of water than to manage any other natural resource. In the process, we contaminate water, move water across and out of natural basins, deplete surface and groundwater; modify the timing and amount of flow in rivers, straighten or build dikes to

constrain rivers, and alter natural flood patterns. We change the amount, timing, and chemistry of fresh water reaching coastal regions, and we dry up wetlands, lakes, and inland seas. Our demands are outrunning supplies of this nonrenewable resource, and the scale of our transformations risks altering the planetary water cycle.

Physical alterations of the Earth's waters, combined with massive industrial, agricultural, and residential pollution, have taken a heavy toll on nonhuman aquatic life. By 2005 as much as one-fifth of the world's coral reefs had been destroyed, and only 30% remained healthy. Globally, the number of oceanic dead zones, where little or no dissolved oxygen exists, tripled during the last 30 years of the 20th century. The biota of freshwater systems has fared no better. A 4-year survey of the freshwater fishes inhabiting Malaysian rivers in the late 1980s found only 46% of 266 known Malaysian species. Some 40% of North America's freshwater fishes are at risk of extinction; two-thirds of freshwater mussels and crayfishes and one-third of amphibians that depend on aquatic habitats in the United States are rare or imperiled.

Humans use at least 54% of the Earth's accessible water runoff, a figure that is likely to grow to 70% by 2025. By then, more than a third of the world's population could suffer shortages of fresh water for drinking and irrigation. Groundwater aquifers in many of the world's most important crop-producing regions are being drained faster than they can be replenished: a study published in 2010 found that the rate of groundwater depletion worldwide had more than doubled from 1960 to 2000. Natural flood regimes, as in the Nile River basin, no longer spread nutrient-rich silt across floodplains to nourish agriculture; indeed, the High Dam at Aswan traps so much silt behind it that the Nile delta, essential to Egypt's present-day economy, is slumping into the sea. Whole inland seas, such as the Aral Sea in Uzbekistan, are drying up because the streams feeding them contain so little water. In addition to eliminating habitat for resident organisms, the sea's drying is bringing diseases to surrounding human populations. Indeed, diseases caused by waterborne pathogens are making a comeback even in industrialized nations.

In the past five or six decades, the number of large dams on the world's rivers grew more than seven times, to more than 40,000 today. The mammoth Three Gorges Dam across China's Yangtze River, completed in 2006, created a 660-km-long serpentine lake behind it. The dam displaced more than 1 million people and may force the relocation of another 4 million from the reservoir region, which, at 58,000 km^2, is larger than Switzerland. The dam has greatly altered ecosystems on the Yangtze's middle reaches, compounding perils already faced by prized and endemic fish and aquatic mammals. The sheer weight of the water and silt behind the concrete dam raises the risk of landslides and strains the region's geological structure, while water released from the dam eats away at downstream banks and scours the bottom. And by slowing the flow of the Yangtze and nearby tributaries, the dam drains the river's ability to flush out and detoxify pollutants from upstream industries.

Soil Depletion

Not just dirt, soil is a living system that makes it possible for raw elements from air, water, and bedrock to be physically and chemically assembled, disassembled, and reassembled with the aid of living macro- and microorganisms into the green cloak of life above ground. Accumulated over thousands of years, soil cannot be renewed in any time frame useful to humans alive today, or even to their great-grandchildren.

Humans degrade soils when they compact it, erode it, disrupt its organic and inorganic structure, turn it too salty for life, and cause desertification. Urbanization, logging, mining, overgrazing, alterations in soil moisture, air pollution, fires, chemical pollution, and leaching out of minerals all damage or destroy soils. Thanks to removal of vegetative cover, mining, agriculture, and other activities, the world's topsoils are eroded by wind and water ten to hundreds of times faster than they are renewed (at roughly 10 tons ha^{-1} $year^{-1}$).

Soils constitute the foundation of human agriculture, yet agriculture, including livestock raising, is the worst culprit in degrading soils. Agricultural practices have eroded or degraded more than 40% of present cropland. Over the last half century, some 24,000 villages in northern and western China have been overrun by the drifting sands of desertification. Besides topsoil erosion, the damage includes salting and saturation of poorly managed irrigated lands; compaction by heavy machinery and the hooves of livestock; and pollution from excessive fertilizers, animal wastes, and pesticides.

Living, dead, and decomposing organic matter is the key to soil structure and fertility. Soil depleted of organic matter is less permeable to water and air and thus less able to support either aboveground plants or soil organisms. The linkages between soil's inorganic components and the soil biota – naturalist Wilson's (1987) "little things that run the world" – are what give soil its life-sustaining capacity. A clear-cut forest patch whose soil biota has been damaged beyond recovery can no longer sustain trees, no matter how many are planted; another clear-cut patch whose soil community is intact – especially the close associations among fungi and tree roots – will support new tree growth. Destroying soil biota unleashes a whole series of impoverishing biotic effects both below and above ground.

Chemical Contamination

In 1962, Rachel Carson's landmark book, *Silent Spring*, alerted the world to the pervasiveness of synthetic chemicals produced since World War II. As many as 100,000 synthetic chemicals are in use today. True to one company's slogan, many of these have brought "better living through chemistry," providing new fabrics and lighter manufacturing materials, antibiotics, and life-saving drugs.

But industrial nations have carelessly pumped chemicals into every medium. Chemicals – as varied as prescription drugs flowing out of sewage plants, pesticides, heavy metals, and cancer-causing by-products of countless manufacturing processes – now lace the world's water, soil, and air and the bodies of all living things, including humans. Chemicals directly poison organisms; they accumulate in physical surroundings and are passed through and, in many cases, concentrated within portions of the food web. Chemicals cause cancer, interfere with hormonal systems, provoke asthma, and impair the functioning of immune systems. They have intergenerational effects, such as intellectual impairment in children whose mothers have eaten contaminated fish.

What's more, over half a century of pesticide and antibiotic overuse has bred resistance to these chemicals among insects, plants, and microbes, giving rise to new and reemerging scourges.

Many chemicals travel oceanic and atmospheric currents to sites far from their source. Sulfur emissions from the US Midwest, for example, fall to earth again as acid rain in Europe, killing forests and so acidifying streams and lakes that they, too, effectively die. China's burning of soft coal sends air pollution all the way to northwestern North America; the heavy haze hanging over China's chief farming regions may be cutting agricultural production by a third. Chlorofluorocarbons (CFCs), once widely used as refrigerants, have damaged the atmospheric ozone layer, which moderates how much ultraviolet radiation reaches the Earth, and opened ozone holes over the Arctic and Antarctic.

But even more alarming is an unprecedented acidification of the oceans that has only recently attracted the attention of major scientific research consortiums. Acid added to the world ocean by human activity has lowered the ocean's pH; it is lower now than it has been in 20 My, which translates into a 30% increase in sea-surface acidity since industrialization began. The future of marine life in an ocean acidifying at this speed and intensity looks bleak. As the concentration of hydrogen ions rises, the calcium carbonate in the shells or skeletons of organisms such as tropical corals, microscopic foraminifera, and mollusks begins to dissolve; further, more hydrogen ions combine with the calcium carbonate building blocks the organisms need, making it harder for them to extract this compound from the water and build their shells in the first place.

Although many of the most obviously deadly chemicals were banned in the 1970s, they continue to impoverish the biosphere. Polychlorinated biphenyls – stable, nonflammable compounds once used in electrical transformers and many other industrial and household applications – remain in the environment for long periods, cycling among air, water, and soils and persisting in the food web. They are found in polar bears and arctic villagers; they are implicated in reproductive disorders, particularly in such animals as marine mammals, whose long lives, thick fat layers where chemicals concentrate, and position as top predators make them especially vulnerable.

The agricultural pesticide DDT, sprayed with abandon in the 1940s and 1950s, even directly on children, had severely thinned wild birds' eggshells by the time it was banned in the United States. Populations of birds such as the brown pelican and bald eagle had dropped precipitously by the 1970s, although they have recovered enough for the species to be taken off the US endangered species list, the bald eagle in 2007 and the brown pelican in 2009. Reproduction of central California populations of the California condor, in contrast, continues to be threatened by DDT breakdown products, which, 40 years after the pesticide was banned, are still found in the sea lion carcasses the birds sometimes feed on.

Carson's book revealed the real danger of chemical pollutants: they have not simply perturbed the chemistry of water, soil, and air but harmed the biosphere as well. The list of chemicals' effects on living things is so long that chemical pollution equals humans' environmental impact in most people's minds, yet it is just one form of biotic impoverishment.

Altered Biogeochemical Cycles

All the substances found in living things – such as water, carbon, nitrogen, phosphorus, and sulfur – cycle through ecosystems in biogeochemical cycles. Human activities modify or have the potential to modify all these cycles. Sometimes the results stem from changing the amount or the precise chemistry of the cycled substance; in other cases, humans change biogeochemical cycles by changing the biota itself.

Freshwater use, dams, and other engineering ventures affect the amount and rate of river flow to the oceans and increase evaporation rates, directly affecting the water cycle and indirectly impoverishing aquatic life. Direct human modifications of living systems also alter the water cycle. In South Africa, European settlers supplemented the treeless native scrub, or fynbos, with such trees as pines and Australian acacias from similar Mediterranean climates. But because these trees are larger and thirstier than the native scrub, regional water tables have fallen sharply.

Human activity has disrupted the global nitrogen cycle by greatly increasing the amount of nitrogen fixed from the atmosphere (combined into compounds usable by living things). The increase comes mostly from deliberate addition of nitrogen to soils as fertilizer but also as a by-product of the burning of fossil fuels. Agriculture, livestock raising, and residential yard maintenance chronically add tons of excess nutrients, including nitrogen and phosphorus, to soils and water.

The additions are often invisible; their biological impacts are often dramatic. Increased nutrients in coastal waters, for example, trigger blooms of toxic dinoflagellates – the algae that cause red tides, fish kills, and tumors and other diseases in varied sea creatures. When huge blooms of algae die, they fall to the seafloor, where their decomposition so robs the water of oxygen that fish and other marine organisms can no longer live there. With nitrogen concentrations in the Mississippi River two to three times as high as they were 50 years ago, a gigantic dead zone forms in the Gulf of Mexico every summer; in summer 2010 this dead zone covered 20,000 km^2.

The burning of fossil fuels is transforming the carbon cycle, primarily by raising the atmospheric concentration of carbon dioxide. With other greenhouse gases, such as methane and oxides of nitrogen, carbon dioxide helps keep Earth's surface at a livable temperature and drives plant photosynthesis, but since the Industrial Revolution, atmospheric carbon dioxide concentrations have risen nearly 40% and are now widely thought to be disrupting the planet's climate. In addition, the effects of catastrophic oil spills like the one that followed the April 2010 explosion of the Deepwater Horizon drilling rig in the Gulf of Mexico – and the effects of the chemicals used to disperse the resulting plumes of oil – may reverberate for decades.

Global Climate Change

In its 2007 report, written and reviewed by more than 3500 scientists, the Intergovernmental Panel on Climate Change (IPCC) concluded that climate warming was "unequivocal" (p. 2), that most of the average global warming over the previous 50 years was "very likely due to anthropogenic GHG

[greenhouse gas] increases" (p. 5), and that this warming has likely had "a discernible influence at the global scale on observed changes in many physical and biological systems" (p. 6).

The concentrations of heat-trapping gases in the atmosphere are at their highest level in more than 200,000 years. The 20th century in the Northern Hemisphere was the warmest of the past millennium; the first decade of the 21st first century, and the first 7 months of 2010, were the warmest on record.

Higher global temperatures set in motion a whole series of effects, making the study of climate change, and of humans' role in it, complex and controversial. Spring arrives at least 1 week earlier in the Northern Hemisphere. Polar glaciers and ice sheets are receding. The Arctic is warming twice as fast as the rest of the planet, and arctic sea ice melted at a near-record pace in 2010. With the sun heating the newly open water, winter refreezing will take longer, and the resulting thinner ice will melt more easily the following summer.

The large-scale circulation of global air masses is changing and, with them, the large-scale cycles in ocean currents, including the periodic warming and cooling in the tropical Pacific Ocean known as El Niño and La Niña. As a result, the distribution, timing, and amount of rain and snow are also changing, making the weather seem more unpredictable than ever. Unusually warm or cold winters, massive hurricanes like those that devastated the US Gulf Coast in 2005, and weather-related damage to human life and property are all predicted to increase with global warming. In the United States in 2005, weather-related damage totaled nearly $97 billion.

Where other nutrients are not limiting, rising carbon dioxide concentrations may enhance plant photosynthesis and growth. Rising temperatures may shift the ranges of many plants and animals – both wild and domestic – rearranging the composition and distribution of the world's biomes, as well as those of agricultural systems. The resulting displacements will have far-reaching implications not only for the displaced plants and animals but also for the goods and services humans depend on from these living systems.

Direct Depletion of Nonhuman Life

From their beginnings as hunter-gatherers, humans have become highly efficient, machine-aided ecosystem engineers and predators. We transform the land so it produces what we need or want; we harvest the oceans in addition to reaping our own fields; we cover the land, even agricultural land, with sprawling cities. All these activities directly affect the ability of other life-forms to survive and reproduce. We deplete nonhuman life by eliminating some forms and favoring others; the result is a loss of genetic, population, and species diversity.

We are irreversibly homogenizing life on Earth, in effect exercising an unnatural selection that is erasing the diversity generated by millions of years of evolution by natural selection. One species is now determining which other species will survive, reproduce, and thereby contribute the raw material for future evolution.

Overharvest of Renewable Resources

In the 1930s, so many sardines were scooped from the waters off Monterey's Cannery Row in California that the population collapsed, taking other sea creatures and human livelihoods with it; after rebounding somewhat in the first decade of the 2000s, the species has still not recovered fully. According to the US National Marine Fisheries Service, nearly 80% of commercially valuable fish of known status were overfished or fished to their full potential by 1993. Atlantic commercial fish species at their lowest levels in history include tuna, marlin, cod, and swordfish. Overfishing not only depletes target species but restructures entire marine food webs.

Marine mammals, including whales, seals, sea lions, manatees, and sea otters, were so badly depleted by human hunters that one species, Steller's sea cow (*Hydrodamalis gigas*), went extinct; many other species almost disappeared. In the 19th century, Russian fur traders wiped out sea otters (*Enhydra lutris*) along the central California coast. With the otters gone, their principal prey, purple sea urchins (*Stronglyocentrotus purpuratus*), overran the offshore forests of giant kelp (*Macrocystis pyrifera*), decimating the kelp fronds and the habitat they provided for countless other marine creatures, including commercially harvested fishes. Thanks to five decades of protection, marine mammal populations began slowly recovering – only to face food shortages as regional marine food webs continue to unravel because of fishing, changing oceanic conditions, and contamination.

Timber harvest has stripped vegetation from the Amazonian rainforest to mountainsides on all continents, diminishing and fragmenting habitat for innumerable forest and stream organisms, eroding soils, worsening floods, and contributing significantly to global carbon dioxide emissions. In the Northern Hemisphere, 10% or less remains of old-growth temperate rainforests. The uniform stands of trees usually replanted after logging do not replace the diversity lost with the native forest, any more than monocultures of corn replace the diversity within native tallgrass prairies.

Habitat Fragmentation and Destruction

A great deal of human ecosystem engineering not only alters or damages the habitats of other living things but often destroys those habitats. Satellite-mounted remote-sensing instruments have revealed transformations of terrestrial landscapes on a scale unimaginable in centuries past. Together, cropland and pastures occupy 40% of Earth's land surface. Estimates of the share of land wholly transformed or degraded by humans hover around 50%. Our roads, farms, cities, feedlots, and ranches either fragment or destroy the habitats of most large carnivorous mammals. Mining and oil drilling damage the soil, remove vegetation, and pollute marine areas. Grazing compacts soil and sends silt and manure into streams, where they harm stream life.

Landscapes that have not been entirely converted to human use have been cut into fragments. In *Song of the Dodo*, writer Quammen (1996) likens our actions to starting with a fine Persian carpet and then slicing it neatly into 36 equal pieces; even if we had the same square footage, we would not have 36 nice Persian rugs – only ragged, nonfunctional fragments. And in fact, we do not even have the original square footage because we have destroyed an enormous fraction of it.

Such habitat destruction is not limited to terrestrial environments. Human channelization of rivers may remove whole segments of riverbed. In the Kissimmee River of the US

state of Florida, for example, channelization in the 1960s transformed 165 km of free-flowing river into 90 km of canal, effectively removing 35 km of river channel and drastically altering the orphaned river meanders left behind.

Wetlands worldwide continue to disappear, drained to create shoreline communities for people and filled to increase cropland. The lower 48 United States lost 53% of their wetlands between the 1700s and mid-1980s. Such losses destroy major fish and shellfish nurseries, natural flood and pollution control, and habitat for countless plants and animals.

The mosaic of habitats in, on, or near the seafloor – home to 98% of all marine species – is also being decimated. Like clear-cutting of an old-growth forest, the use of large, heavy trawls dragged along the sea bottom to catch groundfish and other species flattens and simplifies complex, structured habitats such as gravels, coral reefs, crevices, and boulders and drastically reduces biodiversity. Studies reported on by the National Research Council of the US National Academy of Sciences have shown that a single tow can injure or destroy upward of two-thirds of certain bottom-dwelling species, which may still not have recovered after a year or more of no trawling.

Habitat fragmentation and destruction, whether on land or in freshwater and marine environments, may lead directly to extinction or isolate organisms in ways that make them extremely vulnerable to natural disturbances, climate change, or further human disturbance.

Biotic Homogenization

"The one process ongoing … that will take millions of years to correct," Wilson (1994, p. 355) admonishes us, "is the loss of genetic and species diversity by the destruction of natural habitats. This is the folly our descendants are least likely to forgive us." Both deliberately and unwittingly, humans are rearranging Earth's living components, reducing diversity and homogenizing biotas around the world. The present, continuing loss of genetic diversity, of populations, and of species vastly exceeds background rates. At the same time, our global economy is transporting species worldwide at unprecedented scales.

The globe is now experiencing its sixth mass extinction, the largest since the dinosaurs vanished 65 Ma; present extinction rates are thought to be on the order of 100–1000 times those before people dominated Earth. According to the Millennium Ecosystem Assessment, a 5-year project begun in 2001 to assess the world's ecosystems, an estimated 10–15% of the world's species will be committed to extinction over the next 30 years. Approximately 20% of all vertebrates, including 33% of sharks and rays, are at risk of extinction.

At least one of every eight plant species is also threatened with extinction. Although mammals and birds typically receive the most attention, massive extinctions of plants, which form the basis of the biosphere's food webs, undermine life-support foundations. The mutualistic relationships between animals and plants, particularly evident in tropical forests, mean that extinctions in one group have cascading effects in other groups. Plants reliant on animals for pollination or seed dispersal, for example, are themselves threatened by the extinction of the animal species they depend on. Not surprisingly, some scientists view extinction as the worst biological tragedy, but extinction is just another symptom of global biotic impoverishment.

Ever since they began to spread over the globe, people have transported other organisms with them, sometimes for food, sometimes for aesthetic reasons, and most often inadvertently. With the mobility of modern societies and today's especially speedy globalization of trade, the introduction of alien species has reached epidemic proportions, causing some scientists to label it biological pollution.

Aliens – zebra mussels (*Dreissena polymorpha*) and tamarisks, or saltcedar (*Tamarix* spp.), in North America; the Red Sea sea jelly *Rhopilema nomadica* and the common aquarium alga *Caulerpa taxifolia* now choking the Mediterranean Sea; and Leidy's comb jelly (*Mnemiopsis leidyi*) of northeastern America in the Black Sea, to name just a few – are present everywhere, and they usually thrive and spread at the expense of native species. On many islands, for example, more than half the plant species are not native, and in many continental areas the figure reaches 20% or more.

The costs of such invasions, in both economic and ecological terms, are high. In the United States, for example, annual economic losses due to damage by invasive species or the costs of controlling them exceed $137 billion per year – $40 billion more than the nation's losses from weather-related damage in 2005. Furthermore, alien invasions cause extinctions and, when added to other extinctions and the deliberate monocultures of agricultural crops, homogenize biotas still more. Introduced species are fast catching up with habitat fragmentation and destruction as the major engines of ecological deterioration.

Genetic Engineering

Humans have been manipulating their crop plants and domesticated animals for 10,000 years or so – selecting seeds or individuals and breeding and cross-breeding them. The goal was something better, bigger, tastier, hardier, or all of the above; success was sometimes elusive, but the result was crop homogenization. Of the myriad strains of potatoes domesticated by South American cultures, for example, only one was accepted and cultivated when potatoes first made it to Europe. The new crop made it possible to feed more people from an equivalent area of land and initially staved off malnutrition. But the strain succumbed to a fungal potato blight in the 1800s. Had more than one strain been cultivated, the tragic Irish potato famines might have been averted.

In the last few decades of the 20th century, people began to manipulate genes directly using the tools of molecular biotechnology, even cloning sheep and cows from adult body cells. US farmers routinely plant their fields with corn whose genetic material incorporates a bacterial gene resistant to certain pathogens. More than 40 genetically altered crops have been approved for sale to US farmers since 1992, with genes borrowed from bacteria, viruses, and insects. The United States accounts for nearly two-thirds of biotechnology crops planted globally. Worldwide in 2003, 67.6 million hectares in 18 countries on six continents were planted with genetically modified crops, as compared with 1.7 million hectares in six countries in 1996 – a 40-fold expansion in 8 years.

Biotechnologists focus on the potential of this new-millennium green revolution to feed the growing world

population, which has added almost 1 billion people in the past decade alone. But other scientists worry about unknown human and ecological health risks; these concerns have stirred deep scientific and public debate, especially in Europe, akin to the debate over pesticides in Rachel Carson's time.

One worrisome practice is plant genetic engineers' technique of attaching the genes they want to introduce into plants to an antibiotic-resistant gene. They can then easily select plants that have acquired the desired genes by treating them with the antibiotic, which kills any nonresistant plants. Critics worry that the antibiotic-resistant genes could spread to human pathogens and worsen an already growing antibiotic-resistance problem. Another concern arises from allergies humans might have or develop in response to genetically modified foods.

Although supporters of genetic engineering believe that genetically altered crops pose few ecological risks, ecologists have raised a variety of concerns. Studies in the late 1990s indicated that pollen from genetically engineered Bt corn can kill monarch butterfly caterpillars. Bt is a strain of bacterium that has been used since the 1980s as a pesticidal spray; its genes have also been inserted directly into corn and other crops. Ecologists have long worried that genetically engineered plants could escape from fields and cross-breed with wild relatives. Studies in radishes, sorghum, canola, and sunflowers found that genes from an engineered plant could jump to wild relatives through interbreeding. The fear is that a gene conferring insect or herbicide resistance might spread through wild plants, creating invasive superweeds that could potentially lower crop yields and further disturb natural ecosystems.

In fact, herbicide-resistant turf grass tested in Oregon in 2006 did escape and spread, and transgenic canola has been appearing throughout the US state of North Dakota, which has tens of thousands of hectares in conventional and genetically modified canola. According to the scientists who discovered the transgenic escapees growing in North Dakota – far from any canola field – the plants are likely to be cross-pollinating in the wild and swapping introduced genes; the plants' novel gene combinations indicate that the transgenic traits are stable and evolving outside of cultivation.

Genetically engineered crops do confer some economic and environmental benefits: for farmers, higher yields, lower costs, savings in management time and gains in flexibility; for the environment, indirect benefits from using fewer pesticides and herbicides. But it is still an open question whether such benefits outweigh the potential ecological risks or whether the public will embrace having genetically modified foods as staples in their diet.

Direct Degradation of Human Life

Human biotic impacts are not confined to other species; human cultures themselves have suffered from the widening circles of indirect and direct effects humans have imposed on the rest of nature. Over the past hundred years, human technology has cut both ways with regard to public health. Wonder drugs controlled common pathogens at the same time that natural selection strengthened those pathogens' ability to resist the drugs. Reservoirs in the tropics made water supplies more reliable for humans but also created ideal environments for human parasites. Industrialization exposed human society to a remarkable array of toxic substances.

Although man's inhumanity to man has been both fact and the subject of discourse for thousands of years, the discussions have mostly been removed from any environmental context. Few people today regard social ills as environmental impacts or humans as part of a biota. But diminished societal well-being – whether manifest in high death rates or poor quality of life – shares many of its roots with diminished nonhuman life as a form of biotic impoverishment.

Emerging and Reemerging Diseases

The intersection of the environment and human health is the core of the discipline known as environmental health. Among the environmental challenges to public health are the direct effects of toxic chemicals; occupational health threats, including exposures to hazardous materials on the job; and sanitation and disposal of hazardous wastes.

Exploitation of nonrenewable natural resources – including coal mining, rock quarrying or other mining operations, and petroleum extraction and refining – often chronically impairs workers' health and shortens their lives. Farmworkers around the world suffer long-term ills from high exposures to pesticides and herbicides. Partly because of increased air pollution, asthma rates are rising, particularly in big cities. Synthetic volatile solvents are used in products from shoes to semiconductors, producing lung diseases and toxic wastes. Nuclear weapons production starting in World War II, and associated contamination, have been linked to a variety of illnesses, including syndromes neither recognized nor understood at the time and whose causes were not diagnosed until decades afterward. The grayish metal beryllium, for example, was used in weapons production and was found decades later to scar the lungs of workers and people living near toxic waste sites.

Infectious diseases have challenged human populations throughout history, playing a significant role in their evolution and cultural development over the past 10,000 years, with 75% of human diseases linked to wildlife or domestic animals. The 20th century brought major successes in eradicating such infectious diseases as smallpox, polio, and many waterborne illnesses. But toward the century's end, emerging and reemerging diseases were again reaching pandemic proportions.

Infectious diseases thought to be on the wane – including tuberculosis, malaria, cholera, diptheria, leptospirosis, encephalitis, and dengue fever – have begun a resurgence. In addition, seemingly new scourges – Ebola virus, hantavirus, HIV/AIDS, Lyme disease, and West Nile fever – are also spreading, often, it appears, from wild animal hosts to humans as people encroach further upon previously undisturbed regions. A number of studies over the decade before 2010 show complex connections between biodiversity and disease: biodiversity loss often increases disease transmission, as in Lyme disease and West Nile fever, but diverse ecosystems also serve as sources of pathogens. Overall, however, the studies indicate that preserving intact ecosystems and their endemic biodiversity often reduces the prevalence of infectious diseases.

Human migrations – including their modern incarnation through air travel – also accelerate pathogen traffic and launch global pandemics, such as the 2003 outbreak of severe acute respiratory syndrome and the 2009 swine flu outbreak caused by the H1N1 virus. Even something as simple and apparently benign as lighting can become an indirect agent of disease. Artificial lighting, especially in the tropics, for example, can alter human and insect behavior in ways that speed transmission of insect-borne diseases, such as Chagas's disease, malaria, and leishmaniasis.

In addition, especially in highly developed countries such as the United States, diseases of affluence and overconsumption are taking a toll. Heart disease is the number one cause of death in the United States; overnutrition, obesity, and diabetes due to sedentary, technology-driven lifestyles, particularly among children, are chronic and rising. One estimate put the share of US children considered overweight or obese at one in three. This rise in obesity rates has been stunningly rapid. As recently as 1980, just 15% of adults were obese; by 2008 the rate had hit 34%, and two-thirds of Americans are now considered either overweight or obese.

Loss of Cultural Diversity

Although not conventionally regarded as elements of biodiversity, human languages, customs, agricultural systems, technologies, and political systems have evolved out of specific regional environments. Like other organisms' adaptive traits and behaviors, these elements of human culture constitute unique natural histories adapted, like any natural history, to the biogeographical context in which they arose. Yet modern technology, transportation, and trade have pushed the world into a globalized culture, thereby reducing human biological and cultural diversity.

Linguists, for example, are predicting that at least half of the 7000 languages spoken today will become extinct in the 21st century. With the spread of Euro-American culture, unique indigenous human cultures, with their knowledge of local medicines and geographically specialized economies, are disappearing even more rapidly than the natural systems that nurtured them. This loss of human biodiversity is in every way as troubling as the loss of nonhuman biodiversity.

Reduced Quality of Life

The effects of environmental degradation on human quality of life are another symptom of biotic impoverishment. Food availability, which depends on environmental conditions, is a basic determinant of quality of life. Yet according to the World Health Organization, nearly half the world's population suffers from one of two forms of poor nutrition: undernutrition or overnutrition. A big belly is now a symptom shared by malnourished children, who lack calories and protein, and overweight residents of the developed world, who suffer clogged arteries and heart disease from eating too much.

Independent of race or economic class, declining quality of life in today's world is manifest in symptoms such as increased asthma in the United States caused by environmental contaminants and the high disease rates in the former Soviet Bloc after decades of unregulated pollution. Even with explicit legal requirements that industries release information on their toxic emissions, many people throughout the world still lack both information and the decision-making power that would give them any control over the quality of their lives.

Aggrieved about the degraded environment and resulting quality of life in his homeland, Ogoni activist Ken Saro-Wiwa issued a statement shortly before he was executed by the Nigerian government in 1995 saying, "The environment is man's first right. Without a safe environment, man cannot exist to claim other rights, be they political, social, or economic." Kenyan Maathai (2009, p. 249), 2004 winner of the Nobel Peace Prize, has also written, "[I]f we destroy it, we will undermine our own ways of life and ultimately kill ourselves. This is why the environment needs to be at the center of domestic and international policy and practice. If it is not, we don't stand a chance of alleviating poverty in any significant way."

Having ignored this kind of advice for decades, nations are seeing a new kind of refugee attempting to escape environmental degradation and desperate living conditions; the number of international environmental refugees exceeded the number of political refugees around the world for the first time in 1999. Environmental refugees flee homelands devastated by flooding from dam building, extraction of mineral resources, desertification, and unjust policies of national and international institutions. Such degradation preempts many fundamental human rights, including the rights to health, livelihood, culture, privacy, and property.

People have long recognized that human activities that degrade environmental conditions threaten not only the biosphere but also humans' own quality of life. As early as 4500 years ago in Mesopotamia and South Asia, writings revealed an awareness of biodiversity, of natural order among living things, and of consequences of disrupting the biosphere. Throughout history, even as civilization grew increasingly divorced from its natural underpinnings, writers, thinkers, activists, and people from all walks of life have continued to see and extol the benefits of nature to humans' quality of life.

Contemporary society still has the chance to relearn how important the environment is to quality of life. It is encouraging that the United Steelworkers of America in 1990 released a report recognizing that protecting steelworker jobs could not be done by ignoring environmental problems and that the destruction of the environment may pose the greatest threat to their children's future. It is also encouraging that the 2007 Nobel Peace Prize was awarded to a political figure and a group of scientists for their work on climate change.

Environmental Injustice

Making a living from nature's wealth has consistently opened gaps between haves and have-nots, between those who bear the brunt of environmental damage to their home places and those who do not, and between the rights of people alive now and those of future generations; these disparities too are part of biotic impoverishment. Inequitable access to "man's first right" – a healthy local environment – has come to be known as environmental injustice.

Environmental injustices, such as institutional racism, occur in industrial and nonindustrial nations. Injustice can be overt, as when land-use planning sites landfills, incinerators, and hazardous waste facilities in minority communities, or when environmental agencies levy fines for hazardous waste

violations that are lower in minority communities than in white communities.

Less overt, but no less unjust, is the harm done to one community when unsound environmental practices benefit another, as when clear-cut logging in the highlands of northwestern North America benefits logging communities while damaging the livelihoods of lowland fishing communities subjected to debris flows, sedimentation, and downstream flooding.

The plight of the working poor and the disparities between rich and poor are also examples of biotic impoverishment within the human community. According to the United Nations Research Institute for Social Development, the collective wealth of the world's 358 billionaires equaled the combined income of the poorest 2.4 billion people in 1994. *Forbes Magazine* put the number of billionaires in early 2010 at 1011, with a total worth of $3.6 trillion.

In the United States during the last decade of the 20th century, the incomes of poor and middle-class families stagnated or fell, despite a booming stock market. The Center on Budget and Policy Priorities and the Economic Policy Institute reported that between 1988 and 1998, earnings of the poorest fifth of American families rose less than 1%, while earnings of the richest fifth jumped 15%. By the middle of the first decade of the 21st century, Americans' income inequality had become the widest among industrialized nations, with the wealthiest 20% of the population holding 85% of the wealth. The wealthiest Americans continued to prosper even during the global recession late in that decade, while the less well-off kept losing ground.

But perhaps the grossest example of human and environmental domination leading to continued injustice is the creation of a so-called third world to supply raw materials and labor to the dominant European civilization after 1500 and the resulting schism between today's developed and developing nations. Developing regions throughout the world held tremendous stores of natural wealth, some of it – like petroleum – having obvious monetary value in the dominant economies and some having a value invisible to those economies – like vast intact ecosystems. A 2010 United Nations study (TEEB) estimated that even today, Earth's ecosystems account for roughly half to 90% of the source of livelihoods for rural and forest-dwelling peoples; the study calls this value the gross domestic product (GDP) of the poor.

Dominant European civilizations unabashedly exploited this natural wealth and colonized or enslaved the people in whose homelands the wealth was found. But the dominant civilizations also exported their ways of thinking and their economic models to the developing world, not only colonizing places but also effecting what Wangari Maathai has called a colonization of the mind. Although dominant 21st century society tends to dismiss ancient wisdom as irrelevant in the modern world, perhaps the cruelest impoverishment of all is the cultural and spiritual deracination experienced by exploited peoples worldwide.

Exploitation of poor nations and their citizens by richer, consumer countries – and in many cases by the same governments that fought for independence from the colonists while adopting the colonists' attitudes and economic models – persists today in agriculture, wild materials harvesting, and textile and other manufacturing sweatshops. In the mid-1990s, industrial countries consumed 86% of the globe's aluminum, 81% of its paper, 80% of its iron and steel, 75% of its energy, and 61% of its meat; they are thus responsible for most of the environmental degradation associated with producing these goods. Most of the actual degradation, however, still takes place in developing nations.

As a result, continuing environmental and social injustice – environmental and social impoverishment perpetrated by outsiders and insiders alike – pervades developing nations. Such impoverishment can take the form of wrenching physical dislocation like the massive displacements enforced by China's Three Gorges Dam. It can appear as environmental devastation of homelands and murder of the people who fought to keep their lands, as in the Nigerian government–backed exploitation of Ogoniland's oil reserves by the Shell Petroleum Development Corporation. After Saro-Wiwa's execution, the Ogoni were left, without a voice, to deal with a scarred and oil-polluted landscape.

Despite great advances in the welfare of women and children over the past century, poverty still plagues both groups. Children from impoverished communities, even in affluent nations, suffer from the lethargy and impaired physical and intellectual development known as failure to thrive. Poverty forces many children to work the land or in industrial sweatshops; lack of education prevents them from attaining their intellectual potential. This impoverishment in the lives of women and children is as much a symptom of biotic impoverishment as are deforestation, invasive alien organisms, or species extinctions.

Little by little, community-based conservation and development initiatives are being mounted by local citizens to combat this impoverishment: Witness Maathai's Green Belt Movement, which began with tree planting to restore community landscapes and provide livelihoods for residents, and the rise of ecotourism and microlending (small loans made to individuals, especially women, to start independent businesses) as ways to bring monetary benefits directly to local people without further damaging their environments. Ultimately, one could see all efforts to protect the ethnosphere and the biosphere as a fight for the rights of future generations to an environment that can support them.

Political Instability

Only during the last two decades of the 20th century did environmental issues find a place on international diplomatic agendas, as scholars began calling attention to – and governments began to see – irreversible connections between environmental degradation and national security. British scholar Myers (1993), noting that environmental problems were likely to become predominant causes of conflict in the decades ahead, was one of the first to define a new concept of environmental security. National security threatened by unprecedented environmental changes irrespective of political boundaries will require unprecedented responses altogether different from military actions, he warned. Nations cannot deploy their armies to hold back advancing deserts, rising seas, or the greenhouse effect.

Canadian scholar Homer-Dixon (1999) showed that environmental scarcities – whether created by ecological constraints or sociopolitical factors including growing populations, depletion of renewable resources such as fish or timber, and environmental injustice perpetrated by one segment of a population on another – were fast becoming a permanent, independent cause of civil strife and ethnic violence. He found that such scarcity was helping to drive societies into a self-reinforcing spiral of dysfunction and violence, including terrorism. Environmental and economic injustices worldwide leave no country immune to this type of threat.

Typically, diplomacy has stalled in conflicts over natural resources: arguments over water rights have more than once held up Israeli-Palestinian peace agreements; fights over fish erupted between Canada and the United States, Spain, and Portugal. In contrast, in adopting the Montreal Protocol on Substances That Deplete the Ozone Layer in 1987, governments, nongovernmental organizations, and industry successfully worked together to safeguard part of the environmental commons. The treaty requires signatory nations to curb their use of CFCs and other ozone-destroying chemicals and has been, according to former United Nations secretary general Kofi Annan, perhaps the most successful international agreement to date.

Cumulative Effects

If scientists have learned anything about the factors leading to biotic impoverishment, they have learned that the factors' cumulative effects can take on surprising dimensions. As scholars like Fagan (1999) and Diamond (2005) have chronicled, the multiple stresses of global climatic cycles such as El Niño, natural events like droughts or floods, resource depletion, and social upheaval have shaped the fates of civilizations.

Societies as far-flung as ancient Egypt, Peru, the American Southwest, and Easter Island prospered and collapsed because of unwise management of their environments. The city of Ubar, built on desert sands in what is now southern Oman, literally disappeared into the sinkhole created by drawing too much water out of its great well. In modern Sahelian Africa, a combination of well digging and improved medical care and sanitation led to a threefold population increase; sedentary ways, heavy taxes imposed by a colonial government, and an impoverished people took the place of a nomadic culture evolved within the desert's realities.

During the first decade of the 21st century, numerous natural disasters befell nations around the world: wildfires in Australia, Bolivia, Canada, and Russia; flooding in China, India, Romania, and West Africa; devastating hurricanes and typhoons in the Caribbean, Philippines, Taiwan, and southeastern United States; catastrophic landslides and floods in China, Guatemala, Pakistan, and Portugal; and destructive earthquakes in Chile, China, Haiti, Indonesia, and Pakistan. Neither the rains nor the earthquakes were caused by human activity, but the cumulative effects of human land uses and management practices – from dikes separating the Mississippi from its floodplain to deforestation in Haiti – made the losses of human life and property much worse than they might have been otherwise.

Root Causes of Human Impact

The ultimate cause of humans' massive environmental impact is our individual and collective consumptive and reproductive behavior, which has given us spectacular success as a species. But the very things that enabled humans to thrive in nearly every environment have magnified our impacts on those environments, and the technological and political steps we take to mitigate our impacts often aggravate them. Too many of us simply take too much from the natural world and ask it to absorb too much waste.

Fragmented Worldviews, Fragmented Worlds

For most of human history, people remained tied to their natural surroundings. Even as agriculture, writing, and technology advanced, barriers of geography, language, and culture kept humans a diverse lot, each group depending on mostly local and regional knowledge about where and when to find resources necessary for survival. Their worldviews, and resulting economies, reflected this dependency.

For example, in northwestern North America starting about 3000 years ago, a native economy centered on the abundance of salmon. At its core was the concept of the gift and a belief system that treated all parts of the Earth – animate and inanimate – as equal members of a community. In this and other ancient gift economies, a gift was not a possession that could be owned; rather, it had to be passed on, creating a cycle of obligatory returns. Individuals or tribes gained prestige through the size of their gifts, not the amount of wealth they accumulated.

This system coevolved with the migratory habits of the salmon, which moved en masse upriver to spawn each year. Because the Indians viewed salmon as equals to themselves, killing salmon represented a gift of food from salmon to people. Fishers were obligated to treat salmon with respect or risk losing this vital gift. The exchange of gifts between salmon and humans – food for respectful treatment – minimized waste and overharvest and ensured a continuous supply of food. Further, the perennial trading of gifts among the people effectively redistributed the natural wealth brought each year by fluctuating populations of migrating fish, leveling out the boom-and-bust cycles that usually accompany reliance on an uncertain resource.

In modern times, the gift economy along with the egalitarian worldview that accompanied it, has been eclipsed by a redistributive economy tied not to an exchange of gifts with nature but to the exploitation of nature and to the technologies that enhance that exploitation. Nature became a resource for humans, rather than an equal to humans. In economic terms, natural resources fell under the heading of 'land' in an economic trinity comprising three factors of production: land, labor, and capital. Land and resources, including crops, became commodities – expendable or easily substitutable forms of capital – whose value was determined solely by their value in the human marketplace.

In 1776 Adam Smith published his famous *Inquiry into the Nature and Causes of the Wealth of Nations*, in which he argued that society is merely the sum of its individuals, that the social good is the sum of individual wants, and that the

market (the so-called invisible hand) automatically guides individual behavior to the common good. Crucial to his theories were division of labor and the idea that all the factors of production were freely mobile. His mechanistic views created an economic rationale for no longer regarding individuals as members of a community linked by ethical, social, and ecological bonds.

About the same time, fueling and fueled by the beginnings of the Industrial Revolution, the study of the natural world was morphing into modern physics, chemistry, geology, and biology. Before the mid-19 century, those who studied the natural world – early 19 century German biogeographer Baron Alexander von Humboldt and his disciple Charles Darwin among them – took an integrated view of science and nature, including humans. Both scientists regarded the understanding of the complex interdependencies among living things as the noblest and most important result of scientific inquiry.

But this integrated natural philosophy was soon supplanted by more atomistic views, which fit better with industrialization. Mass production of new machines relied on division of labor and interchangeable parts. Like automobiles on an assembly line, natural phenomena too were broken down into their supposed component parts in a reductionism that has dominated science ever since. Rushing to gain in-depth, specialized knowledge, science and society lost sight of the need to tie the knowledge together. Disciplinary specialization replaced integrative scholarship.

Neoclassical economics, which arose around 1870, ushered in the economic worldview that rules today. A good's value was no longer tied to the labor required to make it but derived instead from its scarcity. A good's price was determined only by the interaction of supply and demand. As part of 'land,' natural resources therefore became part of the human economy, rather than the material foundation making the human economy possible. Because of its doctrine of infinite substitutability, neoclassical economics rejects any limits on growth; forgotten are the classical economic thinkers and contemporaries of von Humboldt, including Thomas Malthus and John Stuart Mill, who saw limits to the growth of human population and material well-being.

Consequently, the 19th and 20th centuries saw the rise to dominance of economic indicators that fostered the economic invisibility of nature, misleading society about the relevance of Earth's living systems to human well-being. Among the worst offenders in this regard are gross national product (GNP) and its cousin, GDP. GNP measures the value of goods and services produced by a nation's citizens or companies, regardless of their location around the globe. GDP, in contrast, measures the value of goods and services produced within a country's borders, regardless of who generates those goods and services.

In effect, both GNP and GDP measure money changing hands, no matter what the money pays for; they make no distinction between what is desirable and undesirable, between costs and benefits. Both indicators ignore important aspects of the economy like unpaid work or nonmonetary contributions to human fulfillment – parenting, volunteering, checking books out of the library. Worse, the indicators also omit social and environmental costs, such as pollution, illness, or resource depletion; they only add and do not subtract. GDP math adds in the value of paid daycare or a stay in the hospital and ignores the value of unpaid parenting or being cared for at home by friends. It adds in the value of timber sold, but fails to subtract the losses in biodiversity, watershed protection, or climate regulation that come when a forest is cut.

Efforts have been made in the past few decades to create less blinkered economic indicators. Social scientists Herman Daly and John Cobb in 1989 developed an index of sustainable economic welfare, which adjusts the United States' GNP by adding in environmental good things and subtracting environmental bad things. Public expenditures on education, for example, are weighted as "goods" while costs of pollution cleanup, depletion of natural resources, and treating environment-related illnesses are counted as 'bads.' Unlike the soaring GDP of recent decades, this index of sustainable economic welfare remained nearly unchanged over the same period.

Still other work aims to reveal nature's worth in monetary terms by assigning dollar values to ecological goods and services. A 1997 study by ecologist David Pimentel and colleagues calculated separate values for specific biological services, such as soil formation, crop breeding, or pollination; by summing these figures, these researchers estimated the total economic benefits of biodiversity for the United States at $319 billion – 5% of US GDP at the time – and for the world at $2928 billion. A 2000 analysis by Pimentel and colleagues reported that the approximately 50,000 nonnative species in the United States cause major environmental damage and reparation costs amounting to $137 billion a year.

As part of the United Nations International Year of Biodiversity in 2010, several studies have translated the value of the world's ecosystems into dollar values. One report estimated the worth of crucial ecosystem services delivered to humans by living systems at $21–72 trillion per year – comparable to a world gross national income of $58 trillion in 2008. Another study reported, among other things, that as many as 500 million people worldwide depend on coral reefs – valued between $30 billion and $172 billion a year – for fisheries, tourism, and protection from ocean storms and high waves, services threatened by warmer and more acidic seas.

Although a monetary approach does not create a comprehensive indicator of environmental condition, it certainly points out that ecological values ignored by the global economy are enormous.

Too Many Consuming Too Much

The US Census Bureau predicts that by 2012, the global human population will top 7 billion – having grown by nearly 1 billion people during the decade between the first and second editions of this encyclopedia. From the appearance of *H. sapiens* about 200,000 years ago, it took humans until 1804 to reach their first billion, 123 years to double to 2 billion, and 33 years to achieve 3 billion. Human population doubled again from 3 to 6 billion in about 40 years – before most post-World War II baby boomers reached retirement age. Even with fertility rates declining in developed countries, China, and some developing countries where women are gaining education and economic power, and with pandemics like AIDS claiming more lives, the Census Bureau predicts that world population will reach 9 billion by 2044.

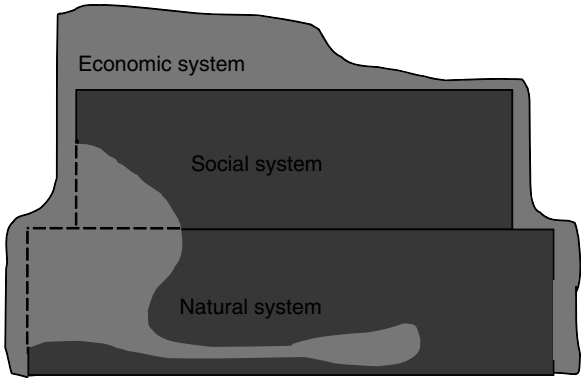


Figure 1 Relationships among the natural, social, and economic systems on Earth. Human economies may be thought of as icing atop a two-layer cake. The economic icing is eroding the human social and natural layers beneath it, threatening the foundation and sustainability of all three systems. Modified from Karr JR and Chu EW (1995) Ecological integrity: Reclaiming lost connections. In: Westra L and Lemons J (eds.) *Perspectives in Ecological Integrity*, pp 34–48. Dordrecht, Netherlands: Kluwer Academic.

Humans appropriate about 40% of global plant production, 54% of Earth's freshwater runoff, and enough of the ocean's bounty to have depleted 63% of assessed marine fish stocks. In energy terms, one person's food consumption amounts to 2500–3000 calories a day, about the same as that of a common dolphin. But with all the other energy and materials humans use, global per capita energy and material consumption have soared even faster than population growth over the past 50 years. Now, instead of coevolving with a natural economy, global society is consuming the foundations of that economy, impoverishing Earth's living systems, and undermining the foundations of its own existence (**Figure 1**).

Measuring Environmental Impacts

For most of the 20th century, environmental measurements, or indicators, tracked primarily two classes of information: counts of activities directed at environmental protection and the supply of products to people. Regulatory agencies are typically preoccupied with legislation, permitting, or enforcement, such as the numbers of environmental laws passed, permits issued, enforcement actions taken, or treatment plants constructed. Resource protection agencies concentrate on resource harvest and allocation. Water managers, for example, measure water quantity; they allocate water to domestic, industrial, and agricultural uses but seldom make it a priority to reserve supplies for sustaining aquatic life, to protect scenic and recreational values, or simply to maintain the water cycle. Foresters, farmers, and fishers count board-feet of timber, bushels of grain, and tons of fish harvested. Governmental and nongovernmental organizations charged with protecting biological resources also keep counts of threatened and endangered species.

As in the parable of the three blind men and the elephant – each of whom thinks the elephant looks like the one body part he can touch – these or similar indicators measure only one aspect of environmental quality. Counting bureaucratic achievements focuses on actions rather than on information about real ecological status and trends. Measurements of resource supply keep track of commodity production, not necessarily a system's capacity to continue supplying that commodity. And measuring only what we remove from natural systems, as if we were taking out the interest on a savings account, overlooks the fact that we are usually depleting principal as well.

Even biologists' counts of threatened and endangered species – which would seem to measure biotic impoverishment directly – still focus narrowly on biological parts, not ecological wholes. Enumerating threatened and endangered species is just like counting any other commodity. It brings our attention to a system already in trouble, perhaps too late. And it subtly reinforces our view that we know which parts of the biota are most important.

Society needs to rethink its use of available environmental indicators, and it needs to develop new indicators that represent current conditions and trends in the systems humans depend on (**Table 2**). It particularly needs objective measures more directly tied to the condition, or health, of the environment so that people can judge whether their actions are compromising that condition.

Such measures should be quantitative, yet easy to understand and communicate; they should be cost-effective and applicable in many circumstances. Unlike narrow criteria tracking only administrative, commodity, or endangered species numbers, they should provide reliable signals about status and trends in ecological systems. Ideally, effective indicators should describe the present condition of a place, aid in diagnosing the underlying causes of that condition, and make predictions about future trends. They should reveal not only risks from present activities but also potential benefits from alternative management decisions.

Most important, these indicators should, either singly or in combination, give information explicitly about living systems. Measurements of physical or chemical factors can sometimes act as surrogates for direct biological measurements, but only when the connection between those measures and living systems is clearly understood. Too often we make assumptions – when water managers assume that chemically clean water equals a healthy aquatic biota, for example – that turn out to be wrong and fail to protect living systems.

General Sustainability Indexes

As environmental concerns have become more urgent – and governmental and nongovernmental organizations have struggled to define and implement the concept of sustainable development – the effort has grown to create indicator systems that explicitly direct the public's and policymakers' attention to the value of living things. Moving well past solely economic indexes like GDP, several indexes now integrate ecological, social, and economic well-being.

The index of environmental trends for nine industrialized countries, developed by the nonprofit National Center for Economic and Security Alternatives, incorporates ratings of air, land, and water quality; chemical and waste generation; and

Table 2 Plausible indicators of environmental quality[a]

Indirect depletion of living systems through alterations in physical and chemical environments

1. Degradation of water (chemical contaminant concentrations, river flows, rainfall, runoff)
2. Soil depletion (erosion rates, desertification rates, salt accumulation in soils)
3. Chemical contamination (pollutant and toxic emissions, pollutant and toxic concentrations in air, water, soil, and living organisms)
4. Altered biogeochemical cycles (river flows and lake levels; amount of nutrients going into water bodies, or nutrient loading; nutrient concentrations in water bodies; chlorophyll concentrations, reflecting nutrient-triggered algal blooms; oxygen depletion in water bodies; trophic status of lakes; changes in air and soil chemistry; atmospheric greenhouse gas concentrations)
5. Global climate change (atmospheric greenhouse gas concentrations, change in atmospheric temperatures, distribution and intensity of severe storms or droughts)

Direct depletion of nonhuman life

1. Overharvest of renewable resources such as fish and timber (tons of fish harvested; for a given anadromous fish population, number of adult fish returning to rivers to spawn; hatchery fish released and recovered; board-feet of timber harvested; forest regrowth rates; quantity of standing timber; ecological footprints)
2. Habitat fragmentation and destruction (area of remaining grassland, wetland, and other habitats; landscape connectivity; rates of habitat destruction)
3. Biotic homogenization (number of extinct, threatened, and endangered taxonomic groups; spread of nonindigenous species; local or regional diversity; damage and reparation costs due to invasions or extinctions; major relocations in species distributions)
4. Genetic engineering (diversity among cultivated crop strains, genetic diversity within strains, escape of genetically engineered organisms or traits to wild populations)

Direct degradation of human life

1. Emerging and reemerging diseases (death or infection rates caused by diseases, geographic spread of diseases, recovery rates, frequency and spread of resistance to antibiotics and other drugs)
2. Loss of cultural diversity (extinction of cultures, death of languages)
3. Reduced quality of life (population size and growth; starvation, malnutrition, and obesity rates; infant mortality rates; teen pregnancy rates; literacy rates; rates of stress and other diseases of affluence; length of work week; child or other forced labor; changes in death rates or average life spans)
4. Environmental injustice (siting of toxic waste dumps or waste emissions relative to resident communities, economic exploitation of certain groups, worker strikes, wage and income gaps, unemployment rates for different economic sectors)
5. Political instability (frequency of domestic and international strife, environmental terrorism rates, number of environmental refugees, genocide)
6. Cumulative effects (frequency of catastrophic natural disasters; costs of weather-related property damage; human death tolls; government subsidies of environmentally destructive activities such as fishery overcapitalization, below-cost timber sales, water projects, and agricultural supports; replacement costs for ecological services; pricing that reflects environmental costs; ''green'' taxes; rise in polycultural agricultural practices; number of organic farms)

[a] These indicators have been or could be used to monitor status and trends in environmental quality, including dimensions of biotic impoverishment. Without a full spectrum of indicators, however, and without coupling them to direct measures of biological condition, only a partial view of the degree of biotic impoverishment will emerge.

energy use since 1970. By its 1995 rankings, environmental quality in the United States had gone down by 22% since 1970, while Denmark had declined by 11%.

In 2000, world leaders, supported by the United Nations Development Programme, defined a set of eight millennium development goals to be attained by 2015, which combine poverty, education, employment, and environmental sustainability. They include human rights and health goals – such as universal primary education, gender equality, and combating AIDS and other diseases – as well as ensuring environmental sustainability. In the decade since the program began, progress has been made in reducing poverty; expanding educational opportunities for children, especially in Africa; slowing deforestation; and providing better water for many people, among others. But progress controlling climate change, halting biodiversity loss, and achieving gender equality has lagged.

A later environmental performance index spearheaded by Yale and Columbia universities ranks 163 countries on 25 performance indicators in 10 well-established policy categories to determine whether and how well countries are meeting established environmental goals. This index incorporates measurements addressing environmental stresses to human health plus ecosystem health and natural resource management, collectively referred to as ecosystem vitality. The categories representing environmental health include the environmental burden of disease and the effects of air and water quality on human health. Those representing ecosystem vitality include air and water resources for ecosystems; biodiversity and habitat; and climate change, among others.

Although hampered by data gaps and the inability to track changes in countries' performance through time, the index provides a sense of countries' relative performance with regard to the environmental challenges faced by all. Top performers in the 2010 rankings included Iceland, Switzerland, Costa Rica, and Sweden. The United States ranked 61st well below Japan and nations in the European Union.

Ecological Footprints

A resource-accounting approach pioneered in the 1990s by geographers Rees and Wackernagel (1996) translates humans' impact on nature, particularly resource consumption, into a metaphorical ecological footprint. This accounting estimates the area required by a city, town, nation, or other human community to produce consumed resources and absorb generated wastes; it then compares the physical area occupied by a city or country with the area required to supply that city or country's needs. The 29 largest cities of Baltic Europe, for example, appropriate areas of forest, agricultural, marine, and wetland ecosystems that are at least 565–1130 times larger than the areas of the cities themselves.

According to the Global Footprint Network, national ecological footprints in 2010 ranged from a high of 10.7 hectares per person for the United Arab Emirates to 0.4 hectares per person for Timor-Leste and 0.6 for Afghanistan and Bangladesh. The United States' ecological footprint – 8.0 hectares per person – tied for fourth among 152 nations with populations of at least 1 million. One-hundred and four of these nations operate under ecological deficits; that is, their

consumption exceeds the biological capacity of their lands and waters to provide needed resources and absorb their wastes. At their present rates of consumption, these nations are therefore overexploiting either their own resources or those of other nations.

By ecological footprint accounting, raising the nearly 7 billion people on Earth in the early 21st century to living standards – and thus ecological footprints – equal to those in the United States would require four planets more than the only one we have. Clearly, humans are consuming more resources, and discarding more waste, than Earth's living systems can produce or absorb in a given time period. This gap is the global sustainability gap the world now faces.

Measuring the State of Living Systems

Most environmental indexes and accounting systems are still human centered; they still do not measure the condition of the biosphere itself. We may know that biodiversity's services are worth huge sums of money and that our hometown's ecological footprint is much bigger than the town's physical footprint, but how do we know whether specific actions damage living systems or that other actions benefit them? How do we know if aggregate human activity is diminishing life on Earth? To answer this question, we need measures that directly assess the condition of living systems.

Biological assessment directly measures the attributes of living systems to determine the condition of a specific landscape. The very presence of thriving living systems – sea otters and kelp forests off the central California coast; salmon, orcas, and herring in Pacific Northwest waters; monk seals in the Mediterranean Sea – says that the conditions those organisms need to survive are also present. A biota is thus the most direct and integrative indicator of local, regional, or global biological condition. Biological assessments give us a way to evaluate whether monetary valuations, sustainability indexes, and ecological footprints are telling the truth about human impact on the biosphere. Biological assessments permit a new level of integration because living systems, including human cultures, register the accumulated effects of all forms of degradation caused by human actions.

Direct, comprehensive biological monitoring and assessment has been done for many aquatic systems; measures are less developed for terrestrial systems. The index of biological integrity (IBI), for example, was developed in 1981 to assess the health of streams in the US Midwest and has since helped scientists, resource managers, and citizen volunteers understand, protect, and restore rivers in at least 67 countries worldwide (Karr 2006). Borrowing from the same page as more recent sustainability indexes, IBI takes the concept behind economic indexes like GDP or the consumer price index – that of multiple indicators integrated into a multimetric index – and applies it to animals and plants in bodies of water or other environments. The specific measurements (**Table 3**) are sensitive to a broad range of human effects in waterways, such as sedimentation, nutrient enrichment, toxic chemicals, physical habitat destruction, and altered flows. The resulting index combines the responses of biological parts such as species, as well as processes such as food web dynamics, to human actions.

Table 3 Biological attributes in two indexes of biological integrity

Benthic invertebrates	*Fish*
Total number of taxa	Number of native fish species
Number of mayfly taxa	Number of riffle-benthic insectivore species
Number of stonefly taxa	Number of water-column insectivore species
Number of caddisfly taxa	Number of pool-benthic insectivore species
Number of intolerant taxa	Number of intolerant species
Number of long-lived taxa	Relative abundance of omnivores
Number of clinger taxa	Relative abundance of insectivores
Relative abundance of tolerant taxa	Relative abundance of tolerant taxa
Relative abundance of predators	Relative abundance of top carnivores
Dominance	Relative abundance of diseased or deformed individuals

Indexes of biological integrity have been developed for a number of aquatic and terrestrial environments; the most widely used indexes for assessing rivers examine fishes and benthic (bottom-dwelling) invertebrates. These groups are abundant and easily sampled, and the species living in virtually any water body represent a diversity of anatomical, ecological, and behavioral adaptations. As humans alter watersheds and water bodies, changes occur in taxonomic richness (biodiversity), species composition (which species are present), individual health, and feeding and reproductive relationships.

Sampling the inhabitants of a stream can tell us much about that stream and its landscape. Biological diversity is higher upstream of wastewater treatment plants than downstream, for example, whereas, at the same location, year-to-year variation is low (**Figure 2**). Biological sampling can also reveal differences between urban and rural streams. For instance, samples of invertebrates from one of the best streams in rural King County, in the US state of Washington, contain 27 kinds, or taxa, of invertebrates; similar samples from an urban stream in the city of Seattle contain only 7. The rural stream has 18 taxa of mayflies, stoneflies, and caddisflies; the urban stream, only 2 or 3. When these and other metrics are combined in an index based on invertebrates, the resulting benthic IBI (B-IBI) ranks the condition, or health, of a stream numerically (**Table 4**).

A benthic IBI can also be used to compare sites in different regions. Areas in Wyoming's Grand Teton National Park where human visitors are rare have near-maximum B-IBIs. Streams with moderate recreation taking place in their watersheds have B-IBIs that are not significantly lower than those with no human presence, but places where recreation is heavy are clearly damaged. Urban streams in the nearby town of Jackson are even more degraded, yet not as bad as urban streams in Seattle.

Nation-specific biological assessments also can be and are being done. The US Environmental Protection Agency (2006), for example, performed a nationwide survey of stream condition using an IBI-like multimetric index. The survey found that 28% of US stream miles were in good condition in

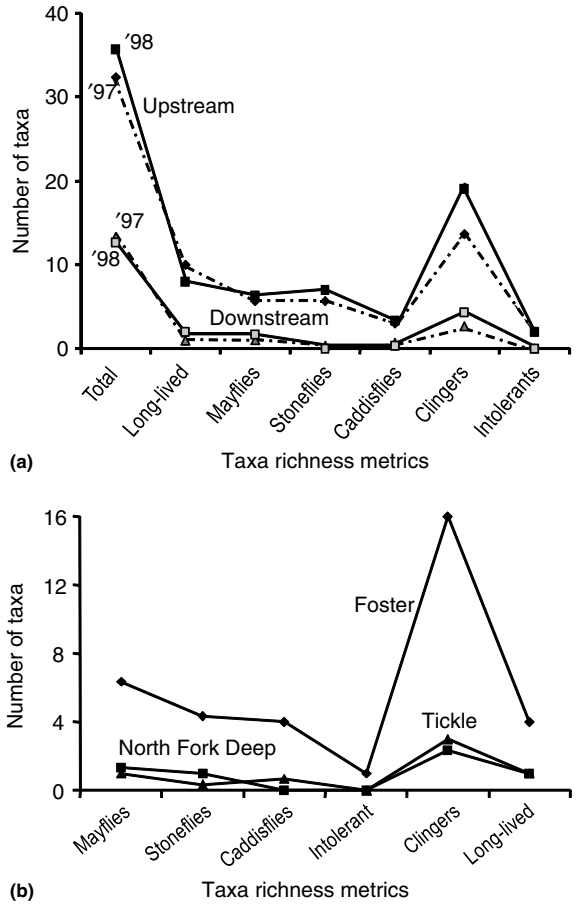


Figure 2 (a) Biodiversity is higher at sites upstream of wastewater treatment outfalls than downstream. At Tickle Creek near Portland, Oregon (United States), taxa richness differed little between years but differed dramatically between sites upstream of a wastewater outfall and sites downstream. (b) Taxa richness also differed between two creeks with wastewater outfalls (Tickle and North Fork Deep) and one creek without an outfall (Foster). All three streams flowed through watersheds with similar land uses.

comparison with least-disturbed reference sites in their regions, 25% were in fair condition, and 42% were in poor condition (5% were not assessed). The agency has been expanding this effort to include other water resource types, including coastal waters, coral reefs, lakes, large rivers, and wetlands.

Since 2000, the Heinz Center (2008) has published two editions of its report on the state of US ecosystems, which seek to capture a view of the large-scale patterns, conditions, and trends across the United States. The center defined and compiled a select set of indicators – specific variables tracking ecosystem extent and pattern, chemical and physical characteristics, biological components, and goods and services derived from the natural world – for six key ecosystems: coasts and oceans, farmlands, forests, fresh waters, grasslands and shrublands, and urban and suburban landscapes.

Among the many conclusions of the 2008 report were that the acreage burned every year by wildfires was on the increase; nonnative fish had invaded nearly every watershed in the lower 48 states; and chemical contaminants were found in virtually all streams and most groundwater wells, often at levels above those set to protect human health or wildlife. On the plus side, ecosystems were increasing their storage of carbon, soil quality was improving, and crop yields had grown significantly.

The massive international UN Millennium Ecosystem Assessment (2005) remains the gold standard for synthesizing ecological conditions at a variety of scales. From 2001 through 2005, the project examined the full range of global ecosystems – from those relatively undisturbed, such as natural forests, to landscapes with mixed patterns of human use to ecosystems intensively managed and modified by humans, such as agricultural land and urban areas – and communicated its findings in terms of the consequences of ecosystem change for human well-being.

The resulting set of reports drew attention to the many kinds of services people derive from ecosystems, specifically, supporting services, such as photosynthesis, soil formation, and waste absorption; regulating services, such as climate and flood control and maintenance of water quality; provisioning services, such as food, wood, and nature's pharmacopoeia;

Table 4 Biological responses to different land uses

Region	*Land use*	*B-IBI*[a]
King County, Washington, USA	Rural	46
	Urban Seattle	12
Grand Teton Region, Wyoming, USA	Little or no human activity	48
	Light to moderate recreation	44
	Heavy recreation	32
	Urban Jackson Hole	21
Clackamas County, Oregon, USA[b]	Upstream of wastewater treatment plant	
	Tickle Creek up (1997, 1998)	40, 42
	Foster Creek	34
	Downstream of wastewater treatment plant	
	Tickle Creek down (1997, 1998)	14, 16
	North Fork Deep Creek	10

[a]Benthic index of biological integrity: the highest possible score is 50, the lowest is 10.
[b]See **Fig. 2** for graphs of selected B-IBI metrics at these sites.

and cultural services from scientific to spiritual. In addition, the reports explicitly tied the status of diverse ecosystems and their capacity to provide these services to human needs as varied as food and health, personal safety and security, and social cohesion. Even while recognizing that the human species is buffered against ecological changes by culture and technology, the reports highlighted our fundamental dependence on the flow of ecosystem services and our direct responsibility for the many faces of biotic impoverishment.

Among other findings, the assessment found that 60% of the services provided by ecosystems are being degraded at levels that derail efforts to stem poverty, hunger, and disease among the poor everywhere. Government subsidies that provide incentives to overharvest natural resources are a leading cause of the decline in renewable natural resources. Such declines are not limited to coral reefs and tropical forests, which have been on the public's radar for some time; they are pervasive in grasslands, deserts, mountains, and other landscapes as well. Moreover, the degradation of ecosystem services could grow worse during the first half of the 21st century, effectively blocking achievement of the United Nations' eight millennium development goals.

The core message embodied in ecological, especially biological, assessments is that preventing harmful environmental impacts goes beyond narrow protection of clean water or clear skies, even beyond protecting single desired species. Certain species may be valuable for commerce or sport, but these species do not exist in isolation. We cannot predict which organisms are vital for the survival of commercial species or species we want for other reasons. Failing to protect all organisms – from microbes and fungi to plants, invertebrates, and vertebrates – ignores the key contributions of these groups to healthy biotic communities. No matter how important a particular species is to humans, it cannot persist outside the biological context that sustains it. Direct biological assessment objectively measures this context and is fundamental to advancing the kinds of syntheses like those of the Millennium Ecosystem Assessment.

Recognizing and Managing Environmental Impacts

Every animal is alert to dangers in its environment. A microscopic protozoan gliding through water responds to light, temperature, and chemicals in its path, turning away at the first sign of something noxious. A bird looking for food must decide when to pursue prey and when not, because pursuit might expose the bird to predators. The bird might risk pursuit when it is hungry but not when it has young to protect. Animals that assess risks properly and adjust their behavior are more likely to survive; in nature, flawed risk assessment often means death or the end of a genetic line.

Humans, too, are natural risk assessors. Each person chooses whether to smoke or drink, fly or take the train, drive a car or ride a motorcycle and at what speeds. Each decision is the result of a partially objective, partially subjective internal calculus that weighs benefits and risks against one another.

Risk is a combination of two factors: the numerical probability that an adverse event will occur and the consequences of the adverse event. People may not always have the right signals about these two factors, however, and may base their risk calculus on the wrong clues. City dwellers in the United States generally feel that it is safer to drive home on a Saturday night than to fly in an airplane, for example. Even though the numerical odds of an accident are much higher on the highway than in the air, people fear more the consequences of an airplane falling out of the sky.

Human society also strives to reduce its collective exposure to risks. Governments seldom hesitate to use military power to defend their sovereignty or, albeit more reluctantly, their regulatory power to reduce workplace risks and risks associated with consumer products like automobiles. But people and their governments have been much less successful in defining and reducing a broad range of ecological risks, largely because they have denied that the threats are real.

Policies and plans generated by economists, technologists, engineers, and even ecologists typically assume that the lost and damaged components of living systems are unimportant or can be repaired or replaced. Widespread ecological degradation has resulted directly from the failure of modern society to properly assess the ecological risks it faces. Like the fate of Old Kingdom Egypt or Easter Island, our civilization's future depends on our ability to recognize this deficiency and correct it.

Risk assessment as formally practiced by various government agencies began as a way to evaluate the effects of toxic substances on human health, usually the effects of single substances, such as pollutants or drugs, from single sources, such as a chemical plant. During the 1990s, the focus widened to encompass mixtures of substances and also ecological risks. Ecological risk assessment by the US Environmental Protection Agency (1998) asks five questions: Is there a problem? What is the nature of the problem? What are the exposure and ecological effects? (A hazard to which no one or nothing is exposed is not considered to pose any risk.) How can we summarize and explain the problem to affected parties, both at-risk populations and those whose activities would be curtailed? How can we manage the risks?

Even though these are good questions, ecological risk management has not made any visible headway in stemming biotic impoverishment. Its central failing comes from an inability to correctly answer the second question, What is the nature of the problem? Our present political, social, and economic systems simply do not give us the right clues about what is at risk. None of society's most familiar indicators – whether GDP or number of threatened and endangered species – measure the consequences, or risks, of losing living systems.

If biotic impoverishment is the problem, then it only makes sense to direct environmental policy toward protecting the integrity of biotic systems. Integrity implies a wholeness or unimpaired condition. In present biological usage, integrity refers to the condition at sites with little or no influence from human activity; the organisms there are the products of natural evolutionary and biogeographic processes in the absence of humans. Tying the concept of integrity to an evolutionary framework provides a benchmark against which to evaluate sites that humans have altered.

Directing policy toward protecting biological integrity – as called for in the United States' Clean Water Act, Canada's

National Park Act, and the European Union's Water Framework Directive, among others – does not, however, mean that humans must cease all activity that interferes with a "pristine" earthly biota. The demands of feeding, clothing, and housing billions of people mean that few places on Earth will maintain a biota with evolutionary and biogeographic integrity. Rather, because humans depend on living systems, it is in our interest to manage our activities so they do not compromise a place's capacity to support those activities in the future; that capacity can be called ecological health.

Ecological health describes the preferred state of sites heavily used for human purposes: cities, croplands, tree farms, water bodies stocked for fish, and the like. At these places, it is impractical to set a goal of integrity in an evolutionary sense, but we should avoid practices that damage these places to the point where we can no longer derive the intended benefits indefinitely. For example, agricultural practices that leave behind saline soils, depress regional water tables, and erode fertile topsoil faster than it can be renewed destroy the land's biological capacity for agriculture; such practices are unhealthy in both ecological and economic terms.

Biological integrity as a policy goal redirects our focus away from maximizing goods and services for the human economy toward ways to manage human affairs within the bounds set by the natural economy. It begins to turn our attention away from questions such as, How much stress can landscapes and ecosystems absorb? to ones such as, How can responsible human actions protect and restore ecosystems? In contrast to risk assessment, striving to protect biological integrity is more likely to lead away from mandated corrective actions in the form of technological fixes for environmental problems and toward practices that prevent ecological degradation and encourage ecological restoration.

Leopold (1949, pp. 224–225), in *A Sand County Almanac*, was the first to invoke the concept of integrity in an ecological sense: "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise." Managing for biological integrity requires the kind of ethical commitment inherent in Leopold's words. We are called to curb consumerism and limit population size, to embrace less-selfish attitudes toward land stewardship, and to understand that the biosphere matters. Instead of calling on human technical and spiritual wellsprings to manage resources, we have to call on them for managing human affairs.

We have to find and use appropriate measurements for all the factors contributing to biotic impoverishment, be they climate change, overharvesting, agriculture, or environmental injustice. Measurement of environmental impact founded on the evolutionary idea of integrity allow us to directly assess biotic condition and to compare that condition with what might be expected in a place with little or no human influence. At least then we can make an informed choice: continue with activities that degrade biotic condition or think of an alternative.

The ecological world is a complex, variable, and often unpredictable system. When managing for ecological risks, people and their governments need to expect the unexpected and develop formal, yet flexible means of coping with environmental surprises. Rather than plunge ahead with projects entailing ecological risks because they can be done, decision makers should follow the precautionary principle, which holds that regulators should act to prevent potential environmental harm even in the absence of certainty. It acknowledges the existence of uncertainty rather than deny it, and it includes mechanisms to safeguard against potentially harmful effects.

Although inappropriate ecological risk assessment and management are more often the norm today, modern institutions are capable of recognizing ecological threats correctly and responding to them in time, as they did with the Montreal Protocol. A decade after the agreement's adoption, satellite measurements in the stratosphere indicated that ozone-depleting pollutants were in fact on the decline. Given the treaty's success, some policy experts have suggested using it to help curb global warming. Specifically, negotiators at the 2010 annual meeting of signatory parties were considering a proposed expansion of the ozone treaty to phase out the production and use of industrial chemicals called hydrofluorocarbons, which have thousands of times the global warming potential of carbon dioxide. Even though the Montreal Protocol was not designed to fight climate change, policymakers in favor of the new proposal say that it could and should be used to achieve broader environmental objectives.

Reuniting the Fragments

Early in the 20th century, two sciences of "home maintenance" began to flourish: the young science of ecology (from the Greek *oikos*, meaning home) and a maturing neoclassical economics (also from *oikos*). Ecology arose to document and understand the interactions between organisms and their living and nonliving surroundings – in essence, how organisms make a living in the natural economy. In fact, Ernst Haeckel, who coined the term in the 1860s, defined ecology in an 1870 article as the body of knowledge concerning the economy of nature. Neoclassical economics, in contrast, reinforced humans' self-appointed dominion over nature's wealth. It brought unparalleled gains in societal welfare in some places, but it also divorced the human economy from the natural one on which it stands (see **Figure 1**).

By now it is clear to economists and ecologists alike that human actions and their effects have reached scales unprecedented in the history of life. We have altered Earth's physical and chemical environment, changed the planet's water and nutrient cycles, and perturbed its climate. We have unleashed the greatest mass extinction in 65 My and distorted the structure and function of nonhuman and human communities worldwide. In trying to make our own living, we have jeopardized the Earth's capacity to sustain other species and our own species as well. We are losing the *bio* in biosphere.

Now, faced with these unprecedented losses, we need to understand – not deny – the ecological consequences of what we do. We urgently need a new science and art of home maintenance, one that sees the human species' role as ecosystem engineer for what it has become and reconnects human and natural economies. This new science can help us understand and correctly interpret the consequences of

human-driven change. By using indicators that measure what matters for sustaining living systems, this new science can make nature visible again and shed new light on the value of the ancient heritage we share with the larger biosphere. It can help us reunite the fragments of our worldview and re-create ethical, social, and ecological bonds that were put aside two centuries ago in the name of progress. And such a science can enable us to reengineer our own social, political, and economic institutions instead of ecosystems. This we must do – now – before we impoverish the biosphere and ourselves for all time.

See also: Biodiversity and Human Health. Biodiversity as a Commodity. Biogeochemical Cycles. Climate Change and Ecology, Synergism of. Conservation and the World's Poorest of the Poor. Domestication of Crop Plants. Ecological Footprint, Concept of. Economic Value of Biodiversity, Measurements of. Ecosystem Services. Ecotoxicology. Endangered Ecosystems. Energy Use, Human. Environmental Ethics. Human Impact on Biodiversity, Overview. Hunter-Gatherer Societies, Ecological Impact of. Justice, Equity and Biodiversity. Market Economy and Biodiversity

References

Davis W (2009) *The Wayfinders: Why Ancient Wisdom Matters in the Modern World*. Toronto: House of Anansi Press.

Diamond J (1997) *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W.W. Norton.

Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature* 418: 700–707.

Diamond J (2005) *Collapse: How Societies Choose to Fail or Succeed*. New York: Viking.

Fagan B (1999) *Floods, Famines, and Emperors: El Niño and the Fate of Civilizations*. New York: Basic Books.

Heinz Center (2008) *The State of the Nation's Ecosystems: Measuring the Lands, Waters, and Living Resources of the United States*. Washington, DC: Island Press. Available online at: http://www.heinzctr.org.

Homer-Dixon TF (1999) *Environment, Scarcity, and Violence*. Princeton, NJ: Princeton University Press.

Intergovernmental Panel on Climate Change (IPCC) (2007) *Climate Change 2007: Synthesis Report: Summary for Policymakers*. Geneva: IPCC. Available online at: http://www.ipcc.ch/publications_and_data/publications_and_data.shtml.

Karr JR (2006) Seven foundations of biological monitoring and assessment. *Biologia Ambientale* 20(2): 7–18.

Leopold A (1949) *A Sand County Almanac: And Sketches Here and There*. New York: Oxford University Press.

Maathai W (2009) *The Challenge for Africa*. New York: Pantheon Books.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press. Available online at: http://www.maweb.org/en/index.aspx.

Myers N (1993) *Ultimate Security: The Environmental Basis of Political Stability*. New York: W.W. Norton.

Quammen D (1996) *Song of the Dodo: Island Biogeography in an Age of Extinction*. New York: Simon and Schuster.

TEEB (The Economics of Ecosystems and Biodiversity) (2010) *The economics of ecosystems and biodiversity: Mainstreaming the economics of nature: A synthesis of the approach, conclusions and recommendations of TEEB*. United Nations Environment Programme. Available online at: http://www.teebweb.org/.

U.S. Environmental Protection Agency (1998) *Guidelines for Ecological Risk Assessment*, EPA/630/R095/002F. Washington, DC: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency (2006) *Wadeable Streams Assessment: A Collaborative Survey of the Nation's Streams*, EPA 841-B-06-002. Washington, DC: U.S. Environmental Protection Agency. Available online at: http://water.epa.gov/type/rsl/monitoring/streamsurvey/index.cfm

Wackernagel M and Rees WE (1996) *Our Ecological Footprint: Reducing Human Impact on the Earth*. Gabriola Island, BC: New Society Press.

Wilson EO (1987) The little things that run the world. *Conservation Biology* 1: 344–346.

Wilson EO (1994) *Naturalist*. Washington, DC: Island Press.

Woodwell GM (1990) *The Earth in Transition: Patterns and Processes of Biotic Impoverishment*. Cambridge, UK: Cambridge University Press.

Wright R (2004) *A Short History of Progress*. Toronto: House of Anansi Press.

Estuarine Ecosystems

G Carleton Ray and Jerry McCormick-Ray, University of Virginia, Charlottesville, VA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is revised from the previous edition by G. Carleton Ray, volume 2, pp 579–591, © 2001, Elsevier Inc.

Glossary

Biological diversity (biodiversity) The collection of genomes, species, and ecosystems occurring in a geographically defined region (NRC, 1995).

Coastal zone Zone whose terrestrial boundary is defined by (a) the inland extent of astronomical tidal influence or (b) the inland limit of penetration of marine aerosols within the atmospheric boundary layer and including both salts and suspended liquids, whichever is greater; the seaward limit is defined by (a) the outer extent of the continental shelf (approximately 200 m depth) or (b) the limits of territorial waters, whichever is greater (Hayden *et al.*, 1984).

Estuary Semienclosed coastal body of water that has a free connection with the open sea and within which seawater is measurably diluted with freshwater derived from land drainage (Pritchard, 1967).

Functional diversity Variety of different responses to environmental change, especially the diverse time and space scales with which organisms react to each other and to the environment (Steele, 1991).

Metapopulation An abstraction of the population to a higher level at which individuals frequently move from one place (population) to another, typically across habitat types that are not suitable for their feeding and breeding activities, and often with substantial risk of failing to locate another suitable habitat patch in which to settle (Hanski and Gilpin, 1991).

Introduction

Some of the steepest environmental gradients on planet Earth occur in the coastal zone, where land, sea, and atmosphere uniquely interact and exchange energy and materials. The dynamic linkages among biological, physical, and chemical systems are exceptionally strong in estuaries, characterized by cyclic changes that occur at different frequencies – such as for tides, salinity cycles, freshwater inputs, light, and temperature stratification. Estuaries also bear the brunt of extreme events, such as flooding, storms, hurricanes, and impacts of seasonal sea ice. All of these features are important for organisms, which have evolved suites of adaptive mechanisms to cope.

Estuaries have usually been considered as transitional areas between fresh and saltwater environments. However, relatively few species are totally confined to estuarine conditions, even though various stages of many species' life cycles are estuary-dependent. This raises questions about whether estuaries can be considered as transitional or as more-or-less autonomous ecosystems. The distribution of biodiversity provides important information toward the resolution of this apparent dichotomy, which needless to say is essential with respect to conservation and management.

Our present knowledge about estuary-dependent biodiversity is sparse. Fundamental questions remain about species distributions in estuaries, in what ways species are adapted to estuaries, and how some species may affect others by means of structural or functional interrelationships. Furthermore, the diversity of estuaries relative to other ecosystems remains to be clarified. These questions require both ultimate, historical–evolutionary explanations and proximate, functional–ecological considerations.

Estuaries are located in the critical portion of Earth called the "coastal zone," which accounts for a disproportionate amount of global ecological functions. For example, the coastal zone (modified from Pernetta and Milliman, 1995; see also Crossland *et al.*, 2005):

- occupies only 18% of the surface of the globe, 8% of the ocean surface, and 0.5% of ocean volume;
- but provides for up to 50% of global denitrification, 80% of global organic matter burial, 90% of global sedimentary mineralization, 75–90% of the global sink of suspended river load and its associated elements/pollutants, and more than 50% of present-day global carbonate deposition;
- also supplies approximately a quarter of global primary production, around 14% of global ocean production and 90% of the world fish catch.

Despite this coastal zone association, the *Global Biodiversity Assessment* (Heywood and Watson, 1995) contained no sections specifically devoted to estuaries, but did characterize biodiversity as comprising three disciplines, which also apply to estuaries: (1) taxonomy: provides the reference system and depicts the pattern or tree of diversity for all organisms; (2) genetics: gives a direct knowledge of the genetic variations found within and between species; and (3) ecology: provides knowledge of the varied ecological systems in which taxonomic and genetic diversity are located, and also provides information on functional components. Evolutionary biology brings these fields together, as it "provides explanations of how biodiversity arose, and the processes, such as speciation and extinction, by which it continues to change." Nevertheless, a comprehensive understanding of the functional biodiversity

of estuaries remains a future goal, and is the focus of this article. In this respect, it is worthwhile to note that, even today, estuarine science continues to be organized along disciplinary lines. Although the study of land–seascape ecology of estuaries remains in its infancy, there are extensive publications on geomorphology, land–sea interactions, coastal zone management, and other disciplines from which to gain an integrated understanding of estuaries, but which are beyond the scope of this review.

Ironically, estuaries became topics of intensive concern and research only in the mid-twentieth century, whereas humans have lived in close proximity to estuaries and have been dependent on them and their biological resources for millennia. Reasons for human proximity to and dependence on estuarine environments are both social and ecological, for estuaries are ecologically diverse and productive, making possible the sustainment of large and sophisticated human societies. Indeed, it is more than coincidental that among the first known city-states were those of the lower reaches of the Tigris and Euphrates Rivers of Mesopotamia.

The distribution of estuaries corresponds to regional and coastal characteristics; that is, they tend to be extensive, large, and numerous where coastal plains are wide and flat, but are relatively small where coastal plains are steep and narrow. Particularly in the former, estuaries and associated lagoons constitute a much higher percentage of the coasts than is generally recognized. In fact, many of the world's largest cities (London, New York, Karachi, Amsterdam, Alexandria, Tokyo, etc.) have been built on or near drained marshes or filled land adjacent to estuaries. In the USA 80–90% of the Atlantic and Gulf Coasts and 10–20% of the Pacific Coast consist of estuaries and lagoons (Emery, 1967).

As noted above, estuaries are best understood in the context of the coastal zone, definitions of which vary. Ketchum (1972) was among the first to take a functional perspective, that the coastal zone "is the broad interface between land and water where production, consumption, and exchange processes occur at high rates of intensity." NERC (1992), however, defined the coastal zone as: "An indefinite zone of land and sea that straddles the shoreline; includes all land that is the product of, and/or at risk from (Holocene) marine processes, and extends seaward from the shoreline to water depths of about 30 m." The key element is "marine processes" and, from that point of view, it seems best to adopt Ketchum's broader view. Accordingly, Hayden *et al.* (1984) adopted Ketchum's definition (*see* Glossary), which makes sense of such interactions as the existence of coastal vegetation under the influence of aerosols, sedimentation induced by freshwater flows and atmosphere–ocean processes, and the coastal distribution of aquatic biota worldwide. With respect to the latter, Nelson (1984) estimated that of about 21,700 described species of fishes, about 8400 (39%) occur in freshwater and 2700 (12%) are oceanic. Nearly half of these fishes (10,200 species, or 47%) are coastal, that is, occur from estuaries to the outer extent of the continental shelf. This proliferation of fish diversity is powerful evidence of the functional importance and the extent of the coastal zone. (Later editions of Nelson (1984), in 1994 and 2006, predict that the total species of fish might reach more than 28,000 and 32,000 species, respectively – greater than the combined total of all tetrapods – but the proportions originally given remain approximately the same.)

Within this coastal zone context, Pritchard's (1967) definition of "estuary" is most generally accepted (*see* Glossary). Additionally, Pritchard's definition takes account of Pleistocene rises and falls in sea level, as well as of terrestrial processes, such as sedimentation, which clearly affect distributions of biota. However, other definitions include that of Mann (1982), who defined an estuary as "a region where river water mixes with, and measurably dilutes, sea-water." However, this definition could include semienclosed seas (e.g., the Baltic), plumes of large rivers, and diluted water off open coasts, making difficult any geographic analysis of estuarine biodiversity or function.

In this coastal zone context, estuaries are subject to rapid environmental, structural-functional change, which has major consequences for biodiversity. Hydrological, biological, and sedimentary processes and events may substantially alter or destroy estuaries. Thus, the estuaries that we now see are the result of the latest major episodes of sea level fall and rise; in fact, the age of the present estuaries is only about 1% of the age of the continental shelf (Emery, 1967). It is reasonable to assume that the communities of estuarine biota that exist today are as young and equally subject to change, most notably as a result of human activities.

Many estuaries around the world have been studied in some detail. The North American bias in this article reflects the considerable body of research that has been conducted on North American estuaries during the past few decades, motivated unfortunately by the depleted, overenriched, polluted, and overpopulated states of many of them, some aspects of which will be examined *in* A Case Study: The Chesapeake Bay, below.

Estuarine Classification

Classification is essential as a reference system among estuaries. Various physical classifications, or typologies, of estuaries facilitate comparisons, but no typology has yet been directed specifically to estuarine biodiversity. To our knowledge, the first classification was the so-called "Venice system" ("V" – Anonymous, 1959), in which estuaries were divided into salinity zones. This system was later modified by on the basis of species' salinity tolerances ("B" – Bulger *et al.*, 1993). These two schemes align rather closely and may be compared as follows (ppt = parts per thousand): Limnetic (freshwater), 0.5 ppt (V), 4 ppt (B); oligohaline 0.5–5 ppt (V), 2–14 ppt (B); mesohaline 5–18 ppt (V), 11–18 ppt (B); polyhaline 18–30 ppt (V), 16–27 ppt (B); euhaline 30 ppt–full marine (V); 24–ppt marine (B).

The reason for the differences in salinity ranges between these two systems is that the former was derived from salinity, whereas the latter was derived analytically from species' salinity tolerances. In both cases, however, the compartments are oversimplistic, as estuaries exhibit many characteristics that influence biotic distribution and the distinction of estuarine zones, variably identified as "upper reaches," "upper-middle reaches," "lower reaches," and so forth. Nor do salinity-derived systems distinguish zones according to variations in bottom

type, water movement, volume of flow, and other attributes important to the biota.

Another classification concerns basin geomorphology, which is of obvious importance with respect to circulation patterns. Classification on this basis appears in many texts and may be summarized as:

- Coastal plain estuary (drowned river valley): Usually confined to areas with a wide coastal plain where seawater has invaded existing rivers because of sea level rise since the Pleistocene. Generally the up-estuary limit is where chlorinity is about 0.06% (salinity about 0.1%); above this point there may be a portion of tidal freshwater.
- Fjord: Generally U-shaped in cross section due to glacial scouring, in which the sides are steep. May be fed by a river, have a deep basin, and a shallow sill may be present near the mouth.
- Bar-built: Occurs in flat, low-lying areas, where sand tends to be deposited in bars lying parallel to the coast. Usually shallow and wind-mixed. Can be a composite of drowned river valleys and embayments, and occurs where offshore sand barriers are built between headlands into a chain enclosing a body of water. May be fed by multiple rivers, but the total drainage area is usually not large.
- Tectonic: A miscellaneous category including estuaries formed from faults or folding of Earth's crust. Often have an excess of freshwater flow.

The interchange of freshwater and seawater provides yet another classification. The inlet (mouth) must be of sufficient dimension to allow mixing of seawater and freshwater, and the dilution of seawater provides the density gradients that drive characteristic circulation patterns. In terms of this interchange, the general classification is:

- Salt wedge: wherein a layer of relatively freshwater flows out at the surface.
- Partially mixed (moderately stratified): wherein tidal flow, turbulence, and mixing are increased, tending to erase the salt wedge.
- Vertically homogeneous: wherein tidal flow is strong, river runoff is weak, and all stratification is broken down.

Combinations of these typologies are possible; that is, it may be possible to find a stratified or a mixed bar-built estuary, or a fjord with a salt wedge or not. Furthermore, the extents of salinity zones can vary considerably for all categories. Such combinations of structure and hydrologic process result in highly varied conditions in the distributions of sediment, phytoplankton, submerged aquatic vegetation, and fishes and invertebrates. Additionally, variations in freshwater inputs, circulation, turbulence, and mixing can modify the typology.

A final classification concerns estuarine evolution. Roy (1984) described estuaries of three successional types for estuaries of New South Wales, Australia: drowned river valleys, barrier estuaries, and saline coastal lakes. Infilling during relatively short time spans affected their size, configuration, the invasion of mangroves and other aquatic vegetation, and fish communities. Faunal population densities and species diversity increase with ecological complexity, but as infilling becomes more advanced the estuary becomes simplified and biological diversity declines. Therefore, estuarine geology, hydrology, and biology form a hierarchical succession and biodiversity maxima are reached in intermediate stages.

Estuarine Biodiversity

From the foregoing discussion, the impression may be gained that estuaries are simply transitional and, therefore, not biologically diverse. Indeed, Sanders (1968) found that estuaries are relatively nondiverse biologically, while noting: "What is significant is that each environment seems to have its own characteristic rate of species increment." This is to say that although salinity, for example, is an important determinant of the distribution of biota, estuaries exhibit high and dynamic habitat and land–seascape diversity, a consequence of which is high variability among the biota and a high degree of biotic interaction. Thus, estuarine biotic communities would be expected to be especially complex, contrary to earlier impressions of estuarine biological and ecological simplicity. Additionally, their biota have evolved resiliency to natural disturbance. This is expressed at species, community, and ecosystem levels, leading to the simplistic impression that estuarine species are facultative with respect to estuaries as preferred environments. These characteristics have resulted in a tendency to describe any species that enters estuaries, or those that tolerate brackish waters, as "estuarine," a characterization that can be misleading. However, many estuarine fishes and invertebrates move in or out of estuaries and appear to be transients. Conversely, some species seem to be restricted to estuarine and near-shore environments, at least at some life-history stage. A notable example concerns temperate oysters (*see* A Case Study: Chesapeake Bay, below) that build extensive reefs in estuaries and lagoons and nowhere else, providing habitat for dozens of species, representative of almost every animal phylum. Therefore, it has been challenging to define an "estuary-dependent" organism.

Carriker (1967) noted that estuarine biota have adapted in different ways to estuarine conditions; for example, oligohaline organisms disappear below the head of the estuary; euryhaline species constitute the majority of the estuarine biota, as they can tolerate salinities as low as 5 ppt, as well as full salt water; and stenohaline species do not tolerate salinities of <25 ppt and are found only at the mouths of estuaries or on open seashores. This leaves "true estuarine organisms" – those relatively few species that are restricted to estuaries and that are best represented in the upper and middle reaches. Carriker concentrated mainly on benthic invertebrates, but concluded that an "estuarine biocenose" may be justified as a discrete functional aggregation of interdependent, regularly recurring, dominant, benthic populations that are strongly represented numerically. Carriker acknowledged that much needs to be learned of species' ecology and life histories, and that the estuarine biotope appears to be more than "just a simple overlapping of factors (an ecotone) extending from the sea and the land, but is characterized by a unique set of its own factors arising from within the estuary from the materials and forces contributed by its bounding environments."

Some of the dominant, or "true," macroscopic biota of estuaries that Carriker (1967) named are plants – *Spartina alterniflora*, *Zostera marina*, *Ruppia maritima*, *Cymodocea mamatorium*, *Rhizophora mangle*, and *Avicennia nitida*, and invertebrates – *Nereis diversicolor*, *Balanus improvisus*, xanthid mud crabs, *Uca pugnax*, *Callinectes sapidus*, *Mya arenaria*, *Mytilus edulis*, *Modiolus demissus*, and *Crassostrea virginica*. Additionally, Carriker noted that characteristic estuarine habitats include tidal marshes, mangrove swamps, seagrasses, oyster reefs, soft clam–clam worm flats, and others. Finally, Carriker stated that: "Little is known of the sum of these effects on community structure, but they do emphasize the need to consider benthic organisms in the context of the total ecosystem rather than as an independent benthic biocenose." This statement, made almost a half century ago, has yet to be fully realized.

Fishes are the best known of aquatic groups in a general sense, mostly due to their commercial value. Therefore, insights into "estuarine dependency" may be best revealed through their study. One reason for this is their mobility in which various life-history stages inhabit quite different environments. Winemiller (1995) reviewed fish ecology and made the point, first, that fishes are by far the most diverse vertebrates, inhabiting an incredibly wide range of aquatic habitats from pole to pole. Second, fishes are ecologically diverse, with a wide variety of food habits, behaviors, reproductive habits, physiologies, and morphologies. Third, fishes exhibit a range of life-history strategies that result from tradeoffs among various attributes, including clutch and egg size; these strategies can be classified as opportunistic, periodic, and equilibrium, but a range of intermediate strategies also exist. Finally, fishes and their diversity in ecosystems can be used as "indicators" of environmental conditions.

Much attention has now been directed toward the early life histories of fishes, as this is closely related to recruitment and, therefore, of much interest to fish ecologists and to fisheries. Selection factors are of special importance Houde (1997), and in this regard Able and Fahay (1998) extended studies on juvenile stages of fishes to define "estuarine dependence," observing that numbers of permanent estuarine residents are relatively low, presumably because estuaries exhibit extremes in environmental conditions. Estuarine fish diversity is augmented by transients, such as freshwater species that occasionally occur in estuaries and marine species that spawn at sea but whose young use estuaries as nurseries. Therefore, estuarine fish fauna includes both residents and transients and a wide range of sizes, ages, and adaptations. In addition, those species that have successfully invaded estuaries usually inhabit only a small number of broad niches, implying that larger estuaries have larger numbers of species owing to increased habitat and niche complexity.

Able and Fahay found that, of the species for which good information is available, 60% are transients, 28% are residents, 6% are infrequent, and 6% are unclassified. For juveniles, they suggested the following adaptive groups:

- Group I. Facultative estuarine breeders, species whose nurseries are either in estuaries or on the inner shelf (e.g., *Centroptristis striata* and *Brevoortia tyrannus*).
- Group II. Seasonal residents, species whose adults migrate into estuaries to spawn in spring or summer (e.g., *Menidia menidia* and *Mustelus canis*).
- Group III. Anadromous species, species whose adults migrate through estuaries in order to spawn in freshwaters (e.g., *Morone saxatilus* and *Alosa* spp.).
- Groups IV–VI. Early users, delayed users, and distant spawners, species that spawn exclusively in the ocean, but the location, timing, and manner of use of estuaries by young-of-the-year juveniles vary (e.g., *Pollachius virens*, *Prionotus carolinus*, and *Mugil cephalus*).
- Group VII. Expatriates, species whose estuarine larvae come from distant spawning (e.g., *Chaetodon ocellatus* and *Monacanthus hispidus*).
- Group VIII. Summer spawners, the largest group, represented by shallow-water spawners whose larvae develop in the immediate vicinity of spawning sites (e.g., *Cyprinodon variegatus* and *Fundulus heteroclitus*).
- Group IX. Winter–spring spawners, a few species that spawn in the winter or spring (e.g., *Pseudopleuronectes americanus*).
- Group X. Migrating spawners, species that undergo spawning migrations within the estuary (e.g., *Morone americana*).
- Group XI Species difficult to classify: species for which some populations appear to be estuarine and other populations do not (e.g., *Tautogolabrus adspersus*).

But, of these, which are truly "estuary dependent"? Able and Fahay (1998) caution that, for fishes at least, "estuarine dependence" depends on the resolution of three areas of research: (1) the need to sample well-defined areas thoroughly for habitat evaluation; (2) assessment of the effects of habitat loss; and (3) more detail on temporal and spatial use of habitats where early stages are collected. In short, a coherent understanding of the life-history factors that control the early life histories of fishes remains to be accomplished. The same no doubt holds for invertebrates. For macroscopic plants, the situation is perhaps less uncertain, as their life histories are simpler and assessments are more easily accomplished.

In sum, most truly estuarine species should typically be resistant to the extreme environmental conditions of estuaries, and/or take advantage of favorable situations when they occur; consequently, such species may not appear to have strong habitat associations. This makes difficult the strict establishment of a definition of "estuarine dependency." Also, the seaward boundary of an "estuary" is often blurred, so that the definition of "dependency" is hampered by lack of comparative, quantitative data from offshore habitats. The easiest distinctions are for those species for which at least one stage is shown to be physiologically or behaviorally obligate, but better natural history and experimental data for most species are required for this determination. Therefore, the question "What is an estuarine species?" remains elusive for many, if not most, species. In addition, the oft-made contention that estuaries with similar habitats may support similar species assemblages seems reasonable, but may be misleading if assumptions of estuarine dependency are based on occurrence rather than in an adaptive-evolutionary sense.

Ecological Function of Biodiversity

In addition to genome, species, and ecosystem aspects of biodiversity, a fourth category must be considered, namely, "functional diversity" (Steele, 1991; see Glossary), which concerns ecological functions with respect to environmental maintenance and change. Ecological functions within the coastal zone and its estuaries are complex and variable, and they must be understood before interpretation of the composition and patterns of biodiversity is possible. Holligan and Reiners (1992) listed a number of factors that underlie the biological diversity of the coastal zone and its estuaries, first for natural processes:

Exchanges of Materials

Riverine and atmospheric export and import, groundwater exchange, and ocean–land material transport operate at various levels, but are presently poorly understood. The role of anadromous fishes in transport of organic matter has hitherto been neglected but may be essential for some systems (Hesslein *et al.*, 1991; Bilby *et al.*, 1996; Garman and Macko, 1998).

Physico–Chemical Properties

The coastal zone is a region of high energy exchange due to interactive oceanic and atmospheric forcing associated with topographical discontinuities, density gradients caused by freshwater inflows, and seasonal heat exchanges. Deltas, estuaries, and lagoons are the major sites for transformation and accumulation of organic matter and sediment, and all are highly variable spatially and temporally, so that their average conditions are not good indicators of net fluxes. Estuaries, in particular, are "sites of complex interactions, related to salinity gradients, phase transformation involving particle–water reactions, and to biological processes that cause biogeochemical transformations."

Biological Properties

Favorable conditions of light and nutrients in the coastal zone maintain high rates of primary productivity that are several times greater than for the open ocean, and even greater than for certain coastal upwelling areas; some coastal systems, such as salt marshes, mangrove swamps, mudflats, beds of aquatic vegetation, and coral reefs, exhibit even higher productivity. Some areas act as sources, others as sinks, and the nature of the coupling of primary productivity to the benthos or to open waters may determine community structure and function.

Biogeochemical Processes

Organic matter is readily reoxidized in coastal waters, but some poorly drained areas may become anaerobic. This is especially apparent in the bottom water of estuaries in summer, when temperatures are high.

Holligan and Reiners (1992) also identify present-day human activities that influence ecological functions and biological diversity:

Altered Delivery of Freshwater

Freshwater impoundment by damming has decreased total discharge into estuaries and coastal seas by about 15% from the 1950s to the 1990s, an amount equivalent to a change in sea level of -0.7 mm year^{-1}. Seasonal flows have also been altered; alteration in the residence time of water in estuaries may have far-reaching effects on chemical processes.

Changes in the Transport and Fate of Suspended Matter

Coastal subsidence, sediment starvation and consolidation, and nutrient levels have all been altered by human interventions. Land clearing, especially on steep slopes, has increased sedimentation.

Chemical Modification

Nutrients, eutrophy, and blooms have become widespread and their frequency seems to be increasing. Contaminants that are of most concern include heavy metals, synthetic organic compounds, radionuclides, and hydrocarbons.

Ecosystem Modification

This takes many forms, from physical change and habitat loss to depletion of resources. The worst-affected areas are those with high human population densities, such as Southeast Asia, and along temperate coasts that have significant sources of pollutants, such as the Baltic Sea.

Finally, longer-term processes that influence biodiversity are the effects of climate change, especially in response to global warming, should that continue to occur:

Natural Variations in Climate

Many climate-change studies describe possible alterations in biotic distributions. However, rather subtle climatic changes can induce large ecological changes that reflect the sensitive nature of marine food chains to climate and to climate-dependent factors such as nutrient levels and salinity. The direct effects of climate are often difficult to distinguish from those incurred by humans.

Temperature

The largest climate changes are expected in the higher latitudes. Thus, the poleward extension of some climate-sensitive species is to be expected in case of global warming. Temperature changes can also affect behavior and physiology (e.g., reproduction, feeding and food availability, predation, and migration), so that predictions are destined to be speculative.

Wind

Wind strongly influences upwelling and stratification, thus affecting productivity through nutrient and light availability. According to most climate change scenarios, wind intensity is expected to increase.

Extreme Events

Short time-scale events are also expected to increase with climate warming, and these may induce dramatic, long-term changes. A single storm lasting <5 days can result in sand transport equivalent to two-thirds of the total for an average year. Tsunamis have had the greatest effects recorded to date.

Changes in Sea Level

Presently, sea level is rising faster than the rate during the late Holocene due to a combination of thermal expansion of seawater and melting of polar ice as climate warms. Impacts of sea level rise on deltas and estuaries are already apparent, partly because these areas are low-lying, strongly perturbed by humans, and exhibit enhanced erosion and subsidence. Natural communities of plants and animals play a crucial role in determining the response of the coastal zone to changes in sea level.

This array of effects requires the development of research programs to address hypotheses that are relevant to the ecological function of estuarine biodiversity. Among many possibilities, the following hypotheses seem fundamental (slightly modified from Solbrig, 1991):

- For species: no aspect of life history has any influence on extinction probability.
- For communities: keystone species are essential for maintaining species richness in communities under all environmental conditions.
- For ecosystems: removal or addition of functional or structural groups that produce changes in temporal or spatial configuration of landscape elements will have no significant effect on ecosystem properties over a range of time and space scales.

These hypotheses can be clarified by means of a case-by-case examination (see the Section A Case Study: The Chesapeake Bay). For example, some species seem very alike in their life histories. However, redundancy in species function may mean that diversity and function are somewhat independent of one another. Many species of benthic infauna and epifauna are extremely abundant and ecologically important in estuaries. Many feed on sediments, and those with complete alimentary canals can consolidate organic residues into often long-lived, sculptured pellets. The question is: Are species that have similar ecological requirements replaceable?

With respect to physical structure, Roy (1984) stated that the ecology of an estuary depends on the geological stage it has reached in its evolutionary progression, and that the rate and direction of natural change provide a yardstick to assess impacts induced by humans. However, as Roy emphasized, factors influencing estuary development include (1) inherited factors, mainly of a geological nature, that control the size and shape of the basin and the nature of the sediment supply, and (2) contemporary factors of a process nature (such as tides, river discharge, waves, etc.) that influence modes of sedimentation, hydrodynamics, and the biota. This prompts the question: to what extent are structure and biodiversity related?

Mann (1982) observed that, in general, estuaries are more productive than adjacent shelf systems, bringing up relative rates of nutrient flushing; that is, many estuaries tend to act as nutrient traps and others may be enriched by pollution. The New York Bight is a spectacular example of enrichment of a large shelf area well beyond the Hudson River's mouth. Within 600 km^2 of sea at the apex of the Bight, phytoplankton production amounted to about 370 g C m^{-2} day^{-1}, compared with only 100 g C m^{-2} $year^{-1}$ at the edge of the shelf. Mann and Lazier (1991) also noted that the dynamics of coastal waters, including estuaries, are made complex by: (1) shallowness, resulting in relatively mixed water that may extend to the bottom, and dead biological material that may accumulate to release nutrients that are carried rapidly to surface waters; (2) tidal currents that create turbulent mixing, which has especially marked effects on food particles, fertilization of planktonic eggs, and larval dispersal; and (3) barriers to convection imposed by coastlines, meaning that wind drives surface water away from the coast, and upwelling is the only way for it to be replaced, bringing nutrients to the surface. The question here is: to what extent are enrichment and/or pollution and circulation related to biodiversity?

Turning to larval transport, a variety of organisms have adapted to the seaward flow of low-salinity water and a compensatory landward flow of bottom water in estuaries. Organisms can make vertical migrations to maintain themselves in an estuary or to enter or to leave it seasonally. For example, estuarine larval transport and retention mechanisms are evident on two scales: large-scale, regional circulation patterns and small-scale, local water motion. There is evidence that oyster larvae (*Cr. virginicus*) rise into the water column to be carried upstream, and that this is cued by increasing salinity associated with increasing upstream flow; larvae of the blue crab (*Ca. sapidus*), however, occur in maximum numbers in surface waters at the mouth of Chesapeake Bay at night as the salinity falls on the ebb tide (Boicourt, 1982). From this and other evidence, it has been concluded that the crab larvae develop offshore, and then reinvade as larvae or juveniles. Fishes have also been shown to vary their depths, some rising into surface waters during flood to remain in the estuary, and others doing the opposite to be taken out to sea. Thus, many invertebrates and fishes utilize the two-layered estuarine structure for retention or dispersal, and this may not be entirely passive, as has often been assumed. Despite improved knowledge, Boicourt's conclusion remains pertinent, that larval transport and retention "stands at the state of the art in both physical and biological fields," which become particularly relevant to recruitment into fisheries. The question is: does recruitment depend on return or retention as the operative process, and to what extent do larvae determine their own fates?

As another example of the importance of functional diversity, juveniles of the five species of Pacific salmons (*Oncorhynchus* spp.) vary in time spent in estuaries, but for all of them a high proportion of prey tends to be detritus feeders (Healey, 1982). This means that the configuration of the estuary and entrapment efficiency of detrital matter are important elements of juvenile salmon habitat. Retention of detritus is enhanced by restricted exchange with the ocean and low bed-load transport. Marshes and submerged aquatic vegetation are efficient detritus traps, and these habitats also shelter salmon from predation. Thus, it may be hypothesized that the complex of intertidal marshes, tidal creeks and secondary river

channels, lower intertidal and subtidal weed beds, and basin morphology all contribute to the carrying capacity of the estuary for young salmon, and that appropriate estuarine configurations must be conserved if salmon production is to be maintained. The question here concerns how the complexity of the land–seascape enhances biodiversity, and how this may operate differently for closely related species.

From these examples, it is apparent that, insofar as ecosystem functioning is concerned, the addition or deletion of species, structural groups, or essential processes can have profound effects on the capacity of an estuary to maintain its biodiversity. This is especially true for "keystone" species, which have influences out of proportion to their density or biomass. Likewise, the fragmentation and/or simplification of habitats and of land–seascapes may have profound effects on estuaries, since these impacts strongly affect ecological complexity and community structure and function. Furthermore, it is likely that the functional autonomy of estuaries depends on their size and the time intervals of various processes. That is, the degree to which an ecological system may be autonomous depends on the extent to which it is independent of the ecological dynamics outside its domain. Of course, no ecosystem can be completely independent owing to the climatic, ecological, and geological connections among all portions of Earth. However, the larger the domain, the more it may tend to be autonomous during restricted time spans of investigation. Obviously, the elucidation of autonomy for a domain of a given size is a complex matter. However, the simple fact is that under many management regimes, autonomy may be incorrectly assumed.

A Case Study: The Chesapeake Bay

Chesapeake Bay is one of Earth's largest estuaries. Its origin is that of a drowned river valley. This is the case for many estuaries associated with coastal plains, wherein the dominant processes are sedimentary and erosional and whereby the bottom is largely soft sand and mud.

The largest estuary in the U.S. and among the most complex and most productive in the world is Chesapeake Bay. This funnel-shaped estuary has the highest land-drainage to surface-area ratio of almost any other coastal water body; its 166,000 km^2 watershed dwarfs the 6500 km^2 water surface. Eight major river systems feed freshwater and terrigenous products into subestuarine systems that empty into the Bay, each exhibiting unique features. Extensive marshlands, seagrasses, and shellfish beds support a diversity of fish and shellfish that historically made the Bay one of the world's most commercially productive aquatic areas. Abrupt climatic events, seasonal change, changing mixes of salt and fresh water, and a varied biota have created what has appropriately been termed a "complex adaptive system" (Levin, 1998).

The Algonquian native word "Chesepiooc" became interpreted to mean Great Shellfish Bay, singling out its major historic feature – the oyster fringing reef (**Figure 1**), formed by a single species, the eastern oyster (*Cr. virginicus*). From a structural point of view, oyster reefs represent unique and dominant biogenic structures. These reefs' distribution and structural features in the 1870s (McCormick-Ray, 1998) indicate their ecological importance and functional diversity as nodes of intense biological activity, connectance through hydrological corridors, and as habitat attractors for estuarine species (McCormick-Ray, 2005).

Numerous scientists for more than a century of observation and study have cautioned against the depletion of oysters, the loss of which could impact biological, hydrological, erosional, and sedimentary patterns and processes, with major influences on functional biological diversity. A review by Rothschild *et al.* (1994) stated that "considerable concern is voiced regarding Chesapeake Bay water quality and the effects of disease on oysters" and that "the effects of a diminished oyster population abundance certainly must have changed the 'ecology' of Chesapeake Bay, and these effects must have become evident at the time of maximum stock decline (1884–1910)." Modern studies provide evidence of functional impact (Dame *et al.*, 1989; Ulanowicz and Tuttle, 1992; Lenihan *et al.*, 2001; Lehnert and Allen, 2002; Grabowski, 2004; Kirby and Miller, 2005).

To better understand the ecosystem effects of oysters and oyster reefs, one must begin at the regional scale, wherein the coastal zone is conceived as a nested hierarchical system (**Figure 2**). The regional scale is that of biogeographic and physiographic provinces. The mesoscale is represented by major regional subdivisions, such as watersheds, estuaries, coastal islands, lagoons, and coastal–ocean fronts that separate major marine regimes. The smallest scale is that of the interacting mosaics of land–seascapes, for example, wetlands, hard and soft bottoms, and water masses distinguished by salinity, temperature, and density. This hierarchy indicates top-down physical "controls" and bottom-up "feedbacks," and places estuaries in a central role. The biogeographic province (and/or "region") is an area whose limits are defined by the relative homogeneity of the biota. For example, the traditionally accepted boundaries for the biogeographic Virginian Province are Cape Cod, MA, and Cape Hatteras, NC. These capes are significant points of deflection for major ocean currents, principally the warm, north-flowing Gulf Stream and the cold, south-flowing Labrador Current. At these capes, dramatic changes occur in coastal characteristics, such as water temperatures and circulation patterns, and these physical features play major roles in determining the biogeographic ranges of biota. One major feature of the Virginian Province is the presence of very large estuaries, such as the Chesapeake and Delaware bays.

Species' ranges respond to these large-scale ocean boundaries, as well as to species' physiological and behavioral adaptations. Fishes are a case in point. Of the almost 1100 East Coast fish species, 556 species presently occur in the Virginian–Carolinian region (Ray, 1997; Ray *et al.*, 1997). Estuary-dependent species are drawn from this species pool. As discussed earlier (*see* Estuarine Biodiversity), "estuary-dependent" usually has been interpreted very broadly. Occurrence and even abundance of fishes in estuaries do not necessarily infer "dependence." Rather, a species must be truly "obligate" in an evolutionary, adaptive sense for this definition to apply; that is, if estuaries were removed, "dependent" species, i.e., those that require estuaries as habitat at some life-history stage, would be at risk of significant depletion, even to the point of local or regional extirpation. According to

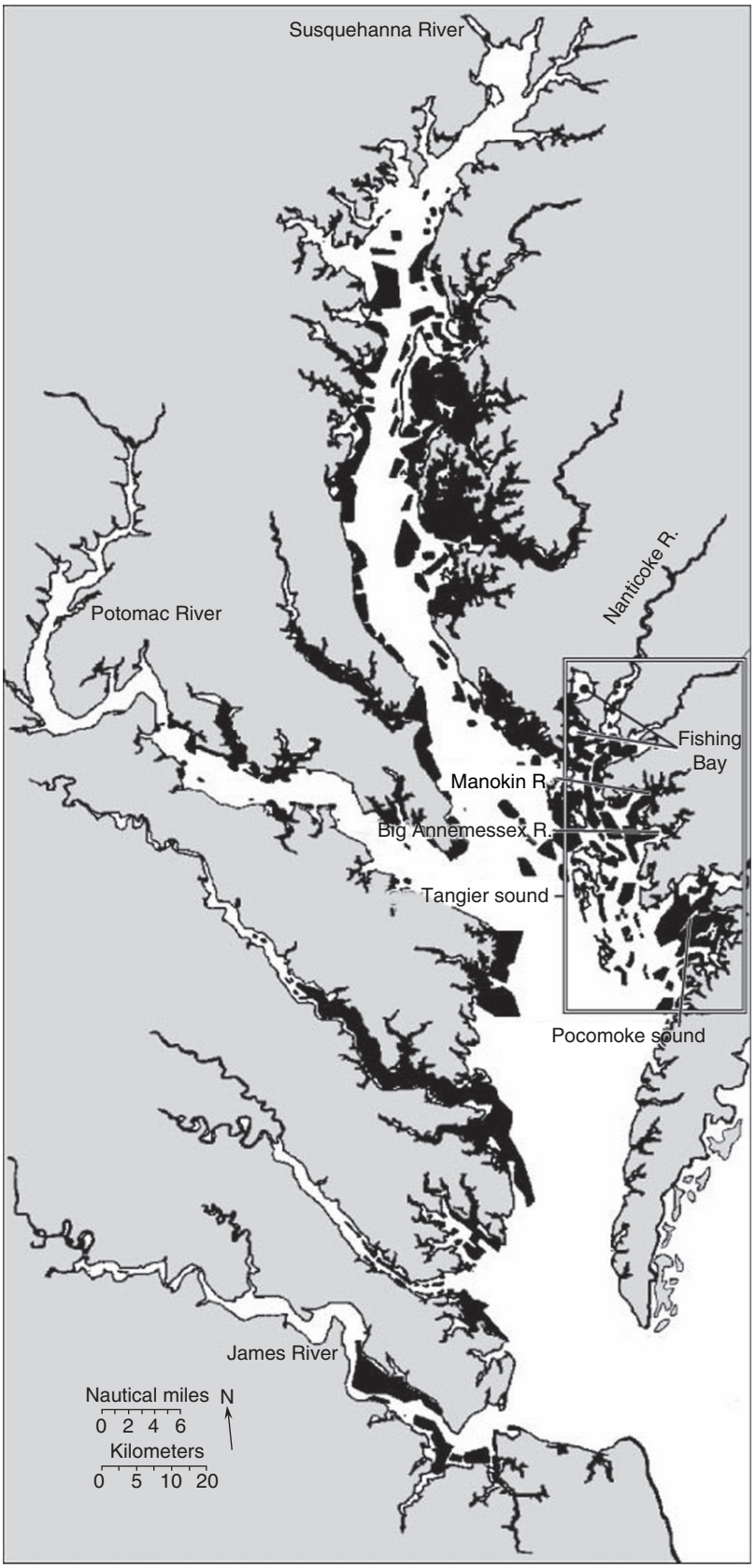


Figure 1 Historical fringing reef of Chesapeake Bay, with the Tangier Sound complex located in box. Reproduced from Ray GC, Hayden BP, McCormick-Ray MG, and Smith TM (1997) Land–seascape diversity of the U.S. east coast coastal zone with particular reference toestuaries. In: Ormond RFG, Gage JD, and Angel MV (eds.) *Marine Biodiversity: Causes and Consequences*, pp. 337–371. Cambridge, United Kingdom: Cambridge University Press.

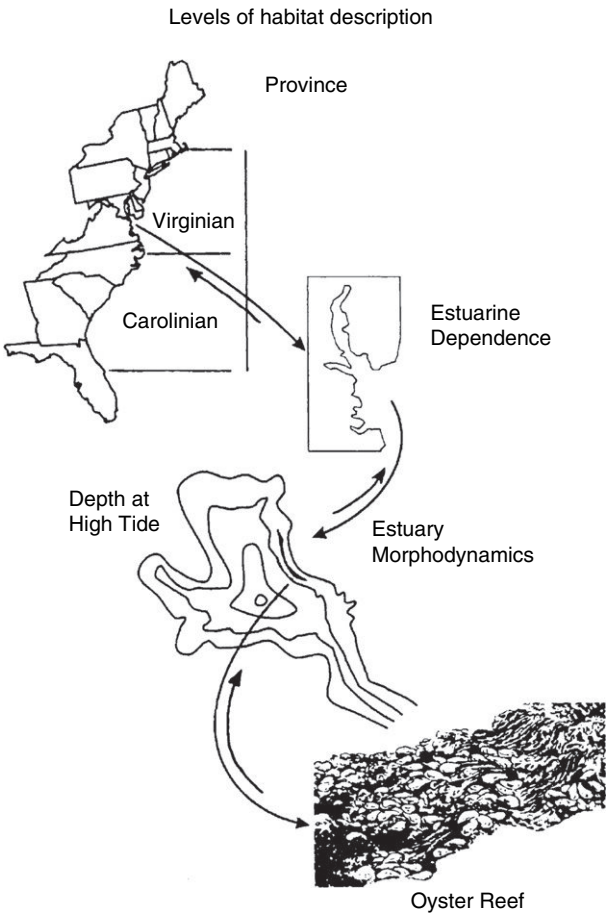


Figure 2 A hierarchical concept of coastal zone relationships, showing top-down "controls" and bottom-up "feedbacks" of coastal zone interactions, involving levels from biogeographic provinces, to estuaries, to the oyster reef. The Virginian biogeographic province provides the species pool from which estuaries may draw "estuary-dependent" representatives. This biota is influenced by the morphometrics of individual estuaries, leading to different species communities among the estuaries of the biogeographic region. The oyster plays the role of a "keystone" species both biologically and ecologically, as the reef it builds provides habitat and may influence the morphometrics of the estuaries in which it occurs. Overharvesting of oysters in the Chesapeake Bay and elsewhere is suggested to have had major effects on estuarine function, structure, and probably biodiversity. Reproduced from Ray GC, Hayden BP, McCormick-Ray MG, and Smith TM (1997) Land–seascape diversity of the U.S. east coast coastal zone with particular reference to estuaries. In: Ormond RFG, Gage JD, and Angel MV (eds.) *Marine Biodiversity: Causes and Consequences*, pp. 337–371. Cambridge, UK: Cambridge University Press.

this definition, 151 species (27% of 556 Virginian species) qualify as "estuary dependent." This figure is remarkably consistent with the results of Able and Fahay (1998: *see* Estuarine Biodiversity). This finding brings up the question: How might changes in estuaries, human-caused or not, influence the composition of these fish assemblages? For insight into an answer, we must examine the dynamics of estuaries themselves.

Many factors interact to characterize an estuary. Among these are drainage and basin area, salinity, tides and mixing, depth, dimension, water column stratification, floods, habitat types, anoxia, and many others. A principal components analysis (Ray *et al.*, 1997) revealed five components that may influence biological diversity: estuarine dimensions, dominance of marine processes, codominance of marine and freshwater processes, fjord-like attributes, and surface area. The interplay of these factors may be used to classify estuaries into the following types: (1) those that are long and wide with extensive catchment areas; (2) large, embayed, well-stratified estuaries with extensive seawater zones; (3) marine-dominated, deep, and well-stratified estuaries; (4) long and narrow, fjord-like estuaries, with large tidal prisms; and (5) estuaries with large surface areas. Chesapeake Bay falls somewhere between categories (1) and (2).

Based on what we know of the natural histories of biota, it seems reasonable to assume that these estuarine types would host different communities of species, and that different disturbance regimes would differentially affect estuarine types and their species' communities. The obvious conclusion is that biotic communities differ among estuaries, and seasonal or weather-related changes in salinity and other factors will be reflected by variable biotic pattern. Also, because the great majority of estuarine fishes also occur over the continental shelf, fluctuations of estuarine fish communities would also be reflected by alterations of shelf-fish communities. But this cannot explain what might be the ecological consequences of these alterations, or whether the ecological role that any organism may play might be lost following that species decline. Clearly, the decline or removal of a foundation species (e.g., oyster reefs), will influence the biological diversity of the system by influencing environmental conditions through environmental feedbacks across scales. However, what might be the effect of declining fishes?

Hypothetically, changes in fish communities have also had ecological effects; how might their roles be conserved? For example, Chesapeake Bay is host to eight abundant species of filter-feeding Clupeid fishes: Atlantic thread herring (*Opisthonema oglinum*), gizzard shad (*Dorosoma cepedianum*), Atlantic menhaden (*Br. tyrannus*), Atlantic herring (*Clupea harengus*), American shad (*Alosa sapidissimus*), hickory shad (*Alosa mediocris*), alewife (*Alosa pseudoharengus*), and blueback herring (*Alosa aestivalis*). These fishes seem to have evolved opportunistic, metapopulation life styles comprising a guild in which these species complement one another in terms of responses to environmental conditions, and in which the ecological function of planktivory of this guild is not likely to be severely compromised under normal scenarios of environmental change (Ray, 2005). Another guild for sciaenids (benthic-feeding croakers) is likely. Thus, it seems reasonable to visualize Chesapeake Bay and other estuaries as composed of diverse species guilds, such that ecological function is conserved, and so is the resiliency of the estuarine system – at least, until severe environmental impacts occur.

For the Chesapeake Bay, oyster reefs are especially critical because they form hard bottom structures that interact with hydrology and influence biodiversity at many levels (McCormick-Ray, 1998). The location of these reefs is not

accidental, reflecting patterns in salinity, hydrology, and geometry of the estuarine basin, tidal stream channels and meanders, and other basin-wide factors. Reefs also influence estuarine development, sedimentation, and water clarity, and thus the formation of other habitats (e.g., submerged aquatic vegetation, marshes, soft bottoms, and hard bottoms) for a host of organisms. And as climatic, interannual, and interseasonal changes mobilize variable hydrologic conditions, the role of rooted and cemented estuarine habitats such as oyster reefs, seagrasses, and marshlands together play interdependent roles in providing shelter, stability, and predictable habitat. This may be of fundamental importance to functional development of the estuarine land–seascape. In sum, the eastern oyster appears to be a classic example of a "foundation" species at the level of the estuarine ecosystem. Similarly, the role of fish communities must not be overlooked.

Another type of functional feedback concerns the fact that many species exist in a number of separate populations, either as young or as adults, that mix together seasonally especially in nonreproductive periods. A common pattern of estuarine-associated fishes, for example, is estuarine or coastal spawning followed by juvenile egress from estuaries or larval ingress into estuaries with settlement by juveniles in nursery habitats. As estuarine-associated species use a diversity of habitats for reproduction, growth, and recruitment, these patchy populations at other times assemble as metapopulations or contingent memberships (Secor and Rooker, 2005; Secor, 1999). For example, an estuary-dependent species such as menhaden, which is distributed patchily among individual estuaries as juveniles, assembles over the shelf to form one or more metapopulations as adults (**Figure 3**). These shelf metapopulations join those of other species to become part of the shelf "metacommunity." It follows that fluctuations of any one metapopulation within any one estuary will affect the total "metacommunity" to a greater or lesser extent (Ray, 1997). This form of large-scale biodiversity centers on community composition, not necessarily on the presence or absence of individual species, but may be strongly affected by functional alterations of estuaries. Observing the collection of species in patterns of communities leads to the conclusion that at the scale of the large, regional ecosystem, each estuary with its community components may be conceived in terms of connectivities to the sum total of estuaries responsible, to a greater or lesser degree, for the overall functional dynamics of the biogeographic region. This large-scale perspective fuses scientific concepts of landscape ecology with metapopulation theory.

The concepts presented in the case of the Chesapeake Bay suggest controls and feedbacks among organisms and the environment at several scales. One fundamental factor seems clear: east coast estuaries have been perturbed in many ways, and one of the most dramatic for the Chesapeake Bay has been the depletion of oyster reefs and some fishes, for example, shad, and the practical eradication of their functional ecosystem roles. Although data are lacking to explain beyond doubt what changes have been perpetrated by these biodiversity losses, it seems apparent that at the very least they have left a marked effect on the entire Bay system. Furthermore, it may be that these effects have cascaded up-scale to affect the adjacent continental shelf.

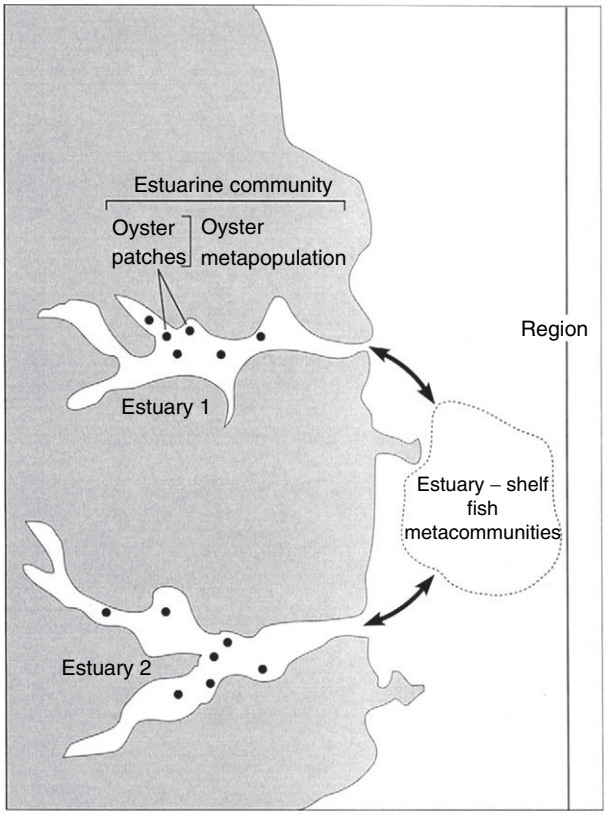


Figure 3 The concept of estuarine metapopulations and shelf metacommunities. Oyster reef metapopulations influence estuarine morphometrics and biodiversity. Consequently, the fish biota of various estuaries influence the fish metacommunity of the shelf.

Future Challenges

We make three points in conclusion. The first concerns the need for greatly increased attention to the natural histories of estuarine and shelf species, which underlie both theory and management practice. The minimal requirements for informed conservation and management are descriptions of species' life histories in the context of their environmental relationships.

Second, many estuarine organisms range widely and form metapopulations over the shelf, as components of estuary–shelf communities. Thus, the minimal scale for sustainability of biodiversity becomes that of the biogeographic region. Quantitative, landscape-level descriptions of the regional coastal zone, including estuarine habitats, are a necessary prerequisite for conservation and management.

Third, it has become a truism in ecology that no one scale adequately describes ecosystem phenomena. Rather, the interaction among phenomena on different scales must become the centerpiece of research and management. This strongly suggests that explanations for fluctuations in biodiversity, including those within biotic communities and at regional scales, will continue to be obscure until multiscale ecosystem functions are better understood. Ecosystem-based management is the logical outcome of interdisciplinary, multiscale knowledge. This recognizes that understanding the

ecology and diversity of coastal zone biota depends in large part on understanding land–sea and estuarine interactions, and also on the joint application of metapopulation and land–seascape theory and methods.

The National Research Council (NRC, 1995) stated that a major future research objective is "to understand the patterns, processes and consequences of changing marine biological diversity by focusing on critical environmental issues and their threshold effects, and to address these effects at spatial scales from local to regional." This objective cannot be met absent a specific consideration of estuaries as major, scale-dependent pathways of biotic and abiotic interchanges. Estuarine biodiversity, structure, and function have been severely modified by humans around the globe. Nevertheless, many estuaries remain either good candidates for restoration and/or relatively rich, productive, and resilient. Documentation of impacts is severely hampered by lack of long-term baseline information, inadequate assessment of biodiversity, lack of trained taxonomists, and difficulty in sampling. Nevertheless, an extensive estuarine literature is now available, illustrating that control of pollution, development, excessive natural resource extractions, and changes in ecosystem function urgently need to be addressed. Problems may not be eliminated, only ameliorated, but increased understanding is essential for the future sustainability of estuaries. Carriker (1967) put the matter boldly more than four decades ago: "There is consequently an urgency to study estuaries before unenlightened defacement obliterates them and before it becomes expedient to investigate them primarily as outdoor pollution laboratories."

See also: Coastal Beach Ecosystems. Freshwater Ecosystems. Intertidal Ecosystems. Lake and Pond Ecosystems. Marine Ecosystems. River Ecosystems. Wetlands Ecosystems

References

Able KW and Fahay MP (1998) *The First Year in the Life of Fishes in the Middle Atlantic Bight*. New Brunswick, NJ: Rutgers University Press.

Anonymous (1959) Archivo di Oceanografia e Limnologia, vol. II, supplemento. *Symposium on the Classification of Brackish Waters*. 8–14 April 1958, Venice, Italy.

Bilby RE, Fransen BR, and Bisson PA (1996) Incorporation of nitrogen and carbon from spawning coho salmon into the trophic system of small streams: Evidence from stable isotopes. *Canadian Journal of Fisheries and Aquatic Science* 53: 164–173.

Boicourt WC (1982) Estuarine larval retention mechanisms on two scales. In: Kennedy VS (ed.) *Estuarine Comparisons*, pp. 445–457. New York: Academic Press.

Bulger AJ, Hayden BP, Monaco ME, Nelson DM, and McCormick-Ray MG (1993) Biology-based estuarine salinity zones derived from a multivariate analysis. *Estuaries* 16(2): 311–322.

Carriker MR (1967) Ecology of estuarine benthic invertebrates: A perspective. In: Lauff GH (ed.) *Estuaries*, pp. 442–487. Washington, DC: American Association for the Advancement of Science Publication No. 83.

Crossland CJ, Baird D, Ducrotoy J-P, and Lindeboom H (2005) The coastal zone – A domain of global interactions. In: Crossland CJ, *et al.* (eds.) *Coastal Fluxes in the Anthropocene*, pp. 1–37. Berlin, Heidelberg/New York: Springer-Verlag.

Dame RF, Spurrier JD, and Wolaver TG (1989) Carbon, nitrogen and phosphorus processing by an oyster reef. *Marine Ecological Progress Series* 54: 249–256.

Emery KO (1967) Estuaries and lagoons in relation to continental shelves. In: Lauff GH (ed.) *Estuaries*, pp. 9–11. Washington, DC: American Association for the Advancement of Science Publication No. 83.

Garman GC and Macko SA (1998) Contribution of marine-derived organic matter to an Atlantic coast, freshwater tidal stream by anadromous clupeid fishes. *Journal of the North American Benthological Society* 17(3): 277–285.

Grabowski JH (2004) Habitat complexity disrupts predator–prey interactions but not the trophic cascade on oyster reefs. *Ecology* 85: 995–1004.

Hanski I and Gilpin M (1991) Metapopulation dynamics: Brief history and conceptual domain. *Biological Journal of the Linnaean Society* 42: 3–16.

Hayden BP, Ray GC, and Dolan R (1984) Classification of coastal and marine environments. *Environmental Conservation* 2: 199–207.

Healey MC (1982) Juvenile salmon in estuaries: The life support system. In: Kennedy VS (ed.) *Estuarine Comparisons*, pp. 315–341. New York: Academic Press.

Hesslein RH, Capel MJ, Fox DE, and Hallard KA (1991) Stable isotopes of sulfur, carbon and nitrogen as indicators of trophic level and fish migration in the lower Mackensie River basin Canada. *Canadian Journal of Fisheries and Aquatic Sciences* 48: 2258–2265.

Heywood VH and Watson RT (eds.) (1995) *Global Biodiversity Assessment*. Cambridge, UK: Cambridge University Press. Published for the United Nations Environment Programme by.

Holligan PM and Reiners WA (1992) Predicting the responses of the coastal zone to global change. *Advances in Ecological Research* 22: 211–255.

Houde EW (1997) Patterns and consequences of selective processes in teleost early life histories. In: Chambers RC and Trippel EA (eds.) *Early Life History and Recruitment in Fish Populations*, pp. 173–196. London: Chapman and Hall.

Ketchum BH (ed.) (1972) *The Water's Edge: Critical Problems of the Coastal Zone*. Cambridge, MA: Massachusetts Institute of Technology Press.

Kirby MX and Miller HM (2005) Response of a benthic suspension feeder (*Crassostrea virginica* Gmelin) to three centuries of anthropogenic eutrophication in Chesapeake Bay. *Estuarine, Coastal and Shelf Science* 62: 679–689.

Lehnert RL and Allen DM (2002) Nekton use of subtidal oyster shell habitat in a southeastern U.S. estuary. *Estuaries* 25: 1015–1024.

Lenihan HS, Peterson CH, Byers JE, Grabowski JH, Thayer GW, and Colby DR (2001) Cascading of habitat degradation: Oyster reefs invaded by refugee fishes escaping stress. *Ecological Applications* 11: 764–782.

Levin SA (1998) Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* 1: 431–434.

Mann KH (1982) *Ecology of Coastal Waters a Systems Approach*. Berkeley: University of California Press.

Mann KH and Lazier JRN (1991) *Dynamics of Marine Systems: Biological–Physical Interactions in the Oceans*. Boston: Blackwell Scientific Publications.

McCormick-Ray J (2005) Historical oyster reef connections to Chesapeake Bay: A framework for consideration. *Estuarine, Coastal and Shelf Science* 64: 119–134.

McCormick-Ray MG (1998) Oyster reefs in 1878 seascape pattern – Winslow revisited. *Estuaries* 21: 784–800.

Nelson JS (1984) *Fishes of the World*, 2nd edn New York: Wiley–Interscience.

NERC (1992) *Land–Ocean Interaction Study (LOIS): Science Plan for a Community Research Project*. Swindon, Wiltshire, UK: Natural Environment Research Council.

NRC (1995) *Understanding Marine Biodiversity*. Washington, DC: National Research Council.

Pernetta JC and Milliman, JD (eds.) (1995) Land–ocean interactions in the coastal zone: Implementation plan. *Global Change Report No. 33*. Stockholm: International Geosphere–Biosphere Programme of the International Council of Scientific Unions (ICSU).

Pritchard DW (1967) What is an estuary: Physical viewpoint. In: Lauff GH (ed.) *Estuaries*, pp. 3–5. Washington, DC: American Association for the Advancement of Science Publication No. 83.

Ray GC (1997) Do the metapopulation dynamics of estuarine fishes influence the stability of shelf systems? *Bulletin of Marine Science* 60: 1040–1049.

Ray GC (2005) Connectivities of estuarine fishes to the coastal realm. *Estuarine, Coastal and Shelf Science* 64: 18–32.

Ray GC, Hayden BP, McCormick-Ray MG, and Smith TM (1997) Land–seascape diversity of the U.S. east coast coastal zone with particular reference to estuaries. In: Ormond RFG, Gage JD, and Angel MV (eds.) *Marine Biodiversity: Causes and Consequences*, pp. 337–371. Cambridge, UK: Cambridge University Press.

Rothschild BJ, Ault JS, Goulletquer P, and Heral M (1994) Decline of the Chesapeake Bay oyster population: A century of habitat destruction and overfishing. *Marine Ecology Progress Series* 111: 29–39.

Roy PS (1984) New South Wales estuaries: Their origin and evolution. In: Thom BG (ed.) *Coastal Geomorphology in Australia*, pp. 99–121. Orlando, FL: Academic Press.

Sanders HL (1968) Marine benthic diversity: A comparative study. *American Naturalist* 102: 243–282.

Secor DH (1999) Specifying divergent migrations in the concept of stock: The contingent hypothesis. *Fisheries Research* 43: 13–34.

Secor DH and Rooker JR (2005) Editorial, connectivity in the life histories of fishes that use estuaries. *Estuarine, Coastal and Shelf Science* 64: 1–3.

Solbrig OT (1991) From genes to ecosystems: A research agenda for biodiversity. *Report of a IUBS–UNESCO Workshop June 1991, Harvard Forest, MA*. Paris: International Union of Biological Sciences.

Steele JH (1991) Marine functional diversity. *BioScience* 41: 470–474.

Ulanowicz RE and Tuttle JH (1992) The trophic consequences of oyster stock rehabilitation in Chesapeake Bay. *Estuaries* 15: 298–306.

Winemiller KO (1995) *Fish Ecology, Encyclopedia of Environmental Biology*, vol. 2. Orlando, FL: Academic Press. pp. 49–65.

Ethical Issues in Biodiversity Protection

Philip J Cafaro, Colorado State University, Fort Collins, CO, USA
Richard B Primack, Boston University, Boston, MA, USA

Published by Elsevier Inc.

Glossary

Anthropocentrism Position that only human beings have moral worth or intrinsic value.
Ecosystem health Condition of an ecosystem, whether natural, managed, or human-dominated, that is free from influences that would damage or destroy its characteristic structures and functions.
Ecosystem integrity Condition of an ecosystem that is largely free from human interference and possesses a species composition and functional organization comparable to those of natural ecosystems in the region.
Instrumental value Value of something relative to human interests or desires.
Intrinsic value Value of something independent of its value to people.
Rights Justified claims that people respect a being's existence and important interests, and act accordingly.
Sustainability Ability of a society, institution, business, or any individual or collective human activity to continue indefinitely without depleting resources or damaging the environment.
Wilderness Large area that remains essentially unmanaged and unmodified by human beings.

Conservationists' Ethical Consensus

There is widespread disagreement in modern societies regarding the proper human relationship to the rest of the natural world. Conservationists and conservation biologists, however, mostly agree on the following ethical principles: the diversity of organisms is good; ecological complexity and natural evolution are good; the untimely extinction of populations, species, and biological communities is bad; biological diversity has great value both to people and in its own right; and human beings have both strong altruistic and strong self-interested reasons for preserving biodiversity (Primack, 2010).

Conservationists justify these principles in very different ways. One may value biodiversity because it is God's creation and testifies to God's glory; another because it is the natural culmination of evolution and global forces that have occurred over hundreds of millions of years; a third because of its value to science, perhaps humanity's noblest pursuit; a fourth because of its beauty and the enjoyment that current and future generations may take in it; a fifth because its intrinsic value grounds a strong duty to protect it, regardless of human interests; a sixth because preserving biodiversity helps preserve human life-support systems. Such justifications may be more or less anthropocentric – centered on human interests. However different their ultimate justifications, conservationists tend to accept these ethical principles and work with others to preserve biodiversity.

Difficult ethical and practical issues emerge when we ask two further kinds of questions. First, what do these general principles entail? Given a high valuation of nature and a sense of personal responsibility to help protect it, what exactly is required of us as scientists and citizens? Second, how does our concern to protect biodiversity mesh with our obligations to humans? What if they clash? Such questions come up in numerous practical contexts and must be answered, although we cannot wait for certainty or complete consensus in order to act.

Conservation Goals: Health, Integrity, Sustainability

For much of the 20th century, public and private land management goals have been primarily economic. It has become increasingly clear, however, that an exclusive focus on economic productivity leads to environmental degradation: pollution, the extinction of species, and the creation of a progressively simplified landscape (Daly, 1996). In response, laws have been passed directing government land managers to preserve the health and integrity of ecosystems. A focus on short-term economic returns also tends to undermine the long-term economic productivity of biological systems, such as forests and fisheries. Recognizing this, governments have begun to promote sustainability as an ideal in the use of natural resources. Conservationists hope that health, integrity, and sustainability specify goals concrete enough to guide us in our actions while commending themselves as basic values that a wide variety of people will share.

Health

Responding to widespread criticism of its forestry practices, Canada's Provincial Forest Ministers agreed in 1992 that "our goal is to maintain and enhance the long-term health of our forest ecosystems, for the benefit of all living things both nationally and globally" (Westra and Lemons, 1995). Government agencies around the world have similarly embraced the notion of land health. Participants in a major symposium on ecological health defined it as follows:

> An ecological system is healthy and free from "distress syndrome" if it is stable and sustainable – that is, if it is active and maintains its organization and autonomy over time and is resilient to stress (Costanza *et al.*, 1992).

Just as a healthy human being is free from disease and able to perform his or her characteristic functions well, a healthy

forest or stream is free from air or water pollution, siltation, or invasions of exotic organisms. Such stresses impede or radically alter the systems' natural functions (such as photosynthesis or net primary productivity) and structure (often eliminating sensitive species and simplifying biological communities) (**Table 1**).

Healthy ecosystems provide both natural biodiversity and a wide array of products and services valued by humans. For this reason, economic productivity (as measured in board feet, fish caught, or crops yielded) can sometimes stand as one measure of overall ecosystem health. Rapidly declining fish catches in Lake Victoria, due to the introduction of the Nile perch, signaled a decrease in the system's health. However, economic productivity may sometimes be increased by a radical simplification of an ecosystem, as when diverse, natural forests are converted to single-species, even-aged pine plantations, or when natural prairies are converted to cattle ranches (**Figure 1**). Such simplified ecosystems are not necessarily economically productive in the long run, but if they are, then whether they are seen as unhealthy depends on whether preservation of native biodiversity is part of our definition of land health.

If simplified ecosystems are managed well, they may stay healthy for humans and a reduced biota, remain stable and resilient to stress, not pollute surrounding ecosystems (e.g., not dump excessive silt or fertilizer in streams), and continue to perform valuable ecosystem services. Because some relatively intensive land use is necessary for human survival and this can be done better or worse, it makes sense to apply concepts of land health to simplified ecosystems. At the same time, we must recognize that preserving as much biodiversity as possible in managed systems (croplands, lakes, forests) is desirable and that excessive simplification across a large landscape leads to extinction of species and loss of characteristic natural communities. Ecosystem health must be supplemented with the ideal of ecosystem integrity – at least in some areas (Pimentel *et al.*, 2000) (*see* Conservation Goals: Health, Integrity, Sustainability, below).

We should also remember an important difference between human health and ecosystem health. Individual human beings grow old, decay, and die, and while we fight this we also accept it. In contrast, we want healthy ecosystems in perpetuity, for their own good and the good of our descendants. Thus ecosystem health must also be supplemented with the ideal of ecological sustainability (*see* Conservation Goals: Health, Integrity, Sustainability, below).

Integrity

Amendments to the U.S. Clean Water Act of 1972 first set ecological integrity as a management goal by calling for the restoration and maintenance of "the chemical, physical, and biological integrity of the nation's waters." Rolston (1994) noted that "both integrity and health are combined fact-value words. Both convey the idea of wholeness and of unbroken functioning." But ecosystem integrity implies a greater measure of freedom from past and current human manipulation and a closer approximation to natural structure and functioning than does ecosystem health. Karr and Dudley (1981) defined ecological integrity as "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitats of the region."

Integrity is a key environmental ideal because it encompasses the full preservation of biological diversity, including individual organisms, species, natural communities, and the ecosystem processes that have created and sustained them (Westra *et al.*, 2008). It has taken a long time for this ideal to be acknowledged, even in supposedly fully preserved areas. For example, wolves, mountain lions, and other predators were routinely shot, trapped, and poisoned in US national parks up until the 1960s. It is only in the last 10 years that wolves, once extirpated, have been reintroduced into Yellowstone National Park. For most of this century, fires were

Table 1 Characteristic response of ecosystems to stress

Types of stress	*Nutrient pool*	*Primary productivity*	*Size distribution*	*Species diversity*	*System retrogression*
Harvesting of renewable resources					
Aquatic	*	*	−	−	+
Terrestrial	−	−	−	−	+
Pollutant discharges					
Aquatic	+	+	−	−	+
Terrestrial	−	−	−	−	+
Physical restructuring					
Aquatic	*	*	−	−	+
Terrestrial	−	−	−	−	+
Introduction of exotics					
Aquatic	*	*	*	−	+
Terrestrial	*	*	*	*	+
Extreme natural events					
Aquatic	*	*	−	−	+
Terrestrial	−	−	−	−	+

Note: Signs (+ or −) indicate direction of change compared with normal functioning of relatively unstressed systems. An asterisk indicates that a characteristic response was not sufficiently determined.

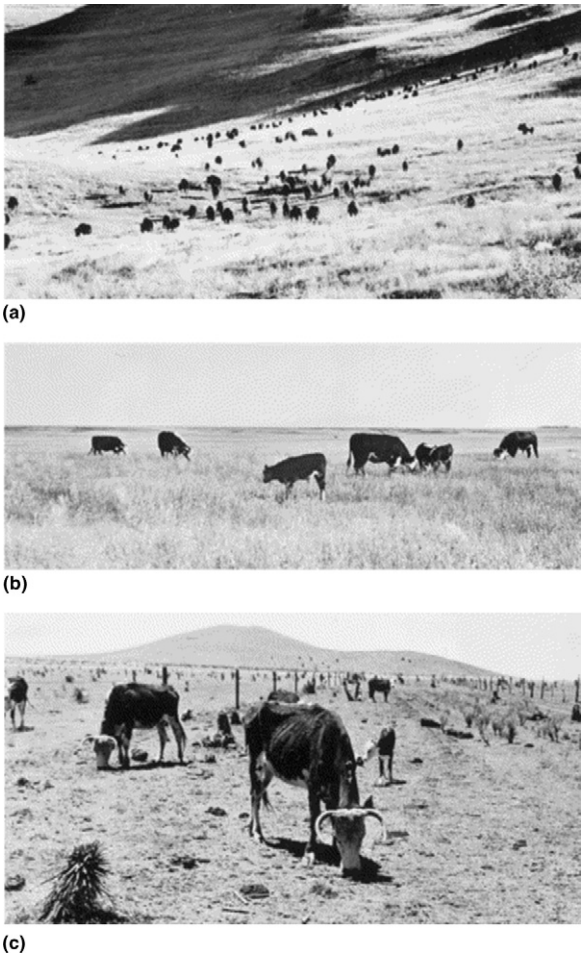


Figure 1 Healthy and unhealthy rangelands. (a) A natural grassland with numerous native species on the National Bison Range, federally protected land in the state of Montana. (b) Cattle graze on natural grassland. (c) Overgrazed grassland takes on the appearance of a desert and native species are eliminated. Photographs courtesy of the U.S. Fish and Wildlife Service and the U.S. Forest Service.

suppressed throughout the national parks; they were perceived as dangerous to both people and forests. Over time, however, fire ecologists documented the historical role of fires in creating the natural landscape and preserving fire-dependent species of plants and animals. In response, land managers have introduced controlled burning into parks and allowed some lightning-ignited fires to burn. Improved ecological understanding and an attempt to distinguish between what is good for the ecosystem and what is comfortable or familiar for people have shown that wolves and fire belong in the parks.

The fact that ecosystems are not as strongly integrated, clearly bounded, or stable as individual organisms complicates attempts to specify ecological integrity in particular cases. It is sometimes difficult to stipulate which outside influences represent assaults on integrity rather than mere changes. Recent study suggests that many ecosystems are relatively loose assemblages of species; that these assemblages may be relatively recent and transient creations; and that even left to themselves many natural communities will not necessarily reach a particular, invariant climax state but may instead reach any one of a number of more or less stable states, or no stable state whatsoever. If species assemblages are always changing naturally, why distinguish species that are extirpated or introduced by humans, as when conservationists reintroduce wolves in the Rockies or eradicate wild boars in the Great Smoky Mountains? If there is no one natural end point to succession, why assume that human-created early-successional stages (fields of shrubs after clear-cuts) have less integrity than naturally occurring late-successional stages (old-growth forests)?

Still, ecosystems typically go through characteristic successional stages and support characteristic (if not invariant or exact) species assemblages. Though natural species assemblages change, this usually occurs on a time-scale that allows for much stability, the development of complex interactions between organisms, and the increase of biological diversity at the landscape level. In contrast, human-induced changes typically decrease biodiversity and always lead to a landscape that is partly our creation. This loss of independence arguably marks a qualitative change in an ecosystem's natural history and a corresponding loss of ecosystem integrity (Foreman, 2004). Applying this complex concept to particular ecosystems thus involves attention to scale, knowledge of an ecosystem's particular history, and comparison with the structures and functioning of similar ecosystems. Maximizing ecological integrity involves balancing the sometimes conflicting goals of freedom from human interference and preservation of historical natural communities.

Despite these complications, the desire to restore degraded ecosystems demands robust conceptions of ecosystem integrity and health. Ecologists attempting to restore ruined pasture-lands to native prairie in Wisconsin, worked-out strip mines to forests in Appalachia, or drained lands back to wetlands in Florida, must set specific objectives (**Figure 2**). As with land managers seeking to limit their effects on relatively pristine areas, restorationists have taken health and integrity, defined in relation to natural baselines, as goals for restored landscapes.

Sustainability

"Live Sustainably!" has become a rallying cry in the environmental movement. Like health and integrity, sustainability is a term that blends facts and values, helping us act on our environmental convictions. At a minimum, sustainability involves preserving resources for future generations. But this minimum will not satisfy conservation biologists, who insist that part of what must be sustained is the full complement of biological diversity, for its own sake and for the benefit of humans.

Wide differences exist here. For instance, in 1987 the United Nations World Commission on Environment and Development (the Brundtland Commission) defined sustainable development as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs." The focus here is wholly on human beings. On this definition the mass extinction of species is sustainable, provided future generations of people can meet their self-defined needs. Noss summarized the



(a)



(b)

Figure 2 Restoration of ecological integrity. (a) In the late 1930s, members of the Civilian Conservation Corps participated in a University of Wisconsin project to restore the wild species to a Midwestern prairie. (b) The prairie as it looked 50 years later. (Photographs from the University of Wisconsin Arboretum and Archives).

shortcomings of such narrow, anthropocentric accounts of sustainability as follows:

> A failure of those who promote sustainability to consider environmental and social limits to growth; an unwillingness to address the unsustainability of the current human population, much less its expected growth; a reluctance to confront the implications of the lifestyles of average citizens of the more affluent societies . . . a failure to recognize the claims of other species to their share of the planet's resources (Westra and Lemons, 1995).

Contrast the Brundtland definition with Barbier's more generous definition of sustainable development: "to maximize simultaneously the biological system goals (genetic diversity, resilience, biological productivity), economic system goals (satisfaction of basic needs, enhancement of equity, increasing useful goods and services), and social system goals (cultural diversity, institutional sustainability, social justice, participation)" (Munda, 1997). Such a definition implies a different conception of human development and a more restrained treatment of the nonhuman world than has prevailed up to this point in human history. Satisfying the basic needs of all people and creating just and flourishing societies are part of this goal; limitless wealth creation, unbounded consumerism and ever more people are not (Daly and Farley, 2011).

Conservationists are particularly concerned to keep the definition of sustainability based in ethics and biology and not solely in traditional economics (Daly, 1996). Economically based definitions often define sustainability in terms of indefinite economic growth, but in the crowded, fossil-fuel-propelled world of the 21st century, continued emphasis on economic growth will inevitably undermine ecosystem health and integrity worldwide (Speth, 2009). We cannot sustain increased consumption and increased human populations and protect biodiversity at the same time.

Management and Nonmanagement

Setting aside parks and wilderness areas where human beings are not permitted to use resources has been conservationists' most effective tool in conserving biological diversity. Here the results from conservation biology are clear. Wild areas preserve greater numbers of native species than do areas managed for agriculture or forestry. Large protected areas are more likely than smaller ones to preserve full complements of native species. Some species need large areas of habitat to preserve minimum viable populations. Some species tolerate little human disturbance. Larger areas are less likely to suffer "edge effects" that render otherwise good habitat unusable by some species (Nelson and Callicott, 2008).

A dilemma arises, however, because in order to preserve certain wild species or communities, we may have to actively manage natural areas. Rare, threatened plant populations may have to be fenced off from browsers, or new populations may need to be created from seeds or cuttings. Small, widely scattered populations of an endangered mammal may have to be gathered together to preserve genetic variety or create viable populations. Controlled burning may have to be carefully introduced into a forest or field to preserve fire-dependent species (**Figure 3**).

Most conservationists accept reintroductions of indigenous species into wild areas and other manipulations that have as their goal the preservation or recreation of indigenous flora and fauna. But they sharply distinguish actions that have this goal from actions that support further resource extraction or tourism development (road building, tree harvesting, construction of lodges and restaurants), which they continue to oppose in wild areas. Further, they see value in the absence of manipulation itself and tend to view the need for heroic measures in pristine areas as an indication of what has already been lost.

As a practical matter, hands-on management will often be necessary to preserve biodiversity. Even our largest natural areas are increasingly "islands" of diversity surrounded by development. These islands will likely lose species over time, as predicted by conservation biology theory and confirmed by empirical study; more native biodiversity may be preserved by augmenting populations and managing habitat. Other semi-wild areas will continue to be utilized by humans; these areas support wild species that may also need management in order to survive. Some manipulation will be needed, in perpetuity, if we are going to preserve as much biodiversity as possible.

It should be noted, however, that heroic measures and intrusive management often fail to preserve species. Skeptics



(a)



(b)

Figure 3 Conservation management: Intervention vs leave-it-alone. (a) Heathland in protected areas of Cape Cod, Massachusetts, is regularly burned to maintain open habitat and protect wildflowers and other rare species (Photograph by Dunwiddie P). (b) This old-growth stand in the Olympic National Forest in Washington is the result of many years of keeping human disturbance to an absolute minimum (Photograph by Thomas Kitchin/Tom Stack & Associates).

also warn that a casual acceptance of manipulation may lead to a loss of biodiversity, if all sorts of human purposes are allowed to vie with nature preservation on our park and forest lands. Too often, management of nature becomes a substitute for management of ourselves (Rolston, 1994). Rather than close popular trails or campgrounds in Yellowstone National Park, for example, the U.S. National Park Service exterminates "problem" grizzly bears, despite their status as an endangered species.

Arguably, too, wilderness is itself a part of biodiversity and not just a means to preserve it. Conservation biologists commonly define biodiversity to include species diversity, genetic diversity, and community diversity: the latter is defined as different biological communities and their associations with the physical environment (the ecosystem). But if biodiversity includes biological communities and ecosystems, their disappearance or development into something essentially different constitutes a loss of biodiversity.

So, conservationists must make hard choices to protect biodiversity. For example, some models of global climate change suggest that more than 10% of the plant species in many U.S. states will not be able to survive new climatic conditions in the next few centuries – they will have to migrate northward or die out. If species are in danger of going extinct in the wild because of climate change, the last remaining individuals may have to be preserved by us; hopefully, wild populations of these species can then be reestablished in new protected areas where the climate is suitable for them. In these cases of ecological triage, we should arguably move from a wild species – not our creation, not domesticated – to a wild species, living and continuing to evolve in the wild. The goal should be a future in which species survive without having to be moved, monitored, or supported by us and in which we rein in the human overpopulation, overdevelopment, and overconsumption that make such biodiversity management necessary. By managing ourselves more wisely, we can limit the need to manage wild nature.

Biodiversity Protection and Animal Rights

Conservationists typically argue that preserving species and whole biological communities should take precedence over preserving individual animals. In practice, this means that exotic animals may be sacrificed when necessary to preserve native species and overpopulous native or exotic animals may be sacrificed to preserve land health. These positions have provoked arguments with animal rights advocates, who might seem to be conservationists' natural allies.

For example, U.S. government agencies judged the continued existence of the endangered plant Santa Barbara liveforever (*Dudleya traskiae*) to be more valuable than the common rabbits that had been introduced on its island home. The rabbits, which fed on the plant's fleshy leaves, were killed to stop the destruction of this fragile plant species. In another example, state officials sanctioned the culling of hundreds of deer in the Quabbin Preserve in central Massachusetts to prevent over-browsing as a way to allow tree regeneration and protect water quality. Protection of the overall health of the forest and continuation of basic ecosystem functioning (forest regeneration, water purification) outweighed the interests of individual deer.

Such practices seem mere common sense to many conservationists, but many animal rights advocates argue that they are misguided, for two reasons. First, animals (at least some higher animals) have rights or interests based on their consciousness or sentience (their ability to feel pleasure or pain). This means that we cannot sacrifice them casually for other goals, such as a healthy forest. Second, species and forests do not have rights or interests, because these complex aggregates are too loosely organized (see earlier discussion). To sacrifice the real interests of a higher animal to the bogus interests of a species or biological community is thus ethically mistaken (Schmidtz and Willott, 2011).

Conservationists' answer that as we learn more about biodiversity, we realize that its greatest value lies not in individuals but in the wholes that those individuals help constitute. It is the species that persists and evolves rather than the individuals, which come and go. It is the natural community that sustains and generates new forms of life. While an ethics based on individual rights best specifies appropriate human interactions, a less individualistic ethics is needed to properly value the natural world (Westra *et al.*, 2008).

This ethical shift seems to be necessary for effective conservation. However, animal rights advocates have given strong arguments for respecting individual higher animals such as cats, rabbits, and wild horses. Conservationists, who usually value species over individuals, may concede that individual animals should not be sacrificed casually, but only if it truly is necessary to preserve species or the health of biological communities. They might add that individual animals should always be culled humanely, that is, with a minimum of suffering. This would rule out most trapping and much sport hunting, although these activities are routinely defended on conservation grounds.

For example, Quabbin Preserve managers might carefully monitor deer populations and tree regeneration and authorize hunting where appropriate. All deer kills would have to be performed humanely, and in some years no hunts might be authorized. Whereas a traditional wildlife management position would authorize hunting game birds and trapping beavers, mink, and muskrats as long as this was done sustainably, a position responsive to animal welfare might not allow any of these activities, unless they could be justified with reference to the good of the forests or particular species.

Conservation and animal rights ethics both typically lead beyond anthropocentrism – the belief that only human interests matter – and they are usually complementary. For example, a concern for individual animals leads logically to vegetarianism, which prevents much animal suffering and bloodshed. But since "eating lower on the food chain" is less energy, water, and land-use intensive, it also helps conserve resources and preserve biodiversity. Limiting consumption and distinguishing between needs, comforts, and luxuries are precisely what are needed to carry out both ethics. Nothing prevents individuals from living highly consistent lives that incorporate both ideals. We may act to limit animal suffering and preserve wild nature so that the entire realm of biodiversity can flourish.

Biodiversity Protection and Human Rights

Conservationists often argue that individual organisms, species, and biological communities have intrinsic value – that is, value in addition to their usefulness to human beings – that obliges us to preserve them. This means human interests should sometimes be sacrificed to preserve biodiversity. For example, a developer should not build a housing project that destroys critical habitat for an endangered songbird species – whether or not it is legal to do so. The continued existence of this species is more important than the developer's profits. His less-than-vital interests should be foregone or pursued in ways that do not sacrifice something infinitely more valuable: a natural species. In many cases, however, preserving biodiversity and furthering human interests are complementary goals, especially if we take a broad view of human interests. Present and future generations have strong material, esthetic, scientific, recreational, and spiritual interests in preserving biodiversity.

Owls and People

The coniferous forests of the US Pacific Northwest are some of the most majestic on earth, with thousand-year-old trees towering over crystal clear streams filled with salmon. As commercial timber harvesting escalated after World War II, the big trees fell and species dependent on old-growth forests such as the northern spotted owl (*Strix occidentalis caurina*) (**Figure 4**), the marbled murrelet (*Brachyramphus marmoratus*), and various salmon species (*Oncorhynchus* spp.) declined dramatically. In response, environmentalists staged logging blockades and tree sittings during the 1980s and successfully sued the U.S. Forest Service under the Endangered Species Act, resulting in drastic limitations on tree harvests in the 1990s.

This contentious political battle was often billed as "owls versus people," but such a simplistic formulation was misleading. Reduced harvests resulted in the loss of thousands of logging jobs, but some reduction was inevitable, given unsustainably high harvest levels based on the one-time cutting of thousand-year-old trees. Logging reductions helped preserve and create jobs in fishing and tourism and enhanced watershed maintenance, flood control and other objectives. A calculation based on the full spectrum of long-term human interests would likely have resulted in substantially decreased timber harvesting. Still, old-growth forests do have great economic value as timber, and preserving more in reserves rather than cutting the big trees and moving to high-yield forest plantations may indeed lead to lower profits for industry and fewer overall jobs.

Are such trade-offs ever justifiable to prevent the extinction of species or important biological communities? Forestry industry executives and loggers' unions said no. The interests of a logger in supporting his family should supersede the interests of animals or trees. Logging restrictions to protect other species are wrong because they lead to real human hardship for loggers and their communities. Environmentalists insisted that more logging jobs and higher corporate profits did not outweigh the permanent loss of species and the old-growth temperate rain forest itself. They also noted that industry and government had made no efforts to preserve logging and mill jobs when technological changes, log exports, or overharvesting had led to job losses in the past (see **Figure 4**).

These considerations suggest a solution to the dilemma: harness the economic system to maximize both human interests and biodiversity protection. Whether we focus on intrinsic or instrumental value, the human communities of the Pacific Northwest have an interest in preserving their remaining old-growth forests. Fortunately, the regional economy generates enough wealth to provide laid-off workers with guaranteed health care, generous severance pay and job retraining for those who want it. The big trees could be preserved and displaced workers supported by their fellow citizens. Such a course would not reconcile all conflicting interests or solve all problems:



(a)



(b)

Figure 4 Natives. (a) The northern spotted owl is an indicator species for old-growth forests of the Pacific Northwest, a habitat coveted for its rich timber stands. (b) Demonstrators protest government and corporate policies that destroy forests and threaten species extinctions.

economic winners would have to be taxed to support economic losers, some loggers would have to give up a valued livelihood, and some businesses would have to forego profits or even fold. But many of the costs of environmental protection would be borne by society as a whole, which benefits from the preservation of biodiversity.

Property Rights

When environmentalists won suits to limit logging in the Pacific Northwest, the resulting bans affected both private and public lands. To some observers, the restrictions on private lands were unjustified. In their view, the right to own and control property is essential. Dictating to individuals or corporations what they can or cannot do on their own land is seen as intolerable, abridging their freedom. In devising a plan to continue tree harvesting in the region, the US government focused preservation efforts on public lands, allowing private landowners greater freedom.

Does private property ownership give owners unlimited control over their land, even if their actions contribute to the extinction of a rare or endangered species? Historically it often has, but morally it does not. Both the intrinsic value of species and duties to our fellow citizens to preserve a common biological heritage argue against such actions, regardless of how much profit must be foregone. Land ownership confers both rights and responsibilities.

What then of a small landowner who needs to cut an old-growth stand in order to send a child to college or keep the land in the family? Clearly as the need becomes greater, the ethical justification increases, even if some loss of biodiversity occurs. Hopefully, though, this small owner is not so poor that he must view his land solely as a money-generator. The ideal is some compromise in which the land is used in responsible ways that also preserve biological diversity.

Finally, it is important to distinguish between individual small owners and the large timber corporations that rival some national governments in the size of their landholdings. Arguments for preserving the freedom of small forest owners do not justify laissez-faire policies for large corporate landowners. Corporate managers' personal freedoms are not at stake in the same way and their management decisions have vastly greater effects on natural and human communities. No government that cares about preserving its citizens' biological heritage will fail to regulate the environmental impacts of large corporate landowners.

Wilderness in Less-developed Countries

Recently, some writers have argued that biodiversity preservation is a specifically North American or Western preoccupation, whose promotion in less-developed nations amounts to cultural imperialism. According to Ramachandra Guha, for example, wilderness preservation is inappropriate and unnecessary in these countries, whose peoples face more pressing environmental issues centered on meeting basic human needs. He argues that:

> the setting aside of wilderness areas has resulted in a direct transfer of resources from the poor to the rich [in the less-developed nations]. Thus, Project Tiger, a network of parks hailed by the international conservation community as an outstanding success, sharply posits the interests of the tiger against those of poor peasants living in and around the reserve. The designation of tiger reserves was made possible only by the physical displacement of existing villages and their inhabitants; their management requires the continuing exclusion of peasants and livestock … transplanting] the American system of national parks onto Indian soil. In no case have the needs of the local population been taken into account, and as in many parts of Africa, the designated wildlands are managed primarily for the benefit of rich tourists (Schmidtz and Willott, 2011).

There is some truth in Guha's claims. Several of Project Tiger's reserves were built around old hunting preserves from which the poor had long been excluded, while others were set up on former state or communally owned lands and displaced numerous villages, causing real hardship to thousands.

Grazing and tree cutting in the core areas of the reserves were prohibited, both to limit human–tiger conflicts and to preserve a more complete flora and fauna.

On the other hand, India's wild tigers were clearly headed for extinction when Project Tiger was undertaken: between 1900 and 1972, the year it was initiated, the tiger population fell from tens of thousands to just 1800 individuals. Strictly limiting tiger hunting had not halted this decline. Habitat conservation was obviously necessary to protect the tigers over the long term and, given the requirements of tigers and their inevitable conflicts with humans, this had to involve some displacement and restrictions on local inhabitants.

Should Project Tiger have been attempted? Guha suggests not: he believes it represented an unjust appropriation of the resources of poor people. At the same time, a focus on tigers went hand in hand with ignoring "environmental problems that impinge far more directly on the lives of the poor – for example, fuel, fodder, water shortages, soil erosion, and air and water pollution . . . [which are] far more pressing environmental problems within the Third World" (ibid.). Others disagree, arguing that habitat preservation and restrictions on human economic use of this habitat were necessary and appropriate responses to the imminent extinction of the Bengal tiger. This would have been a great cultural loss to the Indian people and an unjust destruction of an intrinsically valuable species. Many conservationists do not accept the extinction of species as a less pressing problem than those mentioned by Guha, because they do not accept his exclusive focus on human interests, nor his narrow definition of those interests (**Figure 5**).

In recent decades, wildlife managers around the world have been experimenting with ways to give local people a stake in the success of wildlife and park conservation, with encouraging results. In Africa, these efforts have included disbursing a percentage of park revenues directly to local villages, increasing efforts to hire local inhabitants as guides and forest wardens, compensating for crop damage done by wildlife straying outside of parks, and giving local people input into protected areas management. In India, beginning in the late 1980s, the Ranthambhore Foundation and other groups began promoting sustainable development in the villages on the periphery of the Project Tiger reserves. This has included replanting denuded forests and improving fodder – to decrease pressure to graze and cut firewood within the reserves – as well as efforts to educate local children in natural history and conservation principles.

Justice demands that local people be treated with respect and their interests considered in all programs to protect wildlife and wildlands. Such respect and consideration will not solve all problems, however, particularly in a country as overpopulated as India, whose human population has increased from 350 million people at independence in 1947 to 1.15 billion people in 2010. In some instances, if we want to preserve wildlife, we will have to sacrifice human interests.

Overall, securing human rights and furthering essential human interests should strengthen efforts to preserve biodiversity. The 1972 United Nations Conference on the Human Environment affirmed that all human beings have fundamental rights to clean air, clean water, pure food, and a healthy environment generally. Preserving these environmental goods tends to benefit wildlife and natural systems as well. Environmentalists have also argued that future generations have a right to know, explore, and celebrate an undiminished natural heritage (Carson, 1962). Respect for this right would halt many environmentally destructive development projects. Natural areas are often threatened by remote governments and corporate managers, while their strongest defenders tend to be local inhabitants who know and love them. Securing rights and political power for local inhabitants should help preserve these areas. Safeguarding human rights and preserving biodiversity are only compatible, however, if people are willing to limit their numbers and their demands on nature. Empowering ever more people with a belief in their right to ever higher consumption is a recipe for ecological disaster.



Figure 5 Valuing biodiversity in less-developed countries. Buddhist priests in Thailand offer prayers and blessings to protect communal forests and sacred groves from commercial logging operations.

The Ideal of the Conservation Biologist

Four hundred years ago, modern science's great pioneer and prophet Francis Bacon located the primary value of science in its creation of a powerful technology for controlling nature. "The End of our Foundation is the knowledge of Causes, and secret motions of things," Bacon wrote in *The New Atlantis*, "and the enlarging of the bounds of Human Empire, to the effecting of all things possible." Without denying science's utilitarian benefits, conservation biologists attempt to move in a different direction, developing the knowledge needed to preserve nature rather than to change, control, or exploit it.



Figure 6 Role models. Primatologists Dian Fossey (left), Jane Goodall (center), and Birute Galdikas began by studying animal behavior, but eventually devoted themselves to conservation activism.

Both historical and contemporary examples attest that science can be carried out in this spirit. Aldo Leopold and Rachel Carson, Paul Ehrlich and E. O. Wilson, and many others from a wide variety of disciplines and backgrounds have combined contributions to science, literature, education, public policy, and conservation. The challenge for the scientist is to personally cultivate a loving knowledge of nature and to speak about nature in a way that promotes its celebration and protection (**Figure 6**).

Professional Decisions

If conservation biologists are to help successfully preserve biodiversity, they must take on a variety of active roles in addition to their scientific pursuits. For a start, they must become more effective educators in public forums. Conservation biologists often teach college students and write technical papers addressing environmental issues, but they need to reach a wider range of people by also speaking in villages, parks, elementary and secondary schools, and neighborhood gatherings. They must spend more time engaging all forms of mass media: writing articles and editorials for newspapers and magazines, blogging on popular websites, and speaking on radio and television. Because tenure, promotion, and professional prestige often accrue for technical work and peer-reviewed science but not for more practical or popularizing efforts, conservation biologists may face difficult choices when deciding whether to devote time to the latter.

Similarly, research projects describing habitat requirements or methods of transplanting an endangered species may have important conservation implications, but little potential to make basic advances in ecological or biological theory. Again, hard choices between professional glory and conservation usefulness may be necessary. As conservation biology becomes more established as a discipline, the system of rewards in place for practical conservation work will hopefully improve. Tenure and promotion committees at universities should acknowledge the value of community outreach and research that addresses local conservation issues.

Conservation biologists must also become politically active. Involvement in the political process allows conservation biologists to help pass new laws to preserve biological diversity, or argue against harmful legislation. Though politics may be time-consuming and tedious, it is the only way to accomplish major conservation goals, such as creating new parks and reserves (Lear, 1997). Conservation biologists also need to master the language and methods of the legal process and form effective alliances with environmental lawyers, citizen groups, and politicians. With their detailed knowledge of specific organisms and ecosystems, biologists are well placed to alert the general public to threats to biodiversity. They may also present this natural heritage in an appealing and inspiring way that leads to its preservation.

Such outreach and activism take time away from the pursuit of pure science. But many believe that at this point in human history, we no longer have the luxury to pursue knowledge for its own sake. To learn about biological diversity today is to learn about the threats facing it (Wilson, 2003). These threats demand an active response.

Fortunately, there is room for many types of work within conservation biology. Field and laboratory biologists who perform detailed natural history studies and genetic analyses of endangered species are necessary; so are hands-on managers who put together species recovery programs and coordinate interagency partnerships to facilitate them; so are teachers and writers who popularize wild nature; and so are political firebrands who alert the public to biodiversity losses and corporate environmental crimes.

Personal Choices

What about conservation biologists' private lives? Should a conservation biologist own a sports utility vehicle, for example? It gets terrible gas mileage, yet it might greatly facilitate her field work. Should a conservation biologist take a spectacular eco-tourist vacation to Amazonia? The trip uses prodigious resources, yet he might gather information and take pictures to educate audiences back home about important environmental issues. Then again, such reasoning may hide our real motivations: to drive in maximum comfort, to travel and enjoy ourselves. There is nothing wrong with comfort and enjoyment per se, but such high-consumption activities take a great toll on world biodiversity. Perhaps a conservation biologist should set a more restrained example.

Integrity is a key human virtue. It means living in conformity with our values and resolutely striving toward worthwhile goals. As spokespeople for biodiversity protection, conservation biologists will make more converts if they practice conservation in their own lives. Above all, this means limiting their personal consumption of resources through owning fewer things, recycling, avoiding unnecessary air travel, and taking other appropriate actions (Alexander, 2009). However, it is not necessary to put on a hair shirt or engage in

a quest for absolute purity. Showing that a life devoted to conservation can be stimulating and enjoyable is itself an important conservation message.

Should a conservation biologist dive into political debates? Give up time in the lab to meet with a group of concerned citizens? Drive a car to work? Eat meat? Have more than one child? These difficult professional and personal questions must be faced by conservation biologists and other committed conservationists. Responses may legitimately differ, but some are much better than others; indeed, some are inspiring (Wilson, 1994).

Ultimately, the fate of world biodiversity depends on two kinds of efforts. First, direct work such as habitat preservation, species reintroductions, ecosystem restoration, environmental litigation on behalf of other species, and the like. Second, preserving wild nature depends on creating societies that limit human numbers, consumption and economic growth, in order to enable us to set aside room and resources for other species. The creation of such societies is the greatest challenge facing biodiversity proponents today.

See also: Biodiversity as a Commodity. Conservation Biology, Discipline of. Conservation Efforts, Contemporary. Environmental Ethics. Human Impacts on Ecosystems: An Overview. Property Rights and Biodiversity. Sustainability and Biodiversity. Wildlife Management

References

Alexander S (ed.) (2009) *Voluntary Simplicity: The Poetic Alternative to Consumer Culture*. Whanganui, NZ: Stead & Daughters.

Carson R (1962) *Silent Spring*. Boston: Houghton Mifflin.

Costanza R, Norton B, and Haskell B (eds.) (1992) *Ecosystem Health: New Goals for Environmental Management*. Washington, DC: Island Press.

Daly H (1996) *Beyond Growth: The Economics of Sustainable Development*. Boston: Beacon Press.

Daly H and Farley J (2011) *Ecological Economics: Principles and Applications*, 2nd edn. Washington, DC: Island Press.

Foreman D (2004) *Rewilding North America: A Vision for Conservation in the 21st Century*. Washington, DC: Island Press.

Karr J and Dudley D (1981) Ecological perspective on water quality goals. *Environmental Management* 5: 55–68.

Lear L (1997) *Rachel Carson: Witness for Nature*. New York: Henry Holt.

Munda G (1997) Environmental economics, ecological economics, and the concept of sustainable development. *Environmental Values* 6: 213–233.

Nelson M and Callicott JB (eds.) (2008) *The Wilderness Debate Rages on: Continuing the Great New Wilderness Debate*. Athens: University of Georgia Press.

Pimentel D, Westra L, and Noss R (eds.) (2000) Ecological integrity: Integrating environment. *Conservation and Health*. Washington, DC: Island Press.

Primack R (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates.

Rolston H III (1994) *Conserving Natural Value*. New York: Columbia University Press.

Schmidtz D and Willott E (2011) *Environmental Ethics: What Really Matters, What Really Works*, 2nd edn. New York: Oxford University Press.

Speth G (2009) *The Bridge at the Edge of the World: Capitalism, the Environment, and Crossing from Crisis to Sustainability*. New Haven: Yale University Press.

Westra L, Bosselmann K, and Westra R (eds.) (2008) *Reconciling Human Existence and Ecological Integrity*. UK: Earthscan.

Westra L and Lemons J (eds.) (1995) *Perspectives on Ecological Integrity*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Wilson EO (1994) *Naturalist*. Washington, DC: Island Press.

Wilson EO (2003) *The Future of Life*. New York: Vintage Press.

Ethnobiology and Ethnoecology

Gary J Martin, The Global Diversity Foundation, Marrakech, Morocco

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 609–621, © 2001, Elsevier Inc.

Glossary

Analysis, emic and etic Concepts derived from the linguistic terms "phonetics" (representing speech sounds by precise and unique symbols and by technical descriptions of articulation, as practiced by trained linguists) and "phonemics" (characterization of speech through a minimal number of symbols, typically recognized by the speakers of a language). By extension, etic refers to the external explanation of cultural knowledge and practice (such as the use of Linnean taxonomy or scientific nomenclature to describe local useful plants), whereas emic denotes the internal perspective of local people (e.g., ethnobiological categories and nomenclature).

Economic botany As originally conceived, a branch of applied botany that arose during the colonial period to identify and characterize economically important plants and the products derived from them. Currently, it is a scientific endeavor that seeks to document the properties of useful plants through agronomic, archaeological, ecological, ethnobotanical, genetic, historical, phytochemical, and other empirical approaches. It overlaps broadly with ethnobiology because both fields have witnessed a similar development in theory and methodology in recent years.

Ecosystems, anthropogenic and natural Communities of organisms and their environment formed either through human action or through natural processes. In practice, it is difficult to establish the extent to which an ecosystem is anthropogenic or natural, reflecting the current and historical impact of people on the environment.

Ethnobiology A term coined in 1935 that has been defined as the study of the reciprocal interactions between people and the biological organisms in their local environment and, recently, as the study of biological sciences as practiced in the present and the past by local people throughout the world. Many researchers consider that ethnobiology comprises numerous subfields, such as ethnobotany, ethnoecology, ethnoscience, and ethnozoology, but there is no consensus on this point.

Ethnobotany and ethnozoology Approaches to studying the reciprocal interactions between people and the plants and animals in their local environment. This definition has been criticized as broad and openended, but it captures the common goals of analyzing traditional biological knowledge and assessing human impact on the environment. These approaches include subfields such as paleoethnobotany and paleoethnozoology, which evaluate archeological evidence on the past interactions between people, plants, and animals.

Ethnoecology Typically defined as the study of local knowledge and management of ecological interactions. Recently, some researchers have proposed an alternate definition, considering ethnoecology as an emerging field that focuses on local peoples' perception and management of complex and coevolved relationships between the cultural, ecological, and economic components of anthropogenic and natural ecosystems. It is concerned with the interaction between knowledge, practice, and production, and it is oriented toward applied research on conservation and community development.

Ethnoscience Arose as a minor subfield of ethnography concerned with recording in great detail local peoples' knowledge of biological organisms and the physical environment. Later, the term came to be used in a more restricted sense by cognitive and linguistic anthropologists to refer to local classificatory systems (as an object of study) and their semantic analysis (as a methodological approach). In France, the term is used to refer to ethnobiological studies in general.

Indigenous, local, and traditional Adjectives used by anthropologists, ethnobiologists, and other academics to describe people, practices, and knowledge. Indigenous denotes people (and their cultural practices and knowledge) who claim to be the original or long-term inhabitants of a particular place, in contrast to more recent colonizers. Traditional refers to established lifestyles, practices, and beliefs that guide cultural continuity and innovation – a definition that recognizes that traditions are always in a process of adaptation and change. Local, preferred by many researchers because it is the broadest and least value-laden term, indicates cultures that are found in a specific part of the world. It is commonly used to refer to people, whether long-term residents or recent arrivals (rural or urban), who make a living from the land and are knowledgeable about the biological resources in their environment.

Although the 100th anniversary of ethnobotany (coined in 1896), the golden anniversary of ethnobiology (first used in 1935), and the silver anniversary of ethnoecology (appearing in 1954) have passed, there is no consensus on the precise definition of these fields. This is explained in part because of their relatively recent origin and the current surge in their theoretical, conceptual, and methodological refinement. Disagreement over definitions is typical of multidisciplinary

fields; in the words of Brent Berlin, ethnobiology "combines the intuitions, skills, and biases of both the anthropologist and the biologist, often in quite unequal mixtures."

In one sense, ethnobotany, ethnobiology, and ethnoecology are new terms for old practices. People have been exploring the usefulness of diverse plants, animals, and ecosystems since the dawn of humanity. Documentation of local people's perception of the environment emerged slowly over thousands of years as scholars from many cultural traditions recorded local ways of classifying and using plants and animals. The onset of European colonization of Africa, Asia, the Pacific, and the New World gave added impetus to the study of local knowledge of tropical and temperate organisms and ecosystems.

Toward the end of the nineteenth century, academics began to use the prefix *ethno-* to refer to the way that local people view the natural world, in contrast to the perspective of natural scientists trained in universities. They coined terms such as "ethnobotany" (first used in print by Harshberger in 1896) and "ethnozoology" to describe these emerging fields of study that crossed the boundaries of natural and social sciences. Interest in traditional environmental knowledge continued apace in the early twentieth century, and in 1935 Castetter coined the term "ethnobiology," setting as its agenda the systematic analysis of data collected by ethnobotanists and ethnozoologists to achieve a deeper understanding of local peoples' knowledge and lifestyles. Economic botany gained importance as a parallel field focused on useful plants and the products derived from them. In 1954, Harold Conklin proposed the term "ethnoecology," originally conceived as a holistic and integrated approach to understanding local ecological knowledge and practice on their own terms, even while drawing on the concepts and methods of diverse scientific disciplines. A focus on classificatory systems and the linguistic and anthropological methods used to analyze them gave high visibility to an approach called ethnoscience.

In the 1980s and 1990s, further development of these various lines of research gave rise to new definitions, innovative theoretical orientations, and sophisticated qualitative and quantitative methodological approaches applied to local knowledge of the environment. In addition, ethnobiology expanded beyond its original geographical borders as the field gained importance in countries such as China, India, and Mexico. There is currently a new synthesis emerging – as yet without consensus – that defines ethnobiology as the study of biological sciences as practiced by local people throughout the world, comprising both empirical knowledge *(savoir)* and technical know-how *(savoir-faire)*, and inclusive of subfields such as economic botany, ethnobotany, ethnoecology, and ethnozoology.

Historical Development of Ethnobiology

Despite the insights provided by archaeological and historical linguistic studies, setting an even approximate date of the emergence of local biological knowledge is a matter of opinion. A detailed understanding of the natural world was key to the independent emergence of plant and animal domestication over a period ranging from 8500 B.C. in southwest Asia to 2500 B.C. in the eastern United States. However, environmental knowledge reaches even further back into history, when hunting and gathering dominated subsistence activities.

Some researchers would place the beginning of human ecological knowledge at the dawn of humanity, approximately 7 million years ago. Early human ancestors, who lived on the African continent 2.5 million years ago, apparently fashioned stone tools for harvesting and processing food, probably allowing them to adapt to new environmental conditions. It is widely assumed that humans have been observing natural phenomena, distinguishing between biological organisms and discovering their uses ever since the emergence of early *Homo sapiens* approximately 500,000 years ago. The archeological record reveals that by 50,000 years ago, Cro-Magnons had developed technologies for construction, fishing, gathering, and hunting that were dependent on a detailed understanding of plants, animals, and other elements of the natural environment.

Just as no one knows exactly when ecological knowledge appeared on the cultural landscape, there is no clue when the original precursors of ethnobiologists came on the scene. The first critical observations of other peoples' ways of perceiving nature are probably as old as culture contact itself. Because these observations went unrecorded, the origin of the study of traditional biological knowledge is lost in history.

In the absence of other evidence, the historical roots of ethnobiology can be seen emerging over a period of several thousand years, when students of natural history from Greek, Roman, Egyptian, Chinese, Indian, Arabic, Native American, European, and other cultures began to record popular beliefs in scholarly texts. Original studies focused on medicinal botany, agriculture, and horticulture – activities that drew heavily on the knowledge of local people.

An overview of these early texts and later works reveals that ethnobiology and natural history have evolved – much like the biological species and ecosystems that are their focus – through a process of punctuated equilibrium. Certain historical periods are marked by an intensive effort to expand empirical knowledge of natural phenomena, often by incorporating local lore, whereas other epochs are characterized by an unquestioned acceptance of published works.

Early Scholars in Europe

Academics often trace their intellectual history to the era of Greek philosophers, who lived more than 2300 years before our time. This is when classical botany and zoology were brought into existence by scholars such as Aristotle, who sought to summarize all current knowledge about plants and animals in encyclopedic works. It was as part of this endeavor that scholars in Europe first made a systematic study of what local people knew about the environment. In part, Aristotle and other early naturalists such as Theophrastus – who, as author of *De Historia Plantarum* and other works, is considered the father of botany – rejected many local supernatural beliefs in their quest to understand the natural world. Simultaneously, they drew on commonsense explanations and empirical knowledge of local people when describing the classification and use of plants and animals.

Among the people who followed in the footsteps of these early naturalists is Dioscorides, a military physician born in Asia Minor in the first century A.D. He wrote *De Materia Medica*, a treatise on medicinal plants which was the standard reference of botanists, medical doctors, and other scholars in Europe for 1500 years. Apart from drawing on previous herbals, Dioscorides learned much about herbal remedies by interacting with local people he encountered during his wide-ranging travels with the Roman army in the Mediterranean region. Pliny the Elder, a Roman scholar who was one of Dioscorides' contemporaries, recorded extensive plant lore in his 37-volume encyclopedia called *Historiarum mundi* or *Natural History*. He devoted 9 volumes to medicinal plants, making frequent reference to traditional practices and knowledge.

The Doldrums of the Middle Ages

The documentation of local knowledge that marked the origin of biological thought in Greek and Roman Antiquity was much less evident in the Middle Ages in Europe. The decline of the Roman Empire virtually halted scholarly research on natural history and resulted in the destruction of much existing literature of the epoch. Throughout the Middle Ages, Europeans based their studies of medicinal plants almost entirely on the works of Theophrastus, Pliny, Dioscorides, and other early naturalists. Physicians from across the continent relied heavily on *De Materia Medica*, often trying unsuccessfully to match the local flora to the approximately 600 Mediterranean species described by Dioscorides instead of documenting the popular knowledge of their own region. As anthropologist Scott Atran (1990) summarized,

> After Aristotle, the practice of copying descriptions and illustrations of living kinds from previous sources superseded actual field experience in the schools of late antiquity. Well into the Renaissance, scholastic "naturalists" took it for granted that the local flora and fauna of northern and central Europe could be fully categorized under the Mediterranean plant and animal types found in ancient works. Herbals and bestiaries of the time were far removed from any empirical base.

The Golden Age of the Moors

Despite this stagnation on the part of European scholars, general knowledge of medicinal plants was enriched by the flow of information coming from the Arab world, particularly through Spain (Andalucia), Sicily, and North Africa. Although dedicated in part to translating the works of Aristotle, Theophrastus, Dioscorides, and other classical writers, Moorish and other medieval scholars in these regions pursued empirical research that they applied to practical ends in agriculture, astronomy, botany, mathematics, medicine, and other fields. Although there are records of Arabic writings on botany dating to the ninth century, it was particularly in the twelfth and thirteenth centuries that scholars became prolific in recording precise original observations on plant and animal biology, conducting experiments on agricultural crops, and attempting to classify plants systematically. One of the key early scholars was Abû-l-Kheyr Al-Ichbili, apparently the "anonymous botanist of Seville" of the latter half of the twelfth century who wrote the *'Umdat at-tabîb*, a botanical treatise that contains detailed descriptions of the anatomy, habitat, and local names of plants. Maimonides, a Jewish scholar resident in southern Spain during the twelfth century, contributed works such as *Book Explaining Medicinal Drugs* and *Treatise on Poisons*. The thirteenth century brought the works of Rachid-eddin Ibn Es-Sûri in the Machreq (eastern Arabic region), who is credited with the first herbal containing illustrations, which were made by a collaborating artist from both living and pressed plants. Ibn Es-Sûri traveled widely, describing formerly undocumented plants from Syria, Palestine, and Egypt that he discovered by interacting with local people of the region. His efforts were matched in the Maghreb (western Arabic region) by Abû-l-'Abbas En-Nabati – who documented local plants and their uses during his extensive travels in what is now Spain, North Africa, Syria, and Iraq – and later scholars such as his student Ibn Al-Baytar, author of the *Treatise on Simples* or *Jami' al-mufradat*, which contains information on the synonymy (including Berber and other local names), description, properties, and uses of approximately 1400 species. Throughout this period, descriptions of specific plants were drawn from multicultural sources, including Berber plant knowledge and the traditional practices of Jewish pharmacists who lived throughout Europe and North Africa, passing their profession from one generation to the next. As the Moors were forced from Andalucia and other parts of Europe, Arab science fell into decline. Later botanical treatises, including the *Hadîqat al-azhâr* of the sixteenth century Moroccan scholar and medical doctor Al-Wazir Al-Ghassani, were largely based on the *'Umdat at-tabîb* and other early works.

Scholars from Other Cultural Traditions

A similar pattern of initial empirical discovery and cross-cultural learning mixed with centuries of uncritical acceptance of written works is evident in other cultural traditions. In China, the first scholarly studies of traditional biological knowledge are thought to date at least to the fifth century B.C., approximately 200 years before early Greek philosophers began recording their ideas about botany and medicine. During this epoch, the Chinese philosopher Confucius is said to have encouraged his students to study ancient knowledge, including traditional names of plants and animals.

In the first or second century A.D., about the time that Dioscorides was afoot in the Mediterranean, Chinese scholars published the first of many *materia medica*, referred to generally as *bencao* (from a combined term for tree and grass) in Chinese, which contain information on minerals, plants, and animals used traditionally in the treatment of illness. Medical doctors of the fifth century A.D. revised these early *bencao*, providing a major reclassification of the various types of natural medicines used at the time. The resulting pharmacopoeia went unchanged and unchallenged for many centuries, much as the classic work of Dioscorides in Europe. At the end of the sixteenth century, a Chinese doctor named Li Shizhen began to emulate the practice of empirical observation that Confucius had advocated many centuries before. His major work, the *Bencao Gangmu* or *Compendium of Materia Medica*, contains information on more than 10,000 herbal remedies that he

obtained not only by studying ancient texts but also by traveling to the countryside to talk with people.

Ayurveda, a system of medicine which putatively began in India during the sixth century B.C. and spread to Sri Lanka by the third century B.C. and into Tibet by the seventh century A.D., was partially based on traditional knowledge. According to folklore, shepherds and forest dwellers familiar with the types and properties of medicinal plants first discovered the remedies used in this oriental medical practice. Their knowledge was discussed in various literary religious works called *vedas* (from the Sanskrit word for knowledge), which were apparently written in India approximately 3200 years ago, after millennia of oral transmission. Ayurvedic scholars later compiled additional empirical observations in a series of books referred to as the *Nighants*, or Vedic glossaries. During the subsequent period of foreign domination and internal conflict that brought innovation and documentation of local knowledge to a standstill, these standard texts of ayurveda remained unchanged.

Other ancient written sources that document local biological knowledge were in part the product of culture contact and changes in political and economic dominance. In the New World, for example, the Aztecs broadened their own sophisticated knowledge of medicine and agriculture as they sought tribute and learned of new useful plants from the different Mesoamerican cultures they conquered. The Aztecs cultivated many newly discovered species in extensive highland botanical gardens tended by people from various geographical regions of Mesoamerica. The depth and richness of the preconquest indigenous knowledge of the natural world are demonstrated by scholarly works, including the *Badianus Manuscript*, an illustrated herbal written in 1552 by two Aztecs who had been educated by Catholic missionaries. One author, Martin de la Cruz, was an indigenous physician who had acquired his medical knowledge empirically. The Mayas and Incas had similar literate traditions and they doubtlessly recorded some aspects of the ecological knowledge of the various ethnic groups they dominated at the height of their political power and cultural development. Many of these New World written sources of local knowledge were victims of the conquest, destroyed by overzealous missionaries and conquerors who wished to impose European culture, languages, and religion on the people of the New World.

The Renaissance and Exploration

During the sixteenth and seventeenth centuries, Renaissance botanists began to emulate the methods that Dioscorides had applied approximately 1500 years previously, bringing an end to the intellectual stagnation that characterized the Middle Ages in Europe. They carefully observed plants in the field and inquired about their local names and uses in Germany, Holland, Italy, and other parts of Europe. This experience served them well when faced with the influx of exotic species from areas of the world discovered and colonized by Europeans during this period. The diversity of biological organisms discovered by explorers stimulated Linnaeus, Darwin, and other natural scientists to formulate many of the concepts that are the building blocks of modern-day systematics and evolutionary studies. Although Linnaeus left some notes and sketches on the use of plants by local people, his greatest contribution to the future field of ethnobiology was the incorporation of notions of folk biology and nomenclature, including the concept of morphological affinity as a criterion for defining taxa, in the scientific classification of plants.

This Renaissance was the golden age of the European herbals. The shift from manuscripts (produced by hand) to wood-cut and metal-engraved herbals published in large numbers allowed new botanical knowledge to be disseminated widely.

The quest to exploit local knowledge and economically important species which went along with colonization inspired adventurers, missionaries, and natural historians to record their observations on traditional biological knowledge in many parts of both the New World and the Old World. As ethnobotanist Richard Ford (1978) described,

> A rapid progression of expeditions came to North America to discover and to colonize, and the chronicles of adventure are a record of the utilitarian value of an unfamiliar landscape and the use the indigenous people made of it. Its economic potential certainly had priority to any interest in attitudes about the land. The observations… provided the first natural history of North America and the bases for the beginning of ethnobotany.

From the sixteenth century onwards, researchers began to focus increasing attention on the biological wealth of tropical countries and the benefits it promised for Europeans. To this end, scholars drew on the knowledge of local people, who continually experiment with cultivated and managed species in anthropogenic ecosystems and wild plants harvested in natural ecosystems. Scholars consulted both written sources, such as Ayurvedic works and Chinese pharmacopoeias, and oral history to produce extensive encyclopedias of useful plants from around the world, ranging from the *Coloquios dos simples e drogas da India*, written by the Portuguese explorer Garcia ab Orta in 1563, to the 12-volume *Hortus Malabaricus* of Van Rheede published in the late 1700s and A *Dictionary of Economic Products from the Malay Peninsula* produced in 2 volumes by Burkhill in 1935.

Another notable example is the herbal of Rumphius, a seventeenth-century natural historian from Germany who spent nearly 50 years in Asia working for the Dutch East-Indies Company. Increasingly released from his administrative duties but still under the employ of the company, he focused his attention on studying useful plants, animals, and minerals in various regions that today constitute Indonesia. He provided descriptions of more than 700 medicinal or toxic plants, published posthumously in the six volumes of the *Herbarium Amboinense*.

Creation of New Fields

Although Rumphius and his counterparts approached natural history as a holistic phenomenon, scholarly activity in later centuries began to reveal a fragmentation of research into distinct disciplines, marking the beginning of reductionism in the sciences. Subdivisions of science became more clearly defined, and professional practitioners began to specialize in

specific subfields. Theoretical science was increasingly applied to technological innovation, stimulating the growth of educational institutions and commercial enterprises.

For example, pharmacognosy (the study of naturally occurring compounds that can be used medicinally and in other ways) became recognized as a separate field of endeavor in the early nineteenth century. It focused on the identification, preparation, and commercialization of drugs, which mostly came from natural sources, especially plants. This enterprise, combined with tropical exploration, yielded such novel cures as quinine, which was successfully used to control malaria in Europe and other regions.

As research on traditional biological knowledge expanded in the late nineteenth and twentieth centuries, several lines of research became apparent, reflecting the diversity of researchers who began to appreciate the importance of the subject. The historical development of these various fields is intertwined, and the definition of each is in flux. There may be an emerging consensus – one supported in this article – that ethnobiology is the broadest field, comprising ethnobotany, ethnozoology, ethnoecology, and economic botany as subfields. Other colleagues would posit ethnoecology or even ethnoscience as the most inclusive term. In practice, this debate on terminology is less significant than the realization that these fields broadly overlap, and that the slight differences of opinion on definition are dwarfed by the general agreement on theory, concepts, and methods.

Ethnobotany and Related Fields

By the end of the nineteenth century, researchers began to recognize the study of traditional biological knowledge as a separate discipline. John W. Harshberger, a professor of biology, initiated the fashion of using the prefix *ethno* to indicate the study of local people's natural history. In 1896, he used the term ethnobotany in print, and it began to replace names such as "aboriginal botany" and "botanical ethnography" that had been used previously by other authors. In the words of Richard Ford (1978), after a "half century of scientific attention and an even longer history of casual observations," the study of other people's interaction with nature finally had a name and recognition as a distinct line of academic endeavor.

The emergence of ethnobotany, ethnozoology, and related fields coincided with important developments in the natural and social sciences toward the end of the nineteenth century. The diverse elements of natural history, including botany, zoology, pharmacognosy, and other fields, began to mature into distinct disciplines, each with separate methods and goals. Scott Atran (1990) characterized this as the "breakaway of science," a time when natural historians began to leave behind commonsense descriptions of natural phenomena – drawn in part from local peoples' perception and classification of nature – in order to embrace rigorous experimental methods. Social scientists began to focus on separate aspects of human society and culture, with the consequent emergence of fields such as anthropology, linguistics, and sociology.

Economic Botany

Nineteenth-century botanists, who focused primarily on the utility of plants and only secondarily on local culture, began to refer to their approach as economic botany. The goal of their research has been to document local uses of plants and to organize the resulting data according to the global system of plant classification. They have produced detailed works on plants employed by local people for food, medicine, textiles, utensils, and many other purposes (Schultes and Raffauf, 1990).

Research on the commercial value and utility of plants expanded as botanists from the United States and Europe explored the New and Old World tropics in search of products that would increase the wealth of developed countries and the well-being of people in general. Today, economic botanists continue to search for marketable products in tropical forests and elsewhere, but they are increasingly interested in how the commercialization of these resources can contribute to resolving the poverty, malnutrition, and diminished social status of local people as well as spurring economic development in developing countries. An increasingly important offshoot of this enterprise is bioprospecting (a term derived from "biodiversity prospecting"), the search for useful and novel products (including chemical constituents of medicinal value) from plants, animals, fungi, and other biological organisms. Other economic botanists are concerned with developing theoretical and methodological approaches to understanding the subsistence and commercial value of plant resources, now and in the past.

Ethnoscience

While botanists were establishing economic botany, anthropologists and other social scientists were developing a different perspective. In the tradition of ethnography developed by anthropologist Franz Boas, ethnoscience emerged as a minor subfield dedicated to recording in minute detail local peoples' knowledge of biological organisms and the physical environment. The subfield underwent a further transition in the 1950s and 1960s, when cognitive and linguistic anthropologists began to focus on the empirical categories, social rules, symbolic systems, and modes of behavior that reflect how local people perceive the natural world. These early anthropological studies formed the foundations for a new ethnoscientific approach that advocated rigorous analyses of ethnobiological knowledge, with particular emphasis on systems of ethnobiological classification. Thus, whereas economic botany emerged as a utilitarian practice firmly rooted in commerce and development (and later developed theoretical frameworks), ethnoscience arose as an intellectual endeavor oriented toward a deeper understanding of human culture and cognition.

Ethnoecology

Even though he is most associated with the development of the ethnoscientific approach, Conklin is credited with coining the term ethnoecology in 1954. Given the precedent set by terms such as ethnobotany and ethnozoology, it would be natural to assume that ethnoecology would refer to the study of local perceptions of ecological processes, such as nutrient cycling, vegetational succession, or the interactions between plants and animals. An increasing number of researchers propose a different definition, using the term to refer to local peoples' perception and management of the complex and

coevolved relationships between cultural, ecological, and economic components of anthropogenic and natural ecosystems. This emerging subfield, much as the broader field of ethnobiology, is concerned with the interaction between knowledge, practice, and production and is oriented toward applied research on conservation and community development. Mexican ecologist Victor Toledo stated that the aim of ethnoecology should be the ecological evaluation of the intellectual and practical activities that people carry out during their appropriation of natural resources.

Although the definition of ethnobiology includes a reference to knowledge and know-how (*savoir* and *savoir-faire*), for ethnoecologists the distinction is between an ethnobiological *corpus*, local peoples' repertories of concepts, perceptions, and symbolic representations of nature, and *praxis*, the art, science, and skill of appropriating nature and biological resources. The interrelationship between knowledge and practice is manifested in production, as people apply their intellectual understanding of nature to the everyday tasks of farming, gathering, and hunting for subsistence and commercial purposes. In order to understand these complex interactions, ethnoecologists seek to elucidate how the management of anthropogenic and natural ecosystems – and the biological organisms they harbor – has arisen through a process of coevolution between the environment, knowledge, technology, social organization, and values of local peoples.

Although attractive conceptually, the development of this conception of ethnoecology has been limited by the lack of a unifying theoretical framework and a practical methodology. This distinguishes it from ethnobiology, which is developing a central organizing theory, an orientation toward hypothesis testing, and an increasingly elaborate set of qualitative and quantitative methods, drawn in part from ethnobotany, ethnozoology, and economic botany.

Ethnobiology

Clément (1998) proposes that the starting point for ethnobiology – as the field which integrates related approaches such as economic botany, ethnobotany, ethnozoology, and ethnoscience – is the 1860s, when the first designations for the field began to be used by American and European scientists. In a historical sketch that spans a period of more than 130 years, he discusses the origins, key theories, and methodological approaches of the main trends of ethnobiology. Although no such historical framework of a scientific discipline is without controversy and potential modification, Clément's synopsis is a serious effort to provide a detailed historical analysis of ethnobiology.

Clément divides the development of the discipline into three eras and seven periods (**Table 1**). The preclassical period, from 1860 to 1953, is dedicated to gathering empirical data on the uses of plants and animals from an etic perspective and to the first syntheses that begin to define the scope of the discipline. During the classical period (1954–1980), there is a shift to studies carried out from an emic perspective and a particular focus on ethnobiological classification. An increase in collaborative work between academic specialists and local people and the formation of professional associations of ethnobiology characterize the postclassical period, from 1981 to the present. Later in the period, there is an increased focus on the appropriation and management of plant and animal resources and a concern for application of research results to the resolution of environmental and social problems. This historical review provides an appropriate starting point for considering the current trends in basic and applied ethnobiological research.

Current Trends in Ethnobiology

Recent studies in ethnobiology can be classified in three general, interrelated areas. Documentation and analysis of uses of plants and animals is the oldest aspect of ethnobiology. A focus on knowledge, which became especially popular in the 1950s and 1960s, characterizes the study of perception and classification of the natural world. In the 1980s and 1990s, attention shifted to local management of biological resources and the environment, often drawing on ecological concepts and methods. These areas are combined in various degrees in the following major trends in ethnobiological research.

Cognitive Mapping

Ethnobiologists have documented local peoples' spatial conceptualization, including their ability to locate biological resources, discern landscape features, and identify different types of vegetation. This geographical literacy is linked to an aptitude for assessing the potential productivity of the environment, often through the recognition of plant species that indicate fertility or sterility of soils. This expertise derives in part from local peoples' sophisticated perception of how the various elements of the ecosystem (organisms, soils, climate, topography, etc.) form an interdependent whole. In addition, it is derived from their classification of key geographical landmarks that are labeled by specific toponyms, or geographical place names. Local views of the landscape are being integrated into geographical information systems, which in turn are useful in creating biological resource maps and management plans for specific areas. When united with ethnobiological inventories and studies of local categorization of ecological succession, community mapping allows researchers to test hypotheses on whether the highest number of useful plants and animals come from primary or secondary forest, or anthropogenic versus natural ecosystems.

Local peoples' mastery in selecting plant populations that yield the best fruit, the most potent medicine, or the best materials for construction is attributed in part to their grasp of the landscape. These mental maps also explain their adeptness at selecting the best place to cultivate the earth, create human settlements, or leave natural areas that maintain soil fertility, water purity, and other environmental benefits. Researchers have demonstrated that these skills can degenerate when local people find themselves pressured by economic needs to overexploit the resources available to them or are forced onto lands for which they are maladapted culturally.

The ability of some local people to integrate and recall complex information on the local environment often gives them special proficiency as stewards of community reserves and other protected areas. Their ability to assess the quality of

Table 1 Features of Various Historical Periods and Stages in Ethnobiology

Period	*Stage*	*Dates*	*Features*
Preclassic (1860–1954)	Economic uses	1860–1899	Studies of biological resources and their utility carried out by researchers affiliated with major museums and universities; general lack of appreciation of the sophistication of local knowledge and subsistence systems from an emic perspective
	Information gathering	1900–1931	Greater empirical depth in research but continued emphasis on economic uses of plants and animals; better appreciation of complexity of local knowledge and use of plants and animals, especially as reflected in systematic attempts to record local terminology, myths and beliefs, and knowledge of anatomy and behavior; emergence of comparative studies and standard methods
	The first syntheses	1932–1953	Emergence of ethnobiology as a distinct field of enquiry and appearance of the first syntheses that delimit its scope; increasing distinction between economic botany and ethnobotany, with the latter emphasizing the systematic documentation of local knowledge and management of plants; continued lack of recognition of scientific aspects of traditional biological knowledge
Classic period (1954–1980)	Emic knowledge	1954–1968	Emergence of ethnoscience, leading to a focus on the organization of knowledge systems from the local perspective, with insights from linguistics and empirical anthropological methods; relegation of the study of plant and animal resources to secondary importance; beginning of interest in ethnobiological classification and appreciation of the scientific basis of traditional knowledge
	Classification	1969–1980	Focus on ethnobiological classification, including principles of categorization and nomenclature, and the analysis of correspondence between scientific and local classifications; accumulation of evidence for the scientific basis of local biological knowledge; increasing interest in ethnobiology beyond the United States and Europe, especially in Latin America and the Pacific
Postclassic (1981 to present)	Associations	1981–1992	Production of major empirical works based on close collaboration between academic and local researchers; development of theoretical approaches beyond classification, including gender relations in resource use, cultural significance of plants, and historical reconstruction of ethnobiological knowledge systems; emergence of academic societies and specialized journals of ethnobiology, especially in developing countries
	Resource management	1993 to present	Publication of standard methods manuals, quantitative techniques, and innovative empirical studies; emergence of concern about applying ethnobiology to conservation and development; renewed interest in economic botany, including nutritional and medicinal benefits of plants, but incorporating novel theoretical and methodological approaches and informed participation by local people.

Source: Adapted from Clément D (1998) L'Ethnobiologie/ethnobiology. *Anthropologica* 40: 7–34.

useful biological resources and to locate areas where these organisms are found in greatest density makes them indispensable members of research teams that seek to identify priority areas for conservation and management. Foresters and conservation biologists draw on this expertise when deciding how to zone natural areas according to various land-use options. Local peoples' classification of land units also plays a key role in justification for claims of ancestral domain and other forms of ownership of the lands they have long occupied.

Resource Management and Valuation

In order to ensure subsistence production and to earn a living, local people draw on their detailed ecological knowledge to manage the diverse microenvironments and biological organisms in their communities. Recent studies have shown that the process of plant domestication through selection of preferred varieties occurs not only in cultivated fields but also in other parts of the anthropogenic landscape. From the cultivation of these domesticated plants to the harvesting of wild useful species, local people engage in many ecological practices that are often energy efficient and sustainable, at least under traditional conditions. They capitalize on the consumptive use value of natural resources when they harvest plants and hunt animals for subsistence purposes and on the productive use value when they barter or sell agricultural and forest products. They also benefit, as does the whole world, from the nonconsumptive value that comes from ensuring the viability of ecosystem function in general, including watersheds, nutrient cycles, climate, soils, and other elements.

By investigating the link between local knowledge, practices, and production, ethnobiologists assess the value of anthropogenic and natural ecosystems and the rationality of resource harvesting decisions made by small-scale farmers and gatherers of forest products. Much of this research is carried out in 1-ha plots, agricultural fields, home gardens, and other measured study sites, resulting in quantitative assessments of sustainability and value. These methods allow researchers to evaluate the hidden costs of tropical forest destruction, the economic benefits derived from both subsistence and commercial use of wild species, and the environmental advantages of maintaining forest cover as a way of buffering local climate and preserving the purity of local air, water, and soil. Through these perspectives, conservation biologists,

development specialists, and communities monitor the sustainability of current productive practices, propose new methods of managing fields and forests, and select new biological species that can be domesticated, cultivated, or gathered locally. Studies of local systems of resource management can also enrich the work of ecologists, who seek to restore the diversity and value of forest ecosystems damaged by mismanagement or natural catastrophes. In addition, evidence of long-term management of biological organisms and ecosystems reinforces local peoples' claims for traditional resource rights, including just compensation when novel biological resources are commercialized.

Scientific Covalidation

A key activity of ethnobiologists is to understand the rationale behind the way local people interact with the natural environment. In research laboratories, scientists carry out a broad array of analyses that seek to corroborate the efficacy of local uses of plants and animals, ranging from the identification of active compounds in medicinal plants to appraisal of the tensile strength of natural fibers and assessment of the nutrient content of wild foods. In the field, ecologists assess how the yields of agroecosystems compare with those achieved by large-scale monocultures and to what extent traditional methods of wildlands management are sustainable. Ethnobiologists compare ethnobiological categories with scientific taxa, judging the extent to which local biological classifications correspond to biosystematics. This range of activities is aimed at revealing the logic, from a scientific perspective, of the thought and practices of local people. Because much is understood about the classification of specific cultural domains, such as plants, animals, soil, climatic zones, and vegetation types, there is a call for ethnobiologists to study local peoples' perceptions of ecological interactions, the reciprocal relationships of various elements of the ecosystem.

Scientific covalidation allows researchers to understand how to optimize the value – and also to ensure the safety and efficacy – of plants and animals that are consumed or commercialized locally. In addition, covalidation provides insights on how traditional knowledge can be incorporated in the management of protected areas. These studies reveal the breadth of local people's ecological knowledge and the wealth of resources available in natural areas, highlighting their potential value on the world market and their contribution to local subsistence. In addition, this research plays an important role in convincing protected area managers of the value of including local people in conservation and development projects.

Ethnobiological Classification

Continuing a trend initiated by ethnoscientists in the 1950s, ethnobiologists are documenting how local people classify diverse elements of the natural environment. Based on fieldwork in diverse cultures, they describe complex interrelated sets of categories for plants, animals, soils, climates, vegetation, illnesses, food, and other cultural domains and natural phenomena.

Much attention has focused on describing universal similarities in the ways in which local people perceive the natural world. Many generalizations on the categorization, naming, and identification of plants and animals are now widely accepted, whereas others continue to provoke controversy, especially among anthropologists who place emphasis on cultural relativity or the uniqueness of each ethnic group. In particular, there is disagreement over why people are motivated to classify various elements of the environment. Some researchers seek a utilitarian explanation, suggesting that people enhance their ability to fulfill their basic subsistence needs by naming and classifying useful plants, animals, soils, and other natural features. Those who follow an intellectualist line of reasoning argue that there is a universal human tendency to categorize plants and animals according to their overall appearance or symbolic role. These researchers note that local people tend to group organisms with a similar morphology or behavior, regardless of their cultural utility. In a similar vein, some colleagues propose an ecological rationale, noting that some aspects of ethnobiological classification can be elucidated by reference to the role and interaction of plants and animals in diverse ecosystems. Although there is an emerging consensus that classification is motivated by a combination of these factors, there continues to be much debate about which is the most important dimension.

Studies of ethnobiological classification contribute to understanding – in part by reference to the global system of plant systematics – how local people perceive and manage natural resources. When carried out in a participatory way, ethnobiological inventories provide an opportunity for local people and researchers to work together to document the distribution, management, and use of biological resources locally and globally. They produce the baseline data needed to produce bilingual and bicultural manuals that compare and contrast different ways of classifying, managing, and using biological organisms. These ethnobiological manuals contribute to applied programs of conservation and development by highlighting culturally significant species that local people are harvesting from the wild, managing in anthropogenic landscapes, or cultivating in gardens and fields. Often, these species are selected for use in initiatives that promote reforestation, sustainable harvesting of minor forest products, or cultivation of useful plants that contribute to the well-being and income of local people.

Knowledge Variation

There are significant differences in the way local people perceive and use biological resources. Ethnobiological studies reveal that some plants and animals are known by a majority of the population, but others are the domain of curers, the elderly, women, or members of another social group. There are significant differences in the biological knowledge of people living in separate communities, belonging to distinct ethnic groups, speaking different languages, or subsisting in diverse ecological zones. These differences can be explained in part by the fact that each person's knowledge is correlated with sociological characteristics such as age, gender, occupation, education, social status, and zone of residence. Perception and

management of plants and animals are also affected by each individual's life experiences – for example, if he or she has suffered from a major illness, migrated to other communities, or worked as an apprentice to a plant specialist.

Awareness of these differences is essential when assessing the depth and breadth of traditional knowledge and the diversity of ecological practices in a particular area. It is now recognized that researchers run the risk of obtaining a biased perspective of local biological knowledge if they work with few informants. In contrast, consultation with a representative cross-section of local people can provide a relatively accurate portrayal of perception of the natural world. Ethnobiologists are analyzing patterns of agreement among different individuals, producing a quantitative method of identifying the most culturally significant plant and animal species, soil types, forest zones, and other elements of the landscape. These results ensure accuracy when preparing natural resource management plans and popular manuals produced with local communities. They can also play a role in understanding the knowledge held by specialist user groups and how they are transmitting it from one generation to the next.

Mechanisms of Change

Another principle that has emerged in recent studies is that knowledge about the natural environment and ways of managing biological resources are not static. They change as people move from one region to another, as youth reinterpret what they have learned from elders, and as cultures come into contact with each other. Ethnobiologists are analyzing the dynamics of these cultural changes and assessing how they can enrich or impoverish local knowledge systems.

Archaeology, linguistic reconstruction, archival research, and oral history open a window to the past, allowing researchers to discover how ecological knowledge and resource management have evolved during decades and centuries of political, demographic, and economic change. Studies of historical ecology are revealing which species, ethnobiological categories, and ecological practices have persisted, changed, or disappeared over time, including those that have become widespread through borrowing across cultural boundaries. Research that focuses on world events in recent centuries allows historians to assess the impact of colonization and other forms of culture contact on local ways of perceiving and managing the natural world.

Many regions of the world are currently in an intense period of change characterized by not only the destruction of wildlands and loss of biological diversity but also the transformation of traditional biological knowledge. In some regions, young people are not learning what their elders know about the environment, particularly as traditional ways of using biological organisms and managing natural areas fade away. Specialized knowledge held by only a few curers or spiritual leaders is lost when no apprentice is found to carry on traditional medical or religious practices. This loss of knowledge is often linked to increasing contact with national and international cultures and is often exacerbated as local people incorporate formal education, major religions, and migration into their lifestyles.

Ethnobiologists are playing a practical role in assessing the extent of cultural transformation and searching for ways of promoting the survival of local ways of classifying, using, and managing natural resources. These actions are based on the assumption that participation in the joint management of ecosystems and resources stimulates local people to retain and build on empirical knowledge and practices acquired during the tens, hundreds, or thousands of years that they and their ancestors have resided in the region. Studying the evolution of ethnobiological knowledge elucidates mechanisms of cultural resistance, allowing us to understand why local ecological knowledge persists and how we can reinforce the mechanisms of resistance. These studies also highlight the impact that local people have had on the natural environment over time, often providing evidence that they are responsible for maintaining biological diversity and stimulating innovative ideas for conservation in the future.

Ritual, Religion, and Symbolism

Plants and animals play an important role in rituals and spiritual practices and are a recurring element in myths, legends, and stories. Some researchers assert that local people, because of their cosmology or understanding of the universe, relate to the natural environment in ways fundamentally different from that found in, for example, European or American culture. A central element in this conception is that traditional lifestyles link people to nature in a way that provides them with a special understanding of nature. A common way for outsiders to gain access to the world of ritual, magic, and religion of local peoples is through apprenticeship with spiritual leaders, often accompanied by the use of psychoactive plants.

Whether or not these generalized notions of differences between global knowledge systems and traditional knowledge are valid in all cases, it is undeniable that spiritual beliefs about the forest and associated taboos on the use of natural resources are important elements of conservation and sustainable use of natural resources in many cultures. When people begin to abandon these beliefs, traditional controls on resource exploitation often disappear, potentially leading to devastation of formerly protected areas such as sacred groves. Ethnobiologists are increasingly aware of the need to record these belief systems and to verify empirically what impact they have on conservation and use of natural resources. In areas in which traditional cosmology is still a viable element of local culture, conservation biologists are exploring ways of integrating it with the management of protected areas.

The Internationalization of Ethnobiology

Although ethnobiology as an academic discipline originated in Europe and the United States, it has now been embraced by researchers in many developing countries who have subsequently adapted the techniques and concepts to their own goals and local conditions. The emergence of professional societies of ethnobotanists in developing countries, ranging from the Indian Society of Ethnobotanists in 1980 to the Asociación Mexicana de Etnobiología in 1993 and the

Sociedade Brasileira de Etnobiologia e Etnoecologia in 1997, is evidence of this trend.

The internationalization of ethnobiological research and training has resulted in new directions in theory and application, enriching the field. In India, the tradition of conducting ethnobotanical inventories in various tribal areas has continued, but it is now supplemented by innovative approaches to studying the harvest of nontimber forest products in joint forest management schemes and practical strategies to create community biodiversity registers. Researchers in China have contributed studies on ecological succession in swidden fields, marketing of useful plants, and analysis of agroforestry practices. Equally impressive are developments in Mexico, where ethnobiologists have focused on the management of anthropogenic and natural ecosystems as well as the process of domestication of botanical resources.

The New Synthesis

The proliferation of labels and orientations for the study of local biological knowledge and practice is likely to continue. Consensus on precise definitions will be difficult to achieve because researchers are approaching the field from a variety of academic disciplines, bringing with them a wealth of new concepts and methods. Despite this dynamic development, there is some semblance of an agreement on a typology for the discipline.

Distinct approaches to gathering empirical data on the reciprocal interactions between people and biological organisms will continue to be referred to by terms such as economic botany, ethnobotany, ethnoecology, and ethnozoology. Ethnobiology is becoming the preferred term for an integrative discipline that draws on all these approaches to analyze traditional biological knowledge and practices throughout the world. It is unified by a central theory that local peoples' systematic knowledge and management of biological organisms and ecosystems can be classified as biological sciences, covalidated by qualitative and quantitative research methods. Perceived in this way, ethnobiology blends conventional studies carried out by economic botanists, ethnobotanists, ethnozoologists, and ethnoscientists that present a limited vision of local people's interaction with the natural environment. This provides an opportunity for reintegration of various disciplines of natural science, counteracting the impact of the reductionism of global scientific knowledge.

Ethnobiology seeks not only to integrate these various lines of scientific research but also to focus them on supporting community development and biodiversity conservation, which are clearly multidisciplinary endeavors. American ethnobiologist Darrell Posey argued for this style of "advocacy" or "applied ethnobiology," whose goal is to reform the economic, environmental, and social policies that are at the root of many problems which affect people in rural and urban settings. This approach has stimulated a new generation of researchers to blend scientific research with an awareness of political and ecological problems, including the loss of biotic and genetic resources, indigenous struggles for land and resource rights, and negative aspects of globalization. The ultimate goal is sustainable development, as defined by cultural, ecological, and economic parameters.

The yen for integration goes beyond creating an interdisciplinary, applied field of study. Although ethnobiologists have tended to concentrate on the empirical side of local biological knowledge, there is renewed interest in symbolic and other interpretive approaches that could give a broader view of how people perceive their natural surroundings. Many ethnobiologists advocate adopting a participatory approach through which the entire study, from research design to application of the results, is conceived as a collaborative effort between local people and researchers. These developments represent a significant achievement for ethnobiology, a field that has always sought to blend perspectives from many cultural traditions.

See also: Agriculture, Traditional. Historical Awareness of Biodiversity. Hunter-Gatherer Societies, Ecological Impact of. Indigenous Peoples and Biodiversity. Social and Cultural Factors

References

Alexiades MN (1996) *Selected Guidelines for Ethnobotanical Research: A Field Manual*. New York: New York Botanical Garden.
Atran S (1990) *Cognitive Foundations of Natural History: Towards an Anthropology of Science*. Cambridge, UK: Cambridge Univ. Press.
Balée W (1994) *Footprints of the Forest. Ka'apor Ethnobotany – The Historical Ecology of Plant Utilization by an Amazonian People*. New York: Columbia Univ. Press.
Balick MJ and Cox PA (1996) *Plants, People and Culture: The Science of Ethnobotany*. New York: Scientific American.
Berlin B (1992) *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton, NJ: Princeton Univ. Press.
Clément D (1998) L'Ethnobiologie/ethnobiology. *Anthropologica* 40: 7–34.
Cotton CM (1996) *Ethnobotany. Principles and Applications*. London: Wiley.
Diamond J (1998) *Guns, Germs and Steel: A Short History of Everybody for the Last 13,000 Years*. London: Vintage.
Ellen RF, Parkes PSC, and Bicker A (eds.) (2000) *Indigenous Environmental Knowledge and Its Transformations, Studies in Environmental Anthropology*. Amsterdam: Harwood.
Ford RI (ed.) (1978) *The Nature and Status of Ethnobotany, Anthropological Papers No. 67*. Ann Arbor: Univ. of Michigan. Museum of Anthropology.
Johns T (1990) *With Bitter Herbs They Shall Eat It: Chemical Ecology and the Origins of Human Diet and Medicine*. Tucson: Univ. of Arizona Press.
Lévi-Strauss C (1966) *The Savage Mind*. Chicago: Univ. of Chicago Press.
Martin G (1995) *Ethnobotany*. London: Chapman & Hall.
Plotkin M and Famolare L (1992) *Sustainable Harvest and Marketing of Rain Forest Products*. Washington, DC: Island Press.
Posey D (ed.) (1999) *Cultural and Spiritual Values of Biodiversity*. London: Intermediate Technology.
Prance GT, Chadwick DJ, and Marsh J (1994) *Ethnobotany and the Search for New Drugs*. Chichester, UK: Wiley.
Schultes RE and Raffauf RF (1990) *The Healing Forest: Medicinal and Toxic Plants of the Northwest Amazonia*. Portland, OR: Dioscorides Press.
Simpson BB and Connor-Ogorzaly M (1995) *Economic Botany: Plants in Our World*, 2nd edn. New York: McGraw-Hill.

Eukaryotes, Origin of

B DeRennaux, University of Tennessee, Knoxville, TN, USA

© 2001 Elsevier B.V. All rights reserved.

This article is reproduced from Encyclopedia of Ecology, volume 2, pp 1428–1432, © 2001, Elsevier BV.

Introduction

Eukaryotes are cells commonly identified by the presence of a nucleus. Eukaryota is one of the three domains of life (bacteria and archea are the others and collectively referred to as prokaryotes) and encompasses single-celled organisms as well as all multicellular life. The evolution of eukaryotes is rather unique because it is substantially non-Darwinian; where Darwinian evolution largely concerns itself with beneficial genes arising by mutation and then spreading through differential reproductive success, the story of eukaryotes is one of significant horizontal gene flow (the exchange of genes across species) and of distinct organisms fusing into entirely new organisms.

Eukaryotes contain a variety of structures that distinguish them from prokaryotes (archea and bacteria), most notably the nucleus and a host of organelles (endoplasmic reticulum, Golgi bodies, peroxisomes, mitochondria, chloroplasts, etc.). They also possess flagella of unique structure (containing microtubules) and cytoskeletal structures made of microtubules and microfilaments (neither of which is found in prokaryotes). Eukaryotes also often organize their DNA around histones (proteins that act like spools) and package it into chromosomes. Some eukaryotes also participate in some form of sexual reproduction though many retain the ability to reproduce asexually.

Timing of the emergence of earliest eukaryotes is greatly debated. The first evidence of eukaryotes puts their origin at about 2.3–1.6 billion years ago. This range is derived primarily from phylogenetic analyses. Multicellular algae fossils are dated at 1.2 billion years, thus providing an absolute lower bound while fossils of single-celled organisms resembling algae are dated to around 2.2 billion years old, though scientists debate whether these represent eukaryotes or are merely similar in appearance.

Most calculations place the oldest eukaryotes close to the first global glaciation approximately 2.3 billion years ago. It is believed that a rise in the number in photosynthetic prokaryotes led to an increase in oxygen, which resulted in global cooling. This increase in oxygen is believed to have decimated populations of anaerobic prokaryotes. Some scientists believe that eukaryotes existed in low abundance before the rise in oxygen while others believe they originated shortly after. Scientists do agree however that the rise in oxygen is likely responsible for their increase in abundance and rise to multicellularity.

It is important to note that much of the theories regarding the origins of eukaryotes are speculative and it is unlikely that any one hypothesis will ever be championed above all others. However, as we sequence more genomes (particularly those of proposed early eukaryotes, and complete genome sequences of organisms from all domains of life) and develop more sophisticated phylogenetic techniques, it will be possible to narrow the likely possibilities.

Eukaryotes as Chimera

Eukaryotes are chimeric in that their DNA is the fusion of more than one genome. Phylogenetic analyses of eukaryotic genes result in many conflicting results, some eukaryotic genes appear most recently derived from archea, while others from bacteria. If these conflicts were infrequent, horizontal gene flow could be an adequate explanation. However, these conflicts are frequent and appear to fall into distinct categories. Genes involved in the maintenance and manipulation of the genetic code appear to be most recently derived from archea, whereas genes that are involved in metabolic processes are most similar to those of bacteria. Thus, it appears that at least two organisms are responsible for the major contributions to the eukaryotic genome and thus any theory on the evolution of eukaryotes must be one that can account for this pattern.

Curiously, genes believed to be of bacterial origin are sometimes grouped most closely with Gram-negative bacteria (like *Rickettsia*), while others are grouped most closely with Gram-positive bacteria (suggesting at least three major contributors). Debate exists however concerning whether these two groups are each monophyletic (consisting of a common ancestor and its descendents).

The Archezoa were (as many of these organisms are now being placed elsewhere) a group containing various eukaryotes that lacked mitochondria (amitochondriates) and often lacked or possessed 'primitive' versions of some eukaryotic traits. They are largely parasitic, inhabiting low oxygen environments, and were believed to descend from an ancient eukaryote before the acquisition of mitochondria. Recent phylogenetic analyses however have shown that many of these species are relatively recent with highly simplified structures resulting from their parasitism and anoxic environments. Structures similar to mitochondria have been found in many (though they may lack DNA) as well as nuclear genes normally associated with mitochondrial DNA.

Before the demise of Archezoa scientists had a rather gradual interpretation of eukaryotic evolution from simple cells to more complicated eukaryotes gradually with eukaryotic lineages diverging along the way acting as living fossils of earlier versions. With recent discoveries however, scientists are now of the opinion that most of the major features of eukaryotes (nucleus, cytoskeleton, ability to endocytose, and mitochondria) evolved very rapidly. It is now believed that if any early diverging lineages existed, they are now extinct and did not contribute to modern lineages.

No matter the theories concerning the origin of eukaryotes, there are hypotheses that can account for their chimeric nature. Woese has proposed that early in the evolution of eukaryotes horizontal gene flow was much more common, and as genomes became more complex and integrated these rates decreased. Doolittle posits that the advent of the cytoskeleton and ability to phagocytoze gave eukaryotes the potential to incorporate DNA from ingested organisms. Over time, numerous organisms contributed to the genome of early eukaryotes. While it is difficult to test either of these hypotheses, if future comprehensive phylogenetic analyses across a great many organisms point to 'few' donors for the origin of eukaryotes, then these ideas would be disproved.

Origin of Eukaryotes

An important feature of eukaryotes when discussing their evolution is that they are believed to have originally lacked cell walls or a rigid membrane as found in most prokaryotes. These early eukaryotes would have needed some kind of cytoskeleton to maintain integrity and shape. Cytoskeletal elements are also involved in eukaryotic reproduction, mobility, and the ability to endocytoze. Curiously, no solid evidence exists for the evolution of the cytoskeleton from some prokaryotic precursor (though actin and tubulin have weak similarity to archeon proteins). Thus, it seems likely that the cytoskeleton evolved largely *de novo* within the early eukaryote and also that many of the structural and functional characteristics of modern eukaryotes were possibly present in some of the earliest eukaryotes. Reasons for the loss of a cell wall are unclear though it may have been in response to bacteria-secreted antibiotics that often target cell walls.

Probably the simplest theory for the origin of eukaryotes is that the proto-eukaryote diverged from an archeon and that it later derived its bacterial characteristics from its proto-mitochondrial symbiote; however, another popular theory holds that a bacteria and archea genomes fused prior to the acquisition of the mitochondrion ancestor. In these scenarios it is unclear whether the early eukaryote possessed a nucleus before the mitochondrial event or whether it developed afterwards. Some hypothesize that the nucleus must have been in place prior to prevent harmful genetic interactions with the proto-mitochondria genome, others that the nucleus evolved in response to the acquisition.

Other theories hold that the nucleus is the result of a host bacteria fusing with an archeon or virus (something resembling mimiviruses which have rather large genomes and are less dependent on their hosts cellular machinery; referred to as viral eukaryogenesis). In some versions the bacterial host is RNA-based and the genes controlling the manipulation of genetic code were replaced by those of the DNA-based archeon or virus, leaving the other cellular machinery bacterial.

Figure 1 shows the four major competing theories for the origin of early eukaryotes resulting from ambiguity of two major events. First, have they diverged from an archeon or are they the result of a fusion between an archeon and a bacterium, and second, did eukaryotes exist in an amitochondriate phase during their evolution?

The 'hydrogen hypothesis' proposed by Martin and Muller holds that a methanogen archeon (which metabolizes hydrogen and carbon dioxide and releases methane) endocytozed the future mitochondria and that it provided the host cell with hydrogen and carbon dioxide as by-products of anaerobic respiration. This idea was inspired by hydrogenosomes which are simplified mitochondria often found in anaerobic eukaryotes and are believed to have evolved independently multiple times.

Lynn Margulis has proposed that eukaryotes originated when an archeon developed an ectosymbiotic relationship with a spirochete (a type of bacteria) and eventually their genomes fused; this places an exogenous origin for the flagellum and occurs before acquiring a bacterial endosymbiont. This idea was inspired when observing that some eukaryotes rely on symbiotic spirochetes for their motility (parabasalids within the termite gut). This theory has many bacteria-like genes originating from the spirochete and not the later mitochondria. However given that there is little similarity between spirochetes and eukaryotic flagella, this theory is currently rather unpopular.

Endosymbionts

Two of the more observable organelles in eukaryotes are mitochondria and plastids (plastids which contain chlorophyll *a* and *b* are chloroplasts). Mitochondria are often referred to as cellular 'power plants' and are responsible for oxidative phosphorylation; this enables aerobic respiration (which yields 15 times as much ATP from glucose than does anaerobic respiration alone), while chloroplasts allow for photosynthesis. The acquisition of these organelles would clearly have been an energetic boon to the eukaryotes possessing them.

Unlike theories on the origins of the eukaryotes, the 'endosymbiotic theory' has remained relatively unchanged for decades, has a host of strong evidence, and is almost universally accepted. The endosymbiotic theory dates back to 1883 when Andreas Schimper observed that chloroplasts divided similar to cyanobacteria (blue-green algae which are bacteria despite their common name). In 1905 Konstantin Mereschkowsky suggested that plastids were originally endosymbionts (beneficial organisms living within another organism), and in the 1920s, Ivan Wallin suggested the same for mitochondria. These ideas were based on visual similarity when viewed under light microscopy and were largely ridiculed until the later half of the 1960s, after which Lynn Margulis helped popularize the 'endosymbiotic theory' in her book *Origin of Eukaryotic Cells* (1970).

The 'endosymbiotic theory' states that mitochondria and plastids evolved from an engulfed bacteria and cyanobacteria, respectively. It is believed that a host cell engulfed an anaerobic bacterium (perhaps as prey or a parasite) and that over time a symbiosis arose, eventually becoming an obligate symbiosis (a mutually beneficial interaction required by both participants) before evolving into mitochondria. Later a mitochondria-harboring host engulfed a cyanobacterium, which similarly evolved into plastids. The mitochondrial event is believed to have occurred near the origin of eukaryotes

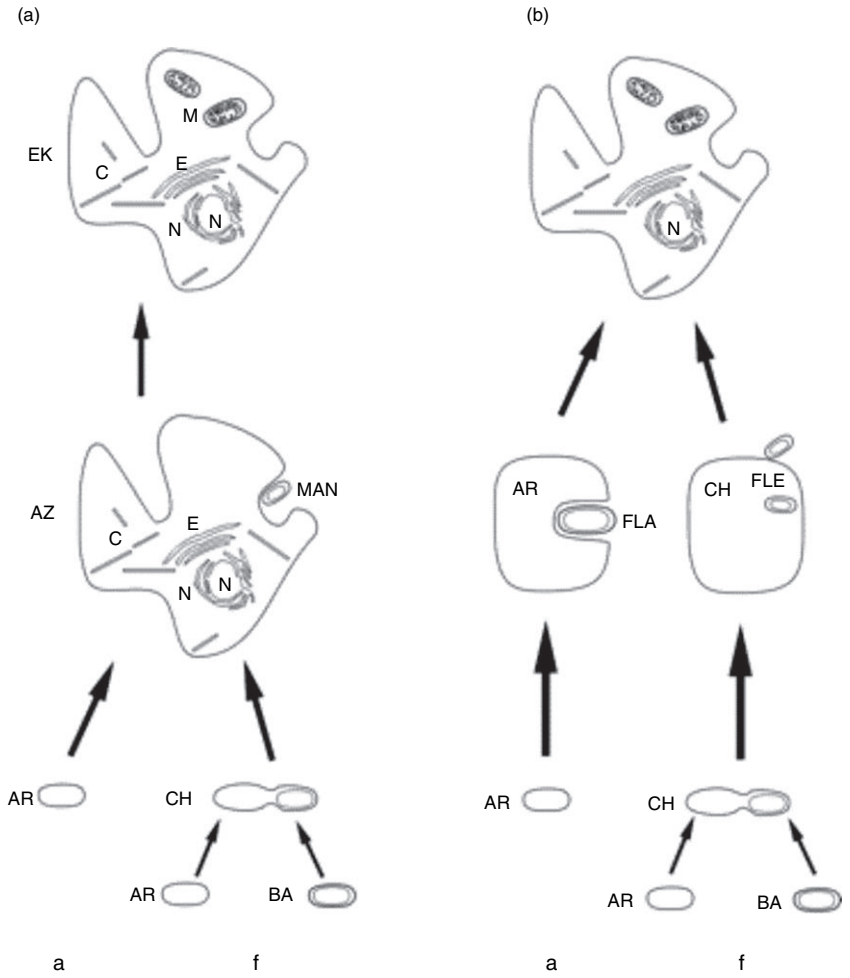


Figure 1 The main competing theories of eukaryotic origin. Schematic diagrams describing the Archezoa (a) and anti-Archezoa (b) hypotheses, and their archaeal (a) and fusion (f) versions as envisioned from genomic and biochemical perspectives. AR, archaeon; BA, bacterium; CH, chimeric prokaryote; AZ, archezoon; EK, eukaryote; MAN, mitochondrial ancestor; FLA, free-living α-proteobacterium; RLE, rickettsia-like endosymbiont; N, nucleus with multiple chromosomes; E, endomembrane system; C, cytoskeleton; M, mitochondria. Reproduced from Emelyanov VV (2003) Mitochondrial connection to the origin of the eukaryotic cell. *European Journal of Biochemistry* 270(8): 1599–1618, with permission from Wiley.

themselves, and the plastid event about a half billion years later.

Evidence

Much of the evidence for the endosymbiotic theory comes from the structure and handling of these organelles' genetic codes. Both mitochondria and plastids have DNA sequences in circles as that of bacteria. Their DNA also lacks histones (proteins that the DNA is wrapped around) which are present in eukaryotes and some archea. In addition, mitochondria and plastid transcription begin with the amino acid fMet (formylmethionine) as in bacteria, not Met (methionine) as in eukaryotes.

Ribosome sizes are questionable evidence for the endosymbiotic theory. Bacteria usually have ribosomes of 70s (Svedberg units) and eukaryotes usually have around 80s in their cytoplasm. While the mitochondrial and plastid ribosomes are usually of around 70s, they do in fact vary among species from around 60s to 80s, thus overlapping both bacterial and cytoplasmic eukaryote ribosome sizes.

Other evidence for the endosymbiotic theory comes from the two membranes usually surrounding these organelles. The inner membrane belongs to that of the original bacteria and outer membrane presumably a result from the original engulfment. The outer membrane has approximately a 1:1 protein–lipid ratio by dry weight, similar to many eukaryotic cytoplasmic membranes, while the inner membrane (which is made of two layers) has approximately a 3:1, similar to many bacteria. These organelles and bacteria also both utilize electron transport enzymes lacking elsewhere in eukaryotes.

Some of the best evidence for the endosymbiotic theory however comes from bioinformatics. Phylogenetic analyses of various bacteria, mitochondria from various hosts from various kingdoms, and nuclear DNA from those hosts usually place mitochondria as most related to a group of bacteria known as proteobacteria, often placed closest to *Rickettsia*

and other α-proteobacteria. The α-proteobacteria as a group are almost entirely symbiotic or parasitic which may have predisposed the mitochondrial ancestor to an existence within its host. Chloroplasts are most often placed next to cyanobacteria and both contain thylakoids and chlorophyll *a*; cyanobacteria are also involved in a number of symbioses including lichens and corals.

Problems

Some inconsistencies do exist between bacteria and plastids/mitochondria. Mitochondria possess the ability to fuse with other mitochondria (this is believed to be the result of simplification that has occurred within the mitochondrial membranes over time) and the genetic information within a mitochondrion or plastid also differs in some ways to that of bacteria. These organelles possess higher concentrations of introns (DNA that will be spliced out prior to translation) than bacteria and intron types not found in bacteria. The genomes of these organelles are also much smaller than those of bacteria.

Gene flow between the nucleus and the organelles is believed to be responsible for the genetic discrepancies. Gene flow from the nucleus to an organelle first requires a copy of a segment of DNA from the organelle to insert into the nuclear DNA and gain expression. Loss of function can then occur in the original gene. A similar process can explain migration from the nucleus to the organelles. While organelle DNA fragments within nuclear DNA are rather common, nuclear DNA fragments within organelle DNA are rare (the exception being plant mitochondria which has experienced a rather large increase in genome size over time). Such genetic transfers operate with a frequency similar to that of mutation.

The introns within the organelles and the increase in the plant mitochondria genome over time are thought to be due to a transfer from the nucleus to the organelles. The size discrepancy between the genome sizes of chloroplasts and mitochondria is believed to be a result of selection for genes originally possessed by the organelles to reside in the nucleus.

Selection for the migration of genes to the nucleus comes in many possible forms. First, it is believed to be physically more difficult for DNA to enter these organelles as compared to the nucleus. Population genetics also tells us that sexual reproduction (which many eukaryotes posses) enables the more rapid spread of advantageous genes, and overcomes Muller's ratchet (the idea that asexual reproduction lacks an effective means of ridding the genome of deleterious mutations that gradually build up). Also, chloroplasts and mitochondria perform redox (reduction–oxidation) reactions, which are prone to producing deleterious mutations. Finally, it has been proposed that smaller organelle genomes would replicate faster than those with redundant genes and thus be selected for.

The reasons for maintenance of organelle DNA are not as clear. Some reasons might include, that the remaining genes code for proteins that cannot be easily transported into the organelle (possibly membrane-spanning or strongly hydrophobic proteins) or that these genes are harmful if expressed in the cytoplasm. Another potential reason may be the slight differences in organelle versus nuclear transcription/translation.

Secondary Endocytosis of Plastids

Green algae and plants are believed to be descendents of the original plastid-harboring eukaryote. However, some protists contain plastids that are believed to be from other eukaryotes. These protists have within their endoplasmic reticulum an endosymbiont resembling a eukaryote. This endosymbiont possesses its own plasma membrane and ribosomes (which are both similar to those of eukaryotes) as well as the remnant of a nucleus containing its own DNA (a nucleomorph), and finally plastids. Thus, these organisms have four sets of DNA: mitochondrial, plastid, nuclear, and nucleomorph. Some protists may even contain plastids from tertiary endocytosis.

Other Possible Symbiotic Events

Christian de Duve proposed that peroxisomes (which contain enzymes for oxidative reactions) may also have originated as endosymbionts; however, because of their single layer membranes, lack of DNA, and the recent observation that peroxisomes can be formed from the endoplasmic reticulum, it is now generally accepted that they arose endogenously.

Lynn Margulis proposed in 1967 that along with mitochondria and plastids, the eukaryotic flagella was also the result of symbiosis (*see* Origin of Eukaryotes). However, this fell largely out of favor when flagella were found not to possess DNA, and subsequently that spirochetes are rather dissimilar to eukaryotic flagella.

References

Brown JR and Doolittle WF (1997) Archea and the prokaryote-to-eukaryote transition. *Microbiology and Molecular Biology Reviews* 61: 456–502.

Emelyanov VV (2003) Mitochondrial connection to the origin of the eukaryotic cell. *European Journal of Biochemistry* 270(8): 1599–1618.

Gupta RS and Golding GB (1996) The origin of the eukaryotic cell. *Trends in Biochemical Sciences* 21: 166–171.

Katz LA (1998) Changing perspectives on the origin of eukaryotes. *Trends in Ecology and Evolution* 13: 493–497.

Katz LA (1999) The tangled web: Gene genealogies and the origin of eukaryotes. *American Naturalist* 154: S137–S145.

Roger AJ (1999) Reconstructing early events in eukaryotic evolution. *American Naturalist* 154: S146–S153.

Europe, Ecosystems of

Ladislav Mucina, Curtin University, Perth, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biome A large-scale complex of ecosystems sharing similar climate and large-scale natural disturbance regime, resulting in high similarity of vegetation structure.
Ecological biodiversity Variety of biotic communities (plant, animal, and microbial communities) and their complexes (ecosystems, landscapes, and biomes).
Formation Large-scale (subcontinental or continental) vegetation complex defined primarily on the basis of a combination of dominating life-forms (hence, vegetation structure).
Habitat A place of dwelling of a biotic community (or a complex thereof) showing a particular combination of ecological factors occupying a certain area in a certain period of time.
Orobiome Mountain range characterized by a particular climatic pattern and a characteristic sequence of vegetation zones.
Syntaxon (plural syntaxa) A category of vegetation typology based on the floristic–sociological approach (known also as the Braun–Blanquet approach); the basic syntaxon rank is association, which further groups into alliances, which group into orders, and orders group into classes.
Vegetation mapping Scientific activity aimed at capturing the spatial extent of the vegetation types on a map.
Vegetation megazone Large-scale vegetation complex characteristic for a zonobiome.
Vegetation survey Product of research activity aimed at the description and classification of vegetation cover on various levels of complexity in a certain geographic area using various field and data-evaluation methods.
Zonobiome Broad ecological topographical unit characterized by a certain climatic pattern.

Introduction

Hardly any place in the world supports in a relatively small area such a variety of countries, peoples, cultures, histories, languages, political views, cuisine, and types of cheese and wine as does Europe. There might be more languages and people in India, higher species diversity of plants in the tropical rain forest of Columbia or Brazil, or more diverse cuisine in China, but there is only one Europe.

Europe's nature is showing many faces, both pristine ones in the form of tracks of tundra and forests and created (or miscreated) ones – those carrying the signs of human activities. Europe is a patchwork of ecosystems connected by an intricate net of gradients of ecological factors ranging from those controlling continental patterns of vegetation megazones and biomes to local gradients creating small-scale mosaics of biotic communities.

This article discusses the variability of European ecosystems and a major framework of biodiversity patterns at various scales of complexity.

Europe is a continent, a large chunk of land; therefore, the author discusses this subject from the viewpoint of the diversity of large-scale ecological units such as vegetation megazones and biomes. The terrestrial habitats (and their complexes) will dominate this discourse and the author emphasizes the diversity of vegetation assemblages, which are traditionally considered the core of the structure and functioning of all terrestrial ecosystems.

Europe is approximately 10 million km^2 and spans 35–81° N of latitude and 60° E–10° W of longitude. Although Europe is the second-smallest continent (7% of the world's land surface), the basic classification of ecosystem diversity can be discussed only at a very large scale of biome. Not only the limited extent of this article but also the extraordinary wealth of scientific knowledge collected over centuries and the diversity of the subject impose constraints on the level of detail and accuracy of this discourse.

Ecosystem Diversity: Concepts and Approaches

Ecosystem, often defined as the union of biotic community and its environment, presents an obvious difficult scaling problem. Ecosystems as real entities occur in space and time. Depending on the conditions, one can recognize natural or artificial borders allowing the classification of ecosystems, an exercise meant to set a framework for simplifying the complexity of ecosystems and featuring them in a synoptic way.

Classification of Vegetation

Most European ecosystems are obviously terrestrial ecosystems accompanied by semiterrestrial and aquatic ones. Vegetation (defined as the unity of plants occupying a certain area in a certain time) is the major biotic element of terrestrial ecosystems. Hence, the problem of classifying terrestrial ecosystems is a problem of classification of vegetation.

The approaches to the classification of vegetation are manifold and largely depend on the criteria (and their weighting), aims, and means (Mucina, 1997a). Several basic approaches to the classification of vegetation have emerged during the past century (Whittaker, 1978). Among these, the floristic–sociologic approach (Westhoff and van der Maarel, 1978) became a standard communication tool among

European (and other) vegetation scientists involved in or harvesting fruits of vegetation classification exercises.

Floristic–Sociologic Approach to Vegetation Classification

The basis of the floristic–sociological approach to the classification of vegetation is the notion of the total floristic composition and the notion that some species indicate the environmental conditions better than others. These are called diagnostic species (Braun-Blanquet, 1964). Vegetation stands are selected by following various criteria, among which the so-called representativity of the studied vegetation type is the leading one. This approach is aimed at the recognition of plant communities – plant assemblages recurrent in space and in time and showing a distinct floristic composition that reflects certain combinations of current and past habitat conditions. The basic vegetation unit of this approach is the so-called association, a theoretical construction that is a result of the abstraction classification process involving many vegetation stands. A hierarchy of vegetation units based on groups of species having similarity between plant communities and their groups is also a vital part of the floristic–sociological approach. Floristically and ecologically similar associations are grouped into alliances, the alliances are grouped into orders, and the orders are grouped into classes. The vegetation classes can be less formally grouped into divisions. Particular associations, alliances, orders, classes, and divisions are termed syntaxa (Westhoff and van der Maarel, 1978) of various ranks and build the syntaxonomic hierarchy. An elaborate, highly formalized system of rules has been introduced to govern the formation of names of the syntaxa.

Dominance Approach to Classification of Vegetation

Another approach to the classification of vegetation emphasizes the role of the dominant species or dominant growth forms, marking a departure from using plant taxonomy as the currency. Without discussing this approach in great detail (Whittaker, 1978), the notion of formation is the core tool of the dominance approach. The formation is the vegetation component (thus, in terrestrial ecosystems it is the leading element) of the concept of biome. Although the floristic–sociological approach is well suited for the classification and description of vegetation on small geographic scales, formation and biome are, for the same purpose, well suited on large scales (subcontinental and continental).

Classification of Habitats

Despite much critical challenging of some subjective points regarding the criteria for selection of stands and criteria of weighting of species in the classification process and sometimes awkward nomenclature, the floristic–sociological approach is aimed at enhancing the effectiveness of communication when addressing units of vegetation cover. It is therefore not surprising that it was the terminology and concepts of this approach that were adopted as the basis of the majority of the units of the habitat systems developed for various purposes by the former European Communities (now European Union) authorities.

The classification of habitats passed several stages of development, spanning CORINE habitat classification (Commission of European Communities, 1991), Palearctic habitat classification (Devillers and Devillers-Terschuren, 1996), and EUNIS3 habitat classification (Davies and Moss, 1998). Obviously, the purpose of habitat classification is a practical one. It should serve (and serves) important purposes by the delimitation and evaluation of land-use characteristics, projection and management of nature reserves of other areas of special nature conservancy interests (e.g., the Natura 2000 network of Special Areas of Conservation), etc. Not surprisingly, such a habitat classification scheme is a legal standard within the borders of the European Union.

Biodiversity of Habitats and Vegetation

Sources

The current vegetation landscape of Europe is a result of manifold forces forming and reforming the habitat networks during the past few thousands of years; undoubtedly, some features of this habitat and vegetation patchwork are very old. One can hardly avoid using oversimplification when talking about the sources of community (including vegetation and animal communities) and ecosystem diversity of present-day Europe. The author distinguishes four major prerequisites to the community and ecosystem diversity (**Table 1**):

1. Taxon diversity: Plant and animal communities are composed of individuals belonging to several (often many) taxa. Speciation and associated processes of within-taxon diversification in addition to migrations of taxa – either ancient, natural ones or current ones largely driven by human activities – are the major sources of taxon diversity.
2. Habitat diversity: Habitats are homes of biotic communities and the greater the variety of habitats, the greater the variety of biotic communities populating them. The nature of habitats is determined primarily by ecological factors – their characteristics and dynamics. Regardless of whether one considers spatial or temporal scales, one can recognize geological, geomorphologic, and soil-formation processes accompanied by hydrological dynamics and

Table 1 Prerequisites and sources of the ecosystem and community diversity

Prerequisite	*Sources*
Taxon diversity	Speciation
	Extinction
	Migrations
Habitat diversity	Geological processes
	Geomorphological processes
	Soil-formation processes
	Hydrologic dynamics
	Climate change
Diversity of biotic interactions	Competition and related negative interactions
	Facilitation and related positive interactions
Diversity of human interference	Disturbance and removal of habitats
	Creation of new habitats

climate changes as the major natural sources of the habitat diversity.

3. Diversity of biotic interactions: The individuals representing different taxa interact within the biotic communities in many ways. They may compete for resources, facilitate each other's growth and reproduction by a plethora of positive interactions, and may behave indifferently. Undoubtedly, the biotic interactions form the face of each community and thus contribute to ecological diversity.
4. Diversity of human interference: Humans have made themselves the center of the universe and from this point of view have also become the dominant source of the disturbance (up to complete removal) but also creation of new habitats and new biotic communities through manipulating ecological factors and facilitation of species migrations.

Patterns

Biomes of Europe

Europe has six (of the world's nine) zonobiomes as defined by Heinrich Walter (Walter and Breckle, 1991; pp. 22–25). These include, from south to north, Mediterranean zonobiome (IV), warm-temperate zonobiome (V), nemoral zonobiome (VI), continental zonobiome (VII), boreal zonobiome (VIII), and polar zonobiome (IX). Furthermore, many zonoecotones mediate between the particular zonobiomes and these are designated as IV–V, V–IV, V–VI, VI–VII, VII–VI, VII(III), VIII–VI, and IX–VIII. The zonoecotone III–IV, mediating between subtropical zonobiome (not represented on the continent) and Mediterranean zonobiome, is found on the Canary Islands.

Following the combined approach to define biomes using climate and vegetation structure, one can tentatively recognize 10 biomes in Europe, the Canary Islands and the Azores. (**Table 2**).

Surprisingly little attention has been paid to the definition of extant European biomes in comparison with other parts of the world. This can, to a huge extent be attributed to traditions in classification of ecosystems in Europe. The dominance-based approach to the classification of vegetation leading directly to the definition of biomes when exercised on a large scale is the predominant approach in parts of the world in which there is a lack of advanced knowledge on flora. Europe is undoubtedly a region in which the flora-based approach has a long and firmly rooted tradition.

Diversity of Vegetation Types

To list all syntaxa described in Europe is impossible at this stage and here. A recent account of the high-rank syntaxa (Mucina, 1997b; Rodwell *et al.*, 2002; and an unpublished checklist of high-rank syntaxa by Mucina L and 24 collaborators) revealed that the European vegetation can be classified into 88 classes (of which eight are endemic to Macaronesian Islands), 248 orders, and 945 alliances (these counts exclude the Azores and Caucasus). **Table 3** provides a list of the classes and comments on their zonality status. The number of associations is unknown, although one estimate suggests there might be as many as 3000. For further details on the diversity of syntaxa occurring in various European countries, refer to the many national vegetation surveys (and transnational vegetation accounts) listed in **Table 4**.

Several interesting biodiversity patterns can be observed within the groups of syntaxa, including classes, orders, and alliances (**Table 3**):

1. The zone of the Temperate Deciduous Forests is the dominating biome of Europe, both spatially and in the number of syntaxa characterizing the natural, postglacial forests and associated scrub, but also secondary vegetation types (such as grasslands and man-made vegetation) replacing the forests in cultural landscapes of Western, central, and Eastern Europe.
2. The Mediterranean region (supporting the Mediterranean biome and Mediterranean orobiome) shows a high concentration of classes and subordinated syntaxa. This can be partly ascribed to high gamma diversity (richness of regional floras), high regional and endemism as a result of the insular character of many types of Mediterranean habitats and last but not least to due to the high concentration of refugia across the Region. Not only are numerous true marine islands very abundant within the Mediterranean but also the mountain summits form an archipelago of their own, demonstrating their own rates of evolution and combination of ecological and biogeographical factors leading to specific plant assemblages.
3. The high-latitude regions (arctic and boreal zones) show lower diversity of vegetation types partly as a result of low extant gamma diversity (regional species pools), being a result of repeated glacial-cycle disturbance resulting in

Table 2 Preliminary scheme of European biomes with corresponding Walter's zonobiomes

Biome	*Zonobiome or zonoecotone*	*Comment*
Arctic Desert Biome	IX in part	Northern Europe, Arctic islands
Tundra Biome	IX in part; IX–VIII	Scandinavia, northern Russia
Boreal Forest Biome	VIII; VIII–VII; VIII–VI	Scandinavia, Russia
Steppe Biome	VII; VI–VII	Ukraine and Russia
Temperate Deciduous Forest Biome	VI; VII–VI	Most of western and Central Europe
Mediterranean Biome	IV; IV–V	Entire Mediterranean basin
Warm-Temperate Evergreen Forest Biome	V; V–VI	Macaronesia, Colchis (Transcaucasus)
Warm-Temperate Semidesert Biome	III–IV	Canary Islands, southern Spain?
Continental Semidesert Biome	VII(III)	Around Caspian Sea

Table 3 Zonality and azonality status of European phytosociological classes and their correspondence with the biome and orobiome classification scheme

Classes	*O*	*A*	*Diagnosis*	*Note*
I. Zonal vegetation				
I.A. Tundra Biome				
Loiseleurio-Vaccinietea	1	7	Arctic-boreal tundra scrub and (sub)alpine dwarf-scrub	Relictual in Nemoral Orobiome
Carici-Kobresietea	2	6	Circumarctic dwarf-scrub and graminoid chionophobous tundra and relict wind-exposed alpine short grasslands on base-rich substrates	Relictual in Nemoral Orobiome
Salicetea herbaceae	1	6	Arctic and alpine-subnival vegetation of long-lasting snow-beds and slopes irrigated by melt waters	Relictual in Nemoral Orobiome
I.B. Boreal Forest Biome				
Vaccinio-Piceetea	3	9	Holarctic coniferous forests on acidic soils of boreal zone and high-altitudes of Nemoral Orobiomes	Also in Nemoral Orobiome and relictual in Mediterranean Biome
Pyrolo-Pinetea	2	2	Euro-Siberian (sub)continental thermophilous pine forests on deep sands on exposed rocky outcrops	Relictual in realm of Nemoral Biome
Roso-Pinetea mugo	1	3	Subalpine krummholz of mountains of Central and south-western Europe	
I.C. Temperate Deciduous Forest Biome (Nemoral Biome)				
Carpino-Fagetea	6	28	Mixed broad-leaved deciduous forests on eutrophic, anhydromorphic soils temperate zone of Europe	
Quercetea roboris	2	9	Acidophilous oak, oak-birch and beech forests on nutrient-poor soils	
Quercetea pubescentis	3	24	Thermophilous (semi)deciduous oak forests of submediterranean provenience	
Brachypodio-Betuletea pendulae	1	3	Hemiboreal pine and birch-pine open grassy forests on fertile soils of the South Urals and southern Siberia	
Erico-Pinetea	4	9	Relictual pine forests and related scrub on calcareous and ultramafic substrates	
Rhamno-Prunetea	5	23	Shrubland vegetation seral or marginal to broad-leaved forests of the nemoral zone of Europe	
Lonicero-Rubietea plicati	1	2	Acidophilous scrub, hedges and scrub of forest clearings on dry sandy, nutrient-poor soils of Western Europe	
Calluno-Ulicetea	3	18	Dwarf-shrub and mat-grass heaths on acidic soils from the planare to montane belts	Also marginally associated with Boreal biome
Koelerio-Corynephoretea	6	24	Pioneer vegetation on primitive soils and rocky outcrops in regions with mild winter climate	Also marginally associated with Boreal biome
Molinio-Arrhenatheretea	12	36	Secondary mesic and wet meadows on fertile soils	Also marginally associated with Boreal biome; relictual or azonal within Mediterranean biome
Trifolio-Geranietea	3	13	Thermophilous woodland fringe and tall-herb vegetation of nutrient-poor sites in the submediterranean to subboreal zones of Europe	Also marginally associated with Boreal, Steppe and Warm-Temperate Evergreen Forest biomes
I.D. Steppe Biome (Temperate Grassland Biome)				
Festuco-Brometea	11	49	Dry grassland and steppe vegetation of mostly base-rich soils of the submediterranean to the hemiboreal zones of Europe	Relictual in realm of Nemoral and Mediterranean biomes
Festucetea vaginatae	1	6	European continental steppes and dry grasslands on sandy soils	Relictual in realm of Nemoral Biome
Festuco-Puccinellietea	6	29	Saline steppes and secondary steppic grasslands of the continental regions of Eastern and Southern Europe	Relictual in realm of Nemoral Biome (especially in the leso-steppe zone)

(*Continued*)

Table 3 Continued

Classes	*O*	*A*	*Diagnosis*	*Note*
I.F. Mediterranean Biome				
Quercetea ilicis	3	23	Mediterranean maquis, pine, and oak forests and associated sclerophyllous scrub	Relictual in realm of Nemoral Biome
Rosmarinetea officinalis	5	17	West and Central Mediterranean scrub (tomillares, esplegueras, romerales, matorrales, garrigue) on calcareous substrates	
Cisto-Micormerietea	2	7	Thermo- to supramediterranean phrygana and garrigue of the Central and Eastern Mediterranean	
Cisto-Levanduletea	2	7	West and Central Mediterranean scrub (jarales, matorrales, garrigue, phrygana) on acidic siliceous and ultramafic substrates	
Onosmo-Ptilostemetea	2	4	Mediterranoid garrigues on raw flysch and lime-rich substrates of the Euxinic seabords	Endemic to the Black Sea region
Cytisetea scopario-striati	2	9	West Mediterranean (mainly Iberian) broomy scrub (retamares, piornales, escobonales) and related mantle communities	
Oleo-Rhamnetea crenulatae	1	5	Macaronesian dry forest, matorral, and tomillar	Endemic to Macaronesia
Cytiso-Pinetea canarensis	1	2	Canarian pine forests and related juniper scrub	Endemic to Macaronesia
Thero-Brachypodietea	4	15	Mediterranean pseudosteppe grasslands on calcareous rocks and relic steppes on deep clayey soils	
Stipo-Agrostietea castellanae	2	4	Ibero-Atlantic and Madeirean meso- to supra-mediterranean perennial grasslands	
Stipo-Trachynietea	3	19	Mediterranean calciphilous communities of annual and ephemeroid herbs and grasses	
Helianthemetea guttati	2	18	Mediterranean acidophilous communities of annual and ephemeroid herbs and grasses	
I.G. Warm-Temperate Semidesert Biome				
Kleinio-Euphorbietea	1	1	Macaronesian succulent scrub on semidesert lava beds (taibal and cardonal)	Endemic to Macaronesia (possibly Late Tertiary relict)
Pegano-Salsoletea	5	15	Thermo-mediterranean and Macaronesian halo-nitrophilous coastal scrub of (semi)arid regions	Also associated with Mediterranean biome in azonal habitats
I.H. Continental Semidesert Biome				
Artemisietea lerchianae	2	5	Aralo-Caspian semideserts	
Petrosimonio-Kalidietea	2	3	Continental hypersaline scrub of edges inland saline lakes and seaboards of Eastern Europe and Central Asia	Also associated with Steppe Biome (relictual?)
I.I. Warm-Temperate Evergreen Forest Biome (Laurisilva Biome)				
Pruno hixae-Lauretea	3	10	Macaronesian laurisilva and related scrub	Endemic to Macaronesia (including Madeira); Late Tertiary relict
Lauro azoricae-Juniperetea	2	3	Azorean broad-leaved evergreen laurisilva and related matle and thickets	Endemic to Azores; Late Tertiary relict
II. Vegetation of orobiomes				
II.A. Boreal and Nemoral Orobiomes				
Elyno-Seslerietea	5	22	Alpine and subalpine calcareous swards of nemoral mountain ranges	
Juncetea trifidi	5	22	Alpine acidophilous grasslands of Central European and Balkan mountains and Arctic and boreo-arctic rush swards	Also associated with Tundra Biome
Mulgedio-Aconitetea	5	20	Scrub and tall-herb vegetation at high altitudes, moistened and fertilized by percolating water	Also associated with Boreal Biome

(*Continued*)

Table 3 Continued

Classes	*O*	*A*	*Diagnosis*	*Note*
II.B. Mediterranean Orobiomes				
Pino-Juniperetea	2	13	Dry relict oromediterranean and submediterranean orotemperate pine forests, juniper woodlands and related scrub of the Mediterranean	
Festucetea indigestae	2	5	Iberian and North African supra- to cryomediterranean xerophilous silicicolous fescue grasslands	Endemic to Spain
Festuco-Ononidetea	2	10	Dry basiphilous supra-oromediterranean and oro-temperate submediterranean grasslands and related dwarf scrub of south-western Europe	
Carlino-Genistetea lobelii	1	2	Cyrno-Sardean oromediterranean cushion-tragacantic scrub and related grasslands	Endemic to Corsica and Sardinia
Rumici-Astragaletea	3	4	Siculo-Calabrian oromediterranean and upper mesomediterranean pulvinate scrub and related grasslands	Endemic To Italy
Daphno-Festucetea	2	6	East-Mediterranean oromediterranean calcareous xeric grasslands and tragacanthic phrygana	Endemic to Greece
II.C. Semidesert Orobiomes				
Spartocytisetea supranubii	1	1	Canarian high-altitude volcanic semidesert scrub	Endemic to Canary Islands
Violetea cheiranthifoliae	1	1	Canarian herbaceous high-altitude tallus vegetation	Endemic to Canary Islands
III. Azonal vegetation				
III.A. Azonal aquatic vegetation				
III.A.1. Azonal aquatic saline vegetation				
Zosteretea	3	3	Eel-grass swards on muddy and sandy substrates in the sea sublittoral and eulittoral zones	
Ruppietea maritimae	1	2	Submerged rooted communities of brackish waters	
III.A.2. Azonal aquatic freshwater vegetation				
Lemnetea	1	1	Free-floating duckweed communities of still, relatively nutrient-rich, fresh waters in warmer regions	
Potametea	2	8	Communities of rooted, floating or submerged plants in mesotrophic and eutrophic fresh waters	
Charetea fragilis	1	2	Submerged macro-algal (stonewort) swards	
III.B. Azonal amphibian vegetation				
III.B.1. Azonal amphibian saline vegetation				
***Spartinetea maritim*ae**	1	1	Pioneer intertidal perennial grasslands	
Thero-Salicornietea	1	5	Pioneer intertidal communities dominated by annual succulent halophytes, rarely occurring also on banks of inland salt pans	
Salicornietea fruticosae	2	10	Mediterranean and thermo-atlantic perennial salt-marsh scrub	
Juncetea maritimi	5	12	Perennial grasslands and herblands of coastal and inland salt-marshes and sea-cliffs	
Crypsidetea aculaetae	1	1	Pioneer ephemeral dwarf-grass vegetation of periodically flooded saline habitats in submediterranean and (sub)continental Eurasia	
III.B.2. Azonal amphibian freshwater vegetation				
Bidentetea tripartiti	1	2	Summer-annual, pioneer vegetation of periodically flooded edges of water bodies	

(*Continued*)

Table 3 Continued

Classes	*O*	*A*	*Diagnosis*	*Note*
Oryzetea sativae	1	1	Weed communities of rice fields	
Littorelletea	2	6	Pioneer ephemeral hairgrass swards and bladderwort communities in oligotrophic waters	
Isoeto-Nano-Juncetea	2	9	Pioneer ephemeral dwarf-cyperaceous vegetation on periodically flooded (freshwater) habitats	
Montio-Cardaminetea	2	10	Moss-rich vegetation of water springs	
Phragmiti-Magnocaricetea	3	10	Swamp and reed vegetation of fresh or slightly brackish waters	
Scheuchzerio-Caricetea	2	8	Vegetation of transitional mires, fens, and bog hollows	
Oxycocco-Sphagnetea	2	4	Holarctic ombrotrophic bog and wet heathland vegetation on acid oligotrophic peat	
III.B.3. Azonal alluvial and mire forests and scrub				
Populetea albae	2	9	Riparian gallery forests of Eurosiberian and Mediterranean regions	
Salicetea purpureae	1	8	Riparian willow scrub and thickets of Eurosiberian and Mediterranean regions	
Nerio-Tamaricetea	2	9	Mediterranean riverine scrub	
Molinio-Betuletea pubescentis	2	4	Azonal forest and shrub communities of mesotrophic mires	
Vaccino uliginosi-Pinetea	1	2	Forest and shrub communities of oligotrophic mires	
Alnetea glutinosae	1	1	European alder carr forests	
Franguletea	1	1	Willow carr communities of Western and sub-Atlantic Central Europe	
III.C. Azonal terrestrial vegetation				
III.C.1 Azonal coastal vegetation				
Cakiletea maritimae	3	4	Pioneer vegetation of nitrophilous annuals on strandlines of sandy and shingle beaches of Atlantic, Mediterranean and Euxinic coasts	
Crithmo-Staticetea	2	13	Herb-rich and dwarf-scrub vegetation of coastal salt-splashed cliffs and walls Atlantic and Mediterranean seabords	
Saginetea maritimae	2	6	Ephemeral winter-annual vegetation of disturbed saline coastal habitats and in inland saline badlands	
Ammophiletea	2	8	Coastal dune vegetation of temperate and mediterranean seaboards of Europe, North Africa, Middle East and Caspian Sea	
Honckenyo-Elymetea	1	1	Vegetation of coastal shingle, boulders or rocky cliffs, enriched with organic detritus of boreal and arctic shores of Europe	
Polycarpo-Traganetea	1	2	Canarian, Cabo Verdan and West-Saharian oceanic halophilous coastal desertic dune scrub	Macaronesia and North Africa
III.C.1 Azonal chasmophytic vegetation				
Asplenietea trichomanis	16	78	Vegetation of crevices and surfaces of rocky cliffs and walls	
Adiantetea	1	2	Fern-rich rocky communities of water-splashed habitats	
Thlaspietea rotundifolii	13	48	Vegetation of scree habitats and pebble alluvia of Europe and Arctic islands	
Aeonio-Greenovietea	1	3	Macaronesian chasmophytic vegetation of exposed volcanic rocks	Endemic to Macaronesia
III.C.1 Synanthropic (anthropogenic) vegetation				
Stellarietea mediae	8	44	Annual weed communities of arable crops, gardens and waste places	

(*Continued*)

Table 3 Continued

Classes	*O*	*A*	*Diagnosis*	*Note*
Polygono-*Poetea annuae*	1	3	Therophyte-rich vegetation of trampled habitats	
Artemisietea vulgaris	4	14	Perennial and thistle-rich (sub)xerophilous ruderal communities of temperate and Mediterranean regions	
Galio-Urticetea	4	12	Tall-herb seminatural saum vegetation of nutrient-rich fringes of forests and scrub and of water courses of temperate Europe	
Epilobietea angustifolii	1	2	Spontaneous pioneer tall-herb vegetation of acidic soils of forest margins and clearings	

The regions covered in this list includes Europe proper, Iceland, Svalbard, Canary Islands, Madeira, and Azores. The nomenclature of classes follows so far unpublished checklist of high-rank syntaxa of Europe by Mucina *et al.* (in prep.). O, number of phytosociological orders; A, number of phytosociological alliances. The classes printed in bold (Groups I and II) are zonal (normal font) or intrazonal (italics). For correspondence of the Biomes with the Walter's zoniobiome scheme see **Table 2**.

dramatic impoverishment of the flora and also because of the large-scale occurrence of uniform habitat complexes, and due to adversity of climate.

4. Among the azonal (and intrazonal) units, the *Asplenietea trichomanis* and *Thlaspietea rotundifolii* (classes comprising vegetation of special habitats such as rock faces, fissures, and screes) show the highest diversity of orders and alliances – reflecting the high diversity of ecological–biogeographic patterns. These two classes probably house most European local endemic species.
5. The dry grasslands of the *Festuco-Brometea* (and partly also of the *Trifolio-Geranietea*) show extensive primary and secondary distribution in the steppe and leso-steppe zones (the latter constituting the transitional zone between the steppes and the forest biomes) and comprise the most species-rich plant communities of Europe. Russian steppes, for instance, can support more than 60 species per square meter. Large-scale distribution and rich regional species pools are prerequisites to high beta diversities manifested in a high number of vegetation types.
6. The synanthropic (man-made and man-controlled) vegetation of Europe is also very diverse in terms of the number of syntaxa. The class of *Stellarietea mediae* comprises 44 alliances. The same pattern applies to the *Molinio-Arrhenatheretea*, comprising 36 alliances classifying secondary mesic grasslands. These numbers are indicative of the high diversity of disturbance factors related to human activities such as agriculture, silviculture, building activities, and unintentional transport of seeds and fruits over continental borders. The flora of Europe was enriched by many alien species adding to the diversity of vegetation types. Yet, personal classification bias could have inflated the number of syntaxa to an extent and more work is needed to establish the identity of these units.
7. Mediterranean classes such as *Rosmarinetea officinalis* and *Quercetea ilicis* also show high diversity with regard to the number of alliances. Still higher diversity becomes obvious on the level of associations reflecting the relict nature of habitats and diversity of disturbance regimes in the Mediterranean.
8. The lowest diversity of vegetation types is encountered in the classes typifying vegetation of the freshwater, marine, and coastal vegetation. Often, only a few orders or alliances are found within classes such as *Lemnetea*, *Potametea*, *Ruppietea*, *Spartinetea*, and *Hockenyo-Elymetea*. However, on the whole, the nonterrestrial complex of habitats is very diverse, which is reflected in the high number of phytosociological classes.

Vegetation Mapping

Vegetation cover is a spatial phenomenon. Regardless of which criteria (physiognomic, structural, ecological, floristics and the like; see Whittaker, 1978) are used to classify (hence simplify) vegetation, vegetation types can be mapped – featured as a two-dimensional graphical model called a *vegetation map*. While a scientifically sound vegetation floristic-ecological classification system should carry a message on ecological (and evolutionary) assembly of the classified vegetation (and flora), a vegetation map places this message in a spatial context. It is therefore not surprising that vegetation maps are sought for as prime sources of information for making informed decisions on conservation and for providing information on biodiversity and other natural resources.

Vegetation mapping in Europe has a long tradition (reviewed by Küchler, 1963; Bohn *et al.*, 2000–2004); its development has been spurred by progress in vegetation classification. Despite political differences between countries East and West of the infamous Iron Curtain, an extensive group of vegetation scientists form almost all European countries, under the leadership of Trautmann W, Ozenda P (France), Lavrenko EM (then Soviet Union), and later joined by Neuhäusl R (then Czechoslovakia) and Bohn U (Germany), embarked in 1975 on a project to collate the existing vegetation map, support construction, and publish new ones. As a result of this unprecedented cooperation, first, a comprehensive and scientifically sound vegetation map of Europe was published 25 years after the project was conceived (Bohn *et al.*, 2000–2004). The interpretational manual to the Map (so far available only in German) contains not only a description of all more than 700 mapping units, but also the most exhaustive list of vegetation maps ever published in Europe. The publication of this Map marks an important methodical transition marked by the advent of the GIS methodology and later increased availability (and accessibility) of satellite imagery aiding the production of modern

Table 4 Major monographs and accounts of diversity of vegetation types and ecosystems in Europe

Region	Year	Source	Comment
Pan-European vegetation surveys			
Europe	1852	Henfrey (1852)	A classical narrative about vegetation of selected regions of Europe
	1933	Various authors (1933–1940)	Unfinished (seven issues published) series of syntaxonomic monographs; edited by Braun-Blanquet J
	1962	Lohmeyer *et al.* (1962)	List of high-ranked syntaxa of Europe
	1973	Various authors (1973–1981)	Unfinished series of syntaxonomic monographs; edited by Tüxen R
	1976	Ernst (1976)	Syntaxonomic survey of vegetation of heavy-metal substrates
	1977	Various authors (1977–1999)	Accounts on major ecosystems of the World; many review papers feature the European ecosystems; edited by Goodall D
	1984	Mayer (1984)	An overview of the forest vegetation of Europe
	1985	Polunin and Walters (1985–1992)	Popular account of major biomes and ecosystems of Europe (1992: Italian edition)
	1997	Mucina (1997)	Conspectus of vegetation classes of Europe and Canary Islands
	1998	Rodwell *et al.* (1998)	Syntaxonomic calibration of the EUNIS units; including survey of high-rank syntaxa of Europe
	2000	Bohn *et al.* (2000–2004)	First comprehensive vegetation map of Europe; 2000: legend, 2003: text and maps, 2004: interactive CD
	2001	FAO (2001)	Map of vegetation zones based on the Map of Natural Vegetation of Europe (Bohn *et al.*, 2000–2004) accompanied by brief descriptions of the zones
	2002	Rodwell *et al.* (2002)	First published comprehensive survey of high-rank syntaxa of Europe; containing a crosswalk to habitats units of EUNIS
	2010	Klötzli *et al.* (2010)	A milestone monographic account of vegetation of Europe
Transnational vegetation surveys			
Alps	1928	Schröter (1928)	First comprehensive description of the vegetation of the Alps
Western Europe	1939	Tansley (1939)	Monographic vegetation account of the British isles (including UK and Ireland)
Central Europe	1943	Braun-Blanquet and Tüxen (1943)	First brief account of high-rank syntaxa of Central Europe
Mediterranean	1943	Rikli (1943–1946)	Classical three-volume opus featuring vegetation of the mediterranean-type climate regions of the world (including the European Mediterranean)
Central Europe	1944	Klika and Hadač (1944)	Annotated checklist of syntaxa of Central Europe
Central Europe	1957	Soó (1957–1963)	Series of papers featuring syntaxa of vegetation of the Pannonian Basin
Central Europe	1963	Ellenberg (1963–1998)	Series of editions of a famous vegetation monograph of Central Europe (1988): English edition; see also Ellenberg and Leuschner (2010)
Central Europe	1967	Hartmann and Jahn (1967)	Account of mountain forest vegetation of Central Europe north of the Alps
Central Europe	1974	Hartmann (1974)	Account of forest vegetation of Central Europe
Eastern Europe	1974	Walter (1974)	A large-scale monograph of vegetation of northern Eurasia, with an account of vegetation of European part of Russia and other states of the former Soviet Union
SE Europe	1974	Horvat *et al.* (1974)	Detailed account of vegetation of SE Europe and the Balkans
Alps	1983	Ozenda (1983)	Vegetation map of the Alps
Central Europe	1986	von Hübschman (1986)	Survey of European cryptogamic syntaxa
Western Europe	1994	Ozenda (1994)	Account of vegetation of Europe (with emphasize on Western Europe)
Alps	1995	Theurillat *et al.* (1995)	Survey of high-rank syntaxa of the Alps
Northern Europe	1996	Dierßen (1996)	Detailed authoritative account of the vegetation of northern Europe
Mediterranean	1998	Rivas-Martínez *et al.* (1998)	List of syntaxa of Spain (including Canary Islands) and the continental Portugal
Mediterranean	1998	Rivas-Martínez *et al.* (1998–2002)	Series of papers featuring syntaxa of vegetation of the Iberian Peninsula, Balearic Islands and Canary Islands; also available on a website
Europe	2000	Ozenda and Borel (2000)	Ecological map of Europe in paper describing the mapping rationale
Central Europe	2001	Pott (2001)	Account of ecosystems and vegetation of Central Europe
Northern Europe	2003	CAVM Team (2003)	Wall map of the circumpolar arctic regions of the World (including Europe); data available on a website
Europe	2004	Rivas-Martínez *et al.* (2004)	Map of thermoclimatic belts of Europe
Northern Europe	2005	Walker *et al.* (2005)	Paper featuring the methods, products, and use of the Circumpolar Arctic Vegetation Map

(*Continued*)

Table 4 Continued

Region	*Year*	*Source*	*Comment*
Central Europe	2010	Ellenberg and Leuschner (2010)	Updated version (by Leuschner C) of the legendary Ellenberg's account of vegetation of Central Europe
National vegetation surveys			
Albania	2002	Dring *et al.* (2002)	Syntaxonomic overview of Albania
Andorra	2003	Carreras *et al.* (2003)	Manual to the map of habitats (and vegetation types) of Andorra
Austria	1985	Wagner (1985)	Brief description of vegetation of Austria (memoir to a vegetation map)
	1993	Mucina *et al.* (1993)	Three-volume describing the syntaxonomic units of Austria
	2007	Willner and Grabherr (2007)	Vegetation survey of forests and related shrublands in Austria
Bulgaria	1997	Apostolova and Slavova (1997)	Account of vegetation units described between 1891 and 1995 in Bulgaria
	2009	Tsonev *et al.* (2009)	List of syntaxa of Bulgaria
Croatia	2008	Trinajstić (2008)	Account of vegetation of Croatia
Czech Republic	1998	Moravec (1998–2003)	Four volumes of the Czech Vegetation Survey (all featuring forest vegetation)
	1998	Neuhäuslová *et al.* (1998)	Map of potential natural vegetation; accompanied by text manual (Czech version)
	2001	Neuhäuslová *et al.* (2001)	Map of potential natural vegetation; accompanied by text manual (English version)
	2003	Chytrý and Tichý (2003)	Formalized diagnostic species calibration of high-rank syntaxa of the Czech Republic
	2007	Chytrý (2007–2011)	New series (so far three published volumes) featuring detail accounts of plant communities of the Czech Republic
Denmark	2004	Lawesson (2004)	First annotated checklist of Danish syntaxa
Estonia	1965	Laasimer (1965)	Account of vegetation of Estonia
	1997	Paal (1997)	Account of vegetation of Estonia
Former Czechoslovakia	1968	Mikyška *et al.* (1968)	Map of reconstructed natural vegetation of then Czech Lands (now Czech Republic); accompanied by text manual
	1983	Moravec *et al.* (1983–1995)	Two editions of a list of plant communities, featuring their conservation status of today's Czech Republic
	1984	Mucina and Maglocký (1984)	Checklist of high-rank syntaxa of Slovakia
	1985	Mucina and Maglocký (1985)	Checklist of syntaxa of Slovakia
	1988	Michalko *et al.* (1988)	Text manual and series of maps of the reconstructed natural vegetation of Slovakia; Slovak and English versions of the manual published separately in the same year
Former GDR	1972	Various authors (1972–1982)	Series of papers featuring survey of plant communities of southern parts of East Germany (then GDR)
	1996	Passarge (1996–2002)	Series of three monographs featuring vegetation of NE Germany (part of then GDR)
	2001	Schubert *et al.* (2001)	Identification key to plant communities of central and northern Germany (area roughly corresponding to the former GDR)
Former Soviet Union	1942	Walter (1942)	Brief account of zonal vegetation of the European part of then Soviet Union (+map of vegetation zones)
	1980	Gribova *et al.* (1980)	Account of vegetation of the European part of the former Soviet Union, accompanying a vegetation map
	1991	Korotkov *et al.* (1991)	Annotated checklist of vegetation units of the former USSR (including the European part)
	1997	Solomeshch *et al.* (1997)	Syntaxonomic survey of high-rank syntaxa of the former Soviet Union (including the European part)
Former Yugoslavia	1978	Lakušič *et al.* (1978)	List of syntaxa of Bosnia and Herzegovina (then part of Federal Republic of Yugoslavia)
	1986	Zupančić (1986)	List of syntaxa of the former Yugoslavia
	1986	Jovanović *et al.* (1986a–c)	Vegetation map of the former Yugoslavia; accompanied by text manual
	1994	Rexhepi (1994)	Vegetation account of Kosovo region (then part of former Yugoslavia, then Serbia, and today recognized by some countries as an independent state)
France	1947	Braun-Blanquet (1947)	Survey of high-rank syntaxonomic units of France for mapping purposes
	1952	Braun-Blanquet *et al.* (1952)	Account of the high-rank syntaxa of the Mediterranean France
	1993	Julve (1993)	Annotated checklist of high-rank vegetation units of France
	2004	Bardat *et al.* (2004)	Annonated checklist of syntaxa; also available on a website

(*Continued*)

Table 4 Continued

Region	*Year*	*Source*	*Comment*
Germany	1937	Tüxen (1937–1979)	First and the second edition of the Tüxen's classical regional vegetation survey (NW Germany)
	1955	Tüxen (1955)	Overview of vegetation types of NW Germany
	1957	Oberdorfer (1957–1993)	Legendary multivolume opus of syntaxa of southern Germany (then part of West Germany; BRD); most of the volumes are in second edition; only the forest volume is in the third edition
	1992	Pott (1992–1995)	A survey of major vegetation types of Germany; two editions
	1995	Schubert *et al.* (1995)	Identification key to plant communities of Germany
	1996	Various authors (1996–2001)	Series of syntaxonomic surveys of vegetation units of Germany; edited by Dierschke H; so far eight volumes published
	2001	Berg *et al.* (2001–2004)	A detailed vegetation survey of the German federal state of Mecklenburg-Vorpommern; data volume and text volume published separately
	2002	Rennwald (2002)	Annotated list of plant communities of Germany, featuring their conservation status
Hungary	1964	Soó (1964–1980)	Syntaxonomic accounts embedded within the monumental six-volume series on Hungarian flora and vegetation; vol. I (1964) is devoted exclusively to vegetation systematics; vol. III (1968) contains a syntaxonomic checklist and addenda
	1971	Soó (1971)	Checklist of vegetation units described from Hungary
	1986	Borhidi (1986)	Annotated checklist of syntaxa of Hungarian vegetation
	1996	Borhidi (1996)	Survey of Hungarian vegetation units
Iceland	1985	Glawion (1985)	Vegetation account of Iceland accompanied by vegetation maps
Ireland	1982	White and Doyle (1982)	Catalogue of vegetation units of the Republic of Ireland
Italy	2001	Various authors (2001–2007)	Two special issues of the journal *Fitosociologia* featuring cross-referenced regional lists of syntaxa described from Italy
Latvia	2001	Jermacāne and Laiviņš (2001)	List of syntaxa of Latvia
Liechtenstein	1998	Schmider and Burnand (1998)	Survey of forest plant communities of Liechtenstein
Lithuania	1997	Rašomavičius (1997)	First volume (meadows) of national survey of Lithuania
	1998	Balavicienē *et al.* (1998)	Account of vegetation of Lithuania
Malta	1993	Schembri (1993)	Brief account of vegetation of Maltese Islands
Norway	1983	Vevle (1983)	Brief survey of higher syntaxa of Norway
	1991	Fremstad and Elven (1991)	Account of the vegetation mapping units of Norway
	1997	Fremstad (1997)	Monographic survey of major vegetation types of Norway
Poland	1959	Szafer (1959–1966)	Monographic survey of vegetation of Poland in two volumes (1966: English edition)
	1981	Matuszkiewicz (1981–1988)	Key to identification of vegetation units of Poland (several editions)
	1984	Matuszkiewicz (1984)	Map of potential natural vegetation of Poland; accompanied by text manual
	2001	Matuszkiewicz (2001)	Survey of forest plant communities of Poland
Romania	1991	Coldea (1991)	Survey of plant communities of the Romanian sections of the Carpathians
	1992	Doniţă *et al.* (1992)	Vegetation map and manual to the mapping units of Romania
	1997	Coldea (1997)	First volume of national vegetation survey of Romania
	2002	Sanda (2002)	Survey of vegetation units of Romania
	2008	Sanda *et al.* (2008)	Survey of vegetation units of Romania
San Marino	2004	Biondi and Vagge (2004)	Syntaxonomic survey and description of major vegetation types of the Serenissima Reppublica di San Marino
Serbia	1998	Kojić *et al.* (1998)	Syntaxonomic survey of vegetation of Serbia
Slovakia	1996	Various authors (1996 *et seq.*)	Series of detailed syntaxonomic accounts of the vegetation of Slovakia; edited by Valachovič M; so far four volumes published
	2008	Janišová (2008)	Expert system (CD and manual) to grassland vegetation of Slovakia
	2008	Jarolímek and Šibík (2008)	Formalized diagnostic species calibration of high-rank syntaxa of Slovakia
Slovenia	2002	Marinček and Čarni (2002)	Vegetation map at scale 1:400,000 of the forest vegetation units of Slovenia; accompanied by text manual
	2006	Marinček *et al.* (2006 *et seq.*)	Series of electronic maps (on CD) and brief manuals to forest vegetation of Slovenia; so far three sheets published
Spain	1987	Peinado Lorca and Rivas-Martínez (1987)	Brief regional account of the vegetation of Spain
	1987	Rivas-Martínez (1987)	Manual describing vegetation mapping units of Spain

(*Continued*)

Table 4 Continued

Region	*Year*	*Source*	*Comment*
	2006	Del Arco (2006)	Vegetation map of the Canary Islands; accompanied by manual and CD
Switzerland	1933	Rübel (1933)	First survey of vegetation units of Switzerland
	1972	Ellenberg and Klötzli (1972)	Detailed survey of forest plant communities of Switzerland
	2008	Delarze and Gonseth (2008)	A phytosociological overview at alliance level of Switzerland
The Netherlands	1946	Westhoff *et al.* (1946)	Survey of plant communities of the Netherlands
	1969	Westhoff and Den Held (1969)	Classical systematic account of vegetation units of the Netherlands
	1995	Schaminée *et al.* (1995–1999)	Monumental completed series describing the plant communities of the Netherlands in five volumes
	2000	Weeda *et al.* (2000–2005)	Series of four volumes featuring grid-based distribution of plant communities of the Netherlands
Ukraine	1996	Solomakha (1996–2008)	Annotated account of vegetation units (in three editions) of Ukraine
	2000	Various authors (2000 *et seq.*)	Series of monographic accounts of the vegetation types of Ukraine (so far four volumes published: high-altitude vegetation, aquatic vegetation, halophytic vegetation, meadows)
United Kingdom	1949	Tansley (1949)	Monograph of the vegetation of the British Isles
	1964	Burnett (1964)	Monograph of the vegetation of Scotland
	1991	Rodwell (1991–2000)	Authoritative, detail account of plant communities of United Kingdom (excl. Northern Ireland) in five volumes

A complete bibliography of the featured surveys and checklists can be requested from the author.

vegetation maps. The consistent conceptual approach of the zonality/azonality approach applied by the Map facilitated a definition of the zonobiomes of Europe in great detail (**Figure 1**).

The European Vegetation Map has not been the last word in vegetation mapping in Europe. More fine scale projects were conceived and slowly chip away in several countries (e.g., in Czech Republic, Catalonia, Slovenia; see (**Table 5**). An international team of the Circumpolar Arctic Vegetation Map produced goods in 2003 and mapped the vegetation of the Arctic tundra in Eurasia and North America. A team of Circumboreal Vegetation Map (see http://www.iavs.org) is underway to write a new article of the Nordic vegetation mapping.

Diversity of Habitat Types

Although vegetation forms, at least in the terrestrial ecosystems, the major component of these ecosystems, the habitat typology cannot and does not reflect only vegetation typology. For instance, as found by Rodwell *et al.* (1998), 60% of the EUNIS3 habitat classification were characterized in phytosociological (syntaxonomic) terms, whereas the other units are largely abiotic and carry geomorphologic and hydrological features or they are associated with fauna or nonvascular plants.

The judgment on the diversity of habitats appears more difficult than that based on the classification of vegetation. First, there are many incommensurable variables to be considered by the definition of a habitat type. Second, the spatial and temporal scaling of the habitat types is very complex. This complexity is reflected in all habitats with habitat complexes recognizable on the classification systems of habitats know until today. To cope with the complexity, the habitat classification systems are hierarchical, but often on the same level of hierarchy, thus mixing small-scale habitats with habitat complexes recognizable on the landscape level. Obviously, the calibration of the habitat units and the classification system leaves much room for improvement (Rodwell *et al.*, 2002).

Outlook

In Europe, as elsewhere, taxonomic diversity is the leading topic of biodiversity research and conservation. Yet, because of deeply rooted research traditions and the awareness of national governments, surveys on higher levels of biological complexity received considerable attention in the past.

Almost every European nation has national vegetation or land-use maps. Almost all European countries have developed at least a checklist of ecosystems, habitats, or vegetation types and detailed national vegetation surveys have been recently completed or are under way in many crucial European countries (see Tables 4 and 5). The surveys of biotic communities, ecosystems, landscapes, and biomes have been recognized as important tools of both biodiversity conservation and sustainable use of natural resources. With Europe growing together as a union of nations sharing political agendas and an economic future, these tools are becoming standardized, more sophisticated, and effective. Pan-European habitat classification scheme EUNIS3, Habitat Directive of the European Union, vegetation map of Europe (Bohn *et al.*, 2000–2004), unified checklist of the syntaxonomic terminology, and building of standardized vegetation databases (Schaminée *et al.*, 2009) are a few prominent examples. Technology-driven progress in the data handling, GIS-assisted vegetation mapping using automatic data collection (remote-sensing), and analysis is boosting the development of the standardized vegetation and ecosystem survey.

The fall of the Iron Curtain in 1989 and 1990 posed a new challenge for research agendas in ecosystem diversity – unification of scientific standards and application tools along the West–East gradient within Europe. The European Vegetation

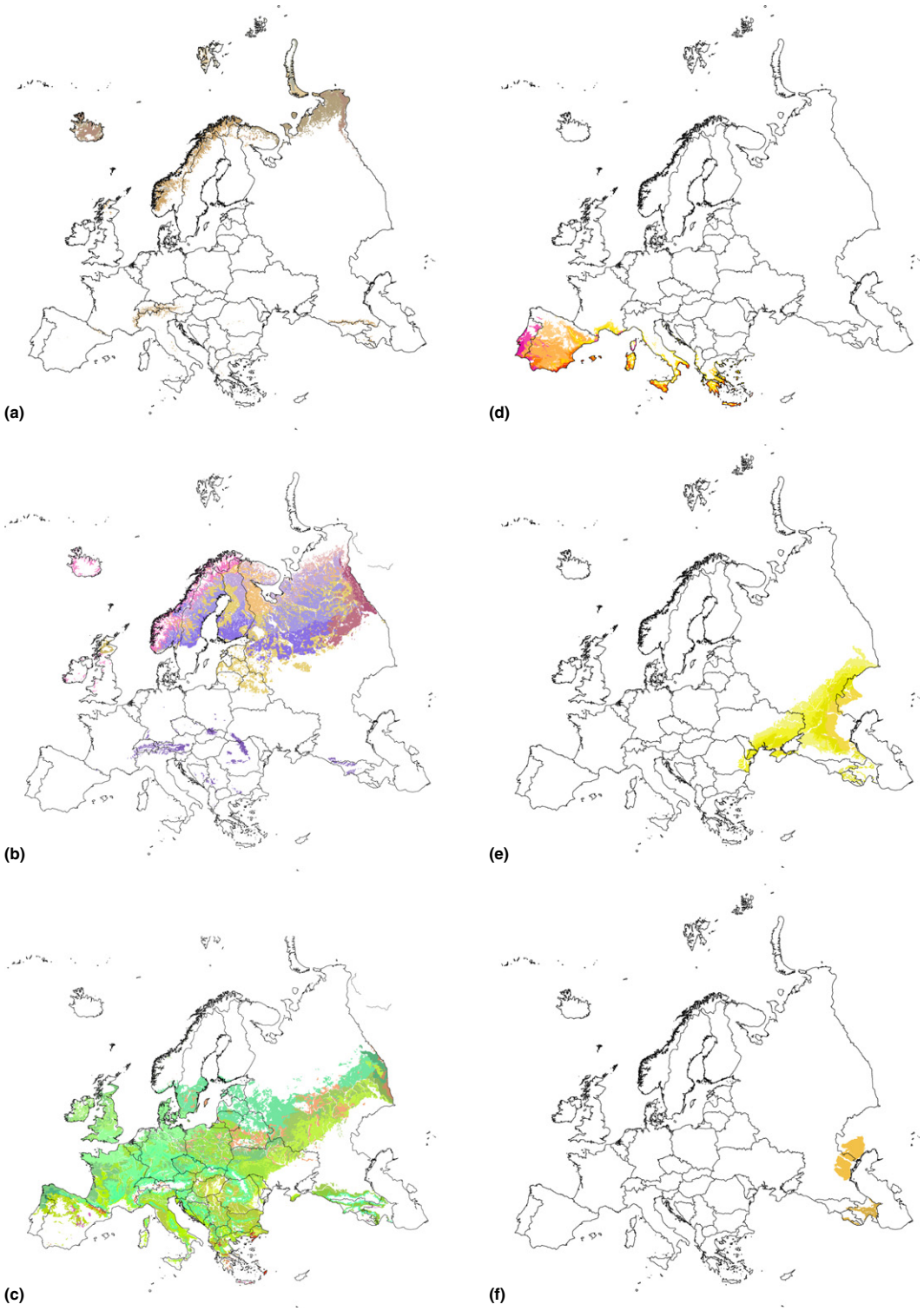


Figure 1 Biome maps of Europe derived from the interactive CD-ROM featuring the Map of the Natural Vegetation of Europe (Bohn *et al.*, 2000–2004) using the following key: (a) Polar Desert & Tundra biomes (Units A and B as defined and described in Bohn *et al.* (2000–2004)); (b): Boreal Forest biome (Units C1–2, D1–4, D5.1–5.2); (c) Temperate Deciduous Forest biome (Units D3, D5.3, F, G, K, L; hemiboreal forests units and the montane needle-wood forests of the Alps, Carpathians, and the Caucasus are included as well); (d): Mediterranean biome (Units J); (e) Steppe biome (Units M); and (f) Continental Desert biome (Unit D). The azonal vegetation (P, R, S, T, and U) are not shown. The color codes follow strictly those of Bohn *et al.* (2000–2004).

Table 5 Major vegetation-mapping sources published on and after 2000

Year	*Source*	*Region*	*Comment*
2000	Bohn *et al.* (2000–2004)	Europe	First comprehensive vegetation map of Europe (2000: legend, 2003: text and maps, 2004: interactive CD)
	Ozenda and Borel (2000)	Europe	Ecological map of Europe in paper describing the mapping rationale
2001	Neuhäuslová *et al.* (2001)	Czech Republic	Map of potential natural vegetation, with interpretation manual English version
	FAO (2001)	Europe	Map of vegetation zones based on the Vegetation Map of Europe (Bohn *et al.*, 2000–2004)
2002	Marinček and Čarni (2002)	Slovenia	Vegetation map on CD-ROM at scale 1:400,000 (and accompanying manual) of the forest vegetation units of Slovenia
2003	CAVM Team (2003)	Northern Europe	Poster-size map of the circumpolar arctic regions of the World (including Europe); data available on a website
	Centre de Biodiversitat (2003)	Andorra	Map of Andorran habitats on CD-ROM
2004	Rivas-Martínez *et al.* (2004)	Europe	Map of thermoclimatic belts of Europe
2006	Marinček *et al.* (2006 *et seq.*)	Slovenia	Series of electronic maps (on CD) and brief manuals to forest vegetation of Slovenia; so far three sheets published
	Del Arco (2006)	Spain	Vegetation map of the Canary Islands; accompanied by manual and CD

The full bibliographic references are available from the author on request. See Bohn *et al.* (2000–2004) for a comprehensive bibliography of vegetation mapping works published before 2000.

Survey (a working group of the International Association for Vegetation Science; see http://www.iavs.org) has been successfully tackling this task since 1991, and yet, a long road still lies ahead.

See also: Africa, Ecosystems of. Asia, Ecosystems of. Australia, Biodiversity of Ecosystems. Ecosystems of South America. North America, Patterns of Biodiversity in

References

Bohn U, Neuhäusl R, Gollub G, *et al.* (eds.) (2000–2004) *Karte der natürlichen Vegetation Europas/Map of the Natural Vegetation of Europe. Maßstab/Scale 1:2,500,000. Teil 1–3/Parts 1–3 and Interactive CD-ROM.* Münster: Landwirtschaftsverlag.

Braun-Blanquet J (1964) *Pflanzensoziologie. Grundzüge der Vegetationskunde*, 3rd edn. Wien: Springer.

CAVM Team (2003) *Circumpolar Arctic Vegetation Map. Scale 1:7,500,000. Conservation of Arctic Flora and Fauna (CAFF) Map No. 1.* Anchorage: U.S. Fish and Wildlife Service.

Commission of European Communities (1991) CORINE Biotopes. *The Design, Compilation and Use of an Inventory of Sites of Major Importance for Nature Conservation in the European Communities.* Luxembourg: Office for Official Publications of the European Communities.

Davies CE and Moss D (1998) *EUNIS Habitat Classification. Final Report.* Paris: European Topic Centre on Nature Conservation.

Devillers P and Devillers-Terschuren J (1996) *A Classification of Palearctic Habitats.* Strasbourg: Council of Europe.

Küchler AW (1963) Vegetation mapping in Europe. *Geographical Review* 43: 91–97.

Mucina L (1997a) Classification of vegetation: Past, present and future. *Journal of Vegetation Science* 8: 751–760.

Mucina L (1997b) Conspectus of classes of European vegetation. *Folia Geobotanica et Phytotaxonomica* 32: 117–172.

Rodwell JS, Dring J, Pignatti S, Schaminée J, and Mucina L (1998) *Scientific background to the EUNIS Habitat Classification. Phytosociological Relationships of EUNIS Habitats.* Paris: Agence Européenne de l'Environment, Centre Thématique Européen pour la Conservation de la Nature.

Rodwell JS, Schaminée JHJ, Mucina L, *et al.* (2002) *The Diversity of European Vegetation: An Overview of Phytosociological Alliances and their Relationships to EUNIS Habitats.* Wageningen: National Centre for Agriculture, Nature Management and Fisheries.

Schaminée JHJ, Hennekens SM, Chytrý M, and Rodwell JS (2009) Vegetation-plot data and databases in Europe: An overview. *Preslia* 81: 173–185.

Walter H and Breckle S-W (1991) *Ökologie der Erde. Band 1: Ökologische Grundlagen in globaler Sicht*, 2nd edn. Stuttgart: G. Fischer.

Westhoff V and van der Maarel E (1978) The Braun-Blanquet approach. In: Whittaker RH (ed.) *Classification of Plant Communities*, pp. 287–399. The Hague: Dr W. Junk.

Whittaker RH (ed.) (1978) *Classification of Plant Communities.* The Hague: Dr W. Junk.

Eutrophication and Oligotrophication

JoAnn M Burkholder, North Carolina State University, Raleigh, NC, USA
Patricia M Glibert, University of Maryland Center for Environmental Science, Cambridge, MD, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by JoAnn M. Burkholder, volume 2, pp 649–670, © 2001, Elsevier Inc.

Glossary

Algae (sing., alga; adj., algal) Primitive plant-like organisms that generally photosynthesize, here including cyanobacteria (sometimes called blue-green algae). These organisms can be unicellular, filamentous, or colonial microscopic forms (microalgae), or they can be macroscopic (macroalgae) consisting of a primitive plant body (thallus) that lacks vascular tissue in most species.

Anoxia (adj., anoxic) Condition in which there is negligible or no dissolved oxygen (DO) in the water (<0.1 mg of DO l^{-1}).

Biochemical oxygen demand (BOD) Measure of the DO required by the microbial community to decompose organic matter by aerobic metabolism.

Harmful algae Algae that are undesirable to humans because they: (1) become too abundant , or "bloom," in response to nutrient over-enrichment and then, at night, use most or all of the oxygen in the water for their respiration, so that the fish and other desirable organisms suffocate or become seriously physiologically stressed; (2) become too abundant in response to nutrient enrichment, and overgrow beds of desirable rooted vegetation so that the beneficial plants cannot receive enough light to survive; (3) cause or promote disease in desirable plants or animals; and/or (4) produce toxins that hurt or kill finfish, shellfish, and other higher trophic levels including humans. More recently, the term has been used to also include organisms that are not photosynthetic, primitive plantlike organisms – for example, certain nontoxic animal-like dinoflagellates, which cause fish disease (e.g., *Amyloodinium ocellatum*); and certain toxic animal-like dinoflagellates (e.g., toxic *Pfiesteria piscicida* and *Pfiesteria shumwayae* described by Burkholder and colleagues (2005), Marshall and colleagues (2006), and Moeller and colleagues (2007)), which do not have their own chloroplasts for photosynthesis, but which resemble plantlike toxic dinoflagellates in appearance and certain other general features. The term also commonly includes cyanobacterial (blue-green-algal) blooms (proliferations), despite the fact that in taxonomy they are now formally classified as bacteria rather than algae. They are included because cyanobacteria are primary producers and, so, function ecologically as photosynthetic algae.

Hypoxia (adj., hypoxic) Condition in which the water has depressed DO levels that generally are too low to sustain healthy populations of most fish species for extended periods (usually considered as >0.1 to <2 mg DO l^{-1}; sometimes <4 or <4.5 mg DO l^{-1}). Note that hypoxic levels of 3–4 mg DO l^{-1} can stress or kill sensitive egg and larval stages of some finfish and shellfish species, and that many fish actively avoid hypoxic areas.

Mixotrophy (adj., mixotrophic) Form of nutrition involving both phototrophic (photosynthetic) and heterotrophic carbon acquisition.

Nutrient (ecological or elemental) stoichiometry The study of changes in the relative proportions of critical nutrients available in the water, relative to differences in the allocation of these elements in organisms (*sensu* Sterner and colleagues). It may also refer to the relative availability of critical elements in ecosystems, as both an influence on and an overall result of interactions among biota and system biogeochemistry. Nutrient or ecological stoichiometry dictates that the relative availability of critical elements determines which of those elements most constrains the ecological dynamics (*sensu* McIntyre and Flecker). More specifically, nutrient stoichiometry compares nutrient ratios (especially N, P, carbon, and silica) in solution versus biomass.

Salt wedge Water from the ocean, with higher salt content, that moves into an estuary along the bottom of the water column, beneath less dense freshwater that has moved into the same area from a river. The salt content makes the ocean water heavier than the fresh (riverine) water, so that under calm conditions, the estuarine water becomes density-stratified. This "salt wedge" of bottom water can become somewhat isolated from the overlying fresh or less brackish water. The longer the period in which the total water column is not mixed by winds or storms, the more distinct the two water layers or strata. Salt wedges most often develop in warm seasons when plantlike phytoplankton production is high in the surface waters and respiration by heterotrophs (bacteria, fungi, protozoans, animals) is also high, especially in the lower water column and sediments. Nutrients typically are higher within the salt wedge than in the overlying water because of decomposition processes. At the same time, the bottom-water salt wedge can become hypoxic or even anoxic, while overlying waters can be saturated or supersaturated with oxygen from phytoplankton photosynthesis.

Stoichiotrophic imbalance A unique, forced trophic state in an aquatic ecosystem that develops when one nutrient, generally nitrogen (N) or phosphorus (P), but not both, is altered, either due to enrichment or management-related nutrient control. Stoichiotrophic imbalance can occur throughout the eutrophication process, depending on the enrichment sources or timing of N or P controls.

Trophic status Ranking system for aquatic ecosystems, based on the amount of organic production and nutrient (N, P) levels. Qualitatively, consists of three categories: Oligotrophic – characterized by relatively high phytoplankton species diversity but low phytoplankton

production, low nutrient concentrations and loadings, low decomposition (with little organic material available to decompose), and plentiful DO throughout the water column. Mesotrophic – characterized by moderate phytoplankton production and moderate nutrient concentrations and loadings. Eutrophic – characterized by relatively low phytoplankton species diversity but high phytoplankton production (often dominated by cyanobacteria in lakes, and by dinoflagellates or other flagellates in estuaries), high nutrient concentrations and loadings (especially of nitrogen (N) and phosphorus (P), high decomposition in the bottom water and surficial sediments (with abundant organic materials available for this process), and bottom-water DO deficits, sometimes with occasional to frequent fish kills. Eutrophic lakes typically are shallow with well-developed littoral zones (area where light penetration is sufficient to support growth of rooted plants), sometimes extending across most or all of the bottom area. Based on these classic definitions, the major component that is usually considered in assigning trophic status is phytoplankton, but this approach can be misleading. For example, some lakes are classified as oligotrophic because water-column nutrients and phytoplankton production are low, despite the fact that benthic primary production (e.g., of rooted angiosperms and their algal epiphytes) may be high. Quantitative rankings have also been developed, such as the Trophic State Index developed for lakes by Carlson (1977).

Introduction

Historic Overview Across Aquatic Habitats and Present Realities

From freshwater lakes to estuaries and marine coasts, human-derived sources of nutrient pollution have rapidly changed water quality and aquatic community structure within the past *ca*. 230 years of industrialization and rapid population growth (Cloern 2001, Wetzel 2001, Boesch 2002). The predominant theme throughout most of the world is increased nutrient enrichment or cultural eutrophication, rather than nutrient decreases or oligotrophication; thus, eutrophication is emphasized here although both processes are addressed.

Despite the advances in treatment of human sewage in some countries during the late twentieth century, bans on use of phosphorus in certain domestic or industrial practices, declines in agricultural fertilizers in some geographic regions, and modest improvements in environmental education in localized areas, the massive recent and ongoing increase in global human population growth has increased nutrient loadings to aquatic ecosystems (Figure 1). These increases are greatest in estuarine and coastal marine areas where population growth has been highest, and where nearly two-thirds of the people of the world now reside. As a reflection of this trend, estuaries have been reported to receive more nutrient inputs per unit surface area than any other type of aquatic ecosystem (Figure 2).

Cultural eutrophication promotes major shifts in the structure of both plant and animal communities, generally affecting dominant components of every trophic level from microbial decomposers to macrofauna. There is clear, compelling evidence of altered aquatic community structure and significantly reduced biodiversity from nutrient over-enrichment in many freshwater, estuarine, and marine ecosystems. Surface waters across the earth are now sustaining such impacts; even the open oceans are no longer sufficiently isolated to avoid nutrient pollution from atmospheric deposition. Similar trends have been demonstrated from the microfossil records of lakes and estuaries, showing a dramatic increase in nutrients and associated organic carbon deposits and a sharp, sustained decrease in the diversity of aquatic species.

Nutrients are essential for primary production by phytoplankton, benthic micro- and macroalgae, and aquatic

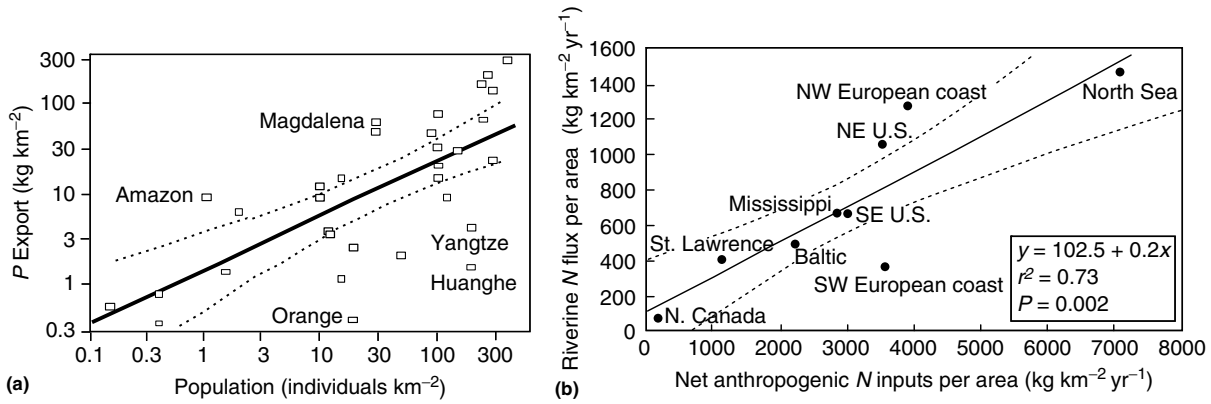


Figure 1 (a) The relationship between population density in watersheds and export of soluble reactive phosphorus (SRP) in river water, considering 32 major rivers (Caraco, 1995). (b) Export of total nitrogen from watersheds surrounding the North Atlantic Ocean, as a function of net anthropogenic inputs to their watersheds. Net anthropogenic inputs are defined as industrial N fertilizer + N fixation by legume crops + atmospheric inputs of oxidized N + net imports of N in food and feedstock. From Howarth *et al.* (1996), *Biogeochemistry* 35, 181–226 Figure 5A, with permission from Kluwer Academic Publishers.

angiosperms that directly or indirectly support aquatic food webs. In freshwaters, phosphorus historically has been least abundant among the nutrients needed in large quantity (macronutrients) by primary producers (Hecky and Kilham, 1988; Wetzel, 2001). Thus, it often has been the first element to become limiting to primary productivity in many freshwater systems, with nitrogen being secondary in importance

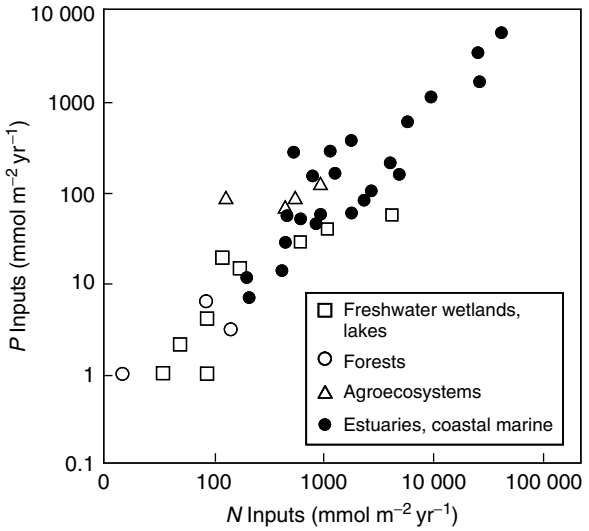


Figure 2 Phosphorus and nitrogen inputs to various types of ecosystems, showing the highest nutrient enrichment for estuaries and coastal waters. Reprinted with permission from Wassmann (2005) Eutrophication, primary production and vertical export, Figure 9.1. In: Wassmann P and Olli K (eds.) *Drainage Basin Inputs and Eutrophication: An Integrated Approach*, pp. 126–138. Norway: University of Tromso. Available at: http://lepo.it.da.ut.ee/~olli/eutr/Eutrophication.pdf

(**Figure 3**). Total phosphorus (TP) is typically considered in evaluating freshwaters for P limitation, rather than inorganic P (Pi – phosphate, PO_4^{-3}) which is the form most often used by algae. TP is used because algae commonly take up and store more P than they need in a process called luxury consumption. In natural systems or surface waters that are affected only by low anthropogenic inputs, PO_4^{-3} ions are rapidly taken up by algae and bacteria as fast as they are released, sometimes within seconds, so PO_4^{-3} concentrations typically appear to be low ($<5\ \mu g\ l^{-1}$) yet offer only a "glimpse" of the P actually available for algal growth. However, in areas that are strongly influenced by anthropogenic nutrient enrichment – such as sewage discharges from treatment plants with only primary or secondary treatment, or discharges from large phosphate mines – PO_4^{-3} concentrations can be 0.5–2 mg l^{-1} or more. Phosphorus also has been the most important nutrient-limiting plant growth in many tropical coastal marine waters and certain pelagic ocean waters (Downing, 1997; Burkholder *et al.*, 2007, and references therein). In temperate and polar coastal marine environments, nitrogen historically has been the most important nutrient that first limits primary production. Inorganic N (Ni), especially nitrate (NO_3^-) and ammonia (NH_3) and its ionized form, ammonium (NH_4^+), is directly taken up by phototrophs, although certain organic N forms such as urea can also be important N sources (Hecky and Kilham, 1988; Smith, 2006). Similarly as for Pi, in some cases Ni concentrations may appear to be low in the water column, but the total availability can be much higher due to rapid regeneration and biological recycling. At the estuarine interface between marine and freshwater habitats, both N and P can colimit plant production, especially in late winter-spring seasons of high precipitation and accompanying high inorganic N inputs. Other nutrients, notably silica and iron,

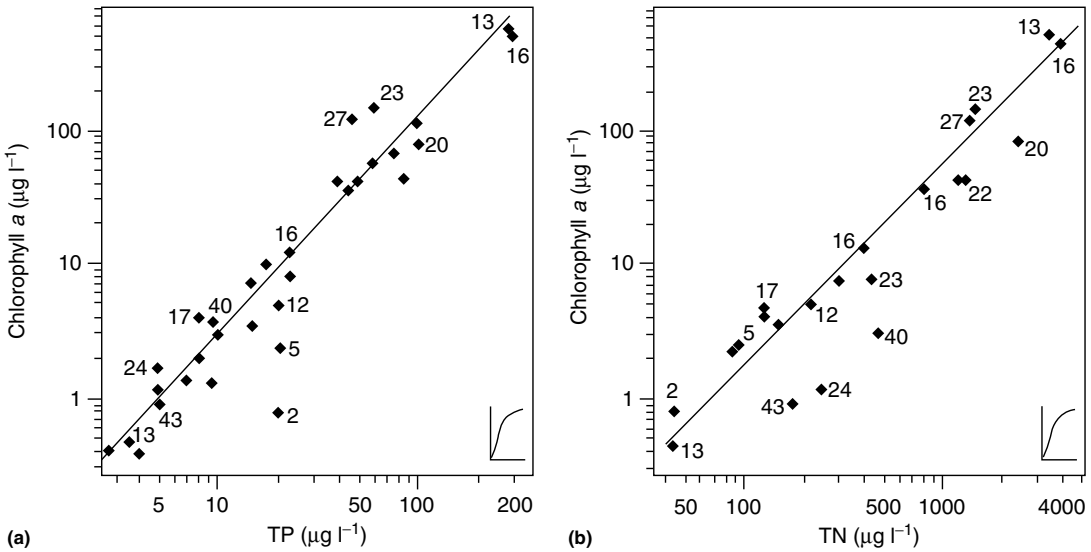


Figure 3 Relationship between summer maximum chlorophyll *a* content and (a) TP or (b) TN in surface and near-surface waters of some lakes in Japan. Numbers next to points represent the N : P ratio (by mass). Note that lakes with exceptionally low and exceptionally high N : P ratios are outliers in the TP-chlorophyll *a* and TN-chlorophyll *a* relationships, respectively. Similar log–log relationships between TP and chlorophyll *a*, and between TN and chlorophyll *a*, are characteristic of many natural lakes in the world that have low abiotic turbidity. Modified by Sakamoto (1966) in Kalff J (2002) *Limnology*, Prentice Hall, Upper Saddle River, NJ, with permission from Jacob Kalff and Pearson Education.

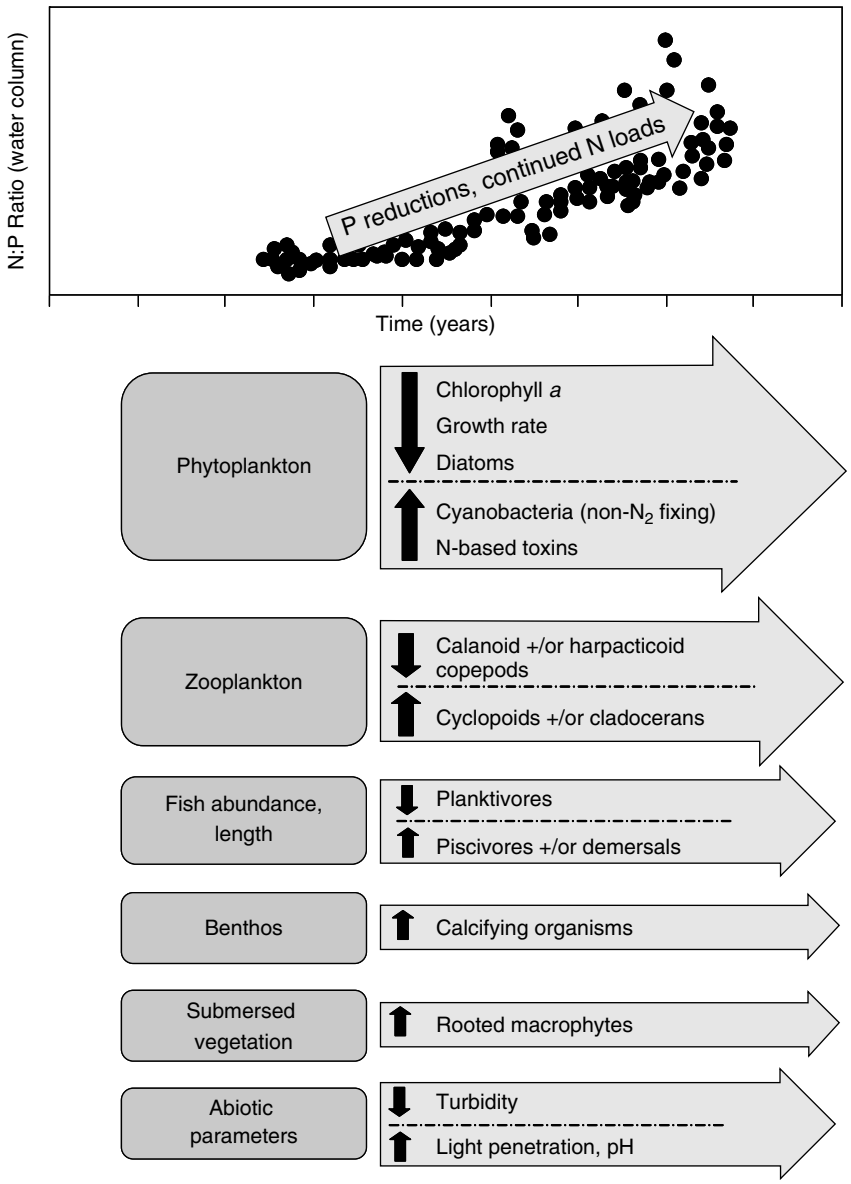


Figure 4 The pervasive influences of changes in ecological (nutrient) stoichiometry, showing expected changes in various components of a generalized aquatic food web as the water-column N : P ratio increases, based on ecological stoichiometry principles. The upper panel depicts the trend in water-column N : P ratios over time. Modified from Glibert *et al.* (2011) Long-term changes in nutrient loading and stoichiometry and their relationships with changes in the food web and dominant pelagic fish species in the San Francisco Estuary, California. *Reviews in Fisheries Science* 18: 358–417.

can sometimes significantly influence the outcome of species dominance and the structure and abundance of phytoplankton assemblages under cultural eutrophication.

Although scientists still evaluate surface waters in terms of N or P limitation in this way, it is important to note that these key nutrients are, in fact, no longer limiting in many freshwaters, estuaries, and coastal marine ecosystems. Instead, as Glibert *et al.* (2011) have pointed out, it is the *excess* of N and P loads that is at issue in present-day waters, not the lack of supply. The apparent strength of N limitation in marine waters, or of P limitation in freshwaters, is overestimated under conditions of P and N enrichment, respectively. Moreover, from an analysis across terrestrial and aquatic ecosystems, Elser *et al.* (2007) found that there tend to be strong positive synergistic effects of combined N and P enrichment; and that, because N and P supplies are closely balanced, single enrichments (of N but not P, or of P but not N) tend to rapidly promote limitation by the alternative nutrient. Nutrient or ecological stoichiometry relates changes in the relative composition of N and P in cells and tissues of aquatic organisms versus the water column. At the level of community dynamics and structure, changes in nutrient ratios, especially the relative proportion of N and P, alter metabolism, species composition, and food webs (**Figure 4**). Systems that are manipulated stoichiometrically by adding or reducing N or P, but not both nutrients, are neither eutrophic or oligotrophic. Application of traditional eutrophication indices will lead to an erroneous conclusion that such systems are not nutrient-stressed when,

in reality, they have been forced out of balance. They are in a state of stoichiotrophic imbalance.

In moderation, nutrients promote beneficial increases in phytoplankton and benthic algal production and, in turn, higher production of zooplankton, macroinvertebrates, shellfish, and finfish. But when added in excess, nutrient pollution can cause overgrowth of micro- and macroalgae, leading to dissolved oxygen (DO) depletion in bottom waters and sometimes even throughout the water column, and also causing alterations in food web dynamics. The classic symptoms of eutrophication are an increase in suspended algal chlorophyll (indicator of algal biomass), high-biomass algal blooms, a decline in submersed aquatic vegetation (SAV), and increased DO deficits and increased variation in DO over a diel (24-h) period. Schindler (1987) noted that among the earliest responses to anthropogenic stressors such as nutrient pollution is the loss of sensitive species. Although fish kills can be an obvious sign of acute impacts from cultural eutrophication, associated more subtle, chronic impacts are much more damaging to aquatic communities over the long term because they cause large-scale alterations in habitat and major shifts in aquatic food webs. Sustained loss in biodiversity is foremost among these chronic impacts, resulting from underlying mechanisms that scientists are only beginning to understand.

Factors Influencing Impacts from Nutrient Enrichment

Among the most important factors that influence the extent of adverse impacts from nutrient enrichment is water exchange rate, or the time required to completely replace the water in the system with new water. Waters that can self-cleanse or rapidly flush, such as fast-flowing rivers or run-of-river impoundments, are less sensitive to elevated nutrient loading than more contained or enclosed waters such as lakes and estuarine lagoons. Similarly, waters along open, wave-swept marine coasts are less sensitive to nutrient loading than quiet, sheltered upper embayments. However, such flushing may only displace the nutrient impacts by moving them downstream or to another area. Thus, countering the old adage that 'the solution to pollution is dilution', dilution is not the solution to eutrophication. The occurrence of other pollutants along with nutrients – for example, suspended solids that can reduce light for growth of aquatic plants, or toxic substances and harmful bacteria that can stress or kill some aquatic species – can exacerbate impacts from nutrient enrichment. The nutrient form can also be important. For example, swine wastes are much richer in organic nutrients (N, P, and C) than untreated human sewage. Various harmful algae prefer organic nutrient forms over inorganic sources.

The timing and frequency of the source inputs are important, as well. During high-precipitation periods in winter, sewage is more effectively diluted and cold temperatures discourage the growth of many nuisance algae. Also, in some waterways, agriculture is the major source of nutrients annually, but during low-flow periods in warmer seasons, sewage can contribute half or more of the river volume and the loading of nutrients – at a time when the readily available nutrients stimulate noxious algal overgrowth of the system.

The initial aquatic community structure also influences the overall effects of nutrient loads to aquatic ecosystems. Systems with low nutrient enrichment rely more on recycled nutrients than on introduction of new nutrient sources, and this may be a driving evolutionary force leading to greater specialization and diversity, as suggested by Howarth (1988). Diverse communities, characteristic of oligotrophic or nutrient-depauperate waters, tend to be dominated by sensitive species that thrive only within a narrow range of environmental conditions. Oligotrophic ecosystems are more sensitive (or less resistant), overall, to stress from nutrient enrichment. The most predictable outcomes are loss of sensitive species and increased abundance of generalist or opportunistic species that apparently are more resistant to the undesirable water quality changes and other stresses resulting from nutrient pollution. As an ecosystem becomes more eutrophic (that is, increasingly stressed by eutrophication), it becomes more resistant to further change.

Increasingly eutrophic lakes, lower rivers, and estuaries are characterized by seasonal low oxygen stress in the bottom waters, sometimes extending throughout half or more of the water column. Low oxygen stress can also occur within dense algal blooms near the water surface. Acute hypoxia or anoxia can result from allochthonous or external inputs of organic wastes, such as sewage discharges or runoff from animal feedlots, whereas chronic low-oxygen conditions can result from within-system (autochthonous) processes such as respiration by dense algal blooms or decomposition of dead remains of those blooms. Total organic carbon (TOC) is often used as a sum parameter for eutrophication (**Figure 5**); high organic content usually translates into an increase in the growth of microorganisms that will decompose the wastes using aerobic respiration, often causing oxygen depletion in the system as more and more TOC is produced because of nutrient overenrichment. In general, the higher the rate of biological metabolism (photosynthesis, respiration), the higher the TOC. Organic contaminants from agricultural runoff and domestic and industrial wastes also add TOC that contributes to the oxygen demand required by decomposers.

Hypoxia and anoxia develop when oxygen consumption exceeds supply, as a result of two major factors. In ecosystems with sufficient light to support their growth, the algae – often in densities of millions to billions of cells ml^{-1} of water – typically are the first to be stimulated by nutrient enrichment. They are net producers of oxygen through photosynthesis during the day, but at night they consume oxygen for respiration. Dense populations of algae in the upper and mid-depths of the water column can consume most of the DO. Low-oxygen stress in the system also occurs because of increased decomposition. As algal blooms senescence and die, and plants and animals die and settle out over the growing season, their remains are decomposed by bacteria and fungi in oxygen-demanding processes that can rapidly deplete the oxygen from the lower water column, sometimes extending to mid-depths or shallower. Oxygen solubility decreases with increasing temperature. Thus, low-oxygen stress tends to be most pronounced in warmer seasons when algal biomass, decomposition rates, and respiratory rates (and oxygen requirements) of fish and other animals are generally high.

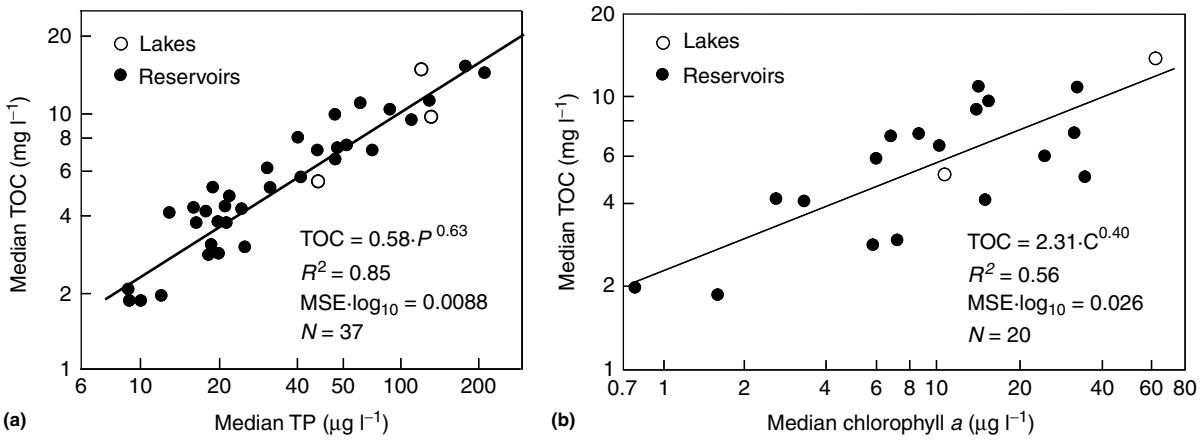


Figure 5 (a) Relationship between TP and TOC in some US reservoirs ($n = 34$) and natural lakes ($n = 3$). (b) Relationship between suspended algal biomass (as chlorophyll *a*) and TOC in some US reservoirs ($n = 18$) and natural lakes ($n = 2$). Modified from Figures 2 and 3 in Walker Jr. WW (1983) Significance of eutrophication in water supply reservoirs. *Journal of the American Water Works Association* 75: 38–42.

Low-oxygen conditions are exacerbated when bottom waters become somewhat isolated from the upper water column that receives oxygen during the day from phytoplankton growth, as well as throughout the diel cycle from the overlying air as it diffuses into the water. The isolating effect occurs when the water column becomes thermally (density) stratified, as in the bottom water layer (hypolimnion) of a stratified lake; or salinity (density) stratified, as in the lower water-column salt wedge in an estuary. As a result, sessile bottom-dwelling animals become physiologically stressed and may suffocate unless they can slow their metabolism until oxygen is replenished. Motile animals that rely on the bottom waters as critical nursery areas for their young, or a refuge area to escape predation, can physically avoid the low-oxygen areas but may nevertheless suffer population declines because of critical habitat loss. NH_3 that is released from decomposed organic matter, and its ionized form NH_4^+, can accumulate to toxic levels (*see* Aquatic Fauna and Inorganic Nitrogen Toxicity, below) in anoxic waters at depth, which further contributes to habitat degradation.

Organic matter in poorly treated sewage and animal waste contributes to low-oxygen stress in aquatic systems, as another source of material for decomposition or biological oxygen demand (BOD) (National Research Council, 1993; Mallin *et al.*, 2006). Chemical oxygen demand may further deplete oxygen, through oxidation of high levels of ammonia and other inorganic reduced compounds in the wastes. These oxygen-demanding processes often are relatively localized problems, for example, near sewage outfalls or animal waste spills, although they can affect the lower water column of eutrophic systems for months or even as a chronic condition. In addition, certain oxidized nutrients such as NO_3^- are highly soluble in water and, therefore, can be transported considerable distances as shown, for example, work by Mallin *et al.* (1993, 2005) or studies of NO_3^- from agricultural lands of the upper midwestern US contributing to an expansive low-oxygen "dead zone" in the Gulf of Mexico by Turner *et al.* (1998). Thus, transport creates a time lag between the introduction of nutrients into one area and adverse impacts at some distance from the source. Oxygen consumption from decomposition of excess phytoplankton production thus can occur on a large scale that is difficult to track or to relate to one specific, original source.

Changes in Flora and Fauna

Microalgae

Species Shifts Across Nutrient Gradients

Phytoplankton dominate the flora of oligotrophic systems. They respond quickly to nutrient inputs because the tiny plantlike organisms are immersed in the enriched medium, in contact with it over all surfaces. Phytoplankton with optimal growth at elevated nutrient concentrations are especially stimulated or selected for by nutrient enrichment, and they eventually overgrow and replace species that grow best at lower nutrient levels (Hecky and Kilham, 1988). Along a gradient from oligotrophic to highly nutrient-enriched, the phytoplankton assemblage structure gradually shifts from low abundance of many species and dominance by small flagellates and picoplankton (algae <2 µm in diameter) – with much of the energy flow channeled through a microbial loop of bacteria and small flagellates rather than directly up the food chain to herbivorous zooplankton – to high abundance of relatively few species consisting mainly of large cells or large colonies along with seasonally abundant flagellates. As lakes, lower rivers, and estuaries become more eutrophic, diatom (Heterokontophyta) species with higher N and P optima predominate in colder periods, certain colonial green algae (Chlorophyta) with high N optima become abundant in early summer, and dinoflagellates (Dinophyta: *Gymnodinium* spp., *Peridinium* spp.) and filamentous and colonial cyanobacteria with high P optima dominate in late summer. If dissolved silica, needed by diatoms to make their cell walls, is limiting in colder seasons, then flagellates (as examples, dinoflagellates, cryptomonads, euglenoids) become more abundant.

In shallow freshwater lakes, estuaries, and lagoons where open water habitat is limited, high phytoplankton densities can occur in response to nutrient enrichment (e.g., high

flagellate densities that historically occurred in response to excrement from duck farms in western Long Island Sound (Rhyther, 1989)). More commonly, benthic algae known as epiphytes (growing on the leaves of SAV), or floating drift macroalgae (in sheltered estuarine and marine coastal embayments), are rapidly stimulated by the nutrient increases and restrict the light needed by underlying plants (Figures 6 and 7). Increased nutrient enrichment generally promotes a decrease in species diversity of benthic algae, and a shift in dominance from diatoms to filamentous green algae or cyanobacteria. The planktonic and (more so) benthic habitats in such systems are characterized by a combination of low light and abundant food resources in the form of smaller organisms and rich dissolved organic substrates. These conditions also select for mixotrophic microalgae, including an array of harmful taxa that can become abundant and diverse under eutrophic conditions (Burkholder *et al.*, 2008; Jeong *et al.*, 2010).

Many regions of the world have landscapes dominated by run-of-river impoundment or reservoirs, rather than natural lakes. Similar trends in impacts from nutrient loading on aquatic communities can occur over time in reservoirs as in lakes, with the exception of depauperate rooted macrophyte populations (*see* Macrophytes, Aquatic Vascular Plants, below) in reservoirs with variable depth imposed by controlled drawdowns in potable water supply management or flood control. Impacts of cultural eutrophication in reservoirs can be mitigated by rapid water exchange, which generally occurs in weeks to months, rather than in years as for most natural lakes (Wetzel, 2001, and references therein). Moreover, the high turbidity from sediment loading or resuspension, characteristic of many reservoirs and estuaries as well as some natural lakes, often makes available light – rather than nutrients – the primary resource limiting the productivity of phytoplankton and other flora in the systems. Nutrients act as secondary factors controlling plant growth under such conditions, although they can become primary limiting factors if the water clears (for example, between storm events). Such systems often

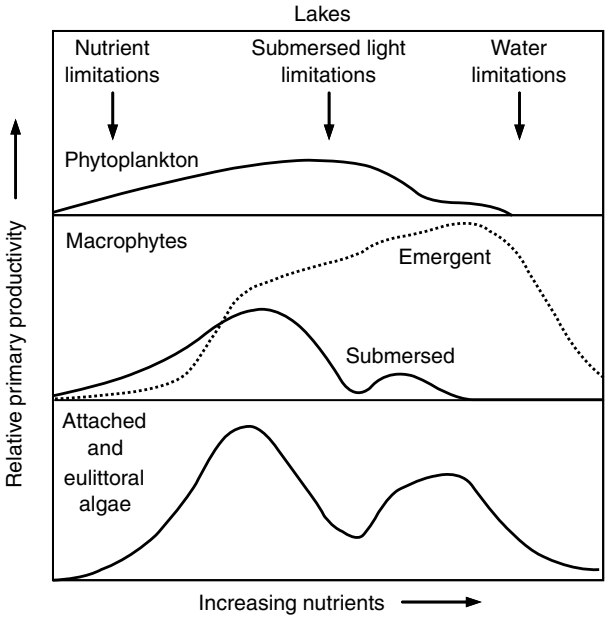


Figure 6 Generalized relationship of primary productivity of major plant groups in lakes under increasing nutrient enrichment. Nutrients are the primary resources limiting plant production under oligotrophic conditions, and phytoplankton (low in abundance and relative primary productivity, but high in species diversity) predominate over other phototrophs. As eutrophication progresses with increasing nutrient enrichment, light becomes more important than nutrients as the primary resource limiting primary production. In mesotrophic conditions prior to the onset of light limitation, submersed macrophytes can dominate, but they eventually are eliminated because of low light availability and emergent plants predominate. In late stages of the aging process, lakes become increasingly shallow and gradually function as wetlands. Plant production is then limited primarily by the availability of water. Reprinted with permission from Wetzel RG (2001) *Limnology*, 3rd edn., New York: Academic Press; originally from Wetzel RG (1979) The role of the littoral zone and detritus in lake metabolism. *Archives für Hydrobiologie Beiheft Ergebnisse der Limnologie* 13: 145–161.

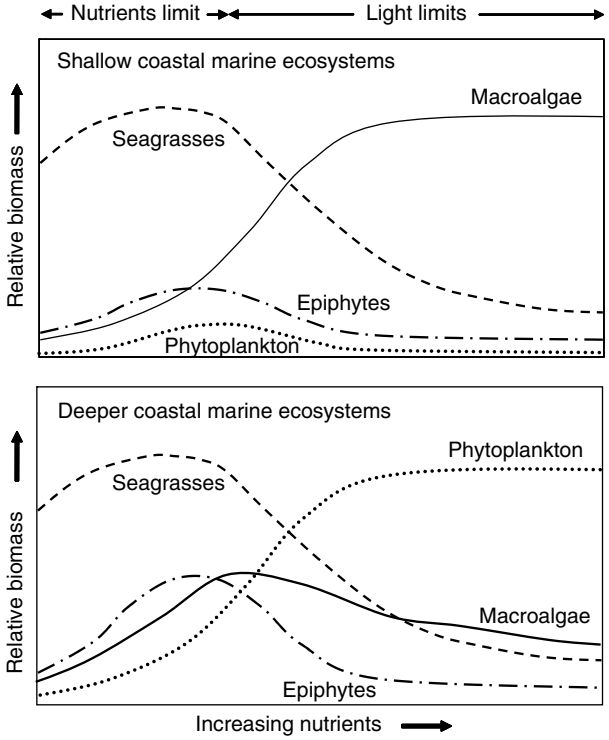


Figure 7 Generalized shift in biomass of major groups of primary producers with increasing nutrient enrichment to shallow and deeper coastal marine waters (upper and lower panels, respectively). In this generalized format, nutrients limit primary production under oligotrophic conditions. As eutrophication progresses with increasing nutrient enrichment, light becomes the primary limiting factor. Macroalgae (in shallow waters) or phytoplankton (in deeper waters) dramatically increase and become dominant, while seagrasses decline. Direct underlying mechanisms include competition for light and nutrients and also, for some species, NO_3^- inhibition or NH_4^+ toxicity. Reproduced from Burkholder et al. (2007) Journal of Experimental Marine Biology and Ecology 350: 46–72, Elsevier. (Note: The upper panel was Figure 3 in the previous version of this chapter. It was originally from Harlin (1993) Changes in major plant groups following nutrient enrichment. In: McComb (ed.) Eutrophic Shallow Estuaries and Lagoons, pp. 173–187. Boca Raton, CRC Press, Inc.)

can sustain higher nutrient (especially P) loading than clear, natural lakes while supporting less phytoplankton biomass, because of more rapid water exchange and light limitation from high suspended sediments. However, it must be underscored that in such cases, the nutrients that are not retained in the system (e.g., by settling out to the bottom sediments) will be transported downstream, where they can support increased production and lead to eutrophication impacts (e.g., Houser and Richardson, 2010; Jacobson *et al*., 2011).

"Bottom–up" control by nutrients interact with grazing pressure from zooplankton and other fauna in "top–down" trophic level effects (Carpenter, 1988; Burkholder *et al*., 2007, and references therein). Such effects are well documented in freshwater lakes and streams, and can also be operative in estuaries and coastal marine waters. Many field and laboratory studies have demonstrated that proliferation of phytoplankton and benthic microalgae under nutrient enrichment (e.g., seagrass epiphytes as described by Neckles *et al*., 1993) can be significantly reduced during periods when grazers are abundant. Thus, herbivores can sometimes alleviate eutrophication impacts by holding algal production in check. However, one effect of grazing pressure is to transform the form of nitrogen available to primary producers. Animals excrete nitrogen in the form of NH_4^+ (or, in some cases, urea), that is used by primary producers are in a different rate than is NO_3^-. Different algal species can be stimulated by NH_4^+ than by NO_3^-.

Decreased grazing pressure can allow higher algal biomass to develop in response to excess nutrients, but decreased grazing can help promote algal blooms only where nutrient inputs are sufficiently high to support such blooms. Therefore, grazing pressure – especially in present-day aquatic ecosystems – is generally regarded as a secondary factor controlling phytoplankton production under cultural eutrophication. Moreover, human hunting practices have caused severe reductions in what likely were once major influences of large herbivores in aquatic food webs, such as waterfowl across the salinity gradient, and green turtles (*Chelonia mydas*, sirenians) in marine food webs. Thus, in considering many present-day temperate regions in particular, Valentine and Duffy (2006) described large grazers mostly as "ghosts" because their once-major influence in many seagrass ecosystems has been reduced to such an extent that it is now "extremely episodic, negligible, or absent."

Long-Term Human Influences of Cultural Eutrophication

Long-term datasets have provided two lines of compelling evidence in support of major impacts from cultural eutrophication on phytoplankton assemblage structure in aquatic ecosystems, with potentially serious ecological and economic ramifications. First, the geological record in sediment cores taken from freshwater lakes and estuaries clearly shows that long-term major shifts have occurred under increased nutrient enrichment – progressing from a balance between species-rich planktonic (mostly centric) and benthic (mostly pennate) diatom assemblages, to dominance by planktonic diatoms with low species diversity (**Figure 8**). Planktonic diatoms tend to be selected because they are the first recipients of the increasing water-column nutrient sources. Benthic diatoms encounter decreased light because of the overlying plankton growth, eventually leading to decreased production.

The second line of evidence concerns species shifts in extant phytoplankton assemblages that have been related to changes in supply ratios of the two most limiting nutrients, again underscoring the importance of ecological stoichiometric considerations. Under nutrient-limiting conditions, species with similar optima in other environmental factors (e.g., temperature or light) that maintain faster growth compete more successfully for the available nutrient resource (Figures 9–11). Elegant work by Tilman (1977) and Tilman *et al*. (1982) examined the response of freshwater diatom species to shifts in Si : P ratios, and Rhee (1982) extended these concepts to controlling influences of N : P ratios. Low molecular N : P ratios (7–15 in those studies; up to *ca*. 29 : 1 by weight in Smith (1983)) favored tested cyanobacteria and diatom species, whereas higher N : P ratios favored green algae. Many green algae grow optimally at high N_i concentrations, whereas many cyanobacteria and dinoflagellates have a high P requirement. Diatoms, unlike the other algae, require major supplies of silica as well as N and P. When Si : N and Si : P ratios are high, the available Si favors growth of diatoms that can effectively compete for N and P resources. However, as Si : N and Si : P ratios decrease, silica can become limiting for diatom growth, and more N and P remain available for growth of flagellates and other algae that do not require silica. These concepts were extended to natural lakes by Tilman *et al*. (1982) (for diatoms and other algae), Smith (1983) (for cyanobacteria), and others.

Silica is only slowly (years) made available for new diatom growth through dissolution of dead diatoms and other natural silica sources, although it can also be affected by installing dams that can trap sediments and reduce silica availability. In contrast, N and P cycles are rapidly affected by anthropogenic inputs from sewage, animal wastes, atmospheric pollution, and other sources. As N and P enrichment increases, the Si : N and Si : P supply ratios are depressed. Eutrophication can effect a decrease in dissolved Si abundance by initially stimulating high growth of diatoms to the point that they reduce or deplete the dissolved silica pool for use by developing diatom populations in subsequent seasons. Long-term datasets on estuarine, coastal, and freshwater phytoplankton communities indicate that shifts to dominance by flagellated algae or cyanobacteria – including some harmful bloom-forming species – have coincided with decreased abundance of diatoms and decreased Si : N and Si : P ratios. Such trends have been documented by Schelske *et al*. (1986), Smayda (1989), Parsons *et al*. (2002), and others in temperate and sub-Arctic waters such as the Great Lakes, New England coastal waters, Chesapeake Bay estuarine waters, the Black Sea, and the coasts of northern Europe. In the Gulf of Mexico near the mouth of the Mississippi River, increasing N (nitrate)- to- Si ratios have been related to an increase in weakly silicified diatoms, including potentially toxic *Pseudo-nitzschia* species, which like other diatoms, are stimulated by nitrate pollution but they also require less silica for their cell walls relative to many other diatoms. In oligotrophic subtropical and tropical marine waters, mixotrophic dinoflagellates can predominate and, symbiotic interactions are common in both planktonic and benthic communities. As these systems undergo cultural

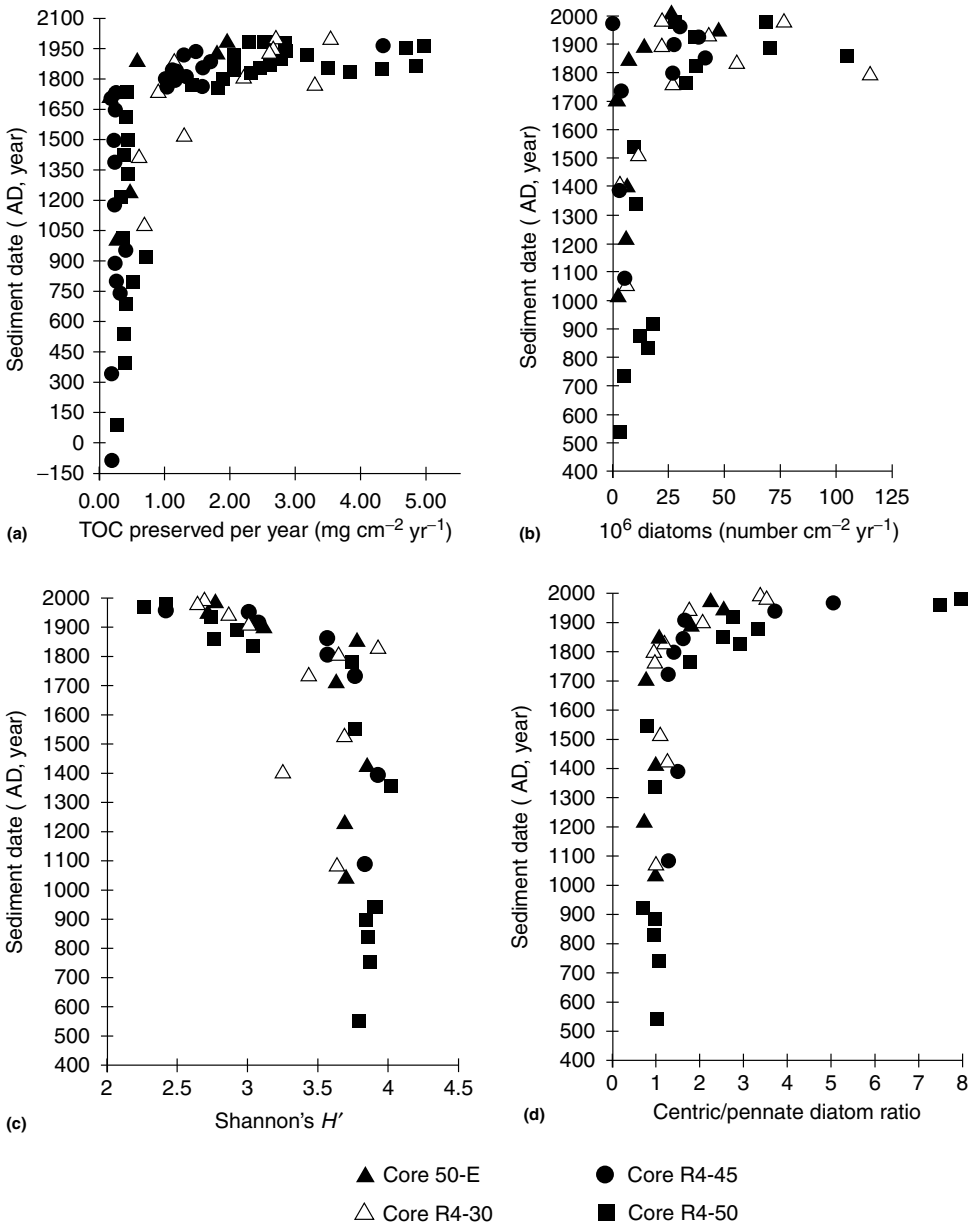


Figure 8 The geological record of cultural eutrophication in Chesapeake Bay, showing major changes since European settlement. The data from four sediment cores are graphed by the average date assigned to each sediment sample (depth layer) according to radiocarbon and pollen methods. (a) Total organic carbon (TOC, indicating total system productivity) preserved over time (historic record from AD 150 to AD 1990). (b) Diatom cell numbers per year (AD 400–AD 1990). (c) Diatom assemblage diversity calculated as Shannon's H (AD 400–AD 1990). (d) Centric/pennate diatom ratios (AD 400–AD 1990). TOC, diatom numbers, and the centric/pennate diatom ratios all showed a significant and abrupt increase following the time of European settlement in the 1700s. The total diatom community diversity, in contrast, significantly decreased post-1700s, relative to pre-1700s diversity. Graphs (a)–(d) were modified from Cooper SR © (1995) Chesapeake Bay watershed historical land use: impact on water quality and diatom communities. *Ecological Applications* 5: 703–723, with permission from American Association for the Advancement of Science.

eutrophication, small cyanobacteria become abundant; in parts of the Florida Keys, for example, the formerly clear water over coral reef areas has become an opaque emerald green from picoplanktonic cyanobacteria blooms. Trace metals such as iron have also been shown to be limiting to phytoplankton growth in some estuarine and marine waters where N and P are at levels that would otherwise be expected to support more algal production.

As eutrophication progresses, the previously described shift in temperate-zone phytoplankton assemblage structure from certain diatoms to other diatom species, flagellates, and cyanobacteria causes subtle but important, undesirable changes for the food web that can adversely affect secondary production. For example, as reviewed by Kilham *et al.* (1997), some diatom species produce high quantities of certain lipids that are essential for zooplankton reproduction. Algal species

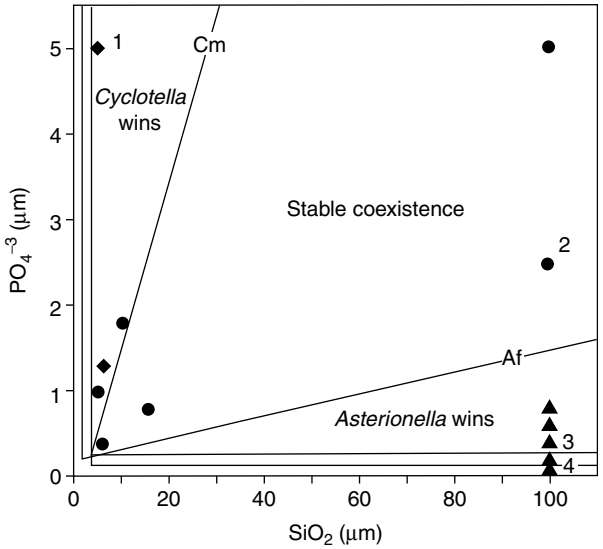


Figure 9 The outcome of competition between diatoms for limiting nutrients (silica or PO_4^{-3}) across a gradient of Si : P ratios, comparing two freshwater planktonic diatoms, *Cyclotella meneghiniana* and *Asterionella formosa*, under growth-limiting conditions (temperature, light, and other conditions held constant). These chemostat culture experiments indicate (a) conditions in which *A. formosa* is dominant because it can outgrow *C. meneghiniana* at very low P (as micromolar concentrations or micromoles per liter of phosphate, µM PO_4^{-3}), (b) conditions wherein the two species can coexist (circles) because one is limited by P, and the other is limited by Si (as silicon dioxide concentrations, µM SiO_2); and (c) conditions wherein *C. meneghiniana* can dominate because it can outgrow *A. formosa* at lower Si. Reprinted from Tilman D (1982) *Resource Competition and Community Structure*. Princeton: Princeton University Press, with permission from Annual Reviews.

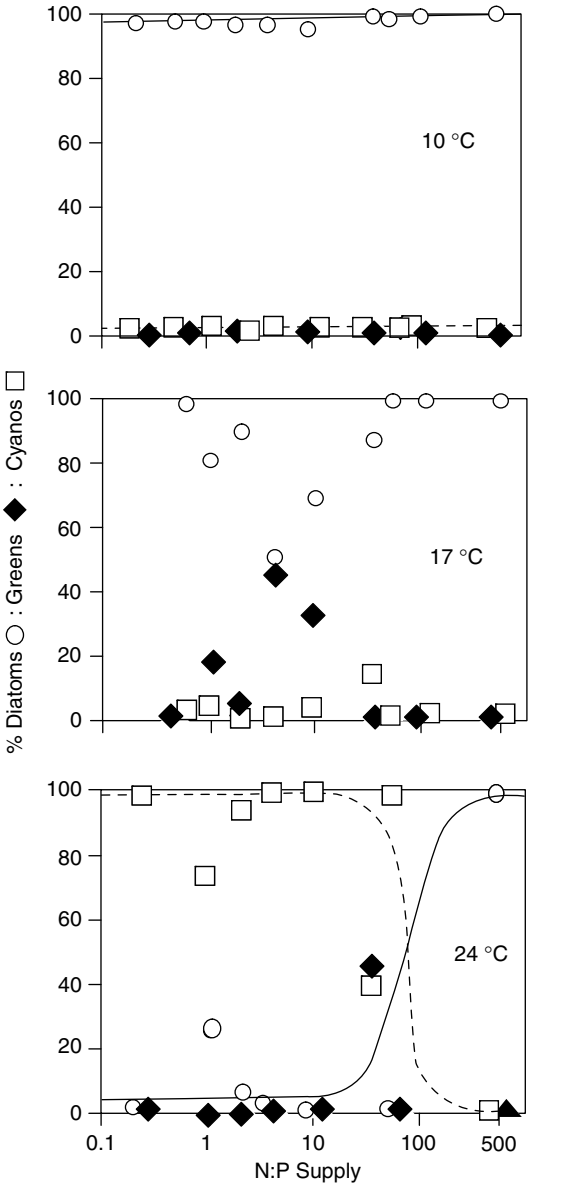


Figure 10 The outcome of competition among diatoms, green algae, and cyanobacteria (blue-green algae) for limiting nutrients (nitrogen or phosphorus) across a gradient of N : P ratios under silica-replete (nonlimiting) conditions, showing an additional influence of temperature. The comparison includes representative species of freshwater lake diatoms (white circles), green algae (black diamonds), and cyanobacteria (white squares) in chemostat culture under growth-limiting conditions for N and P (but not for Si). Note that at the coldest temperature indicative of conditions during spring and fall seasons in this north temperate lake, the cold-optimal diatoms outcompeted the greens and cyanobacteria across all N : P ratios. That is, temperature, rather than N or P, was the most important condition limiting the growth of green algae and cyanobacteria relative to diatoms. The mid-range temperature represented a transitional area where some growth could occur for species representing each group. The highest temperature favored warm-optimal cyanobacteria (which tend to have high P optima), especially at low N : P ratios when P was more abundant.

that may replace these diatoms under increasingly eutrophic conditions do not produce these lipids, or produce much less of them. Analogous phenomena occur in freshwaters, estuaries, and coastal environments. For example, Starr *et al.* (1990) reported that spawning of green sea urchins (*Strongylocentrotus droebachiensis*) and blue mussels (*Mytilus edulis*) apparently is triggered by a heat-stable metabolite that is released in high abundance by certain phytoplankton, especially certain diatoms such as *Skeletonema costatum* (Bacillariophyceae). This substance is not produced, or is produced in much lower quantities, by flagellate algae that replace these diatoms under cultural eutrophication. Different algal species can also affect zooplankton production and composition because they have different N : P stoichiometry which, in turn, affects the stoichiometry of the grazers.

Harmful Algal Blooms and Anthropogenic Nutrient Enrichment

Among the algal species that thrive in nutrient over-enriched or eutrophying waters are many noxious species such as cyanobacteria that are potentially toxic to zooplankton, fish, and wildlife in freshwaters and certain estuaries (Baltic Sea and Australia), and certain dinoflagellates and other flagellates that are potentially toxic to finfish and shellfish in estuarine and

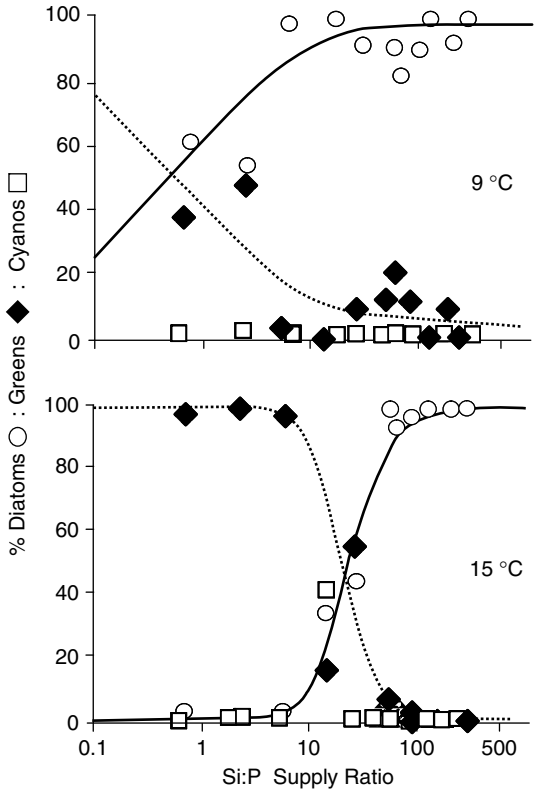


Figure 11 The outcome of competition among diatoms, green algae, and cyanobacteria (blue-green algae) for limiting nutrients (phosphorus or silica) across a gradient of P:Si ratios, showing the influence of temperature. The outcome of competition is indicated among freshwater lake diatoms (white circles), green algae (black diamonds), and cyanobacteria (white squares) in chemostat cultures under growth-limiting conditions for Si (needed only by the cold-optimal diatoms) and P at two temperatures. At the lower temperature, cold-optimal diatoms dominated across all Si : P ratios. However, at the higher temperature they were able to dominate over green algae only at high Si : P ratios, when their Si requirements were alleviated by the relatively high Si. The cyanobacteria species were able to outcompete the other groups only at one transitional Si : P ratio, even at the higher temperature which was still suboptimal for the cyanobacteria.

marine coastal areas worldwide (Burkholder 1998, 2009; Burkholder *et al.*, 2008). It has long been recognized that freshwater algae, most notably a wide array of noxious bloom-forming cyanobacteria species, are strongly stimulated by P and also by N pollution. In contrast, intensive study, for many decades, of a small number of harmful marine species that grow in oligotrophic waters largely resulted in a longstanding dogma that harmful estuarine and marine algae generally are not stimulated by nutrient pollution in estuarine and marine habitats. This dogma was finally laid to rest: a landmark paper published by Heisler *et al.* (2008) reported broad scientific consensus that many harmful estuarine and marine algae are indeed stimulated by N and P over-enrichment, and by changes in nutrient forms or their relative availability.

The fact remains that some species of harmful algae are found in oligotrophic habitats and historically have not been thought to be favored by elevated nutrients, and are believed instead to have optimal growth in low-nutrient conditions. Under low nutrient conditions they would be dependent on rapidly recycled nutrients, as the fundamental requirement of nutrients for cells is generally not different under low-nutrient conditions. Recent studies have provided evidence, however, that certain of these species can be directly or indirectly stimulated by nutrient enrichment. For example, the invasive diatom *Didymosphenia geminata* forms extensive benthic mats in light-replete oligotrophic streams, and its abundance has been negatively correlated with high N and P. However, a series of stream-side experiments conducted by Bothwell and Kilroy (2011) showed that addition of PO_4^{-3}, alone or together with NO_3^-, resulted in a prolonged increase in cell production; thus, the cells of this oligotrophic species divided faster under P enrichment.

As a second example, the potentially toxic dinoflagellate, *Karenia brevis*, causes ecologically and economically damaging "red tides" along the coasts of Florida and Texas, USA. For many years this organism was thought to be a strict phototroph, relying entirely on photosynthesis for its nutrition. Within the past decade, however, *K. brevis* has been found to act as a grazer of the small cyanobacterium *Synechococcus* (Jeong *et al.*, 2005), which is abundant in nutrient over-enriched coastal waters in the Gulf of Mexico. In fact, Glibert *et al.* (2009) reported that *K. brevis* consumed up to 84 *Synechococcus* cells per hour, fulfilling up to 40% of its cellular N requirement, especially under low light. These authors hypothesized that *Synechococcus* blooms may be stimulated by nutrient pollution, in turn providing an additional source of nutrients for enhanced growth of *K. brevis*, and for its red tides to be sustained by grazing the cyanobacterial prey.

Changing nutrient supplies and stoichiotrophic imbalance from altered nutrient supply ratios have resulted in major, diverse effects on harmful algae, exemplified here by the potentially toxic mixotroph *Prymnesium parvum* (Haptophyta). *P. parvum* blooms generally occur in eutrophic waters that are imbalanced in nutrient supplies and supply ratios relative to unenriched habitats. Cell production can be stimulated by PO_4^{-3} enrichment. In contrast, N or P limitation, as well as certain other stressors and suboptimal conditions, significantly enhance production and release of an array of toxins by this organism. When these toxins are released, they form micelles and apparently require activation by cofactors that can include other dissolved plant nutrients such as potassium, calcium, and magnesium. The nutrient stress connection operates as an after-effect for blooms that develop, in the first place, because of nutrient-rich conditions. The toxins of *P. parvum* include a mix of PO_4^{-3}-containing proteolipids, with hemolytic activity, which are thought to be precursors of structural membrane components, or products of imbalanced cell membrane metabolism. These toxins accumulate within cells during exponential growth phase, and are released into the surrounding medium during stationary phase when growth is limited or stressed. High toxin activity is hypothesized to occur under PO_4^{-3} limitation because formation of phospholipids for membrane synthesis is disrupted, resulting in leakage of cellular metabolites including toxins (Burkholder, 2009, and references therein). Research by Granéli and colleagues (e.g., Granéli and Johansson, 2003) has shown that nutrient stress, in the form of either inorganic P or

inorganic N limitation, significantly increases *P. parvum* toxin production or release.

As additional species of harmful algae are evaluated for their responses to increasing nutrient supplies from cultural eutrophication, it will be important to consider that the concentration of a nutrient at any given time may not be correlated with its actual bioavailability, because many phytoplankton can grow for up to several generations on internally stored "luxury-consumed" nutrient pools. Moreover, harmful algal blooms can sometimes develop in regions that are spatially displaced (e.g., downstream) from nutrient sources, partly because as nutrients move through aquatic systems, they are recycled and transformed to forms that are more readily available for algal use. This phenomenon has sometimes made it difficult to link blooms with the upstream nutrient sources. Major progress will continue to be made in understanding the complex direct and indirect influences of nutrient enrichment in harmful algal blooms by studying additional species to characterize their luxury consumption and subsequent use of inorganic N, P, and C under changing nutrient stoichiometry as their habitats become more eutrophic; their cell budgets for these nutrients; and their reliance on mixotrophy as well as photosynthesis to meet their nutritional needs.

Microbial Heterotrophs in Low-Light Environments

Algae and other phototrophs are common "first responders" to nutrient enrichment, and eutrophication has sometimes been defined as an increase in organic carbon to an aquatic ecosystem caused by primary productivity. This definition is too restrictive, however, because increased organic carbon from nutrient pollution is not always a result of primary productivity. The distinction is important because it means that nutrient pollution can detrimentally affect beneficial aquatic life through means other than increase algal growth. For example, from an analysis of blackwater streams in the Southeast, Mallin *et al*. (1993, 2005) reported that within-stream heterotrophic (bacterial) processes, stimulated primarily by P and secondarily by N enrichment, were the major influences on BOD. Blackwater systems generally do not provide sufficient light to enable algae to proliferate in response to nutrient enrichment. However, humic materials in the blackwater can alter nutrient biogeochemistry, generally sequestering nutrients but sometimes also making favored nutrients more available.

Macroalgae

Freshwater Communities

Macroalgae are the major flora of the littoral zones of some lakes, and of many shallow estuarine and marine waters, and they include some notorious 'aquatic weeds' that respond to cultural eutrophication. Temperate-zone freshwater and oligotrophic hardwater lakes may contain abundant benthic charophytes and, rarely, a few species of small brown macroalgae (Heterokontophyta). Oligotrophic softwater lakes can have populations of small, benthic red macroalgae (Rhodophyta), especially near spring-fed areas with bubbling carbon dioxide. As eutrophication progresses, populations of certain filamentous green algae and cyanobacteria that would otherwise be considered as microalgae develop visual biomass accumulations, especially under high P enrichment and also under N enrichment. For example, *Cladophora glomerata* (Chlorophyceae) forms masses of long hair-like growth, and cyanobacteria (e.g., *Anabaena, Aphanizomenon, Microcystis, Plectonema wollei*) form tufts or amorphous masses of greenish, bluish-green, or reddish gray slime.

Macroalgae are the dominant autotrophs of mid-order streams where hard substrata are available for colonization and light is moderate to high, especially after leaf fall in colder seasons. Stream-inhabiting macroalgae are often much more diverse and abundant than lake macroalgal floras. Nutrient-poor (often softwater) stream segments are colonized by diverse, abundant cold-optimal algae such as the diatom *Eunotia pectinalis*; the heterokontophyte *Tetrasporopsis*; chlorophytes such as *Spirogyra*, *Tetraspora*, and *Oedogonium*; red algae such as *Batrachospermum* and *Paralemanea*; and cyanobacterial mat formers like *Phormidium* spp. (for example, see Sheath and Burkholder (1985)). As eutrophication progresses, softwaters may develop larger populations of filamentous green algae such as certain species of *Oedogonium* and *Mougeotia*, or dense growth of colonial *Hydrodictyon* and *Stigeoclonium*. The latter taxon is also known to tolerate high metal concentrations that may be found in poorly treated sewage. Hardwater streams under high P enrichment, as well as the littoral zones of hardwater lentic systems, commonly become choked with massive growth of *Cladophora* spp. where light is sufficient for their proliferation (*see* Oligotrophication – Reversing the Impacts of Eutrophication? below, for information about the famous case of *Cladophora* in the lower Laurentian Great Lakes).

Estuarine and Marine Communities

Macroalgae also dominate the flora of many shallow estuaries, lagoons, and upper embayments, coral reefs, and rocky intertidal/subtidal habitats. Nutrient enrichment leads to reduction in the diversity of macroalgae and associated fauna. Opportunistic species of green algae within the genera *Ulva* (also found in saltwater lakes such as the Great Salt Lake of Utah) and *Cladophora*, and within the brown algal genus *Ectocarpus*, have rapid growth rates and proliferate because they can more quickly take advantage of the elevated nutrient levels and shade out other species. Repeated, pulsed nutrient enrichment to temporarily elevated levels – characteristic of many eutrophic waters – encourages these rapidly growing species, whereas more slowly growing, perennial species tend to store large quantities of nutrients (e.g., N). Macroalgal overgrowth in brackish/marine habitats commonly has been related to elevated N_i (e.g., Lapointe, 1997; Lapointe *et al*., 2005). NH_4^+ is more readily used (less energetically costly) than NO_3^-, but either N_i form has significantly stimulated growth of opportunistic macroalgae in field research and mesocosm experiments. PO_4^{-3} can also stimulate macroalgal growth when N is abundant.

For example, in Boston Harbor, Massachusetts, USA, sea lettuce (*Ulva*) formed dense populations for many years near outfalls of poorly treated sewage. The massive seaweed growth

reduced water flow, smothered shellfish, and affected shoreline areas with an intolerable stench of hydrogen sulfide from decomposition of senescent or dead material. In the Peel–Harvey Estuary, Australia, dense mats of green macroalgae and cyanobacteria have developed in response to P loading. In seagrass meadows worldwide, such overgrowth has been documented to reduce light and promote declines in various seagrass species (Burkholder *et al.*, 2007, and references therein). Along the intertidal area of the Baltic Sea, sewage inputs have been related to overgrowth of formerly dominant brown seaweeds (*Fucus* spp.) by opportunistic macroalgae.

In subtidal outfalls, nutrient enrichment from sewage and other wastes has stimulated overgrowth by 'weedy' red algae such as certain species of *Polysiphonia*. The red alga *Laurencia poiteaui*, the green alga *Cladophora fuliginosa*, and the brown alga *Cladosiphon occidentalis* proliferated over seagrass beds in eutrophying sewage-influenced areas of Florida Bay, USA, with maximal biomass at ~300 g dry weight per square meter. The exotic green macroalga *Caulerpa taxifolia* is a subtidal, subtropical weed that has invaded colder waters of the Mediterranean, and sewage appears to further stimulate its robust growth. This organism can produce metabolites that discourage predation; sea urchins typically avoid consuming *C. taxifolia* and starve to death if it is the only available food. Subtidal forests of giant kelp (*Macrocystis pyrifera*) died back and failed to reproduce during the 1960s–1970s near outfalls of sewage discharged from Los Angeles. The kelp bed loss apparently resulted from toxic substances in the poorly treated sewage.

Sensitive oligotrophic coastal marine coral reefs have been overgrown and smothered by macroalgae after nutrient input from sewage. In Kaneohe Bay, Hawaii, for example, phytoplankton near sewage outfalls first removed the N_i, and were transported to the central area of the coral reef where they decomposed and released N_i for stimulation of the opportunistic green macroalga, *Dictyosphaeria cavernosa*. Inorganic nutrient concentrations are frequently below detection in natural coral reef waters. Lapointe's (1997) research indicated that even minor sewage-related increases in N_i (>1 μM) and P_i (0.1–0.2 μM SRP) off the coast of Jamaica, together with decreased herbivory from natural and fishing-related disturbance, were sufficient to stimulate blooms of the green eutrophic indicator macroalga, *Chaetomorpha linum* and cyanobacteria that have overgrown coral reefs in that area.

Reduced light from such excessive nutrient-stimulated macroalgal (and sometimes phytoplankton) growth affects coral growth by decreasing the productivity of zooxanthellae, symbiotic photosynthetic dinoflagellates inside the coral tissue, that provide much of the corals' nutrition. Nutrient over-enrichment can also shift species dominance within the coral community. As the corals are out-competed for space, filter-feeding taxa such as sponges may be stimulated by increased phytoplankton food resources, and replace corals as dominant species. Sedimentation of decomposing phytoplankton, and other disturbances such as destructive techniques for fish harvesting, can exacerbate the impacts of cultural eutrophication on coral growth and survival.

Macrophytes, Aquatic Vascular Plants

Freshwater Communities

Floating, submersed, and emergent aquatic vascular plants, commonly called macrophytes, respond to nutrient enrichment differently because the competitive forces that they encounter vary with plant growth habit. *Rooted plants* – including those with specialized floating leaves, submersed growth, or emergent growth – generally do not compete with phytoplankton and other algae for nutrients, because they obtain most nutrients (except carbon, usually taken from the water) from the nutrient-rich sediment, with leaf uptake of nutrients from the less enriched overlying water as a secondary nutrient source. In contrast, floating plants at the water surface, such as the Lemnaceae (duckweed family) and the exotic weed *Eichhornia crassipes* (water hyacinth), must compete with suspended algae for N and P, but rely on the overlying air for their carbon. Floating submersed plants must compete with suspended algae for all major nutrients. They tend to have well-developed root systems to aid in nutrient acquisition, and in shallow waters with sparse open-water habitat, they can outcompete phytoplankton for light by forming dense surface populations that reduce or eliminate light availability in the underlying water. SAV is usually eliminated in highly eutrophic systems primarily through light reduction by phytoplankton and other algal overgrowth, as described by Phillips *et al.* (1978) and Wetzel (1978, 2001).

Like stream-inhabiting macroalgae, rooted macrophytes in river systems reach maximal abundance and species diversity in midorder segments. But, in contrast to lake-inhabiting macroalgae, submersed freshwater macrophyte communities attain highest species richness and abundance in mesotrophic lakes, often with dominance by *Potamogeton* spp. The native macrophyte beds help to decrease turbidity from incoming suspended sediments by slowing water motion and trapping sediment particles; their root systems also help to stabilize the lake bottom and reduce sediment resuspension. Light plays a major role in progressive SAV decline as lakes become increasingly eutrophic. In early stages of eutrophication, nutrient enrichment can stimulate macrophyte growth through luxury uptake and tight recycling, sometimes without an increase in phytoplankton. Over time, however, epiphytic algae take advantage of the greatly increased surface area for colonization afforded by the macrophytes, and they can both severely shade the underlying plants and restrict their carbon acquisition.

Increased epiphyte production stimulates growth of invertebrates and, in turn, can lead to increased abundance of cyprinids (e.g., bluegill sunfish – *Lepomis macrochiris*) and other invertebrate-feeding fish. These fish can reduce zooplankton densities, which relieves grazing pressure on phytoplankton that, in turn, proliferate and cloud the water so that light becomes yet more limiting for the macrophytes beneath. Under accelerated eutrophication, dramatic declines in native macrophyte populations typically occur over a relatively short period (several years). The loss of most native submersed macrophytes is often a critical turning point in the eutrophication of a lake or estuary. Without the habitat that provided cover for fish and substrata for littoral-zone invertebrates, cyprinid fish increase their grazing pressure on zooplankton, and their removal of zooplankton effectively decreases grazing

pressure on the phytoplankton. The increased anoxic/hypoxic conditions from yet-higher phytoplankton production cause further reductions in benthic invertebrate species and abundance, leading to more intense cyprinid grazing pressure on zooplankton. Thus, highly eutrophic lakes contain dense phytoplankton and food-limited (stunted) cyprinids.

Emergent macrophytes at the littoral fringe of lakes, rivers, estuaries, and marine coasts can maintain enhanced growth and biomass over a wide range of nutrient inputs. Under certain conditions with sustained high nutrient loading, however, emergent macrophytes can decline. The nutrients stimulate increased growth, greater stem density and, often, accelerated longitudinal growth, leading to higher intraspecific competition for light and weakening of stems. The elevated nutrients also promote higher suspended algal growth as well as higher growth of epiphytes on the submersed portions of the macrophyte stems. The increased weight makes the plants less well anchored, leading to dieback at the water's edge. Elevated N_i also promotes reduction in supporting vascular tissues, and loss of stem strength.

Estuarine and Marine Communities

Brackish waters are colonized by rooted, mostly freshwater species with moderate salt tolerance (as examples, certain *Potamogeton* spp., *Valisneria americana*, *Zanichellia*). In such habitats, light is often the primary resource limiting growth. Species with broader salt tolerance such as *Ruppia maritima* also can be abundant. Only about 50 species of angiosperms, mostly close relatives of freshwater *Potamogeton* spp., have the salt tolerance needed to thrive in marine habitats. Nearly all of these seagrasses grow in muddy substrata of shallow coastal lagoons and quiet embayments. Estuarine and marine macrophytes tend to be highly sensitive to light reduction and, thus, susceptible to eutrophication-related turbidity from algal (phytoplankton, epiphyte, and macroalgae) overgrowth and sediment loading/resuspension. Such shading causes gradual dieback and loss of most SAV, leading to dramatic declines, in turn, in the diversity and abundance of many species that depend on the habitat provided by these plants.

More sensitive SAV species are replaced by others that are tolerant to eutrophic conditions – for example, among subtropical seagrasses, field observations and limited experiments have indicated that turtle grass (*Thalassia testudinum*) is more sensitive to eutrophication than shoal grass (*Halodule wrightii*) and manatee grass (*Syringodium filiforme*). Other regions may not have additional seagrass species available to replace more sensitive species and, even where such species are present, they typically offer less desirable fish nursery habitat than the former dominant. Thus, in seagrass meadows under nutrient over-enrichment, more oligotrophic seagrass species are replaced by less sensitive species when available. As eutrophication progresses the seagrasses are eliminated, and rapidly growing macroalgae and phytoplankton become the dominant flora.

Although light reduction is considered the major mechanism for seagrass decline under cultural eutrophication, excessive nutrients can act independently of light to promote seagrass loss (**Figure 12**). The dominant north temperate species, eelgrass (*Zostera marina*), apparently lacks a physiological mechanism to inhibit NO_3^- uptake through its leaves,

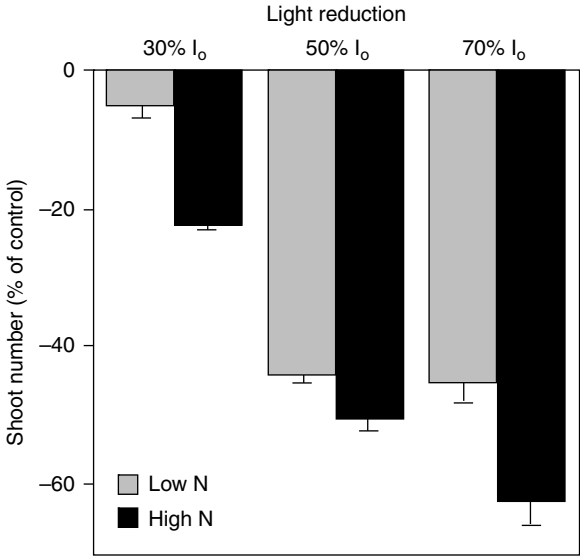


Figure 12 The effects of water-column NO_3^- enrichment and light reduction on shoot production of the seagrass, *Zostera marina*. From the first author's outdoor mesocosm experiments, indicated as the percent decrease from shoot production of control plants that did not receive water-column NO_3^- additions or light reduction (except that plants in controls and treatments all received an additional 30% light reduction for 3 h at 09.00, 12.00, and 15.00 h on a 3-day rotation using neutral density screens to stimulate conditions during high tide). Treatments were imposed for 10 weeks during the fall growing season for *Z. marina*. Controls were maintained at ambient natural light (except during simulated high tide) and NO_3^- (<30 μg NO_3^- N l^{-1}). Treatments included low N (at 50 μg NO_3^- N l^{-1}, or 3.5 μM NO_3^-N) and high N (at 100 μg NO_3^- N l^{-1}, or 7 μM NO_3^-N; low and high NO_3^- additions were added in the morning as a pulse of enrichment) at each of three imposed light levels as 30, 50, or 70% reduction of ambient surface light (I_o, accomplished using the neutral density screens, with additional shading to simulate high tide as noted). Water exchange (5–10% of the tank volumes per day replaced with fresh estuarine water) was done in late afternoons; see Burkholder et al. (1992), for further details about the mesocosm system. In all treatments with water-column NO_3^- enrichment, *Z. marina* declined in shoot production relative to shoot production of control plants, and the NO_3^- inhibition effect was exacerbated by light reduction (means $+1$ standard error; $P<0.05$, $n=3$ mesocosms for controls and 3 for each treatment). These effects were not caused by algal overgrowth, which was maintained at low levels in controls and all treatments throughout the experiment.

as indicated by the research of Roth and Pregnall (1988) and Burkholder *et al.* (1992). Most plants take up NO_3^- during the day with the energy from photosynthesis. In contrast, *Z. marina* takes up water-column NO_3^- day or night if it becomes available, as shown by Touchette and Burkholder (2000) (**Figure 13**). This species is thought to have evolved in N-poor coastal waters, and sustained NO_3^- uptake under temporary enrichment may have developed as a once highly advantageous competitive mechanism or "strategy". However, as coastal waters have become more eutrophic from sewage, septic effluent leachate, agricultural sources, and other anthropogenic sources, sustained uptake of water-column NO_3^- likely has become a disadvantage. NO_3^- enrichment of the sediments, under control by an abundant microbial

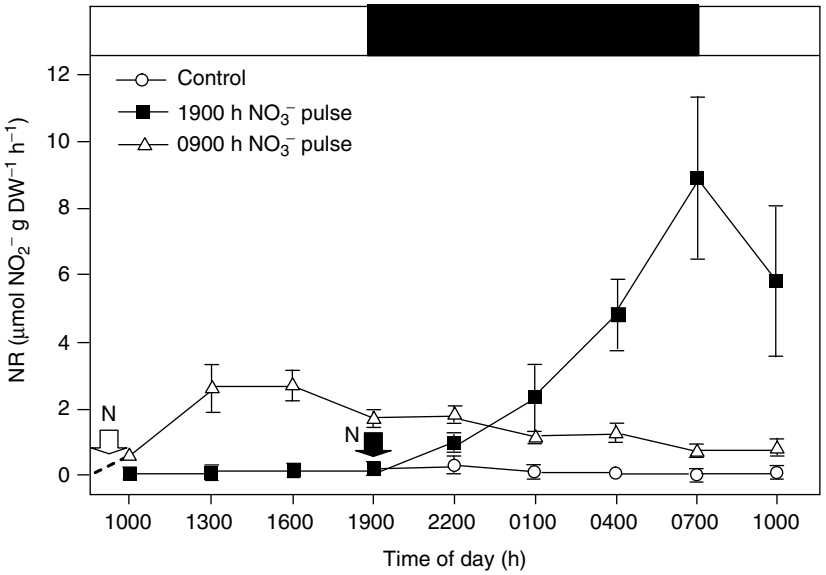


Figure 13 The response of the seagrass, *Zostera marina*, to pulsed water-column NO_3^- enrichment in light and dark periods. Plant NO_3^- uptake is indicated as leaf activity of the enzyme (nitrate reductase (NR), used to actively take up NO_3^-) of previously unenriched shoots (means ± 1 standard error). A spike of NO_3^- (110 µg NO_3^- N l^{-1}) was added in the morning (white arrow) or, to a subsample of plants from the same population, at night (black arrow). NR activity (plotted as micromoles of nitrite product produced per gram dry weight of plant leaf tissue per hour) indicated that *Z. marina* took up NO_3^- day or night, whenever a pulse was detected. In fact, maximal NR activity was significantly higher when NO_3^- was added during the dark period. Reprinted from Touchette BW and Burkholder JM (2000) Overview of the physiological ecology of carbon metabolism in seagrasses. *Journal of Experimental Marine Biology and Ecology* 250: 169–205.

consortium, does not cause a similar effect and, instead, can be mildly stimulatory.

NO_3^- uptake is a metabolically expensive process, requiring high cellular energy. Research by Touchette and Burkholder (2000) indicated that sustained water-column NO_3^- uptake by *Z. marina* can promote severe internal carbon imbalances, apparently from the need to shunt C skeletons from photosynthesis for use in high amino acid synthesis to prevent internal accumulation of toxic products such as NH_3. The physiological mechanism of an internal "carbon drain" from sustained NO_3^- uptake was earlier documented for algae by Turpin *et al.* (1991). A common trait of *Z. marina* shoots under excessive water-column NO_3^- enrichment is structurally weakened growing regions, perhaps analogous to the above-mentioned loss of stem strength that has been reported in certain freshwater emergent macrophytes under NO_3^- enrichment. Excessive N_i enrichment has also promoted seagrass attack by pathogens (e.g., the slime mold *Labrynthula zosteroides*), hypothesized to occur because N and C are internally shunted to amino acid production rather than to production of alkyloids and other antimicrobial compounds.

Another seagrass that has been examined for the NO_3^- inhibition phenomenon, *Halodule wrightii*, and certain macroalgae (e.g., *Ulva lactuca*) have shown depressed growth in response to NO_3^- enrichment, although at much higher N levels (*ca.* 1.4 mg NO_3^-N l^{-1}, pulsed daily for 4–5 weeks) than for *Z. marina* (50–110 µg NO_3^-N l^{-1}, pulsed daily for 5–8 weeks). *Ruppia maritima* has been experimentally stimulated by high water-column NO_3^- but inhibited by elevated N_i as NH_4^+; and *Z. marina* has been experimentally inhibited by high NH_4^+ levels, as well. In mesocosm experiments, light reduction has been shown to exacerbate the inhibitory effects of water-column NO_3^- enrichment on shoot production in *Z. marina*. Warm temperatures exacerbate water-column NO_3^- enrichment impacts on root growth of *Z. marina*, as well, suggesting that warming trends in climate change may be expected to interact with eutrophication to adversely affect this beneficial habitat species.

Invasive macrophytes such as *Hydrilla verticillata* (hydrilla) or *Egeria densa* (Brazilian elodea) are also a concern in ecosystems that have been altered by nutrient loading. These plants appear to be most common in systems that have received excessive N loading relative to P, or systems in which P has been removed but N loads remain high. Examples of ecosystems invaded by these plants after P removal include the Potomac River, a tributary of Chesapeake Bay, and the San Francisco Bay Delta. Both systems receive (or historically received) high nutrient loads from sewage discharge, but have sustained reduced P loading due to improved effluent treatment and removal of P from laundry detergents. Highly productive invasive plants like *H. verticillata* and *E. densa*, have been characterized by Yarrow *et al.* (2009) as "ecological engineers." They can thrive in turbid waters but also tend to trap sediments, thereby reducing turbidity. They also provide habitat for zooplankton and fish, and through their productivity, they alter pH and both water-column and sediment nutrient chemistry.

Microfauna

Freshwater Communities

Whereas plants and mixotrophic algae respond directly to nutrient enrichment, animals are generally indirect recipients

of eutrophication impacts. Nevertheless, all are significantly affected by nutrient control of the quality (including size, taste, and ease of filtering) and quantity of their algal, bacterial, or detrital food (Harper, 1992). Freshwater zooplankton consist of three main groups: protozoa (mainly ciliates and flagellates) and rotifers form the microzooplankton (maximum dimension *ca*. 45 μm to less than 70 μm); and crustaceans form the macrozooplankton (mainly cladocera, calanoid, and cyclopoid copepods, *ca*. 70 μm to less than 500 μm).

Protozoa diversity and abundance increases with the amount of available organic matter that becomes directly or indirectly available through eutrophication, until the fauna are restricted by low-oxygen conditions. Oligotrophic protozoan plankton can be dominated in number by small-bodied ciliates, whereas large ciliates can contribute most of the biomass (especially members of the Oligotrichida such as *Plagiopyla nasuta* and *Paramecium trichium*). Mesotrophic and eutrophic lakes are more commonly dominated in both number and biomass by small-bodied species (e.g., member of the Scuticociliatida). Protozoans are relatively sparse in oligotrophic lakes in comparison to the species-rich and abundant fauna that develop in moderately eutrophic waters. Aggregates of bacteria and detritus form in the water column as organic materials decompose, and these microhabitats can support protozoa that rival shallow benthic communities in diversity. Subtropical lakes show similar trends, but have been reported to support higher ciliate abundance and diversity at a given trophic status than temperate lakes.

Heterotrophic flagellates in the water column are bacteriovores, whereas larger ciliates and heliozoa commonly prey on mixed populations of heterotrophic and autotrophic flagellates, and other algae. In low-oxygen waters, heliozoan amoebae and ciliates such as *Coleps* and *Euplotes* retain endosymbiont green algae (zoochlorellae) that photosynthesize and generate oxygen for their hosts. As nutrient enrichment increases, bottom waters and surface sediments become more organically enriched, but the physiological demands on protozoans and other microfauna such as certain rotifers are greater. Survival generally depends on tolerance of low-oxygen and often co-occurring acidic conditions. In organically overloaded, hypoxic sediments, protozoa species diversity is very low relative to that in well-aerated sites. Bacteria and dissolved organic matter form the major food sources, consumed by species such as *Paramecium*, *Chilomonas*, and *Astasia*. Anoxic sediments become colonized by specialized protozoa such as pelobionts and diplomonads, as well as certain ciliates and amoebae.

Rotifers include many algal herbivores as well as detritivores and a few carnivorous species. They generally attain highest species diversity and abundance with increasing small-celled phytoplankton in moderately eutrophic lakes and lower rivers. These microzooplankton generally are favored in abundance over macrozooplankton under increasing eutrophication. However, rotifer species diversity declines as eutrophication progresses to dominance of the phytoplankton by cyanobacteria, especially large mucilaginous colonial species that clog the filtering apparati of the animals. Rotifers can be adversely affected, as well, by the toxins from certain cyanobacteria that are seasonally characteristic of eutrophic lakes.

Among macrozooplankton in freshwaters, cyclopoid copepods (raptorial feeders) generally feed most efficiently on larger particles (e.g., algae with biovolume $\leq 1000\ \mu m^3$), whereas calanoid copepods consume particles $\leq 100\ \mu m^3$, and cladocerans eat small particles at *ca*. $10\ \mu m^3$. Changes in biodiversity under cultural eutrophication depend, in part, on differences among these species in efficiencies of feeding at certain particle (algal or detrital) concentrations and size ranges, and in the responses of life history stages to food limitation. Changes in biodiversity under eutrophication can also result from fundamental differences in the nutrient stoichiometry required to make the biomass of one type of organism compared to another. Cylopoid copepods and cladocerans, for example, are generally more P-rich in their biomass than calanoid copepods. Nutrient enrichment alters food particle size and abundance which, in turn, leads to competitive species displacement. Larger-bodied individuals are often more efficient filter feeders than smaller fauna. However, larger species may not be competitively superior because their juveniles tend to be more vulnerable to food limitation than adults of smaller species. Smaller species withstand food depletion by reducing metabolism, growth, and egg production. Moreover, their specific food ingestion rate (milligrams food ingested per milligram of zooplankton biomass) is usually higher than that of larger zooplankton.

Calanoid copepods (e.g., *Eudiaptomus*) often dominate the microfauna of oligotrophic and mesotrophic lakes. These organisms feed efficiently on larger algal cells in those systems, with higher ingestion efficiencies at low food density. As algal prey densities increase under eutrophication, larger cladoceran species become abundant under low to moderate predation by fish, whereas high predation tends to select for smaller cladocerans such as *Bosmina* and *Ceriodaphnia* over *Daphnia*. The larger cladoceran species tend to have higher population growth rates than calanoid copepods. Their population cycles during warmer months have been shown to be mainly controlled by the relative proportions of edible and inedible algae, temperature, and predation. Like rotifers in the presence of abundant food, cladocerans can reproduce parthenogenically, thus allowing for rapid growth rates. They are considered more efficient filter feeders at moderate to high algal densities because they ingest more food for the same amount of energy expended, relative to ingestion under food-limited conditions.

Macrozooplankton biomass increases under eutrophication in both temperate and tropical lakes, with increasing dominance by small-bodied species. High-efficiency bacterial feeders are selected for, as bacterial abundances increase under accelerated eutrophication. However, abundant co-occurring cyanobacteria with copious mucilage clog the filtering apparati of large-bodied microzooplankton. Moreover, some cyanobacteria are directly toxic to these fauna. In highly eutrophic lakes, the more selective feeding of calanoid copepods, and the seizing behavior of herbivorous cyclopoid copepods, may afford one or both groups a competitive advantage. Smaller cladocerans such as the littoral-zone chydorid *Chydorus* may become abundant during dense cyanobacteria blooms because of their high efficiency in filtering the extremely small, solitary eubacteria and cyanobacteria particles associated with the blooms while being able to avoid filter apparatus clogging

by large and mucilaginous cyanobacteria colonies. Eutrophication can cause other impacts on zooplankton species, apart from changes in the available food. For example, as bottom waters become increasingly hypoxic, survival is depressed for zooplankton eggs that sink to the bottom as part of the life cycle. Thus, nutrient enrichment can impair zooplankton recruitment as an indirect effect.

Estuarine and Marine Communities

Impacts of anthropogenic nutrient loading have been more difficult to generalize from the complex flow/water exchange environments of estuaries and from large-scale marine environments with high physical and biological variability. Nevertheless, substantial progress has been made in the past two decades in understanding the effects of cultural eutrophication on estuarine and marine fauna (**Figure 14**). Analyses by Michell (1999) and others have indicated that nutrients generally enhance phytoplankton biomass and carnivores depress herbivore biomass. The strength of the couplings between trophic levels (e.g., phytoplankton and zooplankton) varies, and is strongest in more poorly flushed or closed systems such as upper embayments and coastal lagoons.

Brackish and marine habitats differ from freshwaters in having fewer rotifers or cladocera, more extensive representation by protozoans (especially ciliates and foraminiferans), and often-abundant planktonic nauplii (young life history stages) of sessile adult fauna ranging from molluscs and malacostracan crustaceans to vertebrates. Holoplankton spend their entire lives in the water column. Among these are microzooplankton including larval forms of certain macrozooplankton, as well as tintinnid and nonloricate ciliated protozoans, heterotrophic flagellates, and amoebae. In benthic habitats, the 'meiofauna' (similar in size to microzooplankton) include nematodes, harpacticoid copepods, many turbellarians, and several minor phyla with diverse feeding habits and life styles, mostly consumers of other microbes as prey. Macrozooplankton include calanoid copepods, especially of the genus *Acartia*, cyclopoid copepods and planktonic harpacticoids, and non-copepod crustaceans (especially carideans and mysids) and chaetognaths (arrow worms). Meroplankton, which spend only part of their lives in the plankton as larval stages, may include immature forms of benthic invertebrates and tunicates; eggs, larvae, and juveniles of shrimp, crabs and fish; and sexual stages of hydrozoan and scyphozoan cnidarians (jellyfishes).

The microzooplankton are among most important herbivores on phytoplankton among estuarine and marine zooplankton and, thus, are strongly influenced by changing phytoplankton food quality and abundance under eutrophication, with trends that are somewhat analogous to those described for freshwater microzooplankton. Zooplankton species diversity is highest in moderately nutrient-enriched waters, but significantly declines as nutrient enrichment becomes more excessive with accompanying shifts to dominance of unpalatable algal species and pronounced bottom-water oxygen deficits. For example, in the Peel-Harvey Estuarine

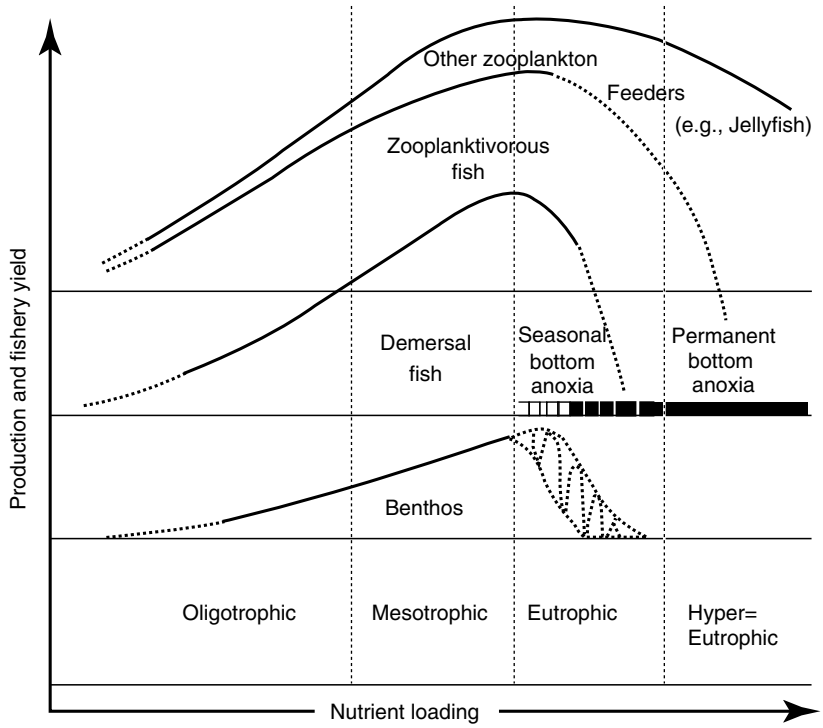


Figure 14 Comparative evaluation of fishery response to nutrients based on data from around the world. Each curve represents a natural guild of species and their reaction to increasing nutrient supplies. The vertical dotted lines separate general categories of organic production that result from different levels of nutrients. Modified from Diaz RJ (2001), Figure 1, who modified and redrew the original figure from Caddy's (1993) model relating fish yield to nutrients supplied to estuarine and marine ecosystems (Caddy J (1993) Toward a comparative evaluation of human impacts on fishery ecosystems of enclosed and semi-enclosed seas. *Reviews in Fisheries Science* 1: 57–96).

System of Australia, large populations of calanoid copepods *Sulcanus* and *Gladioferans* grazed winter diatom blooms, but were rapidly eliminated when the noxious (potentially toxic) filamentous cyanobacterium, *Nodularia spumigena*, became abundant (McComb and Lukatelich, 1995). Predation of carnivorous macrozooplankton by planktivorous fish such as menhaden (e.g., young life history stages of Atlantic menhaden, *Brevoortia tyrannus*) can locally favor small-bodied zooplankton. Such fish become more abundant in localized areas such as small tributaries and are known to track plankton blooms. Later in the growing season these fish switch to phytoplankton prey and, thus, directly compete with herbivorous zooplankton for food resources.

Under high nutrient enrichment, changes in food quality/availability also adversely affect larval stages of many species represented in the meroplankton. Tunicates can increase at the expense of planktonic crustacea, and gelatinous meroplankton (cnidarians, especially jellyfish) can increase at the expense of fish. Jellyfish are favored under increasing eutrophication (Vasas et al., 2007). They are predators on zooplankton and juvenile fish, and are avoided by most pelagic fish, which do not consume them. Low-oxygen conditions in bottom-water habitat decrease survival of zooplankton eggs, and of eggs and larvae of many meroplankton species that settle to the bottom sediments as part of their life cycle.

Invertebrate Macrofauna

The freshwater zoobenthos or bottom macrofauna (>0.5 mm, or 500 μm) consist mostly of insect larvae (with terrestrial adult stages), crustacea, worms, and molluscs. Eutrophication in the littoral zone of lakes and rivers, with accompanying organic pollution, causes similar impacts on the invertebrate macrofauna of these systems. In early stages of eutrophication, oligochaetes, chironomids, gastropods, and sphaerids increase, and mayfly nymphs such as *Hexagenia* decrease (Harper, 1992). As eutrophication continues, the major pattern with progressive hypoxia/anoxia in overlying waters and surface sediments is the decline and then disappearance of additional oxygen-sensitive species such as stoneflies (*Diura*) and certain taxa of mayfly nymphs (e.g., *Baetis, Rhithrogena*) as well as certain caddisfly larvae (e.g., *Rhyacophila, Hydropsyche)* and bivalve molluscs.

Eutrophication increases organic matter and bacterial decay, and depresses oxygen concentrations while increasing settlement of organic detritus to benthic organisms such as certain flatworms (e.g., *Polycelis*) that use it for food. In the profundal sediments underlying deeper waters of lakes, the macroinvertebrate biomass increases but is comprised of low oxygen-tolerant species such as certain chironomid larvae and oligochaete worms (e.g., the oligocheate *Tubifex tubifex* which, in one study, survived, grew, and reproduced under continuous hypoxia/anoxia for 10 months). The species diversity of tubificid oligochaetes decreases with advanced eutrophication and organic enrichment, and associated oxygen deficits. However, if oxygen is periodically available, the rich food supply in combination with the lack of more oxygen-sensitive competitors allows robust growth. Declines in chironomid communities occur most rapidly in the change between nutrient-poor to moderately nutrient-enriched (mesotrophic) waters; and numbers of oligochaetes relative to chironomids increase as organic enrichment increases. In the intermediate zone between the littoral and profundal, the increased supply of fresh littoral detritus is consumed by large-bodied detritivores such as the bivalve mollusc *Dreissena*.

Estuarine and marine coastal invertebrate macrofauna, including molluscs, polycheates, decapods, and other crustacea, and nemerteans, spend their adult lives buried beneath the sediment surface. They are highly diverse in food acquisition as filter feeders, nonselective deposit feeders, selective deposit feeders, or raptorial/other predators, making generalizations difficult. Increasing organic matter from accompanying nutrient enrichment tends to cause similar impacts as in freshwater systems, namely, an increase in macroinvertebrate abundance under moderately eutrophic conditions, and a decline in species diversity and abundance as eutrophication and associated hypoxia/anoxia progress. Organisms that burrow into anaerobic sediments must be able to gain access to oxygen in the overlying water. Increasing eutrophication leads to elimination of burrowing organisms as the anaerobic zone moves closer to the sediment surface. In highly over-enriched areas, only worms such as *Capitella* may survive. Oxygen does not need to be completely absent for damage to occur – hypoxic waters with 3.0–4.3 mg DO l^{-1} have been related to mortality of some benthic invertebrate species, and to loss of habitat for some shellfish species that require higher oxygen availability.

Changes in food quality and quantity also can reduce the species diversity and abundance of estuarine and coastal marine invertebrate macrofauna under highly eutrophic conditions. For example, an extensive early study by Filice (1959) in San Francisco Bay demonstrated that in domestic sewage outfall areas, few species survived. Species diversity was higher but still depressed in surrounding areas that received dilute sewage. However, in those areas, some species (e.g., the clams *Gemma gemma, Mya arenaria,* and *Macoma inconspicua*; the polychaete worm *Polydora uncata*, and the barnacle *Balanus improvisus*) apparently took advantage of new energy and material resources, and became highly abundant relative to their abundance in controlled areas without sewage influence. More recently, an extensive analysis by Glibert (2010) and Glibert *et al.* (2011) examined major trophic shifts, from phytoplankton composition to mid-trophic levels to fisheries, in the San Francisco Bay Delta over the past *ca.* 30 years. This analysis related the major trophic shifts to significant alterations in N : P ratios that especially resulted from ammonium-enriched sewage discharges. These changes included apparent significant effects on food quality for various fauna. In the Great South Bay complex of Long Island, New York, previously mentioned duck farms along the bay tributaries fertilized the water with nutrients and organic wastes, and stimulated dense blooms of the small algae, *Nannochloris* sp. and *Stichococcus* sp. ($>10^6$ cells ml^{-1}). These algae were very different from the previous phytoplankton assemblage in the area, which consisted of mixed species that are needed to support oysters. Following these changes, oyster populations significantly declined because they were unable to thrive on a diet consisting only of these small algae.

Vertebrate Macrofauna

Cultural eutrophication initially reduces and then eliminates sensitive lake fishes (e.g., salmonids and coregonids), by eliminating oxygen-replete, colder bottom-water habitat, and well-oxygenated spawning areas (Harper, 1992). Fish can avoid low-oxygen waters, but the cold-optimal species encounter warmer waters as they are forced to move from deep areas into the shallows. Most temperate-zone fishes breed in the littoral. Their eggs are more vulnerable to developing short-term (e.g., nightly) oxygen deficits, and to the low-oxygen micro-environment of the increased detritus. Thus, the critical habitat of these sensitive species is destroyed by nutrient over-enrichment, and spawning and recruitment are depressed. Overall, nutrient enrichment generally increases fishery yields to a certain level, but beyond that level fish production declines.

Analogous impacts occur in rivers and estuaries. For example, in parts of the Baltic Sea, cod eggs laid in well-oxygenated surface waters die when they sink to anoxic bottom waters (Diaz, 2001). Oxygen levels in the bottom waters of the Baltic's deep basins are negatively correlated with juvenile codfish abundance. Hypoxia in estuaries has been linked to depressed survival of larval fish, mortality of certain benthic invertebrates used as fish prey, and loss of habitat for mobile species of finfish such as cod that require high oxygen availability. Hypoxia/anoxia represents a growing problem for many estuaries and coastal marine waters, such as the Chesapeake Bay, the Baltic Sea, the Black Sea, the Pamlico Estuary, Long Island Sound, the North Sea, and the Gulf of Mexico by the mouth of the Mississippi River.

The composition of fish communities also changes as the water-column nutrient stoichiometry changes. In addition, planktivorous fish generally differ from demersal (bottom-dwelling) or piscivorous fish in their nutrient stoichiometry. For example, making bone is a P-requiring process; cyprinid fishes are bonier and thus require more P relative to small, planktivorous fish. Thus, nutrients can have both direct (mediated through the food web) and indirect effects on fishes.

Overall population sizes and biomass of fish usually increase with nutrient over-enrichment, often regarded as a beneficial effect in early to mid-stages of eutrophication. However, dominance shifts – for example, in freshwaters the community changes from species such as lake trout to cyprinids, bullheads (*Ictalurus* spp.), and other coarse fish that can tolerate low oxygen concentrations (**Figure 15**). In lakes with well-developed littoral zones of rooted aquatic vegetation, high predation pressure from piscivorous fish such as largemouth bass (*Micropterus salmoides*) or pike (e.g., *Esox lucius*) maintains low biomass of cyprinid fish and, thus, low predation pressure on large-bodied zooplankton. These large zooplankton, in turn, control phytoplankton biomass through strong grazing pressure, so that there is high visibility for visually feeding piscivorous fish, and abundant light for macrophyte growth.

As eutrophication progresses, the littoral-zone macrophytes disappear and the carrying capacity of the system for piscivorous fish is reduced. Thus, (zoo)planktivorous fish are freed from high predation. Their larvae tend to selectively feed on the largest herbivorous zooplankton species (based on the size-efficiency hypothesis of Brooks and Dodson (1965)), causing an overall reduction in herbivore size. Smaller herbivores cannot exert enough grazing pressure to control blooms of cyanobacteria and other noxious algae, which proliferate in eutrophic lakes and lower rivers in the absence of the large-bodied zooplankton. Turbidity can increase from higher cyprinid feeding activities and from loss of the littoral-zone filtering effect of macrophytes, further impairing visual predation by remaining piscivorous fish. Photosynthetic activity from the high algal biomass can elevate the pH to levels that adversely affect gill function in sensitive fish species.

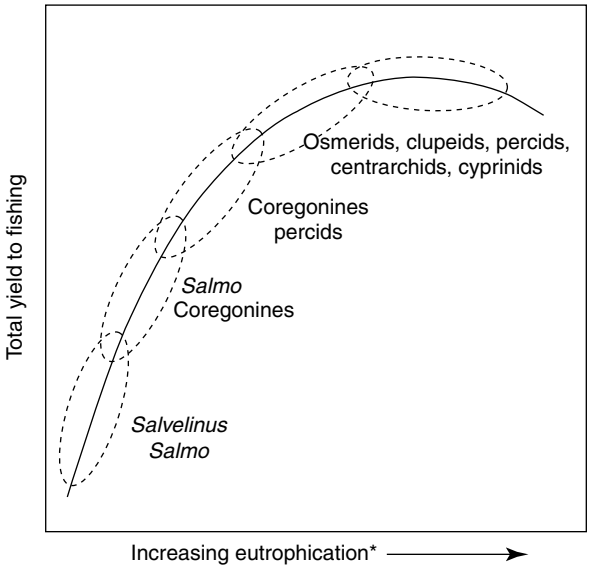


Figure 15 Trends in fish yields and taxonomic composition with increasing eutrophication, indicating total yield (as fish caught) of representative taxonomic groups from various North American lakes.

Similar impacts on fish populations from loss of desirable littoral zone species have been documented in estuarine and coastal waters. For example, along the Swedish coast of the Baltic Sea, excessive nutrient enrichment has been related to increases in nuisance cyanobacteria and other filamentous species and a decrease in formerly dominant *Fucus vesiculosus* (Phaeophyceae), probably because of reduced light availability (Kautsky *et al.*, 1991, and references therein). This seaweed had provided critical habitat for herring spawns, and the shift to dominance by filamentous macroalgae led to decreased egg hatch in the herring populations.

Although aquatic birds and other vertebrates can move among aquatic environments with relative independence, some general effects of eutrophication on these macrofauna also have been described. Increased macrophyte vegetation under moderate nutrient enrichment in mesotrophic systems has been linked to increased numbers of herbivorous waterfowl such as moorhen (*Gallinula chloropus*) and whooper swan (*Cignus cignus*). Some piscivorous birds such as grebes and herons have also increased. As eutrophication progresses, however, bird species may decline following undesirable changes in food and habitat. For example, pochards (diving ducks, *Aythya ferina*) decreased on Lake Constance as their main food supply, the macroalga *Chara*, declined under nutrient enrichment. General macrophyte habitat loss under

nutrient over-enrichment has also been linked to the gradual disappearance of many waterfowl such as black and mute swans (*Cygnus atratus, Cited olor strepera*), Canada geese (*Branta canadensis*), coots (*Fulica atra*), teal (*Anas crecca*) and gadwells (*Anas strepera*). Sensitive life history stages of amphibians (e.g., frog eggs) have been killed in increasing hypoxic events within littoral zones of eutrophic lakes.

It should be noted that waterfowl and other wildlife can also act as eutrophying agents, especially in small lakes and coastal lagoons or poorly flushed embayments. Where waterbirds are abundant, exemplified by the increasing problem created by exotic (nonindigenous) Canada geese (*Branta canadensis*) in small ponds and other waters within urbanizing and agricultural areas of the US, their excreta can be a major nutrient source. For example, a small lake draining agricultural lands in the upper Midwest was used by Canada geese and other waterfowl (mostly mallards, *Anas platyrhynchos*, another exotic species). An analysis by Manny *et al.* (1994) estimated that the lake received ~70% of all external carbon sources, 27% of all N, and 70% of all P from the waterbirds. Overall, however, based on review of available studies, Hahn *et al.* (2008) reported that in most lakes, bird-mediated nutrient input is generally of minor importance compared with agricultural or sewage contributions.

Because of human disturbance, wildlife have also become important contributors of fecal bacteria and associated pollutants such as nutrients to some marine shoreline waters as a localized effect (Siewicki *et al.*, 2007). Coastal urbanization has increased impervious surface area, increasing runoff that is channeled into retention ponds or directly into waterways. These changes have increased the wildlife carrying capacity of the landscape, while also pushing wildlife (deer, raccoons, waterfowl, shorebirds) away from developed areas and into marsh edges near shellfish beds.

Aquatic Fauna and Inorganic Nitrogen Toxicity

Humans have dramatically altered the Earth's nitrogen cycle (Smil, 2001), and high levels of inorganic N enrichment to surface waters from fertilizer application and runoff, air pollution, human and animal wastes and other sources have been shown to impair the growth, reproduction, and survival of aquatic animals. Aquatic fauna can directly take up unionized ammonia (NH_3), NH_4^+, nitrite (NO_2^-), nitrous acid (HNO_2), and NO_3^- from the surrounding water. Of these, following the excellent review of Camargo and Alonso (2006), NH_3 is the most toxic and NH_4^+ and NO_3^- are the least toxic, but even NO_3^- has a wide range of toxic effects at much lower concentrations that are commonly added to waterways by agriculture and other practices. As a generalization, estuarine and marine fauna appear to be comparatively more tolerant of the toxic effects of inorganic nitrogenous compounds than freshwater fauna, likely because of ameliorating effects of chloride and other ions.

NH_3 is highly toxic especially to fish and is thought to act through one or more of the following mechanisms: damage to gill epithelium, causing asphyxiation; stimulation of glycolysis, and suppression of the Krebs cycle, resulting in acidosis and reduced capability of the blood to carry oxygen; uncoupling of oxidative phosphorylation, inhibiting ATP production and depleting ATP in the basilar region of the brain; disrupting blood vessels and osmoregulation, impairing liver and kidney functions; and suppressing the immune system, increasing susceptibility to disease. NH_4^+ ions can exacerbate NH_3 toxicity by reducing internal sodium ion concentrations. Environmental conditions such as increased pH, warmer temperatures, and low DO can increase fish susceptibility to NH_3 toxicity, and mixtures of NH_3 with other chemical pollutants such as certain heavy metals can result in additive or synergistic toxic effects. Freshwater invertebrates such as molluscs and planarians, and salmonid fishes appear to be most sensitive to NH_3 toxicity, and have sustained adverse impacts from chronic exposures to as little as 50 μg NH_3 N l^{-1}. Water quality recommendations developed in the US and Canada to protect sensitive aquatic life have ranged from 50–350 μg NH_3 N l^{-1} for acute (short-term) exposures and 10–20 μg NH_3 N l^{-1} for chronic (long-term) exposures. Such levels starkly contrast with concentrations commonly discharged in effluent from secondary sewage treatment across the US (10^3–10^4 μg NH_3 N l^{-1}), animal waste spills (10^4 μg NH_3 N l^{-1}), etc.

NO_2^- also inhibits nitrifying bacteria, which can result in its increased accumulation and intensified toxic effects. The main mode of action, especially in fish and crayfish, is conversion of oxygen-carrying hemoglobin or hemocyanin pigments into methemoglobin or methemocyanin forms that can no longer carry oxygen, leading to hypoxia and death. Other effects of NO_2^- toxicity in fish and crayfish have included depletion of chloride levels, leading to severe electrolyte imbalance; potassium ion imbalances that adversely affect membrane potentials, neurotransmission, skeletal muscle contractions, and heart function; formation of mutagenic and carcinogenic N-nitroso compounds; damage to mitochondria in liver cells, resulting in oxygen shortage; and immune system supression. Elevated concentrations of certain ions such as chloride in brackish and marine waters are protective because they inhibit NO_2^- uptake. The most sensitive aquatic invertebrates tested so far have included decapod and amphipod crustaceans, ephemeropteran insects, and salmonid and cyprinid fishes. Based on acute toxicity data, it has been recommended that NO_2^- N concentrations should be maintained below 80–350 μg l^{-1} (depending on the organisms of concern) to protect sensitive aquatic fauna during short-term exposures.

The main mode of action of NO_3^- on aquatic animals is similar to that of NO_2^-. In experiments with aquatic organisms such as fish, reptiles, and amphibians, NO_3^- has also been shown to decrease immune response; to induce hematological and biochemical changes; and to adversely affect many metabolic processes by acting as an endocrine disruptor (Guillette and Edwards, 2005; Camargo and Alonso, 2006; McGurk *et al.*, 2006). Because fish and crayfish have lower branchial permeability to NO_3^- ions than to NO_2^- ions, uptake is more limited. The aquatic animals tested as most sensitive to NO_3^- toxicity are freshwater caddisflies, amphipods, certain amphibians, and salmonid fishes. For example, chronic NO_3^- toxicity for freshwater invertebrates can occur at values as low as 230 μg NO_3^- N l^{-1} (0.23 mg l^{-1}). Lowest chronic toxicity levels reported for adult freshwater invertebrates were 2.8–4.4 mg l^{-1} for two amphipod species; early instar caddisfly larvae sustained adverse effects from chronic toxicity at

1.4–2.4 mg l^{-1}. A threshold concentration of 1.1 mg NO_3^- N l^{-1} significantly increased egg mortality of salmonids, and 1.6 mg l^{-1} caused delayed hatching and development, and reduced weight in larval lake trout.

NO_3^- toxicity to aquatic life, historically, was dismissed as irrelevant, but that view has changed in consideration of the excessive NO_3^- concentrations that are now routinely added to surface waters and subsurface groundwaters by human activities (Stanley and Maxted, 2008, and references therein). Recent work suggests, for example, that excessive concentrations in agricultural areas such as the Corn Belt of the upper midwestern US (concentrations up to 25 mg NO_3^- l^{-1} in receiving surface waters, and commonly above 5 mg l^{-1} during crop fertilizer periods; Stanley and Maxted, 2008) are contributing to amphibian declines due to impaired swimming ability, decreased body size, and reduced fecundity and survival.

Oligotrophication – Reversing the Impacts of Eutrophication?

It is possible to reverse nutrient loading impacts in some aquatic ecosystems, at least partially, by reducing the nutrient inputs. This process can occur naturally when, for example, a major storm such as a hurricane causes high flooding or sedimentation that effectively reroutes a nutrient-rich tributary away from a receiving lake or estuary (such as an oxbow lake that is cut off from the main river); when natural flooding destroys a dam and eliminates a run-of-river impoundment; or when a raised bog system becomes isolated from nutrient-rich stream and groundwater sources. As nutrient loading declines, species diversity generally increases while the system production decreases. In ombrotrophic bog systems, there is also a shift over time to more acid-tolerant species, as the system's sole source of new water is low-pH precipitation.

Human influences are generally more pervasive than the natural forces that can contribute to oligotrophication. "Cultural eutrophication" is an acceleration of a long-term, natural process. Resilience, defined as the rate of recovery to the pre-disturbance state, depends on the initial status of the system in the natural eutrophication process, and on the degree of nutrient loading sustained. Ecosystem resilience is generally considered to increase with increased nutrient loading rate, and to decrease with increasing food chain length. Analysis by Cottingham and Carpenter (1994) suggested a modification of this relationship for pelagic food webs of north temperate lakes. In that analysis, planktivore-dominated (short-length) food webs were more resilient at baseline P loading rates of 0.1–1.0 μg l^{-1} day^{-1}. Piscivore-dominated (long-length) food webs were more resilient at high baseline P loading rates (2.0 μg l^{-1} day^{-1}), apparently because the additional nutrients were incorporated into the biota more rapidly due to their different growth requirements.

Eutrophic systems, dominated by generalist or opportunist species that are insensitive to the adverse impacts from nutrient over-enrichment, are considered to be relatively resistant to further change or stress as mentioned. Moreover, the adverse impacts of cultural eutrophication in such systems can be partially reversed in a relatively shorter period of time (several years). Less nutrient enriched systems (e.g., in early phases of moderate nutrient enrichment or mesotrophy) have more specialized species, more opportunity for biotic adjustment of elemental cycles, and tighter coupling among element cycles. The diverse communities of such systems, comprised of more sensitive species, are relatively sensitive to further stress imposed by nutrient enrichment under accelerated eutrophication; and reversal or recovery to a pristine oligotrophic state is difficult to accomplish.

The degree of restoration success also depends on certain physical features of the system, especially mean depth and flushing rate. Aquatic ecosystems tend to retain most nutrients from loading events in the bottom sediments. PO_4^{-3} is especially "sticky"; it readily adsorbs to the iron oxyhydroxides of sediments, and to clays and other sediment particles. In deep lakes, most of the nutrient inputs from the previous years are unavailable to phytoplankton of the next growing season because the lower water column remains permanently stratified and, thus, isolated from the upper water column where the majority of the viable phytoplankton occur. In shallow systems, wind and internal currents mix the water column and surficial sediments completely, sometimes frequently over an annual cycle, so that some portion of the nutrients from the relatively rich bottom waters and sediments repeatedly becomes available for phytoplankton growth. Reservoirs, rivers, well-flushed estuaries, and wave-swept, marine coastal waters are usually less sensitive to nutrient loading and more easily "reversed" in the eutrophication process than lakes or poorly flushed coastal lagoons with long water exchange times (on the order of months to years), because the latter systems cannot flush the nutrient-laden water through and self-cleanse.

As other more practical considerations, the extent to which the acceleration of the eutrophication process can be reversed would be expected to depend on the feature of the aquatic ecosystem that is targeted for improvement. Nutrient reductions would exert the most direct effects on plant communities such as freshwater/estuarine phytoplankton or estuarine/coastal macroalgae. Thus, a goal of reduced incidence of algal blooms may be more rapidly achieved than the goal of improved fish communities, because fish growth is indirectly (mediated by changes in habitat) as well as directly (mediated by the food web) affected by the nutrient inputs. The economic feasibility of controls that can be exerted on anthropogenic nutrient sources is also important. Pragmatically, the highest prognosis for success would be expected for natural waters that are affected mostly by sewage and other point (pipe) source nutrient dischargers, because they are much easier to control than nonpoint (diffuse) sources.

One of the most famous early reports of successful reversal of cultural eutrophication involved removing sewage discharges from Lake Washington within metropolitan Seattle, WA, USA (Edmondson, 1970). This large, deep lake (128 km^2, maximum depth 59 m, mean depth 18 m) historically had shown water quality degradation in response to sewage inputs. In 1922, a diversion was created to carry the raw sewage from 30 outfalls away from the lake (into nearby Puget Sound). Algal blooms and fish kills soon abated. However, in 1930, sewage effluents began to be discharged into the lake from treatment plants in outlying communities, and noxious cyanobacteria blooms and hypoxia again increased. About 76 million liters of sewage without inorganic nutrient removal

were discharged daily into the lake by 1962. Nearly a decade of effort by a courageous limnologist, Dr T. Edmondson, led the city to support zero discharge of sewage into the lake by 1968 (Hampton *et al.*, 2006). By 1970, the phytoplankton growth had decreased to levels that had not been seen since the early 1950s, and the improvements in water quality and aesthetics were heralded by the citizenry as a rapid and remarkable recovery. This recovery was aided by the fact that the lake is relatively deep, and in addition the major nutrient sources to be controlled were sewage pipes. It is unlikely that the degree of recovery reversed the water quality and aquatic communities to a *ca*. 1930, more pristine status. However, the partial recovery (partial oligotrophication) represented a major improvement.

Other partial reversals of cultural eutrophication have been reported worldwide. In freshwaters they have been mostly achieved by targeting P reductions (but see below); in estuaries, both P and N have been reduced or "co-managed" for best results. Studies in north temperate estuaries by Fisher *et al.* (1992) and Chesterikoff *et al.* (1992) indicate that if only P – but not N – is removed from sewage inputs to upstream freshwaters, the P removal can actually exacerbate N-related eutrophication problems in downstream estuaries. The freshwater P reductions decrease the riverine algal growth that, if present, would also have removed a substantial portion of the N_i in the sewage before it reached the estuary. Soluble NO_3^-, in particular, is transported downstream at higher concentrations than if freshwater algal blooms had been available to consume it, and increased N_i is thus available to stimulate higher phytoplankton growth in the receiving estuary.

There is increasing recognition of the need to co-manage N and P in freshwaters as well as estuaries and marine coastal systems (Glibert *et al.*, 2011, and references therein).

> As Glibert et al. (2011, p.359) wrote:
>
> "A common practice has been to reduce P in point sources without concomitant reductions in N while overlooking the fact that nutrient loading results in large sediment deposits of nutrients that influence the overall system for an extended period (years) after loading rates are reduced. Reductions in anthropogenic P loads can initially result in a decline in phytoplankton biomass, but the sediment "pump" of stored P replenishes P supplies in the water column, promoting benthic productivity, which, in turn, has multiple effects on the food web. If the system additionally receives N, especially in the form of ammonium (NH_4^+), it can be expected to shift to undesirable dominant species among primary producers, with ramifications extending to higher trophic levels."

P removal without N removal can also lead to major alterations in food web dynamics compared to systems where both N and P are removed, as organisms that are more efficient at sequestering P are favored. When one nutrient is altered relative to another, the stoichiometric balance in the water column is altered, and this change in turn moves through the food web. While less chlorophyll *a* may be supported in the water column, and waters may become clearer, increased non-native SAV growth (particularly invasive species such as *Hydrilla* and *Egeria*) can occur, leading to altered sediment nutrient chemistry, shifts in the dominant grazers, fish, and other changes. Partial reversals can achieve highly desirable results for resource managers and the general public, even if the recovered community is not the same as it was prior to eutrophication. Importantly, however, there can be unintended as well as intended consequences from attempts to partially reverse eutrophication. For example, what is desirable for one constituency may not be for another: clearer water may be more desirable to the general public than floating algal scums, but decreased abundance of certain fish species, altered fish or shellfish community composition that coincide with the decrease in algae and the "clearing" effect may be undesirable for fishermen.

An example of partial reversal in eutrophication that achieved desirable effects for both the general public and fishermen is the case of Lake Erie in the 1960s–1970s (Ashworth, 1986). The degradation of fish communities in Lake Erie reached its most extreme level in the 1960s from a combination of eutrophication, overexploitation of fishery resources, extensive habitat modification, and other pollution. Beginning in the 1970s, fishery management strategies and pollution abatement programs contributed to a dramatic reversal. Lake Erie walleye fisheries rebounded to world-class status, and point-source P loading significantly declined, especially after sewage treatment was improved at the major point source discharger to the west basin of the Lake (the Detroit metropolitan wastewater treatment plant), and after mandated use of detergents without PO_4^{-3}. The P reductions effected in a dramatic decrease in the abundance of nuisance phytoplankton species and of zooplankton biomass, as well as a decline in the abundance of pollution-tolerant oligochaetes and an overall shift in macroinvertebrates to more pollution-intolerant taxa (but see below). Similarly, in the Bay of Quinte on Lake Ontario, P loading reductions after 1977 led to a decline in the abundance and biomass of oligochaete worms, sphaeriid molluscs, isopod crustacea, and some chironomids. Dominance in both chironomid and oligochaete communities changed toward species less tolerant of eutrophic conditions.

Nutrient reductions have led to success stories in estuaries and coastal lagoons. For example, industrial point source nutrient reductions to the Seto Inland Sea in Japan during the 1980s led to a noticeable decline in the frequency and magnitude of toxic dinoflagellate blooms (Okaichi, 1997). Nutrient loadings to Cockburn Sound, Australia, from industrial point sources, and reduction in nutrient loadings to Chesapeake Bay from sewage point sources have decreased phytoplankton blooms and promoted an increase in seagrass meadow habitat in some reaches of this estuary (Burkholder *et al.*, 2007, and references therein). Sewage discharges to the shallow estuary, Mumford Cove, in Connecticut, USA, were rerouted to another waterway in the late 1980s, and within two years massive nuisance blooms of the macroalga, *Ulva lactuca*, were eliminated (Harlin, 1993). The lagoonal New River Estuary, North Carolina, was one of the most eutrophic estuaries in the southeastern US during the 1980s–1990s. Because of high nutrient loads from a major military base and industrialized animal production, the estuary was characterized by dense harmful algal blooms, bottom-water anoxia and hypoxia, and fish kills. Upgrades in wastewater treatment systems during the late 1990s led to 57% and 71% reductions in N and P, respectively, to the estuary from point sources. Mallin *et al.* (1993, 2005) described the clear improvements as a result: major winter blooms of the harmful dinoflagellate *Prorocentrum minimum* disappeared, phytoplankton biomass

(as chlorophyll *a*) significantly decreased, inorganic N and P significantly decreased, and bottom-water DO and water clarity significantly increased, thereby improving habitat for the growth of sessile shellfish and seagrasses. Thus, partial "oligotrophication" from removal of sewage can have fairly rapid, positive results even in shallow systems where the sediments could provide some degree of nutrient replenishment to the overlying water.

Standing in marked contrast to these successes, however, are other situations such as the sobering recent developments regarding the noxious green macroalga, *Cladophora* in the Laurentian Great Lakes (see Burkholder 2009, and references therein). In the 1960s, benthic *Cladophora glomerata* (likely mixed with other *Cladophora* species) focused international attention especially on the west basin of Great Lake Erie, where it proliferated especially in response to P pollution from sewage and agricultural wastes. Its growth in relatively shallow nearshore waters appeared to be limited only by its relatively high light optimum for photosynthesis. Even so, its major seasonal die-offs that drifted into shore in rotting masses sometimes were measured in tons of fresh weight. The adverse effects of *Cladophora* overgrowth in Lakes Erie, Ontario, Michigan, and Huron included displacement of beneficial benthic plants, reductions in invertebrate densities and fish spawning, reduction in biodiversity, anoxia from decaying mats that results in killing of other aquatic life, retention of pathogenic microbes such as *Escherichia coli* and enterococci fecal bacteria as a potential human health hazard, negative effects on recreational use of affected shorelines and nearshore areas, and depressed property values of adjacent homes. Concerted multibillion dollar efforts that banned P use in detergents and improved P removal in sewage treatment during the 1970s led to a dramatic decrease in *Cladophora* in the Great Lakes by the early 1980s.

The success was painfully short-lived. Dissolved PO_4^{-3} concentrations again increased in nearshore lake waters since the late 1980s from altered sediment P fluxes (that is, increased P coming up from the sediments) due to altered benthic biogeochemistry and feces and other metabolic wastes of massive invasions of exotic zebra mussels (*Dreissena polymorpha*) and quagga mussels (*Dreissena bugensis*), which now dominate the nearshore benthic environment along with *Cladophora*. In fact, the macroalgal biomass is again similar to the situation sustained by the Great Lakes during the 1960s–1970s. Through their filtering activity, the huge populations of exotic invasive mussels also increase water clarity during the spring growing season for *Cladophora*, while their excreta, together with the high algal productivity/decomposition, alter sediment P fluxes and fertilize the benthic and shallow-water nearshore habitats. Importantly, removal of P from external loads tends to accelerate the dynamics of P recycling from the benthos into the water column, and can result in a "cascade" of ecosystem changes. *Cladophora* growth now covers a broader range of depths, an expansion that has been attributed to the increased available light and within-lake P supplies. Thus, and ironically, within-lake P loads contributed by the exotic invaders and the altered sediment P fluxes are now even more important than external P loads in sustaining massive *Cladophora* blooms that have returned to the Great Lakes.

In a recent review, Duarte *et al.* (2008) assessed the recovery of a group of ecosystems after nutrient removal, found that the trajectories of response are variable and complex, and attributed this variation to "shifting baselines" caused by other changes such as exotic species invasions, extinctions, overfishing, and climate change. It was concluded that because of these additional changes, reductions in nutrient loads alone cannot be expected to allow systems to return to conditions that existed decades earlier. Reductions in P or N have reduced total phytoplankton biomass in many systems, but such shifts to an alternate stable state do not necessarily return the system to its pre-eutrophic state because the system may be more susceptible to other stressors such as exotic species invasions. Reduction of one nutrient but not the other, changes the N : P ratio which, in turn, alters metabolism, species composition across trophic levels, and food webs (Glibert *et al.*, 2011, and references therein). Thus, from an overall ecosystem standpoint, the extent to which aquatic habitats can be "reversed" in the cultural eutrophication process depends on a more complex "endpoint" than variables such as reductions in nuisance algal growth, recovery of beneficial macrophyte beds, or increases in desirable fish species. The full range of chronic impacts from variables that frequently accompany nutrient enrichment remain to be resolved before the full extent of impacts from cultural eutrophication on aquatic ecosystems, and the potential degree of reversing those impacts, can be evaluated.

See also: Agriculture, Nutrient Management, and Water Quality. Biodiversity of Harmful Marine Algae. Estuarine Ecosystems. Freshwater Ecosystems, Human Impact on. Lake and Pond Ecosystems. Marine and Aquatic Communities, Stress from Eutrophication. Marine Ecosystems. Plankton, Status and Role of. River Ecosystems. Seagrasses. Trophic Cascades

References*

Ashworth W (1986) *The Late, Great Lakes, an Environmental History*. New York: Alfred A. Knopf.

Barrett SCH (1989) Waterweed invasions. *Scientific American* 261: 90–97.

Boesch DF (2002) Challenges and opportunities for science in reducing nutrient over-enrichment of coastal ecosystems. *Estuaries* 25: 886–900.

Bothwell ML and Kilroy C (2011) Phosphorus limitation of the freshwater benthic diatom *Didymosphenia geminata* determined by the frequency of dividing cells. *Freshwater Biology* 56: 565–578.

Brooks JL and Dodson SI (1965) Predation, body size, and composition of plankton. *Science* 50: 28–35.

Burkholder JM (1998) Implications of harmful microalgae and heterotrophic dinoflagellates in management of sustainable marine fisheries. *Ecological Applications* 8(supplement 1): S37–S62.

Burkholder JM (2009) Harmful algal blooms. In: Likens GE (ed.) *Encyclopedia of Inland Waters*, vol. 1, pp. 264–285. Oxford, UK: Elsevier.

*Burkholder JM, Glibert PM, and Skelton HM (2008) Mixotrophy, a major mode of nutrition for harmful algal species in eutrophic waters. *Harmful Algae* 8: 77–93.

Burkholder JM, Gordon AS, Moeller PD, *et al.* (2005) Demonstration of toxicity to fish and to mammalian cells by Pfiesteria species: Comparison of assay methods and multiple strains. *Proceedings of the National Academy of Sciences (U.S.A)* 102: 3471–3476.

Burkholder JM, Mason KM, and Glasgow HB (1992) Water-column nitrate enrichment promotes decline of eelgrass *Zostera marina*: Evidence from seasonal mesocosm experiments. *Marine Ecology Progress Series* 81: 163–178.

*Asterisk indicates comprehensive secondary references on aspects of this topic; other references were cited in the text, as well.

*Burkholder JM, Tomasko D, and Touchette BW (2007) Seagrasses and eutrophication. *Journal of Experimental Marine Biology and Ecology* 350: 46–72.

*Camargo JA and Alonso A (2006) Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment. *Environment International* 32: 831–849.

*Carpenter SR (1988) *Complex Interactions in Lake Communities*. New York: Springer-Verlag.

Carlson RE (1977) A trophic state index for lakes. *Limnology and Oceanography* 22: 361–369.

Chesterikoff A, Garban B, Billen G, and Poulin M (1992) Inorganic nitrogen dynamics in the River Seine downstream from Paris (France). *Biogeochemistry* 17: 147–164.

Cloern JE (2001) Our evolving conceptual model of the coastal eutrophication problem. *Marine Ecology Progress Series* 210: 223–253.

Caraco NF (1995) Influence of human populations on P transfers to aquatic systems: A regional scale study using large rivers. In: Tiessen H (ed.) *Phosphorus in the Global Environment*, pp. 235–247. New York: John Wiley & Sons Ltd. SCOPE 54.

Cottingham KL and Carpenter SR (1994) Predictive indices of ecosystem resilience in models of north temperate lakes. *Ecology* 75: 2127–2138.

Diaz RJ (2001) Overview of hypoxia around the world. *Journal of Environmental Quality* 30: 275–281.

Downing JA (1997) Marine nitrogen: Phosphorus stoichiometry and the global N : P cycle. *Biogeochemistry* 37: 237–252.

Duarte CM, Conley DJ, Carstensen J, and Sánchez-Camacho M (2008) Return to Neverland: Shifting baselines affect eutrophication restoration targets. *Estuaries and Coasts* 32: 29–36.

Edmondson WT (1970) Phosphorus, nitrogen, and algae in Lake Washington after diversion of sewage. *Science* 169: 690–691.

*Elser JJ, Bracken MES, Cleland EE, *et al.* (2007) Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters* 10.http://dx.doi.org/10.1111/j.1461-0248.2007.01113x.

Filice FP (1959) The effects of wastes on the distribution of bottom invertebrates in San Francisco Bay estuary. *Wasmann Journal of Biology* 17: 1–17.

Fisher TR, Peele ER, Ammerman JW, and Harding Jr. LW (1992) Nutrient limitation of phytoplankton in Chesapeake Bay. *Marine Ecology Progress Series* 82: 51–63.

Glibert PM, Burkholder JM, Kana TM, Alexander J, Skelton H, and Shillings C (2009) Grazing by *Karenia brevis* on *Synechococcus* enhances growth and may help to sustain blooms. *Aquatic Microbial Ecology* 55: 17–30.

*Glibert PM (2010) Long-term changes in nutrient loading and stoichiometry and their relationships with changes in the food web and dominant pelagic fish species in the San Francisco Estuary, California. *Reviews in Fisheries Science* 18: 211–232.

*Glibert PM, Fullerton D, Burkholder JM, Cornwell JC, and Kana TM (2011) Ecological stoichiometry, biogeochemical cycling, invasive species and aquatic food webs: San Francisco Estuary and comparative systems. *Reviews in Fisheries Science* 19(4): 358–417.

Granéli E and Johansson N (2003) Increase in the production of allelopathic substances by Prymnesium parvum cells grown under N- or P-deficient conditions. *Harmful Algae* 2: 135–145.

Guillette Jr. LJ and Edwards TM (2005) Is nitrate an ecologically relevant endocrine disruptor in vertebrates? *Integrative and Comparative Biology* 45: 19–27.

Hahn S, Bauer S, and Klaassen M (2008) Quantification of allochthonous nutrient input into freshwater bodies by herbivorous waterbirds. *Freshwater Biology* 53: 181–193.

Hampton SE, Scheuerell MD, and Schindler DE (2006) Coalescence in the Lake Washington story: Interaction strengths in a planktonic food web. *Limnology and Oceanography* 51: 2042–2051.

*Harlin MM (1993) Changes in major plant groups following nutrient enrichment. In: McComb AJ (ed.) *Eutrophic Shallow Estuaries and Lagoons*, pp. 173–187. Boca Raton: CRC Press, Inc.

Harper D (1992) *Eutrophication of Freshwaters – Principles, Problems and Restoration*. New York: Chapman and Hall.

*Hecky PE and Kilham P (1988) Nutrient limitation of phytoplankton in freshwater and marine environments: A review of recent evidence on the effects of enrichment. *Limnology and Oceanography* 33: 796–822.

Heisler J, Glibert P, Burkholder J, *et al.* (2008) Eutrophication and harmful algal blooms: A scientific consensus. *Harmful Algae* 8: 3–13.

Houser JN and Richardson WB (2010) Nitrogen and phosphorus in the Upper Mississippi River: Transport, processing, and effects on the river ecosystem. *Hydrobiologia* 640: 71–88.

*Howarth RW (1988) Nutrient limitation of net primary production in marine ecosystems. *Annual Review of Ecology and Systematics* 19: 89–110.

Jacobson LM, David MB, and Drinkwater LE (2011) A spatial analysis of phosphorus in the Mississippi River basin. *Journal of Environmental Quality* 40: 931–941.

Jeong HJ, Park JY, Nho JH, *et al.* (2005) Feeding by red-tide dinoflagellates on the cyanobacterium *Synechococcus*. *Aquatic Microbial Ecology* 41. 1331–2143.

Jeong HJ, Yoo YD, Kim JS, Seong KA, Kang NS, and Kim TH (2010) Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. *Ocean Science Journal* 45: 65–91.

Kautsky H (1991) Influence of eutrophication on the distribution of phytobenthic plant and animal communities. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 76: 423–432.

Kilham SS, Kreeger DA, Goulden CE, and Lynn SG (1997) Effects of algal food quality on fecundity and population growth rates of *Daphnia*. *Freshwater Biology* 38: 639–647.

Lapointe BE (1997) Nutrient thresholds for bottom-up control of macroalgal blooms on coral reefs in Jamaica and southeast Florida. *Limnology and Oceanography* 42: 1119–1131.

Lapointe BE, Barile PJ, Littler MM, and Littler DS (2005) Macroalgal blooms on southeast Florida coral reefs: II. Cross-shelf discrimination of nitrogen sources indicates widespread assimilation of sewage nitrogen. *Harmful Algae* 4: 1106–1122.

Mallin MA, Johnson VL, Ensign SH, and MacPherson TA (2006) Factors contributing to hypoxia in rivers, lakes, and streams. *Limnology and Oceanography* 51: 690–701.

Mallin MA, McIver MR, Wells HA, Parsons DC, and Johnson VL (2005) Reversal of eutrophication following sewage treatment upgrades in the New River Estuary, North Carolina. *Estuaries* 28: 750–760.

Mallin MA, Paerl HW, Rudek J, and Bates PW (1993) Regulation of estuarine primary production by watershed rainfall and river flow. *Marine Ecology Progress Series* 93: 199–203.

Manny BA, Johnson WC, and Wetzel RG (1994) Nutrient additions by waterfowl to lakes and reservoirs: Predicting their effects on productivity and water quality. *Hydrobiologia* 279/280: 121–132.

Marshall HG, Hargraves PE, Burkholder JM, *et al.* (2006) Taxonomy of *Pfiesteria* (Dinophyceae). *Harmful Algae* 5: 481–496.

McComb AJ and Lukatelich RJ (1995) The Peel-Harvey estuarine system, Western Australia. In: McComb AJ (ed.) Eutrophic Shallow Estuaries and Lagoons, pp. 5–18. Boca Raton, FL: CRC Press.

McGurk MD, Landry F, Tang A, and Hanks CC (2006) Acute and chronic toxicity of nitrate to early life stages of lake trout (*Salvelinus namaycush*) and lake whitefish (*Coregonus clupeaformis*). *Environmental Toxicology and Chemistry* 25: 2187–2196.

Michell F (1999) Eutrophication, fisheries, and consumer-resource dynamics in marine pelagic ecosystems. *Science* 285: 1396–1398.

Moeller PDR, Beauchesne KR, Christopher SJ, Riggs-Gelasco P, and Gelasco AK (2007) Metal complexes and free radical toxins produced by *Pfiesteria piscicida*. *Environmental Science and Technology* 41: 1166–1172.

Neckles HA, Wetzel RL, and Orth RJ (1993) Relative effects of nutrient enrichment and grazing on epiphyte-macrophyte (*Zostera marina* L.) dynamics. *Oecologia* 93: 285–293.

Okaichi T (1997) Red tides in the Seto Inland Sea. In: Okaichi T and Yanagi Y (eds.) *Sustainable Development in the Seto Inland Sea – From the Viewpoint of Fisheries*, pp. 251–304. Tokyo, Japan: Tera Sci. Publ. Co.

Parsons ML, Dortch Q, and Turner RE (2002) Sedimentological evidence of an increase in *Pseudo-nitzschia* (Bacillariophyceae) abundance in response to coastal eutrophication. *Limnology and Oceanography* 47: 551–558.

Phillips GL, Eminson D, and Moss B (1978) A mechanism to account for macrophyte decline in progressively eutrophicated freshwaters. *Aquatic Botany* 4: 103–126.

*Rhee G-Y (1982) Effects of environmental factors on phytoplankton growth. In: Marshall KC (ed.) *Advances in Microbial Ecology*, pp. 33–74. New York: Plenum Press.

Rhyther JH (1989) Historical perspective of phytoplankton blooms on Long Island and the green tides of the 1950s. In: Cosper EM, Bricelj VM, and Carpenter EJ

(eds.) *Novel Phytoplankton Blooms*, pp. 375–383. New York: Springer-Verlag. Coastal and Estuarine Studies No. 35.
Roth NC and Pregnall AM (1988) Nitrate reductase activity in Zostera marina. *Marine Biology* 99: 457–466.
Sheath RG and Burkholder JM (1985) Characteristics of softwater streams in Rhode Island. II. Composition and seasonal dynamics of macroalgal communities. *Hydrobiologia* 128: 109–118.
*Schindler DW (1987) Detecting ecosystem responses to anthropogenic stress. *Canadian Journal of Fisheries & Aquatic Sciences* 44(supplement 1): 6–25.
Schelske CL, Stoermer EF, Fahnenstiel GL, and Haibach M (1986) Phosphorus enrichment, silica utilization, and biogeochemical silica depletion in the Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 43: 407–415.
Siewicki TC, Pullaro T, Pan W, McDaniel S, Glenn R, and Steward J (2007) Models of total and presumed wildlife sources of fecal coliform bacteria in coastal ponds. *Journal of Environmental Management* 82: 120–132.
Smayda TJ (1989) Primary production and the global epidemic of phytoplankton blooms in the sea: A linkage? In: Cosper EM, Bricelj VM, and Carpenter EJ (eds.) *Novel Phytoplankton Blooms*, pp. 449–484. New York: Springer-Verlag. Coastal and Estuarine Studies No. 35.
Smil V (2001) *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food*. Cambridge: The MIT Press.
Smith VH (1983) Low nitrogen to phosphorus ratios favor dominance by blue-green algae in lake phytoplankton. *Science* 221: 669–671.
*Smith VH (2006) Responses of estuarine and coastal marine phytoplankton to nitrogen and phosphorus enrichment. *Limnology and Oceanography* 51: 377–384.
Starr M, Himmelman JH, and Therriault J-C (1990) Direct coupling of marine invertebrate spawning with phytoplankton blooms. *Science* 247: 1071–1074.
Stanley EH and Maxted JT (2008) Changes in the dissolved nitrogen pool across land cover gradients in Wisconsin streams. *Ecological Applications* 18: 1579–1590.
*Tilman D (1977) Resource competition between planktonic algae: An experimental and theoretical approach. *Ecology* 58: 338–348.
*Tilman D, Kilham SS, and Kilham P (1982) Phytoplankton community ecology: The role of limiting nutrients. *Annual Review of Ecology and Systematics* 13: 349–372.
*Touchette BW and Burkholder JM (2000) Overview of the physiological ecology of carbon metabolism in seagrasses. *Journal of Experimental Marine Biology and Ecology* 250: 169–205.
Turner RE, Qureshi N, Rabalais NN, *et al.* (1998) Fluctuating silicate: Nitrate ratios and coastal plankton food webs. *Proceedings of the National Academy of Sciences (USA)* 95: 13048–13051.
Turpin DH (1991) Effects of inorganic N availability on algal photosynthesis and carbon metabolism. *Journal of Phycology* 27: 14–20.
Valentine JF and Duffy JE (2006) The central role of grazing in seagrass ecology. In: Larkum AWD, Orth RJ, and Duarte CM (eds.) *Seagrasses: Biology, Ecology and Conservation*, pp. 463–501. Dordrecht, the Netherlands: Springer.
Vasas V, Lancelot C, Rousseau V, and Jordán F (2007) Eutrophication and overfishing in temperate nearshore pelagic food webs: a network perspective. *Marine Ecology Progress Series* 336: 1–14.
*Wetzel RG (2001) *Limnology*, 3rd edn. New York, NY: Academic Press.
Yarrow M, Marin VH, Finlayson M, Tironi A, Delgado LE, and Fischer F (2009) The ecology of Egeria densa Planchón (Liliopsida: Alismatales): A wetland ecosystem engineer? *Revista Chilena de Historia Natural* 82: 299–313.

Evaluation of Ecosystem Service Policies from Biophysical and Social Perspectives: The Case of China

Jianguo Liu, Michigan State University, East Lansing, MI, USA
Zhiyun Ouyang, Chinese Academy of Sciences, Beijing, China
Wu Yang, Michigan State University, East Lansing, MI, USA
Weihua Xu, Chinese Academy of Sciences, Beijing, China
Shuxin Li, Michigan State University, East Lansing, MI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Cost-effectiveness analysis (CEA) A kind of economic analysis that allows comparison of the relative effects and costs of two or more actions. Typically the CEA is expressed as the ratio of effects to costs. The effects in CEA do not need to be in monetary values, which is the distinction from cost-benefit analysis (CBA) in which the benefits are assigned with monetary values.

Coupled human and natural systems (CHANS) Integrated systems in which human components interact with natural components.

Ecosystem services Also referred as environmental services, ecological services, or ecosystem goods and services, ecosystem services are the benefits obtained directly and indirectly from nature, ranging from freshwater, timber, non-timber forest products, through carbon sequestration, water and soil conservation, to tourism and nutrient cycling.

Interactive effect Refers to the effect of an interaction term in statistical analysis, a situation in which the simultaneous effect of two variables on a third one is not additive. In other words, the effect of one variable on another variable depends on a third variable.

Payments for ecosystem services (PES) Also known as payments for environmental services or benefits, PES is a kind of conservation practice that offer incentives (e.g., cash, grain) to participants (e.g., farmers) in exchange for their forgone economic benefits or management efforts to protect one or many types of ecosystem service.

Introduction

China is endowed with immense reserves of natural capital and ecosystem services that flow from it (Ouyang, 2007; Task Force for Eco-Compensation Mechanisms and Polices in China, 2007). Unfortunately, natural capital in many areas has been degraded or lost due to a variety of reasons, such as rapid economic development (Task Force for Eco-Compensation Mechanisms and Polices in China, 2007), increasing human population size (Liu and Diamond, 2005), even faster household proliferation (Liu *et al.*, 2003a), and inappropriate governance (Liu and Diamond, 2008; Liu and Raven, 2010). The resulting degradation and loss of ecosystem services have contributed to large disasters, such as massive flooding in 1998, and other huge impacts on human well-being (e.g., economic losses and harm to human health) (Millennium Ecosystem Assessment, 2005). Having realized these problems, China has developed and implemented a series of large-scale policies to protect and restore natural capital and ecosystem services.

These policies include the Key Shelterbelt Construction Program, Beijing–Tianjin Sandstorm Control Program, Nature Reserve System (NRS) (Liu *et al.*, 2008; Ouyang, 2007), Forest Eco-Compensation Program, Grassland Eco-Compensation Program, Wetland Restoration Program, Natural Forest Conservation Program (NFCP; also known as Natural Forest Protection Program), Grain-to-Green Program (GTGP; also known as the Farm-to-Forest Program or the Sloping Land Conversion Program) (Liu *et al.*, 2008; Ouyang, 2007; Xu *et al.*, 2006a), and Ecosystem Function Zones (EFZs) (Ministry of Environmental Protection and Chinese Academy of Sciences, 2008). These programs have generated important biophysical effects (e.g., biodiversity conservation; and mitigation of climate change, desertification, droughts, floods, soil erosion, and water runoff) and socioeconomic effects (e.g., poverty alleviation and economic development) (Liu and Diamond, 2005; Liu *et al.*, 2008; Xu *et al.*, 2006a). They also have provided significant benefits, such as carbon sequestration, at the global level (Liu and Diamond, 2005; Liu *et al.*, 2008).

The purpose of this article is to evaluate three major programs – NRS, NFCP, and GTGP – from biophysical and social perspectives. These programs are national in scope and have significant global implications. For example, NFCP and GTGP are among the biggest payments for ecosystem service programs in the world (Liu *et al.*, 2008). For each program, this article offers an overview of background and goals, illustrates their biophysical and socioeconomic effects, and discusses future opportunities, challenges, and needs. The programs are presented in the order of their inception times (1956, 1998, and 1999, for NRS, NFCP, and GTGP, respectively).

NRS

Background and Goal

China is among the world's most biodiverse countries, probably behind only Brazil, Colombia, and Indonesia (Liu and Raven, 2010). In the temperate Northern Hemisphere, China has the richest assemblage of biodiversity. For instance, it

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00335-X

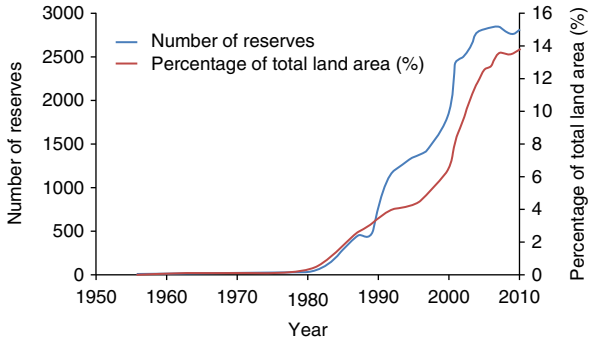


Figure 1 Growth of nature reserves in China. Reproduced from Liu J and Diamond J (2005) China's environment in a globalizing world. *Nature* 435: 1179–1186; Ministry of Environmental Protection (2011) *Bulletin of China's Environmental Status 2010*, and National Bureau of Statistics of China (1985–2010) *China Statistical Yearbooks*. Beijing: China Statistics Press.

contains 16 main biomes according to plant functional types (Liu and Raven, 2010). China has almost 10% of terrestrial vertebrate species (at least 2340 species out of ~25,000 named species worldwide) and ~9% of known vascular plant species (~31,500 out of 350,000 total worldwide) (Liu and Raven, 2010). So far, only a small proportion of species have been recorded. For the eukaryotic species, China may have ~1 million species (out of a global estimate at >12 million (Liu and Raven, 2010)), but <150,000 are actually recorded (Liu and Raven, 2010). However, much of China's biodiversity is under severe threat, such as habitat destruction, environmental pollution, poaching, global climate change, and invasive species (Liu *et al.*, 2003b; Liu and Raven, 2010). For instance, although the number of endangered wild giant pandas is small (~1600) (Viña *et al.*, 2008), the number of remaining wild tigers in China is tiny (~50) ("WWF News," http://www.worldwildlife.org).

To protect biodiversity, China has taken a series of actions (Liu *et al.*, 2003b), including both *ex situ* and *in situ* conservation actions. The former include colony relocation (e.g., National Nature Reserves for Milu (*Elaphurus davidianus*) in Shishou City of Hubei Province and in Dafeng City of Jiangsu Province), botanical gardens, zoos, gene banks, germplasm banks, and breeding facilities (e.g., the Breeding Center for Giant Pandas (*Ailuropoda melanoleuca*) in Wolong Nature Reserve in Sichuan Province). The first germplasm bank in China was established in 2007 at the Institute of Botany of the Chinese Academy of Sciences in Kunming. Plans include the collection of 6450 species of plants by the end of 2012 and 19,000 by the end of 2022 ("China Germplasm Bank of Wild Species," http://www.genobank.org).

The *in situ* conservation actions include policies prohibiting poaching and gathering, restricting economic development, and establishing the NRS. The NRS is particularly impressive. In the half-century since the first reserve, Dinghushan Nature Reserve, was created in 1956 in Guangdong Province, China has designated 2588 nature reserves (through 2010) (**Figure 1**). These nature reserves account for 1.49 million km^2, or ~15% of China's territory (Ministry of Environmental Protection, 2011). The goal of the NRS is to protect biodiversity of all major forms in China.

Distribution

The nature reserves are distributed across the country (**Figure 2**), but the majority of the land area (~56%) is in western China, concentrated in Qinghai Province as well as Tibet and Xinjiang Autonomous Regions, although the numbers in these regions are relatively small (Wu *et al.*, 2011). The nature reserves vary greatly in size, ranging from 0.01 to 298,000 km^2. The largest one is Qiangtang Nature Reserve in Tibet, established in 1993 primarily for the protection of Tibetan antelope (*Pantholops hodgsonii*), wild yaks (*Bos mutus*), and their habitat of desert highlands.

Nature reserves are managed at different institutional levels (i.e., national, provincial, city, and county levels). At the national level, there were 319 nature reserves covering 926,756 km^2 in 2010 (Ministry of Environmental Protection, 2011). These national nature reserves are managed by national government agencies. The majority of them (237) are managed by the State Forestry Administration. The remaining are managed by the Ministry of Environmental Protection (46), State Oceanic Administration (12), Ministry of Agriculture (9), Ministry of Land and Resources (11), Ministry of Water Resources (3), and Chinese Academy of Sciences (1) (Ministry of Environmental Protection, 2010).

Most nature reserves are managed by provincial and county governments. Compared to national nature reserves, provincial and local reserves usually receive less support and are less known. For these reasons, many provincial and local governments have put forth much effort to elevate their reserves to the national level (Xu and Melick, 2007).

Although the vast majority of nature reserves are within a particular administrative or political boundary, some cross those boundaries. At the national level, a number of nature reserves cross country boundaries. For example, China and Russia jointly established the Khanka Lake International Nature Reserve for red-crowned crane (*Grus japonensis*) in 1992. China, Russia, and Mongolia cocreated the Dauria International Nature Reserve in 1994. China and Russia are also considering the establishment of more cross-boundary nature reserves to enhance the conservation of the Siberian tiger (*Panthera tigris*), Amur leopard (*Panthera pardus*), and other rare and endangered species ("Xinhua News," http://news.xinhuanet.com).

Biophysical Effects

NRS includes areas for the protection of a variety of geological and paleontological relics, ecosystems, floras, faunas, natural scenery, and natural coastal environments and resources (Ministry of Environmental Protection, 2010). Protected ecosystems include forests, grasslands, desert highlands, inland wetlands, and marine and coastal areas. These ecosystems provide a variety of services such as food and water supply, water and air purification, carbon sequestration, flood and drought regulation, ecotourism and recreation, and refuges for endangered species. Of the total 53 ecoregions represented in China, 29 have more than 10% of their land areas within the current NRS (Wu *et al.*, 2011). Of 502 important bird areas in China, 340 are within the current NRS (Wu *et al.*, 2011). Approximately 81% of China's natural vegetation communities are within at least one nature reserve (Wu *et al.*, 2011).

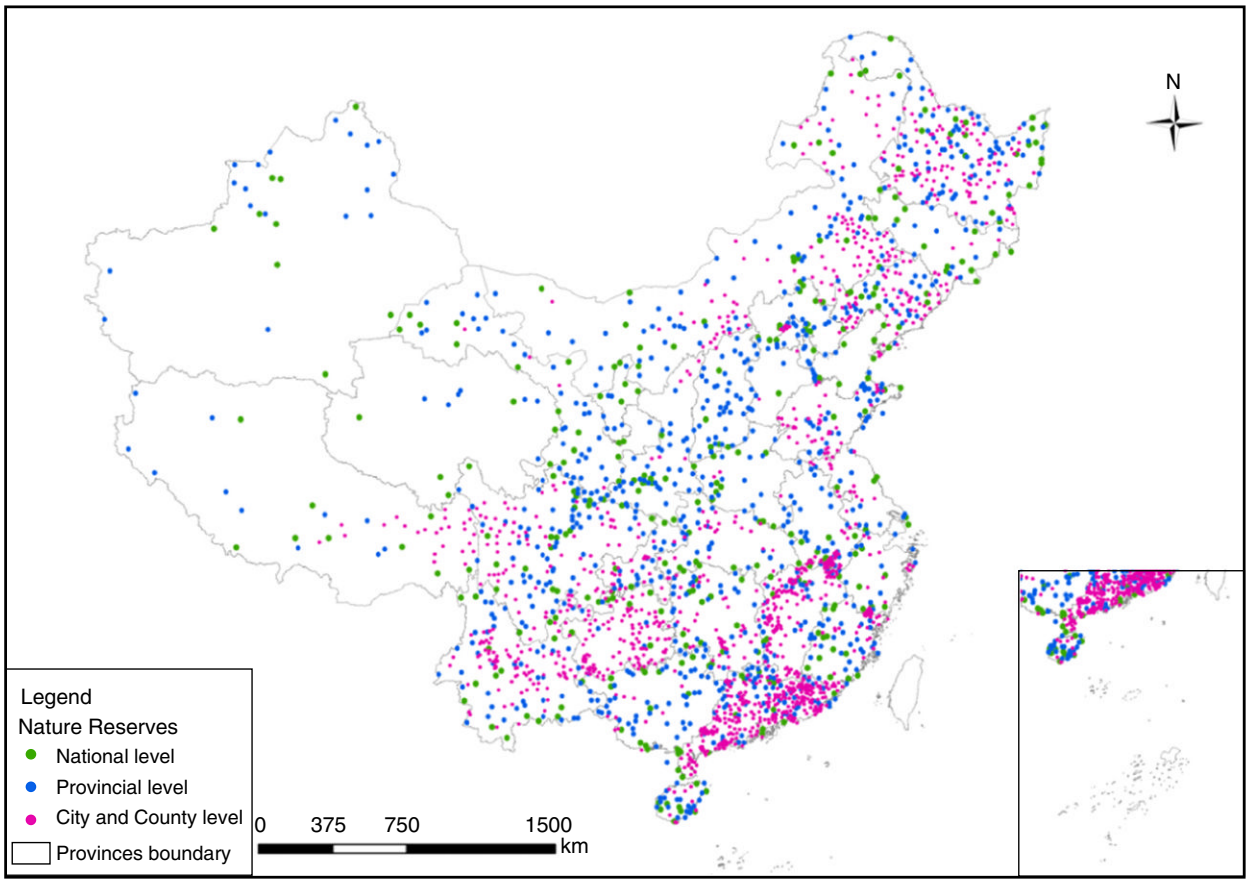


Figure 2 Spatial distribution of nature reserves in China. The map is provided by Nanjing Institute of Environmental Sciences, Ministry of Environmental Protection of China.

Many nature reserves are established with the primary goal of conserving specific species. For example, there are 63 nature reserves whose primary goal is to conserve the endangered giant pandas (Viña *et al.*, 2010). Fifteen nature reserves have been established for Chinese dove tree (*Davidia involucrata*), seven for Chinese white dolphin (*Sousa chinensis*), and 34 for Chinese giant salamander (*Andrias davidianus*) (Ministry of Environmental Protection, 2010). Besides the focal species, many other species and associated ecosystems in these nature reserves also benefit (Viña *et al.*, 2010).

Since 1997, both biophysical and economic benefits of China's ecosystem services have been assessed (Ouyang and Wang, 1997; Ouyang *et al.*, 1999). However, relatively few rigorous ecosystem services assessments of nature reserves in China have been conducted, and many ecosystem services, such as carbon sequestration, are rarely quantified. An ecosystem service assessment of the Changbaishan Nature Reserve (total size 1965 km^2) in Jilin Province, for example, found that annually a total of 105 million m^3 of water were stored, 1.2 million tons of carbon were fixed, and 17,021 tons of nutrients were accumulated (Xue *et al.*, 1999).

Owing to inadequate management, the establishment of nature reserves also has caused some negative biophysical effects. In Nuozadu Nature Reserve of Yunnan Province, almost half of its forests were lost due to illegal extraction and open-access farming with poor management (FCCDP, 1998). Similarly, in Wolong Nature Reserve of Sichuan Province, the panda habitats continued to be lost even after the establishment of the reserve due to rapid population growth and household proliferation, economic development inside the reserve, and extraction of timber and non-timber forest products (Liu *et al.*, 2001). Some studies have also recorded the adverse effects of tourism development (e.g., road construction and hiking) on vegetation, wildlife, and their habitats (Fan *et al.*, 2011; Li *et al.*, 2005; State Forestry Administration, 2006).

Socioeconomic Effects

Globally, the overall benefit–cost ratio of an effective wilderness conservation program is at least 100 (Balmford *et al.*, 2002). In China, taking the Yancheng Biosphere Reserve in Jiangsu Province as an example, a very conservative estimation suggests that the benefit–cost ratio of managing the reserve was 10 without considering the huge revenue from tourism development (Lu *et al.*, 2007). If the benefits of tourism development are taken into account, the benefit–cost ratio would be much larger than 100 for many reserves with tourism (e.g., Jiuzhaigou Nature Reserve in Sichuan Province and Zhangjiajie Nature Reserve in Hunan Province).

The average total annual investment for nature reserves from governments at different levels in China was ~200 million yuan over the past two decades (1 USD = 6.8 yuan as of 2010) (Yuan *et al.*, 2008). The total annual investment per square kilometer was approximately US$53 in the early 2000s, compared to US$2058 and US$157 in developed and other developing countries, respectively (Yuan *et al.*, 2008). Moreover, this number continued to decrease with the rapid expansion of the NRS and relatively slow increase in investment. The gap of insufficient funding was filled through revenue-raising activities such as tourism development and the use of natural resources in the reserves (Ouyang *et al.*, 2002; Xu and Melick, 2007; Yuan *et al.*, 2008).

Since 1982, more than 80% of nature reserves in China have launched tourism development programs (Yuan *et al.*, 2008). These programs have produced socioeconomic impacts on local residents, companies, governments, and tourists inside and outside China. First, every year, tourism development produces billions of yuan in revenue for the government. For example, the Zhangjiajie Nature Reserve in Hunan Province received as many as 1.4 million tourists and created as much as 5.6 billion yuan in revenue in 2010 (Statistic Bureau of Zhangjiajie City, 2011). Of those tourists, 555,500 were from outside China. Second, local residents may also benefit from the nature reserves through government subsidies and participation in tourism activities. In Jiuzhaigou Natural Reserve of Sichuan Province, all of the local households have intensively participated in and benefited from tourism business, with the per capita income increasing by more than fourfold after the initiation of tourism (Li *et al.*, 2006b). However, studies have shown that in most reserves, local residents had received only a very small proportion of the produced economic benefits, with the major proportion divided by tourism-related companies and the government (He *et al.*, 2008; Xu and Melick, 2007; Yuan *et al.*, 2008). Besides direct economic benefits, tourism development has also dramatically improved the construction of infrastructure (e.g., roads, hospitals, and hotels), has produced job opportunities, and has changed the income structure of local households and governments (He *et al.*, 2008; Li *et al.*, 2006b; Xu and Melick, 2007; Yuan *et al.*, 2008).

The establishment and management of nature reserves also have some negative socioeconomic consequences such as conflicts over land use, tourism resources, and restrictions of natural resource use between the administrative bureaus of reserves and local residents as well as local governments. For instance, the establishment of many reserves incorporates land that had been allocated to local households or communities under the Contract Responsibility System in the early 1980s; however, many reserves were delimited on maps before determining the property rights and compensation measures for land and associated resources (e.g., forests), which unavoidably triggers conflicts over the use and management of land and associated resources (Miao and West, 2004). In the core zone of Wolong Nature Reserve of Sichuan Province, local residents were prohibited from collecting bamboo shoots and Chinese herbal medicines, which were traditional practices before the establishment of the reserve and constituted a large proportion of their income (Liu *et al.*, 1999).

Future Opportunities, Challenges, and Needs

China continues to expand the NRS. By the end of 2010, China had already achieved its designed goal for 2050 in terms of the number of nature reserves (2500), but had not yet met the goal for 2010 in terms of the percentage of total land area (16.1%, 16.8%, and 18.0% as planned for 2010, 2030, and 2050, respectively) (State Forestry Administration, 2001). Although the rapid increase in quantity is impressive, the total cover area and performance of nature reserves in protecting biodiversity are what really matter. Future expansion efforts should pay special attention to areas rich in endemic biodiversity, areas vulnerable to human activities and climate change, and connections between some existing nature reserves.

Although expansion efforts are needed, it is crucial to invest more resources in existing nature reserves to enhance their management. The current nature reserves are managed by different administrations at different institutional levels, of which the functions and responsibilities are often ambiguous or conflicting. Many reserves are listed only on paper because they are not effectively managed due to the lack of sufficient funding and low engagement of stakeholders (e.g., local residents). Some of these reserves are even "empty" because conservation targets have become extinct since the establishment of the reserves (Li *et al.*, 2010). For example, in Jiuzhaigou Nature Reserve of Sichuan Province, there is no recent evidence of giant pandas, probably due to the skyrocketing number of tourists (State Forestry Administration, 2006). To improve the effectiveness and efficiency of the NRS, it is necessary to rethink its goal, reevaluate and redefine current nature reserves, and establish an integrated administration.

To make the management of nature reserves most effective, it is critical to improve the well-being of local residents. Although tourism has generated much revenue, local residents receive only a small proportion of the economic benefits (He *et al.*, 2008). Thus, economic benefits should be more fairly distributed to local residents in order to provide incentives for local residents to cooperate for effective nature reserve management.

Another aspect for consideration is the potential impacts of global climate change. Although there has been discussion of the potential impacts of global climate change on biodiversity in general, little quantitative work has been conducted in China. Such quantitative research is a foundation to inform the development and implementation of effective climate adaptation strategies.

Information on the biophysical and socioeconomic effects of nature reserves has been largely fragmented so far. Much information (e.g., detailed environmental monitoring and financial data) is held by government agencies and is inaccessible to the public. Thus, systematic and integrated studies as well as timely and transparent information disclosure are needed to better conserve China's biodiversity and secure the provision of ecosystem services.

NFCP

Background and Goal

From the 1950s to 1990s, China's natural forests declined in area by 70% and stocking of natural forests per unit area



Figure 3 Spatial distribution of NFCP and GTGP. Reproduced from Liu J, Li S, Ouyang Z, Tam C, and Chen X (2008) Ecological and socioeconomic effects of China's policies for ecosystem services. *Proceedings of the National Academy of Sciences USA* 105: 9477–9482.

declined by 32% (Zhang *et al.*, 2000). The decline of natural forests was mainly due to the population explosion and soaring demand for timber. Human population in forested areas increased fivefold and timber harvests increased threefold (Zhang *et al.*, 2000). The dramatic decline in the quantity and quality of natural forests resulted in extensive desertification, soil erosion, floods, droughts, carbon emission, and damage to wildlife habitat, and so forth (Liu and Diamond, 2005). In response to failures of previous forestry policies, the severe droughts in 1997, and the massive floods in 1998, the central government decided to implement a series of conservation policies, including NFCP (Liu *et al.*, 2008; Yin, 2009; Zhang *et al.*, 2000).

The goal of the NFCP is to conserve and restore natural forests through logging bans and afforestation because most of the natural forests in China were lost or degraded due to human activities (e.g., logging) and because deforestation was believed to have contributed to the 1998 massive floods (Liu *et al.*, 2008). Specifically, the NFCP aimed to completely ban logging of natural forests in the upper reaches of the Yangtze and Yellow rivers as well as in Hainan Province (**Figure 3**) by 2000 and substantially reduce logging in other places (Xu *et al.*, 2006a). Furthermore, the NFCP aimed to reduce timber harvests in natural forests from 32 million m^3 in 1997 to 12 million m^3 in 2003, and to afforest 31 million ha by 2010 through aerial seeding, artificial planting, and mountain closure (i.e., prohibition of human activities such as fuelwood collection and grazing) (Xu *et al.*, 2006a).

Distribution

In 1998, right after the large floods, China started the NFCP pilot study in 12 provinces and autonomous regions/municipalities (Liu *et al.*, 2008). By 2000, the NFCP was expanded to 18 provinces, including the upstream regions of major river systems, especially the Yellow and Yangtze rivers, which had suffered massive ecological degradation (**Figure 3**). The NFCP target areas are classified into two priorities: the first priority is the state-owned forests, and the second priority is the community-owned forest region. Under different priorities, the levels of financial support from the central government vary, 20% of all costs for the second priority area and 100% for the first priority area (Liu *et al.*, 2007a). In 2011, NFCP was renewed for the 10-year second phase, which also added another 11 counties around Danjiangkou Reservoir in Hubei and Henan Provinces, the water source for the middle route of the South-to-North Water Diversion Project ("China Internet News Center," http://fangtan.china.com.cn).

Biophysical Effects

It is a common belief that achieving the NFCP's goal can lead to many biophysical benefits, such as soil erosion reduction, water retention, and flood control. Most indicators of biophysical effects studied so far are immediately observable measures, such as changes in harvested timber, newly forested area, and degree of soil erosion (Liu *et al.*, 2008), partially because it takes time for forests to recover or regenerate. However, some recent

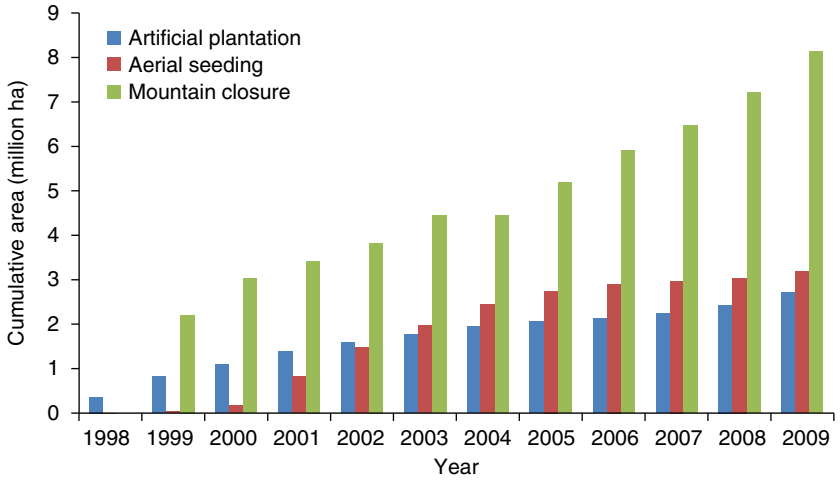


Figure 4 Areas of NFCP by different implementation approaches. Reproduced from State Forestry Administration (2000–2010) *China Forestry Development Report*. Beijing: State Forestry Administration; State Forestry Administration (2001–2009) *Bulletin of Key Forestry Projects*. Beijing: State Forestry Administration, and State Forestry Administration (2006–2010) *Statistical Analysis of National Forestry*. Beijing: State Forestry Administration.

studies have measured changes in forest cover using field data and remotely sensed data (Viña *et al.*, 2011).

Commercial harvesting of natural forests in 13 provinces had stopped, and the unlogged area had amounted to 8.9 million ha by 2000 (Liu *et al.*, 2008). There was a rapid increase in the area under mountain closure and plantation, and the total area had expanded to more than 14 million ha by 2009 (**Figure 4**). Between 2000 and 2010, NFCP protected 107 million ha forest area and increased forest area by 10 million ha ("China Green Times," http://www.greentimes.com).

The NFCP also has reduced soil erosion. For example, during 2000–2007, sediment concentration in Yellow River had declined by 38%. In Chongqing City alone, the total soil erosion area had reduced in 91% of its administrated counties ("China Green Times," http://www.greentimes.com). In Sichuan Province, from 1998 to 2003, the reduced amount of soil erosion was estimated to be 1.5 billion tons (Zhang, 2006).

Forest cover and wildlife habitat also have been recovering. For instance, since the NFCP started in Wolong Nature Reserve of Sichuan Province, illegal harvesting of natural forests rarely occurs (Bearer *et al.*, 2008) and the habitat of the endangered giant pandas has been recovering (Viña *et al.*, 2007). The NFCP also was helpful in reducing the ecological impact of the devastating 2008 Wenchuan earthquake (Viña *et al.*, 2011). Without the NFCP, much more forest cover would have disappeared due to the earthquake. In the Sichuan Giant Panda Sanctuary and Qinling Mountain Area, the percentage of forest cover has increased since the NFCP was initiated (Li, 2010).

In addition, the NFCP also has had significant ecological effects at the global level. One is increased carbon sequestration. Between 1998 and 2010, the NFCP increased carbon sequestration by an estimated 1.3 billion tons and reduced timber harvest by an estimated 220 million m^3 ("China Green Times," http://www.greentimes.com). However, reduced domestic timber production has prompted an increase in the import of timber from other countries (Liu and Diamond, 2005). In 2005, for example, 29.4 million m^3 of logs were imported to China, an increase of 10.4% from 2004. Twenty-five percent (or 7.4 million m^3) of these logs were from tropical forests.

Socioeconomic Effects

NFCP-related activities received a total commitment of 93.7 billion yuan between 1998 and 2009 (**Figure 5**). The central government contributed ~80% of this amount, whereas local governments provided the remainder. Most of the money was used to offset economic losses of forest enterprises caused by the transformation from logging to tree plantations and forest management (Liu *et al.*, 2008; Ouyang, 2007). The payments varied with specific tasks: aerial seeding (750 yuan ha^{-1}), artificial planting (3000 and 4500 yuan ha^{-1} in the Yangtze and Yellow river basins, respectively), regenerating forests through mountain closure (1050 yuan ha^{-1}), and protecting forest (10,000 yuan per worker per 340-ha forest) (Liu *et al.*, 2008; Xu *et al.*, 2006a).

Important measures have been taken to create alternative jobs for people previously in forest enterprises and have largely changed forestry's economic and employment structure. By the end of 2002 (3 years after NFCP implementation), approximately two-thirds of the 1.2 million logging and processing workers impacted by the NFCP had retired or had been transferred to other sectors (Liu *et al.*, 2008; Xu *et al.*, 2006a). The average income from the tertiary industry (e.g., hotels, restaurants, and entertainment) in 32 forest enterprises jumped from 8.5% in 1997 to 20.1% in 2003 (Liu *et al.*, 2008; Zhang, 2006).

Forest management and plantation farming have become the dominant sources of employment since logging was stopped. For example, the proportion of the staff for forest management in the Chuannan Forestry Bureau and Ebian County of Sichuan Province increased from 0% and 13.1% in 1997 to 52.6% and 76.7% in 2001, respectively (Liu *et al.*, 2005, 2008). Even though the income from forestry in some areas such as Longmin Township in Pengzhou County of Sichuan Province declined, there was an increase in the total income, thanks to income from other sources such as tourism (Qiao *et al.*, 2006).

Government data from 35 key state-owned forest enterprises showed that the gross forestry product had continued to

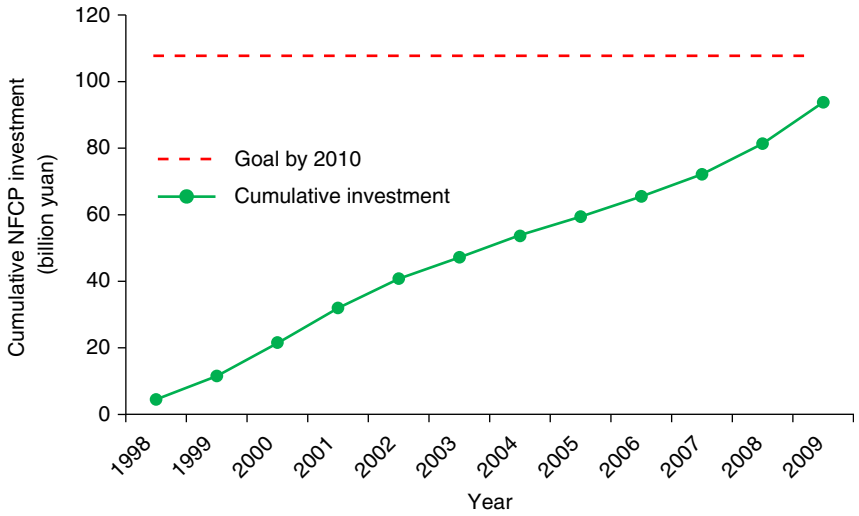


Figure 5 Investment in NFCP. Reproduced from State Forestry Administration (2000–2010) *China Forestry Development Report*. Beijing: State Forestry Administration; State Forestry Administration (2001–2009) *Bulletin of Key Forestry Projects*. Beijing: State Forestry Administration, and State Forestry Administration (2006–2010) *Statistical Analysis of National Forestry*. Beijing: State Forestry Administration.

decline from 1997 to 2000, and then began to increase in 2001. On average, the annual income of employees within NFCP zones had increased from 4437 yuan in 2000 to 12,645 yuan in 2008. Taking the largest NFCP implementation province, Sichuan Province, as an example, the gross forestry product increased more than 10 times from 1997 to 2009. Annual income from forest farmers also increased more than three times from 2000 to 2009 ("China Green Times," http://www.greentimes.com). However, independent studies by scholars have also found adverse effects of the NFCP. For example, in Sichuan Province, 1172 wood-related industry enterprises and 154,000 employees who depended on income from timber harvesting were negatively affected (Zhou, 2006). The tertiary industry in the Chuannan Forestry Bureau of Sichuan Province suffered a big reduction of income (from 5.0 million yuan in 1997 to 1.1 million yuan in 2001) as a result of reduced wood-related activities (Liu and Zhou, 2005). Approximately 55,000 people in Taijiang County of Guizhou Province had a loss of 6 million yuan, placing some forestry workers below the poverty line (Yang, 2004). Some enterprises could not pay back their loans or pay salaries (Huang, 2005). By 2001, these loans amounted to 12.9 billion yuan, and unpaid salaries reached 860 million yuan.

Local governments also had budgetary burdens because they lost revenue from wood-related industries and were responsible for providing matching funds for the NFCP at the same time (Zhou, 2006). For instance, from 1998 to 2001, the Yanbian County Forestry Bureau of Sichuan Province lost 9.7 million yuan in revenue and had to provide matching funds for NFCP (13% of total investment from the central government) (Liu *et al.*, 2005). In northwestern China, 34.9%, 47.0%, and 59.8% of farmers, livestock grazers, and forest workers, respectively, reported on a survey that their livelihoods had been ruined by the implementation of the NFCP, and the poorer the respondents were, the greater the likelihood they believed they had suffered (Cao *et al.*, 2010).

Future Opportunities, Challenges, and Needs

For the second phase (2011–2020), the NFCP plans to increase forest cover by 5.2 million ha, capture 416 million tons of carbon, provide 648,500 forestry jobs, further reduce soil erosion, and enhance biodiversity ("China Internet News Center," http://fangtan.china.com.cn).

In response to the financial burden on local governments of matching ~20% of central government funds during the first phase, that requirement has been eliminated for the second phase of the NFCP for western provinces. The planned cumulative investment for the second phase is 244.0 billion yuan, with 219.5 billion from the central government and 24.5 billion from local investment ("China Internet News Center," http://fangtan.china.com.cn). Although this adjustment may allow local governments to spend more money on local socioeconomic development, it potentially increases the burden of the central government and the financial risk of the program. The program benefits many stakeholders such as hydropower stations, relevant companies and business sectors, people in the Yangtze and Yellow river basins, and even other countries (e.g., Japan, South Korea, and the US would benefit from the mitigation of sandstorms) (Liu *et al.*, 2008). Thus, market-based mechanisms could be attempted to reduce such financial burdens and risks so as to sustain the program in the long run.

The second phase also increases the payment levels for artificial plantation, mountain closure, aerial seeding, and forest monitoring and management. The payment scheme has been adjusted from a fixed payment level through the year to a payment level that will be adjusted with inflation. However, this planned payment scheme still ignores the spatial heterogeneity of socioeconomic conditions across NFCP implementation areas. Increasing the payment levels simultaneously for all implementation areas may cause overpayment in some areas, whereas some areas still receive unreasonably low payments. To improve the fairness and efficiency of the program,

the payment scheme should consider the spatial heterogeneity of different implementation areas (Chen *et al.*, 2010).

The second phase also emphasizes improvement in people's livelihoods. It plans to improve social insurance (including pension, medical care, unemployment, injury, and maternity), housing conditions, and job opportunities of formal employees in state-owned forest enterprises. Although the increase in NFCP payment levels may also benefit other NFCP-related workers who are not formal employees in state-owned forest enterprises, it will not improve those people's social insurance, housing conditions, and job opportunities. This discriminatory policy not only is unethical but may also induce many social conflicts, because a large number of NFCP-related workers do the same work as, if not more than, formal employees but are not registered as formal employees under the current forestry institutional system.

In addition, many reports on forest recovery have been based only on the statistics of government agencies. More scientifically sound and quantitative studies are needed to more rigorously evaluate the impact of the NFCP on forest quantity and quality changes in both the short and long terms. For example, advanced techniques such as remote sensing are available for comprehensive monitoring (Viña *et al.*, 2011). The complexity of policy implementation in coupled human and natural systems should also be considered (Liu *et al.*, 2007b, c). Instead of assessing various conservation programs separately, further studies should also evaluate their interactive effects (Liu *et al.*, 2008). Interdisciplinary studies are also urgently needed to assess the tradeoffs of NFCP implementation on ecosystem services and human well-being.

GTGP

Background and Goal

Much of China's cropland was on steep slopes, causing serious soil erosion. For example, Yangtze and Yellow river basins had almost 4.3 million ha of cropland on slopes of $\geq 25^{\circ}$ (Liu *et al.*, 2008). Although natural forest protection and afforestation efforts of the NFCP are important to reduce soil erosion, another important driving force behind soil erosion is farming on sloping lands. Therefore, a year after the pilot implementation of the NFCP, China initiated the GTGP in 1999. The goal of the GTGP is to convert cropland on steep slopes to forests or grasslands. The main criterion for choosing land plots for inclusion in the GTGP is slope steepness, $\geq 15^{\circ}$ in northwestern China and $\geq 25^{\circ}$ in other parts of the country (Uchida *et al.*, 2005).

The major aim of the GTGP was to increase vegetative cover by 32 million ha by 2010, with 14.7 million ha of cropland conversion to forest and grassland and the remaining portion through afforesting barren land and mountain closure (**Figure 6**) (Xu *et al.*, 2004). It was expected that by 2010 the increase in vegetative cover could control soil erosion across 86.7 million ha and desertification of 102.7 million ha (Yin, 2009). In addition to the primary goal of reducing environmental impacts, the GTGP also aims to alleviate poverty and advance local economic development (Xu and Cao, 2002).

Sloping lands can be converted into ecological and/or economic forests and grasslands, but ecological forests should account for 80% of the total converted land (Liu *et al.*, 2007a). Here, ecological forests refer to forests with high ecological benefits (e.g., carbon sequestration, water and soil retention), while economic forests represent forests with relatively short-term economic returns (e.g., orchards or plantations of trees with medicinal value for commercial use). To offset the lost agricultural revenue due to land conversion, the government provides participating households with a payment in the form of grain and/or cash. The durations of payments are 8 years if the cropland is converted to ecological forests using ecologically important trees, 5 years if it is converted to economic forests using fruit trees, and 2 years if the cropland is converted to grassland (Xu *et al.*, 2004). The payment levels differ by region, with 2250 and 1500 kg of grain, or 3150 and 2100

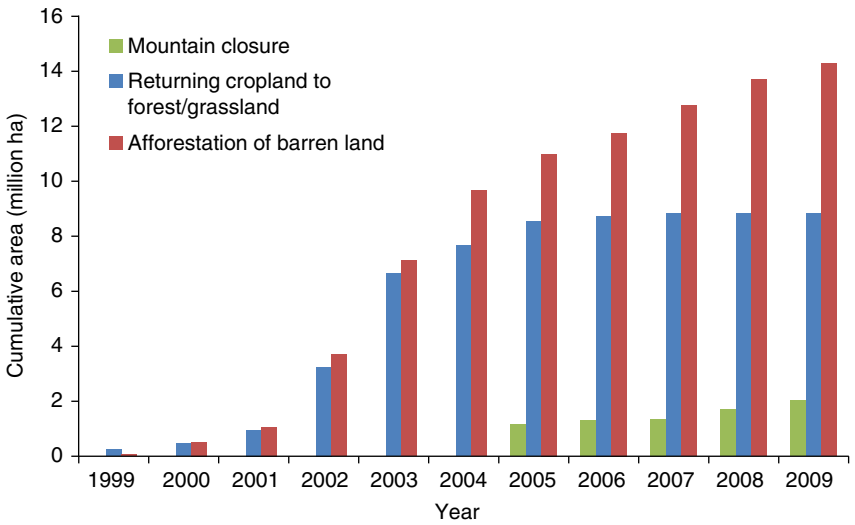


Figure 6 Areas of GTGP by different implementation approaches. Reproduced from State Forestry Administration (2000–2010) *China Forestry Development Report*. Beijing: State Forestry Administration; State Forestry Administration (2001–2009) *Bulletin of Key Forestry Projects*. Beijing: State Forestry Administration, and State Forestry Administration (2006–2010) *Statistical Analysis of National Forestry*. Beijing: State Forestry Administration.

yuan at 1.4 yuan per kg of grain for each hectare of converted cropland annually in the upper reach of the Yangtze River basin and in the upper and middle reaches of the Yellow River basin, respectively (Liu *et al.*, 2008). Furthermore, GTGP offers farmers additional subsidies (300 yuan ha^{-1} for miscellaneous expenses every year and 750 yuan ha^{-1} for seeds or seedlings in the first year) (Feng *et al.*, 2005; Xu *et al.*, 2004).

In 2007, after 8 years of implementation, payment contracts for many GTGP-enrolled lands began to expire. Since the long-term mechanism for addressing the livelihoods of GTGP-participating households had not been established, the central government decided to extend the program for another 8 years. However, the payment levels have been halved (Liu *et al.*, 2008).

Distribution

Compared to the NFCP, the GTGP is broader in geographic scope (**Figure 3**). The pilot program started in three provinces – Sichuan, Shaanxi, and Gansu – in 1999 (**Figure 3**). After initial success, it was extended to 17 provinces by 2000 and finally to 25 provinces by 2002 (**Figure 3**). As **Figure 3** shows, the GTGP covers all provinces in western China, which is ~80% of the total area with soil erosion (>360 million ha), including the headwaters of the Yangtze and Yellow rivers (**Figure 3**). Western China also contains the most desertification-prone area (174 million ha), three quarters of the cropland with a slope >25° (600 million ha), and 60% of the population below the poverty line (Liu *et al.*, 2008; Ouyang, 2007).

Biophysical Effects

Like the NFCP, indicators of the GTGP biophysical effects are often those immediately observable: amount of land converted and afforested, reduction in soil erosion, and reduction in water surface runoff. Ecosystem service changes on large scales, such as flood control, are mainly inferred from changes in immediately observable factors (Liu *et al.*, 2008).

By the end of 2009, the program had cumulatively increased vegetative cover by 25 million ha, with 8.8 million ha of cropland being converted to forest and grassland, 14.3 million ha barren land being afforested, and 2.0 million ha of forest regeneration from mountain closure (**Figure 6**). The statistics of the State Forestry Administration suggest that forest cover within the GTGP region has increased 2% during the first 8 years (Liu *et al.*, 2008).

Soil erosion and surface runoff have been reduced under GTGP. For example, in Hunan Province, surface runoff was reduced by ~20% and soil erosion declined by 30% from 2000 to 2005 (Li *et al.*, 2006a). Investigations in 14 counties of Sichuan Province indicate that from 1998 to 2003, the area affected by soil erosion declined 10% (Bao *et al.*, 2005). In Zigui County of Hubei Province, over a period of 5 years, converted plots lowered surface runoff by 75–85% and soil erosion by 85–96% in comparison to unconverted cropland (Wang *et al.*, 2007). In the Yellow River basin, it was estimated that surface runoffs would be reduced by 450 million m^3 from 2000 to 2020, which is equivalent to 0.76% of the total surface water resources (Jia *et al.*, 2006).

Water resources have been conserved and desertification has been reduced under the GTGP. For example, Minqin County of Gansu Province saved 516,000 m^3 of water in 2003 by reducing irrigation on 4300 ha of GTGP land (Ma and Fan, 2005). In the meantime, desertification has slowed down due to increased vegetation, an increase in air humidity by 15–25%, and a decrease in wind speed on the soil surface by 30–50% of GTGP land (Hou and Zhang, 2002).

The GTGP also enhances soil structure and lowers nutrient loss. In Guizhou Province, GTGP plots had 35–53% less loss of phosphorus than non-GTGP plots (Liu *et al.*, 2002). In Wuqi County of Shaanxi Province, the Chaigou Watershed had 48% and 55% higher soil moisture and moisture-holding capacity in GTGP plots than in non-GTGP plots, respectively (Liang *et al.*, 2006; Liu *et al.*, 2002).

Although vegetation cover has increased due to the GTGP (Liang *et al.*, 2006; Yang, 2006), tree species planted on GTGP land have low diversity. Different regions have different species, but a single or a few species often dominate one region. For example, during 2000–2005 in Henan Province, 40% of GTGP land was planted with poplar, 58% with fruit trees, and less than 2% with other species. In Jiangxi Province, camellia oil was planted in 60% of GTGP land in 2006 (Liu *et al.*, 2008). In addition, some scholars are concerned that afforestation with water-intensive species (e.g., poplar) in the semiarid and arid northwestern regions may not improve the environment; it might even deteriorate it (Cao, 2008).

Socioeconomic Effects

By the end of 2009, the total investment in the GTGP had exceeded 200 billion yuan (**Figure 7**) and more than 120 million farmers in 32 million households had participated in the GTGP. However, the socioeconomic effects of the GTGP vary from one area to another. In some areas, the vast majority of households were satisfied with GTGP (Hu *et al.*, 2006; Xu and Cao, 2002), which has helped alleviate poverty and reduce income inequality (Li *et al.*, 2011; Uchida *et al.*, 2005; Xu *et al.*, 2006b). It has helped many households shift from farming to nonfarming activities and thus change their income structures (Xu *et al.*, 2010; Zhang *et al.*, 2008). In Wuqi County of Shaanxi Province from 1998 to 2003 alone, 15,000 farmers changed their activities from farming to construction, transportation, and restaurant businesses (Ge *et al.*, 2006). The average household net income for GTGP participants has significantly increased by 75% in Ningxia and 8% in Guizhou (Uchida *et al.*, 2005). After their cropland was converted to GTGP land, many farmers had little to do in rural areas and went to cities to work as migrant workers. For example, in Guizhou Province, the number of migrant workers increased almost 50% over a period of 5 years (from 2.2 million in 2000 to 3.1 million in 2005) (Yang, 2006). In some other areas, farmers complained that participation is not voluntary as the central government claims, and the payment is low compared to agricultural revenue from the land (Xu and Cao, 2002; Xu *et al.*, 2010). In contrast to the optimism of the government, most households expressed concerns about their future livelihoods when the program ceases (Zhang *et al.*, 2008).

Similar to the NFCP, the GTGP has also generated financial difficulties for many local governments because local governments do not receive tax revenues from GTGP land (Huang,

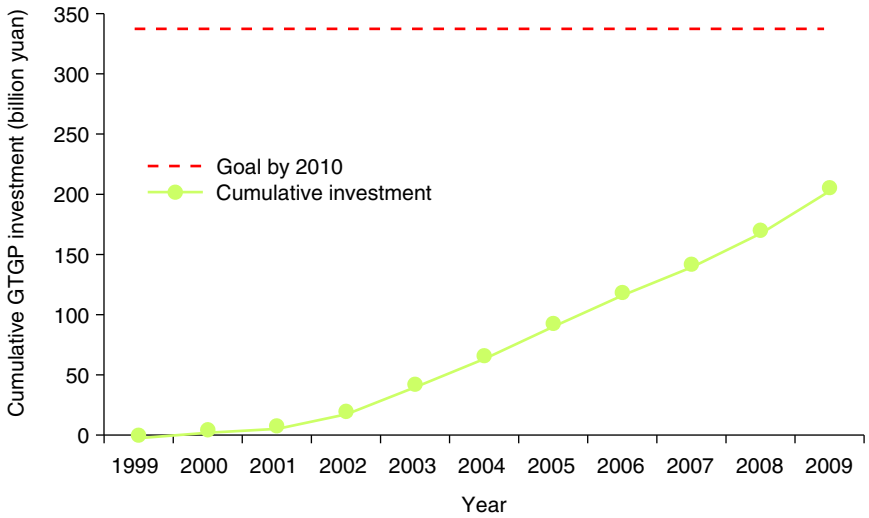


Figure 7 Investment in GTGP. Reproduced from State Forestry Administration (2000–2010) *China Forestry Development Report*. Beijing: State Forestry Administration; State Forestry Administration (2001–2009) *Bulletin of Key Forestry Projects*. Beijing: State Forestry Administration; State Forestry Administration (2006–2010) *Statistical Analysis of National Forestry*. Beijing: State Forestry Administration, and Xu J, Dao R, and Xu Z (2004) Grain to Green Program: Effectiveness, effects of structural adjustment, and economic sustainability. *China Economic Quarterly* 4: 139–162 (in Chinese).

2007; Xu *et al.*, 2006a). Although local governments also receive partial subsidies from the central government, they have to cover other expenses for implementing the GTGP (e.g., monitoring its effectiveness and transporting grain to farmers) (Liu and Zhou, 2005). For example, the local government of Kangding County in Sichuan Province lost 28% of its revenue between 1999 and 2001 (Dong, 2003).

Future Opportunities, Challenges, and Needs

Overall, like the NFCP, the GTGP has led to many accomplishments in terms of both biophysical and socioeconomic effects. However, the program also has induced negative effects on local livelihoods, financial burdens on local governments, and social conflicts between local households and governments over the use of land and associated resources (Miao and West, 2004).

Although the program has been extended for another round, so far there has been no systematic assessment of its effectiveness. Overall assessments were based only on government-reported data with few immediately observable indicators (Liu *et al.*, 2008; Yin, 2009). Considering its tremendous investment and wide distribution, it is crucial to systematically evaluate the biophysical and socioeconomic impacts of this program in both the short and long terms. Although proxies (e.g., slope) could be useful, they may not be sufficient for quantifying erosion severity or other ecological conditions (Yin, 2009). Therefore, further direct measures of environmental conditions, costs, and benefits are important.

The cost-effectiveness of the program can also be enhanced by targeting priority households and lands (Chen *et al.*, 2010; Uchida *et al.*, 2005). This targeting approach is especially useful to balance the environmental and poverty alleviation goals. Currently, in some areas, the GTGP is mandatory rather than voluntary, which has induced enrollment of flat, fertile, and profitable lands, both violating the goals of the program and harming the program in the long run (Xu and Cao, 2002; Yin, 2009; Zhang *et al.*, 2008).

To improve the performance and ensure the sustainability of the program, greater efforts should be made to improve the program design, engagement, institutional capacity building, and guidance on alternative income sources for participants. Since local communities and households are more able to recognize their needs and constraints, they should be given more power in the program participation, design and implementation. Voluntary participation would avoid enrolling flat and profitable lands and ensure that farmers acquire adequate compensation (e.g., no less than their opportunity costs). A pilot practice may be attempted to decentralize the design and implementation processes to local governments. Decentralization could give local governments more flexibility to adapt to local conditions, reduce the transaction costs for program enforcement, and reduce the financial burden on local governments (Bennett, 2008; Liu *et al.*, 2007a). Furthermore, like many other conservation and development policies in China, the GTGP is implemented separately and the interactive effects between the GTGP and other policies are rarely quantified (Liu *et al.*, 2008). However, for example, the GTGP has had positive impacts on rural household income and rural-to-urban and interior-to-coastal migrations in western China, which implies that it does interact with other policies such as the Western China Development Policy (Li *et al.*, 2011; Zhang *et al.*, 2008).

In addition, the government should improve the provision of alternative income sources for participants to improve their livelihoods in the long run. Although this was a requirement of the GTGP implementation plan at the very beginning, few effective approaches have been taken, and in some places the livelihoods of participants even became worse due to the implementation of the GTGP (Yin, 2009). Remedial actions should also be adopted to offset induced negative impacts on both the environment and the people, especially targeting

priority households such as the poor, minorities, and households with low social capital (Zhang *et al.*, 2008). Another focus area of future research would be how to establish a long-term management mechanism to improve the livelihoods of participants when the program expires.

Concluding Remarks

Despite differences in specific goals and implementation approaches, all three programs – NRS, NFCP, and GTGP – are remarkable ecosystem service policies in terms of investment, scale, and biophysical and socioeconomic effects. Overall, they have dramatically improved China's environmental conditions and ecosystem services, and thus have mitigated the unprecedented ecological degradation in China since the 1950s (Liu, 2010). As these programs continue and ecosystems recover, their ecological impacts will be even larger in the future. Besides the positive biophysical and socioeconomic outcomes, they have also induced some negative consequences. Some of these negative effects (e.g., decrease in revenue due to structural changes in forestry and agriculture) may benefit the government, forest enterprises, farmers, and other stakeholders in the long run. But there are also some common issues in these programs.

The reported effectiveness of these programs is largely based on government-reported data with only a few immediately observable indicators (e.g., implementation area). Rigorous, scientifically sound studies are mostly fragmented and often focus on only one aspect or a few aspects in a small area and for a short period (partly due to data accessibility and costs). Measurements of environmental benefits are mostly based on proxies (e.g., slope of land) rather than on direct observations (e.g., the change in soil erosion). Thus, integrated assessment and direct measurements of environmental conditions are needed to assess these programs.

To maximize the outcomes and minimize costs, the cost-effective targeting approach should be incorporated into the design and implementation processes of the programs (Chen *et al.*, 2010). Decentralization could be attempted to give local governments, communities, and households more autonomy in order to reduce transaction costs, avoid social conflicts, and enhance the performance and fairness of the programs. The targeting and decentralization approaches could also address the spatial heterogeneity across different regions and avoid overpayment for ecosystem services in some areas and underpayment in other areas.

Besides the three programs, some other ecosystem service policies have also been implemented recently. One major policy is the designation of EFZs (Ministry of Environmental Protection and Chinese Academy of Sciences, 2008). Based on the integrated analysis of ecological conditions and problems, evaluation of ecological sensitivity, and levels of importance of ecosystems services, 216 EFZs were identified in China. Of them, 50 national key EFZs, covering 23.4% of China's land surface, were identified on the basis of importance for ecosystem services provision, including biodiversity conservation, water resource conservation, soil maintenance, sandstorm prevention, and flood control (Ouyang, 2007). To maintain and improve ecosystem services in EFZs, the financial payments from the central government to local governments have been carried out since 2008. In 2008 alone, 6 billion yuan was transferred to 221 counties in EFZs. The total payment increased to 12 billion yuan for 372 counties in 2009 and to 24.9 billion yuan for 451 counties in 2010. Although many biophysical and socioeconomic effects have not yet been assessed, these large amounts of payments in EFZs are expected to improve China's ecosystem services and benefit the rest of the world in the long run.

For these and other payment for ecosystem service programs in China, the central government is the dominant buyer for ecosystem services. With the rapid economic development and increase of financial revenue during the past three decades, the central government can afford to pay for these programs. In the long run, however, if these programs continue and new programs emerge, the financial cost will be very high. Moreover, the government compensation approach also has limitations that are difficult to overcome, including the lack of flexibility, the difficulty in defining payment levels, and high transaction costs (Task Force for Eco-Compensation Mechanisms and Polices in China, 2007). To secure the financial sustainability of current and future programs, it is important to diversify the fund sources by engaging both public and private funds as well as funds from other countries and international organizations (Liu *et al.*, 2008).

The interactive effects among these programs and other conservation and development programs also should be studied. It should be recognized that all programs affect both ecosystem services and human well-being. Thus, a coupled human and natural systems perspective (Liu et al., 2007b, c) would be helpful to understand the complexity of policies and their impacts, and to establish long-term management mechanisms to improve the livelihood of participants in these programs and other ecosystem service policies in both China and many other parts of the world.

Acknowledgments

Financial support was provided by the National Science Foundation, National Aeronaut1ics and Space Administration, National Natural Science Foundation of China, National Key Basic Research Program of China, and Michigan AgBioResearch.

See also: Biodiversity and Cultural Ecosystem Services. Biodiversity and Ecosystem Services. Biodiversity, Human Well-Being, and Markets. Biodiversity in Logged and Managed Forests. Biodiversity-Rich Countries. Climate Change and Wild Species. Commons, Concept and Theory of. Commons, Institutional Diversity of. Conservation and People. Conservation and the World's Poorest of the Poor. Conservation Biology, Discipline of. Conservation Efforts, Contemporary. Conserving Biodiversity Outside Protected Areas. Deforestation and Land Clearing. Economic Value of Biodiversity, Measurements of. Economics of Agrobiodiversity. Economics of the Regulating Services. Endangered Ecosystems. Endangered Mammals. Endangered plants. Endangered Terrestrial Invertebrates. Environmental Ethics. Environmental Impact, Concept and Measurement of. Ethical Issues in Biodiversity Protection. Framework for Assessment and Monitoring of Biodiversity. Grazing, Effects of.

Habitat Loss and Fragmentation. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Impact of Ecological Restoration on Ecosystem Services. *In Situ*, *Ex Situ* Conservation. Indigenous Peoples and Biodiversity. International Organizations and Biodiversity. Justice, Equity and Biodiversity. Landscape Ecology and Population Dynamics. Land-Use Patterns, Historic. Loss of Biodiversity, Overview. Mammals, Conservation Efforts for. Market Economy and Biodiversity. Natural Reserves and Preserves. Population Stabilization, Human. Poverty and Biodiversity. Property Rights and Biodiversity. Reforestation. Remote Sensing and Image Processing. Restoration of Biodiversity, Overview. Role and Trends of Protected Areas in Conservation. Social and Cultural Factors. Social Behavior. Soil Conservation. Sustainability and Biodiversity. Systematics, Overview. The Value of Biodiversity. Threatened Species: Classification Systems and Their Applications. Tourism, Role of. Traditional Conservation Practices. Trends in Nature Recreation: Causes and Consequences. Valuing Ecosystem Services. Wildlife Management

References

Balmford A, Bruner A, Cooper P, *et al.* (2002) Economic reasons for conserving wild nature. *Science* 297: 950–953.

Bao J, Tang D, and Chen B (2005) Socioeconomic effects of Grain to Green Program in Sichuan Province. *Sichuan Forestry Exploration and Design* 1: 26–32 (in Chinese).

Bearer S, Linderman M, Huang J, An L, He G, and Liu J (2008) Effects of timber harvesting and fuelwood collection on giant panda habitat use. *Biological Conservation* 141: 385–393.

Bennett MT (2008) China's sloping land conversion program: Institutional innovation or business as usual? *Ecological Economics* 65: 699–711.

Cao SX (2008) Why large-scale afforestation efforts in China have failed to solve the desertification problem. *Environmental Science & Technology* 42: 1826–1831.

Cao SX, Wang XQ, Song YZ, Chen L, and Feng Q (2010) Impacts of the Natural Forest Conservation Program on the livelihoods of residents of Northwestern China: Perceptions of residents affected by the program. *Ecological Economics* 69: 1454–1462.

Chen XD, Lupi F, Viña A, He GM, and Liu JG (2010) Using cost-effective targeting to enhance the efficiency of conservation investments in payments for ecosystem services. *Conservation Biology* 24: 1469–1478.

Dong J (2003) A comparative study of values between sloping fields and forestland – Effects of Grain to Green Program. *China Population, Resources and Environment* 13: 81–83 (in Chinese).

Fan J, Li J, Quan Z, Wu X, Hu L, and Yang Q (2011) Impact of road construction on giant panda's habitat and its carrying capacity in Qinling Mountains. *Acta Ecologica Sinica* 31: 145–149.

FCCDP (1998) Forest conservation and community development project rapid rural appraisal report: Nuozadu Nature Reserve and its surrounding communities. Kunming: Forest Conservation and Community Development Project.

Feng Z, Yang Y, Zhang Y, Zhang P, and Li Y (2005) Grain-for-green policy and its impacts on grain supply in West China. *Land Use Policy* 22: 301–312.

Ge W, Li L, and Li Y (2006) On sustainability of Grain to Green Program – A survey of Wuqi and Zhidan counties in Shaanxi Province. *Forestry Economics* 11: 33–49 (in Chinese).

He GM, Chen XD, Liu W, *et al.* (2008) Distribution of economic benefits from ecotourism: A case study of Wolong Nature Reserve for Giant Pandas in China. *Environmental Management* 42: 1017–1025.

Hou J and Zhang S (2002) Evaluating effects of Grain to Green Program in the Loess Plateau Area. *Bulletin of Soil and Water Conservation* 22: 29–31 (in Chinese).

Hu C, Fu B, Chen L, and Gulinck H (2006) Farmers' attitudes towards the Grain-for-Green programme in the loess hilly area, China: A case study in two small catchments. *International Journal of Sustainable Development and World Ecology* 13: 211–220.

Huang Q (2007) On Grain to Green Program – The case of northwestern China. *Ecological Economy* 2: 60–63 (in Chinese).

Huang Y (2005) Countermeasures to the main problems in Natural Forest Conservation Program in Guizhou. *Guizhou Forestry Science and Technology* 33: 49–51 (in Chinese).

Jia Y, Zhou Z, Qiu Y, *et al.* (2006) The potential effect of water yield reduction caused by land conversion in the Yellow River basin. In: Zhang L, Bennett J, Wang XH, Xie C, and Zhao JC (eds.) *A Study of Sustainable Use of Land Resources in Northwestern China*. Beijing: China National Forestry Economics and Development Research Center.

Li D, Bo F, and Tao J (2006a) Achievements in and strategies for Grain to Green Program in Hunan Province. *Hunan Forestry Science & Technology* 33: 1–5 (in Chinese).

Li J, Feldman MW, Li SZ, and Daily GC (2011) Rural household income and inequality under the Sloping Land Conversion Program in western China. *Proceedings of the National Academy of Sciences of the United States of America* 108: 7721–7726.

Li WJ, Ge XD, and Liu CY (2005) Hiking trails and tourism impact assessment in protected area: Jiuzhaigou Biosphere Reserve, China. *Environmental Monitoring and Assessment* 108: 279–293.

Li WJ, Zhang Q, Liu CY, and Xue QF (2006b) Tourism's impacts on natural resources: A positive case from China. *Environmental Management* 38: 572–579.

Li Y (2010) *Effects of Conservation Policies on Forest Cover Change in Panda Habitat Regions, China*. East Lansing: Department of Fisheries and Wildlife, Michigan State University.

Li Z, Zimmermann F, Hebblewhite M, *et al.* (2010) *Study on the Potential Tiger Habitat in the Changbaishan Area, China*. Beijing, China: China Forestry Publishing House.

Liang W, Bai C, Sun B, Hao D, and Qi J (2006) Soil moisture and physical properties of regions under Grain to Green Program in the Gullied Rolling Loess Area. *Soil and Water Conservation in China* 3: 17–18 (in Chinese).

Liu C, Meng Q, Li Y, and Lu J (2005) A case study on ecological and socioeconomic benefit evaluation of Sichuan provincial Natural Forest Conservation Program. *Acta Ecologica Sinica* 25: 428–434 (in Chinese).

Liu C, Wang S, Zhang W, and Liang D (2007a) Compensation for forest ecological services in China. *Forestry Studies in China* 9: 68–79.

Liu F, Huang C, He T, Qian X, Liu Y, and Luo H (2002) Role of Grain to Green Program in reducing loss of phosphorus from yellow soil in hilly areas. *Journal of Soil and Water Conservation* 16: 20–23 (in Chinese).

Liu J (2010) China's road to sustainability. *Science* 328. 50.

Liu J, Daily GC, Ehrlich PR, and Luck GW (2003a) Effects of household dynamics on resource consumption and biodiversity. *Nature* 421: 530–533.

Liu J and Diamond J (2005) China's environment in a globalizing world. *Nature* 435: 1179–1186.

Liu J and Diamond J (2008) Science and government: Revolutionizing China's environmental protection. *Science* 319: 37–38.

Liu J, Dietz T, Carpenter SR, *et al.* (2007b) Complexity of coupled human and natural systems. *Science* 317: 1513–1516.

Liu J, Dietz T, Carpenter SR, *et al.* (2007c) Coupled human and natural systems. *Ambio* 36: 639–649.

Liu J, Li S, Ouyang Z, Tam C, and Chen X (2008) Ecological and socioeconomic effects of China's policies for ecosystem services. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9477–9482.

Liu J, Linderman M, Ouyang Z, An L, Yang J, and Zhang H (2001) Ecological degradation in protected areas: The case of Wolong Nature Reserve for giant pandas. *Science* 292: 98–101.

Liu J, Ouyang Z, Pimm S, *et al.* (2003b) Protecting China's biodiversity. *Science* 300: 1240–1241.

Liu J, Ouyang Z, Taylor WW, Groop R, Tan Y, and Zhang H (1999) A framework for evaluating the effects of human factors on wildlife habitat: The case of giant pandas. *Conservation Biology* 13: 1360–1370.

Liu J and Raven PH (2010) China's environmental challenges and implications for the world. *Critical Reviews in Environmental Science and Technology* 40: 823–851.

Liu Y and Zhou Q (2005) Shortcomings in the incentives of Grain to Green Program. *China Population, Resources and Environment* 15: 104–107 (in Chinese).

Lu HF, Campbell D, Chen J, Qin P, and Ren H (2007) Conservation and economic viability of nature reserves: An emergy evaluation of the Yancheng Biosphere Reserve. *Biological Conservation* 139: 415–438.

Ma Y and Fan S (2005) Ecological-economic effects of Grain to Green Program in desertification areas. *Journal of Natural Resources* 20: 590–596. in Chinese.

Miao GP and West RA (2004) Chinese collective forestlands: Contributions and constraints. *International Forestry Review* 6: 282–298.

Millennium Ecosystem Assessment (2005) *Ecosystems & Human Well-Being: Synthesis*. Washington, DC: Island Press.

Ministry of Environmental Protection (2010) *The List of Nature Reserves in China by the End of 2009*. Beijing: Environmental Science Press.

Ministry of Environmental Protection (2011) *Bulletin of China's Environmental Status 2010*. Beijing: Ministry of Environmental Protection of China.

Ministry of Environmental Protection and Chinese Academy of Sciences (2008) *National Ecosystem Function Zones*. Beijing: Ministry of Environmental Protection of China.

Ouyang Z (2007) *Ecological Construction and Sustainable Development in China (in Chinese)*. Beijing: Science Press.

Ouyang Z and Wang R (1997) A primary study on the indirect value of biodiversity in China. In: Editorial Committee of State Report on Biodiversity of China (ed.) *State Report on Biodiversity of China*. Beijing: China Environmental Science Press.

Ouyang Z, Wang X, and Miao H (1999) A primary study on Chinese terrestrial ecosystem services and their ecological-economic values. *Acta Ecologica Sinica* 19: 607–613.

Ouyang Z, Wang X, Miao H, and Han N (2002) Problems of management system of China's nature preservation zones. *Science & Techonology Review* 49–52 (in Chinese).

Qiao R, Gao J, and Zhang A (2006) Effects of Natural Forest Conservation Program on farmers' income – The cases of Hubei, Sichuan and Chongqing. *Research of Agricultural Modernization* 27: 40–43 (in Chinese).

State Forestry Administration (2001) China wildlife conservation and nature reserve construction master plan. Beijing: State Forestry Administration of China.

State Forestry Administration (2006) *The Third National Survey Report on Giant Panda in China*. Beijing: Science Press.

Statistic Bureau of Zhangjiajie City (2011) 2010 *National Economic and Social Development Statistical Bulletin of Wulingyuan District*. Zhangjiajie: Statistic Bureau of Zhangjiajie City.

Task Force for Eco-Compensation Mechanisms and Polices in China (2007) *Eco-Compensation Mechanisms and Polices in China*. Beijing: Science Press.

Uchida E, Xu JT, and Rozelle S (2005) Grain for green: Cost-effectiveness and sustainability of China's conservation set-aside program. *Land Economics* 81: 247–264.

Viña A, Bearer S, Chen X, *et al*. (2007) Temporal changes in giant panda habitat connectivity across boundaries of Wolong Nature Reserve, China. *Ecological Applications* 17: 1019–1030.

Viña A, Bearer S, Zhang H, Ouyang Z, and Liu J (2008) Evaluating MODIS data for mapping wildlife habitat distribution. *Remote Sensing of Environment* 112: 2160–2169.

Viña A, Chen XD, McConnell W, *et al*. (2011) Effects of natural disasters on conservation policies: The case of the 2008 Wenchuan Earthquake, China. *Ambio* 40: 274–284.

Viña A, Tuanmu MN, Xu WH, *et al*. (2010) Range-wide analysis of wildlife habitat: Implications for conservation. *Biological Conservation* 143: 1960–1969.

Wang Z, Wang X, Shi Y, Pan L, Yu X, and Tang Z (2007) Effects of Grain to Green Program on soil and water conservation in Zigui County of the Three Gorges Reservoir Region. *Science of Soil and Water Conservation* 5: 68–72 (in Chinese).

WU RD, Zhang S, Yu DW, *et al*. (2011) Effectiveness of China's nature reserves in representing ecological diversity. *Frontiers in Ecology and the Environment* 9: 383–389.

Xu J and Cao Y (2002) On sustainability of converting farmland to forests and grasslands. *International Economic Review* Z2: 56–60 (in Chinese).

Xu J, Dao R, and Xu Z (2004) Grain to Green Program: Effectiveness, effects of structural adjustment, and economic sustainability. *China Economic Quarterly* 4: 139–162 (in Chinese).

Xu J, Yin R, Li Z, and Liu C (2006a) China's ecological rehabilitation: Unprecedented efforts, dramatic impacts, and requisite policies. *Ecological Economics* 57: 595–607.

Xu JC and Melick DR (2007) Rethinking the effectiveness of public protected areas in southwestern China. *Conservation Biology* 21: 318–328.

Xu JT, Tao R, Xu ZG, and Bennett MT (2010) China's Sloping Land Conversion Program: Does expansion equal success? *Land Economics* 86: 219–244.

Xu Z, Xu J, Deng X, Huang J, Uchida E, and Rozelle S (2006b) Grain for green versus grain: Conflict between food security and conservation set-aside in China. *World Development* 34: 130–148.

Xue D, Bao H, and Li W (1999) A valuation study on the indirect values of forest ecosystem in Changbaishan Mountain Biosphere Reserve of China. *China Environmental Science* 19: 247–252.

Yang L (2004) Natural Forest Conservation Program and forest eco-compensation system: The case of Taijiang in Guizhou Province. *Journal of Mountain Agriculture and Biology* 23: 158–163 (in Chinese).

Yang S (2006) Policy recommendations for Grain to Green Program in the "Eleventh Five-Year Plan" – Thoughts from studying the program in Sichuan and Guizhou provinces. *Forestry Economics* 9: 7–10 (in Chinese).

Yin R (ed.) (2009) *An Integrated Assessment of China's Ecological Restoration Programs*. New York: Springer.

Yuan JQ, Dai LM, and Wang QL (2008) State-led ecotourism development and nature conservation: A case study of the Changbai Mountain Biosphere Reserve, China. *Ecology and Society* 13: 55.

Zhang L, Tu Q, and Mol APJ (2008) Payment for environmental services: The sloping land conversion program in Ningxia autonomous region of China. *China & World Economy* 16: 66–81.

Zhang P, Shao G, Zhao G, *et al*. (2000) Ecology: China's forest policy for the 21st century. *Science* 288: 2135–2136.

Zhang Z (2006) Review of the Tenth Five-year Plan and Perspective on the Eleventh Five-year Plan in relation to Natural Forest Conservation Program. *Forestry Economics* 1: 49–52 (in Chinese).

Zhou X (2006) Countermeasures to problems in implementing Natural Forest Conservation Program in Sichuan Province. *Forest Resources Management* 2: 19–23 (in Chinese).

Evolution in Response to Climate Change

Julie R Etterson, University of Minnesota, Duluth, MN, USA
Ruth G Shaw, University of Minnesota, St Paul, MN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptation Increase in fitness due to response to natural selection.
Gene flow Dispersal of individuals (or pollen) between populations, resulting in a genetic contribution from one population to another.
Genetic drift Random change in the frequency of alleles or genotypes within a population.
Genetic variation Variation in traits attributable to genetic differences among individuals in a population.
Habitat fragmentation Subdivision of wild habitat, coupled with conversion of wild habitat for use by humans (e.g., agriculture, urbanization).
Heritability Genetic variation expressed as a proportion of the overall variance in a trait.
Inbreeding Mating between relatives. This results in progeny that carry two copies of the same allele at many loci throughout the genome.
Mutation A heritable change in the nucleotide sequence of an organism.
Natural selection Variation among individuals in their contribution of offspring to the next generation. To the extent that this variation in offspring number is associated with genetic differences among individuals in the population, natural selection results in change in the population's genetic composition and tends to increase its average fitness.
Phenotypic plasticity Response of traits to the environment in which an organism develops.
Provenance test An experiment, particularly in forestry, in which plants from different source locations are grown together in common conditions (also referred to as common gardens), often in multiple locations, as a basis for assessing genetically based differences in trait values.

Introduction

Research on evolutionary responses of wild populations to climate change has been stimulated by accumulating evidence demonstrating exceptional rapidity of recent climate change and predictions of rapid change in climate into the future. However, it is challenging to demonstrate empirically that observed changes in wild organisms are due to changing climate *per se* because of the complex dynamics of both climate and evolution.

Climate is Complex

Climate comprises many attributes (i.e., temperature maxima and minima, rainfall, humidity, insolation, all varying over the seasons) and is manifest over extended periods, rather than instantaneously. Consequently, climate cannot be imposed or manipulated experimentally, though individual components of it can be, on a limited spatial scale. In nature, climatic variation is correlated with other aspects of the environment; for example, temperature regimes vary latitudinally along with photoperiod; characteristics of soils vary, in part due to the effects of water percolating through them (Baldwin *et al.*, 1938). Moreover, erratic variability in climate is superimposed on the regularity of annual climate cycles; forecasts indicate that climate variability will increase, such that extremes occur more frequently (Alley *et al.*, 2007).

Evolution is Complex

The complexity of climate is matched by that of evolution, defined as genetic change in populations over generations. Genetic change results from any or all of four evolutionary processes: natural selection, genetic drift, mutation, and gene flow between populations that have diverged genetically (Hartl and Clark, 2006). However, some changes that are observed as environment changes, for example, in flowering time, or migration date, may reflect direct effects of environment on individuals during their development, i.e., phenotypic plasticity (Bradshaw, 1965). Thus, conclusive assessment or prediction of evolutionary change entails multigenerational studies designed to distinguish genetic from environmental responses. Evolutionary theory, developed over nearly a century, ameliorates these challenges by grounding and guiding evolutionary explanation and prediction.

Populations Have Evolved in Response to Climate Change in the Past

Despite the inherent complexities of both climate and evolution, which impose formidable challenges to study of evolutionary responses to climate change, it is well established that populations have adapted to different climates in the past. Climate has resulted in evolutionary convergence of morphological and physiological traits in unrelated taxa as illustrated by the increase in wind pollination during the arid Eocene period (Crepet, 1989) and rise of C4 photosynthesis during a period of low CO_2 concentration during the Miocene

(Ehleringer *et al.*, 1991). It is clear that on a geological time scale, natural selection due to climate has powerfully shaped the attributes of organisms.

Populations Will Evolve in Response to Climate Change in the Future

As climate continues to change, natural selection is expected to change such that different trait values will be favored, compared to the past or present. For example, natural selection may favor traits associated with earlier phenology or enhanced tolerance of drought or heat. If populations harbor genetic variation for traits under selection (and other genetic and demographic aspects of the population are conducive to evolutionary change, *see* Evolutionary Rates), populations are expected to adapt in response. Beyond altering natural selection, changes in climate are also likely to influence other evolutionary processes such as gene flow, genetic drift, and inbreeding.

What is this Article About?

Here the authors review basic concepts in evolutionary genetics with an emphasis on how they pertain to studies of the adaptation to climate change. They focus on evolutionary responses of wild populations inherently subject to the fundamental evolutionary processes; they do not consider domesticated populations, in which genetic change is dominated by decisions of breeders. The authors provide an overview of approaches that have been employed to elucidate adaptive evolution in natural populations in response to both geographical variation in climate in the past and to contemporary climate change. These approaches include observational studies conducted across spatial and temporal gradients, experimental manipulation of genetic and environmental factors, prediction of future evolution based on standing levels of genetic variation and selection gradients, and simulation modeling of evolutionary change.

Overview of Evolutionary Genetics

Populations are constantly changing. An increasing number of studies have documented changes in natural populations, even over a few generations (Grant and Grant, 1993; Reznick *et al.*, 1995; Bradshaw and Holzapfel, 2001; Réale *et al.*, 2003; Franks *et al.*, 2007). However, it is important to recognize that not all changes that have been observed in recent decades (e.g., Walther *et al.*, 2002) in natural populations reflect evolutionary responses to natural selection. A clear distinction needs to be made between "environmental response" and "evolutionary response."

Environmental versus Evolutionary Response

Environmental response of organismal traits, or phenotypic plasticity, reflects the direct dependence of organisms' physiology, growth, development, and behavior on the environmental conditions they experience through their lifetimes (Bradshaw, 1965). Plastic responses do not involve changes in the genetic composition of a population but manifest environmental influences on the expression of genes (e.g., flowering time, Aikawa *et al.*, 2010). In contrast, evolutionary change is the change over generations in the genetic composition of a population. All populations are continually subject to evolutionary change. It results from several processes, including natural selection, i.e., variation among genotypes in survival or reproductive success. Natural selection is the basis for adaptation. Genotypes that express adaptive phenotypes make a greater genetic contribution to the subsequent generation; in this way, alleles that confer higher fitness in a given generation gain in frequency in the population. Natural selection may also favor plasticity that maintains high fitness in a variable environment. For example, a genotype that produces thin leaves in mesic conditions but thicker leaves when it develops in more arid conditions may have a survival or reproductive advantage over genotypes that do not respond in this way. If there is variation among genotypes in their phenotypic plasticity and this variation is under selection, plasticity itself can evolve (Via, 1993).

In the context of climate change, phenotypic changes such as timing of flowering and breeding that correspond statistically with recent climatic trends are well documented (Parmesan and Yohe, 2003). Observation of such changes does not reveal the basis of these responses, whether plasticity in response to immediate environmental condition, adaptive evolution in response to selection mediated by environmental change, or a combination of these. Phenotypic plasticity may buffer populations against natural selection in the short term. However, the limits of phenotypic plasticity in maintaining fitness over a range of environments are poorly understood, and, over the long term, ongoing climate change is expected to exceed those limits (DeWitt *et al.*, 1998). Thus, as climate continues to change, it is likely that natural selection will alter plasticity of traits as well as their expression in particular environments, given climate model predictions that include both directional change and increased variability.

What is Natural Selection?

Natural selection occurs when genotypes in a population differ in fitness, their contribution of progeny, and the genes they carry, to subsequent generations. Over generations, natural selection is expected to increase the average fitness of populations in the environments in which they undergo selection. This process of adaptive evolution proceeds over generations at a rate that depends on the severity of genetic selection, the genotypic variation in fitness.

Climate conditions that are novel relative to previous ones at a particular location are likely to impose a typically strong directional selection on traits, such as rooting depth, that are related to survival in dry conditions. This would be expected even if climatic changes were smooth. However, the prediction that climate is becoming more erratic implies that bouts of intense selection in one direction (e.g., favoring greater drought tolerance) will be interspersed with weaker selection or even selection in the opposite direction. Evolutionary dynamics may be governed by such extreme selective events (Grant and Grant, 2002; Gutschick and BassiriRad, 2003).

Evolutionary Rates

The rate of evolutionary response to selection depends on both genetic variation and environmental conditions and is most simply illustrated by the *breeder's equation* (1). The expected amount of phenotypic change (R) in a single trait per generation in response to selection depends on the strength of selection on that trait as measured by the selection differential (S), a statistical measure of the association between fitness and trait values, and also on genetic variation for the trait as expressed by heritability (h^2):

$$R = h^2 S \quad [1]$$

Heritability (h^2) estimates the fraction of the total phenotypic variance (V_P) due to additive genetic variance (V_A):

$$h^2 = V_A / V_P \quad [2]$$

V_A is the variance in a trait that is attributable to segregation of alleles at all the genes that influence that trait, and the resemblance of relatives due to sharing of alleles at all these genes is proportional to h^2 (Falconer and MacKay, 1996). V_A, V_P, and S are not static quantities. The expression of additive genetic variance and the total phenotypic variance depend on the immediate environmental conditions, as does the selection differential. Thus, estimates of heritability and selection are context dependent and are applicable to prediction of selection only in the specific environment for which they have been estimated.

The single-trait breeder's equation is not adequate to describe response to selection when multiple traits are under selection and are genetically correlated with each other. Genetic covariance between pairs of traits, each subject to selection, can promote the response of each to selection, if, for example, the sign of the genetic covariance is positive and both traits are selected in the same direction. Conversely, a positive genetic covariance between traits impedes response of both traits to selection if higher values of one trait and lower values of the other are associated with higher fitness. The genetic variances of traits and the genetic covariances between pairs of traits depend on direct effects of environment on organisms and the traits they express; genetic variation in traits and fitness may be reduced (Hoffmann and Parsons, 1991) or enhanced (e.g., Shaw *et al.*, 1995) in harsh environmental conditions.

Genetic correlations between traits can be included in predictions of selection response using a multivariate extension of the breeder's equation:

$$\Delta\bar{\boldsymbol{z}} = G\boldsymbol{\beta} \quad [3]$$

where the predicted change in the means of a set of traits (expressed in a vector, $\Delta\bar{\boldsymbol{z}}$) across one generation depends on the additive genetic variance–covariance matrix ($\boldsymbol{G}$) and the vector of selection gradients ($\boldsymbol{\beta}$) (Lande, 1979). As in eqn [1], use of this expression to obtain predictions of selection response requires taking the product of two estimates, each of which has substantial statistical uncertainty due to sampling and measurement error. Thus, interpretation is not straightforward. Alternatively, the response to selection can be estimated directly as the additive genetic covariance between relative fitness and traits (Price, 1970):

$$\Delta\bar{\boldsymbol{z}} = Cov_A[w, \boldsymbol{z}] \quad [4]$$

where w is individual relative fitness (absolute fitness divided by mean fitness), and $\boldsymbol{z}$ is the vector of trait values (Shaw and Shaw, 1994; Etterson and Shaw, 2001; Hadfield *et al.*, 2010).

What Processes Influence Genetic (co)variance?

As with additive genetic variances, additive genetic covariances also depend on a population's genetic composition, and strong selection over the extended period of climate change may radically deplete genetic variation. Spontaneous mutation steadily augments genetic variation; under changing conditions, novel mutations may more likely contribute to ongoing adaptation. Furthermore, influx of genes from genetically divergent populations increases genetic variation. Consequently, $\boldsymbol{G}$ (eqn [3]), which governs the immediate response to selection, i.e., from the current generation to the next, is expected to change with ongoing climate change, though the net effect of all the processes that influence $\boldsymbol{G}$ is not readily predictable.

In populations where climatic conditions become so severe that population size drops dramatically, genetic variation is expected to decline due to genetic drift. By this process of random sampling, genes that could contribute to adaptation can be lost (e.g., Weber and Diggins, 1990). Moreover, as populations decline in size, inbreeding increases, reducing fitness itself (i.e., inbreeding depression) further impairing populations' capacity for increase. Inbreeding may also slow the response to selection (e.g., Shaw *et al.*, 1998). Moreover, the number of new mutations in a population declines directly with the population's size, with the consequence that particular mutations, for example, those that could support adaptation to changing climate, are less likely to arise. These considerations underscore the intimate connection between the demography of populations and changes in their genetic composition. Evolutionary theory before 1990 tended to ignore this connection, but such simplification is especially likely to mislead in circumstances when conditions are so severe that populations become small and are at risk of extirpation. In many cases, such populations are likely to die out. For others, immigration can play the important demographic role of increasing the size of a population, enhancing its chance of persisting, in addition to contributing to its genetic variation (Holt and Gomulkiewicz, 1997).

Evolution of Plasticity

The tendency for natural selection to favor adaptive plastic responses will depend on the extent to which climate increases in variability. In the simplest case, if climates were to change monotonically (e.g., becoming progressively warmer or drier), then natural selection may favor adaptive evolution via fixed genetic changes that result, for example, in earlier breeding or greater thermal/drought tolerance. However, in the more realistic case in which climate variability increases (Alley *et al.*, 2007), and hence variability in natural selection does also, then adaptation involving plasticity may enhance fitness

across a range of environments. However, reversals in genotypic fitness, either within the lifetime of an organism or between generations, are expected to impede evolutionary increase in population mean fitness because genotypes with high fitness under both sets of conditions will be rare (Via and Lande, 1985; Gillespie and Turelli, 1989; Van Tienderen, 1991; Gomulkiewicz and Kirkpatrick, 1992; Gavrilets and Scheiner, 1993a, b). Adaptive plasticity is also less likely to evolve in a highly erratic and unpredictable climate marked by extreme events.

Approaches to Understanding Evolution in Response to Climate Change

Adaptation to Climate in the Past

Past evolution in response to climate is evidenced by the current adaptation of organisms to their home environments, encompassing climate, among many other aspects. A high degree of adaptation to an environment is reflected in high fitness (comprising individual survival and reproduction) that contributes to persistence of a population in that environment. This implies that, in previous generations, the population has undergone natural selection in those (or similar) conditions and, as a consequence, its genetic composition has changed, with increased frequency of genes conferring greater fitness than would otherwise be the case.

Range Limits and Climate Variables

The limits of species' ranges commonly coincide with abrupt spatial gradients in climatic attributes, and the geographic extent of a species' range can sometimes be predicted with high confidence from climate data taken in a subset of its range (Peterson *et al.*, 1999). The geographic ranges of species have shifted during past periods of climate change and shifts in range are ongoing. Shifts of the range of species may be an important component of the biological response to climate change in the future. However, it is not clear that range shifts of populations will proceed at sufficient rates, such that they will be subject to similar climatic conditions as climate changes. For example, it is often suggested that severe fragmentation of contemporary habitat is likely to impede dispersal. Moreover, interactions within and between species may thwart establishment of a population in locations beyond the current range where climate has become suitable for it. Beyond the role of geographic shift in a species range, the evolution of range limits is expected, even in the absence of climate change (Kirkpatrick and Barton, 1997).

Geographic Variation among Populations within a Species

Species that occur across environmental gradients frequently express clinal patterns in phenotype that may suggest underlying local adaptation. However, these patterns do not necessarily reflect genetic differentiation, but may result from phenotypic plasticity, which may or may not be adaptive. The extent to which clinal patterns in phenotype can be attributed to genetic differentiation, phenotypic plasticity, or some combination of these can be ascertained through "common garden" experiments in which populations are sampled from an environmental gradient and reared in common conditions (e.g., Reinartz, 1984; Lacey, 1988). By subjecting the populations to the same environment, such experiments demonstrate that differences detected among them in morphology, physiology, and phenology, in the common garden setting, must be genetically based. Furthermore, if this differentiation is related to the gradient from which they were sampled, they will retain clinal patterns in a common environment (e.g., Etterson, 2004a). Clinal patterns that disappear in a common garden can be attributed to phenotypic plasticity (e.g., Maherali *et al.*, 2002). Expanded forms of the common garden approach involve multiple growing locations, including reciprocal transplant studies (Clausen *et al.*, 1940; Etterson, 2004a, b) and provenance trials (e.g., Rehfeldt *et al.*, 1999). These experimental designs make possible comparisons of the relative adaptation of populations under different climatic conditions, and in relation to the climates at the locations from which the populations were drawn. Manipulation of some particular aspect of climate as an added experimental component of a common garden may yield yet stronger evidence of past adaptation to prevailing climatic conditions.

Observations Over Time

Evolutionary response to climate change has already been documented for a few wild organisms for which long-term studies of field populations have been conducted. For example, over the past 30 years mosquitoes that breed in pitcher plants in the eastern US have evolved different genetically based photoperiodic cues for breaking dormancy that correspond to increases in the length of the growing season in recent decades (Bradshaw and Holzapfel, 2006). In Australia, 20 years of temporal sampling of fruit fly populations has demonstrated latitudinal shifts of clinal variation of genetic traits (Umina *et al.*, 2005). Rapid evolution of clinal variation has also been observed for several invasive species since the time of establishment (e.g., Huey *et al.*, 2000). It has also been inferred that the timing of breeding has evolved in a population of red squirrels in Canada. In the last decade, spruce trees have been producing their cones progressively earlier (Réale *et al.*, 2003), and this has exerted selection on the squirrel population resulting in earlier breeding. By tracing the inheritance of timing of breeding from parents to offspring for several generations, these researchers undertook to distinguish genetic change from phenotypic plasticity and ultimately attributed the observed changes to both. The authors note, however, that Hadfield *et al.* (2010) have called the conclusions of this and some other studies into question because the statistical methods used do not fully account for sampling variance. Given the difficulty in obtaining high-quality empirical data sets of this kind that are sufficiently large to allow robust methods of analysis, it is not surprising that there are few unequivocal examples of genetically based evolution in response to contemporary climate change (Gienapp *et al.*, 2008).

Resurrection Ecology

In a few cases, direct demonstration of the nature of contemporary change in wild organisms has been possible because propagules (e.g., seeds in tundra soil or eggs in lake sediments) were recoverable or fortuitously available such that ancestors could be revived and compared side-by-side with their descendants (Bennington *et al.*, 1991; Vavrek *et al.*, 1991; Hairston *et al.*, 1999; Kerfoot *et al.*, 1999; Franks *et al.*, 2007). This "resurrection approach" (Davis *et al.*, 2005; Franks *et al.*, 2008) has permitted phenotypic and genetic comparisons of populations representing different times. The work of Franks *et al.* (2007) illustrates this approach. They compared the date of first flower of plants that were revived from seeds stored in 1997 with that of plants grown from the seed of their descendants sampled in 2004 following an intervening drought. This comparison yielded estimates of rates of evolution and comparison with changes predicted from quantitative genetics theory.

Experimental Manipulation of Population Genetics

Artificial selection involves experimental imposition of selection to answer questions about potential rates of evolutionary change. In practice, a researcher selects from a population individuals that express phenotypes of interest, for example, drought tolerance or earlier flowering. The selected individuals are interbred with the expectation that, if the trait is heritable, its phenotypic mean will differ in the offspring of the subsequent generation. From eqn [1] we can see that if the strength of selection (S) is experimentally determined and the response to selection (R) is observed, we can estimate the realized heritability (h^2) of the trait of interest. An alternative approach is to expose a population to an environment (e.g., higher temperatures) in which genetic variation in fitness is expressed, such that natural selection proceeds (Bennett *et al.*, 1992). Following several to many generations, fitness and other traits are compared between the selected and control populations to determine the extent of evolutionary change. Experimental evolution by either of these approaches may demonstrate evolutionary change in numerous traits.

Information obtained from studies of experimental evolution has improved our understanding of biotic response to climate change in numerous ways. For example, artificial selection studies have shown that it is possible that shifts in flowering time in herbaceous plant species that have been observed in nature (Parmesan and Yohe, 2003) may be due, in part, to evolutionary change (Burgess *et al.*, 2007). In bacteria, studies where organisms evolved under experimental conditions have resulted in dramatic adaptive evolution in thermal tolerance (Lenski, 2001). It is particularly notable that adaptation proceeded in populations that initially lacked genetic variation; thus the adaptation that occurred depended entirely on newly arising mutations, and proceeded even under conditions of fluctuating selection. Mongold *et al.* (1999) have further shown that genes that can support adaptation beyond the presumed lethal thermal limit of *E. coli* fail to increase in frequency in competitive conditions. Consequently, only when populations decline in abundance does adaptation to such extreme thermal conditions proceed, a clear example of the intimate connection between demographic and genetic change. In contrast to the previous examples, some experiments to evaluate potential for adaptive evolution in response to aspects of global change have detected only limited evolutionary responses. Potvin and Tousignant (1996) manipulated both temperature and CO_2 concentration and found that responses of wild mustard following seven generations in those conditions were primarily through nonadaptive plasticity. In summary, experimental evolution is a powerful approach that will continue to elucidate the rate of evolutionary change, the presence of genetic variation for traits that are under selection, and the role of genetic correlations in evolutionary change.

Prediction of Future Evolutionary Trajectories

To predict adaptation in response to climate change, we need estimates for three fundamental pieces of information: the strength and direction of selection in future climates, the genetic variances of traits that are under selection in altered environments, and the extent to which these traits are genetically correlated with other traits.

Selection regimes currently impinging on populations at lower elevation or latitude may simply shift to higher elevations or latitudes as climate warms and the geographical ranges of species track these shifts. However, populations may change in more complex ways because the ecological context, both abiotic and biotic, will be altered. Paleoecologists, having extensively studied assemblages of species in the past via the pollen record, have frequently noted that these assemblages often do not have present-day analogs, which suggests that environmental conditions are not closely replicated in different eras. A direct way to assess trends in selection is to do phenotypic selection analyses on data collected from long-term studies of field populations. However, data sets that are complete enough to document such temporal changes in selection on wild populations are rare (but see Grant and Grant, 1995). An alternative approach is to compare natural selection in current environments to selection in experimental conditions similar to those predicted for the future. However, most studies that have manipulated environmental conditions, such as temperature, precipitation, and CO_2 have focused on changes in species composition rather than on patterns of natural selection, although there are a few exceptions (Totland, 1999). Insight into temporal changes in selection may also be obtained by characterizing changes in selection along environmental gradients that encompass a range of conditions similar to those predicted for a given location in the future (i.e., a chronosequence, Etterson, 2004a). This approach may provide an incomplete picture of future selection because native species composition will also be altered by climate change and invasive species, pests, and diseases may invade new territories and alter patterns of selection. Furthermore, patterns of selection may be more erratic in the future if climates become more prone to extreme events such as drought, heavy precipitation, heat waves, and intense tropical cyclones (Alley *et al.*, 2007).

Theory for Evolutionary Response to Changing Environment

The persistence of populations in changing environments depends on both evolution and demography. Theory that couples evolutionary and demographic dynamics of closed populations has shown that even if adaptive responses are possible, densities may fall below some critical level such that populations are susceptible to extinction by stochastic processes (Bürger and Lynch, 1995; Gomulkiewicz and Holt, 1995). Genetic variation in a trait subject to changing selection as environment changes can enhance the persistence of populations in a changing environment whether the population shifts in its range in conjunction with the changing environment (Pease *et al.*, 1989) or not (Bürger and Lynch, 1995).

Levels of genetic variation and therefore, the potential for adaptive evolution, are not likely to be uniform among populations across the species range (Antonovics, 1976; Lenormand, 2002). As populations become increasingly fragmented, with reduced dispersal among subpopulations, their potential for local adaptation is expected to increase (Dickinson and Antonovics, 1973; Garcia-Ramos and Kirkpatrick, 1997). Nevertheless, even if marginal populations are not swamped by gene flow from central populations, adaptive evolution may be slowed due to strong selection causing genetic variance to be depleted (Fisher, 1930). Genetic drift may also erode genetic diversity, especially in small and isolated populations near the periphery of the species range or in remnant populations (Wright, 1931; Nei, 1973). Thus, fitness may be positively or negatively influenced by gene flow between populations. Adaptation may be more rapid with gene flow, both because it increases genetic variation in traits subject to selection (Swindell and Bouzat, 2006) and because it impedes genetic drift (Lopez *et al.*, 2009), even in an environment that is constant through time. In populations spanning gradual environmental gradients, gene flow between adjacent populations is not expected to inhibit local adaptation because populations are similarly adapted. Marginal populations evolve in response to extreme conditions expanding their tolerance limits and geographic ranges. However, if the environmental gradient is steep, the process of local adaptation is swamped by gene flow between populations that are adapted to substantially different environments (Kirkpatrick and Barton, 1997).

Summary

Evolution proceeds unceasingly in all biological populations. Rapid adaptation to human-mediated environmental change was evident long before climate warming became apparent. Consequently, there is no doubt that natural populations will continue to evolve and, in particular, that adaptation to changing climate is possible, in principle. However, it remains unclear how the several evolutionary processes – natural selection, mutation, gene flow, and genetic drift – will jointly alter any single population or the biota collectively. Even though natural selection is likely to proceed, there is insufficient understanding to predict for how many species and which species the rate and extent of adaptation will be adequate to support persistence, given the very high rates of environmental change expected. Such predictions will require further empirical investigation (experimental, as well as observational) of genetic variation and natural selection, directly accounting for increasing variability of climate. Moreover, the remaining evolutionary processes and their interplay, all of which are sensitive to demographic aspects of populations, are expected to bear importantly on genetic change in nature. Research that accounts for these processes, as well as natural selection, will yield more reliable prediction of evolution and species persistence in the face of changing climate.

See also: Climate Change and Extinctions. Climate Change and Wild Species. Plant Phenology Changes and Climate Change

References

Aikawaa S, Kobayashic MJ, Sataked A, Shimizuc KK, and Kudoha H (2010) Robust control of the seasonal expression of the Arabidopsis FLC gene in a fluctuating environment. *Proceedings of the National Academy of Science* 25: 11632–11637.

Alley R, Berntsen T, Bindoff NL, *et al.* (2007) Climate Change 2007: The Physical Science Basis. Summary for Policy Makers. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change Secretariat, Geneva.

Antonovics J (1976) The nature of the limits to natural selection. *Annals of the Missouri Botanical Garden* 63(224): 247.

Baldwin M, Kellogg CE, and Thorp J (1938) *Soil Classification, Soils and Men. Yearbook of Agriculture US Department of Agriculture*, pp. 979–1001. Washington, DC: US Government Printing Office.

Bennett AF, Lenski RE, and Mittler JE (1992) Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution* 46: 16–30.

Bennington CC, McGraw JB, and Vavrek MC (1991) Ecological genetic variation in seed banks. II. Phenotypic and genetic differences between young and old subpopulations of *Luzula parviflora*. *Journal of Ecology* 79: 627–644.

Bradshaw AD (1965) Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics* 13: 115–155.

Bradshaw WE and Holzapfel CM (2001) Genetic shift in photoperiodic response correlated with global warming. *Proceedings of the National Academy of Sciences* 98: 14509–14511.

Bradshaw WE and Holzapfel CM (2006) Evolutionary response to rapid climate change. *Science* 312: 1477–1478.

Bürger R and Lynch M (1995) Evolution and extinction in a changing environment: A quantitative-genetic analysis. *Evolution* 49: 151–163.

Burgess KS, Etterson JR, and Galloway LF (2007) Artificial selection shifts flowering phenology and other correlated traits in an autotetraploid herb. *Heredity* 99: 641–648.

Clausen J, Keck DD, and Hiesey WM (1940) *Experimental Studies on the Nature of Species. I. Effect of Varied Environments on Western North American Plants. Publication #520*. Washington, DC: Carnegie Institution of Washington.

Crepet WL (1989) History and implications of the early North American fossil record of Fagaceae. In: Crane PR and Blackmore S (eds.) *Evolution, Systematics, and Fossil History of the Hamamelidae, vol. 2: Higher Hamamelidae*, pp. 45–66. Oxford: Clarendon Press.

Davis MB, Shaw RG, and Etterson JR (2005) Evolutionary responses to climate change. *Ecology* 86: 1704–1714.

DeWitt TJ, Sih A, and Wilson DS (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution* 13: 77–81.

Dickinson H and Antonovics J (1973) Theoretical considerations of sympatric divergence. *American Naturalist* 107: 256–274.

Ehleringer JR, Sage RF, Flanagan LB, and Pearcy RW (1991) Climate change and the evolution of C4 photosynthesis. *Trends in Ecology and Evolution* 6: 95–99.

Etterson JR (2004a) Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change: I. Clinal patterns of selection along an environmental gradient in the Great Plains. *Evolution* 58: 1446–1458.
Etterson JR (2004b) Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change: II. Genetic architecture of three populations reciprocally planted along an environmental gradient in the Great Plains. *Evolution* 58: 1459–1471.
Etterson JR and Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science* 294: 151–154.
Falconer DS and Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. Essex, UK: Longman.
Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford, United Kingdom: Clarendon Press.
Franks SJ, Avise JC, Bradshaw WE, *et al.* (2008) The resurrection initiative: Storing ancestral genotypes to capture evolution in action. *Bioscience* 58: 870–873.
Franks SJ, Sim S, and Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences USA* 4: 1278–1282.
Garcia-Ramos G and Kirkpatrick M (1997) Genetic models of rapid evolutionary divergence in peripheral populations. *Evolution* 51: 21–28.
Gavrilets S and Scheiner SM (1993a) The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *Journal of Evolutionary Biology* 6: 31–48.
Gavrilets S and Scheiner SM (1993b) The genetics of phenotypic plasticity. VI. Theoretical predictions for directional selection. *Journal of Evolutionary Biology* 6: 49–68.
Gienapp P, Teplitsky C, Alho JS, Mills JA, and Merilä J (2008) Climate change and evolution: Disentangling environmental and genetic responses. *Molecular Ecology* 17: 167–178.
Gillespie JH and Turelli M (1989) Genotype–environment interactions and the maintenance of polygenic variation. *Genetics* 121: 129–138.
Gomulkiewicz R and Kirkpatrick M (1992) Quantitative genetics and the evolution of reaction norms. *Evolution* 46: 390–411.
Gomulkiewicz R and Holt RD (1995) When does evolution by natural selection prevent extinction? *Evolution* 49: 201–207.
Grant BR and Grant PR (1993) Evolution of Darwin's finches caused by a rare climatic event. *Proceedings of the Royal Society B* 251: 111–117.
Grant PR and Grant BR (1995) Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49: 241–251.
Grant PR and Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's Finches. *Science* 296: 707–711.
Gutschick VP and BassiriRad H (2003) Extreme events as shaping physiology, ecology, and evolution of plants: Toward a unified definition and evaluation of their consequences. *New Phytologist* 160: 21–42.
Hadfield JD, Wilson AJ, Garant D, Sheldon BC, and Kruuk LEB (2010) The misuse of BLUP in ecology and evolution. *American Naturalist* 175: 116–125.
Hairston NG, Lampert W, Cacres CE, *et al.* (1999) Rapid evolution revealed by dormant eggs. *Nature* 410: 446.
Hartl DL and Clark AG (2006) *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates, Inc.
Hoffmann AA and Parsons PA (1991) *Evolutionary Genetics and Environmental Stress*. Oxford: Oxford University Press.
Holt RD and Gomulkiewicz R (1997) The influence of immigration on local adaptation: A re-examination of a familiar paradigm. *American Naturalist* 149: 563–572.
Huey RB, Gilchrist GW, Carlson ML, Berrigan D, and Serra L (2000) Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287: 308–309.
Kerfoot WC, Robbins JA, and Weider LJ (1999) A new approach to historical reconstruction: Combining descriptive and experimental paleolimnology. *Limnology and Oceanography* 44: 1232–1247.
Kirkpatrick M and Barton NH (1997) Evolution of a species' range. *American Naturalist* 150: 1–23.
Lacey EP (1988) Latitudinal variation in reproductive timing of a short-lived monocarp, *Daucus carota* (Apiaceae). *Ecology* 69: 220–232.
Lande R (1979) Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. *Evolution* 33: 402–416.
Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution* 17: 183–189.
Lenski RE (2001) Testing Antonovics' five tenets of ecological genetics: Experiments with bacteria at the interface of ecology and genetics. In: Press MC, Huntly NJ, and Levin S (eds.) *Ecology: Achievement and Challenge*, pp. 25–45. Oxford: Blackwell Science.
Lopez S, Rousset F, Shaw FH, Shaw RG, and Ronce O (2009) Joint effects of inbreeding and local adaptation on the evolution of genetic load after fragmentation. *Conservation Biology* 23: 1618–1627.
Maherali H, Williams BL, Paige KN, and Delucia EH (2002) Hydraulic differentiation of Ponderosa pine populations along a climate gradient is not associated with ecotypic divergence. *Functional Ecology* 16: 510–521.
Mongold JA, Bennett AF, and Lenski RE (1999) Evolutionary adaptation to temperature. VII. Extension of the upper thermal limit of *Escherichia coli*. *Evolution* 53: 386–394.
Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA* 70: 3321–3323.
Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 42: 37–42.
Pease CM, Lande R, and Bull JJ (1989) A model of population growth, dispersal and evolution in a changing environment. *Ecology* 70: 1657–1664.
Peterson AT, Soberon J, and Sanchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. *Science* 285: 1265–1267.
Potvin C and Tousignant D (1996) Evolutionary consequences of simulated global change: Genetic adaptation or adaptive phenotypic plasticity. *Oecologia* 108: 683–693.
Price GR (1970) Selection and covariance. *Nature* 227: 520–521.
Réale D, McAdam AD, Boutin S, and Berteaux D (2003) Genetic and plastic responses of a northern mammal to climate change. *Proceedings of the Royal Society B* 270: 591–596.
Rehfeldt GE, Ying CC, Spittlehouse DL, and Hamilton Jr. DA (1999) Genetic responses to climate in *Pinus contorta*: Niche breadth, climate change, and reforestation. *Ecological Monographs* 69: 375–407.
Reinartz JA (1984) Life history variation of common mullein (*Verbascum thapsus*): I. Latitudinal differences in population dynamics and timing of reproduction. *Journal of Ecology* 72: 897–912.
Reznick DN, Shaw FH, Rodd FH, and Shaw RG (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* 275: 1934–1937.
Shaw RG, Byers DL, and Shaw FH (1998) Genetic components of variation in *Nemophila menziesii* undergoing inbreeding: Morphology and flowering time. *Genetics* 150: 1649–1661.
Shaw RG and Shaw FH (1994) Quercus programs published electronically, available via anonymous ftp from evolution.umn.edu.directory path pub/quercus.
Shaw RG, Platenkamp GAJ, Shaw FH, and Podolsky RH (1995) Quantitative genetics of response to competitors in *Nemophila menziesii*: A field experiment. *Genetics* 139: 397–406.
Swindell WR and Bouzat JL (2006) Gene flow and adaptive potential in *Drosophila melanogaster*. *Conservation Genetics* 7: 79–89.
Totland Ø (1999) Effects of temperature on performance and phenotypic selection on plant traits in alpine *Ranunculus acris*. *Oecologia* 120: 242–251.
Umina PA, Weeks AR, Kearney MR, McKechnie SW, and Hoffmann AA (2005) A rapid shift in a classic clinal pattern in Drosophila reflecting climate change. *Science* 308: 691–693.
Van Tienderen PH (1991) Evolution of generalists and specialists in spatially heterogeneous environments. *Evolution* 45: 1317–1331.
Vavrek MC, McGraw JB, and Bennington CC (1991) Ecological genetic variation in seed banks. III. Phenotypic and genetic differences between plants from young and old seed subpopulations of *Carex bigelowii*. *Journal of Ecology* 79: 645–662.
Via S (1993) Adaptive phenotypic plasticity: Target or by-product of selection in a variable environment? *American Naturalist* 142: 352–365.
Via S and Lande R (1985) Genotype–environment interaction and the evolution of phenotypic plasticity. *Evolution* 39: 505–522.
Walther G-R, Post E, Convey P, *et al.* (2002) Ecological responses to recent climate change. *Nature* 416: 389–395.
Weber KE and Diggins LT (1990) Increased selection response in larger populations. II. Selection for ethanol vapor resistance in *Drosophila melanogaster* at two population sizes. *Genetics* 125: 585–597.
Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97–159.

Evolution, Theory of

Gregory C Mayer, University of Wisconsin-Parkside, Kenosha, WI, USA
Catherine L Craig, Harvard University, Cambridge, MA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Catherine L. Craig, volume 2, pp. 671–681, © 2001, Elsevier Inc.

Glossary

Adaptation Characteristics of an organism that suit it to its conditions of existence, accounting for the similarity of organisms leading the same way of life; under the theory of descent with modification, a derived feature of an organism that increases the survival and reproduction of the organism by virtue of the possession of that feature.

Anagenesis Evolutionary changes occurring within an ancestor–descendant lineage.

Cladogenesis Splitting of a lineage into two or more descendant lineages.

Natural selection Consistent differential survival and reproduction of heritable variants.

Phylogenetic tree A tree-like diagrammatic representation of ancestry and descent and the branching of lineages; Darwin thought a coral might be more apt, but in either case, the form represents the actual physical continuity of the transmission of materials and information from one generation to the next.

Unity of type Similarities among organisms leading diverse ways of life under diverse conditions of existence, going beyond any functional need for similarity; under the theory of descent with modification, these similarities result from common ancestry.

Introduction

What a Theory is

Before embarking on an exposition of the theory of evolution, it is necessary to clear up a common misunderstanding that arises from a confusion of the vernacular and technical meanings of the word "theory." In everyday speech, "theory" is often taken to mean a guess, an hypothesis, or a speculation – something opposed to the more surely known "facts." In the context in which the word is used in the phrase "evolutionary theory," however, and indeed in much of science, it signifies a series of interconnected high-level generalizations based on a large body of evidence. Thus, one speaks of "atomic theory," "quantum theory," or the "germ theory of disease." Far from indicating that these sets of connected propositions are mere speculations, it indicates that they are well supported by a considerable corpus of data.

All scientific theories, of course, are provisional (as indeed are all scientific "facts") and subject to revision in the light of further data. It is always salutary to bear this in mind, and also that some of the ideas and findings are more likely to be subject to revision than others. But it is important to know that the word "theory" does not partition scientific propositions into those more or less likely to be revised. Rather, it attaches to some of the most well-supported propositions, as in the examples just discussed. It is in this sense that this article refers to the "theory of evolution."

What Evolution is

If "evolution" means "change over time," there are many things that evolve. Solar systems, for example, evolve. Stars condense, planets form, comets come and go, and stars age in predictable sequences. Languages also evolve. Within historical times, Latin has given rise to French, Spanish, Italian, and the other Romance languages, while contributing much vocabulary to the development of English. But the biological theory of evolution does not encompass cosmology or historical linguistics; rather it is restricted to organic evolution, that is, changes in living beings over time. This formulation is still too broad, since individual organisms grow and develop over time (i.e., have an ontogeny), but this is not evolution in the modern sense. The changes in life over time that concern humans are those that persist beyond the lifetime of a single individual and are transmissible to offspring (i.e., are heritable). So organic evolution consists of changes in living beings that transcend a single generation, and which, in principle, can be transmitted through an indefinite number of generations.

The theory of evolution, as considered here, is primarily concerned with the mechanisms of evolutionary change. The study of evolutionary mechanisms, or the causes of evolution, however, is but one of the two great branches of evolutionary biology (Futuyma, 1998). The other is the study of the history of life – phylogeny in the broad sense. The biodiversity observed today, and in the fossil record of past times, is the product of evolutionary mechanisms in tandem with changes in environment and geography acting on organic beings over long periods of time. This history of life and its present condition and causes, of both particular branches of the phylogenetic tree and biodiversity in general, are the subject of the greater part of this Encyclopedia.

Genetics, ecology, and developmental biology are among the disciplines that contribute largely to the study of evolutionary processes, whereas systematics and paleontology are among those that contribute to evolutionary history and the estimation of the evolutionary chronicle (O'Hara, 1988). All these disciplines, however, interact in ways that do not allow a neat separation of their contributions to the two sides of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00048-4

evolutionary biology. As Dobzhansky (1973) noted, nothing in biology makes sense unless it is studied in an evolutionary context, and all fields of biology provide insight into the processes of evolution and the history of life.

Brief History of Evolutionary Thought

Prior to Darwin

Charles Darwin (1809–1882) is the central figure in the development of evolutionary biology (Browne, 1995, 2002), with the publication of *On the Origin of Species* in 1859 being especially crucial (Browne, 2007). It is convenient to consider the state of biological thought before and after his contributions.

Prior to Darwin's work, a number of fundamental building blocks for evolutionary biology had been established. Chief among these were conclusions from Earth sciences: Earth was old (although no reliably exact estimate was made till well after Darwin's death), some ancient forms of life had become extinct, and the fossil record of life on Earth was progressive (i.e., the various forms of life existing at different times differed from time to time, and younger fossils more closely resembled the modern biota) (Young, 2007). These facts showed that life had a long history, but did not establish the causes of that history, nor what the connections were, if any, between the successive life forms.

A number of speculations were made that are now termed "evolutionary" (at the time, "evolution," from a Latin root meaning unfolding, was applied most often to ontogeny), but the first prominent and consistent scientific evolutionist was Jean-Baptiste Lamarck (1744–1829). Lamarck proposed that species originated continually via independent events of spontaneous generation and subsequently progressed "up" the scale of nature, from less complex to more highly organized beings. He viewed this as an inherent attribute of living things. As a subsidiary mechanism of transformation, and, in particular, of branching of forms from the inherent upward path, he supposed that the effects of individual striving and activity in response to the environment could be passed on, resulting in adaptation to the conditions of existence. The inheritance of acquired characteristics was thus a secondary mechanism for Lamarck (as, indeed, it would be for Darwin). Lamarck did not succeed in convincing many of his contemporaries; luminaries such as the paleontologist Georges Cuvier (1769–1832) and geologist Charles Lyell (1797–1875), who otherwise disagreed on many issues of Earth's history, nonetheless agreed in their condemnation of Lamarck's "transformism" (Corsi, 2005) and insisted on the fixity of species (**Figure 1**).

Another key building block of evolutionary biology was the recognition of the great diversity of life and its geographic distribution, with new and distinctive forms from around the world flooding the museums of Europe in the late eighteenth and early nineteenth centuries. The botanist Carolus Linnaeus (1707–1778), through his system of nomenclature, provided a common language for discussing this diversity, but it did not imply any genealogical connectedness among the forms so classified.

Darwin

In the first part of the nineteenth century, the two key problems facing biology were unity of type and adaptation. "Unity of type" referred to the similarities of organisms that led different – and sometimes very different – ways of life. All tetrapods, whether they be flying birds, walking turtles, grasping monkeys, or swimming ichthyosaurs, have a forelimb composed of a proximal humerus, then the radius and ulna, then the wrist and palm bones, and finally the phalanges – a pattern of "one bone, two bones, many bones" repeated throughout the amphibians, reptiles, birds, and mammals (Shubin, 2008). These bones vary in size, shape, and sometimes in number, but the generality of the pattern is striking. Similarly, vertebrate embryos of particular stages are strikingly alike, even though the adults differ starkly, leading the embryologist Karl Ernst von Baer (1792–1876) to formulate his

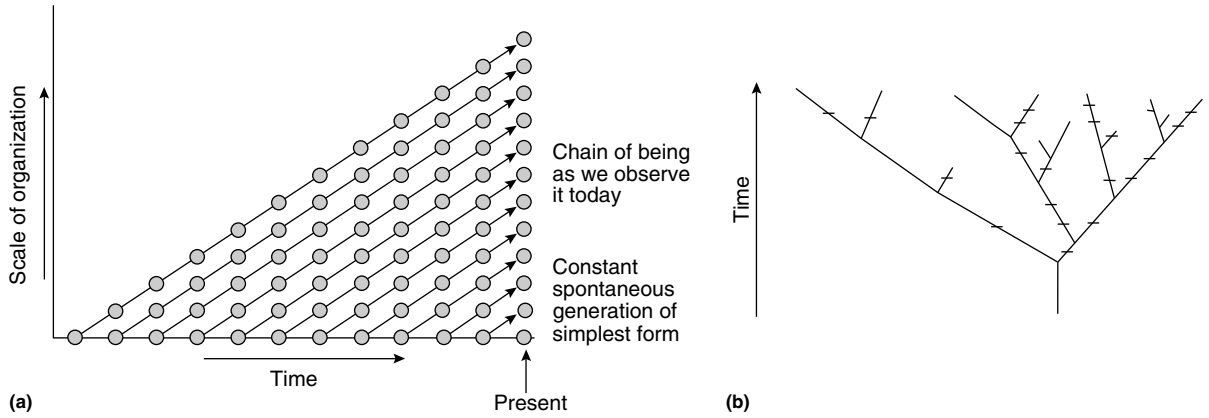


Figure 1 Proposed trees of ancestor–descendant relationships visualized by Lamarck (a) and Darwin (b). (a) Lamarck proposed that organisms originated spontaneously and continuously and evolved "up" a chain of natural being. (b) Darwin proposed that all species evolved from a common ancestor through the mechanism of natural selection. Reproduced from Futuyma DJ (1998) *Evolutionary Biology*, 3rd edn. Sunderland, MA: Sinauer.

law that, within a group, development proceeds from a common plan and progressively diverges, so that a pig embryo can be recognized to be that of a vertebrate early in development, but acquires the distinctive features of mammals later, and those of pigs later still.

"Adaptation" referred to the suitability of organisms to the conditions of existence – the fit of the organism to the environment. A hummingbird, which feeds on the nectar of flowers with long corollas, is thus suited to its conditions of existence by the possession of a long bill and tongue. A bananaquit, which feeds on nectar by snipping into the base of a flower, is suited to its conditions of existence by possessing a short, sharp bill. The birds' bills fit them to their conditions of existence and are thus adaptations to the conditions of existence.

In similar environments, similar adaptations may be found, even though the organisms exhibiting them are otherwise quite different. Thus, dolphins, ichthyosaurs, and sharks all possess a streamlined shape, dorsal fins, compressed tails, and steering forelimbs, yet one is a mammal, one a reptile, and one a fish, which differ from one another in their circulatory, respiratory, reproductive, and other systems in striking ways, despite their close similarity of overall appearance due to features that suit them to their conditions of existence.

The genius of Darwin was to propose a unified explanation for both unity of type and adaptation: descent with modification. Descent implied the well-known phenomenon of parent-to-offspring inheritance, whereas modification provided for changes from generation to generation. Unity of type was thus explained by descent from common ancestors: the resemblances of general body plan of organisms living diverse ways of life are due to the inheritance of the shared features from a common ancestor. Inheritance was understood in general (although not in its details until after Darwin), but what was most needed was a mechanism of modification and, furthermore, a mechanism that could account for adaptation. The need for an explanation of adaptation was especially felt in Britain, where the tradition of natural theology, exemplified in the work of William Paley (1743–1805), held that the often remarkable and complex fit between organisms and their conditions of existence was evidence of divine providence. The mechanism of modification that could account for adaptation proposed by Darwin (and also independently by his great collaborator Alfred Russel Wallace, 1823–1913) was natural selection, which he regarded as the chief, though not the only, means of modification.

Because many (though not all) of Darwin's proposals have been borne out and expanded by modern evolutionary biology, details of some of these proposals will be presented in the section Darwin's Five Theories.

The Modern Synthesis

Although Darwin rapidly succeeded in convincing the scientific community of the truth of descent with modification, acceptance of his view that the chief mechanism of modification (and the one most sure to produce adaptation) was natural selection acting on variation within populations was much delayed. Acceptance of the latter proposition was one of the chief accomplishments of the period of expansion in evolutionary thought from the 1920s to the 1950s, which Huxley (1942; 1887–1975) termed the "modern synthesis" (Young, 2007; Bowler, 2009).

In the initial phase of the synthesis, mathematical population geneticists, prominently Fisher (1930; 1890–1962), Wright (1931; 1889–1988), and Haldane (1932; 1892–1964), showed that Darwinian natural selection was compatible with Mendelian inheritance, and worked out the dynamics of genes in Mendelian populations under the influence of mutation, selection, and population structure.

In the second phase, the synthesis of Darwinism and Mendelism was expanded to include diverse subdisciplines of biology. Key to this development was the Russian–American empirical population geneticist Theodosius Dobzhansky (1900–1975). His work, growing out of the Russian school associated with Sergei Chetverikov (1880–1959), which stressed the study of natural populations, made the developments in both theoretical and empirical genetics available to a wider audience. Dobzhansky's (1937) seminal *Genetics and the Origin of Species* was followed by important synthetic works by the zoologists Mayr (1942; 1904–2004) and Rensch (1947; 1900–1990), the morphologist Schmalhausen (1949; 1884–1963), and the botanist Stebbins (1950; 1906–2000). The integration of paleontology with population genetics by Simpson (1944, 1953; 1902–1984) demonstrated that the synthetic theory of evolution was fully consistent with the fossil record. Alternative theories of the mechanisms of evolutionary change that had been formulated since Darwin but prior to the evolutionary synthesis, such as neo-Lamarckism and the view that biological systems were creative and purposeful, were demonstrated to be inconsistent with the fossil record.

The major achievement of the evolutionary synthesis was thus to fully integrate genetics and Darwinian evolution, and to demonstrate that evolutionary patterns and processes seen in natural populations and the fossil record were consistent with Darwinian natural selection, the hereditary mechanisms discovered by laboratory geneticists, and the theories of mathematical population genetics (Futuyma, 1998, 2009).

Postsynthesis Developments

As noted in the section What Evolution is, there are two central questions in evolutionary biology: the first asks, what are the mechanisms of evolution; that is, what are the causes of the history of life? The second asks, what is the history of life, as defined by the history of species, their origins, and extinctions, and what are the steps through which organismal features evolved? Although the major principles of evolutionary biology and major causes have been established, new questions are continually posed and many long-standing questions remain to be fully answered. In the postsynthesis period, these questions have been addressed in a broadening array of disciplines that have contributed to understanding evolutionary history and mechanisms, including such fields as evolutionary ecology, behavioral evolution, evolutionary paleontology, evolutionary developmental biology, evolutionary

physiology and morphology, evolutionary systematic biology, evolutionary genetics, molecular evolution, and human evolution. Research in all of these fields, and the interdisciplinary overlaps between them, have accelerated in the past 30 years. Two areas of explosive research deserve special mention.

First, the field of molecular evolution, from being scarcely existent in 1950, has become a burgeoning field comprising a large part of the research program of evolutionary biology (Higgs and Attwood, 2005; Edwards, 2009). It was realized early on that, once available, molecular sequence data, of either amino acids or DNA, would be "documents of evolutionary history" (Zuckerkandl and Pauling, 1965) and that branching sequences, ancestral states, and rates of evolution could be estimated from such data. This realization has proven abundantly true, and molecular data now form the basis of the phylogenetic estimates for large swaths of the tree of life (Felsenstein, 2004; Hall, 2011; Hillis, 2010).

Molecular data have contributed to the understanding of not only evolutionary history but also evolutionary mechanisms. In the 1960s, protein electrophoresis revealed high levels of genetic variation in natural populations, which answered old questions about the amounts of natural genetic variation and posed new ones about the maintenance of variation (Lewontin, 1974). The advent of DNA sequencing has revolutionized evolutionary molecular genetics (Kreitman, 1983), and such data now allow hypotheses about the relative importance of drift and selection on particular loci and regions to be effectively tested (Charlesworth and Charlesworth, 2010). As well, the sequencing of whole genomes of an ever-expanding variety of organisms has led to the emergence of comparative genomics, in which entire genomic structures can be compared, and their evolutionary histories and the forces shaping them studied (Lynch, 2007; Edwards, 2009).

The second is the field of evolutionary developmental biology, or "evo devo" as it is known, another field that scarcely existed in 1950. It too has become a vibrant research program, focused on the study of the developmental pathways by which organismal features are produced, the genetic basis of alterations in these pathways that are evolutionarily incorporated within lineages, and the biases or constraints that may be imposed on phenotypic evolution by these alterations. Most of evolution is the modification of preexisting structures, and these structures arise in the organism via a process of epigenetic (in the original sense: Haig, 2004) development. Thus, most of evolution is the modification of preexisting developmental programs. To understand phenotypic evolution, one must understand the variations that alterations of the developmental program can give rise to, their natures, and frequency, and these studies are the domain of evo devo (Carroll, 2005; Carroll *et al.*, 2005; True, 2009; Wray, 2010; Stern, 2011).

One of the most exciting developments in this field has been the discovery of a number of developmental regulatory genes, such as Hox genes, that regulate the spatial expression of genes in developing animal embryos, and which are conserved among taxa as disparate as arthropods and vertebrates. The two major groups of bilaterian animals – the protosomes (Ecdysozoa and Lophotrochozoa) and the deuterostomes (Chordata and Echinodermata) – share a common genetic regulatory repertoire and are characterized by an extensive cluster of Hox genes that bind to DNA and control interacting networks of developmental regulators and key structural genes (Knoll and Carroll, 1999) (**Figure 2**). This repertoire of genes, which was present in the common ancestor of the Bilateria, has been referred to as a "genetic toolkit" for the development and evolution of basic animal morphologies, and understanding their action and evolution will aid in understanding such features as germ layers, coelom formation, and spatial organization (Carroll *et al.*, 2005).

Darwin's Five Theories

Although Darwin often used expressions such as "my theory," he actually proposed a number of distinct, though related, theories. Some, such as his theory of inheritance, termed "pangenesis," were wrong and are of historical interest only. But a cluster of other theories has been borne out by subsequent developments in evolutionary biology, and it is useful to consider these ideas here. Mayr (1991, 2001) has identified a set of these ideas as "Darwin's five theories," and his usage has been followed and adapted by later authors (Futuyma, 1998, 2009; Coyne, 2009), and the following account is adapted from this approach.

The first is evolution as such, the simple idea that later organisms are the modified descendants of earlier organisms. This idea had also been advocated by Lamarck. Darwin convinced essentially all of his scientific contemporaries of the truth of this proposition. The evidence for this may be found in any evolution textbook (e.g., Freeman and Herron, 2007; Hall and Hallgrimsson, 2008; Futuyma, 2009), and in more popular expositions by Coyne (2009) and Dawkins (2009).

The second is common ancestry, which had been rejected by nontransformists such as Cuvier, who held that the four "embranchements" of animal life (Vertebrata, Radiata, Articulata, and Mollusc) were so essentially different that they could not be connected by any possible transformation (Young, 2007). Common ancestry may seem to follow from evolution as such, but it does not: Lamarck, it will be recalled, proposed continual spontaneous generation, with each origination progressing up the chain of being, and thus not being genealogically related to earlier or later originations. Darwin (1859, p. 490) proposed that life had been "breathed into a few forms or into one," and thus most or all organic beings would be connected by a single tree of life. This contention has been abundantly borne out by phylogenetic studies within eukaryotes, but the eukaryotes appear to be chimeric in their origin, incorporating by lateral or horizontal gene transfer both Archaebacterial and/or Eubacterial genes in their nuclear and organellar genomes. A phylogenetic tree based on small subunit RNA places the eukaryotes as sister to the Archaebacteria (Knoll, 2003), but large-scale genomic analyses are in disagreement about the nature, extent, and phylogenetic significance of the shared genomic components. Although some whole genome incorporation via endosymbiosis (e.g., mitochondria) is well established, the possibility of extensive lateral transfer of genes among prokaryotes and early eukaryotes has led Lynch (2007) to ask if life is a tree, ring, or web; resolution of the relationship among the domains of life remains a major task for phylogenetics (Gribaldo *et al.*, 2010) (**Figure 3**).

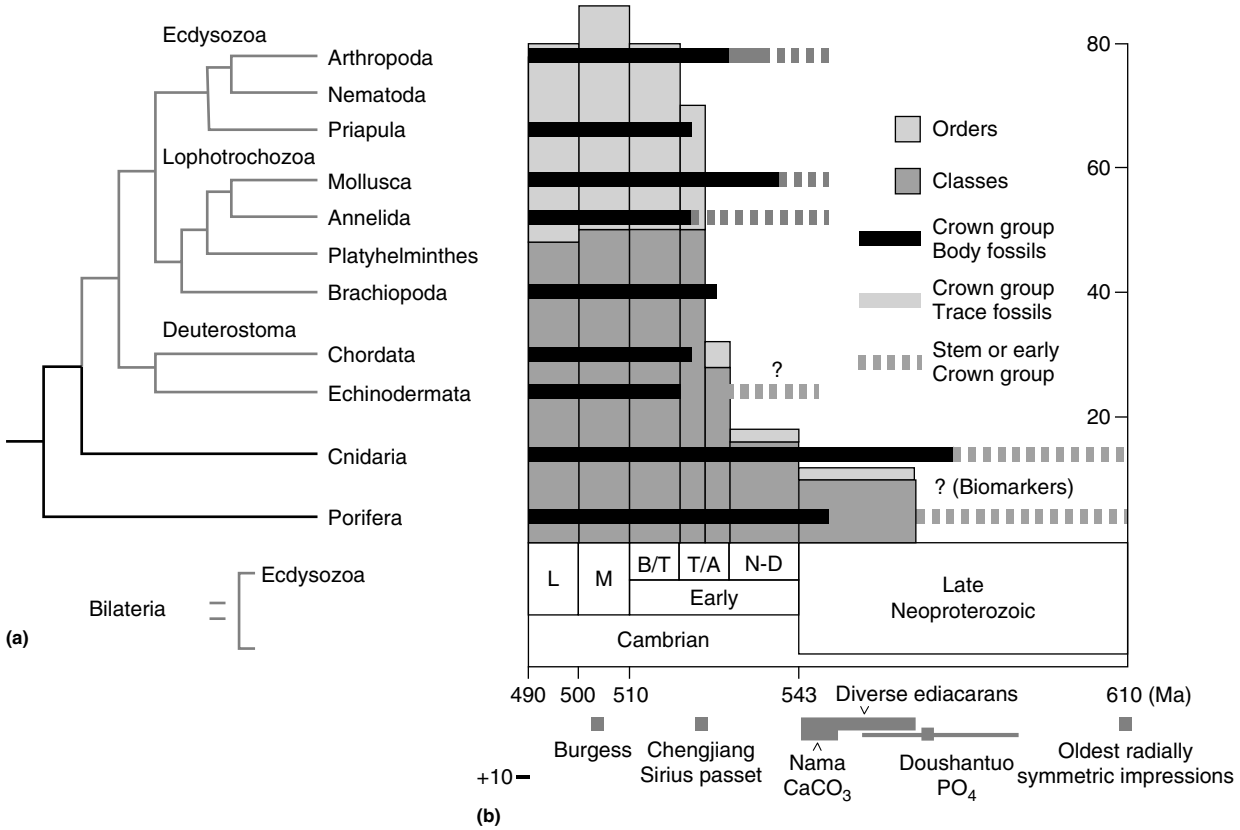


Figure 2 Animal diversity across the Proterozoic–Cambrian transition. By 510 Ma, all higher-level taxa of animals indicated had diverged. Bilaterians are characterized by triploblasty, a rostral–caudal axis, a dorsoventral polarity, and an extensive cluster of Hox genes. Reprinted from Knoll AH and Carroll SB (1999) Early animal evolution: emerging views from comparative biology and geology. *Science* 284: 2129–2137.

The third is gradualism, the idea that evolutionary divergence accrues incrementally by small steps. For Darwin, large changes that could produce or maintain adaptation would be akin to miracles. Although there is still debate (Stern, 2011), the modern consensus agrees with Darwin in rejecting saltational evolution, for a variety of reasons. Prominent among these are that mutations of large effect are usually accompanied by deleterious pleiotropic effects, which would make such mutations selected against, and the many fossil sequences documenting the origin of higher taxa via numerous intermediate steps. For example, that tetrapods evolved from lobe-finned fishes has long been known, and the discovery of fossils near the transition such as *Panderichthys* and *Tiktaalik* is filling in the structural differences between fish and amphibian, so that one can see exactly how, among other changes, legs gradually arose from fins. From the evidence of at least vertebrate paleontology (Prothero, 2007; Shubin, 2008; Laurin, 2010), one can say that most of evolution is the gradual, adaptive modification of preexisting structures or, better, their underlying developmental programs.

The fourth is that evolutionary change is populational or variational. Lewontin (1983) has argued that this is the most distinctive aspect of Darwin's thought. Rather than consisting of the transformation of individuals, Darwin proposed that individuals in a population have varying heritable properties and that evolution consists of changes in proportions of these heritable properties within the population. Some properties persist or increase in frequency, whereas others disappear, so that the hereditary constitution of the population as a whole changes.

Finally, there is natural selection, which Darwin considered the chief, though not the only, means of changing the frequency of the hereditary properties in populations and the evolutionary force most able to produce that adaptation to the conditions of existence, the absence of a compelling explanation for which had been a chief stumbling block for evolutionary hypotheses. Darwin proposed that, if the hereditary properties of individuals conferred on those bearing them varying propensities to succeed in the struggle for existence and reproduction, then those properties that conferred higher survival and/or reproduction (i.e., higher fitness) would increase in frequency in subsequent generations.

Evolutionary Mechanisms

Variation

Because the theory of evolution is variational, asking how much variation there is in natural populations is one of the key questions of evolutionary biology (Lewontin, 1974; Charlesworth and Charlesworth, 2010). And the answer is that

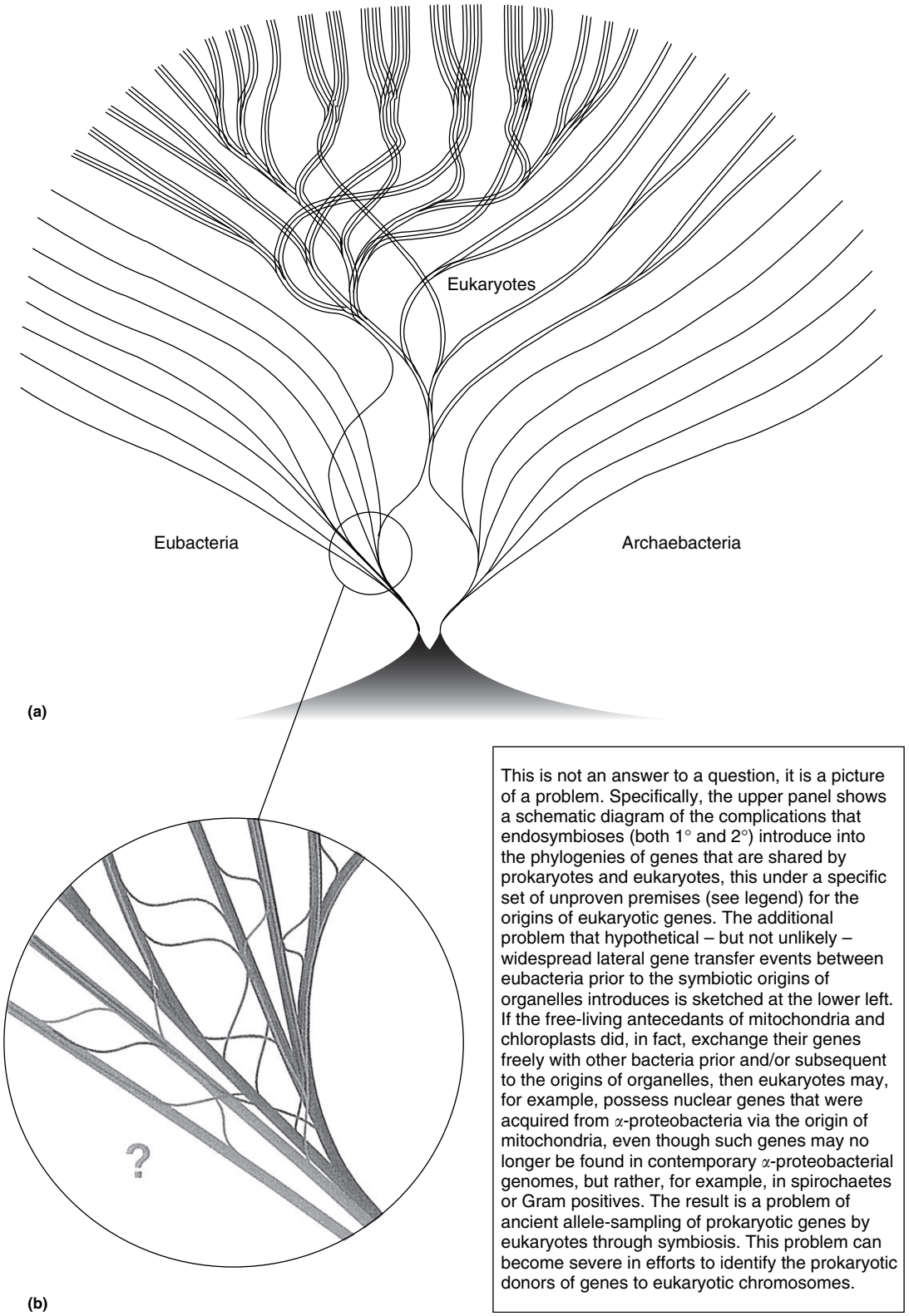


Figure 3 (a) The three "domains" of living organisms. The origin of the Eukaryotes (all multicellular organisms) has involved the combination of genomes from the Archaebacteria (methane-producing bacteria and bacteria living in extreme environments) and the Eubacteria (all remaining single-celled organisms). Lateral transfer (b) of genetic information among lineages complicates representation of the phylogeny, since it is not a pure branching, but also a reticulating, process. Reproduced from Martin W (1999) Mosaic bacterial chromosomes: A challenge en route to a tree of genomes. *BioEssays* 21: 99–104, with permission from Wiley-Liss Inc., a subsidiary of Wiley & Sons Inc.

variation in natural populations is ubiquitous. Since the nineteenth century, it has been clear that there is abundant variation within species in morphological characteristics such as size, shape, color, pattern, number, and so on, and this observation has been extended to behavioral and physiological traits. An important step in the synthesis was the demonstration of the genetic, and specifically Mendelian, basis of much of this variation (Young, 2007). As well, the results of animal and plant breeders had demonstrated an abundance of heritable variation within species.

In the twentieth century, the ability to directly examine proteins and then DNA sequences has confirmed the ubiquity of variation and extended it to the molecular genetic level. A species' genetic composition does not consist of a wild-type genotype with rare mutations, but rather a very large number of genotypes, with polymorphism at the protein level on the order of 30% of loci and 10% of individuals heterozygous. In a sexual species, virtually every individual is genetically unique (Freeman and Herron, 2007; Futuyma, 2009).

The Causes of Evolution

Mutation

A population may be characterized by the frequency within it of its various genotypes or genes. Since evolutionary change is variational – change in the distribution of heritable properties in populations –the causes of evolution are those forces that can change the frequencies of heritable properties in populations. One way in which these frequencies can change is to have new genotypes produced by changes in the genetic material during reproduction. Such changes may be broadly termed "mutations" and may result from errors in the genetic copying process or from environmental insults such as radiation or mutagenic substances. Recombination during sexual reproduction will also produce new genotypes. Mutations may be genic, affecting small sequences of DNA by nucleotide substitution, insertion, or deletion, or they may be chromosomal, affecting larger segments or whole chromosomes by rearrangement of segments and/or changes in chromosome number.

Mutation is of fundamental importance as the ultimate source of new genetic variation. Mutations are random, meaning that the probability of a mutation occurring is unrelated to whether or not it fits the organism to its conditions of existence (i.e., whether or not it is beneficial or deleterious). Mutations thus do not drive adaptation. The randomness of mutations does not mean that all mutations are equiprobable or that any conceivable phenotypic effect is achievable by mutation. A conceivable structure, without a preexisting developmental or structural basis, is difficult to get by mutation. Only phenotypes that can result from alteration of an existing developmental process are mutationally accessible. A Pegasus-like winged horse would be mutationally difficult, because horses (as do tetrapods generally) have four limbs, not six. An existing limb, however, can easily be modified: made longer or shorter, thicker or thinner, bony elements reduced or (more rarely) increased.

Despite being an evolutionary force of fundamental importance, mutation is nonetheless a weak evolutionary force, because observed rates of mutation are inadequate to cause rapid evolutionary change on their own, and even less so if selection acts counter to mutation (Futuyma, 2009).

Migration

The distribution of heritable material in a population may change if individuals enter the population, thus bringing potentially different genes or genotypes into the population. This movement of individuals is called migration, or gene flow. If the genetic constitution of incoming migrants differs from that of the recipient population, then the genetic composition of the population will change, and the effects can be substantial. Human populations have been especially well documented genetically, and the major effects of migration on gene frequencies have been documented and mapped (Cavalli-Sforza *et al.*, 1994); since the advent of rapid intercontinental travel around 1500, even more changes in human genetic population composition have occurred, and parts of the Americas especially are now inhabited by populations strongly shaped by migration.

The chief effect of migration is to retard divergence among populations between which migration is occurring: it leads to the genetic homogenization of populations. It thus tends to oppose the forces promoting differentiation, which is what drift and selection tend to do. Indeed, Futuyma (2010) has proposed that the main constraint on adaptive divergence is that such divergence, in either time or space, is ephemeral in the face of migration from nondivergent or alternatively divergent populations, which erases the evolved differences.

A second way in which different genes or genotypes can enter a population is not from migrants of the same species, but from other species, sometimes quite distantly related, through lateral or horizontal gene transfer. The importance of hybridization in, especially, plant evolution has long been known (Stebbins, 1950), and more recent work has shown the fundamental importance of lateral transfer in prokaryotes and the origin of eukaryotes (Knoll, 2003; Lynch, 2007), but examples are now known of adaptive evolution by lateral transfer involving higher eukaryotes, including fungus-to-animal transfer (Moran and Jarvik, 2010; Zhaxbayeva and Doolittle, 2011).

Genetic Drift

Even if there is no mutation or migration, the genetic composition of the offspring will differ from that of the parents, because reproduction is a probabilistic sampling process. Suppose two genetic variants are present in a population in equal frequency (50% each). It would then be expected that the same proportions be present in the next generation. Suppose the total population is 40. The expectation would be 20:20, but it wouldn't be surprising if it is 19:21, or even 23:17. It is like flipping a fair coin: one would expect 50% heads only in a very large number of tosses; in a few tosses, one might get mostly heads or mostly tails. The same is true of populations: the larger the population, the nearer to the expected proportions; the smaller the population, the more frequent and greater the departures from expectation. The direction of change in frequency of the variant – up or down – is random, and once changed, there is no tendency to move back toward the initial frequency. Thus, the frequency of the genetic variants will wander randomly, and this wandering is called genetic drift.

The strength of drift is inversely proportional to population size: it is more important in smaller populations. Drift has two main effects on the genetic composition of populations. First, it decreases variability within populations

(because some variants will drift to extinction), and second, it will cause divergence among populations (because different variants will drift to high or low frequency independently in the various populations) (Charlesorth and Charlesworth, 2010). For variants that are neutral with respect to selection (i.e., that do not differ consistently in their effects on the survival or reproduction of their bearers), drift will be an important force, and, if populations are small enough, even nonneutral variants can be strongly affected by drift (Graur and Li, 2000).

Selection

Drift is random differences in survival and reproduction. Sometimes, though, there will be good, deterministic reasons why variants differ in survival and reproduction, such that there is a predictable propensity for some variants to be represented at higher frequency in subsequent generations due to their greater survival and/or reproduction. This consistent (i.e., nonrandom) differential survival and/or reproduction is natural selection. It is important to note that this is not how selection acts, or the means of selection, or the result of selection: consistent differential survival and/or reproduction is selection.

Natural selection is the only evolutionary force that can produce adaptation – the fit between organisms and their conditions of existence. In modern evolutionary theory, one can define an adaptation as a derived feature of an organism that increases the survival and/or reproduction (i.e., fitness) of the organism by virtue of the possession of that feature. This definition distinguishes features that are adaptations from features that do not increase fitness (effects, such as the red color of blood due to iron in hemoglobin, as opposed to the adaptive nature of the oxygen-binding capabilities of hemoglobin) and features that may be useful, but which are already present (e.g., the skull sutures of mammals that are not adaptations for passage through the birth canal, but rather an inheritance from reptilian ancestors; Williams, 1966; Futuyma, 2009).

Natural selection is well documented as a cause of evolutionary change in populations (Endler, 1986; Freeman and Herron, 2007; Grant and Grant, 2008; Futuyma, 2009). There is widespread agreement that natural selection can produce adaptations of organisms through differential survival and/or reproduction of genes or individuals. Whether there can be selection among groups, populations, or species, so as to produce adaptations at these levels of organization, is much debated (Gould, 2002; Coyne and Orr, 2004; Futuyma, 2009).

Speciation

Mutation supplies the raw material of variation. Migration, drift, and selection are the forces that effect the fate of that variation within lineages; with respect to nonneutral variants, it can be said mutation proposes, selection disposes. The four evolutionary forces cause evolutionary changes, and these changes within a single ancestor–descendant lineage are called anagenesis. But if evolution consisted only of anagenesis, there would be only one kind of organism – changing over time, but just one lineage. To produce the great contemporaneous diversity of organic beings that inhabit Earth, there must be a multiplication of lineages. The generation of new lineages is called cladogenesis or speciation (Coyne and Orr, 2004; Price, 2008).

For the study of evolutionary mechanisms, including speciation, the most fruitful concept of species, at least for sexual forms, is the biological species concept (Coyne and Orr, 2004). Under this concept, species are groups of actually or potentially interbreeding populations in nature, which are isolated from other such groups by reproductive isolating barriers. Reproductive isolating barriers include phenomena such as courtship behavior, pollinator specificity, and hybrid sterility, all of which act as barriers to the exchange of genes between species. Under the biological species concept, the question of how species originate is the question of how reproductive isolating barriers arise.

One way to achieve this is to divide a population into two or more spatially separated populations, so that migration – a homogenizing force – is no longer possible between the spatially separated descendant populations. Each population initiates an independent lineage, within which mutation, migration, drift, and selection will convert individual variation into variation in time and, because the independent lineages coexist in time, into variation in space. If this independent anagenesis leads to the populations changing so much that they are reproductively isolated (e.g., their mating calls change to the point where they no longer recognize one another as potential mates), then the descendant populations will be able to coexist in the same place without losing their identity through interbreeding: they will have become separate species.

This scenario of geographic isolation is one mode of speciation, called allopatric speciation. It is the conversion of an initially extrinsic isolation (geographic separation) into intrinsic reproductive isolating barriers. There are other possible scenarios, which involve such possibilities as selection acting to reinforce reproductive isolating barriers, or little or no initial geographic separation (parapatric and sympatric speciation). Allopatric speciation is widely accepted (Mayr, 1942), whereas there remains considerable debate about at least some aspects of other modes of speciation (Coyne and Orr, 2004; Price, 2008).

As speciation is repeated over time, there is a potentially exponential increase in the number of species, and the ramifying, branching history of life over time is the phylogenetic tree. Different branches of the tree will speciate and go extinct at different rates. Some branches will undergo adaptive radiations, in which the varied species will come to occupy a diversity of stations in the economy of nature on scales from the archipelagic (Grant and Grant, 2008; Losos, 2009) to the continental (Lull, 1917). The lineages of the various branches will interact with one another and with their environments, and will disperse over, and move about with, the wandering continents and oceans. This constitutes biotic evolution, the evolutionary history of the set of all organic beings, and this is the history of biodiversity on Earth.

Appendix

List of Courses

1. Introductory Biology
2. Evolutionary Biology

3. Paleontology
4. History of Biology
5. Philosophy of Science

See also: Biodiversity, Evolution and. Cladogenesis. Cooperation, Evolution of. Darwin, Charles (Darwinism). Differentiation. Ecological Genetics. Evolution in Response to Climate Change. Genetic Diversity. Geologic Time, History of Biodiversity in. Life History, Evolution of. Mass Extinctions, Concept of. Origin of Life, Theories of. Phylogeny. Population Genetics. Speciation, Process of. Speciation, Theories of

References

Bowler PJ (2009) *Evolution: The History of an Idea*, 4th edn. Berkeley: University of California Press.

Browne EJ (1995) *Charles Darwin: A Biography 1: Voyaging*. Princeton, NJ: Princeton University Press.

Browne EJ (2002) *Charles Darwin: A Biography 2: The Power of Place*. New York: Knopf.

Browne EJ (2007) *Darwin's Origin of Species: A Biography*. New York: Grove Press.

Carroll SB (2005) *Endless Forms Most Beautiful: The New Science of Evo Devo*. New York: W. W. Norton.

Carroll SB, Grenier JK, and Weatherbee SD (2005) *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, 2nd edn. Malden, MA: Blackwell.

Cavalli-Sforza LL, Menozzi P, and Piazza A (1994) *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Charlesworth B and Charlesworth D (2010) *Elementary Evolutionary Genetics*. Greenwood Village, CO: Roberts.

Corsi P (2005) Before Darwin: Transformist concepts in European natural history. *Journal of the History of Biology* 38: 67–83.

Coyne JA (2009) *Why Evolution is True*. New York: Viking Penguin.

Coyne JA and Orr HA (2004) *Speciation*. Sunderland, MA: Sinauer.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: Murray.

Dawkins R (2009) *The Greatest Show on Earth*. New York: Free Press.

Dobzhansky T (1937) *Genetics and the Origin of Species*. New York: Columbia University Press.

Dobzhansky T (1973) Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* 35: 125–129.

Edwards SV (2009) Evolution of genes and genomes. In: Futuyma DJ (ed.) *Evolution*, 2nd edn., pp. 523–551. Sunderland, MA: Sinauer.

Endler J (1986) *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.

Felsenstein J (2004) *Inferring Phylogenies*. Sunderland, MA: Sinauer.

Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Freeman S and Herron JC (2007) *Evolutionary Analysis*, 4th edn. Upper Saddle River, NJ: Pearson Prentice Hall.

Futuyma DJ (1998) *Evolutionary Biology*, 3rd edn. Sunderland, MA: Sinauer.

Futuyma DJ (2009) *Evolution*, 2nd edn. Sunderland, MA: Sinauer.

Futuyma DJ (2010) Evolutionary constraint and ecological consequences. *Evolution* 64: 1864–1884.

Gould SJ (2002) *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.

Grant PR and Grant BR (2008) *How and Why Species Multiply: The Radiation of Darwin's Finches*. Princeton, NJ: Princeton University Press.

Graur D and Li W-H (2000) *Fundamentals of Molecular Evolution*, 2nd edn. Sunderland, MA: Sinauer.

Gribaldo S, Poole AM, Daubin V, Forterre P, and Brochier-Armanet C (2010) The origin of eukaryotes and their relationship with the Archaea: Are we at a phylogenetic impasse? *Nature Reviews Microbiology* 8: 743–752.

Haig D (2004) The (dual) origin of epigenetics. *Cold Spring Harbor Symposia on Quantitative Biology* 69: 1–4.

Haldane JBS (1932) *The Causes of Evolution*. London: Longmans Green.

Hall BG (2011) *Phylogenetic Trees Made Easy: A How-to Manual for Molecular Biologists*. 4th edn. Sunderland, MA: Sinauer.

Hall BK and Hallgrimsson B (2008) *Strickberger's Evolution*, 4th edn. Sudbury, MA: Jones and Bartlett.

Higgs PG and Attwood TK (2005) *Bioinformatics and Molecular Evolution*. Malden, MA: Blackwell.

Hillis DM (2010) Phylogenetic progress and applications of the tree of life. In: Bell MA, Futuyma DJ, Eanes WF, and Levinton JS (eds.) *Evolution Since Darwin: The First 150 Years*, pp. 421–449. Sunderland, MA: Sinauer.

Huxley JS (1942) *Evolution: The Modern Synthesis*. London: Allen and Unwin.

Knoll AH (2003) *Life on a Young Planet*. Princeton, NJ: Princeton University Press.

Knoll AH and Carroll SB (1999) Early animal evolution: Emerging views from comparative biology and geology. *Science* 284: 2129–2137.

Kreitman M (1983) Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* 304: 414–417.

Laurin M (2010) *How Vertebrates Left the Water*. Berkeley: University of California Press.

Lewontin RC (1974) *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Lewontin RC (1983) The organism as the subject and object of evolution. *Scientia* 118: 65–82.

Losos JB (2009) *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. Berkeley: University of California Press.

Lull RS (1917) *Organic Evolution*. New York: Macmillan.

Lynch M (2007) *The Origins of Genome Architecture*. Sunderland, MA: Sinauer.

Mayr E (1942) *Systematics and the Origin of Species*. New York: Columbia University Press.

Mayr E (1991) *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought*. Cambridge, MA: Harvard University Press.

Mayr E (2001) *What Evolution is*. New York: Basic Books.

Moran NA and Jarvik T (2010) Lateral transfer of genes from fungi underlies carotenoid production in aphids. *Science* 328: 624–627.

O'Hara RJ (1988) Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Systematic Zoology* 37: 142–155.

Price T (2008) *Speciation in Birds*. Greenwood Village, CO: Roberts.

Prothero DR (2007) *Evolution: What the Fossils Say and Why it Matters*. New York: Columbia University Press.

Rensch B (1947) *Neuere Probleme der Abstammungslehre: Die transspezifische Evolution*. Stuttgart: Enke. English translation (1959) *Evolution Above the Species Level*. New York: Columbia University Press.

Schmalhausen II (1949) *Factors of Evolution*. Philadelphia: Blakiston.

Shubin N (2008) *Your Inner Fish*. New York: Pantheon Books.

Simpson GG (1944) *Tempo and Mode in Evolution*. New York: Columbia University Press.

Simpson GG (1953) *The Major Features of Evolution*. New York: Columbia University Press.

Stebbins GL (1950) *Variation and Evolution in Plants*. New York: Columbia University Press.

Stern DL (2011) *Evolution, Development, & The Predictable Genome*. Greenwood Village, CO: Roberts.

True JR (2009) Evolution and development. In: Futuyma DJ (ed.) *Evolution*, 2nd edn., pp. 553–584. Sunderland, MA: Sinauer.

Williams GC (1966) *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Wray GA (2010) Embryos and evolution: 150 years of reciprocal illumination. In: Bell MA, Futuyma DJ, Eanes WF, and Levinton JS (eds.) *Evolution Since Darwin: The First 150 Years*, pp. 215–239. Sunderland, MA: Sinauer.

Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97–159.

Young D (2007) *The Discovery of Evolution*, 2nd edn. Cambridge: Cambridge University Press.

Zhaxbayeva O and Doolittle WF (2011) Lateral gene transfer. *Current Biology* 21: 242–246.

Zuckerkandl E and Pauling L (1965) Molecules as documents of evolutionary history. *Theoretical Biology* 8: 357–366.

Relevant Websites

http://darwin-online.org.uk/
The Complete Works of Charles Darwin Online.

http://evolution.berkeley.edu/
Understanding Evolution.

Extinction, Causes of

Richard B Primack, Boston University, Boston, MA, USA
Rachel A Morrison, Institution of Oceanography, La Jolla, CA, USA

Published by Elsevier Inc.

This article is a revision of the previous edition article by Richard B. Primack, volume 2, pp 697–713, © 2001, Elsevier Inc.

Glossary

Air pollution Lowering of air quality due to release of toxic materials by factories, automobiles, fires, and other human activities.
Disease Infections by parasitic organisms that can cause weakness, decreased reproduction, and death.
Exotic species Species that occurs outside its natural range owing directly or indirectly to human activity.
Global climate change Current and predicted changes in global temperature, rainfall, and other aspects of climate due to increased human production of carbon dioxide and other greenhouse gases.
Habitat fragmentation Process by which a continuous area of habitat is divided into two or more fragments by roads, farms, fences, logging, and other human activities.
Overexploitation Harvesting of a natural resource, such as fish or timber, at a rate more rapidly than it can be naturally replenished.
Water pollution Lowering of water quality due to input of sewage, pesticides, agricultural runoff, and industrial wastes that can harm aquatic plants and animals.

If species and natural communities are adapted to local environmental conditions, why should they be faced with extinction? Shouldn't species and communities be able to persist in the same places that they have for thousands of years? Why are species going extinct now? The answers to these questions have become clear in recent decades: Massive disturbances caused by people have altered, degraded, and destroyed the natural landscape on a vast scale, drastically upsetting the balance between natural rates of speciation and extinction. Extinction rates have been increasing since 1650, and we are now in the midst of a sixth mass extinction episode, this one caused by human activities rather than by a natural disaster. The process of evolution will eventually create new species, but this will take thousands, if not millions, of years. And numerous unique species such as pandas, elephants, and cheetahs will be gone forever.

An Expanding Human Population

The major human-induced threats to biological diversity are habitat destruction, habitat fragmentation, habitat degradation (including air and water pollution), overexploitation of species for human use, exotic species introductions, and increased spread of disease (**Figure 1**). Most threatened species face at least two or more of these threats, speeding their way toward extinction and hindering conservation efforts. Typically, these threats develop so rapidly and on such a large scale that species are not able to adapt genetically to the changes or to disperse to a more hospitable location. These threats will continue to increase as the human population grows, development and overexploitation continue, remaining natural habitats disappear, and as global climate continues to change.

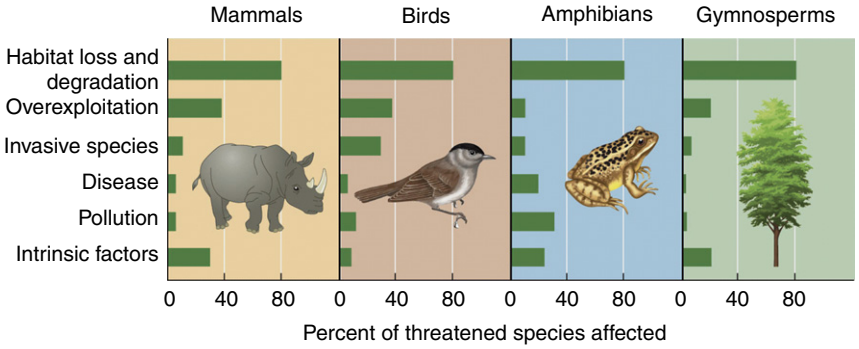


Figure 1 Habitat loss and degradation are the greatest threats to the world's species, followed by overexploitation. Groups of species face different threats; birds are more threatened by invasive species, whereas amphibians are more affected by disease and pollution. Percentages add up to more than 100% because species often face multiple threats. Reproduced with permission from Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates; based on data from the International Union for Conservation of Nature (IUCN), 2004.

These threats to biological diversity are all caused by an ever-increasing use of the world's natural resources by an expanding human population. In Europe, only 15% of the land area remains unmodified by human activities, and the amount of many specific habitat types remaining is below 10%. The greatest destruction of biological communities has occurred during the last 150 years, during which the human population exploded from 1 billion in 1850 to 2 billion in 1930, and to 6 billion in 2001. World population will reach an estimated 7 billion by the end of 2011 and 10 billion by 2050 (**Figure 2**). Birth rates have remained high whereas mortality rates have declined, particularly during the last century, as a result of both medical discoveries (specifically, disease control) and more reliable food supplies. Population growth has slowed in industrialized countries but is still high in many areas of tropical Africa, Latin America, and Asia, where the greatest biological diversity is also found. The increase in human population is predicted to cause an additional 14% of bird and mammal species to be threatened with extinction by 2050.

People use natural resources, such as fuelwood, wild meat, and wild plants, and convert vast amounts of natural habitat for agricultural and residential purposes. Agricultural systems now occupy one-fourth of the Earth's land surface. Because some degree of resource use is inevitable, population growth is partially responsible for the loss of biological diversity. All else being equal, more people equals less biodiversity. Some scientists have argued strongly that controlling human population size is the key to protecting biological diversity; however, overconsumption of resources is also responsible for habitat destruction and species extinctions. The rise of industrial capitalism and materialistic modern societies has greatly accelerated demands for natural resources, particularly in developed countries. Inefficient and wasteful use and overconsumption of natural resources are major causes for the decline in biological diversity. For example, if fewer paper products were used and more paper was recycled, then there would be less need to cut down forests to manufacture new paper.

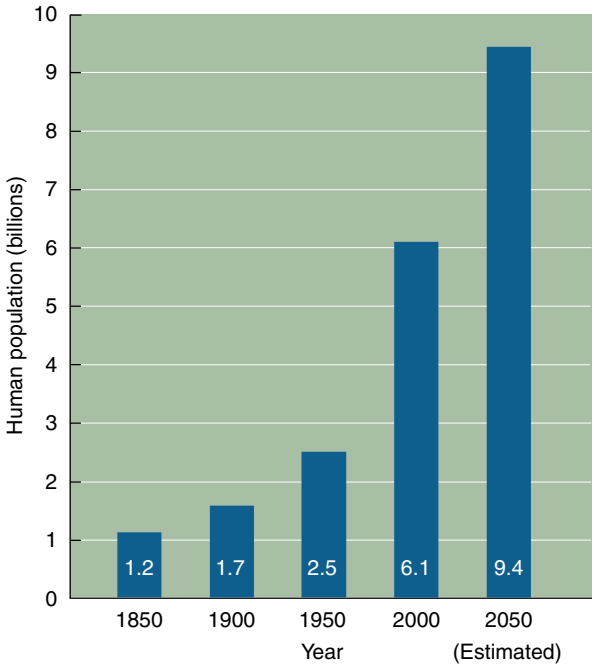


Figure 2 The human population in 2011 is almost 7 billion. The World Resources Institute estimates current annual population growth at 1.1%, but even this modest growth rate will add more than 72 million people to the planet in the next year. This number will escalate each year as the increase is compounded. Reproduced from Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates; based on data from the US Census Bureau.

Unequal Use of Natural Resources

In many countries there is extreme inequality in the distribution of wealth, with a small percentage of the population controlling, owning, and consuming much of the wealth and natural resources such as good farmland, livestock, and timber resources. In many developing countries, large landowners and business interests commonly force local farmers off their land, a pattern often reinforced by the government, the police, and the army. Political instability, lawlessness, and war also displace farmers into remote, undeveloped areas where they feel safer. Most practice shifting cultivation, a form of agriculture involving cutting down forest, burning the plant material, and planting crops in the nutrient-rich ash, because it is the simplest way to make a living when they may be obliged to move again within a short time. Landless farmers and their families often exploit natural resources in their surroundings just to stay alive; often these resources are components of species-rich biological communities.

The unequal use of natural resources worldwide also drives the destruction of biological diversity in species-rich tropical areas. People in industrialized countries (and the wealthy minority in developing countries) consume a disproportionate share of the world's energy, minerals, wood products, and food. Each year, the US, which has 5% of the world's population, consumes roughly 25% of the world's natural resources. The average US citizen annually uses 23 times more energy and 79 times more paper products than does the average citizen of India. The excessive consumption of wealthy individuals, cities, and countries thus leaves a large ecological footprint, meaning that a wide area of the world must supply their needs. This unsustainable resource consumption will cause massive environmental disruption if adopted by the expanding middle class in the developing world.

Large-Scale Development Projects

In many cases, the factors causing habitat destruction, particularly in the developing world, are the large industrial and commercial activities associated with a global economy – mining, cattle ranching, commercial fishing, forestry, plantation agriculture, manufacturing, and dam construction – and initiated with the goal of making a profit. Many of these projects are funded by national governments and international development banks and are touted as sources of jobs, commodities, and tax revenues. Others are initiated and funded by large multinational corporations. However, this exploitation of natural resources often is neither efficient nor cost-effective because

the emphasis is on short-term gain, often at the expense of long-term sustainability and generally with little regard for the local people who depend on the resources.

Habitat Destruction

Increasing human populations and their activities use even greater proportions of the world's terrestrial and marine environments and associated natural resources, resulting in the inevitable destruction of genetic variation, species, habitats, and ecosystem processes.

Habitat Loss

Habitat loss is the primary threat to the majority of vertebrate species currently facing extinction, a generalization that is certain to be true for threatened invertebrates, plants, and fungi as well. In many countries of the world, particularly on islands and in locations where human population density is high, most of the original habitat has been destroyed. Agriculture, commercial development, water projects, livestock grazing, pollution, infrastructure and roads, logging, and outdoor recreation threaten habitats of endangered species. More than 50% of the habitat has been destroyed in 49 of 61 Old World tropical countries. In tropical Asia, 65% of the primary forest habitat has been lost, with particularly high rates of destruction reported for Bangladesh (96%), Sri Lanka (86%), India (78%), and Vietnam (76%). Similarly, sub-Saharan Africa has lost about 65% of its forests, with losses most severe in Gambia (89%), Ghana (82%), and Rwanda (80%). Two biologically rich nations, Zimbabwe and the Democratic Republic of Congo (formerly Zaire), are relatively better off with about half of their forests remaining, although it is too soon to say how the recent civil war in the latter country has harmed its wildlife population. In the Mediterranean region which has been densely populated for thousands of years, only 10% of the original forest remains. Present rates of deforestation vary considerably among countries, with particularly high annual rates of 1.5–2% for tropical countries such as Vietnam, Côte d'Ivoire, Mexico, Paraguay, and Costa Rica.

For many important species, the majority of habitat in their original range have been destroyed, and a very little of their remaining habitat is protected. For certain Asian primates, such as the Javan gibbon, more than 95% of the original habitat has been destroyed. The orangutan, a great ape that lives in Sumatra and Borneo, has lost 63% of its habitat and is protected in only 2% of its range. Such habitat losses inevitably lead to extinctions.

Rain Forest Loss

Tropical rain forests occupy just 7% of the Earth's land surface but are estimated to contain over 50% of its species. Their destruction has therefore come to be synonymous with the loss of species. These evergreen to partly evergreen forests occur in frost-free areas below about 1800 m in altitude and experience at least 100 mm (4 in) of rain per month in most years. Rain forests are characterized by great species richness and complex species interactions and specializations unparalleled in any other community. Based on current patterns of rainfall and temperature, the original extent of tropical rain forests and related moist forests has been estimated at 16 million km^2. Only 11.5 million km^2 remained in 1990, and an additional 2.4% was lost between 2000 and 2005 (**Figure 3**).

Most rain forest destruction results from small-scale agriculture and firewood collection, mostly by poor farmers who have moved to forest areas to practice shifting cultivation out of desperation and poverty. Other major causes include

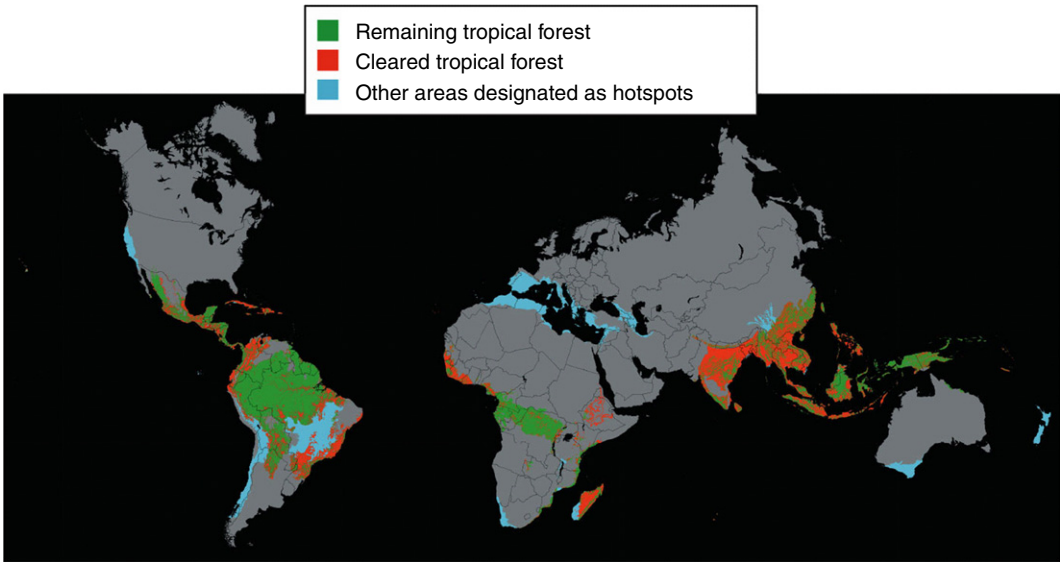


Figure 3 This Fuller projection, which distorts the sizes and shapes of continents less than typical maps do, shows the current extent of tropical forests and areas that have been cleared. Note the extensive amount of deforested land in northern and southeastern South America, India, Southeast Asia, Madagascar, and western Africa. Reproduced from Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates, and Pimm SL and Jenkins C (2005) Sustaining the variety of life. *Scientific American* 293(33): 66–73, with permission from Sinauer Associates.

clearing for commercial logging, cattle ranches, cash crop plantations (oil palm, cocoa, rubber, etc.), road building, and mining. At the current rate of destruction (approximately 140,000 km^2 per year), there will be no large blocks of tropical forest left after the year 2040, except in the relatively small national parks and protected areas and a few remote areas of the Brazilian Amazon, central Africa, and the islands of Borneo and Papua New Guinea (**Figure 4**). The situation is actually grimmer than these projections indicate because the world's population is still increasing, and poverty is rising in many developing tropical countries, putting ever-greater demands on the dwindling supply of rain forest.

Other Threatened Habitats

The plight of the tropical rain forests is perhaps the most widely publicized case of habitat destruction, but other habitats are also in grave danger.

Tropical Deciduous Forests

Tropical deciduous forests contain many species and in some places rival the diversity of the tropical rain forests. The land occupied by tropical deciduous forests is more suitable for agriculture and cattle ranching than the land occupied by tropical rain forests. Moderate seasonal rainfall, in the range of 250–2000 mm per year, allows nutrients to be retained in the soil, where they can be taken up by plants. These forests are also easier to clear and burn than rain forests. Consequently, human population density is five times greater in dry forest areas of Central America than in adjacent rain forests. Today, the Pacific coast of Central America has less than 2% of its original extent of deciduous dry forest remaining, and many species have been lost or are threatened with extinction.

Grasslands

Temperate grasslands have also been almost completely destroyed by human activity, with a consequent loss of species. It is relatively easy to convert large areas of grassland to farmland and cattle ranches. Illinois and Indiana, for example, originally contained 15 million ha (37 million acres) of tall-grass prairie, but now only 1400 ha (3500 acres) of this habitat – one-ten-thousandth of the original area – remain undisturbed. The rest have been converted to farmland, and the remaining area of prairie is fragmented and widely scattered across the landscape. Though widespread efforts to restore prairies in many areas of the world are underway and should be encouraged, it is impossible to bring back species that have already been lost.

Wetlands and Aquatic Habitats

Wetlands are critical habitats for fish, aquatic invertebrates, aquatic plants, and birds. They also provide services to humans such as flood control, drinking water, and power production. Wetlands are often filled in or drained for development, or they are altered by channelization of watercourses, dams, chemical pollution, and siltation. This habitat destruction increases extinction risk for aquatic species, especially one with limited distributions. During the last 200 years, over half of the wetlands in the US have been destroyed. The Florida Everglades, a premiere wildlife refuge, is now on the verge of ecological collapse. Approximately 40–50% of the freshwater snail species in the southeastern US are either extinct or endangered. In California's San Diego County, more than 97% of the vernal pools, which support a unique endemic biota, have been destroyed. Furthermore, massive development projects in other parts of the industrialized world, such as the Three Gorges Dam in China, are destroying aquatic ecosystems on an unprecedented scale.

Mangroves

Mangrove forests are among the most important wetland communities in tropical areas. Composed of species that are among the few saltwater-tolerant woody plants, mangrove forests occupy coastal areas with saline or brackish water, typically where there are muddy bottoms. Such habitats are similar to salt marshes in the temperate zone. Mangroves are extremely important breeding grounds and feeding areas for shrimp and fish. They also reduce storm damage. Despite their great economic value, mangroves are often harvested for timber and charcoal production and cleared for development. In recent years, mangroves have been increasingly cleared for rice cultivation and commercial shrimp hatcheries, particularly in Southeast Asia, where as much as 15% of the mangrove area have been removed for aquaculture. Over 35% of the world's mangrove ecosystems have been destroyed, and 40% of vertebrates endemic to mangroves are threatened with extinction.

Coral Reefs

Tropical coral reefs contain an estimated one-third of the ocean's fish species in only 1% of its surface area. Already 20% of all coral reefs have been destroyed, and an additional 20% have been degraded. The most severe destruction is taking place in the Philippines, where a staggering 90% of reefs are dead or dying. The main culprits are pollution, which either kills the coral directly or allows excessive growth of algae; sedimentation following the removal of forests; overharvesting of fish, clams, and other animals; climate change; invasive species; and, finally, fishermen blasting with dynamite and releasing cyanide and other poisons to collect the few remaining living creatures.

Extensive loss of coral reefs is expected within the next 40 years in tropical East Asia, around Madagascar and East Africa, and throughout the Caribbean (**Figure 4**). In the Caribbean, a combination of overfishing, hurricane damage, water pollution, and disease is responsible for a dramatic decline of a large proportion of the coral reefs and their replacement by fleshy macroalgae. Elkhorn and staghorn corals, which were formerly common and gave structure to the community, have already become rare in many locations.

Desertification

Many biological communities in seasonally dry climates are degraded into man-made deserts by human activities, a process known as desertification. These communities include tropical grasslands, scrub, and deciduous forests, as well as temperate shrublands, such as those found in the Mediterranean region, southwestern Australia, South Africa, central

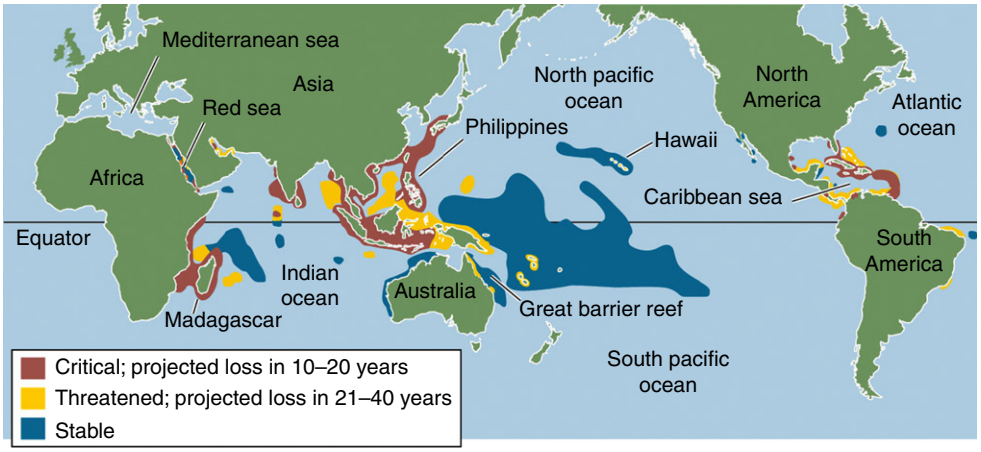


Figure 4 Extensive areas of coral will be damaged or destroyed by human activity over the next 40 years unless conservation measures can be implemented. Reproduced from Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates, and Bryant D, Burke L, McManus J, and Spalding M (1998) *Reefs at Risk: A Map-Based Indicator of Threats to the World's Coral Reefs*. Washington, DC: World Resources Institute, with permission from Sinauer Associates.

Chile, and California. Approximately 10–20% of these dry areas are degraded, with more than 25% of the productive capacity of their plant growth lost. Although these areas may initially support agriculture, repeated cultivation, especially during dry and windy years, often leads to erosion and loss of the soil's capacity to retain water. Land may also be chronically overgrazed by domestic livestock, such as cattle, sheep, and goats, and woody plants may be cut down for fuel. The result is a progressive and largely irreversible degradation of the biological community and loss of soil cover. Ultimately, the original species are lost, and the region takes the appearance of a desert. The process of desertification is most severe in the Sahel region of Africa, which is estimated to have 2.5 times more people than the land can sustainably support.

Habitat Fragmentation

In addition to outright destruction, habitats that formerly occupied wide, unbroken areas are now often divided into two or more fragments by roads, fields, farms, houses, industries, fences, power lines, and a broad range of other human constructs and activities. Fragmentation almost always occurs during a severe reduction in habitat area, but it can also occur when area is reduced to only a minor degree wherein the original habitat is divided by railroads, canals, fences, oil pipelines, fire lanes, or other barriers to the free movement of species. The island model of biogeography is applicable to this situation: the fragments, often isolated from one another by a highly modified or degraded landscape (**Figure 5**), may be considered as habitat islands in an inhospitable human-dominated sea. Fragments differ from the original habitat in three important ways: (1) fragments have a greater amount of edge for the area of habitat, (2) the center of each habitat fragment is closer to an edge, and (3) a formerly continuous habitat hosting large populations is divided into pieces, with smaller populations.

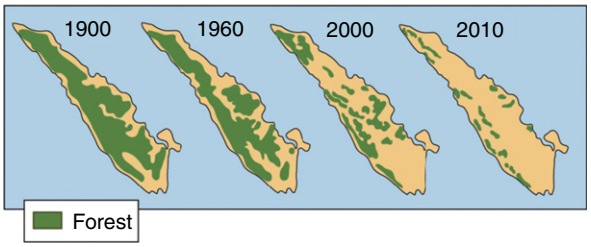


Figure 5 The forests of tropical Asia have experienced massive deforestation and fragmentation in recent decades. Intense habitat destruction in Sumatra, a large island of Indonesia, has occurred over the past 100 years and was predicted to continue through 2010. Reproduced from Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates, and Bradshaw CJA, Sodhi NS, and Brook BW (2009) Tropical turmoil: A biodiversity tragedy in progress. *Frontiers in Ecology and the Environment* 7: 79–87, with permission from Sinauer Associates.

Barriers to Dispersal

Fragmentation may limit a species' potential for dispersal and colonization. Many bird, mammal, and insect species of the forest interior will not cross even very short distances of open area, partly due to the high risk of predation in edge and open habitats. Habitat fragmentation also creates barriers to normal dispersal and the colonization processes. In an undisturbed environment, seeds, spores, and animals move passively and actively across the landscape. When they arrive in a suitable but unoccupied area, a new population begins to develop at that site. Over time, population of a species may build up and go extinct on a local scale as the species disperses from one suitable site to another and the biological community undergoes succession. Because habitat fragmentation limits this dispersal, the species may gradually die out.

Habitat fragmentation also reduces the foraging ability of individual animals. Many species, either as individuals or

social groups, need to move freely across the landscape to feed on widely scattered resources. However, fences and other barriers may prevent the natural migration of animals, such as wildebeest or bison, forcing them to overgraze an unsuitable habitat, eventually leading to starvation and habitat degradation. Barriers to dispersal can restrict the ability of widely scattered species to find mates, causing a loss of reproductive potential for many animal species. Plants also may reduce seed production if butterflies and bees are less able to migrate among habitat fragments to pollinate flowers. Finally, dividing a formerly large population into several smaller ones means that the new, smaller populations will be more vulnerable to genetic and other problems associated with small populations, increasing their risk of extinction.

Edge Effects

Habitat fragmentation often changes the microenvironment at the fragment edge, resulting in increased light levels, higher daytime temperatures, higher wind speeds, and lower humidity. Each of these edge effects can have a significant impact on the vitality and composition of the species in the fragment. Species that are sensitive to humidity such as amphibians, many insects, and herbaceous plants, will be eliminated from the forest fragments. Also, increased wind, lower humidity, and higher daytime temperatures make fires more likely in forest fragments. Fires may spread into habitat fragments from nearby agricultural fields that are being burned regularly, as in sugarcane harvesting, or from the irregular activities of farmers practicing shifting cultivation. In the process, many species will be eliminated.

Interspecific Interactions

Habitat fragmentation also increases the vulnerability of the fragment to invasion by exotic and native pest species. Omnivorous native animals, such as raccoons, skunks, and blue jays, and introduced animals, such as rats, may increase in population size along forest edges, where they can eat food found in both undisturbed and disturbed habitats. These aggressive feeders may seek out the nests of interior forest birds, often preventing successful reproduction of many bird species hundreds of meters from the nearest forest edge. Domestic cats can also be important predators in settled areas with fragmented landscapes. Weedy plant species and exotic herbivores can eliminate native plant species along the edges of fragments, and disease can similarly spread into the interior of habitat fragments.

Habitat Degradation and Pollution

Communities and species in a habitat can be profoundly affected by human activities, even if the habitat itself is not. Biological communities can be damaged and species driven locally or globally extinct by external factors that do not change the structure of dominant plants in the community, so that the damage is not immediately apparent. For example, in temperate deciduous forests, frequent and uncontrolled ground fires might not kill the mature trees, but the rich perennial wildflower community and insect fauna on the forest floor would gradually become impoverished. Keeping too many cattle in grassland communities gradually changes the biological community, often eliminating many native species and favoring exotic species that can tolerate grazing. Boat hulls, anchors, and divers' flippers can crush fragile species on coral reefs. Trawling destroys 15 million km^2 of seafloor communities each year, 150 times greater than the area of forest cleared annually. The subtlest form of environmental degradation is pollution, commonly caused by pesticides, sewage, fertilizer runoff from agricultural fields, industrial chemicals, and wastes, emissions from factories and cars, and sediment deposits from eroded hillsides. The general effects of pollution on water quality, air quality, and even global climate are causes for great concern, not only because of the threats to biological diversity, but also because of their effects on human health.

Pesticides

The dangers of pesticides were brought to world's attention in 1962 by Rachel Carson's influential book *Silent Spring*. Carson described a process, now known as biomagnification, through which (dichlorodiphenyltrichloroethane DDT) and other organochlorine pesticides become concentrated as they ascend the food chain. These pesticides, at that time widely used to kill agricultural pests and mosquito larvae were harming wildlife populations, especially birds that ate large amounts of insects, fish, or other animals exposed to DDT and its byproducts. Birds with high levels of pesticides in their tissues, particularly raptors such as hawks and eagles, became weak and tended to lay eggs with abnormally thin shells that cracked during incubation. As a result of failure to raise young and the outright death of many adults, populations of these birds showed dramatic declines throughout the world.

Recognition of this situation in the 1970s led many industrialized countries to ban the use of DDT and other chemically related pesticides. The ban eventually allowed the partial recovery of many bird populations, most notably peregrine falcons (*Falco peregrinus*), ospreys (*Pandion haliaetus*), and bald eagles (*Haliaeetus leucocephalus*). Nevertheless, the continuing massive use of pesticides and DDT in other countries is still a cause for concern, not only for endangered animal species, but also for the potential long-term effects on people, particularly the workers who handle these chemicals in the field and the consumers of agricultural products treated with these chemicals. These chemicals are widely dispersed in the air and water and are persistent in the environment even in countries where their use has been outlawed.

Water Pollution

Water pollution destroys important food sources and contaminates drinking water with chemicals that can cause immediate and long-term harm to human health. Water pollution also often severely damages aquatic ecosystems. Rivers, lakes, and oceans are used as open sewers for industrial and residential waste. Pesticides, herbicides, oil products, heavy

metals (such as mercury, lead, and zinc), detergents, and industrial wastes can kill aquatic organisms outright or make the environment so inhospitable that species can no longer thrive. For instance, water pollution is a threat to 90% of the endangered fishes and freshwater mussels in the US.

Unlike terrestrial dumps, whose effects are primarily local, toxic wastes in aquatic environments diffuse over a wide area. Many aquatic environments are naturally low in essential minerals, such as nitrates and phosphates, and aquatic species have adapted to the natural scarcity of minerals by developing the ability to process large volumes of water and to concentrate these minerals. When these species process polluted water, they concentrate toxic chemicals along with the essential minerals, which can eventually poison the plant or animal. Species that feed on these aquatic species then ingest these high concentrations of toxic chemicals.

Essential minerals that are beneficial to plant and animal life can become harmful pollutants at higher levels. Human sewage, agricultural fertilizers, detergents, and industrial processes often release large amounts of nitrates and phosphates into aquatic systems, causing cultural eutrophication. For instance, humans release as much nitrate into the environment as do all natural processes, and this input is expected to increase in tandem with the increasing human population. Even small amounts of these nutrients can stimulate plant and animal growth, and high concentrations often result in thick blooms of algae at the surface of ponds, lakes, and coastal areas. These algal blooms may be so dense that they outcompete with other plankton species and shade out bottom-dwelling plant species. As the algal mat becomes thicker, its lower layers die and sink. Bacteria and fungi then decompose the dying algae, absorbing all the oxygen in the water. Without oxygen, much of the remaining animal life dies off, sometimes visibly in the form of masses of dead fish floating on the water surface. The result is a greatly impoverished and simplified community consisting of only those species tolerant of polluted water and low oxygen levels. The spreading dead zone where the Mississippi River enters the Gulf of Mexico is an example of the direct consequences of water pollution. Such dead zones are increasing in number and size around the world as a result of human activities.

Air Pollution

In the past, people assumed that the atmosphere was so vast that materials released into the air would be widely dispersed and their effects would be minimal. But today, several types of air pollution are so widespread that they damage whole ecosystems.

Acid Rain

Acid rain is created when nitrates and sulfates released into the air by the burning of fossil fuels combine with atmospheric water to form acids that fall as rain. Acid rain lowers the pH of soil moisture and water bodies. Increased acidity alone damages many plant and animal species; for instance, acid rain has been blamed for the death of large number of trees in Europe and North America. As the acidity of water bodies increases, many fish either die or fail to spawn (**Figure 6**). Both increased acidity and water pollution are two likely causes of the dramatic decline in amphibian populations throughout the world.

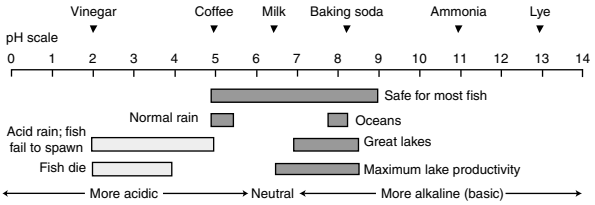


Figure 6 The pH scale, indicating ranges at which acidity becomes lethal to fish. Studies indicate that fish, amphibians, and invertebrates are disappearing from heavily acidified lakes. Reproduced from Cox GW (1993) *Conservation Ecology: Biosphere and Biosurvival*. Dubuque, IA: W. C. Brown; based on data from the US Fish and Wildlife Service.

Ozone Depletion and Ultraviolet Radiation

As a result of human use of chlorofluorocarbons (CFCs) and other ozone-depleting chemicals, the atmospheric ozone layer has been significantly reduced. Ozone plays an important role in filtering out harmful ultraviolet radiation in sunlight. With less atmospheric ozone, more solar ultraviolet radiation reaches the Earth's surface. In humans, exposure to this UV radiation increases the risk of skin cancer. This increased UV radiation will also possibly have a significant negative impact on animals and plants exposed to direct sunlight, such as amphibian eggs on the water surface.

Ozone and Smog

Cars, power plants, and other industrial activities release hydrocarbons and nitrogen oxides as waste products. In the presence of sunlight, these chemicals react with the atmosphere to produce ozone and other secondary chemicals, collectively called photochemical smog. Although ozone in the upper atmosphere is important in filtering ultraviolet radiation, a high concentration of ozone at ground level damages plant tissues and makes them brittle, harming biological communities and reducing agricultural productivity. Ozone and smog are detrimental to both people and animals when inhaled, so controlling air pollution benefits both humans and biodiversity.

Effects on Lichens

Even when communities are not destroyed by air pollution, species composition may be altered as more susceptible species are eliminated. Lichens, symbiotic organisms composed of fungi and algae that can survive in some of the harshest natural environments, are particularly susceptible to air pollution. Because each lichen species has a distinct tolerance to air pollution, the composition of the lichen community can be used as a biological indicator of the level of air pollution.

Global Climate Change

Scientists are now intensively studying atmospheric carbon dioxide, methane, and other greenhouse gases that are

transparent to light but that absorb heat. During the past century, global levels of carbon dioxide (CO_2), methane, and other trace gases have been steadily increasing, primarily as a result of burning coal, oil, and natural gas. Clearing forests to create farmland and burning firewood for heating and cooking also contribute to the rising atmospheric concentration of CO_2, which has increased from 290 to 387 parts per million (ppm) over the last 100 years and is projected to double in the latter half of this century. Humans currently release about 70 million tons of CO_2 into the atmosphere every day.

Climate Change Predictions

Many scientists agree that these increased levels of greenhouse gases have affected the world's climate already, and that these effects will intensify in the future. The best evidence suggests that world climate has warmed by 0.6° Celsius (°C) over the last 100 years. Temperatures at high latitudes have increased even more, by 2–4 °C. Over the last 50 years, the Atlantic, Pacific, and Indian oceans have increased in temperature by 0.06 °C. Even with all of the available weather data, simulation models, and supercomputers, predicting future weather patterns is extremely complex and difficult. However, the consensus among leading scientists is that the global climate will increase in temperature by an additional 2–4 °C by 2100 as a result of increased levels of carbon dioxide and other gases in the atmosphere (**Figure 7**). The increase could be even greater if carbon dioxide levels rise faster than predicted, but it could also be slightly less if all countries agree to reduce emissions of greenhouse gases. The increase in temperature will be greatest at high latitudes and over large continents. Many scientists also predict an increase in extreme weather events such as hurricanes, flooding, snowstorms, fires, and regional drought.

Extinctions and Climate Change

As a result of global climate change, climatic regions in the northern and southern temperature zones will be shifted toward the poles, forcing species to migrate. It seems likely that many species will be unable to disperse rapidly enough to keep up with the changing climate. Habitat fragmentation caused by human activities may further slow or prevent many species from migrating to new habitats or escaping rising sea levels. Many species with limited distributions and poor dispersal ability will undoubtedly go extinct, with new communities favoring widely distributed and easily dispersed species. Endemic mammals that are restricted to isolated mountain peaks or fish species found in a single lake are examples of species that will not be able to easily cross inhospitable terrain to reach a new habitat. The best hope for many species will be to migrate higher on mountain slopes or to disperse along valleys, rivers, and coastlines that are aligned north to south. Warming waters and rising sea levels are already affecting marine species as well.

Concerns about global climate change, as important as they are, should not divert our attention from the massive habitat destruction that is the principle current cause of species extinction. The preservation of intact communities and the restoration of degraded communities are the most important and immediate priorities for conservation.

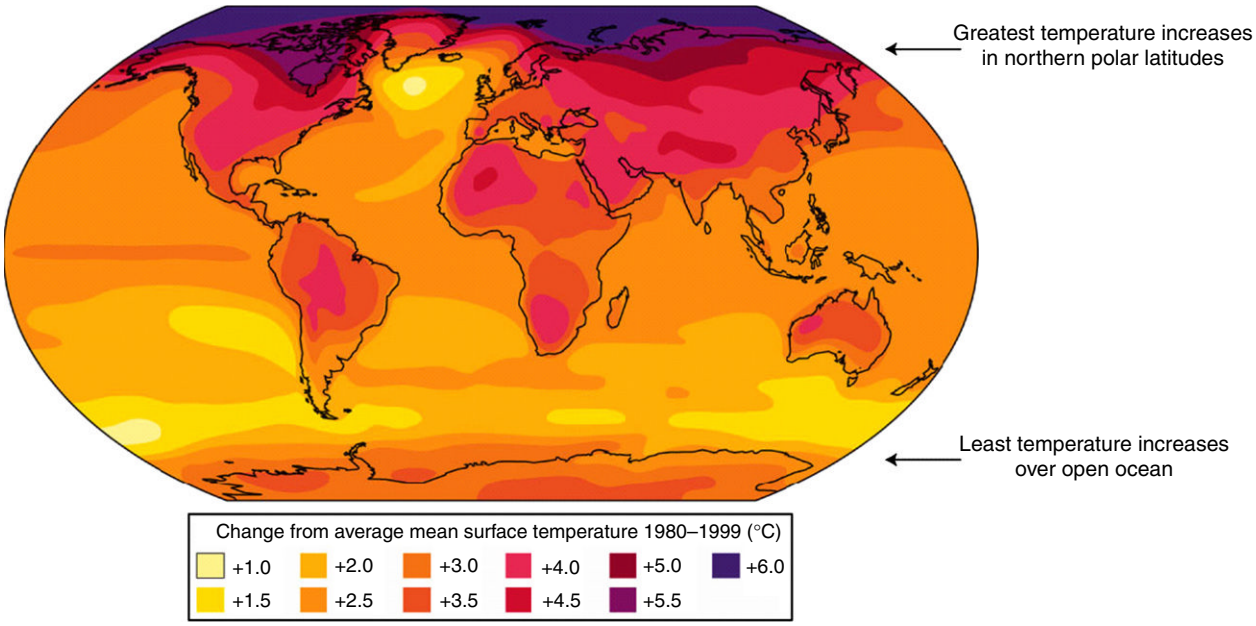


Figure 7 Computer models of global climate predict that temperatures will increase significantly when CO_2 levels double in the mid- to late part of this century. Predicted temperature increases for 2080–2099 are shown, indicated as the amount of deviation (in °C) from mean surface temperatures recorded for 1980–1999. Reproduced from Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates, and Intergovernmental Panel on Climate Change (IPCC) (2007) Climate change 2007: The physical science basis. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press, with permission from IPCC.

Overexploitation

Exploitation in Traditional Societies

People have always hunted and harvested the food and other resources they need in order to survive. As long as human populations were small and the methods of collection simple, harvesting and hunting were sustainable. In traditional societies, restrictions were often imposed to prevent overexploitation of natural resources. For example, the rights to specific harvesting territories were rigidly controlled; hunting in certain areas was banned; there were often prohibitions against taking females, juveniles, and undersized individuals; certain seasons of the year and times of the day were closed for harvesting; and certain efficient methods of harvesting were not allowed. These restrictions, which allowed traditional societies to exploit communal resources on a long-term, sustainable basis, are almost identical to the rigid fishing restrictions imposed on and proposed for many fisheries in industrialized nations. Among the most highly developed restrictions were those of the traditional or artisanal societies of Micronesia and Polynesia. However, there are also numerous cases of large bird and mammal species being hunted to extinction by traditional people using simple methods of hunting.

Exploitation in Modern Societies

As human populations have increased in number and size, their use of the environment has escalated, and harvesting methods have become dramatically more efficient. This has depleted large animals almost completely from many biological communities, leaving strangely empty habitats. In tropical rain forests and savannas, guns are now used instead of blowpipes, spears, or arrows. In the oceans, fish are caught by enormous factory ships and sold in the global market. Even small-scale local fishermen now have outboard motors on their canoes and boats, allowing them to harvest from a wider area more rapidly.

In much of the world today, resources are exploited opportunistically. If a market exists for a product, local people will search their environment to find and sell it. Whether people are poor and hungry or rich and greedy, they will use whatever methods are available to secure that product. Sometimes traditional groups will sell the rights to a resource, such as a forest or mining area, for money to buy desired goods. In rural areas, traditional controls that strictly regulated the extraction of natural products have generally weakened, if not eliminated by substantial human migration, civil unrest, or war. As a result, species are more easily exploited to the point of extinction.

Trade in Wildlife

The legal and illegal wildlife trade is responsible for the decline of many species. Worldwide trade in wildlife is valued at over $10 billion per year, not including edible fish. The fur trade has reduced the chinchilla (*Chinchilla* spp.), vicuña (*Vicugna vicugna*), giant otter (*Pteronura brasiliensis*), and numerous cat species to low numbers. Overharvesting of butterflies by insect collectors; of orchids, cacti, and other plants by horticulturists; of marine mollusks by shell collectors; and of tropical fish for aquarium hobbyists are other examples of targeting whole biological communities to supply an enormous international demand (**Table 1**). For instance, it has been estimated that 350 million tropical fish valued $1 billion are sold worldwide for the aquarium market, and many times that number are killed during collection and shipping.

Besides a surprisingly large legal trade, billions of dollars are involved in the illegal trade of wildlife. A black market links poor local people, smugglers, corrupt customs officials, rogue dealers, and wealthy buyers who do not question the sources they buy from. This trade has many of the same

Table 1 Major groups targeted in the Global Wildlife Trade

Group	*Number traded annually*[a]	*Comments*
Primates	40,000	Mostly used for biomedical research; also for pets, zoos, circuses, and private collections.
Birds	4 million	Zoos and pets. Mostly perching birds, but also legal and illegal trade of about 80,000 parrots.
Reptiles	640,000	Zoos and pets; also 10–15 million raw skins. Reptile extracts are used in some 50 million manufactured products (mainly from the wild, but increasingly from farms).
Ornamental fishes	350 million	Most saltwater tropical fish come from the wild, caught using illegal methods that damage other wildlife and the surrounding coral reefs.
Reef corals	1000–2000 tons	Reefs are being destructively mined to provide aquarium decor and coral jewelry.
Orchids	9–10 million	Approximately 10% of internationally traded orchids are from the wild. These are sometimes deliberately mislabeled to avoid regulations.
Cacti	7–8 million	Approximately 15% of the traded cacti come from the wild, with smuggling a major problem.

[a]Numbers refer to the number of individuals unless otherwise specified.

Source: Based on data from World Resources Institute (WRI) (2005) *World Rresources 2005: The Wealth of the Poor – Managing Ecosystems to Fight Poverty*. Washington, DC: World Resources Institute; Karesh WB, Cook RA, Bennett EL, and Newcomb J (2005) Wildlife trade and global disease emergence. *CDC Emerging Infectious Diseases* 11: 100–1002.

characteristics and practices, and sometimes the same criminal players, as the illegal trade in drugs and weapons. Confronting these illegal activities has become a dangerous job for international law enforcement agencies. Clearly, people involved in the illegal trade of wildlife are not concerned about species going extinct, unless it affects their profits.

Overfishing

In the North Atlantic, one species after another has been overfished to the point of diminishing returns. The Atlantic bluefin tuna, for example, has experienced a 97% population decline since 1960. Similar grim scenarios can be recounted for other prized large fish, such as the swordfish (*Xiphias gladius*). One of the most dramatic cases of overexploitation in recent years is related to the booming demand for shark meat and shark fins, caused in large part by the popularity of shark fin soup in Asian restaurants. Shark fishing has become a lucrative alternative to targeting severely depleted commercial fish populations. But many shark populations are now declining dramatically as well, because most species have a relatively slow reproductive cycle. Heavily fished shark populations in the Atlantic, for instance, have declined by 40–99% over the last 20 years, and some species there and elsewhere may soon go extinct.

Another striking example is the enormous increase in demand for seahorses (*Hippocampus* spp.) in China, which is tied to the nation's economic development. The Chinese use dried seahorses in their traditional medicine because it resembles a dragon and is believed to have a variety of healing powers. Approximately 45 t of seahorses are consumed in China per year – roughly 16 million animals. Seahorse populations throughout the world are being decimated to supply this ever-increasing demand, and international trade is now carefully regulated.

Exotic Species

Humans have radically altered the patterns of species distributions by deliberately or accidentally transporting species throughout the world. The extent of this modern movement of human-transported species is unprecedented and has been described by Elton (1958) as "one of the great historical convulsions of the world's flora and fauna." Many areas of the world are strongly affected by exotic species. For instance, the US currently has over 20 species of exotic mammals, 53 species of exotic reptiles and amphibians, 88 species of exotic molluscs, 97 species of exotic birds, 138 species of exotic fish, 4500 species of exotic insects and other arthropods, and 5000 species of exotic plants.

The great majority of exotic species do not become established or dominant because the new environment is not suitable. However, some species do establish themselves in their new homes, and many of these become abundant at the expense of native species. Exotic plants frequently displace native species because they are better suited to the new conditions created by people, such as increased fire and introduced grazing animals. Exotic perennials completely dominate many North American wetlands: purple loosestrife (*Lythrum salicaria*) from Europe overgrows marshes in eastern North America, whereas Japanese honeysuckle (*Lonicera japonica*) forms dense tangles in bottomlands of the southeastern US.

Exotic species are also often able to thrive because their populations are not held in check by any of the local parasites or predators. These exotic species may displace native species through competition for limited resources, they may kill and eat native species to the point of extinction, or may alter the habitat so that many natives are no longer able to persist. The effects of exotic insects on the native insect fauna can be devastating. At some localities in the southern US, insect species diversity has declined by 40% following the invasion of exotic fire ants, which attack, consume, or outcompete other insect species. Many bird species have dramatically declined after fire ants entered their habitat, because of direct attack or loss of insect prey.

Exotic species are considered the most serious threat facing the biota of the national park system in the US. The effects of habitat degradation, fragmentation, and pollution can potentially be corrected and reversed in a matter of years or decades as long as the original species are present. It may be expensive, difficult, or even impossible, however, to eliminate exotic species that have become abundant, widely dispersed, or thoroughly integrated in the community.

Exotic Species on Islands

Because they have evolved in the absence of mainland herbivores and predators, island species are particularly vulnerable to exotic species; the introduction of one exotic species to an island may cause local extinction of numerous native species.

On Santa Catalina Island off the coast of California, 48 native plant species have been eliminated, primarily due to grazing by introduced goats, pigs, and deer. One-third of the plant species currently found on the island are exotics. Several native plant species have reappeared after goats were removed from part of the island.

Birds of the Pacific are also vulnerable to exotic species. The brown tree snake (*Boiga irregularis*; **Figure 8**) has been introduced onto a number of Pacific islands, where it is devastating endemic bird populations by eating eggs, nestlings, and



Figure 8 The brown tree snake (*Boiga irregularis*) has been introduced onto many Pacific islands, where it devastates populations of endemic birds. This adult snake has just swallowed a bird. (Photograph by Julie Savidge).

adults. On Guam alone, the brown tree snake has driven 10 of 13 endemic bird species extinct. Recent visitors have remarked on the absence of birdsong: "Between the silence and the cobwebs, the rain forests of Guam have taken on the aura of a tomb" (Jaffe, 1994).

Exotic Species in Aquatic Habitats

Exotic species can have severe negative effects on vulnerable freshwater communities, particularly lakes and isolated stream systems. Shipping is a major cause of introductions, moving species that are attached to ship hulls or contained in ballast water. There is also a long history of introductions, many accidental, of exotic commercial and sport fish species into lakes. These exotic fish, often larger and more aggressive than the native fish fauna, may eventually drive the local fish species to extinction. Aquatic plants, invertebrates, and disease organisms can also become aggressive exotics outside their normal range.

One of the most alarming recent invasions in North America was arrival of the Eurasian zebra mussel (*Dreissena polymorpha*) in the Great Lakes in 1988. Within two years, zebra mussels had reached densities of 700,000 individuals per square meter in parts of Lake Erie, choking out native mussel species. They have subsequently been found in the Detroit, Cumberland, and Tennessee Rivers. As it spreads south, this exotic species is causing enormous economic damage to fisheries, dams, power plants, water treatment facilities, and boats, as well as devastating the aquatic communities it encounters.

Invasions also occur in marine and estuarine systems, with 84% of marine areas worldwide affected by at least one invasive species. In the Mediterranean, the invasive green alga *Caulerpa taxifolia* is outcompeting native algae species and reducing fish abundance. Because *Caulerpa* is used as a decorative plant in marine aquariums, this is an example of one negative human impact (the trade in aquarium and ornamental species) resulting in another. One solution would be to ban harmful invasive species from the aquarium trade and provide a list of noninvasive alternatives.

Disease

Disease caused by internal parasites is a natural control mechanism that reduces populations when they reach high densities. However, levels of disease can often increase in populations as a result of human activity. When animals are confined to habitat fragments at abnormally high densities, disease may spread more easily among individuals. Also, animals that are stressed or weakened by living in a degraded or polluted environment may be more susceptible to disease. Furthermore, as habitats are fragmented by human activities, disease can spread more easily from domestic animals into wild populations. At Tanzania's Serengeti National Park, at least 25% of the lions (*Panthera leo*) have recently been killed by canine distemper, a viral disease apparently contracted from one or more of the 30,000 domestic dogs living near the park. For endangered species, such outbreaks can do phenomenal harm: The last population of black-footed ferrets (*Mustela nigrepes*) known to occur in the wild was destroyed by canine distemper virus. Ironically, some conservation efforts may also increase susceptibility to disease; for instance, species in captive breeding programs may contract diseases from other animals or even humans.



Figure 9 Populations of flowering dogwood (*Cornus florida*) are declining in eastern North American forests due to anthracnose disease, which is caused by the introduced fungus *Discula destructiva*. (Photograph by Jonathan P. Evans).

Diseases transported by people to new parts of the world can also decimate species. North American chestnut trees (*Castanea dentata*), once common throughout the eastern US, have been virtually obliterated by an ascomycete fungi carried on Chinese chestnut trees imported to New York City. Introduced fungal diseases are also eliminating elm trees (*Ulmus americana*) and flowering dogwoods (*Cornus florida*) from North American forests (**Figure 9**). Introduced diseases have particularly powerful adverse effects on endemic island species. Many endemic Hawaiian bird populations have been decimated and even driven to extinction by introduced avian malaria protozoans spreading from introduced bird species through introduced mosquitoes.

Increased rates of disease in animal populations can also negatively impact humans. The risk of diseases such as West Nile virus, Lyme disease, and bird flu being transmitted from animals to humans is increasing with increased movements of humans and domestic and wild animals around the world, as well as with warming temperatures that may be allowing disease-causing organisms to expand their ranges.

Multiple Factors

A combination of factors acting simultaneously or sequentially can overwhelm a species, as illustrated by the case of the large freshwater mussel *Margaritifera auricularia*. This species was formerly known from Western Europe to Morocco, but now it occurs in only one river and its adjoining canals in Catalonia, Spain. Humans have used its attractive shell and pearls as ornaments since the Neolithic Age. The main reason for its decline, overharvesting, originally led to its disappearance from rivers in Central Europe in the fifteenth and sixteenth centuries. In more recent times, pollution, destruction of freshwater habitats, and overharvesting continue to reduce

its range. The mussel is also affected by the loss of other species, because its larval stage needs to attach to certain species of fish to complete its life cycle. Unless strict conservation measures are implemented to prevent overharvesting, control water quality, maintain fish stocks, and protect mussel habitat, this culturally important species will soon be extinct. Such comprehensive conservation strategies are often needed to deal with the multiple threats to species.

Threats to biological diversity come from a number of different directions, but their underlying cause is the same: the magnitude of destructive human activity. It is often easy to blame a group of poor, rural people, or a certain industry for the destruction of biological diversity, but the real challenge is to understand the local, national, and international linkages that promote the destruction and to find viable alternatives. These alternatives must include stabilizing the size of the human population, finding livelihoods for rural people that do not damage the environment, providing incentives and penalties that will convince industries to value the environment, and restricting trade in products that are obtained by damaging the environment. Yet an equally important part of the solution is to increase the willingness of wealthy and middle-class people in both developed and less-developed countries to reduce their consumption of the world's resources and to pay fair prices for products that are produced in a sustainable, nondestructive manner.

See also: Air Pollution. Comparing Extinction Rates: Past, Present, and Future. Deforestation and Land Clearing. Desertification. Diseases, Conservation and. Environmental Impact, Concept and Measurement of. Extinction in the Fossil Record. Habitat Loss and Fragmentation. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Introduced Species, Impacts and Distribution of. Latent Extinction—The Living Dead. Loss of Biodiversity, Overview. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of. Modern Examples of Extinctions. Natural Extinctions (not Human Influenced). Pesticides, Uses and Effects of

References

Bradshaw CJA, Sodhi NS, and Brook BW (2009) Tropical turmoil: A biodiversity tragedy in progress. *Frontiers in Ecology and the Environment* 7: 79–87.

Bryant D, Burke L, McManus J, and Spalding M (1998) *Reefs at Risk: A Map-Based Indicator of Threats to the World's Coral Reefs*. Washington, DC: World Resources Institute.

Cox GW (1993) *Conservation Ecology: Biosphere and Biosurvival*. Dubuque, IA: W. C. Brown.

Elton CS (1958) *The Ecology of Invasions by Animals and Plants*. London, UK: Methuen.

Intergovernmental Panel on Climate Change (IPCC) (2007) Climate change 2007: The physical science basis. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

Jaffe M (1994) *And No Birds Sing: A True Ecological Thriller Set in a Tropical Paradise*. New York: Simon and Schuster.

Karesh WB, Cook RA, Bennett EL, and Newcomb J (2005) Wildlife trade and global disease emergence. *CDC Emerging Infectious Diseases* 11: 100–1002.

Pimm SL and Jenkins C (2005) Sustaining the variety of life. *Scientific American* 293(33): 66–73.

Primack RB (2010) *Essentials of Conservation Biology*, 5th edn. Sunderland, MA: Sinauer Associates.

World Resources Institute (WRI) (2005) *World Resources 2005: The Wealth of the Poor – Managing Ecosystems to Fight Poverty*. Washington, DC: World Resources Institute.

Extinction in the Fossil Record

Jeffrey S Levinton, State University of New York, Stony Brook, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Background extinction A distinctly lower rate of extinction, more typical of most of the fossil records.

Extinction rate Number or proportion of taxa becoming extinct per unit time or after an important geological temporal boundary.

Mass extinction Extinction occurring over a short period of time that is of large magnitude, wide biogeographic impact, and involves the extinction of many taxonomically and ecologically distant groups.

Introduction

Why the Fossil Record?

Many species are threatened by impending extinction and attempts have been made to assess population declines and to enact policies of recognizing endangerment by rules, such as the International Conservation Union's rule of three successive years of 80 percent decline. It is difficult to develop a measure of extinction rates of entire flora or fauna, only because we usually have scant knowledge of the species pool before the impact. This is particularly a problem in species-rich tropical habitats, where cryptic species abound, yet have not been identified completely. Surveys of specially rich fauna in tropical wet forests are only being organized now, and the decline of the areal extent of these forests in recent years makes it nearly impossible to measure extinction rates, except by means of indirect estimates of species–area relationships. The same applies to species-rich marine communities such as coral reefs.

Extinction is very much the domain of a paleontologist. It is believed that we are now possibly living through a mass extinction caused by human disturbance of high diversity tropical habitats. But very little is known about the extinction of species, except by fairly obvious mechanisms such as hunting. Can knowing that the Dodo or the Passenger Pigeon was hunted to extinction help very much with understanding climatically induced changes in key structural groups such as forest trees and reef corals, and their dependent species? Could such spotty knowledge be used to extrapolate to the broad sweep of geological time? Paleontological data has the advantage of large banks of "before and after" data on biodiversity on time scales that can be related to global climate change. Its weakness, however, is in associating extinctions with unique causes, as shall be seen.

Invasions have caused extensive extinctions on oceanic islands, particularly when alien predators overwhelmed small populations of endemic species in a matter of decades. Extinction on larger time scales, even over hundreds of years, is much more difficult to track. Unfortunately, the time scale for larger-scale changes over 100,000 years or more is probably unapproachable by the neontologists, who can only observe "normal" extinction, and don't have much understanding of what occurs normally. It may be that fine-scale studies of the fossil record may eventually give us more insight into species-level extinction than neontological studies ever will. After all, the durations of animal species life spans range from the order of 10^5 (land vertebrates) to 10^6–10^7 years (marine species). Even with millions of living species, we are not likely to be able to document many cases of typical extinction of living animal species. The fossil record is probably our only hope of a model for study of extinction rates, especially on the scale of ocean basins and continents.

Although its coverage of the total potential living biota is incomplete, the fossil record affords us a more complete glimpse of extinction rates of a number of readily fossilizable groups, both marine and terrestrial. We have a reasonably complete database that has stabilized over the years and one can readily trace extinctions across geological time horizons. A great advance is the Paleobiology Data Base (http://paleodb.org), which records as of the year 2011 nearly a million fossil occurrences with a wide variety of synthesis of data. The fossil record moreover gives us a deeper insight into what extinction really means. After all, we would like to produce a prospectus of the biological future of living communities following an extinction. Does the loss of a species have a disproportional importance, resulting in the extinction of many associated species? Following an extinction event, is there enough redundancy that allows the surviving species to evolve a new diverse fauna? Does the extinction of certain species cause the snowballing of a larger extinction event? With some judicious reasoning we can infer the answers to some of these questions with the use of the fossil record.

Measures and Types of Extinction Rates

Many extinctions in the fossil record appear to be precipitous, and occur over short time period of hundreds of thousands to a few million years. Impacts of extraterrestrial objects may have caused changes in a year or less. Such length of time is short when you consider the length of the record of the Phanerozoic Era (542 My).

It is possible to quantify the extent of the extinction with the following data:

1. The total pool of taxa before the extinction;
2. the number of taxa that became extinct;
3. the time span over which the extinction occurs.

We could calculate the number of taxa that become extinct or the percent of the former pool of taxa that became extinct. Time could be measured in years, but often we only have segments of relative geological time units such as geological stages. (A set of stages comprises a geological series, and a set of series comprises a geological period (e.g., Cambrian, Cretaceous).) In many parts of the fossil record, the absolute time represented by a stage is not accurately known, and different geological stages are often of great difference in temporal extent.

Charles Lyell developed an ingenious technique to estimate extinction rate by charting the gradual diminution of living species as one went back in geological time. This type of analysis can give longevities and extinction rates. Such Lyellian curves demonstrate, for example, that the diminution of bivalve species on the Pacific coast of the USA is at a steady pace whereas a more precipitous extinction occurred in the Atlantic.

Problems in Measuring Extinction Rates

Taxon-Level Bias

We typically think of extinction rate as a measure of the loss of species. To create a database for paleontology, the species level is very difficult to trust with any degree of confidence; most paleontologists tend to trust the genus and higher taxonomic levels in identifications. In recent years, more and more effort has been directed toward accounting for the ranges of all named species in the fossil record, but most analyses have been done at the family or genus level. The large-scale database we now employ owes its existence to the dedicated work of David Raup and especially the late Jack Sepkoski, who continuously sought to produce a more complete database of the geological ranges of all fossil groups. Initially, the compilation was at the level of taxonomic order but subsequent analyses have moved down the taxonomic hierarchy to the family and generic levels.

Can extinctions of higher-level taxa be used to estimate species-level extinctions? To estimate species richness using numbers of higher-level taxa (e.g., orders), we assume that taxonomic diversity at higher taxonomic levels is correlated with species richness, but they are not necessarily correlated in this manner. This can be seen clearly where changes in ratios of one taxonomic level to another occur over broad spans of time. If the ratios change, then higher taxonomic units might be flawed estimators of changes in species diversity. For example, the ratio of taxonomic orders to families decreased significantly from the Paleozoic to the Mesozoic Era.

Over short periods of time, the number of taxa at a higher taxonomic level (e.g., level of family) might have a regular relationship with a lower taxonomic level, such as species. To estimate species-level extinction from family-level extinction, David M. Raup used a rarefaction technique based on the sampling curve that relates the number of species collected at random to the number of families recovered. This method has been modified and refined to be used in many estimates of standing fossil diversity.

The rarefaction approach is the best we have so far. Nevertheless, we must be careful in applying it. The biggest problem is the potential change in the relationship over geological time. For example, the ratio of families to species decreases by a factor of two from the Mesozoic to the Cenozoic, and other cases are known of changing ratios of species to genera. Selective extinction can also bias our conclusions. For example, certain families may be much more prone to extinction, due to their presence in a particularly vulnerable habitat (e.g., coral reefs during a cooling event). This might overestimate total extinction, if these are added to a larger species list. It is also difficult to get sufficient data to calculate good rarefaction curves for all but the most abundant fossil groups.

Raw Data for Analyses

Any set of diversity data involves a count of the number of taxa. But this number is strongly modulated by the intensity of sampling. Thus diversity could be inflated simply because more studies have been done in a given locality at a given time interval. Thus raw data must be corrected for sampling intensity. In most instances, the genus is the level used for diversity estimates, so the number of genera must be estimated by some sort of sampling process that corrects for the total number of studies, the number of localities, and the number of fossils that are collected at any one locality and time horizon. Two examples of these potential biases are the possible monograph effect of studying intensively one part of geological time and a bias that inflates the diversity of parts of the geological record just before the present, because species are still alive and can be sampled and studied more completely than their antecedents.

Biased Preservation and Convergence

Estimates of extinction rates may be biased by preservation and abundance at the time of extinction. Preservation of appropriate habitats during an extinction may be greatly reduced. Thus a species might have survived, but there are no opportunities to see it because its usual facies of occurrence has not been preserved.

A common change in probability of preservation occurs when a systematic change in rock preservation occurs, as in the reduction of deposition during a regression phase of the sea, as at the end of the Permian and just before the end of the Cretaceous (regression refers to a lowering of sea level in a given area; transgression refers to a rise in sea level). Suppose the ranges of a group of species all ended at the very terminus of the Cretaceous. A gradual reduction of deposition would, by sampling error alone, give the impression of a gradual disappearance of the fossil species. Even if deposition does not decline, previously rarer species would be difficult to sample for presence during a general decline in abundance during extinction, just because we would be unlikely to find them. These biases have come to be known as the Signor–Lipps effect, or "backward smearing," because a sharp extinction might appear to be gradual from fossil sampling. Only abundant forms would be sufficiently "findable" that we could assess their total geological range with confidence, especially up to the time of their extinction. Even at smaller levels of geological resolution, the problem of a spurious component of estimated diversity due to the preservation of different amounts of rock of a given age is a major source of bias. The problem is compounded by cases where some environmental factors

simultaneously affect both the amount of sediment accumulation for a given time period and biodiversity.

Accurate Estimate of Fossil Ranges

To estimate extinction rates one must have an accurate accounting of the geological ranges of species. Then at any time horizon one would be able to account for the number of taxa that disappear from below to above the horizon.

Preservation and rock distributions may strongly bias our perception of geological ranges of taxa. All geological ranges of fossil groups are incomplete, due to lack of appropriately preserved habitats and poor preservation. As incompleteness, or gaps, between fossil occurrences in a vertical section increases, it stands to reason that the actual temporal range of a taxon is greater than what the record would indicate. The number of gaps in preservation may also be combined with a fossil recovery potential curve, which might correct for a change in the probability of preservation during the history of the taxa in question. On the grand scale, the volume of rock correlates positively with the total number of fossil taxa recovered in both marine and terrestrial environments; this suggests that incompleteness of preservation of environments may give us a false impression of true diversity. Although methods have been developed to relate the degree of incomplete preservation to true geological range, the estimated error may sometimes make it very difficult to determine whether, for example, many species became extinct in the same short time period. Some proportion will appear to have disappeared earlier due to incomplete preservation.

Extinction Must be Compared with Origination Rates

If extinction occurs over a very short period of time, one can count the number of species before and after an event and assume that the decline can be explained by extinction alone. But speciation may be occurring continuously, which means that a decline in species richness may just as easily stem from a drop in speciation rate as an increase in extinction. Alternatively, speciation might keep pace with extinction, resulting in no loss of biodiversity. When the speciation and extinction rates are equal, an equilibrium exists, meaning the number of taxa remains constant. At the end of the Devonian, for example, large speciation rates balanced high extinction rates. But speciation rates collapsed during the latest Frasnian (Upper Devonian), which precipitated a severe reduction of marine species diversity. A dramatic extinction of mollusk species occurred in the subtropics of the western north Atlantic at the end of the Pliocene (*ca.* 3 million years ago) but this loss was more than compensated by the origin of new species. Problems also arise when species appear and disappear during an interval. Counts before and after this interval will therefore ignore a significant fraction of the appearances and extinctions.

Pseudoextinction

In many cases, paleontologists have followed lineages through a geological column and have named successions of species, which are recognized by a variety of character transformations. Thus, even though a lineage may not become extinct, the morphological changes result in an arbitrary extinction or pseudoextinction. Pseudoextinctions are a significant fraction of the total disappearances of taxonomic names from one geological horizon to the next.

Phylogenetic Approaches

We are mainly considering taxic approaches to extinction, where the number of taxa appearing and becoming extinct are to be estimated. But much can be learned when considering the evolutionary position of taxa on a tree of evolutionary relationships. If we had a complete evolutionary tree, complete with living species and all preserved species that became extinct, we could estimate the extinction rate as the number of species that became extinct divided by the total of the branch lengths of the phylogeny. Because most fossil groups are incompletely sampled, it is usually not possible to take such a simple approach. Instead, models incorporating estimates of appearance and extinction rates in relation to tree structure and independent estimate of clade age (e.g., molecular clock estimates) can be used to estimate the extinction rate.

Other Biases

Paleontologists are accustomed to dealing with a wide range of problems in preservation. Occasionally, an exquisitely preserved fossil biota, such as the Middle Cambrian Burgess Shale, demonstrates that most of the remaining fossil records have not preserved a wide variety of soft-bodied species and even a number of skeletonized taxa. Such unevenness of preservation also works at smaller scales and therefore preservation strongly biases our estimates of diversity. Monographic studies of the fossil record are also uneven and descriptions of species are often strongly correlated with the intensity of study by specialists, either between fossil groups or between time horizons. Recent studies have attempted to correct these problems by normalizing diversity estimates by the number of monographs produced for a given group at a given time.

Mass Extinction

Definitions and Identification of Mass Extinction

Although diversity in the fossil record was first characterized in 1860 by John Phillips, strong temporal changes in taxon turnover were quantified first in the 20th century by paleontologist Norman Newell, who found peaks of activity in the Ordovician, Carboniferous, and Jurassic. Declines in standing taxon richness were simultaneous and relatively rapid among distantly related taxa, although increases were not so obviously coordinated. We can see in **Figure 1** five conspicuous and precipitous drops in diversity, the most dramatic occurring at the end of the Permian. David M. Raup and John J. Sepkoski Jr analyzed the overall extinction rate of marine taxa at the family level and found that four events fell outside of a one-sided 99 percent confidence interval from the mean extinction rate trend: Ashgillean (Upper Ordovician); Frasnian (Late Devonian); Guadalupe-Dzhulfian (Late Permian); and Maastrichtian (Late Cretaceous). The Norian (Upper Triassic) fails this test but its wide-spread occurrence forces us to include it in the "big five" (**Table 1**). A recent compilation of stratigraphic ranges over a wide variety of taxa produced by a wide variety of specialists. Inspection of extinction rates revealed peaks at the

Table 1 Percent extinction ((extinctions/standing taxon richness) × 100) at the five major mass extinctions in the fossil record (after Jablonski 1994)

Mass extinction	*Families*		*Genera*	
	Observed extinction	*Calculated species-level extinction*	*Observed extinction*	*Calculated species-level extinction*
End-Ordovician 439 Ma	26	84	60	85
Late Devonian 367 Ma	22	79	57	83
End-Permian 243 Ma	51	95	82	95
End-Triassic (Norian) 208 Ma	22	79	53	80
End-Cretaceous 65 Ma	16	70	47	76

same times identified by Sepkoski, and, indeed, by paleontologists traditionally.

The big drops, mass extinctions, have been distinguished from background extinction, which refers to the remainder and overwhelming majority of extinctions. Although the big five are conspicuous, other mass extinctions have been recognized. About half of the marine genera disappeared in the Lower Cambrian and archaeocyathid reefs were decimated, perhaps due to wide-spread marine anoxia. Another possible anoxic event caused a major extinction at the Cenomanian–Turonian boundary (Upper Cretaceous), although lowered productivity and global cooling may have contributed. John Alroy's analysis of the Phanerozoic fossil record suggests that perhaps one cannot justify a distinct class of mass extinctions by means of numerical changes alone.

Statistical tests of mass extinction do not inspire confidence, because they combine many taxonomic groups of complicated taxonomic structure, reify them to independent data points, and usually analyze them using the assumptions of parametric statistics. Because the groups are enmeshed in a phylogenetic tree structure, it is not easy to perform such analyses and distributions of extinction events are usually skewed toward many events of extinction rates of a few percent. There may be as many as 12 mass extinctions.

It may seem inappropriate to fix on mass extinctions, which could be atypical end members, but if we cannot characterize these events, will we be able to explain the smaller extinctions that were far more common in the history of life? If mass extinctions are more or less larger scale or even global versions of what might happen on a more local scale (extinctions stemming from, for example, local tectonism, anoxia, regional sea-level change) then maybe we can extrapolate what we learn about them to smaller scales, and *vice versa*. If we focus on times of heightened turnover in taxon richness that are confined to basins, we might see mass extinctions writ small. There is some reason, however, to believe that the "big five" were distinctive and the effects of extinctions during these times transcended those of more mundane times.

If an extinction event is a statistical outlier, then how could we justify a separate category, and would this category require a set of extinction mechanisms that differ qualitatively from background extinction? This question opens up a can of worms. There has been a good deal of debate about what a statistical outlier really is and whether mass extinctions are really different from lesser periods of extinction. If we assembled extinction rates into a frequency distribution curve, we might argue that the mass extinctions sit squarely on the tails of some expected probability distribution, or a kill curve. The distribution of the risk of marine genera consists of groups with mainly low risk and some of much higher risk. With no mechanistic model in mind, it is not clear whether the somewhat bumpy distribution of extinction rates of marine genera is smooth or discontinuous, and we have no idea of the distribution for species. If we have to wait 100 My for a mass extinction, are we waiting for an intense version of the same stuff or a truly distinctive event?

In the context of the time that they occur, mass extinctions are clear and major drops in taxon richness, distinct from extinction rates in the time periods before and after. The Permian appears as a sharp trough after an Early Paleozoic time of expansion and then stabilization of numbers of taxa. The same can be said for the End-Cretaceous extinction. But numbers really tell only part of the story. A criterion based on a high extinction rate alone would stretch the confidence we have in our statistical assessments too far, and should only be a means of screening for candidates. To qualify, mass extinction events must have the following features:

1. The number of taxa becoming extinct is significantly greater than times of other extinctions.
2. The decline is concentrated in a small fraction of the Phanerozoic, for example, less than a geological series, or at most a few million years.
3. The extinction is broad-based taxonomically, affecting many distantly related taxa that have not arisen in the same time period as the decline.
4. The extinction affects many different biomes, perhaps not equally. For example, a mass extinction would not be confined to epibenthos on hard surfaces of a region, as opposed to coeval soft-bottom benthos living in the same area.
5. The extinction is geographically widespread, most likely global, in extent.
6. Mass extinction may be caused by mechanisms qualitatively different from background extinctions, but it may be caused merely by "much more of the same" (e.g., anoxia, climate change, unusually large extraterrestrial impacts).
7. Mass extinction affects taxa differently than during other extinctions; small-scale influences on the degree of extinction, such as geographic range are swamped during mass extinctions. David Jablonski found that during ordinary extinctions, species with greater geographic ranges were less prone to become extinct.
8. The recovery period following the mass extinction marks the rise of either new taxonomic groups, the expansion of formerly rare groups, or complete reorganization of ecosystem structure.

Table 1 demonstrates two important issues in quantifying the degree of extinction. First, the extinction event is characterized as a loss. But the nature of loss is unclear, as it may result from declining speciation rates, increasing extinction rates, or both. Secondly, in the case of mass extinctions, speciation usually declines and extinction rate increases precipitously. Table 1 shows the startling results: About 95 percent of the marine (readily fossilizable) species became extinct at the end of the Permian, and the others of the "big five" took similarly big hits. The Permian was also bad for taxonomic families (*ca.* 50 percent loss), but the others hovered around 20 percent. The lower loss of families in many extinctions suggests that there might be survivors of many families that could recover and proliferate new species following the extinction.

A compilation of major changes in both terrestrial and marine fossil groups, coordinated by Michael Benton, reveals some important features of the fossil record. For instance, extinction rate can be very high, with not much overall effect on total diversity, because originations may be high or even higher. Thus if one considers extinction rate separately, the Cambrian must be added to our roster of mass extinctions. In terms of percentage extinction, it looms over the rest of the fossil record, even the Permian. There is a possibility, however, that this extinction is more apparent than real. Some have argued that poor Upper Cambrian preservation biases our perception of Cambrian diversity. The analysis reveals a number of extinction peaks not easily seen in a plot of diversity alone. Most notable are strong extinctions in the Carboniferous, Jurassic, and Mid-Cretaceous periods.

The quantitative aspects of mass extinction should not obscure some of the major qualitative effects, causing irreversible changes in the world's biota and therefore major reorganization of the structure of life. At the end of the Permian, marine communities were reorganized completely due to the end of dominance by brachiopods and the extinction of long-abundant forms such as trilobites. In most parts of the world, the end of the Permian was also the end of a sea bed dominated by relatively sluggish and often sessile passive suspension-feeding groups, leading to replacement by sea-bed communities with more active roving individuals. The end of the Mesozoic witnessed the demise of the long-abundant carnivorous ammonites, leaving the modern world with a pitiful representation of the former glory of externally-shelled cephalopods. Bivalves such as the inoceramids and rudists dominated Cretaceous shallow-water seas but they disappeared. Of course, the dinosaurs also became extinct at this time, which presaged the evolution and diversification of the modern orders of mammals.

It is often difficult to estimate the time span over which extinction occurred, due to strong uncertainty in the time span of geological stages, which is often the crucial level at which extinction is assessed. This is particularly a problem in the early and middle part of the Paleozoic, where time estimates of stage lengths are difficult to estimate with certainty. Within stages, rates are especially difficult because of uneven rates of sedimentation, which makes it invalid to assume a linear relationship between meters of geological section and time. For example, during the final flood stage of transgression of the sea, sedimentation rates are often very low, and much time may be compressed in very little geological section.

The estimate of extinction is plagued by other factors as well. Worst of all, the incompleteness of the fossil record imposes biases that cloud an assessment of the tempo and degree of a mass extinction. During the Permian, an enormous regression of the sea resulted in the deposition of few marine deposits and correspondingly few marine fossils. As a result, a number of taxa appeared to become extinct, but they reappeared in the Triassic, much as Lazarus was raised from the dead. Lazarus taxa may be explained by poor preservation in all facies, resulting in poor sampling of species, even those that really existed during the low point at the end of the extinction. Even if appropriate sedimentary rocks are widespread, reductions in population size may make any species more resistant to successful sampling. Alternatively, there may be localized havens in which such taxa may survive, but these refuge environments might escape preservation. In any case, extinction is overestimated. The problem is compounded by so-called Elvis species, which evolved after a mass extinction and converge by means of natural selection to resemble premass extinction morphotypes, much as Elvis impersonators now clutter the landscape, at least if you frequent Las Vegas. In Triassic reefs, sponges may be mistaken for Permian taxa, but they are unrelated.

Preservation and rock distributions may strongly bias our perception of geological ranges of taxa. A common change in probability of preservation occurs when a systematic change in rock preservation occurs, as in the reduction of deposition during a regression phase, as at the end of the Permian and just before the end of the Cretaceous. Suppose the ranges of a group of species all ended at the very terminus of the Cretaceous, a gradual reduction of deposition would, by sampling error alone, give the impression of a gradual disappearance of rarer fossil species. Thus a sudden ending of many taxa can be made to appear gradual merely by a gradual reduction of percent preservation.

Patterns of taxon survivorship of mollusks indicate that mass extinction may or may not be a qualitatively different phenomenon from background extinction. David Jablonski found that during normal periods in the Cretaceous period, extinctions of mollusks were correlated with planktotrophic larval development and geographic range; as might be expected, species with longer dispersal abilities and greater geographic range had greater survival potential. Clade survivorship was positively correlated with species richness. During the End-Cretaceous event, however, none of these generalizations were held, and clade survival was correlated only with the geographic extent of the clade. After the event, the correlations found previously were again obtained. In the End-Cretaceous extinction of planktonic diatoms, however, a different pattern emerged, as diatom species with benthic resting stages survived far better than those with no resting stages. Foraminifera species with specialized morphologies and larger sizes were eliminated and simpler morphologies were favored. Sea urchins suffered extensive extinctions across the K-T boundary, but bulk sediment processors and shallow-water herbivorous species suffered more extinction than omnivores or selective deposit feeders, which suggests a relationship between high extinction and starvation. Here, properties that normally would be related to survival of individuals can be extrapolated to taxon survival. During the Permian mass

extinction, gastropod success did not differ especially from periods of more subdued extinction, but groups with planktotrophic larvae and geographically restricted groups suffered more than average (Erwin, 1993). Thus, as extinction intensity increases, some qualitative changes may emerge for some taxa and biogeographic/dispersal properties, but not for all.

Causes of Mass Extinctions

Mass extinctions are associated in time with major environmental changes. The problem, of course, is that other times of no mass extinction also mark the times of environmental change, and it is fair to say that we could not easily predict all mass extinctions with nonfossil data alone. If environmental forcing, which transcends the abilities of species to survive or adapt, is a major cause of mass extinction, what are the factors? We can list them but finding smoking guns is often another matter.

1. Impact or a series of impacts of extraterrestrially derived objects.
2. Volcanism.
3. Climate change.
4. Lowering of sea level, which reduces available habitats for marine species.
5. Anoxia, especially transgressive spread of deep-anoxic waters onto the continental shelves.
6. Methane hydrate release, resulting in extreme global warming.

These causes, stem more from associations in time between inferred geological events and extinctions, and not from a solid model linking environmental change to extinction. The best example of the latter is the Permian mass extinction. The vast marine regression may have been the driving force behind a variety of environmental changes, including a rise in carbon dioxide, which led to increased temperature and oceanic anoxia. At the end of the Permian, sea level dropped, perhaps about 200 m, which was followed by a transgressive rise of sea level in the Lower Triassic of similar magnitude in just 2 My. Seasonality and reduction of habitat complexity during the regression may also have begotten environmental instability, beyond the adaptive ranges of a number of specialized groups. Robert Berner produced a solid model that demonstrated a remarkable drop in atmospheric oxygen from the end of the Permian to the beginning of the Triassic. The drop may have been stimulated by a period of extensive volcanism, which in turn caused dry climates and the wide-spread drying of the planet, which reduced burial of carbon in swamps and released carbon dioxide to the atmosphere. This might have caused extensive warming and temperature stress. The reduction of oxygen might have been the trigger for extinction both on land and sea. If oxygen in the ocean declined, hydrogen sulfide might have appeared, which would be poisonous to most marine life. Volcanism might be a minor contribution to climate change at the end of the Permian, because calculations preclude much of a change in the large ^{13}C deviations at this time, due to outgassing. However, the extensive volcanism in Siberia might have risen to the surface and heated up carbonates and coal deposits, liberating lethal methane, which might have triggered extinctions and caused larger ^{13}C deviations. The Siberian traps cover an enormous area of about two million square kilometers. Paleontologists Norman Newell and Anthony Hallam have implicated sea-level change in a number of extinctions throughout the Mesozoic, but they are also often combined with other events, such as bolide impacts, anoxia, and temperature change. Douglas Erwin likened this multicomponent explanation to Murder in the Orient Express by Agatha Christie, where twelve culprits were ultimately found to have conspired to murder the victim. Great for murder mysteries but maddening for science. Even this cast of characters ignores the hypothesis of global cooling triggered by glaciation, but this may be discounted as the glacial evidence can be dated much before the extinction begins.

The Pace of Mass Extinctions

The end of the Cretaceous is not the most dramatic mass extinction in the Phanerozoic (**Figures 1** and **2**). At the time, however, both major terrestrial and marine elements were lost, the fauna was sufficiently modern to be understood ecologically, and some of our favorites, such as dinosaurs and ammonites bit the dust. Luis, Walter Alvarez, and colleagues set off a debate that has yet to flag by suggesting that a massive asteroid impact caused the extinctions by blanketing the earth with dust spread along ballistic trajectories outside the atmosphere. Such catastrophes had been suggested before by paleontologists, but here was the first tangible evidence.

The fact of an End-Cretaceous impact is supported by a world-wide anomaly of high concentrations of the element iridium in rocks just at the End-Cretaceous boundary (K-T boundary). Although there still is some controversy about this, extraordinarily high iridium concentrations indicate an extraterrestrial origin for some of the materials in the rock. Shock structures on quartz crystals suggest that an enormous crater should be present on a continent. A possible piece of ejecta has been found in a core of the Pacific at the K-T boundary, which indicates that the bolide was likely a typical metal- and sulphide-rich carbonaceous chondrite rather than derived from cometary materials.

The site of impact has probably been located in the megacrater at Chixulub in the Yucatan of Mexico. The crater harbors an armory of smoking guns, including shocked breccia clasts similar to shocked rock fragments found world-wide, tektite-like glasses, a pronounced iridium anomaly, and a radiometrically estimated geological age of 65.2 Ma, which matches the ages of world-wide K-T boundary samples with tektites. The crater suggests a bolide of some 10 km in diameter. If the impact were at an angle, presumably more material would be spattered into the atmosphere, but it is clear from the world-wide iridium anomaly that winds could have spread the calamity throughout the earth.

We at present can only speculate the possible biological consequences. The dust cloud would exist for a time sufficient to severely disrupt climate by cutting off all light, and temperature might have been expected to drop precipitously. A stable oxygen isotope anomaly at the boundary gives evidence for a sudden temperature change. The impact should therefore

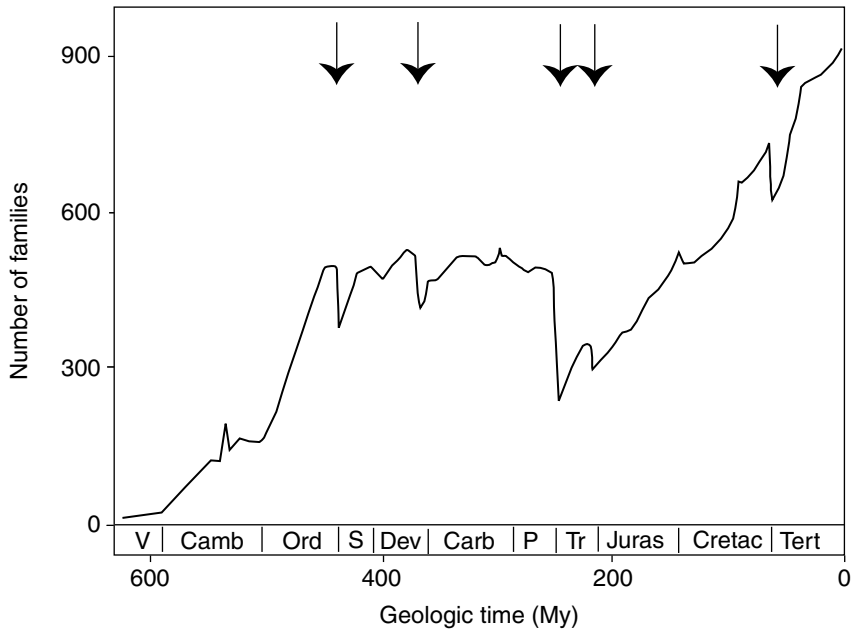


Figure 1 Change in the number of marine and terrestrial taxonomic families throughout Phanerozoic time. Arrows point to times of mass extinction. Reproduced from Sepkoski Jr. JJ (1984) A kinetic model of Phanerozoic taxonomic diversity III. Post-Paleozoic families and mass extinctions. *Paleobiology* 10: 246–267. This estimate of standing diversity is biased with regard to variations in intensity of sampling, publications, and representation of specific time periods in terms of rock volume.

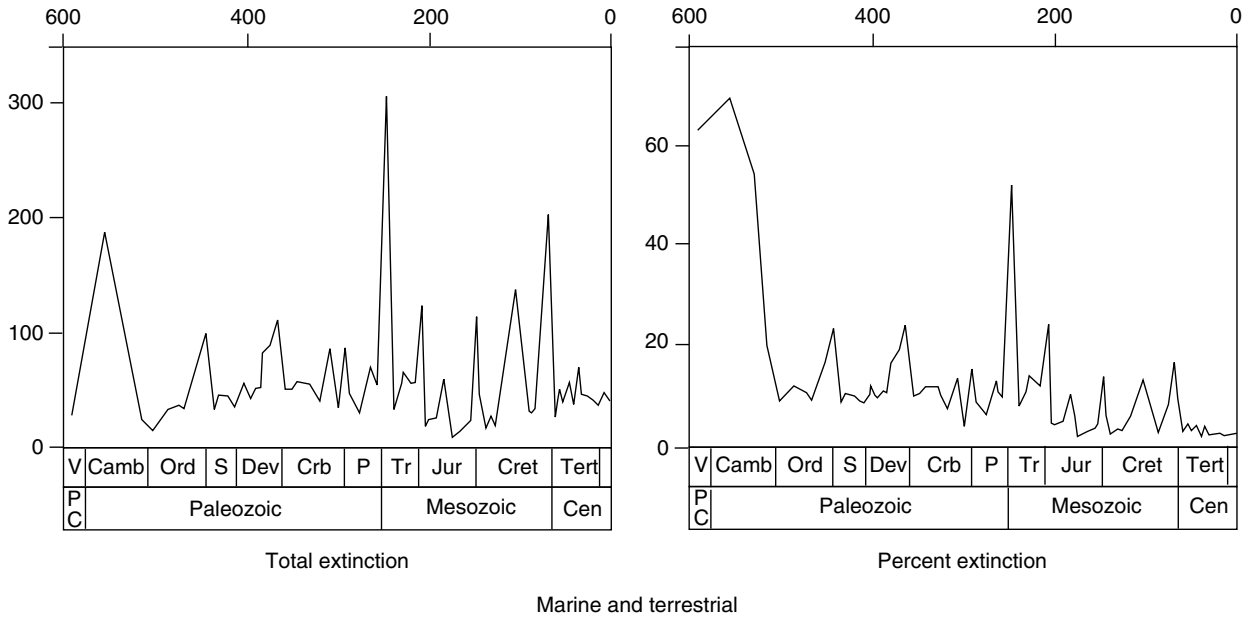


Figure 2 A more recent analysis of extinction rate of combined marine and fossil taxonomic families, based on recent compilations of a broad range of specialists. Reproduced from Benton MJ (1995) Diversity and extinction in the history of life. *Science* 268: 252–258.

have affected all organisms dependent on light and warm temperature. Deep-water forms not so dependent on light or warm temperature, such as nuculid bivalves, would be expected to survive. Alternatively, a hot plume emanating from the impact site could have accelerated the production of nitric acid, causing a world-wide shower of acid rain that might have poisoned the upper ocean.

Can extraterrestrial impacts be used to explain other major extinctions? The results are mixed. Positive and negative evidence for an iridium anomaly has been found for the Frasnian-Famennian (Late Devonian) mass extinction, but George McGhee and colleagues found that the extinction itself was spread over at least 7 My and climatic effects are evident. Solid evidence for impacts unfortunately postdate the

Frasnian-Famennian event. There seems to be no iridium anomaly associated with the terminal Ordovician extinction. A similar iridium anomaly in sediments of 34 My of age occurs simultaneously with the disappearance of five dominant radiolarian species, and at the general time of a mammalian extinction, but the larger picture of biotic change across this boundary is gradual, with no suggestion of a catastrophe. We are therefore left with the End-Cretaceous extinction to consider.

The Alvarez theory has one strong and other weaker predictions. (1) Extinctions must follow or coincide with the impact. (2) One might also expect many groups to die off instantaneously, but a less catastrophic change in temperature and light might have an effect that would be prolonged. (3) Finally, groups more prone to light stress or temperature increase would be more vulnerable (e.g., phytoplankton vs. deep-water deposit-feeding benthos), because an impact might likely spew dust into the atmosphere, lowering world temperature. The response of the sensitive groups should be geologically instantaneous.

As in most other mass extinctions, the end of the Cretaceous was preferential as to organisms affected. Groups associated strongly (Foraminifera, coccolithophorids) or weakly (ammonites) with the water column suffered most strongly, whereas benthic forms (e.g., bivalve mollusks) generally suffered less. Members of food webs less dependent on plant material (marine deposit feeders, scavengers, stream inhabitants, and small insectivorous mammals) suffered less than the strict herbivores. The relative success of sediment-feeding invertebrates relative to suspension feeders may be due to their occurrence in deeper waters.

There is also an apparent thermal bias in extinction at the K-T boundary. S. M. Stanley found that mollusks and Foraminifera in the tropical Tethyan sea suffered large-scale extinction, and were replaced by higher-latitude contemporaries. The question of timing is more confusing. Coccolithophores and non-globigerinoid Foraminifera disappeared so precipitously (and simultaneously with the iridium anomaly) that chalks give way to clastic sediments in a knife edge contact in several geological sections. In the chalk of Denmark, the Maastrichtian fauna, dominated by brachiopods, disappeared abruptly, with no prior warning in terms of reduced diversity, or early extinction of specialized forms. The sediments above the chalk are clayey and indicative of anoxic conditions. They also have a spike of iridium. Turbidity, loss of an appropriate sediment, and anoxia may all have contributed to the abrupt extinction. Radiolitid and hippuritid rudists bit the dust during a period of flourishing radiation. The current evidence suggests that Cretaceous vertebrates also bit the dust at the boundary, but preservation is probably too spotty to tell whether it is sudden in any respect.

Unfortunately, the larger story is not nearly so simple. Some fossil groups, including land plants, inoceramid bivalves, and ammonites experienced major extinctions or reduced speciation rates several million years before the impact occurred. The rudistids, cone-shaped bivalves that often formed Cretaceous reefs, died relatively suddenly, but well before the very end of the Cretaceous. Most embarrassing for the impact theory, the fresh-water biotas seem to have emerged unscathed. One would have thought that organisms such as turtles and crocodiles would be most vulnerable to a major extraterrestrial impact.

The dinosaurs represent an interesting case. The dinosaur fauna of the Late Maastrichtian included fewer than 20 species in 15 genera and 10 families, chiefly in the North American Western Interior. There is no good evidence, however, that the dinosaurs were declining steadily toward this low number in the last 9 My of the Cretaceous. It may well be that, although the dinosaur fauna was a remnant at the end of the Cretaceous, their demise was nevertheless caused by the impact.

The biogeography of extinction in the End-Cretaceous provides some insight. Maastrichtian planktonic foraminifera disappeared suddenly at the K-T boundary in middle and low latitudes. At high latitudes, however, a number of groups survived unscathed into the Danian, the beginning of the Paleocene epoch. The rudistid bivalves, associated largely with tropical and sub-tropical waters became extinct toward the end of the Cretaceous, but otherwise there are no differences with latitude in bivalve mollusk extinctions.

The sharpness of the boundary for any group is clouded by the imperfections of preservation. If the ranges of certain groups fail to extend all the way to the K-T boundary, then it is possible that perfect preservation would have given us a far different picture. Many western Tethyan ammonites appear to become extinct below the boundary but Charles Marshall and Peter Ward demonstrated that the confidence limits of a number of lineages allows the possibilities that poor preservation is the reason why some fossil ranges fail to continue right to the K-T boundary. Of course, this does not necessarily prove that the species became extinct at this time, they could have become extinct before or after. The analysis does exclude being sure that the geological range should be read literally. N. MacLeod performed a similar analysis with Upper Cretaceous foraminifera and found out that the ranges are compatible with a sudden extinction. They also, however, are compatible with many other possible scenarios. Some groups penetrate the boundary and it appears that there is good evidence that the fossils were not reworked up into the Danian by erosion and bioturbation.

In a mass extinction, we would love to have a single cause to explain extinctions, but, like the Permian, there may be several interacting and a succession of climatic changes that caused a range of extinctions.

1. In the Late Maastrichtian, sea level decreased by 150–200 m, making a hypothesis of increased terrestrial seasonality compatible with the ultimate disappearance. Just below the K-T boundary, sea level suddenly rises.
2. Temperature dropped.
3. The Deccan volcanics in India, enormous in scope, probably spewed a variety of substances that strongly affected the atmosphere and might produce effects resembling those of a bolide impact, but the activity probably predated the K-T boundary.

The evidence suggests that there may be more than one process at work in causing major extinctions. The evidence from land plants, dinosaurs, and many mollusks points to a change in conditions well before the Cretaceous-Tertiary boundary. Most notable are groups such as inoceramid bivalves and some of the ammonites that seem to be tracking an environmental deterioration before the termination of the Maastrichtian, or final Cretaceous stage. This would violate the

fundamental prediction that extraterrestrial events should be followed by and not preceded by, extinctions. But the iridium layer, plus its associated faunal disappearances, cannot be reconciled with any hypothesis of gradual climatic deterioration. There clearly was a sudden extinction at the K-T boundary, but it was focused on lower latitude microplankton; we can only speculate about suddenness of extinction in other groups. We are left with a compound hypothesis, at least for the proximate cause of the totality of End-Cretaceous extinctions. It has been suggested that the majority of strong extinction events in the Mesozoic and Cenozoic were at times when the major episodes of volcanism coincided with a bolide impact.

Complexity also characterizes the Permian extinction. It makes sense that sea-level drop was an important factor, operating by the effect of reduced habitable area and reduced environmental heterogeneity, but we have to reconcile the precipitous change in sea-level coverage in the last stage of the Permian, with the pattern of extinction, which was initiated earlier than the sea-level drop. Also, the extinction seemed to concentrate on certain ecological groups, particularly tropical forms at the end, but it affected high-latitude groups earlier in the Permian. Most discouraging of all, what if the extraterrestrial influence occurred in a series of impacts, as has been argued to be possible in the Frasnian-Famennian (End-Devonian) extinction, rather than as one big bang? Without a series of signals (e.g., a series of definitive iridium spikes of extraterrestrial origin), such a hypothesis is very speculative.

In conclusion, there is credible evidence for the role of an extraterrestrial object in causing extinction in the End-Cretaceous. Although other extinction factors might have been at work in the Late Cretaceous, extinctions occurring before the time corresponding to an extraterrestrial iridium anomaly do not falsify the impact hypothesis. At any time in the record, *some* groups must be declining. Even if we accept the impact as a source of some extinction, the Cretaceous still appears to be a compound event, as environmental change and the pace of extinction both accelerate before the K-T boundary. Isn't it bad luck that a large asteroid happened to smack into the earth just as sea level was changing as much as it did in the whole of the Phanerozoic? Several other mass extinctions also seem to be complicated, and were probably associated with changes in climate and sea level (that are probably interrelated). Anoxia may also be an important cause of mass extinctions, as witnessed by the extensive development of black shales in certain periods. What is lacking at present is a credible evaluation of the relative effects of these factors.

Long-Term Consequences of Mass Extinctions

In one obvious respect, mass extinctions had long-term consequences. Major groups of organisms were lost, meaning that ecosystems would never be the same. At the end of the Paleozoic, we lost trilobites, rugose and tabulate corals, orthid and productid brachiopods, and other groups, which changed the character of marine communities forever. Of course, we lost the ammonites and dinosaurs at the end of the Mesozoic. As mentioned above, the loss of relatively simple Paleozoic bottom communities gave way to Mesozoic communities which were dominated more by burrowers and effective predators.

After major mass extinctions, paleontologists have noticed a brief period of acceleration of diversification, as if ecological space was being filled with diverse groups that replaced recent extinctions. Still, local diversity after the Permian extinction was often sparse. Andrew Krug and colleagues demonstrated that the End-Cretaceous extinction was followed by a brief acceleration of diversification, which was in turn followed by an uptick in origination rates of bivalve mollusk groups that persists to the present day. The explanation for this increase is still obscure but might relate to further increases of appearances of predators during the Tertiary period.

Biogeographic Aspects of Extinction

Some of the Phanerozoic changes in taxon richness can be related to a large degree to changes in the degree of provinciality. If we take the species–area effect as given and constant, it is easy to calculate that, as one large province is divided into two, reductions of the area are more than compensated by the increased total species richness as long as the dispersal between areas is limited. Thus a temporal increase in provinciality results in an increase of total biodiversity and a reduction results in a decrease. Major world-wide deterioration of climate (e.g., world-wide cooling, increase of seasonality) might be an example of a time when provinciality might decrease.

The Silurian period, for example, was one of extreme cosmopolitanism, with one province of approximately 90 articulate brachiopod genera in the North Silurian Realm. In the Ludlow (Upper Silurian), two provinces can be delineated, with about 90 genera in each province. In the Devonian, Arthur Boucot found that the number of provinces increased to six; the total number of articulate brachiopod genera increased to about 350 on average. During the Frasnian (Middle Devonian), this provinciality decreased relatively suddenly, and generic richness returned to 93.

The onset of the Permian extinction was also marked by a decrease in numbers of provinces, and the Early Triassic marks a nadir of provinciality in the Phanerozoic. During the end of the Paleozoic, geographically restricted bivalve genera succumbed before more wide-spread genera, suggesting that the overall environmental change was filtering out those forms that define provinciality in the first place. Norman Newell argued that the extinction was related to the major fall in sea level. Shallow marine seas were reduced from a coverage of 40 percent of their possible extent in the Early Permian to less than 15 percent in the latest Permian and then expanded to 34 percent in the Early Triassic. James Valentine and Eldredge Moores speculated that reduced rates of sea-floor spreading may have been responsible for a lowering of ridge activity, depression of deep-sea bottoms, and the consequential large-scale marine regression. The significant reduction in area, coupled with continental assembly of Pangaea at the end of the Permian may have increased extinction rates, and would have homogenized the fauna due to the possible presence of more inter-shelf dispersal possibilities. In contrast, the Pleistocene reduction of area covered by the sea was far lower and on the basis of area alone; the modest marine extinctions

are therefore not surprising from this point of view. Area reduction itself might not be a potent agent of extinction. Sea-level drops would hardly affect the shallow-water habitat distribution of oceanic islands, where most modern families are widely distributed. Sea-level drop may just be a correlate of another change.

The changing spatial relationships generated by continental drift and sea level fluctuations must have had important influences on climate. James Valentine's theory of climate change generated by continental assembly and fragmentation attempted to relate climate and sea level to sea-floor spreading. Periods of continental assembly were envisioned as times when interior continental climates were severe, affecting the continental shelf faunas. In contrast, times of fragmentation were times when the continents' climate was more moderate due to ameliorating marine conditions; this permitted the buildup of shallow-water diversity. Although the post-Permian expansion may fit this pattern, evidence from the Paleozoic does not seem to show an increase in continental fragmentation during the Early-mid-Paleozoic. Indeed, the continents were maximally fragmented and arrayed along the equator during the Cambrian. Continental drift and arrangement nevertheless have had profound effects on climate and probably extinction. During the Ordovician and Silurian Periods, Gondwana drifted southward from its Cambrian position at the equator, and came to rest on the geographic south pole. This coincides with the Late Ordovician glacial tillites that have been found in North Africa, and a large reduction in the degree of marine provinciality relative to the Early Ordovician. In the Cenozoic, the spatial arrangements of the continents about the Pacific and Atlantic Ocean made for a quite different climatic history. The North Atlantic was a more enclosed basin and was far more severely affected by the Late Cenozoic polar cooling. The Pleistocene initiated severe enough climates to cause a major molluskan extinction in the southeastern US Shelf, whereas Pacific American faunas showed no increased extinction.

The effects of increasing access between biogeographic realms can be illustrated by the large-scale interchange of mammals between North and South America after the Pliocene establishment of the Isthmus of Panama, following the disappearance of the Bolivar Trough marine barrier. Before the interchange, there was long-term stability in a number of mammalian families. As a probable result of North America's initial higher taxon richness, more taxa moved from north to south than in the reverse direction. In South America, where taxon richness now exceeded previous "steady state" levels by more than 50 percent, there was about a 70 percent increase in extinction rates. Descendants of the North American invaders participated in an evolutionary radiation, resulting ultimately in an overall richness higher than previous levels. Mammalian diversity is now higher in South America, in contrast to the situation previous to the exchange. This suggests that area does have an effect on regulating diversity, but evolutionary changes can impose a significant overprint on diversity.

Periodicity in Extinction, or Just Ups and Downs?

Periodicity of extinction or climatic change predicted by astronomical or geophysical theories would be the most convincing way to establish a terrestrial or extraterrestrial cause of extinction. If extinctions are measurably periodic, it may be that only one credible cyclic theory would fit in the available pattern. The precedent for such an approach lies with the longstanding theories of the periodicity of Pleistocene glaciations. The Yugoslav astronomer Milankovitch theorized that Pleistocene glacial advances and retreats might be regulated by changes in high-latitude insolation, caused by cyclic changes in the earth's orbital eccentricity, tilt, and time of perihelion. A power spectrum analysis of temporal changes of abundance of Pleistocene planktonic fossils in oceanic cores corresponded well to climate changes estimated by stable oxygen isotopes and to periodicity peaks predicted by the Milankovitch theory.

A number of studies in recent years have taken up this theme and related these cycles to sedimentary cycles, including some of the classic mid-continent alternations of carbonate and mudstone. Many of these cycles occurred during times when there was no significant amount of continental glaciation, and represent transgressive–regressive cycles (rises and falls of sea level). For example, sedimentary cycles in the lacustrine Early Mesozoic supergroup correspond to periodicities of approximately 25,000; 44,000; 100,000; 130,000; and 400,000 years. These periodicities, in turn, correspond to those expected from celestial processes, such as the precession of the equinoxes, the obliquity cycle, and the eccentricity cycle. Cyclic processes such as the precession of the equinoxes may have driven continental heating cycles that rearranged wind and climate.

Milankovitch climatic rhythms also appear in Mid-Cretaceous black shale sedimentary cycles. These cycles consist of alternations of carbonate and shale, with intervals of highly oxidized (red) and highly reduced (black) strata. They are particularly interesting, as they occur in marine sequences and must have reflected periods of ocean bottom anoxia, alternating with vigorous bottom mixing and high productivity in the water column. On an even smaller scale, El Niño-La Niña cycles and the North Atlantic Oscillation whose forcing mechanisms of periodicity are not well understood, are known to cause cycles of benthic abundance in coastal communities and in small bays and fjords, such as the Swedish Gullmar Fjord.

The earth's history has been dominated by large-scale changes in climate, arrangement of continents, volcanism, and sea level. Alfred G. Fischer developed a theory connecting physical conditions with the overall pattern of Phanerozoic life. Global sea level was relatively high in both the Mid-Paleozoic and Mesozoic. Periods of continental breakup, when dispersed and thinner continents resulted in smaller ocean basins, would be associated with higher sea levels. Periods of continental aggregation, when continental crust was bunched up due to collisions and ocean basins were therefore more commodious, which resulted in lower stands of sea level. The temporal variation in granite emplacement matches the sea-level curve. This suggests a causal link between active continental fragmentation, volcanism, and sea level, an environmental condition of obvious importance to the world marine biota.

Fischer, a pioneer in global climate thinking with regard to the evolution and extinction of life, speculated on the presence

of a causal connection between changes in terrestrial vulcanism and global climate through the greenhouse effect. Increased volcanism may have liberated carbon dioxide into the atmosphere. As these periods were of higher sea-level stand, erosion would have been minimal, and loss of CO_2 in weathering would be suppressed. During times of low sea level, low volcanism would reduce the liberation of CO_2, and increased weathering would consume CO_2. Thus, the Mid-Paleozoic amelioration was associated with high CO_2, which, in turn, caused greenhouse effect and an increase of surface temperature. The end of the Paleozoic witnessed the termination of such conditions, and an "icehouse effect" resulted in a deterioration of climate mainly at high latitudes. It is not clear whether these cited fluctuations are irregular temporal changes or regular oscillations.

In time of extraordinary paleontological excitement, that, regrettably, has passed, David M. Raup and John J. Sepkoski reported a periodicity of about 26 My in the occurrence of extinction peaks of taxonomic families. Analyses of genera produced an even better periodic signal. To consider an extinction important, Raup and Sepkoski used a threshold level of 2% minor variations on this criterion change the periodicity to an average time of as much as 30 My between peaks, but the time between specific peaks varies substantially.

A number of celestial cycles have been suggested to explain the cyclicity. Of course, the most interesting ones are those that would cause rains of extraterrestrial objects on the earth or major changes in climate. So far, no theory works very well and it has also been suggested that a number of random models can explain the presence of cycles. This issue has not been settled yet and if the cycles are real, there is great hope that they can be related to an extraterrestrial source. Right now, the variation in extent of extinction and the average time between extinctions are not definitive enough to corroborate any models, which usually involve impacts of extraterrestrial objects such as asteroids or comets. An independent survey of extinctions supervised by Michael Benton failed to corroborate the presence of extinction periodicity. Alroy's more recent analysis also fails to find such periodicity.

Background Extinction and Turnover

Normal Extinction?

The great spans of geological time between mass extinctions also witnessed significant appearances and extinctions, but at lower frequencies. The temporal pattern of such extinctions is not clear. Arthur Boucot suggested that periods of several million years are often dominated by a set of ecologically distinct species, whose coordinated extinction might be followed by the invasion or evolution of a new set of ecologically similar forms. In recent years, a number of studies have shown that turnover (including extinction) is common at this temporal scale, but the exact pattern is variable. In some cases, periods of relatively low extinction are punctuated by high rates of a wide range of distantly related taxa, which are correlated with local environmental change such as basinal sea-level change. Other studies, however, show rather high extinction and appearance rates with no punctuations in extinction. This is a field that needs to be explored with far more data collection and studies of environmental reconstruction before we find any pattern.

At this smaller scale of extinction, some of the same mechanisms as mass extinction may be in effect. Sea level rise and fall, climatic change, and other factors can operate on a smaller scale to cause extinction. At this scale, however, biological factors may be of great importance. Biological factors in extinction might include

1. Competitive displacement by an invader;
2. elimination by an overwhelming predator or herbivore;
3. spread of disease.

Some of these factors may have been involved in a worldwide extinction of the so-called megafauna, a group of mammals and large flightless birds that disappeared at the end of the last glacial advance. The extinction involved large marsupial mammals and large flightless birds, the renowned Irish Elk, large elephantine forms, saber-toothed marsupial, and placental cats, among others. It is possible that human hunting is the cause of these extinctions, but mobile human populations may also have brought novel diseases as they spread around the planet.

Declining Extinction Rates

Sepkoski's extraction from the fossil database of statistical entities known as evolutionary faunas produced the fascinating result that the so-called evolutionary faunas (EF) are less and less prone to extinction as we approach the present. In the Ashgillean and Frasnian extinctions, for example, the more ancient Cambrian EF suffers more than the Paleozoic EF. In the Permian and Norian extinctions, the relatively older Paleozoic EF suffers more than the Modern EF. The successive evolutionary faunas also have progressively lower turnover (appearance plus extinction), which may make for increasing stability.

A class of distinctly lower family-level extinction rates decline during a long Paleozoic period of fairly constant taxon richness. To keep a steady state, a decline in extinction must be matched by an overall decline of originations, which has also been found. Why should family-level extinction and origination rates decline over geological time? There is no simple solution, although it is tempting to believe that taxa over time have evolved more and more resistance to extinction including reduced competition with other groups. Given the vagaries of extinction and the fact that extinction is usually an overwhelming process, driven by habitat loss, wide-spread marine anoxia, and other factors, this idea appears to be far fetched.

There may be an explanation to declining extinction that is a bit more mundane. Extinction might have declined as a result of the ratio of species numbers per family, which has been increasing steadily since the Mesozoic. If a family's representation in the world biota increases in number of species and its consequent ecological and geographic coverage, then the probability of family-level extinction may decline.

This explanation still does not provide a satisfactory answer to the decline in originations, which also decline over the same time period. Such a decline implies a long-term

reduction in the production of novelties sufficient to define taxonomic families. In other words, the rate of origin of morphological diversity has decelerated over time. Two concomitant processes may have contributed to this decline in origins of basic morphological diversity. A general filling in of resource space may have made it difficult for wholly new forms to take root later and spread. Our world may very well be the tangled bank conceived by Darwin.

Extinction and its Relation to Diversity

One of the most fascinating questions in ecology is the limit to which species numbers interact with the appearance of new competitors. The ecologist G. Evelyn Hutchinson first posed this problem in his landmark paper "Homage to Santa Rosalia or Why are There So Many Kinds of Animals?" Do all of the species that currently exist fill up available ecological space, creating a maximum of biological diversity? If so, what would happen if a biological crisis caused an extinction? Would diversity increase and fill in the gaps of ecological exploitation of resources generated by the extinction? This problem would arise for both mass extinctions and background extinctions. A very interesting observation by Arnold Miller and John Sepkoski showed that the number of genera of bivalve mollusks were increasing steadily from the Mid-Paleozoic through the Tertiary. But there were conspicuous times when diversity dropped only to seemingly recover to match the ongoing trend. This may not support the idea that there is really a limit to the number of bivalve species that can be supported, because they were steadily increasing, but it does suggest that diversification is dependent on extinction and accelerated following occasional major drops. In a study of the relationship of diversity levels to rates of diversification during several mass extinctions, Michael Foote has shown that the rate of diversification increases at lower levels of diversity, which suggests that the diversification of life is dependent on standing diversity as Hutchinson would have argued. But Foote could find no obvious pattern relating extinction rate to diversification, despite other previous contrary results found in another study. This subject requires much further study. Alroy's general analysis suggests that extinction might play a role in diversification. (**Figures 3** and **4**).

To summarize, extinction in the fossil record reveals the following main points.

1. Mass extinctions caused world-wide precipitous loss of species, over a wide variety of taxonomic groups and habitats.
2. Environmental change in mass extinctions overwhelm the ability of a species to survive by profound changes in the environment over the whole species range.

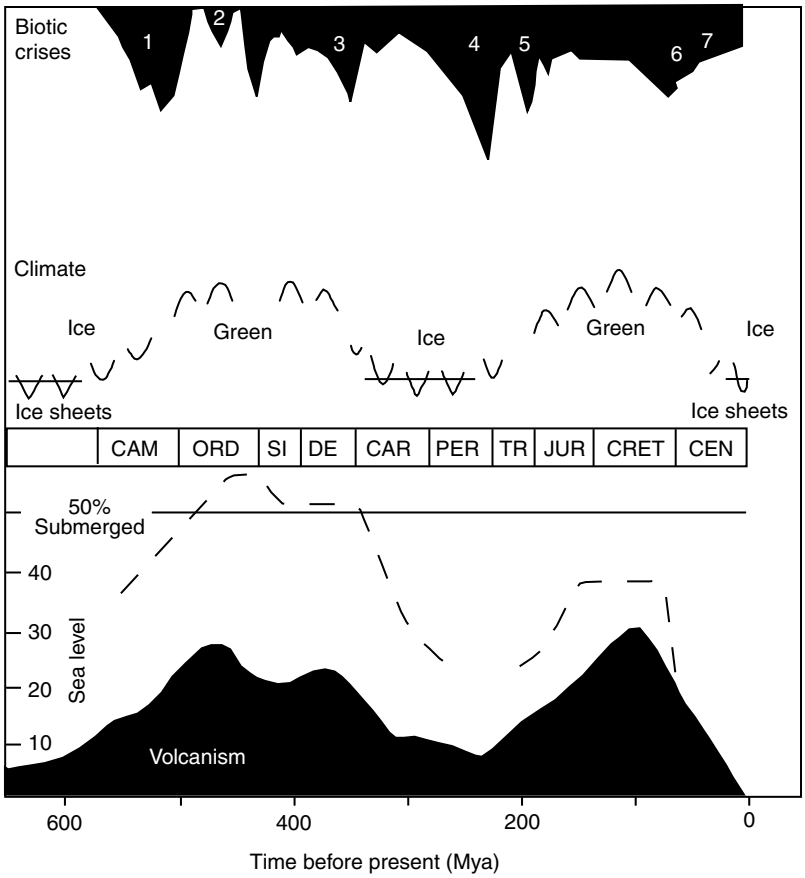


Figure 3 "Supercycles" of the Phanerozoic, postulated by Fischer. Sea-level curve is superposed on a diagram of granite emplacement, times of glaciation, times of biotic crises (numbered) as determined by N. D. Newell, and a general estimate of climate, characterized by either icehouse (***I***) or greenhouse (***G***) conditions. (modified from Fischer, 1984).

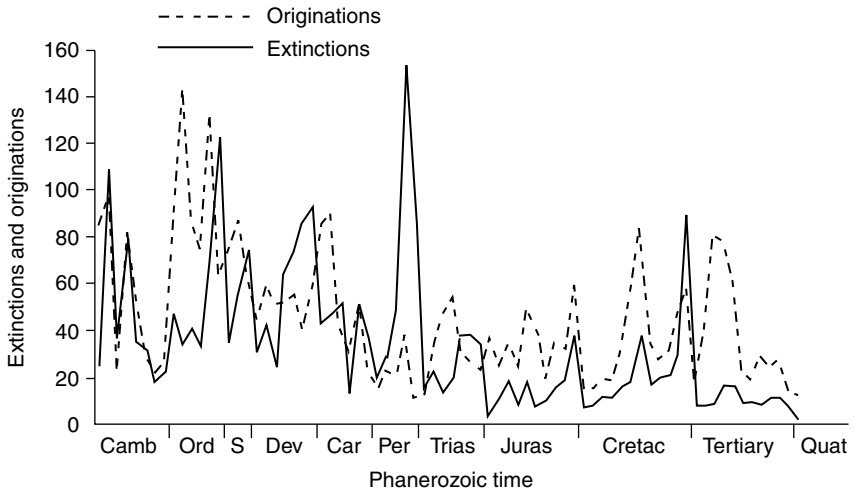


Figure 4 Total numbers of marine animal family originations and extinctions per geologic stage (Reproduced from Hoffman A and Ghiold J (1985) Randomness in the pattern of 'mass extinctions', and 'waves of origination'. *Geological Magazine* 122: 1–4.).

3. The major mass extinctions eliminated a very large majority of the species on the planet and caused major reorganization of the world's biota.
4. In many cases, extinction appears to be selective with regard to ecological characters of the groups that survive, but in others one cannot identify any traits that make a group more resistant to extinction than others.
5. A distinctly lower level of extinction can be extracted from an analysis of the fossil data. This type of extinction might be a smaller version of mass extinctions, with relatively low rates of extinction punctuated by major regional environmental change driven by sea-level changes. Lower-level extinction rates appear to decline toward the present day. The explanation for this is unclear, but it may just be an artifact of the way taxonomic groups are classified into species. More recent taxonomic families have more species and more species-level extinctions would therefore be required to see the loss of a fossil family that occurred closer to the present day.

References

Alroy J (2008) Dynamics of origination and extinction in the marine fossil record. *Proceedings of the National Academy of Science, USA* 105: 11536–11542.

Benton MJ (1995) Diversity and extinction in the history of life. *Science* 268: 252–258.

Erwin DH (1993) *The Great Paleozoic Crisis*. New York: Columbia University Press.

Erwin DH (2006) *Extinction: How Life on Earth Nearly Ended 250 Million Years Ago*. Princeton: Princeton University Press.

Foote M (2010) The geological history of biodiversity. In: Bell MA, Futuyma DJ, Eanes W, *et al*. (eds.) *Evolution Since Darwin: The First 150 Years*, pp. 479–510. Sunderland, MA: Sinauer Associates.

Hallam A and Wignall PB (1997) *Mass Extinctions and Their Aftermath*. Oxford: Oxford University Press.

Hutchinson GE (1959) Homage to Santa Rosalia or Why are there so many kinds of animals? *American Naturalist* 93: 145–159.

Jablonski D (1994) Extinctions in the fossil record. *Philosophical Transactions of the Royal Society, Series B* 344: 11–17.

Krug AZ, Jablonski D, and Valentine JW (2009) Signature of the End-Cretaceous mass extinction in the modern biota. *Science* 323: 767–771.

Levinton JS (2000) Genetics. *Paleontology, and Macroevolution*, 2nd ed. New York: Cambridge University Press.

McGhee Jr. GR (1996) *The Late Devonian Mass Extinction*. New York: Columbia University Press.

Purvis A (2008) Phylogenetic approaches to the study of extinction. *Annual Review of Ecology, Evolution, and Systematics* 39: 301–319.

Raup DM (1986) *The Nemesis Affair*. New York: WW Norton.

Raup DM (1991) *Extinction: Bad Genes or Bad Luck?* New York: WW Norton.

Wagner PJ, Kosnik MS, and Lidgard S (2006) Abundance distributions imply elevated complexity of Post-Paleozoic marine ecosystems. *Science* 314: 1289–1292.

F

Feeding the World and Protecting Biodiversity

Paul C West, University of Minnesota, St. Paul, MN, USA
Reinette Biggs, Stockholm University, Stockholm, Sweden
Bruce A McKenney, The Nature Conservancy, Charlottesville, VA, USA
Chad Monfreda, Arizona State University, Tempe, AZ, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ecosystem services Benefits that nature provides people, such as food, water purification, climate regulation, pollination, waste decomposition, and spiritual values.

Invasive species Plant or animal species that disrupt an ecosystem's composition, structure, function, or the processes that maintain it.

Regime shift A large, persistent, often abrupt change in the structure and function of an ecosystem associated with a change in dominant feedback processes.

Scenarios Alternative outcomes developed to identify and manage uncertainty in complex problems.

Tipping point A threshold separating two alternate regimes or configurations in the structure and function of an ecosystem (also see regime shift).

Agriculture is the biggest driver of land-use change on the planet. The strain of agriculture on biodiversity is only increasing as demand for food and other agricultural goods continue to increase. How will we manage to feed a future world and protect biodiversity? How much can changes in management practices alone meet future food demands while decreasing the environmental costs? What new developments are needed? And what are the mechanisms for protecting biodiversity on a planet that is increasingly managed more intensively to sustain and improve human livelihoods? This article summarizes the scope of agriculture, its effect on biodiversity, and strategies for feeding the world while maintaining biodiversity.

The Scope of Global Agriculture

Nearly 40% of the ice-free land on earth has been converted to pastures or croplands to feed people (Foley *et al.*, 2005; Ramankutty *et al.*, 2008; **Figure 1**). Pastures and croplands cover ~28 million km^2 and ~15 million km^2, respectively (Ramankutty *et al.*, 2008). To put these numbers in perspective, the areas of pasture and croplands are nearly equivalent to Africa and South America, respectively. Croplands cover 20% of temperate regions, compared with ~10% of tropical regions (West *et al.*, 2010). Wheat, rice, and maize are the most dominant crops, covering over a third of the total crop area.

Several projections estimate that agricultural production will likely need to double (or more) by 2050 to meet the growing demands for agricultural production. According to the United Nation's Food and Agriculture Organization, agricultural lands were expanding at a rate of ~10 million ha each year between 1980 and 2007. This expansion is driven by three main factors. First, food demand is increasing to feed a growing population. Second, more people are eating meat as a substantial part of their diet as wealth and preferences change. Depending on the animal species, it takes 2–30 cal of feed to produce a calorie of meat, reducing the efficiency of the number of calories "delivered" from each area of cropland. Finally, many countries are increasing crop production for biofuels as a means of reducing greenhouse gas emissions. Recent concerns over the trade-off of food security and energy make the rate of increase in biofuel production uncertain.

In addition to the land footprint, a tremendous amount of water is used to produce food. During the past 50 years, the area of irrigated croplands has doubled and accounts for 70% of total water withdrawals (Gleick, 2003). Much of this water is "consumed," meaning that it either evaporates from the ground or is transpired by plants, and is not immediately available in rivers, lakes, or groundwater. Irrigation is most intense in arid regions, which most commonly have stressed water supplies for people and biodiversity as well as crops.

Fertilizer use has also increased dramatically in the past 50 years. The development and commercialization of producing synthetic nitrogen gave farmers easier access to fertilizers to increase crop yields. As a result, yields have doubled in

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00338-5

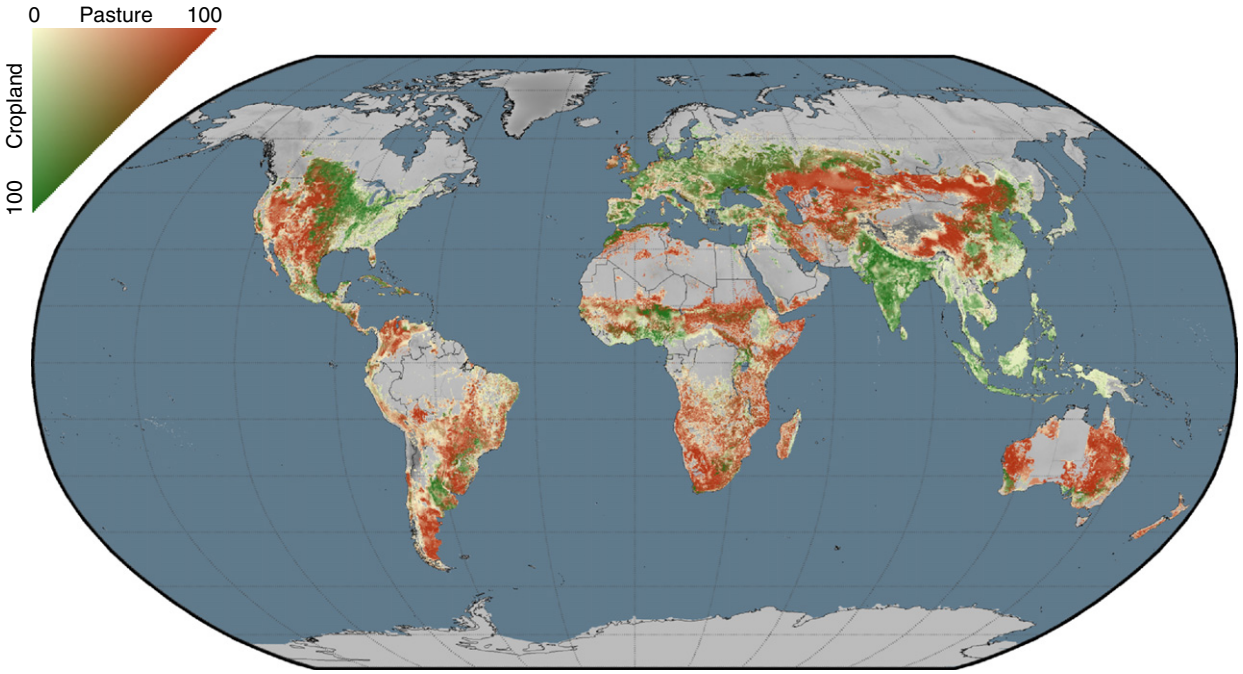


Figure 1 Distribution of croplands and pastures. In 2000, ~15 million ha were used for croplands and ~28 million ha were used for pastures. Data from Ramankutty N, Evan AT, Monfreda C, and Foley JA (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles* 22: GB1003. doi:10.1029/2007GB002952.

many places. Unfortunately, these nutrient additions have not only increased crop yields, but they have also increased the amount of excess nutrients in the environment. The annual flows of nitrogen and phosphorus have doubled their natural levels as a result of fertilizer inputs (Matson *et al.*, 1997; Smil, 2000).

Effects of Agriculture on Biodiversity

Species loss is occurring 100 times faster than the average rate in the fossil record (Dirzo and Raven, 2003). Much of this extinction is attributed to habitat loss from conversion to croplands and pastures. Although it is critical to consider extinction because it is irreversible, land conversion for agriculture has a much broader effect on ecosystem degradation through habitat loss and fragmentation. Populations and ecosystems (as well as other levels of biodiversity) are also altered through indirect effects on climate, water quantity, water quality, and invasive species.

Habitat Loss

Clearly the most direct effect of agriculture on biodiversity is converting natural ecosystems for agriculture. While there are some exceptions, such as planting coffee trees within a selectively cleared forest or grazing on natural grasslands, the species composition of agricultural lands rarely resembles the natural ecosystem that it replaced. However, some agricultural systems, such as pastures and forage crops, provide surrogate habitat for species groups in decline, such as grassland birds.

The distribution of croplands and pastures is a function of climate, technology, crop selection, and cultural traditions. Grasslands have experienced the most conversion, primarily because they exist in climates suitable for many crops, are easily used as pastures, can be converted to croplands much easier than forests, and generally have productive soils. Other ecosystems have also had substantial conversion to croplands and pastures (**Table 1**).

Although extensive areas of many natural ecosystems remain, their habitat quality has been degraded. Agricultural areas have fragmented ecosystems into smaller blocks, which favor species that are tolerant of (or thrive in) disturbed environments at the expense of species restricted to higher quality areas. These simplified ecosystems are less resilient to large disturbance events because fewer components of the former system are able to rebuild, or reorganize, following a major shock.

Climate Change

Clearing land for agriculture has indirect effects on biodiversity through climate change. Deforestation accounts for 12% of global CO_2 emissions (van der Werf *et al.*, 2009). The climate effects of clearing land are particularly strong in the tropics. Nearly all CO_2 emissions from deforestation (98%) originate in tropical regions (DeFries and Rosenzweig, 2010), where agricultural lands are expanding fastest. In addition to the high rate of expansion, tropical ecosystems typically store more than twice as much carbon per hectare as temperate ecosystems (West *et al.*, 2010; **Figure 2**). Unfortunately, these high carbon losses are coupled with loss of habitat in tropical areas that typically have high species diversity and endemism.

Table 1 Percentage of biomes converted for agriculture

Biome	*Potential biome area (millions of ha)*	*Percent converted to pasture (%)*	*Percent converted to cropland (%)*
Tropical evergreen broadleaf forest/woodland	1675.1	10.8	8.5
Tropical deciduous broadleaf forest/woodland	584.0	27.0	23.8
Temperate evergreen forest/woodland	113.1	24.2	18.6
Coniferous forest	361.0	20.0	9.8
Temperate deciduous forest/woodland	482.9	30.0	16.2
Mixed forest	597.6	1.5	1.7
Boreal evergreen forest/woodland	221.0	1.7	1.8
Boreal broadleaf forest/woodland	1494.4	7.7	4.5
Savanna	1912.2	15.7	32.5
Grassland/steppe	1424.9	19.2	49.2
Dense shrubland	597.5	17.9	29.5
Open shrubland	1190.6	7.3	42.3
Tundra	703.9	0.6	12.9
Desert	1530.0	0.9	7.9
Polar desert	124.3	0.1	1.7

Source: Using data on croplands and pasture distribution (Ramankutty *et al.*, 2008) and natural biomes (Ramankutty and Foley, 1999), we calculated the percentage of each natural biome converted to pasture and cropland.

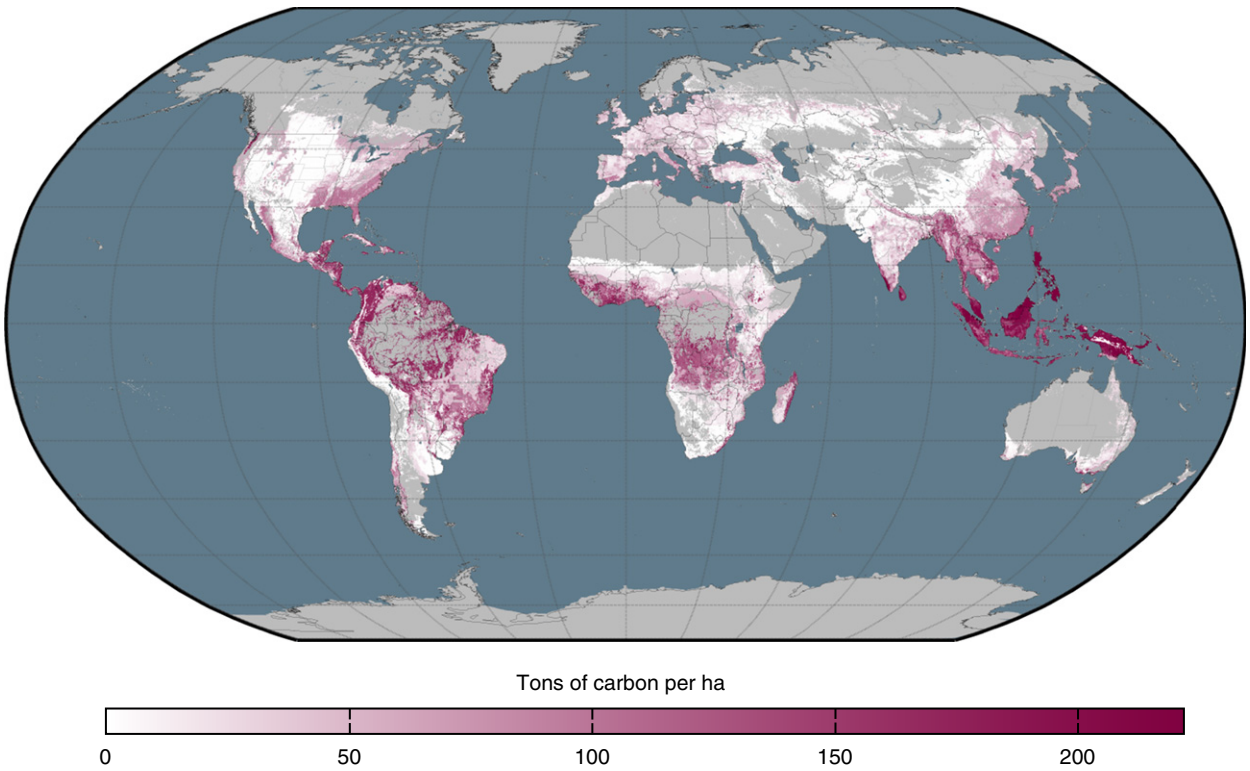


Figure 2 Carbon loss from clearing natural ecosystems. Clearing land for agriculture releases carbon stored in natural vegetation as carbon dioxide, a greenhouse gas. The map above illustrates carbon loss for areas currently utilized for croplands. On average, clearing land in the tropics releases twice as much carbon as in temperate regions ($\sim$120 vs. $\sim$64 ton ha^{-1}). Tropical rainforests commonly store 175–200 ton ha^{-1} of carbon. Data from West PC, Gibbs, HK, Monfreda C, *et al.* (2010) Trading carbon for food: Global comparison of carbon stocks vs. crop yields on agricultural land. *Proceedings of the National Academy of Sciences of the United States of America* 107: 19645–19648.

Agriculture is also a major contributor of other greenhouse gases, methane and nitrous oxide emissions, accounting for 50% and 60%, respectively, of annual global emissions (Smith *et al.*, 2007). The primary sources of these emissions are from rice production, livestock indigestion, and excess fertilizers. Combined, agriculture accounts for 30–35% of total annual greenhouse gas emissions (Smith *et al.*, 2007; van der Werf *et al.*, 2009; DeFries and Rosenzweig, 2010).

In addition to greenhouse gas emissions, clearing natural ecosystems can alter regional climate through changes in how

the Sun's incoming energy is absorbed and released by the land surface. When forests are cleared for croplands and pastures, more sunlight is reflected from the land back to the atmosphere. In the tropics, this creates a warming effect because such a large portion of the absorbed sunlight is released by the land in the form of sensible heat, which warms the air. In the boreal forest region, land clearing results in a cooling effect because several months of snow cover reflects so much more energy than the dark forests that were cleared. The effects of land clearing are less dramatic in many other biomes, but the magnitude of effects on regional climate in tropical and boreal regions is as strong as predicted climate responses to greenhouse gas emissions (West *et al.*, 2011).

Water Quantity

In addition to its effects on climate, agriculture has a large effect on water availability. To provide the right amount of water for crops, many farmlands have managed to increase or decrease the water quantity on the fields. As previously mentioned (*see* The Global Scope of Agriculture), more water is consumed for agriculture than any other use on the planet. Withdrawals in some regions are so extreme that major rivers such as the Colorado and the Yellow no longer reach the ocean every year. In another extreme example, in 2007 the Aral Sea was only 10% of its size when irrigation projects began in the Soviet Union in the 1960s. All three of these examples occur in regions with intensive surface water diversions. Other regions are using groundwater for irrigation much faster than it is being replenished, leading to longer term water shortages.

In other regions, croplands are managed to reduce water on the fields. Management practices such as levees are designed to keep rivers from their floodplains in order for them to be farmed. Wet soils are drained through tile drainage and ditching practices. In tropical regions, replacing forests with crops generally increases stream flows and potential flooding because the deep-rooted evergreen trees that transpire water throughout the year are replaced with season annual plants. For example, average daily stream flow substantially increased after replacing tropical savannas forest and savanna with soybeans in the southern Amazon (Costa *et al.*, 2003; Hayhoe *et al.*, 2011).

These water management practices not only improve crop yields, but also alter river flows. Many ecological processes and species have adapted to the natural flow regime, or flood pulse. The quantity, timing, and duration of flow events are all important for biodiversity. Too much or too little disrupts the ecosystem. Changing the flow regime alters nutrient cycling as well as many aquatic species that have aspects of their life cycle that coincide with the natural ebb and flow of water (Junk *et al.*, 1989; Richter *et al.*, 1997).

Water Quality

High fertilizer application rates are common in many places with crop or fertilizer subsidies. Such practices result in excess nutrients, leading to degraded water quality. Excess phosphorous in freshwater lakes and rivers promotes eutrophication. Since phosphorous is a limited nutrient in freshwater ecosystems, additional input stimulates algal growth, which reduces oxygen and sunlight for other species higher in the food web. Although it is not a limiting nutrient in freshwater systems, nitrogen has similar oxygen-depleting effects in coastal areas. Since the 1960s, the number of hypoxic or "dead zones" resulting from nutrient-rich rivers has doubled from 200 to more than 400 today (Diaz and Rosenberg, 2008).

Other major sources of water quality degradation from agricultural lands include sediment loss and pesticides. Soil tillage and intensive grazing cause soil loss from fields and pastures. If not deposited elsewhere in the field or pasture, eroded soil loss accumulates in aquatic ecosystems. Excessive sediment loads increase turbidity of the water, limiting light penetration and thus plant growth. In addition, nutrients and pesticides are commonly attached to sediment particles. These water quality effects primarily degraded aquatic ecosystems by promoting disturbance-tolerant species at the expense of conservative species, thus simplifying the ecosystem and making it more susceptible to further disturbance.

Invasive Species

Several species introduced to improve agricultural production have had negative effects on natural ecosystems. For example, fast growing, foreign grasses introduced to improve forage production commonly invade natural grasslands. In other cases, plants introduced to control erosion from steep slopes also establish in native ecosystems and replace native species. If we take a broader perspective on agriculture, trees introduced for agroforestry, timber products transported across regions, and fish introduced for aquaculture are all major vectors for invasive species introductions.

Thresholds of Biodiversity Loss

Degradation in response to ecosystem stresses is not always gradual. Regime shifts are large, abrupt, persistent changes in the structure and function of a system (Scheffer *et al.*, 2001). Such shifts occur when a tipping point or threshold is reached, triggering a switch in the dominant feedback processes structuring a system. Regime shifts have been studied in mathematical models and empirically documented in a variety of terrestrial and aquatic systems, including coral reefs, aquatic food webs, savannas, and the climate system (**Table 2**). Tipping points are a major concern for scientists, managers, and policymakers because of their potentially large impacts on biodiversity, ecosystem services, and human well-being, and because they tend to be difficult to predict, hard to control once they begin, and slow and expensive to reverse once they have occurred.

Human demands for food, fiber, and energy play a key role in driving many regime shifts especially through the conversion of natural and seminatural ecosystems to farming and the overexploitation of marine resources. While global biodiversity assessments have emphasized the significance of these drivers, the potential importance of thresholds, amplifying feedback and time-lag effects leading to regime shifts have received less attention. Most regime shifts result from complex feedback mechanisms or interactions between two or more drivers that

Table 2 Documented regime shifts and associated key drivers and ecosystem service impacts

Regime shift	*Key drivers*	*Ecosystem service impacts*
Freshwater eutrophication	Fertilizer runoff	Freshwater, fisheries, water purification, pest and disease regulation, recreation, and aesthetics
Coral–algae shifts	Fishing, disease	Fisheries, natural hazard protection, tourism, and aesthetics
Coral bleaching	Sea surface temperature	Biodiversity, fisheries, natural hazard protection, tourism, and aesthetics
Soil salinization	Clearing of deep-rooted vegetation	Food crops, water purification, and regulation of soil erosion
Bush encroachment	Overgrazing, fire suppression, and exclusion of browsers	Livestock production, woodfuel, and climate regulation
Forest – savanna biome shifts	Deforestation and forest degradation	Biodiversity, freshwater, food crops, livestock, timber, woodfuel, climate, and water regulation
Arctic summer sea ice	Greenhouse gas emissions	Wild animal and plant foods, climate regulation, and water regulation
Thermohaline circulation	Freshwater influx due to climate warming	Fisheries, food crops, livestock, and climate regulation

are not well accounted for in global change models. Because of this, the risk of substantial biodiversity losses, such as widespread Amazon forest dieback resulting from interactions between deforestation and climate change, have been substantially underestimated in earlier global change projections (Leadley *et al.*, 2010). Similarly, recent tipping point analyses indicate that rising atmospheric CO_2 concentrations and climate change could lead to important regime shifts at levels near or below the 2 °C global warming defined by the International Panel on Climate Change (IPCC) as "dangerous," including widespread coral reef bleaching, large shifts in marine plankton community structure, and extensive invasion of tundra by boreal forest (Leadley *et al.*, 2010).

Regime shifts are associated with species extinctions, declining species abundance, or shifts in species and biome distributions, as well as large impacts on ecosystem services and human well-being. For example, widespread degradation and loss of natural coastal habitats due to pollution, habitat destruction, changes in sedimentation, and sea-level rise are accompanied by increased risk of coastal damage by waves and storm surges as well as the loss of productivity of coastal fisheries. Biodiversity loss and the erosion of the capacity of ecosystems to deliver services often respond in similar ways to shared drivers; however, the relationship between them is not simple and may be different for the various dimensions of biodiversity. For example, the links between local species extinctions and reduced capacity to deliver ecosystem services remain, in many cases, elusive.

Experiments, observations, and models indicate that changes in ecosystem services are more tightly coupled to changes in the abundance and distribution of dominant or keystone species than to species extinctions. This calls for increased awareness of the importance of shifts in species distribution and changes in local abundance as the principal drivers of change in ecosystem services. While it is certain that regime shifts will occur in the future, the dynamics in most cases cannot yet be predicted with enough precision and advance warning to allow for secure and adequate approaches to avoid or mitigate impacts. This reality argues for a precautionary approach to human activities that are known to drive biodiversity loss, since lags in the socio-economic, biological, and physical systems of the Earth mean that many regime shifts are essentially irreversible over decadal to centurial time scales.

Creating a Sustainable Future

Most future planning efforts are based solely on our current and past needs and constraints. Yet these too are dynamic. In general, increases in food production will need to come from existing agricultural lands if we are to both feed a growing population and protect biodiversity. Below we summarize a set of future scenarios that addressed the trade-offs of agricultural production and biodiversity, the agricultural practices that have the greatest effect on biodiversity, and mechanisms for affecting change through policy and market-based incentives.

Alternative Scenarios

Over the past decade a number of large, integrated scenarios have been developed to explore alternative future development trajectories of the world and their implications for the environment. These include the scenarios developed by the Millennium Ecosystem Assessment (MA, 2005), the Global Biodiversity Outlook 2, the Global Environmental Outlook 4 (UNEP, 2007), the IPCC Special Report on Emission Scenarios (IPCC, 2000), and the International Assessment for Agricultural Science, Technology and Development (McIntyre *et al.*, 2009). These exercises typically start with the development of a limited number (3–5) of "storylines," which are then quantified using a variety of models to explore their consequences for agricultural development, biodiversity loss, etc. A summary of the Millennium Assessment scenarios and their impacts are given in **Table 3**.

Synthesizing the results from these studies indicates that land-use change, modification of river flow, freshwater pollution, and exploitation of marine resources are currently the most important drivers of biodiversity change and are projected to remain so over the coming century. Climate change and ocean acidification will also become increasingly

Table 3 Millennium Ecosystem Assessment scenarios

	Storyline summary	*Agricultural land in 2050 (million km^2)*[a]	*Loss of vascular plant diversity 1970–2050 (%)*[b]
Global orchestration	Global economic liberalization with strong policies to reduce poverty and inequality, substantial investment in public goods, for example, education	55	13
Order from strength	Economies become more regionalized, nations emphasize their individual security	60	16
Adapting mosaic	Regionalized economies, but with an emphasis on multiscale, cross-sectoral efforts to sustain ecosystem services	54	12
TechnoGarden	Globalized economy with substantial investments in sound environmental technology, engineered ecosystems, and market-based solutions to environmental problems	54	12

[a]Pasture and cropland. Extent of agricultural land in 1970 = 46 million km^2 and in 2000 = 51 million km^2.

[b]Relative loss when populations reach equilibrium at reduced habitat levels.

The table above summarizes possible agricultural land extent and vascular plant diversity in 2050 under four scenarios. Note that moderate biodiversity losses at the global level can translate to dramatic losses at the regional level or for particular species groups.

important (Leadley *et al.*, 2010). Projections of the impacts of these drivers on biodiversity show continuing and, in many cases, accelerating species extinctions and loss of natural habitat, as well as significant changes in the distribution and abundance of species and biomes over the twenty-first century. The greatest impacts involve large, highly visible modifications of ecosystems, including widespread conversion of tropical forest to pastures and croplands, climate-induced invasion of tundra by boreal forest, and reductions in the abundance of top predators in marine systems (Pereira *et al.*, 2010).

However, all is not doom and gloom. Analysis across different scenario studies highlights that there exist plausible development pathways with low greenhouse gas emissions and low land conversion rates, which could lead to much lower biodiversity impacts. These optimistic scenarios require fundamental changes in development paradigms, but are consistent with known constraints on economics, resource use, and human development goals (Pereira *et al.*, 2010).

However, if greenhouse gas emissions continue along current trajectories, several new models indicate that this will result in far greater climate-induced transformations of terrestrial biomes and marine biota than previously estimated. Furthermore, lags in the underlying socio-economic, climate, and biogeochemical drivers make increased biodiversity impacts inevitable over the next several decades. These lags mean that mitigation and adaptation measures aimed at reducing biodiversity impacts must be taken well before large impacts on biodiversity are observed.

Focusing Agriculture-Related Conservation Strategies

Strategies that aim to both increase food production and protect biodiversity need to focus on reducing the main negative effects of agriculture on biodiversity summarized above (*see* Effects of Agriculture on Biodiversity). Foley *et al.* (2011) advocate four strategies that would greatly increase food production while improving the environment. First, stopping agricultural expansion will reduce the rate of biodiversity loss as well as reduce greenhouse gas emissions. Second, improving yields on existing croplands increases food production and reduces the need for expansion. Third, water availability and quality can be improved by increasing the resource efficiency of agricultural production. Many gains in efficiency can be made through existing practices, such as changing fertilization rates and timing, irrigation methods, and livestock feed (reducing methane gas). Genetic improvements are a longer term solution that may generate more nutrient and water-efficient crop varieties. Finally, shifting our diets to be more plant based would reduce the need for agricultural expansion because a greater portion of the crops grown would be directly consumed by people, bypassing the inefficiencies of growing feed to produce livestock.

Meeting Agriculture and Conservation Goals: Mechanisms for Change

Implementing these four strategies will require a mix of policy and market-based incentives at multiple levels. The agricultural and conservation communities have made substantial progress increasing crop yields and protected areas, respectively. Regarding crop yields, agricultural investments in regions with low yields and inefficient use of water, fertilizers, and pesticides offer major opportunities for simultaneously enhancing food production. Underinvestment in agriculture, particularly in Africa, Latin America, and Eastern Europe, have led to significant yield gaps when compared with similar growing regions elsewhere around the world (Licker *et al.*, 2010; Neumann *et al.*, 2010). By one estimate, worldwide crop production could be increased by 60% if the yield gaps were closed for 17 major crops, thereby alleviating the demand to put new areas into cultivation (Johnston *et al.*, 2009). Investments in increased production must also enhance resource efficiency and farm management, including improvements in irrigation management, water harvesting and conservation, and integrated pest management (IAASTD, 2009; ten Brink *et al.*, 2010).

Historically, conservation policy has focused on increasing protected areas by keeping people apart from nature – limiting hunting, farming, grazing, and other human activities in protected areas. Today, over 105,000 protected areas cover about 21 million km^2, or 11%, of the world's land surface (West *et al.*, 2006). There is no doubt that this system of modern protected areas, begun in the nineteenth century with the national park systems of USA and great game reserves of colonial Africa, has been a critical tool for conservation. Many conservation groups and international policy focus their efforts, by targeting biodiversity protection in globally important regions (e.g., Olson and Dinerstein, 1998; Myers *et al.*, 2000).

Strategies to shift dominant patterns of production and consumption, however, are also needed in a world where the growth of population, resource use, and international trade create powerful incentives to convert sensitive habitat into agriculture and intensify production on existing farmlands. Monitoring and enforcement costs, as well as ethical objections to the criminalization of traditional livelihoods, make the search for alternative models all the more pressing (Karieva and Marvier, 2007). While protected areas that regulate land use directly will continue to play a role in conservation, indirect strategies will create new options for how people use the land.

Strategies focused solely on agricultural goals or solely focused on conservation goals will continue, but the fact is that food production and biodiversity conservation are mutual goals across a large part of the planet. Increasingly, strategies are needed that support agricultural production and biodiversity conservation, promoting smarter approaches to achieving these twin goals and managing necessary trade-offs more effectively. The following paragraphs provide examples of a number of potential mechanisms for change within agricultural policy and market-based approaches.

Agricultural Policies

Policies aimed at decreasing the environmental effects of agriculture have been in place for decades. In the USA, for example, the 1985 Conservation Reserve Program furthered policies begun in the 1950s to make direct payments to farmers to keep their land out of production. More recently, a number of states, provinces, and countries have adopted new policies that aim to maintain some semblance of the natural flows of rivers as part of a comprehensive water management strategy. These efforts improve upon earlier policies designed to maintain a minimum flow in rivers, with the intent of managing the river flows within the ecological limits of hydrologic alteration (Poff *et al.*, 2010).

Opportunities also exist for maintaining and enhancing on-farm agrobiodiversity. Some farmlands house a surprising range of nonfarm species, while maintaining high levels of domesticated plant and animal diversity (Daily *et al.*, 2001). Yet, agrobiodiversity is in decline. Single-minded policies intent on increasing production have favored biodiversity-poor monocultures growing genetically homogeneous cultivars. Increasing support for agrobiodiversity could resist this trend through both *ex situ* conservation in seed banks through organizations like the Consultative Group on International Agricultural Research (CGIAR) and the Svalbard Global Seed Vault (Global Crop Diversity Trust, 2007), as well as the *in situ* conservation of traditional cultivars and farming systems like the Potato Park in Peru (Argumedo, 2008).

Policies to shift agricultural economic incentives offer promising means to steer land use in a way that better navigates the trade-offs between agriculture and biodiversity. Currently, decisions made on what, where, and how to farm are primarily influenced by economic signals on the agricultural commodities that can be sold in markets. Immediate benefits to biodiversity could be gained by eliminating public subsidies that make the market price for crops artificially high. Entrenched economic and political interests, however, mean this tactic is difficult, time-consuming and, at times, at odds with the goals of social welfare and rural development.

Market-Based Approaches

As biodiversity and ecosystem services become more measurable and transparent, there will be increasing opportunities to incorporate these values in market-based approaches – internalizing costs that historically have been considered free and external to market systems. Several market-based mechanisms could facilitate this change for the agriculture sector, including payment for ecosystem services (PES) approaches, sustainability measures and certification schemes, and mitigation programs employing biodiversity offsets.

PES mechanisms can improve land-use and resource management in a manner that benefits service users and providers and supports conservation goals. While PES programs can differ substantially, the general approach involves service users compensating service providers for practices that support provision of a defined service (Wunder *et al.*, 2008). For example, many PES projects have been established to support improved watershed management. Downstream users of water (i.e., cities and industrial users) benefit from upstream land-use management, often implemented by agricultural communities. Such management practices may include erosion and sedimentation control, forest protection and management, and other steps to ensure water quality and supply. This approach can benefit both parties and nature conservation, as downstream users secure service improvements and upstream providers are compensated for the opportunity costs of changes to their management practices (e.g., preventing grazing on steep slopes).

The most significant PES policy related to agriculture and biodiversity currently being negotiated is a proposal for reducing emissions from deforestation and forest degradation (REDD). Initially conceived as a tool to reduce carbon emissions from tropical rainforests by transferring payments for forest conservation from developed to developing countries, the concept has evolved to recognize biodiversity, poverty reduction, and other benefits a global forest conservation policy could potentially provide. To work, REDD will need to devise a system to ensure accountability, compliance, and payment mechanisms that will change incentives for deforestation on the ground.

Certification programs and sustainability metrics are likely to play a stronger role in the agricultural sector as consumer demand increases for information about food production, businesses increase their focus on accountability within their supply chains, and measures improve for biodiversity and

ecosystem services impacts and uses. Certification schemes evaluate environmental and social performance to distinguish sustainable production practices from "conventional" agriculture, on the premise that consumers will prefer to buy or pay more for certified goods and services – organic, shade-grown, etc. (Bishop *et al.*, 2009). However, there are significant challenges to certification mechanisms, including costs of becoming certified (which can especially disadvantage small-scale producers), risks of changing farming practices, and confusion about the proliferation of sustainability metrics and programs (e.g., ISO standards, Global Reporting Initiative, Roundtables on sustainable biofuels, palm oil, soy, and other commodities). Nonetheless, with consumers increasingly wanting to support responsible brands, and companies seeking to enhance their brands and customer relations, the role of certification is likely to grow as an important means for conveying a brand's environmental and social performance.

Mitigation programs increasingly employ biodiversity offsets to address the footprint impacts of large-scale projects, such as those from mining, energy, infrastructure, and commercial agriculture projects. The aim is to support smarter development and mitigation decision making in order to maintain the conservation values and ecosystem services upon which people and nature depend. Biodiversity offsets compensate for residual environmental impacts of planned developments after appropriate steps have been taken to avoid, minimize, and/or restore impacts on site. Offsets are emerging as widely employed mechanisms advanced within the mitigation policies of a wide range of countries (McKenney and Kiesecker, 2010), in the sustainability policies of a number of major companies, and through science-based frameworks such as the Nature Conservancy's *Development by Design* (Kiesecker *et al.*, 2010). Although these mitigation programs are more commonly applied for energy, mining, and infrastructure projects, they have great potential to support sustainability goals for commercial agriculture, especially where it is expanding at the forest frontier (e.g., Brazil and Indonesia).

As illustrated above, there are many promising approaches for managing the trade-offs of agricultural production and biodiversity conservation. These interventions need to be targeted at specific threats in particular places in order to achieve the greatest benefit. Although biodiversity conservation or agricultural production may be the primary goal in some places, the challenge remains to manage the synergies and trade-offs between conservation and agricultural goals across the majority of the landscape. Finding the right mix of policies and incentives will depend on many factors, including food security, economics, and cultural practices.

Conclusion

Protecting biodiversity while increasing food production will require strategies that focus on the main drivers of biodiversity loss and degradation by agriculture on biodiversity: habitat loss and degradation, climate change, water quantity and quality, and invasive species. Particular emphasis should be placed in the tropics, where crop yields are lowest, the food security need is high, agricultural lands are expanding at the fastest rate, and the effects on species loss and climate change are highest. Most of the existing policies and market-based strategies are targeting decreasing habitat loss and the associated greenhouse gas emissions. New and improved strategies are needed to manage lands and waters to increase food production while decreasing the environmental impacts.

Appendix

List of Courses

1. Global Change
2. Conservation Biology
3. Introduction to Environmental Sciences
4. Sustainable Agriculture
5. Agroecology

See also: Agriculture, Sustainable. Ecosystem Services. Grazing, Effects of. Land-Use Issues. Loss of Biodiversity, Overview. Oil-Palm Plantations in the Context of Biodiversity Conservation. Slash-and-Burn Agriculture, Effects of. Sustainability and Biodiversity

References

Argumedo A (2008) The potato park: Conserving agrobiodiversity in an Andean indigenous biocultural heritage area. In: Amend T, Brown J, Kothari A, Phillips A, and Solton S (eds.) *Protected Landscapes and Agrobiodiversity Values*. Heidelberg, Germany: Kasparek-Verlag.

Bishop J, Kapila S, Hicks F, Mitchell P, and Vorhies F (2009) New business models for biodiversity conservation. *Journal of Sustainable Forestry* 28: 285–303.

Costa MH, Botta Al, and Cardille JA (2003) Effects of large-scale changes in land cover on the discharge of the Tocantins River, Southeastern Amazonia. *Journal of Hydrology* 283: 206–217.

Daily G, Ehrlich P, and Sánchez-Azofeifa A (2001) Countryside biogeography: Use of human-dominated habitats by the avifauna of southern Costa Rica. *Ecological Applications* 11: 1–13.

DeFries R and Rosenzweig C (2010) Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences of the United States of America* 107: 19627–19632.

Diaz RJ and Rosenberg R (2008) Spreading dead zones and consequences for marine ecosystems. *Science* 321: 926–929.

Dirzo R and Raven PH (2003) Global state of biodiversity and loss. *Annual Review of Environment and Resources* 28: 137–167.

Foley JA, DeFries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.

Foley JA, Ramankutty N, Brauman KA, *et al.* (2011) Solutions for a cultivated planet. *Nature* 478: 337–342.

Gleick PH (2003) Water use. *Annual Reviews of Environment and Resources* 28: 275–314.

Global Crop Diversity Trust (2007) Global strategy for the *ex situ* conservation and utilization of maize germplasm. http://www.croptrust.org/documents/web/Maize-Strategy-FINAL-18Sept07.pdf (accessed: 1 May 2011).

Hayhoe SJ, Neill C, Porder S, *et al.* (2011) Conversion to soy on the Amazonian agricultural frontier increases streamflow without affect streamflow dynamics. *Global Change Biology* 17: 1821–1833.

IAASTD (2009) Agriculture at a crossroads. *Global Report. International Assessment of Agricultural Knowledge*. Washington, DC: Science and Technology for Development.

IPCC (2000) *Special Report on Emissions Scenarios*. Cambridge: Cambridge University Press.

Johnston M, Foley JA, Holloway T, Kucharik C, and Monfreda C (2009) Resetting global expectations from agricultural biofuels. *Environmental Research Letters* 4(1): Article 014004. doi:10.1088/1748-9326/4/1014004.

Junk WJ, Bayley PE, and Sparks RE (1989) The flood pulse concept in river–floodplain systems. In: Dodge DP (ed.) *Proceedings of the International Large River Symposium. Canadian Special Publications of Fisheries and Aquatic Sciences* 106: 110–127.

Kareiva P and Marvier M (2007) Conservation for the people. *Scientific American* 297: 50–57.

Kiesecker JM, Copeland H, Pocewicz A, and McKenney B (2010) Development by design: Blending landscape-level planning with the mitigation hierarchy. *Frontiers in Ecology and the Environment* 8: 261–266.

Leadley PW, Pereira HM, Alkemade JRM, *et al.* (2010) *Biodiversity Scenarios: Projections of 21st Century Change in Biodiversity and Associated Ecosystem Services*. Montreal: Convention on Biological Diversity.

Licker R, Johnston M, Foley JA, *et al.* (2010) Mind the gap: How do climate and agricultural management explain the 'yield gap' of croplands around the world? *Global Ecology and Biogeography.* http://dx.doi.org/10.1111/j.1466-8238.2010.00563.x.

Matson PA, Parton WJ, Power AG, and Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 277: 504–509.

McIntyre BD, Herren HR, Wakhungu J, and Watson RT (eds.) (2009) *Agriculture at a crossroads: International Assessment of Agricultural Knowledge, Science and Technology for Development*. Washington, DC: Island Press.

McKenney BA and Kiesecker JM (2010) Policy development for biodiversity offsets: A review of offset frameworks. *Environmental Management* 45: 165–176.

Millennium Assessment (MA) (2005) In: Carpenter SR, Pingali PL, Bennett EM, and Zurek MB (eds.) *Ecosystems and Human Well-being: Scenarios*. Washington, DC: Island Press.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Neumann K, Verburg P, Stehfest E, and Müller (2010) The yield gap of global grain production: A spatial analysis. *Agricultural Systems* 103: 315–326.

Olson DM and Dinerstein E (1998) The Global 200: A representative approach to conserving the earth's most biologically valuable ecoregions. *Conservation Biology* 12: 502–515.

Pereira HM, Leadley PW, Proenca V, *et al.* (2010) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496–1501.

Poff NL, Richter BD, Arthington AH, *et al.* (2010) The ecological limits of hydrologic alteration (ELOHA): A new framework for developing regional environmental flow standards. *Freshwater Biology* 55: 147–170.

Ramankutty N, Evan AT, Monfreda C, and Foley JA (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles* 22: Article GB1003. doi:10.1029/2007GB002952.

Ramankutty N and Foley JA (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13: 997–1027.

Richter BD, Baumgartner JV, Wigington R, and Braun DP (1997) How much water does a river need? *Freshwater Biology* 37: 231–249.

Scheffer M, Carpenter SR, Foley JA, Folke C, and Walker BH (2001) Catastrophic shifts in ecosystems. *Nature* 413: 591–596.

Smil V (2000) Phosphorus in the environment: Natural flows and human interfaces. *Annual Reviews in Energy and Environment* 25: 53–88.

Smith P, Martino D, Cai Z, *et al.* (2007) Agriculture. In: Metz B, Davidson OR, Bosch PR, Dave R, and Meyer LA (eds.) *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, USA: Cambridge University Press.

ten Brink B, van der Esch S, Kram, *et al.* (2010) *Rethinking Global Biodiversity Strategies: Exploring Structural Changes in Production and Consumption to Reduce Biodiversity Loss*. The Hague, Netherlands: Netherlands Environmental Assessment Agency.

UNEP (United Nations Environment Programme) (2007) *Global Environment Outlook 4*. Nairobi, Kenya: United Nations Environment Programme.

van der Werf GR, Morton DC, DeFries RS, *et al.* (2009) CO2 emissions from forest loss. *Nature Geoscience* 2: 737–738.

West P, Igoe J, and Brockington D (2006) Parks and peoples: The social impact of protected areas. *Annual Review of Anthropology* 35: 251–277.

West PC, Gibbs HK, Monfreda C, *et al.* (2010) Trading carbon for food: Global comparison of carbon stocks vs. crop yields on agricultural land. *Proceedings of the National Academy of Sciences of the United States of America* 107: 19645–19648.

West PC, Narisma GT, Barford CC, Kucharik CJ, and Foley JA (2011) An alternative approach for quantifying climate regulation by ecosystems. *Frontiers in Ecology and the Environment* 9: 126–133.

Wunder S, Engel S, and Pagiola S (2008) Taking stock: A comparative analysis of payments for environmental services programs in developed and developing countries. *Ecological Economics* 65: 834–852.

Fires, Ecological Effects of

William John Bond, University of Cape Town, Rondebosch, South Africa

© 2013 Elsevier Inc. All rights reserved.

Glossary

Fire regime The type of fire, mean and variance in fire frequency, intensity, severity, season, and areal extent of a burn in an ecosystem.

Fire-stimulated recruitment Seedling recruitment in the first one or two growing seasons after a burn.

General circulation model (GCM) Computer models developed to simulate global climate and widely used for global climate change predictions.

Prescribed burning Fires intentionally lit for management purposes.

Serotiny Seeds stored on the plant with dispersal triggered by fire.

Fire in Earth History

Fire is an enormously influential disturbance over very large areas of land in the modern world. Vegetation burns because the earth's atmosphere contains sufficient oxygen (>15%) to support combustion. Should oxygen levels rise above 30%, fires would be so frequent that dense forest vegetation, even in persistently wet climates, would be incinerated. There is an almost continuous record of fossil charcoal over the past 350 million years indicating that the atmosphere supported combustion for most of terrestrial plant evolution. Oxygen levels reached maxima in the Upper Carboniferous, 300 million years ago (Ma), when abundant fossil charcoal indicates frequent fires. Fires were also common during the Cretaceous (135–165 Ma) when flowering plants (angiosperms) first began to spread. Fossil flowers, with fine structure beautifully preserved as charcoal, are common and widespread in Cretaceous deposits. At these and other times, frequent fires may have played a significant part in the ecology and evolution of paleo-ecosystems (Bowman *et al.*, 2009).

Broadleaved forest, analogous to present-day tropical and temperate forests, first became globally widespread in the Eocene (55–35 Ma) a warm wet period. Fossil evidence for fire is rare at this time but dated molecular phylogenies indicate that fires were continuing to burn in Eocene landscapes. Grasslands are the most flammable vegetation that has existed in earth history. Tropical (C4) grasslands and savannas are the most extensive flammable biomes today occupying one-fifth of the world's land surface. Though C4 grasses are ancient (30 Ma), the savanna biome first began to spread from the late Miocene (8 Ma). Charcoal from marine sediments increased dramatically during the past 10 Ma as fire-promoting ecosystems, including savannas, began to spread.

Hominids have used lightning-ignited fire for perhaps as long as 1–1.5 My but first began to ignite their own fires from 200 to 400 ka. Fire was used as a tool for managing the environment to promote the growth of edible plants in hunter–gatherer communities and to attract animals to hunting grounds. Fire was (and is) a tool used by farmers to clear new lands and to prepare sites for swidden-type farming. However, the historical impact of human use of fire on the environment may be less than previously assumed. Increases in charcoal in south-east Asia (50 ka) and Australia (60–45 ka) are coincident with the arrival of fire-using modern humans. However, charcoal records from Europe reveal that climate change, rather than changes in human use of fire, best correlates with fire activity from 70 to 10 ka (Daniau *et al.*, 2010). Charcoal records from around the world for the past two millennia indicate little impact of humans on biomass burning until as late as the mid 1700s with an abrupt decline in burning after 1870 (Marlon *et al.*, 2008). Thus, the area burnt by vegetation fires may be at its lowest level for the past few millennia.

World Biomes and Fire Incidence

Satellite imagery is beginning to reveal the vast global extent of fires. Between 1997 and 2008, the average global burnt area was 3.7 million km^2 or 2.8% of the total unglaciated land area (Giglio *et al.*, 2010). Nearly one-third of the land surface experienced fire activity between 2001 and 2006. Africa accounts for about 70% of annual burnt area with the remaining 30% primarily in Australia, followed by South America and Central Asia. Fires are rare only at the extremes of the climatic continuum. The most humid tropical and temperate forests and the driest deserts have the smallest proportion of burnt area annually. However, between these two extremes, fire has influenced the extent and composition of a great diversity of ecosystems, including tropical grasslands and savannas, temperate grasslands and steppe, boreal forests, dry conifer forests, temperate woodlands, Mediterranean-type shrublands, heathlands, and eucalypt woodlands. Forests with mast-flowering bamboo understories are also prone to burning after the bamboos flower and die, creating massive fuel loads. Humans have changed landscape patterns of burning. Even humid tropical forests are beginning to burn as a result of logging and deforestation fires. All these biomes experience fires of widely differing frequency and severity which help shape ecosystem structure and function.

Given its wide geographic extent, fire inevitably influences the distribution and abundance of many species. Some ecosystems are dominated by species that depend on fire to complete their life cycles. Others are dominated by species that tolerate burning but have no direct dependence on fire.

Ecosystems that seldom or never burn, except when disturbed by human activity, contain mixtures of species that fortuitously tolerate burning and species extremely intolerant of fire. The impact of burning on biodiversity varies greatly among these different types of ecosystems and species-response patterns.

Determinants of Fire

The incidence of natural fires depends on much the same combination of ingredients as a campfire: dry fuel, flammable material (finely branched plant material that dries quickly), continuity of fuels, and a source of ignition. Suitable weather conditions and winds to carry the fire and preheat flammable material are also the key factors. The flammability of vegetation varies with the moisture content of plant biomass. Dead matter has the lowest moisture content, whereas live leaves will burn more easily if their moisture content is low. The shape, size, and arrangement of plant parts influence moisture content and flammability. Plants with narrow leaves or thin branches dry rapidly and burn readily. Ecosystems that accumulate slow-decomposing litter are highly flammable. Leaves with large amounts of oils, fats, waxes, and terpenes also burn readily. Volatile substances enhance burning because they are released from leaves, burn fiercely, and thus dry and heat adjacent material.

Because fire depends very much on moisture content of vegetation, climate and weather conditions exert marked influences on the timing of burns. The length of a warm, dry period needed for ignition depends greatly on vegetation properties. A few days of hot, dry weather are sufficient to dry tall grasslands enough to sustain a fire, whereas months of extreme hot and dry conditions are needed for fire to burn pristine humid tropical forests. Therefore, large fires in woody ecosystems are generally associated with rare drought events, such as those produced by El Niño conditions. In contrast, large fires in arid grassy ecosystems are limited by the availability of fuel and are more common after high rainfall years. Fires generate their own heat that, under hot, dry, and windy conditions, creates a positive feedback, increasing the area burnt and the severity of the fire by drying the vegetation before the fire front.

A source of ignition is needed to start fires. Today, most fires are ignited by humans, except in sparsely populated regions. Lightning fires are still common in many landscapes. Lightning often accounts for large numbers of fires but small areas burnt because many are dowsed by subsequent rain or controlled by fire crews attuned to the risk of thunderstorm fires. Rockfalls also ignite fires, accounting for the spate of burns often triggered by earthquakes. Volcanic activity is of local significance, especially on islands in which lightning is unusual. The contingent requirements of ignition, suitable weather conditions, and contiguous flammable vegetation make for a high degree of uncertainty in predicting landscape fire patterns.

The Fire Regime

The ecological effects of fire depend on the fire regime rather than the occurrence of a single fire. The fire regime is produced by the combined effects of climate, fuel properties of vegetation, and ignition frequency. It is characterized by the type of fire, mean and variance in fire frequency, intensity, severity, season, and areal extent of a burn. Types of fire include ground fires that burn the organic layers of the soil, surface fires that burn just above the ground, crown fires that burn in the canopies of trees, and mixed fire regimes where fires vary in severity with both surface and crown fire patches. Ground fires occur only in peaty soils in which they can be extremely damaging, destroying roots and completely altering soil properties. Crown fires are high severity fires typical of low productivity sites including Mediterranean-type shrublands, and North American boreal forests. Surface fires are common in many woodlands and forests where litter is a major fuel. Surface fires with herbaceous fuels predominate in grasslands and savannas and can burn at annual or even subannual intervals where productivity is high. When crown fires do occur in forests, they produce massive stand-replacing fires. Fire exclusion in national parks, such as the giant sequoia forests of California, has led to an increase in young trees which now act as bridging fuels turning surface to crown fires with extremely damaging consequences.

Fire frequency is estimated from maps of fires, records of fire scars, or patterns of charcoal deposition in sediments. The mean and variance of fire return interval are important descriptors of historical disturbance patterns. Changes in fire frequency are an important cause of changes in ecosystem structure and function. Fire intensity is measured as energy released per meter of fire front. It is widely used in firefighting operations. Fire severity is a distinct but related concept defined as the ecological impact of a single fire and usually estimated from the amount of plant biomass consumed. A fast-moving fire that consumes little biomass and a slow-moving fire that consumes more can have the same fireline intensity but different severity. Fire severity is highly variable, literally depending on the weather during a burn, wind conditions, and the preburn condition of the vegetation. Fire season is largely dictated by the moisture content of flammable biomass. Where the vegetation dries out quickly, fires can burn in almost any season. Seasonal timing of burns can cause significant changes in species composition and ecosystem structure. Continuity of flammable vegetation, especially at the landscape scale, strongly influences the spread of fires. Habitat fragmentation can lead to a reduction in fire frequency of isolated fire-prone ecosystems or an increase in fire-excluding forests surrounded by flammable vegetation. Land abandonment in some countries has led to successional changes producing large, contiguous, highly flammable vegetation. In the Mediterranean region, reduction of pastoral activities has led to the conversion of grasslands to highly flammable shrublands. This process has contributed to an increase in the area burnt annually from a few thousand hectares in the 1960s to hundreds of thousands of hectares in recent years.

Fire Regimes and Phytogeography

Flammable ecosystems differ greatly from one another in the mix of growth forms and their fire-adaptive traits. This diversity is linked to differences in fire regime. Fire regimes are the

product of a synergy between external physical factors and the properties of vegetation. In the absence of humans, key physical preconditions for wildfires were climatic conditions conducive to lightning strikes, periods dry enough for ignition to occur, and enough oxygen in the atmosphere to sustain burning. Where fire influences the geography of plant communities, another physical precondition is landscapes free of major barriers to fire spread. Such barriers include rivers, lakes, ice and snow, gravel beds, and other areas with sparse plant growth. Within these physical constraints, the vegetation itself is the major contributor to the fire regime.

The way in which the plants themselves contribute to the fire regime is most apparent where contrasting ecosystems, with entirely different fire regimes, occur in the same landscape. These have been described as alternative stable states with each ecosystem state maintained by a different set of positive feedbacks. Examples include mosaics of fire-resistant tropical forests and flammable savannas occupying extensive areas of Africa, South America, and Australia. In some instances, mosaics of alternative ecosystem states occur where both states are flammable but with contrasting fire regimes. In western North America, for example, chaparral shrublands with crown fire regimes occur in landscape mosaics with conifer forests with litter-fueled surface fire regimes. Changes in the fire regime can be brought about by changes in external conditions, such as climate change, but also by changes in the growth form mix, producing changes in flammability or productivity. The shortest interval between successive fires is constrained by the time taken to build sufficient, continuous flammable biomass for fires to spread. This depends on the plants present and site productivity. Changes in plant growth form dominance can have major impacts on fire regimes such as those caused by introduction of highly flammable invasive grasses into shrublands and woodlands.

Species Response to Burning

Plants

Ecosystems subjected to similar fire regimes have convergent vegetative and reproductive traits. There are clear distinctions, for example, between fire-adaptive traits characteristic of crown fire regimes fueled by woody plants as opposed to surface fire regimes fueled by litter or herbaceous plants (Keeley *et al.*, 2012). In crown fire regimes, a fire event triggers flowering, seed dispersal, or seed germination in fire-dependent plant species. Fire-stimulated flowering is common among perennial grasses and herbs, including orchids, lilies, and other bulb plants. These flower prolifically after they have been burnt with some species showing a facultative response (continuing low levels of flowering in unburnt vegetation) and others an obligate response with flowering cued by smoke (e.g., *Cyrtanthus* spp. in Africa, *Xanthorrhea* in Australia). Fire-stimulated recruitment also occurs when fires stimulate seed release from woody species with serotinous cone-like structures which store seeds on the plant for years between fires. Serotiny is common in conifers of North American boreal forests and Mediterranean-climate regions and also among diverse groups of flowering plants in Australia and South Africa. Some species are polymorphic for the trait, with serotinous forms increasing in populations that regularly experience large severe crown fires. Fire-stimulated seed germination from soil seedbanks is also common in crown fire regimes. Dormant seeds in the soil show heat-stimulated seed germination, especially in legumes and other clades with hard seed coats (e.g., members of Rhamnaceae). Thick seed coats prevent imbibition of water until cracked by the heat of a fire. Smoke-stimulated seed germination has been reported for many species in fire-prone shrublands of South Africa, Australia, and California. However, smoke-stimulated germination has also been reported for plants that do not come from fire-prone ecosystems. Regardless of the nature of the germination cue, the appearance of numerous seedlings after a fire event is characteristic of fire-dependent species in ecosystems with an evolutionary history of crown fire regimes.

Vegetative features of plants also vary with fire regime. Thick bark and/or the ability to resprout from insulated buds allow many plants to survive burning. Sprouting is a common fire survival mechanism, either from the root stock or from branches above the ground. Some species possess large swollen burls or lignotubers which are thought to act as bud banks or storage reserves. Paradoxically, many woody plants in crown fire regimes cannot resprout and are killed by fire. These nonsprouting plants often have higher seed production and higher seedling growth than related sprouting species. In some lineages, sprouting is the ancestral feature and loss of sprouting is viewed as an adaptive response to fire. Nonsprouting shrubs are particularly common in chaparral and similar shrublands and require fire to release seeds from serotinous cones or to stimulate germination. Among trees in fire-prone forests, many conifers do not sprout and a few eucalypts are also killed by fire. Nonsprouting species are particularly prone to local extinction if recruitment fails after burning.

The occurrence of species with flammable morphologies in crown fire regimes has led to the suggestion that flammability has evolved to promote burning. However, vegetative traits that promote flammability may also serve other functions. Some woody plants accumulate highly flammable fuel by retaining dead branches while also requiring fire for recruitment. In western North America, serotinous pine species retain dead branches and recruit seedlings only after crown fires have destroyed the trees. These species appear to both promote the spread of crown fires and benefit from fire-stimulated reproduction. This strategy contrasts strongly with self-pruning, thick-barked pines (e.g., *Pinus ponderosa*) which tolerate frequent surface fires and do not require fire for seedling recruitment.

Surface fire regimes fueled by herbaceous plants select for a completely different suite of plant traits. Grasses are among the most fire resistant of all plant growth forms. The buds of new shoots are insulated by either layers of leaf sheaths or the soil where species have underground rhizomes. Grasslands recover from burning more rapidly than woody plants and can carry very frequent fires (1–3 years) in productive sites. Although many grasslands burn readily, few species have an obligate dependence on burning. Fire-stimulated flowering is rare but has been reported in many mostly temperate tussock grass species, including species of *Chionochloa* in New Zealand.

Several very widespread warm-climate (C4) grasses (e.g., *Themeda triandra* and *Andropogon gerardi*) decline rapidly in the absence of burning. These species become locally extinct if fires are suppressed for more than a decade since litter accumulates, shades, and kills the grass plants. Unlike crown fire systems, all woody plants in savannas resprout after fire and seedling recruitment is not fire dependent. Instead, savanna trees have developed a remarkable ability to tolerate very frequent grassland fires. These create a particularly hostile environment for juvenile stages. Tree seedlings survive by rapidly acquiring the ability to sprout, and juveniles slowly develop food reserves in swollen roots and eventually produce bolting stems that place foliage above flame height. This peculiar life history occurs in several pine species (e.g., *Pinus palustris*) growing in grassy habitats and also in many savanna trees. A tropical Asian species, *Pinus merkusii*, produces persistent sprouting juveniles in frequently burnt populations but nonsprouting forms in infrequently burnt island populations. Juvenile stages of savanna trees can tolerate repeated burning, suffering repeated killing of the stem parts, for decades before they die or escape the flame zone to become adults.

Animals

The direct effects of fire on wildlife are often surprisingly small. Agile animals flee to unburnt refugia, moving across the fireline to places of safety. Soil is a very effective insulator so that many animals survive in crevices and cracks or in burrows in the soil. Mortalities of large mobile vertebrates, including humans, do occur but only in the most severe fires. Reptiles and slow-moving invertebrates can suffer higher mortalities and their carcasses provide a food source to scavenging birds and other creatures in the first few days after a burn. The threatened bald ibis of South Africa makes extensive use of recently burnt grasslands, as does the endangered whooping crane in its Texan winter feeding grounds.

The indirect effects of burning are generally far more important than fireline mortality, especially changes in habitat attributes as vegetation recovers from a fire. A large crown fire in a forest causes drastic structural change and local extirpation of all faunal elements that depend on unburnt forest habitat. Postburn stages are colonized by a new suite of species. Different successional stages support different suites of animals. Even frequently burnt grasslands, such as those of the South African highveld, have distinct bird assemblages which turn over with successive years of regrowth after burning.

The pattern of fires across a landscape imposes a mosaic of patches of different successional ages. The size and configuration of patches influence rates of local extinction and patch recolonization (metapopulation structure). For example, nectar-feeding birds in shrublands of Australia and South Africa lose their food source, shrubby members of the Protea family, after a burn and have to seek unburnt stands for food. The landscape configuration of old stands with flowering proteas, and young stands with immature shrubs, necessitates a highly mobile bird assemblage. Changes in the spatial pattern of fires may change extinction risks in different faunal elements. Some species of Australian honeyeaters are threatened with extinction because changes in the fire regime no longer produce the right mix of mature and immature populations of nectar plants. The decline of many bird species in Australian savannas has been attributed to the development of homogeneous landscapes through the systematic burning of very large tracts of land. Smaller, patchier burns are thought to have prevailed under aboriginal burning practices. The extinction of many Australian small to medium-sized (50 g to 5 kg) mammals since European settlement (23 species) has been attributed to similar changes in the fire regime, but predation by feral animals is a major factor.

Fires in tropical rain forests, caused by human impacts on forest structure, have had devastating effects on the forest fauna. In Sumatra, for example, primary forest specialists, including squirrels, hornbills, and other fruit-eating and frugivorous birds, and some primate species disappear altogether from burnt and adjacent forests. The increasing risk of fire in humid tropical forests poses serious threats to survival of the forest fauna in addition to those caused by direct forest clearing.

Ecological Effects of Fire

Ecosystem Structure

The consequences of increasing fire frequency and severity for ecosystem structure are to reduce vegetation height (tall forests to shorter ones and woodlands to shrublands); to replace woody vegetation by grasslands; to promote flammable species or communities (low litter decomposition rates, more xeromorphic leaves, and finer twigs/branches); and to reduce biomass. Both tropical and temperate landscapes contain mixtures of fire-prone grassland or shrubland communities and closed forests which tend to exclude fire. Savannas and closed forest are well-studied examples of such alternative ecosystem states. The relative proportions of flammable savannas and fire-resistant forests vary across precipitation gradients and, locally, with soil type. Boundaries between states are typically abrupt with fires excluded from forests, which shade out grasses. In some instances, paleoecological studies have shown that these sharp boundaries have remained stable for millennia. Simulations using physiologically based global vegetation models suggest that forests would at least double in extent in the absence of fire, particularly in the flammable savanna biome. Replacement of flammable communities by fire-resistant forest elements frequently occurs when fires are suppressed. In southern Africa, forests have replaced savannas after 10–30 years of fire suppression. However, counterexamples exist in which fire suppression does not lead to changes in savanna/forest boundaries. Stable boundaries often coincide with different soil types, with forests occurring on the better drained or more fertile soils.

Changes from fire-resistant to flammable ecosystems may also be rapid. In the Brazilian Amazon, fires in closed-canopy forests spread as a "thin, slowly creeping ribbon of flames a few tens of centimeters in height." Despite the low severity of an initial fire, burning causes structural changes, opens up the forest canopy, dries out the understorey and contributes to an increase in flammable understorey biomass, increasing the risk of a second fire. Weedy vines and grasses quickly colonize

twice-burned forests, further adding to the flammable biomass. Positive feedbacks of this kind are estimated to reduce a forest to scrubby vegetation resembling recently abandoned farmland in 20–30 years.

Ecosystem Function

The immediate effect of fire is gaseous loss of carbon and nitrogen from material that burns. More biomass burns in more severe fires and the nutrient losses are greater. Strong winds accompanying fire lead to losses of phosphorus and cations blown away in ash. Cation nutrients in ash tend to be mobile and in a plant-available form. Their presence leads to increases in soil pH – large increases in acid forest soils and smaller increases in neutral or alkaline soils in grasslands or savannas. Increased solar radiation, decreased evaporation, and higher pH lead to increased microbial activity, increased rates of mineralization, and increased availability of nutrients after a burn. After a chaparral burn, for example, nitrate increased more than 20-fold relative to unburnt controls. Short-term increases in nutrient availability can be offset by long-term decreases where fire frequencies are high and inputs to the system between fires are not high enough to replace losses. Severe fires can lead to nitrogen shortages. Many ecosystems have nitrogen-fixing organisms as major components of postburn vegetation which replace nitrogen losses in a few years.

Fire can lead to changes in ecosystem processes at landscape scales. The reduction in biomass caused by burning and changes in soil properties lead to temporary hydrological changes in patterns of stream flow. Severe fires can lead to increased soil erosion. The Yellowstone fire of 1988 led to significant increases in sediment loads and altered the geomorphology of river systems. Debris spread over a distance of 12 km in one valley bottom.

Species and Populations

At the local scale, and within flammable ecosystems, species respond to differences in fire frequency, season, and severity. Variation in the fire interval is an important determinant of population trends. In crown fire regimes characteristic of woody fuels, the effect of fire on population growth depend on key demographic attributes of the species. Population size of nonsprouting species fluctuates more than that of sprouting species, and local extinction is not uncommon after a single fire. Species that are slow to mature are particularly vulnerable where populations are burnt before they have first flowered and set seed. Populations are also negatively affected where intervals between fires exceed the life span of a species or its seedbank. C4 grasses are also sensitive to variation in fire frequency with the dominant species in some grasslands disappearing after a decade or more of fire exclusion. The rich forb diversity in upland grasslands in Africa and South America is also dependent on frequent fires and fire exclusion has been shown to lead to loss of long-lived perennial species, especially those with large underground storage organs. This contrasts with North American prairies where fire exclusion promotes forb diversity. Manipulation of the fire interval is a key tool for influencing biodiversity of vegetation stands. In flammable woody ecosystems, information on the reproductive status of plants, especially the size of the viable seedbank, at different postburn stages has been extensively used to help determine optimum fire frequencies to maintain particular species.

Recovery of plant populations also depends on the unique combination of circumstances on the day a fire burns. These "event-dependent" effects can be as important as fire frequency in influencing biodiversity in some ecosystems. Fire season has a marked influence on recruitment of serotinous members of the Proteaceae in Western Australia and South African fynbos. Spring burns can reduce protea populations to less than one-tenth of their preburn density, whereas autumn burns can result in a 10-fold or greater increase in plant densities. Fire season also influences recovery of sprouting plants where the size of root reserves varies seasonally, affecting the vigor of resprouting. Manipulation of fire season is sometimes the only effective tool for managing densities of sprouting shrubs. Stem density of clonal species of hazel (*Corylus* spp.) in the understorey of temperate forests increased fivefold with four successive spring burns but was halved by successive summer burns. In Zambian woodlands, annual burns in the early dry season caused a 10-fold increase in tree seedlings and halved the adult mortality rate over a decade. Grassland composition can also be very sensitive to fire season. In a long-term burning experiment in the Kansas prairies, late-spring burns caused a halving of *Andropogon scoparius* biomass relative to fires burnt a few weeks earlier. The effects of fire season on species and ecosystem recovery are poorly known in many ecosystems. They may be unimportant where the fire season is short because of climate constraints.

By definition, severe fires cause the most extreme biomass losses in ecosystems. In eucalyptus and conifer forests, intense crown fires kill all aboveground plant growth. Where trees are incapable of sprouting, these fires cause complete replacement of canopy trees. Sprouting plants, especially shallow-rooted species, can be killed by high severity burns. The effects of fire severity on recruitment from seeds vary among species. Legumes and other plants with hard, dormant seeds in fire-prone shrublands do not germinate unless a burn heats the soil sufficiently. For example, the Australian shrub, *Acacia suaveolens*, will not germinate unless soils are heated to more than 50 °C. Variation in fire intensity can directly affect species composition in fire-prone shrublands, incinerating some seeds and stimulating germination in others. Key species, such as legumes, may fail to germinate after low-intensity fires that are applied for safety reasons.

Fire intensity is an important factor in savanna ecology. Where grass growth is sufficient to carry fires at frequent intervals, burning kills the aboveground parts of juvenile trees and shrubs. The amount of dieback depends on the intensity of the burn. In mesic savannas, fires are so frequent and so intense that juvenile trees may spend decades trapped in the grass layer. The frequency and intensity of fire are important determinants of tree biomass (and habitat structure) in savanna ecosystems.

The relative sensitivity of plants to fire season and fire severity varies among species. This makes general prediction of population trends under different fire cycles difficult without

species-specific information. In Mediterranean shrublands, species with diverse responses to frequency, season, and intensity of burn occur in the same community suggesting a history of fires that vary in these factors. It is a considerable conservation challenge to incorporate variability into fire regimes to maintain the full diversity of species.

Interactions between Fire and Other Ecological Processes

Fire and Herbivory

Fire interacts with, and is influenced by, other agents of disturbance that can also influence ecosystem structure on a large scale. Herbivores influence the distribution and biomass of plants and therefore the attributes of fire regimes. Heavily grazed savanna grasslands do not burn so that animals capable of grazing grasses too short to burn, such as wildebeest, white rhino and prairie dogs, can reduce landscape fire activity. Extirpation of short-grass grazers can, in turn, promote more frequent larger fires reducing landscape heterogeneity. Persistent heavy grazing by cattle often leads to an increase in tree densities because of the reduction in fire frequency. In Africa, elephants open up woodlands, enhancing grass growth which promotes more frequent severe fires. The combination of elephants and grass fires can cause a marked reduction in tree densities. In miombo woodlands of Zimbabwe, changes in woodland structure under the combined influence of elephants and fire markedly reduced bird diversity and led to local extinction of four endemic woodland bird species.

Insect herbivory also influences fire regimes, especially in northern ecosystems. In Balsam fir (*Abies balsamea*) and red spruce (*Picea rubens*) forests of eastern North America, fires are rare. Large-scale tree mortality is caused by spruce budworm outbreaks which inhibit the spread of fire by changing vegetation structure and fuel properties. In general, however, plants that provide fuel for burning make poor food for herbivores and vice versa. Fires burn readily where decomposition is slow, causing dead matter to accumulate. Slow decomposition is associated with high C:N ratios, high fiber contents, and high leaf-specific weight, all of which inhibit herbivore food intake. Thus, the most fire-prone vegetation tends to be least edible and vice versa.

Fire and Landscape Fragmentation

The spread of fire in a landscape is sensitive to the continuity of flammable vegetation. Landscape fragmentation can have major impacts on the fire regime, in turn affecting survival of species within fragments. Species losses from North American prairie fragments during a 50-year period were greater among plants of early postburn environments. Similar patterns of local extinction occur in fynbos fragments surrounded by nonflammable forests in South Africa. The causes of species loss in both prairies and fynbos are attributed to infrequent burning due to isolation of the fragments. Small forest patches in extensive areas of fire-prone shrublands or grasslands are also more likely to suffer local extinction of species intolerant of burning. In the tropical forests of Guyana, fire-tolerant forest tree species with thick, fissured bark, and small seeds are unusually common near savanna boundaries and human settlements.

Fire and Invasives

The interplay between fire and invasive species can cause alarming ecosystem transformation. Direct effects of invasive plants can be minor relative to the indirect effects on fuel properties and fire regimes. Grass invasion into woody ecosystems has particularly damaging consequences. In Hawaii, invasion of tall nonnative grasses has transformed the unique forests to grasslands entirely alien to the island ecosystem by fueling frequent fires. In South America, invasion of logged-over tropical forests by fire-promoting grasses can cause elimination of the forest ecosystem and its rapid replacement by the alien grass. In southwestern Australia, species-rich heathlands are fire dependent but they have also been invaded by nonnative grasses which burn so frequently that the heathlands are transformed into a species-poor savanna with scattered relictual trees. The reverse pattern, that of invasion of fire-prone grasslands by plants that do not burn easily, may also be a problem. For example, *Lantana camara* is invading fire-dependent grasslands in South Africa but burns much less readily than the native vegetation.

Managing Fires

Because wildfires are such a widespread feature of world vegetation, managing fires for particular objectives, including conservation of biodiversity, is a major concern. Attitudes and actions to wildfire are never neutral. Fires are actively suppressed in some ecosystems but purposely ignited in others. Wildfires commonly arouse public and media concern. The spread of the urban interface as more people flock to cities has increased threats of wildfire to people and properties. Smoke management has become a health and safety issue. Enormous effort and expense are often expended on fire management. Appropriate fire management policies continue to be a central management headache in protected areas as different as Yellowstone National Park in a coniferous ecosystem and Kruger National Park in African savannas. The expense of maintaining firefighting teams consumes a significant fraction of conservation agency resources. Litigation, when fires move beyond the borders of protected areas, can also strain conservation budgets. Arson burns are not uncommon, sometimes as a protest against state authority. Fire management consumes a great deal of time and is a major expense in fire-prone ecosystems.

Fire Management Policies

There is no consensus on how to manage fires in protected areas or outside them. For the first half of the twentieth century, complete fire suppression was a common policy, and it still is in many parks. Suppression policies have slowly changed, partly because of their cost, partly because they do not work, and partly because of changes in ecological

thinking. Fire suppression leads to the buildup of dead biomass in fire-prone ecosystems which produce very severe fires when they do burn. Suppression policies were also relaxed following the discovery of fire-dependent features of plants. In the Cape fynbos, the beautiful marsh rose (*Orothamnus zeyheri*) declined to a handful of plants before managers realized that the species had an obligate dependence on fire to stimulate germination of its seeds. There has also been increasing recognition of disturbance as a natural process in ecosystems. Outside of rain forests, complete fire suppression is rarely the aim of fire management. However, decades of suppression have led to pronounced changes in ecosystem structure that now pose major challenges regarding how to reintroduce fires without causing more problems. Many conifer forests were maintained as open parklands by frequent surface fires. Fire suppression has allowed numerous young trees to establish, creating "bridging fuels" and a switch from surface to crown fires which destroy the whole forest structure. In some savannas, fire suppression has led to invasion of grasslands by forest species in a process that is very difficult to reverse except by costly manual clearing of trees to restore large mammal habitat.

Prescription burns are fires intentionally lit for management purposes. Safety for the fire crew, and for property, is always an important consideration. Therefore, prescribed burns often cause significant changes to the fire regime, especially to season and intensity but also to fire frequency. Changes to the fire regime can lead to significant population declines in sensitive species such as legumes, which depend on intense heat for seed germination.

Prescription burning calls for clear management objectives. Because fire is so influential a force on community structure and composition, decisions have to be made regarding the desirable objective and what kind of burning pattern should be used to achieve it. In the South African savanna parks, fire policy changed from complete fire exclusion to prescribed burning at fixed intervals (promoting large mammals) to fire regimes that "maximize biodiversity." The intention of the latter is to create and sustain a landscape mosaic of different successional ages or habitat types that can maintain viable populations of most species. What to do in practice is more difficult.

Another commonly considered policy is to recreate "natural" burning regimes. This policy permits fires that have been ignited by lightning, rockfalls, or other nonhuman agents but suppresses fires of human origin. It is difficult to apply in practice because vegetation has been fragmented by roads, buildings, croplands, and settlements so that fires can no longer spread unhindered across the landscape. Recreating natural fire regimes has also been criticized for not including aboriginal influences on fire regimes which may have significantly influenced landscape patterns over millennia. A variation on natural burning policies is therefore to recreate aboriginal burning practices to try to reproduce preagricultural landscapes. Implementation of these policies is constrained by lack of biological knowledge, techniques for fire management, and safety-related considerations.

In practice, fire management increasingly involves a mix of fire suppression, prescription burns, and controlled natural fires depending on safety concerns and conservation objectives in different parts of the landscape.

Fire and Global Change

Effects of Global Change on Fire Ecology

The impact of global change on fire regimes is likely to be quite different depending on whether fire activity is limited primarily by weather or by sufficient fuel to burn. Weather-dependent fires are characteristic of woody ecosystems such as conifer forests, eucalypt woodlands, and flammable shrublands. Fires in these ecosystems are most severe and burn the largest areas following long hot dry periods. In contrast, fires in grass-fueled ecosystems burn most extensively after unusually wet years. The implication is that climate change will have very different effects depending on the nature of the ecosystem. In the boreal forests of Canada, burned area has increased over the past half century. This is attributed to global warming with warmer temperatures generating larger fires (Flannigan *et al.*, 2009). Global warming in the circumboreal region is predicted to cause a doubling of area burned along with a 50% increase in fire occurrence by the end of this century. Similar dire predictions have been made for the conifer forests of Yellowstone with such frequent and severe fires that the forests may be eliminated and replaced by low biomass ecosystems by midcentury. Global warming might also increase the frequency of high severity fires entering tropical forests from adjacent savannas. However savannas, which currently account for more than half the world's annual burnt area, are also influenced by other global change factors that might cause large reductions in fire activity. Climates are predicted to be drier for much of Africa resulting in less grass biomass to fuel fires. In addition, increasing atmospheric CO_2 may favor the spread of less flammable grasses than those that currently dominate these highly flammable ecosystems. Increasing CO_2 is also thought to be contributing to increased woody cover in savannas, and the expansion of forest into savannas despite frequent fires. Increasing human population densities are leading to fragmentation of fuels and reduction of burnt area through increased livestock densities, land conversion to crops, expanded road networks, and growing urbanization. The intrinsic properties of plants that fuel fire are also being altered by invasion of exotic species. Introduced grasses have promoted increased fire activity in many ecosystems but shrubs, such as *L. camara* and *Chromolaena odorata*, have reduced fire activity in others. Thus the future of fires, and of savannas, is very uncertain. Uncertainties in global change predictions for fire regimes are being addressed by the development of physically based fire spread models coupled with physiologically based global vegetation models that incorporate both climate and CO_2 effects on future vegetation. Ultimately, however, the challenge is to devise innovative fire management options that reduce risks to lives and property while maintaining landcover that meets diverse human needs.

Fire as a Source of Greenhouse Gases

Vegetation fires contribute significantly to greenhouse gas emissions thereby influencing global climate change. Satellite data have greatly improved estimates of global fire emissions but there are still many uncertainties, not least of which is interannual variability in fire activity. Global fire emissions

from 1997 to 2009 averaged 2.0 Pg C year^{-1} compared with ~7.2 Pg C year^{-1} from fossil fuel combustion (van der Werf *et al.*, 2010). Major contributions to global fire emissions are from Africa (52%), South America (15%), Equatorial Asia (10%), the boreal region (8%), and Australia (7%). Savannas and grasslands contributed 60% of C emissions but most of these emissions would be compensated by rapid postburn regrowth. Annual emissions ranged from lows of 1.5 Pg C to a high of 2.8 Pg C in 1998, following an El Nino, with major fires in Indonesia. About a quarter of fire carbon emissions (0.5 Pg C year^{-1}) were attributed to deforestation, burning of tropical peatlands, and degradation in the period 1997–2009 and are probably a net source of CO_2 to the atmosphere. Unlike savanna fires, deforestation fires also emit large quantities of reduced trace gases like CO and CH4. Deforestation, degradation and tropical peat fires together account for about half of global CH4 (methane) emissions. Boreal peatlands are a major C pool storing as much as 270–370 Pg C. Increased fires, and a change from surface to ground fires, in boreal forests could potentially release enormous quantities of C into the atmosphere. The increasing frequency of fire in humid tropical forests is also of concern. Tropical forests are estimated to store one-fifth of the world's carbon. If burning of tropical forests continues, very large areas could be converted to flammable secondary scrub or grassland releasing this carbon into the atmosphere.

Fire is also an important source of aerosols. Aerosols decrease regional and global irradiation through backscattering of incoming solar radiation. Smoke aerosols can also increase or decrease cloud cover in complex and nonlinear ways that are not yet adequately quantified. Fire can also influence radiative forcing by altering the albedo. Black soot immediately after a burn heats the surface by reducing albedo. However, fire-induced reduction in tree cover can cause cooling by extending snow cover in boreal forests or replacing dark forests with more reflective vegetation (e.g., savanna grasses) elsewhere. The net effects of changing fire regimes on global warming are complex and uncertain. Public concern regarding atmospheric impacts of burning is leading to public pressure regarding the use of fire for conservation purposes. This may have positive effects if it leads to adoption of logging practices that reduce fire hazard in humid forests. Public pressure to suppress fires could have negative impacts on naturally flammable ecosystems and their fire-dependent species.

Appendix

List of Courses

1. Fire Ecology
2. World Vegetation or Global Biogeography
3. Conservation Biology
4. Global Change Biology

See also: Carbon Cycle. Climate Change and Ecology, Synergism of. Disturbance, Mechanisms of. Greenhouse Effect

References

Baker WL (2009) *Fire Ecology in Rocky Mountain Landscapes*. San Francisco: Island Press.

Bowman DMJS, Balch JK, Artaxo P, *et al.* (2009) Fire in the earth system. *Science* 324: 481–484.

Bradstock RA, Gill AM, and Williams RJ (eds.) (2012) *Flammable Australia: Fire Regimes, Biodiversity and Ecosystems in a Changing World*, 2nd edn. Canberra: CSIRO Publishing.

Cochrane M (ed.) (2009) *Tropical Fire Ecology: Climate Change, Land Use and Ecosystem Dynamics*. Berlin: Springer Verlag.

Daniau A-L, d'Errico F, and Sánchez Goñi MF (2010) Testing the hypothesis of fire use for ecosystem management by Neanderthal and Upper Palaeolithic modern human populations. *PLoS ONE* 5(2): e9157.http://dx.doi.org/10.1371/journal.pone.0009157.

Flannigan M, Stocks B, Turetsky M, and Wotton M (2009) Impacts of climate change on fire activity and fire management in the circumboreal forest. *Global Change Biology* 15: 549–560.

Giglio L, Randerson JT, van der Werf GR, *et al.* (2010) Assessing variability and long-term trends in burned area by merging multiple satellite fire products. *Biogeosciences* 7: 1171–1186.

Johnson EA and Miyanishi K (eds.) (2007) *Plant Disturbance Ecology: The Process and the Response*. Amsterdam, the Netherlands: Elsevier/Academic Press.

Keeley JE, Bond WJ, Bradstock RA, Pausas JG, and Rundel PW (2012) *Fire in Mediterranean Ecosystems: Ecology, Evolution and Management*. Cambridge: Cambridge University Press.

Marlon JR, Carcaillet C, Gavin DG, *et al.* (2008) A climate-driven decline in global biomass burning during the past two millennia. *Nature Geosciences* 1: 697–702.

Pyne SJ (2001) *Fire: A Brief History*. Seattle, WA: University of Washington Press.

van der Werf GR, Randerson JT, Giglio L, *et al.* (2010) Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics Discussion* 10: 16,153–16,230.

Fish Conservation

Carl Safina and Alan Duckworth, Blue Ocean Institute, Cold Spring Harbor, NY, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Carl Safina, volume 2, pp 783–799, © 2001, Elsevier Inc.

Glossary

Aquaculture Commercial farming of aquatic organisms, including seaweeds, raising captive-bred fish, and raising wild-born fish in captivity. Mariculture refers specifically to marine (saltwater) aquaculture.

Ballast Water taken in by ships to balance them after they have unloaded their cargo.

Bycatch Any living thing caught unintentionally in fishing gear; sometimes called bykill because so many such creatures are discarded dead.

Depletion In fisheries, reduction in population levels low enough to reduce or threaten future productivity.

Fishery A collective effort to gather, collect, or catch wild aquatic wildlife or plants for recreational or commercial purposes. Fisheries extract large numbers of wild fish, sea urchins, corals, seaweeds, shrimp, snails, clams, scallops, squids, turtles, whales, and other creatures.

High seas Parts of the ocean outside national boundaries, usually beyond 200 miles of any nation's coast.

Keystone species Species whose removal causes a chain or cascade of ecological effects among other species.

Marine reserves Designated areas where no fishing, mining, or other consumptive use is allowed, usually for purposes of replenishing nearby fishing grounds or maintaining normal evolution, growth, and fecundity.

Overfishing Extracting marine organisms faster than they can reproduce.

Pollution Introduction of substances in quantities that are threatening to living resources, biological processes, and human health and activities.

Outline

Overfishing and habitat degradation have driven many fish populations to historic lows. Poor fisheries management, increasing human population pressures, and habitat deterioration from several factors have caused this situation. But many fish are resilient and can recover within a decade or two if given viable habitat and a respite from overfishing.

Introduction

An ocean of water covers more than 70% of the Earth, and travelers from another world might more logically assume that this planet would be named Ocean. Ninety-seven percent of Earth's water is in the ocean (2% is locked up as ice, and 1% is in surface freshwater or groundwater). Moreover, 99% of the living space on the planet in which life can exist – the biosphere – is in the sea. The basis of most life in the sea are the plant-like single-cells called phytoplankton that create food from sunlight (through photosynthesis) and drift in the upper 1% of the ocean's volume. Plants attached to the bottom can only live in shallow coastal areas where they can get enough sunlight. Most life in the other 99% of the sea relies on food coming from that thin upper layer. Of 36 living animal phyla (the category that reflects the different basic body plans of living things), 34 are found in the ocean, 16 exclusively so. This article focuses on the world's ocean fishes. (For a synopsis of biodiversity issues facing freshwater fish, see **Box 1**).

The World Ocean Fishing Situation

Many factors affect life in the sea, and though climate warming and changes in ocean acidity due to fossil fuel combustion causes deep and fundamental change, fishing has caused the largest and most noticeable changes in ocean wildlife populations to date. The United Nations' Food and Agriculture Organization has stated that modern fishing "is globally nonsustainable."

During the last century, annual landings of wild ocean fish increased 25-fold, from 3 million metric tons to a plateau of over 80 million metric tons (FAO, 2010). In the 1950s and 1960s, fishing technologies exploded as fishing fleets adapted war technologies such as radar, sonar, and LOng RAnge Navigation (LORAN) to peaceful efforts of food gathering. Radar allowed boats to keep fishing in total fog, and sonar could find schools of fish deep beneath the sea's surface. LORAN and, later, Global Positioning System (GPS) turned the trackless sea into a grid and allowed boats to find and return to precise spots where fish gather. Satellite-generated maps faxed directly to boats in mid-ocean now track movements of water temperature fronts, showing where to find the fish (Safina, 1995).

Overfishing, scarcely recognized in the early 1900s, is now prevalent in most major fishing areas. The proportion of overfished populations is increasing, from 10% in the early 1970s to over 30% today (FAO, 2010). Not all of the catch is used for food. About a third of the world's catch becomes fertilizers, animal feeds, and industrial products, and about 7 million tons each year is unwanted and shoveled overboard.

World fish landings peaked in the late 1990s, but many fisheries reached the limit decades earlier. Indeed, most regions in the Atlantic, Mediterranean, Pacific, and even Antarctic have declining catches. In some regions where catches peaked as long ago as the early 1970s, catches have declined by half. Declines in some individual fisheries have been catastrophic. Newfoundland cod, which supported one of the world's largest fisheries for over 400 years, crashed 99% between the early 1980s and early 1990s. Many other fish

Box 1 Synopsis of freshwater biodiversity issues

Worldwide, lakes and rivers contain at least 10,000 fish species, roughly 40% of known fish species on Earth, and almost 20% of all vertebrates (freshwaters support almost one-quarter of the planet's known biodiversity, in only 0.01% of the planet's water; however, the oceans hold vastly larger populations).

Lakes are isolated habitats, leading to a high rate of evolution of species of fish and other animals. For instance, three-quarters of the 2000 plant and animal species in Russia's Lake Baikal are found nowhere else. The lakes of Africa's Rift Valley have produced explosive speciation – 99% of the 500 cichlid fish in Lake Malawi, for instance, live nowhere else. And Lake Tanganyika, the least-species-rich lake in the Rift Valley lake chain, has 25% more species of freshwater fishes than all of Europe.

Freshwaters are being degraded and species eliminated at a rate probably comparable to that occurring in tropical rain forests. Habitat loss and introduced species are the two greatest problems for freshwater biodiversity. Introduced predatory fish have already wiped out nearly 70% of Lake Victoria's cichlid fishes, and greatly reduced total biomass and productivity in the lake. In the USA, half the rivers and streams are significantly polluted, and 98% of US rivers outside Alaska are partially blocked by dams. Consequently, 40% of fish species and almost 70% of mussel species in US freshwaters are endangered or have become extinct, in contrast with only 15% of the mammals and birds in the USA.

Of the world's estimated 10,000 freshwater fish species, about 20% are in serious decline or have gone extinct. Ecologist Norman Myers has called the freshwater fish situation "the greatest extinction spasm of vertebrates in recent times." Climate Change will further negatively impact freshwater fish, and it's likely that many species will be unable to adapt or disperse to favorable environmental conditions.

populations have declined by 80% or more (Casey and Myers, 1998). Some major fishing grounds are now closed in hopes that the fish will recover. Most important ocean wildlife used as seafood are considered fully exploited or overfished (FAO, 2010), and no major untapped fishery resources remain in the world. Humanity's former faith in the sea's inexhaustibility was wrong.

Social and Economic Concerns

Worldwide, 180 million people depend on fishing and aquaculture for their livelihoods (FAO, 2010). Marine fishing supplies 16% of all animal protein consumed worldwide. Annual global fish consumption increased from 10 kg per person in the 1960s to over 17 kg today.

In the last few years, many thousands of fishers around the world have experienced severe decline or loss of their source of income. Depletion, caused primarily by industrial overfishing and exacerbated by coastal habitat degradation, threatens tens of millions of jobs as well as major food sources. As preferred species are fished out (Pauly *et al.*, 1998), and less-preferred fish are targeted, prices rise for these less-preferred fish. Price increases eventually remove the species from the tables of the poor. Distant markets and increasingly globalized trade networks allow and encourage regional overfishing or irresponsible aquaculture that exhausts local resources. This results in serial depletion and habitat degradation along faraway shores. For example: Bluefin tuna are depleted in the Atlantic and Mediterranean, and off Australia and New Zealand, because of the demand in Tokyo, where they have become the world's most expensive sushi (Safina and Klinger, 2008); appetites in the Northern Hemisphere support destructive shrimp farms in the tropics; and shark populations are declining in many places around the world because of demand for their fins, primarily in China.

Management Problems

In most of the world, intense political lobbying by the fishing industry usually causes fishery managers to disregard scientific assessments and scientific advice to bring catches within sustainable limits.

While no group of fishes is exempt, Atlantic bottomfishes such as cod, flounders, and Atlantic halibut have been among the worst managed. Various species of groupers and wild salmon are depleted throughout most of their ranges. Overfishing has driven many large fishes in the Atlantic to all-time lows. The Atlantic tuna commission's mismanagement, labeled an "international disgrace" by an independent review (ICCAT, 2009), has resulted in systematic overfishing of many tuna and marlin populations. The world's giant fishes – the tunas, billfishes (like swordfish and marlins), and sharks, among the most magnificent of the ocean's wildlife – are also among the most mistreated. In much of the world, management in these fisheries is lacking or ineffective. Sharks, subject to intense overfishing in many regions, are easily depleted because they are generally slow growing with low reproductive rates. Fishing annually kills tens of millions of sharks worldwide, mostly just for their fins, which are prized as a thickener in "shark fin soup." Only 7 of 20 major shark-fishing countries have management and research programs.

In many other cases, poor monitoring or outright cheating enlarges the gap between fishery science and fishing activities. In one 7-year period, foreign fleets fishing the Grand Banks removed sixteen times the quotas that had been set for cod, flounders, and redfish by the Northwest Atlantic Fisheries Organization. A seized Spanish boat kept two log books – one showing actual-catches and one for reporting to the authorities.

Few countries are successful in fisheries management whereas most are hardly trying. International cooperation on migratory populations and fishing outside national waters – on the high seas – is particularly difficult, and in some cases a country unhappy with the restrictions of an agreement has simply ignored them outright or through lack of enforcement.

Fish Fights: International Fishing Conflicts

Fish conflicts, many of them violent, have erupted around the world. Norway, Iceland, the United States, Canada, Indonesia, Taiwan, Nicaragua, Russia, China, the Philippines, Japan, France, Britain, Spain, and many others have been involved in international fishing disputes. In disputed waters off the Philippines, Chinese vessels were seized and scores of crew members were jailed. In the Galápagos, troops were deployed after

fishers seized a biological station that was trying to stop rampant illegal fishing inside a marine reserve. In France, fishers protesting fish imports rioted, causing $4.5 million in damage. NATO allies have pointed guns at each other in disputes over cod, and a Russian Navy boat fatally shot Chinese poachers. Hundreds of Indonesian fishermen have been imprisoned for fishing illegally in Australian waters. More trouble is likely as hungry boats scour the oceans for dwindling resources.

Fish as Commodities/Fish as Wildlife

Marine creatures are wild animals still hunted commercially on a large scale. And we usually treat fish as mere commodities, forgetting that fish are wildlife. For instance, we speak of a population of fish as "stock," as in "the New England groundfish stock." We would not speak of the "Serengeti mammal stock" or the neighborhood "woodpecker stock." By preventing us from seeing fish as wildlife in natural ecological communities, the language used helps facilitate their overexploitation.

Probably the most abused word throughout fisheries is harvest. Until recently, humanity has distinguished between hunting and gathering wild things that we do not grow, and agriculture, in which we raise and harvest crops that we do grow. Nowadays, many people try to blur this distinction by using harvest to describe hunting and gathering operations as a way to put a better face on taking large quantities of individuals from wild populations, usually for profit. We usually do not say we have harvested agricultural animals like cows, pigs, and chickens when we slaughter them for meat, even though we raise them for those purposes. So why should we say that we harvest wild animals like cod or sharks? Industry people even speak of harvesting whales to avoid using the obviously correct word, hunting, for the pursuing and killing of wild mammals, in this case the largest animals that have ever lived. When a wild fish or whale or a naturally growing tree is simply taken from a wild community, harvest is the wrong word. Appropriate words include: catch, fish for, take, extract, land, gather, collect (e.g., oysters), boat (i.e., they boated many salmon), cut, or hunt.

Intensity and Limits of Fisheries

Humans exert tremendous pressure on the seas. The coasts, where nutrients continually enter the marine environment primarily from the land, are the most productive parts of the oceans by far. The coastal areas, representing only 10% of the oceans' total area, produce 90% of the world catch (Pauly and Christensen, 1995).

Many industrialized fishing ships deploy large-scale fishing gear such as longlines that are 50 miles long with thousands of baited hooks, drift nets 40 miles long (now subject to a United Nations ban), and bag-shaped trawl nets sometimes large enough to engulf 12 Boeing 747 jetliners.

Since 1970, the world's industrialized fishing fleet increased at twice the rate of the catch, doubling in number and tonnage. The fleets of the world, numbering 4.3 million vessels (FAO, 2010), now wield twice the fishing power needed to catch what the oceans can produce. If the fleets had not grown at all since 1970, they would be able to catch the same number of fish. Fishing has either reached or exceeded natural limits virtually everywhere in the ocean.

Kinds of Overfishing

Overfishing is the main reason fish populations have declined. Overfishing means catching marine creatures faster than they can reproduce. The global proportion of overfished fisheries has steadily increased, from 10% in 1974 to 32% in 2008. "Fisheries" extract large numbers of wild fish, shrimp, snails, clams, scallops, squids, sea urchins, corals, seaweeds, turtles, whales, and other kinds of creatures. Different kinds of overfishing have been identified (see **Box 2**), and various factors that can influence a species' vulnerability to overfishing is seen **Box 3**.

Overfishing causes: (1) massive depletion of many species; (2) loss of breeders, thus fewer young produced and increased risk of reproductive failure in times of poor environmental conditions (e.g., unusual ocean temperatures); (3) declines in average sizes of fish and other marine creatures; (4) loss of genetic diversity; (5) genetic change toward less desirable characteristics like smaller size potential; (6) disruption of natural communities; and (7) disruption of human communities. Simultaneous overfishing of many species leads to functional loss of species or species groups. Keystone species are those whose removal causes a chain or cascade of ecological effects. For example, removal of algae-eating fishes can cause coral reefs to be killed by algae overgrowth, a large problem in many tropical regions.

Until recently, one of the main assumptions in fishery management was that the number of young produced is generally not related to the number of breeders, and many practicing managers who were taught this in school still believe it. The reasoning was that environmental factors cause a different percentage of young to survive from year to year, regardless of the number of eggs laid. But because any surviving fraction (say, 1%) of a large number of eggs is a higher

Box 2 Different kinds of overfishing affect many species

"Growth overfishing" is when too many fish are caught while very small. In species like groupers that change sex with age, a critical shortage of one sex can result.
"Recruitment overfishing" is reproductive failure due to depletion of breeders.
"Demographic overfishing" turns a population with many age groups into a population with only one or two significant age classes doing most of the breeding, making the population vulnerable to years when natural fluctuations cause poor survival of young.
"Genetic overfishing" is when intense fishing changes a population's gene pool.
"Serial overfishing" is the depletion of one targeted species after another.
"Ecosystem overfishing" causes great changes in species composition and functional loss of key species, and can result in long-term community changes.
Overfishing related directly to the increasing human population is sometimes called "Malthusian overfishing."

Box 3 Consideration of these factors can help us assess the vulnerability of a species to overfishing:

1. Inherent vulnerability: Does the species have low growth rate, low fecundity, high catchability (especially while immature), vulnerable behavior (e.g., spawns in groups), vulnerability to environmental changes, increasing vulnerability due to changes in fishing technology, or a poorly understood life cycle?
2. Human-related environmental risk: Does the species suffer destruction of key habitats, widespread effects of pollution, or conflicts with introduced species?
3. Is the population large or small, lightly or heavily exploited?
4. Is there management? Does it involve long-term conservation and sustainability as a goal? Does it benefit from independent and objective scientific advice? Does it have effective mechanisms for monitoring and enforcement and data collection?

number of survivors than the same percentage of a small number of eggs, the assumption that number of young are not related to number of parents is illogical. In 1996, Ransom Myers and Nicholas Barrowman proved this assumption wrong (Myers and Barrowman, 1996). They studied nearly 400 data sets from different species and asked the following questions: (1) Are the largest groups of young fish entering a fishery produced by the highest populations of spawners? (2) Are the smallest groups of young fish entering a fishery produced by the lowest populations of spawners? (3) Are above-average groups of young fish entering a fishery produced by above-average populations of spawners? Their findings: The answer to all three questions is almost always 'yes'.

When fishing is reduced, many depleted populations can recover. When fishing ceased in the North Sea during the world wars, fish populations increased significantly. On the East Coast of North America, a spectacular recovery of striped bass followed strict limits on fishing. Subsequent limits on fishing for several other species have also resulted in increases in those populations. Protecting fish until they reach average spawning age is one of the simplest, fastest, most effective ways to allow recovery.

Bycatch

Every fishery catches unintended, unwanted creatures, called bycatch. Bycatch includes nontarget and juvenile fishes, seabirds, marine mammals, sea turtles, and any other creature that the fishers are not trying to catch. Many die. This source of mortality currently threatens several species of dolphins and albatrosses with extinction.

By weight, about 8% of the animals caught worldwide are discarded (FAO, 2010). By number, a much higher percentage are thrown away, because usually it is smaller fish in the catch that are discarded. In many fisheries, bycatch exceeds target catch. Shrimp trawling incurs more bycatch than any other fishery. Dead discarded catch can outnumber shrimp catch by 15 to 1. Shrimp trawls are the largest human source of mortality to adult sea turtles. Gulf of Mexico shrimp fishery annually kills and discards 25–45 million juvenile red snappers, about 10 times the number of fish caught in directed commercial and recreational red snapper fisheries and several thousand tons of juvenile sharks. Total discard weight of all animals from the Gulf of Mexico shrimp fishery is almost 500,000 tons. But waste is not the only issue. Using everything caught would not fix the biological effects of killing millions of young fishes, as well as breeding-age seabirds, turtles, dolphins, and other mammals.

Lost fishing gear also catches, kills, and wastes sea life (this is called ghostfishing). Gill nets are frequently lost. For instance, the drift gill net (drift net) fishery of the North Pacific, which set approximately 30,000 miles of netting per night before the United Nations banned it, lost roughly 6000 miles of netting annually. A Norwegian study concluded that lost gill nets continued killing fish for at least 7 years; a New England study found them still catching fish 3 years after they were lost. Lost fishing gear frequently tangles seabirds and marine mammals even in areas as seemingly remote as the Antarctic (see **Boxes 4** and **5**).

Subsidies and Economics

Because there are too many boats (4.3 million in 2008), fisheries are said to be globally overcapitalized, and many are unprofitable. Worldwide, to catch $94 billion dollars' worth of fish, governments provide subsidies conservatively estimated at US $30–34 billion per year (Myers, 1997; FAO, 2010). Such massive subsidies arise from governments' efforts to preserve employment. Subsidies include fuel tax exemptions, price controls, low-interest loans, free access, and outright grants for gear or infrastructure.

Subsidies are responsible for building more fishing power than the resources can support, which increases political and social pressure for continued high catches. But high subsidies can only lead to greater and greater economic distress and further depletion. The United Nations Environment Program estimates that overfishing and overcapitalization results in economic loss of $50 billion each year. The ecologist Norman

Box 4 Albatrocities

Bycatch from commercial fishing operations is the most serious threat facing albatross populations. Each year over one hundred million hooks are set on longlines in the Southern Ocean. A longline can bear as many as 3000 baited hooks. Albatrosses often congregate in the same areas where boats fish. As the line is being played out from the moving vessel, albatrosses sometimes pick up the bait just before the weight of the sinking line pulls them underwater. This results in the death of tens of thousands of albatrosses and other seabirds each year (Brothers *et al.*, 2010). According to the World Conservation Union (IUCN), 17 of the 21 Albatross species are now threatened with extinction.

Box 5 Dolphin-safe's downside

After several Pacific dolphin populations fell 50–80% due to drowning in tuna nets, the US adopted a "dolphin-safe policy" in 1990, which barred the import of tuna caught by setting nets around dolphins (yellowfin tuna and dolphin swim together in the eastern tropical Pacific). Consequently, dolphin kills dropped 99%, but other bycatch increased. One study found that the average bycatch from 1000 tuna net sets around dolphin herds includes 500 dolphins, 52 billfish, 10 sea turtles, and no sharks. In contrast, average bycatch from 1000 sets around floating objects (where tuna also gather), the preferred fishing method in the wake of the "dolphin-safe" regulations, includes only 2 dolphin, but also 654 billfish, 102 sea turtles, and 13,958 sharks. In addition, large numbers of the tuna caught under floating objects are juveniles.

Myers (1997) concludes, "Subsidies are far and away the principal cause of overfishing."

Fishery depletion currently costs the US economy $1 billion and tens of thousands of jobs each year. Rebuilding depleted populations of just four fish species in the US would generate an additional $570 million per year. Many studies conclude that what are needed are reductions in fishing power – fewer and smaller-scale boats, each employing more people per unit of horsepower – if we are to let the fish populations rebuild.

Other Factors Affecting Fish Conservation

As discussed so far, fishing is currently the major agent of change in the oceans, but a fuller compilation of the major threats to marine biodiversity includes:

- above-discussed fishing operations, which cause depletion, catch large numbers of nontarget species, degrade seafloor and coral reef habitats, and cause major changes in living communities throughout the world's oceans;
- distant markets, which exert tremendous pull on resources from around the world, cause economic, social, and ecological effects of local depletion, then switch to new sources;
- pollution from those chemicals that cause toxic effects and hormonal disruption;
- pollution from excessive nutrients, which increase the frequency and severity of harmful algal blooms such as red tides and brown tides, and cause oxygen-depleted dead zones;
- physical habitat alterations, including development, wetland alteration, and on-land deforestation, which sends clouds of sediments into rivers and coral reefs;
- introductions of species to areas where they are not native, causing competition, predation, and spread of disease;
- aquaculture or "fish farming," which too often entails extensive habitat damage, pollution, and introductions of alien species and diseases;
- debris, which tangles, traps, or is ingested by large numbers of marine fishes, mammals, birds, and other creatures;
- dams in rivers, which prevent fishes such as certain salmon, sturgeons, river herrings, and others from reaching spawning areas, or kill their young on their way to the sea;
- atmospheric warming from increasingly concentrated carbon dioxide and other greenhouse gases resulting mainly from burning fossil fuels, which stresses corals, lowers overall ocean productivity, melts polar ice, threatens to raise sea level and inundate coastal wetlands and low islands, and may change ocean current patterns, resulting in changes in abundance and distributions of marine wildlife;
- ocean acidification, resulting from the oceans absorbing high levels of carbon dioxide, causing seawater pH to drop affecting all marine life; and
- ozone thinning, which lowers ocean productivity by killing significant amounts of plankton.

We should bear in mind that gravity takes things from land to sea. Almost everything that enters a storm drain or river goes into the ocean, from used oil to litter. The fewer chemicals people use on their lawns and farms, the less of it we will swim in and eat in our seafood. More oil enters the ocean from urban runoff, including automobile leaks, than from large tanker spills. About three-quarters of ocean pollution comes from the land.

Human Population Growth

By 2050, 9 billion people will be joining you for dinner, and the seas will feel their impact directly. Each year the world adds 70 million new human appetites. To keep pace for just the next 20 years, we will have to find 27 million more tons of seafood (FAO, 2010). A third of the world catch is now used for animal feed and fertilizer, but even if it were simply eaten by people and not fed to livestock, it would maintain present consumption levels per person for only 20 years. Widespread recoveries of wild fish in the face of increasing demand seem unlikely. Improved conservation would have difficulty keeping pace with the population growth. We have likely hit the limit of the oceans' natural capacity to feed us.

Coastal habitat disappears in proportion to population growth, and growth in coastal areas is four times the US national average. Seventy percent of major US cities are located along the coasts, and about half of the human race lives within 120 miles of the sea. Diminished habitat means diminished productivity.

Habitat Issues

Coastal habitat losses are very costly to fisheries. Human alteration of aquatic habitats is a major cause of fishes' declines – the major reason in freshwater, and the second-largest factor (after fishing) in estuarine and continental shelf habitats. Marine habitat is physically altered by development, aquaculture, mining, and fishing activities. Development often causes: extensive filling and diking of critical wetland breeding and nursery areas; dredging to deepen shallow estuaries for shipping; shoreline stabilization; and dams that alter normal

patterns of sedimentation, erosion, and water flow. Dams also block migrations to and from spawning areas, and divert to agriculture, industries, and homes, the freshwater flows needed by wetland and estuarine creatures. Aquaculture often destroys local wetlands and mangrove areas. One-quarter of the world's mangrove forests (about 50,000 square kilometers), for example, have been cleared to free land for food production, often for shrimp culture. Loss of this important nursery habitat has eliminated an estimated 5 million tons of annual wild fish catch, which is about 6% of total world landings. Mining removes coral and minerals, often adding toxic chemicals and sediments that destroy seafloor creatures. Agriculture and deforestation put large amounts of sediments into coastal waters, smothering spawning beds, seagrass flats, and coral reefs. Trawl nets and dredges dragged for fish, clams, and other species extensively damage seafloor structures and living communities. To catch fish from coral reefs, where nets are impractical, fishers often use explosives or cyanide that fragment or poison corals. In the Pacific Northwest of the United States and Canada, intensive deforestation and water diversion destroyed thousands of river miles of salmon spawning and nursery habitat, eliminating hundreds of salmon runs and thousands of related jobs.

Coral reefs support Earth's most diverse fish communities. Coral mining (for construction), fishing with explosives and poisons, destruction from boat anchors and ships running aground, and silt from farms and clear-cut logging cause major damage to coral-dependent communities. Sediment flowing off deforested land is a major threat to coral reefs, killing corals by clogging them, blocking sunlight, and preventing larval settlement.

Introduced Species

Aquaculture, the pet trade, sport fishing, and shipping intentionally and accidentally introduce species to new parts of the world. When ships empty their cargo, they often rebalance by taking on water. In the water are larvae, eggs, living cells, and small animals. When the ships arrive and discharge this ballast water, these hitchhiking species – it is estimated that more than 3000 species are in motion in ships on any given day – may be introduced to a part of the world where they have never existed. Often they arrive without the predators, competitors, or pathogens that limit them at home. Sometimes, organisms are introduced to places where the native species cannot escape or compete with them.

Invasive species greatly alter ecological communities. In San Francisco Bay, a tiny Chinese clam displaced 95% of the bottom community, and now consumes so much plankton that native plankton-eating fishes and, in turn, predatory fish may be affected. In Eastern Europe's Black Sea, introduction of an American jellyfish reduced plankton abundance by 90% in less than a decade, causing small fishes like anchovies to crash. Pacific Lionfish, accidently released in the early 1990s, have exploded throughout the Caribbean, the Bahamas, and Florida Keys, devouring millions of small fish and changing the community structure of coral reefs. Reef fishes intentionally introduced into Hawaiian waters for sport fishing have affected the local fish communities. The list of examples is long, and future invasions are inevitable. In one study, five of 80 ships entering Australia carried toxic phytoplankton, which can produce red tides and cause shellfish to become poisonous. Even transporting salmon relatively short distances from one river system to the next for aquaculture or into hatcheries, as often happens, can move diseases into new areas, or expose fish to habitats where they are doomed by diseases for which they have no immunity.

Aquaculture

Many are counting on aquaculture – farming aquatic organisms – to fill the gap between nature's bounty and the hunger of people. Aquaculture is exploding, however. For example, shrimp farming increased 17-fold over the last 25 years, with global value now exceeding $9 billion annually. Aquaculture now produces almost 50% of humanity's fish. In freshwater, aquaculture exceeds wild catch production by 300% (FAO, 2010). In the oceans, it produces about 40% of shrimp and 80% of salmon in commerce.

Aquaculture does not solve all problems, and causes some new ones. Most saltwater fish farming causes environmental degradation, displaces local people reliant on wild fish and their natural habitats but provides fewer jobs. Aquacultural products are often expensive, and commonly exported to rich countries rather than used locally to ease hunger of the poor people. Crowded fish farming operations pollute by releasing pesticides, antibiotics, and oxygen-robbing feed and feces into semi-enclosed waterways. Shrimp farming has a record of boom and bust caused by pollution and disease outbreaks. In Asia, shrimp ponds rarely remain viable for more than 5–10 years before being abandoned because of contamination (Naylor *et al.*, 1998). Thousands of Atlantic salmon have escaped in the Pacific Northwest. The state of Washington listed the Atlantic salmon as a pollutant because of its potential to cause problems for the several native Pacific species of salmon. Atlantic salmon have been discovered breeding in the wild in British Columbia. Even in the Atlantic, escaped Atlantic farmed salmon interbreeding with wild salmon cause genetic degradation, because wild salmon are genetically finely adapted to characteristics in the specific rivers where they spawn. Moreover, ponds and other open systems attract wild fish-eating birds such as cormorants and herons, often bringing them in range of farmers' guns.

Most importantly, most aquaculture consumes vast quantities of wild fish as feed – and is thus a net loss of food from the oceans. Most believe that aquaculture contributes to the supply of human food, and this is true for the few herbivorous species. But fish are not cabbages; they do not grow on sunlight. Most fish and shrimp are carnivorous, and feeding them uses two to four times as much wild fish as the farms produce. Farming such species actually contributes to the ocean's depletion.

Aquaculture does not take pressure off the seas. Wild larvae of species that do not breed in captivity are sometimes intensively caught for pen-rearing. Increased shrimp and salmon farming has not lessened fishing intensity even for those species, and has brought trouble to wild shrimp and salmon through diseases and habitat destruction.

Aquaculture can be done responsibly but as usually done now it destroys naturally productive habitats that support ocean fisheries and wild communities, and introduces alien

species, parasites, and diseases that threaten local species' populations. To be responsible, aquaculture enterprises should not be sited where natural habitat is affected, their wastewater should be treated, and they should focus on plant-eating species, raised in closed systems to prevent escape.

Pollutants

Pollution is the introduction of substances in quantities that are threatening to living resources and human health and activities. Chemical pollution can cause spectacular mass mortalities or cause subtle changes in population composition, impaired sexual development and reproductive success, impaired growth rates, deteriorated seafood quality, tumors and other diseases, and outbreaks of harmful algal blooms such as red tides, *Pfiesteria*, or normally innocuous algae that overgrow corals or deplete waters of oxygen. Chemical pollutants tend to concentrate in surface waters, where larvae and eggs also concentrate. Chemicals enter the seas from sewage, industrial outfall, agricultural runoff, ocean dumping, aquaculture, accidental spills, and from the air (acid rain has significantly damaged freshwater fishes in North America and Europe). Major chemical pollutants include insecticides and herbicides, detergents, PCBs, elements such as chlorine and heavy metals, petroleum products (much of the gasoline/oil mixture run through 2-cycle outboard engines exits unburned), mining wastes, fuel ash, radioactive materials, and excessive nutrients from sewage, farm animals, and fertilizers.

When algal blooms caused by excessive nutrients use up those nutrients and die, their subsequent decomposition can rob the water of oxygen, suffocating marine creatures. The surprising fact that fertilizer or other nutrients in excessive amounts can kill aquatic organisms is sometimes called "the paradox of enrichment."

About three-quarters of marine pollution comes from the land, thus the people causing the pollution do not feel its effects directly or immediately, and the people feeling it (e.g., fishers) cannot directly affect its origin. Spreading from the mouth of the Mississippi River, nutrient runoff from farm land produces a dead zone greater than 20,000 square kilometers in some years, reducing oxygen concentrations to dangerous levels and affecting important shrimp and oyster fisheries. Over the long term, this is more a threat to the Gulf than the spectacular Deepwater Horizon oil blowout of 2010. New York City's harbor gets more than 27 billion gallons of storm-water, which includes raw sewage, every year. That is one reason why laws are necessary. In the US clean water legislation has helped keep some fisheries economically viable by controlling pollutants.

Global Atmospheric Change and Fish Conservation

Major atmospheric changes, like global warming and rising CO_2 levels, have significant implications on marine life. Earth's climate has changed dramatically as ice ages have come and gone, but these changes took many thousands of years, allowing life-forms long periods in which to adapt. Human-caused climate change is happening much faster, and some habitats and species will probably not change or move fast enough to survive.

The burning of fossil fuels has increased atmospheric levels of carbon dioxide, methane, and other heat-trapping gases, and most scientists agree that this is intensifying the greenhouse effect, the warming of Earth's atmosphere. Oceans slow the buildup of greenhouse gases by absorbing about a third of the carbon produced by burning, but substantial climate change is likely over the next few decades. The effects of such changes are not straightforward, but planetary warming is melting polar ice caps; causing population declines in ice-dependent krill, the shrimplike key prey species for Antarctic fishes, marine mammals, penguins, and other seabirds; causing changes in the distribution of many temperate species; and raising sea levels throughout the world. This rise is expected to cause some flooding and loss to coastal marshes, mangrove areas, low-lying islands, and their critical associated fish-nursery habitats, as well as to some cities. Global warming is directly affecting fish abundances and forcing latitudinal migrations of some populations. The Atlantic Cod population in the North Sea is moving northwards, while Cod populations in the Celtic and Irish Seas will likely disappear by 2100.

Warming is likely to alter ocean currents. Such changes may intensify the recurrent El Niño phenomenon. Other possible changes include alteration of currents such as the Gulf Stream and Kuroshio Current, and weakening of major upwellings off South America and Africa. Such changes would affect the production, distribution, and survival of fishes and other creatures in the world's oceans, and alter the fish availability among countries.

Changing sea temperatures will exert major influence on the survival and distribution of the oceanic populations. In parts of the California Current, planktonic animals declined 70% over the last 40 years, a decline many scientists think may be linked to global warming.

Our addiction to fossil fuels is also driving ocean pH to dangerous levels. Our oceans absorb 22 million tons of carbon dioxide (CO_2) daily, much from industrial processes, which when dissolved in seawater forms carbonic acid in a process called "ocean acidification." Ocean pH has dropped from a preindustrial value of 8.18 to 8.07 today, and is expected to reach 7.82 by 2100 (Orr *et al.*, 2005). pH is measured on a logarithmic scale, so a 0.1 decrease means our oceans are 30% more acidic today than 250 years ago. This is having a dramatic and wide ranging impact on fish populations and marine ecosystems. For some fish species, increasing acidity decreases larval survival and replenishment of reef fishes, interferes with their olfactory system so they do not flee predators, and reduces growth. Ocean acidification is reducing survival and growth of organisms low on the food chain, with potentially severe ripple-effects through the entire food web.

Some ecosystems may simply dissolve. Coral reefs, made from the calcium carbonate ($CaCO_3$) skeletons of corals and coralline algae, require waters super-saturated in aragonite (one form of $CaCO_3$) to exist and grow. While more than 98% of coral reefs were bathed in aragonite-saturated waters before the industrial revolution, ocean acidification has already put 40% of reefs into unfavorable conditions (Cao and Caldeira, 2008). If CO_2 levels increase as predicted, by 2050 aragonite levels will be too low to support the formation of coral reefs

anywhere. At this stage, coral reefs will stop growing, and start dissolving. All coral reef organisms including the 4000 species of fish that feed on and live in this fragile ecosystem will lose their habitat. Loss of corals reefs will directly affect about 100 countries and territories that rely on them for fisheries, tourism, and coastal protection.

Certain human-generated compounds such as chlorofluorocarbons rise into the stratosphere and destroy the ozone that shields Earth from the sun's ultraviolet radiation (UV). The biologically damaging UV is called UV-B, and it penetrates many meters below the sea surface. It has increased under the Antarctic ozone hole and elsewhere. Also, recent volcanic eruptions have reduced total atmospheric ozone by as much as 10%, leading to increased UV-B of up to 20% in latitudes as far from the poles as Florida and the Bahamas. This has caused significant damage to plankton (small drifting animals and plants, including fish eggs and fish larvae), corals, and bottom-living organisms. Controls on manufacture of chlorofluorocarbons have stabilized this situation and scientists predict recovery of ozone densities later in this century.

Marine Fish Conservation in a Biodiversity Context

Biodiversity is the diversity of living things, and can be thought of as occurring on three levels: genetic diversity, which is the genetic variability among members of the same species; species diversity, the variety of species found in a community or an ecosystem; and community diversity, the varieties of biological communities on Earth. We will use these categories to examine the fish conservation issues introduced above.

Marine Fish Biodiversity

The greater the diversity of habitats, and the more isolated habitats are from each other, the greater is the diversity of species adapted to those habitats. (Though freshwater covers only 1% of Earth and ocean fish populations are vastly larger, 40% of the world's fish species live in freshwaters because they are so isolated from each other that they evolve differently.) Because most ocean habitat diversity is coastal, so is most ocean fish diversity – there may be 10 times as many coastal fish species (roughly 13,000) as truly open-ocean fishes (about 1200 species).

Genetic Diversity within Species

Conservation of biodiversity is virtually never an official management goal in fisheries. (Indeed, if we wanted to try to exterminate the fish, we could hardly have gone about it differently than we have gone about fishing.)

Fishing can induce genetic changes through unintentional selective breeding because of intensive, nonrandom fishing mortality. In intentional artificial selection (as for farm animals) we choose and breed individuals with traits that we desire. But in fishing, we remove from the breeding population those individuals with the traits we desire.

Fishing activities can cause evolution (Sutherland, 1990). Evolution is a change in the relative frequency of genes. It need not result in progress or new species; it merely shapes species for better survival under prevailing, often changing, conditions (if change is extensive, new species can result). Fishing is so intense in so many places that many fishes have a higher chance of dying by being caught by people than by natural causes (for instance, in some populations 80–90% of adult cod or salmon are killed by fishing).

Fishing can alter inherited characteristics of a population in two ways: (1) by applying a selective pressure that kills individuals based on certain traits, such as size or age; and (2) by applying a random pressure so intense that the population is depleted low enough to lose genes from the pool. Note that fishing can do these things, and it does both of them sometimes, but not all fishing results in genetic changes; whether it does is determined by fishing intensity.

Genes related to size, growth rate, and age of sexual maturity are most likely to be affected by fishing (Sutherland, 1990). Growth overfishing prevents fish from reaching large size, thus it incidentally results in populations of fish genetically preprogrammed to mature and breed younger and at a smaller size than was normal in the absence of fishing pressure. Fish that reproduce pass their genes on. The genes of fish that get killed before breeding will begin dying out. Fish genetically programmed to breed at a slightly younger age than average are also likely to be fish programmed to be smaller as adults. Thus, intense fishing that allows few individuals to live longer will cause normally maturing fish to make a disproportionately small contribution to the next generation, and allow early-maturing fish to make a disproportionately large contribution. In a study modeling the contribution of three age groupings of cod to overall reproduction, the late-maturing subgroup made virtually no contribution to reproduction after only 7 years of intense fishing. This process likely operates on wild populations subject to intense fishing. So fishing can unintentionally breed a population of younger-reproducing, shorter-lived, smaller fish.

Intensive fisheries probably create conditions necessary for genetic change, but genetic changes can be difficult to confirm. Reduced average sizes and younger ages at first breeding do not necessarily reflect accompanying genetic change. They can result from changes in densities and competition for food, too. Whether genetic or not, examples abound of population-level effects of overfishing. In a rock fish (*Sebastes alutus*), lightly exploited populations had a modal age of 30 years, and 73% of the fish were over the age of 20; heavily exploited populations had a modal age of 12 years, with only 7% of the fish over the age 20.

Average sizes of many US Southeast species declined about 75% between the early 1970s and the mid-1980s: red porgies went from 2.6 pounds to 1.3 pounds, red snapper from 18.0 to 4.4 pounds, snowy grouper from 17.6 to 4.4 pounds, speckled hind from 19.1 to 6.6 pounds, scamp from 10.1 to 3.3 pounds, and gag grouper from 18.0 to 4.4 pounds. For species such as groupers, which change sex in older age, intense removal of larger individuals can dramatically affect sex ratios and mate supply, greatly lowering the population's reproduction. Jim Bohnsack of the National Marine Fisheries Service has said that if almost all the big ones are taken and the little ones left, "The result is a race of miniatures."

Overfishing reduced the California sardine population's age structure from five spawning age classes to only two, and

two consecutive years of poor oceanic conditions led to spawning failure and collapse. However, even when vast populations of naturally small, short-lived fishes such as sardines, anchovies, and herrings collapsed to one-thousandth of their former numbers, high enough numbers have remained to maintain genetic diversity. Yet in some animal populations, older or larger individuals are more likely to carry more diversity in the form of heterozygosity – differing forms of particular genes. Larger individuals often spend more time on breeding areas than younger ones, so fisheries targeting spawning groups, such as those for groupers, orange roughy, and many others, can subject the most genetically diverse individuals to the most intense mortality. This kind of genetic diversity loss is believed by some to be a common consequence of heavy exploitation, even without reducing populations to near-extinction levels.

Theory suggests that genetic change from fishing is likely to be common, and several studies have found convincing evidence. Perhaps the species best-studied before and after commercial exploitation is a long-lived fish called the orange roughy. Off New Zealand, only 6 years of heavy commercial exploitation reduced populations by 70% and significantly reduced genetic diversity within those populations, probably because older individuals in this very long-lived species were more heterozygous. In one North Sea cod population, intensive fishing reduced the chances of a young cod surviving to breeding age by 95%. Over several decades, average age of sexual maturation declined from about 10 to 7–8 years, apparently through genetic change. Some populations of flounder and haddock also show convincing evidence of genetic change. In a population of Atlantic salmon, average age at maturity had dropped after two decades of heavy fishing. In a study of chinook salmon, average spawning age fell by 2 years. Other studies on salmon showed other changes. In sum, some studies found no apparent changes attributable to fishing, some showed changes that were inconclusively genetic, and other studies showed genetic changes that were driven by intense fishing.

Species Diversity

Human activities significantly change species composition and abundance, and predator–prey and competitive relationships. Naturally evolved numerical and functional relationships among species are sometimes referred to as the community's ecological integrity.

One indication that many regions' ecological integrity has greatly diminished is the changing world catch. Since 1950, many fisheries, forced to work lower on the food chain as they deplete large fish, have shown a gradual transition from catching mostly large, long-lived, fish-eating bottomfish such as cod toward catching small, plankton-eating, open-ocean fishes like herring, as well as short-lived, low-on-the-food-chain invertebrates like squid (Pauly *et al.*, 1998). Daniel Pauly, the eminent ecologist who helped discover this trend, remarked, "If things go unchecked, we might end up with a marine junkyard dominated by plankton."

Various marine communities have been changed by overfishing. A common pattern – particularly acute in coral reef systems but also seen in kelp communities and elsewhere – is selective removal of the largest species first, then of large individuals of smaller species. People removed most large groupers and sea basses from many coral reefs and kelp forests, and manatees, dugongs, and sea turtles from seagrass communities, before scientists ever studied these habitats. Though no one really understands how those communities are evolved to function, they were certainly very different. Since Columbus landed in the Bahamas, for instance, Caribbean sea turtle populations have declined 99% (Jackson, 1997).

Depletion of fish populations changes patterns of abundance, distribution, and competition among plants, urchins, corals, sponges, tunicates, and other creatures. Alarming declines in North Pacific seabirds, Steller's sea lions, and other wildlife may be caused by heavy fishing. When fishing depleted Barents Sea herring and capelin, cod failed for lack of food. Squid sometimes increase following the collapse of their overfished predators. Off New England and maritime Canada, overfishing shifted communities dominated by cod, haddock, and flounder to domination by spiny dogfish (a small shark) and skates (a type of ray). But not even skates are safe. Fishing has driven the commonest skate in the northeast Atlantic, *Raja batis*, to local extinction in the Irish Sea.

Vulnerability of Marine Fishes to Extinction

Few fully marine creatures, and perhaps no fully marine fish species, have gone extinct because of human activities. Many marine fishes have large distributions and a greater chance for replacement by neighboring populations if they are locally wiped out. But the important thing is to prevent, not document, endangerment and extinctions. The time to save a species is when it is still common. Waiting until biodiversity declines, until populations falter, and until more species names get added to various endangered lists will only ensure three things: (1) more species will get into serious trouble, (2) the value to humanity from extremely important natural resources such as fishes will be seriously diminished, creating more of the economic, social, and nutritional problems already suffered by people in too many communities throughout the world, and (3) fixing the problems will be much harder, more expensive, and prone to failure.

Numerous fish have already been listed on the U.S. Endangered Species List, on the U.S. Overfished Species List, and under the world Convention on International Trade in Endangered Species (CITES) and the World Conservation Union's Red List. The Red List notes that hundreds of marine fish have suffered major and rapid population declines or significant local extinctions (extirpations). Most vulnerable are those that breed in freshwater but mature in salt water – the anadromous species – such as salmon and sturgeons, because they suffer overfishing plus the effects of habitat disruption in rivers that are heavily dammed, subject to intensive agricultural water withdrawals, or subject to massive siltation from destructive logging practices. Hundreds of North American salmon runs are already extinct, and the trends for sturgeons and totoaba (a large fish of the croaker family that is dependent on Colorado River flow into the Gulf of California) suggest that biological extinction may become a real possibility. Groupers are also vulnerable; they are generally slow-growing fishes whose spawning groups are easily targeted by

fishers. Long-lived, slowly reproducing species such as large sharks and sawfishes are particularly vulnerable, and many such species are now depleted and threatened. Several large sharks, including the great white, are now totally protected in some regions, an acknowledgment of their threatened status.

No doubt endangerment to fishes is underappreciated simply because their populations are not well studied. Most of the 300-plus species listed on the World Conservation Union's Red List got there simply because the London Zoological Society and World Wide Fund for Nature convened a workshop to compare known information about fish population changes against the Union's listing criteria. Similarly, the barndoor skate became endangered without anyone realizing it, simply because no one had analyzed decades of already-existing data.

Not just fish are affected. Wild abalone populations off California have been greatly reduced by fishing, and the white abalone is almost certain to go extinct in the wild because probably too few individuals are left to breed effectively. Who would have thought that fishing would be intense enough to drive a snail to extinction?

Kinds of Extinction

Overfishing and habitat degradation can lead to a four-step process of extinction, in which people suffer the major effects long before the last animals vanish. From social, economic, and genetic standpoints, extinction is a process rather than an event; fishing economies usually go extinct and gene pools lose diversity long before the last fish dies.

The first stage of the extinction process is depletion, wherein the fish population is reduced below its most productive level; fish are smaller and fewer than they could be, and less likely to spawn a strong year class of young fish. Consequently, the population's ability to support fishing is reduced, and profit margins decline. Other members of the ecosystem may experience food shortages or an unnatural relaxation of predation pressure or competition.

The next stage is ecological extinction, where the animal's population drops so low that the species no longer effectively functions as prey, predator, or competitor in the community. If it is an important keystone species affecting many others, entire marine communities may undergo profound shifts in numerical and functional relationships, and species less valued by people may increase. Animals that have become ecologically extinct – essentially removed – from the southern California kelp community, for example, include sea otters, black sea bass, white sea bass (though it has been recovering), large groupers, all the large kelp bass, sheephead, rays, flatfish, rock fish, lobsters, abalones, sea urchins, sea cucumbers, and others.

Another stage, commercial extinction, occurs when the species becomes too rare to be fished profitably. During this phase, fishery managers sometimes close large areas to fishing, as off New England and Newfoundland, or in the salmon fisheries off the Pacific US coast.

If commercial and ecological extinction are not reversed, total extinction may become a possibility. But by the time total extinction becomes an issue, all the other practical effects of the creature's disappearance from commerce and from the ecosystem have already been suffered by people. One cannot ignore the community effects of overfishing on the top predators – humans. Through dislocation, malnutrition in remote local villages, job loss, and loss of social identity, humans suffer the major effects of overfishing long before the animals themselves completely vanish.

To date, total extinctions are rare in the oceans, but this may not always remain so. We may unfortunately see more fish on endangered species lists in the future. If we do, it may not be too late to save them. But it will likely be too late to save the fishing jobs and coastal communities that once depended on them. A recovery vision should always include regaining the numerical and functional relationships that had naturally evolved among species.

Ecosystem Diversity and Fishes

Though the ocean may seem like one big bathtub, to the creatures in it and the people making a living from it, the sea is more like a mosaic of habitat types. Water circulation patterns can create discrete habitats because temperature alone can be a boundary for sea life. Currents also cause drifting seaweeds to accumulate in places like the Sargasso Sea, an area in the middle of the Atlantic characterized by vast mats of sargassum weed used as habitat by many creatures. The sediments of the seafloor vary in texture, from boulders to fine silt and clay, all creating different habitats. The area between high and low tides – the intertidal zone – likewise contains particular habitats for specially adapted creatures that live on rocks, in sand, or in tide pools. Along coasts, shallow estuaries, where freshwater from rivers mixes with salty seawater, have their own suite of habitats: seagrass flats, salt marshes, bays between outer beaches and mainlands, and tidal inlets. Tropical and subtropical shores often support mangrove forests growing in salt or brackish water. Near warm shorelines, coral reefs support the most spectacular diversity of fishes and other creatures in the sea.

The distribution pattern of these major habitat types is what we speak of with the term ecosystem diversity. Just as fishery management has generally ignored genetics, it has also generally ignored ecosystem considerations. Fishing activities, aquaculture, coastal development (for instance, damming rivers and filling wetlands), and atmospheric warming have resulted in changes in the distribution and functioning of habitats on which marine communities rely – that is, changes in ecosystem diversity.

Aquaculture has significantly altered coastal and estuarine habitats in many parts of the world. The conversion of mangrove habitats into fish and shrimp farms over extensive areas is a form of ecosystem depletion. Coastal development has also destroyed habitat in ecologically significant areas. For instance, about half of the wetlands in the US have been lost to practices like filling, which has led to a significant shift in the distribution of aquatic habitats. Atmospheric warming is both melting Antarctic ice and stressing coral reefs. This is causing change in these habitats on a global scale.

Coral reefs and continental shelves are two major kinds of habitats experiencing profound changes due to human activities. Small increases in temperatures in tropical seas cause coral bleaching, whereby corals lose the helpful algal cells living inside them, turn white, and die. Rates of coral bleaching also increase with decreasing pH, thus the combined effects of global warming and ocean acidification will

damage all reef corals to some degree. Widespread death of corals, of course, has potentially severe implications for coral-dependent fish and human communities.

Coral reefs are the oldest and largest structures built by living things. When scientists first surveyed them in the 1950s and 1960s, most coral reefs were poised on the edge of profound change in the composition and abundance of species and the functioning of coral reef communities. For example, Caribbean overfishing had already depleted herbivorous fishes to very low levels when algae-eating sea urchins crashed from a disease (Roberts, 1995). Without herbivores, algae overgrew corals, killing them over large areas, and altering the community. Conversely, heavy fishing off East Africa depleted fish that prey on coral-eating sea urchins. With their predators largely fished-out, the urchin populations exploded to densities that were 100 times the normal, and proceeded to inflict significant damage to reefs. The depletion of predator fishes has also been implicated in destructive outbreaks of the coral-eating sea star known as the crown-of-thorns. Such ecological changes affect large communities of organisms.

Fishing often kills corals directly when people use explosives to kill fish, or pound the coral with rocks tied to lines to scare fish toward nets, or use cyanide to stun fishes. All of these methods cause extensive destruction of coral reefs, through either breakage or, in the case of cyanide, death to corals by poisoning.

Coral reef systems are perhaps more sensitive to environmental change than most other marine ecosystems. For various reasons, species of algae may overtake a damaged reef preventing corals from recovering or recolonizing. Many coral reefs are now degraded by human activities, their fish communities modified, and the living corals replaced by algae growing on dead corals. The Caribbean region has lost 80% of its live coral since the 1970s because of overfishing, diseases, algal overgrowth, sewage overfertilization, and smothering silt running off deforested land areas. Some scientists believe these changes may be very long-lasting.

Aside from coral reef areas, continental shelves nearly everywhere are affected by bottom trawling – the towing of large bag-shaped nets or shellfish dredges over the seafloor. Trawling is the fishing method that takes half of the world catch. The vast majority of the world's seabed is encrusted or honeycombed by living things, and trawling causes enough incidental damage of nontarget invertebrates, plants, shellfishes, and fishes to cause major community changes over large areas. Trawling crushes, kills, exposes, and removes these living sources of nourishment and hiding places, making life dangerous for young fishes and almost certainly lowering the habitat's capacity to produce abundant fish populations. In one study of a scallop fishery, only 12% of the scallops in an area were actually caught, but many of those not caught were crushed by scallop dredges.

Until recently, trawlers were unable to work on high-relief, bouldery, or rubble-strewn bottom habitats without risking hanging up and losing their nets and gear. But nowadays, new kinds of trawls make every kind of seabed – whether silt, sand, clay, gravel, cobble, boulder, rock reef, worm reef, mussel bed, seagrass flat, sponge bottom, or coral reef – vulnerable to trawling. Furthermore, trawls hit many areas repeatedly each year. The seafloor structures that juvenile fishes rely on are often easily destroyed by bottom trawls, which kill or disperse the organisms that create those structures. Our focus on extracting edible fishes at all costs has effectively blinded fishery managers to the essential food and shelter that these fish themselves require. Trawling is not uniformly bad for all species or all bottom habitats, and in fact some species do better in trawled than in undisturbed habitats (just as a few species do better in logging clear-cut areas). But most species are not helped by trawling, and some communities are seriously damaged. Recovery of certain bottom communities could require up to a century even if trawling was stopped today. Writing in the journal *Nature*, Daniel Pauly and Villy Christensen observed that, "Contrary to some terrestrial systems such as rainforests, of which large undisturbed tracts still exist,… the overwhelming bulk of the world's trawlable shelves are impacted by fishing, leaving few sanctuaries where biomasses and biodiversity remain high."

The Course Ahead

Marine Protected Areas

In times past, fisheries benefited from *de facto* refuges: undiscovered locations or places too deep or remote to probe. But now, because of new technologies, when fish run they cannot hide. Facing this reality, some scientists have proposed that simply closing some areas to fishing and resource extraction, to allow them to function naturally, while leaving adjacent areas open, is the best way to manage fishing. These closed areas have been suggested especially where precise estimates on population sizes and sustainable catch levels are lacking and where gathering lots of data or intensive management and monitoring are impractical.

Such marine reserves allow fish to go through their normal patterns of growth, abundance, reproduction, genetic diversity, and community structure. Marine reserves can (1) protect critical spawning adults, (2) maintain natural size and age structure, (3) maintain genetic diversity, (4) prevent serial overfishing, where one species and then the next and the next are depleted, (5) maintain natural communities, while allowing fishing to continue nearby, and (6) provide scientific areas for the study of more naturally functioning systems (Norse, 1993). A global review (Halpern, 2003) of marine protected areas (MPA) determined that within reserves, fish sizes are 30% larger, densities are double, and biomass levels are triple compared with unprotected areas. Over 5000 MPAs exist, but they cover less than 1% of the world's ocean. And even in many reserves, fishing is allowed. Fishing, for example, is allowed in all US National Marine Sanctuaries. Marine reserves have been established in several parts of the world, and as populations recover, fish reach their normally evolved maximum size, and their offspring begin wandering outside the reserve, they can increase fish yields in adjacent areas. In the Philippines, catches adjacent to one reserve tripled within two years of the reserve's designation. In a small Caribbean reserve, overall biomass of commercially important species increased 60% and snappers increased 220% in two years; but groupers did not increase, probably because severe depletion eliminated any source of larvae.

Eggs and larvae of fishes and other marine creatures may drift many miles from their source, so reserves must either be selfsustaining sources of fish or areas where incoming juveniles can grow and reproduce. Ideally, reserve networks should be designed to maintain genetic and community diversity over large regional areas. Large adults are disproportionately fecund – a single female red snapper 24 in (61 cm) long produces as many eggs (about 9 million) as *two hundred* 16-inch (41 cm) females (Bohnsack 1996). One study estimated that if 20% of the red snapper's Gulf of Mexico habitat was protected, total egg production would be 1200% greater than under current fishing pressure. Total fish catch would increase even though a fifth of the area was off-limits to fishing. For more migratory fish, seasonal and area closures of areas such as spawning and nursery grounds could be helpful. For example, longlines catch mostly immature swordfish in certain areas. Those areas should be closed to long-lining during times where breeders or juvenile fish concentrate. Reserves are generally opposed by fishers at first, but in New Zealand they worked so well that 10 years after their establishment nearly 80% of fishers wanted more reserves.

Other Solutions, and Reasons for Optimism

The current situation is poor, but not bleak. There are reasons for optimism. For one, we know enough about many human-induced problems, especially overfishing and habitat alteration, to fix them. Clearly one of the most important things that could be done for overfishing and bycatch is to remove the artificial tax breaks, supports, and all the other subsidies that are propping up fisheries incapable of existing off the resources. And many fishes (though not sharks) have high reproductive potential and may be able to repopulate relatively quickly if fishing and habitat degradation are controlled. Where this has been done, protected populations have often shown the ability to increase rapidly.

Some areas should be closed to trawling even if left open to stationary fishing gear such as traps or hook-and-line, which do not destroy the habitat. New Zealand and Australia have closed areas to bottom trawling, as have some US states. Trawls should be allowed on shallow sandy bottoms that are relatively resistant to disturbance, and barred from harder, higher-relief, and deeper bottoms where their damage is much more serious. Government and industry should create incentives for developing fishing gear that does not degrade the very habitat on which the fishing communities depend. Intelligently designed financial incentives for encouraging new problem-solving technology could tap fishers' inherent inventiveness in constructive ways. Reducing the problems of species introductions will require methods for safely sterilizing ship ballast water, and controlling aquacultural movements of species and their hitchhikers and pathogens.

Public interest and commitment regarding the oceans are much higher in just the last few years than ever before, and conservationists have discovered that fish are wildlife. Consumers hold tremendous power to direct change, but they have hardly been informed about how they can "vote with their wallets" in the marketplace to favor sustainable fishing. Initiatives like the Blue Ocean Institute's seafood guides and the Marine Stewardship Council (fishery certification program wherein participating seafood producers receive an on-pack logo saying this product is sustainably produced) help empower consumers. Much more can be done in this area.

Many discussions of fisheries highlight the need for more information in order to manage wisely. But one does not need to fully understand ecosystems to limit obviously detrimental human behavior. Most of what limits our ability to manage is political resistance and lack of enforcement – better knowledge will not necessarily lead to better management (see **Box 6**). Reducing the undue influence of the irresponsible segments of the fishing industry in management arenas, and their incessant political pressure to allow higher and unsustainable catches, is a necessary first step.

Fish Conservation

At present, in almost all parts of the marine realm, exploitation reigns and conservation begs indulgence. The burden

Box 6 Smuggling Whales

Though whales are mammals, not fishes, whaling is sometimes considered a fishery and is similar to fishing in many aspects. Originally, whales were hunted for their meat, but during the seventeenth century commercial whaling focused on another product: whale oil. For over 200 years, more than 50,000 whales were caught annually, driving many species like the blue whale and North Atlantic right whale to near extinction. In 1986, the International Whaling Commission banned commercial whaling, but whaling to supply meat still occurs in several countries for traditional purposes by indigenous tribes, or under the guise of "scientific whaling" and other loopholes.

Japan kills hundreds of minke, fin, and humpback whales in Antarctic waters each year, ostensibly for scientific purposes. But the samples go to food markets, and the markets serve as continued incentive to hunt whales illegally. Iceland and Norway kill hundreds of minke and fin whales in northern waters annually, with the meat eaten locally or exported to Japan.

The export or international trade of whales is prohibited by Convention on International Trade in Endangered Species (CITES), but Japan, Iceland, and Norway maintain reservations on the CITES treaty for whales, essentially a loop-hole to allow trade of whale meat. Whale meat cannot be legally traded with other countries, but a recent study found that whale meat was available for purchase in the USA and South Korea (Baker *et al.*, 2010). Sometimes whale meat is mislabeled, listed as the more common "minke whale" when it is really from endangered species like fin and sei whales. Falsifying catch records is problematic in commercial whaling, with the Soviet Union underreporting 100,000 whales caught in the Southern Hemisphere prior to the 1970s. This dishonesty is why the arguments about sustainable hunting of whales are flawed and sometimes fraudulent, and why any whaling may prevent recovery of endangered whales in the vast areas of the world's oceans where they are still vanishingly rare.

The Atlantic gray whale was apparently hunted to extinction, but whales stand a good chance of recovery when effectively protected. This has been proven by increases of several great whale species in the Northwest Atlantic, humpbacks that breed in Hawaii, blue whales off California, and the recent full recovery of the Pacific gray whale – one of the great successes of the U.S. Endangered Species Act.

of proof must shift, so that conservation and restoration are the rules and exploitation becomes an effectively regulated privilege. Otherwise, the kingdom of the sea will sink deeper into poverty, creating more and more problems and hunger along the world's increasingly crowded coasts. Of course, human overpopulation and overconsumption drive or worsen each of the problems discussed here.

Anyone interested in conservation can join organizations working in this area, write letters to agencies and government officials, speak at meetings and hearings, and write opinion pieces and letters to editors of various publications. Making your voice heard is the most important thing. Public education is vital in conserving fish populations, and today there are many services, from wallet guides to cell phone applications and texting products, that provide instant information about the health of a seafood species or the environmental impact of catching them.

The solutions to overfishing and marine habitat destruction depend on political will, but politics reflects public opinion, and that means politics can be changed. Already, because of changing political "seascapes," several formerly unlikely long shots have been achieved – including the global ban on drift netting, strengthening of US fisheries law toward ending overfishing and hastening the rebuilding of depleted populations, several new international fishing treaties and agreements, and increasing the acceptance of the "precautionary principle" that says we should act conservatively, consider future needs, and so avoid irreversible change. Only by increasing our political involvement and effectiveness will we be able to improve the ecological health and protect the biodiversity of the world's oceans. Conservation itself is a political activity in the service of living nature.

See also: Aquaculture. Coastal Beach Ecosystems. Corals and Coral Reefs. Estuarine Ecosystems. Fishes, Biodiversity of. Fish Stocks. Greenhouse Effect. Keystone Species. Salmon

References

Baker CS, Steel D, Choi Y, *et al.* (2010) Genetic evidence of illegal trade in protected whales links Japan with the US and South Korea. *Biology Letters* 6: 647–650.

Bohnsack J (1996) Marine reserves, zoning, and the future of fishery management. *Fisheries* 21: 14–16.

Brothers N, Duckworth AR, Safina C, and Gilman EL (2010) Seabird bycatch in pelagic longline fisheries is grossly underestimated when using only haul data. *PLoS ONE* 5(8): e12491.http://dx.doi.org/10.1371/journal.pone.0012491.

Cao L and Caldeira K (2008) Atmospheric CO_2 stabilization and ocean acidification. *Geophysical Research Letters* 35: L19609.

Casey JM and Myers RA (1998) Near extinction of a large, widely distributed fish. *Science* 281: 690–692.

Crawford CR, Steele P, McMillen-Jackson AL, and Bert TM (2011) Effectiveness of bycatch-reduction devices in roller-frame trawls used in the Florida shrimp fishery. *Fisheries Research* 108: 248–257.

Creel L (2003) *Ripple Effects: Population and Coastal Regions*. Washington, DC: Population Reference Bureau.

FAO (2010) *The State of World Fisheries and Aquaculture 2010*. Rome: FAO.

Halpern BS (2003) The impact of marine reserves: Do reserves work and does reserve size matter? *Ecological Applications* 13: 117–137.

ICCAT (2009) *Report of the Independent Performance Review of ICCAT*. Madrid: International Commission for the Conservation of Atlantic Tunas (ICCAT).

Jackson J (1997) Reefs since Columbus. *Coral Reefs* 16: S23–S32.

Myers N (1997) *Perverse Subsidies; Their Nature, Scale and Impacts*. Chicago: The MacArthur Foundation.

Myers RA and Barrowman NJ (1996) Is fish recruitment related to spawner abundance? *Fishery Bulletin* 94: 707–724.

Naylor RL, Goldburg RJ, Mooney H, *et al.* (1998) Nature's subsidies to shrimp and salmon farming. *Science* 282: 883–884.

Norse EA (1993) *Global Marine Biological Diversity*. Washington, DC: Island Press.

Orr JC, Fabry VJ, Aumont O, *et al.* (2005) Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* 437: 681–686.

Pauly D and Christensen V (1995) Primary production required to sustain global fisheries. *Nature* 374: 255–257.

Pauly D, Christensen V, Dalsgaard J, Froese R, and Torres F (1998) Fishing down marine food webs. *Science* 279: 860–863.

Roberts CM (1995) Effects of fishing on the ecosystem structure of coral reefs. *Conservation Biology* 9: 988–995.

Safina C (1995) The world's imperiled fish. *Scientific American* November: 46–53.

Safina C and Klinger D (2008) Collapse of Bluefin Tuna in the Western Atlantic. *Conservation Biology* 22: 243–246.

Sutherland WJ (1990) Evolution and fisheries. *Nature* 344: 814–815.

Fishes, Biodiversity of

Gene S Helfman, University of Georgia, Athens, GA, USA

Published by Elsevier Inc.

Glossary

Adipose fin A small, fleshy fin without supporting spines or rays, set far back on the dorsal surface of many catfishes, characins, salmons, and other groups.

Anadromy A form of diadromy where a fish is born in fresh water, matures in the ocean, and returns to fresh water to spawn (lampreys, sturgeons, salmons). Contrast with catadromy.

Ancestral The taxon from which descendant species are derived, often synonymized as primitive or generalized. Ancestral traits or conditions are those which appear in an ancestor.

Catadromy A form of diadromy where a fish is born in the ocean, matures in fresh water, and returns to the ocean to spawn (freshwater eels, some mullets). Contrast with anadromy.

Depauperate Of low diversity, lacking in species; opposite of speciose.

Derived Later-appearing taxa within a lineage, often synonymized as advanced or specialized. Derived traits are those which appear in a descendant species and are changed from the ancestral condition.

Endemic Restricted or native to a geographically defined area.

Extant Living; opposite of extinct.

Speciose Of high diversity, having many species in a group or area.

Taxon (plural = taxa) A group of evolutionarily related species.

Taxonomic Diversity

Overview

Approximately 32,000 known fish species inhabit the earth's oceans, estuaries, rivers, lakes, and streams. This incredible diversity exceeds that of all other vertebrate groups combined (Box 1). It forms a wealth of biological wonder for ichthyologists (fish scientists), but such large numbers can be overwhelming to someone unfamiliar with the many taxonomic groups and their names. However, the different taxonomic groups are logically arranged according to well-studied evolutionary relationships, with those more closely related to groups that evolved earlier in geologic time (so-called ancestral or primitive taxa) placed earlier in lists and those groups that evolved relatively recently (derived or advanced taxa) placed later in lists (Nelson, 2006).

The most primitive of the living fishes are the jawless fishes. These arguably include 22 species of marine lancelets, and less arguably about 115 species of marine hagfishes and freshwater or anadromous (migratory between freshwater and marine) lampreys. Cartilaginous sharks, skates, rays, and chimaeras include approximately 10 orders, 43 families, and 1300 species of almost entirely marine, relatively large-bodied predators. Skates and rays are more diverse than sharks, constituting about 55% of all cartilaginous fishes. Chimaeras are cartilaginous fishes distantly related to sharks that consist of 3 families and 46 species.

Bony fishes make up the vast majority of living fish species, exceeding all other groups in species, habitat, reproductive, and feeding diversity (Figure 1). Bony fishes vary in length from the 8-mm pygmy gobies and Indonesian minnows to the 8-m-plus long Oarfish, 900-kg marlins, and 1000-kg-plus Ocean Sunfish. There are perhaps 45 orders and 435 families among the approximately 31,000 species of bony fishes, ranging from the relatively primitive lungfishes, coelacanths, sturgeons, bichirs, gars, and Bowfin to the more advanced

Box 1 What is a fish and what are fishes?

Generally defined, a fish is a cold-blooded, aquatic-living chordate with fin-like appendages, a body covered with scales, and that breathes using gills. Exceptions to and variations in all traits are common (Paxton and Eschmeyer, 1998). Some sharks, tunas, and billfishes are warm-blooded. Fins can be unsegmented and spiny or segmented and soft rayed (and soft rays can be hardened). Tail types include: (1) the heterocercal tail of sharks and sturgeons, in which the notochord or vertebral axis extends considerably into the upper tail lobe; (2) the abbreviate heterocercal tail of gars and Bowfin, in which the vertebral axis extends only slightly into the upper lobe; (3) the leptocercal or diphycercal tail of lungfishes, the coelacanths, and rattails, in which vertebrae or tail rays extend through the middle of the tail, forming a pointed tail; and (4) the homocercal tail of most advanced bony fishes, in which the vertebral column ends at the tail base (the urostyle) and fin rays form a symmetrical, two-lobed tail. Scales also vary in terms of the number of layers of bony material that constitute them and the extent of spiny projections that cover their surface; more primitive fishes generally have heavier scales, and more advanced fishes have lighter scales, often with more projections. Scale types include the placoid scales of sharks, the ganoid scales of gars and bichirs (and sturgeons), the cycloid scales of lower teleosts, and the ctenoid scales of higher teleosts. But unlike familiar groups of vertebrates such as birds and mammals that can trace their roots back to some common ancestor, "fishes" include a tremendous diversity of groups with different ancestries. Hence our definition can include lancelets, hagfishes, lampreys, sharks, lungfishes, and bony fishes, each evolving from a different ancestral group.

To ichthyologists, "fish" refers to one or more individuals of a single species, whereas "fishes" refers to more than one species, regardless of how many individuals are involved. Hence, this article is about fishes (Moyle and Cech, 2004; Barton, 2006; Helfman *et al.*, 2009).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00054-X

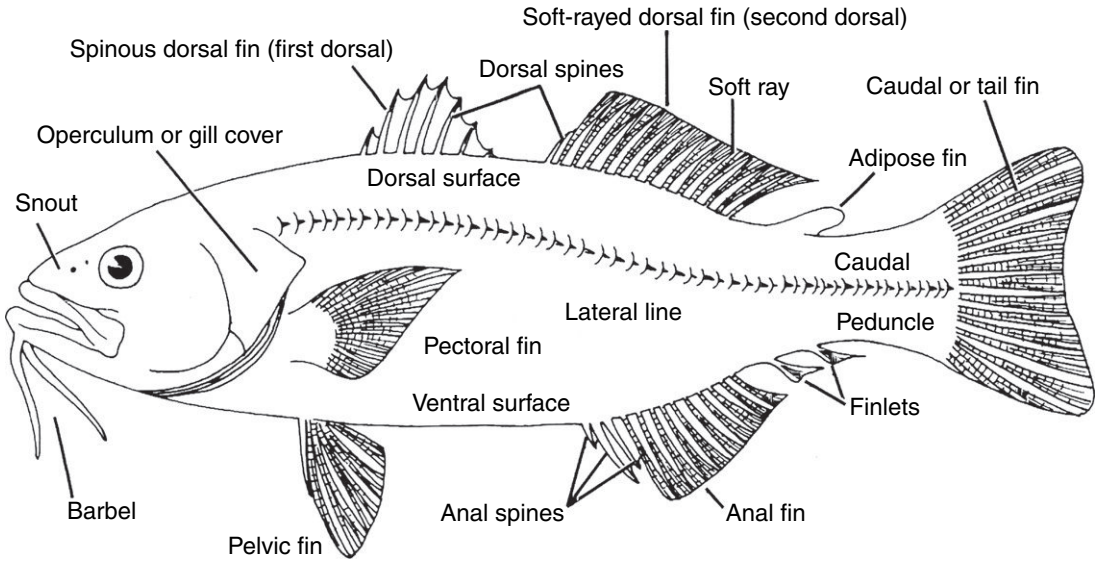


Figure 1 Anatomical features of a hypothetical bony fish. Reproduced from Squire Jr. JL and Smith SE (1977) *Angler's Guide to the United States Pacific coasts.* Washington, DC: National Marine Fisheries Service.

teleostean (higher bony fish) groups that include bonytongues, eels and tarpons, herring-like fishes and the so-called true teleostean groups of minnows, salmons, various deepsea taxa, and cods. The spiny-rayed teleosts are the most evolutionarily advanced of the fishes and include mullets, silversides, scorpionfishes, perch-like fishes, tunas, flatfishes, and triggerfishes (Helfman and Collette, 2011).

Phylum Chordata Subphylum Cephalochordata; order Amphioxiformes (lancelets, two families, 30 species, marine, tropical and temperate) Subphylum Vertebrata Superclass Agnatha Class Myxini; order Myxiniformes (living hagfishes; one family, 70 species, temperate marine) Class Cephalaspidomorphi; order Petromyzontiformes (living lampreys; three families, 38 species, temperate fresh water and anadromous).

Cephalochordates

Lancelets may not really be fishes because they lack scales, fins, and gills. However, their evolutionary and anatomical affinities are similar to what most workers believe characterized the ancestors of fishes, and lancelets are studied primarily by ichthyologists. Lancelets possess a dorsal nerve tube, in common with all vertebrates. They also have a notochord, which is a cartilaginous rod that runs the length of the body and is shared with all embryonic vertebrates as well as with the adults of many primitive living fishes. Lancelets are small (to 3 cm), slender organisms that as adults occupy sandy, usually shallow bottoms in all major tropical and temperate oceans. A commercial fishery for lancelets involving bottom dredging exists in southern China. The fishery is in apparent decline.

Agnathans

The first fishes lacked jaws. Modern jawless fishes – hagfishes and lampreys – look approximately similar, with slippery, eel-like bodies and jawless heads. The fossil record, however, indicates that they have been separated evolutionarily for hundreds of millions of years. Hence, most similarities are due to convergent or parallel evolution.

Hagfishes

Hagfishes, known also as slime eels or slime hags, produce copious mucus from many pairs of slime glands (**Figure 2**). A couple of disturbed 2-ft-long hagfish can fill a 5-gallon bucket with slime. However, a hagfish covered in its own slime will suffocate. To rid itself of slime, a hagfish ties a knot in its tail and passes the knot forward until the slime is pushed off. Hagfishes are nocturnal predators on small invertebrates but are better known for their scavenging behavior, which involves burrowing into a dead or dying fish and consuming the prey from the inside.

Hagfishes occur almost worldwide in temperate and cold-temperate oceans, usually in water deeper than 30 m. Hagfishes can reach high densities upwards of 0.5 m^{-2} on soft-bottom marine areas, which is the most abundant habitat type in the world. Hence, hagfishes could be ecologically important as predators and scavengers. They are also common in the diets of seals and sea lions. Hagfishes are commercially important because their hides are the popular "eelskin" of wallets, purses, and briefcases. Overfishing has depleted hagfish stocks in Korea and Japan, and new fisheries are being exploited, and probably overexploited, in the eastern Pacific and western Atlantic. This unfortunate chain of events has characterized marine fisheries worldwide (Helfman, 2007). Drastic reductions in hagfish populations brought on by overfishing could potentially disrupt a widespread ecosystem; not enough is known about hagfish ecology to predict these impacts.

Lampreys

Lampreys also have a notochord rather than vertebrae. Lampreys go through a long-lived larval phase; the free-living, blind, toothless larva lives in silty streambeds where it filters



Figure 2 Hagfishes, also called slime eels, produce prodigious amounts of mucus when disturbed. Photo by Meyer JL.

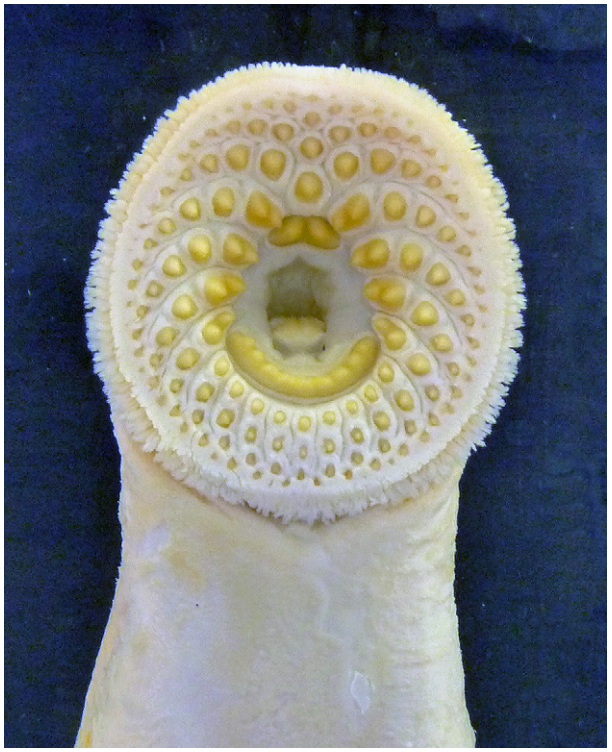


Figure 3 Mouth region of a lamprey. Lampreys attach to the sides of larger fish and scrape a hole with their tongue, feeding on body fluids.

microscopic organisms from the water for up to 7 years before transforming into an adult. In brook lampreys, larvae transform into nonfeeding adults, live for 6 months, spawn, and die. Other species transform into feeding adults and live for 1–3 years as parasites on other fishes. They rasp holes in their host's skin and live off its body fluids (**Figure 3**). Accidental introduction of the parasitic Marine Lamprey into the Great Lakes of North America has contributed to the decline of Lake Trout, whitefishes, and Blue Pike.

Lampreys are cool-water species (30° north and south latitude or higher). Most lampreys live in fresh water, but some parasitic species are anadromous. Brook lampreys typically live in headwater streams, an ecosystem type frequently disrupted by human activities. Hence, several US brook lampreys are imperiled. North America's smallest lamprey, the Miller Lake Lamprey, was poisoned almost into extinction because it parasitized introduced trout in its only habitat – Miller Lake, Oregon.

Cartilaginous Fishes

Elasmobranchs (sharks, skates, and rays) include perhaps 1300 species that live in all the world's oceans; a few live in fresh water (Carrier *et al.*, 2004, 2010). Elasmobranchs are characterized by a cartilaginous skeleton hardened by calcium deposits and usually five (sometimes six or seven) gill slits. They lack lungs or gas bladders but instead have large, oil-filled livers which may aid in buoyancy. Their teeth and pedestal-like placoid scales develop from the same embryonic structures. Teeth are continually lost and replaced, and a shark may produce as many as 30,000 teeth during its lifetime. All elasmobranchs have internal fertilization, and many bear live young that are nourished by the mother *via* a complex umbilicus and placental structure analogous to that found in mammals. Slow growth, late maturation, and low reproductive output make many elasmobranchs exceptionally vulnerable to human exploitation.

Class Chondrichthyes: cartilaginous fishes Subclass Elasmobranchii: shark-like fishes Infraclass Euselachii: sharks and rays Division Neoselachii Subdivision Selachii sharks Superorder Galeomorphi Order Heterodontiformes: 1 family, 8 species, bullhead and horn sharks Order Orectolobiformes: 7 families, 32 species, including wobbegongs, nurse, and Whale sharks Order Carcharhiniformes: 8 families, 224 species, including catsharks, requiem, and hammerhead sharks Order Lamniformes: 7 families, 15 species, including ragged tooth, Tiger, Megamouth, thresher, Basking, and mackerel sharks Superorder Squalomorphi Order Hexanchiformes: 2 families, 5 species, Frill and cow sharks Order Echinorhiniformes: 1 family, 2 species Order Squaliformes: 6 families, 97 species, including sleeper and dogfish sharks Order Squatiniformes: 1 family, 15 species, angel sharks Order Pristiophoriformes: 1 family, 5 species, saw-sharks Subdivision Batoidea skates and rays Order Torpediniformes: 2 families, 59 species, torpedo electric rays and numbfishes Order Rajiformes: 4 families, 295 species, guitarfishes, skates Order Myliobatiformes: 10 families, 183 species, stingrays, butterfly rays, eagle rays, and manta rays.

Sharks

Approximately 465 shark-like species are alive today. Sharks are generally large (>1 m), predatory fishes. The diverse requiem or ground sharks (carcharhiniforms) include the Tiger,

Gray Reef, Bull, Blue, lemon, and hammerhead sharks. Lamniform mackerel sharks are primarily offshore, pelagic inhabitants, although White Sharks spend considerable time close to beaches and small islands, especially where seals and sea lions abound. The squaliform dogfishes, the second largest shark order, are most successful and abundant in the North Atlantic, North Pacific, and deepsea regions.

Sharks range in size from the 15-g, 16- to 20-cm Dwarf Dogshark to the 12,000-kg-plus, 12-m-long Whale Shark, the largest fish in the world. White Sharks as large as 6 m and 3324 kg are known; larger individuals are suspected. Other large sharks include Basking Sharks (9.8 m), Great Hammerheads (5.5 m), Greenland Sharks (6.4 m), Tiger Sharks (5.9 m), and Megamouth (5.4 m).

Sharks inhabit all oceans except the Antarctic. Depth records for sharks are held by the Portuguese Shark at 3690 m and an unidentified dogfish at 4050 m. A few carcharhinid sharks enter fresh water; Bull Sharks have been captured 4200 km up the Amazon River and 1200 km up the Mississippi River. Large pelagic sharks may cross entire ocean basins; Blue Sharks have been tracked across the North Atlantic Ocean and back, a distance of 16,000 km, and a White Shark moved from South Africa to Western Australia and back to South Africa over a 9 month period, a distance of 22,000 km.

Most sharks are predatory on large prey, but three of the largest sharks – the Basking, Megamouth, and Whale sharks – feed on zooplankton. A small, 40-cm-long, midwater species, the Cookie-Cutter Shark, is an ectoparasite on the sides of tunas, dolphins, whales, an occasional Megamouth Shark, and even rubber sonar domes of nuclear submarines. Some sharks use structures other than jaw teeth to capture prey. Thresher sharks use the long upper lobe of their tails to stun schooling prey. Saw-sharks (and rajiform sawfishes) have elongate, blade-like snouts studded with lateral teeth which they slash laterally to disable prey. Hammerheads use their broadened hammer-shaped head to pin stingrays against the bottom before biting chunks out of the rays' wings.

Sharks are sensitive to chemicals, able to detect 1 part fish extract per 10 billion parts seawater. Sharks have good vision, although they tend to be slightly myopic (farsighted). Sharks are also highly sensitive to sounds, including infrasonic sound below 10 Hz, and can localize the direction from which underwater sounds originate, something humans cannot do. Sharks can also locate prey by detecting the weak electric fields their prey emit. This electro-sensitivity may also allow them to navigate using the earth's geomagnetic fields. Many sharks have relatively large brains, with brain to body weight ratios comparable to those of some birds and mammals.

Sharks grow slowly and live long: Spiny Dogfish live 70–100 years and Lemon Sharks 50–60 years. Sharks also reproduce slowly. Lemon Sharks may not mature until they are 24 years old, and Spiny Dogfish may not mature until they are 35 years old. Sharks produce relatively few, large young with a long gestation period. Clutch size varies from 1 or 2 live young (Ragged Tooth, threshers, and makos) to 300 in the Whale Shark. Gestation periods average 9–12 months but may be 2 years in Spiny Dogfish and 3.5 years in Basking Sharks. Bullhead and nurse sharks lay eggs, but most sharks bear live young. The requiem sharks have the most advanced developmental pattern, in which an umbilical cord connects mother and embryo, transporting nutrients and oxygen to the embryo and carrying metabolic wastes to the mother.

Because many sharks mature late, reproduce at long intervals, and have low reproductive output, shark populations are easily and frequently overfished (**Figure 4**). Thresher sharks, School Sharks, Spiny Dogfish, Porbeagles, Basking Sharks, Bull Sharks, and Soupfin Sharks are all examples of shark stocks that have been overexploited. White Sharks are protected in Australia, South Africa, and the US; capture and sale of White Sharks, Whale Sharks, and Basking Sharks is restricted under the Convention on International Trade in Endangered Species (CITES).



Figure 4 Ragged Tooth Sharks, also called grey nurse and sand tiger sharks, have suffered declines because of fishing and as by-catch in nets. Ragged Tooth sharks are protected in several countries. "Raggies" are relatively docile, adapt well to captivity, and are often featured in public aquaria.

Skates and Rays

The batoid skates and rays are 535 species of mostly benthic (bottom-living), mostly marine forms. In skates and rays, the pectoral fins are fused to the sides of the head and the five gill slits are under the head. Skates are most diverse in deep water and at high latitudes, whereas stingrays are most diverse in tropical, inshore waters. Some batoids live much or all of their lives in fresh water. Largetooth Sawfish frequently swim up rivers in Central and South America. Two stingray families contain entirely freshwater species – the river stingrays of South America and several species in the large stingray family Dasyatidae. The latter inhabit African, Southeast Asian, and New Guinea rivers.

Skates and rays feed mostly on benthic invertebrates, except for the huge (up to 6-m wide) manta rays, which capture small crustaceans and fishes in the water column. Torpedo rays stun prey with powerful electrical discharges (50 V and 50 A = 1 kW output). Batoids reproduce by either laying eggs (skates) or bearing live young (rays). Embryonic skates develop inside "mermaid purse" egg cases for as much as 15 weeks (Figure 5).



Figure 5 A mermaid purse (egg case) of a Big Skate, *Raja binoculata*, from the North Pacific. Each case holds a single embryo, which can be seen because part of the egg shell has been removed. Some skates are overfished because they are born at a size that is already too large to escape through the mesh of drag nets.

Skates in some locales are actually increasing in number because of overexploitation of competing bony fishes, such as cod in the North Atlantic. However, the giant Barndoor Skate of the northwest Atlantic and its relative, the ironically named Common Skate of the northeast Atlantic, are caught incidental to bottom trawling for bony fishes; they have been seriously depleted and may face extinction. Largetooth Sawfish in Nicaraguan lakes have been drastically overfished.

Chimaeras

Class Chondrichthyes Subclass Holocephali Order Chimaeriformes: 3 families, 46 species, chimaeras.

Chimaeras, also known as rat- or rabbitfishes, share a cartilaginous skeleton and other features with elasmobranchs. They differ by having (1) the upper jaw permanently attached to the braincase, (2) continually growing tooth plates in the jaws instead of replaceable teeth, (3) a single gill flap instead of five or more gill slits, and (4) no scales. Chimaeras swim by flapping their pectoral fins (something sharks never do) and by undulating their bodies (Figure 6). All chimaeras are egg-layers, the egg being protected by a horny shell. Adult chimaeras range in size from 60 to 200 cm. Chimaeras are cool-water, marine fishes that live at shallow to moderate depths between 80 and 2600 m, where they usually swim just above the bottom. Chimaeras eat predominantly hard-bodied benthic invertebrates, which they crush with their tooth plates. Surprisingly little is known about their general biology and natural history.

Bony Fishes

Modern bony fishes, often referred to as Osteichthyes (literally "bony fishes"), consist of seven major taxonomic orders. Many living members belong to groups that were much more diverse during Paleozoic and Mesozoic eras. Several of these primitive but modern fishes are classified as "bony" even though they have cartilaginous skeletons. Their skeletal condition is actually an advanced, specialized trait; their immediate ancestors were bony.



Figure 6 A Pacific Ratfish, *Hydrolagus colliei*, showing the beak-like mouth, flexible pectoral fin, and lack of scales that set chimaeras apart from sharks and rays. The large, reflective eye is characteristic of fishes that live in deep, dark waters.

Grade Teleostomi (more commonly Osteichthyes) Class Sarcopterygii Subclass Coelacanthimorpha Order Coelacanthiformes: 1 family and 2 species, the coelacanths Subclass Dipnoi Superorder Ceratodontimorpha Order Ceratodontiformes: 1 family and 1 species, the Australian lungfish Order Lepidosireniformes: 2 families, 5 species, South American and African lungfishes Subclass Tetrapoda (amphibians, reptiles, birds, mammals) Class Actinopterygii Subclass Cladistia Order Polypteriformes: 1 family, 16 species, bichirs and Reedfish Subclass Chondrostei Order Acipenseriformes: 2 families, 27 species, sturgeons and paddlefishes Subclass Neopterygii Order Lepisosteiformes: 1 family, 7 species, gars Order Amiiformes: 1 family and species, the Bowfin.

The Coelacanths

Coelacanths were thought to have gone extinct at the end of the Cretaceous, 65 million years ago, until a live one was trawled up in 1938 off South Africa. Today, one small population of 200–600 coelacanths, *Latimeria chalumnae*, lives at 100- to 500-m depths off two small volcanic islands in the Comore Archipelago, between Madagascar and Mozambique. Other populations of the same species have been discovered along Africa's eastern coast from South Africa to Tanzania. In 1998, another species, *L. manadoensis*, was discovered at similar depths in northern Indonesia. The living coelacanths have fleshy pectoral, pelvic, anal, and second dorsal fins (= the lobed fins that define the class Sarcopterygii); a symmetrical, three-lobed tail with a central extension; hollow neural spines (hence "coelacanth" or "hollow spines"); a unique unconstricted notochord; a joint in the dorsal braincase that aids jaw opening; relatively large, thick, bony scales; and live young.

Coelacanths are large (to 180 cm, 95 kg), old (probably 40–50 years), and produce relatively few, live young (five to 26 young per clutch). The gestation period is about 13 months. Replacement rate in a population is therefore slow. Coelacanths are captured primarily as by-catch in the hook-and-line fishery for Oilfish, and it is unlikely that the Comore population can sustain even the current by-catch rate of five to 10 animals per year. Counts from small submarines indicate the Comoran species is declining. The Comoran government has outlawed its capture, and trade in coelacanths is outlawed by the Convention on the International Trade in Endangered Species (CITES).

Lungfishes

Lungfishes today are represented by three families and six species, all located on the former Gondwanan continents of Australia, South America, and Africa. Lungfishes lack jaw teeth but have unusual tooth plates on the mouth roof and floor. The Australian species, *Neoceratodus forsteri*, is limited to four river systems of northeastern Australia. It is large (1 m), with large scales, flipper-like fins, a broad tail, and a single lung. It crushes benthic crustaceans, mollusks, and small fishes with its tooth plates. It can, but does not have to, breathe atmospheric oxygen. Its young lack external gills. *Neoceratodus* populations have declined dramatically and the fish is protected; species recovery efforts include transplantation into several Queensland reservoirs and rivers.

The one South American and four African lungfishes have eel-like bodies; slender, almost-filamentous paired fins; lack scales; have paired lungs; have larvae with external gills; and must breathe air to survive. The four African species occur across central and south Africa, often in swampy areas that frequently experience drought (**Figure 7**). When a swamp dries up, African lungfishes dig a burrow and can wait 4 years for rains to return. The South American species occurs in swampy regions of the Amazon and Parana river basins. Comparatively little is known about its biology.

Tetrapods

Recent paleontologic, anatomic, and genetic discoveries indicate that the 26,734 species of tetrapods (modern amphibians, reptiles, birds, and mammals) are in reality a subclass of the fleshy-finned sarcopterygians. An exciting, confirming fossil discovery, the "fishapod" *Tiktaalik roseae*, was unearthed in 2005 in Nunavut Territory of Arctic Canada. Tetrapods are therefore a sideline taxon of fishes that moved onto land and into the air, with some groups such as turtles, whales, and seals returning to an aquatic existence. We are all fish.

Bichirs and Reedfish

The polypteriform bichirs and Reedfish of west and central tropical Africa have been difficult to assign to any particular taxonomic group. Their gilled larvae, lobe-like fins, thick ganoid scales, and modified heterocercal tail suggest affinities with several living groups, particularly the chondrostean sturgeons and paddlefishes. However, their unique dorsal, caudal, and paired fins and unusual chromosomes place them apart from all extant groups. Their placement near the base of the bony fish lineage is justified by egg structure, genetics (mitochondrial, nuclear, and other DNA and amino acid sequences), and head skeleton. Taxonomists now place them in their own subclass, the Cladistia.

Ten species of polypteriforms are called bichirs (*Polypterus*) and an 11th, elongate species is the Reedfish or Ropefish,



Figure 7 African lungfishes, *Protopterus*, estivate, burrowing into mud and remaining inactive for years while waiting for rains to return. Photo courtesy of Chapman L and Chapman C.

Erpetoichthyes. Bichirs reach 120 cm in length and Reedfish 90 cm. All are predatory and inhabit shallow, swampy regions, where they breathe atmospheric oxygen with their paired lungs. Bichirs, also called "flagfins," have unique dorsal and pectoral fins. Each dorsal finlet consists of a vertical spine with attached horizontal rays, looking like a flagpole with streaming banners. In more usual ray-finned fishes, dorsal fin rays emerge vertically from the body of the fish. The polypteriform pectoral fin has an internal arrangement that involves a wishbone-shaped, flattened plate, again unlike that in any other fish. Bichirs are not particularly well studied; no evidence exists to suggest they are imperiled.

Sturgeons and Paddlefishes

The next-most primitive actinopterygian (ray-finned) fishes are the chondrostean sturgeons and paddlefishes. All 25 species of sturgeons live in the Northern Hemisphere. All spawn in fresh water, although some species move seasonally between marine and fresh water. North American freshwater species include the Lake Sturgeon and three river sturgeons. Anadromous species include the Atlantic and White sturgeons, the latter being an occupant of west coast bays and rivers. White Sturgeons attain the largest size of any North American freshwater fish (3.8 m, 630 kg). The world's largest freshwater fish is the Beluga Sturgeon of eastern Europe and Asia, *Huso huso*, at 8.6 m and 1300 kg (**Figure 8**).

Sturgeons have four barbels ahead of a ventrally located mouth, five rows of large bony shields on an otherwise scaleless body, and a heterocercal tail. They are exceptionally long-lived (118 years for Beluga and 70–80 years for White Sturgeon), mature slowly (as late as 30 years old), and spawn infrequently (every 3–5 years). They migrate up rivers to spawn in clean sand and gravel areas; hence, dam building and siltation of rivers impede their reproduction. A spawning female can be worth thousands of dollars for her caviar alone, and many sturgeon stocks have been reduced 99% from historical levels. The Shortnose Sturgeon of Atlantic coastal rivers is listed as endangered, and Lake Sturgeon have been extirpated from a large part of their native range. Three species endemic to the Aral Sea may be extinct due to extensive drying of that once huge body of water, and several species in the former Soviet Union are fished mercilessly.

Two species of paddlefishes occur in large rivers of North America and China. Paddlefishes also have a heterocercal tail, unconstricted notochord, largely cartilaginous skeleton, and scaleless body. The most distinctive feature of the paddlefishes is their paddle or spoonbill, which is flat and rounded in the North American Paddlefish and elongate and almost spear-like in the Chinese Paddlefish. Paddlefishes are water column swimmers that feed on zooplankton and fishes. The paddle of the North American species may help direct food into its mouth, but abundant electroreceptor cells on the paddle suggest an additional, unknown function. North American Paddlefish may live 30 years and grow to be 2.2 m in length and 83 kg in weight. Late maturation (at 10 years of age), infrequent spawning (every 2–5 years), lack of clean gravel spawning habitat, and overfishing have all contributed to population reductions and range contraction. The exceedingly rare and poorly studied Chinese Paddlefish is larger, reaching more than 3 m in length. This paddlefish may be the most endangered fish in China because of overfishing, habitat destruction, pollution, and dam construction that blocks spawning migrations in the Yangtze River basin where it lives. Recent construction of the Three Gorges Dam in China may have sealed the fate of the Chinese species.

Gars

The lepisosteiform gars are seven species of elongate, predaceous fishes that occur in eastern North America, Cuba, and Central America. They typically inhabit backwater areas of lakes and rivers, such as oxbows and bayous, and breathe atmospheric oxygen using a highly vascularized gas bladder. Gars have bony skeletons, but their vertebral centra are unique, being convex anteriorly and concave posteriorly. In most fishes, the vertebrae are concave on both surfaces. Gars have an abbreviate heterocercal tail (vertebral column extends into the top lobe of the tail) and hinged, diamond-shaped, interlocking ganoid (heavy and bony) scales. Gars are the only freshwater fish in North America with poisonous eggs. Alligator Gars can be 3 m long and weigh 140 kg (**Figure 9**). In recent years, Alligator Gars have come under intense commercial fishing, and concern for their well-being is increasing.

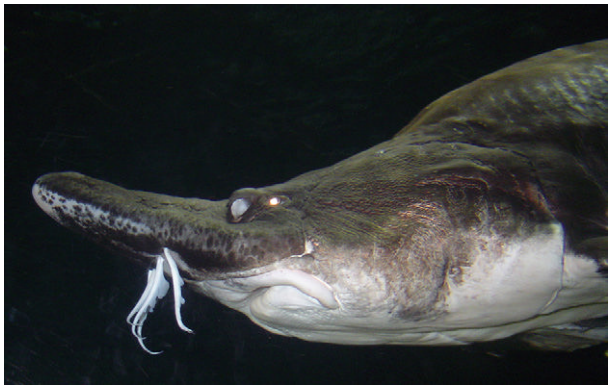


Figure 8 Beluga Sturgeon, *Husa huso*, may be the world's largest and most valuable freshwater fish. The eggs or caviar from a single, large female can be worth thousands of dollars.



Figure 9 An Alligator Gar, *Atractosteus spatula*, one of the largest freshwater fishes in North America. Photo courtesy of Jean-Francois Healias, www.anglingthailand.com.

The Bowfin

The Bowfin, *Amia calva*, is the only living member of its genus, family, and order. The Bowfin has the abbreviate heterocercal tail and spiral valve intestine of the gars but also has teleost-like biconcave vertebrae as well as cycloid scales, a relatively light scale type also possessed by many teleosts. The Bowfin's head is exceptionally bony and the throat is covered by a distinctive large bone, the gular plate. Bowfin swim slowly forwards or backwards by passing undulations back and forth along their long dorsal fin. Bowfin occur throughout much of eastern North America in backwater, often swampy areas; they also have a highly vascularized gas bladder which functions as a lung. They are relatively large and robust (to 1 m and 9 kg) and predatory on anything that moves. Bowfin males guard the young vigorously until they are relatively large (10 cm).

Teleosts

The division Teleostei ("perfect bone") contains most living fishes. Teleosts are not only taxonomically diverse but also ecologically diverse, occupying every aquatic habitat type and niche imaginable. The 27,000 living teleostean species are placed in 4278 genera, 448 families, and 40 orders. This incredible diversity is generally organized into four taxonomic subdivisions that reflect patterns of evolution that date back to the Mesozoic. These four main subdivisions are the osteoglossomorphs (bonytongues), elopomorphs (eels and tarpon), otocephalans (herrings and minnow relatives), and the euteleosts, with the latter group containing the vast majority of modern bony fishes.

Class Actinopterygii

Subclass Neopterygii Division Teleostei Subdivision Osteoglossomorpha Order Hiodontiformes: 1 family, 2 species, mooneyes Order Osteoglossiformes: 6 families, 218 species, including bonytongues, African knifefishes, and elephantfishes.

Osteoglossomorphs derive their name "bonytongue" from the teeth on their tongue that forms part of their bite. These freshwater fishes occur on all major continents except Europe. The hiodontiforms comprise two species, the Mooneye and Goldeye, both of which occur in major river systems of northern North America. The osteoglossiforms are much more diverse. The Asian Arowana or Golden Dragonfish, *Scleropages formosus*, has been depleted in the wild due to overcollecting and is now protected under Appendix I of the Convention on International Trade in Endangered Species (**Figure 10**). The Arapaima of South America is one of the world's largest freshwater fishes, reaching a length of 2.5 m. The African mormyrid elephantfishes produce and detect weak electric fields, have large cerebellums, and have a brain size: body weight ratio comparable to that of humans.

Subdivision Elopomorpha

Order Elopiformes: 2 families, 8 species, ladyfishes and tarpons Order Albuliformes: 3 families, 30 species, including bonefishes and spiny eels Order Anguilliformes: 15 families, 738 species, including freshwater, moray, cutthroat, and conger eels Order Saccopharyngiformes: 4 families, 28 species, including swallower and gulper eels.



Figure 10 The Asian Arowana or Golden Dragonfish (*Scleropages formosus*, Osteoglossidae), an internationally protected species. Desirable color varieties have sold for as much as $5000. Photo by Marcel Burkhard, Wikimedia Commons, http://en.wikipedia.org/wiki/Image:Arowanacele4.jpgfile

Elopomorphs all have ribbon-shaped leptocephalus ("pointed head") larvae. The Atlantic Tarpon is a highly prized gamefish that reaches a length of 2.5 m and a mass of 150 kg. Albuliform bonefishes are also popular gamefishes that occupy sandy flats in shallow tropical waters, The 15 families of anguilliform "true" or freshwater eels are distinguished from the approximately 45 other families of "eel-like" fishes that have independently evolved an elongate body. Anguillid eels are catadromous, spawning at sea but spending most of their lives in fresh water. Muraenid moray eels and their relatives are marine, tropical and warm-temperate, predatory species. Synaphobranchid cutthroat eels include an endoparasitic species, the Snubnose Parasitic Eel, which has been found in the heart of a mako shark. The saccopharyngiform deepsea gulper and swallower eels have giant mouths but lack many head bones, scales, and fins found in most other fishes.

Subdivision Otocephala

Superorder Clupeomorpha Order Clupeiformes: 5 families, 364 species, including anchovies and herrings.

Clupeomorphs are small, schooling, silvery, pelagic marine and occasionally freshwater feeders on zooplankton and phytoplankton. Herrings, round herrings, shads, alewives, sprats, sardines, pilchards, and menhadens are extremely important commercial species. Anchovies range in size from a 2-cm Brazilian species to a piscivorous, riverine, 37-cm New Guinea anchovy. The largest clupeids are the Indo–Pacific chirocentrid wolf herrings, which reach a length of 1 m and have fang-like jaw teeth. Anadromous shads, Alewife, and herrings occasionally establish landlocked populations in rivers, lakes, and reservoirs.

Superorder Ostariophysi Order Gonorhynchiformes: 4 families, 37 species, including milkfish Order Cypriniformes: 6 families, 3268 species, including minnows, barbs, algae eaters, suckers, and loaches Order Characiformes: 18 families, 1674 species, including freshwater hatchetfishes, tetras, and characins Order Siluriformes: 36 families, 2867 species, including North American freshwater, airbreathing, electric,

sea, upside-down, parasitic, callichthyid armored, and suckermouth armored catfishes Order Gymnotiformes: 5 families, 134 species, including glass, ghost, and naked-back knifefishes and Electric Eel.

Rivers, lakes, and streams worldwide are dominated numerically and ecologically by members of the superorder Ostariophysi. Ostariophysans include minnows, carps, barbs, suckers, loaches, piranhas, tetras, catfishes, and electric eels. Two distinctive traits unite this otherwise disparate assemblage: (1) the Weberian apparatus, which is a series of modified anterior vertebrae that link the gas bladder to the inner ear and aid in hearing, and (2) production and reaction to chemical alarm substances that are released when a fish is injured and lead to a stereotyped escape response in school members.

The gonorhynchiform Milkfish, *Chanos chanos,* is an important food fish in the Indo–Pacific region and is often cultured in brackish fish ponds. The Cypriniformes make up the largest order in the superorder. The Cyprinidae is the largest family of freshwater fishes and contains 3268 species of minnows, shiners, carps, barbs, barbels, gudgeons, chubs, dace, pikeminnows, Tench, Rudd, bitterlings, bream, southeast Asian "sharks" (Redtail Black Shark and Bala Shark), Goldfish, Koi (domesticated common carp), danios, and rasboras. Cyprinids are most diverse in Southeast Asia, followed by Africa, North America (270 species), and Europe, but are missing from South America and Australia. The world's smallest freshwater fish is an Indonesian cyprinid, *Paedocypris progenetica*, that matures at 7.8 mm. The largest minnow in North America is the endangered, piscivorous Colorado Pikeminnow, *Ptychocheilus lucius*.

Gyrinocheilid algae eaters scrape algae off rocks in swift, flowing waters. The catostomid suckers (i.e., buffaloes, Quillback, carpsuckers, Blue Sucker, redhorses, jumprocks, and the extinct Harelip Sucker) include 72 species of North American fishes, with 1 species in eastern China and another in Siberia. Loaches (Cobitidae) are 177 species of predominantly Eurasian stream fishes, including popular aquarium fishes such as the kuhli, Clown, and skunk loaches, the weatherfishes, and the Golden Dojo. Weatherfishes (*Misgurnus*) become restless when barometric pressure decreases preceding a storm.

The characiforms are a speciose group of primarily tropical ostariophysans characterized (usually) by a rayless adipose fin and mouths armed with replacement dentition (e.g., piranhas) (**Figure 11**). Body size ranges from very small (13 mm) tetras to large (1.5-m-long) tigerfishes. Numerous aquarium fishes are included (headstanders, freshwater hatchetfishes, blind characins, pencilfishes, tetras, and silver dollars), as well as important food fishes (*Prochilodus*, *Colossoma*, and *Brycon*). Most characins (ca. 1300+ species) are South American, about 200 are African, a few live in Central America, and 1 species, the Mexican Tetra *Astyanax mexicanus*, extends naturally into southwestern Texas.

Catfishes (Siluriformes) are surprisingly diverse, with 36 families and more than 2900 species. Catfishes usually have barbels (whiskers) and sometimes toxic spiny fins, and they are almost entirely freshwater, nocturnal, and benthic (a notable exception being the marine Striped Catfish, *Plotosus lineatus*, which is active by day). Catfishes are most diverse in South America (e.g., loricariid suckermouth armored catfishes, 684 species; trichomycterid pencil catfishes, 201 species, and callichthyid armored catfishes, 177 species). Large species include the European Wels (5 m, 330 kg), the Asian Mekong Giant Catfish (3 m, 300 kg), and a 3-m-long whiskered pimelodid of South America. The largest catfishes in North America are the Flathead and Blue catfishes at about 1.5 m and 50–68 kg. Some small catfishes are notable, such as the parasitic catfishes (Trichomycteridae) of South America, which normally parasitize the gills of fishes but one species, the Candiru, is known to swim up the urethra of male bathers and lodge there, necessitating surgical removal.



Figure 11 Piranhas (subfamily Serrasalminae) belong to the incredibly diverse order of characiform fishes that dominate South American fresh waters. Although certainly capable of dismembering prey, most reported attacks on humans involved drowning victims.

The gymnotiform South American knifefishes are unusual ostariophysans that produce and detect weak electric impulses. They have elongate, compressed bodies; an extremely long anal fin; and electrogenic tissue usually derived from modified muscle cells. Their electrical output is constant at high frequencies, whereas the osteoglossomorph mormyrids produce a pulsed, low-frequency output. Both groups detect objects that disrupt their electric fields. The 2-m-long Electric Eel (*Electrophorus*), a gymnotid, produces a weak field for electrolocation and strong pulses for stunning prey or deterring predators.

Subdivision Euteleostei

Approximately 2/3 of the teleosts are placed in the Euteleostei or true teleosts, a vast assemblage of more advanced fishes. This subdivision contains 346 families, 2935 genera, and 17,419 species – all in nine superorders.

Superorder Protacanthopterygii Order Argentiniformes: 6 families, 202 species, including barreleyes, pencilsmelts, and slickheads Order Salmoniformes: 4 familIes, 154 species, including smelts, galaxiids, Salamander Fish, whitefishes, graylings, chars, trouts, and salmons Order Esociformes: 2 families, 10 species, pikes and mudminnows.

Protacanthopterygians are a mixed agglomeration of marine, freshwater, and diadromous fishes. Argentiniforms are

mostly deepwater marine species. Osmeriforms are small, silvery, elongate, water column dwelling fishes.

Salmoniforms are important commercially, ecologically, and esthetically. Whitefishes and ciscoes are relatively large-scaled, zooplanktivorous salmonids of high-latitude North American and Eurasian lakes. Several North American species have been decimated due to introduced predators, competitors, and parasitic lampreys. Graylings are riverine fishes with a flowing dorsal fin. At least one species, the Michigan Grayling, is extinct. The subfamily Salmoninae contains seven Eurasian and North American genera. The Siberian Taimen, *Hucho taimen*, is the world's largest salmonid at 2 m and 70 kg. North American Salmoninae include the chars (Lake, Brook, and Bull trout, Arctic Char, and Dolly Varden). Arctic Char live farther north than any other freshwater fish. The remaining salmonines are the Atlantic basin salmon and trout (e.g., Atlantic Salmon and European Brown Trout), and the 11 species of Pacific basin trouts and salmons in the genus, two of which are endemic to Japan. Pacific trouts and salmons include Golden, Cutthroat, and Gila trouts and the spectacularly anadromous Coho, Chinook, Chum, Pink, and Sockeye salmons, some of which undergo oceanic migrations of thousands of kilometers before returning to their birth river to spawn and die (**Figure 12**). The Rainbow Trout is a sixth species of Pacific salmon and is called a Steelhead Salmon if it spends part of its life at sea. The actual number of genetically distinct races of Pacific salmons is unknown because many stocks are reproductively isolated in small river systems. Evidence suggests that as many as 1000 stocks exist, more than 100 of which have gone extinct and an additional 300+ of which are imperiled. Osmerids include commercially important species such as capelins, eulachons, Asian Ayu, and smelts. The order Salmoniformes also includes the Southern Hemisphere salamander fish and the galaxiids. Salamander fish inhabit seasonal ponds of southwestern Australia, burying in drying mud and reemerging with the next rains. Galaxiids have suffered numerous extirpations and extinctions as a result of the stocking of nonnative trouts.

Esociform pikes, pickerels, and mudminnows are Northern Hemisphere predators, including the 1.4 m Muskellunge; the Northern Pike has the largest geographical distribution of any Northern Hemisphere fish, occurring across the northern portions of North America, Europe, and Asia. Mudminnows can survive winters in high-latitude lakes by breathing from air bubbles trapped under the ice.

Superorder Stenopterygii Order Stomiiformes: 5 families, 391 species, including bristlemouths, marine hatchetfishes, and barbeled dragonfishes.

Stenopterygians are deepsea fishes, often with long teeth and large mouths. Gonostomatid bristlemouths may be the most abundant and widely distributed vertebrates on Earth. Idiacanthine black dragonfishes have a larva with eyes at the ends of elongate stalks.

Superorder Ateleopodomorpha Order Ateleopodiformes: 1 family, 12 species, jellynosefishes.

The jellynosefishes are elongate deepwater marine fishes that swim just above the bottom. Their large, pointed head and pointed tail are traits they share with other deepsea fishes such as chimaeras, spiny eels, halosaurs, eucla cods, rattails, and grenadiers.

Superorder Cyclosquamata Order Aulopiformes: 15 families, 236 species, including telescopefishes, tripodfishes, lizardfishes, and lancetfishes.

Cyclosquamates are also deepsea forms, including the bizarre giganturid telescopefishes with large tubular eyes, a huge mouth, flexible teeth, and an expandable stomach. Deepsea tripodfishes have long pectoral, pelvic, and caudal rays that they use for resting on soft sediments of the deep ocean floor.



Figure 12 A spawned-out, dead Sockeye Salmon, *Onchorhynchus nerka*, in an Alaskan stream. Spawning runs of Sockeye and other Pacific salmon species play a pivotal role in the ecosystems of the Pacific Northwest. Bears, eagles, etc. feed on spawning fish, and dead fish provide nutrients for food webs, including promoting growth of the large trees that shade the streams essential for survival of young salmon.

Shallow representatives are the synodontid lizardfishes, which are common benthic predators on coral reefs worldwide. Alepisaurid lancetfishes are large (to 2 m) mesopelagic predators with a sail-like dorsal fin of unknown function.

Superorder Scopelomorpha Order Myctophiformes: 2 families, 246 species, including lanternfishes.

Scopelomorphs include the abundant, commercially important lanternfishes, which are identified based on species-specific photophore (light organ) patterns. Lanternfishes occur at middle depths from the Arctic to the Antarctic. They are important in the diets of many fishes as well as of marine mammals.

Superorder Lampriomorpha Order Lampridiformes: 7 families, 21 species, including opahs, tube-eye, ribbonfishes, and Oarfishes.

Lampriforms are generally open-water, oceanic fishes. Opahs are relatively large (1.8 m, 70 kg), oval-shaped, pelagic predators on squids and other fishes. The 30-cm-long tube-eye (*Stylephorus*) can increase the volume of its mouth 40-fold during feeding – a record among vertebrates. The elongate Oarfish, *Regalecus*, may attain 12-m length and is the longest living teleost. It has a bluish-silvery body, scarlet head crest, and deep red fins. It is thought to be responsible for many "sea serpent" sightings (**Figure 13**).

Superorder Polymixiomorpha Order Polymixiiformes: 1 family, 10 species, beardfishes.

Superorder Paracanthopterygii Order Percopsiformes: 3 families, 9 species, troutperches, pirate perch, and cavefishes Order Gadiformes: 9 families, 555 species, including rattails, hakes, and cods Ophidiiformes: 5 families, 385 species, including pearlfishes, cuskeels, and viviparous brotulas Order Batrachoidiformes: 1 family, 78 species, toadfishes Order Lophiiformes: 18 families, 313 species, including goosefishes, frogfishes, handfishes, batfishes, and deepsea anglerfishes.

Paracanthopterygians are primarily benthic, marine, nocturnally active fishes; many live in the deep sea or in caves. Percopsiforms are small (<20 cm), freshwater fishes, most of which live in eastern North America. The anus of the swamp-dwelling aphrododerid Pirate Perch is located in the throat region of adults for functionally mysterious reasons. Amblyopsid cavefishes are often blind and scaleless forms highly adapted for cave life.

The gadiforms include the cods, haddocks, hakes, pollocks, and whitings, which are some of the world's most important commercial fishes. True cods (Gadinae) have three dorsal fins and two anal fins. Many species have chin barbels. The Burbot, *Lota lota*, of high-latitude, Northern Hemisphere lakes is the only freshwater species in the group. The commercially important Atlantic Cod is the largest species (1.8 m, 90 kg), but fish more than 10 kg are rare due to drastic overfishing. One of the world's largest food fisheries is for North Pacific Walleye Pollock.

Ophidiiforms often live in holes or even inside other animals. Carapid pearlfishes live inside the body cavities of starfishes, sea cucumbers, clams, and sea squirts; some feed on the internal organs of their hosts. Ophidiid and bythitid cuskeels and brotulas include blind cave species in freshwater systems of Caribbean and Galapagos Islands as well as coral reef species that hide deep within crevices. The neobythitine cusk eel, *Abyssobrotula galatheae*, holds the depth record for a fish at 8370 m in the Puerto Rico Trench.

Batrachoidiforms include the midshipmen, which have hundreds of photophores, an unusual trait for a shallow dweller. Many batrachoidids produce sounds by vibrating their gas bladders. Venomous toadfishes have dorsal and opercular spines that can inject a powerful toxin. Three South American toadfishes are restricted to fresh water.

Lophiiforms are a diverse and often bizarre-looking order of marine fishes that include benthic, shallow-water forms as well as highly modified, open-water, deepsea forms. Many use a modified first dorsal spine as a lure for catching smaller fish (see **Figure 20**). The meter-long western North Atlantic Goosefish, *Lophius americanus*, has a huge mouth with long, recurved teeth that it uses to catch fishes and even diving seabirds. Antennariid frogfishes also rest on the bottom or walk over it with their pectoral and pelvic fins. The ogcocephalid batfishes walk on their pectorals, but they can also



Figure 13 The Oarfish, *Regalecus glesne*, is probably the longest bony fish in the world and is responsible for many reports of sea monsters. This individual died at a shoreline in Mexico. Photo courtesy of Michael Kanzler.

swim *via* jet propulsion by shooting water out their round, backward-facing opercular openings. The ceratioid anglerfishes include 11 families of strange appearing bathypelagic predators, many of which have very small males that fuse to and become parasitic on the larger females. The endemic Australian handfishes include a Tasmanian species, the Spotted Handfish, that was once common but is now critically endangered due possibly to egg predation by an introduced starfish.

Superorder Acanthopterygii

Most bony fishes belong to a single superorder, the Acanthopterygii, which contains about 14,800 species in 267 families. Two small and one large taxonomic groupings, called series, are recognized, with the vast majority in the third series, the Percomorpha.

Series Mugilomorpha Order Mugiliformes: 1 family, 72 species, mullets.

The mullets are a family of near-shore, marine and freshwater fishes of considerable economic importance. Many mullets feed on organic silt and minute plants, an unusual food type among fishes.

Series Atherinomorpha Order Atheriniformes: 6 families, 312 species, including rainbowfishes and silversides Order Beloniformes: 5 families, 227 species, including needlefishes, flyingfishes, and halfbeaks Order Cyprinodontiformes: 9 families, 1013 species, including topminnows, killifishes, livebearers, and pupfishes.

Atherinomorphs are shallow-water, marine or freshwater fishes that live near the surface. Many atherinomorphs bear live young. Atheriniforms include the melanotaeniid rainbowfishes of Australia and New Guinea, in which males have brighter colors and longer fins than females, traits that make them popular aquarium species. Silversides are widespread, freshwater and marine schooling fishes and include the grunions of southern and Baja California that ride waves up beaches on dark nights to spawn in wet sand biweekly during the summer. Beloniforms are predominantly silvery, marine fishes active at and sometimes above the surface of the water. The lower lobe of the tail in flyingfishes is relatively long and is used to scull rapidly during takeoff. Many cyprinodontiforms, although basically freshwater fishes, tolerate considerable salinity and hence occur in streams on isolated oceanic islands. The rivulines of South America and Africa live only 1 year, laying eggs that survive in the dried bottoms of pools and that hatch with the next season's rains. *Kryptolebias marmoratus* of south Florida and the West Indies is the only truly hermaphroditic fish, fertilizing its own eggs. The livebearers include the mollies, platys, guppies, and swordtails of the aquarium trade. Some species are all female. Many cyprinodontid pupfishes are tolerant of extreme water conditions and consequently can live in saltmarsh and desert conditions. However, they cannot tolerate total desiccation, which has endangered many desert species that have to compete with humans for water. The Devil's Hole Pupfish has the smallest known range of any fish species – one shallow shelf in a single spring in Death Valley, California. Other pupfish relatives inhabit Lake Titicaca, which at an elevation of 4570 m in the Andes Mountains is the highest natural lake with fishes.

Series Percomorpha Order Stephanoberyciformes: 9 families, 75 species, including whalefishes.

Order Beryciformes: 7 families, 144 species, including flashlight fishes, roughies, and squirrelfishes.

Order Zeiformes: 6 families, 32 species, including dories and boarfishes.

Order Gasterosteiformes: 11 families, 278 species, including sticklebacks, pipefishes, seahorses, seadragons, trumpetfishes, and shrimpfishes.

Order Synbranchiformes: 3 families, 99 species, including swamp and tiretrack eels.

Order Scorpaeniformes; 25 families, 1477 species, including scorpionfishes, rockfishes, sea robins, sablefishes, greenlings, sculpins, Baikal oilfishes, and lumpfishes.

Order Perciformes (10,000 species).

Suborder Percoidea: 79 families, 2860 species, including snooks, temperate basses, sea basses, centrarchid sunfishes, black basses, darters, perches, cardinalfishes, bluefishes, remoras, dolphinfishes, jacks, pompanos, snappers, grunts, croakers, drums, goatfishes, archerfishes, butterflyfishes, and angelfishes.

Suborder Elassomatoidei: 1 family, 6 species, pygmy sunfishes.

Suborder Labroidei: 6 families, 2200+ species, including cichlids, surfperches, damselfishes, wrasses, and parrotfishes.

Suborder Zoarcoidei: 9 families, 340 species, including eelpouts, gunnels, and wolffishes.

Suborder Notothenioidei: 8 families, 125 species, including icefishes and Antarctic dragonfishes.

Suborder Trachinoidei: 13 families, 212 species, including sand lances, weeverfishes, and stargazers.

Suborder Blennioidei: 6 families, 818 species, including clinids and blennies.

Suborder Icosteoidei: 1 family and species, Ragfish.

Suborder Gobiesocoidei: 1 family, 140 species, clingfishes.

Suborder Callionymoidei: 2 families, 182 species, dragonets.

Suborder Gobioidei: 8 families, 2211 species, including sleepers and gobies.

Suborder Kurtoidei: 1 family, 2 species, nurseryfishes.

Suborder Acanthuroidei: 6 families, 129 species, including spadefishes, scats, rabbitfishes, Moorish Idol, and surgeonfishes.

Suborder Scombrolabracoidei: 1 family and species, the Scombrolabracid.

Suborder Scombroidei: 4 families, 114 species, including barracudas, mackerels, tunas, Swordfish, and billfishes.

Suborder Stromateoidei: 6 families, 70 species, including driftfishes and butterfishes.

Suborder Anabantoidei: 3 families, 120 species, gouramies.

Suborder Channoidei: 1 family, 29 species, snakeheads.

Suborder Caproidei, 1 family, 11 species, boarfishes.

Order Pleuronectiformes: 14 families, 678 species, including flounders and soles.

Order Tetraodontiformes: 9 families, 357 species, including triggerfishes, boxfishes, trunkfishes, cowfishes, puffers, porcupine fishes, burrfishes, and ocean sunfishes.

The percomorphs constitute by far the largest taxonomic group of fishes, far too many to deal with in any detail. What follows is a very brief overview of some of the more interesting orders, suborders, and families.

Beryciforms are shallow- to moderate-depth, often red, almost always nocturnal fishes, including the reef-dwelling squirrelfishes. Also included is the commercially important Orange Roughy, *Hoplostethus atlanticus*, of high-latitude, southern ocean regions. Orange Roughies are being overexploited because they are slow growing and long-lived, taking more than 20 years to mature and reaching ages of more than 100 years. Zeiforms include commercial species such as the European John Dory, *Zeus faber*.

Gasterosteiforms are small marine, estuarine, and freshwater fishes with dermal armor plating. Sticklebacks are well-studied fishes that frequently form distinct, isolated populations characterized by unusual spines, plates, and behavior. The suborder Syngnathoidei includes the bizarre pegasid seamoths and syngnathid pipefishes, seadragons, and seahorses. Syngnathid pipefishes and seahorses are the only vertebrates in which the female impregnates the male by laying eggs in his brood pouch, which he then fertilizes and raises until hatching. Seahorses are heavily overfished for medicinal uses and the aquarium trade.

Synbranchiforms are primarily freshwater, eel-like, often airbreathing fishes. Swamp eels have recently been introduced into the southeastern United States and are of major concern as potential invading predators in river systems.

Scorpaeniforms are predominantly marine fishes, with spiny heads and sometimes venomous fin spines (i.e., stonefishes, scorpionfishes, and lionfishes) (**Figure 14**). The sebastine rockfishes are a diverse, commercially important, long-lived, and overfished group of the temperate North Pacific, some of which live to be over 200 years old. Hexagrammid greenlings are littoral zone (inshore and shallow water) and kelp-associated fishes endemic to the North Pacific, including the highly edible Lingcod, *Ophiodon elongatus*. The suborder Cottoidei includes many freshwater species, including the cottid sculpins of North American streams and the highly divergent (pelagic and livebearing) comephorid oilfishes of Lake Baikal in northern Asia. The Cabezon of the Pacific coast of North America is unusual in having toxic eggs, whereas the lumpfish of the North Atlantic produces valuable caviar and has consequently been overfished.

The Perciformes are the largest order of percomorphs, with 160 families and over 10,000 species, including most marine and freshwater fishes of littoral zones. Perciforms reach their greatest diversity on coral reefs, but they are also highly diverse in rivers, streams, and lakes. Coral reef perciforms include six of the eight largest fish families (gobies, wrasses, sea basses, blennies, damselfishes, and cardinalfishes). Two other large families, cichlids and croakers, are characteristic of tropical lakes and near-shore temperate marine habitats, respectively.

The largest perciform suborder is the Percoidei. Many percoids are bass-like fishes. Centropomids are large predatory fishes and include the snooks of tropical America, the Barramundi of Australia, and the Nile Perch of Africa. Nile Perch are an introduced predator in Lake Victoria, in which they are thought to have extinguished perhaps hundreds of endemic cichlids. Moronid temperate basses include the Striped and White bass of North America. The sea bass family Serranidae is one of the largest families (475 species) that range in size from small (<5 cm) anthiines and hamlets to giant groupers (3 m long, 400 kg) (**Figure 15**). Sea basses also include commercially important hinds, coneys, Gag, and Scamp. Many serranids are hermaphroditic, usually maturing first as female and then later becoming male, although some hamlets are both male and female simultaneously.

The Centrarchidae contains the sunfishes, crappies, rockbasses, and black basses of North American fresh waters. Largemouth Bass have been introduced extensively as a sport fish. At least 201 species make up the family Percidae, 187 of

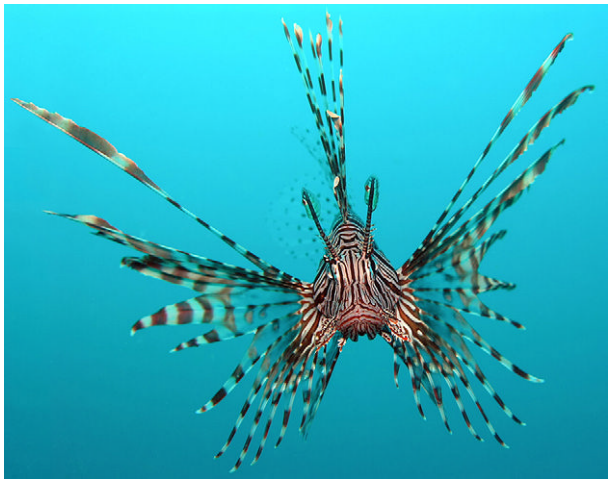


Figure 14 Lionfishes, *Pterois*, belong to the order of scorpionfishes, many of which have toxic spines. Lionfishes are native to Indo–Pacific regions. They have few natural predators, which has contributed to their success as an invasive introduced predatory species on tropical western Atlantic reefs. Photo by Jens Petersen, from Wikimedia Commons, File:Pterois_volitans_Manado-e_edit.jpg.

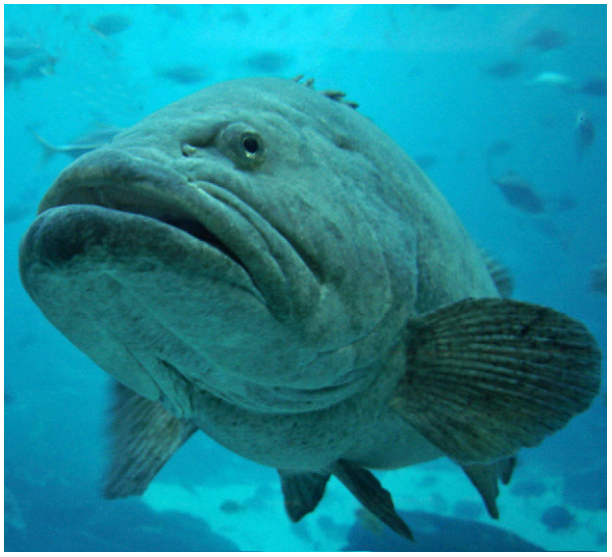


Figure 15 Large groupers, such as this Giant Grouper, *Epinephelus lanceolatus* from the Indo–Pacific, grow as large as 2.7 m long and can weigh over 400 kg.

which occur in North America. Percids include the Yellow and Eurasian perch, Walleye and Sauger (=pikeperches), and about 184 species of small, stream-dwelling, spectacularly colored, and often imperiled darters (Jelks *et al.*, 2008).

Apogonid cardinalfishes are diverse (273 species), small (<10 cm), nocturnal coral reef fishes. Cardinalfishes mouthbrood their eggs, an unusual trait among reef species. The voracious, schooling pomatomid Bluefish, *Pomatomus saltatrix*, occurs in most temperate and semitropical oceans except the eastern Pacific. In echeneid remoras or sharksuckers, the modified first dorsal fin forms a suction disk that is used to cling to various hosts. The coryphaenid dolphinfishes or mahi mahis are two species of open-water, pelagic predators that often associate with floating structure. They are also one of the few marine pelagic fishes that is successfully aquacultured. Carangid scads, jacks, pompanos, and amberjacks are a large family (140 species) of tropical near-shore predators.

Lutjanid snappers include 105 species of generally carnivorous marine fishes. Snappers are usually near bottom dwellers (i.e., Gray Snapper, Red Snapper, and Mangrove Snapper), but some are water column zooplanktivores (Vermilion Snapper).

Haemulid grunts are moderate-sized coral reef fishes and are most diverse in the New World tropics. They are seen most often in their daytime resting schools around coral heads; at night they disperse to feed in surrounding reef and grass areas. Sciaenid croakers (270 species) have chin barbels and a muscularized gas bladder used for sound production. Sciaenids are a widespread family that is particularly diverse in the southeastern US, but representatives occur widely in tropical marine and freshwater habitats. Species include Red Drum (spot tail bass), Black Drum, croakers, weakfish, sea trouts, kingfishes, White Sea Bass, corbinas, and the endangered Mexican Totoaba, one of the few unquestionably imperiled marine fishes. The range of the Freshwater Drum, *Aplodinotus grunniens*, includes much of eastern North America into Central America. Another bottom-oriented family is the tropical reef-dwelling mullid goatfishes, which have movable, muscularized chin barbels. Toxotid archerfishes are Indo–Pacific, brackish-water fishes well-known for their ability to shoot droplets of water that knock insects out of overhanging vegetation.

Two closely related, colorful reef families are the chaetodontid butterflyfishes (122 species) and the pomacanthid angelfishes (74 species). Both families are most diverse in the Indo–Pacific region. Butterflyfishes feed on coral polyps, small invertebrates, tube worms, or zooplankton, whereas angelfishes eat attached invertebrates such as sponges, tunicates, and anthozoans.

Elassomatoid pygmy sunfishes are a suborder of miniature (20- to 45-mm), colorful freshwater swamp dwellers of the southeastern United States.

Several very speciose families belong to the suborder Labroidei. Tropical marine families include the pomacentrid damselfishes (315 species), which are small, colorful, usually herbivorous, territorial reef dwellers. Some are zooplanktivores (*Chromis* and anemonefishes). Several occupy temperate regions (e.g., the Garibaldi of California, see **Figure 19**). The largest labroid family is the mostly tropical, reef-dwelling wrasses and parrotfishes (Labridae, 500 species). Wrasses range in size from the 5-cm-long cleaner wrasses, Slippery Dicks, and Blueheads to the 2.3-m-long Maori or Humphead Wrasse of the Indo–Pacific, which is hunted unmercifully for the live fish restaurant trade. Cool temperate species include the California Sheephead and Senorita, western Atlantic Tautog and Cunner, and the eastern Atlantic Cuckoo Wrasse. Many wrasses change sex from female to a more colorful male. The parrotfishes (83 species) are almost exclusively coral reef dwellers, best known for their fused parrot-like teeth that are used for biting off algal and coral pieces, which are then crushed in the massive pharyngeal (throat) jaws. The embiotocid surfperches are 23 species of deep-bodied, temperate (mostly eastern Pacific) fishes associated with kelp beds and rocky reefs. They are livebearers, feeding on zooplankton or small invertebrates. The largest labroid family is the freshwater cichlids, with more than 2000 species (Barlow, 2000). Cichlids are chiefly tropical, South American and African fishes, with a few species that occur further north (the Rio Grande Cichlid is found in south Texas). Central and South American species include freshwater angelfishes, discus, Oscars, Convict Cichlids, and Peacock Bass. Most cichlids occur in Africa, where they are particularly speciose in the African Great Lakes and are threatened by introduced predators such as Nile Perch (**Figure 16**). African tilapias and other cichlids have been deliberately or accidentally introduced into Florida, California, and Hawaii as a byproduct of aquaculture activities.

Fishes of the suborder Zoarcoidei are all eel-like, bottom-living, marine, cool- to cold-water species. They range in size from the small, intertidal pricklebacks and gunnels to the livebearing eelpouts, some of which live 3000 m below the surface. The large (to 2.5 m long) anarhicadid wolffishes and wolf eels of shallow North Pacific and Atlantic waters are anatomically and ecologically similar to moray eels.

Icefishes (suborder Notothenioidei) are mostly Antarctic, mostly benthic fishes that live under the ice and have antifreeze compounds in their blood. The crocodile icefishes lack red blood cells, hemoglobin, and myoglobin, and hence have colorless blood and flesh.

Trachinoids are marine, generally benthic fishes that tend to bury themselves in sand. Ammodytid sand lances are small, elongate, and abundant zooplankton feeders that spend their nights buried. Trachinid weeverfishes occur in the eastern Atlantic and Mediterranean and have venomous opercular and dorsal spines. Uranoscopid stargazers emit strong pulses of electricity (up to 50 V) from highly modified eye muscles.

Blennioids are small marine fishes that usually associate with structure. Chaenopsid pikeblennies and tubeblennies often live inside corals and worm tubes. Combtooth blennies are diverse (360 species) small fishes in tropical and subtropical waters; they scrape algae with their comb-like teeth.

The suborder Icosteoidei consists of the peculiar North Pacific Ragfish, *Icosteus aenigmaticus*, which has a largely uncalcified, cartilaginous skeleton and is a preferred food of sperm whales. The small, marine and freshwater gobiesocid clingfishes are shallow-water and even amphibious fishes often found in high-energy wave zones. Their pelvic fins are modified into a sucking disc.

Gobioids are usually small, benthic, often abundant fishes. The eleotrid sleepers are small to medium (to 60 cm)



Figure 16 Cichlid fishes in the Great Lakes of Africa are highly diverse ecologically and anatomically, with more than 1000 species in the major lakes. These cichlids at the Georgia Aquarium in Atlanta are representative of what one might see diving in Lake Malawi, Africa.

estuarine and stream fishes in tropical and subtropical areas, often on islands. The gobiid gobies are incredibly diverse (about 2200 species). Many gobies have a suction disk formed by fused pelvic fins. The family includes the amphibious mudskippers. Gobies range in size from tiny pygmy gobies (i.e., 8- to 10-mm adults) to a comparative giant, the western Atlantic Violet Goby, which is a purplish eel-like fish 50 cm long. Round Gobies have recently been introduced into North American Great Lakes from southern Europe and are spreading rapidly; other introduced goby species are now extremely common in San Francisco Bay.

The suborder Acanthuroidei includes the ephippid spadefishes of Atlantic reefs and beaches as well as the Indo–Pacific scatophagid scats, which get their name from feeding on the feces of other animals. Rabbitfishes are Indo–Pacific reef, grassbed, and estuarine herbivores that are convergent in many ways with some butterflyfishes, a description that also applies to the Moorish Idol (Zanclidae). The 72 species of usually herbivorous acanthurid surgeonfishes, unicornfishes, and tangs have a knife blade on the caudal peduncle.

The suborder Scombroidei contains some of the largest and most spectacular marine fishes. Twenty species of barracudas inhabit tropical and subtropical oceans almost worldwide. The gempylid snake mackerels (24 species) are pelagic and deep-water predators, including the cosmopolitan Oilfish, *Ruvettus pretiosus*, a large (1.8 m, 45 kg) predator of moderate depths. An active fishery for oilfish in the Comoro Islands captures endangered Coelacanths as by-catch. The scombrid mackerels and tunas are quintessential open-sea predators, with streamlined bodies and a physiology geared to a high-speed lifestyle. They range from relatively small, 50-cm mackerels to giant Bluefin Tuna (4 m, 500 kg). Most are schooling fishes of tremendous commercial importance; a single, 340 kg Bluefin Tuna sold for $400,000 in Tokyo in 2011. The temperate and warm-temperate xiphiid Swordfish and the more tropical istiophorid sailfishes, spearfishes, and marlins have an elongate upper jaw bone that forms the bill. It is used as a spear, a cutlass, or a billy. Swordfish grow to 530 kg, whereas Blue and Black marlin grow to 900 kg. Swordfish have been heavily overfished, particularly in the Atlantic, although revised management plans have reduced fishing pressure and the species has shown signs of a comeback.

Labyrinth fishes (Suborder Anabantoidei) have an auxiliary breathing structure in the gill chamber for aerial respiration. Anabantid climbing gouramies are African and Asian freshwater fishes that can move across wet ground and reportedly up wet tree trunks. The Kissing Gourami is the sole member of the family Helostomatidae. The belontiid gouramies, fighting fishes (bettas), and paradise fishes have elongate pelvic fin rays that serve as feelers. Bettas (Siamese Fighting Fish) have been bred to battle like fighting cocks, placing them among the few fishes that have been cultured for purposes other than food, appearance, or research.

Pleuronectiform flatfishes are distinctive, compressed, benthic fishes that have both eyes on the same side of the head. Many flatfishes are important commercially (e.g., dab, flounders, halibuts, plaice, sole, tonguefishes, turbots, and whiffs). Paralichthyids include the Summer Flounder and California Halibut, the latter reaching 1.5 m and 30 kg. The pleuronectid righteye flounders include the Atlantic and Pacific halibuts. Pacific Halibuts may live 40 years and attain lengths of 3 m and masses of 200 kg. The fishery for Pacific Halibut in the North Pacific is a well-regulated, sustainable enterprise.

The most advanced bony fishes are in the order Tetraodontiformes, an almost entirely marine order of medium-sized fishes with thick, leathery skin and with scales often

modified into spines or bony plates. In balistoid triggerfishes and filefishes, the long, rigid first dorsal spine is locked erect by an interaction with the shorter, second spine. Ostraciid boxfishes are encased in a triangular or rectangular bony box, with just the fins and caudal peduncle emerging. Puffers and ocean sunfishes lack true teeth. Instead, the jaw bone has a cutting edge that looks like separated teeth or is fused into a parrot-like beak. Diodontid porcupine fishes inflate their body by filling the stomach with water, a process that also helps erect and interlock their body spines. Tetraodontid puffers concentrate a powerful and potentially fatal toxin, tetrodotoxin, in their viscera, which adds to the allure of eating puffers in licensed Fugu restaurants in Japan. The Ocean Sunfish, *Mola mola*, is one of the world's heaviest fishes at 1000–2000 kg, producing as many as 300 million eggs. All this biomass is supported on a diet of jellyfishes (**Figure 17**).



Figure 17 A Giant Ocean Sunfish, *Mola mola*, being cleaned by an adult Emperor Angelfish, *Pomacanthus imperator*, on a reef in Bali, Indonesia. Photo courtesy of Ali Watters, www.mytb.org/Ali

Geographic Diversity

Overview

Fishes occur just about everywhere water occurs in its liquid state, is available through most of the year, and remains below 40 °C. A major zoogeographic distinction can be made between marine and freshwater fishes, with substantial overlap occurring where intermediate salinities occur. Many fishes are restricted to pure fresh water (little or no salinity), many are restricted to normal oceanic salinity (about 35 parts per 1000 salt in water), some occur in both habitats at different times of their lives or of the year, and some occur and are even restricted to areas of intermediate salinity, such as estuaries.

In terms of numbers, about 58% of all fishes are marine and 41% live in fresh water, with the remaining 1% moving regularly between the two salinity designations (**Figure 18** and **Box 2**). Among the 13,000 freshwater species, 80% are primary or obligatory freshwater fishes and are intolerant of even moderate salinities. The remaining 20% can tolerate some salinity and hence inhabit upper estuarine areas or can cross through near-shore ocean regions to move from one river

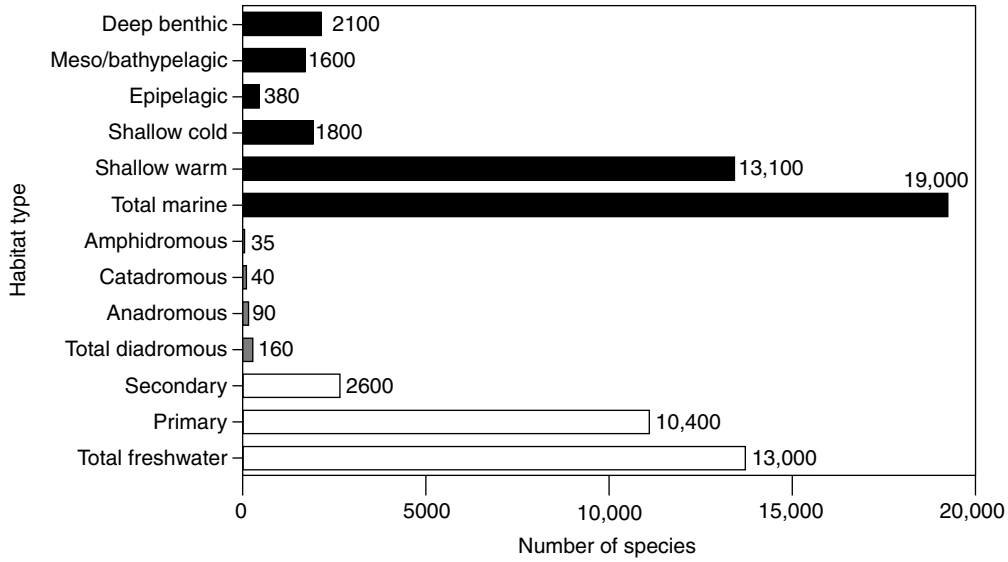


Figure 18 The global biodiversity of marine, diadromous, and freshwater fishes. Approximate numbers of species are given for each habitat type and its subdivisions.

Box 2 Why are there so many freshwater fishes?

The high global diversity of freshwater fishes (**Figure 18**) is at first surprising. Fresh waters make up only about 0.009% of the earth's water, which means that almost half of all fish species live in less than 1% of the world's water. This 7,500-fold discrepancy in biodiversity per unit volume is probably best explained by the relative productivity and isolation of freshwater bodies. Most freshwater habitats are relatively shallow and receive ample sunlight as well as nutrients running off from adjacent land. Hence freshwater habitats are relatively productive and capable of sustaining abundant life. Most of the ocean, in contrast, is deep, dark, and nutrient-poor. Given that 81% of marine diversity occurs in shallow regions, a relationship between water depth and diversity in fresh waters is not surprising. Adding to the influence of available food is the comparative isolation of most freshwater habitats. Lakes are often created and affected by climatic and geologic forces (e.g., drought, floods, landslides, earthquakes, uplifts) that separate them from other systems. Every lake then can be relatively isolated from other lakes, which means that genetically distinct populations can evolve into new species and that little genetic mixing occurs between lakes. Small streams are separated from each other by larger rivers which are barriers to the movement of small fishes, and large rivers are separated by oceans. As a result, freshwater habitats are perfect habitats for the speciation process. Oceans in contrast are largely continuous habitats that are connected by currents, and ocean fishes typically produce larvae that float for several weeks or months on these currents. Hence genetic exchange is common and opportunities for speciation are not as great. Oceanic basins have relatively distinct faunas, but connectedness within basins discourages the kind of genetic isolation needed for speciation of the sort seen in lakes and rivers.

basin to another. Among the 19,000 marine fishes, the vast majority (69%) live in shallow, warm areas such as coral reefs. The remaining marine species are divided fairly evenly among shallow, cold, deep, open-ocean, and deep-bottom areas (about 10%–15% each). About 2% of marine fishes live in near-surface, open-sea (pelagic) habitats. The approximately 160 diadromous species that live in different salinity regions at different times of their lives are divided among three groups. Anadromous fishes (54%) live most of their lives in the ocean but then migrate to fresh water to spawn; this group includes lampreys, sturgeons, shad, and salmons. Catadromous fishes (25%) spend most of their lives in fresh water and migrate to the sea to spawn; included here are freshwater eels, mullets, and temperate basses. Amphidromous fishes (21%) move between fresh and salt water, but migration to the spawning habitat occurs long before the fishes actually spawn; examples include gobies, sleepers, and galaxiids.

Freshwater Diversity

The world's freshwater habitats occur in six major zoogeographic regions or realms that correspond approximately to continental distributions, with important exceptions. Each region has a fairly distinct fish fauna (again with some exceptions and shared elements) (Berra, 2007).

The Nearctic Region

The Nearctic region consists of subtropical, tropical, temperate, and arctic North America. The region stretches from the Mexican Plateau to northern Canada and Alaska. The Nearctic contains 14 families of primary freshwater fishes, with about 1000 species. The most diverse families are minnows, suckers, North American catfishes, perches (and darters), and sunfishes. Other important families include the lampreys, gars, salmons (many of which are anadromous), and whitefishes; sculpins, which are freshwater species in a primarily marine family (= marine derivatives); pickerels and mudminnows; killifishes; and livebearers. The Nearctic is further subdivided into three subregions: the Arctic–Atlantic (with six provinces), the Pacific (with seven provinces), and the Mexican Transition subdivision. Eleven major river systems drain the region; major lakes are abundant, the largest being the five Laurentian Great Lakes (Ontario, Erie, Huron, Michigan, and Superior).

The Neotropical Region

The Neotropical region contains South America and Middle America. It is the most speciose region of the world in terms of freshwater fishes, with 32 families and more than 4500 species. Particularly diverse groups include the colorful characiforms (1500 species of tetras, piranhas, characins, and freshwater hatchet fishes), 14 families and 2500 species of catfishes, 5 families and 134 species of gymnotiform South American electric knifefishes, and more than 150 species of cichlids. Several secondary freshwater and marine derivative groups are included: freshwater stingrays, herrings, silversides, needlefishes, killifishes, and croakers. Many species remain to be discovered and described, particularly in South America. The Neotropical region has been further divided into eight subdivisions with fairly distinctive faunas. Eight major river systems drain the region; major lakes include Lake Titicaca, the world's highest fish-containing lake.

The Palearctic Region

The Palearctic region encompasses Eurasia, including Europe, northern Africa, and Asia north of the Oriental region. Twenty-seven families and about 600 species of temperate freshwater fishes occur in the region, dominated by minnows and loaches but also perches, pickerels, sturgeons, salmons, sculpins (including the Lake Baikal endemics), and 10 species of catfishes in four families. Endemism increases to the south, as is also the case in the Nearctic region (Kottelat and Freyhof, 2007). The Palearctic and Nearctic regions share numerous families and genera (sturgeons, paddlefishes, minnows, smelts, salmons, pikes, mudminnows, and perches) but only a few species occur in both (i.e., Northern Pike, Longnose Sucker, Burbot, Threespine Stickleback, and Fourhorn Sculpin). The region is sometimes subdivided into six subregions based on faunal groupings. Ten major river systems drain the region; major lakes include the Black and Caspian Seas and Lake Baikal, the world's oldest and deepest lake.

The African or Ethiopian Region

The African or Ethiopian region is second to the Neotropics in freshwater fish diversity, with 36 families and more than 2000 species of primary and secondary freshwater fishes. The African region includes all the African continent south of the Sahara Desert, plus the large island of Madagascar

with its endemic fauna. Half of the fishes are in the superorder Ostariophysi, including 300 minnows, 190 characiforms, and 360 catfishes in six families. Other diverse groups include killifishes and topminnows, elephantfishes and other osteoglossiforms, and cichlids. As many as 1000 cichlid fishes may occur in the three African Great Lakes of Lake Victoria, Lake Tanganyika, and Lake Malawi, with more cichlids in smaller surrounding lakes and rivers (see **Figure 16**). Four lungfishes and all 11 polypteriform bichirs occur in Africa. Ten to 12 zoogeographic provinces are recognized, with six major river drainages and numerous lakes including the African Great Lakes of Victoria, Malawi, and Tanganyika.

The Oriental Region

The Oriental region includes eastern Iran, India and Sri Lanka, China south of the Yangtze River, Southeast Asia, and the large island regions of Taiwan, the Philippines, and the East Indies/Indo–Malayan Archipelago. The Oriental region contains 43 families of primary and secondary freshwater fishes. Most diverse are the minnows, loaches, and 12 families of catfishes; clariid walking catfishes and bagrid catfishes are particularly diverse. Other important groups include algae eaters, river loaches, snakeheads, spiny eels, labyrinth fishes and gouramis, a few cichlids, and archerfishes. The Oriental region shares many families with the Palearctic to the north and the Ethiopian to the west but shares few families with the Australian region to the southeast. The Oriental is often subdivided into two major subregions: Peninsular India with more than 700 species and Southeast Asia with more than 1000 species. Each subregion has two major river drainage systems; large lakes are uncommon. Southeast Asia is sometimes divided further into five zoogeographic regions.

The Australian Region

The Australian region (New Guinea, Australia, New Zealand, and Oceania) is relatively depauperate in true freshwater fishes, and in fact all but three of the freshwater fishes in the region are members of families obviously derived from marine groups. The northwestern border of the region, and the practical limit of primary freshwater fishes, is dramatically delineated by an ocean boundary that lies southeast of Java, Borneo, and Sulawesi and is known as Wallace's or Weber's Line. Nineteen families and about 210 species occur primarily in fresh water in the region, but only the Australian lungfish and two species of bonytongue saratogas are true freshwater fishes (another 33 families and 150 species of marine fishes frequently enter fresh waters in Australia). Other important families, many with species endemic to specific regions, include lampreys, river eels, herrings, two families of catfishes, southern smelts and graylings, Salamander fish and other galaxiids, silversides, rainbow fishes, Barramundi, grunters, glassfishes, temperate perches or basses, sleepers and gobies, and Torrentfish (in New Zealand). One major river system occurs on New Guinea (the Fly) and two on Australia (Darling and Murray); permanent, large lakes are rare.

Marine Diversity

Delimiting zoogeographic regions in the world's oceans is complicated by depth, currents, and geographic locales; different faunal breaks occur depending on near-shore, pelagic, or deepsea environments. The greatest fish diversity and the greatest geographic differentiation occur in near-shore, continental shelf (to about 100 m depth) regions. These regions are separated by continents, by large expanses of open ocean, and by currents that differ in temperature from that of the region in question. Temperature zones divide the seas into tropical, temperate, boreal, and polar regions. In addition, different faunal groupings apply to pelagic fishes and to fishes of the deep sea (Briggs, 1995).

The Indo–West Pacific Region

The Indo–West Pacific region includes shallow tropical seas that extend from South Africa and the Red Sea eastwards through the Indo–Malayan area and Australia to Hawaii and Easter Island; it also includes Micronesia, Melanesia, and Polynesia. The Indo–West Pacific is by far the most species-rich marine area, containing 3000–4000 tropical fish species. It is similarly diverse in sea snakes and many invertebrate taxa such as reef-building and soft corals, mollusks, tube worms, and echinoderms, although these numbers describe life in the upper 50–75 m of the reef that are accessible to scuba divers. As new exploration technologies allow deeper exploration, hundreds more fish species have been discovered. Regardless, the Indo–West Pacific is considered the center of evolution for many of the common coral reef fish families that occur in other tropical regions. Only a few families are endemic to the Indo–West Pacific (e.g., sillaginid whitings and rabbitfishes). Common families include moray eels, squirrelfishes, sea basses, grunts, snappers, cardinalfishes, butterflyfishes, angelfishes, damselfishes, wrasses, parrotfishes, surgeonfishes, gobies, triggerfishes, and boxfishes. Barriers to movement of Indo–West Pacific fishes are cool waters to the north and south and a vast expanse of open, deep water to the east known as the Eastern Pacific Barrier. Eight different provinces have been recognized in the Indo–West Pacific, separating it into a huge Indo–Polynesian region and seven smaller areas.

The Eastern Pacific Region

The Eastern Pacific region, with approximately 800 fish species, runs from southern Baja California to Ecuador, its northern and southern limits defined by the coldwater California and Peru currents. Despite its location in the Pacific Ocean, the Eastern Pacific is faunistically more similar to the tropical Atlantic, containing many species that are almost indistinguishable from Atlantic forms. The two oceans mixed before the Panamanian Isthmus formed and the two areas still share 12 species, despite 3 million years of physical separation. Most families are less diverse here than in the Western Atlantic, with the exceptions of sea catfishes, croakers, and herrings. Dactyloscopid sand stargazers occur here and in the Atlantic, but not in the Indo–West Pacific. Sixty-two Indo–West Pacific species have managed to cross the Eastern Pacific Barrier. Three provinces – Mexican, Panamanian, and Galapagos – are recognized.

The Western Atlantic Region

The Western Atlantic region is the second most diverse oceanic area, containing 1200 fish species. It includes Bermuda (which, although at 32° N, sits in the tropical Gulf Stream), southern Florida, the Bahamas Bank, the Caribbean Sea, and tropical and temperate portions of South America. Most of the families that occur in the Indo–West Pacific also occur in the Western Atlantic; a few families are more diverse here, such as grunts and toadfishes. Strong currents of warm water separate the Western Atlantic fauna from colder waters along much of its boundaries. It is subdivided into Caribbean, Brazilian, and West Indian provinces.

The Eastern Atlantic Region

The Eastern Atlantic region is a relatively small area along the west coast of Africa from Senegal to Angola and extending out to oceanic islands such as Ascension and St. Helena. Tropical marine fishes here are limited by cool-water currents impinging from both the north and the south as well as by substantial freshwater runoff and sediments from several major west African rivers, all factors that discourage coral reef growth. The region contains "only" 500 shore fishes; most coral reef families occur but are represented by only a few species. Porgies are particularly diverse. No subdivisions are recognized, with the possible exception of the warm-temperate Mediterranean. The Mediterranean Sea contains 540 species, with many species in the same families as those in the Eastern Atlantic. The Mediterranean has the dubious distinction of being the most heavily invaded tropical marine area in the world, with at least 52 alien fish species having moved in from the Red Sea (an Indo–West Pacific subregion) *via* the human-made Suez Canal.

The Arctic Region

The Arctic region encompasses high-latitude (above 60° N) waters of both the Pacific and Atlantic Oceans. It reaches from Nunivak Island, Alaska, northward and across the polar region to Newfoundland and Norway in the northern Atlantic. Of the two polar areas, the Arctic is more diverse. Successful groups include skates, herrings, salmons, smelts, cods, eelpouts, greenlings, sculpins, poachers, snailfishes, pricklebacks, wolffishes, gunnels, and righteye flounders. Diversity within many of these groups is greater in the Pacific than in the Atlantic portions of the region. A total of 415 species occur here. Distribution of many of these families appears to be limited by temperature, with warmer waters and currents to the south determining species' boundaries.

The Antarctic Region

The Antarctic region (above 60° S) has its own distinctive fauna that is restricted to Antarctic waters and the surrounding Southern Ocean, including the cold waters of Australia, New Zealand, and nearby oceanic islands. Forty-nine families and 274 species occur here, 13 families and 174 species of which are identified as Antarctic continent species. A particularly successful group is the notothenioid icefishes and relatives, which account for 55% of Antarctic species. Families include bovichthyids, cod and crocodile icefishes, plunderfishes, and dragonfishes. Non-notothenioids include skates, snailfishes, eelpouts, lantern fishes, eel cods, deepsea cods, and southern flounders.

Temperate Regions

To the north of the Indo–West Pacific lie cooler temperate waters with their own characteristic fish faunas. This area can be divided into four fairly distinct regions according to location and temperature: Japanese warm-temperate and Californian warm-temperate regions and Eastern and Western boreal regions. The warm-temperate areas (from about Hong Kong to Tokyo in the west and from lower Baja California to central California in the east) contain a fauna that fluctuates seasonally, as tropical species move north in the summer and boreal species move south in the winter. Notable families to the west include lizardfishes, flyingfishes, mullets, jacks, sea basses, and croakers and to the east include endemic silversides, sea basses, croakers, damselfishes, wrasses, and flatfishes (**Figure 19**). The more northerly boreal regions (approximately north from central California in the east and Korea in the west) contain similar families but different species. Important families include migratory salmonids, sculpins, rockfishes, snailfishes, greenlings, gunnels, pricklebacks, and righteyed flounders. In the southern Pacific, coldwater currents create at least three temperate faunal regions, with centers of distribution around Australia, New Zealand, and South America.

In the Atlantic, three northern temperate areas occur: the western and eastern Atlantic boreal regions and the Atlantic warm-temperate or Carolinian region. The boreal regions (Newfoundland to Cape Hatteras in the west and British Isles



Figure 19 The garibaldi, *Hypsypops rubicundus*, is a temperate member of the otherwise largely tropical family of pomacentrid damselfishes. It is common in kelp beds of southern California and is the state marine fish.

to northern Europe and Scandinavia in the east) share a fauna of salmonids, cods, sticklebacks, poachers, sculpins, wolffishes, and righteyed flounders, with occasional strays from more southerly waters during warm months. The Carolinian extends from Cape Hatteras south to Florida and also to the Gulf of Mexico, with southern Florida housing tropical species. Common groups are clupeids, sea robins, pipefishes, silversides, needlefishes, killifishes, croakers, lefteyed flounders, and puffers. Temperate faunas also occur in the southern Atlantic, but their areas and diversity are less than those of the southern temperate faunas of the Pacific Ocean. Two recognized regions are the eastern South American and southern African warm-temperate regions. The former region's fauna includes sea catfishes, croakers, herrings, gobies, scorpionfishes, and sea basses; the latter area has many colder water members of Indo–West Pacific families.

Pelagic Regions

The 350 species of pelagic fishes occur in ocean surface waters to a depth of 200 m. This habitat type can be divided into 10 different regions based on faunal differences, with more joining of Southern Hemisphere areas because of the relative lack of large land masses. These regions are Arctic and Antarctic polar, North Pacific cold temperate, North Pacific warm-temperate, tropical Indo–Pacific, North Atlantic boreal, North Atlantic warm-temperate, Atlantic tropical, southern warm-temperate, and southern cold temperate. Many of the world's most important fisheries species occur in pelagic regions, including numerous sharks, sardines and herrings, salmons, codfishes, pollocks, hakes, haddocks, sauries, mackerels, and tunas. In addition, about 100 species of mostly pelagic fishes have a worldwide distribution. This group includes pelagic sharks (White, Whale, Tiger, and perhaps Megamouth), Swordfish, and ocean sunfishes.

The Deep Sea

Waters deeper than 200 m are as much a zoogeographic as a habitat entity. The deepsea is generally divided into three major regions based on depth: the open-water mesopelagic (200–1000 m) and bathypelagic (1000–4000 m) regions and the bottom-associated benthal (200–1000 m) region. Benthal fishes are further divided into benthopelagic fishes that hover just above the bottom and benthic fishes that rest in contact with the bottom. Deepsea fishes are most common in these regions between 40° N and 40° S latitude, approximately between San Francisco and Melbourne in the Pacific and between New York City and the Cape of Good Hope in the Atlantic. Abyssal and hadal (trench) regions deeper than 4000 m are relatively depauperate.

Each region has a characteristic and relatively diverse bony fish fauna consisting of fishes from many different taxonomic groups. The mesopelagic region worldwide contains about 750 species in seven different superorders and nine orders. Despite their lack of relatedness, deepsea fishes share many anatomical and physiological features, suggesting independent, convergent evolution of adaptations to deepwater existence. Mesopelagic fishes typically have photophores (light organs) on their silvery bodies; have relatively large, oftentubular eyes; undergo daily migrations to surface waters to feed at night; and have large mouths and long teeth. Common names of mesopelagic fishes reflect these traits: barreleyes, bristlemouths, dragonfishes, sabertoothfishes, lanternfishes, tubeeyes, and swallowers.

The bathypelagic region is the largest habitat space on the earth, accounting for 88% of oceanic volume. The five superorders, nine orders, and 200 species in the cold, dark bathypelagic region share some traits with mesopelagic fishes but possess them in the extreme. Photophores are concentrated on lures used to attract prey; eyes are often small; mouths are extremely large and teeth very long; stomachs are expandable; bodies are black; and body musculature, bones, and scales are greatly reduced. These traits reflect greater habitat space and increasingly rare feeding opportunities with increasing depth, which select for an increasing need to conserve energy and to be able to take advantage of feeding opportunities. Again, bathypelagic fishes have names indicative of their adaptations: sawtooth eels, gulper eels, swallower

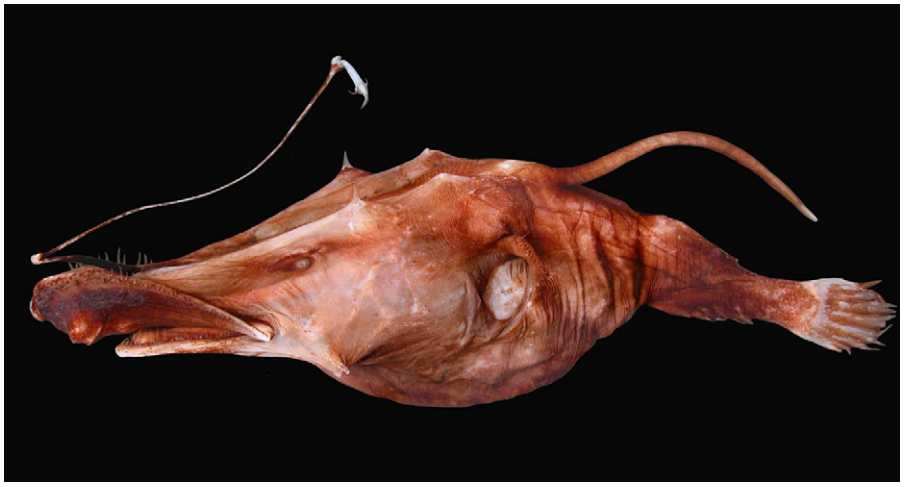


Figure 20 A 15-cm long female Wolftrap Anglerfish, *Lasiognathus amphirhamphus*. Anglerfishes are among the more diverse groups in the bathypelagic region. Many anglerfishes possess a lure at the tip of their modified first dorsal spine used to bring prey within striking distance of a large, expandable mouth. Photo courtesy of Pietsch TW.

eels, dragonfishes, anglerfishes, seadevils, and fangtooths (**Figure 20**. Anglerfish).

Benthal fishes include about 1000 bony fish species in four superorders and nine orders, plus chimaeras and squaloid sharks. Different families inhabit bottom compared to open-water regions. Benthal fishes include greeneyes, tripodfishes, hakes, grenadiers, cuskeels, batfishes, snailfishes, and eelpouts. Although diversity decreases below 1000 m, grenadiers and rattails live between 1000 and 4000 m, tripodfish have been found to 6000 m and snailfishes to 7000 m, and some cuskeels have been found as deep as 8000 m.

Although some differences in species composition occur in different ocean basins or in association with different water masses, deepsea species are relatively cosmopolitan, occurring in several different oceans. One trend is for fishes to occur deeper at lower latitudes, such that species that are bathypelagic near the equator may be mesopelagic at middle latitudes and even epipelagic at the poles.

See also: Fish Conservation. Fish Stocks. Lake and Pond Ecosystems. Ocean Ecosystems. Pelagic Ecosystems

References

Barlow GW (2000) *The Cichlid Fishes: Nature's Grand Experiment in Evolution*. Cambridge, MA: Perseus Publishing.

Barton M (2006) *Bond's Biology of Fishes*, 3rd edn. Stamford, CT: Thomson Brooks/Cole.

Berra TM (2007) *Freshwater Fish Distribution*. Chicago: University of Chicago Press.

Briggs JC (1995) *Global Biogeography*. Amsterdam: Elsevier.

Carrier J, Musick J, and Heithaus M (eds.) (2004) *The Biology of Sharks and their Relatives*. Boca Raton, FL: CRC Press.

Carrier J, Musick J, and Heithaus M (eds.) (2010) *Sharks and their Relatives II. Biodiversity, Adaptive Physiology and Conservation*. Boca Raton, FL: CRC Press.

Helfman GS (2007) *Fish Conservation: A Guide to Understanding and Restoring Global Aquatic Biodiversity and Fishery Resources*. Washington, DC: Island Press.

Helfman GS and Collette BB (2011) *Fishes: The Animal Answer Guide*. Baltimore, MD: Johns Hopkins University Press.

Helfman GS, Collette BB, Facey DE, and Bowen B (2009) *The Diversity of Fishes: Biology, Evolution, and Ecology*, 2nd edn. Oxford, UK: Wiley-Blackwell.

Jelks HL, Walsh SJ, Burkhead NM, *et al*. (2008) Conservation status of imperiled North American freshwater and diadromous fishes. *Fisheries* 33(8): 372–407.

Kottelat M and Freyhof J (2007) *Handbook of European Freshwater Fishes*. Cornol, Switzerland: Publications Kottelat.

Moyle PB and Cech JJ (2004) *Fishes: An Introduction to Ichthyology*, 5th edn. Upper Saddle River, NJ: Prentice-Hall.

Nelson JS (2006) *Fishes of the World*, 4th edn. Hoboken, NJ: John Wiley & Sons.

Paxton JR and Eschmeyer WN (eds.) (1998) *Encyclopedia of Fishes*, 2nd edn. San Diego, CA: Academic Press.

Fish Stocks

Rainer Froese, IFM-GEOMAR, Kiel, Germany
Daniel Pauly, University of British Columbia, Vancouver, BC, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biomass Collective weight or mass of all the members of a given population or stock at a given time, or, on the average, over a certain time period.

Bioquads Occurrence record of organisms, serving as key units for biodiversity research and consisting of four elements (species names, location, time, and source).

Catches The fish (or other aquatic organisms) of a given stock killed during a certain period by the operation of fishing gear. This definition implies that fish not landed, that is, discarded at sea, or killed by lost gear (ghost fishing), should be counted as part of the catch of a fishery.

Ecosystem Area where a set of species interact in characteristic fashion, and generate among them biomass flows that are stronger than those linking that area to adjacent ones.

Recruitment Entry of juvenile fish into the (adult) stock. Recruitment is distinguished from reproduction, because the eggs and larvae that result from fish spawning usually suffer tremendous and largely unpredictable mortalities, thus uncoupling spawning from recruitment.

Trophic level A number indicating the position of a species within an ecosystem through the number of steps linking it to the plants. By definition, plants are TL = 1, herbivores are TL = 2, and so on. Note that trophic levels do not need to be whole numbers; intermediate values occur among omnivorous consumers.

The major adaptations of fishes which determine their spatial distribution pertain to their specific anatomy, reproductive biology, and respiratory physiology. Further, fishing has become increasingly important to the biodiversity of fish, either through its direct impacts (changes of stock size and age structure, and overall biomass reductions, down to extirpation of populations), or by modifying the ecosystems in which fish are embedded. Research on biodiversity of fish (or other organisms) must interpret the huge amount of error-prone observational data in the context of the environmental preferences of the species and of their known native range. Management regimes aiming at preserving fish biodiversity at the level of species, populations, and genes, will have to include much stricter regulation of fishing and the establishment of no-take areas.

Major Adaptations of Fishes

Anatomy and Physiology

With about 32,000 recognized species in over 500 families, fish are the most diverse vertebrate group. However, their watery habitat, while failing to protect them from modern fishing gear, makes it difficult to fully appreciate this diversity, and the extent to which it is now threatened. It is even more difficult for us, as air breathers, to perceive the constraints under which fish, as water breathers, were forced to evolve.

Water is an extremely dense medium, 775 times heavier and 55 times more viscous than air. Further, water contains 30 times less oxygen than air, and this oxygen diffuses 300,000 times more slowly than in air. These physical constraints, which shaped all early life forms, including the jawless predecessors of the fish, the agnathans, are best visualized by describing the major evolutionary trends leading from agnathans to modern fish (**Figure 1(a)**).

The first of these trends was the evolution of jaws from the first upper and lower gill arches of agnathans. This built on the intimate connection, in the most primitive vertebrates, between the feeding apparatus (i.e., the mouth) and the respiratory apparatus (i.e., the gills adjacent to slits on both sides of the anterior part of the alimentary canal). Water-breathing invertebrates lack this close connection between feeding and breathing, one possible reason why even the largest among them (giant squids) cannot reach the mass of the largest fish (34 metric tons, for the whale shark *Rhincodon typus*).

The reorganization of the head of early fish allowed larger gills to evolve, which allowed the higher metabolic rates required for swimming in open waters. This transition was assisted by the gradual loss of the heavy armor protecting the slow, bottom-slurping agnathans. The fine teeth covering the bodies of sharks are vestiges of this armor.

Fast swimming in open water required better fins, both for propulsion and for steering. Propulsion is provided in most fish by oscillations of a caudal fin whose aspect ratio (**Figure 1(c)**) gradually increased toward tunas and other derived, fast-swimming groups with very large gills. Steering, however, is provided by dorsal, pectoral, and anal fins. These fins are stiffened for precise action by hard, bony rays in the most derived fish, the teleosts, whose evolutionary success was further enhanced by a complexly built protrusile mouth that enables capture of a wide range of food items (**Figure 1(b)**).

Subtle anatomical changes in fish can thus create more niches for increasing the numbers of specialists, which then occupy increasing numbers of closely packed niches. Ecosystems in which these changes have run for long periods, undisturbed by physical changes, therefore contain very large numbers of fish species. Their numbers are even larger in areas such as the Great Lakes of Africa and the tropical Indo-Pacific, where changes of water levels have repeatedly isolated basins

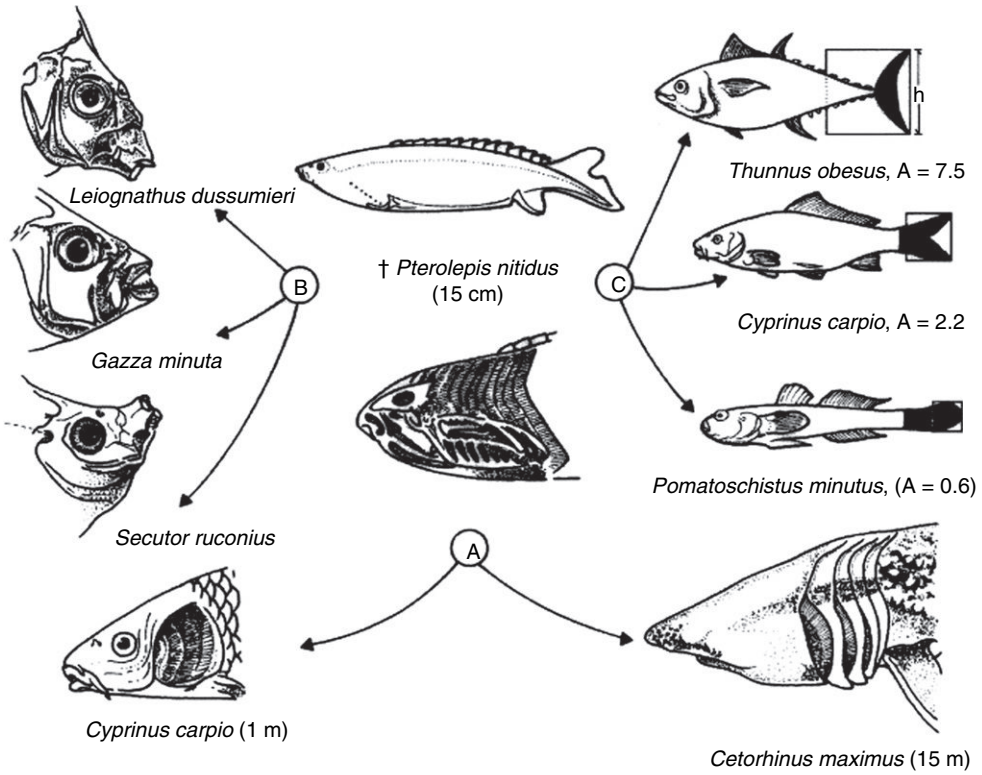


Figure 1 Major evolutionary trends from agnathans to extant fishes. (Note that no direct ancestor–descendant relationships are implied among the groups depicted.) (A) Trends toward larger gills; (B) trends toward efficient jaws; (C) trends toward effective paired and unpaired fins. (Note the aspect ratio of the caudal fin, defined by $A = h^2/s$, where h is the height and s the surface (in black) of the caudal fin, and of which high values define fast, large-gilled continuous swimmers, and conversely for low values.)

and subpopulations, thereby accelerating species differentiation (**Figure 2**).

Reproduction and Recruitment

Though many ancient fishes such as sharks and rays or the coelacanth *Latimeria chalumnae* practice internal fertilization and produce few large eggs or live offspring, more recently evolved fishes produce numerous small eggs that are fertilized externally and develop as part of the plankton, without parental care. The larvae that emerge from these eggs, after less than one day in warm tropical waters and up to two weeks (and more for larger eggs in cold temperate waters) are usually elongated, as befit small, finless zooplankton feeders.

The average zooplankton concentrations that these larvae encounter, even during spawning seasons attuned with zooplankton production cycles, are usually far too low to allow survival of fish larvae, and the overwhelming majority of such larvae perish. Those that tend to survive usually happen to have hatched within plankton-rich water layers. These layers are usually a few centimeters thick and last for only a few days of calm, between wind-driven or other mixing events, such as storms or upwelling pulses that enrich surface waters with nutrients from deeper waters. This implies that large biomasses of fish can build up only when and where the local oceanographic conditions take the form of triads defined by (1) nutrient enrichment, such as generated by wind-driven mixing, (2) high plankton concentration, such as generated by various mechanisms including fronts, and (3) retention of larvae, required to prevent these weak swimmers from drifting away from suitable habitat. In pelagic fishes that build high biomass, for example, the anchovies and sardines in coastal upwelling systems off northwestern and southwestern Africa, Peru, and California, these triads occur only when the coastal winds range from 4 to 6 m s^{-1}. Weaker winds do not generate enough enrichment, and stronger winds disperse the larvae offshore.

Fish have developed several strategies to deal with the uncertain recruitment that results from the triad requirements. One is being small, short-lived, and capable of quickly building up large biomass under favorable environmental conditions. The other is being large, long-lived, and capable of weathering long series of recruitment failures through repeated spawning by old, large, and highly fecund adults. An example of the former strategy is provided by the Peruvian anchovy *Engraulis ringens*, whereas the northern cod, *Gadus morhua*, provides an example of the latter. Yet another strategy is to reduce the dependence on environmental conditions by various forms of parental care, such as nesting and guarding (e.g., in catfishes, family Clariidae), mouth-brooding (e.g., in cardinal fishes, family Apogonidae), and live-bearing (e.g., in ocean perches, genus *Sebastes*).

Another important feature of fish populations is that, contrary to earlier assumptions of homogeneity, most appear to consist of well-differentiated individuals, each aiming to

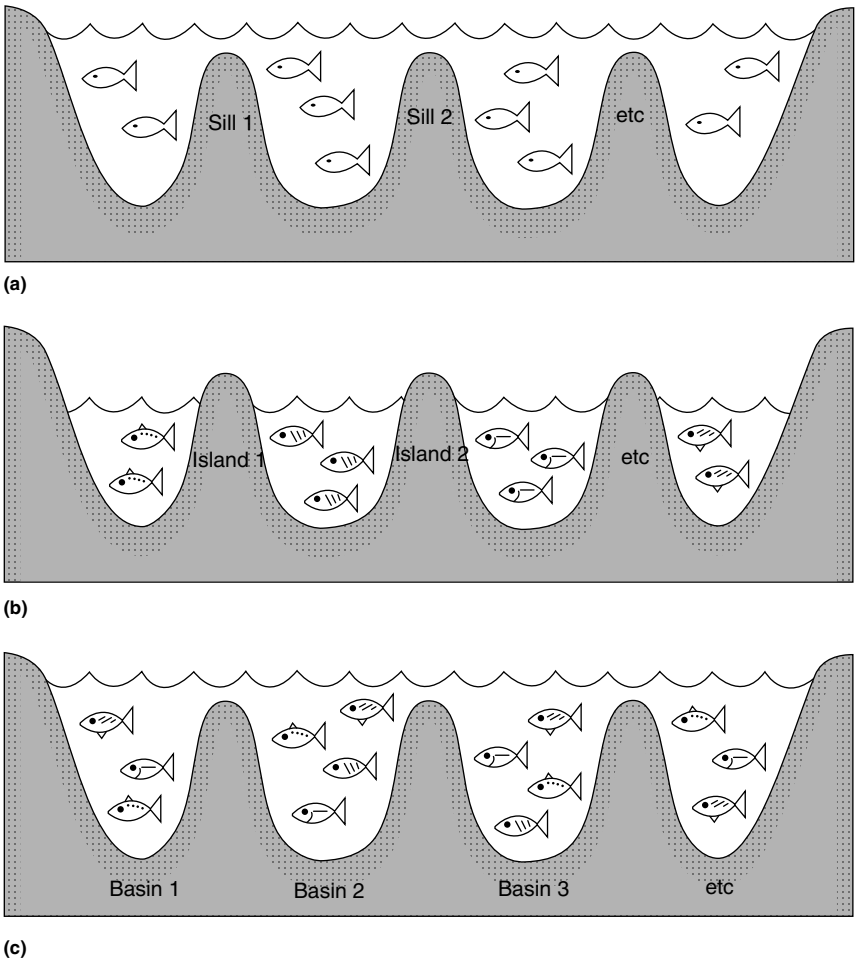


Figure 2 Schematic representation of how changes in water level can multiply, by creating isolated subpopulations, the number of species occurring in a given area. Such a mechanism, driven by repeated climatic changes, is thought to explain the large number of fish species in Southeast Asian marine waters and in the Great Lakes of Africa.

reproduce at the very place where it was hatched. Or, in other words, most migratory fish tend to home. This behavior, well documented only in Pacific and Atlantic salmon (*Oncorhynchus* and *Salmo*, respectively), implies that individual fish, when reproducing, do not seek optimal sites, but rather spawn as close as possible to the site at which they hatched, and to which they are imprinted. This reproductive strategy has proven successful in evolutionary time scales. However, it is not helpful in cases where, for example, climate change or pollution have impacted traditional spawning grounds, or where a stock recovering from overfishing has to rediscover spawning sites that were abandoned during the depleted phase.

Respiratory Constraints to Growth and Related Processes

Basic Geometrical Constraints

Fish growth, as in other animals, requires both food and oxygen, the latter being required to synthesize the substance (adenosine triphosphate (ATP)) that serves as fuel to all organisms. For oxygen to be metabolically available, it must be inside the body of the fish, that is, it must have passed through its gills. Thus, since oxygen cannot be stored inside the fish body (contrary to food, which can be stored as gut contents and as fat), the metabolic and growth rate of fish are largely proportional to the surface area of their gills. So fish that quickly reach large sizes have gills with large surface areas (as in swordfish, billfish, and tunas), and conversely in slow-growing fishes (like groupers). Moreover, gill area per unit of body mass declines with size, because the two-dimensional gill area cannot keep up with the three-dimensional increase in body mass. Hence larger fish dispose of relatively less oxygen to supply their metabolism, the reason why they ultimately stop growing. Furthermore, environmental factors that tend to increase metabolic rate – especially elevated temperature, and other forms of stress – have the effect of reducing the maximum size that the fish of a given population can reach (**Figures 3(a)** and **3(b)**). This is why tropical fish tend to be smaller than their respective cold-water relatives.

Adaptation to Respiratory Constraints

Fish have evolved various strategies and tactics to overcome respiratory constraints. One strategy, illustrated in **Figure 1(b)**,

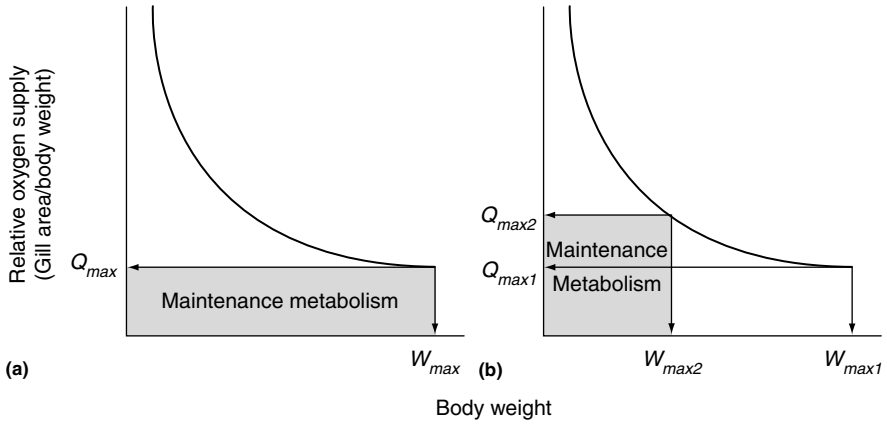


Figure 3 Schematic representation of the relationships linking, in fish, respiratory area (and hence metabolic rate) and maximum body size. (a) As body size increases, gill area per body weight decreases, down to a level when it suffices only for maintenance metabolism. This defines the maximum size that can be reached. (b) Any environmental factor increasing oxygen demand for maintenance (such as elevated temperature) reduces the maximum size that fish can reach.

is to evolve large gills, a route taken by numerous open-water (pelagic) species, culminating not only in tunas but also in plankton feeders such as the basking shark. Another strategy is the evolution of life cycles in which the juveniles migrate to deeper, cooler waters as they grow and then, on maturing, produce eggs that quickly float up to the warmer surface layers, out of reach of the often cannibalistic adults. Such typical cycles are completed by an onshore drift of the larvae to coastal areas, and productive shallow nurseries for the early, voracious juveniles, which again migrate into deeper waters as they grow.

A tactic to accommodate metabolic stress, which is particularly useful in areas with strong seasonal temperature oscillations, is for the feeding adults to store fat during the warmer part of the season (late summer to early fall). Fat requires far less oxygen for maintenance than protein of muscle and other tissues. The accumulated fat is partly burned during the time of the year when food is scarce, and partly converted into other tissues, notably gonads, whose contents are shed in spring. As a result, body mass is reduced and the relative gill area increased when temperatures again start to rise. These cycles, which use fat as protection against respiratory stress, are the reason why temperate fish tend to contain more muscle fat and visceral fat than tropical species, where temperatures, although high, do not fluctuate much in the course of a year.

Another tactic that delays respiratory stress is associated with ontogenetic shifts in diet composition. Here, the young fish feed on a diffused, small prey (e.g., invertebrate zooplankton), whereas the adults, due to their sheer size, can capture energy-rich prey such as other fish, which are acquired at a lesser cost by the predator.

Relationships Between Growth, Mortality, and Maturity

The constrained growth of fishes with age is well described by the von Bertalanffy growth function. This function has an inflexion point where the increase in body weight per unit time is maximum. This is also the area where the expected body weight, that is, the product of body weight and the probability of reaching that weight, has a maximum. A species which, for a single life-time spawning event, transforms a certain fraction of its body weight into gonads, maximizes its expected reproductive output and thus its fitness if it matures, spawns, and dies at the size and age of maximum growth rate. A species with multiple spawning events, all of which are likely to have the same average success rate, maximizes its fitness by maturing in such a way that the maximum growth rate falls roughly into the middle of the mean duration of the reproductive phase. However, a species where successful reproduction depends on rare and unpredictable environmental conditions will maximize its fitness by a trade-off between reproductive output and number of spawning events, that is, by maturing early. These different strategies are shown schematically in **Figure 4**. To maximize fitness, one-time spawners (A) will aim to mature at the peak of the blue curve, which represents the expected reproductive output. Bearers, nesters, and guarders (B), whose parental care ensures a certain success at every spawning event, increase their fitness by maturing such that their average reproductive phase (dotted line) maximizes the area under the blue curve. Highly fecund nonguarders (C), for which most spawning events will be unsuccessful, maximize their fitness by maturing early and thus increasing the number of spawning events. This maturation framework is confirmed by the data on length at first maturity shown in **Figure 5**. One-time spawners, bearers, and guarders mature close to and slightly before 0.67 asymptotic length, which is the size range where the growth rate in weight and the expected reproductive output have their maximum. In contrast, nonguarders mature at a significantly smaller size.

These considerations are highly relevant for the biodiversity of fish stocks, because exploitation typically commences before maturation and high fishing mortality reduces the average duration of the reproductive phase to only one spawning event. From **Figure 4** it should be clear that this dramatically reduces the expected reproductive output of individuals and stocks, exerting unnatural selection for early maturation and provoking recruitment failures. Decrease in size and age at maturation and increase in recruitment failures are regularly observed in heavily exploited stocks.

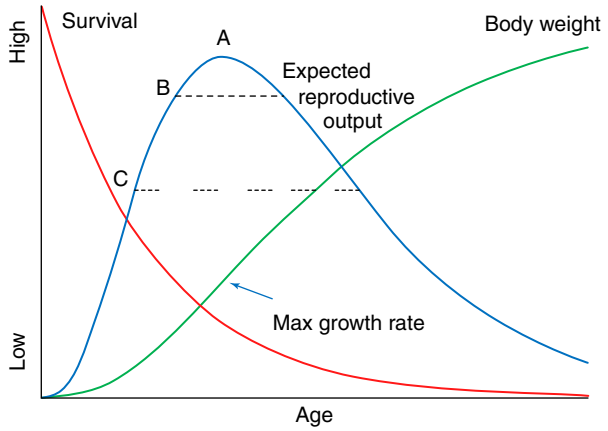


Figure 4 Schematic representation of reproductive strategies in relation to probability of survival (red curve) and increase in body weight (green curve). The blue curve indicates the expected reproductive output. Strategy A represents single-spawners, such as salmon or eels. Strategy B represents multiple-spawners with parental investment, such as live-bearers, nesters, or guarders. Strategy C represents highly fecund nonguarders such as cods, sardines, or tunas. The dotted horizontal lines indicate the necessary duration of the reproductive phase.

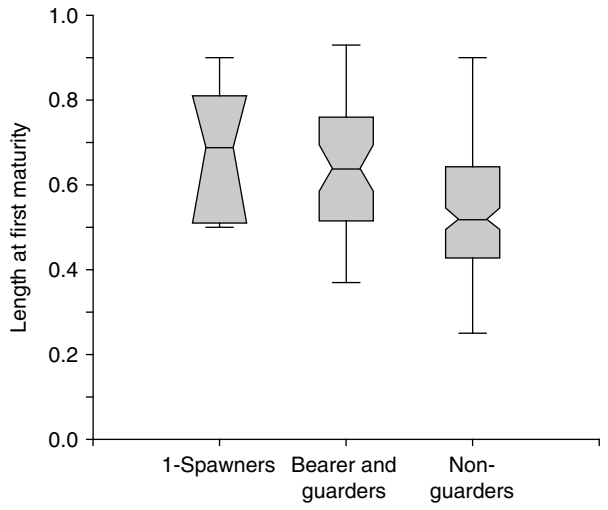


Figure 5 Length at first maturity relative to asymptotic length for species that spawn only once in their life time (one spawners, five studies), bearers or guarders (49 studies), and nonguarders (178 studies). The horizontal lines within the boxes represent the median and the notched area the 95% confidence limits. The boxes contain 50% of the data and the extended lines indicate the spread of the data. The difference between bearers and guarders and nonguarders is significant.

Distribution of Exploited Fish Stocks

Overall Distribution Ranges

Although mostly confined to water, fish occur in a wider range of habitats than any other vertebrate or invertebrate group. Thus, fish range from the upper reaches of streams in high mountain ranges (e.g., river loaches, Balitoridae) to the mouths of temperate and tropical rivers (e.g., gray mullets, Mugilidae). In the marine realm, fish range from the intertidal to the ocean's abyss, either as predators in their desert-like expanses (e.g., skipjack tuna, *Katsuwonus pelamis*) or as components of the rich, newly discovered deep-sea vent ecosystems (e.g., some live-bearing brotulas, Bythitidae). Environmental adaptations include the ability to deal with an enormous range of pressures (from about one to hundreds of atmospheres), temperatures (from $-1.8\,^{\circ}C$ in polar waters to about $40\,^{\circ}C$ in hot springs, tolerated by some tilapias), and salinities (from close to distilled water preferred by the discus fish, *Symphysodon discus*, of Amazonia to about 10%, e.g., in West African hypersaline coastal lagoons inhabited by the blackchin tilapia, *Sarotherodon melanotheron*), to list only three environmental factors. No single fish species or family, however, spans more than small fractions of these ranges. Rather, these various adaptations are exhibited by a bewildering variety of forms, ranging from minute gobies that are fully grown at close to 1 cm (e.g., *Mystichthys luzonensis*) to the 15 m reached by whale sharks (*R. typus*). These two species, incidentally, are exploited for food in the Philippines. The former, despite its turnover rate, is in danger of extinction in the small lake where it is endemic because of overfishing and pollution. The latter is now legally protected, but enforcement remains problematic.

Adaptations to Open-Ocean Habitats

Fish have different strategies to deal with the low production of the oceans. Tuna have adopted a high-energy strategy, where their tightly packed schools quickly move from one food patch to the other, essentially hopping from one oasis to the next and minimizing the time spent in the intervening desert-like expanses. Others, notably the lantern fishes (Myctophidae), occur in scattered populations that, at dawn, migrate from 1000 m to the surface waters, and get back at dusk. These different strategies imply very different biomasses: tens of millions of metric tons for the major tuna species (prior to their recent depletion by various longline, purse seine, and other fisheries) against an estimated global biomass of one billion metric tons for the lantern fish and associated communities. The latter number is often viewed as a promising figure, from which various estimates of potential yields have been derived. Most of these estimates, however, do not consider the extremely dilute nature of this biomass (usually less than 1 g per metric ton of water).

Shelf Communities

Definition of Neritic Stocks

Most fish stocks are neritic, that is, occur above the continental shelves, the productive areas of shallow waters (down to 200 m) around the continents, from which about 90% of the world marine fisheries catches are extracted. Shelves may have rocky or soft (sandy or muddy) substrates, and usually support two weakly connected fish communities, one species-rich and consisting of bottom or demersal fishes, the other consisting of fewer species of open-water or pelagic fishes. The fish of demersal communities are those exhibiting the specialized fins

and mouths mentioned earlier, enabling utilization of distinctive food sources, particularly on reefs in both temperate and tropical regions.

On coral reefs, this fine partitioning of resources culminates in hundreds of fish species sharing a single reef, with dozens of specialists for each of its food resource types, from the filamentous algae consumed, for example, by damselfishes (Pomacentridae), the encrusting algae consumed by parrot fishes (Scaridae), the coral themselves, consumed by butterfly fishes (Chaetodontidae), to the small invertebrates consumed by, for example, wrasses (Labridae). A vast array of predators such as groupers (Serranidae) and sharks (Carcharhinidae) regulate the number of these smaller fishes. Hard-bottom shelves and, in tropical areas, the coral reefs that occur down to 30 m are also exploited wherever they occur. The fishing gear used over hard bottoms are mainly traps and handlines (the latter both sport and commercial), which are rather selective gears that would have relatively minor impacts were it not for their excessive numbers.

Demersal Fish Stocks

The demersal fish living in, on, or just above shelf soft bottoms consist of specialized flatfishes and rays and numerous generalized teleosts feeding on bottom invertebrates (the zoobenthos) and smaller fishes. The complex communities thus formed can reach very high biomass, at shallow depth in the tropics (20–50 m) and deeper in colder waters. In the warm waters of the tropics, bacteria induce a quick remineralization of the dead organic matter (detritus) falling out of the lighted part of the water column. This allows very little detritus to become available for consumption by the zoobenthos. In cold water, however, the short but intensive burst of algal production occurring in the spring is consumed only partly by the zooplankton of the upper water layers. Most of the remainder is consumed as detritus after falling down to the sea bottom as marine snow. Thus, cold-water soft-bottom communities can occur in very deep waters, down to the shelf slopes (200–300 m) and well beyond. Indeed, the latest trend in fisheries development is the exploitation of deep-sea stocks of cod-like fish (order Gadiformes), orange roughy (*Hoplostethus atlanticus*), and other fish, down to depths of 1000 m or more, through ventures that even in principle could never be managed so as to achieve sustainability.

Wherever they occur, soft-bottom shelves are nowadays invariably subjected to bottom trawling, a very unselective fishing method that is environmentally damaging. This involves dragging a heavy, chain-studded net over the sea bottom and catching, that is, removing all that it encounters. Not surprisingly, this procedure has often been compared to harvesting crops with a bulldozer. Trawler catches thus consist of targeted species (usually shrimps in the tropics and subtropics) plus a vast number of nontarget species, often the juveniles of demersals with large adult sizes, and literally parts of the habitat of bottom-fishes, notably sessile invertebrates and chunks of reefs lifted from the sea bottom. Nontarget species and debris are then discarded, and it is therefore trawlers that contribute most to the global discarding problem. Presently, about 10 million metric tons of various fish species are reported to be discarded, down from about 30 million metric tons two decades ago. This is still a very high discard rate when compared to the 80 million metric tons that appear in global landing statistics.

The contribution of trawlers to habitat destruction, including conversion of richly structured bottom habitats into featureless expanses of mud, is well recognized, and can only be compared in terms of scale with global deforestation and the ensuing trend toward desertification.

Pelagic Fish Stocks

The pelagic communities over most shelf areas previously consisted of both major and minor stocks and stocklets of herrings, sardines (Clupeidae), anchovies (Engraulidae), and their relatives, and of their predators, notably mackerels and tunas (Scombridae) and various jacks (Carangidae). In many parts of the world, pelagic fisheries have eliminated the minor stocks and stocklets, and now depend wholly on annual recruitment to the remaining major stocks. The overfishing of old, highly fecund adults in these remaining stocks explains much of their volatility. Indeed, the present emphasis of much fisheries research on variability is thus devoted largely to a secondary phenomenon created by the activity itself. It is true, however, that pelagic stocks, feeding lower in the food web, often closely track environmental changes, such as the decline of the Peruvian anchovy *E. ringens* during El Niño events, and their subsequent rebuilding, mainly from recruits produced off northern Chile.

Pelagic fish tend to form tightly structured, dense schools, which protects them from predators and facilitates detection and herding of scattered food patches. The fisheries rely on this behavior when deploying purse seines, which can surround and catch such schools in one go, often with associated predators such as dolphins. Large pelagics such as billfish (Xiphiidae and Istiophoridae) are caught by arrays of longlines, set by the thousands along shelf edges, which also capture, besides the target species, large amounts of by-catch (notably sharks). These sharks were previously left on the spot, but are now finned before the carcasses are discarded. Longlines are indeed as unselective as the now banned giant driftnets that, in the 1980s, erected "walls of death" that were hundreds of kilometers long across the migratory routes of fish in the North Pacific and the Atlantic.

Overall Status of Neritic Stocks

When combined, the demersal and pelagic fisheries of shelves and adjacent waters represent major threats to fish biodiversity. Particularly endangered are groupers and other slow-growing bottomfish, and pelagics such as bluefin tuna and various species of sharks and billfish.

Besides the fisheries, one factor contributing to this endangerment is the traditional separation of research devoted to fisheries management (stock assessments) from that devoted to conservation and to ecosystem research. Both lines of research are separated institutionally, in terms of their methods and publication outlets, and in terms of what they perceive as their mandates. Overcoming this separation is crucial if fish biodiversity is to be maintained in the face of the onslaught by fisheries. Key needs are the development of tools and concepts for integrating information on fish biodiversity and ecosystem function with the knowledge gained through a century of applied, single-species fisheries research. Before

considering these, however, evidence for fisheries impacts on ecosystems will be presented.

Ecosystem Impacts of Fisheries

Historical Trends

The earliest fishing gear so far identified by archeologists are bone harpoons that were recovered, along with other evidence of systematic fishing, from a site 90,000 years old, in the present-day Democratic Republic of Congo (formerly Zaire). Tellingly, the main species that was targeted appears to have been a now extinct, very large freshwater catfish.

This pattern of fisheries exterminating the stocks on which they originally relied, then moving on to other species, is now understood to be common. This contradicts earlier perceptions of the ocean's quasi-inexhaustible resources, as expressed among others by Victorian grandees such as the geologist Charles Lyell and the zoologist Thomas Huxley. They were misled by the then prevailing abundance of various stocks of coastal fish (notably herring, *Clupea harengus*), and by what may be called "Lamarck's Fallacy:" the notion that "animals living in the waters, especially in sea-water … are protected from the destruction of their species by Man. Their multiplication is so rapid and their means of evading pursuit or traps are so great that there is no likelihood of his being able to destroy the entire species in any of these animals."

The industrialization of the fisheries, first in Northern Europe and then in North America at the end of the nineteenth century, quickly showed these predictions to be wrong. Most coastal stocks of herring and other small pelagics were extirpated, and faded even from memory, there by coastal stocks of demersal fishes soon followed, after the introduction of bottom traveling.

The practical response to this was the introduction of bigger boats with bigger engines, fishing farther offshore. Another response was the creation of research bodies (such as the International Council for the Exploration of the Sea, founded in 1902) to assess the reason why the resources were declining. Further, several countries (notably Norway and the US) initiated costly programs where juvenile cod and other fish were raised in hatcheries and then thrown into the sea, in the vain hope that they would replenish the stocks rather than be eaten by happy predators (which they were, unhappily).

Emergence of the Sustainability Concept

The First World War put an end to most stocking programs. It also established that a strong reduction in fishing effort, as caused by the drafting of fishers and vessels into the war effort, and the spiking of major fishing grounds by underwater mines (thus creating the first marine protected areas), would lead to a recovery of depleted fish stocks. Yet the Second World War, and another demonstration of stocks rebuilding themselves when subjected to less fishing, was required for the notion of sustainable fishing to establish itself. This notion implies that some appropriate level of fishing effort (number of vessels or gear, mesh size) exists such that catches (or yield) can be maintained at high levels – hence the concept of maximum sustainable yield (MSY). This led to the emergence of fish population dynamics and stock assessments, wherein mathematical models of single-species fish stocks and of their response to targeted fishing became the mainstay of fisheries research. Beverton, Holt, and Gulland in England, Ricker in Canada, and Schaefer in the US proposed most of these still-used models during an extremely creative period lasting from the early 1950s to the mid-1970s.

Yet in spite of these advances, the fisheries never became sustainable. One obvious reason was that, given a resource to which access was essentially open, the fisheries could never limit their collective effort at the level supposed to generate MSY. Rather, effort levels increased well beyond that, permitting some fleet owners to increase their stakes even as the aggregate rent from the fisheries declined. Subsidization of expanding offshore and distant water fleets has aggravated these economic issues, enabling commercial profits to be gained even from strongly overexploited stocks. These developments are so widespread that they have rendered obvious the impacts which fisheries have on ecosystems.

Biodiversity Trends in Global Catch Data

Since 1950, the Food and Agriculture Organization of the United Nations (FAO) has collected seafood catch data reported to them by the governments of the World. These data have been criticized as being incomplete, biased toward industrial fisheries, overaggregated both spatially and taxonomically, and unreliable as exemplified by some countries' continued reporting of high catches even after a typhoon had destroyed their fishing fleet. Yet, this is the only available global data set on fisheries, and it forms the basis for global fisheries policy. Attempts to predict global trends from a few hundred stocks for which complete assessments are available are flawed because of their spatial and survivorship biases, that is, the fact that such assessments are only available from developed countries for stocks that have withstood exploitation for decades and thus are more resilient to fisheries than the many stocks that have been quickly depleted, not meriting a full assessment.

The nominal catches of seafood organisms have reached about 80 million tons in 1988 and fluctuated around that level since then (**Figure 6**). This has been interpreted as a phase of stability, as if global fisheries had settled at a level that could be sustained indefinitely. A closer examination of the dynamics in the composition of these catches leads, however, to a very different interpretation, revealing severe impacts on the global fish stock biodiversity.

The number of exploited stocks has increased continuously and has more than tripled since 1950, meaning that from the beginning the increase in global catches did stem not only from increased exploitation of existing stocks, but also from exploitation of new stocks. It also means that the impact of fishing on the biodiversity of fish stocks has tripled in half a century.

During the same period, the number of depleted stocks has increased from close to zero to about one-third of all stocks in 2009. For the phase of perceived stability in total catches after 1988, this means that there is an underlying process where the

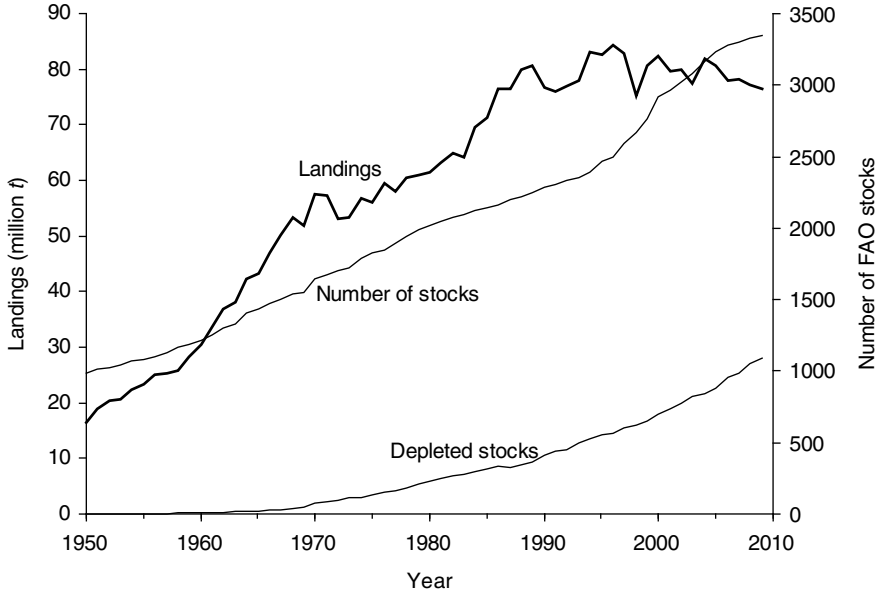


Figure 6 Underlying trends in global catch statistics. The bold curve shows the reported landings of seafood organisms. The number of exploited stocks, defined here as species by statistical marine area, is indicated by the upper thin curve. The number of depleted stocks, defined here as those producing less than 10% of their maximum contribution in the time series, is indicated by the lower thin line.

loss in catch from stock depletion is made up by new catches from new stocks. But the number of new stocks suitable for exploitation is limited, and if current trends continue we can expect to see an accelerated increase in depleted stocks and an accelerating decrease in global catches. The only remaining option for halting this trend is the rebuilding of depleted stocks, a process that has started in New Zealand, Australia, and the USA, but which has yet to reach a level where it makes a difference in global statistics.

Managing Fish Biodiversity Information

Biodiversity As a Conceptual Challenge

There is a widespread perception that the main obstacle to the conservation of fish stocks and of fish biodiversity is lack of data, a notion strengthened by public statements of biologists worried about the lack of funding for relevant research. However, simple lack of data cannot be the problem, not after the 250 years since Linnaeus created the taxonomic standards required for biodiversity research, after 100 years of applied fisheries research, and after at least 50 years of advances in ecosystem research. Rather, the problem here is the fragmentation of the database collected so far. Indeed, many studies conducted in recent years on the status of various stocks fail to consider previous knowledge on their relative abundance and distribution, and thus contribute to shifting baselines, wherein only the most recent and usually low estimates are used as reference for conservation or rebuilding efforts.

One reason for this reluctance of biologists to consolidate existing data into comprehensive, global databases may be due in part to the perception that biological data are too difficult to standardize, or are useless once standardized. Addressing these issues will be a key task of biodiversity research, and the authors now present a few ideas related to this.

There is consensus that the objects of biodiversity research are genes, populations, species, and ecosystems. However, there is little consensus as to what distinguishes biodiversity from the existing disciplines of fisheries biology, ecology, biogeography, population genetics, or taxonomy. As a result, the array of data being claimed to be essential for biodiversity studies reads like a composite list of the data traditionally used in the older disciplines, with few attempts at integration or prioritization. Such integration and prioritization are possible, however, by giving emphasis, in biodiversity studies, to data that are: (1) relevant to current research issues (e.g., richness, rarity, distinctiveness, representativeness, threat, function, and utility of species); (2) part of the data traditionally collected in taxonomy, biogeography, population genetics, and ecology; (3) widely available, in sufficient quantity; (4) pertinent to past, present, and most likely future trends; (5) easy to collect; (6) easy to standardize; (7) easy to verify; and (8) suggestive of new lines of research.

Bioquads As Primary Biodiversity Data Sets

A minimum core of biodiversity information that fulfills these eight criteria is provided by bioquads (from quads, short for quadriads), consisting of (1) the scientific name of a taxon, usually a biological species or other evolutionarily significant unit; (2) the locality where a specimen of this taxon has been encountered; (3) the date (time) of the encounter; and (4) the authority or source reporting (1)–(3). A standard for scientific names (1) has been developed by the Catalogue of Life initiative (www.catalogueoflife.org), which provides an authoritative index of scientific names for more than 1.4 million species of the 1.8 million species that are thought to have been

formally described by taxonomists during the past 250 years. Standards for (2)–(4) have been set by the Global Biodiversity Information Facility (GBIF) and applied so far to more than 250 million bioquads, which are freely accessible from the GBIF portal at www.gbif.org. The number of bioquads is expected to increase rapidly as observations of lay persons are integrated into the system. The challenge now is to interpret these large amounts of data and to derive insights on marine biodiversity and the diversity of fish stocks. This task has been taken on by the AquaMaps initiative (www.aquamaps.org), which has published the first comprehensive global map of marine biodiversity (**Figure 7**). Although the map is based on only 11,500 of the estimated 250,000 species living in the oceans, it already shows the expected trends in global species richness, such as exponential decline in species numbers from the equator to the poles, higher diversity on the continental shelves, and the center of marine biodiversity in the Malaysian-Indonesian-Philippine triangle.

AquaMaps can also be used to depict changes in catches on fish stocks, which in many cases are driven by changes in abundance. For example, **Figure 8** shows catches of Atlantic cod (*G. morhua*) in 1968 and in 2007. Available stock assessment data confirm that the visible strong decline in catches is a result of the strong decline in biomass, which itself is a result of previous overfishing.

Species Databases As Tools for Management of Biodiversity Information

Knowing the correct scientific name and the native range is a minimum requirement for a species to be included in one of the two biological databases available for marine organisms, FishBase (www.fishbase.org) for fishes and SealifeBase (www.sealifebase.org) for all other organisms. Both databases extract and standardize key information from the scientific literature, such as diet composition, growth, reproduction, morphology, and physiology. They also record human use and the resilience of species. FishBase has been utilized extensively for understanding and management of fish biodiversity, with more than 1000 citations in the primary literature, and about half a million visits per month to the FishBase portal. Recent changes in legislation, for example, in the USA, require fisheries managers to provide reference points and assessments for all fished stocks, including many cases where no stock-specific data are available. To fill these gaps, FishBase is exploring Bayesian methods to derive priors from related stocks and species. It also provides empirical equations for preliminary estimates on, for example, resilience or size at first maturity. Following a general trend to preserve scientific data, FishBase is considering to store primary life history data, such as weight-at-age or fecundity-at-length, in addition to the published models fitted to such data. This will enable the reuse of such data with other models and for different questions.

Preserving Fish Biodiversity

Traditional Approaches to Stock Management

None of the foregoing considerations will however, help, if fisheries are allowed to continue undermining their resource base, which they will if fisheries management continues to rely on the panoply of approaches so far deployed. These traditional approaches include, among other things: (1) mesh size restriction; (2) restriction on the amount or species of fish that may be legally landed; (3) effort limitation, for example, through caps on the vessel tonnage that may be deployed; and (4) seasonal closures.

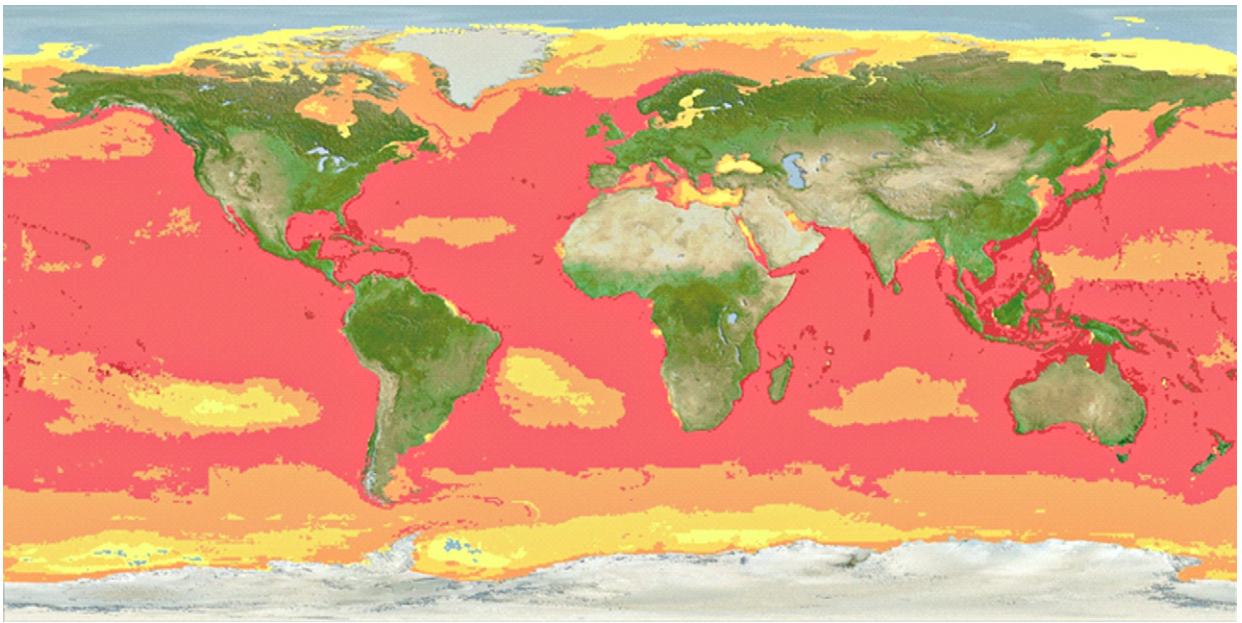


Figure 7 Marine species richness based on individual range maps for 11,500 species of fishes, marine mammals, and invertebrates. Species richness is depicted on a log scale from low (= yellow) to high (= dark red).

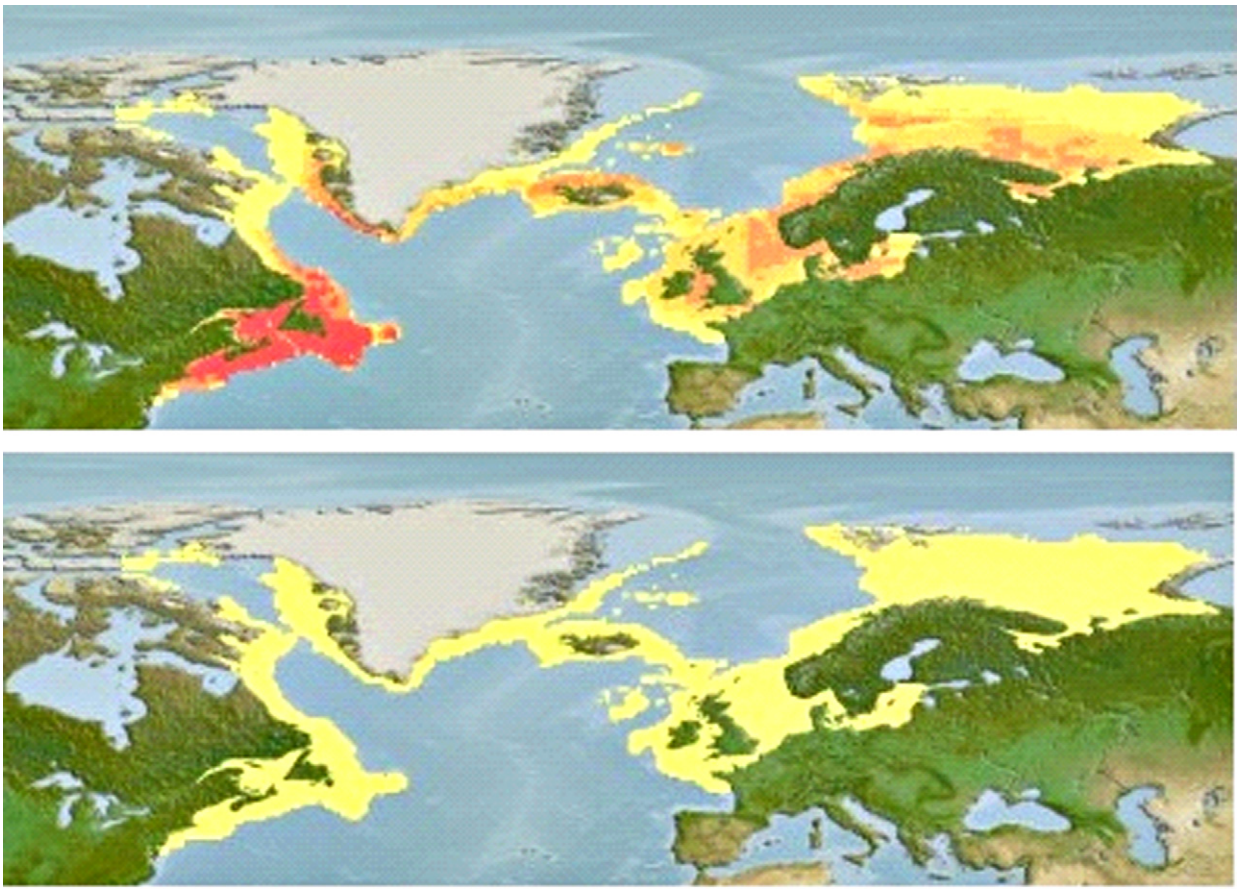


Figure 8 Catches of Atlantic cod distributed according to suitability of habitat. The upper map refers to 1968 whereas the lower map refers to 2007. Light yellow indicates catches of 1–400 t whereas dark-red areas yielded 1600–2100 t per year and half-degree cell.

Besides being extremely hard to enforce, these approaches – which are invariably conceived in the context of single-species assessments – fail to address the ecosystem effects mentioned earlier (*see* Ecosystem Impacts of Fisheries). Thus, mesh sizes above a certain limit, meant to protect the young of a given species, do not prevent associated species form being caught. Indeed, when combined with restrictions on total allowable catch (TAC), and on the landing of bycatch (as is often the case), mesh size restrictions become the very reason for discarding the young of both the targeted species and the nontarget species. Limits on nominal fishing effort are subverted by technological developments, such as improved gears and navigation instruments (e.g., GPS), which increase the catching power of fishing vessels. Thus, government-run vessel retirement schemes often end up subsidizing the modernization of fishing fleets. Finally, seasonal closure of various areas usually has negligible ecological impacts, because the fishing effort expended during the open season is sufficient for the sea bottom to be scraped up numerous times by trawlers, and for the stocks of long-lived fishes to be severely impacted.

Marine Protected Areas

There is an emerging consensus among fisheries scientists and conservationists that an important fisheries management tool that will allow the recovery of damaged stock and ecosystems is the establishment of Marine Protected Areas (MPAs), including permanent No-Take zones as their core. Such core zones are easy to enforce – at least relative to the task of enforcing mesh sizes or TACs. Further, technology-driven increases in fishing effort can be ignored, and there is assurance that the long-lived organisms of seafloors and their associated fish communities can gradually return to a semblance of their original configurations. However, considerable research will have to be devoted to identifying the optimal size and location of MPAs, particularly for migratory stocks.

Traditional fisheries management, aimed at limiting effective fishing effort, will still have to continue around MPAs, lest they become marine larders or fish-attracting rather than fish-producing zones from where resources are drained by fisheries operating at their very periphery.

Finally, the social context of fisheries will have to change: fisheries do not harvest crops they have sown. Rather, they exploit the natural productivity of wildlife; thus there are inherent limits to global fish catches, and future fisheries will not meet the demand of an ever-increasing human population if these limits are ignored. Indeed, the massive ecosystem changes already described indicate that these limits have been reached in most parts of the world, and that sustainable fisheries must be embedded in some form of ecosystem management.

See also: Adaptation. Fish Conservation. Marine Ecosystems

References

Froese R and Pauly D (eds.) (2000) *FishBase 2000: Concepts, Design and Data Sources*. Manila: ICLARM. www.fishbase.org

Hawksworth DL, Kirk, PM, and Clarke SD (eds.) (1997) *Biodiversity Information: Needs and Options*. Wallingford, UK: CAB International.

Miller RI (ed.) (1994) *Mapping the Diversity of Nature*. London: Chapman & Hall.

Mooney P (ed.) (1998) Ecosystem Management for sustainable fisheries. *Ecological Applications*. Supplement 8(1).

National Research Council (1999) *Sustaining Marine Fisheries*. Washington, DC: National Academy Press.

Nelson J (2006) *Fishes of the World*, 4th edn. New York: John Wiley & Sons.

Pauly D (1994) *On the Sex of Fishes and the Gender of Scientists: Essays in Fisheries Science*. London: Chapman & Hall.

Paxton JR and Eschmeyer WN (eds.) (1998) *Encyclopedia of Fishes*. San Diego: Academic Press.

Reaka-Kudla, ML, Wilson DE, and Wilson EO (eds.) (1997) *Biodiversity* II. *Understanding and Protecting our Biological Resources*. Washington, DC: Joseph Henry Press.

Flies, Gnats, and Mosquitoes

Brian V Brown, Natural History Museum of Los Angeles County, Los Angeles, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 815–826, © 2001, Elsevier Inc

Glossary

Diptera Group of insects to which the flies, including gnats and mosquitoes, belong.

Larva (pl. larvae) Immature stage of flies, often called maggots.

Introduction

The insect Order Diptera, with more than 124,000 currently described, extant species, ranks as one of the worlds largest groups of organisms. Along with the other insect megadiversity groups—Coleoptera (beetles; 350,000 species), Lepidoptera (butterflies and moths; 120,000 species), and Hymenoptera (ants, bees, wasps, and sawflies; 130,000 species)—they form the largest aggregation of species on the planet. Each of these megadiversity groups has more species assigned to it than to any other group of organisms except plants (300,000 species). Like other insects, most species of Diptera are still undescribed, and the actual number could range as high as 1 million or more species.

Diptera are found on every continent, including Antarctica. The relative percentage of Diptera within the insect fauna rises with latitude (and elevation) as other, less cold-adapted taxa are lost; for instance, over one-half of all insect species recorded from the Canadian high arctic are Diptera. The extant species of all biogeographical regions have been cataloged (Table 1), as have the fossil Diptera. Some catalogs are badly outdated and the number of species described for each region is often more indicative of the amount of taxonomic activity directed at a given fauna than the true diversity. For instance, the Palearctic Region is the best studied and has more described species than the relatively poorly known, but presumably more diverse, Neotropical Region. Therefore, conclusions about dipteran biogeography based on these numbers will be highly inaccurate.

Diptera are a well-established monophyletic group, with the most obvious defining character being the reduction of the hind wings to small, club-shaped organs called halteres. The insect orders considered to be most closely related to Diptera are Siphonaptera (fleas) and Mecoptera (scorpionflies), although recent molecular research indicates that the highly aberrant Strepsiptera (twisted-wing parasites) may be their closest relatives.

Fossil Diptera are common in the amber faunas of the world, which however extend back in time only to the early Cretaceous. The oldest dipteran fossil is believed to be from at least the mid-Triassic.

Like the other megadiversity groups, the Diptera are holometabolous insects, with a separate egg, larva, pupa, and adult stage. Generally, most feeding takes place in the larval stage, whereas the adult is usually specialized for reproduction and dispersal. Some adult Diptera are voracious feeders, however, requiring substantial supplemental feeding to mature their eggs or power their flight.

Flies are common in natural, disturbed, and urban habitats. Larvae are found on land and in freshwater; there are relatively few marine or brackish water species. Some species are synanthropic and have been transported around the world with human activities. Diptera are the most important vectors of human and animal disease, and a few plant-feeding species have become agricultural pests. The life history of most species of flies, however, is unknown. The group is so large, and there are so many undescribed species, that the science of dipterology is still in its relative infancy.

Table 1 Biogeographical regions and the number of described diptera species

Region	*Number of species*
Nearctic	25,000
Neotropical	20,000
Palearctic	29,000
Afrotropical	16,318
Australasian	15,764
Oriental	15,964
Antarctica	60

Compiled from the various catalogs.

Major Subdivisions

Table 2 presents a general list of the families of Diptera and some of the higher taxa that contain them. This list is based on only one of several possible classifications, however, and cannot be considered the final word in dipteran groupings. Common names, where well established, are given in Table 2. Many families lack common names, including the second largest family in the order (Tachinidae). That a group with more described species than the mammals lacks a common name is a good indication of the lack of general appreciation for the importance and ubiquity of the Diptera. The number of described species in each family is given, but in many instances these numbers are badly out of date and should be considered a bare minimum. Certainly, in most families of Diptera there are a large number of undescribed species awaiting the attention of specialists. It is estimated that we

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00167-2

Table 2 Simplified list of major subgroups and families of Diptera

	Common name	*Number of species*
Lower Diptera (= Nematocerous groups)		
Infraorder Ptychopteromorpha		
Ptychopteridae	Phantom crane files	61
Tanyderidae		42
Infraorder Culicomorpha		
Superfamily Culicoidea		
Culicidae	Mosquitoes	3,000
Dixidae	Meniscus midges	175
Corethrellidae		61
Chaoboridae	Phantom midges	50
Superfamily Chironomoidea		
Ceratopogonidae	Biting midges	5,360
Chironomidae	Midges	5,000
Simuliidae	Black files	1,475
Thaumaleidae		121
Infraorder Blephariceromorpha		
Blephariceridae	Net-winged midges	300
Deuterophlebiidae	Mountain midges	14
Nymphomyiidae		7
Infraorder Bibionomorpha		
Axymyiidae		5
Bibionidae	March flies	675
Cecidomyiidae	Gall midges	4,600
Mycetophilidae	Fungus gnats	3,000
Pachyneuridae		4
Sciaridae	Dark-winged fungus gnats	1,000
Infraorder Tipulomorpha		
Tipulidae	Crane flies	14,000
Trichoceridae	Winter crane flies	110
Infraorder Psychodomorpha		
Anisopodidae	Wood gnats	100
Perissomatidae		5
Psychodidae	Sand flies, moth flies	2700
Scatopsidae		255
Synneuridae		4
Brachycera (an unranked taxon)		
Infraorder Xylophagomorpha		
Xylophagidae		111
Infraorder Tabanomorpha		
Athericidae		90
Pelecorhynchidae		49
Rhagionidae	Snipe flies	520
Tabanidae	Horseflies, deer flies	3,000
Vermileonidae	Ant lions	31
Infraorder Stratiomyomorpha		
Pantophthalmidae		20
Stratiomyidae	Soldier flies	2,500
Xylomyidae		107
Infraorder Muscomorpha		
Superfamily Nemestrinoidea		
Acroceridae	Small-headed flies	500
Nemestrinidae	Tangle-veined flies	250
Superfamily Asiloidea		
Apioceridae	Flower-loving flies	145
Asilidae	Robber flies	5,600
Bombyliidae	Bee flies	4,800
Mydidae	Mydas flies	353
Scenopinidae	Window flies	330
Therevidae	Stiletto flies	850

(*Continued*)

Table 2 Continued

	Common name	*Number of species*
Superfamily Empidoidea		
Dolichopodide	Long-legged flies	5,100
Empididae	Dance flies	3,500
Cyclorrhapha (an unranked taxon of Muscomorpha)		
"Lower Cyclorrhapha" (= Aschiza of previous authors)		
Ironomyiidae		1
Lonchopteridae	Spear-winged flies	35
Opetiidae		3
Phoridae	Humpbacked flies, scuttle flies	3,200
Pipunculidae	Big-headed flies	600
Platypezidae	Flat-footed flies	215
Sciadoceridae		2
Syrphidae	Hover flies, flower flies	5,800
Schizophora		
"Acalypterates" (possibly not monophyletic)		
Superfamily Nerioidea		
Cypselosomatidae		30
Micropezidae	Stilt-legged flies	500
Neriidae	Cactus flies	110
Superfamily Diopsoidea		
Diopsidae	Stalk-eyed flies	153
Megamerinidae		13
Nothybidae		8
Psilidae	Rust flies	200
Somatiidae		7
Strongylophthalmyiidae		27
Syringogastridae		8
Tanypezidae		18
Superfamily Conopoidea		
Conopidae	Thick-headed flies	800
Superfamily Tephritoidea		
Lonchaeidae		700
Otitidae		800
Pallopteridae		52
Piophilidae	Skipper flies	71
Platystomatidae		1,000
Pyrgotidae		200
Richardiidae		170
Tachiniscidae		3
Tephritidae	Fruit flies	4,000
Superfamily Lauxanioidea		
Celyphidae	Beetle flies	90
Chamaemyiidae	Aphid flies, silver flies	250
Eurychoromyiidae		1
Lauxaniidae		1,550
Superfamily Sciomyzoidea		
Coelopidae	Seaweed flies	30
Dryomyzidae		30
Helosciomyzidae		23
Sciomyzidae	Marsh flies	515
Ropalomeridae		30
Sepsidae	Black scavenger flies	240
Superfamily Opomyzoidea		
Suprafamily Clusioinea		
Acartophthalmidae		3
Clusiidae		217
Suprafamily Agromyzoinea		
Agromyzidae	Leafminer flies	3,500
Fergusoninidae		25
Odiniidae		50

(*Continued*)

Table 2 Continued

	Common name	*Number of species*
Suprafamily Opomyzoinea		
Anthomyzidae		49
Marginidae		2
Opomyzidae		40
Suprafamily Asteioinea		
Asteiidae		100
Aulacigastridae		25
Neurochaetidae	Upside-down flies	5
Periscelididae		19
Teratomyzidae		5
Xenasteiidae		8
Superfamily Carnoidea		
Australimyzidae		5
Braulidae	Bee lice	7
Canacidae	Beach flies	113
Carnidae		41
Chloropidae	Frit flies	2,000
Cryptochaetidae		25
Milichiidae		190
Tethinidae		126
Superfamily Sphaeroceroidea		
Chyromyidae		40
Heleomyzidae (s.l.)		520
Mormotomyiidae		1
Sphaeroceridae	Lesser dung flies	2,500
Superfamily Ephydroidea		
Camillidae		11
Curtonotidae		36
Diastatidae		22
Drosophilidae	Vinegar flies, pomace flies	2,900
Ephydridae	Shore flies	1,300
Calyptratae		
Superfamily Hippoboscoidea		
Glossinidae	Tsetse flies	22
Hippoboscidae	Louse flies	200
Nycteribiidae	Bat flies	260
Streblidae	Bat flies	220
Superfamily Muscoidea		
Anthomyiidae		1,100
Fanniidae		265
Muscidae		3,880
Scathophagidae		250
Superfamily Oestroidea		
Calliphoridae	Blow flies	1,000
Mystacinobiidae		1
Oestridae	Bot flies, warble flies	42
Rhinophoridae		100
Sarcophagidae	Flesh flies	2,500
Tachinidae		9,200
TOTAL		124,390

Source: Adapted from Yeates DK and Wiegmann BM (1999) Congruence and controversy: Toward a higher-level phylogeny of Diptera. *Annual Review of Entomology* 44: 397–428, and McAlpine JF (ed.) (1989) *Manual of Nearctic Diptera*, vol 3. Agriculture Canada Monograph No. 32.

Most important taxa are in bold. Approximate number of described species from various sources.

have described only 10% or fewer of the species in many families.

The Diptera are traditionally organized into three suborders: the Nematocera, the orthorrhaphous Brachycera, and the cyclorrhaphous Brachycera. Of these, it is likely that only the Cyclorrhapha is a monophyletic, or natural, group. Although a consensus on the higher relationships within the Diptera is not yet available, it is generally agreed that the "Nematocera" is a paraphyletic assemblage of relatively primitive dipteran families and that some subgroup of this assemblage is more closely related to the Brachycera than to other nematocerans. Similarly, within the Brachycera, a group

widely considered to be monophyletic, the orthorrhaphous families are now believed to be a paraphyletic assemblage relative to the Cyclorrhapha. Finally, within the Cyclorrhapha is a group of flies traditionally called the Aschiza (here called the lower Cyclorrhapha), which is probably paraphyletic with respect to the monophyletic Schizophora. The details of the phylogeny of the Diptera has been reviewed extensively by Yeates and Wiegmann and their discussion of various groups is highly recommended for understanding the current status of dipteran phylogeny.

Within the Diptera, higher taxa show repeated patterns of relatively primitive, nondiverse grades of organization with relatively highly derived, speciose sister taxa. Thus the lower Diptera are collectively much less diverse than the Brachycera, with the notable exception of the incredibly large family Tipulidae. Within the Brachycera, the Muscomorpha is by far the largest infraorder and the Schizophora has four times more species than the lower Cyclorrhapha. The numbers are less disparate in the Schizophora, with about 26,000 acalypterates and 19,000 calypterates, but the monophyly of the acalypterates is highly contentious, making such a comparison questionable.

The distribution of species among the families of Diptera is extremely divergent, with the largest number found in the Tipulidae and the Tachinidae. Together, these families account for almost 20% of the species in the order.

The Tipulidae, or crane flies, are elongate, long-legged, somewhat fragile flies that are found nearly everywhere on earth. The adults are sometimes found at lights but are most often seen resting on vegetation. The larvae have a variety of habitats, including terrestrial—in soil, mosses, and decaying wood—freshwater, and intertidal. They are scavengers, herbivores, or predators.

In contrast to the eclectic habits of the Tipulidae, larvae of Tachinidae have a single way of life: parasitism. All known species are internal parasitoids of other arthropods, mostly other insects and, within the insects, mostly larval Lepidoptera. Eggs are either laid on or in the host or are broadcast in suitable areas. Larvae hatch from broadcast eggs and wait in ambush for hosts. Some tachinids produce microtype eggs that are designed to be ingested by the hosts; these eggs hatch inside the host and penetrate the gut wall to enter the body cavity. Adult tachinids are stout, bristly, housefly-like flies that are seen frequently on vegetation or flowers.

There are another 21 families of Diptera that can be considered large, possessing about 2000 or more described species. Together, these 23 largest families (including tipulids and tachinids) comprise about 100,000 species, or approximately 80% of the Order Diptera. In order of decreasing number of species, these other 21 large families are as follows:

- Syrphidae: hover flies, flower flies (5800 spp.). Adults of this family often are brightly colored mimics of Hymenoptera (bees and wasps). The larvae are saprophagous, predatory, or herbivorous. Some saprophagous species are called rat-tailed maggots because of their elongate posterior breathing tubes. Many of the predatory species live exposed on plants, feeding on aphids.
- Asilidae: robber flies (5600 spp.). The adults are voracious predators that usually attack their prey while in flight, stabbing them with their heavily sclerotized, swordlike proboscis and injecting digestive fluids. The larvae are also predatory.
- Ceratopogonidae: biting midges (5360 spp.). The tiny adult females of this family require blood meals to mature their eggs. Some species bite vertebrates, other feed on large insects, and still others are predatory on small insects. The larvae are predacious, living in damp terrestrial or freshwater habitats.
- Dolichopodidae: long-legged flies (5100 spp.). These flies are often metallic green in color and are found commonly on undergrowth or on tree trunks. The larvae and adults are predatory.
- Chironomidae: midges (5000 spp.). These flies are among the most abundant benthic invertebrates in freshwater environments; some are also terrestrial or intertidal. The adult males often form enormous mating swarms.
- Bombyliidae (including Mythicomyiidae): bee flies (4800 spp.). Species of this family are most diverse in dry areas, including deserts, and are nearly absent from tropical rain forests. The larvae are parasitoids of the immature stages of various other insects or predatory on grasshopper egg pods.
- Cecidomyiidae: gall midges (4600 spp.). Cecidomyiids are generally considered one of the largest, yet most poorly known groups of Diptera. Larvae often form galls on various plants, but there are non-gallforming species and many mycophagous forms as well.
- Tephritidae: fruit flies (4000 spp.). Larval tephritids are phytophagous, attacking a wide variety of plants and sometimes forming galls. The adults often have color patterns on their wings.
- Muscidae: houseflies and relatives (3880 spp.). A few extremely well-known muscid species are synanthropic, especially *Musca domestica* and *Stomoxys calcitrans*. Larvae of most species live in decaying organic material, where they are either saprophagous or predatory. Some species are saprophagous in early larval instars, becoming carnivorous later.
- Agromyzidae: leaf-miner flies (3500 spp.). Along with some other insects, larvae of many species of agromyzids feed within plant leaves, excavating the distinctive, light-colored tunnels called mines. The adults are small, usually dark-colored flies; some species are marked with yellow.
- Empididae: dance flies (3500 spp.). Adults of this family are found mostly in damp terrestrial habitats, often near water. Common and diverse in temperate regions, they are less prevalent in tropical lowland forests. Larvae and adults of both sexes are predatory; larvae of a few species are parasitoids of caddisfly larvae.
- Phoridae: humpbacked flies, scuttle flies (3200 spp.). Phorids are common and diverse nearly everywhere except Antarctica. Larvae can be predators, parasitoids, true parasites, herbivores, or scavengers. Many species are associated with social insects.
- Tabanidae: horseflies, deer flies (3000 spp.). Adult female tabanids are well-known blood feeders, although some species feed only on nectar. The larvae are predatory, usually found near water.
- Culicidae: mosquitoes (3000 spp.). Adult females bite vertebrates to obtain blood meals, and they often transmit diseases. Larvae are aquatic.

- Mycetophilidae: fungus gnats (3000 spp.). Adults are found mostly in humid, forested areas. The larvae feed mainly on fungi, although some species spin webs to capture insect prey and a few species are parasitoids of flatworms.
- Drosophilidae: vinegar flies, pomace flies (2900 spp.). This family is best known for *Drosophila melanogaster*, the ubiquitous model organism for genetic research. Although adults are commonly found around overripe fruit or on mushrooms, the larvae of this family have a variety of lifestyles, from saprophagy to parasitism and predation.
- Psychodidae: moth flies (2700 spp.). The scavenging larvae of these flies are usually found in moist conditions, in soil, rotting wood, or other decaying vegetation. Adults of one subfamily are blood feeders and transmit the disease leishmaniasis to humans.
- Sarcophagidae: flesh flies (2500 spp.). Sarcophagid larvae are saprophagous, parasitoids, predators or commensals in the nests of solitary Hymenoptera. The adults of some species are associated with filth.
- Sphaeroceridae: lesser dung flies (2500 spp.). Larvae and adults of this family are commonly found on dung, carrion, and decaying organic material.
- Stratiomyidae: soldier flies (2500 spp.). The adults of this family are often brightly colored and conspicuous on flowers. The larvae are often found near water, in decaying organic material, or under bark. They are saprophagous, herbivorous, or predatory.
- Chloropidae: frit flies (2000 spp.). Larvae of chloropids have nearly every conceivable way of life, from scavenging to predation, parasitism and herbivory, although there are many more plant feeders in this family than the similarly diverse Phoridae.

Nematocerous Families

The nematocerous families are the relatively primitive members of the Diptera, characterized by long, unconsolidated antennae consisting of many segments. There are 26 families in this group and approximately 40,000 species included. Many are associated with aquatic habitats and some, such as mosquitoes, black flies, and biting midges, are voracious blood feeders.

Brachycera

Most Diptera belong to the Brachycera, a group characterized by the reduction or fusion of antennal segments to eight or fewer and by modifications to the larval head and mouthparts. With about 80,000 described species, this group contains many of the best known flies, such as houseflies and fruit flies.

The lower Brachycera includes several lineages constituting approximately 30,000 species. The larvae of most species are predatory, although there are a few parasitoid groups as well. The most familiar are the large families Asilidae, Bombyliidae, Tabanidae, Dolichopodidae, and Empididae.

The Cyclorrhapha, with about 50,000 species, includes a few primitive lineages and the Schizophora. The primitive groups are relatively small, with the exceptions of the large families Phoridae and Syrphidae. Within the Schizophora, there are a plethora of smaller acalypterate families that are rare, but many, such as the Tephritidae, Agromyzidae, Drosophilidae, Sphaeroceridae, and Ephydridae, are abundant and commonly encountered. The Calyptratae includes the familiar houseflies, flesh flies, and blow flies as well as the speciose parasitoids of the Tachinidae. Also included are a number of mammal parasites, including some truly bizarre, spider-like bat parasites in the Nycteribiidae and Streblidae.

Life History Diversity

Larvae

The larvae of flies are abundant and widespread in most terrestrial and aquatic habitats. They are often encountered in soil (including sand), dead wood, dung, carrion, decaying vegetation, and among the refuse of social insect colonies. Relatively few are found living exposed on vegetation, including some herbivorous tipulids and many aphidophagous syrphids. Aquatic forms are found on the bottom and in the water column of lakes, streams, and ponds. Larval Diptera are predators, scavengers, herbivores, parasitoids, or even true parasites.

Scavengers

This is the lifestyle most commonly associated with Diptera larvae, especially those of the Brachycera. More than half of the 128 families recognized herein have larvae that feed on decaying organic material or organic detritus. Most receive their nutrition from bacteria and other microorganisms of decay, not from the main substance on which they are found. They concentrate these organisms and other suspended particles with a sievelike pharyngeal filter.

Among the most obvious terrestrial scavengers are the larvae of blow flies (Calliphoridae) that are found on newly dead animal carcasses. Early stages of decay are characterized by large numbers of calliphorid and muscid larvae, followed later by drosophilids, fanniids, phorids, piophilids, sepsids, sphaerocerids, and others as decay proceeds. The fauna of buried carrion is different, with the calliphorids largely excluded. Instead, the muscid genus *Muscina*, various phorids, and sphaerocerids predominate. The fauna of carrion immersed in water has also been studied and found to differ from that on dry land. There has been considerable study of the succession of scavenging Diptera larvae on dead animals, leading to the potential for their use in forensic entomology.

Decaying vegetable matter is also rich in scavenging Diptera, especially muscids, sphaerocerids, sciarids, and others. Some muscids are obligate thermophiles, requiring the heat generated by the decay of large piles of compost. Decaying seashore vegetation supports coelopids, sphaerocerids, and anthomyiids. Rotting fruit is the food of drosophilids, stratiomyids, and some phorids.

In aquatic environments, many larval Diptera feed on small organic particles in the water or on the substrate. Larvae of Culicidae and Simuliidae filter particles from the water with their brushlike labral fans.

Herbivores

Diptera that feed on living plants (including algae and fungi) are found in 37 of the 128 families recognized herein. Some, such as Agromyzidae, Anthomyiidae, Bibionidae, Cecidomyiidae, Chloropidae, Phoridae, Psilidae, Sciaridae, Tephritidae, and Tipulidae, include species considered to be pests to human agriculture. Fly larvae attack all parts of plants, including fruits, flowers, stems, leaves (as leaf-miners), and roots.

Predators

Predators are organisms that kill more than a single host organism for their feeding. There are 35 families with this way of life, including nearly all of the noncyclorrhaphan Brachycera. Most have extremely active larvae that attack other invertebrates as their major food source. The larvae of tabanids have been known to kill frogs, an interesting reversal of the usual chain of events. Many predatory dipteran larvae are beneficial to humans in controlling insects considered to be pests. Examples include syrphids attacking aphids and predatory muscid larvae that kill larvae of other muscids, such as houseflies.

Parasitoids

Parasitoids develop on and kill a single host. Twenty-two families of Diptera have this way of life, which was reviewed recently by Feener and Brown. All species of the second largest family of Diptera, the Tachinidae, are parasitoids. All dipteran parasitoids attack other invertebrates, usually other insects, but unusual hosts include terrestrial flatworms, mollusks, earthworms, millipedes, spiders, and scorpions.

Parasites

Parasites feed on a single host, but do not normally kill it. True parasites include the bot flies, whose larvae live under the skin or in the nasal cavities of various mammals, including humans. Some bot fly larvae live in the stomachs and alimentary tracts of horses, elephants, zebras, and rhinos. Other parasites of vertebrates are calliphorids, chloropids, piophilids, and muscids. True parasites of invertebrates are less well known but occur in a few families, such as Phoridae.

No Free-Living Larvae

The larvae of the four families of the Hippoboscoidea—Glossinidae, Hippoboscidae, Nycteribiidae and Streblidae—are retained in the female abdomen and nourished by secretions of the accessory glands. They are deposited by females as fully mature, third-instar larvae, which quickly pupariate.

Adults

Most adult Diptera receive the majority of their nutrition as larvae and do not feed extensively. Many need carbohydrates to power their flight, however, and feed on the nectar in flowers, or on honeydew, the sweet secretions of Homoptera. Some flies require nitrogen for nourishing their eggs and are thus avid flower visitors in search of pollen. Flower foraging for nectar and pollen makes Diptera adults important pollinators of plants.

At least some parasitoids feed on the hemolymph of their hosts after oviposition, and many dipteran species feed on dead insects, carrion, dung, or rotting vegetation. A few adult Diptera, including Deuterophebiidae and some Oestridae, have vestigial mouthparts and do not feed.

Adults of a number of families have species that are well-known blood feeders, especially most or all species of Ceratopogonidae, Culicidae, Glossinidae, Hippoboscidae, Nycteribiidae, Simuliidae, Streblidae, and Tabanidae. Other families with fewer blood-feeding species are Muscidae, Rhagionidae, and Athericidae.

Some families of Diptera have species that are predatory as adults. This lifestyle is especially well developed in the Asilidae, Dolichopodidae, and Empididae but also occurs in some other families, such as Muscidae and Phoridae.

Special Associations

Aquatic Diptera

The larvae of many nematocerous families, as well as some Brachycera, are found in freshwater habitats. Among the most consistently aquatic forms are those families in the Culicomorpha and Blephariceromorpha, including such well-known families as Culicidae (mosquitoes), Simuliidae (black flies), and Chironomidae (midges). Chironomidae in particular can be exceedingly abundant, and various species assemblages are often used for assessment of water quality. The larvae of Simuliidae and Blephariceromorpha are found almost exclusively in clean, running water, with the Deuterophlebiidae and Blephariceridae being especially adapted to fast-flowing streams. Larval Culicidae are often found in ephemeral ponds that appear after snowmelt in temperate regions and after heavy rains elsewhere.

Phytotelmata

Phytotelmata are structures of plants that allow accumulations of water. They occur in various parts of plants, including leaves, leaf axils (especially of bromeliads and bananas), stems (especially of bamboos), fruits, and specialized structures (such as pitchers in pitcher plants). Among the approximately 20 families of Diptera that utilize phytotelmata, the Culicidae are the most prominent, being the most regular and most numerous larvae present. The other most frequently encountered families are Chironomidae, Ceratopogonidae, and Psychodidae, but further research has found that other families, such as Phoridae and Tipulidae, might also be common. Our knowledge of the fauna of phytotelmata is still fragmentary and much further work needs to be directed at these habitats.

The insect trapping structures of pitcher plants (Nepenthaceae and Sarraceniaceae) frequently harbor dipteran larvae. For instance, a distinctive fauna of about 80 species of aquatic Diptera is specialized to live in the carnivorous pitcher plants of the genus *Nepenthes* in Southeast Asia. The species found in this habitat are mostly from aquatic groups, such as mosquitoes, but some are from families more usually found in terrestrial habitats—Calliphoridae, Chloropidae, and Phoridae, for example. Nearly all feed on the trapped, drowned

insects, or filter microorganisms associated with the decay of such insects from the water.

Social Insects

Social insects are the ants, bees, wasps, and termites that have organized societies. They have one or a few females responsible for all the egg laying, while other members of the colony (usually sterile females) gather food and do other tasks. Diptera are associated with social insects as scavengers, predators, parasitoids, and parasites. In the following treatment, only those species intimately associated with social insects are discussed, whereas opportunistic, generalized predators, such as robber flies (Asilidae), will be ignored.

Most flies known to be associated with social insects are brachycerans. Among the few nematocerous records of such associations is the astonishing observation that some Southeast Asian mosquito adults (genus *Malaya*) feed on honeydew carried in the mouthparts of ant workers.

Among the Brachycera associated with social insects, the most diverse are the Phoridae. These small flies are commonly found in association with ants, especially army ants (both New World and Old World), termites (especially in Africa and Southeast Asia), and, to a lesser extent, social bees and wasps. Many are parasitoids, laying their eggs inside the bodies of their adult hosts, usually ants, termites, and stingless bees; some are also known to parasitize the immature stages of ants. Many are scavengers, living in the refuse piles of the large colonies of army ants and leaf-cutter ants, where they are joined by scavenging larvae of other families. Some phorid larvae are predatory on ant brood, as are the larvae of some syrphids. The females of phorid species living in social insect nests are often remarkably modified, with reduced wings, eyes, and body sclerotization. One remarkable genus from Southeast Asia has adult females that mimic the larvae of their army ant hosts. Other females are heavily armored and have a rounded, teardrop shaped body form. This limuoid body form allows the females to escape damage when accosted by aggressive host ants, bees, or termites.

Other dipteran larvae found in social insect nests belong to the families Braulidae, Calliphoridae, Fanniidae, Sarcophagidae, Sphaeroceridae, and Syrphidae. Most are scavengers, although some are predatory.

Army ant raids are spectacular tropical phenomena that provide a number of opportunities for dipteran associates. Tachinids and conopids hover or perch near the raid front, darting down to parasitize the crickets, cockroaches, and other insects flushed by the foraging ants. Closer to the ground, parasitic phorids dart at their respective hosts: either the army ants themselves or other ants that are victims of the ant raids. Sarcophagids swarm on the leaves of nearby undergrowth, feeding on the droppings of birds attracted to the insects flushed out by the ant raids. Often, the raid front is best identified by listening for the loudest buzzing from the activity of hundreds of flies.

Kleptoparasites

Many Diptera exploit other insects or invertebrates that sequester or reserve food for long periods of time. By waiting before feeding, or by feeding slowly, these hosts provide a window of opportunity for larvae or adults of flies. The relationship between the host and kleptoparasite is often developed to such an extent that the flies live permanently associated with their food provider. Phoresy (transportation of the kleptoparasite by the host) often occurs in these associations.

A commonly observed example of kleptoparasitism and phoresy in Diptera is the association of large spiders and flies. Spider webs provide a continuous supply of food, and often a source of stored insect carcasses for later feeding. This warehouse is exploited by a number of fly families, including Ceratopogonidae, Cecidomyiidae, Milichiidae, Chloropidae, Lonchaeidae, Phoridae, and Empididae. In many instances, the flies perch on the bodies of the spider, waiting for their next meal to arrive.

Another, much more restricted association occurs between dung beetles of the family Scarabaeidae and flies of the family Sphaeroceridae. The flies ride on the bodies of the dung beetles, waiting until the beetles find and bury feces for their own larvae to feed upon. The flies briefly hop off the beetles, lay their own eggs on the dung, and then rejoin their food-providing host.

Kleptoparasites also occur in the nests of social Hymenoptera, where fly larvae are fed by deceived workers or attack the provisions left for a developing bee, wasp, or ant larva. A full review of kleptoparasitism is given by Sivinski *et al.* (1999).

Swarming

Swarms of adult Diptera are a common sight in most environments. Usually these are aggregations of males, allowing females to easily find a mate. Often, swarms are seen in the canopy of forests, under overhanging branches, in sunlight "pools" in forests, near fire towers emerging above tree level, or at the summit of tall hills or mountains ("hilltopping").

Most nematocerous families engage in swarming, as do about 15 families of Brachycera. Flies that engage in this behavior often have associated structural modifications, including a well-developed anal lobe of the wing and enlarged compound eyes. The development of the anal lobe probably allows for better maneuvering and hovering within the swarm, whereas the increased size of the eye allows male flies to more precisely place themselves, and assess their place, within the swarm. As females choose dominant males within the swarm, based on their relative position, assessment of position is critical for male mating success.

Species of dance flies (subfamily Empidinae) have elaborated on this basic pattern. Males of many species catch prey (usually smaller flies and other soft-bodied insects) and carry them in the swarm. They offer the prey to the nonhunting females as "nuptial gifts" to be fed upon during mating. A few species have reversed the trend of male-dominated swarms: instead, females form the aggregation, which prey-bearing males visit to selectively mate with the most desirable individuals (such female-dominated swarms are also known for some Phoridae). In some dance flies, the prey items themselves have changed, with males presenting already fed-upon prey, inanimate objects such as plant seeds, or even inedible bodily secretions. In these species, the giving of nuptial gifts apparently has become ritualized, losing all of its functional basis.

Conservation Biology of Diptera

Because of the perception of almost all Diptera as disease-carrying filth flies, there is little public sympathy or interest in their conservation. Also, most Diptera are extremely poorly known, and the study of many families is still contingent on the recognition and description of the many undescribed species.

An exception to these statements is the Syrphidae, a group that is popular among amateur collectors, especially in Europe. There, Red Lists of endangered species exist for many regions, including Britain and parts of the mainland. When species that are at the edge of their distribution are eliminated from consideration, it appears that most endangered syrphids are either saproxylic or associated with wetlands.

Saproxylic insects of all orders are considered to be among the most endangered insects in Europe. These are species that are obligately associated with rotting wood, a habitat that is largely absent from young, even-aged tree plantations or from older forests that are sanitized or managed by removal of dead trees and wood. Most of the saproxylic habitats are afforded by old-growth forests with adequate numbers of injured, bleeding, moribund, or dead trees available. Different species of insects are associated with different types of decay, including whether or not the trees are still standing. Other variables include whether the trees are large or small, whether they are exposed to light and low humidity or are shaded throughout the day, and how long the process of decay has proceeded. Rot holes (some of which contain water), sap runs, dead branches on otherwise healthy trees, and loosened bark are all microhabitats that have specialized insect (including Diptera) faunas. The saproxylic faunas of other regions have not been assessed, but as deforestation proceeds throughout the world, Diptera associated with this habitat will likely be threatened.

Wetlands have been greatly reduced in many areas of the world, in many instances for the express purpose of eliminating biting fly vectors of disease. The loss of such habitats, however, also eliminates populations of other aquatic Diptera that have important ecological roles in the environment.

Another group of Diptera whose conservation needs are relatively well known is the endangered fauna of Hawaii. There are several species of *Drosophila* that have been proposed for listing as endangered species.

Because of their poor public image, Diptera are usually not considered flagship species for conservation projects. In the southwestern United States, however, an endangered species of mydid fly has been used to spearhead efforts to halt the final destruction of an endangered habitat in the Los Angeles area. The Delhi Sands Giant Flower-Loving Fly (*Rhaphiomidas terminatus abdominalis*) is the largest and most identifiable of a number of threatened taxa that live in this habitat, which is reduced to just a few acres in extent. The other subspecies of this taxon, *R. t. terminatus*, previously went extinct when its habitat in the coastal dunes near the Los Angeles International Airport was almost completely destroyed by urbanization. Other species of *Rhaphiomidas* are also considered endangered, as are other Diptera living in small, isolated, sandy habitats.

In general, the conservation status of Diptera on a worldwide scale is unknown. Undoubtedly many species are lost to deforestation, but only those in a few well-known groups (like syrphids) or in habitats that are of interest to some people (wetlands, sand dunes) have been studied. The situation is unlikely to change until we know much more about the systematics of flies.

See also: Beetles. Butterflies. Hymenoptera. Insects, Overview. Moths

References

Cumming JM (1994) Sexual selection and the evolution of dance fly mating systems (Diptera: Empididae: Empidinae). *Canadian Entomologist* 126: 907–920.

Evenhuis NL (1994) *Catalogue of the Fossil Flies*. Leiden: Backhuys.

Feener Jr. DH and Brown BV (1997) Diptera as parasitoids. *Annual Review of Entomology* 42: 73–97.

Ferrar P (1987) A Guide to the breeding habits and immature stages of *Diptera Cyclorrhapha. Entomonograph 8*, parts 1 & 2. Leiden, Copenhagen: E. J. Brill/ Scandinavian Science Press.

McAlpine JF (ed.) (1989) *Manual of Nearctic Diptera*, vol 3. Agriculture Canada Monograph No. 32.

McAlpine JF and Munroe DD (1968) Swarming of Ionchaeid flies and other insects, with descriptions of four new species of Lonchaeidae (Diptera). *Canadian Enomologist* 100: 1154–1178.

Sivinski J, Marshall S, and Petersson E (1999) Kleptoparasitism and phoresy in the Diptera. *Florida Entomologist* 82: 179–197.

Smith KGV (1989) An introduction to the immature stages of British flies: Diptera larvae, with notes on eggs, puparia and pupae. *Handbooks for the Identification of British Insects* 10(14).

Yeates DK and Wiegmann BM (1999) Congruence and controversy: Toward a higher-level phylogeny of Diptera. *Annual Review of Entomology* 44: 397–428.

Food Webs

Gary R Huxel, University of Arkansas, Fayetteville, AR, USA
Gary A Polis†, University of California, Davis, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Community The most practical definition is a set of species that interact at a given location.
Connectivity web This type of food web illustrates only feeding links without reference to strength of interaction or energy flow.
Detrital shunts Energy and nutrients from the saprovore web re-enter the plant herbivore predator food web when detritivores are eaten by predators that also eat plants, herbivores, or other predators.
Donor control Consumer population growth is affected by their resources but consumers do not affect the renewal rate of these resources and hence cannot depress their resources.
Ecosystem Set of one or more communities and their abiotic environment.
Energetic web This type of food web quantifies the amount of energy (or material) that flows across links joining species.
Food chain Representation of the links between consumers and their resources, for example, nutrients → plant → herbivore → carnivore. In these representations, energy or material flows up the chain in a linear fashion. In addition, a food chain can be a linear set of species within a food web.
Food or biomass pyramid Graphic representation of the energy or biomass relationships of a community, in which the total amount of biomass, or total amount of energy available, at each successive trophic level is proportional to the width of the pyramid at the appropriate height.
Food web Representation of feeding relationships in a community that includes all the links revealed by dietary analysis.
Functional or interaction web This type of food web quantifies the strength of interaction between species linked using data from manipulative experiments.
Recipient control Consumers substantially depress populations of their resources.
Spatial subsidies Input from other habitats of organic carbon, nutrients, and prey or the movement of consumers. These resources can influence greatly the energy, carbon, and nutrient budget of recipient habitats. In general, nutrient inputs (nitrogen, phosphorus, and trace elements) increase primary productivity; detrital and prey inputs produce numerical responses in their consumers.
Trophic level An abstract classification to describe subsets of species that acquire energetic resources in a similar way on a subset of species (e.g., top carnivores feed on primary carnivores which feed on herbivores which feed on primary producers). In natural systems, most species do not feed strictly on the "trophic level" below them, making the trophic level concept a difficult term to assign operationally to species.

Introduction

Food webs occupy a central position in community ecology. Charles Darwin introduced the concept of an entangled bank in which he envisioned many kinds of species interdependent on each other in a complex manner governed by "laws acting around us." In the simplest context, food webs incorporate the two factors that, a priori, one would consider most fundamental to the success of any one species: resources and enemies. All species must acquire resources (food or nutrients) and suffer energy losses or mortality from predators (Figure 1). The abundance and success of any species is thus a product of these feeding interactions. This inclusion of such "bottom–up" (productivity and resources) with "top–down" (consumption) factors largely determines the distribution and abundance of almost every species on the planet. In particular, freshwater ecologists have enjoyed notable success by concurrently studying the interaction between these variable factors on the regulation of plant and animal abundance and thus the structure of freshwater communities. This research shows the rich dynamical outcomes that can occur when predation and productivity vary and interact within a food web (Figure 2).

Many important advances have arisen from analyses that concurrently incorporate more than one interaction in a food web: keystone predation and herbivory, the intermediate predation and disturbance hypotheses, the size-efficiency hypothesis, trophic cascades, intraguild predation, apparent competition, and the recognition of the importance of indirect effects. The outcome of virtually all interactions within a community can be modified, directly and indirectly, by other members of the food web. This insight penetrates to all areas of community ecology. For example, the results of experiments must be interpreted carefully for at least two reasons. First,

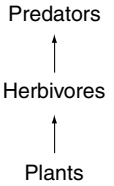


Figure 1 Food chain.

†Deceased.

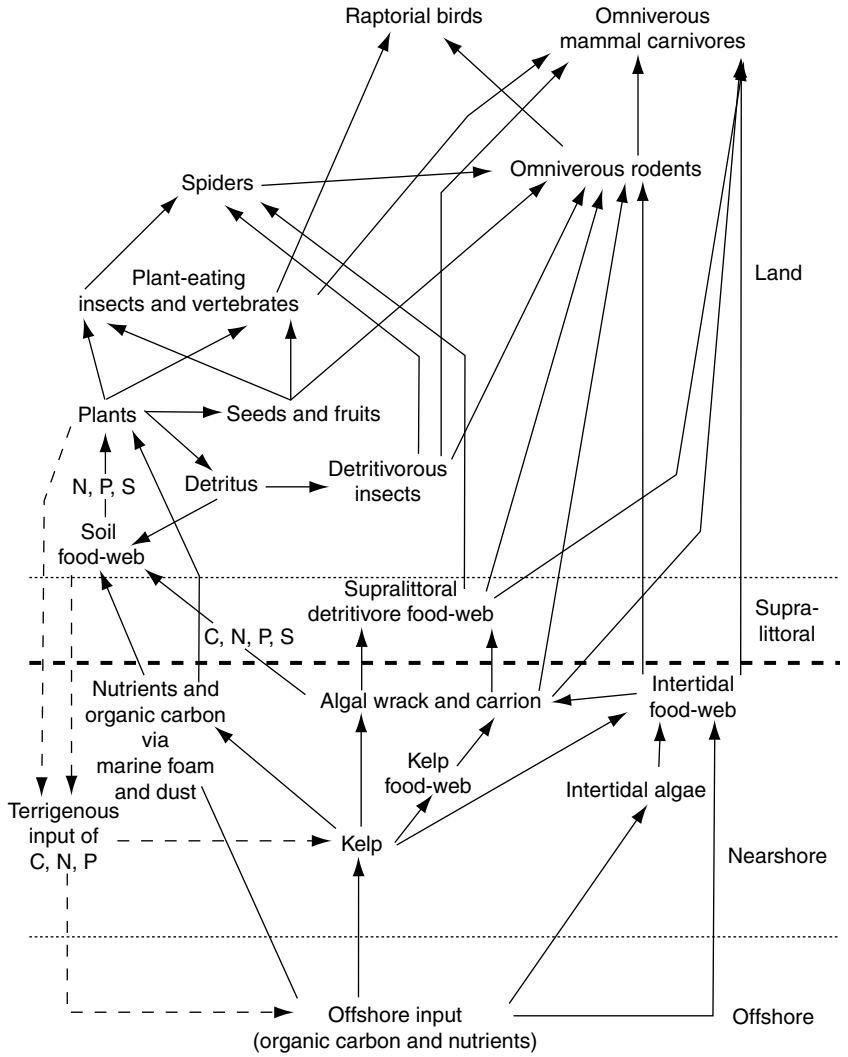


Figure 2 Food web.

indirect effects, moderated by other species in the web, may exert large and sometimes contradictory effects to the direct effects of the manipulation. Thus, under some food web configurations, removal of a predator may directly increase the level of its prey or may actually cause the prey to decrease because of indirect interactions. Second, changes in species dynamics putatively caused by one factor may actually be a product of a second process.

Types of Food Webs

Food web research has grown at a tremendous rate and taken a diversity of forms. Not surprisingly, ecologists have diverged in their methods, emphases, and approaches. Nevertheless, trophic relationships in communities can be delineated in three basic ways. Paine (1980) and Polis (1991) distinguished three types of food web that evolved from the ecological studies (**Figure 3**). The first is the classic food web, a schematic description of connectivity specifying feeding links. Such connectivity webs simply demonstrate feeding relationships. Examples of these are the early food webs of Forbes and Summerhayes and Elton (**Figure 3**). The second web type is also descriptive, quantifying the flow of energy and matter through the community. These energetic webs quantify the flow of energy (or materials) between trophically connected species. Examples of this type of food web include intertidal communities in Torch Bay, Alaska, and Cape Flattery, Washington (Paine, 1980). The third type use experiments to dissect communities to identify strong links and dynamically important species. Such interaction of functional webs demonstrates the most important connections in an ecosystem (**Figure 3**). These food webs depict the importance of species in maintaining the integrity and stability of a community as reflected in its influence on the growth rates of other species. They require experimental manipulations of the community (e.g., by removal or addition of particular species). In the following sections, we discuss the strengths and weaknesses of each approach. Of the three, only the last two have contributed substantially to our understanding of natural systems.

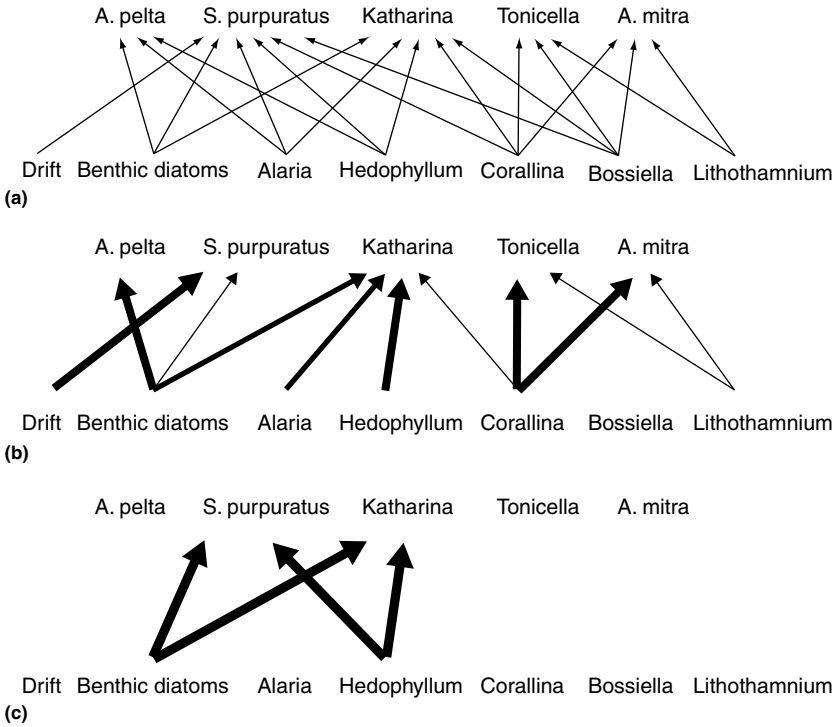


Figure 3 Three conceptually and historically different approaches to depicting trophic relationships, illustrated for the same set of species. The connectedness web (a) is based on observation, the energy flow web (b) on some measurement and literature values, and the functional web (c) on controlled manipulation. Reproduced from Blackwell Scientific Publications.

Connectivity Webs

Connectivity webs are representations of "who eats whom" without inference to the strength or type of interaction and energy flow (**Figure 3**). Early food webs were constructed for essentially two reasons: (1) to depict the interconnectivity of natural systems and (2) to examine issues of "the balance of nature," i.e., to analyze how harmony is maintained through complex predatory and competitive interactions within communities (Forbes, 1887). Such an approach was applied to agricultural systems to examine pests and possible food web manipulations to control pests. As early as the 1880s, beetles were introduced into the US to control agricultural pests. Such control then benefited crop plants *via* an indirect interaction (predator pest prey crop) (following the success of Vedalia, a coccinellid beetle, in controlling cottony-cushion scale in California in 1888, about 50 more coccinellids were introduced in the 1890s).

The knowledge required to construct connectivity webs is straightforward: An approximate, qualitative knowledge of who eats whom is all that is necessary to produce a simple food web, whereas experimental manipulations or quantitative measurements are necessary to construct webs of interaction or energy flow. Consequently, connectivity webs most frequently represent trophic interactions in communities and have received the most attention. Hundreds of such webs slowly accumulated over a century. They were useful to illustrate, in a totally nonquantitiative manner, the feeding interactions within a specific community. Different scientists constructed webs of different diversity, complexity, and resolution, depending on their knowledge of the system and bias or understanding of particular groups. For example, some may emphasize birds and lump all insects as one group. Others will divide the insects into scores of groups and represent one or two bird species.

In the 1970s and 1980s, many theoretical and statistical studies were performed on connectivity webs cataloged from the literature to determine similarities and natural patterns among them. Empirical generalizations were abstracted from data of published connectivity webs. These "natural patterns" largely agreed with predictions made by early food web models. These models showed that food webs were constrained to be quite simple: Each species ate few species and had few predators; the total length of the number of links in a typical food chain was short, usually two or three; omnivory was very rare; and there were a few other patterns. Early modelers argued that the congruence of patterns from the cataloged webs validated the predictions of their models. They thus claimed that their Lotka–Volterra models were heuristic and represented processes that structure real communities. For example, the addition of omnivory to model food webs causes webs to be unstable dynamically and exhibit relative low persistence (time before species are lost). Thus, these models make the prediction that omnivory should be relatively rare in those webs which persist in nature. Comparison of omnivory in cataloged webs relative to its frequency based on chance, shows that omnivory is statistically rare in real webs, as predicted by models. The same general approach was used to

validate other predictions of model webs, for example, short chain lengths.

Thus, modelers soon "explained" these empirically derived patterns. Although these studies, and the connectivity approach, make good food web diagrams, they are flawed to such a great degree that today such analyses are viewed as providing little understanding of natural communities. There are many reasons why this is so, of which only a few are mentioned here.

1. Most species are vastly under-represented in food webs or lumped in taxonomic or functional groups. Most communities have hundreds to thousands of species, but these webs would represent <10–30 species on average. As a consequence, most connectivity webs have severe problems with "lumping" species and taxonomic biases. Some trophic levels are distinguished by species (e.g., birds or fish), whereas other groups suffer a high degree of aggregation, for example, all species of insect or annual plants are represented as one super-species – "insects" or "plants" (**Figure 4**).
2. Most species are highly omnivorous, feeding on many resources and prey so that each has a distinct trophic history and are often at different trophic levels. Because diet is very difficult to delineate, most connectivity webs greatly under-represent the true nature of omnivory. This poses several fundamental problems.
3. Connectivity webs typically only offer a static view of the world and webs are usually idealized representations that show all linkages that occur over large spatial and temporal scales. Therefore much of the important variability and changes due to local environmental conditions are lost. However, studies that compare changes in connectivity over time and space and across environmental gradients (such as those by Mary Power and her group on the Eel River) can provide important insight into community structure and dynamics. One can view connectivity webs as a first step in examining the interactions in communities (i.e., performing "natural history" studies), to be followed by quantification of the fluxes of energy and nutrients (as in energetic webs).

Energetic Webs

Starting with the classic studies of Elton, Summerhayes, and Lindeman, food web studies turned toward quantifying flows of energy and nutrients in ecosystems and the biological processes that regulate these flows. This approach is an alternative to connectivity webs to describe trophic connectedness within communities. This "process-functional" approach explicitly incorporates producers, consumers, detritus, abiotic factors, flow out of a system, and the biogeochemical recycling of nutrients. It views food webs as dynamic systems in time and space. Such an approach necessitated analyzing energy and material fluxes in order to understand the behavior of ecosystems. Thus, a typical analysis would quantify the amount of energy or matter as it travels along different pathways (e.g., plants → consumers → detritus → decomposers → soil). For example, the tracking of energy and dichlorodiphenyltrichloroethane (DDT) through a food web in a Long Island estuary enabled researchers to study bioaccumulation effects on top predators.

The use of energetic webs has provided a rich understanding of the natural world and allowed us to understand much about ecosystems. Several important processes are included in energetic webs. First, they quantify energy and material pathways and key species or processes that facilitate or impede such flows. Second, they include an explicit recognition of the great importance of detritus, a subject virtually ignored in connectivity webs. (10 to $>90\%$ of all primary productivity from different habitats immediately becomes "dead" organic detritus rather than being eaten by herbivores). Third, this approach recognized a great amount of energy, nutrients, and prey that originated outside the focal habitat, which is a key insight to understand natural communities. Thus, energetic webs show how ecosystems function and which species dominate biomass and energy.

Beginning with Lindeman, researchers began to examine the efficiency of transfer from prey species to predator species. It was found that energy transfer is generally inefficient with only about 5–15% of the energy of prey species being converted to energy of predators. Yodzis (1989) used this information to suggest that the length of food chains within a

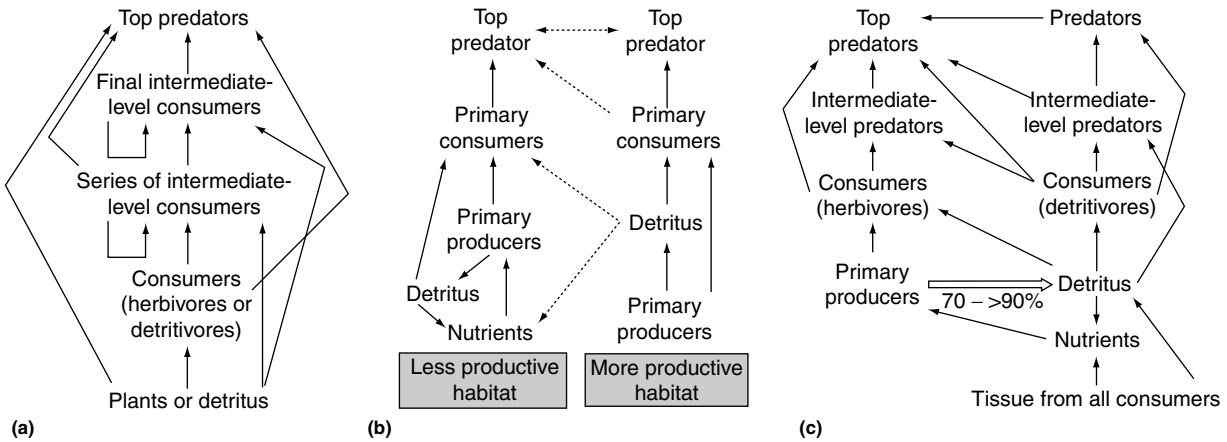


Figure 4 Food web showing aggregation within some trophic levels but not others. (a) The dynamics of omnivory; (b) spatial subsidy; (c) detrital shunts.

community would be set by the amount of energy entering into the base of the chain. This argument was in opposition to Pimm and Lawton's suggestion that food chain length is set by the resilience of the chain. By resilience, Pimm and Lawton, using Lotka–Volterra models, meant the estimated time for model food chains to recover from some disturbance. They argued that frequent disturbances (relative to growth rates of species) would result in shorter food chain lengths. Furthermore, early studies examining the influence of primary productivity (thus, the amount of energy entering a food chain) did not support the hypothesis that food chain length was governed by energy transfer efficiency. However, recent reexaminations of Pimm and Lawton's work suggest that two factors influenced their results – density-dependent regulation of the basal trophic level and food chain structure (the lack of omnivory in their models). Moreover, recent studies of the role of energy efficiency have found that decreases in productivity result in shorter maximum food chains. Thus, the relative role of resilience versus energy transfer in regulating the length of food chains is still debated.

One outcome of the argument for the role of energy transfer as the main governing factor of food chain length is a body of work that examines differences in energy efficiency among organisms. For example, carnivores are found to have greater efficiency than herbivores. Additionally, invertebrate ectotherms have greater efficiencies than vertebrate ectotherms, which in turn are more efficient than endotherms. Yodzis and Innes used this information (and relative body sizes) to parameterize nonlinear predator–prey models.

In summary, the analysis of energy and matter flow is necessary and central to understanding the dynamics of populations and communities. The success of a population is always strongly related to the energy and biomass available to it. Consequently, it is difficult or impossible to understand the dynamics and structure of food webs and interacting populations without incorporating energy flow from below. However, this energetic approach per se, although necessary, is not sufficient by itself to understand the dynamics of communities because energy flow and biomass production are functions of interactions among populations within the food web. The transfer of energy and matter becomes complicated as they pass through many consumers that populate community food webs. For example, increasing the amount of nutrients to plants may increase the biomass of each consumer in the web or may just increase the biomass of a subset of consumers (e.g., only the plants, plants and herbivores, or only the herbivores), depending on the relationship between consumers and their resources. Because of these considerations, pathways must be placed in the context of "functional" food webs to understand the dynamics of energy and material transfer.

Functional or Interaction Webs

Functional or interaction webs use experiments to determine the dynamics within a community. Starting with Connell and Paine, empiricists began to use experiments to examine communities and food webs to discover which species or interaction most influenced population and community dynamics. They manipulated species that natural history or energetic analyses suggested were important. They used either "press" (continual) or "pulse" (singular) experiments to manipulate populations of single species and then followed the response of other species within the food web. The philosophy of these studies was to simplify the complexity of natural systems with the assumption that many species and links between species were unimportant to dynamics. Paine tested this assumption and found that indeed many links between species were weak (essentially zero).

Experimental analyses of food webs are designed to identify species and feeding links that most influence population and community dynamics. These alone are placed into an "interaction web" that, in theory, encompasses all the elements that most influence the distribution and abundance of member species. However, unlike connectivity webs, key species are identified through experiments rather than diet frequency or energy transfer. The initial process of choosing certain species and interactions for experiments and excluding others is subjective, optimally based on strong intuition and a rich understanding of natural history. As the researcher learns more, some elements are discarded and others are subject to further experimentation. Eventually, the community is distilled into an interaction web, a subset including only species that dominate biomass and regulate the flow of energy and matter.

This approach has been used by experimental and theoretical ecologists to produce a rich understanding of the processes that most influence their communities. They have been remarkably fruitful and have introduced many food web paradigms that go to the center of ecology, for example, keystones species, the intermediate disturbance or predation hypothesis, the size–efficiency hypothesis, top–down and bottom–up control, trophic cascades, and apparent competition.

However, this approach is not without limitations. Three major problems stand out. First, many statistical shortcomings can beset experimental manipulation of food webs. For example, replications are commonly difficult (time-consuming and expensive) and therefore experiments often lack the statistical power necessary to avoid type II statistical errors (significant biological differences exist among treatments but low sample size precludes their detection statistically). Second, the number of possible experiments is almost infinite. Which ones should be conducted, and which species should be manipulated?

The third and perhaps most troublesome problem is that experiments isolate a subset of species and links from the community food web, largely ignoring how manipulations interact with the remainder of the community. Thus, unobserved, indirect, or higher order interactions may exert important effects on the dynamics of experimental species and, in theory, make the outcome of experiments indeterminate. For example, predators are thought typically to suppress their prey. However, if a predator is omnivorous, not only eating the prey but also consuming a more efficient predator on the same prey (i.e., it is an "intraguild predator,") it may actually relax the predation load on their shared prey, thus increasing the shared prey's abundance. For example, guilds of biological control agents must be carefully structured because some species eat not only the host but also other predators/parasitoids and thus their presence decreases the number of control agents and increases target pest populations. Many other

cases exist in which consumers, *via* such intraguild predation, may indirectly facilitate its prey while concurrently exploiting it *via* direct consumption. Another example of indirect effects mediated by other than studied "focal species" is shown by the interaction between Australian bell miners and their homopteran food ("lerp.") After these birds were removed experimentally, the insects first increased greatly in number and then vanished when other bird species invaded the now undefended miner territories. Thus, the apparent effect of leaf miners on lerp insects (here, suppression or facilitation) depends on when the insects were surveyed. Such complications have undoubtedly interfered with clear interpretation of many experiments. The caveat is clear: Experiments can be indeterminate, producing contradictory, counterintuitive, or no results, depending on the relative strengths of the direct and indirect effects.

These problems can be anticipated and partially negated with the application of good intuition of the natural history of the system and important mechanisms. Such intuition is a product of intimate empirical knowledge gained through observation and guided by a conceptual awareness of which interactions are potentially important. Initially, this process is essential to design the appropriate experiments and identify, which species and trophic links may be dynamically important. At the end, experiments must be interpreted in a food web context to assess possible indirect and higher order effects. Experimental results must be complemented with good descriptive, mechanistic, and comparative data to produce a deep understanding of the system. This is one role for energetic and dietary data. Experiments in the absence of natural history often do not succeed and may mislead.

The important messages from this section are that the complex food webs of natural communities can be simplified and understood by isolating key species and links into "interaction webs;" experiments are absolutely necessary for this process, and experiments must be designed and interpreted with sound intuition based on natural history and theory.

Omnivory and the Structure of Food Webs

It is necessary to discuss feeding connections in more detail. Empirical research and logic have shown that the vast majority of consumers on this planet are very omnivorous, feeding on many types of food throughout the entire food web. This is not to say that all species are so catholic in their diets. Specialists abound, for example, many herbivores or parasites consume only specific plants or hosts. However, these form a minority of consumers. The ubiquity of omnivory carries many implications for our efforts to produce theory and models to understand how food webs operate in and shape natural systems.

Omnivory occurs ubiquitously when consumers eat prey from general classes of prey, such as arthropods, plankton, soil fauna, benthos, or fish. The existence of multiple trophic types within these classes causes consumers to feed on species from many trophic levels. For example, "arthropodivores" eat whatever properly sized arthropods are available (e.g., predaceous spiders and insects and insect parasitoids, herbivores, and detritivores) without pausing to discriminate among their prey according to trophic status. For example, in the Coachella Valley desert delineated by Polis (1991) in more than 10 years of study, predaceous and parasitoid arthropods formed 41% of the diet of vertebrate and 51.5% of invertebrate arthropodivores, with the remainder of the diet being herbivorous and detritivorous prey. Similarly, inspection of diet data of planktivores, piscivores, "insectivores," carnivores, or benthic feeders reveals that such different channel omnivory is almost universal with the exception of those few taxa that specialize on a few species of prey. Many other systems are dominated by omnivores, for example, Arim and Marquet (2004) found omnivorous links between 86.7% and 72.8% of intermediate and top predators, respectively, in a set of natural food webs. For intermediate predators this was significantly greater than the 46.2% predicted from null models of the same food webs, but for the top predators the difference was not significant (expected = 67.7%), but is still large.

Another important type of omnivory occurs when consumers eat whatever resources are available or abundant at a particular time or place, regardless of their trophic history. When analyzed, the diet of a single species usually shows great differences through time (e.g., seasonally) and space (patches or habitats). Prey exhibit three general phenologies: pulsed (population eruptions lasting a few days or weeks), seasonal (present for 2–4 months), and annual (available throughout the year). Feeding on prey from all three phenologies produces diet changes over time for almost all nonspecialist consumers. Furthermore, many (most?) vertebrates opportunistically switch from plant to animal foods seasonally. For example, granivorous birds, rodents, and ants primarily eat seeds but normally feed on the abundant "arthropods" (= insects from all trophic levels and spiders) that appear during spring. Alternately, many omnivorous, arthropodivorous, and carnivorous species consume significant quantities of seeds or fruits. In the Coachella Valley, 79% of 24 primary carnivores eat arthropods and plants; for example, coyotes eat mammals (herbivorous rabbits, rodents, and gophers; arthropodivorous antelope and ground squirrels; carnivorous kit foxes and other coyotes), birds (including eggs and nestlings, e.g., carnivorous roadrunners; herbivorous doves, and quails), snakes, lizards, and young tortoises as well as scorpions, insects, and fruits. In New South Wales, 15 of 27 ant species are "unspecialized omnivores" eating nectar, seeds, plant parts, and a broad range of living and dead insects, worms, and Crustacea. Overall, it appears that most consumers eat whatever is available and whatever they can catch.

"Life-history" omnivory describes the great range of foods eaten during growth and ontogeny by most species (the "age structure component" of dietary niche breadth). Such omnivory includes abrupt diet changes in species undergoing metamorphosis (e.g., many marine invertebrates, amphibians, and holometabolic insects) and gradual diet changes in "slowly growing species" (e.g., reptiles, fish, arachnids, and hemimetabolic insects). Changes at metamorphosis can be great; for example, 22% of the insect families in the Coachella Valley desert community undergo radical change in diet – larvae are predators or parasitoids and adults are herbivores. Although not as dramatic, significant changes characterize slowly growing species so that differences in body size and resource use among age classes are often equivalent to or greater than

differences among most biological species. Life history omnivory expands the diet of species throughout the entire animal kingdom with the exception of taxa that use the same food species throughout their lives (e.g., some herbivores) and those with exceptional parental investment (e.g., birds and mammals) so the young do not forage for themselves.

"Incidental omnivory" occurs when consumers eat foods in which other consumers live. Thus, scavengers and detritivores not only eat carrion or organic matter but also the trophically complex array of microbes and macroorganisms that live within these foods. Frugivores and granivores commonly eat insects associated with fruits and seeds. Predators eat not only their prey but also the array of parasites living within the prey. In each case, consumers automatically feed on at least two trophic levels.

These types of omnivory are widespread and common. Their ubiquity poses many questions. First, how does omnivory affect food web structure? Most obviously, it increases complexity and connectivity. Second, can we ignore omnivory in the analyses of food webs? By its very nature, omnivory causes consumers to have a great number of links, each of which may be numerically unimportant in the diet. For many reasons delineated later, we cannot arbitrarily ignore apparently minor diet links if we hope to understand dynamics.

Patterns of Biomass and Energy in Food Webs

Primary productivity is among the most fundamental biological processes on the planet, transferring the energy locked in light and various inorganic molecules into forms useful to sustain producers and the diversity of consumers. What factors control primary productivity and regulate its distribution among plants, animals, and microbes? How do changes in primary productivity work their way through a food web to alter the abundance and biomass of herbivores to predators and detritivores? As discussed later, such key questions are best assessed using a food web approach. However, considerable controversy exists regarding the exact way that food web structure influences community and ecosystem dynamics.

Trophic Levels, Green Worlds, and Exploitative Ecosystems

Ecological research has amply demonstrated that food webs in nature contain hundreds to thousands of species, reticulately connected *via* multiple links of various strength to species in the autotroph and saprophagous channels and in the same and different habitats; omnivorous, age-structured consumers are common. Nevertheless, much food web theory still relies on the idealization of trophic levels connected in a single linear chain (plant herbivore carnivore). Here, we evaluate this simplification and some of its implications. In particular, we focus on two grand theories whereby food webs are considered to be central to community organization.

The trophic level ideal in a simple linear food chain has had great appeal. Trophodynamics sought to explain the height of the trophic pyramid by reference to a progressive attenuation of energy passing up trophic levels, envisioned as distinct and functionally homogeneous sets of green plants, herbivores, primary carnivores, and, sometimes, secondary carnivores. This is a bottom–up community theory based on the thermodynamics of energy transfer. In counterpoint, Hairston, Smith, and Slobodkin's green-world hypothesis (GWH; Hairston *et al.*, 1960) is primarily a top–down theory, with abundance at each level set, directly or indirectly, by consumers at the top of the chain. Thus, carnivores suppress herbivores, which releases green plants to flourish. These and earlier theoretical studies attempted to simplify food webs greatly to find generalities among them. GWH reduced complex webs to food chains in which species were pigeonholed into specific trophic levels. This allowed for predictions on how higher trophic levels (e.g., predators) influenced the dynamics of lower trophic levels (e.g., primary producers).

Oksanen *et al.*'s (1981) exploitation ecosystem hypothesis (EEH) generalizes GWH to fewer or more than three trophic levels. Trophic cascades are examples of food chains that behave approximately according to EEH. Trophodynamics and EEH each rely on the integrity of trophic levels and the existence of a single, albeit different, overwhelming mechanism that imposes structure on ecosystems. EEH proposes a conceptual framework of "exploitation ecosystems" in which strong consumption leads to alternation of high and low biomass between successive levels. Even numbers of "effective" trophic levels (two or four levels) produce a low-standing crop of plants because the herbivore population (level 2) flourishes. Odd numbers (one or three levels) result in the opposite effect: Herbivores are suppressed and plants do well. Proponents of EEH differ on subsidiary points, the first being the role of bottom–up effects in which primary productivity sets the number of effective levels. The most productive systems support secondary carnivores and therefore have four levels and low-standing crops of plants. Low-productivity systems (e.g., tundra) support only one effective level – plants. More productive habitats (e.g., forests) have three. Productivity is never high enough to support more than three effective levels on land or four in water. Other studies argue that physical differences between habitats, by affecting plant competition and consumer foraging, cause three levels on land and four in water.

EEH definitions of trophic levels are distinctive and adopt the convention that trophic levels occur only if consumers significantly control the dynamics or biomass of their food species. Without top–down control, consumers do not comprise an effective trophic level regardless of biomass or number of species involved. Supporters of EEH have noted that only when grazers regulate plants are grazers counted (as a trophic level), and only when predators regulate grazers they are fully counted. Thus, considerations of food chain dynamics do not become stranded in the immense complexity of real food webs. Whereas, GWH trophic levels are based on energy deriving from primary productivity. Thus, "trophic level interactions … weight particular links in the food web for their energetic significance." A trophic level is "a group of organisms acquiring a considerable majority of its energy from the adjacent level nearer the abiotic source." Despite these differences, both EEH and GWH theory argue that variability in the number of trophic levels exerts profound consequences on community structure and dynamics.

Considerable controversy exists as to the validity of GWH and EEH. The consensus has swung against these grand

theories. Numerous arguments and empirical observations suggest that such processes operate occasionally in water but never on land. Basically, the complexity observed in natural systems does not conform to the reality of simple trophic levels. It appears that the notion that species clearly aggregate into discrete, homogeneous trophic levels is a fiction, arising from the need of the human mind to categorize. Especially in speciose systems, groups of species with diets of similar species do not occur. Omnivory, ontogenetic and environmentally induced diet shifts, and geographical and temporal diet heterogeneity all obscure discrete trophic levels. Even plants do not easily form a single level; higher plants have diverse crucial trophic and symbiotic connections with heterotrophs and many phytoplankton are mixotrophic, obtaining energy *via* photosynthesis, absorption of organic molecules, and ingestion of particles and bacteria. With increasing diversity and reticulation in webs, trophic levels blur into a trophic spectrum rather than a level. These species-individualistic and continuous "trophic spectra" are a reasonable alternative to the simplistic construct of homogeneous trophic levels.

Complex Food Webs, Multichannel Omnivory, and Community Structure

Polis and Strong (1996) offered a framework in the context of functioning community webs as an alternative to theories based on discrete trophic levels. Substantial evidence indicates that most webs are reticulate and species are highly interconnected, most consumers are omnivorous on foods (frequently on both plants and animals) across the trophic spectrum during their life history, most resources are eaten by many species across the trophic spectrum, plants are linked to a variety of species *via* trophic mutualism, most primary productivity becomes detritus directly, detrital biomass reenters the autotroph channel of the web when detritivores and their predators are eaten by consumers that also eat species in the herbivore channel, and species are often subsidized by food from other habitats.

They proposed that such trophic complexity pervades and generally underlies web dynamics. High connectance diffuses the direct effects of consumption and productivity throughout the trophic spectrum. Thus, consumer and resource dynamics affect and are affected by species at multiple positions along the trophic spectrum rather than interacting only with particular trophic levels. Consumer density is elevated and they often persist by eating resources whose abundance they do not influence (i.e., the interaction is "donor controlled.")

Such dynamics are illustrated by focusing on top–down interactions. Some consumers exert "recipient" control on some resources and, occasionally, produce trophic cascades. Polis and Strong (1996) suggest that such control is often enabled by omnivorous feeding and various consumer subsidies which are usually donor controlled. Here, the transfer of energy and nutrition affects dynamics; numerical increases in consumer abundance occur from eating diverse resources across the trophic spectrum in the autotroph channel, from detritivores and detritus, the saprovore channel, other habitats, and across their life history. Consumers, so augmented, exert recipient control to depress particular resources below levels set by the nutrition traveling through any particular consumer–resource link (analogous to the effects of apparent competition). Top–down effects arising from such donor-controlled, "multichannel" omnivory are depicted in Figures 2 and 4. Strong consumer-mediated dynamics occur precisely because webs are reticulate and groups of species do not form homogenous, discrete entities.

Multichannel omnivory has two essential effects on the dynamics of consumers, resources, food webs, and communities. First, it diffuses the effects of consumption and productivity across the trophic spectrum rather than focusing them at particular trophic levels: It increases web connectance, shunts the flow of energy away from adjacent trophic compartments, alters predator–prey dynamics in ways contra to EEH assumptions, and thus disrupts or dampens the ecosystem control envisioned by EEH. For example, Lodge showed that omnivorous crayfish can depress both herbivorous snails (consistent with GWH and EEH) and macrophytes (inconsistent).

Second, omnivory can affect dynamics in a way analogous to apparent competition. Feeding on "non-normal" prey can increase the size of consumer populations (or sustain them during poor periods), thus promoting top–down control and depression of "normal" prey. Frugivory, herbivory, granivory, detritivory, and even coprophagy form common subsidies for many predators. Vertebrate carnivores consume amply from the lower web without markedly depleting these resources. Does energy from fruit help carnivores depress vertebrate prey (e.g., herbivores)? Arthropodivory by seed-eating birds is the norm during breeding, with insect protein crucial to nestlings. Arthropodivory by granivores (and conversely, granivory by arthropodivores) must enhance bird populations and thus reduce seeds (arthropods) to a greater degree than if diets were not so augmented.

Trophic Cascades or Trickle

One prediction of GWH and EEH is that communities are structured by trophic cascades. Trophic experiments to test cascades use two methods: a bottom–up approach by increasing a resource (e.g., nitrogen or phosphorus) or a top–down approach that adds a top predator to a system. In the former, trophic cascades lead through a set of intermediate steps to increase densities of particular species or trophic groups higher in the web. In the latter, the top predator suppresses the trophic level below leading to increased densities two levels below. Thus, the expected responses should follow GWH/EEH predictions where alternating trophic levels are arranged with opposite densities (common – rare – common). For example, in a tritrophic (three-level) food chain, an increase in nutrients results in increases in the primary producer (plant) trophic level, decreases in the primary consumer (herbivore) level, and an increase in the top consumer level.

Proponents GWH and EEH suggest that strong trophic cascades occur in numerous food webs whereby entire trophic levels alternate in abundance *via* cascading food web interactions. However, empirical evidence shows that such cascades rarely or never occur on land and are apparently only present in a few aquatic communities. What determines whether a strong trophic cascade occurs or food web interactions weaken

to become a trophic "trickle"? One major consideration is the efficiency of energy and resource transfer up the food chain. Highly efficient transfers lead to large numbers of top predators/consumers that would affect top–down control and strong cascades. Any factors that decrease the efficiency of energy/resource transfer would lessen the top–down control. In accordance with Polis and Strong's (1996) multichannel omnivory, an increasing list of factors have been examined to explain the differences between GWH/EEH expectations and experimental results and observations of natural communities that generally show weak or no trophic cascades. These factors include omnivory, ontogenetic shifts, edibility, food quality, ecological stoichiometry, cannibalism, disease, body size refuges (for prey), allochthonous resources, seasonality, life-history characteristics, predator avoidance behavior, and spatial and temporal heterogeneity in the availability of resources.

Current Topics/Trends in Food Web Studies

Here, relatively under-studied aspects of food webs perceived to be central to understanding populations, communities, and ecosystems are identified. Some of the topics are now focal points for food web research, both empirical and theoretical.

Food Webs as Open Systems

Recent methods of tracing stable isotopes through a food web can provide much information on feeding relationships and on the sources of productivity that drive communities. For example, using stable isotopes or diet data, one can determine whether a community utilizes resources that originate in the benthic or pelagic zones of lakes or both.

Virtually all natural systems are open and can exhibit tremendous spatial heterogeneity. Great spatial heterogeneity exists and nutrients and organisms ubiquitously move among habitats to exert substantial effects. However, food web studies have tended to focus on communities at a given site without regard to potential interactions with the surrounding habitat. Thus, little attention has been given to the fact that food web structure and dynamics are influenced by the movement of resources and organisms across habitat boundaries. Trophic linkage between habitats depends on the degree of differentiation in habitat structure and species composition. Systems that are moderately different tend to have broader transition zones and greatly overlap in species composition; these include grassland–forest, littoral–sublittoral, and benthic–pelagic zones. Habitats that have significant and abrupt changes in structure and species composition occur at the land–water interface.

Moving resources (energetic or nutrients) can be utilized by different trophic types and the organisms that move across boundaries may also differ trophically (e.g., predators and prey). Studies of communities on island systems have shown that most of the allochthonous inputs (i.e., input from other habitats) from the ocean are available to detritivores, predators, and scavengers. Such movement of nutrients, detritus, food, prey, and predators is absolutely ubiquitous, occurring in virtually all communities and across all habitats. Some systems heavily dependent on allochthonous inputs include caves; mountaintops; snowfields; recent volcanic areas; deserts; marine filter-feeding communities in currents; soil communities; the riparian, coastal areas; lakes; rivers; and headwater streams that receive watershed inputs. However, all systems depend on allochthonous inputs. For example, recent work shows that plant productivity in both the Hawaiian Islands and the Amazon forest is dependent on phosphorus input from thousands of miles away (China and Africa, respectively). The migrations (e.g., songbirds or geese) and movement of herbivores (e.g., wildebeest or hippopotamuses) can also result in large energetic flows across habitats.

Allochthonous inputs into the top level include carrion or carcasses, the movement of prey species into the habitat, and movement of predators across habitats. For example, the Allen paradox describes cases in which secondary production within streams is insufficient to support levels of fish production in them. Similarly, studies of coyote populations along the coast in Baja California demonstrate that they are highly subsidized by inputs from the ocean (about half of their diet) and are able to maintain three to more than 10 times higher density than in adjacent inland areas. Predators moving along the interface between ecosystems (i.e., shorelines, riverbanks, and benthic and pelagic systems) can utilize resources across habitat. The river continuum concept argues that allochthonous resources entering into small headwater streams provide much of the productivity for organisms downstream in larger order streams. These allochthonous resources include prey, dissolved and particulate organic matter, and litter fall. Such inputs also power estuarine systems in which rivers carry allochthonous inputs into estuaries. Similarly, runoff from terrestrial systems into aquatic systems (and vice versa) provides litter, dissolved and particulate organic matter, and prey.

Spatial coupling can be key to dynamics. For instance, arboreal anole populations, subsidized by insects imported from light gaps, increase so as to suppress some predators and herbivores. Abundant detrital kelp from the sublittoral zone promotes dense intertidal limpet and urchin populations that then graze noncoralline algae to low cover. Allochthonous subsidies commonly influence stream systems: Leaf fall subsidizes herbivores, which in turn depress algae. Spiders that live along the coasts of streams, rivers, lakes, or the ocean are often very dense because they feed on aquatic insects. These spiders can then depress herbivores and thus increase the success of plants on which they live. Such spatial subsidies appear to be the foundations of most of the well-known trophic cascades. All these interactions are donor controlled: Consumers do not affect the rate of import, availability, or dynamics of the allochthonous resources. However, subsidies allow consumers to be more abundant than if supported solely by *in situ* resources, with consequent suppression of *in situ* resources decoupled from *in situ* productivity.

A common thread that has begun to link most thinking on food webs is that they are dynamical systems that vary over space and time. This approach has been liberating to ecologists, both empirical and theoretical. Recent empirical studies have found that communities and food webs contain multiple pathways that allow them to respond to environmental change and disturbance.

Detritus

Little of the energy fixed by plants passes directly into the grazing food chain – herbivores eating plants and then eaten by carnivores. Most of this primary productivity is uneaten by herbivores (median $>80\%$ on land, $\sim 50\%$ in water). What happens to this dominant chunk of the world's productivity? Is the detrital web a self–contained sink internally recycling energy and nutrients or a link that affects the population dynamics of the larger species?

Uneaten plants (and animals) enter the detrital web, in which they are processed by microbes, fungi, and some animals. Although some ecosystems are net accumulators of undigested biomass (e.g., carboniferous bogs and forests that supply today's oil and gasoline), most ecosystems do not accumulate plant biomass. Rather, it is soon digested by detritivores, with nutrients and energy passing through "functional compartments" composed of diverse microbes and animals. Several factors regulate the flow and availability of detritus to detritivores and then onto other consumers. A major question is rather whether the detrital community is a sink that metabolizes most of this energy or a link that passes this energy up the food chain.

An unknown fraction of detrital energy and nutrients re-enter grazing food chains when some detritivores are eaten by predators that also eat herbivores (e.g., a robin eats an earthworm). Such "detrital shunts" are common, interweaving energetics and dynamics of biophages and saprophages. Bypassing herbivores, this linkage can affect herbivore regulation in a manner analogous to the spatial subsidies to consumers discussed previously. Predator populations, subsidized by detritivorous prey, can increase and suppress other predators or herbivores.

The exact effect of detrital shunts depends on the relative benefits for each species and where detritus re-enters (to producers, herbivores, and intermediate or higher consumers). For example, nutrients from detritus greatly influence plant productivity; models show that a 10% reduction in detritus can cause a 50% reduction of plant biomass. The dynamics of consumer control within the detrital web and those produced by infusion of detritivores into the grazing web are undoubtedly crucial to community structure and dynamics. For example, detrital shunts to predators in the grazing chain can create the appearance of a simple linear trophic cascade, but with the difference that nutrition from detritivores sustains or elevates predators to levels sufficient to suppress herbivores.

Age Structure Effects in Food Webs

Almost all species display complex life cycles, marked by moderate to radical changes in diet and habitat; such life histories fundamentally must affect every species with which they interact. However, our understanding of how age- and stage-structured processes affect food webs and communities is embryonic.

Life history omnivory describes shifts in diet during development; often they are accompanied by ontogenetic changes in habitat. Diet can change substantially either discontinuously (e.g., at metamorphosis) or slowly with growth. Such life histories are widespread; an estimated 80% of all animal species undergo metamorphosis. Changes in resource use can be dramatic (e.g., predaceous juveniles, plant-feeding adults in parasitoids and many other insects, and herbivorous tadpoles and predaceous frogs and toads), with prey size variation as great as three or four orders of magnitude. Even among nonmetamorphic species, diets change greatly with age, with diet differences among age classes often more distinct than those among most species. Thus many species can be classified as life-history omnivores, which are species that feed on different trophic levels over their various life stages.

Overall, complex life histories and age structure omnivory can exert diverse and profound effects on the dynamics of populations and food webs. For example, they can either impede consumer control or amplify resource suppression *via* dynamics similar to those of spatial subsidy or detrital shunts.

The Roles of Nutrients and Stoichiometry

Animals require both energy and a variety of "nutritional requisites" to grow, complete their life cycle, and reproduce. Important nutrients include nitrogen, phosphorus, some trace elements, fatty acids, and vitamins. Nitrogen is an integral component of many essential compounds: It is a major part of amino acids, the building blocks of proteins, including the enzymes that control virtually all cellular processes. Other nitrogen compounds include nucleic acids and chlorophyll. Phosphorus is used for adenosine triphosphate (ATP, the energy currency of all cells), nucleic acids (DNA and RNA), and phospholipids, particularly in cell membranes.

The availability of nutritional requisites constrains growth and reproduction in virtually every species. Nitrogen and phosphorus are particularly important. The ratio of carbon to nitrogen (C:N) in plants ranges from 10:1 to 30:1 in legumes and young green leaves to as high as 600:1 in some wood. The C:N ratios in animals and microbes are much lower, ordinarily between 5:1 and 10:1. Such differences in C:N ratios between plants and their consumers lower the rate of decomposition by microbes. There is an ample evidence that heterotrophs chronically lack adequate nitrogen to grow or reproduce optimally. The importance of nutritional restriction is reinforced by the foraging literature that clearly shows that herbivores choose their foods based on nutrient as well as energy content.

In many cases, phosphorus availability constrains herbivore success. The Redfield ratio describes the approximate stoichiometric mix (110 C:250 H:75 O:16 N:1 P) of elements found in marine systems. In particular, the N:P ratio crucially determines productivity and species composition. Thus, energy (C−C bonds) and nitrogen could be abundant, but neither individuals nor populations grow maximally because phosphorus is insufficient. Because phosphorus is essential to cell division (and thus reproduction), a high N:P ratio especially limits the growth of organisms that have high potential r_{max}, such as most herbivores and detritivores. These organisms are key to the potential regulation of plant biomass (and "detritus.") Evidence suggests that high N:P ratios can impede trophic cascades. For example, *Daphnia*, a key to many lake cascades, respond sufficiently rapidly to phytoplankton productivity to depress plant biomass. In lakes with inadequate phosphorus, slower growing copepods replace *Daphnia*; these copepods do not have the reproductive capacity to depress phytoplankton biomass.

Ecologists are beginning to understand how stoichiometry and nutritional balance affect population and food web dynamics. Nevertheless, it is extremely likely that herbivore growth is often less than maximal solely because their environment does not provide sufficient quantities of all key nutritional requisites. In fact, the greatest disparity in biochemical, elemental, and stoichiometric composition in the entire food web occurs at the link where herbivores convert plant material into animal tissue. The implication is clear: Even in a world full of green energy, many or most herbivores cannot obtain enough requisite resources to grow, survive, or reproduce at high rates. Nutritional shortages regulate herbivore numbers and often limit their effects on plant biomass.

Recent theoretical studies of the role of food quality in terms of edibility and nutrient content show that low food quality can greatly influence consumer–resource interactions. This has two important consequences. First, low food quality reduces the growth rate of the consumer, making that interaction more stable. Second, in systems in which multiple resources could be limiting, the addition of large amounts of a single resource (such as nitrogen or phosphate) may increase that resource to a level at which it is no longer limiting; however, a second resource would become limiting and so on. This sequential limiting of resources means that the addition of a single resource would not push the system into highly unstable dynamics, reducing the probability that the "paradox of enrichment" occurs. Rosenzweig introduced the concept of the paradox of enrichment to explain the addition of a resource leading to the collapse of a consumer–resource interaction. This happens because the addition of the resource drives the population of the consumer to a higher level that results in overcompensation by the consumer (predator) driving the resource (prey) extinct. However, most systems have several potentially limiting resources. For example, Leibold's study of ponds found that nitrogen additions do not lead to strong trophic cascades or the paradox of enrichment because light becomes limiting with relatively modest nitrogen additions.

Interaction Strength

One goal of functional webs is the quantification of interaction strengths within food webs. Various definitions have been used for "interaction strengths." In Lotka–Volterra models, interaction strengths are due solely to the direct interactions between species pairs and are measured on a per capita basis. Estimations of the strength of these direct interactions are fraught with difficulties. Measurements in artificial systems may not allow for behavioral responses. For example, Sih has shown that prey species have different escape mechanisms or routes depending on the species of predator. Thus, when in the presence of two predators, the response of a prey may result in its increased susceptibility to one or the other predator due to a behavior that is not evidenced when only the one predator is present.

Measurements in natural systems are also problematic because they may not account for indirect interactions. Many studies have elucidated the interaction strength among pairs of species. However, indirect effects may play a strong role in determining the realized interaction strength. Thus, Paine has argued that interaction strengths should always be measured in the field with the full complement of natural species present and that these measurements should incorporate all indirect effects. The realized interaction strength accounts for all direct and indirect interactions. For example, predator–prey interactions are functionally negative due to the direct effect. However, the indirect effect of a predator may reduce the number of competitors of the prey species, thus resulting in an overall positive interaction strength (direct + indirect effects). Therefore, potentially strong indirect effects can make mechanistic interpretation of experimental results among species difficult.

Path analysis, a new statistical method, has been used to evaluate causal hypotheses concerning the strengths of interactions in many systems. Path analysis is essentially a multiple regression on each species in which specific causal relationships (e.g., alternative food web configurations), specific experimental treatments, and other interactions are diagrammed in a community interaction web. The community interaction is essentially a food web to which nonconsumptive interactions, such as pollination, competition, and mutualisms are added. Hypotheses for the causal relationships between pairs of species not directly linked can become quite complicated. However, path analysis can test different hypothesized community web structures by accounting for both direct and indirect relationships. Then, experimental manipulations (e.g., species removals or additions) can test predictions of the path analysis.

Can Energetic Webs Provide Insight into Population and Community Dynamics?

A problem in food web studies is how to connect the great amount of quantitative information in energetic webs to population and community dynamics described by functional webs. Much progress would occur if we could determine the dynamical importance of a particular species or feeding link from an inspection of the magnitude of energy transfer or diet composition. Unfortunately, no clear answer is forthcoming. In fact, it appears that even highly quantified information such as the number of calories passed along a certain pathway or the frequency of prey in the diet of a consumer conveys little information about the dynamics of interacting populations because these descriptive parameters do not correlate with interaction strength.

There is no clear rationale to argue that food web dynamics and energetics are necessarily correlated; indeed, logic and evidence suggest that dynamics often cannot be predicted from data on diet or energy flow. The degree of resource suppression is not a function of energy transfer. Consumer regulation of populations need involve little energy transfer and few feeding interactions. For example, removing predatory rats from New Zealand islands increased lizard abundance 3–30 times although lizards formed $<3\%$ of rat's diet. Key regulatory factors may produce much less overall mortality than other factors. Brief, intense predation episodes may net little energy for the predator but may be central to prey dynamics. The consumption of young stages (seeds, eggs, and larvae) may provide trivial energy to a consumer but can

greatly depress prey abundance. Pathogens and parasites form an extreme example: They take little energy, even when they decimate their host populations. In a well-studied food web of the marine benthic community in the Antarctic, Dayton showed that the species apparently exerting the strongest effects on the structure and dynamics of this community would be deemed unimportant from analyses of diet, energy transfer, or biomass.

Such discoveries have stimulated many to argue that, without experimentation, one cannot a priori decide which are strong or weak links. An apparently weak link (in terms of diet or energy transfer) can be a key link dynamically, and an important energetic link may affect dynamics little. No necessary concordance of dynamics with either dietary or energetic measures exists. This insight counters the use of energetics to recognize strong interaction links.

Modeling Food Webs

To many ecologists, early food webs of Forbes, Summerhayes, and Elton and those of Lindeman emphasized the overwhelming complexity of natural systems and the need to simplify them into distinct trophic groups. This perspective was culminated in the green-world hypothesis of Hairston *et al*. (1960). Oksanen *et al*.'s (1981) EEH expanded this view for ecosystems that had fewer or more than three trophic levels and for which the exact number of trophic levels was set by productivity. The top level would then regulate the one below it and this would release the one below it, etc. In this sense, both GWH and EEH suggested that all ecosystems are essentially regulated from the top–down by predation.

Lindeman envisioned the food web (or as he called it, the "food-cycle") as a dynamic system in which energy and nutrients are transferred from one trophic level to the next and recycled. This was an important departure from simply determining feeding connectedness (and from the GWH) in that ecosystems could be regulated from the bottom–up by the flow of energy and materials from the level below. However, much more information and data are required to quantify the transfer of energy (and material) through food webs, but this view allows for a more analytical approach.

MacArthur focused the attention of ecologists on the trophic–dynamic approach with his hypothesis that increasing complexity of community organization leads to increasing dynamic stability. The reasoning was simple: When predators have alternative prey, their own numbers rely less on fluctuations in numbers of a particular species. Where energy can take more routes through a system, disruption of one pathway merely shunts more energy through another, and the overall flow continues uninterrupted.

MacArthur's analytical approach linked community stability to species diversity and food web complexity and it stimulated a flurry of theoretical, comparative, and experimental work. This work may be divided into two contemporary approaches that use food webs to study community structure. The first approach involves the study of the properties of food web diagrams with the goal of uncovering general patterns that suggest mechanisms of community stability. This is done both by comparing food webs from natural communities and by the use of simulation and mathematical modeling to study hypothetical food webs. This research has yielded much of the terminology now associated with food webs and generated a body of food web theory that includes many hypotheses about community structure.

The second approach, which grew from early theoretical and experimental community studies, involves the dynamical analysis of food webs to determine not only the pattern of interactions among the populations in the community but also the relative strengths of those interactions. Dynamic food web analysis also seeks to reveal interactions that are not obvious from simple food web diagrams, so-called indirect interactions. This approach requires the careful merging of experimental and theoretical approaches.

The simplicity of the GWH enabled it to be a reasonable starting point to examine the dynamics of food webs. In general, dynamical models are rooted in a tradition based on the application of Lotka–Volterra equations to communities and advocated by May (1973). One of the major conclusions from these phenomenological models is that complexity (e.g., omnivory and long chains) causes instability in model systems. This conclusion was viewed with skepticism by empiricists because observations from field studies (such as work by MacArthur) suggested that increased complexity should result in increased stability. Recent theoretical investigations into the relationship between stability and complexity have found that assumptions and structure of earlier models may have biased them toward decreased stability with increasing complexity.

Early theoretical studies of interactions and consequences of these interactions in food webs were based on equilibrium dynamics of Lotka–Volterra models. The assumption that ecological systems or species populations have some "equilibrium" around which they fluctuate is totally unrealistic. Furthermore, these early models ignored the central belief of many empiricists that most interactions between species were weak. The outcome of many of these theoretical studies went against common sense intuition and the findings of empirical studies, including that omnivory was destabilizing and therefore rare and that complexity (greater diversity) was also destabilizing. Recent studies that incorporated the findings of mostly weak interactions and nonequilibrium dynamics have found that omnivory and complexity may actually stabilize food webs. This agrees with both the intuition and the current arguments of empiricists who find that many weak interactions occur within food webs and these promote stability.

Recent theoretical studies suggested three factors as important to reduce stability in earlier models: (1) linear Lotka–Volterra equations, (2) using equilibrium solutions to these equations, and (3) the distribution of interaction strengths overly estimated the number of strong links. Many studies have shown that many predator–prey relationships are not linear, but instead predators exhibit saturation such as described by a Holling's type II functional response. Current models take advantage of this and use energetic uptake rates that saturate based on body size relationships. Also, equilibrium solutions to Lotka–Volterra relationships can give biologically unrealistic results because the assumption of equilibrium does not appear to hold in many predator–prey relationships. May and others used a uniform distribution in randomly created model food webs, which resulted in their

webs having an over-representation of strong interaction compared to natural systems. This convention was based on the few early studies that examined the distribution of interaction strengths and suggested that there is a bias for weak interactions. May acknowledged that if the distribution of interaction strengths was not uniform, his results may not hold. Furthermore, recent theoretical studies also suggested that omnivory can stabilize food webs. Paradoxically, researchers using Lotka–Volterra models have found that although on average omnivory decreased stability, those systems in which omnivorous links persisted had the greatest stability. This increased stability may occur in Lotka–Volterra models when randomly created omnivorous links are weak.

In modeling food webs, a key consideration is the functional relationship between a consumer and its resource. As noted previously, Lotka–Volterra consumer–resource relationships are linear (type I). This assumes that the predators do not become saturated and can consume all available prey. Holling introduced nonlinear consumer–resource functional relationships with his disk (now called type II) functional response. This functional response assumes that the capture and consumption/digestion time of prey by the predator limits the amount of prey taken by a predator in a given amount of time. Holling also introduced a third type of functional response (type III) to simulate a predator switching capture of prey when a target prey species becomes rare to a more abundant prey species. Various other functional responses have been introduced, including Ginzburg and Arditi's ratio-dependent functional response. Ratio dependence assumes that the growth rate of the predator is dependent on the ratio of prey and predator densities, whereas in types I–III predator growth rates are dependent only on the prey densities (prey dependent). Ratio-dependent models predict that all trophic levels increase proportionately, whereas prey-dependent models predict the alternating pattern of GWH and EEH. The arguments against ratio dependence arise from the lack of a mechanistic basis for the model. Proponents of ratio dependence argue that the formulation accounts for the different time-scales of reproduction and behavior. They also argue that the formulation is simpler and can account for essential dynamics of food webs without added complexity. Detractors argue that using more mechanistic, albeit more complex, models that can account for realistic interactions is the correct way to proceed. Regardless of these arguments, using intermediate levels of complexity based on realistic mechanisms is the current trend in the food web theory.

Intermediate Levels of Complexity

Community ecology has focused on interactions (mainly competition and predation) between pairs of species that are fundamentally important in food webs. However, these interactions, taken out of context of the larger web, may result in misleading information (due to indirect effects). Highly complex food webs, however, are unwieldy and intrinsically difficult to study in model systems. Thus, Holt and others suggested investigating the dynamics of intermediate (between species pairs and whole food webs) levels of complexity – so-called "community modules" – that are defined as small subsets of species that are characterized by strong interactions. These modules are also more representative of the levels of complexity (i.e., number of species) examined in experimental studies.

Recent theoretical studies have taken advantage of intermediate complexity by focusing on food web interactions that typify interactions found in real food webs but are common to many food webs. This allows one to examine how indirect effects interact with direct effects to structure food webs. Common types of interactions among sets of species and their resources (modules) are apparent competition, intraguild predation, omnivory, cannibalism, and spatial subsidies.

This modular approach has allowed for various theoretical studies to examine stability of food webs using mathematical approaches. For example, McCann *et al.* (1998) found that adding relatively weak interactions among species could enhance food web stability. They found this to be true for apparent competition, intraguild predation, omnivory, cannibalism, and spatial subsidies.

Are weak interactions typical of food webs? The answer, from the few studies that have specifically examined this question, is yes. For example, in a study on intertidal food webs, Paine found that most interaction strengths are weak. Moreover, knowing that most predators eat 10 to $\sim$100 species of prey suggests that most of these interactions are weak.

One may then ask, what about strong interactions? The answer goes to the heart of one major problem with earlier food web studies. Strong interactions may occur and be a regular component of food webs. However, in almost every case, they appear to be enabled by "multichannel omnivory" (i.e., feeding on many weak links; Polis and Strong, 1996) or are restricted temporally and spatially because they are inherently unstable. However, time and effort constraints and tradition have caused the vast majority of food web studies to ignore the many weak interactions and the spatial and temporal aspects that characterize all systems.

In another theoretical study of food web processes that took advantage of the modular view, Huxel and McCann examined the flow of the allochthonous energetic resources. They found that allochthonous resources may spread evenly throughout the community or may become compartmentalized. High levels of allochthonous resources decreased stability, whereas low levels increased stability. Thus, again a weak link tended to increase stability.

See also: Diversity, Community/Regional Level. Ecosystem, Concept of. Energy Flow and Ecosystems. Keystone Species. Predators, Ecological Role of. Species Interactions. Trophic Levels

References

Arim M and Marquet PA (2004) Intraguild predation: A widespread interaction related to species biology. *Ecological Letters* 7: 557.

Forbes S (1887) The lake as a microcosm. *Bulletin of the Illinois State Natural History Survey* 15: 537.

Hairston Sr. NG, Smith F, and Slobodkin L (1960) Community structure, population control, and competition. *American Naturalist* 94: 421.

May R (1973) *Stability and Complexity in Model Ecosystems*. Princeton, NJ: Princeton University Press.

McCann K, Hastings AM, and Huxel GR (1998) Weak trophic interactions and the balance of nature. *Nature* 395: 794.

Oksanen L, Fretwell S, Arruda J, and Niemela P (1981) Exploitation ecosystems in gradients of primary production. *American Naturalist* 118: 240.

Paine RT (1980) Food webs: Linkage, interaction strength, and community infrastructure. *Journal of American Ecology* 49: 667.

Polis GA (1991) Complex trophic interactions in deserts: An empirical critique of food web theory. *American Naturalist* 138: 123.

Polis GA and Strong DR (1996) Food web complexity and community dynamics. *American Naturalist* 147: 813.

Yodzis P (1989) *Introduction to Theoretical Ecology*. New York: Harper & Row.

Forest Canopies, Animal Diversity

Terry L Erwin, Smithsonian Institution, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Arboricolous Living on the trees, or at least off the ground in shrubs and/or on tree trunks.
Emergent A very tall tree that emerges above the general level of the forest canopy.
Epiphytic material Live and dead canopy vascular and nonvascular plants, associated detritus, microbes, invertebrates, fungi, and crown humus.
Hectare Metric equivalent of 2.47 acres.
Microhabitat A small self-contained environmental unit occupied by a specific subset of interacting species of the forest (or any other community).
Scansorial Using both the forest floor and canopy for movement and seeking resources.
Terra firme forest Continuous hardwood forest of the nonflooded or upland parts of the Amazon rainforest.

Introduction

The forest canopy is arguably the most species-rich environment on the planet and hence was termed the "last biotic frontier," mainly because until very recently it had been studied less than any other place, with the exception of the deep ocean floor and outer space (Erwin, 1983). The reason for lack of study of the canopy was accessibility; and the evidence of the incredible species-richness, mainly of tropical forests, is primarily the abundance of insects and their allies. This hyperdiverse and globally dominant group has adapted to every conceivable niche in the fine-grained physical and chemical architecture of tree crowns. In the past three decades of the twentieth century, canopy biology became a mixed scientific discipline in its own right that is rapidly gaining sophistication of both approach and access. Tropical arboricolous (tree-living) arthropods were observed in the early 1800s in the "great forests near the equator in South America" and later that century were described by Henry Walter Bates. Even though Bates observed, described, and commented on the canopy fauna (as viewed from the ground and in recently felled trees), more than a century passed before Collyer designed an insecticide application technique that allowed a rigorous sampling regime for canopy arthropods. William Beebe and collaborators early in the twentieth century recognized that the canopy held biological treasures, but "gravitation and tree-trunks swarming with terrible ants" kept them at bay. Frank Chapman, a canopy pioneer (of sorts), viewed the treetops from his "tropical air castle" in Panama in the 1920s, but his interest was vertebrate-oriented, his perch was a small tower, and his observations of insects and their relatives were casual. By the mid-1960s and early 1970s, a few workers in both basic and applied sciences were seriously investigating canopy faunas of temperate and tropical forests in both the Western and Eastern Hemispheres. From the early 1980s until now, many workers have been improving methods of access and other techniques used to register, sample, and study the fauna (see reviews by Basset, Erwin, Malcolm, Moffett, and Lowman, Munn and Loiselle, and Winchester in Lowman and Nadkarni, 1995; Moffett, 1993; Mitchell, 1987). Some of these workers have found that arthropods by far make up the fauna of the canopy (Erwin, 1982, 1988, 2004; Erwin *et al.*, 2004). Visiting and nesting birds, mammals, reptiles, and amphibians represent a mere 1% or less of the species and even lesser in the abundance of individuals in these groups (Robinson, 1986). There are no adequate measures of canopy nematodes, mollusks, or other nonarthropod microfauna groups.

What is meant by the forest canopy? Generally, the canopy, or tree crown, is thought of as that part of the tree including and above its first major lateral branches. The canopy of a single tree includes the crown rim (the leaves and small twigs that face main insolation from the sun) and the crown interior (the main trunk and branches that gives a tree its characteristic shape). The canopy fauna is that component of animal life that inhabits the tree canopy and uses resources found there, such as food, nesting sites, transit routes, or hiding places. Hence, the forest canopy is collectively all the crowns of all the trees in an area. The canopy is often thought of as being stratified into emergent individuals, one to three regular canopy strata, and an understory of smaller trees living in the shade of a more or less continuous over story. All types of forests have their own describable characteristics, from the spruce forests of Alaska and northern Canada to the pine forest of Honduras, the dry and rain forests from México to Bolivia, and the *Rinorea* and *Mauritia* forests of the upper Manu River in Perú. It is through "whose eyes" one views the community, habitat, or microhabitat that determines the scale of investigation and subsequent contribution to the understanding of the environment – the beetles, the rats, the birds, the ocelots, the investigators, or perhaps even the trees.

Canopy Architecture, Animal Substrate

A temperate forest is composed of both broad-leaved and coniferous trees, with one or the other sometimes occurring in near-pure stands depending on the latitude and/or altitude and also on soil and drainage conditions. Normally, there are few canopy vines or epiphytes and perhaps some wild grape or poison ivy vines. Soil and organic debris caches are few or absent in the tree crowns, except for tree holes which provide homes to numerous arthropod groups but few vertebrates. Temperate forests are subjected to cold and hot seasonal climate regimes, as well as wet and dry periods. Great expanses of

the forest lose their leaves in the winter months, sap ceases its flow, and the forest "metabolism" comes to a slow resting state.

The temperate forest seemingly provides a great variety of substrates for the canopy fauna, but faunas are depauperate compared with those in tropical forests. Virtually no mammals are restricted to temperate forest canopies – only a few frogs and lizards. However, many bird species are restricted to the canopies, as they are in tropical forests. Among insects, for example, the beetle family Carabidae has 9% of its species living arboricolously in Maryland, 49% in Panama, and 60% or more at the equator in South America.

Tropical forests, in contrast, have few if any coniferous trees; only forests at higher elevations and/or located closer to subtropical zones have coniferous trees. Tropical canopies are often (but not always) replete with vines and epiphytes, tree holes, and tank bromeliads, and there are soil mats among the roots of orchids, bromeliads, and aroid plants. In the early 1990s, Nadkarni and Longino demonstrated that epiphytic material is fraught with macroinvertebrates, and Coxson and Nadkarni later showed that epiphytic material is important in the acquisition, storage, and release of nutrients.

Lowland tropical forests are subjected to mild temperatures, without frost, but have both wet (sometimes severe) and dry seasons. Individual species of trees may be deciduous, but in general tropical forests are always green and there is a perpetual growing season. Substrates are constantly available for the fauna. Often, some microhabitats with their substrates are temporary in the sense that they remain in place for a season or two, but then their architectural structure collapses into a jumbled pile of organic detritus on the forest floor. Such microhabitats (e.g., a suspended fallen branch with its withering leaves) provide a home resource to thousands of arthropods in hundreds of species, many found only in this setting. Eventually, such a branch loses its dried leaves and crashes to the forest floor. However, a short distance away, another branch breaks from a standing tree and the process begins again. The arthropods of the old, disintegrating branch move to the new one. The microhabitat and its substrates are forever present across the forest; each individual branch being ephemeral. The faunal members occupying such microhabitats are good at short-range dispersal.

Exploring the Past Biotic Frontier

Until recently, it was impossible to study the forest canopy well. Getting there was the limiting factor, and even after getting there (e.g., *via* ropes) it was difficult to find the target organisms. Modern devices such as aerial walkways (e.g., ACEER, Tiputini Biodiversity Station; **Figure 1**), one- or two-person gondolas maneuvered along crane booms (e.g., in Panama at Smithsonian Tropical Research Institute (STRI)), and web-roping techniques (see review by Moffett and Lowman in Lowman and Nadkarni, 1995) now allow real-time observations, sampling, and experiments anywhere in the canopy. Inflatable rafts that suspend mesh platforms resting on the upper crown rims of several trees have provided access from above, although this technique seems more suited to botanical work or leaf-mining insects, especially epiphytes and



Figure 1 The rainforest canopy of the western Amazon Basin from the canopy walkway of the Los Amigos Biological Station.

lianas. Insecticidal fogging techniques allow passive sampling of all arthropods resting on the surfaces of canopy plants (Erwin, 1995), and suspended window/malaise traps collect the active aerial fauna. Many of these techniques have been used. However, often they were simply used as collecting devices to garner specimens for museums and/or for taxonomic studies, and for this purpose they are excellent. In some cases, ecological studies were desired, but the techniques were not properly applied and the results disappointing. It is important to first ask the questions and then design the experiments; in some cases, current canopy techniques can be powerful tools for answering questions. Unfortunately, although sampling is relatively easy, sample processing is time consuming and laborious. For canopy-fogging studies, after the sampling effort, an average of 5 years was required before published products were achieved (Erwin, 1995). The main reason for this is a lack of funding for processing the results of fieldwork, even though the field studies were readily funded. Without processing, the data inherent for each specimen are unavailable for taxonomy or ecology studies. This is an historical funding problem and one of the reasons most studies examine only a few species from few samples.

Results of Studies

Invertebrates

Recent findings by Adis in the central Amazon Basin and by Erwin in the western part of the basin demonstrated that there

are as many as 3.4×10^{10} terrestrial arthropods per hectare. A recent 3-year study of virgin terra firme forest near Yasuni National Park in Ecuador by Erwin found an estimated 100,000+ species per hectare in the canopy alone. This figure was determined by counting the actual species in the samples of several well-known groups and comparing their proportions in the samples with their known described taxonomic diversity. The predatory beetle genus, *Agra* (**Figure 2**), has more than 2000 species found only in Neotropical forest canopies and scattered remnants of subtropical forest canopies in southern Texas and northern Argentina. The herbivorous weevil genus, *Apion*, likely has more than 10,000 species. In only 109 m^2 samples of canopy column from a 1-ha plot of virgin terra firme forest near Yasuni National Park in Ecuador, more than 700 species of the homopteran family Membracidae were found along with 462 species of the beetle family Carabidae, 368 species of the beetle family Mordellidae, and 178 species of the spider family Theridiidae.

"Biodiversity" by any other name is "terrestrial arthropods" – that is, insects, spiders, mites, centipedes, millipedes, and their lesser-known allies. Forest canopy studies of terrestrial arthropods are few (Erwin, 1995). Many of these studies currently concentrate on host specificity as a herbivore or parasite that eats only one other species of plant or animal. However, there is another class of specificity that is very important in understanding biodiversity that has received almost no study: "where" species hide and rest. This is not random, but rather species-specific (Lucky *et al.*, 2002).

Terrestrial arthropods are found in "hotels" and "restaurants" or "in transit" between the two (**Figure 3**). Often, insects and their allies eat, mate, and oviposit in the restaurant, that is, the food source, for example, on fungi or in suspended dry palm fronds. These insects may hide during the day under debris or under barks near the fungus or on the palm debris, but they never roam far from the vicinity of the food source, except to locate new food sources when the old one has been depleted. Members of other species eat in one place and then move to cover for a resting period, that is, the hotel. An example of this is the subfamily Alleculinae of the beetle family Tenebrionidae. These beetles feed on lichens and moss on tree trunks at night and spend the day (hiding, resting, and possibly sleeping) in suspended dry leaves elsewhere in the forest. Many species found in the forest canopy during the day (utilizing leaves, fruits, and/or flowers) hide and rest at night in the understory (e.g., various pollen-feeding beetles and the larger butterflies).

Insects particularly, and some of their allies, have adapted to nearly every physical feature of the planet, and the canopy is no exception. Many beetles have special feet for walking on leaves; some even have modified setae on their feet to slow them down on landing from rapid flight (**Figure 4**). Because they are in an environment with raptorious birds, lizards, and frogs, many insect species have evolved camouflage coloration. Climate is the main constraint on terrestrial invertebrates. In the temperate zones, it is the winter cold and dryness; in the equatorial tropics, it is the dry season for some and the rainy season for others, with the temperature far less of an influence than it is in the far north or south. Many herbivores must contend with plants that produce toxic chemicals or other defensive systems. All insects must also deal with other insects that predate, parasitize, or carry bacteria, fungi, or other insect diseases. Hammond, Stork, and colleagues, in their studies of insects in the Sulawesi dipterocarp forests, and Miller, Basset, and colleagues in New Guinea found much less insect diversity and richness than Erwin and his teams in the Neotropical forests. Hammond also found in southwest Asia that the canopy fauna was not as delimited from the understory fauna as it is in the Amazon Basin. Unfortunately, all these teams used different methodology; hence, much of their results are not comparable. It is certain, however, that the Old World tropical forests are not as biodiverse as those in the New World, nor are the forests of Costa Rica and Panama as diverse as those of the Amazon Basin. Disparate regional richness is one of the main problems in estimating the number of species on the planet. Another is the incredible richness of terrestrial arthropod species and the fact that scientists likely know <3–5% of them if published estimates of 30–50 million extant species are close to reality. Stork (1988) has even gone so far as to suggest that there could be 80 million species on the planet. New data from the Yasuni region of eastern Ecuador (Erwin *et al.*, 2004) suggest that even 80 million is a conservative estimate.



Figure 2 *Agra grace* Erwin, a species of Peru and Ecuador, South America.

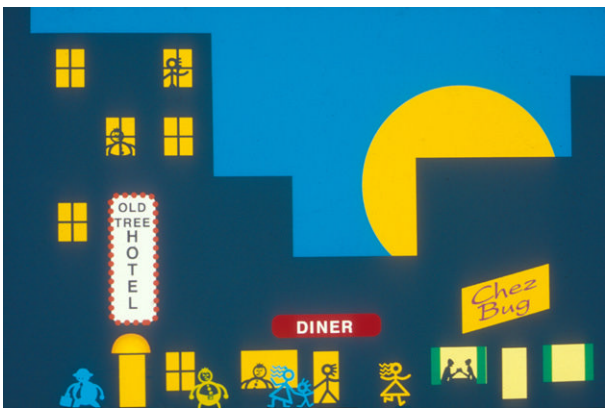


Figure 3 Humorous depiction of where "bugs" live and eat.

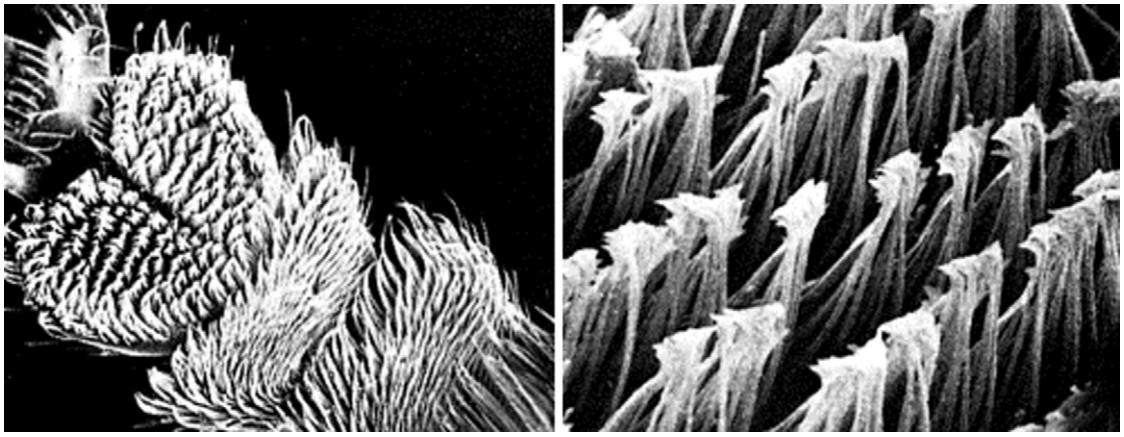


Figure 4 Setae of an arboreal beetle's tarsi used for landing and stopping quickly.

Vertebrates

Availability of food year-round constrains vertebrates from living strictly in canopies (see reviews by Emmons and Malcolm in Lowman and Nadkarni, 1995). Only in evergreen rainforests is there a continuous supply of food (albeit somewhat dispersed and sporadic) for phytophagous and insectivorous vertebrates. In deciduous forests, most species also forage on the ground or hibernate when food supplies are short. Almost all canopy mammals live in evergreen tropical forests, but even there most are scansorial. Timing and distribution of food resources are the critical controlling factors. Among all nonflying vertebrates, anurans and lizards and to a lesser extent snakes are the most important truly canopy creatures. Birds and bats are also exceedingly important components. All these groups except snakes account for vertebrate predator-driven evolution of the far more dominant invertebrates of the canopy. For example (as Blake, Karr, Robinson, Servat, Terbourg, and colleagues have shown), throughout the tropics ~50% of birds are strictly insectivores, whereas another 8% take insects and nectar. Morphological adaptations that allow canopy life include feet that can firmly grip the finely architectured substrate of twigs, leaves, and scaly bark. Emmons, in her many articles on Neotropical mammals, demonstrated that among these animals, those with the ability to "jump" avoided wasting energy and time by descending and climbing new trees to find resources; hence, more true canopy species have this ability. This is certainly true also of frogs and lizards. However, it is the flying forms – birds and, to a lesser extent, bats – that account for most of the treetop vertebrate fauna. Physiological adaptations that allow vertebrate canopy life include the ability to subsist on diets of fruit, flowers, leaves, or insects and their allies. Among mammals, fruit eaters are dominant.

As shown by Duellman, Dial, and colleagues, among canopy anurans and lizards, nearly all are primarily insect predators. Birds are overwhelmingly insectivorous in the canopy fauna, with ~40% in the upper Amazonian and 48% at Costa Rica's La Selva Biological Station. Malcolm, in summarizing the few articles on the subject, estimates that 15% of mammal species are arboreal-scansorial in temperate woodlands, whereas between 45% and 61% exhibit this behavior in tropical forests. In Duellman's 1990 list of anurans and reptiles from Neotropical forest, 36% are strictly arboricolous, whereas 8% are scansorial. Among birds, Blake and colleagues found that scansorial species using the understory and ground were more numerous than strictly canopy species (51% and 42%, respectively), at their site in Costa Rica.

In summary, although canopy vertebrates are important in driving part of invertebrate evolution in the forest canopy, they have not overwhelmingly radiated into or made use of the canopy, as the invertebrates have. For example, the total vertebrate fauna known at Cocha Cashu, Perú, is ~800 species (~45% of which are arboricolous or scansorial), whereas at a nearby location there are nearly 900 species of the beetle family Carabidae, of which more than 50% are strictly arboricolous. In Ecuador, near Yasuni National Park, there are in excess of 600 species of the homopteran family Membracidae in a single hectare, 100% of which are strictly arboricolous.

Conclusions

Although animals may use the air for dispersal, they live on substrate. Here, they eat, mate, hide, and transit. Forest canopies are rich in species because they offer a three-dimensional array of varying substrates that directly receive the sun's energy with little filtering. Although much has been and is being accomplished by faunal studies of the forest canopy, there is still much to do. There are missing data links between vertebrates and invertebrates and between both of these and the plant food and plant architecture on which they depend, and data are also missing on the influence of the canopy's physical features on the fauna such as microclimates (see Parker's review in Lowman and Nadkarni, 1995). Each subsystem is receiving at least some attention, but the new discipline of canopy biology is in its infancy. Is it too late? The forests and their species-rich canopies are rapidly disappearing (World Resources Institute, 1993).

Topics of current investigation include canopy insect β-diversity and measures of host specificity, the latter particularly in leaf-feeding beetles. Both areas of study were

driven by earlier, somewhat naive estimates of millions of species extant on the planet (Erwin, 1982; Stork, 1988; May, 1990; Casson and Hodkinson, 1991; Gaston, 1991). Although some of these studies may have been internally consistent within the parameters set for the estimations, no one had really got a handle on the true meaning of "host" specificity, biocomplexity of tropical forests, the influence of tropical biotope mosaics, β-diversity or what is known as species turnover in space and/or time, or the disparities of richness among continents, or even the disparity among regions within continents.

Even so, our current rudimentary knowledge indicates that we are losing hundreds, even thousands, of invertebrate species with "scorched Earth" programs such as that in Rondonia, Brazil, clear-cutting of Borneo and other southern Asian forests, and other losses in Haiti, Puerto Rico, Hawaii, the western Amazon Basin, Madagascar, and so on.

Conservation strategies are currently dominated by data on vertebrates (Kremen *et al.*, 1993; Samways, 1994); however, invertebrates are rapidly becoming sufficiently known to include them in analyses that are directed toward preservation of forest communities; to this end, the collective human conscience will soon be dealing with real extinction processes equivalent to those in the past, from the Permian to the Cretaceous. We are living at the beginning of the so-called sixth extinction crisis *sensu* Niles Eldridge of the American Museum of Natural History. Amelioration of the impact of this crisis rests on a better knowledge of the natural world around us and the development of conservation strategies that consider what we, Earth's managers (whether we like it or not), want future evolution to look like, as so well described by Quammen (1998).

See also: Amazon Ecosystems. Arthropods (Terrestrial), Amazonian. Beetles. Forest Canopies, Plant Diversity. Forest Ecology. Invertebrates, Terrestrial, Overview. Tropical Forest Ecosystems

References

Casson DS and Hodkinson ID (1991) The Hemiptera (Insecta) communities of tropical rain forests in Sulawesi. *Zoological Journal of the Linnean Society* 102: 253–275.

Erwin TL (1982) Tropical forests: Their richness in Coleoptera and other Arthropod species. *Coleopterists Bulletin* 36: 74–75.

Erwin TL (1983) Tropical forest canopies, the last biotic frontier. *Bulletin of the Entomological Society of America* 29(1): 14–19.

Erwin TL (1988) The tropical forest canopy: The heart of biotic diversity. In: Wilson EO (ed.) *Biodiversity*, pp. 123–129. Washington, DC: National Academy Press.

Erwin TL (1995) Measuring arthropod biodiversity in the tropical forest canopy. In: Lowman MD and Nadkarni NM (eds.) *Forest Canopies*, pp. 109–127. San Diego: Academic Press.

Erwin TL (2004) The biodiversity question: How many species of terrestrial arthropods are there? In: Lozwman MD and Rrinker HB (eds.) *Forest Canopies*, 2nd edn., pp. 259–269. San Diego: Academic Press.

Erwin TL, Pimienta MC, Murillo OE, and Aschero V (2004) Mapping patterns of β-diversity for beetles across the western Amazon Basin: A preliminary case for improving conservation strategies. *Proceedings of the California Academy of Sciences Series 4* 56(supplement I) (7): 72–85.

Gaston KJ (1991) The magnitude of global insect species richness. *Conservation Biology* 5: 283–296.

Kremen C, Colwell RK, Erwin TL, Murphy DD, Noss RF, and Sanjayan M (1993) Terrestrial arthropod assemblages: Their use in conservation planning. *Conservation Biology* 7(4): 796–808.

Lowman MD and Nadkarni NM (eds.) (1995) *Forest Canopies*. San Diego: Academic Press.

Lucky A, Erwin TL, and Witman JD (2002) Temporal and spatial diversity and distribution of arboreal Carabidae (Coleoptera) in a Western Amazonian Rain Forest. *Biotropica* 34(3): 376–386.

May RM (1990) How many species? *Philosophical Transactions of the Royal Society of London Series B* 330: 293–304.

Mitchell A (1987) *The Enchanted Canopy: Secrets from the Rainforest Roof*. London: Fontana/Collins.

Moffett M (1993) The High Frontier-Exploring the Tropical Rainforest Canopy. Cambridge, MA: Harvard University Press.

Quammen D (1998) Planet of weeds, tallying the losses of Earth's animals and plants. *Harpers Magazine*, October: 57–69.

Robinson MH (1986) The fate of the tropics and the fate of man. *Zoogoer* 5: 4–10.

Samways MJ (1994) *Insect Conservation Biology*. London: Chapman & Hall.

Stork N (1988) Insect diversity: Fact, fiction, or speculation. *Biological Journal of the Linnean Society* 35: 321–327.

World Resources Institute (1993) *World Resources 1992–1993*. New York: World Resources Institute/Oxford University Press.

Forest Canopies, Plant Diversity

Nalini M Nadkarni and Mark C Merwin, The Evergreen State College, Olympia, WA, USA
Jurgen Nieder, Botanische Institüt der Universität Bonn, Bonn, Germany

© 2013 Elsevier Inc. All rights reserved.

Glossary

Accidental epiphyte Plant that normally grows terrestrially but that occasionally grows to maturity in a tree crown, usually in terrestrial-like microsites such as the crotches of branches.
Bryophyte Nonvascular plant of the division Bryophyta (a moss, liverwort, or hornwort).
Cryptogam Plant that reproduces by spores or gametes rather than seeds; includes bryophytes and lichens.
Epiphyll (folicolous) Plant that grows on the leaf surface of another plant.
Epiphyte Nonparasitic plant that uses another plant as mechanical support but does not derive nutrients or water from its host.
Facultative epiphyte Plant or lichen that commonly grows epiphytically and terrestrially, usually exhibiting preference for one or the other habit in a particular habitat.
Homoiohydry Ability to maintain a constant internal water balance independent of fluctuating environmental conditions.
Lichen Composite organism consisting of a fungus (the mycobiont) and an alga or a cyanobacteria (the phycobiont) that live in a symbiotic relationship.
Mistletoe Woody parasite that taps the xylem of a tree, but is capable of photosynthesis.
Obligate epiphyte Plant that always grows on another plant for structural support, but derives no nutrients from the host.
Parasite Woody or nonwoody plant that taps into the vascular system of a host plant and derives energy or nutrients from it, often to the detriment of the host.
Poikilohydry Condition of internal water balance varying with changes in ambient humidity.
Primary hemiepiphyte Plant that begins its life cycle anchored in a tree crown and ultimately becomes rooted in the ground (e.g., strangler fig).
Secondary hemiepiphyte Plant that begins its life cycle as a terrestrial seedling, ascends a tree, and can later lose root connections with the ground, including (a) lianas, woody climbing plants with relatively thick stems that generally grow in mature habitats, and (b) vines, herbaceous climbing plants that regularly grow in disturbed habitats or forest edges.

Introduction

Definition of the Forest Canopy

The forest canopy has been called "the last biotic frontier" (Erwin, 1998). It presents a habitat conducive to the evolution of literally thousands – perhaps millions – of species of plants, microorganisms, insects, birds, and mammals that are rarely or never encountered on the forest floor. Although forest canopies have been among the most poorly understood regions of our planet, their mysteries are being explored by increasing numbers of biologists. Canopy communities are now believed to be important in maintaining the diversity, resiliency, and functioning of the forests they inhabit.

The forest canopy is a structurally complex and ecologically important subsystem of the forest. It is defined as "the aggregate of all crowns in a stand of vegetation, which is the combination of all foliage, twigs, fine branches, epiphytes as well as the interstices (air) in a forest" (Parker, 1995). The forest canopy is the primary site of gas exchange between the atmosphere and vegetation and fosters many ecosystem processes that are crucial to the maintenance and diversity of the forest as a whole (Lowman and Nadkarni, 1995).

Scope of this Article

Trees are the most obvious structural component of forest canopies. Their trunks, branches, and leaves constitute the infrastructure of the canopy and provide mechanical support for thousands of species of arboreal plants and animals. Tree species diversity is discussed elsewhere.

Although much has been published on canopy plants, the question of the global importance of epiphytes for the biodiversity of tropical and temperate forests in general and the canopy in particular has only recently been explicitly addressed. Previous reviews of canopy biodiversity have primarily dealt with arthropods (Erwin, 1998; Nieder *et al.*, 2001; Ozanne *et al.*, 2003; Ambrose, 2004; Lindo and Winchester, 2007), or with specific subgroups of canopy plants (Kress, 1986; Rhoades, 1995). Here, the authors describe the diversity of many types of canopy-dwelling plants. Included in this article are epiphytic vascular plants, epiphytic cryptogams (nonvascular plants that include lichens and bryophytes), primary and secondary hemiepiphytic vascular plants (lianas and vines), and arboreal parasitic vascular plants. Arboreal fungi and free-living algae are so poorly known that there is little to review (Stone *et al.*, 1996). The authors place greatest emphasis on obligate and facultative epiphytes, and exclude accidental epiphytes from this article.

The term "diversity" in the following will be based on species as the unit of biological diversity, since assessment of other aspects of biodiversity is virtually nonexistent in the case of canopy plants. The authors first review the systematic distribution of canopy taxa and provide species counts based on the state of current knowledge. The authors discuss gradients

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00158-1

of canopy plant diversity of microsites within the canopy at various spatial scales, spanning a single microsite within a tree (e.g., twig, branch bifurcation) to regional and global levels. Biogeographical analyses of canopy-dwelling taxa are then considered, as well as some of the major evolutionary elements that have influenced their distribution and abundance. "Habit diversity" (the diversity of morphological and physiological features) of arboreal taxa will then be described. Finally, the authors discuss conservation efforts that involve canopy plants and suggest future research possibilities.

Historical Roots and Sources of Information

In 1832, Charles Darwin first described what he termed the great diversity and profusion of "parasitical plants" (that the authors now understand to have been epiphytes), which he encountered in abundance in the coastal forests of Brazil. In the late nineteenth century, the German botanist A. F. W. Schimper first described epiphytes and outlined their importance to tropical botany.

Historically, canopy studies have been dominated by people who sought the thrill of climbing and followed the lure of discovering new species. Early European explorers hired climbers and trained monkeys to collect specimens of "exotic" air-plants that grew out of reach. Pioneering work in old-growth forests of the Pacific Northwest contributed to the application of mountain-climbing techniques for safe and reliable access to the canopies of tall trees. Since 1980, the innovation of high-strength and low-cost canopy access equipment has made canopy study more viable as an option for scientific research. There are now a wide variety of access tools from which to choose, depending on the questions being addressed and the available budget (Lowman and Nadkarni, 1995). With the development of effective technological climbing methods such as the "canopy raft" and the canopy crane, and of ground-based methods such as insecticidal fogging, researchers now spend less time working on how to prudently work in the treetops and more time pondering the difficulties in recording meaningful canopy data, analyzing it, and interpreting the results.

A remarkable burgeoning of scientific interest in the canopy has occurred within the last decade. This is related to increasing concerns with such conservation issues as biodiversity, global atmospheric change, and management of tropical and temperate rain forests. The number of scientific publications on canopy structure has grown at a disproportionately rapid pace relative to the general field of biology (Nadkarni, 1994). Aspects of the canopy have been the focus of many recent symposia, scientific books, and popular articles and media.

Reviews of vascular epiphyte, hemiepiphyte, and parasite diversity have been compiled (Madison, 1977; Calder and Bernhardt, 1983; Kress, 1986; Benzing, 1990; Putz and Mooney, 1991; Williams-Linera and Lawton, 1995; Lowman and Nadkarni, 1995; Benavides *et al.*, 2005). Studies about the biodiversity of nonvascular plants and lichens have increased rapidly in the last decade (Diaz *et al.*, 2010; Rhoades, 1995; Pentecost, 1998; Keller, 2004; Snell, 2001). The biodiversity of canopy nonlichen fungi has not been well documented (D. Reynolds, personal communication).

For this article, the authors compiled the foregoing sources and searched the primary literature for additions and modifications. The authors also consulted on-line databases and communicated with numerous taxonomists and specialists to ensure that the information presented is current and accurate. To place canopy plant diversity into the context of biodiversity in a given study area, the authors compared inventories of epiphytes and nonepiphytes. The epiphyte quotient (sometimes called "epiphytic index") is defined as the percentage of epiphytes out of the total number of vascular plants in an area.

Categories of Canopy Plants

Vascular Epiphytes

Forest canopies support extensive flora that include more than 24,000 species, or about 10% of all of the tracheophytes (Kress, 1986). Vascular epiphytes differ greatly in structure, function, and fidelity to their degree of dependence on canopy versus terrestrial habitats. Ecologists recognize their important roles in nutrient cycling and in providing arboreal and terrestrial animals with food, water, and nesting materials (Nadkarni, 1994). Ecophysiologists recognize the varied structures and mechanisms that protect vascular epiphytes from drought (Benzing, 1990).

Some of the characteristics for regular occurrence on bark and associated aerial substrates are obvious (e.g., holdfast roots and wind-dispersed propagules), but others are more subtle. In an extensive review of vascular epiphytism, Benzing (1990) outlined a variety of characteristics that are exhibited by vascular epiphytes (**Table 1**).

Vascular epiphytes are mainly restricted to the low latitudes and within the tropics. They reach their greatest abundance and diversity at low to mid-montane elevations (Madison, 1977; Benzing, 1990). Ferns occur in higher latitudes along the margins of the Pacific, and a few hardy bromeliads and orchids occur in the mild north and south temperate zones (e.g., *Epidendrum rigidum, Polypodium polypodioides, Tillandsia usneoides, Fascicularia bicolor*). The most extensively colonized temperate forests are those of southeastern Australia, New Zealand, and Chile, where a variety of vascular epiphytes grow in areas protected from frost by nearby ocean currents.

Nonvascular Epiphytes

In a review by Rhoades (1995), nonvascular (or cryptogamic) epiphytes were categorized into three groups: lichens, bryophytes, and free-living algae. Although the phylogeny and composition of the two plant groups considered here are very different – lichens are symbiotic fungi and algae, and bryophytes are plants – they occupy similar habitats and are often studied together. They have been the focus of an increasing number of studies in the upper canopy (relative to vascular epiphytes), in northwestern North America (Rhoades, 1995; Peck and McCune, 1997; Sillett, 1995), the eastern deciduous forest of North America, the boreal forest of Canada, and the forests of the southern hemisphere (Diaz *et al.*, 2010; Wolf, 1994).

Bryophytes (phylum Bryophyta) are plants that lack true vascular tissues and organs. In canopy habitats they include

Table 1 Characteristics of vascular epiphytes

1. Reproduction	
A. Pollination	Exclusively zoophilous, flowers tend to be showy, pollinators highly mobile (Benzing, 1990, Chapter 5)
B. Breeding systems	Little studied, although many orchids appear to be allogamous (Benzing, 1990, Chapter 5)
C. Population structure	Little studied (ferns, Hooper and Haufler, 1997)
D. Seed dispersal	Most families endozoochorous, most species anemochorous (because of the dominance of Orchidaceae)
E. Life history	Almost all iteroparous, long-lived perennials
2. Vegetative	
A. Foliage	Usually evergreen, often succulent, and xeromorphic generally
B. Habit	Woody (wet forests) to herbaceous (wet and dry forests)
C. Shoot architecture	Various
D. Roots	Adventitious, specialized for holdfast, often reduced
E. Special features[a]	Impounding shoots (e.g., Bromeliaceae) and root masses (e.g., ferns), velamentous roots and absorptive foliar trichomes to prolong contact with precipitation and canopy washes, often lack capacity to grow in earth soil
3. Mineral nutrition	
A. Mycorrhizas	Possibly significant in Orchidaceae and Ericaceae, probably relatively unimportant elsewhere compared to terrestrial flora
B. Myrmecotrophy[a]	Nearly exclusive to epiphytes
C. Carnivory	Underrepresented in arboreal flora
D. Saprotrophy[a]	Phytotelm and trash-basket types
E. Special features[a]	Tolerance for low pH (wet forests), effective nutrient scavengers (dry forests), frequent reliance on organic substrates for nutrient ions
4. Photosynthesis/water balance	
A. Photosynthetic pathways	CAM overrepresented (tropical only) no typical C4 types, much interesting detail probably remains underdescribed
B. Water economy	Often very high
C. Moisture requirements	Various
D. Other[a]	Much flexibility, for example, facultative CAM, CAM-C4 intermediates

[a]These characteristics distinguish arboreal from terrestrial flora more than the others.

Source: Modified from Benzing DH (1990) *Vascular Epiphytes*. Cambridge, UK: Cambridge University Press.

the mosses (with about 10,000 species worldwide) and leafy liverworts (leafy hepatics, 7200 species). Thallose (strap- or fan-shaped) liverworts and hornworts are usually restricted to moist, lower trunks. (Rhoades, 1995).

Lichens are important components of canopy biodiversity and of ecosystem processes (e.g., nutrient cycling, providing food for wildlife). Lichens are not a single taxonomically distinct category, but rather are symbiotic organisms, the association of a fungus (the mycobiont) and a photosynthetic partner (the photobiont), which are usually members of the Chlorophyta and Cyanobacteria. The degree of photobiont specificity varies among lichens. Generally, the mycobiont gives a lichen its overall form and provides the bulk of the biomass, outer protective layer, and a looser, inner layer that functions in physical absorption and storage of water and nutrients. The photobiont is usually restricted to the layer just below the protective covering of the mycobiont.

These cryptogams are poikilohydric, that is, they depend on an atmospheric supply of water and inorganic nutrients from precipitation, dew, or fog interception. In general, they absorb water rapidly and lack the water-resistant coverings or cuticles of vascular plants. Bryophyte growth forms have been described and discussed to understand their relationship to water use and conservation (Schofield, 1992; During, 1979). The gametophytes (vegetative bodies) of some species form tight cushions or spherical balls that expose a reduced surface area to retain water; others have a pendant, creeping habit that exposes them to maximal amounts of bark surface water. The sporophytes of many bryophytes are adapted to the periodically xeric nature of epiphytic habitats and can distribute their spores over very wide ranges (Gradstein *et al.*, 1989).

Hemiepiphytes

Hemiepiphytes have been defined as plants that have, at some point in their lives, an "umbilical" connection to the ground. Whether roots or stems, these connections buffer hemiepiphytes from problems of water and nutrient supply that are faced by obligate epiphytes (Williams-Linera and Lawton, 1995). Hemiepiphytes begin their life cycle either as epiphytes and eventually send roots or shoots to the ground (primary hemiepiphytes) or as terrestrially established seedlings that secondarily become epiphytic by severing all connections with the ground (secondary hemiepiphytes) (Kress, 1986).

Hemiepiphytes exhibit a tremendous variety in growth form, impact on their hosts, and degree of dependence on hosts. They range from being erect and treelike in form to species that grow in scandent, clambering heaps. Their impacts on hosts range from lethal (e.g., strangler figs) to benign (e.g., shrubby Ericaceae in tropical cloud forests) (Williams-Linera and Lawton, 1995).

Parasites

The mistletoes, which are woody shrubby parasites, are an ecologically distinctive group of canopy-dwelling plants. They

have received a great deal of attention from botanists because of their ability to tap into the vascular system of their hosts, as well as from foresters, who have been concerned with reducing the damage they wreak through timber loss and mortality of desirable trees and shrubs, and also because they confer benefits, such as nesting habitats and food for birds. Parasitic mistletoes tend to show a greater tendency for host specificity than do the epiphytes (Calder and Bernhardt, 1983).

Canopy Plant Taxa Diversity

Vascular Epiphytes

Although the global species richness of plants is probably in the region of 270,000, neither their exact number nor their global diversity pattern is known. It has been estimated that possibly as many 24,000 vascular plant species are epiphytes (Kress, 1986), so they constitute a major part of the global biodiversity in the forest canopy.

Vascular epiphytes account for 10% of the total vascular plant diversity. Most extant epiphytes are angiosperms, representing about 9% of all angiosperm species (**Table 2**). Many vascular plant families (84) have adapted to life in the canopy, but relatively few taxa have radiated successfully. Within the angiosperms, approximately 31% of the monocots are epiphytic, whereas only 3% of the dicotyledons occupy the epiphytic niche. The Orchidaceae constitute approximately two-thirds of all epiphyte species (Kress, 1986). Other important monocotyledon families are Bromeliaceae and Araceae. The important canopy-dwelling dicotyledon families are Cactaceae, Ericaceae, Gesneriaceae, Melastomataceae, Moraceae, Piperaceae, and Rubiaceae (**Table 2**; Kress, 1986). There are some large taxonomic groups of plants that contain no epiphytes or very few epiphyte species, for example, the Asteraceae, Leguminaceae, and Poaceae (Benzing, 1987). Less than 1% of the gymnosperms are known to be epiphytic. The Pteridophytes (ferns) are another important group of epiphytic plants, of which 29% are epiphytes (Kress, 1986).

The epiphyte quotient (proportion of an entire flora that is epiphytic) varies widely both geographically and among forest types. This ratio has been measured directly in only a few study sites (**Table 3**). Calculated epiphyte quotients based on published floristic studies in the Neotropics are known for Panama (12%), Peru (10%), Ecuador (22%), Costa Rica (26%), Venezuela (50%), and Florida (3%). In the Paleotropics, epiphyte quotients have been calculated from sites in Java (12%), West Malaysia (9%), Sri Lanka (4%), and Japan (0.5%). African forests seem to be much poorer in relative epiphyte species richness. In Ghana, a typical epiphyte quotient in forest plots is 8%; one direct measurement in central Africa (Rwanda and Zaire) was 3%.

It is generally regarded that the New World supports greater vascular plant diversity than the Old World. The number of vascular families containing at least one epiphyte species is very similar in the Paleotropics (43) and Neotropics (42). Within the Paleotropics, the representative families do not exhibit a homogeneous distribution. All 43 of the families occur in Australasia, but only 15 are found in Africa and

Table 2 Taxonomic distribution of vascular epiphytes

Major group	*Taxonomic category*	*Number of taxa containing epiphytes in each category*	*Percentage of taxa containing epiphytes in each category*
All vascular plants	Classes	6	75
	Orders	44	45
	Families	84	19
	Genera	876	7
	Species	23,456	10
Ferns and allies	Classes	2	67
	Orders	5	50
	Families	13	34
	Genera	92	39
	Species	2593	29
Gymnosperms	Classes	2	67
	Orders	2	33
	Families	2	13
	Genera	2	3
	Species	5	<1
Angiosperms (dicots)	Subclasses	6	100
	Orders	28	44
	Families	52	16
	Genera	262	3
	Species	4251	3
Angiosperms (monocots)	Subclasses	4	80
	Orders	9	47
	Families	17	26
	Genera	520	21
	Species	16,608	31

Source: Modified from Benzing DH (1990) *Vascular Epiphytes*. Cambridge, UK: Cambridge University Press.

Table 3 Epiphyte quotients of South American forests that have been directly measured by J. Nieder and his colleagues

Study site	*Elevation (m)*	*Precipitation (mm yr^{-1})*	*Number of epiphyte species*	*Epiphyte quotient (area of reference)*
Sehuencas, Bolivia	2100–2300	5000	230 spp.	37% (0.1 ha)
Otonga, Ecuador	1700–2200	2500	196 spp.	–
Rio Guajalito, Ecuador	1800–2200	2700	166 spp.	28% (400 ha)
Carbonera, Venezuela	2100–2300	1500	192 spp.	45% (360 ha)
Surumoni, Venezuela	100	2800	53 spp.	6
			(crane plot only);	–
			112 spp.	
Matá, Colombian Amazonia	242	3060	213 spp.	–

Madagascar. Vascular epiphytes are most diverse in the Neotropics, and less so in tropical Asia and Africa. There is approximately a twofold increase in species diversity in the Neotropics compared with Australasia, and a sixfold increase compared with Africa (Madison, 1977; Gentry and Dodson, 1987). Vascular epiphytes exhibit their greatest diversity in the montane cloud forests of Latin America. The temperate regions support considerably fewer species than tropical areas. Likewise, the temperate regions generally support more vascular epiphytes than do boreal areas.

Nonvascular Epiphytes

In general, bryophytes account for 9–10% of the total species diversity of the plant kingdom. However, no one has calculated how many species of nonvascular plants are obligate epiphytes, as "the idea of an obligate epiphyte is a slippery concept" (D. Griffin, personal communication). Rhoades (1995) has ably summarized the results of regional floristic studies (**Table 4**).

The standard growth forms of lichens are arbitrary, but have often been used to describe functional groups in canopy habitats. "Foliose" refers to leaflike, "fruticose" refers to thalli without distinctive dorsoventral arrangements, and "crustose" refers to thalli firmly cemented to a substrate. According to the International Code of Botanical Nomenclature, lichen species are given the name of their mycobiont; photobiont names are subsidiary. Morphology (sexual structures, asexual structures, and vegetative surface characters) and thallus chemistry are important species characters. Only a few studies have focused on the worldwide biogeography of bryophytes (Schofield, 1992) or lichens (Rhoades, 1995). For many inventories, the crustose lichens have been lacking or incomplete, which is unfortunate as they are the dominant cryptogamic form in outer canopies. The proportion of lichens that grow arboreally is unknown.

The bryophytes are a very old group of plants, perhaps dating as far back in the fossil record as the Devonian period when the first land plants are known to have existed. The combination of a long history and small airborne dispores has allowed several bryophyte families to show wide geographic ranges. The tropical regions of Australia and Asia generally have more endemic genera of mosses, whereas in the Neotropics endemic liverwort genera are richer (Schofield, 1992). The cosmopolitan families are not restricted by latitude, but may show local altitudinal variation in some parts of their range. Representative moss families in the tropics include Bryaceae, Dicranaceae, Fissidentaceae, Funariaceae, and Hypnaceae. Pantropical moss families include Calymperaceae, Pteroryaceae, Racopilaceae, and Rhizogoniaceae. Important temperate families are Aulacomniaceae, Encalyptaceae, Grimmiaceae, and Polytrichaceae. Representative tropical liverwort families include Frullaniaceae, Lejeuneaceae, Lophocoleaceae, Plagiochilaceae, and Radulaceae. Species-rich liverwort families in temperate regions are the Marsupellaceae and Scapaniaceae. All the species of hornworts are in a single family, Anthocerotaceae, which is most diverse in tropical ecosystems (Schofield, 1992).

Hemiepiphytes

The phylogenetic distribution of hemiepiphytes suggests that this habit has evolved independently a number of times (Putz and Mooney, 1991; Williams-Linera and Lawton, 1995). Twenty-five families and 59 genera contain hemiepiphytes (**Table 5**), with more than 820 species of primary hemiepiphytes and 650 species of seconary hemiepiphytes. These make up 1% of the total vascular plant species diversity, and 1% of the total canopy-dwelling vascular plant species. This is probably an underestimate, especially for woody hemiepiphytes ("lianas"), which are the most undercollected major canopy plant group. The stranglers most commonly occur in Moraceae and Clusiaceae, but are also found in Araliaceae, Rubiaceae, and Myrtaceae. The hemiepiphytic habit may have arisen from plants growing on rocks.

All of the hemiepiphytic monocotyledonous plants are secondary hemiepiphytes in the families Araceae and Cyclanthaceae. Secondary hemiepiphytes also occur in the dicotyledonous family Marcgraviaceae. Primary hemiepiphytes are represented by 20 families of dicotyledons. The majority of primary hemiepiphyte species are found in the families Araceae, Clusiaceae, and Moraceae. The Moraceae contain the most species of hemiepiphytes, with approximately 500 species in the genus Ficus. Primary hemiepiphytes (whose aerial roots eventually reach the ground) represent about 0.8% of all epiphytes with almost 2000 species (Gentry and Dodson, 1987).

As with vascular epiphytes, the contribution of hemiepiphytes to the diversity of the tropical forest canopy varies among forests. The percentage of trees colonized by hemiepiphytes has been reported for study sites in Venezuela (10% and 13%), Zimbabwe (13%), French Guiana (17%), and the Ivory Coast (21%). In neotropical lowland forests, stranglers and large hemiepiphytes can occur on 10–15% of the trees. Stranglers can occur in much higher densities in some

Table 4 Species richness of epiphytic cryptogams in worldwide forest types[a]

Location/forest type	*Latitude (° N)*	*Number of trees sampled*	*Mosses*	*Liverworts*	*Total bryophytes*	*Macrolichens*
Guyana; dry evergreen Eperua spp.	5?	11	28	53	81	33
French Guyana; mixed lowland rain forest	5	4	43	61	104	21
Colombia; montane rain forest, 1500 m	5	4	22	36	58	49
Colombia; montane rain forest, 2550 m	5	4	33	102	135	51
Colombia; montane rain forest, 3510 m	5	4	19	63	82	37
Guyana; mixed lowland rain forest	7?	5	28	60	88	19
Oregon, United States; low, mixed coniferous forest	44	11	11	6	17	37
Wisconsin, United States; mixed conifers and hardwoods	46	Many	14	3	17	29
Montana, United States; old-growth Abies	48	5	4	1	5	34
Montana, United States; managed, second-growth Abies	48	5	1	0	1	37
Washington, United States; low-elevation fir forest on lava flow	49	Many	8	5	13	53
Sweden; deciduous forest	56	Many	78	17	95	–
Chiloéa Island, Chile; coastal temperate rainforest[b]	41°S	3	15[c]		15[d]	3[d]

[a]Modified from Rhoades FM (1995) Nonvascular epiphytes in forest canopies: Worldwide distribution, abundance, and ecological roles. In: Lowman MD and Nadkarni NM (eds.) *Forest Canopies*, pp. 353–408. San Diego, CA: Academic Press.

[b]From Diaz IA, Sieving KE, Pena-Foxon ME, Larrain J, and Armesto JJ (2010) Epiphyte diversity and biomass loads of canopy emergent trees in Chilean temperate rain forests: Aneglected functional component. *Forest Ecology And Management* 259: 1490–1501.

[c]Category includes liverwort.

[d]Epiphytes identify mostly to genus.

neotropical palm savannas. A rich hemiepiphytic flora is typical of mountain forest and cloud forest sites, but wet lowland forest can also show high percentages of these species (25% in the case of La Selva Biological Station in lowland Costa Rica). In dry forests, hemiepiphytes are usually not present (Williams-Linera and Lawton, 1995).

Woody lianas are distinct features of tropical forests, rarely or never occurring in temperate forests. Lianas account for approximately 10% of the tropical flora worldwide. They occur in greatest density in Madagascar and Africa, and less so in neotropical and Australian forests. At some sites in Madagascar, an average of 122 lianas with >2.5 cm dbh per 0.1 ha is reported, whereas lianas average only five individuals per 0.1 ha in northern temperate forests.

Lianas are more abundant in south temperate forests than in north temperate forests. The important families in north temperate forests are Anacardiaceae, Araliaceae, and Vitaceae. South temperate forests support more than twice as many families. Families such as the Bignoniaceae, Gesneriaceae, Gramineae, Lardizabalaceae, Saxifragaceae, and Vitaceae account for the majority of the climbing species (Putz and Mooney, 1991).

Parasites

About 1400 species of mistletoe occur in forests, woodlands, and shrublands on every continent except Antarctica, with most species in the tropics. Less than 1% of the total vascular plant species are mistletoes, and this group accounts for less than 1% of the total canopy-dwelling vascular plant species.

Mistletoes occur in two plant families. The Loranthaceae contain approximately 900 species in 65 genera, and the Viscaceae contain 400 species in 7 genera. The most species-rich genera in the Viscaceae are Phoradendron (170) and Viscum (100). The Viscaceae contain four genera restricted to the Old World, two genera that occur only in the New World, and one predominately New World genus is also widespread through Eurasia and Africa. In the Viscaceae, the New World genera Dendrophthora and Phoradendron contain about half of the 397 species of the family.

Diversity of Habitats of Canopy Plants

Canopy Microclimate

The values of light intensity and quality, temperature, wind, moisture content, and concentrations of various gases and aerosols are strongly modified by canopy structure in several ways. Canopy surfaces act as passive bodies for the absorption of wind energy, the dissipation of turbulence, and the sorption of heat and radiation. They also actively participate in exchanges of biologically important compounds, such as CO_2 and water vapor, which in turn may have an impact on regional, and even global, climate. Canopies also act as "filters" that remove small-scale turbulence, but allow large eddies to penetrate (Parker, 1995). Canopy structure therefore has a direct effect on the climate surrounding individual leaves, on the modification of microclimate through the layers of the forest, and on the large-scale environment of forest regions (Ambrose, 2004; Cardelus and Chazdon, 2005).

Table 5 Families and genera that contain hemiepiphytes[a]

Secondary hemiepiphytes
Monocotyledonae
1. Araceae
Amydrium Schott, 4/4 Malaysia
Anthurium Schott, 200/550 Neotropics
Caladiopsis Engl., 2/2 South America
Epipremnum Schott, 15/15 Indomalaya
Heteropsis Schott _6/21, Central America
Monstera Adans, 24/25 Neotropics
Pedicellarum Hotta, 1/1 Borneo
Philodendron Schott, 133/275 Neotropics
Porphyrospatha Engl., 3/3 Neotropics
Pothos L., 25/75 Indomalaya and Pacific
Rhaphidophora Hassk., 100/100 Indomalaya and Pacific
Rhodospata , 3/45 North and Central America
Syngonium Schott, 18/25 Neotropics
2. Cyclanthaceae
Asplundia Harling, 20/82 Neotropics
Carludovica Ruiz &. Pav., 1/3 Central America
Ludovia Brongn., 2/2 South America
Sphaeradenia Harling, 7/38 Neotropics
Thoracocarpus Harling, 1/1 South America
Dicotyledonae
3. Marcgraviaceae
Caracasia Szyszyl., 2/2 Venezuela
Marcgravia L., 50/55 Neotropics
Norantea Aubl., 20/35 Neotropics
Souroubea Aubl., 20/25 Neotropics
Ruyschia Jacq., 2/10 Neotropics
Primary hemiepiphytes
4. Araliaceae
Didymopanax Decne. &. Planch, Neotropics
Oreopanax Decne. & Planch., Neotropics
Pentapanax Seem., 2/15 Java to Formosa
Polyscias J.R. & G. Forst, 5/80 Malaya to New Zealand
Raukahua 2/9, Neotropics
Schefflera J.R. & G. Forst, 60/200 Pantropics
Sciadophyllum P. Br., 5/30 South America and West Indies
Tupidanthus Hook.f. & Thoms., 1/1 Indomalaya
5. Bignoniaceae
Schlegelia[b]
6. Burseraceae
Bursera, 1 Costa Rica
7. Celastraceae
Euonymus L., 2/175 Himalayas
8. Clusiaceae
Clusia L., 85/145 Africa, Madagascar, Neotropics
Clusiella Planch. & Triana, 3/7 South America
Havetiopsis Planch. & Triana, 3/7 South America
Odematopus Planch. & Triana, 1/10 South America
Quapoya Aubl., 1/3 South America
Renggeria Meisn., 1/3 Brazil
9. Cunoniaceae
Ackama A. Cunn., 1/3 New Zealand
Weinmannia L., 3/170 New Zealand and Neotropics
10. Dulongiaceae
Phyllonoma Willd. ex Schult., 1/8 Neotropics
11. Ericaceae
Cavendishia (2 spp.), Neotropics
Gonocalyx, Neotropics
Disterigma, South America
Sphyrospermum, South America
12. Euphorbiaceae

(*Continued*)

Table 5 Continued

Secondary hemiepiphytes
Schradera (2 supp.)
13. Gesneriaceae
Drymonia (2 sp.), Central America
14. Griseliniaceae
Griselinia Forst.f., 3/6 New Zealand and Chile
15. Melastomataceae
Blakea P. Br., 60/70 Neotropics
Topobea Aubl., 20/50 Neotropics
16. Moraceae
Coussapoa Aubl., 20/45 Neotropics
Ficus L., 500/800 Pantropics
17. Myrsinaceae
Grammadenia Benth., 6/15 Neotropics
18. Myrtaceae
Metrosideros Banks ex Gaertn., 3/60 New Zealand
19. Potaliaceae
Fagraea Thunb., 20/35 Malaysia-Pacific
20. Rubiaceae
Posoqueria Aubl., 1/15 Neotropics
Cosmibuena Ruiz & Pav., Neotropics
21. Rutaceae
Zanthoxylum, Central America
22. Saxifragaceae
Hydrangea, Neotropics
23. Solanaceae
Markea,[b] Neotropics
24. Violaceae
Melicitus
25. Winteraceae
Drimys

[a]Modified from Williams-Linera G and Lawton RO (1995) The ecology of hemiepiphytes in forest canopies. In: Lowman MD and Nadkarni NM (eds.) *Forest Canopies*, pp. 255–283. San Diego, CA: Academic Press.
[b]From Putz F and Mooney HA (1991) *The Biology of Vines*. Cambridge, UK: Cambridge University Press.

Canopy conditions are generally typified by more intense sunlight, greater extremes of relative humidity, higher water stress, and a smaller, more pulse-supplied pool of nutrients than on the forest floor. Sunlight attenuation can be as great as 98% between the tops of emergent trees and the levels reaching the forest floor. Rates of evaporation in the canopy have been recorded that are comparable to those occurring in open savannas. Relative humidity can range from nearly 100% at night to less than 30% during midday in the dry season. Differences in canopy versus forest floor wind speeds can also be extreme, especially in tropical cloud forests. In one Costa Rican ridge cloud forest, wind speeds within the canopy (10 m) were clocked at 11.3 m s^{-1}, while forest floor (2 m) speeds were only 4.0 m s^{-1} (Williams-Linera and Lawton, 1995).

Spatial Scales of Canopy Plant Diversity

The forest canopy is a three-dimensional subsystem of the forest itself. Canopy plants need relatively little space in order to develop a striking diversity. In an Ecuadorian montane forest, for example, 109 epiphyte species occurred on just 20 m^2 of branch surface, compared to only 67 terrestrial plant

species on a 100-m^2 ground plot of elongated shape in the immediate vicinity of the phorophyte. The amazing concentration of epiphytes on single trees has often attracted the attention of naturalists. For example, 66 epiphyte species were found on one specimen of Decussocarpus rospigliosii in the Carbonera Forest in Venezuela.

The canopy offers its occupants a wide variation in water, light, and nutrient regimes compared to the understory and the forest floor, and this variety undoubtedly contributes to arboreal plant diversity. For example, soil-like deposits and litter in the canopy function as a medium for canopy-dwelling plants that have well-developed root systems, such as vascular epiphytic shrubs. This material has a high organic content and is derived principally from leaf litter, feces, and other faunal remains. Also present is a small mineral component derived from fine particles carried there by wind, fog, and rain.

In tropical America, epiphytic bromeliads increase the volume of arboreal soil and litter by creating water-filled tanks ("phytotelmata") in which litter accumulates and soil forms. Although these arboreal epiphytes and their associated soils are patchily distributed, they are linked by climbing vegetation, by percolating rainwater, and probably by the movement of animals, and so they can be likened to a three-dimensional matrix of interconnected islands.

Within a forest, microsite differences exist at many different spatial scales: within a single branch; between branches at different heights of the tree; between trees of different architecture; and within stands of differing topography and aspects. Some studies have described variation in the distribution of canopy plants within single crowns. In a West African rain forest, for example, more than 75% of the orchid species grow on the inner branches, 48% in a middle zone, and only 4% in the outer canopy. Some research has linked differences in microsite water regimes and levels of sunlight input to differential distribution of certain canopy plants between outer and inner crowns; other arboreal species display a "generalize distribution."

Host Tree Specificity

Studies of the mechanisms influencing host tree specificity are scant. It has been suggested that bark texture and pH influence patterns of colonization. For example, the orchid *Cymbidiella pardalina* obligately grows exclusively on the fern *Platycerium madagascarense*, whereas other epiphytes are found on numerous host tree species. This phenomenon warrants further investigation.

Diversity of Growth Habits in Canopy Plants

Canopy-dwelling plants exhibit a great diversity of ecological adaptations, which is most likely a result of their diverse phylogenetic origins and the possibilities for adaptive specialization in the canopy habitat. The one unifying feature of these mechanically dependent plants is their occurrence in the canopy (Benzing, 1987). For vascular epiphytes, the different "habits" can be classified by several criteria, including degree of dependence (obligate vs. facultative), nutritional dependence (parasitic vs. commensalistic), degree of light demand (heliophiles vs. sciophytes), architecture (tank vs. atmospheric), substratum (e.g., ant-gardens, humiphiles), or carbon fixation pathway (Crassulacean acid metabolism (CAM) vs. C_3) (Benzing, 1990). Similarly, the great diversity of habits within the hemiepiphytic species ranges from stranglers, which ultimately become freestanding trees, to epiphytes that have only one root connecting with the ground.

Diversity of Physiology

The C_3 photosynthetic pathway is more typical of epiphytes inhabiting the canopies of cloud forests and cool, shaded, humid microsites. Forest canopies with more arid conditions favor CAM plants. The occurrence of a C_4 epiphyte has not been documented. Some families containing C_4 plants such as Asteraceae, Cyperaceae, Orchidaceae, and Poaceae also contain epiphytic species, but none is known to exhibit C_4 photosynthesis.

Diversity of Modes of Resource Acquisition and Retention

Nutrient acquisition in canopy plants occurs though many modes, including rainwater, bark and leaf leachate, nitrogen-fixing cyanobacteria, airborne particles, carton nests, crown humus, and decomposition of the host.

Humus Epiphytes

Ecologically, humiphily is the most common form of nutrient acquisition and supports the greatest diversity of epiphytes. An overwhelming majority of families containing at least one epiphytic species have at least one humiphile species, and most families contain only humiphile species (Benzing, 1987). Humus-rooted epiphytes include Ericaceae, Gesneriaceae, Melastomataceae, Piperaceae, and Rubiaceae, as well as the hemiepiphytes in Moraceae and Araliaceae. Many of the humus epiphytes are facultative. The obligate species exhibit a diversity of xeromorphic adaptations, including leaf succulence, a flattened pendulous growth form, and poorly developed root systems with strong mycorrhizal associations.

Tank and "Trash-Basket" Epiphytes

Some bromeliads, ferns, and orchids collect water, airborne particles, and leaf litter in the rosette created by overlapping fronds or leaves. Ferns in the genus *Asplenium*, for example, have roots that grow in the form of a trash basket to gain access to nutrients.

Ant-Associated Epiphytes

Myrmecophily is a common feature of vascular epiphytes and may be strictly a canopy phenomenon. Nest garden species include Araceae, Bromeliaceae, Cactaceae, Gesneriaceae, Marcgraviaceae, Orchidaceae, Piperaceae, and Rubiaceae. In Australia, Asclepiadaceae, Melastomataceae, and Rubiaceae occur in less-studied ant-garden symbioses. Asclepiadaceae, Bromeliaceae, Melastomataceae, Orchidaceae, Polypodiaceae, and Rubiaceae contain specialized ant-garden species. Rubiaceae appear to be the most specialized ant-garden species.

Some ferns and orchids have hollow rhizomes (Polypodiaceae), hollow tubers (*Solanopteris* spp.), and hollow pseudobulbs (*Schomburgkia* and *Laelia*) that provide domatia for ants.

Bark Epiphytes

To inhabit the bark substrate, epiphytic plants must cope with very low levels of water and nutrient availability. Therefore, many of the bark epiphytes are obligate epiphytes, including many specialized orchids.

Atmospheric Epiphytes

Some bromeliads (e.g., *Tillandsia*) have special hairs (trichomes) that allow them to absorb water from the atmosphere over the entire surface of their leaves.

Canopy Plant Biodiversity and Conservation Biology

General Considerations of Canopy Plant Conservation

Because of their small size, high degree of endemism, and frequent microsite specificity, epiphytes may be more vulnerable to human-induced disturbance than terrestrial plants. Methods to conserve existing epiphyte populations and floras have been discussed (e.g., Lowman and Nadkarni, 1995). Studies have shown the value of older trees in forests as habitats for certain sensitive species.

Effects of Forest Fragmentation and Habitat Conversion

The effects of forest fragmentation and habitat conversion on canopy plant diversity are poorly documented, especially in the tropics. It is generally accepted by researchers that secondary bryophyte and lichen communities are very different from those in primary forests (Gradstein *et al.*, 1989). Most studies indicate a decrease in species richness between secondary habitats and primary forests, and even disturbances at small spatial scales (within a branch) are reported to result in a decrease in diversity.

Shade epiphytes growing in the understory are more affected by habitat conversion than the sun epiphytes of the canopy, but not all sun epiphytes are able to recolonize following disturbance. The available data from investigations of the regeneration rates of temperate and subtropical canopy plants indicate that many species are slow to recover. The rates for bryophytes range from 25 years in Australia to 80–100 years in California. In Britain, it is estimated that lichens may require 500 years to successfully regenerate.

Effects of Global Environmental Change

Water stress is a major limiting factor for plants inhabiting the crowns of trees. A rise in global temperatures may have an impact locally on the relative humidity of some forest canopies. Experimental work along an altitudinal gradient indicates that the species composition of canopy plant communities may be altered by such changes in temperature and humidity (Nadkarni and Solano, 2002).

Vulnerability of Canopy Plants to Extinction and Invasion

With one exception (an orchid, Sosa and Platas, 1998), there are no records of a specific canopy plant extinction in modern times. However, numerous endemic species are endangered or threatened by habitat conversion (Gradstein *et al.*, 1989). A great deal of information indicating that lichens are very susceptible to air pollution and metal ion deposition has accumulated over the past several decades (Rhoades, 1995). Gradstein *et al.* (1989) suggest that relatively small reserves containing a diversity of life zones should suffice to conserve cryptogam biodiversity if the reserve is large enough to maintain a viable population of host trees. However, they warn that these recommendations are based on very preliminary data and more inventory data and taxonomic work are needed to better define species ranges and to determine which species are locally rare or endemic. Mistletoes are generally very susceptible to environmental changes (Calder and Bernhardt, 1983).

Areas for Further Study

General Considerations

Habitat loss and climate change are growing threats to plant communities. Arboreal plants provide many opportunities and challenges for biologists from many disciplines, and because these plants have no access or sporadic access to terrestrial soil, they make excellent experimental subjects to study physiology and stress. Canopy plants warrant attention for the roles they play in forest dynamics, which affect biodiversity, productivity, and nutrient cycling. A list of research questions was created for vascular epiphytes (**Table 6**); these questions can also be related to the study of other types of canopy plants.

Monographs and Inventories

There is a pressing need for extensive and intensive work on plants that live in the canopy. However, the lack of resident plant taxonomists is a serious concern. There are many more taxonomists in the more developed countries where the resources and infrastructure exist to train students, but there are relatively few specialists in less developed countries where many of the biological resources exist.

However, efforts to create monographs and inventories of canopy-dwelling plants have been increasing. For example, botanists at the Missouri Botanical Gardens and their collaborators have compiled inventories of regional floras. To date, Peru, Panama, Venezuela, and the Guianas have received a great deal of attention, and the study of other floras (e.g., Colombia and Chile) is increasing. To fully represent regional biodiversity, it is crucial that botanists collect plants in the canopy. Likewise, to fully understand global biodiversity, the generally undercollected groups must be collected in the canopy (e.g., lianas and cryptogams). In terms of cryptogams, work needs to continue on broad regional inventories of all tropical species and of crustose lichens worldwide. The bryoflora of Australia is particularly poorly known.

Table 6 Research questions and opportunities for canopy plants

Subject	*Obvious*	*Questions remaining*
1. Fidelity to canopy versus other substrates	Occurrence on trees ranges from accidental to obligate.	What factors differentiate canopy from terrestrial substrates for the obligate epiphyte? How has specialization for arboreal life compromised capacity to survive on the ground?
2. Requirements for specific types of arboreal substrates	Specific epiphytes typically colonize only subsets of the many types of substrates present in occupied tree crowns.	What plant characteristics determine microsite requirements for twig, bark, humus, ant-nest garden, etc., epiphytes?
3. Plant adjustments to the often transitory and relatively unpredictable supplies of moisture in forest canopies	Broadly occurring accommodations to drought (e.g., CAM, xeromorphy) are particularly well developed among the epiphytes.	What is the nature of the moisture supply in forest canopies and how are mechanisms such as photosynthetic pathways, osmotic balance, and stomatal behavior fine-tuned to reduce risk and maximize effective use of available moisture?
4. Plant adjustments to the absence of mineral soil	A variety of organic substrates, including the products of mutualistic biota, serve in lieu of earth soil as sources of nutritive ions.	How is impounded litter processed for phytotelm epiphytes? How substantially do ant mutualists contribute to the nutrient budgets of associated epiphytes? How are the more oligotrophic epiphytes (e.g., atmospheric bromeliads) equipped to scavenge scarce ions and use them economically?
5. Impacts of arboreal ants	Some epiphytes require ants for dispersal and to provide rooting media.	How much arboreal flora beyond the obvious ant-nest garden and myrmecotrophic species are dependent on ants for dispersal, substrates, and defense?
6. Epiphytic vegetation as a resource for canopy fauna	Much arboreal fauna, particularly invertebrates, use epiphytes as resources.	What is the full extent of this dependence and what are the broader consequences of these dependencies for the forest community?
7. Epiphyte involvements in nutrient cycles	Nutritional piracy exists. Epiphyte biomass sometimes contains much of the nutrient capital present in a forest ecosystem.	To what degree and under what conditions does the presence of an epiphyte load have an impact on the nutritional status of a phorophyte?
8. Impacts on community productivity and patterns of resource use	Resources present in canopy soil (e.g., N and P) at least sometimes yield photosynthetic returns at different rates than those of supporting soil-rooted vegetation.	How does the presence of substantial canopy soil affect aggregate forest productivity and help determine overall resource-use efficiency?
9. Conservation	Because many epiphytes occupy narrow ranges (especially orchids), often in regions of rapid development, endangered status is correspondingly common.	What conservation strategies are likely to preserve the greatest diversity of epiphytes?
10. Indicators of habitat quality and global change	Some epiphytes possess characteristics that impart extraordinary utility as air quality monitors.	How can epiphytic vegetation be more effectively used to monitor changing conditions in the troposphere?
11. Succession	Presumed serai stages identified.	Do species displace one another on bark? If so, by what mechanisms?
12. Community organization	Species often co-occur in predictable assemblages, but often distribution and spacing among individuals are random.	Are the factors responsible for the distributions and combinations of species on bark primarily density dependent or density independent?

Source: Modified from Benzing DH (1990) *Vascular Epiphytes*. Cambridge, UK: Cambridge University Press.

Herbaria and Databases

Certain herbaria have significant canopy plant collections (Madison, 1977; Kress, 1986). Herbarium studies have been conducted at the Harvard University Herbaria, the Marie Selby Botanical Gardens, the Huntington Botanical Gardens, the Herbario Nacional Colombiano, the State University of Utrecht herbarium, and the University of Florida.

Several major botanical gardens have produced useful databases. For example, The Missouri Botanical Gardens has developed an on-line database (TROPICOS) for the floras of several Neotropical countries. Such databases can be of great use to systematic biologists and conservationists by providing the most up-to-date information available. With the increasing attention being paid to canopy plant ecology (e.g., Lowman and Nadkarni, 1995), perhaps it is time to initiate a canopy plant biodiversity database.

Experimental Fieldwork

Experimental field studies to investigate the potential effects of forest harvesting on plant community composition and species richness should be conducted, especially in tropical regions, so as to include biodiversity objectives in forest management practices. Humus epiphyte communities growing in bryophyte mats, for example, are ideal for experimental fieldwork because entire moss mat communities are easily transplanted with minimal disturbance to the rooting medium. Transplanting epiphytes along an altitudinal gradient is useful in helping to predict the effects of environmental change. The experimental removal of epiphytic mats is also useful in monitoring epiphyte succession and recolonization.

Ethnobotany of Canopy Plants

Ethnobotanical knowledge and usage of canopy plants is widespread in cultures around the world (Lowman and Nadkarni, 1995). However, the pharmaceutical potential of canopy plants has only begun to be investigated. In particular, the bryophytes and lichens have not been rigorously explored in this regard. The potential importance of canopy plants for human use may spark resources needed to learn more about both applied and basic aspects of these diverse organisms.

Conclusions

Due to the extremely high diversity of epiphytes, as well as the relatively unknown nature of their response to disturbance and resilience at all spatial scales, canopy plants constitute an entity that requires the highest level of awareness for good stewardship and conservation. Although many epiphytic taxa and their environments are clearly healthy and vibrant, others may be vulnerable to changes in conditions wrought by human causes: deforestation, forest fragmentation, pollution, and climate change. Attention should be paid to the maintenance of epiphyte biodiversity and the ecosystem services they provide not only in primary forests, but also in secondary forests, pasture relict trees, and plantations. Such areas are especially critical in tropical regions and in temperate rainforests where land areas of primary forests are now dwindling and in decline.

Acknowledgments

This summary of canopy plant diversity was supported by the National Science Foundation. The authors thank the Monteverde Cloud Forest Reserve and the Tropical Science Center in San José, Costa Rica for maintaining and protecting the Research Area. Funds were provided by Evergreen State College Faculty Sponsored Research, the National Science Foundation (Ecology Program (BSR 86-14935, 89-18006, and DEB-0956301), Long-Term Studies Program (DEB 96-15341), Database Activities Program (BIR 96-30316)), the Whitehall Foundation, the National Geographic Society, and the Evergreen Foundation. Many plant taxonomists provided information on epiphyte floristics. The authors thank A. S. Taylor (Zamiaceae), T. Croat (Araceae), R. Faden (Commelinaceae), H. Luther and B. Hoist (Bromeliaceae), H. Kenedy (Marantaceae), L. Thein (Winteraceae), P. Berry and P. Hoch (Onagraceae), P. Lowry and D. Frodin (Araliaceae), M. Endress (Apocynaceae), and W. J. Kress (for many families). Camila Tejos and Dennis Aubrey provided critical help in updating an earlier version of this article.

See also: Forest Canopies, Animal Diversity. Plant Biodiversity, Overview. Tropical Forest Ecosystems

References

Ambrose AR (2004) *Water-Holding Capacity of Canopy Soil Mats and Effects on Microclimates in an Old-Growth Redwood Forest, Arcata*. California: Master of Science Thesis at the Humboldt State University.

Benavides D, Duque MAJ, Duivenvoorden J, Vasco A, and Callejas R (2005) A first quantitative census of vascular epiphytes in rain forests of Colombian Amazonia. *Biodiversity and Conservation* 14: 739–758.

Benzing DH (1987) Vascular epiphytism: Taxonomic participation and adaptive diversity, Ann. *Missouri Botanical Gardens* 74: 183–204.

Benzing DH (1990) *Vascular Epiphytes*. Cambridge, UK: Cambridge University Press.

Calder M and Bernhardt P (eds.) (1983) *The Biology of Mistletoes*. Sydney, Australia: Academic Press.

Cardelus CL and Chazdon RL (2005) Inner-crown microenvironments of two emergent tree species in a lowland wet forest. *Biotropica* 37: 238–244.

Diaz IA, Sieving KE, Pena-Foxon ME, Larrain J, and Armesto JJ (2010) Epiphyte diversity and biomass loads of canopy emergent trees in Chilean temperate rain forests: A neglected functional component. *Forest Ecology And Management* 259: 1490–1501.

During HJ (1979) Life strategies of bryophytes: A preliminary review. *Lindbergia* 5: 2–18.

Erwin TL (1998) The tropical forest canopy: The heart of biotic diversity. In: Wilson EO (ed.) *Biodiversity*, pp. 123–129. Stanford, CA: Stanford University Press.

Gentry AH and Dodson CH (1987) Diversity and biogeography of neotropical vascular epiphytes. *Annals of the Missouri Botanical Gardens* 74: 205–233.

Gradstein SR, van Reenen GBA, and Griffin D (1989) Species richness and origin of the bryophyte flora of the Colombian Andes. *Acta Botanica Neerlandia* 38(1989): 439–448.

Hooper E and Haufler C (1997) Genetic diversity and breeding system in a group of neotropical epiphytic ferns (Pleopeltis; Polypodiaceae). *American Journal of Botany* 84: 1664–1674.

Keller H, Skrabal M, Eliasson U, and Gaither. T (2004) Tree canopy biodiversity in the Great Smoky Mountains National Park: Ecological and developmental observations of a new myxomycete species of Diachea. *Mycologia* 96: 537–547.

Kress WJ (1986) The systematic distribution of vascular epiphytes: An update. *Selbyana* 9: 2–22.

Lindo Z and Winchester N (2007) Oribatid mite communities and foliar litter decomposition in canopy suspended soils and forest floor habitats of western redcedar forests, Vancouver Island, Canada. *Soil Biology & Biochemistry* 39: 2957–2966.

Lowman MD and Nadkarni NM (1995) *Forest Canopies*. San Diego, CA: Academic Press.

Madison M (1977) Vascular epiphytes: Their systematic occurrence and salient features. *Selbyana* 2: 1–13.

Nadkarni NM (1994) Diversity of species and interactions in the upper tree canopy of forest ecosystems. *American Zoologist* 34: 70–78.

Nadkarni NM and Solano R (2002) Potential effects of climate change on canopy communities in a tropical cloud forest: An experimental approach. *Oecologia* 131: 580–586.

Nieder JJ, Prosperi, and Michaloud G (2001) Epiphytes and their contribution to plant diversity. *Plant Ecology* 153: 51–63.

Ozanne C, Anhuf D, Boulter S, *et al.* (2003) Biodiversity meets the atmosphere: A global view of forest canopies. *Science* 11(301): 183–186.

Parker GG (1995) Structure and microclimate of forest canopies. In: Lowman MD and Nadkarni NM (eds.) *Forest Canopies*. San Diego, CA: Academic Press.

Peck JE and McCune B (1997) Remnant trees and canopy lichen communities in western Oregon: A retrospective approach. *Ecological Applications* 7: 1181–1187.

Pentecost A (1998) Some observations on the biomass and distribution of cryptogamic epiphytes in the upper montane forest of the Rwenzori Mountains, Uganda. *Global Ecology and Biogeography Letters* 7: 273–284.

Putz F and Mooney HA (1991) *The Biology of Vines*. Cambridge, UK: Cambridge University Press.

Rhoades FM (1995) Nonvascular epiphytes in forest canopies: Worldwide distribution, abundance, and ecological roles. In: Lowman MD and Nadkarni NM (eds.) *Forest Canopies*, pp. 353–408. San Diego, CA: Academic Press.

Schofield WB (1992) Bryophyte distribution patterns. In: Bates JW and Farmer AM (eds.) *Bryophytes and Lichens in a Changing Environment*. Oxford, UK: Oxford University Press.

Sillett SC (1995) Branch assemblages in the forest interior and on the clearcut edge of a 700-year-old Douglas fir canopy in western Oregon. *The Bryologist* 98: 301–312.

Snell K and Keller H (2001) Biodiversity of tree canopy cryptogams in the Great Smoky Mountains National Park. Transactions of the Missouri Academy of Sciences. *Senior Division* 35: 53.

Sosa V and Platas T (1998) Extinction and persistence in rare orchids. *Conservation Biology* 12: 451–455.

Stone J, Sherwood M, and Carroll G (1996) Canopy microfungi: Function and diversity. *Northwest Science 70: Special Issue* 37–45.

Williams-Linera G and Lawton RO (1995) The ecology of hemiepiphytes in forest canopies. In: Lowman MD and Nadkarni NM (eds.) *Forest Canopies*, pp. 255–283. San Diego, CA: Academic Press.

Wolf JHD (1994) Factors controlling the distribution of vascular and non-vascular epiphytes in the northern Andes. *Vegetatio* 112: 15–28.

Forest Ecology

Timothy J Fahey, Cornell University, Ithaca, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Disturbance An event that results in a radical change in environment, usually as a consequence of death of the dominant plants.
Food web A set of species that live together and a specification of which species "eat" which other species.
Forest stand A relatively homogeneous forest landscape unit that can be distinguished from neighboring units by forest age or composition.
Forest structure The arrangement of all the parts of the forest stand – stems, branches, leaves, roots, and so on.
Keystone species A species that performs some crucial activity in the ecosystem and for which there is little or no redundancy.
Succession Sequential change in the relative dominance of various species in the biotic community at a location.

Introduction

Forests occur naturally wherever environmental factors permit the upright woody growth form of trees to be a successful strategy. The primary advantage that trees realize in such settings is the ability to gain a competitive edge over any smaller neighbors in gaining access to solar radiation for photosynthesis. Thus, forests dominate the natural landscape in all regions except where the upright woody growth form is not possible or where other mitigating environmental factors eliminate any competitive advantage of being tall. For example, in the Great Plains of North America grass fires were usually too frequent to allow trees to become established and outcompete the grasses for sunlight.

Ecologists have classified the forests and other vegetation of the world into biological regions, or biomes, on the basis of the physiognomy (outward form) and species composition of the dominant plants. The biome classification represents ecological patterns at a very large scale. However, the patterns that ecologists endeavor to explain occur at a wide range of spatial scales, from global to local. Global scale maps of the world's biomes have been constructed to illustrate the broad distributional patterns of the biota. Ecological patterns at smaller scales are represented by subdividing the various biomes to provide more detailed classifications and maps. In essence, the forest ecologist regards the global ecosystem as consisting of a nested series of ecological associations, at smaller and smaller spatial scales, down to the level of the relatively uniform forest stand.

The focus of ecological classification systems on the dominant plants is not just a convenience; the dominant plants in terrestrial ecosystems (e.g., the trees in the forest) play a crucial role in setting the stage for all the other organisms because they convert solar energy into food for the food web and they form the three-dimensional structure that constitutes the habitat of most associated biota. Hence, it is of primary importance in studying the ecology of forests to understand what controls the distribution and abundance of trees. However, this is not to say that the other organisms are not important to the functioning of the ecosystem; in fact, many of the less prominent organisms play important roles in regulating the distribution and abundance of the trees.

Underlying these ecological classification systems and maps at the broad scale is the principle that climate is the environmental feature that exerts primary control over the distribution of organisms on the earth's surface. This principle was realized long ago by European geographers who observed during their explorations the coincidence between the broad patterns of climate and vegetation physiognomy on earth. The key components of climate, temperature and precipitation, exert this control through their effects on the growth of plants. Different plant traits, expressed in part as whole plant physiognomy, prove to be most suitable for growth and survival under different combinations of temperature and precipitation. Hence, climatic patterns largely coincide at the broad scale with the patterns of distribution of forest biomes. Expressed most simply, the effect of climate on vegetation distribution can be plotted in terms of mean annual temperature and precipitation (**Figure 1**); forest vegetation is restricted to relatively moist climates. However, annual averages cannot account for the important effects of climatic seasonality – variations in temperature and precipitation through the year – in determining biome distributions. More complicated systems of expressing the influence of climate account for the important effects of seasonal drought and of subfreezing temperatures on plant growth and activity.

The principle that climate exerts primary control over the distribution of organisms fails to account for the observation that some equivalent climates support vegetation with differing composition and physiognomy. For example, the temperate rain forests of the Pacific Northwest of North America are dominated by evergreen needleleaf trees whereas in some equivalent climates of the southern hemisphere (New Zealand and South America) evergreen broadleaf forests may dominate. These differences reflect the fact that the flora and fauna that have developed by biological evolution in the various regions of the world are quite distinct because they have been geographically isolated from one another. Several biogeographic zones have been identified by botanists and zoologists, and these zones reflect geologic history, especially continental drift. For example, plant geographers typically recognize four plant domains whose floras are distinct from one another: (1) paleotropical (Old World Tropics),

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00058-7

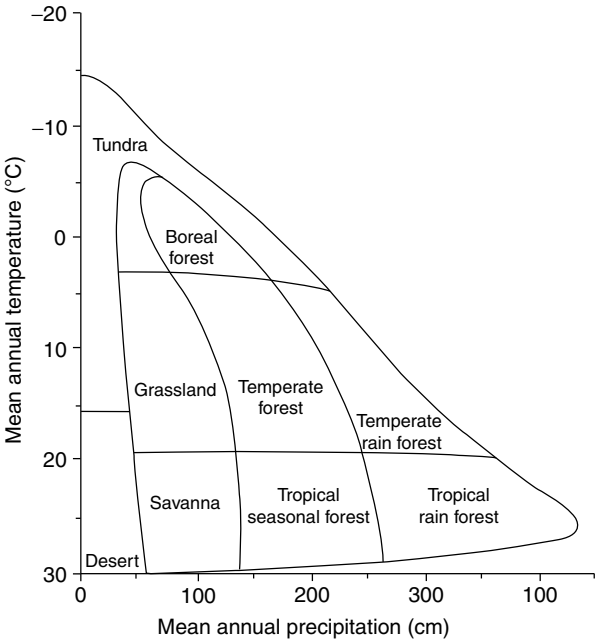


Figure 1 The effects of mean annual precipitation and temperature on vegetation. Whittaker (1975).

(2) neotropical (New World Tropics), (3) north temperate, and (4) south temperate. To expand on our earlier principle, then, climate together with the available flora and fauna exert primary control over the distribution and abundance of organisms on earth. This principle points toward the likelihood that human effects on climate and the introduction of exotic species across the globe will have profound consequences for ecology and biodiversity.

At a more localized scale, species distribution and abundance are also affected by secondary environmental factors, topography and soils. Topography exerts its influence in part by locally modifying climate; for example, in the northern hemisphere south-facing slopes are warmer and drier than north-facing slopes. Also, topography influences the environment through the action of gravity moving matter downhill (e.g., water, soil particles). Soil supplies the essential environmental resources, water and mineral nutrients, for tree growth. Soil properties reflect the combined effects of geologic, climatic, and biologic forces, emphasizing the complex web of cause and effect that underlies ecological patterns and processes.

Geologists have divided the earth's landscape into physiographic provinces that reflect the effects of geological processes on the earth's surface features. Within each physiographic province, relatively orderly and recurring patterns of topography and soils are observed that differ fundamentally from those of neighboring provinces. For example, in the coastal plain province of the eastern United States, sandy sediments have emerged from the receding ocean to leave a gently rolling landscape of porous, sandy soils. Forest patterns reflect subtle variations in topography and drainage. In contrast, directly to the west, the ridge-and-valley province is marked by recurring combinations of bold narrow ridges of resistant rocks and intervening valleys underlain by softer substrates like limestone and shale. The forest patterns in this province directly mirror the striking gradients in environment that result from the combination of topography and soils determined by the bedrock geology of the province. An understanding of the ecology of these two regions begins with the recognition of the underlying differences in geological forces that have shaped their physiography.

One additional factor that strongly regulates the composition of the biotic community is the legacy of disturbance events. Disturbance is a natural phenomenon in all ecosystems and is defined as any event that results in a change in environmental conditions and resource availability, usually as a consequence of death of the dominant plants. Among the most prominent natural disturbance agents in forest ecosystems are fire, windstorms, and irruptions of pests and pathogens. When disturbances occur at a spatial scale that is much larger than the area occupied by individual dominant plants (i.e., large-scale disturbances), they so profoundly alter the environment that the suite of plants that subsequently colonizes the disturbed site may be quite different from the original community. These large-scale disturbances initiate the ecological process of succession – that is, successive changes in the composition of the biotic community occur as a result of progressive changes in environmental factors. Hence, the actual composition of the biotic community at any particular time and place depends not only on climate, flora and fauna, topography and soils, but also on the time interval since the last large-scale disturbance that initiated successional change – as well as the nature and intensity of that disturbance event. That disturbance has played a major role in shaping ecological patterns and processes is evident from the many traits of the flora and fauna that reflect the selective force associated with disturbance.

The natural disturbance regime that characterizes any particular ecosystem depends on exogenous factors that act as disturbance agents (e.g., the combination of drought, lightning, and wind that favors the occurrence of fire; or frequent exposure to hurricanes or tornadoes in certain geographic regions) as well as endogenous factors, such as the traits of the plants themselves, that influence the frequency or intensity of the disturbance. For example, pine forests are more prone to fire disturbance than deciduous broadleaf forests in part because the fuels produced by pine trees are more flammable. And forests on sandy soils are more prone to fire because coarse soils dry out more rapidly than fine-textured soils. Thus, the influences of disturbances, environment, and biota may be mutually reinforcing in shaping ecological patterns and processes.

The prevalence of disturbance and succession in natural forest ecosystems complicates the task of classifying and mapping forest distributions, particularly at smaller scales, because forest composition is continually changing. Recognizing this problem, ecologists have defined the climax forest association as the assemblage of species that would persist under any particular combination of environmental factors in the absence of large-scale disturbance. In many regions, however, the recurrence of large-scale disturbances is naturally so frequent that climax forest associations rarely develop, and the practical value of such a strict climax concept is somewhat

limited. However, in some regions where large-scale disturbances are infrequent, particularly under humid climates (where fires are rare) and areas not often exposed to intensive windstorms (i.e., the humid tropics and some temperate areas), climax forest associations probably were common prior to the advent of anthropogenic influences. Natural disturbance regimes in these forests consisted of small-scale events resulting from the death of individuals or small groups of trees in a patchwork mosaic. Although the composition and structure of various patches would differ, at a larger scale the average composition of the forest ecosystem would remain relatively steady (i.e., it would exhibit a shifting-mosaic steady-state). These observations illustrate the importance of the scale of observation to our understanding of ecological patterns and processes.

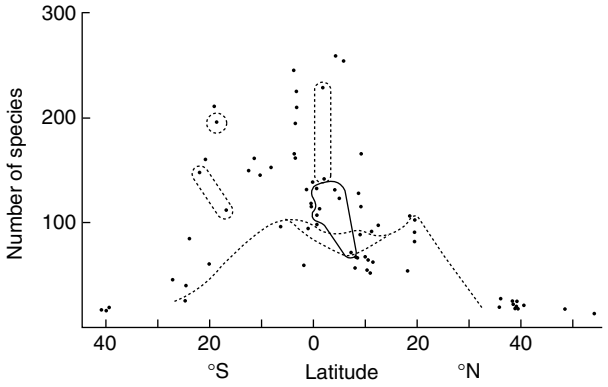


Figure 2 Species richness of 0.1-ha samples of lowland (<0.1000 m) forest as a function of latitude. Dashed line separates dry forest (bottom) from moist and wet forest (top) with intermediate sites (moist forest physiognomy despite relatively strong dry season) indicated by alternate lines. X, the anomalous Coloso and Loma de los Colorados sites in northern Colombia. Reprinted from Gentry AH (1995) Diversity and floristic composition of neotropical dry forests. In: Bullock SH, Mooney HA, and Medina E (eds.) *Seasonally Dry Tropical Forests*, pp. 146–194. Cambridge, England: Cambridge University Press.

Biodiversity in Forest Ecosystems

General Patterns

The number of species observed in any particular forest varies markedly across the earth's forest biomes. Biodiversity can be conceived as consisting of three distinct elements, termed the gamma, alpha, and beta diversity. The total diversity in a large area, the gamma diversity, can be partitioned into two components, the local (alpha) diversity in a single habitat or forest stand, and the turnover of species between stands, the beta diversity. High gamma diversity could be associated either with high alpha diversity, high beta diversity, or both. Ideally, considerations of the patterns of biodiversity across forest regions need to account for the contributions of these different elements of biodiversity. Unfortunately, studies of forest biodiversity patterns generally have not provided samples of sufficient detail to resolve these elements, and more systematic approaches are needed. Nevertheless, many valuable insights into the patterns and causes of biodiversity variation across forest regions have emerged from research to date.

The most striking pattern of forest biodiversity is the latitudinal gradient: Much higher biodiversity is observed both locally and regionally in tropical than in temperate forests. Because observations are most comprehensive for tree species, the following description focuses on these taxa.

The general relationship between latitude and the alpha diversity of trees is illustrated by a plot of the number of species in 0.1-ha samples taken from around the world (**Figure 2**). At one extreme are lowland tropical rain forests in the upper Amazon basin where on average every second tree in a sample belongs to a different species; at the other extreme, a variety of natural, mono-specific forests are common in many temperate and boreal regions. Within the tropical region considerable variation in alpha diversity has been observed (**Figure 2**). The two most prominent features correlated with this variation are annual precipitation and biogeographic province or region. For example, the average alpha diversity (number of tree species in 0.1-ha samples) of neotropical lowland wet and moist forests is 152 species, whereas for seasonally dry forests the average is 65. And the alpha diversity of trees in Africa is generally much lower than for eastern Asia and America (**Figure 3**). The extremely high alpha diversity of lowland tropical rain forest is not paired with equally high beta diversity. For example, the 307 tree species identified on a single 1-ha plot in the Ecuadorian Amazon constituted about 16% of the total tree flora (trees $\geq$5 cm diameter) of Amazonian Ecuador and a single 50-ha plot at Pasoh, Malaysia contained 830 species, 20–30% of the total tree flora of this country.

By comparison, tree species diversity in the temperate latitudes is much lower: The enormous area of temperate zone forests harbors only 1166 tree species, not much more than the 50-ha plot at Pasoh! As for the tropics, however, striking geographic differences in the diversity of tree taxa are observed in the temperate zone. The highest gamma diversity is in east-central Asia, with intermediate levels in eastern North America and lowest in Europe (**Table 1**).

A comprehensive explanation of these global patterns in tree species diversity remains elusive, but a growing consensus on the role of historical or biogeographic factors and local physical habitat factors is emerging. The overall lower diversity of temperate than tropical regions probably is explained in part by the physiological constraints on colonization imposed by the need to tolerate subfreezing conditions and by the smaller contiguous area in temperate regions. Regional differences between biogeographic provinces in the temperate zone appear to owe in large part to higher extinction rates in Europe and North America during the Pleistocene glacial epochs as well as greater access of the east Asian region to dispersal routes from the tropics. Similarly, the combination of geographic area and climatic differences probably explains the contrasts in diversity between tropical Africa versus America and east Asia and between wet and dry forests. As noted by Latham and Ricklefs (1993), "further resolution of the causes of diversity patterns will require new paleontological, biogeographical and taxonomic data and synthesis" (p. 310).

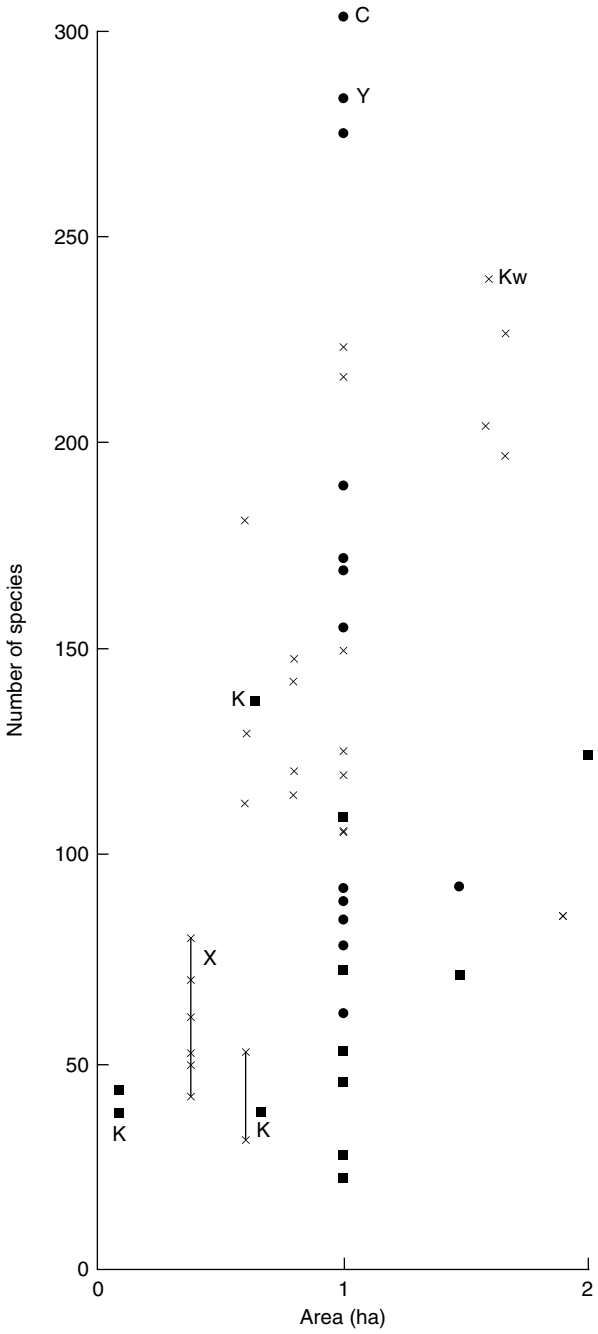


Figure 3 Number of species shown among trees of at least 0.1 m in diameter on small plots in tropical lowland rain forest. ●, America; x, Eastern tropics; ■, Africa. Lines connect sample plots that lie close together. Reprinted from Whitmore TC (1998) *An Introduction to Tropical Rain Forests*, 2nd edn. Oxford, England: Oxford University Press.

Forest Structure and Pattern and Disturbance Regimes

The local biodiversity in particular forests depends on the complex suite of factors that characterize the habitats of individual species. These factors include such components as the species composition, phenological timing, structural complexity, and horizontal patterning of the vegetation, which in turn depend on environment and the legacy of disturbances. The dominant plants (i.e., trees in forests) play a pivotal role in defining the habitats for associated organisms by providing food and shelter and by regulating the local microenvironment. For some organisms (e.g., many insects) host-specific interactions with particular dominant plants result in strong correlations in their distributions, so that forest composition is the key factor influencing the composition of associated species. For other organisms, the structural and horizontal patterning of the dominant vegetation may be more important than composition alone.

The importance of the three-dimensional spatial arrangement of the branches and leaves of the plants in defining animal habitats was realized by pioneers in the study of ecology and evolution. MacArthur and MacArthur (1961) demonstrated that the diversity of birds in forests of the eastern United States could be predicted by the structural complexity of the vegetation (**Figure 4**): in habitats with high foliage-height diversity (FHD; defined by the formula $\mathrm{FHD} = \%_i p_i$. ln p_i where p_i is the proportion of total foliage area in the ith layer), bird diversity was much higher than in habitats with low FHD. Moreover, this relationship largely transcended differences in plant species diversity. Similarly, for ground-dwelling organisms, the complexity of habitat at the soil surface – including vegetation cover, litter, rocks, fallen logs, and moisture – is strongly correlated with local biodiversity. By extension the below-ground structural complexity might influence diversity of subterranean organisms, but little work on this topic has been accomplished.

The horizontal pattern of forest vegetation, particularly the arrangement of relatively uniform forest stand units across the landscape, influences biodiversity of the region and is the subject of a newly emerging branch of ecology – landscape ecology. The edges or ecotones (zones of rapid change in plant species composition) between forest stand units or forest associations often provide qualitatively different habitat than the interiors of those units; hence, the size, shape, and spatial arrangement of forest stands in the landscape influence the populations of associated species. Edges and sharp ecotones between stands may arise because of environmental discontinuities (e.g., topographic depressions where soil water accumulates) or because of the legacy of past disturbances. Thus, the local biodiversity of forests reflects both the compositional and structural diversity of the plants as well as the arrangement of units of relatively uniform composition and structure that we define as forest stands.

The structure of any particular forest stand traditionally has been defined as the distribution among age or size classes of the trees in a forest stand. Even-aged stands arise as the result of large-scale disturbances in which all or most of the large trees in an area are killed by a natural disturbance agent (e.g., crown fire, hurricane, fungal pathogen) or by human activity. The forest that arises on such a site consists of trees of roughly the same age, from new seedlings that colonize the site or from advanced regeneration (preexisting seedlings and saplings that escaped the disturbance). The process of stand development following large-scale disturbance results in gradual changes in the structure of the forest, as the growth, mortality, and recruitment of new individuals proceeds. These changes in forest structure include not only age structure, but also the

Table 1 Summary by taxonomic level and region of moist temperate forest trees in the Northern Hemisphere

Taxonomic level	*Number of tree taxa characteristic of moist temperate forests in:*				
	Northern, central, and eastern Europe	*East-central Asia*	*Pacific slope of North America*	*Eastern North America*	*Northern Hemisphere (total)*
Subclasses	5	9	6	9	10
Orders	16	37	14	26	39
Families	21	67	19	46	74
Genera	43	177	37	90	213
Species	124	729	68	253	1,166
Families excluding those of predominantly tropical	18	37	18	29	41
Distribution (% of total)	(86%)	(55%)	(95%)	(63%)	(55%)
Genera excluding those of predominantly tropical	41	121	35	77	149
Distribution (% of total)	(95%)	(68%)	(95%)	(86%)	(70%)
Species exclusive of predominantly tropical genera	122	570	66	236	987
(% of total)	(98%)	(78%)	(97%)	(93%)	(85%)

Source: Data from Latham RE and Ricklefs RE (1993) Continental comparisons of temperate-zone tree species diversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 294–314. Chicago: University of Chicago Press.

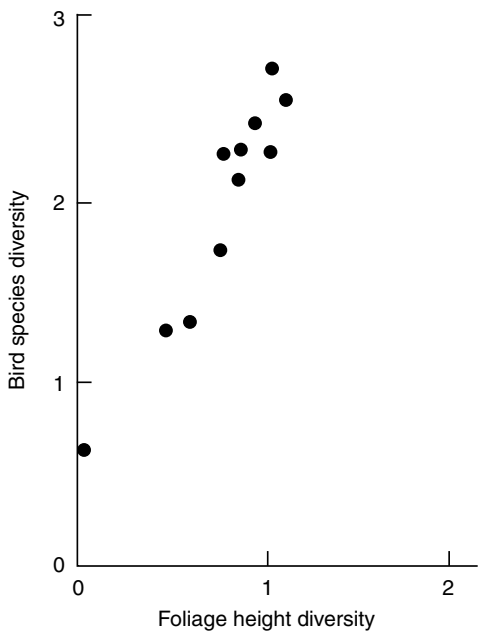


Figure 4 Relationship between foliage height and bird species diversity in areas of eastern North American deciduous forest. Modified from MacArthur RH and MacArthur JW (1961) On bird species diversity. *Ecology* 42: 594–598.

spatial arrangement of the stems, branches, foliage, and roots of the plants that define the habitat of other biota in the forest. Moreover, changes in species composition of the forest usually accompany stand development, as species capable of growing in the shaded understory replace the pioneer species that colonize the disturbed site.

A typical sequence of forest stand development following large-scale disturbance has been characterized by Oliver and Larson (1990) as consisting of four stages: (stage 1) stand initiation, (stage 2) stem exclusion, (stage 3) understory reinitiation, and (stage 4) old growth. In the stand initiation stage, trees colonize the disturbed area. The time interval of this stage varies markedly depending on the severity of the disturbance, environmental factors at the site (e.g., climate and soils), and often herbivory. This stage concludes when the forest canopy becomes closed or when some soil resource (often water) becomes limiting to further increases in the leaf area of the forest. During the stem exclusion stage resources like light and soil water are so limited that suppressed understory trees die and regeneration is severely restricted. In this stage there is usually a continual reduction in the density (number of stems/area) of the original cohort of trees. The structure of the canopy is exceptionally simple at this stage as the individual trees grow in height to co-opt the light resource from neighbors.

In the understory reinitiation stage the overstory begins to break up as canopy trees die and the differential height growth of various species or individuals results in more complex arrangements of the foliage. Increased light reaching the understory favors the establishment and growth of new cohorts of species most capable of surviving in the highly competitive understory environment. The old growth or late successional forest stage is attained as overstory trees age and the canopy develops even greater complexity of structure. Gaps form in the canopy as a result of injury or death of the large, mature individuals, and the previously suppressed individuals are released from severe competition and grow in height. Decaying, coarse, woody debris accumulates on the ground and dead tree snags also provide new habitats for animals. Obviously, this idealized model of forest stand dynamics exhibits myriad local variations depending on the nature of the disturbance, environmental factors, and the tree species that dominate the area.

The changes in forest structure (e.g., foliage-height distributions) that accompany stand development following large-scale disturbance result in consequent gradual shifts in the quality of the habitat for different animals and plants. In the example cited earlier, MacArthur and MacArthur (1961) observed that maintenance of high numbers of bird species in eastern deciduous forests of the United States depended on the adequate provision of three layers of foliage, corresponding roughly to ground vegetation (0–2 ft), shrubs and small trees (2–25 ft), and overstory trees (>25 ft). If nearly all the foliage of the forest is in just one layer, as in the stand initiation and stem exclusion stages of stand development, bird diversity is much lower than in the old-growth stage, when canopy stratification becomes prominent.

The nature and degree of vertical stratification differ among forests. For example, temperate deciduous forests in the mature stage often exhibit the three strata just identified: An overstory stratum occupied by the canopy trees, an intermediate stratum represented by the crowns of saplings and understory species like dogwoods and hornbeams, and a ground stratum of low shrubs and forbs. In contrast, in the lowland tropical rain forest, very tall "emergent" trees with broad shallow crowns overtop the main canopy, which may be subdivided into two or more additional strata above the understory layers. And in the boreal forest the low stature and conical crowns of the conifers often preclude the formation of strong vertical stratification. These structural differences contribute to the contrasts in the number of distinct habitats provided for associated biota.

As noted earlier, many forest regions are only rarely affected by large-scale disturbances because they are both too moist to carry fires and not subjected to catastrophic windstorms. In these regions the steady-state forest is characterized as a shifting-mosaic landscape of small patches of different sizes and shapes, each patch reflecting the legacy of disturbance caused by the death of individuals or small groups of overstory trees. Each of the patches may follow a sequence of development analogous to that outlined for large-scale disturbances, but the overall structure of the forest is dependent on the arrangement of the tapestry of patches that comprise the larger forest stand. The edges between these patches and the vertical distribution of structural elements represent important dimensions of the habitat variability that permits species coexistence in forests.

Recognizing the importance of forest structure and pattern for biodiversity, ecologists are developing new, more sophisticated approaches for quantifying these parameters. Because the structural features that are important in determining animal and plant habitats differ among taxonomic groups and forest types, it is unlikely that any single approach will provide a universal standard by which forest structure and biodiversity can be related. Current efforts are utilizing new tools in the areas of spatial statistics, computer modeling, and remote sensing to provide suitable protocols for evaluating the connections between forest structure and pattern, management activities, and biodiversity of various groups of biota. These efforts will provide a better basis for understanding how forest ecology and biodiversity are related.

Food Webs and Community Organization in Forests

A food web is a set of species that live together and a specification of which species "eat" which other species. Plants provide the base of the food web by converting solar energy into biomass; herbivores and detritivores utilize living and dead biomass, respectively, to build their own tissues; and these organisms are in turn consumed by predatory species. Although a connection between the structure of food webs and biodiversity seems axiomatic, the exact nature of this connection is extremely complex. The biodiversity of a particular community is expressed in three different elements of food web structure: (a) the food chain length – the number of trophic links in the food web (producer–l°consumer–2°consumer–3°consumer, etc.); (b) the number of distinct trophospecies – the set of all species that share some particular set of predators and prey – within a trophic level; and (c) the diversity of species constituting each trophospecies.

In forests, as distinct from most other biomes, the first link in the food web is dominated energetically by the detritivores rather than the herbivores. That is, most of the biomass in forests is consumed after the plant tissues die and are added to the soil as detritus. Thus, the detrital food web dominates the energetics of forest ecosystems and much of the complexity of the detrital food web in forests remains to be explored. These energetic considerations are not translated in a simple way to biodiversity in the respective food webs. The variety of different food sources available to herbivores in forests (leaves, stems, fruits, seeds, flowers, and roots of different plant species) provide numerous niches for their diversification, whereas most of the biomass energy available to detritivores is in the form of woody tissues which are structurally and biochemically so similar from species to species of trees that the diversification of wood decay organisms is somewhat limited. For example, one taxonomic group of insects, the higher termites (family Termitidae), overwhelmingly dominates in the comminution and decomposition of plant biomass in many tropical and warm temperate forest biomes. Nevertheless, in comparison with herbivore-dominated biomes (e.g., grasslands, aquatic ecosystems), the role of detritivores in forest biodiversity is probably relatively high.

The interactions of species in the forest community are not entirely competitive and predator–prey in nature. Some of the most fascinating interactions provide selective benefits to both of the interacting individuals or species populations – mutualistic and symbiotic relationships. Mutualism refers generally to a relationship in which two interacting species enhance their survival, growth, or reproduction, while symbiosis refers more specifically to two such organisms living together in close association. The most important symbiotic mutualism in forests is the mycorrhiza, an association between the mycelia of fungi and the roots of trees. The tree roots provide a supply of food to the fungus, which in turn increases the capability of the plant to acquire soil nutrients and water. All forest trees are mycorrhizal, and each tree species may harbor dozens of different fungal species in this mutually beneficial relationship. Two distinct types of mycorrhizae are common in forest trees – the ectomycorrhizae and endomycorrhizae. These types differ taxonomically, anatomically, and physiologically. The possible role of these mycorrhizal associations in regulating

the diversity of forests was pointed out by Connell and Lowman (1989). They observed that pockets of low-diversity forest are found within the matrix of high-diversity tropical rain forest in all tropical regions. These low-diversity forests are composed of trees with ectomycorrhizal associations, whereas most of trees in the high-diversity forest are endomycorrhizal. Functional differences between the mycorrhizal types in soil nutrient acquisition or transfer could maintain or reinforce competitive interactions between individuals in these distinct forest types.

Most mutualisms between plants and animals have developed around the successful completion of the reproductive cycle of the plant – pollination and seed dispersal. Although many trees simply disperse their pollen to the wind, this method of pollination is unreliable when individuals are widely scattered in the forest, as in the species-rich tropical rain forest. Insects, nectivorous birds, and bats visit the plants to exploit them as a source of food and in the process carry pollen from one flower to another. Similarly, plants with seeds too heavy to be dispersed by the wind rely on animals to carry the seeds away from the mother plant. Although in most cases these mutualistic interactions are nonobligate and facultative (at least on one side), many remarkable examples of highly intricate, obligatory interactions have evolved, especially in the tropics. These interactions promote specialization and increased biodiversity.

Human Activity and Forest Biodiversity

Few forests have escaped the effects of human activity. Ancient civilizations decimated forests locally as a source of fuel and fiber and as sites for intensive agricultural production. In the modern era the pervasive influence of industrial civilization on forests has expanded to the regional and global scale through the additional effects of species introductions, air pollution, and likely climatic change. Insights from forest ecology provide a basis to evaluate the implications of these human influences on biodiversity.

Introduction of Alien Species

Alien insects and pathogens typically wreak havoc on host trees because these hosts have not developed adequate defenses through the process of evolution. The result is widespread decline of the host trees throughout their range even to near the point of extinction. The consequences of great reductions in the abundance of such declining species for the wider biotic community are not well understood and undoubtedly vary depending on the characteristics of the declining species (discussed later). Similarly, introductions of other species in different taxonomic groups or at other positions in the food web will have consequences for forest biodiversity that depend on their particular role in the community. Invasive trees may displace congeneric species from the forest community; and introduced herbivores that lack natural controls on their populations may decimate populations of their favored food plant species. Because the pace of species introductions has increased very rapidly in recent years, the ultimate consequence for forest ecosystems and biodiversity will be played out over the coming century.

Forest Harvest

The consequences of tree harvesting for forest biodiversity depend on particular features of the forest and the methods of harvest. Because all forests are regularly subjected to disturbance, if forest harvest practices mimic the natural disturbance regime, then consequences for biodiversity should be minimal. However, the exigencies of the financial bottom line result in harvest practices that do not mimic natural disturbances, and the consequences of actual forest harvest practices for biodiversity may be substantial. Most serious are (a) logging practices that result in the failure of the cut-over site to regenerate (e.g., because of severe damage to soils); (b) the coincident harvest of extensive areas, so that most of the landscape is in a single stage of forest stand development; and (c) recurring harvest on short rotation intervals. Also, forest harvest differs fundamentally from natural disturbance in that wood products are removed from the site; any species that depends on decaying wood for its habitat will be harmed by harvest practices that do not recognize this dependency. Finally, some species are believed to be old-growth obligates (i.e., they depend on old forests to complete their life cycles). These species are threatened when the great majority of natural, old forests in a region enters the harvest-regrowth system of industrial forestry, leaving little old-growth habitat.

Forest Conversion and Fragmentation

Permanent or semipermanent conversion of forested areas to other land uses has more severe consequences for biodiversity than forest harvest. Many forests occur where climate and soils are suitable for permanent agriculture and where the expansion of urban communities gobbles up native vegetation. Maintenance of biodiversity in such regions depends on having protected forest areas large enough to harbor the native flora and fauna. However, general rules to guide forest preserve planning for biodiversity protection are complicated by variations in the habitat requirements of different species. Forests in most agricultural regions occur as small fragments dispersed across the landscape, and how effectively these fragments can maintain viable populations of forest species is a topic of great concern (Schelhas and Greenberg, 1996).

Pollution

Local declines of forests has been associated conclusively with point-source releases of air pollutants, especially sulfur dioxide, fluoride, and toxic metals from smelters. Broad scale, regional effects of air pollution on forests have been more difficult to demonstrate. Regional pollution – by ozone smog in the southwestern United States and by acidic deposition in the eastern United States and Europe – probably has contributed to documented forest declines. Although improvements in emission controls and regulations in these regions are likely to reduce the chances of further damage, rapid industrialization without adequate

emission controls in other regions of the world threatens forest health and biodiversity.

Rapid Climatic Change

The Pleistocene epoch was marked by dramatic climatic shifts that profoundly affected forests and biodiversity. The rapid rise in greenhouse gas concentrations is likely to bring about similar climatic shifts in coming decades or centuries. Of course, the consequences of rapid climate change for forests and biodiversity will depend on a combination of species' natural responses (e.g., dispersal, colonization, natural selection) and human mitigation efforts. In many forest regions certain species, like the dominant trees and particular wildlife populations, are likely to be controlled by management efforts because of their relatively high value to humans. For relatively low-valued species and forest regions, the maintenance of biodiversity may depend on natural mechanisms or heroic human efforts that recognize nonmarket values of species. The current level of understanding of the physiological and population ecology of many forest-dwelling taxa is insufficient to predict the effects of rapid climatic change, but species with limited capacity for dispersal and colonization (e.g., soil invertebrates, perennial herbs) may be most sensitive. The implication of the loss of these species from forest communities will vary depending on the role they play in forest ecosystem function.

Biodiversity and Forest Ecosystem Function

Forests regulate energy flow and cycling of materials in the landscape, collectively known as ecosystem functions. The effects of biodiversity, expressed in terms of species richness, on forest ecosystem functions are not yet clear and apparently not very straightforward. There is great interest in these possible effects as ecologists probe the implications of loss of species diversity for the integrity of ecosystems. Will species extinctions result in destabilization of ecosystem functions and possible feedbacks in the form of undesirable shifts in dominant vegetation types?

Although the role of species richness *per se* in regulating forest ecosystem functions remains unclear, it is well known that loss of particular species from the biota of a community can have important ramifications for energy flow, material cycling, and maintenance of stable biotic composition. That is, all species in the ecosystem are not equal in terms of their quantitative influence on ecosystem function. In particular, for some species there appears to be little or no redundancy with respect to their role in the ecosystem; if such a species performs some crucial activity, its loss from a forest can create havoc for the normal functioning of the ecosystem. These species are known as keystones.

In forests keystone species are represented among many different taxonomic groups or food web positions. For example, at the primary producer level, a nitrogen-fixing tree like *Alnus rubra* in N-poor conifer forests is a keystone species; elephants appear to be keystone herbivores in semiarid Africa; beavers are keystone "ecosystem engineers" in northern forests; and jaguars that prey on seed predators in neotropical forests may be keystone carnivores. The loss of these species has consequences for ecosystem structure, function, and composition that are out of proportion from their individual abundances. Important efforts to conserve forest ecosystem functions in the face of biodiversity loss are focused on the identification of keystone species and ways of maintaining stable populations of keystone species.

Although the broader effects of overall reductions in biodiversity on forest ecosystem function remain more obscure, a specific example will illustrate that this is also a cause for concern. During the past several decades, excessive inputs of nitrogen (from air pollution) to forests in northern Europe have resulted in striking reductions in the abundance of mushroom species. These mushrooms are the fruiting bodies of ectomycorrhizal fungi discussed earlier. Lilleskov *et al.* (2001) have shown that a single host tree, white spruce (*Picea glauca*), maintains associations with about 90 different mycorrhizal fungi in natural, N-poor forests in Alaska, whereas in adjacent N-polluted forests only about five fungal associates are found. The high diversity of this mycoflora in the natural forest certainly represents some degree of functional redundancy. However, if dozens of species are lost from the mycorrhizal fungal flora in temporarily N-polluted regions, some long-term effects on forest ecosystem function are likely. The fungi in the N-rich forests appear to utilize only the mineralized nitrogen sources NH_4^+ and NO_3^-, which are abundant there; with a return to normal, low-mineral N availability, in the absence of mycorrhiza that access organic N forms, the productivity and nutrient cycling in the forests could be altered profoundly. Analogous situations probably apply in other aspects of both the detrital and grazing food webs of forest ecosystems around the world.

See also: Boreal Forest Ecosystems. Deforestation and Land Clearing. Disturbance, Mechanisms of. Fires, Ecological Effects of. Food Webs. Keystone Species. Rainforest Loss and Change. Reforestation. Succession, Phenomenon of

References

Connell JH and Lowman MD (1989) Low diversity tropical rainforests: Some possible mechanisms for their existence. *American Naturalist* 134: 88–119.

Gentry AH (1995) Diversity and floristic composition of neotropical dry forests. In: Bullock SH, Mooney HA, and Medina E (eds.) *Seasonally Dry Tropical Forests*, pp. 146–194. Cambridge, England: Cambridge University Press.

Latham RE and Ricklefs RE (1993) Continental comparisons of temperate-zone tree species diversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 294–314. Chicago: University of Chicago Press.

Lilleskov E, Fahey TJ, and Lovett GM (2001) Ectomycorrhizal fungal aboveground community change over an atmospheric nitrogen deposition gradient. *Ecological Applications* 11: 397–410.

MacArthur RH and MacArthur JW (1961) On bird species diversity. *Ecology* 42: 594–598.
Oliver CD and Larson BC (1990) *Forest Stand Dynamics*. New York: McGraw Hill.
Schelhas J and Greenberg R (1996) *Forest Patches in Tropical Landscapes*. Washington, DC: Island Press.
Whitmore TC (1998) *An Introduction to Tropical Rain Forests*, 2nd edn. Oxford, England: Oxford University Press.
Whittaker RH (1975) *Communities and Ecosystems*, 2nd edn. New York: Macmillan.

Fossil Record

Sean R Connolly, James Cook University, Townsville, QLD, Australia
Matthew A Kosnik, Macquarie University, North Ryde, NSW, Australia

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Sean R. Connolly, volume 3, pp 53–62, © 2001, Elsevier Inc.

Glossary

Correlation Determining the relative age of strata.
Extant Having some living representatives (i.e., not extinct) at a particular time.
Ma An abbreviation of the Latin "mega-annum" used by paleontologists to indicate dates millions of years before present.
Molecular phylogenetics The study of the history of evolutionary relationships according to similarities and differences among (usually extant) organisms in molecular characters (such as DNA).
Parametric statistics Statistical methods that assume data have certain characteristics, such as a particular kind of distribution (Poisson, normal, binomial, etc.).
Stage Normally the lowest-ranking unit in the hierarchy of terms for stratigraphic intervals that can be recognized on a global scale.
Stratigraphic interval All (and only those) rocks formed during a specific interval of geologic time.
Stratigraphic range The strata through which a taxon is known to have been extant.
Stratum (plural:strata) A unit of sedimentary rock distinguishable from other units on the basis of properties and/or relative position.
Taxon (plural:taxa) A group of organisms that share a name.

Obtaining reliable estimates of biological diversity in the fossil record is critical to addressing key questions about changes in biological diversity over long timescales and the causes and consequences of those patterns. On what scales is the history of life characterized by periods of stasis punctuated by relatively brief periods of rapid evolutionary change? On what scales is it a more gradual, continuous process? What are the causes of these changes? How often are major changes in biodiversity associated with major changes in the physical environment, such as extraterrestrial impacts? How have human impacts changed biodiversity? Unfortunately, the fossil record provides an incomplete picture of the history of biological diversity. This article reviews some important features of the fossil record that could cause the apparent history of biodiversity to differ quite markedly from its true history. It then introduces several methods that have been applied in an attempt to account for these features, it discusses advantages and disadvantages associated with each of them, and it reviews how their application in specific cases has altered earlier views about changes in diversity through time. Throughout this article, the terms "diversity," "biodiversity," and "biological diversity" will refer to taxonomic richness: the number of taxa present at a particular time or place. Although alternative metrics of diversity have been proposed, most paleobiologists have focused on taxonomic richness, at least when addressing regional or global patterns in biodiversity.

The Nature of the Fossil Record

The fossil record is an incomplete sample of the original living community, and, like most biological samples, some taxa are under-represented (e.g., soft-bodied worms from desert environments) relative to others (e.g., clams from soft-sediment marine habitats). Fossil assemblages are accumulations of organisms that share a common grave, but that may not have been alive at the same time or place. The ages of individual fossils are rarely determined, and, typically, fossil assemblages are assigned to a time interval that can range from thousands to millions of years. Any paleontological study larger than the local scale must place fossils from different locations onto a common time scale, which limits temporal resolution because the ages estimated for fossils include some degree of uncertainty, and coarser temporal resolution ensures that samples are put in the correct temporal sequence.

The Fossil Record is Incomplete

The fossil record is incomplete: Only a small fraction of individuals are fossilized; of those, very few are collected and identified. As a result, the biological diversity recorded in the fossil record is less than total diversity over the region and time interval from which it is a sample, and the observed stratigraphic ranges of taxa are shorter than their true ranges. This is because the first appearance of a taxon in the fossil record occurs sometime after it actually originated, unless it's very first representative was fossilized and subsequently sampled (a highly unlikely prospect). Similarly, the last appearance of a taxon occurs sometime before it actually went extinct. Biased estimates of stratigraphic ranges are important for the study of trends in biodiversity because such ranges have traditionally been used to estimate biodiversity. Thus, rather than estimating diversity as the number of taxa that actually occur in samples from a particular stratigraphic interval, diversity has been estimated as the number of taxa whose stratigraphic ranges encompass the interval. The use of

stratigraphic ranges to estimate diversity leads to underestimates of biodiversity, because true times of origination precede the first appearance of a taxon in the fossil record, and true times of extinction come after a taxon's last appearance in the fossil record.

The Quality of Preservation Varies

If incomplete sampling were the only problem with the fossil record, it would still be possible to accept relative trends through time at face value – that is, an increase in diversity in the fossil record would indicate a real increase in biological diversity, even if the true number of taxa could not be determined. Unfortunately, however, the degree to which the fossil record is incomplete varies in both space and time. Despite notable exceptions, such as the Burgess Shale or Solnhofen Limestone, most fossil assemblages record only hard parts. Thus, organisms with hard parts are much more likely to be preserved in the fossil record than those without them. Clams, for instance, have a better fossil record than nematodes. One consequence of this is obvious: the difference between apparent and true diversity tends to be greater for the latter group than for the former. However, estimates of diversity trends through time can be affected as well, if the proportion of taxa with body parts that are readily fossilized changes. Indeed, some have argued that the "Cambrian explosion" represents, not a diversification of multicellular life, but a rapid and extensive proliferation of hard parts. This hypothesis is supported by molecular phylogenetic studies predicting that the major animal phyla diverged long before the early Cambrian ($\sim$542 Ma). The timing of this divergence remains controversial; nevertheless, the very fact that it has received considerable attention illustrates just how profoundly sampling effects may influence the fossil record of major events in the history of life.

Differences in abundances among taxa can also affect their preservation in the fossil record. It is individuals (or parts of individuals) that are fossilized; thus, more abundant taxa are, on average, likely to have more complete fossil records than will rare taxa. Further, fossil diversity is likely to be higher for intervals during which abundances were higher, on average, than during other intervals. This, too, has important consequences for inferences about diversity trends. For instance, several major episodes of diversification, as recorded in the fossil record, coincide with geophysical changes that probably increased rates of nutrient supply to the biosphere. Some workers have argued that these geophysical changes were important causes of the coincident biological diversification. However, if increases in rates of nutrient supply also allowed taxa to sustain higher abundances, then increased probabilities of preservation during those intervals could be contributing to the increased diversity. Thus, inferences about the causes of trends in biodiversity, in addition to inferences about the trends themselves, can be influenced by sampling effects.

Differences in abundance or geographic range can also affect estimates of stratigraphic ranges. One of these effects is illustrated by the previous example: when comparing two intervals, more first appearances will be found in the interval with greater average abundance or geographic range size, even if there were, in fact, an equal number of originations during the two intervals. Assessing the severity of this effect is complicated by the fact that there are sound biological reasons for increases in abundance to facilitate originations of new taxa: to the evolutionary biologist, a correlation between abundance and rate of speciation may be precisely what is expected!

Yet another characteristic of organisms that can affect inferences about biodiversity trends is habitat. In particular, individuals in some habitats are more likely to become fossils than individuals in others. For instance, marine soft-bottom habitats are likely to provide more complete records of their inhabitants than will rocky shores. In the former case, individuals will much more readily be buried shortly after (or even before) they die. In the latter, wave action may render the remains of individuals unidentifiable before currents carry them to a location where they might be buried and preserved. Indeed, the fossil record of rocky shore communities is among the poorest in the marine realm. Again, the difficulties associated with these effects has led to disagreements about key events in the history of life. For instance, fossils record an explosive diversification of mammals in the early Tertiary, following the mass extinction that ended the Cretaceous period (best known for catastrophic extinctions among dinosaurs; 65.5 Ma). Most workers believe that this reflects the true pattern, at least qualitatively. However, others, using molecular phylogenetics, have argued that much of this diversification occurred in the Cretaceous, prior to this mass extinction. One explanation for this discrepancy has been that these early Cretaceous mammals occupied habitats with lower potential for preservation (such as forest interiors) than for their Tertiary descendants, who expanded into new habitats to occupy niches vacated by the extinction of dinosaurs. Whether this discrepancy between molecular and fossil data is primarily due to a poor fossil record for Cretaceous mammals or to biases and uncertainties in the molecular methods remains controversial.

As the discussion of the effects of habitat preference suggests, the probability of an individual being preserved in the fossil record depends, in part, on the sediments themselves. The existence and availability of fossiliferous rock, too, can have profound consequences on inferences about diversity patterns. For instance, in a classic paper, Raup (1976) noted a systematic increase in the volume of sedimentary rock through time. The implication was that the probability of individuals being preserved in the fossil record becomes progressively greater through time. Indeed, he presented a graph of sedimentary rock volume through time that looked strikingly similar to a graph showing biological diversity in the fossil record through time. After removing the effect of rock volume on diversity, he found no evidence for a long-term increase in species diversity through time. This line of research has been revived by Peters and others, particularly with regard to the effects of changes in rock volume on origination and extinction rates (Peters, 2006). Recent attempts to account for such effects have substantially reduced the magnitude of the diversity increase over time and changed our understanding of the timing of extinction events, relative to data that ignores such systematic changes in the fossil record.

The extent and quality of the fossil record varies over smaller timescales as well. Geophysical transitions can dramatically affect the probability that individuals will be preserved as fossils. When these transitions are global in scope, or occur in regions that receive disproportionately large amounts of attention from paleontologists, then apparent changes in diversity can result. Just such a possibility has complicated analysis of what was probably the most extensive mass extinction in the history of life, an event that marks the end of the Permian and the beginning of the Triassic. This event coincided with a major loss of marine benthic habitat (due to decreasing sea level). While it is certainly reasonable to suspect that biodiversity would decrease as available habitat decreases, it also reduces the probability of sampling taxa that remain, a phenomenon known as the Signor–Lipps effect. This and other sampling effects associated with characterizing this extinction event and its aftermath are extensively reviewed by Erwin (2006).

The Extent of Sampling Varies

Decisions paleontologists make can also introduce, or exaggerate, differences in the quality of the fossil record. For instance, North America and Europe have been more extensively sampled than many other regions of the world. This reflects, in part, the fact that many paleontologists live on these two continents. The logistical problems associated with fieldwork are generally less severe, and the cost of travel lower, when fieldwork is conducted near a home institution. In addition, fieldwork tends to be easier in regions with a well-developed infrastructure. As a result, more of the preserved biodiversity is actually sampled and recorded, and thus the apparent diversity of well-sampled regions is higher than that of poorly sampled regions. In the fossil record this is especially problematic as certain areas preferentially record specific time intervals and environments. For example, the marine fossil record of North America mainly represents tropical marine environments during the Paleozoic (251–542 Ma) but temperate marine environments in the Cenozoic (0–65 Ma).

Similarly, the questions that one can ask of paleontological data, and the degree of confidence that can be assigned to answers, depends in large measure on the quality and quantity of data that can be collected. For instance, more workers focus on the most recent era of time, the Cenozoic, than on the previous era, the Mesozoic (65–251 Ma), despite the fact that the latter era was longer in duration. This tends to exaggerate the effects of more abundant Cenozoic fossils. Likewise, more paleontologists study the Mesozoic than the earlier Paleozoic, again potentially amplifying the effects of greater availability of Mesozoic fossiliferous rock.

Correlation and Dating

Yet another factor affecting estimates of diversity patterns is correlation and dating. When fossils are collected, their age is generally inferred according to the particular strata in which they were found. Because direct age estimates are not available for most locations, age assignments are generally based on the presence of particular indicator taxa. That is, strata containing particular indicator taxa are assigned to a particular time interval (not a specific point in time) based on the presence of the indicator taxa in other dated strata. Since new taxa originate at a point in space as well as time, there is some uncertainty due to the rate at which the indicator taxa increase their geographical ranges. In practice, paleontologists attempt to minimize this problem by using multiple indicator taxa. They also select indicator taxa that are likely to have high probabilities of preservation in the fossil record and whose spread is likely to have been rapid, such as planktonic foraminifera.

Because data are assigned to intervals rather than points in time, the number of taxa found in a particular interval may be greater than the number of those taxa that were actually extant at any particular point in time during that interval (assuming that extinctions or originations occurred during the interval). The longer the interval lasts, and the higher the rates of origination and extinction, the more apparent diversity is inflated. Because it is the proportional difference in length that is critical here, this problem becomes more severe as intervals are more finely divided. The shorter the interval, the greater the proportional uncertainty associated with the estimated ages of the intervals' boundaries. For instance, overestimating the duration of a 10 million year interval by 1 My inflates the estimated duration of that interval by 10%. However, an overestimate of the same magnitude on the duration of a 3-My interval represents a 33% inflation. Fortunately improved dating techniques and concerted efforts by geologists and paleontologists have led to ongoing improvements in the geologic time scale (Gradstein *et al.*, 2005).

If one is interested in estimating global diversity or macroevolutionary rates, then the temporal resolution of investigation is limited by the ability to correlate stratigraphic intervals. At the global scale temporal resolution is typically limited to ~90 internationally recognized stages which average ~6 Myr long, but can vary in length by more than a factor of 10. If the origination and extinction of taxa occurred at roughly constant rates, then variation in the length of international stages could be responsible for a 10-fold difference in apparent diversity! The Paleobiology Database (http://paleodb.org/) has attempted to address this problem by grouping shorter stages and creating 48 time intervals, each approximately 11 My long. These intervals are more uniform than stages, but they can still vary by a factor of three due to difficulties subdividing a few time intervals.

Statistical Estimation of Diversity Trends

Until recently, nearly all large-scale studies of paleontological diversity took the first and last appearances of taxa in the fossil record as a true record of its times of origination and extinction. In taxa with relatively poor sampling, but widely-accepted phylogenetic trees, such as dinosaurs, first occurrence times have been extended based on the estimated times of origin obtained from the phylogenetic trees. However, for the reasons described in the previous section, this practice is now widely recognized as inadequate. More recent compilations of data, initially by individual researchers but increasingly by large teams, such as the Paleobiology Database, are beginning

to make large-scale analyses possible, by recording the spatial and temporal pattern of fossil occurrences. That is, rather than simply a time of first and last appearance, there is a record of the occurrences of a particular taxon in the fossil record over time and space. Such occurrence patterns contain valuable information about the completeness of the fossil record. For instance, consider the taxa that are found both before and after, but not within, a particular time interval. While not sampled these taxa must have been present during that interval. The proportion of taxa that do not have fossil occurrences in that interval but must have occurred during the interval, provides information about the completeness of the fossil record during that interval. The accessibility of large datasets has led to substantial growth in the use of statistical methods that seek to use occurrences in various ways to obtain better estimates of diversity, or at least of changes in diversity.

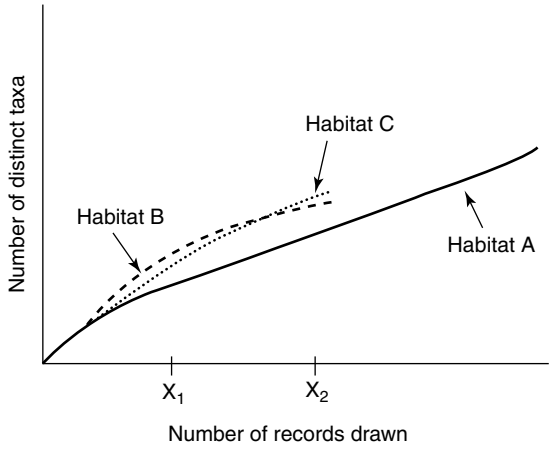


Figure 1 Hypothetical rarefaction curves for three samples representing three different habitats, designated A (solid line), B (dashed line), and C (dotted line). Note that habitat A, whose sample contains the most taxa in its entirety (note that its rarefaction curve ends at a higher diversity than the others), is not the most diverse habitat when its diversity is compared with the others for a particular sample size (X_1 on the horizontal axis). Also note that the rarefaction curves for samples B and C cross. Thus, habitat B is more diverse when there are X_1 records in the sample, but habitat C is more diverse when there are X_2 records.

Sampling Standardization

One way to control for sampling effects in making inferences about diversity trajectories is to compare taxonomic richness among locations or strata with samples that have been standardized to be equivalent in extent, according to some reasonable measure. The simplest standardization technique, called rarefaction, works as follows: A sample is collected from each of a set of habitats (for instance, a certain volume of sediment is obtained) using an identical sampling scheme. All of the individuals in each sample are identified, and their taxonomic identity is recorded. The next steps can be visualized by imagining placing all of the records for a particular sample in a bowl, stirring them, and then randomly selecting records from that bowl until none are left. Each time a record is picked, the taxonomic identity of that record is noted. From this sequence of random draws, one constructs a graph with the number of records on the horizontal axis and the number of distinct taxa on the vertical axis. What is plotted on that graph are the results of that sequence of record selection. This curve must intersect the points (0, 0) and (1, 1). That is, before the experiment begins, 0 records have been picked from the bowl, so 0 distinct taxa have been selected. When one record is chosen, exactly one distinct taxon has been found. When two records have been chosen, either one or two distinct taxa have been found. (If the second record has the same taxonomic identity as the first, the next point is (2, 1); if it has a different identity, the next point is (2, 2)). Once all of the points have been plotted, a curve is fit to them: this is the rarefaction curve. This process is repeated for each sample collected, so that each sample has a rarefaction curve associated with it. Biodiversity in the different samples is then compared by examining the number of distinct taxa encountered for a given number of records selected. That number must be equal to or less than the number of records in the sample with the fewest records. This comparison is illustrated in **Figure 1**.

Figure 1 also illustrates something else: Rarefaction curves sometimes cross. This means that the rank order of diversities of a set of samples can change, depending on sample size. These relative changes depend upon the relative frequencies of occurrence of different taxa. For instance, if taxa have similar sampling probabilities (e.g., because they have similar abundances), then the rarefaction curve will tend to quickly accumulate all available taxa. However, if occurrences are unevenly distributed among taxa, then initially diversity will increase very slowly, as the dominant species are sampled, but very rare species will still be occurring for the first time even as the standardized sample approaches the size of the original sample. One might therefore ask what rarefied diversity means, and, further, what it means to say that one site has a greater rarefied diversity than another. Hurlbert (1971) proposed that rarefied diversity measures the number of distinct taxa encountered, on average, by an individual organism in a particular habitat over the course of a particular number of encounters with other individuals (provided that individuals are not encountered multiple times). The fact that one habitat has a greater rarefied diversity than another for a given number of occurrences of individuals does not mean that that habitat actually has a greater taxonomic richness (i.e., contains more taxa) than another.

Although rarefaction can be applied to the fossil record in the manner just described, it is often applied somewhat differently. When comparing diversities over very large scales (e.g., global diversity during two stratigraphic intervals), a sample enumerating all of the individuals in those intervals will not be available. Rather, data will consist of a number of distinct samples for each of the intervals being compared. Further, while a list of the taxa found in each sample may be available, information on the abundances of those taxa may not. Such data have been rarefied in two different ways. In one approach (hereafter called rarefaction by occurrence), a rarefaction curve is constructed by randomly selecting occurrences of taxa from the set of available samples. The rarefaction curve is then a plot of the number of distinct taxa sampled against the number of occurrences of taxa drawn

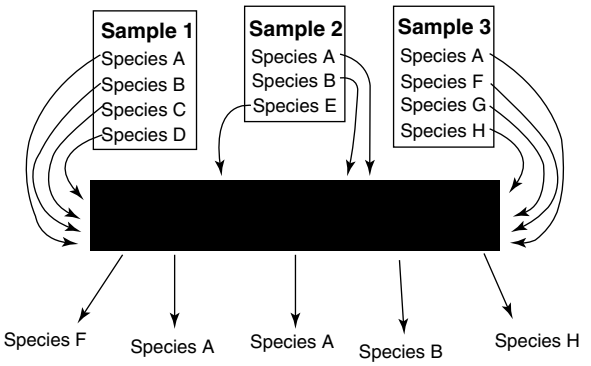


Figure 2 A graphical illustration of rarefaction by occurrence. Occurrences of a taxa within samples are treated independently, and these occurrences are drawn individually at random from the set of occurrences in all samples. This can be visualized as adding each occurrence individually (regardless of the sample from which it comes) into a pool (arrows into the black box), then randomly choosing these occurrences from that pool (arrows out of the black box). In this figure, the rarefied diversity after five records have been drawn is four distinct species.

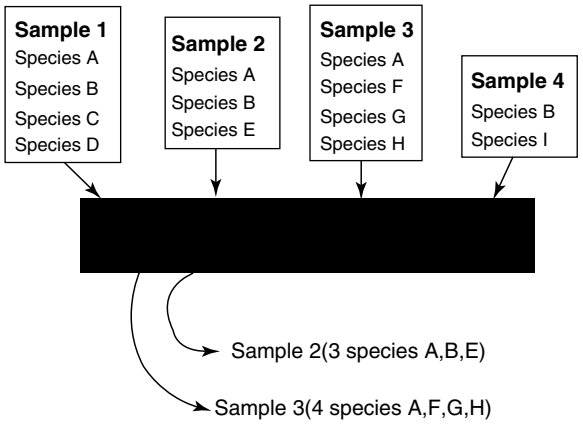


Figure 3 A graphical illustration of rarefaction by list. In this case, an entire paleontological sample is drawn at random from the population of samples that constitute the data set, and rarefied diversity is the total number of unique taxa contained in those lists for a given number of samples drawn. In this figure, the rarefied diversity after two samples have been chosen at random is six species (A, B, E, F, G, and H).

from the available samples. This is illustrated in **Figure 2**. Diversities are compared by constructing these rarefaction curves for each interval. An alternative approach (hereafter called rarefaction by list) involves randomly selecting the entire list of taxa present in a given sample as a unit, rather than selecting occurrences of taxa within those samples. In this case, the rarefaction curve for a particular interval consists of the number of distinct taxa detected plotted as a function of the number of samples drawn. Note that in this instance, the curve is not constrained to pass through the point (1, 1), since several taxa may appear in a single sample. However, it still must pass through the point (0, 0). This method is illustrated in **Figure 3**. As with rarefaction by occurrence, rarefaction curves can cross in this case also, depending on the pattern of alpha and beta diversity (i.e., whether diversity is due mainly to there being many taxa present in individual locations, or due to different locations having very different suites of taxa).

In a quest to exploit the information contained at multiple levels (e.g., published references, lists within references, occurrences within lists), paleobiologists have explored a variety of ways of sampling occurrences while ensuring that those occurrences are not dominated by a small number of very extensive lists, at the expense of the greater geographical dispersion of samples that one would get using rarefaction by list, or by published reference (see Alroy *et al.*, 2008, for an example). For instance, the probability of sampling occurrences from larger lists may be down-weighted, relative to those from smaller lists, thus increasing the probability that a standardized sample of occurrences is drawn over a broader geographical area. A variety of such algorithms have been tried, which essentially weight alpha, beta, and gamma (i.e., local, among-habitat, and among-region) diversity differently. For example, one of the more commonly used methods assumes that alpha diversity does not change through time. Most recently, more complicated approaches have been proposed. For instance, in "quorum subsampling," each species is assigned a value based on the frequency of its occurrences in the original data. As new taxa appear in the subsample, a cumulative sum of these frequencies is updated until some specified level of coverage is achieved (Alroy, 2010).

Paleontologists have typically been very cautious in their interpretation of rarefied diversities. There is a reason for this caution: the biological meaning of a standardized diversity measurement (or a relationship between two standardized diversity measures) is, in most cases, unclear. The interpretation offered by Hurlburt (described earlier) is appropriate only when (1) the sampling protocol described in the first paragraph of this section has been followed or (2) each occurrence of a taxon is a truly independent, random sample from the habitat about which one wishes to make inferences. Generally, neither is true for large-scale paleobiological diversity estimates, because paleontologists use a wide variety of sampling methods and sampling effort has not been distributed randomly. However, there have been some recent advances in this area, particularly recent work initiated by Bush (e.g., Bush *et al.*, 2004) that attempts to determine, from first principles, how different sampling standardization methods respond to different patterns of alpha, beta, and gamma diversity. Even where such insights are unavailable, however, rarefaction can provide useful information. In particular, discrepancies between rarefied and overall diversity patterns indicate where marked increases or decreases in apparent diversity are likely to be artifacts of sampling.

Nonparametric Estimates

In contrast to sampling standardization, a variety of methods attempt to estimate the actual underlying taxonomic richness in a particular interval have been explored, although these techniques are far less prevalent than sampling standardization. Among the most common (although typically not applied at very large spatial scales) are nonparametric methods that utilize the frequency distribution of fossil occurrences for a particular interval of time. That is, the raw material for

the diversity estimate is the number of taxa occurring in only one sample, f_1, the number occurring in two samples, f_2, and so on, as well as the total number of occurrences in all samples.

Burnham and Overton (1979) utilized a statistical approach known as the jackknife. The mathematics of the derivations are too complex to review here, but with this approach they obtained a series of possible estimators. The simplest of these utilizes only the number of samples and the number of taxa occurring in only one sample:

$$\hat{D}_1 = D_{\text{obs}} + f_1 \frac{k-1}{k}, \quad [1]$$

where $\hat{D}_1$ is the first-order jackknife estimate of diversity, D_{obs} is the number of distinct taxa appearing in the sample, f_1 is the number of taxa occurring in exactly one sample, and k is the number of samples. They developed additional estimators by incorporating the number of taxa occurring in more than one sample (for instance, their second-order jackknife uses the number of taxa appearing in exactly two samples, as well as the number appearing in just one). Nichols and Pollock (1983) applied these models to a molluscan dataset from the Middle Miocene (16–12 Ma) of Denmark. They found that, even under a relatively intensive sampling regime, sampled diversity was as much as 30% lower than estimated diversity.

An alternative model, proposed by Chao (1987), uses the number of taxa occurring in either exactly one or exactly two samples:

$$S_2 = D_{\text{obs}} + \frac{f_1^2}{2f_2} \quad [2]$$

Wing and DiMichele (1995) used this estimator, usually termed "Chao-2," to compare regional vegetation diversity in the late Paleozoic and early Cenozoic for North American river and delta floodplains. Somewhat surprisingly, they found similar biodiversity levels during the two periods for these regions, despite the markedly higher apparent diversities for the latter interval at the global level. The same formula has also been applied to estimate local diversity, but with the number of individuals representing each taxon in a single sample used rather than the frequency of occurrences in a set of multiple samples. That is, f_1 is the number of taxa represented by only one individual in a sample, f_2 is the number represented by two individuals, and so on. Wing and DiMichele (1995) used this approach to examine local vegetation biodiversities in the Paleozoic and Cenozoic floodplain data set discussed previously. They found that, on average, local diversity in the late Paleozoic was similar to local diversity in the early Cenozoic. This was consistent with their findings for regional diversity using the Chao-2 estimator. However, they did find greater variation in diversity levels among sites in the Cenozoic; in particular, the most diverse Cenozoic sites were much more diverse than the most diverse Paleozoic sites.

Yet another approach, developed by Chao and Lee (1992), utilizes the entire frequency distribution of occurrences in a set of samples, rather than just the number occurring in only one or a few samples. Wang and Dodson (2006) used this method to estimate that vertebrate paleontologists have discovered and named only ~29% of an estimated total of ~1850 nonavian dinosaur genera. Anderson *et al.* (1996) estimated South African late Triassic (228–200 Ma) plant and insect diversities with this method.

These approaches have some advantages over sampling standardization. Perhaps foremost of these is that it is clear what is being estimated – the number of taxa in the sampling universe from which the occurrences have come. Secondly, the estimators are explicitly derived from a set of statistical assumptions, so it is possible to quantitatively assess how well those assumptions are met. Unfortunately, some of these assumptions are violated in fossil data, often quite severely, so they are likely to produce biased estimates in many applications. For instance, the jackknife and Chao estimators assume that each taxon has an equal probability of occurring in each sample. That is, taxon A may have a different probability of being present in a sample than taxon B, but that taxon-specific probability is the same for every sample. This assumption may be violated if samples differ in extent or quality. However, even if great care is taken to minimize this problem, the assumption may still be violated. For instance, if relative abundances of particular taxa differed among sampling locations, then the associated relative probabilities of sampling may differ accordingly.

Parametric Estimates

Nonparametric methods seek to make minimal assumptions about the true underlying probability distribution of abundance or occurrences. However, parametric approaches assume that these distributions follow a particular mathematical form, they estimate the parameters of this distribution for a particular sample or set of samples (hence the term parametric), and they then use those parameter estimates to infer the fraction of species predicted not to be sampled at all. Various distributions have been proposed as good fits to fossil data, and it has been suggested that the best fit distributions may even change through time reflecting ecological changes in the biosphere. In one such case, such a fit has been used to estimate unsampled diversity. The particular probability distribution used was the Generalized Inverse Gaussian-Poisson (GIGP) distribution. Anderson *et al.* (1996) fitted this model to plant and insect data mentioned in the previous section. They used these estimates to argue that late Triassic plant and insect diversity in the sampled habitats was comparable to those of the present day, again in contrast to apparent global diversity patterns, which record increasing biodiversity levels through time.

The primary limitation of this approach is that it assumes that the underlying frequency distribution of occurrences follows a particular statistical distribution. As a general rule, parametric methods are more precise than nonparametric analogs when the assumptions about the underlying distribution are met. When they are not, however, the estimates can be very inaccurate. Another limitation, shared by the nonparametric estimates, is that the uncertainty associated with the estimated diversity increases as the proportion of unsampled taxa increases. That is, when the probability of a taxon appearing in a sample is low on average (or when

there are few samples), the estimates are especially prone to error. For this reason, the fact that a method works well as an estimate of present-day diversity does not necessarily mean that it will work well when applied to the fossil record, except perhaps where that record is unusually complete.

In ecology, parametric models have been used to infer local and regional diversity from observed distributions of occurrences or abundance, sometimes with good results. One potential problem with the application of such methods to the fossil record is that differences among taxa in their probability of entering the fossil record, and subsequent loss of fossiliferous rock (at rates that may vary among regions and through time), may make distributions of fossil occurrences very different from distributions of occurrences of living organisms. However, to date, models that are commonly used by ecologists, such as the lognormal and the gamma distribution, have yet to be applied to the fossil record.

Turnover-Based Estimates

Paleontologists typically break taxa down into five categories relative to their occurrence before, in, and after the time interval in question: (1) those sampled before, in and after; (2) those sampled before and after (but not in); (3) those sampled before and in (but not after); (4) those sampled in and after (but not before); (5) those sampled only in the interval (not before or after). In traditional stratigraphic range based estimates of diversity, all of these taxa contribute to the diversity in the time interval, although taxa only occurring in a single time interval are often removed from paleontological analyses, because such "singletons" are particularly sensitive to between-interval variation in sampling intensity. An alternative, known as "boundary crosser" estimates of diversity, count diversity at interval boundaries, rather than within intervals. For instance, diversity at the bottom of an interval includes all of the taxa crossing the bottom boundary (found sometime before and after the boundary, including those not actually sampled in the interval itself). Both traditional and boundary-crosser approaches have been widely used in paleontological analyses, and several other counting methods have been used occasionally as well.

It has been suggested that better estimates of diversity (along with rates of origination and extinction) can be obtained by only counting taxa that occur in two consecutive intervals (two timers), three consecutive intervals (three timers), or the first and third but not the second interval (part timers). Although *ad hoc*, these data selection criteria appear to be relatively robust to the biases typically associated with fossil data. Thus, while these counting methods underestimate diversity even more than standard range-through approaches, they are likely to be more robust as measures of relative diversity across multiple time intervals.

In the mid-1980s, Nichols and Pollock (1983) proposed that the information contained in presence–absence sequences in the fossil record should be used more systematically to estimate diversity. Specifically, they proposed that "open-population" statistical methods used to estimate abundances (along with birth and death rates) from capture–recapture data in population biology could be adapted, by analogy, to estimate taxonomic diversity (along with origination and extinction rates – births and deaths of taxa) in the fossil record. When one conducts an "open-population" capture–recapture study, individual organisms are captured during discrete sampling occasions over a period of time. Each individual captured is given a unique mark, so that its capture history can be constructed. That is, if one constructs a matrix, the columns j of which represent different sampling occasions and the rows i of which correspond to each individual captured at least once, then each element a_{ij} in the matrix will be either 1 or 0, indicating whether individual i was captured on occasion j. Similarly, fossil data are collected during discrete sampling periods, and one can simply list the taxa sampled in each interval. In this case, the matrix elements a_{ij} are 1 or 0 according to whether taxon i was found in interval j. In this context, taxonomic diversity (the total number of distinct taxa) is analogous to population size (the total number of distinct individuals). Thus, each species' sampling history is a sequence of presences and absences. This set of sampling histories can be used to estimate probabilities of sampling, origination, and extinction, and to estimate total taxonomic diversity, using standard statistical theory.

Like nonparametric and parametric diversity estimates that utilize replicate samples within an interval of time, these turnover-based methods make assumptions that are rarely fully met by fossil data, and it is not always clear which results are robust to violations of assumptions and which are not. However, a major advantage of these approaches over sampling standardization approaches is that it is possible to test for, and estimate the magnitude of, violations of those assumptions. When Nichols applied a turnover-based estimate to late Eocene (37–34 Ma) mammals from the Big Horn Basin, Wyoming, goodness of fit statistics indicated rejection of the model. However, when applied to Middle Miocene molluscan diversity, the model provided an adequate fit. Similarly, goodness of fit testing was used by Connolly and Miller (2001) to show that violation of assumptions was unlikely to have severe effects on estimates of origination and extinction (in other words, that the uncertainty in the estimates was likely to be fairly large, relative to any biases due to assumption violations).

Some of the most questionable assumptions in turnover-based methods involve the assumption that probabilities of origination, extinction, and sampling are constant for all taxa that are grouped together in an analysis. To minimize the potential biases associated with violations of these assumptions, researchers have adopted several approaches. One is to divide taxa into subgroups believed to have different sampling or evolutionary turnover rates, and estimate parameters for them separately. An alternative is to relax that assumption, and use characteristics of individual taxa believed to be positively or negatively correlated with sampling, origination, or extinction rates, and to estimate the parameters of that relationship. For instance, Connolly and Miller (2001) allowed sampling probabilities to differ among Ordovician (488–461 Ma) marine invertebrate genera according to differences in the apparent geographical and ecological distribution of those genera.

Conclusions

As we have seen, many features of the fossil record make assessing diversity trends difficult. Some of these difficulties can be eased, at least in principle. For instance, unequal distribution of sampling effort by paleontologists can be reduced by emphasizing undersampled regions and strata in future fieldwork. Similarly, the expanding palette of tools for dating fossils and their surrounding sediments, their increasing precision, and the development of more robust statistical methods for chronostratigraphic correlation is progressively improving the fossil record's temporal accuracy and precision. However, some problems are less tractable. For instance, fossiliferous rock is progressively lost as it ages. Thus, the fossil records of large regions (and thus the record of biodiversity in the associated habitats) may simply not exist. This problem is particularly acute for older time intervals. Similarly, organisms in some habitats are simply less likely to be preserved than organisms in others, and those habitats will thus have poorer records of their biodiversity history.

Several approaches have been used to account for these sampling effects in the estimation of diversity or (in the case of rarefaction) to minimize the effects of variation in sampling on estimates of relative diversity. All of these approaches were originally designed to estimate local diversity; thus, their application to regional and global patterns is problematic, and workers have interpreted results cautiously. Nevertheless, these results can sometimes be used to eliminate some explanations for particular biodiversity trends. For instance, Raup's use of rarefaction confirmed that the increased diversity of echinoids through time was not solely due to the greater number of paleontological samples available for younger strata. When the results of these approaches differ from one another, or from diversity trends apparent from the fossil record in its entirety, however, things become complicated. Do the estimates indeed account for sampling effects and reveal true diversity trends, or does violation of model assumptions render the estimated trends even less reliable than the uncorrected diversity trends they are intended to improve?

The future of diversity estimation will no doubt involve considerable effort on several fronts. One promising approach is to apply several methods, then identify biodiversity trends that are robust to these alternative methods. Another is to investigate directly how different estimates are biased when particular assumptions are violated, then devise means of minimizing these biases. Diversity estimation methods that have not yet been applied to the fossil record may be incorporated into the paleobiological research program, and new estimates will undoubtedly be forthcoming as well. Nonparametric methods based on the frequency distribution of occurrences have been identified as a promising area for further progress by many biostatisticians, so increasingly robust metrics are likely to become available. Parametric estimates of diversity may also come into greater use, as ecologists increase their understanding of the shapes of frequency distributions of abundance and occurrence over different scales of time and space. The development of computationally intensive methods of model fitting may allow the use of "mixture models," which would allow sampling, origination, and extinction rates to vary randomly among taxa in turnover-based statistical models, rather than being assumed constant or specified to vary only in response to particular explanatory variables. In recent decades, quantitative approaches have been rapidly growing in popularity and sophistication among paleobiologists. This movement is likely to produce an increasingly clear picture of the history of biodiversity.

See also: Biodiversity, Evolution and. Biodiversity, Origin of. Extinction in the Fossil Record. Mass Extinctions, Notable Examples of. Measurement and Analysis of Biodiversity. Paleoecology. Species–Area Relationships

References

Alroy J, Aberhan M, Bottjer DJ, *et al.* (2008) Phanerozoic trends in the global diversity of marine invertebrates. *Science* 321: 97–100.

Alroy J (2010) The shifting balance of diversity among major marine animal groups. *Science* 329: 1191–1195.

Anderson J, Anderson H, Fatti P, and Sichel H (1996) The Triassic explosion(?): A statistical model for extrapolating biodiversity based on the terrestrial Molteno formation. *Paleobiology* 22: 318–328.

Burnham KP and Overton WS (1979) Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60: 927–936.

Bush AM, Markey MJ, and Marshall CR (2004) Removing bias from diversity curves: The effects of spatially organized biodiversity on sampling-standardization. *Paleobiology* 30: 666–686.

Chao A (1987) Estimating population size for capture–recapture data with unequal catchability. *Biometrics* 43: 783–791.

Chao A and Lee S-M (1992) Estimating the number of classes via sample coverage. *Journal of the American Statistical Association* 87: 210–217.

Connolly SR and Miller AI (2001) Global Ordovician faunal transitions in the marine benthos: Proximate causes. *Paleobiology* 27: 779–795.

Erwin DH (2006) *Extinction: How Life on Earth Nearly Ended 250 Million Years Ago*. Princeton: Princeton University Press.

Gradstein FM, Ogg JG, and Smith AG (eds.) (2005) *A Geologic Time Scale 2004* Cambridge: Cambridge University Press.

Hurlbert SH (1971) The nonconcept of species diversity: A critique and alternative parameters. *Ecology* 52: 577–586.

Nichols JD and Pollock KH (1983) Taxonomic diversity, extinction rates, and speciation rates from fossil data using capture–recapture models. *Paleobiology* 9: 150–163.

Peters SE (2006) Genus extinction, origination, and the durations of sedimentary hiatuses. *Paleobiology* 32: 387–407.

Raup DM (1976) Species diversity in the Phanerozoic: An interpretation. *Palebiology* 2: 289–297.

Wang SC and Dodson P (2006) Estimating the diversity of dinosaurs. *Proceedings of the National Academy of Sciences (USA)* 103: 13601–13605.

Wing SL and DiMichele WA (1995) Conflict between local and global changes in plant diversity through geological time. *Palaios* 10: 551–564.

Framework for Assessment and Monitoring of Biodiversity

Francisco Dallmeier, Robert C Szaro, and Alfonso Alonso, Smithsonian Conservation Biology Institute, Washington, DC, USA
James Comiskey, National Park Service, Fredericksburg, VA, USA
Ann Henderson, National Geographic Channel, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Ann Henderson, James Comiskey, Francisco Dallmeier, Alfonso Alonso, volume 1, pp 1–11, © 2001, Elsevier Inc.

Glossary

Adaptive management A systematic, cyclical process for continually improving management policies and practices based on lessons learned from operational programs.

Analysis A process of exploring and estimating parameters of data, gathered with statistical procedures, to test previously established hypotheses.

Biodiversity assessment The identification and classification of species, habitats, and communities, within a given area or region. The overall purpose is to provide information needed to evaluate whether management is necessary to conserve biological diversity. Assessments also provide data and information that can be applied for programs of scientific inquiry.

Indicators Species or communities that enable an evaluation of environmental conditions and detect changes. Indicator species are normally surrogates of other species in the area of interest and are usually sensitive to environmental change. Environmental variables may also be used as indicators.

Monitoring The repeated collection and analysis of observations and measurements to evaluate changes in populations of species and environmental conditions. Monitoring also helps in assessing progress toward meeting a management objective. Monitoring can serve as a warning system, alerting managers that changes in biodiversity may require changes in biodiversity management regimes to ensure protection of biological resources.

Sampling Refers to the process of data collection. A protocol must be developed to direct sampling that includes sampling objectives, parameter to be measured, sample unit determination, permanent or temporary sampling, number of samples, and method of sample selection.

Uncertainty Describes the condition whereby managers have a lack of data, or unreliable, error-prone data of biodiversity that prevents them from defining the best course of management action. Uncertainty can be overcome through a well-designed adaptive management program.

Why are we interested in assessing and monitoring biodiversity? Accurate information on the status of components and processes of ecosystems is essential for the long-term maintenance of biodiversity, and its sustainable use. The selection of biotic indicators observable or linked across different spatial and temporal scales, are important to detect trends that can result in management practices (Feld *et al.*, 2009). Biodiversity assessment across spatial and temporal scales, at different levels of biological organization, and in terms of various attributes, processes, and functions of biodiversity, is critical to our understanding of ecosystem functioning and formulation of management objectives (Ash, 2009). In addition, most ecosystems are under severe pressure, either from increasing primary and secondary human impacts, or from climatic changes (Dallmeier *et al.*, 2010).

Biodiversity monitoring and assessment programs (BMAP) provide a process which collects and reports scientific information to be used for natural resource management. The first task in developing the BMAP program is to define the objectives. Then, a baseline biological assessment is often necessary to determine what kinds of species and habitats need to be considered as part of the program. After the assessment has been completed, monitoring the selected indicators to answer questions using the scientific method will provide biodiversity trends in the ecosystem. Monitoring activities leads to insights on issues such as those requiring management strategies, for example, the need for protection of threatened or endangered species, or the eradication of nonnative invasive species. Monitoring is also used to assess progress made toward meeting a management objective and is an essential component of a successful adaptive management program.

Introduction

Natural resource managers and other decision makers working at the conservation and development interface confront challenging environmental issues that require science-based understanding of the status and trends of habitats and species. The primary role of the BMAP is to improve natural resource management and conservation and to build institutional knowledge. BMAPs are also designed to determine status and trend monitoring of species and habitats or baseline monitoring that is not focused on a particular management action but aims at increasing our understanding of natural processes and also acts as an early warning system. Reports and recommendations to managers based on the sampling design and data collection are the principal products of the BMAP. Sound data management is necessary to provide quality

information to long-term programs. Sound BMAP programs provide information that managers need to make decisions either alerted to an impending problem, for example, early detection of invasive exotic species or pathogen, or in response to management actions that they are implementing, for example, effectiveness of invasive exotic plant management efforts, or effectiveness of restoration of degraded habitats.

BMAP Spatial and Temporal Scales

Effective BMAP frameworks require successful integration of different monitoring components to produce information that is more useful than the individual components. This integration requires spatial and temporal considerations at the ecosystem and landscape levels.

Monitoring ecosystem structure and function at the landscape level requires careful consideration for the selection of appropriate indicators. The monitoring strategy will utilize several individual indicators that collectively will measure the status and changes within the entire ecosystem. An ecosystem can be very large and most monitoring programs, depending on their design, can only make inferences to the areas that are considered as part of the sampling strategy. Results may be representative of larger areas but must be used cautiously. Often a suite of indicators is selected at various hierarchical levels of ecological organization, such as community, population, genetic levels (Noss, 1990) or from species or habitats of conservation concern and of value to society. Integration of spatial information requires the establishment of measurements made at different areas and scales within the study areas. This strategy requires understanding of spatial sampling design frameworks.

Timelines for sampling indicators also need to be considered based on the characteristics and temporal variations that these indicators may show. For example, monitoring forest dynamics and regeneration often requires 5-year measurements of tree size classes and their distribution in permanent plots, whereas water quality may require continuous measurements. The integration of monitoring programmatic activities requires coordination and communication among BMAP management organizations and other stakeholders to promote wide participation in the use of the resulting information.

BMAP Goals and Objectives

The purpose of BMAP is to develop science-based information on the current status and trends in the composition, structure, and function of habitats and species, and to link that information to management practices and policy decisions. Scientifically based information from BMAP increases the confidence level of a manager in predicting the outcome of his decisions and to operate more effectively. The overall goals of the BMAP are site specific and related to the issues and concerns addressed by the program.

Overall goals and objectives of BMAP include:

- Assess species and habitats to determine presence and relative abundance.
- Select the species and habitats that will be used as surrogates for the area under study.
- Establish the status and trends of selected species and habitat indicators to allow managers to work more effectively to minimize potential impacts on the resource and provide adaptive management strategies.
- Provide early warnings on the condition of species and habitats to help management in developing mitigation strategies.
- Monitor selected species and habitats to understand their natural and human-induced dynamics and conditions. First assessment should serve as a baseline reference to determine desired state.
- Integrate the BMAP in an adaptive management cycle for planning, management, and decision making.
- Provide data-driven information to understand the dynamics of species and habitats through comparison with similar natural and human-managed systems.
- Generate information to meet regional, national, and international environmental laws and regulations.
- Provide timelines and deliverables toward achieving the specific goals of the BMAP.
- Disseminate information generated by the program to multiple stakeholders to engage them in management and conservation efforts.

Based on these goals, it is important to develop a conceptual ecosystem model to define the study areas, identify geographical issues, habitats and species, and describe the ecosystem structure. Designing the assessment and monitoring programs requires the selection of indicators to monitor, and the development of the monitoring protocol. The final components are the implementation of the protocols, synthesizing information, and the application of recommendations as part of the adaptive management process (**Figure 1**).

The BMAP is based on the scope of the project, specifications to meet the established goals and contain variables that can be measured so that the information collected addresses questions and hypothesis. Managing objectives provides the appropriate focus to address the change/trend objective, for example, to increase the density of Species A by 30% or decrease the frequency of invasive species by 30% at a particular location (Elzinga *et al.*, 1998).

Sampling strategy is based on the sampling objectives of the BMAP that will address the monitoring questions. These questions may be formulated as part of the process and lead to other projects using experimental design to test hypotheses. They include levels of precision and magnitude of variation expected to be detected by the BMAP (see the BMAP protocol design section). The BMAP objectives provide additional details about the monitoring program or sampling protocol and the limits of the program. For example, are the objectives measureable and achievable? Is the monitoring and sampling location specific enough? Are the species and habitats being monitored, clearly selected, and specified? Will the data and information user be able to anticipate what the data and information will look like? **Table 1** provides some specific examples of monitoring objectives for BMAP sampling protocols.

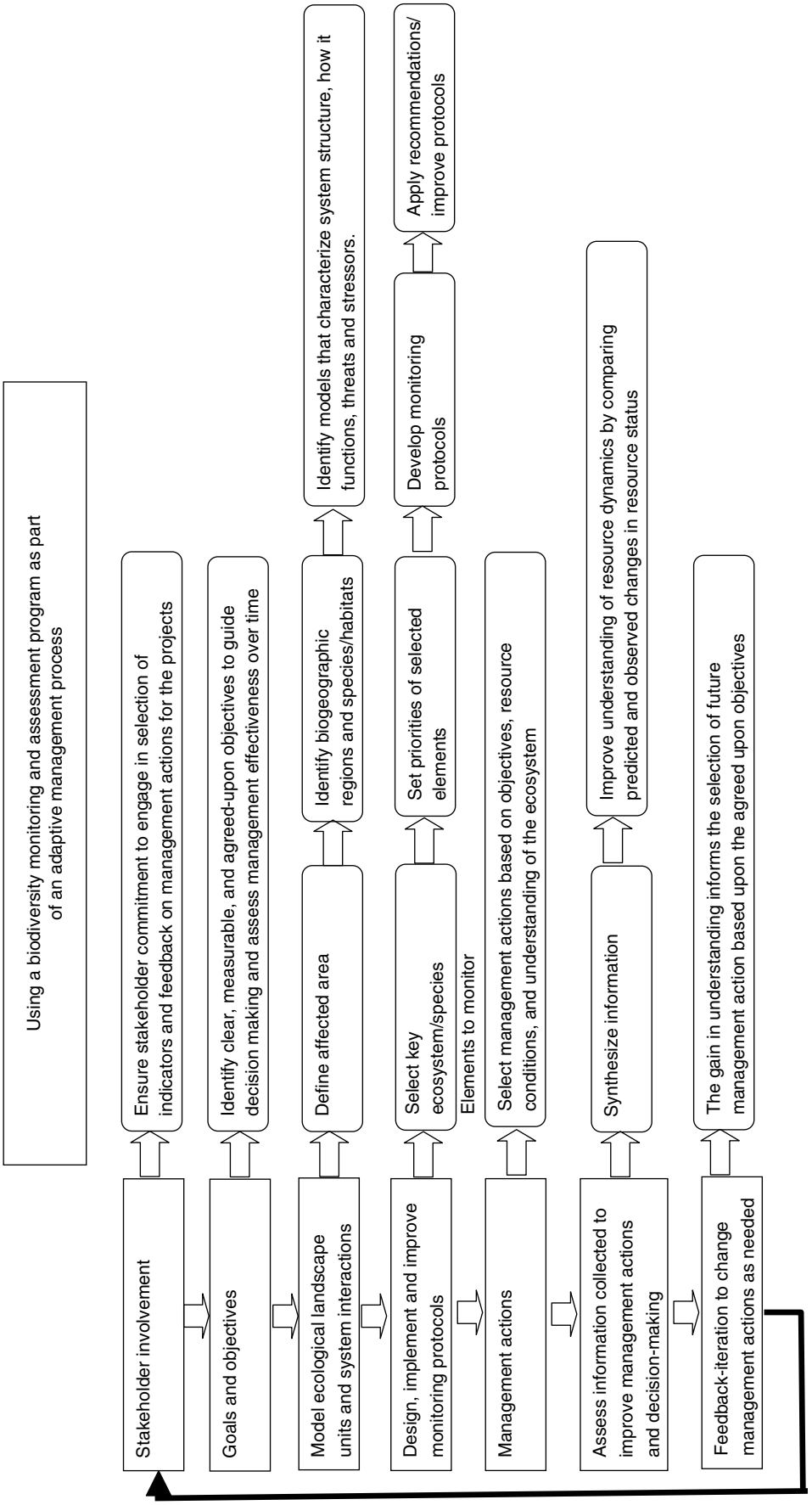


Figure 1 Biodiversity monitoring and assessment program as part of an adaptive management process.

Table 1 Examples of BMAP monitoring objectives

BMAP protocols	*Measurable objectives*
Wetland vegetation	1. Determine if species composition and abundance is different in areas up-stream and down-stream from park road. 2. Determine if species composition and abundance is the same in areas around the road where road improvements were made.
Fishes	1. Determine fish species diversity, composition and abundance in high altitude freshwater lagoons. 2. Associate these parameters to water quality in areas close and far from access roads. 3. Determine trends in the distribution and abundance of native species of fishes in areas where exotic species are introduced.
Birds	1. Determine annual status and trends in territory occupancy of cloud forest birds in areas near and far from the human development. 2. Determine annual status and trends in nest success of endemic species of birds in areas near and far from the human development.
Bats	1. Determine the distribution and abundance of desert bat species in the coastal pacific desert area of a National Park. 2. Determine short- and long-term trends in the presence of bat species and flower and fruit production by cacti species in the area.
Squamata	1. Determine the species of lizards that are found in the area of influence of new road and infrastructure development. 2. Determine the distribution and abundance of lizards in the area of new road and infrastructure development. 3. Evaluate the mitigation measures and the short and long-term trends of the distribution and abundance of lizards in the restored right of way of new road and infrastructure development and in areas control.
Amphibians	1. Determine the species of amphibians that are found in the area of influence of new development. 2. Determine the distribution and abundance of amphibians in water bodies that are close and far away from the new development. 3. Evaluate the presence of the chytrid fungus in the populations of amphibians close and far away from the new development. 4. Determine the short and long-term trends of the distribution and abundance of the chytrid fungus.
Grasslands bio-restoration	1. Determine the composition and abundance of grassland communities. 2. Monitor short and long-term trends in species composition and abundance of grass land species in areas that have been restored to reduce road erosion.
Endemic tree species	1. Determine the distribution and abundance of the endemic tree species with respect to the area of influence of development. 2. Determine the short- and long-term trends in annual recruitment.
Benthic marine community	1. Determine short- and long-term trends in percent cover of sessile marine benthic invertebrates and algae around a new jetty. 2. Determine short- and long-term trends of non-native marine communities that may have been brought by incoming boats.
Marine fish	1. Determine long-term trends in the composition and abundance of fish species of importance to the local communities in areas close to a newly constructed jetty and its area of impact.

BMAP Guiding Principles

The BMAP is based on the following guiding principles: (1) science-based, (2) peer reviewed protocols, (3) biological sampling, (4) link species/habitat to management objectives, (5) focus on species/habitats of conservation concern, (6) comply with national and international regulatory frameworks, and (7) communicate findings and implement recommendations. A small suite of indicators (Simberloff, 1998; Zacharias and Roff, 2001) are part of the basic measures of the BMAP, including biological (e.g., key species abundance and distribution), physical, and chemical elements and processes (e.g., precipitation, water temperature and pH) that represent the entire array of ecosystem features. These features, when measured over time, reflect temporal and spatial changes in habitat and species related to ecosystem structure and function. They also serve as 'leading indicators' of system integrity, stability, and capacity for self-renewal.

Choosing Appropriate Indicators

According to the U.S. National Academy of Sciences (2000, p. 1) the purpose of indicators and monitoring is succinctly described as:

> Developing indicators and monitoring them over time can help to determine whether problems are developing, whether any action is desirable or necessary, what action might yield the best results, and how successful past actions have been. To develop and implement sound environmental policies, data are needed that capture the essence of the dynamics of environmental systems and changes in their functioning.

Sampling all components of biodiversity in a given area is generally an impractical and costly task, but biodiversity surrogates enable monitoring of ecosystem functions (Noss, 1990, 1999). The concept of 'biodiversity indicators' is widespread and a key component of most biodiversity inventory and monitoring systems (Lee *et al.*, 2005). As examples, one might choose an oak tree species affected by gypsy moth defoliation, or a species whose recruitment and mortality is impacted by extended droughts. During the planning stage, it is recommended to identify potential indicator species. Important criteria in selecting indicator species include the following:

(i) The species, populations, or observed physical or chemical phenomena should be good measures of questions that the monitoring program was designed to answer.
(ii) The indicators should be able to detect a condition in advance to assist in solving the problem or else they may have a limited role in achieving the monitoring goals.
(iii) Whenever possible, indicators should be selected for which experimented controls are present (e.g., populations under different management intensities).

(iv) It should be possible to monitor the indicators within realistic budgets.
(v) The species should be selected based on their potential for impacting management decisions (e.g., charismatic species are more likely to facilitate management changes than less-attractive species).

Several considerations play a role in meeting these criteria. First, keystone species are those on which many other species may rely at some point in their life cycles. In tropical regions, for example, nectar-feeding bats are considered as a keystone species because the reproductive success of many plant species through pollination depends on them. Key common species are those that are wide-ranging, easily observed and studied, long-lived, and generally occurring at high-population densities (e.g., oak species in the eastern United States). A sudden increase, decline, or disappearance of these species in certain habitats serves as a warning that may require management attention.

Next, species or taxonomic groups that have sensitive life histories may be good indicators for biodiversity monitoring. Amphibians depend on water for reproduction and will be affected by drought or water pollution. Information on amphibian numbers, diversity, sex ratio, age, and size structure will indicate changes that may affect other components of biodiversity. And, monitoring the abundance and diversity of selected tropical frog populations can provide valuable data on the health of those populations, and hence the habitat, from one year to the next.

Finally, key habitats may also be important indicators. For example, the aspen and wet meadow habitats of the Colorado Rocky Mountains in the United States are good indicator habitats because of the species that depend on them.

Legal and Institutional Framework

The BMAP needs to provide full compliance with applicable legal requirements and project commitments (i.e., government, corporate, lenders, stakeholders, others). The principal requirements and commitments relating to biodiversity and monitoring are embodied in the international, national, state and local laws and regulations, specific project, environmental and social impact assessment (ESIA) and environmental and social management plans (ESMP), Ecological Management Plans, Biodiversity Action Plans, International Performance Standards on Social and Environmental Sustainability, and other national and international conventions, standards, and policies.

Conceptual Ecosystem Model to Support the BMAP

Conceptual ecosystem and landscape models provide an important framework to integrate current information and knowledge of an area to identify the important elements of the system, species, processes, dynamics, and conservation issues, facilitate the communication of complex subjects, and demonstrate connections within the ecological system (Gross, 2003). Models play an important role in virtually all applications of structured decision making, whether adaptive or otherwise, by linking potential management actions to ecological consequences (McGowan, 2011; Moore and McCarthy, 2010). To make informed decisions, it is important to compare and contrast management alternatives in terms of their costs, benefits, and resource consequences. Models typically express benefits and costs in terms of management inputs, outputs, and outcomes through time. Importantly, they allow one to forecast the resource impacts of management.

Development of conceptual models is recommended before initiating the assessment and monitoring programs. This process improves the collective understanding of the system and the models are important throughout the formulation of the BMAP as they provide the scientific framework for the selection of the elements to monitor. Conceptual models are a combination of diagrams and narratives that communicate interactions among ecosystem processes and dynamics, identify links among system stressors and system responses, assist in the selection of assessment and monitoring variables, and communicate the processes to diverse audiences (**Table 2**).

The recommended process (Gross, 2003) to develop project specific conceptual models includes the following activities:

- Goal setting. The primary goal of the model is to provide a synthesis of the ecosystem dynamics that will help identify the species, habitats, and other indicators that will be monitored. It will also show the relationship between indicators and ecosystem variables. The model also facilitates communication among professional and technical personal and external stakeholders.
- Set boundaries of the system. The model clearly defines the boundaries of the system illustrating the special and temporal issues for monitoring and management. It includes the variations of habitats to be assessed and monitored such as altitudinal variations, forest, grasslands, wetlands, and rivers.
- Identify system stressors. Natural and human-induced stressors, such as pollution, road construction, water use, poaching, logging, land-use changes, and many others, impact ecological systems with different intensity. Stressors are these natural or anthropogenic physical, chemical, or biological impacts to a system that are applied to the system at extreme or deficient levels (Barrett *et al.*, 1976). It is important that the model communicates the relationship among stressors, environmental responses, and the characteristics of the ecosystem. During the process of evaluating this relationship it is important to record the possible assessment and monitoring questions as well, and alternative hypotheses to be considered by the program. It is also important to prioritize, at this time, the list of possible indicators for the BMAP.
- Model reviews and revisions. As new information is generated and goals change over time, the model will need to be revised and updated.

Design Monitoring Protocols for Species and Habitats

Assessment and Monitoring Components

Biodiversity assessment involves compiling inventories and providing baseline information for the selection of the

Table 2 Examples of monitoring plans and conceptual models

Monitoring Plans and Conceptual Models	*References*
Biodiversity Monitoring Strategies	Beever (2006)
	Payet *et al.* (2010)
	Regan *et al.* (2008)
	Reynolds *et al.* (in press)
	Scott *et al.* (2009) – Forests
	Watson and Novelly (2004)
Vital signs monitoring plans	Davis (2005)
	Comiskey and Callahan (2008)
	Garrett *et al.* (2007)
	Gitzen *et al.* (2010)
	Jean *et al.* (2005)
	Mau-Crimmins *et al.* (2005)
	Mitchell *et al.* (2006)
	Patterson *et al.* (2008)
	Route and Elias (2007)
Multiple species monitoring and assessment	Dallmeier *et al.* (2011)
	Ludwig *et al.* (1997, Ludwig and Tongway (2000)) – Rangelands
	Manley *et al.* (2000) – Forests & rangelands
	McKelvey *et al.* (2009)
	Noon *et al.* (1999)
Mammals	Wilson *et al.* (1996)
Birds	Ralph *et al.* (1993)
	Welsh (1995)
Reptiles and Amphibians	Dodd (2003) – Amphibians
	Heyer (1994) – Amphibians
	Mattfeldt and Grant (2007) – Salamanders
	Patton *et al.* (2003) – Pond breeding amphibians
	Zylstra *et al.* (2010) – Desert Tortoise
Fish	Isaak *et al.* (2009) – Bull Trout
Plants	Dallmeier and Comiskey (1998, 1999)
	Comiskey *et al.* (2009) – Forest Monitoring
	Elzinga *et al.* (1998)
	Woodward and Beever (2010)
Arthropods	Bowser and Morton (2009)
Impacts	Abbott and Le Maitre (2010) – Climate change
	Lutes *et al.* (2006) – Fire
	Knutson *et al.* (1999) – Fragmentation
	Martin and Murray (2010) – Invasive Species
	Randolph *et al.* (2009) – Disease
Sustainability and Integrity	Aguirre-Bravo *et al.* (2006)
	Harwell *et al.* (1999)
Uncertainty Modelling	Rodríguez *et al.* (2007)

elements to monitor. These assessments can be comprehensive for multiple species and habitats or selective to target individual or group of species and habitats. The assessment can aid in identifying species, habitats, and communities that are important due to their rarity or those that could be used as biological indicators, or those that are important to local human communities. This helps determine species and areas that require conservation or restoration actions. The assessment process includes literature reviews, field surveys, stakeholder consultation, and inventories to gather data relevant to the site-specific monitoring program, and to identify knowledge gaps and target indicator variables (Fancy *et al.*, 2009).

Before monitoring or management activities are implemented, an assessment process should be conducted. The assessment consists of an inventory of species and resources in the area of interest, as well as an evaluation of the priorities and threats to those resources. For example, acid deposition is a stressor that can influence both terrestrial and aquatic systems in the mid-Atlantic region of North America and land-use change can increase fragmentation and alter ecosystem processes in the lowland rainforests of eastern Brazil. Without an understanding of the potential effects of stressors, it is difficult to evaluate the best management options and define the indicators that we need to monitor. Stakeholder consultation from subject matter experts is integral to this process to ensure clear understanding of effects, processes, and potential outcomes.

Assessment objectives generally include information on the location, the taxonomic groups, species or indicator to be recorded, attributes to be measured, the desired state of the species or habitat, and any measures on the status to be evaluated. Monitoring objectives also include the management action, threshold values, levels of change occurring before management activities occur, the measurable quantity, direction and degree of change expected from the management, and a time frame for the monitoring activity. As the sampling design is developed, a sample objective is stated that includes specifics such as the size of the change you aim to detect and levels of precision. Before objectives are finalized, stakeholders should be consulted and a thorough review of relevant literature and background knowledge of the study ecosystem and processes should be completed.

Selecting the elements to assess and monitor can be complex and challenging. The evaluation process requires the identification of the reliable measurements that will generate the best species and habitat information for the least cost. It will also require identifying the measures that will be essential to provide early warnings of conditions that require management intervention. The conceptual model described above helps to break down these diverse and complex elements of the system into categories to simplify the task and assure that all critical elements are considerate. The study area may be divided into an exhaustive list of mutually exclusive categories geographically, climatically, ecologically (life zones), or by scientific disciplines (geology, hydrology, climatology, biology). More commonly, most of the monitoring programs are organized around scientific disciplines as practical and cost effective approach to engage experts and trained specialists to implement the BMAP (see **Table 2**).

BMAP Selection Criteria

Selection criteria of habitats and species and associated parameters to be measured for each are essential first steps after

the development of the conceptual model. The selection criteria include: (1) a broad array of species representing a wide array of ecological roles; vegetation, herbivores, long-lived and short-lived species, and mobile grazers and predators; (2) species and habitats representative of common and dominant ecosystems; (3) species and habitats of conservator concern and special legal status such as threatened or endangered species, endemic species, invasive or alien species and species of social, economical and cultural importance; and (4) abiotic parameters, such as sea temperature, precipitation, and meteorological measures are also complementary important criteria to the biotic system. Based on the information from the baseline assessments, the selection criteria for the BMAP is often based on site-specific priorities such as landscape changes; habitat destruction, overhunting, overgrazing, soil compaction, erosion, degradation of water quality, and the conservation status of certain species and habitats.

BMAP Protocols Design

A key component of a credible monitoring program is the protocols for data collection (Holthausen *et al.*, 2005). A protocol establishes how data are to be collected at each sample point. Protocols are detailed plans for each stage of the monitoring process and should be developed based on the management and monitoring objectives (Atkinson *et al.*, 2004; Oakley *et al.*, 2003; Szaro *et al.*, 1999a). Protocols are necessary to ensure that changes detected from the monitoring data reflect actual changes in the environment and are not just a result of measurements taken by different observers using slightly different methods or approaches. Well-designed protocols allow comparisons of data between organizations and localities, and lend credibility to your program during peer-review processes. Protocols should be developed for all steps of the monitoring project, including sampling design, field methods, personnel training, data management, quality assurance and quality control, data analysis, interpretation, and reporting. The US National Park service has an online database of existing protocols that are available for download: http://science.nature.nps.gov/im/monitor/VitalSigns/BrowseProtocol.aspx.

Monitoring and assessment protocols are designed to frequently collect and analyze observations of selective elements of the system to evaluate the status and changes of these elements and the resource and to measure progress toward meeting the goals and objectives of the BMAP (Oakley *et al.*, 2003; Elzinga *et al.*, 1998). Precise monitoring and assessment protocols are necessary to measure variables and observations in time by different individuals. Protocols often include quality assurance and quality control measures to ensure that they are being implemented correctly and to provide the levels of confidence for the credibility of the program, which allows comparisons of data across space and time sampling intervals.

It is often useful to pose the following questions related to the variables to measure: Are the results of each specific technique well integrated with the overall program? Do the methods that are used ensure reliable, timely, and effective data analysis? Are the collected data subject to the appropriate techniques for data management and analysis? Could the data gathered be coupled with new technologies for analysis and management? What mechanisms exist or can be developed to allow timely transfer of data and information to managers and decision makers?

Oakley *et al.* (2003) provides the guidelines for the preparation of long-term monitoring protocols that include the Standard Operating Procedures that specifically describe how all the components of the monitoring program will be carried out. Periodical formal reviews of individual protocols and the monitoring program are an important component of the overall quality assurance of the program.

Developing the Sampling Design

Fundamentals

Sound sampling designs provide guidelines for the most cost-efficient and effective way to gather and analyze data while maintaining high quality control standards (Smyth and James, 2004). The sampling design should address the following questions: (1) What do researchers need to know about the study area (e.g., data on forest composition, structure, and diversity as well as site-specific information such as the effects of acid deposition, drought, fires, or other disturbances)? (2) How will data from the program be used (managerial and scientific uses)? (3) Is the site representative of the habitat of interest? (4) Will the design be sensitive enough to detect changes? (5) What limits of change are expected and are important to detect (e.g., mortality rates of forest stands should not exceed 3–5% per year)? (6) What is the degree of confidence expected from the results? (7) Will the sampling design produce results that will be representative of what is happening in the entire study area?

Sampling Objectives

Sampling objectives should be developed in conjunction with management objectives (Elzinga *et al.*, 1998). Sampling (or statistical) objectives specify information such as target levels of precision, power, and the magnitude of change that one wants to detect. Precision is a measure of repeatability, or how close two repeated measurements are to each other, and is frequently obtained by calculating the standard deviation of the estimated mean from which confidence limits are calculated. Sampling objectives also identify the level of change to be detected and the risk of missing a real change or, alternatively, detecting a false change. These factors are related to the variability of the population and the number of samples used, all of which will determine the power of the sampling approach. The sampling objective should also define the population of interest. This target population consists of the complete group of sample units (e.g., individual animals and quadrats), about which inferences are to be made. It is important to distinguish the target population from the population actually sampled (the sample population). Statistical inferences can only be made about individuals and areas that have a chance of being sampled. For example, the target population may be all individuals of a specific ant species within a reserve. However, if sampling to the roadside is limited, then the ants in areas close to roads become the sample population and inferences can only be made about ants inhabiting areas close to roads. In some cases, it may be

necessary to define a new sample population for logistical reasons, but this can be done in such a way that the inference is not reduced. For example, if the target population covers a very large geographic area, sample sites need to be randomly placed within the target population and sampling should be done only within these. As the sample sites have been randomly placed within the target population, inference still extends to the entire target population.

Defining Spatial and Temporal Scales

Identifying the spatial and temporal scale of the study is essential at the planning stage because it has important implications in defining effective sampling design and resource requirements. The scale often depends on the scientific goals of the program or on the managerial regions under study. Often, the scale for assessment and monitoring is based on the geographical boundaries of the protected area or conservation unit, or on subsets of such areas. Biodiversity assessment and monitoring at local and regional levels can provide decision makers with high-quality data and cost-effective choices. At local levels, studies usually concentrate on specific communities, and are chosen because of the degree of threat, because of species' locations, or due to organizations' legislative boundaries; but the results may not be representative of the landscape-level biodiversity. Monitoring populations at the landscape level provides broader information sets, but often requires sampling under different land management conditions, and may call for special permits and a large budget. The time frame depends on the biology of the species (e.g., short-lived species will respond more quickly than long-lived ones), the intensity of management (intense management produces rapid changes), and the level of specified change (the smaller the change, the sooner it will be detected). Short-term responses may benefit from more frequent evaluation of the management objectives, and results can be presented relatively fast. Nevertheless, sufficient time must be allocated to detect changes as many impacts may not be detected until as much as a decade after the initial activity. Conclusions were obtained only after a few years indicating that no effects of management could be deceptive. In some cases, monitoring objectives may be best met by focusing on more than one scale. One approach for answering multiscale questions is to select indicators at multiple hierarchical levels of ecological organization (e.g., genetic, individual, population, community, and landscape). Using multiple indicators can also increase knowledge gained as each indicator may respond differently to the same stressor or management activity. Integration of these levels will often require nesting of more intensive and frequent sampling within the context of less-frequent sampling. Objectives with multiple time frames also benefit from frequent assessment of monitoring directions and protocols within the context of long-term goals.

Parameters and Sampling Unit

- *Choice of parameter*: A parameter is a measurable attribute such as plant height or animal mortality. The parameter should be sufficiently sensitive in detecting the desired level of change and capable of distinguishing between natural fluctuations and human-induced change. The change registered by the parameter should be biologically meaningful and lead to a logical management response. The variability in observer error inherent in measuring the parameter should be minimal. In addition, the cost of measuring the variable must be within the budget, and it is necessary to identify the expertise and technical ability needed to measure the parameter. Some of these issues may be addressed by conducting pilot studies.
- *What is a sample unit?* A sample unit is often determined by the choice of parameter. For example, if the parameter is a plant part such as the number of flowers per plant, then your sample unit will be the individual plant. If your parameter is plant density, a quadrat will be the appropriate sample unit choice. Important considerations in choosing sample units include independence from each other, randomization in sample unit selection, and sufficient replication (sample size) of units (Elzinga *et al.*, 1998). When this is the case, statistical estimates of your parameter can be produced along with a reliability estimate. Randomization of the selection process is imperative to produce unbiased information with inference to the population as a whole. Selectors of sample units, based on judgments of 'representative' areas, are often heavily biased.
- *Sampling unit size and shape*: Some sample units have predetermined unit sizes (e.g., individual lizards), and thus the unit sizes and shapes are not factors in the design. Others, such as pitfall traps, could be set at any unit size or shape that the design dictates, and are thus determined by the researcher. As increasing sample sizes (i.e., the number of samples) increases the power to detect changes, smaller sample units may be advantageous by allowing the sampling of more units. However, if the units are too small, the sample will be highly variable, reducing the precision of the estimates. Conversely, in very large plots, the variability among units will be low, but there may be too few plots to compare. This is also dependent on the inherent variability in the community or population being monitored. If the variability in your sample is low, then you may be able to achieve the same power to detect changes with fewer and smaller sample units, than if the variability is high. In plot-based sampling, the shape of the plot must also be determined. Rectangular plots have a greater "spread" over the area than square or circular plots, and are thus more likely to incorporate individuals if they are clumped or aggregated. This decreases the variation among the plots and increases the precision of the estimates.
- *Sampling unit locations*: Positioning of sample units within the population should be done in such a way as to (1) incorporate a random component to the selection process and (2) achieve good interspersion of units across the population. There are a number of different types of sampling methods that can be used; the most common ones being the following:
  a. Simple random sampling can be employed for relatively small study areas with homogenous habitats. This type of sampling keeps analysis simple, but it may lead to some areas within the target population being underrepresented. It tends to be an inefficient design for clustered populations, and it can be logistically difficult

as it can result in considerable travel time between sample units in larger study areas.

b. Stratified random sampling is useful when the parameter of interest is influenced differently across some variables such as habitat or climatic differences. A simple random sample is taken within each stratum, and the number of sample units within each stratum does not need to be equal. For example, if the strata are of differing sizes, the number of units per stratum may be proportional to the size of the strata. If there is a strong influence on the parameter by the strata variable, this design can result in more efficient population estimates than simple random sampling.
c. Systematic random sampling, when samples are placed in a grid or regular pattern, is useful for most sampling situations, except when the number of possible samples is low. It provides good interspersion of sample units and is an efficient design for data collection. Two necessary components of this design are (1) a randomly selected starting point and (2) units that are far enough from each other to be considered independent. Another approach, generalized random tessellation stratified (GRTS) sampling, combines random and systematic sampling in a spatially balanced design allowing for addition or subtraction of sample sites/units in the future while still maintaining spatial balance.
d. Restricted random sampling is setup by determining the number of sampling units that are needed (n), and then dividing the population into n equal-sized segments. A single sample unit is randomly positioned or selected within each segment. This method provides good interspersion across the population and is an alternative to systematic sampling if the number of potential samples is low. When the potential number of samples is 425–430, systematic sampling is more efficient and should be chosen instead.
e. Cluster sampling is used when it is difficult to take a random sample from the population. Clusters of units are identified, and then clusters are randomly chosen instead of randomly choosing from the units themselves. Every unit within the cluster is then measured. This is a cost-efficient design but more complex calculations are required.
f. Two-stage sampling can be used in place of cluster sampling if there is a large number of units within each cluster. In two-stage sampling, a second sample of units is taken within each cluster, instead of sampling all the units in the cluster. Cluster sampling and two-stage sampling are generally the only efficient methods for estimating a parameter associated with individual plants.
g. Should permanent or temporary sample units be used? As monitoring occurs over time, it is necessary to decide whether to measure the same sample units at each time period, or to randomly reselect sites at each time period. This decision partially depends on whether the aim of the project is to estimate the status or the trend. If the main goal of the program is to determine the status of a population, selecting a new sample is appropriate. Conversely, if the objective is to estimate the change or trend in a population, it is best to use permanent units. Permanent units outperform temporary ones for trend assessments, as statistical tests for detecting change between one time period and the next are more powerful when based on permanent units owing to the removal of variation between different plots. However, permanent units are often more costly as units need to be marked for relocating at later time periods, and marking may be infeasible and difficult (e.g., sand dune systems). As monitoring programs often aim to determine statuses and trends, an option is to incorporate both aspects into the design. For example, an augmented rotating panel design includes some units that are resampled every year and others that are reselected on a rotating basis. Another option is the augmented serially alternating design in which some units are always sampled and the others are sampled on a rotating basis. However, data analysis becomes more complex with these designs, and this fact should be taken into account during the planning stage.
h. How many sample units do we need? A sufficient sample size gives you the power to assess whether your management objective has been achieved. However, as costs increase with your sample size, a balance needs to be met. There are a number of issues involved in determining the appropriate sample size. The initial consideration is the level of precision stated in the objectives. For example, if the objective is to increase the population abundance of a species by 20%, a sample size with sufficient precision to detect a change of that magnitude is needed. It is important to keep in mind that the increase in precision is not proportionate to the increase in sample size. The statistical benefits of increasing sample size generally diminish after a certain sample size has been achieved. There are a number of power analysis equations that can be used to determine sample size, and, for these calculations information (such as variability in the measurements) needs to be gathered during a pilot study. Researchers need to be familiar with the assumptions of these formulas and the effects they have on the calculations before proceeding. Assumptions include random selection of sampling units, an infinite population, and an approximately normal distribution of sample means.

Data Analysis

Once the scientific questions are defined and the objectives of the monitoring established, the methodology for the data analysis needs to be considered. Each monitoring protocol contains detailed information on analytical tools and approaches for data analysis and interpretation, including the rationale for a particular approach, and advantages and limitations of each procedure. A number of graphical methods are used to reveal patterns in the data that are not evident by calculating summary statistics such as means and standard deviations (Ellison, 2001). The type of data analysis that is most appropriate should be determined during the planning

stage of the project. Analysis will depend on the objectives and design.

Exploring Data

The first important step in data analysis is to do an initial exploration of the data. This is very important when assessing pilot studies as it will help in understanding the variability in the data as well as whether the assumptions of the parametric tests apply to the data (*see* Parameter Estimation and Significance Tests). There are a number of graphical methods that reveal patterns in the data that are not evident by calculating summary statistics such as means and standard deviations (Ellison, 2001). A normal probability plot allows you to assess how closely the data fits a normal distribution. The observed values are plotted against the values that would be expected if the data came from a normal distribution. A plot showing a straight line indicates normality. Density plots such as histograms are useful for displaying the distribution of data and thus can reveal issues such as data skewness. However, histograms have the disadvantage of grouping the data into arbitrarily determined bars or bins. The number and width of these bins can alter the shape of the histogram. Dit plots are similar to histograms, where all data points are presented. Box plots are another commonly used form of data exploration. This plot displays more information about the data than a histogram. The median, 25th and 75th percentiles, data spread, and outliers are represented.

Parameter Estimation and Significance Tests

The type of data analysis that is most appropriate should be determined during the planning stage of the project. Analysis will depend on the objectives and design. For example, parameter estimation is generally needed for a monitoring project with target/threshold objectives, and significance tests are often used in projects with change/trend objectives. Parameter estimation involves generating a sample statistic such as a mean or total that estimates the true population value. An associated confidence interval is also calculated. This provides an estimate of precision around the sample statistic that specifies the likelihood that the interval includes the true value. The sample statistic can then be compared with the target or threshold value to assess whether the objective has been reached or if further management activity is necessary. For example, the management objective is to maintain a population of at least 5000 individuals of a certain species in a reserve over a period of 10 years. If the lower bound of the confidence interval is above this threshold, there is confidence that the objective has been achieved (based on the desired confidence level). Conversely, if the upper bound of the confidence interval is below the threshold of 5000, there is confidence that the objective has not been achieved. The results are less clear when the confidence interval overlaps the threshold. In these cases, management decisions are often based on the perceived level of risk associated with implementing or not implementing further action. Tests of significance are used to assess whether there has been a change in the parameter of interest between the time periods of the monitoring program. In the same way, they can also be used to compare specific parameters in different spatial areas. These tests produce a probability estimate of whether a change seen is real or simply due to random variation from different samples. The next step is to test for significance and thus a null hypothesis needs to be expressed. This usually asserts that there has been no change in the parameter of interest. The hypothesis will be rejected if the significance test concludes (with a desired level of confidence) that there has been a change in the parameter between the sampling periods. The null hypothesis is evaluated by calculating a test statistic that quantifies the difference between the samples. The confidence level is important as this determines, on a probabilistic basis, how large the test statistic should be before the null hypothesis is rejected. There are two types of errors that can be made in arriving at a decision regarding the null hypothesis. The first, Type I error, involves rejecting the hypothesis when it is actually true. The larger the predetermined confidence level, the more likely it is that the null hypothesis will be falsely rejected. The second, Type II error, involves failing to reject the null hypothesis when it is actually false. Type II error is related to the power of your test. The most commonly used significance tests for monitoring questions are *t*-tests (when comparing two samples), analysis of variance tests (when comparing more than two samples), and w2 tests (when comparing proportions). The first two tests here are parametric tests as they estimate population parameters such as means or totals. Parametric testing requires several assumptions to be met for the results of the test to be valid: (1) The population must follow a normal distribution. (2) The variances must be relatively similar between the samples. (3) The sample must be randomly selected from the population. There are tests that evaluate departures from assumptions 1 and 2, but in most cases graphical review of data is the best method. If deviations from these assumptions are found, increasing sample size or transforming data may resolve the issue and allow you to proceed. If not, there are alternate methods (nonparametric statistics and resampling methods) that do not require these assumptions to be met.

Nonparametric Statistics

When the assumptions of parametric tests cannot be met, or due to the nature of the objectives and data, nonparametric statistics may be an appropriate tool for data analysis. Many nonparametric tests focus on the order or ranking of data, not on the numerical values themselves. Other nonparametric tests are useful for data for which ordering is not possible, such as categorical data. These tests generally focus on the differences between samples in medians instead of their means, as seen in parametric tests. Nonparametric tests commonly used for monitoring questions are w2 tests, Mann–Whitney *U*-test, Wilcoxon's signed rank test, and McNemar's test. The advantages of nonparametric tests are (1) they may be the only alternative when sample sizes are very small, unless the population distribution is known exactly, (2) they make fewer assumptions about the data, (3) they are useful in analyzing data that are inherently in ranks or categories, and (4) they often have simpler computations and interpretations than parametric tests. The main disadvantage of nonparametric tests is that they are generally less powerful than their parametric analogs.

Data and information management goes hand in hand with data collection. The raw data are the basis for the

analysis, synthesis, and modelling of the monitored species and habitats that will generate the interpretation for decision making. For these reasons, data need to be properly recorded, analyzed, reported, archived, documented, and catalogued using a proper information management system. Data management within the information management system needs to ensure that the data are readily available, unverified data are not released, data distributed is accompanied by metadata, sensitive data (i.e., potential commercial value of plant species) are identified and protected from unauthorized access, and data dissemination records are maintained.

Decision Making

Analysis and evaluation elicit answers to the questions underlying the project's objectives, thus allowing for the generation of management recommendations and the calibration of the program. In addition, it allows an evaluation of the process. It is often useful to pose the following questions: Are the results of each specific technique well integrated with the overall BMAP? Do the methods used ensure reliable, timely, and effective data analysis? Are the collected data subject to the appropriate techniques for data management and analysis? Could the data gathered be coupled with new technologies for analysis and management? What mechanisms exist or can be developed to allow timely transfer of data and information to managers and decision makers? Management approaches should be viewed as a means to reach operational goals. Through evaluation of the monitoring data, managers receive timely feedbacks as hypotheses are tested. Thus, evaluations are the tools for improving management by checking on management actions and providing guidelines for improvement. For example, when a predetermined degree of change is detected, appropriate action is taken and the results are evaluated. All the preceding steps lead to decision making regarding the need or lack, thereof, to adjust the management practices and monitoring program. If the findings determine that biodiversity trends are within the expected values, monitoring will continue without substantial alterations. If significant changes in the trends are observed, managers or decision makers need to design the most appropriate response in an adaptive management framework. The reasons for monitoring can be evaluated at this stage. Is the monitoring still required, and, if so, do the objectives still remain the same? Inconclusive results require adjustment of the objectives and sampling approaches to increase the degree of precision. Therefore, careful planning and design of the monitoring program can reduce the risk of inconclusive information.

Data and Information Management

Analysis of the data generated by the BMAP will support the program objectives, the sampling design, and the information for management and stakeholders. There are several levels of data analysis generated by assessment and monitoring protocols: (1) descriptive summaries of basic statistical analysis; (2) baseline descriptive information on status of the monitored resource; (3) trends and condition of the monitored resource; and (4) statistical synthesis of resource trends along spatial and temporal scales. Often descriptive analyses are performed at any time following data collection and input to calibrate sampling methodologies and detect early variations. One approach is to ensure that results are reported on a regular basis – may be annual or less frequent depending on the protocol. The idea is that data is put out in some form to stakeholders. These more frequent reports are low level data summaries. Status and trends reports go into more depth and involve more review, but are less frequent. A long-term trend information synthesis is subject to scientific peer review as appropriate.

The US National Park Service (2008) has provided Data Management Guidelines for Inventory and Monitoring Networks: http://science.nature.nps.gov/im/datamgmt/standards/data_standards_summary_20100610.pdf. In addition, national (NPS, 2008) and network data management plans (e.g., Callahan and Wakamiya, 2008) are developed and made available. The strategy for data and information management for BMAP includes: describing the long-term data and information management for the BMAP with the overall BMAP goals and strategies; defining the specific procedures and work practices for effective data management; training the BMAP personal in sound data management practices; encouraging effective data management practices as a part of project management so that all data are available and usable for management decisions at all times; and promoting data sharing, software development, applications and analyses. The products that are often part of the data and information management system include:

- Compiled data – GIS information, maps, species lists, data files, and analyzed records.
- Documents – BMAP protocols and protocol updates, metadata, information on reference collections and photographs.
- Reports – Periodical or final reports and publications.
- Management and administrative records – BMAP contracts and agreements, BMAP operational timelines, research permits, and memorandum of understandings.

Information Management Goals

The information management goals for the BMAP are: (1) to confirm that the BMAP personnel leading the protocols, data collection and data management, and collaborators understand the essentials of data and information management; (2) to implement effective data management practices by integrating quality control procedures throughout the project; (3) to generate communication protocols to reach multiple stakeholders; and (4) to develop a framework to integrate the monitoring information into recommendations that can be implemented by adaptive management.

Communications

Communicating the BMAP results is essential to disseminate knowledge about the resource and its management, and to keep stakeholders informed of the activities performed by the program. There are many approaches to develop communication plans including the establishing of mechanisms and timing for information dissemination, developing the information and tools to convey the messages to specific audiences,

creating timelines for knowledge and information dissemination, assessing the cost to implement the communication plan, and defining the messages to be communicated to each audience.

BMAP and the Adaptive Management Approach

Adaptive management is a systematic and iterative approach for improving resource management by emphasizing learning from management outcomes (Holling, 1978; Bormann *et al.*, 2007). The underlying assumption is that adaptive management is not simply changing management direction in the face of failed policies but is a planned approach to reliably learn how to improve policies or management practices over time in the face of uncertainty (Bormann *et al.*, 2006, 2007; Burgman, 2005). Adaptive management is not an end in itself, but a means to more effective decisions and enhanced benefits through monitoring the impacts of management practices at achieving the desired outcome. A fundamental premise of adaptive management is that complete knowledge of species and ecological systems is not only incomplete but almost impossible to fully determine (Stankey *et al.*, 2005). As a result, there is a growing realization that decisions need to be made without complete understanding and that expanding knowledge through traditional scientific inquiry will always be limited by resources and time. Adaptive management is one way that understanding and learning can directly inform decision-making and policy processes. In the case of biological resources, the desired outcome is to maintain biodiversity in an optimally functioning state. Human activities can have drastic impacts on biodiversity, impacts that often are irreversible or require long and costly periods of recovery. Managing biodiversity conservation through the adaptive management process can help to avoid or mitigate those impacts.

Adaptive management can be described as a cycle of iterative activities, where each step builds on the learning experiences of previous steps, as the calibration of the goals and objectives is conducted through a feedback process determined by assessing and analyzing monitoring data (**Figure 1**). Each step is calibrated periodically to assure that the appropriate information feeds the next level. The cyclical nature of the process is very important in validating the results of the separate steps. One useful way to describe the implementation of adaptive management is in terms of a setup or planning phase in which its key components are put in place, and an iterative phase in which they are linked together in a sequential decision process (Williams *et al.*, 2009). The iterative phase utilizes the elements of the setup phase in an ongoing cycle of learning about system structure and function, and managing based on what is learned.

Stakeholder Involvement

Project definition, development, and implementation require continuous stakeholder involvement and feedback. Stakeholders are individuals representing local, regional, national, and international communities, agencies, nongovernmental organizations (NGOs), industry, and any other organization with an interest in the project or responsibility within the area of the project's influence. Consulting stakeholders throughout different stages of the project is very important. Of particular importance is the participation of stakeholders in assessing the resource problem and reaching agreement about its scope, objectives, and potential management actions (recognizing that differences of opinion about system responses may exist even with consensus on these issues). By defining the operating environment of a project, stakeholders directly influence both decision-making and the opportunity to learn. The breadth and extent of stakeholder involvement can vary greatly among projects, and both are influenced by the scale and complexity of the intended project. Stakeholder participation in some cases will be restricted to an internal process when the organizations' goals are driven by legal requirements and institutional/organizational policies.

Management Objectives as Related to the BMAP Objectives

Objectives, resource status, and learning all influence the choice of management interventions in adaptive management. Objectives also play a crucial role in evaluating performance, reducing uncertainty, and improving management through time. It is especially important to have clear, measurable, and agreed-on management objectives at the outset, to guide decision making and assess progress in achieving management success. It is often the case in adaptive decision making that there are multiple objectives. In such a situation it is important to weigh different objectives in terms of their perceived importance, so as to facilitate the comparison and prioritization of management alternatives (Burgman, 2005). Initially, management objectives are formed to frame the question and goal.

Management Actions

Like any iterative decision process, adaptive decision making involves the selection of an appropriate management action at each decision point, given the status of the resources being managed at that time. Resource managers and stakeholders, typically working with scientists, have the responsibility of identifying the set of potential actions from which this selection is made.

The management alternatives in an adaptive management project constitute a key element in its operating environment, and they can strongly influence strategy selection. Just as the choices made in daily life depend on one's available options, so are strategy choices in an adaptive management project constrained by the set of available options. If these options fail to span a reasonable range of management activities or fail to produce recognizable and distinct patterns in system responses, adaptive management will be unable to produce effective and informative strategies. This argues for careful thinking about the potential actions to be included in a project.

Decision-Making

At each decision point, an action is chosen from the set of available management alternatives. Management objectives

are used to guide this selection, given the state of the system and the level of understanding when the selection is made. The appropriate action is likely to change through time, as understanding evolves and the resource system responds to environmental conditions and management actions. That is, management is adjusted in response to both changing resource status and learning. It is the influence of reduced uncertainty (or increased understanding) on decision making that renders the decision process adaptive.

Follow-up monitoring

Monitoring is used to track system behavior, and in particular to track the responses to management through time. In the context of adaptive management, monitoring is seen as an ongoing activity, producing data to evaluate management interventions, update measures of model confidence, and prioritize management options in the next time period.

Evaluation, Decision-Making, and Feedback

At any given time, the gain in understanding from monitoring and assessment is used to inform the selection of management actions (Szaro *et al.*, 1999b). As understanding evolves, so too does the decision making that is influenced by improved understanding. In this way, the iterative cycle of decision making, monitoring, and assessment leads gradually to improved management as a consequence of improved understanding.

Management approaches should be viewed as a means to reach operational goals. Through evaluation of the monitoring data, managers receive timely feedback as hypotheses are tested. Thus, evaluations are the tools for improving management by checking on management actions and providing guidelines for improvement. For example, when a predetermined degree of change is detected, appropriate action is taken and the results are evaluated.

The need to better understand and characterize the process elements of adaptive management often becomes more pressing as adaptive management rolls forward through time. Thus, stakeholder perspectives and values can shift as the adaptive process unfolds, and previously unanticipated patterns in resource dynamics can arise that require an adjustment of objectives, alternatives, and other elements of the process. In this sense, learning focuses on changes in institutional arrangements and stakeholder values as well as changes in the resource system.

All the preceding steps lead to decision making regarding the need, or lack thereof, to adjust the management practices and monitoring program. If the findings determine that biodiversity trends are within the expected values, monitoring will continue without substantial alterations. If significant changes in the trends are observed, managers or decision makers need to design the most appropriate response in an adaptive management framework. The reasons for monitoring can be evaluated at this stage. Is the monitoring still required, and, if so, do the objectives still remain the same?

Inconclusive results require adjustment of the objectives and sampling approaches to increase the degree of precision. Therefore, careful planning and design of the monitoring program can reduce the risk of inconclusive information.

Learning from Experience

The explicit hypotheses about the relationships between strategies and objectives and the emphasis on learning are some of the basic elements of adaptive management (Price and Daust, 2009). Assessment and monitoring programs are widespread in conservation organizations, agencies, and industry as a means to improve understanding and future management. Many have had mixed results, suffering from problems that may have been foreseen if other approaches and programs had been reviewed (Stem *et al.*, 2005; Reid, 2001). Although each monitoring program has differing objectives and thus needs to be tailored individually, there is much to be learned from previous projects. Common issues include: poorly defined monitoring questions and objectives, poorly trained field crews, incomplete baseline information and assessments, sampling plans that are not designed to meet objectives, delays in analyzing data, inadequate monitoring durations, poor data management and interpretation, and lack of integration of the monitoring results into management. In many cases, utilizing management in an experimental context, such as that implemented in adaptive management, may be the only feasible way to gain the system understanding needed to improve that management.

All these can be overcome by paying greater initial attention in identification of specific monitoring questions, goal development, and project planning. Strict development of protocols for each step of the process increases quality control and allows the identification of potential issues at the outset. A more integrated approach to biodiversity management across sectors, administrative boundaries and at landscape and seascape scales would be an important step forward in conservation (European Environment Agency, 2010). To accomplish this, there is a need for greater collaboration within the conservation community to promote learning from experience, reduce duplication of efforts, and strengthen projects by combining resources.

See also: Ecosystem, Concept of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Function, Principles of. Population Viability Analysis

References

Abbott I and Le Maitre D (2010) Monitoring the impact of climate change on biodiversity: The challenge of megadiverse Mediterranean climate ecosystems. *Austral Ecology* 35: 406–422.

Atkinson AJ, Trenham PC, Fisher RN, *et al.* (2004) Designing monitoring programs in an adaptive management context for regional multiple species conservation plans. In: *U.S. Geological Survey Technical Report*, USGS Western Ecological Research Center, Sacramento, CA.

Barrett GW, Van Dyne GM, and Odum EP (1976) Stress ecology. *BioScience* 26: 192–194.

Beever EA (2006) *Great Lakes I&M Network Monitoring Plan*. National Park Service, US Department of the Interior.

Bormann BT, Haynes RW, and Martin JR (2007) Adaptive management of forest ecosystems: Did some rubber hit the road? *Bioscience* 57: 186–191.

Bormann BT, Lee DC, Kiester AR, Busch DE, Martin JR, and Haynes RW (2006) Adaptive Management and Regional Monitoring. Chapter 10. In: Haynes RW, Bormann BT, and Martin JR (eds.) *Northwest Forest Plan – the First Ten Years (1994–2003): Synthesis of Monitoring and Research Results*. Portland, OR: Pacific Northwest Research Station. PNW GTR 651, USDA Forest Service.

Bowser ML and Morton JM (2009) Monitoring and modeling terrestrial arthropod diversity on the Kenai National Wildlife Refuge. In: McWilliams W, Moisen G, and Czaplewski R (eds.) *Forest Inventory and Analysis (FIA) Symposium 2008*. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station October 21–23, 2008; Park City, UT. Proc. RMRS-P-56CD.

Burgman M (2005) *Risks and Decisions for Conservation and Environmental Management*. Cambridge, UK: Cambridge University Press.

Callahan K and Wakamiya S (2008) Mid-Atlantic Network Data Management Plan. In: *Natural Resource Report NPS/MIDN/NRR – 2008/XXX*. National Park Service, Fort Collins, CO.

Comiskey JA and Callahan KK (2008) Mid-Atlantic Network vital signs monitoring plan. In: *Natural Resource Report NPS/MIDN/NRR – 2008/071*. National Park Service, Fort Collins, CO.

Comiskey JA, Schmit JP and Tierney G (2009) Mid-Atlantic Network forest vegetation monitoring protocol. *Natural Resource Report NPS/MIDN/NRR – 2009/119*, National Park Service, Fort Collins, CO.

Dallmeier F and Comiskey JA (eds.) (1999) *Forest Biodiversity in North, Central, South America and the Caribbean: Research and Monitoring,* vol. 21, pp. Carnforth, Lancashire, UK: UNESCO and the Parthenon Publishing Group 768 pp.

Dallmeier F and JA Comiskey (eds.) (1998) *Forest Biodiversity Research, Monitoring and Modelling: Conceptual Background and Old world Case Studies*. Man and the Biosphere Series, vol. 20.

Dallmeier F, Fenech A, MacIver D, and Szaro R (2010) *Climate Change, Biodiversity and Sustainability in the Americas: Impacts and Adaptations*. Washington, DC: Smithsonian Institution Scholarly Press.

Dallmeier F, Langstroth R, Davis G, *et al.* (2011) Assessing and monitoring biodiversity across human-modified landscapes: PERU LNG biodiversity monitoring and assessment program (BMAP). Smithsonian Conservation Biology Institute, Center for Conservation Education and Sustainability, Lima, Peru and Washington, DC.

Davis GE (2005) National BMP Stewardship and 'Vital Signs' Monitoring: A case study from Channel Islands National BMP, California. *Aquatic Conservation: Marine & Freshwater Ecosystems* 15: 71–89.

Dodd KC Jr. (2003) Monitoring amphibians in Great Smoky mountains National Park. *U.S. Geological Survey Circular 1258*, Tallahassee, FL.

Ellison AM (2001) Exploratory data analysis and graphic display. In: Scheiner SM and Gurevitch J (eds.) *Design and Analysis of Ecological Experiments*. New York, NY: Chapman & Hall.

Elzinga CL, Salzer DW, and Willoughby JW (1998) Measuring and monitoring plant populations. USDA-Bureau of Land Management Technical Reference 1730-1, Denver, CO.

European Environment Agency (2010) Assessing biodiversity in Europe – the 2010 Report. EEA, Copenhagen, Denmark.

Fancy S, Gross J, and Carter S (2009) Monitoring the condition of natural resources in US national parks. *Environmental Monitoring and Assessment* 151: 161–174.

Feld CK, Martins da Silva P, Paulo Sousa J, *et al.* (2009) Indicators of biodiversity and ecosystem services: A synthesis across ecosystems and spatial scales. *Oikos* 118: 1862–1871.

Forest Service, Rocky Mountain Research Station (2006) In: Aguirre-Bravo C, Pellicane PJ, Burns DP, and Draggan S (eds.) *Monitoring Science and Technology Symposium: Unifying Knowledge for Sustainability in the Western Hemisphere. Proceedings RMRS-P-42CD*. Fort Collins, CO: U.S. Department of Agriculture. Forest Service, Rocky Mountain Research Station. 990 p.

Garrett, LK, Rodhouse TJ, Dicus GH, Caudill CC, and Shardlow MR (2007) Upper Columbia Basin Network vital signs monitoring plan. In: *Natural Resource Report NPS/UCBN/NRR—2007/002*. National Park Service, Fort Collins, CO.

Gitzen RA, Wilson M, Brumm J, *et al.* (2010). Northern Great Plains Network vital signs monitoring plan. *Natural Resource Report NPS/NGPN/NRR – 2010/186*. National Park Service, Fort Collins, CO.

Gross JE (2003) *Developing Conceptual Models for Monitoring Programs*. Ft Collins, CO: NPS Inventory and Monitoring Program.

Heyer WR (ed.) (1994) *Measuring and Monitoring Biological Diversity. Standard Methods for Amphibians*. Washington, DC: Smithsonian Books. 384pp.

Holling CS (ed.) (1978) *Adaptive Environmental Assessment and Management*. New York: Wiley.

Holthausen R, Czaplewski RL, DeLorenzo D, *et al.* (2005) Strategies for Monitoring Terrestrial Animals and Habitats. *Gen. Tech. Rep. RMRS-GTR-161*. Fort Collins, CO, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 34 p.

Isaak D, Rieman B, and Horan D (2009) A watershed-scale monitoring protocol for bull trout. Gen. Tech. Rep. RMRS-GTR-224. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 25 p.

Jean C, Schrag AM, Bennetts RE, Daley R, Crowe EA, and O'Ney S (2005) *Vital Signs Monitoring Plan for the Greater Yellowstone Network*. Bozeman, MT: National Park Service, Greater Yellowstone Network. 107 pp.

Knutson MG, Sauer JR, Olsen DA, Mossman MJ, Hemesath LM, and Lannoo MJ (1999) Effects of landscape composition and wetland fragmentation on frog and toad abundance and species richness in Iowa and Wisconsin, U.S.A. *Conservation Biology* 13: 1437–1446.

Lee W, Matt McGlone M, and Wright E (2005) Biodiversity inventory and monitoring: A review of national and international systems and a proposed framework for future biodiversity monitoring by the Department of Conservation. *Landcare Research Contract Report: LC0405/122*, Wellington, New Zealand.

Ludwig JA, Tongway DJ, Freudenberger DO, Noble JC, and Hodgkinson KC (eds.) (1997) *Landscape Ecology, Function and Management: Principles from Australia's Rangelands*. Collingwood, Victoria, Australia: CSIRO Publishing.

Lutes DC, Keane RE, Caratti JF, *et al.* (2006) FIREMON: Fire effects monitoring and inventory system. Gen. Tech. Rep. RMRS-GTR-164-CD. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Manley PN, Zielinski WJ, Stuart CM, *et al.* (2000) Monitoring ecosystems in the Sierra Nevada: The conceptual model foundations. *Environmental Monitoring and Assessment* 64: 139–152.

Martin LJ and Murray BR (2010) A predictive framework and review of the ecological impacts of exotic plant invasions on reptiles and amphibians. *Biological Reviews* 86: 407–419.

Mattfeldt S and Grant E (2007) Are two methods better than one? Area constrained transects and leaf litterbags for sampling stream salamanders. *Herpetological Review* 38: 43–45.

Mau-Crimmins T, Hubbard A, Angell D, Filippone C, and Kline N (2005) Sonoran Desert Network Vital Signs Monitoring Plan. *Technical Report NPS/IMR/SODN-003*. National Park Service. Denver, CO.

McKelvey KS, Cushman SA, Schwartz MK, and Ruggiero LF (2009) Wildlife monitoring across multiple spatial scales using grid-based sampling. In: McRoberts RE, Reams GA, Van Deusen PC, and McWilliams WH (eds.) *Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium; 2006 October 16–19; Monterey, CA. Gen. Tech. Report WO-79*. Washington, DC: U.S. Department of Agriculture. Forest Service, pp. 137–142.

Mitchell BR, Shriver WG, Dieffenbach F, *et al.* (2006) Northeast Temperate Network Vital Signs Monitoring Plan. *Technical Report NPS/NER/NRTR – 2006/059*. National Park Service, Northeast Temperate Network, Woodstock, Vermont.

Moore AL and McCarthy MA (2010) On valuing information in adaptive-management models. *Conservation Biology* 24: 984–993.

National Academy of Sciences (USA) (2000) *Ecological Indicators for the Nation*. Washington, DC: National Academy Press.

NPS (2008) Data management guidelines for inventory and monitoring networks. *Natural Resource Report NPS/NRPC/NRR – 2008/035*. National Park Service, Fort Collins, CO.

Noon BR, Spies TA, and Raphael MG (1999) Conceptual basis for designing an effectiveness monitoring program. In: The strategy and design of the effectiveness monitoring program for the Northwest Forest Plan, USDA Forest Service Gen. Tech. Rept. PNW-GTR-437, ch. 2.

Noss RF (1990) Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355–364.

Noss RF (1999) Assessing and monitoring forest biodiversity: A suggested framework and indicators. *Forest Ecology and Management* 115: 135–146.

Oakley KL, Thomas LP, and Fancy SG (2003) Guidelines for long-term monitoring protocols. *Wildlife Society Bulletin* 31: 1000–1003.

Patterson ME, Atkinson AJ, Witcher BD, *et al.* (2008) South Florida / Caribbean Network vital signs monitoring plan. *Natural Resource Report NPS/SFCN/NRR – 2008/063*. National Park Service, Fort Collins, CO.

Patton PWC, Timm B, and Tupper T (2003) Monitoring Pond-Breeding Amphibians: A Protocol for the Long-term Coastal Ecosystem Monitoring Program at Cape Cod National Seashore. Long-term Coastal Ecosystem Monitoring Program Cape Cod National Seashore Wellfleet, MA.

Payet K, Rouget M, LaGabrielle E, and Esler KJ (2010) Measuring the effectiveness of regional conservation assessments at representing biodiversity surrogates at a local scale: A case study in Réunion Island (Indian Ocean). *Austral Ecology* 35: 121–133.

Price K and Dave Daust D (2009) Making monitoring manageable: A framework to guide learning. *Canadian Journal of Forest Research* 39: 1881–1892.

Ralph CJ, Geupel GR, Pyle P, Martin TE, and DeSante DF (1993) *Handbook of Field Methods for Monitoring Landbirds*. Albany, CA: U.S. Department of Agriculture. Forest Service, Pacific Southwest Research Station. Gen. Tech. Rep. PSW-GTR-144. 41 p.

Randolph KaDonna, Bechtold W, Morin R, and Zarnoch S (2009) From detection monitoring to evaluation monitoring – a case study involving crown dieback in northern white-cedar. In: McWilliams W, Moisen G, and Czaplewski R (Compilers). *Forest Inventory and Analysis (FIA) Symposium 2008*; October 21–23, 2008; 13 p. Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Regan HM, Hierl LA, Franklin J, *et al.* (2008) Species prioritization for monitoring and management in regional multiple species conservation plans. *Diversity and Distributions* 14: 462–471.

Reid LM (2001) The epidemiology of monitoring. *Journal of the American Weather Research Association* 37: 815–820.

Reynolds JH, Thompson WL, and Russell B (2011) Planning for success: Identifying effective and efficient survey designs for monitoring. *Biological Conservation* 144(5): 1278–1284.

Rodríguez JP, Brotons L, Bustamante J, and Seoane J (2007) The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions* 13: 243–251.

Route B and Elias J (eds.) (2007) Long-term Ecological Monitoring Plan: Great Lakes Inventory & Monitoring Network. *Natural Resource Report NPS/GLKN/NRR—2007/001*. National Park Service, Fort Collins, CO.

Scott CT, Renate B, and Brewer K (2009) Design tool for inventory and monitoring. In: McWilliams W, Moisen G, and Czaplewski R (Compilers). Forest Inventory and Analysis (FIA) Symposium 2008; October 21–23, 2008; Park City, UT. *Proc. RMRS-P-56CD*,7 p. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Simberloff D (1998) Flagships, umbrellas, and keystones: Is single-species management passé in the landscape era? *Biological Conservation* 83: 247–257.

Smyth AK and James CD (2004) Characteristics of Australia's rangelands and key design issues for monitoring biodiversity. *Austral Ecology* 29(1): 3–15.

Stankey GH, Clark RN, and Bormann BT (2005) Adaptive management of natural resources: Theory, concepts, and management institutions. PNW-GTR-654, USDA Forest Service, Pacific Northwest Research Station, Portland, OR.

Stem C, Margoluis R, Salafsky N, and Brown M (2005) Monitoring and evaluation in conservation: A review of trends and approaches. *Conservation Biology* 19: 295–309.

Szaro RC, Cooperrider A, Correll CS, Lint J, and Malk A (1999a) Data collection, management and inventory. In: Johnson N, Malk A, Szaro RC, and Sexton WT (eds.) *Ecological Stewardship: A Common Reference for Ecosystem Management, vol. 1: Key Findings*, pp. 231–239. Oxford, UK: Elsevier Science.

Szaro RC, Maddox D, Tolle T, and McBurney M (1999b) Monitoring and evaluation. In: Johnson N, Malk A, Szaro RC, and Sexton WT (eds.) *Ecological Stewardship: A Common Reference for Ecosystem Management, Volume 1: Key Findings*, pp. 223–230. Oxford, UK: Elsevier Science

Watson I and Novelly P (2004) Making the biodiversity monitoring system sustainable: Design issues for large-scale monitoring systems. *Austral Ecology* 29: 16–30.

Welsh DA (1995) An Overview of the Ontario Forest Bird Monitoring Program in Canada. In: Ralph CJ, Sauer JR, and Droege S (eds.) *Monitoring Bird Populations by Point Counts. Gen. Tech. Rep. PSW-GTR-149*, pp. 93–98. Albany, CA: U.S. Department of Agriculture. Forest Service, Pacific Southwest Research Station.

Williams BK, Szaro RC, and Shapiro CD (2009) Adaptive Management: The U.S. Department of the Interior Technical Guide. Adaptive Management Working Group, U.S. Department of the Interior, Washington, DC.

Wilson DE, Cole FR, Nichols JD, Rudran R, and Foster MS (1996) *Measuring and Monitoring Biological Diversity: Standard Methods for Mammals (Biodiversity Handbook)*. Washington, DC: Smithsonian Books.

Woodward A and Beever EA (2010) *Framework for ecological monitoring on lands of Alaska National Wildlife Refuges and their partners*. Anchorage, AK: U.S. Geological Survey. Open-File Report 2010-1300, 94 p.

Zacharias MA and Roff JC (2001) Use of focal species in marine conservation and management: A review and critique. *Aquatic Conservation: Marine and Freshwater Ecosystems* 11: 59–76.

Zylstra ER, Steidl RJ, and Swann DE (2010) Evaluating Survey Methods for Monitoring a Rare Vertebrate, the Sonoran Desert Tortoise. *Journal of Wildlife Management* 74: 1311–1318.

Freshwater Ecosystems

Robert G Wetzel, University of Alabama, Tuscaloosa, AL, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 75–87, © 2001, Elsevier Inc.

Glossary

Benthos Nonplanktonic animals associated with substrata within sediments or closely above the sediment–water interface.
Detritus Nonliving organic matter in both soluble and particulate forms.
Littoral zone Region of a lake or river between the land and the open water (pelagic zone) that is colonized by emergent, floating-leaved, and submersed aquatic plants and their attendant sessile microbiota (periphyton).
Pelagic zone Open-water portion of a lake or reservoir beyond the littoral zone.
Periphyton Bacteria, fungi, algae, and sessile microfauna growing attached to substrata (sediments, rock, plants, animals, sand).
Phytoplankton Small photosynthetic plankton, largely algae and cyanobacteria.
Plankton Small organisms with no or limited powers of locomotion that are suspended in the water and largely dispersed by turbulence and other water movements.
Production Amount of new organic biomass formed over a period of time, and includes any losses from respiration, excretion, secretion, injury, death, and predation.
Zooplankton Small animal plankton, usually denser than water, that sink by gravity to greater depths.

Fresh Waters: Physical Ecosystem Structure

Most water on the Earth's surface is saline in oceans (97.61%), and most of the remainder is polar ice (2.08%) and ground water (0.29%). The remaining water (only 0.009%) exists in freshwater ecosystems as temporary storage in lakes and reservoirs on land. These surface waters have much shorter renewal times than those of the oceans (**Table 1**). The volume of water flowing from these storage sites to the sea in streams and rivers is very small (0.00009% of total water), with a mean residence time of about two weeks. During high-precipitation periods, surface waters of rivers increase and often recharge adjacent groundwater aquifers. Return flows from ground water to rivers occur during periods of low flow or drought, and usually maintain a base flow in river channels.

The storage of water in lakes and retention times are altered by shifts in the balance between inputs from all sources and water losses. Lakes receive water from precipitation directly on the surface, from surface inflows from the drainage basin, and from subsurface groundwater seepage. Lakes lose water by flow from a usually single outlet (*drainage lakes*), by seepage through the basin walls into the ground water (*seepage lakes*), evaporation, and evapotranspiration from higher aquatic plants. Saline lakes, which constitute nearly half of inland surface waters (see **Table 1**), occur in *closed* basins with no outflow except by evaporation. That evaporation results in a gradual marked increase in residual concentration of dissolved salts that were imported with influent water.

Natural lakes are concentrated in the subarctic and temperate regions of the Northern Hemisphere, whereas reservoirs are constructed predominately in the subtemperate and subtropical regions. Some 40% of the total volume of surface freshwater is contained in seven great lake basins of Siberia (Lake Baikal), North America (the Laurentian Great Lakes: Superior, Michigan, Huron, Erie, and Ontario), and eastern Africa (Lake Tanganyika). Most of the millions of lakes and reservoirs, however, are very much smaller and relatively shallow, usually <15 m in depth (**Figure 1**). Mean depths are even more shallow (<5 m). As a result, light commonly penetrates to over half of the sediments within the basins of a large percentage of lakes. Photosynthesis can thus occur not only by algae that are suspended in the water column but can extend to higher aquatic plants and attached algae. Major flowing-water ecosystems occur in low-gradient tropical and subtropical regions and are posited to support high photosynthetic productivity in extensive floodplain and land–water interface regions.

The productivity and internal metabolism of aquatic ecosystems are driven and controlled by energy from solar radiation acquired by photosynthesis. Inland waters receive organic products of photosynthesis directly from their aquatic flora and indirectly from their drainage basins as particulate and dissolved organic matter from terrestrial and wetland plants imported by stream water, storm runoff, ground water, and the atmosphere.

Light is attenuated exponentially with increasing depth in water. As the light is attenuated, ultraviolet wavelengths are absorbed strongly by dissolved organic compounds in the water. The blue portion (500–600 nm) of the visible spectrum penetrates most deeply in relatively clear waters. Infrared wavelengths are absorbed rapidly, largely as a result of the molecular structure of water molecules, and much of this energy is dissipated as heat. Because water becomes less dense as heat content increases above 4 °C, the less dense warmer water floats upon the more dense cooler water. This density stratification results in separation of major strata within lakes, where a warmer, less dense stratum (*epilimnion*) overlies a cooler, denser zone (*hypolimnion*). The interface zone (*metalimnion*) is a region of rapid thermal discontinuity, often as much as several degrees change per meter, between the epilimnion and the hypolimnion.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00060-5

Table 1 Water in the biosphere[a]

	Volume (thousands of km^3)	*Percentage of total*	*Renewal time*
Oceans	1,370,000	97.61	3100 years[b]
Polar ice, glaciers	29,000	2.08	16,000 years
Groundwater (actively exchanged)[c]	4000	0.295	300 years
Freshwater lakes	125	0.009	1–100 years[d]
Saline lakes	104	0.008	10–1000 years[d]
Soil and subsoil moisture	67	0.005	280 days
Rivers	1.2	0.00009	12–20 days[e]
Atmospheric water vapor	14	0.0009	9 days

[a]Wetzel (1983, 2000). Slightly different values are given by I. A. Shiklomanov, but ratios are similar.
[b]Based on net evaporation from the oceans.
[c]Kalinin and Bykov estimated that the total groundwater to a depth of 5 km in the earth's crust amounts to 60×10^6 km^3. This is much greater than the estimate by the U.S. Geological Survey of 8.3×10^6 km^3 to a depth of 4 km. Only the volume of the upper, actively exchanged groundwater is included here.
[d]Renewal times for lakes vary directly with volume and mean depth, and inversely with rate of discharge. The absolute range for saline lakes is from days to thousands of years.
[e]Twelve days for rivers with relatively small catchment areas of less than 100,000 km^2; 20 days for major rivers that drain directly to the sea.

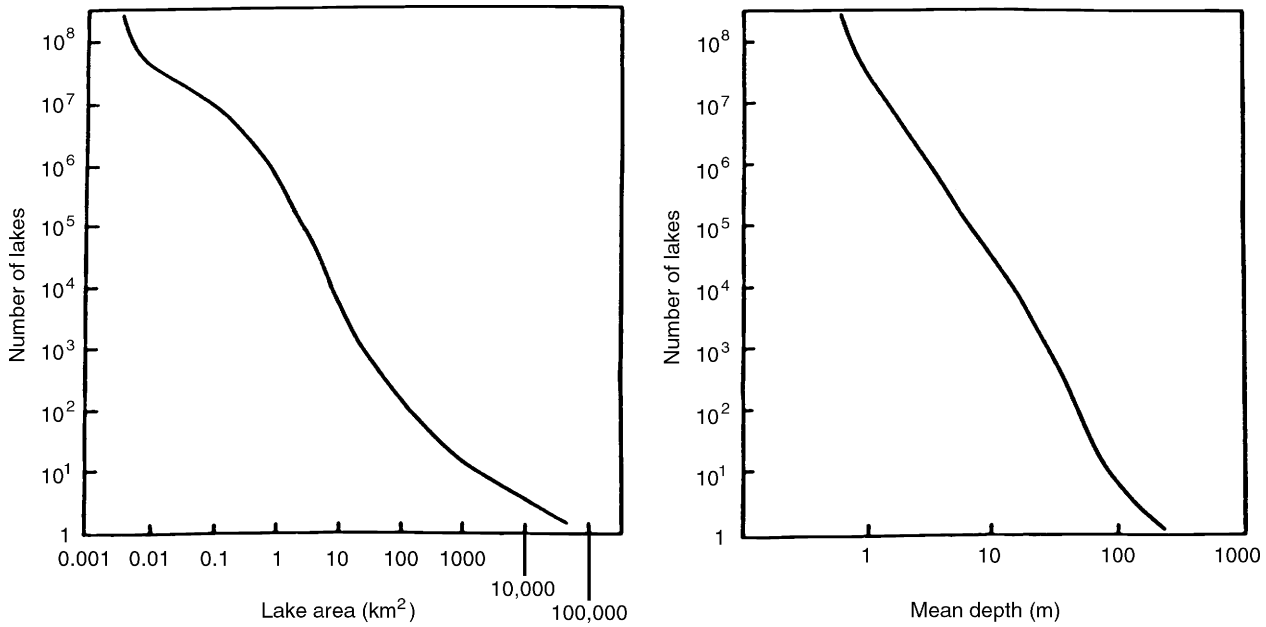


Figure 1 Approximate number of lakes of the world in relation to lake area and approximate mean water depth. (Wetzel (1990)).

The resulting density stratification affects not only the thermal structure and water mass stratification, but also the hydrodynamics of lakes and reservoirs. Heat is distributed and altered by the physical work of wind energy, currents and other water movements, basin morphometry, and water losses. Patterns of density-induced stratification influence physical and chemical properties and cycles both spatially within the lake and seasonally. These characteristics structure the aquatic habitats and have marked attendant effects on all chemical cycles, metabolic rates, and the population dynamics of organisms and their productivities.

The total *salinity* of inland waters nearly always consists of eight most abundant ionic species, of usually four major cations (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}) and four anions (HCO_3^-, CO_3^{2-}, SO_4^{2-}, Cl^-). The salinity of fresh waters has a world average concentration of about 120 mg/L, but varies among continents and with the lithology of landmasses (**Table 2**).

Of the major constituents of the cellular protoplasm of organisms (C, H, O, N, P, and S), the biogeochemical cycles control the physical and metabolic availability of phosphorus, nitrogen, and several minor nutrients. These elements often limit biotic development, particularly photosynthetic generation of organic matter. Phosphorus, in comparison to other macronutrients required by biota, is least abundant geochemically in a large portion of the global landforms and drainage. As a result, phosphorus is commonly the first element to limit biological productivity, especially of photosynthetic producers and heterotrophic microbiota. When phosphorus is in adequate supply, nitrogen availability

Table 2 Mean composition of river waters of the world (mg L^{-1})[a]

	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	CO_3^{2-} (HCO_3^-)	SO_4^{2-}	Cl^-	NO_3^-	Fe (as Fe_2O_3)	SiO_2	Total
North America	21.0	5	9	1.4	68	20	8	1	0.16	9	142
South America	7.2	1.5	4	2	31	4.8	4.9	0.7	1.4	11.9	69
Europe	31.1	5.6	5.4	1.7	95	24	6.9	3.7	0.8	7.5	182
Asia	18.4	5.6	5.5	3.8	79	8.4	8.7	0.7	0.01	11.7	142
Africa	12.5	3.8	11	—	43	13.5	12.1	0.8	1.3	23.2	121
Australia	3.9[b]	2.7	2.9	1.4	31.6	2.6	10	0.05	0.3	3.9	59
World	15	4.1	6.3	2.3	58.4	11.2	7.8	1	0.67	13.1	120
Cations (μ eq 1^{-1})	750	342	274	59	—	—	—	—	—	—	1425
Anions	—	—	—	—	958	233	220	17	—	—	1428

[a]Wetzel (1983, 2000).
[b]Values of calcium are likely less, on the average, than Na and Mg in Australian surface waters.

invariably limits productivity. Standing inland waters have been categorized into various trophic scales on the basis of ranges of major nutrients and algal productivity (**Table 3**).

Functionally similar organisms of biological communities can be grouped into *trophic levels* based on similarities in patterns of organic matter production and consumption. Energy is transferred and nutrients are cycled within an overall ecosystem trophic structure. The productivity of each trophic level is the rate at which energy enters the trophic level from the next lower level. Because organisms expend considerable energy for maintenance and since death of an organism routes much energy and nutrients into the detrital pool of dead organic matter, only a portion of the energy of one trophic level is available for transfer and use by higher trophic levels. Available energy decreases progressively at higher trophic levels, so that rarely can more than five or six trophic levels be supported. The efficiency of energy transfer from one level to the next is low (5–15%), and often decreases as trophic level increases.

The organic matter of freshwater ecosystems is primarily (>90%) dead organic matter, and most (usually 80–90%) of this organic detritus occurs as soluble or colloidal organic compounds. Oxygen, however, is very insoluble in water, and most diffuses to the atmosphere. When organic matter in lakes is metabolized and organic compounds respired by heterotrophic organisms, the amount of organic matter to be oxidized can be much larger than the dissolved oxygen available for oxidation. This limitation can become acute when lakes are stratified by density differences, and exposure of deep strata to oxygen of the atmosphere may be restricted for long periods of time (months).

In large lakes, the volume of water with sufficient light to support photosynthesis (*euphotic zone*) is small relative to the total volume of water containing stored dissolved oxygen that can be used in heterotrophic respiration. In small and relatively shallow lakes, however, the proportion of water supporting photosynthesis is large relative to total lake volume, and respiratory demands can exhaust oxygen dissolved in the lower strata of the lake. The process of eutrophication exacerbates the exhaustion of dissolved oxygen in large portions of lake ecosystems. *Eutrophication* is the increased loading of organic matter to an aquatic ecosystem. Often in lakes, eutrophication results from increased photosynthetic production of organic matter in response to excessive loading of nutrients, particularly phosphorus, to the water from external sources.

When dissolved oxygen is depleted from water strata, alternate electron acceptors are utilized by microbial metabolism of dissolved organic compounds. Bacterial denitrification occurs by biochemical reduction of oxidized nitrogen anions (NO_3^- and NO_2^-), concomitant with the oxidation of organic matter by many genera of facultative anaerobic bacteria. When nitrate and nitrite are depleted, alternative oxidized ions can function as electron acceptors, although with decreasing efficiencies—manganese, iron, and particularly sulfate, the latter being reduced to hydrogen sulfide. Often the same compound can serve as a hydrogen acceptor or donor, depending on the environmental conditions. When these electron acceptors are essentially depleted, organic matter can be anaerobically degraded by fermentation to methane and CO_2 in two stages. Facultative and obligate anaerobic bacteria convert proteins, carbohydrates, and fats primarily to fatty acids by hydrolysis and fermentation. Obligatory anaerobic methanogenic bacteria then convert the organic acids to methane and CO_2.

Diversity of Habitats

Freshwater ecosystems consist of entire drainage basins as water moves from land and in groundwater runoff to stream and river channels, and to recipient lakes or reservoirs. The nutrient and organic matter content of drainage water from the catchment area is modified in each of the terrestrial soil, stream, and wetland–littoral components as water moves downgradient to and within the lake or reservoir itself (**Figure 2**). Photosynthetic productivity of organic matter is generally low to intermediate in the terrestrial components, highest in the wetland–littoral interface regions between the land and water, and lowest in the open water (*pelagic*) zone. The same productivity profile emerges in the gradient from land to river channels, where the greatest productivity occurs in the marginal floodplain regions. Autotrophic productivity in river channels is generally low, as is also the case in the pelagic regions of lakes. Most of the organic matter utilized by heterotrophic communities in running water is imported from floodplain and terrestrial sources as particulate and especially dissolved and colloidal organic compounds.

Table 3 General ranges of primary productivity of phytoplankton and related characteristics of lakes of different trophic categories[a]

Trophic type	*Mean primary productivity (mg C m^{-2} day^{-1})*[b]	*Phytoplankton density (cm^3 m^{-3})*	*Phytoplankton biomass (mg C m^{-3})*	*Chlorophyll*[a] *(mg m^{-3})*	*Dominant phytoplankton*	*Light extinction coefficients (η m^{-1})*	*Total organic carbon (mg L^{-1})*	*Total P (μg L^{-1})*	*Total N (μg L^{-1})*	*Total inorganic solids (mg L^{-1})*
Ultraoligotrophic	<50	<1	<50	0.01–0.5		0.03–0.8		<1–5	<1–250	2–15
Oligotrophic	50–300		20–100	0.3–3	Chrysophyceae, Cryptophyceae,	0.05–1.0	<1–3			
Oligomesotrophic		1–3			Dinophyceae, Bacillariophycee			5–10	250–600	10–200
Mesotrophic	250–1000		100–300	2–15		0.1–2.0	<1–5			
Mesoeutrophic		3–5						10–30	500–1100	100–500
Eutrophic	>1000		>300	10–500	Bacillariophyceae, Cyanobacteria,	0.5–4.0	5–30			
Hypereutrophic		>10			Chlorophyceae, Euglenophyceae			30–>5000	500–>15,000	400–60,000
Dystrophic	<50–500		<50–200	0.1–10		1.0–4.0	3–30	<1–10	<1–500	5–200

[a]Reproduced from Wetzel RG (1983) *Limnology*, 2nd ed. Philadelphia: Saunders College Publ, and Wetzel RG (2000) *Limnology: Lake and River Ecosystems*. San Diego: Academic Press, after many authors and sources.
[b]Referring to approximately net primary productivity, such as measured by the ^{14}C method.

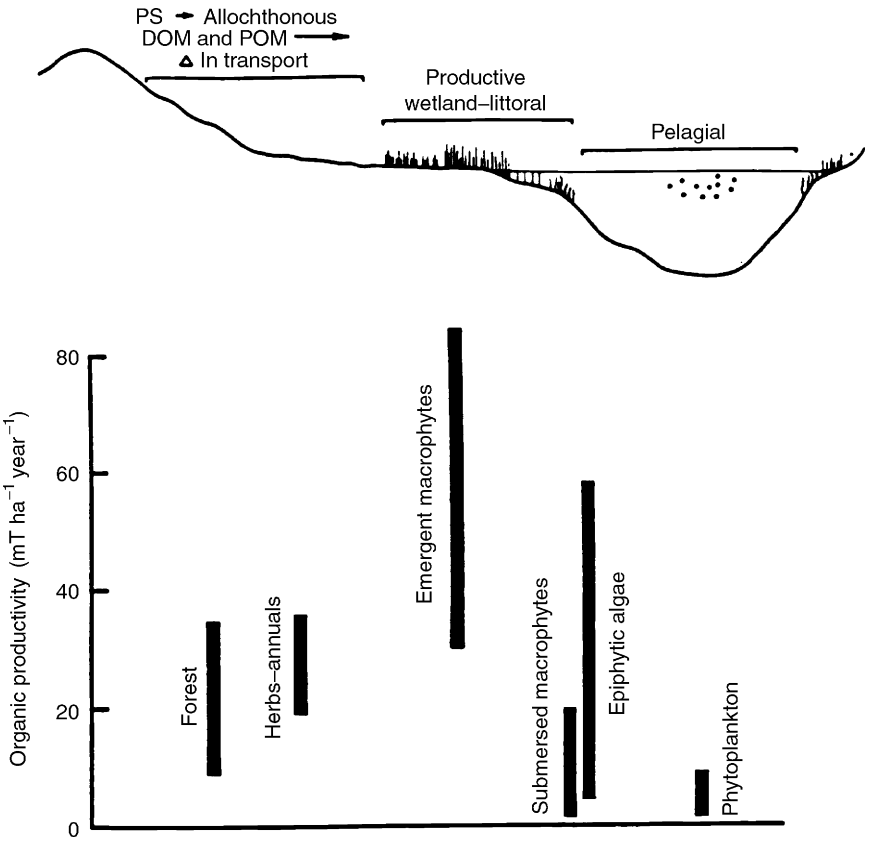


Figure 2 The lake ecosystem showing the drainage basin with terrestrial photosynthesis (PS) of organic matter, movement of nutrients and dissolved (DOM) and particulate (POM) organic matter in surface water and groundwater flows toward the lake basin, and chemical and biotic alteration of these materials en route, particularly as they pass through the highly productive and metabolically active wetland–littoral zone of the lake per se (net organic productivity in metric tons per hectare per year). (Wetzel (1990)).

The interface region between the land and water is always the most productive per unit area along the gradient from land to the open water of lakes, reservoirs, and streams. Because most aquatic ecosystems occur in geomorphologically mature terrain of gentle slopes and are small and shallow, the wetland–littoral components usually dominate in both productivity and the synthesis and loading of organic matter to the systems. The region of greatest productivity is the emergent macrophyte zone. Emergent aquatic plants have a number of structural and physiological adaptations that not only tolerate the hostile reducing anaerobic conditions of saturated sediments but also exploit the high nutrient conditions and water availability of this habitat. Nutrients entering the zone of emergent macrophytes from external sources tend to be assimilated by bacterial and algal microflora of the sediments and detrital organic particles, and are then recycled to the emergent macrophytes. Dissolved organic compounds released from decomposition of plant detrital materials dominate the export of organic matter from the emergent plant zone.

Submersed macrophytes are limited physiologically by slow rates of diffusion of gases and nutrients in water within the boundary layers surrounding the leaves and by reduced availability of light underwater. Internal recycling of resources, particularly of gases (CO_2, O_2) of metabolism and of critical nutrients, is important to the abilities of submersed plants to function and grow as well as they do in underwater conditions of chronic light and gas limitations. Despite these adaptive mechanisms, growth and productivity of submersed plants are less than those of emergent and floating macrophytes.

The second most productive component of the wetland–littoral community is the microflora attached to aquatic plants epiphytically and to other surfaces, both living and dead. The surfaces provided by aquatic plants in lakes and rivers can be very large, often exceeding 25 m^2 per square meter of bottom sediments. High sustained growth of attached microflora results from their recycling of essential gases (CO_2, O_2) and dissolved nutrients within the attached communities. Nutrient uptake from the surrounding water is directed primarily to the high net growth of attached microflora and is responsible for the high capacity of wetland–littoral areas to improve the quality of water passing through these communities.

The wetland–littoral complex of higher plant and microbial communities produces the major sources of organic matter and energy of many freshwater ecosystems, including the marginal floodplains of many rivers. Most of the particulate organic matter is decomposed within these interface regions. Organic matter is exported predominantly from these marginal regions as dissolved organic matter to the recipient lake or river (**Figure 3**).

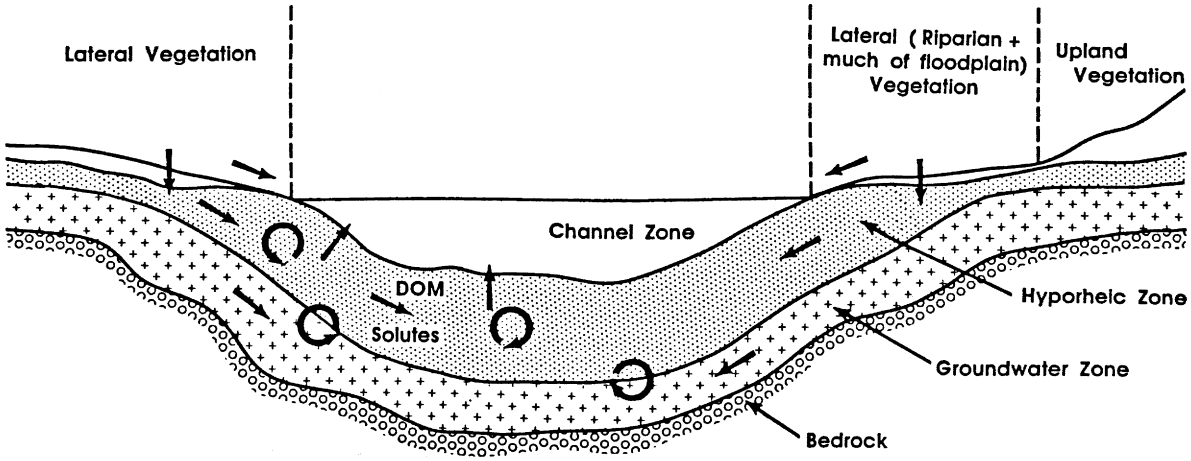


Figure 3 Lateral and vertical boundaries of flowing-water ecosystems. The stream ecosystem boundary is defined as the hyporheic/goundwater interface and thereby includes a substantial volume beneath and lateral to the main channel. Vegetation rooted in the hyporheic zone is therefore part of stream ecosystem production. Arrows indicate flow pathways of dissolved organic matter and inorganic solutes derived from plant detritus within the stream ecosystem. (Reproduced from Wetzel RG and Ward AK (1992) Primary production. In: Calow P and Petts GE (eds.) *Rivers Handbook*, pp. 354–369. Oxford: Blackwell Scientific, UK, with permission from Wiley.)

The deep-water pelagic zone of lakes is the least productive along the gradient from land to water (see Figure 2), regardless of nutrient availability. Growth of phytoplanktonic algae of the pelagic zone is limited by sparse distribution in a dilute environment where efficient nutrient recycling is restricted by the sinking of senescent phytoplankton below the depth of photosynthesis. When nutrient recycling and availability are increased, greater phytoplankton cell densities attenuate underwater light and reduce the volume of water in which photosynthesis occurs. Despite low productivity per unit area, pelagic productivity can be collectively important in large lakes and for higher trophic levels that depend on this source of organic matter.

Higher trophic levels of communities in freshwater ecosystems consist of zooplankton (dominated by four major groups of animals: protozoa/protista, rotifers, and the crustaceans cladocera and copepoda) and benthic invertebrates. In the pelagic zone, small fishes, fry of larger fishes, and predatory zooplankton, which collectively comprise a third trophic level (primary carnivores), consume a portion of these generally herbivorous organisms. A fourth trophic level may consist of medium-sized piscivorous fishes, and the fifth level of large predatory piscivorous fishes. Higher trophic levels are rare in freshwater ecosystems.

The species composition of the higher trophic levels affects the pathways of energy utilization from lower trophic levels. Environmental factors that selectively influence the populations of the communities can alter the pathways and strengths of energy fluxes from subordinate trophic levels. For example, efficiency of consumption of primary production by zooplankton is often appreciably greater in the absence of zooplankton-feeding fishes than in their presence. The population structure of the phytoplankton community responds variably to grazing impacts in concert with their available resources (light, nutrients, and organic constituents). The phytoplankton community may or may not be able to compensate for grazing losses in overall primary production, but generally is able to shift quite quickly to an alternative, less vulnerable species composition.

Ecosystem Functions: Interactive Regulatory Mechanisms

Most of the organic carbon of aquatic ecosystems, and hence much of the energy potentially available for metabolism of heterotrophic biota of the ecosystem, exists as dead organic matter (*detritus*), in both dissolved and particulate forms. Dissolved organic matter and much of the particulate organic matter, usually expressed as organic carbon, move with the water. Much of the particulate organic matter that is not decomposed ultimately settles by gravity and is deposited at the bottom of static water bodies. Dissolved organic compounds can also sediment if adsorbed to inorganic or organic particulate matter or if polymerization and aggregation occur into particulate form. Because much of the highest photosynthetic production of organic matter occurs in shallow regions of lakes and streams, and because most lakes are small and shallow, most heterotrophic decomposition occurs in benthic regions as organic matter. Therefore, because much of the organic matter is displaced to the sediments, most of the total heterotrophic decomposition is also largely displaced from sites of production to the sediments.

The specific composition of organic matter varies greatly. The organic matter of pelagic phytoplankton, which can dominate in-lake production in very large, deep lakes, is relatively labile. Much of the phytoplanktonic algal and cyanobacterial organic matter is decomposed by microbiota either in the water column within the water strata of production or below as it sediments to the bottom. Some highly variable portion, usually much less than 50% on an annual basis, of the phytoplanktonic organic matter is ingested and partially digested by zooplankton. An appreciable portion, usually $>50\%$, of that particulate organic matter ingested by zooplankton is egested as dissolved and as particulate organic matter, some of which also settles as feces to the sediments.

The composition of organic matter of the wetland and littoral regions of lakes and streams tends to be considerably more recalcitrant to rapid microbial degradation than that of

phytoplankton because of the littoral dominance of lignocellulose-based organic compounds of structural tissues of higher plants. Because of the much higher rates of production and slower rates of decomposition of both the particulate and dissolved organic matter from these shallow-water sources, and because of the shorter distances and times for particulate organic matter to be displaced to the sediments, organic matter accumulates on and in the sediments at greater rates in shallow regions than in those of the pelagic region. The deposition of organic matter in the sediments encourages intensive bacterial metabolism. As a result, intrusion of dissolved oxygen from the overlying water is consumed rapidly; nearly all sediments in standing freshwater ecosystems are anaerobic and highly reducing chemically.

Although the specific composition of organic matter varies greatly, detritus inevitably carries most of the energy of the ecosystem from its points of photosynthetic origins to places of transformation by heterotrophic organisms. Most of that heterotrophic transformation occurs by microbes in areas where organic matter is concentrated. Even though greater than 70% of the very large pool of dissolved organic matter of the pelagic region of lakes and rivers consists of relatively recalcitrant humic and fulvic acids that originated from higher plants, these compounds are slowly metabolized by bacteria, often being catalyzed by partial photolytic degradation by ultraviolet light (both UV-A and UV-B). In contrast, much of the particulate organic matter of the productive littoral areas accumulates on the bottom, largely under anaerobic conditions particularly in sediments and in detrital mats.

Transformation of particulate organic matter of the pelagic zone also occurs by digestion in predatory metazoan animals as well as by pelagic microbes. Productivity of the microbes is maximized within the resources available from dead organic matter and nutrients. Microbial heterotrophic utilization is highly dynamic and changes with sufficient rapidity to allow many generation turnovers before predators with much slower turnover rates can respond reproductively to changes in availability. Viral mortality of microbes, as well as ingestion of bacteria by protists (the "microbial loop"), diverts much organic matter from animal trophic levels. Bacterial, viral, and much of protistan heterotrophic metabolism as respiration represents a major output of carbon, largely as CO_2, from the aquatic ecosystem. Although this output is a loss of carbon from the ecosystem, it is not an energetic loss from the ecosystem.

That productivity that is not ingested by metazoan animals is of major importance and dominates the energy flows within the pelagic region. Nonpredatory death and metabolism by prokaryotic and protistan heterotrophs usually dominate energy flows within the pelagic region and completely dominate energy flows of the composite wetland–littoral, benthic, and pelagic components that collectively constitute freshwater ecosystems. Even ignoring the benthic portions of lakes where most of the organic carbon metabolism occurs and considering only the pelagic metabolism of lakes, for example, most of the organic carbon entering the system does not reach higher trophic levels.

The microbial heterotrophy is of primary importance to higher trophic levels by means of feedback processes, both positively (e.g., nutrient recycling and utilization by primary producers) and negatively (e.g., oxygen consumption and production of toxic fermentative metabolic end products). The abundance, distribution, and microbial decomposition of dissolved and particulate detritus both regulate and stabilize energy metabolism and nutrient availability in aquatic ecosystems. Because of the very large magnitudes and relative chemical recalcitrance of these detrital sources, the large but slow metabolism of detritus provides an inherent ecosystem stability that energetically dampens the ephemeral, volatile fluctuations of numbers and biomass of the biota of higher trophic levels.

Biodiversity Within Freshwater Ecosystems

The biodiversity of most microbial, plant, and animal groups of freshwater stream, lake, and wetland ecosystems is very poorly known on a global basis. Moreover, it is likely erroneous to believe that biodiversity of many taxa in temperate ecosystems is appreciably less than that in tropical freshwater ecosystems. For example, the rivers and streams of Alabama, a "hot spot" within the United States, contain 43% of all gill-breathing snails, 52% of freshwater aquatic or semiaquatic turtles, 60% of mussels, and 38% of freshwater fishes of North America (Lydeard and Mayden, 1995). Over half of the species in some of the groups mentioned are either threatened or endangered under the U.S. Endangered Species Act of 1973.

The fresh waters of the world are collectively experiencing markedly accelerating rates of degradation. Major sources of disturbance impact the biodiversity of freshwaters in many ways. Direct chemical toxicants released into surface waters and ground waters are common. Many forms of heavy metals, inorganic reducing agents, and organic compounds enter the environment and eventually fresh waters. Although some are inactivated, such as by chemical precipitation, or oxidized, many have long residue resident times. Despite dispersion and dilution in the aquatic environment, bioconcentration of both metals and organic compounds is common, by which toxicity can be increased exponentially. The biological effects of many of these compounds are unknown. Elemental and compound toxicity is increased markedly when the pollutant substances are radioactive or acidic. Increased acidity in poorly buffered fresh waters can increase concentrations and potential biotic availability and toxicity of metals, such as aluminum, that are not normally abundant in soluble, reactive states.

Nutrients, particularly phosphorus and nitrogen, can lead to well-understood enhancements of plant and other organic productivity in fresh waters. This eutrophication process often results in enhanced rates of decomposition and in chemical conditions that greatly reduce or eliminate suitable habitat for many species of plants and animals. Similar excessive loading of organic matter and enhanced rates of degradation result from organic sewage from human populations, industry, and agriculture. A further common pollutant that markedly reduces habitat availability is the suspension of finely divided organic and inert inorganic matter. Erosion and transport of such suspensions are increasing as large areas of forest and former wetland interface zones between land and water are eliminated, primarily for agricultural expansion. As a result, flow patterns of surface waters are often altered and benthic habitats of surface waters can be obliterated by sedimentation.

Changes in biodiversity in freshwater ecosystems can arise from many other disturbances. Introduction of certain competitively superior species can result in marked losses of biodiversity. The infamous example of introduction of the Nile perch into Lake Victoria of East Africa resulted in the extinction of over 200 species of its endemic cichlid fish taxa in two decades. There are many other examples of introductions of exotic plant and animal species that resulted in either direct destruction of prey or inferiorly competitive species or indirect alteration of habitats required by many species. Dense, floating macrophyte communities and other eutrophication-associated excessive plant productivity often result in deoxygenation and reduction in habitat and elimination of many plant and animal species.

Largely based on terrestrial studies, the Eltonian diversity–stability hypothesis suggested that, because of the many different traits of multiple species, ecosystems with more diverse habitats would likely have species that will survive and expand during and following an environmental disturbance and compensate for those species that are reduced by the disturbance. Therefore, a more species diverse freshwater ecosystem should be more resilient to disturbances than a less biodiverse system.

Because genetically based physiological differences and tolerances among species can be small, the individual interactive strengths of some species in a freshwater ecosystem can become saturating at high biodiversity. A point can be reached where increasing species may be functionally redundant and have reduced individual impact on the ecosystem processes. On the basis of both theoretical and experimental grounds, only a small fraction of species manipulations have strong influences on food web structure. Species redundancy implies that an appreciable functional resiliency exists in which the ecosystem can compensate in its collective metabolism and biogeochemical cycling when disturbed. Although the population dynamics become progressively less stable as the biodiversity and the number of competing species increases, biodiversity can enhance the resiliency of many community and ecosystem processes in the rate that the system metabolism returns to equilibrium states following a disturbance.

There is very little storage capacity for organic carbon within the higher trophic levels. Low residence times among the higher trophic levels results in rapid cycling of carbon and nutrients of food web components. Such rapid cycling and recycling result in a reduction in the resiliency of the higher trophic levels. Most of the storage of organic carbon occurs in the dissolved organic carbon compartment in the open water and in the paniculate organic carbon deposited in the sediments. In both of these compartments, the soluble organic carbon of the pelagic areas of lakes or running water of streams and the organic carbon of the sediments, the cycling of carbon is slowed. That rate of cycling is slowed in the pelagic by the recalcitrant chemical composition of the dissolved organic carbon emanating largely from higher plants. In the sediments, cycling is further impeded by the anoxic conditions that prevail almost universally among aquatic sediments. The reduced rates of cycling and recycling result in an inherent increase in resilience stability of the ecosystem.

Any factor that influences the rates of nutrient and carbon cycling in freshwater ecosystems will influence the resilience of the ecosystem and its biodiversity to disturbances. Changing sources of organic matter, as discussed in the following, and hence bacterial metabolism and nutrient cycling thus change resilience and biotic stability.

A wealth of limnological data from a spectrum of hundreds of lake ecosystems of differing productivity suggests that with a shift in nutrient loadings, concomitant shifts occur in the development of photosynthetic producers and loadings of organic matter. During the common sequential development of lake ecosystems over long periods of time (centuries, millennia), shifts in the ratios of higher vegetation versus algal dominance can occur. Increased relative organic loading from higher vegetation results in proportionally greater loading of recalcitrant dissolved organic carbon, which can suppress nutrient cycling and increase the resilience of the ecosystem.

In addition, the development of higher vegetation in littoral and wetland combinations increases the habitat heterogeneity enormously, often by a factor of 10 or more, in comparison to lakes with limited littoral development. Species diversity nearly always increases under these circumstances by at least an order of magnitude among nearly all major groups of organisms, particularly among the lower phyla.

Greater biodiversity may have a greater collective effect by improving the capacity of ecosystems to recover from large disturbances. Reduced biodiversity increases vulnerability by reducing the total collective physiological tolerances of the community to large habitat changes. Recovery after a major or catastrophic disturbance would be slower with a reduced aggregation of residual physiological ranges within the remaining species. Recovery then must depend to a greater extent on slower fortuitous methods, such as importation of species, rather than generation from residual surviving species, or slow recolonization processes such as from remnants in resting stages or seed banks. In some cases, such as in many ponds, streams, and reservoirs in clay-rich regions where high turbidity often occurs with successive rain events, photosynthetic productivity within the water is intermittently but repeatedly suppressed. Biodiversity is likely also suppressed under these conditions or restricted to species with high reproductive potential that utilize improved conditions in periods between turbidity events.

As indicated earlier, disturbance to freshwater ecosystems can occur in many forms and to different extents. Certain perturbations can be catastrophic, such as overwhelming a lake or stream with an organic or inorganic poison in which most of the biota are eliminated. Many disturbances, however, are more gradual over long periods of time (months, year), such as nutrient enrichment, or irregularly episodic and often of short duration (days, weeks), such as severe flooding and the scouring of a section of river. Biodiversity is coupled to ecosystem stability and the type and extent of disturbances.

A model of the responses of organic productivity of lake ecosystems to changes in nutrient loading from the drainage basins has been abundantly verified by nutrient and comparative primary productivity data of phytoplankton, attached algae, and macrophytes from hundreds of lakes in different stages of ontogeny (**Figure 4**). The differences in plant productivity result in very different amounts and types of chemical composition of organic matter loaded to lakes. Because of the large amounts and relative chemical recalcitrance of dissolved

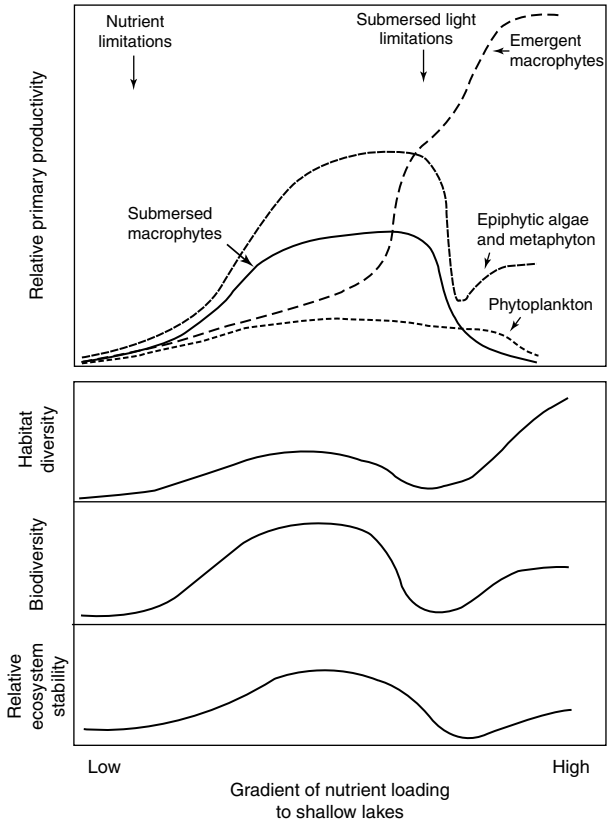


Figure 4 Relative changes in the primary productivity of phytoplankton, macrophytes, and attached microflora, and habitat diversity, species diversity, and ecosystem stability along a gradient of nutrient loading to lake ecosystems. (Reproduced from Wetzel RG (1999) Biodiversity and shifting energetic stability within freshwater ecosystems. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* 54: 19–32, with permission from Cornell University Press.)

and particulate organic detrital sources from higher plants, decomposition of this pool is slowed. The resulting large but slow metabolism of organic detritus provides an inherent ecosystem stability that energetically buffers the organic matter reserves from rapid exploitation and cycling of nutrients contained in the organic matter. Forested watersheds, undisturbed riparian floodplains, and wetland–littoral zones are therefore important to the metabolic stability within the lakes and streams as a result of both the chemical quality and the large quantity of their inputs of organic matter. The organic carbon couplings between the drainage basin and land–water interface zones are metabolically as important to lake and stream energetic and carbon utilization stability as are the traditionally studied nutrient (phosphorus, nitrogen, silica) loading relationships to food web productivity. As the sources of organic matter change, the chemical recalcitrance to degradation can change. Alterations of rates of utilization, nutrient recycling, and energetic resiliency can ensue as a result. These changes directly affect biodiversity by altering the chemical and physical habitat characteristics.

A common disturbance to lake and reservoir ecosystems involves progressive increases in loading of nutrients. These loadings, in excess of losses to sites of temporary or permanent inactivation such as to the sediments, result in enhanced nutrient availability for phytoplankton and other autotrophs and increased rates of growth and productivity. If nutrient loading increases to an oligotrophic lake, increased productivity is rapid. Similarly, if the disturbance is brief (i.e., the duration of increased nutrient loading is relatively short), nutrient cycling is rapid, the ecosystem will recover rapidly, and productivity would be reduced proportionally to the load reduction. The return time of such an oligotrophic ecosystem is high, but the resiliency is low. Species diversity of the plankton tends to be high in oligotrophic lakes, but because of limited physical habitats, particularly associated with littoral areas and surfaces, composite biodiversity of the lake ecosystem would be low (see **Figure 4**).

As nutrient loading increases, particularly among shallow lakes that predominate globally, a marked shift in the productivity occurs from the pelagic to attached surfaces associated with living aquatic plants and the particulate detritus of senescing macrophyte biomass (see **Figure 4**). Under these conditions, total primary productivity and biomass increase greatly. Habitat diversity among the massively dissected surfaces of submersed aquatic plants increases exponentially, and biodiversity among attached biota would increase by at least an order of magnitude. It is likely that among all autotrophic and heterotrophic microbial groups, as well as for most of the smaller metazoans, over 90% of the total freshwater ecosystem species are sessile in association with surfaces.

As nutrient loading increases further, phytoplanktonic productivity per unit volume increases, but self-shading by the suspended algae restricts the depth of light penetration and hence the depth of the photogenic zone. Phytoplanktonic productivity per unit area declines precipitously under these eutrophic conditions of light limitation and is usually accompanied by a marked decrease in planktonic biodiversity as well (see **Figure 4**). Submersed light limitations also reduce or eliminate submersed macrophyte growth and attendant surfaces for microbial growth. The elimination of photosynthesis of attached microbiota decreases collective productivity markedly. Biodiversity of the attached biota would decline precipitously as a result. The losses of attached microbial communities and their interdependent metabolism cause a massive reduction in the capacities of the lake ecosystems to retain loaded nutrients and dissolved organic compounds. Because of the great accumulation and recycling of nutrients in the sediments from a predominantly planktonic eutrophic ecosystem, the time for recovery of the water quality of a lake from a reduction of nutrient loading would be long and not proportional to load reductions. Under these conditions after the nutrient loading perturbations end, the internal nutrient loading and recycling could be greater than external nutrient loading.

Continual high nutrient loading and resulting hypereutrophic phytoplanktonic conditions generate large areas of anaerobic reducing conditions and slower rates of decomposition of organic matter. Production exceeding decomposition leads to rapid sedimentation and generation of increased shallow habitat conducive to colonization by emergent macrophytes. The high productivity of emergent macrophytes increases the proportion of lignocellulose supporting tissues that are relatively recalcitrant to rapid decomposition, particularly under reducing conditions. The collective result is markedly increased productivity and habitat

diversity. However, although habitat diversity is high among very shallow wetland-dominated waters, environmental fluctuations are also more extreme than in continually submersed habitats. Quite different community structures occur and biodiversity is likely tempered as a result, though comparative data to support this statement are few.

Many descriptive and experimental studies suggest that greater biodiversity results in a higher efficiency of utilization of nutrients. Greater biodiversity likely entails more physiological diversity that can better cope with natural vagaries in environmental parameters. Competition for resources, nutrients in this specific case, is thus intensified. Efficiency of nutrient retention in an ecosystem would be maximized under conditions of greatest microbial community diversity. That microbial diversity (algae, fungi, protists, and bacteria) is maximized in the attached communities where habitat diversity and microenvironment differences are the greatest. Physical constraints of boundary layers and mucopolysaccharide matrices in which the attached microbiota live mean that movements of ions and gases to organisms living within the attached communities occur predominantly by diffusion. Nutrients, once acquired, are intensively recycled among the attached biota and conserved. Resources from external sources can then be utilized largely for new growth and reproduction. Nutrient retention is very high within the micro-communities and collectively within the attached habitats.

In certain productive lakes, piscivore consumption of planktivorous fishes can lead to enhanced development of cladoceran zooplankton. The high grazing rates of cladoceran zooplankton can result in selective reduction of larger algae for brief to moderate intervals of time. The effects of these changes among the larger forms are, however, poorly translated to smaller microbiota, and only minor changes are seen at the microalgal and bacterial levels. For example, removal of portions of the larger algae can decrease competition for nutrient and light resources. Often the smaller forms with shorter generation times increase in productivity. Nutrient recycling likely also increases, particularly as decomposition of these algae by bacteria is accelerated by protistan microconsumers (heteroflagellates, ciliates, and related organisms) and tightly retained among the microbiota. Little is known concerning changes in biodiversity under these conditions, but it appears that biodiversity among the microbiota would increase and likely compensate for the losses of larger forms by the crustacean grazing. A shift in energy fluxes occurs, here from larger planktonic forms to smaller forms with much higher turnover rates and increased rates of resource turnover. Compensatory mechanisms likely also appear in the composite biodiversity of the smaller biota. Resiliency to disturbances within the pelagic food web likely declines as a result.

See also: Freshwater Ecosystems, Human Impact on. Invertebrates, Freshwater, Overview. Lake and Pond Ecosystems. Pelagic Ecosystems. Wetlands Ecosystems

Reference

DeAngelis DL (1992) *Dynamics of Nutrient Cycling and Food Webs*. London: Chapman & Hall.

Francko DA and Wetzel RG (1983) *To Quench Our Thirst: Present and Future Freshwater Resources of the United States*. Ann Arbor: University of Michigan Press.

Lydeard C and Mayden RL (1995) A diverse and endangered aquatic ecosystem of the southeast United States. *Conservation Biol.* 9: 800–805.

Tilman D (1996) Biodiversity: Population versus ecosystem stability. *Ecology* 77: 350–363.

Wetzel RG (1983) *Limnology*, 2nd ed. Philadelphia: Saunders College Publ.

Wetzel RG (1990) Land–water interfaces: Metabolic and limnological regulators. *Verh. Int. Verein. Limnol.* 24: 6–24.

Wetzel RG (1995) Death, detritus, and energy flow in aquatic ecosystems. *Freshwater Biol.* 33: 83–89.

Wetzel RG (1999) Biodiversity and shifting energetic stability within freshwater ecosystems. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* 54: 19–32.

Wetzel RG (2000) *Limnology: Lake and River Ecosystems*. San Diego: Academic Press.

Freshwater Ecosystems, Human Impact on

Kaj Sand-Jensen, University of Copenhagen, Copenhagen, Denmark

© 2013 Elsevier Inc. All rights reserved.

Glossary

Detritus Dead organic matter.
Denitrification Bacterial nitrate respiration of organic matter to elemental nitrogen under oxygen-free conditions in soils and aquatic sediments.
Oxic and anoxic conditions Environmental conditions with and without oxygen, respectively.
Red lists Compilations of recently extinct, threatened, and vulnerable species for a country or a larger region.
Secchi depth Measurement of the transparency of water by determining the depth at which a white Secchi disk (diameter 20 cm) lowered from the surface disappears out of the sight of the observer.

Freshwaters: Opportunities and Conflicts

Why are humans so interested in freshwater and natural freshwater ecosystems? The answer is obvious – humans cannot live without it! Freshwater is an essential resource for daily life and for the production of vital food and industrial products.

Freshwater is used for drinking and for cooking. It is also used for bathing, washing clothes, and removing waste products. Securing clean tap water and good sanitary conditions through installation of water closets and sewer pipelines has been important in reducing infectious diseases and infant mortality. In their daily lives, people also enjoy natural lakes and streams, which provide variation in the landscape and opportunities for bathing, fishing, sailing, and observing animal and plant life. Therefore, for human beings, the availability of clean and uncontaminated freshwater habitats in large numbers should be of high priority.

Freshwater is also required for irrigation, in many industrial processes, and for the transportation of waste products. Because of the magnitude of these uses, the pressure on groundwater and surface water is immense in regions with dense populations and intense agriculture and industry. Unfortunately, many of the uses deteriorate water quality, thereby affecting water use for other purposes. The growing number of uses and users generates conflicts among neighboring countries and different user interests, and further exacerbates the dilemma of rising demand and the need and desire for high-quality aquatic habitats. Overall, freshwater ecosystems have been more restricted, manipulated, and polluted than any other ecosystem on Earth, and conflicts will intensify in the future in the light of growing human populations and contemporary environmental trends.

The Small, but Numerous Freshwater Habitats

Earth has much water. A deep, continuous ocean covers 70% of Earth's surface and contains 97% of all water. Only 0.65% of all water is found as freshwater on the continents, and most of this is groundwater (0.62%). Even smaller proportions are found in streams (0.0001%) and lakes (0.017%).

Lakes and streams cover a variable proportion of the land surface. In dry regions, surface waters cover less than 0.1% of the land surface. In wet environments such as the tundra, the boreal forest, and the rainforest, plenty of water is found in shallow pools, lakes, and streams, which may occupy 5–10% of the land surface. A map of Great Britain reveals that surface waters cover less than 0.1% of the land in the south and more than 5% in the rainy uplands of Wales and northwest Scotland; the average water coverage is 1–2%.

Because inland waters are mostly small, shallow, and numerous, they form an intimate contact with the terrestrial environment. This contact has been essential for the historic and contemporary development of species and life stages of amphibious plants and animals. Despite the small volume of lakes and streams compared to the ocean's volume, freshwater environments have been a vital platform for the evolution of many lines of algae, plants, and animals.

It is possible to quantify the contact between terrestrial and freshwater environments. In a small country like Denmark, inland waters occupy 1.7% of the terrestrial area, and on average, 1 km^2 of land area is in contact with 3.6 km shoreline of freshwater lakes and streams and 0.2 km of the sea. Before agriculture removed numerous shallow lakes, pools, and streams, the freshwater contact was 2–3 times longer.

The greater land contact to freshwaters than to the sea is even more striking for the world, because land areas are joined in large continents. As land masses grow in size, the contact zone to the sea increases only with the periphery and, thus, with the square root of the land area, whereas the contact zone with inland waters approximately increases in proportion to the land area. For the continents, the estimated contact zone is 100–1000 times longer between land and freshwater than between land and sea.

The transition between land and freshwater is gradual, gentle, and suitable for organisms, because physical forces and disturbance are weaker than in the transition from land to sea. The transition zone from small lakes and streams to land has a closed vegetation cover; that of the oceanic coasts faces strong winds, waves, and moving sand with only scattered vegetation. It has been much easier for organisms to cross between land and freshwaters than between land and the sea. The many new freshwater bodies that are formed during glaciations or appear transiently during wet periods also offer opportunities for the

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00369-5

emigration and development of new species without the intense competition and predation caused by well-established species, as is typical of the sea. In contrast, shallow lakes disappear after only a few thousand years, because they are filled with particles eroded from the land and organic matter produced in the lake. In essence, freshwaters provide many opportunities, which constantly come and go, but individual lakes lack the long-term stability of the sea that has lasted for billions of years.

This scenario should have stimulated the selection of species with rapid evolution and efficient means of dispersal allowing them to colonize new freshwaters as their original habitats disappear. Indeed, apart from species of fish, molluscs, and crustaceans associated mainly with large ancient lakes (e.g., Lakes Baikal, Titicaca, and Victoria) and large ancient river systems (e.g., the Danube, Amazonas, and other tropical rivers), most freshwater species are widespread within and even among continents. Freshwater microorganisms among the bacteria, algae, and protozoans are generally both locally abundant and cosmopolitan. Most aquatic rooted plants are also widespread, though a few species are endemic because underwater dispersal restricts their spread.

Species Evolution and Richness in Freshwater

As a result of the intimate contact, the long coastline, and the suitable transition zone, there has been a lively exchange of species from freshwaters to terrestrial environments and back again. Plants have evolved from a special group of freshwater green algae (Charophyceae, Coleochaetales) and diversified under the highly variable conditions on land. Insects have evolved from groups of arthropods in the transition between freshwater and land. Amphibians and reptiles have evolved from special fish in freshwater and brackish wetlands. Many evolutionary lines have been followed at different times and places.

Freshwaters have formed a corridor for the two-way dispersal of organisms between land and sea. Freshwaters also share many taxonomic groups of algae, plants, and animals with the terrestrial environment and the sea, whereas divergences are stronger between the land and the sea. Freshwater environments are surrounded by large surface areas of terrestrial and oceanic environments, which promote the emigration of marine and terrestrial species and their adaptation to freshwater environments.

Among major groups of land plants, several have secondarily returned to freshwaters, though few have reached the sea. Among liverworts, true mosses, bog mosses, horsetails, and ferns, there are many freshwater species, but no marine representatives. Among flowering plants, 1000–1400 truly aquatic freshwater species have been described from many plant families, illustrating that the secondary return from land to freshwater has occurred independently and repeatedly. The return process continues today, with more than 4000 species living an amphibious double life in the transition between land and freshwater. The sea includes only about 60 species of flowering plants with a restricted taxonomic diversity.

The diverse freshwater insects (>45,000 species) live a double life as eggs, larvae, and pupae in water and as flying imagines dispersing and mating on land. Mayflies live for only a few days on land without taking food, whereas adult dragonflies live a long and active terrestrial life. Freshwater insects, therefore, require a suitable environmental quality in both freshwater and adjacent terrestrial environments for their sustained survival and development, though most evaluations of habitat quality in relation to species composition and richness of insects focus only on the aquatic zone.

No phylum – the highest taxonomic entity of organism below the kingdom – is found exclusively in freshwaters, but all large phyla apart from the echinoderms are represented. Virtually all types of photosynthetic organisms, including cyanobacteria, the great variety of algae, mosses, and vascular plants, are represented in freshwaters (**Table 1**). In contrast, the sea has very few species of gold algae, yellow–green algae, plants, and insects, and it lacks some groups of green algae (e.g., desmids and stoneworts). The sea outnumbers freshwaters only with respect to photosynthetic species among red algae, brown algae, and dinoflagellates.

The three main determinants of species richness for a specific type of environment are surface area, habitat heterogeneity, and the time history in terms of durability and evolutionary development. Considering that freshwaters cover only a few percentage of the surface area relative to that of the sea and the terrestrial environment, and considering that inland surface waters mostly have a short life time, the freshwater biodiversity is surprisingly high and must arise from the high habitat variability among individually confined water bodies and the suitable conditions for emigration and establishment of species from the large, adjacent terrestrial and marine environments.

Taking the vertebrates as an example, there are about 10,000 named species of freshwater fish and 15,000 species of marine fish. Virtually all birds and mammals depend on freshwater bodies for drinking, but a large proportion of the species also are dependent on freshwater lakes, streams, and wetlands for breeding and feeding. In Europe, about 25% of the bird species and 11% of the mammal species live their entire life or part of it in freshwaters.

Table 1 Estimated number of algal species in freshwaters and total number of all species in soils, freshwaters, brackish waters, and the sea[a]

Taxon	*Freshwater species*	*All species*
Cyanophyta (blue–green algae)	1500	2000
Rhodophyta (red algae)	150	5000[c]
Chrysophyceae (gold algae)	190[b]	2000
Xanthophyceae (yellow–green algae)	550[b]	600
Bacillariophyceae (diatoms)	5000	10,000
Phaeophyceae (brown algae)	Few	2000[c]
Cryptophyta	100	200
Dinophyta (dinoflagellates)	200	2000[c]
Euglenophyta	700[b]	800
Chlorophyta (green algae)	7000[b]	8000
Zygnematophyceae (e.g. desmids)	6000[b]	6000
Charophyceae (stoneworts)	400	400

[a]Values are presented for major taxonomic groups.
[b]Freshwater groups.
[c]Marine groups.

Species in Freshwater Sediments

A global overview of organisms associated with freshwater sediments yields approximately 175,000 described species (Table 2), but the true number of species is much higher than this. The most speciose groups in freshwater sediments are insects, nematodes, and crustaceans. Among nematodes and rotifers living between the particles in freshwater sediments, there are probably many thousands of undescribed species.

Compilation of local species richness and taxonomic diversity (see Table 2) yields equally high, or often even higher, values in freshwater environments than in marine or terrestrial environments. There are apparently no systematic differences in local species richness between lakes and streams, but upstream sections of streams lack the variety of phytoplankton, zooplankton, and fish that are typical of open waters, and instead support higher species richness in the surface sediments and in the deep sediments below the streambed. The downstream parts of streams have a greater resemblance to lakes, so the similarity of species composition and richness will depend on whether entire stream systems or just certain stream sections are compared with the lakes. The biota in porous groundwater is much deprived in species due to lack of light, degradable organic matter, and dissolved oxygen. Specialized species of protozoan and small invertebrates live here in local species numbers that are typically 10–20-fold lower than those encountered in lakes and streams (see Table 2).

Comparisons of species richness of plants and macroinvertebrates (>1 mm) for a standard area typically show higher local and regional diversity of rivers than streams attributed to the greater habitat complexity in larger watercourses (Table 3, Williams *et al.*, 2003). Including other water bodies in the comparison for an 80 km^2 English region reveals that ponds per standard area have the second highest and ditches the lowest local species richness compared to rivers and streams. Regional species richness and the presence of rare species, however, are higher for ponds than for the other water bodies (Table 3). Ponds have small catchment areas and highly individual physical–chemical environments depending on local geology and land use. Rivers have extensive catchments, and local variations in soils chemistry are homogenized by the continuity of the flowing water. The ease of dispersal through rivers and streams should also lead to more uniform vegetation and fauna, whereas ponds are relatively isolated and should display greater community heterogeneity due to stochastic processes of colonization and extinction.

Table 2 Number of species described globally from freshwater sediments and the typical local range of species in lakes, streams, and groundwaters

Taxon	*Global*	*Lake/stream*	*Groundwater*
Bacteria	>10,000	>1000	>100
Fungi	600	50–300	0–10
Algae	14,000	0–1000	0
Plants	1000	0–100	0
Protozoans	<10,000	100–800	0–20
Crustaceans	8000	25–150	5–60
Insects	45,000	50–500	0–10
Molluscs	4000	0–50	0–10
Other invertebrates	>12,000	30–70	5–70

Source: Data compiled by Palmer MA, *et al.* (1997) Biodiversity and ecosystem processes in fresh-water sediments. *Ambio* 28: 571–577.

Human Impacts on Freshwater Ecosystems

When evaluating the human impacts on freshwaters, this article tends to concentrate on the numerous examples of water pollution. Over the past 150 years, the environmental issues have gradually changed as new problems have appeared. Since the mid-1800s, organic pollution of streams and lakes with organic wastes from households and domestic animals has been of major concern in Europe and North America. Since 1945, the stronger focus has been on cultural eutrophication of inland and coastal waters with nitrogen and phosphorus

Table 3 Site characteristics and ranges of nitrate and total phosphorus in rivers, ponds, streams, and ditches of the 80–km^2 River Cole catchment, England For macroinvertebrates and plants local species richness (mean and range), total number of species and number of unique species are shown

Variables	*Rivers*	*Ponds*	*Streams*	*Ditches*
Site definition	Width > 8.25 m	Area: 20–20,000 m^2	Width < 8.25 m	Human–made
Length or number	17.3 km	65 sites	28.8 km	76.6 km
Nitrate (mg N l^{-1})	3.2–9.7	0.1–38	2.3–19.8	0.4–17.7
Total phosphorus (mg P l^{-1})	0.06–1.30	0.002–2.49	0.05–5.7	0.002–0.88
Local species richness				
Macroinvertebrates	45.3 (18–66)	32.6 (5–69)	18.7 (3–50)	12.9 (2–35)
Plants	10.7 (6–19)	10.1 (2–17)	7.3 (1–17)	6.1 (1–14)
Total species number				
Macroinvertebrates	152	173	124	90
Plants	49	67	39	30
Number of unique species				
Macroinvetebrates	26	50	6	8
Plants	9	24	9	11

Source: Adapted from Williams P, *et al.* (2003) Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in southern England. *Biological Conservation* 115: 329–341.

from agriculture, towns, and industries. After 1960, acidification of low-buffered inland waters came onto the agenda due to increasing concentrations of sulfuric and nitric acids in the precipitation and changes in land use. The latest chemical concerns include trace metals and an enormous range of synthetic organic compounds of largely unknown behavior and ecological effect.

Many physical changes in the catchments are, however, more important to the existence, environmental quality, and biodiversity of surface waters than the direct water pollution. The most significant influence on terrestrial and freshwater environments is the removal of natural vegetation and the cultivation of land, which lead to immediate, profound alterations of the hydrology and nutrient cycling. These alterations are grossly enhanced when soils are drained and streams are canalized. The intimate linkage between natural wetlands and streams, which has been important for the evolution and contemporary diversity of plant and animal life, is also disrupted when surplus water directly flows to the stream through drain pipes rather than slowly percolating through the wet, sponge-like organic soils.

Most tropical countries now face the unregulated impact from all pollution sources. The most widespread problem appears to be the rising organic pollution of streams and lakes by untreated domestic sewage. However, in some regions, heavy pollution takes place from (1) extensive application of pesticides and nitrogen in plantations of bananas, cocoa, and oil palms, (2) large oil spills (e.g., Ecuador, Venezuela, and Nigeria), (3) acidification and pollution with heavy metals (e.g., copper and mercury) from mining areas and tanneries, and (4) outlet of phenolics from the timber industry. A fundamental problem is the erosion of soils deriving from the massive deforestation of tropical rainforests. Many streams have become chocolate colored from the heavy load of solids eroded by the strong rainfall. Erosion is a severe problem in agriculture, in which the fine, fertile top soils are lost. It also causes environmental problems as hydroelectric dams are quickly filled with sediments and when suspended loads in streams prevent light penetration and thereby the growth of algae, which are food for many invertebrates and fish. Furthermore, stream bottoms get clogged with fine sediments that destroy the habitats of invertebrates and the spawning banks of fish.

In many tropical countries, physical alterations of streams are still relatively few. In Ecuador, for example, streams still run in natural streambeds with meanders and rapids surrounded by strips of riparian vegetation even in densely populated areas. The hydrological cycle has been altered, however, because water is removed for irrigation, leading to artificially low discharge and current velocity during the dry season. As a consequence, the stream biota is impoverished due to oxygen depletion and smothering of the streambed following organic pollution.

Restriction of Area and Variability of Freshwater Habitats

The scene of physical changes and areal reductions of freshwater ecosystems changes among countries and with time. Changes of freshwater ecosystems have been greatest over the past 100–200 years in densely populated countries and in regions with intense agriculture and industry. Thus, the combination of population density, resource use, and powers of technology is a suitable measure of human impact on the biosphere in general and the freshwaters in particular.

Petts (1984) defined four phases in the recent era of river modification in Europe. Phase 1, from 1750 to 1900, includes ambitious regulations of the major rivers for the purpose of navigation, flood control, and cultivation of the river valley. Phase 2, from 1900 to 1950, marks the first major technological period, during which large dams and power plants were built across major rivers. In many European lowlands, extensive drainage of wetlands and shallow lakes and channelization of streams took place, and continuous management of the dimensions of stream channels and cutting of aquatic plants were initiated. Management intensified in phase 3 from 1950 to 1980 by the use of specialized machines. During the recent phase 4, from 1980 onward, the intensity of regulation works and dam building has gone down, because most watercourses have already been exploited and public resistance has increased because of rising environmental concerns. However, most countries of the world are still in the most exploitive phase 3.

An overview of the 69 major rivers in Europe and Russia shows that most of them are strongly (40) or moderately (10) affected by dams and regulations, whereas only 19 rivers located in Arctic and northern boreal regions have remained relatively unaffected. River regulation has been undertaken to the greatest extent in Western and Southern Europe. In Belgium, Denmark, England, and Wales, the percentage of river reaches that are still in a natural state is less than 20%. In Estonia, Norway, and Poland, the rivers still have 70–100% of their reaches in a natural state.

In southern Canada and the United States, rivers are heavily regulated; those in northern Canada and Alaska have remained relatively pristine. The history of the Willamette River in Oregon is an example of how expansion of agriculture and construction of 11 dams has transformed a complex multichannel river into a simple one- or two-channel river. Over a 25-km-long stretch in the floodplain, the length of the shoreline has declined from 250 km in 1854 to only 64 km in 1967.

In addition to the shortening of rivers and streams, watercourses have become more uniform and physically disturbed. When streams are channelized, the natural variability in depth, width, current velocity, and sediment composition disappears between straight reaches and meanders and between riffles and pools. The natural dynamics of the flow channel, characterized by spatial variations of erosion and sedimentation, formation of new meanders, and the cut-off of oxbow lakes, also vanishes. Drainage of the floodplain reduces its storage capacity of water and the ability to buffer storm surge in the rivers. Moreover, drainage reduces baseflow during dry periods. Peak discharge during rainy periods introduces high physical stress and erosion in the stream channel, and low baseflow during drought increases the risk of high water temperatures, oxygen deficiency, and insufficient currents for the well-being of invertebrates and fish. All three major changes – (1) smaller area and length of the streams, (2) lower habitat diversity, and (3) greatly fluctuating flow and

higher physical, oxygen, and temperature stresses – have drastically reduced species number and population density of plants and animals.

Numerous shallow lakes and ponds in European farmland have disappeared due to drainage and filling-in. Many ponds were originally dug between 1800 and 1930 to extract calcareous sediments for the liming of acidic fields, clay for bricks, and sand and gravel for construction works and to provide watering places for cattle. In the United Kingdom, 1.2 million ponds present in 1880 declined to 0.3–0.4 million in 1990–2000. In Denmark, about 40% of lakes larger than 10 ha disappeared from 1700 to 1960 and several regions lost all large lakes. After 1980, one-third of this number have been formed as new or reestablished lakes. Similar to ponds in the United Kingdom, about 70% of the Danish ponds disappeared between 1884 and 1974. Detailed Danish studies of 2500 ponds revealed that this decline has continued recently as 40% of the ponds present in 1940 had disappeared in 1990 and only 5% had been reestablished. The net decline was larger (65%) for localities inhabited by common frogs and newts, and the decline was almost 95% for localities inhabited by rare and less robust species. The remaining ponds have been deteriorated by organic pollution and eutrophication, and retracting populations of rare frogs and newts have faced reduced dispersal that had to be overcome by intentional spread. Rare species have survived 2–3 times better in ponds that have been cleaned for nutrients and excessive plant growth compared with ponds that have not.

Large, deep lakes have experienced less physical disturbance and restriction of areal cover than streams and ponds. Exceptions are artificial reservoirs, which face great fluctuations in water level over the seasons. In Europe there are now more than 10,000 reservoirs, and about 4000 large reservoirs have dams higher than 15 m. Dams have broken the natural connection and dynamics of watercourses for the movement of water, solutes, and living organisms. However, reservoirs have generated habitats for lake organisms and new recreational opportunities.

Changing Direction: Restoring and Creating New Freshwater Habitats

Although areal reduction and the manipulation of inland waters continue in most countries, public and political priorities have changed in some wealthy countries with the goal of ameliorating some of the environmental damages caused by intensive agricultural and industrial practices.

Norway experienced serious demonstrations when a hydroelectric power plant was built across the former pristine Alta River. In Spain, the dam across the Esla has remained unused, because the inhabitants of the valley to be expropriated have refused to move. In Iceland, local "farmers" stopped early attempts to build a dam across the famous Laxa River. In Denmark, new laws now give higher priority to the development of diverse landscapes, diverse biological communities, and trout fishing. After water quality had been improved following the extensive purification of domestic sewage, new management procedures and structural changes have been implemented in many stream reaches to improve habitat variability, to prevent catastrophic disturbance, and to strengthen the linkages within the stream course and between the stream and the floodplain. These Danish initiatives include the following: (1) removing all weirs and dams blocking the natural migration of animals; (2) bringing the streams that had been directed underground in culverts to the open surface; (3) reestablishing the variation in width, depth, and substratum between riffles and pools; (4) re-creating the former meandering patterns of straightened channels; and (5) establishing strips of natural vegetation free of pesticide spraying along the streams. Populations of brown brook and sea trout have benefitted strongly from these initiatives leading to an increase in the mean density of naturally produced trout fry per 100 stream sections in some regions from less than 1 in 1967 to more than 50 in 2008. Another initiative in northern Europe has been to restore shallow lakes and ponds that had been drained or were facing disappearance due to lack of water, filling-in, or overgrowth by littoral vegetation. Restored shallow lakes have developed a rich bird life in cases where high habitat heterogeneity has been obtained by securing variable depths, small protected coves, and islands unavailable to foxes. Large areas with shallow waters (0–30 cm) have benefitted high densities of waders. New ponds have been created in farmlands with the purpose of regaining the patchwork of small, suitable biotopes for aquatic plants, insects, and amphibians. Despite successful establishment of new ponds with amphibians, this has not compensated for the overall decline of ponds and populations of rare amphibians in Denmark, though it has prevented the extinction of some of the rarest species. The diversity of plants and macroinvertebrates in agricultural ponds has shown greater sensitivity to eutrophication than to terrestrial habitat destruction around the ponds or to an invasion of alien species. Reduction of biodiversity, however, can be predicted more accurately by combining all three environmental stressors.

Changes in the Hydrological Cycle

It is often overlooked how much the hydrological cycle has been changed by humans and the environmental implications of these alterations. Archeology, paleolimnology, and historical studies of old maps and written records in Europe reveal that the groundwater level was much higher and there used to be many more open waters in the past than today. Previously, water moved slowly and diffusively through the large water reservoirs in the soil and the wetlands before reaching streams and lakes. As a result of cultivation of soil and establishment of ditches and drains, water has taken a more direct and rapid course from the fields to the streams, resulting in huge temporal variations in discharge. Water for irrigation of crops and domestic supply to metropolitan cities is abstracted from groundwater and surface water, and this water demand has markedly increased the risk of wetlands drying up and critically low flows developing in the streams during dry periods, thereby diminishing all components of biodiversity. Because of low summer flow, about 10% of Danish streams are unable to attain the projected biodiversity; in dry regions this percentage is much higher.

Breaking the long contact between water and soil has dramatically reduced the natural purification of water by preventing the removal of surplus nitrogen and phosphorus through binding to soil particles, incorporation in new organic material, precipitation as phosphate minerals, and denitrification. In addition to the increased cultivation of crops, application of fertilizers, and discharge of sewage, changes in the hydrological cycle have resulted in the explosive eutrophication of inland and coastal waters. A few countries have tried to reestablish some of the wetlands that have been lost so extensively, thereby reinstalling the cleaning sponge between cultivated soils and streams. In most other countries, however, wetlands continue to disappear at a high rate. Spain, for example, has lost two-thirds of its wetlands since 1965 with the support of European Union subsidies to farmers.

Organic Pollution of Inland Waters

The discharge of organic wastes from towns and from livestock operations creates the classic pollution problem of inland waters. Human wastes are derived from bathing, cooking, laundry, and the flushing of feces and urine in lavatories, and discharges from agriculture include manure from the animals and food spills. Intense organic pollution can also occur from breweries, dairy factories, slaughterhouses, sugar refineries, and countless other industrial sources. In regions where water quality otherwise is high, the farming of fish, prawn, and crayfish can be common. The trout farms of Northern Europe mainly use cold spring water of superb quality, but these farms in turn pass on substantial organic pollution from food spills and fish excreta.

With the increase in domesticated animals, human populations, and the installation of water closets and sewers, there has been a profound increase in organic loading of inland waters in Europe and North America from 1850 to 1950. In several cases, the external loading of lakes has increased more than 10-fold (**Figure 1**). Substantial purification of domestic sewage has been established over the past 60 years. Meanwhile, wet slurry from burgeoning animal farms has presented a new source of pollution. In poor countries of the world, organic pollution has continued to grow and few attempts have been made to ameliorate the problem through the use of natural wetlands with selfpurification capacities or the construction of costly sewage treatment plants.

The organic pollution of streams increases with the density of human beings and animal livestock and the consumption of oxygen by organic matter in the water. An acceptable water quality can be reached by extensive purification, as shown by a 12-fold reduction of the biological oxygen demand at the outlet from sewage treatment plants and a 3–4-fold reduction in the stream water in Denmark from 1970 to 2000. Otherwise, the European inland waters with the low oxygen demands and highest oxygen concentrations are found in the water-rich regions in the north, where population densities are low and large areas remain uncultivated. In contrast, streams with the poorest water quality are located in intensely cultivated regions of Middle and Southern Europe, where purification remains insufficient and water is in short supply during the summer. To illustrate the differences in environmental

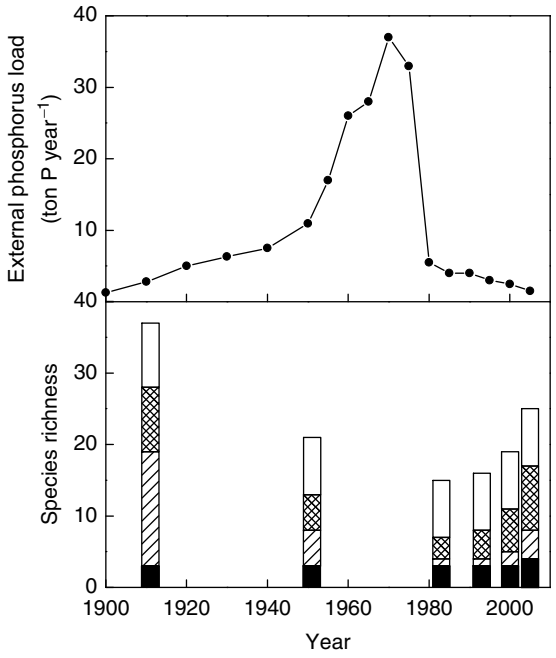


Figure 1 Historical development of eutrophication in Lake Fure in Denmark from 1900 to 2005. Above: Annual phosphorus loading. Inhabitants in the watershed increased 10-fold over the period. Water closets and sewers were installed in the 1930s, detergents applied in the 1950s, and sewage was diverted or underwent tertiary treatment after 1970. Below: Number of submerged macrophytes among macroalgae (black), mosses plus characeans (hatched), small (gray) and large rooted plants (white) on five occasions. Adapted from Sand-Jensen K, *et al.* (2008) 100 years of vegetation decline and recovery in lake Fure. *Journal of Ecology* 96: 260–271.

state in the 1990s, oxygen conditions were regarded as good in 97–99% of the streams in Scotland and Iceland and 64–77% of the streams in Wales, but in less than 20% of the streams in Belgium, Poland, and countries in the Balkans. Most European countries have experienced an overall decline of organic pollution in the recent decades.

Agricultural and Industrial Chemicals

An immense variety of chemical products are being manufactured, used, and released by agricultural and industrial activity. Some inorganic compounds include acids, alkalis, ammonia, chlorine, radionuclides, and heavy metals (e.g., cadmium, copper, iron, mercury, and zinc). Organic compounds are grouped under different names such as chlorinated hydrocarbons, hydrocarbons, pesticides, phthalates, and phenolic compounds. Perhaps 20,000–50,000 substances are manufactured or applied within industrial countries, and a few thousand are added each year. In most European countries, between 120 and 530 active pesticides are approved for agricultural use today. The annual usage is usually between 1 and 14 kg ha^{-1} of arable land. Substances that are applied within a country are also detected (or their degradation products) when tested for in surface waters and groundwaters. The most common pesticides in groundwater (i.e., atrazine, desethylatrazine, lindane, and simazine) are often found in

concentrations exceeding the maximum allowable threshold (0.1 $\mu g\, l^{-1}$ in the European Union).

A major obstacle to pollution control is the need to first recognize the presence of potentially harmful substances in potentially harmful concentrations in relevant ecological situations. It requires a lot of money, appropriate chemical skills, and advanced analytical methods to conduct an adequate survey of the distribution and concentrations of just a small number of these trace organic chemicals. It is then even harder to evaluate the biological consequences under natural conditions. As more traditional and obvious pollution problems in developed countries are stabilized or reduced, these organic substances may perhaps become the key pollution problem in many water bodies. The pollution effects may become more apparent as the application of these new organic compounds spreads and they accumulate over time in groundwater. A Danish overview suggests that biodiversity is significantly reduced by industrial chemicals in 2% of the streams, whereas 4% of the streams are severely impacted by outlets of unknown composition. Because their biological effects are subtle, chronic, and extremely costly to verify in the complex blend of numerous compounds, living organisms, and environments, the only effective solution to the pollution problem is not to discharge the industrial pollutants at all, or reduce their toxicity, magnitude, and rate of application as far as possible. In less-developed countries, it will be impossible to monitor the release and biological effects of these trace organic compounds. Some frightening pollution events have been reported from the heavy misuse of pesticides in tropical crop production.

Cultural Eutrophication

Cultural eutrophication – predominantly due to increasing loads of nitrogen and phosphorus – leads to profound changes in the composition, biomass, and productivity of algae and plants. Lake eutrophication results in phytoplankton blooms, untransparent water, and oxygen deficiency. Eutrophication spoils the quality of bathing water and threatens the survival of bottom animals and fish. Algal blooms can include toxic algae and thereby harm animal life and become a public health risk.

Eutrophication of streams can also enhance the growth of attached macroalgae and flowering plants. Most European lowland streams have long passed the threshold at which nutrient concentrations limit plant growth. Streams in sparsely populated regions of Northern Europe and Canada can still hold such low concentrations that plant growth is enhanced at sites of elevated nutrient input. Stimulation of algal growth by eutrophication is also important in most tropical streams.

Nutrient input to watercourses has increased dramatically during the past 150 years, and it has intensified over recent decades. The sources of nitrogen and phosphorus input include (1) towns and industries, (2) scattered settlements, (3) agriculture, and (4) a background input deriving from precipitation and runoff from uncultivated areas. Input from towns and agriculture usually dominates the overall nutrient budget, but all four sources have increased because of anthropogenic impacts.

The classic examples of lake eutrophication have been documented in the vicinity of metropolitan towns. In the case of Lake Fure close to Copenhagen, annual phosphorus input increased 30-fold from 1900 to 1970 due to a 10-fold increase in the population density in the catchment, the installation of sewers, and the use of phosphorus-rich detergents (**Figure 1**).

Agriculture is the other major source of nutrient loading. Paleolimnological studies document the increase in accumulation rates of mineral particles, phosphorus, and organic matter in lake sediments due to increased erosion and runoff following the early cultivation of watersheds. Though nutrient input with domestic sewage has recently declined in several industrial countries, thanks to tertiary sewage treatment, agricultural sources of nitrogen have either continued to rise, leveled off, or declined slightly, but, nonetheless, occupy an increasing proportion of both nitrogen and phosphorus because of the heavy application of fertilizers and import of food to pigs and cattle (**Figure 2**). A large proportion (60–70%) of nitrogen applied as food and fertilizers in efficient industrial agriculture production is lost to the environment, emphasizing that highly productive agricultural systems inevitably lead to profound eutrophication of surrounding natural habitats. Nitrogen is mostly lost from agricultural fields as organic matter and nitrate in accordance with the 2–6-fold increase in nitrate concentrations in a series of English streams and groundwater reservoirs from 1920 to 1990.

An assessment of European streams reveals that nitrate concentrations have increased from 1970 to 1990 because of

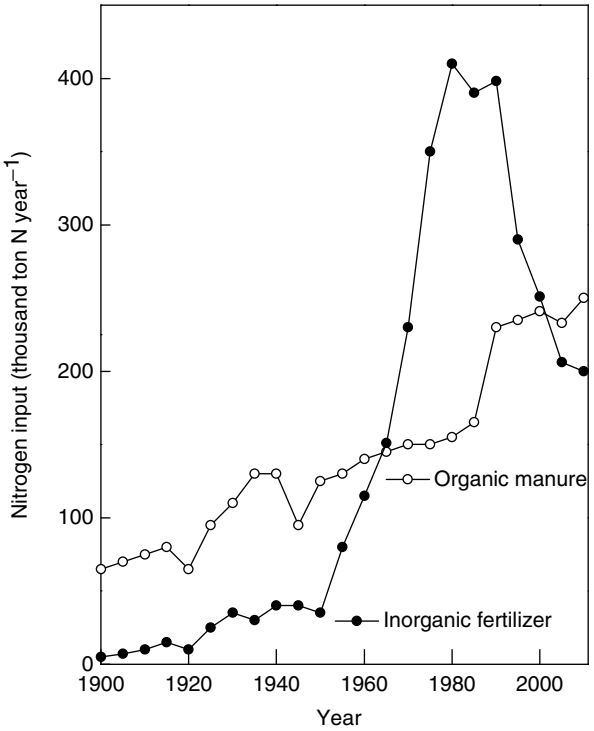


Figure 2 Historical changes in the annual nitrogen input in fertilizer and organic manure to Danish farmland between 1900 and 2010. Reproduced from Journal of Ecology 96: 260–271.

the strong agricultural input, whereas phosphate concentrations have declined, thanks to better cleaning of domestic and industrial sewage (Kristensen and Hansen, 1994). With unregulated application of phosphorus fertilizers, phosphorus pools in agricultural soils have increased, and phosphorus will eventually get lost to inland waters by erosion of soil particles and by leaching as the binding capacity to soil minerals is surpassed. This development represents a new risk to the control of lake eutrophication in countries attempting to control phosphorus through tertiary treatment of domestic sewage. In streams exposed to strong human impact, median concentrations of nitrogen and phosphorus are typically 10-fold higher than in the few pristine streams in mountain ranges and forest regions. However, even oligotrophic waters in uncultivated areas have experienced a profound enrichment of nitrogen through increasing atmospheric deposition up to 1980. Atmospheric deposition of nitrogen has declined between 1980 and 2010, but still remains markedly above pre-industrial levels.

Anthropogenic Acidification and Recovery

Anthropogenic acidification of precipitation and surface waters commenced with the Industrial Revolution, but accelerated in Europe and North America particularly after the 1950s. Burning of fossil fuels releases sulfur oxides, which are converted to sulfuric acid in the atmosphere. High combustion temperatures in car engines and power plants release nitrogen oxides, which are converted to nitric acid. Annual wet deposition of sulfur over England increased five-fold from 1850 (approximately 5 kg S ha^{-1}) to 1977 (25 kg S ha^{-1}), but afterward has declined steeply to levels in 2010 resembling those in 1850 because of the use of cleaner fuel gases and cleaning of the emission from power plants. Also nitrate deposition, contributing to acidification, increased from 1850 (1–2 kg N ha^{-1}) to 1977 (7–8 kg N ha^{-1}) and has declined since then. Deposition has followed the same time course throughout Europe and the United States, though rates are highly variable among regions. The initial increase from 1860 to 1980 was much greater in Norway and Sweden (more than 10-fold) than in the United Kingdom, and the decline from 1985 to 2008 has been greater, reaching 70% for sulfur and 20% for nitric acid.

Acidification of inland waters has been substantial in the northeastern United States and southeastern Canada. Acidified regions of Europe include large parts of Finland, Norway, and Sweden; mountains in Great Britain, Poland, and the Czech Republic; and carbonate-poor, well-leached sandy soils in Belgium, Denmark, and the Netherlands. In Norway and Sweden, acidification has represented a significant national problem as thousands of inland waters have been acidified and lost their resident invertebrate species and fish. Vast forested areas have also been acidified to the extent that species diversity and elemental cycling have been grossly disturbed.

With the 70% reduction of sulfur emission during the past 30 years, acidification has been reverted in many European softwater lakes. As a typical Swedish example, sulfate in Lake Bolmen peaked in 1975–1977 (15 mg l^{-1}) concomitant with the lowest recorded acid-neutralizing capacity (50 μeq l^{-1}) and water color (15 mg Pt l^{-1}). Since then, sulfate has declined by about three times and acid-neutralizing capacity and water color have increased three-fold. In the heavily affected Lake Saudlandsvatn, Norway, acid-neutralizing capacity has increased from −60 μeq l^{-1} in 1980 at pH 4.7 to 20 μeq l^{-1} at pH 5.7 in 2008. Before industrial acidification commenced, acid-neutralizing capacity was around 80 μeq l^{-1} and the estimated pH was 6.7, so full "chemical" recovery has not yet been attained (Hesthagen *et al.*, 2011).

Although the increase in acid-neutralizing capacity and pH accompanying the reduced deposition of sulfur from 1980 to 2010 is easy to explain, the nationwide increase in brown-colored humic substances, particularly in the formerly most impacted regions, has been attributed to changes in land use, climate, and acid deposition. The spatial patterns, however, suggest that increasing loss of humic acids from formerly heavily acidified soils have been driven principally by reduced acid deposition during the past 30 years. This has favored solubilization of humic substances that are formed continuously and, in part, had accumulated in the acids soils during 1960–1980. These mechanisms are particularly important in soils where bicarbonate buffer systems are weaker than organic acids, soils are wet and nutrient poor, and contents of organic matter and C : N ratios are high (Clark *et al.*, 2010). As sulfur deposition eventually approaches background levels, it is likely that temperature and hydrology will dictate the direction of future trends of dissolved organic carbon as they already dictate seasonal changes.

The biological effect of acidification is in part due to the direct reduction in pH, which is intolerable to many organisms most notably bicarbonate-using plants, fish and crustaceans, mussels, and snails with exoskeletons of carbonate (see **Figure 5**). Associated effects arise from an altered ionic balance, high metal concentrations, and immobilization of vital nutrients. It is, therefore, expected that these harmful effects are ameliorated and biodiversity reestablished with some delay following the recent rise in acid-neutralizing capacity and pH depending on the extent and duration of acidification of the individual lake and region. In addition to the physicochemical effects, recovery is influenced by the biological recruitment. Indeed, the massive growth of acid-tolerant aquatic mosses and the reduction of isoetid plants during acidification have been reversed during recent years in Danish lakes experiencing a pH rise from about 5.2 to 6.5. In Norway, brown trout disappeared from more than 8000 inland waters, and although the most highly acidified regions have yet to show signs of recovery, this is evident in lakes when pH rises to 5.5–6.0. Also species of daphnids and net-spinning caddisflies reappear at this pH when acid-neutralizing capacity exceeds 20 μeq l^{-1} (Hesthagen *et al.*, 2011).

The rising levels of brown-colored humic substances during alkalinization have caused concerns about carcinogenic byproducts of drinking water treatment and ecosystem effects. The profound humification inevitably leads to loss of submerged vegetation, altered prey–predator relations between fish species at reduced visibility, and low water quality for recreation. It remains to be seen whether brownification is a chronic problem or whether it will decline naturally during the coming decades.

Other conditions may contribute to the acidification of inland waters. Drainage and mining can expose metal sulfides (notably pyrite) to oxygen, starting one of the strongest acidifying processes that has sulfuric acid and ochre (iron oxyhydroxides) as the end products. Ochre pollution of streams leads to untransparent waters, smothering of sediments, and a steep decline in the diversity of plants, algae, macroinvertebrates, and fish. Eggs of trout and salmon suffocate from lack of oxygen due to reduced water flow through spawning banks, and adults suffer from starvation and toxicity of ferrous iron ($>1\,\mathrm{mg\,Fe\,l^{-1}}$). In streams draining pyrite-rich soils, pH may decline (3–4) and iron may increase to the extent that most multicellular organisms disappear.

Biological Quality of Freshwater Ecosystems

Anthropogenic impacts on freshwater ecosystems lead to loss of lakes and streams, changes in the hydrological cycle, physical disturbance, and water pollution. Streams have presumably experienced the entire range of physical and chemical influences, whereas major lakes mostly have experienced acidification, eutrophication, and pollution by organic matter and industrial chemicals.

The resulting changes in species composition, species richness, and elemental cycles in freshwater ecosystems are complex, have several causes, and can be evaluated at different levels. Major changes have taken place in almost every stream due to regulation, management, and pollution, and in many regions and countries, it is impossible to find pristine streams to use as a reference for baseline conditions and high biological quality.

The natural variability among lake types has dropped. Historic and paleolimnological studies of Danish lakes, for example, show that the former full range of lake types from oligotrophy to eutrophy has been restricted owing to widespread eutrophication such that there are very few oligotrophic lakes left. Though probably more than 80% of Danish lakes, belonging to the oligotrophic, mesotrophic, and eutrophic categories, had a rich submerged vegetation 100–150 years ago, less than 15% of all lakes today have sufficiently clear water to permit submerged plant growth. With the exception of Finland, Norway, Scotland, and Sweden, oligotrophic, clear-water lakes have become rare across the entire European continent.

Less variation in lake and stream types will restrict beta and gamma diversity and, in particular, reduce population densities of those species that prefer undisturbed, oligotrophic habitats. Likewise, in European terrestrial vegetation, many oligotrophic and disturbance-sensitive species characteristic of heathers, *Sphagnum* bogs, and nutrient-poor grasslands have become rare and are threatened by local or regional extinction. In many individual lakes and streams, profound alterations of species composition and abundance can be documented over time. Recent changes are small, however, if one evaluates only whether the individual species still exist within the national borders or have survived globally. A few refugia of relatively undisturbed lakes and stream reaches are still likely to exist, and they may support small populations of those species that otherwise have declined dramatically in geographic range size and local abundance. Species extinction is a slow process, and long before the last individual finally dies, the species has lost its significance in the ecosystems and in human perception of nature. The continued decline of *Potamogeton* species in regulated streams is a clear warning that extinction can continue although recent steps have been taken to improve water quality (see **Figure 3**).

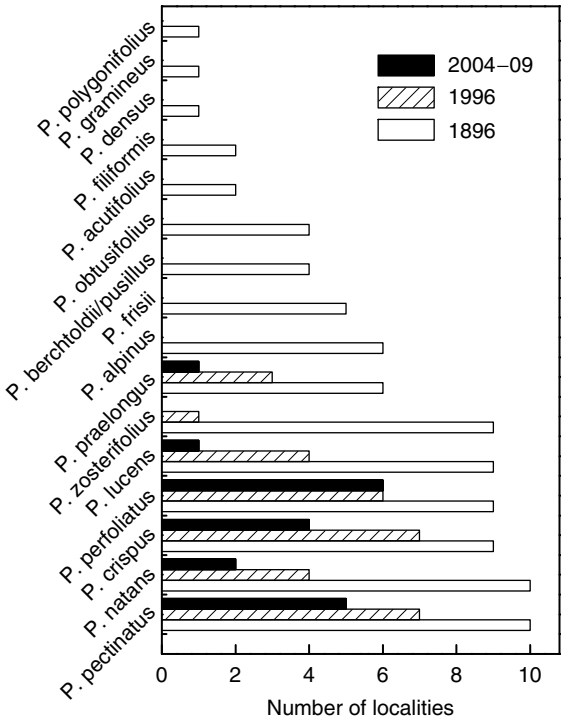


Figure 3 Historical changes in the number of Danish stream localities with different *Potamogeton* species during three studies of the same 13 localities in 1896, 1996, and 2004–2009. Adapted from Figure 1 in Riis and Sand-Jensen (2001), Freshwater Biology 46: 269–280, with permission from John Wiley & Sons.

Decline in Freshwater Species in Europe and North America

A large proportion of extinct, threatened, and rare species in Europe and North America live in freshwaters. Species in streams and ponds have been most strongly affected by human impact. In North America, 11–14% of the mainly terrestrial birds, mammals, and reptiles are among the extinct, threatened, and rare species (**Table 4**). Among freshwater amphibians, fish, crayfish, and mussels, the percentages are particularly high (28–73%). No less than 103 species of North American freshwater fish are classified as endangered, 114 are threatened, and 147 deserve special attention. These species represent about one-third of all fish species. Twenty-seven species have gone extinct over the past 100 years from habitat loss, chemical pollution, introduction of exotic species, hybridization, and overfishing (Allan, 1995).

Table 4 Percentage of species within selected groups of freshwater or mainly terrestrial animals on the Red List in North America (by 1990) and Denmark (by 1996)[a]

Group	*North America*	*Denmark*
Freshwater habitats		
Amphibians	28	36
Fish	34	39
Crayfish	65	–
Unionid mussels	73	–
Caddisflies	–	32
Dragonflies	–	42
Mayflies	–	50
Stone flies	–	40
Terrestrial habitats		
Birds	11	37
Mammals	13	30
Reptiles	14	25
Beetles	–	26
Butterflies	–	49
Moths	–	16

[a]Species on the Red List are recently extinct, threatened, vulnerable, or rare. Red Lists focus on particular vulnerable groups often highly valued by humans.

The European status of freshwater species is worse because of a long history of strong human impact. Europe is relatively poor in species of freshwater fish (approximately 250) compared with North America (approximately 850), Africa (1800), and the Amazon region (2000 species). Species numbers are particularly low in Northern Europe (e.g., 34 indigenous species in Denmark), whereas they are much higher in Middle and Southern Europe. The Volga and the Danube include 60–70 species each, or 25% of the entire European fish fauna. In European countries, about one-third of all fish species are on the IUCN Red List, much like in North America. In the global status, 20 European species are threatened, susceptible, or rare.

On both continents, many genetically isolated stocks of salmon and trout confined to certain stream systems are threatened by eradication, or they have already been lost from a variety of threats such as river regulation and construction of dams that prevent upstream migration and that destroy spawning grounds. Discharge of wastewater from agriculture, industry, and towns has also contributed to the loss, as have acidification in areas with poorly buffered waters, interbreeding with hatchery-reared individuals, and overfishing in the ocean.

The Atlantic salmon was once very common along its range of distribution from Iceland to Portugal. It has now disappeared from many major rivers on the continent, and the annual catch has dropped profoundly. The large salmon population in the Rhine supported an annual marketed catch of more than 100,000 individuals in the late 1800s. The species declined during the 1900s and went extinct in 1957, but was later reintroduced from an artificially reared stock. Atlantic salmon was lost from Danish rivers during the 1900s, apart from very small indigenous populations surviving in four larger rivers in western Jutland.

DNA studies of preserved salmon scales from the extinct and extant Danish populations have shown that the individual rivers had genetically distinct populations. These differences were much larger between populations in Denmark, Scotland, and Sweden than between populations in neighboring rivers. Thus, restocking with foreign salmon was abandoned, whereas initiatives were taken to secure and perhaps disperse the small, national salmon stock, which is probably better adapted to the local environmental conditions and food sources than are the foreign salmons. Although Atlantic salmon may still be part of the fauna in several countries of continental Europe, it has lost, at least for the moment, its role as an enjoyable catch and an important part of the ecology of the streams. Moreover, the original high genetic diversity among the many local populations of individual rivers is definitely gone. Nonetheless, systematic rearing of the small surviving salmon stocks from individual Danish rivers during the past 20 years has been highly successful and reestablished healthy, very substantial salmon populations that are now exploited by sportfishing.

Red Lists and Historical Development in Streams of Denmark and the United States

Red Lists focus on threatened, vulnerable, and rare species and usually include those species that have disappeared recently. Among the five studied groups of freshwater insects, the percentage of red-listed species of all national species in Denmark ranges from 32% for caddisflies to 50% for mayflies (**Table 4**). Freshwater fish have 39% and amphibians have 36% of all national species on the Red List. Among terrestrial insects, the percentages vary from 16% for moths to 49% for butterflies. Birds and mammals have 37% and 30%, respectively.

Overall, the percentage of species on Red Lists is high for three reasons. First, the most vulnerable groups are more often included than the least vulnerable groups. Second, many species are naturally rare. Third, many species are present in very small numbers, because natural habitats have experienced profound areal restriction and deterioration. Although the percentage of red-listed species tends to be higher for freshwater than for terrestrial species, this tendency disappears if more common groups of freshwater invertebrates such as dipterans, oligochaetes, and polychaetes were included.

Consequently, Red Lists have several weaknesses because they are selective and qualitative, and the intensity of search for rare species has increased over time. Red Lists do not give quantitative data on the abundance of species, and they can describe only the temporal development in crude ways, provided evaluations are repeated at suitable intervals.

Many evaluations of species development, however, suffer from the lack of suitable historical description of species distribution and abundance. If such studies indeed exist, they did not use exactly the same methods and survey intensity and are therefore open to critique, even though differences may be very profound and without reasonable doubt are real.

A historical analysis of *Potamogeton* (a large, aquatic plant genus) in 13 localities in Danish lowland streams, for example, revealed the existence of 6.7 species per locality and 16 species altogether in 1896 (**Figure 3**). In 1996, the mean number was 2.8 species per locality and only seven species were present. The decline has continued, resulting in only 1.9

species per locality and six surviving species in 2009, which implies that recent improvements in water quality and managements have been insufficient or recolonization of species from undisturbed river or lake refugia is no longer possible. Several of those species that have disappeared from the 13 localities have become extremely rare throughout the country. The same type of development has taken place in other lowland regions of northwestern Europe such that some species (e.g., *Potamogeton acutifolius*, *Potamogeton filiformis*, and *Potamogeton zosterifolius*) have become rare over wide geographic areas. The *Potamogeton* species that are still common are fast-growing eutrophic species of high dispersal capacity that have best been able to cope with the high disturbance and nutrient richness of contemporary streams. It is noteworthy that plant species in streams display an overall positive relationship between geographic range size and local abundance, resembling the pattern for terrestrial herbs and animals. Species of low geographic range size and low local abundance face a double jeopardy of extinction because they grow in just a few places and are infrequent at sites where they do occur. Stochastic loss of some habitats and degradation of others should therefore have a strong impact on these species, because their few and small populations make them susceptible to further losses and reduce their ability to disperse to new suitable places. In contrast, widespread species of high local abundance have a double security, because they have a higher probability of surviving stochastic changes and spreading to new habitats.

The channelization of streams has been shown to have a significant influence on riparian plant assemblages by reducing local diversity associated with a steeper bank slope and reduced flooding. Moreover, undisturbed streams are also spatially more heterogeneous and richer in submerged species than streams that are cut regularly.

Freshwater insects have undergone a similar decline of species richness, number of ecological types, and number of occupied habitats. About 20 species of the 285 Danish species of caddisflies, dragonflies, mayflies, and stone flies have gone extinct between 1900 and 1989. A larger number of 76 species are either threatened, vulnerable, or rare. The red-listed species often require high water quality and have long life cycles that are sensitive to disturbance. The historical development has, therefore, led to a more stereotypic composition of both plant and macroinvertebrate communities with the same few robust species dominating at most stream sites. In contrast to the *Potamogeton* vegetation, however, insect diversity has improved in Danish streams following reduced organic pollution since 1980 and species have spread from local surviving populations both within and between neighboring river courses. Physical impact by chanellization and intense management of streams still remains a main factor for depriving diversity of macroinvertebrates in about 23% of Danish streams, whereas about 46% of the streams are still deprived of species by organic inputs from scattered settlements, farms and fish farms in open areas, and outlet from sewer systems during heavy rainfall.

There are very few similar historical evaluations of species distribution and abundance. What come closest are studies of agricultural areas of the United States Midwest. These demonstrate profound reduction of habitat quality, species richness, and abundance of stream invertebrates and fish between 1950 and 2000 as a result of removal of most of the natural riparian forest vegetation, stream regulation, and alteration of water chemistry. Indices of stream quality decline with the intensity of cultivation in the catchment, and this is accompanied by a predictable decline in species richness and feeding types of invertebrates and fish. Predominantly robust, generalist species have survived in the stream communities. Fish that specialized on insects have disappeared, whereas omnivorous fish have survived. The historical development of plants and animals, therefore, has many similarities in the lowland streams of Europe and the United States, because all macroscopic organisms have been strongly influenced by cultivation and agricultural practices.

Biological Consequences of Organic Pollution and Agricultural and Industrial Chemicals

Organic pollutants give rise to deoxygenation and release of ammonia and other mineral nutrients in the water due to high microbial activity. Organic pollution leads to unhealthy aquatic conditions, high nutrient levels, untransparent water, and smothering of sediments. These conditions are detrimental to the survival of invertebrates and fish, and highly oxygen-sensitive groups of insects such as beetles, caddisflies, mayflies, and stone flies, as well as salmonid fish, may disappear almost entirely. Also, the diversity of feeding groups of macroinvertebrates declines as stony substrata for algal grazers are smothered, coarse terrestrial detritus for shredders becomes insignificant relative to the accumulation of organic mud for deposit feeders, and large predators with long life cycles find it hard to sustain episodes of oxygen depletion. Thus, organic pollution is accompanied by a general loss of species richness, taxonomic diversity, and functional diversity of macroinvertebrates, in addition to the appearance and dominance of particularly robust species (e.g., tubificid polychaetes and chironomids) in exchange for the loss of many more sensitive species (**Figure 4**).

The biological changes are principally the same in temperate and tropical streams. The macroinvertebrate indices originally developed to describe organic pollution in temperate streams work well in tropical streams. The obvious reason is that the taxonomic composition of orders and families of macroinvertebrates does not differ greatly between temperate and tropical streams. Taxonomic similarity is particularly high between temperate lowlands and tropical highlands. The warm tropical lowland streams may have a high sensitivity to organic pollution because of low oxygen solubility, fast microbial degradation, and high invertebrate metabolism. Tropical highland streams also may be more sensitive to organic pollution than temperate streams because of the reduced partial pressure of oxygen at higher elevations, whereas temperatures resemble those in temperate lowland streams.

Domestic wastewater contains many chemical stressors in addition to the oxygen-consuming organic material. Sensitivity of organisms to low oxygen often goes hand-in-hand with the sensitivity to industrial chemicals. Similarly, sensitivity to acid waters usually changes in parallel with the sensitivity to heavy metals. Chemical stressors are lethal at high

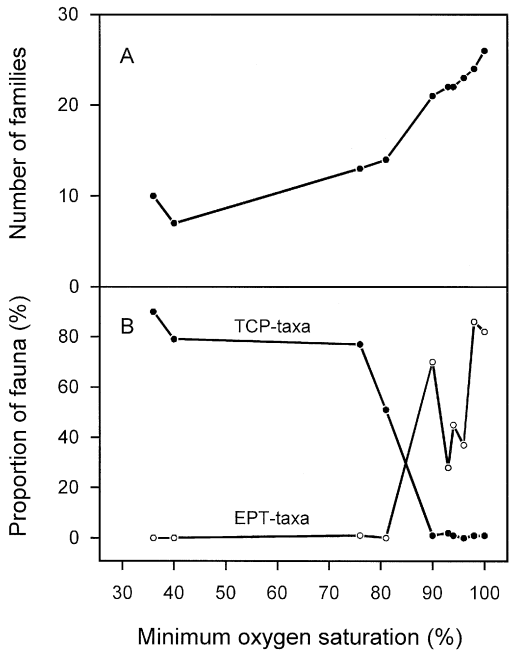


Figure 4 Composition of macroinvertebrate communities in small highland streams in tropical Ecuador as a function of organic pollution from villages, resulting in a lower minimum oxygen saturation in the water: (a) number of macroinvertebrate families; (b) proportion of oxygen-requiring taxa of caddisflies, mayflies, and stone flies (EPT taxa) to taxa that require less oxygen (TCP taxa), such as Tubificidae and species of *Chironomus*. Data from Jacobsen D (1996) *Ecology and Environmental Status of Tropical Streams in Ecuador*. Copenhagen: Freshwater Biological Laboratory and Danida, Danish Foreign Ministry.

concentrations, and sublethal disturbance of body form, reproduction, neural function, and behavior occurs at lower concentrations. Among *Chironomus* species, for example, the proportion of individuals with deformed mouthparts increases with the level of pollution. Among net-spinning caddis larvae, the proportion of abnormal, incomplete nets increases with the concentration of toxic aluminum. These sublethal effects will eventually lead to a decline of populations due to reduced growth and survival compared with more resistant species. However, under very severe oxygen deficiency or high concentrations of toxic chemicals, no species of higher life forms among fish, invertebrates, and plants are capable of surviving, and only microorganisms will remain. Chemically polluted streams and ponds without any higher life include habitats with high concentrations of cadmium, copper, zinc, and ferrous iron.

The immune and hormone systems of animals can be disturbed by exposure to organochlorine compounds (e.g., dichlorodiphenyltrichloroethane (DDT) and polychlorinated biphenyls (PCBs)), nonylphenols and octylphenols of industrial detergents, and an estradiol component from degradation of contraceptive pills. In water bodies with a high proportion of sewage effluents, such compounds can apparently lead to the development of hermaphrodite individuals of roach and rainbow trout, abnormally small testicles in alligators, and feminization of populations of terns and gulls feeding on aquatic animals. Estrogenic effects have received great attention because human populations have been showing increasing frequencies of testicle cancer and decreasing numbers of viable sperm cells over the past 50 years in Scandinavia.

The great variety of pesticides now being found in surface waters are also potentially dangerous because they have been manufactured to be deadly poisonous in trace amounts to fungi, insects, and plants. Insecticides target the most species-rich group of animals on land and in freshwater. For the time being, however, the broad-scale environmental significance of pesticides is largely unknown. There are suggestions that intense application of pesticides in some regions can account for an impoverished insect fauna in streams, where other pollution sources have largely been eliminated. This is the situation in 2–6% of Danish streams.

Biological Consequences of Lake Eutrophication

Lake eutrophication influences ecosystem structure and functioning as well as the distribution and abundance of most species. This fundamental influence is the result of the primary role of nutrients in boosting the biomass and productivity of phytoplankton, setting the scene for many structural and functional regulations.

The attitude toward cultural eutrophication changes with its magnitude and the public interests in different regions. In lowland regions of Europe and North America, substantial eutrophication has shifted the natural wide range of lakes to a higher nutrient level that encompasses conditions from mesotrophy to hypereutrophy. Water purification aims at improving hypereutrophic lakes and moving them to eutrophy, such that most lakes may occupy the narrow range from mesotrophy to eutrophy with no lake being unpolluted and no lake being extremely polluted. As a consequence, species assemblages have become more stereotypic as the susceptible species of oligotrophic habitats, and in some cases also the extremely robust species of hypereutrophic habitats, have become very rare. Full control of pollution and the ecological disturbance associated with agriculture, fish farming, forestry, and recreational use, which may be needed to maintain pristine, oligotrophic, clear-water lakes, has rarely been implemented.

Social and environmental priorities are different in other societies. In Southeast Asia, for example, the focus is more on fish production and less on esthetics and the protection of rare species. A small lake that is heavily fertilized by village sewage may be a valuable source of fish that feed on the thick algal soup and that are tolerant to periods of deoxygenation.

The Case of Lake Fure and Other Hardwater Lakes

The 100-year development of Lake Fure in Denmark can serve as an example of the widespread, accelerated nutrient loading of temperate lakes followed by recent reduction of loading rates (see **Figure 1**). The long-term evaluation for this lake matches the long water residence time (10 years) and the slow numerical response of macrophyte communities in this naturally mesotrophic, hardwater lake.

Annual external phosphorus loading of the lake increased from 1.3 tons of P in 1900 to 37 t in 1969.

Until the late 1940s, the lake had managed to tolerate the increased loading relatively well because of the efficient sediment binding of phosphorus to calcium carbonates and iron oxyhydroxides. The bottom waters had remained oxic and iron-bound phosphorus had stayed in the sediments. With the accelerating phosphorus load during the 1950s and 1960s, extensive phytoplankton blooms developed, leading to oxygen-free bottom waters and release of iron-bound sediment phosphorus. Following diversion and tertiary treatment of the sewage in 1969, the external phosphorus loading has gradually declined to 2–3 tons of P in 2008.

The mean Secchi depth during summer was about 5–6 m from 1900 to 1940, but it has fluctuated between 1.5 and 3.2 m in the 1970–1980s. With the reduced phosphorus loading in the 2000s, the mean Secchi depth has increased to 3–5 m. The number of phytoplankton species has remained approximately constant between 70 and 85. However, species within particularly nutrient-demanding groups typical of eutrophic lakes (e.g., blue–green and chloroccalean green algae) have increased from 17 to 39, and species typical of oligotrophic lakes (e.g., chrysophytes, desmids, and certain dinoflagellates) have declined from 32 to 13. The same changes in dominant algal groups are found in broad-scale comparisons among oligotrophic and eutrophic lakes.

Species richness of submerged macrophytes in Lake Fure before eutrophication commenced was 37 species, including 18 rooted flowering plants, 11 characeans, five mosses, and three macroalgae. Tall flowering plants grew from shallow water to the depth limits at 8 m, but dominated the vegetation down to 5 m depth. Small flowering plants, characeans, and mosses formed a mixed carpet of vegetation below the canopy of large flowering plants, whereas characeans dominated from 5 to 8 m.

In 1983 and 1993 only 12–13 of the original 37 species remained, whereas three new species of pollution-tolerant green macroalgae had appeared. Most small flowering plants and characeans and all mosses had vanished. Eight large species of canopy-forming flowering plants survived in the turbid lake because they can compensate for the poor light conditions by exposing the leaves close to the water surface. Depth distribution of all original submerged macrophyte species was restricted from 8 to 2 m because of the reduced light penetration.

During the recent oligotrophication of the lake, species richness increased to 25 in 2005, depth limits grew to 4–7 m, and a few characeans and mosses reappeared, whereas bottom mats of the pollutant-tolerant macroalgae remained dominant in the now nutrient-richer sediments.

Although internal phosphorus concentrations, phytoplankton biomass, Secchi transparency, and macrophyte depth penetration and species number are obviously reversible following reduced external phosphorus loading, macrophyte species composition is not reversible, at least not in the time frame of 30 years. About 10 small oligotrophic species susceptible to shading and competition had vanished and did not reappear. Their reestablishment is either impossible or just extremely slow because widespread eutrophication has eradicated them across large regions and they can recolonize only from small, distant refugia.

Broad-Scale Comparisons and Whole-Lake Experiments

Other individual lake studies basically tell the same history of eutrophication. In the large lakes, however, the role of the littoral zone and bottom plants is small, and emphasis is more on the increase of phytoplankton biomass, loss of bottom fauna, and alteration of fish communities. In most lakes, phosphorus is the main nutrient-limiting phytoplankton biomass, but in some lakes, the importance of phosphorus and nitrogen as limiting nutrients alternates with the seasons or nitrogen plays the major role.

In lakes with a small magnitude or a short period of cultural eutrophication, rapid recovery was observed upon reduction of nutrient loading, because only small nutrient pools had accumulated in the lake bottom. Recovery also is fast and more complete when water renewal is high, and when large proportions of unpolluted water are available from undeveloped mountain and forest areas such that nutrients can be flushed from the lake (e.g., Lake Washington, United States).

Case studies of individual lakes, comparisons among many lakes, and controlled whole-lake fertilization experiments all show strong positive relationships among nutrient availability, phytoplankton biomass and productivity, zooplankton and fish production, light attenuation in the water column, and risk of anoxia in bottom waters. Predictions of phytoplankton biomass based on external phosphorus loading and water renewal are often excellent for lakes within a region of the same climate and soil geology and for lakes where nutrient inputs are precisely known. These results suggest that biotic differences do not play a major role for the biomass of phytoplankton. It has therefore been argued that the lower accuracy of predictions observed among lakes from different regions and studies may result from methodological differences and uncertainties rather than from the influence of biological differences in the food webs. In shallow lakes, however, wide shifts can take place from transparent lakes of low phytoplankton biomass and high macrophyte cover to more turbid lakes of higher phytoplankton biomass and marginal macrophyte cover, even though phosphorus loading remains approximately the same. In shallow lakes, rooted macrophytes and attached animals have the potential to exert considerable control on the communities in the open waters – in part through internal alterations of nutrient cycling. The same intimate interactions between organisms and processes at the bottom and in the open water are not possible in large, deep lakes.

Resource and Predatory Control in Food Webs

In many lakes, phosphorus stripping and the diversion of sewage have reduced nutrient input and greatly improved lake quality (**Figure 1**). In some lakes, improvements of lake quality have been small presumably because of high internal nutrient circulation. To reduce sediment release and internal water concentrations of phosphorus, several methods have been attempted. One principle is to remove P-enriched cultural sediments by pumping them up. Another principle is to increase the P-binding capacity in the sediment by adding aluminum. The immediate effect is very profound, but over time, the binding capacity of aluminum flocks declines and much phosphorus is released again. Iron addition conducive

to efficient phosphorus binding to iron oxyhydroxides is problematic because permanent oxic waters are required through, for example, continuous aeration of the bottom waters.

In shallow lakes, biomanipulation has had the objective of reducing the number of planktivorous fish, which eat the large zooplankton (especially daphnids), and thereby hopefully reduce the biomass of phytoplankton through enhanced zooplankton grazing. The additional goal is to increase the cover of submerged vegetation, which can directly reduce sediment release of nutrients, impede phytoplankton development by shading, and provide refuge to daphnids.

The success of attempts to reduce phytoplankton blooms and improve water clarity by internal manipulation of sediments and food webs is variable, and some positive short-term effects for up to 10 years by removal of planktivorous fish have not been sustainable in the long term. Addition of predatory fish (e.g., pike) with the aim of reducing planktivorous fish and phytoplankton has been unsuccessful. The successes and failures of biomanipulation are widely debated. The view that grazing control on phytoplankton is high in lakes with an even number of trophic levels in the food web and low in lakes with an odd number of levels has not only received some support, but also attracted growing opposition because of the following: (1) predictions based on the concept are poor, (2) trophic levels and numbers of links are difficult to define, and (3) many species are omnivorous, show ontogenetic and opportunistic changes in food preference, and cover more than one trophic level and habitat (i.e., both water and sediment). Moreover, there is a strong positive relationship from nutrients and resource levels to biomass and productivity of phytoplankton, zooplankton, and fish because of the need to transfer energy and matter from low to higher trophic levels.

Many studies clearly demonstrate a strong role of predators on size composition and species abundance of prey organisms, and they show behavioral alterations of predators and prey to increase their own growth and survival. These flexible biological responses, however, do not imply that the main control of the combined biomass and productivity of all the species at each trophic level is exerted by predation. Overall, predatory fish have a small cascading trophic effect through the entire food web and a minor influence on phytoplankton biomass and water transparency. As a consequence, there is no easy technical or biological fix to ameliorate lake eutrophication, but various techniques may assist the effect of sustained reduction of external nutrient input with the aim of reducing troublesome phytoplankton blooms.

Species Richness and Lake Trophy

Species richness increases with habitat area and heterogeneity. The increase in species richness (S) with habitat area (A) is often predicted by the equation $S = \text{constant} \times A^z$. The z value depends on the organisms, the spatial range of the study, and the habitats, and there is no single mechanism that can account for the variability of z. However, low z values are consistent with high immigration rates, low extinction rates, and a low rate of increase in additional habitat with increasing area. As a consequence, very small z values are predicted for bacteria, microalgae, protozoa, and zooplankton. Indeed, the z values predicting crustacean zooplankton species richness are only 0.054 for European and 0.094 for North American lakes. A low z value of 0.10 for submerged macrophytes in Scandinavian lakes is also consistent with their widespread occurrence.

In contrast, species richness of fish in lakes of four different regions in Canada and the United States shows much higher z values (0.16, 0.22, 0.36, and 0.37) consistent with the more restricted geographic range size, smaller local abundance, and greater risk of extinction of fish following disturbance or restriction of natural habitats. As a consequence, risks of local, regional, and global extinction are high among fish and low among microorganisms and plankton organisms, with aquatic plants and insects probably holding an intermediate position. Some large crustaceans and molluscs presumably resemble the fish by having a low rate of dispersal and including several species of restricted geographic distribution (e.g., the freshwater crabs and prawns on tropical oceanic islands).

Changes in ionic composition and nutrient concentrations of freshwaters are superimposed on changes in habitat area and heterogeneity. Water chemistry is not independent of habitat area and heterogeneity, since small lakes include a higher proportion of low pH, low-calcium, and nutrient-poor waters than do large lakes. In many studies of the influence of pH and lake trophy on species richness, it is not possible to fully compensate for the influence of habitat area and heterogeneity.

The most diverse conditions and highest species richness often are found under mesotrophic conditions, where many species can coexist at different sites and depths within the lakes. Thus, high habitat heterogeneity is probably important for the common peak of species richness under mesotrophic conditions. For example, many free-living and attached species adapted to different conditions of light, temperature, and exposure can replace each other along depth gradients from shallow to deep waters. Under very oligotrophic conditions, many algae and plant species are nutrient limited, and low primary production places an energetic restriction on the diversity and density of invertebrates and fish. In contrast, under hypereutrophy, many species are restricted by the lack of light and oxygen in deep waters, and special adaptations are required for survival in very muddy sediments.

Two other conditions contribute to the decline of species diversity under hypereutrophic conditions. First, the natural local variability between unfertilized and fertilized sites disappears with the overall nutrient enrichment at all sites accompanying eutrophication. Second, hypereutrophy is exceptional for most species, which have evolved over many millions of years of no or weak human impact. Species richness is expected to be highest in those common habitat types that have had the most widespread and long-term natural occurrence, because there has been time and room for speciation. Hypereutrophic lakes, which are common today, have been exceedingly rare during most of the development of freshwater species.

It is noteworthy that lake ecosystems are much more sensitive to environmental deterioration and catastrophic declines of organisms by eutrophication than terrestrial ecosystems for at least four reasons. First, nutrients tend to stay

in circulation within the lake boundaries. Second, phytoplankton responds by a steep increase in biomass and productivity. Third, photosynthesis and growth of algae and plants are restricted to the uppermost surface waters because of the impoverished light conditions. Fourth, the risk of oxygen depletion is very high because of low oxygen solubility in the water. Therefore, species richness within many groups of organisms declines in hypereutrophic lakes because of reduced habitat heterogeneity, restriction to the distribution of organisms, and development of stressful and highly variable environmental conditions with respect to light, oxygen, pH, and sulfide.

Biological Consequences of Acidification

Acidification up to 1980 initiated complex chemical and biological changes in surface waters and catchments with poorly weathered rocks and thin soils. The acidification of freshwater habitats was also accompanied by acidification of terrestrial ecosystems in the catchments, leading to a widespread regional decline of biodiversity.

Whole-lake experiments involving acidification to pH 5 demonstrate that damage to ecosystem functioning is of secondary concern compared to impoverishment of the biotic communities. Although many phytoplankton algae disappear, the remaining algae can maintain biomass and primary production at the same level, largely regulated by the input of phosphorus and water renewal. The rate of decomposition is governed by primary production and not by pollutants. However, the complete disappearance of large invertebrate detritivores upon acidification may lead to the accumulation of detritus and over the long term reduce plant production based on recirculated nutrients. Overall, nutrient cycling was not disrupted in whole-lake experiments, but nitrification declined at pH values below 5.4.

Acidification of the watershed changes ionic composition and may reduce the availability of phosphorus in soils and surface waters. Soil acidification increases the leaching of base cations, and it has specific effects on the solubility of aluminum and phosphate. Chemical weathering of aluminum minerals is enhanced under acid conditions. However, soluble forms of aluminum ions may precipitate with phosphate at the proper pH, leading to "oligotrophication." As a result, lakes in acidified watersheds may become very clear from the enhanced phosphorus limitation of phytoplankton growth, and certain acid-tolerant aquatic mosses (e.g., *Sphagnum* spp.) and filamentous green algae (e.g., *Mougeotia* spp.) may spread across the lake bottom to greater depths. Only a few vascular plant species can tolerate highly acidic conditions (e.g., *Juncus bulbosus*).

Species diversity and abundance decline within most groups of organisms upon acidification, though some groups are much more sensitive than others and distinct differences exist among species in their tolerance to acidification (**Figure 5**). In acidified regions of Europe and North America, many lakes lost 30% or more of the species in some taxonomic groups before 1980. Among planktonic and benthic microalgae, desmids – a group of green algae – prefer buffer-poor waters of low pH and low-calcium content, and they are often accompanied by a few acid-tolerant species of blue–green algae, dinoflagellates, diatoms (e.g., *Eunotia exigua*), green algae (*Chlamydomonas acidophila*), and euglenophytes (*Euglena mutabilis*). In contrast, most species of diatoms, chlorococcalean green algae, and blue–green algae prefer well-buffered waters of neutral pH. Among planktonic rotifers, copepods, and daphnids, only a few species tolerate acid waters (less than 4–5), so the planktonic food webs become simple. The tolerant species can utilize the vacant ecological niches left by others, thus maintaining a relatively constant zooplankton biomass despite the decline in species richness.

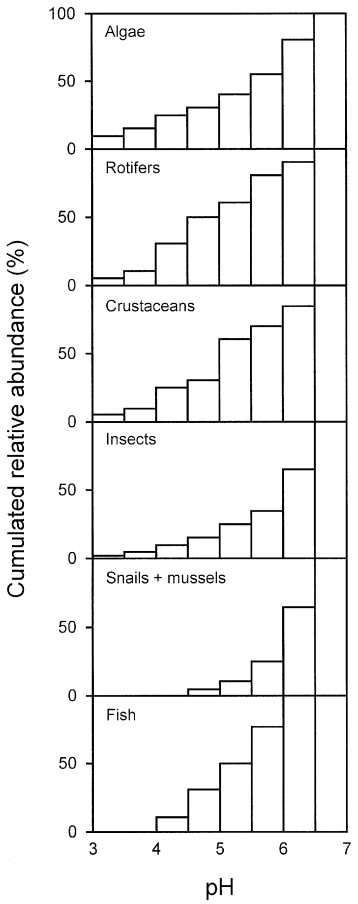


Figure 5 Cumulative frequency distribution of minimum field observations of pH for 46–370 aquatic taxa within different taxonomic groups. Adapted from Schindler DW (1991) Aquatic ecosystems and global ecology. In: Barnes RSK and Mann KH (eds.) Fundamentals of Aquatic Ecology, 2nd edn, pp. 108–122. Oxford, United Kingdom: Blackwell.

Many snails, mussels, and crustaceans (e.g., crayfish) are highly sensitive to acidification due to problems in maintaining the ionic regulation of sodium, chloride, and potassium and in forming the exoskeleton of calcium carbonate or calcium-chitin when pH and calcium concentrations are low. Higher calcium concentrations can help in maintaining membrane integrity, reducing gill permeability to salt loss, and ensuring the construction of shells. Several metal toxins such as aluminum mobilized under acidic conditions have the opposite influence.

Among insects, the most acid-tolerant species include surface-breeding beetles with a hard exoskeleton and certain stone flies and caddisflies with a low metabolism and mostly lacking thin-walled body parts. In contrast, mayflies with their large, thin gills are often highly sensitive to acidification. The decline in species richness of insects is not as pronounced when acidic conditions occur in waters rich in humic substances, which can bind and detoxify the metal ions mobilized at low soil pH.

A common result of acidification is the dramatic loss of fish. Former acid-stressed regions of Scandinavia, the Adirondack Mountains (USA), and La Cloche Mountains (Canada) showed a steep decline in fish species as lake pH declined from 8 to 4, and many lakes with pH less than 5 became fishless. In Norway, brown trout disappeared from more than 8000 lakes. With the reversal of acidification, fish populations are slowly coming back. The decline in fish species richness with reduced pH is much less pronounced in subtropical Florida lakes and in tropical lakes. The presence of more acid-tolerant fish species in warm lakes may be due to a much slower and more gradual decline in pH over time and, accordingly, better opportunities of genetic adaptations to low-pH waters. Also, Florida lakes have lower concentrations of aluminum and free mineral acids, longer growing seasons, and no episodic input of acids resembling the massive acid input of north temperate lakes during snowmelt.

Fish studies in north temperate lakes showed that eggs, embryos, and larval stages of fish are more sensitive to acidification than the adult stages. As a consequence, young generations may vanish upon acidification, whereas old fish may remain for a while. Populations will eventually die out unless young fish are continuously stocked. Salmonids and cyprinids with a high metabolism and a large gill surface appear to be more susceptible to acidification than fish with a low metabolism and a low gill surface (e.g., eels). A critical phase for the well-being of fish and invertebrates is the maintenance of efficient oxygen uptake and ion regulation across the gill surfaces.

Ecosystem Functions and Explanations

Ecosystem functions have proven to be more resistant to acidification than species changes. Species assemblages are often plastic enough to maintain ecosystem functions under quite extreme stresses, although only a few species survive. For example, plant productivity continued in a dense monospecific vegetation of *J. bulbosus* growing in a highly acidic, temperate stream (pH 2–3) from a lignite mining area with intense pyrite oxidation, whereas neighboring alkaline streams (pH 6–7.5) contained a mixed assemblage of about 10 rooted plant species. Similarly, phytoplankton biomass and production were unaffected in lakes polluted with strong acids and trace metals from smelters. Phytoplankton and other microorganisms are also locally abundant and are easily dispersed, such that highly pH-tolerant species will rapidly show up and multiply in recently acidified localities.

Some ecosystem processes, however, may decline or disappear entirely under very acidic conditions. At pH values below 5.4–5.7, ammonium begins to accumulate due to the cessation of nitrification. Therefore, the loss of internally produced nitrate through denitrification will stop. However, the enhanced atmospheric deposition of nitric acid and sulfuric acid will permit the microbial activity of denitrifiers and sulfate reducers in oxygen-free sediments and, thereby, generate alkalinity in the water bodies.

The biotic communities of lakes and streams are strongly damaged and impoverished by acidification. The most vulnerable species appear to be organisms with year-long life cycles and poor dispersal, such as large invertebrates and fish. These groups are also the less speciose, particularly in northern water bodies, which means that there is little redundancy in the food webs to prevent pH-sensitive species of large detritivores, plankton-eating fish, and predatory fish from being replaced by pH-tolerant species. Fish and crayfish, which are highly valued by humans, are therefore easily damaged. Even bird populations such as dippers, osprey, and pied flycatchers may vanish from the lack of food or the formation of thin-shelled eggs while feeding on calcium-deficient and aluminum-rich prey.

Overall, surface waters of very low pH represent a special environment to which very few species have become adapted as a result of (1) the direct stress at low pH, (2) the unsuitable ionic composition, (3) the toxic metals, and (4) the rarity of these habitats during the evolution of aquatic organisms, particularly in temperate and subarctic regions.

Conclusions

Freshwaters of high quality are essential to humans for domestic and agricultural uses. Freshwater ecosystems provide many other highly valued and critical ecological services, and they support a very high diversity of species and taxonomic groups considering their relatively small volumes and surface areas.

Human impact has caused numerous stream reaches and shallow lakes to disappear, disturbed and polluted many others, and disrupted the close contact between the floodplain and the streams. Consequently, the diversity and abundance of numerous freshwater species have declined due to the profound restriction of areal cover and habitat diversity of freshwater environments. Moreover, the systematic loss of pristine, unpolluted freshwaters has threatened or eradicated those species that require such environments.

The overall disruption of the linkage between freshwater and terrestrial environments has far-reaching consequences for global hydrology, elemental cycling, and biodiversity. Among those consequences are enhanced risks of flooding and drought, greater risks of land erosion and marine siltation, and reduced opportunities for the countless plants and animals living a natural double life between freshwater and land to survive in the future.

Accompanying all of this disturbance and pollution has been the accidental or deliberate introduction of non-native organisms such as crayfish, plants, sport fish, and zebra mussels, which may monopolize freshwater habitats and drive the native, endemic species in formerly undisturbed lakes and

river systems to extinction. Several species of freshwater fish have gone extinct during the past few centuries, and many additional species of plants, large invertebrates, fish, and frogs now exist in extremely small and geographically isolated populations. These species face a great risk of extinction, and they can no longer provide their accustomed ecosystem functions. Nor can these species be enjoyed within most of their native area. Unfortunately, they will find it exceedingly difficult to recolonize newly restored aquatic habitat and contribute to future speciation.

See also: Agriculture, Industrialized. Endangered Freshwater Invertebrates. Freshwater Ecosystems. Human Impacts on Ecosystems: An Overview. Wetland Creation and Restoration

References

Allan JD (1995) Stream Ecology. Structure and Function of Running Waters. London: Chapman & Hall.

Clark JM, Bottrell SH, Evans CD, *et al.* (2010) The importance of the relationship between scale and process in understanding long-term DOC dynamics. *Science of the Total Environment* 408: 2768–2775.

Hesthagen T, Fjellheim A, Schartau AK, Wright RF, Saksgård R, and Rosseland BO (2011) Chemical and biological recovery of Lake Saudlandsvatn, a formerly highly acidified lake in southernmost Norway, in response to decreased acid deposition. *Science of the Total Environment* 409: 2908–2916.

Kristensen P and Hansen HO (eds.) (1994) *European rivers and lakes – Assessment of their environmental state*. Copenhagen: European Environment Agency. EEA Environmental Monographs, vol. 1.

Petts GE (1984) *Impounded Rivers*. Chichester, United Kingdom: John Wiley & Sons.

Williams P, Whitfield M, Biggs J, *et al.* (2003) Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in southern England. *Biological Conservation* 115: 329–341.

Functional Diversity

David Tilman, University of Minnesota, St Paul, MN, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 109–120, © 2001, Elsevier Inc.

Glossary

Diversity-productivity hypothesis The proposal that greater diversity would lead, on average, to greater total biomass or productivity.

Diversity-stability hypothesis The proposal that ecosystems containing more species would be more stable.

Ecosystem composition The list of species or functional groups that are present in a given ecosystem.

Ecosystem functioning The rate, level, or temporal dynamics of one or more ecosystem processes such as primary production, total plant biomass, or nutrient gain, loss, or concentration.

Functional diversity The range and value of those species and organismal traits that influence ecosystem functioning.

Functional group A set of species that have similar traits and that thus are likely to be similar in their effects on ecosystem functioning.

Niche differentiation Differences in the morphology, physiology, or behavior of species that can influence their abundances, dynamics, and interactions with other species, including the ability of various competing species to coexist.

Sampling effect The hypothesis that diversity might influence an ecosystem process because of the greater chance that a given species trait would be present at higher diversity, and the effect of its presence on ecosystem functioning.

Measurement of Functional Diversity

Because of the large number of traits that each species possesses, the large number of different species that exist in most habitats, and the incomplete knowledge of which species traits influence various ecosystem processes, there is, as yet, no simple way to measure functional diversity. Rather, items that are more easily measured than functional diversity are used as indices or correlates of functional diversity. The most common of these indices is the number of species present in a habitat, which is called the species richness or species number of the habitat. All else being equal, habitats with greater species richness should also have greater functional diversity. This occurs because species differ in their traits. Sites that contain more species should thus also contain, on average, a greater range of species traits, which is greater functional diversity. Species diversity indices, such as the Shannon diversity index, are similarly used as indirect measures of functional diversity. Another commonly used index of functional diversity is the number of different functional groups (defined later) that exist within a given community or ecosystem. This is also called functional group diversity. Assuming that organisms can be categorized as belonging to groups that differ in traits relevant to ecosystem functioning, greater functional group diversity should correlate with greater functional diversity. However, variations among species within a given group could also contribute to functional diversity. Observational, experimental, and theoretical studies indicate that functional diversity, as measured by any of these three means, is one of several important factors that determine ecosystem functioning. Because there is, as yet, no clear way to measure functional diversity, one or more of these three indices will be used as a proxy for functional diversity in this chapter.

Before reviewing the research linking functional diversity to ecosystem processes, which is the focus of the remainder of this chapter, it is important to introduce and define some terms.

Explanation of Concepts and Terminology

Functioning

As they are used by ecologists, the words *function*, *functional*, and *functioning* are not meant to imply that an ecosystem process has any underlying goal or purpose. Indeed, to try to minimize any such implications, it has become standard practice to refer to "ecosystem functioning" or "ecosystem process" rather than the "function of an ecosystem." The latter might be misinterpreted as meaning that an ecosystem exists to perform a given function, which is inconsistent with our knowledge of the process of evolution. Rather, functioning refers solely to the way in which an ecosystem operates.

Ecosystem Processes

Ecologists study many different aspects of the functioning of communities and ecosystems. The three most frequently considered ecosystem processes are productivity, stability, and resource dynamics. Productivity refers to the rate of production of biomass within a given trophic level. The production of plant biomass is called primary production, the production of biomass of herbivores is called secondary production, and that of predators is called tertiary production. Stability has a wide range of definitions, including the degree to which an item is resistant to change when experiencing a

single perturbation, the degree to which an item fluctuates in response to an ongoing suite of small-scale perturbations, and the dynamics of return to its prior state after a single perturbation. Stability can be measured at the level of populations, communities, or ecosystems. The resource dynamics of an ecosystem are measured by the rates of supply and loss of limiting nutrients, by the efficiency with which organisms use limiting resources, and by the proportion of limiting resources that the organisms living in an ecosystem are able to capture.

Functional Groups

Each species has a large number of morphological, physiological, and behavioral traits, many of which might influence the abundance of species and ecosystem functioning. One way to deal with such complexity has been to identify traits that seem more likely to influence ecosystem processes. Chapin *et al*. (1997) suggested that the species traits with the greatest effects on ecosystem functioning were those that (a) controlled the acquisition, use, and availability of limiting resources; (b) modified the feeding structure of food webs; and (c) affected the occurrence and magnitude of disturbances. Such traits can be used to classify organisms into different functional groups. For instance, species can be divided, first, into functional groups based on their position in a food web: photosynthetic plants, herbivores, predators, parasites, parasitoids, decomposers, and so on. Organisms within each of these groups can be further subdivided based on their acquisition and use of their limiting resources. For grassland plants, for instance, this might be based on the time, within the growing season, when each plant was maximally active (cool-season versus warm-season plants), and on its carbon (C-3 or C-4 photosynthetic pathway) and nitrogen physiology (high nitrogen use efficiency, low nitrogen use efficiency, ability to fix atmospheric nitrogen). Such considerations might lead, for instance, to the classification of grassland plants into six functional groups: C-3 grasses, C-4 grasses, C-3 forbs, C-4 forbs, legumes, and woody plants. The assumption inherent in making such a classification is that species within a class are highly similar, and those in different functional groups differ markedly from one another.

Diversity versus Composition

It has long been recognized that the functioning of an ecosystem depends on which species the ecosystem contains (i.e., on it species composition). Interest in species diversity as an alternative or additional explanation for ecosystem functioning means that it is necessary to define species diversity, especially functional diversity, in a way that distinguishes diversity from species composition. This requires a definition that is more restricted than that traditionally used. In particular, effects should be attributed to diversity only once there has been simultaneous control for effects of composition, and effects should be attributed to composition only once there has been simultaneous control for effects of diversity. To achieve this in an experimental, theoretical, or observational study, it is necessary (a) to hold composition constant via randomization (numerous communities with randomly-chosen compositions) while changing diversity, (b) to hold diversity constant while changing composition, (c) to simultaneously vary both in an appropriately randomized and replicated design, or (d) to control for each statistically, such as via multiple regression, which is most appropriate for observational studies.

Early Work on Functional Diversity and Ecosystem Processes

Effects of diversity on ecosystem processes were first recognized by Darwin in *The Origin of Species*. Darwin noted that it was well-known that increased plant diversity led to greater primary productivity in pastures. The British ecologist, Charles Elton, hypothesized in his 1958 book titled *The Ecology of Invasion by Animals and Plants* that diversity would impact many aspects of ecosystem functioning. In particular, he suggested that greater diversity would lead to greater ecosystem stability, an idea that was further developed by the leading ecologists of that era, including Robert MacArthur, Gene Odum, and Ramon Margalef. Elton also suggested that greater diversity would decrease the susceptibility of an ecosystem top invasion by other species and would decrease the incidence of outbreaks by diseases and pests.

Elton's diversity-stability hypothesis was called into question, though, by the mathematical theory of May (1972), which predicted that the linear stability of communities of competing species would, in general, decrease as the diversity of the communities increased. The general consensus reached after publication of May's book was that other factors were likely to be more important than diversity as determinants of ecosystem processes. This view led ecologists to focus more of their attention on other issues, with much of that effort dedicated to better understanding the mechanisms of species interactions and the effects of species composition on ecosystem processes.

Recent explorations of the potential effects of diversity on ecosystem processes were inspired, to a great extent, by the publication of *Biodiversity and Ecosystem Functioning* (Schulze and Mooney, 1993). In a chapter in that book, Vitousek and Hooper hypothesized that many ecosystem processes, like primary productivity, should increase as diversity increased, and they stressed that the most important component of diversity might be functional group diversity. Agricultural studies were reviewed in a chapter by Swift and Anderson, who noted that mixed crops, especially those containing a legume and a grass, were often more productive than either crop species growing alone, supporting the diversity-productivity hypothesis. A chapter by McNaughton reviewed and evaluated a large number of observational and small-scale experimental studies in which stability was greater for ecosystems containing more species and highlighted data supporting Darwin's diversity-productivity hypothesis. These and other contributions in this book set the stage for a burst of work that has included development of additional mathematical theories, field and laboratory experiments, and observational studies.

The Effects of Functional Diversity

Functional Diversity, Productivity, and Nutrient Dynamics

Theory and Concepts

The potential effects of functional diversity on productivity have been described by two qualitatively different models, reviewed in Tilman (1999). The first is the sampling effect model, simultaneously proposed in 1997 by three different authors (L. Aarssen; M. Huston; and D. Tilman, C. Lehman, and K. Thomson). The sampling effect model hypothesizes that species differ in their competitive abilities, and that species that are better competitors are also more productive. Given these assumptions, communities that have greater diversity should, on average, be more productive because they are more likely to contain one or more species that are more productive.

A formal mathematical treatment of the sampling effect, provides some deeper insight into the way that functional diversity can impact ecosystem processes. For this treatment, let R^* be the level to which a limiting resource is reduced by a species when growing alone. As shown both theoretically and in numerous competition experiments (Grover 1997), the best competitor would be the species with the lowest R^*. The R^* value of the species can be used to rank them from good to poor competitive ability (i.e., from the lowest to the highest R^* value). Assume that the species composition of a community is determined by random draws (sampling) from the infinite pool of species with all possible R^* values between a minimum (R^*_{min}) and maximum (R^*_{max}) On average, the functional diversity of a community would depend on the number of species drawn, N, which is the initial diversity. The number of species in a community, N, is a good measure of functional diversity in this model because the range in the values of the relevant species trait (R^*) is higher in communities containing more species. These assumptions of sampling effect yield a simple equation that relates the long-term average biomass of a plant community, $B_{(N)}$, to its original plant species diversity, N:

$$B_{(N)} = aQ\left(S - \left(R^*_{min} + \frac{R^*_{max} - R^*_{min}}{N+1}\right)\right) \qquad (1)$$

Here a is the rate of resource mineralization, Q the coefficient of resource conversion into biomass, and S is the rate of resource supply in the habitat.

The sampling effect model predicts that total community biomass, a measure of primary productivity, increases with plant diversity, as shown in **Figure 1a** The trend predicted is one in which added diversity leads to large increases in productivity when diversity is low, but has progressively smaller impacts when diversity is higher. This simple model demonstrates that the magnitude of the effect of functional diversity, as measured by N, on ecosystem functioning depends on the range of interspecific differences in the species pool—that is, on the term ($R^*_{max} - R^*_{min}$) in Equation 1. This gives basis to the intuitive concept that diversity effects ecosystem processes because ecosystems with greater diversity have a greater range in those species traits that influence functioning.

The sampling effect model also predicts that the average quantity of unconsumed resource should decrease as diversity increases (**Figure 1b**). Indeed, in the sampling effect model, the increased biomass at higher diversity is caused solely by the more complete utilization of the limiting resource that occurs, on average, at higher diversity.

The model also illustrates the importance of species composition. Each point in the two graphs of **Figure 1** represents the response of a community with a different randomly determined species composition. Thus, the variability among plots with the same diversity measures the impact of composition, and the variability among diversity levels represents the impact of diversity. Both diversity and composition are strong determinants of productivity and resource levels in the sampling effect model.

The other major type of models that have been proposed to relate productivity to diversity are niche differentiation models. In essence, such models assume that a habitat is spatially or temporally heterogeneous, that species differ in the traits that determine their response to this heterogeneity, and that each species is a superior competitor, and thus is more productive, for some subset of the heterogeneous habitat conditions. These assumptions can allow a large number of species to coexist and assure that ecosystem productivity increases, on average, as diversity increases. For instance, two factors, such as soil pH and temperature, might limit plant abundance. Each species could have some combination of these factors at which it performed best. Such niche differentiation would mean that each species did best in a part of the habitat, but that no species could fully exploit the entire range of conditions.

The essence of such niche models can be captured by making the simple assumptions that each species has a circular area of radius r in which it can live and be a good competitor (**Figure 2a**), that all species attain comparable abundances per unit habitat occupied, and that competition similarly reduces abundances of all overlapping species. If the values for one limiting factor range from 0 to a r and the other from 0 to b r, where a and b measure habitat heterogeneity for factors 1 and 2, and if species are drawn at random from all those that could live at some point in the habitat, then total community biomass (i.e., the proportion of environmental conditions "covered" by one or more species) would be

$$B_{(N)} = 1 - \left(1 - \frac{\pi}{ab + 2(a+b) + \pi}\right)^N \qquad (2)$$

here N is species diversity. $B_{(N)}$ is an increasing function of species diversity (**Figure 2b**). The amount of unused habitat decreases as diversity increases, much as the concentration of unutilized resource was decreased for the sampling effect model. As for the sampling effect model, the variance within a given level of diversity is caused by differences in species composition, and differences between diversity levels is caused by diversity.

In addition, the niche model predicts that greater habitat heterogeneity (i.e., greater values of a and b) requires greater diversity in order to achieve a given level of productivity. In general, heterogeneity should increase with habitat size, leading to the prediction that greater biodiversity is required

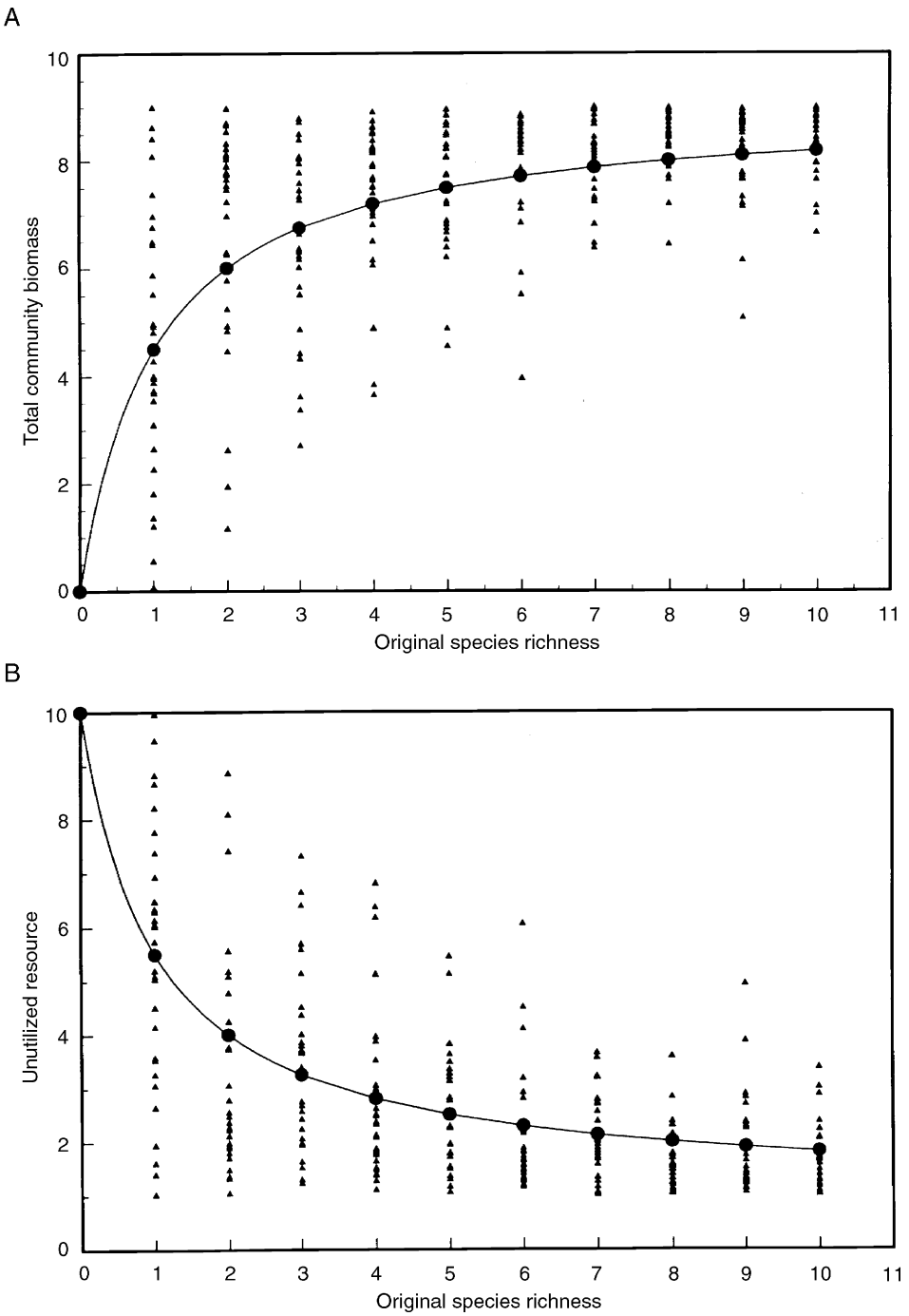


Figure 1 (A) The sampling effect model predicts that productivity should be greater at greater functional diversity, here measured by the number of species present. The variation within a given level of species richness is caused by different species compositions. (B) Productivity is higher in plots with greater functional diversity because of greater capture of the limiting resource. The concentration of unutilized resource is predicted to decline as diversity increases.

to attain a given level of productivity in larger habitats. For instance, for small, relatively homogeneous habitats ($a = b = 1$), only six species are needed to attain 95% of maximal productivity. However for spatially heterogeneous habitats ($a = b = 10$), a diversity of 135 plant species is needed to achieve this level.

A comparison of the sampling effect model with the niche differentiation model reveals a major difference in the expected pattern of the dependence of productivity on diversity. For the sampling effect model, there are no higher diversity plots that are more productive than the most productive monoculture. In contrast, for the niche model, there are

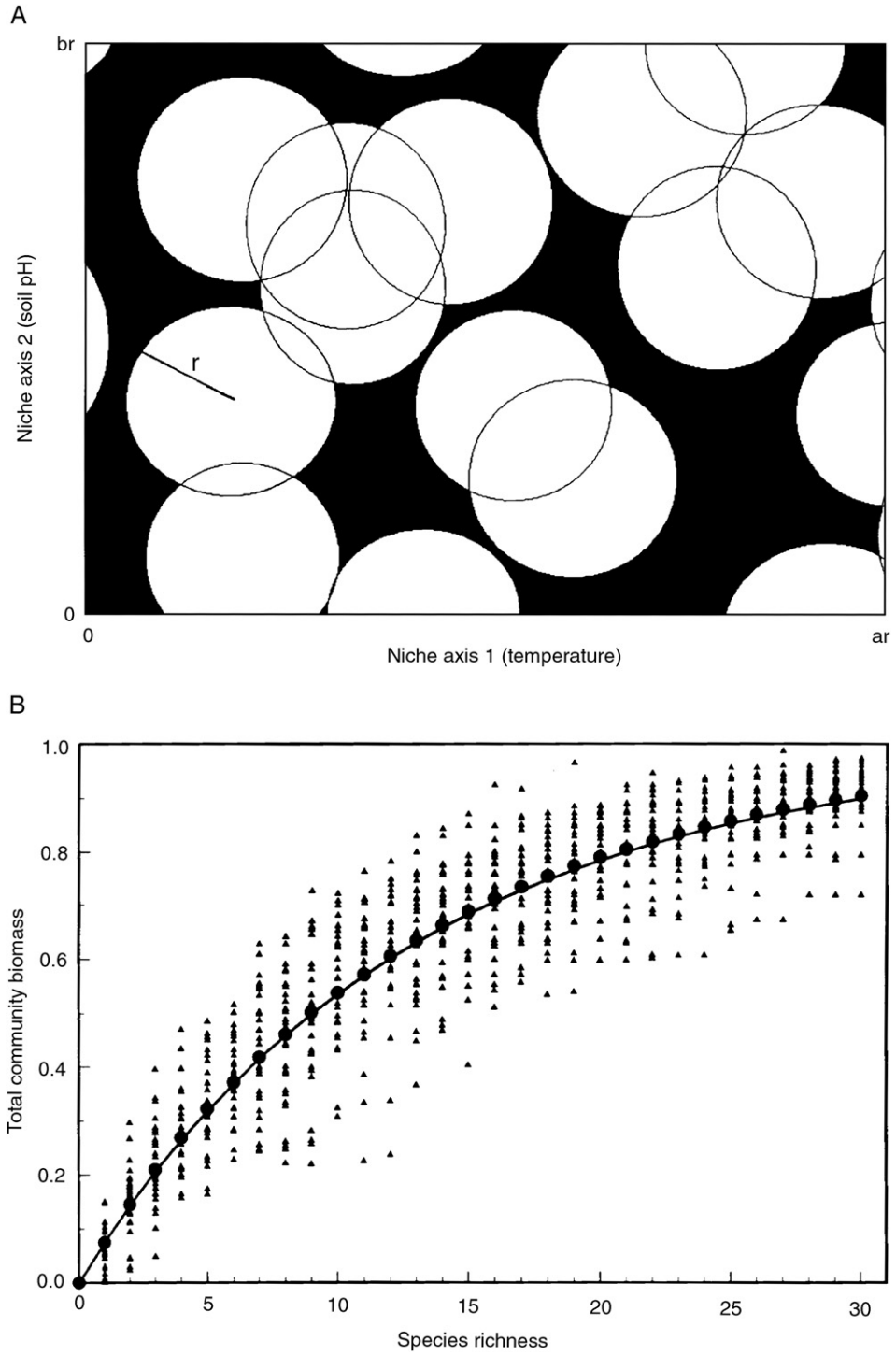


Figure 2 (A) A graphical illustration of a niche differentiation model. Here each circle represents the range of environmental conditions in which a given species can live, and the full rectangle shows the range of environmental conditions that occur in a given habitat. This model and similar niche differentiation models predict that productivity should be an increasing function of diversity. (B) The predicted effects of diversity on productivity for the model illustrated in part (A).

two-species plots that are more productive than the most productive monoculture, three-species plots that are better than the best two-species plot, and so on. For ecosystems that meet the assumptions of the sampling effect model, which might occur for highly productive agricultural fields, there might be situations in which judicious choice of the right species and variety could lead to as great productivity from a monoculture as would be possible for a highly diverse mixture

of species. In contrast, for habitats with spatial or temporal heterogeneity, which should occur for almost all natural ecosystems and for all but the most intensively managed ecosystems, niche differentiation models are more likely to hold. In such cases, increased diversity is expected to lead to greater productivity and to more complete use of limiting resources.

Although these models, and the models of Michel Loreau, have predicted that greater diversity can lead to greater ecosystem productivity, this need not always be the case. For instance, if the assumptions of the sampling effect model were modified to have progressively better competitors be progressively less productive, productivity would be a decreasing function of diversity. This suggests a more general principle: if species differ in their competitive abilities, and if higher competitive ability is correlated with some other traits, then these traits will, on average, be better represented in more diverse communities, thus biasing the functioning of these communities in the direction determined by these correlated traits.

Experimental Studies

Darwin suggested that it was common knowledge among farmers that a greater diversity of pasture plants would lead to a greater production of herbage in pastures. In his 1993 chapter, McNaughton cited this and presented more recent examples in which greater plant diversity led to greater productivity, as did Swift and Anderson. Indeed, earlier work reviewed in Harper's 1977 book showed that pairs of coexisting species often yield more than either species did when living by itself. As reviewed in the 1993 chapter by Vitousek and Hooper, some of the first evidence linking higher plant diversity to greater retention of soil nutrients came from a field experiment in Costa Rica by Ewel as collaborators. They found that communities planted to many tropical species generally retained more soil fertility than those planted to monocultures.

The first published direct test of the diversity-productivity hypothesis came from a greenhouse experiment by Naeem *et al.* (1995). By growing various randomly chosen combinations of 16 plant species 1, 2, 4, 8, or 16 at a time in a greenhouse, they found that community biomass was greater at higher plant diversity (**Figure 3a**). This team performed another experiment in a series of growth chambers and also had results suggesting that greater diversity leads to higher productivity (Naeem *et al.*, 1994). Next came results from a large-scale field experiment begun in Minnesota in 1993 (**Figure 4**). Its 147 plots, each 3 m × 3 m, were planted to contain 1, 2, 4, 6, 8,12, or 24 plant species randomly and independently chosen from a set of 24 prairie-grassland species (reviewed in Tilman, 1999). It found highly significant effects of plant diversity on both productivity (**Figure 3b**) and on the soil concentration of the limiting resource, nitrate (**Figure 3c**). By the fifth year of this experiment, its results supported niche differentiation models more than the sampling effect model as the major cause of the effects of diversity on the measured ecosystem processes. Indeed, the most productive plot in 1998 was a 24-species plot that had 65% greater total biomass than the most productive monoculture. A second experiment, adjacent to this Minnesota experiment (reviewed by Tilman, 1999), controlled for both species diversity and functional group diversity (**Figure 5**). Its results were similar to those of the first experiment and showed highly significant effects of species diversity, functional group diversity, and functional group composition on primary productivity and nutrient dynamics. In both of the Minnesota grassland diversity experiments, the vast majority of species coexisted in all plots to which they had been added, further supporting niche differentiation models.

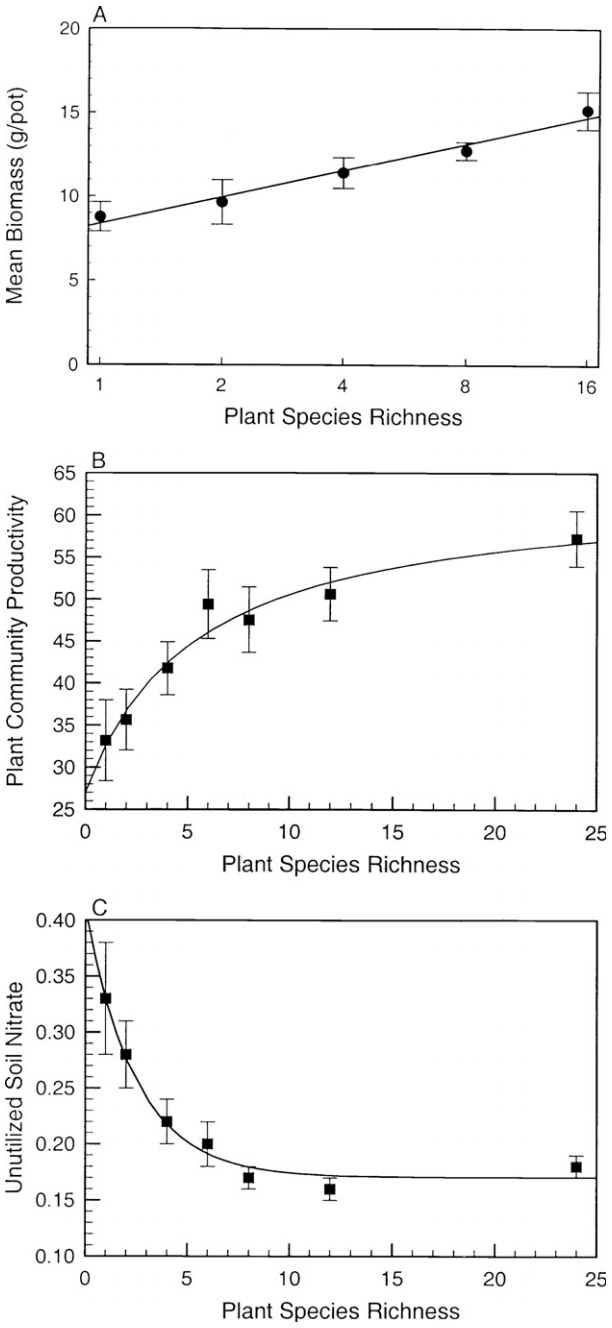


Figure 3 (A) The observed effect of plant diversity on the productivity of plant communities in the greenhouse experiment of Naeem and collaborators. (B) Effects of diversity on productivity for the Minnesota field experiment in which grassland diversity was experimentally controlled in 147 plots. (C) Effects of diversity on the concentration of unutilized soil nitrate for the Minnesota experiment.

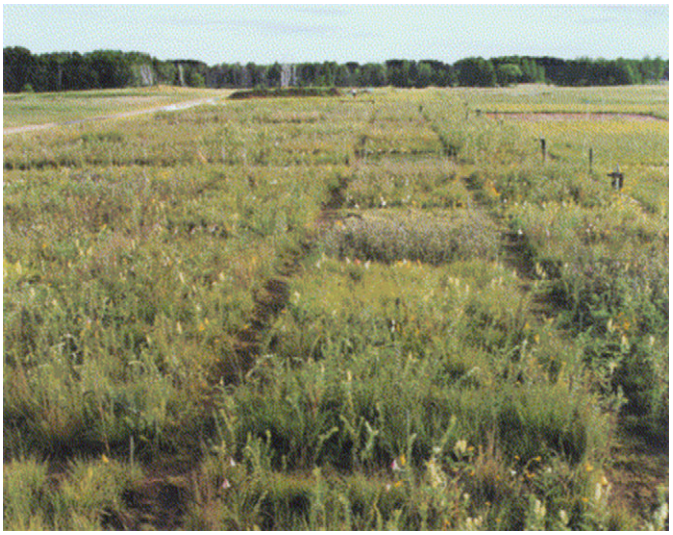


Figure 4 The smaller of the Minnesota biodiversity experiments, shown here, has demonstrated that plant diversity has a strong effect on ecosystem productivity and nutrient dynamics. The experiment has 147 plots, each being 3 m × 3 m (about 10 feet by 10 feet) in size.



Figure 5 The larger of the Minnesota biodiversity experiments uses about 270 of the 342 plots shown. It has shown strong effects of plant species richness, plant functional group richness, and plant functional group composition on ecosystem processes. Each plot is 13 m × 13 m (about 40 feet by 40 feet).

Knops *et al.* (2000) recorded the number of non-planted species that invaded the Minnesota diversity experiment plots, and their biomass at the time when they were removed from the plots. They found that significantly fewer species invaded higher diversity plots and that the total biomass of invading species was lower in higher diversity plots. Further analyses suggested that the effect of diversity on invasions was caused by the lower levels of soil nitrate in higher diversity plots. This provides one simple mechanism whereby diversity may influence the extent to which an ecosystem is invaded by other species and suggests that levels of unconsumed limiting resources may, in general, be an important determinant of the success of an invading species.

For native, undisturbed grasslands close to the two Minnesota biodiversity experiments, plant abundances were greater and soil nitrate was lower in more diverse plots (see Tilman, 1999), which is consistent with the experimental results and with the predictions of theory. However, correlational patterns must be interpreted carefully because they could be confounded by other correlated variables. Michel Loreau used a model that linked environmental factors, biodiversity, and ecosystem functioning to explore this point. The

model illustrated that correlational field data could be misinterpreted easily because of a confusion of cause-and-effect relationships. Just such issues cloud the interpretation of the possible effects of island diversity on ecosystem processes for a study of 50 Swedish islands. In an intriguing study that showed links between island size and the frequency of wildfire, David Wardle and collaborators found that a suite of ecosystem traits were correlated with both island size and plant diversity. However, it is unclear if diversity caused the observed differences in ecosystem processes or if both these processes and diversity were controlled by fire frequency.

Hooper and Vitousek (1998) performed a field experiment, planted in 1992, in which they controlled plant functional group diversity and composition using plants common to California grasslands. After a year of growth, they found that functional group composition had a much greater effect on plant community biomass than functional group diversity, but that the utilization of soil nutrients increased significantly as diversity increased.

In a 4-month greenhouse experiment, Symstad *et al.* (1998) found that total plant biomass was significantly higher at higher diversity and that most of this effect was attributable to the presence of legumes. They also determined the effects of the deletion of individual species on total biomass and found that the strength and direction of these effects depended on which species were present and which was deleted.

In an experiment that was replicated at eight different sites across Europe, ranging from Scotland and Ireland to Portugal and Greece, Hector *et al.* (1999) found that greater plant diversity led to greater primary productivity. An important finding of this unique experiment was that the quantitative effect of diversity on primary productivity was the same across all eight sites. In combination with the other field and laboratory experiments, the European experiment suggests that there is a general, repeatable effect of grassland diversity on primary productivity.

In total, these studies show that plant productivity is greater at greater diversity and that this also corresponds with greater utilization of limiting soil resources. In general, short-term experiments showed weaker effects of diversity on productivity and soil nutrients than longer-term experiments. This is expected because diversity should impact ecosystem processes via changes in plant abundances mediated by competition, and such interactions can require several years to occur. Further work is needed on other trophic levels and in other communities to determine the extent to which the patterns observed to date apply to other trophic levels (e.g., herbivores, predators) or to other communities (e.g., marine fisheries, forest ecosystems, coral reefs).

Functional Diversity and Stability

Theory and Concepts

A large number of authors, including Charles Elton, Robert May, Stuart Pimm, and Sam McNaughton have contributed considerable insights into the effects of diversity on stability. May (1972), for instance, showed that the abundances of individual species become progressively less stable as the diversity of the community in which they live increases. Several recent papers have explored the effects of diversity on the stability of communities of competing species (Doak *et al.*, 1998; Ives *et al.*, 1999; Tilman, 1999). The first two of these papers showed that the temporal variability of an ecosystem process, such as ecosystem productivity, is expected to be lower when the ecosystems contain more species. This can occur for the same reason that a portfolio composed of many different types of stock tends to be more stable than one containing stock of a single company. An additional factor that can cause ecosystem functioning to be more stable for more diverse ecosystems is competition. When some disturbance harms one species, the species with which it interacts experience less competition. This allows these competitors to increase in abundance. Their greater abundance partially compensates for the decreased abundance of the first species, thus stabilizing the functioning of the ecosystem. Ives *et al.* (1999) showed that increased diversity only led to increased stability when the species differed in their responses to habitat fluctuations and disturbances. Because such differences are a direct measure of functional diversity, the work of Ives *et al.* (1999) showed that increases in functional diversity lead to greater stability. For a thorough treatment of theory relating diversity and stability, see "Stability, Concept of."

Experimental and Observational Studies

The evidence that led Elton to propose the diversity-stability hypothesis was anecdotal. In his 1993 chapter, and in earlier papers, McNaughton defended the diversity-stability hypothesis by citing several observations and experiments in which greater diversity was associated with greater stability. A variety of other studies, summarized in Tilman (1999), also have found effects of diversity on stability. For instance, a study by Frank and McNaughton of eight grassland sites within Yellowstone National Park found that those with greater plant species diversity had smaller shifts in plant community compositions during a severe drought. Two British ecologists, Taylor and Woiwod, performed a long-term project in which they monitored the abundances of hundreds of insect species at a large number of sites. The data they collected provide evidence that supports the hypothesis that more diverse insect communities should be more stable. The greater stability is expected because of the statistical averaging (or portfolio) effect pointed out by Doak *et al.* (1998). Specifically, because the temporal variances in the abundances of individual species in this community scales as their abundance to a power of about 1.6, the portfolio effect should cause more diverse insect communities to have lower temporal variability.

Several authors have found that greater oak tree diversity stabilizes the population density of an animal, the acorn woodpecker, that feeds on the seeds of the trees (see Koenig and Haydock, 1999). Acorn woodpeckers are highly dependent on acorns as a source of food, but oaks produce acorns as a mast seed crop. Masting means that there is great year-to-year variability in the rate of acorn production. There is a striking decrease in the year-to-year variability of acorn woodpecker abundances for woodpeckers living in habitats containing a greater diversity of oaks. Thus, greater oak diversity led to more stable acorn woodpecker populations.

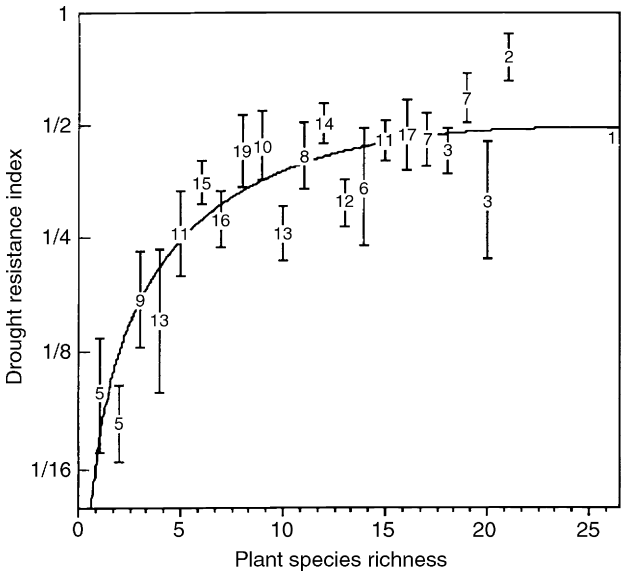


Figure 6 The resistance of Minnesota grassland ecosystems to drought was highly dependent on their plant biodiversity. Ecosystems containing a large number of plant species had their productivity fall to about half of its predrought levels during a severe drought, but those containing only one or two plant species had it fall to about 1/8 to 1/12 of the predrought level.

Moreover, acorn woodpecker densities were much lower for areas with a single oak species than for those with several.

A long-term experiment in Minnesota provides additional evidence suggesting that greater plant diversity leads to greater stability (reviewed in Tilman, 1999). In a series of 207 plots annually monitored from 1982 to 1999, total plant community biomass was found to be more stable in plots containing more species. Both in response to a major disturbance, a severe drought (**Figure 6**), and in response to normal year-to-year variation in climate (Tilman, 1999), plots with greater diversity had lower year-to-year variability in their total plant biomass. In particular, the severe drought caused plant biomass to fall to half of its predrought level in plots with about 15 or more species, but caused it to fall to 1/8 to 1/12 of its predrought levels in plots containing one or two plant species (**Figure 6**), Similarly, year-to-year variation in total biomass fluctuated about twice as much in low diversity as in high diversity plots (Tilman, 1999). Although total community biomass was more stable at higher diversity, analyses of the stability of individual species showed that these declined slightly but detectably, at higher diversity. Thus, diversity stabilized total community biomass at the same time that it destabilized the abundances of individual plant species. Plant diversity and composition were confounded in this experiment because both changed in response to nitrogen addition. Multiple regression, used to control for this confounding, found highly significant effects of diversity on stability for both cases. These analyses also showed that species composition and functional group composition also had significant effects on stability.

McGrady-Steed, Harris, and Morin (1997) found, in a laboratory study of the effects of diversity in microbial communities, that the temporal variability was significantly smaller at higher diversity. Indeed, a four-fold increase in diversity led to about a three-fold decrease in the temporal variability of whole-community net respiration, a measure of ecosystem activity. The rate of microbial decomposition of particulate organic matter also increased with diversity in this study. Finally, they found that greater diversity led to lower susceptibility to invasion by another species, but that invader success was highly dependent on community composition. Naeem and Li (1997) similarly found that greater diversity led to greater reliability, which was measured as the lower variability in total community biomass among communities of identical diversity. This effect was also apparent in the greenhouse experiment that Naeem and collaborators had performed earlier.

In total, these studies provide strong evidence that communities with greater diversity are more stable and suggest that individual species in such communities may be less stable. Theory, experiment, and observation are in general agreement, but this topic merits additional exploration.

Conclusions

The research performed to date illustrates that a variety of different ecosystem processes are impacted by the number and kinds of species living in the ecosystem. This work illustrates that species differ in traits that influence ecosystem functioning and suggests that ecosystem processes depend on the range in those traits represented in the ecosystem. However, there are, as yet, no clear demonstrations of the specific traits that are relevant to particular ecosystem processes and no simple ways to directly measure functional diversity. Rather, correlates of functional diversity, such as species richness or functional group richness, remain the best, albeit indirect, way to measure functional diversity.

See also: C_4 Plants. Ecosystem Function, Principles of. Functional Groups. Habitat and Niche, Concept of. Stability, Concept of

References

Chapin FS III, Walker BH, Hobbs RJ, *et al.* (1997) Biotic control over the functioning of ecosystems. *Science* 277: 500–504.

Doak DF, Bigger D, Harding EK, Marvier MA, O'Malley RE, and Thomson D (1998) The statistical inevitability of stability-diversity relationships in community ecology. *Am. Nat.* 151: 264–276.

Grover JP (1997) *Resource Competition*. London: Chapman & Hall.

Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

Hooper DU and Vitousek PM (1998) Effects of plant composition and diversity on nutrient cycling. *Ecological Monographs* 68(1): 121–149.

Ives AR, Gross K, and Klug JL (1999) Stability and variability in competitive communities. *Science* 286: 542–544.

Knops JMH and Tilman D (2000) Dynamics of soil nitrogen and carbon accumulation for 61 years after agricultural abandonment. *Ecology* 81: 88–98.

Koenig WD and Haydock J (1999) Oaks, acorns, and the geographical ecology of acorn woodpeckers. *Journal of Biogeography* 26: 159–165.

May RM (1972) Will a large complex system be stable? *Nature* 238: 413–414.

McGrady-Steed J, Harris PM, and Morin PJ (1997) Biodiversity regulates ecosystem predictability. *Nature* 390: 162–165.

Naeem S and Li S (1997) Biodiversity enhances ecosystem reliability. *Nature* 390: 507–509.

Naeem S, Thompson LJ, Lawler SP, Lawton JH, and Woodfin RM (1994) Declining biodiversity can alter the performance of ecosystems. *Nature* 368: 734–737.

Naeem S, Thompson LJ, Lawler SP, Lawton JH, and Woodfin RM (1995) Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Philos. Trans Royal Soc. London B* 347: 249–262.

Schulze ED and Mooney HA (1993) *Biodiversity and Ecosystem Function*. Berlin: Springer Verlag.

Symstad AJ, Tilman D, Willson J, and Knops JMH (1998) Species loss and ecosystem functioning: Effects of species identity and community composition. *Oikos* 81: 389–397.

Tilman D (1999) The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80: 1455–1474.

Functional Diversity Measures

Norman WH Mason, Landcare Research, Hamilton, New Zealand
David Mouillot, James Cook University, Townsville, Queensland, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Beta diversity Portion of regional diversity due to differences between local communities in their taxonomic, functional, or phylogenetic composition.

Community assembly processes Processes which influence the assembly of species to form communities – particularly those processes that influence the probability of occurrence and the abundance of species in local communities.

Ecosystem function The rate, level, or temporal dynamics of one or more ecosystem processes, such as primary productivity.

Environmental filtering Constraints on the presence and abundance of species in local communities or regional species pools due to abiotic conditions.

Evenness The equitability in abundance of the species present in a local community or region.

Functional trait space Space where species are located according to their functional traits. Functional trait space may be multidimensional, with each functional trait representing a separate dimension.

Hierarchical partitioning of diversity Process of partitioning regional diversity between alpha (local) and beta components.

Local community Collection of interacting or potentially interacting individuals or species. The spatial scale covered by a local community depends on the scale on which interactions occur.

Niche complementarity Enhancement of taxonomic diversity and ecosystem functioning through differences in functional traits between species that co-occur in local communities.

Regional species pool Species occurring within a defined geographical area or in a collection of samples from within such an area.

Introduction

Biodiversity is a multifaceted concept, and this is perhaps best illustrated through the definition supplied by Gaston (1996) as "the variety of life on Earth at all its levels, from genes to ecosystems, and the ecological and evolutionary processes that sustain it." Quantification of such a broad concept is obviously challenging, but we cannot begin to understand the causes of biodiversity loss and the consequences of this loss for ecosystem functioning unless we can devise methods for the measurement of biodiversity. The increasing interest in the quantification of functional diversity (FD) arises from the observation that species differ both in their response to disturbance or environmental stress and in their influence on ecosystem functioning. This observation has led to the general realization that traditional measures of taxonomic diversity are inadequate for understanding the relationship between biodiversity, environmental filters, and ecosystem functioning. FD measures provide a means of examining the mechanisms by which functional differences between species influence the response of biodiversity to environmental stress (and disturbance) and the effect of biodiversity on ecosystem functioning. This article provides a definition of FD from which independent measures can be derived, identifies key criteria required for valid FD measures, reviews existing FD measures, and provides some examples of their application.

Defining Functional Diversity

Quantification of functional diversity requires a definition of FD from which measureable concepts can be derived. Many definitions have been provided by different authors, but these generally do not provide much guidance in deriving FD measures. Most definitions agree that FD is the diversity in the functional traits of the species in an assemblage. However, this general definition contains two terms that need to be clarified further: (1) diversity as applied to functional trait data and (2) which traits are functional traits. For many years, taxonomic diversity has been separated into two independent components – species richness and evenness in species abundance. Mason *et al.* (2005) demonstrated that these components have their analogs in FD (i.e., functional richness and functional evenness), while also identifying a third independent component – functional divergence (**Figure 1**). They gave a definition of FD as the distribution of the species and abundance of a community in functional trait space, including:

1. the amount of functional trait space filled by species in the community (functional richness);
2. the evenness of abundance distribution in filled trait space (functional evenness); and
3. the degree to which abundance distribution among species maximizes divergence in functional traits (functional divergence).

Defining FD in this way helps to guide the development of measures for its quantification. Each of the components represents a mathematical concept that can be quantified using indices that vary on a continuous rather than a categorical scale. This definition of FD as three independent components also shows that no single index can give a complete quantification of FD. This definition also provides a means for

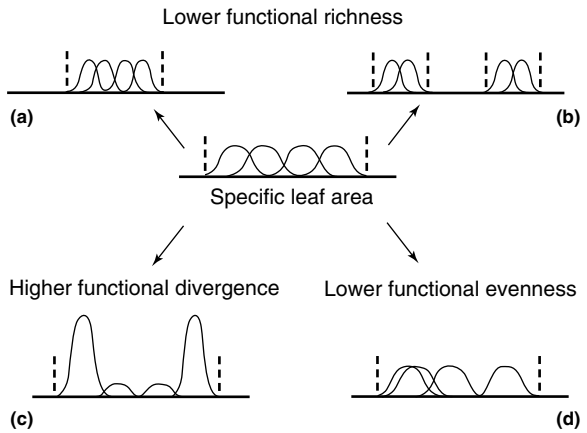


Figure 1 The primary components of functional diversity. The vertical axes represent abundance (e.g., biomass). The bell-shaped curves indicate the distribution of the abundance of individual species in functional trait space. The vertical dotted lines indicate the amount of niche space filled by the community as a whole. Comparison of the central diagram with those in the corners demonstrates how changes in the distribution of species and abundance in functional trait space affect each component. Functional richness decreases either if the range of functional trait values is narrower (a) or if there is unoccupied trait space between the maximum and minimum trait values (b). Functional richness can decrease without a change in functional evenness or divergence (a). Similarly, functional evenness and functional divergence can increase or decrease without a change in functional richness if the amount of niche space filled is unchanged (c and d). In (c) the increase in functional divergence is associated with a decrease in functional evenness, but these components generally vary independently of each other (see Mouchet *et al.*, 2010). Functional evenness is lower in (d) as the decrease in the evenness of spacing between species causes abundance to become more concentrated at the lower end of the trait gradient and more sparsely distributed at the upper end (thus resulting in a less even distribution of abundance in trait space).

arranging the growing range of published FD measures into a typology. This is helpful for ecologists attempting to choose indices for the quantification of FD in their study assemblages. For a full quantification of FD, a separate index is required for each component. Because indices that estimate the same component are almost always highly correlated, only a single index is required for each component.

Lavorel *et al.* (1997) provide a succinct definition of a functional trait as: "A feature of an organism, which has demonstrable links to the organism's function." Most studies seem to conform to this definition, with functional traits considered to be characteristics of individual organisms that act as indicators of functions occurring at the scale of individual organisms, though sometimes the functions themselves are considered functional traits (Violle *et al.*, 2007). Some confusion has been caused by studies that consider properties measured at the scales larger than an individual to be functional traits (e.g., net primary productivity). Violle *et al.* (2007) provide a typology accommodating the variety of senses in which the term has been applied, but this article considers only functional traits measured at the individual scale.

There is a growing body of primary research demonstrating that measureable attributes of individual organisms are linked to functions at the individual level. These processes might include growth rate, response to environmental stress, the types of resources used, where and when resources are acquired, and how resources are allocated within the organisms. For example, specific leaf area (SLA) is an indicator of maximum rates of growth in plants. Mouth size is an indicator of prey size in fish. Leaf tissue density indicates individual resource turnover rates and tolerance of drought stress. The most important consideration in choosing which traits to measure is that they are indicators of functions relevant to the question FD indices are being used to answer. For example, to test if more functionally diverse plant communities have higher primary productivity, traits that indicate growth rates and complementarity in resource use should be measured (e.g., SLA for growth rate and leaf inclination for complementarity in light capture).

Criteria Required for Valid Functional Diversity Indices

The primary criterion for all FD measures is that they should be independent of existing biodiversity measures, such as species richness. If FD indices merely replicate the information provided by other measures, then they will do nothing to enhance our understanding of ecological processes. Several authors have suggested differing sets of criteria in addition to this for FD indices. However, given that indices intended to measure each FD component will necessarily have differing behavior, not all criteria will be applicable to all FD indices. **Table 1** provides a list of criteria, with their applicability to each component of FD.

A Review of Existing Measures

A range of existing FD measures is reviewed, though this should not be considered an exhaustive list. Further reading is available in existing reviews of FD measures by Schleuter *et al.* (2010), Pavoine and Bonsall (2010), and Mouchet *et al.* (2010).

In reviewing existing measures, we group them according to each of the three FD components. This is not possible for all indices, because some are a combination of functional richness and functional divergence. There is also an assessment of the validity of each index against the criteria listed. Although the original definition of FD components by Mason *et al.* (2005) explicitly incorporates intraspecific variation in functional traits, few indices make an attempt to do so. Increasingly consideration is being given to the inclusion of intraspecific trait variability in deriving new FD indices (Cianciaruso *et al.*, 2009; De Bello *et al.*, 2010b; Schleuter *et al.*, 2010), and indices that do this are noted where they are available.

Functional Richness

Although this component contains perhaps the largest number of indices, they are based primarily on four major approaches. First of all is the functional group richness approach, where species are categorized either into *a priori*

Table 1 Criteria required for valid functional diversity indices. Yes and No indicate whether or not criteria are relevant for each functional diversity component

Criterion	*Functional diversity component*		
	Richness	*Evenness*	*Divergence*
1: Be constrained to a 0–1 range (for convenience) and use that range well	Yes	Yes	Yes
2: Reflect the range of character values present	Yes	No	No
3: Reflect the contribution of each species in proportion to its abundance	No	Yes	Yes
4: Decrease when the abundance of a species with an extreme trait value decreases	No	No	Yes
5: Be unaffected by the units in which traits are measured	Yes	Yes	Yes
7: Be unaffected by the units in which the abundance is measured	Yes	Yes	Yes
8: Be unaffected by the number of species	Yes	Yes	Yes
9: Be unaffected when a species is split in two (i.e., is replaced by two with the same character value, with the same total abundance)	Yes	Yes	Yes
10: Have set monotonicity (a set of species is as or more diverse than any subset drawn from it)	Yes	No	No
11: Have concavity (average diversity of local communities is less than the diversity of regional species pool)	Yes	No	No

groups, or groups defined by functional traits. Functional richness is measured as the number of functional groups present. Since this approach is dealt with at length in the "functional diversity" entry of this encyclopedia, it is not considered further here.

Second, there are a variety of dendrogram-based FD measures, the first of which was published by Petchey and Gaston (2002). These measures begin by building a matrix of pairwise distances between species in functional trait space. They then use hierarchical classification techniques to cluster species together based on these distances. Usually, a single dendrogram is built for all the species within a study area, or species pool. FD values for local communities (or other subsets of the species pool) are estimated as the total branch length required to link the species present in the community on the dendrogram (**Figure 2**). Taking this approach allows dendrogram FD measures to satisfy the set monotonicity and concavity criteria, since FD of local communities will always be lower than for the entire assemblage, and FD will increase monotonically with species richness. Dendrogram FD measures can be constrained to a 0–1 range if raw FD values for subsets of the species pool are expressed as a proportion of the total branch length of the dendrogram including all species in the pool. The dendrogram approach has recently been extended to include intraspecific variation, but the behavior of this approach has been little explored (Cianciaruso *et al.*, 2009).

One of the problems with dendrogram-based measures is that distances between species on the dendrogram are often poorly correlated with species distances in functional trait space. This arises as dendrograms are an attempt to reduce the dimensionality of functional trait space to allow relationships between species to be visualized. This necessarily entails some distortion of species distances in functional trait space. One attempt to minimize this distortion has been to choose a dendrogram from a range of potential distance metric and clustering method combinations, including "consensus trees" (which build a composite dendrogram that retains only the nodes supported by the majority of the dendrograms considered; Mouchet *et al.*, 2008). However, in most instances,

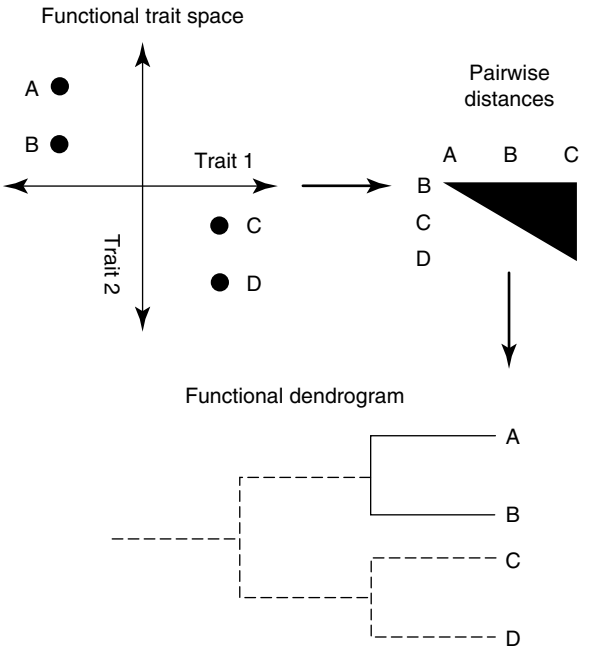


Figure 2 A simplified example illustrating how species distribution in two dimensional functional trait space is converted to a matrix of pairwise distances in trait space between species, which is then used to build a functional dendrogram. In a community containing only species A and B, functional richness is the summed length of the solid lines. The functional richness of the species pool is the sum of all branch lengths. Reproduced from Poos MS, Walker SC, and Jackson DA (2009) Functional-diversity indices can be driven by methodological choices and species richness. *Ecology* 90: 341–347, with permission from Wiley.

even the strongest correlation between distances in functional space and on dendrograms is relatively low ($r<0.4$). This discrepancy between functional dendrograms and species distribution in functional trait space brings into question whether dendrogram measures are suitable as FD indices following the definition above. Another issue with dendrogram-based FD

measures is that they tend in practice to be very strongly correlated with species richness. The effect of species richness on dendrogram-based FD dominates variation in FD between communities, so that differences in FD between communities generally reveal nothing more than the fact that those communities differ in species richness (Poos *et al.*, 2009).

The third major approach to the estimation of functional richness is convex hull volume (CHV, Cornwell *et al.*, 2006). This approach defines the minimum volume of functional trait space required to include the species within a local community. For a single functional trait, CHV is simply the range of trait values (i.e., the FRi of Mason *et al.*, 2005), but when two or more traits are involved, a complex algorithm is required to define CHV. CHV conforms to the set monotonicity criterion, since the addition of new species will result in either no change, or an increase in hull volume. CHV also conforms to the concavity criterion, since the hull volume of the regional species pool will always be greater than or equal to that of any local community. CHV can be converted to a 0–1 range if values for local communities are divided by those for the entire species pool.

The convex hull approach tends to be less sensitive to species richness than dendrogram-based approaches (Mouchet *et al.*, 2010). However, one potential drawback is that CHV is undefined when the number of species is less than or equal to the number of traits. For example, in a two-dimensional functional trait space, at least three species are required to form a convex hull. This means that the number of dimensions in functional trait space must be less than the species richness of the least species-rich community. This can cause a tradeoff between the inclusion of more traits and the inclusion of species-poor communities. One solution to this problem has been to reduce the dimensionality of functional trait space using ordination techniques (**Figure 3**), such as principal components analysis (PCA) or principal coordinates analysis (PCoA). However, reducing the number of dimensions runs the risk of losing some of the information from the raw trait data.

The final approach (FRis and FRim) differs from dendrogram and convex hull based estimates of functional richness in that it incorporates intraspecific trait variability. This permits estimation of not only the volume of trait space occupied by species, but also how completely this occupied trait space is filled (Schleuter *et al.*, 2010). For one-dimensional trait space the index FRis uses intraspecific trait variability to find the union of the trait ranges of individual species (i.e., the combined range of functional trait values occupied by individual species, **Figure 4**). In multidimensional trait space, the index FRim uses intraspecific trait variability to define a membership probability function, which estimates the probability that a species "belongs to" a given point in trait space. The functional richness of the community is then computed as the integral of the maximum membership probability values across the species present in the local community (illustrated for a single trait in **Figure 4**). This index conforms to both the set monotonicity and the concavity criteria since the functional richness of the regional species pool will always be greater than or equal to that of local communities or other subsets of species. Both the single and multidimensional versions of this approach can be converted to a 0–1 range if values for local communities are divided by those for the entire species pool.

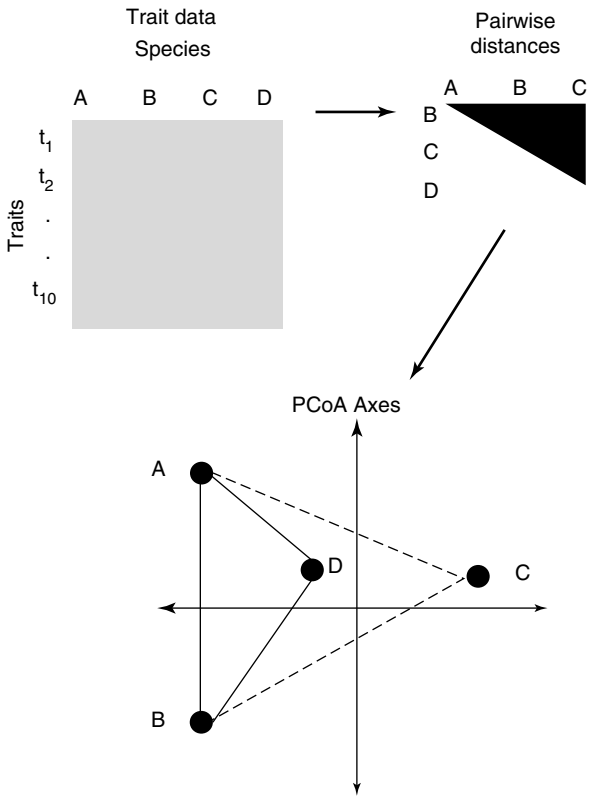


Figure 3 A simplified example illustrating the need for ordination to reduce dimensionality in functional trait space for the calculation of convex hull volume when the number of traits is greater than the number of species. In this example we have 10 traits and four species. The trait data are used to calculate pairwise distances between species using all 10 traits, then principal coordinates analysis (PCoA) is used to derive ordination axes from these distances. Reducing the dimensionality of functional trait space to the first two PCoA axes allows us to calculate convex hull volume for communities containing at least three species. The functional richness of a local community containing species A, B and D is the volume enclosed by the solid lines. The functional richness of the species pool is the volume contained in the lines connecting species A, B and C.

The main advantage of these indices is that in considering not only the volume of functional trait space spanned by the species present, but also how much of that volume is occupied, they provide a closer representation of the functional richness concept as defined by Mason *et al.* (2005). However, because of the novelty of these indices, it remains unclear whether they are more strongly affected by species richness than ecological processes. Further tests of these indices following the methods of Mouchet *et al.* (2010) and Poos *et al.* (2009) will help to better assess the practical usefulness of these indices. One drawback of these indices is that there is a degree of arbitrariness in defining the volume of functional trait space occupied by individual species. Schleuter *et al.* (2010) use the 5th and 95th percentile of trait values to define this, though other options, such as the minimum and maximum trait measurements, would be possible. It remains

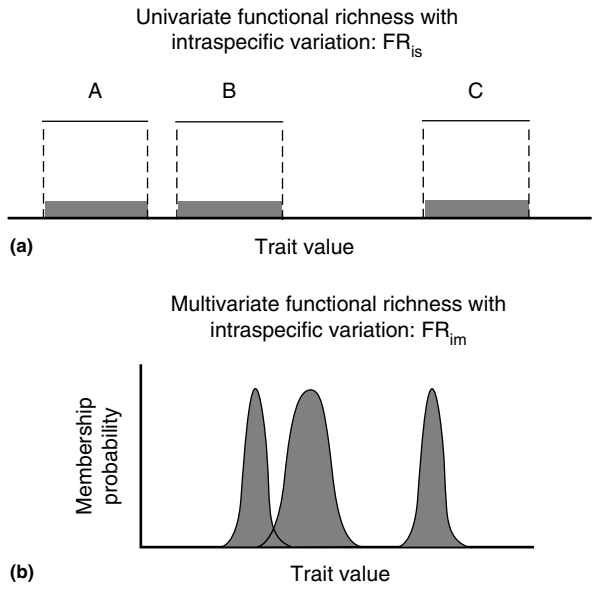


Figure 4 Functional richness measures incorporating intraspecific variation for either (a) single traits or (b) in multidimensional trait space (demonstrated using a single trait for simplicity). In (a), functional richness is the sum of the distances between the dotted lines, which indicate the range of trait space occupied by each species. In (b) the bell shaped curves indicate the membership probability (i.e., the probability that a species belongs to a particular point in functional trait space). The functional richness of the community in (b) is the integral of the maximum membership probability amongst the species present, which is represented by the gray shaded area.

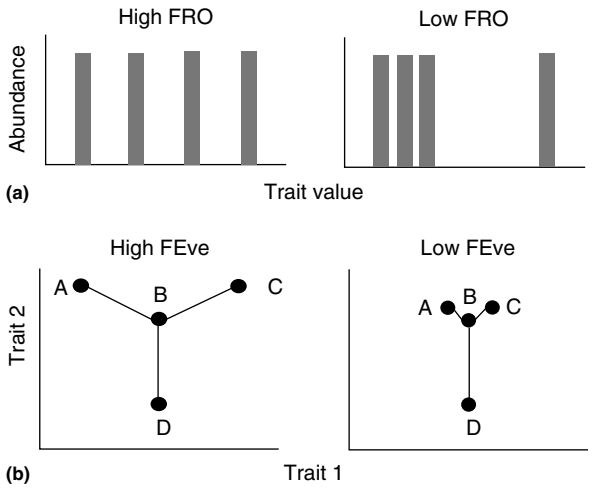


Figure 5 Illustration of communities with comparatively high and low values for the functional evenness estimator FRO (A) and its multivariate extension FEve (B). In the high FRO example, the perfect evenness in abundance and perfect regularity in spacing between nearest neighbor species along the trait gradient would give an FRO of 1. In the low FRO example, the variation in distances between nearest neighbors would cause a decrease in FRO. In the two FEve examples, species abundances are proportional to the diameter of the circles. In the high FEve example, the perfectly even abundances and the identical distances between nearest neighbor species in two dimensional trait space (as defined by the minimum spanning tree) would yield an FEve value of 1. The low FEve example would yield a decreased value due to the variability in nearest neighbor distances.

unclear whether this choice may have an effect on the outcome of functional richness comparisons between communities. Another issue is that these indices are unable to deal with categorical trait data, since they require intraspacific trait variation.

Functional Evenness

Compared with functional richness, there are comparatively few indices of functional evenness. The two existing indices are single and multidimensional measures of "functional regularity" (FRO, Mouillot *et al.*, 2005a), which is an extension of the Bulla (1994) index for evenness in species abundance. The single-dimension version measures the evenness in distances between nearest-neighbor species, weighted by relative abundance (**Figure 5**). This equates to an abundance-weighted measure of how evenly spaced species are along a functional trait gradient. This index was extended to consider multidimensional trait space with the index FEve, by using a minimum spanning tree algorithm to calculate the distances between nearest-neighbor species by Villeger *et al.* (2008). The original FRO index was then used to calculate the abundance-weighted regularity in these distances.

It should be noted that the concept of functional evenness is not *a priori* dependent on the regularity of species spacing in functional trait space. Rather it is solely dependent on how evenly abundance is distributed in trait space. However, there are no indices available that quantify this directly. Consequently, *FRO* and FEve represent indirect estimates of functional evenness. Nevertheless, both these indices conform to all the criteria relevant to an index of functional evenness, except that when a species is split into two functionally identical species with the same total abundance, FRO decreases. This will mainly be a problem when communities contain a large number of cryptic species (i.e., morphologically similar species that are difficult to separate in the field). Where there is a high level of confidence in species identification it will not be an issue. For the moment, the lack of an alternative approach to functional evenness measurement means that this is the best option available.

Functional Divergence

Since the original definition of functional divergence by Mason *et al.* (2005), there has been some confusion about how to measure it. Initially, Mason *et al.* (2005) proposed FDvar (*see* Indices Combining Functional Richness and Functional Divergence) as a measure of functional divergence. However, FDvar is affected by the range of functional trait values, so it is dependent on functional richness. As noted (*see* Defining Functional Diversity), the intention of dividing FD into three components was to separate it into independent, measureable concepts, so it is undesirable for indices measuring one component to be affected by another. Since then, several other measures have been presented, which are in effect a combination of functional richness and functional divergence, and these are reviewed below (*see* Indices Combining Functional Richness and Functional Divergence).

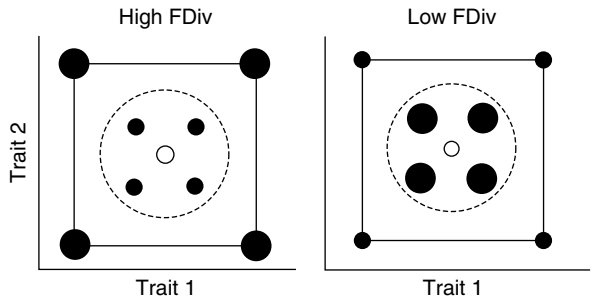


Figure 6 Illustration of the multivariate FDiv index for estimating functional divergence. The solid lines indicate the convex hull containing all the species present and the species connected by the solid line are the vertices of the convex hull. The diameter of the solid circles is proportional to species abundance. The open circle is the unweighted centre of gravity (i.e., group centroid) of the species forming the vertices of the convex hull. The dotted circle is the unweighted mean (taken across all species) of the species distances from the centre of gravity. In the example with a high FDiv value, distance of the most abundant species from the centre of gravity is greater than the mean distance. In the low FDiv example, the distance of the most abundant species from the centre of gravity is less than the average distance.

The only pure indices of functional divergence are the FDiv of Villeger *et al.* (2008) and the FDs of Schleuter *et al.* (2010).

FDiv measures the degree to which the abundance of a community is distributed toward the extremities of occupied functional trait space. It does this by first finding the convex hull for the species present (see **Figure 3**). It then finds the coordinates of the center of gravity for the convex hull by calculating the nonweighted mean across the species forming the vertices of the convex hull (i.e., the group centroid of the species forming the convex hull, **Figure 6**). Finally, it applies a series of equations to compare the observed sum of abundance-weighted Euclidean distances of species from the center of gravity relative to the maximum possible value. FDiv conforms to all the relevant criteria for a measure of functional divergence. The major drawback of FDiv is that it depends on the construction of a convex hull around the species in functional trait space. As noted above (*see* Functional Richness), this may constrain either the maximum number of functional traits or the minimum species richness allowable for calculation of FDiv, since the construction of a convex hull requires species richness to be greater than the number of trait dimensions. However, FDiv remains the only available multidimensional measure of functional divergence.

FDs is the abundance-weighted interquartile range of trait values, divided by the total range of trait values. In the original presentation Schleuter *et al.* (2010) provide for abundance weighting in finding the first and third quartile trait values by assuming species abundances are measured on a discrete scale (i.e., as counts of individuals). In this case, the quartile values are found simply by replicating species mean trait values in accordance with the number of individuals counted for each species. Obviously there are many instances where measurement of species abundance using counts of individuals is neither possible, nor desirable. In these cases, it will still be possible to calculate FDs since continuous abundance measurements may be used to weight quartile values as well. FDs conforms to all the criteria relevant for a measure of functional divergence. Its only limitation is that no multidimensional analog has been found.

Indices Combining Functional Richness and Functional Divergence

In recent years several indices have been developed, which incorporate both functional richness and divergence, since they are affected both by the range of functional trait space occupied and the degree to which the abundance of a community is distributed toward the extremities of this occupied space. Most of these indices are variants on the theme of abundance-weighted deviance of species trait values from the weighted community mean trait value. The other approach is Rao's quadratic entropy, which was applied as a FD measure by Botta-Dukat (2005).

The first index quantifying the abundance-weighted squared deviance of trait values across species was the FDvar of Mason *et al.* (2003). This index first calculates the abundance-weighted mean of log-transformed functional trait values across species (log-transformed values are used to ensure the index is independent of the units in which trait values are expressed). It then calculates the abundance-weighted sum of squared deviances of species log-transformed trait data from this weighted mean. Finally, the arc tangent of the resulting sum of squares is taken to ensure the index is bound between 0 and 1. FDvar conforms to all the relevant criteria for an index of functional divergence, except that it is sensitive to the range of functional trait values (**Figure 7**). FDvar does not conform to the set monotonicity and concavity criteria for an index of functional richness. Apart from this, the main practical drawback of FDvar is that it cannot deal with trait values of 0 (except by adding an arbitrary amount), since the log of 0 is undefined. Leps *et al.* (2006) proposed a modified version of FDvar, which does not log transform trait values or use the arc tangent of the abundance-weighted squared deviance in trait values to constrain the index between 0 and 1. This obviously has the drawbacks that index values may be affected by the units in which traits are measured and observed values of the index are not readily interpretable, since it is not bound between 0 and 1. The problem of comparing index values between traits measured in different units can be circumvented by standardizing trait values so they have a mean of 0 and variance of 1, or by expressing trait values relative to the range of observed values.

Laliberte and Legendre (2010) presented a multivariate extension of the Leps *et al.* (2006) modification of FDvar, which they termed functional dispersion. This is essentially the abundance-weighted sum of deviances in multivariate trait space taken across species. The functional dispersion index first finds the abundance-weighted group centroid by calculating the abundance-weighted mean for each trait. It then sums the abundance-weighted deviance from this centroid across species. This index conforms to all the criteria relevant for measures of functional divergence, except that it is not bound between 0 and 1 (though this is only important for ease of interpretation of observed values). It does not conform to the set monotonicity criterion and no mathematical proof

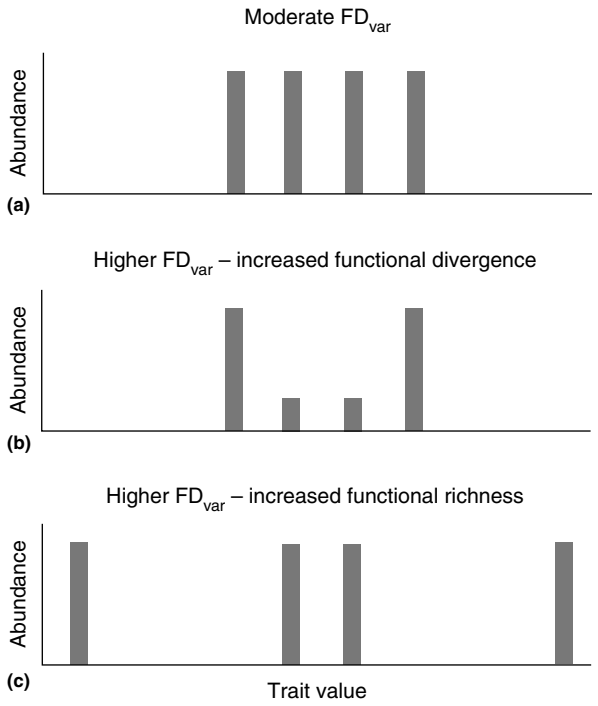


Figure 7 Illustration of the dependence of FDvar on both functional richness and functional divergence. Comparison of panel A with panel B demonstrates that FDvar increases when the most abundant species have extreme trait values (a feature of functional divergence). Comparison of panels A and C shows that FDvar also increases as the range of trait space occupied increases (a feature of functional richness). The same applies to adaptation of FDvar proposed by Leps *et al.* (2006) and to the index of multivariate dispersion presented by Laliberté and Legendre (2010), which is essentially a multivariate extension of the index presented by Leps *et al.*

has been provided to show that it conforms to the concavity criterion. In practice, functional dispersion appears to be very strongly correlated with quadratic entropy.

Rao's quadratic entropy is the sum of abundance-weighted pairwise distances between species in functional trait space. Since this index is sensitive both to the volume of functional trait space occupied and the degree to which abundance is distributed toward the extremities of this occupied trait space, it incorporates both functional richness and functional divergence (**Figure 8**). It does not conform to the set monotonicity criterion but does have concavity if regional abundances are calculated appropriately (Ricotta, 2005). This index also conforms to all the criteria relevant for a measure of functional divergence, except that it is not always bound between 0 and 1 (although using distance metrics such as Gower's distance Rao will be bound between 0 and 1, it will not necessarily use the entire range of values within these bounds). Other than this, quadratic entropy does not have any obvious drawbacks.

Recommendations

CHV and FR_{im} are emerging as the most promising measures of functional richness. Their strength is that they are tightly linked to species distribution in functional trait space, whereas dendrogram measures of functional richness are a distortion

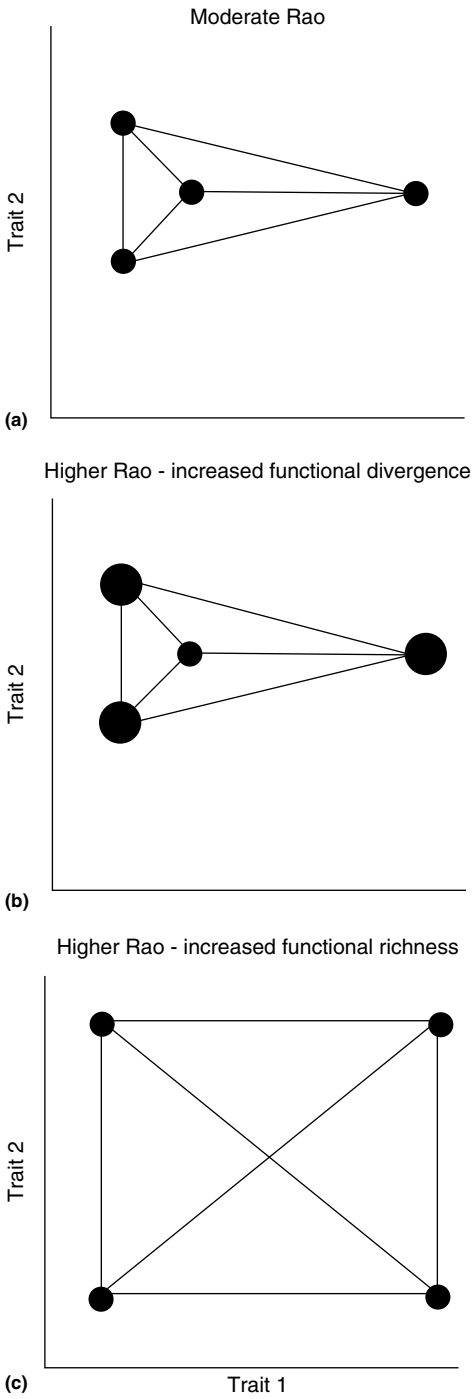


Figure 8 Illustration in two dimensional functional trait space of the dependence of Rao's quadratic entropy (Rao) on both functional richness and functional divergence. Rao is the sum of the abundance-weighted pairwise distances between species in functional trait space (the solid lines connecting the filled circles). The diameter of the filled circles is proportional to species abundance. Comparison of A with B shows that Rao increases when the most abundant species occur at the extremities of occupied functional trait space (a feature of functional divergence). Comparison of A with C shows that Rao also increases as the volume of trait space occupied increases (a feature of functional richness).

of this distribution. FR_{im} has the advantage that it does not require species richness to be greater than the number of functional traits, so ordination techniques are not required to reduce the dimensionality of functional trait space. It also has the advantage of measuring how completely species fill the volume of trait space occupied, and so provides a closer representation of the functional richness concept. Because both indices are relatively new, simulation and field studies are required to determine whether one or the other has greater power to reveal ecological processes. For functional evenness, choices are limited to FRO and its multidimensional analog FEve. Its failure to satisfy all criteria is nontrivial, and its use should be restricted to situations where species identifications can be made with a high degree of certainty. Despite this, FRO has proven useful in revealing ecological patterns (Mason *et al.*, 2008). For functional divergence, FDiv provides a useful multivariate measure, and its ability to reveal ecological processes has recently been demonstrated (Mouillot *et al.*, 2011). FDs has yet to be applied to real ecological questions, so its practical usefulness remains unclear, although it is mathematically sound. Among the composite measures of functional richness and functional divergence, Rao quadratic entropy is the most suitable because it satisfies most of the important criteria for both types of index. In particular, the fact that it conforms to the concavity criterion opens up interesting possibilities for exploring ecological questions.

Application of FD Measures to Answer Ecological Questions

FD measures have a dual role as potential indicators of ecosystem functioning, such as primary productivity, and of the processes that govern the assembly of species into communities, since complementarity between species in resource use and acquisition may both increase rates of ecosystem function and facilitate coexistence between species. In addition to this, indices that conform to the concavity criterion open up the possibility that hierarchical partitioning of FD measures may be used to estimate FD at local (alpha), regional (gamma), and beta scales. Below we provide an example demonstrating the link between FD measures and ecosystem function in experimental grassland communities and an example showing how FD measures can reveal community assembly processes in lacustrine fish communities. Finally, we provide an example of how hierarchical partitioning can reveal changes in FD at multiple scales along a gradient of declining soil phosphorous. Because development of continuous FD measures has been so rapid in recent years, there are only a limited number of examples of their application to answering ecological questions relative to the number of indices available. We anticipate a wide variety of innovative applications of FD measures in the future. Consequently, these examples are not intended to provide a guide for the application of FD measures, but to highlight some of their potential applications.

Example 1: Ecosystem Function in Experimental Grassland Communities

The niche complementarity hypothesis states that differences between species in their functional traits will enhance ecosystem functioning by allowing greater complementarity in resource use between species, which will in turn increase the total amount of resources available to the community. In plant communities, niche complementarity predicts that primary productivity rates should be enhanced by increasing FD. Enhanced primary productivity, by increasing the energy and nutrients available to consumers, is likely to increase consumer diversity, which may increase rates of function (especially litter decomposition) in consumer communities. Mouillot *et al.* (2011) tested these ideas by comparing the relative influence of species richness, evenness in species abundance, functional identity (community trait means weighted by relative abundance), and FD measures on a measure of ecosystem multifunctionality at an experimental grassland near Bayreuth, Germany. This site was part of the pan-European project BIODiversity and Ecological Processes in Terrestrial Herbaceous ecosystems (BIODEPTH), where communities differing in plant diversity were established from seed.

Ecosystem multifunctionality was calculated as the mean performance of communities over four processes (aboveground biomass production, litter and cotton decomposition, and assimilation of nitrogen in live aboveground vegetation) after standardizing each community process (mean of 0 and standard deviation of 1) in order to give them the same weight. The FD measures used were CHV for functional richness, FEve for functional evenness, and FDiv for functional divergence, all of which are described in Villeger *et al.* (2008). Six functional traits were measured in total, containing a mixture of continuous and categorical variables. To allow the inclusion of the categorical variables in FD calculations, PCoA was performed. This resulted in three independent functional trait axes. The distribution of species in functional trait space was thus described by species positions along each of these axes. For each community, FD was calculated using PCoA coordinates instead of raw functional trait measurements with species aboveground biomass as the measure of abundance. Analyses were restricted to communities with four or more species since CHV and FDiv may only be calculated when species richness is greater than the number of dimensions in functional trait space. Community weighted mean (CWM) values were calculated as the abundance-weighted mean coordinate value for each PCoA axis. Species richness was fixed at the number of species sown in the original community, whereas evenness in species aboveground biomass was estimated using the Pielou index.

These eight indices of community structure were used as predictor variables in a structural equation model (SEM) to predict ecosystem multifunctionality. The SEM showed that FDiv had a strong positive effect on multifunctionality, independent of the other community structure variables (**Figure 9**). It also showed that CWM values for PCoA axis 1 had a positive and those for PCoA axis 3 a negative relationship with multifunctionality. Overall, the SEM revealed a synergistic effect of FDiv and functional identity (CWM values) on ecosystem function, indicating that highest rates of function were achieved by communities that had particular CWM values and where the most abundant species were complementary in their functional traits. This example shows how SEMs, in combination with field experimental data, can reveal the impact of FD on ecosystem function. Multiple

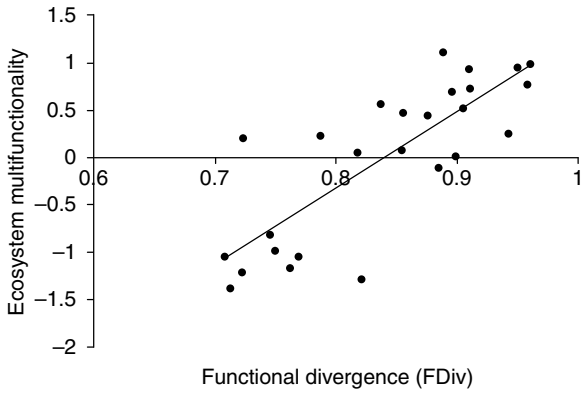


Figure 9 Relationship between ecosystem multifuntionality and functional divergence (as estimated by FDiv) Multifunctionality was estimated as the mean of standardised values taken across four measures of ecosystem function (above-ground biomass production, litter and cotton decomposition and assimilation of nitrogen in live above-ground vegetation). This explains why some of the plots have negative values for multifunctionality.

regression analyses show that FDiv had a strongly significant positive effect on all four of the ecosystem processes examined, suggesting that increased functional divergence in plant communities can enhance not only primary productivity but also rates of decomposition and nutrient cycling within grassland ecosystems.

One complication of this example is that the PCoA axes are not immediately functionally interpretable. Subsequent analyses show that the effect of FDiv was largely due to differentiation along PCoA axes 1 and 3. Axis 1 is largely a contrast between plants with vertical leaves and plants with horizontal or inclined leaves. Axis 3 is a contrast between plants with evergreen and nonevergreen leaves. Thus, the results of Mouillot *et al.* (2011) suggest that ecosystem function is enhanced by complementarity in resource use between species differing in leaf inclination and phenology. Differences in leaf inclination allow complementarity in light capture between species, which can enhance the total amount of light absorbed in photosynthesis for the community as a whole, whereas differences in phenology allow temporal differentiation in resource use between species, which enhances annual above-ground biomass and litter production. This enhancing effect of FDiv on productivity may be responsible for its relationship to decomposition rates, by increasing the flux of nutrients available to decomposer communities. Thus, although the use of ordination to reduce the dimensionality of trait space introduces some complexity in interpreting FD–ecosystem function relationships ecologically, it is still possible to find an ecological basis for the observed patterns.

Example 2: Assembly Processes in Lacustrine Fish Communities

Mason *et al.* (2008) used FD measures to explore mechanisms explaining species richness–energy relationships in French lacustrine fish communities. The study included 116 lakes covering most of France and including the island of Corsica. The lakes were divided among 11 major drainage basins, spanning an altitudinal range from 4 to 2600 m (a.s.l.) and ranging in area from 1.4 to 5299 ha. The species composition dataset consisted of a matrix of species presence or absence within each lake. Sixteen functional traits were measured on up to 10 individuals for each species. They indicate three major aspects of resource acquisition/resource type in fish species – (1) mode of food acquisition, (2) position in the water column or swimming behavior, and (3) preferred food type (Dumay *et al.*, 2004), thus providing information on aspects of the species niche that are relevant for coexistence in the local community. Functional richness was calculated using the FRi of Mason *et al.* (2005, which is the same as CHV for a single trait) functional evenness was estimated using the FRO index of Mouillot *et al.* (2005a), whereas niche overlap was also calculated following the method of Mouillot *et al.* (2005b). FD and niche overlap values were calculated for each trait separately, with the mean value then being taken across traits. Values were expressed relative to those expected when species occurrences were randomized to ensure that the observed relationships were not simply due to chance.

The first test was whether the FD indices had a linear or saturating nonlinear relationship with species richness. Next the effect of mean air temperature, lake surface area, maximum water depth, surface area of the upstream drainage basin, regional species richness, and FD of the regional species pool on species richness, FD, and niche overlap was assessed using multiple regression. The intention was to test whether energy (temperature), habitat diversity (lake area and water depth), and immigration potential (catchment size and regional species richness and FD) had a dominant effect on local FD, niche overlap, and species richness.

The quadratic model provided the greatest amount of explanatory power for the relationship between functional richness and species richness. This indicates that the ratio of observed to expected functional richness (i.e., volume of niche space occupied) either peaked or saturated at moderate levels of species richness (**Figure 10**). The power function had the greatest explanatory power for niche overlap, yielding a relationship where overlap decreased sharply initially with increasing species richness, and then showed only a slight decrease with subsequent increases in species richness (**Figure 10**). There was strong evidence for a linear relationship between functional evenness and species richness with only slight signs of saturation until the very highest levels of species richness (**Figure 10**). These results indicate that although the volume of trait space occupied did not increase at the highest levels of species richness, an increase in overlap was avoided because species became more evenly spaced within the occupied trait space. This in turn suggests that increased niche specialization facilitated species coexistence at the highest levels of species richness.

Mean air temperature was the variable with the greatest explanatory power, for species richness, all measures of FD, and niche overlap in multiple regression (**Figure 11**). The effects of all other explanatory variables were minor in comparison. There was some evidence for a nonlinear relationship between species richness and mean annual temperature, though there was not enough evidence to reject the linear model (**Figure 11**). The linear model received the greatest support for the functional evenness–temperature relationship.

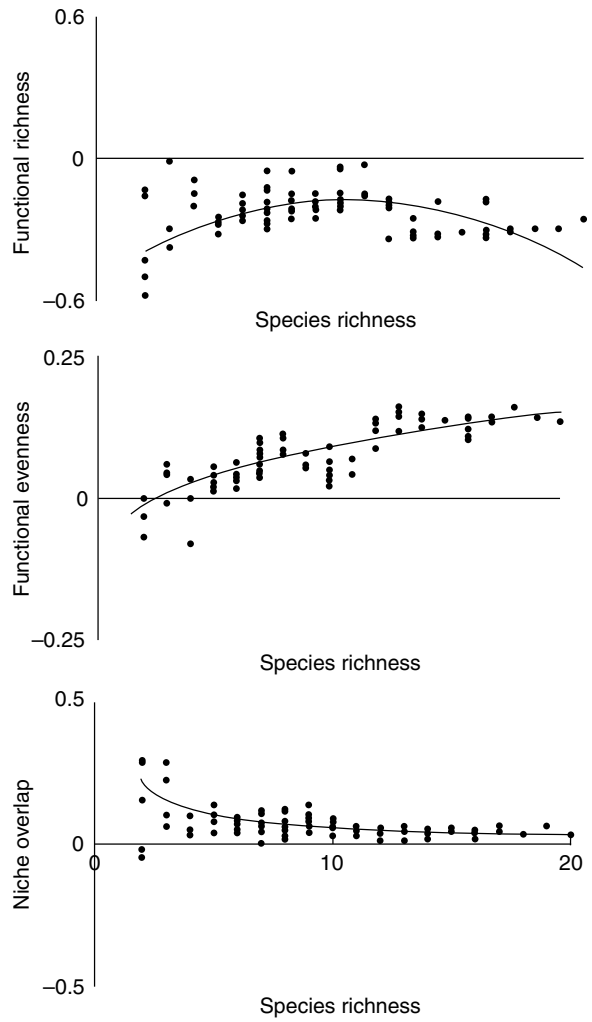


Figure 10 Relationships between functional diversity, niche overlap and species richness in French lacustrine fish communities. Functional richness was measured using the FRi of Mason *et al.* (2005) while functional evenness was estimated using the FRO of Mouillot *et al.* (2005). Functional diversity and niche overlap values are expressed relative to those expected when species co-occurrences are randomised.

These results, combined with the relationships between functional evenness, niche overlap, and species richness, suggest that increased energy inputs may permit species coexistence in the most species-rich communities by allowing species with specialized niches to maintain viable populations.

Low-energy lakes are dominated by generalist carnivores that may feed efficiently on other fish and insects at all positions in the water column (e.g., trout – *Salmo trutta* – and other salmonid species). High-energy lakes include specialist piscivores (e.g., the pikeperch *Sander lucioperca* and Wels catfish, *Silurus glanis*), specialist benthic feeders (e.g., the bream, *Abramis brama*), and species with strong herbivorous tendencies (e.g., the common carp). Primary productivity and zooplankton density increase with temperature in lake ecosystems. However, it is uncertain whether increasing temperature causes an increase in the availability of resources used by specialists. Evidence for the "niche specialization" hypothesis would be provided by finding that the density of resources used by specialists (e.g., fish prey species, benthic invertebrates, or detritus and macrophytes) increases with increasing temperature. This evidence would be strengthened by a positive relationship between the probability of occurrence or population density of niche breadth specialists within lakes and the density of the resources, which they are specialized to exploit. Full confirmation would require demonstration that increased density of these resources was associated with a decrease in the probability of local extinction (or periods of low population density) for specialist species. Thus the example of Mason *et al.* (2008) demonstrates that although FD measures may reveal patterns that are suggestive of community assembly processes, additional work is generally required to confirm the link between observed patterns and inferred processes.

Example 3: Hierarchical Partitioning of FD along a soil Phosphorous Gradient

Mason *et al.* (2011) examined changes in FD at local, regional, and beta scales along a soil phosphorous gradient associated with a soil chronosequence generated by a retreating glacier in a super-humid region of New Zealand. The intention was to test whether the relative influence two contrasting assembly processes (environmental filtering and niche complementarity, Mason *et al.*, 2011) varied as soil P declined at any of these scales. Theoretically, the influence of environmental filtering on community assembly should increase as nutrient stress intensifies, so that FD should decrease with declining soil P. Four traits (leaf N and P, leaf thickness, and density) that indicate species resource use strategy (i.e., fast and leaky resource acquirers versus slow and tight resource retainers) were measured for the woody species occurring at each of eight sites along the chronosequence. Diameter measurements were used to estimate species basal area within three circular plots of 10 m radius at each site. Local (alpha) FD was calculated as the mean of Rao's quadratic entropy (taken across plots within each site), using basal area as the abundance weighting. Regional diversity was calculated at the site scale, with Rao values calculated using species mean proportional abundance taken across the three plots as an abundance weight. Beta FD was estimated by subtracting local Rao from the corresponding regional value. All calculations followed the methods of De Bello *et al.* (2010a). In addition to examining patterns for raw Rao values, Mason *et al.* (2011) compared observed local Rao values with those expected when species abundances were randomly allocated within communities. This method has proven very powerful in revealing evidence for community assembly processes in plant, fish, and plankton communities. Here, higher Rao than expected provides evidence for niche complementarity, whereas lower Rao than expected provides evidence for environmental filtering.

There was no evidence of a decline in raw local Rao values when weighting by basal area (**Figure 12**). However, there was a strongly significant decline in local Rao values expressed relative to random expectation (**Figure 12**). This result suggests that constraints on the functional traits required to

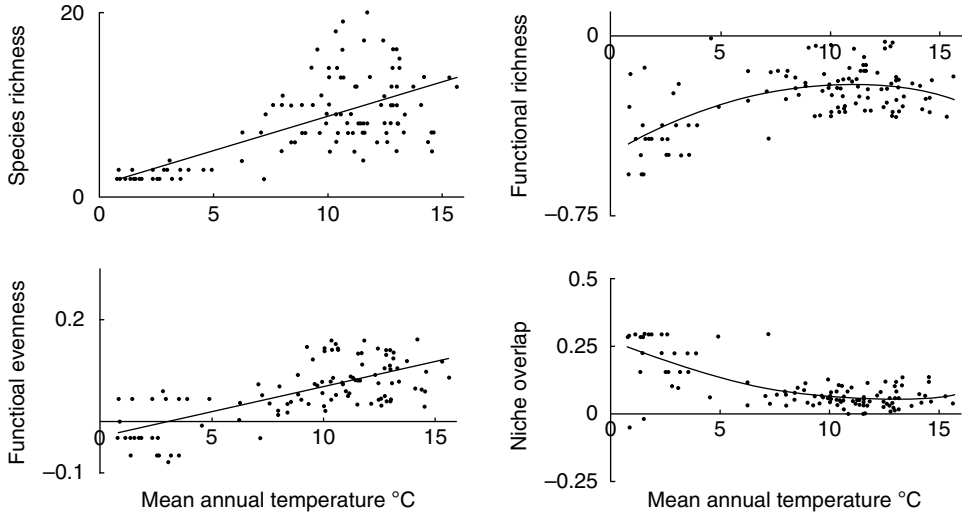


Figure 11 Relationships between functional diversity, niche overlap and mean annual temperature in French lacustrine fish communities.

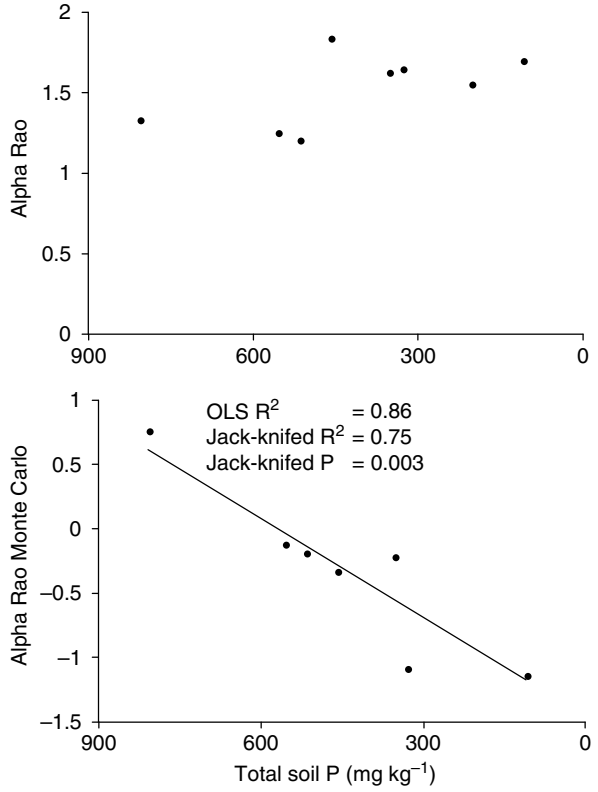


Figure 12 Observed Alpha Rao quadratic entropy and Alpha Rao relative to random expectation (Alpha Rao Monte Carlo) for each site along the chronosequence versus total soil P. Each point represents a single site along the chronosequence. OLS R2 is R-squared from ordinary least squares regression. Other statistics are from Jack-knifed linear regression.

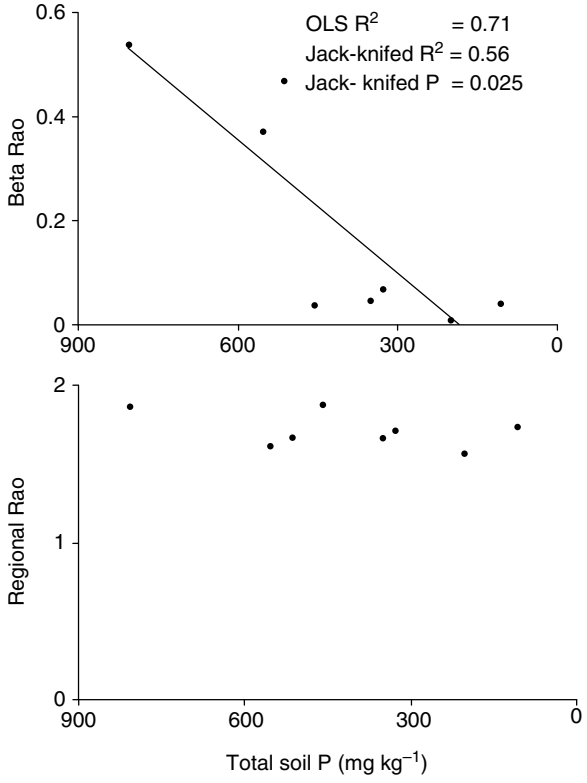


Figure 13 Regional (site) and beta (between plots but within sites) Rao quadratic entropy for each site along the chronosequence versus total soil P. Each point represents a single site along the chronosequence. OLS R^2 is R-squared from ordinary least squares regression. Other statistics are from Jack-Knifed linear regression.

achieve high abundance became much more severe as soil P declined. There was no relationship between regional (site) Rao and total soil P, but Beta Rao was significantly negatively correlated with total soil P (Figure 13). This indicates that communities became functionally more homogeneous as nutrient stress increased.

The fact that a decline in local FD with declining soil P was only observed when raw Rao values were compared with

random expectation indicates that the ability of FD indices to reveal evidence for ecological processes may be highly dependent on how they are expressed. The decline in beta diversity suggests that the functional composition of local communities became much more deterministic with declining fertility, indicating that the limitation imposed by abiotic stress on local FD also increased homogeneity in functional composition across local communities. This example shows that the comparison of raw FD values with random expectation and the hierarchical partitioning of FD provide two potentially powerful methods for the application of FD indices in revealing community assembly processes.

See also: Biodiversity, Definition of. Birds, Biodiversity of. Competition, Interspecific. Complementarity. Defining, Measuring, and Partitioning Species Diversity. Diversity, Community/Regional Level. Diversity, Taxonomic versus Functional. Ecosystem, Concept of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Function Measurement, Terrestrial Communities. Ecosystem Function, Principles of. Ecosystem Services. Fishes, Biodiversity of. Functional Diversity. Global Species Richness. Guilds. Habitat and Niche, Concept of. Limits to Biodiversity (Species Packing). Measurement and Analysis of Biodiversity. Plant Biodiversity, Overview. Resource Partitioning. Species Assemblages, Macroecology, and Global Change. Species Coexistence. Species Diversity, Overview. Stability, Concept of

References

Botta-Dukat Z (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science* 16: 533–540.

Bulla L (1994) An index of evenness and its associated diversity measure. *Oikos* 70: 167–171.

Cianciaruso MV, Batalha MA, Gaston KJ, and Petchey OL (2009) Including intraspecific variability in functional diversity. *Ecology* 90: 81–89.

Cornwell WK, Schwilk DW, and Ackerly DD (2006) A trait-based test for habitat filtering: Convex hull volume. *Ecology* 87: 1465–1471.

De Bello F, Lavergne S, Meynard C, Leps J, and Thuiller W (2010a) The spatial partitioning of diversity: Showing Theseus a way out of the labyrinth. *Journal of Vegetation Science* 21: 992–1000.

De Bello F, Lavorel S, Albert CH, *et al.* (2010b) Quantifying the relevance of intraspecific trait variability for functional diversity. *Methods in Ecology and Evolution* 2: 163–174.

Dumay O, Tari PS, Tomasini JA, and Mouillot D (2004) Functional groups of lagoon fish species in Languedoc Roussillon, southern France. *Journal of Fish Biology* 64: 1–14.

Gaston KJ (1996) *Biodiversity: A Biology of Numbers and Difference*. Oxford: Blackwell Science.

Laliberte E and Legendre P (2010) A distance-based framework for measuring functional diversity from multiple traits. *Ecology* 91: 299–305.

Lavorel S, Mcintyre S, Landsberg J, and Forbes TDA (1997) Plant functional classifications: From general groups to specific groups based on response to disturbance. *Trends in Ecology & Evolution* 12: 474–478.

Leps J, De Bello F, Lavorel S, and Berman S (2006) Quantifying and interpreting functional diversity of natural communities: Practical considerations matter. *Preslia* 78: 481–501.

Mason NWH, De Bello F, Dolezal J, and Leps J (2011) Niche overlap reveals the effects of competition, disturbance and contrasting assembly processes in experimental grassland communities. *Journal of Ecology* 99: 788–796.

Mason NWH, Irz P, Lanoiselee C, Mouillot D, and Argillier C (2008) Evidence that niche specialization explains species–energy relationships in lake fish communities. *Journal of Animal Ecology* 77: 285–296.

Mason NWH, Macgillivray K, Steel JB, and Wilson JB (2003) An index of functional diversity. *Journal of Vegetation Science* 14: 571–578.

Mason NWH, Mouillot D, Lee WG, and Wilson JB (2005) Functional richness, functional evenness and functional divergence: The primary components of functional diversity. *Oikos* 111: 112–118.

Mouchet M, Guilhaumon F, Villeger S, Mason NWH, Tomasini JA, and Mouillot D (2008) Towards a consensus for calculating dendrogram-based functional diversity indices. *Oikos* 117: 794–800.

Mouchet MA, Villeger S, Mason NWH, and Mouillot D (2010) Functional diversity measures: An overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology* 24: 867–876.

Mouillot D, Mason WHN, Dumay O, and Wilson JB (2005a) Functional regularity: A neglected aspect of functional diversity. *Oecologia* 142: 353–359.

Mouillot D, Stubbs W, Faure M, *et al.* (2005b) Niche overlap estimates based on quantitative functional traits: A new family of non-parametric indices. *Oecologia* 145: 345–353.

Mouillot D, Villeger S, Scherer-Lorenzen M, and Mason NWH (2011) Functional structure of biological communities predicts ecosystem multifunctionality. *PLoS ONE* 6: e17476.

Pavoine S and Bonsall MB (2010) Measuring biodiversity to explain community assembly: A unified approach. *Biological Reviews* 4: 792–812.

Petchey OL and Gaston KJ (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters* 5: 402–411.

Poos MS, Walker SC, and Jackson DA (2009) Functional-diversity indices can be driven by methodological choices and species richness. *Ecology* 90: 341–347.

Ricotta C (2005) Additive partitioning of Rao's quadratic diversity: A hierarchical approach. *Ecological Modelling* 183: 365–371.

Schleuter D, Daufresne M, Massol F, and Argillier C (2010) A user's guide to functional diversity indices. *Ecological Monographs* 80: 469–484.

Villeger S, Mason NWH, and Mouillot D (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* 89: 2290–2301.

Violle C, Navas ML, Vile D, *et al.* (2007) Let the concept of trait be functional! *Oikos* 116: 882–892.

Functional Groups

Robert S Steneck, University of Maine, Orono, ME, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 121–139, © 2001, Elsevier Inc.

Glossary

Convergent evolution Distant or unrelated organisms evolve the same anatomical, morphological body plan characteristics, or ecological function.

Ecosystem The combined plant and animal communities plus their physical environment.

Ecosystem function The energy flow, productivity, element cycling, and resilience of ecosystem structure; synonymous with "behavior of ecological systems" and "ecological processes."

Ecosystem structure The organisms, their communities, biodiversity, and habitats that comprise an ecosystem.

Functional groups Polyphyletic suites of species that share ecological characteristics and play equivalent roles in natural communities and ecosystems. Commonly, organisms with convergent anatomical, morphological, physiological, behavioral, biochemical, or trophic characteristics are grouped together.

Guilds Organisms that use similar resources in similar ways. Depending on the application, guilds can be synonymous with functional groups.

Functional Groupings and their Utility

More than 40 years ago, ecologist and evolutionary biologist G. E. Hutchinson wondered why there were so many kinds of species. This question engaged many of the most powerful minds in ecology and the natural sciences. It launched decades of research, brought us the concept of ecological niche and considerable ecological and evolutionary theory, and is the focus of many articles in this encyclopedia. Much less has been written about an equally compelling question: Why are there so few kinds of morphologies or physiologies, and what might the ecological consequences of such limits be? It also raises new questions, such as the following: Are there limits to the possible number of variations on any given body plan? Are there similar limits to the functional attributes of species and suites of species? Can functional attributes of species be abstracted such that their ecological roles can be inferred? Can several species fulfill the same functional roles? In other words, are species replaceable such that there are ecologically equivalent or redundant species? The answers to these questions have profound implications to how we view the natural world and possibly how we can conserve it. For example, does the extinction or extirpation of any given species degrade natural ecosystems in tangible ways?

Biodiversity includes the diversity of organic life on Earth and how it functions. How organisms "make their living," how they interact and contribute to the ecosystem, relates to the functional properties of biodiversity. Functional groupings are polyphyletic suites of species that share ecological characteristics. Commonly, organisms with convergent anatomical, morphological, physiological, behavioral, biochemical, or trophic characteristics are grouped together. Functional properties may correspond to different ecological characteristics depending on the question being asked. They can include the habitats in which they live, the biomass or structure they impart to a community, the rates at which they grow (their productivity), the rates at which they lose biomass (i.e., disturbance), the organisms they consume, the organisms that consume them, and their numerical or mass-specific importance to their communities or ecosystems. A functional group analysis can be applied more broadly in space for making biogeographic comparisons and in time for reconstructing paleocommunities than is possible at the level of species or of related taxa.

I explore functional groupings first from the perspective of the evolution of functional similarities, and then I discuss what functional groupings might mean to the structure and function of communities and ecosystems. I draw heavily, but not exclusively, on examples from the marine realm, in which there is the greatest phyletic diversity, longest evolutionary history, and clearest fossil record of how organisms and assemblages have evolved. Other extensive works on the topic of functional groups, such as that by Schluze and Mooney (1994a), are exclusively terrestrial in their approach.

Functional distinctions can be based on any characteristic shared among organisms, such as morphology, physiology, or behavior. These are not independent features of organisms. An organism's morphology can influence its behavior and its physiology. Functional groupings would be unnecessary if organisms' taxonomic relationships corresponded well with their functional roles. However, functional characteristics are often decoupled from even closely related species, and thus the relationship between taxonomic relatedness and functional groupings warrants consideration.

Decoupling Relatedness from Functional Similarities

Species that are closely related evolutionarily are not necessarily morphologically or functionally similar. Conversely, just

because organisms are distantly related does not mean they are morphologically or ecologically different. There are examples of unrelated, or distantly related, organisms having similar morphologies, anatomies, behaviors, physiologies, and functional roles in natural communities. Such convergent functional properties are well-known and operate at many levels. Obviously, bats are flying mammals that exploit some of the same habitats and foods as some birds and moths. Squid are swimming mollusks that exploit some of the same habitats and foods as fish. Although bats are not perfect substitutes for birds, nor are squid identical to fish, their phyletic relationships tell us little about the ecological roles or functional characteristics they possess.

Although most closely related organisms share traits with the clade from which they evolved, and thus will tend to be functionally similar, this is not always the case. For example, hermit crabs are relatively small, anomuran decapods with an exposed, asymmetric abdomen that requires protection from unoccupied snail shells. They differ morphologically from true crabs, the brachyurans that have abdomens protected by a shield-like cephalothorax or carapace that also covers their walking legs (**Figure 1**). True crabs can attain large size. However, hermit crabs are very closely related to the morphologically distinct anomuran king crab (**Figure 1**). King crabs convergently evolved many of the characteristics of true brachyuran crabs; they are large and their abdomen is protected by a broad shield-like cephalothorax. Therefore, phyletically related king and hermit crabs are morphologically different, whereas phyletically distant king and true crabs are morphologically similar.

Morphological similarities can translate to functional similarities. In shallow coastal habitats of the Gulf of Maine, there are four species of large, clawed, decapods that live together and feed on the edible blue mussel, *Mytilus edulis* (**Figure 2**). The two most closely related species are *Cancer irroratus* and *Cancer borealis*. Both are brachyuran crabs in the family Cancridae. The third brachyuran crab, *Carcinus maenas*, is in the Portunidae family. However, the two distantly related crabs, *Carcinus maenas* and *C. irroratus*, share functional characteristics of small body and chelae ("claws") size while being behaviorally flexible, quick, and dexterous. They characteristically chip the perimeter of mussel shells at many locations to gain access to the food. Conversely, *C. borealis* and the lobster *Homarus americanus* share functional characteristics of large body and chelae size while being behaviorally constrained, slow, and strong (**Figure 2**). These two decapods characteristically crush mussel shells in the middle of the shell by applying brute force to the center of the shell. Therefore, the two most closely related crabs are functionally dissimilar and the two sets of most distantly related crabs (*Carcinus maenas* and *C. irroratus*) and other decapod groups (*C. borealis* and *H. americanus*) are both functionally similar.

Even within species, large functional differences exist. Amphibians such as frogs begin life as tadpoles with diets and habitats different from those of adults. Ferns and algae have ecologically distinct haploid and diploid phases. The heteromorphic range in some marine algae is striking. For example, the brown (Phaeophyta) algal kelp *Macrocystis pyrifera* can exceed 40 m in length and forms forests in its sporophyte ($2n$) phase but is a microscopic filament in its gametophyte ($1n$) phase. Some red algae (Rhodophyta) alternate between encrusting and erect phases; some green algae (Chlorophyta) alternate between endolithic microscopic and filamentous phases. Some habitats only support one phase. Each of these different ontogenetic and ploidy phases has strikingly different ecological properties, and differences within species can be greater than some differences among species. Therefore, knowing the species does not always provide insight into how the organism functions ecologically.

There are distinct limits to variety in morphology or important ecological function that are independent of biodiversity. Some characteristics, such as body size, are biomechanically limited; therefore, depending on the biomaterials available to an organism (i.e., cellulose, chitin, calcium carbonate, or bone), their morphology and the environment in which they live may limit their ecological

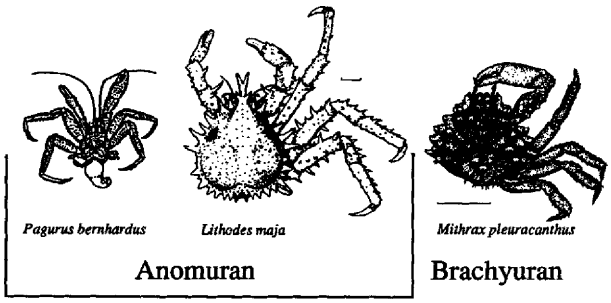


Figure 1 Phylogenetically related anomuran hermit crab (*Pagurus*) and king crab (*Lithodes*) are morphologically dissimilar. The distantly related brachyuran crab (*Mithrax*) is morphologically similar to the king crab (Reproduced from Cunningham CW, Blackstone NW, and Buss LW (1992) Evolution of king crabs from hermit crab ancestors. *Nature* 355: 539–542.).

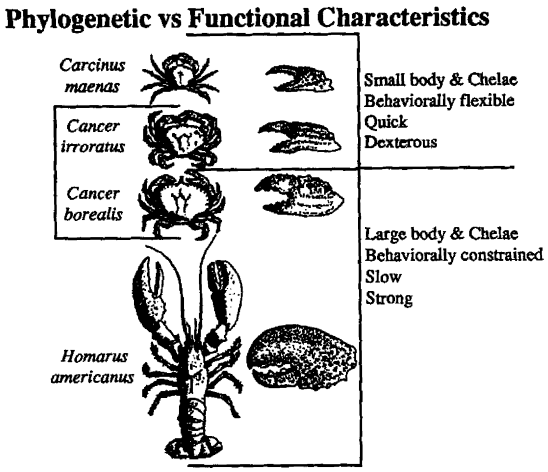


Figure 2 Phylogenetically related *Cancer* species are functionally different. The phylogenetically distant *Cancer borealis* and *Homarus americanus* are functionally similar (Reproduced from Moody KE and Steneck RS (1993) Mechanisms of predation among large decapods crustaceans of the Gulf of Maine coast: Functional vs phylogenetic patterns. *J. Exper. Marine. Biol. Ecol.* 168: 111–124.).

functions. These relationships are well described in works such as Thompson's (1966) *On Growth and Form*, Peter's (1983) *The Ecological Implications of Body Size*, or Niklas's (1994) *Plant Allometry*. Arguably, size and shape are prime determinants of many functional groupings of both plants and animals. Predators generally scale with, or are larger than, their prey. Thus, organisms constrained to small size are likely to be more vulnerable to predators and rarely if ever are "apex predators" in an ecosystem or community. Insects are a good example. They also illustrate the disconnect between species diversity and morphological variety. Insects account for most of Earth's biodiversity, but they are all relatively small. In contrast, most vertebrates are much larger (especially mammals and dinosaurs) and their biomass distribution does not overlap that of insects. Despite their long evolutionary history, there are no cow-sized insects, nor are there insect-sized cows. Body size constraints among insects occur because, among other reasons, their respiration is relatively inefficient and limited by their open circulatory system and trachea. Other factors, such as the composition of their exoskeleton, further constrain morphological evolution and the habitats into which they can and have radiated. Although insects have radiated impressively into terrestrial systems, they are virtually nonexistent in the marine realm. Thus, three-fourths of all species are insects, but three-fourths of the habitable space is devoid of them. There are morphological and functional limits to diversity.

Functional Changes in Assemblages over Evolutionary Time

Ultimately, the environment controls the distribution, abundance, and body plans that live in it. Evolution at several levels generates diversity of form and function, but the environment filters it. Constraints to diversification are many and relate to the environment as it changes and how organisms deal with it.

The history of life on Earth is punctuated with significant functionally novel additions to biotic assemblages. An excellent example of a functional group effect on global-scale ecosystem function is the cyanobacteria that began oxygenating the atmosphere more than 3 billion years ago. Biologically available oxygen began increasing approximately 1.4 billion years ago and is coincident with (and possibly necessary for) the evolution of the first metazoans. Still higher levels of oxygen may have been necessary for the Cambrian "explosion" of small shelly fauna. The sudden polyphyletic and global appearance of hard skeletal material (chitin and calcium carbonate) allowed for the evolution of large body sizes, efficient mobility, teeth, and defenses against predators. The structure and functioning of life on Earth changed at this time. Organisms functionally similar to those alive today first appeared. Since then, during the past 600 million years, there have been several biotic revolutions.

Decoupling phyletic evolution from functional group evolution is important. All but one of the major phyla may have evolved at or near the Precambrian–Cambrian boundary in as few as 25 million years. Alternatively, some argue that phyla evolution may have occurred much earlier (1 billion years ago) as minute and embryonic or larval-like organisms and thus not recognizable by today's criteria. If this latter theory is correct, it would be a clear case of phyletic evolution being disconnected from the functional/morphological (phenetic) evolution that occurred 400 million years later. Despite the Cambrian (or earlier) origin of most major phyla, functional evolution and the filling of unexploited habitats ("ecospace" *sensu* Bambach, 1985) took considerably more time. Throughout the Paleozoic, morphological evolution and the filling of unoccupied habitats occurred episodically and polyphyletically, beginning with small surface deposit feeders and low-profile suspension feeders, progressing to greater depths into sediment and new epifaunal heights. The progressive utilization of three-dimensional space called "tiering" is thought to have contributed substantially to the global diversification in the early Paleozoic (**Figure 3**). The diversifying continued into the pelagic realm and into the sediment habitats (infaunalization) and recently included the radiation of large-bodied predators.

This pattern of diversification is composed of three relatively distinct faunas, each characterized by its own characteristic dominant taxa (**Figure 3A**), initial rates of diversification, and "equilibrium" diversity. Each fauna was functionally different. Surface deposit feeders such as trilobites and monoplacophorans dominated the Cambrian fauna, whereas the Paleozoic fauna radiated into and relied heavily on the pelagic realm. For example, the taller suspension-feeding crinoids of the Paleozoic fauna joined near-benthos suspension-feeding brachiopods, anthozoans, and stenolomate bryozoans and pelagic groups such as cephalopods and graptolites as dominants of this fauna. There were no known macroalgal grazers in the Paleozoic, and large stands of macroalgae were probably common. Sand habitats were virtually devoid of infauna. In contrast, the modern fauna included the rise to dominance of large invertebrate and vertebrate predators and herbivores. These included predatory neogastropod snails, malacostracan crab and lobsters, and at least four groups of swimming vertebrate predators among the bony fishes (Osteichthyes), dinosaurs (Reptilia), sharks (Chondrichthyes), and cetaceans (Mammalia, i.e., dolphins and whales). Most modern groups of fish (e.g., most modern reef fish) evolved at this time. Explanations for why large-scale changes occurred at the family level are complex. Organisms such as small deposit feeders may have suffered higher extinction rates as more mobile and more pelagic organisms with larger gene pools replaced them. A global crisis in the water column contributed to the most severe extinction event at the end of the Permian. Therefore, suspension feeders, pelagic organisms, and organisms with pelagic larvae suffered the highest extinction rates. This extinction caused a sharp decline in several of the dominant families of the Paleozoic fauna and marked the transition to the dominance of the modern fauna (**Figure 3B**). Prior to the dominance of the modern flora there were virtually no shell-crushing carnivores. Several groups of shell crushers, such as true brachyuran crabs, evolved in the Cretaceous and became important predators. Deep-grazing herbivores also evolved among the modern fauna. They created a uniquely intense level of grazing pressure which may have denuded many shallow marine habitats of macroalgae.

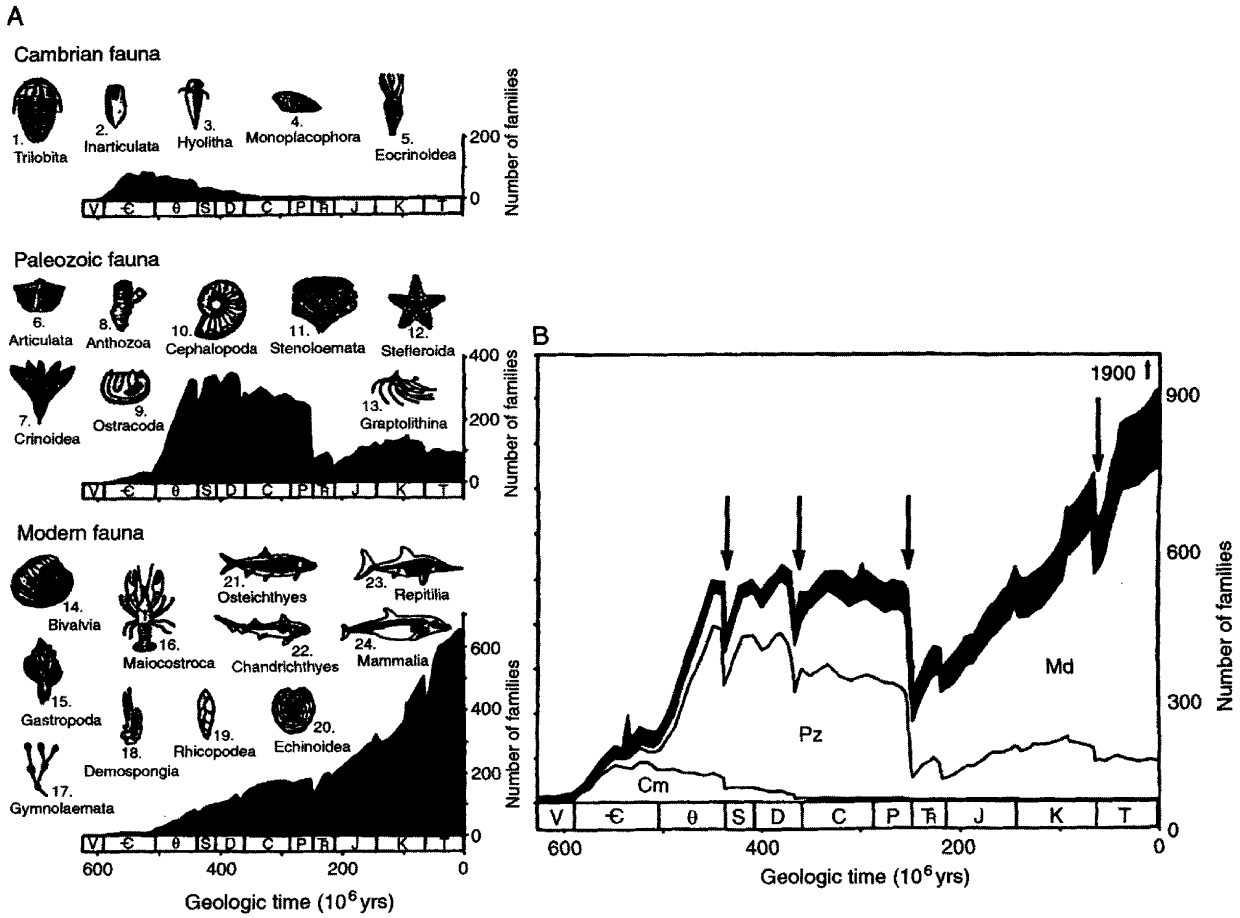


Figure 3 Faunal changes during the past 600 million years. (A) Three faunas with representative dominant animals. (B) Cumulative pattern of diversification of the three faunas. Arrows identify mass extinction events (Sepkoski, 1984). Each fauna represents dominance by a functionally distinct group. The Cambrian fauna were surface deposit feeders, the Paleozoic fauna were dominated by suspension feeders, and the modern fauna were dominated by large predators and protected prey. Note that the Paleozoic fauna was particularly susceptible to mass extinctions resulting from disruptions of the water column.

The Sudden Origin of Functionally Distinct Organisms and Then Functional Stasis

Functional evolution in the marine realm is striking. The transitions from surface-dwelling deposit feeders to a range of suspension feeders, to organisms using vertical space, and to big predators and herbivores represent significant functional changes over evolutionary time. There were also terrestrial revolutions in organism function. Angiosperms displaced functionally similar conifers and evolved functionally unique weeds. Pollinating insects and angiosperms coevolved and diversified. Obviously, the history of life is punctuated with the sudden appearance of functionally unique organisms that often change the biotic world.

Often, functionally unique groups change little over evolutionary time. Among marine invertebrate herbivores, there are only a few capable of biting into limestone or calcium carbonate. These herbivores leave a grazing trace fossil that allows for functional interpretations to be made over evolutionary time. Three groups known to excavate calcium carbonate are regular echinoids (sea urchins), true limpets (Patellogastropoda), and chitons (Polyplacophora). Trace fossils from the Cretaceous (approximately 70 million years ago) reveal that each of these groups grazes distinctively with the same graze marks and to about the same depths into calcium carbonate as their modern counterparts (**Figure 4**). Although considerable variation exists within groups of herbivores, the differences among groups are much greater. Echinoid sea urchins have a characteristic intensity of grazing that is much deeper than that of chitons and limpets which are very distantly related but have similar shallow bites into calcium carbonate. In all cases, the first trace fossils known show these invertebrates biting the surface of calcium carbonate with the pattern and the depth into the substrate as they produce today. There appears to be little or no functional change in how they graze after the group first evolved. There are other examples. Predatory gastropods that drill their shelled prey have not changed their mode of attack since they first evolved in the Triassic more than 200 million years ago. Many limb and running characteristics between terrestrial vertebrate predator–prey interactions have also been shown to be remarkably stable over evolutionary time. This suggests that among these consumers and within established clades, there is functional stasis over evolutionary time.

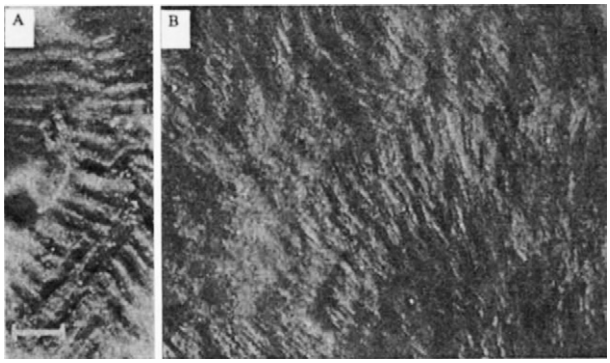


Figure 4 Comparison of modern and ancient limpet graze marks on calcium carbonate substrates. (A) Modern limpet, *Tectura testudinalis*, grazing calcareous algae (scale bar = 0.5 mm) (Reproduced from Steneck RS (1982) A limpet coralline alga association: Adaptations and defenses between a selective herbivore and its prey. *Ecology* 63: 507–522.). (B) Cretaceous (more than 65 million years ago) limpet, *Radulichnus* sp. (scale is the same as in A) (Reproduced from Crimes TP and Harper JC (1977) *Trace Fossils 2, Geological Journal*, Special Issue No. 9. Liverpool, UK: Seel House Press.).

Functional Convergences

Convergent evolution results in distantly related organisms converging on the same body plan or ecological function. Charles Darwin noted in his *Origin of Species* (1859) that

> for animals belonging to two most distant lines may have become adapted to similar conditions, and thus have assumed a close external resemblance; but such resemblances will not reveal—will rather tend to conceal—their blood relationship (page 463).

Relatively few convergently evolved anatomical, morphological, or physiological characteristics are the cornerstones of functional groupings.

Groupings among Mobile Organisms

Convergent evolution is well-known and documented in the terrestrial realm. Marsupial and placental mammals have converged to similar morphologies and ecological function (**Figure 5**). The limited variations on the mammalian body plan are evident in the wolf and catlike carnivores, the arboreal gliders, fossorial herbivores, anteaters, and subterranean insectivores that evolved independently in Australia for the marsupials and on the other continents for the placentals.

Convergent functions can be found among dissimilar-looking organisms. Distantly related marine molluskan herbivores, chitons and limpets, provide an excellent example. Chitons (class Polyplacophora) and snails (class Gastropoda) evolved in the Cambrian at the beginning of the molluskan diversification. The true limpets, Patellogastropoda, evolved from the Archaeogastropoda, much later in the Triassic. However, the radula (the teeth) of these two groups are functionally similar (**Figure 6**). Both groups have relatively few teeth that contact the substrate, and those that do are hardened by mineralization of iron or silica compounds (note the black teeth in **Figure 6**). Both groups have strong buccal musculature for applying downward forces. Within the molluskan body plan, only chitons and limpets have such a large foot area-to-mass relationship and an excavating-type radula. These morphological and anatomical characteristics, along with their size and mobility, allow species of these groups to specialize on large and expansive macrophytes such as sea grasses, kelp, and encrusting coralline algae. Although species diversity is much higher in other groups of mollusks such as nonlimpet gastropods, it is the functional characteristics of these two groups that make them capable of consuming and even trophically specializing on the tough or limestone-imbedded cells of kelp or coralline algae, respectively.

Convergent functions are numerous among terrestrial organisms. For example, a diversity of flowering plants and their pollinators possess similar morphological and anatomical characteristics despite significant phyletic separation among the plants and the pollinators. The geometry of the flowers, such as the length and width of the floral tube, as well as the placement of nectaries are common features among the plants. The flying characteristics and mouthparts are convergent among the pollinating insects, butterflies, moths, bats, and birds. It is surprisingly easy to find functional similarities among distantly related organisms.

Groupings among Sessile Organisms

Convergent morphology and anatomy is common among sessile organisms on land and in the sea. I briefly consider terrestrial plant communities that have many of the same convergences described previously for animals. Families of cactus-like plants in the Old World (Euphorbiacea) and New World (Cactaceae) dominate arid environments but are unrelated. Much has been written about the remarkably convergent forest and biomes that develop under similar environmental conditions (Huston, 1994). There appear to be relatively few adaptive solutions to common environmental conditions. Often, the suite of characteristics that improve ecological function under specific conditions define functional groupings.

Functional groupings among sessile organisms were studied by terrestrial ecologists, who grouped plants according to their size and the location of their growth and regenerative structures. Raunkiaer (1934) showed that particular functional groups (called "life-forms") persisted and dominated the vegetation under similar environmental conditions. Using this scheme, Arctic regions were clearly distinguished from arid or tropical regions. Similar floras, defined at the functional group level, dominated under similar conditions in different regions even though there were often no species in common. This early focus on the size and placement of growth and regenerative structures was based on the realization that gradients in environmental stress and productivity relative to disturbances are critical to the success of specific life-forms. Plants with protected meristems and perenniating structures persisted in stressful or highly disturbed habitats. Where environmental limitations were relaxed, plant size and competitive ability were more important to the flora.

Recently, the field has become popular among marine ecologists and paleontologists who study sessile marine organisms. Again, morphological and functional convergences

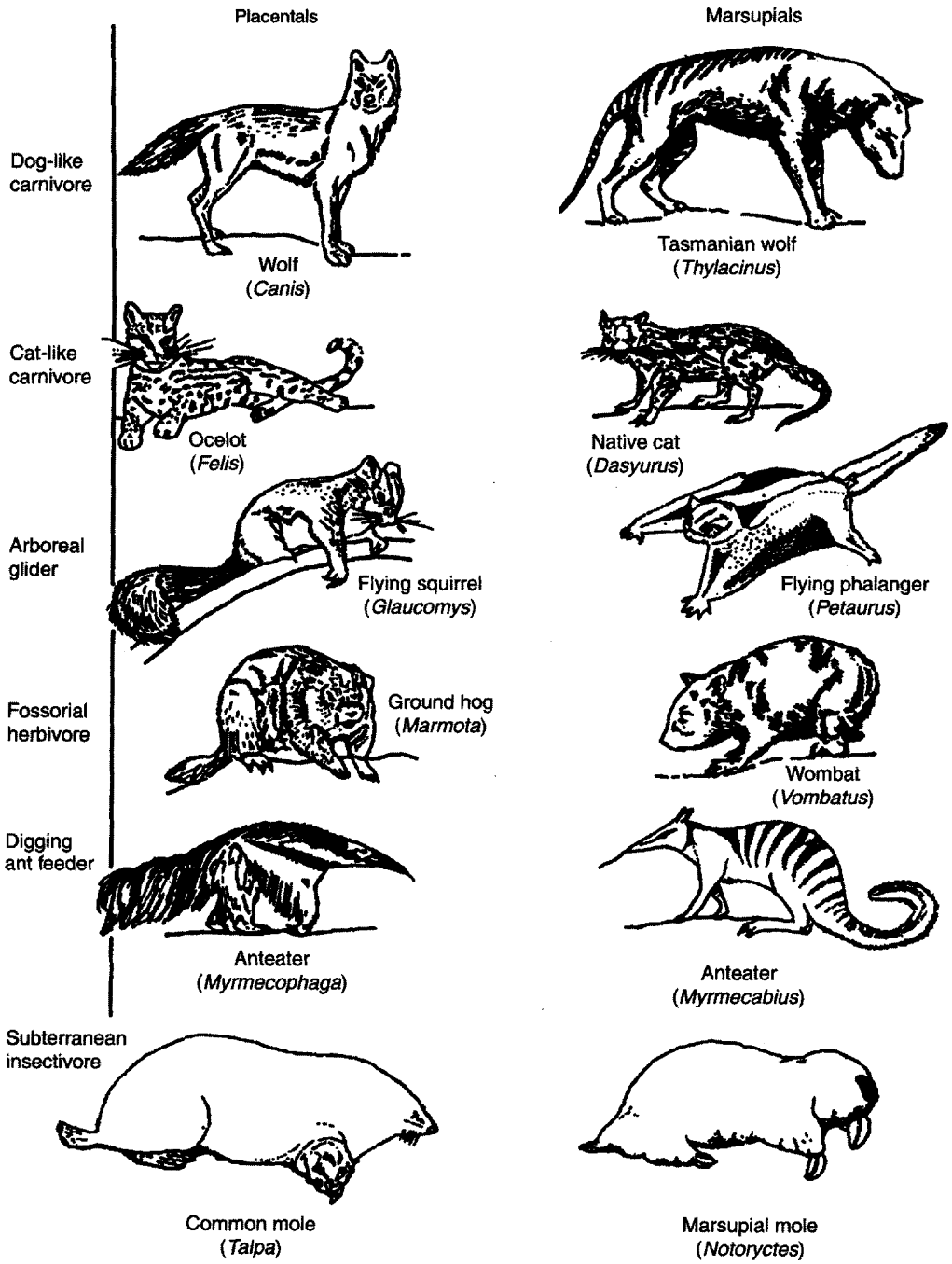


Figure 5 Convergent evolution among placental (left) and marsupial (right) mammals (Reproduced from Begon M, Harper JL, and Townsend CR (1986) *Ecology, Individuals, Populations and Communities*. Sunderland, MA: Sinauer.).

are viewed more as the rule than the exception. Several growth forms have evolved convergently among widely disparate groups, such as marine algae, Cnidaria, and Bryozoa (Figure 7). Most extant and extinct species within those higher taxa can be placed within the groups illustrated in Figure 7. The growth forms represent a morphological progression. The simplest growth form is solitary and can be a single cell as is the case of a diatom or a solitary organism. This group is often composed of small organisms. The next step in a morphological progression of growth forms simply involves a linear series of the single cells or modules, creating simple filaments. These can grow prostrate along the substrate as recumbent algal filaments or as runners or vines among sessile invertebrates. Similar morphologies growing vertically escape the benthic boundary layer. They experience greater water flow, and most of the cells of these groups are in contact with their aqueous environment. The more massive multiserial or multiseriate organisms allow three-dimensional space to be exploited. Several ecological processes, such as productivity and competition, vary predictably as sessile organisms become larger and more

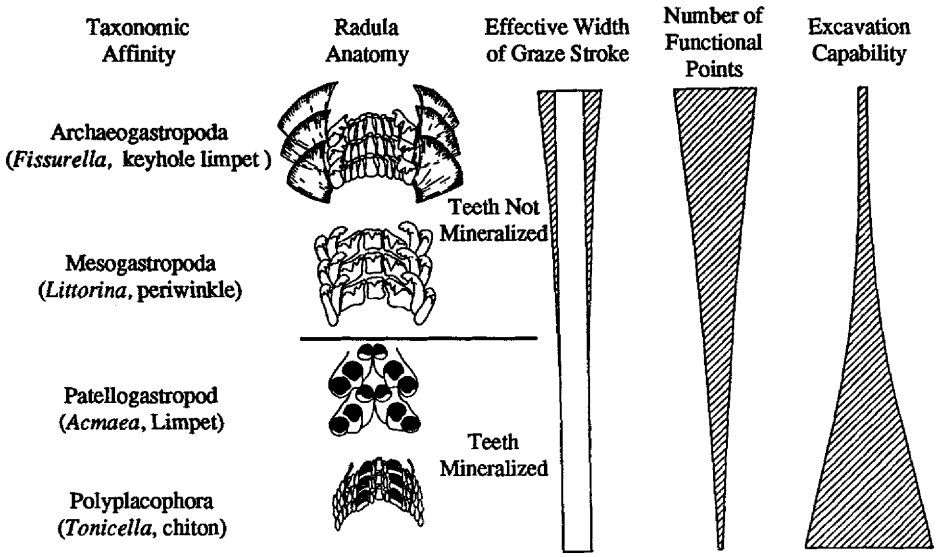


Figure 6 The teeth (i.e., radulae) of herbivorous mollusk functional groups. Black teeth are heavily mineralized, making them harder than the limestone they graze (Reproduced from Steneck RS and Watling LE (1982) Feeding capabilities and limitations of herbivorous molluscs: A Functional group approach. *Mar. Biol.* 68: 299–319.).

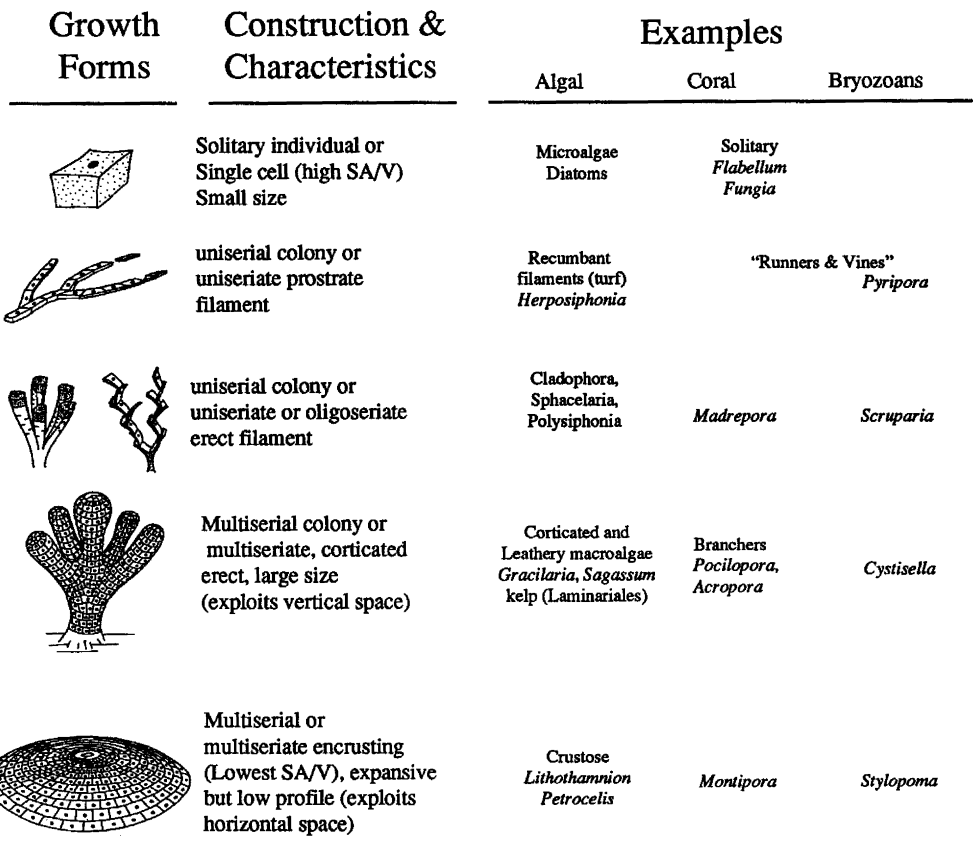


Figure 7 Growth forms of sessile marine organisms. Examples include several divisions of marine algae and Cnidaria and Bryozoa phyla. Other groups having some of the same growth forms include the extinct tabulate and rugose corals and modern sponges (Reproduced from Coates AG and Jackson JBC (1985) Morphological themes in the evolution of clonal and aclonal invertebrates. In: Jackson JBC, Buss LW, and Cook RE (eds.) *Population Biology and Evolution of Clonal Organisms*, pp. 67–106. New Haven, CT: Yale Univ. Press.).

massive. With increasing mass, fewer cells directly contact the marine environment. Fewer still are in contact with the environment among the encrusting forms because half of their surface area is affixed to primary substrate. The growth forms in Figure 7 exhibit progressively increasing size and generally decreasing surface area to volume ratio.

For benthic marine algae of all three major divisions (=phyla), the progression from single-celled microalgae to

filamentous, foliose, larger erect forms and finally to crustose algae (Figure 7) corresponds to a reduction in productivity per unit biomass (Figure 8A; Steneck and Dethier, 1994). In general, microalgae and filamentous forms with a high surface area to volume ratio are capable of the highest mass-specific productivity. This decreases progressively from filamentous forms to the more massive erect forms including kelp and is lowest among the crustose forms. Thallus longevity (Figure 8B) increases along the same continuum of growth forms. Canopy height, however, shows somewhat of a bell-shaped curve, with the tallest algal forms such as kelp and fucoid algae having intermediate levels of productivity and longevity (Figure 8C). These emergent ecological properties were shown for cold-water floras of the Atlantic and Pacific as well as for tropical floras of the Caribbean (Figure 8).

Morphologically and functionally similar organisms have evolved similar ways to overcome similar problems. For example, as organisms become more massive, fewer cells are in direct contact with the marine environment. That is, they have a plumbing problem. Sessile organisms such as plants, algae, and invertebrates evolved physiological integration or a degree of interdependence and cooperation among cells within an organism or modules within a colony. Often this involves sharing cytoplasm or photoassimilates within the organisms and among modules of a colony.

Physiologically integrated colonial organisms have an advantage over physiologically isolated aclonal organisms by sharing cytoplasm among members of the colony. Sessile organisms can respond from a stimulus by growing at a place distant from the stimulus. Wound healing can be facilitated and mortality of the genet reduced by physiological integration. Physiological integration evolved iteratively among sessile organisms on land and in the sea. This allows land plants such as vines to colonize the inhospitable habitats of sand dunes or tree canopies by being physiologically integrated. The most abundant marine algae, the red encrusting coralline algae, are integrated and this is thought to have contributed to their ecological and evolutionary success over unintegrated or poorly integrated algae. Arguably, physiological integration in algae is essential for their morphological diversity. The least physiologically integrated group, Chlorophyta or green algae, have the lowest morphological diversity despite having the same long period of evolutionary history dating back to the Precambrian.

Physiological integration may be a prerequisite for the development of coral reef ecosystems. It evolved independently among cnidarian reef builders: the tabulate, rugose, and finally the scleractinian corals. In each case, the proportion of species with a high degree of physiological integration increased during times of significant reef-building activity. Coral reef ecosystems require high rates of coral growth to keep up or catch up with rising sea level. Reef-building corals today are physiologically plants, with endosymbiotic algae producing carbohydrates necessary for growth. Transport of photosynthetic products throughout the coral colony requires a high

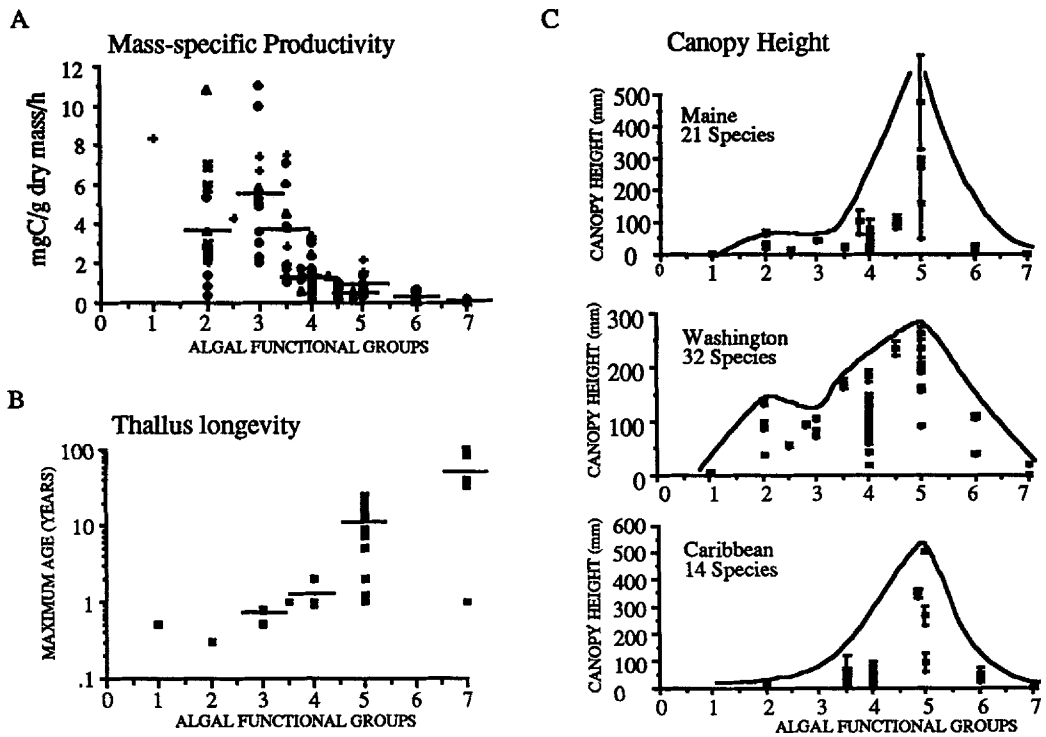


Figure 8 Emergent ecological properties of benthic marine algal functional groups. (A) Mass-specific productivity of algae from southern California, the Caribbean, and Hawaii. (B) Thallus longevity of 27 species from 25 published studies. (C) Canopy heights measured in the field for dominant species in Maine, Washington, and the Caribbean. Lines envelop the maximum canopy heights recorded for each algal functional group and points represent average canopy heights per species. Error bars represent one standard deviation. Reproduced from Steneck RS and Dethier MN (1994) A functional group approach to the structure of algal-dominated communities. *Oikos* 69: 476–498.

degree of physiological integration. Without high rates of coral growth translating to high rates of reef growth, coral reef ecosystems would stand little chance of remaining in the shallow turbulent zone in which they thrive. In fact, it is likely that many coral reef ecosystems that have dominated shallow seas for much of the past 600 million years require the framework-building corals to be physiologically integrated.

Physiological integration must be functionally important to sessile organisms because it evolved independently many times. It is necessary for the similar functional groups found among the corals, algae, and bryozoans (**Figure 7**). Colonial integration, or the degree of interdependence and cooperation among modules within a colony, is important for all these groups because it increases colonial function. It also allows the development of nonfeeding defensive or reproductive functions within the colony or clone, and it allows greater structural integrity and morphological diversity. The important functional advantages of physiological integration may account for why it has persisted and become ubiquitous among so many unrelated sessile, colonial, and clonal organisms.

The Structure and Functioning of Natural Communities and Ecosystems

There exists a broad spectrum of ways to describe patterns. Historically, the emphasis in ecology has been on species. As Naeem (1998) pointed out, species are often viewed as being phyletically "singular" and likely to be "autecologically singular." This was the basis of the important works by Hutchinson and MacArthur (1959), who stressed the uniqueness of species. That no two species can occupy the same ecological niche was the foundation of decades of studies on the structure and function of natural communities and ecosystems. Functional group approaches are different. They stress similarities among unrelated species that share critical organismal features. Convergent evolution in many cases results in phyletically distant organisms sharing ecological properties. The utility of using shared species characteristics for describing communities rests on the idea that in any given environment there are relatively few species attributes important to the structure and functioning of natural communities and ecosystems. These attributes are largely independent of the biodiversity of the ecosystem.

The importance of any species or group can relate to its abundance (measured as population densities, biomass, areal cover, or structural height) or its impact on other species. Whittaker (1965) argued that species' importance should be measured by its productivity or effect on productivity. Obviously, the production of biomass must be approximately matched to its degradation or "disturbance" (*sensu* Grime, 1981). In some systems, species importance depends more on the role of disturbance (e.g., predation) than on productivity. For both disturbance and productivity, properties of the organisms and properties of the environment must be considered simultaneously and independently. Holdridge's (1947) classic "life zone system" for classifying terrestrial vegetation considered only precipitation and temperature as driving productivity and thus creating discrete vegetative communities and ecosystems. Others have shown how disturbances such as storms and fires additionally shape terrestrial plant communities. Although debates continue regarding the primacy of bottom-up or productivity-driven vs top-down or disturbance-driven structure of natural communities, in most systems both components are viewed as important. Therefore, the debate is largely over relative importance.

Productivity and disturbance are the two variables most often identified as environmental structuring agents of natural communities and ecosystems. Huston (1994) reviewed the work of others to develop a model that compared the frequency of disturbance (defined as patch or gap formation) as a function of "potential plant productivity." Relative to these two axes, he showed consistent patterns in the terrestrial realm in species diversity, patch formation, dominance, and competitive ability. General models of community dominants have been advanced by Grime (1981), Southwood (1988), and Steneck and Dethier (1994) (**Figure 9**). Specifically, Grime (1981) suggested that, in environments with a high productivity potential (low stress) and low disturbance potential, large, canopy-forming trees will dominate the system (called a "competitive strategy"). Under conditions that favor high productivity with increasing productivity, perennial herbs and finally weedy annual herbs will come to dominate the system (called a "ruderal strategy"). Under low disturbance and low productivity potential environments, stress-resistant groups such as lichens will dominate the system (**Figure 9A**).

Several studies applied Grime's (1981) model to the marine realm (**Figure 9C**). Strong similarities between the two realms surfaced, including that the highest biomass and largest canopy-forming organisms (trees and kelp) grow under high productivity and low disturbance potentials (**Figure 9C**). Furthermore, low biomass and canopy height dominated by encrusting growth forms are characteristic of stress-tolerant assemblages in both realms. A difference might occur among the organisms that dominate under conditions of chronically high levels of disturbance. Grime (1981) assumed that it is impossible to withstand such conditions and thus weedy, "ruderal strategies" would persist, whereas Steneck and Dethier (1994) argued that an additional group of organisms tolerant of such disturbances (the coralline algae) predictably dominate such systems. Although it is common to find disturbance-resistant plants under conditions of high stress in the terrestrial realm, they can be found under conditions of low stress (high productivity potential) in the marine realm. It is unclear whether this functional difference between the marine and terrestrial realms is real or perceived. Nevertheless, one value to a functional group approach is that this dialog can occur at all.

Ecosystem Diversity, Stability, and Redundancies: Examples from Coral Reefs

Much has been written on this topic, but most of the literature draws from terrestrial studies such as O'Neill *et al.* (1986); Schulze and Mooney (1994a), and Huston (1994). However, rather than reiterating these studies and their myriad examples, I will use coral reef ecosystems as examples in this

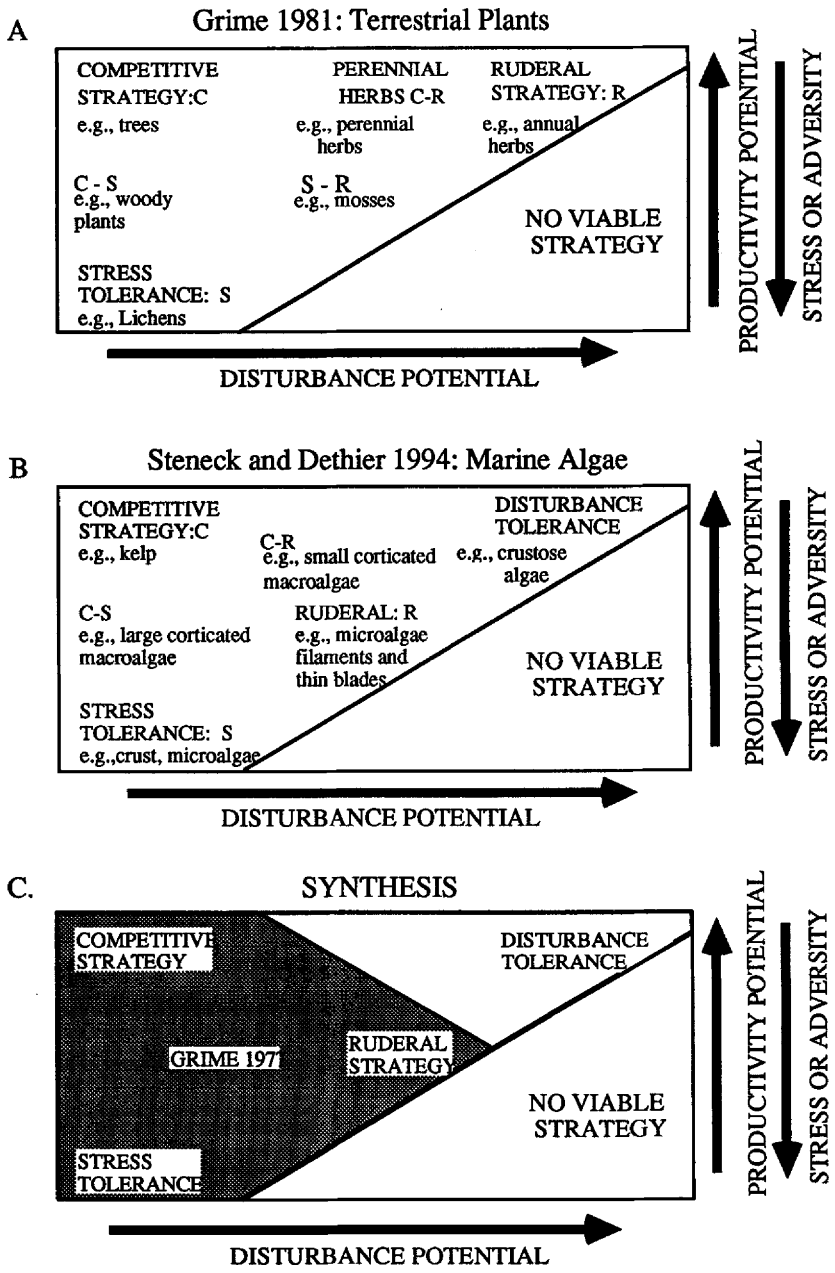


Figure 9 Generalized model of community dominants that compares Grime's (1981) model for terrestrial plants (A) with that of Steneck and Dethier (1994) for marine algae (B). (C) A synthesis of A and B. Grime's primary "strategies" or groupings of "competitive," "ruderal," and "stress tolerance" are denoted by the letters C, R, and S, respectively.

section. Coral reefs provide excellent, albeit often overlooked, examples. They are spatially discrete and smaller in vertical and horizontal structure than most of their terrestrial counterparts. Most of their functional components are readily visible (there are few below-surface functions), and major alterations and resilience can be measured in years and decades rather than centuries for observing ecosystem changes. Biogeographically, they vary widely in species diversity and thus are ideal for comparing ecosystem structure and function relative to their biodiversity.

If relatively few environmental factors can structure communities and ecosystems at the functional group level, then what role does species diversity play? Here, I consider biodiversity and its role in the stability, structure, and functioning of coral reef ecosystems. If functional groups are composed of ecologically equivalent species, then perhaps diversity within these groups contributes to the redundancy of important functions and thus may be ecosystem insurance for stability.

Diversity and Stability

Ecosystem and community stability was once thought to correspond directly with species diversity. MacArthur (1955)

theorized that the stability of a community will increase as the number of links in its food web increases. It followed that with the loss of each species "the integrity of the biosphere will degrade in a small but significant way" (Lawton and Brown, 1994, page 255). Alternatively, species richness may be relatively unimportant as long as structuring processes controlled by primary producers, consumers, and decomposers are maintained and function well with very few species. In fact, it is now widely believed that "ecosystem processes often have considerable redundancy built into them" (Lawton and Brown, 1994, page 266). Functional groups can be composed of the redundant species that perform essential ecosystem processes.

There is good empirical evidence that diverse ecosystems are not stable. Coral reefs are among the most diverse and most productive ecosystems in the world, but they are not characteristically stable. In the Caribbean and Indo-Pacific reef coral cover, abundance and dominance suddenly changed as the result of changes in abundance of a single echinoderm species in each of the two regions. In 1983, the mass mortality of the Caribbean sea urchin, *Diadema antillarum*, caused a rapid "phase shift" that increased macroalgal biomass and contributed to the decline of live coral cover in Curacao, St. John, Jamaica, and St. Croix. In all cases, algal assemblage dominance shifted from turf algae (**Figure 8A**, Nos. 1–3) to macroalgae (**Figure 8A**, Nos. 4–6), and this caused a decline in mass-specific primary productivity. It also precipitated changes in the abundance of other herbivores. The abundance of all reef-building organisms (corals, hydroids, and coralline algae) declined sharply in shallow zones where carbonate production is normally highest.

In the Indo-Pacific, periodic, localized outbreaks of the predatory sea star *Acanthaster planci* caused coral cover declines of all species of between 85 and 94% as reviewed by Connell (1997). Other disturbances, such as storms, sedimentation, toxins, and fishing, have also been shown to cause significant reductions in coral cover. In the cases of very rapid ("acute or pulse") mortality from sea star predation, rapid recovery at decadal timescales was observed. Other longer lasting ("chronic or press") mortality events had slower rates of recovery. However, in the scale of years to decades these highly diverse ecosystems were demonstrably unstable.

Ecosystem resilience is sometimes considered a correlate of stability. Rapid recovery from disturbances was more commonly observed in reefs of the Indo-Pacific than in reefs of the Caribbean, where no recoveries have been documented. Reef recovery relates more to the nature and duration of the disturbances than to the biodiversity of the system. One of the best examples of recovery from collapse is in some reefs of Hawaii, although the species diversity there was as low as that found throughout the Caribbean (**Table 1**).

Diversity and Ecosystem Structure

The basic structure of coral reef ecosystems does not vary with biodiversity. Coral reefs throughout the Indo-Pacific, Caribbean, and at Clipperton Atoll (eastern Pacific) have similar morphological zonation of corals despite a two orders of magnitude difference in coral species richness (**Table 1**). Branching corals dominate shallow zones, mound-shaped forms at greater depths, and platy corals at still greater depth. At Clipperton Atoll, one of the most species-depauperate reefs in the world, Glynn and coworkers (1996) measured vertical growth rates of coral similar to those found for similar morphologies at similar depths elsewhere in the Indo-Pacific. Steneck (1988) found similar patterns of zonation, distribution, and abundance among marine algae and herbivores at the functional group level in most studied reef systems of the Indo-Pacific and Caribbean. Others showed that functional characteristics of reef fish, such as size, activity rhythms, and relative abundances among ecological and trophic categories, were similar regardless of the species richness of the fish fauna.

Diversity and Ecosystem Function

Species diversity is always low at high trophic levels. Often a single, "apex" predator is critical to the structure of some systems. However, at lower trophic levels, diversity increases, and it is at these levels (among herbivores and primary producers) that the relationship between ecosystem diversity and function is best examined at the level of functional groups.

Herbivores

Reef-dwelling herbivores have been grouped according to their grazing abilities on algal communities (**Table 2**). Deep grazing (i.e., carbonate scraping herbivores), parrotfishes, and certain sea urchins (e.g., *Diadema*) have the greatest impact on algal abundance. Other fishes, such as surgeonfish (acanthurids), certain damselfishes and blennies, and some gastropods, are "denuding" herbivores in that they can significantly reduce fleshy algal biomass when at sufficiently high densities. This group does not bite into limestone structures of reefs. They cannot feed on crustose corallines and have a limited ability to consume articulated algae and large leathery macroalgae. Nondenuding herbivores have little or no ability to reduce algal biomass and include territorial damselfishes and smaller herbivores, such as amphipods and polychaetes. Each of these herbivore functional groups performs distinctly different ecological roles in the ecosystem, and all were found on the species-depauperate Clipperton Atoll. The abundance of species within a functional group compensates for low diversity

Table 1 Comparison of Coral Reef biodiversity[a]

Region	*Coral species*	*Reference*	*Fish species*	*Reference*
Indo-Pacific	350	Karlson and Cornell (1998)	3000	Lieske and Myers (1994)
Caribbean	50	Karlson and Cornell (1998)	750	Lieske and Myers (1994)
Hawaii	50	Hoover (1998)	680	Hoover (1993)
Clipperton (Eastern Pacific)	7	Glynn *et al.* (1996)	101	Robertson and Allen (1996)

[a]Species richness of scleractinian corals and reef fishes.

Table 2 Functional groups of herbivores common on Coral Reefs[a]

Functional group	*Specific group*	*Common reef-dwelling genera*
Scraping	Parrotfish	*Scarus, Sparisoma*
	Sea urchins	*Diadema, Echinometra, Echinothrix, Lytechinus*
	Limpets, chitons	*Acmaea, Acanthopleura, Acanthochitona, Choneplax*
Denuding	Surgeonfish, signids,	*Acanthurus, Naso, Zebrasoma, Siganus*
	Few pomacentrids	*Microspathodon*
	Blennies, kyphosids	*Ophioblennius*
Nondenuding	Damselfish	*Stegastes, Pomacentrus, Eupomacentrus*
	Snails (excluding limpets)	*Nerita, Cittarium, Tectarius, Astraea, Fissurella*
	Amphipods	*Amphithoe*, tanaids
	Polychaetes	Eunicids, syllids.

[a]Based on more than 30 published studies (after Steneck, 1988).

within the group. For example, high rates of scraping herbivory were evident at Clipperton, but all such grazing resulted from a single parrotfish species and several species of diadematid urchins. The abundant crustose coralline algae, which are often found on reefs experiencing high rates of grazing, were covered with graze marks from scraping herbivores (**Table 2**). High rates of grazing, low fleshy algal biomass, and a high abundance of both corals and coralline algae are characteristics of relatively pristine reefs. That these conditions were found in as species depauperate a system as Clipperton Atoll suggests that ecological roles of functional groups can be independent of species diversity.

Primary Producers

Primary productivity is a significant component of ecosystem functioning. A functional approach to ecosystem study often "implies that energy flow and nutrient cycling are somehow more important or more fundamental than the biotic entities performing the function" (O'Neill *et al.*, 1986, page 10). Although there may be no species in common, the same functional groups of marine algae are found on virtually all coral reefs. Since the productive components of reefs (algal turfs, corals, and macroalgae) all occur in approximately similar proportions, ecosystem-level productivity is fairly consistent. Thus, Caribbean reefs with only a fraction of the species richness of the western Indo-Pacific reefs have similar rates of gross primary productivity. Other factors, such as the geological history of the reefs and their proximity to land, have much greater impacts on ecosystem production than do reef biogeography or species diversity per se. Note that these patterns of consistent rates of productivity among unstressed reefs do not negate the well-known pattern of low diversity in stressed habitats.

Equivalent Species and Ecosystem Redundancy

Species within a function group, by definition, are ecologically equivalent and as such provide a degree of redundancy to the ecosystem. One test of functionally equivalent species is if their abundance compensates as other species within the functional group change. Morphological and functional equivalency is evident within turf algae on coral reefs. Many genera of different divisions look similar. They often grow as prostrate branching filaments, with relatively few laterally divided "pericentral" cells. Many species among these genera (**Figure 10A**) have about the same high rates of productivity (**Figure 8A**) and they respond similarly to injuries. In the fore reef of St. Croix, dominance and abundance of algal species changed dramatically from season to season at the species level but very little at the functional group level (**Figure 10B**). That algal community biomass remained relatively constant throughout the year, as species composition fluctuated greatly, suggests that species within functional groups compensate for species fluctuations.

It is necessary to sort out functionally important changes from redundant species "noise." On the fore reef of St. Croix (**Figure 10**), seasonal changes among species were so great that it was impossible to describe it as having a characteristic dominant species. However, at the functional group level, a clear pattern of filamentous and microalgal (i.e., turf algal assemblage) dominance was evident. These algae readily coexist with reef-building corals. In contrast, macroalgae comprised a minor component of this flora and appeared only in winter and spring. The functional distinction between macroalgae and the dominant micro- and filamentous algae is important. The algal turf assemblage was composed of species which have the highest mass-specific productivity (i.e., algal functional group Nos. 1 and 2 in **Figure 8A**). Trophically, these algal communities support very high levels of grazing. The herbivorous fish and urchins cannot distinguish species within the diverse algal turfs, and thus no species preferences are possible and there are no know chemical deterrents within this assemblage. Any of the filamentous algal species could have gone locally extinct (many did seasonally) without functionally changing the system. As such, many of the species within these functional groups are redundant.

Compensation is also evident among ecologically equivalent functional groups of herbivores. In shallow reef zones, the grazing effects from parrotfish and sea urchins are functionally similar (**Table 2**). They are both capable of grazing deeply into limestone substrates and can control the structure of algal communities on reefs. Several studies have reported that deep-grazing sea urchins and fish compete and show compensatory changes in relative abundance when one group is excluded or removed. Hay (1984) found that reefs where fishing was important were dominated by sea urchin grazing, whereas with little fishing pressure reefs were dominated by herbivorous

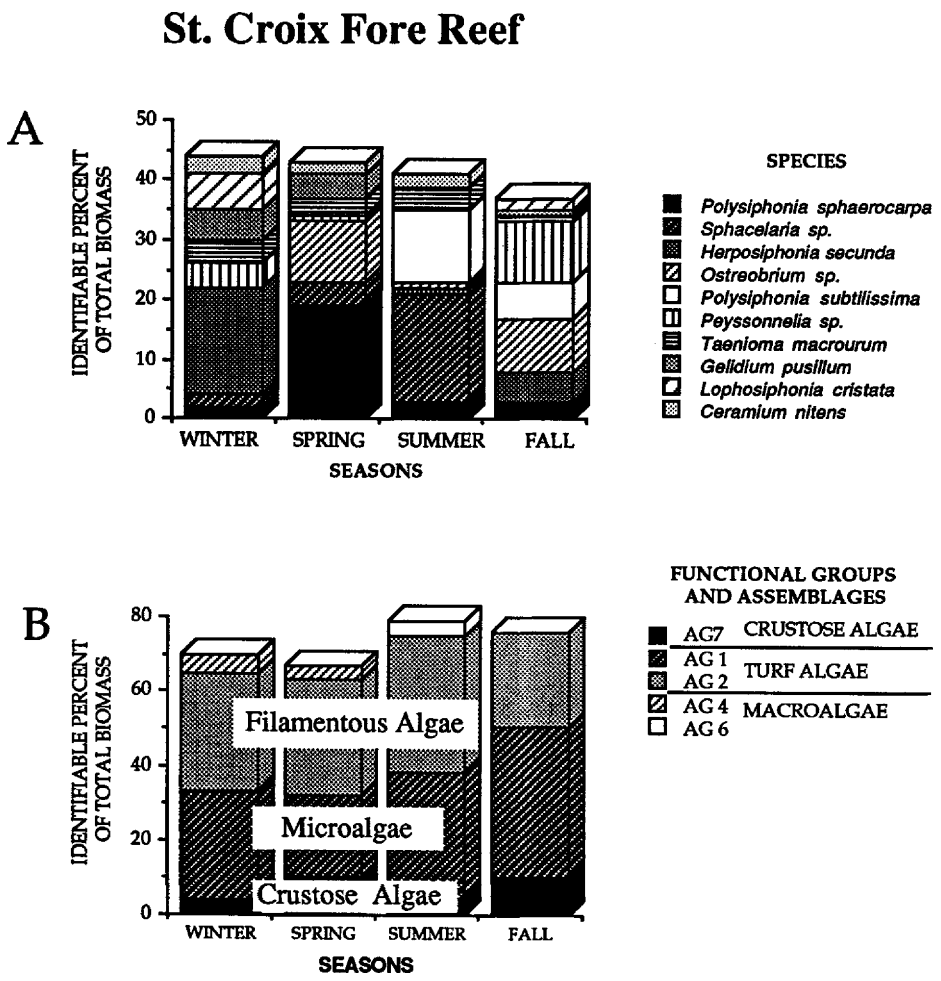


Figure 10 Seasonal stability of the 10 most abundant algal species (A) and functional groups (B) on a fore reef of St. Croix. Differences in the total percentage represented among seasons for species (A) and functional groups (B) resulted from the different proportions of unidentifiable components when samples were analyzed at species *vs* functional group levels. Functional group numbers correspond to the ones plotted in **Figure 8**. Functional groups were further combined into common tropical algal assemblages (i.e., crustose algae, turf algae, and macroalgae). Reproduced from Steneck RS and Dethier MN (1994) A functional group approach to the structure of algal-dominated communities. *Oikos* 69: 476–498.

fishes (e.g., scraping and denuding fishes in **Table 2**). Subsequent studies at several locations in the Caribbean and Indian Ocean found that when urchins were removed experimentally or naturally due to a disease-induced mass mortality, fish abundance and grazing increased as urchin abundance declined. In regions of the Caribbean traditionally having a relatively low abundance of *Diadema antillarum*, another urchin, *Echinometra viridis*, was abundant and filled *Diadema's* functional role.

Ecosystem Insurance

Species diversity within functional groups buffers ecosystems against species loss, but often there is "a threshold of change that will overwhelm the damping effect of biodiversity, with an associated break point of ecosystem function to quite different levels" (Schulze and Mooney, 1994b, page 501). Unfished reefs contain high within-functional group diversity at several levels. There are several species of large carnivores (e.g., groupers and snappers), denuding herbivores (e.g., surgeonfish), and scraping herbivores (e.g., parrotfish) that, when abundant on reefs, control the reef's overall structure and functioning. Fishing affects reefs by first removing large-bodied carnivores and then herbivores. Reefs under particularly intense fishing, such as in Jamaica, are left with few or very small fishes. Compensatory population increases in the sea urchin *D. antillarum* probably occurred (**Figure 11**), causing it to become hyper-abundant and the prime grazer of the reefs of Jamaica and elsewhere in the Caribbean. Levitan (1992) showed that *Diadema's* increasing role as a grazer throughout the Caribbean corresponded with increasing human populations and associated fishing pressure. On many Caribbean reefs this herbivore became the primary grazer within the functional group of scraping herbivores (**Table 2**). Arguably, the diversity within this functional group had decreased to one species (**Figure 11**). When *D. antillarum* suffered a mass mortality in 1983, the "break point" had been exceeded and reef ecosystems throughout the Caribbean collapsed.

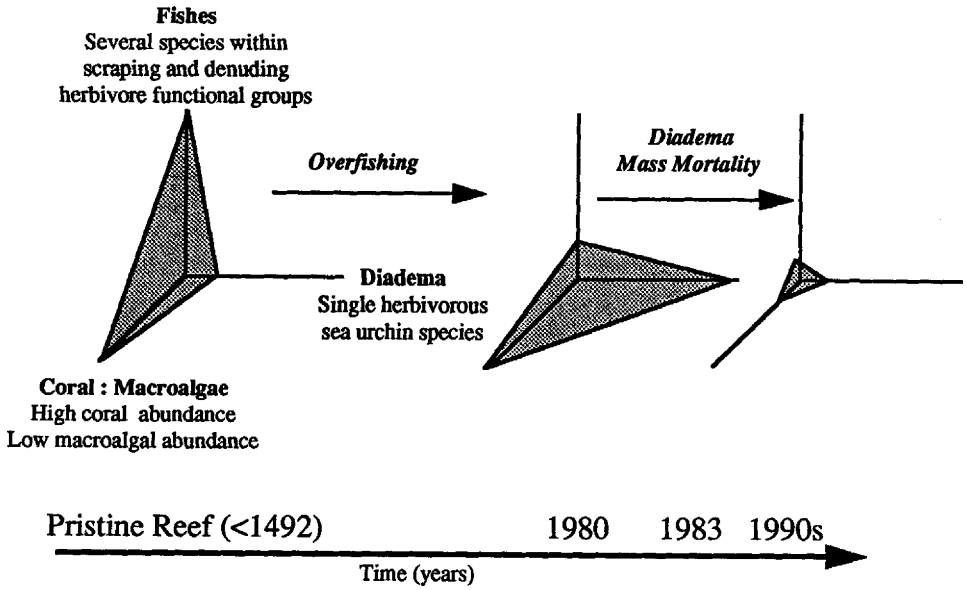


Figure 11 Temporal changes in Jamaica's coral reef ecosystem in coral to macroalgal abundance relative to herbivorous fish and sea urchin, *Diadema antillarum*. After herbivorous fishes were extirpated by overfishing, grazing remained high due to compensatory increases in *Diadema* abundance. The reef ecosystem collapsed soon after the mass mortality of the urchin (see text) (Adapted from Done TJ (1995) Ecological criteria for evaluating coral reefs and their implications for managers and researchers. *Coral Reefs* 14: 183–192.).

Keystone Species to Functional Groups: Some Species Are More Equal Than Others

Some species have a disproportionately great impact on ecosystems. Hurlbert (1997) noted, "The general functional importance of a species (can be) defined as the sum, over all species, of the changes (sign ignored) in productivity which would occur on removal of the particular species from the (system)" (page 369). Most communities and ecosystems are dominated by a few species that are both abundant and important to that system (**Figure 12**). The few tree species that dominate given forests, corals that dominate zones of reefs, and specific grasses that comprise specific prairies need not be functionally grouped to be easily understood. The loss of community or ecosystem dominants has a large impact on natural systems. Diseases can eliminate dominant organism with impacts registering throughout the system. American chestnut blight eliminated chestnuts from forests in eastern North America. White-band disease killed most of the dominant elkhorn and staghorn corals and a pathogen caused the *Diadema* decline in shallow Caribbean reefs. Rinderpest decimated wildebeest populations in Africa, and paramoebae caused a mass mortality of the green sea urchin in Nova Scotia. Each of these dominant or "foundation" (*sensu* Dayton, 1972) species underwent a radical decrease in abundance, causing their respective ecosystems to change. In these cases, the functional importance of these species is unquestioned.

Keystone species have been redefined by Paine (1995) and Power *et al.* (1996) to include organisms that are functionally important but in relatively low abundance in the system. Pathogens aside, this limits keystone species to apex predators. Classic keystone predators such as sea otters or the predatory sea star, *Pisaster*, are the only species in their system and thus define their ecological function. However, many systems do not have an obvious apex predator. In contrast with the marine systems of the eastern North Pacific in which sea otters and *Pisaster* roam, the western North Atlantic has no single apex predator. Top coastal predators include cod, haddock, hake, and wolffish. When all these top predators were extirpated simultaneously by overfishing, the system changed. Therefore, whether there is a single apex predator (i.e., a keystone species) or a functional group filling the same role matters little. In other words, the ecological function can be carried out by a single species or by a suite of species, but the ecosystem will not change unless the ecological function changes significantly. In the end, the ecological function, independent of the diversity, may be most important to the structure of ecosystems.

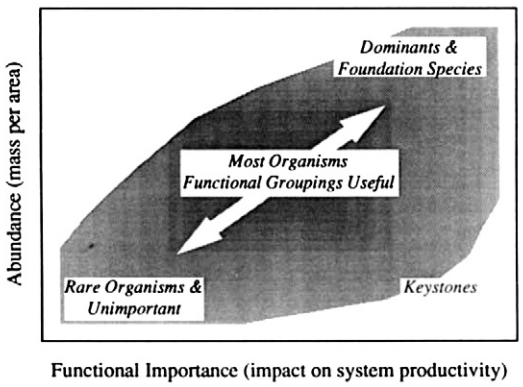


Figure 12 Functional importance and abundance of species. The relatively low diversity of dominant/foundation and keystone species as well as the relatively low ecosystem importance of rare organisms make these groups poorer candidates for functional groupings (Modified from Hurlbert SH (1997) Functional importance vs keystoneness: Reformulating some questions in theoretical biocenology. *Austr. J. Ecol.* 22: 369–382.).

The utility of functional groupings relates directly to the diversity of the system, community, or trophic level of interest. It will be most useful in sorting out the functional importance in diverse systems among moderate to very abundant organism (i.e., the center of **Figure 12**). Functional groups lose their utility among the rare, dominant, and keystone species (i.e., the periphery of **Figure 12**). Not only is species diversity low at high trophic levels but also it is characteristically low in certain stressed habitats, such as deserts, the rocky intertidal zone, estuaries, and high-altitude or high-latitude habitats. Functional group utility is likely to be greatest in diverse systems in which many species perform similar ecological functions.

Conclusions

Despite Earth's remarkable biodiversity, the range of functionally different organisms is surprisingly low. Phyletic constraints on evolution relative to biomechanical limits of what works in nature limit the number of functional groups in any system to a management level. In highly diverse ecosystems there are many redundant species within functional groups. These redundancies provide a buffer against ecosystem collapse should individual species within the group become rare or extinct. Ultimately, functional groups may provide a low-resolution tool for accurately predicting ecosystem change. Research to improve how we identify and define functional groups is likely to increase our understanding of the structure and functioning of natural communities and ecosystems.

Acknowledgments

Drafts were read by E. Annis, A. Leland, J. Vavrinec, and S. Zimsen, with additional help from P. Yund and L. Watling. Greg Welch helped with manuscript preparation. To all, I am grateful.

See also: Biodiversity Generation, Overview. Cladogenesis. Coevolution. Corals and Coral Reefs. Guilds. Mammals, Biodiversity of

References

Bambach RK (1985) Classes and adaptive variety: The ecology of diversification in marine faunas through the Phanerozoic. In: Valentine JW (ed.) *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, pp. 191–253. Princeton, NJ: Princeton Univ. Press.

Begon M, Harper JL, and Townsend CR (1986) *Ecology, Individuals, Populations and Communities*. Sunderland, MA: Sinauer.

Coates AG and Jackson JBC (1985) Morphological themes in the evolution of clonal and aclonal invertebrates. In: Jackson JBC, Buss LW, and Cook RE (eds.) *Population Biology and Evolution of Clonal Organisms*, pp. 67–106. New Haven, CT: Yale Univ. Press.

Connell JH (1997) Disturbance and recovery of coral assemblages. *Coral Reefs* 16(Suppl.): S101–S113.

Crimes TP and Harper JC (1977) *Trace Fossils 2, Geological Journal*, Special Issue No. 9. Liverpool, UK: Seel House Press.

Cunningham CW, Blackstone NW, and Buss LW (1992) Evolution of king crabs from hermit crab ancestors. *Nature* 355: 539–542.

Darwin C (1859) *The Origin of Species by Means of Natural Selection*. New York: Random House. (Reprinted by Modern Library).

Done TJ (1995) Ecological criteria for evaluating coral reefs and their implications for managers and researchers. *Coral Reefs* 14: 183–192.

Glynn PW, Veron JEN, and Wellington GM (1996) Clipperton Atoll (eastern pacific): Oceanography, geomorphology, reef-building coral ecology and biogeography. *Coral Reefs* 15: 71–99.

Grime JP (1981) *Plant Strategies and Vegetation Processes*. New York: Wiley.

Hay ME (1984) Patterns of fish and urchin grazing on Caribbean coral reefs: Are previous results typical? *Ecology* 65: 446–454.

Holdridge LR (1947) Determination of world plant formations from simple climatic data. *Science* 105: 367–368.

Hoover JP (1993) *Hawaii's Fishes: A Guide for Snorkelers, Divers, and Aquarists*. Honolulu, HI: Mutual.

Hoover JP (1998) *Hawaii's Sea Creatures: A Guide to Hawaii's Marine Invertebrates*. Honolulu, HI: Mutual.

Hurlbert SH (1997) Functional importance vs keystoneness: Reformulating some questions in theoretical biocenology. *Austr. J. Ecol.* 22: 369–382.

Huston MA (1994) *Biological Diversity. The Coexistence of Species on Changing Landscapes*. Cambridge, UK: Cambridge Univ. Press.

Hutchinson GE and MacArthur R (1959) A theoretical ecological model of size distributions among species of animals. *Am. Nat.* 93: 117–126.

Karlson RH and Cornell HV (1998) Scale-dependent variations in local vs regional effects on coral species richness. *Ecol. Monogr.* 68: 259–274.

Lawton JH and Brown VK (1994) Redundancy in ecosystems. In: Schulze ED and Mooney HA (eds.) *Biodiversity and Ecosystem Function*, pp. 255–268. New York: Springer.

Levitan DR (1992) Community structure in times past: Influence of human fishing pressure on algal urchin interactions. *Ecology* 73: 1597–1605.

Lieske E and Myers R (1994) *Coral Reef Fishes: Caribbean, Indian Ocean, and Pacific Ocean Including the Red Sea*. Princeton, NJ: Princeton Univ. Press.

MacArthur RH (1955) Fluctuations of animal populations and a measure of community stability. *Ecology* 35: 533–536.

Moody KE and Steneck RS (1993) Mechanisms of predation among large decapod crustaceans of the Gulf of Maine coast: Functional vs phylogenetic patterns. *J. Exper. Marine. Biol. Ecol.* 168: 111–124.

Naeem S (1998) Species redundancy and ecosystem reliability. *Conserv. Biol.* 12: 39–45.

Niklas KJ (1994) *Plant Allometry: The Scaling of Form and Process*. Chicago: Univ. of Chicago Press.

O'Neill RV, DeAngelis DL, Waide JB, and Allen TFH (1986) *A Hierarchical Concept of Ecosystems*. Princeton, NJ: Princeton Univ. Press.

Paine RT (1995) A conversation on refining the concept of keystone species. *Conserv. Biol.* 9: 963–964.

Peters RH (1983) *The Ecological Implications of Body Size*. Cambridge, UK: Cambridge Univ. Press.

Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *BioScience* 46: 609–620.

Raunkiaer C (1934) *The Life Forms of Plants and Statistical Plant Geography*. Oxford: Clarendon.

Robertson DR and Allen GR (1996) Zoogeography of the shorefish fauna of Clipperton Atoll. *Coral Reefs* 15: 121–131.

Schulze E and Mooney HA (eds.) (1994a) *Biodiversity and Ecosystem Function*. New York: Springer.

Schulze E-D and Mooney HA (1994b) Ecosystem function of biodiversity: A summary. In: Schulze E-D and Mooney HA (eds.) *Biodiversity and Ecosystem Function*, pp. 497–510. New York: Springer.

Sepkoski JJ Jr. (1984) A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* 10: 246–267.

Southwood TRE (1988) Tactics, strategies and templets. *Oikos* 52: 3–18.

Steneck RS (1982) A limpet coralline alga association: Adaptations and defenses between a selective herbivore and its prey. *Ecology* 63: 507–522.

Steneck RS (1988) Herbivory on coral reefs: A Synthesis. *Proc. 6th Int. Coral Reef Symp. Austr.* 1: 37–49.

Steneck RS and Dethier MN (1994) A functional group approach to the structure of algal-dominated communities. *Oikos* 69: 476–498.

Steneck RS and Watling LE (1982) Feeding capabilities and limitations of herbivorous molluscs: A Functional group approach. *Mar. Biol.* 68: 299–319.

Thompson D'arcy W (1966) *On Growth and Form*. Cambridge, UK: Cambridge Univ. Press.

Fungi

Thomas J Volk, University of Wisconsin-La Crosse, La Crosse, WI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Exoenzymes Digestive enzymes excreted by fungi into the environment to digest materials externally.
Fruiting body Sexual reproductive structure of a fungus.
Fungus A member of the kingdom Fungi. Members of this kingdom are heterotrophic (requiring a preformed organic source, i.e., not able to make their own food), eukaryotic, have walls of chitin, and reproduce by means of spores.
Hyphae The feeding (vegetative) threads of most fungi. The hyphae of one organism are sometimes referred to as a mycelium.
Mushroom Vernacular word for a large, fleshy fruiting body consisting of a cap with gills or pores on the undersurface (or sometimes flat or with teeth or folds), usually on a stalk, and producing sexual spores. The term is usually reserved for members of the Basidiomycota. A poisonous mushroom is sometimes called a toadstool.
Mutualism Two organisms living together for the mutual benefit of both.
Mycorrhiza Part of a mutualistic relationship between a fungus and a host plant. The two types are ectomycorrhizae (usually a Basidiomycota, but sometimes an Ascomycota) in which hyphae aggregate as an extra surface around roots of the plant, and endomycorrhizae (a Glomeromycota), in which no sheath is formed, but nutrient transfer occurs with the aid of highly branched arbuscules.
Parasite Organism that obtains its nutrition from a living host, and in so doing harms the host.
Spore One- to several-celled propagule of totipotent cytoplasm with cell walls, produced by cell division with concomitant meiosis or mitosis, that may serve for dispersal or overseasoning, but that does not contain an embryo.

Introduction

When someone mentions "fungi" you may think immediately of mushrooms on pizza or maybe moldy food in your refrigerator. But in fact fungi are everywhere and affect our lives every day, from edible mushrooms to industrially important products to plant helpers, plant pathogens, and human diseases. Many fungi are detrimental in causing a large number of plant diseases that result in the loss of billions of dollars worth of economic crops each year. They also cause a number of animal diseases, including many human maladies. On the other hand, there is a long and rapidly growing list of useful fungi. Fungi have been used in the preparation of foods and beverages for thousands of years, and there are many mushrooms that are edible. Industry has used other fungi in the manufacture of many valuable organic compounds, including organic acids, vitamins, antibiotics, and hormones. They have been used in the research laboratory to study metabolic pathways, mineral nutrition, genetics, and a variety of other problems. But perhaps their greatest contribution has been, and continues to be, their role in recycling carbon and other essential elements in the ecosystem. Because all fungi are heterotrophic (*see* Fungi in the Environment), they rely on organic material, either living or dead, as a source of energy. Thus, many are excellent scavengers in nature, breaking down dead animal and vegetable material into simpler compounds that become available to other members of the ecosystem.

"Mycology" comes from the Greek words *mykos*, which means "fungi," and "-logy," which means "the study of." Mycologists study many aspects of the biology of fungi, usually starting with their systematics, taxonomy, and classification, and continuing on to their physiology, ecology, pathology, evolution, genetics, and molecular biology. There are quite a few fields of applied mycology, such as plant pathology, human pathology, fermentation, mushroom cultivation, and many other fields.

What are Fungi?

Characteristics of Fungi That Separate Them from the Other Kingdoms

The most significant characteristics of fungi that place them apart from the plant, animal, and other kingdoms are the following:

- Fungi are eukaryotic.
- They are nonvascular organisms, meaning they have no specialized transport tissues.
- Most reproduce by means of spores, usually wind-disseminated, and occasionally by insects in some species.
- Both sexual (meiotic) and asexual (mitotic) spores may be produced, depending on the species and conditions.
- Fungi are typically not motile, although a few (e.g., chytrids) have a motile phase.
- Like plants, sexually reproducing fungi have an alternation of generations, although the generations may be different from those of plants.
- The fungal vegetative body may be unicellular (yeasts) or, more often, composed of microscopic threads called hyphae.
- Fungal cell walls are similar in structure to those of plants but differ in chemical composition – fungi cell walls are composed mostly of chitin, which are β-1,4 linkages of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00062-9

Table 1 Taxonomic endings for the fungi

Taxonomic level	*Taxonomic ending*	*White button pizza mushroom*	*The morel*
Domain	Eukarya	Eukarya	Eukarya
Kingdom	Fungi	Fungi	Fungi
Phylum	-mycota	Basidiomycota	Ascomycota
Class	-mycetes	Hymenomycetes	Discomycetes
Order	-ales	Agaricales	Pezizales
Family	-aceae	Agaricaceae	Morchellaceae
Genus	–	*Agaricus*	*Morchella*
Species	–	*Agaricus bisporus*	*Morchella esculenta*

n-acetylglucosamine. In contrast, plant cell walls are composed mostly of cellulose, which are β-1,4 linkages of glucose. Many plants also contain lignin in their secondary walls.

- Fungal cytoplasmic ultrastructure is broadly similar to that of plant cells, but differs significantly in the kinds of organelles and their structures.
- Fungi are heterotrophic (i.e., "other feeding" in that they must feed on preformed organic material), not autotrophic (i.e., "self-feeding" in that they make their own food by photosynthesis, like most plants and algae).
- Unlike animals (which are also heterotrophic), which ingest and then digest, fungi first digest and then ingest. Fungi produce exoenzymes to accomplish this.
- Most fungi store their food as glycogen (like animals) – plants store food as starch.
- Fungal cell membranes have a unique sterol, ergosterol, which replaces the cholesterol found in mammalian cell membranes.
- The lysine biosynthesis pathway in fungi is different.
- The microtubule protein formed during nuclear division is different from that of all other organisms.
- Most fungi have very small nuclei, with little repetitive DNA, usually with few chromosomes.
- Mitosis and meiosis are generally accomplished without dissolution of the nuclear envelope.

There are about 75,000 named species of fungi and this is believed to be about 5% of the total number of species that exist in nature. If this is the case, 95% of all fungal species are unknown to science and do not yet have names.

In the outdated two-kingdom system, fungi were included in the plant kingdom, as were almost all walled organisms that are not motile. In almost all systematic schemes today, the heterotrophic, eukaryotic fungi are placed in their own kingdom, cleverly called the kingdom Fungi (Alexopoulos *et al.*, 1996).

The eukaryotes, organisms with nuclei and membrane-bound organelles, can be divided into five kingdoms: Fungi, Plantae, Animalia, Protista, and Stramenopila. Recent DNA evidence suggests that fungi are more closely related to animals than to plants. Prokaryotes, without nuclei and membrane-bound organelles, include the Bacteria and the Archaea.

As with all taxonomy, the names of various taxa of fungi each have a specific ending that refers to their taxonomic level. The fungal kingdom has its own endings for many taxa; these include a suffix containing -myc. It is of considerable advantage to be familiar with these endings so that one knows immediately what taxonomic level is being referred to when confronted with a particular name. The names of fungi are "regulated" by the International Code of Nomenclature for Algae, Fungi, and Plants, and thus endings for the order and family levels are the same as those of plants (**Table 1**).

Fungal Roles in the Ecosystem

Lack of chlorophyll profoundly affects the lifestyle of fungi. Since they are not dependent on light, they can occupy dark habitats and can grow in any direction. The exoenzymes allow them to invade the interior of a substrate with absorptive filaments. Fungi may gain their nutrition from dead organisms, in which case they are called saprophytes. Some fungi derive their nutrition from living organisms; these are called symbionts. Symbionts can be further divided on the basis of whether or not they harm their host. Parasites cause harm to the host, whereas mutualists engage in a reciprocally beneficial association with their host. These categories are further described in the section Fungi in the Environment.

Importance of Studying Fungal Biodiversity

Fungi affect human lives in many and varied ways, so it is important to know something about the fungi to be able to control or exploit them for our own purposes. For example, more than 90% of known fungal species have never been screened for antibiotics or other useful compounds. However, even more important is the role that fungi play in the ecosystem. They are a vital part of the links in the food web as decomposers and pathogens and are important in grassland and forest ecosystems alike. Fungi have many different kinds of associations with other organisms, both living and dead. To learn more about the impact that fungi have on our lives, we must learn a lot more about them.

Many fungi are harmful to human interests. They can cause human disease, either directly or through their toxins, including mycotoxins and mushroom poisons. They can also cause diseases of plants and animals (e.g., crops, fruit trees, farm animals). Very often fungi cause rot and contamination of foods – most of us probably have something green and moldy in the back of our refrigerator right now. They can destroy almost every kind of manufactured goods, with the exception of some plastics and some pesticides.

On the other hand, many fungi are very useful to humans (Hudler, 1998). Of course, there are many edible mushrooms. Yeasts have been used for baking and brewing for many

millennia. Antibiotics such as penicillin and cephalosporin are produced by fungi. The immunosuppressive antirejection transplant drug cyclosporine is produced by the ascomycete *Tolypocladium inflatum*. Steroids and hormones – and even birth control pills – are commercially produced by various fungi. Many organic acids are also commercially produced with fungi, for example, citric acid in cola and other soda pop products is produced by an *Aspergillus* species. Some gourmet cheeses such as Roquefort and other blue cheeses, brie, and camembert are fermented with certain *Penicillium* species. Stone-washed jeans, strange as it sounds, are actually softened by *Trichoderma* species, not by old ladies with babushkas beating the jeans on rocks in a stream. There are likely many more potential uses that have not yet been explored.

Fungi are also important experimental organisms. They are easily cultured, occupy little space, multiply rapidly, and have a short life cycle. Since they are eukaryotes and more closely related to animals, their study is more applicable to human problems than is the study of bacteria. Fungi are used to study metabolite pathways, to study growth, development, and differentiation, for determining mechanisms of cell division and development, and for microbial assays of vitamins and amino acids. Fungi are also important genetic tools; the "one gene one enzyme" theory in *Neurospora* won George W Beadle and Edward L Tatum the Nobel Prize for Physiology or Medicine in 1958. The first eukaryotic genome to have its DNA sequenced was that of the bakers' and brewers' yeast, *Saccharomyces cerevisiae*. Lee Hartwell and his colleagues used *Saccharomyces cerevisiae* to study mitosis genes and won the Nobel Prize in 2001.

How Fungi Grow

Biology of Hyphae and Yeast Forms

A fungus is more than just the visible mushroom structure. In fact the mushroom, more properly called a fruiting body, is a very small portion of the individual life cycle and is mainly used for reproduction. The major portion of the life cycle, or the vegetative growth form, in the great majority of fungi consists of a system of threadlike, walled, more or less cylindrical hyphae (singular, hypha) making up what is called a mycelium (plural, mycelia) (**Figure 1**). The Ascomycota and Basidiomycota have crosswalls called septa (singular, septum) separating compartments of the mycelium. An exceptional group is the yeasts, which consist of about 800 species that have a single-celled vegetative form. Note that yeast is a morphological term and has no taxonomic significance; yeasts and yeastlike forms can be found in all of the fungal phyla.

Exoenzymes and the Heterotrophic Lifestyle

Exoenzymes are the most important reason why fungi are so successful. Fungi excrete exoenzymes at the tips of the growing hyphae into their surrounding environment, where they play a major role in breaking down the substrate. Simpler molecules can then move into the hyphae by diffusion. Exoenzymes are the reason that fungi can grow into the center of a solid log, into your brain and lungs, or into a wide variety of other substrates.

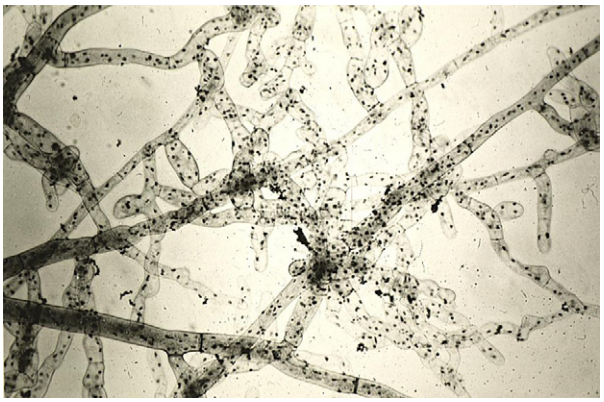


Figure 1 Hyphae with dark-stained nuclei.

Reproductive Biodiversity of Fungi

Fungal Life Cycles

The major events of any sexual life cycles are plasmogamy (cell and protoplast fusion), karyogamy (nuclear fusion), and meiosis (**Figure 2**). In most other familiar types of organisms, such as plants and animals, plasmogamy and karyogamy occur in rapid succession and are usually referred to as the single event of fertilization. In the fungi, however, plasmogamy and karyogamy may be separated in time by several minutes, several hours, several days, several years, or even several centuries! Thus the dikaryon, the $n + n$ stage, is a major component of the life cycle of fungi, especially in the Basidiomycota and Ascomycota. Nuclear cycles of all the members of the various phyla can be placed within this generalized nuclear cycle, differing mainly in the amount of time spent in each of the phases.

Besides this sexual cycle, many fungi, commonly called molds, can also reproduce asexually (mitotically) in the absence of karyogamy and meiosis. Many of them produce specialized structures that bear the asexual spores. As in much of biology, there is some "competing" terminology here – the asexual state is also known as the anamorph or mitosporic state. Asexual reproduction can take place at any point in the life cycle (haploid, diploid, or dikaryon), depending on the species and conditions. The sexual state is also known as the teleomorph or meiosporic state.

Phyla of fungi

Based primarily on variation in their sexual reproductive structures, the kingdom Fungi is usually divided into five major phyla.

- Chytridiomycota – sexual and asexual spores are motile, with posterior flagella.
- Zygomycota – sexual spores are thick-walled resting spores called zygospores, and asexual spores called sporangiospores

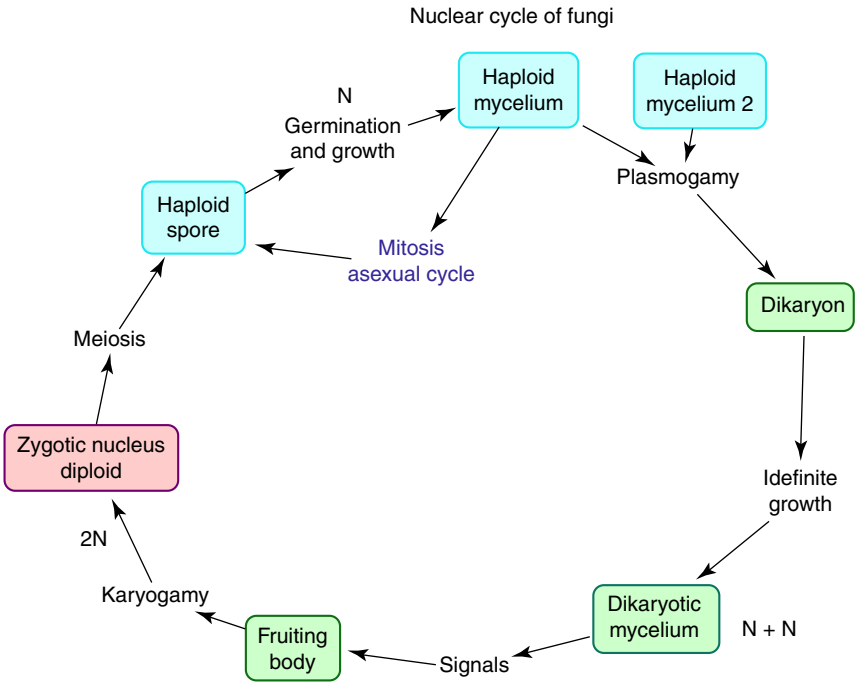


Figure 2 Generalized nuclear cycle for fungi.

(when present) are borne internally in structures called sporangia.

- Glomeromycota – form endomycorrhizae or AMF with 70% of the world's plants. Asexual reproduction is in the form of large underground spores, while sexual reproduction is unknown.
- Ascomycota – sexual spores (ascospores) are borne internally in a sac called an ascus, and asexual spores called conidia (when present) are borne externally on structures called conidiophores.
- Basidiomycota – sexual spores (basidiospores) are borne externally on a club-shaped structure called a basidium; most species do not have asexual spores, but when present they often take the form of conidia.

There is an additional group, the deuteromycetes, or "Fungi Imperfecti," for which there is no known sexual state. Its members have affinities to members of at least three of the phyla, especially the Ascomycota.

Chytridiomycota

The Chytridiomycota, commonly called the chytrids, are a group of mostly water-inhabiting organisms, although some are plant pathogens. In aquatic environments they mostly form scanty filaments with sporangia. Some examples of the Chytridiomycota are *Allomyces*, a water mold, *Synchytrium endobioticum*, a pathogen of potato, and *Neocallimastix*, a chytrid that lives symbiotically in the gut of herbivores, such as cattle. *Batrachochytrium dendrobatidis* has been implicated as an infection associated with the worldwide decline in frog populations. This chytrid infects the skin of amphibians, rendering it too thick for water and gas exchange, and the frogs die. Many species of frogs, particularly in the tropics, are thought to have gone extinct because of this quickly-spreading pathogen. It is unclear whether this is an introduced pathogen or whether it has always been there, but possibly becoming more prominent because of some environmental pollutant hampering frogs' immune systems.

There is some controversy about the placement of certain fungi into the Chytridiomycota, and there is some evidence to support that the multicellular forms, such as *Allomyces* and other members of the Blastocladiales, should be placed in a new phylum, the Blastocladiomycota. Additionally, there is some molecular evidence that the cow chytrid and related species should be placed into a new phylum, the Neocallimastigomycota. Clearly, systematics of this group is in a state of flux as new evidence continues to be discovered.

Zygomycota

Commonly called the bread molds, the Zygomycota are terrestrial fungi whose fruiting bodies are mostly microscopic in nature, although their asexually produced sporangia can reach greater than 5 cm tall in some species (**Figure 3**). Under certain conditions they may sexually produce thick-walled resting spores called zygospores (**Figure 4**). Some, such as *Rhizopus*, *Mucor*, and *Phycomyces*, can grow on a wide variety of substrates, and a few can act as human pathogens. Many species, such as *Pilobolus*, inhabit dung and pass through the guts of herbivores to get into more dung.

Glomeromycota

Members of the Glomeromycota, are responsible for forming mutualistic associations called endomycorrhizae with the roots of about 70% of the world's plants. These endomycorrhizae are

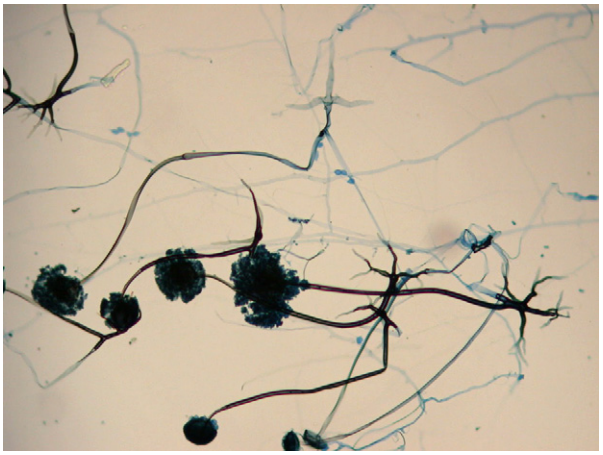


Figure 3 *Rhizopus* sporangia.

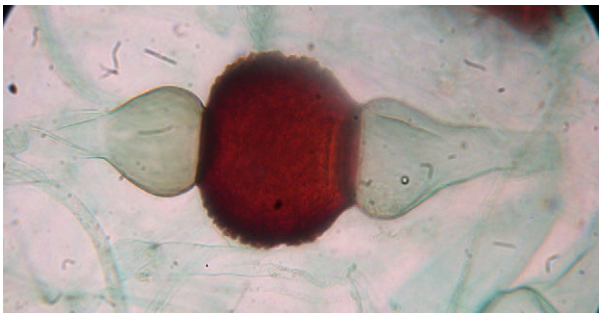


Figure 4 *Rhizopus zygospore* between suspensors.



Figure 5 *Morchella esculenta*, the morel, a prime edible mushroom.

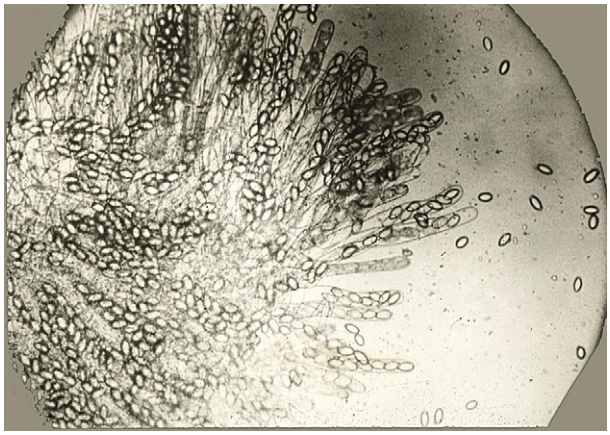


Figure 6 Morel asci containing ascospores.

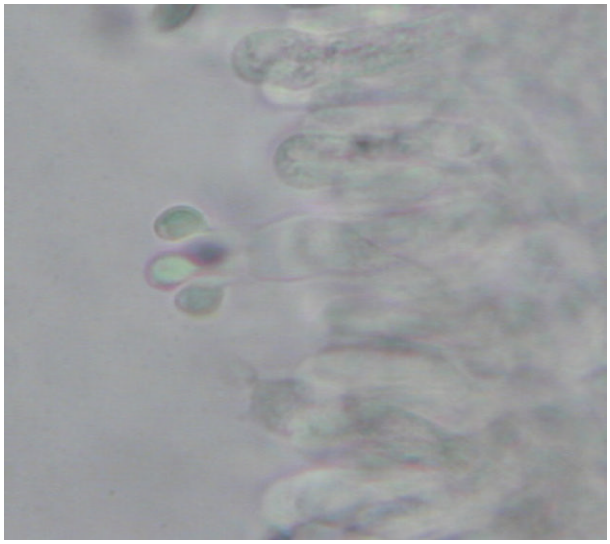


Figure 7 Basidium showing the four basidiospores produced by meiosis and pinching off of the spores from the basidium.

also known as arbuscular mycorrhizal fungi, abbreviated AMF. Glomeromycota form relatively large (up to 1 mm) asexual spores in the soil, but sexual reproduction is unknown. The Glomeromycota were once considered to belong to the order Glomerales (= Glomales) of the Zygomycota, but molecular data, along with the unique morphology and ecological niche, have confirmed that the Glomeromycota should be considered a separate phylum. See the section Mycorrhizae below for further discussion.

Ascomycota

The Ascomycota bear their sexual spores (ascospores) internally in sacs called asci, which are usually cylindrical. Many members also form conidia as asexual spores. Familiar members of this phylum include the morels (**Figures 5** and **6**) (Weber, 1988; Kuo, 2005) and other cup and saddle fungi, truffles, powdery mildews, the industrial yeast *Saccharomyces cerevisiae*, the incitant of chestnut blight (*Cryphonectria parasitica*,) the cause of Dutch elm disease (*Ophiostoma ulmi*), and a

variety of other plant pathogens. About 1/3 to 1/2 of the species of Ascomycota are associated with algae or cyanobacteria in the form of lichens (*see* Lichens), while a growing number of genera, such as *Morchella*, *Helvella*, and *Leotia*, are being discovered to form ectomycorrhizae with host plants.

Basidiomycota

The Basidiomycota bear their sexual spores externally on a usually club-shaped structure called a basidium, which is often borne on or in a fruiting body called a basidiocarp or basidiome (**Figure 7**). This phylum includes the well-known mushrooms, both edible and poisonous, as well as boletes, puffballs, shelf fungi, jelly fungi, and coral fungi (**Figure 8**). Many of the most prized edible species, such as king boletes, matsutake, and chanterelles, form ectomycorrhizae with host plants. (*see* Mycorrhizae) The species that produce fruiting bodies exhibit various methods of increasing their surface area, as discussed in the section Surface Area and Reproduction.

The Basidiomycota also include perhaps the most important crop plant pathogens, the rusts and the smuts, which are responsible for billions of US dollars every year in lost yield. These fungi do not produce macroscopic fruiting bodies, but instead bear their spores on the stems, leaves, and flowers of host plants. However, remember that the mycelium is internal and "sucks" the nutrients out of the plant using its exoenzymes. Effects on the plant range from a reduced yield to death. Rusts are also interesting because of their very complicated life cycles, often requiring two unrelated host species to complete their growth stages. For example, the agent of

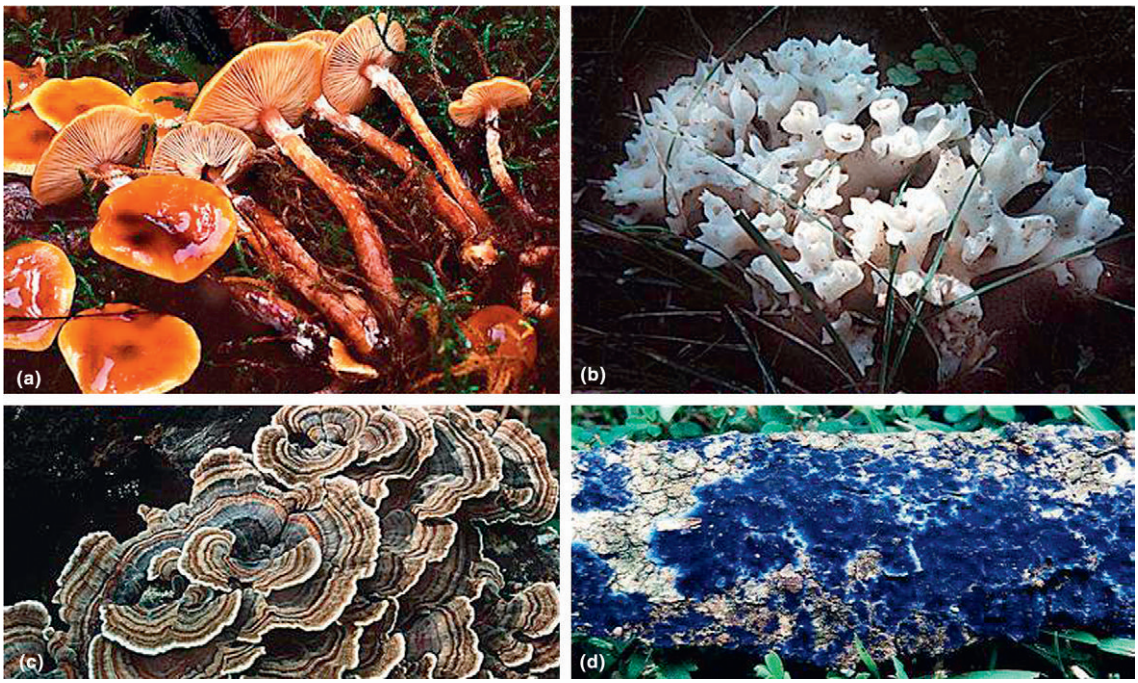


Figure 8 (a) *Armillaria nabsnona*, (b) *Tremella reticulata*, (c) *Trametes versicolor*, (d) *Pulcherricium caeruleum*.

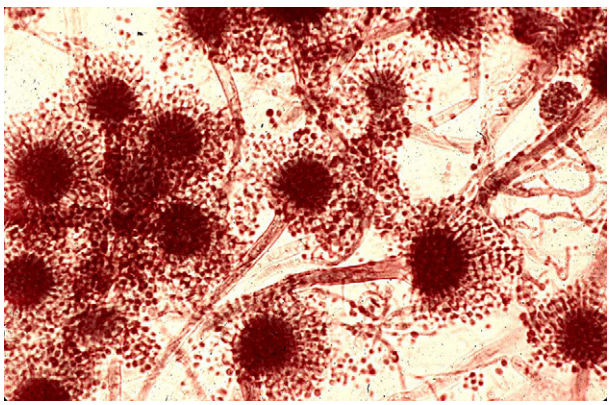


Figure 9 *Aspergillus* conidia.

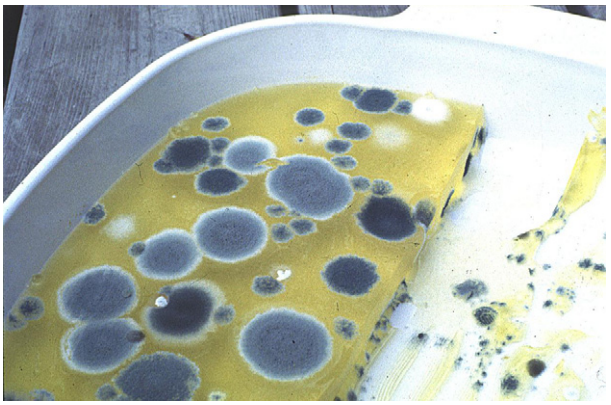


Figure 10 Jell-O mold, a favorite dish for picnics.

black stem rust of wheat alternates between wheat and barberry, taking a year to complete its life cycle.

"Deuteromycetes," the Fungi Imperfecti

The deuteromycetes, commonly called molds, are "second-class" fungi that have no known sexual state in their life cycle, and thus reproduce only by producing spores *via* mitosis, This asexual state is also called the anamorph state. Despite its "-mycetes" ending the deuteromycetes are not a class; rather this term is now used as a common name for this dumping-ground group of fungi. About 90% of these have affinities to the Ascomycota. Most food spoilage (**Figures 9** and **10**) and fungal human diseases are caused by members of this group. They are also known as the fungi imperfecti, because of their "imperfect" lack of sex. When the "perfect state" of one of these organisms is discovered, as happens every year, the fungus is more properly classified with the teleomorph name, although in practicality, like for people, it is difficult to remember to switch to their "married" name. Notice that this group is not classified as one of the phyla. It is just a loose assemblage of organisms that we have not been sure where to place accurately in the taxonomic order. Although we can accurately place most of these fungi into phyla using molecular biology techniques such as PCR and DNA sequencing, mycologists maintain this group and its genera for convenience and tradition (St-Germain and Summerbell, 1996).

Excluded Taxa: Former Members of the Kingdom Fungi

The organisms included here were once considered to be true fungi, but have been excluded because they do not exactly fit our modern classification scheme of the true fungi in various ways. They are now considered to belong in other kingdoms, such as Protista or Stramenopila.

Oomycota: Water Molds and Downy Mildews

The Oomycota have been excluded from the kingdom Fungi primarily because their cell walls are made of cellulose rather than chitin. They also have swimming zoospores with two flagella (one whiplash and one tinsel), large nuclei, large egglike oospores, and various other nonfungal-like features. They are closely related to the brown algae (kelps) and diatoms, but of course lack chloroplasts. These organisms are now placed as heterokonts in the kingdom Stramenopila (or the catch-all dumping-ground kingdom Protista according to some authors). Ecologically many of its members act like fungi, especially the plant pathogens. *Phytophthora infestans* causes a disease called late blight of potato, which was the cause of the Irish potato famine in the 1840s, in which more than a million Irish people perished and another million emigrated. The downy mildews *Peronospora parasitica* and *Plasmopara viticola* cause diseases of members of the cabbage family and of grapes, respectively. *Pythium* species cause damping-off disease of seedlings in agricultural practice. However, many other species are innocuous saprophytes that decompose debris in water. A few of these so-called water molds (some species of *Saprolegnia* and *Achlya*) are opportunistic fish parasites, especially in aquariums and fish hatcheries (**Figures 11** and **12**). They are a particular problem in trout and salmon hatcheries.

Myxomycota: True Slime Molds

The Myxomycota are the true slime molds, also known as the plasmodial slime molds, and are considered members of the kingdom Protista or Amoebozoa, closely allied with other garden-variety amebae. They exist in nature as a plasmodium – a blob of protoplasm without cell walls and only a cell membrane to keep everything in (**Figure 13**). It is really just a large ameba and feeds much the same way, by engulfing its food (mostly bacteria) with pseudopodia, in a process called phagocytosis. So the slime mold ingests its food, then digests it. True fungi have a cell wall and digest their food with



Figure 11 *Saprolegnia* oogonium containing oospores.



Figure 12 *Saprolegnia* sporangium containing small zoospores.



Figure 13 Plasmodium of a slime mold.



Figure 14 Some slime mold fruiting structures (a) *Lycogala epidendrum*, (b) *Comatricha typhoides*, (c) *Badhamia utricularia*, (d) *Dictydium cancellation*.

exoenzymes before ingesting it. When the plasmodium runs out of food, it can form fruiting bodies. Most slime mold fruiting bodies are quite small, 1–4 mm in height (**Figure 14**), but some can be up to 25 cm in diameter. In such cases a large plasmodium, a meter or more diameter, may be seen crawling along the ground, inspiring science fiction movies such as "The Blob". These slime molds have been traditionally studied by mycologists because their small, delicate fruiting bodies tend to be fungal in appearance. Most slime mold fruiting bodies are quite beautiful, especially under a dissecting microscope.

Dictyosteliomycota: Cellular Slime Molds

The Dictyosteliomycota are the cellular slime molds or "social amebae" and are among the most bizarre of microorganisms (Raper, 1984). These members of the Protista or Amoebozoa are free-living amebae with no cell walls, indistinguishable from garden-variety amebae until they begin to run out of food. At that point they signal to one another using cyclic AMP (a small nucleotide molecule) and begin to aggregate to form a "slug" or pseudoplasmodium. This slug, composed of hundreds of amebae, acts as a single organism (although it does not feed) and can actually migrate along a light or temperature gradient. Eventually the slug stops migrating, rounds up, and forms a sorus, a kind of sporangium containing spores on a stalk. Not all of the cells become spores; some of them "sacrifice" themselves to become stalk cells to raise the spores up into the air for a better position for wind dispersal. So formerly free-living organisms act like a single organism for the good of the species. This "altruism" seems to fly in the face of the selfish gene hypothesis. Very strange indeed.

Plasmodiophoromycota: Endoparasitic Slime Molds

The Plasmodiophoromycota are an odd group of endoparasites that live almost their entire life cycle inside a cell of another organism. They lack cell walls in the assimilative state and reproduce by means of swimming spores (with two unequal whiplash flagella), the only part of their life cycle that does not occur inside a cell. *Plasmodiophora brassicae* causes club root in crucifers, and *Spongospora subterranea* causes powdery scab of potatoes. Scientists really do not know where to place these organisms taxonomically; they have been allied with the Oomycota, the Myxomycota, or various other protists, but in reality they are not closely related to any other known group of organisms. (Castlebury and Domier, 1998) No one is even yet sure what kingdom to place them in!

Surface Area and Reproduction

Many fungi have very specialized habitats. For example, the basidiomycete *Suillus americanus* can grow only in association with the eastern white pine, *Pinus strobus*. To infect a new white pine, the immobile mycelium of the fungus must produce spores to move to a new host. These spores are carried by the wind, not by some specific insect or other animal vector. Thus the fungus must produce enormous numbers of spores so that



Figure 15 Different ways of increasing the surface area for bearing spores. Clockwise from upper left: pores, upright branches (coral), teeth, and gills.

a few of them will land on the "correct" substrate. The underside of the *Suillus* mushroom has small pores that are lined with microscopic spores. This increases the surface area for bearing spores by more than 100 times. More commonly, mushrooms have gills underneath the cap to increase surface area. Other fungi increase their surface area by forming upright coral-like branches, whereas others form downward-pointing teeth or spines (**Figure 15**). Others (not shown) have blunt ridges and some have folds, and some increase the time of sporulation rather than the spore-bearing physical space. Analogous surface area modifications are made by members of other phyla of fungi.

Fungi in the Environment

Fungi occupy many different niches in the environment, although any one species usually occupies only a single niche. Many niches have not yet been explored for fungi, which is the major reason why mycologists believe there are high numbers of fungi yet to be discovered. Fungi can be divided into groups based on their nutritional status and the nature of their relationship with their host. Saprophytes use nonliving organic material and are important scavengers in ecosystems. Along with bacteria, fungi are important in recycling carbon, nitrogen, and essential mineral nutrients. Parasites use organic material from living organisms, harming them in some way. Their hosts range from single-celled diatoms to other fungi, plants, animals, and humans. Fungi are the major parasites (pathogens) of plants. Mutualists are fungi that have a reciprocally beneficial relationship with other living organisms, in which both organisms benefit. The three main types of mutualistic associations involving fungi are mycorrhizae, which are associations of fungi with plants roots, lichens, associations with algae or cyanobacteria and endophytes, associations inside leaves, stems, and roots of plants. There are also a few commensal fungi that use other organisms as merely a place to live; these fungi derive no nutrition from their host.

Fungi as Saprophytes

Along with bacteria, fungi are the major decomposers and recyclers in the environment. For every sort of dead material present, there is usually at least one fungus that can degrade that material. A few exceptions include some pesticides and some types of plastics; no fungi have yet developed exoenzymes capable of digesting these synthetic materials, although Gusse *et al.* (2006) have shown that biodegradation of some plastics is possible. Fungi are important in breaking down carbon- and nitrogen-containing compounds into

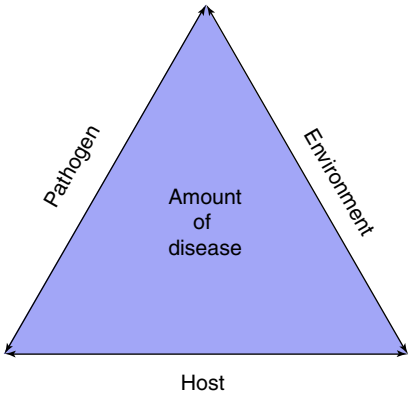


Figure 16 The disease triangle, showing the three factors that determine the disease severity.

Table 2 Types of plant diseases

Type of disease	*Symptoms*
Blight	Extensive necrosis and rapid death
Dieback	Dead branches protruding from apparently healthy tissue
Lesion	Localized cell death
Canker	Lesion on a woody stem
Rot	Extensive decay of tissue
Vascular wilt	Blockage of the vascular transport system, resulting in starvation or desiccation
Gall	Localized tumor like growth
Stunting	Reduction in overall growth size
Rusts	Rust-colored lesions on various parts of the plant
Smuts	Drastic malformations, usually of reproductive parts

Table 3 Some important forest pathogens

Pathogen	*Disease*
Cryphonectria parasitica	Chestnut blight
Ophiostoma ulmi	Dutch elm disease
Discula destructiva	Dogwood anthracnose
Cronaritum ribicola	White pine blister rust
Ceratocystis fagacearum	Oak wilt
Armillaria spp., especially *A. mellea* and *A. solidipes*	*Armillaria* root rot

components that they and other organisms can use. Fungi are especially important in the breakdown of the wood components cellulose and lignin (discussed in the section Fermentation Products for Food Use).

Fungi as Plant Parasites

About 90% of plant diseases are caused by fungi, resulting in billions of dollars in crop and forest losses each year (Farr *et al.*, 1989). A parasite is referred to as a pathogen if it elicits a recognizable and persistent disease. The most severe pathogens are those that are imported accidentally from other continents and that cause severe problems with the native populations and cultivated plants. The severity of any plant disease is determined by three factors, known as the disease triangle, consisting of the interaction of the host (conditions favoring susceptibility), pathogens (conditions affecting virulence), and the environment (conditions favoring disease) (**Figure 16**). All three factors must work in conjunction to produce a disease and determine its harshness.

Fungi in Crop Pathology

The science of plant pathology deals mostly with organisms, especially fungi, that cause plant disease. Plant diseases are generally classified on the basis of what kinds of symptoms occur in which part(s) of the plant (**Table 2**). Some important fungal pathogens of crop plants include *Puccinia graminis* (black stem rust of wheat), *Erysiphe graminis* (powdery mildew of grasses), *Claviceps purpurea* (ergot), and *Ustilago maydis* (corn smut).

Fungi in Forest Pathology

There have been a number of fungal diseases of forest trees that have caused great problems in North America (**Table 3**). Forest pathogens often need a longer time to kill their larger, perennial hosts, sometimes living inside the tree for several decades before the host finally succumbs.

Cryphonectria parasitica, *Ophiostoma ulmi*, and *Discula destructiva* are introduced pathogens that have devastated their hosts in North America. In their native Asia, these pathogens coevolved with their hosts and caused them only minor annoyance. However, when they were accidentally introduced into North America, either on live plants or on logs for sawing, the fungus escaped and subsequently devastated the North American tree species because there was not enough time for the host to develop resistance. *Cronartium ribicola*, *Ceratocystis fagacearum*, and *Armillaria* species are native American fungi, but they cause severe diseases nonetheless. A good pathogen does not kill its host right away, but keeps it alive as long as possible to continue deriving nutrients from it.

Fungi as Animal Parasites and Other Harmful Relationships

Fungi Affect Humans

Fungal effects on humans are classified by whether they can grow on the body (mycoses) or whether they cause problems by their ingestion. Mycetismus refers to the eating of poisonous mushrooms, while mycotoxicosis refers to the ingestion of toxins produced by the fungus, not involving eating the fungus itself. Mycoallergies, allergic reactions to fungi and their metabolites, are now defined enough that they should be considered a separate category.

Mycosis: Fungi Growing Directly on Human Tissues

There are a number of fungi that can grow directly on human tissue. Most of them require some debilitation of host defenses, especially the immune system. Fungal diseases have gained in importance over the last few decades because of increases in the numbers of such patients, including those with AIDS, people on corticosteroid therapy or immunosuppressive drugs (such as the transplant drug cyclosporine),

Table 4 Mycoses can be characterized by their location on the body

Type of mycosis	*Location*	*Some diseases*	*Some fungi causing disease*
Superficial mycoses	Infections of the hair shaft or dead outer layer of skin (stratum corneum)	Pityriasis versicolor, tinea nigra palmaris, piedra	*Malassezia furfur, Trichosporon beigelii, Cladosporium, Piedraia hortae*
Cutaneous mycoses	Dermatophytes – infections of skin, hair, and nails	Ringworm, jock itch, athlete's foot, tinea corporis, tinea capitis, and many others	*Microsporum Trichophyton*, and *Epidermophyton* spp., occasionally a *Candida* sp.
Subcutaneous mycoses	Chronic localized infections of the skin and subcutaneous tissues	Sporotrichosis, chromoblastomycosis, phaeohyphomycosis	*Sporothrix schenckii, Phialophora, Cladosporium*
Systemic mycoses (deep mycoses), true human pathogens	Fungal infections of the body caused by dimorphic fungal pathogens, usually entering through the lungs	Histoplasmosis Blastomycosis Coccidioidomycosis Paracoccidioidomycosis	*Histoplasma capsulatum* *Blastomyces dermatitidis* *Coccidioides immitis* *Paracoccidioides brasiliensis*
Systemic mycoses, opportunistic pathogens	Fungal infections of the body caused by common fungi becoming pathogenic owing to patient debilitation	Aspergillosis Cryptococcosis Candidiasis Zygomycosis *Pneumocystis* pneumonia Penicilliosis	*Aspergillus* spp. *Cryptococcus neoformans* *Candida* spp. *Mucor, Rhizopus* *Pneumocystis carinii* *Penicillium marneffei*

diabetics, and people undergoing chemotherapy treatment for cancers.

About 175 human pathogens are recognized among the approximately 75,000 known species of fungi. Around 20 are regularly isolated from cutaneous infections (dermatophytes and yeasts), a dozen are associated with severe subcutaneous localized disease, and about 20 may cause systemic infections (**Table 4**). In addition, there are many opportunistic pathogens that cause disease in debilitated or immunosuppressed patients.

Probably the most common fungal infection in North America is yeast infection, caused by *Candida albicans*. Three out of every four women will get a yeast infection sometime in her lifetime. This yeast can grow on outside portions of the skin or in any area of the body that is moist and warm. It is especially prevalent in the genital area of women. Oddly enough *Candida* is part of the normal flora of the body and can be isolated from almost everyone at any time. Only when conditions get out of balance does the yeast flourish and cause disease.

However, the systemic true pathogens are the most feared, since they do not require a debilitation of the host defenses to become pathogenic. These dimorphic fungi get around host defenses by changing their form from a mycelium to a yeast upon change in temperature, thus evading even the healthiest of immune systems. Fortunately they are all geographically restricted. (see also Rippon, 1988).

Mycetismus: Eating Poisonous Mushrooms

Of the 75,000 species of fungi, there are about 250 species of delicious edible mushrooms and about 25 species that will kill you. In addition there are perhaps 200 species that will make you wish you were dead (Volk and Rippon 2001). Most of the other species are innocuous or unpleasant tasting. Some people call poisonous mushrooms "toadstools," a word probably derived from the old German *Tödstuhle,* meaning "death chair." But how can you tell the difference? There is no easy way to tell if a mushroom is poisonous. Old wives' or old husbands' tales about silver spoons or whether it grows on wood or whether animals eat the mushrooms do not work! The only way to be sure if a mushroom is edible is to identify it to species by consulting books or knowledgeable persons (Barron, 1999; Bessette *et al*., 1997; Lincoff, 1981; Gilbertson and Ryvarden 1987). The old saying is "there are old mushroom hunters, and there are bold mushrooms hunters, but there are no old, bold mushroom hunters." The best way to learn about edible mushrooms is to join a local mushroom club. Do an internet web search for the North American Mycological Association for a list of these clubs. (Note: The pictures and descriptions in this article are NOT adequate to identify mushrooms for eating. We have a report of a woman who nearly died from eating wild mushrooms that she identified only from an encyclopedia drawing. Don't let this be you! Be absolutely sure of your identifications!)

As shown in **Table 5**, unrelated species in various genera may contain the same toxin, and one or more toxin groups may be found in the same genus (Volk and Rippon, 2001). To complicate things further, many genera contain both poisonous and nonpoisonous species, although several genera are more or less homogeneous with respect to poisoning ability. Most species of fungi have not yet been analyzed for toxins (Benjamin, 1995). Evidence suggests that most mushroom toxins have not evolved for protection. For example, what good does it do a fungus such as *Amanita bisporigera*, known as the death angel, if an animal eats the mushroom and then dies two days later? The toxins appear to be merely waste products, usually from nitrogen metabolism, that happen to be poisonous to animals or humans.

Table 5 Seven major classes of mushroom toxins

Class of toxin	*Fungi where found*	*Main symptoms*	*Mode of action*
Cyclopeptides (amatoxins and phallotoxins)	*Amanita bisporigera* (Death angel), *A. phalloides* (destroying angel), *A. verna*, *A. ocreata*, *Galerina autumnalis*, and a few others	Violent cramps, diarrhea, nausea, jaundice, coma, death; it normally takes about 12–36 h for symptoms to appear	Attacks RNA polymerase in the liver, eventually destroys liver
Monomethyl hydrazine (gyromitrin)	*Gyromitra* species (false morels)	Nausea, vomiting, diarrhea, cramps, jaundice, convulsions, coma, death; cancers produced in mice	Similar in structure to solid rocket fuel, destroys red blood cells, attacks central nervous system
Coprine (tippler's bane)	*Coprinus atramentarius*	Hot, flushed, metallic taste, palpitation, vertigo, vomiting headache like a bad hangover	Blocks alcohol dehydrogenase so alcohol cannot be completely metabolized; similar in structure and function to Antabuse
Muscarine	*Inocybe* spp., *Clitocybe dealbata* (the "sweater")	Induces PSL (perspiration, salivation, lacrimation) symptoms, cramps, blurred vision; contraction of pupils, hypotension	Anticholinergic – antagonist to action of parasympathetic nerve fibers
Ibotenic acid, muscimol	*Amanita muscaria* and *A. pantherina*	Results in "expanded perception," talking to God, macropsia (perceiving objects as enlarged), rapid heartbeat, dry mouth	Hallucinogenic, psychoactive, action on nervous system, acts as a neuropeptide receptor
Psilocybin, psilocin, psychedelic mushrooms	Many *Psilocybe* spp., some *Panaeolus* spp.	Trancelike state induced, hallucinations	Serotonin (neurotransmitter) analog
Gastrointestinal irritants	Many different kinds, found in hundreds of species	Nausea, vomiting, no effect on other organs, most just cause discomfort or distress to varying degrees	Various, depending on the species of mushroom

Mycotoxicosis: Eating Toxins Produced by Fungi, not Necessarily from Eating the Fungi Themselves

Mycotoxins are usually produced by molds growing on foods. As the molds grow, they metabolize the food product and excrete their waste products back into the substrate. Some of these waste products are highly toxic, and when a person eats the toxin (even after cutting the mold off), there is usually some harm to the person. For example, the common contaminants *Aspergillus flavus* and *A. parasiticus* can produce aflatoxins, especially in peanuts. If you have ever eaten a whole peanut that tastes so bitter that you spat it out, it was probably contaminated with aflatoxin. This compound is highly carcinogenic at about 1 part per billion, and kills very rapidly at higher concentrations. There are legal limits for the amount of aflatoxin allowed in peanut butter sold in the United States. Many other toxins are produced by molds, especially in the genera *Aspergillus*, *Penicillium*, and *Fusarium*.

Mycoallergies, Allergic Reactions to Fungi and Their Metabolites

Allergies to fungi have been an increasingly recognized problem for many people. These allergies usually involve hypersensitivity to inhalation of the spores. One often hears about the "mold count" on a summer day. The main group of species that cause inhalation allergy problems are from the genus *Alternaria*, which has allergenic proteins on its surface.

Other kinds of fungal spores may germinate in the sinuses and secrete their waste products, which become a source of allergic reactions for other people. Allergic fungal sinusitis is now being successfully treated with antifungal drugs. Previous treatments suppressed the immune system, which allowed the fungi to grow better and produce more allergenic metabolites.

Sick Building Syndrome (SBS) often makes news headlines. People find molds growing in the homes, schools, and workplaces, where many people seem to develop symptoms such as headaches, nausea, and fatigue. The species that gets most of the blame for SBS is *Stachybotrys chartarum*, the so-called "black mold" (although there are dozens of common species of molds that are black). However there is yet no scientific evidence that *Stachybotrys* causes the symptoms that people describe. Further experimentation awaits.

Fungi Affect Other Animals

Nonhuman animals are affected by fungi in much the same way that humans are, with some differences in certain species. For example, dogs are susceptible to cyclopeptide poisoning while cats apparently are not. There are numerous other examples of this discrepancy – so do not pick and eat a mushroom just because you see an animal eat it. Another important point to consider is that you do not know what happened to that squirrel after it ate that mushroom.



Figure 17 Ectomycorrhizae on pine roots.

Animals are also affected by many of the same fungal diseases as are humans. For example, dogs are particularly susceptible to blastomycosis and are often used as a warning sentinel, like a canary in a mine, for alerting humans to possible risk.

Fungi as Mutualists with Other Organisms

Rather than being harmful, some fungi benefit their host in some way while receiving nutrients from them. The three most common mutualistic associations are mycorrhizae, an association between a fungus and the roots of a plant, lichens, associations between a fungus and either an alga or a cyanobacterium or both, and endophytes, associations of fungi inside the leaves, roots, and stems.

Mycorrhizae

According to Harley and Smith (1983), a mycorrhiza is defined as "an association between a fungus and a host plant in which destructive disintegration of the host does not occur and which is a prevalent and usual condition of the host plant in natural habitats and as such is very common and widespread." *Myco*, of course, means "fungus," and *rhiza* is "root," so mycorrhiza literally means "fungus root." More than 90% of plants in nature have a mycorrhizal symbiont. The only groups of plants that regularly lack mycorrhizae are some crucifers, sedges, and some legumes.

There are several types of mycorrhizae. In ectomycorrhizae, the fungus, usually a Basidiomycota or sometimes an Ascomycota, forms a sheath outside the root (**Figure 17**). Exchange of nutrients takes place in transfer cells called a Hartig net, which penetrates between the cells of the root cortex, but does not penetrate the cells themselves. In endomycorrhizae, also called vesicular arbuscular mycorrhizae (VAM) or sometimes simply arbuscular mycorrhizae, no sheath is formed. This fungus is always a member of the Glomeromycota. Nutrient exchange takes place in highly branched hyphae called arbuscules, which penetrate into the cortical cells, but do not penetrate the cell membrane. Orchid mycorrhizae and ericoid mycorrhizae are special types that are found with plants in the Orchidaceae and the Ericaceae, respectively, and they differ significantly in their structure and life strategies.

In all types of mycorrhizae, hyphae extend from the root into the surrounding soil, greatly increasing the surface area for absorption of nutrients, particularly phosphate, nitrogen, and potassium. In return for shunting some of these nutrients into the plant, the fungus receives some sugars from plant photosynthesis. Thus both organisms benefit. Mycorrhizal fungi are abundantly represented in fossils from the Devonian and later periods and apparently coevolved with their hosts. It has been hypothesized that these fungi were necessary for the movement of water plants onto land.

As an interesting sidelight, there are also several hundred species of nonphotosynthetic plants (such as *Monotropa uniflora*, the Indian pipe) that get their energy as parasites of fungi that are mycorrhizal with photosynthetic plants. Radioactive carbon has been used to trace nutrient flow from the host plant through the mycorrhizal fungus and into the achlorophyllous plant. This type of ménage à trois is surprisingly common.

Endophytes

Only recently having been recognized for their importance, endophytes are microscopic fungi that live almost entirely within the leaves, roots, and stems of apparently healthy host plants. When they were discovered in leaves, they were assumed to be parasites, even though their presence caused no visible signs of infection. Although the hyphae are usually pervasive in the leaf, there is a lack of destruction of most cells and tissues. Further studies using radiolabeled compounds have shown that there was nutritional and chemical cycling between the host and fungus. Furthermore, both the fungus and the plant show increased vigor over plants without such relationships. In most cases that have been studied, the fungus receives sugars from the plant while producing bad-tasting toxins that deter herbivores from eating the plant. Thus the relationship should be classified as mutualistic, not parasitic. The best studied endophyte systems are in grasses and conifers (where endophytes apparently help the plants keep their needles longer), but endophytic fungi are probably found in most plants. Their significance in nature is greatly underappreciated.

Lichens

A lichen is a dual organism that consists of a mutualistic relationship between a fungus (the mycobiont) and an alga or cyanobacterium (the photobiont). Usually neither can survive on its own. Most of the fungi involved are Ascomycota, though a few are Basidiomycota. There are about 16,000 species of lichens, many of which can grow in very inhospitable environments – on rocks, sides and branches of trees, and gravestones, from the tropics to deserts to the Arctic (**Figure 18**). Lichens are very sensitive to air pollution, especially sulfur and nitrogen, and so they are natural indicators of

air quality. However, even under optimal conditions, lichens grow extremely slowly, usually 1–2 mm per year. The main ecological importance of lichens is their capacity to break down rocks into soil. They are an important food source for caribou and reindeer on the tundra. One of the first indications to the outside world that there had been a nuclear accident at Chernobyl, then in the Soviet Union, was the accumulation of radioactivity first in the lichens in Scandinavia, then in the milk given by the reindeer. Some lichens have been used as natural dyes, such as tweed and in litmus paper. In less enlightened times, a bright yellow lichen called *Letharia vulpina* was used by "pioneers" as a wolf poison.

Industrial Uses for Fungi

Mushroom Cultivation

The most widely available mushroom produced in the United States is the white button mushroom, *Agaricus bisporus* (**Table 6**). Commonly used on pizza and at salad bars, this mushroom is a secondary decomposer that is grown on composted cow or horse manure. Some brown forms of *A. bisporus* are currently being cultivated; brown buttons are sold as crimini (or baby bellas) and opened mushrooms are sold as portabella mushrooms.

There are a number of steps in the commercial production of *Agaricus bisporus* (Volk, 2011):

- Manure is placed in large concrete "runways."
- Composting (breakdown of the substrate into simpler components) occurs with the bacteria and fungi naturally present in the manure.
- Large machines turn the compost weekly, otherwise the center of the pile gets too hot because of metabolic heat and kills the composting bacteria and fungi. This also ensures even composting. The different species present naturally shift as composting progresses.
- After a couple weeks, the odorless compost is ready and is placed into large trays, 6 ft × 6 ft × 2 ft.
- Mycelium of *Agaricus* is inoculated into the composted substrate and allowed to grow for a few weeks.
- When the substrate is colonized, a sterile layer of nutrient-poor casing soil is placed over the top of the substrate.

A few weeks later, the mycelia send up rhizomorphs (hyphal aggregations) through the casing layer and form mushrooms on the surface. Contrary to popular belief, most mushrooms require light to initiate fruiting body formation. Light is often used as a signal to the mycelium that it is outside the substrate and that fruiting bodies can be formed. Some fungi form fruiting bodies in response to reaching outside air, where the concentration of carbon dioxide is lower. A few fungi fruit only when they run out of available nutrients. *Agaricus bisporus* has been bred so that it requires no light for fruiting.

There are a number of other specialty mushrooms, such as shiitake (**Figure 19**), oyster mushrooms, and lion's mane mushrooms, that are being grown throughout the world and that are just becoming available in North American supermarkets (Volk *et al.*, 1997). Almost all these newly available mushrooms are primary decomposers of wood or other cellulose-containing substrates. These specialty mushrooms have gained in popularity in the past 10–20 years as consumers discover that these specialty mushrooms have great flavor and



Figure 18 British soldier lichens, with their red caps for reproduction.

Table 6 Some commonly cultivated edible mushrooms

Common name	*Scientific name*
White button pizza mushroom, crimini, portabella	*Agaricus bisporus*
Shiitake, shiang-gu	*Lentinula edodes*
Oyster mushrooms	*Pleurotus ostreatus* and other species
Enoki, velvet stem, winter mushroom	*Flammulina velutipes*
Hen-of-the-woods, Maitake, sheepshead	*Grifola frondosa*
Cloud ear, wood ear, black mushroom	*Auricularia auricula,*
Chinese paddy straw mushroom	*Volvariella volvacea,*
Pom-pon or Lion's Mane	*Hericium erinaceus*
Reishi	*Ganoderma lucidum*
Wine cap	*Stropharia rugoso-annulata*
Morels	*Morchella* spp.



Figure 19 Shiitake mushrooms fruiting on an artificial log.

other interesting qualities. In Japan, China, and Korea, very little *Agaricus* is being grown (Stamets, 1993); growers concentrate on mushrooms with more robust flavors and interesting textures. Most of these specialty mushrooms are now grown on artificial sawdust logs by the following method:

- Place sawdust or wood chips with supplementary bran and millet and the appropriate amount of water into clear polypropylene bags. Autoclave or sterilize the filled bags, then allow to cool.
- Inoculate spawn (usually grain or sawdust with mycelium of the fungus growing on it) into medium. Mix thoroughly by hand or mechanically.
- Place bags in growth room. Allow spawn to grow rapidly through the substrate. Fungus colonizes bag in 30–60 days; sawdust is easier than solid wood to colonize because of increased surface area-to-volume ratio, abundance of air spaces, and uniform distribution of nutrients.
- During this period the loose medium is joined together into a coherent synthetic log.
- Depending on the species, at this time the plastic is removed, and the synthetic log can be handled like a natural log. Fruiting usually occurs within 90–120 days after inoculation.

Antibiotics and Other Drugs

Penicillin is the first antibiotic that was discovered to fight bacteria that cause human disease. It is naturally produced by *Penicillium chrysogenum* and related species as way of killing bacterial competitors in their environment. Cephalosporins are another class of antibiotics produced by *Acremonium* and related species. For a fungal infection of the fingernails and toenails, one prescribed drug is griseofulvin, produced by *Penicillium griseofulvum*. Ergotamine, produced by *Claviceps purpurea*, is used to facilitate the delivery of babies and can also be used to relieve migraine headaches. Another chemical found in *Claviceps* is a precursor to the hallucinogen LSD (lysergic acid diethylamide) that has the same effects as that illegal drug. The steroids in birth control pills are produced industrially by the fungus *Rhizopus nigricans*, as are the steroids cortisone and prednisone. People who have had an organ transplant usually take the antirejection drug cyclosporine, which is produced by the fungus *Tolypocladium inflatum* (= *Cordyceps subsellis*). The cholesterol lowering drugs, the statins, are all produced by various mold fungi. The pharmaceutical industry is constantly searching for new antibiotics and drugs to counteract microbes that become resistant to frequently used medications. Many fungi, such as *Lentinula edodes* (shiitake), *Grifola frondosa* (maitake), *Ganoderma lucidum* (ling-zhi), and *Ophiocordyceps sinensis* (caterpillar fungus), are used in Oriental herbal medicine to cure a wide variety of illnesses and to boost the immune system. Western medicine is just starting to examine these eastern remedies and finding that they probably do really work.

Wine and Beer Making

Yeasts, especially *Saccharomyces cerevisiae*, have played an important role in the development of human culture. Since brewing is an ancient art going back for millennia, yeasts have been our allies for many years. Virtually anything that contains simple sugars is fermentable by yeasts. In wine making, the grapes are pressed and yeast is added. Fermentation breaks the sugars down into carbon dioxide and ethyl alcohol. Since barley is the main carbohydrate source in beer making, there is an added step before fermentation, that of allowing the barley to germinate to produce sugar from its starch. Various other ingredients, such as hops, oatmeal, and flavorings, are added to make the wide variety of beers that are available. The aging process in alcoholic beverages also adds to their distinct flavors.

Bread Baking

Bread making also uses *Saccharomyces cerevisiae* to ferment sugars into carbon dioxide and ethyl alcohol. However, bakers are only interested in the carbon dioxide, which makes the bread rise. The alcohol evaporates rapidly on baking.

Ethanol Production for Fuel

Saccharomyces cerevisiae, the alcohol and carbon dioxide-making fungus used to brew beer and to make bread rise, can also be used to make fuel from simple carbohydrates. In the United States, about 10% of fuel is ethanol, while in Brazil, that number is close to 25%. Unlike petroleum, bioethanol is considered to be a renewable resource. The major problem with further success in the United States is that the use of corn as the substrate has led to an increase in the price of food made from corn. It would be more ideal to use cellulosic waste, such as that from sugar cane, corn cobs and stems, wheat and rice straw, or even wood. The major problem is that *S. cerevisiae* cannot ferment cellulose, and no one has yet found an organism that can do so on an industrial scale. Most of the technology development now focuses on using fungal enzymes to break down cellulose into simple carbohydrates that can be fermented by *Saccharomyces*.

Fermentation Products for Food Use

Fungi are often used in the large-scale fermentation of liquid or solid substrates. Fermentation vats can sometimes be several hundred thousand liters in size. For example, citric acid in cola drinks is produced by large-scale vat fermentation of the deuteromycete *Aspergillus niger*. Yeasts are sometimes grown in large fermenters and used directly as food supplements. Vitamin B_2 (riboflavin) in enriched flour is produced by the ascomycete *Ashbya gossypii*. Authentic soy sauce is fermented in a three-step process with the fungi *Aspergillus oryzae* and *Zygosaccharomyces rouxii*, as well as the bacterium *Pediococcus halophilus*. Tempeh, a soybean product popular in Indonesia, is partially fermented with a species of *Rhizopus*. Many good cheeses, such as blue cheese, camembert, and brie, are ripened through the action of fungi to obtain their distinctive flavors. Blue cheeses such as Roquefort, Gorgonzola, and Stilton are ripened by *Penicillium roquefortii* – the blue color is caused by sporulation of the fungus! The white crust on the outside of brie and camembert is the mycelium of *Penicillium*

camembertii. Fungi are also essential for chocolate: the naturally bitter cocoa beans are processed into a sweet tasty candy by a "fermentation" (sensu food-scientists) using *Candida krusei* and *Geotrichum*.

Biopulping and Bioremediation

Several important industrial process, which have the potential to be great boons to ecosystem health, are in the pilot stage: biopulping and bioremediation. The lignin-degrading enzyme system of *Phanerochaete chrysosporium* is special for these two uses. One of the biggest energy and pollution expenditures in paper making comes from removal of the brown lignin from wood so that the white cellulose is all that is left to make paper. What if paper companies could use the enzymes of a fungus to remove the lignin? This could result in a savings in both energy and time and avoid the polluting wastes that are commonly dumped out of the mills. This process is known as biopulping. There are several products in the pilot stage, but no large-scale biopulping is yet being done.

To understand this system, you must know that wood consists primarily of cellulose, which is white, and lignin, which is brown. Members of *Phanerochaete* and other genera cause a white rot of wood. That is, the fungus decays the lignin and leaves the cellulose behind. There are also fungi that cause a brown rot, digesting the cellulose and leaving the lignin behind. Many kinds of fungi cause a white rot, but *P. chrysosporium* has several features that might make it very useful. First of all, unlike some white rotters, it leaves the cellulose of the wood virtually untouched. Second, it has a high optimum temperature (about 40 °C), which means it can grow on wood chips in compost piles, which attain a high temperature. These characteristics point to some possible roles for this fungus in biotechnology applications. The basis for all of *Phanerocaete*'s powers come from its enzymes that can degrade lignin. This same enzyme system can also degrade chemicals that are chemically similar to lignin.

Some of the lignin-degrading enzymes of *P. chrysosporium* will also degrade toxic wastes, such as PCBs (polychlorinated biphenyls), PCPs (phencyclidines), TNT (trinitrotoluene), and similar chemicals. The structure of these chemicals is similar to that of lignin, and the ligninase enzymes will work on them. The fungus performs well on the laboratory bench, but as with many industrial bioprocesses, there are difficulties in scaling up the process. Nonetheless, this procedure has the potential to clean up some industrial and toxic waste sites.

In my lab, Adam Gusse, Paul Miller and Volk (2006) published a paper on the biodegradation of phenolic resin plastics with *Phanerochaete chrysosporium*. We demonstrated the ability of this white-rot fungus to degrade phenolic resin, a previously nonbiodegradable industrial polymer (plastic-like) that is not commercially recycled. These phenolic resin polymers, first known as Bakelite, are found gluing layers of plywood together, providing the binding matrix to particleboard, or laminating the surface of Formica™ counter tops. They are also used in constructing rotary telephone casings, bowling balls, toilet seats, motor casings, and many other everyday products. The annual production of phenolic resin is 2.2 million tons per year. Formerly considered to be nonbiodegradable and contributing greatly to landfills, our work has demonstrated, through three lines of independent evidence, that phenolic resin polymers can be degraded by *Phanerochaete chrysosporium*. (See TomVolkFungi.net for further information.)

Threatened or Endangered Fungi

Like the proverbial canary in the coal mine, fungi may be the first indicators of things going wrong in ecosystems. Remember that fungi are the threads that tie the whole food web together, since they are the primary decomposers and aid most plants as mycorrhizae in the absorption of minerals and water. Thus if something goes wrong with the fungi – if they disappear – there may be dire consequences for any plants and animals that are dependent on them. Lichens have already been proven to be accurate indicators of air quality, in both their quantity and diversity.

Several European countries maintain "red lists" of threatened or endangered fungi. One fungus that may be a candidate for the endangered species list in the United States is *Bridgeoporus nobilissimus*, a polypore fungus with a very large, perennial fruiting body (**Figure 20**). For a long time this fungus was in the *Guinness Book of World Records* as the largest known fruiting body of a fungus, at over 160 kg (300 pounds)! There are just eight known sites in Oregon, Washington, and California at which *B. nobilissimus* is now known to occur. It is considered by many to be a rare and probably endangered fungus. The main reason for this designation is that it is restricted to very large specimens of noble fir (*Abies procera*) and occasionally Pacific silver fir (*Abies amabilis*) with a diameter at breast height (dbh) of 1–2 m. Trees of this diameter are not very common. This fungus is considered endangered because its habitat is endangered. Unfortunately, it is not clear whether the U.S. Endangered Species Act applies to fungi. In this zoocentric world, most people are more interested in the "charismatic megafauna" than in some "lowly" fungus. However, the fungi have many important roles to play in the ecosystem and should not be ignored. No ecosystem could exist for long without fungi!



Figure 20 The author with *Bridgeoporus nobilissimus* on a large host tree of *Abies procera* (noble fir).

Two states have thus far recognized the importance of fungi by naming a state mushroom. Minnesota was the first, naming the morel (*Morchella* species) as their state mushroom. Some 10 years later, Oregon declared the Pacific golden chanterelle (*Cantharellus formosus*) as their state mushroom. Surprisingly, some forests can yield more income annually from wild mushroom harvesting than from lumber harvesting. Moreover, intact forests continue to produce edible mushrooms yearly, whereas lumber harvesting can occur only once every 50 years or so. Mushrooms for human food is only one of the many contributions that fungi make to our lives, for their fundamental decomposer role supports almost all ecosystems and so helps to provide the essential ecological services that we take for granted, as well as the recreational opportunities of enjoying nature. Fungi have contributed a great deal to our standard of living by making most of the living world possible.

See also: Biodiversity as a Commodity. Conservation Biology, Discipline of. Eukaryotes, Origin of. Forest Ecology. Parasitism. Restoration of Animal, Plant, and Microbial Diversity. Timber Industry

References

Alexopoulos CJ, Mims C, and Blackwell M (1996) *Introductory Mycology*. New York: John Wiley & Sons.

Barron G (1999) *Mushrooms of Ontario and Eastern Canada*. Edmonton, Alberta, Canada: Lone Pine Publishing (also published as Mushrooms of Northeast North America).

Benjamin DR (1995) *Mushrooms: Poisons and Panaceas*. New York: W. H. Freeman & Company.

Bessette A, Bessette AR, and Fischer D (1997) *Mushrooms of Northeastern North America*. Syracuse, NY: Syracuse University Press.

Castlebury LA and Domier LL (1998) Small subunit ribosomal RNA gene phylogeny of Plasmodiophora brassicae. *Mycologia* 90: 102–107.

Farr DF, Bills GF, Chamuris GP, and Rossman AY (1989) *Fungi on Plants and Plant Products in the United States*. St. Paul, Minnesota: APS Press.

Gilbertson R and Ryvarden L (1986–1987) *North American Polypores* 2 vols., Oslo, Norway: Fungiflora.

Gusse A, Miller P, and Volk T (2006) White-Rot Fungi Demonstrate First Biodegradation of Phenolic Resin. *Environmental Science & Technology* 40: 4196–4199.

Harley JL and Smith SE (1983) *Mycorrhizal Symbiosis*. London: Academic Press.

Hudler G (1998) *Magical Mushrooms, Mischievous Molds*. Princeton, NJ: Princeton University Press.

Kuo M (2005) *Morels*. Ann Arbor: University of Michigan Press.

Lincoff G (1981) *The Audubon Society Field Guide to North American Mushrooms*. New York: Alfred Knopf.

Raper KB (1984) *The Dictyostelids*. Princeton, NJ: Princeton University Press.

Rippon JW (1988) *Medical Mycology. The Pathogenic Fungi and the Pathogenic Actinomycetes*, (3rd edn.) Philadelphia: W. B. Saunders Company.

St-Germain G and Summerbell R (1996) *Identifying Filamentous Fungi: A Clinical Laboratory Handbook*. Belmont, CA: Star Publishing Company.

Stamets P (1993) *Growing Gourmet and Medicinal Mushrooms*. Berkeley, CA: Ten Speed Press.

Volk Thomas J and Rippon John W (2001) Eating mushrooms: Death, ecstasy or gourmet's delight. In: Misra JK and Bruce W (eds.) *Trichomycetes and Other Fungal Groups: Robert W. Lichtwardt Commemoration Volume*, pp. 365–383. Horn. Enfield, NH: Science Publishers. Chapter 17.

Volk TJ (1995–2011) Tom Volk's Fungi. http://TomVolkFungi.net

Volk TJ, Kozak ME, and Krawczyk J (1997) Ecological guides to the cultivation of edible mushrooms. *Mushroom News* 45(5): 26–36.

Weber NS (1988) *A Morel Hunter's Companion: A Guide to the True and False Morels of Michigan*. Lansing, Michigan: Two Peninsula Press.

G

Gene Banks

Simon H Linington and Hugh W Pritchard, Royal Botanic Gardens, London, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 165–181, © 2001, Elsevier Inc.

Glossary

Accession Usually a sample (e.g., seed lot) but may be a set of genetically related samples.
Collection One or more accessions.
Gene pool The genetic diversity contained within a population, species, or crop. The primary gene pool of a crop represents the biological species, the secondary gene pool includes species that can be crossed with it, allowing at least some transfer of genes, and the tertiary gene pool includes related wild species where such gene transfer involves specialized techniques.
Genotype The genetic makeup of an individual.
Seed lot A sample of seeds with a common harvest and post-harvest history.
Wild or nondomesticated species Those species that have not been brought into regular cultivation. Many may, however, have well-known uses.

A gene bank is an *ex situ* (or offsite) collection of genetic material held for long-term security or for ease of access. The genetic material is usually in the form of live propagules or gametes though, increasingly, pure DNA is held. A bank can also be a collection of full-grown plants representing the diversity of certain species (a field gene bank) or a set of cultures (cell, tissue, embryo, or microorganism). The widest definition might also embrace collections of plants grown primarily for purposes of education and horticulture in botanic gardens and animals represented in zoological gardens or parks.

Role Within Conservation Strategies

Ex situ (or off site) conservation underwrites conservation measures *in situ* (on site or in the natural state) and provides centers from which genetic material (germplasm) can be accessed for research, breeding, or (re-) introduction. Nearly any *ex situ* collection might be included under the loose definition of a "gene bank." As with a financial bank, deposits are made with the intention of later withdrawal. Furthermore, while in the gene bank, safety and integrity will be paramount and value will be added through study. However, the emphasis placed on withdrawal varies considerably between gene banks. For some collections, short-term utilization is the sole motivation, though inevitably conservation will be served to some extent. In others, long-term insurance against loss *in situ* is the stronger driving force. Given the increasing risks of genetic erosion (i.e., a narrowing of the gene pool leading to species being condemned to extinction) as a result of human development pressures and the not inconsiderable difficulties of managing *in situ* populations, the likelihood is that many such collections will be called on in the future.

While *in situ* conservation should be seen as the primary goal of all conservationists, having a second line of defense is a prudent measure particularly if it assists in other ways. For example, by acting as a supply source, *ex situ* collections reduce the pressure of repeated sampling of wild populations. Importantly, through the use of such collections, access to national genetic resources can also be controlled. Furthermore, because research is facilitated by such collections, knowledge about both the conservation biology and socioeconomic value of wild populations is improved, thereby increasing the likelihood of survival *in situ*. It is sometimes asserted that *ex situ* conservation of certain species reduces political will to conserve the habitats in which they occur. Should this be true, it would seem essential to ensure against the loss of, at the very least, flagship species within the habitats so that they can be reintroduced. Additionally, it should be apparent that habitat complexity may limit reconstruction opportunities from *ex situ* collections. In summary and as often stated, *ex situ* conservation is complementary to that carried out *in situ* (see also the entry for "*Ex situ*, *in situ* conservation").

The main types of animal *ex situ* collections are zoological gardens and parks, sperm and ova banks, and DNA banks. Due to controlled breeding programs, genetic resource databanks also play a more prominent role in the management of such collections than is perhaps the case with those for plants and particularly those in botanical garden collections. Microbe collections are mainly cultures though the possibility exists for fungal spore banks.

When considering plant species, the main types of *ex situ* collections are botanic gardens specimen collections (each normally comprising many species and few individuals per species); field gene banks (usually few species, many individuals); seed, pollen, and spore gene banks (potentially many species, many individuals); *in vitro* cultures (relatively few species, many individuals); and DNA banks (potentially many species, many individuals—though the latter are rarely held separately). The collections vary in the degree of technical input required, their ability to effectively conserve inter- and intraspecific variation, their longevity, and the ease with which gene products can be obtained. Seed gene banks (in this entry abbreviated to "seed banks") nicely balance these factors and offer a very effective form of *ex situ* conservation for many plant species. For this reason and the fact that most references to gene banks relate to seed banks, this chapter will firstly consider this type of gene banking and in particular the technology.

Seed Banks

Seed Banks are an Effective Form of *Ex Situ* Conservation

What makes seed banks such an effective *ex situ* conservation technique is that the methodology can be applied to a wide range of species in a universal and straightforward way and that large amounts of intraspecific diversity can be conserved and for long periods of time without intervention.

The fecundity of most plants means that a small seed sample can be removed from the annual seed rain with limited effect either on that population's survival or on the seed yield. One seed lot, if carefully sampled, can hold the majority of alleles found within that population. Moreover, if carefully germinated and then grown up under the same conditions, a seed lot has the potential to recreate the original population. In the meantime it can be held within a relatively small volume. For instance, a million tobacco seeds, each one with the potential to grow into a genetically distinct plant, occupy about a fifth of a liter.

Additionally, germinating seeds to obtain fully grown plants is relatively simple compared with obtaining plants from *in vitro* stored material. Plants recovered from banked seeds can also be compared with natural populations from which the material was harvested years before and which may have been subjected subsequently to environmental change (e.g., as a result of global warming).

The Historical Context

Seed banks might be seen as an invention of the 20th century, though the exploitation of crop seeds' ability to store in a dry state has almost certainly played a significant role in early human civilization. Early forms of seed storage probably included burying the seeds in ash or sealing them in adobe huts for the following season. By comparison, modern seed banks use controlled drying facilities and deep freeze stores and aim for storage lives of decades, centuries, or more. Modern seed banks have a lineage built on plant exploration and the development of botanic gardens that dates back nearly 5 millennia.

A collection of medicinal plants was established in China in 2800 B.C. Three hundred years later, the Sumarians were collecting vine germplasm from Asia Minor. During the Middle Ages, the Islamic world made great strides forward in the study of plants and their properties. There followed a proliferation of gardens of medicinal plants during the late Middle Ages and Renaissance of Europe. Acquisition and exchange of crop plants gathered pace with the colonial activities of the European powers. For instance, Columbus took wheat and other seeds to the New World (Plucknett *et al.*, 1987). Following such relocation of crops and their subsequent adaptation and diversification, the current patterns of food species use were established. Current collections of crop genetic resources reflect a strong degree of interdependence across the world on genetic diversity born out of this era of colonization. For example, one study suggests that North America now relies nearly totally on, and two-thirds of developing countries now obtain more than 50% of their crop production from, crops domesticated external to their region (FAO, 1996).

European colonization spurred on the development of many of the world's botanic gardens. For instance, the French established the first tropical botanic garden at Pamplemousses in 1735 to act as a staging post for germplasm movement and to screen for use. The Royal Botanic Gardens, Kew, was established in 1759 and in turn helped develop botanic gardens elsewhere in the world (e.g., Singapore). These subsequently acted as conduits for distribution of germplasm around the world. Many famous plant collectors date from this period of European colonial expansion and include Sir Joseph Banks, Richard Spruce, and George Forest (see Plucknett *et al.*, 1987). Not surprisingly, many botanic gardens now hold important collections of both endemic and exotic nondomesticated species including those known to be used locally for medicines and timber plus many ornamentals.

Changed agricultural practice and increased pressure on the natural world during the 20th century has led to increased erosion of genetic diversity both in crops and wild species. Currently, the most important causes of genetic erosion are considered to be replacement of local varieties, land clearance (e.g., fuel-wood collection), and overexploitation; some of the greatest concern relates to diversity found in the arid and semiarid regions (FAO, 1996). In 1949 there were about 10 thousand wheat varieties in China; in the 1970s the number had reduced to a thousand. Erosion of the genetic base of crops is not new. For instance cultivated carrot color was infinitely more diverse in Medieval Europe than it is now. However, the rate of erosion has accelerated in the 20th century with the drive toward greater genetic uniformity of crop varieties. Some highly uniform varieties have been sown across huge areas. In 1983 a single variety was sown across two-thirds of the wheat fields in Bangladesh and the following year across nearly a third of all those in India. This uniformity has productivity advantages within agricultural systems but is not without risk. For instance, there was a severe shortfall in winter wheat production in the Soviet Union when a variety grown over 15 million ha did not have sufficient cold tolerance to survive the winter of 1972. Less dramatic was the abandonment in 1975 of all of the United Kingdom's white clover varieties due to susceptibility to the disease *Sclerotinia trifoliorum* (FAO, 1996).

The more modern uniform varieties have been selected from genetically heterogeneous old varieties and landraces (primitive varieties) that they have then displaced. Therein lies the conundrum. For these new varieties to be developed further they need to draw on the genetic diversity that they have displaced. Half of the production increases of the Green Revolution were brought about by the use of plant genetic resources for plant breeding.

The Current Status of Seed Banks for Food and Agriculture

Although the outstanding Russian genetic resource scientist Nikolai Vavilov had assembled a collection of 50,000 cereal and pulses during the 1920s and 1930s, which were maintained by an annual grow-out, the first impetus to develop modern seed banks occurred in the United States. Here it was noted in the late 1940s that less than 10% of 160,000 plant accessions in the national system since 1898 could still be found (Plucknett *et al.*, 1987). Following the establishment of U.S. regional genetic resource facilities to address this problem, the National Seed Storage Laboratory (NSSL) was established at Fort Collins, Colorado, in 1958. With about 360,000 accessions, it and the Institute of Crop Germplasm in Beijing with about 316,000 accessions are the world's largest seed banks.

A significant proportion of the accessions held in crop genetic resource collections (including seed banks) date from the 1960s onward when the extent of loss of diverse landraces was becoming apparent. Catalyzed by the Food and Agriculture Organization of the United Nations (FAO) and subsequently by the International Board for Plant Genetic Resources (IBPGR, now the International Plant Genetic Resources Institute, IPGRI), a great amount of crop genetic diversity was saved in the 1970s and 1980s through the organization of collecting missions. Not surprisingly, conservation efforts have concentrated on the 30 crop species that provide 95% of dietary energy or protein. Of these species, rice, wheat, and maize provide 26, 23, and 7%, respectively. Consequently, about 40% of the estimated 6 million accessions conserved in the 1308 national and regional collections are cereals (of which 14% are wheat). A further 15% are food legumes. Other foods such as roots/tubers, vegetables, and fruits each comprise less than 10% of the total. Because many of the species in the collections are harvested for their seeds, it is not surprising that most accessions (about 90%) are held within seed banks. The remainder are held *in vitro* or within field gene banks (see later). Of the accessions for which there are data, perhaps half are advanced cultivars or breeders' lines and a third are landraces and old cultivars. The tertiary (plus much of the secondary) gene pool within wild and weedy species comprise only 15% of the accessions (FAO, 1996). The narrow range of species is indicated by the fact that forages and forest species are considered to have minimal coverage. There is no good estimate of the number of wild species held under internationally acceptable gene bank conditions. There are probably more than 10,000 species, but almost certainly less than 20,000 species, held. The National Plant Germplasm System in the United States accounts for some 8720 species (see ten Kate and Laird, 1999). This should be set against a background of about a quarter of a million scientifically described higher plants. Obviously, the *ex situ* conservation of wild plant diversity remains a major gap that needs to be filled within the genetic resource system.

The international institutes within the Consultative Group on International Agricultural Research (CGIAR) system account for 10% of the 6 million accessions. These institutes include the following (with number of accessions, most of which are held in seed banks, shown in brackets; see FAO, 1996):

- The Centro International de Mejoramiento de Mais y Trigo (CIMMYT) in Mexico (136,637)
- The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India (110,478)
- The International Centre for Agricultural Research in the Dry Areas (ICARDA) in Syria (109,029)
- The International Rice Research Institute (IRRI) in the Philippines (80,646)
- The Centro Internacional de Agricultura Tropical (CIAT) in Colombia (70,940).

Just under half of the 6 million accessions are held in 12 national collections. To some degree this is a function of the early establishment of their genetic resource collections. The collections include those in Russia and the United States (noted earlier), Japan (established in 1966), Germany (1970), Canada (1970), and Brazil (1974). Of the 1308 national or regional collections currently noted by FAO (1996), only 397 within 75 countries are held in long- or medium-term seed banks. Medium-term storage might be assumed to be in the order of 10 or more years.

Whilst 496 collections are within Europe and 293 are within Asia, some areas of the world have very few such facilities. FAO (1996) noted that there is only one long-term store within the Caribbean and this is in Cuba. Relatively few of the collections covered by the FAO survey are held by private companies, though many breeders hold smaller working collections.

The remits for crop banks are very varied. Some cover wide ranges of the crops that are important at the national or regional level, others cover very specific material. For instance, facilities at Horticulture Research International in Wellesbourne (United Kingdom) and the Asian Vegetable Research and Development Centre (Taiwan) have remits for vegetable germplasm.

While the science and technology underpinning these seed banks is generally well founded (discussed later), problems such as unreliable electricity supply mean that only the nine CGIAR banks, those in 35 countries and four regional centers, meet international standards as established by FAO and IPGRI (1994). To some extent this problem is obviated by the high level of duplication within the world's 6 million accessions. It is estimated that perhaps 1 to 2 million of these are genetically unique.

The way in which seed bank and associated genetic resource work is organized both at a regional and national level varies considerably. In Europe, the European Cooperative Programme for Crop Genetic Resources (ECP/GR) works through crop networks to coordinate effort. In Southern Africa, work is coordinated through a network set up through the Southern Africa Development Corporation (SADC). At a

national level, in India, genetic resources activities are formally centered around the National Bureau for Plant Genetic Resources. Within the United Kingdom, genetic resources activities are less centralized with responsibility residing with a number of institutes and their respective ministries. Advice to government is provided through the U.K. Plant Genetic Resources Group, which draws members from all of the relevant organizations.

Seed Banks for Wild Species

Increasingly, botanic gardens are adopting seed banking technology. Although FAO (1996) note that 150 botanic gardens (about 10% of the world total) have seed banks, relatively few of these have embraced international standards (discussed later). Those that have include the ones at the Berry Botanic Garden, Oregon (United States), the National Botanic Garden of Belgium, the Royal Botanic Gardens, Kew (Wakehurst Place), and the network of seed banks within the Spanish botanic gardens (notably Cordoba). Most botanic garden seed banks were born out of a need to more effectively manage the annual seed exchange. However, many now concentrate on the conservation of undomesticated species and interspecific variation. Additionally, several banks not in botanic gardens concentrate on wild species. These include the Threatened Flora Seed Centre established by the Department of Conservation and Land Management in Western Australia, the forest tree seed banks (e.g., the National Tree Seed Programme in Tanzania), and the seed bank of the Universidad Politecnica de Madrid. The latter bank was established in 1966 to conserve Crucifers and species from around the Mediterranean. Together with the Royal Botanic Gardens, Kew Seed Bank that has concentrated on wild species from the world's drylands, it has helped pioneer the application of seed bank technology for the conservation of wild species. Faced with the acceleration in wild plant genetic erosion and extinction, the Kew Seed Bank has taken on the ambitious objective of facilitating the conservation of seed from 10% of the higher plant species by the year 2010. Termed the Millennium Seed Bank Project, the work is a collaborative exercise involving botanical organizations from around the world. One of the aims of the Project is to maximize the efficiency of the banking work through an active seed research and technology transfer program. Other than the NSSL at Fort Collins, the project will have one of the largest seed research groups associated directly with a seed bank. Certain other botanic gardens, such as that in Rio de Janeiro, also carry out seed research work applicable to seed banking.

Community Seed Banks

A valuable role in seed banking is played by nongovernmental organizations (NGOs) at the community level. There is an increasing use of community seed banks within the developing world that involve donation, storage (usually on a short-term basis), and use of seeds at the local level. Examples include a set of community seed banks set up in North Eastern Brazil. The Plant Genetic Research Centre in Ethiopia is an example of a national program integrating with a community-based system (Heywood, 1995). Seed Savers in the United States and the Henry Doubleday Research Association in the United Kingdom are examples of NGOs active at the national level in the conservation of traditional vegetable varieties through their heritage seed libraries.

Collection Value

A financial value can be placed on biodiversity in one of two ways (see FAO (1996)):

- Direct value, or a willingness of the market to pay for it
- Indirect value, including its value as an insurance policy, the value to a product derived from the input of specific genetic material or the travel costs people would accept to collect it or to visit it for tourism

Obviously, there are other ways of considering the value of biodiversity such as its "service" value to biosystems and its value as a heritage to each human generation. However, debate has not surprisingly concentrated on its economic value and consequently the issue of access to the genetic wealth contained within seed banks is an increasingly controversial subject. Behind this debate is an attempt to reconcile the financial contribution of samples collected in one country to the economy of other countries. A fundamental concern that is often expressed is the degree of control over agricultural systems, particularly within the developing world, by multinational agrochemical/seed companies. Concern revolves around the use of the following:

- Uniform F1 hybrid varieties (where the seed, if harvested from the crop by the farmer, subsequently produces a heterogeneous crop of lower yield)
- Varieties that work best within systems when treated with certain agrochemicals
- The use of genetically modified organisms (GMOs)
- Plant breeders' rights and, increasingly, patents

Set against a backdrop of a $15 billion estimated annual turnover for the seed industry in Organization for Economic Cooperation and Development (OECD) countries, the debate has polarized on an "Industrial North–Developing South" divide. This has led to a useful reassessment of the sharing of benefits, though inevitably there has been some over expectation of the value of individual germplasm collections. Undoubtedly, there are examples of large financial contributions to the commercial world from plant germplasm. The value of genes for high sugar content from *Lycopersicon chmielewskii* have been estimated to contribute $5 to 8 million annually to the tomato industry (see FAO, 1996). However, it is worth bearing in mind that these are the exceptions to the rule. Relatively few advances in plant breeding have been spectacular and few have resulted from the incorporation of single genes. Most have occurred from the gradual improvement of characteristics governed by many genes (polygenic characters). Of course, advanced genetic manipulation techniques may change this picture and, indeed, may have an impact on the perceived value of traditional crop genetic resources.

In contrast to the seed industry, the pharmaceutical sector is estimated to be valued at $235 billion (FAO, 1996). Included within this turnover are plant-derived drugs, many of

which relate back to plants collected from developing countries. However, large financial returns from plant-derived drugs are unusual because major breakthroughs in the identification of new bioactive compounds from plants are infrequent. It is true that 34 out of the top 150 prescription drugs in the United States have a plant-based origin. This aside, despite the estimated tens of thousands of species investigated in bioprospecting programs, only 90 plant species have yielded 119 compounds considered to be important drugs (see ten Kate and Laird, 1999).

The debate has driven the establishment of agreements covering the access to and the sharing of benefits derived from plant genetic resources. Consequently, seed bank curation needs to accommodate not only national and international legislation but compliance with increasingly strict conditions attached to the material by the country or organization originally providing the material to the seed bank.

The Convention on Biological Diversity (CBD) came into force on December 29, 1993, following its opening for signature at the Rio Earth Summit in 1992. It has provided party governments with a framework within which access to genetic resources and the sharing of benefits arising from their utilization can be negotiated. Many countries have either developed, or are preparing national legislation to implement, access and benefit-sharing provisions of the CBD. Fundamental concepts that are often incorporated into national access legislation are as follows:

- The need to obtain *prior informed consent* for acquisition of genetic resources from the appropriate national authority
- The need to reach *mutually agreed terms*, which implies a negotiation leading to a form of bilateralmaterial transfer agreement specifying the terms and conditions under which material is transferred

In addition to the provisions of the CBD and implementing legislation at the national or regional level, certain seed bank collections may fall within the ambit of the International Undertaking (IU) on Plant Genetic Resources for Food and Agriculture (PGRFA). This is currently being revised in order to establish a *multilateral* access and benefit-sharing system for PGRFA that is compatible with regimes established under the CBD. It also recognizes the contribution made by indigenous farmers, in particular, to the development of plant varieties.

Seed banks, particularly the larger ones such as those within the CGIAR system (discussed previously), have responded to the emerging legal framework on access to genetic resources and benefit-sharing by reviewing the terms under which they acquire genetic resources and supply them to third parties. In many cases, seed banks have adopted the use of standard material transfer agreements that entitle the providing organization (or the country of origin of the material) to a share of any benefits arising from the utilization of the material. They may also restrict the subsequent transfer and use of the material. A pioneering worldwide project by 16 botanic gardens, which like the agricultural seed banks have traditionally supplied genetic resources to a wide range of users, has also attempted to find broadly standard terms under which material may be transferred to third parties.

Particular challenges faced by seed banks include the following:

- The acquisition of new material on terms that enable its subsequent distribution and use, so far as possible, on a common basis
- The development of a supply policy flexible enough to ensure that terms and conditions of any bilateral agreement (e.g., with the country of origin) are met, while enabling plant genetic resources to be also made available for food and agriculture under any multilateral system to be adopted in the revised IU
- The development of procedures that enable curators to clearly identify whether a proposal to acquire or supply material should be dealt with on a multilateral or bilateral basis. For instance, due to the single source nature of many successful pharmaceutical "hits," such work may fall into the latter category

Finally, it is important to note that the benefit of the annual investment in PGFRA conservation and utilization compared with its estimated total annual cost of $1 billion (FAO, 1996) is considered to be very high. For this reason it is essential that the future operations of seed banks are not threatened by financial insecurity.

The Scientific Principles Underlying Seed Banking

Viable seeds of many species when maintained in a dry and cold state are capable of being germinated many years later. This capability means that the long-term *ex situ* conservation of many higher plants is a realistic possibility.

Seed Storage Conditions

The science of seed storage is not a new one and dates back, at least, to China in the 6th century. Advances in the quantification of seed longevity under different storage conditions were made in the second half of the 20th century through the work of Harrington in the United States and Ellis and Roberts in the United Kingdom (see review in Hong, Jenkins, *et al.*, 1998). Critical factors that determine seed longevity are the seed's moisture content, temperature, and gaseous environment; its initial viability; and its genetic background. With respect to the latter, differences between species would appear to be much greater than those within. Genetics particularly influences the relationship between seed longevity and seed moisture content. Most species produce seeds that can be dried to low moisture contents (e.g., where less than 5% of the seeds' fresh weight is water) without loss of viability. The seeds of such species are termed "orthodox." Furthermore, orthodox seed longevity increases in a predictable fashion as the moisture content is reduced. Within limits, there is a straight-line relationship of negative slope between longevity and seed moisture content (up to about 15 to 25%) when both are expressed on a logarithmic scale. More simply, Harrington's rule of thumb states that a 1% reduction in moisture content roughly doubles seed longevity. Ellis *et al.* in 1989 found that drying seeds below a certain moisture content (equating to that in equilibrium with air at about 11% relative humidity at room temperature) did not increase longevity in most species

studied. However, in a few species, a decrease in longevity was noted.

Temperature also has a predictable effect on seed longevity over the range −13 °C to 90 °C (Dickie *et al.*, 1990) such that there is a quadratic relationship between longevity on a logarithmic scale and temperature—although an Arrhenius relationship also fitted the data well. Importantly, the relative effect of temperature on the seed longevity of eight species was identical. In its simplest form, Harrington's rule of thumb states that a 5 °C reduction in temperature doubles seed longevity.

In 1980, Ellis and Roberts developed a predictive model for orthodox seed storage that incorporated the determinants of longevity (except gaseous environment that has minimal effect at the low moisture contents over which the model is applicable). Based on acceleration of the aging process in seed lots by increasing their moisture content and temperature, the model allows extrapolation of longevity at cooler and drier conditions. For instance, rice seeds at 5% moisture content and −20 °C (fairly typical seed bank conditions) have a theoretical potential to survive at least 1900 years (see Hong *et al.*, 1998). Such data lend weight to occasional press reports of the germination of seeds of great antiquity, many of which cannot be confirmed due to the lack of supporting archaeological evidence. Nonetheless, there is, for example, evidence of longevity greater than 100 years at room temperature for dry wheat seeds (Steiner and Ruckenbauer, 1995). The seed storage model indicates the independence of temperature and moisture content on longevity over a wide range of conditions (Ellis *et al.*, 1995). It has been suggested that this may not necessarily be the case, particularly at very low moisture contents (Vertucci and Roos, 1990). The model also appears to be influenced by developmental stage (Hay *et al.*, 1997). These deliberations do not, however, alter the current seed bank storage recommendations (details of which are given later).

Even if these longevity predictions prove to be unattainable for all species, they do imply that, for some at least, storage in seed banks may be sufficiently long to carry the current heritage of genetic material to new eras when other technologies will be available. Predicted longevities of this magnitude also imply that the amount of regeneration of seed stocks inherent in many gene bank programs currently may be more a function of preferred working practice (e.g., regular study of collections), or poor application of the technology, rather than a failure of the technology itself.

In a study of the seed storage behavior of 6700 species, Hong *et al.* (1998), found that 91% had orthodox (desiccation tolerant) seed. The remainder had either desiccation susceptible or "recalcitrant" seed or had characteristics intermediate between the recalcitrant and orthodox types. Also certain genera (e.g., *Acer*) contain both orthodox and recalcitrant species. There is little evidence, however, that individual species produce both orthodox and recalcitrant mature seeds. It should, though, be noted that orthodox seed passes through a desiccation susceptible phase during development. The true percentage of species with orthodox seed within the plant kingdom is probably significantly less than 91% as the sample was biased toward species in trade. Although recalcitrant seed behavior is loosely associated with long-lived perennial species producing large fleshy fruits, there is great uncertainty of the number of wet tropical forest species producing such seeds. Seed storage behavior is usually assessed by quantifying the survival of seed lots of a species after drying to different moisture contents. A more rapid biochemical diagnosis of potential survival of drying is clearly of interest. Most recalcitrant seeds, including oaks and rubber, die below a value of about 40% moisture content. Because they need to be kept moist (and aerated) they have short life spans, generally of a few months, and are not suitable candidates for seed banking. Other forms of genetic conservation must be employed. The same is true for species, such as coffee, which produce "intermediate" seeds that cannot be dried much below about 10% moisture content. They also tend to be susceptible to storage at cold temperatures (−20 and 0 °C). Work on other species, such as orchids, suggests the use of other subzero temperatures may be appropriate for the conservation of species sometimes included in this storage category (Pritchard *et al.*, 1999).

Genetic Considerations

The genetics of seed bank storage is an important issue. A criticism occasionally leveled at seed banks is that there is selection in storage. Selection can occur during collection through a biased sample (e.g., for early or late flowering genotypes). It can also occur when samples are grown out under conditions that differ dramatically from those where the seeds were harvested. There is, however, little evidence to show that, compared with, say, room conditions, the seed bank environment does other than slow down the (normally distributed) times for individuals to die. In other words there is no greater risk of selection out of any individual by the conditions applied than under natural conditions.

Another criticism is that seed lots in a seed bank are protected from evolutionary pressures that maintain fitness. The extent of this problem will depend on the generation time of the species involved and the severity of the selection pressures (at the site from which it was collected) during the time the seed lot was in storage. It might be argued that those species with short generation times will quickly adapt to all but the most extremely changed environments (i.e., the gene frequencies in the seed lot will quickly be selected to meet the new needs). Of course, maintaining the original genes sampled from within the population may have longerterm benefits of fitness for the species if pathogen attack has altered the natural population in the intervening period. An additional consideration is the prospect of genetic contamination of natural populations by chance hydridization with exotic material including genetically modified organisms. This concern is also partially alleviated by the existence of long-term stored germplasm in its original state.

Seed Bank Management

Procedure

The basic elements of the seed banking procedure (more or less in order) are as follows:

- Collection planning and permission seeking
- Seed (and pressed specimen) collecting and field data recording

- Shipment of the seeds
- Creation of a data record about the accession
- Seed cleaning (sometimes preceded by initial drying and sometimes accompanied by X-ray analysis and quantity determination)
- Main drying
- Seed moisture determination
- Initial germination test (sometimes left until after banking)
- Packaging and banking and security duplication
- Characterization (including verification of identity in the case of wild species) and evaluation (where appropriate)
- Distribution of stocks to users (through time)
- Germination retests (through time)
- Regeneration/multiplication (as required)

Although many of the well-established seed banks broadly follow these procedures, most modify them to meet their own specific needs. Some banks carry out very little field collecting and some multiply the seed sample on arrival. Costs for each stage vary considerably between seed banks. The RBG Kew Seed Bank, which has an international collecting program, estimated, in 1997, that the ratio of costs between collecting, processing, and annual maintenance for a collection was in the order of 100:50:1.

Seed Collecting

Set against the background of the CBD and the IU, collecting should only be carried out with the permission of the national and local authorities. Foreign collectors should work collaboratively with local scientists and clear agreements on benefit sharing should be in place. One immediately tangible benefit is the sharing of collections and the information relating to them. Other international regulations need to be adhered to. These include the Convention on International Trade in Endangered Species (CITES) and national quarantine laws.

Seed collecting methodology and genetic resource exploration have been thoroughly covered by Guarino *et al.* (1995). In most instances, random and even sampling of wild plant or crop populations is recommended, including careful note taking of the sample method (in often less than perfect conditions). Objective data recording is essential as is accurate recording of location. This latter aspect is now facilitated by the use of Global Positioning System receivers that help fix latitude, longitude, and even altitude using satellites.

Harvesting seed that is close to maturity and keeping the seed alive in the field should be paramount. Also it is essential that harvested seed is returned to the seed bank facility as securely and rapidly as possible. Prestorage conditions influence seed viability at the start of storage. Because the longevity of a species stored under specific seed bank conditions is substantially fixed, the only other key longevity factor that can be varied is initial seed storage viability. Delay in drying the seed properly or unduly harsh cleaning methods can severely reduce initial viability; in turn, this can dramatically reduce both seed longevity and seed bank efficiency. Where space is not limiting within a cold store and the longer seed lots can be kept alive, the less the unit fuel cost per species.

It is worth noting that seed lots of the same species should not be mixed either within or between years. Such action could diminish both the unique genetic makeup and reduce the longevity (through lower initial seed quality) of a seed lot.

Seed Cleaning

Seed lots harvested from some populations of wild species in particular contain large numbers of "empty" (or aborted) and insect-infested seeds. They may be outwardly similar to the "competent" seed and often are thus not removed during the cleaning process. Their presence can be detected by X-ray analysis or a simple cut-test applied to a subsample. Once detected, such incompetent seeds need to be allowed for when distributing seed to users and when carrying out germination tests. Although the insects may not be removed at this stage, nearly all adult and most larval forms of insects are killed by standard seed bank practice. However, insect eggs along with fungal and viral contaminants can survive seed banking. Thus, appropriate quarantine procedures need to be in place when the seed is removed from store.

Seed Bank Storage Standards

No two banks are the same. Traditionally, seed banks have been classified into the following categories:

- *Base* collections that are for the long-term storage of seed lots and from which seeds are not normally sent to users (though this is not always the case).
- *Active* collections from which seeds are made available to users. Often such stores are maintained under less optimal storage conditions compared to those holding base collections though this need not be the case.FAO/IPGRI (1994) recommends that storage lives of 10 to 20 years might be appropriate.

Base and active collections can occur at the same or different genetic resources centers or, as is the case in the SADC genetic resources network, the active facilities in different countries are linked to a regional base facility. In theory, the collections held within the active store need occasionally to be refreshed by the same samples held in the base store. Because active stores are likely to more frequently regenerate (discussed later) their samples, selection and genetic drift could cause a genetic divergence from the same samples held in the base store.

The FAO/IPGRI standards (1994) for base collection storage requires fresh seed to be dried within the range of 3 to 7% moisture content, packaged in a moistureproof way, and placed at subzero (and preferably $-18\ ^{\circ}C$) temperatures. Although many of the world's seed banks aspire to such conditions, as has been noted, relatively few are able to emulate them. Nonetheless, the robust nature of seed bank technology means that substantial longevities are achievable even with partial fulfillment of the standards that have been discussed and using a low technology approach.

The standards set out in 1994 are important as they set the benchmark for standard operating procedures by seed banks. Consolidation of these standards and increased use of a common format for data would facilitate exchange of information between facilities. Such unification of data standards could ultimately embrace all specimen banks including those

specializing in the storage of samples for environmental monitoring.

Seed Drying

Drying is perhaps the key to success in seed banking of orthodox-seeded species. Not only is storage prolonged in orthodox-seeded species but germination is prevented, the risks of insect and mite attack are reduced, and the seeds are protected from freezing damage. FAO (1996) found good drying facilities to be a limiting factor in a number of seed banks. Drying involves the manipulation of the water potential gradient between the inside and the outside of the seed. To all intents, the water potential is the difference between the activity of water molecules in any system compared with pure water. When the water potential of the air is lower (more negative) than that of the seed, there is a net movement of water out of the seed and drying occurs. As the water potential gradient reduces, drying slows down. When no gradient exists, the seed moisture content is said to be in equilibrium with the surrounding air conditions. Consequently, factors that influence the water potential of the air surrounding the seed are important to drying. Lowering the relative humidity or increasing the temperature of the air lowers its water potential. By keeping air moving over the seed, low humidity can be maintained by preventing moist air from accumulating around the surface of the seed. Seed size and seed depth affect rate of drying. Migration of moisture from the center of a large seed to the outside will take longer than from the center of a small seed. The same principles apply to large and small seed sacks. Shape, affecting surface to volume ratios, and seed structure will also have an effect on drying rate as can seed maturity.

It is worth noting that at any particular water potential, the seed moisture content will depend on its chemical composition. Seeds with a higher oil content will have a lower moisture content.

Most seed banks are located where the ambient relative humidity is not sufficiently low to allow the seeds to dry to the moisture content levels set out in the previous standards. High temperature gives rapid drying. However, although shade and oven drying can be used for drying, there are seed aging dangers of placing wet seeds at high temperatures and of leaving dried seeds too long under such regimes. Similarly, sun drying has risks associated with radiant heat gain. A degree of caution needs to be exercised if these methods are used. Consequently, seeds are usually placed in a drying environment where the relative humidity has been artificially lowered. Most often, this is achieved by sorption or occasionally by refrigeration drying systems. At their simplest, sorption systems can consist of a closed container into which the wet seeds are placed and the air dried by a quantity of silica gel or dried rice. Such methods require a degree of experimentation to achieve the desired results. Greater control over the extent of drying can be achieved by use of a suitable saturated salt solution (e.g., that of lithium chloride) that will maintain a set relative humidity (about 11% at room temperature) within a closed container. However, many of the larger seed banks have now adopted controlled drying rooms that allow for large samples to be dried in thin layers. Air from the chamber is dried using dryers containing lithium chloride or silica gel, cooled, and then returned. In effect, this is a closed system though most facilities allow for some fresh air intake. Ducting the air to and from the dryer on opposite sides of the chamber encourages air movement. Conditions within the chamber are those necessary to achieve the desired moisture contents at equilibrium for long-term storage. Such conditions may be 10 to 15% relative humidity at 10 to 25 °C, and equilibrium is usually achieved in about one month.

Packaging

Maximizing orthodox seed longevity depends on drying and freezing. Cooling ambient air increases its relative humidity and hence its water potential. Placing the dried seeds directly inside a refrigerator will cause the seeds to absorb moisture. Therefore seed banks need to carefully package the dried samples. In 1996, IPGRI reviewed the types of container used in conventional long-term seed stores. Types of container used by banks include laminated aluminum foil bags, sealed steel cans, sealed glass tubes, screw-top glass bottles, and levertop fruit preserving jars. All containers have some limitations but, carefully managed, risks of moisture ingress can be minimized especially if some monitoring system is in place. The type of container chosen will partly depend on the frequency of access required. The main concern with all is that the seal through ingress of moisture through the fabric of the material over long periods is a risk where foil is punctured or cans have poor seams. Glass offers the advantage that the contents are visible but is of course at risk from breakage. A number of banks double-pack for added security and others add a desiccant. A number of facilities dry the air in the cold store and store the collections in paper or cloth bags. Consequently, these facilities must have adequate generator backup. This is less of a problem where seeds are held in moisture-proof containers and there is a loss of electrical power. In such circumstances there is little evidence to suggest that more than a few days storage life is lost per disruption.

Storage Temperature

Many long-term seed banks store seed under deep freeze conditions using either purpose-built prefabricated cold rooms or domestic deep freezers. To reduce staff time at subzero temperatures, a few seed bank cold rooms, such as one at the National Institute of Agrobiological Resources (NIAR), Japan, have mechanized banking/retrieval systems. The use of such systems have implications to energy consumption by the bank.

Use of permafrost has been considered for long-term duplicate storage of seed in places such as Svalbard. Although the dependence on electricity is cut, such stores are usually unable to match the lowering of temperatures possible in conventional base storage conditions. In 1997 the Japanese-based Biological and Environmental Specimen Time (BEST) Capsule 2001 Project discussed the possibilities of long-term storage of flagship samples under Antarctic ice at -58 °C (which incidentally is not sufficiently low for animal tissue preservation) or even on the dark side of the moon at -230 °C.

More usually, seed storage at ultra-low temperatures (cryopreservation) is achieved using liquid nitrogen. Seed samples are normally held in polypropylene (or similar) screw-cap containers placed in the vapor phase above liquid

nitrogen (about $-160\,^{\circ}C$). Preferably samples of dry seed should be cooled and rewarmed at a relatively slow rate (about $10\,^{\circ}C\ min^{-1}$) to reduce problems of rapid thermal contraction and expansion that can cause physical injury to the seeds, such as cracking of the embryo tissue. The largest cryogenic seed bank is operated by the National Seed Storage Laboratory (United States) where there are more than 37,000 accessions stored over liquid nitrogen. More than 11,000 of these are also stored under conventional seed bank conditions at $-18\,^{\circ}C$. The National Bureau for Plant Genetic Resources in New Delhi also conserves seeds under conventional and cryogenic storage conditions. Moreover, seeds of a significant number of North American and Australian wild species have been tested successfully for their tolerance of cryopreservation (Pence, 1991; Touchell and Dixon, 1993). It is worth bearing in mind when considering setting up a cryogenic facility for seed material that such low temperatures are not necessary to achieve practical periods of long-term storage (Pritchard, 1995), that the setup costs are relatively expensive, and the storage volume is less efficient than a conventional seed bank. However, these additional costs may be acceptable when creating the ultimate "base collection" for material that is in short supply, is inherently short-lived, or is of particular commercial value (unique genotypes).

Monitoring Seed Lot Viability

Perhaps one of the most important parameters of seed bank effectiveness is the result of germination monitoring. Germination is the preferred test for seed lot viability. Providing such information to those using the seed is helpful. Additionally, other viability tests such as vital staining using tetrazolium solution have a greater element of subjectivity about them. This staining test is, however, sometimes used to help distinguish between dead and dormant seeds among those that did not germinate under a given test regime. Two problems relate to the germination monitoring of seed bank accessions. First, because seeds are tested soon after arrival at the bank and then at regular intervals (often every 5 to 10 years) during their storage life, the tests need to be repeatable and operator independent. Second, in order to recover as many genotypes as possible represented within a seed lot, it is necessary to break seed dormancy. This can be a particular problem in seed of wild species and where the seed is freshly harvested. Key techniques include scarification of hard seed coats to facilitate water or oxygen permeation, imbibed chilling at 5 to $10\,^{\circ}C$, and later incubation at diurnal alternating temperatures with fluorescent light (i.e., rich in red light) provided only in the higher temperature phase. Tests have to be seed lot specific as most seed dormancy is not strongly genetically inherited and the form it takes depends on the conditions under which the seed matured. Once determined, the same treatments can be used during the monitoring of that seed lot through time.

Seed banks use a variety of media for germinating the seeds, such as filter paper and sterilized sand. However, 1% (w/v) plain water agar is increasingly popular especially as it reduces the risk of imbibition damage to dry seeds.

Obviously through time these monitoring tests can consume a significant proportion of each seed lot. Ideally, when the seed viability has fallen to a level where regeneration needs to take place, sufficient seeds should remain in the collection to make several attempts at growing out the collection. Because base collections may have a projected storage life of as much as 200 years, collection size needs to be large. The international standards (discussed earlier) recommend at least 1000 seeds per seed lot.

Where seed banks handle seed with unknown storage behavior, the initial germination could be delayed until after drying and banking. Survival will indicate orthodox characteristics. Few banks will have sufficient staff time to do more than brief tests prior to this on suspected recalcitrant or intermediate seeded species.

Duplication

One of the main advantages of seed banks is that they centralize collections of genetic material making them more easily accessed and studied. Indeed, some seed banks might be seen as some of the world's greatest plant diversity hot spots with more individuals and, in some banks, more species per square meter than anywhere else on the planet. This centralization poses a risk to all but the most carefully located and constructed facilities. Potential catastrophic loss, which of course threatens plants conserved both *ex situ* and *in situ*, mean that duplication of collections and their associated data is an important element of seed bank safety. The FAO report (1996) indicates that the level of security duplication of plant genetic resources for food and agriculture still needs to be improved and is at best uncertain.

Characterization and Evaluation

Characterization can vary from accurate naming of the species or subspecies represented by the collection through to more detailed recording of characters governed by genes that are little modified by environmental factors (major genes). Such information, published in the form of descriptor lists, is of great value to plant breeders wishing to narrow their choice of material from, often, vast collections. Similarly, the concept of core collections has been established to facilitate use by breeders. A core collection genetically represents a limited set of accessions of a crop gene pool with the minimum of repetition.

Increasingly, characterization is taking the form of more detailed molecular techniques such as screening by Amplified Fragment Length Polymorphism. By contrast to characterization, evaluation records data on traits such as yield that are strongly influenced by the environment in which the plants are grown. Such data are thus site and year specific and are perhaps of less use to breeders.

Distribution to Users

A very important element of seed bank work is to make the seed available wherever possible. In evidence of the scale of such dispatch, ten Kate and Laird (1999) quote an annual distribution of nearly 120,000 samples by the U.S. National Plant Germplasm System of which 65% are sent abroad, many requested by reference to the Germplasm Resources Information Network (GRIN) available on the Internet. Furthermore, the usage rate of crop banks by plant breeders is probably less than the rate of request from banks holding

broader plant diversity collections where uses include a wide array of pure and applied research in addition to field trials. During 1994–1996, there was a 50% request rate for seeds offered through an extensive list offered by the Kew Seed Bank.

Several elements need to be considered concerning the distribution of seed to users. The recipient should be provided with accurate information about the collection and how to germinate the seeds. It should also be remembered that certain species require a symbiont for effective growth (e.g., legumes and *Rhizobium*), and that the user may need to draw on germplasm for both plant and symbiont. The seed sample must be dispatched having fulfilled all necessary plant health and, where appropriate, CITES requirements. Finally, to meet the needs of the CBD and to clarify the conditions under which the material can be used, all germplasm samples are increasingly dispatched under material supply agreements.

Regeneration

Seed bank accessions are grown out for the purposes of regeneration of seed stock (either when seed numbers are low or when viability has reduced), for characterization and for evaluation. Many banks have a regeneration standard below which the germination of a seed lot should not fall. This is usually set at 85%. This high value limits the risk of accumulated genetic damage that is associated with seed aging. Even though falling levels of seed germination are correlated with falling levels of field establishment, many banks have adopted lower standards. This may in part be due to the backlog of regeneration work that in some national facilities highlighted by FAO (1996) is nearly 100% of the collection. By collecting high-quality seed lots in good quantity, other banks have reduced the necessity for regeneration that can be time and labor consuming and that can have adverse effects on the genetics of the collection. Samples regenerated under conditions different from where they originated can experience selection. If too small a sample is regenerated, genetic drift may occur in which rarer alleles are lost through chance. Under some circumstances, recollection, if possible, may the more desirable option.

Seed Bank Design

Having considered the aspects of seed bank management, a brief consideration of seed bank design is appropriate (also see Cromarty *et al.*, 1985). The location of the bank is important from political, practical, and security aspects. Potential risks have to be considered be they earthquake, flooding, or radiation fallout. Some facilities are placed underground such as the seed bank at Krasnodar in Russia and the Millennium Seed Bank in the United Kingdom. Others such as the NSSL are located on the first floor to limit possible impact from structures above resultant from seismic activity. The size of most banks should be dictated by peak annual intake (seed drying and cleaning facilities), projected capacity before a rebuild is practical (seed storage), and annual collection maintenance (germination, field, and greenhouse facilities). Cold storage facilities vary from a few domestic deep-freezers up to large rooms such as one of 140 m^2 (with capacity for 150,000 samples) at NIAR in Japan.

Other Types of Gene Bank

This section provides summaries of the current status of nonseed gene banks, starting with dry propagules (pollen and spores), which can be stored under conditions similar to those used for seeds, and covering normally hydrated tissues that can also be preserved under a different set of controlled conditions. Finally, field gene banks are covered and the role of botanical and zoological gardens is briefly mentioned.

Pollen

There are many practical reasons for storing pollen: to support work on allergenic responses, plant hybridization, and fertility; haploid plant production; and genetic transformation systems with isolated pollen (or gametes). Optimal storage conditions for pollen are similar to those for seeds (i.e., at low moisture content and subzero temperatures). In addition, there is some evidence (e.g., in maize and *Impatiens*) that longevity in dry storage is enhanced in anoxic atmospheres. Moreover, pollen has also been stored in a vacuum-dried state. The ease of storage though relates in part to the cellular and physiological nature of the pollen. Bicellular pollen, as found in Liliaceae, Orchidaceae, Solanaceae, and Rosaceae, generally tolerates desiccation to about 10% moisture content and is relatively long-lived. By comparison, tricellular pollen, as found in Graminae and Compositae, is relatively short-lived and is much more sensitive to desiccation. Most work on the long-term storage of pollen has focused on fruit tree or forestry species, for which 10 years storage at conventional gene bank temperatures ($-20\ ^{\circ}C$) is easily attainable. Pollen of at least 30 species are known to survive liquid nitrogen temperatures. Although pollen banking is evidently possible for many species, there does not appear to be any large-scale gene bank operation using such material.

Spores

Spores of many species of both pteridophytes and bryophytes are stable for months or years when dried, and this time can be extended with storage at cold or freezing temperatures. The longevity of short-lived (chlorophyllous) spores of some species can be extended significantly by drying and freezing in liquid nitrogen (Pence, 2000). Although there are data on fungal spore storage (e.g., work by Hong *et al.* in 1998), the majority of fungal germplasm appears to be conserved in culture or through cryopreservation of hyphae (see later).

Somatic and Zygotic Embryos of Plants with Nonbankable Seeds

Nonbankable seeds can nonetheless be stored using alternative approaches. Usually, these revolve around the use of rapid, partial desiccation of embryos or embryonic axes to about 20% moisture content and subsequent transferal to liquid nitrogen temperature or the use of other subzero storage temperatures. Recovery levels may be improved by pretreatment of embryos with cryoprotectants, encapsulation of the material in alginate beads, or careful manipulation of the

in vitro recovery environment. The embryos of more than 50 species have been successfully cryopreserved. Other parts of the plant can also be used to establish *ex situ* gene banks for species with nonbankable seeds, as described in the following section.

Vegetative Parts of Plants

Vegetative tissues of both pteridophytes and bryophytes can be banked for germplasm preservation (Pence, 2000). Gametophytes of many bryophyte species are naturally adapted to desiccation stress and can be cryopreserved after sufficient drying. Gametophytes of pteridophytes and some desiccation-intolerant bryophytes can also be frozen when provided with some cryoprotection, such as encapsulation in alginate beads followed by dehydration or the use of abscisic acid and the amino acid, proline, as a pre-treatment. Shoot tip freezing of sporophytes of pteridophytes has also been demonstrated. It is estimated that fewer than 200 taxa of bryophytes and pteridophytes combined are currently banked worldwide using vegetative tissues, primarily at the Cincinnati Zoo and Botanical Garden and the University of Kansas in the United States, but there is significant potential for increasing this number.

The National Seed Storage Laboratory of the U.S. Department of Agriculture (USDA) also cryopreserves about 1700 apple lines using dormant scion sections, which are retrieved by grafting (i.e., no culture of meristems). The lines are mainly from *Malus x domestica*, plus 10 to 15 other apple species. Some pear and cherry species (Towill and Forsline, 1999) are also banked in this way.

Apical shoot tips and other meristems/buds are the most popular vegetative materials for cryopreservation. Lines from about 50 species are now routinely cryopreserved. Initial studies used shoot tips from cold hardy, temperate zone species (apple, pear), but cryopreservation methods have been extended to tropical zone species (banana, pineapple). Successful cryopreservation depends on defining the physiological adaptation of the stock plant, the explant size and type, and its water content, the steps in the preservation process (cryoprotectant concentrations and rates of addition/removal; cooling/warming rates), and the recovery system. Two-step cooling procedures have been useful for some species, but vitrification procedures (solution-based systems and encapsulation/dehydration systems) are more favored because of the technical simplicity (Sakai, 1993). All methods are designed to reduce ice crystal growth in the specimen. Other vegetative material that has been cryopreserved using similar methodological approaches include cell suspensions and callus (more than 40 species), protoplasts (more than 10 species), and root cultures (5 species).

It should be noted that cryopreservation of vegetative germplasm overcomes the problem of genetic instability during storage as all cellular divisions and metabolic processes are stopped. In contrast, two other methods of *in vitro* preservation, normal and slow growth techniques, run the risk of genetic changes (somaclonal variation) in the conserved germplasm that may result in the loss of distinct genotypes.

Species are stored under normal growth conditions (e.g., *Coffea* at 27 °C) for short-term purposes only. The explant (usually meristem or nodal cutting) is frequently transferred (subcultured) to fresh nutrient medium with the risk of microbial contamination, or loss through human error. To retard growth and hence extend the subculture interval, temperature and light intensity are reduced. For example, 0 to 5 °C and 1000 Lux are generally used for cold tolerant species, and 15 to 22 °C and reduced light intensity for tropical species. Alternatively, growth can be slowed down by the addition to the medium of chemicals to induce mild osmotic stress (e.g., mannitol) or hormonal retardants (paclobutrazol, abscisic acid). Also, maintenance of tissue under reduced oxygen (e.g., under mineral oil or liquid medium) slows growth. Under the appropriate conditions subculture intervals can be extended to one year or longer. The slow growth technique is now routinely used for the medium-term conservation of a number of species such as banana, potato, yam, cassava, and strawberry.

Although *in vitro* culture without cryopreservation poses considerable threat of genetic drift, the propagation of plant material in an aseptic environment ensures the production of disease-free stock material, which is readily accessible internationally because it satisfies most country quarantine requirements. Undoubtedly, *in vitro* culture is a valuable complementary approach to field conservation and is particularly useful when applied to species that are predominately propagated vegetatively (banana, potato, and pear), produce non-bankable or highly heterozygous seeds, and have a particular gene combination (i.e., elite genotypes; see Ashmore, 1997). The impact of these positive features of the technique is such that FAO estimates that 37,600 accessions of plant material (vegetative and embryos) are conserved *in vitro* (including cryopreservation) worldwide.

Animal Germplasm Samples

A majority of *ex situ* animal germplasm is maintained in zoological gardens and institutes around the world. The NIAR in Japan holds 621 accessions of animal germplasm, including silk worms in the living state. More than 100 of these accessions, mostly sperm, are cryopreserved (for a general methodology, see the discussion presented later). Similarly, the main gene bank methodology for ova is cryopreservation. At present, however, procedures for sperm and ova preservation are not well developed for wild species, even though a number of reported successes with artificial insemination and frozen semen can be found in the literature, especially for ungulates such as deer and antelopes. The concept of gamete rescue from tissues has considerable value for spermatozoa, where epididymal spermatozoa are readily obtainable post mortem and can be frozen using glycerol as a reasonably standard cryoprotectant. Oocyte cryopreservation has only been achieved in the hamster, rat, rabbit, and cow and is therefore not a practical proposition at present. Interest in freezing ovarian tissue, and then culturing follicles and oocytes by various methods after thawing, has recently been resurrected and progress has included the birth of a lamb originating from ovarian tissue autotransplanted into the donor-recipient after freezing and thawing. In another recent study, isolated rat spematogenic cells were transferred to a mouse testis, where they displayed the ability to develop into

Table 1 Collections Maintaining Microorganisms

Collection	*Location*	*Number of strains*	*Type of material*
Agricultural Research Service Culture Collection, USDA	Peoria, IL (US)	78,010	Algae, bacteria, fungi, yeasts, actiomycetes
American Type Culture Collection (ATCC)	Rockville, MA (US)	53,615	Algae, bacteria, fungi, yeasts, protozoa, cell lines, hybridomas, viruses, vectors, plasmids, phages
CABI Bioscience UK Centre (formerly (IMI)	Egham, UK	21,000	Fungi, bacteria, yeasts
Centraalbureau voor Schimmelcultures (CBS)	Baarn, The Netherlands	41,300	Fungi, yeasts, lichens, plasmids
Culture Collection, University of Goteborg (CCUG)	Goteborg, Sweden	28,100	Bacteria, fungi, yeasts

spermatozoa. The testicular cells were frozen-thawed prior to transplantation and development, and thus there may be some merit in exploring the cryopreservation of testicular cell suspensions as an alternative or adjunct to the preservation of spermatozoa. Cell suspension from genetically important animals could be used to populate the testes of common species, thus permitting the eventual harvesting of spermatozoa. It should be stressed that this technique is still only in its infancy.

Unlike animal species in which reproductive cells and tissues are stored to conserve the gene pool, in humans they are stored for the use of the couple/woman electing to have them cryopreserved. Some embryos are subsequently donated for the treatment of others. Storage is limited by law in some countries.

The storage of embryos is common practice in *in vitro* fertilization (IVF) clinics worldwide, with about 70 offering the service in the United Kingdom alone. Oocyte cryopreservation is poorly developed and few clinics store them other than for research. An increasing number are offering tissue storage.

Despite sporadic reports of cryopreservation using nonequilibrium rate cooling, the vast majority of clinics prefer conventional slow cooling rate procedures with the samples loaded into "straws" (embryos and oocytes) or plastic vials (ovary). The first step involves the addition of the cryoprotectant (about 1.5 M; 1–2 propanediol $\pm$ sucrose (pronucleate and early cleavage stage embryos), dimethyl sulfoxide (4–8 cell embryos, and ovary) or glycerol (blastocyst stage embryos)). Ice formation is then induced at -5 to -7 °C and the samples are cooled further at rates of about 0.3 to 0.5 °C per minute to various subzero temperatures before storage in liquid nitrogen. After warming at appropriate rates (of about 20 to more than 300 °C per minute), the samples are returned to isotonic conditions stepwise, with or without the addition of sucrose to the diluent.

Microorganisms

The most widely applicable preservation method for the preservation of microorganisms that retains viability and stability is cryopreservation. However, for convenience and ease of transport freeze-drying is preferred for most bacteria, viruses, and sporulating fungi (see Hunter-Cevera and Belt, 1996).

There are currently 497 collections from 60 countries maintaining microorganisms registered with the World Data Centre for Microorganisms in Japan (see Sugawara and Miyazaki, 1999). Their on-line database lists species held and the expertise and services provided by the collections along with contact addresses and links to collection websites. There are about 11,500 species held. Over 25% of the strains are held by 5 of the 497 collections (see **Table 1**).

DNA Banks

DNA banks have been established in several places worldwide (Adams, 1997), the largest (more than 140,000 clones) being for plant and animal material at the NIAR in Japan. This collection is mainly constituted of rice clones (about 36,000 comprising Random Fragment Length Polymorphisms, cDNA, YAC) and pig clones (about 106,000 comprising cosmids, BAC, cDNA). The most diverse DNA bank for plants is at the Royal Botanic Gardens (RBG) Kew, which currently holds more than 10,000 DNA samples from a wide range of species. Standards of quality of preservation differ between banks and some, such as those at the Missouri Botanical Gardens (St. Louis, Missouri, in the United States) and several zoos, store only frozen tissue. Others extract DNA and purify it to varying degrees. Purposes for these banks differ as well, with some established to hold samples of a particular country or region with the intent of using these in conservation genetic studies, whereas others focus more on taxonomic and systematic studies, such as that at RBG Kew. Most banks are prepared to consider sharing aliquots of DNA or small samples of tissue with researchers at other institutions.

Field Gene Banks

Field gene banks are *ex situ* collections of mainly agricultural or forest species. They should be contrasted with what might be termed "farm gene banks" where crop germplasm is, in effect, conserved *in situ* by the farmer. Field gene banks normally comprise considerably more individuals per accession than is the case in botanic gardens. Their particular use is for the conservation and utilization of species with the following traits:

- Have nonbankable seeds
- Have long life cycles where growing up material for regular study from a seed collection is impractical
- Are normally vegetatively propagated

Essentially, they are not a new idea as the Kayapo people of Brazil maintain germplasm collections of tuberous plants in hillside gardens protected from flood (see Plucknett *et al.*, 1987). Most countries have at least one field gene bank and FAO (1996) estimate some 527,000 accessions are conserved in this way worldwide. Examples of field gene banks include the National Fruit Collection in the United Kindgom, one for cassava at CIAT, Colombia, one for sugarcane at the Centro Nacional de Pesquisa de Recursos Geneticos e Biotecnologia (CENARGEN), Brazil, and the potato collection at the Centro Internacional de la Papa (CIP) in Peru. Such facilities are considered by FAO to be particularly important in small island developing states. It should also be noted that the Nordic Gene Bank, which operates on a regional basis, unlike many seed banks, includes within its remit the *in situ* conservation of wild crop relatives.

While they offer the opportunity for characterization and evaluation, such collections are labor intensive and are susceptible to catastrophic events. For example, a field collection of yams was lost in St. Lucia during 1994 as the result of cyclone damage. This is one of the reasons that a number of field gene bank collections are now backed up *in vitro*.

Botanical and Zoological Gardens

Botanical and zoological gardens may be seen as types of gene bank with relatively few individuals per accession. While zoos have embraced careful breeding programs that help maximize the genetic value of the limited collections across the world, this is much less true of collections in botanical gardens. Here species may be represented within the "botanical gardens flora' by a single genotype such is the clonal exchange of material.

Acknowledgements

The authors would particularly wish to acknowledge the following for their assistance in providing information: Mr. M. Ambrose (John Innes Centre, United Kingdom); Dr. D. Astley (HRI, United Kingdom); Dr. M. Fay, Mr. R. Smith and Mr. M. Way (RBG Kew, United Kingdom); Dr. W. Holt (Institute of Zoology, London, United Kingdom); Dr. S. Miyazaki and Mr. A. Yamamoto (NIAR, Japan); Dr. B. Panis (Leuven, Belgium); Dr. V. C. Pence (CREW, Cincinatti, United States); Dr. D. Smith (CABI Bioscience, United Kingdom); Dr. S-H Tan (CBS, Netherlands); Dr. L. Towill and Dr. C. Walters (NSSL, United States); Dr. M. Wood (London, United Kingdom); and Dr. X-Y. Yang (XTBG, Peoples Republic of China).

See also: Biodiversity in Plant Breeding. Breeding of Animals. Captive Breeding and Reintroduction. Genetic Diversity. *In Situ*, *Ex Situ* Conservation. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Zoos and Zoological Parks

References

Adams RP (1997) Conservation of DNA: DNA banking. In: Callow JA, Ford-Lloyd BV, and Newbury HJ (eds.) *Biotechnology and Plant Genetic Resources: Conservation and Use*, pp. 163–174. Wallingford: CAB International.

Ashmore SE (1997) *Status Report on the Development and Application of in Vitro Techniques for the Conservation and Use of Plant Genetic Resources*. Rome: IPGRI.

Cromarty AS, Ellis RH, and Roberts EH (1985) *The Design of Seed Storage Facilities for Genetic Conservation*. Rome: IBPGR.

Dickie JB, Ellis RH, Kraak HL, Ryder K, and Tompsett PB (1990) Temperature and seed storage longevity. *Annals of Botany* 65: 197–204.

Ellis RH, Hong TD, and Roberts EH (1995) Survival and vigour of lettuce (*Lactuca sativa* L.) and sunflower (*Helianthus annuus* L.) seeds stored at low and very-low moisture contents. *Annals of Botany* 76: 521–534.

FAO (Food and Agriculture Organization of the United Nations) (1996) *The State of the World's Plant Genetic Resources for Food and Agriculture*. Rome.

FAO/IPGRI (1994). *Genebank Standards*. Rome.

Guarino L, Ramantha Rao V, and Reid R (eds.) (1995) *Collecting Plant Diversity. Technical Guidelines*. CAB International.

Hay FR, Probert RJ, and Smith RD (1997) The effect of maturity on the moisture relations of seed longevity in foxglove, (*Digitalis purpurea* L.). *Seed Science Research* 7, 341–349.

Heywood VH (ed.) (1995) *Global Biodiversity Assessment*. Cambridge: Published for the United Nations Environment Programme by Cambridge University Press.

Hong TD, Jenkins NE, Ellis RH, and Moore D (1998) Limits to the negative logarithmic relationship between moisture content and longevity in conidia of *Metarhizium flavoviride*. *Annals of Botany* 81: 625–630.

Hong TD, Linington S, and Ellis RH (1998) *Compendium of Information on Seed Storage Behaviour*. 2 volumes. Royal Botanic Gardens, Kew, United Kingdom in collaboration with the University of Reading and IPGRI.

Hunter-Cevera JC and Belt A (eds.) (1996) Preservation and Maintenance of Cultures used in Biotechnology and Industry. San Diego, CA: Academic Press.

ten Kate K and Laird SA (1999) *The Commercial Use of Biodiversity: Access to Genetic Resources and Benefit-Sharing*. Oxford, UK: Earthscan Publications.

Pence VC (1991) Cryopreservation of seeds of Ohio native plants and related species. *Seed Science and Technology* 19: 235–251.

Pence VC (2000) *Ex situ* conservation methods for bryophytes and pteridophytes. In: Havens K, Guerrant E, and Maunder M (eds.) *Saving the Pieces: The Value, Limits and Practice of Offsite Plant Conservation*. In preparation.

Plucknett DL, Smith NJH, Williams JT, and Anishetty NM (1987) *Gene Banks and the World's Food*. Princeton, NJ: Princeton University Press.

Pritchard HW (1995) Seed cryopresevation. In: McLellan MR and Day JG (eds.) *Methods in Molecular Biology Volume 38: Cryopreservation and Freeze-Drying Protocols*, pp. 133–144. Totowa, NJ: Humana Press.

Pritchard HW, Poynter ALC, and Seaton PT (1999) Interspecific variation in orchid seed longevity in relation to ultra-dry storage and cryopreservation. *Lindleyana* 14(2): 92–101.

Sakai A (1993) Strategies for survival of plant cultured cells and meristems cooled to −196 °C. In *Cryopreservation of Plant Genetic Resources*, pp. 1–16. Japan International Cooperation Agency.

Steiner AM and Ruckenbauer P (1995) Germination of 110-year-old cereal and weed seeds, the Vienna Sample of 1877. Verfication of effective ultra-dry storage at ambient temperature. *Seed Science Research* 5: 195–199.

Sugawara H, Ma J, and Miyazaki S (eds.) (1999) *World Directory of Collections of Cultures of Microorganisms*, 5th edition Japan: WFCC World Data Center on Microorganisms.

Touchell DH and Dixon KW (1994) Cryopreservation for seedbanking of Australian species. *Annals of Botany* 74: 541–546.

Towill LE and Forsline PL (1999) Cryopreservation of sour cherry (*Prunus cerasus* L.) using a dormant vegetative bud method. *CryoLetters* 20: 215–222.

Vertucci CW and Roos EE (1990) Theoretical basis of protocols for seed storage. *Plant Physiology* 94: 1019–1023.

Genes, Description of

Michael F Antolin and William C Black IV, Colorado State University, Fort Collins, CO, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 1–11, © 2001, Elsevier Inc

Glossary

Allele A unique nucleotide sequence of a gene, which may or may not affect phenotypic expression.
Chromatin A complex of DNA and associated proteins that when assembled make up chromosomes of eukaryotes.
Chromosome A thread-like structure which can be visualized when condensed during cell division; consists of DNA and proteins.
DNA Deoxyribonucleic acid, the molecule of inheritance which stores genetic information passed from one cell to another and from one generation to succeeding generations, composed of four nucleotides: adenine (A), cytosine (C), guanine (G), and thymine (T).
Eukaryotes Organisms with genetic material organized into chromosomes contained within a membrane-bound nucleus in the cell—eukaryotic cells undergo mitosis and meiosis during cell division, ensuring the equal division of chromosomes among daughter cells.
Exons Parts of a gene sequence transcribed and translated into a protein; usually interrupted by introns which are spliced out of messenger RNA before translation.
Genetic code The language that specifies how DNA will be translated into amino acid sequences by means of three nucleotide "words" (codons) that specify the 20 amino acids and regulators of transcription (start and stop codons).
Heterochromatin Regions of chromosomes that in general do not include coding DNA, always remain condensed during a cell's life cycle.
Introns Generally noncoding segments of genes, situated between exons and removed before translation of messenger RNA to proteins.
Locus A precise location in a genome, whether a gene is found there or not; in former times "locus" was used interchangeably with "gene," but the definition has become more specific in the era of molecular genetics.
Nucleotide Subunit of DNA and RNA, composed of a ringed five-carbon sugar, a ringed nitrogen-rich base, and phosphates; nucleotides are often referred to as base pairs because individual types form complementary hydrogen bonds (G-C and A-T in DNA, G-C, and A-U in RNA) to make double-stranded DNA and RNA molecules.
Organelle Membrane-bound structures within eukaryotic cells that carry out specific functions and often contain their own DNA; typical organelles are mitochondria and chloroplasts.
Posttranscriptional modification Cutting and splicing of messenger RNA, in some cases to produce alternative transcripts and proteins with different structural or regulatory properties (e.g., sex determination in *Drosophila*).
Prokaryotes Single-celled organisms defined by having their DNA arranged as a circular molecule not contained within a nucleus, and which reproduce by simple fission.
Repetitive DNA Regions of DNA that include the same DNA sequence repeated several hundred or thousand times; regions with repeated segments that involve only 2–5 base pairs of DNA are called microsatellites.
RNA Ribonucleic acid, composed of nucleotides like DNA, but in RNA the base uracil (U) replaces the base thymine (T), RNA molecules form important structural and regulatory parts of cells.
Transposable elements Also called mobile elements, jumping genes, or insertion sequences (in prokaryotes); fragments of DNA that contain genes that enable the DNA fragment to excise itself and then relocate into a new location in the genome.

Introduction

Understanding the mechanisms of heredity is a key to the study of biodiversity for at least three reasons. First, genetic variation provides the material for evolution—descent with modification—including adaptation within populations, diversification of lineages, and the formation of new species. Second, standing genetic variation within populations may reflect the recent history of populations and indicate prospects for change in the future. Third, variation in molecular genetic markers can be used to track the relationships among living groups of organisms and the taxonomic status of individual populations. Heritability of both adaptive and neutral genetic variation depends upon genes, defined as sequences of DNA that specify cell structure, including proteins and several types of RNA. Genes also code for enzymes and proteins that bind to DNA or RNA to control gene expression and cellular physiology. In former times, the term "locus" was used interchangeably with "gene," but today locus simply means a particular location in a genome, whether the DNA codes for a gene product or not. Locus may still be used to describe neutral molecular markers, where allelic variation at the DNA level is used to study population genetics, relatedness within populations, parentage, and genome mapping.

Genes and genomes are both complex and continually modified. Organisms share a common biochemical mechanism of heredity, but with characteristic differences that allow us to make hypotheses about the history of life. Examples

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00063-0

include the bacterial origins of mitochondria and chloroplasts in eukaryotic plant and animal cells. Variation in genomes is pervasive in nature, and continually arises by the inevitable and inexorable process of mutation, both on the level of individual nucleotides within DNA molecules and the level of chromosomal gene rearrangements. One of the most significant (and initially unappreciated) discoveries in genetics is that genomes are not static within organisms; transposable elements, for instance, can excise themselves from one locus and reinsert in another place in the genome. The original discovery of transposable elements was in maize by Barbara McClintock, who in the 1930s correlated regular breaks in chromosomes with the loss of purple pigment in corn kernels. Regaining the pigment was associated with similar breaks at other places in the genome. We now know these as "jumping genes," which carry with them genes that control their ability to excise, duplicate, and reinsert into multiple places within genomes, sometimes carrying other genes with them, and sometimes disrupting genes at their insertion sites. Dozens of transposable elements are now known, and more continue to be discovered in an ever-widening sample of organisms (Li, 1997). As a consequence, it is no longer possible to view genomes as static entities, with genes aligned in precise locations that forever remain stable.

In this article, we describe genes in the context of entire genomes. In prokaryotes (bacteria and blue-green algae), genomes are organized as naked circular DNA molecules within the cytoplasm. In eukaryotes, most genetic material is in chromosomes—long, linear arrays of DNA bound with proteins—found within the membrane-enveloped nucleus. Genes in eukaryotes can also be found in prokaryote-like genomes inside cellular organelles like mitochondria and chloroplasts (the photosynthetic organelles of plants). All genomes contain regions of noncoding single-copy DNA and repetitive DNA, which are used extensively in population genetic and phylogenetic studies as molecular markers. Avise (2004) provides an in-depth overview of the use of molecular markers in the study of evolution.

Structure of DNA and RNA

What follows may serve as a primer for those who are not conversant with molecular genetics. A full description of genes is not possible in the space of a few pages; those seeking more detail should consult other sources (e.g., Hartl and Jones, 2005; Lewin, 2004). Molecular genetics is a mushrooming field, but the salient features of genes are universal, and genes and the genetic code can be understood on an intuitive level without extensive knowledge of biochemistry. In general, genes are carriers of information working in a nested series from individual nucleotides, to short regulatory DNA or RNA sequences, to relatively large RNA molecules and enzymes composed of hundreds of nucleotides or amino acids, to enormous cell structure proteins comprising thousands of amino acids folded into complex three-dimensional shapes.

First and foremost, in this hierarchy of information are nucleic acids, DNA and RNA, which are the molecules of heredity for most life forms (one notable exception being prions, which are infectious proteins and cause the famed "mad-cow disease" and a number of neurological disorders in humans, Prusiner and Scott, 1997). Both DNA and RNA are made up of four nucleotide subunits linked together to form long chains (**Figure 1**). Nucleotides themselves are made up of three parts: ringed sugars, phosphates, and ringed nitrogen-and-carbon bases. Nucleotides are joined to each other to

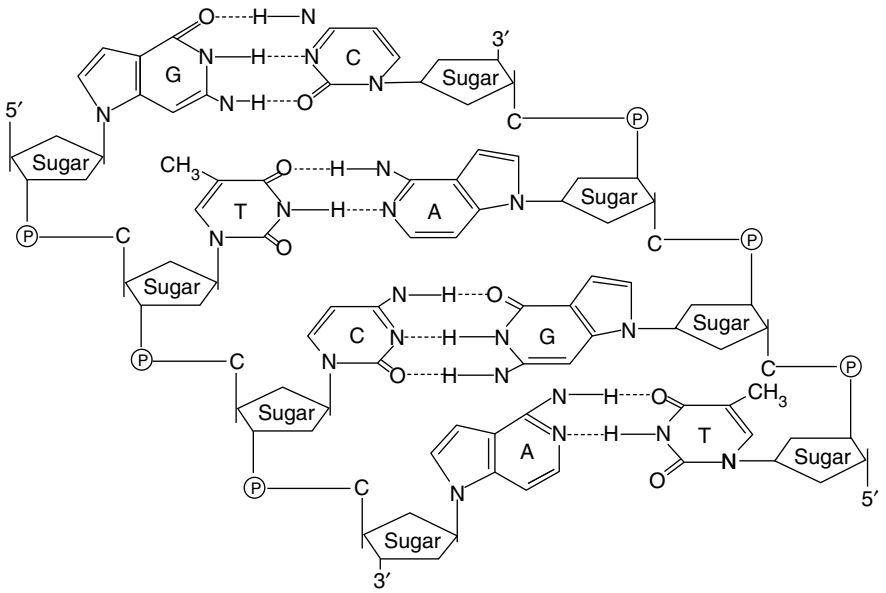


Figure 1 DNA (and RNA) are composed of four nucleotide subunits, A, C, G, and T (U replaces T in RNA), each covalently bonded to a sugar, with sugars linked to each other by a phosphate bond (circled P) between the 3′ and 5′ carbon atoms of the adjacent sugars. Complementary strands of DNA are held together by hydrogen bonds (shown as broken lines). Of the two strands, the lower is in the 5′–3′ direction, showing how nucleic acid sequences are usually written (GTCA).

form chains by bonds between phosphates and sugars. Nucleotides are complementary to each other in terms of the numbers of hydrogen bonds they form between bases: A (adenine) and T (thymine) form two hydrogen bonds, G (guanine) and C (cytosine) form three bonds. Two complementary chains intertwine to form double-stranded DNA or RNA by hydrogen bonding. Double-stranded DNA molecules form the classic "double-helix," where two chains of nucleotides joined by hydrogen bonds spontaneously form twists of $\sim 36^\circ$ so that ~ 10 base pairs make a complete 360° turn. DNA can take on a number of other forms, such as supercoiled, depending upon temperature, pH, and the nucleotide sequence.

RNA can also form a double helix, but one with a broader diameter. An additional feature of RNA is that single-stranded molecules include complementary sequences that allow the single-stranded RNAs to fold back on themselves, forming double-stranded helical stems and other three-dimensional grooves, knobs, and loops (Gutell, 1996). Three-dimensional structure is a vital characteristic of the long ribosomal RNA (rRNA) subunits of ribosomes, and the shorter transfer RNA (tRNA) molecules that are part of the translation machinery for processing messenger RNA (mRNA) into protein.

Genes

The number of genes within organisms varies tremendously between viruses, bacteria, and multicellular eukaryotic organisms. For instance, it is estimated that humans have $\sim$25,000–40,000 genes, the fruit fly *Drosophila melanogaster* 13,767, the nematode worm *Caenorhabditis elegans* $\sim$19,000, the mustard *Arabidopsis thaliana* 25,500, the yeast *Saccharomyces cerevisiae* 6275, the bacterium *Escherichia coli* 4286, and the human immunodeficiency virus (HIV) RNA retrovirus 9. For the organisms whose entire genomes have been characterized and whose complete nucleotide sequences are known (*Drosophila*, yeast, worm, *E. coli*, and HIV), the number of protein-coding genes is known with accuracy. For others, the numbers represent estimates based on incomplete sequencing and lack of similarity to previously identified genes from other species. For instance, early estimates of the number of genes in humans have been as high as 140,000 genes. The number of identified genes is bound to increase within each kind of organism, with the possible exception of viruses. In part, this is because novel mechanisms of inheritance and gene regulation by RNAs continue to be discovered (Pearson, 2006).

The flow of information within a cell generally follows a pathway, with genes coded in DNA transcribed into three types of RNA, followed by translation of the nucleotide sequence in mRNA into chains of amino acids (proteins and enzymes). This process of information transfer, which has been called the "central dogma of molecular biology," is depicted in **Figure 2**. Two types of genes result from the transfer of information from DNA to proteins. The first type may encode enzymes such as DNA polymerases, soluble enzymes, and transcription factors that control the metabolic and biochemical processes within a cell. The second type encodes the physical structures of the cell, including cytoskeleton proteins like tubulin and actin that affect cell shape, and subunits of rRNA that comprise the machinery of protein translation. The majority is expressed as proteins and enzymes, and the majority of proteins and enzymes is involved in cellular regulation rather than cell structure.

Protein and enzyme coding is specified by the genetic code (**Table 1**), which contains the language of protein synthesis. The genetic code has a series of three-base-pair codons that specify 20 amino acids as well as codons that signal translation to "start" and "stop." The genetic code is almost universal, although there are a number of differences between nuclear and mitochondrial DNA (mtDNA) in eukaryotes and there are unique codons in a few bacteria and Protista (Li, 1997). The code functions by specifying three-base-pair codons of mRNA that match three-base-pair anticodons found on a loop of the tRNA. Each tRNA is specific to a particular amino acid. The complex dynamic molecular structure involved in translation of the genetic code contains mRNA, ribosomes (rRNA plus proteins), and tRNAs.

Because DNA contains specific information, transcription and translation must proceed in only one direction for the information to be communicated accurately. This is analogous to the conventions followed in most written Indo-European languages, where words form sentences that are read on a page from left to right. The direction along any strand of DNA or RNA can be identified by which carbon atoms of the ringed-sugar-and-phosphate backbone are bonded together. Carbons in the ribose sugar are numbered 1–5, and bonds occur between a phosphate attached to the 5′ carbon of one sugar with a hydroxyl group attached to the 3′ carbon of the next. Consequently, directionality can be specified along each nucleic

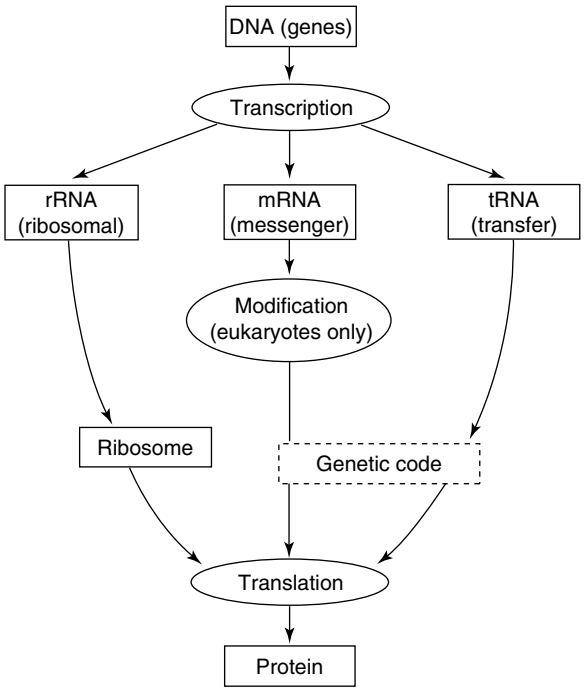


Figure 2 The flow of information from DNA to proteins. DNA is transcribed into three types of RNA, which then combine to read the genetic code and translate the information into proteins (or enzymes).

Table 1 The genetic code that shows how 20 amino acids are specified by three-base-pair codons[a]

First position (5′ end)	*Second position*				*Third position (3′ end)*
	U	*C*	*A*	*G*	
U	UUU	UCU	UAU	UGU	U
	UUC PHE	UCC	UAC TYR	UGC CYS	C
	UUA	UCA SER	**UAA**	**UGA STOP**	A
	UUG LEU	UCG	**UAG STOP**	UGG TRP	G
C	CUU	CCU	CAU	CGU	U
	CUC	CCC	CAC HIS	CGC	C
	CUA LEU	CCA PRO	CAA	CGA ARG	A
	CUG	CCG	CAG GLN	CGG	G
A	AUU	ACU	AAU	AGU	U
	AUC ILE	ACC	AAC ASN	AGC SER	C
	AUA	ACA THR	AAA	AGA	A
	AUG MET	ACG	AAG LYS	AGG ARG	G
G	GUU	GCU	GAU	GGU	U
	GUC	GCC	GAC ASP	GGC	C
	GUA VAL	GCA ALA	GAA	GGA GLY	A
	GUG	GCG	GAG GLU	GGG	G

[a]The three stop codons (UAA, UAG, and UGA) and the start codon for methionine (AUG) are shown in bold. Most amino acids are specified by more than a single codon, with several (arginine, leucine, and serine) specified by six codons. There is a weakly positive general relationship between the frequency of incorporation of an amino acid into proteins and the number of codons specifying the amino acid. In most cases, a change in the third position in a codon does not result in a change in the amino acid because most tRNA anticodons will still bind to mRNA codons if the first two positions on the mRNA match the tRNA. This is the so-called wobble effect and redundancy in the genetic code.

acid molecule, and the genetic code must follow this direction for the language of transcription and translation to be faithfully transmitted. For instance, RNA polymerases transcribe by adding new nucleotides to the 3′ end of a growing RNA molecule, and so create the RNA in the 5′–3′ direction. However, transcription is from a complementary DNA chain that is read in the 3′–5′ direction. In general, reading frames of DNA are written by placing the 5′ end on the left, and protein sequences are written in the same left-to-right direction specified by the nucleotide sequence (e.g., **Table 2**).

Prokaryotic Genomes

The genomes of most prokaryotes (bacteria and blue-green algae) are composed of one double-stranded circular DNA molecule attached to a central core of proteins in a series of supercoiled loops emanating from the center like the cotton fibers of a dust mop. Although the nucleoid "chromosome" is not enclosed within a nuclear membrane, it tends to be found in a particular part of the cytoplasm. An additional feature of prokaryotes is their ability to acquire smaller circular DNA molecules, called plasmids. Plasmids contain numerous genes, including genes found in the chromosome. Transfer of plasmids into and out of bacterial cells makes them useful in DNA-based biotechnology and is responsible for the rapid transfer of antibiotic resistance genes between bacterial strains.

Genes in prokaryotes are collinear and uninterrupted, in that one to several genes are expressed in the order in which they follow a promoter. Transcription of mRNA leads directly to translation with little modification (**Figure 3**). This arrangement of several genes being expressed as a group is sometimes referred to as polycistronic, and is restricted to prokaryotic and organelle genomes.

The control of gene expression in prokaryotes is through either positive or negative regulation. Under negative regulation, like that found in the tryptophan operon of *E. coli*, a repressor protein combines with a tryptophan molecule and then binds to the repressor region of the gene, turning off transcription when tryptophan becomes common within the cell. Under positive control, a transcription factor must bind with the regulatory site to initiate transcription. This is part of the *lac* operon of *E. coli*, one of the earliest metabolic pathways described at the molecular level by Jacques Monod and colleagues at the University of Paris. The *lac* operon is mainly under the negative control of a repressor that is removed when lactose is present in the cell. However, transcription also depends upon a second protein that must bind for transcription of RNA. This double control allows the bacterial cell to turn off the *lac* operon when glucose, a preferred food source to lactose, is present in the cell.

Eukaryotic Genomes—Chromosomes

The genomes of eukaryotic organisms are characterized by DNA arranged in long linear molecules and associated proteins in a complex structure known as chromatin, which is contained within a membrane-bound nucleus. Chromatin is in a diffused state during interphase of the cell cycle when transcription and translation occur. When the chromatin coils and condenses during cell division, it forms itself into microscopically visible threads in the nucleus. These threads are called chromosomes. The number of chromosomes found in different eukaryotic groups varies considerably, but each group displays a characteristic range of chromosome numbers. For example, humans and great apes have haploid chromosome numbers of $N = 23$ and $N = 25$, respectively, while

Table 2 Types of mutations: (A) chromosomal mutations (very small fragments are illustrated; chromosomal mutations usually affect large chromosomal fragments, and (B) gene mutations

A. Chromosomal mutations		
	DNA sequence (of a chromosome)	Type of mutation
Original sequence	AGTTCGTAGGGTACCTGATCGACG	
Mutant sequences	AGTTCGTACCTGATCGACG	Deletion of AGGGT
	AGTTCGTAGGGTACCTGAT**CATGAG**CGACG	Insertion of CATGAG
	AGTTCGTAGG**TCCATG**GATCGACG	Inversion of GTACCT
	GACGGTAGGGTACCTGATC**AGTTC**	Translocation of AGTTC and GACG
	AGTTCGTAGGGTA**CCTGACCTGA**TCGACG	Duplication of CCTGA
B. Gene mutations		
	Three-base-pair codons (top line), and the proteins specified (lower line)	Type of mutation
Original gene	AUG UGG UGU GAG UAC AUU CGA GAG AAG UAG met-trp-cys-glu-tyr-ile- arg-glu-lys- stop	
	Mutations at the third codon	
Mutant genes	AUG UGG UG**C** GAG UAC AUU CGA GAG AAG UAG met-trp-cys-glu-tyr-ile- arg-glu-lys-stop	Synonymous (no amino acid change)
	AUG UGG U**C**U GAG UAC AUU CGA GAG AAG UAG met-trp-ser- glu-tyr-ile- arg-glu-lys-stop	Missense (amino acid change)
	AUG UGG UG**A** GAG UAA AUU CGA GAG AAG UAG met-trp-stop	Nonsense (mutation to a stop codon)
	AUG UGG **U**UG AGA GUA AAU UCG AGA GAA GUA G met-trp-leu- arg-val-asn-ser-arg-glu-val-	Nonsense (insertion and shift of reading frame)
	AUG UGG GAG AGU AAA UUC GAG AGA AGU AG met-trp-glu-ser-lys-phe-glu- arg-ser-	Nonsense (deletion and shift of reading frame)

Source: Reproduced from Li W-H (1997) *Molecular Evolution*. Sunderland, MA: Sinauer Associates.



Figure 3 General structure of prokaryotic genes, where a repressor, promoter, and operator act together to control expression of three genes. Repressors, operators, and promoters are binding sites on the DNA where enzymes attach to begin or prevent transcription.

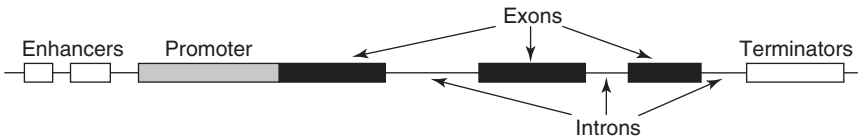


Figure 4 General structure of a eukaryotic gene, including the upstream promoter and enhancer sequences where protein transcription regulators can bind, and the downstream terminator where transcription ends, and where the 3′ UTR controls the longevity of mRNA in the cell. Eukaryotic genes are interrupted by exons, which are translated or transcribed, as well as nontranslated introns, which are removed during the process of posttranscriptional modification.

different groups of parasitic Hymenoptera (wasps) have $N = 5$ and $N = 10$, an ant species has $N = 1$ chromosome, many moths and butterflies have $N = 50$–80 and a species of fern has $N = 630$!

Eukaryotic genes contrast sharply with prokaryotic genes in both structure and regulation of expression (**Figure 4**). First, while several prokaryotic genes are grouped together under control of a single promoter, each eukaryotic gene is expressed individually under control of a promoter specific to that gene. Second, eukaryotic genes are often interrupted, with nontranslated introns interspersed between exons that are translated into proteins. The result is that after transcription, introns are excised and the pieces are spliced back together as a mature mRNA for translation into proteins. The number of introns can be vast. The human titin gene encodes a long, elastic protein that occurs in the sarcomeres of muscle and has 234 exons, the longest single exon (17,106 nucleotides), and the longest coding sequence (80,781 base pairs = 26,927 amino acids). The longest primary transcript in humans is from the dystrophin gene (2.4 million base pairs).

One consequence of posttranscriptional modification is that the same gene can be spliced into several different mature mRNAs by including or removing different exons. A classic example is the development of gender in animals by alternative splicing and cascades of gene expression leading to male versus female characteristics (Hodgkin, 1990; Marin and Baker, 1998). In *D. melanogaster*, for instance, the male and female forms of the gene *Doublesex* both have the first three

exons, but differ in the particular 3′ exons included after splicing.

Eukaryotic gene expression is regulated by transcription factors that bind to promoters and enhancers. Transcription requires the binding of RNA polymerases and protein transcription factors to a promoter, a DNA sequence on the 5′ end of the gene, to form a transcriptional complex. Many eukaryotic promoters, but by no means all, contain a TATA box (sequence TATAAA), which typically lies within 50 bases of the transcriptional start site. Another commmon example is the E-box sequence CACGTG. Enhancers are DNA sequences that function to increase transcription of a gene over basal level, and are usually responsible for tissue-specific gene transcription. Specialized transcription factors bind to the enhancer sequence and transcriptional complex, and are regulated by whether the genes coding for the enhancer have been expressed, by environmental signals, and by signals from other cells. Transcription factors can be classified according to the structure of their DNA binding domains; these include zinc finger proteins, helix–turn–helix proteins, leucine zipper proteins, helix–loop–helix proteins, and steroid receptors.

Once, transcribed, mRNAs carry terminal sequences in the 3′ untranslated region (3′ UTR) that predetermine their life span in the cell, and in turn regulate how much of a protein is produced. The sequence AUUUA in the 3′ UTR is a signal for early degradation, and the more often that this sequence occurs, the shorter will be the life span of the mRNA.

Another common feature of eukaryotes is gene duplication. Some genes that are important for basic housekeeping, like rRNA, may have hundreds or thousands of copies of the gene within the genome, which results in functional redundancy within the genome. Another form of duplication is represented by genes coding for blood proteins (globins) in vertebrates. These gene families include several kinds of duplications, where gene clusters code for a diversified set of proteins that are part of fetal blood versus adult blood on the one hand, and a series of nonfunctional pseudogenes on the other.

Noncoding Repetitive DNA Sequences

Large portions of the genomes of eukaryotes and some prokaryotes consist of repetitive DNA, which is mostly untranscribed and consists of tandemly arranged sequences with identical nucleotide composition. Each DNA sequence can be repeated up to several million times within the genome. Some repetitive DNA is tied up in the heterochromatic parts of the genome that make up the structural elements of chromosome such as centromeres (where sister chromatids bind during cell division) and telomeres (the ends of chromosomes). Otherwise, repetitive DNA may be interspersed throughout the genome. Repetitive elements can be very large ($>10{,}000$ nucleotides) or of moderate size (100–800 nucleotides), but the majority of highly repetitive sequences consists of repeats smaller than 100 nucleotides. Generally speaking, repetitive DNA makes up as little as 10% (fruit fly), and as much as 90% (humans) of eukaryotic genomes, and much of the repetitive DNA is now thought to arise as a by-product of transposable elements moving within genomes (Li, 1997).

Repetitive DNAs are highly variable in number and location among individuals in a population, and are generally classified according to the size of the repeat unit, the number of repeat units per array, and the genomic location of the tandem arrays. No consensus for classifying simple tandem repeats currently exists, and terminology varies among authors. The term "microsatellite" (also simple sequence repeat (SSR)) denotes any tandem array of repeats with a unit length of 2–5 nucleotides. The most common human microsatellites are dinucleotide arrays of $(CA)_n$, which means n repeats of CA. The human genome has $\sim 50{,}000$ $(CA)_n$ arrays, or one array every 30,000 nucleotides. It has been estimated that there are 300,000 trinucleotide and tetranucleotide microsatellites in the human genome or, on average, one array in every 10,000 nucleotides of genomic DNA.

Microsatellites are used extensively in population genetics, DNA fingerprinting for forensics or paternity analysis, and genome mapping. Because a large number of alleles are usually found at each microsatellite locus, a collection of several microsatellite loci can lead to a huge number of possible combinations or DNA profiles. This feature makes these loci useful for identity analysis and forensics. For example, in the investigation of a crime where a biological specimen has been obtained from the crime scene, DNA from the specimen and from a suspect is typed for size at several unlinked microsatellite loci. If the profiles do not match, the suspect is excluded. If they match, the probability of this match occurring by chance can be calculated from the frequencies of alleles in the population. This probability is usually extremely low because of high polymorphism at microsatellite loci. When primers for microsatellites are well designed, and when each pair of microsatellite primers amplify only a single locus, the high polymorphism makes them very useful for genetic mapping, studies of population structure, detecting loss of heterozygosity, and studies of behavior.

Organelles and their DNA

The most likely origins of organelles within the cells of eukaryotic organisms are ancient bacteria or blue-green algae taken up as endosymbionts into the cells of early eukaryotes. The strongest line of evidence for an endosymbiotic origin is that organelle genomes closely resemble those of prokaryotes, both in general and in the details of DNA sequences within genes. As in prokaryotes, organelle genomes are circular double-stranded molecules with little repetitive DNA. Additionally, organelle genes do not contain introns, are transcribed as a group (polycistronically), and have relatively simple prokaryote-like promoters (Mayfield *et al.*, 1995). Although organelles carry out their own transcription and translation, gene expression and regulation in organelles depend upon transcription factors that originate from nuclear genes.

Mitochondrial DNA in Animals

The animal mtDNA is a single circular, double-stranded DNA molecule that ranges in size from 14,000 to 42,000 base pairs.

With few exceptions, the animal mitochondrial genome contains 37 genes that encode 13 protein subunits, 22 tRNAs, and a small and a large ribosomal subunit RNA. Several compilations of primers that are useful for polymerase chain reaction (PCR) amplification of segments of the mitochondrial genomes of animals are currently available (Avise, 2004; Simon *et al.*, 1994; Palumbi, 1996; Roehrdanz and Degrugillier, 1998).

The analysis of mtDNA has become one of the most powerful tools for studying animal populations. The mitochondrial genome is maternally inherited as a haploid genome: it does not recombine, and it mutates at a faster rate than the nuclear genome in most animal groups. Sequence differences arising from mutations in mtDNA haplotypes (haploid alleles) record the phylogenetic histories of female lineages within and among species. Variation in frequencies of different mtDNA haplotypes can be used to estimate effective migration rates among populations, and the genetic diversity within and among populations. This level of variation is also used for estimating phylogenetic relationships among recently evolved taxa. A number of other unique attributes distinguish animal mtDNA, including within-individual homoplasmy (predominance of a single mtDNA sequence) despite between-individual sequence differences and the rapid pace of nucleotide substitution. Although inheritance of animal mtDNA is usually maternal, it is biparental in some mollusks.

Plant mtDNA

Plant mtDNA is extremely different from animal mtDNA, which has prompted several investigators to suggest independent symbiotic origins. The mtDNA genome in plants ranges from 200,000 to 2,400,000 nucleotides in circumference, and typically exists as a collection of different-sized circles arising from extensive recombination within individuals that convert between a "master" allele and subgenomic circles. Inheritance is not always maternal. Although plant and animal mtDNAs are similar with regard to gene content and general function, their evolutionary patterns differ. Plant mtDNA appears to evolve rapidly with respect to gene order, but slowly in nucleotide sequence (perhaps 100-fold slower than animal mtDNA). Reasons for the slow mutation rate are not understood, but plant mtDNA may have relatively error-free DNA replication, or perhaps highly efficient enzymes for repair of DNA damage. The lack of nucleotide-level variation and regular recombination has limited the utility of plant mtDNA in population genetics.

Chloroplast DNA

Chloroplast DNA (cpDNA) in photosynthetic land plants is also a circular genome, which varies in size from about 120,000 to 247,000 nucleotides, largely because of a large inverted repeat that includes genes for the rRNA subunits. Each chloroplast contains from about 22 to 900 cpDNA copies and each encodes 123 genes. These include four genes for rRNA, 20 genes for ribosomal proteins, 30 genes for tRNAs, many of the photosynthetic proteins, six of the nine genes for ATPase, and chloroplast RNA polymerase. Unlike mtDNA, 15 of the cpDNA genes contain introns. Most chloroplast proteins are encoded by the nucleus, chloroplast ribosomes consist of 52 proteins but only 19 of them are encoded by plastid genome.

cpDNA is transmitted maternally in most flowering plants, biparentally in a few, and paternally in gymnosperms. cpDNA genes have been shown to transpose to the nucleus and there is good evidence that mtDNA, cpDNA, and nuclear genomes exchange genes. The rate of cpDNA evolution generally appears slow both in terms of primary nucleotide sequence and in terms of gene rearrangement. Because of the large size of the cpDNA genome, most systematic treatments have involved restriction site or sequence determinations for particular genes or have monitored the taxonomic distributions of unique cpDNA structural features across higher-level plant taxa. Nonetheless, some studies have uncovered considerable intraspecific cpDNA variation as well, and this suggests that portions of the cpDNA should be useful in population genetics. Again, several primers for amplifying cpDNA genes in a number of plant species are available (e.g., ndhF: Olmstead and Sweere, 1994; rbcL: Palumbi, 1996).

Mutation

Mutation is the ultimate source of genetic variation, and is generally defined as any change in genetic material, whether phenotypic effects caused by the mutation are large, small, or nondetectable. The causes of mutation include direct insults from environmental effects like ultraviolet light, chemicals, X-rays and other atomic particles, but many mutations arise from errors in DNA replication and DNA repair during cell division. RNA viruses, in fact, depend upon error-prone RNA polymerases to generate novel variants during every replication of the viral genome.

One class of mutations arises from gross changes at the level of chromosomes, where chromosomes break and are then modified by deletions, insertions, inversions, translocations, and duplications (**Table 2**). Although the term "chromosome" is used here, similar mutations occur in the genomes of prokaryotes as well. Chromosomal mutations can result in changes in the total amount of DNA found within and among different species, and in the case of gene duplications or deletions, can change the number of genes. Mutations like these can have drastic phenotypic effects. Individual genes can be disrupted if insertions occur within them, or if part of a gene is lost in a deletion. Transposable elements can cause deleterious mutations by inserting themselves into genes.

Inversions of chromosomal fragments, where a fragment is flipped 180° from its original orientation and reinserted at the same place, will cause complex looping of the chromosomes when they pair during meiosis and will reduce recombination. Inversions are common in some organisms, like the Diptera, especially fruit flies, blackflies, and malaria mosquitoes.

Translocation of pieces of chromosomes from one place to another within the genome will result in a change in the size or number of chromosomes. The effects of translocations can be dramatic, often resulting in sterility because of the inability of the altered chromosomes to pair during meiosis. Thus,

translocations are a potential mechanism of isolation of a single species into incompatible breeding groups, and eventually the formation of new species.

The magnitude of phenotypic effects of duplications depends on the size of the affected chromosomal fragment, and thus the number of genes that are duplicated. An example is trisomy 21, which results in Down syndrome (mental retardation, flat facial features, and heart defects). Trisomy results from a failure of proper segregation of chromosomes during meiosis so that one gamete contributes two complete copies of chromosome 21, resulting in three copies in the zygote. On the other hand, duplications are an integral part of the evolution of multigene families, like the various globin genes and nonfunctional pseudogenes in vertebrates. Duplication is the main mechanism by which entirely new genes arise (Li, 1997).

Mutations can also occur at the level of single-nucleotide changes, including some that affect the reading of the genetic code (**Table 2**). Single-nucleotide substitutions are classified as transitions if they occur between similar nucleotides, between purines (A↔G) or pyrimidines (C↔T), and transversions if mutations substitute purines for pyrimidines and vice versa (e.g., A↔T). This distinction is important because the transitions are much more common than transversions, sometimes occurring 20 times more frequently. Further examination of the genetic code (**Table 1**) reveals that where a substitution occurs within a codon will determine the severity of the mutation. Most mutations in the first nucleotide position will result in an amino acid change (missense) or a mutation to a stop codon (nonsense). Mutations at the second position always result in a change (missense and nonsense). Third-position mutations are usually synonymous ("silent"), resulting in no change of amino acid. Nucleotide insertions and deletions will result in a shift of the "reading frame," so that all amino acids are changed after the mutation.

Mutations occur at very low rates, and are usually measured as the number per meiosis in eukaryotes, or as the number per cell division for prokaryotes. The numbers of new mutations are commonly <1 in 10,000 (10^{-4}) for nucleotide substitutions, and are as low as 1 in 100,000,000 (10^{-8}) for some visible mutations like coat colors. In an evolutionary sense, mutation is considered a very slow process. Mutation can still produce considerable variation within populations because the numbers of new mutations produced in each generation in a population will be a function of population size. In addition, if the number of genes in a genome is taken into account, even low mutation rates can generate high numbers of new mutations in each individual within a population.

See also: Adaptation. Diversity, Molecular Level. Ecological Genetics. Speciation, Process of

References

Avise JC (2004) *Molecular Markers; Natural History, and Evolution*, 2nd ed. Sunderland, MA: Sinauer Associates.

Gutell RR (1996) Comparative sequence analysis and the structure of 16S and 23S rRNA. In: Zimmerman RA and Dahlberg AE (eds.) *Ribosomal RNA: Structure, Evolution, Processing, and Function in Protein Biosynthesis*. Boca Raton, FL: CRC Press.

Hartl DL and Jones EW (2005) *Genetics Principle and Analysis*, 6th ed. Sudbury, MA: Jones and Bartlett Publishers.

Hodgkin J (1990) Sex determination compared in *Drosophila* and *Caenorhabditis*. *Nature* 344: 721–728.

Lewin B (1999) *Genes VII*. Oxford, UK: Oxford University Press.

Lewin B (2004) *Genes VIII*. Upper Saddle River, NJ: Pearson Prentice Hall.

Li W-H (1997) *Molecular Evolution*. Sunderland, MA: Sinauer Associates.

Marin I and Baker BS (1998) The evolutionary dynamics of sex determination. *Science* 281: 1990–1994.

Mayfield SP, Yohn CB, Cohen A, and Damon A (1995) Regulation of chloroplast gene expression. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46: 147–166.

Olmstead RG and Sweere JA (1994) Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. *Syst. Biol.* 43: 467–481.

Palumbi SR (1996) Nucleic acids II: The polymerase chain reaction. In: Hillis DM, Moritz C, and Mable BK (eds.) *Molecular Systematics*, 2nd ed. Sunderland, MA: Sinauer Associates.

Pearson H (2006) What is a gene? *Nature* 441: 399–401.

Prusiner SB and Scott MR (1997) Genetics of prions. *Annu. Rev. Genet.* 31: 139–175.

Roehrdanz RL and Degrugillier ME (1998) Long sections of mitochondrial DNA amplified from fourteen orders of insects using conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 91: 771–778.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, and Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 87: 651–701.

Genetic Diversity

Eviatar Nevo, University of Haifa, Haifa, Israel

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 195–213, © 2001, Elsevier Inc.

Glossary

Electrophoresis Separating charged molecules (such as polypeptides or polynucleotides) between the two poles of an electric field.
Genome A complete single set of genes of an organism or organelle; also the basic haploid chromosome set.
Genomics The study of the molecular organization of genomic DNA and physical mapping.
Heterozygosity The average number of different heterozygotes across loci divided by all loci studied.
Polymorphism Different alleles at a gene locus within a population.

Genetic diversity (i.e., molecular hereditary differences within or between populations) is the basis of evolutionary change. The nature of genetic diversity among organisms has always been the basic problem of evolutionary genetics (Darwin, 1859; Lewontin, 1974; Kimura, 1983) as well as of domestication, agriculture, and medicine. However, despite its cardinal role in evolutionary theory and application, the maintenance of genetic diversity remains largely enigmatic, notwithstanding the dramatic discoveries of molecular biology, which revealed abundant genetic diversity in nature.

The Problem

Historical Background

The era of molecular evolution was ushered in by three major discoveries in the late 1950s and 1960s: (i) recognition by Markert and Moller in 1959 of molecular diversity of enzymes and their importance in genetics, physiology, development, and evolution; (ii) Identification by Zuckerkandl and Pauling in 1965 of protein sequence variation between species and the resulting postulate of the molecular clock; and (iii) evaluation in 1966 of enzyme variation (isozymes and allozymes) in *Drosophila* by Lewontin and Hubby and Johnson *et al*. and in humans by Harris. Molecular biology permitted the characterization of genetic diversity among individuals, populations, and species—the three cornerstones of evolution. It did so first by unraveling relationships between genes and proteins (Lewontin, 1974). Second, it did so by elucidating, at the extranuclear and nuclear coding and noncoding DNA regions, the structure, expression, function, mechanism, and evolution of genes, intergenic spacers, and multigene families by employing recombinant DNA methodologies (Watson *et al*., 1987; Avise, 1994).

Advancements in molecular evolution have followed the introduction of new laboratory techniques (Avise, 1994). Among the most influential methods have been protein electrophoresis in the late 1960s and 1970s, restriction fragment length polymorphism (RFLP) analyses of nuclear and mtDNA in the late 1970s and 1980s, DNA fingerprinting in the mid to late 1980s, and polymerase chain reaction (PCR)-mediated DNA sequencing in the 1990s. Comparative genomics and genome prospecting (*Science* **286**, October 15, 1999) promise to highlight the genetic basis of evolutionary change and the nature, expression, meaning, transfer, and regulation of information in biological systems, thus unraveling the blueprint and evolutionary forces driving life.

The Enigma of the Maintenance of Genetic Diversity

Early and later molecular–genetic studies revealed massive levels of genomic diversity, surpassing any prediction. In addition, a comprehensive compilation of dense genetic maps was published of viruses, bacteria, plants, animals, and humans. In 1998, Deloukas *et al*. physically mapped 30,000 encoded human genes for proteins of known functions. The ongoing human genome project (HGP), the largest biological project ever (Bodmer and McKie, 1994), ushered in a new era in the life sciences in which complete genomes of prokaryotes (Casjens, 1998) and eukaryotes became available. Genetic maps, current complete genome sequences and those expected in the near future, and tens to hundred of thousands of protein and DNA sequences have contributed dramatically to elucidating genome diversity, organization, expression, dynamics, and evolution (Karlin *et al*., 1998). They form the basis for future studies of genetic diversity, when more genomes within species will be analyzed.

However, despite the ease of measuring and deciphering genetic diversity and the dramatic advances in comparative genomics, the evolutionary forces that generate and maintain segregating genetic diversity, preventing allele fixation or random elimination in nature, remain elusive.

Theories of Molecular Evolution: Selection versus Neutrality

Theoretically, balancing selection could account for protein polymorphism (Gillespie, 1991). In contrast, the neutral theory of molecular evolution (Kimura, 1983) suggests that most of the molecular–genetic diversity within and between species is neutral (i.e., non-selective) or "non-Darwinian." The neutralist–selectionist debate has been one of the major controversies in evolutionary biology since the late 1960s. How much of the genetic diversity at single and multilocus

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00065-4

structures is adaptive, processed by natural selection and contributing to differences in fitness? The problem of distinguishing between deterministic and stochastic forces in evolution has pervaded evolutionary biology at all levels, genotypic and phenotypic, and is now focused on DNA polymorphisms. I recognize the contribution of the neutral and nearly neutral theories of molecular evolution, primarily by representing a null hypothesis to selection. Nevertheless, by ignoring the ecological heterogeneity and stress in evolution, neutral and nearly neutral theories have stripped genetic diversity from nature. I believe that in-depth understanding of genetic diversity in nature is intimately linked to the interface between ecology and genetics; hence, to ecological genetics and now to ecological genomics. I submit that only this essential interface can meaningfully highlight the dynamic evolution of genetic diversity in nature.

The Structure of This Review

This review comprises three unequal perspectives of molecular–genetic diversity of protein and DNA in nature: methodology, evidence, and theory. The main focus is on the huge amount of protein and DNA evidence from nature in diverse groups of organisms across phylogeny from bacteria to humans and across diminishing geographical scales (global, regional, and local). The evidence naturally leans heavily on results of genetic diversity in plants and animals obtained during the past 25 years at the Institute of Evolution, University of Haifa, Israel, that demonstrate the pervasiveness of ecological determinants in genetic differentiation. In theory too, I focus on theoretical results obtained during the past decade at the Institute of Evolution that show how selection in cyclical environments can maintain genetic polymorphism in nature, thus preventing drift. I mainly refer to books which cite the primary literature.

Methodologies

The methodologies analyzing protein and DNA diversity have been extensively described (Lewontin, 1974; Avise, 1994; Mitton, 1997), so they are only briefly reviewed here.

Protein Polymorphism

Technical Innovation

New techniques at the protein level permit an increasingly subtle resolution of allelic variation at a locus. These techniques include isoelectric focusing in various media; high-resolution, two-dimensional electrophoresis; thermal and urea denaturation analysis; and sequential gel electrophoresis (SAGE), in which electromorphic classes were retested with other pHs, buffer systems, and gel pore sizes (Ramshaw *et al.*, 1979).

Hidden Variation

"Alleles" detected by routine gel electrophoresis are essentially phenotypes or "electromorphs" (i.e., internally heterogeneous, genetically involving "hidden variation"). For example, xanthine dehydrogenase has many more alleles than are visible routinely (Ramshaw *et al.*, 1979). Recent pulsed-field gel electrophoresis for separating large DNA fragments, recombinant DNA techniques, methods of genomic "hopping," and new DNA sequencing strategies and genomic analysis reinforce the powerful field of molecular evolution, climaxing in the HGP (Bodmer and McKie, 1994). Extensive comparative analyses across genomes of model organisms (Karlin *et al.*, 1998) became the focus of genomic and proteomic studies, highlighting diversity in nature (Avise, 1994). These new horizons partly unraveled the molecular structure, function, and evolution of life. Bioinformatics became increasingly important as many genes and genomes became analyzable. Comparative genomics, using dense genetic maps based on coding genes and on microsatellite and single nucleotide polymorphisms (SNPs), permit precise gene homolog alignment across taxa, the unraveling of gene function and regulation, highlighting of genome organization and evolution, and the determination of the genetic basis of speciation and adaptation.

DNA Polymorphism

RFLP, PCR, and Mini- and Microsatellites

Many assays reveal DNA diversity (Avise, 1994). Before the revolutionary PCR was established, DNA fingerprinting technology had been revolutionized by RFLP analysis, a very efficient technique but quite laborious and not suitable for high-throughput applications. Hundreds of restriction enzymes, originating in bacteria, cleave duplex DNA at particular oligonucleotide sequences, usually 4–6 base pairs (bp) in length. Since its development, PCR has been a powerful tool for DNA fingerprinting and for effectively measuring genetic diversity within and between populations. These molecular markers include minisatellite, also called variable number of tandem repeats (VNTR), consisting of core sequences 10- to 15-bp long and extending from 16 to 64 bp, and microsatellites (2–6 bp), also called simple sequence repeats (SSRs) (Goldstein and Schlotterer, 1999). Microsatellites and minisatellites differ in size, mutation processes, and chromosomal distribution, but the boundary between the two classes is not defined.

Random Polymorphic DNA

A novel PCR-based strategy involving the use of arbitrary primers (AP-PCR) to amplify random polymorphic DNA (RAPD) fragments has been developed using short primers ($\approx$10 bp). RAPDs generate polymorphic markers using very small amounts of starting DNA or RNA, independently of any prior knowledge of the target DNA sequence. This feature makes RAPD a useful tool in genetic analysis. Compared with protein polymorphisms, DNA markers reveal more genetic diversity and permit choice among sets of loci with different functions and patterns of inheritance (Mitton, 1997; Avise, 1994). Use of molecular markers dramatically widened the perspectives of evolutionary biology (Avise, 1994).

Amplified Fragment Length Polymorphism: A New, Powerful Technique for DNA Fingerprinting

A major novel DNA fingerprinting technique called amplified fragment length polymorphism (AFLP) was recently described. AFLP is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. AFLP can analyze genetic diversity in many hundreds of genes in both coding and noncoding regions across the genome if several enzyme combinations are applied, allowing high-resolution genotyping of fingerprinting quality. The time and cost efficiency, replicability, and resolution of AFLPs are superior or equal to those of other markers (allozymes, RAPD, RFLP, and microsatellites), except that AFLP methods primarily generate dominant rather than codominant markers.

Nucleotide Polymorphism

Studies of nucleotide polymorphism have three distinct advantages: (i) detection of all types of sequence changes (single nucleotide substitutions, insertion/deletions, and copy number variation in nucleotide repeat motifs), of which SNPs are the most common; (ii) variant detection in coding (cSNPs and noncoding DNA; and (iii) combined detection and genotyping using a single method.

Summary

DNA techniques revealing DNA diversity include RFLP, RAPD and AFLP, mini- and microsatellites, SNP and sequencing; all are relevant for analyzing DNA diversity. PCR is currently the method of choice for amplifying DNA segments for detecting polymorphism. Generally, DNA diversity is higher than protein diversity (see **Figure 4**).

Comparative DNA Analysis across Diverse Genomes

The era of comparative genomics is dramatically advancing (Casjens, 1998; Karlin *et al.*, 1998; *Science*, October 15, 1999), including the analysis of complete genomes for various organisms and imminent completion of the HGP. Genes, individuals, and species are becoming comparable. Sequence polymorphism will highlight the generation, maintenance, and function of genetic diversity responsible for controlling normal development, physiologic homeostasis, and disease processes. This information explosion will revolutionize molecular evolutionary studies. Comparative analysis of complete genomes includes assessments of genomic compositional contrasts based on di-, tri-, and tetranucleotides relative abundance values; identification of rare and frequent oligonucleotides; evaluations and interpretations of codon biases in several large prokaryotic genomes; and characterizations of compositional asymmetry between the two DNA strands in certain bacterial genomes. Comparative analysis also allows identification of alien (e.g., laterally transferred) genes and detection of potential specialization islands in bacterial genomes and the assessment of DNA curvature. It can compare genomes within species and across life and cluster organisms according to linguistics based on many words in large DNA stretches.

Microarrays: Biotechnology's Discovery Platform for Functional Genomics

Advances in microarray technology make possible massive parallel mining of biological data, based on PCR, with biological chips providing hybridization-based expression, monitoring, polymorphism detection, and genotyping on a genomic scale, as reported by Schena *et al.* (1998). Microarrays containing sequence representative of all the genes of an organism may soon permit the expression analysis of the entire human genome in a single reaction. These "genome chips" will provide unprecedented access to genomic diversity of many thousands of genes to large-scale gene discovery as well as polymorphism screening and mapping of genomic DNA clones on a massive scale—critical for science and application in agriculture and medicine. Oligonucleotide microarray (DNA chip)-based hybridization analysis is a promising new technology which potentially allows rapid and cost-effective screens for all possible mutations and sequence variations in genomic DNA. Currently, it is performed in humans, mice, and *Arabidopsis*, but use in additional species has been reported. The increasing use of relatively inexpensive microarrays is expected to revolutionize genomic, proteomic, and other biological research projects.

Encyclopedia of Genes

Extensive ongoing programs in plants and animals are generating a large database of expressed sequence tags (ESTs) that can provide rapid access to numerous genes and their diversity. The development of a comprehensive database of ESTs for *Arabidopsis*, corn, soybean, rice, barley, wheat, mouse, and human has been reported. As of October 23, 1998, 352,040 sequences had been generated in the mouse and annotated and deposited in dbEST, in which they comprised, according to Marra *et al.* (1999), 93% of the total ESTs available for the mouse. EST data are versatile and have been applied to gene identification, comparative sequence analysis, comparative gene mapping and candidate disease gene identification, genome sequence annotation, microarray development, and the development of gene-based map resources. Large-scale exploration of the genic diversity of the genome is becoming possible through the UniGene database (http://www.ncbi.nlm.nih.gov/UnGene).

Evidence: Genetic Diversity Within and Among Species

Protein Diversity

Electrophoretic results have revealed large amounts of genic polymorphism in natural populations (Lewontin, 1974; Nevo, 1978, 1988, 1998; Hamrick *et al.*, 1979; Nevo *et al.*, 1984; Avise, 1994; Mitton, 1997). This is true for both eukaryotes and prokaryotes. Currently, theory appears to lag behind evidence. The review and reanalysis of approximately 1100 plant and animal species (Nevo *et al.*, 1984) revealed an average heterozygosity, H, of 0.073 (SD 0.076) and an average proportion of loci polymorphic, P, of 0.284 (SD 0.197). Across

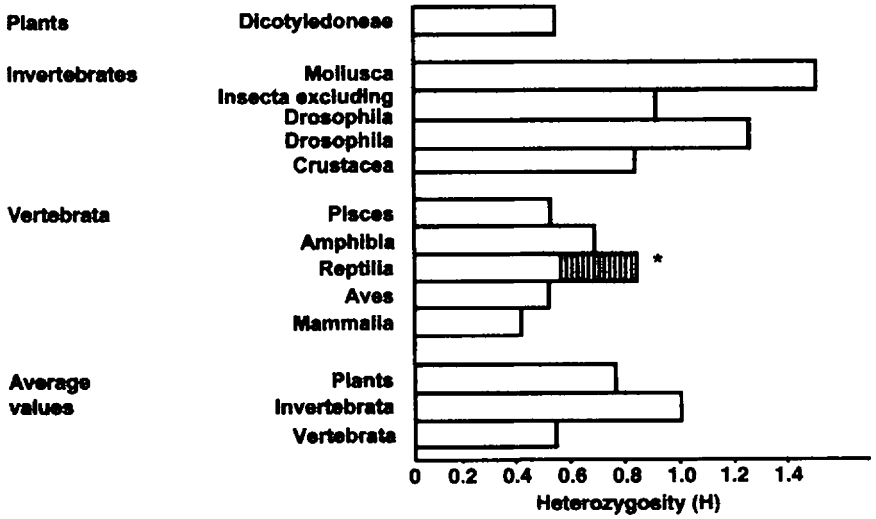


Figure 1 Levels of genetic (allozymic) diversity among higher taxa based on a mean of 23 gene loci per species, in 1111 species. *Hatched region indicates parthenogenetic species (reproduced from Nevo E, Beiles A, and Ben-Shlomo R (1984) The evolutionary significance of genetic diversity: Ecological, demographic and life history correlates. In: Levin S (mang. ed.) *Lecture Notes in Biomathematics* and Mani GS (ed.) *Evolutionary Dynamics of Genetic Diversity*, vol. 53, pp. 13–213. Berlin, Heidelberg: Springer-Verlag).

species, the coefficient of correlation between H and P was $r = 0.793$, $p < 0.001$. These estimates are based on the average of 23 electromorphically detectable gene loci (minimum 14 loci) and an average of 199 individuals per species (minimum 10 individuals) (**Figure 1**). This set of data was assembled up to 1983, and many hundreds of species have since been tested (see Figure 2.2 in Avise, 1994, that summarizes heterozygosity estimates of 1803 species: 648 vertebrate, 370 invertebrates, and 785 plant species), all demonstrating diverse levels of polymorphisms (Mitton, 1997). Genetic diversity varies dramatically among species. Fruit flies, marine mussels, and conifers have much genetic diversity, whereas large vertebrates have much less (Mitton, 1997).

The aforementioned genetic indices appear to be lower than the real ones since routine horizontal gel electrophoresis underestimates diversity and protein coding genes represent only a small portion of the genome. Regarding levels of diversity, protein diversity does not represent the noncoding genome, which amounts to 90–98%, involving invariant and variant portions whose general diversity levels in natural populations appear to be high. Studies of "hidden variation" indicate that most protein diversity is detectable (Ramshaw *et al.*, 1979), thus increasing H of enzymatic loci. Moreover, at the DNA sequence level of individuals, most loci may prove heterozygous. Nevertheless, there are significant merits to the estimates of H and P, derived from routine protein electrophoresis. First, although relative, low, and limited genomically, they are commensurate estimates of genetic diversity among numerous populations and species living in varied ecologies. Second, H and P are highly correlated in global (Nevo *et al.*, 1984) and regional (Nevo, 1988, 1998) analyses. Thus, although hidden variation is regrettably lacking in studies of routine protein electrophoresis, in most species the high correlation of H and P makes for a profitable analysis since P is largely independent of the high resolution achievable by SAGE and other new techniques. Third, the relatively large number of electromorphic protein loci currently available in Drosophila and the many species analyzed permit a sound statistical analysis of isozyme phenotypes in an ecological context rather than only in genetic terms (Nevo *et al.*, 1984; Nevo, 1978, 1988, 1998).

The Adaptive Evolution of Enzyme Kinetic Diversity

Kinetic studies of hemoglobin, haptoglobin, and transferin proteins and at least a dozen enzyme polymorphisms typically reveal biochemical kinetic differences among the gene products of alternative genotypes at a locus (Mitton, 1997). These include single gene effects such as lactate dehydrogenase in killifish, leucine aminopeptidase in blue mussel, and phosphoglucose isomerase (PGI) in Collias butterflies. In the latter, Watt and colleagues discovered in the 80s that PGI heterozygotes fly over a greater range of temperatures and produce 33% more eggs. Additional examples are alcohol dehydrogenase (ADH) in fruit flies, salamanders, and barley; glutamate pyruvate transaminase in cope-pods; and esterase, glucose-6-phosphate, 6-phospho-gluconate dehydrogenase, and superoxide dismutase in fruit flies. In all these cases, biochemical–kinetic studies reveal differences among phenotypes, either within or between species, that have measurable effects of alternative genotypes on the physiology of whole individuals. They result in fitness differences distinguished by natural selection in accord with their alternative spatiotemporal environments (Mitton, 1997). Usually, the heterozygotes are intermediate between homozygotes, but sometimes molecular overdominance for enzyme kinetics is found.

Patterns of Variation Among Loci

Proteins vary genetically. Regulatory enzymes appear to be more variable than nonregulatory enzymes, and enzymes that work on many substrates appear to be more variable than

enzymes that utilize a single substrate. Genetic diversity tends to decrease from monomer to dimer to tetramer as the steric restrictions on the molecule increase. The number of alleles at a locus and the heterozygosity at a locus tend to increase with the subunit size of the protein.

Genetic Population Structure

Geographical clines may reflect the action of natural selection on genetic polymorphism at local, regional, and global scales. In *Drosophila melanogaster*, several latitudinal clines occur for many characters such as allozymes, inversions, and quantitative traits. The identical nature of these clines on the various continents, in both the Northern and Southern Hemispheres, strongly suggests adaptation to specific stress factors, primarily climatic selection. Polymorphism for stress-resistance genes abounds in natural populations. The ADH polymorphism shows high frequencies of the S allele in tropical regions and this declines with latitude. The reasons for this cline are difficult to determine because of the entanglement with other polymorphisms varying with latitude. In 1977, Van Delden and Kamping reviewed the tentative connections with other polymorphisms such as α-GPDH, in (2L)t inversion, body size, and development time with respect to the possible environmental stress factors involved. They concluded from these results, and also from recent experiments, that the (2L)t plays a dominant role in resistance to high temperature and is partly responsible for the ADH cline. They are currently studying the specific selective forces acting on ADH, focusing on the physiological and life history aspects. Many plant and animal species distributed in several climatic zones—tropical, temperate, and arctic—display a decline in gene diversity, both allozyme and DNA, towards the Arctic (Nevo *et al.*, 1984). This is highlighted by postglacial colonizing populations. Regional and local clines also abound in nature.

Partition of Genetic Diversity

An important analysis is that of genetic partition within and between populations. In most analyzed outcrossers, genetic diversity is chiefly within populations (e.g., 85% in humans). In contrast, in some selfers (e.g., wild emmer wheat) only 40% of genetic diversity is within and 60% is between populations. This partition is shaped by both external (gene flow and selection at the population level) and internal (genomic and genic) evolutionary forces.

DNA Diversity

DNA technologies revolutionized population genetics by providing an unprecedented amount of genetic diversity for critical analysis and hypothesis testing. I briefly review these dramatic developments, from molecular (extranuclear, mitochondrial, chloroplast, and nuclear) DNA markers to sequence polymorphism.

MtDNA Diversity

MtDNA is largely maternally nonrecombiningly inherited and demonstrates a rapid pace of evolution and an extensive high level of intraspecific restriction. PCR and sequence polymorphism derive primarily from base substitution with a preponderance of transitions, length variation, sequence rearrangements, and duplications of coding sequences (Avise, 1994). Hence, its relevance is not only to large-scale phylogenies but also to micro-geographic divergence, i.e., to intraspecific phylogeography, the recently established bridge between population genetics and systematics. The amount of mtDNA polymorphism within species and sequence heterogeneity is striking. Remarkably, a different scenario emerged for plant mtDNA from that in animals. Whereas animal mtDNA ranges in size from 14 to 26 kb in length, that of plants (e.g., maize) is approximately 30 times larger. Most polymorphism in plant mtDNA is attributable to major reorganizations of sequence. The level of mtDNA polymorphism in higher animals is several-fold higher than that of single-copy nuclear DNA. The maternal inheritance and high polymorphism of mtDNA and the assay of whole mtDNA by long PCR provide unique opportunities for population evolutionary studies of animals and plants on both micro- and macrogeographic scales. They contribute to natural history, population structure and signature, gene flow, hybridization, biogeography, phylogeny, and biological conservation.

Mitochondrial genomes are increasingly being used to study ancient divergences among animal groups. Recent studies by Curde and Koder (1999) of complete mitochondrial DNA sequences reached somewhat heretical conclusions, raising questions about the use of mitochondrial gene sequences for studying the relationships among highly divergent lineages. Other studies have documented convergent evolution of mitochondrial gene order, casting doubt on the use of these characters for phylogenetic analysis. The use of mitochondrial genomes for studying such deep divergences is coming under increased scrutiny, and these novel results need to be confirmed with data from nuclear genes.

Comparative Summary of Genetic Distances in the Vertebrates from the Mitochondrial Cytochrome b Gene

Mitochondrial cytochrome b is among the most extensively sequenced genes to date across the vertebrates. Johns and Avise (1998) employed approximately 2000 cytochrome b gene sequences from GenBank to calculate and compare levels of genetic distance between sister species, congeneric species, and confamilial genera within and across the major vertebrate taxonomic classes. The results of these analyses parallel and reinforce some of the principal trends in genetic distance estimates derived from multilocus allozymes. In particular, surveyed avian taxa on average show significantly less genetic divergence than do same-rank taxa surveyed in other vertebrate groups, notably amphibians and reptiles.

Chloroplastid Genome Comparisons

Chloroplast genome (cpDNA) evolution is also very dynamic, indicating gene transfer to the nucleus. Among the 210 different protein-coding genes contained in the completely sequenced chloroplast genomes from a glaucocystophyte, a rhodophyte, a diatom, an euglenophyte, and five land plants, Martin and colleagues identified in 1998 the set of 45 genes common to each and to a cyanobacterial outgroup genome.

Phylogenetic influence, with an alignment of 11,039 amino acid positions per genome, surprisingly indicates that independent parallel gene losses in multiple lineages outnumber phylogenetically unique losses by more than four to one. They identified homologs of 44 different plastid-encoded proteins as functional nuclear genes of chloroplast origin, providing evidence for endosymbiotic gene transfer to the nucleus in plants. cpDNA diversity in lodgepole pine was high within populations, with little (>5%) differentiation among populations, in contrast to the mtDNA pattern as shown by Hamrick and Godt in 1990 (as cited in Mitton, 1997, p. 71). Extensive intraspecific cpDNA diversity may sometime exceed interspecific diversity, reflecting dynamic ecologies and seriously affecting phylogenetic conclusions, as in the pine *Draba* species.

Simple Sequence Repeats (SSRs)

SSRs consist of tandem repeats of relatively short nucleotide motifs, such as TCCTCCTCCTCC. Microsatellite repeat number can range from two $(TA)_2$ or three $(GA)_3$ to a few dozen $(GCAA)_{11}$, whereas minisatellites often consist of many dozens or even hundreds of repeated motifs. Significantly, SSRs experience mutation at notably higher rates than do nonrepetitive sequences: 10^{-2} to 10^{-3} per locus, per gamete, per generation, which leads to their high polymorphism. Replication slippage, sister chromatid exchange, unequal crossing-over, and gene conversion may cause microsatellite diversity. Replication slippage seems to play a major role in producing new alleles at microsatellite loci (Goldstein and Schlotterer, 1999). SSRs are characterized by high site-specific and reversible rates of gain or loss in the number of tandem repetitions of a short DNA motif. Importantly, recent critical population genetics and specific chromosome data derived from Li and colleagues in 1999 at our laboratory indicate massive genomic non-random chromosomal and environmental distributions of microsatellites, suggesting that they are subjected to natural selection, generating adaptive complexes in contrast to neutral theory expectations.

The functional properties of SSRs suggest evolution's effect on mutability, as recently concluded by King and Soller (as cited in Wasser, 1999). Many SSRs are functionally integrated into the genome and exert a quantitative regulatory effect on gene transcription activity affecting phenotype and fitness. Genes associated with SSRs may be favored by indirect selection whenever quantitative variation in the affected traits can provide a population with genetic resilience for adaptation, especially in stressful, fluctuating, or heterogeneous environments. Such "adjustable genes" may provide a prolific and evolutionarily significant source of quantitative genetic variation.

Nucleotide diversity of the human nuclear genome has been estimated to be approximately 0.1%. For two randomly selected sequences, this number translates into one polymorphic site for 1000 nucleotides (or, in a large sequence sample, one polymorphic site is expected for every 200–500). Since the human nuclear genome contains approximately 3 billion nucleotides, several million polymorphic sites are expected to exist.

Single Nucleotide Polymorphism

SNPs are the most frequently found DNA sequence variations in animal and plant genomes, usually followed by SSRs, RAPDs, and allozymes. Recent SNP surveys in humans reported different rates of polymorphism among classes of sites within genes (noncoding, degenerate, and nondegenerate) as well as between genes. Of all coding SNPs, 54% lead to predicted change in the protein sequence. As expected, the coding SNPs that alter amino acid sequence of the encoded protein are found at a lower rate and with lower allele frequencies than silent substitutions. This was interpreted as a reflection of selection against deleterious alleles during human evolution. Determination of ancestral alleles from human SNP polymorphisms became available using high-density oligonucleotide arrays. A densely packed map of human SNP sites could efficiently identify disease-associated genes by linkage disequilibrium between sets of adjacent markers and highlight human history. In plants, SNP polymorphism has been associated with transcript efficiency, nonconcensus splice sites, and gene expression. The following sections discuss the results of genetic diversity and divergence at global, regional, and local scales based primarily on results derived at the Institute of Evolution since 1979.

Global Analysis of Genetic Polymorphisms

Global Allozyme Diversity Across Phylogeny

We used the entire globe as a large-scale ecological genetic laboratory. We analyzed the correlates of biotic factors involving ecological, demographic, and life history variables with the level of genetic diversity in natural populations of animals and plants (Nevo *et al.*, 1984). This review involved 1111 species studied for allozymic variation with an average of 23 gene loci in each species and a biotic profile characterized by 21 variables (7 ecological, 5 demographic, and 9 life history and other biological characteristics). We then (i) estimated the levels of genetic diversity, indexed by heterozygosity and polymorphism, for all species, three major taxa (vertebrates, invertebrates, and plants), 10 different higher taxa, and the categorized 21 biotic factors; (ii) correlated the levels of genetic diversity with the biotic factors; and (iii) matched some of the evidence obtained with theoretical predictions.

The following results were obtained:

1. The levels of genetic diversity vary nonrandomly among populations, species, and higher taxa and also among ecological parameters (life zone, geographical range, habitat type and range, and climatic region) (**Figure 2**), demographic parameters (species size and population structure, gene flow, and sociality), and life history characteristics (longevity, generation length, fecundity, origin, and parameters related to the mating system and mode of reproduction).
2. Generally, genetic diversity is higher (i) in species living in broader environmental spectra; (ii) in large species with a patchy population structure and limited migration as well as in solitary or social species; (iii) in species with small body size, annuals or long-lived perennials, older in time, with smaller diploid chromosome numbers; and (iv) in plant species primarily out-crossed, reproducing sexually,

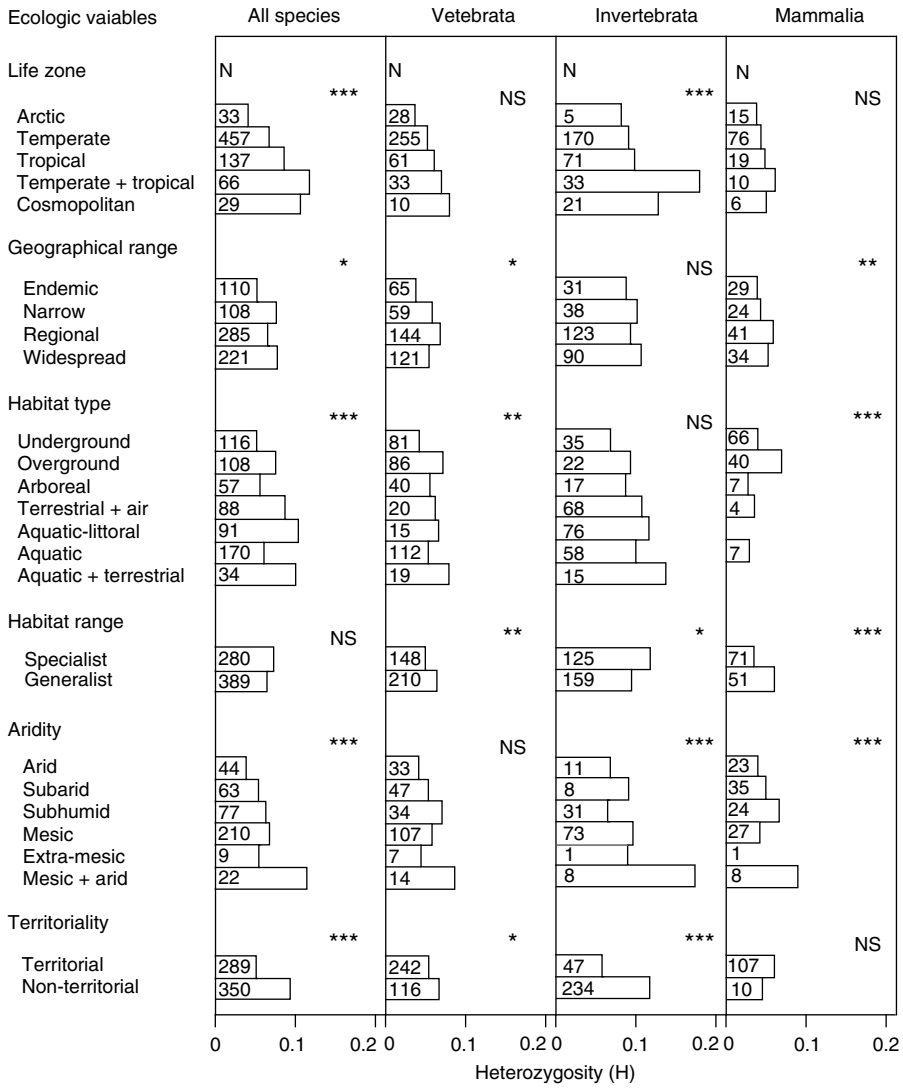


Figure 2 Global analysis of genetic diversity: levels of heterozygosity of ecological factors. Numbers in the bars indicate species. Significance NS = $p > 0.05$; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$ (reproduced from Nevo E, Beiles A, and Ben-Shlomo R (1984) The evolutionary significance of genetic diversity: Ecological, demographic and life history correlates. In: Levin S (mang. ed.) *Lecture Notes in Biomathematics* and Mani GS (ed.) *Evolutionary Dynamics of Genetic Diversity*, vol. 53, pp. 13–213. Berlin, Heidelberg: Springer-Verlag).

and pollinated by wind. Diversity among species has also been summarized by Mitton (1997). Higher genetic diversity characterizes species with high fecundity, large populations, broad geographic ranges, and a high speciation rate. In contrast, low genetic diversity characterizes endemic species with low fecundity and low speciation rate.

3. Genetic diversity is partly correlated and predictable by a three- or four-variable combination of ecological, demographic, and life history variables, largely in that order. Ecological factors account for the highest proportion of the 20% explained genetic variance of all species compared with demographic and life history factors (90, 39, and 3.5%, respectively). Within individual higher taxa, the explained portion of genetic diversity increases considerably (mean of 44% and maximum of 74% in mollusks). Neutrality theory would be substantiated if demographic variables, rather than ecological variables, could better explain the variation in genetic diversity. The opposite, however, was often found, substantiating selection as an evolutionary driving force of genetic diversity in nature. A more detailed global analysis of genetic differentiation in amphibians was performed in 1991 by Nevo and Beiles and in small mammals across the globe (Nevo, 1999). In both cases, genetic diversity, and in amphibians also genome size, is predictable by ecological factors: The higher the environmental heterogeneity and stress, the higher the heterozygosity. Similar conclusions were drawn for 480 species of plants in 1990 by Hamrick and Godt and for diverse animal species (Mitton, 1997). As expected, theoretically the global evidence exhibits a positive relationship between environmental diversity and genetic diversity, as was also shown in cage experiments. Clearly,

environmental diversity enhances genetic diversity. Levins (1968) predicted that environmental grain determines the level of polymorphism. Finegrained species (highly mobile) would evolve a monomorphic strategy, whereas coarse-grained species (highly sedentary) would maintain polymorphism. This theoretical prediction is largely supported by evidence (Nevo *et al.*, 1984).

The patterns and correlates of genetic diversity revealed in the global analysis for many unrelated species, subdivided into different abiotic and biotic regimes, strongly implicate selection in the genetic differentiation of species. Natural selection in several forms, but most likely through the mechanisms of spatiotemporally varying environments at the various life cycle stages of organisms, appears to be an important evolutionary force causing change at the molecular level. Other evolutionary forces, including mutation, migration, and genetic drift, certainly interact with natural selection, either directly or indirectly, and thereby contribute differentially, according to circumstances, to population genetic differentiation at the molecular level. However, the final orientation of the evolutionary process is determined by natural selection (Bell, 1997).

Regional Analysis of Genetic Polymorphisms

Regional: Allozyme Diversity across Phylogeny in Israel and the Near East

Israel as an Ecological Genetic Laboratory of Increasing Aridity Southwards

Our regional analyses of genetic diversity extended over Israel and the Near East. We used Israel, with its remarkable physical and biotic diversities, as a medium to large-scale ecological genetic laboratory (Nevo, 1988, 1998). In 1988, Nevo and Beiles conducted an ecological test of protein polymorphism in 13 unrelated taxa of plants, invertebrates, and vertebrates involving 21 species, 142 populations, and 5474 individuals (**Figure 3**), following the extensive studies of allozyme diversity in 38 species in Israel.

Each individual, population, and species were tested, on average, for 27 enzymatic gene loci. These species varied in population size and structure, life histories, and biogeographical origins, but they largely shared geographically short and ecologically stressful gradients of increasing aridity in Israel, both eastward (70 km) toward the Syrian and Jordanian deserts and (mainly) southward (260 km) towards the Negev desert. We found genetic parallelism across most taxa

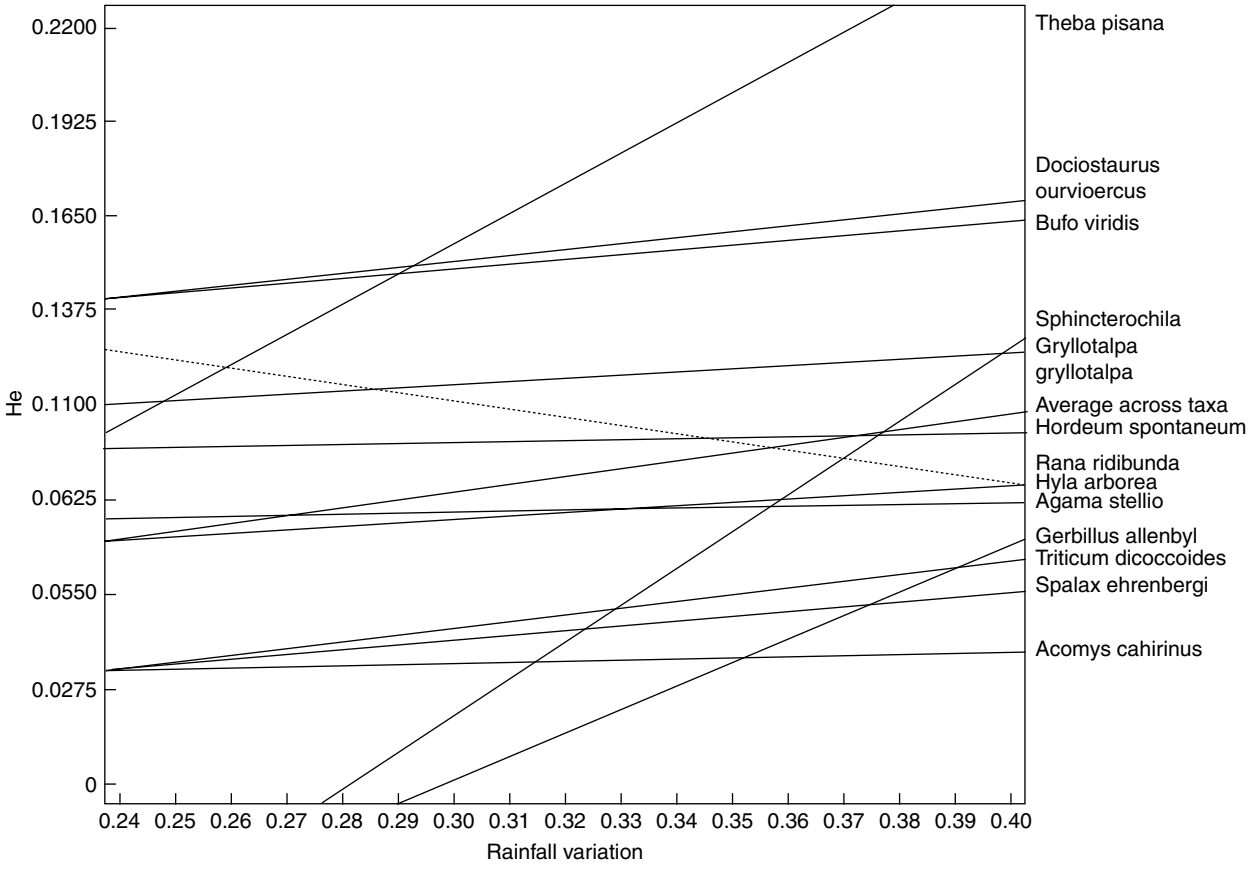


Figure 3 Parallel genetic patterns in the level of gene diversity, H_e, of 13 enzymatic systems averaged across all taxa studied here. The average regression line is $H_e = -0.0556 + 0.4803$ RV. The only change is that *Gryllotalpa gryllotalpa* has been divided into two species, *G. tali* ($2n = 19$) and *G. marismortui* ($2n = 23$). See Broza *et al.*, 1998, cited in Nevo, 1999. The regression line represents *G. tali* (reproduced from Nevo E and Beiles A (1988) Genetic parallelism of protein polymorphism in nature: Ecological test of the neutral theory of molecular evolution. *Biol. J. Linn. Soc.* 35: 229–245).

and most loci. Observed average heterozygosity, H, and gene diversity, H_e, were positively and overall significantly correlated with rainfall variation, which is lowest in the northwest and highest in southern deserts (i.e., with a wider, fluctuating, climatic temporal niche in the desert) (**Figure 3**). Similar trends were found for several DNA systems in subterranean mole rats in Israel (Nevo, 1999). This result corroborates the environmental theory of genetic diversity, which regards ecological heterogeneity and stress as major determinants enhancing genetic divergence.

Notably, in the regional Israeli studies, heterozygosity, H, is negatively correlated with effective population size, N_e; H increases, whereas N_e declines drastically toward the deserts. This contradicts a basic postulate of neutrality theory, which predicts positive correlation between H and N_e (Kimura, 1983). Our results are clearly inconsistent with the genetic drift theory postulated in 1931 by Sewall Wright or the neutral theory of molecular evolution, even in its milder form of near neutrality (Ohta and Gillespie, 1996). Our results suggest that natural selection, through environmental range and stress, in space or time or both appears to be an important genetic differentiating evolutionary force at the protein and DNA levels (i.e., driving molecular evolution) (Nevo, 1998). Similar results were obtained in extensive ecological genetic studies in wild wheat in 1989 by Nevo and Beiles and in subterranean mole rats across the Near East and Asia Minor (Nevo, 1999). Genetic diversity either in the inbreeding wild cereals aboveground or in outbreeding mole rats underground is primarily determined by climatic selection. Similar conclusions—that is, that natural selection on allozyme polymorphism may be intense with selection coefficients in the range of 0.1–0.8—have also been reached for many model organisms (Mitton, 1997). Remarkably, the abundant genetic diversity found in natural populations of plants and animals provides immense genetic resources for crop improvement in agriculture, biotechnology, and gene therapy in medicine.

Genetic Parallelism and Ecological Stress in Subterranean Mammals

Parallel genetic patterns of increasing diversities (primarily southward and secondarily eastward) were shown in 1996 by Nevo and colleagues in Israeli subterranean mole rats for diploid chromosome numbers, nuclear allozyme, mtDNA, RAPDs, minisatellites, chiasma frequency, and organismal traits and also for approximately 200 genes by AFLP (in preparation). Clearly, genomic molecular evolution of nuclear and extranuclear diversities of proteins and DNA, coding and noncoding sequences, in subterranean mole rats, generally assumed to be neutral by proponents of the neutral theory of molecular evolution, is *nonrandom* and correlated positively with increasing aridity stress and climatic unpredictability. The increase in heterozygosity is primarily toward the southern and eastern deserts, where temporal stressful climatic fluctuations and unpredictability climax. Adaptive climatic selection appears to be the prime driving force in both molecular and organismal evolution of mole rats. Natural selection through fluctuating environmental stress (i.e., wider climatic niche) appears to be a major cause of molecular evolution and genome organization in subterranean mole rats, similar to the pattern found in aboveground organisms facing stressful unpredictable climates, across the same transect of increasing aridity stress (Nevo, 1999).

Local Population Genetics Studies across Sharp Ecological Contrasts

Past Microgeographic Studies in Plants and Animals

Microgeographical studies in nature, particularly those involving sharp contrasts (climatic, thermal, edaphic, geologic, topographic, and chemical), provide remarkable long-term experiments, representing small-scale natural laboratories, for analyzing ecologic genetic population dynamics. Since 1975, we have conducted several microgeographic studies in nature (Nevo, 1998). We employed thermal (high vs low temperature) and chemical (polluted vs clean areas) stresses in marine balanids and conducted an extensive research program on inorganic and organic pollution and its effect on allozyme diversity in marine organisms. Likewise, we examined edaphic (terra rossa vs basalt soil and rock vs deep soil) and microclimatic (sun vs shade and high vs low solar radiation) stresses in wild barley (*Hordeum spontaneum*), wild emmer wheat (*Triticum dicoccoides*), and *Aegllops peregrina*. The conclusions of these studies all point inferentially to natural selection as a major differentiating factor of qualitative and quantitative patterns of genetic diversity at single loci, but primarily at multilocus structures and genome organization (Nevo, 1998, 1999).

The following is the summary of our extensive studies in wild cereals (wild barley and wild emmer wheat) that have been studied spatiotemporally since 1975 at the Institute of Evolution at the protein and DNA levels on micro- and macrogeographical scales. We found significant climatic and soil divergence at four microsites as well as in our macrogeographic studies in Israel and the Near East (Nevo, 1998). All the foregoing studies illustrate massive genetic *nonrandom* divergence at single-locus, two-locus, and multilocus structures of genome organization, with specific and unique alleles and levels of genetic diversity correlated with ecological heterogeneity, niche breadth, stress, and the largely parallel patterns of allozymes, RAPDs, and microsatellites despite increasing differences in levels of genetic diversity in this order (**Figure 4**). Recently, we have shown the operation of edaphic selection on microsatellites across the microsites of Yehudiyya, Tabigha, and Ammiad (**Figure 5**). Neither random drift nor gene flow can explain the levels and divergence of genetic diversity at the protein level or at the coding and noncoding DNA regions. The data strongly suggest that natural selection overrides the homogenizing effect of gene flow, maintains molecular polymorphisms in accordance with divergent ecologies, and orients molecular evolution. How general is this genetic divergence pattern of wild cereals, which are predominantly selfers, across life as a whole, primarily in outcrosser plants and nonsedentary animals?

In 1992, we embarked upon a long-term multidisciplinary project at the "Evolution Canyon" microsite. It aims to elucidate the causes of biodiversity and genome evolution at the molecular and organismal levels in diverse groups of organisms across phylogeny from cyanobacteria to mammals, representing a microcosm of life.

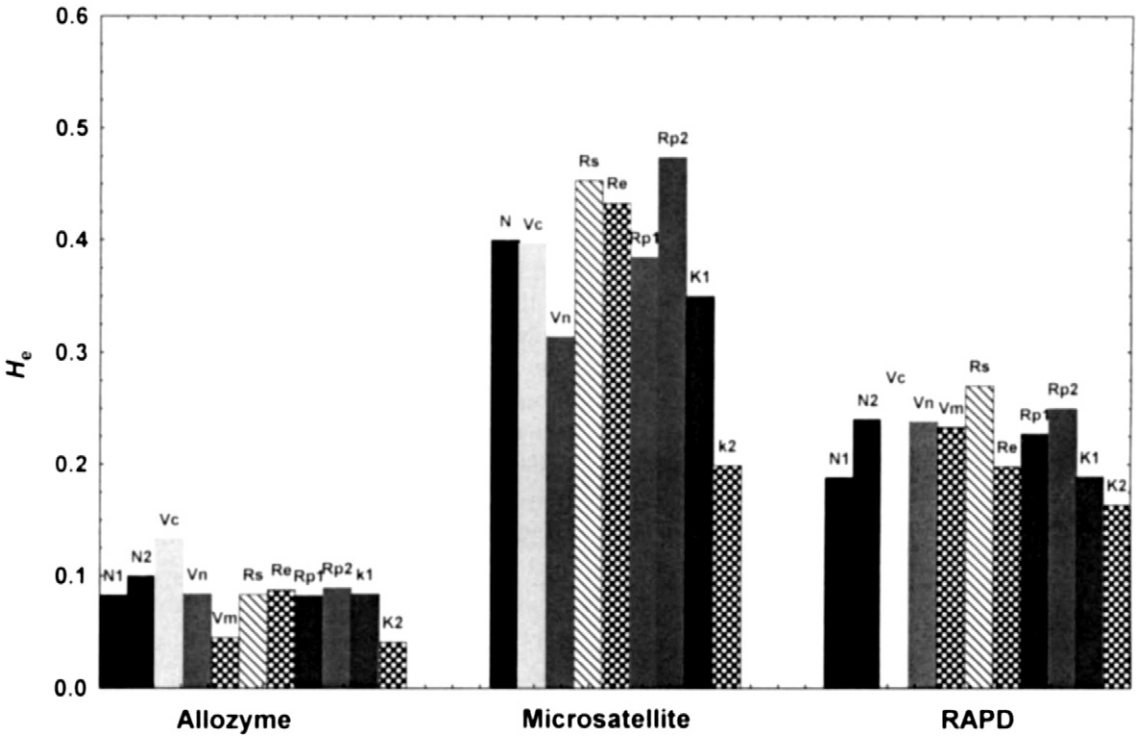


Figure 4 Genetic diversity, H_e, of wild emmer wheat, *Triticum dicocoides*, in four habitats (N, north; V, valley; R, ridge; K, Karst) divided into 11 microhabitats in the Ammiad microsite, Upper Galilee, as reflected in allozymes, microsatellites, and RAPDs (Li *et al.*, 2000a). Note the much higher levels of H_e in microsatellites and RAPDs compared with allozymes; likewise, note the similarity of all systems, but primarily note the high and significant correlation between microsatellites and RAPDs ($r_s = 0.85$, $p = 0.003$; see text). VC, main valley center; VM, main valley margins; VN, narrow valley; N_1, upper north-facing (moderate) slope; N_2, middle north-facing (steep) slope; RE, ridge, east-facing slope; RP_1, ridge, shoulder of plateau; RP_2, ridge, top of plateau; RS, ridge, south-facing slope; K_1, upper karst; K_2, lower karst (reproduced with permission from Li YC, Krugman T, Fahima T, *et al.* (2000a) *Parallel Microgeographic Patterns of Genetic Diversity and Divergences Revealed by Allozyme, RAPD, and Microsatellites in Triticum dicoccoides at Ammiad, Israel*. Netherlands: Springer).

Evolution in Action across Phylogeny caused by Climatic Stresses: ''Evolution Canyon''

Biodiversity and genome evolution, and the relative importance of forces driving evolution, are critically tested at the "Evolution Canyon" microsite (Lower Nahal Oren, Mount Carmel), a dynamic ongoing microscale research program (Nevo, 1997). Our aim is to draw generalizations across life in genotyes and phenotypes and to highlight controversial and unresolved problems of biological evolution. The opposed slopes of "Evolution Canyon" display dramatic physical and biotic contrasts on a microscale (**Figure 6**). Higher solar radiation (up to 600% more) on the south-facing slope (SFS) makes it warmer, drier, and spatiotemporally more heterogeneous and fluctuating than the north-facing slope (NFS). The two slopes are separated by only 100 m (at bottom) and 400 m (at top). The SFS represents an Afroasian savanna park forest, whereas the NFS represents a dense Euroasian live oak maquis forest.

The tropical Afroasian warm and xeric SFS savanna displays wider ecological heterogeneity and higher stress for temperate terrestrial organisms. As predicted, genetic diversity was higher on the ecologically more heterogeneous (wider niche) and stressful SFS in 9 of 11 tested temperate model species (lichen, wild barley, 2 species of land snails, earthworm, diplopod, 3 beetles, and 2 rodents) (**Figure 7**). In 6 species, heterozygosity was negatively correlated with population size, i.e., population abundance was lower, whereas genetic diversity was higher on the SFS. This negates the positive correlation expected between H and N_e by neutrality theory (Kimura, 1983). We have shown in wild barley, *H. spontaneum*, that RAPD and sequence-tagged site PCR analysis mirror allozyme interslope patterns.

Remarkably, heritable mutation and recombination rates were several-fold higher in the soil fungus *Sordaria fimicola* and in the fruit fly *D. melanogaster* on the more stressful SFS. Importantly, in terrestrial taxa species richness was higher on the SFS, and interslope adaptive complexes were demonstrated in land snails, *Drosophila*, wild barley, and mice, contributing to differential fitness in morphology, physiology, behavior, and life history. Transplant experiments showed inter-and intraslope-specific fitness components matching the differential ecological stresses (all studies mentioned in this section are cited in Nevo, 1997). Similar results abound in local genetic differentiation in plants (Linhart and Grant, 1996).

We hypothesized that microclimatic diversifying natural selection, overriding migration and stochasticity, appears here as a major evolutionary driving force of genotypic and phenotypic evolution. The "Evolution Canyon" microsite provides a fertile microgeographic critical testing model of biodiversity evolution due to its sharp microscale ecological slope contrasts, both in niche width and in climatic stress.

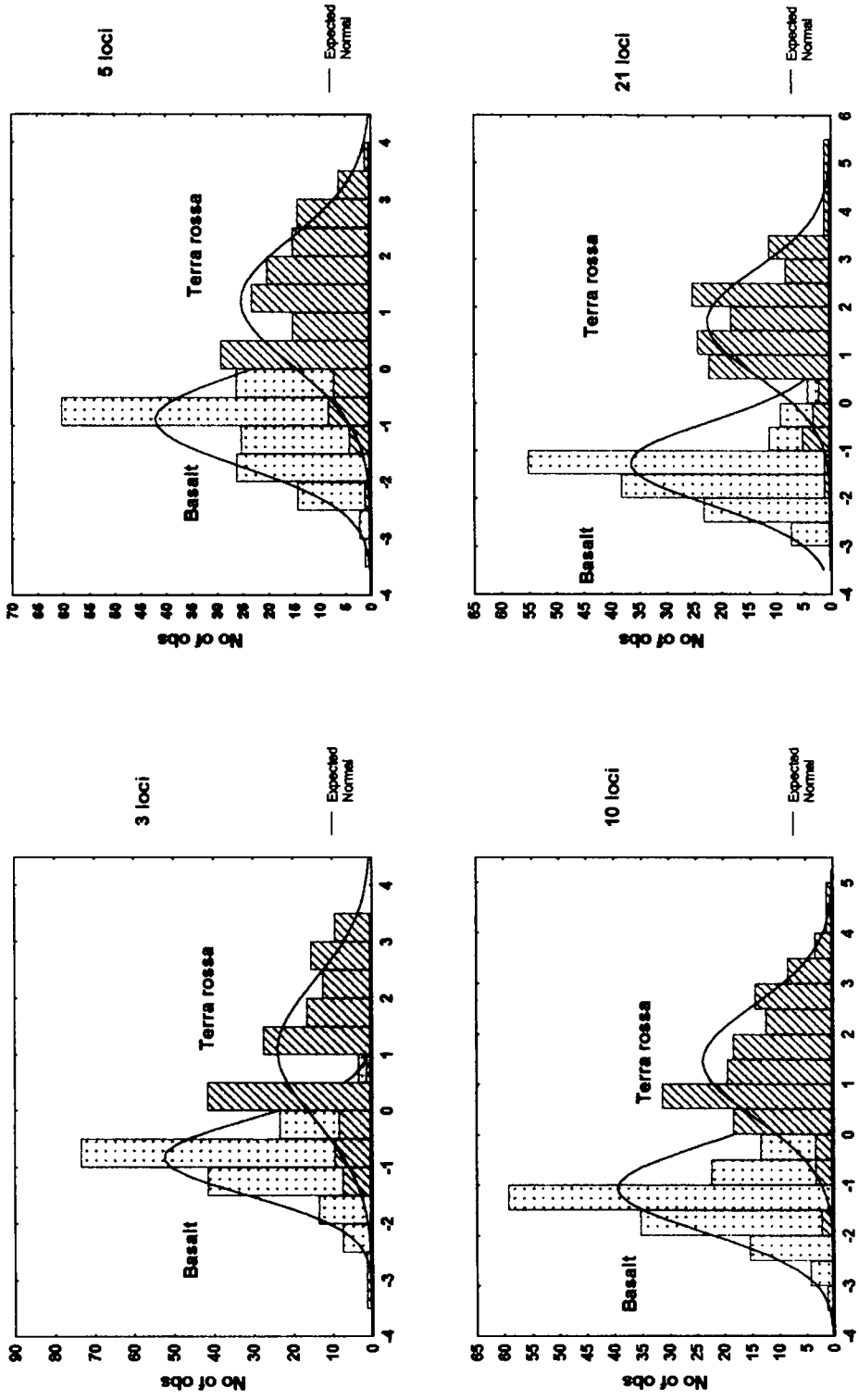


Figure 5 Discriminant analysis between two soil types, basalt and terra rossa of wild emmer wheat (*Triticum dicoccoides*), across three microsites (Yehudiyya, Tabigha, and Ammiad) north of the Lake of Galilee, Israel. This analysis is based on 2, 5, 10 and microsatellite loci, classifying correctly 21, 83, 88, 90, and 91% of plants into their original soil types (based on data from Li YC, Fahima T, Korol AB, *et al.* (2000b) Microsatellite diversity correlated with ecological-edaphic and genetic factors in three micro-sites of wild emmer wheat in north Israel. *Mol. Biol. Evol.* 15: 5).

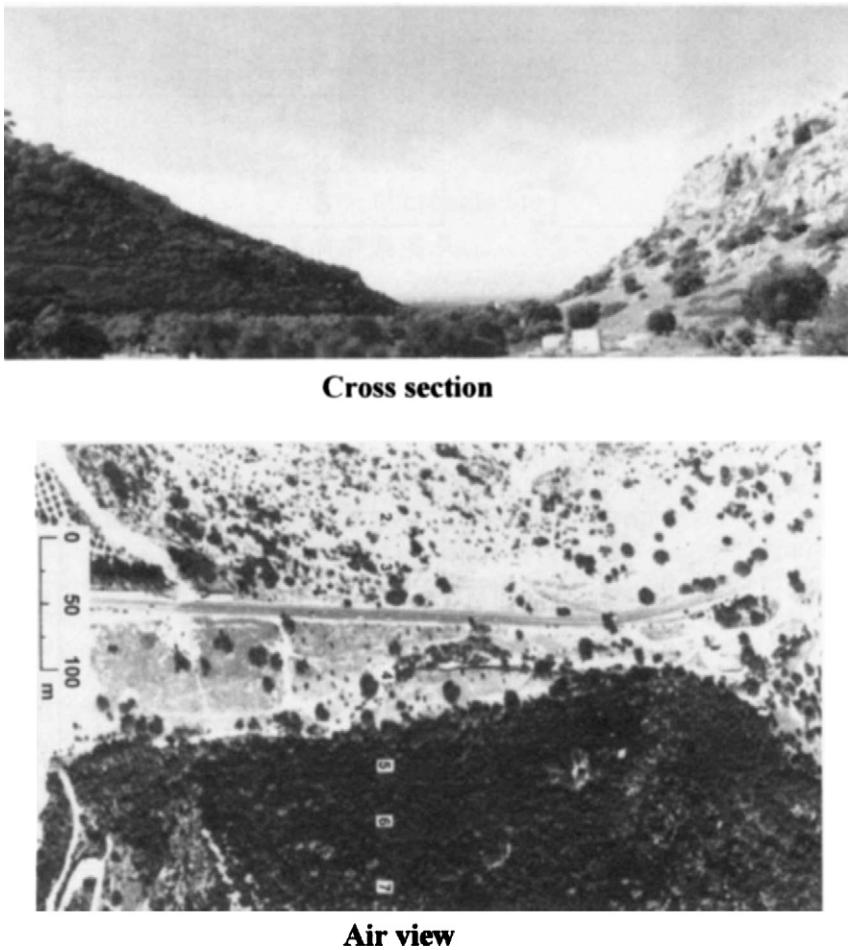


Figure 6 "Evolution Canyon," Lower Nahal Oren, Mount Carmel, Israel Note the plant formation on opposed slopes. The lush, green "European" live oak maquis forest on the temperate, cool-mesic, north-facing slope sharply contrasts with the open park forest of warm-xeric, tropical "Afroasian" savanna on the south-facing slope (SFS). (Top) Cross section; (bottom) air view with the seven experimental stations—three on the SFS (1–3), one at the valley bottom (4), and three on the NFS (5–7) (reproduced from Nevo E (1997) Evolution in action across phylogeny caused by microclimatic stresses at "Evolution Canyon". *Theor. Population Biol.* 52: 231–243).

The Nature of Allozyme and DNA Diversities

Critical laboratory experiments on marine model organisms have shown that allozymic diversity is adaptive and responds rapidly to environmental inorganic and organic pollution stresses (Nevo, 1998); a large amount of allozymic diversity can be gained or lost rapidly. Our experiments demonstrated unequivocally that ecological stresses of thermal and chemical pollution affect the level and pattern of allozyme allele frequencies and heterozygosity of marine organisms, thereby corroborating the environmental theory of genetic diversity. The survivorship values of heterozygotes and homozygote allozyme genotypes vary in accordance with the pollutant concentration. The broader the ecological niche, the higher the heterozygosity. Furthermore, heterozygote sites mutate more frequently than equivalent homozygous sites, possibly because mismatch repair between homologous chromosomes during meiosis provides extra opportunities to mutate. At medium levels of pollution, heterozygotes seem to be superior, whereas at high pollution levels specific homozygotes become superior. Our laboratory results on mercury pollution were confirmed in the sea. Remarkably, Schulman and colleagues described dynamic and rapid changes in genome size mediated by an increase or decrease in copy number of the retrotransposon BARE-1 in wild barley, *H. spontaneum*, both regionally and locally in Israel were described in 1999 by Schulman and colleagues.

Evolutionary Significance of Molecular Polymorphisms: Summary and Evidence

The foregoing macro- and microscale field and laboratory evidence across phylogeny displays massive genetic parallelisms to environmental heterogeneity and diverse stresses, physical (thermal, climatic, and edaphic) and biotic (pathogens, competitors, vegetation, and habitats). In general, higher levels of genetic polymorphisms, at both the protein and the DNA levels, occur under more stressful and variable environments. This is highlighted by the enormous diversity recorded in the vertebrate major histocompatibility complex (Bodmer

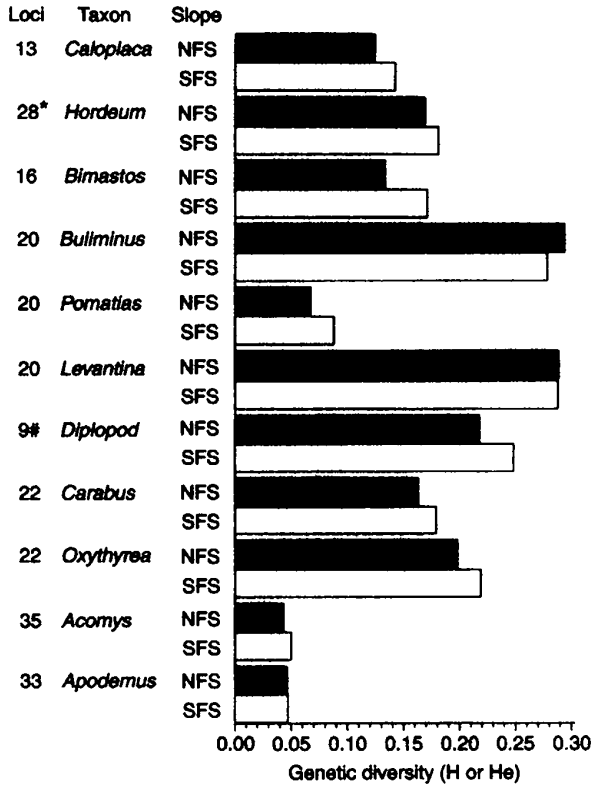


Figure 7 Genetic (allozyme) diversity in major plant and animal taxa on the opposed slopes of "Evolution Canyon"—the mesic, temperate, and mild north-facing slope (NFS) and the xeric, tropical, and stressful south-facing slope (SFS). *Caloplaca aurantia* (lichen), *Hordeum spontaneum* (wild barley), *Bimastos syriacus* (earthworm), *Buliminus labrosus*, *Pomatias olivieri*, and *Levantina caesareana* (land snails), *Tetrarthrosoma syriaca* (diplopod), *Carabus hemprichi* and *Oxythyrea noemi* (beetles), and *Acomys cahirinus* and *Apodemus mystacinus* (rodents) are presented. The genetic diversities represent allozyme heterozygosity observed (H) or expected (H_e). *Comparison between the highest stations (H_e); ♯, *Tetrarthrosoma syriacum* (reproduced from Nevo E (1997) Evolution in action across phylogeny caused by microclimatic stresses at "Evolution Canyon". *Theor. Population Biol.* 52: 231–243).

and Bodmer as cited in Wasser, 1999) and immunoglobulin gene families, resistance genes in plants, against the immense pathogen diversity, and in the regulation of clock genes. Linkage disequilibria and genome organization (Nevo, 1998) are correlated with higher environmental heterogeneity (also called the niche width variation hypothesis by Van Valen, which predicts positive correlation between niche width and morphologic variation) and stress (**Figures 3** and **7**). Similarly, mutation and recombination rates in the soil fungus *Sordaria fimicola* and in male *Drosophila melanogaster* at "Evolution Canyon" were both higher on the ecologically more heterogeneous and stressful SFS. Ecological heterogeneity and stress appear to be major determinants of the level of genetic diversity in nature, as was demonstrated in numerous organisms (Mitton, 1997). Physical and biotic ecological variables are better predictors of genetic diversity than demographic variables of population and species size. In fact, the levels of genetic diversity are often negatively correlated with effective population size, negating neutrality theory predictions. Generally, genetic polymorphism is positively correlated with niche width, or the level of ecological heterogeneity and stress, as emerged from our global, regional, and local analyses and from critical laboratory experiments. A brief theoretical discussion follows.

Theory

Mechanisms of Generating Genetic Diversity

It is beyond the scope of this article to describe the mutation mechanisms that generate diversity. Suffice it to say that a host of mechanisms provide a permanent input of diversity into natural populations, including point mutations, deletions, additions, recombination, gene conversion, and mismatch repair. Mutation rates also vary dramatically among organisms and genes and under diverse ecological stresses. Evidence from bacterial genetics (Casjens, 1998) suggests that organisms can resist stress through mutational changes or acquisition of preevolved functions via horizontal transfer. Activating mutagenic response and inhibiting antimutagenic activities (e.g., mismatch repair) could be important in recruiting genetic diversity as an adaptation to stress. Stress-induced increases in mutation rates enhance genetic polymorphism, primarily in genomic "hot spots."

Significance of Genetic Diversity

The functional value of genetic diversity within populations has been confirmed in a series of studies. Clearly, this rich pool of diversity provides the resource for continuous selection of adapted genotypes. Reduced genetic diversity may not only compromise the capacity of an impacted population for genetic adaptation in the face of further environmental challenge but also may result in increased energy requirements, lower production efficiency, and reduced homeostasis and reproductive output. These metabolic consequences of reduced genetic polymorphism would further lower that population's potential for survival under lethal conditions of contaminant exposure and also affect the genetic makeup of populations through differential reproduction under conditions of sublethal stress. In laboratory studies with population cages, higher levels of allozyme and additive genetic diversity are generally maintained in cages with greater heterogeneity, as is also generally true in nature (Mitton, 1997). Genetic diversity makes it possible to establish genetic distances within and between species, identify species and strains by unique and specific genetic profiles, establish phylogenetic trees, reinforce conservation biology and management programs, and have profound implications for plant and animal husbandry in agriculture and medicine.

Maintenance of Genetic Diversity in Nature

Explaining the maintenance of genetic diversity in natural populations has been a central problem of evolutionary genetics since the discovery of abundant protein polymorphisms in nature (Lewontin, 1974; Nevo, 1978, 1988, 1998; Kimura, 1983; Gillespie, 1991; Avise, 1994; Mitton, 1997). It is now

clear that stochastic and bottleneck explanations, relating heterozygosity primarily to population size effects, are not realistic. Populations are rarely at equilibrium because environmental conditions fluctuate with abiotic (e.g., climate and seasons) and a host of biotic (e.g., parasites, pathogens, and competitors) factors. Correspondingly, genetic diversity can be gained or lost rapidly in dynamically changing natural populations due to ecological stresses. The major models explaining genetic diversity in nature relate primarily to the levels of ecological heterogeneity (niche breadth) and stress, or multiniches, and to environmental grain (Levins, 1968) rather than to effective population size or gene flow. Environmental variability enhances genetic diversity. Natural selection appears to be a major driving evolutionary force maintaining genetic diversity in nature (Bell, 1997). However, heterosis (i.e., overdominance or heterozygote advantage) alone is not a mechanism for maintaining many alleles segregating at a locus, as argued in 1978 by Lewontin, Ginzburg, and Tuliapurkar; hence the resort to spatiotemporal driving forces at multilocus structures.

Natural Selection

Theoretically, spatial and temporal variations of selection ("diversifying selection") could maintain and enhance genetic polymorphisms, although the conditions of their applicability are strongly limited (see cited papers by Levene, Haldane, Karlin, Lande, Felsenstein, Hedrick, Hoekstra, Maynard Smith in Ewens (1979) and Nevo (1998)). Spatial variation appeared more effective than temporal variation, although their joint action could reinforce the maintenance of polymorphism. Most results related to selection of variation in time were derived from the one-locus case. Polymorphism maintenance may be reinforced in the case of two-locus or multilocus structures (Gillespie, 1991; Kirzhner *et al.*, 1998). Fitness components (viability, growth rate, fecundity, mating success, and developmental stability) increase with heterozygosity, suggesting that selection balances diversity at protein polymorphisms (Mitton, 1997). Some of the correlations between heterozygosity and fitness components can be attributed to allozyme loci or loci in strong linkage disequilibrium with them. This positive relationship between heterozygosity and fitness is not only expected theoretically but also abundantly documented empirically (Mitton, 1997). The selective mechanism is much more effective in promoting genetic diversity if carriers of the alternative alleles are able to select the niche in which their fitness is greatest as argued in 1977 by Taylor and Powell and by Nevo and colleagues in 2000 for mole crickets. For selection theories of molecular evolution and polymorphism based on the combined forces of natural selection, genetic drift, and mutation, see Gillespie (1991). For a discussion on the development of neutral and nearly neutral theories, see Ohta and Gillespie (1996). A series of articles dealing with natural selection and Darwinian fitness in nature appear in Wasser (1999).

In a series of articles and books during the period 1931–1978, Sewall Wright developed his "adaptive landscape" theory. In this model, populations are imagined to spend most of their time on selective peaks, with genetic drift providing the push when a population jumps from one peak to another. In contrast, Fisher's (1958) model on the nature of adaptation assumes that a population is never exactly at its optimal phenotype. Likewise, in Gillespie's (1991, pp. 291–305) random environment selection theories, there are no analogs to adaptive peaks because the adaptive landscape is changing faster than the genetic system: "The population is always running uphill, but the peak is always two steps ahead. All the population ever sees, in effect, is the side of the mountain. Should it stop evolving, it will face extinction." This view is related to Fisher's model and to the red queen hypothesis of Van Valen (1973), which assumes that if a population does not continue to adapt at the same rate as its competitors or environmental deterioration it may become extinct.

Stabilizing Selection in Cyclical Environments

Stabilizing selection for an intermediate optimum is generally considered to rapidly deplete genetic diversity in quantitative traits, with increased number of loci. In contrast to previous conclusions, we found in both haploids and diploids, in the case of an additive two-locus model, that stabilizing selection with cyclically moving optimum may be an efficient factor in protecting polymorphisms for linked loci additively affecting the selected trait (Korol *et al.*, 1994; Kirzhner *et al.*, 1998). We proved that within the same class of fitness functions, nonequal gene action and/or dominance effect for one or both loci may lead to local polymorphism stability with substantial polymorphism-attracting domain. A higher intensity of selection could result in two forms of polymorphic limiting behavior: (i) the usually expected forced cycles with a period equal to that of environmental changes and (ii) "supercycles," which are nondamping autooscillations with a period composed of hundreds of forced oscillations.

We have demonstrated (Kirzhner *et al.*, 1998) that a multilocus system subjected to stabilizing selection with cyclically moving optimum can generate ubiquitous complex limiting behavior, including supercycles, T cycles, and chaotic-like phenomena. This mode of multilocus dynamics far exceeds the potential of complex dynamics attainable under ordinary selection models resulting in simple behavior. It may represent a novel evolutionary mechanism increasing genetic polymorphism over long-term periods (Kirzhner *et al.*, 1998).

Selection versus Genetic Drift in Small Isolated Populations

Remarkably, our findings of high-level heterozygosity in small (several dozen or 100 individuals) desert isolates of the subterranean mole rats, *Spalax ehrenhergi* superspecies ($2n = 60$), in the northern Negev (Nevo, 1999) contradict both the Wrightian notion of genetic drift in small populations that was elaborated in 1931 and the extreme version of genetic drift, i.e., the neutral theory of molecular evolution (Kimura, 1983). The latter predicts positive correlation between effective population size (N_e) and heterozygosity H, whereas our finding demonstrated the opposite (i.e., the highest level of H in the smallest populations, or a negative correlation between H and N_e). A negative correlation between population size and heterozygosity was also found in several species tested at the "Evolution Canyon" micro-site (Nevo, 1997). Current theoretical models predict fast gene fixation in small panmictic populations without selection, mutation, or gene inflow. Using simple multilocus models, we demonstrated that

moderate stabilizing selection (with stable or fluctuating optimum) for traits controlled by additive genes could oppose random fixation in such isolates during thousands of generations (Nevo *et al*., 1997a, as cited in Nevo, 1999). We also showed that in selection-free models polymorphism persists only for a few hundred generations even under high mutation rates. Our multichromosome models challenge the hitchhiking hypothesis of polymorphism maintenance for many neutral loci due to close linkage with a few selected loci.

Genetic Ecological Diversity and Stress

The foregoing discussion suggests several explanations for the maintenance of genetic diversity subjected to ecological diversity and environmental stress. Spatial and temporal ecological variation, which predominates in nature, is of prime importance in maintaining and enhancing genetic diversity in natural populations. This may be true because different genotypes display varying fitnesses in variable environments and stresses. Recombination frequencies and mutation rates tend to increase under stressful conditions (Hoffmann and Parsons, 1991; Korol *et al*., 1994; Nevo, 1997). Rates of evolutionary change are therefore enhanced in adverse environments, as was demonstrated under controlled laboratory experiments in the case of mercury pollution (Baker *et al*. as cited in Nevo, 1998), under regional aridity stress across the physically stressful Israeli environment, and locally at "Evolution Canyon" because of high solar radiation on the SFS (Nevo, 1997).

Developmental instability may be enhanced under both environmental stress and genomic stress. This could increase genetic diversity under stressful conditions (Hoffmann and Parsons, 1991). Heterozygote advantage tends to increase diversity with stress up to extreme points. However, heterosis alone is not a mechanism for maintaining many alleles segregating at a locus. It is much more likely that stable equilibria for multiple alleles will be best explained by multiple-niche selection. Models of sexual reproduction as an adaptation to resist parasites proposed by Hamilton and colleagues in 1990 may also contribute to sex evolution, recombination, and polymorphism. Finally, the simple model advanced by Kirzhner and colleagues in 1999 of genetic interaction between multiple species on a trait for trait basis governed by abiotic and biotic selection for multilocus quantitative traits opens wide horizons for the evolution of genetic diversity due to species dynamic interactions in nature.

Ecological heterogeneity and stress appear to enhance genetic polymorphisms, particularly in dynamically cycling environments (which can generate complex dynamic-like supercycles and chaotic-like behavior). This mode of multilocus dynamics far exceeds the potential for maintaining genetic polymorphism attainable in ordinary selection models. It may represent a novel evolutionary mechanism increasing genetic polymorphism over long-term periods. This novel mechanism could contribute to the observation that biological diversity has increased over geological time despite the well-known massive extinctions, providing ever-increasing genetic diversity and thus enhancing the evolution of biodiversity.

Acknowledgments

I thank Abraham Korol and Avigdor Beiles for critically reading and commenting on the manuscript. This study was supported by the Israel Discount Bank Chair of Evolutionary Biology and the Ancell–Teicher Research Foundation for Genetics and Molecular Evolution.

See also: Ecological Genetics. Evolution, Theory of. Genes, Description of. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms. Phenotype, a Historical Perspective

References

Avise JC (1994) *Molecular Markers. Natural History and Evolution*. New York: Chapman & Hall.
Bell G (1997) *Selection. The Mechanism of Evolution*. New York: Chapman & Hall.
Bodmer W and McKie R (1994) *The Book of Man. The Quest to Discover Our Genetic Heritage*. London: Abacus.
Casjens S (1998) The diverse and dynamic structure of bacterial genomes. *Annu. Rev. Genet.* 32: 339–377.
Curole AP and Kocher TD (1999) Mitogenomics: Digging deeper with complete mitochondrial genomes. *Trends in Ecology and Evolution* 14: 394–398.
Darwin C (1859) *On the Origin of Species by Means of Natural Selection*. London: Murray.
Ewens WJ (1979) *Mathematical Population Genetics*. Berlin/New York: Springer-Verlag.
Fisher RA (1958) *The Genetical Theory of Natural Selection*. New York: Dover.
Gillespie JHL (1991) *The Causes of Molecular Evolution*. Oxford: Oxford Univ. Press.
Goldstein DB and Schlotterer C (1999) *Microsatellites: Evolution and Applications*. London: Oxford Univ. Press.
Hamrick JL, Linhart YB, and Mitton JB (1979) Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annu. Rev. Ecol. Syst.* 10: 173–200.
Hoffmann AA and Parsons PA (1991) *Evolutionary Genetics and Environmental Stress*. Oxford: Oxford Univ. Press.
Johns GC and Avise JC (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology and Evolution* 15: 1481–1490.
Karlin S, Campbell AM, and Mrazek J (1998) Comparative DNA analysis across diverse genomes. *Annu. Rev. Genet.* 32: 185–225.
Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge, UK: Cambridge Univ. Press.
Kirzhner VM, Korol AB, and Nevo E (1998) Complex limiting behaviour of multilocus genetic systems in cyclical environments. *J. Theor. Biol.* 190: 215–225.
Korol AB, Preygal IA, and Preygal SL (1994) *Recombination Variability in Evolution—Algorithms of Estimation and Population Genetics Models*. London: Chapman & Hall.
Levins R (1968) *Evolution in Changing Environments*. Princeton, NJ: Princeton Univ. Press.
Lewontin RC (1974) *The Genetic Basis of Evolutionary Change*. New York: Columbia Univ. Press.
Li YC, Fahima T, Korol AB, *et al*. (2000b) Microsatellite diversity correlated with ecological-edaphic and genetic factors in three micro-sites of wild emmer wheat in north Israel. *Mol. Biol. Evol.* 15: 5.
Li YC, Krugman T, Fahima T, *et al*. (2000a) *Parallel Microgeographic Patterns of Genetic Diversity and Divergences Revealed by Allozyme, RAPD, and Microsatellites in Triticum dicoccoides at Ammiad, Israel*. Netherlands: Springer.
Linhart YB and Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annu. Rev. Ecol. Syst.* 27: 237–277.
Marra M, *et al*. (1999) An encyclopedia of mouse genes. *Nature Genetics* 21: 191–194.
Martin W, Stoebe S, Goremykin V, Hansmann S, Hasegawa M, and Kowallik K (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393: 162–165.
Mitton JB (1997) *Selection in Natural Populations*. Oxford: Oxford Univ. Press.

Nevo E (1978) Genetic variation in natural populations: Patterns and theory. *Theor. Population Biol.* 13: 121–177.
Nevo E (1988) Genetic diversity in nature: Patterns and theory. *Evol. Biol.* 23: 217–247.
Nevo E (1997) Evolution in action across phylogeny caused by microclimatic stresses at "Evolution Canyon". *Theor. Population Biol.* 52: 231–243.
Nevo E (1998) Molecular evolution and ecological stress at global, regional and local scales: The Israeli perspective. *J. Exp. Zool.* 282: 95–119.
Nevo E (1999) *Mosaic Evolution of Subterranean Mammals: Regression, Progression and Global Convergence.* Oxford: Oxford Univ. Press.
Nevo E, Beiles A, and Ben-Shlomo R (1984) The evolutionary significance of genetic diversity: Ecological demographic and life history correlates. *Lect. Notes Biomath.* 53: 13–213.
Ohta T and Gillespie JH (1996) Development of neutral and nearly neutral theories. *Theor. Population Biol.* 49: 128–142.
Ramshaw JAM, Coyne JA, and Lewontin RC (1979) The sensitivity of gel electrophoresis as a detector of genetic variation. *Genetics* 93: 1019–1037.
Schena M, Heller AR, Theriault TP, Konrad K, Lachenmeier E, and Davis RW (1998) Microarrays: biotechnology's discovery platform for functional genomics. *Trends in Biotechnology* 16: 301–306.
Van Valen L (1973) A new evolutionary law. *Evol. Theory* 1: 1–30.
Wasser SP (ed.) (1999) *Evolutionary Theory and Processes: Modern Perspectives.* Dordrecht: Kluwer.
Watson JD, Hopkins NH, Roberts JW, Steitz JA, and Weiner AM (1987) *Molecular Biology of the Gene,* Vols. 1 and 2, Menlo Park, CA: Benjamin/Cummings.

Geologic Time, History of Biodiversity in

James W Valentine, University of California, Berkeley, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 215–231, © 2001, Elsevier Inc.

Glossary

Alpha diversity Diversity of species within a single habitat.

Beta diversity Measure of the rate and extent of change in species along a given habitat or physiographic gradient.

Disparity Range of diverse morphological architectures present in higher taxa such as classes, phyla, and kingdoms; it is used to explain the origins and maintenance of this diversity of life-forms and body plans.

Taxonomic richness Number of species, genera, and families from a given time period or fossil excavation.

Biodiversity in the Fossil Record

There are two distinctive though related aspects to paleobiodiversity studies. One approach focuses on the number of taxa present during a certain geologic time, or over a certain geologic time span. The taxonomic categories involved are usually species, genera, or families, and this sort of diversity is commonly termed taxonomic *richness*. The ultimate aim of most such studies is to understand and explain the origin and maintenance of patterns of species richness in space and time; when genera or families are used it is usually as proxies for species. The evolution of species richness has been studied since Darwin's time and continues to be a highly productive area of research. The second aspect deals with diversity, not of numbers of species, but of different types of organisms. The taxonomic categories involved are the higher ones, such as kingdoms, phyla, and classes, and the aim is to understand and explain the origins and maintenance of *disparity*—the range of diverse morphological architectures that is present. The major patterns of disparity over geologic time have been described, chiefly using taxonomic categories, as proxies for morphological distance. However, the evolution of that disparity has received much less attention than the evolution of richness, partly because it was thought that disparity could be explained by extrapolation from the processes that create richness. With the rise of molecular developmental studies, however, the genetic basis for disparity is becoming understood, and it has become possible to decouple studies of the origin of disparity from those of richness.

The ecological regulation of local biodiversity is not well understood, but empirical patterns of species distributions and richnesses indicate some of the factors governing species accommodation on regional and global scales. These large scales are of particular importance to tracing biodiversity through time. A basic measure of species richness is *alpha diversity*, the number of species found at a given locality or in a single sample. A second sort of species richness, *beta diversity*, is a measure of the additional species that are found at a second locality or in a second sample. Alpha diversity was taken to indicate the amount of species packing (reflecting ecological niche partitioning) at a locality, and beta diversity to indicate the extent of habitat differences between localities. In practice, however, different species are found at a given locality at different times, varying with season, weather, or other factors that affect successful reproduction or colonization at the locality. Alpha diversity can therefore vary significantly from time to time, in which case beta diversity is affected as well.

The global environment is quite heterogeneous, a mosaic of habitats separated by gentle gradients in some places and by sharp boundaries in others and further broken into disjunct regions, such as different continental masses or remote islands. Beta diversity will vary according to whether two samples are from adjoining areas within the same general habitat type (i.e., represent the same biotic community), or are from distinctive habitat types (represent different communities), or are from distinctive climatic or physically disjunct regions (represent different biotic provinces). Though beta diversity is sensitive to differences among samples within a community, it is not as useful a measure between communities; rather, a more appropriate measure is the number of species added as entire communities are combined—a sort of super beta diversity. To capture global biodiversity patterns, a hierarchy of beta-type measures is required; differences between bioprovinces can be evaluated by the numbers of species added when combining one entire bioprovince with the next, or similar evaluations can evaluate differences between bioprovinces can be made between entire realms. The pattern of environmental heterogeneity in the world has varied continuously over geologic time, thus affecting the pattern of accommodation of biodiversity.

Evolution and Biodiversity

Richness

Speciation is produced by isolation of populations, severing the flow of genes between them and permitting the evolution of differences in their features, underlain by differences in genes, gene frequencies, and gene organizations that arise between the populations. Once sister population crosses are unable to produce fertile offspring, speciation is complete. Environments favoring more isolation among populations are therefore more likely to produce rich biotas, other things

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00066-6

being equal. A world broken into many continents favors global species richness, with a nearly separate biota on each continent. A continental landscape consisting of a great variety of climates and habitat types favors more species richness than does a monotonous landscape, for species tend to become well adapted to particular habitats, and the more habitat types available, the more species. A landscape in which any given habitat type is scattered among other types, which therefore provide barriers to gene flow, favors more species richness than a landscape where each habitat type occurs in a large tract. In short, environmental heterogeneity clearly promotes isolation and therefore species richness through speciation. Other factors also promote speciation, but they are not so clear nor so well understood. For example, low-latitude, tropical regions hold many more species than midlatitude temperate regions, and these latter hold more than high-latitude arctic/antarctic regions, but the mechanisms that link the obvious climatic differences to this latitudinal gradient in species richness are not yet understood. It is clear that the reduction of environmental heterogeneity by human activities is lowering the capacity of the globe to support species.

Disparity

The genetic differences producing morphological disparity are not based so much on differences in genes and gene frequencies as on gene organization. Individuals belonging to different higher taxa, such as phyla, differ from their very early stages, producing different adult body plans as they develop. In animals, the developmental regulatory genes that are responsible for organizing the many disparate body plans are surprisingly similar. The animal phyla originated over half a billion years ago, so similarities in those developmental control genes have been conserved since that time. Even members of other kingdoms, such as fungi and plants, share some similarities with animals in their developmental control genes, which therefore have been conserved for over a billion years. Some key developmental genes have been found in all phyla that have been investigated, though they may play a variety of roles. These genes are responsible for regulating still other genes. The regulatory genes govern cascades of gene expression that mediate cellular differentiation and organogenesis, the hallmarks of multicellular body plans, and control the final body architecture. The evolution of disparity has involved the evolution of these basic developmental control systems. Thus, although the genes are similar among phyla, many of the gene expression pathways are different within each phylum and class and so produce distinctive body plans.

In animals, the use of genes that mediate the development of numbers of cell types and morphologies seems to have originated in a spongelike organism. Animal genomes contain two to three times as many genes as most unicellular protists, suggesting that one or more early, massive gene duplication event(s) provided the genetic resources that were necessary to regulate the development of complex morphologies. Vertebrates have approximately four times as many genes as an average invertebrate, suggesting that some massive gene duplication events in early chordate lineages provided a genome of the requisite size to permit the evolution of the very complex vertebrate bodies. Regulatory genetic systems have evolved in land plants as well, quite independently, mediating the disparity found in plant anatomy, but as yet these systems are not as well known as those of animals.

Capturing Fossil Biodiversities

Problems of Sampling Fossil Biodiversity

It is not easy to become a fossil; the remains of only an exceedingly small fraction of all the organisms that have lived have survived to provide information on the nature of life through geologic time. Therefore we must treat the fossil record of biodiversity as a huge sampling problem. One approach would be to estimate how many samples from how many places it would take to adequately represent present biodiversity, and then to compare this ideal sampling pattern with the sorts of fossil samples with which nature has presented us. To estimate today's biodiversity we would need to sample each community in each province on each continent in each major realm, such as terrestrial plains and moutains, continental shelves, and each major oceanic water mass and deep-sea region. We would not have to find every last species, but we would have to understand the patterns of species' distributions well enough to calculate biodiversity from our samples. Although valiant attempts have been made, the problem is so large that estimates of modern biodiversity are not closely constrained. Assessing fossil samples is not easy either. We need to know how densely the biota has been sampled regionally and ecologically by nature in presenting us with the fossil assemblages. And for a given fossil assemblage, we must estimate how completely the living biota was sampled, without knowing how diverse it was in the first place. We also need to know how different ancient biotas were from region to region, for it is not uncommon for large regions to lack fossils for given time periods, and their diversities must then be estimated from assumptions about global patterns. Finally, we must be able to estimate the temporal density of the fossil record and make allowances for intervals of time that lack records.

Completeness of the Rock Record

The rocks from which fossils have been recovered are nearly all exposed on land and have therefore been subjected to erosion, and many fossiliferous formations have been stripped away and lost. The first question to ask when reconstructing fossil biodiversities is how much of the original rock record remains. The completeness of a sequence of rocks is independent of its age, relating rather to its depositional history. In general, sedimentary sequences contain many small gaps but increasingly fewer gaps of longer durations. For a given sedimentary sequence, then, fossil taxa with very short geologic ranges may have been present only during a gap, whereas longer-ranging forms can "jump" most gaps and may be captured as fossils. Fortunately, the gaps in one region are commonly represented by sediments elsewhere, though a few periods of time are poorly represented in most regions. For example, few marine sediments are preserved on continents from times when sea

levels were particularly low, such as at the end of the Permian and beginning of the Triassic periods.

Terrestrial sedimentary rocks, deposited as they are above sea level, are subjected to very active processes of erosion, which remove much of the sediment in rivers or in winds and deliver it to the sea. Preservation of terrestrial formations is thus restricted chiefly to downwarped epicontinental basins or interior continental platforms, where there is some protection from the main erosional forces. Such settings tend to be more localized and geologically shorter-lived than nearshore marine sites. Furthermore, terrestrial animal remains are also exposed to decay and erosion in the subaerial environment and tend to be destroyed more easily than the shells of marine forms. Thus, terrestrial rock sequences are less complete than marine ones, and terrestrial animals are not as well sampled by the fossil record as marine animals.

Incompleteness of the Fossil Record

Organisms differ greatly in their potential for preservation as fossils. In general, entirely soft-bodied, or small, or rare, or geographically restricted species of short durations have fewer opportunities for preservation than their opposites: a large, abundant, widespread species that lived for many millions of years and has a very durable skeleton is most likely to be represented by fossils. Some phyla, such as Mollusca (clams and snails, etc.) are well represented in marine sediments, whereas others, such as Platyhelminthes (flatworms), though at least as old as the Mollusca and quite rich in species, have left no recognized body fossils. The biodiversity of the living biotas from which fossil assemblages are recruited must usually have been considerably greater than the numbers of fossil taxa that are found. However, it seems that most species with durable skeletons are in fact originally captured in sediments. For example, a study of fossilization of living species of marine bivalves (Mollusca) of the Californian bioprovince found that about 80% of the species, 84% of the genera, and 90% of the families were represented by fossils in sediments of Pleistocene age (the last 1.6 million years). As expected, the missing species were small, thin-shelled, and rare forms, and the missing genera and families were made up of such species.

In richness studies of taxa with high preservation potentials, such as shelled mollusks, fossil assemblages actually have an advantage over samples of the living biota. The fossils are usually time-averaged, that is, they represent the accumulation at a given locality of many generations and may include individuals of species that have occupied the region for only some fraction of the time represented. Thus, a fossil assemblage can sample the species richness that was found in a given area far better than a sample of alpha diversity in a living community. The fossils represent a sampling through time, often over hundreds to thousands of years.

Foote and Sepkoski (1999) devised a method that estimates the probability of preservation of fossil higher taxa from a sort of frequency distribution of the observed geologic range lengths of their members. When such an estimate is made for marine bivalves (clams), the probability of preservation of a genus within a stratigraphic interval representing about 5 million years is around 50% (**Figure 1**). The bivalve faunas preserved in the sediments have been significantly degraded. Although some elements of the fossil faunas may have been destroyed by dissolution within the sediments, it seems likely that most of the missing genera were originally present in sediments that were then either removed during intervals of erosion or are not now exposed for collection and study. In other words, although fossils commonly represent a good sample of living associations of easily preservable taxa, not enough such associations are preserved to directly count ancient diversities over lengthy intervals of geologic time.

The completeness of the record of a number of other higher taxa is given in **Figure 1**, where the probability of their generic preservation is plotted against the proportion of the living families that have a fossil record. There are two groups that deviate from the general trend. The Cephalopoda (Mollusca) fall well above the trend because they are represented by rich fossil assemblages of extinct higher taxa and have few living representatives. By contrast, Chondrichthyes (Chordata) fall well below the trend because their living families have a good fossil record, chiefly of teeth found in relatively young sediments, but the earlier fossil genera, represented by bodies, are mostly known only from single localities and thus have low preservation potentials. Thus the two outliers are readily explained, and the correlation otherwise evident between these totally independent data sets suggests that they are reasonably accurate. Because the marine sedimentary rocks of continental shelves combine high completeness and availability, the marine invertebrates with the highest preservation potentials provide the longest, most complete fossil record of any major groups of organisms.

In establishing the geologic range of a fossil taxon, it is almost certain that we have not found the earliest members to evolve or the last individuals to become extinct, because the fossil record is so incomplete. It is possible, however, to judge the probable extent of a taxon's range beyond its actual record by how commonly it is found between its first and last occurrences: if it occurs rather continuously within its known range, then when it disappears it is likely to be absent; if it occurs rarely, then it may have existed well before we see it and continued to exist well after our last record. The number of a taxon's occurrences in a given time interval thus permits a statistical estimate of the likely extent of its true geologic range. By taking account the observed preservation potentials of a taxon, estimates of the numbers of that taxon present in a given interval can be augmented by granting range extensions beyond the fossil occurrences. Corrections of this sort will produce better approximations of the richness within geologic intervals than the raw numbers of fossil taxa in collections.

Partial Remedies for the Spotty Fossil Biodiversity Record

The taxonomic hierarchy, based on Linnaean principles, produces many paraphyletic taxa, but it is useful for purposes of biodiversity reconstruction. Even within taxa with the highest preservation potentials, a large proportion have been lost to erosion and other factors. Genera, on the other hand, contain on the average several species, so that a genus is usually more widespread, occurs in more environments, and contains many more individuals than any one of its species. As a

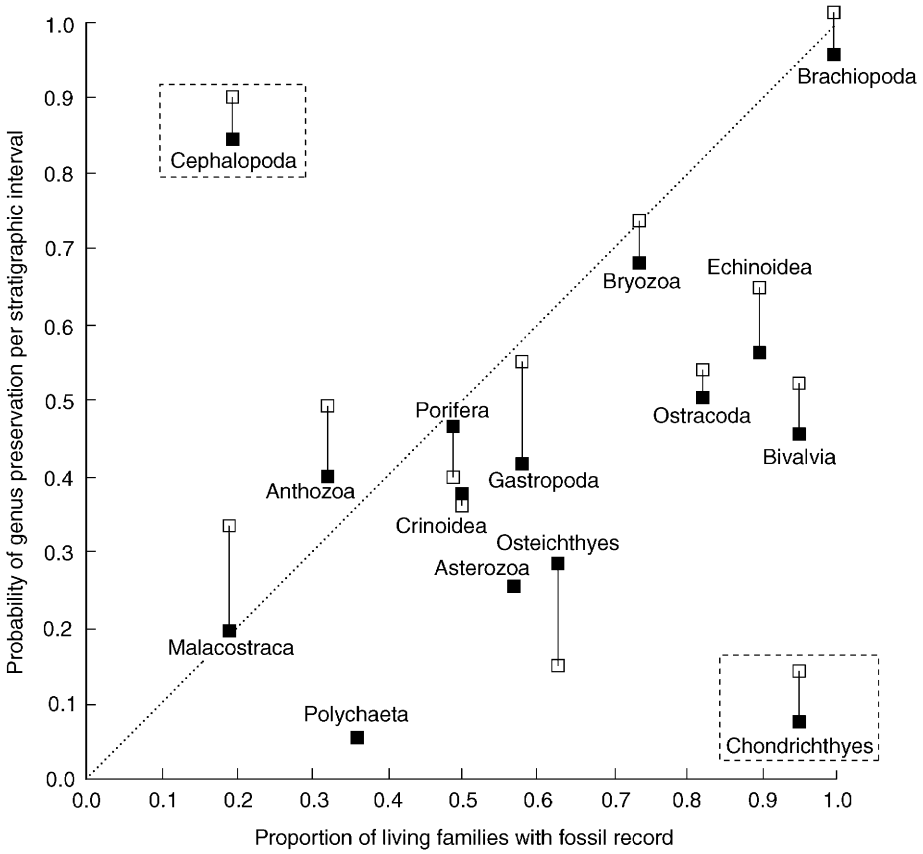


Figure 1 Two measures of completeness for taxa of marine animals: the vertical scale is the calculated probability of the preservation of a genus of the taxa per stratigraphic interval; the horizontal scale is the proportion of living families of the taxa with a fossil record (a measure of the preservation potential of the taxa). There are two points (boxes) for each taxon because two different timescales (with different bins) were used to evaluate binning biases. The significant correlation of these two measures indicates that the estimated completenesses are generally accurate. The Cephalopoda and Chondrichthyes are outliers; see text for discussion. Reprinted from Foote M and Sepkoski Jr. JJ (1999) Absolute measures of the completeness of the fossil record. *Nature* 398: 415, with permission from Nature.

consequence, the fossil representation of genera is more complete than that of species. Families usually include several genera and so are even more completely represented as fossils. Thus by compiling biodiversity data on the generic level, one produces a significantly more complete and somewhat more accurate record than is possible at the species level, and the record at the family level is even better.

There is a price for this taxonomic remedy. The lower levels of taxonomic biodiversity are the more volatile through time, being more sensitive to the environmental changes that affect biodiversity. For example, a wave of extinction may decimate a biota at the species level, but to be recorded in a compilation of family diversity an extinction must extirpate every last species of a family. A small extinction wave might not even be noticed at the family level, considering that fossil data are rather noisy anyway. Of course the time of an extinction, whether recorded at the family, generic, or species level, will be registered in precisely the same place on the geologic timescale. Because species contain reproducing populations, they are in direct ecological contact with environmental parameters, and the volatility of their richness is the most sensitive measure of environmental changes. Higher taxa record the outcome of environmental change more indirectly, that is, when their species first appear, or when all of them disappear from the record. Nevertheless, their behavior in the face of environmental change presumably reflects the broad similarities in physiological and behavioral responses of their species, inherited from common ancestral species. While generic and family biodiversities can make quite useful proxies for species diversity in the fossil record, they are in fact measuring different features, and this should be taken into account in interpretations of fossil richness.

Most individual fossil samples represent a particular ecological community, perhaps somewhat intermixed with rare species that are chiefly found in other communities, where they are common. In some cases, though, the samples represent mixtures of common faunal elements that represent different communities but that lived in close proximity. It is sometimes possible to establish biodiversities for metacommunities by taking many samples of similar, approximately contemporaneous fossil assemblages at different localities throughout a region, using the rule that species that occur together most often as fossils are most likely to have lived together. Because two fossil localities are likely to vary in composition even if drawn from the same parental biota, an increasingly full picture of regional biodiversity is gained

when numbers of localities are pooled. Such studies can lead to understanding the gross diversity patterns among biotic provinces. Finally, sampling of approximately contemporaneous localities can be made at the global level, in this way providing a sample of global biodiversity for the realm(s) involved.

Because most fossil localities are not directly dated by the more accurate radiometric techniques, correlation between them can only be approximate, particularly when the fossil assemblages represent different communities or provinces. This difficulty can be mitigated by using fairly coarse time units, within which most fossil localities can be placed with some confidence. Global biodiversity studies using families and genera usually employ temporal bins of from 5 to 8 million years in duration. Some studies have used geologically based time subdivisions, such as Stages; in these cases the diversity must be normalized to the differing Stage lengths.

The various approaches to help in overcoming the incompleteness of the fossil record yield binned samples of adequate size and quality to establish general levels and trends of biodiversity within well-skeletonized groups on global scales over long periods of time. The times of major diversifications or extinctions and periods of high and low biodiversities are captured by such data. Changes in the relative dominance of different major clades are also clearly demonstrated. However, the price for binning is that fine details of diversity changes are lost. Yet preliminary studies indicate that the generic record captures the same events as the family record and indeed is more sensitive to diversity changes. Furthermore, increasingly accurate dating techniques are permitting the use of ever-shorter time intervals. These ongoing refinements have confirmed the general trends described from the coarser data, while permitting the recovery of more detailed diversity information that reflects more closely the history of species richness.

Microbial Diversity of the Archaean and Proterozoic Eras

The earliest rocks known on earth date to about 4.6 billion years ago. The earliest eras in earth history, long thought to be barren of indications of life, were named the Archaean (about 4.6–2.5 billion years ago) and the Proterozoic (about 2.5 billion–543 million years ago). The most recent era was named the Phanerozoic (visible life) because it contained a fossil record. However, microbial fossils have now been found to date well back into the Archaean. The earliest are filamentous prokaryotic microbes resembling cyanobacteria (“blue-green algae”), recovered from chert beds that are over 3.5 billion years old. There are a few other Archaean records of microbial fossils that may be from other bacterial groups, representing both benthic and planktonic realms. However, bacteria and archaea have such extremely low preservation potentials that it is impossible to reconstruct their history of richness from fossil evidence. According to molecular phylogenies, a number of major bacterial branches, equivalent at least to phyla in a hierarchical system, arose before the cyanobacteria. It appears likely that there was a rapid early diversification within the bacterial and, probably, archaeal domains that produced a highly disparate microbiota well over 3 billion years ago. Considering the way in which lineages branch during evolutionary diversifications, it is likely that the microbiota was quite rich during much of Archaean and Proterozoic times. Recorded fossil diversity is much greater in the Proterozoic, but sampling is much better in rocks of that age and the recorded fossil richness simply tends to correlate with the number of formations and samples studied. One exception is in the latest Proterozoic, when sampling is highest but when the richness of fossil microbiotas declines, suggesting a real diversity drop at that time. The fossil record of richness of the Proterozoic microbiota is given in **Figure 2**.

Unicellular fossils belonging to the Eukarya are difficult to separate from those of other domains. Relatively large fossil algal ribbons (1 m long, 2 mm in diameter) about 2.1 billion years of age are likely to be the first known eukaryotes. It is possible that Eukarya evolved during the Archaean. The best Proterozoic records of unicellular Eukarya are of spheroidal plankters with organic walls, sufficiently tough to be preserved in some quantity. They increase in number, richness, and morphological complexity throughout the Proterozoic. As noted earlier, great bacterial disparity is likely to have been achieved in the Archaean, and the major richness trend may be attributed chiefly to an improving fossil record rather than to increased biodiversity.

Phanerozoic Marine Biodiversity

Neoproterozoic Animals

The record of metazoan fossils dates from at least 570 and possibly 600 million years ago, in the late Proterozoic (the “Neoproterozoic”), but it has proven to be very difficult to relate those early fossils to living groups, which first appear near the beginning of the Phanerozoic. There are two chief modes of fossilization: trace fossils, which are trails, burrows, and similar structures left in sediments by animal activities; and body fossils, which register the gross morphology of the animals. Most Proterozoic body fossils are restricted to body impressions, skeletons being absent until latest Neoproterozoic time. The early trace fossils are simple curved trails, and though traces become increasingly complex and diverse as time passes, they remain small throughout the Neoproterozoic, mostly 1 mm or less in width, though a few range to 5 mm. The early body fossil impressions (the “Vendian” or “Ediacaran” fauna) are chiefly of frondose and discoidal organisms that are large (some reach 1 m) and bear a general resemblance to cnidarians, although they differ from living forms in important structural details. Ediacaran faunas also include rare sponges and a few forms that resemble bilateral animals but that cannot be assigned to any living group. Estimating the living diversity of Neoproterozoic times from these fossils is difficult; body fossils of the trace-makers, for example, are not known, and many kinds of organisms can leave a given type of trace fossil. When body fossils appear during the Cambrian (see the following section), traces become larger (some are centimeters wide), much more varied, and more common. Therefore the best guess for

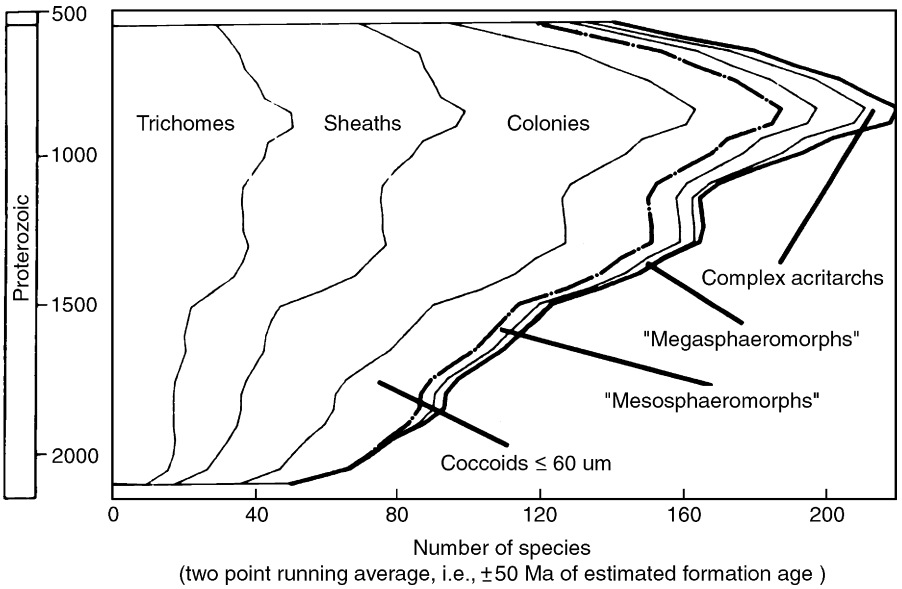


Figure 2 Species richness of several groups of unicellular microfossils as represented in Proterozoic sediments. The morphologic categories to the left of the heavy dashed line are presumptive Bacteria; those to the right are Eukarya. The record is unlikely to mirror the history of either richness or disparity, although the decline late in Proterozoic time may be real. Reprinted with permission from Schopf JW and Klein C (eds.) (1992) *The Proterozoic Biosphere.* Cambridge, UK: Cambridge University Press.

Neoproterozoic animal diversity is that it was lower than that of the Cambrian, beginning sometime before 570 million years ago (Ma), possibly tens to hundreds of millions of years earlier. This guess is likely to be correct as far as body plan disparity is concerned, but it remains possible that Neoproterozoic seas supported a rather high richness of species that had such low preservation potential that they have not yet come to light.

The Cambrian Explosion Establishes Metazoan Disparity

About 543 Ma there was a notable increase in the complexity, sizes, and abundances of trace fossils, one of which (*Triptichnus*) marks the basal Cambrian boundary by international agreement. Rare, minute, mineralized skeletons of uncertain affinities had appeared just prior to the first appearance of *Triptichnus*, and continued to appear in increasing kinds and numbers, along with an increasing trace diversity, in younger rocks. At about 530 Ma the earliest undoubted skeletons of living animal groups appeared and, during the next 10 million years or so, 11 animal phyla made their first appearances, a period known as the "Cambrian explosion" because of the geologically sudden appearance of diverse, abundant skeletal remains. Invertebrate skeletons are then continuously present for the remainder of geologic time. Among the body plans to appear during the explosion are complex representatives of all major animal clades, including annelids, arthropods, and chordates (**Figure 3**). Judging from the phylogenetic tree of animals, all living phyla had originated by the close of this interval, although some living phyla with low preservation potentials are still completely unknown as fossils.

Many of the explosion animals include features that are not present in their living allies, and thus extend the known morphological ranges of their clades. Some of these forms have such distinctive body plans that they are considered to be phyla of their own, whereas others can be nested within living phyla but not within living classes. These forms indicate that the disparity among the body plans of Early Cambrian animals was at least as great as it is today. For example, the relatively few Cambrian arthropod taxa are just as morphologically disparate as are living arthropods, which have had the benefit of over half a billion years of subsequent evolution and consist of million of species today. Thus disparity and richness are evolutionarily decoupled. The Early Cambrian burst of disparity in the fossil record suggests that there was relatively rapid evolution of the developmental regulatory gene systems at that time.

Phanerozoic Diversity Patterns in the Marine Environment

Standing Diversity of Marine Invertebrates

The combination of completeness, disparity, richness, and duration of the marine fossil record is superior to that of other realms, especially for marine invertebrates, and thus is best suited as an introduction to major diversity patterns. **Figures 4** and **5** depict the overall pattern of family richness for marine animals, almost entirely shallow-water (continental shelf depth) forms. It is likely that the range of body plan disparity declined somewhat after the Cambrian, and probably showed a significant drop and then a recovery following the Permian–Triassic extinction, but there are no data that incorporate the entire marine fauna. However, Jack Sepkoski has complied an excellent data set on family richness, which has certainly varied significantly.

During the Cambrian, family richness rose abruptly and seemed to be reaching a plateau as the period ended. However, the Ordovician witnessed a dramatic increase in family richness, which reached a plateau near 400 families by the end of

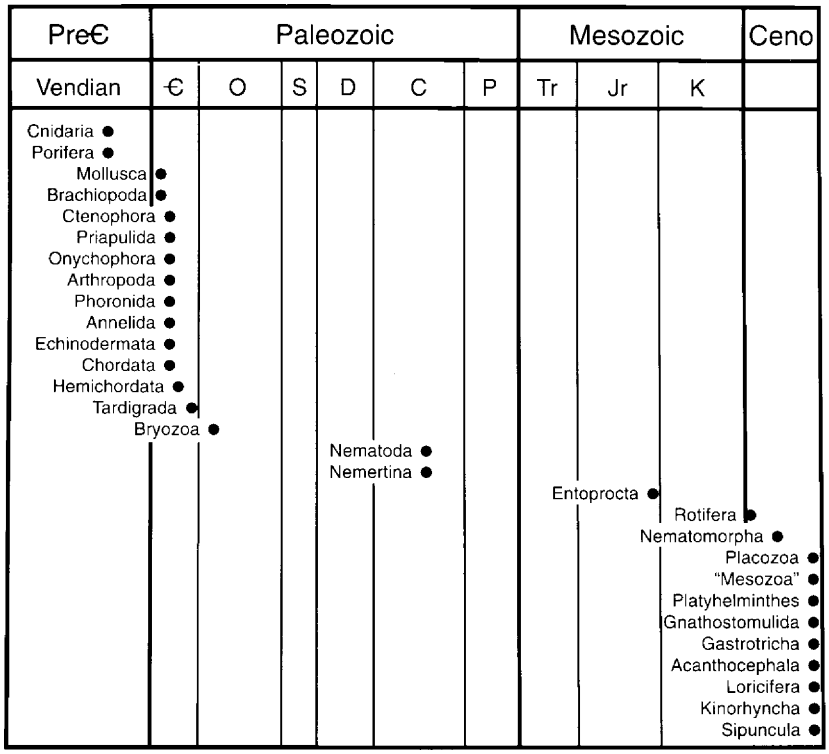


Figure 3 The earliest appearance of body fossils of living phyla. The earliest records of small, soft-bodied forms with low fossilization potentials are scattered through time, being found at localities with unusual preservation, such as in amber, or having no fossil records at all. The only phylum with well-mineralized skeletons that does not appear in the Cambrian is the Bryozoa, which, however, has numbers of nonmineralized forms today and may have had no mineralized representatives during the Cambrian. Є, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; Jr, Jurassic, and K, Cretaceous. Modified from Valentine JW, Jablonski D, and Erwin DH (1999) Fossils, molecules and embryos: New perspectives on the Cambrian explosion. *Development* 126: 851.

the period (see **Figure 4**). This richness plateau lasted for about 300 million years, during which there were two major extinctions that lowered family richness for a relatively brief time, followed by recoveries to the plateau level until the close of the Permian period. At that time, the plateau was abruptly terminated at about 250 Ma by the greatest extinction on record, near the Permo–Triassic (P/T) boundary. About half of the marine families were lost. From that low point during the depths of the extinction, family richness has climbed at a rate quite similar to the Cambro–Ordovician rate, and for a far longer time, peaking at the present. This last richness rise is interrupted twice by major extinctions, the last of which, near 65 Ma, is the famous Cretaceous–Tertiary (K/T) extinction that eliminated the dinosaurs. Despite these interruptions, family diversity today far exceeds that of the long Paleozoic plateau.

Perhaps as few as one-third of living marine families have durable skeletons; the rest have very poor preservabilities. Nevertheless, there is evidence to suggest that the major trends in the fossil record are real, and that the ratio of easily to poorly preservable forms has been similar enough, at least following the Cambrian explosion, that the general magnitudes of family diversities can be reconstructed from the preserved fossils. For example, the diversity patterns of easily and poorly preservable taxa are similar throughout the shallow ocean today, so far as we can tell. Further, the ratio of poorly preserved to easily preserved families that are found (see **Figure 5**) holds reasonably steady through the Phanerozoic, increasing slightly toward the present, suggesting that the record is sampling a fairly consistent fraction of the poorly preserved clades, and the ratio improves slightly in younger sediments as would be expected. Also, the presence of a long richness plateau during the Paleozoic is unlikely to be a matter of chance, and suggests some stability in sampling success.

Standing Diversity of Marine Chordates

The phylum Chordata includes an invertebrate class, the Cephalochordata, that is known from the Early Cambrian and is represented today by amphioxus, (*Branchiostoma* spp.), but this class is not known to ever have been diverse. There are primitive groups within the class Vertebrata, such as lampreys and hagfish, that in fact lack vertebral columns. A group of marine vertebrates at about this level of organization, known as the Conodonta from toothlike feeding structures, appeared during the Middle Cambrian and left a significant Paleozoic fossil record. Conodonta became extinct during the Triassic. Nonconodont jawless fishes (Agnatha) appeared in the late Upper Cambrian and radiated into four distinctive, disparate clades during the Silurian; many agnathans had bony armor that has provided most of their fossil record, but the agnathans seem not to have been rich in species. They were joined by jawed fishes by Devonian time. The jawed forms radiated to produce several disparate types—arthrodires and placoderms, with bony head shields and plated armor, and several chondrichthian groups (sharks and rays), which were the richest of

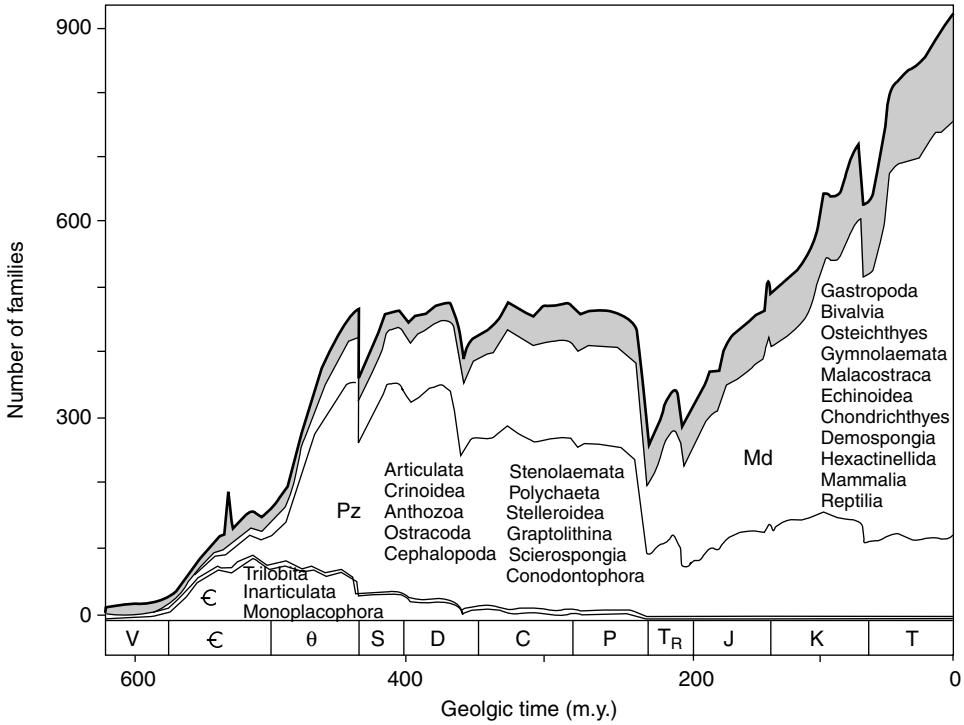


Figure 4 Marine animal family richness known from the fossil record. Geologic period abbreviations are as in **Figure 3**. Major faunas are labeled €, (Cambrian Fauna), Pz (Paleozoic Fauna), and Md (Modern Fauna). Shaded area at top of curve indicates animals that lack durable skeletons. Reprinted with permission from Sepkoski Jr. JJ and Miller AI (1985) Evolutionary faunas and the distribution of Paleozoic marine communities in space and time. In: Valentine (ed.) *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, pp. 153–190. Princeton, NJ: Princeton University Press.

fish clades in Late Paleozoic seas. Also arising at least by Early Devonian time was a clade, the osteichthyes or bony fishes, which was of secondary importance during the Paleozoic but came to be the most richly diverse vertebrates in the sea during the Mesozoic and quite dominate marine vertebrate faunas today. Thus the history of marine fishes can be summed up as beginning modestly in the Cambrian and finally achieving significant disparity during Silurian and perhaps Early Devonian radiations that established most major groups, which then waxed and waned in richness, with bony fishes finally coming to dominance. Reptiles entered the sea during the earliest Mesozoic and mammals early in the Cenozoic, but neither group became very rich in marine species.

Taxonomic Turnover

Relative clade richnesses vary considerably across Phanerozoic time. In **Figure 4**, the changing assemblages of dominant clades are classed into three major faunas. The Cambrian Fauna (labeled €), composed dominantly of trilobites (Arthropoda), does not diversify during the Ordovician and is overtaken by the expanding clades of the Paleozoic Fauna (labeled Pz), dominated by articulates (Brachiopoda) and crinoids (Echinodermata). The Paleozoic Fauna is decimated by the P/T extinction and does not increase significantly thereafter, and is overtaken by the diversifying Modern Fauna (labeled Md), which is dominated by mollusks and fishes. Elements of all of the faunas are present throughout the Phanerozoic.

Figure 5 depicts the great variety of family richness histories recorded for marine classes and phyla; the dominance of the characterizing clades during each of the three faunas is clearly evident. There are some general features that link the diversity patterns to evolutionary rates that seem to be inherent in the clades. The dominant clades of the Cambrian Fauna have very high evolutionary turnover rates—high rates of both diversification and extinction. The dominant clades of the Paleozoic Fauna have moderately high turnover rates, and those of the Modern Fauna have the lowest turnover rates of the dominant clades. As a result, the extinction rate for all families combined, excluding the major extinction events, declines steadily throughout the Phanerozoic. Evidently, a strategy of being resistant to extinction is superior to having a high speciation rate, perhaps because slight inclement environmental fluctuations can force a high-turnover lineage too close to extinction to recover. Note that the presently dominant bivalve and gastropod mollusks, slow-turnover clades, increase their relative representations inexorably throughout the Phanerozoic, and are not greatly affected even by the P/T extinction. A few clades that have fair preservation potentials can be shown to have survived for hundreds of million of years despite having very low diversities; the inarticulate brachipods and the scaphopod mollusks are classic examples (see **Figure 5**). These are low-turnover clades; if they had high turnovers they would presumably have been swept away by even a small extinction event. Both speciation and extinction rates are probably tied to population parameters that of course evolved without respect to clade longevity or dominance,

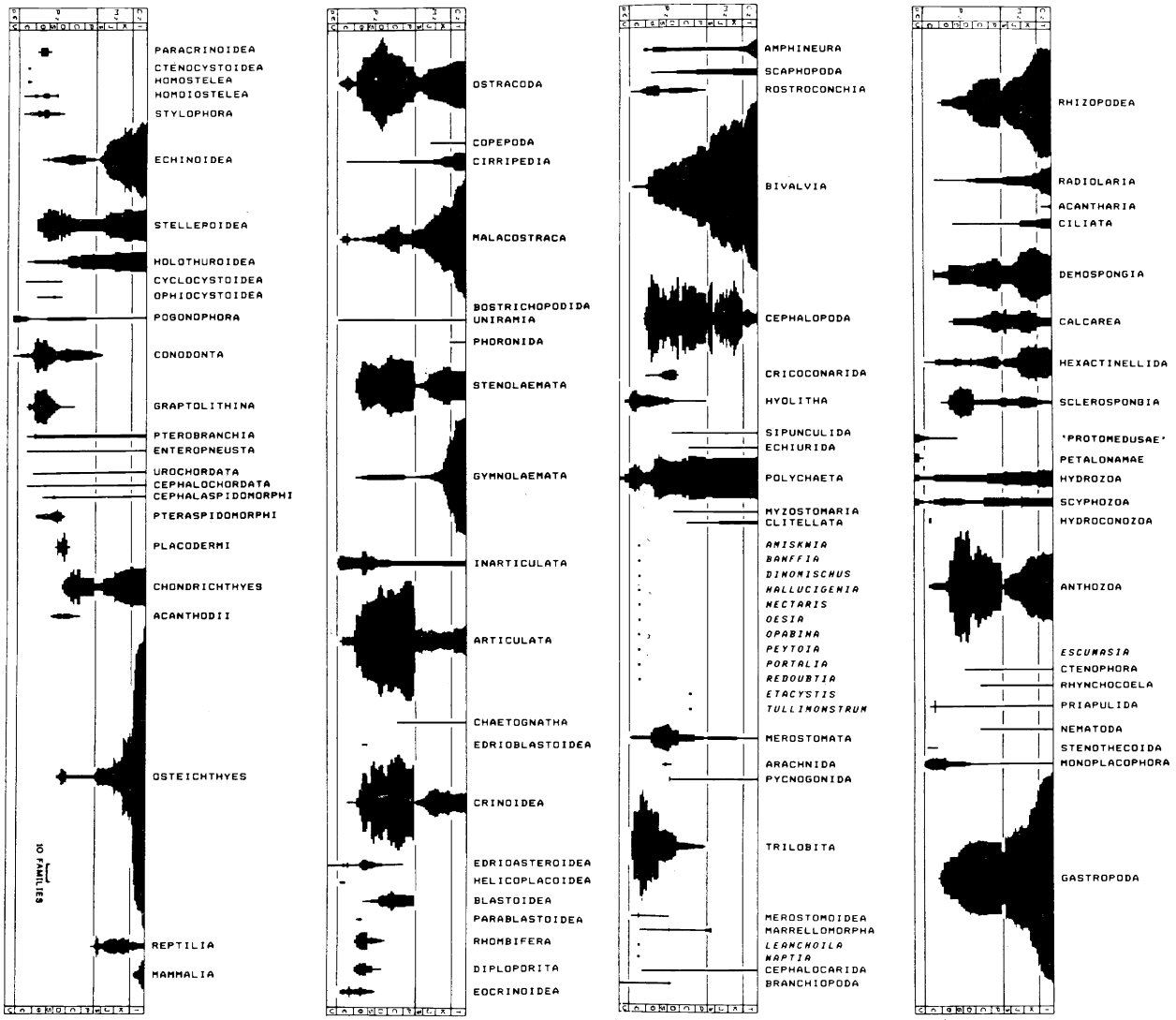


Figure 5 Phanerozoic diversity in the marine fossil record, showing animal family richness by phylum or class. Geologic periods are as in **Figure 3**. Reprinted with permission from Sepkoski Jr. JJ and Miller AI (1985) Evolutionary faunas and the distribution of Paleozoic marine communities in space and time. In: Valentine (ed.) *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, pp. 153–190. Princeton, NJ: Princeton University Press.

which in a sense are merely side effects, though they profoundly influence the composition of the world's biotas.

Phanerozoic Terrestrial Biodiversity

Plant Diversity

"Land plants" include two major groupings: forms such as mosses and liverworts, collectively termed bryophytes, which lack a water-transporting system; and the tracheophytes, which possess vascular tissues that conduct water vertically and permit significant upward growth. It seems likely that tracheophytes arose from among the mosses, perhaps more than once. The earliest fossil record of plants on land may be furnished by spores of Middle Ordovician age that are possible bryophytes; tracheophytes probably date from Early Silurian, but their vegetative elements first appear in the Middle Silurian. From then through the Devonian, a plexus of early plant types diversified, and from this plexus the modern groups of lycopods, horsetails, ferns, and gymnosperms evolved by latest Devonian time; there were forests before the close of the Devonian. This early burst of disparity is similar to that found among animals, but the richest of living plant groups, the angiosperms or flowering plants, appear much later, arising perhaps in the Late Jurassic but spreading during the Late Cretaceous and dominating Cenozoic plant associations. There is no animal analog for the late origination of a dominant major taxon, except for the class Mammalia, also terrestrial of course. However, the disparity indicated in the pollen record of angiosperms has been investigated, and it shows an early burst of disparity increase in the Early Cretaceous that gradually tapers off to a plateau by late in the Cretaceous. Evidently angiosperm "body" plans were evolved early in the history of the group and the rise of species richness followed later, the same pattern that is found within many

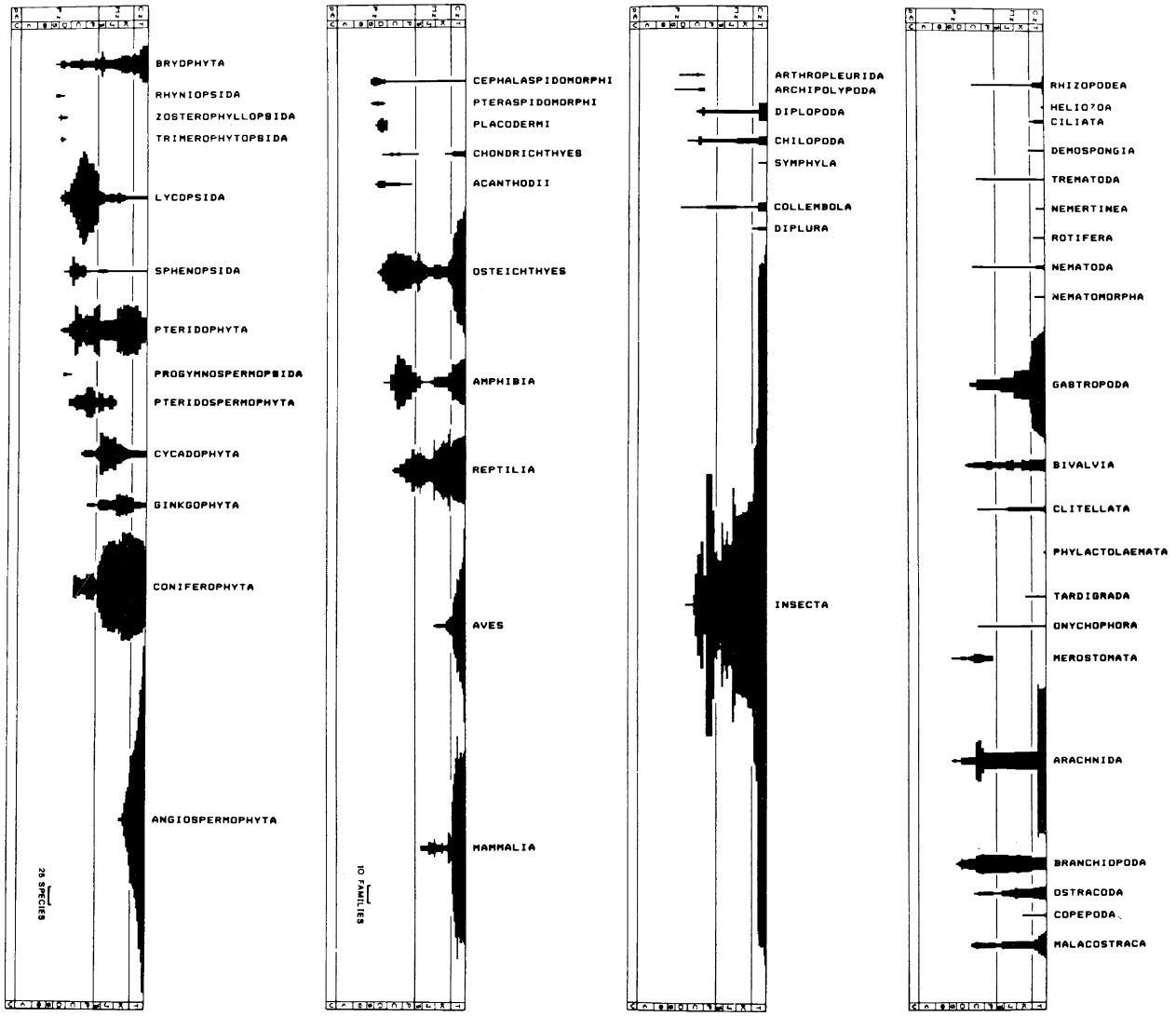


Figure 6 Phanerozoic diversity in the terrestrial fossil record, showing plant and animal family richness by phylum, class, or other higher taxon. Geologic periods are as in **Figure 3**. Reprinted with permission from Sepkoski Jr. JJ and Miller AI (1985) Evolutionary faunas and the distribution of Paleozoic marine communities in space and time. In: Valentine (ed.) *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, pp. 153–190. Princeton, NJ: Princeton University Press.

marine invertebrate groups and in numbers of vertebrate groups. **Figure 6** indicates the relative richness of major land plant groups through time as judged from their fossil records.

Animal Diversity

Invertebrates

Animals came ashore with the earliest plants, so far as can be told. It is quite possible that early terrestrial animals fed on bacteria, fungi, or protistans and were ashore before tracheophytes; early arthropod burrows are reported in Ordovician soils. However, the earliest undisputed terrestrial animal remains, of arthropods, are from the Late Silurian. Because these fossils are predators (primitive centipedes and an arachnid), there must have been by that time an extensive fauna of primary consumers and a fairly complex ecosystem, evolved in previous epochs. There is little evidence to judge how quickly and to what levels terrestrial invertebrate diversity rose during those times.

Insects evidently arose in the Early Devonian, and they have a relatively rich fossil record for a terrestrial group, being richer than the vertebrate record. The same species proportions found among living orders are also found among Tertiary fossils, suggesting that the insect record, while certainly incomplete, is not heavily biased taxonomically, and that at least their major fossil trends probably reflect actual paleobiological trends. During their earliest, Devonian history, insects are spottily represented; most of them are predators, whereas the primary consumers seem to have been chiefly detritus feeders; little herbivory is found. Insects underwent a significant expansion in the Carboniferous, and species richness rose to a high just before the major Permo–Triassic extinction, which affected insects significantly (see **Figure 6**).

However, the insects rebounded, rising steadily to their present overwhelming richness. Probably the rise was more evenly distributed in time than indicated in **Figure 6**, for there are some spectacularly rich insect faunal localities in the Late Tertiary, while equivalent faunas are not yet known from earlier deposits. As with marine invertebrates, significant morphological disparity was reached very early in insect history, preceding the greatest rises in richness. Though insect richness seems closely tied with the richness of the angiosperm flora, insect disparity, even including the evolution of the array of mouthparts and feeding structures present today, was largely accomplished well before the origin of angiosperms.

Chordates

All the primitive groups of jawed fishes are found in terrestrial aquatic deposits during the Middle Paleozoic, but on land as in the sea it was the bony fishes (Osteichthyes) that radiated to produce rich faunas that dominated in fresh waters, from the Devonian onward. The richness of terrestrial fossil fish faunas declined significantly in the Mesozoic, perhaps partly owing to a poor aquatic fossil record from that Era.

Tetrapods first appear in the Devonian, possibly Early Devonian, in effect as part of the general radiation of fishes, but with limbs that eventually permitted locomotion on land and led to the evolution of the Amphibia. The fossil record of this group is very sparse, but it is clear that the amphibians diversified extensively to reach a level of family richness comparable to today's by the close of Carboniferous time. The early amphibians were rather different from living forms, however, and included large animals, which were eventually replaced ecologically by reptiles. Three important groups arose from amphibian stocks, one leading to reptiles, one to mammals, and one to dinosaurs and birds. The early histories of these groups are so poorly known that their patterns of diversification cannot be accurately reconstructed. It is possible that both reptiles and dinosaurs gradually increased in richness and diversity through most of the Mesozoic era. The richness of these clades was reduced drastically during the extinction at the close of the Cretaceous.

While mammals must have diversified as well during the Mesozoic, it is in the early Cenozoic that their fossil record is marked by a burst of new appearances and, by about 10 million years after the start of the Cenozoic, all of the 17 modern orders of placental mammals had appeared. Only two of these orders are known from the Cretaceous, and while it is likely that most trace their roots to a few tens of millions of years earlier, mammals were clearly a minor though probably not unimportant part of the tetrapod fauna before the Cretaceous–Tertiary (K/T) extinction. It is usually assumed that mammals replaced dinosaurs, which became extinct at the K/T boundary, in ecological roles that permitted large body sizes. The only dinosaur descendants today are birds, which have a poor fossil record. Modern bird orders are unknown before the Cenozoic, suggesting that there was a significant K/T extinction of birds, leading to the radiation that produced modern types. There must indeed have been a significant diversification of birds during the Cenozoic, but as the roots of the modern lineages are uncertain and the bird record is poor, the pattern of bird diversity increase has not been established.

Principal Biodiversity Factors

Plate Tectonics and Global Heterogeneity

The geographic pattern of land and sea is constantly changing due to the processes of plate tectonics, which create and destroy the crust of the earth. At present the crust is divided into about 6 major and 14 minor crustal plates. New crust is added at one plate margin, a rift marked by deep-sea ridges, from sources of molten rock in the earth's interior. At the opposite plate margin, the crust plunges back into the interior, its descent marked by deep-sea trenches or "subduction zones"; this crust then remelts. Continents or islands on a crustal plate therefore move across the earth's surface, from the constructional margin toward the destructional margin, as if on a conveyor belt. Continents separated by a rift will diverge as they are carried apart on their respective plates, and continents will fragment if a rift cuts across them. On the other hand, continents separated by a subduction zone will converge, and if they meet they can weld together. The rates of movement of the plates are only several centimeters per year. Nonetheless, because plate tectonic processes have been incessant since well before the origins of animals and plants, the continents and oceans have changed their numbers, sizes, and geographic configurations over the many millions of years since those organisms first evolved.

As continents drift across the earth's surface they enter new climatic zones, and as oceans enlarge or shrink or change their shapes, the oceanic circulation changes, with warm or cool currents directed to different regions. The biota of any given area will evolve to adjust to the changing conditions, and may become enriched, perhaps if moving into the tropics, or depauperate, perhaps if moving into high latitudes. At the same time, migration routes are opened or closed by the shifting geography, permitting invasions of biotas into some regions while isolating the biotas of other regions. The biodiversity patterns both on land and in the sea reflect the environmental patterns created by plate tectonic processes.

In addition to its effects on rearranging regional biodiversity patterns, changes in global geography can significantly affect global biodiversity levels. The most impressive example is in the rise of biodiversity during the last 200 million years or so, from the Middle Mesozoic era through the Cenozic era. At the beginning of the Mesozoic, the continents were welded together into a supercontinent, called Pangaea. Although Pangaea was certainly environmentally heterogeneous, with large climatic and topographic variations, the climatic zones tended to be occupied by widespread biotas that could spread across what are now separate continents. In the seas, the shelf encircled the single supercontinent in an essentially uninterrupted band. While the distribution of species was limited by climatic zonation and perhaps regional habitat peculiarities, even shallow-water species tended to be relatively widespread and of course were largely confined to the one shelf, though perhaps some were present on oceanic islands, of which little is known. As the Mesozoic passed, however, the continents were traversed by rifts and began to break up into smaller landmasses, between which were arms of the sea that finally widened into oceanic expanses. By the

close of Early Cenozoic time, all of the present continents that are now separated had broken from each other.

Species that could once range widely across Pangaea became isolated on separate continents. As the continents dispersed into different climatic zones and developed different topographies, their terrestrial biotas each evolved distinctively in response to local conditions and events. For example, North America, Eurasia, and Australia each developed distinctive temperate biotas, and the tropics of the Americas became distinct from the Old World tropics and those of the western Pacific. In the sea, the growing geographic differentiation was even more profound, as new shelves appeared on each side of the new oceans, significantly raising the number of isolated shelf segments. As on land, each continent evolved a distinctive biota in response to the unique set of environmental conditions and events that developed in each region. The heterogeneity of world environments thus increased greatly during this period, on a global scale. These plate tectonic effects are responsible for a significant part of the important rise in standing global marine diversities during the Mesozoic and Cenozoic as represented at the family level in **Figure 4**; an accompanying rise occurred in the continental biota.

Climate Change

Climate is another major factor governing global biodiversity. For example, earth underwent a general cooling trend during the Cenozoic, although there were shorter warming and cooling events superimposed on the generally falling temperature curve. The temperature change was greatest in high latitudes, thus increasing the temperature differential between the poles and the equator. The margins of the tropics were shifted toward the equator, the tropics became compressed, and cooler climates came into existence in high latitudes, shifting equatorward themselves as cooling continued, to be replaced poleward by still cooler climatic regimes.

In the sea, the increasing temperature gradient made for a markedly increased provincialism. Tropical species whose biology was related to climate became progressively more restricted to lower latitudes, with some occasional, relatively short-lived reversals in which species tolerant of lower temperatures arose and diversified in the more poleward regions. Because the major rifting that produced the Cenozoic continental pattern was primarily in a north–south direction, most major continental shelves trend north–south in the Cenozoic, as they do today, and so there was a rise in species richness on each shelf as the latitudinal thermal gradient increased, multiplying the effect on global marine biodiversity by the number of isolated shelves. The Cenozoic global marine rise in family richness shown in **Figure 4** is compounded partly of the increase in the number of isolated shelves following continental rifting and partly of the increased provincialism on each shelf created by an increasing thermal gradient; together they have formed a powerful engine of global species enrichment.

In terrestrial environments, family richness also increased during those Cenozoic events, and for the same reasons. The continental expanses provided for a three-dimensional array of habitats (including high mountains and basins) unlike the rather two-dimensional array along the narrow, shallow continental shelves, and therefore the biotic response has been correspondingly more complicated in terrestrial than in marine environments. Not only did high-latitude climate zones open up, but high-altitude regions became cooler, increasing the contrasts between mountain and lowland environments. A lowering of rainfall in some continental interiors produced semiarid plains and steppes. The biota responded to these increases in environmental heterogeneity by producing, for example, alpine and grassland plants and animals, thus enriching both the flora and fauna. The species richness of angiosperms among plants and of insects among animals profited the most from these events.

During the Late Cenozoic, as cooling has continued in high latitudes, massive ice accumulations on a subcontinental scale have produced a series of glacial ages interrupted by warmer interglacial periods, most pronounced during the Pleistocene, which began about 1.6 million years ago. Some regional extinctions occurred during the early onset of cold periods, but in general there has been little extinction, and thus no significant lowering of global biodiversity, associated with these climatic swings. Instead, the biotas of both land and sea have tended to migrate with their climatic zones, southward during glaciations and northward during interglacial times. Thus regional diversities have changed as species have migrated, but overall diversity has not been significantly affected. Indeed, the Recent biosphere may be the richest in species during the entire history of animals and plants, at least prior to the deleterious influences of human activities on biodiversity.

Extinction Events

Changing environments have often provided opportunities for the origin of species, some of which have clearly led to major morphological novelties and enhanced morphological disparity. At the same time, changing conditions have led to the extinction of species, which has occurred more or less continually over Phanerozoic time. As noted earlier, extinction rates vary among taxa, whose histories tend to reflect their rates of turnover. The species richness of a taxon results from an interaction between its rates of speciation and extinction. Thus when there has been some unusually profound disturbance to the global environment, and extinction rates have become unusually high, species richness has fallen dramatically in events called mass extinctions. The most drastic such event, recorded in marine fossils, was the Permian–Triassic (P/T) extinction about 250 million years ago, at the close of the Paleozoic era. As many as 90% of marine species may have become extinct across the P/T boundary (see **Figures 5** and **6**). This extinction was highly differential; the groups that did best in surviving the extinction and in rediversifying were those that dominate the marine fauna today. Disparity as well as richness was affected; some entire classes and orders of marine invertebrates disappeared. Brachiopods, crinoids, and other groups that were dominant before the extinction were reduced to minor roles. The causes of this extinction, and the extent of its effects on the terrestrial biota, are not yet clear.

There have been a number of other mass extinctions (five major events in all) and numerous regional extinctions that were severe locally. The most famous of the mass extinctions, the K/T, occurred about 65 million years ago at the close of the Mesozoic era, although it did not have as profound an effect on the marine environment as the P/T event. The K/T extinction strongly affected terrestrial as well as marine life. It probably resulted from the impact of an extraterrestrial object, such as a meteor, with the earth.

The "background" extinctions that account for the general turnover of taxa are accompanied by "background" speciations, so that species richness levels do not change much during normal turnover and tend to be significantly affected by background effects only during trends that last tens of millions of years or more. However, the more abrupt, severe regional and mass extinctions are not immediately compensated by correspondingly massive speciations, and indeed the recovery of species richness usually requires millions of years following a mass extinction. For phyla, classes, or orders that are lost to extinction, there is no recovery at all, and the unique gene regulatory systems involved in their architectures are gone forever.

Clearly, biodiversity data from the fossil record support the notion that species richness is largely associated with environmental heterogeneity, and thus the preservation of species diversity can be promoted by preserving habitat diversity. There are entire animal phyla that are represented today by very few species: the phylum Phoronida by only 12 or so, and the phylum Priapulida by perhaps 17, for example. These phyla harbor uniquely organized genomes that have been present for over half a billion years of earth history, and that are largely unstudied. Their preservation should be a priority, for they represent major elements in the disparity of the present biosphere.

See also: Archaea, Origin of. Biodiversity, Origin of. Biogeography, Overview. Cladogenesis. Eukaryotes, Origin of. Genes, Description of. Microbial Biodiversity. Taxonomy, Methods of

References

Bengtson S (ed.) (1994) *Early Life on Earth*. New York: Columbia University Press.

Foote M and Sepkoski Jr. JJ (1999) Absolute measures of the completeness of the fossil record. *Nature* 398: 415.

Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* 276: 734.

Schopf JW and Klein C (eds.) (1992) *The Proterozoic Biosphere*. Cambridge, United Kingdom: Cambridge University Press.

Valentine JW (ed.) (1985) *Phanerozoic Diversity Patterns*. Princeton, New Jersey: Princeton University Press.

Valentine JW, Jablonski D, and Erwin DH (1999) Fossils, molecules and embryos: New perspectives on the Cambrian explosion. *Development* 126: 851.

Global Declines of Amphibians

Kellie Whittaker, Michelle S Koo, and David B Wake, University of California, Berkeley, CA, USA
Vance T Vredenburg, San Francisco State University, San Francisco, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Vance T. Vredenburg and David B. Wake, pp 1–9, © 2001, Elsevier Inc.

Glossary

Cryptic diversity Populations that are nearly identical to others in morphology but have significant differences in their DNA sequences, indicating isolation for an extended period of time; these represent unique evolutionary lineages that should be preserved.

Emerging disease An infectious disease that has newly appeared in a population or that has been known for some time but is rapidly increasing in incidence or geographic range.

Exotic species Organisms living in habitats where they do not occur naturally.

Metapopulation Collection of local populations linked through emigration and dispersal whose long term survival depends on the shifting balance between local extinction and recolonization.

Taxonomist Scientist whose work involves the discovery, description, and classification of life forms, based on the Linnean hierarchical system.

Introduction

Global declines of amphibians refer to the phenomenon of the population declines and even extinctions of amphibian species around the world. Assessments of the world's amphibians (Stuart *et al.*, 2004, 2008) found that nearly a third of the known species of amphibians are globally threatened with extinction and that at least 42% of known amphibian species are experiencing population losses. The rapidity and extent of these declines, far more dramatic than those described for birds, mammals, or reptiles, forecast impending extinction of numerous amphibian species during the coming decades. Many of these declines have occurred in protected areas where the causes have remained enigmatic. Other declines are due to obvious reasons, particularly habitat destruction, but the number of species experiencing enigmatic declines is increasing and these have caused the greatest alarm. Although climate change has been invoked as a potential contributing factor (e.g., Pounds *et al.*, 2006), evidence is mounting that epidemics of infectious disease may be primarily responsible for many of these catastrophic declines (Collins and Crump, 2009). The greatest threat is presented by a virulent fungal pathogen that can cause the infectious disease chytridiomycosis, which has decimated entire assemblages of amphibians worldwide. This pathogen, *Batrachochytrium dendrobatidis*, is responsible for what may well be the greatest disease-caused loss of biodiversity in recorded history, having caused population crashes or extinctions (often within a single year) of at least 200 species of frogs, even in relatively undisturbed, remote habitats. Compounding the crisis, amphibians suffer from relatively low levels of conservation effort compared to other vertebrate taxa. Although amphibians have survived a number of mass global extinctions throughout their 200 million year history, the speed and magnitude of this loss in amphibian biodiversity supports the idea that the earth may be now in the midst of a sixth mass extinction, similar to that which wiped out the nonavian dinosaurs.

Amphibian Biodiversity

The world's living amphibians include over 6800 species placed in three distinct clades, the frogs and toads (Anura), salamanders (Caudata), and caecilians (Gymnophiona). Of the three groups, frogs and toads exhibit the most varied reproductive modes and habitat associations and comprise the majority of amphibians (over 6000 species). Salamanders and caecilians, also diverse, have fewer species and are more restricted in distribution, but are still widespread (more than 600 species and nearly 200 species, respectively, based on the online resource AmphibiaWeb 2010). Most of the world's amphibian diversity occurs in the tropics, especially in Central and South America; other amphibian biodiversity hotspots include sub-Saharan Africa, Madagascar, Sri Lanka, SE Asia, New Guinea, and Australia (**Figure 1**). Salamanders are generally found in North Temperate regions, where all 10 families occur, but the largest family (Plethodontidae) is well represented in tropical America, where more than 40% of all salamander species occur. Salamanders are especially abundant in North America, whereas caecilians are restricted to tropical regions. In contrast, frogs range in distribution from the Arctic zone to the southern tips of South America and Africa.

Amphibians are often characterized as tetrapods with aquatic larvae and terrestrial adults, but alternative life histories are common. Some species of the three main amphibian clades are permanently aquatic. In contrast, other members of all three clades (including a majority of the species of salamanders and caecilians and more than a thousand frogs) are strictly terrestrial and lack aquatic larvae; eggs well provisioned with yolk are laid on land and develop directly into miniatures of adults. Some members of all three amphibian clades are live-bearers, giving birth to metamorphosed young that have been nourished during development in the reproductive tract of the female. Some salamander species never metamorphose, but remain aquatic in gilled or semi-gilled states throughout

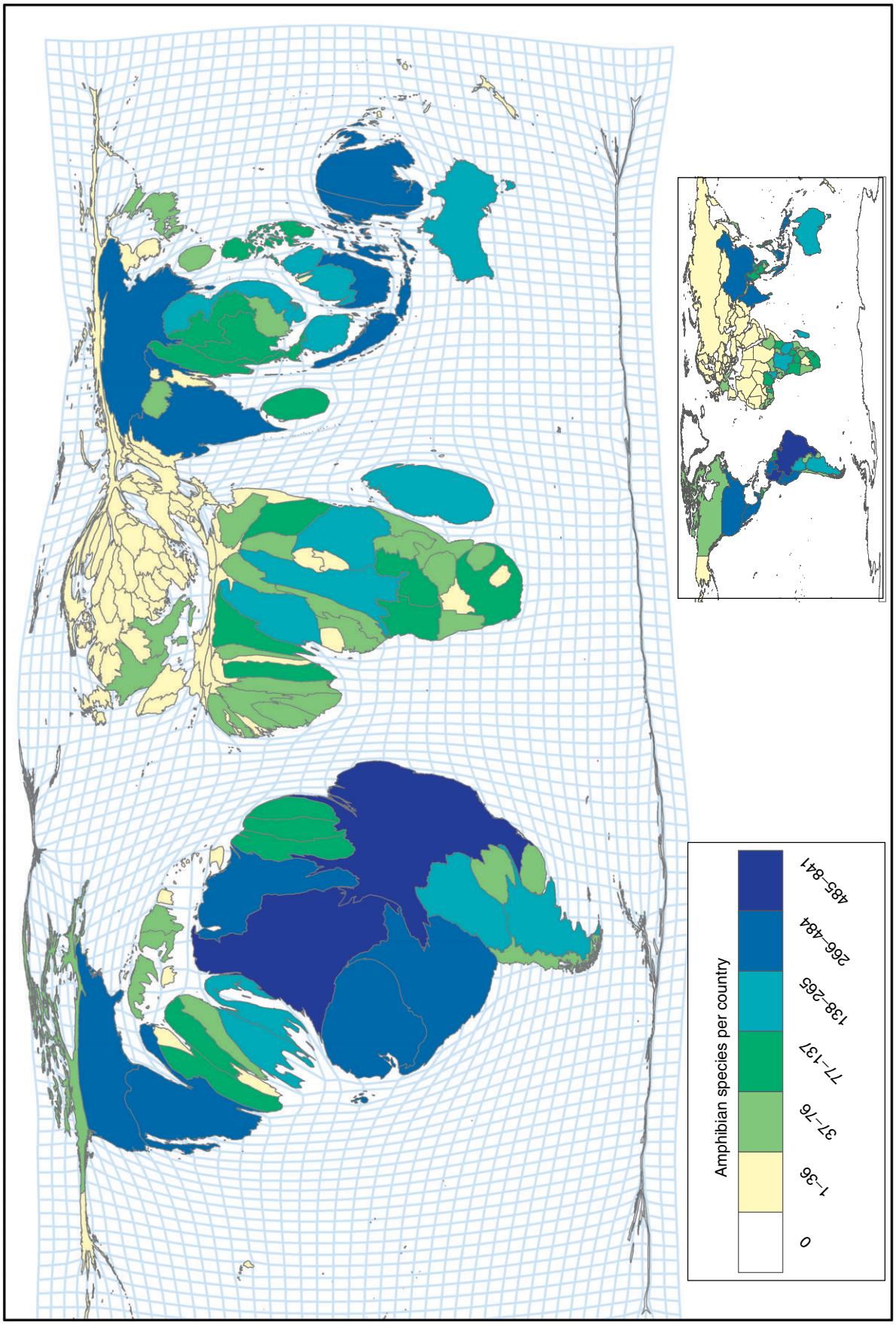


Figure 1 Global amphibian species diversity by country, visualized using density equalizing cartograms. Each country's shape is purposely distorted in proportion to the number of amphibian species per country area (square kilometer) using the cartogram technique after Gastner and Newman (2004). (Data source: AmphibiaWeb (website accessed 17 November 2010).)

life. Frogs have evolved many unusual life histories that avoid aquatic habitats for depositing eggs or larvae, such as transferring eggs to male vocal sacs, to pouches in the skin of the back of females, or even to the stomach.

Most salamanders have the structure of a generalized tetrapod with four legs, a relatively short but flexible trunk, and a tail, but some are extremely elongated with very small limbs or only forelimbs, and some reach very large size – in excess of 1.5 m. Caecilians are limbless and their eyes are covered by skin or bone. Caecilians have high numbers of trunk vertebrae and are very elongated, but they either lack entirely or have an exceedingly short tail. Frogs are also tetrapods but have a more specialized form that generally enables jumping (although some species walk, hop, or burrow), because they lack a tail, have a very short and stiffened trunk, and have relatively elongated legs and feet.

Adult amphibians are effective predators and both salamanders and frogs have tongues specialized for rapid, long-distance prey capture. Caecilians generally feed on subterranean prey such as earthworms. Amphibians are represented in diverse aquatic and terrestrial ecosystems and frequently are important components of communities and food webs. In some ecosystems, they are the dominant predators, both in terms of numbers and total biomass.

Dimensions of the Problem

Geography

The geographic extent of amphibian declines is worldwide (see Fisher *et al.*, 2009 for a longer review). Areas most strongly affected are located in Central and South America, the Caribbean, the wet tropics of eastern Australia (**Figure 2**), and western North America (Stuart *et al.*, 2008). The islands of the Caribbean and Pacific have some of the highest percentages (up to 100%) of endemic species that are threatened or endangered, but have relatively few species. Less is known about the status of species in Africa and Asia. Nonetheless, some species-rich countries have alarmingly high percentages of threatened and endangered species (e.g., 70% for Sri Lanka, based on the online resource IUCN 2010). The first reports of massive collapse of amphibian faunas came from montane areas in Central America and Australia. The loss from the protected Monteverde Cloud Forest Reserve in Costa Rica of more than 50% of the amphibian species within the course of a single year (1987) was a profound shock, and included the first prominent extinction (the golden toad, *Bufo periglenes*). Collapse of amphibian fauna in montane and lower montane Central America and South America is ongoing. In Australia, declines began in the 1970s in southern Queensland and Tasmania and accelerated in the 1980s. In Europe, die-offs were observed beginning with Spain in the mid-1990s. Concern has been expressed over declines of frogs in the western US for many years, and in the 1980s and early 1990s population collapses were observed in relatively undisturbed areas such as Yosemite National Park. Amphibian declines have since been reported from many parts of the world.

Ecology

In characterizing the ecology of species that have declined, most attention has focused on high- or mid-altitude species that are associated with streams, have small ranges (implying higher habitat specificity), and low reproductive rates, but there are many exceptions. Species that have aquatic breeding habits and stream-dwelling tadpoles have generally experienced sharper declines than species that lay eggs on land and that develop without a larval stage (direct developers). However, direct-developing species are not exempt; in Puerto Rico, for instance, many species of direct-developing frogs have declined and some may be extinct.

Amphibians are important members of ecosystems, and their declines are likely to have diverse and as yet not fully understood impacts on communities and ecosystems. Amphibians are unique among vertebrates in that many have biphasic lifecycles that include both fully aquatic (larvae) and terrestrial or semi-terrestrial (adult) forms. This makes amphibians key components in both terrestrial and aquatic food webs. Aquatic systems are more productive than the surrounding terrestrial systems in some areas, and amphibians help to link these two types of habitats. Although larval salamanders are carnivorous, tadpoles (larval frogs) are generally herbivorous and play important roles in controlling vegetation levels in both lentic (still water) and lotic (flowing water) ecosystems. For example, streams that have lost tadpoles can become choked with aquatic vegetation. Tadpoles compete with aquatic insect larvae for food, and loss of tadpoles alters the macroinvertebrate community composition. Adult amphibians may prey on a variety of items depending on their size and lifestyle; burrowing caecilians generally consume earthworms, whereas terrestrial frogs and salamanders may consume anything from tiny terrestrial leaf-litter dwelling insects such as springtails, mites, and ants to other vertebrates (including other amphibians). Amphibian loss affects predator populations (e.g., snakes) and may also affect the population dynamics of prey taxa such as insects.

Systematics

Amphibian systematics is in a state of flux, adding to the challenge of assessing and tracking the scope of global decline. There are insufficient numbers of amphibian taxonomists, which hinders fundamental conservation assessments. Currently, amphibian taxonomists are concentrated in the more species-poor but wealthier countries; there is increasing recognition of the need for established resident taxonomists in species-rich regions. A new generation of taxonomists is developing in Asia and South America, but there are gaps in expertise in species-rich regions such as Africa. The marked recent increase in amphibian biologists has accelerated field studies in remote areas and discovery of many new species. Ironically, during a time of crisis for the world's amphibians, the number of described species has risen from 4000 in 1985 to over 6800 today.

Phylogenetic patterns of amphibian declines are not well understood. Some clades appear to be especially susceptible to declines. For example, frogs of the genus *Atelopus* from the highlands of lower Central America and the mountains of

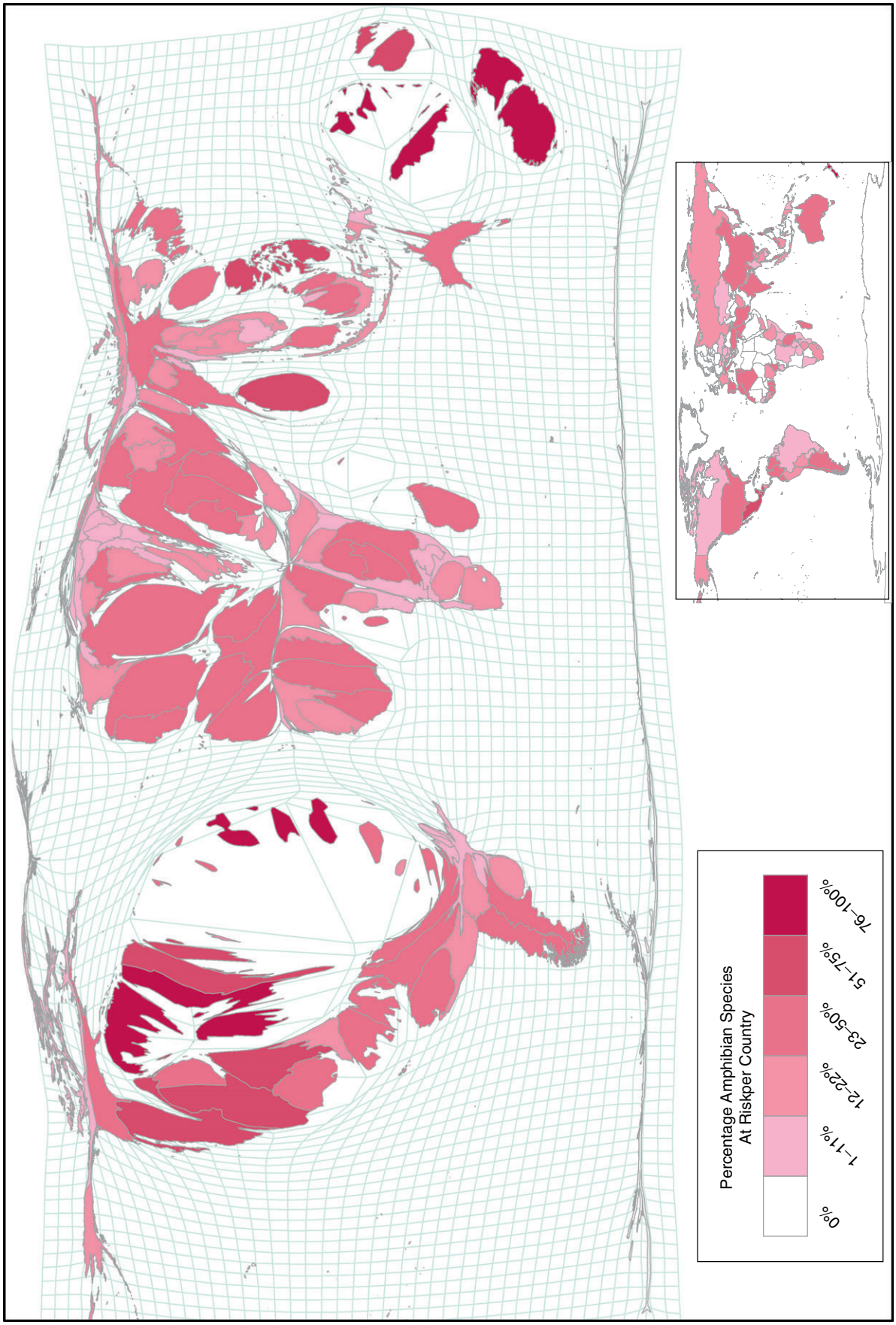


Figure 2 Percent of threatened and endangered amphibian species by country visualized using density-equalizing cartograms. Country size is purposely distorted in proportion to the percentage of threatened and endangered amphibian species, not including introduced species using the cartogram technique after Gastner and Newman (2004). (Data sources: AmphibiaWeb (website accessed 17 November 2010) and IUCN (website accessed 30 September 2010).)

northwestern South America have been hard-hit by declines; nearly 80% of described *Atelopus* species have been classified as critically endangered or extinct. These species have stream-adapted larvae. Elsewhere stream adaptation is also associated with declines, suggesting a phylogenetic bias within the clade containing *Atelopus*. Although salamanders seem to show less evidence of declines than frogs, neotropical salamanders (both ambystomatids and plethodontids) are experiencing severe declines. The vast majority of tropical salamanders are plethodontids, members of a single clade known as bolitoglossines, and declines of Mexican and Guatemalan species have been well documented (Rovito *et al.*, 2009). Although a number of studies have found a phylogenetic component to amphibian declines, data from one of the best-studied amphibian mass mortality events in Panama indicated that decline and extirpation were random with respect to phylogeny (Crawford *et al.*, 2010). Assessing the amount of phylogenetic diversity lost in declines and extinctions is also complicated by the fact that the amphibian fauna is not fully known in many species-rich areas.

Factors Responsible for the Declines

Potential causes for the widespread declines of amphibians can be grouped into two major categories: (1) factors general to the overall biodiversity crisis, including habitat destruction, alteration and fragmentation, introduced species, and overexploitation and (2) factors associated with amphibians that might account for declines in relatively undisturbed habitats. The first category includes relatively well understood ecological phenomena, whereas the second includes less well understood agents with complex underlying mechanisms, such as infectious diseases, climate change, increased UV–B radiation, chemical contaminants, and the causes of deformities. Habitat loss has played a major role, as has the infectious disease chytridiomycosis, but amphibian populations often face many simultaneous threats and synergistic interactions between different threats may exacerbate declines.

Habitat Degradation and Land Conversion

Perhaps the most obvious factor in the loss of amphibians is habitat degradation and land conversion. Although some amphibians are able to not only persist but thrive in urbanized or agricultural areas, many species require undisturbed (or relatively undisturbed) habitat. Loss of wetlands is a major factor in temperate zones, whereas removal of forests is the most serious threat in the tropics. Habitat fragmentation, where forest remnants persist but are separated by disturbed habitat such as pasture, is a serious issue for widespread species and in those that have natural metapopulation structures. Many amphibians, especially terrestrial species with direct development and no larval stages, are typically highly structured genetically and have cryptic diversity; in this case, habitat fragmentation leads to extensive loss of biodiversity as populations die out, even if species fragments persist. Human-induced disconnection between habitats used by different life history stages makes breeding migrations risky and may account for some of the reason why species with aquatic stages are declining more than strictly terrestrial species. Some habitat changes are subtle and have unexpected effects. Fire suppression in parts of western North America has led to encroachment of forests into breeding sites, with the resulting increase in shade having negative effects on tadpole development in some species.

Habitat degradation from pollution from agricultural herbicides, insecticides, and fertilizers impacts amphibians with aquatic life stages, either directly or by synergistic interactions with other factors (Hayes *et al.*, 2010). Chemical pollutants can directly kill larvae at certain concentrations, but even sublethal concentrations can affect behavior, reproduction, and life history, and can reduce the food supply (Relyea and Diecks, 2008). In some species, stress can magnify the effects of chemical doses that would normally not be lethal. We remain far from having a full understanding of the impacts of additives by humans to the environment on amphibians.

Impact of Exotic Species

The establishment and spread of exotic species are major threats to worldwide biodiversity and there are many examples of amphibians being affected. Exotic species can affect amphibians as competitors, predators, and as vectors for parasites and diseases. Nonnative amphibians may have contributed to many of the declines of native amphibian species. The bullfrog (*Rana catesbeiana*) is native to eastern North America but has been transported around the world by humans, mostly for human consumption. In California, wild populations of nonnative bullfrogs now eat and out-compete red-legged frogs (*Rana draytonii*) and foothill yellow-legged frogs (*Rana boylii*), both of which are in decline. In the last several decades, nonnative bullfrogs have escaped from commercial frog farms, particularly in South America, and established wild populations in areas where native frog populations collapsed. Although many native amphibians are dying from chytridiomycosis, bullfrogs appear to be resistant and are thought to act as carriers for the disease.

In Australia, sugar cane farmers introduced cane toads (*Bufo marinus*) in an effort to control insect pests, but the nocturnal cane toads do not control the largely diurnal pests. Since the first introductions in the late 1800s, cane toads have spread rapidly throughout eastern and more recently northern portions of Australia where they act as both competitors and predators of native amphibians, as well as apparently introducing pathogens and parasites. Cane toads produce toxins at all life stages, and ingesting cane toad eggs, larvae, or metamorphs is fatal to both larval and adult native Australian amphibians.

Another invasive amphibian species is *Xenopus laevis*, the African clawed frog, which is now widely established around the world. These fully aquatic frogs are prolific, highly invasive, and prone to *Bd* infection but resistant to chytridiomycosis. Like bullfrogs, African clawed frogs have been implicated in the spread of both chytridiomycosis and ranaviruses to native amphibians. They also out-compete native

amphibians, which decrease reproduction in ponds invaded by feral *Xenopus laevis*.

Fish are generally dominant species in aquatic systems and human-mediated introduction of fishes to new aquatic habitats has often had devastating consequences, especially for those amphibian species that evolved without fish predators. Fish-naive amphibian species may fail to recognize predator cues or lack behavioral responses to avoid fish predation. Introduced fish such as mosquitofish, carp, and trout have been shown to prey on amphibian eggs and larvae. Trout are native to cold mountain streams in North America and Europe but have been widely introduced to every continent except Antarctica. In the Sierra Nevada of California, more than 99% of the lakes and ponds above 2,100 m were fishless prior to the mid-1800s. Since then, trout have been introduced into more than 90% of Sierra Nevada lakes for sport fishing. Mountain yellow-legged frogs, once the most common vertebrates in these high-elevation ponds and lakes, are adapted to life in environments without any fish and have declined dramatically with the introduction of trout. Field experiments showed that removal of introduced trout from entire lakes could lead to recovery of local mountain yellow-legged frog populations. However, soon after this discovery chytridiomycosis was detected in the Sierra Nevada and has spread rapidly, further endangering the last remaining populations of these species. Paedomorphic salamanders (which retain certain larval traits and remain aquatic into adulthood) can also be particularly vulnerable to fish predation. Another introduced aquatic predator known to feed on amphibian larvae is the red swamp crayfish (*Procambarus clarkii*, which is native to North America but has been introduced on all continents except for Antarctica and Australia) and is associated with amphibian population collapses in Portugal.

Fish compete with amphibians (and birds) for aquatic insect prey. Presence of introduced insectivorous fish such as trout in montane lakes of northern California was correlated with decreased diversity of native aquatic insects and far less insect emergence. Removing introduced trout from northern California lakes led to the recovery of invertebrate populations.

Introduced species can exert effects on native amphibians indirectly by altering the habitat. Exotic fish such as carp or tilapia can increase water turbidity in shallow water bodies, which decreases productivity and is thought to be one factor in the decline of native axolotls in Mexico.

Infectious Diseases

Infectious diseases have been associated with collapsing amphibian populations almost everywhere that amphibians occur (Fisher *et al.*, 2009; Gray *et al.*, 2009). Most of the reported amphibian mass mortality events have been associated with the emerging fungal pathogen *Batrachochytrium dendrobatidis* and viruses of the family Iridoviridae, but other agents such as bacteria, water molds, and trematode parasites have also been associated with varying levels of mortality and population decline.

Chytridiomycosis, an amphibian disease caused by the pathogenic fungus *Batrachochytrium dendrobatidis* (*Bd*), plays a major role in amphibian population collapses in both protected and disturbed habitats around the world and has been called "perhaps the most virulent threat affecting vertebrates to emerge in recent years" (Hoffman *et al.*, 2010). Catastrophic *Bd*-related declines have occurred in Central and South America, Spain, Australia, and California (Collins and Crump, 2009). Chytridiomycosis is of great concern in general because it is the first emerging disease that causes declines, and perhaps even extinctions, of many species not otherwise threatened. Scientists debate whether *Bd* is a novel pathogen, with the spread to new host species and new geographical areas mediated by humans, or an endemic pathogen that has increased in virulence, possibly because amphibians have been rendered more vulnerable by environmental changes. When the fungal load reaches a critical threshold, it fatally compromises electrolyte transport, causing cardiac arrest (Voyles *et al.*, 2009). The ecology of *Bd* remains poorly understood but we do know that density-dependent host–pathogen dynamics are key components (Briggs *et al.*, 2010; Vredenburg *et al.*, 2010). Many questions remain unanswered: Is reproduction of *Bd* ever sexual? Are there resting life stages? How is *Bd* transported?

Viruses belonging to the family Iridoviridae (ranaviruses) have been associated with mass mortality in both frogs and salamanders, in wild populations as well as in captivity. Like *Bd*, spread of these viruses appears to have been at least partly mediated by humans via the global amphibian trade. Ranaviruses appear to have undergone recent host shifts from fish to amphibians.

Other agents contributing to amphibian declines include gram-negative bacterial pathogens; the best-known of these is *Aeromonas hydrophila*, which causes red-legged disease and has been responsible for local extirpations of wild ranid frogs in California and Rhode Island. Pathogenic water molds (*Saprolegnia* sp.) have been found to be responsible for local extinctions in some North American frogs, as well as massive egg mortalities in some European frogs, and to increase mortality in at least one North American salamander species. Trematode infestation has been implicated in limb deformities in the Pacific treefrog (*Pseudacris regilla*) and several other species of amphibians, and to cause tadpole mortality, but so far no major population declines have been tied to trematode infestations. Declines in the Wyoming toad (*Bufo baxteri*) have been attributed to both *Bd* and another parasitic fungus, *Basidiobolus ranarum*.

Human Exploitation

The impact of harvesting wild species for the global pet trade and for human consumption is of increasing concern (Schloegel *et al.*, 2009; Gratwicke *et al.*, 2010). Solutions to these global problems, which often are tied intimately to local economic conditions and traditions, ultimately depend on effective public education.

Factors Associated with Global Climate Change

Amphibians may be more sensitive to climate change than other vertebrates such as birds because amphibians are more

likely to show habitat and microhabitat specialization and because they are significantly less vagile. Many of the effects of climate change might be indirect, involving such factors as manifestations of disease or pathogen-host dynamics, and may involve interactions with other stressors (Blaustein *et al.*, 2010), making direct detection of impacts difficult.

Elevated UV–B radiation

Global atmospheric changes caused by anthropogenic activities are well documented and one result is a reduction of stratospheric ozone, leading to an increase in the amount of biologically damaging ultraviolet radiation (UV–B) reaching the Earth's surface. Increase in UV–B has been hypothesized to contribute to increased mortality rates in amphibians with the idea that this might explain enigmatic declines in protected areas. Tests of this hypothesis have focused mainly on comparing egg-hatching rates in species that lay their eggs in shallow, exposed breeding sites that should be subjected to high levels of UV–B. Field experiments have concluded that many species are sensitive, but others are not, even at high-elevation sites where the ambient UV–B dose is high. Interactions between UV–B and abiotic factors such as water chemistry are important in some instances. Synergism between UV–B exposure and stress from the risk of predation can increase tadpole mortality in a sensitive species by two-fold over the additive effects of exposure to either stressor alone. Also, not all species that are known to be sensitive to UV–B are in decline. Although UV–B may be involved in synergistic interactions with other factors (Bancroft *et al.*, 2008; Blaustein *et al.*, 2010), at this point it does not appear to be a major component of amphibian declines in general.

Temperature Variability and Amphibian Immunity

Elements of climate change are thought by some to make conditions more favorable for amphibian diseases and parasites (Pounds *et al.*, 2006), but the causal link between climate change and disease virulence is unclear and debated. The rapid spread of *Bd* appears to be sufficient to cause mass amphibian mortalities without invoking climatic change (Lips *et al.*, 2008). However, stress from climate change (in particular, increased temperature variability) may decrease frog immunity, making it easier for pathogens to cause death (Rohr *et al.*, 2008; Rohr and Raffel, 2010).

Threats from Shifts in Weather Patterns

Some species react to changes in climate by shifting their distributions either latitudinally or altitudinally (Rovito *et al.*, 2009). Amphibians, especially those in tropical regions, are narrowly distributed; often they are restricted to specific and narrow elevational belts on mountains that are themselves isolated. There may be no place to go when climates are warm or dry, except higher up the mountain. Species already at the top of such habitats literally are pushed off the mountain, into extinction. Other species may be limited in movements to the north or south by ranges fragmented by habitat conversion or by barriers such as rivers that effectively stop range expansion.

A documented effect of climate change has been the earlier breeding of some species in the northern parts of their ranges. The common frog (*Rana temporaria*) in Great Britain must advance the date of first spawning by three to five weeks in order to accommodate increasing temperatures, but will only be able to advance by five to nine days (Phillimore *et al.*, 2010), so adaptation for this species and other amphibian species may fall well short of the challenge.

In Central America, the cloud line has risen several hundred meters, which apparently precipitated a drought of unusual severity in a high-elevation cloud forest in Costa Rica (Monteverde) in the late 1980s (Pounds *et al.*, 1999). This drought has been implicated in the disappearance of 20 of the 50 species of amphibians known from the site, including a local endemic, the golden toad, *Bufo periglenes*. But it is not clear whether this was the result of an exceptional event (an unusual dry season) or a more general effect, and chytridiomycosis has also been implicated in the disappearance of Monteverde amphibians.

Synergistic Effects

Many factors by themselves pose severe threats to amphibian survival, but synergistic interactions between factors can magnify the negative effects on amphibians. Many different synergistic effects have been suggested, likely posing major threats to the continued existence of numerous amphibian species (Pounds *et al.*, 1999; Relyea and Diecks, 2008; Blaustein *et al.*, 2010). For instance, the effects of infectious diseases may be greater in the presence of unusual weather conditions, elevated UV–B, or chemical pollutants, which may compromise amphibian immune systems. Unusual weather conditions (particularly drought) might enhance the impact of different stressors. The link between climate change and pathogen growth conditions in Central America may be enhanced by the presence of high environmental loads of pesticides in the area, which may be an additional stressor. Hierarchical approaches to defining the causes of declines (Hayes *et al.*, 2010) hold promise for identifying critical factors that might be subjects for further study.

Challenges and Opportunities for the Future

Many of the most dramatic instances of declines have occurred in protected areas, such as the great national parks of the Sierra Nevada of California, the Monteverde Cloud Forest Preserve in Costa Rica, and protected areas in Australia, to give three prominent examples. Thus the standard conservation approach of purchasing and protecting land and habitats may not ensure survival, but remains an essential strategic component for maintaining ecosystem function. Conservation strategies must involve both specialist and generalist researchers with diverse talents in infectious disease ecology, reproductive biology, endocrinology, immunology, and pollution ecology, as well as knowledge of the natural history of affected species. *Ex situ* strategies may be essential in particular cases, but captive breeding is difficult and expensive and careful thought must be given to selection of candidate species (e.g., phylogenetic, ecological, and behavioral diversity should be represented). A greater understanding of synergistic effects is required in order to counteract the diverse threats facing this ancient group of organisms.

Mitigation of Habitat Changes

Habitat fragmentation is a cause of declines that might be mitigated by re-establishing habitat corridors and connectivity, such as vegetated freeway overpasses, reforestation, and road tunnels to facilitate wildlife movement.

Wetland restoration holds promise for recovery of some amphibian species. Wetlands are important components of the earth ecosystems, providing vital ecosystem services (i.e., water purification). Mitigation in wetlands not only benefits humans directly by helping restore clean water systems, but also provides important habitat for many species of amphibians.

Removal of Exotic Species

The removal of exotic species has been effective in restoring amphibian populations in some instances. For example, removal of exotic fishes has had positive benefits for amphibians in Spain, Chile, and in several areas within the United States. In California, removal of nonnative trout from entire lakes in the Sierra Nevada led to rapid recovery of threatened mountain yellow-legged frog populations. However, whether these population recoveries will translate into species persistence is undetermined; only 1 of 209 amphibian species examined (the Mallorcan midwife toad, *Alytes muletensis*) improved in conservation status due to mitigation of invasive species introductions (Hoffman *et al.*, 2010).

Attenuation of Infectious Agents

Disease ecology seeks to understand the mechanisms that lead to disease outbreaks in natural systems. Global trade, and thus the vast increase in connectivity between continents, has been linked to many of these outbreaks (Schloegel *et al.*, 2009; Gratwicke *et al.*, 2009). Precautions necessary to lower the probabilities of future outbreaks include limiting transport of diseases by use of disinfection and quarantine programs and by stopping human movement of disease vectors (i.e., live bullfrogs). Chytridiomycosis and ranaviral infections are now on the list of notifiable diseases (World Organization for Animal Health, Office International des Epizooties), meaning that all 177 member countries must report on the status of *Bd* and ranaviruses within their borders every six months.

There are a few glimmers of hope for amphibians. Despite the fact that most populations of mountain yellow-legged frogs have been decimated, a few with known infections of *Bd* have persisted for years. Some species of Australian frogs that survived epidemics of chytridiomycosis appear to be re-establishing themselves and to have relatively robust antimicrobial defenses.

Captive Breeding

Although some biologists view captive breeding as a last-resort conservation action, the IUCN endorses captive breeding as a proactive conservation measure, one that should be initiated while a species is still available. IUCN, through its Amphibian Specialist Group, sponsors the Amphibian Ark, which tracks *ex situ* captive breeding activities for amphibians around the world. However, captive breeding is costly and many hurdles must be overcome (including getting a species to breed successfully in captivity and ensuring that disease is not transmitted) before captive-bred individuals can be reintroduced to the wild (Griffiths and Pavajeau, 2008).

Implications for the Biodiversity Crisis in General

Amphibian declines represent the leading edge of the extinction wave (Wake and Vredenburg, 2008). As humans continue to alter the environment on a global scale, amphibian declines may be the window into the future of what we can expect for global biodiversity: extinctions on a massive scale, well beyond anything we have ever experienced. Despite some local successes, current conservation efforts are insufficient to offset global biodiversity losses. Meanwhile, amphibian declines continue. Although the past decade has seen a dramatic increase in research directed to the multiple and diverse factors associated with declines and significant progress has occurred in many areas, challenges remain. Human activities are a primary source of the problems facing amphibians and, ironically, also the main hope for the future.

See also: Amphibians, Biodiversity of. Captive Breeding and Reintroduction. Captive Breeding and the Evolutionarily Significant Unit. Climate Change and Extinctions. Diseases, Conservation and. Endangered Amphibians. *In Situ*, *Ex Situ* Conservation. Global Species Richness. Human Impact on Biodiversity, Overview. Indigenous Peoples and Biodiversity. Introduced Species, Impacts and Distribution of. Mass Extinctions, Notable Examples of. Ultraviolet Radiation. Wetland Creation and Restoration. Wetlands Ecosystems

References

Bancroft B, Baker N, and Blaustein A (2008) A meta-analysis of the effects of ultraviolet B radiation and its synergistic interactions with pH, contaminants, and disease on amphibian survival. *Conservation Biology* 22: 987–996.

Blaustein AR, Walls SC, Bancroft BA, *et al.* (2010) Direct and indirect effects of climate change on amphibian populations. *Diversity* 2: 281–313.

Briggs CJ, Knapp RA, and Vredenburg VT (2010) Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. *Proceedings of the National Academy of Sciences USA* 107: 9695–9700.

Collins JP and Crump ML (2009) *Extinction in our times: Global amphibian declines*. Oxford: Oxford University Press.

Crawford AJ, Lips KR, and Bermingham E (2010) Epidemic disease decimates amphibian abundance, species diversity, and evolutionary history in the highlands of central Panama. *Proceedings of the National Academy of Sciences USA* 107: 13777–13782.

Fisher MC, Garner TWJ, and Walker SE (2009) Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. *Annual Review of Microbiology* 63: 291–310.

Gastner MT and Newman MEJ (2004) Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences USA* 101: 7499–7504.

Gratwicke B, Evans MJ, Jenkins PT, *et al.* (2010) Is the international frog legs trade a potential vector for deadly amphibian pathogens? *Frontiers in Ecology and the Environment* 8: 438–442.

Gray MJ, Miller DL, and Hoverman JT (2009) Ecology and pathology of amphibian ranaviruses. *Diseases of Aquatic Organisms* 87: 243–266.

Griffiths RA and Pavajeau L (2008) Captive breeding, reintroduction, and the conservation of amphibians. *Conservation Biology* 22: 852–861.

Hayes TB, Falso P, Gallipeau S, *et al.* (2010) The cause of global amphibian declines: a developmental endocrinologist's perspective. *The Journal of Experimental Biology* 213: 921–933.

Hoffman M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science* 330: 1503–1509.

Lips KR, Diffendorfer J, Mendelson JR III, *et al.* (2008) Riding the wave: reconciling the roles of disease and climate change in amphibian declines. *Public Library of Science Biology* 6: 441–454.

Phillimore AB, Hadfield JD, Jones OR, *et al.* (2010) Differences in spawning date between populations of common frog reveal local adaptation. *Proceedings of the National Academy of Sciences USA* 107: 8292–8297.

Pounds JA, Bustamante MR, Coloma LA, *et al.* (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439: 161–167.

Pounds JA, Fogden MPL, and Campbell JH (1999) Biological response to climate change on a tropical mountain. *Nature* 398: 611–615.

Relyea RA and Diecks N (2008) An unforeseen chain of events: lethal effects of pesticides on frogs at sublethal concentrations. *Ecological Applications* 18: 1728–1742.

Rohr JR and Raffel TR (2010) Linking global climate and temperature variability to widespread amphibian declines putatively caused by disease. *Proceedings of the National Academy of Sciences USA* 107: 8269–8274.

Rohr JR, Raffel TR, Romansic JM, *et al.* (2008) Evaluating the links between climate, disease spread, and amphibian declines. *Proceedings of the National Academy of Sciences USA* 105: 17436–17441.

Rovito SM, Parra-Olea G, Vásquez-Almazán CR, *et al.* (2009) Dramatic declines in neotropical salamander populations are an important part of the global amphibian crisis. *Proceedings of the National Academy of Sciences USA* 106: 3231–3236.

Schloegel LM, Picco AM, Kilpatrick AM, *et al.* (2009) Magnitude of the US trade in amphibians and presence of *Batrachochytrium dendrobatidis* and ranavirus infection in imported North American bullfrogs (*Rana catesbeiana*). *Biological Conservation* 142: 1420–1426.

Stuart S, Chanson JS, and Cox NA (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786.

Stuart S, Hoffmann M, Chanson J, *et al.* (eds.) (2008) *Threatened Amphibians of the World. USA: Lynx Edicions, IUCN, and Conservation International.* Spain: Barcelona.

Voyles J, Young S, Berger L, *et al.* (2009) Pathogenesis of chytridiomycosis, a cause of catastrophic amphibian declines. *Science* 326: 582–585.

Vredenburg VT, Knapp RA, and Tunstall TS (2010) Dynamics of an emerging disease drive large-scale amphibian population extinctions. *Proceedings of the National Academy of Sciences USA* 107: 9689–9694.

Wake DB and Vredenburg VT (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences USA* 105: 11466–11473.

Globalization Effects on Common Plant Species

Thomas J Stohlgren, U.S. Geological Survey, Fort Collins, CO, USA
Petr Pyšek, Academy of Sciences and Charles University, Prague, Czech Republic
John Kartesz and Misako Nishino, Biota of North America Program, Chapel Hill, NC, USA
Aníbal Pauchard, Universidad de Concepción and Institute for Ecology and Biodiversity (IEB), Concepción, Chile
Marten Winter, Helmholtz Center for Environmental Research - UFZ, Leipzig, Germany
Joan Pino, University of Barcelona, Barcelona, Spain
David M Richardson, Stellenbosch University, Matieland, South Africa
John Wilson, South African National Biodiversity Institute (SANBI), Claremont, South Africa
Brad R Murray and Megan L Phillips, University of Technology Sydney, Sydney, NSW, Australia
Laura Celesti-Grapow, Università La Sapienza, Rome, Italy
Jim Graham, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alien species A species introduced from another region or ecosystem.
Globalization The process engaging financial markets to operate internationally, largely as a result of increased trade, transportation, and communications.
Homogenization A blending of species throughout the globe largely as a result of trade and transportation.
Invasive species A species from another country or ecosystem that causes economic or environmental harm, or harms human health.
Native species A species that evolved and is naturally distributed in the region.

Introduction

The trade of goods by humans has bridged the continents, in effect restoring the old supercontinent of Pangaea. In the past century, humans have been responsible for an exponential increase in plant migrations, moving plant species around the globe for food, fuel, forage, horticulture, landscaping, and medicines. Trade within and among continents is breaking geographic barriers and providing long-range dispersal for seeds and propagules at unprecedented rates (Richardson *et al.*, 2000; Wilson *et al.*, 2009). Here, the authors provide a brief review of the globalization effects on common plant species and "homogenization" of the world's plant communities.

It's All About Trade

Globalization is a process that occurs when the trade in goods and service is no longer limited by the physical and administrative barriers that separate countries. Local trading has been a common feature within many cultures. Through human expansion around the globe, increasingly reliable and quick forms of transport, and fewer restrictions on human movement, trade now operates at a global scale. The dawn of long-distance merchant ships in the Minoan, Greek, and Roman civilizations connected Europe, Africa, and Asia. As long ago as 67 BC, the Roman Empire had cargo ships carrying 70,000 kg of grain and other supplies from far and wide (Kessler and Temin, 2007). Indeed, many current invasive plant species in Italy can be traced to early trade. Trade in crops accelerated in the Old World in the mid-eighth to mid-thirteenth centuries AD, with Arab people bringing a range of food items to the Mediterranean. In the New World, Native Americans were trading obsidian, shells, tobacco, maize, squash, and dozens of other plant products.

In the fifteenth to seventeenth centuries AD, the Age of Discovery accelerated the global exchange of agricultural products. Foods such as potatoes, tomatoes, maize, and other agricultural crops, for instance, moved from the New World to the Old World (Crosby, 1972). In the late eighteenth century, the isolated southern-hemisphere continent of Australia was brought into somewhat permanent connection with the rest of the world, when it was established by the British as a penal colony. Since then, over 26,000 alien plant species have been introduced to Australia with periods of high influx over this time (Phillips *et al.*, 2010).

International trade skyrocketed between 1870 and 1913 and after 1950, the volume of world trade merchandise increased seventeen-fold (Flemming, 2004). In the modern day, most of the goods consumed by humans stem from imports (**Figure 1**). As the global economy expanded, so did trade in live plant materials, with ornamental horticulture an important source of introduced invasive plants (Reichard and White, 2001). For example, Maki and Galatowitsch (2004) found that prohibited alien species in Minnesota were acquired in 92% of orders they placed with 34 aquatic plant vendors across the United States. Additionally, 93% of orders contained a plant or animal species not specifically requested, including several additional prohibited taxa. The advent of e-commerce has turned every home and office computer into a virtual port of entry. Live plants and seeds can be mailed to most addresses in many countries.

The comparison between regions with similar climates but with different history of trade may help to elucidate the causes

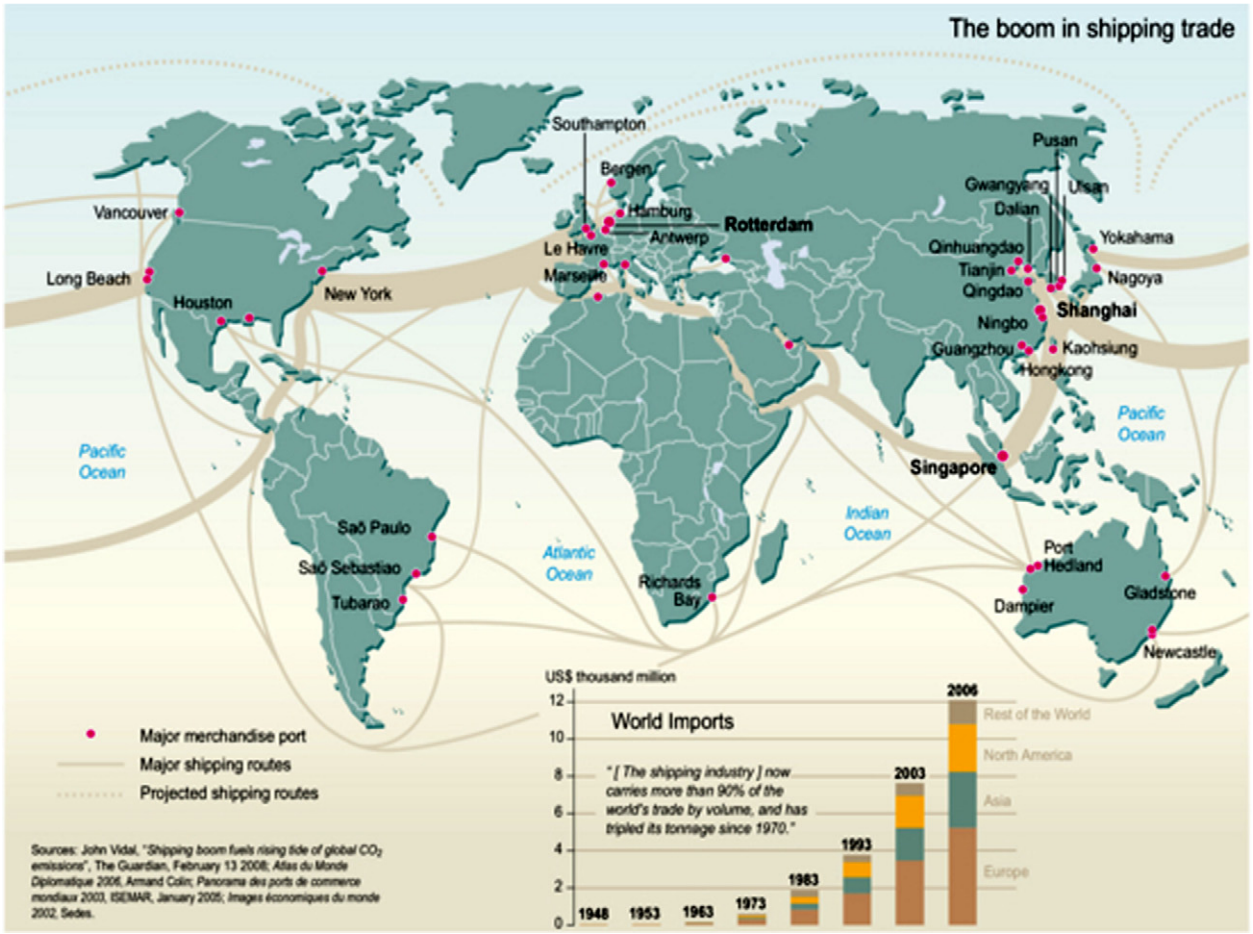


Figure 1 The boom in the shipping trade, reproduced from the website (http://maps.grida.no/go/graphic/the-boom-in-shipping-trade1), and based on John Vidal, "Shipping boom fuels rising tide of global CO_2 emissions," The Guardian, February 13, 2008; Atlas du Monde Diplomatique 2006, Armand Colin; Panorama des ports de commerce mondiaux 2003, ISEMAR, January 2005; Images économiques du monde 2002, Sedes.

of current patterns of homogenization. For example, although central Chile and California share a similar Mediterranean climate, the history of massive introductions in the last two centuries have favored a much more diverse alien flora in California (*ca.* 1200 species) (Jiménez *et al.*, 2008). Interestingly, the Chilean alien flora (*ca.* 600 species) is almost completely a subset of its counterpart, including 25 species native to California. Central Chile has a much stronger influence of European species, and because of its higher commercial trade and diverse immigration history, California has a more cosmopolitan flora with a higher proportion of species from Asia, Africa, Australia, and other regions of the Americas.

Homogenization of Biota

Today, plant species from the modern nursery trade often do not stay in their port of entry or primary destination (Reichard and White, 2001). Wind, water, birds, and other animals (including humans), carry propagules to many destinations, where they grow and reproduce. The naturalization of alien species is further facilitated by urbanization, agriculture, and grazing by domestic livestock. The combination of trade, human-caused land use change, and extinction of native species result in biotic homogenization across the globe (Sax *et al.*, 2002). Homogenization is defined as the increasing similarity of species composition due to the concurrent extirpation of rare or even unique native species, and the increase of common, widespread alien species primarily due to the effects of modern humans on the rate of change (McKinney and Lockwood, 1999). Generally, local species assemblages lose their uniqueness, affecting also their conservation value (Pino *et al.*, 2009; Winter *et al.*, 2009). Researchers are just beginning to explore patterns of homogenization at intercontinental and global scales using species frequencies and abundances to provide a more complete view of biotic homogenization (Winter *et al.*, 2010). However, in most regions of the world alien invasions exceed regional extirpations, and hence species richness increases (Stohlgren *et al.*, 2008).

Uneven Results of Globalization

So to what extent have floras become globalized? The degree to which a flora is composed of alien species has most often been measured by the relative numbers of native and alien

species. However, how widespread alien species are in invaded regions is an equally important measure of their success. Recently, a team of scientists from around the world collated the best available data from several countries and regions to see if alien plant species were as widely distributed as native plant species (Stohlgren *et al.*, 2011).

There are, of course, profound differences in how native and alien floras have developed. Native species have had evolutionary time to disperse into the available ecological niches in their respective regions. If alien plant species became equally as widely distributed as native species in a much shorter time, this would indicate that alien plant species and the introduction processes are functioning, in terms of distributional range, in a similar way to native plant species. However, even if this was true, there would still be a principle difference in that these processes would have been operating much faster than in native species, possibly reflecting extraordinary dispersal abilities of aliens and their high competitiveness in colonizing new habitats, for reasons believed to determine the success of alien invasive species (Pyšek and Richardson, 2007). If aliens are more widely spread than natives, this would indicate that part of the invasion process is operating differently to background distributional range expansion and occupancy (i.e., the alien species are different, the environment has changed to favor aliens, or human-assisted dispersal enables aliens to gain wider distributions). If aliens have narrower distributions than natives, this suggests that either aliens tend to not succeed widely across a new range of environments, or that these species have not reached their distributional limits, and that we may only be on the cusp of a homogenization tsunami.

The team of scientists gathered information on the distributions of the 120 most widely distributed plant species in various region of the world, based on their frequency in defined subregions (ecological zones or administrative units) or grid cells (typically 10 km × 10 km). Species were characterized as "native" or "alien," the latter being plant species not native to the biogeographical region or country, having arrived by means of human activities. The initial cutoff point of 120 species was arbitrary, but based on general knowledge of plant frequency distributions (Stohlgren, 2007). The sample sites differed greatly in area and species richness, so they also evaluated the top 120 species in three US states (California, Texas, and Florida), and three European regions (Great Britain and Ireland, the Czech Republic, and Catalonia in Spain) to see if patterns held at smaller spatial extents. What they found surprised them.

First, there were large differences between areas in the percentage of the widespread floral species that was alien (**Figure 2**; Stohlgren *et al.*, 2011). In North America, over half of the most widely distributed species were from other continents, primarily Eurasia. Alien plant species were also well represented in regions of Australia, South America, and South Africa, where they contributed between 22% and 43% to the total number of most widely distributed species, but European regions contained few widespread aliens, <2%, despite thousands of years of long-distance trade in plants. Second, individual subregions in Europe and North America varied considerably in the percentage of aliens on the 120 most widely distributed species lists (**Figure 2**). In subregions, the percentage of aliens can be greater or less than that of the surrounding region. For example, California with 5688 total native plant species had 20% aliens on the list, while Florida, with 3284 total native plant species had only <1% (Stohlgren *et al.*, 2011). Third, the most surprising and consistent finding, was that aliens on the lists were as widely distributed as their native species counterparts, regardless of scale and extent of the study areas. Aliens were significantly more widely

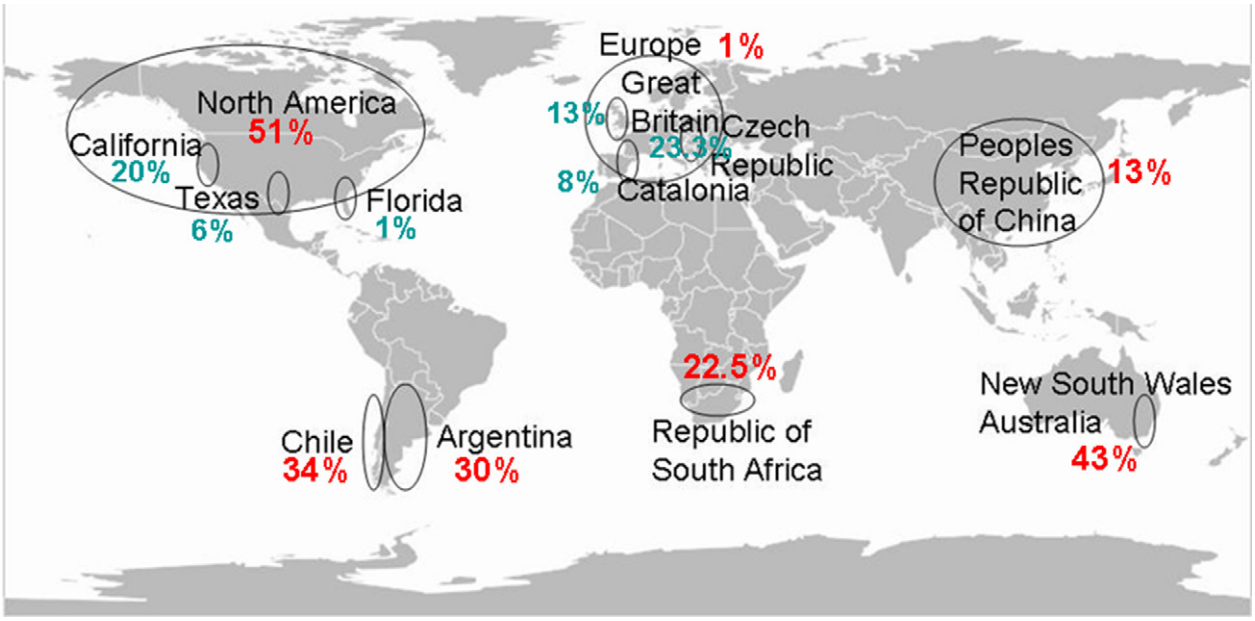


Figure 2 Relationship of percent alien species in the 120 most widely distributed species in a region. Reproduced from Table 1 in Stohlgren TJ, Pyšek P, Kartesz J, *et al.* (2011) Widespread plant species: Natives vs. aliens in our changing world. *Biological Invasions*. doi: 10.1007/s10530-011-0024-9, with permission from Springer.

distributed than native plant species in North America, Argentina, Chile, and the Republic of South Africa, and equally widely distributed as native species in California, Texas, Florida, and throughout Europe.

Donor and Receptor Regions

To evaluate donor and receptor regions for common plants around the world, it is helpful to assess how common plant families overlap across continents. Broadly spread plant families include Asteraceae, Brassicaceae, Poaceae, and Fabaceae. Likewise, many families often overlap between native and alien species within countries or regions. In the aforementioned study by Stohlgren *et al.* (2011), the Poaceae and Asteraceae were more broadly spread in North America, South Africa, and New South Wales (across origin status). The Rosaceae family included widely spread native and alien species in Great Britain and Ireland.

In terms of particular regions, Europe and Asia contributed many species to lists of widely distributed species in the New World (Stohlgren *et al.*, 2011). As such, Europe appears to be a better donor than receptor of plant species (Pyšek, 1998; Winter *et al.*, 2009; **Figure 2**), while North America proves to be a melting pot of its flora. However, every region on the globe must be considered a potential donor of alien species.

Stohlgren *et al.* (2011) also found many examples of widely distributed species, which help us to understand donor and receptor regions for plants. For example, *Chenopodium album*, contains varieties from Europe (*Chenopodium album* var. *album*) that are now widely dispersed in North America. Likewise, *Cirsium vulgare*, a native species in Europe, is a widely dispersed alien species in other parts of the world. In these cases, widely distributed native species became widely distributed alien species. The scientists also found that some species were more widely spread in their invaded habitat compared to their native homelands. *Senecio vulgaris* ranges broadly in Catalonia, and less so in Great Britain and Ireland, but was widespread throughout North America. Likewise, *Agrostis stolonifera*, a species native to the United States and temperate Europe, was found to be narrowly ranged in Great Britain and Ireland, but less so in Catalonia, and a widespread weed in Chile. However, *Bromus tectorum*, a well known Eurasian invader in North America, did not appear on any other New World lists. Likewise, *Opuntia ficus-indica*, is the most widely distributed alien invader in South Africa, and is now common in Argentina, but is not as relatively widespread in its native Mexico.

Selected Examples

There are many examples of common plant species now spreading around the world. Knowledge of the distributions of native and alien species around the globe provides the opportunity to understand the traits of highly invasive species, and thereby develop effective monitoring and prevention programs against the harmful spread of invasive species. The following examples serve to illustrate the global and continental scales of invasions (**Figure 3**).

Water Hyacinth (*Eichhornia crassipes*)

Water hyacinth is a perennial, free-floating aquatic plant native to tropical regions of South America, and now present on all continents except Antarctica. Plants rapidly increase biomass and form dense mats, reproducing from stolons (i.e., vegetative runners). Water hyacinth can completely cover lakes and wetlands, outcompeting native aquatic species, reducing oxygen levels for fish, and creating ideal habitat for disease-carrying mosquitoes. Large infestations of water hyacinth can prevent river transport, fishing, damage bridges, and clog dams. Lake Victoria, Africa, and the water-ways of Papua New Guinea are prime examples where massive populations have limited transportation and fishing, and increased the incidence of diseases (Masifwa *et al.*, 2001).

Within the US, water hyacinth is thought to have been introduced to Louisiana in the 1880s, and released later in the St. John's River in Florida. Enthusiasts distributed it around the world and water hyacinth was naturalized in Egypt, India, Australia, and Java well before the end of the nineteenth century. It is now found across the tropics and in some subtropical countries, for example, New Zealand and Portugal. Despite regional bans on its transport, and various control efforts, it has invaded many new areas particularly in Africa. Water hyacinth is a popular aquarium plant, and can be purchased at many aquatic plant nurseries and online distributers around the world.

Castor Oil Plant (*Ricinus communis*)

Native to Africa and the Middle East, Castor oil plant is a member of the spurge family, Euphorbiaceae. It produces seeds containing between 40% and 60% oil rich triglycerides. The plant is a fast growing shrub or tree that can reach a height of 12 m or more in tropical areas. Today, the plant is widespread in tropical regions of the world, largely as a crop. Castor seed production is around one million tons per year, with major growers in India, China, Brazil, and Ethiopia (www.faostat.fao.org). Horticulturists have created and sold several cultivars as landscaping trees and shrubs, adding to its spread. The seeds of this species, which are commercially widely available, are dispersed by both wind and ants over long distances. Once established, castor oil plant can invade natural areas such as those found in Hawaii, Florida, and elsewhere, creating problems for land managers trying to contain this fast-growing plant.

Alligator Weed (*Alternanthera philoxeroides*)

Alligator weed, native to South America, is an amphibious plant species in the Amaranthaceae family. It has become highly invasive in China, Australia, New Zealand, Thailand, and the United States. Like water hyacinth, it grows quickly to reduce water flow, canopy light, and oxygen levels in the water column, thus reducing habitat quality for various wetland species. However, alligator weed can also invade dry lands and agricultural fields. In the US, it has become a species of great concern in south Texas, and throughout most of Florida. In the US, it has now spread to 337 counties and 15 states, with recent records extending as far north as Virginia and Illinois. It

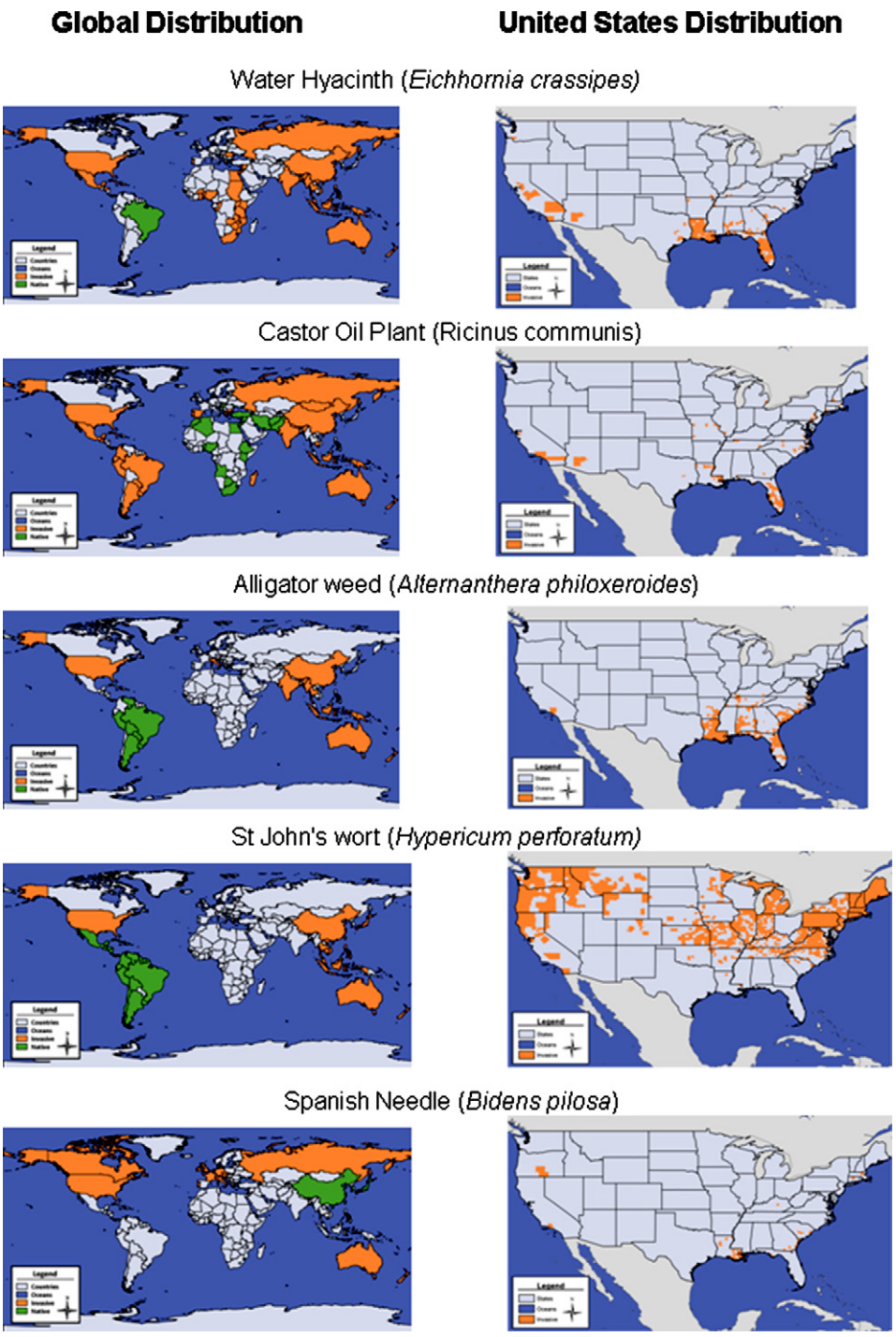


Figure 3 Maps of selected alien species at global scales, and in the United States.

spreads easily as small stem fragments can quickly produce new plants.

St. John's Wort (*Hypericum perforatum*)

St. John's wort is a yellow-flowering perennial herb native to Europe, grows wild in North America, Asia, India, Australia, South Africa, and many islands. It is a member of the Hypericaceae family (formerly placed in the Clusiaceae family), with a long history of human uses including warding off evil, and serving as an antidepressant. St. John's wort is currently treated as a noxious weed in many areas, because high concentrations of hyperforin and hypericin make it toxic to livestock. Throughout the US, populations of St. John's wort can be found in rocky open ground, wooded areas, and meadows. Seeds, which adhere to wildlife and livestock for easy dispersal are commercially available for wildflower plantings.

Spanish Needle (*Bidens pilosa*)

Spanish Needle (*Bidens pilosa*), an annual member of the Asteraceae family, has long been called Hairy Beggarsticks (in the US). The species is native to South America, but it is now commonly established throughout tropical and subtropical areas of the world. It is also common on many islands in the Indian Ocean. Spanish Needle can dominant plant cover in cultivated fields, plantations, and other cultivated and ruderal areas. Considered one of the most noxious annual weeds in East Africa, it invades fields of over 30 crops and is now well established in more than 40 countries (Holm *et al*., 1991). The plant is thought to produce allelopathic toxins that affect a number of crops. The colorful white and yellow flower produces 1.8 cm long barbed seeds that adhere to the legs of humans and other mammals, representing an effective means of dispersal. In some parts of the world, such as in sub-Saharan Africa, the plant is used for food or medicine (Grubben and Denton, 2004).

Giant Hogweed (*Heracleum mantegazzianum*)

One of the most impressive invader weeds of Europe is the Asian hogweed genus (*Heracleum*) of the family Apiaceae. Several members of the genus were introduced as garden ornamentals or as fodder crops outside their native range. The most distinctive characteristic of these closely related species is their conspicuous appearance; they can attain heights of up to 4–5 m, which ranks them among the tallest and largest herbs in Europe. The most widely distributed member of the genus, giant hogweed (*Heracleum mantegazzianum*) produces phototoxic sap, dangerous to humans. The first known record of this species was reported in Kew Botanic Gardens, London in 1817. In 1828, the first naturalized population was recorded in the wild in Cambridgeshire, England. Soon afterwards, the plant began to spread rapidly across Europe; and is currently recorded from 19 European countries. A likely reason for its widespread distribution in Europe is due to multiple introductions, as suggested *via* genetic analysis of invading and native populations of this species. Giant hogweed has a number of characteristics that contribute to its invasiveness: including its extremely high fecundity, rapid growth rate, capability of selfpollination, extended germination period by means of short-term persistent seed bank, high germination, and negligible impact of natural enemies (Pyšek *et al*., 2007).

Implications

For reasons still unknown to us, some geographic areas are more easily invaded than others (Richardson *et al*., 2005). General observations suggest that areas with high numbers of native species are vulnerable to invasions due to high resource availability and high environmental heterogeneity – a pattern that holds up generally well over large portions of the earth from local and regional scales (Stohlgren *et al*., 1999; 2003) to global scales (Lonsdale, 1999). Increasing trade and transportation diminishes dispersal limitations, and increases the number of introductions and species overlap among regions and nations (Pyšek *et al*., 2010). The number and overlap of alien taxa among countries or regions (Stohlgren *et al*., 2011) is a testament to trade and transportation. However, the paucity of widely distributed alien species in Europe remains an aberration. Very few of the 1600 plant species that were introduced to Europe since AD 1500 have become widespread (Lambdon *et al*., 2008). The approximately 11,000 native plant species seem to be holding their ground.

Climate, level of disturbance, time since introduction, and trade volume are often given as reasons for differences in invasion patterns. However, geographic areas within the Peoples Republic of China, North America, and portions of Europe include temperate zones with very different invasion patterns. All these areas are heavily affected by land use change. The early plant invaders in Europe (pre AD 1500) showed that as a group, they are equally distributed despite widely different times since introduction (Pyšek *et al*., 2002; Stohlgren *et al*., 2011). Increased volume of trade and transportation with Peoples Republic of China, a country very rich in species diversity, may create additional future challenges for other countries and regions. Europe, especially its Mediterranean region, has been traditionally considered a donor of invasive species to other parts of the world due to historical reasons and its long association of plants and animals with humans since the beginning of agriculture some 10,000 years ago. This is reflected by the majority alien species on other continents being of Eurasian origin (Pyšek, 1998).

Regional and national plant inventories and databases are essential to understand patterns of invasions. Many areas of the globe may lose their floristic uniqueness as plant species continue to invade. Trade and transportation and changing patterns of widely distributed species will most likely dominate future patterns of biodiversity. We may be witnessing accelerating effects of the great reshuffling of species on the globe as geographic isolation diminishes over time. Human-assisted migrations, and unintentional species establishment followed by natural dispersal by wind, water, and animals may further spread alien assemblages from local to global scales. While the scientific community is making great strides in tracking global invasive species, it might be prudent to begin thinking about the tracking of widely distributed native species as likely donor populations.

See also: Dispersal Biogeography. Economic Control of Invasive Species. Global Species Richness. Human Impact on Biodiversity, Overview. Introduced Species, Impacts and Distribution of. Loss of Biodiversity, Overview. Mediterranean-Climate Ecosystems. Plant Invasions. Species Coexistence. Species Distribution Modeling

References

Crosby AW (1972) *The Columbian Exchange: Biological and Cultural Consequences of 1492*. CT: Greenwood Publishing Group.

Flemming L (2004) *Excel HSC Business Studies*. Glebe, New South Wales, Australia: Pascal Press.

Grubben GJH and Denton OA (eds.) (2004) *Plant resources of tropical Africa 2. Vegetables*. Wageningen, Netherlands/Leiden: PROTA Foundation/Backhuys Publishers.

Holm GL, Plucknett DL, Pancho JV, and Herberger JP (1991) *The World's Worst Weeds. Distribution and Biology.* Malabar, FL: Krieger.

Jiménez A, Pauchard A, Cavieres LA, Marticorena A, and Bustamante RO (2008) Do climatically similar regions contain similar alien floras? A comparison between the Mediterranean areas of central Chile and California. *Journal of Biogeography* 35: 614–624.

Kessler D and Temin P (2007) The organization of the grain trade in the early Roman Empire. *Economic History Review* 60: 313–332.

Lambdon PW, Pyšek P, Basnou C, *et al.* (2008) Alien flora of Europe: Species diversity, temporal trends, geographical patterns and research needs. *Preslia* 80: 101–149.

Lonsdale WM (1999) Global patterns of plant invasions and the concept of invasibility. *Ecology* 80: 1522–1536.

Maki K and Galatowitsch S (2004) Movement of invasive aquatic plants into Minnesota (USA) through horticultural trade. *Biological Conservation* 118: 389–396.

Masifwa WF, Twongo T, and Denny P (2001) The impact of water hyacinth, Eichhornia crassipes (Mart) Solms on the abundance and diversity of aquatic macroinvertebrates along the shores of northern Lake Victoria, Uganda. *Hydrobiologia* 452: 79–88.

McKinney ML and Lockwood JL (1999) Biotic homogenization: A few winners replacing many losers in the next mass extinction. *Trends in Ecology and Evolution* 14: 450–453.

Pino J, Font X, de Cáceres M, and Molowny R (2009) Floristic homogenization by native ruderal and alien plants in north-east Spain: The effect of environmental differences on a regional scale. *Global Ecology and Biogeography* 18: 563–574.

Phillips ML, Murray BR, Pyšek P, *et al.* (2010) Plant species of the Central European flora as aliens in Australia. *Preslia* 82: 465–482.

Pyšek P (1998) Is there a taxonomic pattern to plant invasions? *Oikos* 82: 282–294.

Pyšek P, Cock MJW, Nentwig W, and Ravn HP (eds.) (2007) *Ecology and management of Giant Hogweed (Heracleum mantegazzianum).* Wallingford, UK: CAB International.

Pyšek P, Jarošík V, Hulme PE, *et al.* (2010) Disentangling the role of environmental and human pressures on biological invasions. *Proceedings of the National Academy of Sciences USA* 107: 12157–12162.

Pyšek P and Richardson DM (2007) Traits associated with invasiveness in alien plants: Where do we stand? In: Nentwig W (ed.) *Biological Invasions*, pp. 97–126. Berlin: Springer.

Pyšek P, Sádlo J, and Mandák B (2002) Catalogue of alien plants of the Czech Republic. *Preslia* 74: 97–186.

Reichard SH and White P (2001) Horticulture as a pathway of invasive plant introductions in the United States. *BioScience* 51: 103–113.

Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta FD, and West CJ (2000) Naturalization and invasion of alien plants: Concepts and definitions. *Diversity and Distributions* 6: 93–107.

Richardson DM, Rouget M, Ralston SJ, Cowling RM, Van Rensburg BJ, and Thuiller W (2005) Species richness of alien plants in South Africa: Environmental correlates and the relationship with indigenous plant species richness. *Ecoscience* 12: 391–402.

Sax DF, Gaines SD, and Brown JH (2002) Species invasions exceed extinctions on islands worldwide: A comparative study of plants and birds. *The American Naturalist* 160: 766–783.

Stohlgren TJ (2007) *Measuring Plant Diversity: Lessons from the Field.* New York, NY: Oxford University Press.

Stohlgren TJ, Barnett D, and Kartesz J (2003) The rich get richer: Patterns of plant invasions in the United States. *Frontiers in Ecology and the Environment* 1: 11–14.

Stohlgren TJ, Barnett DT, Jarnevich CS, Flather C, and Kartesz J (2008) The myth of plant species saturation. *Ecological Letters* 11: 313–326.

Stohlgren TJ, Binkley D, Chong GW, *et al.* (1999) Exotic plant species invade hot spots of native plant diversity. *Ecological Monographs* 69: 25–46.

Stohlgren TJ, Pyšek P, Kartesz J, *et al.* (2011) Widespread plant species: Natives vs. aliens in our changing world. *Biological Invasions.* doi: 10.1007/s10530-011-0024-9.

Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, and Richardson DM (2009) Something in the way you move: Dispersal pathways affect invasion success. *Trends in Ecology and Evolution* 24: 136–144.

Winter M, Kühn I, La Sorte FA, Schweiger O, Nentwig W, and Klotz S (2010) The role of non-native plants and vertebrates in defining patterns of compositional dissimilarity within and across continents. *Global Ecology and Biogeography* 19: 332–342.

Winter M, Schweiger O, Klotz S, *et al.* (2009) Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proceedings of the National Academy of Sciences USA* 106: 21721–21725.

Global Species Richness

Kevin J Gaston, University of Exeter, Exeter, UK, and University of Sheffield, Sheffield, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Homonymy The same scientific name applied to two or more species.

Synonymy Two or more scientific names applied to the same species.

Significance

Responses to attempts to answer the question of how many extant species there are on Earth take a variety of forms. Some regard these attempts as meaningless exercises, given the lack of resolution of fundamental issues of the definition and recognition of species. Others find them futile, given the extraordinary difficulties in providing an accurate answer. More positively, some suggest that the question provides an important challenge to the state of understanding of global biodiversity, and that the attempts to answer it provide a valuable test of that understanding. Others go further, and argue that the answer is a basic fact that we should know, on a par perhaps with estimating the global quantities of the elements, the rate of drift of the continents, or the masses of atomic particles. Finally, some maintain that sound knowledge of global species richness will facilitate better answers to other important questions, such as the extent to which the carrying capacity of diversity on Earth has been attained, the rate at which species are becoming globally extinct, and the scale of the task faced by species conservation.

Fundamental Issues

Determining the number of species in an area is one of a general, and often difficult, set of problems in which an estimate of S is desired, where a statistical population, finite or infinite, is partitioned into S classes (Bunge and Fitzpatrick, 1993). A number of basic assumptions have typically to be made in such circumstances. These almost invariably become more problematic as the area concerned becomes larger, and thus particularly significant when attempting to estimate global species numbers.

First, the S classes should be defined in a consistent fashion. Unfortunately, this constitutes a major problem when estimating global species richness. For some higher taxa, known species (a sample s of S) have arguably been recognized on a reasonably uniform basis, usually employing a morphological species concept and definition. However, for other higher taxa, this basis has changed both as knowledge of key characters discriminating between species and as available technologies have developed, and as different species concepts (e.g., biological, ecological, morphological, and phylogenetic) have been favored. Moreover, for known species in different higher taxa, the same basic concept has been applied in different ways, or different species concepts have predominantly been applied.

Second in two samples, s_i of S, both the number of classes in common and the number unique to each sample i should be estimable. Whereas for better-known higher taxa, determining the overlap in species composition is a relatively trivial exercise, for others it is not. This is because of the huge challenge provided by the geographic variation exhibited in the traits of many species. The use of molecular approaches to species identification is reducing this problem, although this of course introduces a basis for species identification that differs from that which has been applied historically.

Third, S is assumed to remain constant. However, current species extinction rates have been estimated, and forthcoming rates predicted, to exceed the background rates recorded in the fossil record by orders of magnitude. Thus, an estimation of S is a moving target. Given the extremely rough and ready nature of most attempts to estimate global species richness, arguably this is not yet a major problem. However, if estimates were to become sufficiently refined, it could become so.

Fourth, the size of a sample s of S is commonly assumed to constitute a reasonable proportion of S. However, although across all higher taxa of eukaryotes, known species may constitute perhaps a fifth of the global number of extant species, it is quite possibly considerably less. Indeed, for some individual higher taxa, this figure is very much smaller. A typical large areal sample may contain hundreds or thousands of species, and thus fractions of a percent of global species richness.

Finally, it is commonly assumed that it is possible to take a sample s that is randomly drawn from S. Given the diversity of habitats and species (and the different challenges they pose for sampling), this is in practice impossible at a global scale.

Methods of Estimation

Counting

On the face of it, the best way of finding out how many species (S) there are would simply be to count them. However, other issues aside, the diversity of life is so great that at a global scale this presents a truly formidable task, and one that has never risen sufficiently high on the agenda of humankind to be given serious consideration. Nonetheless, the question of how insurmountable the obstacle would be if substantial resources, technology, and ingenuity were brought to bear remains unanswered. Some believe that it is attainable in a matter of decades, but others are unconvinced. For a small number of species-poor higher taxa, reasonable estimates of global richness have indeed been obtained simply

by counting. However, in few cases, are these estimates likely to be definitive. Even for the better-known groups, new species continue to be discovered every year (e.g., birds and mammals), and the boundaries between species continue to be revised. Even when all the extant species in a group have been encountered, lists of described species are prone to two kinds of errors. First, the same species name may have been attributed to more than one species, so-called homonymy. Second, more than one species name may have been attributed to the same species, so-called synonymy. The balance of these two kinds of errors leans strongly toward synonymy, with many thousands of species names thought to be synonyms. Indeed, the actual level of synonymy has proven to be a key issue in, for example, attempts to determine the global species richness of plants, a group for which surprising uncertainty and disagreement exists around such estimates. Here, the highly variable synonymy rate among different groups of plants that have been well studied in this regard makes the assessment of the overall level of synonymy across all the known species difficult.

Canvassing Experts

Rather than take lists of described species at face value, estimates of overall species richness can potentially be improved based on the opinions of those taxonomic experts who have studied the relevant groups over long periods, and have gained an understanding of the frequency with which new species are encountered and described, of the levels of homonymy and synonymy in species lists that have been recognized but remain unresolved (their formal resolution is time consuming and is not presently a high priority in systematic studies of many groups), and of likely hidden levels of homonymy and synonymy (Gaston, 1991). Indeed, many informal estimates of the richness of particular higher taxa have effectively been generated in this fashion. This approach to estimation makes the entirely untestable assumption that the experts know the taxonomic groups they investigate sufficiently well to make reliable estimates. At the very least, it seems likely that such estimates will be of rather variable accuracy. They may also suffer from systematic biases. The more speciose and geographically widespread the taxon the greater these problems are likely to be.

Temporal Patterns of Species Description

Overall numbers of species in some groups have been estimated by extrapolating to infinite time the growth in the cumulative numbers of taxonomically described species through time ($S = f(s(t))$, where t is time). This assumes that temporal patterns of species description follow some relatively simple pattern (much debate surrounds the most appropriate model, although the pattern is commonly assumed to be approximately sigmoidal, with low early and late rates of description separated by a period of rapid increase), and that this pattern can be identified. Unfortunately, for the majority of higher taxa, relatively small proportions of species have been described, and insufficient information is available in the temporal patterns of description of those species to provide any indication as to what might be an appropriate model on which basis to predict the future history of descriptions. Moreover, temporal patterns of description often tend to be highly idiosyncratic. This is because they can be profoundly influenced by individual taxonomists (a high proportion of species in a group are formally described by a relatively small number of taxonomists), by changes in spatial patterns of collecting effort (which are shaped by opportunity as well as biological considerations), by changes in the features by which species are distinguished or the species concept employed, by the development or application of new technologies which facilitate or enable those changes, and by funding possibilities (description of species in some groups has at various times collapsed through lack of support). Indeed, these issues have meant that estimates of the number of species remaining to be discovered even in rather well-known taxonomic groups generated using temporal patterns of species description have sometimes quickly proven wildly inaccurate.

Proportions of Undescribed Species

An alternative approach to estimating global species richness is to determine the ratio of previously unknown to previously known species in large samples of specimens, and then to extrapolate from the overall numbers of known species (i.e., $S = S_K(s_t/s_k)$, where S is the total number of extant species, S_K the number of those that were previously known, s_t the total number of species in a sample, and s_k the number of species in the sample that were previously known). This assumes that the samples are representative. However, unfortunately, in practice samples tend to be much more serendipitous, and biased both in space and by higher taxon. Species from temperate regions tend to be much better sampled than are those from tropical and megadiverse regions, as are those that, on average, are larger bodied, more abundant (locally or regionally), more widely distributed, occupy a larger number of habitats or life zones, or are economically more important (e.g., of agricultural or medical significance). Nonetheless, arguably, this approach toward estimating overall global species richness or that of major higher taxa has not been sufficiently exploited, and coupled with a carefully designed sampling program could go far toward improving those estimates, if it were sufficiently well resourced.

Extrapolating From Well-Studied Areas

Overall, numbers of species globally or in very large regions have been estimated by extrapolating from those, typically rather small, areas for which numbers of species are reasonably well known, usually exploiting a species–area relationship ($S = f(s(A))$, where A is area). This assumes that the areas for which overall species numbers are better known are representative of those for which they are not. Of course this is not so, with the former being predominantly temperate and with well-developed infrastructures. The method also assumes that the form of the species–area relationship is sufficiently well established and that area is a sufficient predictor of the numbers of species present in a region. A major complication

is that over a wide span of variation in area the species–area relationship is complex, with a slope that varies between small areas and large ones, and commonly characterized as triphasic. Ways of estimating the slope at large areas from that at small ones have been explored, based on the change in similarity of species composition among small areas with increases in the distance by which they are separated. Extrapolation from well-studied areas is of limited value in estimating overall global species richness because for no area in the world has species composition of all groups been entirely inventoried.

Extrapolating From Well-Studied Groups

This involves estimating overall numbers of species based on the global numbers in well-known groups and estimates of the ratio of the numbers of species in these groups to others in those, typically few, regions where the latter are reasonably well known (i.e., $S_A = S_B(s_a/s_b)$, where S_A is the total number of extant species in the less well-known group, S_B the total number of extant species in the better-known group, s_a the number of species of the former group in the well-known region, and s_b of the second group). This assumes that these ratios of numbers of species in well-known and other groups remain reasonably constant across space. This is usually not the case. Thus, although this approach has played an important role in estimates of global species richness, those estimates have large associated uncertainties.

Species–Abundance Distributions

If reasonable estimates can be made of the global numbers of individuals of a higher taxon, and the shape of the species–abundance distribution (the frequency of species with different numbers of individuals) can be assumed ($S = f(s(N))$, where N is abundance), then the global numbers of species can be estimated. Such an approach has particularly been employed in attempts to estimate the species richness of microbes, for which the numbers of individuals are plainly enormous, and has provided useful insights into their diversity (Curtis *et al.*, 2002). A lognormal species–abundance distribution tends to be assumed, although it is well established that species–abundance distributions cannot be strictly lognormal.

Species–Body Size Distributions

In general, a much higher proportion of large-bodied species are known to science than small-bodied. If the overall shape of the species–body size distribution (the frequency of species of different body sizes) of all species, or a major component thereof, could be simply characterized ($S = f(s(M))$, where M is body size), then one could extrapolate from the known part of this distribution (that for the larger species) to the unknown part (that for the smaller species) to estimate total species numbers (May, 1988). The obvious difficulty here lies in knowing the shape of the regional or global species–body size function, and the extent to which this is approximated by that of species–body size distributions from smaller areas. Such limitations have largely thwarted effective use of this approach to date.

Trophic relations

If the numbers of species of a taxonomic group consuming a particular kind of resource can be determined, and their level of specialism to that resource and the number of kinds of resources are known, the number of species of the group can be estimated. Such an approach provided the starting point for Erwin's (1982) dramatic, and much cited, upward revision of the estimated global numbers of arthropod species. The debate that this engendered centered foremost on the assumed level of specialism, in this case the specificity of tropical beetle species to particular tree species, which can be hard to determine in systems with large numbers of resource types, but has a powerful influence on the resultant estimates. In this instance, subsequent detailed studies have generally served to suggest lower levels of specialism (or increased levels of generalism) than those originally assumed.

Current Estimates

More or less formally, various combinations of these approaches have been used to provide assessments of how many extant species there are globally in large taxonomic groups and, by summation of these estimates, how many extant species there are overall. Inevitably, given the previous remarks, some of these estimates are highly speculative, and all will be subjected to further (typically marked) revision. One recent summary for eukaryotes is provided in **Table 1**. The working figure for overall numbers is around 10 million species, an estimate with which many workers seem broadly comfortable.

The particular pattern of change in global species richness since the emergence of life 4–3.5 billion years ago remains contentious. However, it is plain that there has been a dramatic net increase. If more than 98% of species are now extinct, then based on estimates of numbers of extant species, and accepting that the error in such a calculation is likely to be substantial, 500 million eukaryote species may have ever lived.

The Greatest Uncertainties

The greatest uncertainties in estimates of global species richness lie in understanding the biodiversity associated with a relatively limited number of kinds of environments and of higher taxa.

Environments

Tropical Forest Canopies

Despite their large extent (>11 million km^2), tropical forest canopies long remained poorly explored. This has changed through the application of a variety of techniques that have much improved access to a previously largely inaccessible environment (e.g., rope-climbing techniques, aerial walkways,

Table 1 Estimates (in thousands), by different taxonomic groups, of the global numbers of extant eukaryote species that have been described and overall. Substantial uncertainty is associated with most of these figures (see text)

	Described species	*Overall species*	*Accuracy of overall species estimate*
Protoctista	29	1000	Very poor
Plants	310	391	Good
Fungi	99	1500	Moderate
Nematodes	25	500	Very poor
Arthropods	1176	5892	Moderate
Mollusks	85	200	Moderate
Chordates	65	81	Good
Others	101	364	Moderate
Total	1890	9928	

Source: Reproduced from Chapman AD (2009) *Numbers of Living Species in Australia and the World*, 2nd edn. Canberra: Australian Biological Resources Study (ABRS), and evaluations reproduced from May RM (2000) The dimensions of life on Earth. In: Raven PH and Williams T (eds.) *Nature and Human Society*, pp. 30–45. Washington, DC: National Academy Press.

cranes, and balloons). This has revealed much higher levels of richness of some groups than had been anticipated, and lower richness of others, but overall has served to confirm the significance of tropical forests for global species numbers.

Soils

Despite the much greater ease of access, understanding of the species richness of soils has been severely hampered by the small body size of many of the associated species, and the poor efficiency of many extraction methods. Detailed studies of soils at a single site in the Scottish borders generated high species richness estimates, particularly among small organisms, with 500–5000 bacteria, 365 protozoa, and 400–500 nematodes (Fitter *et al.*, 2005). Spatial turnover in species composition remains poorly explored, so the basis for estimations of regional or global species richness and patterns is largely lacking.

Deep Ocean

Areas at depths below 1 km in the oceans cover more than half of the planet, and the associated biodiversity remains one of the great imponderables. As with tropical forest canopies, gaining better access is important for improving this situation, as to date samples have been obtained from only a miniscule proportion of this extent. Thus, alongside developments in remotely operated vehicles, bottom landers, submarines, sonar, and video, for example, the identification of new kinds of communities of organisms, such as hydrothermal vents, cold seeps, and cold-water coral reefs, has occurred. Studies on the fauna of deep-sea floors in the Atlantic and Pacific uncovered a high level of species richness, leading to suggestions that there may be 10 million species in the deep sea (Grassle and Maciolek, 1992). Although this seems likely to be a marked overestimate, it is undoubtedly the case that large numbers remain to be discovered.

Coral Reefs

In contrast to tropical forests, soils, and the deep ocean, the global area covered by coral reefs is only about 600,000 km^2 (*ca*. 0.18% of total area of oceans). Nonetheless, it has been estimated that up to one quarter of all marine species and one-fifth of known marine fish species live in coral reef ecosystems, leading it to be seen as the marine equivalent of tropical forests.

Higher Taxa

Viruses

If one accepts that they exhibit at least some of the primary features common to all life forms, then the numbers of species of viruses are likely to constitute a significant issue in determining global species numbers. Most forms of life have some sort of susceptibility to viruses, they may exhibit significant levels of host specificity, and many types of viruses are known from some hosts. Moreover, previously unknown assemblages and types of virus are currently being found on a massive scale. However, what constitutes a species of virus, and the comparability of any species concept with those applied to other groups of organisms, are difficult issues. At present, it seems unlikely that these could be satisfactorily resolved, particularly given evidence that many viruses may be continually and randomly recombining with DNA that they encounter.

Bacteria

Basic understanding of the number of species of bacteria (and archaea) is complicated by frequent difficulties in applying standard species concepts to these creatures (resort is usually made to operational taxonomic units (OTUs)), by the difficulty of culturing the vast majority of them and thereby applying classical identification techniques, and by the unimaginable numbers of individuals that exist. The numbers of species estimated to occur in even very small areas can vary by several orders of magnitude, depending on the approach taken to estimation. Use of DNA reassociation techniques has revealed that in pristine soils and sediments with high organic content, samples of 30–100 cm^3 correspond to *ca*. 3000–11,000 different genomes, and may contain 10,000 different prokaryotic species of equivalent abundances (Torsvik *et al.*, 2002). Globally, the diversity of bacteria, both in terrestrial and marine systems, may be far larger than many had previously imagined, and has been argued to comprise possibly millions of species, and by some many orders of

magnitude more than that. If this is correct, then the vast majority of microbial diversity remains unexplored and undocumented.

Protozoa

A critical issue in understanding the extent of microbial diversity, whether it be bacterial or protist, is the extent to which individual microbial species are cosmopolitan in their distribution. In particular, it has been argued that driven by huge population sizes microbes are able to disperse widely and so occur wherever their habitat requirements are met, resulting in high local richness but relatively low global richness (Finlay, 2002). The evidence that this is so for protists, and thus how large is their global richness, is contentious. The debate rests heavily on how well morphological features (which have long been used in protist taxonomy) serve to discriminate among species, the interpretation of observed patterns of genetic variation, and the extent to which documented examples of apparently locally endemic species are products of undersampling.

Fungi

Somewhat in contrast to microbes, the high local species richness that has been documented at some sites for fungi has been argued to provide evidence for their high global species numbers. A working global estimate of 1.5 million species of fungi, based primarily on extrapolation from temperate studies, has been widely cited (Hawksworth, 2001). On the one hand, some tropical studies and some of endophytic species suggest that this may constitute a substantial underestimate. On the other hand, it has been argued that the frequency of discovery of previously unknown species in areas whose fungi are not well studied suggests that the figure may be a substantial overestimate.

Nematodes

The local densities of individual nematodes can be extraordinarily high, and they are extremely widespread. However, how this vast abundance translates into numbers of species remains unclear. Figures of 1–100 million extant species have been suggested. But recent analyses have cast severe doubt on the more extreme upper estimates (Lambshead and Boucher, 2003). For example, suggestions of immense numbers of species in the deep sea have been challenged by data from the equatorial central Pacific, which suggest richness values more akin to those of moderately speciose coastal offshore regions.

Arthropods

Estimates of the global number of arthropods have ranged particularly widely. Much of the uncertainty rests on the numbers of species that are to be found in the tropical rain forest canopy, the proportion that are restricted to this environment, and the degree of host specificity of herbivorous species in such forests (Novotny *et al.*, 2006; Hamilton *et al.*, 2010). Estimates have tended to fall as evidence has accrued that both restriction of species to the canopy and their host specificity tend to be lower than was previously assumed.

Among those higher taxa about which uncertainty in species numbers contributes most to uncertainty in overall global species richness, many, if not all, of their species are parasites, exploiting free-living organisms or other parasites for their resources. This has led to a lively debate as to whether parasitism is the most common lifestyle on Earth, and whether the majority of species are parasitic rather than free-living.

Prospect

If the past is anything to go by, long into the future prevailing estimates of global species richness will continue periodically to be revisited, as fresh data and improved methodologies become available. This is a healthy, albeit frustrating, comment on understanding of the magnitude of life on Earth.

See also: Corals and Coral Reefs. Loss of Biodiversity, Overview. Marine Ecosystems. Measurement and Analysis of Biodiversity. Microbial Biodiversity. Nomenclature, Systems of. Parasitism. Protozoa. Species, Concepts of

References

Bunge J and Fitzpatrick M (1993) Estimating the number of species: A review. *Journal of the American Statistical Association* 88: 364–373.

Chapman AD (2009) *Numbers of Living Species in Australia and the World*, 2nd edn. Canberra: Australian Biological Resources Study (ABRS).

Curtis TP, Sloan WT, and Scannell JW (2002) Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences USA* 99: 10494–10499.

Erwin TL (1982) Tropical forests: Their richness in Coleoptera and other arthropod species. *Coleopterists Bulletin* 36: 74–75.

Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061–1063.

Fitter AH, Gilligan CA, Hollingworth K, *et al.* (2005) Biodiversity and ecosystem function in soil. *Functional Ecology* 19: 369–377.

Gaston KJ (1991) The magnitude of global insect species richness. *Conservation Biology* 5: 283–296.

Grassle JF and Maciolek NJ (1992) Deep-sea species richness: Regional and local diversity estimates from quantitative bottom samples. *American Naturalist* 139: 313–341.

Hamilton AJ, Basset Y, Benke KK, *et al.* (2010) Quantifying uncertainty in estimation of tropical arthropod species richness. *American Naturalist* 176: 90–95.

Hawksworth DL (2001) The magnitude of fungal diversity: The 1.5 million species estimate revisited. *Mycological Research* 105: 1422–1432.

Lambshead PJD and Boucher G (2003) Marine nematode deep-sea biodiversity – hyperdiverse or hype? *Journal of Biogeography* 30: 475–485.

May RM (1988) How many species are there on Earth? *Science* 241: 1441–1449.

May RM (2000) The dimensions of life on Earth. In: Raven PH and Williams T (eds.) *Nature and Human Society*, pp. 30–45. Washington, DC: National Academy Press.

Novotny V, Drozd P, Miller SE, *et al.* (2006) Why are there so many species of herbivorous insects in tropical rainforests? *Science* 313: 1115–1118.

Torsvik V, Ovreas L, and Thingstad TF (2002) Prokaryotic diversity – magnitude, dynamics, and controlling factors. *Science* 296: 1064–1066.

Government Legislation and Regulations in the United States

Kathryn A Saterson, U.S. Environmental Protection Agency, NC, USA

Published by Elsevier Inc.

Glossary

Case law Interpretations of the US Constitution, statutes, and regulations provided by the judicial branches of US federal or state government. Such interpretations are provided when two or more parties disagree as to the meaning of a law in a specific context and bring it to the courts to decide.

Constitution A statement of the principles governing a nation or state.

Critical Habitat Habitat required to maintain a threatened or endangered species.

Executive order A directive from the US president, or a state governor, specifying actions by government officials and agencies. State and federal executive orders often describe how to administer a provision of a statute, treaty, or the constitution and usually have the force of law.

Precautionary principle An approach to regulations and treaties based on consideration of the fact that when there is a threat of serious or irreversible damage, lack of full scientific certainty about outcomes should not be used as a reason to postpone measures to avoid that environmental damage.

Regulations Rules and administrative codes required by statutes and issued by local, state, and federal government agencies. Regulations have the force of law since they are adopted under the authority granted by statutes.

Statutes or legislation Laws passed by the legislative branch of US federal or state governments. The US Congress passes federal laws and state legislatures pass state laws. Local laws are usually called municipal ordinances.

Treaties Multilateral agreements between or among nations regarding the actions of each national government. Only rarely do treaties address the actions of private institutions. International treaties, often called "conventions," are signed (entered into) in the US by the President but must be ratified by the Senate for the US to be an official party to a treaty. Treaties supersede federal, state, and local laws that might have contradictory goals.

Introduction

For centuries, customary or traditional laws have governed the use of biological resources throughout the world. The systems of customary law in traditional cultures often dictate limits on how much of a specific biological resource can be used and in what seasons. In many societies, an individual's rights to biodiversity are not always linked to land tenure. As traditional societies are becoming more integrated into national and global markets and political systems, traditional knowledge and laws are being lost. As human population levels increase in all societies, voluntary actions governed by cultural practices are frequently insufficient to protect biodiversity, leading to an increasing reliance on the creation and enforcement of government legislation and regulation to protect biodiversity.

A variety of motivations have led to creation of legislation intended to help conserve biodiversity. Early legislation was primarily aimed at protecting biodiversity as an economic resource for humans (e.g., for hunting or agriculture). Recent laws, and amendments to earlier laws, have included other reasons, such as the value of biological diversity to overall ecosystem health and the intrinsic value of species. Although US national parks were originally created primarily for scenic and recreational values, the 1916 National Park Act identified conservation of wildlife as an important goal of parks.

Laws and regulations to protect biological diversity can aim to influence actions undertaken by a range of actors (government agencies, private corporations, and individual citizens) and to affect a range of resource types (animals, plants, water, land, and specific habitats and ecosystems) with differing ownership and resource control (public and private). The success and evolution of legislation and regulation in the US are in part a reflection of the actions of nongovernmental organizations, corporations, and others who lobby for (and against) new or amended legislation, oversee how agencies implement legislation, and use the judicial system to help ensure that laws are enforced and interpreted appropriately.

All laws and regulations in the US are based on rights found in the US Constitution. The federal government derives its authority to regulate wildlife from its constitutional authority to regulate commerce, to protect its own property (federal lands) and to make treaties. When federal laws conflict with state laws, the federal laws generally take precedence.

Many advocates of the right to private property, also supported by the US Constitution, view legislation to protect biodiversity as sometimes conflicting with their right to private property. The Fifth Amendment to the Constitution requires that when private property is taken for public use (whether to benefit public safety, create roads, protect wildlife, or for any other purpose) there must be "just compensation." Most state and federal case laws indicate that restrictions on killing wildlife are not a sufficient taking of private property to require compensation by the government (Bean and Rowland, 1997).

Despite the large number of environmental laws and regulations that now exist, and the many positive impacts that they have had, these positive steps have been insufficient to prevent the loss of biodiversity. Some argue that federal and state governments give greater weight to protecting current economic growth and private property rights than to protecting biodiversity because of a lack of constitutional authority to protect future generations. Some conservationists

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00186-6

advocate an amendment to the US Constitution that obligates the government to protect biodiversity for future generations. Many other nations, and a number of states in the US require conservation of nature in their constitutions.

One of the controversial issues confronting the agencies and courts in implementing all the legislation described in this article is determining how the laws apply to American Indian reservations, which are covered under separate treaties, and to traditional subsistence use of biodiversity by Alaskan natives. Balancing the needs of individuals who depend on biological resources for food, shelter, and other benefits with the need to protect biodiversity for current and future societies is an issue with both legal and ethical challenges.

This article will summarize only legislation and regulations that aim directly to conserve biodiversity. **Table 1** presents a summary of selected laws with provisions that are intended to help conserve biodiversity. There are many additional environmental and other laws that can have more indirect positive impacts on biodiversity, for example, controlling air pollution (Clean Air Act), water pollution (Clean Water Act), hazardous waste (Resource Conservation and Recovery Act), global climate (Global Climate Change Convention), trade (General Agreement on Tariffs and Trade) or creating conservation incentives through tax policies (Internal Revenue Code). This article will also not address the specific federal, state, and local programs (often resulting from legislation) that can help to conserve biodiversity.

International Treaties

International agreements are important for conserving biodiversity because many species have very large natural ranges or migrate between countries, most habitats and ecosystems

Table 1 Selected legislation intended to help protect biological diversity

Law (year)[a]	*Purpose*
Multilateral International Agreements	
International Convention for the Regulation of Whaling (1946)	To conserve whales for the whaling industry. Created the International Whaling Commission which designates protected whale species and hunting limits
Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES, 1973)	Prohibits and regulates international trade of specified plants and animals
Convention on Biological Diversity (CBD, 1993)	Encourages nations to conserve, use, and share the benefits from conservation of genetic, species, and ecosystem levels of biodiversity. As of 2011, the US had signed but not ratified the CBD
Convention on wetlands of international importance (Ramsar Convention 1971)	Encourages nations to designate wetlands of international importance and to conserve and wisely use wetlands
International Undertaking on Plant genetic Resources (2000)	To ensure the conservation of plant genetic resources used for food and agriculture and the equitable sharing of benefits
U.S. Federal Laws	
Lacey Act (1900)	Prohibits commerce in wildlife injurious to agriculture and in wild animals and some plants taken in violation of US and foreign laws
National Park Service Act (1916)	Requires conservation of scenery, natural and historic objects, and wildlife in national parks and other protected areas for public benefit
Migratory Bird Treaty Act (1918)	Implements treaties with Canada, Mexico, Japan, and the Soviet Union to establish hunting seasons and protect habitat for migratory birds
Bald Eagle Protection Act (1940)	Prohibits killing or possessing bald and golden eagles, their nests, or parts
US Agricultural Research and Marketing Act (1946)	Requires collecting, preserving, and disseminating genetic material important to US agriculture
Foreign Assistance Act (1962)	Requires that US foreign aid to developing nations helps to protect biodiversity and tropical forests, and that development projects generally avoid adverse impacts on biodiversity
National Environmental Policy Act (1969) (signed by the president in January 1970)	Requires Environmental Impact Statements for federal programmatic and site-specific actions
Marine Mammal Protection Act (1972)	Creates a moratorium on importing marine mammals and limits taking of US marine mammals
Endangered Species Act (1973)	Prohibits or restricts taking of endangered or threatened species, requires federal agencies to protect their critical habitat, and regulates their trade; requires evaluation of the impact of federal projects on endangered and threatened species
Fishery Conservation and Management Act (1976)	Requires plans from state/federal councils to manage fisheries in a 200 mile offshore exclusive economic zone to ensure a maximum sustainable yield
National Forest Management Act (1976)	Requires management plans for all national forests, including conservation of a diversity of plant and animal communities
Wild Bird Conservation Act (1992)	Limits imports of exotic birds to protect wild populations in country of origin and reduce inhumane treatment of birds
National Wildlife Refuge System Improvement Act (1997)	Requires that National Wildlife Refuge System conserve fish, wildlife, plants, and their habitat while maintaining the biological integrity, diversity, and environmental health of the system

[a]This is the year legislation was first passed. All laws shown have been amended subsequent to passage. The stated purpose may include objectives added in the subsequent amendments or may not include all the purposes of the law.

do not follow national boundaries, and the earth is increasingly globalized in trade and movements of people. Under international treaties, the ratifying countries agree to take specific actions at the national level. Although the earliest international treaties reflected concerns that one nation might overexploit a species that many nations found economically important, recent treaties have acknowledged the importance of conserving all levels of biodiversity, whether or not there is an immediate economic value. Ratification of a treaty by each country is voluntary, and enforcement is often difficult. The US is signatory to many international treaties relating to biodiversity, some dating back to the early 1900s.

International Convention for the Regulation of Whaling (1946)

The major nations then involved in whaling (including the US, Japan, and the Soviet Union) were party to the 1946 Convention to Regulate Whaling, which prohibits killing calves, immature whales, or females accompanied by calves and encourages using as much of the killed whales as possible. The convention created the International Whaling Commission and gave it the authority to designate protected species and to establish hunting methods and seasons, size limits, and take limits. The convention had the conflicting goals of both conserving whale stocks and further developing the whaling industry. Most species of whales continued to decline dramatically until a 10-year moratorium on commercial whaling was agreed to by signatory countries in 1982 and went into effect in1986. By that time, only Japan, the Soviet Union, and Norway were still whaling. The moratorium is still in effect, although it is not adhered to by Norway, and many countries continue to take whales for the "research purposes" allowed for under the convention. Iceland withdrew from the convention in 1992 (Nagle and Ruhl, 2006). The International Whaling Commission, established by the Convention, sets the aboriginal subsistence catch limits for aboriginal communities in the US, Canada, Greenland, and St.Vincent and the Grenadines.

The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) (1973)

The Convention on International Trade in Endangered Species (CITES), which was introduced in 1973 and entered into force in 1975, prohibits international trade in endangered species by assigning each protected species to one of three lists. Appendix I lists "all species threatened with extinction which are or may be affected by trade." They cannot be traded primarily for commercial purposes, and trade in the species cannot be detrimental to the species survival. Trade in Appendix I species requires both import and export permits approved by the "management authority and scientific authority" of the nations involved (the secretary of interior in the case of the US). The permit has to certify that the specimen was obtained legally, it can be shipped without risk of cruel treatment or harming the health of the specimen, and the trade is not detrimental to survival of the species.

Species listed in Appendix II are not necessarily currently threatened, but unregulated trade could jeopardize their survival. Two-thirds of the parties to CITES must agree in order to add or remove a species from Appendix I and II lists. Any nation can unilaterally add or remove a species from Appendix III if it believes that cooperation from other nations is needed to control trade in a vulnerable species under their jurisdiction. Only export permits are required for Appendix II and III species. The precautionary principle was explicitly endorsed at the Ninth Conference of the parties to CITES in 1994 (Dickson, 1999). By 2011, more than 175 countries had ratified the CITES treaty.

Convention on Biological Diversity (1993)

The International Union for the Conservation of Nature began encouraging the creation of a global biodiversity treaty in the early 1980s. The United Nations Environment Program led the process of creating the convention, which was agreed to in 1992 and entered into force in 1993. By 2011, 193 countries had ratified the treaty. The US signed the Convention on Biological Diversity (CBD) in 1993, but it has not yet been ratified by two-thirds of the Senate. Until ratification, federal executive agencies are expected to conform with the treaty to the extent possible, but it is not recognized as law in federal and state courts.

The CBD outlines objectives for the conservation, sustainable use, and equitable sharing of the benefits from biodiversity. Signatory countries are called on to develop national strategies for protection of biodiversity, establish a system of protected areas, conduct environmental assessments of projects in order to prevent adverse impacts on biodiversity, enact laws regulating conditions of access to genetic resources, and support a financial mechanism for developing countries to obtain grants to assist in meeting the terms of the treaty. The convention supports conservation of agrobiodiversity (the plants and animals that contribute to food security for humans) as well as wild biodiversity. The importance of habitat conservation and of controlling alien species is acknowledged in the treaty.

Some of the most discussed provisions of the treaty (and the basis for the US reluctance to ratify it as of 1999) are those concerning access to genetic resources and the "intellectual property rights" resulting from commercial use of genetic resources (Bass, 1999). Countries with large biotechnology industries (e.g., pharmaceutical and agribusiness) supported continued open access to genetic resources (often obtained from flora in other nations), whereas countries supplying biodiversity to the industry wanted to share the profits to help ensure continued protection of biodiversity. The treaty both acknowledges each nation's sovereign control of its biodiversity and encourages access to genetic resources by other contracting parties.

The 2010 meeting of the 10th Conference of the Parties (COP) for the CBD supported the establishment of the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES), which will be coordinated by the United Nations Environment Program. The IPBES is a mechanism proposed to further strengthen the science-policy interface on

biodiversity and ecosystem services. By evaluating the scientific consequences of real policy options for biodiversity, the IPBES will contribute to improved legislation and regulation to protect biodiversity (Perrings *et al.*, 2011).

National Legislation

Most national laws discussed here are primarily concerned with actions influencing biodiversity in the US. A few, including the Foreign Assistance Act, legislate how actions of the US government relate to conservation in other countries. The US states were solely responsible for conserving wildlife until states urged the federal government to pass the Lacey Act in 1900. States still retain primary responsibility for biodiversity. The role of the federal government is limited to managing federal lands; conserving economically important genetic resources; protecting migratory waterfowl, marine species, birds of prey, and endangered and threatened species; and enforcing international treaties. The following discussion describes US laws that support conservation at the genetic, species, and ecosystem levels according to their primary purpose. Such divisions are arbitrary to a degree since genes are best conserved in the organisms in which they exist, and species depend on their habitats for survival.

Conservation of Genetic Resources

Federal laws regarding genetic resources are primarily concerned with conservation of varieties of economically important domesticated plants and animals, with some attention to "wild genetic resources," for example, wild relatives of plants or animals known to be economically important. For wild species, legislation is more concerned with access to genetic resources and sharing of benefits if those resources end up being economically valuable. In 1990, an amendment to the 1946 US Agricultural Research and Marketing Act modified the National Genetic Resources Program in the Agricultural Research Service in the Department of Agriculture. The purpose of this program is to collect, preserve, and disseminate genetic material that is important to American agricultural production. This amendment expanded the types of genetic resources in the program to include silvicultural species, animals, and aquatic and microbiological organisms. Congress authorized creation of the National Seed Storage Laboratory in Colorado by the US Department of Agriculture. Although this gene bank is world renowned for the variety of agricultural species it contains, it has been criticized for inadequately maintaining seed from wild relatives of important crops.

Conservation of Species

The earliest federal wildlife legislation in the US was concerned with species, both migratory and nonmigratory, with ranges that went beyond state boundaries. Most early conservation legislation emphasized regulation of consumptive uses of species, such as hunting and fishing, to protect charismatic species such as eagles or bison. By the 1970s, Congress was aware that habitat degradation is a significant threat to biodiversity; this is reflected in the habitat conservation provisions of the Endangered Species Act and in the increase in laws concerning habitats described in the Section Conservation of Habitats and Ecosystems.

Lacey Act of 1900

The Lacey Act was created to strengthen the ability of states to protect wildlife by regulating interstate commerce in game birds and wild birds. The Lacey Act now also prohibits importing all wildlife and some plants that were taken contrary to the laws of another country. The original act also aimed to protect US agricultural and horticultural interests by prohibiting the import of injurious foreign wildlife, including mongooses, English sparrows, and other birds and animals. Later amendments broadened the definition of injurious wildlife to include wild birds, wild mammals, fish, amphibians, and reptiles and their eggs, and the protected interests were also expanded to include humans, forestry, and US wildlife resources. The 1981 amendments also increased the maximum penalties and lowered the standard of proof for a violation, making the Lacey Act one of the laws most frequently used by federal law enforcement officials to protect wildlife (Bean and Rowland, 1997).

Laws Controlling the Introduction of Nonnative Species

One of the major threats to biodiversity is the introduction of nonnative species to new habitats, intentionally or by accident. The federal authority to regulate interstate commerce led to creation of the Federal Noxious Weed Act of 1974 and the Nonindigenous Aquatic Nuisance Prevention and Control Act of 1990. The Noxious Weed Act gives the secretary of agriculture authority to regulate both foreign imports and interstate commerce in potentially harmful plants, defined as plants of foreign origin that can directly or indirectly injure agriculture, natural fish, and wildlife resources, or public health. Amendments in 1990 gave authority to the Fish and Wildlife Service, the National Park Service, and the Bureau of Land Management to control undesirable plants on their lands.

The accidental introduction of the economically destructive zebra mussel into US waters when a ship's ballast water was released led Congress to create the Nonindigenous Aquatic Nuisance Prevention and Control Act in 1990. This act covers any plant, animal, or other "viable biological material," such as a virus, that disperses to an aquatic ecosystem in which it is not historically found. In contrast to the Lacey Act provisions, a nonindigenous species does not have to be from a foreign country. To be termed a nuisance, a nonindigenous species must threaten the abundance or diversity of native species or the ecological stability or commercial productivity of the infested waters. The act mandates creation of a task force to implement a program to prevent the introduction and dispersal of aquatic nuisance species, but it does not specify who determines which species are a nuisance (Bean and Rowland, 1997).

Birds

Migratory Bird Treaty Act of 1918

This act was passed to implement a treaty signed in 1916 between the US and Great Britain with the aim of conserving

birds that migrate between the US and Canada. The treaty establishes closed seasons for hunting and prohibits taking of nests or eggs (except for scientific purposes) for three groups of migratory birds (game, insectivorous, and nongame – a classification that includes almost all birds). The act also implements treaties signed with Mexico in 1936, Japan in 1972, and the Soviet Union in 1976. The latter two treaties include provisions for protecting bird habitat, for example, from pollution. Numerous cases have resulted in judicial interpretations that "taking" under the treaty act does not include habitat destruction. The act was successful in supporting the 1991 requirement that migratory waterfowl not be hunted with lead shot. The courts have rejected landowner claims that the hunting prohibitions under this act are unlawful "takings" of their private property (Bean and Rowland, 1997).

Bald Eagle Protection Act of 1940

The Bald Eagle Protection Act prohibits killing or possessing bald eagles, their nests, or any part of the eagle. The secretary of interior could approve exceptions for scientific and educational purposes. A 1962 amendment added the same protections for golden eagles and authorized the secretary to permit taking of bald and golden eagles by Indian tribes only for religious purposes. A 1972 amendment of the act expanded the definition of prohibited taking to include poisoning.

Wild Bird Conservation Act of 1992

The US is the largest importer of wild birds in the world. It is estimated that as many as 50% of captured exotic wild birds die even before they leave the country of origin, and another 15% die in transit to the US (Bean and Rowland, 1997). The goal of the Wild Bird Conservation Act is to promote conservation of birds not indigenous to the US by (1) limiting or prohibiting imports of nonindigenous birds in order to reduce depletion of wild populations and reduce inhumane treatment of birds in transit, (2) ensuring that trade is biologically sustainable, and (3) assisting in wild bird conservation programs in the country of origin. Congress found that CITES was not effective in decreasing the rate of loss of wild bird populations because many signatory exporting countries were unable to implement it adequately. The act allows the secretary of interior to declare a moratorium on importing any exotic bird listed on any CITES appendix and to permit exceptions for scientific purposes.

Marine Mammal Protection Act of 1972

The Marine Mammal Protection Act (MMPA) replaced individual state authorities and programs concerning marine mammals and created a moratorium on importing marine mammals to the US and on taking of US marine mammals. Taking was defined to include attempts to harass, hunt, capture, kill, and (after later amendments) feed marine mammals in the wild. Responsibility for implementing the act rests with the secretary of commerce for (1) incidental take of mammals during commercial fishing, (2) all Cetacea (whales and porpoises), and (3) all Pinnipedia (seals) except for walruses. The secretary of interior implements the MMPA for all other marine mammals, including manatees, polar bears, sea otters, and walruses. Like most environmental legislation, the MMPA reflects a compromise between many interest groups: Commercial interests wanting to protect an economic product, scientists who valued marine species and their role in marine ecosystems, and animal welfare groups wanting to protect marine mammals for their individual and intrinsic value.

The MMPA outlines management principles that include maintaining the health and stability of marine ecosystems and the "optimum sustainable population" for all marine mammals as well as the "maximum sustainable yield" for commercially exploited species. The taking prohibitions allow many exceptions; one that received the most attention was the controversial allowance for "incidental taking" during commercial fishing operations. At the time the act was passed, more that 5 million dolphins had died accidentally by drowning in purse-seining nets used to catch tuna (Bean and Rowland, 1997). Beginning in 1976, a series of regulations set quotas on how many dolphins could be killed incidentally by US tuna fishing boats. In 1984, amendments sought to ensure that countries exporting tuna to the US had kept incidental dolphin kills as low as those by the US fleet; otherwise, a moratorium could be imposed on importing tuna from that country.

The 1992 International Dolphin Conservation Act required no import of any tuna caught in association with incidental dolphin mortality to the US after 1994(Bean and Rowland, 1997). This and similar laws have been challenged by other nations for inhibiting the free trade agreed to under the General Agreement on Tariffs and Trade.

Endangered Species Act of 1973

The 1973 Endangered Species Act (ESA) is one of the strongest, and most controversial, federal biodiversity conservation laws. ESA has a broad mandate to restrict the taking of species that are at risk of becoming extinct, protect and acquire the habitat necessary for those species to survive, regulate their trade, and force all federal agencies to evaluate the impacts of their activities on endangered species and to avoid jeopardizing a species' continued existence. This act was intended to enable the recovery of the populations of species it protects and the habitats that those species depend on. An endangered species is defined by the act as one that is in danger of extinction throughout all or a significant portion of its range. A threatened species is considered likely to become endangered in the foreseeable future throughout all or a significant portion of its range. The ESA is administered by the Interior Department's US Fish and Wildlife Service (FWS) and the Commerce Department's National Marine Fisheries Services (NMFS). The FWS is responsible for terrestrial and freshwater organisms and the NMFS is responsible for marine wildlife and anadromous fish such as salmon. The act outlines detailed procedures for the Secretary of Interior or Commerce to follow in listing or delisting a species as threatened or endangered (Houck, 1993). The decision must be based on objective scientific information on the biological status of the species and not on its economic or other value. Until a species is listed, it does not receive most of the protections of the act and the Department of Interior has a large backlog of species in the listing process.

The ESA outlines many duties and prohibitions that apply to all people subject to US laws (in Section 9) and other duties

that apply only to federal agencies (in Section 7). It is illegal to take any endangered or threatened animal species; take means to harm, harass, pursue, hunt, shoot, wound, kill, trap, or capture. Regulations further define harm to include habitat degradation to the point that wildlife is injured or killed. Although the original act did not prohibit taking of plants, later amendments prohibited taking plants on federal lands and prohibited removing or destroying plants on nonfederal lands where such action violated other laws.

Federal agencies are prohibited by Section 7 from approving, funding, or directly undertaking actions that are likely to jeopardize the continued existence of an endangered or threatened species or result in the destruction or adverse modification of its designated critical habitat. These agencies must consult with the Department of Interior or Commerce to determine the potential for jeopardy to the species or harm to its critical habitat and reasonable ways to avoid such harm. There is no direct legal obligation under ESA for nonfederal agencies or persons to avoid adverse modification of critical habitat.

In 1978, the Supreme Court ruled that the Tellico Dam in Tennessee could not be completed in order to prevent extinction of the endangered snail darter fish since the dam's reservoir would inundate the snail darter's habitat. Soon after this decision, Congress amended the ESA to allow the USFWS to take economic considerations into account when designating critical habitat and to allow a process for seeking exemption from the act's restriction on federal agencies. In 1982, Congress again amended the ESA to add section 10(A), allowing the Department of Interior to permit "incidental" private takings of endangered or threatened species if it was part of an approved Habitat Conservation Plan that would help reduce the risk to the species overall.

Section 7 also requires designation of critical habitat for species listed as endangered or threatened. Critical habitat is defined as an area critical to the survival of the species; this could include a portion of the area occupied by a listed species or an area beyond currently occupied habitat. Although the requirement to protect critical habitat has been applauded by many, it has been difficult to implement, and designations of critical habitat by the Department of Interior or Commerce have been rare. The requirement has resulted in a great deal of litigation brought both by those challenging critical habitat designations and by those challenging the agencies for their failure to make them.

The ESA can be enforced by the government through civil and criminal penalties and by citizens through citizen suits provided for in the act. Those who knowingly violate the ESA for an endangered species can be fined up to $50,000 and put in prison for up to 1 year. For a threatened species, a violation can result in a $25,000 fine and imprisonment for 6 months. Additional civil penalties can be imposed. States can participate in enforcement of ESA by signing cooperative agreements with the secretary of interior and are also eligible for federal financial assistance for enforcement.

An international mandate is also included in the ESA. The act directs the secretary of interior to encourage other countries to protect endangered species and authorizes financial support and assistance from US wildlife officers for such programs. Implementation of CITES is also authorized by ESA.

As of March 2011, the FWS had listed 1967 species worldwide as endangered or threatened; 1372 of those species occur in the US. The goal of the ESA is to help a species recover a large enough population to ensure that it is no longer at risk of extinction and can be taken off the threatened and endangered list. As of 2005, 40 species had been delisted but just 17 of those had recovered; the remainder were delisted for reasons including taxonomic changes and extinction (Nagle and Ruhl, 2006).

The ESA is one of the most controversial federal environmental laws passed during the past century. Its' very specific, substantive, and legally binding provisions make it a target for continuous challenge from a variety of perspectives. Challenges have come from those who argue about whether ESA really helps protect species, inhibits use of private land, or applies to specific federal and private actions. The act forces attention to solving the problem of threatened and endangered species, with its wider attention to habitat conservation. This focus on solutions has increased discussion of how humans develop and use broad landscapes for economic production in areas with important biodiversity. More procedurally oriented laws are able to avoid controversy by not requiring the threats to biodiversity to be addressed. At the same time, those who believe that the irreversible nature of biodiversity loss warrants strong use of the precautionary principle argue that the ESA is too weakly precautionary because it requires so much scientific evidence to support listing and when species are finally listed as threatened or endangered it is often too late for populations to recover (Saterson, 2011).

Fishery Conservation and Management Act (Magnuson Act) of 1976

Beginning in the early 1900s, international agreements between nations fishing in the same waters attempted to establish control of exploitation of ocean fish; by the 1970s the US was a party to approximately 20 such agreements (Bean and Rowland, 1997). Although waiting for agreement on a single international convention on fishing, the US passed the Fishery Conservation and Management Act in 1976. The act established a 3-mile-wide territorial sea off the coast of the US and a 200-mile-wide exclusive economic zone (EEZ). Within the EEZ, the US claimed exclusive management authority over fish and all other forms of marine animal and plant life except birds, mammals, and tuna. Tuna was added to this authority in 1990 amendments. The US also claimed authority to manage certain sedentary species found in places where the continental shelf extends beyond 200 miles from shore.

The act required eight Regional Fishery Management Councils, which include state and federal officials, to develop comprehensive plans for conserving and managing the fisheries off their coasts. These plans have to comply with standards that include preventing over-fishing and ensuring a maximum sustainable yield. The 1996 reauthorization of the act created new standards for minimizing "bycatch" (the accidental harvest of fish species not intended for harvest) and required that management plans try to minimize adverse impacts on fish habitats. Implementation authority for the act lies with the National Marine Fisheries Service in the Commerce Department. An important and controversial issue is

the extent to which international trade agreements, such as the General Agreement on Tariffs and Trade or the North American Free Trade Agreement, have the potential to conflict with, and undermine, national laws aimed at protecting biodiversity.

Conservation of Habitats and Ecosystems

During the past 40 years, legislators, administrators, and the general public have come to understand that species are not effectively conserved without protecting their habitats and ecosystems. The ESA reflected this by requiring conservation of critical habitat. Implementation of ESA has increasingly focused on habitat conservation plans as a way to conserve species. Many federal agencies have developed ecosystem management approaches to conserving lands and ecosystems in their entirety in order to complement species-oriented laws. Although an ecosystem management act was introduced in Congress and some have proposed an Endangered Ecosystem Act, a single regulatory approach to conserving US ecosystems is not legislated. The European Union's 1979 Wild Bird's Directive, supplemented by the 1992 Habitat Directive requires EU member states to protect specific habitats and all bird species before there is evidence of population decline, creating one of the first applications of the precautionary principle to protected areas (Saterson, 2011).

One important and threatened habitat type, wetlands, receives protection through the US Clean Water Act of 1972. The dredge and fill permit program in Section 404 of the act is a significant force in protecting aquatic ecosystems by regulating the physical alteration of "waters of the US," including wetlands, estuaries, and streams. In 2001, the US Supreme court ruled that this authority extends to wetlands that are adjacent to or have an effect on navigable waters of the US, but does not extend to isolated wetlands.

Federal Lands

Legislation and regulations addressing conservation of habitats and ecosystems in most countries focus primarily on creation of protected areas on publicly owned lands. On average, less than 5% of a country becomes legally designated as protected, and enforcement of that limited area of protection is often weak. Many protected areas allow some consumptive use of plant and animal species. An increasing number of countries have legislation that requires special permits to clear native vegetation or special habitat types outside of protected areas.

In the US, approximately 30% of the land area (more than 700 million acres) is owned by the federal government and managed primarily by four agencies. Congress has enacted separate laws outlining how each agency should manage its lands. Historically, public land management agencies have given high priority to production of a commodity, such as timber, grazing land, or waterfowl production, and put less emphasis on maintaining the health of an ecosystem and its biodiversity. The 1964 Wilderness Act authorized Congress to designate portions of any federal land as wilderness areas, in which new commercial enterprises and permanent roads are prohibited and natural ecological processes are supposed to be allowed to occur without interference. During the early 1990s, federal agencies were directed to begin using "ecosystem management" to administer their lands. A level of production of a particular biological resource is to be considered in light of the health of the entire ecosystem.

Only the National Wildlife Refuge System (created in 1966) and the National Marine Sanctuary System (created in 1972) have as their primary goals the conservation of wildlife and marine resources, respectively. The National Marine Sanctuary Act of 1972 authorizes the secretary of commerce to designate sanctuary areas to "maintain, restore, and enhance living resources" by providing critical habitat for endangered species and areas where ecosystem structure and function are maintained. The National Wildlife Refuge System Improvement Act of 1997 consolidated refuge-related authorities and stated the mission of the National Wildlife Refuge System. The system's mission is to administer a national network of land and water "for the conservation, management, and restoration of fish, wildlife, and plant resources and their habitats" and to maintain the system's biological integrity and diversity.

National Park Service Act (1916)

The National Park Service Act outlines the goals and administrative guidelines for conserving the scenery, natural and historic objects, and wildlife in the national park system (which includes parks, monuments, preserves, wild rivers, and lakeshores) in a manner that does not impair those resources for future generations. There have been many instances in which the National Park Service has had to balance the (occasionally) conflicting goals of wildlife management and human recreation. Amendments to the act in 1978 provided the secretary of interior with some authority to protect park resources from outside threats, such as logging on private lands adjacent to parks. Courts have also upheld National Park Service authority to prohibit or regulate fishing, trapping, and hunting within park lands.

National Forest Management Act of 1976

National forest lands in the US began to be set aside in the late 1800s, with the primary purpose of protecting water supplies and timber production. The 1960 Multiple-Use Sustained-Yield Act (MUSY) expanded the purpose of national forests to also include outdoor recreation, grazing, timber, watersheds, fish, and wildlife. Concerned that the US Forest Service was giving timber production more attention than other multiple uses, Congress passed the 1976 National Forest Management Act (NFMA) which required the creation of management plans for each national forest. The plans are required to integrate physical, biological, and economic issues in order to provide for a diversity of plant and animal communities, including viable populations of native vertebrates. The plans are supposed to outline how timber harvests will mitigate impacts on biological diversity.

One of the most famous ecosystem management plans resulting from this legislation was for the old-growth forest in the Pacific Northwest. The Forest Service, after several court cases, was forced to develop a plan for old-growth forest that managed for a diversity of plant and animal species, including the spotted owl. Once the spotted owl was listed under the

ESA, the management plan had to comply with both ESA and NMFA.

Private Land

Individuals and institutions privately own approximately 60% of the US land area. Many of the previously discussed laws, such as the ESA and Section 404 of the Clean Water Act, aim to regulate actions with adverse impacts on biodiversity on private lands. Other federal laws aim to create positive financial incentives for conservation on private lands.

The 1985 Food Security Act included creation of a Conservation Reserve Program with the goal of encouraging farmers to voluntarily remove marginal lands from intensive agricultural production and plant them in perennial grass or tree cover for at least 10 years. A farmer enters into a contract to keep land in the reserve program for 10–15 years in exchange for cash or commodities. In addition, the farmer might receive up to 50% of the cost of certain conservation practices, such as creating wildlife corridors.

Biodiversity on private lands can also be conserved through donation or purchase of conservation easements. An easement is an interest in land that restricts the owner's use of the property in some way. The easement holder, usually a private interest or a government, has the right to enforce the restriction in perpetuity or for an agreed on time period. Since the 1970s, the Fish and Wildlife Service has been negotiating easements on wetlands as waterfowl production areas. The 1990 Food, Agriculture, Conservation, and Trade Act created a Wetlands Reserve Program through which the government purchases easements on wetlands from farmers, who then restore and protect the functional values of the wetland. The farmer is eligible to receive a share of the costs of implementing the land management plan for the easement. A similar Environmental Easement Program was created at the same time. Under this program, the government purchases easements on areas critical for wildlife on farms and ranches. The Forest Legacy Program, created in 1990, allows the secretary of agriculture to purchase conservation easements on forest areas that are threatened with development for other uses.

The 1969 US Internal Revenue Code began permitting landowners to take a charitable tax deduction for qualified conservation easements, thus creating an incentive for private landowners to protect their lands. Protection of natural habitat for plants and animals is one conservation purpose that qualifies an easement for tax deduction. This legislation has led to a dramatic increase in donation of conservation easements on private land to nonprofit land conservation organizations.

The 1985 and 1990 Farm Bills contained a strong incentive, in the "swampbuster" provisions (Highly Erodible Land Conservation and Wetland Conservation Compliance provisions), for farmers to maintain wetlands and not convert them to agricultural production. Anyone who drained, dredged, filled, or altered a wetland in order to produce an agricultural product would lose their eligibility for federal price supports, crop insurance, the Conservation Reserve Program, and all other federal benefits for all of their land and agricultural products. Other provisions of the Farm Bill also encourage protection of wetlands.

Other National Laws

National Environmental Policy Act of 1969

The most significant provision of the National Environmental Policy Act (NEPA) for biodiversity is the requirement that federal agencies prepare an Environmental Impact Statement (EIS) for every federal action or major legislative proposal that can have a significant impact on the environment. The EIS is expected to address potential impacts of proposed programmatic or site-specific actions on the environment, including biodiversity. An important feature of NEPA is the requirement that there be a public comment period on the draft EIS, and that public concerns be addressed in the final EIS. Most federal agencies evaluate the effects of their programs and projects on threatened and endangered species and sensitive habitats. However, many EISs have given inadequate attention to impacts on less threatened species, to overall habitat impacts, and to cumulative impacts on biodiversity.

Foreign Assistance Act of 1962

The 1962 Foreign Assistance Act (FAA) authorized the creation of the US Agency for International Development (USAID) and had little to say about environmental issues. Many of the areas of highest species diversity are located in developing tropical nations with insufficient economic resources to adequately protect biodiversity. As scientists and US citizens became aware of the global loss of biodiversity in the late 1970s, they began to lobby the US Congress to amend the FAA to ensure that US foreign aid to other countries helped to conserve biological diversity and tropical forests. Beginning in 1982, amendments to the FAA have required that the US assist developing nations with protection of habitats in protected areas, antipoaching measures for species, and research on biodiversity. Many of the most important and threatened national parks in developing countries now receive a significant portion of their operating budgets from bilateral assistance provided by developed nations. Sections 118 and 119 of the FAA require that USAID produce an annual report on the status of tropical forests and biodiversity in nations receiving US bilateral assistance.

Beginning in 1979, Section 216 of the FAA required that all US-funded development projects conduct environmental impact assessments to ensure that tropical rain forests and other biologically important habitats were not adversely impacted. USAID is also required to monitor the impacts of other multilateral donor projects, particularly those funded by the World Bank, and to provide input to the US representative to the World Bank about projects of environmental concern.

State and Local Legislation

The relative brevity of the following description of state and local laws should not imply that state and local legislation and regulations have less potential to contribute to conservation of biodiversity. In fact, the opposite is true. The following brief sections merely skim the surface of the laws for 50 states and thousands of municipalities in the US.

State Laws

All of the international and national laws and regulations outlined previously in this article have their impacts at local levels. Primary responsibility for stewardship of biological resources is reserved to states by the 10th Amendment to the US Constitution, which states that all powers not given to the federal government by the Constitution are retained by the states. Federal laws usually define minimum standards for the states to follow. States can, and sometimes do, pass laws that are more restrictive than federal laws in protecting plant and animal species, wetlands, forests, and other ecosystems.

States are responsible for managing the wildlife resources within their borders. Historically, most states have focused their wildlife conservation activities on a relatively small number of hunted or fished species. States retain authority to regulate and legislate many issues related directly to conservation of biodiversity; for example, managing the natural and biological resources on state-owned lands, setting hunting limits on deer and other resident, nonmigratory species, and approving permits for development and other projects that affect many habitats and species on privately owned land. State environmental law related to air and water quality and waste management can help to protect biodiversity indirectly. Perhaps the largest impact of state laws and local ordinances on biodiversity relates to how well the laws regulate the approval of land development plans so as to protect natural habitats and mitigate the impacts of development on habitat fragmentation and degradation.

State planning laws related to growth management and to land use planning can contribute to conservation. Thirteen US states have adopted growth management laws that require state plans to consider conservation of natural resources, wildlife, forests, and critical natural areas (Environmental Law Institute and Defenders of Wildlife, 2003). Many states also require that municipalities develop comprehensive plans, which might address conservation of biological resources along with land use and development issues. A number of state land use planning laws include mandatory elements such as natural resources, open space, wildlife habitat, and critical and sensitive areas (Environmental Law Institute and Defenders of Wildlife, 2003)

Only a small number of states have passed laws and regulations specifically addressing conservation of genes, species, and ecosystems, although the number is increasing. In 1992, the state of Michigan passed a Biological Diversity Conservation Act that required creation of a state biodiversity conservation strategy. A 1993 statute in New York State established a Biodiversity Research Institute with a mandate to coordinate state and private efforts to collect information about the state's biodiversity. The law also directed the New York Department of Environmental Conservation to implement programs to conserve rare and endangered species on state lands. In 1996, the state of Wisconsin passed a bill that revised the state's forest statute to ensure that management of state forests is consistent with maintaining native biodiversity and sustainable ecosystems. Many states have statutes that authorize creation of conservation easements on private lands. Hawaii created a statutory tax incentive that provides owners of forest or water reserve lands with relief from all property taxes in exchange for allowing the state to manage their lands. Hawaii also has strong statutes to control the introduction of exotic species.

Many states, including California and Tennessee, have created executive orders and memoranda of understanding regarding conserving biodiversity, but these do not have the force of law. Many other states have policies aimed at conserving biodiversity by creating land acquisition programs for areas with important biodiversity, prohibiting introduction of exotic plant and animal species, and requiring environmental impact assessments in order to minimize and mitigate the impacts of development projects on biodiversity.

Many states are involved in legislatively mandated, regional initiatives to conserve ecosystems that cross state boundaries. The 1987 Chesapeake Bay Agreement was signed by Virginia, Pennsylvania, Delaware, Maryland, the District of Columbia, the Chesapeake Bay Commission, and the US Environmental Protection Agency. The signatories agreed to work on a program to conserve the biological resources and water quality of the Chesapeake Bay system.

Local Ordinances

Local laws can be enacted by municipalities, towns, cities, counties, townships, or boroughs, depending on their regulatory authority and on what local government units are called in different parts of the US. Although ultimately all biodiversity is conserved or lost at a local level, rarely are local laws primarily concerned with biodiversity per se; rather, they tend to focus on biological resources that indirectly influence public health, safety, or welfare (e.g., protecting wetlands or wooded slopes to control floods or regulating hunting for safety or to maintain economic benefits). Increasingly, however, concerned citizens have been working to ensure that local laws do consider conservation at some scale. Local governments can address biodiversity through broad municipal comprehensive planning and through municipal ordinances that address specific human activities that affect biodiversity on a more site-specific level.

Whether required by state law or not, many municipalities create comprehensive plans that address future land uses in the context of economic growth and quality of life, including conservation of biological resources. Many comprehensive plans include sections on natural or environmental resources. Some municipalities are beginning to prepare separate plans for "open space" – undeveloped land that often contains natural habitats. Most open space plans identify specific environmentally important habitats and recommend specific conservation strategies for those habitats.

Municipalities have adopted a variety of ordinances intended to minimize impacts on species and natural habitats; often, these ordinances support the resource conservation goals established through comprehensive or open space plans. An increasing number of municipalities are enacting ordinances that offer stronger protection to stream and wetland ecosystems than the requirements of state and federal law. Some municipal ordinances focus on integrating protection and restoration of habitats with land development project approvals. For example, some local ordinances have required that development projects maintain or create natural vegetated

buffers of more than 25 ft along wetland boundaries and more than 75 ft along each side of a stream.

An increasing number of municipalities have passed forest conservation ordinances that limit development in woodlands and penalize developers who harm those habitats. Some municipalities will not approve logging permits for sensitive woodlands without implementation of a forest management plan that demonstrates sustainable practices such as leaving buffers along streams, restricting logging to seasons when birds are not nesting, or leaving dead trees standing for wildlife habitat. Forest conservation ordinances have required developers to demonstrate that their plans will minimize unnecessary forest fragmentation and will not exceed the percentage of any woodland allowed to be cut. Plans that exceed the legal cutting percentage might only be approved if the developer mitigates the impact through reforestation or off-site forest conservation. Some ordinances seek to minimize tree removal by setting a limit on clearing around buildings and roads (e.g. 25 ft).

Many local municipalities in agricultural areas have ordinances that prohibit growing of introduced "weeds" that might decrease the productivity of agricultural crops. In some residential municipalities, weed laws have been replaced with natural landscaping ordinances that promote and govern natural alternatives to lawn grasses, such as wildflower meadows and reforestation.

Conclusions

The evolution of government legislation and regulation of biodiversity during the past 100 years reflects both the dramatic increases in human impacts on biodiversity in every part of the earth and the increased understanding of the structure and function of the natural world and how it responds to human activity. Legislation has evolved from individual nations focusing on single species of economic value to multinational agreements aimed at protecting entire habitats and ecosystems. The five greatest current threats to biodiversity are habitat alteration and destruction; over-exploitation of species; invasive alien species; climate change; and nutrient loading and pollution. The laws summarized in this article address these threats to varying degrees. Despite improvements in legislation and regulations aimed at conserving biodiversity, the earth continues to lose genes, species, habitats, and ecosystems at a rapidly increasing rate.

There are many challenges to the effectiveness of the laws and regulations described in this article. Perhaps the greatest challenge is ensuring that existing laws are implemented as intended and then enforced. Many laws require scientific information on the status of species and ecosystems; such information is often unavailable or extremely costly to obtain. Many federal, state, and local government agencies charged with implementing conservation laws are challenged by insufficient resources to do so. Federal, state, and local implementing agencies are often pressurized by special-interest groups to compromise conservation mandates in favor of short-term economic gains. The laws in many countries go unenforced due to lack of oversight from private citizens and advocacy groups. Finally, most laws are based on prohibiting actions that harm biodiversity; such laws can be supplemented by additional legislation that creates positive incentives for conserving and enhancing biodiversity. Legislation that creates increased positive incentives could be particularly beneficial for biodiversity on private land since private landowners control 60% of the land in the US.

Although the laws and regulations regarding conservation of biodiversity have generally improved significantly during the past 30 years, it is important to acknowledge that they can be amended and weakened whenever the various levels of government, from international to local, choose to do so. Equally important, many laws and regulations in other sectors create incentives to destroy biodiversity, such as logging subsidies, underpriced grazing rights, subsidies for economic development, and international trade policies. An additional challenge in regulating species, habitat, and ecosystem conservation is the difficulty in separating natural rates of change in populations and ecosystems from changes that result from human activities that can be regulated. The goal is to conserve the natural process of change and evolution and to determine the "signals" for when human activity is leading to adverse impacts on species and ecosystems so that those impacts can be prevented before they occur.

Disclaimer

This article does not reflect the views or policies of the US Environmental Protection Agency.

See also: Conservation Efforts, Contemporary. Conservation Movement, Historical. Endangered Ecosystems. Natural Reserves and Preserves. Property Rights and Biodiversity. Resource Exploitation, Fisheries

References

Bass S (1999) *Protecting Biodiversity: Legal Mechanisms Concerning Access to and Compensation for the Use of Genetic Resources in the United States of America*. Washington, DC: Environmental Law Institute Research Report.

Bean MJ and Rowland MJ (1997) *The Evolution of National Wildlife Law*, 3rd edn. Westport, CT: Praeger Press.

Dickson B (1999) The precautionary principle in CITES: A critical assessment. *Natural Resources Journal* 39(2): 211–218.

Environmental Law Institute and Defenders of Wildlife (2003) *Planning for Biodiversity: Authorities in State Land Use Laws*. Washington, DC: Environmental Law Institute.

Houck O (1993) The Endangered Species Act and its implementation by the U.S. Departments of the Interior and Commerce. *University Colorado Law Reviews* 64(2): 277–370.

Nagle JC and Ruhl JB (2006) *The Law of Biodiversity and Ecosystem Management*. New York: Foundation Press.

Perrings C, Duraiappah A, Larigauderie A, and Mooney H (2011) The Biodiversity and Ecosystem services science-policy interface. *Science* 331: 1139–1140. March 4, 2011.

Saterson KA (2011) Chapter 9: Biodiversity conservation. In: Wiener JB, Rogers MD, Hammitt JK, and Sand PH (eds.) *The Reality of Precaution. Comparing Risk Regulation in the United States and Europe*, pp. 201–222. Washington, DC: Resources for the Future Press.

Grasshoppers and their Relatives

Piotr Naskrecki, Harvard University, Cambridge, MA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Piotr Naskrecki, volume 3, pp 247–264, © 2001, Elsevier Inc.

Glossary

Brachypterous Having tegmina and wings shorter than the abdomen but overlapping or touching each other on the dorsum.
Cerci Paired, usually not segmented structures at the end of the abdomen, sometimes used by males during mating to grasp the female's abdomen.
Hemimetabolous Having incomplete metamorphosis, that is, showing gradual change from molt to molt, with externally developing wing pads, and lacking any larval and pupal stages.
Hypognathous Position of the head when the mouth is directed toward the ventral side of the body.
Macropterous Having wings that are fully developed, reaching or exceeding the end of the abdomen.
Micropterous Having wings that are greatly shortened, not overlapping or touching on the dorsum.
Prognathous Position of the head when the mouth opening is directed forward.
Spermatophore A membranous package containing sperm that is transferred from the male to the female during copulation.
Stridulatory apparatus An organ of sound production based on the mechanism of rubbing one part of the body against another.
Tegmina (singular: tegmen) Thickened forewings.

Overview of Orthoptera

Introduction

Orthoptera are hemimetabolous insects, with nymphs resembling adult forms in their general appearance but lacking fully developed wings and reproductive organs. Mouthparts of orthopterans are of the chewing/biting type (mandibulate). The head is hypognathous, rarely prognathous; the antennae are usually long and threadlike, consisting of fewer than 10 to several 100 segments. The part of the body immediately behind the head, or the pronotum, is usually large, often shieldlike, and in extreme cases covers a large part (many katydids) or the entire body of the insect (pygmy grasshoppers).

The front and middle legs are cursorial (i.e., adapted for walking); yet in some cases, the front pair of legs may be modified for digging (mole crickets, pygmy mole crickets, and false mole crickets), or both the front and middle pairs may be modified for grasping (predatory katydids). In some orthopterans (most katydids and crickets), the front legs have tibial auditory organs (the ear). The hind legs of most orthopterans are saltatorial (i.e., modified for leaping), with large, muscular femora and long, slender tibiae. Some grasshoppers can perform repeated leaps of 2.6 m without any obvious signs of fatigue. This is possible primarily because of the presence of a protein called resilin in their hind legs. Resilin has superb elastic properties, with 97% efficiency in returning stored energy. This allows for explosive release of energy that catapults the insect, a task impossible with muscle power alone. Certain groups of orthopterans, especially those leading a subterranean life, have lost their ability to jump and their hind legs resemble typical cursorial legs. In some grasshoppers and certain Ensifera, the inner surface of the hind femur is modified for sound production (stridulation).

The wings of orthopterans are either fully developed or reduced to various degrees. Wing polymorphism, or the occurrence of individuals with well-developed and reduced wings within the same species, is common. The forewings are somewhat thickened, forming leathery tegmina. In most katydids and crickets, parts of the tegmina are modified for stridulation. The hindwings, when present, are fanlike, hidden under the first pair in the resting position. Often the hindwings are longer than the tegmina and protrude behind their apices. Wing buds of nymphal stages are always positioned in such a way that the second pair of wings overlaps the first one, whereas in adult individuals of micropterous and brachypterous species, the first pair of wings always overlap the second pair, despite their nymphal appearance. The base of the abdomen in grasshoppers has lateral auditory organs known as abdominal tympana. Females of most orthopterans have a prominent ovipositor at the end of the abdomen, derived from the eighth and ninth abdominal segments. Katydids and crickets usually have a well-developed ovipositor – sword-, sickle-, or needle-shaped ovipositor, whereas females of grasshoppers and their relatives usually lack a long, external ovipositor.

The number of described species of Orthoptera is approximately 25–600. However, this number probably represents only a half or less than the actual number of species of Orthoptera present on the Earth today. Tropical regions of South America, Africa, and Asia still remain virtually unexplored in terms of their grasshopper and katydid fauna, and many thousands of species are expected to be described in the future. The Australian orthopteran fauna is the best studied of all tropical regions of the world, yet more than 1000 Australian species still remain to be formally described despite being already recognized as new to science.

Members of the order Orthoptera inhabit virtually all terrestrial habitats of the world, from rock crevices of the littoral

zone of the oceans, subterranean burrows, and caves, to treetops and peaks of the alpine zones of mountain ranges. Both deserts and grasslands as well as dense forests have rich and unique orthopteran faunas. There are few purely aquatic forms, but many are associated with marshes and other semiaquatic habitats. Orthopterans are important members of nearly all terrestrial ecosystems, both in the role of consumers and prey. Massive outbreaks of some species of grasshoppers (and less frequently katydids and crickets) can cause enormous losses for the food industry and forestry. Locusts (a vernacular name for certain grasshopper species that tend to produce large seasonal outbreaks, and not a taxonomic entity) have been a part of human history from the very beginning of our agricultural tradition. They still pose a great risk for agriculture in many parts of the world, although they pose less of a problem now than a few hundred years ago, thanks mostly to better understanding of their population dynamics and application of various chemical and biological control measures.

Sound Production

The single characteristic most frequently associated with grasshoppers and their relatives is their ability to produce sounds. Although less widespread than generally believed, this ability is nonetheless quite common in some groups of orthopterans. The role of sound production is threefold and similar in some respects to that of birdcalls: (1) attraction of mates, (2) territoriality, and (3) disbursement of release calls (alarm calls produced when seized by a predator). The calls of orthopterans are usually species specific and play a very important role in species recognition. The information in the call may be coded in the form of frequency modulation (the pitch of the call changes through time, a mechanism best known in birds), time modulation (the pitch of the call remains the same throughout its duration but its temporal pattern is unique to the species), or both modes combined.

The dominant mechanism of sound production in Orthoptera is stridulation, which involves rubbing one modified area of the body against another. Contrary to the popular belief, no orthopterans (or any insect for that matter) produce sound by rubbing their hind legs against each other. Katydids (Tettigoniidae) and crickets (Gryllidae) produce sound by rubbing a modified vein (the stridulatory vein) of one tegmen against a hardened edge of the second tegmen (the scraper). The stridulatory vein is equipped with a file-like row of teeth, the number of which varies from a few to a few hundred. In most katydids, the stridulatory area is situated at the base of the tegmina, except in brachypterous species where it covers their entire surface. In crickets, virtually the entire surface of the tegmina is modified for stridulation. As a rule, in katydids, the stridulatory file is situated on the left tegmen and the scraper on the right one, whereas in crickets, the situation is reversed. A membranous area at the base of the tegmen, the mirror, amplifies the sound. In addition, some katydids (e.g., *Thoracistus*) use their enlarged, shield-like pronotum as an additional sound amplifier. Crickets, lacking the enlarged pronotum, use other methods of sound amplification, such as singing from burrows, the shape and size of which is attuned to boost certain frequencies (*Gryllotalpa*), or using the surface of a leaf for the same purpose (*Oecanthus*). The ability to stridulate is restricted almost exclusively to males, although in a few groups of katydids (Phaneropterinae and Ephippigerinae), females respond to the male's calls by stridulating as well. Their sound apparatus is not homologous to that of the males and is usually quite simple, lacking the sophisticated mechanism for sound amplification. In addition to tegminal sound apparatus, a few groups of katydids have developed other mechanisms of stridulation. For example, all members of the Australasian subfamily Phyllophorinae lack the typical wing stridulation and produce sound, instead, by rubbing their hind coxae against modified thoracic sterna. Mandibular sound production occurs in some members of Mecopodinae.

Grasshoppers use the same principle of stridulation, but instead of rubbing their tegmina against each other, these insects produce sound by rubbing the inner surface of the hind femur against one of the veins of the tegmen. In the slant-faced grasshoppers (Gomphocerinae), the inner surface of the femur possesses a file of small knobs, whereas the vein on the tegmen acts as the scraper. In band-winged grasshoppers (Oedipodinae), the vein has a row of pegs and the femur plays the role of the scraper. In addition to these two principal mechanisms, some grasshoppers stridulate by rubbing their hind legs against the sides of the abdomen (Pamphagidae) or by kicking their legs/feet against a modified area at the apex of the tegmen (*Stethophyma*). The males of South African bladder grasshoppers (Pneumoridae) produce sound by rubbing their hind legs against a greatly enlarged, balloon-like abdomen, and the sound produced in this way can be heard from a distance of at least 2 km. Australian false mole crickets (Cylindrachetidae) have a stridulatory file at the base of their maxillary palps, and some species of pygmy mole crickets (Tridactylidae) produce sound by rubbing a modified vein on the dorsal side of the tegmen against another vein at the base of the hindwing.

The sound frequencies produced by orthopterans during stridulation vary from a few kilohertz (most crickets and grasshoppers) to well above 100 kHz (some katydids). Crickets' calls are characterized by their tonal purity, with most energy of the call allocated within a narrow range of frequencies. Katydid calls vary from tonally pure (although often well above the human hearing range) to broad, noise-like signals. Grasshoppers produce mostly broad-spectrum, noise-like calls. Unlike many vertebrate calls, many orthopterans produce time-modulated rather than frequency-modulated signals. Crickets are a notable exception and most species produce melodious, bird-like, frequency-modulated chirps (although their calls are also time modulated.)

In addition to stridulation, some grasshoppers crepitate in flight. In this case, the sound is produced by hitting certain veins of the wings against other veins. This behavior is especially common among band-winged grasshoppers (Oedipodinae) and plays an important role in courtship and territorial displays.

A few members of normally acoustic orthopterans have lost their ability to produce airborne signals and instead have developed a number of substitute mechanisms of substrate-borne communication. Males of the oak katydids (*Meconema*)

lack the typical tegminal sound apparatus and instead produce sound by drumming with their hind legs against the bark of trees. Similar drumming behavior, although still accompanied by typical stridulation, is a component of the courtship behavior of the pitbull katydid (*Lirometopum*). The male crickets of the genus *Phaeophilacris* have lost their ability to stridulate and instead signal by rapidly flicking their tegmina back and forth while holding them in the vertical position. The near-field motion is detected by a female's cerci, rather than by her ears. Despite having a fully developed stridulatory apparatus, many neotropical members of the katydid subfamilies Pseudophyllinae and Conocephalinae spend little or no time stridulating, relying instead on substrate-borne tremulations. In this case, a male stands rigidly on a leaf or stem of a plant and violently shakes his entire body. The low-frequency waves are transmitted along the branches of the plant. The reason for this behavior is unclear, although a few hypotheses trying to explain it have been proposed. The most widely accepted one is the avoidance of predation by foliage-gleaning bats that are known to use insect sounds to locate their prey. Others include eluding satellite males (nonsinging males of the same species trying to intercept a female), avoidance of parasitoid flies, and helping females locate males on multibranched plants. Thanks to the rapid development of recording techniques in recent years, it is now known that many groups of orthopterans previously believed to be silent appear to employ a number of techniques of substrate communication.

Reproduction and Growth

The courtship and mating behavior of orthopterans provide some of the most complex and fascinating spectacles of the insect world. In addition to sound production, as discussed earlier, many species employ visual, tactile, and olfactory signals in their mating strategies. Visual communication is especially well developed in grasshoppers, where males often have bright, species-specific markings on different parts of their bodies, displayed in carefully choreographed sequences during courtship. Grasshoppers of the genus *Syrbula* are definite champions in this respect; and males of some species, in addition to calling, perform a dance consisting of 18 distinct movements. Visual signals employed by many diurnal grasshoppers include flight displays, where males flash their colorful hindwings (this is sometimes accompanied by crepitation), flagging with distinctly colored hind legs, and displays involving brightly colored, and often enlarged antennae. The courtship of katydids and crickets relies less on visual signals and more on sound and chemical cues, which are more appropriate for these mostly nocturnal animals. In both groups, males sometimes produce two different types of calls, a long-range advertisement call and a quieter, courtship song, performed only in the presence of a female. Females in some species may reply using either airborne signals or tremulation.

Chemical communication in Orthoptera is little studied, but there is evidence that at least some species employ it during courtship. Females of the New Zealand giant weta *Deinacrida rugosa* produce a musky substance that males use to locate females, whereas males of camel crickets of the genus *Ceuthophilus* posses thoracic glands that may also play a role in courtship. Field crickets *Teleogryllus commodus* use pheromones covering the female antennae to initiate courtship. Some other crickets use airborne pheromones in locating members of the opposite sex.

Copulation in orthopterans involves transfer of a sperm sac (spermatophore), which in some groups is accompanied by a large packet of nutritious proteins, the spermatophylax. Males of some orthopterans also allow females to feed on parts of their own bodies during copulation. Males of haglids (*Cyphoderris*) have their hindwings modified into thick, fleshy lobes, the sole purpose of which is to be eaten by the female during copulation. Females of tree crickets (*Oecanthus*) feed on males' thoracic glands during copulation; and in some crickets of the subfamily Nemobiinae, the females feed on enlarged spines on males' hind tibia. New evidence suggests that nuptial feeding from special glands on the male's body is also present in some katydids. Males of other orthopterans, lacking such tasty incentives, must rely on their strong grasp or modified cerci at the end of the abdomen to hold the female during copulation.

Oviposition takes place in a variety of substrates, such as soil, plant tissues, or rock crevices. In some cases, eggs are protected from desiccation by a foamy mass produced by the female. Nymphs usually hatch within a few weeks or months, but sometimes the eggs undergo a year-long, or longer, diapause. Few orthopterans display any kind of parental behavior, although some crickets (*Anurogryllus*) lay eggs in burrows guarded by the female. Female mole crickets (*Gryllotalpa*) not only lay eggs in special egg chambers underground but also actively care for the eggs by licking and removing fungal spores from their surfaces. The hatchlings stay with their mother for a few weeks before dispersing.

Food and Feeding

Orthopterans are extremely diverse in their food preferences and feeding techniques. Virtually all Caelifera (grasshoppers and their allies) are strictly herbivorous, very rarely engaging in cannibalistic behavior, and doing so only under crowded conditions. Most grasshopper species seem to be polyphagous (feeding on a wide variety of plant species), but some are oligophagous (feeding on a narrow spectrum of plant species) or monophagous (feeding on only one species of plants). The last are often associated with toxic, alkaloid-rich plant species, and these substances make the insect themselves inedible to many potential predators. Pygmy grasshoppers (Tetrigidae) are some of the few insects that feed on mosses and lichens.

Ensifera (katydids, crickets, and their relatives) range from herbivorous to omnivorous to strictly predaceous. Some katydids specialize on rather unusual food sources. Members of the Australian genus *Zaprochilus* feed exclusively on pollen and nectar of flowers. However, most katydids feed on a wide range of organic material. For example, the Central American Rhinoceros katydid (*Copiphora rhinoceros*) is known to feed on flowers, fruits, hard seeds, caterpillars, other katydids, snails, frog eggs, and even small lizards. Strictly predaceous katydids employ both the "sit-and-wait" strategy (Saginae) or actively forage and hunt living insects (Listroscelidinae). Some raspy crickets (Gryllacrididae) also actively search for insect prey by rapidly running along branches and grasping any sitting insect

they encounter. Crickets and cave crickets tend to be generalists in their dietary preferences, but rarely exhibit tendencies to feed on live prey. Some mole crickets have a behavior unique among orthopterans, and insects in general, of gathering and storing germinating seeds in circular chambers below the ground for later consumption.

Phylogeny and Taxonomy

Despite a long history of taxonomic research on orthopteroid insects and despite their economic importance, there is still little agreement as to the taxonomic position of many groups of the Orthoptera, and large portions of the current classification system is still based on nineteenth-century works. Older classifications also included groups, such as cockroaches (Blattodea), praying mantises (Mantodea), walking sticks (Phasmida), and earwigs (Dermaptera), in this order, each of which is now considered a separate, albeit closely related order. For the purpose of this review, the classification system proposed by Eades et al. (2011) has been adopted. This system divides the order Orthoptera into two suborders: Ensifera, with superfamilies Tettigonioidea (katydids), Hagloidea (grigs), Grylloidea (crickets), Schizodactyloidea (splay-footed crickets), Stenopelmatoidea (Jerusalem crickets and relatives), and Rhaphidophoroidea (camel crickets); and Caelifera, with superfamilies Eumastacoidea (monkey grasshoppers), Pneumoroidea (bladder grasshoppers), Tanaoceroidea (desert long-horned grasshoppers), Acridoidea (grasshoppers and locusts), Pyrgomorphoidea (gaudy grasshoppers), Trigonopterygoidea (leaf grasshoppers), Tetrigoidea (pygmy grasshoppers), and Tridactyloidea (pygmy mole crickets and relatives). The following differences generally allow Ensifera and Caelifera to be distinguished (note that there are numerous exceptions):

Ensifera (katydids, crickets, and relatives)	*Caelifera (grasshoppers and relatives)*
Antennae long and thin, with more than 30 segments	Antennae short, with less than 30 segments
Anterior tibiae usually with tympanal organs	Anterior tibiae without tympanal organs
Ovipositor long, sword- or needlelike	Ovipositor very short
Sound produced by tegminal stridulation	Sound produced by rubbing hind legs against tegmina

The monumental task of describing all living species of Orthoptera is far from complete. As the exploration of tropical regions of the world progresses, more and more species waiting to be formally described accumulate in museum collections worldwide. Unfortunately, at the same time, the number of specialists capable of such work is dwindling rapidly.

Paleontology

The Orthoptera is an old group, dating back to the Carboniferous period. Along with Titanoptera and Phasmida (Phasmatoptera), they represent the superorder Orthopterodea. Both Titanoptera and Phasmida branched off of the main stock of Orthopterodea in the early Permian period. Titanoptera disappeared at the end of the Triassic period, whereas Phasmida (walking sticks and their relatives) and Orthoptera still flourish today. The two main lineages of Orthoptera, Ensifera and Caelifera, probably separated in the late Carboniferous period. Most modern families of Ensifera appeared between the early Jurassic and the early Triassic periods. The oldest, still extant family of Ensifera, the Prophalangopsidae, appeared in the early Jurassic period. The oldest family of Caelifera, the Eumastacidae, appeared in the middle Jurassic period, followed by the Tetrigidae and the Tridactylidae at the beginning of the Cretaceous period.

Tettigonioidea (Katydids)

Introduction

The superfamily Tettigonioidea includes only one family, the Tettigoniidae (katydids or bush crickets), although some authors have elevated several subfamilies of katydids to the status of separate families (**Figure 1**). It is the second largest group of Orthoptera (grasshoppers being more species rich), with more than 6800 described species assigned to nearly 1200 genera and 20 subfamilies. Tettigonioidea are characterized by the presence of well-developed tegminal stridulatory apparatus in the males; long, thread-like antennae, usually longer than the body; four-segmented tarsi; short, nonarticulated cerci in the males; and a large external ovipositor in the females. The wings, when present, are usually held in a roof-like fashion over the abdomen. Virtually all members of the superfamily have bilateral tympanal organs on the front tibia as well as an additional acoustic trachea on the thorax behind the pronotum.

The Tettigonioidea are distributed throughout the world, but like most other orthopterans, they show their greatest diversity in the tropical and subtropical regions. Although many groups of katydids are quite diverse in temperate regions of the world, they are less common in areas with long, cold winters. A large proportion of the Tettigonioidea are arboreal but many also inhabit grasslands, savannas, deserts, or rocky



Figure 1 Tettigonioidea: Katydid *Celidophylla albimacula* from Costa Rica.

mountain tops. Most species are nocturnal or crepuscular, although several major groups consist almost exclusively of diurnal species.

Major Lineages

Tettigoniidae (Katydids or Bush Crickets)

Katydids are a large and diverse group of orthopterans, elevated by some authors to the rank of a superfamily with numerous families. More than 6900 species have been described, placed in nearly 1200 genera and 20 subfamilies. Most katydids are medium- or large-sized insects. The largest katydid, *Siliquofera grandis*, from New Guinea, has a wingspan of more than 250 mm. On the other hand, members of the Australian subfamily Microtettigoniinae have a wingspan of less than 10 mm.

Males of almost all species have a well-developed stridulatory apparatus at the base of the tegmina. The exceptions are the subfamilies Phyllophorinae, Phasmodinae, and some members of Meconematinae and Mecopodinae. Males of the Australasian subfamily Phyllophorinae, despite having fully developed wings, lack the stridulatory apparatus and instead produce sound by rubbing their hind coxa against specially modified thoracic sterna. Australian Phasmodinae lack wings altogether and are presumably totally silent. This is also the only subfamily of katydids having greatly reduced tympanal tibial structures (''the ear'') and lacking the thoracic auditory spiracle. All katydids have long, thin antennae, often longer than the body, in extreme cases exceeding its length three to four times. The head of katydids is hypognathous (mouth directed down), rarely prognathous (mouth directed forward) (Phasmodinae and Zaprochilinae); and sometimes bears distinct, cuticular, horn-like processes between the antennae. Their function is uncertain, but is likely related to the fact that many predators (especially bats and birds) tend to seize their katydid victims by the head. The pronotum is enlarged, saddle shaped, and in some forms greatly expanded, covering part of the wings and the abdomen. In the male, the enlarged pronotum often acts as a call amplifier. Curiously, the group with a particularly enormous, box-like pronotum, the Phyllophorinae, lacks the ability to produce sounds with their wings. The pronotum is also sometimes armed with sharp spines and processes. The wings of katydids can be fully developed or reduced to a varying extent, in some lineages (Tettigoniinae) showing significant intraspecific polymorphism in the degree of their development. The tegmina are sometimes reduced to scale-like structures, the only function of which is sound production. Hind wings in such taxa are either totally missing or are greatly reduced. Many katydids use their wings to enhance their remarkable mimicry of leaves, twigs, or bark to a degree that can fool professional biologists. Members of the neotropical tribe Pterochrozini are particularly convincing mimics of both live and dead leaves, perfectly imitating the look and feel of leaves, including simulated traces of herbivory, with fake lichens and mosses ''growing'' on them. A few katydids are Batesian mimics of wasps.

The legs of katydids are usually long and slender. Predaceous species frequently have rows of long spines along the ventral edges of the front and middle legs, using them for grasping and holding prey. In some Listroscelidinae, the spines are particularly long, forming a kind of net, used for scooping small flies and other insect prey. The hind legs are saltatorial, except for the members of some groups living in burrows or under rocks (Hetrodinae and Bradyporinae). The abdomen of males has a pair of cerci, in some groups modified to grasp the end of female's abdomen during mating. The ovipositor of females is usually long and sickle or sable shaped. The reduction of the ovipositor to short processes appears independently in ground-dwelling forms (Hetrodinae) and some arboreal ones (Phaneropterinae), apparently related to the fact that the eggs are laid on the surface of the soil or leaves, rather than inserted into them, as most katydids do.

The significant male investment in offspring is one of the characteristics of katydids that attracts many researchers to study their behavior and evolution. During copulation, males of many katydids produce a large, protein-rich spermatophylax, which is eaten by the female after copulation. The size of the spermatophylax can approach 60% of the male's body mass, making it an extremely costly and significant contribution to egg production. This causes the males of many species to be quite choosy in selecting their mating partners; and under certain circumstances, the females may compete for males, a role reversal remarkably rare in the animal world.

The geographic distribution of katydids reveals several large centers of endemism, the largest being Australia, where 5 of the 13 subfamilies known from the continent are unique to it. The neotropics have the highest number of described species (approximately 1800) but no endemic subfamilies, followed by Eurasia (approximately 1350 species), the Indo-Malaysian region (approximately 1150), and Africa (approximately 850 species and 2 endemic subfamilies). However, these numbers are very likely to be multiplied in the future if taxonomic work in these regions is intensified.

Major Subfamilies

Bradyporinae This is likely the most basal subfamily of katydids and is characterized by having a large, stout body, greatly reduced wings, and the tarsus with a metatarsal pulvillus. Members of this family are restricted in their distribution to southwestern Palearctic. Often grouped with several other subfamilies (Hetrodinae and Ephippigerinae), they nonetheless appear to be a separate clade of their own. Approximately, only 50 species have been described. All species of the subfamily are dark colored, often black, and resemble giant crickets rather than typical, graceful katydids. Most live on the ground or low vegetation. Some species are known to produce defensive autohemorrhage and squirt their hemolymph from orifices on their body if the insect is seized by a predator.

Phaneropterinae This is the largest subfamily of Tettigoniidae, with more than 2200 species described from all continents of the world (other than the polar regions). Members of this subfamily are medium- to large-sized insects, usually green and leaf-like katydids, characterized by a lack of lateral grooves on the tarsi, and rather primitive venation of the wings. Wing reduction is widespread in several lineages of the subfamily but the stridulatory apparatus is always present. Female stridulation is known in a few genera. Virtually

all species are phytophagous, although opportunistic cannibalism has been observed in a few species. Many species are exclusively arboreal. Eggs are laid in the soil, plant tissues, or on the surfaces of leaves and bark. Immature stages of many species mimic ants, tiger beetles, and even spiders, whereas adults of the same species usually mimic leaves or blades of grass. Several South American genera are superb mimics of pompilid wasps.

Acridoxeninae This is a small, aberrant subfamily of large Central African katydids, with only one species, *Acridoxena hewaniana*. Its relationship to other katydids is uncertain as it displays a mixture of very advanced and very primitive characteristics. The most interesting feature of *Acridoxena* is its astonishing mimicry of a dried, spiny plant. All parts of the body look like shriveled, twisted, dry leaves or twigs. Nothing is known of its biology or behavior.

Ephippigerinae This is a small Palearctic subfamily, with approximately 140 described species. All species have greatly reduced, scale-like wings, usually hidden under an enlarged, saddle-shaped pronotum. Females of many species stridulate and certain species have long been model organisms for studies on acoustic communication and courtship in insects. Some species occasionally cause minor agricultural damage.

Pseudophyllinae This is the second largest subfamily of Tettigoniidae, with approximately 1100 described species, distributed in tropical and subtropical regions of the globe. A notable exception is the true katydid (*Pterophylla camellifolia*) of eastern USA, which is both the northermost member of the subfamily and the source of the name ''katydid.'' Its characteristic, loud call resembles (although some listeners disagree) the syllables ''ka-ty-did.'' Most pseudophyline katydids are found in the tropical areas of South America and Southeast Asia. Many are spectacular mimics of leaves (tribes Pterochrozini and Pseudophyllini) and bark (tribes Pleminini and Cymatomerini). Secondary loss of stridulation and widespread presence of tremulation is characteristic of many neotropical members of the subfamily. On the other hand, Old World members of the subfamily belong to the loudest night singers of tropical forests. This disparity in their acoustic behavior has been explained by different hunting strategies of insectivorous bats in the New and the Old World, which in turn shaped different defensive strategies among acoustic insects. Nearly all species of the Pseudophyllinae are arboreal or at least associated with tall vegetation. A notable exception is the genus *Callimenellus*, which is known to inhabit marine littoral rock crevices in Hong Kong. All species seem to be phytophagous although opportunistic insectivory has been observed in a few Central American species. Virtually all species are nocturnal.

Tettigoniinae It is a large subfamily, with nearly 1000 described species. Most of the species of this subfamily occur in the temperate regions of the world, and only a handful of species are present in the tropics. Tettigoniinae achieved the greatest diversity in the regions of Europe, western North America, South Africa, and Australia characterized by the Mediterranean type of vegetation. Many species have a large, shield-like pronotum, hence the common name ''shield-backed katydids.'' Females always have a long, sword-like ovipositor and lay eggs in soil or the stems of herbaceous plants. Males stridulate loudly and many species are active during the day. A few species of shield-backed katydids are agricultural pests, the best known being the Mormon cricket (*Anabrus simplex*) of western USA. However, some species in China and Japan have been kept as pets for hundreds of years, and their pleasing calls have made a remarkable impact on the poetry and other arts of these countries.

Conocephalinae The conehead katydids are a large subfamily of Tettigoniidae, with more than 1100 described species worldwide. Many are characterized by a prominent fastigium of the vertex, forming a characteristic ''horn'' on the head. The diet of many of these katydids is restricted to grasses and their seeds, although quite a few species are predaceous. The conehead katydids are quite common in temperate regions of North America and Europe, although they reach their greatest diversity in the tropical areas of South America. Some species of Conocephalinae may become agricultural pests and a few have been known to form large, locust-like swarms (*Ruspolia*). Male investment in offspring can be significant in some neotropical conehead katydids, and an elaborate courtship behavior is common in such species.

Phasmodinae and Zaprochilinae These two, closely related subfamilies are restricted to Australia. Phasmodinae contains only one genus, *Phasmodes*, with three species restricted in their distribution to Western Australia. These insects are a good example of convergent evolution, resembling walking sticks (Phasmida) to an extraordinary degree. Both sexes are completely wingless, and females lack tibial tympana as well as thoracic auditory spiracles. The entire body is extremely elongate and thin, and the head is prognathous. A simple way of telling these insects apart from walking sticks is looking at the proportions of thoracic segments. In real walking sticks, the prothorax is extremely elongated, whereas in *Phasmodes*, it is the mesothorax that has such modification. These interesting katydids feed on both leaves and flowers of the heath habitats.

Zaprochilinae They have similarly elongated bodies and prognathous head, but all species have wings. They are strongly reduced in the genus *Kawanaphila*, but fully developed in the remaining three genera of the subfamily. Males of all 28 described species produce short, ultrasonic calls. Several species of the subfamily have been extensively studied with regard to the parental investment of males, which is significant and may lead to courtship role reversal. All members of the subfamily feed primarily on pollen and nectar, and they play some role in pollination of the flowers on which they feed.

In addition to the above-described subfamilies, the family Tettigoniidae also includes the following subfamilies: Austrosaginae, Hetrodinae, Lipotactinae, Listroscelidinae, Meconematinae, Mecopodinae, Microtettigoniinae, Phyllophorinae, Saginae, Hexacentrinae, and Tympanophorinae.



Figure 2 Hagloidea: Grig *Cyphoderris strepitans* from Wyoming, USA.



Figure 3 Grylloidea: Field cricket *Gryllus veletis* from northeastern USA.

Hagloidea (Grigs or Hump-Winged Crickets)

This ancient superfamily, dating back to the early Jurassic period, is represented by only eight extant species placed in a single family Prophalangopsidae (**Figure 2**). They are characterized by a rather primitive stridulatory apparatus lacking a proper mirror, head with antennae inserted near the lower margin of the eyes, and the metatarsus with a well-developed pulvillus. The largest of these Mesozoic relics is *Prophalangopsis obscura*, with the wingspan of more than 100 mm, found on the Tibetan Plateau. The remaining eight species belong to four genera distributed in western North America (*Cyphoderris*) and northeastern Asia (*Paracyphoderris*, *Aboilomimus*, and *Tarragoilus*). The biology of *Cyphoderris* is quite well known and some species of the genus are often used in studies on mating behavior and parental investment in offspring. The male's stridulatory apparatus is almost perfectly symmetrical, that is, males have fully functional files and scrapers on both tegmina and can produce sound with either the left or right tegmen (hence the name "ambidextrous crickets," is sometimes applied to these insects). The hindwings are reduced to fleshy lobes, which are devoured by the female during copulation. Males who had already mated once and are missing these courtship "snacks" must resort to other methods of holding the female's attention and instead use the "gin trap," a complex system of cuticular modifications whose role is to hold the female's abdomen firmly in place during copulation. These insects are also unique in their ability for being active, and even call during winter, from snow-covered bushes, when the temperature oscillates at approximately 0 °C.

Grylloidea (Crickets and Mole Crickets)

Introduction

The superfamily Grylloidea includes nearly 5300 described species placed in more than 670 genera and four families (**Figure 3**). Crickets are small- to medium-sized insects, the smallest species being only approximately 1.5 mm long (Myrmecophilidae) and the largest approximately 60 mm (Gryllotalpidae). The family is cosmopolitan in its distribution and some species occur even within the polar regions, albeit usually only in association with human dwellings. Most members of the subfamily are brown or black, very rarely green. The antennae are usually thin, threadlike, and longer than the body (with the exception of certain burrowing forms). The head is usually large and almost globular, although it may be somewhat elongate and prognathous in some lineages (Oecanthinae and Gryllotalpidae). The pronotum is usually quadrate, rarely produced backward as to cover the wings (a condition common in katydids). The tarsi of the legs are three segmented (superficially four segmented in Oecanthinae), and the front legs are sometimes strongly modified for digging. The front tibia typically have well-developed bilateral tympanal organs. Hind legs are usually saltatorial but are occasionally short and not adapted for leaping. Wings are usually present but may be reduced to various degrees or totally absent (Myrmecophilidae). The tegmina are held flat on the dorsum and their anterior (lateral, in the resting position) margins are bent downward at an angle, forming a box-like structure. The stridulatory apparatus of males is well developed in most groups, although it may be absent even in some fully winged taxa (some Pentacentrinae, Trigonidiinae, and Eneopterinae). The left and right tegmina of males have similar venation, but stridulation is performed virtually always with the right tegmen rather than with the left tegmen. The hindwings, when folded, often project far beyond the tegmina. Many taxa lack the hindwings and wing polymorphism is common. Male cerci are usually long, flexible, often superficially annulated, and never grasping. The ovipositor is usually long and needlelike, although it may be reduced or even absent in some crickets (Gryllotalpidae). Most species of crickets are nocturnal or crepuscular. They occur in almost all terrestrial habitats, from treetops to underground burrows, and many species are associated with aquatic and semiaquatic environments (such forms are usually very good swimmers). The great majority of crickets are opportunistic feeders and few seem to be exclusively predaceous.

Major Lineages

Gryllotalpidae (Mole Crickets)

This interesting family, which includes six genera and more than 100 described species of usually large, robust crickets, are

distributed worldwide. All species of the family spend their lives in underground tunnels, excavated with their extremely modified, shovel-like front legs. The head is prognathous and somewhat elongated, and so is the pronotum. The antennae are short and somewhat thickened, an apparent adaptation for their subterranean lifestyle. The wings are usually fully developed, rarely absent. The stridulatory apparatus of males is simple, lacking the mirror. The females' tegminal venation resembles that of the males, but lacks proper stridulatory organs. Cerci in both sexes are long and flexible, acting as an extra pair of antennae while the insects are moving backward in their underground corridors. Females lack an external ovipositor.

Mole crickets emerge from their tunnels only during courtship, although calling takes place underground. To broadcast their signals as far as possible, males build Y-shaped tunnels, the length and diameter of which are perfectly suited to amplify the dominant frequencies of their call. Females stridulate as well, but their call seems to have a territorial rather than a courtship role. Eggs are often laid in special brooding chambers and actively cared for by the female. Newly hatched nymphs stay in the chamber for a few weeks, feeding on humus and tender rootlets protruding into the burrow.

Several species of mole crickets are serious crop pests, not only because they feed on roots of plants but also because they damage entire root systems while digging tunnels. A species of *Gryllotalpa* recently introduced accidentally to Australia has quickly become one of the major threats to the Australian golfing industry by damaging carefully manicured putting greens.

Gryllidae (True Crickets and Tree Crickets)

The family Gryllidae is the largest lineage of the superfamily, with nearly 4700 species, assigned to nearly 600 genera in 21 subfamilies (there are several alternative systems of classification of crickets, and the number and status of higher categories varies among different authors). The body size of true crickets ranges from small (less than 5 mm) to large (approximately 50 mm). Most members of the family are rather stout insects, with short, thick legs. Phalangopsinae are a notable exception, having extremely long and slender legs and other appendages, making them resemble large spiders rather than typical crickets. Delicate bodies with slender legs also characterize most tree crickets (Oecanthinae). The head is generally globular, except in tree crickets, which have an elongated, prognathous head. The pronotum is generally quadrate. The wings and stridulatory apparatus may be fully developed, reduced, or totally absent. The fully winged forms often lack the stridulatory apparatus, whereas forms with greatly reduced tegmina can have a fully developed stridulatory apparatus. Species with well-developed stridulatory apparatus usually produce melodious, frequency- and time-modulated calls, and many species have been kept as singing pets in Asia and Europe for many centuries. Rarely, the call is short and clicklike (*Eneoptera*). Courtship behavior is often complex and involves acoustic, tactile, and olfactory signals. The courtship song, produced only in the close proximity of the female, is quieter and of a different structure than the advertisement call. It is usually produced by rubbing only a part of the stridulatory file against the scraper. Some spider crickets (Phalangopsinae) additionally drum with their legs. During copulation, males produce spermatophores, which in some species contain a large spermatophylax. Males of some species allow the female to feed on different parts of their bodies. Males of tree crickets have thoracic glands, which the female licks during copulation, and females of some pygmy field crickets (Nemobiinae) feed on males tibial spines during copulation.

The ovipositor in true crickets is usually long, needlelike, sometimes laterally flattened, and swordlike (Trigonidiinae) or reduced (*Brachytrupes*). Opposition takes place in soil or plant tissues. Maternal care is rare but sometimes quite well developed. Females of *Anurogryllus muticus* excavate extensive subterranean burrows as a nursery for their eggs, which are aggressively protected against all intruders. Newly hatched nymphs stay with their mother, who feeds them with unfertilized eggs, produced by her for the sole purpose of feeding her nymphs.

Major Subfamilies

Gryllinae This subfamily includes the true crickets, placed in more than 1100 species and more than 100 genera. They are distributed worldwide, and a few species closely associated with human habitats are virtually cosmopolitan. Morphologically, Gryllinae are rather uniform, having stocky bodies, with a globular head and thick legs. Most have well-developed wings and a stridulatory apparatus. True crickets are some of the first singing insects to be heard in the spring in the temperate areas of Europe and North America, as they often overwinter as late-instar nymphs and mature as soon as the temperature allows them to become active. North America has a particularly interesting fauna of true crickets, with several complexes of morphologically nearly identical, so-called "cryptic species." Some of the species appear to be chronospecies, that is, very closely related species that avoid hybridization by being sexually mature during different times of the year.

Most true crickets are ground dwellers and many are territorial, vigorously defending their burrows. In some species, a certain percentage of the males (satellite males) consistently exhibit a kind of sexual parasitism by intercepting females attracted to the call of another male. Several species form large swarms, and some (e.g., *Gryllus bimaculatus* in Africa) can become serious agricultural pests. The house cricket (*Acheta domesticus*) is a cosmopolitan insect associated with human dwellings and when occurring in large numbers can cause damage to stored food and other material.

Brachytrupinae This subfamily, often treated as a subset of the true crickets, includes about 20 genera and more than 220 species distributed mostly in the tropical and subtropical regions of the Old World, with relatively few species in the New World. The wings in some species are greatly reduced or absent, and the female's ovipositor is reduced. Some of these crickets are very large and produce calls of an intensity directly proportional to their size. Some species are gregarious and live in small underground colonies consisting of adults and nymphs of different ages. Some species of the

subfamily exhibit extraordinary maternal care for their eggs and nymphs.

Phalangopsinae The spider crickets are characterized by having a rather robust body supported, in most cases, by very long, thin legs. The antennae and palps in such forms are also extremely long and slender. Wings are frequently reduced or absent, and some winged species lack the stridulatory apparatus in the male. On the other hand, some species are loud singers, with a pleasing call, and males of several Asian species are greatly prized for their songs and are kept in special little cages as pets. The subfamily includes about 380 species in nearly 100 genera, achieving the greatest diversity in the tropics of Central and South America. Many species inhabit caves and rock crevices as well as spaces between buttress roots of trees in tropical forests. Others prefer tree trunks, leaf litter, and decaying wood.

Oecanthinae The tree crickets are rather unusual members of the family Gryllidae, having a somewhat elongated, prognathous head and a rather long pronotum. The wings can be either fully developed or completely absent. Winged forms have a very well-developed stridulatory apparatus, capable of producing loud, pure sounds. Males of the genus *Oecanthus* enhance the range of their call by positioning themselves in a hole chewed out in a leaf, with their tegmina aligned during singing with the surface of the leaf. By doing this, they significantly increase the area from which the sound radiates, using the same principle on which speakers in radios and other musical appliances are built. Thanks to this technique, these small insects are capable of becoming the dominant singers in many environments.

Females of tree crickets lay eggs in the stems of plants, and by doing so may damage young trees in nurseries and orchards, thus becoming pests. On the other hand, many species feed on aphids, thus balancing the negative effect of their reproductive behavior.

Mogoplistidae (Scaly Crickets)

Scaly crickets are small, rarely exceeding 15 mm in length. Their name is derived from the fact that their entire body is covered with minute scales, reminiscent of the scales on the wings of butterflies and moths. Similar to these insects, the scales on the body of scaly crickets often form beautiful color patterns. However, unlike butterfly wings, the wings of scaly crickets are not covered with scales and are strongly reduced, often entirely concealed under a somewhat elongated pronotum. About 370 species have been described, placed in 30 genera. They are distributed worldwide, but Australia and the Indo-Malaysian region seem to have the largest share of known species.

Myrmecophilidae (Ant-Loving Crickets)

This small family seems to be closely related to Mogoplistidae and probably does not deserve the status of a separate family. Only about 10 genera and fewer than 70 species are known, and all of these are closely associated with colonies of social insects, mostly ants, and less frequently termites. All species are minute, sometimes less than 1.5 mm, with an oval, wingless body and a small head with greatly reduced eyes. The hind femora are short and enormously enlarged.

Ant-loving crickets seem unable to live without a close association with ants or termites. Some species have a wide range of host species, whereas others are restricted to only one host species. The nature of this association is unclear. At least one myrmecophilous species, *Myrmecophila manni* from North America, appears to be able to mimic ant behavior to an extent that allows it to be fed by workers of its host, *Formica obscuripes*. Another species, *M. oregonensis*, feeds on residues of the substance that lubricates ants' bodies, left on the walls of passageways. A few species of these crickets appear to be parthenogenetic. Their eggs are relatively large and their embryonic development is long, sometimes lasting over a year.

Schizodactyloidea (Splay-footed Crickets)

The unusual superfamily Schizodactyloidea, which includes the single family Schizodactylidae, is unique in the peculiar modifications of the hind tarsi, which bear distinctive lobe-like processes that help the insects move on soft, sandy substrates. Species of the South African genus *Comicus* (**Figure 4**) are entirely wingless, whereas those of the southern Palearctic genus *Schizodactylus* have fully or partially developed wings. The Indian *S. monstrosus* has particularly long hindwings, which are coiled into tight, vertical spirals, a feature unique among all insects. All species of splay-footed crickets are very agile and feed on a variety of insects. A recently discovered fossil of splay-footed cricket indicates that forms remarkably similar to modern species were already present in Early Cretaceous period (approximately 108 million years ago.)

Stenopelmatoidea (Camel Crickets, Jerusalem Crickets, and Raspy Crickets)

Introduction

The taxonomy and phylogeny of this group of the Orthoptera seems to be in a greater disarray than that of any other lineage of the order. Many different names have been applied to this group of insects, and frequently various subgroups have been



Figure 4 Schizodactyloidea: Splay-footed cricket *Comicus capensis* from Namibia.

elevated to the levels of superfamilies or even infraorders. The superfamily Stenopelmatoidea includes about 980 species and more than 140 genera assigned to four families: Anostostomatidae (wetas), Cooloolidae (Cooloola monsters), Stenopelmatidae (Jerusalem crickets), and Gryllacrididae (raspy crickets) (**Figure 3**). Members of this superfamily are characterized by the lack of the tegminal stridulatory apparatus even in fully winged species. Tegmina, if present, are soft and characteristically curved around the abdomen. The head is usually large and bullet-like, and some forms have greatly enlarged mandibles. The antennae tend to be very long and threadlike, but in some subterranean forms (Cooloolidae), it may be strongly shortened. Tibial auditory structures are usually absent, and the thoracic auditory spiracle is small and simple. Acoustic communication in this group is not as common as in other orthopterans and is restricted to femoral stridulation (Gryllacrididae) and abdominal drumming (Stenopelmatidae). Body size varies from approximately 10–100 mm, and some species of New Zealand wetas are considered the heaviest insects in the world, reaching the weight of 70 g – nearly three times as much as an average house mouse. Coloration of species of Stenopelmatoidea is often drab, usually brown, black, or yellow, but never green.

The distribution of the superfamily is worldwide, with the exception of the Cooloola monsters, which are known only from northeastern Australia. Nearly all species of Stenopelmatoidea are nocturnal or crepuscular. Their habitats include treetops in tropical forests, savannas, and deserts, but a great proportion of species dwells in caves or underground burrows.

Major Lineages

Anostostomatidae (Wetas)

Approximately, 200 species placed in nearly 40 genera belong to this family. The body of these insects is usually fairly large and bulky, and New Zealand wetas are among the largest living insects (*Deinacrida heteracantha* reaches the weight of 70 g). The head is large and globular, with relatively short antennae. In some species, the powerful hind legs can serve as a weapon, and these species are capable of delivering strong and painful kicks. Wetas are generally wingless, but some neotropical species (*Anabropsis*) may have fully developed wings (**Figure 5**). Tree wetas (*Hemideina*) exhibit complex social behavior, where males maintain harems of females in their refuges in holes and cracks in trees. Other males frequently try to take over the harems, and fights involving head-on clashing and gaping of their jaws are quite common among males. New Zealand and South African wetas often display a strong enlargement of the male mandibles, whereas males of South African *Henicus pattersoni* carry long, sharp "tusks" below their eyes. These structures are used primarily in combat with other males and territorial displays; but some species, such as the South African Parktown Prawn (*Labidonasidus vittatus*), use them for digging deep burrows. Some North American species (*Cnemotettix*) produce silk with their mouthparts, which is used to line their burrows in coastal sand dunes.



Figure 5 Stenopelmatoidea: *Anabropsis marmoratus* from Costa Rica.

Cooloolidae (Cooloola Monsters)

Some of the most peculiar insects belong to this small family, consisting of only one genus and four species. They are restricted in their distribution to central and northern Queensland in Australia and were discovered only in 1980. Their appearance is so unusual that the first specimens ever collected were suspected of being fabricated as an entomological joke. The body of Cooloola monsters is so bulky that these insects, and especially the females, are almost incapable of walking on the surface of the ground. They spend their entire lives underground, digging through sand with their extremely thick and powerful legs. They are virtually blind and their short antennae consist of only nine segments. Their mandibles are elongate and sharp, and the maxillae (the second part of mouthparts) bear two dagger-like processes, which probably act as piercing organs, a situation absolutely unique within Orthoptera. The wings are vestigial in males and usually completely absent in females. The internal anatomy of these insects is also unusual in the fact that the foregut extends backward into the posterior part of the abdomen.

Little is known about biology of Cooloola monsters. They appear to "sand swim" through sand and do not construct burrows. They feed probably on body fluids of other subterranean animals.

Stenopelmatidae (Jerusalem Crickets)

This family includes about 40 described species, placed in 7 genera. They are known primarily from North and South America, South Africa, and southeast Asia. The legs are thick, powerful, and frequently well adapted for digging. Wings in Jerusalem crickets are usually absent and are rarely developed well enough for the insects to be able to fly. They do not have stridulatory organs on the tegmina, but many species communicate with elaborate percussive signals produced by drumming their abdomen against the substratum. The courtship behavior of Jerusalem crickets is an elaborate spectacle, involving a synchronized and precise positioning of the partners. If the female does not recognize the male as an acceptable partner, she devours him, thus preventing copulation. If the male is found to be acceptable, she allows him to mate and deliver to her a large spermatophore, and she attempts to eat him only afterward. This behavior may account for the rarity of male specimens of Jerusalem crickets in entomological collections. Most species of Stenopelmatidae

spend their lives on or close to the ground or underground. Species of the genus *Oryctopus* from India include wingless, robust, and entirely subterranean forms that closely resemble Australian Cooloola monsters; whether the resemblance is the result of convergence due to a similar lifestyle or these two lineages are closely related is still unclear. Many species are predaceous, but most seem to be opportunistic generalists. Jerusalem crickets have been reported to cause minor injuries to crops by gnawing on roots and parts of the plants close to the ground.

Gryllacrididae (Raspy Crickets and Leaf-Rolling Crickets)

This large family includes nearly 800 described species placed in over 90 genera. However, it is likely that these numbers represent only a fraction of the actual number of species, as the faunas of Gryllacrididae of most tropical areas are virtually unknown. Most species of the family have been described from southeast Asia and Australia, but they are also known from Africa and the neotropics. North American fauna of Gryllacrididae is limited to only one species, *Camptonotus carolinensis*, and they are absent from Europe.

Species of Gryllacrididae are characterized by the globular head with often extremely long antennae, which in some species may exceed the body length four to five times. The wings are usually present but may be reduced or absent in some species. The ovipositor of females is long and thin, and in nymphs held curved backward over the body. This morphological modification may allow them to fit more easily into their burrows or rolled leaves. Leaf-rolling behavior is quite common in this family. All stages of some species, from the tiniest nymphs to adults, are capable of producing silk, which they use to bind shelters made of leaves, twigs, sand, and other material. Burrowing species living in hot, arid zones are known to insulate themselves from the heat by sewing shut the entrances to their burrows. Tegminal stridulation is absent in the Gryllacrididae but many species produce raspy sound (hence the common name) by rubbing the inner side of the femora against the sides of the abdomen. This behavior does not seem to play any role in courtship but is rather a defensive strategy, as all life stages of both sexes are known to display it. During courtship, some species of this group produce loud drumming noises with their legs or abdomen. All species of Gryllacrididae are nocturnal. They are probably unique in the insect world in possessing individually recognizable pheromones, which allows a cricket to return to its own burrow, distinct from one of its conspecifics, after a night of foraging. Most members of the family are predators and feed on small insects. Some are extremely agile and actively hunt for prey by rapidly running along branches, searching for insects.

Rhaphidophoroidea (Cave Crickets and Camel Crickets)

Cave and camel crickets are represented by approximately 590 described species assigned to nearly 80 genera in a single family Rhaphidophoridae (**Figure 6**). They are distributed worldwide and at least one species is cosmopolitan. Species of this family are rather uniform in their appearance. All are completely wingless and have smooth, dorsally convex bodies. Most have long, spidery legs and similarly long and slender antennae and palps. The ovipositor is laterally flattened and often has distinctly serrated edges.



Figure 6 Rhaphidophoroidea: Camel cricket *Ceuthophilus maculatus* from northeastern USA.

Virtually all species are associated with dark and humid habitats, such as caves and holes in the ground. Several species sometimes invade cellars and basements of houses, and one species (*Diastrammena asynamora*) is cosmopolitan in greenhouses and other man-made environments. Most species appear to be gregarious, and many individuals of different ages are found in a single colony. Many species are associated with burrows of larger animals, mostly rodents or foxes. They feed on a great variety of organic material, and some species are associated with bat colonies, feeding on guano and the bodies of dead bats. At least one species, *Pallidotettix nullarhorensis* from Australia is known to cause death of unprotected chicks of nesting birds by swarming them and feeding on their tissues.

Eumastacoidea (Monkey Grasshoppers or Eumastacids)

Introduction

Most members of this large superfamily occur in the tropics, although a few are found in temperate regions (**Figure 7**). Nearly 1300 species have been described and placed in approximately 300 genera in eight families. As in all Caelifera, the antennae of eumastacids are very short, but unlike other members of the suborder, one of the last few segments of the antennae always bears a small tubercle of unknown function. The thorax can be short (Eumastacidae) to extremely long (Proscopidae); and in some groups, it can have a distinct cuticular crest. The hind femora typically are long and slender, although in some forms they may be greatly shortened and not adapted for leaping. In forms with long, leaping hind legs, they are characteristically splayed outward from the body when the insect is at rest. Wings are frequently reduced or absent, but when present they are very long and characteristically widened toward the end of the body. The tegmina are frequently transparent or semitransparent. The abdominal



Figure 7 Eumastacoidea: Monkey grasshopper *Lophothericles* sp from South Africa.



Figure 8 Pneumoroidea: Bladder grasshopper *Bullacris* sp from South Africa.

auditory organ is absent in most species of the superfamily. Little is known about the biology of members of this superfamily. None of the species appear to stridulate. Many species are arboreal, but some live on grasses in savanna-like environments and frequently have strongly elongated bodies, perfectly mimicking blades of grass (Australian Morabidae). Eggs are laid in batches in soil; but unlike those of other grasshoppers, they are never bound together by a foamy mass.

Major Lineages

Eumastacidae (Monkey Grasshoppers)

The larger of the two families in this superfamily, the Eumastacidae includes more than 230 species in nearly 50 genera, distributed in the Americas, Madagascar, and southern Asia. Body size in this family is small (10 mm) to medium (45 mm). The pronotum is always short, and the front and middle legs are relatively short. Hind legs are saltatorial and usually held widely spread, nearly horizontally. The male genitalic structures are very complex and of great value in taxonomic research. The ovipositor of females is short, with strongly serrated edges.

Proscopiidae (False Walking Sticks or Proscopids)

Proscopids are restricted in their distribution to South America, barely reaching the southernmost part of the isthmus of Panama. Approximately 215 species are known, placed in 33 genera. This family is readily distinguishable from other grasshoppers by their greatly elongate, twig-like appearance. Some species may reach a length of 165 mm, whereas the smallest are about 25 mm long. The head of these insects is unusually long and pear shaped, with large eyes situated near its narrow tip. The thorax is long and slender, and the front and middle legs are well separated. Femora of the hind legs are characteristically thickened at their bases and are not adapted for jumping. Virtually all species are completely wingless. Most species live on bushes and trees, and a few in savanna-type habitats. One species has occasionally been injurious to crops.

Pneumoroidea (Bladder grasshoppers or Pneumorids)

This small group of grasshoppers includes some of the most unusual orthopterans living today (**Figure 8**). The superfamily includes only 17 described species, placed in 9 genera of the single family Pneumoridae, restricted in their distribution to southern and eastern Africa. Bladder grasshoppers are large (50–100 mm) insects, with the wings either fully developed or somewhat reduced. The pronotum is greatly enlarged, forming a triangular helmet covering a large part of the body. The most unusual feature of the members of the superfamily is the shape and size of the male's abdomen, which in most species is grossly inflated, creating a balloon-like, semitransparent resonator. The stridulatory apparatus consists of a dozen or so cuticular ridges on the third abdominal segment. Thanks to the abdominal resonating chamber, the sound made by the males is among the loudest sounds produced by any insect. Females are capable of producing sound as well, but lack the abdominal stridulatory apparatus. Common among some species of this group of grasshoppers is the occurrence of sexual parasitism, where so called "satellite males" do not stridulate and develop the appearance of short-winged females. This allows them to remain close to a calling male without eliciting a territorial, aggressive response from him, and intercept females attracted to the song of the caller. Long-winged species are sometimes called "flying gooseberries" as they are frequently attracted to lights. Males who have the misfortune of being attracted to campfires explode like balloons if they fall into the fire.

Tanaoceroidea (Desert Long-horned Grasshoppers)

This small superfamily includes only three species of small grasshoppers, living on shrubby vegetation in deserts of southwestern North America. They differ from other grasshoppers in the development of their antennae, which are exceptionally long, as long as or longer than the body of the insect (other Caelifera have antennae that are shorter than half of the body length). All three known species are completely



Figure 9 Acridoidea: Lubber grasshopper *Taeniopoda reticulate* from Costa Rica.

wingless and are active during winter and early spring months, but virtually nothing is known about their biology.

Acridoidea (Grasshoppers and Locusts)

Introduction

Grasshoppers and locusts form the largest lineage of the Orthoptera, with more than 7750 species in more than 1700 genera assigned to 11 families (Figure 9). They are distributed worldwide, and although they are most diverse in tropical and subtropical regions of the world, they also occur in cold subpolar regions and high in the alpine zones of the mountains. The body size of these insects ranges from less than 10 mm to more than 120 mm, with a wingspan of 250 mm. The body form and the head shape are diverse, ranging from short and bulky to extremely elongate and slender. The antennae are always short and relatively thick, but often modified to form comb- or club-like structures. The hind legs are always distinctly thickened toward the base and always adapted for leaping. The wings vary from fully developed to completely absent. Fully winged forms have relatively narrow tegmina that are not widened toward their apices. Stridulatory organs are present only in some families and usually consist of a modified vein on the tegmen, against which the inner side of the hind femur is rubbed. The abdomen usually has lateral auditory structures, even in silent and wingless species. The female's ovipositor is always very short, with the upper and lower portions (valves) distinctly divergent at the apex. Eggs are laid primarily in the soil and are always protected with a foamy mass produced with special glands in the female's abdomen.

Most grasshoppers are diurnal, although nocturnal species are also known. They inhabit nearly all terrestrial habitats and a few species are aquatic (Paulininae). Open steppes and savannas of warmer regions of the world have particularly rich faunas of grasshoppers, but tropical rain forests have their share of unique forms too. Grasshoppers are also relatively diverse in temperate zones, with some species crossing the polar circles. A number of alpine species are adapted to activity in very low temperatures. Some of these forms exhibit the interesting behavior of changing their body color depending on air temperature. In cold weather, they become much darker, and thus more likely to absorb sun rays (*Kosciuscola*).

Virtually all species of grasshoppers are exclusively phytophagous. Opportunistic cannibalism has been observed in a few rare instances, but only under crowded conditions. Most grasshoppers are polyphagous, feeding on a number of plant species, although some grasshopper species are monophagous, feeding on only one species of plant throughout their life. Such species tend to feed on plants rich in toxic secondary compounds, and in doing so, they become unpalatable to potential predators themselves. They often advertise their toxicity with bright, aposematic coloration, and there are confirmed cases of human fatalities after eating certain African species of the family Pyrgomorphidae.

A few grasshopper species, mostly of the subfamilies Cyrtacathacridinae, Oedipodinae, Calliptaminae, and Melanoplinae, are serious agricultural pests and cause enormous losses to crops every year. They form huge swarms, which on descending on fields can devour the entire annual crop within minutes. Single swarms of the African plague locust *Schistoccrca gregaria* may fan out over hundreds of square miles and consist of 50 billion (50,000,000,000) individuals, weighing approximately 70,000 tons. Densities may reach 200 million locusts per square mile (500 million km^{-2}). In most species of locusts, this kind of behavior is seasonal, and some have two distinct morphological forms, the so-called solitary and gregarious phases. Sometimes the differences between the solitary and gregarious phases are so dramatic that they have been described as different species. The solitary forms occur in low densities, and usually have shorter wings and cryptic coloration. Under certain conditions, especially following rains, the densities of these grasshoppers increase dramatically, prompting great physiological, morphological, and behavioral transformations. Their coloration changes, usually to black and yellow, and they exhibit a tendency to form tightly packed groups that start marching across the land. On reaching the adulthood and the ability to fly, they take off in huge swarms, usually following the prevailing winds of a cold front that could lead them to areas of fresh vegetation.

Controlling massive outbreaks of grasshoppers is difficult, and many different chemical and biological agents have been used. As one can imagine, controlling a cloud of insects stretching across hundreds of miles is a nearly impossible task. Therefore, the most effort is directed toward controlling the young, freshly emerged bands of hoppers or destroying eggs laid in areas that are likely to originate massive outbreaks. Fungal pathogens have been quite successful in combating locusts, but the survival of fungal spores in arid conditions, where many locust species live, is limited.

Major Lineages

Pamphagidae (Earth Hoppers or Pamphagids)

Members of this family are large, robust grasshoppers, seldom smaller than 30 mm and often up to 90 mm long. More than

550 species in 94 genera have been described, mostly from arid regions of Asia and Africa, with a few representatives in southern Europe. Sexual dimorphism is often marked, with males in general being significantly smaller and often fully winged, as opposed to completely wingless females. The surface of the body is frequently rough and uneven, and many species blend superbly among rocks and pebbles. Some South African species (*Trachypetrella*) have been dubbed "living rocks."

Defensive stridulation is common in both nymphs and adults of many species, and it is achieved by various mechanisms (often more than one stridulatory mechanism is found in a single species). The typical advertisement stridulation is rare but, if present, the sound produced by males by rubbing their hind legs against the tegmina is exceptionally loud (*Lobosceliana*).

Romaleidae (Lubber Grasshoppers)

This family includes nearly 500 species in more than 100 genera, distributed mostly in warmer regions of North and South America. Romaleidae are large grasshoppers, usually strikingly colored. Species of *Tropidacris* and *Titanacris* are some of the largest orthopterans, with wingspan reaching 250 mm. The pronotum in the members of this family is often strongly keeled or crested. Wings, if present, have dense venation, and the hindwings often have a unique stridulatory mechanism formed by numerous fine teeth on one of the veins. Sound is produced by rubbing this vein against the underside of the tegmen. In addition to stridulation, some species produce a hissing sound by expelling air through their thoracic spiracles. Lubber grasshoppers have little economic significance, although a few species can defoliate trees on tropical plantations. In the USA, the best-known representative of this group is the Florida Lubber Grasshopper (*Romalea microptera*), a large and colorful grasshopper frequently used to study insect internal anatomy and external morphology.

Acrididae (True Grasshoppers and Locusts)

This is the largest family of Acridoidea, with nearly 6500 species in more than 1420 genera and 25 subfamilies. They occur in all regions of the world, and in nearly all terrestrial and semiaquatic habitats. Their body size ranges from less than 5 mm to approximately 100 mm. Body form and shape are varied, from robust and stocky to exceptionally slender and elongate. Wings can be fully developed, reduced, or entirely absent. Wing polymorphism is common in some groups (Gomphocerinae). Abdominal tympanal organs are generally present, absent only in some wingless forms. Many species stridulate by rubbing their hind femora against modified veins on the tegmina. Sometimes, an expanded area on the tegmen forms a membranous "speculum," which acts as a resonator.

All species of locusts and most other economically important grasshoppers belong to this family. Also, most grasshoppers common in temperate zones of the world belong here, and nearly the entire grasshopper fauna of Australia is the result of a massive adaptive radiation within one of the lineages of this family (Catantopinae). Members of the small subfamily Paulininae are unique among other grasshoppers in their nearly fully aquatic lifestyle. They are capable of skating on the surface of the water, diving, and swimming, and their eggs are laid on stems of submerged aquatic plants. The hind tibiae are flattened and widened at their apices, forming effective swimming paddles. Only two genera with five species belong to this family, and their original distribution was restricted to South America. Recently, one species (*Paulinia acuminata*) has been deliberately introduced to Africa in an effort to control the noxious aquatic plant *Salvinia*. This is probably the only example thus far of a truly beneficial species of grasshopper.

Pyrgomorphoidea (Gaudy Grasshoppers)

Approximately 470 species are known in this family, placed in more than 150 genera. They are distributed mostly in the tropics of the Old World, with only a few species known from Mexico and Central American countries.

The body is usually large, often reaching 90 mm. The pronotum is quite variable in form, sometimes adorned with granules or spine-like processes. Pyrgomorphidae do not use sound in their courtship behavior, but some make rustling noises with their wings when alarmed. Many members of this family have warning (aposematic) coloration, a characteristic frequently associated with their toxic properties (**Figure 10**). Some species eject protective froth or irritating fluids as a defense mechanisms, and quite a few species are poisonous. Some species are gregarious, especially during their early nymphal stages. A few have been known to cause serious damage to crops in Africa (*Zonocerus*).

Trigonopterygoidea (Leaf Grasshoppers and Wingless Desert Grasshoppers)

This small group of grasshoppers includes only 21 species and seven genera, divided into families Trigonopterygidae (leaf grasshoppers) from southeast Asia and Australia and Xyronotidae from southwestern North America (desert wingless grasshoppers). Leaf grasshoppers are associated with humid, dense forests, and in their morphology resemble dry, brown



Figure 10 Pyrgomorophoidea: Gaudy grasshopper *Zonocerus elegans* from South Africa.



Figure 11 Tetrigoidea: Pygmy grasshopper *Afrolarcus aequalis* from Guinea.



Figure 12 Tridactyloidea: False mole cricket *Ripipteryx* sp from Guyana.

leaves. Desert wingless grasshoppers are found in desert habitats, and they differ from other grasshoppers in the unique structure of the male copulatory organs (cerci), which resemble enlarged, multibranched antlers.

Tetrigoidea (Pygmy Grasshoppers or Grouse Locusts)

This interesting lineage of the Orthoptera includes approximately 1660 species in nearly 260 genera, placed in one family, Tetrigidae (some authors divide this superorder into two families: Tetrigidae and Batrachideidae) (**Figure 11**). Most species of pygmy grasshoppers are tropical but a few occur in temperate zones, often at very high elevations. Their body size is generally small, seldom exceeding 20 mm. The pronotum is always greatly enlarged, covering the entire body and often extending far beyond the end of the abdomen. Sometimes the pronotum is vertically expanded, forming a large crest or a leaf-like lobe. A few species display a pronounced polymorphism in the development of the pronotum and wings. Wings are usually present but may be reduced or absent in some forms. The tegmina are always reduced to small, oval lobes, with greatly reduced venation. The hindwings are well developed, fanlike, and usually completely concealed under the pronotum. The hind femora are very stout and all tarsi lack the arolium between claws. Members of Tetrigoidea have neither stridulatory organs nor the abdominal auditory tympana. Most species are associated with moist or semiaquatic habitats. Many can swim very well and dive when threatened. Some species are arboreal and many live on bark of trees in tropical forests. Pygmy grasshoppers feed on a variety of small plants, such as mosses, as well as on lichens and algae. They have little economic importance, although some may cause some damage in rice plantations.

Tridactyloidea (False Mole Crickets and Sandgropers)

This superfamily includes approximately 230 species in 17 genera divided among three families (**Figure 12**). Most species are small, 4–15 mm in length, although some sandgropers (Cylindrachetidae) can reach the length of approximately 40 mm. Families Tridactylidae and Ripipterigidae have characteristically modified hind tibiae, which bear a pair of long, apical flaps, whereas the tarsi are strongly reduced. Their tegmina, if present, are short and leathery and sometimes have a row of stridulatory pegs. The Tridactylidae are distributed worldwide, whereas the Ripipterigidae are restricted to Central and South America. They inhabit wet and semiaquatic habitats and some of them make dense networks of shallow burrows in the mud. The Cylindrachetidae are restricted to Australia and New Guinea, and one species is known from Patagonia. The body is strongly elongated and completely wingless. The pronotum is greatly elongated and cylindrical, with its lateral edges strongly curved downward, nearly meeting on the underside of the thorax. The front legs are modified for digging, and the second and third pairs are shortened. Sandgropers spend their entire lives underground in sandy soils. They are omnivorous, feeding on plants as well as insects and their larvae. Sandgropers produce odorous defensive secretion from a pair of abdominal glands.

See also: Insects, Overview

References

Eades DC, Otte D, Cigliano MM, and Braun H *Orthoptera Species File Online*. Version 2.0/4.0. 1 October 2011. http://Orthoptera.SpeciesFile.org

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF
Simon A Levin
Princeton University, Princeton, NJ, USA

VOLUME 4



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xv

Contributors to Volume 4 xxi

Foreword to the First Edition xxv

Preface to the First Edition xxvii

Preface to the Second Edition xxix

How to use of the Encyclopedia xxxi

VOLUME 4

Grassland Ecosystems 1
Osvaldo E Sala, Lucía Vivanco, and Pedro Flombaum

Grazing, Effects of 8
Mark E Hay and Cynthia E Kicklighter

Greenhouse Effect 18
Jennifer A Dunne, Stacy C Jackson, and John Harte

Guilds 33
Richard B Root

H

Habitat and Niche, Concept of 39
Kenneth Petren

Habitat Loss and Fragmentation 50
Heather Bird Jackson and Lenore Fahrig

Habitat Selection 59
Robert A Montgomery and Gary J Roloff

Hemiparasitism 70
David Smith, Todd J Barkman, and Claude W dePamphilis

Herbaceous Vegetation, Species Richness in 79
JP Grime and Jason D Fridley

Herbicides 87
Jodie S Holt

High-Elevation Andean Ecosystems 96
Mary TK Arroyo and Lohengrin A Cavieres

High-Temperature Ecosystems 111
Richard G Wiegert

Historical Awareness of Biodiversity 121
David Takacs

Hotspots 127
Ana SL Rodrigues

Human Impact on Biodiversity, Overview 137
Leslie E Sponsel

Human Impacts on Ecosystems: An Overview 153
Paul R Ehrlich, Claire Kremen, and Anne H Ehrlich

Hunter-Gatherer Societies, Ecological Impact of 162
Kathleen A Galvin and Tyler Beeton

Hybridization in Plants 166
Pamela S Soltis

Hymenoptera 177
Norman F Johnson

I

Identifying Conservation Priorities using a Return on Investment Analysis 185
Edward T Game

Impact of Ecological Restoration on Ecosystem Services 199
Holly P Jones

Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity 209
Thomas Kitzberger

Impact of Past Global Warming on Biodiversity 224
Gregory J Retallack

Implications of Urbanization for Conservation and Biodiversity Protection 231
Robert Ian McDonald

Inbreeding and Outbreeding 245
Katherine Ralls, Richard Frankham, and Jonathan D Ballou

Indicator Species 253
John H Lawton and Kevin J Gaston

Indicator Species: Computation 264
Pierre Legendre

Indigenous Peoples and Biodiversity 269
Víctor M Toledo

Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon 279
Giorgini A Venturieri

Indirect Land Use and Greenhouse Gas Impacts of Biofuels 293
Richard J Plevin and Daniel M Kammen

Insecticide Resistance 298
Ian Denholm and Greg Devine

Insects, Overview 308
Brian V Brown

In Situ, *Ex Situ* Conservation 313
Nigel Maxted

International Organizations and Biodiversity 324
Margaret Goud Collins

Intertidal Ecosystems 332
AJ Underwood and MG Chapman

Introduced Plants, Negative Effects of 345
William G Lee

Introduced Species, Impacts and Distribution of 357
Daniel Simberloff

Invertebrates, Freshwater, Overview 369
Margaret A Palmer and Holly L Menninger

Invertebrates, Marine, Overview 379
PJD Lambshead and PH Schalk

Invertebrates, Terrestrial, Overview 392
Olof Andrén

Island Biogeography 395
Dieter Mueller-Dombois

Isoptera 408
Takuya Abe and Masahiko Higashi

J

Justice, Equity and Biodiversity 434
Kai MA Chan and Terre Satterfield

K

Keystone Species 442
Bruce A Menge, Alison C Iles, and Tess L Freidenburg

L

Lake and Pond Ecosystems 458
Christian Lévêque

Landscape Corridors 467
Ellen I Damschen

Landscape Diversity 476
Debra PC Peters, Sarah C Goslee, Scott L Collins, and James R Gosz

Landscape Ecology and Population Dynamics 488
Scott M Pearson

Landscape Epidemiology 503
Felicia Keesing

Landscape Genetics 508
Andrew Storfer

Landscape Legacies 524
Jennifer M Fraterrigo

Landscape Modeling 531
Robert M Scheller

Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production 539
Wallace E Tyner and Farzad Taheripour

Land-Use Issues 555
Kees Klein Goldewijk

Land-Use Patterns, Historic 569
Oliver Rackham

Large-Scale Biodiversity Experiments 583
Andrew Hector and Michel Loreau

Latent Extinction—The Living Dead 590
Daniel H Janzen

Latitudinal and Elevational Range Shifts under Contemporary Climate Change 599
Jonathan Lenoir and Jens-Christian Svenning

Latitudinal Gradients of Biodiversity 612
Michael R Willig and Steven J Presley

Life Cycle Analysis of Biofuels 627
Jason Hill

Life History, Evolution of 631
Derek Roff

Limits to Biodiversity (Species Packing) 642
LB Slobodkin

Literary Perspectives on Biodiversity 649
William Howarth

Loss of Biodiversity, Overview 656
Robert Barbault

M

Macroscopic Patterns in Marine Plankton 667
Angel López-Urrutia

Mammals, Biodiversity of 681
Joshua R Ginsberg

Mammals, Conservation Efforts for 708
EJ Milner-Gulland, Sarah Durant, Rosie Woodroffe, and Richard Young

Mammals (Late Quaternary), Extinctions of 721
Paul S Martin

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 4

Takuya Abe[†]
Kyoto University, Kyoto, Japan

Olof Andrén
Swedish University of Agricultural Sciences, Valsätra-Ultuna, Sweden

Mary TK Arroyo
Universidad de Chile, Santiago, Chile, and Instituto de Ecología y Biodiversidad, Santiago, Chile

Jonathan D Ballou
Smithsonian Conservation Biology Institute, Washington, DC, USA

Robert Barbault
Museum National d'Histoire Naturelle CNRS, Paris, France

Todd J Barkman
Pennsylvania State University, University Park, PA, USA

Tyler Beeton
Colorado State University, Fort Collins, CO, USA

Brian V Brown
Natural History Museum of Los Angeles County, Los Angeles, CA, USA

Lohengrin A Cavieres
Universidad de Concepción, Concepción, Chile, and Instituto de Ecología y Biodiversidad, Santiago, Chile

Kai MA Chan
University of British Columbia, Vancouver, BC, Canada

MG Chapman
University of Sydney, Sydney, NSW, Australia

Margaret Goud Collins
National Academy of Sciences/National Research Council, Washington, DC, USA

Scott L Collins
University of New Mexico, Albuquerque, NM, USA

[†]Deceased.

Ellen I Damschen
University of Wisconsin-Madison, Madison, WI, USA

Ian Denholm
Rothamsted Research, Harpenden, UK

Greg Devine
Queensland Health, Cairns, QLD, Australia

Jennifer A Dunne
Santa Fe Institute, Santa Fe, NM, USA

Sarah Durant
Zoological Society of London, London, UK, and Wildlife Conservation Society, Bronx, NY, USA

Anne H Ehrlich
Stanford University, Stanford, CA, USA

Paul R Ehrlich
Stanford University, Stanford, CA, USA

Lenore Fahrig
Carleton University, Ottawa, ON, Canada

Pedro Flombaum
CONICET and Universidad de Buenos Aires, Buenos Aires, Argentina

Richard Frankham
Macquarie University, North Ryde, NSW, Australia

Jennifer M Fraterrigo
University of Illinois, Urbana, IL, USA

Tess L Freidenburg
California Ocean Science Trust, Oakland, CA, USA

Jason D Fridley
Syracuse University, Syracuse, NY, USA

Kathleen A Galvin
Colorado State University, Fort Collins, CO, USA

Edward T Game
The Nature Conservancy, Brisbane, QLD, Australia

Kevin J Gaston
University of Exeter, Exeter, UK, and University of Sheffield, Sheffield, UK

Joshua R Ginsberg
Wildlife Conservation Society, Bronx, NY, USA

Kees Klein Goldewijk
The Netherlands Utrecht University, Utrecht, The Netherlands

Sarah C Goslee
USDA-ARS Pasture Systems and Watershed Management Research Unit, University Park, PA, USA

James R Gosz
University of Idaho, Moscow, ID, USA

JP Grime
University of Sheffield, Sheffield, UK

John Harte
University of California, Berkeley, CA, USA

Mark E Hay
Georgia Institute of Technology, Atlanta, GA, USA

Andrew Hector
University of Zürich, Zurich, Switzerland

Masahiko Higashi†
Kyoto University, Kyoto, Japan

Jason Hill
University of Minnesota, St Paul, MN, USA

Jodie S Holt
University of California, Riverside, CA, USA

William Howarth
Princeton University, Princeton, NJ, USA

Alison C Iles
Oregon State University, Corvallis, OR, USA

Heather Bird Jackson
Carleton University, Ottawa, ON, Canada

Stacy C Jackson
University of California, Berkeley, CA, USA

Daniel H Janzen
University of Pennsylvania, Philadelphia, PA, USA

Norman F Johnson
The Ohio State University, Columbus, OH, USA

Holly P Jones
University of California, Santa Cruz, CA, USA

Daniel M Kammen
University of California, Berkeley, CA, USA

Felicia Keesing
Bard College, Annandale-on-Hudson, NY, USA

†Deceased.

Cynthia E Kicklighter
Goucher College, Baltimore, MD, USA

Thomas Kitzberger
INIBIOMA-CONICET and Universidad Nacional del Comahue, Bariloche, Argentina

Claire Kremen
University of California, Berkeley, CA, USA

PJD Lambshead
The Natural History Museum, London, UK

John H Lawton
Imperial College, London, UK

William G Lee
Landcare Research, Hamilton, New Zealand

Pierre Legendre
Université de Montréal, Montréal, QC, Canada

Jonathan Lenoir
Aarhus University, Aarhus, Denmark

Christian Lévêque
Institut de Recherche pour le Développement (IRD), Paris, France

Angel López-Urrutia
Instituto Español de Oceanografía, Gijón, Spain

Michel Loreau
McGill University, Montreal, QC, Canada

Paul S Martin
University of Arizona, Tucson, AZ, USA

Nigel Maxted
University of Birmingham, Birmingham, UK

Robert Ian McDonald
The Nature Conservancy, Arlington, VA, USA

Bruce A Menge
Oregon State University, Corvallis, OR, USA

Holly L Menninger
Cornell University, Ithaca, NY, USA

EJ Milner-Gulland
Imperial College London, London, UK

Robert A Montgomery
Michigan State University, East Lansing, MI, USA

Dieter Mueller-Dombois
University of Hawai'i, Honolulu, HI, USA

Margaret A Palmer
University of Maryland, College Park, MD, USA

Claude W dePamphilis
Pennsylvania State University, University Park, PA, USA

Scott M Pearson
Mars Hill College, Mars Hill, NC, USA

Debra PC Peters
United States Department of Agriculture, Agricultural Research Service, Las Cruces, NM, USA

Kenneth Petren
University of Cincinnati, Cincinnati, OH, USA

Richard J Plevin
University of California, Berkeley, CA, USA

Steven J Presley
University of Connecticut, Storrs, CT, USA

Oliver Rackham
University of Cambridge, Cambridge, UK

Katherine Ralls
Smithsonian Conservation Biology Institute, Washington, DC, USA

Gregory J Retallack
University of Oregon, Eugene, OR, USA

Ana SL Rodrigues
Centre d'Ecologie Fonctionnelle et Evolutive, Montpellier, France

Derek Roff
McGill University, Montreal, QC, Canada

Gary J Roloff
Michigan State University, East Lansing, MI, USA

Richard B Root
Cornell University, Ithaca, NY, USA

Osvaldo E Sala
Arizona State University, Tempe, AZ, USA

Terre Satterfield
University of British Columbia, Vancouver, BC, Canada

PH Schalk
ETI, Universiteit van Amsterdam, Amsterdam, Netherlands

Robert M Scheller
Portland State University, Portland, OR, USA

Daniel Simberloff
University of Tennessee, Knoxville, TN, USA

LB Slobodkin
State University of New York-Stony Brook, Stony Brook, NY, USA

David Smith
University of Maryland, College Park, MD, USA

Pamela S Soltis
University of Florida, Gainesville, FL, USA

Leslie E Sponsel
University of Hawai'i, Honolulu, HI, USA

Andrew Storfer
Washington State University, Pullman, WA, USA

Jens-Christian Svenning
Aarhus University, Aarhus, Denmark

Farzad Taheripour
Purdue University, West Lafayette, IN, USA

David Takacs
University of California, San Francisco, CA, USA

Víctor M Toledo
National University of México (UNAM), Morelia, México

Wallace E Tyner
Purdue University, West Lafayette, IN, USA

AJ Underwood
University of Sydney, Sydney, NSW, Australia

Giorgini A Venturieri
Universidade Federal de Santa Catarina, Florianópolis, Brazil

Lucía Vivanco
Universidad de Buenos Aires, Buenos Aires, Argentina

Richard G Wiegert
University of Georgia, Athens, GA, USA

Michael R Willig
University of Connecticut, Storrs, CT, USA

Rosie Woodroffe
Zoological Society of London, UK

Richard Young
Durrell Wildlife Conservation Trust, Trinity, Jersey

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

G

Grassland Ecosystems

Osvaldo E Sala, Arizona State University, Tempe, AZ, USA
Lucía Vivanco, Universidad de Buenos Aires, Buenos Aires, Argentina
Pedro Flombaum, CONICET and Universidad de Buenos Aires, Buenos Aires, Argentina

© 2013 Elsevier Inc. All rights reserved.

Glossary

Convention on biological diversity The Convention was first enacted in June 1992, has been signed by many countries, and its objectives are the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding.

Functional type A group of species that share morphological and physiological characteristics that result in a common ecological role.

Global biodiversity assessment (GBA) The Global Biodiversity Assessment is an independent peer-reviewed analysis of the biological and social aspects of biodiversity commissioned by United Nations Environment Programme.

Niche complementarity Refers to how the ecological niches of species may not fully overlap and complement each other. Consequently, an increase in the number of species that complement each other may result in a larger volume of total resources utilized and in higher rate of ecosystem processes.

Sampling effect Refers to the phenomenon where increases in the number of species increase the probability of including in the community a species with a strong ecosystem effect (Huston, 1997). This phenomenon yields an increase in ecosystem processes with increases in diversity without invoking niche complementarity. *See* Functional Diversity chapter (00061) in this work for ways of distinguishing between niche complementarity and sampling effect.

Extent of Grasslands

Grasslands are one of the major vegetation types in the world accounting for almost 40% of terrestrial surface (Shantz, 1954). They exist in all continents and cover a vast area of 49×10^6 km^2. In North America, grassland is the potential natural vegetation of most of the Great Plains and it reaches from the Chihuahuan desert in the south to the deciduous forests of Canada in the north, and from the Rocky Mountains in the west to the deciduous forest of the eastern United States (**Figure 1**). In South America, grassland is the potential vegetation of the vast pampas and most of the Patagonian steppe. Finally, in Asia, grassland ecosystems cover a huge area from Ukraine to China.

Determinants of Grasslands

Grasslands are water-limited ecosystems, and water availability defines their distribution in space. The amount of water available for plants primarily depends on precipitation amount and temperature. The former is the input of water and the latter controls the loss of water from the ecosystem, since as temperature increases so does the evaporative demand. Increases in temperature result in increases in soil evaporation and plant transpiration; consequently, for a similar precipitation regime, the water balance becomes more negative as temperature increases. In contrast with most biological phenomena, primary production in grasslands decreases with increasing temperature, highlighting the importance of the indirect mechanism of the temperature control on the distribution of grasslands (Epstein *et al.*, 1996).

Grassland ecosystems occur in areas of the world that have an annual precipitation between 150 and 1200 mm and mean annual temperature between 0 and 25 °C (Whittaker, 1975). Along precipitation gradients grasslands are located between forests and deserts. In North America, South America and Asia, clear E–W precipitation gradients exist. In North and South America there are very small changes in elevation along those precipitation gradients and vegetation changes are mostly accounted for by precipitation. At the eastern and wettest end, Tall-grass Prairie is the dominant vegetation, which is replaced by Mixed-grass Prairie and by Shortgrass Steppe at the driest end of the gradient. A similar pattern occurs in Asia and South America (**Figure 1**).

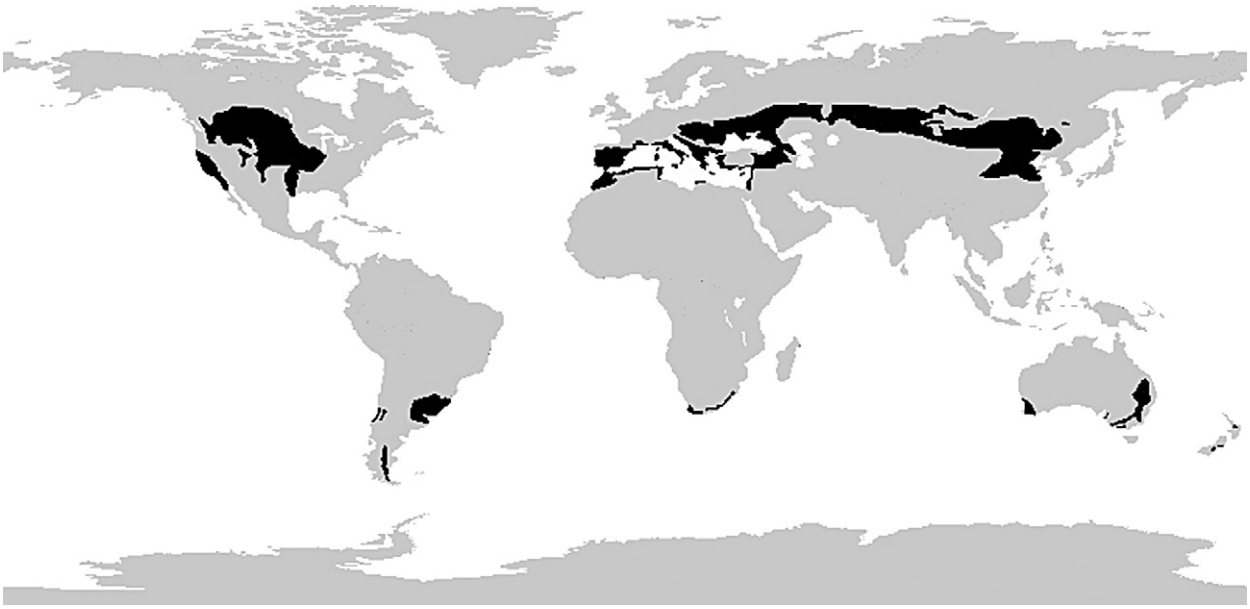


Figure 1 Map of the global distribution of temperate grasslands. Adapted from Figure 7.2 in Bailey RG (1998) *Ecoregions: The Ecosystems Geography of the Oceans and Continents*, 87 pp. New York: Springer.

Although temperature and precipitation are the major determinants of the distribution of grasslands, fire also may play an important role. Fire becomes particularly important in the grass–forest ecotones where the dominance of grasses or woody plants in many cases is determined by the frequency and intensity of fires. For example, in North American Tallgrass Prairie, the area covered by woody plants has increased dramatically in the past 100 years and the human intervention in reducing fire frequency is largely responsible for the change (Briggs *et al.*, 1998). Similarly, data from pollen profiles, tree ring analysis, and photographic sources documented a shift in the grassland–forest ecotone in northern Patagonia with woody vegetation invading grasslands (Veblen and Markgraf, 1988). Again, fire control implemented by land managers was responsible for the forest expansion.

Soil texture also modulates the distribution of grasslands by modifying the soil water-holding capacity and the location of water in the profile. First, water penetrates deeper into the soil profile in coarse-textured soils than in fine-textured soils because soil water-holding-capacity depends on soil texture and is lower in the former. Therefore, the same rainfall event would penetrate deeper in a coarse than in a fine textured soil. Second, grasses and woody vegetation have contrasting rooting patterns with grasses having predominantly shallow roots and shrubs and woody plants having deep roots (Jackson *et al.*, 1996). Consequently, for a given amount of precipitation, grasses dominate in areas with predominantly fine-textured soils (Sala *et al.*, 1997) (**Figure 2**). Finally, seasonality of precipitation and the synchrony between the wet and warm seasons also affect the dominance of grasses and woody vegetation. Locations where precipitation occurs during the cold season, when evapotranspiration is low, result in a deeper distribution of water in the soil profile and these locations would benefit woody vegetation over grasses (**Figure 2**). In synthesis, precipitation and temperature are the major drivers of the distribution of grasslands with their distribution modified at the edges by fire, soil and seasonality of precipitation.

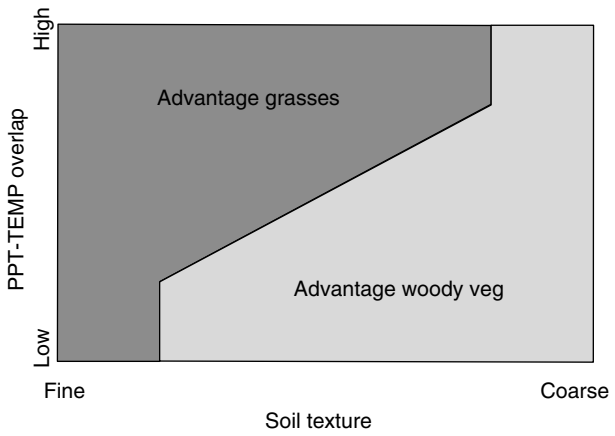


Figure 2 A conceptual model explaining the relative abundance of grasses and woody vegetation as a function seasonality of temperature and precipitation (congruence of warm and wet seasons) and soil texture. The gray shaded area represents conditions that are likely to favor the persistence of grasses, while the speckled area represents conditions that favor woody vegetation. The intersection of the two areas represents points where biotic influences are likely to be most pronounced. Adapted from Sala OE, Lauenroth WK, and Golluscio RA (1997) Plant functional types in temperate semi-arid regions. In: Smith TM, Shugart HH, and Woodward FI (eds.) *Plant Functional Types*, pp. 217–233. Cambridge: Cambridge University Press.

This article focuses exclusively on climatically determined grasslands, in contrast with grasslands resulting from human intervention. Anthropogenic grasslands are located in areas where potential natural vegetation is forest. Humans, in an

attempt to produce forage for domestic animals, have logged forests and have maintained these plots as grasslands by mowing them periodically.

Biodiversity

Biodiversity can be examined in many different ways, and multiple definitions exist for what constitutes "biodiversity." Nevertheless, the authors of the Global Biodiversity Assessment, using definitions originally proposed by the Convention on Biological Diversity, defined biological diversity as "variability among living organisms from all sources" (Heywood and Baste, 1995). Here, we focus on biodiversity in terms of taxonomically defined species, and the vast majority of studies quantifying ecosystem variation have used this measure. However, genetic biodiversity (genetic variation within a single species) and ecological diversity (including landscape diversity and functional group diversity) are also important components of biological diversity. The definition of biodiversity, therefore, will depend to a certain extent on one's objective and scale of interest, ranging from the gene to the ecosystem.

Plant Species Diversity

Floristic diversity in grasslands varies broadly, with many natural types of grassland having a very high level of plant species diversity, at times approaching that seen in mainland tropical forests (Groombridge and Jenkins, 2002). Grasslands are dominated by grasses (family Gramineae excluding bamboos). The Pampa region in Argentina represents some of the highest diversity grassland, with more than 400 species of grasses (Cabrera, 1970). In North America, more than 250 native species are found in Tallgrass Prairie (Freeman, 1998), the vast majority of which are perennial grasses.

Plant species diversity shows a great spatial variation. Grassland communities can be very species rich at fine spatial scales but tend to be similar and structurally simple over large areas (Groombridge and Jenkins, 2002). For example, plant species turnover among stands occupying different landscape position was 50% greater than among communities encompassing two degrees of latitude in the Flooding Pampa grasslands of Argentina (Perelman *et al.*, 2001). Grasslands tend to have low rates of endemisms, however, the climatic and soil gradients within them have led to substantial ecotypic variation and high genetic diversity (Groombridge and Jenkins, 2002). Plant species of grasslands can be categorized into four functional types: grasses, shrubs, succulents, and herbs (Sala *et al.*, 1997). The classification of plant species into functional types only has epistemological value and serves the purpose of facilitating the study of grasslands. This classification can be divided into many new subcategories or aggregated into fewer units depending on the needs of the analysis.

Grassland plant species can also be classified according to their photosynthetic pathway into C_3 and C_4 species. The two groups of species have differences in the physiology of photosynthesis and in the morphology of leaves that result in different ecological characteristics that separate them in time and space. Regional analyses of the distribution of these two types of grass species showed that C_3 species decrease southward in North America and northward in South America and C_4 species show the opposite pattern (Paruelo *et al.*, 1998). Similarly, the abundance of C_3 species increases whereas that of C_4 decreases along an altitudinal gradient (Cavagnaro, 1988). These biogeographical analyses correlate with ecophysiological studies showing that C_4 species have photosynthesis optima at higher temperature, have higher water use efficiency, and are better adapted to low water availability conditions (Kemp and Williams, 1980).

Animal Species Diversity

All major taxonomic groups are represented in grasslands, but despite their large areal extent (40% of the area of terrestrial ecosystems), overall faunal diversity is lower than in many other biomes. The number of bird and mammalian species that are found primarily in grasslands are estimated to be 477 and 245, respectively, representing only 5% of the world's species for each taxonomic group (Groombridge, 1992). Local diversity can be high in specific areas, for example, there are an estimated 208 avian species for Tallgrass Prairie (Kauffman *et al.*, 1998). However, general patterns show lower diversity for most taxonomic groups compared with other ecosystems.

One of the striking features of grasslands in terms of animal diversity is the presence of large herbivores as a prominent component of secondary production. These large grazing mammals have an important impact on the functioning of grasslands, altering patterns of nutrient cycling, primary production, and plant species composition (McNaughton, 1993), although their presence and diversity varies across different continents. For example, in the Great Plains of North America, nearly all the large grazing mammals went extinct during the glaciation of the Pleistocene, but the proliferation of a very few species, particularly *Bison bison* (plains bison), dominated the plant-herbivore interactions until the introduction of domestic cattle at the beginning of the twentieth century (Lauenroth and Milchunas, 1992). In contrast, African grasslands contain a very high level of mammalian diversity of grazers, with up to 20 species coexisting in a single reserve (Cumming, 1982). Finally, many South American grasslands evolved without the presence of large grazers, and their primary herbivory prior to the introduction of sheep and cattle was due to insect species (Bucher, 1982). Thus, although there is variation in the diversity of the large herbivores, their presence and importance is a distinctive characteristic of grassland ecosystems.

Small mammals, birds, reptiles, amphibians, and insects also play an important role in the functioning of grasslands. Species richness of small mammals is actually higher than that for large mammals (168 vs. 77 species overall), and they are mostly granivores or omnivores (Groombridge, 1992). In contrast, in Australian deserts small mammals are mostly insectivorous. Fluctuations in seed supply caused by unpredictable environmental conditions and the infertile soils could be an explanation for these differences (Morton, 1993).

Avian diversity in grasslands represents 5% of the total species of the world species diversity, and again the fluctuating

climate has an important control on this distribution. In this case, birds can migrate to remote areas outside of the grassland biome to seek alternative resources in periods of unfavorable conditions. In North American grasslands, which have a strong seasonality, there are large annual variations of passerines in response to climatic conditions. Additionally, within the grassland ecosystems, there exists a gradient of avian biomass that decreases with precipitation and primary production from the Tallgrass Prairie to the Mixed Prairie and Shortgrass Steppe (Lauenroth and Milchunas, 1992).

Reptiles in grasslands are less diverse than mammals and birds, and amphibians are less diverse than reptiles in the Tallgrass Prairie of North America (Kauffman *et al.*, 1998). Latitude has an effect on the biodiversity of reptiles and amphibians because they are ectothermic organisms, with an increase in the number of species from north to south of the Tallgrass Prairie (Kucera, 1992).

Insects are a diverse element of the terrestrial macrofauna of Tallgrass Prairie (Kauffman *et al.*, 1998), reflecting general patterns of diversity for terrestrial ecosystems in which insects represent more than 50% of the species (Strong *et al.*, 1984). They have a very important role as herbivores, pollinators, predators, parasitoids, and decomposers. Herbivorous insects are probably the most conspicuous functional group in Tallgrass Prairie (Kauffman *et al.*, 1998) and may replace large grazing mammals as the primary consumer in some South American grasslands (McNaughton *et al.*, 1993). Arthropods, constituting the largest proportion of invertebrates in the Shortgrass Steppe and primarily herbivores, take advantage of the large amount of belowground primary production (Lauenroth and Milchunas, 1992). In fact, most grassland invertebrate biomass is found within the soil and may be in the order of 100 to 1000 times as great as vertebrate biomass (Groombridge and Jenkins, 2002). In terms of species numbers, a soil invertebrate study in Tallgrass Prairie showed more than 200 species of nematodes, with fungivores constituting 40% of the nematode species (Ransom *et al.*, 1998) and the nematode biomass was exceeded only by that of bacterial and fungal groups.

Soil Microbial Diversity

Microorganisms are broadly defined as a group of microscopic life forms that include bacteria, archaea, viruses and unicellular eukaryotes like some fungi and protists. The diversity of microorganisms is known in much less detail than that of plants and animals but it constitutes a very important component of the soil biota (van der Heijden *et al.*, 2008). The large biomass of roots and other underground organs in grasslands and the high concentration of organic matter provide substrate for a large variety of microorganisms. The number of prokaryotes (bacteria and archaea) in grassland soils is estimated to be 32×10^{27} cells (Whitman *et al.*, 1998). Soil microbes represent a very diverse group of organisms in grasslands soils with a variety of functional roles. For the Shortgrass Steppe, the relative importance in terms of biomass of the different functional groups is bacteria > fungi > nematodes > protozoa > macroarthropods > microarthropods (Lauenroth and Milchunas, 1992).

There is a tight association between soil microbial communities and plant species composition (Wardle *et al.*, 2004). For example, arbuscular mycorrizal (AM) fungi are obligate root symbionts that are present in most terrestrial ecosystems and have roles in plant mineral nutrition, carbon cycling, and biotic interactions. The number of AM fungal taxa per host plant species is high in grasslands (8.3 fungal taxa per plant species) only higher in tropical forests (18.2 fungal taxa per plant species) (Opik *et al.*, 2006). Also, AM fungal communities greatly differed in their composition in grassland ecosystems. In addition, bacteria and fungi communities differed in soils associated with plant functional types in grasslands. In the Great Plains of North America, the bacterial community that occurred under the mesquite shrub harbored greater levels of richness than those occurring beneath a C_3 perennial grass, a C_4 midgrass or a C_4 shortgrass communities (Hollister *et al.*, 2010).

Biodiversity and Ecosystem Functioning in Grasslands

The relationship between biological diversity and the functioning of ecosystems has been central in ecology and grassland ecosystems have been crucial in testing those ideas. The biodiversity and ecosystem functioning hypothesis indicates that the rate of ecosystem processes, such as primary productivity or nutrient cycling, might increase linearly as species richness increases and that this relationship eventually saturates as ecological niches overlap increases (Vitousek and Hooper, 1993). The increase in ecosystem functioning can be interpreted as an evidence of niche complementarity; that is, the higher the number of species with niches that do not overlap the larger the total volume of resources exploited (Tilman *et al.*, 1997). For example, plots containing just shallow-rooted or deep-rooted species should have lower productivity than plots containing both groups of species that jointly have access to water and nutrients stored in both upper and lower layers of the soil. Alternatively, the same results can be interpreted as resulting from the sampling effect, which results from the increased probability of including species that outperform the others as the number of species in the mix increases (Huston, 1997; Tilman *et al.*, 1997).

Grassland ecosystems played a key role in testing the biodiversity-ecosystem-functioning hypothesis mostly because of the small size and short life span of grasses that made manipulative experiments feasible with few resources and in short periods of time. The most common manipulative experiment type that was used in grasslands is called "replacement series" where, at the beginning of the experiment, treatments differ in the number of species but keep plant biomass or density constant. This type of experiment normally includes treatments where each individual species grows alone. Monocultures are used to estimate biodiversity effect, that is the difference in production between full diversity and monoculture treatments (de Wit and van den Bergh, 1965; Loreau and Hector, 2001). The Relative Yield Total (RYT) compares the performance of a mixture of species with the average of monocultures and indicates the strength of the biodiversity effect (de Wit and van den Bergh, 1965). Values

greater than 1 indicate that biodiversity is responsible for the increase in primary production.

The first large-scale field experiment was located in the North American Tallgrass Prairie and showed a saturating curve with total plant cover and nutrient uptake increasing with species richness up to a level of approximately 10 species (Tilman *et al.*, 1996). Other sets of experiments across Europe and USA basically showed similar results (i.e., Hector *et al.*, 1999; Reich *et al.*, 2004; Roscher *et al.*, 2005). Grassland biodiversity–ecosystem-functioning studies yielded an average RYT of 1.2 with a range between 0.9 and 2.2 (**Figure 3**) (Flombaum and Sala, 2008). The only experiment performed in a natural ecosystem had an RYT significantly higher than all the rest of the experiments that were done using synthetic communities. This result reaffirms the role of biodiversity on ecosystem functioning in grasslands and suggests that the effect of biodiversity on ecosystem functioning might be even larger than previously thought. Consequently, expected global biodiversity loss (*see* The Future of Biodiversity in Grasslands) will have a large negative effect on grasslands primary production and carbon sequestration. It is likely that the magnitude of the biodiversity effect would be positively related to the co-evolutionary history of species in the mix and negatively to the frequency and intensity of disturbance (Sala, 2001), both characteristics would argue for a larger biodiversity effect in natural than in synthetic communities. The effect of the loss of biodiversity in grasslands and its effects on the C cycle would be an example of the synergism of global change drivers where biodiversity loss would constrain grassland ability to cope with the effect of other stressors such as climate change and ozone pollution.

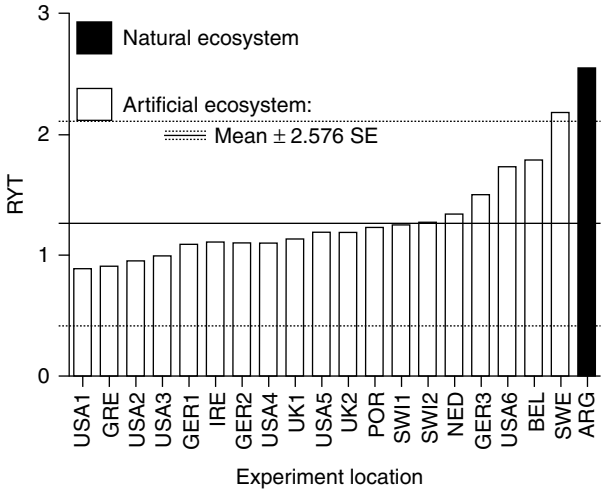


Figure 3 The effect of biodiversity on primary production in natural and artificial ecosystems. The RYT compares the performance of a mixture of species with the average of monocultures. Values greater than 1 indicate that biodiversity is responsible for the increase in primary production. The effect of biodiversity in natural ecosystem is larger than expected using artificial ecosystems Reproduced from Flombaum P and Sala OE (2008) Higher effect of plant species diversity on productivity in natural than artificial ecosystems. *Proceedings of the National Academy of Sciences of the USA* 105: 6087–6090.

A large effort has been made in identifying the relative contribution of niche complementarity and sampling effect. Using analytical tools, Loreau and Hector (2001) found that niche complementarity was the most important mechanism to account for increases in primary production with biodiversity. Long-term field experiments in grasslands further strengthened this argument by showing that the niche complementarity increased while sampling effect decreased (Fargione *et al.*, 2007; van Ruijven and Berendse, 2009). After more than a decade of the first experiments performed in grasslands, a general consensus exists that plant diversity has a linear and positive relationship with net primary production, and that relationship is mostly due to niche complementarity (Hillebrand and Matthiessen, 2009; Hooper *et al.*, 2005).

The Future of Biodiversity in Grasslands

Biodiversity in grassland ecosystems is seriously threatened by human activity. Two independent studies that developed global scenarios of biodiversity change for the next 50 and 100 years highlighted that grassland ecosystems were among the most vulnerable ecosystem types (Sala *et al.*, 2000, 2005). Depending on how interactions among drivers of biodiversity loss were modeled, grasslands ranged from the most threatened biome to the third most threatened behind tropical forests, arctic ecosystems, and southern temperate forests (Sala *et al.*, 2000). According to the most comprehensive regional scenarios, grasslands may lose between 8% and 10% of the vascular plant species by the year 2050 (**Figure 4** and Sala *et al.*, 2005). Differences between these estimates are driven by different socioeconomic scenarios with Order from Strength being the most pessimistic from the biodiversity point of view with 10% loss and Adapting Mosaic the most optimistic with 8% loss. Order from Strength is a scenario driven by security and protection issues that yield a fragmented world with islands of wealth isolated from the rest of the world where population growth and food demand remain highest (**Figure 4a**). High demand for food and low technology transfer from developed to developing nations result in large conversions of grasslands into croplands with the corresponding loss in biodiversity. In contrast, Adapting Mosaic is one of the most optimistic scenarios from the point of view of biodiversity loss where regional watershed-scale ecosystems are the focus of political and economic activity (**Figure 4b**). Local institutions are strengthened and societies develop a strong proactive approach to the management of ecosystems. Perhaps, the most striking feature of these scenario results is the small difference between the most optimistic and the most pessimistic scenarios. Even under the most optimistic scenario, large biodiversity losses are going to occur in grassland ecosystems.

What makes biodiversity in grassland ecosystems so vulnerable to human impact? Are grassland ecosystems particularly sensitive? Or, are they located in areas that will be affected the most? Grasslands are located in parts of the world where ecosystems are going to be hit the hardest by human activity (Sala *et al.*, 2000, 2005). Grasslands are among the biomes that are going to experience the largest conversion in land use because of their mild climate and favorable soil conditions that made them quite suitable for agriculture. The

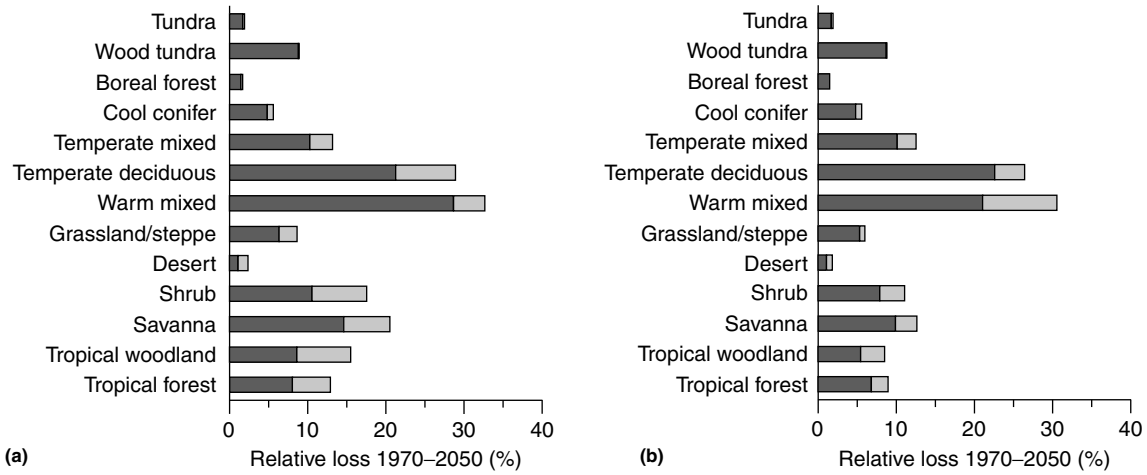


Figure 4 Scenarios of biodiversity change for different biomes for the years 2020 and 2050. Bars represent relative losses of biodiversity of vascular plants through habitat loss for different biomes for two scenarios: (a) Order from Strength and (b) Adapting Mosaic. Losses of biodiversity would occur when populations reach equilibrium with habitat available in 2050 and are relative to 1970 values. Darker bars represent scenarios for 2020 and lighter bars for 2050. Adapted from original Figure 10.6 in Sala OE, van Vuuren D, Pereira H, *et al.* (2005) Biodiversity across scenarios. In: Carpenter SR, Pingali PL, Bennett EM, and Zure kM (eds.) *Ecosystems and Human Well-Being: Scenarios*, pp. 375–408. Washington, DC: Island Press.

most dramatic changes in land use in grasslands are those that result from conversion into croplands. The conversion into agricultural land is not expected to be even across the world but driven by patterns of food demand and population growth that indeed are quite idiosyncratic. For example, the IMAGE2 model (Alcamo, 1994) predicts for the year 2100 a large increase in agricultural area in Africa and a reduction in North America resulting from both an increase in demand and an increase in intensification respectively. Biodiversity losses resulting from conversion to agriculture in one part of the world are not offset by a similar area that will be abandoned and is now reverting to grassland but that is located in a different part of the world. Therefore, total change in grassland area underestimates the impact of land-use change on biodiversity.

The major driver of biodiversity loss in grasslands in the next 50–100 years will be land-use change followed by climate change (Sala *et al.*, 2000, 2005). The third driver of biodiversity change in grasslands will be nitrogen deposition. Densely human-populated regions where nitrogen deposition is the highest are predominantly located in temperate regions where the potential native vegetation is that of grasslands.

See also: Agricultural Invasions. Asia, Ecosystems of. Biodiversity and Ecosystem Services. Climate Change and Extinctions. Climate, Effects of. Desert Ecosystems. Ecosystems of South America. Endangered Ecosystems. Fires, Ecological Effects of. Grazing, Effects of. Herbaceous Vegetation, Species Richness in. International Organizations and Biodiversity. Loss of Biodiversity, Overview. Soil Biota, Soil Systems, and Processes. Terrestrial Ecosystems

References

Alcamo J (1994) *Image 2: Integrated Modeling of Global Climate Change*. Dordrecht: Kluwer Academic Publishers.

Bailey RG (1998) *Ecoregions: The Ecosystems Geography of the Oceans and Continents*. New York: Springer.

Briggs JM, Nellis MD, Turner CL, Henebry GM, and Su H (1998) A landscape perspective of patterns and processes in tallgrass prairie. In: Knapp AK, Briggs JM, Hartnett DC, and Collins SL (eds.) *Grassland Dynamics: Long-Term Ecological Research in Tallgrass Prairie*, pp. 265–279. New York: Oxford University Press.

Bucher EH (1982) Chaco and Caatinga – South American arid savannas, woodlands and thickets. In: Huntley BJ and Walker BH (eds.) *Ecology of Tropical Savannas*, pp. 48–79. New York: Springer.

Cabrera A (1970) *Flora de la Provincia de Buenos Aires: Gramíneas*. Buenos Aires: Instituto Nacional de Tecnología Agropecuaria.

Cavagnaro JB (1988) Distribution of C_3 and C_4 grasses at different altitudes in a temperate arid region of Argentina. *Oecologia* 76: 273–277.

Cumming DHM (1982) The influence of large herbivores on savanna structure in Africa. In: Huntley BJ and Walker BH (eds.) *Ecology of Tropical Savannas*, pp. 217–245. New York: Springer.

Epstein H, Lauenroth W, Burke I, and Coffin D (1996) Ecological responses of dominant grasses along two climatic gradients in the Great Plains of the United States. *Journal of Vegetation Science* 7: 777–788.

Fargione J, Tilman D, Dybzinski R, *et al.* (2007) From selection to complementarity: Shifts in the causes of biodiversity–productivity relationships in a long-term biodiversity experiment. *Proceedings of the Royal Society B – Biological Sciences* 274: 871–876.

Flombaum P and Sala OE (2008) Higher effect of plant species diversity on productivity in natural than artificial ecosystems. *Proceedings of the National Academy of Sciences of the USA* 105: 6087–6090.

Freeman CC (1998) The flora of Konza Prairie. A historical review and contemporary patterns. In: Knapp A, Briggs J, Hartnett D, and Collins S (eds.) *Grassland Dynamics: Long-Term Ecological Research in Tallgrass Prairie*, pp. 69–80. New York: Oxford University Press.

Groombridge B (1992) *Global Biodiversity: Status of the Earth's Living Resources*. London: Chapman & Hall.

Groombridge B and Jenkins D (2002) *World Atlas of Biodiversity*. Berkeley, Los Angeles, London: University of California Press.

Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

Heywood VH and Baste I (1995) Introduction. In: UNEP (ed.) *Global Biodiversity Assessment*, pp. 5–19. Cambridge: Cambridge University Press.

Hillebrand H and Matthiessen B (2009) Biodiversity in a complex world: Consolidation and progress in functional biodiversity research. *Ecology Letters* 12: 1405–1419.

Hollister EB, Schadt CW, Palumbo AV, Ansley RJ, and Boutton TW (2010) Structural and functional diversity of soil bacterial and fungal communities

following woody plant encroachment in the Southern Great Plains. *Soil Biology and Biochemistry* 42: 1816–1824.
Hooper DU, Chapin FS, Ewel JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75: 3–35.
Huston MA (1997) Hidden treatments in ecological experiments: Re-evaluating the ecosystem function of biodiversity. *Oecolgia* 110: 449–460.
Jackson RB, Canadell J, Ehleringer JR, Mooney HA, Sala OE, and Schulze ED (1996) A global analysis of root distributions for terrestrial biomes. *Oecologia* 108: 389–411.
Kauffman DW, Fay PA, Kaufman G, and Zimmerman JL (1998) Diversity of terrestrial macrofauna. In: Knapp A, Briggs J, Hartnett D, and Collins S (eds.) *Grassland Dynamics: Long-Term Ecological Research in Tallgrass Prairie*, pp. 101–112. New York: Oxford University Press.
Kemp P and Williams G (1980) A physiological basis for niche separation between *Agropyron smithii* (C3) and *Bouteloua gracilia* (C4). *Ecology* 61: 846–858.
Kucera CL (1992) Tall-grass Prairie. In: Coupland RT (ed.) *Natural Grasslands: Introduction and Western Hemisphere*, pp. 227–268. Amsterdam: Elsevier.
Lauenroth WK and Milchunas DG (1992) Short-grass Steppe. In: Coupland RT (ed.) *Natural Grasslands: Introduction and Western Hemisphere*, pp. 183–226. Amsterdam: Elsevier.
Loreau M and Hector A (2001) Partitioning selection and complementarity in biodiversity experiments. *Nature* 412: 72–76.
McNaughton SJ (1993) Biodiversity and function of grazing ecosystems. In: Schulze ED and Mooney HA (eds.) *Biodiversity and Ecosystem Function*, pp. 361–383. Berlin, Heidelberg, New York: Springer.
McNaughton SJ, Sala OE, and Oesterheld M (1993) Comparative ecology of African and South American arid to subhumid ecosystems. In: Goldblatt P (ed.) *Biological Relationships between Africa and South America*, pp. 548–567. New Haven: Yale University Press.
Morton SR (1993) Determinants of diversity in animal communities. In: Ricklefs RE (ed.) *Species Diversity in Ecological Communties*, pp. 159–169. Chicago: University of Chicago Press.
Opik M, Moora M, Liira J, and Zobel M (2006) Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. *Journal of Ecology* 94: 778–790.
Paruelo J, Jobbágy E, Sala O, Lauenroth W, and Burke I (1998) Functional and structural convergence of temperate grassland and shrubland ecosystems. *Ecological Applications* 8: 194–206.
Perelman S, León R, and Oesterheld M (2001) Cross-scale vegetation patterns of Flooding Pampa grasslands. *Journal of Ecology* 89: 562–577.
Ransom MD, Rice CW, Todd TC, and Wehmueller WA (1998) Soils and soil biota. In: Knapp A, Briggs J, Hartnett D, and Collins S (eds.) *Grassland Dynamics: Long-Term Ecological Research in Tallgrass Prairie*, pp. 48–66. New York: Oxford University Press.
Reich PB, Tilman D, Naeem S, *et al.* (2004) Species and functional group diversity independently influence biomass accumulation and its response to CO_2 and N. *Proceedings of the National Academy of Science* 101: 10101–10106.
Roscher C, Temperton VM, Scherer-Lorenzen M, *et al.* (2005) Overyielding in experimental grassland communities – irrespective of species pool or spatial scale. *Ecology Letters* 8: 419–429.
Sala OE (2001) Price put on biodiversity. *Nature* 412: 34–36.
Sala OE, Chapin FS, Armesto JJ, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.
Sala OE, Lauenroth WK, and Golluscio RA (1997) Plant functional types in temperate semi-arid regions. In: Smith TM, Shugart HH, and Woodward FI (eds.) *Plant Functional Types*, pp. 217–233. Cambridge: Cambridge University Press.
Sala OE, van Vuuren D, Pereira H, *et al.* (2005) Biodiversity across scenarios. In: Carpenter SR, Pingali PL, Bennett EM, and Zurek M (eds.) *Ecosystems and Human Well-Being: Scenarios*, pp. 375–408. Washington DC: Island Press.
Shantz H (1954) The place of grasslands in the earth's cover of vegetation. *Ecology* 35: 142–145.
Strong DR, Lawton JH, and Southwood RRS (1984) *Insects on Plants: Community Patterns and Mechanisms*. Harvard Press: Cambridge, Massachusetts.
Tilman D, Lehman CL, and Thomson KT (1997) Plant diversity and ecosystem productivity: Theoretical consideration. *Proceedings of the National Academy of Sciences* 94: 1857–1861.
Tilman D, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379: 718–720.
van der Heijden M, Bardgett R, and van Straalen N (2008) The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* 11: 296–310.
van Ruijven J and Berendse F (2009) Long-term persistence of a positive plant diversity-productivity relationship in the absence of legumes. *Oikos* 118: 101–106.
Veblen TT and Markgraf V (1988) Steppe expansion in Patagonia? *Quaternary Research* 30: 331–338.
Vitousek PM and Hooper DU (1993) Biological diversity and terrestrial ecosystem biogeochemistry. In: Schulze ED and Mooney HA (eds.) *Biodiversity and Ecosystem Function*, pp. 3–140. Berlin: Springer.
Wardle D, Bardgett R, Klironomos J, Setälä H, van der Putten W, and Wall D (2004) Ecological linkages between aboveground and belowground biota. *Science* 304: 1629–1633.
de Wit R and van den Bergh JP (1965) Competition between herbage plants. *Netherlands Journal of Agricultural Science* 13: 212–221.
Whitman W, Coleman D, and Wiebe W (1998) Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* 95: 6578–6583.
Whittaker RH (1975) *Communities and Ecosystems*. New York: MacMillan Publishing Co.

Grazing, Effects of

Mark E Hay, Georgia Institute of Technology, Atlanta, GA, USA
Cynthia E Kicklighter, Goucher College, Baltimore, MD, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Associational refuge When a potential prey escapes or deters a consumer by associating with another organism, that interferes with the ability of the consumer to locate or attack the prey.

Bottom-up effects When physical parameters (such as nutrients) allow increased primary productivity and the effects of this cascade up through higher trophic levels. A simple example involves nutrient additions to a lake resulting in increased phytoplankton growth, this increasing resource of food allowing increased number of herbivorous zooplankton, and the abundance of zooplankton leading to increased densities of zooplankton-eating fishes.

Coevolution Joint evolution of two (or more) interacting species, each of which evolves in response to selection imposed by the other.

Competitive release The expansion of a species-realized niche that is associated with the absence or removal of competition from other species.

Grazer A consumer that removes only a part of each prey it attacks and thus rarely kills a prey in the short term. The term is commonly applied to animals that eat plants (herbivores) but can also be argued to apply to nonlethal microbial pathogens and to animals that cause tissue loss, but not death, when they feed from colonial animals such as sponges.

Resource partitioning When organisms differentially utilize resources such as food, space, or nutrients.

Secondary metabolite An unusual compound that does not play a role in primary metabolism (e.g., production and storage of energy). These metabolites were initially argued to be waste products but have often been shown to defend the producer from consumers, pathogens, or competitors.

Top-down effects When feeding by higher trophic-level, organisms has cascading effects on lower trophic levels. A simple example is when fishes (predators) consume zooplankton (herbivores) and, by lowering the numbers of these herbivores, allow increased densities of photosynthetic phytoplankton (primary producers).

Introduction

Herbivores play a major role in determining plant biodiversity by strongly affecting distribution, abundance, and community organization. Because plants are the major primary producers and often constitute the major biogenic structural complexity around which other community members are organized, effects of herbivores on plants commonly produce strong indirect effects on other trophic levels and on the biodiversity and function of the ecosystem as a whole.

Because almost all the energy on Earth is fixed *via* plant photosynthesis, it is an evolutionary mandate that higher trophic levels will need to consume plants and that successful plants will need to resist, tolerate, or escape herbivores, pathogens, and other natural enemies. Humans are dependent on plants for food, clothing, shelter, fuels, pharmaceuticals, and other necessities; thus, we often compete with other consumers (e.g., insect herbivores and pathogens of crop plants) for desirable plant resources. Owing to this continuing struggle to grow and harvest plant resources, plant–herbivore interactions are one of the more intensively studied areas in ecology, evolution, and agriculture. Research on this topic provides a rich resource for evaluating the effects of grazing on biodiversity. In numerous instances, examples from both marine and terrestrial studies show similar patterns. However, because of the more rapid growth of marine plants, the more intense grazing by marine herbivores, and the greater ease of manipulating seaweeds versus trees or shrubs, some of the clearest examples investigating the mechanisms of how herbivores affect plants are from marine communities.

In subtidal marine communities such as coral reefs or kelp beds, it is not uncommon for herbivores to be one of the primary forces determining the distribution and abundance of plants and often the species composition and diversity of the entire community. For example, herbivory by fishes or sea urchins on coral reefs keeps reefs largely devoid of macroalgae and allows corals to flourish by reducing competition from rapidly growing seaweeds, several of which produce chemicals that damage corals on contact. When these herbivores severely declined on the Caribbean island of Jamaica due to a combination of overfishing and urchin disease, coral reefs suffered severe overgrowth by seaweeds, and coral cover declined from more than 50% to less than 5% throughout this entire island nation (Figure 1). Thus, removing herbivores from these habitats converted species-rich coral reefs into a completely different community dominated by a limited number of seaweeds. Unfortunately, this pattern is not limited to Jamaica, coral cover has declined about 80% Caribbean-wide during the last three decades, and this pattern of rapid coral decline now seems to be occurring world-wide through interactive effects of global change and overfishing of reef herbivores.

Equally dramatic changes occur in temperate systems. Numerous ecological and paleontological studies have shown that subtidal communities in the temperate eastern Pacific shift from sea urchin-grazed barrens to lush kelp beds depending on the presence or absence of sea otters, which selectively forage on herbivorous sea urchins, reduce grazer biomass and allow kelps to flourish. If otters are removed from kelp communities due to hunting or predation, urchin densities increase, urchins drive kelps to local extinction, and

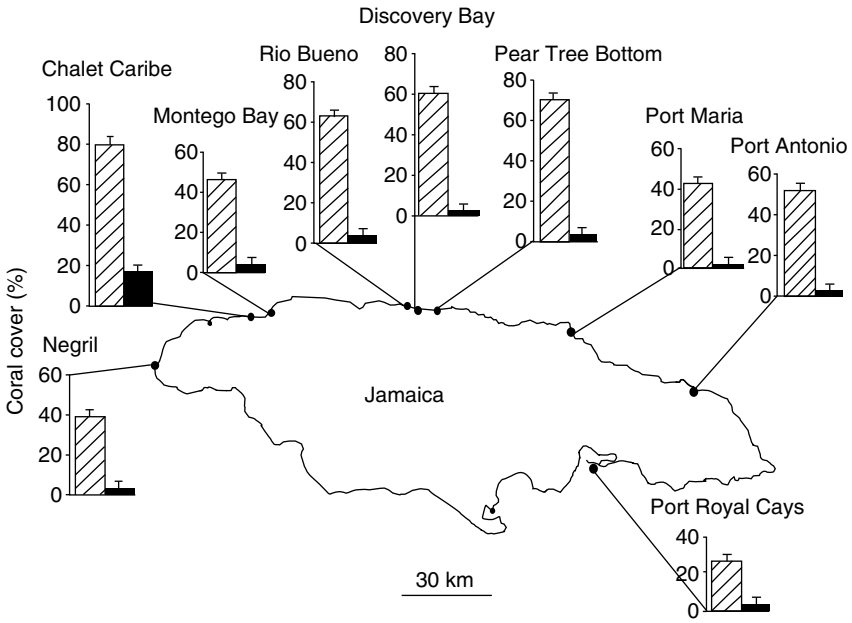


Figure 1 Large-scale changes in percentage cover of live coral at fore-reef sites along >300 km of the Jamaican coastline between 1977 (hatched bars) and 1993 (solid bars). For the island as a whole, coral cover decreased from 50 to <5%. Reproduced with permission from Hughes (1994) Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science* 265: 1547–1550. Copyright 1994 American Association for the Advancement of Science.

fishes, kelp-associated invertebrates, and seals or sea lions that feed on kelp bed fishes also decline dramatically. Thus, in the tropical example, grazers increase diversity by removing seaweeds and allowing corals to produce the biogenic structure that enhances reef biodiversity, whereas in the temperate system grazers directly remove the major biogenic structure (kelps) and cause a decline in overall biodiversity.

Herbivores can also cause major alterations in terrestrial communities. When myxomatosis eliminated rabbits in Britain, many grasslands reverted to scrub woodlands. Similarly, it is argued that grazing elephants damage trees and change closed woodlands or thickets into grassy savannas.

The previously mentioned instances are among the most dramatic examples of herbivore effects, but they serve to show the effects that grazing can have on biodiversity and community structure and to point out that there should be strong selection for plant traits that deter feeding by herbivores.

Grazers as Disturbance Regimes

Charles Darwin noted that one grass species tended to dominate and exclude other species in unmown areas of his lawn, but that numerous species coexisted in areas that were periodically mown. He recognized that grazers might play a similar role in diversifying plant communities by preventing a few rapidly growing species from competitively excluding others. In the late 1970s, Joe Connell formalized this relationship when he demonstrated for both tropical rain forests and coral reefs that intermediate frequencies or intensities of disturbance could lead to coexistence of a larger number of species. When disturbances were severe and frequent, only a limited number of hardy species could occupy a habitat. When disturbances were too infrequent or weak to remove competitive dominants, superior competitors tended to dominate. However, when disturbances occurred at intermediate levels, they removed enough of the competitively superior species to allow invasions by others but were not so severe as to prevent the occurrence of less hardy species. Thus, these intermediate levels of disturbance promoted the maintenance of greater biodiversity within a given community type. One of the most omnipresent forms of disturbance to plant communities is grazing, and it can play the diversifying role suggested by Connell. However, the effects that grazers have on plant biodiversity may also depend on the productivity potential of the environment. A recent review of how grazing affected plant diversity used data from a diverse range of marine, terrestrial, and fresh-water communities and found that grazing almost always increased plant diversity in environments characterized as having high nutrients (and presumably high potential growth rates), but often decreased diversity or had no effect on diversity when the environment was nutrient poor. These findings can be argued to mean that there are important grazer–nutrient interactions that need to be factored into the intermediate disturbance model. Alternatively, these findings simply mean that a given level of disturbance is "high" if it occurs in a habitat in which growth rates (and thus the ability to recover from the disturbance) are low, but it is "intermediate" if it occurs in a habitat in which rapid growth allows quick recovery.

As an example of intermediate grazing diversifying plant communities, work by Hixon and Brostoff (**Figure 2**) demonstrated that fish grazing rates on open areas of coral reefs were so intense that only the most resistant (coralline algae, which are red algae that resemble a thin calcified paint applied to a rock) or most tolerant (small filamentous algae that have basal sections embedded in the coral and that grow very rapidly) seaweeds could persist. In contrast, when fishes were

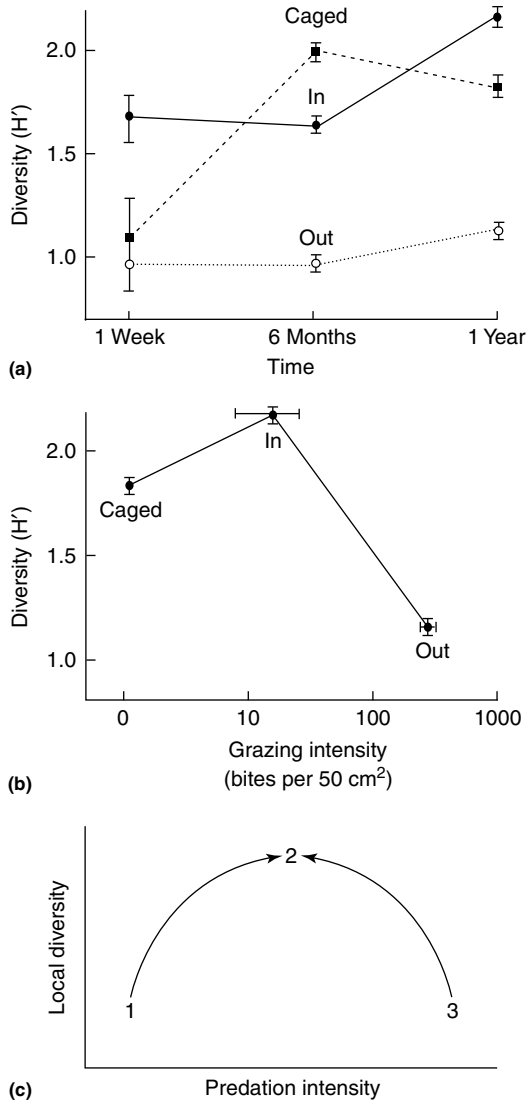


Figure 2 (a) Diversity, as measured by the Shannon–Weiner index H′, of algae on settling plates as a function of exposure time to three treatments: caged, inside damselfish territories (In), and outside territories (Out). (b) Diversity after 1 year (same data as in a) as a function of the intensity of fish grazing. (c) Graph of the intermediate disturbance hypothesis showing that consumers can enhance local diversity by increasing predation intensity from point 1 to point 2 or by decreasing overall predation intensity from point 3 to point 2. Reproduced with permission from Hixon and Brostoff (1983) Damselfish as keystone species in reverse: Intermediate disturbance and diversity of reef algae. *Science* 220: 511–513. Copyright 1983 American Association for the Advancement of Science.

excluded by cages, substrates were dominated by a few species of competitively superior macroalgae. However, in defended territories in which damselfishes drove most other herbivores away, bite rates on the substrate were intermediate between those experienced on open versus caged substrates, and at these intermediate rates of herbivory there were many more species of seaweeds. Similar patterns have been documented in Mediterranean grasslands in which plant species diversity is maximized if livestock densities are maintained at intermediate levels.

Direct Versus Indirect Effects

Grazers may indirectly affect biodiversity *via* trampling, tunneling, seed dispersal, nutrient regeneration, and selection for plant traits that have cascading effects on other processes. As herbivorous prairie dogs dig their burrows, they create patches of newly disturbed ground that can be colonized by plant species that are rare in undisturbed habitats. Their activities also create mosaics of vegetation structure and nutrient status that result in both buffalo and antelope differentially using centers or edges of prairie dog colonies versus areas remote from colonies. Prairie dogs also clip inedible bushes to low heights so that the bushes do not provide ambush sites for predators or obscure the prairie dog's field of view. Lowering bush height decreases competitive effects and allows increased coexistence of other plants. As animals graze, they release nutrient-rich urine and feces back into the environment; this release affects nutrient dynamics, the spatial patchiness of plant production, and thus plant diversity. Such effects occur in habitats ranging from open-ocean gyres to African savannas and livestock rangelands.

Initial feeding on plants can also produce a cascade of indirect effects on habitat parameters or plant traits that then affect use by other species. For example, when infestations of beetles feed on goldenrod, they not only affect their target plant but also their activities alter light penetration to the ground, soil water content, and soil nitrate content. Following these changes in the physical habitat, areas near attacked goldenrod are invaded by other plant species, increasing local biodiversity. These indirect effects also occur in marine systems. When the seaweed *Fucus vesiculosus* is attacked by the specialist smooth periwinkle, *Littorina obtusata*, it induces greater chemical defenses; these defenses are not induced when it is attacked by the generalist common periwinkle, *Littorina littorea*. Because Fucus serves as food and habitat for numerous grazers, chemical changes in response to one herbivore could affect many others, which occur in this instance. The induced plants are less palatable to three other consumer species, and less frequently colonized by these species in the field. Thus, one consumer had a negative indirect effect on a second, co-occurring consumer *via* induced changes in a shared host. Indirect effects like these are likely common in all ecosystems.

The Roles of Spatial and Temporal Scales of Grazing

Many investigations of herbivore effects on diversity have concentrated on the maintenance of within-habitat diversity due to selective feeding on dominant plants or on the effects of small-scale disturbances that create patches and produce numerous successional states within a single community. In contrast to these studies, several investigations from marine systems show that species richness of seaweeds is often maintained by mosaics of herbivore impact that are relatively predictable in space and persistent in time. Many seaweeds are predictably found only within certain physical habitats. This restricted distribution was initially assumed to be due to physiological limitations or to fine-scale resource partitioning along physical gradients; however, experimental

studies demonstrated that many of these restrictions were herbivore induced and did not result from habitat partitioning based on competitive interactions or differing physiological requirements. When spatial variations in herbivory are experimentally reduced, habitats that differ dramatically in seaweed composition often become more uniform, usually resulting in a lowering of total species richness due to the loss of between-habitat diversity. In most cases, species that dominate following herbivore removal are palatable species that exclude, or significantly reduce the abundance of, herbivore-resistant plants. This is consistent with the hypothesis that the evolution of antiherbivore defenses can be achieved only by diverting energy and nutrients from other needs. Thus, defenses appear to be costly, and in the absence of herbivory, less defended individuals or species will have higher fitness than more heavily defended individuals or species. This relationship is depicted graphically in **Figure 3**. When herbivores are rare, plant communities are dominated by competitively superior, but palatable, plants. When herbivores are common, defense is favored over rapid growth and competitive ability. Under these conditions, communities become dominated by unpalatable plants that depend on the herbivores to remove better competitors. If herbivores selectively graze competitive dominants, they tend to free resources for other species and increase biodiversity. However, if herbivores selectively graze competitively inferior species, they may hasten local extinction of poorer competitors and lower biodiversity. This latter scenario appears to be relatively uncommon.

In the following sections, we discuss how herbivores affect plant diversity over a range of spatial scales. Patterns occurring over a scale of hundreds of kilometers are called geographic patterns. Patterns that occur over a scale of several meters to many kilometers and involve habitats that clearly differ in physical traits are called between-habitat patterns. Patterns occurring over only a few meters and within areas that appear relatively homogeneous are called within-habitat patterns. Patterns on a scale of centimeters or millimeters are called microhabitat patterns.

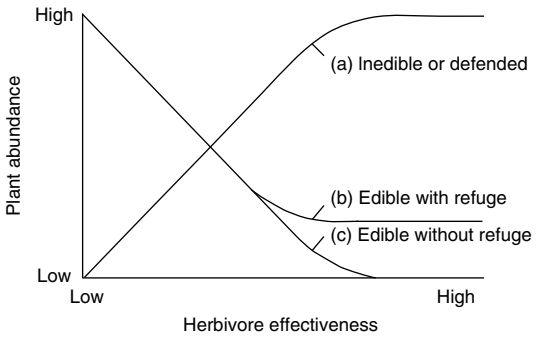


Figure 3 The hypothesized relationship between plant abundance and herbivore effectiveness for palatable and unpalatable plants. (a) Inedible or defended plants; (b) palatable plants with a refuge; (c) palatable plants without a refuge. Spatial patchiness of herbivory within a habitat allows multiple points on the *x*-axis to exist within the same habitat, thus increasing local biodiversity. Reproduced with permission from *Annual Review of Ecology and Systematics*, volume 12, Copyright 1981, by Annual Reviews, www.annualreviews.org.

Geographic Patterns

Large-scale patterns in plant distribution are often assumed to be due to physical parameters such as temperature. However, numerous recent studies suggest that some portion of geographic constraints on both specific species (or phenotypes within species) and morphological types may be due to geographic variance in herbivory. Because rates of herbivory are higher at lower latitudes, one might predict that tropical plants would be better defended than temperate plants. The few direct tests of this hypothesis suggest that it is generally true. Recent investigations demonstrated that seaweeds from tropical reefs were significantly less palatable and significantly better defended chemically than the similar seaweeds from temperate reefs. Similarly, studies of many species of salt-marsh plants across latitudinal gradients demonstrated that individuals from lower latitudes were subjected to greater rates of herbivory and were less palatable to a range of common herbivores than individuals of the same species from higher latitude regions. An investigation of freshwater plants and grazers shows similar trends. These patterns suggest that herbivore-generated processes shown in **Figure 3** could function on geographic and local scales and could enhance global biodiversity by creating differing selective regimes across these latitudinal scales.

Because tropical herbivores encounter more chemically defended prey than do temperate herbivores, one might expect that tropical herbivores would be more tolerant of plant chemical defenses than would temperate herbivores. This contention has not been extensively tested, but initial tests of a few tropical versus temperate fishes, sea urchins, and marine crustacea are supportive of this hypothesis, but indicate that there may also be considerable within-region variance in both plant defenses and consumer tolerance for those defenses.

Between-Habitat Patterns

Near-shore reefs often support more seaweed biomass and a different community of seaweeds than offshore reefs. Although both physical parameters and herbivores differ between these habitats, field experiments demonstrate that these differences in seaweed community structure are determined to a significant degree by differential herbivore activity. Early studies in the Caribbean demonstrated that seagrass beds, unstructured sand plains, and shallow unstructured reef flats supported different plant communities than did immediately adjacent reefs. Although the different physical regimes of these habitats appeared to be an adequate explanation for the differing seaweed communities, caging and transplant studies demonstrated that removal of herbivores allowed a more homogeneous distribution of seaweeds across these different habitats and resulted in lowered biodiversity of the reef system as a whole. In several studies, seaweed species typical of shallow reef flats invaded and dominated deeper areas on the reef once herbivores were removed. For example, when large fishes were excluded from parts of a coral reef in Belize, large macrophytes that were typical of nearby reef flats (but normally absent from the subtidal reef) invaded in large numbers and significantly decreased the abundance of corals and herbivore-resistant seaweeds during a 10-week period. Once

barriers were removed, these 10 weeks of algal growth were totally consumed in less than 48 h. A common response in herbivore exclusion experiments has been for palatable seaweeds to competitively suppress or exclude less palatable (and presumably better defended) seaweeds or other space occupiers (e.g., corals), again suggesting a tradeoff between defense and competitive ability. Recent work of this type on the Great Barrier Reef demonstrated that excluding grazers also caused large fields of algae to establish in areas formerly dominated by corals and that these algal beds suppressed the recruitment and survival of coals.

Within-Habitat Patterns

By aggressively excluding other herbivores from their algal mats, territorial damselfishes generate patches of intermediate grazing intensity in which algal species richness, evenness, and diversity are increased relative to either areas that are available to all grazers or areas from which grazers are excluded by cages. Several species that are normally excluded from unprotected areas due to grazing are found only within these mats. Thus, these territorial herbivores create distinct biotic patches on an otherwise more homogeneous background, raising species richness of the community as a whole.

In a similar manner, predators can strongly impact the spatial patterns of habitat use by herbivores and indirectly create mosaics of grazing intensity. In streams, algal-eating minnows avoid both shallow areas where they are susceptible to terrestrial predators and areas near larger predatory fishes. In both cases, these behaviors create patches of increased algal mass separated by heavily grazed habitats with minimal algal mass. Similar mosaics are created in terrestrial habitats when desert rodents forage more heavily under bushes than they do in more open areas where they are more exposed to predators. Within grasslands, excavating mammals, such as badgers, prairie dogs, and pocket gophers, provide newly bared soils for fugitive plants that are unable to successfully invade undisturbed plots. These within-habitat mosaics produce a series of successional states within the community and enhance plant diversity.

Microhabitat Patterns

In marine hard-substrate communities, small cracks and crevices serve as microhabitat escapes from some herbivores. Under these conditions, herbivory on exposed surfaces may select for defended seaweeds, whereas selection favors competitive ability in cracks and indentations. Work in experimental microcosms has shown that algal diversity decreases with increasing numbers of herbivorous fishes when substrate is smooth, but species number changes little with grazer densities if the substrate is more topographically complex. Thus, the presence of microhabitat escapes provided spatial refuges for palatable species and prevented their exclusion from the system.

Some types of plants can generate microhabitat escapes. There are many instances in agriculture and in the plant–insect literature in which interspersion of host plants with nonhost species interferes with the ability of insect herbivores to find or effectively utilize their host. Spiny desert plants sometimes function as ''nurse plants'' where less well-defended juvenile plants escape grazing vertebrates until they become large enough to escape because of their size. Plankton communities also exhibit these associational refuges; herbivorous copepods may quit feeding on palatable species of phytoplankton when they are mixed with adequate densities of chemically noxious species.

The effects of associational refuges on species richness have been addressed experimentally in several marine communities. Numerous seaweeds that are commonly driven to local extinction by grazers persist in herbivore-rich communities by growing on or beneath their herbivore-resistant competitors. The brown seaweed *Stypopodium zonale* produces cytotoxic compounds that deter feeding by Caribbean reef fishes and urchins. Numerous species of seaweeds are significantly more common near the base of *Stypopodium* plants than several centimeters away. When the chemically deterrent plant is removed, these less defended species are rapidly removed by herbivores, decreasing local species diversity. When plastic mimics of *Stypopodium* are placed in the field, they provide a partial refuge for palatable species but they are less effective than the real plants, suggesting that associational refuges are generated in part by the physical presence of a nonfood plant but that the plant's chemical repugnance makes the associational refuge more effective.

In temperate communities, palatable seaweeds can reduce losses to herbivores by growing on or near unpalatable seaweeds. Growing in close association with these unpalatable competitors drastically depresses the growth of palatable species, but the associational benefits, in terms of reduced herbivory, can more than offset this competitive cost. Thus, palatable species can be dependent on their unpalatable competitors to produce spatial refuges from herbivory and prevent their exclusion from the community. Both field and mesocosm studies indicate that removal of common unpalatable competitors can cause extinction, rather than competitive release, of associated competitors that are more palatable. These associational refuges were initially interpreted as arising from simple visual crypsis, but more detailed investigations suggest that chemistry plays a significant role.

As a final example, sulfuric acid can constitute up to 18% of the dry mass of brown alga *Desmarestia*; this concentration is sufficient to dissolve barnacles from coastal rocks when this alga is deposited in the intertidal by waves. In Chilean kelp beds heavily grazed by sea urchins, the palatable kelp *Macrocystis* cannot successfully colonize unless it invades an area encircled by *Desmarestia* plants, which appear to act as acid brooms that prevent urchins from entering the area.

The associational escapes discussed previously are opportunistic rather than coevolved. As such, they may be used by many organisms in a wide variety of situations. There are, however, more intimate associational refuges or defenses that may be coevolved. Many microbes are predictably associated with specific species of macroorganisms. Because of the broad ability of microbes to produce bioactive secondary metabolites, many host organisms could be coevolved with certain microbes because the microbes produce compounds that defend the host from natural enemies. For example, certain fungi infect host grasses, produce toxins, and by doing so make the

grasses much more resistant to herbivores. Similarly, marine cyanobacteria grow in host sponges and produce bioactive secondary metabolites that can protect the host. There are also instances in which marine bacteria that are specialized to certain host surfaces produce metabolites that chemically defend their host from microbial pathogens. In such cases, pathogens or consumers may be selective for specialized microbial associates that defend their host. Such associations could increase microbial diversity and the diversity of macroorganisms by allowing hosts to persist in new habitats, thus facilitating both the evolution and the retention of increased species diversity.

Host Chemistry as a Promoter of Consumer Diversity

On some tropical reefs, fishes have been reported to bite the bottom in excess of 150,000 times/m^2/day. In these areas, small herbivores such as amphipods, polychaetes, and crabs (collectively called mesograzers because of their size) would live short lives if they occupied plants that were preferred by fishes. Selection should therefore favor sedentary mesograzers that live on and eat seaweeds that are chemically repellent to fishes. Patterns supporting this notion have been documented for divergent types of mesograzers in several of the world's oceans. In the temperate Atlantic, some herbivorous amphipods and polychaetes live in tubes, which they attach to the seaweeds they consume. These mesograzers selectively live on and feed from brown algae that are chemically defended from fishes. By living in association with these seaweeds, which are seldom visited by fishes, they lower their susceptibility to predation.

The hypothesis that sedentary mesograzers minimize predation by specializing on toxic hosts has been tested more broadly using (1) a specialist Caribbean amphipod that eats, and builds a mobile domicile from, a chemically defended alga; (2) several species of crabs and sea slugs from both the Caribbean and tropical Pacific that each live on and feed from only one species of chemically defended alga; (3) nonherbivorous amphipods that live on and build homes from only one species of chemically noxious seaweed; (4) a decorator crab that minimizes predation by selectively decorating with a seaweed that is chemically repellent to fishes; and (5) contrasts between the palatability and susceptibility to predation or parasitism of specialist versus generalist herbivorous insects. In all these cases, predation on, or palatability of, the mesograzers was reduced as a consequence of their association with chemically noxious hosts. Additionally, mesograzers were generally stimulated or unaffected by plant compounds that deterred feeding by larger herbivores or predators.

There are numerous examples indicating that seaweeds or sessile invertebrates that evolve effective defenses against reef fishes may become evolutionary targets for specialized mesograzers that can escape or deter their own consumers by evolving a resistance to these compounds and living on, feeding from, and in some cases morphologically mimicking or sequestering defensive compounds from their toxic hosts. Thus, once hosts effectively deter common fish predators, these hosts can serve as valuable spatial escapes from predation for small consumers that specialize on these noxious hosts. As in many of the earlier examples, this creates spatial mosaics of selection within otherwise more uniform habitats and allows for the evolution and retention of more species within the system.

The Potential Importance of Paleo-Patterns

Plant–herbivore interactions occurring today are the result of both modern ecological forces and an evolutionary history that has been operating for many millions of years. Major changes in plants or in herbivores can select for fundamental and often cascading changes in plant–herbivore interactions. For example, in modern African communities, grazing by elephants causes repeated deforestation of savannas, prevents larger plants from replacing grasses, and facilitates retention of other herbivore species that require open-grazing habitats. It has been hypothesized that loss of elephants and other megaherbivores from Europe and the Americas at the end of the Pleistocene may explain the dramatic loss of other grazing mammals in these regions. It can be argued that elephant grazing prevented open savannas from converting to woodlands and grazing by rhinoceros and hippopotamus prevented short grasslands from converting to less nutritious tall grasslands. The loss of these large grazers would have caused mosaics of parklike woodlands and grasslands to convert to more homogeneous forests and less productive tall grasslands. The loss of the mosaic vegetation structure is predicted to have caused the loss of numerous smaller herbivores that depended on spatial patchiness and the availability of nutrient-rich vegetation. This hypothesis is consistent with the variability in extinction seen among Africa, Eurasia, and the Americas and with the absence of extinctions at previous glacial–interglacial transitions.

The mass extinction of megaherbivores at the end of the Cretaceous also coincided with a dramatic change in the structure and species composition of the plant community. Most of the dominant ferns and cycads went extinct, but angiosperms and gymnosperms flourished. The decreased herbivory following the mass extinction should have allowed vegetation to become much denser, potentially increasing competition among plants. This change is hypothesized to have lead to the evolution of larger seeds as a way of better provisioning juvenile plants. These larger seeds and the need for dispersal may have led to fruits for attracting animal dispersers. These abundant fruits and seeds are thought to have favored the evolution of small, fruit-eating birds and mammals that aided seed dispersal.

The cause–effect relationships hypothesized to have driven the paleo-changes discussed previously must be viewed with caution, given the general inability to rigorously test hypotheses associated with evolutionary time. However, it is clear that major changes in vegetational structure often coincide with extinctions of large herbivores and *vice versa*. There is also broadly accepted support for the notion that there has been an evolutionary arms race between vascular plants and their insect herbivores – the plants evolving chemical defenses against the insects, the insects responding by evolving resistance to some of these defenses and then tending to specialize on plants with those chemical signatures, and the cycle then

repeating. It is hypothesized that this diverse range of plant chemical defenses serves as a rich and heterogeneous resource gradient that has promoted the dramatic diversification of herbivorous insects.

It is tempting to assume that coincident changes in herbivores and plants are caused by reciprocal and evolved responses. However, such "co-evolved" responses may be difficult to differentiate from fortuitous preadaptations. One marine study with an especially complete fossil record provides an example. Encrusting coralline algae are highly calcified crusts that cover hard substrates and fossilize so well that grazing scars attributable to specific types of herbivores (i.e., gastropods, sea urchins, or fishes) are clearly preserved on their surfaces. These algae have left a long and abundant fossil record. Using fossilized corallines and information on present-day interactions between corallines and reef herbivores, one can convincingly argue that the evolution of herbivorous fishes (especially parrotfishes) dramatically increased herbivory on coral reefs and fundamentally changed the selective regimes affecting seaweeds on reefs and reef community organization in general. Coincident with these changes in grazing pressure, coralline crusts with numerous traits that minimized the effects of herbivorous fishes radiated dramatically while their assumed parent taxon (the solenopores, which lacked these traits) diminished in importance and eventually went extinct. From these coincident changes, it seems reasonable to assume that the escalation of grazing drove the evolution of corallines and selected for the numerous morphological traits that minimized the effects of herbivores on corallines. However, the exceptionally complete fossil record for the corallines allows this assumption to be tested. Virtually all the morphological traits that appear to be "adaptations" to grazing were present in the corallines before the evolution of macroherbivores and hundreds of millions of years before the evolution of parrotfishes. Thus, grazing fishes may well have driven the solenopores extinct and may have changed reef ecology to favor the expansion of the corallines, but the corallines were fortuitously preadapted for these changes – they did not adapt in response to them. Such examples should encourage caution in ecological interpretation of the fossil record.

Using Grazers to Manage and Restore Ecosystems

There is a long history of using grazers to maintain or restore ecosystem structure, function, or diversity. Both productivity and community composition can be affected by manipulating the duration, intensity, and timing of grazing. Herbivore manipulations have been used to increase species diversity, to increase the esthetic value of habitats, to preserve endangered species, and grazers appear to play roles in preventing invasions by exotic species, and to be useful in manipulations to restore community structure and function. Most grazers being used as management tools come from terrestrial grasslands, but this practice might also be useful in a wide variety of habitats ranging from forests and heathlands to coral reefs and lakes. In the following sections, we discuss some examples by habitat type or by critical process that occur across many community types.

Grasslands

Grazers are crucial to the maintenance of grasslands because without them there is an accumulation of plant litter that sequesters nutrients, physically limits vegetative growth, and interferes with seedling establishment. In a 10-year experiment focused on chalk grasslands in the Netherlands, sheep grazing helped rebuild species-rich grasslands that had been degraded due to agricultural use and heavy fertilization. When grazed and ungrazed plots were compared, the grazed plots had about 50% more species and the relative mass of forb was increased threefold. Studies of other grassland systems and grazers have produced similar results. For example, sheep grazing can be used to reclaim heathland from woodland by opening up vegetation, repressing growth of scrub, promoting low-growing plants, and encouraging growth of dwarf shrubs. Experiments with rabbits and bison show similar results. In both cases, when the grazers were removed from grasslands, plant diversity declined.

Although grazing increased biodiversity in the previously mentioned studies, several of these management efforts were designed to have this result, thus potentially providing a biased impression of how grazing will affect plant species richness. Under some conditions, grazers decrease biodiversity. Effects are likely to vary as a function of the evolutionary history of the plants and herbivores, the types of habitats investigated, the levels of grazing employed, and when and for how long these grazing regimes are maintained. As an example of the variable effects that grazing can have on producer diversity, a recent study reviewed 44 comparisons of plant species richness under low versus high grazing pressure in nutrient-rich versus nutrient-poor ecosystems. High rates of grazing reduced plant species richness in 100% of the nutrient-poor situations. In contrast, high rates of grazing increased plant diversity in 56% of the high-nutrient contrasts; diversity was unaffected by grazing in 36% of these contrasts and decreased by grazing in 8% of the contrasts.

Manipulation of herbivores can also be used to alter habitat traits, such as productivity, rates of nutrient cycling, and nutrient, water, and organic levels in the soil. In general, herbivores often increase organic breakdown and the mineralization of potentially limiting nutrients such as nitrogen and phosphorous. In the Serengeti, the aboveground productivity of moderately grazed plots is stimulated to about twofold greater than the productivity of ungrazed plots. In addition, ungrazed grass stands are senescent, whereas those grazed by large herbivores produce younger and more palatable shoots.

Although mowing can crudely substitute for some grazer effects, grazing will generally produce greater diversity due to spatial mosaics produced by patchy grazing (vs. even defoliation by mowing) and localized trampling. Selectivity and patchiness of grazing will also be affected by herbivore morphology and behavior. For example, teeth size and mouthpart morphology are important in determining the degree of selectively exhibited by grazers. For example, cattle are less able to graze individual plants or plant parts than sheep, goats, or horses.

Fire has also been used as a surrogate for grazing or, in conjunction with grazing, to manage plant community

structure and species composition. This is especially true in attempts to prevent exotics and woody species from altering grasslands. Although fire may be a necessary disturbance in some systems, such as tallgrass prairie, fire alone is often not sufficient for restoration of biodiversity. After 9 years, in annually burned tallgrass prairie with nitrogen-addition, species diversity was 48% lower than it was 5 years before burning started and 66% lower than in unburned plots. In contrast, mowed, annually burned, nitrogen-addition plots had more than twice the species diversity than similar unmowed plots. Therefore, mowing prevented a decrease in species diversity that would have occurred with only fire and nutrient addition. When bison grazing replaced mowing, plots that were grazed and burned had the highest species diversity of all plots. Similarly, in a Mediterranean grassland study, cattle grazing led to an increase in species diversity, whereas burning resulted in no significant difference.

Forests

Many tropical forest trees produce large fruits that either fall near the parent tree or are dispersed by birds, primates, and other large mammals to sites remote from the adult tree. Dispersal is so critical for some of these species that their seeds require gut passage through dispersal agents before they can germinate. Because of the reliance on dispersers and the heavy predation on seeds near the parent canopy, it has been suggested that seed predators and dispersers are critical for maintaining tree species diversity in tropical forests. By selectively preying on seeds that fall near parent trees, specialized seed predators may decrease the survivorship of seeds near parents, preventing the occurrence of monospecific patches of forest trees and facilitating a more even and diverse assemblage of species. Because larger animals disperse the seeds of more than half of woody plant species, they play a critical role in removing seeds from sites of high predation and in introducing new species into different patches within a forest. The hunting, poaching, and habitat modifications that have left many forests depauperate in these animal dispersers have been suggested to be a cause of lowered plant diversity in tropical forests. If this is correct, then effective management of tropical forests may require the replenishment of large-bodied seed dispersers. In addition to grazing by larger animals, plants are also attacked by a diverse set of microbial pathogens and parasites. Recent investigations have shown that soil microbes may also play important roles in preventing monospecific patches. If microbes that attack a common plant build up in soils near that plant, they can lower the probability of that species' seedlings being successful in that site and thus promote diversity by favoring unrelated plant species in the vicinity of a long-established adult of a different species.

Livestock introduced to serve as seed dispersers can partially replace native dispersers that disappeared during the Pleistocene megafaunal extinction. The feeding activities of these introduced species can increase the range of some plants that produce large, fleshy fruits. In a lowland deciduous forest in Costa Rica, introduced horses and cattle feed on the fruits of jicaro (*Crescentia alata*) and disperse their seeds. In areas with livestock, jicaro are common. In areas where livestock are absent, jicaro is relatively rare and occurs primarily in small, spatially restricted patches.

Many seed dispersers subsist exclusively on fruit for at least part of each year. This suggests that removal of seed dispersers by anthropogenic activities could have several repercussions. First, decreased seed dispersal could eventually lead to a decrease in the number of mature seed-producing parent plants. This would in turn lead to a decrease of important wildlife food resources, which could then lead to even lower numbers of seed dispersers. This situation can be especially critical if a particular tree species is pivotal in maintaining the health of dispersers during food-limiting seasons. For example, *Casearia* in tropical rain forests produce rich fruit during seasons of relatively low fruit abundance and likely supply much of the diet of local fruit-eating dispersers. In addition, this species supports frugivorous birds that are important dispersers of many other tree species during seasons when fruit is more abundant. The disappearance of animals that feed on *Casearia* fruits, leading to decreased recruitment of *Casearia*, could have far-reaching effects on the forest community. This scenario is thought to have occurred with jicaro. Frugivorous bats also feed on jicaro, and it is thought that a decrease in jicaro following the loss of Pleistocene seed dispersers led to a decrease in bats and then to a corresponding decrease in other fruit-producing plant species on which bats feed.

Lakes

Attempts to regulate the community composition of lakes have produced a mature science in which the separate and interactive effects of both physiochemical and biotic interactions are reasonably well understood and used in novel ways to manage lake ecosystems. Success in these systems has been achieved by fusing an understanding of biotic processes such as competition and predation (traditionally studied by population and community ecologists) with an understanding of the role of physiochemical processes (traditionally studied by limnologists and ecosystem ecologists). Investigations by Steven Carpenter and colleagues have been especially important in stimulating this field.

Problems of harmful algal blooms, fish kills, and general eutrophication of lakes have become increasingly apparent. Initial attempts to explain these problems, and lake productivity in general, as a function of nutrient levels revealed that nutrients could vary considerably in lakes with similar biotic communities and that nutrient levels (also known as bottom-up effects) explained only approximately 50% of the variability in lake productivity. This finding prompted lake ecologists to investigate the effects of trophic cascades (or top-down effects) in structuring these ecosystems. By understanding and manipulating both the bottom-up effects of nutrients and the top-down effects of consumers, lake ecologists have been remarkably successful at altering biotic communities and fundamental processes (e.g., productivity and nutrient cycling) occurring in lakes.

Because lake communities are structured by interactions between physiochemical conditions and biotic processes, initial attempts to manage lakes by managing nutrient input met

with only variable success. When the traditional focus on nutrients was merged with top-down manipulation of trophic structure, greater success was achieved in the control of lake phytoplankton populations and productivity. Mechanisms for this are as follows: As piscivore biomass increases, their feeding causes a decline in the biomass of small fishes that feed on herbivorous zooplankton; with the decline of these planktivorous fishes, the biomass of large herbivorous zooplankton increases and causes a decline in phytoplankton biomass and an increase in water clarity (**Figure 4**). In all these instances, productivity of each group is maximized at intermediate levels of piscivore biomass. This system of trophic cascades allows managers to affect planktivore, herbivore, or phytoplankton biomass and productivity *via* manipulation of larger fishes (which can be done using fish additions, removals, or common fisheries management practices).

Whole lake manipulations have shown that piscivorous fish populations can regulate planktivorous fish populations in both North American and European lakes. For example, in an experiment by Carpenter and Kitchell, one lake had a high bass (piscivorous fish) population and a small minnow (zooplanktivorous fish) population, whereas another had a low bass population and a high minnow population. Ninety percent of the large bass population was placed in the bass-deficient lake, and 90% of the large minnow population was placed in the minnow-deficient lake. When bass were added and minnows removed, the size of herbivorous zooplankton increased due to decreased predation. This led to decreases in chlorophyll concentration and primary production. In the opposite experiment, trophic interactions ultimately led to increase in chlorophyll concentration and primary productivity.

The interaction between trophic cascades and nutrient addition has also been addressed. In whole lake experiments in which phosphorous was added to lakes with a high or low density of piscivorous fishes, only a high density of piscivorous fishes was effective in maintaining low levels of palatable phytoplankton. The added nutrients increased the biomass of herbivores and colonial blue-green (unpalatable) algae but not the biomass of palatable algae. Grazers can control chlorophyll levels at phosphate loading significantly higher than loads that cause eutrophication. However, grazer control of undesirable blue-green algae appears to fail at phosphate loading rates lower than those used in this manipulation.

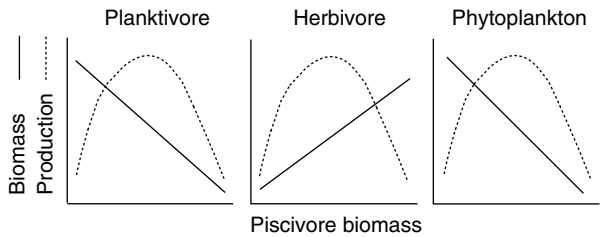


Figure 4 Effects of piscivore (= bass) biomass on biomass (solid line) and productivity (dashed line) of vertebrate zooplanktivores, large herbivores, and phytoplankton. Reproduced with permission from Carpenter et al. (1985) Interactions and lake productivity. *Bioscience* 35: 634–639 Copyright 1985 American Institute of Biological Sciences.

Coral Reefs

Feeding by large generalist herbivores, such as fishes and sea urchins, is critical for maintaining the coral-generated topographic complexity that helps maintain biodiversity on tropical coral reefs. If these herbivores are removed, reef corals are competitively excluded by seaweeds. For example, overharvesting led to severely depleted populations of herbivorous fishes on coral reefs in Jamaica. However, the effects of overfishing were not fully realized for decades because the lowered numbers of predatory and competing herbivorous fishes allowed a build-up of herbivorous sea urchins that largely compensated for the declining fish grazing. When the urchin population was severely reduced by disease, macroalgal cover increased from 4 to 92%, severely overgrowing corals and inhibiting coral recruitment. The few grazers left on the reef (small parrotfishes and surgeonfishes) were not enough to control the algal growth, which has persisted for several decades. Recent work has demonstrated that seaweeds not only suppress coral recruitment, but numerous common seaweeds also damage (or sometimes kill) corals using surface associated chemicals that harm corals on contact. Thus, after disturbances like bleaching events, coral diseases, or storms that damage corals, the remaining corals may suffer additional declines due to contact with seaweeds. On overfished reefs such damage will be extensive and corals may decline further while seaweeds proliferate; on reefs protected from fishing, fishes prevent seaweed blooms and corals tend to recover. From scenarios such as this, it is evident that a higher density and diversity of grazers is required to decrease algal cover and allow coral recovery.

Manipulating these herbivores would therefore be a logical way to manage critical biotic processes on coral reefs. Because of the open nature of marine populations (i.e., the long-distance larval dispersal typical of most reef herbivores means that populations must be managed regionally rather than locally), this is a more challenging proposition than in the lake or grassland systems discussed previously. However, there is considerable evidence that degrading reefs recover, or degrade more slowly, when they are protected from fishing than when fishing is allowed to continue. The increased health of the protected reefs is correlated with an increase in herbivorous fishes. Recent experiments demonstrate that it is not only the abundance of herbivorous fishes that is important, but also the diversity. Different fishes consume different seaweeds and certain mixes of herbivore biodiversity are critical for preventing seaweeds from suppressing corals. On reefs in the Florida Keys, long-term enclosure of different herbivores on reefs demonstrated that herbivore diversity, independent of density and mass, was critical for suppressing seaweeds and facilitating coral survivorship and growth. Managing reefs to facilitate targeted herbivores that are especially important in removing reef macrophytes is a promising tool for aiding coral reef recovery.

Preventing Invasions by Exotic Plants (Data from Many Ecosystems)

Invasive species are a major threat to native biodiversity. Despite the widely held notion that plants become invasive

in new habitats because they have escaped their native herbivores, a recent meta-analysis of how native and nonnative herbivores in numerous terrestrial and aquatic systems affected native versus nonnative plants found just the opposite. Native herbivores selectively attacked exotic plants, and exotic herbivores selectively attacked native plants. Exotic plants seemed to be succeeding because they were following rather than escaping their native herbivores. The documented patterns suggested that as humans invaded new continents, they killed off the native herbivores (thus preventing them from suppressing invasions by nonnative plants) and replaced them with nonnative, domestic herbivores (cattle, goats, sheep, etc.) that selectively attacked native plants. By removing native competitors, these exotic herbivores cleared the way for exotic plants from their own native region that were already evolved to live with these herbivores. These findings suggest that maintaining intact guilds of native, generalist herbivores may be effective in suppressing plant invasions and thus maintaining native plant biodiversity.

Summary

Coexistence of potential competitors within diverse plant communities has often been explained as a result of fine-scale resource partitioning. This is assumed to have resulted from an evolutionary history of competitive encounters and is usually inferred from shifts in abundance that occurs between habitats with different physical characteristics. Both marine and terrestrial plants show predictable patterns of distribution that can be interpreted as this type of habitat partitioning. However, manipulative experiments in marine systems, and to a lesser extent some terrestrial communities, show that many plants are habitat generalists and both between- and within-habitat patterns of distribution are often controlled by herbivores, or other natural enemies. Spatial patterns in herbivore effectiveness create a mosaic of habitats that differ in the degree to which they favor poorly defended but competitively superior species versus well-defended but competitively inferior ones. If herbivory is decreased, the mosaic nature of the habitat is reduced and many well-defended species are excluded by competition. This pattern occurs across a large range of spatial scales and can explain between-habitat differences in plant communities that occur hundreds of meters apart, within-habitat differences that occur only a few meters apart, and microhabitat differences that occur on a scale of centimeters or millimeters. A portion of the microhabitat pattern is created by the plants. For example, consumption of palatable plants can be significantly reduced when they grow on, or are intermixed with, unpalatable plants. These associational refuges increase species richness by allowing palatable forms to invade grazed areas after unpalatable forms establish and create microsites of reduced herbivory. At all of the scales examined, herbivore activities that promote spatial mosaics of herbivore effectiveness, or that differentially affect dominant plants, help to maintain higher biodiversity.

Foraging by herbivores is often constrained by factors (e.g., predation) that do not directly affect the physiological performance of primary producers. For plants, this differential use of space by herbivores often produces a spatial mosaic of selective regimes in an area that would otherwise be treated as one uniform habitat; this interaction commonly results in elevated species richness. The strong effects of herbivores on plants, and the strong effects of plants on community structure in general, mean that manipulations of plant–herbivore interactions can often be used as an effective management tool for promoting the maintenance of biodiversity or ecosystem function.

See also: Cattle, Sheep, and Goats, Ecological Role of. Coevolution. Corals and Coral Reefs. Disturbance, Mechanisms of. Plant–Animal Interactions. Resource Partitioning. Seagrasses. Species Coexistence

Further Reading

Burkepile DE and Hay ME (2008) Herbivore species richness and feeding complementarity affect community structure and function: The case for Caribbean reefs. *Proceedings of the National Academy of Sciences, USA* 105: 16201–16206.

Carpenter SR, Cole JJ, Kitchell JF, and Pace ML (2010) Trophic cascades in lakes: Lessons and prospects. In: Terborgh J and Estes J (eds.) *Trophic Cascades*, pp. 55–69. Washington, DC: Island Press.

Diaz SS, Lavorel S, McIntyre V, *et al.* (2007) Plant responses to grazing – a global synthesis. *Global Change Biology* 13: 313–341.

Fuhlendorf SD and Engle DM (2001) Restoring heterogeneity on rangelands: Ecosystem management based on evolutionary grazing patterns. *Bioscience* 51: 625–632.

Gervasio P, Paruelo JM, Oesterheld M, and Jobbagy EG (2010) Pathways of grazing effects on soil organic carbon and nitrogen. *Rangeland Ecology & Management* 63: 109–119.

Hughes TP, Graham NAJ, Jackson JBC, Mumby PJ, and Steneck RS (2010) Rising to the challenge of sustaining coral reef resilience. *Trends in Ecology and Evolution* 25: 633–642.

Lewinsohn TM, Novotny V, and Basset Y (2005) Insects on plants: Diversity of herbivore assemblages revisited. *Annual Reivew of Ecology, Evolution, and Systematics* 36: 597–620.

Mumby PJ and Steneck RS (2008) Coral reef management and conservation in light of rapidly evolving ecological paradigms. *Trends in Ecology and Evolution* 23: 555–563.

Nunez-Farfan J, Fornoni J, and Valverde PL (2007) The evolution of resistance and tolerance to herbivores. *Annual Review of Ecology, Evolution, and Systematics* 38: 541–566.

Ohgushi T (2005) Indirect interaction webs: Herbivore-induced effects through trait change in plants. *Annual Review of Ecology, Evolution, and Systematics* 36: 81–105.

Parker JD, Burkepile DE, and Hay ME (2006) Opposing effects of native vs. exotic herbivores on plant invasions. *Science* 311: 1459–1461.

Provenza FD, Villalba JJ, Haskell J, MacAdam JW, Griggs TC, and Wiedmeier RD (2007) The value to herbivores of plant physical and chemical diversity in time and space. *Crop. Science* 47: 382–398.

Schmitz OJ (2008) Herbivory from Individuals to Ecosystems. *Annual Review of Ecology, Evolution, and Systematics* 39: 133–152.

Smetacek V, Assmy P, and Henjes J (2004) The role of grazing in structuring Southern Ocean pelagic ecosystems and biogeochemical cycles. *Antarctic Science* 16: 541–558.

Greenhouse Effect

Jennifer A Dunne, Santa Fe Institute, Santa Fe, NM, USA
Stacy C Jackson and John Harte, University of California, Berkeley, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Jennifer A. Dunne, John Harte volume 3, pp 277–293, © 2001, Elsevier Inc.

Glossary

Aerosols Microscopic airborne particles.
Albedo The fraction of light hitting a surface that is reflected.
Anthropogenic Resulting from human activities.
Climate sensitivity Long-term change in global mean surface temperature following a doubling of carbon dioxide (CO_2) concentration in the atmosphere versus its preindustrial (1750) level.
Earth energy balance Difference (if any) between incoming solar energy and outgoing terrestrial radiation for Earth as a whole.
Feedback Change in a system component that triggers effects that eventually change the original component again. Feedbacks can be positive (self-reinforcing) or negative (self-dampening).
Greenhouse gases Atmospheric gases that can absorb and re-radiate infrared radiation.
Litter Dead organic matter, such as leaves and branches, that is deposited as the top soil layer.
Radiative forcing Measure used to express and compare the potential of climate system changes (vs. preindustrial) to perturb the Earth's energy balance; reported in watts per square meter ($W\,m^{-2}$). A positive radiative forcing tends to warm the Earth's surface and a negative radiative forcing tends to cool the surface.
Sink A physical reservoir of a compound that can take up additional quantities of this compound from the air and thus reduce its atmospheric concentration; for example, trees are a sink for CO_2.

Introduction

Earth's climate, from daily average weather events to glacial and interglacial cycles, is driven by the amount of radiation received from the sun and how that radiation is distributed throughout the global Earth atmosphere system. The atmospheric greenhouse effect acts as an important factor in establishing a temperature that is hospitable for life. The basic mechanism is simple and was first detailed by the Swedish physicist Svante August Arrhenius in 1896. Light from the sun follows one of three pathways: absorption by the atmosphere (25%), reflection back into space (30%), or absorption by the earth's surface (45%). The light absorbed by earth's surface and by the atmosphere is converted from energy in the form of shortwave visible light to energy in the form of longwave infrared radiation (i.e., heat). Several gases in the atmosphere, referred to as "greenhouse gases," absorb some of the heat and reradiate it back toward the surface, increasing surface temperature. Without this naturally occurring greenhouse effect, Earth's average surface temperature would be approximately $-18\,^{\circ}C$, about $33\,^{\circ}C$ colder than it is today.

It should be noted that while the atmospheric greenhouse and actual physical greenhouses both trap heat, they do so via different mechanisms. A physical greenhouse traps air inside the building so that its heat cannot escape by convection; the atmospheric greenhouse absorbs the heat in gas molecules and re-radiates it back toward the earth's surface.

Enhanced Greenhouse Effect

The most important naturally occurring greenhouse gases are water vapor, carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and ozone (O_3). Although water vapor plays the biggest role in generating the natural greenhouse effect, anthropogenic emission of the other gases, along with artificially produced halocarbons, are most important in generating an enhanced greenhouse effect. Detailed instrument data show that concentrations of these gases have been increasing since preindustrial times (1750) (**Table 1**; **Figure 1**), particularly in recent decades, largely due to human, industrial, agricultural, and urbanization activities. As the concentrations of greenhouse gases increase in the atmosphere, they continue to trap and re-radiate more heat, resulting in rising surface temperature and other climatic changes.

Increases in CO_2 account for about half of the current direct positive radiative forcing due to anthropogenic emissions (**Figure 2**). The atmospheric CO_2 concentration has increased by 35% since preindustrial times, as a result of emissions from fossil fuel combustion, land conversion, and cement production, and is continuing to increase by 0.5% per year (**Figure 2**). If future emissions of CO_2 were maintained at 2005 levels, CO_2's atmospheric concentration would be roughly double the preindustrial level by the end of twenty-first century. Concentrations of other greenhouse gases, particularly methane and nitrous oxide, also continue to slowly rise. Greenhouse gases tend to remain in the atmosphere for

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00068-X

Table 1 A sample of greenhouse gases affected by human activities

	CO_2	CH_4	N_2O	CFC-12	HFC-134a
Preindustrial concentration	~280 ppmv[a]	~700 ppbv[b]	~270 ppbv	0	0
2005 concentration	379 ppmv	1774 ppbv	319 ppbv	538 ppbv	35 pptv
1998–2005 change	1.9 ppmv y^{-1}	1.6 ppbv y^{-1}	0.7 ppbv y^{-1}	0.6 pptv y^{-1}	3.9 pptv y^{-1}
1998–2005% change	0.5% y^{-1}	0.1% y^{-1}	0.2% y^{-1}	0.1% y^{-1}	23.5% y^{-1}
Atmospheric lifetime[c] (years)	[d]	12	114	100	14
Direct radiative forcing (W m^{-2})	1.66	0.48	0.16	0.17	0.01

[a]CO_2's concentration ranged from 275–285 ppmv in the preindustrial era (AD 1000–1750).
[b]CH_4's concentration ranged from 400–700 ppbv in the 650,000 years preceding the industrial era.
[c]Atmospheric lifetime is the time after emission for the atmospheric concentration of a gas to be reduced by 63% (to 37%, or e^{-1}, of its original concentration); after two lifetimes, the concentration will be reduced by another 63% (to 23%); and so on for three, four, etc., lifetimes.
[d]CO_2 does not follow the exponential decay of other gases because it is taken up by multiple sinks. The first lifetime (time to 37%) is about 80 years; subsequent lifetimes are longer.
Abbreviations: ppmv, parts per million by volume; ppbv, parts per billion by volume; pptv, parts per trillion by volume.
Source: Adapted from Forster P, Ramaswamy V, Artaxo T, *et al.* (2007) Changes in atmospheric constituents and in radiative forcing. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK and New York, NY, USA: Cambridge University Press.

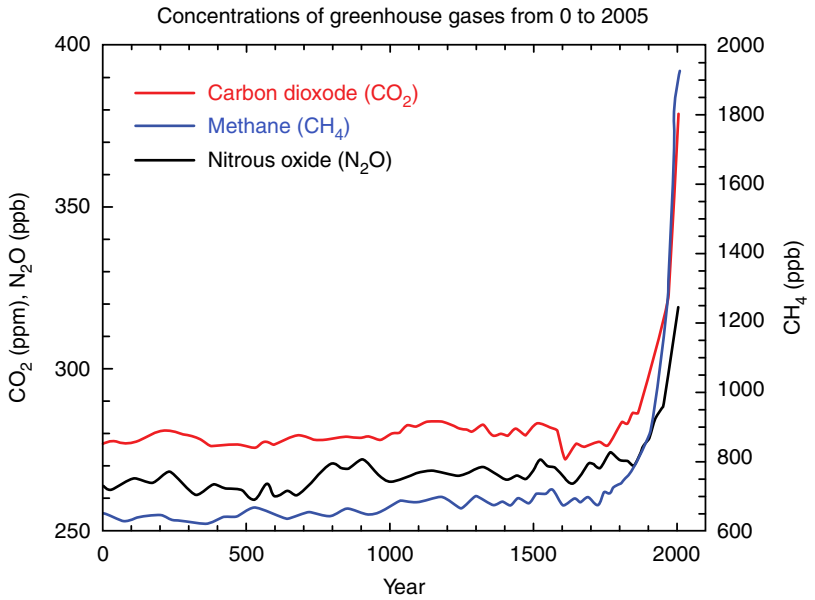


Figure 1 CO_2, CH_4, and N_2O concentrations over the past 2000 years. Reprinted from FAQ 2.1, Figure 1 in Forster PM, Ramaswamy V, Artaxo P, *et al.* (2007) Changes in atmospheric constituents and in radiative forcing. In: Solomon S, Qin D, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press.

many years (Table 1) and continue to affect the climate long after initial emissions.

In addition to greenhouse gases, fine particles suspended in the air called aerosols can also affect the climate. Aerosols can be produced by natural sources such as dust and sea spray, and can also be produced by human activities such as incomplete combustion of fossil fuels and biomass. Aerosols can alter the climate by changing atmospheric albedo. These fine particles can both reflect and absorb solar radiation and alter cloud properties to produce a net cooling effect. The overall effect is usually one of cooling, although atmospheric soot over snow or ice can warm regionally. The current estimate of the combined effect of aerosols is -1.2 W m^{-2} compared to 3 W m^{-2} due to greenhouse gases (Figure 1). Aerosols are very short-lived in the atmosphere (days to weeks) and respond rapidly to changes in emissions.

Climatic Consequences: Global Warming

Past Climate Change

Scientists use deep ice cores drilled from glaciers in Antarctica, Greenland, and South America to examine ancient climate and atmospheric trends over the last several hundred thousand years. The data from these ice cores reveal a strong correlation between temperature and CO_2 concentrations (Figure 3). The historical swings between Ice Ages and warm periods are triggered by periodic, predictable changes in the tilt of Earth's axis, the direction of its axis relative to the sun, and the shape of its orbit. The initial trigger created by these changes produces biological feedbacks that will be discussed in the Section, Feedbacks. These feedbacks change the amount of CO_2 in the atmosphere, and the resulting greenhouse effect amplifies the initial warming or cooling.

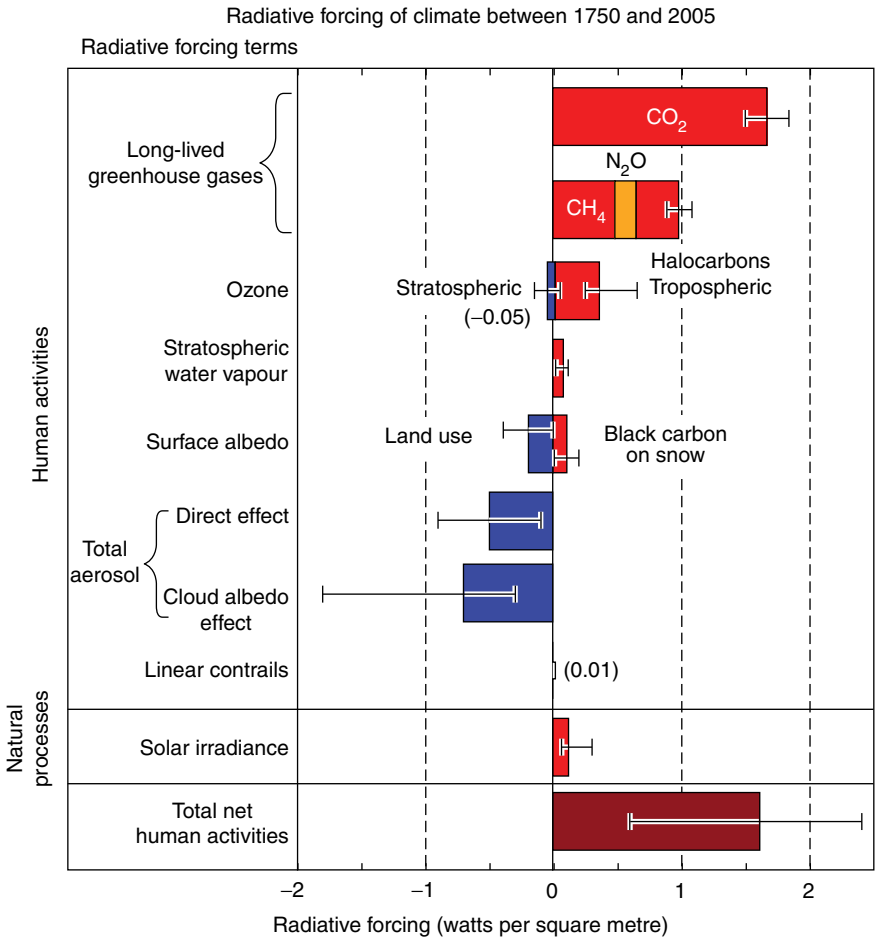


Figure 2 Estimates of the globally and annually averaged radiative forcing due to anthropogenic changes in concentrations of greenhouse gases and aerosols and natural changes in solar output from preindustrial times to 2005. The length of the rectangular bars indicates a midrange estimate of forcing, and the error bars show an estimate of the uncertainty range. The level of scientific understanding is rated by the IPCC as "high" for the long-lived greenhouse gases, "med" for ozone, "med–low" for surface albedo and aerosol direct effect, and "low" for stratospheric water vapor, aerosol cloud albedo effect, linear contrails, and solar irradiance. Reprinted from FAQ 2.1, Figure 2 in Forster PM, Ramaswamy V, Artaxo P, *et al.* (2007) Changes in atmospheric constituents and in radiative forcing. In: Solomon S, Qin D, *et al.* (eds.) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. Caption information from Figure 2.20 in the same report.

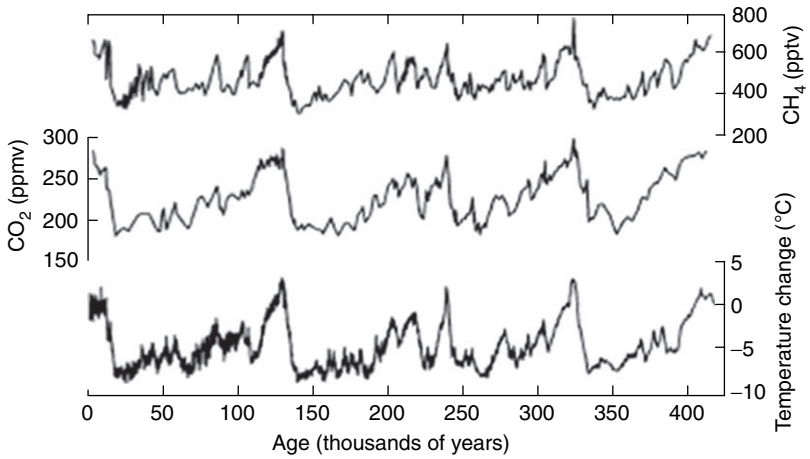


Figure 3 Variation during the last 450,000 years of atmospheric CO_2 and CH_4 concentrations and atmospheric temperature, as derived from the Vostok, Antarctica ice core. Note that the temperature change is that of an average over the higher latitudes of the southern hemisphere (roughly 45 to 90° S) and that polar regions experience larger temperature swings than the earth overall. Reprinted with permission from Figure 2.31 of Wallace JB and Hobbs PV (2006) *Atmospheric Science: An Introductory Survey*. Amsterdam, The Netherlands: Academic Press.

Recent Climate Change

Observational data provides clear evidence of human-induced global warming. Global average surface temperature trends reveal a 0.7 °C temperature increase during the past hundred years, with the 10 hottest years on record between 1880 and 2010 occurring between 1998 and 2010.

How do the recent temperature increases compare to natural climate variability? The key drivers of natural climate variability are orbital patterns, solar cycles, and volcanic eruptions. The first two are predictable and are included in climate model projections. Volcanic eruptions are included in retrospective models. **Figure 4** illustrates the amount of temperature change attributable to natural variation (in blue) relative to the total change observed (black line) and the total change predicted (in pink). This diagram illustrates that most of the change experienced during the past century is attributable to human activity and also shows that models can do a good job of replicating past climate system behavior. This ability of models to faithfully reproduce the past provides evidence that they have some predictive power for the future.

Future Climate Change

What will the future bring? Future climate will depend in part on societal decisions about future emissions and in part on emissions that we have already generated. The climate system includes two important time lags. First, long-lived emissions like CO_2 stay in the atmosphere for decades or centuries after the initial release. Second, Earth's heat builds up gradually over many years, creating a slow rise in Earth's temperature. These two factors explain why the emissions produced today commit the planet to a warming trajectory that extends decades into the future.

General circulation models (GCMs) integrate the physics of radiative forcing, ocean dynamics, and other complex Earth–atmosphere processes in order to simulate and predict the climate trends. **Figure 5** illustrates the global average temperature trajectory for the period 1900–2100 under several different scenarios for anthropogenic emissions. The orange line illustrates the slight temperature increase that would result if atmospheric concentrations were held constant (i.e., emissions were dramatically reduced to a replacement-rate level). The other lines illustrate scenarios with higher levels of emissions. Note that the graph shows only the global average, but temperature change is greater at the poles.

As of 2010, society is on a path that is in line with the upper emissions scenarios. Temperatures are projected to rise by approximately 1.5 °C versus preindustrial by 2030 and by 2 °C versus preindustrial by 2050. The threshold for "dangerous anthropogenic interference" (DAI) is considered to be 2 °C. This threshold has international political significance: the Kyoto Protocol (1997), the Copenhagen Accord (2009), and the Cancun Agreements (2010) are agreements by national governments to take the necessary steps to avoid crossing the threshold of DAI.

The impacts of climate change include effects on agricultural productivity, availability of freshwater, spread of disease, extreme weather events, alteration of ecosystems, and loss of biodiversity (*see* Climate Change and Biodiversity). Climate models point to many climatic impacts of increases in greenhouse gases and aerosols that will be of particular importance to ecosystems and biodiversity. First, surface warming will not be uniform, with greater warming at the poles than at the equator. Winter temperatures are expected to increase more than summer temperatures, and night-time more than daytime temperatures. The incidence of record-breaking hot days will tend to increase in the summer, and fewer frost days are likely to occur in the winter. Second, patterns of precipitation and soil moisture will change. Warming is predicted to increase evaporation and global mean precipitation, and may increase cloud cover. Geographic distribution of precipitation is expected to shift in patterns that are not yet predictable. High and middle latitudes and elevations will experience more winter precipitation falling as rain, earlier

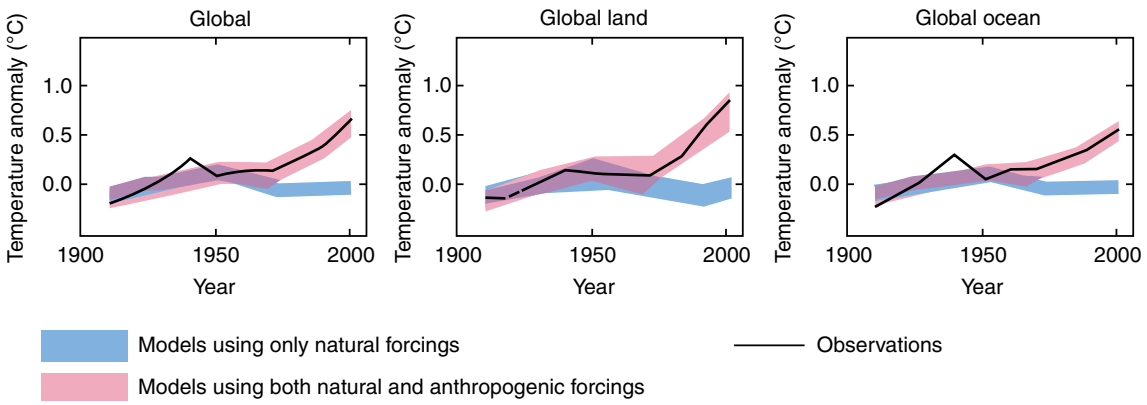


Figure 4 Surface temperature changes caused by natural earth system changes (solar and volcanic activity) (blue) versus temperature changes caused by the combination of anthropogenic and natural changes (pink) versus observed temperature changes (black line). Bands represent 5–95% confidence intervals. Results are taken from five climate models for natural forcings and 14 climate models for combined forcings. Observations are decadal averages. Reprinted from Figure SPM.4 in IPCC (2007) Summary for Policymakers. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press.

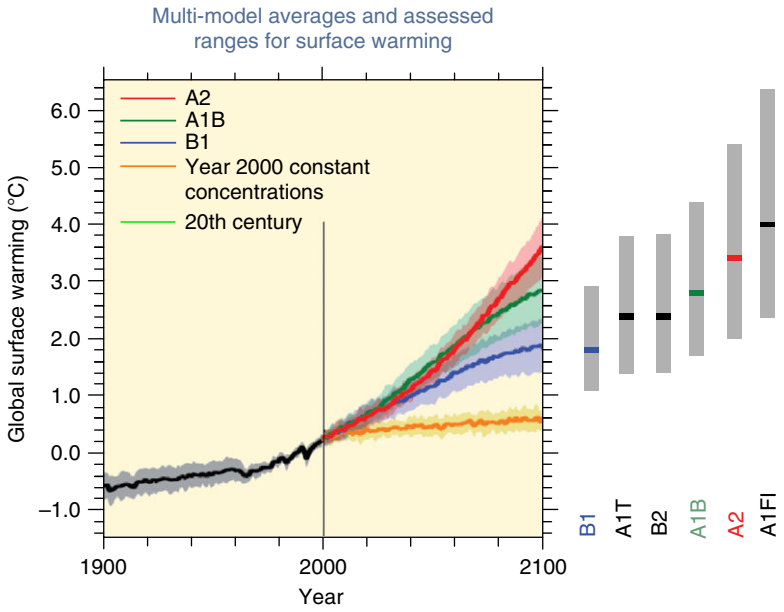


Figure 5 Projections for global surface temperature under a wide range of possible twenty-first century emission scenarios. Shading on graph indicates +/− one standard deviation from individual model results. Bars on the right indicate the likely (66% confidence) range of year 2100 outcomes based on the same models used in the graph, plus independent models and observational constraints. Reprinted from Figure SPM.5 in IPCC (2007) Summary for Policymakers. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press.

snowmelt, and reductions of summer soil moisture in non-coastal areas. Third, the frequency and intensity of extreme weather and disturbance events (e.g., drought, deluge, summer heat waves, hurricanes, and fires) are expected to increase. Fourth, average sea level is expected to rise 0.2–2.0 m during this century from a combination of thermal expansion, glacier melt, and ice sheet melt.

Some of these trends have already been observed. For example, more warming has occurred near the poles (2 × versus the global average), hot days and heat waves have become more frequent, cold days have become less frequent, droughts have increased in length and intensity, snowpack is decreasing and snowmelt is occurring earlier at high latitudes and elevations, glaciers are retreating, and sea levels are rising (sea level rose ~17 cm during the twentieth century, mostly attributable to anthropogenic climate change).

While some climate changes follow a smooth, predictable trajectory (e.g., global average temperature), other changes are more abrupt. Tipping points occur when a major shift in a tipping element results from a small incremental change in the climate. Examples of potential tipping elements include loss of Arctic summer (September) sea ice, a shift in the monsoons (India, Sahara/Sahel, West Africa) or El Nino patterns, a shift in the Atlantic Ocean circulation, loss of the Greenland or West Antarctic ice sheets, or loss of the Amazon rainforest (due to precipitation shifts). Any one of these changes would have large-scale effects on the rest of the climate system *via* feedback loops (discussed below) and on biodiversity. While scientists have some broad estimates of the temperatures at which these tipping points could occur, the exact tipping points and associated time lags are highly uncertain.

Feedbacks

Climate change can trigger a variety of responses to biotic and abiotic processes, which results in changes in the flow of energy and greenhouse gases between the surface and the atmosphere. As a result of these responses, climate change resulting directly from anthropogenic emissions of greenhouse gases triggers further climate change. These feedbacks are reviewed below.

Geophysical Feedbacks

General circulation models include not only mechanisms underlying direct greenhouse gas and aerosol radiative forcing, but also mechanisms underlying the three large geophysical feedback processes: water vapor, snow/ice albedo, and cloud cover. Because the capacity of the atmosphere to hold water vapor increases as it warms and water vapor acts as a greenhouse gas to further increase temperature, a positive feedback to the climate is created that amplifies warming. The snow/ice albedo effect is also a positive feedback – as warmer temperatures melt highly reflective snow and ice at the poles and high elevations, thus lowering surface albedo, Earth absorbs more sunlight, which augments warming.

Cloud formation adds much of the uncertainty to GCM estimates. If clouds form over highly reflective surfaces (e.g., ice and snow), they reduce the planet's albedo. If they form over low-reflective surfaces (e.g., oceans and forests), they increase the planet's albedo. In addition to uncertainty over where clouds will form, their longevity (before raining out),

height, and shape are also highly uncertain. In short, it is unclear whether there will be more clouds and if so, whether they will have characteristics that create warming or cooling. Currently, model estimates of the cloud feedback effect range from a slight negative feedback to a large positive feedback.

Biogeochemical Feedbacks

Biogeochemical feedbacks are those in which the ecosystem responses to climate changes influence the atmospheric concentrations of greenhouse gases and aerosols or the planet's albedo, which in turn influence the climate. Examples of these feedbacks include: higher rates of decay of organic matter at higher temperatures (increases CO_2), release of methane from melting permafrost (increases CH_4), northward shift of shrubs and boreal forest due to amenable growing conditions (reduces albedo), and increased dust storms due to droughts (increase dust aerosols). Over the past few years, GCMs have been expanded to incorporate many biogeochemical feedback effects and their overall effect is to increase the rate of warming (a positive feedback effect). Many biogeochemical feedbacks due to climate change also interact with other anthropogenic stresses, such as deforestation and pollution, to further influence climate change effects at local, regional, and global scales.

Marine Feedbacks

Climate change is likely to drive many complex marine-based feedback processes, of which only a few are mentioned here. With regard to geochemistry, warm water holds less dissolved CO_2 than cooler water; thus the warming of ocean surface waters expected under climate change will result in a positive feedback since a warmer ocean will release CO_2 to the atmosphere. Because oceans hold approximately 50 times as much carbon in the form of dissolved CO_2 and bicarbonate $\left(HCO_3^-\right)$ as there is CO_2 in the atmosphere, this effect could add 1° or 2° C to the predicted equilibrium warming. If the additional surface water warmth travels down the water column, methane could be released from temperature- and pressure-sensitive hydrates that are present in some ocean floor sediments. This would also result in a positive feedback since the released methane, a greenhouse gas, would cause more warming.

With regard to biology, more than a third of annual global primary production occurs in ocean surface waters, primarily by microscopic single-celled organisms called phytoplankton, which form the base of the oceanic food web. Through the process of photosynthesis, CO_2 is fixed by the phytoplankton and thus transferred from the atmosphere to ocean surface waters. Some of this fixed carbon, in the form of dead bodies and fecal matter of phytoplankton and other organisms, sinks into deep ocean layers and sediments and is sequestered there, where it can no longer be exchanged with the atmosphere on short timescales. This process is referred to as the "biological carbon pump" and has been important in maintaining a level of CO_2 in the atmosphere that is currently about 40% lower than it would be in the absence of marine organisms.

Terrestrial Feedbacks

Three to five times as much carbon is stored in terrestrial vegetation and soils than is stored in the atmosphere. Through photosynthesis and respiration, more than one-eighth of atmospheric CO_2 is exchanged each year with terrestrial ecosystems. Changes to terrestrial–atmospheric carbon cycling thus have the potential to produce significant feedbacks to climate change. The feedback pathways for carbon in terrestrial ecosystems are complex, representing both positive and negative feedbacks. Perhaps the most well-known potential carbon cycle feedback is the "CO_2 fertilization effect," which refers to the stimulation of photosynthesis by increased levels of CO_2, which in turn can result in increased plant growth and greater storage of carbon in vegetation. This effect is limited to situations in which water and other nutrients are available in sufficient supply so that CO_2 is the limiting growth factor.

Climatic changes that will accompany higher concentrations of CO_2 make it even more complicated to predict the net effect of climate change on the storage of carbon in vegetation and soil versus the atmosphere. For example, under climate change, changes in water availability and temperature will reduce the CO_2 uptake and growth of some plants while favoring others. Resulting changes in plant community composition can alter the quantity and quality of litter that enters the soil, which can lead to changes in soil carbon storage. Soil microorganisms will not only respond to changes in litter inputs, but will also be directly affected by changes in climate. Microbes and fungi tend to respire more CO_2 to the atmosphere as temperatures increase. However, rates of respiration depend on levels of soil moisture, and different extremes of water availability (both too much and too little) will tend to decrease respiration. The effects of climate change on microorganisms will also alter fluxes of other greenhouse gases such as methane in wetlands (e.g., northern peatlands, which store large amounts of carbon and may be a source of strong positive feedback to warming) and nitrous oxide in moist tropical soils.

In addition to the more direct effects of changes in temperature and moisture on the terrestrial carbon cycle, indirect effects such as the alteration of fire regimes due to climate change may produce significant feedbacks. In general, predicted increases in fire frequency for many ecosystems as a result of climate change will result in reductions of terrestrial carbon storage, a positive feedback. Those same fires, however, will in many cases increase albedo (by reducing dark forest), a negative feedback. The net effect of fire can be a positive or negative feedback in different locations, and can change over time as the plant community composition changes postfire.

Another set of climate change feedbacks related to changes in albedo may result from climate-induced shifts in land cover and vegetation. The drying of soils and increased desertification expected from climate change will add dust to the atmosphere, like anthropogenic aerosols, that can reduce warming through increases in atmospheric albedo. Surface albedo is also expected to change as the boundaries of biomes shift, since different vegetation types can have different reflectivity. For example, the predicted northward expansion of boreal forest into tundra could decrease surface albedo, resulting in increased surface warming. This mechanism may have acted as a strong positive feedback 6000 years ago when

an initial warming at high latitudes as a result of orbital variations appears to have doubled in magnitude owing to changes in surface albedo from boreal forest expansion.

Climate Change and Biodiversity

Introduction

Currently, the largest reductions in biodiversity result from massive deforestation in the tropics, in conjunction with other sources of worldwide habitat destruction. Even as the razing of vast tracts of tropical forests is expected to lead to immediate direct losses of hundreds to many thousands of species per year, the carbon that is released to the atmosphere through deforestation is amplifying the anthropogenic greenhouse effect. Climate change could lead to losses in biodiversity over the next several hundred years that are similar to or greater in magnitude than losses from direct habitat destruction. Dramatic changes in global climate have the potential to disrupt every ecosystem on Earth, leading to a pervasive trend of biodiversity loss due to climate-related habitat alteration, reorganization, and destruction.

Why does anthropogenic climate change present such a threat to biodiversity? There are two main reasons. First, the rate and magnitude of climate change expected over the next several decades to centuries are greater than any changes that current organisms have experienced. Over the last 18,000 years, starting during the last full glacial period and continuing through the current interglacial period that began about 10,000 years ago, average global surface temperature has gradually increased by about 5 ± 1 °C. If we assume conservatively that climate change will increase mean temperature by 5 °C over the next 200 years, this represents a 90-fold increase over the recent natural rate of change. In terms of magnitude, a 3 °C increase would result in the warmest world in 100,000 years, and a 4 °C increase would result in the warmest world in 40 million or more years. Second, climate change will interact synergistically with other anthropogenic stresses such as habitat destruction, pollution, ozone depletion, and alien species introduction to reduce biodiversity by more than just the sum of losses that would occur if each factor occurred independently.

Scientists' ability to specifically predict how biodiversity will be affected by climate change is constrained by large uncertainties associated with local and regional climate change predictions. Much of the following discussion is based not so much on how scientists specifically think biodiversity will change under climate change, but on generic ways in which biota will be affected by climate change that can lead to changes in biodiversity. The term "biota" is used as shorthand to refer collectively to individual organisms, groups of the same types of organisms (populations, species, functional types), and ecological complexes of multiple populations and species (communities, ecosystems). This section summarizes how scientists study the effects of climate change on biota, the types of responses biota can have to climate change, the kinds of biota likely to be harmed by and to benefit from climate change, and evidence for biotic responses to current anthropogenic climate change.

How Scientists Study the Effects of Climate Change on Biota

Species Distribution Models

Climate–vegetation classification systems are types of static models that are based on the hypothesis that climate patterns are the primary determinant of the broad-scale distribution of vegetation types. In 1947, Holdridge developed a "life-zone" concept that used three variables based on temperature and precipitation to predict under what climates, 20 vegetation types should occur. Later researchers refined and added detail to this basic concept by using a wider variety of species and bioclimatic variables that explicitly incorporate drought stress and seasonality (**Table 2**). These types of static models, referred to as "climate envelope models" can be used for single species or for life zones. They are constructed from comparisons of the current locations of species of vegetation or animals and climate and can then be used in conjunction with maps of simulated future climate to predict shifts in distribution of species due to climate change. The Holdridge life-zone classification has been used to predict the conversion of much of today's boreal forest into temperate deciduous forest. While climate envelope models provide a way to look at the potential global-scale shifts in the distribution of plants or animals under climate change, they are limited in their applicability by a number of factors.

- Novel climates, consisting of combinations of, say, temperature and rainfall that are not found in today's climate, may arise in the future.
- Nonclimatic factors, such as soil type, can limit the ability of vegetation to track shifting climate.
- Species interactions, such as those between plants and pollinators, are often critical to the ability of species to survive; if a plant but not its pollinators can disperse to a new geographic location with suitable climate, it will not be able to establish there.
- Climate envelopes are often constructed from locations of adult plants but the climate conditions may be limiting germination or seedling survival, not adult plants.
- Climate envelopes generally assume that all individuals in a species share the same preferences for climate, whereas populations within a species found across a range of climates may be adapted to local climate.

Table 2 Bioclimatic variables used by Box to predict distribution limits of plant types

T_{max}	Mean temperature of the warmest month (°C)
T_{min}	Mean temperature of the coldest month (°C)
D_T	Range between T_{max} and T_{min} (°C)
P	Mean total annual precipitation (mm)
P_{max}	Mean total precipitation of the wettest month (mm)
P_{min}	Mean total precipitation of the driest month (mm)
P_{Tmax}	Mean total precipitation of the warmest month (mm)
MI	Moisture index: the ratio of P to annual potential evapotranspiration

Source: From 1981, Macroclimate and Plant Form, The Hague: Junk Publishers.

- Species face many threats besides climate change, including habitat destruction, pollution such as acid rain, invasion of exotic species, and hunting or fishing pressure, and these threats can act synergistically with climate.

Simulation models, the other major modeling approach, incorporate and link multiple factors such as life-history traits, physiological constraints, biotic interactions, resource availability, and climate in order to predict dynamic changes in biota over the course of time. By using characteristics of individuals, populations, species, or functional types in conjunction with predicted changes in climate, modelers can simulate continuous vegetation responses to climate change at scales that vary from changes within a stand of plants to changes at landscape, regional, and global levels. Vegetation models can be used in turn to look at the changes in habitat and dynamics of other types of organisms. Simulation models not only attempt to predict patterns of change, but also provide means to explore how and why such change might occur. However, these analyses are often constrained by data requirements. Detailed and accurate information may not be available about the biota and environment of interest, and as simulations are scaled up to look at global impacts, specificity is lost. No matter how well models appear to fit observed dynamics, some factors not included in the model may be important for future dynamics, or the model may fit observed dynamics for the wrong reasons, rendering model predictions inaccurate.

Paleobiology

Paleobiologists combine historical climate and biological data to reconstruct past relationships between changes in climate and species distributions. Paleobiological data sources range from ice and soil cores to plant and animal fossils to tree rings and slow-growing corals. Such studies reveal the long-term and integrated direct and indirect effects of past climatic and atmospheric change on past biota and thus facilitate projections about how present-day biota might respond in the long term to anthropogenic climate change. Many crucial insights about previous and potential effects of climate change have emerged from this type of work (*see* Impacts of Climate Change on Biota). However, the generality of these types of studies is limited by the lack of good fossil records for many organisms and strict climatic parallels in the past to both the rate and magnitude of anthropogenic climate change, and by differences in the biology and geology of ancient times compared to the present. In addition, the interactions of climate change with other anthropogenic stresses such as pollution, development, agriculture, and deforestation are novel and not represented in the paleorecord.

Natural Climate Variability

Whereas paleobiologists look to the ancient past for insight into the future, many field biologists interested in climate change look to the present to examine how biota are regulated by and respond to natural climate variability of a magnitude similar to that expected from climate change. Two approaches are used in this type of research. First, scientists can conduct "space-for-time" analyses along elevational or latitudinal gradients. This approach suggests that the effects of climate change over time on particular biota may be represented by current differences between the biota of interest and the same type of biota found at warmer, lower elevations or latitudes. Second, researchers who conduct multiyear studies at particular sites can monitor the response of biota to the natural interannual variability of climate. In particular, biotic responses to climate in more "normal" years can be compared to biotic responses to very warm years, droughts, early snowmelts, and other climatic events that fit predicted changes due to climate change. In some cases, researchers have access to records of climate and biota from several decades ago and can compare them to current records for the same sites.

This type of research has the advantage of actually studying current biota in the field in relation to climate, but it also has several disadvantages. For example, from one year to the next and from one place to another, sites vary not only in their climate but also in many other factors (e.g., land use history, species composition, topography), making it difficult to establish whether particular climate factors and other nonclimatic factors underlie observed patterns. Even if sites appear to differ primarily in climate, they may differ by predicted amounts of temperature but not by predicted changes in precipitation and soil moisture. Additionally, biota have had longer time periods to adjust and adapt to current climate variation at particular sites than they will have to adjust to different climate changes. Finally, this approach is unable to explore the effects of increased atmospheric CO_2, with the exception of some research on natural CO_2 gradients near hot springs.

Manipulations

Manipulations are one of scientists' most potent research tools. By conducting controlled manipulation of various climate, atmosphere, and resource conditions expected to change as a result of climate change, researchers can work toward understanding the role of single or multiple factors and their interactions in changing ecosystem structure and function. This type of research is used to predict both how specific anthropogenic climate change scenarios might impact biota and how resulting ecosystem changes may produce feedbacks to the climate system. In climate change ecology, manipulative research falls into two types of approaches: microcosm experiments and field experiments. Microcosms, which generally take the form of laboratory growth chambers of various sizes, can be used to carefully manipulate particular climate change factors (e.g., temperature, moisture, light, nutrients, atmospheric composition) and to monitor the response of soils, single or multiple organisms, assembled simple ecosystems, or intact ecosystem cores taken from the field. Field experiments also manipulate factors in order to look at interactions between climate change and biota, but do so in intact ecosystems (**Figure 6**).

The strengths and weaknesses of these two approaches are interrelated. Although it is relatively easy to manipulate, control, and replicate microcosms, field experiments may be confounded by ecosystem variability and complexity. In field experiments, typically only a very few experimental variables

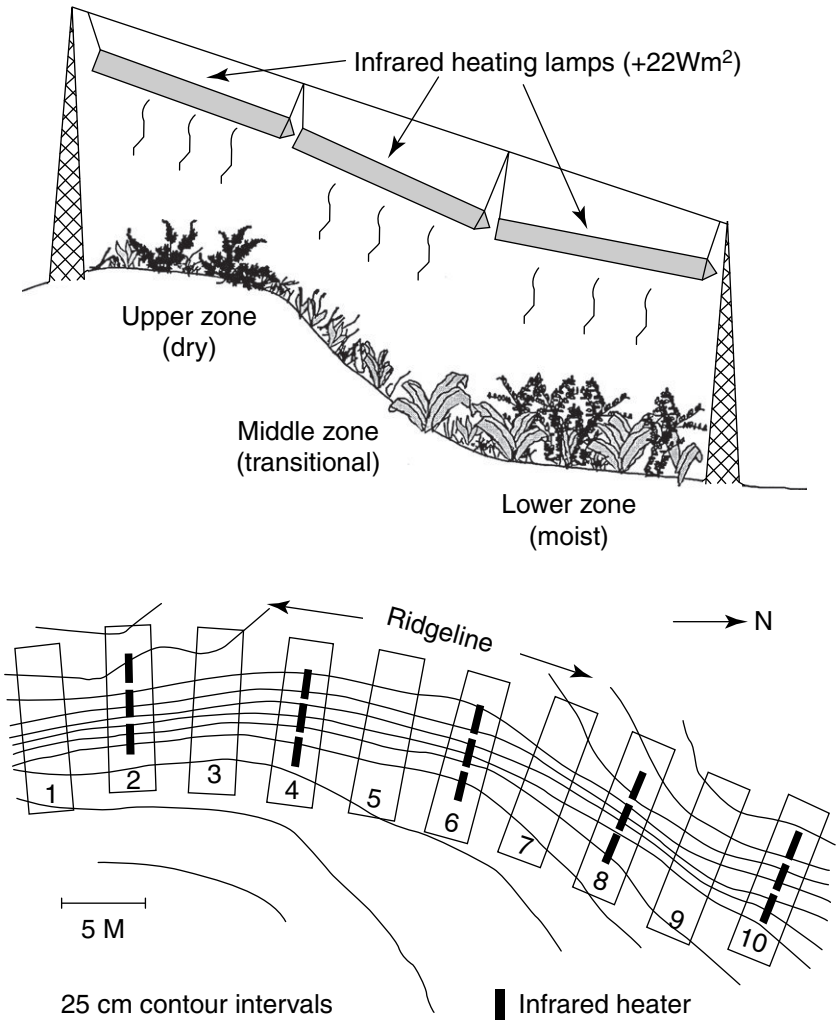


Figure 6 Profile of a heated plot used in a long-term ecosystem warming experiment to investigate the interactions between climate warming and the ecology and biogeochemistry of a subalpine meadow ecosystem at the Rocky Mountain Biological Laboratory in Gothic, Colorado, USA. The heaters increase downward infrared radiation flux ($+22\ W\ m^{-2}$) to a degree comparable to that expected under a $2 \times CO_2$ scenario. Layout of the plots is shown below; even-numbered plots have heaters and odd-numbered plots are control plots that lack heaters. Adapted from Figure 1b, 1c in Harte J and Shaw R (1995) Shifting dominance within a montane vegetation community: Results of a climate-warming experiment. *Science* 267: 876–877.

can be manipulated, controls can be difficult to establish, and adequate replication is often expensive and time-consuming. However, field experiments have the advantage of being conducted in a natural setting at broader scales that may be more useful for drawing conclusions about complex, "real-world" ecosystem dynamics, compared to highly simplified and small-scale microcosm experiments. In both types of experiment, the relatively abrupt, short-term (months to decades) manipulation of climate may not be a good analog of anthropogenic climate change, which is occurring over decades and centuries. In addition, changes to disturbances such as fires and hurricanes may prove to be more important in determining the abundance and distribution of biota in many ecosystems than the usual experimental focus on "average" changes in variables such as temperature, moisture, and CO_2.

Integrated Research

Given the limitations of each type of research, the most productive strategies for exploring interactions between ecosystems and anthropogenic climate change integrate multiple research approaches. For example, responses of biota to natural climate variation can be compared to responses of the same biota to manipulated climate change to see how responses differ or stay the same over multiple spatial and temporal scales. Results from field experiments and gradient studies often suggest mechanisms that can be more thoroughly tested in microcosm experiments. Gradient, field experiment, and paleobiological data sets can be used to parameterize, calibrate, and validate models of biotic response. In general, the thoughtful integration of approaches can build on the strengths and avoid some of the limitations of each type of research, thus helping scientists to develop

more rigorous hypotheses about climate change impacts and ecosystem feedbacks.

Impacts of Climate Change on Biota

Types of Responses

Adjustment

The first level, short-term response of any organism to changes in its environment is adjustment, also referred to as acclimatization. All organisms have some degree of physiological, life-history, or behavioral plasticity that enables them to live in a variable environment. The degree of plasticity with regard to climatic and atmospheric conditions varies widely among different kinds of organisms. Therefore, some types of organisms will be able to adjust to relatively large changes in climate, whereas others will be unable to adjust to even apparently minor increases in temperature or slight variations in precipitation.

An example of climatic adjustment in animals involves thermoregulation in vertebrates. Endotherms such as mammals have built-in physiological mechanisms to cope with body temperature changes. Ectotherms such as reptiles have behavioral traits that help regulate body temperature. Because of traits such as these, initial increases in environmental temperature should be well within the tolerances of many vertebrates. In plants, the concurrent increase of atmospheric CO_2 with surface temperature may augment the ability of some individuals, populations, and species to adjust to and flourish under anthropogenic climate change. Increases in CO_2, especially for plants with the common C3 photosynthesis pathway (e.g., most trees and shrubs), can result, at least initially, in the CO_2 fertilization effect mentioned above. Enhanced CO_2 concentrations can increase the ability of these types of plants to tolerate water stress, higher temperatures, and lower light. Other kinds of plants, particularly those with the C4 photosynthesis pathway (e.g., many low-latitude and low-elevation grasses), have physiological mechanisms that enable them to withstand warm temperatures and low availability of water. Such mechanisms provide a means of adjustment to drought stress that may be associated with increased temperatures and evaporation.

While most biota will have at least some capacity to withstand, and in some cases benefit from, initial changes in climate, the rapid rate and large magnitude of climate change are likely to quickly surpass their capacity to adjust to new climate conditions within their prewarming habitats. Biota that cannot continue to adjust will have to respond through evolution, migration, or extinction.

Evolution

Theoretically, populations and species could develop new adaptive traits as a result of evolution in response to anthropogenic climate change, thus enhancing the long-term survival of current taxa under new climate conditions. However, scientists generally agree that evolutionary responses to climate change are unlikely for most taxa since climate is changing rapidly compared to usual rates of evolutionary change. This view is supported by fossil data that reveal the morphological stasis of many taxa during previous periods of rapid climate change.

In the face of strong selectional pressure there is evidence that some species, especially those with fast generation times, can evolve very rapidly. For example, grass populations grown on soils polluted by heavy metals have shown signs of significant, genetic-based, heavy-metal tolerance within one or two decades. Some insects can evolve increased resistance to pesticides over the course of a few years. These types of responses depend on the presence of appropriate genetic variability in populations and species relative to a strong selective factor. As a result of climate change, populations and species will be exposed to novel environments resulting from climate change and associated shifts in ecosystem structure and function. Since many populations and species have climate-related genetic variability (e.g., differences in high temperature tolerance, drought tolerance), rapid evolution is possible.

Will anthropogenic climate change actually result in directed selectional pressures that are strong enough to drive microevolutionary responses? For at least the next several hundred years, species distributions are likely to be in fairly constant flux, which will tend to disrupt any potentially adaptive trends. Also, rapid evolutionary responses to anthropogenic climate change are unlikely in populations and species that have relatively long generation times, such as trees and many vertebrates. An added constraint on potential microevolutionary responses to climate change is the ongoing reduction in population size and thus genetic diversity of many species as a result of habitat destruction and other stresses. For most biota, other types of responses are far more likely to occur than evolution.

Migration

As current habitat becomes inhospitable owing to direct and indirect effects of climate change, biota will tend to track shifting climate and suitable habitat through dispersal and migration. Consequently, as a result of climate change, organisms are predicted to move generally poleward in latitude and upward in elevation. A rule of thumb is that a 3 °C change in temperature is approximately equivalent to a move of 250 km of latitude or 500 m of elevation. However, migration will be restricted or made impossible to the extent that there are inherent barriers (e.g., low mobility, slow reproductive rates) or external barriers (e.g., mountain ranges, large lakes) to movement.

Fossil records show that for many types of organisms, warming during the last deglaciation induced significant changes in latitude and elevation of species' ranges (**Figure 7**). Those distributional changes sometimes occurred at very rapid rates. For example, peak migration rates for some tree species in North America during the last deglaciation reached 100–500 m per year, probably as a result of haphazard, long-distance transport of seeds by animals, storms, or water. However, even these very fast historic migration rates only translate into 10–50 km per century, whereas anthropogenic climate change will likely require latitudinal shifts of at least 200–300 km over the next century. In some cases, changes in potential range boundaries resulting from climate change due to doubled CO_2 may exceed 1000 km. For example, suitable habitat for beech trees, which currently grow over

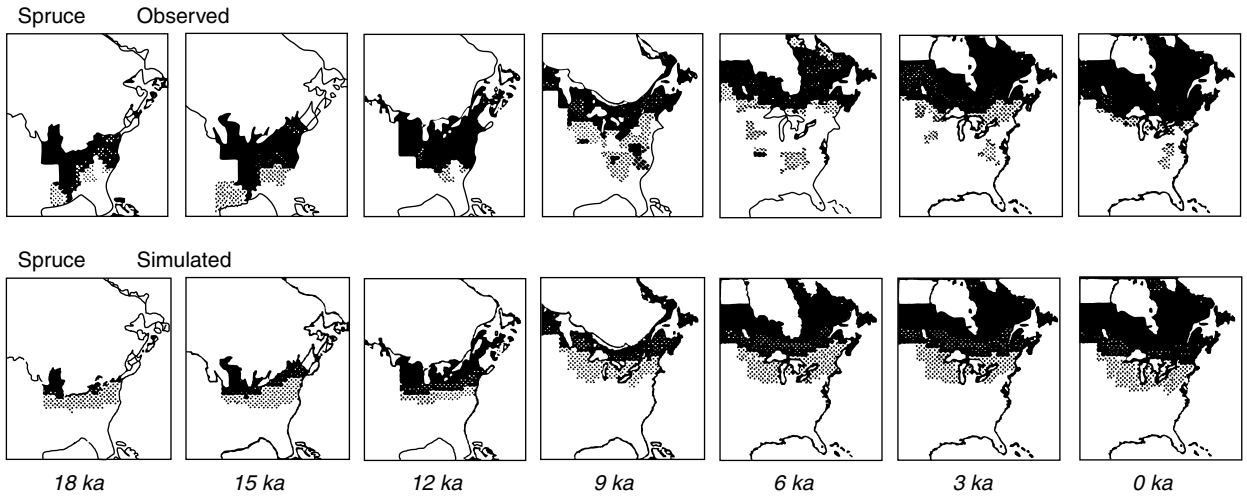


Figure 7 Maps of observed (upper row) and simulated (lower row) percentages for spruce tree pollen in eastern North America over the last 18,000 years. The simulated spruce pollen maps are based on the modern response of spruce pollen percentages to July temperature, January temperature, and precipitation as applied to simulated historical climates. Dark shading indicates the highest abundance of spruce. Reprinted from Webb T III (1992) Past changes in vegetation and climate: Lessons for the future. In: Peters R and Lovejoy T (eds.) *Consequences of Global Warming for Biological Diversity*, pp. 59–75. New Haven, CT: Yale University Press, and adapted from Figure 5 in COHMAP Members (1988) Climatic changes of the last 18,000 years: Observations and model simulations. *Science* 241, 1043–1052.

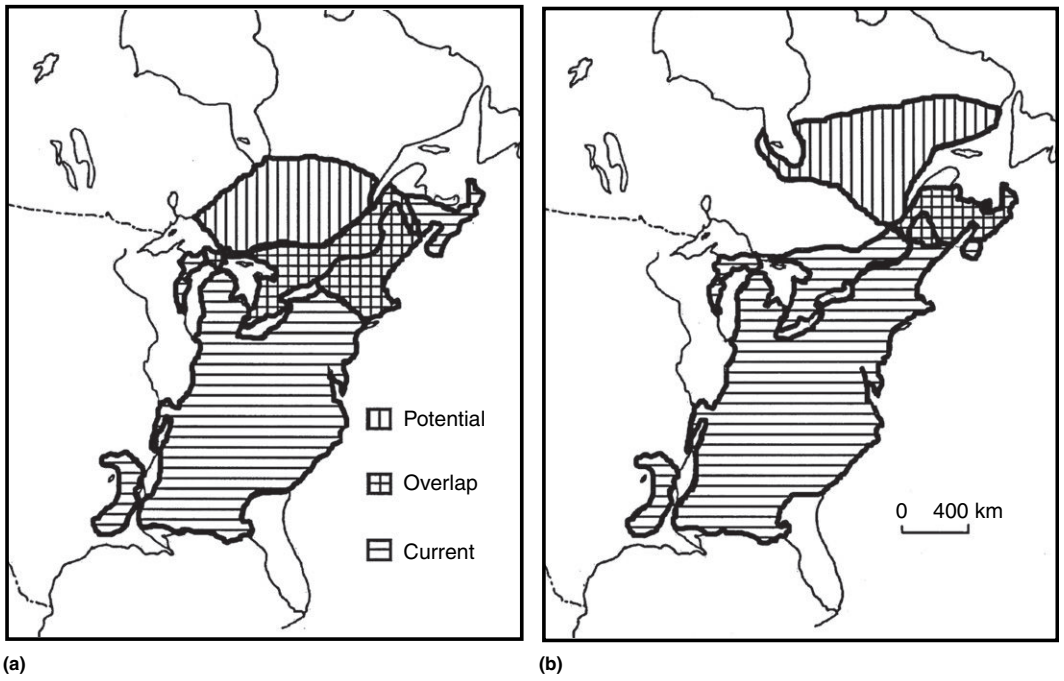


Figure 8 Current geographical range (horizontal lines) and potentially suitable range under doubled CO_2 (vertical lines) for beech trees in North America. Cross-hatched lines indicate areas of overlap between current and potential future ranges. (a) Output for a milder climate change scenario and (b) output for a more severe climate change scenario. Adapted from Figure 5.6 in Davis MB and Zabinski C (1992) Changes in geographical range resulting from greenhouse warming: Effects on biodiversity in forests. In: Peters and Lovejoy, with permission from Yale University Press.

most of the eastern third of the US, could shift almost completely out of the country and into a much smaller area of the Northeastern US and Southeastern Canada as a result of a $2 \times CO_2$ atmosphere (**Figure 8**). The rapid pace of anthropogenic climate change will easily outstrip the capacity of some organisms to move or disperse to a suitable new habitat. In addition, human destruction of habitat will create insurmountable barriers to migration for biota that have difficulty crossing large areas of urban development or agricultural use. In the short run, any migration that does occur may tend to increase local levels of biodiversity in some ecosystems as new species move in before old species completely

migrate or die out. This lag effect will tend to disappear with time.

Extinction

Biota that are unable to adjust, evolve, or migrate are unlikely to survive long in their prewarming habitats. As a result of climate change, organisms may be exposed to increasing physiological stress, abandoned by their mutualists and prey, outcompeted by more flexible neighbors or incoming species, and attacked by new predators and pests. Although microevolutionary responses are possible for some populations, rapid genetic adaptation is likely to occur in very few. Many species will be faced with migration problems. Even in species that successfully migrate, some populations are likely to go extinct, particularly at southern and lower edges of species' ranges, reducing genetic variability. Extinctions attributable primarily to climate change will be few over the next several decades, but will increase dramatically as time passes, climate change intensifies, and biotic response options narrow.

Types of Biota Impacted

Organisms at Risk

Given our knowledge of climate change and biology, the kinds of biota most likely to be at risk from climate change over the next several decades and centuries can be characterized. They include organisms that live in particular types of habitats such as at higher latitudes, on mountaintops, in low-lying coastal areas, or on islands. Scientists know with a high degree of certainty that temperature increases due to climate change will be greatest in polar regions. Therefore, higher-latitude temperate and Arctic/Antarctic ecosystems such as boreal forest, tundra, and peat bogs will experience both rapid and severe temperature increases, resulting in profound biotic change and disruption. Temperature increases will also be greater at higher elevations. Montane biota will tend to move up to cooler elevations. Biota already limited to mountaintops will be at serious risk of local extinction due to alteration of summit climate, the lack of potentially suitable habitat to migrate to, and the encroachment of lower-elevation species. Additionally, even small increases in sea level (i.e., several centimeters) due to climate change can result in altered coastal marine dynamics and flooding of low-lying areas. Rising sea levels will destroy or cause severe damage to ecosystems at the terrestrial/marine interface, such as salt marshes, estuaries, mangroves, and sand dunes, and are likely to disrupt coastal marine food webs.

Biota that lack the ability to readily disperse or move will be at a serious disadvantage. These include plants whose seed or clone dispersal rates and animals whose movement rates lag behind rates of climate change; slow-growing populations and species that will not have time to adjust to new conditions; biota that cannot or are slow to cross geographic barriers, for example, fish in isolated lakes, low-elevation plants bounded to the north by mountain ranges, and tropical forest birds and insects that do not cross unforested areas; organisms that depend on other biota for habitat or food, but that have very different degrees of mobility; species such as monarch butterflies and migratory shorebirds that have multiple habitat requirements; and relic biota that have been left in small, unusual habitats by chance and have no nearby habitat.

Organisms that are sensitive to extreme disturbance events, dependent on particular hydrological regimes, or that have certain kinds of physiological tolerances are also likely to be negatively impacted by climate change. Even though disturbances such as fires and hurricanes are a natural part of ecosystem dynamics, any increases in frequency and intensity of such disturbances due to climate change are likely to disrupt biota. For example, ecosystems such as tropical montane forests may have less time to recover between hurricanes, limiting the development of slow-growing, late-successional species. In terms of hydrological cycles, though it is often uncertain at local and regional scales how and to what degree precipitation and moisture availability will result from climate change, it is quite certain that changes will occur. Such changes could be critical in ecosystems such as tropical forests where the availability of food resources for animals is dependent on the timing of rainfall. In montane areas, many organisms will be very sensitive to changes in the snowpack and snowmelt. Regarding physiological tolerances, some organisms are adapted to living within a narrow range of limits of temperature, moisture, nutrients, light, and atmospheric composition, while others have wider tolerances but are already operating close to a threshold, beyond which their ability to live, grow, and reproduce is severely limited. Climate change may force the environment past physiological limits for some organisms, resulting in severe impacts, especially if dispersal or growth is slow. For example, slight increases (1–2 °C) in surface water temperature can induce bleaching and mortality of coral reefs, which have very slow rates of growth and provide habitat for many marine species.

Populations and species with few numbers, low genetic variability, or limited or unusual ranges will be vulnerable to climate fluctuations and will be at increased risk of extinction. Taxa that have highly specialized relationships with organisms are also at risk. Some species depend entirely on just one or a very few other species for nourishment or reproduction. If species respond very differently to changes in climate than do the species they depend on, and they cannot substitute other organisms to fulfill those roles, they will go extinct. Unusual or unique ecosystems may break apart as populations and species respond in largely individualistic ways to climate change and as feeding and other species interactions are disrupted.

Finally, climate change may negatively impact taxa that are already being challenged by other anthropogenic stresses. Humans engage in many activities that result in deleterious impacts such as habitat destruction, acid deposition, pollution, ozone depletion, and alien species introduction. When organisms are weakened by one of these stresses, they tend to become even more vulnerable to other stresses such as climate change. For example, insect pests can damage vegetation more when pollutants reduce plants' resistance to herbivory and warmer temperatures encourage pest population growth. If those insect pests happen to be alien, they may cause even more damage owing to lack of local predators and because local plants may lack resistance to alien pests. Also, changes in land use such as deforestation can reduce and isolate populations as well as create barriers to migration through habitat fragmentation and destruction.

Organisms Likely to Benefit

Climate change will benefit some biota through increases in abundance and range expansions, often at the expense of more at-risk biota. Types of biota likely to benefit from climate change include those that migrate easily, are opportunistic or generalists, or have high variability and rapid reproduction. Biota that are highly mobile and have rapid dispersal rates, such as some kinds of winged insects, will be equipped to track changing climate. Opportunistic organisms that can colonize disturbed areas will be at an advantage because they will be able to migrate through marginal habitat and establish in climatically disrupted ecosystems. For example, climate change is expected to promote the spread of already weedy introduced plant species, and may facilitate the escape of more garden cultivars into natural ecosystems. Generalist organisms that flourish in a wide variety of environments and have either wide tolerances for variable resource availability and climate or many possible prey items will fare better than highly specialized organisms. Also, populations and species with lots of phenotypic or genetic variation and rapid reproductive rates have the best chances of adjusting and adapting to rapidly changing climate. Finally, some organisms will be favored by new optima that result from climate change. For example, temperature increases are expected to increase parasite and insect development time, allowing parasites to spread with migrating insect hosts and promoting pest infestation and parasite infection of new hosts. These types of responses may lead to range expansions of agricultural pests and disease transmissions, as well as more frequent outbreaks.

Impacts on Communities

A key insight from the fossil record is that species tend to respond to climate change individualistically rather than as a group. Thus, while the general trend for species is to move poleward and to higher elevations as climate warms, particular species can vary quite dramatically in how fast and how much their ranges contract, expand, or move, what directions they move in, and at what rate they move around or over barriers such as mountain ranges. Consequently, past communities repeatedly disassociated and re-sorted into novel combinations. In some cases, although the same species still exist, there are no modern examples of historic species associations. For example, for several thousand years at the end of the last glacial period, spruce trees grew in open parklands in association with sedges. Today spruce is found in a completely different ecosystem type, the closed-canopy boreal forest, in association with birch, alder, and fir. Anthropogenic climate change is likely to result in the reconstitution of communities and ecosystems in unexpected ways.

Biota tend to move individualistically in response to long-term climate change because tolerances to climatic and atmospheric conditions are often specific to the organism, population, or species. Individualistic responses can lead to apparently counterintuitive shifts in range, especially if temperature is not the primary determinant of distribution. For example, the distribution of the gopher tortoise during the most recent deglaciation shifted south, rather than north. One hypothesis for this pattern is that seasonal climate extremes increased with warming, and that these extremes were more important for determining tortoise distribution than warming. In addition to climate and atmosphere, other abiotic factors (e.g., soil type, topography, disturbance regime, site history) and biotic interactions (e.g., mutualism, competition, predation, pollination) can play significant roles in determining the abundance and distribution of biota in both the short and long term. The interplay of all of these factors over time can result in complex and often unpredictable changes in the distribution, abundance, and diversity of biota.

Climate change will also precipitate many asynchronies that reduce the ability of biota to respond effectively to climate change. "Asynchrony" refers to a mismatch in timing or rate of change. One type of asynchrony already discussed is the mismatch between very rapid rates of anthropogenic climate change and slower rates of dispersal and migration for many species. Another type of asynchrony due to climate change is the potential mismatch between required resources and the availability of resources. Organisms are embedded within a network of relationships with other organisms, which they may depend on for sustenance and reproduction. However, different types of organisms may be affected very differently by climate change, which can disrupt biotic relationships that are important for community and ecosystem dynamics. For example, many species of flowering plants depend on specific animal pollinators such as butterflies, hummingbirds, and bees for successful reproduction, and the pollinators depend on those plants for food. In a hypothetical example, if the timing of flowering in a plant is primarily determined by early season temperature, while the emergence and activity of its pollinator is primarily determined by the amount of daylight, the shifting of temperature but not light levels due to climate change could lead to the pollinators being active after the plants flower. This type of asynchrony can reduce the reproduction and abundance of the plants, and reduce the availability of food for the pollinators. If the species involved are highly specialized on each other, they are likely to go extinct. These types of disruptions to species interactions can potentially alter the structure, dynamics, and stability of entire food webs and other ecological networks such as pollination and seed dispersal networks.

Evidence for Current Climate Change Impacts on Biota

During the 1990s, researchers started reporting the first evidence that climate change during the twentieth century has begun to influence populations, species, and ecosystems (see **Box 1** for a bibliography of the sources of information included in this section). Some evidence comes from coastal marine systems. Between 1951 and 1995, zooplankton biomass in coastal southern California waters decreased by 80% over four decades, at the same time the surface water layers warmed more than 1.5 °C in some areas. The surface warming appeared to result in changes to stratification and the thermocline that led to a reduction in upwelling of nutrients and thus primary production by phytoplankton, the ultimate food source of zooplankton. Other research compared changes in abundances of 45 invertebrate species in a central California intertidal community from the 1930s to the 1990s. During that period of time, annual mean shoreline temperature at the study area increased by 0.75 °C, and average summer

Box 1 Selected references for current climate change impacts on biota

Barry JP, Baxter CH, Sagarin RD, and Gilman SE (1995) Climate-related, long-term faunal changes in a California rocky intertidal community. *Science* 267: 672.
Brown JL, Li SH, and Bhagabati N (1999) Long-term trend toward earlier breeding in an American bird: A response to global warming? *Proceedings of the National Academy of Sciences USA* 96: 5565.
Crick HQP and Sparks TH (1999) Climate change related to egg-laying trends. *Nature* 399: 423.
Grabher G, Gottfried M, and Pauli H (1994) Climate effects on mountain plants. *Nature* 369: 448.
Myneni RB, Keeling CD, Tucker CJ, Asrar G, and Nemani RR (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* 386: 698.
Parmesan C(1996) Climate and species' range. *Nature* 382: 765.
Parmesan C, Ryrholm N, Stefanescu S, *et al.* (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* 399: 579.
Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics* 37: 637.
Roemmich D and McGowan J (1995) Climatic warming and the decline of zooplankton in the California current. *Science* 267: 1324.
Root T, Price J, Hall K, Schneider S, Rosenzweig D, and Pounds J (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421: 57.
Thomas CD and Lennon JJ (1999) Birds extend their range northwards. *Nature* 399: 213.

temperature increased by 2.2 °C. Eight of nine species with a southern geographic range showed significant increases in abundance, whereas five of eight species with a northern distribution showed significant decreases. Species with wide distributions showed no strong patterns of change.

At high elevations and latitudes, terrestrial vegetation appears to be responding to the warmer climate. A comparison of 40- to 90-year-old records of plant distribution on high-altitude mountaintops ($\sim$2900–3500 m) in the Alps to 1992 data, given a 0.7 °C increase in mean annual temperature during that time period, found upward migration rates by nine "typical" species of $<$1–4 m per decade. In addition, species richness increased over time, with the increase most pronounced at lower-altitude summits. Evidence from satellite data from 1981 to 1991 suggests that the photosynthetic activity of terrestrial vegetation in northern high latitudes increased by 7–14% over that time period. This increase may indicate an intensification of plant growth associated with a 12-day increase in the growing season.

Insect species are also displaying sensitivity to recent climate change. Historical records and the recent status of 151 populations of Edith's checkerspot butterfly throughout its entire range show that the net extinctions of populations were significantly greater at southern latitudes and lower altitudes, resulting in an observed northward and upward shift in the species' range during the last several decades. In a later study that examined range shifts in 35 nonmigratory European butterfly species during the last century, 63% of the species displayed northward range shifts of 35–240 km, while only 3% shifted south.

Vertebrates also appear to be sensitive to recent climate change. Current evidence suggests that birds are both breeding earlier and shifting their ranges northward in response to warming. From 1971 to 1998, the average timing of first clutch in the Mexican jay in Arizona occurred 10 days earlier. This change was associated with significant increases in monthly minimum temperatures. Twenty UK bird species over 25 years displayed long-term trends toward earlier egg laying in most species. A data set of 36 bird species over 57 years suggest that 86% of the species display significant relationships between timing of egg laying and temperature or rainfall. Furthermore, the breeding distributions of British birds over a recent 20-year period indicate that the northern margins of the species' ranges moved north by an average of 19 km.

In general, the observations of most of these types of studies are consistent with predictions of impacts of climate change on biota, and demonstrate how just a small amount of climate change over brief time periods can lead to significant ecological changes.

Conclusions

The enhanced greenhouse effect, in conjunction with other anthropogenic stresses, is likely to precipitate unprecedented changes to Earth's climate and ecosystems. Though the details of how climate change will affect biodiversity are often hard to predict, there is little doubt that biological impacts will be pervasive and often dramatic. Studying the effects of climate change on biota can help in the formulation of strategies for conserving biodiversity and ecosystem structure and function in the face of potentially massive change and loss. Such knowledge is also crucial for refining predictions of the future rate and magnitude of climate change, since biological responses are likely to produce significant feedbacks that can augment or dampen climate change at local, regional, and global scales. Understanding and addressing the interactions between climate change and biodiversity represents one of the greatest challenges that scientists and policymakers face in the twenty-first century.

See also: Atmospheric Gases. Biogeochemical Cycles. Carbon Cycle. Climate Change and Ecology, Synergism of. Climate, Effects of. Deforestation and Land Clearing. Migration. Soil Biota, Soil Systems, and Processes

References

Clark JS, Fastie C, Hurtt G, *et al.* (1998) Reid's paradox of rapid plant migration. *Bioscience* 48: 13.
Emanuel WR, Shugart HH, and Stevenson MP (1985) Climatic change and the broad-scale distribution of terrestrial ecosystem complexes. *Climatic Change* 7: 29.
Foley JA, Kutzbach JE, Coe MT, and Levins S (1994) Feedbacks between climate and boreal forests during the mid-Holocene. *Nature* 371: 52.
Hartmann DL (1994) *Global Physical Climatology.* San Diego, California: Academic Press.

IPCC (2007) Summary for Policymakers. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press.

Kareiva P, Kingsolver J, and Huey R (eds.) (1993) *Biotic Interactions and Global Change*. Sunderland, MA: Sinauer Associates.

Lashof DA (1989) The dynamic greenhouse: Feedback processes that may influence future concentrations of atmospheric trace gases and climatic change. *Climatic Change* 14: 213.

Lashof DA, DeAngelo BJ, Saleska SR, and Harte J (1997) Terrestrial ecosystem feedbacks to global climate change. *Annual Review of the Energy and the Environment* 22: 75.

Lenton TM, *et al.* (2008) Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences of the United States of America* 105(6): 1786–1793.

Lorius C, Jouzel J, Raynaud D, Hansen J, and Le Treut H (1990) The ice-core record: Climate sensitivity and future greenhouse warming. *Nature* 347: 139.

Markham A (1996) Potential impacts of climate change on ecosystems: A review of implications for policy makers and conservation biologists. *Climate Research* 6: 179.

Markham A, Dudley N, and Stolton S (1993) *Some Like It Hot: Climate Change, Biodiversity and the Survival of Species*. Gland, Switzerland: World Wildlife Fund International.

Peters RL and Lovejoy TE (eds.) (1992) *Global Warming and Biological Diversity*. New Haven, CT: Yale University Press.

Schlesinger WH (1991) *Biogeochemistry: An Analysis of Global Change*. San Diego, CA: Academic Press.

Solomon AM and Shugart HH (1993) *Vegetation Dynamics and Global Change*. New York: Chapman & Hall.

Wallace JM and Hobbs PV (2006) *Atmospheric Science: An Introductory Survey*. Amsterdam, The Netherlands: Academic Press.

Watson RT, Zinyowera MC, and Moss RH (eds.) (1996) *Climate Change 1995 – Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses, Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

Woodwell GM and Mackenzie FT (eds.) (1995) *Biotic Feedbacks in the Global Climate System: Will the Warming Feed the Warming?* New York: Oxford University Press.

Guilds

Richard B Root, Cornell University, Ithaca, NY, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 295–302, © 2001, Elsevier Inc.

Glossary

Adaptive syndrome A coordinated set of adaptations.
Community An assemblage of populations that coexist in an area.
Eltonian niche The role, or occupation, of a species in a community.
Exploitative competition An adverse interaction which results from organisms depleting their shared resources.
Family A category, in the classification of evolutionary lineages, of related organisms that ranks above a genus and below an order. A family usually contains many genera.
Grinnellian niche The requirements and behaviors expressed by a species wherever it normally occurs.
Hutchinsonian niche The set of environmental conditions, or opportunities, that will permit a species to exist indefinitely. The set of opportunities that are available to a guild can be referred to as a "nook."
Individualism of species Refers to the observation that the distributions and abilities of most species do not exactly coincide because each species is the product of a unique evolutionary history. Coevolution and convergence may occur, but rarely are these processes so complete that groups of species are distributed as a unit or superorganism.
Taxon A general term used to indicate groups of related organisms at any level in a taxonomic hierarchy. In the classification of evolutionary lineages, the term can refer to species, genera, families, orders, etc.

The Definition and Properties of Guilds

Guilds are groups of species that exploit the same class of resources in a similar way (Root, 1967). This simple definition has several implications that are best explained by way of an example. Consider the birds that probe for insect prey on the bark of tree trunks (**Figures 1** and **2**). In many parts of North America this guild is composed of small woodpeckers (family Picidae), nuthatches (family Sittidae), and treecreepers (family Certhiidae). The guild also contains the black-and-white warbler (family Parulidae), whose relatives are primarily adapted for gleaning insects from foliage. Guilds thus contain species with very different phylogenetic histories and, as a consequence, with different propensities for dealing with the special requirements associated with exploiting a particular resource. Thus, the woodpeckers and treecreepers have evolved stiffened tails to act as props in climbing tree trunks, whereas the nuthatches and warbler have evolved modifications of the foot to assist in moving on vertical surfaces.

In addition to the core members of the guild which take the bulk of their food by foraging on trunks and limbs, there are many species, such as titmice (family Paridae), that occasionally feed by probing bark. The presence of these infrequent bark probers illustrates that guilds can have a hierarchical structure with members that range from strict specialists to occasional opportunists. As a consequence, the size and makeup of the guild will depend on where we draw the line with respect to the proportion of a species' diet that is obtained while probing bark.

The hierarchical structure of guilds is also reflected in how broadly we interpret the "in a similar way" portion of the definition. For instance, there are ants that forage for insect prey on tree trunks. These insects have the potential of influencing the food supply available to the bark probers. Their foraging style, however, is fundamentally different: Ants rely on their small size to enter bark crevices, whereas the birds employ their relatively large mass to lever out embedded prey and to flake bark. In this example, it can be seen that the scope of a guild's delimitation needs to be matched to the question being asked. The broader definition, which includes ants, might be most useful if our primary concern is competition, and the narrower concept, which excludes ants, might be more appropriate if our primary interest is in the evolution of adaptations. Furthermore, the birds eat the ants, which introduces further complications if the ants are lumped into the same guild with the birds.

During the northern winter, the black-and-white warbler migrates to the American tropics, where it joins another representation of the bark-probing guild which includes the tropical woodcreepers (family Dendrocolaptidae) and barbtails (family Furnariidae) shown in **Figure 2**. The warbler's experience illustrates that guild membership can vary in space and season.

The opportunities provided by the existence of bark-dwelling arthropods, many of which are in a cryptic resting stage, has been exploited by bark probers that evolved from several phylogenetic lineages throughout the world. Thus, the Australian manifestation of the bark-probing guild contains Australasian treecreepers (family Climacteridae) and sittellas (recently placed in their own family, the Neosittidae). On Madagascar, the bark-probing role is filled by the nuthatch-vanga (family Vangidae), and on the Galapagos Islands the woodpecker finch (family Fringilliade) uses cactus spines to probe for prey under bark.

In classifying the ways that organisms exploit resources, guilds hold a rank that is similar to genera in phylogenetic schemes. One may also think of a guild as a group of species that occupy similar niches. Indeed, one of the original

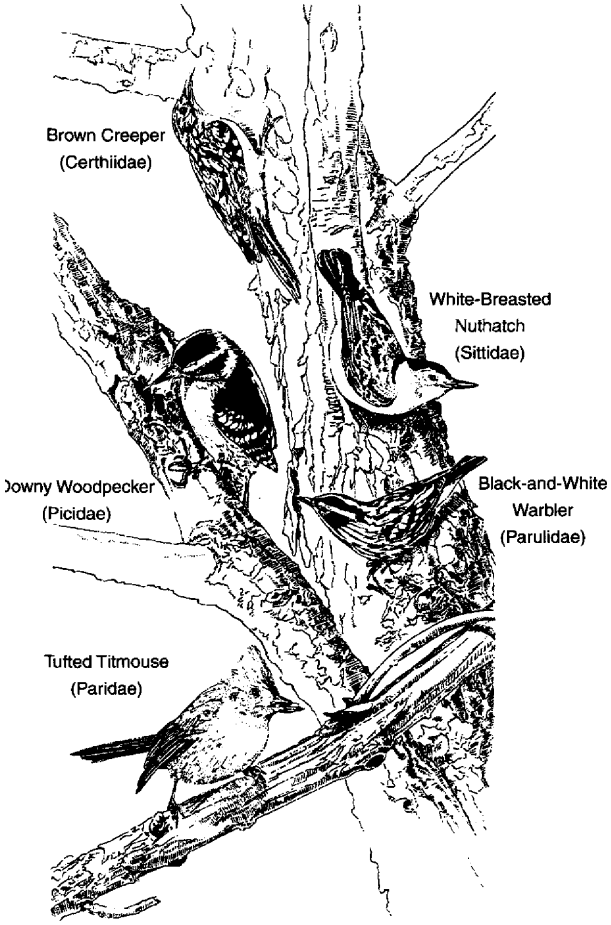


Figure 1 A few members of the bark-probing guild in North America.

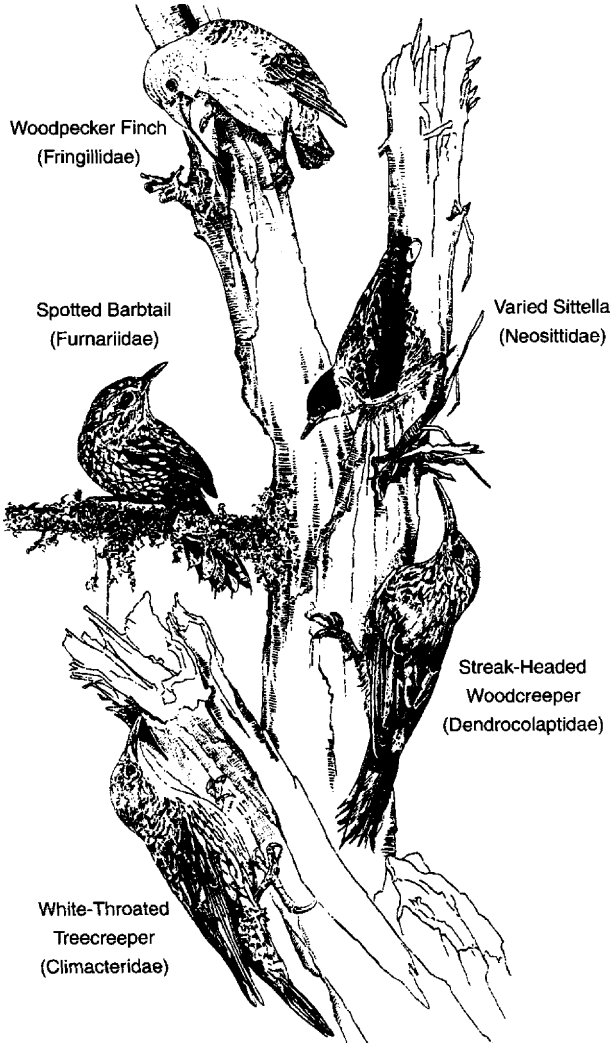


Figure 2 Birds from different regions and lineages that have independently evolved the bark-probing habit. The dendrocolapids and furnaniids live in the American tropics and the climacterids and neosittids live in Australia. The woodpecker finch is one of the Darwin's finches that inhabit the Galapagos Islands.

motivations for introducing the guild concept was to clarify confusion regarding the niche concept, which was used to describe three quite different entities: (i) the Hutchinsonian niche, which refers to the set of conditions that are sufficient for a species to exist in a particular habitat; (ii) the Grinnellian niche, which refers to the requirements and behaviors that are expressed wherever a species normally occurs; and (iii) the Eltonian niche, which refers to the role or occupation of a species in a community. The Eltonian niche was usually viewed as a relatively broad category (e.g., sap-feeding insects and predators of small mammals) which could contain several species and occur in a variety of habitats. In creating the guild concept, we provide an alternative category which clarifies some of the ambiguities associated with the Eltonian niche. For instance, by substituting guild for the Eltonian niche, we can avoid the contradictions that occur when several similar species are said to occupy the same "niche," a category that is supposed to be a property of individual species according to the Hutchinsonian and Grinnellian concepts. Accordingly, species exploit niches and guilds exploit "nooks"—the adaptive space that is presented by resources with similar characteristics.

There are many types of guilds in addition to those that are based on food resources. For instance, animals can be grouped according to their use of tree cavities for nest sites and plants can be grouped according to their shared use of agents for pollination (e.g., moths or hummingbirds) and seed dispersal (e.g., fruit-eating birds or wind). In the case of plant-eating insects, there has been a tendency to define "mixed" guilds that group species according to a combination of overlapping functions. Frequently, the "feeding guild," defined on the basis of the insects' mode of feeding and the plant tissues they ingest, is combined with the "sheltering guild," defined on the basis of the insects' use of different plant structures as lairs for protection from enemies and adverse conditions. As a result, we often find that plant-chewing insects are divided into categories such as "leaf-miners," "leaf-rollers," "shoot-borers," and "stem galls" that have been shaped by the interaction of nutritional and protective functions. In searching for patterns, it is important to take all these mixed functions into full account.

It should also be kept in mind that species can belong to more than one guild. For instance, the woodpeckers

mentioned previously are members of both the feeding guild that probes bark and the nesting guild that utilizes tree cavities. Many insects shift from one guild to another during the course of development because larvae and adults are adapted to perform such different functions. Thus, most butterfly species have caterpillars that are in the leaf-chewing guild and adults that are in the nectar-drinking guild.

The Utility of Guilds in the Search for Organization and Patterns

Guilds Provide a Natural Unit of Manageable size for Comparative and Evolutionary Studies of Species Interactions

Even fairly simple communities, such as old fields in the temperate zone, can harbor more than 1500 species of insects belonging to more than 175 families. Such diversity obscures our view of fundamental processes, such as competition, mutualism, and evolution, because most species interact only rarely. This requires that we sort out the insignificant interactions to appreciate the true intensity of the critical interactions. For instance, most species of plant-eating insects in an old field rarely come into direct contact with one another because they are adapted for feeding on different plant species. As a consequence, the use of standard sampling methods, such as sweeping vegetation with a net or collecting the moths that are attracted to lights, which combine the insects on different plant species, would blur our understanding of competitive interactions between plant feeders. Furthermore, the existence of many types of species presents technical difficulties. Thus, to measure the density of all the plant feeders in a community might involve censusing deer, rabbits, mice, sparrows, snails, grasshoppers, caterpillars, aphids, galls, and tiny mites. Clearly, each of these categories of plant feeders require somewhat different methods based on different assumptions, relating to different spatial scales, and having different biases. To avoid these difficulties, ecologists often restrict their attention to some portion of the community. The guild concept is ideally suited to this purpose.

In defining a guild, we create a community detachment whose members share resources that are exploited in a similar way. In many situations, guilds have an advantage over partitioning on the basis of taxonomy because the members of a taxon may be utilizing a variety of resources in a variety of ways. For instance, a bird community may contain rodent-eating owls and seed-eating sparrows but take no account of rodent-eating weasels and seed-eating mice. This is not to imply that guilds are necessarily invalid if all their members are drawn from the same major taxon; frequently, certain styles of exploiting resources are only "available" to organisms with capabilities that evolved in particular taxa. As discussed previously, birds possess the beak and body mass required to probe tree bark. Focusing on a guild has the added advantage of improving our ability to conduct accurate censuses on all the guild members because by sharing the same resources they are likely to be found in the same microhabitats and, by behaving in a similar way, they are more likely to have similar reactions to being observed or captured. Furthermore, the resources, by virtue of belonging to the same class, can be more accurately measured and compared. As a consequence, we are in a position to draw closer comparisons and conduct more realistic experiments when we are working with a guild.

By observing how a guild changes along spatial gradients, we can seek patterns in the relationships between environmental factors and variations in population density, reproductive rate, body size, dispersal ability, investment in predator or herbivore defenses, and a variety of other traits within a group of species that are comparable because they use similar resources in a similar way. Such comparisons also permit us to examine how species are "packed" onto the guild's resource base in relation to disturbance, elevation, and latitude.

Comparisons of similar guilds can reveal interesting patterns. For instance, there is a much smaller range of sizes of plant-chewing insects that live in abodes, such as mines and leaf rolls, compared to that of plant chewers that feed in exposed locations. This pattern reflects the limitations on body size that are imposed by the small spaces provided by plant structures that are pliant enough to be penetrated or rolled.

Much can be learned by comparing the same guild in regions that have been isolated from one another so that their biotas have undergone independent evolution (**Figures 1** and **2**). In such comparisons, the discovery of close resemblance provides insights into the process of convergent evolution. In particular, we can observe (i) if certain types of characters are more responsive to evolutionary pressures, (ii) if certain combinations of traits are associated with each other in different regions, and (iii) the degree to which convergence is limited by deep ancestral traits in the various phylogenetic lines that make up the guild.

Guilds Partition Biotas into Entities Appropriate for Detecting Community Organization

Ecologists have long debated the significance of exploitative competition in organizing communities. An important line of inquiry in this debate concerns the constancy of a community's functional relationships. The reasoning is as follows: If communities are competitively organized, we expect to find that important functional relationships are predictable in space and time because the waxing and waning of populations of one species would be countered by compensatory changes in the populations of other species that use the same resource. Furthermore, these compensatory responses are expected to be most intense within groups of species that have similar methods of using the shared resources. Finally, we might expect to find that similar types of communities have similar functional profiles because the imperatives associated with using similar resources have sorted and shaped species into convergent groups. Guilds provide a means of partitioning a community into functional entities in which these compensations might be detected.

Much of the search for community organization has focused on plant–insect associations. One reason for doing this is that the basic resources (the biomass of foliage, stems, roots, flowers, and seeds produced by a particular plant taxon) are well circumscribed and relatively easy to measure. The

constancy of the association's functional profile can be judged by changes in the guild spectrum—the proportion of total herbivore biomass that is engaged in exploiting different plant parts in various ways. A matrix for sorting insect herbivores into guilds is presented in **Table 1**. In all likelihood, such a scheme will need to be revised to suit the needs of a particular investigation.

Evidence for community organization based on guilds has been elusive. For instance, Cornell and Kahn found extensive variation in the guild spectra associated with 28 tree species in Great Britain and Root and Cappuccino found no evidence for compensatory changes in the densities of species within guilds that feed on goldenrods. These failures draw us to question our assumption that the intensity of competition is likely to be most severe within guilds. Different types of resources are linked in nature. Thus, in plant–insect associations, the consumption of leaves will reduce the future availability of seeds and the sapping of juices from stems will reduce the nutritional quality of foliage eaten by leaf chewers. Similarly, the caterpillar eaten by a foliage-gleaning bird is unavailable as a moth to bats that forage in the air. Linked resources link guilds and open the possibility for exploitative competition to play out its effects across a diffuse network of species. In other words, the use of a common resource seems to be of primary importance in competitive interactions; the manner in which the resource is exploited is of little consequence unless differences in a guild's behavior permit members to use an exclusive subset of resources. As a consequence, there are many cases of competition between guilds. For example, Schluter found that Darwin's finches in the Galapagos Archipelago consumed more nectar on islands on which nectar-feeding carpenter bees were absent.

Of course, competition could be operating in a diffuse fashion that would be difficult to detect by only studying variations in guild spectra. In plant–insect associations, however, there is increasing evidence from long-term experiments and extensive surveys that plants in natural communities are rarely devastated by herbivores and that competition between plant-eating insects is a sporadic occurrence. These results, considered in conjunction with the highly variable guild spectra that have been reported by Cornell, Root, and Strong, suggest that most plants could support more herbivore species (i.e., plants provide Hutchinsonian niches that have not been filled) and that the structure of these associations is idiosyncratic, largely determined by the characteristics of a few dominant species that happen, for a variety of historical reasons, to have evolved to become members of the community.

Many ecologists have the impression that guilds and communities that are dominated by predators or vertebrates are more likely to be organized by competition and to exhibit more predictable functional profiles. This impression, however, has not been adequately evaluated.

Guilds Provide a Framework for Describing and Comparing the Trophic Structure of Ecosystems

The trophic structure of ecosystems is often described in terms of broadly defined levels consisting of primary producers, primary and secondary consumers, decomposers, and so on. Food webs, on the other hand, are described in terms of individual species. Various attempts have been made to develop a scheme based on guilds that can be used to describe trophic structure with an intermediate degree of refinement; thus, herbivores might be subdivided into browsers of woody plants, grazers of grasses, sap-tappers on succulent foliage, borers in stems, and so on. To accomplish this requires that the entire community be partitioned so that all its members can be placed into a standard set of feeding guilds that have a widespread occurrence in nature. When ordered in this way, it is hoped that insights about energy flows can be gained by drawing comparisons between systems. In practice, however, such schemes have been little used, probably because it is almost impossible to develop a "key" to the guilds of the world that is flexible enough to accommodate a wide array of taxa and yet sufficiently circumscribed that different ecologists will assign species to guilds in the same manner.

The Individualistic Nature of Species and the Vague Boundaries of Ecological Categories

The process of defining guilds raises some fundamental questions. Are guilds distinct entities that reflect the outcome of natural processes or are they merely artificial groupings that we invent to divide diversity into more comprehensible units? In other words, are the opportunities provided by the environment discontinuous so that species are sorted into clearly defined groups on the basis of their requirements and lifestyles? Furthermore, are similar openings (or nooks) available in different regions? Are organisms constrained from filling these opportunities by their evolutionary histories? We can begin to address these difficult questions by comparing the characteristics of the core and marginal members of various guilds.

Table 1 A Provisional matrix for defining guilds of insects that feed on terrestrial plants

Manner of feeding	*Resource*							
	Buds and leaves	*Pliant stems*	*Flowers*	*Pollen*	*Fruits*	*Seeds*	*Wood and bark*	*Crusts of algae, molds, etc.*
Chewers								
Exposed	✓	✓	✓	✓	✓	✓		✓
Concealed	✓	✓	✓		✓		✓	
Sap-tappers	✓	✓	✓		✓	✓		
Gall-makers	✓	✓	✓		✓			
Grazers				✓				✓

If we census the habits of all the species that utilize a particular resource, we usually find that there is a wide variation in the level of their dependency on the resource. The most dependent species—the core users—usually possess obvious specializations. Thus, in the bark-probing guild (**Figure 1**), the treecreepers, woodpeckers, and nuthatches, which take the bulk of their diet from bark, have specialized feet and tails for moving on vertical surfaces (Richardson, 1942). In our census, however, we usually encounter a host of species that use the resource only rarely. These marginal users are either generalists or they are specialized along other lines that compromise their ability to use the resource in question. For instance, bark-probing titmice often forage on horizontal surfaces, such as the tops of limbs and branches, where they take much of their prey by hammering apart acorns, galls, and similar objects that they hold against the perch with their foot. These habits are reflected in the distinctive foraging maneuvers that titmice display during their infrequent forays onto tree trunks. Other species, such as gnatcatchers, vireos, and wood warblers, which are primarily adapted for gleaning insects from tree foliage, make clumsy efforts to take insects from bark on rare occasions.

Here we confront the individualistic nature of species—a fundamental issue, originally raised by Gleason, that complicates all of our efforts to define functional groups of species. Thus, the various taxa that can utilize a particular resource have evolved "individualistically" along independent paths. As a result of their separate histories, these taxa have different constraints and proclivities, which result in different levels and styles of specialization. As a consequence, guilds often have ambiguous boundaries consisting of several species of generalists and inept opportunists that are specialists on other resources. Since the degrees of dependency grade into one another, the line one draws to define the membership in a guild can be somewhat arbitrary.

The indefinite boundaries of guilds are merely one expression of a more general issue confronting ecologists. Community ecology and biogeography are, by their very nature, concerned with levels of organization that consist of multiple species. In their efforts to discover the processes operating at these levels, workers have developed several systems for classifying species into groups on the basis of similarities in their (i) response to physical or site conditions to define communities; (ii) possession of particular adaptive traits to define life-forms and adaptive syndromes; (iii) response to seasonal cues to define phenological aspects; (iv) geographic distributions to define biogeographic provinces, biomes, life zones, and plant formations; and (v) diet to define trophic levels. Ecologists recognize and discuss these entities because they find that several species fall into clusters on the basis of these various types of classification. As with guilds, however, we also encounter species with intermediate characteristics that tend to blur the boundaries of these various categories. This is because all the traits that underlie these classifications are subject to the same individualistic evolution that can produce the marginal members of a guild. Thus, vague boundaries are an inherent property of any species assemblage—a fact that requires ecologists to cultivate certain habits of mind in forming arguments.

Attitudes Concerning Vaguely Bounded Categories

People are generally uneasy about categories, such as guilds, that are subjectively defined because the criteria one chooses to define the operational boundaries will have an impact on our ability to observe patterns. Thus, if a guild is defined too broadly, the characteristics of a miscellaneous collection of marginal species could obscure interesting similarities between the core members. On the other hand, if a guild is defined too narrowly, we could overlook the full range of influences that stem from extracting a particular resource in similar ways. There have been a variety of attempts to address this problem by using quantitative procedures, such as cluster analysis and principal components analysis, to delimit the membership in guilds. The raw material for these analyses comes from "activity censuses" which measure the relative frequency that species utilize different resources in different ways. The groupings that are sorted out by these procedures, however, often seem abstract and artificial. Furthermore, they are probably no less arbitrary than those that are more subjectively defined because the person who is running the analysis must decide how broadly to cluster the species' activities.

Despite these difficulties, the continued and widespread use of subjectively defined categories attests to the need of such concepts in ecology. The best results are obtained when we follow some simple guidelines. In the case of guilds, the classifications that provide the most insights are based on a thorough knowledge of the species' natural history. Open-minded observations and journal keeping are good starting points to form an intuition for distinguishing resources with properties that require special means for their efficient exploitation. For instance, insects on bark constitute an appropriate resource for a guild of bird-sized predators because the core members must move on vertical surfaces and extract prey from narrow crevices or under bark—maneuvers that are difficult for birds that lack the necessary specializations.

After a guild has been tentatively defined, it helps to conduct activity censuses at different sites and in different seasons to discipline intuition and provide a quantitative basis for determining guild membership. (For sedentary organisms such as plants, one would count the incidence of the traits that define the guild along transects laid out in appropriate habitats.) When the investigators are satisfied that a guild "makes sense" with respect to the questions that are being asked, the criteria that will be used to delimit the boundaries should be clearly described and fully justified. At this point, it may be useful to assign classes of membership; for instance, core species might be those that engage in the activities that define the guild on at least 50% of occasions, and marginal or accidental species might be those that engage in such activities on less than 10% of occasions. Many of these steps seem obvious, but they are often left out of guild classifications, especially in cases in which the investigators are attempting to partition entire communities in such a way that each species can be assigned to a single guild. Such classifications are often used to compare the functional organization of communities in which only the most general of categories can accommodate the great diversity of species that must be placed.

Part of the art of becoming an ecologist involves developing a set of attitudes for coping with the complications

that stem from the individualistic nature of species. One of the most important of these is an ability to match the question one is asking with the most appropriate grouping of species—the set that will reveal valid patterns that act in nature. Thus, the guild definition that is most useful for exploring convergent evolution may be quite different from the one that is best for observing compensatory shifts in the densities of interspecific competitors. In addition, we need to be on steady guard against the natural tendency to drift into thinking that the entities we have invented are "real." Guilds are not a fixed feature of nature. They are a convenience that we can use to cope with diversity, reveal patterns, and facilitate discussion. As a consequence, guilds can be modified as long as they provide a valid base for addressing a question and their limits are explicitly defined and justified.

See also: Competition, Interspecific. Functional Groups. Habitat and Niche, Concept of. Trophic Levels

References

Cornell HV and Kahn DM (1989) Guild structure in the British arboreal arthropods: Is it stable and predictable? *J. Anim. Ecol* 58: 1003–1020.

Curtis JT (1959) *The Vegetation of Wisconsin*. Madison: Univ. of Wisconsin Press.

Gleason H (1926) The individualistic concept of the plant association. *Bull. Torrey Botanical Club* 53: 7–26.

Hawkins CP and MacMahon JA (1989) Guilds: The multiple meanings of a concept. *Annu. Rev. Entomol* 34: 423–451.

Richardson F (1942) Adaptive modifications for tree-trunk foraging in birds. *Univ. California Publ. Zool* 46: 317–368.

Root RB (1967) The niche exploitation pattern of the blue-gray gnatcatcher. *Ecol. Monogr* 37: 317–350.

Root RB and Cappuccino N (1992) Patterns in population change and the organization of the insect community associated with goldenrods. *Ecol. Monogr* 62: 393–420.

Simberloff D and Dayan T (1991) The guild concept and the structure of ecological communities. *Annu. Rev. Ecol. Syst* 22: 115–143.

Strong DR, Lawton JH, and Southwood R (1984) *Insects on Plants: Community Patterns and Mechanisms*. Cambridge, MA: Harvard Univ. Press.

Wiens JA (1989) *The Ecology of Bird Communities*. Cambridge, UK: Cambridge Univ. Press.

H

Habitat and Niche, Concept of

Kenneth Petren, University of Cincinnati, Cincinnati, OH, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 303–315, © 2001, Elsevier Inc

Glossary

Community An ecological term referring to a set of species that occur in the same location that have the potential to affect each other either directly or indirectly.

Community structure The web of potential biological interactions among members of a community that may be characterized in terms of diversity, complexity, hierarchy, and stability.

Ecotone A zone of transition between two different habitats that may contain a community of organisms distinct from either habitat.

Habitat structure Analogous to community structure, but limited to the physical structural aspects of a habitat. The structure of habitats may be characterized by such measures as complexity, heterogeneity, regularity, stratification, and fractal dimensionality.

Microhabitat Locations within a habitat where organisms may carry out important aspects of their lives, such as places for harvesting food, nesting, or taking shelter.

Niche overlap The proportion of available resources that are shared by two species. Usually used in the context of a single resource that limits population growth.

Contrasting the Habitat with the Niche

As is the case with many ecological concepts, defining habitats and niches concisely and unambiguously is difficult. Habitat definitions are particularly prone to problems of scale. At the larger end of the scale is the biogeographical term *biome*. Some biomes may be considered the habitat of larger, well-traveled species such as large birds of prey, but usually the habitat is defined more narrowly. On the smaller end of the scale, the term microhabitat is used to describe the places where an organism spends part of its time. For instance, a fish may forage in the microhabitat that occurs near the banks of a river. Thus the concept of the habitat lies somewhere in between the biome and the microhabitat.

An ecological community is a suite of species that occur in the same location that at least have the potential to interact and affect each another either directly or indirectly. These effects are usually measured in terms of population growth or changes in the density of individuals. Different communities of organisms will generally reside in different habitats. If most species in one community have little chance of affecting the populations of other species in a different community, then these communities probably occur in different habitats. In this way, microhabitats refer to places within a community, and biomes encompass a number of different habitat types. Definition of habitat and community boundaries will depend on the exact species under consideration. For instance, it may be reasonable to distinguish the small mammal communities of grassland from nearby riparian forest, though an individual jaguar may prey on species in both habitats.

The habitat concept was originally applied to single species, but because many species share similar habitats, some general habitat descriptions are applied to many species. Habitat descriptions used in this fashion often incorporate the dominant species commonly found in that habitat. In nature, borders between different kinds of habitats are often not very distinct. Generally, one habitat type will gradually give way to another creating a zone of transition. Communities in the transition zone may be qualitatively different from communities that occur within one habitat type or another.

The primary habitat preferences of a species can be displayed graphically. Ordination is the process of characterizing a species with regard to habitat gradients. Gradients can be displayed as axes on a graph. Usually, some measure of density or population growth is plotted, and the portion of the gradient that is occupied by the species is referred to as the habitat breadth of the species. **Figure 1** shows how two species of plankton can be compared according to a single habitat dimension, depth in the water column. There are potentially limitless gradients or factors that may be involved in characterizing the habitat of a species. Graphical representation of more than three variables in a single graph is difficult, but the concept easily extends into multiple dimensions.

The niche is a more abstract term than the habitat. It encompasses all possible interactions that a species has with the environment and other species in the community.

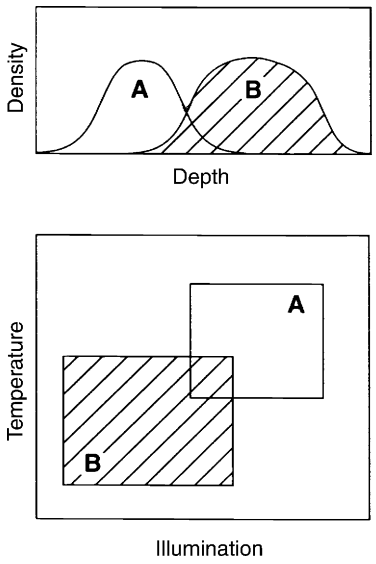


Figure 1 Two hypothetical species of plankton compared along the habitat axis of depth in the water column (top panel). This single habitat axis is correlated with two different niche axes, temperature and illumination. Species B is found in shallower, brighter, and warmer areas.

Conceptually, the niche is richer than the habitat and forms the foundation for much ecological theory. Because of the abstract nature of the concept, the definition is perhaps even more difficult to formulate for all species. There are striking similarities in the way niches and habitats are displayed. Like a habitat, a niche is often defined by axes, and there are multiple axes that can be compared.

The most common type of axes used to characterize a niche include those that define environmental tolerances and resource requirements. Humidity, sunlight, temperature, wind exposure, and pH are examples of environmental axes that help to define a niche. Resource axes may include food resources (insects, seeds, bacteria, nutrients), space requirements (breeding sites, refuges, foraging zones). **Figure 1** also compares the two plankton species in two niche dimensions, illumination and temperature, which are correlated with the habitat axis, depth. This highlights the similarity of habitat and niche axes, which are often difficult to distinguish clearly.

In practice, ecologists focus on a subset of niche dimensions that are the most important in determining the role of a species in the community. Among these are food resources that are essential for survival. These factors *limit* the population size of the species and therefore determine the distribution and abundance of the population.

In spite of the more limited role of the habitat in theory, the physical nature of habitat axes make them useful tools for practical purposes. Habitats can often be adequately described with relatively few variables, whereas to fully document the niche of a species requires careful study and a great deal of time. Yet if the habitat can be accurately captured, it is likely that many of the complex niche requirements and interconnections will be contained within the same habitat. As stated earlier, communities tend to map one to one onto habitats. The utility of this relationship is exploited for management when there is insufficient time to study all aspects of an organism's ecology before an action needs to be taken. To preserve a single species, one must also preserve its niche requirements and the complex interconnections with the rest of the community. Preserving adequate habitat will often achieve this goal. This principle is important not only for managing target species, but for preserving and restoring biodiversity in general.

The primary habitat and niche for a number of species is given in **Table 1**. These brief characterizations are typically used as a shorthand way to describe where to find species and the basic role played by each in the community. Habitat descriptions include geographical information and often refer to the dominant species present (e.g., grassland, coral reef). Niche descriptions for a species usually begin with the main food items consumed.

As one moves through this list of very diverse species, the differences in the scales and roles are readily apparent. Trees may define different successional stages of a forest, each of which may be considered a different habitat. A jaguar may routinely cross into different successional stages on a daily basis, making it difficult to clearly define a jaguar's habitat. In contrast, certain trees, birds, and lichen are limited to only a certain successional stage, while specialized insects, parasites, and pathogens may even be limited to a single species of host. These differences make broad comparisons of niches and habitats difficult, but at any single scale, these concepts can be applied to facilitate meaningful comparisons among similar species.

History of the Habitat and the Niche

Niche Origins

It is tempting to think of a niche as a physical place. This is the common usage of the word, and there are examples in the early ecological literature that use the term *niche* in the purely physical sense. However the origins of the ecological niche reside in the more general observation that no two species are exactly alike.

Many naturalists in the latter part of the 19th century turned their attention toward documenting the traits that distinguish one species from another. When two species appeared very similar, it was thought that eventually differences could be found that would distinguish the unique role of each in the community. This idea was evident even in the writings of Darwin, and over time it has evolved into a very general principle: the principle of competitive exclusion. Exploring the implications of this relatively simple notion has dominated the study of ecology for much of the 20th century.

The principle of competitive exclusion and the concept of the niche developed in parallel through the early 1900s. Theoretical work by Volterra (1926) showed that if certain assumptions are met, only one species should be able to survive on a single resource. Gause (1934) demonstrated this principle experimentally with two species of *Paramecium* feeding on a common resource. Two species that consume the same resource in the same way simply cannot coexist.

Table 1 The primary habitat and major niche components of some representative species

Species	*Habitat/Biome*	*Niche*
California thrasher *Toxostoma redivivium*	Coastal chaparral; western North America	Generalist insectivore; territorial; sometimes a fruigivore
Cuban crown giant lizard *Anolis equestris*	Tropical forest tree crowns; Caribbean (Cuba)	Generalist insectivore; short pursuit predator; territorial
Spotted salamander *Ambystoma maculatum.*	Ephemeral ponds in deciduous forest; temperate North America	Generalist carnivore, breeds in ponds with few large fish
Jaguar *Panthera onca*	Tropical forests of all types, savanna; Central and South America	Small animal carnivore; solitary stalking predator, keystone predator
Plains zebra *Equus burchelli*	Tropical savanna, Africa	Grazer of short grass; gregarious; migratory
Harvester ant *Pogonomyrmex rugosus*	Sandy regions of deserts; southwest North America	Scavenger of small seeds; makes terrestrial, colonial burrows
Tawny mining bee *Andrena fulva*	Temperate fields, gardens; Europe	Nectar/pollen harvester; solitary soil burrower; pollinator
Camel cricket *Ceuthophilus silvestris*	Caves, moist environments; North America	Scavenger, detrivore
Ghost crab *Ocypode quadrata*	Sandy marine beaches; temperate and tropical Nearctic	Scavenger; territorial; nocturnal
Mussel *Mytilus californianus*	Intertidal, rocky marine shores; western temperate North America	Filter feeds on Plankton; sessile; pelagic larvae
Tardigrade (water bear) *Echiniscus spp.*	Moist film on plants, moss, or lichen; cosmopolitan	Plant and moss cell predator; meiofaunal (0.4-1.0 mm)
Guanacaste tree *Enterolobium cyclocarpum*	Tropical savanna, dry forest; Central America	Large canopy tree; pioneer, keystone species
Flowering dogwood *Cornus florida*	Temperate deciduous forest, North America	Shade-tolerant, understory; pollen and nectar producer
Deuteromycota fungi *Arthrobotrys spp.*	Moist soil, pools; cosmopolitan	Sit-and-wait predator of roundworms, hoop-snare structures
Slime mold *Acrasiae spp.*	Moist soils, various habitats; cosmopolitan	Bacteria consumer, single and multicell stages
Schistosome *Schistooma mansoni*	Humans and freshwater Snails; tropical, Africa, South America	Blood parasite in humans, gut Parasite in snails

The term niche was first used in the context of competitive exclusion by Grinnell, a superb naturalist of the North American west, as early as 1914. Grinnell's study of the California thrasher represents one of the first that specifically set out to characterize the ecological niche of a species, though he laid out the concepts of the niche and competitive exclusion 10 years earlier in a study on chickadees.

Grinnell's study of the California thrasher illustrates some of the most basic components of a niche (a) type of food consumed (mostly insects, berries at some times of the year), (b) microhabitat preference (beneath shrubby vegetation), (c) physical traits and behaviors used in gathering food (a long beak thrashed through the top layers of soil and leaves), and (d) resources required for shelter and breeding (dense shrubs for night roosting and nesting). These four basic factors allow one to characterize the basic niche of most animals, and most animals differ with respect to one or more of these factors.

An analogous set of core niche dimensions may be constructed for plants, which tend to partition niche space along resource axes such as available light, soil moisture, and various soil nutrient gradients. Important niche dimensions for marine organisms may include temperature, substrate type, salinity, pH, and exposure to waves.

Thirteen years after Grinnell, the publication of Elton's *Animal Ecology* text (1927) established the term niche in the lexicon of ecology. However it was not until the work of Hutchinson (1957) that the concept of the niche was fully developed as a cornerstone of ecological theory. Hutchinson's niche is an *n*-dimensional hypervolume, implying that there are usually many factors, or dimensions, that define the role of

a species in a given habitat. These included yet extended well beyond the basic niche dimensions described by Grinnell.

Other ecological concepts emerged from this foundation. The resources that a species would use if it were isolated from all potential competitors is part of its fundamental niche. The realized niche is a subset of the fundamental niche that includes the resources actually consumed by the species in nature. The difference between the realized and fundamental niche can be attributed to competition from other species in the community.

Habitat Origins

Originally, the habitat of a species was not viewed as a part of its niche. In the early history of these terms they were treated separately. The habitat was seen more as a guide to the kind of community in which the organism played its role. For instance, the California thrasher occupies coastal chaparral of western North America, but this habitat description can be viewed as distinct from its niche within that habitat. Also, species were viewed as possessing suites of traits and physical tolerances that enabled them to be suitable only for specific habitats.

Over time, habitat axes were compared right alongside common niche axes, and the distinction became obscured. This change may be due in part to parallel development in the field of evolution and adaptation and increased knowledge of the processes involved. Today it is not difficult to imagine that two species may have evolved to occupy different habitats, whereas at a previous time only one species may have occupied both habitats.

The history of the habitat concept can also be traced back to the work of Grinnell and his contemporaries. In one of the earlier studies, Grinnell and Storer characterized the animals of Yosemite according to the elevations at which they occurred. In this mountainous region, species tend to be found in narrow elevational zones or bands.

At about the same time as the Yosemite study, Ramensky wrote that each species was unique and possessed adaptations that enabled it to tolerate a unique set of environmental conditions. This "principle of species individuality" bears strong resemblance to the early influences on the niche concept, except that in the habitat context, competitive exclusion was not at the core. Instead, species were seen to be adapted to certain conditions and habitats, and they independently sorted themselves out accordingly.

Subsequent work has confirmed that habitat boundaries are often determined by physical tolerances. However there are numerous examples where competition and other species interactions play an important role in determining the habitat breadth of a species in nature. This further illustrates how the habitat and the niche have converged over time. Therefore it is appropriate to treat the habitat and the associated gradients and axes as a subset of the niche.

Comparing Habitats and Niches

Multispecies Comparisons

The habitat and niche of any species can be characterized independently, but the true value of these concepts lies in comparing multiple species. The ability to display habitats and niches graphically greatly facilitates comparisons. **Figure 2** shows how species can be ordered along a habitat gradient. Copepod species show different patterns of abundance with respect to the gradient from ocean to land in the intertidal region. While in this instance there is a large amount of overlap, each species has a characteristic distribution in this region.

Habitat boundaries are rarely sharply defined. *Ecotone* is the term used to refer to this zone of transition between habitats. Because this zone is somewhat different from either neighboring habitat, an ecotone may contain several unique

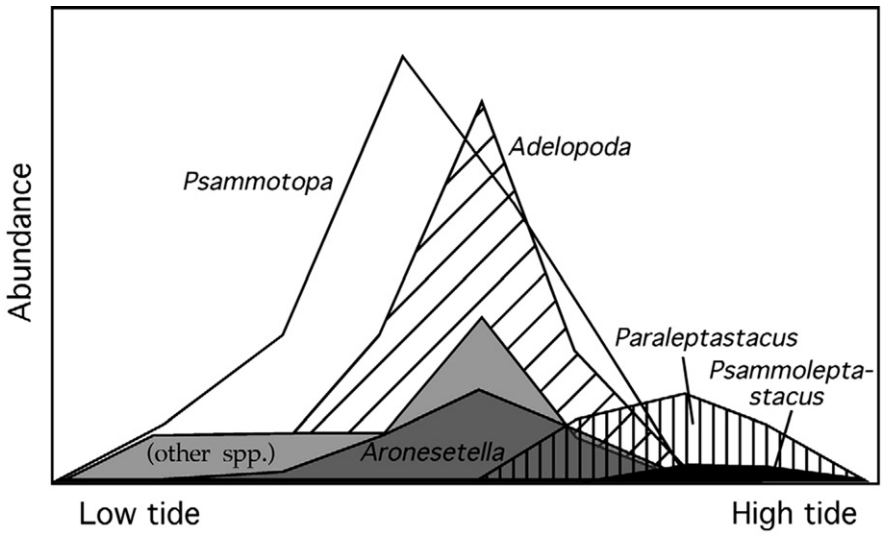


Figure 2 The distribution of copepod species along the sandy intertidal habitat gradient in Massachusetts (only genus names are given). Reproduced from Pennak RW (1951) Comparative ecology of the interstitial fauna of fresh-water and marine beaches. *Ann. Biol. Sér.* 3: 462–463.

species. This concept was proposed to help explain distributions of species that did not appear to coincide with habitat boundaries. In a unique test of this idea, Terborgh and Weske (1975) found little evidence that subsets of bird species along an elevational gradient were responding to ecotones. Instead, it appeared that the elevational ranges of many species were independently defined by the presence of other closely related species. In this case, competition appeared to be more important than habitat gradients for setting the actual limits of species distributions.

The niche framework is typically used to compare groups of closely related species that consumed similar resources but differed in one or a few important ways. Suites of species in a community that utilize the same general resource were termed guilds by Root (1967). Just as a guild of blacksmiths all work with iron yet have different specialties, assemblages of animals that specialize on seeds, fruits, or insects all have different ways of harvesting these resources.

Early studies focused on guilds that differed mainly with respect to body size. Hutchinson (1959) observed that species within a guild often differed by a minimum size ratio. Known as Hutchinsonian ratios, these regular arrangements of species body sizes occur in a number of vastly different communities consuming very different resources.

One example of the kind of pattern Hutchinson was attempting to explain is the assemblage of rodents that consume seeds in the deserts of North America. **Figure 3** shows how similar assemblages of rodents in different communities

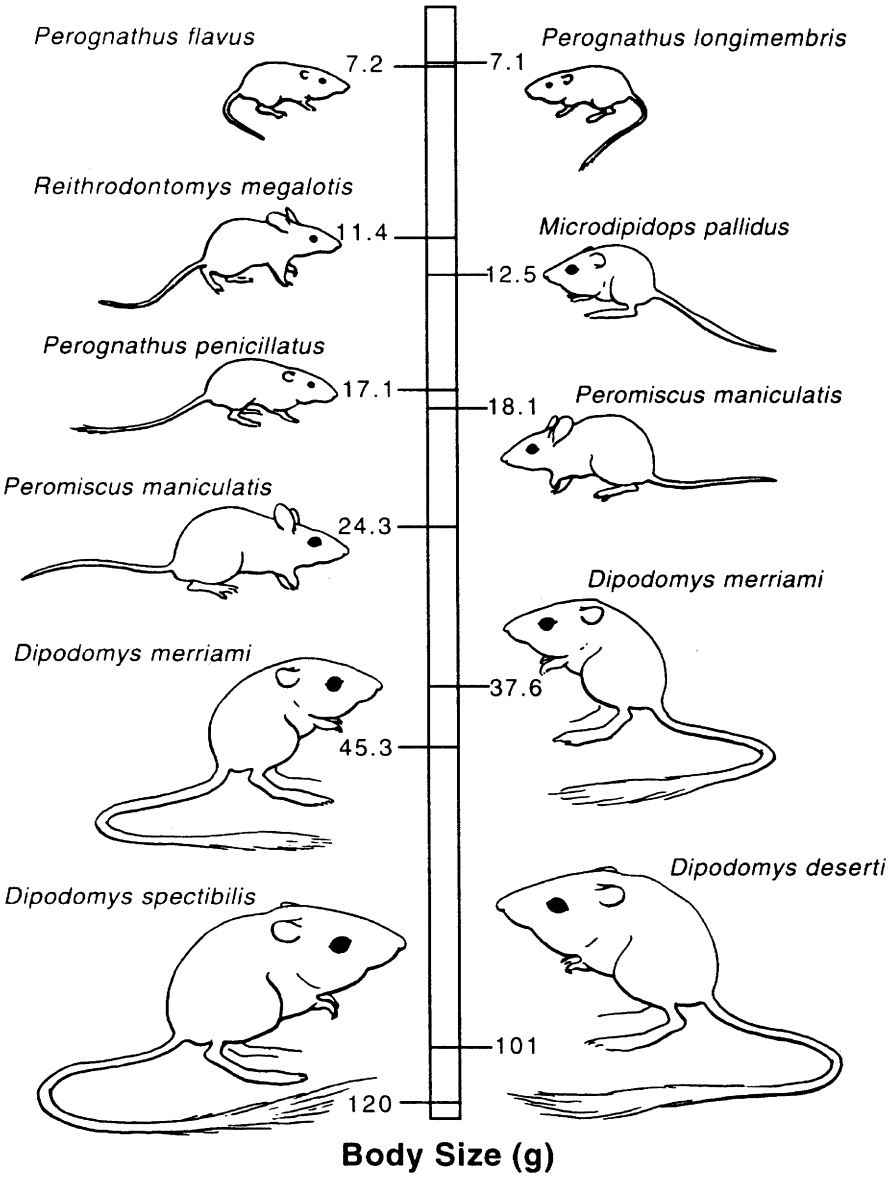


Figure 3 A comparison of the niche relations among Great Basin and Sonoran desert rodent guilds. Different habitats support a different assemblage of rodents that show similar differences in body size. Reproduced from Brown JH (1975) Geographical ecology of desert rodents. In: Cody ML and Diamond JR (eds.) *Ecology and Evolution of Communities*, pp. 315–341. Cambridge: Harvard University Press.

contain a similar set of rodents. Each species differs from others in the same community mainly in terms of body size. Convergence is a term applied to organisms or communities that are not closely related historically that have members with similar physical features because they occupy similar niches. The implication is that parallel evolutionary trajectories were followed independently by different species because of similar selection pressures.

The underlying mechanisms that create such regular patterns can be explained because each species consumes resources of a size class that is proportional to its body size. Smaller rodents consume smaller seeds, while larger rodents eat larger seeds. In a different community, smaller fruit pigeons consume smaller fruits, while larger species primarily consume larger fruit. For a medium sized bird, small fruits take too long to gather and handle, and the bird can not eat them efficiently enough to sustain itself. Larger fruits may be too big to handle or crack open.

Many types of food resources such as seeds, fruit, or arthropods have a wide, flat distribution along a size gradient. These relationships can be displayed graphically with a single niche axis, food resource size (**Figure 4**). One curve represents the available resources, in this case seeds, and the narrower curves underneath represent the sizes and amounts of seeds consumed by each species. This division of available resources among members of a community is termed resource partitioning.

Utilization curves are often bell shaped (normal, Gaussian) to reflect the tapering ability of each species to consume seeds away from the "optimum" size. This type of niche representation allows direct comparison of the amount of overlap in resource consumption. If food is indeed the primary limiting resource, the overlap of species resource utilization curves directly represents the amount of competition that each species is experiencing. It is the regular spacing of curves to minimize overlap, or competition, that creates the pattern of regular size distributions of animals seen in natural communities. This simple framework forms the basis of niche theory.

The concepts of fundamental and realized niches can be clearly visualized within this framework of graphical representation of both habitat and niche dimensions. Mussels are capable of living in a wide range of elevations above the low tide mark, and they will move into these spaces in the absence of other species. However other animals, such as barnacles, are better adapted to occupy lower positions and can outcompete mussels in some zones. That is why in natural communities, the realized niche of mussels in terms of this habitat gradient is usually narrower than the fundamental niche.

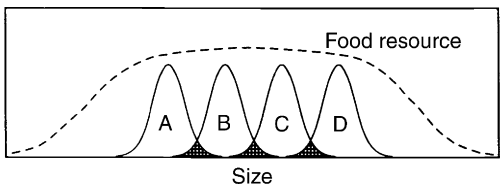


Figure 4 A hypothetical species assemblage. A nearly continuous spectrum of resources that differ mainly in size (top curve) are consumed and partitioned among different species that also differ mainly in size (A–D).

A similar process may be operating in the granivorous rodent example presented earlier. Each rodent may be capable of consuming a wider range of seeds, albeit inefficiently. In nature, either there are fewer of these suboptimal resources available because of competitors, or each species deliberately chooses a narrower range of seeds to consume to avoid competition.

The effect of competition on realized niches can be seen by comparing island and mainland communities. Islands usually support fewer species than equivalent mainland habitats. It has been shown for many kinds of species that in these species-poor island communities, species tend to eat a wider array of foods and therefore have wider niches. This agrees with the expectation based on niche theory outlined earlier: there are fewer potential competitors on islands, and therefore there is less resource overlap with other species acting to constrain their niches.

Species that have very high overlap in many niche dimensions often have little or no overlap in one key dimension. This principle, called niche complementarity, lies at the heart of the guild concept. In previous examples, similar species consumed the same general resource and had high niche overlap, but they showed little overlap with respect to a single dimension: body size. Typical niche axes where complementary species can be found include size, light (nocturnal/diurnal), benthic/pelagic, and canopy/ground or coniferous/deciduous microhabitat.

In many instances niche differences among species are clear, but sometimes niche differences are extremely subtle and require careful study to be revealed. For instance, five species of *Dendroica* warblers can all be found in the very same tree in temperate zone coniferous forest habitat. All glean similar types of prey, primarily lepidopteran larvae, from the surfaces of the tree. Only careful study by MacArthur (1958) revealed that each species favors different parts of the tree, and each tends to move about the tree in different ways (**Figure 5**). Thus it is not only the type of food resource that is limiting the population growth of each species, but where and how that food resource was harvested. There are enough subtle differences in microhabitat use and behavior to allow coexistence of species that at first glance appear to share equally the same limiting resource.

Niches can change through time. For instance, the larval forms of many organisms often have completely different niches from adult forms. Thus, the term ontogenetic niche has been applied to amphibians, arthropods, and a wide variety of marine organisms. The great versatility of the niche concept allows incorporation of these temporal changes by simply adding a temporal axis to the species niche. The realized niches of species in a community can also change through time if the community undergoes change. Extinctions and invasions cause rapid community changes that in turn result in niche shifts among other community members.

Niche theory is the branch of ecology that has taken these informal graphical models and extended them mathematically to model the process of competition and community organization. The origins of niche theory can be traced back to the early work of Volterra, and has grown and undergone dramatic

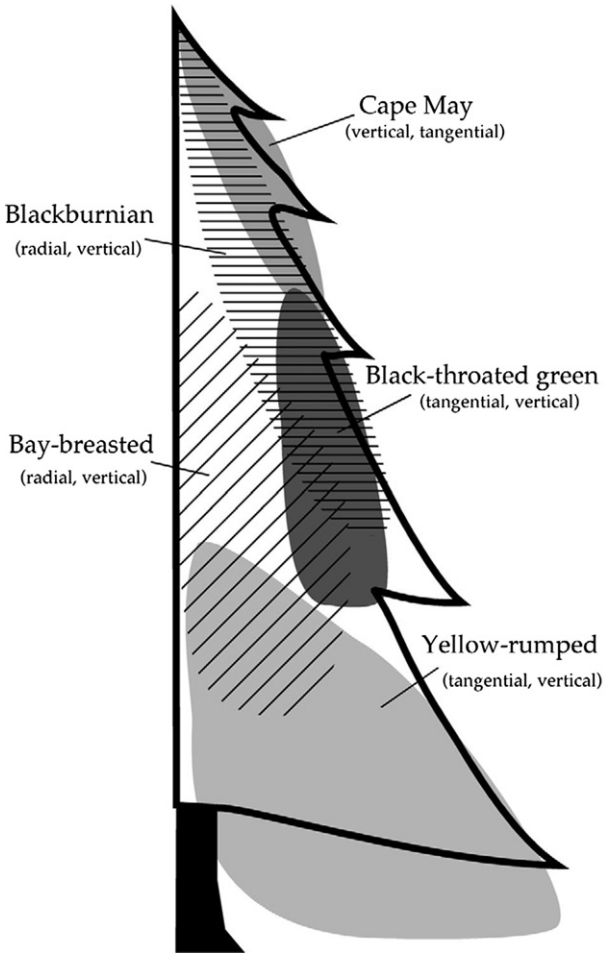


Figure 5 Niche relations among some North American warblers that forage in coniferous forests. Species differ according to microhabitat (parts of the tree most frequently visited) and foraging behavior. Reproduced from MacArthur RH (1958) Population ecology of some warblers of Northeastern coniferous forests. *Ecology* 34: 599–619.

changes. Complex interactions among members of a community are very difficult to measure, and the approach of studying theoretical models has greatly enhanced our knowledge of the factors that influence community dynamics.

Other Niche Dimensions

The classical context of niche comparisons is mainly restricted to niche axes characterizing habitat and food resources. Some ecologists have even considered the definition of a niche to be limited to the food resources consumed. However developments in the field of ecology through the 1900s have made it clear that an expanded definition of the niche that encompasses all the interactions of a species within a community is most appropriate. A broad definition of the niche includes all potential interactions that ultimately have the effect of changing the population density of another species in the community.

Mutualistic interactions are somewhat different from food resources in that they are often not literally consumed, nor are they occupied as in a nesting or sheltering place. However in niche descriptions, mutualistic interactions are treated much the same as other resources in that they may be limiting and therefore may even be competed for. For instance, pollinators interact mutualistically with angiosperms, yet from the point of view of a plant, they may be regarded as any other resource.

Indirect Interactions

Competition for resources has historically played a central role in niche descriptions. More recently, studies have uncovered processes where two or more species coexist because of factors that have little to do with resource niche axes. For example, predation can act to promote coexistence of potentially competing species in the following way: The presence of a predator may reduce the numbers of one prey species and thereby allow other species that have similar resource requirements to coexist. If two or more species are all harvested by a predator that does not show preferences for any specific prey type, then the most common prey will experience a disproportionate amount of predation. In this way, the presence of a predator can actually promote coexistence of species and an increase in biodiversity. Predators that have this effect have been referred to as keystone species because of their disproportional affect on community diversity.

An excellent example of this process can be found in the rocky intertidal communities of the eastern Pacific. Paine (1980) and colleagues conducted a unique set of experiments demonstrating that the presence of a starfish predator (*Pisaster*) enabled the coexistence of a number of species, whereas in the absence of the starfish, the community was dominated by the California mussel. The niche description for a barnacle in this community would therefore be incomplete without incorporating information about not only the presence of the California mussel, but also the relative vulnerability of each of these species to starfish predation. This implies that the niche characterization may change depending on the presence of potential predators and prey and their specific traits.

Predation certainly can be seen as a direct interaction, but the example above falls into the class of indirect interactions. A keystone predator can enhance the abundance of a species by interacting with a third species. Many interactions may cause ripple effects through a community in indirect ways. Elephants change the physical structure of their habitats, and many species excavate holes and nests that are used by other species. Parasites and pathogens can also have indirect effects on entire communities.

Viruses and other pathogens may initially appear to occupy a very simple niche, especially those that are confined to live entirely inside the body of a host species. How can such an organism play an important role in the community through indirect interactions? Pathogens can significantly impact the population size of the host species. This change in abundance may have consequences for other species, such as predators and prey of the infected species.

Some pathogens have complex life cycles that depend on more than one host species, and sometimes these hosts are very different. For instance, in using humans and snails at different stages of its life cycle, Schistosomiasis provides a clear

link between humans and some snail species that are used as hosts for different stages of the life cycle.

Another interaction involving pathogens provides an example of apparent competition. One host species may be relatively unaffected by a pathogen, whereas other species may be severely impacted. In this case, the unaffected species can act as a carrier, spreading the pathogen to other species in the community that are more susceptible, causing their decline. On the surface, this pattern may resemble competition.

Community ecologists have used various methods to incorporate these indirect interactions into model representations of communities. It is certainly not as straightforward as simple resource utilization graphs. Mathematical matrices have been applied to study the dynamics of communities. In this format, all pairwise interactions can be explicitly considered, and broad concepts such as the relationship between community stability and complexity have been studied within this framework. Recently, web theory has been applied to this task because complex interactions within communities can be more explicitly constructed. All of these theoretical constructs of communities are based on the niche concept, thus the utility of the niche concept has grown even as the field of ecology has undergone dramatic changes.

Evolution of the Niche

Ecological studies can help us to understand *why* species occupy unique niches, but *how* did the great diversity of species arise in the first place? The processes underlying the evolution, adaptation, and coexistence of species that have unique niches is a fundamental aspect of understanding biodiversity.

For many organisms the primary mode of species formation, and therefore the primary generating force for biodiversity, is thought to involve differentiation of isolated populations. A terrestrial species may have populations separated by water, or coastal marine populations may be isolated by unsuitable open ocean. Over time, these populations undergo changes and begin to accumulate different characteristics. Some of these differences are due to chance events such as mutation and genetic drift. Traits change over time as the organisms adapt to the unique features of their respective habitats. Subtle environmental differences may result in the evolution of thicker shells in a population of crabs, or a preference for fish over mammals in a population of killer whales.

This differentiation by adaptation to different habitats with different food resouces is in itself one engine for generating biodiversity, however it cannot explain the vast diversity of species that occur within a single habitat. Eventually, populations that have been isolated for a while may reestablish contact, and individuals will attempt to coexist within the same community. It is at this stage of speciation, referred to as secondary contact, that one can see how the concept of the niche is fundamentally related to species diversity.

When secondary contact occurs, one of three outcomes is possible. First, populations may interbreed, and speciation is not completed. Second, if populations do not interbreed, they may either coexist, or, third, one may outcompete the other. Which of these two latter outcomes is realized will depend on the amount of niche overlap between these species. If overlap is high, then one species is likely to outcompete the other. If overlap is low enough, the species may coexist.

In situations where coexistence is possible (or at least extinction is not very rapid), and the species involved still possess similar traits, we would expect evolutionary changes to occur over time. Individuals that consume limiting resources that are shared with another species will experience more competition and may survive or reproduce less than individuals that do not overlap as much. If there is natural variation in resource consumption among individuals in each population, and if the traits that cause this variation are genetically passed to offspring, then the species are expected to diverge over time to minimize niche overlap. This process is called ecological character displacement, and it is a logical extension of the competitive exclusion principle into evolutionary time. **Figure 6** illustrates the process of character displacement. It is the change in a trait of a species, such as body size, that leads to a shift along a niche axis and reduction of overlap.

The empirical evidence for character displacement lies mostly in descriptive, snapshot-like comparisons of species. Typically, closely related communities are compared and character displacement is inferred if the presence or absence of one species is correlated with a change in the traits of another species. These studies infer that it is niche overlap in the use of a single limiting resource that has led to the observed species differences. Lack (1947) was one of the first to infer character displacement in this way in his study of Darwin's finches, **Figure 7**.

Static character differences constitute indirect evidence for competitive processes and have therefore received criticism. To document evolution directly in natural populations is a difficult task because even rapid evolutionary changes may require generations to occur. Nevertheless, there is solid empirical evidence for niche evolution in nature. The Grants and their colleagues (see Grant, 1999) have shown the niche of Darwin's finches evolved over time according to the seeds

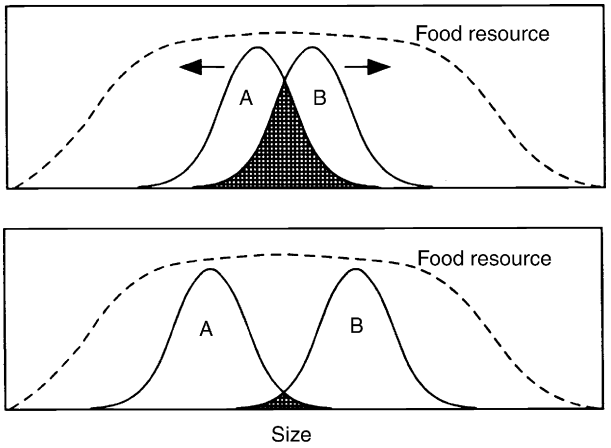


Figure 6 A graphical representation of ecological character displacement. Initially upon secondary contact, species have highly overlapping resource utilization. Over time (b) resource overlap and competition is reduced through evolutionary changes that result in a shifting of the niche.

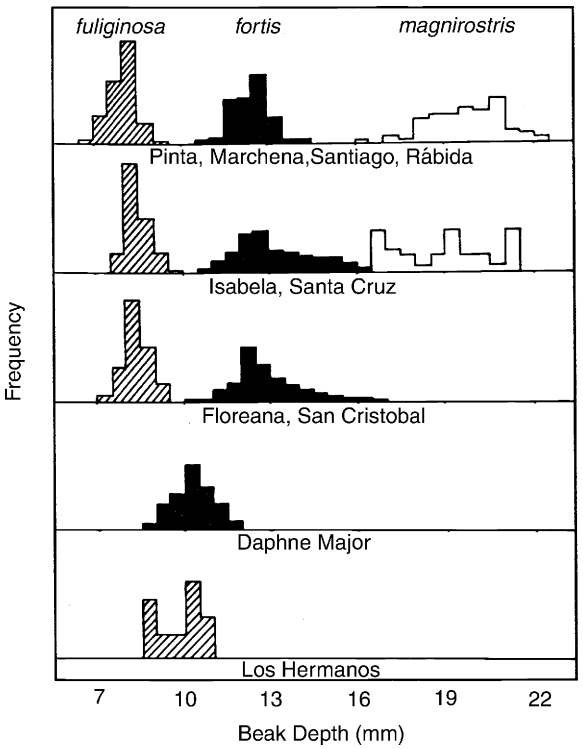


Figure 7 Size variation among three species of Darwin's finches (genus *Geospiza*) that specialize on seeds. Individuals are pooled among islands (given) to generate the histograms of beak depth. Populations on islands with fewer species (lower two panels) are intermediate in size compared to islands with the complete complement of species (upper panels). This is indirect evidence for ecological character displacement caused by competition for seeds; Pennsylvania in the United States. Reproduced from Grant PR and Schluter D (1984) Interspecific competition inferred from patterns of guild structure. In: Strong DS, Simberloff D, Abele LG, and Thistle AB (eds.) *Ecological Communities*, pp. 201–233. Princeton: Princeton University Press.

that are available. In unique experimental tests of character displacement, Schluter (1994) has shown that stickleback fish populations changed over time in response to competition and niche overlap. Natural selection acted against individuals with feeding anatomies that caused them to harvest resources that were shared between species.

Ecologists may differ in opinion about the importance of character displacement in specific instances, and there are surely other very important forces that act to shape community composition and species traits. Nevertheless, few would doubt that character displacement is an important fundamental process in shaping communities of organisms.

The Habitat and Biodiversity

We have traced the habitat concept from its independent origin through its incorporation into the concept of the niche. With growing awareness of the need to conserve biodiversity, the habitat concept has emerged again as an important tool for management. Essentially, habitats contain functioning communities of organisms. If the goal is to protect a focal species, then setting aside appropriate habitat will likely result in preserving the required niche elements of the species as well.

Assessing appropriate habitat has two great advantages over attempting to define niches. First, because of the physical nature of habitats, they are easier to recognize and quantify. Second, ecological communities tend to fall along similar boundaries as habitats, thus by preserving a habitat one is more likely to preserve communities with all niche interconnections intact. This relationship among species, communities, and habitats has become the center of the more recent conservation efforts termed habitat conservation plans.

For example, in San Diego county, a number of endangered or threatened species rely on a unique type of habitat: chaparral, or coastal sage scrub. A major goal of habitat conservation plans in this region is to create contiguous regions of appropriate habitat. In this way, the needs of many species can be met at once.

Habitat fragmentation is a major concern because small refuges of habitat are essentially islands. Islands are prone to a number of processes that degrade biological diversity. Small populations on islands are prone to extinction through random fluctuations in numbers. The geometry of islands gives them a high ratio of edge to interior. Because animals may wander across this boundary, those that rely on the interior island habitat experience more competition and predation from those outside the island. More isolated islands are also less likely to be recolonized in the event that a population has gone extinct. Therefore, natural reserves and habitat conservation plans pay particularly close attention to fragmentation and try to consolidate larger parcels of unbroken natural habitat.

Once a habitat has been severely degraded, there may be ways to return it to its former natural state. Habitat restoration is a challenging field, but one that has made great strides recently. Wetlands have been restored by controlling of the water table and planting appropriate vegetation. Foresters have been restoring habitat for decades by replanting harvested areas. Highway construction projects aim to restore natural vegetation upon completion of a project by spraying a mix of grass, flower, and shrub seeds in a fertilizing matrix onto the bare ground. Another interesting example of habitat restoration can be seen with the construction of artificial reefs. By placing physical structures on the seafloor, a reef community can take hold and flourish. This is an example of how subtle environmental features such as physical structure can have a profound effect on habitat quality and species diversity.

Two general mechanisms act to promote species diversity within structurally diverse habitats. First, biological species diversity leads to more species diversity. A monoculture such as an orchard allows for relatively little niche differentiation of the organisms within. There is little opportunity for species to adapt to specific host species, and this applies to insects, birds, mammals, and parasites. Reduced numbers of species limits the number of possible niches in an orchard when compared to a tropical rain forest.

Artificial reefs belong to a second class of mechanisms that generate biodiversity by acting over and above the biological mechanisms. These are the purely physical features of the

habitat that allow for increased niche differentiation. Dead logs, snags, distinct ground, and shrub layers will provide habitat heterogeneity that allow higher trophic levels to differentiate. Just as warblers partition the physical space within a single tree species, other competitors that use different tactics to forage on similar prey can partition a structurally diverse habitat to a finer scale. Behavior plays a key role as foraging tactics and escape strategies of prey are honed to specific microhabitats within the larger forest. MacArthur and colleagues (1962) were among the first to correlate purely physical aspects of a habitat with species abundances. Since then, physical habitat structure has been shown to augment species diversity in systems as diverse as birds, lizards, stream invertebrates, fish in marine reefs, and microorganisms in beakers.

Keystone species have a disproportional effect on many other species in the environment, and as such are often the focus of conservation efforts. Starfish foraging in the intertidal zone have a positive effect on biodiversity. Jaguars prey on a number of small rodents and animals, and they tend to capture prey in proportion to their abundance. The net effect is to prevent any single species from becoming very common and outcompeting other prey species. The niche of the jaguar and the starfish includes this community-wide role, and they are sometimes referred to as keystone predators.

Management strategies have turned more attention toward keystone species in order to preserve communities. A similar keystone role is played by tree species such as the Guanacaste tree of Central America. In this instance the link between the niche, the habitat and biodiversity is clear. The Guanacaste tree is vital for establishing forest in grassland habitats, and it therefore plays a key role in restoring tropical dry forest. The niche of this species would include its role in developing habitat structure and therefore intersects with the niches of many other species in the community.

The habitat and the niche can interact synergistically with human activity to cause extinctions and loss of biodiversity. Niche complementarity has led to a very general pattern whereby species with similar niches are spatially segregated. Increased human movements on a global scale has had a dramatic effect in bringing these species into contact by introducing exotic species into new geographic regions. If these alien species have high niche overlap with resident species, the resident species can be driven to extinction by a successful invader. Even though there are many cases where introduced species can not establish or outcompete residents, many threatened and endangered species are at risk because of invaders. This threat also extends to introduced predators. Resident species that have adapted to a niche in one community can be driven to extinction by a predator that has evolved in a different community.

Conclusions and New Frontiers

The concepts of the habitat and the niche have grown mainly through study of the organisms most familiar to us as humans. Most conceptual advances were derived in some way from the study of larger animals and plants. Even within these groups, there has been a disproportionate amount of study of birds. Yet there have been significant contributions from other sectors of the biological world. For instance, Gause's study of *Paramecium* lies at the heart of the niche concept. It is fitting then that more attention is returning toward understanding the niches of microorganisms.

Experimental evolution using bacteria is a growing field that holds great promise. The short generation times and relatively simple ecologies of bacteria enable evolutionary experiments on the scale of the community. This is one of the few instances where the interaction of ecological and evolutionary factors can be studied during the process of community formation.

There is an enormous number of microscopic, planktonic and meiofaunal (0.4–1 mm) organisms about which we know very little. These organisms are ubiquitous, speciose, and show an amazing amount of diversity in form. However there are significant obstacles encountered in studying these smaller organisms that even make it difficult to quantify the habitats in which they can be found. Practical problems arise while trying to observe them in nature, while on the other hand, many fail to survive the transfer to laboratory environments. Yet there are compelling reasons to learn more about their ecology.

The niche of a terrestrial vertebrate or marine invertebrate is often defined largely by how it feeds or defends itself. In turn, anatomical structures have evolved that reflect these niche differences. Similar examples can be found in the nearly invisible world of small creatures around us. For instance, ciliates and other protozoans possess anatomical structures that suggest a diversity of roles even greater than those observed in larger organisms, yet there has been comparatively little study of their niches in nature.

Some organisms stretch the concepts of the niche and habitat to their limit. Some plants display an extreme amount of what is referred to as phenotypic plasticity: different individuals develop extremely different body forms depending on the environmental circumstances. Some ciliates have the ability to radically change the form of feeding or defensive appendages even within the lifetime of an individual.

Many bacteria, and even some larger meiofauna like the bearlike tardigrades, can enter a sporelike state to withstand extreme environmental conditions. They may persist for long periods of time in this state and may be very difficult to detect. Much like the seed bank of desert plant communities, these alternative states pose difficult problems in understanding the dynamics of their communities.

The habitat and the niche have grown, developed, and been applied with great success. Observation and theory have played important roles throughout, and experimental investigation has undergone a resurgence. Yet it is intriguing to ponder that most of this development has included only a subset of the biological world. Only time will tell whether the concepts of the habitat and niche will persist as the complex interactions of these diverse communities of small creatures are revealed.

See also: Conservation Biology, Discipline of. Differentiation. Diversity, Community/Regional Level. Ecology, Concept and Theories in. Guilds. Species Diversity, Overview

References

Brown JH (1975) Geographical ecology of desert rodents. In: Cody ML and Diamond JR (eds.) *Ecology and Evolution of Communities*, pp. 315–341. Cambridge: Harvard University Press.

Elton C (1927) *Animal Ecology*. London: Sidgwick & Jackson.

Gause GF (1934) *The Struggle for Existence*. Baltimore: Williams & Wilkins.

Giller PS (1984) *Community Structure and the Niche*. London: Chapman and Hall.

Grant PR (1999) *Ecology and Evolution of Darwin's Finches*, second printing. Princeton: Princeton University Press.

Grant PR and Schluter D (1984) Interspecific competition inferred from patterns of guild structure. In: Strong DS, Simberloff D, Abele LG, and Thistle AB (eds.) *Ecological Communities*, pp. 201–233. Princeton: Princeton University Press.

Grinnell J (1917) The niche-relationships of the California thrasher. Auk. 34: 427–433.

Hutchinson GE (1957) *A Treatise on Limnology. I. Geography, Physics and Chemistry*. New York: Wiley.

Hutchinson GE (1959) Homage to Santa Rosalia, or why are there so many kinds of animals. *Am. Nat.* 93: 145–159.

Hutchinson GE (1978) *An Introduction to Population Ecology*. New Haven: Yale University Press.

Lack DL (1947) *Darwin's Finches*. Cambridge: Cambridge University Press.

MacArthur RH (1958) Population ecology of some warblers of Northeastern coniferous forests. *Ecology* 34: 599–619.

MacArthur RH, MacArthur JW, and Preer J (1962) On bird species diversity II. Prediction of bird censuses from habitat measurements. *Am. Nat.* 96: 167–174.

Paine RT (1980) Food webs: linkage, interaction strength and community infrastructure. *J. Anim. Ecol.* 49: 667–685.

Pennak RW (1951) Comparative ecology of the interstitial fauna of fresh-water and marine beaches. *Ann. Biol. Sér.* 3: 462–463.

Root RB (1967) The niche exploitation patterns of the blue-grey gnatcatcher. *Ecol. Monog.* 37: 317–350.

Schluter D (1994) Experimental evidence that competition promotes divergence in adaptive radiation. *Science* 266: 798–801.

Terborgh J and Weske JS (1975) The role of competition in the distribution of Andean birds. *Ecology* 56: 562–576.

Volterra V (1926) Variations and fluctuations of the number of individuals in animal species living together. In: Chapman RN (ed.) *Animal Ecology*. Reprinted in 1931. New York: McGraw-Hill.

Whittaker RH and Levin SA (1975) *Niche: Theory and Application*. Stroudsburg: Dowden Hutchinson & Ross.

Habitat Loss and Fragmentation

Heather Bird Jackson and Lenore Fahrig, Carleton University, Ottawa, ON, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Cover type A term used by the geographers for a type of mapped feature (e.g., forest, freshwater, or grassland). The dominant natural cover type in a landscape is often used to approximate habitat.

Fragmentation per se The breaking apart of habitat while keeping habitat amount constant.

Habitat The place where an organism normally lives.

Habitat fragmentation The breaking apart of habitat into several smaller pieces. Habitat fragmentation results in both habitat loss and fragmentation per se.

Habitat loss A change to an area that prevents a species from living there (e.g., conversion of forest to crop field).

Human landscape modification An umbrella term used to describe the many complex and related features and processes correlated with human-caused habitat loss and fragmentation.

Landscape complementation The degree to which a landscape provides all necessary cover types for a species that requires multiple cover types to complete its life cycle.

Matrix An area where an organism does not usually live.

Scale of effect The spatial extent over which the habitat and the matrix characteristics in the surrounding landscape influence individual, population, or community processes in a focal area.

Introduction

Habitat loss is occurring at an alarming rate. Agriculture, the major cause of habitat loss (FAO, 2010; **Figure 1**), covers 36% of Earth's potentially suitable land (FAO, 2003). The cover type for which loss is best documented globally is forest (Balmford *et al.*, 2002). Earth's forests underwent a net decrease of 5.2 million hectares per year between 2000 and 2010 with the greatest losses occurring in tropical and subtropical woodlands (FAO, 2010). Although some forest loss has natural causes (e.g., fire, Harrod *et al.*, 1999), most of the current forest loss results from human land use (FAO, 2010). The impact of forest loss on the biodiversity is even larger than expected from the raw number of hectares because forest loss is greatest in the species-rich regions of the tropics and subtropics (Pereira *et al.*, 2010). Furthermore, it is in these areas that the most agricultural growth is expected in the future (FAO, 2003).

Habitat loss has consistent, strong, negative effects on biodiversity. Habitat loss has negative impacts on species richness (Laurance *et al.*, 2002), population abundance (Laurance *et al.*, 2002), and genetic diversity (Aguilar *et al.*, 2008). In addition, habitat loss can shorten trophic chain length; alter species interactions; and reduce successful foraging, breeding, and dispersal (reviewed in Fahrig, 2003). A combination of agriculture and hunting is the greatest perceived threat to mammal, bird, and amphibian populations (Laurance and Useche, 2009). Habitat loss is commonly cited as the greatest threat to wild bee populations (Brown and Paxton, 2009) and is second only to hunting as the major threat to marine fish populations (Dulvy *et al.*, 2003).

Habitat loss affects not only biodiversity but also impacts humans directly by decreasing production of ecosystem goods and services such as pollination (Potts *et al.*, 2010; Ricketts *et al.*, 2008), soil and water management (Bruijnzeel, 2004), and carbon storage (Fargione *et al.*, 2008). After accounting for the potential economic benefits of habitat loss (e.g., agricultural and mineral products), a conservative estimate of the global net economic cost of habitat loss is US$ 250 billion per year (Balmford *et al.*, 2002).

Habitat fragmentation, or the breaking up of habitat into smaller pieces (**Figure 2**), is a second major effect of human land use. Its prevalence is difficult to summarize because it is confounded with habitat loss and can be measured in many different ways. Understanding of the effects of habitat loss and fragmentation on populations has been hampered by a vague conceptualization of habitat fragmentation, but some broad generalizations can be made. The strongest finding from decades of research on this topic is a consistent negative effect of habitat loss on biodiversity, although the strength of this effect depends on species traits and environmental factors.

Problems with the Concept of Habitat Fragmentation

Habitat Fragmentation is a Vague Concept Because it Encompasses Multiple Concepts

The term "habitat fragmentation" has been criticized as "broadly conceived and therefore oversimplified" (Bunnell, 1999); "a conceptually ambiguous and empirically multifaceted term" (Haila, 2002); a "catchall for human-caused habitat changes that have negative effects on biodiversity" (Fahrig, 2003); and a "generic umbrella term" that is "vague and ambiguous, thereby limiting its practical value for conservation managers" (Lindenmayer and Fischer, 2007). Despite these criticisms, the term habitat fragmentation continues to be used to refer to the general effects of human landscape modification, a practice defended by researchers who credit the term with "real heuristic value in explicit recognition of an overarching domain" (Ewers and Didham, 2007).

Habitat fragmentation refers to the subdivision of habitat into smaller pieces (Andrén, 1994; Ewers and Didham, 2006; Fahrig, 2003) and as such can refer to many potentially

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00399-3



Figure 1 An aerial photo depicting a typical agricultural area near Ottawa, Ontario, Canada (OMNR, 2010). To meet the growing demand for human food, forested land (shown in dark green) has been replaced by agricultural crops (usually corn, soy, wheat, or alfalfa in this region). Photo provided by Carleton University under licence from Ontario Ministry of Natural Resources. Reproduced from OMNR (2010) DRAPE: Orthoimagery (computer file). Ontario Ministry of Natural Resources.

confounded processes and concepts. This definition can refer to both habitat loss and fragmentation per se (i.e., a change in habitat configuration), can be measured at the patch scale or the landscape scale, is used to refer to either the loss of habitat or the loss of a cover type, does not require that the multiple cover types that some species require are accounted for, may not account for the scale at which a species interacts with the landscape, and does not usually distinguish between human-caused fragmentation or natural fragmentation. Each of these areas of confusion is described in further detail.

Habitat Loss versus Habitat Fragmentation

Habitat fragmentation involves both habitat loss and a change in the configuration of habitat (Andrén, 1994; Fahrig, 2003). Use of "habitat fragmentation" to describe both processes obscures the fact that the habitat loss has a much stronger negative impact on biodiversity than fragmentation per se, which generally has weak effects that can be both positive and negative for the biodiversity (Fahrig, 2003, *see* Effects of Habitat Loss and Fragmentation: Habitat Loss is More Devastating than Fragmentation per se). The correct emphasis on habitat loss can be particularly important when management decisions (e.g., whether to increase habitat amount or reduce subdivision) are being made (Lindenmayer and Fischer, 2007).

Landscape Scale versus Patch Scale

Island biogeographic theory (MacArthur and Wilson, 1967) greatly increased interest not only in the consequences of habitat fragmentation but also led to an overemphasis on patch-level ("island") measurements rather than landscape-level measurements (Laurance, 2008). Many habitat fragmentation studies (possibly more than 50%, Fahrig, 2003; McGarigal and Cushman, 2002) measure fragmentation using patches as the unit of replication (e.g., they measure isolation and size of individual patches rather than habitat loss over the entire landscape), which can create confusion when making inferences at the landscape scale. If all patches are in the same landscape, then inference regarding the landscape scale effects of fragmentation is impaired by the lack of replication at the landscape scale (Delin and Andrén, 1999). Inferences at the patch scale cannot necessarily be scaled up to inferences at the landscape scale without explicit consideration of landscape scale measurements because qualitatively different patch characteristics can result from the same amount of habitat loss (**Figure 3**; Fahrig, 2003).

Habitat Loss versus Loss of a Cover Type

The term "habitat" is defined as the place where an organism normally lives (Ricklefs, 2008), but is often used to refer to a natural cover type (e.g., forest) whether or not the taxon of interest normally lives there. For example, some species thrive

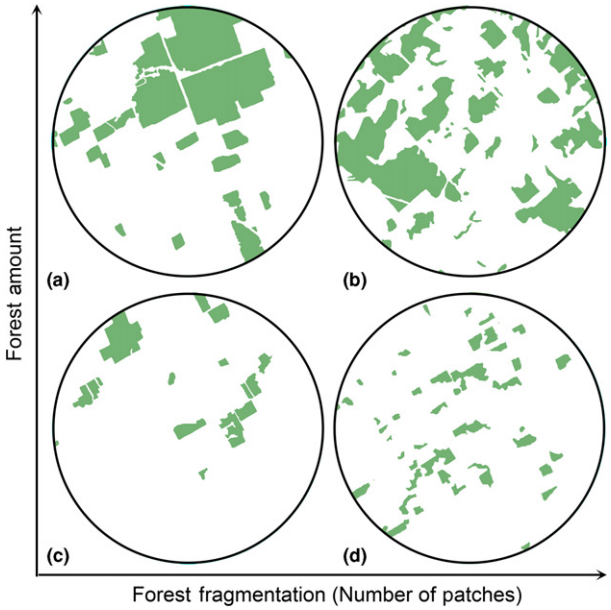


Figure 2 Landscapes near Ottawa, Ontario, Canada, which differ in the amount and fragmentation of forested habitat. Green areas indicate forest. Other habitats (agriculture, roads, and water) are shaded in white. Although most habitat loss involves both reduced amount and increased fragmentation (breaking up into pieces), these landscapes were selected to have independent variation in habitat amount and fragmentation: (a) high amount, low fragmentation; (b) high amount, high fragmentation; (c) low amount, low fragmentation; (d) low amount, high fragmentation. Adapted from Ethier K and Fahrig L (2011) Positive effects of forest fragmentation, independent of forest amount, on bat abundance in eastern Ontario, Canada. *Landscape Ecology* 26: 865–876.

in human-dominated cover such that it constitutes habitat from the species' perspective. When measuring natural cover (e.g. forest amount) to estimate habitat amount one might erroneously conclude that species abundance is improved by "habitat" loss, when that species' habitat is actually the cover type (e.g., agriculture) that increased with human landscape modification.

The use of natural cover types as an approximation for habitat is common because most species decline with decreased natural cover, but the exceptions to this rule can make the use of "habitat" to mean "natural cover" misleading. Species richness declines with loss of natural cover for most groups (bacteria, Gans *et al.*, 2005; birds, butterflies, canopy beetles, canopy ants, and termites, Lawton *et al.*, 1998; forest birds, molluscs, and lichens, Moning and Müller, 2009; and spiders, Prieto-Benítez and Méndez, 2011). However, a significant number of species benefit from human-dominated cover. Edge- and matrix-favoring groups in the Amazon benefit from forest loss (Laurance *et al.*, 2002), and some prey species can increase in abundance following habitat loss, if the habitat loss has a larger negative effect on the predators than on the prey (Ryall and Fahrig, 2005; Ryall and Fahrig, 2006). Measurement of a cover type without reference to a species use of it can lead to confusing conclusions about the effect of habitat loss (Lindenmayer and Fischer, 2007).

Even if a cover type reflects an accurate measure of habitat, it may be insufficiently precise. Holland *et al.* (2005) found that the amount of forest in the surrounding habitat was more strongly related to the abundance of beetles that relied on multiple food sources than the abundance of beetles that relied on a single food source. They hypothesized that this occurred because the specialist feeders were each limited to specific stand types within a forest and the generalist feeders were more tightly associated with forest cover in general.

Landscape Complementation

For some species, measurement of a single cover type may be inadequate because their habitat includes multiple cover types. Leopard frogs, for example, require spring breeding habitat (ponds), summer foraging habitat (grassy meadows), and overwintering habitat (streams or lakes). In a study of landscape effects on local leopard frog abundance, researchers were unable to detect an effect of breeding habitat amount on abundance in focal ponds unless the amount of summer foraging habitat was also included in their model (Pope *et al.*, 2000).

Scale of Effect

The spatial extent at which cover is measured influences the ability to accurately identify the relationship between landscape structure and population or community outcomes. Numerous studies show that the spatial extent over which habitat is measured influences the strength of the relationship between habitat and the response of interest (e.g., abundance), and that this "scale of effect" is species specific (Carr and Fahrig, 2001; Holland *et al.*, 2004; Roland and Taylor, 1997; Steffan-Dewenter *et al.*, 2002). Failure to measure habitat at the scale of effect can lead to the erroneous conclusion that habitat amount is not associated with biodiversity (Holland *et al.*, 2004). When researchers are unsure of the scale at which a species responds to the landscape, a common solution is to measure landscape features at multiple spatial scales to find the one that best correlates with the response of interest (Holland *et al.*, 2004). A simulation study indicates that the maximum dispersal distance recorded for a species is likely to be a good estimate of the scale of effect (measured as the diameter of a circular landscape surrounding a focal area, Jackson and Fahrig, in preparation).

Grain, or the resolution at which landscape structure is defined, may also influence the ability to detect relationships between species and landscape structure. Animals are expected to have a "functional grain" or the smallest spatial scale at which they recognize spatial heterogeneity (Baguette and Van Dyck, 2007). If the grain of a landscape is defined too coarsely (e.g., the finest resolution is larger than a species' average dispersal distance), important relationships between a species and landscape structure may be overlooked.

Human-Caused Fragmentation versus Natural Fragmentation

Habitat fragmentation is sometimes used to refer to both natural fragmentation (e.g., forest patches separated by fire-maintained savannah) and human-caused fragmentation (e.g., forest patches separated by agricultural land). Species in

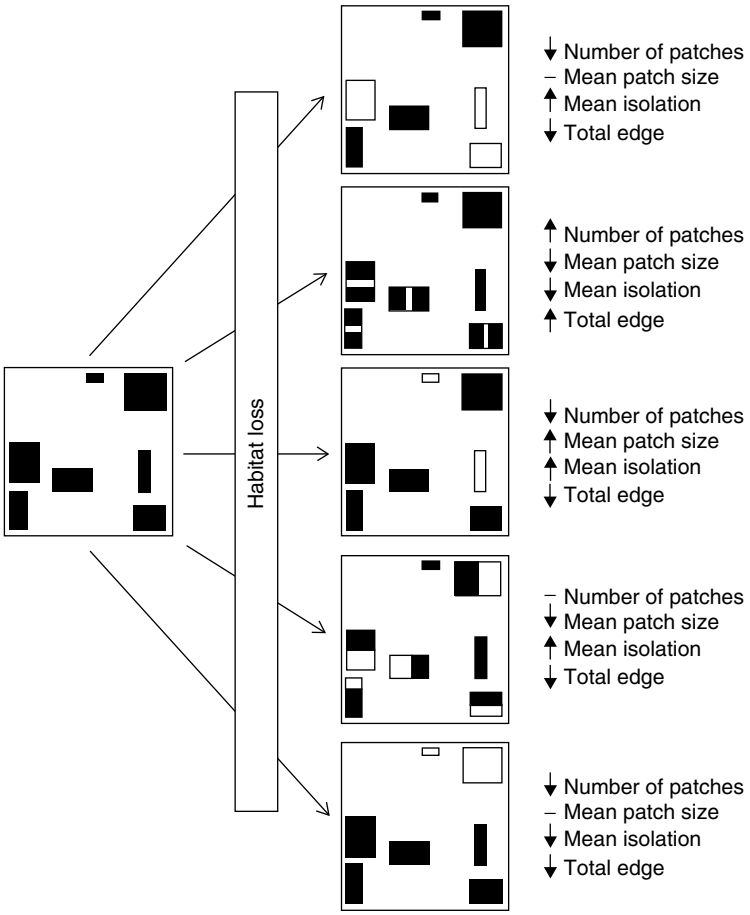


Figure 3 Illustration of how potential effects of habitat loss on measures of habitat fragmentation can be contrary to commonly expected effects. The commonly expected effects of habitat loss are an increase in the number of patches, an increase in mean isolation (here measured as nearest neighbor distance), a decrease in mean patch size, and an increase in total edge. Effects which are contrary to these expectations are italicized. Empty rectangles indicate areas of habitat loss. Adapted from Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34: 487–515.

naturally fragmented habitats have presumably adapted to natural fragmentation such that their movement behavior (Fahrig, 2007) and population response (Bunnell, 1999; Haila, 2002) to fragmentation will be qualitatively different from species confronted with fragmentation in formerly contiguous habitats.

Precise Language Facilitates Informative Science and Clear Management Directives

A vague conceptualization of habitat fragmentation can lead to several problems. Inadequate separation of the processes involved in habitat fragmentation can lead to the false impression that the effects of fragmentation cannot be generalized (Fahrig, 2003). The effects of fragmentation can be generalized when the aspect of fragmentation that is measured is explicitly defined and the other aspects of fragmentation are held constant. In addition, the illusion of a single process of habitat fragmentation has led to the misconception that one measure of the habitat fragmentation is equivalent to another (Fahrig, 2003). Without carefully delineating terms, one might be led to believe that any measure of fragmentation refers in the same way to the broader concept when in fact different metrics measure different processes. Finally, linguistic uncertainty can make management decisions more difficult (Regan *et al.*, 2002).

The main remedy for confusion surrounding the broad conceptualization of fragmentation is precise communication concerning what aspect of human landscape modification is being measured. Fahrig (2003) recommends that "fragmentation" be reserved for the breaking apart of habitat after accounting for habitat loss – fragmentation per se. Fischer and Lindenmayer (2007) also warn against "double-booking" for the term "habitat fragmentation" which can refer to both the broad concept (that includes everything discussed in this section) and the narrow definition of change in configuration after accounting for loss. They recommend that the broad concept be termed "human landscape modification" and accept Fahrig's suggestion that the narrow concept be called fragmentation per se. Ewers and Didham (2007) defend the use of "habitat fragmentation" for the broad

concept as an aid to communicating ideas across related studies but support other authors in their insistence that the specific way in which habitat fragmentation is measured should be clearly defined.

Careful language will improve the communication of ideas, but definitive separation of the various mechanisms conceptualized in the term habitat fragmentation also requires improved study design. Fahrig (2003) recommends that habitat loss and fragmentation per se be controlled, experimentally if possible, but statistically if necessary.

Effects of Habitat Loss and Fragmentation

Habitat Loss is More Devastating than Fragmentation per se

Measures of habitat configuration are almost always strongly correlated with habitat amount (**Figure 4**; Fahrig, 2003; Hargis *et al.*, 1998) making it difficult to discern whether the effects of human landscape modification are due to habitat loss or fragmentation per se. When the effect of fragmentation per se is measured independently from habitat amount (either by statistically or experimentally controlling for habitat amount), the effect size of habitat loss on biodiversity is usually much larger than that of habitat fragmentation per se. When fragmentation per se does have an effect, it is just as likely to be positive as negative (reviewed in Fahrig, 2003).

Habitat Loss is Associated with Greater Extinction Probability and Reduced Species Richness

Large amounts of habitat are expected to result in positive effects for populations and communities for a number of reasons. At the population level, greater amounts of habitat sustain larger population sizes, which are less susceptible to stochastic extinction (MacArthur and Wilson, 1967) and genetic drift (Wright, 1931). At the community level, lower extinction rates in areas with more habitat are expected to lead to greater species richness (MacArthur and Wilson, 1967). In addition, large areas of natural cover are more likely to include a larger variety of habitat types with their associated species (Williams, 1964). Greater species richness may also occur in larger areas simply because larger areas allow for a larger sample of species from a regional species pool, resulting in a greater number of rare species (Connor and McCoy, 1979).

Fragmentation per se Can Have Both Positive and Negative Results

Fragmentation per se is expected to lead to negative effects if it results in smaller, less viable populations or if the increased exposure to edge is detrimental (Bender *et al.*, 1998). Although the negative effects of habitat loss are expected to result from a reduced number of individuals, cover types, and species, fragmentation per se is generally assumed to reduce the probability of successful movement. This is likely a false

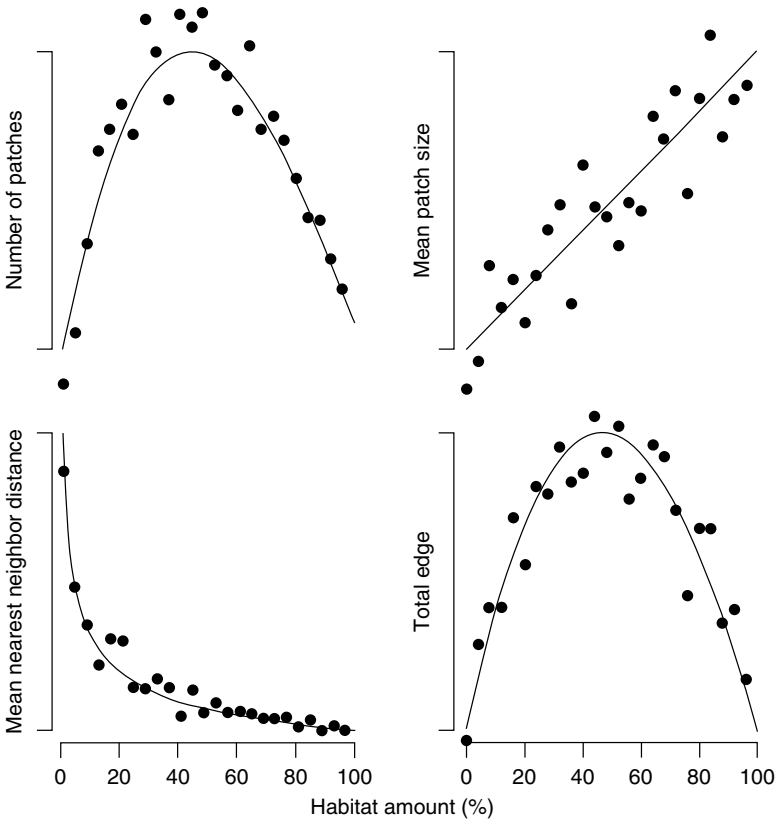


Figure 4 Illustration of typical relationships between habitat amount and four measures of fragmentation. Each data point corresponds to an individual landscape. Adapted from Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34: 487–515.

dichotomy, however, because habitat loss itself reduces dispersal success by increasing the distances among patches (Fahrig, 2007).

Fragmentation per se implies a greater number of smaller patches. If the cost of crossing matrix is too high to allow for regular movement among patches, then individual subpopulations operate more or less independently. As in classic metapopulations (Hanski, 1994), the resulting small, isolated subpopulations are more susceptible to extinction.

A greater number of smaller patches also results in a higher perimeter to area ratio over the landscape. More frequent encounters with edges often increases emigration rates (Kareiva, 1985), thereby increasing the probability of dispersal mortality (Fahrig, 2001). Some matrix types are particularly risky and can have strong effects on population density (e.g., roads, Fahrig *et al.*, 1995).

Edges themselves can represent areas of decreased habitat quality. The Biological Dynamics of Forest Fragments Project documented strong effects of edges in a tropical forest, including the following: enhanced wind disturbance, increased tree mortality, invasion of disturbance-adapted species, altered species composition, lower relative humidity, decreased soil moisture, lower canopy height, decreased canopy foliage, increased temperature, and increased litter (reviewed in Laurance *et al.*, 2002). Negative edge effects are not as well-documented in temperate areas and may be less common (Fahrig, 2003).

Receiving less attention in the habitat fragmentation literature is the fact that fragmentation per se is just as likely to result in positive population responses as negative ones. Fragmentation per se may lead to positive population responses if subdivision of habitat results in (1) greater coexistence of competitors (Chesson, 1985), (2) greater coexistence of predators and their prey (Huffaker, 1958), (3) lower probability of simultaneous extinction of an entire population (Heino *et al.*, 1997), (4) greater immigration into habitat due to larger perimeter to area ratio (Bowman *et al.*, 2002), (5) greater dispersal success due to shorter interpatch distances (Grez *et al.*, 2004), (6) greater landscape complementation (e.g., providing access between foraging and nesting sites for bats, Ethier and Fahrig, 2011), or (7) positive edge effects (e.g., increasing abundance of edge-favored species, Laurance *et al.*, 2002).

Increased Reproductive Rate Decreases Sensitivity to Habitat Loss and Fragmentation

Reproductive rate may be the most important factor influencing the amount of habitat required to sustain a population. Theory predicts the existence of a threshold amount of habitat at which a small loss of habitat propels the probability of population extinction from 0 to 1 (Fahrig, 2001; Lande, 1987; With and King, 1999). The minimum habitat requirement likely depends on the attributes of the species and the landscape. A simulation testing the relative importance of reproductive rate, emigration rate, matrix quality, and fragmentation (randomly distributed vs. clumps of habitat) showed that the most drastic differences in extinction threshold occurred when the reproductive rate was varied (Fahrig, 2001).

Empirical evidence supports the importance of reproductive rate. Among 41 breeding bird species, higher reproductive rate is associated with lower minimum habitat requirements (Vance *et al.*, 2003). Likewise, for 17 mammal species, the sensitivity of abundance to road density declined with reproductive rate independently of body size (Rytwinski and Fahrig, 2011).

Mobility, Habitat Specialization, and Matrix Quality Interact to Alter the Effects of Habitat Loss and Fragmentation

Habitat loss may have particularly severe effects for species that do not avoid matrix. The negative effect is expected to be even stronger if the matrix is dangerous. This is likely because matrix avoidance and/or low mobility (low probability of moving and/or short dispersal distances) is expected to be common among species originating from naturally fragmented habitats surrounded by high-risk matrix (Fahrig, 2007). Species from naturally continuous habitat or with little exposure to high-risk matrix may not have evolved matrix avoidance and may therefore suffer greater mortality than matrix-avoiding species when encountering high-risk matrix resulting from habitat loss (Fahrig, 2007).

Theoretical and empirical studies support the important interaction between dispersal traits and matrix quality. Two studies of variation in matrix avoidance among butterfly subpopulations support the hypothesis that matrix avoidance is selected for when interpatch distances are high and probability of successful dispersal through matrix is low (Hanski *et al.*, 2004; Schtickzelle and Baguette, 2003). A simulation model testing the relative impact of the number of patches, the distance between patches, and matrix avoidance on population density found that matrix avoidance was the most important factor preventing population decline (Tischendorf *et al.*, 2005). A model by Fahrig (2001) showed that populations with increased movement probability (matrix avoidance was not modeled) required more habitat to persist. This result depended on the mortality risk in the matrix – when the risk of mortality in the matrix was low, the effect of movement on minimal habitat requirements was greatly decreased.

Low-contrast matrix (matrix similar in structure to habitat) is generally thought to be better tolerated by native species than high-contrast matrix, but this conclusion may depend on how "matrix" is defined – for habitat generalists, "matrix" may actually provide habitat. Matrix type was more important than fragment size when predicting the recolonization of many Amazonian insectivorous birds after deforestation, probably because the "matrix" in which forest regrowth was more rapid and was most likely to provide birds with foraging and breeding opportunities (Stouffer and Bierregaard, 1995). In other words, the secondary growth served as habitat (technically not matrix). Likewise, dung beetle assemblages largely returned to their pretreatment state 15 years after deforestation when fragments were surrounded by secondary growth forest. Again, this recovery was due to the ability of dung beetles to use resources in secondary growth forest (Quintero and Roslin, 2005).

Low-contrast matrix may be detrimental to populations if it fails to evoke matrix avoidance and provides no resources on its own. A planthopper specialized on cordgrass avoids crossing into natural mudflat matrix but crosses easily into inedible, nonnative brome, which is similar in height to

cordgrass (Haynes and Cronin, 2006). Higher emigration into nonnative brome is associated with decreased density of planthoppers within cordgrass patches (Haynes *et al.*, 2007) and increased probability of planthopper extinction in cordgrass patches (Cronin, 2007).

Matrix heterogeneity can obscure the relationship between habitat loss and dispersal success. In both a simulation study and a field study of two mammal species, large patch size and short isolation distances were strong positive predictors of dispersal success when matrix was homogeneous, but the predictive value of these metrics was greatly reduced when matrix consisted of multiple cover types (Bender and Fahrig, 2005).

The interaction between matrix quality and mobility emphasizes that the important dispersal characteristic for populations is not so much low or high mobility, but high probability of success when dispersal is attempted (Callens *et al.*, 2011; Fahrig, 2007; Van Houtan *et al.*, 2007). For example, a simulation study suggests that high dispersal mortality can deplete genetic diversity more rapidly than complete dispersal avoidance (Jackson and Fahrig, in review). Simulation studies indicate that successful dispersal in patchily distributed habitats is likely to rely on matrix avoidance (Tischendorf *et al.*, 2005), linear movements in matrix (Zollner and Lima, 1999), strong ability to perceive habitat from a distance (Zollner and Lima, 2005), and adequate energy storage (Zollner and Lima, 2005). Species with high rates of successful dispersal may have been less vulnerable during past mass extinction events (Stork *et al.*, 2009), potentially because of their ability to spread the risks of environmental and demographic stochasticity over a large area and to avoid inbreeding depression (reviewed in Bowler and Benton, 2005).

In summary, population health after human landscape modification depends strongly on the interactions among matrix quality, habitat specialization, and the response of species to matrix during dispersal.

The Effects of Habitat Loss are Often Exacerbated by Additional Environmental Threats

Habitat loss usually cooccurs with other environmental threats (Laurance and Useche, 2009). In a study of vertebrate species listed under International Union for Conservation of Nature (IUCN), the most common pair of threats was agriculture (the most common reason for habitat loss) and hunting (Laurance and Useche, 2009). Human landscape modification results in more edge, providing easier access for hunters than is available in continuous forest. In the Amazon, for example, small forest fragments tend to experience much greater hunting pressure than large forest fragments (Peres, 2001), which can result in local extirpation of game species (Cullen *et al.*, 2000).

Habitat loss can lead to increased fire frequency in the tropics. Forested edges tend to be significantly drier than forest interiors (Kapos, 1989) and often abut slash and burn agriculture in the tropics (Cochrane and Laurance, 2002). In one such region, forest fires on forest/agriculture edges occurred every 5 years, whereas the fire interval averaged 120 years in core forest areas (Cochrane and Laurance, 2002).

Habitat loss is also expected to exacerbate the effects of climate change on species. Poleward range expansion is hypothesized to be impaired if human landscape modification limits poleward movement or decreases population growth rate (Opdam and Wascher, 2004). One area of active research is the identification of areas of habitat conservation priority based not only on current habitat use but also on projected future distributions of species under various climate change scenarios (Olson and Lindsay, 2009).

Summary

The global scale of human landscape modification is staggering. Amelioration of its effects depends in part on clear communication. To enable understanding of the effects of habitat loss and fragmentation, practitioners must (1) clearly define which aspect of human landscape modification (habitat loss or some metric of fragmentation per se) they are discussing, (2) make measurements at the scale (patch or landscape) that will yield the most interpretable results (usually landscape scale with replication), (3) clearly define what constitutes habitat to the taxon of interest, (4) measure habitat at the spatial extent appropriate for the taxon of interest, and (5) delineate between natural and human-caused fragmentation. Habitat loss is a bigger problem than habitat fragmentation per se and should be the main focus of study and management. Fragmentation per se is expected to be a problem when it leads to smaller, isolated populations or when negative edge effects are expected to be prominent (e.g., in tropical systems). The effect of habitat loss is modified by species and landscape characteristics. An increase in reproductive rate decreases the amount of habitat required to sustain a population. High mobility is associated with high sensitivity to habitat loss if dispersal mortality is high, but some successful dispersal may be required for long-term persistence. Habitat loss and fragmentation are seldom the only threats in an area – they are often coupled with hunting, fire, and other threats, which should be taken into account.

See also: Deforestation and Land Clearing. Extinction, Causes of. Loss of Biodiversity, Overview. Metapopulations. Population Viability Analysis. Restoration of Biodiversity, Overview. Species–Area Relationships

References

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, and Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Molecular Ecology* 17: 5177–5188.

Andrén H (1994) Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: A review. *Oikos* 71: 355–366.

Baguette M and Van Dyck H (2007) Landscape connectivity and animal behavior: Functional grain as a key determinant for dispersal. *Landscape Ecology* 22: 1117–1129.

Balmford A, Bruner A, Cooper P, *et al.* (2002) Ecology – Economic reasons for conserving wild nature. *Science* 297: 950–953.
Bender DJ, Contreras TA, and Fahrig L (1998) Habitat loss and population decline: A meta-analysis of the patch size effect. *Ecology* 79: 517–533.
Bender DJ and Fahrig L (2005) Matrix structure obscures the relationship between interpatch movement and patch size and isolation. *Ecology* 86: 1023–1033.
Bowler DE and Benton TG (2005) Causes and consequences of animal dispersal strategies: Relating individual behaviour to spatial dynamics. *Biological Reviews* 80: 205–225.
Bowman J, Cappuccino N, and Fahrig L (2002) Patch size and population density: The effect of immigration behavior. *Conservation Ecology* 6: 9.
Brown MJF and Paxton RJ (2009) The conservation of bees: A global perspective. *Apidologie* 40: 410–416.
Bruijnzeel LA (2004) Hydrological functions of tropical forests: Not seeing the soil for the trees? Agriculture. *Ecosystems & Environment* 104: 185–228.
Bunnell FL (1999) Let's kill a panchreston: Giving fragmentation meaning. In: Rochelle JA, Lehmann LA, and Wisniewski J (eds.) *Forest Fragmentation: Wildlife and Management Implications*, pp. vii–xiii. Leiden, The Netherlands: Koninklijke Brill NV.
Callens T, Galbusera P, Matthysen E, *et al.* (2011) Genetic signature of population fragmentation varies with mobility in seven bird species of a fragmented Kenyan cloud forest. *Molecular Ecology* 20: 1829–1844.
Carr LW and Fahrig L (2001) Effect of road traffic on two amphibian species of differing vagility. *Conservation Biology* 15: 1071–1078.
Chesson PL (1985) Coexistence of competitors in spatially and temporally varying environments – A look at the combined effects of different sorts of variability. *Theoretical Population Biology* 28: 263–287.
Cochrane MA and Laurance WF (2002) Fire as a large-scale edge effect in Amazonian forests. *Journal of Tropical Ecology* 18: 311–325.
Connor EF and McCoy ED (1979) The statistics and biology of the species-area relationship. *The American Naturalist* 113: 791–833.
Cronin JT (2007) From population sources to sieves: The matrix alters host-parasitoid source-sink structure. *Ecology* 88: 2966–2976.
Cullen L, Bodmer RE, and Padua CV (2000) Effects of hunting in habitat fragments of the Atlantic forests, Brazil. *Biological Conservation* 95: 49–56.
Delin AE and Andrén H (1999) Effects of habitat fragmentation on Eurasian red squirrel (*Sciurus vulgaris*) in a forest landscape. *Landscape Ecology* 14: 67–72.
Dulvy NK, Sadovy Y, and Reynolds JD (2003) Extinction vulnerability in marine populations. *Fish and Fisheries* 4: 25–64.
Ethier K and Fahrig L (2011) Positive effects of forest fragmentation, independent of forest amount, on bat abundance in eastern Ontario, Canada. *Landscape Ecology* 26: 865–876.
Ewers RM and Didham RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews* 81: 117–142.
Ewers RM and Didham RK (2007) Habitat fragmentation: Panchreston or paradigm? *Trends in Ecology and Evolution* 22: 511.
Fahrig L (2001) How much habitat is enough? *Biological Conservation* 100: 65–74.
Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34: 487–515.
Fahrig L (2007) Non-optimal animal movement in human-altered landscapes. *Functional Ecology* 21: 1003–1015.
Fahrig L, Pedlar JH, Pope SE, Taylor PD, and Wegner JF (1995) Effect of road traffic on amphibian density. *Biological Conservation* 73: 177–182.
FAO (2003) *World Agriculture: Towards 2015/2030. An FAO Perspective.* Rome, Italy: Food and Agriculture Organization of the United Nations.
FAO (2010) *Global Forest Resources Assessment 2010 – Key Findings.* Rome, Italy: Food and Agriculture Organization of the United Nations.
Fargione J, Hill J, Tilman D, Polasky S, and Hawthorne P (2008) Land clearing and the biofuel carbon debt. *Science* 319: 1235–1238.
Fischer J and Lindenmayer DB (2007) Landscape modification and habitat fragmentation: A synthesis. *Global Ecology and Biogeography* 16: 265–280.
Gans J, Wolinsky M, and Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* 309: 1387–1390.
Grez A, Zaviezo T, Tischendorf L, and Fahrig L (2004) A transient, positive effect of habitat fragmentation on insect population densities. *Oecologia* 141: 444–451.
Haila Y (2002) A conceptual genealogy of fragmentation research: From island biogeography to landscape ecology. *Ecological Applications* 12: 321–334.
Hanski I (1994) A practical model of metapopulation dynamics. *Journal of Animal Ecology* 63: 151–162.
Hanski I, Eralahti C, Kankare M, Ovaskainen O, and Siren H (2004) Variation in migration propensity among individuals maintained by landscape structure. *Ecology Letters* 7: 958–966.
Hargis CD, Bissonette JA, and David JL (1998) The behavior of landscape metrics commonly used in the study of habitat fragmentation. *Landscape Ecology* 13: 167–186.
Harrod RJ, McRae BH, and Hartl WE (1999) Historical stand reconstruction in ponderosa pine forests to guide silvicultural prescriptions. *Forest Ecology and Management* 114: 433–446.
Haynes KJ and Cronin JT (2006) Interpatch movement and edge effects: The role of behavioral responses to the landscape matrix. *Oikos* 113: 43–54.
Haynes KJ, Dillemuth FP, Anderson BJ, *et al.* (2007) Landscape context outweighs local habitat quality in its effects on herbivore dispersal and distribution. *Oecologia* 151: 431–441.
Heino M, Kaitala V, Ranta E, and Lindstrom J (1997) Synchronous dynamics and rates of extinction in spatially structured populations. *Proceedings of the Royal Society of London Series B – Biological Sciences* 264: 481–486.
Holland JD, Bert DG, and Fahrig L (2004) Determining the spatial scale of species' response to habitat. *Bioscience* 54: 227–233.
Holland JD, Fahrig L, and Cappuccino N (2005) Body size affects the spatial scale of habitat-beetle interactions. *Oikos* 110: 101–108.
Huffaker C (1958) Experimental studies on predation: Dispersion factors and predator–prey oscillations. *Hilgardia* 27: 343–383.
Kapos V (1989) Effects of isolation on the water status of forest patches in the Brazilian Amazon. *Journal of Tropical Ecology* 5: 173–185.
Kareiva P (1985) Finding and losing host plants by *Phyllotreta*: Patch size and surrounding habitat. *Ecology* 66: 1809–1816.
Lande R (1987) Extinction thresholds in demographic-models of territorial populations. *American Naturalist* 130: 624–635.
Laurance WF (2008) Theory meets reality: How habitat fragmentation research has transcended island biogeographic theory. *Biological Conservation* 141: 1731–1744.
Laurance WF, Lovejoy TE, Vasconcelos HL, *et al.* (2002) Ecosystem decay of Amazonian forest fragments: A 22-year investigation. *Conservation Biology* 16: 605–618.
Laurance WF and Useche DC (2009) Environmental synergisms and extinctions of tropical species. *Conservation Biology* 23: 1427–1437.
Lawton JH, Bignell DE, Bolton B, *et al.* (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature* 391: 72–76.
Lindenmayer DB and Fischer J (2007) Tackling the habitat fragmentation panchreston. *Trends in Ecology and Evolution* 22: 127–132.
MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Princeton: Princeton University Press.
McGarigal K and Cushman S (2002) Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. *Ecological Applications* 12: 335–345.
Moning C and Müller J (2009) Critical forest age thresholds for the diversity of lichens, molluscs and birds in beech (*Fagus sylvatica* L.) dominated forests. *Ecological Indicators* 9: 922–932.
Olson LT and Lindsay KF (2009) Here today, gone tomorrow? Targeting conservation investment in the face of climate change. *Journal of Geography and Regional Planning* 2: 20–29.
OMNR (2010) DRAPE: Orthoimagery (computer file). Ontario Ministry of Natural Resources.
Opdam P and Wascher D (2004) Climate change meets habitat fragmentation: Linking landscape and biogeographical scale levels in research and conservation. *Biological Conservation* 117: 285–297.
Pereira HM, Leadley PW, Proença V, *et al.* (2010) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496–1501.
Peres CA (2001) Synergistic effects of subsistence hunting and habitat fragmentation on Amazonian forest vertebrates. *Conservation Biology* 15: 1490–1505.
Pope SE, Fahrig L, and Merriam HG (2000) Landscape complementation and metapopulation effects on leopard frog populations. *Ecology* 81: 2498–2508.
Potts SG, Biesmeijer JC, Kremen C, *et al.* (2010) Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and Evolution* 25: 345–353.
Prieto-Benítez S and Méndez M (2011) Effects of land management on the abundance and richness of spiders (Araneae): A meta-analysis. *Biological Conservation* 144: 683–691.
Quintero I and Roslin T (2005) Rapid recovery of dung beetle communities following habitat fragmentation in central Amazonia. *Ecology* 86: 3303–3311.
Regan HM, Colyvan M, and Burgman MA (2002) A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications* 12: 618–628.
Ricketts TH, Regetz J, Steffan-Dewenter I, *et al.* (2008) Landscape effects on crop pollination services: Are there general patterns? *Ecology Letters* 11: 499–515.

Ricklefs RE (2008) *The Economy of Nature*. New York: W. H. Freeman and Company.
Roland J and Taylor PD (1997) Insect parasitoid species respond to forest structure at different spatial scales. *Nature* 386: 710–713.
Ryall KL and Fahrig L (2005) Habitat loss decreases predator–prey ratios in a pine-bark beetle system. *Oikos* 110: 265–270.
Ryall KL and Fahrig L (2006) Response of predators to loss and fragmentation of prey habitat: A review of theory. *Ecology* 87: 1086–1093.
Rytwinski T and Fahrig L (2011) Reproductive rate and body size predict road impacts on mammal abundance. *Ecological Applications* 21: 589–600.
Schtickzelle N and Baguette M (2003) Behavioural responses to habitat patch boundaries restrict dispersal and generate emigration-patch area relationships in fragmented landscapes. *Journal of Animal Ecology* 72: 533–545.
Steffan-Dewenter I, Münzenberg U, Bürger C, Thies C, and Tscharntke T (2002) Scale-dependent effects of landscape context on three pollinator guilds. *Ecology* 83: 1421–1432.
Stork NE, Coddington JA, Colwell RK, *et al*. (2009) Vulnerability and resilience of tropical forest species to land-use change. *Conservation Biology* 23: 1438–1447.
Stouffer PC and Bierregaard RO (1995) Use of Amazonian forest fragments by understory insectivorous birds. *Ecology* 76: 2429–2445.
Tischendorf L, Grez A, Zaviezo T, and Fahrig L (2005) Mechanisms affecting population density in fragmented habitat. *Ecology and Society* 10: 13.
Van Houtan KS, Pimm SL, Halley JM, Bierregaard RO, and Lovejoy TE (2007) Dispersal of Amazonian birds in continuous and fragmented forest. *Ecology Letters* 10: 219–229.
Vance MD, Fahrig L, and Flather CH (2003) Effect of reproductive rate on minimum habitat requirements of forest-breeding birds. *Ecology* 84: 2643–2653.
Williams CB (1964) *Patterns in the Balance of Nature*. London: Academic Press.
With KA and King AW (1999) Extinction thresholds for species in fractal landscapes. *Conservation Biology* 13: 314–326.
Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97–159.
Zollner PA and Lima SL (1999) Search strategies for landscape-level interpatch movements. *Ecology* 80: 1019–1030.
Zollner PA and Lima SL (2005) Behavioral tradeoffs when dispersing across a patchy landscape. *Oikos* 108: 219–230.

Habitat Selection

Robert A Montgomery and Gary J Roloff, Michigan State University, East Lansing, MI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Environmental variables Abiotic and biotic elements of an area including vegetation, soil, topography, species competition, geology, prevalence of diseases and hydrologic characteristics, among others. Habitat selection can be positively or negatively related to these environmental variables.

Habitat An area where an organism resides is comprised of abiotic and biotic elements.

Habitat patch A spatially contiguous area in which environmental variables are similar.

Habitat selection The act of choosing the combination of available abiotic and biotic elements for the purpose of fulfilling the life history events of the organism.

Hierarchical habitat selection This concept recognizes that an organism's selection of habitat occurs at varying spatial and temporal scales. The scale at which different selections occur are ordered; First-order (habitat selections defining the species range), Second-order (selections defining the placement of an individual's home range within the species range), Third-order (within home range selections), and Fourth-order (fine-scale daily habitat selections).

Introduction

Habitat is a theoretical construct used to describe the living space of an organism. As such, habitat includes the suite of interacting abiotic (e.g., weather, soils, topography, and hydrology) and biotic (e.g., vegetation structure and composition, inter- and intraspecific competitions, and prevalence of diseases) elements that influence whether an organism uses a particular location. Habitat selection is the act of choosing the combination of available abiotic and biotic elements for the purpose of fulfilling the life history events of the organism (e.g., mating, raising young, and avoiding death). An organism's selection of habitat varies across space and time as certain needs are fulfilled and others pursued. The objective of habitat selection research is to describe why organisms use various combinations of environmental (abiotic and biotic) variables, an understanding of which informs population or species conservation. The concept of habitat selection can be extended to plants and fungi but animals are the most common research subject.

When discussing the concept of habitat selection it is important to differentiate among habitat use, habitat selection, and habitat preference because each concept has important management and conservation implications. Habitat use refers to organisms occurring in an area, not necessarily by choice. Habitat selection implies that an organism made a choice about occurring in a particular location, that is, the organism had multiple options and for some reason (presumably to efficiently satisfy a life history event) chose to occur at a particular location. Habitat preference is applied at the species-level and refers to the habitat that the species prefers above all others, even if that habitat is unavailable in the landscape where a local population resides. Habitat preference is the most difficult concept to quantify because in most areas organisms are not free to choose among an unrestricted set of environmental variables. Thus, researchers tend to focus their efforts on quantifying habitat selection and use. Knowledge of preferred habitat may elude researchers because of the inability to comprehend the exact combination of environmental variables that would be selected if all were available. This reality has conservation implications which could potentially compromise organism, population, or even the species persistence into the future.

The origin of habitat selection research is based on the observation of animals by naturalists. One of the earliest documented naturalists was Aristotle who took a keen interest in animal–habitat associations (Morrison *et al.*, 2006). Since Aristotle, both informally and formally trained naturalists and ecologists have studied how organisms use their environment. This was certainly of concern to nineteenth century naturalists who examined the morphology and relatedness of certain species as the basis for concepts like biogeography and speciation (Wallace, 1876). These naturalists realized that the selection of habitat had important evolutionary ramifications (Darwin, 1859). The interaction of an organism with habitat was later framed as the ecologic niche (Grinnell, 1917). Niche theory developed into a research area unto itself with ecologists being principally concerned with how organism–habitat interactions shape fitness (Petren, 2001).

The twentieth century saw an expansion in research methodologies used to study habitat selection. The collective goal of these methodologies was to identify the location of organisms across spatial and temporal dimensions and to examine the environmental variables associated with those locations. Habitat selection research broadly consists of two phases: (1) Documenting habitat used by an organism of interest and (2) making inferences on why habitat was selected (usually based on a comparison of habitat use and availability). The first step in quantifying habitat use is to map organism locations. Multiple techniques are used to locate organisms in space including direct visual observation, conducting visual or auditory surveys, using traps or other evidence of occurrence (i.e., scat, tracks, and carcasses), and telemetry technology (**Figure 1**). Each of these techniques results in animal location(s) that can be related to a suite of environmental variables which collectively define habitat.



Figure 1 Four techniques used to locate organisms in environmental space, (a) direct visual observation, (b) visual surveys, (c) trapping (camera trapping), and (d) telemetry.

Once animals are located and the environmental variables of interest are measured, habitat selection can be quantified.

Measurement of Animal Locations

Direct Visual Observation

Directly observing animals in their habitat is the oldest technique used to measure habitat selection. This process involves researchers following single or multiple animals through environmental space while making notations about the environmental variables that are used to satisfy various life history events (**Figure 1a**). Though visual observation might appear to be a relatively simplistic means of evaluating habitat use, increasingly complex methods of sampling have been employed in animal behavior research (Altmann, 1974). *Ad-libitum* sampling, focal animal sampling, and all occurrence sampling are common direct visual observation techniques used to assign the location of an organism with the environmental variables that describe habitat.

Ad-libitum sampling is a form of visual observation where the researcher documents habitat use in an *ad-lib*, or opportunistic fashion. As such, *ad-libitum* sampling most aptly describes the technique used by early naturalists. A field notebook is commonly used to record environmental variables chosen by the organism of interest. This sampling scheme is notoriously difficult to quantify with reliability because of the opportunistic fashion in which the data are recorded. Thus this technique has been criticized for being more anecdotal than empirical (Altmann, 1974; Mann, 1999). However, *ad-libitum* sampling can be useful in pilot studies or in situations where a researcher is attempting to document the variety of environmental variables used by an organism over limited spatial and temporal dimensions (Altmann, 1974).

Focal animal sampling involves the detailed observation of a single organism at specified intervals over a period of time. For instance, a researcher would identify an individual to observe and would follow that individual wherever it goes recording all behaviors, including environmental variables that are used. The researcher notes the type of behavior, the duration of the performed behavior, whether or not the behavior involved interactions with other individuals or organisms, and any other environmental variables that were used. This type of sampling would continue for a specified interval (e.g., 5 min, 15 min, 1 h) at which point the researcher would break from recording data.

With all occurrence sampling, the researcher focuses on a specific behavior (such as selection of an environmental variable) and records any time that the behavior is performed throughout the sampling interval. This technique is widely employed in animal behavior research because it provides an

estimate of occurrence rate for a particular behavior. With this sampling technique the researcher can estimate how often a particular food item is selected, for instance, or how often cover habitat is selected per unit of time. All occurrence sampling can be useful in determining the relative dependence that an organism has for a particular environmental variable.

Visual and Auditory Surveys

Another set of research methodologies developed to identify the location of organisms in space at a particular time involves visual or auditory surveys (**Figure 1b**). Visual or auditory detection have been used to estimate species richness, produce population estimates, and characterize habitat selection. Several different methodologies exist but these can be categorized into two broad types; visual encounter surveys and auditory surveys.

Visual encounter surveys are widely utilized in ground, air, and ship-board capacities. In these cases the researcher scans the environment to visually identify individuals of the focal species. Various methods are used to conduct these surveys including line transects, randomized walk designs, and patch-based searches. Line transect surveys involve the designation of routes that researchers follow while searching for organisms (e.g., Otto and Roloff, 2011). Typically, the researcher limits the sampling to an area where organisms can be reliably detected. In random walk surveys the researcher follows a random path, while searching for organisms, for a specified distance (e.g., Gregory *et al.*, 2006). On reaching that distance, another random path is selected and followed after which the process is repeated. The advantage of this technique is that it can be an effective method for searching large areas, if the focal organism can be reliably observed. Patch-based surveys involve sampling all environmental variables within a given area. For example, if the goal of a survey is to identify birthing sites and the abundance and distribution of Hawaiian monk seals (*Monachus schauinslandi*) patch-based surveys would sample all beaches in the Hawaiian archipelago searching for females and their young during and directly after the pupping season (Baker and Johanos, 2004).

To conduct an auditory survey the researcher relies on sounds to identify and locate different organisms or species. Auditory surveys are particularly useful for vocal organisms like some birds and amphibians and include point counts and calling surveys. When employing point count surveys a number of locations or stations are identified where the researcher remains stationary listening for calls (e.g., Ralph *et al.*, 1993; Seavy *et al.*, 2009). The researcher stays on station for a specific period of time, documenting species occurrence through auditory and in some cases visual detection. The stations are typically spread across several different environmental variables that are relevant to the species of interest. Calling surveys involve surveyors searching certain areas with the goal of detecting where different organisms reside. These surveys are predicated on the idea that there are spatial and temporal conditions where certain species vocalize. Frogs and toads are common subjects of calling surveys (Pellet and Schmidt, 2005). The result of these surveys is the identification of environmental variables used by organisms of one or multiple species. These surveys are designed to detect organisms, assign organism locations, and associate organism locations to specific environmental variables.

Trapping and Other Evidence of Occurrence

Another technique used to measure animal locations in the environment involves trapping (**Figure 1c**). There are a wide variety of trapping methodologies that include, but are certainly not limited to, baited traps (e.g., Cypher *et al.*, 2009), drift fences (e.g., Otto and Roloff, 2011), and pitfall traps (e.g., Müller and Brandl, 2009) that restrain animals so their presence, abundance, and location can be documented. Trapping can also be done in noninvasive fashions through the collection of images (photographs or video; e.g., Soisalo and Cavalcanti, 2006; Thorn *et al.*, 2009) or deoxyribonucleic acid (DNA) (hair snares; e.g., Gardner *et al.*, 2010). Trapping designs can include a grid, transect, or web network of traps, with traps spaced accordingly to increase the probability of detecting organisms (i.e., spacing based on home range size or known movement patterns). Like direct visual observation and visual or auditory surveys, trapping techniques can be used to associate organism locations with environmental variables on the landscape.

There are some concerns with trapping. With baited traps, in particular, it is possible that certain organisms will become conditioned to and potentially attracted to the trap. Within this context, study results would be biased if the researcher is repeatedly capturing the same individual or a cohort of individuals that are vulnerable to trapping. Conversely, if organisms associate a negative experience with being trapped, they may purposefully avoid traps and similarly bias study results. Furthermore, if multiple species are captured in the same trap it is possible to have competition and predation events where the presence of one species is being masked because it is consumed by another species. For example, it is well known that some species, such as shrews (family *Soricidae*) and opossums (family *Didelphidae*), will prey on other animals and insects captured in the same trap (Jurzenski and Hoback, 2011).

Researchers are increasingly using noninvasive techniques such as motion-activated cameras to photographically capture organisms or hair snares to collect DNA samples (Cutler and Swann, 1999; Gardner *et al.*, 2009; **Figure 1c**). The design of camera traps is quite similar to the aforementioned trapping procedures. Cameras are placed in a variety of environmental variables throughout the study area with the intention of documenting species occurrence or to produce population abundance estimates using mark-recapture on individuals. Individual identification of organisms is possible through visual acuity or with computer software programs which assist in the recognition of pelage patterns, markings, and other morphological characteristics specific to individuals. These methods effectively eliminate the necessity of restraining organisms and have the benefit of noninvasively documenting organism–habitat associations, although baited camera and hair snare traps are subject to the same biases described in the preceding paragraph for other baited trapping approaches.

Several additional techniques can be used to identify the position of organisms in space. These include locations where

evidence of organism presence persists including scat, tracks, and carcasses, among others. In these cases researchers are generally interested in describing specific life history events and collect data to describe those events. For instance, a researcher may be interested in quantifying the environmental variables that are characteristic of areas where organisms are subject to predation. All areas in the study can be searched in opportunistic or structured fashions to identify predator-killed carcasses (Kauffman *et al.*, 2007; Bump *et al.*, 2009). Once located, carcass sites can be associated with different environmental variables to determine the abiotic and biotic elements that correlate with predation events.

Increasingly complex forms of collecting organism locations are continually employed in habitat selection research. Remote sensing, satellite imagery, and infrared technology are now being used to study habitat selection in new and interesting ways. For instance, to document the distribution and abundance of harbor seals (*Phoca vitulina*) researchers are using aircraft fit with infrared technology to locate heat signatures of seals on ice flows (London and Richmond, 2010). This technology enables researchers to cover vast areas with increased reliability in the estimates of organism–habitat associations. Radar technology, typically used for describing the weather, is now being used to map concentrations of migratory birds (Bonter *et al.*, 2008). Large concentrations of birds at stopover sites along migratory pathways are viewed as selection for the environmental variables that coincide with these aggregations.

Telemetry

Telemetry technology has enabled increasingly complex spatial and temporal assessments of habitat selection. Telemetry, which literally translates to 'remote measurement,' was utilized by the U.S. Government to detect Cold War conflict and was retrofitted for animal research in the middle of the twentieth century (Benson, 2010). In 1960 Dwain Warner and John Tester were field testing Very High Frequency (VHF) radio tags on ruffed grouse (*Bonasa umbellus*). At about the same time John and Frank Craighead started fitting radio collars on grizzly bears (*Ursus horriblis*) in the Greater Yellowstone Ecosystem, a feat which earned notoriety through National Geographic coverage and a CBS documentary (Benson, 2010). The Craighead's research led to publications in major peer-reviewed journals and a 1979 book titled *Track of the Grizzly*. The depth of this research and the notoriety that it garnered propelled telemetry technology to the forefront of habitat selection research. As a result, telemetry has developed into one of the main techniques used to study habitat selection.

Both VHF and Global Positioning System (GPS) telemetry depend on fitting an organism with a tag that transmits a signal and relocating that tag across space and time. For VHF telemetry data organisms are located through triangulating the radio signal (Mech, 1983). Tower-mounted, vehicle-mounted, or hand-held receivers are used to detect the tag signal and once detected a 3-dimensional fix is sought to improve precision in the location. Collecting VHF telemetry locations can be time consuming and logistically challenging. GPS technology has largely alleviated the time-intensive procedures associated with VHF by allowing for more frequent locations (provided that a fix can be acquired) and automated or remote data download. With GPS technology, the radio tag communicates with satellites at preset intervals. Like VHF, at least a three-dimensional fix is required for GPS to precisely locate the radio tag. Regardless of the telemetry system used, the goal is to accurately locate the organism in space and connect the location to the environmental variables that describe habitat.

The main advantages of telemetry technology are the opportunities to seemingly study animal behavior in a non-invasive fashion. The close observation of organisms, characteristic of direct visual observation and some visual and auditory survey methodologies, has been scrutinized for the potential influence of the researcher's presence on natural behavior (see Strum and Fedigan, 2000 for review). Similarly, potential bias associated with trapping can limit inference on habitat selection. In theory, telemetry technology permits collection of data via remote radio towers or portable telemetry units that minimize observer effects on individual organisms. However, in some instances the weight and resistance associated with telemetry tags have been shown to influence both organism behavior and survivability (Marcström *et al.*, 1989; Swenson *et al.*, 1999; Barron *et al.*, 2010). A common limitation of telemetry studies is the ability to capture and instrument enough animals for sound ecological inference on habitat selection.

Measurement of Environmental Variables

The characterization of habitat selection depends not only on the identification of organism locations but also the intersection of these locations with the environmental variables that define habitat. Environmental variables can be collected using field-based sampling or digital remote sensing. With some of the organism location methodologies described above (visual observation and some survey or trapping designs) environmental variables can be collected at the time the organism is located (e.g., Gutzwiller *et al.*, 1994). Though direct observation and measurement of environmental variables is still widely used in habitat selection research, another common technique for describing environmental variables at certain locations is digital remote sensing.

Digital data portraying environmental variables are now widely available for download from numerous reputable governmental and nongovernmental organizations. These data include digital elevation models, bathymetry data, hydrology, soil type, geologic features, vegetation cover, landuse data, climate, and many others. Therefore, some components of habitat can be compiled with relative ease using modern, often freely available data. Caution must be exercised when using widely available digital data so that the research question is properly framed and issues of scale are reconciled (Roloff *et al.*, 2009). For example, the data vintage must be consistent with the spatial and temporal positions of the organism location (Roloff *et al.*, 2009). If vintage of the environmental data are not considered, then the researcher will potentially describe environmental variables that the organism of interest was not necessarily selecting. Furthermore, spatial and temporal resolutions are also important factors to

consider (Roloff *et al.*, 2009). As discussed below (*see* Issues of Scale and Resolution) the resolution of the environmental data can limit the ability to make valid inferences. Spatial resolution refers to the minimum mapping unit in the data layer. Temporal resolution refers to the frequency of data collection or update. Regardless of the technique used for acquiring or compiling data on environmental variables, the resolution of the data must align with the spatial and temporal scales of selection for the organism of interest.

Analytical Techniques

Measurement of habitat use through mapping of animal locations and associated environmental variables is only the first part of describing habitat selection. The next step is to make inferences, or reach conclusions on the basis of evidence and reasoning. Numerous analytical techniques have been used to infer habitat selection. Models have been the cornerstone of habitat selection research for the last 30 years. Common modeling techniques can be based on organism presence/or abundance and the relationship between environmental covariates and organism locations, utilization distributions (UDs) as calculated from organism locations, or animal movements.

Linear and Logistic Regression Models

A common analytical approach in habitat selection research is to regress environmental variables (i.e., predictor variables) against a response variable that is either continuous (linear regression) or binary (logistic regression). Linear regression models typically take the form;

$$Y_i = \mathbf{x}'_i\boldsymbol{\beta} + Z_i$$

where Y is the response variable at location i, $\mathbf{x}$ is a vector of the environmental variables at location i, $\boldsymbol{\beta}$ is a vector of the regression parameters, and Z accounts for the random error term. The random error term can have spatial and temporal dependencies that should be accounted for (Ver Hoef *et al.*, 2001). A common response variable in linear regression models is a count of the individuals that occupy an area. Within these models there is a tacit assumption that more individuals in an area represent greater selection for the corresponding environmental variables associated with that area. The interpretation of the parameter estimates that are output from these models is based on the expected amount of change in the response (organism count) given unit changes in the environmental variables.

Logistic regression is generally used to model the relationships between presence or absence of an organism and environmental variables. Logistic regression models take the form:

$$P(y = 1|\mathbf{x}) = \frac{\exp(\boldsymbol{\beta}'\mathbf{x})}{1 + \exp(\boldsymbol{\beta}'\mathbf{x})}$$

where the binary response variable is modeled given a value of 1 (presence or used), $\mathbf{x}$ is a vector of the environmental variables, and $\boldsymbol{\beta}$ is a vector of the regression parameters. Presence or use is based on measured organism locations whereas absence is often estimated. The creation of absence locations is a topic that has been widely scrutinized (*see* How is Available Habitat Defined?). Absence locations in available habitat tend to be estimated via a random or stratified sampling design (Keating and Cherry, 2004). For example in a random design, absence is randomly assigned to locations in a study area, with the assumption that these locations are not used by the organism or that the locations portray what is available to the organism for selection. In a stratified design, absence is assigned to different locations based on some predetermined stratification, perhaps by vegetation or soil type. Selecting the sampling strategy that best suits the data depends on the research question, the species of interest, and the composition of environmental variables within the study area. Regardless of the technique that is selected, at each presence and absence location environmental variables are measured and model parameters are fit.

A specific type of regression modeling is referred to as resource selection functions. The goal of resource selection functions is to develop probabilities of relative use of environmental variables within the study area (Manly *et al.*, 2002). The classic use versus availability exponential model is expressed as;

$$w(\mathbf{x}) = \exp(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1\mathbf{x}_1 + \boldsymbol{\beta}_2\mathbf{x}_2 \cdots \boldsymbol{\beta}_k\mathbf{x}_k)$$

where $\mathbf{x}$ is a vector of k environmental variables and $\boldsymbol{\beta}$ is a vector of the coefficients at each location. This model is calculated with a logistic regression based on maximum likelihood estimation. The environmental variables can be either continuous or categorical, though similar to logistic regression, resource selection functions are subject to critique for the exact methodology employed to generate available locations (Boyce, 2006).

Utilization Distributions

UDs have been used to model habitat selection by departing from strict point-based modeling (Marzluff *et al.*, 2004). Essentially, UDs stretch the probability of an animal occurrence across space and time. This technique improves on traditional home range estimators (e.g., minimum convex polygons) by recognizing that organisms use habitat nonuniformly. UDs enable the researcher to quantify those environmental variables that are used more often than others.

The creation of a UD is initially informed by the collected organism locations. The organism locations are used to develop density estimates of relative use across spatial and temporal dimensions. Researchers typically divide their dataset of organism locations based on biological relevance and the question being asked (e.g., sex, season, year). The latitude (Y) and longitude (X) coordinates from organism locations are mathematically analyzed to develop the Z dimension in the UD, which represents the probability of space use. This probability of use can then be manipulated to function as the response variable in any modeling design. A variety of techniques can be implemented to measure the environmental variables within the UD. For instance Marzluff et al. (2004) developed a resource utilization function based on the entire UD. However, this procedure could prove computationally challenging for extremely large UDs. Another option is to

implement a strategy for sampling the environmental variables within the UD (Marzluff *et al.*, 2004).

Animal Movement Models

One emerging area of modeling habitat selection is animal movement models. Given the potential for fine-scale temporal sampling from GPS telemetry tags, consecutive organism locations can be linked to simulate animal movement paths. Innovative new models incorporate temporal autocorrelation and can also be modified to integrate habitat availability within the context of an animal's UD while estimating model parameters via a Bayesian framework (Christ *et al.*, 2008). Furthermore, animal movement models can be fit using data collected at regular or irregular intervals (Johnson *et al.*, 2008). Animal movement models represent an exciting area of habitat selection research and have thus far proved integral to understanding migration patterns, responses to various time-sensitive phenomena (weather, temperature, disturbance events), and use of edge habitat (Schick *et al.*, 2008; Forester *et al.*, 2009; Morales *et al.*, 2010).

Inference

Among the scientific community, the term inference is used to describe the process of reaching logical conclusions based on evidence and reasoning. In habitat selection research inference is based on animal location and associated environmental variable data in an effort to understand organism decision-making. Inference from habitat selection research has resulted in formulation of multiple theories and concepts including niche theory (Petren, 2001; Hirzel and Le Lay, 2008), habitat suitability (Hirzel *et al.*, 2006), and specificity of habitat use (Devictor *et al.*, 2008). These theories and concepts have collectively added to our understanding of ecology.

The foundation of niche theory and resource partitioning includes a seminal publication by Joseph Grinnell in (1917) on the organism–habitat relationship for California thrashers (*Toxostoma redivivum*). Grinnell referenced a specific set of environmental variables that California thrashers were adapted to as 'critical conditions.' These conditions included humidity, slope, elevation, and composition of shrubbery. Grinnell (1917) posited that California thrashers were adapted to a specific set of environmental conditions and that the availability of these conditions limited distributional range, thereby defining niche. Later (in 1957), G. Evelyn Hutchison formally articulated the multivariate components of habitat as an *n*-dimensional niche, where the *n*-dimensions correspond to different environmental variables that are used by the organism (Morrison *et al.*, 2006). Niche theory has been a useful construct to guide development of analytical tools for habitat selection research.

The concept of habitat suitability is based on Grinnell's notion that some portions of the niche are better suited for organism fitness than others, and that habitat selection can help reveal those relationships. The concept of suitability contends that habitat varies in quality, that is, some habitat is more conducive to occupancy, reproduction, and survival than others. Numerous analytical approaches have been used to portray habitat quality, including linear and logistic regression, resource selection functions, UDs, and an expert-opinion-based modeling approach called habitat suitability modeling (Singh *et al.*, 2009). These techniques portray disproportionate habitat use, with the assumption that areas receiving more use or having a higher probability of occupancy or use correspond to higher quality. However, animal density or occurrence is not always a reliable correlate to habitat quality (Van Horne, 1983) and researchers have emphasized the importance of using fitness (often quantitatively expressed as viability) as the best indicator of quality (e.g., Roloff and Haufler, 2002; Mosser *et al.*, 2009)

Habitat suitability modeling has been conducted on a variety of species. One well-known example was performed in the mid 1990s. At that time it became clear that the wolf (*Canis lupus*) population in the Great Lakes region of the United States was growing rapidly. Concerned citizens, stakeholders, and scientists wanted to determine how widely wolves could potentially range. Mladenoff et al. (1995) developed a model that was used to map habitat suitability for wolves throughout the Great Lakes region. The model was built from habitat selection data collected on a population of well-studied wolves in northern Minnesota. The model was subsequently extrapolated to other areas in the Great Lakes region to estimate where wolves might colonize and survive. The research of Mladenoff et al. (1995) demonstrated, amongst other things, that wolves were capable of traversing large areas of unsuitable habitat in search of areas that provided the right mix of favorable habitat characteristics, a habitat use pattern that has been realized as wolf populations recovered in this region (Mladenoff *et al.*, 2009).

The occurrence, abundance, and fitness of some species are closely linked to specific environmental variables whereas other species exhibit more latitude in habitat selection. This observed difference in habitat specificity among species has resulted in two broad categories; habitat-specialists and habitat-generalists. Habitat-specialists are so suited to a specific set of environmental variables that they are vulnerable if these variables are altered. Habitat-generalists are more malleable in their dependence on specific environmental variables and can potentially thrive in disturbed habitat. In a world that is rapidly changing as a result of anthropogenic disturbance including human population growth, development, exploitation of natural resources, and accelerated climate change, habitat-specialists can experience significant population losses from habitat disturbance whereas habitat-generalists are more capable of adjusting to environmental change (Brook *et al.*, 2003).

Although the broad categories of habitat specialist and generalist imply autonomy, recent research has demonstrated that habitat functionality also depends on the composition of different species that occupy similar space and time. The term trophic cascade, for instance, describes an event where predators, at the top of the food chain, influence the vegetative community at the bottom of the food chain (Carpenter *et al.*, 1985; Ripple and Beschta, 2007). At the system level, predators are known to influence nutrient cycling through consumptive and nonconsumptive means (Bump *et al.*, 2009; Schmitz *et al.*, 2010). Furthermore, habitat heterogeneity and species diversity are positively associated (Tews *et al.*, 2004). Essentially, more varied environmental variables result in more potential niches and hence, greater species richness.

Limitations on Inference

How is Selection of Habitat Defined?

Point-based habitat selection models are built on the assumption that organism locations represent the full suite of selection decisions. In reality, sampling and technological shortcomings often limit, sometimes in a biased manner, the animal locations that are collected. For example, Beyer and Haufler (1994) demonstrated that inference on habitat selection by elk (*Cervus elaphus*) was different depending on whether the sample contained night-time locations. Technological advances, like modern GPS telemetry technology, have helped alleviate some of the logistical constraints associated with sampling. Indeed, use of GPS technology has revealed new insights into environmental variables that are infrequently used but nonetheless important (e.g., transitory areas between more highly used areas; Beyer *et al.*, 2010). However, GPS telemetry data do not represent a guaranteed solution as vegetation cover, weather, animal movement, and topography affect the likelihood of triangulating a position and the precision of the organism location. For example, topography is known to influence the ability of GPS to triangulate a location (Gantz *et al.*, 2006) and thus, resultant animal locations are biased to those areas where topography did not limit functionality of the GPS technology. By assuming that all organism locations in a data set represent the full suite of selection decisions, the importance of certain environmental variables is likely exaggerated whereas the value of others is likely diminished (Marzluff *et al.*, 2004). Hence, it is important for researchers to critically evaluate their animal location data, understand the spatial and temporal dimensions of the data, and adjust the scope of inference accordingly.

Another factor that limits our ability to make inference from habitat selection studies is the assumption that organisms are making selections for habitat that is ideal (i.e., organisms are free to choose among the full range of habitat conditions). Fretwell (1972) and Fretwell and Lucas (1969) described this process as the ideal free distribution. The concept implies that organisms are ideal in their perception of habitat and that they are free to choose among habitats. In reality, individuals are likely selecting less-than-ideal habitat because external pressures (e.g., dispersal constraints, intra and interspecies competition, source-sink dynamics, and human encroachment on habitat) limit the ability to perceive or move into ideal habitat. By studying organisms that are not behaving according to an ideal free distribution researchers risk identifying less than ideal habitat as ideal. This misidentification can have significant conservation ramifications if management activities inadvertently focus on less than ideal habitat. Herein lies the importance of using fitness as a metric of habitat suitability. Fitness should be highest in ideal habitats.

How is Available Habitat Defined?

Many logistic regression models and resource selection functions require the designation of available or presumably unused habitat. Because of the coarse temporal resolution characteristic of almost all animal sampling efforts, it is difficult to say with certainty that available habitat were unused. For instance, an animal may make use of available habitat between sampling intervals. Thus, the naming convention representing the binary response is traditionally used versus available habitat. It is important to note that this is a problem which does not pertain to the habitat selection research of plants and fungi. Within this context it is possible to identify true absence locations by visiting various habitat and confirming plant presence or absence (Turner *et al.*, 2003). Methods for estimating available habitat in animal habitat selection research have been scrutinized (MacKenzie, 2005; Boyce, 2006). One technique for estimating available habitat is based on randomly sampling locations from the study area where organisms were not recorded (Edge *et al.*, 1987; Aebischer *et al.*, 1993). Another technique uses the organism locations to create a home range, and then samples available habitat within the home range (Anderson *et al.*, 2005; Christ *et al.*, 2008) (**Figure 2**). The aforementioned techniques are similar, yet the latter limits the generation of available habitat to only that within the organism's home range whereas the former allows sampling throughout the study area without considering the organism's site fidelity. The problems associated with identifying available habitat are often based on the delineation of the spatial and temporal boundaries of what is considered available.

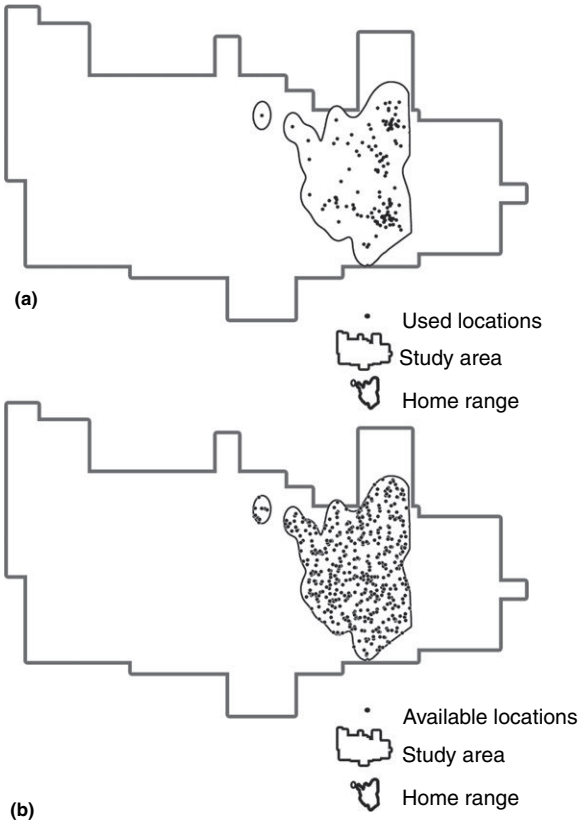


Figure 2 Example of a use vs. availability design with a random sample of available habitat drawn from within the home range; (a) shows the organism locations collected by the researcher that were used to create the home range, whereas (b) displays the available locations which resulted from a random sample of all available habitat within the home range.

Depending on the method used to identify available habitat, the ability to infer ecological relationships can vary substantially. Porter and Church (1987) found that results from the same habitat model could contradict depending on the manner in which available habitat was portrayed. This facet of traditional experimental designs for assessing habitat use and availability injects uncertainty into the inferences derived from this research. The problem is that decisions on how to define habitat availability are commonly made *a priori*. Study area boundaries are identified as a result of funding, logistics, or the delineation of political demarcations, none of which may be relevant to the biology of an organism. Despite this reality, study areas are one of the primary means by which habitat is considered available.

There is also an assumption that areas without recorded animal locations represent unselected habitat. The limitations of different organism location techniques (e.g., organisms can only be observed during certain times of day, measurement error), coarse temporal resolution of some telemetry data, and problems with detection probability influence whether the researcher is able to definitively designate available habitat as definitively unused. While measuring habitat selection, the probability of detecting organisms of interest can vary. Historically, researchers have assumed that detection probability was reliable and constant. In reality, conditions during measurement of habitat selection (vegetation cover, weather, animal behavior, and observer bias) can influence the researcher's ability to detect an organism. Failure to account for imperfect detection probability can impact inferences derived from habitat selection research (MacKenzie *et al.*, 2006). For instance, the importance of habitat where an organism is readily detected can be overestimated whereas habitat where the organism occurred but was not detected would be undervalued. However, once the probability of detecting an organism is calculated, this probability can be incorporated into models developed to quantify habitat selection (MacKenzie *et al.*, 2006). Regardless, near-continuous telemetry sampling or intense visual observation would be required to completely assuage this critique associated with how available habitat is defined.

Furthermore, there is an important spatial and temporal component to identifying what should be considered available habitat which can be referred to as habitat accessibility (Buskirk and Millspaugh, 2006). For example, if a sedentary organism is selecting habitat in the far western portion of its home range, then habitat 5 km away cannot be simultaneously considered available. Researchers must explicitly consider both habitat availability and accessibility in their habitat selection research to increase the validity of their inferences.

Issues of Scale and Resolution

Scale is critical to any examination of habitat selection. The components of scale include extent and resolution. Extent defines the boundary of the analysis; resolution refers to the smallest identifiable unit (also referred to as grain and minimum mapping unit). Inference derived from habitat selection research can vary depending on extent (*see* How is Available Habitat Defined?) and resolution (e.g., see Roloff *et al.*, 2009). Thus, an understanding of scale as it relates to the biology of the study organism and the research question is critical for valid inference (Turner *et al.*, 2001; Roloff *et al.*, 2009). Furthermore, ecological processes exhibit spatial and temporal variabilities and the results and inferences derived from habitat selection research directly depend on the scale of the analysis (Boyce, 2006).

Johnson (1980) articulated that organisms perceive their environment at varying scales and make decisions to satisfy certain life history events in orders that depend on these scales (**Figure 3**). The first-order represents a selection of habitat that defines the range of the species. Selection of habitat by individuals for the placement of home ranges within the species range occurs at the second-order. The third-order of the hierarchy represents within home range selections, for example, denning, feeding, or roosting areas, whereas the fourth-order addresses fine-scale (e.g., microsite, daily) habitat selections.

As organisms use their environment at varying spatial and temporal scales, failure to explicitly consider scale can distort inference. In a habitat modeling framework, scale affects the relationship between the response variable(s) and the environmental variables that describe the habitat. For instance, certain habitat selection studies divide the study area (extent) into a network of polygons, called habitat units (e.g., a grid lattice of polygons delineated at a specific resolution) (Servheen and Lyon, 1989; Boyce *et al.*, 2003). When habitat selection research depends on concepts like habitat units, identifying the scale of those units must be done in careful consideration of the research question, the biology of the organism, and the patterns and processes of the environment. Furthermore, it is recommended that several different scales be examined and model selection methods be utilized to identify the best-approximating model (Anderson *et al.*, 2005; Boyce, 2006).

Spatial and Temporal Autocorrelation

Spatial and temporal autocorrelations are two factors that influence the measurement of habitat selection. Autocorrelation can be viewed as the similarity among observations as a function of the space or time separation among them. Regression models have a statistical structure that requires independence of the environmental variables. As environmental variables (e.g., elevation, vegetation type, and soil structure) are more similar to one another in space and time the fine-scale sampling protocols associated with acquisition of some animal location techniques (e.g., direct visual observation, GPS telemetry) will yield data that exhibit correlated, directional dependencies. Failure to account for these inherent dependencies is a violation of regression model assumptions.

The issue of autocorrelation in space and time has direct ramifications on inference because failure to account for dependencies amongst the environmental variables imputes bias into the model design, potentially inflating the significance of certain relationships. The problem is that a researcher could infer an organism–habitat relationship that is stronger than it is in actuality, simply because a facet of the modeling design was neglected. However, regression models and RSFs can be structured to account for autocorrelations. As mentioned, a progressive area of habitat selection research involves the

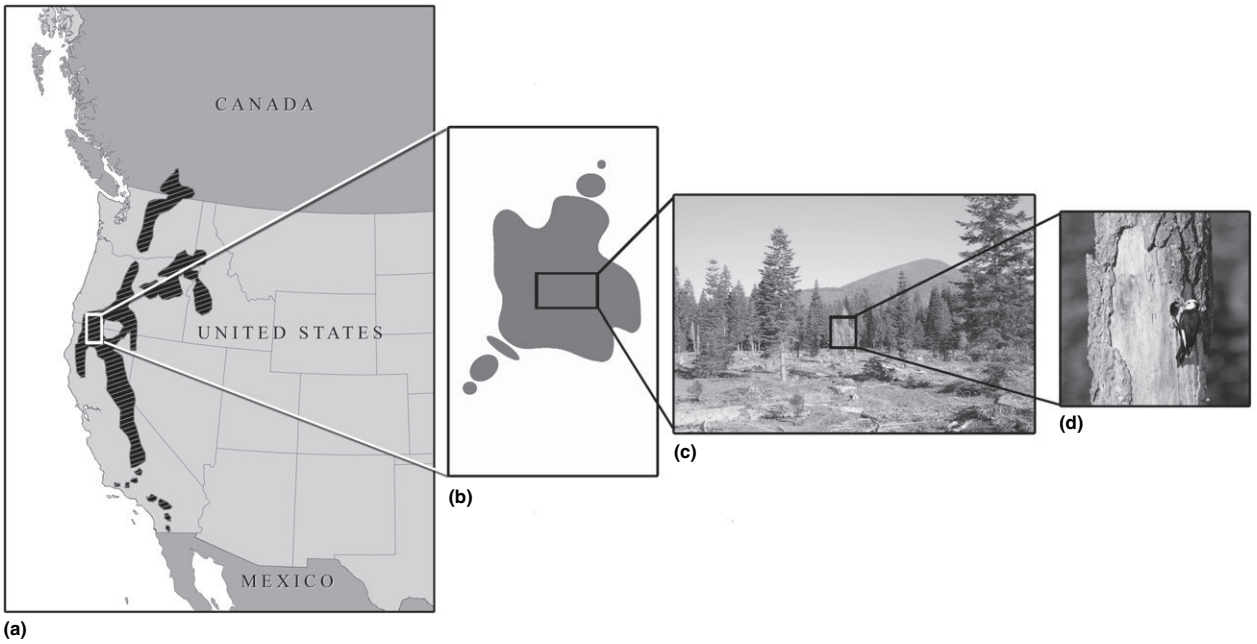


Figure 3 Four orders of hierarchical habitat selection outlined by Johnson (1980); (a) the first-order selection represents habitat selection that defines the range of the species, in this case the range of the white-headed woodpecker (*Picoides albolarvatus*). (b) in the second-order selection an individual white-headed woodpecker makes a selection for the placement of their own home range within the species range. (c) the third-order selection correspond to within-home range habitat selection, and (d) the fourth-order selection relate to fine-scale daily habitat selection.

integration of autocorrelation within animal movement models. When autocorrelation is explicitly included into the likelihood method of estimating model parameters more accurate interpretations of ecological relationships become possible (Boyce *et al.*, 2010).

Locational Imprecision

The unprecedented depth and scope of our ability to locate organisms, particularly with telemetry systems have fostered the assumption that measured organism locations are precise. In actuality, locational data can have considerable amounts of measurement error which is problematic for habitat selection research. In 1990, Gary White and Robert Garrott published an important book which, amongst other things, demonstrated the necessity of explicitly incorporating measurement error into habitat selection models. Given that habitat selection research strives to understand which habitat patches an individual is using in relation to those that are available, measurement error limits our ability to accurately detect true organism–habitat associations. White and Garrott summed up this sentiment when they wrote that habitat selection models "require a precise estimate of an animal's location so that it can be correctly placed in a habitat type" (White and Garrott, 1990:200). Without precise locations habitat selection models have the potential to produce erroneous results. Following White and Garrott (1990) numerous methods (e.g., ignoring the error, rescaling environmental variables, majority sampling, and bivariate-weighted techniques) were developed to explicitly incorporate measurement error into habitat selection models. For several decades these techniques were applied to habitat selection research. However, the ability of these techniques to ameliorate the error associated with organism locations had rarely been tested.

Montgomery et al. (2010) demonstrated that commonly used techniques for incorporating error into habitat selection studies were largely incapable of accurately portraying habitat use. However, by 2010 the primacy of VHF technology had been usurped by GPS telemetry data. GPS data are, on the whole, more accurate than VHF data with errors that are commonly <12 m, depending on topography, canopy cover, and geographic location (Cargnelutti *et al.*, 2007). Because of this documented precision, a common approach was to simply ignore the measurement error and proceed with the modeling design assuming that the organism locations were near perfect. Montgomery et al. (2011) examined the effects of ignoring measurement error consistent with GPS telemetry systems. Their analysis discovered that ignoring measurement error from precise telemetry systems could have considerable ramifications on the ability to make valid habitat selection inferences. Regardless, new models and analytical methodologies are frequently being designed to portray habitat selection with as little bias as possible. The progression of telemetry technology from the 1950s forward demonstrates that habitat selection research is ever-evolving, constantly advancing on what has been done previously with new techniques and innovative methodologies.

Conclusions

It is doubtful that the nineteenth century naturalists who painstakingly recorded organism–habitat associations through

visual observation could have possibly conceived of the technological advances that currently inform habitat selection research. Researchers have moved from hours spent every day in the field collecting data to mere seconds at the computer engaged in the very same activity. Nevertheless, the quest to understand ecology remains the same. The prospect of elucidating the ecology of an individual which might characterize a population or even the species itself has proved completely absorbing. In this effort researchers have made important technological innovations and have developed new and interesting models to describe habitat selection. These collective advances have led to a vast progression of the discipline of habitat selection research.

The relationship between organisms and their habitats is considered integral to species survivability and the inferences garnered from habitat selection research have informed management and policy decisions. This information is particularly useful in the dynamic twenty-first century where threats to environmental change are ubiquitous and persistent. There is no doubt that further technological and quantitative innovations will be an integral part of the continued development of habitat selection research.

See also: Biofuels and Biodiversity, Wildlife Habitat Restoration. Conserving Biodiversity Outside Protected Areas. Habitat and Niche, Concept of. Habitat Loss and Fragmentation. Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity. Landscape Ecology and Population Dynamics. Landscape Modeling. Species Distribution Modeling. Species–Area Relationships. Trophic Cascades

References

Aebischer NJ, Robertson PA, and Kenward RE (1993) Compositional analysis of habitat use from animal radio-tracking data. *Ecology* 74: 1313–1325.

Altmann J (1974) Observational study of behavior: Sampling methods. *Behaviour* 49: 227–267.

Anderson DP, Turner MG, Forester JD, *et al.* (2005) Scale-dependent summer resource selection by reintroduced elk in Wisconsin, USA. *Journal of Wildlife Management* 69: 298–310.

Baker JD and Johanos TC (2004) Abundance of the Hawaiian monk seal in the main Hawaiian Islands. *Biological Conservation* 116: 103–110.

Barron DG, Brawn JD, and Weatherhead PJ (2010) Meta-analysis of transmitter effects on avian behaviour and ecology. *Methods in Ecology and Evolution* 1: 180–187.

Benson E (2010) *Wired Wilderness: Technologies of Tracking and the Making of Modern Wildlife*. Baltimore, MD: Johns Hopkins University Press.

Beyer Jr. E and Haufler JB (1994) Diurnal versus 24-h sampling of habitat use. *Journal of Wildlife Management* 58: 178–180.

Beyer HL, Haydon DT, Morales JM, *et al.* (2010) The interpretation of habitat preference metrics under use-availability designs. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 2245–2254.

Bonter DN, Gauthreaux Jr. SA, and Donovan TM (2008) Characteristics of important stopover locations for migrating birds: Remote sensing with radar in the Great Lakes basin. *Conservation Biology* 23: 440–448.

Boyce MS (2006) Scale for resource selection functions. *Diversity and Distributions* 12: 269–276.

Boyce MS, Mao JS, Merrill EH, *et al.* (2003) Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. *Ecoscience* 10: 421–431.

Boyce MS, Pitt J, Northrup JM, *et al.* (2010) Temporal autocorrelation functions for movement rates from global positioning system radiotelemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 2213–2219.

Brook BW, Sodhi NS, and Ng PKL (2003) Catastrophic extinctions follow deforestation in Singapore. *Nature* 424: 420–426.

Bump JK, Peterson RO, and Vucetich JA (2009) Wolves modulate soil nutrient heterogeneity and foliar nitrogen by configuring the distribution of ungulate carcasses. *Ecology* 90: 3159–3167.

Buskirk SW and Millspaugh JJ (2006) Metrics for studies of resource selection. *Journal of Wildlife Management* 70: 358–366.

Cargnelutti B, Coulon A, Hewison AJM, *et al.* (2007) Testing global positioning system performance for wildlife monitoring using mobile collars and known reference points. *Journal of Wildlife Management* 71: 1380–1387.

Carpenter SR, Kitchell JF, and Hodgson JR (1985) Cascading trophic interactions and lake productivity. *BioScience* 35: 634–639.

Christ A, Ver Hoef JM, and Zimmerman D (2008) An animal movement model incorporating home range and habitat selection. *Environmental and Ecological Statistics* 15: 27–38.

Cutler TL and Swann DE (1999) Using remote photography in wildlife ecology: A review. *Wildlife Society Bulletin* 27: 571–581.

Cypher BL, Bjurlin CD, and Nelson JL (2009) Effects of roads on endangered San Joaquin kit foxes. *Journal of Wildlife Management* 73: 885–893.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London, England: John Murray.

Devictor V, Julliard R, and Jiguet F (2008) Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos* 117: 507–514.

Edge WD, Marcum CL, and Olson-Edge SL (1987) Summer habitat selection by elk in western Montana: A multivariate approach. *Journal of Wildlife Management* 51: 844–851.

Forester JD, Im HK, and Rathouz PJ (2009) Accounting for animal movement in estimation of resource selection functions: Sampling and data analysis. *Ecology* 90: 3554–3565.

Fretwell SD (1972) *Populations in a Seasonal Environment*. Princeton, NJ: Princeton University Press.

Fretwell SD and Lucas Jr. LH (1969) On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical development. *Acta Biotheoretica* 19: 16–36.

Gantz GF, Stoddart LC, and Knowlton FF (2006) Accuracy of aerial telemetry locations in mountainous terrain. *Journal of Wildlife Management* 70: 1809–1812.

Gardner B, Royle JA, and Wegan MT (2009) Hierarchical models for estimating density from DNA mark-recapture studies. *Ecology* 90: 1106–1115.

Gardner B, Royle JA, Wegan MT, Rainbolt RE, and Curtis PD (2010) Estimating black bear density using DNA data from hair snares. *Journal of Wildlife Management* 74: 318–325.

Garrett KL, Martin GR, and Dixon RD (1996) White-headed woodpecker (*Picoides albolarvatus*). In: Poole A (ed.) *The Birds of North America Online*. Ithaca, NY: Cornell Lab of Ornithology. Retrieved from the birds of North America online: http://bna.birds.cornell.edu/bna/species/252doi:10.2173/bna.252. Ithaca, NY. USA.

Gregory CJ, Carthy RR, and Pearlstine LG (2006) Survey and monitoring of species at risk at Camp Blanding training site, northeastern Florida. *Southeastern Naturalist* 5: 473–498.

Grinnell J (1917) The niche-relationships of the California thrasher. *Auk* 34: 427–433.

Gutzwiller KJ, Wiedenmann RT, Clements KL, and Anderson SH (1994) Effects of human intrusion on song occurrence and singing consistency in subalpine birds. *Auk* 111: 28–37.

Hirzel AH and Le Lay G (2008) Habitat suitability modeling and niche theory. *Journal of Applied Ecology* 45: 1372–1381.

Hirzel AH, Le Lay G, Helfer V, Randin C, and Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* 199: 142–152.

Johnson DH (1980) The comparison of usage and availability measurements for evaluating resource preference. *Ecology* 61: 65–71.

Johnson DS, London JM, Lea M-A, and Durban JW (2008) Continuous-time correlated random walk model for animal telemetry data. *Ecology* 89: 1208–1215.

Jurzenski J and Hoback WW (2011) Opossums and leopard frogs consume the federally endangered American burying beetle (*Coleoptera: Silphidae*). *Coleopterists Bulletin* 65: 88–90.

Kauffman MJ, Varley N, Smith DW, *et al.* (2007) Landscape heterogeneity shapes predation in a newly restored predator-prey system. *Ecology Letters* 10: 690–700.

Keating KA and Cherry S (2004) Use and interpretation of logistic regression in habitat-selection studies. *Journal of Wildlife Management* 68: 774–789.
London J, and Richmond E (2010). National Marine Mammal Laboratory Researchers Census Harbor Seals Along Coastal Alaska. Alaska Fisheries Science Center Quarterly Research Report, NOAA Fisheries, Seattle, WA, USA.
MacKenzie DI (2005) What are the issues with presence–absence data for wildlife managers? *Journal of Wildlife Management* 69: 849–860.
MacKenzie DI, Nichols JD, Royle JA, *et al.* (2006) *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. New York, NY: Academic Press.
Manly BFJ, McDonald L, Thomas DL, McDonald TL, and Erickson WP (2002) *Resource Selection by Animals: Statistical Design and Analysis for Field Studies*. Norwell, MA: Kluwer Academic Publishers.
Mann J (1999) Behavioral sampling methods for cetaceans: A review and critique. *Marine Mammal Science* 15: 102–122.
Marcström V, Kenward RE, and Karlbom M (1989) Survival of ring-necked pheasants with backpacks, necklaces, and leg bands. *Journal of Wildlife Management* 53: 808–810.
Marzluff JM, Millspaugh JJ, Hurvitz P, and Handcock MS (2004) Relating resources to a probabilistic measure of space use: Forest fragments and Steller's jays. *Ecology* 85: 1411–1427.
Mech LD (1983) *Handbook of Animal Radio-Tracking*. Minneapolis, MN: University of Minnesota Press.
Mladenoff DJ, Clayton MK, Pratt SD, Sickley TA, and Wydeven AP (2009) Changes in occupied wolf habitat in the Northern Great Lakes Region. In: Wydeven AP, Van Deelen TR, and Heske EJ (eds.) *Recovery of Gray Wolves in the Great Lakes Region of the United States: An Endangered Species Success Story*, pp. 119–138. New York, NY: Springer.
Mladenoff DJ, Sickley TA, Haight RG, and Wydeven AP (1995) A regional landscape analysis and prediction of favourable gray wolf habitat in the Northern Great Lakes region. *Conservation Biology* 9: 279–294.
Montgomery RA, Roloff GJ, and Ver Hoef JM (2011) Implications of ignoring telemetry error on inference in wildlife resource use models. *Journal of Wildlife Management* 75: 702–708.
Montgomery RA, Roloff GJ, Ver Hoef JM, and Millspaugh JJ (2010) Can we accurately characterize wildlife resource use when telemetry data are imprecise? *Journal of Wildlife Management* 74: 1917–1925.
Morales JM, Moorcroft PR, Matthiopoulos J, *et al.* (2010) Building the bridge between animal movement and population dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 2289–2301.
Morrison ML, Marcot BG, and Mannan RW (2006) *Wildlife–Habitat Relationships: Concepts and Applications*, 3rd ed. Madison, WI: The University of Wisconsin Press.
Mosser A, Fryxell JM, Eberly LE, and Packer C (2009) Serengeti real estate: Density vs. fitness-based indicators of lion habitat quality. *Ecology Letters* 12: 1050–1060.
Müller J and Brandl R (2009) Assessing biodiversity by remote sensing in mountainous terrain: The potential of LiDAR to predict forest beetle assemblages. *Journal of Applied Ecology* 46: 897–905.
Otto C and Roloff GJ (2011) Using multiple methods to assess detection probabilities of forest-floor wildlife. *Journal of Wildlife Management* 75: 423–431.
Pellet J and Schmidt BR (2005) Monitoring distributions using call surveys: Estimating site occupancy, detection probabilities and inferring absence. *Biological Conservation* 123: 27–35.
Petren K (2001) Habitat and niche, a concept of. In: Levin SA (ed.) *Encyclopedia of Biodiversity*, pp. 303–315. San Diego, CA: Academic Press.
Porter WF and Church KE (1987) Effects of environmental pattern on habitat preference analysis. *Journal of Wildlife Management* 51: 681–685.
Ralph CJ, Geupel GR, Pyle P, Martin TE, and DeSante DF (1993). Handbook of Field Methods for Monitoring Landbirds. General Technical Report PSW-GTR-149. Pacific Southwest Research Station, Albany, CA, USA.
Ripple WJ and Beschta RL (2007) Restoring Yellowstone's aspen with wolves. *Biological Conservation* 138: 514–519.
Roloff GJ, Donovan ML, Linden DW, and Strong ML (2009) Lessons learned from using GIS to model landscape-level wildlife habitat. In: Millspaugh JJ and Thompson, III FR (eds.) *Models for Planning Wildlife Conservation in Large Landscapes*, pp. 287–320. Amsterdam, Netherlands: Elsevier Science.
Roloff GJ and Haufler JB (2002) Modeling habitat-based viability from organism to population. In: Scott JM, Heglund PJ, Morrison ML, Haufler JB, Raphael MG, Wall WA, and Samson FB (eds.) *Predicting Species Occurrences: Issues of Scale and Accuracy*, pp. 673–685. Washington, DC: Island Press.
Schick RS, Loarie SR, Colchero F, *et al.* (2008) Understanding movement data and movement processes: Current and emerging directions. *Ecology Letters* 11: 1338–1350.
Schmitz OJ, Hawlena D, and Trussell GC (2010) Predator control of ecosystem nutrient dynamics. *Ecology Letters* 13: 1199–1209.
Seavy NE, Viers JH, and Wood JK (2009) Riparian bird response to vegetation structure: A multiscale analysis using LiDAR measurements of canopy height. *Ecological Applications* 19: 1848–1857.
Servheen G and Lyon LJ (1989) Habitat use by woodland caribou in the Selkirk Mountains. *Journal of Wildlife Management* 53: 230–237.
Singh N, Yoccoz N, Bhatnagar Y, and Fox J (2009) Using habitat suitability models to sample rare species in high-altitude ecosystems: A case study with Tibetan argali. *Biodiversity and Conservation* 18: 2893–2908.
Soisalo MK and Cavalcanti SMC (2006) Estimating the density of a jaguar population in the Brazilian Pantanal using camera-traps and capture–recapture sampling in combination with GPS radio-telemetry. *Biological Conservation* 129: 487–496.
Strum SC and Fedigan LM (2000) Changing views of primate society: A situated North American perspective. In: Strum SC and Fedigan LM (eds.) *Primate Encounters: Models of Science, Gender, and Society*, pp. 3–56. Chicago, IL: The University of Chicago Press.
Swenson JE, Wallin K, Ericsson G, Cederlund G, and Sandegren F (1999) Effects of ear-tagging with radiotransmitters on survival of moose calves. *Journal of Wildlife Management* 63: 354–358.
Tews J, Brose U, Grimm V, *et al.* (2004) Animal species diversity driven by habitat heterogeneity/diversity: The importance of keystone structures. *Journal of Biogeography* 31: 79–92.
Thorn M, Scott DM, Green M, Bateman PW, and Cameron EZ (2009) Estimating brown hyaena occupancy using baited camera traps. *South African Journal of Wildlife Research* 39: 1–10.
Turner MG, Gardner RH, and O'Neill RV (2001) *Landscape Ecology in Theory and Practice: Pattern and Process*. New York, NY: Springer Verlag.
Turner MG, Romme WH, Reed RA, and Tuskan GA (2003) Postfire aspen seedling recruitment across the Yellowstone (USA) landscape. *Landscape Ecology* 18: 127–140.
Van Horne B (1983) Density as a misleading indicator of habitat quality. *Journal of Wildlife Management* 47: 893–901.
Ver Hoef JM, Cressie N, Fisher RN, and Case TJ (2001) Uncertainty and spatial linear models for ecological data. In: Hunsaker CT, Goodchild MF, Friedl MA, and Case TJ (eds.) *Spatial Uncertainty in Ecology*, pp. 214–237. New York, NY: Springer-Verlag.
Wallace AR (1876) *The Geographical Distribution of Animals*. London, England: MacMillan and Co.
White GC and Garrott RA (1990) *Analysis of Wildlife Radio-Tracking Data*. San Diego, CA: Academic Press.

Hemiparasitism

David Smith, University of Maryland, College Park, MD, USA
Todd J Barkman and Claude W dePamphilis, Pennsylvania State University, University Park, PA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 317–328, © 2001, Elsevier Inc.

Glossary

Dodder Viney parasites of the genus *Cuscuta* that form haustoria on the stems of their hosts. The unrelated parasitic vine, *Cassytha*, is often referred to as Laurel Dodder.
Endophyte The portion of a parasitic plant that is embedded inside host tissue.
Endophytic parasite Parasites with vegetative bodies that are entirely endophytic and unobservable unless flowering.
Haustorium (pl. haustoria) The organ found in all parasitic plants that penetrates vascular tissue of the host plant and forms a functional bridge for uptake of water and nutrients from the host.
Hemiparasite A parasite that photosynthesizes.
Holoparasite A parasitic plant that lacks chlorophyll and cannot photosynthesize.
Mistletoe Parasites of the closely related families Loranthaceae or Viscaceae with haustoria that originate as root tissue and penetrate stems or roots (in Loranthaceae) of the host plants.
Mycotrophic (=mycoparasitic) plant A plant that obtains energy through mycorrhizal associations with saprophytic or other biotrophic fungi.
Root parasite Parasites that form haustoria on roots of the host plant.
Stem parasite Parasitic plants that form haustoria on stems of the host.

The Natural History of Parasitic Plants

Parasitic plants are strictly defined by the production of specialized feeding structures, known as *haustoria*, that form a functional bridge into their hosts. Through the haustorium, parasitic plants obtain water, mineral nutrients, sugars, and other host metabolic products. Approximately 3000 angiosperm species (1% of named species) are haustorial parasites; however, this total does not include any plants that feed heterotrophically on fungi (*mycotrophic* plants). Most parasitic plant species are *hemiparasites*; they photosynthesize and are only partially heterotrophic. The degree of heterotrophy in parasitic plants varies between autotrophy and the complete loss of photosynthesis, a condition known as *holoparasitism*.

Other than the presence of haustoria, there are few generalizations that can be made about parasitic plants because they display enormous variability in growth form and host dependence. In most species, haustoria are formed from root tissue of the parasite, but in a few species they are initiated from stem tissue. The haustorium, in turn, can penetrate the root or stem of the host plant. The major life-histories among parasitic plants are distinguished by these differences and they have important ecological implications.

Root parasites are extremely variable in growth form, occurring as trees, shrubs, and herbaceous perennials or annuals. Because their haustoria are often below ground, many root parasites do not demonstrate any obvious indications of parasitism. Experimental studies have compared hemiparasitic plants grown in pots with and without hosts. They found enormous variability among species in their ability to set seed in the absence of a host. The degree of parasitism is extremely variable in root parasites and the increase in seed set for any one species is different depending on the host.

Dodders and some *mistletoes* are both called *stem parasites* because they attach to the stems of the host; however, they have very different life histories. Mistletoes are closely related to root parasites and, in fact, their haustoria are derived from root tissue. Upon seed germination, their haustoria immediately penetrate the host's stem. Mistletoes most often grow on one individual, but in some cases they may sprawl along or hang down from the stems of the host plant. In contrast, dodders are closely related to nonparasitic vines. After germination, dodder seedlings begin to sprawl and twine along the stem of the host, forming haustoria from time to time where the new stems of the dodder contact the host. The rudimentary hypocotyl of the dodder eventually dies and it grows extensively beyond the original point of germination, often using multiple hosts.

All parasitic plants are obligate, requiring a host for survival, growth, or seed production, except for some root hemiparasites. The degree of parasitism is a useful concept for comparing the ecology of related species that vary in their dependence on the host. At one extreme, the ecology of facultative parasites is similar to their fully autotrophic relatives. At the other extreme are a few highly derived lineages of nonphotosynthetic *endophytic parasites* that have all of their plant body within the host tissue emerging only to flower and fruit. Among root holoparasites, there are enormous differences in the amount of biomass that is below ground. Some species are almost exclusively subterranean, growing as long-lived underground tubers, producing only flowers and fruits above ground.

Extremely derived holoparasitic plants are some of the great oddities of nature. In the 10th century, Arabic scholars wrote that parasitic plants performed the actions of animals by feeding from the juices of the host plant; they have the "souls of worms." In 1828, a naturalist named Trattanick proposed

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00207-0

that a separate category should be created to encompass them, the taxonomic equivalent of a mental asylum for the manic and delusional. Certainly the morphology of many holoparasites is nonplant like, and this led some early researchers to call them fungi!

The dissonance experienced on first learning about parasitism in plants betrays a common prejudice in some botanists; plants and parasites are naively regarded as opposites. As a result, parasitic plants have been misidentified and misunderstood. The "bearded-grape" was regarded as a monstrosity and assigned to its own genus; however, it was simply a grape infected by *Cuscuta!* Parasitism in most root hemiparasites was unnoticed by naturalists until very recently making the modern view of parasitic plants only about a hundred years old. In 1969, Job Kuijt's book, *The Biology of Parasitic Flowering Plants*, abolished many of the mysteries surrounding parasites. In this seminal work, Kuijt provided a critical analysis of all previous research and proposed a conceptual framework for understanding the evolution, morphology, physiology, and ecology of parasitic plants. While Kuijt's book remains the best authoritative source on the subject, new molecular evidence is being used to resolve the detailed phylogenetic history of parasitism, and a wide range of biochemical, genetic, and molecular approaches are now being applied to better understand parasite biology.

The Parasitic Plant Clades and their Relationships to Other Plants

Phylogenetic relationships between parasitic plant families and their affinities to nonparasitic plants are now well known for most species. A few families of root holoparasites have been difficult to place in a phylogeny because their floral morphologies are unique and they virtually have no vegetative morphology. Molecular phylogenetic study of many parasitic plant groups has been impaired by the fact that photosynthetic and other genes commonly used in plant molecular systematics are missing or evolve at greatly accelerated rate. However, recent studies of plant mitochondrial DNA sequences have resolved 11 or 12 independent origins of haustorial parasitism (**Figure 1**) distributed throughout angiosperm history.

The Sandalwood order (Santalales) are the largest, most diverse clade of parasitic plants in terms of growth habit, but all retain at least some chlorophyll and photosynthetic ability; in fact, some members may not even be parasitic. This group includes almost 2000 species of trees, shrubs, mistletoes, and at least one species of mainly endophytic mistletoe. Olacaceae and Opiliaceae are composed entirely of trees and shrubs that include both free-living as well as root parasitic members. Santalaceae are also trees and shrubs with both root and stem parasites. The mistletoes belong to one of two large families, Loranthaceae and Viscaceae. The stem and root parasites of Loranthaceae are known as showy mistletoes because of their long, colorful, tubular flowers. Viscaceae, including the "dwarf mistletoe," *Arceuthobium*, are all stem parasites. In addition to the two mistletoe groups, two other families of Sandalwoods, Misodendraceae and Eremolepidaceae, are stem parasites.

About 1700 root parasitic annuals and perennials are found in what has traditionally been classified as two closely related families, Scrophulariaceae and Orobanchaceae. As a whole, they represent the most diverse clade of parasitic plants in terms of variation in degree of heterotrophy. Traditional Orobanchaceae includes only holoparasites, while the large family Scrophulariaceae (ca. 4500 spp.) has nonparasites, hemiparasites and even some holoparasites. Molecular phylogenies based on chloroplast gene sequences have found the distinction between the two families to be artificial—a strongly supported clade contains all known parasites from both groups, but most of the rest of traditional Scrophulariaceae are not closely related to this group. In other words, Scrophulariaceae appear to be polyphyletic, while the parasitic lineages are quite clearly monophyletic, suggesting a single origin for parasitism in this group. For this reason, dePamphilis and colleagues have proposed to include all of the parasites (plus a few closely related nonparasitic genera) in an expanded Orobanchaceae. Within this large group, holoparasitism appears to have evolved not less than five times. The agriculturally important genera *Striga* and *Alectra* (both mostly hemiparasitic) and *Orobanche* (holoparasitic) are members of this clade, as are the large hemiparasitic genera *Pedicularis*, *Euphrasia*, and *Castilleja*.

The remaining parasitic plant families are not nearly as large or diverse as Orobanchaceae or Santalales. Plants commonly referred to as dodders are found in two families Cuscutaceae and Lauraceae. Cuscutaceae is a family of about 150 species of one genus, *Cuscuta*, that are distributed nearly worldwide. Lauraceae is a large family comprised of mostly nonparasitic trees that also includes one hemiparasitic genus, the vine *Cassytha*. *Cassytha* has about 20 species, most of which are endemic to Australia or Africa, with one species found throughout the tropics. Although strikingly similar in gross morphology, these two parasitic groups originated independently. *Cuscuta* is related to morning glories (Convolvulaceae), in the order Solanales, while *Cassytha* belongs in Laurales, a basal angiosperm order. *Cuscuta* are usually considered to be holoparasites, but very low photosynthetic activity exists in some species; whether or not the chlorophyll-containing species can use their photosynthetic products is unknown.

Approximately 50 species of endophytic holoparasites are found in the traditionally recognized family, Rafflesiaceae. Although considered to be closely related, molecular phylogenies now indicate up to four independent origins (polyphyly) for the genera making up this collection of extreme parasites. Molecular evidence is in agreement with previous studies of pollen morphology that suggested Rafflesiaceae be divided into four families: Rafflesiaceae, Mitrastemonaceae, Apodanthaceae, and Cytinaceae. While floral morphology is diverse for these families, they were previously thought to be related because they all are endoparasites.

Rafflesiaceae includes the largest flower in the world, *Rafflesia arnoldii*, which measures 1m across. *Rafflesia* species only parasitize species of *Tetrastigma* vines (Vitaceae) and have an unusual reproductive biology. Emerging buds develop for approximately 1 year before they open and become receptive to pollinating flies that are attracted to the often fetid scent. Flowers die within 5 days leaving a narrow window of time for pollination to occur. Because of the short flower life span, and because female and male flowers are on separate plants, sexual

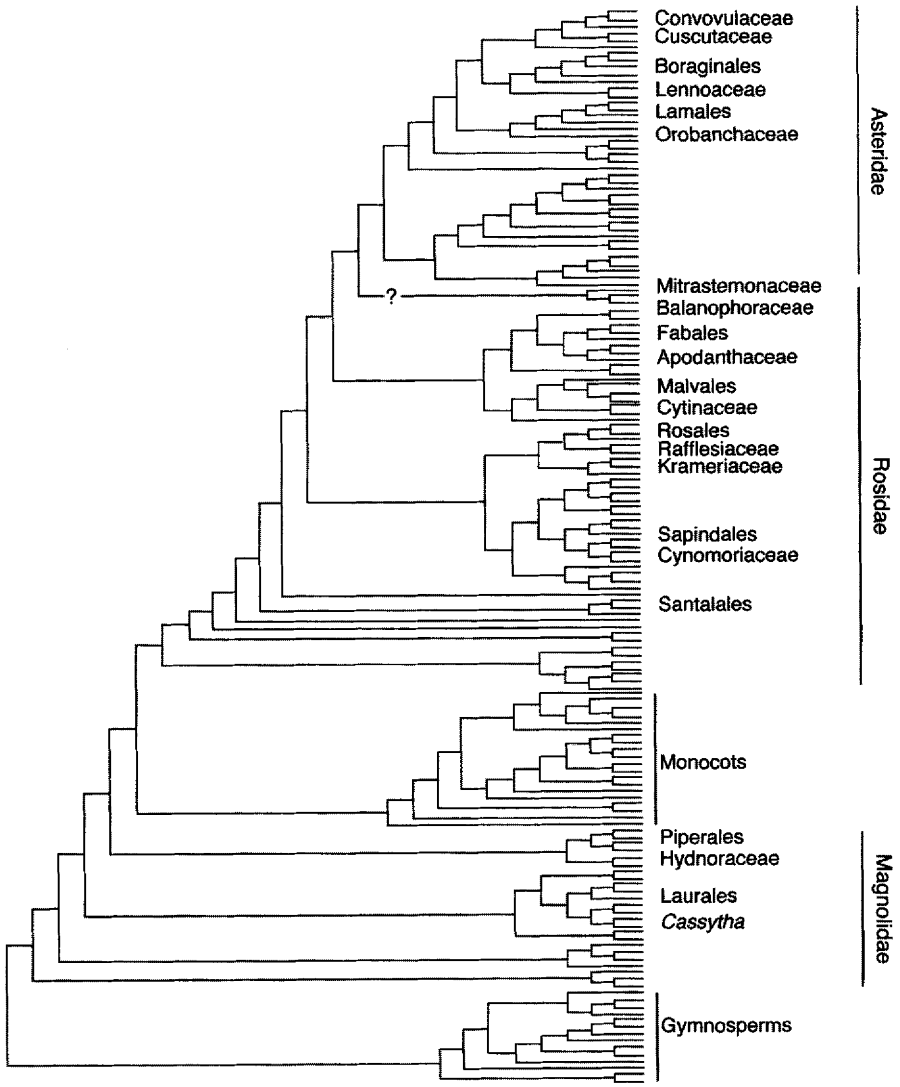


Figure 1 *Phylogenetic origins of parasitism in the flowering plants.* Summary tree of flowering plant phylogeny is based on the Neighbor-Joining analysis of mitochondrial gene sequences (*coxI* and *atp* A) conserved in all parasitic plants regardless of photosynthetic ability (T. J. Barkman and C. W. dePamphilis, unpublished). Note that 12 or 13 independent origins are resolved, with parasitic groups containing holoparasites or only hemiparasites shown near their closest resolved nonparasitic group. Mitrastemonaceae + Balanophoraceae is annotated as uncertain (?) because monophyly of this group depends on choice of gene and method of phylogenetic analysis.

reproduction is rare. Rafflesiaceae species are all exceptionally rare and endangered throughout their range, partly due to their reproductive strategy and partly due to habitat loss. As shown in **Figure 1**, recent mitochondrial DNA sequence data suggest that the closest relatives of Rafflesiaceae include the rose order (Rosales).

In contrast to *Rafflesia*, members of Apodanthaceae, *Pilostyles* and *Apodanthes*, produce tiny flowers only a few millimeters wide. Approximately 20 species are included in Apodanthaceae, which are distributed from the southwestern United States throughout the northwestern tropics with restricted localities in Africa, Australia, and the Middle East. Curiously, Apodanthaceae appear to be most closely related to the order Fabales, which includes their primary hosts; various species of legumes.

Cytinaceae include two genera, *Cytinus* and *Bdallophyton*. *Cytinus*, with 6 to 7 species, is distributed in the Mediterranean region, South Africa, and Madagascar and most commonly parasitizes species of *Cistus* (Cistaceae). *Bdallophyton* has only 1 to 4 species that are found mainly in Mexico and parts of Central America usually parasitizing members of the Burseraceae. The closest relatives of this family are classified within the cotton or hibiscus order, Malvales. Like the previously mentioned families, Rafflesiaceae and Apodanthaceae, Cytinaceae are not known to cause economic loss because they do not parasitize crop plants.

Mitrastemonaceae are specialists on trees in the family Fagaceae found in southeast Asia north to Taiwan and parts of central and South America. *Mitrastemon* is the only genus included in this family that is made up of two named species

that are totally white. Plants are protandrous and may produce large quantities of nectar.

Another small family, Hydnoraceae, is composed of about 20 species of root holoparasites native to tropical and southern Africa, Madagascar, and South America. Hydnoraceae have traditionally been considered close relatives of Rafflesiaceae, based largely on the shared holoparasitic habit with a reduced vegetative body and highly modified floral structures. However, molecular phylogenies now suggest that Hydnoraceae represents an independent origin of parasitism within the basal angiosperms, and that similarities with other holoparasitic groups have evolved convergently.

Balanophoraceae has about a hundred species in 17 genera, all of which are extremely reduced root holoparasites that have some of the smallest flowers in the world. Members of Balanophoraceae are found throughout the tropics and have broad host ranges but do not cause extensive crop damage. Balanophoraceae are sometimes included with the sandalwoods; however, molecular evidence does not support that relationship.

Cynomorium (Cynomoriaceae) is a holoparasite in the Mediterranean region comprised of two species that has sometimes been included within Balanophoraceae. As shown in **Figure 1**, Cynomoriaceae is not closely related to Balanophoraceae; instead it appears to be related to the order Sapindales, which includes the citrus family.

Krameriaceae includes 17 species of root hemiparasitic shrubs and herbs in the single genus *Krameria* that are often found in dry, open habitats in New World deserts and subtropics. *Krameria* species are unusual in that they produce floral oils that are used by oil-collecting bees. Traditionally, *Krameria* has been placed near the legumes (Fabaceae) or milkworts (Polygalaceae) due to a superficial resemblance of their flowers. Molecular evidence, however, unambiguously suggests a close relationship of Krameriaceae with Zygophyllaceae in spite of little morphological similarity between the two families.

Lennoaceae has two genera with five species of root holoparasites restricted to New World deserts in the southwestern United States, northwestern Mexico, and Columbia. All are herbaceous holoparasites of woody plants, including Boraginaceae (their closest known relatives) and Asteraceae. These strange parasites are spread by rhizomes often far beneath the sand or soil surface and are difficult to find unless flowering. *Pholisma sonorae*, also known as sand food, was traditionally eaten by indigenous people.

Finally, one gymnosperm genus, *Parasitaxus* (Podocarpaceae), is parasitic on other free-living Podocarpaceae in New Caledonia.

Evolution of the Degree of Parasitism

The evolution of parasitism has been difficult to study because intermediate evolutionary stages between non-parasitic ancestor and extremely derived holoparasites are often missing. The variability in the degree of parasitism and the multiple origins of parasitism make these plants ideal for studying the evolutionary origins of parasitism and the often rapid canalization of parasitic traits. The degree of parasitism refers to increases in the dependence that parasitic plants have on their hosts. Aspects of the degree of parasitism include the amount of carbon, water, or sugar obtained from the host, as well as major changes in the way that parasitic plants relate to a host. These major shifts include the evolution of stem parasitism or the evolution of an endophytic life history. The evolutionary factors that have influenced these events remain one of the most important questions about parasitic plants.

Changes in the degree of hetrotrophy represents one of the most interesting and best-documented cases of quantitative evolution of parasitism. Some hemiparasitic root parasites and mistletoes are almost entirely autotrophic—that is, they obtain a small fraction of their carbohydrates from their host whereas holoparasitism (100% heterotrophic) occupies a presumably irreversible end-point of this quantitative scale. Most clades of parasitic plants are entirely hemiparasitic or entirely holoparasitic. Krameriaceae are all hemiparasites, while the Hydnoraceae, Rafflesiaceae, Apodanthaceae, Mitrastemonaceae, Cytinaceae, Lennoaceae, Cynomoriaceae, and Balanophoraceae are entirely holoparasitic. *Cassytha* and *Cuscuta* are often called holoparasites, although these species have very low amounts of chlorophyll. *Cassytha* species are yellow-green in color, while *Cuscuta* species are more orange-yellow. In *Cuscuta*, chloroplasts are less than 10% as numerous as in autotrophic plants but the green color is masked by yellow pigments.

Members of the order Santalales are mostly hemiparasitic, but at least one species represents a very advanced stage of parasitism. In *Arceuthobium*, the endophyte is extremely well developed and the leaves are reduced to scales. The most derived parasite in the Sandalwoods is found in the endophytic parasite, *Tristerix aphyllus*. *Tristerix aphyllus* is a parasite of columnar desert cacti, whose endophytic life history may allow the parasite to escape the hot and desiccating desert conditions. Curiously, *T. aphyllus* and all *Arceuthobium* species retain some chlorophyll, and in spite of these extreme advances toward parasitism, true holoparasites are absent in the order.

In contrast to the conditions in Santalales, within Orobanchaceae, holoparasitism has evolved on several occasions. In fact, some species include hemiparasitic and holoparasitic populations and the degree of heterotrophy varies substantially. Peter Atsatt (1970) was able to produce holoparasites in a few generations in artificial selection experiments. This suggests that the complete evolutionary loss of autotrophy, as seen repeatedly in Orobanchaceae, may be relatively common in nature. One obvious conclusion is that holoparasitism is more likely to evolve in root parasites than in stem parasites. Since root parasites are often found in the dark understory, the selective advantage of retaining functional chloroplasts in root parasites may be much weaker than for sunbathed stem parasites. In fact, holoparasitism in beech drops (*Epifagus virginiana*, Orobanchaceae) is clearly irreversible because entire sections of the genome that contain the genetic instructions for photosynthesis have been deleted.

The evolution of the mistletoe life history represents a different kind of innovation in the degree of parasitism that is different from the evolutionary loss of chlorophyll. Stem parasitism may have evolved more than once within the Sandalwoods; two families, Santalaceae and Loranthaceae,

have both stem and root parasites. It is possible that these represent paraphyletic taxa, but it is also possible that study of intermediates between root and stem parasites in Santalales will lead to an understanding of the evolutionary forces that led to the evolution of stem parasitism from root parasitism.

Parasitism and Other Phenomena in Plants

The utility of the strict botanical definition of parasitism is most apparent when parasitic-host interactions are compared with other plant-plant interactions. Almost all plant-plant interactions are characterized by conflict, and many are ecologically parasitic because one species exploits another. In short, each aspect of parasitism is present in another kind of plant-plant interaction.

Parasitic plants acquire resources directly from another plant. Similar phenomena exist in root-grafted forest trees and in mychorrizal associations with fungi that indirectly connect plants. Root grafting is defined by a morphological continuum of the vascular tissue between two individuals that occurs in several species of forest trees. Interestingly, enough materials can be transferred through natural root grafts to sustain girdled trees. While root grafting is most common between closely related species, parasitic plants often form haustoria on a variety of species that are not closely related.

Mycorrhizal associations are very common in nature. Some mycorrhizal plants, such as indian pipe (*Monotropa*, Monotropaceae) have evolved complete heterotrophy and lost functional chloroplasts, much like holoparasites; actually these myco-heterotrophs are often confused for parasitic plants. In general, mycorrhizal associations are not species specific, and networks of mycelia and roots from a number of individuals can be interconnected thereby allowing the exchange of resources. In root grafts as well as mycorrhizal associations, source-sink relationships may exist between individuals; however, the precise relationship depends on ecological differences that may change over time. In haustorial parasitism, the parasitic plant is generally the sink while the host is the source.

Sometimes plants exploit the habitat created by another plant at the expense of the "host." For example, epiphytes use hosts primarily as anchors. Epiphytes intercept light, minerals, and water that might otherwise be available to the host plant. The reduction of available resources, plus the mechanical stresses on the supporting structure, may impose fitness costs on the host plant. A phenomenon related to epiphytism, but potentially more harmful to the "host," involves nurse-plant relationships, in which an established plant creates a microhabitat that facilitates the establishment of another species. For example, vines use the structural wood of trees to reach the forest canopy without investing in support themselves. In an extreme case, strangler figs. twine around the trunk of a host, sending down adventitious roots to eventually support itself. Often the fig grows into a large tree that densely shades and envelops the host and may kill it. In contrast, parasitic plants impose fitness costs by removing water, mineral, and carbon resources from the host plant. They may also physically damage conducting tissues and negatively affect the host's metabolism through hormonal interactions.

Seed Dispersal, Germination, and Prehaustorial Events

Before forming haustoria, the seeds of a parasitic plant must germinate close to a host plant and make contact. Seeds have little control over where they are dispersed and because seedlings are fragile, encountering a host is an important but difficult challenge.

New contacts between hosts and root parasites are initiated by root growth; either the host grows near the parasite or the parasite grows toward the host. Some species use chemicals exuded by host roots as a cue for finding the host such that seedlings grow chemotropically up gradients. In facultative root parasites, the plants can invest in root systems that "search" for hosts. In the meantime, these roots can absorb resources and the plants can usually persist until they contact a host. In contrast, obligate parasites do not invest in elaborate structures; they must make contact immediately or die.

Seed size and number and the degree of parasitism are correlated in parasitic plants. Holoparasites generally produce large numbers of small seeds. This correlation is probably driven by different strategies for managing the risk of finding a host. Producing larger, more provisioned seeds allows the facultative parasite to persist until it finds a host, but the production of many small seeds increases the probability that at least some seeds will land near a susceptible host root. In general there is a negative correlation in all plants between seed size and number, which is thought to be related to resource limitations. Theoretically, resources are unlimited for parasites unless the host is killed. Therefore resource limitations may not be the reason for the seed size-number tradeoff in parasites.

In some species of obligate root parasites, seed dormancy may be prolonged as a way to spread the risks associated with finding a host over time as well as space. In other words, if a seed is dispersed to a suboptimal site, dormancy allows germination to occur at later times if the site becomes appropriate. Seed dormancy is broken when chemical cues indicate the proximity of a host. The first root exudate discovered to induce parasite seed germination was Strigol. This chemical is found in cotton and it stimulates the germination of *Striga* seeds. It is not known how well seeds discriminate between chemical cues of hosts from nonhosts in their local environment, but seeds often make mistakes. For example, cotton has been used as a way of managing *Striga* in agriculture by causing seed germination in the absence of a suitable host (*see* Parasitic Plants as Pests, below).

Mistletoe stem parasites must germinate on the stems of their hosts, requiring that seeds be dispersed directly onto a stem or leaf. Seeds are dispersed in two ways. Ballistic dispersal projects seeds up to 50 ft from the original host at velocities up to 90 ft/s as reported in *Arceuthobium*. Birds also disperse seeds of mistletoes over long distances because fruits are the main source of energy and protein for some birds. After consumption, seeds may be defecated onto hosts or the sticky seeds may attach to the bird's beak and be carried directly to another host. The seed may be wiped onto the host plant by the bird as it tries to remove the seeds from its beak.

The Haustorium

Haustoria are extremely varied structures among parasitic plants. In root parasites, haustoria are easily identified on exposed roots. They appear as swollen tissue at a contact point between parasite and host. In dodders, haustoria superficially resemble pegs or suction cups that connect the host and parasite. In general, haustorial cells occupy intercellular spaces and displace the host tissue, but enzymes also digest the host cell walls. Once a parasitic plant has encountered a host, it must penetrate the cambium and establish an interface. The interface is extremely varied. In most cases, the parasite forms a continuum with the xylem of the host plant, but in others (e.g., *Cuscuta*), the parasitic plant taps into the phloem. Host-derived materials may be transferred through strawlike intrusions into the host vascular tissue, or they are simply absorbed across cell walls. Haustorial cells near host tissue are usually rich in mitochondria and rough endoplasmic reticulum suggesting they are probably actively producing proteins that are likely used to produce digestive enzymes. After forming the initial haustorium, parasites may enhance local root growth to increase the number of haustoria and strength of the connection with the host. In stem parasitic mistletoes, seeds germinate and send out a radicle that grows into the host through a stomate.

The Physiology of Parasite–Host Interaction

Once a haustorium has established, the question remains: Why do materials move from host to parasite? Part of the answer appears to be quite simple for most xylem-tapping species: parasitic plants usually transpire at higher rates than their hosts to maintain a gradient in water potential across the haustorial boundary. The stomata of most plants close at night to conserve water, and they may also constrict during the day to reduce rates of transpiration under drought conditions. However, in many parasitic plants, stomata remain open at night, and they tend to maintain a high rate of transpiration throughout the day, even under drought conditions. Parasitic plants are among the least efficient plants at water use. In extreme cases, some species have glands to exude extra water. The net effect is that water and other dissolved substances move down the water potential gradient via mass flow from the host to the drier parasite.

If all the transfer of resources was facilitated in this manner, the concentration of nutrients should be higher in the parasite, but in the same proportion as the host plant. These general patterns are not found—instead, calcium generally has the same concentration as the host, while nitrogen, phosphorus, and potassium are enriched. Some evidence suggests that there is a component of active transport in parasitic plants. The haustorial cells at the interface have enhanced concentrations of mitochondria and show signs of being able to mobilize energy. It is possible that the main function of these cells is to pump unwanted materials back to the host. The potential benefits of reducing calcium concentrations and the mechanism(s) that make it possible are not fully understood.

The Ecological Advantages of Parasitism

In facultative parasites, the advantages of parasitism can be demonstrated experimentally. When grown with a host, parasitic plants look healthier, grow bigger, have more flowering branches, and produce more seeds. Compared with a patch of soil, the resources available in the vascular tissue of a host are extremely enriched. Parasitism allows plants access to a rich, hydroponic nutrient source. The benefit of parasitism can be mimicked experimentally by growing parasites in nutrient-rich solution culture.

A genus of annual hemiparasites, *Cordylanthus*, blooms during the hottest time of the year in an arid environment. Few other annual plants bloom at that time of year, and it is possible because the parasites use roots of perennial plants to gain access to groundwater. This unusual flowering time may also permit the parasites to have more exclusive access to pollinators.

In addition to nutrients and water, parasitic plants absorb secondary chemicals from hosts. Recent experiments by Michelle Marvier showed that herbivory is slowed and herbivores have lower fitness on parasitic plants that are attached to hosts that deter herbivory. Parasitic plants use a wide range of hosts, and, as a consequence, may accumulate a diverse and changing complement of protective secondary compounds. This strategy may effectively make parasites "moving targets" that are difficult for herbivores to "hit" in ecological or evolutionary time.

Host Specificity

In general, parasitic plants are considered to be host generalists. The advantage of this is quite simple; seeds can not choose where they are dispersed. Typically, a higher fraction of mistletoe seeds successfully establish on the species that is most abundant locally. In addition, dodders preferentially coil toward good hosts. Measuring host selectivity is much more difficult in root parasites than stem parasites. However, careful measurements by Gibson and Watkinson have demonstrated that root parasites are generally selective.

Some species are only found on one species or genus. The entire family, Rafflesiaceae, is composed of host specialists, as discussed earlier. Two closely related species of Orobanchaceae, *Epifagus virginiana* and *Conopholis americana*, are specialized on beech and oak trees, respectively. *Striga* (Orobanchaceae) tends to utilize grasses, and some species are highly host specific, even to the extent of requiring a particular subspecies as a host.

A tempting generalization is that species that are more dependent on their hosts tend to be more host specific. Kuijt rejected this hypothesis because most species of root holoparasites are host generalists. However, the most host specific species, including *Epifagus* and *Conopholis*, Rafflesiaceae, and *Tristerix aphyllus*, all represent extremely derived parasites. It may be that the endophytic lifestyle or the subterranean habit releases these plants from other selective pressures, and not that they become host specific because they are extremely derived. New phylogenetic evidence may provide more rigorous tests of this hypothesis.

Parasitic Plants as Pests

Many species of parasitic plants are economically important because they reduce crop yields. Parasitic plants remove nitrogen, phosphorus, and other mineral nutrients that may reduce the host's ability to grow. Parasitic plants have nutritional requirements that are very similar to their hosts, therefore the damage caused by a parasitic plant is often directly proportional to its biomass. In addition, parasitic plants are less efficient at water use than their hosts, so parasitized hosts will tend to be more water stressed than unparasitized hosts. Dodders do not have the same dramatic effects on water balance as mistletoes or root parasites, but their haustoria remove sugars and stunt developing fruits. In general, crops that are infested with parasitic plants grow slower, have lower yields, and are more susceptible to disease. In addition to crop plants, parasites are important pests of trees. The invasion of mistletoe haustoria into host trees can distort wood or alter growth patterns, thereby decreasing the value of the timber. Some mistletoes and dodders damage fruit trees or ornamental plants. The damage caused by these parasites is usually measured in lower yields, the loss of aesthetic value to ornamentals, or the increased effort required to control them. An enormous amount of research has focused on the root parasitic weeds, mistletoes, and dodders.

Root Parasitic Weeds

The most economically important root parasites are the witchweeds (*Striga*, Orobanchaceae). *Striga* includes approximately 17 species, of which, 11 are known to be pests. Most witchweeds have a tendency to parasitize grasses including sorghum, pearl millet, finger millet, rice, maize, and sugarcane. The most important witchweed species is *S. hermonthica*, which ranges from Ethiopia and Sudan through Sahel, north into Arabia, and south to the Ivory Coast, Nigeria, Angola, Namibia, and the Lake Victoria basin. *Striga asiatica* has a much wider range that includes much of Africa, parts of India, China, and Australia. In addition, *Striga asiatica* was accidentally introduced into North America in the 1950s. One species, *S. gesneroides*, has very different host range from the other witchweeds; it uses a diverse set of broad-leaved hosts. It is best known for the damage it causes to cowpeas in West Africa, but it also damages tobacco in East Africa.

Among root parasites, the holoparasitic broomrapes (genus *Orobanche*, Orobanchaceae) are second to the witchweeds in terms of their economic impact. The geographical ranges of the economically important species are centered in the Middle East, but they are found further west and south into Africa, west and north into Europe, east into India, Pakistan, and Nepal, and north into Afghanistan, and several countries that were once in the Soviet Union. In general, broomrapes have wide host ranges. *Orobanche ramosa* attacks crops in Solanaceae, Brassicaceae, and Fabaceae. It sometimes attacks onions, but in general it is never found on grains (it was once reported on maize). The host range of *O. cernua* is more restricted; it parasitizes only Asteraceae (mainly sunflower) and Solanaceae (tobacco, tomato, and eggplant). *Orobanche crenata* Forsk. has a wide host range, but it is most important as a pest of Fabaceae, especially faba beans and some Apiaceae (carrots).

The genus *Alectra* (Orobanchaceae) is closely related to the witchweeds. Of the 30 species, 4 are notable pests. *Alectra vogelii* and *A. picta* are pests of cowpeas and other pulse crops in semiarid Africa. *Alectra orobanchoides* occasionally attacks sunflowers or tobacco in South Africa and *A. fluminensis* parasitizes sugarcane in Central and South America. Several other genera in Orobanchaceae have agriculturally important species, but they are not nearly as important as the witchweeds, broomrapes, or *Alectra*. Other genera with notable or potentially important pest species include *Buchnera, Ramphicarpa, Odontites, Rhinanthus, Aeginetia, Melampyrum*, and *Christonia*.

Most control methods for root parasites focus on reducing host damage as well as seed set. Witchweeds and broomrapes produce hundreds of thousands of small, dust-like seeds per plant. Furthermore, the seeds can lie dormant in the soil for several years, and therefore reducing the size of the seed bank involves enormous effort. The most effective available method of control involves hand pulling the plants before they set seed. Although most of the damage to the host has been done by the time the parasitic plants emerge, hand pulling can reduce some of the damage. More important, early hand pulling prevents reseeding. If reproductively mature plants are pulled, care should be taken to prevent the seeds from being dispersed, and the adult plants should be burned.

More direct chemical methods can be used to reduce the seed bank. Ethylene gas, or artificial chemicals that mimic germination stimulants, can be applied to break dormancy and induce germination at times when no appropriate hosts are available. Alternatively, fumigants are applied to fields that kill seeds as well as other soil organisms. Raising the temperature of the soil by covering it with black plastic can also kill seeds.

Trap cropping, catch cropping, and crop rotation are other effective ways of managing the land. One option is to let the land lie fallow for 10 to 20 years to reduce the seed bank. This may not be an option where the demands on land use are heavy. Planting an unsuitable host is a more effective alternative that may allow the land to be used. Some alternative crops, called trap crops, induce seed germination but are not suitable hosts. A third option involves catch cropping. The land is seeded with susceptible hosts to induce germination of parasitic plants, but the crops are destroyed before the parasite can reproduce. Mixed cropping involves planting suitable and unsuitable hosts, and it has been demonstrated to reduce the impact of some witchweeds. The mechanisms, however, are not well understood.

Herbicides can also be used to control witchweeds. Herbicides applied directly to the soil after planting but before crop seedlings emerge can reduce or delay witchweed establishment. Systemic herbicides that are applied to the crop immediately after emergence may become translocated and concentrated in the parasites, killing them. The timing of application of systemic herbicides is critical because it must degrade before the crop begins to fruit. Some herbicides can be applied directly to the emerged witchweed to control them.

Damage to crops by witchweeds is minimized when nitrogen fertilizer is applied to host crops. This method has the added advantage of helping increase yields, but the benefits of

using this method alone over a long period of time have not been demonstrated. The mechanisms that allow nitrogen fertilization to decrease the impact of the parasitic plants are not fully understood. It does seem clear, however, that on poor soils, the witchweed infestations increase in intensity as the quality of the soil decreases; the result is a feedback that continues until the land is almost useless and must be abandoned.

One of the most promising control options is the development of resistant varieties of crop plants through plant breeding. As new resistant varieties are found other methods of control become unnecessary. Resistant varieties of many crops have not been discovered, however, therefore this is a very active area of research.

Parasitic plants may pose the largest problem to rural, subsistence farmers. As pressures on land use increase, the ability to use crop rotation methods decreases. Furthermore, subsistence farmers often do not have the means to buy fertilizers, herbicides, or expensive resistant seeds. For these farmers, the feedbacks between poor soils and parasitic plant infestation may eventually cause the farmers to abandon plots of land and move elsewhere.

Mistletoes

Several species of mistletoes are economically important pests of fruit trees, ornamental trees, or timber. Most mistletoes have a similar ecology, so the impact and control measures are easy to generalize. The exceptions are dwarf mistletoes (genus *Arceuthobium*, Viscaceae) that have a much more highly developed endophytic component and inconspicuous, scaly leaves. Because of the highly developed endophyte, physical removal of dwarf mistletoes is more difficult than other mistletoes.

Most species of dwarf mistletoes are extremely host specific. For example, *A. douglasii* is a parasite of Douglas fir, and *A. tsugense* is a parasite of the western hemlock. The dwarf mistletoes are parasites of pine trees in North America, the Himalayas of India, Pakistan, and Bhutan, and in southwestern China. Dwarf mistletoes are common throughout the Rocky Mountains where large regions of forests are infested and the timber volume and quality are affected.

Two other genera in Viscaceae are notable for their economic impact. *Phoradendron* is a genus of about 190 species distributed throughout North and South America. Of these, the most important are *P. serotinum*, found on an extremely wide variety of hosts (but never conifers), and *P. piperoides*, found on cocoa in Costa Rica. *Viscum* is a genus of about 60 species that include *V. album* that has an extremely wide host range including fruit trees, pines, and poplars in Europe and persimmon in China. Another agricultural pest is *V. cruciatum* that parasitizes olives in the Middle East.

Many species of Loranthaceae cause economically significant loss. *Dendrophthoe falcata* is a pest of fruit trees in India and teak in Kerala. *Dendropthoe pentandra* parasitizes rubber and kapok in Indonesia. *Helixanthera mannii* is a pest of citrus and coffee in West Africa, and *H. parasitica* is a pest of citrus in the Himalayas. *Tapinanthus bangwensis* is an important pest of cocoa and cola in Ghana, while *T. globiferus* grows on coffee, citrus, and other fruit trees in Ethiopia. Some other genera that have economically important species are *Amyema*, *Englerina*, *Loranthus*, *Macrosolen*, *Oryctanthus*, *Phragmanthrea*, *Scurrula*, and *Struthanthus*.

Some resistant varieties of poplars, pines, and oaks do exist, but very little variability for resistance to mistletoe infections exists and, in general, breeding for resistance has not been a viable option. Biological control by insects, pathogens, or other mistletoes remains a possibility, but none of these methods have been successful. Instead, control of the leafy mistletoes involves pruning off infected limbs, or chemical control. Pruning infected limbs removes the mistletoe, but the limb must be cut below the point of the infection to remove the entire endophyte. This practice is extremely labor intensive and impractical on a large scale. With *Arceuthobium*, the endophyte is often extensive, and the limb must be severed 20 to 30 cm below the lowest parasite shoots. In extreme cases, it may be more cost-effective to burn the forest. In less extreme cases, early harvest followed by controlled burns may reduce the mistletoes. Some successful chemical control methods have been found. The leaves of mistletoes may be sprayed with ethephon once the host drops its leaves. Alternatively, herbicides can be injected into infected limbs below the point of infection, accumulate in mistletoes, and eventually kill them without severely affecting the host.

Dodders

The economic impact of *Cuscuta* species is greatest in forage crops such as alfalfa and clover; however, crops like citrus, coffee, peach, litchi, flax, linseed, and other crops as well as ornamental plants all suffer from dodder attack. The most economically important dodder is *Cuscuta campestris*. This species attacks alfalfa and has been shown to reduce forage yield by as much as 57% over a 2-year period. It also affects Niger seed in India and many vegetable and flower crops. Of the 20 species of *Cassytha*, the most damaging is *C. filiformis*. It is distributed throughout the tropics and parasitizes a wide range of hosts. It is a problem on citrus in India and Tanzania and on *Pinus massonia* in China.

Mechanical control methods for *Cuscuta* include hand pulling, crop rotation, burning, delaying planting until after *Cuscuta* have germinated, or deep ploughing to reduce the seed bank. Few resistant varieties of crops are known. Chemical control methods include fumigants to eliminate the seed bank, herbicides applied to the soil to prevent seedling growth, and herbicide application after seeds have germinated to prevent establishment. Some methods of biological control have been established with insects and pathogens, but the scope of these has been limited.

Conclusions

Parasitic plants have been understudied, but recent studies have improved our understanding of their evolution, ecology, molecular biology, and physiology. Haustorial parasitism has evolved independently at least 12 times within angiosperms and, surprisingly, complete loss of photosynthesis has occurred multiple times within some lineages. Most parasitic plants exhibit near-normal levels of photosynthesis like their

autotrophic ancestors, whereas others are incapable of photosynthesis such as the derived endophytic holoparasites. Study of parasites that photosynthesize at intermediate levels will be important to understand the evolution of holoparasitism. Development of sophisticated molecular methods will be important for understanding the bases of haustorial development and should ultimately enhance efforts to control economically important crop pests. Further ecological and physiological studies aimed at understanding host-parasite interactions will also enhance efforts to control parasitic weeds.

Acknowledgments

We thank the NSF (grants DEB-91029258, DBI-9604814, and DEB-9811362 to C. W. D.) for funding our research on parasitic plants, and Joel McNeal for critical reading of the manuscript. Other important contributions of data, samples, or discussion, were also made by other members and friends of the dePamphilis lab, past and present, and by numerous botanical gardens and botanists with an interest in parasitic plants.

See also: Coevolution. Parasitism. Parasitoids

References

Atsatt PR (1970) Hemiparasitic Flowering plants: Phenotypic canalization by hosts. *Nature* 225: 1161–1163.

dePamphilis CW (1995) Genes and genomes. In: Press MC and Graves JD (eds.) *Parasitic Plants*, pp.177–205. London: Chapman and Hall.

Finlay RD and Read DJ (1986) The structure and function of the vegetative mycelium of ectomycorrhizal plants. I: Translocation of C-labelled carbon between plants interconnected by a common mycelium. *New Phytologist* 103: 143–156.

Graham BF Jr. and Bormann FH (1966) Natural root grafts. *Bot. Rev.* 32: 255–292.

Hawksworth FG and Weins D (1972) Biology and classification of dwarf mistletoes (*Arceuthobium*). *Agriculture handbook no. 401*. Washington D.C: U.S. Government Printing Office.

Kuijt J (1969) *The Biology of Parasitic Flowering Plants*. Berkeley, CA: University of California Press.

Marvier MA (1996) Parasitic plant-host interactions: Plant performance and indirect effects on parasite-feeding herbivores. *Ecology* 77: 1398–1409.

Musselman LJ (1980) The biology of *Striga, Orobanche*, and other root-hemiparasitic weeds. *Annual Review of Phytopathology* 18: 463–489.

Musselman LJ (1987) *Parasitic Weeds in Agriculture*, Vol. I Striga., Boca, Raton, FL: CRC Press.

Nickrent, DL (2000, January) The parasitic plant connection, http://www.science.siu.edu/parasitic-plants/index.html.

Nickrent DL, Duff RJ, Colwell AE, *et al.* (1998) Molecular phylogenetic and evolutionary studies of parasitic plants. In: Soltis D, Soltis P, and Doyle J (eds.) *Molecular Systematics of Plants, II.* Chapter 14.

Parker C and Riches CR (1993) *Parasitic Weeds of the World: Biology and Control*. Wallingford, UK: CAB International.

Press MC and Graves JD (eds.) (1995) *Parasitic Plants*. London, UK: Chapman & Hall.

Stewart GR and Press MC (1990) The Physiology and biochemistry of parasitic angiosperms. *Annual Review of Plant Physiology and Plant Molecular Biology* 41: 127–151.

Yoder JI (1997) A species-specific recognition system directs haustorium development in the parasitic plant *Triphysaria (Serophalariaceae). Planta* 202: 407–413.

Young ND, Steiner K, and de: Pamphilis CW (1999) The evolution of parasitism in Scrophulariaceae/Orobanchaceae: Plastid gene sequences refute an evolutionary transition series. *Ann. Missouri Bot. Gard.* 86: 876–893.

Herbaceous Vegetation, Species Richness in

JP Grime, University of Sheffield, Sheffield, UK
Jason D Fridley, Syracuse University, Syracuse, NY, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by J. P. Grime, volume 3, pp. 329–337, © 2001, Elsevier Inc.

Glossary

Dominant species Species that have the greatest influence on ecosystem structure and function by virtue of their abundance, biomass, or coverage.

Mass ratio hypothesis Hypothesis stating that ecosystem processes are largely determined by the dominant contributors of the overall plant biomass, that is, dominant species will exert greater influence on processes than will subordinate species.

Species richness Number of species present in a given habitat or ecosystem.

Subordinate species Species that have a minor influence on ecosystem structure and function, presumably because of their lesser abundance and biomass compared to dominant species.

Transient species Species that are present as scattered seedlings or small immature individuals; many of these species occur as dominant or subordinate species in neighboring vegetation associated with different environmental conditions or management regimes.

Introduction

Ecological research into losses in plant diversity has a long history, but in recent decades the agenda has shifted from a primary concern with its mechanisms to an assessment of the extent to which losses impair the functioning of ecosystems and reduce their utility to humankind. It is relatively easy to manipulate the species composition and diversity of herbaceous communities in both natural and model ecosystems, and experimenters have applied themselves with vigor to this new field of inquiry. The objectives in this research are twofold. First, one must identify where and how losses in diversity affect ecosystem functions such as primary productivity, carbon storage, mineral nutrient cycling, and ecosystem resistance and resilience. Second, information is required on the plant constituents necessary for successful ecosystem restorations; it must be established which plant species are irreplaceable and what population sizes and pools of genetic variation are required for their persistence.

As participants in both the current and the preceding phases of research on plant diversity, the authors emphasize the value of connecting theory and data from the earlier phase of research into the causes of plant diversity to more recent efforts that examine its consequences for ecosystems. The main objective of this article is to identify these connections. An additional purpose is to highlight opposing perspectives of whether the traits of dominant species or species richness itself have a primary role in ecosystem functioning.

Controls of Species Richness in Herbaceous Vegetation

Species Richness Peaks at Intermediate Levels of Disturbance and Productivity

The humped-back model (HBM) was proposed (Grime, 1973) as a compact scheme to describe the controlling effects of five factors (productivity, disturbance, dominance, niche-differentiation, and the local species pool) on the species richness of a plant community. Along a gradient of standing biomass, plant species richness increases as rates of disturbance or environmental stress decrease, to a maximum determined by the size of the local species pool and the degree of local environmental heterogeneity. Continuing along the gradient, ecosystems of high productivity and low disturbance rates favor species capable of quickly usurping available resources, which leads to rapid declines in richness from competitive exclusion. Four decades of testing have shown that the HBM is compatible with a large volume of plant community data, particularly at fine spatial grains (Grace, 1999). Unresolved aspects of HBM include the extent to which niche-differentiation is important in species-rich communities (Silvertown, 2004) and the general contribution of dispersal limitation in the context of regional species pools (Helm *et al.*, 2005).

Role of Intraspecific Diversity

More recent research on linkages between species diversity and intraspecific genetic diversity has suggested that the process of competitive exclusion may be slower in communities where populations exhibit high local genetic differentiation (Vellend and Geber, 2005). For example, in the species-rich calcareous grasslands of northern England inhabited by outcrossing perennials that exhibit high local genetic and morphological variation, genetically diverse populations maintained higher species evenness in an experimental setting (Booth and Grime, 2003; Fridley and Grime, 2010), which was due in part to the genotypic dependence of species-level competitive relationships (**Figure 1**; Fridley *et al.*, 2007). The role of intraspecific genetic diversity in maintaining high species richness seems to be greatest when disturbance (e.g., grazing or mowing) reduces the importance of apical dominance as a 'best strategy' for competitive superiority (Herben *et al.*, 2001), but more

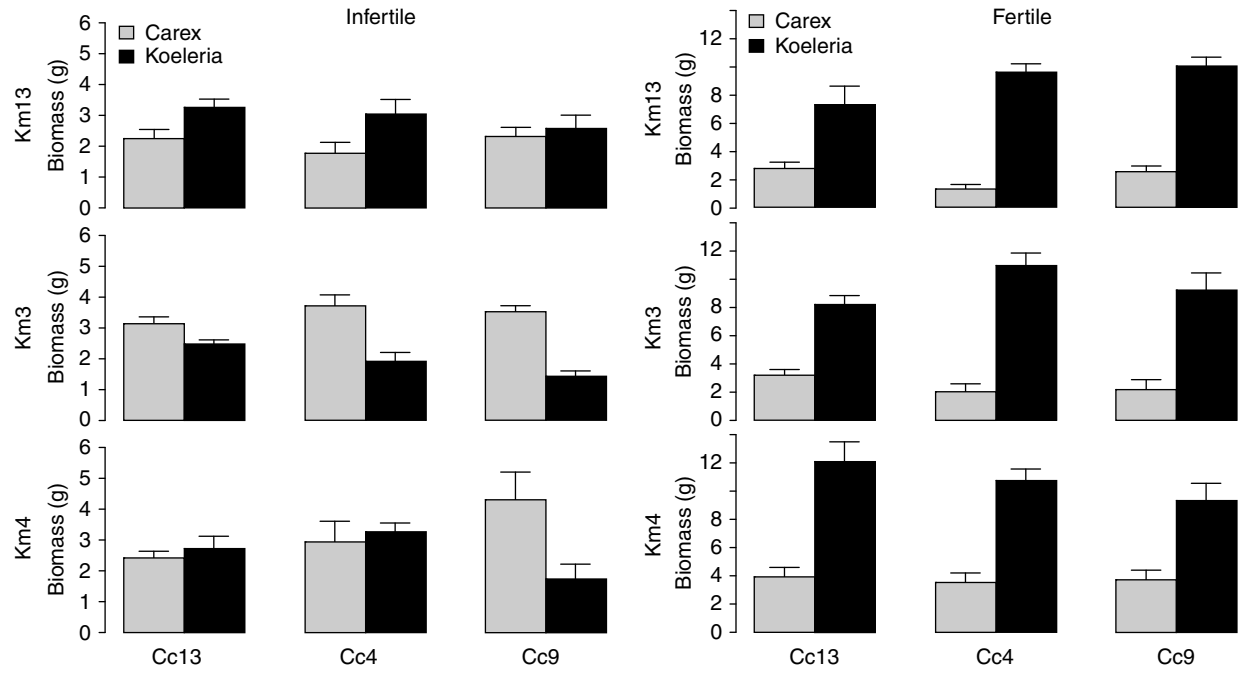


Figure 1 Left panel: The competitive dominant in pairwise competition between a sedge (*Carex caryophyllea*, gray bars) and a grass (*Koeleria macrantha*, black bars) in a pot experiment depended on the genetic identity of each species (rows are three genotypes of *Koeleria*, columns are three genotypes of *Carex*, all propagated from mature individuals growing together in the field) 2007. This suggests competitive exclusion in genetically diverse populations is slow particularly in cases of disturbance (e.g., grazing) and low resource availability. Right panel: the same competitive relationships under more fertile, less disturbed conditions. Here there was no influence of genotypic identity on the outcome of competition. Reproduced from Fridley JD, Grime JP, and Bilton M (2007) Genetic identity of interspecific neighbors mediates plant responses to competition and environmental variation in a species-rich grassland. *Journal of Ecology* 95: 908–915, with permission from Wiley.

studies are needed before the overall significance of intraspecific diversity to species richness is understood.

Plant Traits, Species Richness, and Ecosystem Functioning

Effects of Plant Traits on Ecosystem Properties

Although the majority of investigations in plant ecophysiology have sought to explain how vegetation is determined by environment, there has also been widespread recognition of feedbacks in which plants affect their environments and influence ecosystem properties. This idea is implicit in the writings of the founders of ecosystem theory (Odum, 1963, 1969) and more recently has found expression in the search for plant functional types (Grime, 1979; Chapin, 1980; Smith *et al.*, 1996; Diaz *et al.*, 2004), in which specific plant community and ecosystem properties have been attributed directly to particular plant traits such as potential growth rate, palatability to herbivores, resistance to fire, litter quality, and seed persistence. This approach has been extended further by studies in which the predictive value of plant traits has been tested by examining ecosystem responses to perturbation. An early example is the investigation by Leps *et al.* (1982) in which differences between two neighboring grasslands in their resistance and resilience following the severe drought of 1975 in the Czech Republic were predicted from the life histories and potential growth rates of the main component plant species. More comprehensive tests were conducted by MacGillivray and Grime (1995), in which responses to frost, drought, and fire in five contrasted grasslands were successfully predicted from a set of plant traits measured in a laboratory screening program (Grime *et al.*, 1997).

The Mass Ratio Hypothesis

It is important to note that in tests such as those conducted by MacGillivray and Grime (1995), predictions are weighted according to the abundance of each plant species in the vegetation; this is founded on the assumption that the extent to which the traits of a species affect ecosystem properties is likely to be strongly related to the contributions of the species to processes such as photosynthesis, mineral nutrient capture, transpiration, and provision of substrates exploited by herbivores and decomposers. There is a clear implication here that ecosystem processes are determined to a large extent by the characteristics of the dominant contributors to the plant biomass. As a corollary of this "mass ratio hypothesis" (Grime, 1998), one would not expect minor contributors to the vegetation to exert strong effects on ecosystem properties.

Is it reasonable to conclude that ecosystem functions can be sustained by inputs originating from only a few dominant plant species? Some of the most extensive ecosystems contain very few plant species; this is particularly obvious in grasslands and heathlands on acidic soils, and here the conclusion is

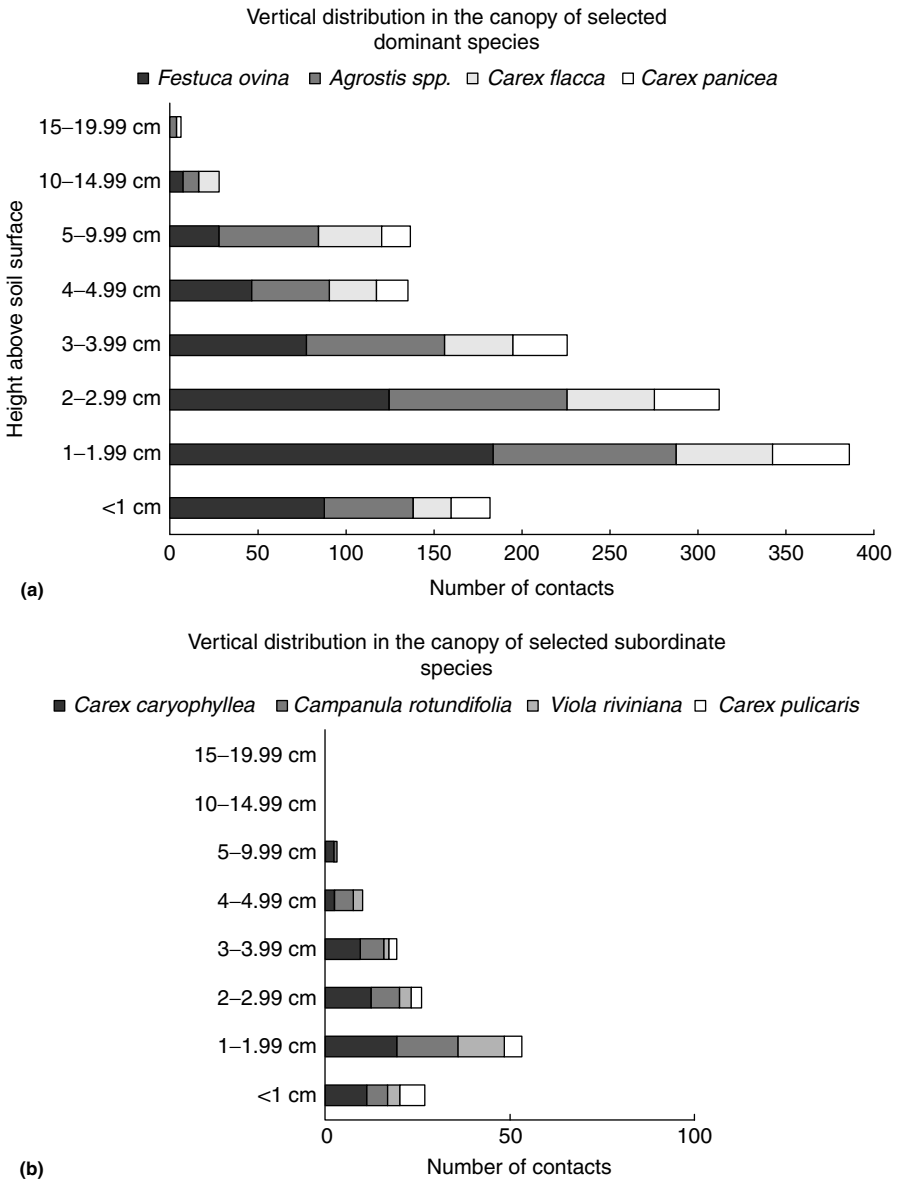


Figure 2 Vertical distribution in the leaf canopy of four dominant and four subordinate component species in an ancient limestone pasture at Buxton, North Derbyshire, England. Canopy distribution in June was estimated by measuring contacts with 375 randomly distributed, vertical pins. Reproduced from Hillier SH, unpublished data.

inescapable that controls on ecosystem processes by vascular plants are mediated through very few species. In herbaceous vegetation on soils of higher base status, analysis is complicated by the presence of more species, and in old calcareous grasslands the densities of herbs can rise to 30–40 species per m^2. Even here, however, it is interesting to note that when quantitative studies are made by harvest methods (Al-Mufti *et al.*, 1977) or point analysis (Mitchley and Grubb, 1986; see also **Figure 2**), it is evident that in comparison with the canopy dominants, many of the subordinate species of such communities account for a very small proportion of the biomass. There must be considerable doubt as to whether such minor components can, even collectively, exercise immediate effects on properties such as productivity, carbon storage, and water relations.

Effects Associated with High Species Richness

Against a theoretical background in which the traits of dominant plants were widely suspected to be acting as the overriding controllers of ecosystem properties, considerable interest and controversy were generated in the 1990s when studies began to demonstrate the immediate benefits to ecosystem properties arising from high species richness in experimental plant assemblages (Naeem *et al.*, 1994; Tilman and Downing, 1994; Hector *et al.*, 1999; Tilman *et al.*, 2001). In each case it was suggested that benefits arose in the species-rich mixtures from the presence of a wider range of morphologies and physiologies, generating either complementary and more complete exploitation of resources (Naeem *et al.*, 1994; Hector *et al.*, 1999; Tilman *et al.*, 2001) or

conferring resistance and resilience in the face of an extreme event (Tilman and Downing, 1994). Interest in these publications extending beyond the realm of ecology was stimulated by commentaries suggesting that studies of this kind provided a justification for the conservation of species-rich ecosystems (Kareiva, 1994, 1996; Naeem *et al.*, 1999).

An alternative perspective of these studies was drawn by others (e.g., Aarssen, 1997; Garnier *et al.*, 1997; Huston, 1997; Grime, 1998; Hodgson *et al.*, 1998), who argued that ecosystem properties attributed to high species richness were in reality due to the presence in the more diverse communities of dominant species with traits attuned to high productivity (Naeem *et al.*, 1994) or drought response (Tilman and Downing, 1994). More recent syntheses of a growing body of experimental studies examining richness-ecosystem functioning relationships have suggested that species complementarity can be important in some systems, but monocultures often outperform species mixtures (Cardinale *et al.*, 2006). Although the idea that plant species richness is a significant control over ecosystem processes like productivity continues to generate debate (Thompson, 2010), the overriding significance of the traits of plant dominants to ecosystem functioning is apparent from the large growth in trait-functioning studies (Lavorel and Garnier, 2002; De Deyn *et al.*, 2008; Suding *et al.*, 2008).

Components of Species Richness: Dominants, Subordinates, and Transients

Dominants and Subordinates

Early efforts to describe and interpret herbaceous vegetation (e.g., Clements, 1905; Tansley, 1939; Ramenskii, 1938) involved listing of the plants present in selected stands of vegetation and estimating the abundance of each species. Even before the widespread adoption of experimental approaches, there was a keen awareness of the potential of certain species to occupy a high proportion of the plant biomass and to control the abundance and fitness of other minor contributors. A classic example is the rapid expansion of tall coarse grasses and the coincident suppression of small herbs noted by Tansley and Adamson (1925) in their study of the consequences of excluding rabbits from chalk grassland.

A more quantitative approach followed as plant ecologists adopted stricter sampling methods and measured the relative abundance of plant species by harvesting, sorting, and weighing vegetation samples. This allowed more formalized attempts to examine the functional relationships between dominant and subordinate members of plant communities (Whittaker, 1965; McNaughton, 1978). Subsequently, it was recognized that the potential to dominate vegetation could be associated with specific plant traits such as canopy height, lateral spread, and the capacity to project shoots forcefully through litter and herbaceous cover (Al-Mufti *et al.*, 1977; Sydes *et al.*, 1981a, b; Grubb *et al.*, 1982; Campbell *et al.*, 1992). A further step in defining the functional differences between potentially dominant and subordinate members of plant communities became possible through the application of foraging theory to plants. From a study of the development of the roots and shoots of isolated plants exposed to standardized patchy environments (Campbell *et al.*, 1991), it became possible to predict the dominance hierarchy that developed when eight herbaceous species were grown together in an experimental assemblage (**Figure 3**). In this investigation it was found that dominance was associated with relatively imprecise foraging both above and below ground. Such coarse-grained foraging allowed the exploitation of a large volume of habitat but was associated with an imprecise concentration in

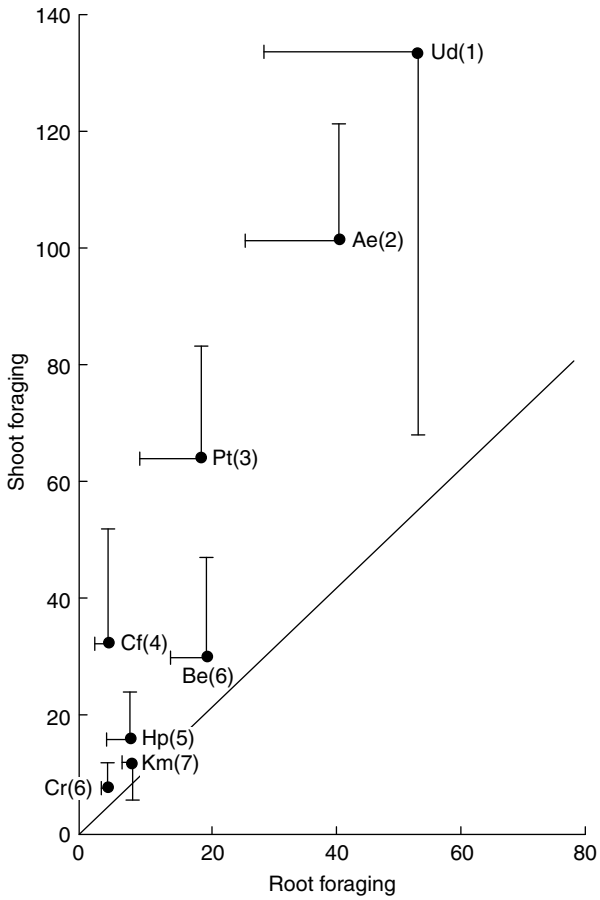


Figure 3 An examination of the relationship between root and shoot responses to resource heterogeneity in eight herbaceous species of contrasted ecology. A description of the methods used to expose the plants to resource patchiness is provided in Campbell, *et al.* (1991). Scales of foraging by the roots and shoots in the foraging assays are expressed at the respective increments of biomass (mg) to two undepleted quadrants, which in both assays constitute 50% of the available volume. The vertical and horizontal bars indicate 95% confidence limits. The numbers in parentheses refer to the species ranking in a conventional competition experiment in which all eight species were grown together in an equiproportional mixture on fertile soil for 16 weeks. Note the tendency for the scale of leaf canopy adjustment to exceed that of the root system. This of course arises from the freedom of movement of leaves in air and the encasement of roots in soil. Key to species: Ae, *Arrhenatherum elatius*; Be, *Bromus erectus*; Cf, *Cerastium fontanum*; Cr, *Campanula rotundifolia*; Hp, *Hypericum perforatum*; Km, *K. macrantha*; Pt, *P. trivialis*; Ud, *Urtica dioica*. Reproduced from Fig. 14 of Grime (2001) *Plant Strategies, Vegetation Processes, and Ecosystem Properties*, published by Wiley.

resource-rich sectors. In marked contrast, subordinate species were characterized by precise but local placements of roots and shoots in resource-rich patches, a foraging pattern permitting temporary coexistence with dominants but apparently committing them to a minor status in the community and leaving them vulnerable to extinction in circumstances where the dominants continued to enjoy unrestricted growth.

These findings raise important questions concerning the ecology and evolution of the very large numbers of plant species that consistently occupy subordinate positions within the hierarchies of herbaceous communities. Are these plants merely "also-rans" in the struggle for existence or are there fitness benefits associated with playing a minor part in plant communities? What is the precise role of subordinates in the long-term functioning and dynamics of ecosystems? These are not trivial questions in view of the fact that it is the subordinates rather than dominants that account numerically for the existence of high species richness in plant communities. Detailed attempts have been made to address these issues (Grime, 1987, 1998); here argument will be reduced to the following propositions:

1. Many subordinates achieve widespread occurrence because they are capable of exploiting similar niches associated with different dominant species. For example, *Poa trivialis*, the commonest vascular plant species of Lowland Britain, occupies shaded microhabitats beneath tall grasses, broad-leaved herbs, and a variety of shrubs and trees.
2. In many species-rich communities, the risks to subordinates of competitive exclusion remain low because potential dominants are restricted in vigor by environmental and biotic factors (e.g., mineral nutrient stress, drought, grazing, fire).
3. Because they do not experience the costs in support materials and construction time incurred during the building of robust root and shoot architectures, subordinates enjoy a temporary advantage when communities are subject to episodes of biomass removal. Where damage is catastrophic, as in tree felling and coppicing or the ploughing and burning of grasslands (Skutch, 1929; Marks, 1974; Platt, 1975; Bormann and Likens, 1979; Pickett and White, 1985; Pons, 1989), this can lead to a temporary but massive expansion of subordinates until such time as the dominants recover.

It can be concluded therefore, that many subordinate members of plant communities achieve high levels of fitness, particularly where natural or human-inspired interventions restrict the vigor of potential dominants either continuously or intermittently. Later in this article, it will be suggested that, in the long term, this rather opportunistic ecology displayed by subordinates may be sufficient to allow them an indirect but crucial involvement in the determination of ecosystem properties.

Transients

Many published records of the species composition of vegetation are incomplete. This can arise from cursory sampling and recording or from a deliberate policy of discounting minor constituents. Where the objective is to recognize recurring plant communities, data analysis often involves procedures that exclude species of low frequency or inconsistent occurrence in vegetation samples. The result of these various decisions can be a divergence between published data and field reality, and it has been suggested (Grime, 1998) that important information relevant to the long-term dynamics of ecosystems may be lost. To examine this hypothesis it is necessary to refer to surveys in which a strenuous attempt has been made to include all the plant species present in the vegetation.

During the period 1965–1975, a series of vegetation surveys was conducted to produce an inventory of the herbaceous plant communities of the Sheffield region in north-central England (see Grime *et al.*, 1988, for details). In each of approximately 10,000 m^2 quadrats, a complete list of species was made and the frequency of each species was recorded. When the resulting data were examined, it became apparent that, in addition to dominants and subordinates, the majority of the samples contained species that were represented only as scattered seedlings and small immature individuals. Most of these species were present as dominants or subordinates in neighboring vegetation associated with different environmental conditions or management regimes. Examples of this phenomenon are provided in **Table 1** for three contrasted sets of vegetation samples; in each case it is evident that the plant communities examined were harboring juveniles of species that occur as persistent, reproductive populations elsewhere in the landscape in vegetation of a different type.

Therefore, it seems reasonable to conclude that in addition to the dominant and subordinate members of a community, a third contributor to species richness can be identified in the form of transients originating either from the seed rain from

Table 1 Ecological classification of the species recorded in three distinctive and highly contrasted habitats sampled widely in an area of 2400 km^2 around Sheffield, UK[a]

Sampled habitat	*Number of m^2 samples*	*Woodland species*	*Grassland species*	*Arable species*	*Others*[b]	*Total*
Woodland on limestone	51	65	23	1	5	94
Meadows	40	7	64	8	0	79
Cereal arable	55	5	38	69	2	114

[a]All the species encountered within a particular habitat were classified in terms of their primary habitat (columns 3–6). Details of the sampling, recording, and habitat classification procedures are provided in Grime JP, Hodgson JG, and Hart R (1988) *Comparative Plant Ecology: A Functional Approach to Common British Species*. London: Uhwin Hyman.
[b]Includes species primarily associated with wetland or skeletal habitats (cliffs, walls, and rock outcrops).

the surrounding landscape or from seed banks occurring as a legacy of previous vegetation types that occupied the site.

Declining Species Richness and Ecosystem Reassembly

If immediate control of the functional properties of an ecosystem rests with dominant plants and if species-richness depends on the numbers of subordinates and transients, it is pertinent to ask "Does a decline in species richness matter?" To address this question, it is necessary to consider the significance of losses of subordinates and transients in the long-term dynamics of plant communities and ecosystems.

Filter Effects of Subordinates?

One mechanism whereby losses in subordinates could affect ecosystems is through alteration of the filter controlling the recruitment, identity, and relative abundance of dominants. To review the opportunities for subordinates to control the admission of dominants into communities, it is necessary to consider the long-term dynamics of vegetation and the regenerative phases in the life cycles of dominants. Studies of vegetation succession conducted earlier this century (e.g., Watt, 1925, 1947) established that continued dominance by particular species is frequently determined by the success of seedling or vegetative reestablishment following disturbance events. As already described here, often the early course of events following a disturbance is a temporary expansion in the cover and vigor of subordinates. This is most obvious in forest clearings where a dense, low cover of shrubs, herbs, and bryophytes characterizes the environment of regenerating trees (Watt, 1925; Skutch, 1929; Marks, 1974; Bormann and Likens, 1979), but similar phenomena have been described for grasslands and heathlands (Oosting, 1942; Keever, 1950; Hillier, 1990). Establishment following disturbances involves complex interactions of seedlings and vegetative shoots with substratum conditions, and contributions to the ground cover by subordinate plants may be expected to have both positive and negative effects (Pickett and White, 1985). Benefits to establishment have been described in circumstances where seedlings survive in the shelter afforded by low-growing shrubs, herbs, and bryophytes. Detrimental effects of shrub, herbaceous, and bryophyte cover on the establishment of grassland and forest dominants have been observed (Niering and Goodwin, 1962; Webb *et al.*, 1972; Pons, 1989), and it is widely recognized that many small-seeded herbs, trees, and shrubs are incapable of establishment in a closed cover of vegetation.

It can be deduced therefore, that there is a potential for subordinate members of a plant community to act as a filter in selecting between different potential dominants during the early phases of recolonization following a disturbance event. Selection could operate on the basis of variation in the seed reserves of dominants and on the capacity of their seedlings to penetrate a low canopy. The filter might also discriminate between dominants that rely on rapid emergence and those that regenerate by persistent juveniles. Subordinates could also control regenerating dominants through more indirect mechanisms, such as provision of sites in which seed predation is reduced, or through more complex phenomena such as the maintenance of critical pests, pathogens, herbivores, or mutualists.

Founder Effects of Transients?

If, as suggested earlier in this article, the sources of the transients are seed banks in the soil and the seed rain from the surrounding landscape, it would appear that they are an index of the pool of potential colonizing species at each site. On this basis, a diversity of transients signifies a high probability that, in the event of habitat disturbance or changes in management, there will be a rapid ingress of different plant functional types, some of which may be capable of exploiting the new conditions. An obvious example is the benefit to woodland development where an abandoned grassland already contains a diverse assortment of tree seedlings.

Current losses in biodiversity in Europe and in many other parts of the world are taking place in a complex landscape mosaic that is continuously disturbed by natural events and by urbanization, arable cultivation, forestry, and various forms of grassland management. Ecosystem sustainability depends, in part, on the continuous movement of populations and the reassembly of vegetation types and ecosystems. The extent to which communities and ecosystems are rapidly reconstituted is likely to be related to the reservoir of colonizers, many of which should be detectable prior to disturbance as transient constituents of the existing vegetation. As Egler (1954) recognized, the authors may suspect that the speed and completeness with which ecosystem reassembly occurs will depend on early colonization by appropriate dominants and subordinates; late arrivals will be delayed in their establishment and some may be excluded completely (Keever, 1950; Niering and Goodwin, 1962; Platt, 1975). It is not difficult to envisage how circumstances could then arise whereby efficiently dispersed plant species with "poor fit" to habitat and management conditions could assume dominance with damaging consequences for ecosystem function. There is an urgent need to discover the extent to which failure in the processes of plant dispersal and ecosystem reassembly can be predicted from the decline through time in the density and species richness of transients in plant communities.

Conclusions

The species richness of herbaceous vegetation is controlled by ecosystem processes related to rates of productivity and disturbance; the balance of evidence suggests that richness itself exerts only a minor control on ecosystem processes. However, this does not mean that losses of plant diversity should be ignored. The priority in the next phase of research on declining plant diversity should be to consider its long-term consequences on ecosystem structure and function. Losses in species richness may be associated with less obvious impacts that operate through failures in filter and founder effects. A progressive loss of ecosystem functions may be predicted in

circumstances where vegetation patch dynamics and ecosystem reassembly continue against the background of a declining pool of colonizing propagules. The effects on the recruitment of dominants, rather than the immediate consequences of declining richness *per se,* deserve our curiosity and attention.

See also: Ecosystem Function, Principles of. Herbicides. Plant Biodiversity, Overview

References

Aarssen LW (1997) High productivity in grassland ecosystems: Effected by species diversity or productive species? *Oikos* 80: 183–184.

Al-Mufti MM, Sydes CL, Furness SB, Grime JP, and Band SR (1977) A quantitative analysis of shoot phenology and dominance in herbaceous vegetation. *Journal of Ecology* 65: 759–791.

Booth RE and Grime JP (2003) Effects of genetic impoverishment on plant community diversity. *Journal of Ecology* 91: 721–730.

Bormann FH and Likens GE (1979) *Pattern and Process in a Forested Ecosystem.* Berlin: Springer-Verlag.

Campbell BD, Grime JP, and Mackey JML (1991) A tradeoff between scale and precision in resource foraging. *Oecologia* 87: 532–538.

Campbell BD, Grime JP, and Mackey JML (1992) Shoot thrust and its role in plant competition. *Journal of Ecology* 80: 633–641.

Cardinale BJ, Srivastava DS, Duffy JE, *et al.* (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443: 989–992.

Chapin FS (1980) The mineral nutrition of wild plants. *Annual Reviewes of Ecology and Systematics* 11: 233–260.

Clements FE (1905) Research Methods in Ecology, University Publishing Company. *Wisconsin*

De Deyn GB, Cornelissen JHC, and Bardgett RD (2008) Plant functional traits and soil carbon sequestration in contrasting biomes. *Ecology Letters* 11: 516–531.

Diaz S, Hodson JG, Thompson K, *et al.* (2004) The plant traits that drive ecosystems: Evidence from three continents. *Journal of Vegetable Science* 15: 295–304.

Egler FE (1954) Vegetation science concepts I. Initial floristic composition, a factor in old-field vegetation development. *Vegetatio* 4: 412–417.

Fridley JD and Grime JP (2010) Community- and ecosystem-level consequences of intraspecific genetic diversity in grassland microcosms of varying species diversity. *Ecology* 91: 2272–2283.

Fridley JD, Grime JP, and Bilton M (2007) Genetic identity of interspecific neighbors mediates plant responses to competition and environmental variation in a species-rich grassland. *Journal of Ecology* 95: 908–915.

Garnier E, Navas M, Austin MP, Lilley JM, and Gifford RM (1997) A problem for biodiversity–productivity studies: How to compare the productivity of multispecific plant mixtures to that of monocultures? *Acta Oecologica* 18: 657–670.

Grace JB (1999) The factors controlling species density in herbaceous plant communities: An assessment, perspectives in plant ecology. *Evolution and Systematics* 2: 1–28.

Grime JP (1973) Competitive exclusion in herbaceous vegetation. *Nature* 242: 344–347.

Grime JP (1979) *Plant Strategies and Vegetation Processes.* Chichester, UK: Wiley.

Grime JP (1987) Dominant and subordinate components of plant communities – Implications for succession, stability and diversity. In: Gray A, Edwards P, and Crawley M (eds.) *Colonisation, Succession and Stability*, pp. 413–428. Oxford, UK: Blackwell Scientific Publications.

Grime JP (1998) Benefits of plant diversity to ecosystems: Immediate, filter and founder effects. *Journal of Ecology* 902–910.

Grime JP, Hodgson JG, and Hart R (1988) *Comparative Plant Ecology: A Functional Approach to Common British Species.* London: Uhwin Hyman.

Grime JP, Thompson K, Hunt R, *et al.* (1997) Integrated screening validates primary axes of specialisation in plants. *Oikos* 79: 259–281.

Grubb PJ, Kelly D, and Mitchley J (1982) The control of relative abundance in communities of herbaceous plants. In: Newman E (ed.) *The Plant Community as a Working Mechanism*, pp. 77–97. Oxford, UK: Blackwell. British Ecological Society No. 1.

Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

Helm A, Hanski I, and Pärtel M (2005) Slow response of plant species richness to habitat loss and fragmentation. *Ecology Letters* 9: 72–77.

Herben T, Hara T, Hadincova V, *et al.* (2001) Neighborhood effects and genetic structure in a clonal grass: The role of the spatial structure of the environment. *Plant Species Biology* 16: 1–11.

Hillier SH (1990) Gaps, seed banks and plant species diversity in calcareous grasslands. In: Hillier SH, Walton DWH, and Wells DA (eds.) *Calcareous Grasslands: Ecology and Management, Bluntisham Books, Huntingdon*, pp. 57–66

Hodgson JG, Thompson K, Bogaard A, and Wilson P (1998) Does biodiversity determine ecosystem function? The Ecotron experiment reconsidered. *Functional Ecology* 843–848.

Huston MA (1997) Hidden treatments in ecological experiments: Evaluating the ecosystem function of biodiversity. *Oecologia* 110: 449–460.

Kareiva P (1994) Diversity begets productivity. *Nature* 368: 686–687.

Kareiva P (1996) Diversity and sustainability on the prairie. *Nature* 379: 673–674.

Keever C (1950) Causes of succession on old fields of the Piedmont. *North Carolina, Ecological Monographs* 20: 229–250.

Lavorel S and Garnier E (2002) Predicting changes in community composition and ecosystem functioning from plant traits: Revisiting the Holy Grail. *Functional Ecology* 16: 545–556.

Leps J, Osbornova-Kosinova J, and Rejmanek M (1982) Community stability, complexity and species life-history strategies. *Vegetation* 50: 53–63.

MacGillivray CW and Grime JP (1995) and the ISP team, Testing predictions of resistance and resilience of vegetation subjected to extreme events. *Functional Ecology* 9: 640–649.

Marks PL (1974) The role of pin cherry (*Prunus pensylvanica* L.) in the maintenance of stability in northern hardwood ecosystems. *Ecological Monographs* 44: 73–88.

McNaughton SJ (1978) Stability and diversity of ecological communities. *Nature* 274: 251–253.

Mitchley J and Grubb PJ (1986) Control of relative abundance of perennials in chalk grassland in southern England I. Constancy of rank order and results of pot- and field-experiments on the role of interference. *Journal of Ecology* 74: 1139–1166.

Naeem S, Thompson LJ, Lawler SP, Lawton JH, and Woodfin RM (1994) Declining biodiversity can alter the performance of ecosystems. *Nature* 368: 734–737.

Naeem S, Chapin FS, Costanza R, *et al.* (1999) Biodiversity and ecosystem functioning: Maintaining natural life support processes. *Issues in Ecology* 4.

Niering WA and Goodwin RH (1962) Ecological studies in the Connecticut Arboretum Natural Area I. Introduction and survey of vegetation types. *Ecology* 43: 41–54.

Odum EP (1963) *Ecology.* New York: Holt, Rinehart and Winston.

Odum EP (1969) The strategy of ecosystem development. *Science* 164: 262–270.

Oosting HJ (1942) An ecological analysis of the plant communities of Piedmont. *North Carolina, American Midland Naturalist* 28: 1–126.

Pickett STA and White PS (eds.) (1985) *The Ecology of Natural Disturbance and Patch Dynamics.* Orlando, FL: Academic Press.

Platt W (1975) The colonisation and formation of equilibrium plant species associations on badger disturbances in a tall-grass prairie. *Ecological Monographs* 45: 285–305.

Pons TL (1989) Dormancy, germination and mortality of seeds in heathland and inland sand dunes. *Acta Botanica Neerlandica* 38: 327–335.

Ramenskii LG (1938) *Introduction to the Geobotanical Study of Complex Vegetations.* Moscow: Selkzgiz.

Ratcliffe DA (1984) Post-medieval and recent changes in British vegetation: The culmination of human influence. *New Phytologist* 98: 73–100.

Silvertown J (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution* 19: 605–611.

Skutch AF (1929) Early stages of plant succession following forest fire. *Ecology* 10: 177–190.

Smith TM, Shugart HH, and Woodward FI (1996) *Plant Functional Types: Their Relevance to Ecosystem Properties and Global Change.* Cambridge, UK: Cambridge University Press.

Suding KN, Lavorel S, Chapin FS, *et al.* (2008) Scaling environmental change through the community-level: A trait-based response-and-effect framework for plants. *Global Change Biology* 14: 1125–1140.

Sydes C and Grime JP (1981a) Effects of tree leaf litter on herbaceous vegetation in deciduous woodland. I. Field investigations. *Journal of Ecology* 69: 237–248.

Sydes C and Grime JP (1981b) Effects of tree leaf litter on herbaceous vegetation in deciduous woodland. II. An experimental investigation. *Journal of Ecology* 69: 249–262.

Tansley AG (1939) *The British Islands and Their Vegetation*. Cambridge, UK: Cambridge University Press.

Tansley AG and Adamson RS (1925) Studies of the vegetation of the English Chalk. III. The chalk grasslands of the Hampshire–Sussex border. *Journal of Ecology* 13: 177–223.

Thompson K (1994) Predicting the fate of temperate species in response to human disturbance and global change. In: Boyle TJB and Boyle CEB (eds.) *Biodiversity, Temperate Ecosystems and Global Change*. Berlin: Springer-Verlag.

Thompson K (2010) *Do We Need Pandas? The Uncomfortable Truth About Biodiversity*. Devon, UK: Green Books.

Thompson K and Jones A (1999) Human population density and prediction of local plant extinction in Britain. *Conservation Biology* 15: 1–6.

Tilman D and Downing JA (1994) Biodiversity and stability in grasslands. *Nature* 367: 363–365.

Tilman D, Reich PB, Knops J, Wedin D, Mielke T, and Lehman C (2001) Diversity and productivity in a long-term grassland experiment. *Science* 294: 843–845.

Vellend M and Geber MA (2005) Connections between species diversity and genetic diversity. *Ecology Letters* 8: 767–781.

Watt AS (1925) On the ecology of the British beechwoods with special reference to their regeneration. Part II, Sections II and III. The development and structure of beech communities on the Sussex Downs. *Journal of Ecology* 13: 27–73.

Watt AS (1947) Pattern and process in the plant community. *Journal of Ecology* 35: 1–22. Full.

Webb LJ, Tracey JG, and Williams WT (1972) Regeneration and pattern in the subtropical rain forest. *Journal of Ecology* 60: 675–695.

Whittaker RH (1965) Dominance and diversity in land plant communities. *Science* 147: 250–260.

Herbicides

Jodie S Holt, University of California, Riverside, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Herbicide Chemical used to suppress or kill plants, or to severely interrupt their normal growth processes.

Herbicide resistance Inherited ability of a plant to survive and reproduce following exposure to a dose of herbicide normally lethal to the wild type. In a plant, resistance may be naturally occurring or induced by techniques such as genetic engineering or selection of variants produced by tissue culture or mutagenesis.

Herbicide tolerance Inherent ability of a species to survive and reproduce after herbicide treatment. This implies that there was no selection or genetic manipulation to make the plant tolerant; it is naturally tolerant.

Integrated weed management Approach for suppressing weeds that combines information on the biology and ecology of the weed with all available control technologies so that no one method is used exclusively.

Postemergence herbicide Herbicide applied after the emergence of the specified weed or crop.

Pre-emergence herbicide Herbicide applied to the soil prior to the emergence of the specified weed or crop.

Selectivity Phenomenon in which some plants are killed with doses of herbicides that have little or no effect on other plants.

Weed Plant that interferes with the growth of desirable plants and is unusually persistent and pernicious. Weeds negatively affect human activities and as a result are undesirable.

Weed control Reducing or suppressing weeds in a defined area to an economically acceptable level without necessarily eliminating them.

Herbicides are used in agricultural, aquatic, forest, and wildland ecosystems to reduce the density of unwanted vegetation (weeds) and to permit the growth of desirable species. Over time, use of herbicides in agroecosystems reduces weed density and may also select species that are adapted to the particular chemicals, thereby reducing weed species diversity. In wildland ecosystems, herbicides are sometimes used to reduce the density of exotic or invasive weed species and thus indirectly increase the diversity of native or desirable species. Where herbicides inadvertently enter the environment as contaminants, detrimental impacts on nontarget plants and other organisms may occur. However, the threat to plant biodiversity caused by habitat loss and the spread of exotic species is far greater than that caused by all forms of environmental pollution, including herbicides. This article reviews concepts of weeds; principles of weed management; and categories, action, and fate of herbicides. Impacts of herbicides on biodiversity of both target (weeds) and nontarget species and the role of weed control in preserving biodiversity are also discussed.

The Concept of a Weed

Weeds are most often defined in human or anthropomorphic terms, that is, as plants growing where they are not wanted or plants out of place. More useful definitions of weeds are those that describe biological traits or characteristics. A list of "ideal characteristics of weeds," developed by Baker (1974), is widely known and cited for its description of traits that confer weediness on plants that possess them (**Table 1**). These traits include germination under a broad range of conditions and over an extended period; rapid growth and prolific reproduction by sexual and asexual means; flexible breeding systems including selfpollination, cross-pollination, and nonspecialized pollinators; effective seed dispersal; plasticity and tolerance of a breadth of environmental conditions; and adaptations for competitiveness. Though no plant species could possess all of these traits, plants considered to be major weeds are likely to possess many of them. Traits alone do not determine whether a plant will be a weed, but Baker's list is a useful tool for categorizing and studying plants that interfere with human activities and, thus, are called weeds.

Impacts of Weeds

By their very definition, weeds affect many activities in which humans are engaged. In agroecosystems their primary effect is

Table 1 Characteristics of the "Ideal Weed"

Germination requirements fulfilled in many environments
Discontinuous germination (internally controlled) and great longevity of seed
Rapid growth through vegetative phase to flowering
Continuous seed production for as long as growing conditions permit
Selfcompatible but not completely autogamous or apomictic
When cross-pollinated, unspecialized visitors or wind utilized
Very high seed output in favorable environmental circumstances
Produces some seed in wide range of environmental conditions; tolerant and plastic
Has adaptations for short- and long-distance dispersal
If a perennial, has vigorous vegetative reproduction or regeneration from fragments
If a perennial, has brittleness, so not easily drawn from ground
Has ability to compete interspecifically by special means (rosette, choking growth, allelochemicals)

Source: Reproduced from Baker HG (1974) The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1–24, with permission from Annual Reviews.

to reduce crop yield and quality through competition for limited resources. Weeds also increase the time and costs required for crop production and interfere with harvesting. Indirectly, weeds affect crops through both positive and negative interactions with insect herbivores and their natural enemies. In rangelands, weeds possessing thorns or barbs pose physical hazards to livestock, and those containing toxins may cause allergies or poisonings of animals or humans. Other negative impacts of weeds include obstructing visibility around roadways, serving as a fire hazard, impeding use of recreation areas, and blocking the free flow of water in waterways, irrigation canals, and drainage ditches. In nonagricultural ecosystems, weeds often comprise the first stage of plant succession on land where the native vegetation has been disturbed. With increasing movement of humans across continental boundaries, exotic (non-native) weed species have invaded many wildland ecosystems where human activities have disrupted the growth of indigenous (native) species. As a result, biological diversity has been reduced in many wildland areas that interface with urban areas. Other impacts of weed invasions in wildlands include alteration of ecosystem processes, support of non-native animals, fungi, or microbes, and hybridization with native species to alter gene pools (Radosevich *et al.*, 2007).

Weed Science

Since the beginning of agriculture, humans have employed various tactics to remove weeds from land where other uses are desired. In the US today, billions of dollars are spent annually for weed removal using chemical and mechanical means. With the discovery of synthetic organic herbicides in 1941, weed science developed as a formal scientific discipline (Anderson, 2007). Weed scientists have been extremely successful during this century in developing techniques to remove weeds from agricultural and other ecosystems. Over the past 50 years, weed science has grown into a multidisciplinary field of study encompassing fundamental research along with applied aspects of weed suppression. Today it encompasses researchers from numerous scientific disciplines, including chemistry, ecology, genetics, morphology, and physiology, as well as applied scientists and practitioners who focus on weed suppression (Radosevich *et al.*, 2007). In recent years, increased awareness of environmental concerns has shifted the emphasis in weed science from a primary focus on herbicides to more integrated, ecological approaches for dealing with weeds (Liebman *et al.*, 2001).

Weed Management

In crop, forest, and rangeland production systems, as well as in wildlands, weed presence must usually be minimized to achieve a desired land use goal. Various tools and methods are used to suppress or remove weeds while not injuring the crop or desirable species. Weed management is a strategy that includes growing or fostering desired or beneficial vegetation while suppressing unwanted plants. For such management to be successful, knowledge of the biology, ecology, life history, and taxonomy of the weed species is required, as well as selection of the proper tools to use for their suppression (Clout and Williams, 2009; Radosevich *et al.*, 2007).

Tools and Methods of Weed Management

The strategy of weed management includes three key components: prevention, eradication, and control (Ross and Lembi, 2009). Prevention is keeping a weed from being introduced into an area where it does not already occur. Common preventive measures include using sanitary practices, eliminating weed spread through seeds and vegetative propagules, using quarantines, and following federal and state weed laws and regulations. Eradication is the total elimination, from a particular area, of a weed species and any plant parts capable of reproducing. Although eradication is often a stated goal of weed management programs, it is seldom achieved owing to the presence of seed banks and vegetative bud reserves in the soil of weed-infested areas. In contrast to eradication, weed control is the suppression or reduction of a weed species to an economically acceptable level. Under weed control programs, complete elimination of the weed is not the goal; instead, weeds are reduced to a level at which the cost of continued suppression does not exceed the value of the land or crop growing on it plus the benefit afforded by weed control. Once a weed is established in an area, weed control is the approach most commonly used to manage vegetation in that area over the long term. Methods used for weed control include biological, chemical, cultural, and mechanical techniques.

Integrated Approaches to Weed Management

With the growing recognition and concerns about the environmental impacts of agricultural practices, particularly the use of herbicides, integrated approaches for weed management have become commonplace. Integrated weed management (IWM) refers to a strategy for weed suppression that combines information on the biology and ecology of the weed with all available control technologies. Using this strategy, a variety of different methods are used in weed control, including nonchemical ones, as well as preventive measures, such that emphasis on herbicides is minimized. An alternative approach to crop production that is currently receiving widespread attention is sustainable agriculture, which refers to production systems in which external inputs, including synthetic fertilizers and pesticides, are minimized or avoided. However, current knowledge about the biology and ecology of weeds, their interactions with crops and wild plants, and nonchemical methods for their control is still limited. Thus, in most agricultural systems today, achieving weed control without herbicides requires extensive mechanical and hand labor, which is very costly. Until viable and economical alternatives to herbicides are available, most crop production and land management systems in the US will continue to depend on some level of herbicide use for weed control (Ross and Lembi, 2009).

Chemical Weed Control

Herbicides are chemicals used to suppress or kill plants or to interrupt their normal growth processes (Cobb and Reade, 2010). Of all groups of pesticides (including insecticides, fungicides, and rodenticides), herbicides are the leading group in terms of tons produced, dollar value from sales, and total acreage treated. Extensive and widespread use of herbicides in agriculture continues because of their high level of effectiveness and low cost relative to other methods of weed control. Use of herbicides has resulted in improved control of weeds that grow within crop rows where cultivation is not possible. Herbicides have replaced frequent tillage operations in some systems, which conserve energy, reduce crop damage, and minimize damage to soil structure. With herbicides, crop production is less dependent on weather and human labor, such that greater flexibility in choice of crops and management methods is possible. In the US today, the abundance of relatively inexpensive food and fiber is due in large part to the benefits afforded by herbicides for weed control in the last half century. However, the use of herbicides also carries risks, including injury to crops and nontarget plants, herbicide residues in soil or water, toxicity to nontarget organisms, and concerns for human health and safety. For this reason, the benefits and risks of each method must be weighed carefully when developing a weed control program, particularly in wildland ecosystems. In the US, all pesticide development and use is subject to strict regulation by the federal government.

Herbicides

There are approximately 150 herbicide active ingredients, which are formulated into hundreds of commercial products (Senseman, 2007). Most are organic compounds, containing carbon, hydrogen, oxygen, and various other chemical elements. Each herbicide has a chemical name that describes its structure and a common name, which is often a simplified version of the chemical name. Formulated herbicides also have a trade name assigned by the manufacturer for marketing purposes. Manufacturers formulate herbicides to enhance their handling and weed control properties. Formulated compounds include the herbicide active ingredient plus inert ingredients such as solvents, diluents, and various additives. When the same herbicide active ingredient is formulated in more than one way, each is assigned a different trade name. Herbicides can be classified in several different ways, which provide users with a convenient means of selecting herbicides for various purposes.

Herbicides are often classified according to similarities in chemical structure, which often, but not always, result in similar effects on plants. A more useful classification scheme is based on where they may be used. In agriculture, herbicides are registered for use in agronomic and horticultural crops, turfgrass, and landscape and ornamental plantings. In noncrop areas, herbicides are used in pastures and rangelands, aquatic habitats, rights-of-way, utility sites, recreation areas, forests, and wildlands. Herbicides are also classified according to the method and timing of application. Application methods include soil and foliage treatments, depending on where in the plant the chemical is most readily absorbed. Timings of herbicide application include preplant (prior to crop planting), preemergence (prior to crop or weed emergence), and postemergence (after crop or weed emergence). A herbicide classification scheme based on method of application is summarized in **Table 2**.

Other classification schemes are based on plant responses to herbicides. Selective herbicides are more toxic to some plant species than others (e.g., monocotyledonous vs. dicotyledonous plants), whereas nonselective herbicides are toxic to all plant species. Selectivity is one of the most important and useful characteristics of herbicides, for it allows applications to be made to weeds without the risk of injury to crops or desirable vegetation. Factors related to herbicide chemistry (structure, formulation), the plant (age, size, surface characteristics, morphology, physiology), and the environment (humidity, temperature, soil moisture) determine the selectivity of a particular herbicide. Other categories of plant response used to classify herbicides are the pathway of herbicide movement in plants and the mechanisms by which herbicides kill plants (see **Table 2**). Systemic herbicides are those that move in plants; movement occurs in the phloem (symplast), xylem (apoplast), or both. Herbicides that do not move in plants, but rather exert their effect at the site of application, are called contact herbicides. Herbicides classified by mechanism of action include growth regulators; inhibitors of photosynthesis, pigments, lipid synthesis, cell wall synthesis, amino acid synthesis, and cell division; and cell membrane destroyers.

Fate of Herbicides in the Environment

Herbicide fate in the environment is an issue of public concern and an important consideration when herbicides are registered for legal use. For herbicides to be effective, they must persist long enough to kill the weeds for which they were intended. Persistence beyond that time, however, may result in injury to nontarget plants and other organisms, residues in crops, and environmental contamination. Herbicides that enter plants generally move to a site of action and cause a toxic reaction. Over time, most herbicides in plants are transformed into relatively less toxic forms by biochemical processes. If a herbicide is not degraded, it may remain in the plant or end up in the soil as a contaminant. Eventually, however, all herbicides that enter a plant, the soil, water, or atmosphere will be degraded by the same chemical and physical reactions that act on biologically derived compounds (Cobb and Reade, 2010).

Once a herbicide reaches the soil, several processes, including adsorption to soil particles, movement to another location, and decomposition, will determine its persistence. Soil and herbicide characteristics regulate adsorption of herbicide molecules onto particles of clay and organic matter. Herbicides that are tightly adsorbed are not available for plant uptake, movement to other sites, or decomposition. Herbicides that are loosely adsorbed or located in the soil solution may move by leaching (vertical and lateral movement by water). The potential for movement of herbicides into groundwater through leaching is a concern for a few very mobile herbicides, which consequently are subjected to strict monitoring and regulation. Herbicides may also enter the environment during application by drift (movement of herbicide particles in air) or from the soil surface by

Table 2 Herbicide classification according to the method of application, movement in plants, and mode of action (examples of herbicide classes or individual herbicides in each category are listed)

Method of application	*Movement in plants*	*Mode of action*	*Chemical class or individual herbicide*
Foliar	Translocated in phloem	Auxin-type growth regulators	Phenoxy acid, benzoic acid, and picolinic acid herbicides; naptalam
		Aromatic amino acid (EPSPS) inhibitors	Glyphosate
		Branched-chain amino acid (ALA/AHAS) inhibitors	Sulfonylurea, imidazolinone, pyrimidinyl oxybenzoate, and triazolopyrimidine sulfonanilide herbicides
		Carotenoid pigment inhibitors	Amitrole, clomazone, fluridone, isoxaflutole, norflurazon
		Lipid biosynthesis (ACCase) inhibitors	Aryloxyphenoxy propionate and cyclohexanedione herbicides
		Organic arsenicals	DSMA, MSMA
		Unclassified herbicides	Asulam, difenzoquat, fosamine, propanil
Foliar	Translocated in xylem	Photosystem II (PSII) photosynthetic inhibitors	*s*-Triazine, phenylurea, and uracil herbicides; metribuzin
		Other photosynthetic inhibitors	Phenylcarbamate herbicides, bentazon, bromoxynil, pyrazon, pyridate,
Foliar	Contact (not translocated)	Photosystem I (PSI) cell membrane destroyers	Bipyridilium herbicides
		Protoporphyrinogen oxidase inhibitors	Diphenylether, oxadiazole, and triazolinone herbicides; flumiclorac
		Glutamine synthesis inhibitors	Glufosinate
Soil	Depends on herbicide	Microtubule/spindle apparatus inhibitors (root inhibitors)	Dinitroaniline herbicides; DCPA, dithiopyr, pronamide
		Shoot inhibitors	Chloroacetamide and thiocarbamate herbicides
		Miscellaneous cell division inhibitors	Bensulide, napropamide, siduron
		Cell wall formation inhibitors	Dichlobenil, isoxaben, quinclorac

Abbreviations: EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; ALS, acetolactate synthase; AHAS, acetohydroxy acid synthase; ACCase, acetyl-coenzyme A carboxylase; DSMA, disodium methanearsonate; MSMA, monosodium methanearsonate; DCPA, dimethyl 2,3,5,6-tetrachloro-1,4-benzenedicarboxylate.

Source: Adapted from Anderson WP (2007) *Weed Science: Principles and Applications*, 3rd edn. Long Grove, Illinois: Waveland Press, Inc.; Radosevich SR, Holt JS, and Ghersa CM (2007) *Ecology of Weeds and Invasive Plants: Relationship to Agriculture and Natural Resource Management*, 3rd edn. Hoboken, NJ: John Wiley & Sons, Inc., and Ross MA and Lembi CA (2009) *Applied Weed Science: Including the Ecology and Management of Invasive Plants*, 3rd edn. Upper Saddle River, NJ: Pearson Education.

volatilization (conversion into a vapor). Both of these processes can be minimized or avoided when proper equipment and application techniques are used.

Herbicide decomposition occurs in soil, air, water, plants, animals, and microorganisms and results in breakdown of the original herbicide molecule and loss of herbicide activity. Decomposition of herbicides occurs by photochemical (breakdown in sunlight), chemical, or microbiological means. Herbicide decomposition also depends on temperature and other environmental factors, plus the concentration of herbicide applied. Products of herbicide decomposition may eventually degrade into simple organic molecules. Because most herbicides are degradable, they do not build up in the soil over time, even after repeated use. Additionally, the soil microbial populations are adaptable such that the application of a pesticide to soil is often followed by an increase in the number of microbes that can degrade it. When processes that regulate herbicide fate are understood, and the legal requirements for herbicide application and use are followed, minimal contamination of the environment should result.

Regulation of Herbicide Use

All pesticides, including herbicides, must be registered with the U.S. Environmental Protection Agency (EPA) before they can be distributed or sold in the US (Whitford, 2002). Two laws, the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), and parts of the Food, Drug, and Cosmetic Act (FDCA) regulate pesticide development and use. FIFRA provides for registration and cancellation of pesticides, maintains a classification system for pesticides based on toxicity, and allows states to regulate pesticide use in a manner consistent with federal regulations. The FDCA mandates the establishment of tolerances for pesticides in food, feed, fiber, and water. These laws were written to ensure that benefits from the use of pesticides are in balance with the concerns about health and environmental impacts. Each state also has laws regulating pesticide use, including worker safety regulations and requirements for use of the most toxic (restricted use) pesticides.

For a pesticide to be registered by the EPA, it must be subjected to over 100 safety and environmental tests. Information required before registration includes chemical and physical properties, environmental fate, amounts of the pesticide in feed and food crops, toxicological properties, and effects on nontarget plants and animals. Such data are usually required both for the pesticide and for its metabolites, or breakdown products. With this information, the relative benefits and risks of each pesticide can be determined.

Pesticide effects on human health are expressed as toxicity, the amount of the chemical that is harmful or lethal, and exposure, the probability of encountering a harmful dose of

the chemical. The combination of acute toxicity plus exposure to a pesticide during its expected use determines the hazard it poses to humans. The EPA uses these data to set a tolerance level for each agricultural pesticide, which is the maximum amount of the chemical allowed on a particular crop. These and many other data, including chronic toxicity, reproductive effects, teratogenicity, and carcinogenicity, are used by the EPA in setting limits for pesticide use to ensure that hazard to humans from the use of pesticides will be at acceptably low levels. Toxicology and exposure studies are also required on certain species of wildlife, including birds, fish, and invertebrates, before pesticides can be registered. These tests utilize studies of pesticide residues in foods that these species may consume, as well as potential concentrations of pesticides in water or air to determine hazard to nontarget species from pesticide use.

Pesticides vary widely in their toxicological properties. In relation to all pesticides, which include insecticides, fungicides, and rodenticides, most herbicides are relatively nontoxic to mammals, since many of the processes or pathways they inhibit in plants are not present in mammalian systems. For the purposes of regulation, pesticides are classified according to their toxicity, which is a relative term used to describe the amount of a chemical that causes harm to a particular species. The most common unit of measurement for toxicity is the lethal dose, LD_{50}, or lethal concentration, LC_{50}, which is the dose or concentration that kills 50% of the test population, respectively. High values for LD_{50} or LC_{50} indicate lower toxicity, as higher doses are required to produce lethal effects. The categories of acute toxicity that must be shown on herbicide labels are listed in **Table 3**. Because all herbicides are toxic to some degree, container labels are required to give specific directions for-use as well as ingredients, properties, hazards, exposure limits, first aid procedures, and other information. When handled according to the label directions, hazard to humans and wildlife from exposure to herbicides can be avoided.

Effects of Herbicides on Biodiversity

Weeds

All weed control practices exert selection pressure on weeds and thus can have short- and long-term effects on the composition, structure, and dynamics of weed communities (Radosevich *et al.*, 2007). Agricultural weed communities tend to have lower species diversity than natural plant communities and are often dominated by a few key species. The primary short-term effect of weed control is a reduction in weed density, particularly of the dominant species, which is desirable in order to improve crop yields or facilitate land use. Over a longer time frame, weed control practices rarely eliminate weeds altogether; rather they generally result in changes in species composition and structure of weed communities. In the case of herbicides that act on specific plant processes, selection pressure over time can eliminate susceptible genotypes and thus cause evolutionary changes in weed populations.

Most research in weed control to date has emphasized reducing weed density and improving crop yields. Only recently the attention has been focused on changes in weed community dynamics as a result of weed control practices. Thus, only a few generalities can be made about specific effects or directions of change in weed diversity caused by biological, cultural, mechanical, and even chemical methods of control. The most information available on the role of weed control in shaping weed communities comes from documented cases of herbicide-resistant weeds, which have been selected by repeated use of the same herbicide or herbicide group (Powles and Holtum, 1994). It is clear that herbicides are a powerful evolutionary force acting on weed community composition. However, herbicide effects on weed diversity are generally transient and occur immediately following application. Over longer time periods, herbicide effects on weed diversity are either inconsistent or not apparent, and in many cases are overshadowed by more pronounced effects of other weed control and cropping practices.

Table 3 Toxicity, signal words, and lethal dosages (LD) used in herbicide labeling

Toxicity	*Signal word*	*Oral LD_{50} (mg/kg)*	*Dermal LD_{50} (mg/kg)*
Very high	Danger	≤ 50	≤ 200
High	Warning	51–500	201–2000
Moderate to low	Caution	> 500	> 2000

Effects of Weed Control Practices

Changes in weed community composition and structure due to agricultural practices have been documented; however, assessments of the effects of these practices on weed species diversity are relatively limited. Clements *et al.* (1994) calculated diversity indices from an array of published data to compare the impacts of conventional and alternative weed management practices on weed species diversity. When compared to weed control by mechanical cultivation only, broadcast applications of herbicides resulted in lower weed species diversity over time. When herbicide use was reduced by placement of applications only in bands over the crop row, higher diversity of weeds resulted than when broadcast herbicide applications were made. Although these findings suggest that use of herbicides reduces weed species diversity, case studies show that this generalization is too simplistic. Impacts on diversity depend on several factors, including the periodicity of tillage practices, weed species susceptibility to the herbicide product and rate (dosage), and variability in weather and its interaction with herbicide efficacy (Grundy *et al.*, 2010). When different herbicides were evaluated, in some cases applications of pre-emergence herbicides (those having residual soil activity) reduced weed species diversity more than the applications of postemergence herbicides (those with no residual activity). By exerting continuous selection pressure on susceptible species from the time of emergence, pre-emergence herbicides can reduce the richness and diversity of weed communities. In contrast, postemergence herbicides are present in the environment only after emergence and for a shorter time period, which may permit a more diverse weed flora to establish both before and after the disturbance of herbicide application is imposed.

Potts *et al.* (2010) monitored changes in the weed flora of cereal crops in England over 38 years and documented persistence and stability of weed abundance over time. Short-term decreases in weed abundance occurred following the use of herbicides and long-term changes in species composition were observed, including a steady increase in the abundance of perennial dicotyledonous (dicot) species. However, other agricultural practices were more important drivers of these long-term changes in species composition than were herbicides. Weed diversity increased over the monitoring period, possibly due to the buffering effect of the soil seed bank.

Even in cases where weed species diversity is relatively unaffected by herbicide applications, however, interspecific selectivity of many herbicides causes a shift within a weed community from species that are susceptible to species, that are naturally more tolerant, to the particular herbicide. A common example of this phenomenon is the shift in relative abundance from dicot weed species to monocotyledonous (monocot, usually grass) weed species following repeated use of the herbicide 2,4-D for control of dicot weeds in cereal (grass) crops. Similarly, in fields where herbicides have been used over many years for control of annual weed species, shifts in the weed flora to predominantly perennial weed species are commonly observed. Weed species compositional shifts also occur when the weeds in a field are taxonomically related to the crop species grown there, since plants often respond similarly to herbicides when they are in the same taxonomic family. Thus, many cases have been documented where repeated use of a particular herbicide in a crop has selected for weeds that are in the same plant family as the crop.

Integrated and alternative weed management methods, which employ a combination of tools to control weeds below a specified threshold, theoretically should not impose strong directional selection on weed populations. Thus, weed species diversity might be expected to increase, decrease, or remain stable, depending on the combination, timing, and frequency of control methods used. Similarly when control methods are rotated or changed over time, selection of well adapted or less susceptible weed biotypes is minimized. Many questions remain to be answered about the potential effects of integrated weed management techniques on diversity of weeds. In addition, the role and potential importance of weed biodiversity in agroecosystems remain poorly understood.

Selection of Herbicide Resistance

Herbicide resistance represents an extreme shift in weed species composition caused by the selection of plants possessing a gene or genes for resistance to a particular herbicide within a species that was formerly susceptible (Powles and Holtum 1994). In cases where a particular herbicide has been used repeatedly over several years, resistant weeds may be selected and come to dominate a weed community such that species diversity declines. Since the 1970s, many cases of the evolution of resistance have been documented as a result of repeated herbicide applications for weed control. **Table 4** summarizes the current worldwide occurrence of resistant weed biotypes to different herbicide groups.

Several characteristics of herbicides and their use contribute to a high probability for selection of resistance in weeds. These include having a single target site and specific mechanism of action, being extremely active and effective in killing a wide range of weed species, and having long soil residual activity and season-long control of germinating weeds. In addition, frequent application of a particular herbicide over several growing seasons without rotating, alternating, or combining with other types of herbicides contributes to a high risk for the evolution of resistance. Some herbicides are thought to pose a low risk for selection of resistance owing to their nonspecific mechanism of action and short or no soil residual activity. Even in these cases, however, repeated use of the same herbicide will exert selection pressure on weeds. Recommendations for preventing or managing herbicide-resistant weeds include practices such as rotating herbicides from different chemical groups to avoid imposing the same selection pressure over time and integrating a combination of weed control methods. The higher level of weed species diversity that presumably would result from these approaches should reduce the potential for propagation of herbicide-resistant genes in weed populations.

Since the discovery of evolved resistance to herbicides in the 1970s and the concurrent advancement of molecular technology, many crops have been genetically engineered for desirable traits, including resistance to herbicides. Transgenic herbicide-resistant crops, particularly cotton, maize, and soybean are now widely grown in the US (Owen, 2008). When this technology was first adopted, one concern was the potential for reliance on a single herbicide to which the crop was resistant and subsequent selection of weeds that are not

Table 4 Worldwide occurrence of resistant weed biotypes to different herbicide classes

Herbicide group	*Resistant weed biotypes(number)*
ALS/AHAS inhibitors	107
Photosystem II inhibitors	68
ACCase inhibitors	37
Synthetic auxins	28
Bipyridiliums	25
Ureas and amides	21
Glycines	21
Dinitroanilines and others	10
Thiocarbamates and others	8
PPO inhibitors	4
Triazoles, ureas, isoxazolidiones	4
Chloroacetamides and others	4
Nitriles and others	3
Carotenoid biosynthesis inhibitors	2
Arylaminopropionic acids	2
4-HPPD inhibitors	1
Mitosis inhibitors	1
Cellulose inhibitors	1
Unknown	1
Organoarsenicals	1
Total number of unique herbicide-resistant biotypes	350

ALS, acetolactate synthase; AHAS, acetohydroxy acid synthase; ACCase, acetyl-coenzyme A carboxylase; PPO, protoporphyrinogen oxidase; HPPD, 4-hydroxyphenylpyruvate-dioxygenase.

Source: Compiled by Heap I (2011) *International Survey of Herbicide-Resistant Weeds*. Available at www.weedscience.org. Data current as of 6 January 2011.

susceptible to that herbicide. Research conducted to date has shown that herbicide-resistance in crops has a benign effect on weed abundance and diversity. Instead, it is the weed control methods, including herbicides, as well as other management practices used in those crops that exert selection pressure for changes in weed abundance, composition, and dynamics, as discussed above.

Other Organisms

By suppressing, removing, or destroying vegetation, weed control modifies the environment and habitat of other organisms. In agroecosystems, crop pests as well as beneficial organisms can be affected by weed removal since weeds can serve as host plants or food sources for many types of organisms, including insects, fungi, and nematodes (Norris and Kogan, 2000). In some cases, a high diversity of weed species in an agricultural field has been shown to reduce the magnitude of insect attacks on crop plants because the weeds serve as alternate food for the insect pests or harbor beneficial organisms that feed on the pests. In other cases, however, increased weed species diversity results in increased insect problems in a crop field because the weeds provide a food source or habitat so that the insect pest can remain in the field even during periods when the crop is absent. In those cases, weed control to reduce the diversity of weeds in a particular field will decrease the incidence of insect pest damage to the crop, to the extent that herbicides affect weed biodiversity, therefore, their use will also indirectly affect insect organisms.

The response of fungal and nematode populations to weed species diversity has not been well studied. In general, weed control practices result in a cleaner crop field, which usually leads to fewer disease and nematode problems. However, it is also possible that large areas of crop monocultures with few weeds may be susceptible to widespread disease epidemics because of the lack of genetic diversity in response to the disease. To date, little information is available to indicate how weed biodiversity or use of herbicides in weed control influence populations of fungal pathogens or nematode pests of crop plants.

As noted above, most herbicides are nontoxic to mammals, which lack most of the biochemical pathways in plants that are targeted by herbicides. Aside from toxicology research required for registration, studies of herbicide effects on animals in field settings are limited. One exception is aquatic communities, which have been relatively well studied since taxon diversity is usually high and pesticides can move more readily than in terrestrial ecosystems. As expected, impacts on animals are pesticide specific. While some insecticides have direct negative impacts on aquatic animals, such as zooplankton, insects, amphibians and others, any negative herbicide effects are generally indirect through a reduction in food or shelter for organisms at higher trophic levels. One study showed direct negative effects on tadpoles of glyphosate herbicide when it was formulated with a surfactant additive; however, this additive is absent in the aquatic forms of the herbicide, which had no detrimental effects on tadpoles or other animals (Relyea, 2005). Thus, in both terrestrial and aquatic ecosystems, herbicides used properly appear to have only indirect effects on nonplant organisms when weeds they rely on for food, habitat, or other uses are controlled.

The question whether weed biodiversity is an asset or a detriment to overall pest management and other ecosystem services in agriculture deserves serious attention by researchers. This issue is particularly important in the UK and other countries where the majority of land is agricultural and wilderness areas are limited. In some regions weeds are increasingly being considered an important component of healthy agroecosystems and worthy of conservation (Marshall *et al.*, 2003). Evidence exists showing that agricultural intensification over the past century, including use of herbicides and inorganic fertilizers, has been accompanied by a reduction in populations of many insect groups and birds, presumably through reductions in weeds that they rely on (Storkey and Westbury, 2007). To understand the importance of weed biodiversity in agroecosystems, more information is needed on the effects of various weed control practices on both weed biodiversity and nonweed organisms. With this information, the costs and benefits of weed control can be weighed against potential costs and benefits of alternative strategies to other organisms that are indirectly affected by weed control. The question is complicated by the fact that in large-scale mechanized agricultural production, increases in weed species diversity complicate weed control efforts. However, this problem could be offset if there were benefits to be gained from maintaining increased genetic diversity of weeds in a field, such as greater buffering against selection for herbicide resistance, enhanced pest control, improved nutrient cycling, or other ecosystem services.

In wildland situations where exotic weeds have replaced native vegetation, weed control is increasingly practiced to reduce weed invasions and restore the abundance and diversity of native plants. The indirect result of these weed control activities is often a restoration of habitat for nonplant organisms, such as birds and small mammals. In these situations, therefore, weed control, including use of herbicides, can result in increased diversity of other organisms that depend on native plant communities for habitat.

Wild Plants

Weeds are managed in many situations to restore or preserve the biodiversity of native plant species. The most common methods used for weed management in wildlands in the US are manual and mechanical weed removal, prescribed fire, release of biological control agents, use of grazing animals, encouragement of native competitors, and judicious use of herbicides (Clout and Williams, 2009). Weed management strategies in wildlands differ from those in agroecosystems because wildland managers must promote or protect large numbers of plant and nonplant species rather than one or a few crop species. Thus, wildland weed management is generally approached with a desire to manage an entire plant community rather than to control a single weed species. Although the same methods can be used in agricultural and nonagricultural habitats, wildland managers must minimize negative impacts to a wider range of nontarget species than must agricultural land managers. Therefore, the methods used

for wildland weed control are often labor-intensive and more environmentally conservative than in agroecosystems.

All weed control methods incur some risk to the environment in which they are used, which must be weighed against the risk of taking no action and allowing weeds to continue to spread. Mechanical weed control disturbs the soil, destroys vegetation, and leaves gaps that may be reinfested with weeds. Biological control agents may attack nontarget species or become adapted and spread in undesirable ways. Herbicides may unintentionally kill nontarget plant species, indirectly impacting habitat for other organisms, and may become environmental contaminants if not used properly. In most cases, the herbicides used in wildland situations are postemergence compounds with very short or no soil residual activity to minimize their effect on nontarget plant species. However, few empirical data exist on either intentional or unintentional effects of herbicides on native plant communities.

A long-term field study on the effects of herbicides on the structure and species diversity of native plant communities showed that herbicide treatments were highly effective in controlling the exotic weed (*Centaurea maculosa*, spotted knapweed) and converting the plant community back to its native composition (Rice *et al.*, 1997). Overall, only small, short-term depressions in species richness and diversity resulted from herbicide use. In another field study comparing various methods of removing an invasive grass (*Microstegium vimineum*, Japanese stiltgrass), native community diversity was over 20% greater following either hand-weeding or postemergence herbicide use as compared to nontreated plots, indicating that the invasion was suppressing the native plant community (Flory and Clay, 2009). As shown in agroecosystems and as discussed earlier, however, the rate, frequency, and timing of herbicide applications are important determinants of plant community responses. Higher rates, more frequent applications, or applications early in the growing season have been shown to reduce the diversity of nontarget species more than lower rates, less frequent applications, or later applications. Similarly, selectivity of the herbicide as well as precision of the application method can be managed to reduce the impacts on nontarget species. These and other studies show that periodic application of appropriate herbicides can be used to restore native plant communities from dominance by exotic species and to maintain their diversity, as well. Transient reductions in diversity that may result from herbicide use are likely to be negligible when compared to the serious impacts on native communities created by exotic plant species.

As in agroecosystems, questions remain in natural systems about the role and management of weed abundance and diversity in wildland ecosystems. Where exotic weed invasions are widespread, weed removal alone may not result in the establishment of native species without additional inputs such as revegetation with desirable species. In some areas, invasive exotic weeds have replaced native species and become an important source of food and cover for native birds and mammals, which must be considered before the weeds are removed. The use of herbicides in wildlands is often controversial because of the risks of environmental contamination, as well as the general public perception that herbicides are dangerous. Just as in agriculture, therefore, the costs and benefits of weed control in wildlands must be weighed before any method is chosen. In cases where herbicides are used, the determination is made that the risks of using herbicides are more than the offset by the benefits gained in the increased or restored diversity of native species due to control of exotic weeds.

Conclusions

Herbicides have been employed for decades to control unwanted vegetation in agroecosystems, yet their impacts on weed biodiversity and the implications of those impacts for weed management have not been well studied except in the case of the evolution of herbicide resistance. Increasing emphasis over the past decade on the importance of maintaining diversity of all genetic resources has rarely been extended to agroecosystems, despite the likelihood that assessments of the nature and importance of weed biodiversity might reveal an important role of weeds in agroecosystem stability and sustainability. Based on experiences with herbicide use in agriculture, deliberate use of herbicides to control invasive weed species in wildlands has become increasingly common, yet the impacts of herbicides on the composition, structure, and dynamics of nontarget native plant communities are poorly understood and usually assumed to be negative. When placed in a comparative context, however, habitat degradation and destruction, particularly by competition with exotic species, is a much more pervasive threat to biodiversity of the endangered plant species than is pollution of all forms, including agricultural pesticides. Thus, it is imperative that research be continued into the effects of all forms of weed control on biodiversity of both target and nontarget species.

See also: Agriculture, Industrialized. Ecology of Agriculture. Herbaceous Vegetation, Species Richness in. Insecticide Resistance. Pesticides, Uses and Effects of

References

Anderson WP (2007) *Weed Science: Principles and Applications*, 3rd edn. Long Grove, Illinois: Waveland Press Inc.

Baker HG (1974) The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1–24.

Clements DR, Weise SF, and Swanton CJ (1994) Integrated weed management and weed species diversity. *Phytoprotection* 75: 1–18.

Clout MN and Williams PA (eds.) (2009) *Invasive Species Management: A Handbook of Principles and Techniques*. Oxford, Great Britain: Oxford University Press.

Cobb AH and Reade JPH (2010) *Herbicides and Plant Physiology*, 2nd edn. Chichester, West Sussex, UK: Wiley-Blackwell.

Flory SL and Clay K (2009) Invasive plant removal method determines native plant community responses. *Journal of Applied Ecology* 46: 434–442.

Grundy AC, Mead A, Bond W, Clark G, and Burston S (2010) The impact of herbicide management on long-term changes in the diversity and species composition of weed populations. *Weed Research*.http://dx.doi.org/10.1111/j.1365-3180.2010.00831.x

Liebman M, Mohler CL, and Staver CP (2001) *Ecological Management of Agricultural Weeds*. Cambridge, UK: Cambridge University Press.

Marshall EJP, Brown VK, Boatman ND, Lutman PJW, Squire GR, and Ward LK (2003) The role of weeds in supporting biological diversity within crop fields. *Weed Research* 43: 77–89.

Norris RF and Kogan M (2000) Interactions between weeds, arthropod pests, and their natural enemies in managed ecosystems. *Weed Science* 48: 94–158.

Owen MDK (2008) Weed species shifts in glyphosate-resistant crops. *Pest Management Science* 64: 377–387.

Potts GR, Ewald JA, and Aebischer NJ (2010) Long-term changes in the flora of the cereal ecosystem on the Sussex Downs, England, focusing on the years 1968–2005. *Journal of Applied Ecology* 47: 215–226.

Powles SB and Holtum JAM (1994) *Herbicide Resistance in Plants: Biology and Biochemistry*. Boca Raton, Florida: Lewis Publishers.

Radosevich SR, Holt JS, and Ghersa CM (2007) *Ecology of Weeds and Invasive Plants: Relationship to Agriculture and Natural Resource Management*, 3rd edn. Hoboken, NJ: John Wiley & Sons, Inc.

Relyea RA (2005) The impact of insecticides and herbicides on the biodiversity and productivity of aquatic communities. *Ecological Applications* 15: 618–627.

Rice PM, Toney JC, Bedunah DJ, and Carlson CE (1997) Plant community diversity and growth form responses to herbicide applications for control of *Centaurea maculosa*. *Journal of Applied Ecology* 34: 1397–1412.

Ross MA and Lembi CA (2009) *Applied Weed Science: Including the Ecology and Management of Invasive Plants*, 3rd edn. Upper Saddle River, NJ: Pearson Education.

Senseman SA (ed.) (2007) *Herbicide Handbook*, 9th edn. Champaign, Illinois: Weed Science Society of America.

Storkey J and Westbury DB (2007) Managing arable weeds for biodiversity. *Pest Management Science* 63: 517–523.

Whitford F (ed.) (2002) *The Complete Book of Pesticide Management: Science, Regulation, Stewardship, and Communication*. New York: John Wiley and Sons, Inc.

High-Elevation Andean Ecosystems

Mary TK Arroyo, Universidad de Chile, Santiago, Chile, and Instituto de Ecología y Biodiversidad, Santiago, Chile
Lohengrin A Cavieres, Universidad de Concepción, Concepción, Chile, and Instituto de Ecología y Biodiversidad, Santiago, Chile

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alpine life zone Life zone that occurs above the climatic high-elevation treeline or its bioclimatic equivalent.
Altiplano Extensive internally drained high-elevation plain found between the Western and Eastern Cordilleras of the Andes at 3800–4100 m above sea level.
Andes Mountain range comprised of various geological units that extends from latitude 12° N in Venezuela to 56° S off the southern tip of South America.
Endemism State of being unique to a defined geographic location, such as an island, nation or other defined zone, or habitat type.
Ma Millions of years.
Páramo High-elevation vegetation zone found from 11° N to 8° S in the South American Andes and on some mountains in Costa Rica and Panama.
Patagonia Region of South America in southern Argentina and Chile extending from the Río Colorado to the Straits of Magellan Fuego and from the Andes to the Atlantic Ocean and often considered to include northern parts of the Island of Tierra del Fuego.
Phylogeny The evolutionary history of a group of organisms or populations.
Southern Andean steppe High-elevation vegetation zone found from latitude 27° S to 56° S in the South American Andes, in Argentina, and Chile.

Introduction

The high-elevation habitats of the major mountain ranges of the world contain a sizable amount of plant biodiversity at local and global scales. Such habitats are intriguing in that most are of recent geological origin and are known to have garnered evolutionary lineages from many different biogeographic sources. Importantly, high-elevation areas provide critical ecosystem services such as the regulation of water flow and slope stability and, on account of their steep altitudinal gradients and opposite-facing slopes offer readily accessible thermal refuges for species affected by climate change. The high-elevation habitat, in addition to being of great scientific interest, thus is important for conservation as well as for the welfare of humanity.

In this article, the authors review plant biodiversity in the páramo, puna, and southern Andean steppe of the South American Andes. A wealth of other vegetation types (including tropical cloud and cool temperate forest, desert communities, scrubland, and grassland) subtend these three high elevation vegetation zones, but for reasons of space, these will not be considered here. The authors first situate Andean high elevation ecosystems in a global vegetation context. They then synthesize the uplift history and topological features of the Andes relevant to present-day biodiversity patterns. Having set the stage, the entire high-elevation flora, considering species richness, and endemism are overviewed, and recent findings on the origin of the flora as revealed by phylogenetic construction are highlighted. Finally, narrative accounts of the páramo, puna, and southern Andean steppe are provided.

High Elevation Habitats of the Andes

The South American Andes extend from latitude 12° N in Venezuela to 56° S in Chilean territory and as such constitute the worlds longest uninterrupted high-elevation mountain corridor (**Figure 1**). The Andes make their first appearance in the Lake Maracaibo region of Venezuela; their last terrestrial vestige is found on the Cape Horn Islands off the southern coast of Tierra del Fuego. Traditionally, there has been a tendency to divide the high-elevation habitat into páramo, puna, and southern Andean steppe (Arroyo *et al.*, 2010). The páramo extends from 11° N in Venezuela to around 8° S in northern Peru (Luteyn, 1999) and is found entirely within tropical latitudes. The puna covers tropical to midtemperate latitudes, occurs from 8° S to roughly 27° S (Luteyn and Churchill, 2000; Arroyo *et al.*, 2010; but see Martínez Carretero, 1995). Southern Andean steppe extends southward and lies entirely within temperate latitudes. Taken together, the páramo, puna, and southern Andean steppe are estimated to cover more than 913,600 km^2, the equivalent of approximately 5.1% of the land area of the South American continent (Arroyo *et al.*, 2010). Páramo covers around 63,600 km^2, puna 684,400 km^2, and southern Andean steppe 165.600 km^2.

In the northern tropical and southern temperate Andes south of 35–37° S, a consistent treeline defines the lower limit of the high-elevation ecosystem. However, over large intervening areas of the Andes a treeline is often lacking, and here the high-elevation ecosystem transitions gradually into shrublands, grasslands, or desert communities. While the páramo, puna, and southern Andean steppe vegetation zones enclose South America's equivalent of alpine vegetation in their upper reaches, it is important to realize that these vegetation types can extend well below the equivalent of the alpine life zone, defined as that vegetation above the upper tree limit. Under the Rivas-Martínez (1993) bioclimatic system for defining vegetation zones the puna as defined in this article, based on work in northern Andes of Chile (Luebert and Gajardo, 2000) and southern Peru (Galán de Mera *et al.*, 2003) is included mostly in the orotropical and cryotropical bioclimatic belts. The páramos of Venezuela traverse the upper part of the mesotropical belt into the orotropical belt

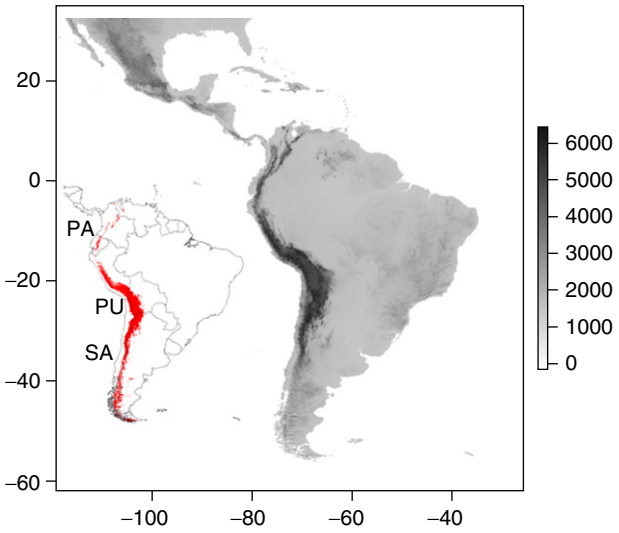


Figure 1 Digital elevation model of South America showing the South American Andes. The inset shows the extent to páramo (PA), puna (PU), and southern Andean steppe (SA).

(Méndez, 2007). Under the Rivas-Martínez (2003) bioclimatic scheme, generally, northern hemisphere subalpine areas fall into the orotemperate belt, while alpine areas fall into the cryotemperate belt. The degree of the mismatch with the alpine life zone can be illustrated by considering Körner *et al.*'s (2011) recently proposed bioclimatic rules, which consider mean growing season temperature (GST) and length of the growing season (GS) for defining the major vegetation belts on mountains. Estimated GST (GS, approximately 2 months) at treeline for the far southern Andes close to Ushuaia (55° S) is 5.1 °C, and fits fairly well with Körner *et al.*'s (2011) bioclimatic definition of the lower alpine belt (GST < 6.4 °C to > 3.5 °C; GS of < 94 days to 54 days). However, in the drier central Chilean Andes (33° S), where the lower limit of southern Andes steppe lies immediately above an open *Kageneckia angustifolia* (Rosaceae) tree line at 2200 m (Cavieres *et al.*, 2000a), estimated GST (GS, approximately 4–5 months) calculated from Cavieres and Arroyo (1999) is still as high as 9.9 °C at 2600 m. Estimated GTS for some puna localities in Argentina and Chile is minimally 6.7–9.7 °C (Cabrera, 1968; Ruthsatz, 1974; Arroyo *et al.*, 1988), while for some páramo localities it oscillates around 8–10 °C (Cavieres *et al.*, 2000b; Rangel, 2000a). Based on this brief analysis, with the exception of those parts of the southern Andes where there is a good treeline, páramo, puna, and southern Andean steppe tend to overlap with Körner *et al.*'s (2011) montane belt (GS temperature of < 10 and > 6.4 °C).

The Physical Setting

To understand biodiversity patterns in the high Andes, a clear picture of the time of appearance of the high-elevation belt and of landscape evolution at different latitudes along the Andes over which clades have developed is essential. In this section, the authors focus on the main geological events and accompanying topological changes that led to the emergence of high-elevation habitats, relying heavily on the comprehensive reviews by Orme (2007) and Young *et al.* (2007).

The Northern Andes, housing the relatively small area of páramo in comparison with puna and southern Andean steppe, form a distinctive tectonic realm combining continental crust deformation, accreted oceanic terranes, magmatism, and strike-slip faulting. They arise in Venezuela in the form of two southward-running parallel ranges on either side of Lake Maracaibo, where the highest peak is Pico Bolívar (5007 m). These ranges merge to become part of the Colombian Cordillera Oriental just inside the border with Colombia. Further south the Cordillera Oriental merges with the parallel ranges of the Cordillera Central and Cordillera Occidental in southernmost Colombia. The Andes in Ecuador are initially characterized by two large cordilleras separated by numerous interconnected intermontane valleys, which eventually become less well-defined southward on account of several deep valleys traversing the mountains. The highest peak is Volcán Chimborazo (6310 m). At the southern end of the páramo is found the Huancabamba Depression, thought by some authors (e.g., Molau, 1988) to comprise an important biogeographical barrier for north–south migration, although the alternate hypothesis that strong climatic differences, north and south of the depression, are responsible for floristic differences should also be considered. Under present climatic conditions and heights of the Andes the páramo is highly fragmented (Josse *et al.*, 2009) in comparison with the situation during the glacial cycles when the treeline became depressed by approximately1500–2000 m (van der Hammen and Cleef, 1986). Postglacial conditions are thought to have favored the evolution of many local endemic species.

The huge extension of puna vegetation, the equivalent of over ten times the land area of páramo, develops mainly in the topographically complex Central Andes yet it shows far better physical continuity than the páramo habitat. There are two main cordilleras (Western and Eastern). In northern Peru, these cordilleras are poorly defined due to several deeply traversing valleys. The highest peak here is Huascarán (6768 m). Southward in northwest Bolivia, northern Chile

and northwest Argentina, the Western and Eastern Cordilleras are separated by the Altiplano, a wide (120–250 km), high elevation (3800–4100 m) plain. The present relief of the Central Andes is largely a result of late Cenozoic uplift, igneous activity, and erosional exhumation. The Western Cordillera is a late Cenozoic magmatic arc with many active volcanoes (e.g., Volcán Sajama, 6542 m), whereas the Eastern Cordillera is mostly comprised of Ordovician sediments intruded by granite batholiths, these supporting extinct volcanoes such as Illampu (6550 m) and Illimani (6457 m). The Altiplano, which lay at sea level until about 60 Ma, today contains approximately 10,000 m of Cenozoic continental deposits (Horton *et al.*, 2001). The high Altiplano is important in that it provides good continuity for the dispersal of organisms between the widely separated Western and Eastern Cordilleras. However, an area of severe aridity on the western flanks of the Andes at 24–25° S constitutes an important biogeographical barrier for north–south migration (Villagrán *et al.*, 1981; Arroyo *et al.*, 1988).

Southern Andean steppe covers close to three times the land area of páramo, and just under a quarter of that of puna. The tectonic history of the Southern Andes (including the Patagonian Andes) includes long periods of extension and basin formation during the Mesozoic and Oligocene-Miocene, alternating with relatively short intervals of mountain building/exhumation during the mid-Cretaceous and Late Miocene. In the northern extreme of southern Andean steppe the Andes are comprised of a principal chain of variable width. As a result of aridity, which slows the erosion–deposition cycle, the terrain is very steep. From 34° S south, there is a sharp decrease in the elevation of the main chain and a switch from V-shaped to U-shaped valleys coincident with a change from dominantly fluvial erosion in the north to dominantly glacial erosion in the south. From the Gulf of Penas (42 °C) southward, the Andes are subdivided into the lower Insular Patagonian Andes that run along the Pacific Ocean and the higher eastern Patagonian Andes (>3000 m). Further south, the Patagonian Andes lose elevation to rise to around 1500 m on the north coast of the Straits of Magellan. On the island of Tierra del Fuego the Andes are represented by the Cordillera de Darwin. There are many high peaks (including volcanoes) in the area of the Andes housing southern Andean steppe: to mention a few, from north to south: Mount Aconcagua (6962 m), Volcán Tinguiririca (4623 m), Cerro Tronador (3478 m), Mount San Valentín (4058 m), Cerro Paine Grande (3240 m), and Mount Darwin (2488 m). North of 38° S the higher elevation of the Andes provides good continuity for organisms found in high-Andean steppe vegetation. However, from here southward, under present-day climates, southern Andean steppe occurs on a myriad of isolated, island-like summits, as occurs in páramo. As in the northern Andes, the Pliocene and Pleistocene glacial cycles were associated with significant fluctuations in the altitude of permanent snow and ice (Thomson *et al.*, 2010). The North and South Patagonian Icefields constitute immense north–south biogeographical barriers in the western Patagonian Andes, whose impact on present-day distributions of high altitude plant species has not been fully studied.

Overall, the evidence indicates that high-elevation habitats are geologically young and their floras are recent. With regards to the second point, the most recently evolved species of *Polylepis* (Rosaceae) are distributed at the highest elevations (Simpson, 1979), and this also appears to be true for some species of *Centropogon* (Stein, 1992) and *Fuchsia* (Godley and Berry, 1995). In relation to geological age, the Northern Andes had reached no more than 40% of their modern heights between 15 and 5 Ma, but as of 2–5 Ma rose rapidly to reach their modern elevations by ~2.7 Ma (Gregory-Wodzicki, 2000). By the Late Pliocene/Early Pleistocene (1–2 Ma), a floristically impoverished version of páramo came to occupy large areas between 2000 and 3000 m (Sklenář *et al.*, 2011). The Altiplano had attained no more than a third of its present elevation by 20 Ma, and no more than half its modern elevation by 10 Ma (Gregory-Wodzicki, 2000). Based on the paleoaltitude projections given by Hartley (2003) the Western Cordillera would have reached about 3500 m by about 8 Ma, whereas the Eastern Cordillera and the Altiplano did not reach an equivalent height until <5 Ma and perhaps as late as 3 Ma.

Less is known about paleoaltitudes in the southern Andes. According to Quinteros *et al.* (2005) the Patagonian Andes attained some 400 m elevation by the early-middle Miocene. Blisniuk *et al.* (2005) suggest that uplift of the Patagonian Andes was already sufficient to produce a west–east rainshadow by 16.5 Ma. From Thomson *et al.*'s (2010) recent work, the Andes at 38–40° S seem to have reached around 1800 m in places by 10 Ma. Here the modern treeline lies at around 1400–1500 m, but it must be remembered that the climate today is cooler than at 10 Ma. In any case, because far less uplift is required to support high-elevation vegetation at the southern extreme of the continent, southern Andean steppe probably appeared earliest (cf. Simpson, 1983; van der Hammen and Cleef, 1986), followed by the puna, and finally the páramo where the treeline develops at much higher altitudes. In accordance with the hypothesis of an older southern Andean steppe, Simpson *et al.* (2009) showed that páramo and puna species of *Perezia* diverged later than their southern Andean counterparts. Hershkovitz *et al.* (2006a) showed that the very large high elevation predominantly puna clade comprising *Chaetanthera* subgenus *Egania* diverged at about 3.5 Ma in comparison with a smaller southern Andean steppe clade in subgenus *Oriastrum*, which dates to about 11 Ma. The genus *Oreobolus* (Cyperaceae) is believed to have colonized into the southern Andes at around 5–5.6 Ma from Australasia, to eventually make its way into the northern Andes (Chacón *et al.*, 2006).

Plant Biodiversity in High Elevation Andean Habitats

Here the current state of knowledge concerning plant biodiversity in the entire South American high-elevation belt is reviewed.

Species Richness

Most floristic work in the high Andes pertains to surveys of local areas (e.g., Méndez *et al.*, 2006), altitudinal transects (e.g., Arroyo *et al.*, 1988), or to a particular vegetation zone (e.g., Luteyn, 1999). Where plot data exist, it has been taken using a variety of plot sizes. The high level of physical continuity along

the Andes reflected in the distribution of many plant genera in more than one, and often across the three vegetation zones, makes it imperative to take a broad view of the high-elevation habitat. The first question that may be asked concerns how many plant species inhabit the entire Andean corridor. Arroyo *et al.* (2010) provided a preliminary estimate of 6700 vascular plant species in 870 genera, indicating a very rich high-elevation flora. This number incorporated all páramo species found in the Andes as per Luteyn's (1999) páramo list. Such high species richness in the high-elevation belt is not surprising given the north–south trending nature of the mountain range associated with a great variety of climates and the high-topological diversity already alluded to. These last factors, in addition to providing an outstanding number of environmental niches in the high-elevation belt itself, are associated with a high diversity of subtending lowland vegetation types (cf. Arroyo *et al.*, 1983 for the range of vegetation types in the Chilean Andes), ranging from closed tropical and temperate forests through Mediterranean-type scrublands to open steppe grasslands and desert communities, all providing an outstandingly rich source of evolutionary lineages from which the high-elevation flora has been partially recruited. Sizable genera in the high-elevation belt include: *Adesmia, Astragalus, Baccharis, Calamagrostis, Calceolaria, Carex, Diplostephium, Draba, Elaphoglossum, Espeletia, Festuca, Gentianella, Lupinus, Geranium, Gynoxys, Halenia, Hypericum, Huperzia Lupinus, Miconia, Nototriche, Oxalis, Pentacalia, Poa, Puya, Senecio, Solanum, Valeriana,* and *Viola*. *Senecio* stands out above all other genera with an estimated 300 species occurring in Andean high elevation habitats. *Senecio* is followed distantly by *Calceolaria, Gentianella, Lupinus, Nototriche, Solanum,* and *Valeriana*. By far the largest high elevation plant families globally for the Andes as well as in each of the three high-elevation zones are the Asteraceae and Poaceae.

Our early effort to estimate the size of the high-elevation flora brought forward the many typical difficulties associated with developing a continental-wide checklist for high Andean habitats – poor habitat information on many plant specimens, information not available in a readily accessible form, numerous synonymy problems. In what is an ongoing refinement process some species have been eliminated as better habitat knowledge has come to hand, while others are being added. Additional páramo records are also possible. In his comprehensive treatment of the Columbian páramos Rangel (2000b) points out that many species are missing from Luteyn's (1999) páramo list. Taking these various factors into consideration, the above estimates can be expected to change in the near future.

Information on other plant groups thus far is only available for páramo, which houses 465 species of lichens, 544 species of mosses, and 291 species of hepatics (Luteyn, 1999). Numbers for these groups can be expected to grow substantially as information for the puna and southern Andean steppe becomes available.

Endemism

Andean high-elevation habitats, not surprisingly given their youthful age, support no endemic plant families. However, at the generic level, 64 genera in 22 families (**Table 1**) for all

Table 1 Sixty-four plant genera considered to be confined (endemic) to the high-elevation vegetation belt in the South American Andes

Apiaceae: *Cotopaxia* (PA), *Laretia* (SAS), *Perissicoelum* (PA), *Pozoa* (SAS)
Asteraceae: *Ascidiogyne* (PA), *Blakiella** (PA), *Calopappus** (SAS), *Espletia* (PA), *Floscaldasia* (PA), *Gamochaetopsis** (SAS), *Huarpea** (PU), *Jalcophila* (PA-PU), *Loricaria* (PA-PU), *Marticorenia** (SAS), *Misbrookea* (PU)*, *Mniodes* (PA-PU), *Novenia** (PA-PU), *Oriastrum* (PU-SAS), *Raouliopsis* (PA), *Werneria* (PA-PU-SAS), *Xenophyllum* (PA-PU)
Brassicaceae: *Aschersoniodoxa* (PU), *Brayopsis* (PA-PU), *Catadysia** (PU), *Dactylocardamum** (PU), *Englerocharis* (PU), *Eudema* (PA-PU-SAS), *Paradiodoxa** (PU), *Petroravenia** (PU), *Weberbauera* (PA-PU-SAS)
Cactaceae: *Oroya* (PU), *Yavia** (PU)
Campanulaceae: *Lysipomia* (PA-PU)
Caryophyllaceae: *Pycnophyllopsis* (PU-SAS), *Pycnophyllum* (PU-SAS), *Reicheella** (PU)
Cyperaceae: *Phylloscirpus* (PA-PU-SAS), *Zameioscirpus* (PU-SAS)
Ericaceae: *Plutarchia* (PA)
Juncaceae: *Distichia* (PA-PU), *Oxychlöe* (PU-SAS), *Patosia** (PU-SAS)
Lamiaceae: *Kurzamra** (PU-SAS)
Malvaceae: *Nototriche* (PA-PU-SAS)
Melastomataceae: *Bucquetia* (PA), *Castratella* (PA)
Orchidaceae: *Myrosmodes* (PA-PU)
Plantaginaceae: *Bougeria** (PU), *Melosperma* (SAS)
Poaceae: *Anthochloa** (PU), *Dielsiochloa** (PU)
Portulacaceae: *Lenzia** (PU-SAS)
Pteridaceae: *Nephopteris** (PA)
Ranunculaceae: *Barneoudia* (PU-SAS), *Callianthemoides** (SAS), *Laccopetalum** (PA-PU), *Oreithales** (PA-PU)
Saxifragaceae: *Saxifragella** (SAS)
Scrophulariaceae: *Aragoa* (PA)
Solanaceae: *Combrera* (SAS)
Valerianaceae[a]: *Aretiastrum* (PA-PU-SAS), *Belonanthus* (PA-PU-SAS), *Phyllactis* (PA-PU), *Stangea* (PU)

[a]These genera are considered by some authors as part of *Valeriana*.

PA, found in páramo sector; PU, found in puna sector; SAS, found in southern Andean steppe sector. *, Monotypic genus. Information for Páramo mainly follows Sklenář, *et al.* (2011), but eliminating those genera that are only or partially distributed in the Central American páramos; information for puna and southern Andean steppe from Arroyo (unpublished).

intents and purposes are restricted to the high-elevation habitat considered as a single unit (Arroyo *et al.*, unpublished data). In addition to this large number of endemic genera, several other plant genera in the Andes come close to being endemic, as, for example: *Chrysactium*, *Laestadia*, *Ocyroe*, *Parastrephia* (Asteraceae), *Lithodraba* (Brassicaceae), *Nastanthus* (Calyceraceae), *Acaulimalva* (Malvaceae), *Hamadryas* (Ranunculaceae), *Lampaya* (Verbenaceae).

Twenty-eight of the endemic genera listed in **Table 1** spill out beyond the limits of the three traditionally recognized vegetation zones. Indeed seven of the endemic genera span all three vegetation zones (**Table 1**). Of the 36 genera that are strictly endemic to one of the three Andean sectors (**Table 1**), 12 are from páramo, 15 from puna, and nine from southern Andean steppe. Counter-intuitively the older southern Andean steppe has the least number of sectorially endemic genera. Overall by far the largest concentration of endemic genera in the high-elevation belt is found in puna (43), followed by páramo (31) and southern Andean steppe (25). Puna not only houses the largest total number of endemic genera but is also the center of diversity for many of the largest endemic genera (e.g., *Nototriche, Pcynophyllum*, *Werneria*, *Weberbauera*, *Xenophyllum*).

Presently it is unknown what percentage of plant species are strictly endemic to the entire Andean high-elevation belt. To ascertain that a species is endemic to a particular ecosystem, it must be shown that it does not occur in any other vegetation type and/or geographical area. Such detailed habitat knowledge is still very fragmentary for many high-elevation species. Luteyn (1992) originally suggested that perhaps as many as 60% of páramo species are endemic to that ecosystem, but later (Luteyn, 1999) came to the conclusion that a realistic approximation of species endemism is not yet possible, for the very reasons expressed above. Briceño and Morillo (2006) concluded that 25% of Venezuelan páramo species are true páramo endemics. There has been no attempt to determine the number of truly endemic puna and southern Andean steppe species at this stage. In any case, if specific endemism were looked at in relation to the entire high Andean corridor, as we have done for generic endemism, it would not be surprising if around 25–30% of all high-elevation species turn out to be endemic. Based on the estimated size of the entire high-elevation flora given above, this signifies that around 1700–2000 high-elevation species in the Andes might be endemic.

Origin and Diversification of the High-Elevation Flora

How the high-altitude flora of the South America Andes has been assembled is a fascinating topic for evolutionary biologists and biogeographers. Phylogenetic studies reveal that multiple sources have been involved: the subtending lowland vegetation types, a wider neotropical source, an austral-Antarctic source and a holarctic source. Phylogenetic evidence also confirms that once clades are able to establish in the high-elevation belt, they are likely to undergo north–south migration (cf. Hughes and Eastwood, 2006). Evidence for more than one independent entry into the high-elevation belt has been found in several genera (e.g., *Chaetanthera*, Hershkovitz *et al.*, 2006a, b; *Halenia*, von Hagen and Kadereit, 2003), leading in such cases to increased phylogenetic diversity.

Phylogenies containing páramo taxa were recently analyzed by Sklenář *et al.* (2011). These authors found that temperate and tropical sources contributed about equally to the páramo flora, but that temperate genera of true northern hemisphere origin contributed comparatively more species to the páramo than their southern hemisphere counterparts, in spite of the fact that the arrival of northern hemisphere groups would have occurred through chance long-distance dispersal.

The well known *Espeletia* complex (>130 species) was originally hypothesized to have an ancestor in the subtending montane belt (Cuatrecasas, 1986), and although important advances have been made in understanding the relationships of the complex (Rauscher, 2002), better taxon sampling is needed to iron out the fine details of the colonization process. Several high Andean clades involving a northern hemisphere origin such as *Lupinus* (Hughes and Eastwood, 2006), *Astragalus* (Scherson *et al.*, 2008), *Gentianella* (von Hagen and Kadereit, 2001), *Halenia* (von Hagen and Kadereit, 2003), and *Valeriana* (Bell, 2004; Bell and Donoghue, 2005) have experienced huge radiations in the páramo and puna, and among these are found some of the highest divergence rates ever recorded (Scherson *et al.*, 2008). Although Sklenář *et al.* (2011) found that temperate and tropical sources have contributed equally to the páramo flora, a huge proportion of the very large wet tropical lowland lineages subtending the páramo seem not to have been able to successfully colonize the páramo, presumably because they are less adaptable for life at high altitudes than incoming lineages of temperate origin. It is not clear at this point whether the same holds for the puna and southern Andean steppe. Here there are many examples of moderate to large-sized radiations in genera of South American origin as found in *Chaetanthera* (Hershkovitz *et al.*, 2006a), *Perezia* (Simpson *et al.*, 2009), *Junellia* (ÓLeary *et al.*, 2009), and especially the endemic high-elevation *Nototriche* (see **Table 1**), which is nested in the *Tarasa* clade (Tate and Simpson, 2003), with many other large genera still to be studied (e.g., *Adesmia*, *Acaulimalva*, *Nassauvia*, *Leucheria*, *Werneria*, *Xenophyllum*). The more gradual vegetation transition into the high-altitude environment in the puna and over a large part of southern Andean steppe could be expected to facilitate easier establishment and radiation of clades of South American origin into the high-elevation habitat.

Apart from major radiations that commonly cover two and even three of the major vegetation zones, the high-elevation flora has been enriched by small numbers of high altitude species evolving at different places in the phylogenies of groups mainly centered at lower to midelevations (e.g., *Tropaeolum*, Hershkovitz *et al.*, 2006b; *Malesherbia*, Gengler-Nowak, 2003; *Ourisia*, Meudt and Simpson, 2006) or from migration along the Andean corridor from the initial point of establishment, as seen in case of the arrival of *Astragalus* in páramo (Sklenář *et al.*, 2011) and *Ourisia* in páramo and puna (Meudt and Simpson, 2006) where northward migration has been involved. Although good phylogenies do not exist to ascertain the directionality of colonization, an important source of enrichment in the far southern Andes concerns a number of small endemic South America genera today occurring in the Patagonian steppe as well as at high elevations

(e.g., *Benthamiella*, *Gamocarpha*, *Oreopolus*, *Hamadryas*, *Huanaca*, *Onurus, Bolax*). These genera, along with larger-radiating genera in the southern Andes such as *Azorella* and *Nassauvia*, could have comprised very early constituents of the original South American high-elevation flora. *Hamadryas* is a recently diverged clade within the *Ranunculus* complex (Hoot *et al.*, 2008) and knowledge about when it diverged would shed important light on this important question. Other sources of enrichment include dispersal from an austral-Antarctic quarter. For example, as was already mentioned, *Oreobolus* is hypothesized to have arrived by long-distance dispersal in the southern Andes from Australasia at 5.5–6 Ma (Chacón *et al.*, 2006) to later colonize into the northern Andes. *Abrotanella* is hypothesized to have existed on Antarctica in the mid-late Tertiary, to disperse northward into the Andes as well as into Australasia at around 3 Ma (Wagstaff *et al.*, 2006). Overall, chance long-range dispersal has played an important role in bringing distant clades into the Andean high-elevation flora. Long to intermediate range dispersal has probably also played an important role in extending the latitudinal distributions of many resident clades, and might explain the outstanding success of the Asteraceae, a family which is generally well adapted for wind dispersal.

Twenty-five high Andean endemic genera are monotypic (**Table 1**). This high Andean biodiversity component is an intriguing one. Phylogenetic studies are now revealing that many of these monotypic genera are turning out to be embedded in other larger Andean clades. For example, southern Andean steppe *Laretia*, a dominant cushion plant in the central Chilean Andes, is sister to *Azorella madreporica* distributed in the puna (Nicolas and Plunkett, 2009). Until recently, *Urbania* was recognized as a monotypic genus of the puna, where it occurs at very high elevations. ÓLeary *et al.* (2009) showed that it is embedded within a well supported clade of *Junellia*. Subsequently, it was transferred to *Junellia* (ÓLeary *et al.*, 2011). The high Andean monotypic cushion genus *Patosia* forms a well-supported clade with *Oxychlöe,* another small Juncaceous cushion genus (**Table 1**) (Záveská Drábková and Vlček, 2009). Monotypic *Oreithales* centered in the puna and páramo is nested within a clade of *Anemone* (Meyer *et al.*, 2010), as is the slightly larger southern Andean steppe endemic high-elevation genus *Barneoudia* (Ranunculaceae). *Laccopetalum* and *Krapfia* (occurring also at lower altitude) are sister to one another, and belong to a terminal branching clade of the core *Ranunculus* group (Hoot *et al.*, 2008); however, taxon sampling in *Krapfia* is presently insufficient to determine whether *Laccopetalum* diverged earlier than *Krapfia*. Overall, phylogenetic results tend to suggest recent origins for many of the high-elevation monotypic endemic genera. However, in other cases, relictual status seems more likely, as seen in the basal position of the endemic monotypic *Caloppapus* (Asteraceae) of the central Chile Andes with respect to a clade of *Nassauvia* and *Triptilion* (Simpson *et al.*, 2009). Likewise, *Huarpea*, which is basal to a clade of *Barnadesia* (Gruenstaeudl *et al.*, 2009) – the latter a genus of montane and lowland forest habitats in the Andes and Brazil – could turn out to be a relic taxon. The rare monotypic austral alpine genus *Saxifragella* was found to be deeply embedded in arctic and northern hemisphere *Saxifraga sensu stricto* and is considered to have arrived in the South American Andes via long-distance dispersal (Soltis *et al.*, 2001), perhaps at an early stage. A fascinating case concerns *Kurzamra* (Lamiaceae), a monotypic dwarf shrub occurring at the upper limits of the vegetation. In a well sampled phylogeny *Kurzamra* occurs in a fairly well-supported clade with species of *Cuminia*, a small woody endemic genus of the Juan Fernández islands (Bräuchler *et al.*, 2010). This relationship suggests that extinction of intermediates in the intervening Atacama desert might have occurred. Alternatively chance of long-distance dispersal into the high Andes associated with rapid morphological evolution might have been involved. To advance in our understanding the evolution of these fascinating high Andean genera, dated phylogenies, and often better taxon sampling are badly needed.

Páramo

Páramo occurs above treeline in Venezuela, Colombia, Ecuador and northern Peru and in outlying locations in Costa Rica and Panama outside the Andean cordillera and occurs between 3000 and 5000 m. It is characterized by remarkable constancy in mean monthly temperature but large diurnal temperature fluctuation (Sarmiento, 1986; Rundel, 1994). In contrast to temperature constancy, precipitation patterns in the páramo are exceedingly complex in terms of amount and seasonality (for detailed reviews see Sarmiento, 1986; Rundel, 1994) varying from approximately 600 to 4400 mm (Sarmiento, 1986; Rangel, 2000a), and from a bimodal pattern in the Venezuelan and Colombian Andes facing the Caribbean Sea to a unimodal in Los Llanos (eastern slopes). Further south in the Ecuadorian Andes, the more pronounced dry season tends to shift to midyear, and the bimodal pattern is almost lost.

While much ecophysiological work has been carried out in páramo, little is known about the pollination and breeding systems of páramo plants. A recent study by Barrios *et al.* (2010) showed that several woody species at a Venezuela subpáramo location are highly dependent on insect pollination and amply visited by bees and other insect groups. In *Espletia schultzii*, a strongly outcrossing insect-pollinated species in the tropical Venezuelan Andes, breeding system remains fairly constant among populations (Sobrevila, 1989). However, wind pollination appears at higher elevations (Berry and Calvo, 1989), suggesting a possible deficit of biotic pollinators at the highest elevations, but maintenance of an outcrossing breeding system. Arroyo *et al.* (1983) showed that high Andean wind-pollinated plant species found in southern Andean steppe have larger geographical ranges that biotically pollinated species. The acquisition of wind pollination at the very high elevations in the Andes, as seen in *Espeletia*, has probably facilitated dispersal within the páramo habitat.

For the relatively small land area it occupies, páramo is the richest and possibly most endemic high mountain ecosystem in the world, with 3399 vascular plant species in 500 genera (Luteyn and Churchill, 2000). The largest plant families are: Asteraceae, Poaceae, Orchidaceae, Scrophulariaceae, Melastomataceae, Gentianaceae, Ericaceae, Bromeliaceae, Rosaceae, and Fabaceae (Luteyn, 1999). The páramos of Venezuela support 1420 species of angiosperms (Briceño and Morillo, 2006), while those of Colombia support 3379 species and subtaxa of spermatophytes (Rangel, 2000b). Local páramo

Table 2 Number of plant species (*N*) reported for a selection of localities in the three high elevation vegetation zones of the South American Andes

Sector and locality	*Lat*	*Alt*	*N*
Páramo			
Páramo de Sumapax, Colombia[1]	3° N	NA	619
Massif de Chingaza, Colombia[1]	4° N	NA	534
Páramos meridionales, Ecuador[2]	3° S	2850–3635 m	216
Páramos El Espino y Palambe, Sallique, Jaén, Peru[3]	5° S	3000–3560 m	252
Ramal de Guaramacal, Venezuela[4]	9° N	2800–3100 m	252
Puna			
Lags. Pomacocha y Habascocha, Peru[5]	11° S	4350–4550 m	100
P.N. del Manu, Peru[6]	12° S	3300–4100 m	355
Cordillera de Vilcanota, Peru[7a]	13° S	?–4900 m	145
Parque Nacional Lauca, Chile[8b]	18° S	1540–5200 m	391
S. Coposa-Cordón Collaguasi, Chile[9]	20° S	3800–4700 m	97
P.N. Llullaillaco, Chile[10]	24° S	3500–6750 m	126
R.B. Laguna Blanca, Argentina[11]	26° S	3200–5500 m	75
Southern Andean steppe			
Lag. Grande, Lag. Chica, Chile[12]	28° S	3200–4300 m	123
Cordillera de Doña Ana, Chile[13]	30° S	>3000 m	161
P. Provincial Aconcagua, Argentina[14]	32° S	2470–4200 m	124
Cordón del Plata, Argentina[15b]	32° S	3000–4500 m	295
Farellones-La Parva, Chile[16]	33° S	2200–4100 m	309
P. Provincial Tromen, Argentina[17c]	37° S	2100–3600 m	180
P.N. Nahuelhapi, Argentina[18]	40° S	>1700 m	232
P.N. Los Glaciares, Argentina[19]	49° S	800–1000 m	178
P.N. Torres del Paine, Chile[20]	50° S	700–1300 m	179
Cerro Santa Lucía, Chile[21]	50° S	900–1700 m	154
Cerro Martial, Argentina[22]	54° S	600–1250 m	80

[a]Only puna and high *Polylepis* forests included.
[b]Includes a small number of prepuna and desert edge species.
[c]Only "piso altoandino" included.
PA, Páramo; PU, Puna; SAS, Southern Andean steppe; Lat, latitude; Alt, Altitude.
Source: [1]Rangel (2000c), [2]Izco, *et al.* (2007), [3]Marcelo Peña, *et al.*, (2006), [4]Cuello, *et al.* (2010), [5]Flores, *et al.* (2005), [6]Young and Cano (1994), [7]Tupayachi (2005), [8]Arroyo, *et al.* (1988), [9]Teillier (2006), [10]Marticorena, *et al.* (2004), [11]Borgnia, *et al.* (2006), [12]Arroyo, *et al.* (1984), [13]Squeo, *et al.* (1994), [14]Méndez, *et al.* (2006), [15]Méndez (2009), [16]Arroyo, *et al.* (1983), [17]Chiapella and Ezcurra (1999), [18]Ferreyra, *et al.* (1998), [19]Iribarren and Ferreyra (2011), [20]Arroyo, *et al.* (1992), [21]Arroyo, *et al.* (1989), [22]Mark, *et al.* (2001).

floras are generally richer than in the puna and southern Andean steppe (**Table 2**). Páramo is characterized by major species radiations as seen in the *Espeletia* complex (>130 species), which includes "Frailejones" (**Figure 2**). These are giant rosettes that can grow to 5 m tall and live >100 years (Monasterio, 1980). The spiral arrangement of leaves and their dense layer of hairs together with the retained dead leaves, provide both thermal insulation and protection against UV radiation (Meinzer *et al.*, 1994; Rada *et al.*, 1985). While the emblematic *Espletia* complex is restricted to páramo, many species of páramo are also found in puna. An interesting biodiversity question here for future research concerns whether southward expansion of páramo clades of neotropical origin, has been more difficult than northward expansion of puna clades.



Figure 2 Páramo in Ecuador, 3600 m dominated by frailejones (*Espeletia* spp.).

Subpáramo

Páramo is subdivided into three main altitudinal zones: superpáramo, páramo, and subpáramo (Cuatrecasas, 1968; Cleef, 1981; Ramsay and Oxley, 1997; Jorgenson and León-Yánez, 1999; Luteyn, 1999). Nevertheless, for the Venezuelan páramos, Monasterio (1980) recognizes two altitudinal zones, in Spanish called 'pisos altitudinales': a high Andean zone or

'Piso Altiandino' (4000–4800 m) and the upper Andean zone or 'Piso Andino Superior' (2800–4000 m). Here the three main altitudinal zones are reviewed mainly based on information given in Luteyn (1999) and some particular communities mentioned by Monasterio (1980) will be highlighted.

The subpáramo constitutes a transition zone between the upper montane forests and true páramo. Found between 3000 and 3500 m it is comprised of grasslands intermixed with isolated thickets of shrubs and dwarf trees. Representative shrubby plants include species of *Aragoa*, *Berberis*, *Bejaria*, *Hypericum*, *Macleania* and *Vaccinium*, *Brachyotum*, *Miconia*, *Gynoxys*, *Diplostephium*, *Hesperomeles*, and the bambusoid grass *Chusquea*. Small trees in the subpáramo thickets include species of *Buddleja*, *Escallonia*, *Gynoxys*, *Hesperomeles*, *Myrsine*, *Oreopanax*, *Schefflera*, and *Weinmannia*. A number of species found principally in montane cloud forests occasionally inhabit the subpáramo, contributing significantly to the overall high diversity of páramo in general.

True Páramo

True páramo, also referred to as grass páramo, extends from approximately 3500 to 4100 m and is dominated primarily by grasses, in particular species of *Calamagrostis* and *Festuca*. Shrubs of *Baccharis*, *Diplostephium*, *Loricaria*, *Pentacalia*, *Hypericum*, *Gaultheria*, *Pernettya*, and *Vaccinium* are common, as well as rosette and cushion plants of the genus *Acaena* and *Plantago*. The *Espeletia* complex is well represented and several species such as *Espeletia barclayana*, *Espeletia brachyaxiantha*, *Espeletia jaramilloi* and *Espeletiopsis corybosa*, *Espeletiopsis guacharaca*, *Espeletiopsis muiska*, and *Espeletiopsis colombiana* are more or less restricted to bunchgrass páramo. Common herbaceous genera include *Acaena*, *Bartsia*, *Bomarea*, *Cerastium*, *Eryngium*, *Gentianella*, *Halenia*, *Jamesonia*, *Lachemilla*, *Lupinus*, *Lycopodium sensu lato*, *Paepalanthus*, *Senecio*, and *Sisyrinchium*.

Superpáramo

Superpáramo usually occurs at above 4000 m. The superpáramo landscape supports a broad array of mosses, lichens, herbaceous vascular plants (such as *Arenaria* spp., *Cerastium floccosum*, *Astralagus geminiflorus*, *Azorella pedunculata*, *Disterigma empetrifolium*, *Draba* spp., *Eudema nubigena*, *Geranium multipartitum*, *Hypochaeris sessiliflora*, *Senecio* spp., *Luzula racemosa*, *Nototriche* spp., *Plantago sericea*, *Valeriana alpifolia* and *Viola pygmaea*), and numerous other grasses. Superpáramo includes a rich collection of endemic vascular plants as typified by *Senecio cocuyanus*, *Senecio adglacialis*, *Senecio supremus*, *Senecio pasqui-andinus*, *Senecio cleefii*, and *Senecio guicanensis*. *Senecio santanderensis* is only known from the Páramo de Santurbán and the Páramo de Almorzadero in Colombia (Rangel, 2000b). The floras of dry and humid superpáramo in Ecuador differ in relation to the importance of tropical and temperate elements, with the latter more abundantly represented in the drier páramos (Sklenář and Balslev, 2006).

Northern Andean Alpine Cushion Formations

Cushion-like plants of the genera *Azorella*, *Arenaria*, and *Montia* can be well developed from 4200 to 4800 m in páramo (Monasterio, 1980). At the highest altitudes these formations may be represented by solitary cushions of *Azorella* spp., which can reach heights of up to 40 cm. In less exposed areas the cushions are accompanied by species such as *Gentiana sedifolia* and *Lysipomia sphagnophila* and various bryophytes such as *Tortula andicola* and *Zygodon pichinchensis*.

Northern Andean Dwarf *Polylepis* Forest

Scattered patches dominated by dwarf trees of *Polylepis* (Rosaceae) are often found in páramos. Arborescent *Polylepis*, comprising 28 species, is endemic to the high tropical and subtropical Andes (Schmidt-Lebuhn *et al.*, 2006). Palynological studies show that the genus has been established in the Andes for at least 2.7 million years, and certain species are considered to grow at altitudes higher than any other arborescent angiosperm in the world (Hoch and Körner, 2005). Species found in the páramo include *P. lanuginosa*, *P. reticulata and P. weberbaueri*, *P. quadrijuga*, and *P. sericea* among others. *P. sericea* has been widely studied from an ecophysiological point of view (e.g., Goldstein *et al.*, 1994; Rada *et al.*, 1996, 2009; Colmenares, 2002).

Puna

Puna occurs from northern Peru south through Bolivia to northern Argentina and northern Chile and is found from 3300–3700 to over 5000 m (Cabrera, 1971; Arroyo *et al.*, 1988; Luteyn and Churchill, 2000). Unlike in páramo, a consistent treeline is lacking throughout much of the lower limit of puna (especially on the western side of the Andes), making the definition of its lower limit for floristic and ecological studies difficult. While diurnal temperature fluctuations are wide, the annual temperature range is low. At Parinacota (18° S) located on the western side of the puna at 4400 m, mean temperature is 1°C, freezing temperatures occur during every month of the year, and diurnal temperature fluctuation can be as high as 30 °C (Torres *et al.*, 1978). North of 24–25° S, the Andean mountains are much drier on the western side than on the eastern side (Sarmiento, 1986; Arroyo *et al.*, 1988), and there is a trend for increasing aridity from north to south. The northern part of the central Andes generates a rainshadow by forcing moisture laden air to rise and cool on their eastern slopes. In central Peru and northern Bolivia, the amount of moisture crossing the eastern cordillera is still relatively high and the vegetation still quite moist on the western flanks. However, in northern Chile and southern Peru, as the air masses descend toward the western Pacific slopes, they undergo further drying on account of the cold surface waters of the Humboldt Current, producing significant aridity above 3000 m in the Western cordillera. Precipitation (mostly in summer as snow) at Parinacota (18° S) on the western side is approximately 160 mm with high variability among years (Rundel *et al.*, 2003); to the east at La Paz (16° S, 4100 m) the puna receives 564 mm per year (Sarmiento,



Figure 3 Puna in Bolivia dominated by *Festuca* spp.

1986). South of 24–25° S the western flanks of the Andes experience a radical shift in climate. Here the scarce puna precipitation is derived from the Polar front and falls mostly during the winter months (Arroyo *et al.*, 1988).

Compared with páramo and southern Andean steppe, the puna landscape is vegetationally less diverse (**Figure 3**) and can be monotonous in the drier areas. Showy flowers are less frequent than in the páramo and southern Andean steppe. While insect pollination occurs, bee pollination is less frequent, and pollination rates are generally lower than in southern Andean steppe (Arroyo and Squeo, 1990). With increasing elevation and aridity wind pollination becomes more important (Arroyo *et al.*, 1983). Three altitudinal vegetation belts are often recognized in puna: (1) the subalpine belt; (2) a low alpine belt, and (3) a high alpine belt where permafrost is usually close to the soil surface and only a few species forming small rosettes and grasses are found (Ruthsatz, 1974; Villagrán *et al.*, 1981, 1983; Arroyo *et al.*, 1988).

Local puna floras, especially in dry and desert puna, are generally more impoverished than in the páramo and in southern Andean steppe (**Table 2**). Currently there is no published figure for the entire vascular flora of puna. An early estimate of the puna flora of northern Chile came up with 865 species (Arroyo *et al.*, 1997) while for equivalent latitudes on the eastern side of the Andes in Argentina the estimate is for 1034 species (Novara, 2003), with Asteraceae, Poaceae, Fabaceae, and Solanaceae dominating, and with many species shared on the two sides of the Andes. It has been estimated that the total puna flora, taking its full extent in Peru, Bolivia, Chile, and Argentina into account, could come close to 3000 species (Arroyo, unpublished data). Large plant families include: Asteraceae, Poaceae, Fabaceae, Malvaceae, Solanaceae, Caryophyllaceae, Gentianaceae, Brassicaceae, Valerianaceae, and Calceolariaceae. Compared with páramo, it can be seen that there are fewer large families with tropical affinities in puna. The authors caution that the puna is by far the most difficult of the three Andean areas to delimit physically, such that the estimate given here is likely to change in either direction on more detailed study. Despite the very different climates of páramo, puna, and southern Andean steppe, and some significant biogeographical barriers, several genera centered on the puna have representatives in páramo and/or southern Andean steppe, among them *Werneria, Xenophyllum, Nototriche, Weberbauera*. Until good phylogenies become available, of course, it cannot be ascertained whether these genera originated in the puna to migrate north and south.

Humid Puna

Humid puna extends from north central Peru adjacent to the páramos (jalcas) in southeasterly direction to along the eastern margin of the Altiplano and Eastern Cordillera of Bolivia, with a small southward easterly extension into Argentina (Cabrera, 1968). Annual precipitation ranges from 400 to 1500 mm. It supports a well-developed plant cover and is often dominated by an uninterrupted carpet of grasses (e.g., *Calamagrostis, Festuca,* and *Stipa*). Other common genera of the humid puna include prostrate, rosette, or cushion species of *Azorella, Daucus, Baccharis, Werneria, Draba, Echinopsis, Gentiana, Geranium, Hypsela, Isoetes, Lilaeopsis, Lupinus, Nototriche, Ourisia, Oxychloe, Oriastrum, Plantago,* and *Pycnophyllum*, low shrubs in *Adesmia, Calceolaria, Mutisia, Satureja, Tetraglochin,* and several species of *Senecio* of diverse habit. A good number of species puna species are shared with páramo.

Dry Puna

Dry puna is found in southern Peru just north of Arequipa, from along the western Altiplano of Bolivia, in northernmost Chile to around 20° S and comprises practically all of the puna found in northwestern Argentina, including the fairly moist kind of puna found in Cumbres Calchaquíes in the Province of Tucuman (Cabrera, 1968). Rainfall ranges from 100 to 400 mm per year and ground cover is usually less than 50%. At the lower elevations, dry puna is dominated by extensive shrublands and thickets of the resinous tola shrubs, such as *Acantholippia, Parastrephia, Fabiana, Chuquiraga, Junellia,* and *Chersodoma, Nardophyllum, Ephedra, Acantholippia*. At higher elevations stiff bunchgrasses of *Festuca orthophylla* become common, interspersed with species of *Nototriche, Werneria, Adesmia, Baccharis, Ephedra, Senecio, Pycnophyllum,* and *Tetraglochin*. The hard cushion plant *Azorella compacta* is commonly found abundantly in dry puna.

Desert Puna

Desert puna is located mostly in the southern part of the Central Andes along the western cordillera of Bolivia and adjacent to the Atacama Desert region of northern Chile. It receives <100 mm annual precipitation. The vegetation is sparse ($<15\%$ cover) and local species richness tends to reach its lowest for puna (**Table 2**); in contrast with more typical elevational gradients, species richness peaks at mid-elevation. Species richness (including the desert edge) drops from 333 species at 18° S in dry puna to 59 at 24° S in desert in the Chilean Andes and is much lower than across the Andes in Argentina (Arroyo *et al.*, 1988). Generally the vegetation is composed of xerophytic plants, including cushion-forming cacti, thorny shrubs, and bunchgrasses such as *Festuca*

orthophylla and *Stipa* spp. Prominent genera include *Adesmia, Mulinum, Anthobryum, Baccharis, Chersodoma, Cortaderia, Fabiana, Lampaya, Lophopappus, Margyricarpus, Oreocereus, Parastrephia,* and *Urmenetea*. A strong floristic break has been detected between desert puna and southern Andean steppe on the Chilean side of the Andes, whose strongest manifestation is seen in the lower puna belts. This has been attributed to extreme aridity at very high elevations at latitude 24–25° S (Villagrán *et al.*, 1983).

Central Andean Dwarf *Polylepis* Forest

More than 20 species of *Polylepis* can be found in humid puna, many with restricted distributions (Kessler and Schmidt-Lebuhn, 2006; Schmidt-Lebuhn *et al.*, 2006). However, species such as *P. tarapacana* found in Bolivia, Peru, and Chile and *P. tomentella* found from Colombia to NW Argentina have wider distributions. *Polylepis* dwarf forests in puna are encountered from about 3700 to 4600 m and in some places they can cover entire high-elevation slopes as on Volcán Sajama in Bolivia. In wet puna the trees are often festooned with mosses, vines such as *Bomarea, Loasa, Mutisia,* and *Passiflora*, mistletoes (*Tristerix* spp.), and many epiphytic flowering plants (Galán de Mera *et al.*, 2003).

Cushion Bogs and Salar Vegetation

Extensive high azonal elevation bogs (bofedales) are found throughout puna becoming progressively smaller, more widely spaced, and floristically impoverished toward the southern puna limit (Arroyo *et al.*, 1988). Bofedales in humid puna tend to intergrade with the zonal vegetation, whereas in dry and desert puna they become more conspicuous on account of their dark green color contrasted against the more zonal puna vegetation. The flora of these bogs tends to be more uniform latitudinally than the surrounding zonal vegetation (Arroyo *et al.*, 1988), and includes cushion species of *Oreobolus, Distichia,* and *Oxychloe,* together with species of *Scirpus, Juncus, Calamagrostis, Lachemilla, Colobanthus, Werneria, Aa,* and *Gentiana*. The large latitudinal distributions of bog taxa have been attributed to good dispersal and the availability of similar azonal conditions along the Andean chain independently of macroclimatic differences.

Large saline deposits representing ancient lakes are also common throughout dry and desert puna (Arroyo *et al.*, 1988). Some of the typical cushion bogs species are able to survive on their edges where there is still some water, but such species as *Distichlis humilis, Muhlenbergia fastigiata, Salicornia pulvinata, Lampaya* spp. *Suaeda foliosa, Tessaria absinthoides,* and *Triglochin maritima* tend now to be more common (Beck, 1985).

Southern Andean Steppe

Southern Andean steppe occurs south of the puna to the extreme end of the continent. The lower altitudinal limit is found at approximately 3200 m at 28° S and at 500–600 m in the extreme south. In northern areas a treeline is entirely lacking. From 31° to 34° S on the western side of the Andes, an open treeline of *Kageneckia angustifolia* (Rosaceae) is found. South of 35–37° S closed-canopy forest forming a well-defined treeline appears. Treeline species can be deciduous or evergreen species of *Nothofagus,* evergreen *Austrocedrus chilensis, Araucaria araucana,* and *Pilgerodendron uviferum,* depending on precipitation level and degree of continentality. *N. betuloides* may form krummholz on the wetter western maritime mountains in the southern canal zone (Wardle, 1998), while *N. antarctica* does so on drier easterly mountains.

Southern Andean steppe develops over a large range of precipitation conditions. Most precipitation (which falls mainly in the winter as snow) derives from a southwesterly Pacific source, and in parts of Argentina, from an Atlantic source. Because of the strong west to east rainshadow, the western side of the southern Andes generally receives more rainfall than the eastern side, and this source increases southward to diminish again in the far south. The western side of the Andes is still quite arid at 30° S (242 mm at 3700 m; Squeo *et al.*, 1993) and 33° S (445 mm at 2500 m; Santibañez and Uribe, 1990) (**Figure 4**), but even more so in adjacent areas in Argentina where mean annual precipitation at Puente del Inca (2720 m) is approximately 200 mm (Méndez *et al.*, 2006). Southward at 40° S precipitation reaches 1746 mm above a treeline situated at 1350 m (estimated from Hijmans *et al.*, 2005) and occurs in both summer and winter. Further south (46° S), but on the eastern side of Patagonian Andes, at an altitude of 1310 m (just above treeline), annual precipitation drops to approximately 717 mm. Based on treeline and climatic characteristics, southern Andean steppe should probably be subdivided into a more northerly section to around 35–37° S and a southerly section extending to the end of the continent.

Southern Andean steppe is mostly dominated by a mixture of perennial herbs, low rounded shrubs and cushions plants, with accompanying grasses, the latter rarely as dominant as in puna. Plants of this high-elevation Andean vegetation zone have developed numerous adaptations for survival, including some the highest tolerance levels to freezing temperatures ever reported, these involving different physiological mechanisms (Squeo *et al.*, 1996; Sierra-Almeida *et al.*, 2009, 2010;



Figure 4 *Tropaeolum polyphyllum* in the subandean vegetation belt, southern Andean steppe, 32° S, 2400 m, Valle de Juncal, Chile.

Sierra-Almeida and Cavieres, 2012). Deep rhizomes or taproots, and fleshy leaves are commonplace. Growth forms are highly diverse. At the higher elevations, many species form flat cushions or dense mats, which can withstand strong winds and cold air-temperatures. Cushions can be herbaceous (e.g., *Oreopolus glacialis*), semiwoody (*Oxalis erythrorhiza*), or woody (*Adesmia caespitosa, A. subterranea, Azorella monantha, Laretia acaulis*). Low rosettes are common (e.g., *Barneoudia major, Pachylaena atriplicifolia, Nastanthus spathulatus*). The high-altitude conditions are also associated with interesting reproductive adaptations. Many species, including on the southern extreme of the continent, have large, brightly colored flowers and have been shown to be insect pollinated (Arroyo *et al.*, 1982; Arroyo and Squeo, 1990; Squeo, 1991). In central Chile species found in the lower part of the southern Andean steppe are mainly biotically pollinated by a rich array of bees, butterflies, and flies (Arroyo *et al.*, 1982, 1983, 1987). With increasing altitude, and colder temperatures bee pollination declines in frequency as do flower visitation rates (Arroyo *et al.*, 1985), while fly and butterfly and wind pollination tend to increase in frequency (Arroyo *et al.*, 1982; Arroyo and Squeo, 1990). Despite the very harsh environmental conditions a surprising number of plant species in the Patagonian alpine are genetically self-incompatible and dioecious (Arroyo and Squeo, 1990).

Knowledge of the floristic composition of southern Andean steppe at different latitudes along both sides of the Andes is now fairly good (**Table 2**). Local richness tends to be higher than in the puna and can approach that at some páramo sites. In spite of good local knowledge, there is no published figure for total species richness in southern Andean steppe. From the authors' ongoing work, it has been estimated that the southern Andean steppe contains around 1400–1500 species. As in the puna, a lack of a treeline over a significant part of the southern Andean steppe often makes decisions as to where a species belongs to this particular vegetation zone very difficult. The largest plant families are: *Asteraceae, Poaceae, Fabaceae, Brassicaceae, Cyperaceae, Apiaceae, Violaceae, Solanaceae, Calceolariaceae, and Boraginaceae.* Important and interesting genera are *Acaena, Adesmia, Anemone, Astragalus, Azorella, Barneoudia, Berberis, Bolax, Calandrinia, Calceolaria, Caltha, Carex, Cerastium, Chaetanthera, Chuquiraga, Colobanthus, Descampsia, Draba, Empetrum, Ephedra, Erigeron, Eudema, Festuca, Hamadryas, Haplopappus, Jaborosa, Juncus, Laretia, Leucheria, Luzula, Montiopsis, Moschopsis, Mulinum, Mutisia, Nassauvia, Nastanthus, Olsynium, Oreopolus, Ourisia, Oxalis, Oxychlöe, Patosia, Perezia, Pernettya, Phacelia, Poa, Ranunculus, Rhodophiala, Ribes, Saxifragella, Saxifragodes, Schizanthus, Stachys, Stipa, Tropaeolum, Viola, Weberbauera,* among others. Some cushion species support rich associations of species (Cavieres *et al.*, 2002, 2006; Arroyo *et al.*, 2003). This so-called "nurse effect" has been shown to result in a net increase in species richness at the community level (Cavieres and Badano, 2009).

Latitudinal Variation and Altitudinal Zonation

In the more northerly sector of southern Andean steppe on the western side of the Andes three altitudinal subdivisions are usually recognized (Arroyo *et al.*, 1981; Squeo *et al.*, 1993; Teillier *et al.*, 1994; Cavieres *et al.*, 2000a), each with a characteristic flora and life forms. On the Argentine side of the Andes vegetation belts are less clearly delimited (Chiapella and Ezcurra, 1999; Ferreyra *et al.*, 1998).

At 33° S in central Chile the subandean scrub belt occurs between 2200 and 2600 m and is characterized by the dominance of prostrate shrubs; a lower Andean belt occurs between 2700 and 3100 m and is characterized by the dominance of cushion plants; and an upper Andean belt occurs between 3200 and 3800 m, and is characterized by a sharp decrease in vegetation cover, with dominance of tussock grasses and small rosette plants. Typical shrub species in the subandean belt are *Acaena splendens, Anarthrophyllum cummingii, Berberis empetrifolia, Chuquiraga oppositifolia, Mulinum spinosum,* and *Nassauvia axillaris*. Annual species such as *Collomia biflora* and *Chaetanthera* spp., and geophytes, such as *Rhodophiala rhodolirion* and *Zoellneriallum andinum* are not uncommon here. Cushion plants such as *A. monantha* and *Laretia acaulis* are the dominant species in the lower Andean belt, where herbs and subshrubs such as *Nassauvia pyramidalis* and *Senecio francisci* also become important in terms of vegetation cover. Typical elements of the upper Andean belt are small herbs and rosette forming species such as *Nassauvia lagascae, Oxalis compacta, Chaetanthera lycopodioides, C. apiculata, C. pusilla, Nassauvia pinnigera, Viola atropurpurea,* and *V. philippii* and cushions such as *Azorella madreporica* and *Oreopolus glacialis*. Many of these same species and others occur on the Argentinean side of the Andes (Chiapella and Ezcurra, 1999; Ferreyra *et al.*, 1998) where local species richness can be very high (**Table 2**).

Further south in the wetter Patagonian Andes two main above-treeline vegetation belts are discernable: (1) a low andean belt from treeline to around 800–950 m (**Figure 5**), dominated by large cushion species such as *A. monantha, Abrotanella* spp., *Adesmia salicornia,* and *Bolax gummifera,* accompanied by the low prostrate shrubs *Empetrum rubrum* and *Pernettya pumila* and perennial, sometimes loose mat-forming herbs such as *Adesmia villosa, Azorella filamentosa, A. fuegiana, Calceolaria uniflora, Ourisia breviflora, Oxalis loricata, Perezia pilifera, Senecio magellanicus,*

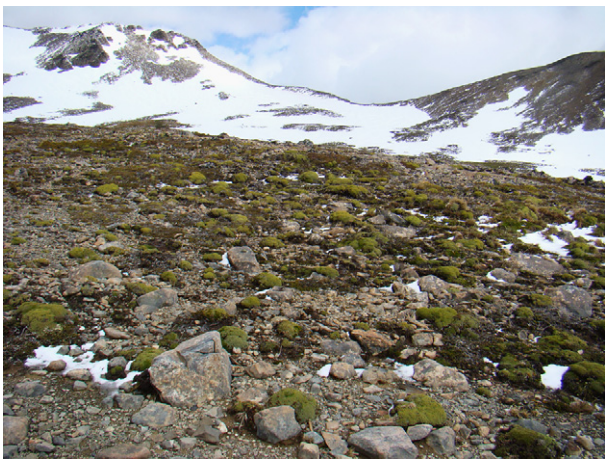


Figure 5 Lower alpine belt in the Patagonian Andes, Cordillera Darwin, Tierra del Fuego, 55° S, approximately 600 m, Chile, dominated by the cushion species *Bolax gummifera* (Apiaceae).

and *Luzula* spp. and (2) a high alpine belt composed of hemicryptophytes, subshrubs, and grasses including *Hamadryas kingii, Moschopsis rosulata, Perezia megalantha, Nassauvia lagascae, Nassauvia pygmaea, Leucheria leontopodioides, M. nordenkjoeldii, Saxifraga magellanica, Poa alopecurus,* and occurring to 1200–1400 m (Arroyo *et al.*, 1992). Very rare plants such as *Epilobium conjugens, Saxifragodes albowiana, Tetrachondra patagonica,* and circumpolar *Azorella selago* may be found in very wet areas. On the drier more continental eastern Patagonian mountains, where a treeline can be lacking, and where the effect of the Pleistocene glaciations is less evident, these same vegetation belts are displaced upward by about 200 m, and are noticeably richer in species than their western counterparts (Arroyo *et al.*, 1989). On these eastern mountains are found many additional species of *Adesmia, Draba, Leucheria, Oxalis,* and especially *Senecio,* as well as several species belonging to genera endemic to Patagonia, among these, *Hamadryas, Benthamiella, Saxifragella,* and *Onurus* (Arroyo *et al.*, 1989; Iribarren and Ferreyra, 2011). Many different taxa have evolved the cushion habit, for example, *Plantago uniglumus, Benthamiella azorella, Benthamiella nordenskjoldii, Hamadryas sempervivens, Oreopolus glacialis, Valeriana (Aretiastrum) sedifolia*. Local richness on these easterly Patagonian mountains is surprisingly high (**Table 2**).

High Andean Bogs and Seeps

On wet sites throughout southern Andean steppe cushion bog communities and seeps often called 'vegas,' or 'mallines' develop. In the northern part of the southern Andean steppe at high elevations, these are dominated by cushion species, namely of the genus *Oxychlöe* and *Patosia clandestina,* which in turn are accompanied by a rich array of perennial herbs and forbs such as *Antennaria chilensis, Azorella trifoliata, Calandrinia caespitosa, C. acaulis, C. affinis, Cardamine glacialis, Caltha appendiculata, Colobanthus quitensis, Eleocharis albibracteata, Luzula* spp., *Nastanthus caespitosus, Phylloscirpus* spp. In the Patagonian Andes, the bogs tend to be softer and are composed of a mix of perennial herbs, forbs and grasses, such as *Acaena antarctica, Acaena magellanica, Caltha appendiculata, Carex vallis pulchrae, Erigeron leptopetalous, Eudema hauthalii, Hamadryas sempervivoides, Juncus scheuzerioides, Plantago barbata, Primula magellanica, Marsippospermum reichei,* and *Trisetum, Deschampsia Agrostis, Poa,* and *Phleum* spp. On the wetter western Patagonian mountains, hard cushion species more typical of Magellanic moorland (e.g., *Caltha dionaefolia, Oreobolus obtusangulus*), can also be found occasionally in alpine bogs.

Concluding Remarks

In what has been a rapid flight along the Andes the authors have attempted to overview the state of knowledge of plant biodiversity in the high-elevation habitat. This has not been an easy task, as most work pertaining to the high Andes has been carried out by researchers from many different countries, with a tendency for them to work independently and often using different concepts to define and divide the high-elevation habitat. Difficult access to high-elevation areas in the Andes has clearly hindered a more systematic approach to describing biodiversity and much of the available floristic information is not appropriate for more quantitative analyses. Nevertheless, as was seen, significant progress has been made in relation to documenting the composition of local floras, in developing a comprehensive continental-scale plant checklist and sectorial lists, and in our understanding of the evolution and diversification of the high-elevation flora. But clearly, much work remains. Developing a unified scheme to describe altitudinal zonation in the high-elevation belt and determining the real extent of overlap of the páramo, puna, and southern Andean steppe with a bioclimatically defined montane belt are two urgent tasks. Likewise, adopting standardized sampling procedures, such as plot size to compare local species richness along and between altitudinal gradients at different latitudes is necessary. In attempting to develop a comprehensive high Andean plant checklist, the authors have been struck with the poor quality of habitat data on many herbarium specimens, and the rarity of online georeferenced plant specimen databases in the Andean countries, where many herbaria and excellent collections exist. Apart from the need for well-resolved phylogenies for many more plant genera, undoubtedly, one of the most promising areas for future research concerns niche evolution of high-altitude plants in the phylogenetic space. With their outstanding range of climatic conditions, diverse flora, and many morphological, reproductive and ecophysiological adaptations for survival in the harsh high-altitude habitat, the high Andes are well positioned to contribute to fundamental integrative research of this kind. Finally, much of the high-elevation habitat in the Andes is still quite well conserved, but certainly not without threat, for example, by the mining industry.

Acknowledgments

Research support from several Fondecyt-Chile grants, the Chilean Millennium Scientific Initiative (F ICM P 05-02) and the Conicyt Base Financing Program (PFB-23) has enabled much of the Chilean and wider Andean research referred to in this article. Interaction with the GMBA program of DIVERSITAS has been very useful.

Appendix

List of Courses

1. Alpine Ecology
2. Biodiversity
3. Biogeography

See also: Alpine Ecosystems. Biodiversity, Origin of. Biodiversity-Rich Countries. Biogeography, Overview. Ecosystems of South America. Endemism. Global Species Richness. Grassland Ecosystems. High-Temperature Ecosystems. Latitudinal Gradients of Biodiversity. Mediterranean-Climate Ecosystems. Neotropical Seasonally Dry Forests. Phylogeny. Pollinators, Role of. Southern

(Austral) Ecosystems. Species–Area Relationships. Species Distribution Modeling

References

Arroyo MTK, Armesto JJ, and Villagrán C (1981) Plant phenological patterns in the high Andean Cordillera of Central Chile. *Journal of Ecology* 69: 205–233.

Arroyo MTK, Primack RB, and Armesto JJ (1982) Community studies in pollination ecology in the high temperate Andes of Central Chile. I. Pollination mechanisms and altitudinal variation. *American Journal of Botany* 69: 82–97.

Arroyo MTK, Armesto JJ, and Primack R (1983) Tendencias altitudinales y latitudinales en mecanismos de polinización en la zona andina de los Andes templados de Sudamérica. *Revista Chilena de Historia Natural* 56: 159–180.

Arroyo MTK, Marticorena C, and Villagrán C (1984) La flora de la Cordillera de los Andes en el área de Laguna Grande y Laguna Chica, III Región, Chile. *Gayana Botánica* 41: 3–46.

Arroyo MTK, Armesto JJ, and Primack RB (1985) Community studies in pollination ecology in the high temperate Andes of Central Chile. II. Effect of temperature and visitation rates and pollination possibilities. *Plant Systematics and Evolution* 149: 187–203.

Arroyo MTK, Squeo FA, and Lanfranco D (1987) Polinización biótica en los Andes de Chile: Avances hacia una síntesis. In: Forero E, Sarmiento F, and La Rotta C (eds.) *Ecología de la Reproducción e Interacciones Planta/Animal, Vol. 2, Anales del IV Congreso Latinoamericano de Botánica*, pp. 55–76. Editorial Guadalupe: Bogotá.

Arroyo MTK, Squeo FA, Armesto JJ, and Villagrán C (1988) Effects of aridity on plant diversity in the northern Chile Andes: Results of a natural experiment. *Annals of the Missouri Botanical Garden* 75: 55–78.

Arroyo MTK, Marticorena C, Miranda P, Matthei O, Landero A, and Squeo FA (1989) Contribution to the high elevation flora of the Chilean Patagonia: A checklist of species on mountains on an east-west transect in the Sierra de los Baguales, latitude 50°S. *Gayana Botánica* 46: 121–151.

Arroyo MTK and Squeo FA (1990) Relationship between plant breeding systems and pollination. In: Kawano S (ed.) *Biological Approaches and Evolutionary Trends in Plants*, pp. 205–227. London: Academic Press.

Arroyo MTK, von Bohlen CP, Cavieres LA, and Marticorena C (1992) Survey of the alpine flora of Torres del Paine National Park, Chile. *Gayana Botanica* 49: 47–70.

Arroyo MTK, Squeo FA, Veit H, Cavieres LA, Leon P, and Belmonte E (1997) Flora and vegetation of northern Chilean Andes. In: González C (ed.) *El Altiplano: Ciencia y Consciencia en los Andes*, pp. 167–178. Santiago: Vicerrectoría Académica y Estudiantil, Universidad de Chile.

Arroyo MTK, Cavieres LA, Peñaloza A, and Arroyo-Kalin M (2003) Positive associations between the cushion plant *Azorella monantha* (Apiaceae) and alpine plant species in the Chilean Patagonian Andes. *Plant Ecology* 161: 121–129.

Arroyo MTK, *et al.* (2010) A possible correlation between the altitudinal and latitudinal ranges of species in the high elevation flora of the Andes. In: Spehn EM and Körner C (eds.) *Data Mining for Global Trends in Mountain Biodiversity*, pp. 39–47. Boca Raton: CRC Press, Taylor and Francis.

Barrios Y, Ramírez N, Ramírez E, Sánchez E, and del Castillo R (2010) Importancia de los polinizadores en la reproducción de seis especies de subpáramo del Pico Naiguatá (Parque Nacional el Ávila-Venezuela). *Acta Botánica Venezuelica* 33: 213–231.

Beck SG (1985) Florula ecológica de Bolivia: Puna semiárida en el Altiplano Boliviano. *Ecología en Bolivia* 6: 1–41.

Bell CD (2004) Preliminary phylogeny of Valerianaceae (Dipsacales) inferred from nuclear and chloroplast DNA sequence data. *Molecular Phylogenetics and Evolution* 31: 340–350.

Bell CD and Donoghue MJ (2005) Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms, Diversity & Evolution* 5: 147–159.

Berry PE and Calvo RN (1989) Wind pollination, self-incompatibility and altitudinal shifts in pollination systems in the high Andean genus *Espeletia* (Asteraceae). *American Journal of Botany* 75: 1602–1614.

Blisniuk PM, Stern LA, Chamberlain CP, Idleman B, and Zeitler PK (2005) Climatic and ecologic changes during Miocene uplift in the Southern Patagonian Andes. *Earth and Planetary Science Letters* 230: 125–142.

Borgnia M, Maggi A, Arriaga M, Aued B, Vilá BL, and Cassini H (2006) Caracterización de la vegetación en la Reserva de Biósfera Laguna Blanca (Catamarca, Argentina). *Ecología Australia* 16: 29–45.

Bräuchler C, Meimberg H, and Heubl G (2010) Molecular phylogeny of Menthinae (Lamiaceae, Nepetoideae, Mentheae) – Taxonomy, biogeography and conflicts. *Molecular Phylogenetics and Evolution* 55: 501–523.

Briceño B and Morillo G (2006) Catálogo de las plantas con flores de los Páramos de Venezuela: Parte II. Monocotiledóneas (Liliopsida). *Acta Botánica Venezuelica* 29: 89–134.

Cabrera AL (1968) Ecología vegetal de la puna. In: Troll C (ed.) Geo-ecology of the mountainous regions of the tropical Americas. *Colloquium Geographicum* 9: 91–117.

Cabrera AL (1971) Fitogeografía de la República Argentina. *Boletín de la Sociedad Argentina de Botánica* 14: 1–42.

Cavieres LA and Arroyo MTK (1999) Tasa de enfriamiento adiabático del aire en el Valle del Río Molina, Provincia de Santiago, Chile Central (33°S). *Revista Geográfica Terra Australis* 44: 79–86.

Cavieres LA, Peñaloza A, and Arroyo MTK (2000a) Altitudinal vegetation belts in the high Andes of central Chile (33°S). *Revista Chilena de Historia Natural* 73: 331–344.

Cavieres LA, Rada F, García-Nuñez C, Azocar A, and Cabrera HM (2000b) Gas exchange and low temperature resistance in two tropical treeline species from the Venezuelan Andes (8°N). *Acta Oecologica* 21: 203–211.

Cavieres LA, Arroyo MTK, Molina-Montenegro M, Torres C, and Peñaloza A (2002) Nurse effect of *Bolax gummifera* (Apiaceae) cushion plants in the alpine vegetation of the Chilean Patagonian Andes. *Journal of Vegetation Science* 13: 547–554.

Cavieres LA, Badano EI, Sierra-Almeida A, Gómez-González S, and Molina-Montenegro M (2006) Positive interactions between alpine plant species and the nurse cushion plant *Laretia acaulis* do not increase with elevation in the Andes of central Chile. *New Phytologist* 169: 59–69.

Cavieres LA and Badano EI (2009) Do facilitative interactions increase species richness at the entire community level? *Journal of Ecology* 97: 1181–1191.

Chacón J, Madriñan S, Chase MW, and Bruhl JJ (2006) Molecular phylogenetics of *Oreobolus* (Cyperaceae) and the origin and diversification of the American species. *Taxon* 55: 359–366.

Chiapella J and Ezcurra C (1999) La flora del Parque Provincial Tromen, provincia de Neuquén, Argentina. *Multequina* 8: 51–60.

Cleef AM (1981) *The Vegetation of the Páramos of the Colombian Cordillera Oriental*. Dissertationes Botanicae Band 61. J. Cramer.

Colmenares M (2002) *Estudio del crecimiento de Polylepis sericea Wedd. en el páramo venezolano*. Mérida: Tesis Magister, Universidad de Los Andes.

Cuatrecasas J (1968) Paramo vegetation and its life forms. In: Troll C (ed.) *Geo-Ecology of the Mountainous Regions of Tropical Americas, Proceedings of the UNESCO Mexican Symposium*. Berlin: Verlag.

Cuatrecasas J (1986) Speciation and radiation of Espeletiinae in the Andes. In: Vuilleumier F and Monasterio M (eds.) *High Altitude Tropical Biogeography*, pp. 267–303. New York: Oxford University Press.

Cuello NL, Cleef AM, and Aymard. G (2010) Phytogeography of the vascular páramo flora of Ramal de Guaramacal (Andes, Venezuela) and its ties to other páramo floras. *Anales del Jardín Botánico de Madrid* 67: 177–193.

Ferreyra M, Clayton S, and Ezcurra C (1998) La flora altoandina de los sectores este y oeste del Parque Nacional Nahuel Huapi, Argentina. *Darwiniana* 36: 65–79.

Flores M, Alegría J, and Granda A (2005) Diversidad florística asociada a las lagunas andinas Pomacocha y Habascocha, Junín, Peru. *Revista Peruana de Biología* 12: 125–134.

Galán de Mera A, Cáceres C, and González A (2003) La vegetación de la alta montana andina del sur del Perú. *Acta Botanica Malacitana* 28: 121–147.

Gengler-Nowak K (2003) Molecular phylogeny and taxonomy of Malesherbiaceae. *Systematic Botany* 28: 333–344.

Godley and Berry PH (1995) The biology and systematics of *Fuchsia* in the South Pacific. *Annals of the Missouri Botanical Garden* 82: 473–516.

Goldstein G, Meinzer FC, and Rada F (1994) Environmental biology of a tropical treeline species, *Polylepis sericea*. In: Rundel PW, Smith A, and Meinzer FC (eds.) *Tropical Alpine Environments, Plant Form and Function*, pp. 129–149. Cambridge: Cambridge University Press.

Gregory-Wodzicki KM (2000) Uplift history of the Central and Northern Andes: A review. *Geological Society of America Bulletin* 112: 1091–1105.

Gruenstaeudl M, Urtubey E, Jansen RK, Samuel R, Barfuss MH, and Stuessy TF (2009) Phylogeny of Barnadesioideae (Asteraceae) inferred from DNA sequence data and morphology. *Molecular Phylogenetics and Evolution* 51: 572–587.

Hartley AJ (2003) Andean uplift and climate change. *Journal of the Geological Society, London* 160: 7–10.

Hershkovitz MA, Arroyo MTK, Bell C, and Hinojosa LF (2006a) Phylogeny of *Chaetanthera* (Asteraceae: Mutisieae) reveals both ancient and recent origins

of high elevation lineages. *Molecular Phylogenetics and Evolution* 41: 594–605.
Hershkovitz MA, Hernández-Pellicer CC, and Arroyo MTK (2006b) Ribosomal DNA evidence for diversification and hybridization in *Tropaeolum* sect. Chilensia. *Plant Systematics and Evolution* 260: 1–24.
Hijmans RJ, *et al.* (2005) WorldClim. Global Climate Data. http://www.worldclim.org.
Hoch G and Körner C (2005) Growth, demography and carbon relations of *Polylepis* trees at the world's highest treeline. *Functional Ecology* 19: 941–951.
Hoot SB, Kramer J, and Arroyo MTK (2008) Phylogenetic position of the South American dioecious genus Hamadryas and related Ranunculeae (Ranunculaceae). *International Journal of Plant Sciences* 169: 433–443.
Horton BK, Hampton BA, and Waanders GL (2001) Paleogene synorogenic sedimentation in the Altiplano plateau and implications for initial mountain building in the central Andes. *Geological Society of America Bulletin* 113: 1387–1400.
Hughes C and Eastwood R (2006) Island radiation on a continental scale. Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America* 103(27): 10334–10339.
Iribarren ML and Ferreyra M (2011) Flora y vegetación altoandina: Parque Nacional Los Glaciares. Zona norte y áreas vecinas. *Final Report*, Administración de Parques Nacionales, Argentina.
Izco J, Pulgar I, Aguirre Z, and Santin F (2007) Estudio florístico de los páramos de pajonal meridionales de Ecuador. *Revista Peruana de Biología* 14: 237–246.
Jorgenson PM and León-Yánez S (1999) Catalogue of the Vascular Plants of Ecuador. *Monographs in Systematic Botany from the Missouri Botanical Gardens* 75: 1–1169.
Josse C, *et al.* (2009) *Ecosistemas de los Andes del Norte y Centro. Bolivia, Colombia, Ecuador, Perú y Venezuela*. Lima: Secretaría General de la Comunidad Andina.
Kessler M and Schmidt-Lebuhn AN (2006) Taxonomical and distributional notes on Polylepis (Rosaceae). *Organisms, Diversity and Evolution* 6: 1–10.
Körner Ch, Paulsen J, and Spehn EM (2011) A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data. *Alpine Botany* 121: 73–78.
Luebert F and Gajardo R (2000) Vegetación de los Andes áridos del norte de Chile. *Lazaroa* 21: 111–130.
Luteyn JL (1992) Páramos: Why study them? In: Balslev H and Luteyn JL (eds.) *Páramo: An Andean Ecosystem Under Human*, pp. 1–14. London: Academic Press.
Luteyn JL (1999) Páramos – a checklist of plant diversity, geographical distribution and botanical literature. *Memoirs of the New York Botanical Garden*, vol. 84.
Luteyn JL and Churchill SP (2000) Vegetation of the tropical Andes: An overview. In: Lentz DL (ed.) *Imperfect Balance: Landscape Transformations in the Precolumbian Americas*, pp. 281–310. Nueva York: Columbia University Press.
Marcelo Peña LM, Sánchez Vega I, and Millán Tapia JF (2006) Estado actual de la diversidad florística del páramo sectores: el espino y Palambe, Sallique, Jaén. Catamarca. Peru. *Ecología Aplicada* 5: 1–8.
Mark AF, Dickinson KJM, Allen J, Smith R, and West CJ (2001) Vegetation patterns, plant distribution and life forms across the alpine zone in southern Tierra del Fuego, Argentina. *Austral Ecology* 26: 423–440.
Marticorena A, Pardo V, Peñaloza A, Negritto MA, Cavieres LA, and Parada M (2004) Adiciones y notas a la flora del Parque Nacional Llullaillaco, II Región, Chile. *Gayana Botánica* 61(2): 49–54.
Martínez Carretero E (1995) La puna argentina: Delimitación general y divisiones en distritos florísticos. *Boletín de la Sociedad Argentina de Botánica* 31: 27–40.
Meinzer FC, Goldstein G, and Rada F (1994) Páramo microclimate and leaf thermal balance of Andean giant rosette plants. In: Rundel PW, Smith A, and Meinzer FC (eds.) *Tropical Alpine Environments, Plant Form and Function*, pp. 45–59. Cambridge: Cambridge University Press.
Méndez E, Martínez-Carretero E, and Peralta I (2006) La vegetación del Parque Provincial Aconcagua (Altos Andes centrales de Mendoza, Argentina). *Boletín de la Sociedad Argentina de Botánica* 41: 41–69.
Méndez E (2009) Biodiversidad de la flora del flanco oriental del Cordón del Plata (Luján de Cuyo, Mendoza, Argentina). *Catálogo florístico. Boletín de la Sociedad Argentina de Botánica* 44: 75–102.
Meudt HE and Simpson BB (2006) The biogeography of the austral, subalpine genus *Ourisia* (Plantaginaceae) based on molecular phylogenetic evidence: South American origin and dispersal to New Zealand and Tasmania. *Biological Journal of the Linnean Society* 87: 479–513.
Meyer KM, Hoot SB, and Arroyo MTK (2010) Phylogenetic affinities of South American *Anemone* (Ranunculaceae), including the endemic segregate genera, *Barneoudia* and *Oreithales*. *International Journal of Plant Science* 171: 323–331.
Molau U (1988) Scrophulariaceae – Part 1. Calceolarieae. *Flora Neotropica Monograph* 47: 1–326.
Méndez E (2007) La vegetación de los altos Andes II. Las vegas del flanco oriental del Cordón del Plata (Mendoza, Argentina. *Boletín de la Sociedad Argentina de Botánica* 42: 273–294.
Monasterio M (1980) Las formaciones vegetales de los páramos de Venezuela. In: Monasterio M (ed.) *Estudios Ecológicos en los Páramos Andinos*, pp. 93–157. Mérida: Ediciones de la Universidad de los Andes.
Nicolas AN and Plunkett M (2009) The demise of subfamily Hydrocotyloideae (Apiaceae) and the re-alignment of its genera across the entire order Apiales. *Molecular Phylogenetics and Evolution* 53: 134–151.
Novara LJ (2003) Catálogo de la flora de la Puna en noroeste Argentina. *Aportes Botánicos de Salta. Serie Misceláneas* 2(1): 1–56.
O'Leary N, Yuan Y, Chemisquy A, and Olmstead R (2009) Reassignment of species of paraphyletic *Junellia* s.l. to the new genus *Mulguraea* (Verbenaceae) and new circumscription of genus *Junellia*: Molecular and morphological congruence. *Systematic Botany* 34: 777–786.
O'Leary N, Peralta P, and Múlgura ME (2011) Sinopsis del género *Junellia* (Verbenaceae). *Darwinana* 49: 47–75.
Orme AR (2007) The tectonic framework of South America. In: Veblen TT, Young KR, and Orme AR (eds.) *The Physical Geography of South America*, pp. 3–22. Oxford: Oxford University Press.
Quinteros J, Ramos VA and Jacovkis PA (2005) A finite element model for the early to middle Miocene evolution of the Patagonian Andes at 47°S, *Sixth International Symposium on Andean Geodynamics (ISAG 2005, Barcelona)*, Extended Abstracts. pp. 582–585.
Rada F, Goldstein G, Azócar A, and Meinzer FC (1985) Daily and seasonal osmotic changes in a tropical treeline species. *Journal of Experimental Botany* 36: 989–1000.
Rada F, Azócar A, Briceño B, González J, and García-Núñez C (1996) Carbon and water balance in *Polylepis sericea*, a tropical treeline species. *Trees* 10: 218–222.
Rada F, García-Nuñez C, and Rangel S (2009) Low temperature resistance in saplings and ramets of *Polylepis sericea* in the Venezuelan Andes. *Acta Oecologica* 35: 610–613.
Rangel JO (2000a) Clima de la región paramuna en Colombia. In: Rangel JO (ed.) *Colombia Diversidad Biótica III. La Región de Vida Paramuna de Colombia*, pp. 85–125. Bogota: Universidad Nacional de Bogota.
Rangel JO (2000b) Flora. In: Rangel JO (ed.) *Colombia Diversidad Biótica III. La Región de Vida Paramuna de Colombia*, pp. 126–378. Bogota: Universidad Nacional de Bogota.
Rangel JO (2000c) Catálogo florístico de los macizos de Chingaza y Sumapaz. In: Rangel JO (ed.) *Colombia Diversidad Biótica III. La Región de Vida Paramuna de Colombia*, pp. 563–598. Bogota: Universidad Nacional de Bogota.
Ramsay PM and Oxley ERB (1997) The growth form composition of plant communities in the Ecuadorian páramos. *Plant Ecology* 131: 173–192.
Rauscher JT (2002) Molecular phylogenetics of the *Espeletia* complex (Asteraceae): Evidence from nrDNA ITS sequences on the closest relatives of an Andean adaptive radiation. *American Journal of Botany* 89: 1074–1084.
Rivas-Martínez S (1993) Clasificación bioclimática de la Tierra. *Folia Botánica Matritensis* 10: 1–23.
Rundel PW (1994) Tropical alpine climate. In: Rundel PW, Smith A, and Meinzer FC (eds.) *Tropical Alpine Environments, Plant Form and Function*, pp. 21–43. Cambridge: Cambridge University Press.
Rundel PW, *et al.* (2003) Ecological and ecophysiological patterns in a pre-altiplano shrubland of the Andean Cordillera in northern Chile. *Plant Ecology* 169: 179–193.
Ruthsatz B (1974) Los arbustos de las estepas andinas del noroeste argentino y uso actual. *Boletín de la Sociedad Argentina de Botánica* 16: 27–45.
Santibañez F and Uribe JM (1990) *Atlas Agroclimático de Chile. Regiones V y Metropolitana*. Santiago: Facultad de Ciencias Agrarias y Forestales, Universidad de Chile.
Sarmiento G (1986) Ecological features of climate in high-tropical mountains. In: Vuilleumier F and Monasterio M (eds.) *High Altitude Tropical Biogeography*, pp. 11–45. New York: Oxford University Press.
Scherson R, Vidal R, and Sanderson MJ (2008) Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. *American Journal of Botany* 95: 1030–1039.

Schmidt-Lebuhn AN, Kessler M, and Kumar M (2006) Promiscuity in the Andes: Species relationships in *Polylepis* (Rosaceae, Sanguisorbeae) based on AFLP and morphology. *Systematic Botany* 31: 547–559.

Sierra-Almeida A, Cavieres LA, and Bravo LA (2009) Freezing resistance varies within the growing season and with elevation in high Andean species of central Chile. *New Phytologist* 182: 461–469.

Sierra-Almeida A, Cavieres LA, and Bravo LA (2010) Freezing resistance of high-elevation plant species is not related to their height or growth-form in the Central Chilean Andes. *Environmental and Experimental Botany* 69: 273–278.

Sierra-Almeida A and Cavieres LA (2012) Summer freezing resistance of high-elevation plant species changes with ontogeny. *Environmental and Experimental Botany* 80: 10–15.

Simpson BB (1979) A revision of genus *Polylepis* (Rosaceae Sanguisorbae). *Smithsonian Contribution to Botany* 43.

Simpson BB (1983) An historical phytogeography of the high Andean flora. *Revista Chilena de Historia Natural* 56: 109–122.

Simpson BB, Arroyo MTK, Sipe S, de Moraes MD, and McDill J (2009) Phylogeny and evolution of *Perezia* (Asteraceae: Mutisieae: Nassauviinae). *Journal of Systematics and Evolution* 47: 431–443.

Sklenář P and Balslev H (2006) Geographic flora elements in the Ecuadorian superpáramo. *Flora* 202: 50–61.

Sklenář P, Dŭsková E, and Balslev H (2011) Tropical and temperate: Evolutionary history of Páramo flora. *Botanical Review* 77: 71–108.

Sobrevila C (1989) Effects of pollen donors on seed formation in *Espeletia schultzii* (Compositae) populations at different altitudes. *Plant Systematics and Evolution* 166: 45–67.

Soltis D, Kuzoff RK, Mort ME, *et al.* (2001) Elucidating deep-level phylogenetic relationships in Saxifragaceae using sequences for six chloroplastic and nuclear DNA regions. *Annals of the Missouri Botanical Garden* 88: 669–693.

Squeo FA (1991) *Estructura de comunidades vegetales andinas en relación con la polinización en la cordillera de los Baguales, Patagonia, Chile.* Doctoral Thesis, Facultad de Ciencias, Universidad de Chile.

Squeo FA, Veit H, Arancio G, Gutiérrez JR, Arroyo MTK, and Olivares N (1993) Spatial heterogeneity of high mountain vegetation in the Andean desert zone of Chile (30°S) Mountain. *Research and Development* 13: 203–209.

Squeo FA, Arancio G, Osorio R, Arroyo MTK, and Veit H (1994) Flora y vegetación de los Andes desérticos de Chile. In: Squeo R, Osorio R, and Arancio G (eds.) *Flora de Los Andes de Coquimbo: Cordillera de Doña Ana*, pp. 1–17. La Serena: Ediciones Universidad La Serena.

Squeo FA, Rada F, Garćia CE, Ponce ME, Rojas AL, and Azócar A (1996) Cold resistance mechanisms in high desert Andean plants. *Oecologia* 105: 552–555.

Stein BA (1992) Sicklebill hummingbirds, ants, and flowers. *Bioscience* 42: 27–33.

Tate JJ and Simpson BB (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* 28: 723–737.

Teillier S, Hoffmann AJ, Saavedra F, and Pauchard L (1994) Flora del Parque Nacional El Morado (Región Metropolitana, Chile). *Gayana Botanica* 51: 13–47.

Teillier S (2006) Catálogo de las plantas vasculares del área altoandina de Salar de Coposa-Cordon Collaguasi, Chile, Región de Tarapacá (I). Chloris Chilensis. Año 2, No. 1. http://www.chlorischile.cl.

Thomson SN, Brandon MT, Reiners PW, Tomkin JH, Vásquez C, and Wilson NJ (2010) Glaciation as a destructive and constructive control on mountain building. *Nature* 467: 313–317.

Torres H, Hernández L, Weber C, Palma R, and von Buch K (1978) Plan para el manejo del Parque Nacional Lauca. *Publicación Técnica No. 6.*. Santiago: Corporación Nacional Forestal.

Tupayachi AF (2005) Flora de la Cordillera de Vilcanota. *Arnaldoa* 12: 126–144.

van der Hammen T and Cleef AM (1986) Development of the high Andean paramo flora and vegetation. In: Vuilleumier F and Monasterio M (eds.) *High Altitude Tropical Biogeography,* pp. 153–200. New York: Oxford University Press.

von Hagen B and Kadereit JW (2001) The phylogeny of *Gentianella* (Gentianaceae) and its colonization of southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. *Organisms Diversity & Evolution* 1: 61–79.

von Hagen B and Kadereit JW (2003) The diversification of *Halenia* (Gentianaceae): Ecological opportunity versus key innovation. *Evolution* 57: 2507–2518.

Villagrán C, Armesto JJ, and Arroyo MTK (1981) Vegetation in a high Andean transect between Turi and Cerro León in northern Chile. *Vegetatio* 48: 3–16.

Villagrán C, Arroyo MTK, and Marticorena C (1983) Efectos de la desertización en la distribución de la flora andina de Chile. *Revista Chilena de Historia Natural* 56: 137–157.

Wagstaff SJ, Breitweiser I, and Swenson U (2006) Origin and relationships of the austral genus *Abrotanella* (Asteraceae) inferred from DNA sequences. *Taxon* 55: 95–106.

Wardle P (1998) Comparison of alpine timberlines in New Zealand and the Southern Andes. *Royal Society of New Zealand. Miscellaneous Series* 48: 69–90.

Young KR and Cano A (1994) Aporte florístico de la puna del Parque Nacional del Manu, Peru. *Boletín de Lima* 16: 381–393.

Young KR, León B, Jørgensen PM, and Ulloa C (2007) Tropical and subtropical landscapes of the Andes. In: Veblen TT, Young KR, and Orme AR (eds.) *The Physical Geography of South America*, pp. 200–216. Oxford: Oxford University Press.

Záveská Drábková L and Vlček Č (2009) Mitochondrial DNA variation within Juncaceae: Comparison of impact of organelles regions on phylogeny. *Plant Systematics and Evolution* 278: 169–186.

High-Temperature Ecosystems

Richard G Wiegert, University of Georgia, Athens, GA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 349–361,
© 2001, Elsevier Inc.

Glossary

Acidophilic Organisms preferring or requiring a low pH environment.

Chemoautotrophs Organisms able to synthesize organic compounds by the oxidation of energy-rich inorganic sources. Light is not used.

Cyanobacteria A group of bacteria containing chlorophyll and capable of photosynthesis.

Eucaryotes Organisms possessing a defined cell nucleus and nuclear membrane. Includes all organisms above the level of the procaryotic bacteria.

Frustule The hard, silica-containing skeleton of diatoms (green algae).

Geothermal heating Water heated at depth in the earth and released to the surface as thermal outflows.

Heterotrophs Organisms dependent for their energy on organic compounds. If oxidized, the process is aerobic. When oxygen is absent, the process is anaerobic (fermentation).

Metazoa Organisms whose individuals have more than one cell (multicellular).

Photoautotrophs Organisms able to synthesize organic compounds from water and inorganic nutrients, using the energy in photons of light.

Procaryotes Organisms without a defined cell nucleus or nuclear membrane.

Protozoa A group of eucaryotic organisms usually classified as animals and regarded as primitive. But some major problems exist with some protozoa such as slime molds and euglenoid forms, which share many of the properties of plants.

Solar heated Water raised to temperatures significantly exceeding the regional temperatures of lakes and streams. Many of the thermally tolerant and thermal-opportunist groups are thought to have evolved in shallow solar-heated ponds and water margins.

Thermal opportunists Organisms that have developed life history characteristics that permit the exploitation of temporary cooler spots in thermal systems.

Thermal systems Outflows of geothermally heated water, usually thought of in terms of those at the surface of the earth (see *thermal vent*) where light is present.

Thermal vent Geothermally heated water issuing from cracks in the ocean floor. Very high temperatures and pressures and the complete absence of light are characteristic.

Thermally tolerant Organisms able to live at temperatures significantly higher than the regional norm. They can also exist at cooler temperatures, but may not compete well in the latter situation.

Thermophily Often restricted to mean those organisms adapted to living at high temperatures and unable to survive at lower temperatures. In this chapter the term is also used in a more general way to characterize any type of adaptation to thermal systems.

Extreme Environments

The diversity of organisms in thermal environments is determined, in common with all biodiversity, by the interactions of the physical/biotic environment with the adaptive abilities of individual species populations. But in some respects thermal environments (ecosystems) are quite different, even unique, compared to ecosystems dominated by weather, geological origins of soils, and patterns of precipitation and drainage. First of all, thermal ecosystems form a class of extreme environments characterized by a mean temperature markedly higher than "normal." (Other extreme environments possess one more environmental variable far outside the "norm"—for example, arctic, desert, brine lakes, low pH systems, etc.) Some of these, such as the arctic/antarctic frozen wastes and extreme deserts, are caused by widespread climatic factors and are typically very large, whereas thermal ecosystems and some other extreme environments are small and result from local changes in the physical environment.

What is "normal" and what effect does the extreme factor or factors have on biodiversity? Here I define an extreme environment as "a condition under which some organisms can grow, whereas others cannot." This is too general for the purpose of this chapter, because each different ecosystem has its adapted flora and fauna. A more narrow definition identifies the extreme as excluding entire higher taxonomic groups, not simply species. Thus the extreme environment is not only low in the diversity of species, but also low in the diversity of higher orders of classification as well.

Thermal Systems—General Characteristics and Definitions

Among extreme environments, thermal ecosystems are defined as those with an elevated temperature, compared with the range of other ecosystems found at the same latitude and elevation. The source of the heat may be (a) solar energy, (b) anthropogenic activities, or (c) geothermal. The organisms inhabiting these heated ecosystems can be classified as thermophilic, thermal tolerant, or thermally opportunistic. The first of these, thermophiles (the general term used in this chapter),

require the high temperature for survival to which they are adapted optimally. Thermally tolerant organisms have evolved mechanisms to enable them to exist at extreme temperatures, but they can also survive at lower temperatures. Indeed, in culture they might do better physiologically at these lower temperatures. Competition with other nonthermal species may, however, restrict their distribution in nature. Thermal opportunists, on the other hand, have evolved mechanisms to avoid the effects of direct immersion in hot water. They seek out and take advantage of temporary refugia by using rapid maturation times, rapid mobility, and short life histories. I will later discuss the evolution of these traits and the groups to which they apply.

Solar heating of vascular plant leaves and of the soil/litter surface is enhanced by dark coloration, abundance of sunlight, and the slope of the land surface. Temperatures of plant leaves in desert environments may reach 60 °C; the same temperature applies to soil/litter systems, but it is commonly much lower, and the soil litter surface cools rapidly at night and during rains. Furthermore, the soil/litter surface (as well as plant leaves) heated by the sun is very thin because of its insulating value, and the system is very dry when heated. Thus motile organisms can move out of these thermal systems during periods of inhospitable temperature regimes. Vascular plants can avoid the effects of heat by being rooted in the cooler subsurface soil, and nonmotile microorganisms resist the thermal effects by forming spores or other heat-resistant quiescent stages. The net result is that there is little effect of solar heating on the biodiversity of these ecosystems compared to adjacent systems with more shading of the soil surface. There may, of course, be a loss of species in the driest of these systems, but as a consequence of lack of water, not of temperature per se. Solar heating may also act on very shallow aquatic systems; the margins of lakes and ponds in temperate and tropical latitudes are examples. This can provide a habitat for certain multicelled animals adapted to warm waters, but it does not exclude animals that simply avoid the shallows during the day and forage there at night. Temperature increases are seldom very great, and biodiversity in these systems is affected very little. A case involving a thermally adapted dragonfly will be discussed later in this chapter.

Anthropogenic activities include steam vents, hot water plumbing and heating systems, and fermentation. Where liquid water condenses around steam vents, certain thermal microorganisms can be found and some of them can also be found in hot water systems. But these are too small and transient to be of interest to the student of thermal biodiversity. Fermentation, however, is a large natural phenomenon, as well as the result of certain human activities, such as composting and piling up tailings from coal mines. As a natural process, the accumulation of fermenting organic matter is seldom large enough or accumulating fast enough for the heat of fermentation to significantly increase the temperature. Plant debris added to the soil litter system during the year is simply decomposed at a relatively constant rate and the heat produced is dissipated to the environment. Humans, however, often pile up organic wastes to a degree that the insulating value of the pile causes a rise in the internal temperature, killing, for example, the seeds of vascular plants in compost heaps and excluding all but microorganisms from the water draining from coal tailing piles. However, the latter effect is often caused more by the extremely low pH of such systems than from elevated temperatures. Although these environments are of little interest from the standpoint of biodiversity, they are of great interest to the student of microbial taxonomy. For example, an entire genus, *Thermoplasma*, is restricted to coal refuse piles.

Geothermally Heated Systems-Classification and Description

The extreme ecosystems characterized by some form of geothermal heating of water are small, both in absolute size (some are only a few square meters or less in extent) and in relation to the surrounding regional climate and soil-dominated areas. Other extreme environments caused by aridity, extreme salinity, and low temperatures stress organisms by posing problems of getting or retaining water. In contrast, geothermal systems have water, indeed they are defined on the basis of hot water. The problems they pose for organisms attempting to survive and grow involve adaptations for tolerating or avoiding the hot water. Biotic diversity is typically low, both in absolute numbers of species and diversity of higher taxonomic groups. This diversity is also usually low relative to the environments surrounding the thermal systems, unless the latter are situated within a larger extreme environment. The deep-sea thermal vents are an example of the latter.

Most thermal ecosystems of the world are caused by the surface emergence of geothermally heated water. The mechanisms of heating involve the percolation of surface water to some depth, where it is heated and pushed again to the surface. The differences in thermal and nutrient characteristics of the emerging water depend on the substrate rocks it saturates after being heated and the length of time the water travels just under the surface before emerging in heated lakes, pools, streams, and spring seeps. Thermal systems are usefully catalogued into four types. These take the form of (a) heated pools or small lakes, (b) small streams, and (c) outflows from thermal springs. These first three types all exist under one atmosphere of pressure or less and the highest temperature found is 100 °C, the boiling temperature at sea level. A fourth type (d) the thermal vent communities form when heated water is ejected at depth on the ocean floor, and temperatures in the hundreds of degrees centigrade are possible. These were first discovered in 1977; they are small, usually less than a few hundred square meters, and may vent water (black smokers) that reaches 450 °C.

The remainder of this chapter focuses on the various manifestations of these heated ecosystems. All of them share the characteristic of being aquatic, but they show differences related to the volume and movement of water flowing out of the ground as well as the mechanisms available to organisms for adapting to the thermal environment. If temperature is regarded as analogous to a toxin, an organism can exist by (a) adapting to exclude the factor (in the case of temperature, this amounts to avoiding lethal levels for a particular stage in the life history), (b) adapting its structure/life history/physiology

to nullify the effects of the factor, or (c) simply adapt in more minor ways to live with the factor. The second and third strategies differ quantitatively but not qualitatively. For example, enzymes can be changed to different temperature optima by relatively small evolutionary steps, but to cope with temperatures that will destroy (denature) the enzyme, a major new structure must be developed.

Geothermally heated lakes and pools characteristically are well mixed thermally, with little or no directional current. This produces, typically, a body of water with a relatively homogeneous temperature regime. The consequences for organisms attempting to colonize such ecosystems is that they cannot escape the thermal load, thus any organisms with limited motility must adapt to survival at the temperature regime of the system. Algal/bacterial mats in such thermal ecosystems, at temperatures where photosynthesis is possible (discussed later), form on the bottom of the lake or pool. There is little or no thermal gradient in these systems, so any eucaryotic consumers must adapt to the prevailing temperature regime or fail to colonize. Thus in Yellowstone National Park in the United States and in other geothermally active areas of the world, the thermal lakes and pools can be ranked from low to higher diversity as an inverse function of the mean maximum annual temperature of the system. Because of the lack of strong currents and replacement of surface water, cooling of these systems is relatively slow and the annual variation in temperature is low, even in the cold temperate latitude and high elevation of Yellowstone Park.

Thermal streams share with thermal lakes and pools the characteristic that the layer of productive organic photo- or chemoautotrophs is on the bottom and organisms attempting to colonize must adapt to tolerate the temperature regime. But the stream has a significant unidirectional flow and hot water emerging from the substrate onto the surface immediately begins to cool. The result is the establishment of a strong thermal gradient from the source downstream, within which different specific conditions for community development can be found. In theory, the cooling curve will be exponential, but other factors intervene, among which the most important are the current turbulence and the temperature, humidity, and the wind speed immediately above the surface of the stream. For example, in a small Yellowstone thermal stream (Gentian Stream, Firehole Lake Drive), I have found that the temperature in summer can fluctuate rapidly as the sun goes behind clouds and wind velocity changes. This effect takes place in minutes, whereas longer lasting weather changes (cloudy days, for example) cause changes in hours to days. In Yellowstone, 250 m from the source, the annual fluctuation in temperature is approximately 20 °C, significant, but also a testimony to the thermal capacity of water since the flow of this small stream is measured in liters per minute. Thus in contrast to the organisms in thermal lakes and pools, the inhabitants of geothermally heated streams are far from living in a constant temperature natural chemostat, unless the stream habitat is very close to the source.

Outflows from thermal springs differ from lakes, pools, and streams in the volume of water discharged, its depth and pattern of flow after discharge, and in the types of communities and organisms that colonize the outflows. In contrast to lakes and pools, spring discharges have strong flows and turbulent mixing, with a relatively steep thermal gradient established from the source to the point where temperature of the water approximates that of ambient air. Outflows do not generally follow a well-defined channel (except near some of the sources), and in any case the depth is shallow. Once the temperature reaches the point that is tolerated by the filamentous cyanobacteria (blue green algae), a mat of microorganisms forms that can thicken and directly affect the flow of the water. This creates a community with distinct cool patches; some are cool on top with hot water flowing underneath, others are devoid of flow. This temperature heterogeneity provides an additional manner in which motile organisms can colonize thermal communities, by avoiding lethal temperatures as a consequence of adapting physiology, life history, or behavior to take advantage of the temporary cooling of the algal mat. Spring outflows exhibit considerable variety. The sources are of different temperatures, primarily a consequence of the water traveling for variable distances in the soil after emergence from the underlying rock and before emergence onto the surface. The chemical characteristics of the emerging waters may differ (although I am eliminating from this discussion those waters with abnormally low pH, caused by excessive sulfur content). Finally, spring outflows are often intermittent, creating yet another form of temperature heterogeneity. These intermittent flow communities vary greatly in their biotic composition, depending not only on the temperature gradient but also on the period and volume of flow, the chemical composition of the source water, and so on. The biodiversity found at temperatures less than 40 °C is, however, low compared to the biodiversity of similar temperature zones in a relatively constant thermal gradient.

Thermal vent communities are formed by the emission of superheated water from fissures in the ocean floor that form along the line marking the meeting of tectonic plates. Because of the pressure of deep water, temperatures are possible that greatly exceed the 100 °C temperatures of water boiling at sea level. Unfortunately, the great depths at which these vent communities form makes their study both difficult and extremely expensive. In general outline, the water is colonized by strains of high temperature heterotrophic bacteria utilizing the organic compounds dissolved in the superheated water. A strong thermal gradient is set up where the heated water meets the cold seawater and several groups of filter-feeding marine invertebrates have evolved to utilize these resources. Whether this evolution involves tolerance to high temperatures, however, is problematic, since samples and temperatures are hard to match up under such difficult sampling conditions. Furthermore, the thermal gradient is so steep that small distances may see a radical change in mean temperature. Numbers of individuals and biomass of the invertebrate consumers are large, because of the richness of the production by heterotrophic microorganisms, but number of species is low, although perhaps similar or even higher than that of the cold, dark, relatively sterile community away from the thermal vents at these depths. Although the microorganisms are clearly living at temperatures far higher than any found at the surface of the earth, the question of whether multicelled eucaryotes have evolved to survive at temperatures above 40 to 50 °C is still an open question.

Temperature Limits to Life

Extreme environments were characterized earlier by the absence of species belonging to large systematic groups. In thermal ecosystems, as noted earlier, not only are these large systematic groups not found, but within groups, species numbers are low, compared with cool water communities. In general, in terms of temperature tolerance, there is a progression from more "primitive" to higher taxonomic groups, and from small to large individual size. Later I consider the arguments for the evolution of the various temperature tolerance ranges. What are the observed limits, based on sampling and observation of thermal ecosystems?

Four different groups of organisms need to be considered: (a) the heterotrophic procaryotic bacteria, (b) the procaryotic photoautotrophs, (c) the eucaryotic microorganisms, both heterotrophic and photoautotrophic, and (d) the eucaryotic metazoans. Much of the literature prior to 1978 on these limits has been summarized. The study of photosynthetically active thermally tolerant organisms really began with the pioneering work of W. A. Setchell. The main body of his work was never published, but it exists as a 215-page manuscript in the archives of the University of California at Berkeley. A brief summary of the temperature limits portion of this work was published in *Science*. Setchell reported the upper temperature for "algae" as 75 to 77 °C, and for bacteria to be 89 °C. A later investigator claimed, on the basis of superficial samples in Yellowstone, that organisms could not grow above 73 °C, but this work relied on the uptake of radioactive phosphorous as the indicator of life; what was very likely being measured was the upper temperature for photosynthesis rather than the upper temperature for life. In other older records of maximum temperatures, unfortunately, some observers confused blue-green algae, procaryotic cyanobacteria, with green algae, eucaryotes. Within the procaryotes, there has also been confusion about cyanobacteria versus filamentous bacteria. Furthermore, it was not realized how steep the thermal gradient could be, thus putting a premium on temperature measurements at precisely the point where organisms are growing in the field. More recently, the discovery of ocean thermal vents (discussed earlier) has reopened some of the controversy regarding these limits.

Heterotrophic procaryotes have colonized habitats at all temperatures up to the boiling point of water (100 °C) at sea level. The most thermophilic of these organisms have not been cultured, but they are easily sampled in the parent thermal communities using glass slides and simple photomicroscopy. On the basis of these findings, the prediction is that life is possible at any temperature at which there is liquid water. This prediction has now been verified by the preliminary explorations of the thermal vent communities on the ocean floor (discussed earlier) with microorganisms living at temperatures far in excess of 100 °C. However, reports of organisms living in water to 300 °C are still provoking argument. As noted earlier, the most extremely thermophilic procaryotes have not even been cultured. This may be in part due to insufficient knowledge about the nutritional requirements of these heterotrophs, but it might also be due to the difficulty of maintaining cultures near the boiling point of water. The problem is exacerbated in the case of thermal vent microorganisms growing at both extreme temperatures and pressures. Many of these species are also obligate anaerobes, for which even extremely low concentrations of oxygen are poisonous. Others are endosymbionts, which are also notoriously difficult to culture. Just getting the samples to the surface in a viable condition is a problem, and special culture techniques are required for the methanogens. In summary, thermophilic heterotrophic procaryotes are apparently not limited by temperature but by the presence of liquid water. At the surface this is approximately 100 °C (depending on altitude), but at depth in the ocean life is found at substantially higher temperatures, limited by suitable nutrients, pH, and energy sources.

The temperature limit for procaryotic photosynthetic autotrophs (both cyanobacteria and photosynthetic bacteria) is substantially below 100 °C on the surface. All of the observations and experimental evidence to date suggests a maximum limit of 73 to 74 °C. Of this group the most thermophilic are the photosynthetic procaryotes, of which the most temperature tolerant is the single-celled species of cyanobacterium (*Synechococcus lividus*) and the filamentous photosynthetic bacterium (*Chloroflexus aurantiacus*). These species are found in nature at temperatures up to 74 °C, although the optimum temperature is 63 to 67 °C for the cyanobacteria and even lower, about 55 °C, for the filamentous bacterium. In general, the filamentous forms of the photosynthetic procaryotes seem to have adapted to much lower optimal temperatures than the single-celled forms such as *Synechococcus* sp.

Eucaryotic microorganisms are found in nature at substantially lower temperatures than are procaryotes. Different reactions and responses to thermal environments are found in (a) fungi, (b) eucaryotic algae, (c) protozoa, and (d) metazoans, including invertebrates, vascular plants, and, vertebrates.

Fungi: In general, thermophilic and thermotolerant fungi are found at temperatures lower than 60 °C and occur at this temperature only in acid thermal waters. In the very common and widespread "alkaline" hot spring communities, where the pH is initially somewhat acid because of dissolved carbon dioxide but rapidly rises when exposed to the air, free-living filamentous fungi are absent, even at temperatures below 40 °C. There are reports of fungi parasitic on cyanobacteria that would probably be found at somewhat higher temperatures, but these have not been found in the "alkaline" thermal effluents of Yellowstone Park and are, in any case, poorly known (although there is a remarkable thermal range of the disease-causing fungus, *Dactylaria gallopava*).

Single-celled and filamentous green algae as well as diatoms become abundant at temperatures around 40 °C and below. Earlier reports in the literature of diatoms growing at higher temperatures relied only on the recovery of the resistant siliceous frustules, without demonstrating that the cells were alive or growing. The maximum temperatures for which diatoms can be proved to be surviving and growing is 43 to 44 °C. This particular diatom, *Achanthes exigua*, has an optimum temperature of 40 °C. The single-celled and filamentous green algae have very similar temperature limits to those of the diatoms. Below 40 °C in the "alkaline" thermal outflows of Yellowstone, the diversity of algal species rises (discussed in a later section). Although some mats of filamentous green algae

may be found in these communities on top of mats of filamentous bacteria and cyanobacteria, where the underlying water is much hotter than 40 °C, this is because the green algae is growing on top of the mat and is not exposed to the higher temperature water. This underscores the care needed when measuring the temperature of the environment of suspected thermally tolerant or thermophilic microorganisms. Strong gradients can occur in millimeters between where temperatures are measured and where the organisms are found. An interesting and important exception to the restriction of the eucaryotic algae to temperatures in the range of 40 °C is that of the acido-philic single-celled asexual green alga, *Cyanidium caldarium*. This organism has been studied intensively in Yellowstone Park acid hot spring effluents. The upper temperature limit is 55 to 57 °C, and the organism can be found down to 35 °C. The high temperature adaptation of this acidophilic eucaryote has been explained as a simple matter of no other competition, but the story may be more complicated than that (see next section on the evolution of thermophily).

Protozoa have not been given the attention of other eucaryotes in terms of their occurrence in thermal environments. Determining the temperature limits of this group is difficult because of erroneous reports in the older literature reporting the growth and survival of protozoa at high temperatures. Protozoa will grow readily at temperatures up to 45 °C. Other studies have reported somewhat higher temperatures (57 to 58 °C seems to be the maximum), but the majority of protozoa in the natural thermal systems are found at lower temperatures. I have not found living protozoa in the "alkaline" thermal outflows of Yellowstone at temperatures greater than 40 to 43 °C

Invertebrates are ubiquitous in and around thermal outflows throughout the world. In surface thermal features many are winged and thus widely distributed. The same genera and species are found in widely separated thermal ecosystems. Some species, such as the water mites, have a juvenile stage parasitic on winged adult insects and are also widely distributed. Those invertebrates that are neither winged nor parasitic are more local in the distribution of species. Most invertebrate inhabitants of thermal ecosystems are motile, thus enabling them to choose, within limits, the space and temperature optimal to growth and survival. Those that are not motile, or only minimally so, such as many of the deep-sea thermal vent invertebrates and the eggs and juveniles of many winged insects are subject to mass mortality when major shifts in current and in the thermal gradient occur. To date, no invertebrates have been found surviving and growing at temperatures greater than the 50 °C claimed for the ostracod, *Potamocypris*. The thermal limits of many of the most abundant and important invertebrate inhabitants of thermal ecosystems are in fact often quite low (well below 40 °C). They exist by quickly colonizing temporary cool spots in the algal mats and growing into motile adults before the thermal environment becomes hot again. These are typical examples of thermal opportunists.

Vascular plants, being rooted in place, can colonize thermal waters at their maximum temperature only if the temperature is extremely stable. Thus there has been little study of their maximum tolerances. Species of *juncus* grow in thermal streams in Yellowstone Park up to about 40 °C and are often so abundant that they block stream flow and cause minor flooding of adjacent soil. Many species of vascular plant grow along the sides of the streams, one of the more conspicuous is the yellow monkey flower (*Mimulus guttatta*). It is difficult to say to what degree this plant is thermal tolerant because it does not grow directly in the hot water; its restriction in thermal basins to the sides of streams may simply reflect the availability of water, since the soil (sinter) beside thermal outflows is very dry. The species is found in wet swampy areas throughout Yellowstone Park and shows a wide range in size, being smallest in the hottest and least fertile environments.

Vertebrates, even more than vascular plants, are absent from thermal environments. One fish, the desert pupfish (*Cyprinodon*) is found in nonthermal (solar-heated) streams of Death Valley, California. Here it tolerates temperatures up to 43.5 °C. But geothermal outflows are too hot, too small, too ephemeral, and vary too much in temperature and flow to support populations of even the small vertebrates. Reports of fish living in very hot water suffer from the same problem of most of the early observations of thermal maxima in organisms—namely, the means of temperature measurement were not sufficient to measure small gradients. Often the natural gradient of the water was such that the fish swam quickly from cool water into warm water to feed and then retreated to their cool refuges. For example, in Yellowstone Park, the many nutrients, particularly phosphorous, added to large, cold water rivers such as the Firehole, create exceptionally fertile and productive rivers. But these inflows also raise the temperature to the point that in summer they create a naturally thermally polluted river. In summer the lower Firehole river often increases in temperature far above the optimum for trout. For example, above the geyser basins in the summer of 1952, the average temperature was 10.2 °C, whereas below the geyser outflows the temperature was 18.7 °C, despite the influx of cold water from nonthermal springs and streams. In September of 1962 the difference was even more marked, 10.6 °C above and 20.0 °C below. That fish are able to not only survive but prosper in this situation is due to the many cold spring and streamside inflows entering the river, where the trout spend the majority of their time, venturing into the warmer currents of the river only to feed for short periods.

The Evolution of Thermophily

Thermophily in general must be thought of as the adaptation of organisms to survive and grow at temperatures (water as used in this chapter) significantly higher than the range of temperature in nonheated waters of the geographic region or climate. I have defined the range of thermal heating and the different types of ecosystem developing from the heated outflows. The types of organisms and their maximal and, where available, their optimum temperatures were discussed earlier. In general, the number of species and the diversity of the higher taxonomic classifications is lowest at the highest temperatures and increases with temperature decrease. The absolute upper temperature for life seems to be set only by the temperature at which liquid water is present. But if heterotrophic procaryotes can evolve to live and grow at these

highest temperatures, why do not the other microorganisms and the metazoans also evolve to exploit higher temperatures? In one of the first (and the most detailed) discussions of the evolution of thermophily, the early arguments about this issue are reviewed and new hypotheses presented. Prior to the mid-1970s, the question of thermophily was regarded primarily as a physiological question, not an ecological one. This despite a wealth of evidence from other types of extreme environments that showed that adaptation to the extreme characteristic did not only have to be physiological, but that competition with other organisms was an important factor.

Several earlier authors regarded the increasing temperature tolerance of metazoa to eucaryotic algae to cyanobacteria (plus fungi) to heterotrophic bacteria as being due primarily to the increasing complexity of the cellular structure. In the fungal eucaryotes and for the eucaryotic green alga, *Cyanidium*, both of which reach 60 °C as the upper maximum, the situation is complicated by the fact that they grow only in low pH hot systems. For the most part the arguments are simple statements that the cell structure has some fundamental property that cannot be overcome by further adaptation. This postulates not only that the fundamental limit exists, but that there are positive selective pressures for the further adaptation to high temperatures. To date there is little evidence for the idea that temperature tolerance is a simple function of cell complexity, nor has the locus of the effect been identified. At first it was thought that the deep sea invertebrates were extraordinarily tolerant of high temperature, and this might have stimulated research into the cellular question. However, recent evidence points to rather low temperature tolerances in these groups. Clearly, no animals have been found living in the "black smokers" where temperatures of 400 °C or more are common. White smokers at about 50 °C may have some fauna, but the majority of the vent fauna is distributed around the vents, in much cooler water at 40 °C or less. Since these temperatures are far above the ambient temperatures of the deep sea, the fauna is, in the sense denned earlier, thermophilic, but will do little to settle the question of why eucaryotes have not evolved to tolerate temperatures endured by thermophilic procaryotes. The essential test of the cell complexity-membrane postulate was to separate the question into: (a) the evolutionary advantage to be achieved above a maximum limit of 62 °C and (b) the evidence that the various eucaryotic genomes have the capacity to respond to the selective pressures if they exist.

Following the preceding argument, the reasoning is that adaptation of eucaryotes to high temperatures could be prevented by some inherent physicochemical limitation or by the lack of benefit in additional resources gained. Two tactics are possible: (a) to shift the range of temperature tolerance at the upper level (the maximum) independently of the lower range and (b) to shift the entire range of temperature tolerance upward. Thermal outflows vary considerably in width, depth, and flow volumes. But the nonlinear cooling process (more rapid at higher temperature differences between water and air/substrate) ensures that the total area of colonization space, and thus of resources (whether solar energy or dissolved nutrients), will become steadily smaller as one proceeds, for example, in 5 °C increments, down the thermal gradient. Thus organisms that can adapt to higher temperatures by moving only the upper tolerance limit, keeping efficiencies at lower temperatures constant, will always have a positive selective pressure for doing so. If, however, the range must remain the same, shifting it upward will result in an overall loss of resource space. In the latter case, there will never be a positive advantage to be accrued in shifting the range unless other factors are involved, such as predation, competition, or the thermal range of a required food species. Therefore, any general explanation of the evolution of temperature tolerance by themophiles must consider not only a possible physicochemical limitation, but also whether resources are present and the relative amounts of resources at different temperatures. The first of these parameters, physicochemical limits due to cellular differences, has been the focus of most of the early literature. The presence of abundant resources for the various forms of eucaryotic organisms found in and around thermal ecosystems has been well established, to the point where it cannot be invoked as a sole or even a major block to the adaptation to higher temperatures. The invertebrate grazers of the algal mat are restricted to lower temperatures than their food organisms. In the case of photosynthetic eucaryotes, there is the same amount of solar energy per square meter at high temperatures and nutrients are at least as abundant as at low temperatures, if not more so. Thus the question devolves to physicochemical limits or amount of space and therefore of resources. To examine the effect of resource availability (read area) the two tactics discussed above are used. Only populations that are genetically homogeneous or whose genotypes are distributed by random mating fit the criteria. Asexual procaryotes do not fit because changes in genotypic abundance are the result of rates of growth and death. New genotypes arise by mutation and some exchange of genetic material, but the extent of the latter was not known in the 1970s and to my knowledge is not known today in thermophiles.

Because water begins to cool once it emerges from depth into the open air, surface thermal systems typically have the highest temperature water in spring-fed pools with relatively little current and little day-to-day or even annual variation in temperature. At temperatures above the (approximate) 75 °C limit for photosynthetic autotrophs, these pools are dominated by monotypic or very low diversity heterotrophic procaryotes and the pools, although small in absolute area, are virtually infinite in size compared to the size of the bacteria colonizing them. These bacteria are often attached rods or sometimes filaments. Hyperthermophilic bacteria also occur in those thermal outflows where temperatures are 80 °C or above. Because of the current, these forms must be attached rods or filaments. Furthermore, the colonies in the hot outflow channels decrease rapidly downstream. Whether this is because of the decrease in downstream temperature or the development of nutrient shadows is still an open question. Experiments in the field made by enriching these channels with nutrients suggest that the latter may be a more reasonable explanation than the former.

The dominant organism in the high temperature outflows, *Thermus aquaticus*, is an obligate aerobe that forms attached filamentous colonies at temperatures up to 80 °C. Discovered and named in 1969, it now appears to be a complex of several different bacterial types. If the attached bacteria in very hot

outflows do show changes in growth and viability with temperature, it might appear as a broadened range of temperature tolerance. However, this could also be explained as the successive downstream colonization of different genotypes, each adapted to a different optimal temperature. Alternatively, if competition for a scarce nutrient presents a major restriction on distribution, low temperature tolerance will be sacrificed to gain higher temperature tolerance. In this case successive adaptations to higher temperature result in a gain in resources providing the new genotype can maintain a sufficient level of efficiency to gain an advantage over its lower temperature-adapted competitor. Since I have suggested that the apparent temperature response could, in fact, be a response to nutrient depletion, the whole question is still an open one. These colonial forms can be cultured in the laboratory and also can be easily used in field experiments. Some efforts to solve the question of nutrient versus temperature would be worthwhile. In contrast, the extreme thermophilic bacteria that exist in the source pools at temperatures up to the boiling point are very refractory to culture. Thus studying their genetics and measuring their response to temperature clines has been impossible.

In the procaryotic photosynthetic microorganism, the evolution of temperature tolerance was operating in the same way, that is, the adaptation of successive genotypes to higher and higher temperatures, where there is a gain in resources, should be observed. Instead, we observe the complete failure of these organisms to colonize water at temperatures greater than 73 to 75 °C (see discussion on limits presented earlier). The two major resources of the photosynthetic cyanobacteria and the flexibacteria are sunlight and a carbon source (phosphate is abundant in these systems and the organisms fix nitrogen). Cyanobacteria in particular respond positively to increases in free carbon dioxide. Sunlight is intense at the Yellowstone altitude in all shallow-water systems, obviously independent of temperature. Sunlight only becomes limited deep in thick algal mats, thus expanding the thermal range upward by the development of higher-temperature clones will neither increase or decrease the availability of this factor to thin new mats of cyanobacteria. In the "alkaline" thermal flows however, the water typically emerges at a pH in the 6 to 6.5 range due entirely to large concentrations of dissolved carbon dioxide. Therefore, any organism capable of forming genetically homogeneous clones should be able to adapt to higher and higher temperatures by gaining resources (carbon dioxide) at hotter temperatures. Such clones are known for both of the most common genera of cyanobacteria in thermal springs, *Mastigocladus* and *Synechococcus*.

The best evidence for an inherent physicochemical limit on the temperature of photosynthesis would be if the previous conditions were met with respect to resources. But adaptation proved unable to extend the range of temperature tolerance without a significant decrease in the ability or efficiency to use the increased abundance of the resource at higher and higher temperatures. This is exactly the case for *Synechococcus*. When growth (doubling time) is measured as a function of temperature, it is five times longer at 70 °C than at 40 °C. Thus adaptation to higher temperatures results in a rapidly decreasing ability to process a limiting nutrient, even though it (presumably carbon dioxide) is more concentrated at the higher temperatures. Thus the failure to evolve tolerances above 73 to 75 °C must be due to a temperature dependent reduction in growth rate or a physicochemical failure in the algal cell.

Eucaryotic microorganisms, including single-celled and filamentous forms plus diatoms and free-living fungi, are seldom found at temperatures above the range of 40 to 45 °C, with the exception of some fungi and the acidophilic genus *Cyanidium*. Collections made from six different hot springs in Montana showed a clear and significant separation between the eucaryotic algae and the cyanobacteria; the former were never collected at temperatures greater than 50 °C and the latter showed more species per sample (2 to 5 at 50 to 58 °C, 10 at 35 °C) than the former. But the diversity of the eucaryotic algae was increasing much faster than the cyanobacteria with decreasing temperature. This separation was present in the thermal springs of both Yellowstone and Mount Rainier National Parks. The eucaryotic alga, *Cyanidium*, which exists at much higher temperatures than the other eucaryotic algae, may be subject to quite different selection pressures since it also occupies a variety of adjacent habitats (wet soil) where the temperatures fluctuate greatly, as do the temperatures in the outflows, which are very cool during snowmelt in the spring and the alga disappears, only to reappear in the water when the temperature rises. The argument is made that the evolution of thermophily in this group must take into account these differences plus the fact that hot acid flows have no procaryotic photoautotrophs that could offer competition to *Cyanidium*. The possibility of genetic mixing due to sexual selection in the eucaryotic algae may, together with the increasing area for colonization as the thermal gradient is descended, render the selective optimum temperature lower in eucaryotes than in sets of competing procaryotic clones. This idea needs testing along with the notion of a major physicochemical limit on the eucaryotic cell. The growth of *Cyanidium* at higher temperatures certainly casts some doubt on the latter as a general explanation for the failure of most eucaryotic algae to adapt to temperatures higher than 50 °C.

Metazoans typically have the lowest maximum thermal tolerances of any group of organisms. Here again, the arguments prevailing before the mid-1970s detailing the failure to adapt to higher temperatures invoked some vaguely defined physicochemical limit. In the case of the metazoa, complexity of organization is added to the supposedly inherent limits of the eucaryotic cell. But, as pointed out for other groups, this is only one of the three lines of evidence that must be considered when pondering the origins of thermophily. The other two are the availability of resources, both space and nutrients, and the interactions of resource abundance and evolutionary mechanisms that act to determine fitness. In other words, whatever the physicochemical limit to the evolution of thermal tolerance, simply measuring maximum tolerance rates in the field does not provide evidence, necessarily, that the physicochemical limit has been reached. Clearly, the larger aquatic metazoans could not sustain viable population densities in the small thermal systems, so there would be no selective pressure for them to colonize even moderate to low temperature thermal systems. Thus no fish are found in geothermally heated waters. The few cases of the evolution of tolerance to elevated temperatures in fish involve small

individual size and solar-heated habitats. Here the maximum temperatures tolerated somewhat exceed 43 °C.

The metazoans characteristic of hot springs are the arachnids (water mites), insects (mostly flies), small mollusks (small pulmonate snails), odonates (dragon flies), and, occasionally, ostracods. These groups are neither characteristic of the cold water streams into which the thermal effluents drain, nor of running water systems in general. Rather, they are related to (sometimes the same species as) the invertebrates found in the shallow muddy margins of lakes and ponds, where ambient daily maximum temperatures, due to solar heating, are of the same order as those in which the thermal inhabitants are found. If these low to moderate temperatures are not set by a physicochemical threshold, then the problem is to explain why, in thermal habitats, the species have not evolved more thermal tolerance.

The second factor affecting distribution (mentioned earlier), competition for scarce resources, seems not to be a factor. In every case studied, the grazers on the algae of thermal ecosystems are less tolerant of high temperature than the cyanobacteria on which they feed. Examples are the ostracods in Hunter Hot Springs, Oregon, and the ephydrid flies of the "alkaline" thermal flows in Yellowstone Park. This leaves the interaction of resources and selection determining fitness. Under random sexual mating (with respect to temperature), the direction of the evolution of temperature tolerance will be determined by the relative numbers of surviving offspring from the upper half of the temperature range relative to those from the lower half. But because of the nonlinear cooling curve of water, the size of the upper temperature part of the range will always be smaller than that of the lower half. Thus under a tactic of shifting range, thermophily cannot evolve unless there is nonrandom mating or fecundity or survival is lower at low temperatures. The randomness of mating with respect to temperature differs in the various groups of hot springs invertebrates, depending on life history and behavior. In the flies and dragonflies, mating occurs after emergence and the animals are highly motile, so randomness with respect to temperature of development is assured. The water mites are also highly dispersed as parasitic larvae, but they do spend their adult lives within a relative narrow temperature range and mate there. The smaller mite, *Partnuniella*, tends to spend most of its adult life in cooler patches of the cyanobacteria/flexibacteria mat where it feeds on the eggs of the ephydrid flies and, as expected, has a rather low temperature tolerance. The large predaceous mite, *Thermacarus*, lives directly in the hot water and has an upper temperature tolerance 5 to 10 °C higher than the smaller species of mite and just lower than the most temperature tolerant metazoan, the ostracod. This group has no widespread dispersal mechanism, thus nonrandom mating is predicted. They also seem to be limited by food. Not surprisingly, these organisms have evolved the highest tolerance to high temperature of any metazoan.

In summary, current evidence supports the following statements concerning the evolution of thermophily: (a) the heterotrophic procaryotic bacteria seem not to have any physicochemical limit, having evolved to inhabit geothermal systems of any temperature as long as liquid water is present; (b) the photosynthetic cyanobacteria seem to have evolved to the point (73 to 75 °C tolerance) where a physicochemical limit has been reached; and (c) eucaryotes vary greatly in temperature tolerance, depending on the group, but none has evolved tolerances equal to the maxima exhibited by the procaryotes. In most cases, the observed limits can be explained on the basis of simple selection/fitness processes, without invoking unknown physicochemical mechanisms.

The Biodiversity of Thermal Ecosystems

Species in, using, or around (living from) geothermal ecosystems throughout the world have been the subject of many studies. The earliest intensive work was on the fauna of the hot springs of Iceland. The algae of several hot springs in western India were cataloged in the 1960s. The algae and fauna of warm springs in New Zealand have also been studied, as has the fauna of thermal springs in the Dutch East Indies. The thermal effluents of Yellowstone National Park in the United States have been the subject of both the largest number and most intensive studies of thermophiles of any area in the world. The original studies concentrated on the fauna but also included some observations of the algae (cyanobacteria). Beginning in 1964, detailed studies of the microorganisms living in the hot spring waters of Yellowstone Park were begun. These became both the most extensive survey and intensive experimental protocol yet done on thermal microbiology. At first the focus was on high temperature systems where only one or a few species were found in any given thermal ecosystem. In 1967, I and my colleagues and students began studying the lower temperature thermal ecosystems with higher diversities of microorganisms and a richer food chain involving species of animal acting as scavengers, predators, and parasites. Additional unpublished information has been obtained from 1983 to the present on the thermal ecosystems developing at temperatures below 45 °C. Thus the ecology of the Yellowstone ecosystems can begin to be understood from the standpoints of species diversity and food-chain relationships. In this final section I review what is known about the diversity of these systems, in their various manifestations of temperature and pH, and compare the results with the thermal biodiversity of the deep sea thermal vents, insofar as the latter are currently understood.

High temperature is the most difficult of extreme environmental conditions to which organisms can attempt to adapt. There are many ways to adapt to exist in extreme cold or aridity and resume growth and activity during short seasonal changes in solar warming and precipitation. High salinity can be avoided or salt excreted so extremes can still support a reasonably high diversity. Indeed, many saline ecosystems may appear simple from the standpoint of vascular plant species and the food chains they support, but be relatively high in the diversity of algae and bacteria in the surface sediments. Thus, in every thermal system investigated to date, the biodiversity decreases rapidly with increases in temperature. At their most diverse, thermal system species are measured in the hundreds; their least diverse manifestations are monospecific.

Temperatures near the boiling point of water in Yellowstone thermal systems and elsewhere (85 °C–100 °C) are most common in source pools fed by deep subterranean springs of hot water. These pools vary widely in chemical characteristics,

depending on the mineral content of the rocks through which the superheated water passes on its way to the surface. All such pools have a large volume of heated water with a relatively small inflow/outflow. Currents are small; almost all mixing is by convection or mild bubbling of various gases, often carbon dioxide. Such systems are natural analogues of chemostats, and they are eventually colonized by adapted heterotrophic bacteria (photosynthesis being impossible at these high temperatures). Not surprisingly, competition generally ensures that they are monospecific, or, if chemoautotrophy is possible, the diversity might rise by one or two additional species. Although thermal source pools may be thought small relative to surrounding mountains, forests, and so on, from the viewpoint of microorganisms, these are virtually constant environmental systems capable of supporting immense numbers of individuals. Current ecological theory supports the prediction of dominance by one or a few species in these circumstances. In contrast, the emergence of superheated water from the thermal vents in the deep sea cannot reproduce these constant conditions. Pools of hot water do not exist, but instead a turbulent mixture of water is emitted at temperatures apparently as high as 350 °C or more, directly into cold ocean water near 2 °C. Thus environmental variation in temperature is undoubtedly high and the higher diversity of microorganisms would be predicted. It is impossible, under such variability in temperature, for one species to monopolize the resources of the system. Thus the students of thermal vent microbiology are finding a number of thermophilic species new to science (although no novel genera have been found). The latter may be due to the culture techniques used, which have not been designed for organisms adapted to high pressures as well as high temperatures.

Food chains in the deep sea vent communities are simple and short, because in the absence of light no photoautotrophy is possible. Thus the vents are surrounded by a diverse array of marine invertebrates that ingest the cells and products from the thermal vents. Almost all of these animals are novel species. As many as 236 species of animals have been collected around thermal vents, of which 223 are new to science. This diversity of animal species is far higher than that of any single category of surface thermal ecosystems, although the data represent many different thermal vent locations. If the worldwide diversity of animal species for all known surface thermal features is used, the animal diversity per system is similar in the two types of system. Thus the thermal vent communities of the deep sea, although of great interest from the standpoint of physiological adaptations, mass mortality, and colonization strategy, appear to offer little that is new to ecological theory.

At temperatures below 75 °C and above about 40 °C to 45 °C, thermal outflows form small streams of hot water. These are colonized throughout the world by unicellular photoautotrophic cyanobacteria and filamentous bacteria. These biotic communities are characteristic of all surface thermal features throughout the world, but most early studies were simply endeavoring to determine maximum temperatures. The first rigorous studies of these systems were initiated in Yellowstone Park. In each ecosystem, there are generally two dominant species or species groups at the upper end of this temperature range (a base mat of the filamentous photosynthetic bacterium, *Chloroflexus*, and an overlying film of the unicellular cyanbacterium, *Synechococcus*). As the temperature decreases, the overlying mat becomes dominated by one or more genera of filamentous cyanobacteria. In relatively constant flows, temperature-adapted strains of *Mastigocladus* are the common dominants; in springs with variable ouflows, diversity is often higher due to the presence of additional genera of filamentous cyanobacteria, commonly *Calothrix* or *Phormidium*. When these filamentous forms are able to grow, the outflow begins to go through a definite cycle, because the filamentous mat is not limited to a thin film as are the unicellular forms. As the mat grows, it thickens unevenly; water flows around these thickened areas, producing cool patches. These patches are populated, almost immediately, by one or more species of brine fly (Ephydridae). Upon hatching, the fly larvae begin to eat and destroy the filamentous integrity of the mat. The pupae metamorphose to adults within a few days. Eventually hot water reenters the cool patch, remaining larvae and pupae are killed, and the cycle starts over in another part of the mat. These flies and their predators are not thermophilic, and in fact are rather thermally intolerant, relative to organisms living in other thermal ecosystems—compost piles, for example. They are an example of what I have termed thermal opportunists, organisms that have adapted their life cycle to take advantage of opportunities provided within the thermal ecosystem. The diversity of the alkaline filamentous mat-fly association is not high for most expressions of this community of 20 to 30 species. Yet a diverse set of food chains and ecological processes is found, photosynthesis, grazing, decomposition, and predation, for example.

When a strongly channeled outflow cools down to 40 °C or below, the diversity goes up rapidly. Metazoan animals can adapt to live directly in the warm water, and eucaryotic algae such as green algae and diatoms invade. Although cyanobacteria are still abundant, they are heavily grazed, and do not form a thick mat. In different expressions of this community are found crustacea, midges (Diptera), and water mites (Hydrachnellae), dominated by a top predator, the larva of the dragonfly, *Erythemis collocatta*. The net effect on diversity is to increase the species numbers to 50 to 60, exclusive of the nonphotosynthetic microorganisms. The latter have not been studied in detail at these temperatures, thus a total diversity of 100 or more would not be unexpected.

Summary

Thermal waters, as extreme ecosystems, exhibit the low diversity expected under such conditions. But high temperature is a particularly difficult factor for adaptive evolution to overcome. Thus, thermal ecosystems appear to have the lowest diversities of any of the extreme environments. Nevertheless, there is considerable variation in diversity, depending on both the temperature of emerging water, the rate of cooling, and the stability of the emerging plume. Generally, there is an increase in diversity as the temperature declines. This can be modified by short-term fluctuations in temperature caused by currents set up in the emerging plume of hot water (in deep sea thermal vents) or by intermittent flows from the source creating cooler periods downstream. Both of these effects increase

diversity. When animals are able to enter the system, a rich ecological food chain (in terms of numbers of ecological processes) often develops.

Acknowledgments

I am indebted for the information in this article to dozens of previous studies of the ecology of thermal organisms. My own work in Yellowstone National Park and elsewhere has benefited from permission from the park administration to do research. I have received financial support from the National Science Foundation through a series of grants, the latest of which, DEB-821–0142, was used to investigate the community forming at temperatures below 40 to 45 °C. I am also indebted to Drs. A. Chalmers and R. Hodson, who provided useful comments on the manuscript. Support for preparation of the manuscript was received as discretionary research funds from the University of Georgia.

See also: Psychrophiles. Thermophiles, Origin of. Vents

References

Brock TD (1967a) Life at high temperatures. *Science* 158: 1012–1019.

Brock TD (1967b) Microorganisms adapted to high temperatures. *Nature* 214: 882–885.

Brock TD (1969) Microbial growth under extreme conditions. *Symp. Soc. Gen. Microbiol* 19: 15–41.

Brock TD (1978) *Thermophilic Microorganisms and Life at High Temperatures*. New York: Springer-Verlag.

Brues CT (1924) Observations on animal life in the thermal waters of Yellowstone Park, with a consideration of the thermal environment. *Proc. Amer. Acad. Arts and Sci.* 59: 371–437.

Brues CT (1927) Animal life in hot springs. *Quart. Rev. Biol.* 2: 181–203.

Brues CT (1932) Further studies on the fauna of North American hot springs. *Proc. Amer. Acad. Arts and Sci.* 67: 185–303.

Castenholz RW (1973) The ecology of blue-green algae in hot springs. In: Carr NG and Whitten BH (eds.) *The Biology of Blue-Green Algae*, pp. 379–414. Oxford: Blackwell.

Childress JJ and Fisher CR (1992) The biology of hydrothermal vent animals: Physiology, biochemistry and autotrophic symbioses. *Oceanogr. Mar. Biol. Annu. Rev.* 30: 337–441.

Collins NC, Mitchell R, and Wiegert RG (1976) Functional analysis of a thermal spring ecosystem, with an evaluation of the role of consumers. *Ecology* 57: 1221–1232.

Kullberg R (1971) Algal distribution in six thermal spring effluents. *Trans. Am. Micr. Soc.* 90: 412–434.

Lutz RA and Kennish MJ (1993) Ecology of deep-sea hydrothermal vent communities: A review. *Reviews of Geophysics* 31: 211–242.

Mitchell R (1974) The evolution of thermophily in hot springs. *Quart. Rev. Biol.* 49: 229–242.

Prieur D (1971) Microbiology of deep-sea hydrothermal vents. *Marine Biotechnology, TIBTECH* 15: 242–244.

Segerer AH, Burggraf S, Fiala G, *et al.* (1993) Life in hot springs and thermal vents. *Origins of Life and Evolution of the Biosphere* 23: 77–90.

Tunnicliffe V (1972) The biology of hydrothermal vents: Ecology and evolution. *Oceanogr. Mar. Biol. Annu. Rev.* 29: 319–407.

Wiegert RG and Fraleigh PC (1972) Ecology of Yellowstone effluent systems: Net primary production and species diversity of a successional blue-green algal mat. *Limnol. Oceanogr.* 17: 215–228.

Wiegert RG and Mitchell R (1973) Ecology of Yellowstone thermal effluent systems: Intersects of blue-green algae, grazing flies (*Paracoenia*, Ephydridae) and water mites (*Partnuniella*, Hydrachnellae). *Hydrobiologia* 41: 251–271.

Historical Awareness of Biodiversity

David Takacs, University of California, San Francisco, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Biologists and Biodiversity Before 1986

Since ancient times, scholars have simultaneously revered the natural world, attempted to discover a natural order in that world or impose rational order on that world, and sought to understand the place of humans in the cosmos based on what they read in the natural world.

Worster (1987), Bowler (1993), and other historians offer extensive treatments of pre-twentieth century attempts by biologists and their intellectual predecessors to understand, categorize, and philosophize the natural world. In the mid-eighteenth century, for example, in order to fathom God's wisdom, Carolus Linnaeus classified the riot of life into a functional taxonomy, the *Systema Naturae*. In "The Oeconomy of Nature" (1749), Linnaeus offered a protoecological treatise on how these life forms fit together in mutual, stable interdependence, with humans at a central nexus in the web. For Linnaeus and his contemporaries, God meant humans to use those species that we found valuable; the complexity and richness of the natural world meant that our use of nature could not upset what God had balanced so exquisitely.

A century later, Charles Darwin built on Linnaeus's and others' ecological and taxonomic groundwork and painstakingly laid out a theory that provided an historical, deterministic explanation for the relatedness of all God's creatures – even if his work made God seem a bit less relevant and humans a bit less central to the machinery of nature. Darwin saw all forms of life as human kin and believed a key to civilization's maturity was the ability to empathize with and respect our extended family. For both Linnaeus and Darwin, the diversity of earth's life forms was a phenomenon to be revered, contemplated, and explained. Neither of them envisioned the object of their study as a commodity that needed protection from human advances, although a contemporary of Darwin's, Marsh (1864/1965), did begin to grasp this in his prescient *Man and Nature*. Marsh warned that human activity was destroying natural functions, and that this destruction threatened to undercut the survival of our own species.

In the twentieth century, some biologists looking for meaning in the diversity of life added to the trio of reverence, imposition of order, and search for deeper meaning about humanity. Biologists came to understand and document that the grand panoply of life forms was threatened as the fabric of ecological connections was rent. This awareness led biologists to make leaders and laypersons aware, and they searched for ways to use their carefully tended scientific authority to make nonscientists care about and conserve the objects of their reverence. Such work carries perils: advocacy threatens to undermine the perception of value neutrality and objectivity that leads laypersons to listen to scientists in the first place.

In *A Sand County Almanac II*, wildlife ecologist Leopold (1949/1970) proposed a "land ethic" that we should follow if we are to revere nature's diverse organisms and protect them to play the precise roles in nature's scheme that ecological scientists were just beginning to elucidate:

> If the biota, in the course of aeons, has built something we like but do not understand, then who but a fool would discard seemingly useless parts? To keep every cog and wheel is the first precaution of intelligent tinkering (p. 190).

Leopold asserted that all species are intertwined in complex interrelationships, and the diversity of organisms and their interrelationships are crucial for a stable, functioning planet. From the "is" of ecology, Leopold derived the "ought" that forms the basis of his land ethic:

> A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise (p. 262).

British ecologist Elton (1958) also viewed "ecological variety" as a threatened commodity. In *The Ecology of Invasions by Animals and Plants*, he foretold the dangers that humans would face as anthropogenic introductions of nonendemic species sweep the globe and threaten the flora and fauna that had previously been left to evolve in splendid isolation:

> It is not just nuclear bombs and wars that threaten us, though these rank very high on the list at the moment: there are other sorts of explosions, and this book is about ecological explosions (p. 15).

Like Elton, Carson (1962/1987) warned that people could no longer take this variety for granted as benign backdrop to human affairs. Rather, ecological science taught that species diversity is essential for ecological and human health and that uncontrolled pesticide-use threatened this diversity. Biologists like Leopold, Elton, and Carson still revered "natural variety," still sought to reveal the secrets of nature's order. But unlike pre-twentieth century natural philosophers, they desired that their readers come to view species diversity as a quantifiable, measurable entity that was inextricably bound up with their own well-being. Species diversity was now a threatened commodity, and humans who caused this threat were simultaneously threatened as this commodity diminished.

Creation of the Term "Biodiversity"

Throughout the 1960s and 1970s, biologists fomented public alarm over the deteriorating environment. Prominent biologists (e.g., Ehrenfeld, 1981; Ehrlich and Ehrlich, 1981; Myers, 1979) helped raise awareness that diversity was a threatened commodity, and US legislatures responded with laws designed to assuage the threat. Most notably, the 1973 Endangered Species Act (ESA) recognized and sought to protect the

paramount value of species diversity (Kohm, 1991). Recognizing that

> species of fish, wildlife, and plants have been so depleted in numbers that they are in danger of or threatened with extinction

and that these forms of life

> are of esthetic, ecological, educational, historical, recreational, and scientific value to the Nation and its people

the ESA commits the nation to preserving these species for current and future generations. While limiting its focus to species, and whittled down from its original ambitions by insufficient budgeting, cautious administrative bureaucrats, and restrictive judicial interpretations, the ESA remains one of the most powerful allies biodiversity has ever had.

Still, biologists were frustrated that efforts to protect diversity were not keeping pace with the furious rate of destruction. Walter G. Rosen, a biologist and senior program officer at the National Research Council (which advises the National Academy of Sciences (NAS)), brought together prominent scientists from the NAS with the clout of the Smithsonian Institution to host the National Forum on BioDiversity in 1986. At the National Forum, biologists and others concerned about imperiled diversity staged a consciousness-raising event that sought, and received, widespread attention from the public. Rosen coined the neologism "biodiversity" for the event as a convenient shorthand, a buzzword that would at once encapsulate biologists' understanding of a chaotic, diminishing natural world, and would raise public awareness about threats to the natural world (Takacs, 1996). In *BioDiversity* (Wilson, 1988), the collection of essays that chronicled the National Forum, Paul Ehrlich, Daniel Janzen, Tom Cade, Lester Brown, Michael Soulé, and other scientists declared the need to rouse public attention on behalf of biodiversity and exhorted their colleagues to adopt that mission.

As biologists promote the term and the complex worldview it represents, "biodiversity" has become a widespread conservation buzzword. Biologists write about it in scientific and popular presses, both exploring its complexity and advocating its protection. Environmental groups focus on it in fundraising efforts; conferences convened in its name occur regularly. Laypersons have joined biologists in attempting to shape the planet's physical, political, and normative environments to make more room for biodiversity.

Problems with Other Foci of Conservation Efforts

Why has "biodiversity" gained prominence as a conservation buzzword, and why have biologists speaking on its behalf had some success in shaping public opinion about threats to the natural world?

Various scholars (Cronon, 1995; Evernden, 1992; Williams, 1980) note that nature is so all-encompassing that what one attributes to it may say more about the speaker than it does about the natural world. Previous attempts to preserve "nature" or "wilderness" are too vaguely defined or deemed elitist in some quarters. For example, Guha (1989) sees traditional advocacy for nature or wilderness preservation as setting the aesthetic desires of the rich against the needs of the poor who need land to survive. Guha's critique has, to some extent, been incorporated into the biodiversity preservation discourse. For example, as people came to appreciate the value of cultural diversity, they might also see diversity in all its forms as a normative good, particularly when efforts to preserve both may be mutually reinforcing (Nabhan, 1997). A document prepared for the 1992 United Nations Conference on Environment and Development notes that

> cultural diversity is closely linked to biodiversity. Humanity's collective knowledge of biodiversity and its use and management rests in cultural diversity; conversely, conserving biodiversity often helps strengthen cultural integrity and values (Reid *et al.*, 1992, p. 23).

Prior to the advent of biodiversity, the most effective conservation efforts in the US focused on endangered species. Biologists' foci changed because biodiversity represents a more sophisticated ecological worldview, and a more sophisticated view of what biologists want preserved and how they want it preserved. Biologists promoting biodiversity conservation also seek to circumvent certain problems arising from efforts to preserve endangered species.

Even though human activities accelerate the rate of species extinction, some opponents of conservation argue that species extinction is a natural process that we should let proceed. Biologists have difficulties defining with precision what constitutes a species (or a subspecies, like the Northern Spotted Owl) in the first place. We are not aware of the economic or ecological benefits some individual species confer, so it is difficult to argue for their conservation in some circles. Many of the insects, bacteria, plants, and other members of what Wilson (1987) calls "the little things that run the world" remain unidentified and unloved; invertebrates comprise only a tiny percentage of organisms protected by the US ESA. Land set aside to protect individual endangered species may prove insufficient if global warming induces species migration (Peters and Lovejoy, 1992). When we focus on species diversity, we sometimes ignore genic, population, community, or ecosystem diversity. Endangered species conservation proves to be nearly impossible in poorly explored areas, in oceans, and in nations without species checklists. Focus on species that are endangered can be a last minute emergency effort to save species that may no longer be playing functional ecological roles, and this strategy may be inferior to proactive efforts to preserve healthy populations in healthy ecosystems. Finally, some view the US ESA as a mixed blessing: its unyielding allegiance to species on the verge of extinction is also a political lighting rod that leaves little room for compromise and has set many citizens against conservation efforts (Mann and Plummer, 1995).

What is "Biodiversity"?

Elsewhere in this encyclopedia, one may read about definitions of biodiversity. In the research conducted by the author (Takacs, 1996), prominent biologists were asked to define the term "biodiversity." Little, if anything, in the natural world is

excluded from these definitions of biodiversity. Responses included that biodiversity is "the complete array of organisms, biologically mediated processes, and organically derived structures out there on the globe" (Jerry Franklin); "the whole package of genes, populations, species, and the cluster of interactions that they manifest" (Daniel Janzen); "the total number of genetic lineages on earth" (Thomas Eisner); "shorthand for all the richness of life" (Reed Noss); or "the sum total of plants, animals, fungi, and microorganisms in the world including their genetic diversity and the way in which they fit together into communities and ecosystems" (Peter Raven). According to Terry Erwin (1991, p. 3), it is

> the sum of earth species including all their interactions and variations within their biotic and abiotic environment in both space and time.

So biodiversity represents a rich, complicated vision of life and a corresponding rich, complicated vision of what biologists want to see preserved. Although few biologists wish to see any species slip into extinction, what they really want – and what biodiversity really represents – is preservation of ecosystems, of vast stretches of land where the evolutionary process may continue relatively unfettered.

But ecosystems are hard to delineate, and even mentioning "evolution" lights a powder keg in some quarters. Kellert (1986, 1996) notes that the public is most likely to rally around "cognitively meaningful" organisms. Habitat has no fur, and ecosystems lack big, expressive eyes that rouse public affection and therefore attention to conservation causes. With biodiversity, biologists can focus attention on the charismatic megavertebrates the layperson finds endearing. These organisms are emotionally and ecologically appealing to the conservation minded: they often are at trophic levels that require large territories to survive. When we preserve the large areas of land they require, we also fortuitously protect a myriad of species and enable continued functioning of ecological and evolutionary processes that biologists value.

Conservation efforts on behalf of biodiversity enable multiple images and multiple strategies to protect and conserve the natural world; many of those strategies place more power in the hands of biologists to realize their conservation values. If definitions of biodiversity seem complex and all encompassing, that is part of why it has been successful as a conservation tool. At once it represents the complexity of biologists' worldview, the whole span of what biologists wish to conserve, and in it each of us can see that part of nature we cherish.

Biologists and the Promotion of Biodiversity

When biologists speak about biodiversity, they simultaneously refer to small parts of the ecological world, the interrelatedness between those parts, the ecological processes that sustain those parts, and the evolutionary processes that gave rise to those parts. While appearing as a purely scientific, objective entity, biodiversity also encompasses political arguments on behalf of conservation and symbolizes much that biologists do not know about the natural world.

Biodiversity represents ecological complexity that, the more they study, the more biologists realize they do not understand. Biologists do not know, within an order of magnitude, how many species live on earth, so they cannot specify how many we are losing and cannot say what the loss of these species represents to the ecosystems or to the humanity's future prospects (Stork, 1997; Tilman, 1999; quotes in Takacs, 1996). Biologists believe that if people are alarmed by such ignorance, they will simultaneously take precautions about dismantling an ecological world that supports us in myriad unknown ways, will support biologists' research efforts to understand that world, and will listen to biologists' policy prescriptions on how and why we ought to save that world. For example, the question that Wilson (1992) is most frequently asked about the diversity of life is:

> if enough species are extinguished, will the ecosystem collapse, and will the extinction of most other species follow soon afterward? The only answer anyone can give is: possibly. By the time we find out, however, it might be too late. One planet, one experiment (p. 182).

In light of this uncertainty, Wilson urges

> prudence. We should judge every scrap of biodiversity as priceless while we learn to use it and come to understand what it means to humanity (p. 351).

In Wilson's view, biologists hold the key to that understanding.

Many who have come to call themselves conservation biologists cite this uncertainty as they attempt to have the public become aware of biodiversity and become concerned about its diminution. They warn that what we do not know about biodiversity will hurt us, and urge us that if we were to come to know it, our lives might have greater meaning. They seek a new ethic where biodiversity would be cherished. Leopold (1949/1970) sought this, and his "land ethic" laid the groundwork for a new way of valuing diversity. Leopold knew that

> no important change in ethics was ever accomplished without an internal change in our intellectual emphasis, loyalties, affections, and convictions. The proof that conservation has not yet touched these foundations of conduct lies in the fact that philosophy and religion have not yet heard of it.

Today, a number of prominent biologists seek to have our intellectual emphases, loyalties, affections, and convictions match theirs. Some biologists see this as a crucial moment in history when they must help raise awareness of biodiversity; they see it as their responsibility to speak out about what they study and love so that subsequent human and natural history will reflect their values. For example, Janzen (1986) exhorts his fellow biologists:

> Set aside your random research and devote your life to activities that will bring the world to understand that tropical nature is an integral part of human life.

Janzen is part of the "mission-oriented" discipline of conservation biology, which was founded not only to study biodiversity, but also to embrace its many values and to promote

findings that will help conserve the objects of study. Built into the discipline's foundations are normative principles or, as discipline founder Soulé (1985, p. 730) says,

> value statements that make up the basis of appropriate attitudes toward other forms of life – an ecosophy.

The pages of the discipline's flagship journal, *Conservation Biology*, report not just the latest research findings; they are also filled with thoughtful articles on conservation education, policy implications of research, and spirited debates on what types of activism are required or appropriate for practitioners. Although some biologists feel they must advocate on behalf of the natural world they study and love, others fear that such advocacy threatens the scientific enterprise: If scientists are no longer perceived as objective and value neutral, why would societies fund their work and listen to their results?

Many scientists ignore the risks of advocacy and proselytize on behalf of biodiversity. Some well-known biologists (e.g., Wilson, Raven, Eisner, and Lovejoy) have spoken before Congress on the value of biodiversity. Others testify in courts as expert witnesses on behalf of biodiversity. Paul Ehrlich and others appear on radio and television to raise awareness about biodiversity. Many biologists write for the popular press or speak to general audiences on the values of biodiversity. They talk to garden clubs, and preach in church pulpits. Lovejoy takes senators, movie stars, and other influential people to the Brazilian Amazon so that they may experience biodiversity

> in a way words can't touch. … And it has never failed to be a truly touching experience for them (Takacs, 1996, p. 153).

By helping laypersons become aware of biodiversity's beauty and fragility – by tireless efforts to, in the words of the "Father of Biodiversity" Wilson, "educate, educate, educate" (quoted in Anonymous, 1994) – biologists hope for widespread transformation of values. That transformation of values might lead to what Ehrlich (1985) calls a "quasireligious transformation" of feelings toward the wonders of the natural world, which would presumably be translated into action to conserve those wonders.

Many prominent North American biologists work in far-flung corners of the earth to raise awareness of biodiversity, particularly in tropical nations that hold the greatest concentrations of diversity, and which face the greatest imminent threats to that bounty. In Costa Rica, a small country with perhaps 5% of the world's species diversity, biologists have spurred a national effort not only to preserve biodiversity in protected areas, but also to catalog biodiversity and to search for the economic wealth contained within. At Costa Rica's Instituto Nacional de Biodiversidad, specimens of plants, insects, and other taxa are sorted and sent to multinational pharmaceutical companies for research. Profits from successful drugs deriving from Costa Rica's biodiversity would be funneled back to the areas where the biodiversity flourishes. In effect, this is an experiment in which biodiversity would pay its own way, with biologists driving the process.

Specimens are collected by a large group of rural-dwelling parataxonomists. As they tromp around the nation's forests, the idea goes, they will lead a movement in "biocultural restoration": by becoming reenchanted with biodiversity, they will be compelled to spread their new knowledge and appreciation to their neighbors. Through renewed awareness – and renewed economic dividends – rural dwellers would be empowered and inclined to protect their biodiversity patrimony (Evans, 1999; Gámez *et al.*, 1993; Janzen, 1988).

The Values of Biodiversity

In Costa Rica, and in many other corners of the earth, biologists are working to raise awareness of the many values of biodiversity. Biologists hope that biodiversity's appeal to many different audiences will spark diverse efforts to preserve it. Some of the values that biologists attribute to biodiversity find their loci in biodiversity itself; in most, though, humans are the value holders.

As we might expect, biologists promote the value biodiversity holds for science, as the raw material for scientists' investigations: it is "Earth's living library" (Lovejoy, 1992). Also unsurprisingly, biologists speak extensively about biodiversity's ecological value. Biodiversity provides numerous "ecosystem services" (or, depending on your definition, "ecosystem services" is part of what constitutes "biodiversity"). That is to say, biodiversity keeps global ecosystems functioning. Biodiversity purifies water, controls agricultural pests, decomposes waste, pollinates plants, stabilizes global climate, creates soil, transports nutrients, controls erosion, and maintains the ecological matrix human society requires to exist at all (Daily, 1997; Tilman, 1999). And as humans attempt to repair damaged ecosystems, we will need reservoirs of intact biodiversity to restore what we have destroyed (Jordan, 1997).

For those who measure value in terms of money, biologists assert that biodiversity has vast economic value. Biologists note that biodiversity's value to functioning ecosystems is priceless – although Costanza *et al.*, (1997) have tried to calculate the incalculable. Costanza's team figured that 17 ecosystem services in 16 biomes contribute between 16 and 54 trillion (US) dollars per year to human economies, and they believe this is a conservative estimate. This, they note, is higher than the global gross national product of 18 trillion dollars per year. Janzen (1986) offers an exhaustive list of the treasures that biodiversity already has directly provided human society. Others (e.g., Wilson, 1992) discuss the lucre that can be gained from the hidden food, medicine, chemicals, fibers, and other goods lurking in the world's natural places. As noted earlier, Costa Rica has invested heavily in its biodiversity; through ecotourism and bioprospecting, the nation is profiting from its biodiversity resources. Biologists (e.g., Janzen, 1990) also tout biodiversity's social amenity value: it can contribute to sustainable development efforts (through direct harvesting or ecotourism), foster national pride, or help those living nearby to lead more intellectually and aesthetically fulfilled lives.

Wilson (1992, 1998) has been the leading proponent of biodiversity's value to fulfill our biophilic impulses. Wilson and others believe that love of nature has been hardwired into our genes, and we can only be truly fulfilled by satisfying that love. Those whose biophilic passions have been rekindled will also be those who work hardest for biodiversity

preservation: others refer to this as biodiversity's "transformative value." Philosopher Norton (1987) explains how interactions with biodiversity can help us reconsider our consumerist impulses and make us convert to lifeways that preserve the biodiversity that provided the impetus for transformation. Biodiversity's aesthetic value – the beauty of an individual organism and its adaptations, of landscapes, of intricate ecological processes, of the sheer riot of different life forms – may effect this kind of transformation.

When biologists suggest that biodiversity may transform us, may reawaken our biophilic impulses, they still place the locus of its value in the human valuer. Some biologists take a different tack and assert that biodiversity has intrinsic value, independent of a human valuer. This concept is difficult to prove empirically and therefore difficult for a scientist to assert. Yet given the preceding definitions of "biodiversity," if biodiversity is the totality of life forms on earth, their interactions, and the processes that gave rise to them and to us, it makes it a more complicated proposition to reject the notion of biodiversity's intrinsic value out of hand. This notion takes a turn for the spiritual, and, in fact, some biologists also discuss biodiversity's spiritual value. Ehrenfeld (1981), who would become the founding editor of *Conservation Biology*, wrote of "The Noah Principle": species

> should be conserved because they exist and because this existence is itself but the present expression of a continuing historical process of immense antiquity and majesty. Long-standing existence in Nature is deemed to carry with it the unimpeachable right to continued existence.

Some biologists (see quotes in Takacs, 1996) agree that the value inherent in biodiversity makes it sacred and others put the locus of value in humans, and they discuss the spiritual nourishment that contact with biodiversity brings.

The multiplicity of meanings of "biodiversity" is reflected in the multiplicity of values biologists find in it. For those biologists who would raise awareness of biodiversity's plight, it makes sense to speak for as many different values – and therefore to as many different audiences – as possible.

International Awareness Grows, Biodiversity Continues to Dwindle

Efforts to raise awareness about biodiversity continue unabated. One-hundred and sixty-eight nations have signed the Convention on Biological Diversity, whose objectives are

> the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources… (Convention on Biological Diversity, 2011a, b).

Despite nearly 20 years of meetings and plans and firm targets (see, e.g., the Strategic Plan for Biodiversity 2011–2020, Convention on Biological Diversity, 2011c), and with the heralding of 2011–2020 as the "United Nations Decade on Biodiversity," the web of life on Earth continues to unravel. The Secretariat of the Convention on Biological Diversity, despite noting some tepid bright spots (e.g., more nations with biodiversity plans and a slower rate of tropical forest loss), nonetheless asserts that a quarter of plant species are threatened, amphibians and coral reefs lead other species groups on the march to extinction, and most habitats of the world are increasingly fragmented or degraded (Convention on Biological Diversity, 2010). Similarly, the European Union has a formal Biodiversity Strategy, with the heading, "Our life insurance, our natural capital" (European Commission, 2011). Recognizing that only 11% of key ecosystems and 17% of species and habitats protected by the EU are in a favorable condition, the Strategy nonetheless recognizes that their previous biodiversity target for 2010 would not be met, but pledges to do better this time around.

Inspired by an idea from Wilson, the "Encyclopedia of Life" seeks to establish "a webpage for every species" (more than 900,000 as of this writing) as a resource "to increase awareness and understanding of living nature" for "the general public, enthusiastic amateurs, educators, students and professional scientists from around the world" (EOL, 2011). The specter of global climate change has provided new urgency for efforts to raise awareness about the manifold threats to biodiversity, with new mechanisms like carbon offsetting for "Reducing Emissions from Deforestation and forest Degradation (REDD)" providing new financial means for preserving tropical biodiversity (Takacs 2010). Biologists are playing a key role in designing mechanisms and standards to ensure that funds invested in carbon offsetting conserve biodiversity; for example, the Climate, Community, and Biodiversity Alliance requires carbon offsets to result in "Net Positive Biodiversity Impacts" and requires biodiversity impact monitoring (CCB, 2011).

Historical Awareness of Biodiversity Redux

We can read the title of this encyclopedia entry in many ways. Smithsonian tropical biologist and biodiversity advocate Erwin says (in Takacs, 1996, pp. 100–101):

> I've always kind of been ahead of the game or ahead of the thinking in my own little field of entomology … I think there will be individuals who can influence the direction of change. And I'd like to be among them, for whatever little part I might be able to play.

Ideas can act as forces of nature. They can change ecologies. They can reshape how we value and therefore how we treat nature. On behalf of biodiversity, biologists are making history: they are changing the course of events so that human history and natural history will unfurl to their liking. This encyclopedia attempts to inform, to raise awareness of both the science of biodiversity and the conscience of those who study it, revere it, and wish to see it preserved.

History suggests that two phenomena will continue: the biological complexity of the earth will continue to diminish, and biologists will continue to look for strategies that will compel us to care about what they care about and support their authority to speak for those entities. Biologists have attempted to raise our awareness of biodiversity: the complexity of real world organisms, species, and processes commingled with biologists' factual, political, emotional, ethical, aesthetic,

and spiritual values of the natural world, all combined to shape public perceptions, actions, and feelings. To become aware of it is to understand not only how biologists understand the riot of life and interconnections, but also to become aware of our own obligations to future evolutionary history.

Appendix

List of Courses

1. History of Life on Earth
2. History of Biology
3. Conservation Biology
4. Environmental Law
5. Environmental Education

See also: Biodiversity, Definition of. Biodiversity-Rich Countries. Conservation Biology, Discipline of. Conservation Efforts, Contemporary. Convention on Biological Diversity. Darwin, Charles (Darwinism). Ethical Issues in Biodiversity Protection. Forest Canopies, Animal Diversity. Stewardship, Concept of. The Value of Biodiversity

References

Anonymous (1994) EO Wilson: An interview with the father of biodiversity. *Nature Conservancy* 44(4): 24–29.
Bowler P (1993) *The Norton History of the Environmental Sciences*. New York: Norton.
Carson R (1962/1987) *Silent Spring*. Boston: Houghton Mifflin.
Community, Climate, and Biodiversity Alliance (2011) "CCB Standards,"http://www.climate-standards.org/standards/index.html (last accessed 5 December 2011).
Convention on Biological Diversity (2010) *Global Biodiversity Outlook 3*.
Convention on Biological Diversity (2011a) "Article 1: Objectives," (last accessed 3 December 2011), http://www.cbd.int/convention/articles/a=cbd-01
Convention on Biological Diversity (2011b) "List of Parties," (last accessed 3 December 2011), http://www.cbd.int/convention/parties/list/
Convention on Biological Diversity (2011c) "Strategic Plan for Biodiversity 2011–2020, including Aichi Biodiversity Targets," (last accessed 3 December 2011), http://www.cbd.int/sp/
Costanza R, d'Arge R, deGroot R, *et al*. (1997) The value of the world's ecosystem services and natural capital. *Nature* 387: 253.
Cronon W (1995) Introduction: In search of nature. In: Cronon W (ed.) *Uncommon Ground: Toward Reinventing Nature*. New York: WW Norton.
Daily GC (ed.) (1997) *Nature's Services: Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.
Ehrenfeld D (1981) *The Arrogance of Humanism*. New York: Oxford University Press.
Ehrlich P (1985) Extinctions and ecosystem functions: Implications for humankind. In: Hoage RJ (ed.) *Animal Extinctions: What Everyone Should Know*. Washington, DC: Smithsonian Institution Press.
Ehrlich PR and Ehrlich AH (1981) *Extinction: The Causes and Consequences of the Disappearance of Species*. New York: Ballantine Books.
Elton CS (1958) *The Ecology of Invasions by Animals and Plants*. London: Methuen.
Encyclopedia of Life (2011) http://eol.org (last accessed 5 December 2011).
Erwin T (1991) A plan for developing consistent biotic inventories in temperate and tropical habitats. Part 1 of Establishing a Tropical Species Co-ocurrence Database. In: Erwin TL (ed.) *Memorias del Museo de Historia Natural, Parts 1–3*. Lima: Universidad Nacional Mayor de San Marcos.
European Commission, Our life insurance, our natural capital: An EU biodiversity strategy to 2020, Brussels, 5 March 2011, COM (2011) 244 final.
Evans S (1999) *The Green Republic: A Conservation History of Costa Rica*. Austin: University of Texas Press.
Evernden N (1992) *The Social Creation of Nature*. Baltimore: Johns Hopkins University Press.
Gámez R, Piva A, Sittenfeld A, Leon E, Jimenez J, and Mirabelli G (1993) Costa Rica's Conservation Program and National Biodiversity Institute. In: Reid WV, Laird SA, Meyer CA, Gámez R, Sittenfeld A, Janzen DH, Gollin MA, and Juma C (eds.) *Biodiversity Prospecting*. Washington, DC: World Resources Institute.
Guha R (1989) Radical American environmentalism and wilderness preservation: A third world critique. *Environmental Ethics* 11: 71–83.
Janzen DH (1986) The future of tropical ecology. *Annual Review of Ecology and Systematics* 17: 305–324.
Janzen DH (1988) Tropical ecological and biocultural restoration. *Science* 239: 243–244.
Janzen DH (1990) "Sustainable society through applied ecology: The reinvention of the village. In: Goodland R (ed.) *Race to Save the Tropics*. Washington, DC: Island Press.
Jordan W (1997) Ecological restoration and the conservation of biodiversity. In: Reaka-Kudla ML, Wilson DE, and Wilson EO (eds.) *Biodiversity II: Understanding and Protecting Our Biological Resources*. Washington, DC: Joseph Henry Press.
Kellert SR (1986) Social and perceptual factors in the preservation of animal species. In: Norton B (ed.) *The Preservation of Species: The Value of Biological Diversity*. Princeton: Princeton University Press.
Kellert SR (1996) *The Value of Life: Biological Diversity and Human Society*. Washington, DC: Island Press.
Kohm K (ed.) (1991) *Balancing on the Brink of Extinction: The Endangered Species Act and Lessons for the Future*. Washington, DC: Island Press.
Leopold A (1949/1970) *A Sand County Almanac*. San Francisco: Sierra Club.
Lovejoy TE (1992) Earth's living library: Check it out. *Washington Post*, 19 March, A27.
Mann CC and Plummer ML (1995) *Noah's Choice: The Future of Endangered Species*. New York: Knopf.
Marsh GP (1864/1965) *Man and Nature*. New York: Charles Scribner.
Myers N (1979) *The Sinking Ark: A New Look at the Problem of Disappearing Species*. Oxford: Pergamon Press.
Nabhan GP (1997) *Cultures of Habitat: On Nature, Culture, and Story*. Washington, DC: Counterpoint.
Norton B (1987) *Why Preserve Natural Variety?* Princeton: Princeton University Press.
Peters RL and Lovejoy TE (eds.) (1992) *Global Warming and Biological Diversity*. New Haven: Yale University Press.
Reid W, Barber C, and Miller K (1992) *Global Biodiversity Strategy: Guidelines for Action to Save, Study, and Use Earth's Biotic Wealth Sustainably and Equitably*. Washington, DC: World Resources Institute, World Conservation Union, and United Nations Development Program.
Soulé M (1985) What is conservation biology? *BioScience* 35: 727–734.
Stork NE (1997) Measuring global diversity and its decline. In: Reaka-Kudla ML, Wilson DE, and Wilson EO (eds.) *Biodiversity II: Understanding and Protecting our Biological Resources*. Washington, DC: Joseph Henry Press.
Takacs D (1996) *The Idea of Biodiversity*. Baltimore: Johns Hopkins University Press.
Takacs D (2010) Forest carbon offsets and international environmental law. *22 Georgetown International Environmental Law Review* 521.
Tilman D (1999) The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80(5): 1455–1474.
Williams R (1980) Ideas of Nature. *Problems in Materialism and Culture*. London: Verso.
Wilson EO (1987) The little things that run the world (the importance and conservation of invertebrates). *Conservation Biology* 1: 344–346.
Wilson EO (ed.) (1988) *BioDiversity*. Washington, DC: Smithsonian.
Wilson EO (1992) *The Diversity of Life*. Cambridge: Harvard University Press.
Wilson EO (1998) *Consilience: The Unity of Knowledge*. New York: Alfred A. Knopf.
Worster D (1987) *Nature's Economy: A History of Ecological Ideas*. Cambridge: Cambridge University Press.

Hotspots

Ana SL Rodrigues, Centre d'Ecologie Fonctionnelle et Evolutive, Montpellier, France

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Norman Myers, volume 3, pp 371–381, © 2001, Elsevier Inc.

Glossary

Biodiversity (or biological diversity) The full diversity of life on Earth, at all levels of biological organization, from molecules to genes, to populations, to species, and to ecosystems. It is a multidimensional concept difficult to measure, and in practice most focus is on diversity at the species level.

Biogeographic regions Large regions with characteristic assemblages of natural communities and species, distinct from other regions. There are several classifications of biogeographic regions, for example, into ecoregions and ecozones.

Endemic A species is said to be endemic to a given area when it cannot be found anywhere else. An area of high endemism may either have high numbers of species that only exist there or high concentrations of species with small ranges.

Irreplaceability The irreplaceability of a region is the degree to which that region is necessary as part of a comprehensive conservation strategy. A region with high levels of endemism has high irreplaceability because it cannot be replaced by others if the endemic species are to be conserved.

Threatened species Species at risk, either of global extinction (globally threatened species) or of extirpation within a given region or country (regionally or nationally threatened species). The most authoritative reference for the identification of globally threatened species is the International Union for Nature Conservation (IUCN) Red List of Threatened Species. It distinguishes three categories of threat (in increasing order of risk): Vulnerable, Endangered and Critically Endangered.

Web of Science An online academic citation index provided by Thomson Reuters. It provides access to a diversity of databases of peer-reviewed scientific publications, which can be searched using multiple criteria (for example, particular keywords within particular years).

Introduction

The term "hotspot" is commonly used in science – as well as in everyday life – to refer to areas of high activity in, or importance to, something. In science, it is used in genetics (e.g., "recombination hotspots"; Smagulova *et al.*, 2011), geophysics (e.g., "hotspot volcanism"; Hieronymus and Bercovici, 2001), and in environmental sciences (e.g., "pollution hotspots"; Spiteri *et al.*, 2005). Within ecology, there are, for example, hotspots of population abundance (e.g., Nur *et al.*, 2011), of animal activity (e.g., Okkonen *et al.*, 2011), and of disease emergence (e.g., Caron *et al.*, 2011).

The term "biodiversity hotspots" refers to areas of high concentration of biodiversity and/or high conservation value. Within this general meaning, multiple definitions have emerged, with different conceptual and practical implications. The term is now widely used in the biodiversity conservation literature (e.g., a Web of Science search for "biodiversity AND hotspot*" yields over 2000 citations), but, perhaps because it seems so intuitive, it is frequently used without being defined, its multiple meanings are often confused, and the resulting literature is replete with ambiguity and misunderstandings. This article reviews and clarifies the definition, history, usage and impact of, and criticism toward, the various concepts of biodiversity hotspots.

Different types of Biodiversity Hotspots

Two main versions of the concept of biodiversity hotspots have become particularly influential in conservation theory and practice. One (referred throughout as "biodiversity hotspots *sensu* Myers/Conservation International (CI)") corresponds to large biogeographic regions that have exceptional levels of plant endemism (>1500 plant species as endemics) and exceptional levels of threat (have lost more than 70% of their original vegetation cover) (Myers *et al.*, 2000, 2001). These correspond to very specific areas of the planet, even if their number and precise boundaries have been evolving over time. This concept has been extended to marine systems to identify regions of exceptional levels of endemism and of threat in tropical reefs (Roberts *et al.*, 2002).

The second main usage of biodiversity hotspots corresponds to areas with particularly high numbers of species (designated below as "biodiversity hotspots *sensu lato*"). Rather than referring to specific areas, this concept is associated with a methodology for highlighting sites of conservation interest within a given study region, and for a given taxonomic group. It was first used to refer to the top 5% sites in Britain with the highest richness in plant and animal species (Prendergast *et al.*, 1993), but has subsequently been extended, for example, to highlight areas with high concentrations of rare (Williams *et al.*, 1996) or threatened species (Dobson *et al.*, 1997), of a diversity of taxonomic groups, at various spatial scales, and following multiple variations on the precise method for selecting the top sites.

Biodiversity Hotspots *sensu* Myers/CI

Myers first proposed the concept of biodiversity hotspots in 1988 (Myers, 1988), defining them as regions that both feature

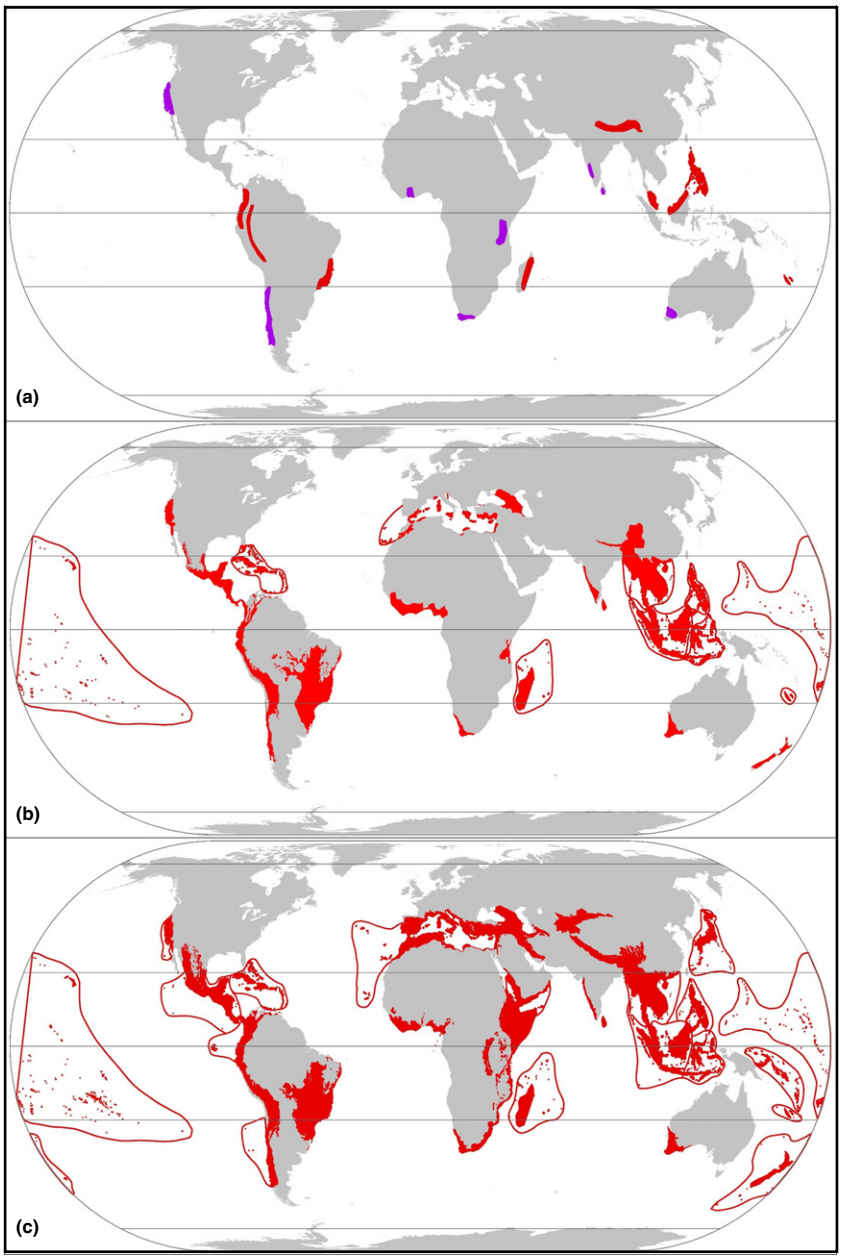
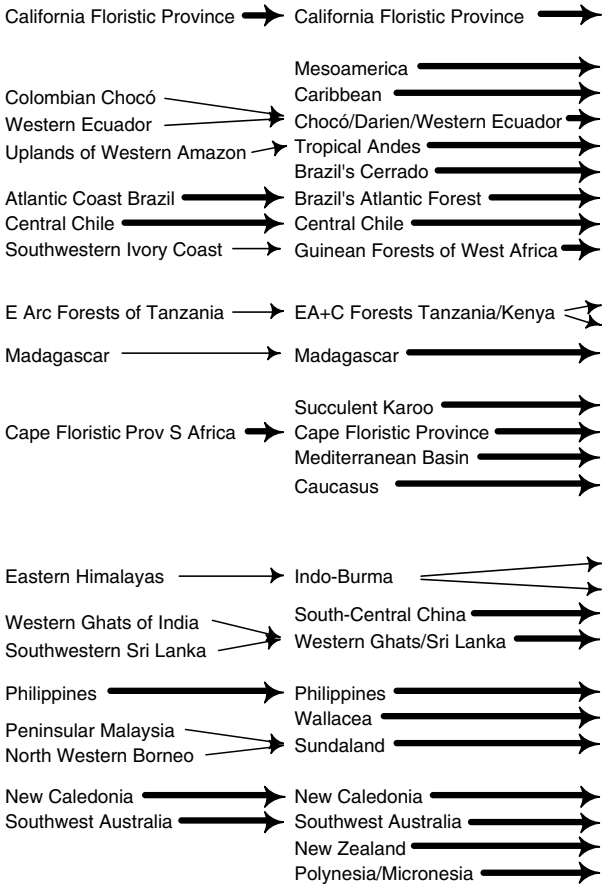


Figure 1 The temporal evolution of the spatial coverage of the biodiversity hotspots *sensu* Myers/Conservation International: (a) in 1988 (red; adapted from Myers (1988) Threatened biotas: "hot spots" in tropical forests. *Environmentalist* 8: 187–208, with permission from University of Chicago Press); and in 1990 (purple; adapted from Myers (1990) The biodiversity challenge: Expanded hot-spots analysis. *Environmentalist* 10: 243–256, with permission from University of Chicago Press); (b) in 1999–2000 (adapted with permission from Mittermeier RA, Meyers N, Robles Gil P, and Mittermeier CG (1999) *Hotspots – Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Chicago, USA: CEMEX/ Conservation International and the University of Chicago Press; and Myers N, Mittermeier RA, Mittermeier CG, Fonseca, GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858); and (c) in 2004 (adapted with permission from Mittermeier RA, Robles Gil P, Hoffmann M, *et al.* (2004) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Mexico City, Mexico: CEMEX, Conservation International and Agrupación Sierra Madre). In (a) and (b) the red lines correspond to the outer limits of the Hotspots, encompassing groups of islands.

exceptional concentrations of species with exceptional levels of endemism, and that face exceptional degrees of threat (**Figures 1(a)** and **2**). Although no thresholds were proposed then for defining exceptional endemism or threat, Myers made it clear that it was the spatial coincidence of these two criteria (and not just one or the other) that made hotspots global conservation priorities for preventing species extinctions. He focused on vascular plant endemism, but predicted these areas would also capture disproportional fractions of global animal diversity, and he evaluated threat as the fraction of original forest estimated to have been lost from each region. The 1988 paper identified 10 tropical forest hotspots (Myers, 1988), complemented 2 years



Original extent (1000 km²)	Remaining habitat (%)	Endemic plants
294	25	2124
461	20	3975
1130	20	2941
230	10	6550
275	24	2750
1543	25	15000
2032	22	4400
1234	8	8000
397	30	1957
620	15	1800
1659	5	2750
1018	11	2356
291	10	1750
600	10	11600
274	25	1900
103	29	2439
79	20	6210
2085	5	11700
533	27	1600
900	15	2500
863	20	1500
741	25	3160
2373	5	7000
262	8	3500
190	23	3049
373	20	1950
297	7	6091
338	15	1500
1501	7	15000
99	30	3000
19	5	2432
357	30	2948
270	22	1865
47	21	3074

Figure 2 The temporal evolution of the list of biodiversity hotspots *sensu* Myers/CI (thick arrows, strong correspondence; thin arrows, major change in boundaries). Key statistics are presented for each of the 34 Hotspots in the 2004 version. Reproduced with permission from Mittermeier RA, Robles Gil P, Hoffmann M, *et al.* (2004) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Mexico City, Mexico: CEMEX, Conservation International and Agrupación Sierra Madre.

later by eight additional areas, four in tropical forests, and four in mediterranean biomes (Myers, 1990; **Figures 1(a)** and **2**).

In 1989, the concept of biodiversity hotspots was embraced by CI, the US-based nongovernmental international conservation organization (Mittermeier *et al.*, 1999). In collaboration with Myers, CI scientists refined the criteria for identifying hotspots by establishing quantitative thresholds for endemism (at least 0.5% or 1500 of the world's 300,000 vascular plant species) and for threat (loss of at least 70% of the primary vegetation). CI placed hotspots at the core of their global conservation strategy, complemented by High Biodiversity Wilderness Areas (what Myers had called "good news areas"; Myers, 1988) – regions of similarly high levels of plant endemism (>1500 plant species) but with most of their original vegetation still intact ($<30\%$ lost) (Mittermeier *et al.*, 1998, 2002, 2003b). A revised set of 24 hotspots was proposed in 1998 (Mittermeier *et al.*, 1998), immediately increased to 25 (Mittermeier *et al.*, 1999; Myers *et al.*, 2000, 2001) (**Figures 1(b)** and **2**). These 25 hotspots were estimated to hold 44% of all species of vascular plants and 35% of all species of terrestrial vertebrates as endemics. Their total area corresponded to 11.8% of the Earth's land surface, but their remaining original vegetation was estimated to occupy only 1.4% of the planet. Myers *et al.* (2000) advocated focusing conservation efforts on these 1.4%, as a "silver bullet" strategy to prevent species extinctions worldwide. This strategy recognizes that such concentrations of endemic species in hotspots are a conservation opportunity, but they also reflect a very high jeopardy: a large fraction of global plant diversity is potentially at risk of extinction by being restricted to regions that have high levels of threat.

The latest list of 34 biodiversity hotspots was published in a 2004 revision (Mittermeier *et al.*, 2004) (**Figures 1(c)**, and **2**). The quantitative criteria remained the same, but improved data on plant endemism and on habitat loss revealed six previously overlooked areas. Boundaries were redefined in order to align hotspots with groups of ecoregions (large biogeographic units with distinct assemblages of species and communities; Olson *et al.*, 2001), resulting in two of the 1999–2000 hotspots being split. More worryingly, one region (East Melanesia) was identified as a newly formed hotspot: despite being considered part of a Wilderness Area in 1998 (Mittermeier *et al.*, 1998), it was found to have lost 70% of its natural forest by 2004. Overall, the 34 hotspots correspond to

15.7% of the global land surface, their original habitat covering only 2.3%. Between them, they are estimated to hold at least 150,000 plant species as endemics (50% of the world's total), and an overall 22,022 terrestrial vertebrate species (77% of the world's total) including 11,980 species as endemics (42% of the world's total).

The list of biodiversity hotspots *sensu* Myers/CI is therefore dynamic, and additional updates are expected as knowledge improves on the distribution of species and on habitat loss, and as landuse changes worldwide. The 2004 update highlighted two regions (Queensland Wet Tropics and Taiwan) that came close to meeting the hotspots criteria, and mentioned three others (Angola Escarpment, rainforests of eastern Australia, and southeastern China) likely to be considered in future revisions (Mittermeier *et al.*, 2004).

Roberts *et al.* (2002) extended the concept of biodiversity hotspots *sensu* Myers/CI to marine systems, by identifying areas of high levels of endemism and of high threat across the world's tropical reefs . Using data on the distribution of 3235 species of reef fish, corals, snails, and lobsters, they mapped species endemism across a 50,000 km^2 degree grid by calculating the sum in each cell of the reciprocal of the range size of each species. They then identified 18 multitaxa centers of endemism from areas that were in the top 10% of the cells for at least two taxa. Finally, 10 of these regions considered under high threat were identified as marine biodiversity hotspots. Threat was assessed using a map-based indicator of risk to coral reefs from coastal development, marine pollution, overexploitation, and inland pollution (Bryant *et al.*, 1998).

Biodiversity Hotspots *sensu lato*

The concept of biodiversity hotspots as areas with high concentration of species was (possibly) first proposed by Prendergast *et al.* (1993) (who cited Myers (1988) as the source for the term, but then used it in a different sense). Their analysis was based on the distribution of butterflies, dragonflies, aquatic plants, breeding birds, and liverworts in Britain, and hotspots were defined as the top 5% cells (10 km^2) with the highest species richness for each taxon. Later in that same year, Pomeroy (1993) independently applied the term hotspots to refer to African countries with a high diversity in plant species, after controlling for variation in country area using the species–area relationship.

The concept of hotspots was extended to areas with high concentrations of rare species by Williams *et al.* (1996), who distinguished between "richness hotspots" (the top 5% sites in terms of diversity) and "rarity hotspots" (the top 5% sites in richness of rare species, defined as the 25% species with the smallest ranges). BirdLife International's Endemic Bird Areas are a global application of the concept of rarity hotspots: there are 218 regions where the distributions of two or more restricted-range bird species (with ranges smaller than 50,000 km^2) overlap (Stattersfield *et al.*, 1998).

The concept of hotspots was further extended to refer to concentrations of threatened species (Dobson *et al.*, 1997), dubbed elsewhere as "threatspots" (Troumbis and Dimitrakopoulos, 1998). The term hotspots has since become a common way of designating areas with high richness of all, rare, or threatened species, either in absolute terms (Orme *et al.*, 2005; Ceballos and Ehrlich, 2006; Grenyer *et al.*, 2006), or after controlling for variation in the area of the analyzed units (e.g., Veech, 2000; Guilhaumon *et al.*, 2008).

The concept of hotspot continues to expand as it is applied to other facets of biodiversity, for example, through the designation of hotspots of functional diversity (Lucifora *et al.*, 2011), or hotspots of beta diversity (Harborne *et al.*, 2006). There is however one key aspect in which hotspots *sensu lato* remain conceptually distinct from hotspots *sensu* Myers/CI: the former apply to peaks of diversity along a particular axis of diversity or of threat, whereas the latter highlights areas that have both high endemism and high threat. The two approaches can meet if hotspots *sensu lato* are defined as concentrations of species that are simultaneously threatened and spatially rare.

Biodiversity Hotspots in the Literature

A review of the usage of the term "biodiversity hotspots" in scientific journals illustrates how the various concepts have permeated the literature. A Web of Science search (June 2011) for publications on biodiversity and hotspots (search term: Topic=((hotspot* OR hot-spot* OR hot spot*) AND biodiversity) in four journals (three high impact journals, *Nature*, *Science*, and *Proceedings of the National Academy of Sciences USA*; and the highest impact journal on biodiversity conservation, *Conservation Biology*) yielded 180 articles. Of these, 23 were excluded: 17 for not using the term "hotspot" at all; four for being news articles without references; and two for using the term in a different sense from biodiversity hotspots (pollution hotspots, Grimm *et al.*, 2008; groundwater recharge hotspots, Herbert *et al.*, 2010). The remaining 157 articles were analyzed. Testifying the extent to which it has become a glamorous concept in the literature, a third of these articles (51/157) used the term hotspots in the title. The concept was used in four (partially overlapping) main ways (**Figure 3**): hotspots *sensu* Myers/CI; hotspots *sensu lato*; following other definitions; and without being defined. These are discussed below.

Biodiversity Hotspots *sensu* Myers/CI

The concept of biodiversity hotspots *sensu* Myers/CI has had substantial impact in the scientific literature. Indeed, the 2000 *Nature* paper (Myers *et al.*, 2000, now with over 3500 citations) is not only the most cited publication on biodiversity hotspots (**Figure 4**), but also one of the most cited studies in *Ecology* (was the most cited paper in *Environment and Ecology* in the 1998–2008 period; Thomson Reuters (2008)). Analysis of a sample of articles provides a deeper perspective of the way in which the concept of hotspots *sensu* Myers/CI has permeated the literature (**Figure 3**). Out of 157 articles analyzed, 99 (63%) used the term hotspots in this sense, including 88 who use it exclusively in this way. The dominant usage (2 in 5 of the 99 papers) was to highlight the value of the paper's study area as a global biodiversity hotspot (e.g., Bossuyt *et al.* (2004) and Delibes-Mateos *et al.* (2008)). Conversely, in a separate search for conservation articles focusing on five regions classified as biodiversity hotspots *sensu* Myers/CI, more than half

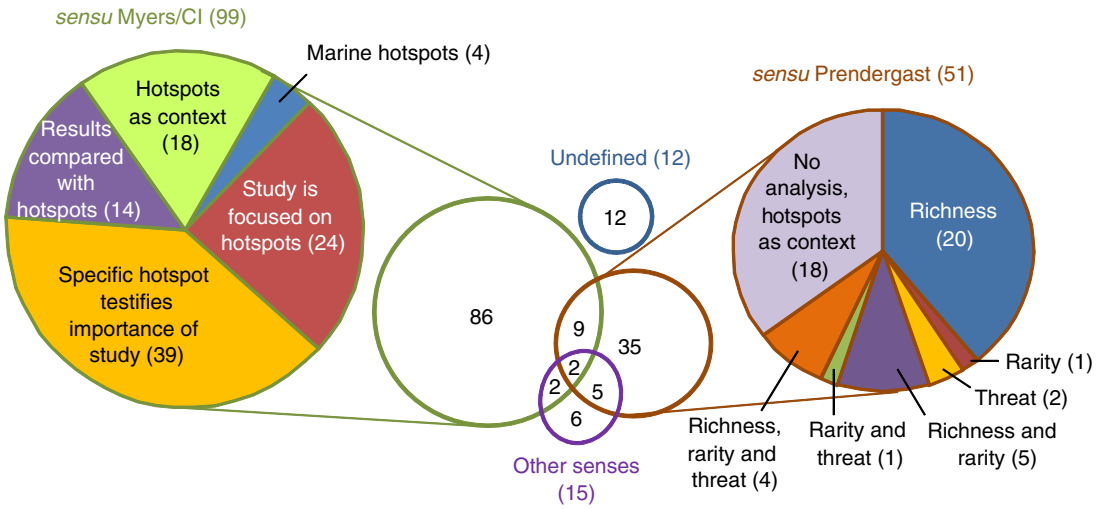


Figure 3 A review of the multiple uses of the term "hotspot" in the biodiversity literature. Data were obtained from a sample of 157 articles on biodiversity and hotspots, published in *Nature*, *Science*, *Proceedings of the National Academy of Sciences USA*, and *Conservation Biology*. They were classified into four (partially overlapping) main groups: those that use the term hotspots *sensu* Myers/CI; those that use it *sensu lato*; those that used it another way; and those that used it in an undefined way. The 99 papers on hotspots *sensu* Myers/CI were further classified into studies on marine hotspots (Roberts *et al.*, 2002) and studies on terrestrial hotspots, with the latter divided into: studies where hotspots themselves were units or subjects of analysis; papers whose study area was within a specific hotspot, this fact being mentioned as evidence of the relevancy of the study to global biodiversity conservation; papers where results were compared to hotspots; and papers where hotspots are mentioned as context. The 51 studies on hotspots *sensu lato* were split into those where the concept is mentioned merely as context, and those where analyses are done to identify hotspots of richness, of rarity, or of threat (or combinations of these three types).

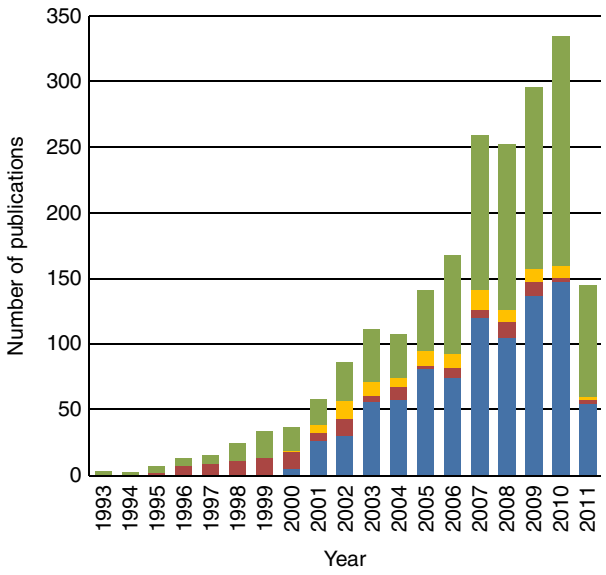


Figure 4 Number of peer-reviewed papers on biodiversity hotspots according to whether they cite Myers *et al.* (2000; blue), or Prendergast *et al.* (1993; red), both (yellow), or none (green). Papers were found through Web of Science using the search term topic=(biodiversity AND hotspot*) (June 2011).

of the papers highlighted the global importance of their study areas by referring to the hotspot designation (Web of Science search for Topic=("Atlantic Forest" OR Cerrado OR Madagascar OR Philippines OR "Western Ghats") AND Publication Name=Conservation Biology; September 2011; 63 references found, of which 27 had one of the five hotspots as a study-area, of which 15 refer to their study area as a biodiversity hotspot *sensu* Myers/CI). This suggests the concept of global biodiversity hotspots *sensu* Myers/CI has been widely embraced by conservation scientists working in these regions.

The second main usage of the concept of hotspots *sensu* Myers/CI (25/99) was in studies that focused on the hotspots themselves, either as spatial units or as subjects of analysis (**Figure 3**), for example, by comparing hotspots to other global conservation priorities (Brooks *et al.*, 2002), by quantifying the human impact on the hotspots (Cincotta *et al.*, 2000), or by investigating the degree to which hotspots represent evolutionary history (Sechrest *et al.*, 2002). In a sizeable fraction of papers using the concept of hotspots *sensu* Myers/CI (14/99), the study's results were compared to hotspots, sometimes as a particular analysis, others as part of the discussion (e.g., Cardillo *et al.*, 2006 compared their predictions of future areas of high extinction risk with Myers/CI biodiversity hotspots), whereas in other publications (18/99) hotspots were mentioned to provide context, typically in the introduction (e.g., Lamoreux *et al.*, 2006 mentions hotspots as one approach that is based on ecoregions as spatial units). Finally, a few (4/99) studies focused on marine hotspots *sensu* Roberts *et al.* (2002).

Biodiversity Hotspots *sensu lato*

Although usage of the term hotspots *sensu* Myers/CI was tightly linked to the respective scientific publications (in 92/99 cases), only 21 out of 51 studies that used the concept of hotspots *sensu lato* quoted Prendergast (1993). And yet, most studies undertook analyses in Prendergast's original sense of

peaks of species richness (29/51), with hotspots of rarity and of threat analyzed, respectively, in 10/51 and 7/51 of studies (**Figure 3**). This approach has been applied to a variety of taxa (e.g., moths, Grand *et al.*, 2004; beetles, Fattorini, 2006), of regions (e.g., South Africa, Maddock and Benn, 2000; Borneo, Meijaard and Nijman, 2003), and of spatial scales (e.g., from 2 km^2 landscapes, Gjerde *et al.*, 2004; to global, Orme *et al.*, 2005). The exact method for identifying the hotspots has also varied, including changes in the thresholds (e.g., the 10% top cells rather than the 5%; Maddock and Benn, 2000) and correction for area when the spatial units vary in size (e.g., Fattorini, 2006).

The other dominant usage of the concept of hotspots *sensu lato* (18/51) was in providing context, for example, by designating concentrations of species as hotspots (e.g., Mace *et al.*, 2003 referred to the description of 140 new amphibian species from Sri Lanka as "a new hotspot of diversity") or by describing results from previous studies (e.g., Lawler *et al.*, 2003 mentioning the lack of hotspot congruence across taxa found by Prendergast).

Other Concepts of Biodiversity Hotspots

A small, but nonnegligible, fraction of the studies analyzed (15/157) proposed other concepts of biodiversity hotspots (**Figure 3**). A third of these (5/157) corresponded to "complementarity hotspots" (perhaps first coined by Araujo and Williams, 2001). These are sets of sites with a complementary species composition, selected not for their exceptional species richness as individual sites, but because together they maximize species representation (Margules and Pressey, 2000). However, this usage of the hotspots term is very unusual within the large literature on complementary reserve selection (*see* Hotspots and Systematic Conservation Planning).

The other concepts of biodiversity hotspots were: areas predicted to concentrate future species extinctions ("latency hotspots"; Cardillo *et al.*, 2006); sites of high environmental diversity ("environmental hotspots"; Chen, 2009); areas of active deforestation ("forest clearing hotspots"; Hansen *et al.*, 2008); areas of high vulnerability to traditional livelihoods are where agroecological systems and knowledge are threatened ("rural hotspots"; Harvey *et al.*, 2008); areas of high population abundance (Winston and Angermeier, 1995); areas where biodiversity is being generated ("evolutionary hotspots of diversification"; Petit *et al.*, 2003); areas of intense reciprocal selection ("coevolutionary hotspots"; Thompson and Cunningham, 2002); and areas of intense conservation activity ("conservation planning hotspots"; Reyers *et al.*, 2010).

Ambiguity in Usage

More than one in four of the 157 studies analyzed applied the term "hotspots" ambiguously. There were three main sources of ambiguity: the term "hotspot" was simply left undefined (in 13/157 studies; e.g., Lindenmayer and Hunter, 2010); the term was used to refer to multiple concepts without clarifying the differences (24/157; e.g., Orme *et al.*, 2005 used hotspots in both *sensu* Myers/CI and *sensu lato*, but treated them as equivalent); or the concept was defined and a reference provided, but the two did not match (5/157; e.g., Bellwood *et al.*, 2004 defines hotspots as "areas of exceptional species richness" while citing Myers *et al.*, 2002).

Ambiguity has been prevalent in the hotspots literature since its early days, starting with Prendergast *et al.* (1993), who credited Myers (1988) for the term "hotspots" while defining them as "species-rich areas". Unfortunately, this ambiguity has been at the root of substantial confusion, particularly when discussing the value and adequacy of hotspots as tools in conservation planning and action.

Biodiversity Hotspots in Conservation Planning and Action

Hotspots and Systematic Conservation Planning

The two main concepts of hotspots are closely intertwined – but in very different ways – with the broader literature on the selection of priority areas for conservation. This literature has expanded substantially since the 1980s, as the field of systematic conservation planning emerged in response to the recognition that the limited resources available to conservation need to be allocated strategically to maximize the cost effectiveness of conservation action (Kirkpatrick, 1983; Pressey *et al.*, 1993; Margules and Pressey, 2000). A key principle in systematic conservation planning is that of complementarity (Pressey *et al.*, 1993): networks of priority areas for conservation should be selected such that sites best complement each other efficiently in terms of their biodiversity composition.

The complementarity approach was explicitly proposed as an alternative to the identification of areas based on ranking (or scoring) algorithms, where the conservation value of all sites is based on a single formula, with areas sorted in descending value order and the top ones being selected (Kirkpatrick, 1983). The drawback of the ranking approach is that the top sites often capture similar mixes of species, communities or habitats, thus being highly redundant among them, and not adequately representing the overall diversity across the region. The concept of biodiversity hotspots *sensu lato* is a type of ranking system, with areas sorted in terms of their overall diversity or richness in rare species, and the top sites (hotspots) being selected. Multiple studies have found that complementarity sets are much more efficient at representing the overall diversity of a region than hotspots *sensu lato* (e.g., Kirkpatrick, 1983; Pressey and Nicholls, 1989; Williams *et al.*, 1996; not all of these studies used the term "hotspots" to designate the top areas selected through the ranking approach). The overall consensus in the conservation planning literature has thus been to discourage the use of hotspots *sensu lato* as a practical conservation planning tool. Hotspots *sensu* Myers/CI have not been designed explicitly to maximize complementarity (are not necessarily the most efficient way of maximizing overall diversity within a given area) but they are by definition complementary among themselves. Indeed, each hotspot has at least 1500 endemic plants (up to 15,000; **Figure 2**) and is therefore not redundant with the other regions.

Another key concept in the systematic conservation planning literature is that of irreplaceability: the extent to which the spatial options for conservation are lost if a given area is

lost (Pressey *et al.*, 1993). Irreplaceability is intrinsically connected to complementarity, because the more irreplaceable a site is the more likely that it needs to be selected as parts of a complementary set representing the overall diversity of the region. For example, if a species occurs in only one site, this is a fully irreplaceable area that will inevitably be part of all complementary networks representing all species. A complementary set often also includes low irreplaceability sites, but these are flexible, in the sense that they can be replaced by others in the region (Pressey *et al.*, 1993; Williams *et al.*, 1996). Hotspots of richness or of threat *sensu lato* are not necessarily highly irreplaceable (if the species they hold are widespread and can be conserved in other areas) but hotspots of rarity typically are. And, accordingly, hotspots *sensu* Myers/CI are regions that are highly irreplaceable at the global scale, as each one of them holds large numbers of species that cannot be found elsewhere.

Conservation implementation typically takes place over a long period of time, during which threats continue to operate and conservation opportunities may be lost. It is therefore crucial to decide not just which areas should be conserved, but which ones must be conserved first (Pressey and Taffs, 2001). Options for future conservation can be compromised in spatial as well as in temporal terms. Irreplaceability is inversely related to the degree of options in space: if a highly irreplaceable site is destroyed or damaged, it will be difficult (or impossible) to find other sites that can replace its conservation value. Site threat (or vulnerability), however, is inversely related to options in time: there is little time to conserve a site that is at high risk, opportunities must be seized soon or they will disappear. Pressey and Taffs (2001) thus proposed giving higher priority in conservation scheduling to areas of simultaneous high irreplaceability and high threat in order to minimize the extent to which conservation targets are compromised before they can be met in new conservation areas. Hotspots *sensu* Myers/CI correspond to regions of simultaneous high irreplaceability (many endemic species that cannot be protected elsewhere) and high threat (past levels of habitat loss presumed to indicate high risk into the future) (Margules and Pressey, 2000; Brooks *et al.*, 2006) and are therefore regions where conservation is urgent if global extinctions are to be avoided. It was this sense of urgency that led Myers to first propose the concept and CI to subsequently adopt it. Hotspots *sensu lato*, however, are less directly related to the extent of spatial or temporal options for conservation, particularly for diversity hotspots. Rarity hotspots correspond to areas of high irreplaceability, and threat hotspots to areas of high threat, but they do not capture the double jeopardy of simultaneously high irreplaceability and high threat.

Conservation Impact of Hotspots

Publications on hotspots *sensu lato* have been mainly academic, rather than proposing specific conservation priorities. Nonetheless, and despite the recent emphasis on complementary reserve selection, having high levels of species richness continues to be a selling point in conservation, and is a near-universal claim in the marketing materials of particular sites (e.g., "home to a stunning array of animals" in Australia's Great Barrier Reef; Great Barrier Reef Marine Park Authority, 2011). In some cases, species richness is still a formal criterion for site selection, for example, in identifying important plant areas (Anderson, 2002), even though the term "hotspots" is not always applied. Whereas it is conceivable that this label may increase the appeal of particular sites, it is not straightforward to quantify the extent to which the concept of hotspots *sensu lato* has had a real impact on biodiversity conservation worldwide.

Hotspots *sensu* Myers/CI, however, have from the start been proposed and advocated as very specific priority regions for conservation investment. Many of these regions are conservation priorities under multiple other global conservation schemes (such as BirdLife's Endemic Bird Areas and WWF's Global 200; Brooks *et al.*, 2006), and it is not easy to quantify the additional value of the "biodiversity hotspots" label for their conservation. Nonetheless, it is clear that the concept of biodiversity hotspots *sensu* Myers/CI has had a substantial impact in mobilizing globally flexible conservation investment from multilateral agencies, bilateral aid, private individuals, and corporations (Myers, 2001). In 1989, the same year Conservation International adopted hotspots as its central strategy, the Chicago-based John D. and Catherine T. MacArthur Foundation adopted hotspots as its primary global investment strategy (Mittermeier *et al.*, 2004). For the decade 2000–2010, they focused on eight hotspots (Madagascar, Melanesia, the Andes, Insular Caribbean, Albertine Rift, Eastern Himalayas, and Lower Mekong), awarding over 400 grants for US$ 110 million (Wells *et al.*, 2010). In 2000, the World Bank and the Global Environment Facility joined CI to establish the Critical Ecosystem Partnership Fund (CEPF), a conservation finance mechanism focused explicitly on the hotspots (Dalton, 2000). They were subsequently joined by the MacArthur Foundation, the Japanese Government, and the French Development Agency. CEPF grant attribution in each hotspot is based on an "ecosystem profile", which sets conservation targets, identifies major threats, and describes the policy, civil society and socioeconomic context, as well as funding gaps and opportunities. As of March 2011, the CEPF had committed US$ 131 million, and leveraged a further US$ 318 million, to over 1600 partners, spanning 53 countries and territories in 20 hotspots (Conservation International, 2011). CI's Global Conservation Fund (GCF) also uses hotspots (along with High Biodiversity Wilderness Areas) to guide its investments. Funded by a 2001 US$ 100 million grant from the Gordon and Betty Moore Foundation, it finances the creation, expansion, and long-term management of protected areas. As of March 2010, the GCF had committed US$ 36 million in grants to biodiversity hotspots, to fund the creation and/or expansion of 33 protected areas, covering 7.1 million hectares (Conservation International, 2010). Overall, more than $750 million have been devoted to biodiversity hotspots *sensu* Myers/CI, perhaps the largest financial investment in any single conservation strategy (Mittermeier *et al.*, 2004).

Criticism

Bodiversity hotspots have attracted substantial criticism, generally directed specifically at one or the other of the two main

types (e.g., Williams *et al.*, 1996; Mace *et al.*, 2000; Brummitt and Lughadha, 2003; Kareiva and Marvier, 2003; Ovadia, 2003), but sometimes in a broader and more ambiguous way (e.g., Orme *et al.*, 2005; Ceballos and Ehrlich, 2006). The main criticisms are reviewed here.

Criticism of the Hotspots *sensu* Myers/CI

Too Large

Hotspots *sensu* Myers/CI have been criticized for being much larger than the scale at which conservation decisions take place (Mace *et al.*, 2000; Brummitt and Lughadha, 2003). However, like other large-scale priority schemes, hotspots are not proposed as areas that should be entirely set aside for conservation (e.g., to be designated as protected areas), but as priority regions for concentrating geographically flexible global conservation investment (from multilateral agencies, bilateral aid, private corporations, foundations and individuals; Brooks *et al.*, 2006). Hotspots have attracted substantial funding (*see* Conservation Impact of Hotspots) but their ultimate conservation impact will depend on how funds are employed within each hotspot. Disbursement of conservation grants for any given hotspot under the Critical Ecosystem Partnership Fund is done after a finer scale analysis (termed an "ecosystem profile") that identifies conservation priorities at the site scale (Key Biodiversity Areas, based on the presence of rare and/or threatened species; Eken *et al.*, 2004).

Duplication of Efforts and Lack of Consensus

Hotspots *sensu* Myers/CI and other large-scale conservation schemes have been criticized for duplicating efforts and producing competing rather than consensual proposals (Mace *et al.*, 2000). Indeed, at least nine global schemes have been proposed as global conservation priority templates, championed by different conservation organizations, following different criteria, and resulting in distinctive but partially overlapping spatial configurations (see Brooks *et al.*, 2006 for a review). Arguably their costs in terms of data acquisition, analyses, publication, and advertisement could have been pooled to produce a better and more consensual global conservation agenda, with a higher probability of adoption by decision makers (Mace *et al.*, 2000). However, different schemes may appeal to different audiences of individuals and organizations (e.g., Endemic Bird Areas to bird lovers; Stattersfield *et al.*, 1998; Frontier Forests to those concerned with carbon sequestration; Bryant *et al.*, 1997), and so multiple schemes may raise more conservation investment than a single consensual scheme would. The different priority schemes can also be seen as cooperating to attract investment to particularly important regions. Indeed, despite their apparent dissimilarity, there is substantial consensus between schemes once a distinction is made between those emphasizing regions requiring an urgent reaction to threats to biodiversity (e.g., Madagascar and Philippines) and those favoring a proactive investment in still mainly intact regions (e.g., Amazon and New Guinea; Brooks *et al.*, 2006).

Focus on a too Small Area of the Planet (Triage)

One of the strongest selling points of hotspots *sensu* Myers/CI is that they concentrate a large diversity of global plant diversity as endemics in a very small fraction of the Earth, but this has raised the criticism that the hotspot strategy advocates a sort of triage, ignoring the rest of the planet (the "coldspots"; Kareiva and Marvier, 2003), including entire ecosystems like the Arctic and the Serengeti that are not represented within hotspots. Hotspots, however, do not prescribe which areas of the planet should be written off from conservation strategies, they advocate which areas should be conserved first (i.e., which ones are priorities) for avoiding species extinctions (Mittermeier *et al.*, 2003a).

Criticism of the Hotspots *sensu lato*

Lack of Efficiency

As described above (*see* Hotspots and Systematic Conservation Planning), hotspots of species richness (and other ranking-based approaches for site prioritization that sort areas in descending value order, with the top ones being selected) have been criticized for being an inefficient way of representing the broader diversity of a given region, because they are often highly redundant in terms of their biological composition. Complementary-based approaches for the selection of networks of priority areas were proposed as a more efficient strategy (e.g., Kirkpatrick, 1983; Pressey and Nicholls, 1989; Williams *et al.*, 1996) and have since become firmly established in the conservation planning literature.

Sampling Bias

Patterns of species diversity are heavily influenced by sampling effort (Freitag *et al.*, 1998; Pautasso and McKinney, 2007), a pattern that may be self-reinforced as perceived high richness stimulates further attention and thus further sampling (Ahrends *et al.*, 2011). This problem affects all data-based conservation planning, and needs to be compensated by an improvement in the quality of the distribution data, or to be statistically corrected.

Broader criticism of Hotspots

The Different Types of Hotspots Do Not Match

Studies of distribution patterns of terrestrial vertebrates at the global scale found that peaks of species richness, of rarity and of threat (i.e., hotspots *sensu lato*) have reduced spatial overlap. This has been interpreted as evidence against allocating global conservation priorities based on hotspots in general (including, ambiguously, hotspots *sensu* Myers/CI; Orme *et al.*, 2005; Possingham and Wilson, 2005; Ceballos and Ehrlich, 2006) and generally bad news to conservation, because different areas are apparently needed to protect different types of species. In contrast, this lack of congruence can be interpreted as good news for conservation if conserving a maximum amount of all species diversity is the ultimate goal. Areas with high concentrations of rare species have limited spatial options for replacement, whereas areas with high levels of threat have limited options in time. Therefore, the greatest conservation challenges correspond to regions where these two

coincide – such as hotspots *sensu* Myers/CI – and it is therefore good news that a relatively small fraction of the planet suffers from this double jeopardy.

Hotspots for One Taxon are Not Necessarily Hotspots for Others

Multiple studies, at a diversity of scales, found that peaks of diversity for multiple taxa do not always overlap (Prendergast *et al.*, 1993; van Jaarsveld *et al.*, 1998; Grenyer *et al.*, 2006). As above (*see* The Different Types of Hotspots Do Not Match), results based on hotspots *sensu lato* have been extrapolated to criticize hotspots *sensu* Myers/CI, which have been assumed to be appropriate only for plant conservation (e.g., in Mace *et al.*, 2000, quoting results from van Jaarsveld *et al.*, 1998; Grenyer *et al.*, 2006). The lack of overlap between peaks of diversity for multiple taxa raises interesting macroecological and biogeographical questions, but it is not necessarily a conservation concern given the limited application of the concept of hotspots *sensu lato* to conservation theory and practice. In contrast, hotspots *sensu* Myers/CI have already mobilized substantial conservation investment, and if they are only effective for plant conservation that is a serious limitation to the hotspots strategy. Few studies have directly tested whether areas of simultaneously high endemism and threat for different taxa coincide, but a study investigating priority areas for expanding the global protected area network for birds, mammals, and amphibians, based on irreplaceability and threat, highlighted areas that were almost completely within the hotspots *sensu* Myers/CI (Rodrigues *et al.*, 2004).

Conclusions

More than two decades since it was first proposed, the concept of biodiversity hotspots is more popular than ever (**Figure 4**), but it means different things to different people. As a vague term for "important areas for biodiversity" it is no more than a sexy label, and widespread confusion about its different meanings can hinder its utility as a theoretical concept and as an applied tool. However, if rigorously defined it can have, and indeed it already has had, a significant impact in the conservation literature and practice.

See also: Biogeography, Overview. Comparing Extinction Rates: Past, Present, and Future. Complementarity. Conservation Biology, Discipline of. Conservation Efforts, Contemporary. Deforestation and Land Clearing. Endangered Ecosystems. Extinction, Causes of. Global Species Richness. Habitat Loss and Fragmentation. Human Impact on Biodiversity, Overview. Loss of Biodiversity, Overview. Species Diversity, Overview. Tropical Forest Ecosystems

References

Ahrends A, Burgess ND, Gereau RE, *et al.* (2011) Funding begets biodiversity. *Diversity and Distributions* 17: 191–200.

Anderson S (2002) *Identifying Important Plant Areas: A Site Selection Manual for Europe, and a Basis for Developing Guidelines for Other Regions of the World.* London, UK: PlantLife International.

Araujo MB and Williams PH (2001) The bias of complementarity hotspots toward marginal populations. *Conservation Biology* 15: 1710–1720.

Bellwood DR, Hughes TP, Folke C, and Nystrom M (2004) Confronting the coral reef crisis. *Nature* 429: 827–833.

Bossuyt F, Meegaskumbura M, Beenaerts N, *et al.* (2004) Local endemism within the Western Ghats – Sri Lanka Biodiversity Hotspot. *Science* 306: 479–481.

Brooks T, Mittermeier R, Mittermeier C, *et al.* (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology* 16: 909–923.

Brooks TM, Mittermeier RA, da Fonseca GAB, *et al.* (2006) Global biodiversity conservation priorities. *Science* 313: 58–61.

Brummitt N and Lughadha EN (2003) Biodiversity: Where's hot and where's not. *Conservation Biology* 17: 1442–1448.

Bryant D, Burke L, McManus J, and Spalding M (1998) *Reefs at Risk: A Map-based Indicator of Potential Threats to the World's Coral Reefs.* Washington DC, USA: World Resources Institute, International Center for Living Aquatic Resources Management, World Conservation Monitoring Centre, United Nations Environment Programme.

Bryant D, Nielsen D, and Tangley L (1997) *The Last Frontier Forests: Ecosystems and Economies on the Edge.* Washington DC, USA: World Resources Institute.

Cardillo M, Mace GM, Gittleman JL, and Purvis A (2006) Latent extinction risk and the future battlegrounds of mammal conservation. *Proceedings of the National Academy of Sciences USA* 103: 4157–4161.

Caron A, de Garine-Wichatitsky M, and Morand S (2011) Parasite community ecology and epidemiological interactions at the wildlife/domestic/human interface/: Can we anticipate emerging infectious diseases in their hotspots? *EcoHealth* 7: 16324.

Ceballos G and Ehrlich PR (2006) Global mammal distributions, biodiversity hotspots, and conservation. *Proceedings of the National Academy of Sciences USA* 103: 19374–19379.

Chen Y-H (2009) Combining the species–area–habitat relationship and environmental cluster analysis to set conservation priorities: A study in the Zhoushan Archipelago, China. *Conservation Biology* 23: 537–545.

Cincotta RP, Wisnewski J, and Engelman R (2000) Human population in the biodiversity hotspots. *Nature* 404: 990–992.

Conservation International (2010) Global Conservation Fund: Facts and Figures. URL http://www.conservation.org/sites/gcf/about/pages/factsandfigures.aspx.

Conservation International (2011) Critical Ecosystem Partnership Fund: Facts and Figures. URL http://www.cepf.net.

Dalton R (2000) Biodiversity cash aimed at hotspots. *Nature* 406: 818.

Delibes-Mateos M, Delibes M, Ferreras P, and Villafuerte R (2008) Key role of European rabbits in the conservation of the western Mediterranean basin hotspot. *Conservation Biology* 22: 1106–1117.

Dobson AP, Rodriguez JP, Roberts WM, and Wilcove DS (1997) Geographic distribution of endangered species in the United States. *Science* 275: 550–553.

Eken G, Bennun L, Brooks TM, *et al.* (2004) Key biodiversity areas as site conservation targets. *Bioscience* 54: 1110–1118.

Fattorini S (2006) Detecting biodiversity hotspots by species–area relationships: A case study of Mediterranean beetles. *Conservation Biology* 20: 1169–1180.

Freitag S, Hobson C, Biggs HC, and van Jaarsveld AS (1998) Testing for potential survey bias: The effect of roads, urban areas and nature reserves on a southern African mammal data set. *Animal Conservation* 1: 119–127.

Gjerde I, Saetersdal M, Rolstad J, Blom HH, and Storaunet KO (2004) Fine-scale diversity and rarity hotspots in northern forests. *Conservation Biology* 18: 1032–1042.

Grand J, Buonaccorsi J, Cushman SA, Griffin CR, and Neel MC (2004) A multiscale landscape approach to predicting bird and moth rarity hotspots in a threatened pitch pine-scrub oak community. *Conservation Biology* 18: 1063–1077.

Great Barrier Reef Marine Park Authority (2011) About the reef: Animals. URL http://www.gbrmpa.gov.au/about-the-reef/animals.

Grenyer R, Orme CDL, Jackson SF, *et al.* (2006) Global distribution and conservation of rare and threatened vertebrates. *Nature* 444: 93–96.

Grimm NB, Faeth SH, Golubiewski NE, *et al.* (2008) Global change and the ecology of cities. *Science* 319: 756–760.

Guilhaumon F, Gimenez O, Gaston KJ, and Mouillot D (2008) Taxonomic and regional uncertainty in species–area relationships and the identification of richness hotspots. *Proceedings of the National Academy of Sciences USA* 105: 15458–15463.

Hansen MC, Stehman SV, Potapov PV, *et al.* (2008) Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceedings of the National Academy of Sciences USA* 105: 9439–9444.

Harborne AR, Mumby PJ, Zychaluk K, Hedley JD, and Blackwell PG (2006) Modeling the beta diversity of coral reefs. *Ecology* 87: 2871–2881.

Harvey CA, Komar O, Chazdon R, *et al.* (2008) Integrating agricultural landscapes with biodiversity conservation in the Mesoamerican hotspot. *Conservation Biology* 22: 8–15.

Herbert ME, Mcintyre PB, Doran PJ, Allan JD, and Abell R (2010) Terrestrial reserve networks do not adequately represent aquatic ecosystems. *Conservation Biology* 24: 1002–1011.

Hieronymus CF and Bercovici D (2001) A theoretical model of hotspot volcanism: Control on volcanic spacing and patterns via magma dynamics and lithospheric stresses. *Journal of Geophysical Research* 106: 683–702.

Kareiva P and Marvier M (2003) Conserving biodiversity coldspots. *American Scientist* 91: 344–351.

Kirkpatrick JB (1983) An iterative method for establishing priorities for the selection of nature reserves – An example from Tasmania. *Biological Conservation* 25: 127–134.

Lamoreux JF, Morrison JC, Ricketts TH, *et al.* (2006) Global tests of biodiversity concordance and the importance of endemism. *Nature* 440: 212–214.

Lawler JJ, White D, Sifneos JC, and Master LL (2003) Rare species and the use of indicator groups for conservation planning. *Conservation Biology* 17: 875–882.

Lindenmayer D and Hunter M (2010) Some guiding concepts for conservation biology. *Conservation Biology* 24: 1459–1468.

Lucifora LO, García VB, and Worm B (2011) Global diversity hotspots and conservation priorities for sharks. *PLoS ONE* 6: e19356.

Mace GM, Balmford A, Boitani L, *et al.* (2000) It's time to work together and stop duplicating conservation efforts. *Nature* 405: 393.

Mace GM, Gittleman JL, and Purvis A (2003) Preserving the tree of life. *Science* 300: 1707–1709.

Maddock A and Benn GA (2000) Identification of conservation-worthy areas in northern Zululand, South Africa. *Conservation Biology* 14: 155–166.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.

Meijaard E and Nijman V (2003) Primate hotspots on Borneo: Predictive value for general biodiversity and the effects of taxonomy. *Conservation Biology* 17: 725–732.

Mittermeier RA, Fonseca GAB, Brooks T, Pilgrim J, and Rodrigues A (2003a) Hotspots and coldspots. *American Scientist* 91: 384.

Mittermeier RA, Meyers N, Robles Gil P, and Mittermeier CG (1999) *Hotspots – Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Chicago, USA: CEMEX/Conservation International and the University of Chicago Press.

Mittermeier RA, Mittermeier CG, Brooks TM, *et al.* (2003b) Wilderness and biodiversity conservation. *Proceedings of the National Academy of Sciences USA* 100: 10309–10313.

Mittermeier RA, Mittermeier CG, Robles Gil P, *et al.* (2002). *Wilderness: Earth's Last Wild Places*. Mexico City, Mexico: CEMEX, Conservation International and Agrupación Sierra Madre.

Mittermeier RA, Myers N, Thomsen JB, da Fonseca GAB, and Olivieri S (1998) Biodiversity hotspots and major tropical wilderness areas: Approaches to setting conservation priorities. *Conservation Biology* 12: 516–520.

Mittermeier RA, Robles Gil P, Hoffmann M, *et al.* (2004) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Mexico City, Mexico: CEMEX, Conservation International and Agrupación Sierra Madre.

Myers N (1988) Threatened biotas: "Hot spots" in tropical forests. *Environmentalist* 8: 187–208.

Myers N (1990) The biodiversity challenge: Expanded hot-spots analysis. *Environmentalist* 10: 243–256.

Myers N (2001) Hotspots. In: Levin SA, (ed.) *Encyclopedia of Biodiversity* vol. 3. pp. 371–381. San Diego, CA: Academic Press.

Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Nur N, Jahncke J, Herzog MP, *et al.* (2011) Where the wild things are: Predicting hotspots of seabird aggregations in the California Current System. *Ecological Applications* 21: 2241–2257.

Okkonen SR, Ashjian CJ, Campbell RG, *et al.* (2011) Satellite observations of circulation features associated with a bowhead whale feeding "hotspot" near Barrow, Alaska. *Remote Sensing of Environment* 115: 2168–2174.

Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* 51: 933–938.

Orme CDL, Davies RG, Burgess M, *et al.* (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature* 436: 1016–1019.

Ovadia O (2003) Ranking hotspots of varying sizes: A lesson from the nonlinearity of the species–area relationship. *Conservation Biology* 17: 1440–1441.

Pautasso M and McKinney ML (2007) The botanist effect revisited: Plant species richness, county area, and human population size in the United States. *Conservation Biology* 21: 1333–1340.

Petit RJ, Aguinagalde I, de Beaulieu J-L, *et al.* (2003) Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* 300: 1563–1565.

Pomery D (1993) Centers of high biodiversity in Africa. *Conservation Biology* 7: 901–907.

Possingham HP and Wilson KA (2005) Turning up the heat on hotspots. *Nature* 436: 919–920.

Prendergast JR, Quinn R, Lawton JH, Eversham BC, and Gibbons DW (1993) Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* 365: 335–337.

Pressey RL, Humphries CJ, Margules CR, Vane-Wright RI, and Williams PH (1993) Beyond opportunism – Key principles for systematic reserve selection. *Trends in Ecology and Evolution* 8: 124–128.

Pressey RL and Nicholls AO (1989) Efficiency in conservation evaluation – Scoring versus iterative approaches. *Biological Conservation* 50: 199–218.

Pressey RL and Taffs KH (2001) Scheduling conservation action in production landscapes: Priority areas in western New South Wales defined by irreplaceability and vulnerability to vegetation loss. *Biological Conservation* 100: 355–376.

Reyers B, Roux DJ, Cowling RM, *et al.* (2010) Conservation planning as a transdisciplinary process. *Conservation Biology* 24: 957–965.

Roberts CM, McClean CJ, Veron JEN, *et al.* (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295: 1280–1284.

Rodrigues ASL, Akçakaya R, Andelman S, *et al.* (2004) Global gap analysis: Priority regions for expanding the global protected-area network. *BioScience* 54: 1092–1100.

Sechrest W, Brooks TM, Fonseca GAB, *et al.* (2002) Hotspots and the conservation of evolutionary history. *Proceedings of the National Academy of Sciences USA* 99: 2067–2071.

Smagulova F, Gregoretti IV, Brick K, *et al.* (2011) Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* 472: 375–378.

Spiteri C, Kalinski V, Rösler W, *et al.* (2005) Magnetic screening of a pollution hotspot in the Lausitz area, Eastern Germany: Correlation analysis between magnetic proxies and heavy metal contamination in soils. *Environmental Geology* 49: 1–9.

Stattersfield AJ, Crosby MJ, Long AJ, and Wege DC (1998) *Endemic Bird Areas of World: Priorities for Biodiversity Conservation*. Cambridge, UK: BirdLife International.

Thompson JN and Cunningham BM (2002) Geographic structure and dynamics of coevolutionary selection. *Nature* 417: 735–738.

Thomson Reuters (2008) The 20 Most-cited Papers in Environment & Ecology, 1998–2008. URL http://sciencewatch.com/inter/aut/2008/08-oct/08oct20PaprsENV/.

Troumbis AY and Dimitrakopoulos PG (1998) Geographic coincidence of diversity threatspots for three taxa and conservation planning in Greece. *Biological Conservation* 84: 1–6.

van Jaarsveld AS, Freitag S, Chown SL, *et al.* (1998) Biodiversity assessment and conservation strategies. *Science* 279: 2106–2108.

Veech JA (2000) Choice of species–area function affects identification of hotspots. *Conservation Biology* 14: 140–147.

Wells MP, Alderman C, Puri J, and Timpson SL (2010) Independent Evaluation of Conservation and Sustainable Development Grants (2000–2009) for the MacArthur Foundation. Chicago, USA: MacArthur Foundation.

Williams P, Gibbons D, Margules C, *et al.* (1996) A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology* 10: 155–174.

Winston MR and Angermeier PL (1995) Assessing conservation value using centers of population density. *Conservation Biology* 9: 1518–1527.

Human Impact on Biodiversity, Overview

Leslie E Sponsel, University of Hawai'i, Honolulu, HI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Anthropocene The name proposed for the geological epoch beginning with the Industrial Revolution characterized by the increasingly dominant influence of humans on the planet's atmosphere, hydrosphere, biosphere, pedosphere, and lithosphere.

Cultural ecology Analysis of how culture influences the interactions between a human population and the ecosystems in which they reside; also called ecological anthropology.

Culture System of socially learned, shared, and patterned ideas, institutions, behaviors, and their material products that distinguishes a particular society.

Diversity principle General geographical coincidence between high concentrations of both biological and cultural diversity, usually in the tropics.

Ecological transition Tendency for societies to be in growing disequilibrium with their biophysical environment as they increasingly deplete natural resources and degrade their habitat, thereby reaching new thresholds of environmental impact.

Historical ecology Transdisciplinary and diachronic analysis of how human societies and ecosystems change and in turn transform one another through time in local and regional landscapes. Data are drawn from geology, archeology, history, and other sources.

IPAT A mathematical formula for calculating human impact on the environment expressed as Impact = Population × Affluence × Technology.

Overexploitation Harvesting a population or resource at a rate greater than its natural replacement rate, which is the opposite of sustainable use.

Swiddening Umbrella term including diverse types of horticulture in which a small section of forest is cut and burned to plant crops in a temporary garden; shifting cultivation is used as a synonym, but slash-and-burn cultivation is now considered to be a pejorative term.

Traditional or local environmental knowledge (TEK) Detailed and accurate knowledge about the environment, including biotic species and ecological processes, that many indigenous and other peoples have developed and accumulated, and apply in their daily and intimate interactions with their habitats and in their system of natural resource use and management. The study of TEK is included within ethnoecology.

Introduction

Throughout human prehistory and history, human impacts on biodiversity have reached progressively higher thresholds. The net impact of humanity has been to reduce biodiversity. However, at the population level, the types and magnitudes of human impacts on biodiversity vary tremendously through space and over time, depending on the specifics of the particular context. Most societies have decreased local biodiversity, but many others have sustained or even increased it. Because the direct and indirect impacts of humans on the environment are nearly ubiquitous almost any biodiversity research and conservation must take the possibilities of human influence into consideration. As Western (2001) observes, "Conservation philosophy, science, and practice must be framed against the reality of human-dominated ecosystems, rather than the separation of humanity and nature underlying the modern conservation movement." Accordingly, this article explores successive thresholds of human impact on biodiversity in terms of biodiversity modification, conversion, commodification, and toxification. It concludes with a discussion of biodiversity futures. (Major surveys of this subject include Marsh, 1864; Thomas, 1956; Turner *et al.*, 1993; McNeely *et al.*, 1995; Goudie and Cuff, 2002; Myers and Kent, 2005; Sodhi and Ehrlich, 2010).

Principles

Human and Environmental Diversity

Not all humans are equal in their impacts on biodiversity. Although it is widely recognized that there is tremendous diversity in nature, there is also considerable diversity in humanity. Nearly 7000 distinctive cultures exist today and there were many times more in the past, probably on the order of at least tens of thousands over some four million years of human existence. Accordingly, there is tremendous diversity in human relationships with and impacts on biodiversity. In examining human impacts, it is imperative to consider the diversity of humans through time (prehistory through history) and space (cultures and regions).

Environmental diversity is another variable in determining the human impact on biodiversity. Some environments are simply much more fragile and vulnerable than others, especially those that are relatively simpler, such as the arctic and deserts, and those that are isolated with a high proportion of endemic species, such as oceanic islands. Also, some regions are simply much less accessible or hospitable to humans than others, including the poles, mountain heights, deserts, and deep oceans. In these cases, the label "wilderness" may still apply – pristine nature that is largely unaffected by humans.

Furthermore, given the dynamism of ecosystems and ecological processes, together with the widespread impact of humans on them, it is misleading to consider most environments as pristine, virgin, primeval, or wilderness. The condition of any ecosystem is the cumulative product of previous conditions, usually including human impacts. The more important question is not whether or not a human society had or has any environmental impact, but its particular forms and magnitude, and if negative then the extent to which it is reversible and allows natural regeneration within a normal period of time.

In studying human impact it cannot be assumed that the previous environment was in a state of equilibrium or climax stage in ecological succession. Furthermore, even when humans are a factor in environmental change, it cannot be assumed that other agencies, such as natural alterations in climate or weather and natural fires, are not also contributing to change.

In any case, it is unrealistic to consider biodiversity anywhere without also considering the possibilities of human influence. Human impact is inevitable because, ultimately, biodiversity is the primary or raw natural resource that all societies rely on for their subsistence and economy. Also, biodiversity is the biotic component of ecosystems and ecological processes on which human survival, adaptation, and welfare ultimately depend. Another term for human impact that has become popular is ecological footprint.

Types of Impact

Human impact on biodiversity, direct or indirect, involves four basic factors: (1) overexploitation of natural resources; (2) habitat modification, conversion, and fragmentation; (3) the introduction of exotic (nonnative) species; and (4) pollution. Any one of these four factors may influence ecosystem composition, structure, and function; ecological processes; and biodiversity, especially through species extinction. Documented extinctions during the past 400 years include 484 animal species and 654 plant species, but this must be only a fraction of the actual total (Heywood, 1995:198). However, in some circumstances, the second and third factors may increase biodiversity locally, because some components of biodiversity are anthropogenic (human-caused). In Britain, for example, there has been a net increase in mammalian diversity as 21 of its 49 species are introductions (Heywood, 1995:757).

Human populations may influence biodiversity at the genetic, population, species, ecosystem, biome, and/or biosphere levels of the ecological spectrum. Because species as well as ecosystems are interrelated and interdependent, human impact on any one species or ecosystem may influence others as well, much like a chain reaction often referred to as a cascade effect.

Species and Population Levels

Globally humankind has increasingly tended toward greater levels of ecological disequilibrium, that is, with many (but not all) populations exceeding the carrying capacity of their habitat, depleting resources, degrading environments, and thereby threatening and reducing biodiversity. There have been at least seven successive thresholds of environmental impact in cultural evolution, their inception approximated in years before present and varying regionally: the use of fire by foragers (hunter–gatherers) (0.5 million), farming (5000–10,000), cities (5000–6000), exponential growth in the human population (1000), European colonialism (500), industrialization (150–200), and globalization (50). These thresholds are part of a continuum or gradual process, although an accelerating one. (These thresholds are somewhat analogous to the states of water as ice, liquid, and gas. The substance – H_2O – is the same, but a gradual quantitative increase in temperature leads to major qualitative changes in its physical properties.) In the case of sociocultural evolution, scale of complexity is a pivotal factor (see **Table 1**).

At the population level, the situation is complicated. Long ago some societies developed relatively sustainable economies that did not markedly endanger or reduce biodiversity, as

Table 1 General trends in sociocultural evolution correlated with increasing thresholds of human impact on biodiversity

1. Population – nomadic to sedentary settlement pattern with increasing population density, nucleation (settlements to cities), and pollution.
2. Food – wild to domesticated foods with shift from foraging to farming, and polycropping to monocropping
3. Energy – somatic (human and animal) to extrasomatic (water, wind, wood, fossil fuel, nuclear) sources of energy for work
4. Land – extensive (horticultural) to intensive (agricultural) land use; land tenure – community/public to private/corporate ownership
5. Economy – subsistence (satisfying basic physiological needs) to market of surplus production to materialist consumerism; local selfsufficient to regionally and then globally interdependent economy (globalization)
6. Waste – organic products that are readily biodegradable to more recently those like metals and plastics that disintegrate very slowly and are toxic
7. Scale – small and decentralized to large and centralized societies (states), the latter with increasing import of natural resources from ecosystems in distant regions
8. Differentiation – egalitarian to hierarchical (stratified) societies, the latter with increasing inequality in access to resources, goods, and services and institutionalized warfare with militarization of institutions and regions
9. Alienation – daily to occasional contact with and feedback (monitoring human impact) from the natural environment; eventually with alienation from nature and other humans, and increasingly decisions made by agents far removed from the locations they affect
10. Worldview – ecocentric to anthropocentric and egocentric worldviews, attitudes, and values; also sacred/moral to secular/amoral and utilitarian orientation to nature; may include shift from biophilia (love of nature) to biophobia (fear of nature)
11. Balance – some degree of dynamic ecological equilibrium with recognition of limits to increasing disequilibrium with the false assumption that there are no limits (i.e., ecological transition)
12. Impact – environmental modification to conversion (natural to cultural landscapes) and fragmentation (remnant patches of nature); also toxification with industrialization; local to global impact on biodiversity and ecosystems

evidenced by their longevity in prehistoric and historic records. Some societies even increased biodiversity. Furthermore, the same society may increase biodiversity in certain ways, yet decrease it in others. Thus, human impact varies in time and space as well as with different species, ecosystems, and biomes. Consequently, any generalizations about human impact must be carefully scrutinized with attention to temporal and geographic scales as well as the specific details of the particular cultural and ecological contexts.

It follows that a holistic anthropological approach, one emphasizing cultural and historical ecology, is indispensable for assessing the possibilities of human impact on biodiversity in any context. It could even be argued that cultural and historical ecology are indispensable for understanding biodiversity itself, because it is so rarely free from some degree of human influence, at least indirect. Indeed, biodiversity as a concept is a Western cultural construct developed at a particular stage in history, even if the associated biophysical phenomena are undeniably real.

Ambivalence

Biologists and others have been ambivalent about the human species. Most would readily consider *Homo sapiens* to be a part of nature as a product of biological evolution, but apart from nature in terms of ecology. However, in much of the past and in many cases to this day, human populations are simply one component in the dynamics of ecosystems and ecological processes. The idea that humans are unique and apart from nature may derive from the anthropocentrism (human centeredness) of Western civilization and of Christianity, Judaism, and Islam. (Anthropocentrism is also a contributing cause of the environmental crisis.) In fact, every species is unique as a closed genetic system and in other respects. However, it may be valid and useful to view *H. sapiens* as the dominant or most important keystone species in many if not most ecosystems. (A keystone species is one that has an unusually important influence on other species and the ecosystem.)

An important example of such anthropocentrism is the common view that human impact on the environment is necessarily an unnatural disturbance. All organisms affect their environment in various degrees and ways; consider the various impacts of trees, fungi, earthworms, leaf-cutter ants, bees, woodpeckers, kangaroo rats, prairie dogs, beavers, wild pigs, primates, bats, elephants, plankton, starfish, and sharks. However, nonhuman organisms that influence their environment are rarely considered a disturbance, except for invasive species or exotic species introduced by humans. Usually only when an organism depletes resources and degrades the habitat beyond its natural capacities or normal period for regeneration might the term disturbance be appropriately applied. Yet it could be argued that some human societies are unnatural or even antinature, especially industrial ones. Other societies appear to be an integral part of nature, such as many more traditional indigenous cultures in the Amazon forests like the Yanomami. The impact of such societies on their environment is usually no less natural than that of other species. This was likely the case for the majority of societies through most of human existence.

Biodiversity Modification

Extinction Hypothesis

Human antiquity markedly varies among regions, and consequently so does cumulative impact. Hominins (the human line) evolved in Africa some four to six million years before the present. Human dispersal into other regions is approximately dated as follows: Europe and Asia (1 million), Australia and New Guinea (50,000–60,000), the Americas (12,000–20,000), and Pacific and other islands (30,000–to recent, depending on the island) (Bodley, 2011).

As humans colonized new regions, supposedly they would have a great advantage over the prey that lacked previous experience with them. Relatively suddenly the hunters became a new top carnivore in the food web, rather than through the usual gradual process of predator–prey co-evolution. Furthermore, their technology provided advantages, particularly projectile weapons like the spear that allow some safe distance from and surprise for prey when making a kill. Indeed, Paul S. Martin and others hypothesize that the dispersal of humans into new areas caused massive faunal extinctions as early as around 50,000 years ago, depending on the region, especially in Australia, the Americas, New Zealand, Oceania, and Madagascar. The main basis for the so-called "blitzkrieg hypothesis" is the apparent coincidence in timing of human arrival and massive megafaunal extinctions. However, with little to support it but circumstantial evidence, this hypothesis remains far from conclusively proven, and it is much more problematic and controversial than advocates usually admit. For example, convincing evidence is lacking in Australia for a temporal coincidence between human dispersal and megafaunal extinctions. Yet even if temporal coincidence were demonstrated, that does not automatically prove a cause–effect relationship nor exclude other potential causal factors such as climatic change instead or in addition. In New Zealand, however, there is no doubt that overhunting and habitat destruction by indigenous people (Maori) caused the extinction of many animal species, including the large moa birds.

Whether examining faunal extinctions or any other aspect of human impact on biodiversity, human diversity must be taken into account. Even in especially vulnerable ecosystems, different cultures may have very different impacts, such as indigenous peoples and subsequent colonists in Hawai̓i. Hawaiians introduced 34 exotic species and supposedly caused the extinction of some 50 endemic species during their more than 1500 years of exclusive occupation. In contrast, European, Asian, and other colonizers introduced 4653 exotic species, caused the extinction of 211 endemic species, and endangered more than 800 others, all in a little over two centuries. Consequently, compared with native Hawaiians, subsequent colonizers caused 137 times as many introductions and four times as many extinctions, all in a mere one-seventh of the time (Sponsel, 2001).

In general, traditional hunter–gatherers and others may seem to have a very limited impact on their environment, given their low population density, high mobility, limited technology, subsistence economy, minimal needs and wants, intimate environmental knowledge and monitoring, and Animistic worldview, attitudes, and values that are nature oriented.

Nevertheless, even small bands of nomadic foragers can have some impact, for example, on seed dispersal; one way archeologists identify prehistoric occupation sites such as hunting camps is by distinctive combinations and concentrations of useful plant species that would not likely occur naturally.

Fire

Natural fires have influenced ecological systems and processes for many millions of years. In contrast, so far evidence indicates that the human use of fire extends back at least half a million years in the Old World to the later portion of the period of *Homo erectus*, but it could be twice or more that length of time. Usually the main reason hunters burn an area is to create fresh plant growth to attract game. In addition, low-intensity fires set by indigenes prevent the accumulation of fuels that might otherwise lead to devastating wildfires, especially during a dry season, drought, or from lightning strikes. (This is a lesson only recently learned by national park managers in the US, Australia, and elsewhere). Controlled burning of different patches of the landscape at different times by indigenes or others may help increase biodiversity, whereas wildfires may decrease it.

The long-term cumulative effect of burning by foragers can be substantial, as in the case of the so-called fire-stick farming by the Aborigines of Australia more than some 50,000 years (Flannery, 1995). Repeated burning of large areas of vegetation over long periods would eventually significantly modify the plant community and consequently also the associated animal community. Sometimes a pyrosere is created, a biotic community resistant or adapted to fire that is prevented from further ecological succession or development by repeated burning. With the regular and widespread use of fire by foragers, surely a new threshold in impact was reached. Some grasslands in temperate and tropical zones, such as the Scottish highlands and Venezuelan llanos, respectively, may be the result of repeated burning by humans at least in part, although in many areas this conclusion is still controversial and climatic change may be a factor as well. Fire can also increase the biodiversity of an area, among other things, by creating a mosaic of plant communities at different stages of succession.

Horticulture

Fire is also an important component of farming technology. In tropical forests, during the dry season a plot of about a hectare or so of vegetation is cut and then, after drying out in the sun, it is burned. The organic ash from the burn provides fertilizer for the growth of crops that are planted just before the rainy season. This ash can be critical because tropical soils are usually poor. The crops are often pioneer or weed species that thrive in disturbed areas. Swiddening in some form is not only ubiquitous throughout tropical forests, but was also practiced by pioneer farmers in Europe and America.

Traditional swiddening may enhance biodiversity. This practice creates a gap in the forest that provides sunshine for crops. They are harvested for a few years and then less intensively as productivity declines with decreasing soil fertility and increasing problems with weed competition and pests. Over years or even decades, the swidden gradually converts into fallow, with successional plants growing from seeds left in the soil and entering from surrounding forest through agencies like wind and birds. Consequently, swiddening leads to a mosaic of plant and associated animal communities at different stages of succession. Swiddening also forms ecotones (transition zones between two environments) that may increase biodiversity (edge effect) by harboring some species from both environments and others specializing on the ecotone. Such environmental heterogeneity increases the potential for higher biodiversity. Gardens and fallows also attract game and other animals, given the concentration of vegetation on the ground that provides cover and a concentration of edible plants, unlike a primary forest in the tropics.

Swiddening enhances biodiversity at the genetic level as well, since most swiddens are polycrops (numerous species and varieties) of domesticated plants. Crop diversity reduces the risk of total failure as a result of weather, disease, or pests, because some species and varieties are more resistant or resilient.

As a result of the variety of human uses and manipulations of forests, it cannot simply be assumed that any are entirely pristine. In the Amazon it is estimated that nearly 12% of the forests beyond the floodplains are to some degree anthropogenic (Balée, 1994). For some two decades, researchers at La Selva, the famous tropical biology research station in Costa Rica, thought that they were studying natural forest. They were unaware that it was to some extent anthropogenic until the discovery of charcoal and pottery fragments in the soils. In Africa, conservationists assumed that forest patches in savanna grasslands were relics of formerly more widespread forest. However, studies demonstrated that many of these forests are not natural relics of high biodiversity, but are established, used, and managed by local villagers (Fairhead and Leach, 1996). In many instances, farming may grade into forestry with agroforestry (mixed tree crops) as an intermediate phase.

Under traditional conditions – low density of the human population, a subsistence rather than market economy, adequate fallow periods, and extensive forest for future gardens – swiddening is usually sustainable, that is, it does not irreversibly deplete natural resources and degrade ecosystems. Indeed, it does not appear that forests like the Amazon were ever endangered, even after centuries or millennia of indigenous activities, until the encroachment of Western civilization and particularly in recent decades. However, in most areas traditional conditions no longer hold.

However, there are cases where swiddening has a negative impact. As an example, changes in plant species diversity over the past 14,000 years in Panama are preserved in phytoliths (plant fossils) from lake sediments. A sharp decrease in this diversity coincides with the onset of swiddening around 7000 years ago and may be causally related (Piperno, 1994). Thus, human impact on biodiversity from swiddening and other activities needs to be assessed on a case-by-case basis.

Still, traditional shifting cultivators contrast markedly with shifted cultivators of recent decades, the latter being immigrants who colonize tropical forests, often as economic refugees. Pioneer swiddeners are much more likely to contribute to deforestation because of their lack of familiarity with farming in a tropical forest and orientation to a market economy rather than mainly subsistence. This is the case with

people moved by the Indonesian government in its transmigration program from islands with high human population density like Java to those with low density like Irian Jaya. A similar situation occurred when the Brazilian government built the Trans-Amazon highway as a way to relocate rural poor mainly from the northeast into the Amazon Basin.

Biodiversity Conversion

Agriculture

The shift from human modification to the conversion of environments and biodiversity pivots on the decline of rotational systems of land and resource use, a process usually associated with increasing population and economic pressures. Traditional foragers, herders, and swiddeners usually tend to maintain sustainable economies by rotating one or more of these factors: types of resources used as well as places, times, techniques, and personnel in resource exploitation. If rotation decreases, then human impact increases markedly and eventually conversion occurs instead of modification, reflecting an emphasis on intensive rather than extensive use of land. By various estimates up to 50% or more of the land surface of the planet has been converted by humans and 23.8% of potential net primary productivity is captured by humans (Haberl *et al*., 2007).

Agriculture usually involves some kind of plow pulled by a domesticated animal, in contrast to horticulture, which usually involves only hand tools like an ax and digging stick powered solely by human muscle. Agriculture is a more intensive and permanent form of land use focused on conversion of the environment, whereas horticulture is a more extensive but temporary or rotational form of land use focused on modification of the environment. Accordingly, agriculture usually simplifies ecosystems and reduces biodiversity, whereas horticulture may sustain or even increase environmental heterogeneity and biodiversity. Agriculture also creates more environmental fragmentation than horticulture, that is, the dividing up of landscapes into isolated or semiisolated patches of natural environments such as forest remnants in farm fields. Again, whether horticulture or agriculture, it is necessary to assess the impact on biodiversity on a case-by-case basis. Nevertheless, since World War II, agricultural expansion and intensification have been major factors generating accelerating deforestation, most of all in the megadiversity countries of Brazil, Congo, Indonesia, and Mexico. However, it is possible for farming to generate diversity in domesticated and wild species.

Rice Paddies

Throughout Asia, large portions of the landscape, especially in the lowlands near rivers and streams, were converted centuries or millennia ago to wet rice paddies. Paddies can also be found in highlands where terracing has been developed. Conversion continues to this day, driven by increasing population and economic pressures on land and resources; it is accelerating in many areas. Conversion includes numerous types of forests, grasslands, and wetlands, and accordingly sacrifices an enormous amount of biodiversity. For instance, various types of forest covered about 70% of Thailand until World War II, whereas today only about 15% of the country is forested, largely because of agricultural expansion and logging (Sponsel, 1998).

Rice paddies are not simple monocrops, however. There may be hundreds or even thousands of varieties of rice in a region. It is estimated that some 50,000 local varieties of rice existed in India until recently. In addition, there may be several hundred species of wild plants and animals associated with paddies, many of them considered edible. Beyond the paddies there are home gardens, fruit orchards, pastures, various types of forest, and agroforestry (tree crops), all of which contribute considerable biodiversity to the regional landscape. For instance, in parts of Bali, home gardens may contain as many as a hundred or more species of wild and domesticated plants.

Raised Fields

Extensive areas of raised or ridged fields are found in many parts of the world, usually along river floodplains or in other wetlands. They elevate land above water for farming. Soil is dug and piled on top of either side of a ditch to form a long mound. The ditch becomes a canal for water drainage and irrigation. On floodplains the ridges are arranged in parallel fashion perpendicular or at an angle to the river. Organic debris and sediments from the canals are periodically dredged and placed on top of the mound as a rich organic fertilizer. Fish, waterfowl, and other aquatic organisms thrive in the canals and are harvested for protein in the diet. Raised fields may increase productivity and biodiversity beyond the level that would otherwise exist in the region.

Vavilov Centers

Vavilov centers are concentrations of genetic and morphological diversity remaining at the several foci of plant and animal domestication from the Neolithic some 5000–10,000 years ago, the exact period depending on the region. These centers, as concentrations of landraces and their wild ancestors, are in effect *in situ* gene banks. (Landraces are species and varieties of domesticated plants and animals that have been genetically improved by traditional farmers and herders, but remain uninfluenced by modern breeding technology.) Prehistoric and historic farmers and herders engineered significant biodiversity in their crops and livestock, in part to reduce risk. For instance, some eight species and 3000 varieties of potatoes were cultivated in the Andes. As another example, today a single species of sheep (*Ovis aries*) includes more than 800 different breeds, many quite ancient (Heywood, 1995:111, 775). Vavilov centers are increasingly endangered and degraded by growing population and economic pressures. Thus, attempts at *ex situ* conservation of their biodiversity and seeds of other plant species may be the main hope for long-term preservation, such as in the Kew Millennium Seed Bank partnership, which has so far saved seeds of 10% of the wild plant species of the world with 60,000–100,000 species under threat of extinction. The loss of diversity in the varieties of domesticated species and their ancestors severely limits the potential for adaptive responses to future environmental changes and perturbations, whether natural or anthropogenic.

Population

It has been estimated that through agriculture and other activities, humans preempt at least 40% of the Earth's total primary biological production annually (Vitousek *et al.*, 1986). This magnitude of resource use surely diminishes the possibilities for many other species and the planet's biodiversity as a whole. Given projections for continued human population growth in this century before any global stabilization, surely this cooption of energy and nutrient sources from other species will get much worse. Indeed, some four dozen of the poorest countries will triple their populations within the next five decades. It is also noteworthy that about 40% of the human population lives within 60 miles of the coast which is often relatively rich in biodiversity, but is being increasingly impacted by sea level rise with global climate change.

The human population explosion is a relatively recent phenomenon. After some four million years of human evolution, the world population had only reached 300 million by AD 1000. Then world population grew to 500 million by 1500, 1 billion by 1900, and now nearly 7 billion. It is projected to be up to 10 billion by 2050 (Myers and Kent, 2005). Correlated with this growth is increasing population density. For example, generally in the tropics, human population densities (individuals per km^2) are: foragers, <1; swiddeners, dozens; and wet rice farmers, hundreds to thousands. There are megacities with a population of more than 10 million. Increasing levels of population mean increasing impacts on biodiversity and the environment more generally (Bodley, 2008; Sponsel, 1998). A mathematical expression of such matters is IPAT, which is an acronym for $I = P \times A \times T$, wherein I stands for impact, P for population, A for affluence, and T for technology (Ehrlich and Holdren, 1971; O'Neill, 2002).

Though increases in the size and density of the human population certainly increase impacts on biodiversity, not all humans create the same impact, and some are grossly disproportionate. For example, upper- and middle-income consumers use resources and produce pollution at levels many times higher than lower-income consumers and the poorest, whether comparing economic classes within a single society or the so-called developed and less-developed countries – the core and periphery in the world economic system, respectively. This is the problem with labeling the entire human species as the dominant animal, it is mostly industrialization, rapacious capitalism, and the wealthy that impact the most on the environment (Ehrlich and Ehrlich, 2009). Because of their greater environmental impact, the wealthy nations, as well as petroleum, pharmaceutical, and other corporations, have a special responsibility to reinvest some of their profits to help fund biodiversity research, education, and conservation.

Biodiversity Commodification

Cities

Market economies and cities are largely responsible for the commodification of nature – the commercial or monetary evaluation of biota and landscapes. Because most urbanites and suburbanites do not produce their own food, they must trade or purchase it and other natural resources as imports from the farmers working in the vast rural hinterland. Thus, commodification has been developing with urbanization for at least 5000–6000 years. Today about half of humanity resides in an urban environment, although this concentration, the megalopolis, and suburban sprawl are recent phenomena. In 1800 only about 3% of humanity lived in cities. Now there are megacities with a population above 10 million. However, it is not always easy to detect a clear boundary between urban, suburban, rural, and natural areas. Still, one thing is clear: urbanization will continue to increase around the world and variously impact biodiversity (Myers and Kent, 2005).

Another result of the commodification of nature, and a major strategy in biodiversity conservation that is mainly advocated by persons from the urban elite, is the economic valuation of environmental products and services, such as medicinal plants and other products from tropical rain forests. Although this approach may be necessary to counteract the economic pressures for large-scale resource extraction from national governments, corporations, and market forces, this sales pitch diminishes biodiversity by ignoring its diverse nonmonetary intrinsic values.

The human body is host to biodiversity in the form of a multitude of internal and external parasites and other microscopic to small organisms that inhabit and visit it, such as bacteria, fungi, mites, lice, mosquitoes, ticks, and leeches. Likewise, human habitations and settlements, especially cities, may be viewed as hosts to a peculiar assemblage of biota, mainly commensals and parasites, such as squirrels, dogs, cats, rats, mice, pigeons, cockroaches, and flies. Though cities are mostly anthropogenic environments, they may also preserve fragments or islands of nature of different sizes in the form of various parks (including botanical gardens, zoos, and aquariums), home yards and gardens, indoor plants, and pets. As an example, there are about 1000 species of plants in the parks and gardens of the city of Singapore. Immigrant and ethnic neighborhoods introduce biodiversity into cities through the distinctive variety of plant species cultivated in their home gardens. In some regions where biodiversity is naturally low, human habitations and settlements may elevate biodiversity beyond prior natural levels.

The microscopic level of biodiversity cannot be ignored. Large urban populations also serve as a reservoir for viruses and a variety of other microbes that transmit and maintain diseases. The emergence and reemergence of new infectious diseases and drug-resistant strains are other aspects of how microbial diversity and history are often linked to cities. The high density of the human population in cities not only creates sanitation and health problems for humans, but also degrades the water quality and life in nearby aquatic ecosystems, including in coastal zones where more than a third of humanity dwells.

Colonialism

Colonialism, in its various forms, has affected biodiversity throughout the world, and continues to do so. During the past 500 years, Europeans introduced plants and animals from the

Old World into their overseas colonies, especially areas that were climatically similar to parts of Europe, like portions of Burma, California, Chile, Kenya, and South Africa. This Europeanization of local ecosystems reduced native biota and created neo-Europes. The biological exchange also flowed in the opposite direction, however, from the New World (Americas) to the Old (Europe and beyond), with the spread of crops originally domesticated by prehistoric native Americans, such as potatoes, corn, squash, tomatoes, tobacco, cotton, and manioc. Indeed, today much of the world's food supply comes from introduced species (Crosby, 1986).

Biodiversity at the microbial level is another aspect of the so-called "Columbian exchange," the biological and cultural consequences of the Old World and New World on one another. Indeed, throughout the Americas epidemics of introduced Old World diseases caused catastrophic depopulation of indigenous communities. Because these populations had been isolated from the Old World for thousands of years, they had no previous experience with the Old World diseases to develop immunological resistance (Crosby, 1986). The devastating crash of their populations relieved pressure on their land, resource base, and local biodiversity. In many areas, forests and other environments regenerated rapidly, indicating that indigenes had not irreversibly degraded them. Moreover, the apparent demographic void and supposed wilderness encouraged European colonization of what appeared to them to be an underpopulated and underused frontier. This process continues to this day in parts of the Amazon and elsewhere (Denevan, 1992).

The Columbian exchange was certainly an unprecedented event in the history of human impact on biodiversity; however, it did not operate on any prior pristine nature. Long before European contact, the peoples of the Americas modified and in some regions converted many ecosystems (Denevan, 1992). For instance, in some areas of the ancient Mayan civilization there was an extensive deforestation. As another example, anthropogenic prehistoric shell middens (mounds) in riverine and coastal zones may have higher diversity and density of useful plant species than adjacent natural sites.

Europeans are not the only colonials, just those most familiar. All empires are built in part on geographic expansion through some combination of long-distance trade, military conquest, and domination and oppression (economic, political, social, cultural, and religious) for the exploitation of the land, resources, and labor of other societies. Some examples are the Aztec in Mexico; the Inca in the Andes; Arabs in many parts of Africa, Europe, and Asia; and Chinese in much of East and Southeast Asia. All empires affect local biodiversity through biotic exchanges and environmental modification and conversion. For instance, in Asia the geographical distribution of some plant species is inexplicable without considering the role of sacred species such as the bodhi or pipal tree (*Ficus religiosa*) in Buddhism and temple yards.

Modern Transportation and International Commerce

Long-distance ocean travel and trade became possible with the development of seaworthy ships some 6000 years ago, and eventually facilitated European colonialism and widespread biotic introductions. However, intentional and unintentional introductions of species between formerly independent or isolated regions of the world have greatly accelerated with the advent of modern transportation and international commerce. More than 600 species of animals and plants are directly threatened by illegal international trade that involves more than ten billion dollars annually (International Union for the Conservation of Nature (IUCN), 2011).

The introduction of exotic species can endanger native species and ecosystems. For example, the introduction of the South American otter (*Myocastor coypus*) into southern Louisiana in the 1940s had profound consequences on the landscape and ecology of the Mississippi delta and marshes making flooding with storms more devastating such as in the case of hurricane Katrina. The inadvertent introduction of the brown tree snake (*Boiga irregularis*) from Papua New Guinea to the island of Guam in Micronesia caused the extirpation (local extinction) of a dozen species of native birds and still threatens other fauna. The zebra mussel (*Dreissena polymorpha*) accidentally spread from Europe into the Mississippi River basin of North America when a tanker dumped ballast waters in the Great Lakes. Among other problems, this exotic mussel now threatens numerous native mussel species. Many other examples could be given, for more than 4500 exotic species have become established in the United States during this century alone. Many exotic species, whether spreading naturally or introduced by humans, become invasive species severely disrupting ecosystems. The effect of species introductions worldwide probably results in a net loss of global biodiversity, and by now alien species have become the greatest threat to biodiversity second only to habitat destruction. However, in some areas, such as New Zealand and Hawai'i, net biodiversity has actually increased as the result of numerous introductions of exotic species, even though they variously threaten, endanger, or extirpate many native and endemic species.

Through European colonialism, Enlightenment ideals of rationalism and individualism, industrialization, capitalism, economic development, modernization, and globalization, the world's environments, lands, and natural resources, including biodiversity, have progressively become objectified and commodified. This desanctification of nature has been facilitated by the spread of some religions, especially the Abrahamic ones, that have usually suppressed if not replaced previous religions like Animism, which is usually more environmentally friendly (Harvey, 2006). All of this has greatly facilitated the overexploitation, degradation, and destruction of biodiversity throughout the world and in its last frontiers such as tropical forests. In turn, this trend has been magnified by the insatiable greed that is inherent in the competition and growth mania of capitalist economies and the associated culture of materialism and consumerism, including the recent processes of globalization. Biotic homogenization is now transpiring throughout the world.

This situation is compounded by the fact that the rich are getting richer at the expense of the poor who are becoming poorer. Wealthy entrepreneurs, multinational corporations, and the so-called developed countries can afford to buy greater access and rights to biodiversity, just as they also exert a

disproportionate impact on it. The poor are able and willing to pay less attention and money for biodiversity conservation than the rich, not necessarily because they are ignorant or unconcerned, but because they have fewer economic assets and alternatives (e.g., Stonich, 1993). During the 1997 economic crisis in Southeast Asia, for example, the harvesting and export of wildlife species and their products accelerated markedly as local people turned to their forests to generate desperately needed cash, as their regular farm crops and jobs no longer provided adequate household income. Material wealth and technology are concentrated in the developed countries in the Northern Hemisphere, whereas population growth, poverty, and biological wealth are concentrated in the so-called less-developed and developing countries of the Southern Hemisphere. Such inequities are a serious obstacle to biodiversity conservation as well as human rights and well-being.

Commercial Farming Industry

The beginnings of various agribusiness industries can be traced back to European colonialism. Most of these industries cause the massive conversion of forests and other ecosystems to monocrop plantations, such as in the tropics with bananas, cocoa, coconut, coffee, cotton, eucalyptus, oil palms, pineapple, rubber trees, sugarcane, tea, and tobacco. As a result, higher thresholds of environmental impact have been reached with extensive biodiversity loss at the genetic, species, and ecosystem levels, especially in fragile tropical islands.

Since the 1960s, the so-called Green Revolution and other forces of globalization have been threatening and diminishing the genetic and species diversity of domesticated plants and animals that have been developing since the Neolithic period. Monocrops from the Green Revolution are genetically engineered for rapid growth and high productivity in response to massive subsidies of chemicals as fertilizers, herbicides, and pesticides. However, because of the high costs of inputs, mechanization, and transportation, these agribusiness industries are grossly inefficient even if highly productive. Such monocrops are also much more vulnerable than the diverse agroecosystems of traditional farms to climatic changes and weather perturbations such as El Niño. Nevertheless, modern agribusiness, biotechnology and genetic engineering, and bio-prospecting are currently revolutionizing food production among other things, but with largely unknown yet probably far-reaching consequences from the erosion of agrodiversity and biodiversity and from chemical and genetic pollution.

The forests and woodlands of the world have declined significantly in area since the origin and spread of agriculture some 10,000 years ago. However, deforestation rates have surged dramatically in recent centuries, and especially in recent decades in the tropics. Although tropical forests cover only about 6% of the Earth's surface, they contain about half of all its biodiversity; thus they are the recent and future hot spots for biodiversity research and conservation. These forests are also increasingly becoming the hot spots of sociopolitical conflict and violence, such as in the Democratic Republic of Congo, Rwanda, Burundi, and Sierra Leone. Many temperate forests are deteriorating from loss of old growth, acid rain damage, and other forces. However, the forest cover in New England has expanded as the farm economy declined over the past 100 years, a hopeful indication of nature's capacity for regeneration.

The combination of causes and the magnitude of deforestation and consequent biodiversity erosion have varied tremendously through time and space depending on the specific details of the particular context. The impact of a stone ax on tropical forest compared with a metal one is quite different, even allowing for cumulative impact with substantial human antiquity in an area. The introduction of more efficient metal tools and a market economy to trade local natural resources for Western manufactured goods have accelerated deforestation throughout the tropics. Moreover, in recent decades newer technologies, including chain saws and bulldozers, together with new economic enterprises and incentives, created unprecedented rates of deforestation. In many areas of the Brazilian Amazon, for instance, cattle ranching made possible by government tax support and subsequent land speculation has been the major cause of deforestation since the 1960s. In much of Central America, forests were converted to pasture for cattle to feed beef for the fast food industry in North America and Europe, the so-called hamburger connection. Banana plantations have been another major cause of deforestation in Central America. An important factor in deforestation in many areas of Southeast Asia and elsewhere is the alienation of local communities from their land and resources by state government. Forests have been converted to monocrop plantations of eucalyptus and other fast-growing trees for export to Japan and elsewhere to supply the paper pulp industry. In South and Southeast Asia rubber tree plantations have been another economic development destroying biodiversity.

Aquaculture is yet another kind of commercial and industrial farming, now producing a third of the fish consumed by humans. Fish farming has been destroying hundreds of thousands of acres of mangrove forests and other wetlands with their biodiversity in recent decades. Shrimp pond farms along the southern peninsula of Thailand have grown exponentially with the investment of outside capital and technology, with the benefits going mostly to outsiders. Although aquaculture is a very ancient, productive, and efficient method for producing quality protein in many parts of Asia and elsewhere, where it has traditionally been sustainable and integrated with other aspects of the economy and ecology, this new economic development is very different. The shrimp are usually much too expensive to be consumed by most local people. Instead, they are exported to distant markets in Japan, Taiwan, North America, and Europe. Yet shrimp ponds are short-lived because waste products accumulate in the sediments until they become toxic to the shrimp or for any other subsequent uses. Consequently, repeatedly aging ponds are abandoned and new ones are constructed. At the same time, this cancerous growth of shrimp farms is degrading and destroying the local economies and biodiversity of many coastal regions. Nevertheless, aquaculture is expanding as a response to the decline of ocean fisheries, and often, although not always, with deleterious environmental consequences.

Biodiversity Toxification

Industry

At least since 1962, with the publication of Carson (1962)'s classic book *Silent Spring*, there has been growing concern about the environmental impact of chemicals and pollution. Because in both ecosystems and the biosphere, ultimately everything is connected to everything else in various ways and degrees, chemicals from agriculture, factories, motor vehicles, and other sources eventually circulate worldwide. Residues of dichlorodiphenyltrichloroethane (DDT) and other pollutants are even found deposited in layers of polar ice and the tissues of marine animals. DDT is still produced in China and India. In the Gulf of Mexico and other areas of the world's oceans, more than 50 dead zones from pollution and other causes have been recognized so far. The accumulation of nitrates and other nutrients from agricultural and household chemicals can generate the explosive growth of algae in lakes and rivers to the detriment of biodiversity, a process called eutrophication. Massive kills of birds, fish, and other species have been reported from this and other contaminants such as oil spills.

One recent possible symptom of widespread pollution, whether direct through water and soil contamination or indirect through increased ultraviolet radiation from the ozone hole, may be the gross deformities observed in frogs, a most alarming phenomenon and perhaps an early warning of impending ecological catastrophe. Pollution may be a direct or indirect factor in the alarming decline recently of bee, bat, and reptile populations. Another result of atmospheric pollutants such as motor vehicle exhaust is the greenhouse effect, which in the twenty-first century is increasingly recognized as a major cause of global climate change and the consequent rise of sea level by up to 2 m or more. The consequences of sea level rise on the biodiversity of coastal and marine ecosystems like wetlands and coral reefs are uncertain, but are likely to be negative and potentially catastrophic. In terms of high biodiversity, coral reefs are the marine analog of tropical rain forests. Global climate change is disrupting ecosystems and biodiversity throughout the planet.

Oil

Industrial society, whose lifeblood is oil and other fossil fuels that are actually the remains of ancient biodiversity, has produced unprecedented types and levels of chemical pollution that undoubtedly endanger and erode living biodiversity. To illustrate, in recent decades many remaining frontier zones like the Amazon of Ecuador and Peru have been one target for increasing oil exploration and extraction. Because of the isolation of these frontiers, the usual environmental safeguards and measures for clean-up are ignored.

During the exploration phase, hundreds of miles of roads and large grids of extensive trails for seismic testing are cut into the forest. The seismic explosions not only scatter wildlife, but also the resulting shock waves can kill hundreds of fish in rivers, lakes, and wetlands. During the production phase, a single oil well platform consumes about six acres of forest and about 2000 trees. (In the rain forest a single giant tree may be inhabited by thousands of insect and other species, most unknown to science.) Adjacent production wastes and treatment chemicals amount to millions of gallons every day for decades. At well sites there is no proper disposal of toxic waste, only open-air pits that eventually overflow into the soils, ground waters, and surface waters, thereby contaminating and killing fish and other wildlife. One gallon of oil can kill the fish living in a million gallons of water and adversely affect aquatic life at concentrations as low as one part per hundred billion (Kimerling, 1991).

Over decades, many hundreds of oil wells and extensive pipelines allow for numerous leaks and spills of black crude when breaks occur through metal aging or earthquakes. The magnitude of pollution that has been occurring in frontiers like the Amazon makes the oil spills associated with the Exxon Valdez, Gulf War, and Deepwater Horizon look like minor irritants in comparison! The costs to biodiversity are incalculable. Furthermore, beyond the corporate irresponsibility and massive environmental destruction, in the Amazon, oil is even being pursued in wildlife reserves and other areas supposedly set aside for conservation by national governments.

Militarization

A grossly neglected type of human impact on biodiversity and the environment are military activities and warfare, even though armament production is the top industry in the world with more than $400 billion in sales annually (Stockholm International Peace Research Institute (SIPRI), 2011). One aspect of warfare that is especially detrimental to biodiversity is the use of scorched earth tactics, which unfortunately are nothing new. One of the largest applications of this tactic was the US military's dumping of some 13 million gallons of Agent Orange to defoliate forests in the Vietnam War during 1962–1971. In recent decades, the US war on drug production in Amazon forests has also involved the use of defoliants. During the Gulf War, the government of Iraq set fire to oil wells in the desert of Kuwait and created oil spills along the coasts. The tolls of such tactics and other impacts of warfare and militarization on local and regional biodiversity have yet to be fully revealed. Yet as Hanson *et al*. (2009) note, "More than 90% of the major armed conflicts between 1950 and 2000 occurred within countries containing biodiversity hotspots, and more than 80% took place directly within hotspot areas."

The rise of nationalism and ethnic conflicts with the end of the Cold War has spread militarization into frontier, border, and other zones. Conflicts and wars are likely to proliferate in the future, according to Malthusian pessimists and others who consider growing resource scarcity and competition as major contributing causes. Such situations will be aggravated by global climate change. If so, then national and international security would be well served by redirecting a significant portion of funds that now go to the weapons, military, and defense industries into biodiversity and environmental research, education, and conservation programs as well as sustainable economic development. A complication is the issue of parks versus people; that is, whether environmental conservation should concentrate on developing protected areas,

developing local economies that alleviate poverty and are sustainable, or somehow integrate these two concerns.

Whereas war may threaten and erode biodiversity, peace may promote it. For instance, the Demilitarized Zone (DMZ) between North Korea and South Korea is a corridor 4 km wide and 250 km long that extends across the peninsula. For about six decades this corridor has been rigidly enforced as a no-man's land. As a consequence, farmlands thousands of years old and degraded forests have both reverted to a natural condition, thus protecting threatened and endangered species of plants and animals, as well as a cross section of the ecosystems of the Korean peninsula (Kim, 1997). The international peace parks between Costa Rica and its neighbors, Panama and Nicaragua, also promote biodiversity conservation by acting as refuges from most human activities.

Anthropocene

The human impact on the planet has reached such a threshold that geologists are now considering declaring a new epoch, the Anthropocene, as the age of human domination. It began two centuries ago with the Industrial Revolution at the end of the previous Holocene epoch from 10,000 to 12,000 years ago. It reflects the increasing visibility of the impact of humankind as a whole on the geology and ecology of the planet, but recognizes the seminal importance of industrialization. The human population explosion is another major contributing factor and has even been compared with cancerous growth (Hern, 1993). During the Anthropocene the impacts of the population and economic explosions, the latter including per capita consumption of natural resources and production of waste, involve urbanization; combustion and depletion of fossil fuels; increased emissions of gases like carbon dioxide, sulfur dioxide, nitric oxide, and methane; substantial alteration of the biogeochemical cycles and of biophysical landscapes; construction of megadams; rampant deforestation; spreading desertification; increased soil erosion and sedimentation rates silting waterways; depletion of freshwater and marine fisheries; mass extinction; and so on (Myers and Kent, 2005).

During the Anthropocene the global geological, climatic, and ecological impacts of *H. sapiens* are of such a magnitude as to be qualitatively distinct from anything prior to the Industrial Revolution. For instance, geologically, some human impacts are of a magnitude similar to earthquakes and volcanic eruptions. Increasingly there are anthropogenic ecosystems that depend on human activities for their creation and maintenance such as various types of rangelands, croplands, managed forests, and dense settlements. However, there is a continuum from natural to anthropogenic biomes and ecosystems, and many landscapes are a heterogeneous mosaic between these two poles instead of one or the other extreme, this mosaic now covering about 75% of the land surface of the planet.

Global climate change is well underway and increasingly threatens biodiversity. It is a serious challenge to protected areas as well as other environments. Populations, species, and ecosystems are already responding in various ways. Many species of plants are adjusting the timing of their life histories (phenology) whereas many species of animals are adapting the timing in their life cycles. The migration patterns and geographic distribution of species are changing (Lovejoy, 2010). Moreover, some estimate that more than a third of all species may become extinct as a result of global climate change (Thomas *et al.*, 2004).

The Anthropocene refers mostly to the total human impact when viewed cumulatively since the beginning of industrialization and collectively throughout the world. The possibility of somehow altering the character, scale, and intensity of population growth, industrialization, resource consumption, pollution, economic development, and related factors to significantly reduce their detrimental impact on biodiversity and other environmental phenomena would appear to be rather doubtful, at least in the short term. Yet the human economy depends on the nature's economy including its ecosystem services. In the long term, if radical transformations to more sustainable and greener lifestyles and societies are not achieved, then it may well be ecocidal for the Earth and suicidal for humanity. If the human species can significantly alter the chemistry of the atmosphere and oceans, then surely it can alter its own behavior, if for no other reason than for the sake of its own survival and well-being. However, dangerous short-term economic and political interests as well as just plain ignorance or denial are major obstacles.

Biodiversity Futures

The Uncertain Future

As those from the green sociopolitical movement and many others recognize, industrial society, capitalism, and economic development are based on the false assumption that infinite growth is possible on a finite base. Here growth refers to both population and economy. Base refers essentially to carrying capacity. The latter involves not only the ability of the land and natural resources to support a certain level of population without essentially irreversible resource depletion and environmental degradation, but also the capacity of ecosystems and the planet as a whole to absorb pollution and other anthropogenic stresses. Increasingly human impacts exceed the resilience of nature to regenerate and recover within a normal time span from any perturbations and stresses, natural or anthropogenic.

Until the demonstrably ecocidal ideas and practices of modern industrial society and related factors are corrected and ecosanity with some modicum of ecological balance is restored, the net impact of humans on biodiversity will be negative and only accelerate and intensify. Accordingly, all life, including that of humanity, will remain endangered. There is not much room for optimism, at least in the near future, given the great momentum of population and economic growth combined with political pressures for the so-called economic development and the elevation of the standard of living throughout the world, all at the expense of the environment and its natural capital of resources, biodiversity, and ecosystem services. Indeed, the ozone hole, acid rain, acidification of the oceans, collapse of oceanic and freshwater fisheries, destruction of coral reefs, soil erosion, desertification, global climate change, and other environmental problems from the local to global levels may be symptoms of the failure of the

experiment of industrial society after just two centuries. No human society is infallible and eternal – the archeological and historical records provide numerous cases of those that became maladaptive, collapsed, and disintegrated, such as Harappa in the Indus Valley, or ancient Greece and Rome, although the exact causes of some like Rapa Nui (Easter Island) are disputed.

Limits of Government

Regional, national, and international governmental and nongovernmental agencies have been making some significant progress on resolving environmental problems and managing the human impact on biodiversity, such as in the implementation of the Convention on Biodiversity in 1993 and the launch of the United Nations Decade on Biodiversity in 2011. Nevertheless, there are, for example, many serious limitations on the efficacy of government-protected areas for biodiversity conservation. First, estimates are they comprise up to12% of the Earth's surface, which is not a very large or representative sample of the tremendous biodiversity of the planet. Second, many are little more than "paper parks" because of inadequate funding, staff, and administration. Third, they will come under increasing attack in many ways with accelerating population and economic pressure in the future. Fourth, even for some 100,000 areas that are supposedly protected by governments, only a fraction of these have been thoroughly inventoried for biodiversity, and it would take the equivalent of the current number of experts several centuries to inventory the remainder. Clearly there are enormous challenges for inventorying, managing, and conserving biodiversity (Myers and Kent, 2005; Naughton-Treves *et al*., 2005).

Indigenous and other local communities can make a significant contribution to biodiversity studies and conservation, and thereby also have a positive impact on biodiversity. This is just beginning to be appreciated; an example is INBio (Instituto Nacional de Biodiversidad), the biodiversity inventory and conservation program in Costa Rica that employs local people as parataxonomists. Also, though today there are some 1600 botanic gardens in the world that help conserve plant diversity *ex situ*, there are nearly 7000 distinct cultures in the world and each may somehow contribute to *ex situ* and *in situ* conservation. Comanagement is usually the ideal, that is, the cooperative sharing between community and government in the design, authority, responsibility, and benefits of natural resource management and biodiversity conservation projects. Among the successes in comanagement are Manu National Park in the Peruvian Amazon and Kakadu National Park in northern Australia. However, when government administrators of protected areas ignore the needs of local people, then conservation efforts usually falter or fail.

Numerous people have grown increasingly skeptical that government, science, technology, education, economic, and other secular approaches are sufficient for resolving the spiraling environmental crises facing the world. They view such attempts as treating only the superficial symptoms of the crises, not the broader underlying causes. Some of these people are turning to their own religion as a source of worldviews, inspiration, motivation, attitudes, and values for developing a more sustainable and meaningful relationship with nature (Dudley, Higgins-Zogib, and Mansourian, 2005, 2009). As the pioneers in spiritual ecology, indigenes can also provide profound insights for such endeavors (Sponsel, 2012).

Indigenous Potential

Indigenous cultures contrast sharply with ecocidal industrial and other modern cultures. Many indigenous societies, especially those that are more traditional, may provide heuristic models for biodiversity conservation through their intimate environmental knowledge (ethnoecology), sustainable economy, natural resource management and conservation practices, spiritual ecology, and protection of sacred places. (To a large extent this is because most indigenes lie to the left on the continua described in **Table 1**.) In these and other respects, it could be argued that many indigenous societies are actually more developed than industrial ones.

Regarding traditional environmental knowledge (TEK), for instance, the Ka'apor people in the Brazilian Amazon recognize at least 768 species of plants from seed to reproductive adult stages. This reflects a tremendous amount and depth of reliable empirical knowledge about the biodiversity in their habitat, and that example is just in their domain of useful plants (Balée, 1994). Indeed, such knowledge of indigenous and other peoples may provide biologists and conservationists with one desperately needed short-cut for the inventory and conservation of local biodiversity (e.g., Ventocilla *et al*., 1995).

Indigenes often consider their environment, including the biotic and abiotic components, to be sacred in various ways and degrees. Such a worldview can encompass a deep respect and reverence for nature that tempers their cultural ecology and resource use, and that may lead to environmental conservation, inadvertently if not intentionally. The role of spiritual ecology and sacred places in biodiversity conservation is just beginning to be recognized and explored by those who have an open mind to such phenomena (Posey, 1999).

Most indigenous and numerous other religions are nature centered, considering certain areas of nature in their habitat to be the foci of spiritual forces and beings. These sacred places are treated with extraordinary care and respect, and frequently taboos restricting resource use are associated with them. However, throughout the world, aggressive Christian missionization and other dominating monotheistic religions have often destroyed sacred places and indigenous religions because they were perceived as Pagan; these religion-inspired actions threaten and erode biodiversity as well. The objectification and commodification of nature also degrade and destroy sacred places, as Australian Aborigines and other indigenes have found with Western mining activities. Nevertheless, numerous and diverse sacred places in nature remain and may have contributed to biodiversity conservation in the past or could do so in the future (Ramakrishnan *et al*., 1998).

Conservationists or Exterminators

The preceding discussion does not support the romantic idea that all indigenes are always in perfect harmony with their

environment, the myth of the so-called "ecologically noble savage," which amounts to little more than a "straw man" argument (Harkin and Lewis, 2007). Instead, it recognizes the fact that many indigenous societies were to some degree environmentally friendly and that some still are, even though others are not or were not. These variants of the human impact on biodiversity have been documented in numerous cases by cultural and historical ecologists. Revisionist advocates who attack indigenous societies as environmentally destructive have yet to realize, let alone adequately resolve, the basic contradiction in their argument – how indigenes can be so knowledgeable about their habitat and interact with and monitor it on a daily basis, yet be so ignorant of, or amenable to, such destructive practices.

A related fallacy is that all humans are environmentally destructive, the so-called "*Homo devastans*" or "humans as the exterminator species" view (Balée, 1998). Such simplistic either–or, all-or-nothing, always-or-never thinking is misleading at best, but it is a surprisingly common defect of advocates of this view. Blaming all of humankind for a negative impact on biodiversity is simply scientifically inaccurate, sloppy scholarship, and professionally irresponsible, one of the problems with the idea of humans as the dominant animal. It needs to be emphasized again that humanity is diverse and so is its impact on biodiversity – some societies decrease it, others sustain it, some enhance it, and others affect it in some combination of these directions. A more careful, contextualized, and nuanced approach to such issues is sorely needed.

Diversity Principle

In general, several authorities have observed independently that the greatest concentrations of biological diversity tend to coincide with those of cultural diversity, especially in tropical forest areas, and most of all in the so-called megadiversity countries of Brazil, Colombia, Mexico, Congo, Madagascar, Indonesia, and Papua New Guinea. The present author calls this general tendency toward a geographical coincidence of high cultural diversity and high biodiversity the *diversity principle*. This coincidence has only begun to be recognized and explored, let alone explained. However, wherever indigenous societies thrive, biodiversity is likely to do so as well.

One thing is certain: in such megadiversity regions, threats to either cultural or biological diversity also threaten the other. Ironically, many indigenous societies that have proven sustainable and adaptive for centuries or even millennia, as well as their environments, are being degraded and destroyed by industrial and other societies that have yet to stand the test of time and increasingly show clear symptoms of maladaptation. The concerns for biodiversity conservation and human rights are interdependent. Furthermore, the degradation and destruction of cultural diversity, like that of biodiversity, seriously endanger the future adaptability of *H. sapiens* as well as biological evolution in general (Maffi, 2001; Maffi and Woodley, 2010).

Conclusions

Globally the net impact of the human species has decreased biodiversity. However, not all humans are equal in their impact on biodiversity because of the tremendous diversity in humankind throughout its temporal and spatial distributions, including cultural diversity. At the population level, clearly some societies may sustain or even enhance biodiversity. In particular, many indigenous societies, especially those who retain some core traditions despite superficial changes, have special potential in their environmental knowledge, worldviews, values, and other attributes to contribute to developing systems for the sustainable use, management, and conservation of biodiversity. The cumulative and collective impact of humans on biodiversity across the world is sufficient to make it imperative that anyone concerned with the biodiversity of any area must consider the possibilities of human influence. Accordingly, research on cultural ecology and historical ecology is usually indispensable.

Current anthropogenic extinction rates are estimated at 1000 to 10,000 times higher than normal background rates. Furthermore, these recent anthropogenic extinctions also involve plants, whereas prehistoric mass extinctions mainly affected animals. This is an alarming fact, among other reasons, considering how fundamental plants are to other life as primary producers in capturing solar energy through photosynthesis. Another distinction of the present extinction spasm is that an increasing proportion of humanity is becoming aware of what they are doing to nature and could change their behavior to reduce their negative impacts. After all, destroying biodiversity and ecosystems is ultimately ecocidal for humanity since they are our life-support systems.

Biodiversity is unlikely to be adequately conserved only by preservationism – by isolating nature from human disturbance in a few areas of supposed wilderness. Much more effort needs to be directed to adequately recognizing and better managing human impacts from the local to the global levels and over the long term. One problem is that decisions and actions that seem reasonable in the short term may have negative consequences in the long term. Thus, adequate environmental and resource management requires, among other things, much better informed politicians and policymakers at all levels from the local to the international who have the political will and morality to consider no less than the integral relationship between humanity and biodiversity for many generations to come. Of course, so far such leadership is grossly inadequate, but not unprecedented. For instance, the Iroquois in North America acted with the seventh generation into the future in mind. Perhaps the Convention on Biological Diversity from the 1992 Rio Summit, Rio +10, the Earth Charter, and other initiatives are a hopeful change.

Biodiversity conservation also depends on a much more informed, concerned, and involved public that understands the nature and consequences of human impact on biodiversity. Considering the gravity and urgency of this subject, environmental and biodiversity education must be advanced at all levels including in the mass media. In the process, environmental ethics must be first and foremost. One of the best places to begin is by exploring Aldo Leopold's (1949) land ethic: "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise." (However, today the term "resilience" is more appropriate than "stability.") Every individual decision potentially has some impact on biodiversity, however

small or indirect; and those of humanity collectively can be synergetic and life threatening.

Finally, because the human impact on biodiversity both locally and regionally is mixed with positive as well as negative influences, there is reason for hope as well as despair for the future. However, with the current magnitude of biodiversity loss and other environmental problems, there is no doubt that the very viability, resilience, and habitability of too many of the world's ecosystems, and thereby the biosphere as a whole, are at risk, as are the very futures of organic and human evolution on this planet. Globally the net impact of *H. sapiens* in reducing biodiversity is a dangerous reversal of the megatrends that held throughout the past 3.75 billion years of organic evolution on Earth – namely increasing diversity, complexity, and adaptability. Furthermore, recovery from this extinction spasm may require thousands or even millions of years (Weisman, 2007). Because of the gravity and urgency of the negative impact of humans on the Earth's life, it is perhaps the single greatest challenge facing humankind and biodiversity studies for the twenty-first century.

See also: Agriculture, Traditional. Agrobiodiversity. Carrying Capacity, Concept of. Commons, Concept and Theory of. Ecological Footprint, Concept of. Historical Awareness of Biodiversity. Human Impacts on Ecosystems: An Overview. Hunter-Gatherer Societies, Ecological Impact of. Indigenous Peoples and Biodiversity. Introduced Species, Impacts and Distribution of. Keystone Species. Land-Use Patterns, Historic. Religious Traditions and Biodiversity. Traditional Conservation Practices

Documentary Films

Cannon J (2005) Radically Simple. Oley: Bullfrog Films, DVD 35 min. http://www.bullfrogfilms.com

Cooper A, *et al.* (2008) *Planet in Peril*. Burbank: Cable News Network, DVD 174 min. http://www.cnn.com/planetinperil

Davis W (2009) *Light at the Edge of the World*. Washington, D.C.: Smithsonian Networks, D.C., DVD 186 min. http://www.smithsonianchannel.com

DiCaprio L (2008) The 11th Hour. Burbank: Warner Bros. Entertainment Inc., DVD 92 min. http://www.11thhourfilm.com

Eberts J (2005) Sacred Planet. Burbank: Buena Vista Home Entertainment, DVD 47 min. http://www.disneydvd.com

Ehrlich P (2008) *The Dominant Animal: Human Evolution and the Environment*. San Francisco: The Long Now Foundation/Whole Earth Films/ San Simeon Films, DVD 90 min. http://dominantanimal.org/

Gore A (2006) An Inconvenient Truth: A Global Warning. Hollywood: Paramount Pictures, DVD 96 min. http://www.climatecrisis.net

Gudmundsson BJ (2006) MountainMourning. Lewisburg: Patchwork Films, DVD 78 min. http://www.PatchworkFilms.com

Maltby C (2008) Human Footprint. Burbank: Warner Home Video, DVD 90 min.

Moyers B (1991) Spirit and Nature. Cooper Station: Mystic Fire Video, VHS 88 min. http://www.pbs.org

National Academy of Sciences (1988) *Biodiversity*. Washington, D.C.: National Academy of Sciences. http://www.nasonline.org

Roszak T, Conn S, and Anthony C (1995) Ecopsychology: Restoring the Earth, Healing the Self. Palo Alto: Foundation for Global Community, VHS, 26 min. http://www.globalcommunity.org

Stone R (2010) *Earth Days*. Boston: WGBH Educational Foundation, DVD 102 min. http://www.pbs.org

Suzuki D (2003) *The Sacred Balance*. Toronto: Sacred Balance Productions/Title House e-Distribution, DVD 216 min. http://www.davidsuzuki.org

Swimme B and Tucker ME (2011) *Journey of the Universe: The Epic Story of Cosmic, Earth and Human Transformation*. Mill Valley, CA: Northcutt Productions, 58 min. http://www.journeyoftheuniverse.org

Vance RS (2003) *World Population*. Washington, D.C.: Population Connection, DVD 7.5 min (2003). http://www.populationconnection.org

Wackernagel M (2005) The Ecological Footprint: Accounting for a Small Planet. Oley: Bullfrog Films, DVD 30 min. http://www.bullfrogfilms.com

Wilson EO (2003) *The Future of Life: Biodiversity in the New Millennium*. Los Angeles: Into the Classroom Media, DVD 47 min. http://www.classroommedia.com/ew1.html

References

Achard F, Eva H, Stibig H, *et al.* (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297: 999–1002.

Agrawal A (1995) Dismantling the divide between indigenous and scientific knowledge. *Development and Change* 26: 413–439.

Archibold OW (2002) Human Impacts on Biota. In: Goudie AS (ed.) *Encyclopedia of Global Change: Environmental Change and Human Society* 1, pp. 602–607. New York: Oxford University Press.

Austin AT (2004) The human footprint in ecology: Past, present and future. *New Phytologist* 164(3): 419–422.

Balée W (1994) *Footprints of the Forest: Ka'apor Ethnobotany – the Historical Ecology of Plant Utilization by an Amazonian people*. New York: Columbia University Press.

Balée W (ed.) (1998) Historical ecology: Premises and postulates. *Advances in Historical Ecology*, pp. 13–29. New York: Columbia University Press.

Beaumont LJ, Pitman A, Perkins S, Zimmermann NE, Yoccoz NG, and Thuiller W (2011) Impacts of climate change on the world's most exceptional ecoregions. *Proceedings of the National Academy of Sciences of the United States of America* 108(6): 2306–2311. Washington, D.C.: National Academy of Sciences.

Berkes F (1999) *Sacred ecology: Traditional ecological knowledge and resource management*. Philadelphia: Taylor & Francis.

Bodley JH (2008) *Anthropology and Contemporary Human Problems*. Lanham: AltaMira Press.

Bodley JH (2011) In: *Cultural Anthropology: Tribes, States and the Global System*, 5th edn., pp. 73–78. Lanham: Altamira Press.

Bond WJ, Woodward FI, and Midgley GF (2005) The global distribution of ecosystems in a world without fire. *New Phytologist* 165: 525–538.

Botkin DB (1990) *Discordant Harmonies: A New Ecology for the Twenty-First Century*. New York: Oxford University Press.

Botkin DB, Saxe H, Araujo MB, *et al.* (2007) Forecasting the effects of global warming on biodiversity. *BioScience* 57(3): 227–236.

Bowman DMJS (1998) The impact of Aboriginal landscape burning on the Australian biota. *New Phytologist* 140: 385–410.

Bratton SP (2002) Ecology and religion: New science, old relationships. In: Clayton P and Simpson Z (eds.) *The Oxford Handbook of Religion and Science*, pp. 207–225. New York: Oxford University Press.

Brauer J (2010) *War and Nature: The Environmental Consequences of War in a Globalized World*. Lanham: AltaMira Press.

Brookfield H, Padoch C, and Stocking M (eds.) (2002) *Cultivating Biodiversity: Understanding, Analyzing, and Using Agricultural Diversity*. London: International Technology Development Group.

Brosius JP, Tsing AL, and Zerner C (eds.) (2005) *Communities and Conservation: Histories and Politics of Community-Based Natural Resource Management*. Walnut Creek: AltaMira Press.

Cajete G (2000) *Native Science: Natural Laws of Interdependence*. Santa Fe: Clear Light Publishers.

Carson R (1962) *Silent Spring*. Boston: Houghton Mifflin.

Catton Jr. WR (1980) *Overshoot: The Ecological Basis of Revolutionary Change*. Chicago: University of Chicago Press.

Chapin FS, Zavaleta ES, Eviners VT, *et al.* (2000) Consequences of changing biodiversity. *Nature* 405(6782): 234–242.

Cincotta RP and Gorenflo LJ (eds.) (2011) *Human Population: Its Influence on Biological Diversity*. New York: Springer.

Clergue B, Amiaud B, Pervanchon F, *et al.* (2005) Biodiversity: Function and Assessment in Agricultural Areas: A Review. *Agronomy for Sustainable Development* 25(1): 1–15.

Cochrane MA (2003) Fire science for rainforests. *Nature* 421: 913–919.

Cocks M (2006) Biocultural diversity: Moving beyond the realm of 'indigenous' and 'local people'. *Human Ecology* 34(2): 185–200.

Corlett RT (2007) The impact of hunting on the mammalian fauna of tropical Asian forests. *Biotropica* 39: 292–303.
Cox GW (1999) *Alien Species in North America and Hawaii: Impacts on Natural Ecosystems*. Washington, D.C.: Island Press.
Crosby AW (1972) *The Columbian Exchange: The Biological and Cultural Consequences of 1492*. Westport: Greenwood Press.
Crosby AW (1986) *Ecological Imperialism: The Biological Expansion of Europe 900–1900*. Cambridge: Cambridge University Press.
Crutzen Paul (2002) Geology of mankind. *Nature* 415(6867): 23.
Culliney J (2006) *Islands in a Far Sea: The Fate of Nature in Hawaii, revised ed.,* Honolulu: University of Hawaii Press.
Daily GC, Ceballos G, Pacheco J, Suzan G, and Sanchez-Azofeifa A (2003) Countryside biogeography of neotropical mammals: Conservation opportunities in agricultural landscapes of Costa Rica. *Conservation Biology* 17: 1814–1826.
Dallmeier F, Fenech A, Maciver D, and Szaro R (eds.) (2010) *Climate Change, Biodiversity, and Sustainability in the Americas: Impacts and Adaptations*. Washington, D.C.: Smithsonian Institution Scholarly Press.
Davis MA (2009) *Invasion Biology*. New York: Oxford University Press.
DeFries RS, Asner GP, and Houghton RA (eds.) (2004) *Ecosystems and Land Use Change*. Washington, D.C.: American Geophysical Union.
Dematteis L, Szymczak K, and Styler T (2008) *Crude Reflections: Oil, Ruin and Resistance in the Amazon Rainforest*. San Francisco: City Lights Publishers.
Denevan WM (1992) The pristine myth: The landscape of the Americas in 1492. *Annals of the Association of American Geographers* 82: 369–385.
Diamond J (2011) *Collapse: How Societies Choose to Fail or Succeed*. New York: Penguin.
Dirzo R and Raven PH (2003) Global state of biodiversity and loss. *Annual Review of Environment and Resources* 28: 137–167.
Dudley N, Higgins-Zogib L, and Mansourian S (2005) *Beyond Belief: Linking Faiths and Protected Areas to Support Biodiversity Conservation*. Gland: World Wide Fund for Nature and Alliance of Religions and Conservation.
Dudley N, Higgins-Zogib L, and Mansourian S (2009) The links between protected areas, faiths, and sacred natural sites. *Conservation Biology* 23(3): 568–577.
Ehlers E, Kraft T, and Moss C (2006) *Earth System Science in the Anthropocene: Emerging Issues and Problems*. New York: Springer.
Ehrenfeld D (2005) The environmental limits to globalization. *Conservation Biology* 19(2): 318–326.
Ehrlich PR and Ehrlich AH (2009) *The Dominant Animal: Human Evolution and the Environment*. Washington, D.C.: Island Press.
Ehrlich PR and Holdren J (1971) Impact of population growth. *Science* 171: 1212–1217.
Fa JE, Peres CA, and Meeuwig J (2001) Bushmeat exploitation in tropical forests: An intercontinental comparison. *Conservation Biology* 16: 232–237.
Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34: 487–515.
Fairhead J and Leach M (1996) *Misreading the African Landscape: Society and Ecology in a Forest–Savanna Mosaic*. New York, NY: Cambridge University Press.
Farnham TJ (2007) *Saving Nature's Legacy: Origins of the Ideas of Biological Diversity*. New Haven: Yale University Press.
Finer M, Jenkins C, Pimm S, Kean B, and Rossi C (2008) Oil and gas projects in the western Amazon: threats to wilderness, biodiversity, and indigenous peoples. *PLoS One*, http://dx.doi.org/10.1371/journal.pone.0002932?
Flannery TF (1995) *Future Eaters: An Ecological History of the Australasian Lands and People*. New York: George Braziller.
Foley JA, DeFries GP, Asner C, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.
Foster JB, Clark B, and York R (2010) *The Ecological Rift: Capitalism's War on the Earth*. New York: Monthly Review Press.
Gardner GT (2006) *Inspiring Progress: Religions' Contributions to Sustainable Development*. New York: W.W. Norton & Company.
Gaston KJ and Spicer JI (2004) Human impacts. In: Gaston KJ and Spicer JI (eds.) *Biodiversity: An Introduction*, pp. 108–137. Malden: Blackwell Publishing.
Gill AM, Bradstock RA, and Williams JE (2002) *Flammable Australia: The Fire Regimes and Biodiversity of a Continent*. Cambridge: Cambridge University Press.
Goudie AS (2005) *The Human Impact on the Natural Environment: Past, Present, and Future*, 6th ed. New York: Wiley-Blackwell.
Goudie AS and Cuff DJ (eds.) (2002) *Encyclopedia of Global Change: Environmental Change and Human Society*. New York: Oxford University Press, Volumes 1–2.
Groombridge B and Jenkins MD (2002) Humans, food and biodiversity. In: Groombridge B and Jenkins MD (eds.) *World Atlas of Biodiversity: Earth's Living Resources in the 21st Century*, pp. 33–70. Berkeley: University of California Press.
Haberl H, Erb KH, Krausmann F, *et al.* (2007) Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 104(31): 12942–12945. Washington, D.C.: National Academy of Sciences.
Halladay P and Gilmour DA (eds.) (1995) *Conserving Biodiversity Outside Protected Areas: The Role of Traditional Agro-Ecosystems*. Gland: IUCN – World Conservation Union.
Halpern BS, Walbridge S, Selkoe KA, *et al.* (2008) A global map of human impact on marine ecosystems. *Science* 319: 948–952.
Hanson T, Brooks TM, Da Fonseca GAB, *et al.* (2009) Warfare in biodiversity hotspots. *Conservation Biology* 23(3): 578–587.
Harkin ME and Lewis DR (eds.) (2007) *Native Americans and the Environment: Perspectives on the Ecological Indian*. Lincoln: University of Nebraska Press.
Harvey G (2006) *Animism: Respecting the Living World*. New York: Columbia University Press.
Hawksworth DL and Bull AT (eds.) (2006) *Human Exploitation and Biodiversity Conservation*. New York: Springer.
Hern WM (1993) Is human culture carcinogenic for uncontrolled population growth and ecological destruction? *BioScience* 43: 768–773.
Heywood VH (ed.) (1995) *Global Biodiversity Assessment*. New York: Cambridge University Press.
Hildebrand H (2004) On the generality of the latitudinal diversity gradient. *American Naturalist* 163: 192–211.
Hornborg A and Crumley C (eds.) (2007) *The World System and the Earth System: Global Socioenvironmental Change and Sustainability since the Neolithic*. Walnut Creek: Left Coast Press.
Houghton J (2009) *Global Warming: The Complete Briefing*, 4th ed. New York, NY: Cambridge University Press.
International Union for Conservation of Nature and Natural Resources (1997) *Indigenous Peoples and Sustainability: Cases and Actions*. Utrecht: International Books.
Jablonski D (1991) Extinctions: A paleontological perspective. *Science* 253(5021): 754–757.
Jensen D and McBay A (2011) *What We Leave Behind*. New York: Seven Stories Press.
Johnson BF (2009) Anthropocene may have preceded Industrial revolution. *Earth* 54(3): 20–21.
Johnston BR (ed.) (2011) *Life and Death Matters: Human Rights, Environment, and Social Justice*, 2nd edn. Walnut Creek: Left Coast Press.
Kannan R and James DA (2009) Effects of climate change on global biodiversity: A review of key literature. *Tropical Ecology* 50(1): 31–39.
Kellert SR (1996) *The Value of Life: Biological Diversity and Human Society*. Washington, D.C.: Island Press.
Kim KC (1997) Preserving biodiversity in Korea's demilitarized zone. *Science* 278: 242–243.
Kimerling J (1991) *Amazon Crude*. Washington, D.C.: Green Ink, Inc.
Klare MT (2002) *Resource Wars: The New Landscape of Global Conflicts*. New York: Henry Holt.
Klesius M (2002) State of the planet. *National Geographic Magazine* 202(3): 102–115. http://ngm.nationalgeographic.com/ngm/0209/resources_planet.html
Lansing JS (1991) *Priests and Programmers: Technologies of Power in the Engineered Landscape of Bali*. Princeton: Princeton University Press.
Laurance WF (2007) Forest destruction in tropical Asia. *Current Science* 93: 1544–1550.
Lee TM and Jetz W (2008) Future battlegrounds for conservation under global change. *Proceedings of the Royal Society Biological Sciences Series B* 275(1640): 1261–1270.
Leopold A (1949) *A Sand County Almanac*. New York: Oxford University Press.
Leslie J (1996) *The End of the World: The Science and Ethics of Human Extinction*. New York: Routledge.
Levi T, Shepard Jr. GH, Ohl-Schacherer J, Peres CA, and Yu DW (2009) Modeling the long-term sustainability of indigenous hunting in Manu National Park, Peru: Landscape-scale management implications for Amazonia. *Journal of Applied Ecology* 46: 804–814.
Lindenmayer DB and Fischer J (2006) *Habitat Fragmentation and Landscape Change: An Ecological and Conservation Synthesis*. Washington, D.C.: Island Press.
Lockwood JL and McKinney ML (eds.) (2001) *Biotic Homogenization*. New York: Springer.

Lovejoy TE (2010) Climate Change. In: Sodhi NS and Ehrlich PR (eds.) *Conservation Biology for All*, pp. 153–162. New York: Oxford University Press. http://www.mongabay.com/conservation-biology-for-all.html
Lovejoy TE and Hannah L (2006) *Climate Change and Biodiversity*. New Haven: Yale University Press.
Lowenthal D (2000) Nature and morality from George Perkins Marsh to the millennium. *Journal of Historical Geography* 26.1: 3–27.
Lunt ID and Spooner PG (2005) Using historical ecology to understand patterns of biodiversity in fragmented agricultural landscapes. *Journal of Biogeography* 32: 1859–1873.
MacDonald C (2008) *Green, Inc.: An Environmental Insider Reveals How a Good Cause Has Gone Bad*. Guilford: The Lyons Press.
Maffi L (ed.) (2001) *On Cultural Diversity: Linking Language, Knowledge, and the Environment*. Washington, D.C.: Smithsonian Institution Press.
Maffi L (2005) Linguistic, cultural, and biological diversity. *Annual Review of Anthropology* 34: 599–617.
Maffi L and Woodley E (2010) *Biocultural Diversity Conservation: A Global Sourcebook*. Washington, D.C.: Earthscan Ltd.
Mann CC (2005) *1491: New Revelations of the Americas Before Columbus*. New York: Knopf.
McClosky JM and Spalding H (1989) A reconnaissance-level inventory of the amount of wilderness remaining in the world. *Ambio* 18: 221–227.
Markussen M, Buse R, Garrelts H, *et al.* (eds.) (2005) *Valuation and Conservation of Biodiversity: Interdisciplinary Perspectives*. New York: Spring.
Marsh GP (1864) *Man and Nature; or, Physical Geography as Modified by Human Action*. New York: Charles Scribner.
McAnany PA and Yoffee N (eds.) (2010) *Questioning Collapse: Human Resilience, Ecological Vulnerability, and the Aftermath of Empire*. Cambridge: Cambridge University Press.
McCallum ML (2007) Amphibians decline or extinction? Current declines dwarf background extinction rates. *Journal of Herpetology* 41(3): 483–491.
Mcdonald RI, Kareiva P, and Formana RTT (2008) The implications of current and future urbanization for global protected areas and biodiversity conservation. *Biological Conservation* 141(6): 1695–1703.
McKibben B (2006) *The End of Nature*. New York: Random House.
McNeely JA (ed.) (1995) *Expanding Partnerships in Conservation*. Washington, D.C.: Island Press.
McNeely JA, Gadgil M, Leveque C, *et al.* (1995) Human influences on biodiversity. In: Heywood VH (ed.) *Global Biodiversity Assessment*, pp. 711–821. New York: Cambridge University Press.
McNeely JA (2003) Conserving forest biodiversity in times of violent conflict. *Oryx* 37(2): 142–152.
McNeely JA and Scherr SJ (2003) *EcoAgriculture: Strategies to Feed the World and Save Biodiversity*. Washington, D.C.: Island Press.
Meybeck M (2003) Global analysis of river systems: From Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 358(1440): 1935–1955.
Miller GH, Fogel ML, Magee JW, *et al.* (2005) Ecosystem collapse in Pleistocene Australia and a human role in megafaunal extinction. *Science* 309: 287–290.
Minnis PE and Elisens WJ (eds.) (2000) *Biodiversity and Native America*. Norman: University of Oklahoma Press.
Minteer BA and Miller TR (2011) The new conservation debate: Ethical foundations, strategic trade-offs, and policy opportunities. *Biological Conservation* 144(3): 945–947.
Murphy BP and Bowman DMJS (2007) The interdependence of fire, grass, kangaroos and Australian Aborigines: a case study from central Arnhem Land, northern Australia. *Journal of Biogeography* 34: 237–250.
Myers N (1992) *The Primary Source: Tropical Forest and our Future*, 2nd edn. New York: W.W. Norton and Company, Inc.
Myers N (1996) *Ultimate Security: The Environmental Basis of Political Stability*. Washington, D.C.: Island Press.
Myers N and Kent J (eds.) (2005) *The New Atlas of Planet Management*, revised ed. Berkeley: University of California Press.
Naiman RJ, Melillo JM, and Hobbie JE (1986) Ecosystem alteration of boreal forest streams by beaver (*Casor canadensis*). *Ecology* 67: 1254–1269.
Naughton-Treves L, Holland MB, and Brandon K (2005) The role of protected areas in conserving biodiversity and sustaining local livelihoods. *Annual Review of Environmental Resources* 30: 219–252.
Nazarea V D (2005) *Hierloom Seeds and their Keepers: Marginality and Memory in the Conservation of Biological Diversity*. Tuscon: University of Arizona Press.
Nazarea VD (2006) Local knowledge and memory in biodiversity conservation. *Annual Review of Anthropology* 35: 317–335.
Nentwig N (2008) *Biological Invasions*. New York: Springer.
Nepstad DC, Verissimo A, Alencar A, *et al.* (1999) Large-scale impoverishment of Amazonian forests by logging and fire. *Nature* 398: 505–508.
Oelschlaeger M (1991) *The Idea of Wilderness: From Prehistory to the Age of Ecology*. New Haven: Yale University Press.
Oldfield ML and Alcorn JB (eds.) *Biodiversity: Culture, Conservation, and Ecodevelopment*. Boulder: Westview Press.
O'Neill BC (2002) IPAT. In: Goudie AS (ed.) *Encyclopedia of Global Change: Environmental Change and Human Society* 1. pp. 702–706. New York: Oxford University Press.
O'Riordan T and Stoll-Kleemann S (eds.) (2002) *Biodiversity, Sustainability and Human Communities*. Cambridge: Cambridge University Press.
Orlove BS and Brush S (1996) Anthropology and biodiversity conservation. *Annual Review of Anthropology* 25: 329–352.
Parenti C (2011) *Tropic of Chaos: Climate Change and the New Geography of Violence*. New York: Nation Books.
Pauly D, Christensen V, Guenette S, *et al.* (2002) Toward sustainability in world fisheries. *Nature* 418: 689–695.
Peres CA (2000) Effects of subsistence hunting on vertebrate community structure in Amazonian forests. *Conservation Biology* 14: 240–253.
Peres CA (2001) Synergistic effects of subsistence hunting and habitat fragmentation on Amazonian forest vertebrates. *Conservation Biology* 15: 1490–1505.
Peres CA and Nascimento HS (2006) Impact of game hunting by the Kayapo of south-eastern Amazonia: Implications for wildlife conservation in tropical forest indigenous reserves. *Biodiversity and Conservation* 15: 2627–2653.
Pfeiffer JM and Voeks RA (2008) Biological invasions and biocultural diversity: Linking ecological and cultural systems. *Environmental Conservation* 35.4: 281–293.
Piperno DR (1994) Phytolith and charcoal evidence for prehistoric slash-and-burn agriculture in the Darien rainforest of Panama. *The Holocene* 4: 34–50.
Ponting C (2007) *A New Green History of the World: The Environment and the Collapse of Great Civilizations*. New York: Penguin.
Posey DA (ed.) (1999) *Cultural and Spiritual Values of Biodiversity: A Complementary Contribution to the Global Biodiversity Assessment*. London: Intermediate Technology Publications/UN Environmental Programme.
Powell JL (2011) *The Inquisition of Climate Science*. New York: Columbia University Press.
Price SV (ed.) (2002) *War and Tropical Forests: Conservation in Areas of Armed Conflict*. Binghamton: Haworth Press, Inc.
Pyne SJ (2003) *Fire: A Brief History*. Seattle: University of Washington Press.
Rajendran CP (2008) The Anthropocene: A human-driven geological epoch on the anvil. *Current Science* 95(1): 18–19.
Ramakrishnan PS, Saxena KG, and Chandrashekara UM (eds.) (1998) *Conserving the Sacred for Biodiversity Management*. Enfield: Science Publishers.
Rands MRW, Adams WM, Bennun L, *et al.* (2010) Biodiversity Conservation: Challenges Beyond 2010. *Science* 329(5997): 1298–1303.
Redford KH (1990) The ecologically noble savage. *Orion Nature Quarterly* 9(3): 25–29.
Redford KH (1992) *Conservation of Neotropical Forests: Working from Traditional Resource Use*. New York: Columbia University Press.
Redford KH and Richter BR (1999) Conservation of biodiversity in a world of use. *Conservation Biology* 13(6): 1246–1256.
Redford KH and Mansour JA (eds.) (1996) *Traditional Peoples and Biodiversity Conservation in Large Tropical Landscapes*. Arlington: American Verde Publications.
Redman CL (1999) *Human Impact on Ancient Environments*. Tucson: University of Arizona Press.
Robinson JG and Bennett EL (eds.) (2000) *Hunting for Sustainability in Tropical Forests*. New York: Columbia University Press.
Robinson JG and Redford KH (eds.) (1991) *Neotropical Wildlife Use and Conservation*. Chicago: University of Chicago Press.
Sanderson EW, Jaiteh M, Levy M, *et al.* (2002) The human footprint and the last of the wild. *BioScience* 52: 891–904.
Sodhi NS, Koh LP, Brook BW, and Ng P (2004) Southeast Asian biodiversity: An impending disaster. *Trends in Ecology and Evolution* 19: 654–660.
Sodhi NS and Ehrlich PR (eds.) (2010) *Conservation Biology for All*. New York: Oxford University Press. http://www.mongabay.com/conservation-biology-for-all.html
Soulé ME and Lease G (eds.) (1994) *Reinventing Nature? Responses to Postmodern Deconstructionism*. Washington, D.C.: Island Press.
Sponsel LE (1998) The historical ecology of Thailand: Increasing thresholds of human environmental impact from prehistory to the present. In: Balee W (ed.) *Advances in Historical Ecology*, pp. 376–404. New York, NY: Columbia University Press.

Sponsel LE (2001) Is Indigenous spiritual ecology a new fad? Reflections from the historical and spiritual ecology of Hawaiì. In: John Grim (ed.) *Indigenous Traditions and Ecology: The Interbeing of Cosmology and Community*, pp. 159–174. Cambridge: Harvard University Center for the Study of World Religions.

Sponsel LE (2006) Yanomamo. In: Birx HJ (ed.) *Encyclopedia of Anthropology* 5. pp. 2347–2351. Thousand Oaks: Sage Publications.

Sponsel LE (2008) Sacred places and biodiversity conservation. In: Cultler J Cleveland, *et al*. (eds.) *Encyclopedia of Earth*. Washington, D.C.: National Council for Science and the Environment, Environmental Information Coalition. http://www.eoearth.org/article/sacred_places_and_biodiversity_conservation

Sponsel LE (2011) The master thief: Gold mining and mercury contamination in the Amazon. In: Johnston BR (ed.) *Life and Death Matters: Human Rights, Environment, and Social Justice*, 2nd ed, pp. 125–150. Walnut Creek, CA: Left Coast Press.

Sponsel LE (2012) *Spiritual Ecology: A Quiet Revolution*. Santa Barbara: Praeger.

Sponsel LE, Headland TN, and Bailey RC (eds.) (1996) *Tropical Deforestation: The Human Dimension*. New York: Columbia University Press.

Stahl PW (1996) Holocene biodiversity: An archeological perspective from the Americas. *Annual Review of Anthropology* 25: 105–126.

Steadman DA (1995) Prehistoric extinctions of Pacific islands birds: Biodiversity meets zooarcheology. *Science* 267: 1123–1131.

Steffen W, Crutzen PJ, and McNiell JR (2007) The Anthropocene: Are humans now overwhelming the great forces of Nature? *Ambio* 36: 614–621.

Stepp JR, Wyndham FS, and Zerger R (eds.) (2002) Ethnobiology and Biocultural Diversity*: Proceedings of the Seventh International Congress of Ethnobiology*. Athens: University of Georgia Press.

Stevens S (ed.) (1997) *Conservation Through Cultural Survival: Indigenous Peoples and Protected Areas*. Washington, D.C.: Island Press.

Stockholm International Peace Research Institute (SIPRI) (2011) http://www.sipri.org

Stonich SC (1993) *"I am Destroying the Land!": The Political Ecology of Poverty and Environmental Destruction in Honduras*. Boulder: Westview Press.

Sutherland WJ, Adams WM, Aronson RB, *et al*. (2009) One hundred questions of importance to the conservation of global biological diversity. *Conservation Biology* 23(3): 557–567.

Takacs D (1996) *The Idea of Biodiversity: Philosophies of Paradise*. Baltimore: Johns Hopkins University Press.

Thomas CD, Cameron A, Green RE, *et al*. (2004) Extinction risk from climate change. *Nature* 427: 145–148.

Thomas WL (ed.) (1956) *Man's Role in Changing the Face of the Earth*. Chicago: University of Chicago Press.

Thrupp LA (1998) *Cultivating Diversity: Agrobiodiversity and Food Security*. Wasington, D.C.: World Resources Institute.

Turner, II BL, Clark WC, Kates RW, and Richards JF (eds.) (1993) *The Earth as Transformed by Human Action: Global and Regional Changes in the Biosphere over the Past 300 Years*. New York: Cambridge University Press.

Ventocilla J, Herrera H, and Nunez V (1995) *Plants and Animals in the Life of the Kuna*. Austin: University of Texas Press.

Verschuuren B, Wild R, McNeely JA, and Oviedo G (eds.) (2010) *Sacred Sites Conserving Nature and Culture*. London: Earthscan.

Vitousek PM, Ehrlich PR, Ehrlich AH, and Matson PA (1986) Human appropriation of the products of photosynthesis. *BioScience* 36: 368–373.

Wackernagel M and Rees W (1995) *Our Ecological Footprint: Reducing Human Impact on the Earth*. Gabriola Island: New Society Publishers.

Weisman A (2007) *The World Without Us*. New York: Thomas Dunne Books.

Westbroek P (1991) *Life as a Geological Force: Dynamics of the Earth*. New York: W. W. Norton & Co.

Western D (2001) Human-modified ecosystems and future evolution. *Proceedings of the National Academy of Sciences of the United States of America* 98(10): 5458–5465.

Williams M (2003) *Deforesting the Earth: From Prehistory to Global Crisis*. Chicago: University of Chicago Press.

Wilson EO (1989) Threats to biodiversity. *Scientific American* 261.3: 108–116.

Wilson EO (2003) *The Future of Life*. New York: Alfred A. Knopf.

Wilson EO (2006) The Creation: An Appeal to Save Life on Earth. New York: W.W. Norton & Company, Inc.

Worm B, Barbier EB, Beaumont N, *et al*. (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.

Wright SJ (2003) The myriad effects of hunting for vertebrates and plants in tropical forests. *Plant Ecology, Evolution and Systematics* 6: 73–86.

Wright SJ and Muller-Landau HC (2006) The future of tropical forest species. *Biotropica* 38: 287–301.

Human Impacts on Ecosystems: An Overview

Paul R Ehrlich, Stanford University, Stanford, CA, USA
Claire Kremen, University of California, Berkeley, CA, USA
Anne H Ehrlich, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Paul R. Ehrlich, Claire Kremen, volume 3, pp 395–409, © 2001, Elsevier Inc.

Glossary

Ecosystem The community of organisms in a defined area combined with the physical factors of the environment with which the community interacts.

Ecosystem goods Commodities, such as timber and seafood, supplied free to humanity by natural ecosystems.

Ecosystem services Functions, such as flood control and pollination, supplied free to humanity by natural ecosystems.

Introduction

Ecosystems are the functional units composed of the elements of biodiversity (plants, animals, fungi, and microorganisms) in a specified area interacting with each other and their nonliving environments. Ecosystems are the "natural capital" that generates a flow "interest" in the form of ecosystem goods and services that are essential to civilization. Natural ecosystems have been the source of all foods consumed by human beings, a large proportion of the medicines, and many important industrial products ranging from natural dyes to cotton and timber. In addition to those ecosystem goods, ecosystems also supply a wide array of ecosystem services. Among other services, ecosystems tend to buffer against sudden changes in the climate, control the hydrologic cycle that brings freshwater, control floods, generate and maintain fertile soils, dispose of wastes and recycle nutrients, pollinate crops, and control the vast majority of pests that would attack crops or carry disease to human beings.

Human impacts on ecosystems precede the dawn of history. Discernable impacts probably began when people (quite possibly *Homo erectus*) first began to use fire as a tool in hunting, but there are almost insurmountable difficulties in determining when that was – perhaps more than a million years ago, perhaps as recently as 50,000 years ago. The skill of modern *Homo sapiens* as hunters is attested to by their probable role in the extermination of the so-called megafauna near the end of the Pleistocene epoch (ice ages starting about a million years ago) – mammoths, cave bears, giant sloths, and the like. In dramatically changing the structure of communities of large animals, many of which were herbivores (plant eaters), they thereby substantially altered plant communities and thus modified entire ecosystems. It was the agricultural revolution, some 10,000 years ago, however, that paved the way for the large-scale impacts that threaten the integrity of ecosystems (and thus of the critical services they supply to humanity) today. Indeed, early agricultural activities seem likely to have initiated the climate disruption that so threatens ecosystems today. In the past 150 years, since the Industrial Revolution hit its stride, the scale of the human enterprise (as measured by overall energy use) has multiplied roughly 20 times, as a result of a fourfold increase in population size and a fivefold increase in *per capita* consumption. This, of course, has resulted in a roughly 20-fold increase in the impact of *Homo sapiens* on ecosystems.

Current Impacts

The result of this expansion is that humanity has become a global geophysical force, altering Earth on a scale as large as or larger than forces such as volcanism, erosion, natural oil seepage into oceans, and, especially, the natural extinction of populations and species of nonhuman organisms. Humanity has now altered every cubic centimeter of the biosphere, at the very least by changing the climate and distributing synthetic organic chemicals, novel radionuclides, and particulate matter globally. Humanity's modification of ecosystems has been primarily mediated through its effects on their living elements.

Habitat Alteration

The most significant way in which people influence ecosystems is by degrading the habitats required by nonhuman organisms. Often this occurs simultaneously through a variety of channels, so that the types of degradation discussed later overlap with one another.

Deforestation

Forest destruction is probably the most widely recognized form of ecosystem degradation, since it so rapidly and dramatically transforms the structure of the habitat. In the past few centuries, about one-half of all tropical moist forests have disappeared, and in southeast Asia, much of the lowland moist forest is already gone. Large areas of temperate forests have been cleared, but in some areas (such as the USA), there has been substantial regrowth. Many ecosystem services are degraded by deforestation, including flood control, maintenance of appropriate gaseous content in the atmosphere (especially through carbon sequestration), regulation of the climate in other ways, control of vectors of disease, and provision of timber. Reduction of populations of dangerous

predators such as tigers is a rare and debatable benefit to *Homo sapiens*.

Forest Fragmentation

The growing patchiness of forests is receiving increasing attention, since brute rates of deforestation underrate the ecosystem impact of deforestation for a number of reasons. First, fragmentation leads to a loss of species diversity due to species-area effects. The number of species that a parcel of habitat can maintain is generally a function of its area. In general, 10 fragments averaging 1 km^2 each will retain less diversity than a single fragment of 10 km^2. Second, fragmentation increases the amount of forest edge relative to forest interior, making forest interior organisms more vulnerable to edge-specializing predators. For example, there is some evidence that predation on the eggs of ground-nesting temperate forest birds increases with proximity to edge. Third, fragmentation, especially of tropical moist forests, may lead to further forest destruction because of the penetration of winds and drying into much of the interior of small patches. Fourth, fragmentation can lead to secondary extinctions resulting from the loss of biotic interactions. For instance, extermination of populations of required pollinators can lead to the loss of plant species still present in the fragments; loss of army ants from a too-small patch can cause the disappearance of obligate ant birds. And fifth, loss of species from the fragments due to inbreeding depression (loss of genetic variability leading to lowered fertility or other deleterious effects) or chance (stochastic) extinctions stems from the small size of residual populations.

Selective Harvesting of Forests

Highgrading is the removal of commercially valuable species while bypassing others; it can dramatically alter the taxonomic composition and structure of forests. Removal of mahogany trees is a classic example; so is removal of key understory elements, such as the rattans now being extracted from the forests of southeast Asia.

Fire Suppression

Intervention to prevent fires in forest and chaparral ecosystems that normally burn periodically, known as fire suppression leads to a buildup of fuels that eventually can result in uncontrollable, extremely hot fires that change the entire system by killing organisms that would ordinarily survive the smaller, cooler fires characteristic of the original fire regime.

Conversion to Farms

Clearing of natural ecosystems to introduce crop agriculture normally results in the near-total destruction of the ecosystems. Virtually all of the larger native plant and terrestrial animal species are directly extirpated, and many of the smaller plants and animals (e.g., insects, reptiles and amphibians, birds and bats) go to extinction because of loss of habitat and changes in the microclimate.

Conversion to Grazing

Conversion of natural ecosystems to grazing can be as destructive as or more destructive than crop agriculture if, for example, large areas of tropical moist forests are cleared to make pastures. If, on the contrary, a savanna ecosystem with abundant natural grazers simply has the composition of the grazing community changed (as when small numbers of cattle are added to the ungulate population of an East African savanna), the damage can be much smaller. Nonetheless, at a minimum the composition and structure of the flora are altered by the introduced herbivores, and often attempts are made to control potential predators on the grazing animals, which may also have cascading effects on the original flora and fauna.

Conversion to Infrastructure

Building structures, highways, malls, and so forth competes with farming to have the greatest impacts per unit area on natural ecosystems. At best, only disconnected fragments of the original system remain, and the delivery of ecosystem services is almost totally disrupted. Conversion to infrastructure, with its destruction or covering of the soils that are normally essential to maintaining a terrestrial ecosystem, is one of the most irreversible forms of human impact on ecosystems (although in the absence of maintenance, altered but "natural" ecosystems will eventually recapture human-modified systems). Division of habitat by roads is also a form of conversion to infrastructure. Roads cause impacts on ecosystems far more than those created by the road-building itself. First, they may form barriers to the normal movement of animals. Second, they serve as conduits for the introduction of invasive organisms . Third, dirt roads may influence the flora and fauna of adjacent areas by deposition of dust (which, for example, functions as an insecticide). And fourth, especially in tropical moist forest areas, they increase penetration, settlement, and recreational use by human beings, all of which generate additional impacts.

Mining

Extraction of minerals varies in its ecosystem impacts, depending on the scale of the mines, whether they are underground, open pit, or placer (hydraulic/washing/dredging) operations, and the material mined. Large-scale open-pit mining tends to be totally destructive and expensive and difficult to ameliorate. Underground mining usually has less impact, but can destroy aquatic systems over large areas through the runoff of toxic drainage from the mines and can convert terrestrial systems into "slime pits" for mine wastes.

Damming and other Water Development

Human attempts to intervene in the hydrologic cycle has massive impacts on the ecosystems of rivers and streams by changing water flow patterns and temperatures and blocking the movements of anadromous fishes (ones that breed in streams but spend much of their lives at sea). Conditions may be created that encourage the spread of disease, as in the increase of Bilharzia in Egypt after the building of the Aswan High Dam. The dam made possible irrigation canals that were ideal breeding grounds for the parasitic worms that cause the disease. Terrestrial ecosystems, of course, are directly destroyed by inundation.

Wetland Drainage

Drying of wetlands eliminates the organisms and services supplied by estuaries, marshes, and bogs (flood control, water

purification, nursery services for commercially important fish species). Loss of wetland habitats also disrupts the migratory patterns of birds and reduces their food supply, including wildfowl valued by hunters.

Siltation of Onshore Ocean Waters

A flow of silt into the oceans results from timbering operations, construction, and other human activities in coastal zones. Siltation leads in the tropics to the destruction of coral reef ecosystems by smothering the coral organisms. This in turn leads to the loss of local fisheries and ecotourism opportunities.

Toxics in Aquatic Systems

Toxification of onshore ocean waters, lakes, and streams may disrupt marine ecosystems and negatively impact fish populations (by direct poisoning or indirectly by eutrophication – overfertilization of the water bodies by nitrogenous runoff from agriculture and urbanization). Toxins may also render fish unfit for human consumption, disrupting the food supply service of aquatic ecosystems. Agricultural runoff and sewage dumping have converted many areas of the oceans into "dead zones" in which oxygen levels are so low that most animals cannot live in them. Toxification, which has influenced the entire planet from pole to pole, is one of the most subtle forms of habitat alteration. Often toxic substances are active in concentrations that are difficult to detect, especially when they function by mimicking hormones that influence the early development of animals. The long-term ecosystemic effects of the adding of tens of thousands of synthetic organic chemicals to the global environment are simply unknown.

Ozone Depletion

Thinning of the ozone layer adds another toxic substance to Earth's ecosystems – an increased flux of dangerous ultraviolet B radiation (UV-B). The impact of such radiation can be judged from the fact that until some 450 million years ago, life could not exist on the land surface because there was no ozone layer to screen out UV-B. Until then, organisms had sheltered in aquatic ecosystems, because water rapidly reduces the UV-B flux. Ozone depletion has been most serious in the region of the "ozone hole" in the Antarctic, where impacts on oceanic ecosystems are already reported. As with other toxic substances, the long-term effects of moderate ozone depletion are difficult to predict. It has been postulated that increased UV-B or other toxic materials are involved in a global decline of amphibian populations, but the data are still too fragmentary to permit definitive judgments to be made.

Recreation

Recreational activities have diverse impacts on ecosystems. Off-road vehicles compact and erode soils and kill organisms that burrow in them, being especially destructive in arid climates. Trails in parks can cause local erosion and disrupt the normal activities of wildlife. Boat anchors and scuba-diving activities contribute substantially to the destruction of coral reefs in some localities. Ecotourism, which can provide economic benefits for conservation, can also be destructive if not properly controlled. Outflow of sewage from coastal hotels pollutes the waters of offshore reefs. Disturbance to animal communities and the native flora by tourist vehicles and lodge garbage in some African game parks is another example of the negative effects of ecotourism.

Climate Change

Alteration of the climate, which may be induced or accelerated by human activities, is a major mechanism for disrupting ecosystems. Global warming is clearly changing both terrestrial and aquatic ecosystems in the boreal regions. Perhaps the most serious immediate threat is thinning of the sea ice in the Arctic, which may lead to ice-free summers in the Arctic Ocean by 2050, if not sooner. Loss of the summer ice pack will have large, if difficult to predict, consequences for polar bears and other aquatic mammals and the food chains in which they participate. But the greatest threat of broad-scale ecosystem impacts likely resides in the potential for changing patterns of oceanic currents. Should, for example, the Gulf Stream fail or be deflected, all of the ecosystems of Western Europe would be severely stressed, to say nothing of the human populations there.

Anthropogenic climate change could be many times faster than the relatively rapid natural changes (probably enhanced by the spread of agriculture) that occurred following the last glaciation. It would certainly cause large-scale population extinctions in some areas. Changes in range and some local population extinctions have already been detected in butterfly species in western North America and Europe. Changing climates might also lead to a substantial loss of species diversity. Much will depend on the speed and nature of the changes, the genetic attributes of organisms and other species with which they are coevolving, and the capability of organisms to migrate past extensive barriers created by near-ubiquitous human development.

Alien Invasions

Alien invasions are little appreciated as engines of ecosystem modification, and yet they are one of the most important.

Islands

Many oceanic islands have lost most of their native floras and faunas to species transported by *Homo sapiens*. The destructive invaders include goats (which long ago played a major role in transforming the ecosystems of the Mediterranean basin), rats, mongooses, Philippine brown snakes (which devoured most of the avifauna of Guam), and mosquitoes (which carried bird malaria to Hawaii, resulting in the extinction of almost all species of native birds).

Continents

Continental ecosystems have also been highly modified by invasions. Most of the grasslands and woodland savannas of California are now nearly devoid of native herbaceous plants, having been substantially taken over by introduced annual grasses and herbaceous weeds from Europe.

Onshore Waters

Not only has the flora of California been greatly modified by exotics, but the coastal waters, especially around the

San Francisco Bay, are also host to a myriad of invaders. Many of these arrived in the ballast water of ships arriving from all over the world. One, the eastern shipworm, was tolerant of the relatively dilute sea water of the bay, as the native western shipworm was not. After it was introduced early in the twentieth century, it proceeded to destroy the wooden piers and bridges that laced the bay, doing as much financial damage as the 1906 earthquake.

Disease Organisms

Pathogens moved about by human beings also greatly influence ecosystems. A classic case was the introduction of a virus disease of ruminant animals, rinderpest, into northeast Africa around 1884. The disease raced south, reaching South Africa in 1896. It not only decimated the cattle herds of peoples like the Masai (causing severe famines), but also infected giraffes, buffalo, and wildebeest, causing a die-off that led to starvation of their natural predators. One consequence was that lions switched to eating people, so farmers deserted their land. In the absence of farmers and native browsing animals, brush and woodlands invaded grasslands, altering the entire ecosystem. Some resistance developed in the animals, and the situation has fluctuated since, with rinderpest remaining a factor in east African ecosystems. More recently, a fungal disease of bats, white-nose syndrome, has killed millions of the tiny mammals in North America, threatening the pollination and insect control services that bats supply.

It is speculated that the fungus was imported from Europe, but this remains unclear.

Overexploitation of Ecosystem Components

Habitat alteration and invasions are not the only ways in which human activities produce impacts on ecosystems. Another important pathway is overexploitation of economically desirable organisms, which in turn alters habitats.

Food

Often overexploitation occurs when natural populations are harvested for food. A classic example was the persecution of the passenger pigeon, which once was the most abundant bird in North America. One flock was estimated to contain 2 billion birds. Flocks moved around, breeding when large crops of beechnuts, acorns, and other forest nuts were found. They were commercially hunted to extinction to supply food markets in eastern cities, the last individual dying in the Cincinnati zoo in 1914. It has been suggested that their absence altered eastern forest ecosystems and greatly increased the food supply of rodents involved in the cycle of Lyme disease. If true, the ecosystemic effect of their extermination had an impact on humanity beyond the loss of a once abundant and easily harvested source of animal protein.

Fishing often alters ecosystems by disrupting normal oceanic food chains, especially when there is overfishing and when there is heavy "by-catch." According to new estimation methods, some 7 million metric tons of marine life are killed and discarded by oceanic fishing fleets, equivalent to about a tenth of the total catch retained, and this does not account for the heavy local mortality caused by dynamite and cyanide fishing on coral reefs. Increasingly, trawls dragged along the bottom are destroying the biotic structure there, with unknown future consequences.

Natural Products

The demand for natural products can also lead to overexploitation with accompanying ecosystem alteration, as discussed under forest clearing and fragmentation above (*see* Forest Fragmentation). But it is often difficult to measure the impact, for example, of the harvesting of specialty woods by "highgrading." Again, far more damage to the forest is sustained by the building of roads, the associated felling of nontarget trees, and the transport of harvested trunks by "skidding," than the small amount of damage from the selective removal of valuable hardwoods. Similarly, although the demand for turtle shell and soup has had serious consequences for turtle populations, little is known about the ways oceanic ecosystems may be altered by the reduction of turtle populations.

Esthetic Resources

The demand for pets frequently leads to the overexploitation of natural populations. Some of the cyanide fishing on coral reefs is aimed at aquarium fishes rather than those sold to grace the plates of rich Hong Kong businessmen. Freshwater tropical fishes are also often overexploited for the aquarium trade. The ecosystem changes being caused by the latter have not been seriously investigated. The pet bird trade has large-scale impacts on some species and in some regions. It is so popular in Indonesia that songbirds are virtually absent from much of Java due to trapping (and overuse of pesticides). In the Pramuka bird market in Jakarta, some 60,000 individuals and 160 species once were for sale in a given day. But even the impact of bird removal on the insect populations of Java is not known, let alone the overall (or long-term) ecosystemic consequences. The story is similar for the removal of giant saguaro cacti in Arizona or the poaching of orchids, chameleons, and geckos around the world – biologists realize the ecosystems are being altered, but there is neither the time nor the personnel required to evaluate these and many other impacts.

Scientific and Educational Use

Finally, sad to say, some organisms are overexploited for scientific and educational purposes. A classic example is the northern leopard frog, *Rana pipiens*, which once was abundant and harvested in large numbers for use in dissections in high school and college biology courses. It is now a rather scarce organism, in part because of overharvesting.

Signs of Ecosystem Deterioration

If humanity is causing substantial impacts on ecosystems, can one detect what they are? In many cases, the answer is yes, although precise measurements are difficult to obtain, mainly because baseline measurements were never taken before impacts occurred. Most, if not all, of the observed and measurable impacts are negative; positive changes have been observed largely in cases where the pressure has been lightened (e.g., the recovery of forests in parts of the eastern US after failing farms were abandoned).

Perhaps the most direct way to measure ecosystem deterioration is to measure the biological productivity or net primary productivity (NPP) of an area, if possible over time. If an area is converted from a natural to a human-directed system (forest to pasture; meadow to farm fields), NPP usually falls. A further decline in NPP (for instance, crop yields) over time indicates deterioration. Such declines, even in the face of increased fertilizer applications, have been widely observed, and it is well known that fertilizer applications can mask losses of micronutrients from soil and declines in soil structure for considerable periods before declining productivity becomes apparent. In such situations, not only is crop or forest productivity lost, but also the ecosystem services of soil replenishment and nutrient/waste recycling (which are also a vital part of the planet's biogeochemical cycles that control the distribution of nitrogen, oxygen, carbon, and water, among other critical substances) are also usually damaged.

Faltering Food Production

This may be the most significant trend in which human impacts on ecosystems play a role. Since 1984, the global grain harvest has failed to increase significantly on a *per capita* basis. Since cereal grains – primarily wheat, maize, and rice – make up more than two-thirds of the human feeding base and equal in weight at least half of all foodstuffs produced by agriculture, trends in grain harvests are the best indicators of food supplies and availability. There are, unfortunately, signs that climate disruption is already limiting the increase in grain production. In addition to ecosystem impacts, economic factors and changes in eating habits are also important in determining sufficiency. More than a third of the world grain harvest is used for feeding livestock today, and a growing fraction in recent years has been diverted into biofuels (including nearly a third of the grain crop in the USA by 2010). Shortages in grain supplies for human food could be somewhat compensated by reducing its use as feed or biofuel, although the higher prices commanded by biofuels make it unlikely that its feedstock will be diverted.

Tight supplies of grains often cause rises in price on the global grain market, which in turn cause prices of meat and other animal products to rise and consumer demand for them to fall. Reduced demand for animal products naturally leads to reduced demand for feedgrains, and vice versa. In industrial countries, more than half the grain consumed is used for feed, whereas in developing nations, by contrast, most grain is directly consumed by people. In the USA, demand for animal products has shifted away from beef to poultry, which can be grown with about a third as much grain per pound of meat produced. But in some developing countries, especially China, demand for animal foods is rising rapidly, and consumption of feedgrains has risen accordingly.

Despite these shifting economic and consumption patterns, the overall trend of food production has been increasingly problematic in recent years, with aggregate production barely managing to keep abreast of population growth, even though that growth has slowed markedly in the past two decades. Food production in sub-Saharan Africa, where populations are still growing rapidly, has fallen especially far behind.

The reasons for the slowdown of food production increases are many and vary from area to area. They include the Green Revolution running out of steam, diminishing returns from fertilizer applications, the decline in genetic diversity of major crops, the continuing loss of prime farmland to urbanization, soil erosion, salinization (land degradation from poorly managed irrigation), desertification, and water shortages, often from depletion of aquifers. Among problems of rising importance are underlying trends of land degradation and the related loss of ecosystem services. Although potential food production has kept up (marginally) with population growth, following are the key issues: (1) how much higher would that potential be if important ecosystem services had been carefully conserved and if climate disruption had not begun to be a factor and (2) can humanity depend on linear extrapolations of twentieth century production trends over a 50-year timescale?

Of ecosystem services that are being impaired or lost, some are basically local – soil erosion and depletion, loss of natural pest and crop disease controls, and loss of pollination services. Others are regional, such as the consequences of damage to nearby watersheds. Thus, deforestation of a watershed area leaves adjacent or downstream areas more vulnerable to floods and droughts. The destruction by unprecedented floods of a substantial portion of the agricultural capacity of Pakistan in 2010 is a case in point. There also is evidence that maintaining natural areas interspersed among intensively farmed areas helps to preserve soil fertility as well as providing sources of natural pest control, pollinators, and impediments to the spread of crop diseases.

The Green Revolution, along with agricultural expansion, a dramatic increase in irrigated lands, and other favorable factors, allowed grain production to rise by more than 73% from 1960 to 1980, whereas the population expanded by 46%. But from 1980 to 2009, grain production increases fluctuated, barely keeping pace with population growth. Up to 1984, grain production increases had exceeded population growth, reaching a *per capita* peak then of 342 kg; since then it has fluctuated between 300 and 340 kg *per capita*.

The degree to which human impacts on ecosystem services are involved in the difficulties of agriculture is hard to evaluate. Some crop damage by pests is clearly due to diminution of natural pest control services, but pesticide use (a major culprit in the destruction of natural pest control) and cultural practices make up for some of that loss. Similarly, declines in natural pollinator services are known to be occurring, but crop failures due to lack of pollination are few, since managed pollinators (generally honeybees) are often brought in to do the job. Honeybees themselves are at risk, due to exotic diseases, loss of floral resources that support honeybee colonies, and pesticide use. Loss of flood control services, due primarily to deforestation of watersheds, often results in serious damage to crops in places as diverse as the Indian subcontinent and Central America. Soil loss to erosion – perhaps 25 billion tons annual more than soil regeneration – also reduces agricultural production. And anthropogenic modification of ecosystems is depleting the genetic library of crop relatives and crop biocontrol agents, which are all critical to maintaining crops in their coevolutionary races with diseases and pests.

It is important to remember that agriculture is largely dependent on the weather and that climate stabilization is an important service of ecosystems. Thus, for example, the severe damage to agriculture due to a massive heat wave in Russia, and huge floods in China and Pakistan in 2010, may have been exacerbated by a decline in this service if anthropogenic global warming played a role in intensifying those unusual weather events. Changing trends in weather, including floods, droughts, and changes in the timing of seasonal warming or rainfall, will inevitably have major impacts on agricultural production around the world. Indeed, climate change has been shown to have reduced global harvests of wheat and corn by about 5% and 4%, respectively, from 1980 to 2008.

Fisheries Decline

Fisheries yields tell an even more dismal story – and at sea there is no doubt that yields have not been keeping pace with demographic expansion. Some 80% of the world's major fisheries are depleted or are being maximally harvested today. A peak in fisheries yield was reached in 1988 at 17.2 kg *per capita*, but yields have since remained below that level. Most of the blame for declining yields can be put on overharvesting, yet more systemic environmental damage has also played a role through the pollution and modification of estuaries and coral reefs, the destruction of mangrove fringes and coastal wetlands, the creation of dead zones through eutrophication, and the devastation of sea-floor communities by repeated benthic trawling. Anadromous fish (such as salmon) have suffered from damming of rivers and oversilting from bank erosion. Major oil spills have taken a toll on seabirds, marine mammals, and shellfish, causing lingering damage to coastal fauna in many areas.

With fisheries yields roughly stagnant for more than 20 years, growing dependence on aquaculture (fish farming) harvests has met demand for seafood. By 2010, fish farming was contributing nearly half of the seafood produced globally. But aquaculture causes new pollution and habitat destruction problems that impact natural fish populations, and fish farms rely extensively on fish products for feed. Thus, they tend to produce a net loss in fish biomass. They also are creating a rising demand for feedgrains and other agricultural products to support their yields, thus competing, along with livestock operations, with human beings for the food they raise.

It is not clear how much sustainable production of food by terrestrial agriculture is being reduced now by overexploitation and ecosystem modification. Genetic improvement of crops and increased use of water and fertilizers (among other technological advances) allowed rather consistent increases in production over the last half of the twentieth century. Despite the heavy promotion of genetically engineered crops, their introduction has not resulted in significant increases in harvests. More often, the gains were in pest control without use of toxic chemicals and in the development of such crop traits as resistance to drought or salt buildups in soil.

Important questions for the future include the following: How much more production in the past (and in the near and long-term future) would have been possible if more care had been taken to protect ecosystem services? How sustainable is the present agricultural production system, given that land degradation continues and the availability of sufficient water seems increasingly problematic as demand keeps rising and watersheds are being compromised by melting glaciers and snowpacks and changing rainfall patterns due to global warming? In addition, depletion of petroleum reserves is likely to lead to higher costs and shortages of fertilizers and pesticides. It is clear that, to meet the continued increases in demand for food by a growing human population and to meet them sustainably, agricultural production methods must undergo a different type of "green" revolution – one that more closely mimics and harnesses the natural processes in natural ecosystems.

Change in the Epidemiological Environment

This may be another sign of ecosystem modification. Anthropogenic ecosystem changes and greenhouse gas emissions from burning fossil fuels are clearly threatening the deterioration of the climate stabilization service. Those changes may have played a role in the warming that has caused dengue fever to move northward in North America, for example – and warming has been predicted to have similar effects on malaria and other tropical diseases.

Decline in Water Quality

Increasing water pollution is often a direct reflection of modification of ecosystems. A classic example is that of the New York City water supply. Around 1900, New York City's water was of such high quality that it was bottled and sold throughout New England. A century later the EPA notified the city that its water had dropped below acceptable quality. The reason was changes in the ecosystem of the Catskill Mountain watershed that served the city. Fertilizer runoff and inadequate local sewage treatment had impaired the natural-water-supply service. It was estimated that constructing a plant to treat the Catskill water would have a capital cost of $6–8 billion, with roughly a $300 million annual maintenance cost. In contrast, restoring the functioning of the natural Catskill ecosystem would cost only $1–1.5 billion. New York made the obvious choice and provided a classic case of ecosystem modification and the loss and restoration of ecosystem services. The New York City case is now being emulated in many parts of the world.

Other Symptoms

There are, of course, many other readily observable symptoms of anthropogenic ecosystem deterioration, some of which have already been mentioned. On land they include flooding, landslides and heavy soil erosion, changes in microclimate, the local disappearance of sensitive species of plants and animals, siltation of streams, and lowered water tables. A general decline of amphibians that seems to be occurring over much of the world suggests ecosystem disruption from multiple causes. In oceanic systems, changes in the mix and abundance of large animal and plant (e.g., kelp and sea grass) and planktonic species are also observable and often, as in cases of algal blooms or precipitously declining fisheries yields, are signals of serious disruption.

Driving Forces

The impacts of human beings on ecosystems are a product of three multiplicative factors: population size, affluence (or *per capita* consumption), and the technologies and social–political–economic relations that supply the consumption. These relationships are often condensed for convenience into the equation **I**mpact = **P**opulation × **A**ffluence × **T**echnology = **I** = **P** × **A** × **T**, the "IPAT equation." If one uses *per capita* energy consumption as a surrogate for A × T, the result is that quoted in the start of this article – a 20-fold increase in human impact on ecosystems in the past 150 years.

Overpopulation and Continued Population Growth

Population growth is probably the single most important factor leading to impacts on natural ecosystems. Early in the nineteenth century, there were only 1 billion people. That number had doubled to 2 billion around 1930, doubled again to 4 billion in 1975, reached 6 billion in 1999, and 7 billion in 2011. It took on the order of 300,000 years for the population of *Homo sapiens* to reach a billion, a single century to add the second billion, 30 years to add the third billion, 15 for the fourth, and 12 each for the fifth, sixth, and seventh.

Simply adding so many people to the population required gigantic alterations of natural ecosystems to bring sufficient land under cultivation to feed them and further destruction as agriculture was intensified in the past century or so. This effect was exacerbated because each individual added to the population tended to generate a disproportionate ecosystemic impact. Naturally, people farmed the most fertile soils first, got their water from the nearest sources, mined the most concentrated ores, and so on. More people required the farming of more marginal land, leading to increased land degradation. The transportation of water over greater distances involved using more water per person to make up for losses in storage and transport, and the disruption of more land per person with dams and other infrastructure. The mining of poorer ores meant that more ecosystemic destruction by mining was caused per person, since more ore must be dug up to provide the same level of metal use for each additional individual.

Since at least 1950, the entire world clearly has been overpopulated by the simple standard that the human population could not be sustained by the flow of resources generated by incoming sunlight mediated through farms, forests, the hydrological cycle, and so forth – the agricultural and natural ecosystems of Earth. Instead of preserving humanity's "natural capital," it has been necessary to expend it. Three elements of that capital are especially important. The first is deep rich agricultural soils that are generated on a timescale of centimeters per millennium and are now in many areas being destroyed at a rate of centimeters per decade. The second is aquifers, many of which were last filled during the ice ages and which are now in many, if not most, areas of the world being pumped at many times the natural recharge rate. The third is biodiversity, the species of microorganisms, plants, and animals that are working parts of ecosystems, which is now being exterminated at a rate unprecedented in the past 65 million years – at perhaps 1000 times its rate of natural regeneration from speciation.

Furthermore, the high rate of additions to the human population in the second half of the twentieth century meant that little concern could be shown for the possible long-term effects of the interventions designed to keep people eating. Programs such as the Green Revolution were driven in part by fear of massive starvation, and their consequences for natural ecosystems and the vital services they provide were largely ignored. So were their social impacts.

Overconsumption

Overconsumption is almost as important as overpopulation as a driver of ecosystem modification. For example, a significant proportion of the ecosystem modification caused by the agricultural enterprise is created by growing grain and feeding it to animals destined for human consumption. This is the case even though in rich societies, consumption of products containing animal fat (and, perhaps, overconsumption of animal protein) is a serious public health problem. The directly toxic and hormone mimicking effluents from the production of myriad products, many of them plastic or packaged in plastic, also are undoubtedly having effects on ecosystems, although data beyond changes in single species are lacking. But one of the most serious overconsumption problems derives from converting societies to dependence on the automobile, combined with the marketing and purchase of cars much too large, heavy, and inefficient for the purposes to which they are put.

Use of Environmentally Malign Technologies

The use of technologies that are unnecessarily environmentally malign is related to overconsumption. The use of inefficient automobiles is echoed in the use of unnecessarily inefficient lighting, heating, appliances, and so on. But the basic technological problem in society at the beginning of the twenty-first century is the persistent overdependence on fossil fuels, and especially coal, as an energy source.

Faulty Economic Arrangements

Market failures are partly to blame for the overdependence on fossil fuels, because the market prices of those fuels do not come close to reflecting their social costs. For example, the costs incurred by society in the form of lung disease from the effluents caused by fossil fuel burning are not factored into the market price, neither, in the USA, are the costs of maintaining large military forces designed to ensure a steady flow of oil into the American economy. More importantly, the present and potential costs to society of climate change induced by carbon dioxide emissions from fossil fuel use are also not factored. If these and other social costs, such as pollution due to oil spills and discarding of lubricants used in automobile maintenance, other emissions, and medical costs deriving from automobile accidents, were captured in market prices, the price of gasoline in the USA might be quadrupled.

Faulty Social Arrangements

Social arrangements often work against the protection of ecosystems. For instance, society has not designed educational systems to apprise its citizens of the biophysical and social realities that it faces. Few people, even "well educated" people in rich societies, can provide a coherent explanation of ecosystem services or the threats to their delivery. Indeed, there is a general failure of the media and standard educational systems to make explicit connections among the many factors and effects of human alteration of ecosystems.

Faulty Political Arrangements

Political errors also play a substantial role in endangering ecosystems. Societies generally lack "foresight" institutions that have the capacity and independence to analyze complex problems and provide competent and independent advice to government and the general public. In addition, there is the age-old problem of disparities in power, which now has increasingly widespread – often global – implications for the sustainability of society. Powerful supporters of politicians in Rome were hard-pressed to weaken the Empire. Today's business interests pressing politicians through campaign contributions and lobbying in the USA, or in other ways elsewhere, to take no actions to ameliorate global warming theoretically could bring down civilization.

Solutions

In light of the myriad uncertainties about both human and ecosystem behavior, it would seem prudent to try to reduce the three factors in the IPAT equation simultaneously, rather than risk a possible catastrophic collapse of services as *Homo sapiens* progressively alters natural ecosystems. What sorts of steps might be taken to avoid such a collapse? Some are proximate – steps that would directly address impacts on ecosystems. Others are ultimate – steps that would deal with the basic drivers of ecosystem deterioration.

Proximate

A nonexhaustive list of proximate steps might include the following:

1. Limit development so that a maximum of the remaining relatively undisturbed ecosystems would be preserved. In virtually every country where new infrastructure is needed, it should be provided as far as possible by intensively redeveloping areas that are already highly disturbed.
2. Rapidly implement agreements and measures to slow climate change, despite the serious political and economic difficulties of so doing. It is entirely in the interest of affluent countries to assist the poorest countries to develop and deploy from scratch energy systems that do not depend on fossil fuels.
3. Establish and protect reserves designed to protect relatively undisturbed ecosystems and the elements of biodiversity they contain. This should be done only when it is possible to integrate the interests of local people into the reserve design – otherwise in the long term the effort is likely to be wasted as population growth pushes people increasingly into reserves that they do not value as such.
4. Limit toxic releases by shifting the burden of proof to those who would claim that the introduction of a novel compound into use (and thus into the environment) carries benefits that clearly exceed the social costs. And reexamine the ~65,000 compounds that have been "grandfathered" as safe for release without any real evidence, just in the USA.
5. Enforce sustainable-yield harvesting for all living resources. Harvesting systems should be designed with substantial attention to the precautionary principle (erring on the side of conservatism), in order to buffer society against the consequences of overoptimistic estimates of what is sustainable.
6. Game ranching is, in some circumstances, able to maintain the diversity of grazing ecosystems much better than a monoculture of cattle or sheep. Problems of game ranching can include public acceptability of the meat, high prices, and conflict with cultural values associated with traditional grazing routines. In one case, on the Athi plain of Kenya, near Nairobi, a game ranch led to restoration of the land and profits from the meat of a variety of antelope species (with different grazing and browsing methods). That, despite resistance from traditional Masai cattle grazers and a bureaucratic prohibition from marketing hides.
7. Apply principles of countryside biogeography. Countryside biogeography is the field that is developing principles for making already disturbed areas most hospitable for biodiversity and thus more likely to provide necessary ecosystem services. One step in doing this would be to pass (or strengthen) laws to protect endangered populations, species, and landscapes and to restore moderately degraded areas, including restoration with useful, native species
8. Develop nonfarm and urban area management regimes that allow for maximal adjustment to changing climates by floras and faunas. This would include keeping possibilities open for migration of animals and plants through wildlife corridors and other mechanisms.

Ultimate

Steps to deal with the ultimate drivers of ecosystem deterioration mean dealing with the factors of the IPAT equation:

Population Reduction

Population growth should be halted as soon as humanely possible, and a slow decline initiated toward a level that can be sustained indefinitely at whatever average level of consumption is selected. Selection of that level, which is closely related to the issue of optimum population size, need not be debated in the near future. That size will change with available resources, technologies, and preferences. The human population is already well above the long-term carrying capacity, so a goal of gradually and humanely reducing population size is likely to be valid for at least a century. Until then there will be

abundant time for scientific study and social debate over the level at which to stabilize human population. Human beings, as far as can be known, have at times and in some places altered vital rates and adjusted population sizes to needs perceived primarily through the economic situations of reproducing couples. In recent decades, there has been a shift to substantial consideration of the needs of society as well, and that trend should be accelerated in the future.

Consumption Control

Controlling consumption will be a more difficult task than controlling population growth, if current trends and attitudes are any guide. This is an area fraught with difficult issues of equity and feasibility. For instance, if the gasoline tax in the USA were raised to the point where gasoline was as expensive as in Europe, a substantial decline in fuel consumption would be likely. But there would be a deleterious impact on poor people that must commute to work by car. Compensatory tax changes would be needed to lighten that burden.

Not only do many of the things human beings now consume seem unnecessarily harmful, but also the quantities consumed per person often seem outrageous. But since one person's outrage is triggered by the use of something that another values greatly, it is difficult to see how consumption patterns can be changed significantly without substantial social conflict. Much, however, can be accomplished by working hard to increase the energy and material efficiency of the processes by which consumption is served – such as using bikes; mass transit; and light, fuel-efficient automobiles to "consumers" of travel; more efficient production of consumer products; reduction of packaging; etc. – and by conducting marketing and education campaigns to influence consumption patterns toward more efficient, more sustainable ones, with reduced levels of input.

There is sometimes confusion about the relative roles of population and *per capita* consumption in human modification of ecosystems. But one can no more separate their impacts than distinguish the multiplicative roles of a rectangle's length or width in contributing to its area. In the past two centuries, population growth and expanding *per capita* consumption have contributed roughly equally to humanity's assault on its life-support systems. Reducing the assault and transitioning to a sustainable society will require action on both factors. But once it is decided to take action, it will take much longer to humanely reduce population size than to alter human consumption patterns. Given appropriate incentives, the latter can be transformed very rapidly, as World War II mobilizations showed dramatically. Moving toward population reduction now in the USA and globally is required. But if overconsumption by the rich continues to escalate, the benefits of population shrinkage will be compromised. And if underconsumption by the poor is allowed to continue or worsen, the cooperation needed to resolve the human predicament is unlikely to materialize. Equity issues and environmental issues are conjoined twins.

Substitution of Technologies

Of course, it is possible to substitute environmentally benign technologies for environmentally malign ones. For example, with proper education, natural pest control services, enhanced by integrated pest management, can be substituted in agriculture for broadcast spraying of pesticides. Buildings can be equipped with "more expensive" energy-saving devices that pay for themselves through energy savings within 1–5 years.

Revising Socioeconomic and Political Systems

All of the steps discussed would benefit from public education and broad social discussion of the need to, as the economists say, "get the prices right." The prices of goods and services need to be adjusted so that they more closely reflect the social costs of each. In other words, externalities should be internalized. The heavy burden put on society due to deterioration of the environment and public health from overuse of large automobiles should be reflected in the prices of both the vehicles and the fuel they consume.

Success at evaluating and implementing the steps that can and should be taken in various nations and diverse economic strata within nations may depend on substantial changes in governmental institutions and the generation of broad social consensus. Substantial revision of socioeconomic and political systems and of human behavior will be required if humanity is to preserve the ecosystem services on which it depends and the biodiversity on which the services themselves depend.

Appendix

List of Courses

1. Environmental Science
2. General Biology
3. Plant Agriculture, Ecology

See also: Countryside Biogeography. Island Biogeography

References

Daily GC (ed.) (1997) *Nature's Services*. Washington, DC: Island Press.

Ehrlich PR and Ehrlich AH (2009) *The Dominant Animal: Human Evolution and the Environment*, 2nd edn. Washington, DC: Island Press.

Ehrlich PR and Ornstein RE (2010) *Humanity on a Tightrope: Thoughts on Empathy, Family, and Big Changes for a Viable Future*. Lanham, MD: Rowman & Littlefield.

Ehrlich PR and Pringle RM (2008) Where does biodiversity go from here? A grim business-as-usual forecast and a hopeful portfolio of partial solutions. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11579–11586.

Postel SL, Daily GC, and Ehrlich PR (1996) Human appropriation of renewable fresh water. *Science* 271: 785–788.

Ruddiman WF (2003) The anthropogenic greenhouse era began thousands of years ago. *Climatic Change* 61: 261–293.

Steffen W, Sanderson RA, Tyson PD, *et al.* (2004) *Global Change and the Earth System: A Planet Under Pressure*. Berlin: Springer.

Turner II BL (ed.) (1990) *The Earth as Transformed by Human Action: Global and Regional Changes in the Biosphere Over the Past 300 Years*. Cambridge, UK: Cambridge University Press.

Vitousek PM, Mooney HA, Lubchenko J, and Melillo JM (1997) Human domination of earth's ecosystems. *Science* 277: 494–499.

Hunter-Gatherer Societies, Ecological Impact of

Kathleen A Galvin and Tyler Beeton, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Kathleen A. Galvin, volume 3, pp 411–415, © 2001, Elsevier Inc.

Glossary

Bands The basic economic, social, and political unit of hunter-gatherer societies.

Cultural change The process by which cultures undergo change in response to social, political, and environmental factors.

Cultures and the global system The analysis of cultural diversity, responses and adaptations of smaller scale societies to emerging global trends.

Exogamy The practice of a person seeking a mate outside of his or her group.

Global environmental sustainability The effect of human social systems and economies on ecosystem processes, their feedbacks, now and into the future.

Human environment interactions The study of the complex systems of interaction between people and their environment.

Patrilocal residence The practice of married couple living in the husband's community.

Introduction

Many people have impressions of hunter-gatherers as people who live in harmony with nature, who are organized into simple societies, and are associated with our "pristine" Paleolithic hunter-gatherer past. Many of these stereotypic impressions are false (cf. Moran, 1991; Cane, 1996). Today, all foragers live in nation-states, have some dependence on either crop cultivation or wage labor, and are not isolated. Hunter-gatherer societies have social systems that are extremely complex and whose interactions with the biodiversity surrounding them are as complicated and variable as was probably the case 10,000 years ago when all humans were foragers. It is no accident that today, areas with the greatest remaining biodiversity are also the areas inhabited by hunter-gatherers. Many hunter-gatherers retreating from land appropriation, settler immigration, and European diseases have occupied the most remote parts of their region. Today, these homelands are often part of or adjacent to conservation areas, parks, or other protected areas.

This article describes traditional hunter-gatherer societies and the adaptations these societies have made to the environment. As hunter-gatherer societies and their environments have undergone continuing changes, issues of biodiversity conservation and hunter-gatherer welfare are discussed within the context of their changing world.

Hunter-Gatherer Societies and Natural Resource Exploitation

Because hunter-gatherers have lived in diverse environments and live next to numerous other cultural groups, they have manifested an incredible diversity of cultures and natural resource-management adaptations. Nevertheless, there are several general characteristics of hunter-gatherer societies: these traits have had a direct impact on the use of natural resources.

Traditional hunter-gatherer societies were comprised of bands, social groups made up of close biological kin and friends. The composition and sizes of bands changed seasonally, depending on the abundance and location of food resources. Bands were led by individual hunters who were respected for particular talents such as singing or dancing well, good storytelling, or hunting prowess. Other features of band organization are small group size, flexible but primarily patrilocal residence, and strong pair bonds between individual men and women. Marriage was exogamous, that is, females were recruited from other groups. These features of hunter-gatherer society were a reflection of a history of ecologic, economic, and social interactions. For example, Efe Pygmy hunter-gatherer men of the Ituri Forest in the former Zaire had very strong relationships with close kin, which facilitated defense of their territories against other cooperative kin groups (Bailey and Aunger, 1989). Moreover, related men could assure women access to valuable resources in neighboring Lese agricultural villages. Also, women were attracted to men who could guarantee long-standing reciprocal economic relationships with Lese villages. Competition for women was high, so close relations with kin may also have helped to obtain marriageable women and provide protection, as some women left Efe society to live in Lese agricultural villages.

Hunter-gatherers were sedentary or nomadic depending on the distribution and dynamics of their resource base and their economic relations with other people. Typically, men hunted and fished while women gathered and collected foods. Sometimes women's work contributed more to the diet and sometimes male hunting and fishing products were most important. Gathering of wild foods tended to contribute more to the diet among people inhabiting tropical and semitropical areas (e.g., Bushmen of the Kalahari) than in northern temperate climates (e.g., the Inuit of Canada) where hunting contributes the bulk of the diet. Staple foods were exchanged with neighbors and more recently in markets (Moran, 1991).

Foragers learned about their environment and resource use through acculturation. Parents taught their children different kinds of ecological knowledge and resource-exploitation strategies. Ecological knowledge is a source of landscape manipulation. For example, the Kayapo Indians of Brazil created forest islands of planted semidomesticated crops of medicinal species, wild yams, and bush bean, as well as domesticated plants such as taro, papaya, and banana (Posey, 1985). A fully grown island had sites that varied in shade and moisture, thereby creating the opportunity for cultivation of different crops. They became, through time, forest patches of varying successional stages within the savanna. Cree Indians of North America rotated their hunting and fishing lands yearly to reduce wildlife disturbance and increase harvests. Biodiversity conservation is, in this case, an indirect effect of resource management. There is evidence that until recently, Indians of Canada used fire to maintain trails and to open up meadows (Lewis, 1989). This provided improved habitat for ungulates and increased hunting success. Australian aborigines used fire to clear trails (of poisonous snakes) and keep game habitat open.

Appropriate use of natural resources was maintained through moral and belief systems of forager societies, which included a strong respect for nature. Through religious belief and social conventions, people revered and exerted some control over their natural resources. These beliefs, however, did not always prevent hunter-gatherers from overusing their resource base. Hunter-gatherers did not always live harmoniously with the environment. Indeed, evidence of escalating overuse is accumulating (e.g., Redford and Mansour, 1996; Lyman, 2003).

Conservation among Hunter-Gatherers

It has been suggested that hunter-gatherer adaptation occurred in environments where resources were freely available to all and were abundant. Thus, the environment was one where subsistence strategies emphasized short-term returns over long-term conservation. But during the Neolithic rise of agriculture, natural ecosystems were compressed and the value of resources increased as relative abundance declined. Some scholars have suggested that self-regulatory mechanisms evolved under resource limitation in some hunter-gatherer societies (Berkes and Folke, 1998).

There has been much written about how hunter-gatherers are actively engaged in conserving resources, especially animal resources. However, the actual data gathered on the subject suggest that subsistence hunters do not conserve prey resources. Most work shows that hunters are concerned about short-term gains and not about resource conservation (Smith and Wishnie, 2000). Small, mobile groups may use resources in a sustainable manner, for example, by maintaining small groups and ranging over a large territory, but this does not necessarily imply that they are consciously conserving resources. Evidence suggests that some resources may be used intensively or even depleted in local areas while other resources are sparingly used. For example, Alvard (1998) showed that the Piro hunters of Peru depleted the large primates in the area around their village but have not yet done so to peccaries. Likewise he showed that the Indonesian Wana nearly depleted their area of macaques (large primates) but hunt pigs in a sustainable manner. Overexploitation of bushmeat in Mozambique (Fusari and Carpaneto, 2006) and in Gabon (Carpaneto *et al.*, 2007) is leading to resource depression of certain species. These and other studies (e.g., among the Inuit of Canada, the Ache of Paraguay, and the Cree of Canada) show that both overexploitation and conservation may be practiced by hunting groups. But the point remains that hunters sometimes reduce prey species to the point of local extinction.

One plausible explanation for resource depletion is that the resources exploited by subsistence hunters are considered to be open-access resources. Open access implies that there are no controls over resource use, which is said to result in the "tragedy of the commons" (Hardin, 1968; National Research Council (NRC), 2002). This concept proposed that deterioration of open-access grazing land is inevitable when individuals see no benefits from resource conservation. Another reason for resource depletion is lack of concern for very abundant resources. Some level of scarcity adds value to a resource relative to when resources are quite abundant. Resource users are motivated to conserve only when they see benefits to nonuse of resources. Thus, it is only when long-term benefits outweigh the short-term benefits that conservation is expected. When tied to a specific resource base and well-defined territories, hunter-gatherers have long-term strategies for natural resource conservation (Alvard, 1998; Arunatoi, 2006). For example, traditional Maine lobstermen have had strong norms of territory ownership, which are enforced through threats of violence and damage to property.

Although foragers may or may not overuse resources, their perception of the land and its value is based on use rights. Local biological diversity is an important element of local survival strategies. This view contrasts with the western view of biodiversity conservation, which is based on Western epistemology. In the western view, nature exists apart from humankind and has value independent of human use. Biodiversity conservation implies no resource use or restraints in resource use.

Processes of Modernization and Hunter-Gatherers

Major changes in hunter-gatherer societies are occurring even in the most remote regions of the world. It is even questionable whether or not hunting and gathering exists today as a distinct economic occupation. These changes are associated with human population growth, land-tenure changes, land-use changes such as agricultural development, infrastructure advancement, resettlement schemes, tree harvesting, mining, and oil exploration, and other types of development. The building of roads makes it easier for outsiders to gain access to remote areas and the resources therein. In addition, hunter-gatherer populations are growing, altering their relationship to the land. The result is that indigenous systems of resource use are changing due to both internal and external pressures. The traditional systems of resource use are not sufficient or are sometimes less effective under current conditions. For example, traditional sanctions to protect, or at least not exhaust, resources are becoming ineffectual as cash income has become increasingly important to individuals interested in cash and

commodities. Hunter-gatherers now have, under these conditions, a growing demand for cash and market goods. Under these conditions, it is less likely that people will give priority to conservation unless it is economically profitable (Daily *et al.*, 2000)

Land Tenure, Institutions, and Biodiversity

One political factor that is almost universally common among hunter-gatherers today is that they often do not own or have control of the land they live on. Until recently, their remoteness meant that they and the resources on which they depended were somewhat protected from outside influences. Thus, resources were locally controlled by informal norms through individual behavior. Now, however, national governments and private entrepreneurs, among others, have put native lands to "productive" use. This means that if the market for some product is strong, it will be exploited or cultivated regardless of environmental impact (Lambin *et al.*, 2001). For example, the strong local demand for *aguaja* (a local plant) in the Peruvian Amazon has led to destructive harvesting. In theory, most hunter-gatherer communities have use rights to their territories but old laws and treaties are continually violated. Legalizing communal resource-use rights is a way of giving hunter-gatherers a long-term stake in conserving the resources on which they depend. Securing rights to resources can occur through various management and development institutions. This means that hunter-gatherers, who formally did not have institutions for collective action in the formal sense, find the need to work with formal western institutions to acquire control over their lands.

The future of biodiversity, conservation, and hunter-gatherer sustainability depends on understanding that there are fundamental differences in the concept of conservation for westerners and for indigenous hunter-gatherers, though this too is changing. Understanding that there are different worldviews toward nature is fundamental to forming a relationship between conservation groups and hunter-gatherer peoples. The reality is that even if hunter-gatherers are using resources, selling wild animals, and cutting down trees, they perhaps remain the most effective conservationists for their region. Therefore, acceptance that there are different ways of knowing the world is a first prerequisite to working with indigenous hunter-gatherer populations. Second, it is necessary to recognize that there are no "pristine" hunter-gatherers and they have needs and aspirations just like the rest of us. Third, securing land tenure for hunter-gatherers and biodiversity conservation is required as a basis of a "sustainable" interaction.

Economic Development and Biodiversity Conservation

Community-based conservation is a concept aimed at involving local people in the conservation of wildlife or protection of biodiversity. The concept developed from the realization that much of the planet's wildlife and biodiversity exist outside protected areas and in regions occupied by rural people in developing countries (Galvin *et al.*, 2008). Models of community-based conservation adhere to the notion that if local communities can derive some value, nominal income, through conserving biodiversity, they will do so. This promising concept has been widely promoted as "the answer" to conservation in developing countries. Thus, several models of community-based conservation have developed (Goldman, 2003; Berkes, 2007) such as Integrated Conservation–Development projects. However, results from community-based conservation projects suggest that there are more failures than successes (Goldman, 2003; Berkes, 2004). Many community-based conservation efforts involve local communities in name only. Locals are neither involved in project identification and planning, nor are they beneficiaries; thus, these projects are not really community-based conservation projects. Other scenarios for failure also have in common insufficient involvement of the local people at all levels in the project. For community-based conservation to work, people need to be considered a component of the system being conserved and brought into the project process from the beginning (Carlson and Berkes, 2005; Reid *et al.*, 2009).

It is Useful to View Hunter-Gatherers and Biodiversity as a Social–Ecological System

One of the fundamental problems with community-based conservation is that hunter-gatherers as well as other indigenous populations are often viewed as an external disturbance to the natural system rather than as integral components of the ecosystem. But hunter-gatherer societies see their relationship with the environment as one: they are part of that environment. Though not a study on foragers, but rather herders who do some hunting and gathering, the South Turkana Ecosystem Project (Ellis and Swift, 1988; Little and Leslie, 1999) is one of the only truly interdisciplinary and long-term projects to study the social behavior, knowledge systems, demography, human biology, and ecology of a group of people. An important goal of this study was to understand how the environment affected human management and how people affected the environment. In this case, people and livestock (camels, cattle, sheep, goats, and donkeys) lived in a harsh, dry, and highly seasonal environment. This assemblage of people, livestock, plants, and other organisms within a semiarid ecosystem produced a remarkably interactive system.

Vegetation structure in this tropical savanna and dry woodlands was shown to be hierarchically constrained by physical factors: by climate at regional scales, by topography and geomorphology at landscape scales, and by water redistribution and disturbance at local and patch scales; livestock and humans played a small role. The pastoralists did influence vegetation composition and cover by burning, woodcutting, and through seed distribution by livestock. These influences were small. Livestock ecology and production followed those of the seasonal dynamics of plants. The different patterns of forage utilization by different herbivores, plus differential habitat use, led to almost complete niche separation among this suite of domestic herbivores; among all five species, they managed to utilize a wide variety of the available plant types in the ecosystem. Thus, physical heterogeneity of the Turkana landscapes ultimately resulted in

spatial and temporal variation in plant production, plant life form diversity, and refuge areas for pastoralists. These, in turn, contributed to social and ecological persistence by reducing variability of ecosystem energy flow and long-term variations in species diversity. Thus, biodiversity was important to ecosystem (which included people) maintenance. These systems' approach to understanding human–environment interactions is a useful way to discern the ecological impact of hunter-gatherers and, more importantly, to derive appropriate management of lands where hunter-gatherers live.

The description presented here shows that indigenous concepts of conservation, ecological knowledge, and moral and religious beliefs are fundamental to understanding how hunter-gatherers use resources. Not all hunter-gatherers conserve their resources, thus whether or not and to what extent hunter-gatherers effect their environment is an empirical question that needs to be investigated, not a notion to be assumed one way or another. It is, however, the case that when hunter-gatherers have short-term strategies for resource use, they may overuse some resources; when long-term goals are in place, they may not. Informal institutions control use of some resources in hunter-gatherer societies, but collective action or formal institutions are generally not well developed. With major changes in and around the lands inhabited by hunter-gatherers, it is becoming increasingly necessary for hunter-gatherers to develop institutions to gain control over their resource base. Alliances between hunter-gatherers and others interested in conservation may facilitate resource-management strategies that reduce the impact of negative changes. Hunter-gatherer natural resource-management strategies that include their social system are important attributes of these social–ecological systems and need to be the fundamental components of any plan to conserve biodiversity.

See also: Biodiversity-Rich Countries. Ethnobiology and Ethnoecology. Indigenous Peoples and Biodiversity. Land-Use Patterns, Historic. Religious Traditions and Biodiversity. Traditional Conservation Practices

References

Alvard MS (1998) Evolutionary ecology and resource conservation. *Evolutionary Anthropology* 72: 62–74.

Arunatoi N (2006) Moken traditional knowledge: An unrecognized form of natural resources management and conservation. *International Social Science Journal* 58: 139–150.

Bailey RC and Aunger Jr. R (1989) Significance of the social relationships of Efe Pygmy men in the Ituri Forest, Zaire. *American Journal of Physical Anthropology* 78: 495–507.

Berkes F and Folke C (1998) *Linking Social and Ecological Systems, Management Practices and Social Mechanisms for Building Resilience*. Cambridge: Cambridge University Press.

Berkes F (2007) Community-based conservation in a globalized world. *Proceedings of the National Academy of Sciences of United States of America* 104(23): 15188–15193.

Berkes F (2004) Rethinking community-based conservation. *Conservation Biology* 18(3): 621–630.

Cane S (1996) Australian aboriginal subsistence in the western desert. In: Bates DG and Lees SH (eds.) *Case Studies in Human Ecology*, pp. 17–54. New York: Plenum Press.

Carlsson L and Berkes F (2005) Co-management: Concepts and methodological implications. *Journal of Environmental Management* 75: 65–76.

Carpaneto GM, Fusari A, and Okongo H (2007) Subsistence hunting and exploitation of mammals in the Haut-Ogooue Province, Soth Eastern Gabon. *Journal of Anthropological Sciences* 85: 183–193.

Daily GC, Söderqvist T, Aniyar S, *et al.* (2000) The value of nature and the nature of value. *Science* 289: 395–396.

Ellis JE and Swift DM (1988) Stability of African pastoral ecosystems: Alternate paradigms and implications for development. *Journal of Range Management* 41: 450–459.

Fusari A and Carpaneto GM (2006) Subsistence hunting and conservation issues in the game reserve of Gile, Mozambique. *Biodiversity and Conservation* 15: 2477–2495.

Galvin KA, Reid RS, Behnke RH, and Hobbs NT (eds.) (2008) *Fragmentation of Semi-Arid and Arid Landscapes. Consequences for Human and Natural Systems*. Dordrecht, The Netherlands: Springer.

Goldman M (2003) Partitioned nature, privileged knowledge: Community based conservation in Tanzania. *Development and Change* 34(5): 833–862.

Hardin G (1968) The tragedy of the commons. *Science* 162: 1243–1248.

Lambin EF, Turner BL, Geist HJ, *et al.* (2001) The causes of land use and land cover change: Moving beyond the myths. *Global Environmental Change* 11: 261–269.

Lewis HT (1989) Ecological and technical knowledge of fire: Aborigines versus park managers in Northern Australia. *American Anthropologist* 91: 940–961.

Little MA and Leslie PW (eds.) (1999) *Turkana Herders of the Dry Savanna: Ecology and Biobehavioral Response of Nomads to an Uncertain Environment*. Oxford: Oxford University Press.

Lyman RL (2003) Pinniped behavior, foraging theory, and the depression of metapopulations and non depression of a local population on the southern northwest coast of North America. *Journal of Anthropological Archaeology* 22: 376–388.

Moran EF (1991) Human adaptive strategies in Amazonian black-water ecosystems. *American Anthropologist* 93: 361–382.

National Research Council (NRC) (2002) The drama of the commons. In: Ostrom E, Dietz T, Dolšak N, *et al.* (eds.) *Committee on the Human Dimensions of Global Change*. Washington, DC: National Academy Press.

Posey DA (1985) Indigenous management of tropical forest ecosystems: The case of the Kayapo Indians of the Brazilian Amazon. *Agroforestry Systems* 3: 139–158.

Redford KH and Mansour JA (1996) *Traditional Peoples and Biodiversity Conservation in Large Tropical Landscapes*. Arlington, VA: America Verde Publications. The Nature Conservancy Latin America and Caribbean Division.

Reid RS, Nkedianye D, Said MY, *et al.* (2009) Evolution of models to support community and policy action with science: Balancing pastoral livelihoods and wildlife conservation in savannas of East Africa. *Proceedings of the Academy of Sciences of United States of America*. www.pnas.org/cgi/doi/10.1073/pnas.0900313106.

Smith EA and Wishnie M (2000) Conservation and subsistence in small-scale societies. *Annual Review of Anthropology* 29: 493–524.

Hybridization in Plants

Pamela S Soltis, University of Florida, Gainesville, FL, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Robert S Fritz, volume 4, pp 659–675, © 2001, Elsevier Inc.

Glossary

Allopolyploid Species with chromosome sets derived from interspecific hybridization and chromosome doubling.

Hybrid An individual that results from crossing between individuals from two or more populations, races, subspecies, species, etc., including individuals that are F1s, F2s, and the set of all backcrosses.

Hybridization Interbreeding of individuals from two or more populations of species, subspecies, or races.

Hybrid swarm Mixture of hybrid genotypes, including F1s, F2s, and backcrosses, due to hybridization between two or more species that co-occur in a locality.

Hybrid zone Area of overlap or point of contact between two populations that are distinguishable based on one or more heritable characters where viable or partially fertile hybrids are formed.

Introgression Permanent transfer of genes from one or more species, subspecies, or race into another species, subspecies, or race via hybridization and backcrossing.

Reticulate evolution Pattern of speciation where new species arise from interspecific hybridization between two species coupled with establishment of reproductive isolation.

Introduction

Despite common use of the term "hybridization" in plant biology, application of the term varies widely. To some, hybridization involves crossing between two species and the production of fertile offspring (although fertile progeny are not required in all definitions). At the opposite end of the genetic spectrum, hybridization may refer to crossing between any two individuals, as opposed to self-fertilization. Still others recognize the intergradation, and sometimes arbitrary distinction, between differentiated populations, races, subspecies, and species and therefore consider hybridization to occur when individuals of differentiated populations or population systems interbreed (see Arnold, 1997; Levin, 2002). These definitions are not merely semantic differences – they have implications for our understanding of the frequency, genetics, and ultimate role of hybridization in generating, or eroding, biodiversity. In this article, interactions between species are focused, with the understanding that species represent an arbitrary level of the biological hierarchy and that species limits are not always clear.

Hybridization has variously been viewed as an engine of biodiversity (Arnold, 1997, 2004; Anderson and Stebbins, 1954) and as evolutionary noise (Stebbins, 1971; Wagner, 1970), and it is likely both in different circumstances. Rare hybridization between two largely allopatric species may have little or no long-term effect on the gene pool of either species, particularly if the hybrids are sterile; and this scenario, which is probably quite common, clearly represents evolutionary noise. In contrast, hybridization, particularly when coupled with polyploidy (genome doubling; whole-genome duplication (WGD)), results in speciation, with the generation of novel phenotypes that occupy novel niches (see Soltis and Soltis, 2009). In this article, an overview of the varying extent of hybridization and its associated evolutionary consequences has been provided, with an emphasis first on homoploid hybridization (i.e., hybridization involving species of the same ploidy, typically both diploids) and then on polyploidy and hybridization.

Homoploid Hybridization

Hybridization between two species may have little or no impact on either species in either the short or long term. However, hybridization may have evolutionarily significant consequences for the participating species, ranging from small genetic exchanges to the origin of species.

Introgression

Introgression – the transfer of genetic material between species following hybridization and backcrossing to the parental species – is a fundamental concept of plant evolutionary biology (Anderson and Hubricht, 1938; Anderson, 1949). However, it is difficult to demonstrate in natural populations, and its extent in nature is therefore unknown. The elegant treatment of the topic by Anderson (1949), coupled with painstakingly described morphological variation in key groups of plants (e.g., Heiser (1949) on sunflowers), led botanists for many years to believe that introgression played a significant role in plant evolution. However, later perspectives on the difficulty of distinguishing between "primary intergradation" (i.e., divergence of a population system into two species) and "secondary intergradation" (hybridization; e.g., Mayr, 1942), along with limited genetic markers for actually testing for hybridization, led many botanists to question how common hybridization and introgression really were.

Then along came deoxyribonucleic acid (DNA) markers, especially the chloroplast genome (cpDNA). In the 1980s, cpDNA restriction site markers were being applied to studies of plant phylogeny. These markers were ideal for reconstructing the evolutionary history of closely related species (e.g., Palmer *et al.*, 1985), such as those classified in the same genus, and tremendous progress was achieved in plant systematics in a very short time. However, these studies did more than clarify relationships among plant species: they also showed the underlying history of interspecific hybridization.

In many plant groups (e.g., *Salix*, willows; Brunsfeld *et al.*, 1992), instances arose where populations of a morphologically uniform species did not appear to be monophyletic (i.e., descended from a single common ancestor) in the phylogenetic trees derived from cpDNA data. Typically, one population, which occurred sympatrically with another species, had the cpDNA of the sympatric species rather than its "own" cpDNA. This phenomenon, termed chloroplast capture (Rieseberg and Soltis, 1991), resulted from hybridization between two species, followed by continued backcrossing to the paternal parent to yield a population with the nuclear genotype – and therefore general phenotype – of one species and the cpDNA (which is maternally inherited) of the other, in other words, introgression of the cpDNA of one species into the other species. Despite the seemingly low probability of these events occurring, hundreds of cases of chloroplast capture have been reported, making chloroplast introgression one of the most common – and significant – outcomes of hybridization.

Novel Genotypes Through Hybridization and Introgression

Introgression, whether of nuclear genes or cpDNAs, provides the opportunity to generate novel genetic combinations on which selection can act. Classic models of hybridization and introgression argue that a hybrid or introgressant requires a novel "hybridized" habitat if it is to survive – it simply cannot compete with its established parents in either of their habitats. For this reason, offspring of hybridization and introgression are predicted to occur (and often do; Stebbins, 1985; Bayer *et al.*, 1991) in habitats that are intermediate between those of its parents.

This sort of novelty is perhaps best seen in several hybrid species of sunflowers. *Helianthus annuus*, the common sunflower, has hybridized extensively with other species of *Helianthus*, and on multiple occasions hybridization has led to species formation (Rieseberg 2006; *see* Hybrid Speciation). One hybrid species, *Helianthus anomalus*, occurs on sand dunes in the deserts of the southwestern USA, an extreme habitat that is differentiated from the prairie habitats of both parents (Rieseberg *et al.*, 2007). Likewise, *Helianthus paradoxus* occupies extremely saline environments, a novel habitat perhaps conquerable because of the genetic novelty of the hybrid genotype (Rieseberg *et al.*, 2007).

Certainly, genetic mechanisms for hybrid novelty can be proposed, such as combinations of genes that yield "hybrid vigor" or heterosis. In addition, novel alleles may be generated from the parental copies through recombination, such that half of an allele may be contributed by one parent and the other half from the second parent (Golding and Strobeck, 1983). Although rare, such mutations may generate the types of genetic combinations that lead to morphological and physiological novelty.

The origin of genetic novelty in hybrids or introgressants may not be limited to the nuclear genome. Although the transfer of cpDNAs is often considered a neutral change, the many examples of chloroplast capture may in fact represent the outcomes of positive selection in which new combinations of chloroplast and nuclear genes have a selective advantage. Although data for the mitochondrial genome (mtDNA) are much more limited than for cpDNA, it is possible that mtDNA is also transferred along with the cpDNA through hybridization and introgression. However, given the relatively slow rate of nucleotide substitution in the mtDNA (Palmer, 1992) and therefore the high genetic similarity between mtDNAs of related species, the transfer of mtDNA may not yield many new genetic combinations, in contrast to the possibilities with cpDNA.

Loss of Novelty Through Hybridization and Introgression

Hybridization may also retard genetic differentiation and lead to loss of genetic and biological diversity. For example, divergence between populations is a function of selection, drift, and gene flow, and the same is true at the interface of two species. Gene flow (i.e., hybridization) can retard or prevent diversification. Furthermore, genetic differences between species can be eroded by gene flow, or hybridization, between them, leading to a breakdown of species barriers and the potential loss of one species or the other.

The loss of the genetic integrity of a species often occurs when a rare species hybridizes with a more widespread relative, often with detrimental consequences for the rare endemic (Levin *et al.*, 1996). For example, human-mediated movement of species, such as introductions to islands, can pose a serious risk to the rare and endangered species of a given area. Hybridization with a more aggressive and more abundant relative may lead to swamping out of the genetic characteristics of the rare species, and possibly extinction. The Catalina Island mahogany (*Cercocarpus traskiae*) has been threatened by hybridization with *Cercocarpus betuloides* for many years, and conservation measures to limit hybridization and its aftermath are underway (Rieseberg, 1991). Similar situations exist for many rare species (Levin *et al.*, 1996).

Hybrid Swarms

Field botanists are more than familiar with the concept of hybrid swarms, which arise when interspecific hybrids are mostly or completely fertile. Following hybridization, fertile hybrids can self-fertilize or cross with other hybrids (and hybrid derivatives) or their parents. These mating opportunities provide a baffling array of genetic outcomes – even when only a single gene is considered. When the polygenic nature of most phenotypic traits is considered, the range of genotypes leading to any given phenotype is mind boggling and leads to the extensive phenotypic variation observed in the field when such hybrid swarms occur. Such swarms provide the ideal

conditions under which introgression can occur. Hybridization between species that can produce fertile offspring therefore leads to extensive genetic novelty on the one hand and perhaps loss of genetic uniqueness, through erosion of parental genotypes, on the other.

Hybrid Speciation

Homoploid hybrid speciation – that is, the origin of a new species via hybridization between two species without a change in chromosome number – is one of the most spectacular outcomes of hybridization (Rieseberg and Willis, 2007). This mode of speciation has been a staple of plant evolutionary biology for decades, but despite its conceptual framework, very few cases of homoploid hybrid speciation have been unequivocally demonstrated. The model of hybridization, followed by the sorting of a unique phenotype that moves into a novel habitat and therefore becomes isolated from its parents, is straightforward. However, the number of putative hybrid species greatly outnumbers those that have been analyzed in detail; and of those analyzed, very few provide convincing evidence of hybrid species formation (perhaps only 20 good examples; Gross and Rieseberg, 2005). Part of the problem is that, even with excellent genetic markers, it is difficult to determine (1) if a population or set of populations arose via hybridization between two (or more) other species and (2) if this set of populations should be recognized as a distinct species (Gallez and Gottlieb, 1982; see discussion in Soltis and Soltis, 2009).

Hybrid species are often expected to be "additive" of their parental traits, with this additivity manifested as intermediacy. However, numerous studies have shown that, for a single trait, hybrids may be intermediate, or they may resemble one parent or the other or exhibit a novel phenotype (McDade, 1990; Rieseberg, 1995). When the many traits of a plant are considered, hybrids are probably never true intermediates, but mosaics of traits that span the range of parental resemblance through intermediacy to novelty. The range of mosaic genotypes – and their expressed phenotypes – may make hybrid species highly polymorphic, with some genotypes at selective advantages and disadvantages.

The best-studied hybrid species in plants occur in *Iris* and sunflowers (*Helianthus*). Studies of the Louisiana *Iris* group (Riley, 1938; Anderson, 1949) served as the basis for the development of Anderson's ideas on introgression. Three hybridizing species – *Iris fulva*, *Iris brevifolia*, and *Iris hexagona* – have contributed to the complex formation of *Iris nelsonii* (Arnold, 1993). This species complex has provided opportunities for studying the genetic architecture of interspecific traits and adaptation to different habitats.

Multiple and repeated examples of homoploid hybrid speciation have occurred in *Helianthus* (*see* Novel Genotypes Through Hybridization and Introgression; Heiser, 1949; Rieseberg *et al.*, 1995, 2006). The widespread cultivated sunflower, *H. annuus*, has hybridized with *Helianthus petiolaris* multiple times, and these hybridization events have yielded three hybrid species: *H. anomalus*, *Helianthus deserticola*, and *H. paradoxus*. Although these species share the same parental species, they reflect separate origins and different combinations of parental genes. The consequence is that each hybrid species is unique, and each occupies a specialized, and environmentally extreme, habitat. *H. anomalous* occurs on sand dunes in the Southwest, *H. paradoxus* occurs in saline wetlands in Texas and New Mexico, and *H. deserticola* is found in the deserts of the Great Basin (Rieseberg *et al.*, 2007). Despite their divergence in ecological tolerances, these three derivative species have undergone concerted chromosomal rearrangements (Rieseberg *et al.*, 1995).

Hybridization and Polyploidy

Although hybridization at the homoploid level (*see* Hybrid Speciation), has significant consequences for plant evolution, hybridization manifests itself most strongly when combined with polyploidy. Both the number of species generated through polyploid origins and the extent of genetic and phenotypic novelty generated through polyploidy are greater than those for homoploid hybrids. Herein, both patterns and processes of polyploid evolution in plants are addressed, emphasizing the role that polyploidy has played in structuring plant genomes and how this complex genome structure results in genetic, phenotypic, and evolutionary novelty.

Types of Polyploids

Polyploids are typically classified as "autopolyploids" or "allopolyploids." Although these terms themselves have a long and complicated history, today, autopolyploids are considered to have arisen from within a single species, and allopolyploids constitute those that arose via the combined processes of hybridization and chromosome doubling. Of course, nature does not follow these strict definitions, and polyploids tend to occupy a continuum from polyploidization within a single individual to interspecific hybridization and polyploidization; furthermore, some polyploids have only slightly differentiated chromosome sets, a situation Stebbins (1947, 1950) referred to as "segmental allopolyploids." Despite these differences, most authors agree that most polyploids are derived from some form of hybridization – either between species or between individuals of the same (or different) population(s).

Ancient Polyploidy

Polyploids are very common among extant plants, with traditional estimates averaging approximately 50% (and some much higher) of all species of flowering plants being of polyploid origin (Stebbins, 1947, 1950, 1971; Grant, 1981; Goldblatt, 1980; Masterson, 1994). Polyploidy is also common in bryophytes and ferns. In fact, based on chromosome numbers, many of the most ancient lineages of land plants may be ancient polyploids. This hypothesis is based on chromosome numbers, and although it has been tested with genetic data, the arrival of genome sequencing promises a new and refreshing perspective on ancient polyploidy in plants.

The green plants are a clade of approximately half a million species with a fossil record of nearly 1 billion years. About 450 million years ago (Ma), the first plants emerged on land, and

their subsequent diversification was rapid and astounding. Today, the embryophytes (land plants) constitute liverworts, mosses, hornworts (these three collectively termed "bryophytes" although they do not form a clade), and a large clade of vascular plants (tracheophytes). The tracheophytes consist of the lycophytes (club mosses, spike mosses, and quillworts), the monilophytes (*Psilotum*, *Equisetum*, and ferns, among others), and the seed plants gymnosperms and angiosperms (flowering plants). The tracheophytes extend back in the fossil record to more than 400 Ma, with the angiosperms representing the most recently originating group, with the oldest fossils dating to approximately 130 Ma.

Ancient Polyploidy in Tracheophytes

Ongoing polyploid formation is common in tracheophytes, with polyploid species complexes (*see* Polyploid Complexes) of diploid and polyploid species in ferns and angiosperms in particular. An interesting observation of many ancient lineages of tracheophytes (e.g., lycophytes, *Psilotum*, *Equisetum*, and most homosporous ferns) is that they have very high chromosome numbers (on average, $n=57$). These high chromosome numbers led to the hypothesis that these groups were ancient polyploids whose diploid ancestors had gone extinct, given the antiquity of these groups (see Klekowski and Baker, 1966). Given the apparently high self-fertilization rates of these plants, Klekowski and Baker (1966) argued that the genetic polymorphism housed in the multiple chromosome sets of an ancient polyploid could offset the detrimental effects of selfing. This elegant hypothesis was favored until the use of genetic markers, developed in the 1960s for analysis of human genetics, made its way to studies of ferns.

The first studies of isozyme genetic markers in ferns clearly showed patterns of gene expression expected for a diploid individual, based on studies of diploid animals and flowering plants: one band for a homozygote and two for a heterozygote. If ferns were ancient polyploids, housing reserves of genetic variation because of their self-fertilizing breeding system, they should have complex genetics, with many isozyme bands, but this was not the case. However, ferns known to be more recent polyploids did have the complex banding expected of a polyploid. Given these results, Haufler and Soltis (1986) proposed that ferns with high chromosome numbers are not, in fact, ancient polyploids, but genetic diploids. These results could be interpreted in the following ways: (1) ferns were not now, and never had been, polyploid – their high chromosome numbers were achieved through some alternative mechanism, such as chromosome fission; or (2) ferns were, in fact, ancient polyploids but their genomes had been "diploidized" such that gene expression patterns were typical of diploids rather than polyploids.

These results for ferns triggered similar analyses in other ancient groups of plants with high chromosome numbers. Lycophytes classified in Lycopodiaceae (*Lycopodium*, *Huperzia*, and relatives; lowest numbers for each genus range from $n=22$ to 67) also have high chromosomes, as do *Psilotum* ($n=52$) and *Equisetum* ($n=108$). Application of isozyme genetic markers to all of these groups also revealed diploid gene expression despite high chromosome numbers (Soltis, 1986; DE Soltis and PS Soltis, 1988; PS Soltis and DE Soltis, 1988). Thus, if these plant groups are truly ancient polyploids, they have lost their polyploid attributes and seem completely diploidized ("completely," i.e., given the small number of genes examined). These plants await more sophisticated analyses, such as those now possible using genomic data. For example, a complete genome sequence has been obtained for a species of *Selaginella*, a spike moss, and no evidence exists for ancient polyploidy. Of course, *Selaginella* and its sister group, *Isoetes*, are among only a few groups of these ancient lineages that do not have the high chromosome numbers ($n=7$ and $n=10$, respectively) thought to be associated with ancient polyploidy. Genomic resources, or a complete genome sequence, for one of the other groups would be useful for addressing the question of whether or not these plants are ancient polyploids.

Ancient Polyploidy in Angiosperms

As in the tracheophytes generally, several lineages of angiosperms have consistently high chromosome numbers: Magnoliaceae (magnolia family), Myristicaceae (nutmeg family), Calycanthaceae (sweetshrub family), Lauraceae (laurel family), Salicaceae (willow family), to name a few. Most of these occur in the magnoliid clade, an ancient group that includes all of these families except Salicaceae. Stebbins (1950) hypothesized that these groups were all ancient polyploids, with additional, more recent polyploids present in several of these groups. In contrast to the data for lycophytes, *Psilotum*, *Equisetum*, and ferns application of isozyme genetic markers to these plant groups demonstrated complex banding patterns consistent with ancient polyploidy (Soltis and Soltis, 1990).

Unlike for tracheophytes, however, additional genomic resources have been developed for several of these putatively ancient polyploid angiosperms, and these allow for further investigation of ancient polyploidy. For example, the entire genome of *Populus* (in the willow family) has been sequenced, and this sequence reveals evidence of ancient tetraploidy (Tuskan *et al.*, 2006). Although not complete genome sequences, genomic resources for members of Magnoliaceae and Lauraceae also demonstrate ancient polyploidy (Cui *et al.*, 2006).

Interestingly, analyses of whole-genome sequences for several species of angiosperms with low chromosome numbers also show evidence of very ancient polyploidy. Even *Arabidopsis*, chosen as the first plant to have its genome sequence because its genome is very small, is actually an ancient polyploid, as are its relatives in the mustard family (Brassicaceae). In fact, a WGD can be traced, using genetic data, to the common ancestor of all angiosperms, with another WGD having occurred in the common ancestor of all seed plants (Jiao *et al.*, 2011).

These genome-based analyses for angiosperms demonstrate that all angiosperms are polyploid, but they have experienced extensive, and sometimes differential, levels of diploidization, followed by additional polyploidization, etc. It will be interesting to apply similar analyses to the older lineages of tracheophytes to see if they, too, harbor evidence of ancient polyploidy. Our expectation would be that this evidence would be even harder to find, given the older ages of these groups.

Polyploid Complexes

Polyploids and their diploid progenitors often form complexes of related and hybridizing species. A group of related diploid species may hybridize and form polyploid derivatives in a variety of combinations. These polyploids may backcross to various diploids, or they may hybridize among themselves and form additional high-level polyploids ($4x \times 2x = 6x$, $4x \times 4x = 8x$, etc.). These species may subdivide the habitat in such a way that new intermediate niches are occupied by the new species, which have been referred to as "fill-in taxa" (e.g., in *Antennaria* (pussytoes); Bayer *et al.*, 1991).

Complex patterns of hybridization and polyploid formation among a group of species may lead to the formation of a "compilospecies" – a group of interacting species that share a gene pool across ploidal levels. Although an interesting concept, few examples of compilospecies have been documented, although recent molecular data for the grass *Paspalum* (bahia grass) in South America are consistent with the compilospecies concept. Extensive gene flow among populations and between ploidal levels in a compilospecies can generate substantial genetic and phenotypic novelty that may confuse scientists but have significant ecological and evolutionary impact.

Polyploid species complexes may also generate asexual lineages that arise via the production of sterile hybrid offspring. Asexual diploid or odd-ploid lineages may perpetuate themselves via a combination of methods, such as vegetative reproduction and apomixis, the production of seeds without sexual reproduction. *Rubus*, which includes raspberries and blackberries, is notorious for forming asexual lineages associated with hybrids and polyploids. Anywhere from a dozen to hundreds of species of *Rubus* have been recognized based on morphological variants, most of which are either apomictic or otherwise asexual. Apomictic polyploid species complexes also tend to occur at high altitude and high latitude, in groups such as *Draba* (in the mustard family) and *Antennaria* (in the sunflower family).

Polyploidy and Diversification

Given the foregoing discussion, it should be obvious that hybridization and polyploidy generate extensive genetic and phenotypic novelty on which selection can act. Some groups seem to undergo unreduced gamete formation – the most common process by which genomes are duplicated – much more frequently than others (Ramsey and Schemske, 1998, 2002), and polyploidy is therefore unequally distributed across even the plant branch of the tree of life.

Analyses of the frequency of polyploids, in the context of their diploid relatives, have yielded estimates on the frequency of polyploid plant speciation. These estimates range from a low of 2–4% of angiosperm speciation events and 7% of fern speciation (Otto and Whitton, 2000) to 15% of angiosperm speciation events and 31% of fern speciation (Wood *et al.*, 2009) – values that yield standing estimates of polyploid frequency much closer to estimates based on chromosome numbers than do the estimates by Otto and Whitton (2000). These estimates, although highly speculative, relate to the origin of polyploid species. However, rates of diversification must take both formation and extinction into account.

One way of examining diversification is by plotting species numbers across a phylogenetic tree – a depiction of the overall branching history of evolution. When investigating a causal factor for increased or decreased diversification (the net of speciation and extinction), species numbers are examined relative to the origin of a specific characteristic. For example, in an analysis of the possible causes of the early diversification of the angiosperms, Davies *et al.* (2004) tested such attributes as pollination syndrome, geographic distribution, dispersal mode, habit, and chromosome number, but none of them, at least alone, is significantly associated with diversification early in angiosperm history. However, although chromosome number was tested as a proxy for ploidy, ploidy itself was not investigated. Because of the dramatic changes that can occur to a genome after polyploidization, leading to a chromosome number that might no longer be recognized as polyploid, perhaps chromosome number itself is not an appropriate character; alternatively, ploidy might not have been significantly associated with diversification early in angiosperm history.

Plotting known genome duplication events on the phylogenetic tree for angiosperms suggests that, in fact, diversification at a deep level may be linked to increases in ploidy (Soltis *et al.*, 2009). A genome duplication event occurred before the origin of the large core eudicot clade (which contains ~70% of all angiosperm species), and additional duplications are associated with the origins of some of the largest families: grasses, legumes, and the potato/tomato family (Soltis and Soltis, 2009). Although more rigorous studies are needed, along with more information on the evolutionary history of WGDs, these analyses suggest that polyploidization events deep in the evolutionary history of angiosperms may have triggered rapid diversification.

In contrast to these results, Mayrose *et al.* (2011) reported that diversification rates among polyploids are lower than those among their diploid relatives, consistent with views by Stebbins, for example, that polyploid genomes provided buffering that retarded the rate of evolution and speciation. Although the results of Soltis and Soltis, 2009 and Mayrose *et al.* (2011) differ, this difference may actually be one of scale. Soltis and Soltis, 2009 examined diversification rates deep in the evolutionary tree, whereas Mayrose *et al.* (2011) focused on the tips. It is very possible that polyploidization events are like other types of mutations – most are lost as detrimental, but some persist as positive changes. Clearly, additional research is needed in this important area.

Recent Polyploidy

Hybridization and polyploidy may generate new biodiversity – often recognized as species – and they can have both positive and negative impacts on existing communities. Understanding the role that these processes play in generating and eroding biodiversity is crucial for establishing sound conservation practices. However, till date, data on hybridization and polyploidization are often anecdotal and not predictive. Furthermore, most polyploids have changed substantially since their

time of formation, making it difficult, if not impossible, to evaluate the contributions that are due to hybridization versus those due to polyploidization versus novelty that may have arisen within the polyploid species.

A handful of polyploid species have originated during the last 150 years. We know their time of origin because their parentage can be traced to human-mediated movement of plants, often transoceanically. Thanks to specimens housed in herbaria and other museum collections, it has been possible to pinpoint within a few years when certain polyploid species originated. The best studied of these polyploids are two tetraploids from the Pacific Northwest of USA. *Tragopogon mirus* and *Tragopogon miscellus* (in the sunflower family) formed in the early 1900s following the introduction of their parental species from Europe (**Figure 1**; Ownbey, 1950; Soltis *et al.*, 2004). These species have provided the opportunity to investigate what happens in a new polyploid species and whether or not there are any principles that may govern how polyploids evolve.

Duplicated genes have several possible fates: (1) both copies of the gene may be retained and both can continue to function as the original gene did; (2) one copy can be lost or silenced, leaving the second copy to carry out the original functions; (3) one copy can retain the original function and the second copy can take on new functions (called neofunctionalization); and (4) the two copies can subdivide the original functions so that each carries out a part of the original function, perhaps in a specific tissue or at a specific time, etc. When an entire genome is duplicated, as in a polyploid, all of the genes are duplicated, and these options then occur for each gene. The resolution of these options makes for a very complicated set of possibilities, particularly when one estimates that a diploid plant has approximately 25,000–30,000 genes (about the same number as humans, mice, and fruit flies). Does each polyploid plant species resolve these options independently, or are there some rules that control which genes are retained in duplicate, which are lost, and which are neo- or subfunctionalized?

The question of what happens to duplicate genes in a new polyploid has been addressed in the two *Tragopogon* tetraploids. Surprisingly, in a series of studies (most recently Buggs *et al.*, 2012 and references therein), both species were found to lose copies of their genes much more frequently than they underwent gene silencing or neo- or subfunctionalization (most genes were retained in duplicate). Amazingly, more than 10% of the duplicate genes surveyed actually lost one member of the pair. Although this might be an expectation for an ancient polyploid that has undergone diploidization, it was not expected in recent polyploid that is only approximately 80 years old. Furthermore, certain types of genes – those encoding particular structural proteins – were more often retained than various types of regulatory genes, in contrast to theory on gene loss and to data for *Arabidopsis*.

The mechanisms for such genetic changes are unknown at this point, but abnormal pairing and breakage of chromosomes could lead to loss of duplicate genes. Interestingly, *Tragopogon* polyploids exhibit extensive chromosomal variation relative to their diploid parents and to all expectations of additivity for an allopolyploid, and this variability may result in the unusual genetic patterns observed (Chester *et al.*, 2012). The phenotypic manifestations of this genetic and chromosomal variation are unknown, but could be substantial (**Figure 2**).

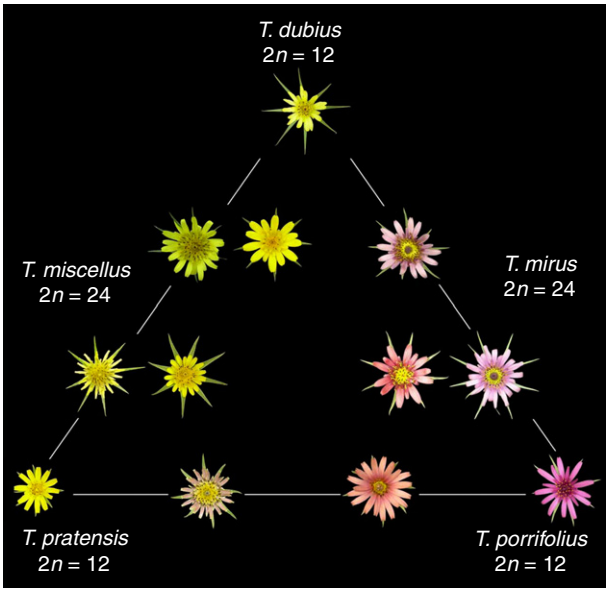


Figure 1 Diagram showing the parentage of the two recent allotetraploids, *T. mirus* and *T. miscellus*, from their diploid parents, *Tragopogon dubius* and *Tragopogon porrifolius* and *T. dubius* and *Tragopogon pratensis*, respectively. The diploids are placed at the points of the triangle, and the synthetic (artificially produced) allopolyploids are on the lines between the parents, each near its maternal parent. The three naturally occurring allopolyploids are located inside the triangle. Reprinted with permission from Cold Spring Harbor Symposia on Quantitative Biology.

Ecological Consequences of Hybridization

Hybridization and polyploidization generate novelty that may be manifested as ecological adaptations, such as the ability to occupy niches separate from those of the parental species. However, the ecological implications of hybridization on biodiversity extend beyond single plant species, or even plant communities. The diversity and structure of plants are used to define all ecological communities. But it is the biological diversity of organisms that live on plants – their herbivores, pathogens, pollinators, dispersers, root and leaf mutualists, and the natural enemies of these organisms – that contributes the most to total biological diversity. Interspecific hybridization of plants contributes to biological diversity in fundamental ways. Beyond generating biological diversity through introgression, speciation, polyploidization, and the processes described in the preceding sections, hybridization in plants increases overall biological diversity through its influence on the abundance and diversity of herbivores and their natural enemies, and their pathogens. In this section, the effects of plant hybrids on the biological diversity and community structure of higher trophic levels, herbivores, their natural enemies, and pathogens are considered. The conceptual model proposed in this section suggests that plant



Figure 2 Capitula of *Tragopogon porrifolius*, *T. dubius*, and representative synthetic diploid hybrids (2x) and allopolyploids (4x) resulting from crosses between these two diploid species. (a) *T. porrifolius*, (b) *T. dubius*, (c–f) *T. porrifolius* (maternal) × *T. dubius* (paternal) – (c) 4x, (d) 4x, (e) 4x, (f) 2x, (g–j) *T. dubius* (maternal) × *T. porrifolius* (paternal) – (g) 2x, (h) 4x, (i) 2x, (j) 2x. Panels (c) and (g) are exact reciprocals (i.e., the same diploid individuals were used in reciprocal crosses). Reprinted with permission from Tate JA, Symonds VV, Doust AN, *et al.* (2009) Synthetic polyploids of *Tragopogon miscellus* and *T. mirus* (Asteraceae): 60 Years after Ownbey's discovery. *American Journal of Botany* 96: 979–988.

hybridization contributes to genetic diversity of plant populations, that genetic diversity increases biological diversity, and that biological diversity creates new opportunities for interaction diversity.

Genetic Basis of Hybrid Plant Resistance

The influence of hybrid plants on biological diversity depends on the resistance of hybrids to phytophages, that is, organisms that feed on plants. Resistance is the ability of a plant genotype to avoid attack or prevent the development of a specific herbivore or pathogen relative to other genotypes. Resistance is a genetic trait that determines the abundance of each species of phytophagous organism on a plant; consequently, the resistance of hybrid plants, relative to their parents, to phytophages contributes to the biodiversity, community structure, and interactions among phytophages.

The resistance of hybrid plants compared with parental species may be a polygenic trait, influenced by many genes, or a single gene, Mendelian trait; consequently, a number of genetic models have been proposed to describe and predict patterns of resistance. These models also vary in the type of hybrids under consideration, as the genetic properties differ depending on whether the hybrid is a homoploid F_1, an F_2, backcross, or later-generation hybrid, or an allopolyploid. Patterns of susceptibility and resistance of hybrids to phytophages can therefore be complex. In some cases, genetic recombination may lead to higher susceptibility relative to the parents and hybrid breakdown, as coadapted gene complexes (perhaps composed of several resistance genes) from either or both parents are broken down. In other cases, the hybrid may combine the best of both parental species' defense mechanisms and be much more resistant; this outcome is particularly likely when hybridization is coupled with polyploidization, and the F_1 additive genotype is essentially stabilized. Furthermore, hybrids may be differentially resistant to generalist versus specialist herbivores. Finally, although the fundamental resistance or susceptibility of a plant to an herbivore or pathogen is genetically determined, the environment may also contribute to whether or not a certain genotype is attacked. The role of environmental variation in determining plant resistance is particularly important because some hybrid zones occur across steep environmental gradients, hybrids may occur at the range limits of each parent

species, and some hybrids are found in unique habitats where abiotic conditions differ. Plant stress, induced by environmental extremes, is known to make plants more susceptible to herbivores; thus, elevated levels of herbivory on hybrid plants in a given setting may be a direct result of stress rather than genetic composition (Whitham *et al.*, 1994; Fritz, 1999).

Communities on Plant Hybrids

Species richness of herbivores and pathogens on hybrids, in fact, varies, as predicted by the diversity of underlying genetic mechanisms of resistance, the type of hybrid, the environment, and the distribution of phytophages on the parental species (**Figure 3**). Species diversity will be greater on hybrids than on either parent if the parent species have species-specific herbivores and hybrids are not completely resistant to them – hybrids will have herbivore communities derived from both parent species. The effects of hybridization on species richness will be small if the parental species share a large proportion of their herbivores, as is the case for hybrids between the willows *Salix sericea* and *Salix eriocephala*. Most of the gall formers, leaf miners, and leaf tiers occur on both plant species and hybrids. In contrast, several herbivores occur on one species and hybrids but not on the other parent. For example, *Pontania gracilis* occurs on *S. sericea* and hybrids and *Pontania pomum* occurs on *S. eriocephala* and hybrids. Likewise, the stem gallers *Rabdophaga rigidae* and *Rabdophaga salicisbrassicoides* occurs on *S. sericea*, but *Rabdophaga strobiloides* occurs on *S. eriocephala*. These herbivores are responsible for the higher diversity on hybrid willows compared with the two parents. Furthermore, a significant number of species occurring on hybrids may be unique in that they are not found on either parent species. On spruce hybrids in two locations in Michigan, 12.2% and 15% of herbivore species were unique to hybrids. About half of species are restricted to one or the other parent species, and 35% and 41.5% of herbivores on hybrids were found on both parent species. This result may reflect the underlying genetic novelty of the hybrids relative to their parents.

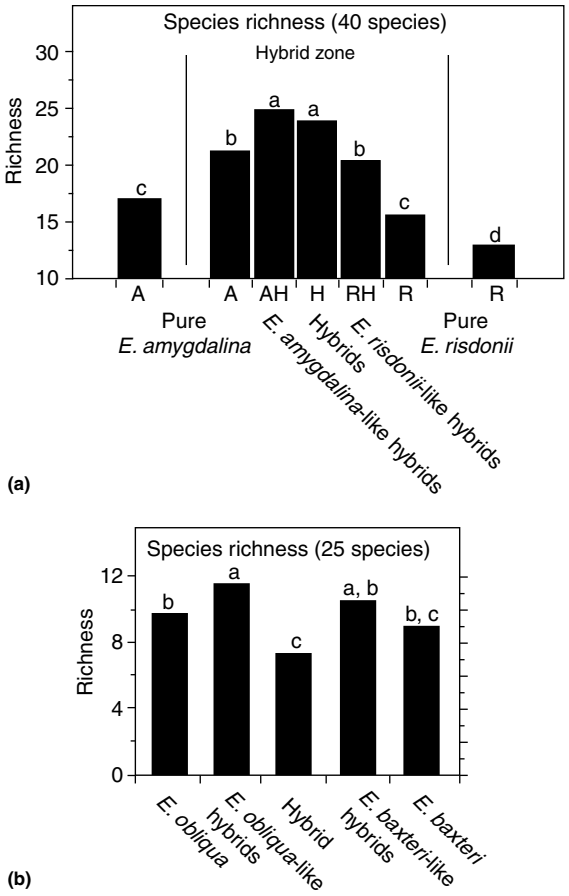


Figure 3 Species richness in a *Eucalyptus amygdalina – Eucalyptus risdonii* hybrid zone (40 insect and fungal species). (a) Species richness in a *Eucalyptus oblique–Eucalyptus baxteri* hybrid zone (25 insect and fungal species). Reproduced from Whitham TG, Morrow PA, and Potts BM (1994) Plant hybrid zones as centers of biodiversity: The herbivore community of two endemic Tasmanian eucalypts. *Oecologia* 97: 481–490. (b) Bars with different letters differ significantly at $p < .05$. Reproduced from Morrow PA, Whitham TG, Potts BM, Ladiges P, Ashton DH, and Williams JB (1994) Gall-forming insects concentrate on hybrid phenotypes of *Eucalyptus*. In: Price P, Mattson W, and Baranchikov Y (eds.) *The Ecology and Evolution of Gall-Forming Insects*, pp. 121–134. St. Paul: North Central Forest Experiment Station, Forest Service, USDA.

The combination of altered resistance of hybrid plants compared with parents, species-specific responses of herbivores to hybrid plants, and new combinations of coexisting phytophages on hybrid plants dictates that community shape (the relative representation of different species) will differ between hybrids and parent species and among genetic classes of hybrids. The ecological consequences of the novel communities of phytophages will depend on species abundances and how species interact with each other, and novel interactions (competition, mutualism, predation, and parasitism) are likely to result.

Species that have large effects on populations of other species in a community are called keystone species. Removing them from the community and measuring the responses of other species can reveal their indirect effects on community structure. Two such cases of indirect effects of keystone species on diversity and relative abundance of other species on hybrids were found for the gall aphid, *Pemphigus betae*, and the leaf beetle, *Chrysomela confluens*, in the *Populus angustifolia–Populus fremontii* (cottonwood) hybrid zone (Whitham *et al.*, 1999). Removal of *P. betae* from hybrid trees reduced species richness by 32% and relative abundance by 55% of other herbivores in the system. The negative effects on diversity were due to elimination of the parasites and predators of the aphids and reduction of other herbivores, which benefit from aphid modification of the host plant. In contrast, removal of the beetle, which preferentially feeds on leaves of young plants, increased species richness of other herbivores by 120% and relative abundance by 75%. Defoliation by the beetle reduces the available foliage for several other herbivores, an intense competitive effect. Thus, when plant hybridization strongly alters resistance to even one herbivore, there may be a cascade of indirect effects on other species.

Parasitoids, predators, parasites, and pathogens of herbivores or plant pathogens typically outnumber herbivores by several fold, and therefore their diversity is an important part

of the biological diversity supported by hybrid plants. Few studies have considered the effects of parasitoids or predators on herbivores occurring on hybrid and parental plants, but variation in abundance and fitness of herbivores on hybrid plants may be explained in part by the impact of this trophic level. Decreased levels of parasitism or predation could explain the higher abundance of some herbivores on hybrid plants compared with parental species. Higher parasitism of herbivores on hybrids may reduce the density and fitness impact of herbivores on hybrids. Parasitism by chalcid parasitoids did not differ among galls of *Andricus californicus* on *Quercus dumosa* and a wide range of hybrid genotypes. Mortality of aphids on hybrids of *Alnus incana* and *Alnus glutinosa* was intermediate between that of the two parents. However, parasitism rates of a seed-eating herbivore of two cattails (*Typha latifolia* and *Typha angustifolia*) and their F_1 hybrid were significantly greater on the hybrid, although densities of the seed predator were significantly lower on the hybrids. In this case, hybridization seems to directly affect herbivore susceptibility to parasites. Parasitism of the leaf miner *Phyllonorycter* sp. on oak hybrids was lower than on parental species, and lower parasitism of *Phyllonorycter salicifoliella* on willow hybrids was found in the field, but was not repeated the next year or on F_1 hybrid plants in a common garden experiment (Fritz *et al.*, 1999). Predation on beetle larvae on willow hybrids was intermediate between that on the parental species. There is a wide range of effects of parasitoids and predators on herbivores of hybrid and parental plant species. The outcome of these interactions is likely to influence the population dynamics of herbivores of hybrids and therefore their effects on hybrid plant fitness. One consequence of unique communities on hybrid plants is that related herbivores that share natural enemies may interact more intensely. Apparent competition, where differential parasitism causes negative density relationships between species, would seem to be more likely on hybrid plants. Coexistence of related species on hybrids might facilitate shifts in host utilization by parasitoids.

The interactions of plants with other species extend beyond herbivores and their natural enemies. Hybridization often results in novel plant architecture. F_1-type hybrids in the cottonwood hybrid zone between *P. fremontii* and *P. angustifolia* have greater numbers of lateral branches than parent species. Compared with parent species and backcrosses to *P. angustifolia*, these hybrids supported significantly greater numbers of bird nests. Northern orioles (*Icterus galbula*) and blackbilled magpies (*Pica pica*) were the most common species using hybrid cottonwoods as nest sites, but other common species using these trees were the American robin (*Turdus migratorius*) and the warbling vireo (*Vireo gilvus*) (Whitham *et al.*, 1999). Plants have mutualist or commensalistic interactions with endophytic fungi, which may have detrimental effects on insect and mammalian herbivores. In the *Quercus grisea* and *Quercus gambelii* hybrid zone, total endophyte frequency and frequency of *Gnomonia cerastis* was highest on *Q. gambelii* and was positively associated with hosts resembling *Q. gambelii*, but the frequency of *Coccochorella quercifolia* was most frequent on *Q. grisea* and on hybrids resembling *Q. grisea*. Although direct effects on the density of a leaf-mining moth (*Phyllonorycter* sp.) were not found, mortality of the tissue-feeding stage of *Phyllonorycter* sp. was positively associated with the presence of *G. cerastis*. This study suggests the possible interaction between endophytes and an insect herbivore mediated by hybrid genotype. Mycorrhizal fungi are ubiquitous mutualists of plants that commonly facilitate nutrient uptake. How hybridization affects these mutualists is unknown, but the effects of high levels of herbivory that occur on some hybrid plants could be detrimental to mycorrhizae. Herbivory on pinyon pine, *Pinus edulis*, by a stemboring moth reduced mycorrhizal levels by as much as 33%, but when moths were removed, mycorrhizae recovered. Thus, interactions between herbivory on hybrids and root mutualists may be antagonistic, resulting in more harmful effects on plant fitness. A rich diversity of ecological interactions (interaction diversity) is likely to occur on hybrid plants. Some interactions on hybrids will vary on themes present on parental species; other interactions will be unique due to unique species combinations on hybrids (e.g., mutualists and competitors). Still other interactions will be unique because novel or transgressive (traits that lie outside the parental range) traits of the hybrids themselves (e.g., morphology, phenology, or chemistry) create novel conditions for evolution among herbivores, pathogens, and their natural enemies.

Conclusions

The last few years have witnessed a resurgence of interest in hybridization and polyploidy, perhaps prompted by the ability to ask both old and new questions with genomic data. Ancient polyploidy seems rampant across plant lineages (except possibly for gymnosperms, but see Burleigh *et al.* (2012)). The question is no longer: is species x diploid or polyploid, but how many rounds of genome duplication are represented in the genome of species x? Given multiple polyploid genomes, how do the many duplicate genes interact to generate genetic and phenotypic novelty? Do consistent patterns emerge, suggesting that a set of principles may control the evolution of polyploid genomes? Are plants burdened by too many genomes and too much DNA (the genetic obesity hypothesis), and what are the mechanisms and pace of diploidization?

The questions raised here are specifically aimed at solving aspects of plant genetics and evolution, but ancient and recent polyploidy characterizes all lineages of eukaryotes, including vertebrates. Although polyploidy is much more common in plants, strong evidence supports multiple rounds of ancient WGD in the ancestry of vertebrates. Therefore, whatever we learn from polyploid plants may be useful for understanding the processes governing our own genomes, and those of many animals and fungi.

Finally, hybridization and polyploidization are sources of genetic (and phenotypic) novelty that may have long-term ecological and evolutionary consequences. Hybridization – including hybrid swarms, introgressants, and allopolyploids – may confer novel conditions that affect not only plant biodiversity but also the diversity of many communities that rely on plant communities. As drivers and shapers of biodiversity, these processes are fundamental for understanding the roles of plants on Earth.

References

Anderson E (1949) *Introgressive Hybridization*. New York: Wiley.

Anderson E and Hubricht L (1938) Hybridization in *Tradescantia*. III. The evidence for introgressive hybridization. *American Journal of Botany* 25: 396–402.

Anderson E and Stebbins GL (1954) Hybridization as an evolutionary stimulus. *Evolution* 8: 378–388.

Arnold ML (1993) *Iris nelsonii*: Origin and genetic composition of a homoploid hybrid species. *American Journal of Botany* 80: 577–583.

Arnold ML (1997) *Natural Hybridization and Evolution*. New York: Oxford University Press.

Arnold ML (2004) Transfer and origin of adaptations through natural hybridization: Were Anderson and Stebbins right? *The Plant Cell* 16: 562–570.

Bayer RJ, Purdy BG, and Lebedyx DG (1991) Niche differentiation among sexual species of *Antennaria gaertner* (Asteraceae: Inuleae) and *A. rosea*, their allopolyploid derivative. *Evolutionary Trends in Plants* 5: 109–123.

Brunsfeld SJ, Soltis DE, and Soltis PS (1992) Evolutionary patterns and processes in *Salix* section *Longifoliae*: Evidence from chloroplast DNA. *Systematic Botany* 17: 239–256.

Buggs RJA, Chamala S, Wu W, *et al*. (2012) Rapid, repeated, and clustered loss of duplicated genes in allopolyploid *Tragopogon* populations of independent origin. *Current Biology* 22: 248–252.

Burleigh JG, Barbazuk WB, Davis JM, Morse AM, and Soltis PS (2012) Exploring diversification and genome size evolution in gymnosperms through phylogenetic synthesis. *Journal of Botany*. Article 292857.

Chester MJP, Gallagher JVV, Symonds W, *et al*. (2012) Extensive chromosomal variation generated in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae). *Proceedings of the National Academy of Sciences, USA* 109: 1176–1181.http://dx.doi.org/10.1073/pnas.1112041109.

Cui L, Wall PK, Leebens-Mack JH, *et al*. (2006) Widespread genome duplications throughout the history of flowering plants. *Genome Research* 16: 738–739.

Davies TJ, Barraclough TG, Chase MW, Soltis PS, Soltis DE, and Savolainen V (2004) Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences, USA* 101: 1904–1909.

Fritz RS (1999) Resistance of hybrid plants to herbivores: Genes, environment, or both? *Ecology* 80: 382–391.

Fritz RS, Moulia C, and Newcombe G (1999) Resistance of hybrid plants and animals to herbivores, pathogens, and parasites. *Annual Review of Ecology and Systematics* 30: 365–391.

Gallez GP and Gottlieb LD (1982) Genetic evidence for the hybrid origin of the diploid plant *Stephanomeria diegensis*. *Evolution* 36: 1158–1167.

Goldblatt P (1980) Polyploidy in angiosperms: Monocotyledons. In: Lewis WH (ed.) *Polyploidy: Biological Relevance*, pp. 219–239. New York: Plenum Press.

Golding GB and Strobeck C (1983) Increased number of alleles found in hybrid populations due to intragenic recombination. *Evolution* 37: 17–29.

Grant V (1981) *Plant Speciation*, 2nd edn. New York: Columbia University Press.

Gross BL and Rieseberg LH (2005) The ecological genetics of homoploid hybrid speciation. *Journal of Heredity* 96: 241–252.

Haufler CH and Soltis DE (1986) Genetic evidence indicates that homosporous ferns with high chromosome numbers may be diploid. *Proceedings of the National Academy of Sciences, USA* 83: 4389–4393.

Heiser CB (1949) Studies in the evolution of the sunflower species *Helianthus annuus* and *H. bolanderi*. *University of California Publications in Botany* 23: 157–196.

Jiao Y, Wickett NJ, Ayyampalayam S, *et al*. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.

Klekowski EJ and Baker HG (1966) Evolutionary significance of polyploidy in the pteridophyta. *Science* 135: 305–307.

Levin DA (2002) *The Role of Chromosomal Change in Plant Evolution*. New York: Oxford University Press.

Levin DA, Francisco-Ortega J, and Jansen RK (1996) Hybridization and the extinction of rare plant species. *Conservation Biology* 10: 10–16.

Masterson J (1994) Stomatal size in fossil plants: Evidence for polyploidy in majority of angiosperms. *Science* 264: 421–423.

Mayr E (1942) *Systematics and the Origin of Species*. New York: Columbia University Press.

Mayrose I, Zhan SH, Rothfels CJ, *et al*. (2011) Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257.

McDade L (1990) Hybrids and phylogenetic systematics I. Patterns of character expression in hybrids and their implications for cladistic analysis. *Evolution* 44: 1685–1700.

Morrow PA, Whitham TG, Potts BM, Ladiges P, Ashton DH, and Williams JB (1994) Gall-forming insects concentrate on hybrid phenotypes of *Eucalyptus*. In: Price P, Mattson W, and Baranchikov Y (eds.) *The Ecology and Evolution of Gall-Forming Insects*, pp. 121–134. St. Paul: North Central Forest Experiment Station, Forest Service, USDA.

Otto SP and Whitton J (2000) Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437.

Ownbey M (1950) Natural hybridization and amphiploidy in the genus *Tragopogon*. *American Journal of Botany* 37: 487–499.

Palmer JD (1992) Mitochondrial DNA in plant systematics: Applications and limitations. In: Soltis D, Soltis P, and Doyle JJ (eds.) *Molecular Systematics of Plants*, pp. 36–49. New York: Chapman and Hall.

Palmer JD, Jorgensen RA, and Thompson WF (1985) Chloroplast DNA variation and evolution in *Pisum*: Patterns of change and phylogenetic analysis. *Genetics* 109: 195–213.

Ramsey J and Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467–501.

Ramsey J and Schemske DW (2002) Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* 33: 589–639.

Rieseberg LH (1991) Hybridization in rare plants: Insights from case studies in *Cercocarpus* and *Helianthus*. In: Falk DA and Holsinger KE (eds.) *Genetics and Conservation of Rare Plants*, pp. 171–181. New York: Oxford University Press.

Rieseberg LH (1995) The role of hybridization in evolution: Old wine in new skins. *American Journal of Botany* 82: 944–953.

Rieseberg LH (2006) Hybrid speciation in wild sunflowers. *Annals of the Missouri Botanical Garden* 93: 34–48.

Rieseberg LH, Kim S-C, Randell RA, Whitney KD, Gross BL, *et al*. (2007) Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica* 129: 149–165.

Rieseberg LH and Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* 5: 65–84.

Rieseberg LH, Van Fossen C, and Desrochers A (1995) Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* 375: 313–316.

Rieseberg LH and Willis JH (2007) Plant speciation. *Science* 317: 910–914.

Rieseberg LH, Wood TE, and Baack EJ (2006) The nature of plant species. *Nature* 440: 524–527.

Riley HP (1938) A character analysis of colonies *of Iris fulva*, *Iris hexagona* var. *giganticaerulea* and natural hybrids. *American Journal of Botany* 25: 727–738.

Soltis DE (1986) Genetic evidence for diploidy in *Equisetum*. *American Journal of Botany* 73: 908–913.

Soltis DE, Albert VAA, Leebens-Mack J, *et al*. (2009) Polyploidy and angiosperm diversification. *American Journal of Botany* 96: 336–348.

Soltis DE and Soltis PS (1988) Are lycopods with high chromosome numbers ancient polyploids? *American Journal of Botany* 75: 238–247.

Soltis DE and Soltis PS (1990) Isozyme evidence for ancient polyploidy in primitive angiosperms. *Systematic Botany* 15: 328–337.

Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate J, and Mavrodiev E (2004) Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): Cytogenetic, genomic and genetic comparisons. *Biological Journal of the Linnean Society* 82: 485–501.

Soltis PS and Soltis DE (1988) Electrophoretic evidence for genetic diploidy in *Psilotum nudum*. *American Journal of Botany* 75: 1667–1671.

Soltis PS and Soltis DE (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60: 561–588.

Stebbins GL (1947) Types of polyploids: Their classification and significance. *Advances in Genetics* 1: 403–429.

Stebbins GL (1950) *Variation and Evolution in Plants*. New York: Columbia University Press.

Stebbins GL (1971) *Chromosomal Evolution in Higher Plants*. London: Addison-Wesley.

Stebbins GL (1985) Polyploidy, hybridization, and the invasion of new habitats. *Annals Missouri Botanical Garden* 72: 824–832.

Tate JA, Symonds VV, Doust AN, *et al*. (2009) Synthetic polyploids of *Tragopogon miscellus* and *T. mirus* (Asteraceae): 60 Years after Ownbey's discovery. *American Journal of Botany* 96: 979–988.

Tuskan GA, DiFazio S, Jansson S, *et al*. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.

Wagner Jr. WH (1970) Biosystematics and evolutionary noise. *Taxon* 19: 146.

Whitham TG, Morrow PA, and Potts BM (1994) Plant hybrid zones as centers of biodiversity: The herbivore community of two endemic Tasmanian eucalypts. *Oecologia* 97: 481–490.

Whitham TG, Martinsen GD, Floate KD, Dungey HS, Potts BM, and Keim P (1999) Plant hybrid zones affect biodiversity: Tools for a genetic-based understanding of community structure. *Ecology* 80: 416–428.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, and Rieseberg LH (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academies of Science, USA* 106: 13875–13879.

Hymenoptera

Norman F Johnson, The Ohio State University, Columbus, OH, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Arrhenotoky Reproductive mode in which unfertilized eggs develop into haploid males and fertilized eggs develop into diploid females.
Eusociality Cooperative behavior among individuals of the same species characterized by reproductive division of labor, overlap of generations, and cooperative nesting.
Holometabola Insects characterized by complete metamorphosis, a wingless larval stage, and an intermediate pupal stage.
Idiobiont A parasitoid that develops on a paralyzed, incapacitated host.
Koinobiont A parasitoid that develops on a mobile, active host.
Monophyletic A group in which all species are descended from a single common ancestor and all descendants of the ancestor are classified in the group; characterized by shared derived characters.
Ovipositor Modified appendages of the seventh and eighth abdominal segments used for egg laying.
Paraphyletic A group in which only some of the species descended from an ancestor are classified together; characterized by shared ancestral characters.
Parasitoid An organism in which the immature stage feeds and develops on a single host arthropod, resulting in the death of the host.
Parthenogenesis Reproduction in which eggs are not fertilized by males.
Phylogeny Branching pattern of evolutionary relationships among organisms.
Phytophagy Plant feeding, herbivory.
Thelytoky Reproductive mode in which unfertilized eggs develop into diploid females.

The insect order Hymenoptera comprises a vast array of species that are familiar to even the most casual observer. The group includes the ants, bees, and wasps as well as less well-known groups such as the chalcids, ichneumons, sawflies, and the wood wasps. Hymenoptera are extremely common on all the continents of the world, except Antarctica. Hymenoptera range in size from microscopic species less than 1 mm in length as full-grown adults to species in which the females are 10 cm or more in total body size. Hymenoptera are holometabolous insects, generally characterized by mandibulate mouthparts, complete metamorphosis, and two pairs of membranous wings. One striking feature is that the vast majority of species are characterized by arrhenotokous parthenogenesis: Fertilized eggs develop into diploid females and unfertilized eggs ultimately develop into haploid males. Adult Hymenoptera generally feed on nectar, honeydew, or other sugar sources, but some are actively predaceous. The larvae have a wider variety of feeding habits: Some feed on plants or fungi, whereas others feed on animals, as predators or parasitoids. The larvae of bees and a few other groups feed on provisions of pollen collected and stored by the parent female. The egg-laying appendages, the ovipositor, are often specialized to place the eggs on or into the plants or arthropods that will serve as the food source for the developing larvae. In one large group, the Aculeata, the ovipositor is modified into the sting, a structure that primarily serves to deliver toxic venoms to paralyze the prey or is used in defense. The Hymenoptera are also noteworthy for the evolution of social behavior in several groups. The most highly developed form, eusociality, has arisen many times within the order among the bees, wasps, and ants.

Classification

Hymenoptera has been traditionally divided into two suborders, the Symphyta and the Apocrita. The Symphyta comprise the sawflies and wood wasps (Gauld and Bolton, 1988; Goulet and Huber, 1993; Hanson and Gauld, 1995; Naumann, 1991). The larvae of the vast majority of species in this suborder consume plant or fungal material, feeding on leaves or stems, feeding within galls produced in the host plant, or boring within stems or trunks of woody plants. Symbiotic fungi comprise a significant part of the diet of many of the wood-boring species. Sawfly larvae that feed externally on leaves are generally eruciform, i.e., caterpillar-like in form, with well-developed thoracic legs, abdominal prolegs, and sclerotized heads.

The suborder Apocrita descends from a symphytan ancestral species. As adults, the first abdominal segment of the Apocrita is intimately associated with the thorax and separated from the following segments by a mobile constriction. This segment is called the propodeum. Thus, the locomotory tagma of the body, bearing the legs and wings, is four segmented. The larvae are generally immobile and highly simplified in structure. Apocrita are generally characterized by larvae that feed on other arthropods, including predatory species in which one larva feeds on several prey items and parasitoids in which the larva feeds on a single host individual. Parasitoids are distinguished from parasites in that the host of the former is usually killed as a result of feeding. Thus, ecologically, parasitoids function as predators, feeding on and ultimately killing their host. The parasitoid life history is also found in the closest living relative of Apocrita within the Symphyta – the

family Orussidae. However, numerous apocritan groups are secondarily phytophagous, consuming seeds or pollen or forming galls in plant tissue within which the larvae feed and develop.

The division of Hymenoptera into two suborders is still generally used, even though recent evidence indicates that the Symphyta are paraphyletic. In the formal taxonomic classification, approximately 22 superfamilies and 89 families are recognized in the extant fauna (Sharkey, 2007). The Tenthredinoidea, Pamphilioidea, Cephoidea, and Xyeloidea comprise the phytophagous sawflies. Wood wasps comprise families within the Siricoidea and Xiphydrioidea. In the Apocrita there are many small, relatively primitive superfamilies of parasitoids: Trigonalyoidea, Ceraphronoidea, Evanioidea, Megalyroidea, and Stephanoidea. The superfamilies Chalcidoidea, Cynipodea, Ichneumonoidea, Proctotrupoidea, Diaprioidea, and Platygastroidea encompass the vast majority of parasitoid species. Many of these are common and extremely abundant. The family Ichneumonidae is estimated to contain more than 22,000 known species, with many others yet to be recognized. The remaining three superfamilies – the Chrysidoidea, Vespoidea, and Apoidea – make up the aculeate, or stinging Hymenoptera. The Chrysidoidea include many relatively small families of parasitoids. The Vespoidea contains the ants, velvet ants, hornets, spider wasps, paper wasps, and potter wasps (Hölldobler and Wilson, 1990). The Apoidea are the bees and predatory wasps (Michener, 2007).

The order Hymenoptera is estimated to contain approximately 115,000 described species (Gauld and Bolton, 1988). This is only a rough indicator of the actual total because hundreds of new species are described each year in the taxonomic literature. The largest families are the Ichneumonidae (more than 20,000 species), Braconidae (approximately 12,000 species), Formicidae (the ants, with more than 12,000 species). Austral disjunct distributions are characteristic of a relatively few families of Hymenoptera, such as Megalyridae and Plumariidae. The major groups are found in all biogeographic realms or perhaps limited to the tropics, such as the fig wasps or Agaonidae. Ichneumonids and sawflies appear to be the most diverse in the temperate realms.

Phylogeny and Fossil Record

The oldest fossil Hymenoptera are found in deposits from the Triassic period of the Mesozoic. These include some species that are closely related to living sawflies. Parasitoid wasps, as inferred from their structure (especially the elongate ovipositor in some species) and their relationships to living species, first appear in the Jurassic. The oldest fossils of social insects, ants and bees, have been found in Cretaceous impressions and amber nodules. Minute parasitic Hymenoptera and ants and bees are commonly found in Cenozoic amber from the Baltic, the Dominican Republic, and Mexico.

Although Hymenoptera are clearly members of the monophyletic group Holometabola, comprising insects with complete metamorphosis, their position within that huge complex is unclear. Most workers place them as a basal group, sometimes as the sister group to all other holometabolous orders. Within the Hymenoptera, the basal position of Symphyta is supported by evidence from morphology, behavior, molecular sequences, and the fossil record. The historical grouping of parasitoid species into a taxon variously called Parasitica or Terebrantes, although of some practical use, does not reflect the pattern of phylogenetic relationships among them. The relationships among the basal Apocrita are still unclear. Many of these species are parasitoids of wood-boring Coleoptera, a life history shared with their closest living relatives in the Orussidae. Numerous hypotheses of relationships and classifications have been offered through the years that have been based on varying amounts of evidence from several sources and analyzed more or less rigorously, but none has yet achieved anything approaching general acceptance. The Aculeata are a well-supported monophyletic group, and a fair amount of progress has been made in elucidating the relationships among its constituents.

Biology

General

Reproduction

Arrhenotokous parthenogenesis, or haplodiploid reproduction, is characteristic of the vast majority of species of Hymenoptera and is generally thought to have been very important in its evolution. In this mode of reproduction, the offsprings are produced from both the fertilized and unfertilized eggs. The eggs that are not fertilized develop into haploid males. The fertilized eggs develop into diploid embryos, and these are usually females. Because of the mechanism of sex determination in Hymenoptera, which is not completely understood, diploid males are rarely produced. Mated adult females store spermatozoa received from the males during mating in a diverticulum off of the reproductive tract, the spermatheca. The female is capable of controlling the release of spermatozoa to fertilize eggs passing through the oviduct, thus controlling the sex of her offspring. The control of the sex ratio of offspring may be advantageously applied, in theory, to maximize the reproductive success of the female. In some parasitoid species, the size of the host can influence the female to lay a male or female egg. In other circumstances, the presence or past presence of other ovipositing females on a host has been shown to influence the relative proportions of male and female eggs laid. Haplodiploidy also affects the proportions of alleles that are shared among the relatives in comparison to the expectations of relatedness that are commonly found among strictly diploid species. In the simplest case of a female mating with one male, a daughter may be generally expected to share a greater proportion of its alleles with her sister (75%) than with her brother (25%). The two female offsprings share one set of chromosomes contributed by the haploid male spermatozoa and, on average, would expect to share 50% of the alleles inherited from the mother. In contrast, the male offspring has only a single set of chromosomes derived from his mother. This asymmetry in relatedness has been thought to be a contributing factor to the multiple evolution of eusociality in the Hymenoptera, but its importance is strongly questioned. Although single matings may be the rule for many species under normal conditions,

one female mating with many males is not uncommon. The mating flight of the honeybee, *Apis mellifera*, is an extreme example. Also, many species of social insects, especially among the ants, have numerous unrelated egg-laying queens within a colony.

Another form of parthenogenesis is relatively common within Hymenoptera: thelytoky. In this case, the female does not mate with a male, and her unfertilized eggs are capable of embryonic development and produce only female offspring. Thelytoky is not uncommon among parasitoid groups, and the funneling of all reproductive output into the production of new egg-laying daughters maximizes short-term reproductive success. It has recently been discovered that in several cases thelytokous parthenogenesis is caused by bacterial infections, principally by the genus *Wolbachia*. Treatment of the wasps with heat or antibiotics can cause the individuals to revert to normal arrhenotokous modes of reproduction. Many strains of *Wolbachia* do not result in thelytokous reproduction but rather in cytoplasmic incompatibility between eggs and spermatozoa. Infections by these bacteria have been discovered in a wide range of insects and evolutionary and ecological significance of the phenomenon is not well understood.

Oviposition Behavior

Hymenoptera are probably most familiar as stinging insects. The sting apparatus is a modification of the egg-laying appendages fundamental to all Hymenoptera. Because egg-laying structures are naturally associated with females, only female wasps can sting. However, the males of some species are extraordinarily adept at mimicking both the structure and behavior of females and can sometimes fool even experienced workers.

In its ancestral condition, the ovipositor is composed of two pairs of elongated appendages, the gonapophyses or ovipositor valves, associated with the seventh and eighth abdominal segments. The gonapophyses are physically linked together and function as a single articulating unit. When not in use, the ovipositor is usually enclosed within a protective pair of sheaths, the gonoplacs. In sawflies, the gonapophyses are laterally flattened, serrate, and used to cut an incision into plant material into which one or more eggs are laid. The similarity in structure and function to a saw gives rise to the common name for most of the Symphyta.

Within the Hymenoptera, the ovipositor and its function have been extended beyond its basic function in egg laying. Among parasitoids, it is used to paralyze hosts; in the Aculeata or stinging Hymenoptera, the ovipositor and venom are often used as a defensive weapon.

Parasitoids

A parasitoid is defined as an animal in which the immature stage feeds on a single host individual, and this feeding activity normally results in the death of the host (Quicke, 1997). Historically, such insects have been called parasites, but this was misleading. True parasites, such as lice, feed on a large host but the feeding activity has only a minor deleterious effect. Despite this distinction in names for the two types of life history, the verb parasitize is still normally used to describe the action of both groups. A parasitoid is essentially a predator, but one that feeds on and kills only a single prey item. They also differ from predators in that it is the parent that finds the prey and the offspring that feed on it. The Hymenoptera are the primary group of insect parasitoids, but other orders, particularly the Diptera and Coleoptera, have species that follow this same pattern.

Parasitoid Hymenoptera attack a very broad range of arthropods. Practically all other insect groups, as well as many arachnids and myriapods, have their suite of parasitoids. Most parasitoids attack the immature stages of their host: the egg, nymph, larva, or pupa. A few groups parasitize adult insects. In some cases, the female parasitoid oviposits on one stage and her adult offspring emerge from another. For example, some species of Braconidae oviposit within the egg of their host, but the wasp larva actually feeds on and emerges from the host larva. After the completion of feeding, the parasitoid pupates on, within, or sometimes near the spent carcass of its host.

The ovipositor is used to first sting the host arthropod, and venom is then injected. The venom may incapacitate the host for only a short period of time or, in more extreme cases, may result in the complete suspension of most normal activities, putting the host in a state of suspended animation. The female wasp then lays one or more eggs on or often within the body of the host. In some species, the female parasitoid attacks exposed hosts – for example, a caterpillar feeding on a leaf. In other cases, though, the hosts are hidden in the soil, enclosed within shelters made of leaves, twigs, debris, or silk, or even bore deep within the trunks of trees. The structural characteristics, sensory capabilities, and behavioral repertoire of the female wasp combine to enable her to locate and successfully oviposit on such protected, cryptic hosts. For example, many species of parasitoid Hymenoptera that attack wood-boring beetles are equipped with elongate ovipositors that may be two or more times the length of the rest of the body. The females are capable of drilling through tens of centimeters of wood with their ovipositor in order to find and parasitize the hidden beetle larva.

The proximate mechanisms by which a female parasitoid locates, identifies, and determines the suitability of hosts within a complex environment are only beginning to be understood. Chemical cues are certainly important in all of these stages of host finding. These chemicals may be produced by the hosts as a by-product of normal activity such as feeding. Often, parasitoid females are attracted to volatile chemicals produced by the plant on which a host is feeding. Other sensory modalities used in finding hosts and assessing their quality include vision, tactile examination, detection of substrate vibrations, and possibly heat detection. The external surface of the body of an adult wasp is studded with an array of sensory structures. For host location and acceptance, the most important of these are found on the antennae, tarsi, mouthparts, and ovipositor.

Among the array of details of life-history strategies, two general patterns are often distinguished, and parasitoids may be classified as idiobionts or koinobionts. Idiobionts are generally characterized by the fact that the development of the host arthropod is arrested through the action of the venom

injected by the parent. The larval parasitoid thus develops on an incapacitated host. In koinobionts, the host resumes feeding and development after the parasitoid female has oviposited. The larval koinobiont parasitoids typically develop within the body of the host, but this may be delayed for some time as the host grows and even pupates. Thus, the parasitoid must have the physiological mechanisms to evade the immune system of its host. Idiobionts typically attack hosts that are found in concealed situations and a broad taxonomic range of hosts may be attacked. Koinobionts, in contrast, may parasitize only exposed hosts and, because they must be attuned much more closely to the host immune and endocrine systems, their host range is typically much narrower.

There is a tremendous range of variation found within the general life-history patterns described. Parasitoid larvae may feed on the host from outside the host's body as ectoparasitoids, or they may live and feed within the body of the host as endoparasitoids. Among solitary parasitoids, a single individual wasp develops from each host individual. The adult that emerges from the pupa then must locate mates, food, water, and new hosts on which to oviposit. Many individual parasitoids develop on a single host in gregarious parasitism; the number of wasps may range from two to thousands. In these cases, males generally emerge first and mate with the later-emerging females on or near their pupation site. In some cases, this leads to a high level of sib-mating (i.e., brothers mating sisters), and in such situations it is not uncommon for the sex ratio to be significantly skewed toward the production of females. In other words, the parent female seems to lay just enough male eggs, sometimes only one, to ensure the insemination of all her daughters. In this manner, she is thought to be able to increase her overall reproductive output. Gregarious parasitism may be effected by numerous wasps attacking a single host, by a single wasp placing more than one egg on or in the host, or by a peculiar phenomenon called polyembryony. In this case, the female parent places one egg within the body of the host, but this single egg produces from two to thousands of identical embryos, each of which eventually develops into a new adult parasitoid. Polyembryony has apparently evolved at least four times within the Hymenoptera because it is found in the families Braconidae, Platygastridae, Dryinidae, and Encyrtidae.

Hyperparasitoids, or secondary parasitoids, are species that are actually parasitoids of other parasitoids. For example, in some species of the family Trigonalidae, the adult female lays large numbers of eggs of foliage near feeding caterpillars. The tiny eggs are ingested by the caterpillar, the thick shell of the egg is disrupted by the caterpillar's mandibles, and the trigonalid larva quickly hatches and bores through the gut of the host to enter the hemocoel. It then lies dormant until the caterpillar is subsequently parasitized by another parasitoid species, typically an ichneumonoid or a tachinid (Diptera). The new parasitoid develops on the caterpillar host but is then attacked and ultimately killed by the trigonalid. Success in such a complicated life history requires a sequence of individually unlikely events: ingestion of the egg; survival of the host caterpillar from disease, predation, starvation, etc.; and subsequent parasitization of the caterpillar by a suitable species. The low probability of such a sequence is compensated by the fact that the female trigonalid lays huge numbers of eggs.

One of the most bizarre types of life history is the heteronomous parasitism found among some species of the chalcidoid family Aphelinidae. Female offspring develop as primary endoparasitoids of sternorrhynchous Hemiptera (e.g., scale insects). The males, in contrast, may develop as ectoparasitoids of the same species of host or as secondary parasitoids, sometimes attacking females of their own species.

Although the parent female of most parasitic Hymenoptera finds and oviposits on the host for her offspring, in a few cases it is the larvae that locate the hosts. In the Perilampidae and Eucharitidae, the adult female wasp lays her eggs on or in plant material, often in large numbers. The first instar larvae that hatch from these eggs are strongly sclerotized and mobile, in contrast to the typical larva of Apocrita. These mobile first instars, called planidia, somehow find their way to their host, attach, and feed. In the eucharitids, it appears in some species that the planidia may attach to thrips, which are then picked up and carried by ants into their nest. Once there, the planidia disengage from their temporary transport and attach to and feed on ant larvae.

Predators

The distinction between hymenopteran larvae that act as parasitoids and those that act as predators is fairly arbitrary, being dependent only on the number of host individuals killed and eaten in the process of larval development. The typical predatory life history is extremely similar to the idiobiont strategy. The parent female wasp finds, stings, and immobilizes the prey item. The prey may already be in a place of concealment, such as beneath a rock or within a burrow, or the female wasp may actually construct a hiding place, often a burrow in the soil or tunnels in wood. More than one prey item may be cached in a chamber in this burrow, and then the female wasp lays an egg on the prey. The larva that hatches feeds on and develops on this store of food that has been provided by its mother. In the case in which only one prey item is available, this is essentially identical to the life history of an external parasitoid. In many species, the female parent may use the same burrow within which to store the prey for several of her larvae, constructing side chambers for each of the offspring or building chambers in a linear series. Thus, from the starting point of ovipositing on a paralyzed prey item within its own hiding place, there is the development of a nest – that is, a structure built by the female parent within which her young are fed and develop. Some wasps gather together all of the prey that an offspring will need for development before ovipositing. Others progressively provision their nests, providing new prey items to the developing larva as it consumes the food available.

Phytophagy

It is generally accepted that the Hymenoptera evolved from a phytophagous ancestor similar in many respects to the extant sawflies. The Apocrita, in turn, are derived from a parasitoid ancestor. However, within the Apocrita, there are numerous cases in which phytophagy, in its broadest sense, has reoccurred. In some cases, such as the seed chalcids of the family

Eurytomidae, the host relationship is fairly straightforward: The female wasp oviposits into the developing seed and the larva feeds on the endosperm. In other cases, their phytophagy is more elaborate or even bizarre.

Gall Makers

Galls are abnormal growths of plant tissue caused by some sort of stimulus from another plant or animal. A wide range of insects are known to cause gall development; within the Hymenoptera it is found among sawflies, the primarily parasitic Braconidae (Ichneumonoidea) and Chalcidoidea, and especially in the gall wasps, the family Cynipidae (Cynipoidea). Gall development is induced either by chemicals injected into the plant by the ovipositing female or by secretions produced by the newly hatched larvae. The plant tissue that is elaborated into the gall may take a variety of shapes and sizes, from mere swellings of the stem to large, ornate structures on the leaves or roots. The gall maker feeds on the plant tissue within the gall, pupates there, and eventually chews an emergence hole through which it escapes. The nutritive tissue, both that of the gall and sometimes that of the developing gall maker, is also made use of by a wide variety of inquiline species which oviposit into the gall and kill the original gall maker.

The life histories of cynipid species may vary greatly. Some are fairly typical of other Hymenoptera, that is, males and females emerge from their respective galls and then mate, and the females then seek out new host plants in which to oviposit. Their eggs develop into either females or males, depending on whether the eggs are fertilized or not. Such species are typically univoltine and attack a wide variety of host plants. Other life histories may be significantly more complicated. Some species have simply abandoned the production of males and are strictly thelytokous. Others alternate between sexual and asexual generations, reproducing via arrhenotokous and thelytokous parthenogenesis, respectively. The sexual generation consists of both males and females. These mate and the inseminated females seek hosts in which to oviposit. All of the offspring of these wasps develop into females. The adults of this new generation in turn oviposit, and some of the eggs develop into males and some into females, thus returning to the sexual generation. The adults of the sexual and asexual generation of gall wasps often strikingly differ in their structure and produce very different types of galls in different positions on the host plant. In some cases, they even attack different species of plants.

Fig Wasps

The fig wasps, of the chalcidoid subfamily Agaoninae, are very specialized forms of gall makers. Female fig wasps burrow their way into the syconium, the inflorescence of the fig plant. Pollen from the body of the female is transferred to the flowers concealed within the syconium and then the female oviposits within some of the flowers. These flowers subsequently swell, and the emerging larva feeds within the tissues, consuming both the embryo and endosperm. The male wasps then emerge first and mate with the females while they are still within the galled flowers of the fig. The females later emerge from the flowers, pick up pollen, sometimes actively storing it in special cavities on the body, and escape through holes chewed by the males. Male fig wasps are extraordinarily aberrant creatures: They lack wings, their eyes and antennae are underdeveloped, but in contrast the legs and mandibles are sometimes enormous. A rich complement of other parasitic Hymenoptera are associated with figs, some developing on fig tissues and others acting as parasitoids of the fig wasps.

Leaf-Cutting Ants

The leaf-cutters belong to the tribe Attini (Formicidae). These ants are found nearly throughout tropical America, north into the southern portions of the US. The long lines of thousands of ants, pruning pieces of vegetation and carrying them back to the nest are a familiar sight throughout the New World. Strictly speaking, the leaf-cutters are not phytophagous because they do not feed directly on the vegetation that they collect. Rather, they use the leaves and stems as the substrate on which they maintain a fungal colony. The ants feed on specialized structures produced by their fungus; the fungal organism, in return, is nurtured and maintained in an environment favorable for growth. When leaving the nest on her mating flight, the new virgin queen takes a bit of the fungal mass to use as an inoculant when she begins here new colony. Colonies of leaf-cutters can be extremely large, numbering in the millions. As a result, they can be extremely destructive to agriculture, forestry, and horticulture.

Pollen Feeders

The life history previously described for predatory stinging Hymenoptera has been modified in at least two cases in which the parent female provisions her nest with pollen rather than arthropod prey. The largest group in which this has occurred, and one of the groups most familiar to the casual observer, comprises the bees. Female bees visit flowers both to gather nectar for their own energetic needs and to gather the pollen on which their offspring will feed. Some bees are oligolectic, i.e., they are fairly restricted in the range of plant species from which they gather pollen. Others, such as the ubiquitous honeybee, are polylectic and gather pollen from a broad range of species. There are many specializations on the bodies of bees that facilitate the transport of pollen. Some bees will ingest the pollen grains and transport them within their crop; others have dense beds of hairs particularly on the legs or on the underside of the abdomen where pollen is packed for transport. Pollination of flowers by bees – that is, the transport of pollen from one flower to another – is often accidental and is irrelevant to the insect but of critical reproductive importance to the plant. Many anatomical features of plants are believed to be specializations for attracting bees or other insects to the nectar and ensuring that pollen grains adhere to their body.

Social Behavior

Social behavior is most simply defined as that of groups of individuals of the same species that cooperate with one another. Simple aggregations of individuals may occur, for example, where some limiting resource is found, such as water or nesting sites. Evolutionarily important social behavior, however, involves some sort of cooperation among individuals leading to reproductive success of some or all of the

participants. Among the insects, the extremes in development of social behavior primarily are found in the termites (order Isoptera) and within the Hymenoptera. Its most highly developed level, called eusociality, is characterized by cooperation among females in nesting, overlap in generations, and reproductive division of labor. This means that some individuals sacrifice production of their own offspring in order to facilitate reproduction by other individuals of the same colony. Eusociality has clearly evolved several times within the order Hymenoptera: once in the ants (Hölldobler and Wilson, 1990), which are primitively eusocial; once in the family Crabronidae; and several times among the bees and the social wasps (i.e., the paper wasps, yellowjackets, hornets, etc.).

Colonies of social insects can be extremely large, both physically and in terms of total numbers of individuals. The life cycle of a colony typically begins with the mating flight of a virgin queen. After mating with one to several males, the queen begins construction of a nest in the soil, a natural cavity, or in some cases in the open. Colony founding is sometimes cooperative (e.g., in some paper wasps), and the determination of which individual will become the primary reproductive one is established through behavioral interactions among the founderesses. Males do not participate in colony founding and only serve to inseminate the new queens, after which they soon perish. The eggs produced by the queen develop into the first worker generation: These are all daughters of the queen and typically do not reproduce. Workers forage for food and nest materials, care for the developing brood, defend and care for the nest, and care for the queen. This range of behaviors is sometimes divided among the workers. In some species, individuals are morphologically specialized into recognizable castes for certain functions such as defense in the case of soldiers. Morphological differentiation into castes results from allometric growth of the individual workers. Individuals of overall larger body size have disproportionately large mandibles, spines, and head capsules. In other cases, individuals may not be morphologically distinguishable but specialize in particular sets of behaviors at different periods of their adult life, usually ending as foragers outside the nest. Once well established, the collective efforts of the individuals of the colony result in the production of a new generation of males and reproductive females. Some social species reproduce by swarming: In the case of the honeybee, the old queen leaves the existing nest, taking with her a portion of the worker force, and reestablishes in a new site. Then one of the newly emerging queens takes over the remainder of the colony.

Communication among individuals is critical to social behavior. Chemical communication is probably of primary importance. Many exocrine glands have been identified on social Hymenoptera from all major parts of the body. These chemicals include alarm and defense pheromones, mating pheromones, trail pheromones, and signals from the queen that suppress the reproduction of workers. Trophallaxis, the exchange of food between individuals, may serve as the medium of exchange of some chemical signals. Tactile behavior is also important, principally involving antennal contact between individuals. The classic example of tactile communication is the waggle dance of honeybees. The location of food sources, including direction and to some extent distance, is transmitted to nestmates by a stylized reenactment within the darkness of the hive of the foraging flight. The direction of the resource in relation to the sun is indicated by the angle formed by the waggle portion of the dance in relation to the vertical axis within the hive.

Eusocial behavior is a fairly rare phenomenon in the animal world. The repeated evolution of such complex behavior within the single order Hymenoptera has led to a great deal of work in search of an explanation. One suggestion is that the asymmetric relationship among siblings and parents may be an important underlying feature that predisposes this group to the evolution of the characteristics of sociality. However, within the group at least two separate paths leading to eusociality have been identified. The parasocial route first envisions communal nesting among related females, followed by cooperative brood care and eventually reproductive division of labor. The subsocial route to eusociality posits first the development of overlap of generations followed by continued reproduction of the parent female in the company of her own daughters.

These characterizations of social behavior are often more simplified than the real situation found in nature. Colonies may be founded by more than one queen; sometimes these are closely related. Queens are probably usually inseminated by several or even many males, thus reducing the relevance of the asymmetry of relationships associated with arrhenotokous parthenogenesis. It is also clear that eusociality, although a highly evolved trait, is not the end point of evolution in Hymenoptera. Local environmental factors are undoubtedly important in the development of eusocial traits and in their subsequent loss. Several cases have been identified in which these behavioral repertoires and both the morphological and physiological features associated with them have been lost. Some of the most extreme cases are social parasitism in which reproductive females of one species live within the nest of another social species. The workers of the host species feed and care for the parasite, and the latter contributes nothing to the welfare of the colony. Social parasites are widely found among the ants, bees, and wasps.

Practical Import

Pests

In comparison to other large orders of insects, the Hymenoptera have relatively few species that qualify as pest species. A few of the phytophagous forms do cause some damage to agricultural crops or to trees. Leaf-cutting ants and leaf-cutter bees can cause damage by removing significant amounts of foliage from plants. In contrast to the ants, leaf-cutter bees do not feed on the leaves, either directly or indirectly, but use the pieces in order to line the brood chambers in their nests. Seed chalcids may destroy seeds of crops such as clover. Many of the phytophagous sawflies are important pests, particularly to forest trees. The family Diprionidae, which includes species such as the red-headed pine sayfly (*Neodiprion lecontei*), are often important pests of conifers. Other families such as the Tenthredinidae (e.g., the larch sawfly, *Pristiphora erichsonii*, and the birch leaf miner, *Fenusa pusilla*) and Cephidae

(e.g., the wheat stem sawfly, *Cephus cinctus*) are occasionally important.

The most significant deleterious role played by Hymenoptera is derived from reaction that humans have to being stung. Although a few species of parasitic Hymenoptera are capable of jabbing their ovipositor through the skin of a person, this usually requires that the person actually hold the wasp in his or her hand. This is not a health concern. Aculeate Hymenoptera, however, inject venom through the sting apparatus into the subject of their attention. In some cases, the venom can cause severe pain by virtue of the properties of the chemicals. The most significant hazard, though, lies with the reaction of the human immune system to the components of the injected venom. The typical response of swelling, pain, and itching may become life threatening for individuals that become hypersensitive to the components of the venom. In extreme cases, the resulting loss in blood pressure and shock can cause death, even from the sting of a single bee.

The so-called killer bees or Africanized honeybees are perceived as a serious health risk in the New World. These bees are not an aggressive exotic species that has invaded the Western Hemisphere, but rather are the same species as the honeybee that is used in apiculture throughout the world (*A. mellifera*). Colonies of the honeybee from Africa were brought to Brazil in order to try to breed some of their useful traits into the colonies of European honeybees being maintained in the tropics. Unfortunately, African bees escaped from confinement and became established as feral colonies. This strain of bees has spread and interbred with both feral and domesticated colonies of honeybees throughout South and Central America and into the southern US. The African subspecies of the honeybee is notorious because the colony more vigorously defends its nest against marauders than the European bees and is generally more sensitive to disturbance. In defense, the bees attack the animal or person disturbing the colony by stinging, just as do European bees. The difference between the two types of bees lies not in a more toxic venom but in the fact that the African forms are quicker to attack and are more persistent in pursuit and stinging. Numerous deaths of humans and livestock have been recorded that result from the high number of stings inflicted and not from hypersensitivity of the person or animal being attacked. In areas in which the honeybees have become more aggressive, greater care is needed by persons approaching the hives.

Biological Control Agents

Although some species of Hymenoptera have negative impacts on humans and their commerce, by and large the order is considered to be very beneficial. One of the most important categories of such beneficial species consists of those that either parasitize or prey on pest arthropods. These species, by eventually killing the other insects, act as biological control agents that reduce the population levels of their hosts. Such control agents have many advantages: Parasitoids are often fairly host specific, and thus their effects are focused on the pest problem; a population of biological control agents is capable of maintaining itself through time, thus providing continuous control; parasitoids are often extremely effective at locating their hosts in a complexly structured environment; and parasitoid populations respond to those of their hosts, increasing when host populations increase and decreasing when the availability decreases. The importance of parasitoids as regulators of the populations of their hosts is most clearly seen when, for example, a phytophagous species is accidentally introduced into a new region without its normal complement of natural enemies. Populations of such species are capable of explosive growth in the new environment, rapidly becoming important pests, for example, gypsy moths (*Lymantria dispar*) in North America and the cassava mealybug (*Phenacoccus manihoti*) in tropical America. Biological control programs involve the discovery of the natural environment and enemies of such pests and the importation of these natural enemies for control. Some examples of biological control have been spectacularly successful, with effective long-term control of pests. The parasitoid families Encyrtidae and Aphelinidae have been particularly effective in biological control programs of many species of sternorrhynchous Hemiptera.

Pollination

The successful development of seeds and fruit of many species of flowering plants depends on pollination – the movement of pollen from one flower to another. Although some species of commercially important plants, particulary grasses, are wind pollinated, a great many others rely on animal agents for the transfer of pollen. Under natural conditions pollinators can include vertebrates, such as bats and hummingbirds, and a wide variety of insects, including flies and beetles. Hymenoptera, however, and in particular the bees, are extremely important as pollinators. Recall that bees provision their nests with plant pollen on which their developing larvae feed. In the process of collecting pollen for their nests and also in collecting nectar from within the flowers, pollen grains that adhere to the hairy body of a bee can be transferred to the style of the flower. Natural pollinators include a vast array of native bees, but for many commercial crops colonies of honeybees are used to effect pollination. The honeybee is a polylectic species, that is, it gathers and its larvae feed on pollen from a wide variety of plant species. Colonies are regularly transported to fields in order to temporarily increase population levels to maximize fruit and seed production.

Honey and Other Bee Products

The large colonies of social bees also store honey within the nests that they construct. Honey is the modified and concentrated product of the nectar gathered principally at flowers and is used as an energy source for the adult bees. The most familiar species that serve as sources of honey for human consumption belong to the genus *Apis*, including the domesticated honeybee, but both bumblebees and stingless bees also produce and store honey. Honey production and the sale of it often provide a significant income supplement for small farmers. Another product from honeybee colonies that is of some value is beeswax, used variously in cosmetics, pharmaceuticals, and candles.

See also: Cooperation, Evolution of. Insects, Overview. Parasitoids. Phylogeny. Pollinators, Role of. Social Behavior. Systematics, Overview

References

Gauld ID and Bolton B (1988) *The Hymenoptera*. Oxford: Oxford University Press.

Goulet H and Huber JT (1993) *Hymenoptera of the World: An Identification Guide to Families, Publ No. 1894/E*. Ottawa: Research Branch, Agriculture Canada.

Hanson PE and Gauld ID (1995) *The Hymenoptera of Costa Rica*. Oxford: Oxford University Press.

Hölldobler B and Wilson EO (1990) *The Ants*. Cambridge, MA: Harvard University Press.

Michener CD (2007) *The Bees of the World*, 2nd edn. Baltimore, MD: Johns Hopkins University Press.

Naumann ID (1991) Hymenoptera. In: Naumann ID (ed.) *The Insects of Australia*, 2nd edn., pp. 916–1000. Ithaca, NY: Cornell University Press.

Quicke DLJ (1997) *Parasitic Wasps*. London: Chapman & Hall.

Sharkey MJ (2007) Phylogeny and classification of Hymenoptera. *Zootaxa* 1668: 521–548.

Identifying Conservation Priorities using a Return on Investment Analysis

Edward T Game, The Nature Conservancy, Brisbane, QLD, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Benefit function (or curve) Function that describes the relationship between the expected conservation gains relative to level of investment.
Complementarity Extent to which a given area or action contributes to the conservation of biodiversity that is unrepresented in existing conservation areas or projects.
Conservation action Any activity aimed at improving conservation outcomes.
Conservation–investment curve Function describing the marginal change in conservation benefit with magnitude of investment.
Conservation prioritization Process aimed at determining how conservation resources can best be used to achieve stated objectives.
Convention on biological diversity International convention aimed at the conservation of biological diversity and the sustainable use of natural resources.
Cost-effectiveness analysis A type of cost–benefit analysis where the benefit cannot be expressed as a monetary value.
Decision analysis Discipline concerned with procedures for formally representing and assessing decisions in order to provide insight for decision makers.
Decision support tools General term for systems, often computer based, that use decision-analysis procedures to provide targeted information to inform decisions.
Ecoregion Generally large geographic areas defined by characteristic biodiversity that is largely distinct from other ecoregions.
Lookup tables Tabular data structures that allow users to describe the result of combining two or more variables. Used to avoid generating or using otherwise complex functions to describe the relationship between variables.
Marginal gain Economic concept that refers to the difference between the increase in benefit from investing in more of something relative to the cost of doing so.
Opportunity cost In conservation, value of alternative activities forgone at site in favor of conservation actions.
Restoration Process of repairing or renewing degraded habitat.
Return-on-investment analysis In conservation, process of comparing options in a way that explicitly considers the cost of each alternative.
ROI Abbreviation or shorthand for return on investment.
Species–area curve Relationship between area of habitat and number of species that it is expected to contain.

What is a Return-on-Investment (ROI) Analysis?

In the field of conservation, return-on-investment (commonly referred to simply as "ROI") has come to be a rather general term for prioritization schemes that explicitly consider the cost of the alternatives being considered. These are in contrast to prioritization schemes based principally on the distribution of biodiversity (e.g., biodiversity hotspots; Myers *et al.*, 2000) or biodiversity and threats to it (e.g., Hoekstra *et al.*, 2005). The first published use of the term ROI referring to conservation prioritization was in 1998 (Bryant, 1998), but it was not until 2003 that the utility of ROI analysis for conservation prioritization began to explored in earnest (O'Connor *et al.*, 2003). However, before ROI entered the popular lexicon, the phrase "getting the biggest bang for our buck" had been appearing in conservation publications with increasing frequency. The common element of both terminologies is that the expected return or outcome from a conservation action (The term "action" is used here in a broad sense to include any allocation of resources aimed at improving conservation outcomes.) should be balanced against the cost of achieving that outcome.

In economic parlance, a return on investment is just that: the ratio of the amount of money you receive relative to the amount you need to invest in some endeavor (e.g., fund into a share portfolio or buying new equipment to improve the output of a business). The comparison between different options based on this ratio between investment and return is known as a cost–benefit analysis. Conservation ROI analyses belong to a general class of economic analysis known as cost-effectiveness analysis. Economists consider something to be a cost-effectiveness analysis rather than a cost–benefit analysis, when the outcome or return side of the equation is not monetized (i.e., expressed as a dollar value). This tends to be the case in conservation where our return might relate to the number of species saved or habitat protected. (The exception to this might be prioritizations involving

ecosystem services (Kareiva *et al.*, 2011)). Cost-effectiveness analysis was largely pioneered and is most widely used in health sciences, particularly looking at the effectiveness of different interventions where the desired metric might be human lives saved (e.g., deciding between passive case detection of tuberculosis or different levels of vaccination coverage; Edejer *et al.*, 2005).

A second use of the term ROI in conservation is associated with retrospective analyses of what conservation outcomes were delivered for a particular investment (e.g., Laycock *et al.*, 2009; Underwood *et al.*, 2009). Such analysis is more for accounting and communication purposes than prioritization of conservation actions, and it is not explored further in this article. Understanding the conservation outcomes that were actually realized as a result of conservation investment does, however, play an important role in more confidently estimating the expected return from future actions.

The Use of ROI Analysis in Conservation Prioritization

The main application of ROI analysis in conservation is in the field of conservation prioritization – tools or analysis that help us decide where and how to use conservation resources. Not all practitioners are comfortable with the use of the term prioritization, but if it is intended to inform or guide a decision about where or how conservation action takes place, then it is a prioritization scheme. ROI analysis has been applied to a number of different types of prioritization or resource-allocation problems. Probably the most common application has been to problems involving the spatial prioritization of conservation activities – in other words, "what places should be priorities for achieving desired conservation outcomes?" This question has been approached through ROI analysis at a global scale (e.g., Carwardine *et al.*, 2008a), regional scale (e.g., Moore *et al.*, 2004; Wilson *et al.*, 2006), and on national scales or below (e.g., Murdoch *et al.*, 2011). Although it is not always recognized as such, systematic conservation assessments (Margules and Pressey, 2000) are often close to ROI analyses, particularly when they are implemented using optimization software such as Marxan (Ball *et al.*, 2009) or Zonation (Moilanen and Kujala, 2008), whose algorithms work to minimize the cost of the set of places required to meet established conservation targets (Game and Grantham, 2008). If the cost data used with such software reflect the relative cost of taking conservation action in particular places, then this is essentially performing an ROI analysis (e.g., Carwardine *et al.*, 2008b).

The spatial analyses just described generally make the assumption that we are looking to apply a particular strategy at a chosen place. Although this assumption is more often than not implicit rather than explicit, it is necessary because it is hard to know the cost of working in a place without some idea of the activity you intend to take there. The most common type of analysis is based on prioritizing the establishment of some kind of conservation area (e.g., National parks). Protected areas make good fodder for analysis as the return or benefit from implementing them in a given place is potentially equal to the sum of the biodiversity present there. ROI analysis can also guide strategic decisions about investment between different activities or strategies. For example, Barlow *et al.* (2011) compare the cost-effectiveness of 18 different actions to reduce human–tiger conflict in Bangladesh. Possible actions include killing problem tigers, collaring tigers, using different kinds of fences and deterrents, and establishing education programs, and their effectiveness is judged by the number of expected human and tiger lives saved per dollar spent. Although this sort of analysis is consistent with the use of cost-effectiveness analysis in health sciences, it has been less common in conservation, in part because conservation problems have rarely been framed clearly enough to allow a convincing benefit or return function to be expressed in the same units for different activities.

More recently, ROI analyses have also been used to help prioritize both strategy and place simultaneously – in other words, answering the question, "what should we do where?" (e.g., Wilson *et al.*, 2007; Klein *et al.*, 2010). The addition of both spatial and strategic decisions inevitably complicates analysis because the universe of potential alternatives is vastly greater, but it remains an equally suitable and perhaps more useful and realistic application of ROI. Of course, the division between these three classes of problem (spatial, strategic, and combined) is somewhat artificial as all three of these can be framed generally as prioritizing between alternative actions.

Enthusiasm for ROI analysis in conservation is growing. The motivation for its application comes from a number of directions. In part, the magnitude of issues facing global biodiversity so dwarfs the resources available that we recognize the need to make hard choices and to be extremely efficient with those resources we do have. Particularly in the nonprofit world, donors want to see that their money is achieving the greatest benefit it can – a point that is hard to argue without applying ROI analysis to investment decisions. The shift toward a more analytical and economic approach to conservation decision making also reflects a change in the skill set and experience of those involved in conservation. Historically, conservation biologists have typically been drawn from empirical ecology, with many having little formal training in decision analysis or economics, whereas today these are both essential pieces of most graduate conservation training. There is also an increasing willingness and, indeed, interest in conservation to emulate approaches from other fields that share similar motivations of efficient resource use (or simply the desire to make the biggest profit) and have dedicated substantial expertise to developing methodologies for doing so. In addition, the influence of an increasing interest in ROI among the academic conservation community should not be underestimated. As is typical, application in the field follows methodological advances in the literature. In the case of ROI, the role of academic publications has been not only to demonstrate and raise awareness about ROI methodologies but also to illustrate the importance of considering cost when prioritizing conservation work (in contrast to a number of high-profile conservation prioritization schemes that ignored cost; see **Table 1**).

The Value of ROI Analysis in Conservation

The ultimate motivation for employing ROI prioritization methods in conservation is more efficient conservation of biodiversity – essentially, more biodiversity protected. However, because the application of ROI to conservation

Table 1 Number of additional species protected for five measures of biodiversity by applying an ROI approach and four alternative priority setting approaches with a conservation budget of $100 million per year over 20 years

Prioritization approach	*Total species richness*	*Vertebrate species richness*	*Plant species richness*	*Endemic vertebrate richness*	*Threatened vertebrates*	*No. of distinct vertebrates*
Return on investment	2524	341	2231	8	14	69
Areas with endemic species	1705	201	1504	6	8	45
Crisis biomes	1910	227	1683	3	10	23
Highly threatened species per dollar	1495	218	1276	4	9	43
Random	1571	208	1383	3	8	27

Source: Reproduced from Underwood EC, Shaw MR, Wilson KA, *et al.* (2008) Protecting biodiversity when money matters: Maximizing return on investment. *PLoS One* 3.

resource-allocation decisions is in its infancy, it is too soon to judge its effectiveness on this front. What makes a good decision is the process by which it was generated, not the ultimate outcome. From a philosophical standpoint, ROI analysis has a number of strengths relative to many existing approaches to conservation prioritization (whether this prioritization is explicit or not). These are mentioned throughout the article but are worth highlighting as they provide the foundation for the role and increasing use of ROI in conservation.

1. It acknowledges the importance of cost. ROI analysis reflects a recognition that the investment required to achieve conservation outcomes is important in deciding where and how we work. The resources available for conservation are limited, and directing them to the places with the most biodiversity does not necessarily lead to greatest protection of biodiversity. Conservation prioritizations have often overlooked this basic and fundamental point. This is despite numerous publications illustrating that ignoring cost in decisions about conservation priority is likely to lead to substantially less biodiversity being conserved than is possible for the same total resources (e.g., Bode *et al.*, 2008; Carwardine *et al.*, 2008a, b; Bottrill *et al.*, 2009). In other words, it is grossly inefficient to ignore cost. Similarly, high variation in the cost of taking conservation action (often orders of magnitude difference between alternatives) emphasizes the importance of considering cost in conservation decision making (Bode *et al.*, 2008; Murdoch *et al.*, 2011).
2. It encourages transparency about assumptions and values. Formulating ROI problems forces us to be explicit about the assumptions we are making. Being explicit about value judgments is a basic tenet of decision theory, and yet many conservation prioritizations and decision-support systems contain hidden assumptions and value judgments. For example, the widely used strategic planning process, Conservation Action Planning (TNC, 2007), uses a series of lookup tables to combine qualitative estimates of different aspects of threat to biodiversity (e.g., scope and severity) and produce an overall threat rank for the conservation feature of interest. These tables use rules like "One High or Two Medium threat rankings yield an Overall Threat Rank of Medium." Such statements contain many hidden values and risk tolerances, and yet their appropriateness is almost never considered before using the Conservation Action Planning process. Although the term "value judgment" is commonly used in a pejorative sense, they are an important and perfectly reasonable part of conservation decision making, provided we are transparent about them. Going through an ROI prioritization exercise has often served to illustrate inconsistencies and biases in conservation preferences that were concealed by earlier prioritization schemes.
3. It encourages the consideration of diminishing returns. ROI analyses typically recognize that there are likely to be diminishing returns on our conservation actions. For example, the more area you protect in a given ecoregion, the smaller the marginal gain in conservation value. Although the notion of complementarity has been an important feature of systematic conservation planning for more than 30 years, the idea that the marginal benefit of conservation decreases even before some ideal target has been reached is relatively new. This is not dissimilar to the concept that biodiversity with the least protection should be a higher priority than those with more protection. The diminishing returns argument has both genuine ecological and value judgment underpinnings. On the ecological side, biodiversity conservation is subject to diminishing returns because of the species–area curve (whereby the number of new species increases at a declining rate with increasing area). On the value side, many believe that for the sake of ecological equality, places with greater protection represent a lower priority than those with less protection.
4. It encourages accountability. Decision-support tools are just that – support tools. They should not be so prescriptive as to limit the ability to capitalize on opportunities. However, an ROI analysis can serve to provide a transparent view of the value of particular opportunities, which will hopefully encourage decision makers to carefully justify decisions that have objectively poor ROI.

Framing ROI Problems

As self-evident and straightforward as it might sound, being clear about the question that an ROI analysis is intended to answer is often one of the most contentious steps in conducting ROI analyses. Clearly framing the problem involves being explicit about what resource-allocation decision the ROI

analysis is intended to support (the question it is trying to answer) and the appropriate conservation objective for this purpose. For example, a conservation organization or agency might want to identify places that maximize the protection of representative biodiversity for an available budget, prioritize projects that represent the most cost effective way to reduce the extinction risk of threatened species, or explore what agricultural policy measures are likely to have the highest biodiversity return per dollar. The universe of possible problems is large, but clarity at this stage is essential for defining appropriate return and investment functions.

Outside of academia, the conservation community has been distinctly poor at framing its work through explicit and well-formulated problems. In part, this is because of the legitimate challenge of distilling the multiple aims and elements that influence conservation decisions into a single objective statement. Other reasons are more culpable: There is a common perception that all conservation actions are working toward good ends and therefore clearly stated objectives are not necessary; we commonly employee less-rigorous and less-transparent methods of prioritization that do not demand well-framed problems; and vagueness about objectives avoids disagreement about mission and also makes accountability more challenging.

Objectives do not have to be a question of science. Rather, objectives should reflect values or the basic reason for caring about the decision at hand (e.g., a belief that all biodiversity should be equally protected, or a desire for the maintenance of particular species, etc.). Sometimes this basic reason will reflect legal obligations, perhaps resulting from the Convention on Biological Diversity, or endangered species legislation. Perceiving that objectives must be supported by scientific understanding has been a common trap for conservation biologists conducting prioritizations. Compared with ecological evidence, value judgments represent a less-disputable and more-legitimate reason for prioritizing one ecological feature over another.

Defining Conservation Return

Despite economic analyses being novel for most conservation biologists, ironically it has been the process of defining and accounting for conservation returns that have proved the most difficult piece of ROI analysis. Ecological systems are undeniably more complicated than financial systems, as are the motives of those involved in conservation. Distilling our substantial and yet incomplete knowledge of ecological systems into metrics that adequately capture the values and motivations of those involved will always be challenging. Poorly defined benefit functions have dogged many conservation prioritization schemes, representing probably the single most common reason for their failure to significantly influence subsequent conservation investment.

Conservation return is essentially the expected benefit toward some predefined objective of taking a particular conservation action. It can be thought of as equivalent to defining a dose–response model (Hajkowicz, 2009). Although the return side of the conservation ROI equation does not have to be expressed in terms of monetary value, it must be in the same units for all alternatives being considered, and it must be possible to reliably estimate the likely return (or a surrogate for it) for each alternative being considered. This highlights two challenges – first, finding an appropriate metric; and second, estimating it. The former depends on having a clearly defined objective (*see* Framing ROI Problems) and on explicitly capturing the value systems of those using the prioritization. The latter will often depend on the creative use of surrogates or on expert opinion.

In the field of conservation, the basic building blocks of a return function are typically some surrogate for biodiversity – for example, the number of species in a particular taxonomic group (birds, mammals, or plants are common choices because they are relatively well studied) or habitat types. For large geographies where consistent data on species or habitats are lacking (like the continent of Africa), biodiversity might be represented by synthetic indices such as the Biodiversity Distinctiveness Index used by WWF (Burgess *et al.*, 2006). For more focused ROI prioritization problems, such as choosing between actions to conserve a threatened species, the basic return metric might relate to extinction risk of that species or the number of viable populations (e.g., Bode and Wintle, 2010) or, where projects across multiple species are being considered, the number of species expected to remain extant over a given time horizon (e.g., Joseph *et al.*, 2009). Recently, some conservation prioritizations have also included ecosystem services (e.g., freshwater provision or carbon sequestration) as fundamental elements of conservation return (e.g., Tallis and Polasky, 2009; Goldman *et al.*, 2010).

Benefit Functions and Diminishing Returns

Almost all conservation prioritizations will want to use more discriminate return functions than simply the number of species or habitat types. For example, we might care about protecting representative habitat types but are also likely to care about whether that habitat is intact or highly fragmented. A whole host of considerations influence how we value different elements of biodiversity, and these should be reflected in the return function as far as possible. To do so, we need to weight our base biodiversity metrics with these additional values but do so in a way that is consistent, transparent, and nonarbitrary. Species might be weighted by things such as taxonomic distinctiveness, rarity, cultural value, or financial value (e.g., Faith, 2008; Joseph *et al.*, 2009). Habitats might be weighted by condition, level of fragmentation, or the percent remaining (e.g., Lipsett-Moore *et al.*, 2010; Murdoch *et al.*, 2011).

Adding another layer of complexity is the fact that these weighted returns will almost never change linearly with changes in the variable under consideration. As an example, we might consider that very fragmented habitats have little conservation value, but once the level of fragmentation is even minimally reduced, conservation value increases rapidly (**Figure 1**). In practice, this means that elements contributing to weighted return functions are often further transformed according to a benefit curve. Defining these weightings and benefit curves is what it means to be explicit about value judgments.

The value of a conservation action is often considered dependent on what conservation has already occurred. For example, adding additional conservation areas to a region that is already well protected might be considered of less benefit than

conserving the same area in a region with very little existing protection. This concept is known as diminishing returns. Diminishing returns are frequently implemented in ROI analysis through the use of a species–area curve (but see, Smith, 2010). The details of this curve and procedure are described below (*see* ROI with Diminishing Returns), but it is noteworthy that the actual form of the species–area curve is relatively unimportant for the function it performs (Bode and Murdoch, 2009).

Probability of Success

In nearly all prioritization situations, there is likely to be some uncertainty in the expected conservation return from a given action. Of particular concern are events or risks that are beyond the control of conservation management and yet are likely to influence the success of conservation interventions. In ROI analyses, these risks are typically expressed as a probability of success. This is the likelihood that a particular chosen conservation alternative will deliver the return expected of it. Probabilities are logically multiplicative, such that the expected return should be calculated as $E = R \times P$, where R is the return function derived from the elements described above and P is the probability of success. The probability of success might be derived from the assessment of experts; for example, Joseph *et al.* (2009) asked experts to estimate the likelihood of different factors that might contribute to the failure of threatened species recovery projects (e.g., operational, legal, political, and social) and then combined these to calculate an overall probability of success. Game *et al.* (2008) used models of the risk of catastrophic bleaching to determine the likely success of protecting different reefs on the Great Barrier Reef, Australia. And Tear *et al.* (unpublished) used the Ibrahim Index of African Governance (http://www.moibrahimfoundation.org/en/section/the-ibrahim-index) as a surrogate for the probability of successful conservation projects in different African countries.

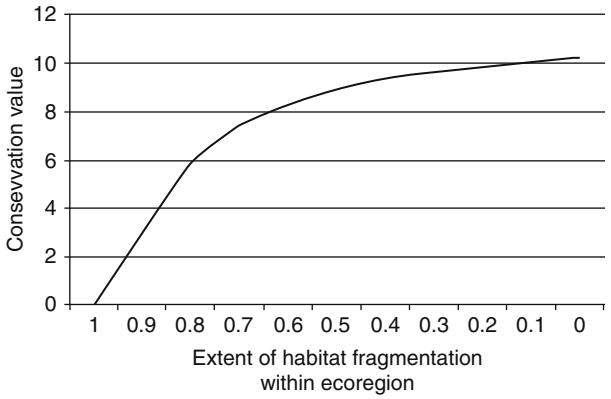


Figure 1 Conservation benefit function applied to habitat fragmentation within African ecoregions during an ROI analysis conducted by The Nature Conservancy (Tear *et al.*, unpublished). To give the conservation return in each ecoregion, a base measure of biodiversity was multiplied by the level of habitat fragmentation transformed according to the modified exponential function, $ae^{(b/F)}$, where F is the inverse of fragmentation extent in the ecoregion, and a and b are constants that reflect the value judgments of The Nature Conservancy staff with respect to habitat fragmentation and conservation value.

Estimating Investment

It goes without saying that a well-thought-out estimate of the costs associated with the alternatives under consideration is a prerequisite for an ROI analysis. Although some ROI papers use terms such as realistic or thorough when setting standards for cost data, these are desirable but not necessary. Most important is that the cost data used reflect something that is relevant to the decisions at hand, and that the relative cost of each alternative under consideration can be reliably estimated.

Numerous elements contribute to the cost of taking a conservation action somewhere (Naidoo *et al.*, 2006). (Conservation biologists tend to use cost in a looser sense than economists would be comfortable with. For example, Klein *et al.* (2010) define the opportunity cost of establishing marine protected areas as equivalent to forgone revenue (quantity × price), a common convention in conservation. In reality, economists would define opportunity cost as net benefits forgone, which is profit not revenue.) Although actual costs are undeniably complicated to estimate, cost surrogates useful for prioritizations are often surprisingly easy to estimate because there are typically limited measures available. At large scales, there will often only be a couple of options. At finer scales, cost data will often depend on either the creative use of surrogates or expert opinion. Ideally, in estimating the investment required of alternative conservation options, we would be able to draw on and extrapolate from actual costs of conservation, and some data sets do (e.g., Balmford *et al.*, 2004). Unfortunately, the field of conservation has not been good at tracking the actual costs associated with different conservation actions in different places, so despite many years of conservation activity we rarely have good data to draw from. **Table 2** identifies some typical types of costs and surrogate measures for them that have been employed during ROI prioritizations.

Table 2 Different types of cost data used during conservation prioritizations, with published examples

Type of cost	*Example data or surrogate*
Management cost	Modeled annual management costs of terrestrial protected areas (Moore *et al.*, 2004; Bode *et al.*, 2008)
	Modeled annual management costs of marine protected areas (Balmford *et al.*, 2004)
	Predicted management costs associated with threatened species recovery projects (Joseph *et al.*, 2009)
Opportunity cost	Expected agricultural productivity (Naidoo and Iwamura, 2007)
	Historical fishing revenue (Game *et al.*, 2008; Klein *et al.*, 2010)
Purchase cost	Current market value of land (Murdoch *et al.*, 2011)
Transaction cost	The number of language groups with which negotiation would need to happen (Lipsett-Moore *et al.*, 2010)

ROI Analysis

The following examples illustrate a variety of ROI analyses, along with the conservation prioritization question they were intended to inform. For each example, there is a description of the problem, the return model, the investment model, and the equation or algorithm used to generate a solution. They are presented roughly in order of increasing analysis complexity.

Prioritizing between Restoration Strategies

The most straightforward ROI analysis is developing a simple ROI rank of alternatives. This is essentially answering the question, "What is the most cost-effective alternative in which to invest the next dollar?" Goldstein *et al.* (2008) describe the use of an ROI analysis to rank six alternative approaches to the reforestation of montane pastureland in Hawaii. Each restoration approach represents a different set of land-use transitions. Based on an average 200 ha parcel in their study area, the authors model conservation return as the expected increase in mean native bird and plant understory density under each of the six alternative restoration approaches. Required investment was modeled as the net present value of expected restoration management over 50 years. To calculate the ROI of each restoration approach, the authors simply divided the expected increase in native bird and understory plant density by the expected cost of that restoration transition pathway (see **Figure 2**).

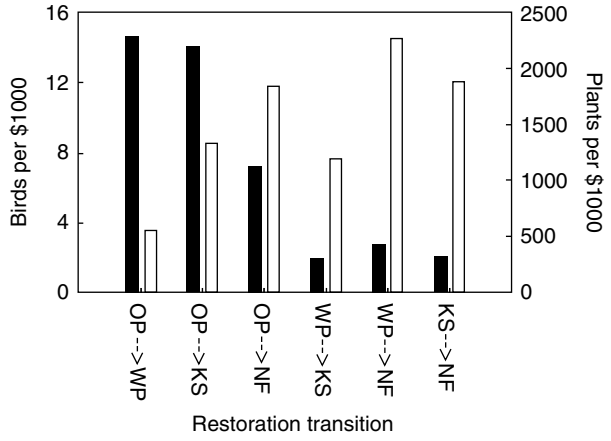


Figure 2 ROI ratio for each restoration transition for native forest birds (filled bars) and understory plants (open bars). Reproduced from Goldstein JH, Pejchar L, and Daily GC (2008) Using return-on-investment to guide restoration: A case study from Hawaii. *Conservation Letters* 1: 236–243, with permission from Wiley.

ROI with Diminishing Returns

To demonstrate the value of even a simple ROI analysis, Murdoch *et al.* (2007) calculated the expected ROI of buying land for conservation in each of the 21 North American ecoregions that are dominated by temperate broadleaf and mixed forests. Conservation return was defined as the number of plant species conserved across all ecoregions. Required investment was assumed to be equivalent to the mean per acre land value in each ecoregion. The expected number of species protected per acre purchased depends on not only how many species are present in each ecoregion but also how much of the ecoregion is already protected (the diminishing returns element). In ecoregions that are already well protected, new protected areas will add fewer additional species. To accommodate these diminishing returns, the authors weighted the expected return by the species–areas curve. Thus, the relationship between the number of species conserved, S, and the area protected, A, is given by $S=\alpha A^z$, where α is an ecoregion specific constant based on the number of species present in the ecoregion, and $z=0.2$, an established value for the rate of accumulation of plant species on continental landmasses. The marginal return (number of species) of protecting additional land in each ecoregion is given by the derivative of the species–area equation with respect to A:

$$\frac{dS}{dA} = z\alpha A^{z-1}$$

With an estimate of species accumulation with area and knowledge of the cost per unit of land area, it is possible to define an ROI curve (**Figure 3**). In the case of Murdoch *et al.* (2007), this is a species–investment curve because species is their unit of return. Based on where each ecoregion starts on that curve, the authors rank each ecoregion

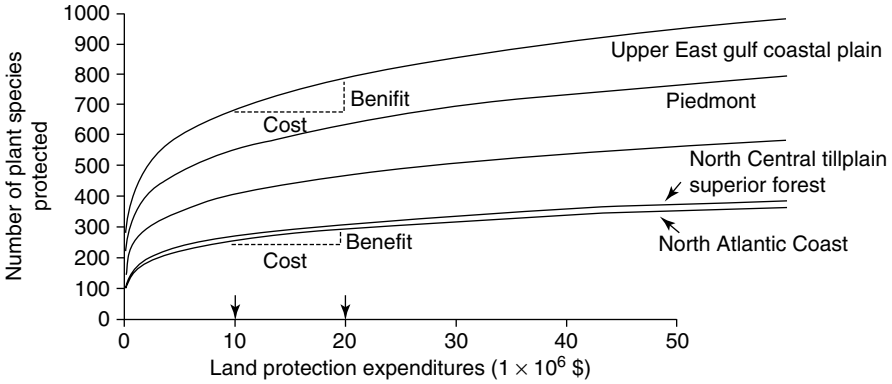


Figure 3 Species–investment curves for five US ecoregions dominated by temperate broadleaf and mixed forests. Reproduced from Murdoch W, Polasky S, Wilson KA, Possingham HP, Kareiva P, and Shaw R (2007) Maximizing return on investment in conservation. *Biological Conservation* 139: 375–388.

Table 3 Return on investment by ecoregion based on the increase in plant species conserved per $1 million investment

Ecoregion	*Plant richness*	*Total area (million ha)*	*Land value ($ ha^{-1})*	*ROI: Plant species conserved per $1 million*	*Rank*	*Current protected area (ha)*	*ROI: Increase in plant species conserved per $1 million investment*	*Rank*
North Central Tillplain	2243	12.316	3990.75	6.94	12	0	258.54	1
Upper East Gulf Coastal Plain	3363	13.708	2050.97	20.26	1	49,348	2.15	2
Ouachita Mountains	1743	4.647	2031.21	10.6	3	29,742	2.09	3
Western Allegheny Plateau	2487	10.782	3083.88	9.96	5	33,425	1.51	4
Piedmont	3363	17.137	4732.07	8.78	7	25,705	1.5	5
Cumberlands and Southern Ridge Valley	2487	12.566	3002.33	10.23	4	35,185	1.45	6
High Allegheny Plateau	1883	6.826	3165.42	7.35	11	34,131	1.2	7
St. Lawarence/ Champlain Valley	1381	6.146	2268.43	7.52	9	42,409	1.06	8
Willamette Valley	1067	1.485	7724.51	1.71	19	10,097	1	9
Chesapeake Bay Lowlands	1488	4.396	6283.89	2.93	18	22,418	0.73	10
Interior Low Plateau	2332	19.329	2925.73	9.85	6	110,175	0.51	11
Ozarks	2332	13.892	2149.82	13.4	2	219,500	0.43	12
Mississippi River Alluvial Plain	1468	10.972	2686.04	6.75	14	141,538	0.32	13
Central Appalachian Forest	2398	9.656	4012.99	7.38	10	207,604	0.27	14
Prairie-Forest Border	1420	15.833	3494.07	5.02	16	134,580	0.23	15
Southern Blue Ridge	2398	3.809	5520.33	5.37	15	276,547	0.19	16
Lower New England/ Northern Piedmont	1695	9.401	13852.73	1.51	20	50,763	0.17	17
North Atlantic Coast	1695	5.138	23830.84	0.88	21	39,561	0.14	18
Superior Mixed Forest	1459	20.76	2142.4	8.41	8	1,123,097	0.07	19
Great Lakes	1459	57.519	3617.62	4.98	17	1,000,831	0.04	20
Northern Appalachian-Boreal Forest	1496	33.454	2718.16	6.8	13	1,361,569	0.04	21

ROI is given both with and without diminishing returns.

Source: Adapted from Murdoch W, Polasky S, Wilson KA, Possingham HP, Kareiva P, and Shaw R (2007) Maximizing return on investment in conservation. *Biological Conservation* 139: 375–388.

based on the increased number of plant species expected to be conserved per million dollars invested in that ecoregion (**Table 3**).

ROI Considering Threat

The expected benefit of a conservation action depends not only on what conservation already exists, but also on what would be expected to happen in the absence of that action. Protection of habitat that prevents clearing that would have happened otherwise is arguably more valuable than protecting a piece of habitat that in unlikely to be cleared regardless of conservation investment. This concept is used by Murdoch *et al.* (2011) in conducting an ROI analysis to establish the most-effective cadastral parcels to purchase in order to conserve 10% of all 77 ecosystem types that compose the grasslands of Argentina. Ten percent of Argentina's grassland is equivalent to 15.5 million ha.

In the analysis of Murdoch *et al.* (2011), the expected conservation return from purchasing a cadastral parcel for conservation depends not only on a diminishing returns function according to how much of that ecosystem type is already protected but also on the risk of that parcel being converted. Because this risk is assumed to be entirely mitigated once the parcel is protected, the return for each parcel is defined as the intrinsic value of the parcel multiplied by the probability that the parcel will be converted if not protected. Conversion risk for each parcel was estimated based on the human footprint dataset. The purchase or investment cost for each parcel was defined as the per hectare cost of land based on mean real estate values for that parcel. The analysis was used to indentify spatial priorities for increasing grassland protection to 15.5 million ha. For illustrative purposes, the authors compared priorities indentified using the ROI approach to those identified using two other plausible prioritization strategies; selecting those areas with the highest conservation return regardless of cost, and selecting the

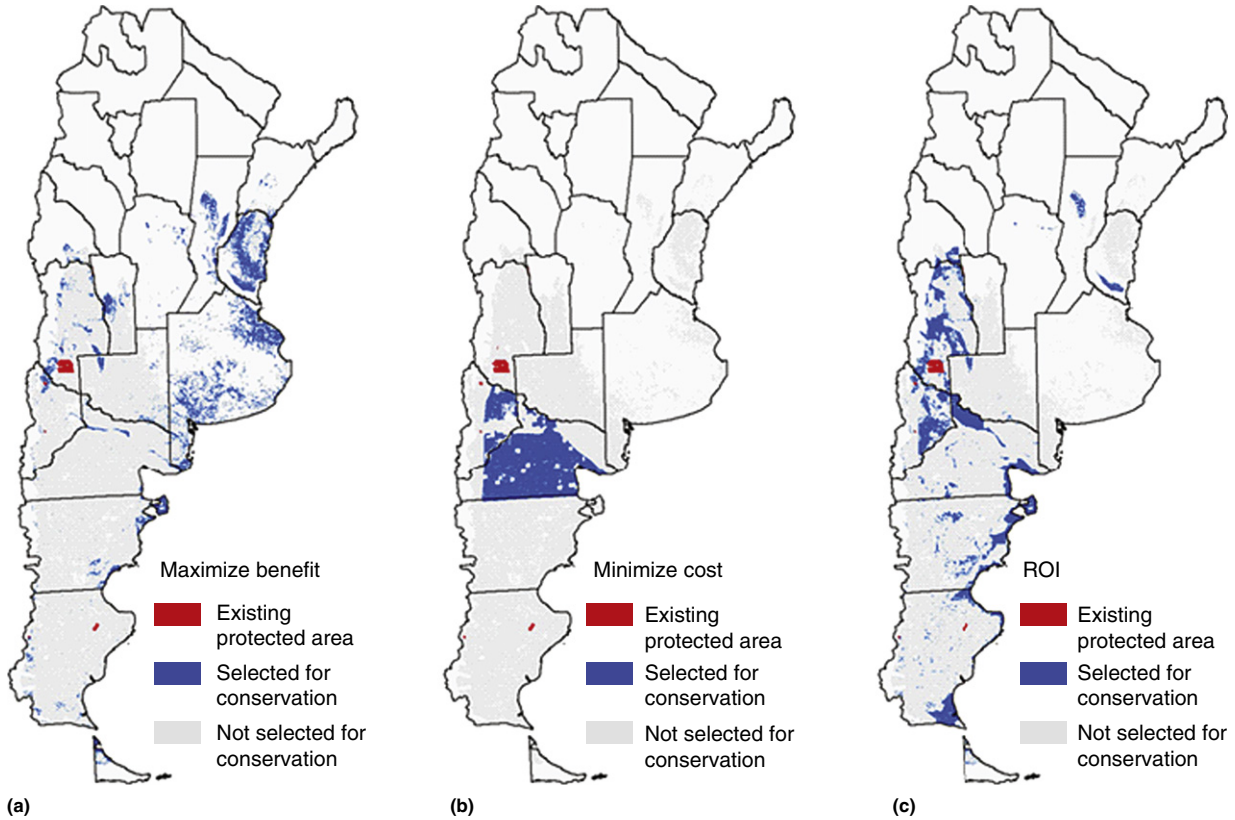


Figure 4 Priority conservation areas resulting from applying each of three prioritization strategies to achieve the goal of 15.5 million new ha conserved (10% of total grasslands. (a) Maximize benefit – selecting those areas with the highest conservation return regardless of cost; (b) minimize cost – selecting the cheapest areas regardless of conservation return; (c) ROI. Existing protected areas (0.5 million ha) are shown in red. Reproduced from Murdoch W, Ranganathan J, Polasky S, and Regetz J (2011) Using return on investment to maximize conservation effectiveness in Argentine grasslands. *Proceedings of the National Academy of Sciences of the United States of America* 107: 20855–20862, with permission from PNAS.

cheapest areas regardless of conservation return. The substantial difference in priority between these three approaches is shown in **Figure 4**.

ROI with Actions and Locations

The preceding examples use ROI analysis either to prioritize actions (restoration approaches) without specific reference to the location where this will occur or to prioritize the location for an already established action (land purchase). ROI analysis can also be used to prioritize both action and location simultaneously, essentially answering the question, "What should we do where?"

The reefs of the Coral Triangle are threatened by both unsustainable and damaging fishing practices and by runoff from the adjacent land containing sediment, nutrients, and other organic pollutants. The threat from fishing can be ameliorated through the establishment of marine protected areas, but mitigating the threat posed from runoff requires terrestrial conservation measures in coastal catchments. Klein *et al.* (2010) describe an ROI analysis to prioritize conservation investment between these two activities across the 16 marine ecoregions of the Coral Triangle. To model conservation return, the authors use a formula developed by Halpern *et al.* (2008) to estimate the overall level of threat from both marine and terrestrial sources being experienced by reefs in each ecoregion. Assuming that these threats reduce linearly as the proportion of the ecoregion protected increases, conservation return per action per ecoregion (or ecoaction, as the authors call them) is defined as the expected reduction in overall threat to reefs either from protecting either an area of reef or land in that ecoregion. The cost of investing in each possible ecoaction is calculated as the sum of the annual management and opportunities costs associated with marine and terrestrial protected areas in those locations. Protected-area management costs were calculated based on modes developed by Balmford *et al.* (2004) and Moore *et al.* (2004) and subsequently generalized by Bode *et al.* (2008), whereas opportunity costs are based on expected agricultural profitability in terrestrial areas (Naidoo and Iwamura, 2007) and expected catch value in marine areas. Each action in each ecoregion was then ranked according to its expected ROI (**Figure 5**).

This ranking provides two useful pieces of information for prioritization. First, it provides an ROI comparison of marine versus terrestrial action (for marine conservation) within each ecoregion. Second, it provides a comparison of the expected ROI (both marine and terrestrial) between ecoregions across the Coral Triangle. With the exception of one ecoregion (northern Philippines), marine protection always has a higher ROI than terrestrial protection within an ecoregion. When ROI ranks are compared across all

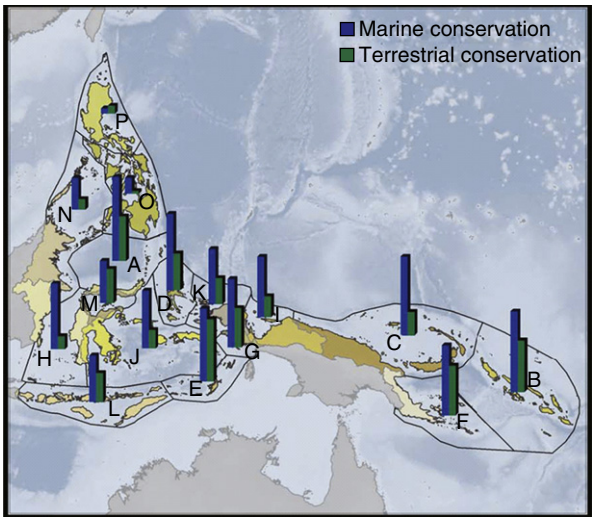


Figure 5 Ecoaction rankings indicate their relative priority for coral reef conservation investment across the Coral Triangle (taller bar, higher rank). Letters labeling ecoregions follow the ranking order for marine conservation (i.e., ecoregion A ranks highest for marine conservation). Adapted from Klein CJ, Ban NC, Halpern BS, *et al.* (2010) Prioritizing land and sea conservation investments to protect coral reefs. *PLoS One* 5: e12431.

ecoregions of the Coral Triangle, however, terrestrial action in some ecoregions often has a higher ROI than marine action in others.

Using ROI Analysis to Develop an Optimal Allocation Schedule

Threatened Species Recovery Projects with a Fixed Budget and Time

Conservation agencies often have fixed annual budgets with which to allocate and no shortage of competing demands. For example, New Zealand has more than 2000 species that are listed as threatened and for which the Department of Conservation has the task of improving conservation status. Joseph *et al.* (2009) described the use of an ROI analysis to help the New Zealand Department of Conservation prioritize threatened species recovery projects, with the objective of improving the security of the greatest possible number of unique species. The term security refers to security from extinction in the wild (achieved when available evidence indicates there is a viable population), and uniqueness of a species is measured as its taxonomic distinctiveness. The analysis compared potential projects for 32 different species for which the current annual budget for management across all species was NZD $512,000. Given the time required for threatened species recovery projects, the analysis looked at the optimal allocation of resources over the next 50 years – a total budget (with discounting) of $20,269,096.

Each species i, was given a weighting W_i, according to its taxonomic distinctiveness. Joseph *et al.* (2009) then used experts to estimate the expected return from projects for each species. The biodiversity benefit of the project, B_i, is the difference between the probability of the species being secure in 50 years with, P_i, and without, P_0, the implementation of that project; $B_i = P_i - P_0$. The same experts were also asked to estimate the probability of success for the project, S_i, and the cost for the project over 50 years, C_i. The ROI for project i (or project efficiency in the authors' language) is calculated as $\text{ROI}_i = (B_i \times W_i \times S_i)/C_i$. The authors use the ROI rank to work out which projects should be funded over the 50 years with the given budget in order to secure the greatest number of threatened species. They compare the results to other allocation approaches (**Table 4**).

Land Purchase Priorities with a Fixed Budget

The preceding examples all rank a series of alternative options based on their current expected ROI. As discussed, however, an available alternative's ROI depends to some extent on what conservation has already taken place. This means that ROI will change as we invest conservation resources (diminish as conservation takes place), and it might lead to changes in the ROI rank of options following some investment. If we have a known conservation budget, we can use the conservation–investment curves for each alternative to calculate the optimal allocation of this budget to different alternatives. In its simplest form, this optimal allocation proceeds by investing in the alternative with the highest ROI until another alternative is expected to have a higher ROI, at which point investment is switched to this alternative – the same process being repeated until the entire budget is allocated. The alternative with the steepest conservation–investment curve according to current investment levels has the highest ROI.

Murdoch *et al.* (2007) extended their temperate broadleaf forest ecoregion analysis (*see* ROI with Diminishing Returns) to define the optimal allocation schedule across ecoregions for five different total budgets. These allocations are those that are expected to protect the greatest number of plant species across all ecoregions for that total budget (**Table 5**).

Allocation Across Actions and Locations

To demonstrate how an ROI framework can be used to inform the optimal allocation of resources across both actions and locations, Wilson *et al.* (2007) calculated the optimal investment of conservation resources to address the threats faced by 17 of the world's Mediterranean ecoregions. Each ecoregion has a different mixture of key threats and different candidate actions to address them. In total, they evaluated 51 ecoregion–conservation action combinations (ecoactions). With the objective of maximizing the total number of Mediterranean vertebrates and vascular plants conserved, the authors allocate a hypothetical budget of US$100 million annually for 20 years to reduce threats to species in the 17 ecoregions. For each ecoaction combination, a variety of data (expert input, literature review, and analysis of regional datasets) was used to estimate not only the number of at-risk species likely to be conserved as a result of that action but also the cost of successfully implementing the action. Together with the expected rate of diminishing returns, these estimates were used to generate conservation–investment curves for each ecoaction combination (**Figure 6**).

At any point in time, the optimal allocation of resources across ecoactions is determined by an interplay between

Table 4 New Zealand threatened species recovery projects ranked by weighted ROI, unweighted ROI, cost, taxonomic distinctiveness, and threat status

Species	*Weighted ROI rank*	*Unweighted ROI rank*	*Cost rank*	*Distinctiveness rank*	*Threat status rank*
Dactylanthus	1[a]	2[a]	6[a]	1	27
Maud Island frog	2[a]	6[a]	15[a]	4	20
Shrubby tororaro	3[a]	9[a]	7[a]	3	26
Hamilton's frog	4[a]	10[a]	16[a]	5	7[a]
North Island brown kiwi	5[a]	18	30	2	28
Climbing everlasting daisy	6[a]	1[a]	3[a]	13	17
Hochstetter's frog	7[a]	13[a]	2[a]	6	31
New Zealand shore plover	8[a]	11[a]	10[a]	10	3[a]
Pittosporum patulum	9[a]	12[a]	21	9	21
Oreomyrrhis sp.	10[a]	3[a]	4[a]	17	18
Pachycladon exilis	11[a]	7[a]	12[a]	12	5[a]
Archey's frog	12	17	23	7[a]	11[a]
Canterbury mudfish	13	16[a]	13[a]	8[a]	19
Carmichaelia hollowayi	14	5[a]	5[a]	24	1[a]
Poa spania	15	4[a]	9[a]	25	2[a]
Chatham Island oyster catcher	16	14[a]	19[a]	18	10[a]
Kaki	17	21	31	14	14
Cardamine cf. *bilobata*	18	15[a]	11[a]	22	4[a]
Black robin	19	20	17[a]	19	8[a]
Pygmy button daisy	20	22	1[a]	21	16
Cook Strait giant weta	21	8[a]	8[a]	32	32
Bignose galaxias	22	27	20[a]	15	29
Long-tailed bate	23	26	26	20	24
Carabid beetle	24	19	14[a]	26	6[a]
Lowland long jaw galaxid	25	29	18[a]	16	9[a]
Orange-fronted parakeet	26	32	32	11	15
Mohua	27	28	25	23	23
Grand skink	28	23	27	27	12
Otago skink	29	24	28	28	13
Short-horned grasshopper	30	25	22	31	30
Chevron skink	31	30	29	29	25
Robust grasshopper	32	31	24	30	22

[a]Species selected with a fixed budget of NZD $20,269,096.

Source: Adapted from Joseph LN, Maloney RF, and Possingham HP (2009) Optimal allocation of resources among threatend species: A project prioritization protocol. *Conservation Biology* 23: 328–338.

(1) the existing level of conservation investment in that ecoaction, (2) the relationship between additional area invested in each ecoaction and the biodiversity benefit, and (3) the cost of this investment. Because of the dependency between the optimal ecoaction and the investment taking place in other ecoactions, solving for a truly optimal allocation schedule through time among such a large number of ecoactions is computationally intractable (Costello and Polasky, 2004; Wilson *et al.*, 2006). With a smaller number of options ($\sim < 8$), it is possible to use methods such as stochastic dynamic programming to find the optimal solution, although "rule of thumb" or heuristic algorithms can be used to solve large problems quickly and have been found to closely approximate the optimal investment schedule (Wilson *et al.*, 2006). This heuristic used by Wilson *et al.* (2007), "maximizes short-term gain" by directing funds each year to ecoactions that provide the greatest short-term increase in biodiversity benefit per dollar invested. Using this heuristic, the authors generate an investment schedule that allocated some money to 30 different ecoactions over the 20-year period; it is expected to result in the greatest number of Mediterranean species conserved (**Table 6**).

Extensions and Variations

The preceding examples represent conventional applications of ROI analysis in conservation prioritization. ROI is, however, a flexible tool that can be used to support a wide range of conservation-prioritization and resource-allocation problems. The following examples illustrate some extensions and variations of ROI analysis for other situations that are likely to be commonly encountered in conservation.

ROI with Dependent Returns

Using ROI to optimally allocate resources through time recognizes that there is a relationship between the value of one action and the investment already occurring in others. However, each conservation option is still assumed to be independent (i.e., the return and required investment do not depend on what other actions are being taken). In reality, such dependencies commonly exist. Biodiversity often faces multiple threats in the same place, and conserving it might require multiple actions to ameliorate these threats. These dependencies can be incorporated into an ROI analysis. Evans *et al.* (2011) explore the situation in which

Table 5 Optimal budget allocation across ecoregions

Ecoregion	*Total budget allocation*				
	$100 million	$200 million	$300 million	$400 million	$500 million
North Central Tillplain	$37,651,522	$49,584,149	$59,853,018	$69,612,043	$79,371,069
Upper East Gulf Coastal Plain	$30,862,887	$53,627,571	$73,218,189	$91,836,145	$110,454,101
Ouachita Mountains	$17,036,561	$30,187,220	$41,504,291	$52,259,476	$63,014,661
Piedmont	$5,880,874	$23,348,979	$38,381,517	$52,667,698	$66,953,880
Western Allegheny Plateau	$5,530,982	$20,502,457	$33,386,468	$45,630,795	$57,875,123
Cumberlands and Southern Ridge Valley	$3,037,174	$17,543,961	$30,028,075	$41,892,360	$53,756,645
High Allegheny Plateau	$5,205,663	$15,339,848	$24,970,877	$34,601,907	
St. Lawarence/Champlain Valley	$5,826,109	$13,120,561	$20,415,012		
Willamette Valley	$2,462,485	$8,010,044	$13,557,603		
Central Appalachian Forest					
Chesapeake Bay Lowlands					
Great Lakes					
Interior Low Plateau					
Lower New England/Northern Piedmont					
Mississippi River Alluvial Plain					
North Atlantic Coast					
Northern Appalachian-Boreal Forest					
Ozarks					
Prairie-Forest Border					
Southern Blue Ridge					
Superior Mixed Forest					

Source: Reproduced from Murdoch W, Polasky S, Wilson KA, Possingham HP, Kareiva P, and Shaw R (2007) Maximizing return on investment in conservation. *Biological Conservation* 139: 375–388.

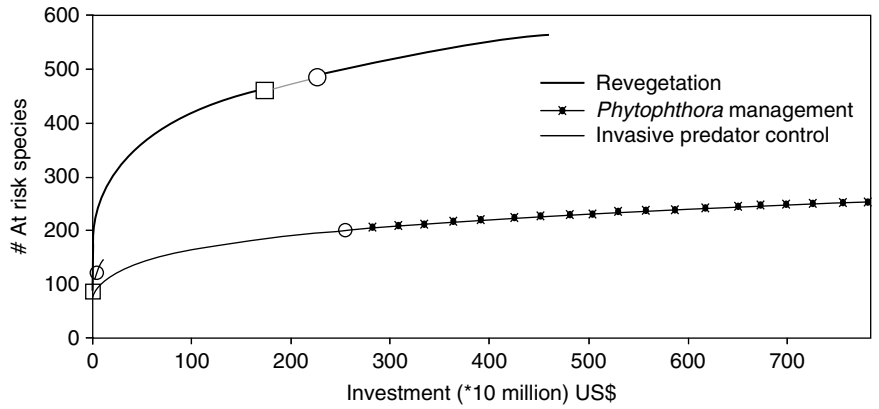


Figure 6 Conservation–investment curves for three conservation actions: *Phytophthora* management, invasive predator control, and revegetation in the Swan Coastal Plain scrub and woodlands ecoregion of Australia. The area of conservation interest is illustrated by the curve to the left of the circles, comprising the area already receiving the action (denoted by the solid lines between the origin and the squares) and the area requiring the action (denoted by the dashed lines between the squares and the circles). Adapted from Wilson KA, Underwood EC, Morrison SA, *et al.* (2007) Conserving biodiversity efficiently: What to do, where, and when. *PLoS Biology* 5: 1850–1861.

some conservation benefits cannot be realized unless two actions are taken in the same place – in this case, protecting threatened Australian species, which requires both foxes and rabbits to be controlled in order to benefit. Although controlling foxes and rabbits on their own will benefit a wide range of threatened species, some species require both. Explicitly incorporating these dependent benefits into an ROI framework resulted in a far more cost-effective conservation prioritization than considering them only independently (**Figure 7**). The results also emphasize the importance of considering the full range of threatening processes when estimating expected conservation returns.

Relationship Between Cost, Risk, and Complementarity

Another type of dependency that is more familiar to conservation planners is the notion of complementarity. When trying to efficiently conserve the range of biodiversity that exists in a region, the value of taking action to conserve a given location depends on what other locations are already protected and the species conserved there. Although this dependency represents a substantial computational challenge, there is freely available conservation software, such as Marxan (Ball *et al.*, 2009) or Zonation (Moilanen and Kujala, 2008),

Table 6 Percent of the total budget (US$100 million) allocated to the 30 ecoactions that receive investment over both 5 and 20 years

Ecoaction	*Percent of budget allocated over 5 years*	*Percent of budget allocated over 20 years*
Land protection and management in the montane fynbos and renosterveld to abate agricultural conversion	21	14
Land protection and management in the montane fynbos and renosterveld to abate urban development	18	4
Land protection and management in the lowland fynbos and renosterveld to abate urban development	11	8
Land protection and management in the lowland fynbos and renosterveld to abate agricultural conversion	8	14
Invasive plant control in the montane fynbos and renosterveld	8	3
Invasive plant control in the Chilean matorral	7	10
Land protection and management in the Albany thickets to abate agricultural conversion	4	5
Land protection and management in the Albany thickets to abate urban development	4	5
Invasive plant control in the lowland fynbos and renosterveld	4	1
Invasive plant control (riparian invasives) in the montane chaparral and woodlands	3	3
Fire suppression in the Chilean matorral	2	1
Invasive plant control in the Albany thickets	2	2
Phytophthora management in the Swan Coastal Plain scrub and woodlands	1	2
Phytophthora management in the Mount Lofty woodlands	1	1
Phytophthora management in the Jarrah-Karri forest and shrublands	1	2
Phytophthora management in the Southwest Australia savanna	1	1
Invasive predator control in the Coolgardie woodlands	1	1
Phytophthora management in the Naracoorte woodlands	1	1
Phytophthora management in the Southwest Australia woodlands	1	1
Phytophthora management in the Eyre and York mallee	1	1
Phytophthora management in the Murray–Darling woodlands and mallee	1	1
Invasive predator control in the Eyre and York mallee	1	2
Phytophthora management in the Esperance mallee	0	1
Invasive predator control in the Naracoorte woodlands	0	1
Invasive plant control (priority noxious weeds) in the montane chaparral and woodlands	0	5
Invasive plant control (priority noxious weeds) in the interior chaparral and woodlands	0	3
Invasive plant control (priority noxious weeds) in the coastal sage scrub and chaparral	0	2
Invasive predator control in the Swan Coastal Plain scrub and woodlands	0	2
Invasive predator control in the Mount Lofty woodlands	0	1
Land protection and management in the montane chaparral and woodlands to abate urban development	0	0.35

Source: Reproduced from Wilson KA, Underwood EC, Morrison SA, *et al.* (2007) Conserving biodiversity efficiently: What to do, where, and when. *PLoS Biology* 5: 1850–1861.

to help find good solutions. These software are also good at finding cost-effective solutions, essentially conducting an ROI analysis that incorporates complementarity. (Although Marxan does find cost-effective solutions, it does not technically maximize ROI because the algorithm does not consider benefits beyond the target amount. The algorithm used in Zonation, however, more closely approximates an ROI optimal solution.) Game *et al.* (2008) describe the use of Marxan to find cost-effective solutions to protecting representative examples of reefs on Australia's Great Barrier Reef that also consider the risk of climate change eliminating the expected conservation return after investment. The results illustrate a relationship between cost and probabilistic risk in which a number of reefs that represented very low climate change risk were not prioritized when conservation budgets were limited because they were valuable to the fishing industry and therefore expensive to convert to protected areas. Instead, limited conservation resources were better spent acquiring a higher number of reefs that were medium to high risk, given the low probability that all will be lost.

ROI Where Costs Are Unknown

All the examples just described depend on knowing the expected investment or cost associated with each alternative being considered. These data, or surrogates for them, cannot always be reliably estimated. This is often the case with

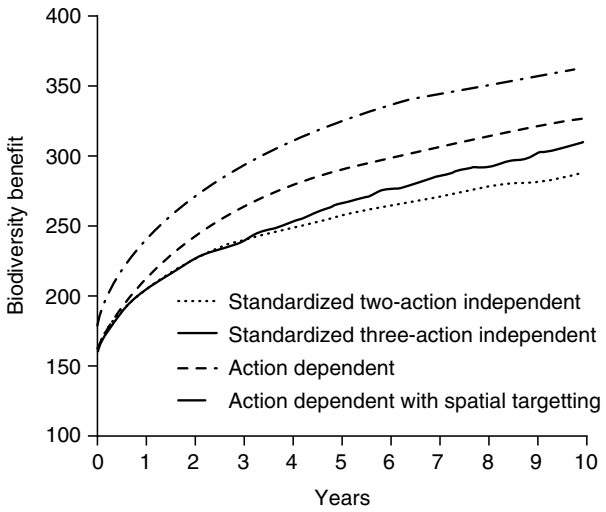


Figure 7 Comparison of the relative performance of ROI analysis recognizing the dependency between actions and their returns. Reproduced from Evans MC, Possingham HP, and Wilson KA (2011) What to do in the face of multiple threats? Incorporating dependencies within a return on investment framework for conservation. *Diversity and Distributions* 17: 437–450, with permission from Wiley.

community-based protected-area initiatives, where the investment cost is largely born by the community that is sacrificing some of its land or sea and numerous cultural, social, economic, and geographic factors influence its willingness to convert part of the customary land or sea to conservation.

Game *et al.* (2011) describe a prioritization process used by The Nature Conservancy and partners in Choiseul Province, Solomon Islands, to help establish a representative network of community-based protected areas. When a community expresses an interest in establishing a protected area, it is presented with a map of the conservation return across its lands and seas. The community is then asked to choose the piece of land or sea with the highest conservation return where it is willing to establish a protected area. After each decision, the map of conservation return for the whole province is updated to consider complementarity with the current set of protected areas. In this way, each decision obtains the highest ROI possible for that decision without having *a priori* knowledge of the investment side of the equation. This is fundamentally equivalent to an ROI approach to conservation prioritization.

Opposition and Challenges

Despite their compelling intention (more conservation for our dollars) and wide array of potential prioritization applications, ROI methods face a series of cultural and philosophical challenges in conservation. General philosophical opposition exists to all prioritization schemes that weigh benefit, cost and probability of success, based on the belief that everything has value and prioritizing means sanctioning the loss of some pieces (e.g., Jachowski and Kesler, 2009). There is also a belief that prioritization schemes are unable to be responsive to opportunities that are the real driver of conservation actions and therefore unreasonably limit flexibility or are not useful. Neither argument is logically sound (for discussion see Bottrill *et al.*, 2009; Game *et al.*, 2011).

There is also a deeper philosophical distaste for the application of financially driven, business-like approaches to biodiversity conservation (a field that undeniably draws great strength from the beliefs and value systems of those involved and has a far more complicated and existential value system than is present in most businesses). This is a legitimate argument, especially given the clear failure of conventional economic systems during the 2008 global financial crisis. This fundamental debate remains challenging, productive, and relatively unexplored philosophical ground.

ROI methods for conservation also face resistance for a number of more pragmatic reasons. In transferring these approaches from theory to practice, it has been difficult to define return functions that convincingly reflect what people care about. People's belief in the conservation value of a place or strategy are composed of many variables that are not necessarily consistently applied, which makes replicating this in a model extremely difficult. Similarly, the available cost data are often either not robust enough or not considered relevant enough to the decisions at hand (even when the consideration of cost is recognized as critical). A study looking at fine-scale conservation-management costs found them to be poorly correlated with commonly used proxies such as the agricultural value of land (Armsworth *et al.*, 2010). Essentially, it is challenging to capture the nuances of conservation value and opportunity in single-return and investment functions. This challenge has frequently resulted in prioritizations that are biased and opaque but that validate existing beliefs about conservation priority.

Not unique to ROI prioritizations, there is also a technical barrier to its successful implementation. Even the relatively modest mathematical familiarity required for the analyses described here are beyond the expertise of many conservation planners. Consequently, the problem-formulation and optimization methods, although rigorous, are often less trusted (the black-box syndrome) than straightforward approaches that suffer serious methodological deficits (such as ranking and additively combining different variables). However, these same challenges were faced by other systematic decision-support methods such as minimum-set reserve design algorithms (e.g., Marxan) that are now considered standard practice and used around the world by conservation practitioners with widely divergent training. Despite these challenges, current trends in both conservation culture and research suggest that ROI analysis will play an increasing role in future determinations of conservation priority.

Appendix

List of Courses

1. Conservation Biology
2. Applied Ecology
3. Environmental Economics

See also: Biodiversity, Human Well-Being, and Markets. Biofuels and Biodiversity: The Implications of Energy Sprawl. Conserving Biodiversity Outside Protected Areas. Convention on Biological Diversity. Development by Design: Using a Revisionist History to Guide a Sustainable Future. Feeding the World and Protecting Biodiversity. Impact of Ecological Restoration on Ecosystem Services. Land-Use Issues. Loss of Biodiversity, Overview. Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data. Natural Reserves and Preserves. Priority Setting for Biodiversity and Ecosystem Services. Restoration of Animal, Plant, and Microbial Diversity. Restoration of Biodiversity, Overview. Species Distribution Modeling. Valuing Ecosystem Services. Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy

References

Armsworth PR, Cantu-Salazar L, Parnell M, Davies ZG, and Stoneman R (2010) Management costs for small protected areas and economies of scale in habitat conservation. *Biological Conservation* 144: 423–429.

Ball IR, Possingham HP, and Watts ME (2009) Marxan and relatives: Software for spatial conservation prioritization. In: Moilanen A, Wilson KA, and Possingham HP (eds.) *Spatial Conservation Prioritization: Quantitative Methods & Computational Tools*, pp. 185–210. Oxford: Oxford University Press.

Balmford A, Gravestock P, Hockley N, McClean CJ, and Roberts CM (2004) The worldwide costs of marine protected areas. *Proceedings of the National Academy of Sciences of the United States of America* 101: 9694–9697.

Barlow ACD, Greenwood CJ, Ahmad IU, and Smith JLD (2011) Use of an action-selection framework for human–carnivore conflict in the Bangladesh Sundarbans. *Conservation Biology* 24: 1338–1347.

Bode M and Murdoch W (2009) Cost-effective conservation decisions are robust to uncertainty in the species–area relationship. *Proceedings of the National Academy of Sciences of the United States of America* 106: E12.

Bode M, Wilson KA, Brooks TM, *et al.* (2008) Cost-effective global conservation spending is robust to taxonomic group. *Proceedings of the National Academy of Sciences of the United States of America* 105: 6498–6501.

Bode M and Wintle B (2010) How to build an efficient conservation fence. *Conservation Biology* 24: 182–188.

Bottrill MC, Joseph LN, Carwardine J, *et al.* (2009) Finite conservation funds mean triage is unavoidable. *Trends in Ecology & Evolution* 24: 183–184.

Bryant R (1998) Nonprofit profit (By integrating conservation and economic development Ecotrust effects an environmental return on investment). *Landscape Architecture* 88: 52.

Burgess ND, D'Amico Hales J, Ricketts TH, and Dinerstein E (2006) Factoring species, non-species values and threats into biodiversity prioritisation across the ecoregions of Africa and its islands. *Biological Conservation* 127: 383–401.

Carwardine J, Wilson KA, Ceballos G, *et al.* (2008a) Cost-effective priorities for global mammal conservation. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11446–11450.

Carwardine J, Wilson KA, Watts M, Etter A, Klein C, and Possingham HP (2008b) Avoiding costly conservation mistakes: The importance of defining actions and costs in spatial priority setting. *PLoS One* 3: e2586.

Costello C and Polasky S (2004) Dynamic reserve site selection. *Resource and Energy Economics* 26: 157–174.

Edejer TT, Baltussen R, Adam T, *et al.* (eds.) (2005) *Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis*. Geneva: World Health Organisation.

Evans MC, Possingham HP, and Wilson KA (2011) What to do in the face of multiple threats? Incorporating dependencies within a return on investment framework for conservation. *Diversity and Distributions* 17: 437–450.

Faith DP (2008) Threatened species and the potential loss of phylogenetic diversity: Conservation scenarios based on estimated extinction probabilities and phylogenetic risk analysis. *Conservation Biology* 22: 1461–1470.

Game ET and Grantham HS (2008) *Marxan User Manual; for Marxan Version 1.8.10*. Brisbane: The University of Queensland and Pacific Marine Analysis and Research Association.

Game ET, Lipsett-Moore G, Hamilton R, *et al.* (2011) Informed opportunism for conservation planning in the Solomon Islands. *Conservation Letters* 4: 38–46.

Game ET, Watts M, Wooldridge S, and Possingham H (2008) Planning for persistence in marine reserves: A question of catastrophic importance. *Ecological Applications* 18: 670–680.

Goldman RL, Benitez S, Calvache A, and Ramos A (2010) *Water Funds: Protecting Watersheds for Nature and People*. Arlington, VA: The Nature Conservancy.

Goldstein JH, Pejchar L, and Daily GC (2008) Using return-on-investment to guide restoration: A case study from Hawaii. *Conservation Letters* 1: 236–243.

Hajkowicz S (2009) The evolution of Australia's natural resource management programs: Towards improved targeting and evaluation of investments. *Land Use Policy* 26: 471–478.

Halpern BS, Walbridge S, Selkoe KA, *et al.* (2008) A global map of human impact on marine ecosystems. *Science* 319: 948–952.

Hoekstra JM, Boucher TM, Ricketts TH, and Roberts C (2005) Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters* 8: 23–29.

Jachowski DS and Kesler DC (2009) Allowing extinction: Should we let species go? *Trends in Ecology & Evolution* 24: 180.

Joseph LN, Maloney RF, and Possingham HP (2009) Optimal allocation of resources among threatend species: A project prioritization protocol. *Conservation Biology* 23: 328–338.

Kareiva P, Tallis H, Ricketts TH, Daily GC, and Polasky S (eds.) (2011) *Natural Capital: Theory and Practice of Mapping Ecosystem Services*. Oxford: Oxford University Press.

Klein CJ, Ban NC, Halpern BS, *et al.* (2010) Prioritizing land and sea conservation investments to protect coral reefs. *PLoS One* 5: e12431.

Laycock H, Moran D, Smart J, Raffaelli D, and White P (2009) Evaluating the cost-effectiveness of conservation: The UK Biodiversity Action Plan. *Biological Conservation* 142: 3120–3127.

Lipsett-Moore G, Game ET, Peterson N, *et al.* (2010) *Interim National Terrestrial Conservation Assessment for Papua New Guinea: Protecting Biodiversity in a Changing Climate*. Brisbane, Australia: The Nature Conservancy.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.

Moilanen A and Kujala H (2008) *Zonation Spatial Conservation Planning Framework and Software V. 2.0, User Manual*. Helsinki: University of Helsinki.

Moore J, Balmford A, Allnutt TF, and Burgess ND (2004) Integrating costs into conservation planning across Africa. *Biological Conservation* 117: 343–350.

Murdoch W, Polasky S, Wilson KA, Possingham HP, Kareiva P, and Shaw R (2007) Maximizing return on investment in conservation. *Biological Conservation* 139: 375–388.

Murdoch W, Ranganathan J, Polasky S, and Regetz J (2011) Using return on investment to maximize conservation effectiveness in Argentine grasslands. *Proceedings of the National Academy of Sciences of the United States of America* 107: 20855–20862.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Naidoo R, Balmford A, Ferraro P, Polasky S, Ricketts TH, and Rouget M (2006) Integrating economic costs into conservation planning. *Trends in Ecology & Evolution* 21: 681–687.

Naidoo R and Iwamura T (2007) Global-scale mapping of economic benefits from agricultural lands: Implications for conservation priorities. *Biological Conservation* 140: 40–49.

O'Connor C, Marvier M, and Kareiva P (2003) Biological vs. social, economic and political priority-setting in conservation. *Ecology Letters* 6: 706–711.

Smith AB (2010) Caution with curves: Caveats for using the species–area relationship in conservation. *Biological Conservation* 143: 555–564.

Tallis H and Polasky S (2009) Mapping and valuing ecosystem services as an approach for conservation and natural-resource management. *Annals of the New York Academy of Sciences* 1162: 265–283.

TNC (2007) *Conservation Action Planning Handbook*. Arlington, VA: The Nature Conservancy.

Underwood EC, Klausmeyer KR, Morrison SA, Bode M, and Shaw MR (2009) Evaluating conservation spending for species return: A retrospective analysis in California. *Conservation Letters* 2: 130–137.

Underwood EC, Shaw MR, Wilson KA, *et al.* (2008) Protecting biodiversity when money matters: Maximizing return on investment. *PLoS One* 3.

Wilson KA, McBride MF, Bode M, and Possingham HP (2006) Prioritizing global conservation efforts. *Nature* 440: 337–340.

Wilson KA, Underwood EC, Morrison SA, *et al.* (2007) Conserving biodiversity efficiently: What to do, where, and when. *PLoS Biology* 5: 1850–1861.

Impact of Ecological Restoration on Ecosystem Services

Holly P Jones, University of California, Santa Cruz, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Active restoration Actively assisting a system in recovery beyond removing the disturbance.

Ecosystem engineers Species that modify, maintain, or create habitats and modify resource supplies through either their physical structure or by transforming living materials from one state to another.

Ecosystem services Those products or services that nature provides society.

Keystone species Species that have a disproportionate effect on ecosystems relative to their biomass.

Nonuse ecosystem services Benefits which are derived from ecosystems without using them (e.g., existence values – the benefit we derive from knowing something exists).

Novel ecosystems Systems that have conditions and combinations of organisms never seen in the historical past.

Passive recovery Allowing a system to recover on its own without intervention beyond removing the disturbance.

Restoration The process of assisting in the recovery of damaged, degraded, or destroyed ecosystems.

Ecological Restoration

As resource use and human populations grow (Haberl *et al.*, 2004; United Nations, 2004), restoration of degraded ecosystems is becoming one of the most important tools in conservation managers' toolboxes (Dobson *et al.*, 1997; Kareiva *et al.*, 2007). Restoration is critical not only to maintain habitat for biodiversity, but also to provide the ecosystem services necessary for human survival (Dobson *et al.*, 1997; Hobbs and Harris, 2001; Foley *et al.*, 2005; MEA (Millennium Ecosystem Assessment), 2005; Kareiva *et al.*, 2007; Benayas *et al.*, 2009). However, restoration can be costly, complex, and labor-intensive and is not always successful, which has traditionally led to reluctance in pursuing large-scale restoration projects (Hobbs and Harris, 2001; Kumar, 2004; van Andel and Aronson, 2006; Jentsch, 2007). If restoration is to succeed in the context of limited funding and capacity, calculating the value of restoration projects is absolutely critical. As our ability to value ecosystem services continues to develop, the return on investment of restoration projects increasingly looks favorable, even for large-scale restoration projects (TEEB, 2010). This article will discuss the concepts of ecological restoration, ecosystem services, and how they are being merged to show society what it stands to gain from restoring earth's damaged ecosystems.

What is Ecological Restoration?

Ecological restoration has been in practice for centuries. For example, Native Americans in the Washington Cascades lit periodic forest fires to sustain natural ecosystem services, such as food crops and game animal forage (Everett *et al.*, 2000). Despite centuries of restoration projects, restoration ecology was not formally recognized as its own scientific discipline until the 1990s (Allen *et al.*, 1997). Until then, grass-roots restoration projects were common throughout the world in the form of invasive species removal (Howald *et al.*, 2007), native species introductions (Towns and Ferreira, 2001), and native plant revegetation (D'Antonio *et al.*, 1992). However, most of these projects did not have a common theoretical framework and consisted of a hodgepodge of strategies, goals, successes, and failures (Hobbs and Norton, 1996).

Starting in the 1990s, this incoherence began to change (Hobbs and Norton, 1996; Allen *et al.*, 1997). The National Science Foundation began to expressly request proposals for restoration ecology, and scientists began devising a theoretical framework from which to practice restoration. Current restoration ecology is not like that of the past, because it is now based around sound ecological principals and it uses theoretical frameworks to design experiments. That said, it is still developing and has much progress ahead of it. In fact, the first textbook for use in classroom teachings was only recently published (van Andel and Aronson, 2006).

The Society for Ecological Restoration defines ecological restoration as the process of assisting the recovery of damaged, degraded, or destroyed ecosystems (SER, 2002). This can include removing the original cause of damage and allowing the system to recover on its own (passive recovery), or actively assisting the system in recovery (active restoration). It is important to distinguish passive recovery and active restoration because each has much different costs, capacity needs, and methodology. If systems can recover passively after the agent of damage has been removed, less money, staff, and commitment will be necessary to repair that system. However, ecosystems that need active restoration to recover will require much more time, money, and effort.

Restoration ecologists have begun to design a framework from which to plan restoration projects to remedy the ubiquitous restoration projects built with little ecological knowledge and oftentimes unclear or unattainable restoration goals. It is now widely accepted that the first step of a restoration project is to define the goals of the project clearly. Project goals of "restoring native vegetation" or "recovering a preperturbation state" are unacceptable; rather, specific goals of restoring species from X to Y abundance or recovering 85% of natural floodplains on a particular watershed are strived for. The second requirement for a restoration goal is that it should

be realistic. With the belief that ecosystems are dynamic entities (Pickett and Parker, 1994), restoration ecologists realize that restoring something to a state in the past may not be possible, and rather, the goal should focus on societal expectations for the future (van Andel and Aronson, 2006). The third requirement of a restoration goal is that progress toward the goal should be measured explicitly.

One thing that distinguishes restoration ecology from most other hard sciences is its explicit inclusion of societal values in its goals. People set restoration goals, meaning goals are often selected according to subjective views of what society needs or wants from a particular ecosystem. The use of the terms like "ecosystem health" and "ecosystem integrity," as definitions for restored systems have been particularly controversial. Ecosystem health is defined as "the state or condition of an ecosystem in which its dynamic attributes are expressed within the normal ranges of activity relative to its ecological state of development" (van Andel and Aronson, 2006), while ecosystem integrity is "the state or condition of an ecosystem that displays the biodiversity characteristic of the reference, such as species composition and community structure, and is fully capable of sustaining normal ecosystem functioning" (SER, 2002). Both definitions take the concepts of ecosystem perturbation, resistance, and resilience and add an aspect of societal expectations (van Andel and Aronson, 2006). Integrating society into ecological terms outrages some ecologists (Lancaster, 2000; Davis and Slobodkin, 2004), but restoration ecology, unlike conventional sciences, aims to explicitly incorporate societal expectations into restoration programs. One promising way of incorporating society into restoration goals is using ecosystem services to calculate the value of restoring nature to meet human needs.

Restoration Targets

One significant issue for restoration ecology is the definition of a preperturbation state, which, unless resolved, can leave enduring confusion and ambiguous restoration targets. Some restoration ecologists have argued against the use of historical reference data to attain restoration goals for several reasons. First, any attempt to restore systems perturbed by human activity must address the question of how far back in history to go (Diamond, 1987; Hobbs, 2005). In addition, historical and paleontological records are often lacking or incomplete. Finally, using historical reference sites may be very difficult or just impossible because under climate change, shifts of carbon cycles, climate regimes, sea level, and climatic events will change the biophysical make-up of ecosystems to make historical systems difficult or impossible to recreate (Hobbs *et al.*, 2006).

Many restoration projects have moved away from the idea of restoring back to "natural" or prehuman ecosystems, and instead, use existing reference systems that have not been exposed to the perturbation (contemporaneous reference sites) as restoration targets (Harris and Diggelen, 2006). The advantages of modern reference systems are that they reflect contemporary climate regimes, have shown resilience to human activity, and have properties that can be explicitly measured to evaluate restoration progress. Another possible target for restoration is the functional target (Harris and Diggelen, 2006). This is the idea of using ecosystem function parameters such as primary productivity, energy flow, and perturbation regime as restoration targets. However, these properties cannot be easily manipulated and thus make difficult goals to achieve without considering what community type would approach these targets. Finally, some projects aim to restore a suite of ecosystem services to society (Harris and Diggelen, 2006). Using ecosystem service targets can help mangers gain society's support for restoration by detailing the benefits of restoration in a currency that everyone understands.

Restoring Entire Ecosystems, a Few Ecosystem Drivers, or Ecosystem Services

There has been significant debate over whether to restore entire ecosystems or just a few ecosystem drivers (Goldstein, 1999; Risser, 1999; Walker, 1999), and it is complicated by divergent views about the nature of communities and ecosystems. Ecosystems can be viewed as stocks and flows of their biotic and abiotic components, which is an entire ecosystem view (Odum, 1969). An alternative view is that ecosystems are composed of communities of species pitted against one another in outright Darwinian evolutionary struggles, and only a few of these influence the entire community, which is an ecosystem driver view (Oksanen, 1988). Both perspectives have advantages and disadvantages for restoration ecology. A restoration project that focuses on ecosystem drivers emphasizes key components of ecosystems, may enhance biodiversity, and can require less time and cost to carry out than whole-ecosystem restorations. However, this focus on single species or small groups of species fails to recognize the importance of processes at the landscape scale, assumes untested links between components of communities, may inadvertently damage species not targeted in restoration, and may divert attention from other important species (Ehrenfeld, 2000).

The ecosystem view, however, recognizes the importance of landscape-scale processes to species persistence, incorporates the view that systems are dynamic, and can help to encourage participation from a variety of agencies and stakeholders. However, attempting to restore entire ecosystems is complicated because ecosystem functions are difficult to define or measure. Even with adequate definitions, functions sometimes do not correlate with each other and often operate on different temporal and spatial scales. This makes realistic restoration goals difficult to set.

The third concept is the ecosystem services view (Ehrenfeld, 2000). Projects with goals of restoring ecosystem services have clearly identified goals and tend to generate public support and funding. However, such projects may have difficult-to-measure services, are contingent upon willingness-to-pay, must overcome trade-offs between equally desirable ecosystem services, and could be compromised if changes in society, technology, or the economy devalue the services that such projects are targeting. Ecosystem services are difficult to distinguish from the ecosystem view because restoring ecosystem services requires restoring the same ecosystem functions that are the targets of whole-ecosystem restoration projects. As such, the ecosystem service viewpoint can be seen as a strategy

to justify restoration rather than a completely different viewpoint.

Each of these views for setting restoration goals has advantages and disadvantages. No one paradigm is correct for restoration – instead, each one should be considered in turn before deciding which paradigm is appropriate for the focal restoration project. In many cases, more than one paradigm will apply. For example, restoring an entire watershed must focus on landscape-scale processes such as hydrologic flows, river diversions, and runoff control. Ecosystem driver restoration goals are unrealistic in this case because its scope is well beyond restoring one or a few species. However, whole-ecosystem goals and ecosystem service goals are both appropriate. Whole-ecosystem frameworks can be used to help define the restoration targets, while ecosystem services can be used to help generate public support for the project. Attempting to create universally applicable laws may not be necessary for restoration as it is such a broad, diverse, and complex discipline (Ehrenfeld, 2000). Instead, using common sense principles to decide which framework works for each unique situation will ensure each restoration project has a better chance of reaching the goals that are appropriate in each context. Below are examples of projects that use each framework as the basis for setting restoration goals.

Restoration of Ecosystem Drivers

Some ecosystems seem to be largely driven by one or a few species, in terms of ecosystem structure or function. Keystone species are one example of ecosystem-driving species and they are so named because they are akin to a keystone that supports an entire arch – without the keystone, the entire arch would collapse (Paine, 1995). Keystone species have a disproportionate effect on ecosystems relative to their biomass and when they are affected or removed by a perturbation, the effects on ecosystems can be tremendous (Paine, 1966). Sea otters are a classic example of a keystone species in ocean kelp forests. Otters prefer to consume herbivorous invertebrates such as sea urchins, which are particularly voracious consumers of kelp. When otters are removed from the system (which occurred extensively along the west coast of North America in the 1800s due to otter hunting for pelts), invertebrate populations explode and quickly consume and destroy the kelp forests, which results in the collapse of an entire system – without the kelp, fish and invertebrate species lose their habitat, and abiotic conditions (light, nutrient availability) fundamentally change (Estes and Duggins, 1995). Kelp forests that have been destroyed by uncontrolled herbivorous invertebrates have been termed urchin barrens because all that is left are large populations of urchins. In this way, the loss or reduction of a single species can result in devastating losses to ecosystem services including opportunities for recreation, commercial fishing nursery grounds, and carbon sequestration.

Other species thought to be the key drivers of ecosystem structure or function are termed ecosystem engineers. These species modify, maintain, or create habitats and modify resource supplies through either their physical structure (e.g., trees in forests), or by their transforming living materials from one state to another (e.g., beavers building dams; Jones *et al.*, 1994). Coral reefs are classic examples of ecosystem engineers. Their structure changes the coastal environment and modifies the abundance and distribution of resources (e.g., light, nutrient availability), and protects entire reef systems and coastal ecosystems from water force and wave action (Anderson, 1992; Moberg and Rönnbäck, 2003). This makes perturbations to corals especially dramatic because they alter entire ecosystems as opposed to just one species. Loss of corals reduces the ecosystem services that corals provide – commercially important fish and invertebrate species lose their habitat, coastal communities are less protected from storm surges and wave action, and recreational and tourist opportunities are diminished.

Restoring systems that have lost their keystone species or ecosystem engineers usually consists of restoring the species to its historical location. For example, over the past century, riparian ecosystems in much of the western US have been lost from large-scale recruitment failures of key riparian tree species like native cottonwoods and aspens. Mounting evidence has suggested that these failures were due to the system losing its top predators – wolves – from human hunting pressures in the 1920s (Vander Zanden *et al.*, 2006). The loss of wolves led elk to change their foraging behavior and forage for cottonwoods and aspens without fear of predation from wolves (Ripple *et al.*, 2001; Beschta, 2003). Wolves were reintroduced in the mid-1990s and their regulation of elk populations and foraging behavior has allowed riparian vegetation to recover. This recovery is expected to improve bird nesting habitat, stabilize stream banks, strengthen ecosystem linkages, and moderate water temperatures (Osborne and Kovacic, 1993; Berger *et al.*, 2001). In this case, reestablishment of one keystone species has carried over to the entire ecosystem. It illustrates the relative simplicity of restoring systems that have lost only one of their key species.

Whole-ecosystem Restoration

Restoring entire ecosystems is an ambitious undertaking. Whole-ecosystem restoration involves repairing or restoring most if not all of the functions of an ecosystem following a perturbation. It most often refers to restoration projects at the landscape scale and is arguably one of the most efficient ways to bolster ecosystem services because restoring such large-scale processes results in provisioning of a wealth of ecosystem services at vast scales. For example, efforts to restore the Mississippi–Ohio–Missouri river basin following eutrophication from farm runoff is projected to reduce hypoxia in the Gulf of Mexico, improve water quality for commercially important fish and invertebrate species, provide large swaths of riparian floodplains to help regulate floodwaters for people and croplands, improve water quality and quantity for local residents, and provide habitat for local wildlife (Mitsch and Day, 2006). This single restoration action, carried out by various organizations and covering multiple states, and millions of hectares, has the potential for a much larger ecosystem service benefit than smaller-scale restorations aimed at one or a few species or habitats. It is also significantly more complex because it covers more habitats, ecosystem processes, species, and jurisdictions than smaller-scale restorations. As such, it is a good illustration of both the immense benefits and challenges of attempting to restore entire ecosystems.

Ecosystem Service Restoration

Since the seminal works defining ecosystem services – also known as natural capital (Daily, 1997) – and valuing them (Costanza *et al.*, 1997), restoration ecology has been increasingly focused on restoring not just species or ecosystems, but also ecosystem service values. Ecosystem services are increasingly called for as a means to justify restoration, set goals that are more realistic than historical targets given continuing climate and global change, and to illustrate the connection between people and nature (Cairns, 2000; Hobbs and Cramer, 2008; Chazdon *et al.*, 2009; Jackson and Hobbs, 2009). Ecosystem services are those products or services that nature provides society. There are many different ways to categorize ecosystem services, but most are blocked into four main types (MEA, 2005): provisioning – those that provide a product such as croplands provide food; regulating – those that provide regulation of function such as coastal wetlands provide wave attenuation for coastal populations faced with tropical storms (Alongi, 2008; Das and Vincent, 2009; Shepard *et al.*, 2011); support – those that support other ecosystem services such as primary production; and cultural – those that provide cultural, aesthetic, or spiritual benefits such as national parks provide recreational opportunities. Ecosystem services are covered extensively elsewhere in this volume, so they are discussed only in reference to restoration in this article.

Ecosystems naturally provide some of these benefits whether they are degraded or intact, but intact systems have higher potential for ecosystem services than do degraded ones (Chazdon, 2008; Benayas *et al.*, 2009). This is where ecological restoration and ecosystem services combine – restoration can help ecosystems regain their function and thereby increase the potential for ecosystems to provide services to humankind. Restoring ecosystem services, then, requires restoring the underlying ecosystem functions that relate to the desired services. Each service relies on an underlying ecosystem process for its production. For example, **Table 1** lists some common ecosystem services, ecosystems, and what ecosystem process is necessary for each service. As the table illustrates, restoring ecosystem services is no different than restoring key ecosystem functions. Instead, using ecosystem services as a target for restoration is a way to communicate how society will benefit from restoring key ecosystem functions, generate public support, and generate funding mechanisms for restoration projects.

It is still difficult to glean examples of restoration projects whose specific goals were to recover ecosystem services. This stems from issues inherent in defining the functions needed to restore the service, an incomplete understanding of ecosystem dynamics, and from the disconnect between ecological and economic understanding of ecosystem service restoration (Kremen, 2005; Palmer and Filoso, 2009; Wegner and Pascual, 2011). Instead, it is more common to use ecosystem service targets as a call to restore. For example, a model was recently developed for The Drakensberg Mountains of South Africa to calculate the economic benefits of restoring degraded grasslands and riparian zones as ways to increase the low base water flows during the winter months (Mander *et al.*, 2008). The restoration would cost around $5.1 million over 7 years, while the ecosystem service benefits of additional water, sediment reduction, and carbon sequestration are valued at $8.4 million per year (Blignaut *et al.*, 2010). The study and many other similar studies (Loomis *et al.*, 2000; Tong *et al.*, 2007; Yu *et al.*, 2009) provide the motivation behind large-scale restoration projects that attempt to align nature with societal goals.

A second way that ecosystem services are incorporated into restoration is through hindsight. Ecosystems throughout the world have been restored, but the concept of ecosystem services is so new that it is more common to calculate the ecosystem service benefits of past restoration projects, or to frame past projects in terms of ecosystem services now that we know how effective ecosystem services can be in gathering public support for restoration (e.g., Aronson *et al.*, 2007). For example, the island-scale ecological restoration on Tiritiri Matangi Island, New Zealand, was conceived as a way to restore native flora and fauna, but the ecosystem service benefits of the restoration project are now being touted as some of the greatest benefits of restoration (Craig and Vesely, 2007). As the practice of restoration catches up with the theory of ecosystem services, we will begin to see many restoration projects with specific goals of restoring ecosystem services.

Other Key Considerations for Ecological Restoration

Novel Ecosystems

The idea that there are novel or synthetic ecosystems – those that have conditions and combinations of organisms never before seen – goes back as far as Odum (1962). Recent work

Table 1 Ten of the most critical ecosystem services for human well-being and the associated ecosystem functions that are critical in providing those services. While ecological restoration can be framed in terms of the ecosystem services restoration provides, it is the underlying ecosystem functions which must be restored to provide those services

Ecosystem service	*Corresponding ecosystem processes*
Air purification	Plant respiration, photosynthesis, toxin filtration, nutrient filtration, nutrient cycling
Carbon sequestration	Carbon storage, carbon cycling, photosynthesis
Climate regulation	Carbon storage, carbon cycling, photosynthesis, light absorption, light reflection, evaporation, transpiration
Coastal defense from storms, sea-level rise	Wave attenuation, water absorption, water retention, water storage, land stabilization
Crop pollination	Pollinator population dynamics
Fertile soils	Nutrient cycling, nutrient mineralization, water absorption, water retention, nutrient fixing
Flood protection	Floodwater absorption, floodwater retention, wave attenuation, water storage
Freshwater provisioning	Water absorption, water retention, water storage
Timber	Forest net primary production, photosynthesis
Water purification	Nutrient filtration, toxin filtration, nutrient retention/cycling, water storage

has termed such systems novel ecosystems (also known as "emerging" or "no-analog" systems) and has defined them as those "containing new combinations of species that arise through human action, environmental change, and the impacts of the deliberate and inadvertent introduction of species from other parts of the world" (Hobbs *et al.*, 2009; Marris, 2011). Novel ecosystems are an ambiguous concept because depending on the timeframe of interest, anything could be considered novel. However, the main proponents of using novel ecosystems to help guide restoration efforts suggest that the accelerated pace and breakdown of biogeographic barriers sets current novel ecosystems apart from the past in terms of the increased rate of novel environment appearance, novel species combinations, and altered ecosystem functioning (Hobbs *et al.*, 2009). The utility of terming systems novel seems to move restoration targets away from historical states and instead argues for ecosystem service and ecosystem function restoration goals (Jackson and Hobbs, 2009). This is laudable but much of the debate on using historical reference sites has already been settled and so the introduction of this new terminology is still somewhat puzzling.

Novel ecosystems have arbitrary definitions and could therefore represent a range of scenarios from a severely perturbed system to one that has been completely replaced by new species. Another component of the novel system framework is the hybrid ecosystem, which also has an arbitrary definition – somewhere between a novel and an historical system – and cannot be clearly distinguished from just perturbed or novel ecosystems. Indeed, a large perturbation such as logging can result in a complete replacement of the previous vegetation with new, early-successional and possibly invasive grasses. Is this then a novel ecosystem because it is different from the historical past and likely made up of new species combinations? Or is it merely a perturbed system in the beginning stages of recovery? Their arbitrary definitions make novel systems impossible to distinguish from just severely perturbed systems and therefore cause difficulty in identifying the extent of their existence, illustrating their impact, and using them as potential restoration targets. Despite these difficulties, the existence of novel ecosystems is certainly another good argument against using historical references in restoration, but focusing our energy on ambiguously defined and difficult-to-identify ecosystems could also take away from sorely needed advancements in using ecosystem services to calculate return on restoration investments.

Restoration and Global Change

It is important to recognize that ecosystem repair from various perturbations is happening against a backdrop of global change that has accelerated at unprecedented rates since the industrial revolution (Baines and Folland, 2007; IPCC, 2007; Maslanik *et al.*, 2007; Belkin, 2009; Gregory *et al.*, 2009). From sea-level rise to heightened hurricane activity, longer and more frequent droughts and floods, and acidification of the world's oceans, ecosystems are undergoing rapid change (Allison *et al.*, 2009; Füssel, 2009). This presents an interesting conundrum: How do we choose restoration targets when much of the world's ecosystem processes are in a state of flux and are likely to continue changing for the foreseeable future? This dynamic and unpredictable climate argues against the use of historical baselines for restoration because they will be increasingly impossible to recreate amidst changing biophysical conditions (Harris *et al.*, 2006). Indeed, many have suggested that restoration of processes, rather than historical conditions, will give restoration the best chances of success in a changing world (Harris *et al.*, 2006). In a changing world, our restoration goals are likely to include ecosystems that were never before seen in historical times because we are predicted to experience climate shifts never before realized in such a short time span. The interaction between ecological restoration and global change is an understudied area that will be of rapidly growing interest as we watch systems change before our eyes.

Climate change also presents an interesting challenge to legislation that protects critical habitat for species of concern like the Endangered Species Act. Changing ecosystem processes may mean that critical habitats will shift according to species' biophysical envelopes, and so, protected habitats will no longer be suitable for the species the protection was created for (Harris *et al.*, 2006). Restoration may provide a critical answer to protecting these species – it may be necessary to restore habitats and create shelter for species that will need to migrate given climate change (Harris *et al.*, 2006).

Ecological Restoration and Ecosystem Services

Ecosystem Restoration Potential

Human population now exceeds seven billion people and this number is projected to grow to eight billion by 2050 (United Nations, 2004). As our population grows, so too do our technological demands, thirst for resources, and exploitation of nature to meet our expanding needs. We have fundamentally altered most ecosystem processes, leaving few areas on Earth untouched by people in some way (Kareiva and Marvier, 2011). Humans have been domesticating nature to meet our needs for centuries (Kareiva *et al.*, 2007), but the rate of ecosystem alteration has spiked since the industrial revolution. As we continue to use nature's capital to meet our needs, we also threaten to destroy the very resources that we depend on for survival. We have transformed around half of Earth's surface area to grazing or croplands (MEA, 2005). More than 50% of the world's forests have been lost through this extreme transformation (MEA, 2005). Although providing food for the masses is a necessary service that ecosystems provide, it is often done at the expense of other ecosystem services such as water quality and quantity regulation, climate regulation, and carbon sequestration. Our insatiable thirst for resources and to control nature has driven the largest mammals on each continent to extinction or severe depletion, fundamentally altering food webs across the globe (Woodroffe, 2000; Lyons *et al.*, 2004). No marine ecosystem on earth is untouched by people and almost half of the world's oceans face multiple threats (Halpern *et al.*, 2008). As such, there is substantial concern that we are leaving a legacy of extinction, biodiversity loss, and unsustainable exploitation of nature to future generations.

Humans are arguably better off now than in previous generations. We are able to feed the masses, control predators, alter natural disturbance regimes, and build infrastructure to help

protect us from storms. So what has our domination over nature done to the world's natural capital? The answer is complex because there are trade-offs between different ecosystem services (Kareiva *et al.*, 2007). For example, food provisioning is a service that nature provides but as we replace ecosystems with cropland, other ecosystem services are diminished. Maximizing crop production requires the input of artificial fertilizer, which has caused massive nutrient leaching into waterways and has led to eutrophication of rivers, lakes, estuaries, and oceans. This has led to toxic algal blooms and has created dead zones around the world (Galloway *et al.*, 2003). Similarly, we rely on timber for our furniture, shelter, paper, and other household products. But logging and deforestation has caused soil erosion, which reduces forests' capacity to absorb floodwaters, reduces water quality, and can alter climate regulation (Daily, 1997; Sweeney *et al.*, 2004; Dobson *et al.*, 2006). We use fossil fuels to travel, heat our homes, produce food, and access water and electricity, but the by-products of this energy demand are increased greenhouse gases and shifting climate regimes that ultimately threaten our very existence (IPCC, 2007). Our marine fisheries and coastal development have depleted 91% of ecologically and economically important coastal marine species to less than 50% of their historical abundances (Worm *et al.*, 2006) and destroyed more than 50% of the world's mangroves (World Resources Institute, 1996). This severe biodiversity depletion has led to a 33% reduction of fish stocks, a 69% reduction in nursery habitat for marine organisms, and a 63% reduction in water filtration ecosystem services (Worm *et al.*, 2006). So while we have done generally well at taming nature to meet our needs, it has come at a price to ecosystems and the other important services they provide. This price threatens to impact our livelihoods, standard of living, ability to withstand environmental perturbations, and at the most basic level, our persistence. All of these threats may not be realized in our lifetime, but they are very real consequences that will be passed on to future generations.

At the same time, there is cause for some optimism. Ecological restoration of ecosystem services offers great promise for restoring the balance between society and nature and to pass on a more hopeful legacy to future generations. As E.O. Wilson so eloquently put it, "Here is the means to end the great extinction spasm. The next century will, I believe, be the era of restoration in ecology." Combining ecological restoration with ecosystem services can be challenging because they are both such nascent disciplines, there are many trade-offs to study and account for, and the two disciplines (restoration and economics) have so often been at odds with each other in the past. However, the imperative is clear: No place on Earth is untouched by people and so restoring ecosystem services to ensure our future survival will be a key focus of ecological restoration moving forward (Kareiva and Marvier, 2011). Restoring ecosystem services will be one of the most efficient ways to repair ecosystems and ensure human prosperity together in one action.

Bolstering Ecosystem Services as a By-Product of Restoration

Many restoration projects that address ecosystem services view them as a co-benefit of restoration rather than a goal of restoration. This is because restoration was historically justified as a moral imperative (Leopold, 1949). With the realization that moral arguments are not as effective as economic ones, restoration ecologists have increasingly looked at past restoration projects to show how they have benefited society. One of the most comprehensive studies to date illustrating the effects of ecosystem restoration on ecosystem services was a meta-analysis that compared ecosystem services for degraded, restored, and unperturbed reference ecosystems (Benayas *et al.*, 2009). The study showed that restoration increased ecosystem services by 125% relative to degraded systems, giving much hope for restoration to play a key role in repairing nature's capital. However, ecosystem services were still lower in restored systems versus unperturbed systems, indicating conservation of unperturbed systems will continue to be an important strategy for maintaining ecosystem services.

Many other studies illustrate what society stands to gain from ecological restoration in myriad ecosystems across multiple scales. Establishing marine protected areas to help fish populations recover from exploitation leads to increased fish biomass, more catch per unit effort for fisheries, and increased tourism opportunities (Worm *et al.*, 2006), which provides hope that passive restoration can be an effective tool to repair ocean fisheries while simultaneously providing cultural ecosystem services. Mangrove reforestation around Bangladesh has increased fishery potential, provided 600,000 m^3 in forest products, bolstered livelihoods, and stabilized coastal land (Moberg and Rönnbäck, 2003). The US Mississippi Delta provides $12–47 billion per year in flood and hurricane protection, water supply, water quality, recreation, and fishery ecosystem services (Batker *et al.*, 2010). Moreover, if levees along the Delta were replaced with restored wetlands, there would be an estimated net gain of $62 billion in ecosystem services annually (Batker *et al.*, 2010). Each of these studies illustrates the enormous potential for ecosystem service restoration in diverse ecosystems throughout the globe.

Using Ecosystem Services as Justification for Restoration

One of the biggest utilities of ecosystem services in ecological restoration is as a justification to carry out restoration projects. In a resource and capacity-limited society, it is increasingly important that restoration projects make the case to society on their worth. Monetary justification is in sharp contrast to the moral justifications Aldo Leopold so vividly articulated in his Land Ethic (Leopold, 1949). While Leopold might disagree with the means, he would undoubtedly approve of the outcome of using ecosystem services to justify restoration because it shows society how connected to nature people are and speaks to society in a currency that everyone understands and appreciates. This is a modern way to help people understand their dependence on nature and bring even those living in industrialized cities (who use the majority of ecosystem services) a little bit closer to the land.

The overwhelming majority of projects that interweave restoration with ecosystem services are in the developing stages and use ecosystem services to calculate the projected benefits of the proposed restoration (e.g., Loomis *et al.*, 2000; Tong *et al.*, 2007; Yu *et al.*, 2009; Batker *et al.*, 2010). For example, one study was done on the benefits of restoring the historical

floodplains along the lower Danube River in Europe. Historical floodplains were converted to dykes for agriculture, aquaculture, and forestry, which reduced the capacity of the floodplain to absorb floodwaters and has increased flood peaks and spurred interest in restoration. The study showed that restored floodplains would provide $700 per hectare in flood control, provision of fish, forestry, and animal fodder, nutrient retention, and recreational benefits. In this case, projected ecosystem service benefits were two-fold higher than the cost of restoring this huge swath of land, which makes a strong case for investing in restoration (Ebert *et al.*, 2009). Another study valued coastal wetlands for a single ecosystem service – storm protection – over the entire US. The authors determined wetland value by calculating the amount of storm damage in places with varying degrees of wetland protection and found that a loss of just one hectare of wetlands cost on average $33,000 dollars in hurricane damage. US wetlands currently provide $23.2 billion annually in storm protection and while no restoration costs were calculated, the study concludes that restoring wetlands is an extremely cost-effective strategy (Costanza *et al.*, 2008). This massive ecosystem service value of coastal wetlands would be even higher if additional ecosystem services such as water filtration, toxin filtration, carbon sequestration, nursery habitat for commercially important fish, and recreational and tourism opportunities were included. This underscores the massive potential for large-scale restoration efforts to provide society with tangible benefits and preserve biodiversity in a single action.

Using Ecosystem Services as Explicit Restoration Targets

Billions of dollars are spent annually restoring ecosystems (Enserink, 1999; Zhang *et al.*, 2000), but with little to no consideration of the ecosystem service returns on the restoration investments (but see Goldstein *et al.*, 2008). With so much money invested into restoring ecosystems, it is puzzling that so few restoration actions are prioritized according to the ecosystem services they can provide. As these two rapidly emerging fields continue to develop and expand, there will undoubtedly be growing numbers of restoration projects with specific ecosystem service goals, or projects that are prioritized by the amount of services they are likely to produce. Most restoration projects have a whole suite of ecosystem service benefits (Benayas *et al.*, 2009). However, a key area for theoretical advancement is looking at the trade-offs between ecosystem services that specific restoration actions would entail. For example, reforesting an area that has been used as cropland will take away the food provisioning that the crops provide but may reinstate key water purification, flood protection, and carbon sequestration services. Studying these synergies and how they interact to restore ecosystems and contribute to human well-being will likely be one of the cutting-edge applications of ecosystem service restoration in the future.

The Future of Restoration and Ecosystem Services

Tackling the Biodiversity Crisis

Opinions differ on how best to stem the enormous biodiversity crisis that we face. The traditional conservation approach has been to preserve "natural" ecosystems by excluding people (Mittermeier *et al.*, 1998; Reid, 1998; Holl and Aide, 2011). Inherent in this approach is the idea that people and ecosystems are separate entities – that humans are not part of nature. The growing consensus is that this viewpoint is antiquated, often produces maladapted conservation outcomes that may be difficult to maintain in perpetuity, and encourages conservation of nature at the expense of people (McNeely, 1994; Holling, 1996; Berkes, 2004; Kareiva and Marvier, 2011). While conserving biodiversity will continue to be an important strategy moving forward, doing so at the expense of people may ostracize man from nature and hinder conservation objectives. Restoring the earth's damaged ecosystem services is one way to entwine two important objectives: Protecting and preserving biodiversity and providing the necessary functions and services for human survival. This is a unique way to reconnect people with nature and to simultaneously justify the need for restoring the damage we have inflicted. This is a relatively new way to engage society into conservation objectives and it has much progress ahead of it to ensure that it continues to be an effective conservation strategy.

Even so, this is an exciting time to be a restoration ecologist. Restoration ecology is and will continue to be one of the most important sciences of the twenty-first century because of the widespread need to repair ecosystems globally. Society has placed great hopes on the ability of restoration to repair Earth's damaged ecosystems and ensure continued survival for future generations (Palmer and Filoso, 2009). This is enormous pressure for such a nascent science but we are equal to the challenge. As restoration ecology continues to incorporate ecosystem services into its frameworks, there are important areas of research that will help to ensure that restoration is operational as a science but also meets society's needs.

Key Challenges

Many studies calculate ecosystem service benefits of conservation, restoration, or management projects. But very few of these examples examine the trade-offs and synergies of different ecosystem service benefits and so it remains difficult to predict how different restoration actions will affect ecosystem services as a whole. To understand these trade-offs, we must gain better understandings of how underlying ecosystem processes change with different restoration options and how ecosystem services change as a result (Palmer and Filoso, 2009). Right now, there are few satisfactory ways to value nonuse ecosystem services such as aesthetic values and so cultural ecosystem services are rarely valued (Portney, 1994; Benayas *et al.*, 2009). Even indirect use (nonmarket) services can be valued in so many different ways (Pearce, 1993) that it becomes difficult to determine the accuracy of different valuation studies without intimate knowledge of economic theory. While this will continue to be a challenge for restoring ecosystem services going forward, it is also a prime opportunity to expand our existing theories and develop them to meet society's needs.

The interactions between ecological restoration, ecosystem service valuation, and climate change present enormous

opportunities for further research. Discussions on climate change and ecological restoration are mostly theoretical (Harris *et al.*, 2006), so there is ample opportunity for empirical studies to inform the discussion. Relatively more attention has been given to the effects of climate change on ecosystem services (Schroter *et al.*, 2005), but there is still a dearth of studies that combine restoration, climate change, and ecosystem services. For example, it will be important to prioritize restoration efforts that help people adapt to climate change but it is unclear which restoration efforts will have the greatest adaptation benefits to people (Jones *et al.*, 2012). Restoring coastal wetlands may be one of the most efficient and economical ways to protect people from sea-level rise and tropical storms (Costanza *et al.*, 2008), but more research is needed to analyze the costs, benefits, and trade-offs of such restoration programs. Other areas may not be worth investing in restoring if they will be inundated by sea-level rise or destroyed by changing temperature regimes. Some species or ecosystems may not be resilient to projected climate change and therefore may be precluded from restoration investment, and others may thrive under climate change scenarios so they may be sought out as restoration targets. Empirical research is needed to examine different scenarios so that we can best prioritize our restoration actions in a changing world.

Integrating Restoration Goals with Societal Values

Society is at a crossroads. We can continue to exploit Earth's ecosystems with little thought to the effects it might have on other species or even future generations of our own species. We can unsustainably exploit resources and leave a legacy of extinction and degradation to future generations. Or we can recognize that our actions are not without consequence, take responsibility for those consequences, and begin making the difficult decisions that are necessary to ensure that future generations will have access to the food, clean water, timber, shelter, and other ecosystem services we all enjoy. Undoubtedly we have damaged ecosystems globally, but we do have the ability to repair damage, restore ecosystems, and ensure ecosystems can continue to provide the services that we all depend on for survival. This may require a large investment in restoration projects, a commitment to scientific research, and a change in our perception of our relationship with nature. But even in its infancy, restoration of ecosystem services can show society how much we stand to gain by making these changes. Using ecosystem service valuation to justify ecological restoration is just one way to reconnect people with nature and to start making more sustainable choices. In current society, it is also arguably the most salient one – the time to try to convince humanity that it is our moral duty to live in balance with the Earth has passed; it is now time to frame the imperative to restore differently, so that it speaks to everyone rather than a selected few.

Making the economic case for restoration will be challenging going forward – ecosystem service valuation and ecological restoration still have much progress ahead of them. This is daunting because it is difficult to join two constantly changing disciplines, but it is also promising because they can both be developed alongside one another to leave them poised to address the profound issues that our society faces. Restoration goals must be explicit, achievable, and take into account the varying views of ecosystems, societal expectations, and the dynamic nature of ecosystems. As we increasingly accept that invasive species, climate change, and unprecedented land transformation will produce novel ecosystems and species combinations never before seen, restoration may have to be directed toward some new system as opposed to a historical baseline (Marris, 2011). Restoration ecology is one science where creativity is necessary, along with scientific intellect, to achieve goals. Incorporating ecosystem services into restoration will require even more careful thought, ingenuity, and creativity. As these fields grow and expand, they will continue to push the boundaries of restoration to what might currently seem impossible. This progress will certainly be necessary in the face of an ever-growing human population, constant global change, and continued stress on the ecosystems that we depend on for survival.

Appendix

List of Courses

1. Restoration Ecology
2. Biodiversity Conservation
3. Ecological Economics

See also: Biodiversity and Cultural Ecosystem Services. Priority Setting for Biodiversity and Ecosystem Services. Restoration of Biodiversity, Overview. Valuing Ecosystem Services. Wetland Creation and Restoration

References

Allen EB, Covington WW, and Falk DA (1997) Developing the conceptual basis for restoration ecology. *Restoration Ecology* 5: 275–276.

Allison EH, Perry AL, Badjeck M-C, *et al.* (2009) Vulnerability of national economies to the impacts of climate change on fisheries. *Fish and Fisheries* 10: 173–196.

Alongi DM (2008) Mangrove forests: Resilience, protection from tsunamis, and responses to global climate change. *Estuarine Coastal and Shelf Science* 76: 1–13.

Anderson RA (1992) Diversity of eukaryotic algae. *Biodiversity Conservation* 1: 267–292.

Aronson J, Milton SJ, and Blignaut JN (eds.) (2007) *Business and Practice.* Washington, DC: Island Press.

Baines PG and Folland CK (2007) Evidence for a rapid global climate shift across the late 1960s. *Journal of Climate* 20: 2721–2744.

Batker D, de la Torre I, Costanza R, *et al.* (2010) *Gaining Ground: The Value of Restoring the Mississippi River Delta.* 100 pp. Tacoma, WA: Earth Economics.

Belkin IM (2009) Rapid warming of large marine ecosystems. *Progress in Oceanography* 81: 207–213.

Benayas JMR, Newton AC, Diaz A, and Bullock JM (2009) Enhancement of biodiversity and ecosystem services by ecological restoration: A meta-analysis. *Science* 325: 1121–1124.

Berger J, Stacey PB, Bellis L, and Johnson MP (2001) A mammalian predator–prey imbalance: Grizzly bear and wolf extinction affect avian neotropical migrants. *Ecological Applications* 11: 947–960.

Berkes F (2004) Rethinking community-based conservation. *Conservation Biology* 18: 621–630.
Beschta RL (2003) Cottonwoods, elk, and wolves in the Lamar valley of Yellowstone National Park. *Ecological Applications* 13: 1295–1309.
Blignaut J, Mander M, Schulze R, *et al.* (2010) Restoring and managing natural capital towards fostering economic development: Evidence from the Drakensberg, South Africa. *Ecological Economics* 69: 1313–1323.
Cairns J (2000) Setting ecological restoration goals for technical feasibility and scientific validity. *Ecological Engineering* 15: 171–180.
Chazdon RL (2008) Beyond deforestation: Restoring forests and ecosystem services on degraded lands. *Science* 320: 1458–1460.
Chazdon RL, Peres CA, Dent D, *et al.* (2009) The potential for species conservation in tropical secondary forests. *Conservation Biology* 23: 1406–1417.
Costanza R, d'Arge R, de Groot R, *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature* 387: 253–260.
Costanza R, Perez-Maqueo O, Martinez ML, *et al.* (2008) The value of coastal wetlands for hurricane protection. *AMBIO: A Journal of the Human Environment* 37: 241–248.
Craig J and Vesely E-T (2007) Restoring natural capital reconnects people to their natural heritage: Tiritiri Matangi island, New Zealand. In: Aronson J, Milton SJ, and Blignaut JN (eds.) *Restoring Natural Capital: Science, Business and Practice*, pp. 103–111. Washington, DC: Island Press.
D'Antonio CM, Halvorson WL, and Fenn DB (1992) Restoration of denuded areas and iceplant areas on Santa Barbara Island, Channel Islands National Park: Final report to National Park Service. Davis, CA, USA: Cooperative National Park Resources Studies Unit, University of California at Davis Institute of Ecology.
Daily G (1997) *Nature's Services: Societal Dependence on Natural Ecosystems.* Washington, DC: Island Press.
Das S and Vincent JR (2009) Mangroves protected villages and reduced death toll during Indian super cyclone. *Proceedings of the National Academy of Sciences* 106: 7357–7360.
Davis MA and Slobodkin LB (2004) The science and values of restoration ecology. *Restoration Ecology* 12: 1–3.
Diamond J (1987) Reflections on goals and on the relationship between theory and practice. In: Jordan WR, Gilpin ME, and Aber JD (eds.) *Restoration Ecology. A Synthetic Approach to Research*, pp. 329–336. Cambridge, UK: Cambridge University Press.
Dobson A, Lodge D, Alder J, *et al.* (2006) Habitat loss, trophic collapse, and the decline of ecosystem services. *Ecology* 87: 1915–1924.
Dobson AP, Bradshaw AD, and Baker AJM (1997) Hopes for the future: Restoration ecology and conservation biology. *Science* 277: 515–522.
Ebert S, Hulea O, and Strobel D (2009) Floodplain restoration along the lower Danube: A climate change adaptation case study. *Climate and Development* 1: 212–219.
Ehrenfeld JG (2000) Defining the limits of restoration: The need for realistic goals. *Restoration Ecology* 8: 2–9.
Enserink M (1999) Plan to quench the everglades' thirst. *Science* 285: 180.
Estes JA and Duggins DO (1995) Sea otters and kelp forests in Alaska: Generality and variation in a community ecological paradigm. *Ecological Monographs* 65: 75–100.
Everett RL, Schellhaas R, Keenum D, Spurbeck D, and Ohlson P (2000) Fire history in the Ponderosa pine/Douglas-fir forests on the east slope of the Washington Cascades. *Forest Ecology and Management* 129: 207–225.
Foley JA, DeFries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.
Füssel H-M (2009) An updated assessment of the risks from climate change based on research published since the IPCC fourth assessment report. *Climatic Change* 97: 469–482.
Galloway JN, Aber JD, Erisman JW, *et al.* (2003) The nitrogen cascade. *Bioscience* 53: 341–356.
Goldstein JH, Pejchar L, and Daily GC (2008) Using return-on-investment to guide restoration: A case study from Hawaii. *Conservation Letters* 1: 236–243.
Goldstein PZ (1999) Functional ecosystems and biodiversity buzzwords. *Conservation Biology* 13: 247–255.
Gregory B, Christophe L, and Martin E (2009) Rapid biogeographical plankton shifts in the North Atlantic Ocean. *Global Change Biology* 15: 1790–1803.
Haberl H, Wackernagel M, and Wrbka T (2004) Land use and sustainability indicators. An introduction. *Land Use Policy* 21: 193–198.
Halpern BS, Walbridge S, Selkoe KA, *et al.* (2008) A global map of human impact on marine ecosystems. *Science* 319: 948–952.
Harris JA and Diggelen V (2006) Ecological restoration as a project for global society. In: van Andel J and Aronson J (eds.) *Restoration Ecology*, pp. 3–15. Oxford: UK: Blackwell Publishing.
Harris JA, Hobbs RJ, Higgs E, and Aronson J (2006) Ecological restoration and global climate change. *Restoration Ecology* 14: 170–176.
Hobbs RJ (2005) The future of restoration ecology: Challenges and opportunities. *Restoration Ecology* 13: 239–241.
Hobbs RJ, Arico S, Aronson J, *et al.* (2006) Novel ecosystems: Theoretical and management aspects of the new ecological world order. *Global Ecology and Biogeography* 15: 1–7.
Hobbs RJ and Cramer VA (2008) Restoration ecology: Interventionist approaches for restoring and maintaining ecosystem function in the face of rapid environmental change. *Annual Review of Environment and Resources* 33: 39–61.
Hobbs RJ and Harris JA (2001) Restoration ecology: Repairing the earth's ecosystems in the new millennium. *Restoration Ecology* 9: 239–246.
Hobbs RJ, Higgs E, and Harris JA (2009) Novel ecosystems: Implications for conservation and restoration. *Trends in Ecology & Evolution* 24: 599–605.
Hobbs RJ and Norton DA (1996) Towards a conceptual framework for restoration ecology. *Restoration Ecology* 4: 93–110.
Holl KD and Aide TM (2011) When and where to actively restore ecosystems? *Forest Ecology and Management* 261: 1558–1563.
Holling CS (1996) Surprise for science, resilience for ecosystems, and incentives for people. *Ecological Applications* 6: 733–735.
Howald G, Donlan CJ, Tershy BR, *et al.* (2007) Invasive rodent eradications on islands. *Conservation Biology* 21: 1258–1268.
IPCC (2007) Impacts, adaptation and vulnerability. *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Cambridge University Press.
Jackson ST and Hobbs RJ (2009) Ecological restoration in the light of ecological history. *Science* 325: 567–569.
Jentsch A (2007) The challenge to restore processes in the face of nonlinear dynamics – on the crucial role of disturbance regimes. *Restoration Ecology* 15: 334–339.
Jones CG, Lawton JH, and Shachak M (1994) Organisms as ecosystem engineers. *Oikos* 69: 373–386.
Jones HP, Hole D, and Zavaleta ES (2012) Harnessing nature to help people adapt to climate change. *Nature Climate Change* 2. doi: 10.1038/NCLIMATE1463.
Kareiva P and Marvier MA (2011) *Conservation Science: Balancing the Needs of People and Nature.* Woodbury: Roberts & Co.
Kareiva P, Watts S, McDonald R, and Boucher T (2007) Domesticated nature: Shaping landscapes and ecosystems for human welfare. *Science* 316: 1866–1869.
Kremen C (2005) Managing ecosystem services: What do we need to know about their ecology? *Ecology Letters* 8: 468–479.
Kumar P (2004) Challenges of restoration. *Bioscience* 54: 5.
Lancaster J (2000) The ridiculous notion of assessing ecological health and identifying the useful concepts underneath. *Human and Ecological Risk Assessment* 6: 213–222.
Leopold A (1949) *A Sand County Almanac and Sketches Here and There.* New York, NY: Oxford University Press.
Loomis J, Kent P, Strange L, Fausch K, and Covich A (2000) Measuring the total economic value of restoring ecosystem services in an impaired river basin: Results from a contingent valuation survey. *Ecological Economics* 33: 103–117.
Lyons SK, Smith FA, and Brown JH (2004) Of mice, mastodons and men: Human-mediated extinctions on four continents. *Evolutionary Ecology Research* 6: 339–358.
Mander M, Blignaut J, Schulze R, *et al.* (2008) *Payment for ecosystem services: Developing an ecosystem services trading model for the Mnweni/Cathedral Peak and Eastern Cape Drakensberg areas.* Development Bank of Southern Africa, Department of Water Affairs and Forestry, Department of Environment Affairs and Tourism, Ezemvelo KZN Wildlife.
Marris E (2011) *Rambunctious Garden: Saving Nature in a Post-Wild World.* New York: Bloomsbury.
Maslanik JA, Fowler C, Stroeve J, *et al.* (2007) A younger, thinner arctic ice cover: Increased potential for rapid, extensive sea-ice loss. *Geophysical Research Letters* 34: L24501.
McNeely JA (1994) Protected areas for the 21st century: Working to provide benefits to society. *Biodiversity and Conservation* 3: 390–405.
MEA (Millennium Ecosystem Assessment) (2005) *Ecosystems and Human Well-being: Synthesis.* Washington, DC: Island Press.
Mitsch WJ and Day JJW (2006) Restoration of wetlands in the Mississippi–Ohio–Missouri (MOM) river basin: Experience and needed research. *Ecological Engineering* 26: 55–69.
Mittermeier RA, Myers N, Thomsen JB, Da Fonseca GAB, and Olivieri S (1998) Biodiversity hotspots and major tropical wilderness areas: Approaches to setting conservation priorities. *Conservation Biology* 12: 516–520.

Moberg F and Rönnbäck P (2003) Ecosystem services of the tropical seascape: Interactions, substitutions and restoration. *Ocean & Coastal Management* 46: 27–46.

Odum EP (1969) The strategy of ecosystem development. *Science* 164: 262–270.

Odum HT (1962) Ecological tools and their use. Man and the ecosystem. In: Waggoner PE and Ovington JD (eds.) *Proceedings of the Lockwood Conference on the Suburban Forest and Ecology*, pp. 57–75. The Connecticut Agricultural Experiment Station, Bulletin 652.

Oksanen L (1988) Ecosystem organization mutualism and cybernetics or plain Darwinian struggle for existence? *American Naturalist* 131: 424–444.

Osborne LL and Kovacic DA (1993) Riparian vegetated buffer strips in water-quality restoration and stream management. *Freshwater Biology* 29: 243–258.

Paine RT (1966) Food web complexity and species diversity. *The American Naturalist* 100: 65–75.

Paine RT (1995) A conversation on refining the concept of keystone species. *Conservation Biology* 9: 962–964.

Palmer MA and Filoso S (2009) Restoration of ecosystem services for environmental markets. *Science* 325: 575–576.

Pearce DW (1993) *Economic Values and the Natural World*. London: Earthscan.

Pickett STA and Parker VT (1994) Avoiding the old pitfalls: Opportunities for a new discipline. *Restoration Ecology* 2: 75–79.

Portney PR (1994) The contingent valuation debate: Why economists should care. *The Journal of Economic Perspectives* 8: 3–17.

Reid WV (1998) Biodiversity hotspots. *Trends in Ecology & Evolution* 13: 275–280.

Ripple WJ, Larsen EJ, Renkin RA, and Smith DW (2001) Trophic cascades among wolves, elk and aspen on Yellowstone National Park's northern range. *Biological Conservation* 102: 227–234.

Risser PG (1999) Examining relationships between ecosystem function and biodiversity: Reply to Goldstein. *Conservation Biology* 13: 438–439.

Schroter D, Cramer W, Leemans R, *et al.* (2005) Ecosystem service supply and vulnerability to global change in Europe. *Science* 310: 1333–1337.

SER (Society for Ecological Restoration Science and Policy Working Group) (2002) The SER primer on ecological restoration.

Shepard CC, Crain CM, and Beck MW (2011) The protective role of coastal marshes: A systematic review and meta-analysis. *PLoS ONE* 6: e27374.

Sweeney BW, Bott TL, Jackson JK, *et al.* (2004) Riparian deforestation, stream narrowing, and loss of stream ecosystem services. *Proceedings of the National Academy of Sciences of the United States of America* 101: 14132–14137.

TEEB (2010) The economics of ecosystems and biodiversity: Mainstreaming the economics of nature: A synthesis of the approach, conclusions and recommendations of TEEB.

Tong C, Feagin RA, Lu J, *et al.* (2007) Ecosystem service values and restoration in the urban Sanyang wetland of Wenzhou, China. *Ecological Engineering* 29: 249–258.

Towns DR and Ferreira SM (2001) Conservation of New Zealand lizards (Lacertilia: Scincidae) by translocation of small populations. *Biological Conservation* 98: 211–222.

United Nations (2004) *World Population to 2300*. New York, NY: Department of Economic and Social Affairs.

van Andel J and Aronson J (2006) *Restoration Ecology*. Oxford, UK: Blackwell Publishing.

Vander Zanden MJ, Olden JD, and Gratton C (2006) Food web approaches in restoration ecology. In: Falk DA, Palmer MA, and Zedler JB (eds.) *Foundations of Restoration Ecology*, pp. 165–189. Washington, DC, USA: Island Press.

Walker B (1999) The ecosystem approach to conservation: Reply to Goldstein. *Conservation Biology* 13: 436–437.

Wegner G and Pascual U (2011) Cost-benefit analysis in the context of ecosystem services for human well-being: A multidisciplinary critique. *Global Environmental Change* 21: 492–504.

Woodroffe R (2000) Predators and people: Using human densities to interpret declines of large carnivores. *Animal Conservation* 3: 165–173.

World Resources Institute (1996) *World resources 1996–1997*. The World Resources Institute, UNEP, UNDP, and World Bank.

Worm B, Barbier EB, Beaumont N, *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.

Yu X, Jiang L, Li L, *et al.* (2009) Freshwater management and climate change adaptation: Experiences from the central Yangtze in China. *Climate and Development* 1: 241–248.

Zhang P, Shao G, Zhao G, *et al.* (2000) China's forest policy for the 21st century. *Science* 288: 2135–2136.

Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity

Thomas Kitzberger, INIBIOMA-CONICET and Universidad Nacional del Comahue, Bariloche, Argentina

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acclimation Reversible physiological or morphological changes an organism experiences in response to changing environmental conditions. Such physiological changes enable the organism to tolerate (or acclimatize to) the new environmental conditions.

Atlantic Multidecadal Oscillation (AMO) A mode of variability occurring in the North Atlantic Ocean and which has its principal expression in the sea surface temperature (SST) field. Controversy exists with regard to its amplitude, and in particular, the attribution of sea surface temperature change to natural or anthropogenic causes, especially in tropical Atlantic areas important for hurricane development.

Bottom up effects Refer to controls on the abundance and/or community structure of organisms that derive from supply of resources (light or nutrients for plants, prey organisms for animals) or from physical factors such as temperature of the environment.

Ecological sorting The process whereby organisms colonize and persist in novel environments, use novel resources or form novel associations with other species as a result of the suites of traits that they carry at the time they encounter the novel condition. It provides an important mechanism whereby new species can fit into an existing community without adaptation.

El Niño Southern Oscillation (ENSO) An occasional shift in winds and ocean currents, centered in the South Pacific region, with worldwide consequences for climate and biological systems.

Functional redundancy Ecosystem property by which more than one species performs a given role so that when one species disappears, the loss of the ecosystem's efficiency as a whole is relatively small; however if several species are lost, the system essentially collapses.

Hysteretic response Phenomenon in which the response of a physical system to an external influence depends not only on the present magnitude of that influence but also on the previous history of the system. Hysteretic responses involve different forward and backwards trajectories of change.

Mast fruiting/seeding The production of many seeds/fruits by a plant every two or more years in regional synchrony with other plants of the same species (community-wide mast fruiting: with other plants of same and different species).

Pacific Decadal Oscillation (PDO) A long-term El Niño-like pattern of Pacific climate variability. While the two climate oscillations ENSO and PDO have similar spatial climate fingerprints, they have very different behavior in time: 20th century PDO "events" persisted for 20–30 years, while typical ENSO events persisted for 6–18 months.

Semelparous An organism producing offspring (flowers and fruits in plants) only once in its lifetime before dying.

Tele-connection Causal link between patterns of weather in two locations, or between two atmospheric occurrences, which are very far apart, such the El Niño Southern Oscillation, where unusually high sea surface temperatures off the west coast of South America can be connected with droughts in Indonesia.

Extremes: A Growing Concern, an Emerging Discipline

The concept that discrete extreme events rather than gradual trends can have disproportionate ecological and societal impacts relative to the temporal scale over which they occur has received increased public and scientific attention in the last few years. In particular, extreme climatic events such as hot spells, megadroughts, hurricanes, and floods involving adverse societal impacts or large economic costs, which often receive extensive media coverage, are increasingly being perceived as indicators of climate change (Easterling *et al.*, 2000a). Within the climate change science community, there is also a consensus based on instrumental and modeling evidence that a set of weather or climate events from the extreme tails of their frequency distributions are becoming and will become more frequent or more intense (or both) in the near future (Easterling *et al.*, 2000b; Meehl *et al.*, 2007). These advances in understanding climatic extremes as key components of the variability in the climate system are accompanied by a growing awareness of the importance of discrete extreme events (whether driven by climate, geology, or biological forces) in influencing ecological systems at all levels of organization from individuals to entire biomes (Gaines and Denny, 1993; Parmesan *et al.*, 2000; Jentsch and Beierkuhnlein, 2008; Smith, 2011).

An indicator of this emerging subdiscipline is increasing appearance of the term *extreme events* within the ecological literature (**Figure 1**) as well as special issues or special reviews on the topic appearing in mayor ecological journals such as *New Phytologist* (160, 2003), *Comptes Rendus Geosciences* (340, 2008), and *Journal of Ecology* (99, 2011).

Dimensions of Driver Extremity: Rarity, Intensity, and Severity

Despite this increased interest, a clear definition of extreme event is still elusive. In climate science, three criteria are often

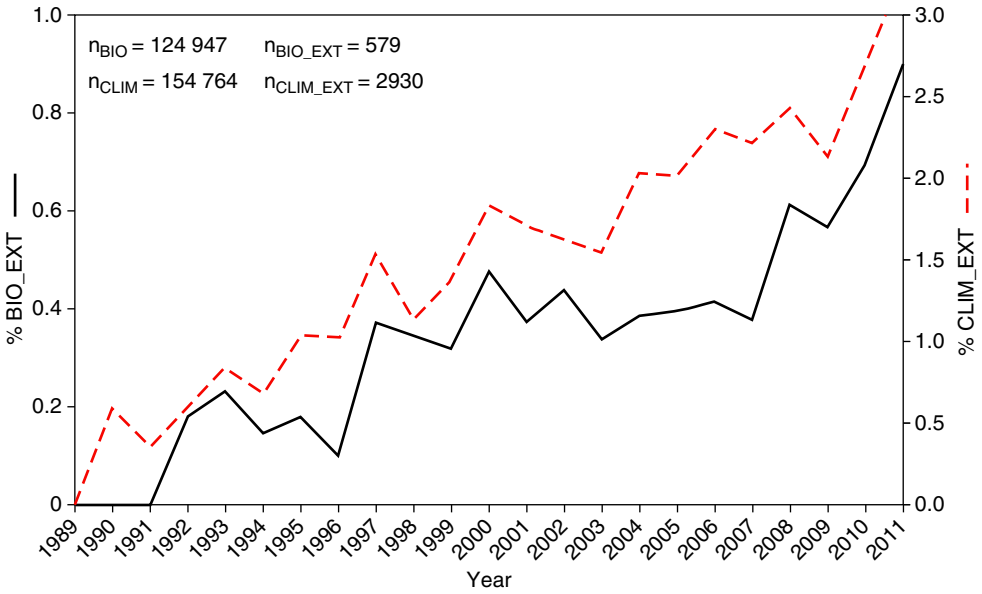


Figure 1 A bibliometric search of the topic "extreme event/s" using ISI Web of Knowledge in June 2011 yielded in 579 papers published in conjunction with the topics "ecosystems" or "biodiversity" (out of a total of 124,947 papers containing the latter two topics) and 2930 papers published in conjunction with the topic "climate" (out of a total of 154,764 papers containing the topic "climate"). The yearly percent of publications using the topic "extreme event/s" in the ecological literature has nearly doubled every decade, a rate that largely mirrors the pace of growth of the concept within the climatological community.

used to classify events as extreme: rarity, intensity, and severity (Beniston *et al.*, 2007).

Rarity

An event may be classified as extreme when it occurs at a particularly low frequency or rate. For example, the IPCC (2001) defines an extreme weather event as "an event that is rare within its statistical reference distribution at a particular place." To date, most discussions of extreme events have focused on the statistical description; however, the frequency and likelihood of a given type of extreme event are often poorly known, incomplete, and, in some cases, based on flawed models (Katz and Brown, 1992).

Intensity

Whether or not an extreme event will have important consequences depends in part on the physical characteristics of the event. Irrespective of whether they are rare or not, events characterized by relatively small or large values in these characteristics (i.e., events that have large magnitude deviations from the norm) will more likely have larger impacts compared to events closer to the mean or central tendency. The way intensity is delivered in space (spatial coverage) and time (timing of onset, duration, continuity, and temporal contingency with other events) are likely to determine the magnitude of the event. Although different research groups tend to focus either on temporal or spatial events, these often interact. Events are intense when the ratio between the greatest impact and a typical impact is large or when the magnitude of an extreme event is comparable to the magnitude of all the others taken together. When analyzing the distribution of intensities of events (e.g., magnitudes of earthquakes or sizes of snow avalanches or wildfires) the frequency often decreases as a power law (frequency$=[\text{size}]^{-\alpha}$), with α, the slope of a linearized (log–log) regression, falling in the range of 1–2 (**Figure 2**).

Many natural hazards such as earthquakes, landslides, and forest fires (Sethna *et al.*, 2001; Turcotte *et al.*, 2002) often display these empirical distributions, indicating that the sizes of the rarest and largest (or most intense) event is comparable to the sum of the sizes (or intensities) of all other events (Décamps, 2008). This suggest that under certain critical conditions thresholds are surpassed (e.g., for fire, a threshold drought or fire load condition) that determine the occurrence of very large or intense (extreme) events (Moritz, 1997; Slocum *et al.*, 2010).

Severity or Magnitude

Events are severe whenever they translate into large persistent ecological consequences. The concept is context-dependent because equally intense or equally rare events can translate in different severities in different ecological or evolutionary contexts (e.g., the effects of frosts can be severe in lowland tropical forests but will have little effect at higher latitudes or altitudes). Severe events occur when intensity surpasses some critical threshold. A system may become vulnerable to extreme events whenever intensity values driving variables approach or surpass critical thresholds. The existence and relative position of a critical threshold in turn depends on acclimatory, evolutionary, or ecological sorting processes the system may have experienced that confer the capacity to avoid, resist, or effectively respond to those changes. At the individual level, physiological acclimation ensures that environmental changes are tolerated within a certain normal range of conditions.

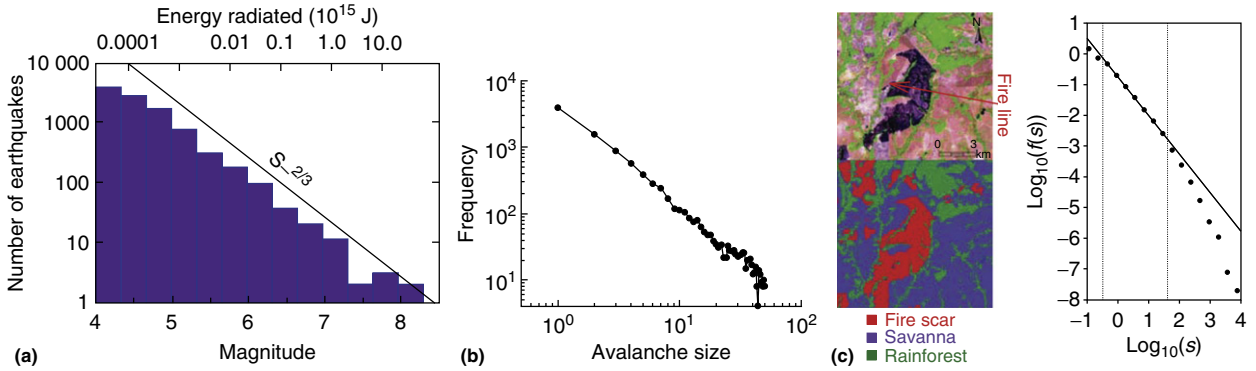


Figure 2 Examples of the distributions of natural events displaying power distributions of (a) earthquakes that occurred in 1995 as a function of their magnitude (logarithmic scale; Sethna *et al.*, 2001, reproduced from James *et al.*, with permission from Nature); (b) snow avalanche size distribution (Solé *et al.*, 2001, reproduced from Trends in Ecology); and (c) Amazonian fire size distribution as detected by "scars" on satellite images (Pueyo *et al.*, 2010, reproduced from Ecology Letters, 2010, with permission from Wiley).

At the population level, because extreme events are considered the effective arenas where natural selection takes place (Gutschick and BassiriRad, 2003), former regimes of extreme events may have selected traits to avoid, resist, or circumvent future events. At the level of communities and ecosystems, functional redundancy (more than one species affecting ecosystem properties in similar ways) may provide the capacity to maintain ecosystem functioning (Chapin *et al.*, 1998).

Toward Organism- and System-Based Definitions of Extremity

Extreme climatic or nonclimatic events of interest to ecology are those that impact physiology, development, or evolution of organisms and ultimately affect all levels of ecological organization. Thus, a more unified and operational definition of extremity should emerge from understanding how events shape individuals, populations, and ecosystems. Gutschick and BassiriRad (2003) propose that extreme events are "those episodes during which the acclimatory capacities of organisms are substantially exceeded in rate or magnitude." Consequently, after the event passes, the organism (or population) exhibits physiological and developmental responses to the environment that are significantly different from normal acclimation. Extremity thus differs from simple stress in that the system carries persistent irreversible aftereffects and displays hysteresis or a lack of capacity to retrace its path as a driving variable varies cyclically. Irreversible or hysteretic responses may result from a driving variable changing at a rate that exceeds acclimation rates or that exceeds an absolute upper or lower limit of acclimation of the organism (Gutschick and BassiriRad, 2003; **Figure 3**).

Although there is no equivalent description of acclimation or adaptation at higher levels of organization, clearly the range of responses of communities or ecosystems will likely be a function of the "rarity" of the event relative to the set of conditions that the system and its component species have previously experienced (Smith, 2011). An extreme event acts either as an external disturbance which a system can resist or as a disturbance exceeding the resilience of an ecosystem – that is, its ability to return to its former dynamic state (Bréda and Badeau, 2008). The concept of hysteretic ecological responses (i.e., responses incapable of retracing back its trajectory) may also be useful to define extremity at higher levels of organization such as communities or ecosystems. Events may produce changes that go beyond the capacity of communities to return to the previous state or may make the recovery period extremely long and thus induce community reorganization and ecosystem restructuring.

Types of Extreme Driving Events

Climatic extremes and their ecological consequences have received the most attention because their rate and magnitude of occurrence may be affected by global change. However, extreme events with important ecological responses may also be nonclimatic such as geologic or geomorphic physical disturbances (e.g., earthquakes or landslide-induced root damage) or resource inputs induced by these events (e.g., volcanic ashfalls, large wildfires, Turner *et al.*, 1997). Extreme events can be of biological origin; prominent examples are episodic resource pulses related to reproductive events such as mast fruiting or seeding in plants, periodic irruptions of palatable insects, or episodic frass input during insect outbreaks (Ostfeld and Keesing, 2000). Other resource pulses can be of mixed biological–climatic origin such as storm-related exports of marine mineral nutrient (subsidies) to terrestrial ecosystems (Rose and Polis, 1998), storm-related mass leaf shedding in forests, or rainfall-event driven biomass or seed production in desert ephemeral plants with bottom-up effects on higher trophic levels (Holmgren *et al.*, 2001). Clearly, these pulses are rare and intense, yet the severity in terms of long-term ecological consequences is far from being totally understood.

Large infrequent disturbances such as fires could qualify as extreme events that are generally driven by extreme climate events (e.g., ENSO related-droughts, heat waves) and can have strong long-lasting impacts on ecological systems (Moritz, 1997). Extreme fires are generally related to infrequent large or severe events capable of mobilizing large quantities of

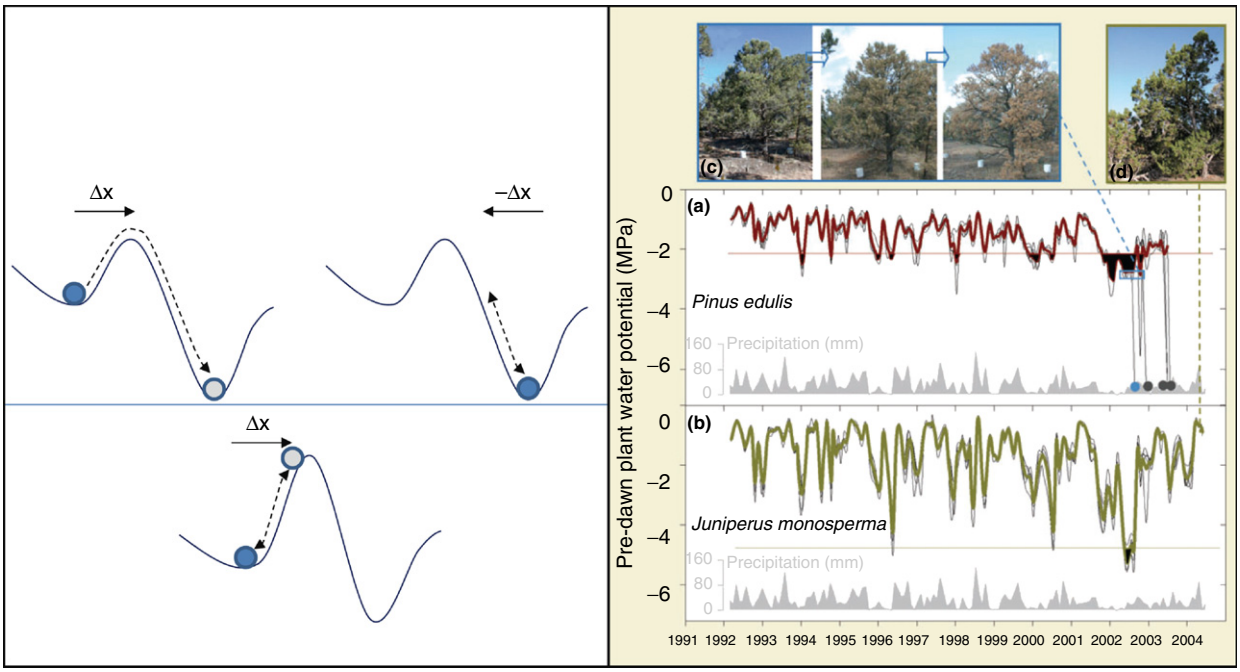


Figure 3 Example of contrasting trajectories of water stress between co-occurring species, depending on acclimation potential. A global change-type drought occurred in southwestern North America during 2001–2003 (Δx) inducing irreversible regional scale die-off of piñon pine trees (*P. edulis*) (a, c) while only producing temporary water stress in juniper (*Juniperus monsperma*) (b, d). Junipers exhibit anisohydric behavior in which low stomatal control allows leaf water potential to frequently decrease as soil water potential decreases. Despite higher variability in leaf water potentials (b), die-off risk is lower due to lower zero-carbon assimilation thresholds (i.e., the water potential at which net assimilation of carbon is zero; approximately -4 MPa) and total stem cavitation threshold (i.e., the water potential at which formation of air bubbles disrupt the xylem water transport, approximately –11 MPa) in juniper. In contrast, piñon pines are isohydric and normally avoid low water potentials by strong stomatal control (a). Piñon pines are more vulnerable to drought because zero-carbon assimilation lies only at approximately -2 MPa and total stem cavitation at approximately -7 MPa. Thus, drought intensity and duration contributed to die-off of piñon pine but not of juniper. Note that with the return to normal hydrological conditions ($-\Delta x$) junipers recovered predrought water potential levels (b) whereas piñon either died or could not recover predrought water potentials (a) (Breshears *et al.*, 2009). Left panel modified from Beisner BE, Haydon DT, and Cuddington K (2003) Alternative stable states in ecology. *Frontiers in Ecology and the Environment* 1: 376–382. Right panel (a–d) reproduced from Breshears DD, Myers OB, Meyer CW, *et al.* (2009) Tree die-off in response to global change-type drought: Mortality insights from a decade of plant water potential measurements. *Frontiers in Ecology and the Environment* 7: 185–189, with permission from ESA.

nutrients, consuming large quantities of aboveground and belowground biomass and organic matter, modifying substrates, resource availability, and the physical environment and thus disrupting population, community, and ecosystem structure and functioning. Finally, biological invasions have been proposed as extreme events (Gutschick and BassiriRad, 2003), yet despite the important impacts on the invaded community, their inclusion is questionable as they remain in the system, chronically affecting it, and thus they may not qualify as discrete events.

Events, Trends, Regime Shifts

Events are naturally discrete in that they involve a certain magnitude of change in drivers and response variables during a certain time, thus defining a certain abruptness (White and Jentsch, 2001). Geologic (e.g., earthquakes), geomorphic (e.g., landslides), hydrologic (e.g., floods), and ecological disturbances (e.g., fire) have in common that they involve a triggering phase, a spreading phase during which large amounts of energy are released and transformed, affecting ecosystems in a certain magnitude, followed by a slowdown and stopping phase. Climatic events are harder to pinpoint because often there are no clear-cut triggering and slowdown phases, and the distinction between discrete and gradual trends is blurred. White and Jentsch (2001) suggest the use of abruptness (the magnitude to duration ratio) to define events; because magnitude is a measure of the system's response, the definition of events is system- and hierarchy-specific and dependent on the response duration, which in turn is related to the longevity of individuals or the successional speed of communities. Thus, using such relative currency to express abruptness allows comparison among organisms with different life spans or communities with different response speeds (Jentsch *et al.*, 2007).

Large Scale-Drivers of Ecological Events

Although ecologists have historically focused on the effects of locally measured weather components, increasing attention has been recently paid to large-scale patterns of climate variability with marked ecological impacts on interannual and

longer time scales (Stenseth *et al.*, 2003). A growing suite of large-scale indices that reduce complex space and time variability into simple measures or "packages of weather" have become widely available, and, paradoxically enough, these have been often found to be better predictors of ecological processes than local climate itself (Hallett *et al.*, 2004).

One of the most prominent features of large-scale climatic patterns is the ability of spatially synchronizing ecological effects, with examples in most ecosystems of the world, both marine and terrestrial (Stenseth *et al.*, 2003; Holmgren *et al.*, 2001). Mechanistically, synchronization occurs because large-scale teleconnections lead to synchronization of climate variables over mesoscales (10^4–10^6 km^2, Swetnam and Betancourt, 1998) and these same climatic variables have strong bottom-up effects on food-web structure and ecosystem processes. Rainfall changes associated with ENSO in semiarid regions may, for instance, produce a highly synchronic pattern of massive annual plant germination and seed production, inducing rodent outbreaks and predator responses (Lima *et al.*, 2001). The ability of climate to synchronize population dynamics over large geographic areas is also known as the *Moran effect* (Ranta *et al.*, 1997). In forests, El Niño Southern Oscillation (ENSO) events may trigger climatic cues for the synchronization of tree fruit and seed production with cascading effects across the food web and consequences on tree demography. For instance, in dipterocarp forests of Indonesia, community-wide mast fruiting associated with ENSO events triggered regional-scale tree seedling recruitment pulses from large amounts of seeds escaped from predation. At the same time, the population sizes of nomadic vertebrate seed predators increased by two to three orders of magnitude both through reproduction and regional movement (Curran and Leighton, 2000). Similarly, ENSO (El Niño phase)-related grass growth in semiarid regions may, in a lagged fashion, induce large synchronous fires over large areas when an immediate phase reversal to dry La Niña produces desiccation of fuels, thus generating a phase locking between climate oscillations and ecological events (Swetnam and Betancourt, 1990; Kitzberger and Veblen, 1997; Kitzberger *et al.*, 2001).

Because climate variation occurs over spectra of frequencies of climatic events, events may therefore be hierarchically embedded within longer trends or waves of low frequency variation; in other words, high-frequency (e.g., interannual) events may be superimposed over lower-frequency events (e.g., multidecadal). A classic example is variability teleconnected to ENSO embedded in decadal or multidecadal regime shifts between warm and cold ocean states like the Pacific Decadal Oscillation (PDO) or the Atlantic Multidecadal Oscillation (AMO). In this way, the occurrence of extreme climatic and ecological events has been related to the modulating (or contingent) action of high- (e.g., ENSO) and low-frequency events (e.g., PDO, AMO; Jackson *et al.*, 2009). These modulating effects of low and high-frequency climatic modes have been documented as drivers for the occurrence of droughts (Enfield *et al.*, 2001; Gray *et al.*, 2003; McCabe *et al.*, 2004), hurricanes (Kossin *et al.*, 2010), wildfires (Schoennagel *et al.*, 2005; Kitzberger *et al.*, 2007), and events affecting tree demography (Gray *et al.*, 2006).

Dimensions of Extremity in Ecological Responses

The occurrence of an extreme event in a potential driver does not necessarily translate into the occurrence of an extreme ecological response. Understanding when, how, and to what extent extreme events translate into substantial ecological responses is still in its infancy and is the basis of an emerging new field in ecology (Smith, 2011). Authors have considered extreme ecological responses when a set of nonmutually conditions are met.

Response Rarity

Ecological responses are extreme if the alteration in structure or function lies outside the bounds of the "natural or historical range of variability" or the range of critical ecological processes and conditions that have characterized particular ecosystems over specified time periods (Landres *et al.*, 1999). This approach typically uses retrospective data (instrumental, historical, dendroecological, palynological, or other paleoenvironmental proxy) to place the condition of a specific event within the envelope of the previously reconstructed distribution of conditions. Rarity is therefore a function of the capacity of the historical or paleoenvironmental proxy to reconstruct the condition back in time. For example, Suarez and Kitzberger (2010), using dendroecological reconstructions along a rainfall gradient in northern Patagonia showed that the 1998–1999 megadrought induced in the dry eastern forest a concatenation of a massive tree mortality event, growth releases of surviving trees, and a pulse of seedling establishment that was unprecedented during the 20th century. In contrast, in the wet western end of the gradient these processes lay within the range of natural occurrence.

Response Abruptness

In many systems, gradual changes in factors might have little effect until a threshold (commonly referred to as a *tipping point*) is reached at which a large (abrupt) shift occurs that might be difficult to reverse (Scheffer *et al.*, 2001). This extreme ecological response may be the response to a regime shift of the driving factor, a progressive and gradual change that surpasses a threshold. A discrete driver event may surpass a threshold either because the factor increased its variability or because the factor shows low variability but slowly approaches the threshold value and eventually surpasses it (Holling, 1973; Maslin, 2004; Scheffer *et al.*, 2001; Williams *et al.*, 2011; **Figure 4**).

Response Characteristics

Irreversible, Slow, Delayed, or Hysteretic Recovery

Ecological responses to extreme events often involve permanent irreversible changes in ecosystem structure and function or changes that require long recovery periods to return to prior structural or functional states. For example, large severe infrequent fires that remove adults and propagules of preexisting dominant vegetation in combination with soil erosion, desiccation, and high radiation levels due to removal of dominant cover may determine persistent vegetation shifts or extremely

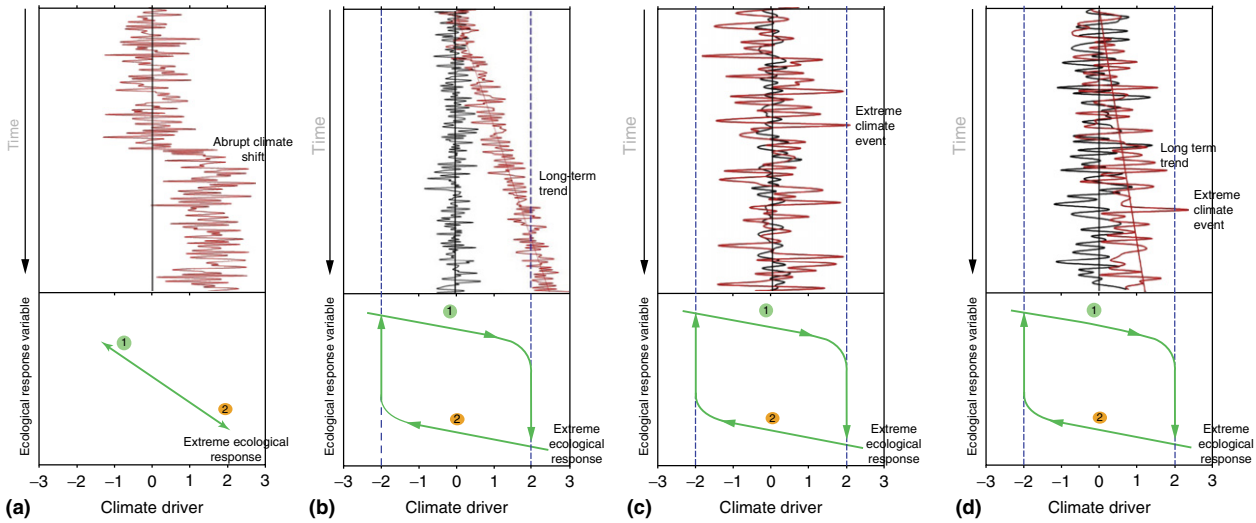


Figure 4 Abrupt ecological change can occur when (a) a climatic forcing abruptly changes in a steplike fashion, inducing an equally abrupt (but reversible) ecological change from state 1 to state 2 (e.g., sudden reduction of trees versus grasses); (b) when a slow-changing forcing variable eventually surpasses a critical threshold (right vertical line) that produces a major ecological shift; (c) when a forcing variable increases its variability and surpasses in an eventlike fashion a critical threshold that produces an abrupt ecological response; or (d) when a forcing variable exhibits a trend that gradually increases the chances of surpassing a critical threshold. Note that ecological trajectories in cases b–d are hysteretic in that even if the driver returns to previous values, the system remains "trapped" in state 2 as long as the lower (left vertical line) threshold is not reached.

slow regeneration of preexisting vegetation (Allen and Breshears, 1998). Large fire events will have persistent consequences whenever thresholds of fire resistance or propagules' dispersal capacities are surpassed (Romme *et al.*, 1998).

Discrete events may push the system to a state or condition that precludes tracing back to the stages prior to the event; in other words, the forward and backward responses may not follow the same path (hysteretic response). At the level of organismal responses, hysteresis is the property that differentiates stress acclimation (a reversible response) from responses to extreme events that have surpassed acclimation capacity (Gutschick and BassiriRad, 2003). For example, xylem conductivity, evaporation, and assimilation in plants that suffered drought cavitation (i.e., the disruption under tension of the capillary properties of xylem water due to air-bubble formation) do not follow the same path of recovery (if any) as stressed plants that deal with drought without cavitating. Hysteresis results because the recovery is delayed (e.g., xylem embolism cannot be repaired instantaneously), so the return of the driving variable (e.g., evaporative demand) to its initial value does not instantaneously lead the plant to its initial state through the same forward trajectory. Hysteretic responses at high levels of organization are also common. For example, in semiarid grasslands, experimental rainfall pulses induced hysteretic responses in functional properties of the ecosystem (photosynthesis, respiration and evapotranspiration). Soil wetting and subsequent drying induced changes in belowground microbial communities that produced soil respiration values different from that observed under the reverse trajectory (i.e., soil drying and subsequent rewetting; Potts *et al.*, 2006; Williams *et al.*, 2009). Severe fires induced by global change type droughts may also show hysteretic properties compared to fires to which the components of the local system are well adjusted (**Figure 5**).

Response Magnitude

Loss of Key Structuring Species, Changes in Physical Conditions, and Indirect Consequences

Extreme ecological responses may involve the local to regional loss of key structural (e.g., habitat-forming) foundation species due to range shifts, ecotone shifts generated by massive mortality, reproductive and behavioral alterations, migration, and so on. These losses produce concurrent changes in environmental conditions (e.g., light regimes, soil water regimes, etc.) that can permanently or over long periods affect subsequent dynamics (Royer *et al.*, 2011). Climate extremes such as droughts have the potential to rapidly alter vegetation structure and transform associated ecosystem properties. For example, drought and heat anomalies have been implicated in tree die-off on every wooded continent (Allen *et al.*, 2010). An immediate consequence of tree die-off in response to an extreme drought event is the loss of canopy cover (Allen and Breshears, 1998; Royer *et al.*, 2011), resulting in increased near-ground solar radiation, a fundamental driver of ecological processes that in turn impact the physiology, phenology, and reproduction of surviving plants, as well as soil water status and nutrient cycling, and facilitation–competition balances (Sthultz *et al.*, 2007). Such microenvironmental changes may impact ecosystems, reducing productivity and diversity (Zhao and Running, 2010), community dynamics, for example, altering forest successional pathways (Suarez and Kitzberger, 2008), or increasing mortality on understory or codominant woody species (Breshears *et al.*, 2005). Lastly, die-off may indirectly impact disturbance regimes (e.g., wildfire, Breshears, 2006; Clifford *et al.*, 2008; insect outbreaks, Floyd *et al.*, 2009), increase invasibility (Kreyling *et al.*, 2008) and accelerate erosion and land degradation rates (Allen and Breshears, 1998).

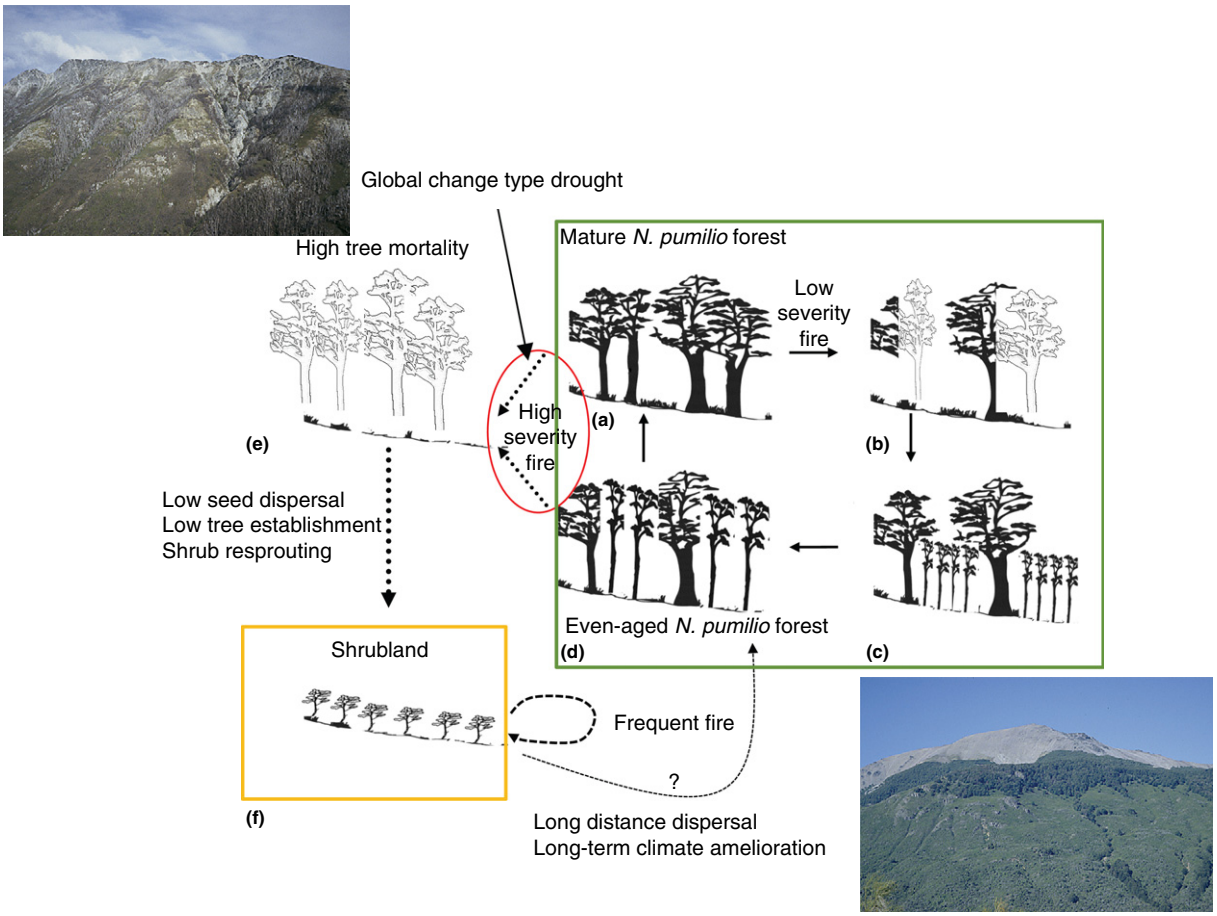


Figure 5 In northern Patagonia, subalpine forests show hysteretic response in relation to drought and fire. Under normal climatic variability, the alternative state *N. pumilio* forests: (a–d) are dominated by obligate seed–dispersed species and shrublands (*N. antarctica-Chusquea culeou* bamboo; (f) are dominated by fire resprouters and have their own internal fire dynamics of low frequency and low severity in the forest and high frequency in the shrublands. Most shrubland fires stop when they reach the forest boundary, creating sharp edges typical of northern Patagonian landscapes (bottom right photography). Eventually, some fires enter the forest as low-severity fires. However, under conditions of global change-type drought, the fire-sensitive *N. pumilio* forests become flammable and high mortality occurs (e), such as in the unusually large and severe fires that occurred in 1999 (top left photography). Under these conditions *N. pumilio* is unable to disperse or establish due to changes in surface conditions, and shrubs and bamboos, formerly confined to the understory, vigorously resprout and dominate. Return to the forest state has not been documented but may occur at very slow rates and under wetter or colder long-term conditions (Mermoz *et al.*, 2005; Kitzberger *et al.*, 2005). (a–f) Reproduced from Maria Laura *et al.*, with permission from Wiley.

Cascading Effect Across Hierarchical Levels Ultimately Affecting Ecosystem Functions

Extreme climate events can produce physiological responses that exceed acclimation capacity of individuals. If severe enough, these responses may translate into fitness of individuals and population demography (e.g., increased mortality); and if occurring in dominant or key structuring species, they may influence community structure and composition and ultimately produce alteration of ecosystem functions such as primary productivity, water regulation, carbon fixation, and nutrient cycling (Jentsch *et al.*, 2011).

This area of research is related to a relatively new area of manipulative event-based experiments that started in the early 1990s and gained momentum during the last decade (Jentsch *et al.*, 2007). Experiments have the advantage over observational studies of ongoing or past extreme events of allowing adequate replication and control of event characteristics (magnitude, duration, timing, etc.), thus enabling the analysis of causal relations and the identification of thresholds beyond which individual allocation, population fluctuations, community structure, or ecosystem functions change significantly (Jentsch *et al.*, 2007 and references therein).

Extreme Ecological Responses: Emerging Patterns and Challenges

Extreme Ecological Responses Are Less Common Than Extreme Driver Events

A pattern that is beginning to emerge from experimental studies is that extreme ecological responses are probably much rarer than the occurrence of extremes in potential drivers, and possibly much less frequent than suggested by observational

studies (Smith, 2011). Not all climate extremes lead to ecological extremes, but some ecological extremes (e.g., regime or compositional shifts) may be triggered by small changes in drivers if complex systems are close to tipping points (Van Nes and Scheffer, 2004, ecological surprises). Clearly, extreme ecological events reported from observational studies suffer from being a "filtered" sample of extreme events. Observational studies suffer from two potential biases: in the first, extreme conditions of drivers do not trigger extreme ecological responses and remain unstudied; in the second, cases of extreme events that do not produce significant ecological responses become underreported due to the file "drawer problem" (Bauchau, 1997). Therefore, our view from observational studies may be biased toward equating climate extremes with extreme climatic events (Smith, 2011). However, the apparent resistance of ecosystems observed in experimental studies may not be of such value in the real world because of the following factors.

Too simple systems: Highly simplified experimental systems leave out complexities that reduce resilience (e.g., effects of herbivorous insects, pollinators).

Too small systems: Experimental plots cannot incorporate larger-scale processes such as species range shifts, dispersal of plants, and insects.

System bias: Because of logistic restrictions, event-type experiments have been carried out mostly in grasslands with small-sized and short-lived plant species compared to woody-dominated ecosystems. For example, between 2008 and 2011, 24 data-based publications (22% out of 110 with topic "climate change" AND "experiment∗" AND "event∗" AND "grassland∗" OR "heathland∗" OR "tundra∗"; Web of Knowledge) appeared, having experimentally manipulated grasslands, heathlands, or tundras in an event-type fashion. In contrast, for the same period, only 10 data papers (7% out of 136 with topic "climate change" AND "experiment∗" AND "event∗" AND "forest∗"; Web of Knowledge) reported on manipulated woody-species dominated ecosystems. Conversely, observational studies are possibly biased toward woody systems because it is in those where the effects – at least in structural, compositional, or demographic terms – are most evident (massive mortality, rapid compositional shifts, sharp structural changes).

Historic blindness: Unless done on artificial communities where historic effects are controlled, experimental studies suffer from ignoring historic contingencies – for example, prior driver regime that may have acclimated species or the prior occurrence of extreme events or disturbances that may have modified functioning influencing the resistance and resilience in the face of the experimental conditions imposed.

Too short measurement periods: Ecological responses may be slow, lagged, or display long-term feedbacks that delay the most important responses, thus preventing experiments from detecting them. For example, in experimentally manipulated grasslands plots, important shifts in composition still occurred 3 years after the ending of treatments (Buckland *et al.*, 2001). Lagged effects are also very important in systems with long-lived species. In trees, for example, the effects of drought might become evident with important time lags related to complex interactions between physiological stress and mortality agents such as insect and diseases, which may trigger tree mortality in a lagged fashion (Cherubini *et al.*, 2002; Bigler *et al.*, 2007; Bréda *et al.*, 2006). Single extreme events may have long-lasting effects on the dynamics of plant production due to internal feedbacks. For example, Haddad *et al.* (2002) suggest that a single drought may be the cause of 9-year-long periodic oscillations in productivity of grassland driven by litter-nitrogen productivity feedback effects.

Too simple treatments: Manipulations that only amplify one aspect of an extreme event may not suffice to produce ecological responses. For example, the timing and duration of drought are often more important than the absolute reduction in rainfall. Combinations with other factors (e.g., temperature and drought) are increasingly being incorporated. More complex treatments such as sequences of events are probably more realistic.

Too unnatural treatments or excessive undesired artifacts: Imposing treatments that mimic climate extremes can be a challenging task. For example classic warming artifacts such as open-top chambers by being passive have the problem of producing variable temperature differences, depending on short-term micrometeorological conditions. Moreover, by being partially enclosed they reduce wind speed and increase relative humidity. A solution to this is the use of modulated infrared lamps that instantaneously maintain constant differences between treatment and ambient temperatures (De Boeck and Nijs, 2011).

Weak treatments: Imposed conditions may have not be sufficiently extreme in comparison to what the system may have previously experienced (Smith, 2011).

Extreme Events Can Alter Components but Generally Higher Level Functions Are Maintained by Compensatory Mechanisms

Ecological systems may display a range of responses to climate extremes, from no response to extreme responses, depending on hierarchically complex, context-dependent sets of interacting mechanisms. Single-species reactions cannot be translated directly to the community and ecosystem levels. Extreme events may produce inconsistent functional responses, or no response at all, at higher hierarchical levels, due to complex complementary compensation mechanisms at lower levels. For example, drought has an important effect on the functional performance (e.g., carbon fixation and nutrient cycling) of individual species or the community, but ecosystem properties related to primary production remain remarkably stable. By changing their relative abundance, drought may change the relative strength of competitive and facilitative interactions among different neighboring plant species, thus buffering primary production changes at the community level without changing community composition (Jentsch *et al.*, 2011).

More Species-rich Communities are not Necessarily more Resistant to Extremes

It has been proposed that species richness may be playing a key role in the ability to withstand perturbations while maintaining ecosystems functions (*resistance*). Richer communities may increase the likelihood that some species or set of species with previous low performance will increase their

performance after the perturbation and compensate for others ("insurance hypothesis," Yachi and Loreau., 1999). However, experiments show conflicting results on the role of species' richness modifying resistance. Some drought or drought and heat experiments produced higher absolute biomass loss or plant mortality in species-rich than in species-poor synthesized grassland communities (Pfisterer and Schmid, 2002; de Boeck *et al.*, 2008; Van Ruijven and Berendse, 2010), presumably because more species-rich systems have a greater chance of including species that perform better under unperturbed conditions but worse under drought conditions ("opposite insurance hypothesis," Pfisterer and Schmid, 2002). Other experiments showed increased resistance with higher species richness (Tilman and Downing, 1994) or lack of relationship (Kahmen *et al.*, 2005; Wang *et al.*, 2007). Possible reasons for these inconsistencies are different relationships between species richness and productivity. When productivity and species richness are positively related, more diverse-productive systems suffer more from droughts than less diverse-less productive ones. In contrast, when diversity and productivity are negatively correlated, more diverse-less productive systems are more resistant to drought. Irrespective of their species richness, communities with lower biomass may be more resistant to drought than more productive ones (Wang *et al.*, 2007).

Importance of Ecological Memory, Recurrent Extreme Events, Sequences of Extreme Events, and Compound Effects of Extremes

Both experimental and retrospective studies are beginning to elucidate the importance of recurrent extreme events as opposed to single extreme events in determining the response and recovery of ecosystems. Outcomes from repeated extreme events can be opposite: each single abiotic stress may weaken organisms and reduce resilience of ecosystems, thus resulting in deteriorated performance and a total loss of resilience when the system is affected by recurrent stressful events (Scheffer *et al.*, 2001). Alternatively, organisms are able to acclimate when faced with abiotic stress, revealing phenotypic plasticity as a response to environmental variability and thus conferring increased resilience (Walter *et al.*, 2011).

Support for the first mechanism can be found in long-lived species such as trees that may witness several extreme events during their lives. Forensic reconstructions using tree rings from trees damaged or killed by extreme drought events and comparison with surviving trees suggest the existence of important negative memory effects from previous droughts (**Figure 6**).

In support of the first mechanism, populations previously weakened by extreme events are *more* vulnerable to future stressing agents such as extreme climatic events (Suarez *et al.*, 2004; Lloret *et al.*, 2004) or pest attacks (Bréda and Badeau, 2008; **Figure 4** and references therein). In support of the second mechanism, physiological adjustment to previous drought events (e.g., root or shoot plastic adjustment, **Figure 6b**; Suarez *et al.*, 2004; photo protection, Walter *et al.*, 2011) have made some systems less vulnerable to subsequent droughts.

Experimental studies of vegetation responses to climate have largely focused on responses to a trend in climate or to a single extreme event but have largely overlooked the potential for complex responses to specific sequences of extreme events. Recent experiments imposing different ordering of opposite hydrological events (drought–flood) show that species responses were dependent on the sequence of events imposed as well as the presence or absence of an interceding baseline (average) condition between extreme events (Miao *et al.*, 2009). Moreover, ecological responses (plant growth and mortality) to sequences were not simply an additive result of the responses to individual events and are not likely to be predictable a priori based on the knowledge of the physiology and distribution of species (Miao *et al.*, 2009).

The importance of sequences of extreme events for ecological responses has been long recognized in retrospective studies that normally reconstruct ecological events and relate them with time series of drivers. Century-long reconstructions of fire regimes and climate suggest that the occurrence of extreme fires in certain forest types where grasses are important is a direct function of a sequence of extremely wet years that stimulate grass growth followed by extremely dry years that desiccate fuels (Swetnam and Betancourt, 1990; Kitzberger and Veblen, 1997). These sequences are often linked to dominant modes of climate variability such as ENSO (a sequence of the wet phase of El Niño followed by the dry phase of La Niña), and thus extreme ecological processes may be synchronized over large areas with dominant modes of variability (Kitzberger *et al.*, 1997). Other examples of ecological events triggered by particular sequences of drivers are establishment pulses of masting plants (i.e., species that have evolved very irregular temporal seeding patterns with extremely high seed production events interceded by periods of low or null seed production). These species generally require an extreme drought to synchronize the massive seed production followed by an abnormally wet period that ensures adequate germination and seedling survival.

This locking to particular sequences of forecastable drivers has important management implications, as it allows some degree of prediction of extreme ecological events (e.g., Kitzberger, 2002) and the prediction of ecological consequences of particular sequences (e.g., flood-drought) that may become common under future climate change (Miao *et al.*, 2009). Finally, compound sequences of extreme events (Paine *et al.*, 1998) composed by climatic and nonclimatic events may generate extreme ecological responses. For example, mass flowering events of semelparous bamboos (species that reproduce gregariously once in their lifetimes, after which they die) generate enormous loads of fuels when culms die off. When these events are followed by strong droughts that desiccate fuels, large and severe fires may occur (Keeley and Bond, 1999 and references therein, Veblen *et al.*, 2008).

Importance of Anthropogenic Amplification in producing Extreme Ecological Responses

Global climate change is projected to yield both higher temperatures and increases in the frequency and intensity of extreme events (Easterling *et al.*, 2000b; IPCC, 2001). Moreover, increased temperatures and droughts may act synergistically to produce extreme events. Unique oceanic and circulation

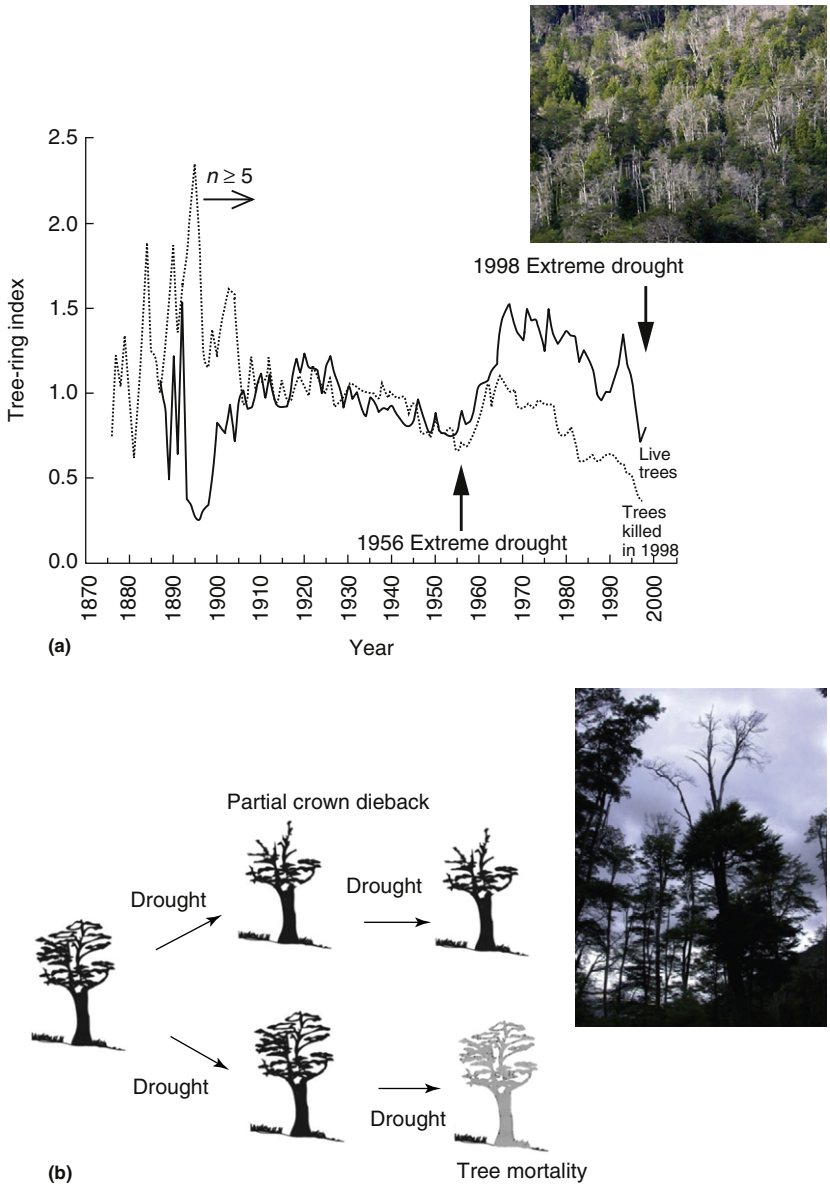


Figure 6 Examples of opposite effects of ecological memory in relation to past drought events. (a) Tree-ring width chronologies from *Nothofagus dombeyi* trees killed during the most recent 1998–1999 extreme drought in northern Patagonia and from trees that survived that drought. Note that whereas tree growth was nearly equal between groups during 1910–1950, after 1956 growth of trees that were later killed during the 1998–1999 drought was 30–50% lower than growth of trees that survived the 1998–1999 drought. This suggests that the fraction of trees that suffered more from the 1956 drought (in terms of reduced growth) were those that during the 1998–1999 drought were finally killed. In contrast, trees that survived the last 1998–1999 drought were not weakened by the previous drought. On the contrary, in the same system and for the same drought events (b), trees that had adjusted to prior drought events by shutting off the top crown, a common syndrome known as *partial crown dieback*, suffered less mortality from the 1998–1999 drought event than trees with large intact crowns (large shoot to root ratios).

conditions that are "ideal" for drought and embedded within a multidecadal warming trend can cause global change-type droughts such as the globally spanning 1998–2002 drought (Hoerling and Kumar, 2004). One of the most prominent examples of extreme ecological responses are massive forest die-off episodes. Although episodic tree mortality may occur in the absence of climate change, it is increasingly evident that under current and future climate changes, forests may become increasingly vulnerable to increased tree mortality rates and die-off in response to warm droughts (Allen *et al.*, 2010). Examples of these events come from more than just semiarid systems such as the 2002–2003 massive die-off of *Pinus edulis* (piñon pine) woodlands in North America (**Figure 2**; Breshears *et al.*, 2005). In environments that are not normally considered water-limited, 30–40-m tall *Nothofagus dombeyi* trees died massively during the warm drought event of 1998–1999 that affected northern Patagonia, Argentina (Suarez *et al.*, 2004).

Similarly, the dynamics of eruptive forest herbivores (i.e., defoliating or boring insects that eventually produce outbreaks) is controlled by complex interactions across different hierarchical levels (Raffa *et al.*, 2008), and sometimes both discrete extreme climate events and trends are needed to produce regime shifts. Severe drought reduces tree resistance and can provide the pulse of stressed trees that beetles need to surpass the eruptive threshold, which in complex interaction with the direct effects of warming on beetle survival caused spruce beetle *(Dendroctonus rufipennis)* outbreaks over vast areas of Alaska and the adjacent Yukon Territory during the 1990s (Berg *et al.*, 2006).

Warming may interact with severe fires to lead to extreme ecological responses. For example, in northern Patagonia during extreme droughts such as the La Niña–related 1998–1999 drought, vegetation firebreaks (mesic less flammable vegetation) and topographic firebreaks (ravines or mountain ridges) that normally control fire spread during average climate lifted during droughts, allowing fire events to seep over large sections of the landscape (Mermoz *et al.*, 2005), thus severely affecting fire-sensitive subalpine *Nothofagus pumilio* forests that are unable recolonize under high-radiation and warmer postfire conditions, leading to a shift from subalpine forests to shrublands (Tercero-Bucardo *et al.*, 2007).

Similarly, introduction of exotic species may reduce resilience and induce extreme ecological responses. In northern Patagonia, invasive European hares prefer browsing in the open areas generated by fire rather than under the forest canopy. Therefore, the presence of this novel herbivore strongly reduces the postfire regeneration of the vulnerable slow-growing seedlings of *N. pumilio*, thus shifting the system toward dominance of shrubs that vigorously resprout after fire, rapidly escaping the browsing action of exotic hares (Kitzberger *et al.*, 2005). In Hawaii, the nitrogen-fixing exotic *Myrica faya* actively invades new surfaces created by large-scale volcanic events and initiates a series of changes including the identity of the dominant tree, nutrient cycling, and productivity (Vitousek, 1990).

Other global anthropogenic changes related to land use may cause loss of resilience in the face of extreme events (Bréda and Badeau, 2008). For example, in the Amazon basin, deforestation causes forest fragmentation and increased edge effects, with changes in microclimate and species composition that make them more susceptible to ENSO-related megadroughts (Laurance and Williamson, 2001) and more prone to carry fires in these fire-vulnerable forests (Cochrane, 2001).

Understanding Extreme Events, a MultiPerspective Endeavor

Three different strategies have been used to analyze the role of events in ecology: look back for events, pinpoint particular ongoing events, or artificially impose events. Retrospective approaches take advantage of temporal depth to capture at least a few events, observational studies that actively pinpoint at individual events capture the event naturally in action, and manipulative experiments on natural or synthetic communities impose conditions to systems that attempt to mimic extreme natural conditions.

Retrospective studies should provide understanding of the long-term statistical distribution of both drivers and responses, which could provide a notion of extremeness of observed or manipulated events. Furthermore, retrospective approaches are providing important insights into the notions of ecological memory, event sequences, compound effects, and anthropogenic amplification. Observational studies deal with extreme events in action in their natural environment. These have the ability to provide clear notions of magnitude as they are the actual evidence of the consequences of extreme events in natural action. By being opportunistic, they generally fail to detect preexisting conditions but are very efficient in monitoring changes related to the event (e.g., surface conditions, structure, composition, functions). Because events operate as natural experiments, any community can be observed, but emphasis has been on woody-dominated ecosystems where changes are more visible because they often imply the mortality of key structural species. Finally, event experiments have the ability to impose precise conditions, modifying intensity, duration, timing, and sequence of events and to incorporate interactive factors. They also have the capacity of modifying the system to explore how diversity, biomass, and other properties affect resistance and resilience in the face of the imposed condition. This approach is rapidly gaining popularity thanks to important technological improvements that reduce undesired artifacts giving more realism to the condition imposed.

Clearly, none of these approaches by itself will provide full understanding of extreme events in ecology (**Table 1**). Instead, complementarity of approaches will be necessary to better understand the hierarchy of processes involved in responding to extreme drivers and producing extreme responses.

Extreme Events as Drivers of Global Change

The study of ecological responses to global change has traditionally focused on the effects of climatic trends such as gradual warming, precipitation change, and carbon dioxide enrichment. Short, rare departures from average conditions have been generally overlooked, possibly under the assumption that only persistent changes in drivers would lead to persistent ecological changes. Yet persistence of conditions or slow changes in drivers give natural systems the ability to buffer and linearly adjust by means of an array of mechanisms such as physiological acclimation, resource reallocation, population genetic changes, species migration, and gradual changes in community composition. In contrast, when certain conditions or resources fluctuate sharply, they exceed thresholds beyond which the capacity of systems to effectively adjust is surpassed and, even if the driver fluctuation has a short duration, the consequences on ecological systems are severe and long-lasting. What is becoming clear is that the rules under which ecosystems respond to slow changes in average conditions are very different from the rules that drive ecological responses to abrupt driver changes.

Projected global changes suggest that neither of these responses should be ignored as both gradual changes in average

Table 1 Approaches for study extreme events in ecology and their associated capacities to detect dimensions of extremity of drivers and ecological responses as treated in this chapter

	Retrospective	*Observational*	*Experimental*
Statistical rarity of driver	Yes	No	No
Intensity of driver	Yes	Yes	No
Severity of driver	Sometimes	Yes	No
Spatial synchronization	Yes	Yes	No
Relationships with large-scale climatic patterns	Yes	Yes	No
Statistical rarity or severity of response	Yes	Overestimated	Underestimated
Response abruptness	Yes	Yes	Yes
Response characteristics	Sometimes	Yes	Yes
Response magnitude	No	Yes	Yes
Measures functional responses	No	Yes	Yes
System oversimplification	No	No	Yes
Accounts for larger processes?	Yes	Yes	No
System bias	No	Woody-biased	Grassland-biased
Captures slow or, lagged responses	Yes	Sometimes	Starting
Can account for timing, duration of event	No	No	Yes
Can account for factor interactions	No	No	Yes
Reveals effects of diversity	No	No	Yes
Reveals effects of memory, recurrence, sequences	Yes	No	Starting
Reveals compound effects	Yes	Yes	No
Reveals anthropogenic amplification	Yes	Yes	No

conditions and increases in extreme event frequency are expected. Moreover, extremes and long-term trends in drivers may be acting together. The joint effect of how gradual changes of baseline conditions and extreme events translate into extreme ecological responses is still far from being understood.

This chapter shows how global change forces *sensu lato* – for example, climatic change, species introductions, land-use changes (Sala *et al.*, 2000) – may be turning systems less resistant and resilient to extreme driver events (e.g., warming trends may stagnate recovery of systems that have suffered from large infrequent disturbances). Warming may also affect the susceptibility of ecosystems to eventual extreme events. Warmer conditions increase susceptibility of systems to suffer the consequences of extreme driver events, such as eruptive beetle outbreaks. Land-use changes, one of the drivers that probably will produce the largest global effects, also probably strongly interact with extreme events to produce substantial ecological responses. Fragmentation, land degradation, and structural simplification all probably decrease resistance to climate extremes such as droughts or drought-related fires by strongly modifying microenvironments, biotic composition, and fuel loads, particularly in structurally complex systems such as rainforests. Loss of genetic, specific, or functional diversity may be another force making ecosystems less resistant and resilient to extremes. However, clearly the relationships between diversity and stability in the context of extreme events are not clear and are probably very system-specific. Operational limitations need to be resolved to generate more experimental work on functional responses to extreme events in systems other than grass-dominated communities. Introduced plants and animals may modify structural and functional recovery of ecosystem that suffered extreme events. More work is needed to understand whether invasion modifies resistance of communities to extremes. It is also unclear if one of the responses to extremes is loss of resistance to invasions. Anthropogenic atmospheric changes (e.g., CO_2 increases, N deposition) may also be interacting to modify vulnerability thresholds to extremes, but these interacting effects are still poorly understood.

Extreme Events in Ecology

It is increasingly clear that extreme events are the scenarios in which individuals, populations, and communities undergo the most significant physiological, ecological, and evolutionary changes, whereas intervals between events are periods of relative quiescence. Therefore, trying to understand system properties based on long interevent (average) conditions is simply not enough. Yet because extreme events are generally also short and rare, standard experimental ecological research programs lasting a few (random) years of the systems are not likely to capture the events in action but rather the results of past events.

Extreme drivers and extreme events may be part of the natural functioning of ecosystems. It is therefore crucial to advance our understanding and predictive capacity of ecosystems and their services in the context of a more variable global climate. In other words, ecosystems as we conceive of and study them today are most likely the actual consequence of past extreme events. Thus, we need to understand extreme responses as part of the natural variability of ecosystems. Ecosystems may often be dependent on such extreme excursions to sustain some long-term ecological functions. For example, partial crown dieback produced by extreme droughts in northern Patagonia (Figure 6) are a necessary condition to the creation of nesting habitats for the Magellanic woodpecker, a cavity-building bird that sustains a whole guild of cavity-using animals in *Nothofagus* forests (Ojeda *et al.*, 2007).

Large fire events triggered by large fuel accumulations produced by die-off events in semelparous bamboos (*see* Understanding Extreme Events, a Multi Perspective Endeavor) are probably one of the rare but important regeneration opportunities for shade-intolerant, nutrient-demanding, long-lived trees in northern Patagonian forests (Keeley and Bond, 1999, Veblen *et al.*, 2008). Rare (sometimes drought-triggered) mast-seeding or fruiting events are key events not only for the dynamics of plant populations but also for the long-term persistence of consumers such as granivores (Curran and Leighton, 2000). Populations of terrestrial consumers such as coyotes may be dependent on extreme storms that bring marine nutrient pulses into coastal areas (Rose and Polis, 1998).

The pulse of terrestrial resources associated with El Niño precipitation has a profound effect on the ecology of many arid and semiarid terrestrial communities (Holmgren *et al.*, 2001). All these examples suggest that extreme events provide an extra source of resources and conditions that generate ecosystem functions and services that would not be possible in the absence of those events. This not only is conceptually important but also has practical management implications. Extreme events may help managers restore degraded functions – for example, ENSO rainfall events may provide temporal windows of opportunity for the restoration of arid and semiarid systems (Holmgren and Scheffer, 2001), something that would not be feasible during interevent periods.

Our understanding of the causes and consequences of extreme ecological events is still in its infancy. Maturation in this emerging discipline will require a better integration of approaches as well as a broader conception that includes extreme events as part of long-term ecosystem dynamics.

See also: Biodiversity and Ecosystem Services. Climate Change and Extinctions. Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. Climate, Effects of. Disturbance, Mechanisms of. Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems. Ecosystem Function, Principles of. Ecosystem Services. El Niño and Biodiversity. Fires, Ecological Effects of. Functional Diversity. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Impact of Past Global Warming on Biodiversity. Landscape Legacies. Loss of Biodiversity, Overview

References

Allen CD and Breshears DD (1998) Drought-induced shift of a forest-woodland ecotone: Rapid landscape response to climate variation. *Proceedings of the National Academy of Sciences* 95: 14839–14842.

Allen CD, Macalady AK, Chenchouni H, *et al.* (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* 259: 660–684.

Bauchau V (1997) Is there a "file drawer problem" in biological research? *Oikos* 79: 407–409.

Beniston M, Stephenson DB, Christensen OB, *et al.* (2007) Future extreme events in European climate: An exploration of regional climate model projections. *Climatic Change* 81: 71–95.

Berg EE, Henry JD, Fastie CL, De Volder AD, and Matsuoka SM (2006) Spruce beetle outbreaks on the Kenai Peninsula, Alaska, and Kluane National Park and Reserve, Yukon Territory: Relationship to summer temperatures and regional differences in disturbance regimes. *Forest Ecology and Management* 227: 219–232.

Bigler C, Gavin DG, Gunning C, and Veblen TT (2007) Drought induces lagged tree mortality in a subalpine forest in the Rocky Mountains. *Oikos* 116: 1983–1994.

Bréda N and Badeau V (2008) Forest tree responses to extreme drought and some biotic events: Towards a selection according to hazard tolerance? *Comptes Rendus Geoscience* 340: 651–662.

Bréda N, Huc R, Granier A, and Dreyer E (2006) Temperate forest trees and stands under severe drought: A review of ecophysiological responses, adaptation processes and long-term consequences. *Annals of Forest Science* 63: 625–644.

Breshears DD (2006) The grassland-forest continuum: Trends in ecosystem properties for woody plant mosaics? *Frontiers in Ecology and the Environment* 4: 96–104.

Breshears DD, Cobb NS, Rich PM, *et al.* (2005) Regional vegetation die-off in response to global-change-type drought. *Proceedings of the National Academy of Sciences* 102: 15144–15148.

Breshears DD, Myers OB, Meyer CW, *et al.* (2009) Tree die-off in response to global change-type drought: Mortality insights from a decade of plant water potential measurements. *Frontiers in Ecology and the Environment* 7: 185–189.

Buckland SM, Thompson K, Hodgson JG, and Grime JP (2001) Grassland invasions: Effects of manipulations of climate and management. *Journal of Applied Ecology* 38: 301–309.

Chapin III FS, Sala OE, Burke IC, *et al.* (1998) Ecosystem consequences of changing biodiversity. *BioScience* 48: 45–52.

Cherubini P, Fontana G, Rigling D, Dobbertin M, Brang P, and Innes JL (2002) Tree-life history prior to death: Two fungal root pathogens affect tree-ring growth differently. *Journal of Ecology* 90: 839–850.

Clifford MJ, Rocca ME, Delph R, Ford PL, and Cobb NS (2008) Drought induced tree mortality and ensuing bark beetle outbreaks in southwestern Pinyon-Juniper woodlands. In: Gottfried GJ, Shaw JD, and Ford PL (compilers) *Ecology, Management, and Restoration of Piñon-juniper and Ponderosa Pine Ecosystems: Combined Proceedings of the 2005 St. George, Utah and 2006 Albuquerque, New Mexico workshops. Proceedings RMRS-P-51.* Fort Collins, CO: U.S.D.A., Forest Service, Rocky Mountain Research Station.

Cochrane MA (2001) Synergistic interactions between habitat fragmentation and fire in evergreen tropical forests. *Conservation Biology* 15: 1515–1521.

Curran LM and Leighton M (2000) Vertebrate responses to spatiotemporal variation in seed production of mast-fruiting dipterocarpaceae. *Ecological Monographs* 70: 101–128.

De Boeck HJ, Lemmens CMHM, Zavalloni C, *et al.* (1998) Biomass production in experimental grasslands of different species richness during three years of climate warming. *Biogeosciences* 5: 585–594.

De Boeck HJ and Nijs I (2011) An alternative approach for infrared heater control in warming and extreme event experiments in terrestrial ecosystems. *Journal of Ecology* 99: 724–728.

Décamps H (2008) Ecosystems and extreme climatic events. *Comptes Rendus Geoscience* 340: 553–563.

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, and Mearns LO (2000b) Climate extremes: Observations, modeling, and impacts. *Science* 289: 2068–2074.

Easterling JL, Evans PY, Groisman R, Karl KE, Kunkel, and Ambenje P (2000a) Observed variability and trends in extreme climate events: A brief review. *Bulletin of the American Meteorological Society* 81: 417–425.

Enfield DB, Mestas-Nunez AM, and Trimble PJ (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophysical Research Letters* 28: 2077–2080.

Floyd ML, Clifford M, Cobb N, *et al.* (2009) Relationship of stand characteristics to drought-induced mortality in piñon-juniper woodlands in Colorado, Arizona and New Mexico. *Ecological Applications* 19: 1223–1230.

Gaines SD and Denny MW (1993) The largest, smallest, highest, lowest, longest, and shortest – extremes in ecology. *Ecology* 74: 1677–1692.

Gray ST, Betancourt JL, Fastie CL, and Jackson ST (2003) Patterns and sources of multidecadal oscillations in drought-sensitive tree-ring records from the central and southern Rocky Mountains. *Geophysical Research Letters* 30.http://dx.doi.org/10.1029/2002GL016154.

Gray ST, Betancourt JL, Jackson ST, and Eddy RG (2006) Role of multidecadal climate variability in a range extension of pinyon pine. *Ecology* 87: 1124–1130.

Gutschick VP and BassiriRad H (2003) Extreme events as shaping physiology, ecology, and evolution of plants: Toward a unified definition and evaluation of their consequences. *New Phytologist* 160: 21–42.

Haddad NM, Tilman D, and Knops JMH (2002) Long-term oscillations in grassland productivity induced by drought. *Ecology Letters* 5: 110–120.

Hallett TB, Coulson T, Pilkington JG, Clutton-Brock TH, Pemberton JM, and Grenfell BT (2004) Why large-scale climate indices seem to predict ecological processes better than local weather. *Nature* 430(6995): 71–75.

Hoerling M and Kumar A (2004) The perfect ocean for drought. *Science* 299: 691–694.

Holling CS (1973) Resilience and stability of ecological systems. *Annual Review of Ecological and Systematics* 4: 1–24.

Holmgren M and Scheffer M (2001) El Niño as a window of opportunity for the restoration of degraded arid ecosystems. *Ecosystems* 4: 151–159.

Holmgren M, Scheffer M, Ezcurra E, Guitérrez JR, and Mohren GMJ (2001) El Niño effects on the dynamics of terrestrial ecosystems. *Trends in Ecology & Evolution* 16: 89–94.

IPCC (2001) Climate Change 2001: Impacts, Adaptation, and Vulnerability. In: McCarthy JJ, Canziani OF, Leary NA, and Dokken DJ (eds.) *Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge UK: Cambridge University Press.

Jackson ST, Betancourt JL, Booth RK, and Gray ST (2009) Ecology and the ratchet of events: Climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Sciences of the United States of America* 106: 19685–19692.

Jentsch A and Beierkuhnlein C (2008) Research frontiers in climate change: Effects of extreme meteorological events on ecosystems. *Comptes Rendus Geoscience* 340: 621–628.

Jentsch A, Kreyling J, and Beierkuhnlein C (2007) A new generation of climate-change experiments: Events, not trends. *Frontiers in Ecology and the Environment* 5: 365–374.

Jentsch A, Kreyling J, Elmer M, *et al.* (2011) Climate extremes initiate ecosystem-regulating functions while maintaining productivity. *Journal of Ecology* 99: 689–702.

Kahmen A, Perner J, and Buchmann N (2005) Diversity-dependent productivity in semi-natural grasslands following climate perturbations. *Functional Ecology* 19: 594–601.

Katz RW and Brown BG (1992) Extreme events in a changing climate: Variability is more important than averages. *Climatic Change* 21: 289–302.

Keeley JE and Bond WJ (1999) Mast flowering and semelparity in bamboos: The bamboo fire cycle hypothesis. *American Naturalist* 154: 383–391.

Kitzberger T (2002) ENSO as a forewarning tool of regional fire occurrence in northern Patagonia, Argentina. *International Journal of Wildland Fire* 11: 33–39.

Kitzberger T, Brown PM, Heyerdahl EK, Swetnam TW, and Veblen TT (2007) Contingent Pacific-Atlantic Ocean influence on multi-century wildfire synchrony over western North America. *Proceedings of the National Academy of Sciences* 104: 543–548.

Kitzberger T, Raffaele E, Heinemann K, and Mazzarino MJ (2005) Effects of fire severity in a north Patagonian subalpine forest. *Journal of Vegetation Science* 16: 5–12.

Kitzberger T, Swetnam TW, and Veblen TT (2001) Inter-hemispheric synchrony of forest fires and the El Niño-Southern Oscillation. *Global Ecology and Biogeography* 10: 315–326.

Kitzberger T and Veblen TT (1997) Influences of humans and ENSO on fire history of Austrocedrus chilensis woodlands in northern Patagonia, Argentina. *Ecoscience* 4: 508–520.

Kitzberger T, Veblen TT, and Villalba R (1997) Climatic influences on fire regimes along a rainforest-to-xeric woodland gradient in northern Patagonia, Argentina. *Journal of Biogeography* 23: 35–47.

Kossin JP, Camargo SJ, and Sitkowski M (2010) Climate modulation of North Atlantic hurricane tracks. *Journal of Climate* 23: 3057–3076.

Kreyling J, Beierkuhnlein C, Ellis L, and Jentsch A (2008) Invasibility of grassland and heath communities exposed to extreme weather events – additive effects of diversity resistance and fluctuating physical environment. *Oikos* 117: 1542–1554.

Landres PB, Morgan P, and Swanson FJ (1999) Overview of the use of natural variability concepts in managing ecological systems. *Ecological Applications* 9: 1179–1188.

Laurance WF and Williamson GB (2001) Positive feedbacks among forest fragmentation, drought, and climate change in the Amazon. *Conservation Biology* 15: 1529–1535.

Lima M, Julliard R, Stenseth NC, and Jaksic FM (2001) Demographic dynamics of a neotropical small rodent (*Phyllotis darwini*): Feedback structure, predation and climatic factors. *Journal of Animal Ecology* 70: 761–775.

Lloret F, Siscart D, and Dalmases C (2004) Canopy recovery after drought dieback in holm-oak Mediterranean forests of Catalonia (NE Spain). *Global Change Biology* 10: 2092–2099.

Maslin M (2004) Ecological versus climatic thresholds. *Science* 306: 2197–2198.

McCabe GJ, Palecki MA, and Betancourt JL (2004) Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proceedings of the National Academy of Sciences* 101: 4136–4141.

Meehl GA, Stocker CM, Bowker TF, *et al.* (2007) Global Climate Projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL (eds.) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Cambridge University Press.

Mermoz M, Kitzberger T, and Veblen TT (2005) Landscape influences on occurrence and spread of wildfires in Patagonian forests and shrublands. *Ecology* 86: 2705–2715.

Miao S, Zou CB, and Breshears DD (2009) Vegetation responses to extreme hydrological events: Sequence matters. *The American Naturalist* 173: 113–118.

Moritz MA (1997) Analyzing extreme disturbance events: Fire in Los Padres National Forest. *Ecological Applications* 7: 1252–1262.

Ojeda VS, Suárez ML, and Kitzberger T (2007) Crown dieback events as key processes creating cavity habitat for magellanic woodpeckers. *Austral Ecology* 32: 436–445.

Ostfeld RS and Keesing F (2000) Pulsed resources and community dynamics of consumers in terrestrial ecosystems. *Trends in Ecology & Evolution* 15: 232–237.

Paine RT, Tegner MJ, and Johnson EA (1998) Compounded perturbations yield ecological surprises. *Ecosystems* 1: 535–545.

Parmesan C, Root TL, and Willig MR (2000) Impacts of extreme weather and climate on terrestrial biota. *Bulletin of the American Meteorological Society* 81: 443–450.

Pfisterer AB and Schmid B (2002) Diversity-dependent production can decrease the stability of ecosystem functioning. *Nature* 416(6876): 84–86.

Potts DL, Huxman TE, Enquist BJ, Weltzin JF, and Williams DG (2006) Resilience and resistance of ecosystem functional response to a precipitation pulse in a semi-arid grassland. *Journal of Ecology* 94: 23–30.

Pueyo S, De Alencastro Graça PML, Barbosa RI, Cots R, Cardona E, and Fearnside PM (2010) Testing for criticality in ecosystem dynamics: The case of Amazonian rainforest and savanna fire. *Ecology Letters* 13: 793–802.

Raffa KF, Aukema BH, Bentz BJ, *et al.* (2008) Cross-scale drivers of natural disturbances prone to anthropogenic amplification: The dynamics of bark beetle eruptions. *BioScience* 58: 501–517.

Ranta E, Kaitala V, Lindström J, and Helle E (1997) The Moran effect and synchrony in population dynamics. *Oikos* 78: 136–142.

Romme WH, Everham EH, Frelich LE, Moritz MA, and Sparks RE (1998) Are large, infrequent disturbances qualitatively different from small, infrequent disturbances? *Ecosystems* 1: 524–534.

Rose MD and Polis GA (1998) The distribution and abundance of coyotes: The effects of allochthonous food subsidies from the sea. *Ecology* 79: 998–1007.

Royer PD, Cobb NS, Clifford MJ, *et al.* (2011) Extreme climatic event-triggered overstorey vegetation loss increases understorey solar input regionally: Primary and secondary ecological implications. *Journal of Ecology* 99: 714–723.

Sala OE, Chapin F, Stuart, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287(5459): 1770–1774.

Scheffer M, Carpenter S, Foley JA, Folke C, and Walker B (2001) Catastrophic shifts in ecosystems. *Nature* 413(6856): 591–596.

Schoennagel T, Veblen TT, Romme WH, Sibold JS, and Cook ER (2005) ENSO and PDO variability affect drought-induced fire occurrence in Rocky Mountain subalpine forests. *Ecological Applications* 15: 2000–2014.

Sethna JP, Dahmen KA, and Myers CR (2001) Crackling noise. *Nature* 410(6825): 242–250.

Slocum MG, Beckage B, Platt WJ, Orzell SL, and Taylor W (2010) Effect of climate on wildfire size: A cross-scale analysis. *Ecosystems* 13: 828–840.

Smith MD (2011) The ecological role of climate extremes: Current understanding and future prospects. *Journal of Ecology* 99: 651–655.

Solé RV, Manrubia SC, Benton M, Kauffman S, and Bak P (2001) Criticality and scaling in evolutionary ecology. *Trends in Ecology and Evolution* 14: 156–160.

Stenseth NC, Ottersen G, Hurrell JW, *et al.* (2003) Studying climate effects on ecology through the use of climate indices: The North Atlantic Oscillation, El Nino Southern Oscillation and beyond. *Proceedings of the Royal Society of London Series B – Biological Sciences* 270: 2087–2096.

Sthultz CM, Gehring CA, and Whitham TG (2007) Shifts from competition to facilitation between a foundation tree and a pioneer shrub across spatial and temporal scales in a semiarid woodland. *New Phytologist* 173: 135–145.

Suarez ML and Kitzberger T (2008) Recruitment patterns following a severe drought: Long-term compositional shifts in Patagonian forests. *Canadian Journal of Forest Research* 38: 3002–3010.

Suarez ML and Kitzberger T (2010) Differential effects of climate variability on forest dynamics along a precipitation gradient in northern Patagonia. *Journal of Ecology* 98: 1023–1034.

Suarez ML, Ghermandi L, and Kitzberger T (2004) Factors predisposing episodic drought-induced tree mortality in *Nothofagus*: Site, climatic sensitivity and growth trends. *Journal of Ecology* 92: 954–966.

Swetnam TW and Betancourt JL (1990) Fire – Southern Oscillation relations in the southwestern United-States. *Science* 249: 1017–1020.

Swetnam TW and Betancourt JL (1998) Mesoscale disturbance and ecological response to decadal climatic variability in the American Southwest. *Journal of Climate* 11: 3128–3147.

Tercero-Bucardo N, Kitzberger T, Veblen TT, and Raffaele E (2007) A field experiment on climatic and herbivore impacts on post-fire tree regeneration in north-western Patagonia. *Journal of Ecology* 95: 771–779.

Tilman D and Downing JA (1994) Biodiversity and stability in grasslands. *Nature* 367: 363–365.

Turcotte DL, Malamud BD, Guzzetti F, and Reichenbach P (2002) Self-organization, the cascade model, and natural hazards. *Proceedings of the National Academy of Sciences* 99: 2530–2537.

Turner MG, Dale VH, and Everham III EH (1997) Crown fires, hurricanes, and volcanoes: A comparison among large-scale disturbances. *BioScience* 47: 758–768.

van Nes EH and Scheffer M (2004) Large species shifts triggered by small forces. *American Naturalist* 164: 255–266.

Van Ruijven J and Berendse F (2010) Diversity enhances community recovery, but not resistance, after drought. *Journal of Ecology* 98: 81–86.

Veblen TT, Kitzberger T, Raffaele E, *et al.* (2008) The historical range of variability of fires in the Andean-Patagonian *Nothofagus* forest region. *International Journal of Wildland Fire* 17: 724–741.

Vitousek PM (1990) Biological invasions and ecosystem processes: Towards an integration of population biology and ecosystem studies. *Oikos* 57: 7–13.

Walter J, Nagy L, Hein R, *et al.* (2011) Do plants remember drought? Hints towards a drought-memory in grasses. *Environmental and Experimental Botany* 71: 34–40.

Wang Y, Yu S, and Wang J (2007) Biomass-dependent susceptibility to drought in experimental grassland communities. *Ecology Letters* 10: 401–410.

White PS and Jentsch A (2001) The search for generality in studies of disturbance and ecosystem dynamics. *Progress in Botany* 62: 399–450.

Williams C, Hanan N, Scholes R, and Kutsch W (2009) Complexity in water and carbon dioxide fluxes following rain pulses in an African savanna. *Oecologia* 161: 469–480.

Williams JW, Blois JL, and Shuman BN (2011) Extrinsic and intrinsic forcing of abrupt ecological change: Case studies from the late Quaternary. *Journal of Ecology* 99: 664–677.

Yachi S and Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences* 96: 1463–1468.

Zhao MS and Running SW (2010) Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* 329: 940–943.

Impact of Past Global Warming on Biodiversity

Gregory J Retallack, University of Oregon, Eugene, OR, USA

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–9,
© 2007, Elsevier Inc.

Glossary

Carbon isotope composition Ratio of common isotopes of carbon (^{12}C and ^{13}C) analyzed in a mass spectrometer and reported in delta notation ($\delta^{13}C$) relative to a standard (usually Peedee belemnite) and reported in parts per thousand (permil), used as a proxy for atmospheric methane abundance.

Chemical index of alteration Molar ratio of alumina over alumina plus soda, potash, lime, and magnesia, determined by whole-rock chemical analysis, and used as a proxy for the residue (alumina) versus solutes (soda, potash, lime, and magnesia) of hydrolytic weathering, used as a proxy for paleoprecipitation.

Methane clathrate Methane (CH_4) gas frozen in ice within deep submarine sedimentary pore space or permafrost.

Oxygen isotopic composition Ratio of common isotopes of oxygen (^{16}O and ^{18}O) analyzed in a mass spectrometer and reported in delta notation ($\delta^{18}O$) relative to a standard (usually Peedee Belemnite) and reported in parts per thousand (permil), used as a proxy for paleotemperature and global ice volume.

Soil Bk horizon Level within a soil of carbonate (calcite or dolomite) enrichment in the form of nodules, mottles, rhizoconcretions, or filaments, used as a proxy for paleoprecipitation.

Soil salinization index Molar ratio of soda and potash over alumina, determined by whole-rock chemical analysis, and used as a proxy for paleotemperature.

Stomatal index Percentage of stomatal openings compared with epidermal cells, used with fossil plant cuticles as a proxy for past atmospheric CO_2 levels.

Introduction

A variety of natural processes appear to have dramatically increased atmospheric greenhouse gases in the geological past: asteroid impact, flood-basalt volcanic eruption, release of submarine or permafrost methane clathrates, and contact metamorphism of coal measures by igneous intrusions (Retallack and Krull, 2006). Abrupt transients of atmospheric CO_2 in the past are discernible from proxies such as the stomatal index of fossil plants (Retallack, 2002) and carbon isotopic composition of pedogenic carbonate (Tabor *et al.*, 2004). Abrupt transients in atmospheric CH_4 are discernible from proxies such as carbon isotopic composition of organic matter (Retallack and Krull, 2006). These indications of atmospheric conditions can be matched with indications of temperature change, such as the oxygen isotopic composition of marine shells (Veizer *et al.*, 2000) and desalinization of soils (Sheldon *et al.*, 2002). Other climatic consequences of CO_2-induced warming include changes in mean annual precipitation from depth to carbonate (Retallack, 2005a) and degree of base depletion of paleosols (Sheldon *et al.*, 2002). These indications of global change can be compared with past records of life, such as biodiversity (Sepkoski, 1996). Indications are already clear that some past global warmings due to increases in greenhouse gases were hardships for life and detrimental to biodiversity.

Proxies for the Past

A critical difference between evidence from the fossil record and studies of the modern environment is the limitation of preservation. Measurements such as air temperature measured directly today must be inferred for the geological past from proxy measurements, which are preserved features of rocks or fossils known to vary with temperature. Such proxy measurements depend for their accuracy on the degree to which they are known to covary with temperature today, and the degree to which they have resisted alteration during burial. There are more such proxies than can be adequately discussed in this short account, which is limited to explaining proxies applied to two time series of global warmings in the Cenozoic (**Figure 1**) and Permian (**Figure 2**).

Paleotemperature

A popular proxy for global marine paleotemperature is the inverse relationship between temperature and oxygen isotopic value ($\delta^{18}O$) of carbonate in shells of marine foraminifera (Zachos *et al.*, 2001), mollusks and brachiopods (Veizer *et al.*, 2000). Water vapor with the light isotope (^{16}O) evaporates from a cold ocean more readily than water vapor with the heavy isotope (^{18}O), thus leaving ocean water and bicarbonate used in shell manufacture isotopically heavier (more positive in standard $\delta^{18}O$ notation permil). This proxy for oceanic paleotemperature is compromised by global ice volumes, which are a separate reservoir dominated by the light isotope (^{16}O). As either a global ice volume or local temperature signal, marine biogenic carbonate oxygen isotopic composition is an important global change indicator (**Figure 1b**). Disentangling the different signals of ice volume and temperature from oxygen isotope values requires independent proxies for ice volume or temperature, which have their own problems. Another problem is burial recrystallization of fossils,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00232-X

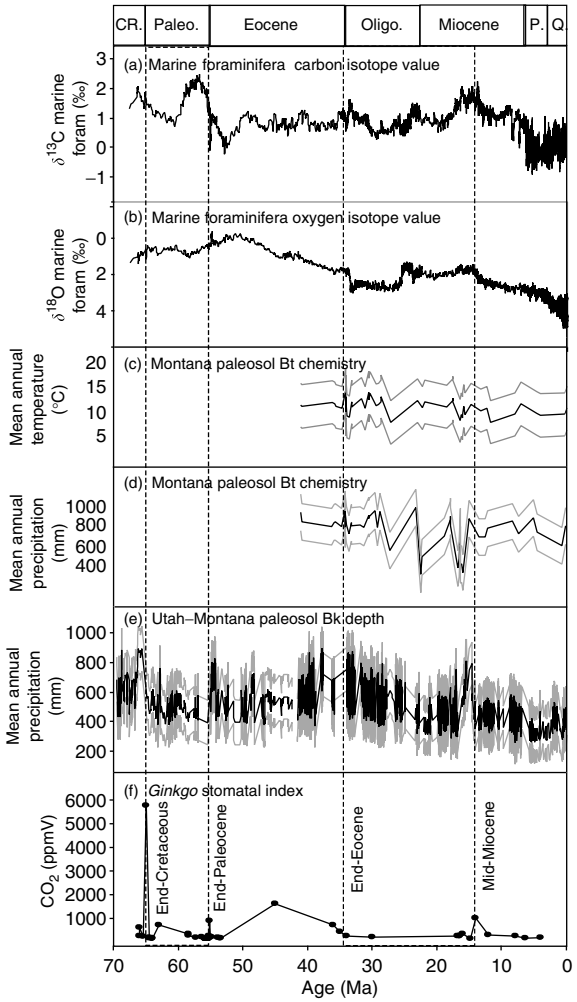


Figure 1 Cenozoic global change indicators: (a) carbon isotopic composition of benthic foraminifera (Zachos *et al.*, 2001), (b) oxygen isotopic composition of benthic foraminifera (Zachos *et al.*, 2001), (c) mean annual temperature of paleosols in Montana inferred from paleosol salinization index (Sheldon *et al.*, 2002), (d) mean annual precipitation of paleosols in Montana inferred from paleosol chemical index of alteration (Sheldon *et al.*, 2002), (e) mean annual precipitation of paleosols in Utah and Montana inferred from depth to Bk horizon in paleosols (Retallack, 2005b), (f) CO_2 concentration in the atmosphere estimated from the stomatal index of fossil *Ginkgo* leaves (Retallack, 2002). Times of high CO_2 are generally warm and wet.

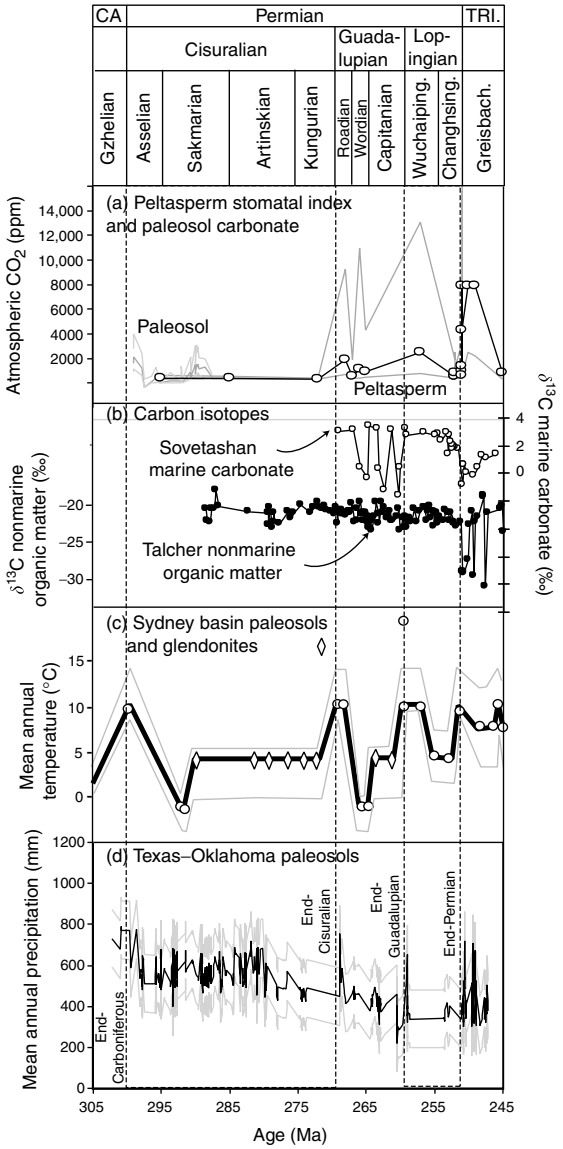


Figure 2 Permian global change indicators: (a) CO_2 concentration in the atmosphere estimated from the stomatal index of fossil *Lepidopteris* leaves (Retallack, 2005c) and carbon isotopic composition of paleosols (Tabor et al., 2004), (b) carbon isotopic composition of marine limestone from Sovetashan, Armenia (Baud *et al.*, 1989) and of organic carbon from Talcher, India (de Wit *et al.*, 2002), (c) paleotemperature indicated by glendonites and paleosols from the Hunter Valley, New South Wales, Australia (Retallack, 2005c), (d) paleoprecipitation from depth to Bk horizons in paleosols of Oklahoma and Texas, USA (Retallack, 2005c). Times of high CO_2 are generally warm and wet, and some are also times of significant negative carbon isotope excursions.

especially small, porous fossils such as foraminifera. Observations of clam or brachiopod shells in petrographic thin sections or the scanning electron microscope for original biogenic microstructure are sufficient to rule out burial alteration (Veizer *et al.*, 2000).

A proxy for paleotemperature on land is based on observations that warmer climates have greater evapotranspiration than colder climates, and thus leach more alkali cations from soils. The soil salinization index, or molar ratio of potash plus soda to alumina in the clay-enriched subsurface (Bw or Bt) horizons of soils, shows a clear inverse relationship with mean annual temperature in North American soils (Sheldon *et al.*, 2002). Such estimates of paleotemperature from paleosols show trends broadly comparable with marine isotopic records (**Figure 1c**). This paleosol proxy for paleotemperature can be compromised by burial alteration of soil clay mineral composition, particularly illitization, which can be assessed by X-ray diffractograms of clay (Retallack, 2001).

Other indications of paleotemperature on land are not so fully quantified, but nevertheless locally useful (**Figure 2c**). Deeply weathered kaolinitic soils (Ultisols) are not found at mean annual temperatures lower than 12 °C, whereas soils

with ice disruption (Gelisols) are found in regions with mean annual temperature no more than 0 °C. Similarly, ikaite crystals ($CaCO_3$ $6H_2O$) or their pseudomorphs (glendonites) are large spindle-shaped structures within sediment of lake or ocean bottoms, and unstable in waters warmer than 5 °C (Retallack, 2005c).

Paleoprecipitation

The increased depth to carbonate nodules (Bk horizon) in modern soils with increasingly humid climates was first documented eastwards from the Rocky Mountains, USA. This relationship was subsequently shown to be a global phenomenon, and attributed to increases in acidifying soil respiration with rainfall (Retallack, 2005a). Long sequences of paleosols show considerable variation in paleoprecipitation (**Figures 1e** and **2d**). This measure of paleosols is compromised by compaction due to burial, but corrections are applied based on unfolding of formerly planar features or on thickness of overburden. Corrections for different levels of atmospheric carbon dioxide can also be made, but these are not very large (40 mm per annum more rainfall for five times preindustrial level of CO_2), because atmospheric variations in carbon dioxide levels are dwarfed by soil respiration in most soils.

Another proxy for paleoprecipitation is based on observations that wet climate soils are more leached of alkali and alkaline earth cations than dry climate soils. The chemical index of weathering without potash, or molar ratio of alumina, lime, magnesia, and soda to alumina in the clay-enriched subsurface (Bw or Bt) horizons of soils shows a clear relationship with mean annual precipitation in North American soils (Sheldon *et al.*, 2002). This geochemical proxy shows trends comparable with the physical proxy of calcic horizon depth in the same paleosols (**Figure 1d**). Leaving out potash, a common soil cation, obviates problems with burial illitization (Retallack, 2001). The chief problems with this proxy are local refreshing of soil fertility, for example, by limestone colluvium or volcanic ash, which are best detected through petrographic thin sections of the paleosols.

Atmospheric CO_2 and CH_4

A useful proxy for atmospheric CO_2 is the stomatal index of fossil leaves. This paleobotanical paleobarometer is based on the observation that plant leaves have fewer stomates when atmospheric CO_2 is high than when atmospheric CO_2 is low (Wynn, 2003). Stomatal index is the number of stomatal openings as a percent of epidermal cell numbers, and is preferable to measures of stomatal density, because it is less affected by competing effects of aridity, salinity, and soil nutrient deficiency. A major limitation of this approach is that sensitivity to carbon dioxide is taxon-specific, and few taxa span long periods of geological time. The living fossil *Ginkgo biloba* has been used as a standard for atmospheric CO_2 estimates from congeneric fossils back some 230 My (Retallack, 2002). Other fossils of *Lepidopteris* in the same 230–200 million-year-old deposits as *Ginkgo* fossils have similar stomatal anatomy and index, and can be used to extend atmospheric CO_2 estimates back 300 My (**Figures 1f**, **2a** and **3**). The principal limitation of this approach is its requirement of preserved plant cuticles preparable in sufficient volume to obtain statistically valid stomatal index counts (Retallack, 2002).

Another proxy for atmospheric CO_2 is the carbon isotopic composition of pedogenic carbonate. Plant-respired CO_2 in soil pores has much more light carbon (^{12}C) than heavy carbon (^{13}C) because of the very marked preference of photosynthesis for the light isotope. Animal and microbe-respired CO_2 in soil pores is also isotopically light because the food chain starts with plant material. In contrast, the isotopic composition of CO_2 in the atmosphere is isotopically heavy CO_2 degassed from volcanoes, and the mixing of these two end-members is a function of atmospheric CO_2 partial pressures. The approach is compromised by effects of soil temperature and productivity, which can in part be compensated by geochemical and petrographic study of paleosols (Tabor *et al.*, 2004). A more serious limitation is temporal variation in isotopic composition of the atmosphere, which can be approximated from analysis of organic matter of the same age as the soil carbonate (Retallack, 2002). This requires large samples because carbonate nodular soils of dry climates are lean for organic matter. This is a potentially voluminous source of past atmospheric CO_2 estimates, because calcareous paleosols are common (**Figure 2a** and **d**).

A proxy for atmospheric CH_4 is the carbon isotopic composition of organic matter, such as fossil leaves, wood, humus, or dispersed organic matter. The carbon isotopic values of methane reflect an unusual dominance of ^{12}C over ^{13}C. Whereas most organic matter ranges from $-22‰$ to $-30‰$ $\delta^{13}C$, thermogenic CH_4 is $-35‰$ to $-55‰$, typically $-40‰$, and biogenic CH_4 is $-35‰$ to $-120‰$, typically $-60‰$ (Clayton, 1998). These measurable differences are preserved in plants because atmospheric CH_4 oxidizes within about a decade to isotopically unusual CO_2, which is then used for photosynthesis, and buried with the death of the plant. The carbon isotopic anomaly at the Permian–Triassic boundary is so profound in both marine and nonmarine rocks worldwide (**Figure 2b**) that release to the atmosphere of some 2000 Gt of methane are required (Berner, 2002). This is an enormous atmospheric injection of a mass of methane equivalent to the carbon in all current terrestrial biomass and soil carbon (1 Gt is 10^{15} g). A limitation with this proxy is that the amount of methane can only be calculated if an original isotopic composition is assumed: Berner's estimate was calculated for biogenic values of –60‰, and is low compared with estimates assuming thermogenic values of –40‰. Sources of such large amounts of methane include outburst from submarine methane clathrates and thermal alteration of coals by volcanic intrusions (Retallack and Krull, 2006).

Biodiversity

Past biodiversity is estimated from the numbers of families, genera, or species of fossils within particular geological periods or stages. For truly global coverage such estimates come from literature compilations, for example the *Treatise of Invertebrate Paleontology*, and now increasingly from computer databases. The most robust of these is the record of skeletonized marine invertebrates (Sepkoski, 1996), but estimates

of plant and vertebrate diversity are also available (Benton, 1995). Such proxies of biodiversity suffer from fundamental limitations of preservational bias and study bias. They are clearly an underestimate of past diversity because many weakly skeletonized and soft-bodied creatures are preserved only in rare amber or black shales. There is also a clear bias between groups, with trilobites, for example, attracting more taxonomic attention than corals. Some periods of time are better represented by rocks than others, because of changes in sea level, and so have inflated diversity compared with less well-represented times.

Correlation Between Biodiversity and Global Warming

Many data on paleotemperature and paleoprecipitation are now available, and no clear relationship emerges between them and biodiversity. This is to be expected because climate is so heterogeneous around the world that climate change may induce diversity loss in one part of the world but diversity gains elsewhere. An exception is atmospheric CO_2 levels and marine generic extinction (**Figure 3**), which shows positive correlation (**Figure 4**). With 42 degrees of freedom and a t value of 6.7, this relationship is highly significant ($p < 0.0001$).

The relationship between CO_2–CH_4 greenhouse and extinction is best understood in extreme cases, such as the End-Permian mass extinction, widely acknowledged to be the greatest life crisis in Earth history (Sepkoski, 1996). An abrupt increase in atmospheric CO_2 is indicated at 250 Ma by stomatal index decline of seed ferns (**Figure 2a**), and contemporaneous transient CH_4 outburst is indicated by markedly lower carbon isotopic composition of organic matter and marine carbonate worldwide (**Figure 2b**). Modeling of this atmospheric pollution event with reducing gases may have resulted in geologically rapid (over some 20,000 years) drawdown of atmospheric O_2 from levels of 35%, which is higher than present levels of 21%, to levels as little as 12% (Berner, 2002). Such an atmospheric redox crisis has suggested several mechanisms of extinction for different kinds of organisms.

In the sea, the most lethal effect would have been excess CO_2 (hypercapnia) and low O_2 (hypoxia). This would adversely affect skeletonized organisms with limited ventilatory capacity such as corals and brachiopods, which were mostly extinct, rather than more muscular bivalves and ammonoids, which survived in great numbers but limited diversity. Corals may also have been affected by postapocalyptic global warming, inducing expulsion of zooxanthellae and coral bleaching. No coral reefs are known for the first 6 My of the Triassic, a conspicuous "reef gap" in the fossil record of coral reefs extending back to the Ordovician.

Among land plants, the most lethal effect would be low O_2 (hypoxia) in soil solution, endangering root respiration, especially in swampy soils already challenged for aeration by waterlogging. This would adversely affect swamp plants such as glossopterids and lycopsids, which became extinct, with less-severe effects on surviving horsetails, true ferns, seed ferns, and conifers. No peat swamps are known for the first 6 My of the Triassic, as indicated by a "coal gap," which is the most profound break in the record of coals since their first appearance in Devonian rocks.

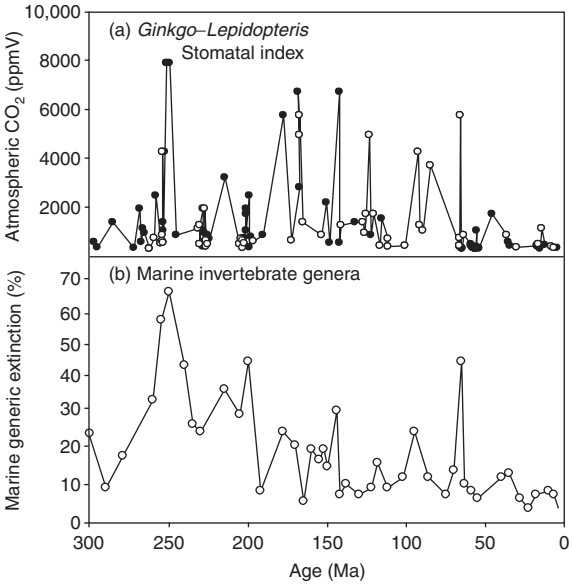


Figure 3 (a) Records of atmospheric CO_2 at geological stage boundaries from stomatal index of *Ginkgo* and *Lepidopteris* and (b) of percentage generic extinction at those same geological stage boundaries from compilations of shelly marine invertebrate taxonomy for the past 300 My.

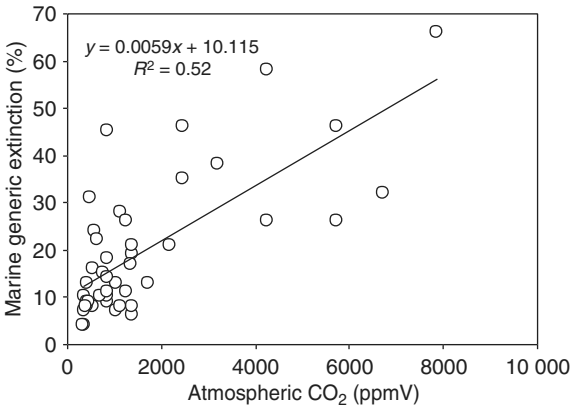


Figure 4 Cross-correlation between CO_2 and extinction levels at geological stage boundaries shown in **Figure 3**.

Among land vertebrates, atmospheric O_2 depletion (hypoxia) would have resulted in conditions at sea level more like those found at high altitude today, leading to death by pulmonary edema. Surviving vertebrates were primarily small burrowing species, preadapted to poorly ventilated spaces, and with bony palate, large chests, and short limbs common in high-altitude vertebrates today (Retallack and Krull, 2006).

The Permian–Triassic boundary crisis was extreme, a global warming and atmospheric redox perturbation of unusual magnitude. Other comparable events of lesser magnitude (**Figures 1–3**) had less-severe effects on biodiversity (**Figure 4**). A base level of about 10% generic extinction at very low levels of atmospheric CO_2 reflects typical differences in biota from one geological stage to another, and may be an artifact of

long spans of geological time used to calculate extinction rates. Nevertheless, background rates of biodiversity loss grade insensibly into global mass extinctions.

Environmental Consequences of Past Global Warming

The 18 events of global greenhouse in the past 300 My (**Figure 3**) give a new perspective on the current anthropogenic greenhouse. Instead of a unique event, current global warming can be considered among an array of earlier global warming events showing a spectrum of severity. Consistent environmental correlates of global warming can now be observed in multiple examples. Past global warmings furnish a variety of environmental expectations for global warmings of the future.

Soil Acidification

Increased atmospheric CO_2 creates greater volumes of carbonic acid (H_2CO_3), which is the chief proton source for hydrolytic weathering. Atmospheric acidity is minor compared with the effect of soil respiration, which can raise soil CO_2 levels to 100 times that of the atmosphere and generate large volumes of carbonic acid. Furthermore, increases in rainfall (**Figure 1d, e,** and **2d**) and temperature (**Figures 1c**, and **2c**) with higher atmospheric CO_2 levels raise both primary and secondary productivity of soils, and thus soil respiration. Thus, a variety of related effects accelerate chemical weathering at times of high atmospheric CO_2, more rapidly depleting soils of alkaline soil nutrient cations such as Ca^{2+}, Mg^{2+}, K^+, and Na^+. The result is the spread of base-poor kaolinitic soils (Ultisols and Oxisols) and soil-weathering products (bauxites) to high paleolatitudes at times of CO_2 greenhouses (Retallack, 2001).

Marine Alkalinization

Increased chemical weathering at times of high CO_2 has the effect of adding more soluble ions to groundwater and ultimately the sea. The ocean is thus buffered with bases, mainly bicarbonate (HCO_3^-), which is the conjugate base to carbonic acid. Also enhanced in groundwater and the ocean are basic cations (Ca^{2+}, Mg^{2+}, K^+, and Na^+), which are important biological nutrients. Moderate levels of oceanic fertilization would have promoted coral and shellfish growth, and thus limestone formation. High levels of fertilization promote seafloor anoxia, and pyritic black shale formation, commonly seen at times of global warming. Extreme alkalinization may explain abiotic precipitation of seafloor fans of calcite (Knoll *et al.*, 1996).

More Humid Climate

Global warming induced by CO_2 increases the water vapor holding capacity of the atmosphere. Water vapor itself is a powerful greenhouse gas, thus amplifying CO_2-induced warming. Unlike CO_2, however, water is regularly rained out of atmosphere and so is not a primary control of greenhouse warming. Thus, it is not surprising to find evidence from both chemical weathering and depth of calcic horizons in paleosols that times of high CO_2 were also times of greater mean annual precipitation (**Figure 1d, e,** and **2d**).

Enhanced Climatic Seasonality

Tropical monsoonal climates have very pronounced dry seasons interspersed with short deluges of rain, and soils of such climate have a distinctive spread of pedogenic nodules throughout the soil profile. Less seasonal climates in contrast have carbonate nodules within well-defined (Bk) horizons. Thickness of soil with nodules is a proxy for the difference between the mean precipitation of the driest and wettest month, and seasonality of precipitation as well as mean annual precipitation both increase during atmospheric CO_2 spikes (Retallack, 2005a).

Increased Runoff

Increased rainfall at times of atmospheric CO_2 spikes would be expected to result in increased stream runoff, with the complication that increased rainfall also appears to increase plant productivity and stature and thus soil moisture retention. In the extreme case of the Permian–Triassic boundary, increased runoff in a postapocalyptic greenhouse of the earliest Triassic is apparent from a change in stream architecture from meandering to braided, and appearance of beds of eroded soil clods (Retallack, 2005b). Such changes are not so obvious in lesser atmospheric CO_2 spikes, such as the middle Miocene warm-wet spike in Montana (**Figure 1c–e**).

Rising Sea Level

Shrinking ice caps and thermal expansion of the ocean leads to marine transgressions at times of high atmospheric CO_2. The record of sea level change through time supports this deduction (Haq and Al-Qahtani, 2005). Especially striking are expansions of large epicontinental seaways on stable continental platforms. In Utah, for example, most Mesozoic rocks are nonmarine, but these red eolian, fluvial, and pedogenic red beds include thin marine limestones: Sinbad Limestone (Smithian, 249 Ma), Virgin Limestone (Spathian, 247 Ma), Carmel Formation (Callovian, 164 Ma), and Tropic formation (Cenomanian, 94 Ma). These were all times of high atmospheric CO_2 (**Figure 3**).

Shrinking Cryosphere

Global warming results in melting of polar and alpine ice caps and glaciers, and surrounding regions of periglacial soils. In accord with this view, glacial tillites, varved shales, and dropped pebbles are limited in temporal distribution to well-defined "Ice Ages." The ice age of the past 1.6 My is presaged by high-latitude ice-rafted debris as old as 30 Ma. With the exception of some possible Valanginian (138 Ma) ice-rafted debris in South Australia there are no reported glacial deposits, even at high latitudes as far back as Late Permian (258 Ma). The Late Devonian to Middle Permian Ice Age, well known especially from the high-latitude southern supercontinent of

Gondwana is at a time of low atmospheric CO_2, judging from stomatal index (**Figure 3**). From then back to another ice age of the Late Ordovician, there is again little evidence of glacial facies (Hambrey and Harland, 1981), and indications from paleosols of high CO_2 (Mora *et al.*, 1996). The fossil record of ice-disrupted soils, or Gelisols, and associated tundra and taiga ecosystems, also supports evidence from glacial facies for a reduced cryosphere at times of global warming (Retallack, 2001).

Biotic Consequences of Past Global Warming

The various environmental effects of global warming outlined above are closely interrelated with a variety of biotic changes recorded in the fossil record. As with environmental consequences, many biotic consequences are interrelated.

Reduced Plant Transpiration

Times of high CO_2 are not only times of low stomatal index (**Figures 1f**, and **2a**), but also times of reduced stomatal aperture and of increased cuticle thickness of leaves (Retallack, 2002). These effects appear to be direct plant responses to high CO_2 levels, and lessened demand for air intake at the stomate. The effect of all these changes is to reduce transpiration of water by vegetation, which in turn lowers atmospheric moisture, raises soil water tables, and increases stream runoff (Beerling, 2005).

Increased Plant Sclerophylly

Times of high CO_2 are marked by an abundance of photosynthetic shoots rather than leaves, of needle-leaves rather than broadleaves, and of densely veined leaves rather than loosely veined leaves (Beerling, 2005; Retallack, 2005c). These various indications of increased sclerophylly may be a response to declining soil nutrient status with soil acidification or to increased plant consumption by animals and pathogens.

Increased Leaf Herbivory

Large collections of fossil leaves commonly contain examples of a variety of insect and fungal damage, but the proportion of damaged leaves is greater at times of high CO_2 (Wilf *et al.*, 2001). Increased abundance of fungi at times of high CO_2 is also indicated by sclerotinite in coals and spores and hyphae in palynofloras (Retallack and Krull, 2006). This may be a secondary effect of the greater primary and secondary productivity ecosystems under warmer and wetter conditions, or perhaps also a consequence of the spread to higher latitudes of animals and pathogens formerly limited by seasonal freezing.

Tropical Plant Geographic Expansion

The spread into polar regions of tropical plants such as palms with large frost-insensitive terminal meristems at times of high CO_2 is probably a consequence of global warming. Such geographic expansion of tropical plants has been especially important for North American broadleaf vegetation, which contains many Asiatic and European taxa which migrated via Beringian and North Atlantic land bridges and islands (Wing, 1998).

Tropical Animal Geographic Expansion

Times of high CO_2 and mild polar climates also enabled migration to high latitudes of tropical animals such as primates and crocodiles (Markwick, 1994). As for plants, such migrations also allowed transcontinental biotic interchange via high-latitude land bridges such as Beringia (Clyde and Gingerich, 1998).

Helical Animal Burrows

Helical animal burrows are rare in the fossil record, presumably requiring more effort than straight ramps. Helical burrows are common at times of high CO_2 and global warming, perhaps because helicoidal airflow is self-ventilating (Meyer, 1999).

Wetland Eutrophication

Wetland paleosols at times of high CO_2 include minerals such as siderite and berthierine indicating unusually anoxic conditions in close association with root traces and burrows (Sheldon and Retallack, 2002). Atmospheric hypoxia may explain these differences, but so could increased microbial productivity under warmer and wetter conditions, because microbial respiration depletes oxygen from stagnant groundwater.

Oceanic Eutrophication

Times of high CO_2 are also times of black, pyritic shales with paper clams and other fossils tolerant of hypoxia (Knoll *et al.*, 1996), reduced area and extinction of tropical coral reefs (Wiedlich *et al.*, 2003), and blooms of microfossil dinoflagellates and diatoms (Siesser, 1995). These geological observations could be consequences of oceanic eutrophication with nutrient cations from increased terrestrial weathering, but also plausible is direct chemical reduction from an hypoxic atmosphere (Retallack and Krull, 2006), coral death by bleaching as zooxanthellae were expelled in warming water (Wiedlich *et al.*, 2003), and phytoplankton blooms tracking increased climatic seasonality (Siesser, 1995).

Conclusions

Anthropogenic global warming due to atmospheric pollution by CO_2 and CH_4 is commonly regarded as unprecedented, but the rock and fossil record indicates many comparable past atmospheric redox crises. Commonalities among these past events allow a better understanding of global warming induced by greenhouse gases. Biodiversity generally declined during geologically rapid global warmings due to CO_2 and CH_4 emissions from volcanic eruption and intrusion, meteorite impacts, or outbursts from submarine and permafrost

methane clathrates. Extreme cases of greenhouse release to the atmosphere are the great mass extinctions of the fossil record, when many creatures died from the effects of hypoxia, hypercapnia, and pulmonary edema. At lower levels of greenhouse gases, such lethal effects were limited in area to high altitudes, stagnant wetlands, oligotrophic coral reefs, and deep ocean basins. Some consequences of global warming such as increased precipitation and warmth extending to high-latitude intercontinental land bridges had the effects of increasing biodiversity. Hardships such as soil acidification and increased seasonality of climate selected for sclerophyllous, thickly cutinized plants. Hardships such as hypoxia selected for greater respiratory scope in animals. Such adaptations added to biodiversity. The effects on biodiversity of global warming by greenhouse gases are complex, but a rich fossil and rock record of comparable events in the past promises new perspectives on future global warming.

See also: Atmospheric Gases

References

Baud A, Magaritz M, and Holser WT (1989) Permian–Triassic of the Tethys; carbon isotope studies. *Geol. Rundsch* 78: 649–677.

Beerling DJ (2005) Evolutionary responses of land plants to atmospheric CO_2. In: Ehleringer JR, Cerling TE, and Dearing MD (eds.) *A History of Atmospheric CO_2 and its Effects on Plants, Animals and Ecosystems*, pp. 114–132. Berlin: Springer.

Benton MJ (1995) Diversification and extinction in the history of life. *Science* 268: 52–58.

Berner RA (2002) Examination of hypotheses for the Permo–Triassic boundary extinction by carbon cycle modeling. *Proc. Natl. Acad. Sci. USA* 99: 4172–4177.

Clayton JL (1998) Geochemistry of coalbed gas; a review. *Int. J. Coal Geol.* 35: 159–173.

Clyde WC and Gingerich PD (1998) Mammalian community response to latest Paleocene thermal maximum: An isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology* 26: 1011–1014.

Hambrey MJ and Harland WB (eds.) (1981) *Earth's Pre-Pleistocene Glacial Record*, 1004 pp. Cambridge: Cambridge University Press.

Haq BU and Al-Qahtani AM (2005) Phanerozoic cycles of sea-level change on the Arabian Platform. *GeoArabia Manama* 10: 127–160.

Knoll AH, Bambach RK, Canfield DE, and Grotzinger JF (1996) Comparative Earth history and the late Permian mass extinction. *Science* 273: 16–17.

Markwick PJ (1994) "Equability," "continentality," and Tertiary "climate": The crocodilian perspective. *Geology* 22: 613–616.

Meyer RC (1999) Helical burrows as a response to paleoclimate: Daimonelix by Paleocastor. *Palaeogeogr. Palaeoclim. Palaeoecol* 147: 191–198.

Mora CI, Driese SG, and Colarusso LA (1996) Middle to late Paleozoic atmospheric CO_2 levels from soil carbonate and organic matter. *Science* 271: 1105–1107.

Retallack GJ (2001). *Soils of the Past: An Introduction to Paleopedology*, 2nd ed., p. 600. Oxford: Blackwell.

Retallack GJ (2002) Carbon dioxide and climate over the past 300 Myr. *Proc. Roy. Soc. Lond. Philos. Trans* A360: 659–674.

Retallack GJ (2005a) Pedogenic carbonate proxies for amount and seasonality of precipitation in paleosols. *Geology* 33: 333–336.

Retallack GJ (2005b) Earliest Triassic claystone breccias and soil-erosion crisis. *J. Sedimentary Res* 75: 679–695.

Retallack GJ (2005c) Permian greenhouse crises. In: Lucas SG and Zeigler KE, (eds.) *The Nonmarine Permian,* vol. 30, pp. 256–269. New Mexico Museum of Natural History and Science Bulletin.

Retallack GJ and Krull ES (2006) Carbon isotopic evidence for termian–Permian methane outbursts and their role in extinction of animals, coral reefs and peat swamps. In: Greb SF and DiMichele WA, (eds.) *Wetlands Through Time,* vol. 399, pp. 249–268. Geological Society of America Special Paper.

Sepkoski JJ (1996) Patterns of Phanerozoic extinction: A perspective from global data bases. In: Walliser OH (ed.) *Global Events and Event Stratigraphy in the Phanerozoic*, pp. 35–51. Berlin: Springer.

Sheldon ND and Retallack GJ (2002) Low oxygen levels in earliest Triassic soils. *Geology* 30: 919–922.

Sheldon ND, Retallack GJ, and Tanaka S (2002) Geochemical climofunctions from North American soils applied to paleosols across the Eocene–Oligocene boundary in Oregon. *J. Geol* 110: 687–696.

Siesser WG (1995) Paleoproductivity of the Indian Ocean during the Tertiary period. *Global Planet. Change* 11: 71–88.

Tabor NJ, Yapp CJ, and Montañez IP (2004) Goethite, calcite, and organic matter from Permian and Triassic soils; carbon isotopes and CO_2 concentrations. *Geochim. Cosmochim. Acta* 68: 1503–1517.

Veizer J, Godderis Y, and François LM (2000) Evidence for decoupling of atmospheric CO_2 and global climate during the Phanerozoic eon. *Nature* 408: 698–701.

Wiedlich O, Kiessling W, and Flügel E (2003) Permian–Triassic boundary interval as a model for forcing marine collapse by long-term atmospheric oxygen drop. *Geology* 31: 961–964.

Wilf P, Labandeira CC, Johnson KR, Coley PD, and Cutter AD (2001) Insect herbivory, plant defense and early Cenozoic climate change. *Proc. Natl. Acad. Sci. USA* 98: 6221–6226.

Wing SL (1998) Late Paleocene–early Eocene floral and climatic change in the Bighorn Basin, Wyoming. In: Aubry M-P, Lucas SG, and Berggren WA (eds.) *Paleocene–Early Eocene Climatic and Biotic Events in the Marine and Terrestrial Records*, pp. 308–400. New York: Columbia University Press.

de Wit MJ, Ghosh JG, de Villiers S, *et al.* (2002) Multiple organic carbon isotope reversals across the Permo–Triassic boundary of terrestrial Gondwana sequences; clues to extinction patterns and delayed ecosystem recovery. *J. Geol* 110: 227–246.

Wynn JG (2003) Towards a physically based model of CO_2-induced stomatal frequencey response. *New Phytol* 157: 391–398.

Zachos J, Pagani M, Sloan L, Thomas E, and Billups K (2001) Trends, rhythms, and aberrations in global climate, 65 Ma to present. *Science* 292: 686–693.

Implications of Urbanization for Conservation and Biodiversity Protection

Robert Ian McDonald, The Nature Conservancy, Arlington, VA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Compact city theory The theory that more compact cities, with greater population density, are more efficient, use less energy per capita, and have less environmental impact per capita than less compact cities.

Ecosystem services The positive benefits that humanity obtains from ecological processes.

Sprawl inequality The idea that the majority of habitat loss from urban expansion is due to a minority of new housing unit creation, typically at the edges of the urban area. Similarly, a minority of households live on parcels that cover the majority of the urban area.

Urbanization Generally, the process of urban growth. Often used more specifically to refer to an increase in the proportion of total population that lives in cities.

Urban sprawl Urban development that occurs at a relatively low population density and/or takes a large amount of land.

Why Think About Cities?

An Urban World

Homo sapiens sapiens is now predominately an urban-dwelling creature. For the first time in human history, a majority (50.4%) of humanity lives in urban areas (**Figure 1**). By 2050, there will likely be 2.8 billion new urbanites (UNPD, 2009), a dramatic rate of urban population growth that, if it could be compressed to one place on the globe, is the equivalent of building a city the size of San Antonio, Texas, every week (cf. McDonald, 2008). In tens of thousands of cities across the globe, bricks are being laid, concrete is being poured, and new businesses and households are being formed. This article aims to answer a simple question: What are the main implications of this massive urban growth on biodiversity, ecosystem services, and conservation?

Demographic Realities

Urbanization, the increase over time in the proportion of a population that lives in urban areas, is strongly correlated with economic development and decrease in fertility. This shift is often called the demographic transition. Poor, predominately rural societies have a high fertility rate. As economic development occurs, a greater and greater proportion of the population lives in cities, and various social and economic incentives encourage a shift to smaller family size. It is generally not clear whether economic development causes urbanization or vice versa; the two processes are often inextricably linked (Montgomery *et al.*, 2003).

Note that urban growth, in terms of an increase in the absolute number of people living in a city, is technically different than urbanization. Urban population change, like that of all populations, is the sum of the natural increase (births minus deaths) and net migration (immigration minus emigration). The growth of most cities' population is due to both natural increase and migration, although the relative strength of the two effects varies by society. However, most cities have lower rates of natural increase than rural areas, so if economic development is to increase the proportion of people that live in cities, significant rural-to-urban migration must occur (Montgomery, 2008).

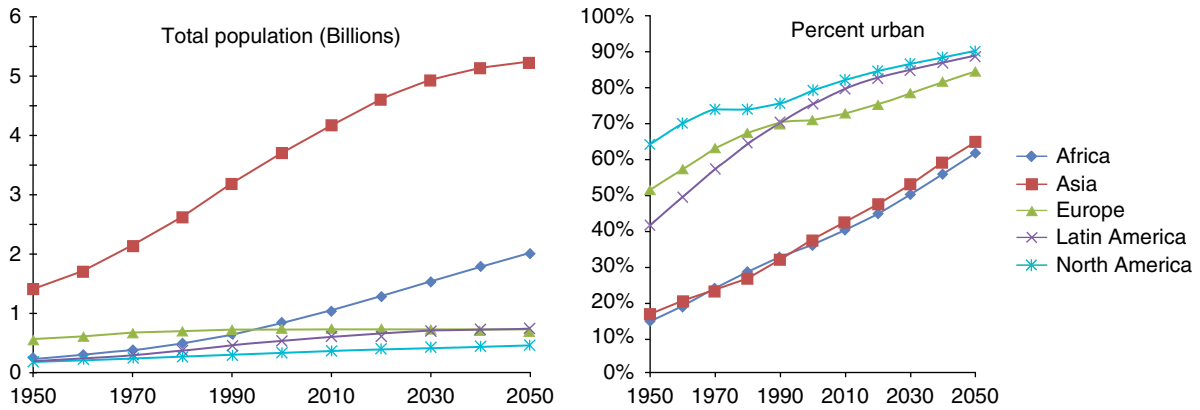


Figure 1 Trends in population over time, by region. Left panel: Total population, both rural and urban, over time. Only Africa and Asia are still substantially increasing. Right panel: Percent of the population in urban areas. Africa and Asia are significantly more rural than other regions. Data from the UNPD (2009) *World Urbanization Prospects*: *The 2009 Revision*. NewYork: United Nations Population Division.

From the perspective of an individual in the countryside, why move to the city? There are generally better employment opportunities and better wages (Anas *et al.*, 1998). At present, there is also often better health care in cities, although this has not always been the case – in the nineteenth century, European dwellers were substantially less healthy than those in the countryside, but in developing countries today the converse is true (Mace, 2008). There are generally better educational opportunities in cities, and children have a greater variety of potential occupation to choose from. Collectively, these drivers of urbanization (Seto *et al.*, 2010) are sometimes called "pull" factors, to contrast with factors that occasionally push people from rural areas to cities (e.g., war, famine, and political instability).

Why set up shop in a city from the perspective of a firm? One important reason is that firms gain tangible economic benefits by being near other firms (Seto *et al.*, 2010). These agglomeration economies are one of the fundamental factors driving urban development (Krugman, 1998). For industries that move significant quantities of inputs or outputs, transport costs can be minimized by being near suppliers or customers (von Thünen, 1875). A factory making automobiles, for example, may benefit by being near plants making steel or aluminum.

Another agglomeration economy is the increased ease of communication that comes with proximity. Some sectors, despite not having significant quantities of physical inputs and outputs, tend to cluster (Sassen, 2001). Government and corporate bureaucracies, for example, find communication easier when employers are near one another. Creative endeavors, like film production, often cluster in certain towns because it is cheaper to shoot a film or produce a play if all the necessary staff are in one place (Florida, 2002). Even writing a script or a play is easier when there are other writers around to trade ideas with.

There are also costs that both individuals and firms consider for being in a city. The cost of renting or buying land is significantly higher in a city (McDonald, 2009). There are also negative effects of crowding, such as the productivity lost while workers are stuck in traffic. The rural versus urban split of individuals and firms is ultimately a result of the balance between the positive and negative economic effects of agglomeration in cities. Similarly, within urban areas, where individuals or firms locate is an attempt to balance these same positive and negative effects (Van Ommeren and Rietveld, 2005). As transportation and building technologies change, this economic balance shifts. For instance, highways and new communication technologies like phones and e-mail have made it more profitable for many firms to be at the edges of urban cores, since they have lower costs of doing business but can still gain some benefit from being near a city (Bairoch, 1988).

So far in this discussion one tricky issue has been skipped: How does one define "urban" and "rural?" The ideal definition might vary depending on the issue under study (Montgomery *et al.*, 2003): a researcher studying employment might define cities based on commuting distance to major urban employment centers, while an ecologist studying habitat loss might use a measure of urban land cover. Sometimes political boundaries are used, although in many parts of the world these do not correspond with the entire metropolitan region. Population density is often used as a definition of urban areas, although population density often varies substantially from an urban core to the suburbs and exurbs of a city. In most metropolitan regions, suburban and exurban land-uses cover the majority of the land area. However, the majority of the population lives at greater density in the core of the city.

What is the Alternative?

To talk sensibly about urbanization and the environment, one has to specify the baseline or counterfactual scenario: What would have happened to the environment without the rural-to-urban migration (McDonald and Marcotullio, 2011)? This involves considering the distribution of population and associated environmental impact, under some baseline scenario. Without migration to cities, there will be less environmental impacts in cities, but more impacts in the countryside. Moreover, less immigration will potentially limit economic development, and potentially result in a larger total population since rural fertility rates are generally higher.

A simple conceptual model may illustrate the environmental effects of more or less rural-to-urban migration over some (relatively short) time period. Assume a simple, two population model, where some number of people live in the countryside, c, and another number of people live in urban areas, u. Assume also that the population growth rate, α, is different in the two locations, with growth rate usually being higher in rural areas. In each time period, some fraction of people, f, in the rural area moves to the urban area. Then, the growth in population from time t in one time step is

$$c_{t+1} = \alpha_c c_t (1 - f)$$

$$u_{t+1} = \alpha_u (u_t + f c_t)$$

Consider an environmental impact, defined here as a flow in one time step, such as the amount of greenhouse gas pollution in one time step. A simple equation for calculating such an Impact is number of People times the Affluence (e.g., energy consumption per capita) times the Technology (e.g., greenhouse gas pollution per unit energy), the so-called $I = \mathrm{PAT}$ equation (cf. Dietz and Rosa, 1997). Assume that the A and T terms are different in the two locations, with A usually being higher in urban locations (city dwellers are richer and can consume more) but T often being lower (city dwellers have more efficient technology, and spatial concentration of population often increases efficiency more). Substituting our population equations above, and multiplying out terms:

$$I_{c,t+1} = \alpha_c c_t A_c T_c - f \alpha_c c_t A_c T_c$$

$$I_{u,t+1} = \alpha_u u_t A_u T_u + f \alpha_u c_t A_u T_u$$

In words, the environmental impact in the countryside is reduced by the amount of impact that those who emigrated would have otherwise caused, given the affluence and technology levels in the countryside. Similarly, the environmental impact in the urban area is increased by the amount of impact

that new immigrants cause, given the affluence and technology levels in the city.

Compare two scenarios, one where there is some migration (i.e., f is greater than zero) and one where there is none (i.e., f equals 0). The increase in environmental impact with migration is then:

$$fc_t(\alpha_u A_u T_u - \alpha_c A_c T_c)$$

The first term is just the number of people who migrate, and is generally nonnegative. The term in parentheses can be positive (migration is a net bad for the environment) or negative (migration is a net good for the environment). In words, if the product of A and T (e.g., greenhouse gases emitted per person) for urban dwellers times the urban birth rate is greater than the product of A and T for rural dwellers times the rural birth rate, the migration increases total impact on the environment.

Another way to look at is that migration increases total impact on the environment when:

$$\frac{\alpha_u}{\alpha_c} \times \frac{A_u}{A_c} \times \frac{T_u}{T_c} > 1$$

The first term is just the ratio of urban-to-rural rate of natural increase, and is generally less than 1. The second term is the ratio of affluence of urban-to-rural areas, and is generally greater than 1. The third term varies greatly depending on the environmental impact under consideration, and the socioeconomic and technological context. For instance (McDonald, 2008), US city dwellers generally use less energy to heat their homes because of efficiency gains in multiunit buildings, and hence emit less greenhouse gas emissions from this heating (i.e., this term is less than 1).

Despite the theoretical interest in studying the counterfactual case where there is no rural-to-urban migration, there is no clear historical example of a country that has economically developed without urbanization. Are there societies that are wealthy and that are predominately rural? In general, no (**Figure 2**). Economic development appears strongly correlated with urbanization. The lack of any historical examples makes evaluation of a hypothetical counterfactual scenario difficult.

There have been considerable policy efforts to slow the pace of urbanization. More than 73% of countries have such policies, which the United Nations Population Fund suggests is the evidence of a growing antiurban bias among policymakers (UNFPA, 2007). Most of these policies have had little effect on rates of urban growth. Many countries, for example, have tried to limit urban growth by denying residency permits to migrants, but have usually just succeeded in creating a large pool of undocumented workers. In short, in the real world there is only little that governments can do to stop urbanization, short of overt violence. A large amount of urban growth is coming as poorer countries, particularly in Africa

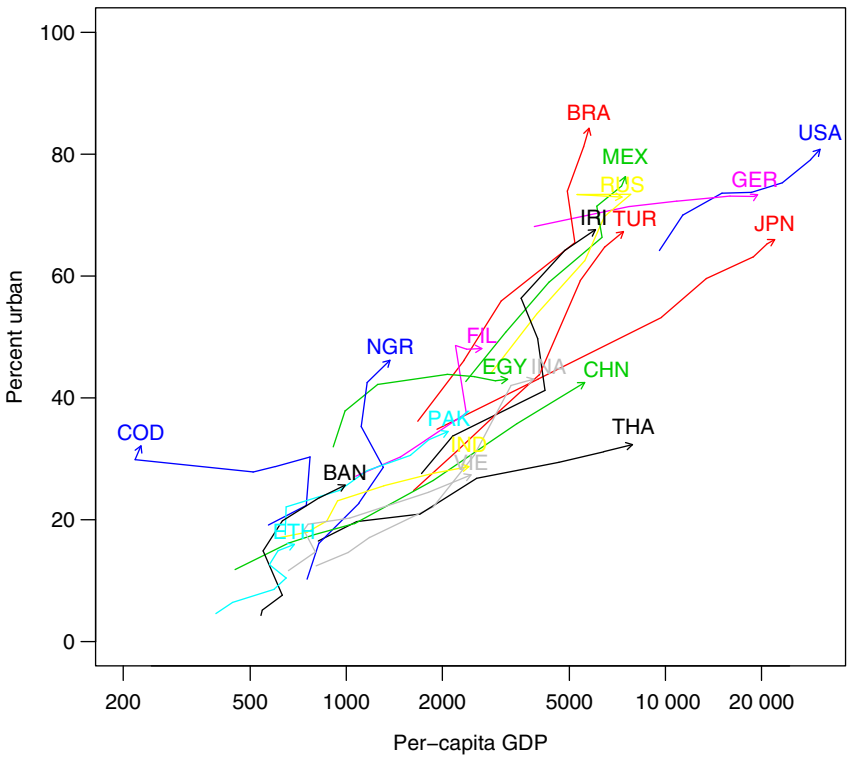


Figure 2 Correlation between per capita GDP and percent urban over time for the 20 countries with the greatest population in 2005. Each line represents one country from the period 1950 to 2005, with the arrowhead at the year 2005. Per capita GDP is taken from Maddison (2001), and is shown in 1990 International Geary-Khamis US dollars, a method correct for different purchasing powers over time and space. Percent urban is taken from the World Urbanization Prospects database (UNPD, 2009). Countries are abbreviated as: BAN (Bangladesh), BRA (Brazil), CHN (China), COD (Congo Kinshasa), EGY (Egypt), ETH (Ethiopia), FIL (Philippines), GER (Germany), INA (Indonesia), IND (India), IRI (Iran), JPN (Japan), MEX (Mexico), NGR (Nigeria), PAK (Pakistan), RUS (Russian Federation), THA (Thailand), TUR (Turkey), USA (United States), and VIE (Vietnam).

and Asia, urbanize. There is no realistic alternative to this urbanization (McDonald and Marcotullio, 2011).

Effects of Cities on the Environment

Cities affect the environment in a myriad of ways, and cataloging all of them in one article is impossible. In a sense, it is not surprising that the predominant place of human habitation should have such a great effect on the environment. Urbanization has profound effects on the economy, but it is rarely thought of as a "problem." What is different about urbanization's effect on biodiversity and ecosystem services?

One key part of the answer is that most facets of the environment are free or undervalued in the economy of most cities (MEA, 2005). The benefits that nature provides people, ecosystem services, are often nonexcludable goods (e.g., it is impossible practically to stop someone from breathing clean air) and nonrival (e.g., my breathing clean air does not reduce the supply of clean air significantly), and thus are undersupplied by the market like most public goods (Kolstad, 2000). Even where an economic value exists (e.g., water from a municipal water supply), political or market failures often leave natural resources undervalued. For all these reasons, the existence of important biodiversity or ecosystem services provision is not fully or adequately incorporated into urban decision-making. There is every reason, therefore, to worry about cities having undesirable impacts on the environment.

Another useful starting point for discussing the impact of cities and the environment is the observation that these impacts operate over a range of scales (McDonald *et al.*, 2009). At one extreme, the direct physical footprint of urban infrastructure fragments natural habitat. At the other extreme, urban forms affect the extent of greenhouse gas emissions, with global implications for climate change. In the subsections that follow, the impacts that span this range of spatial scales are discussed. First, the effect of the cities' infrastructure footprint on natural habitat is examined. Second, the impacts on ecosystem service provision are discussed. Third, impacts of cities on protected areas are examined. Finally, some of the positive effects of having people in close proximity with protected areas are discussed.

Natural Habitat and Biodiversity

Urban areas are currently around 3–5% of the Earth's land surface (Schneider *et al.*, 2009), although the actual figure varies significantly depending on the definition of urban and the spatial grain of analysis (Seto *et al.*, 2010). With urban growth, this will likely double (McDonald, 2008), although the future population of cities is difficult to predict (Seto *et al.*, 2010). However, urban areas are preferentially near coasts and on islands, areas with high species richness, which increases the impact of urban areas on biodiversity (**Figure 3**). Around 10% of terrestrial vertebrates are in ecoregions that are heavily impacted by urbanization (McDonald *et al.*, 2008). Urban-growth impacts are particularly significant in Mediterranean habitat types, where there is a large concentration of cities as well as many restricted-range endemic species.

The amount of land taken up by a city is a function of the population as well as the amount of built-up area per capita (Angel *et al.*, 2005). Poor countries have very compact cities, with $<100\ m^2$ per person, simply because few people have access to automobiles and must live within walking distance of jobs. Rich countries tend to have more dispersed cities (generally $>100\ m^2$ per person), especially those with a great deal of land area available for development (e.g., US cities often exceed $400\ m^2$ per person). Also important in controlling urban density are transportation policy (greater investment in roads tends to enhance dispersion, while greater investment tends to minimize dispersion) and zoning, which can either encourage or discourage dense settlement among neighborhoods (Jackson, 1985). In US cities, small proportions of neighborhoods (and people) cover the majority of the urban area and arguably have a large ecological impact, a phenomena sometimes called "sprawl inequality" (McDonald *et al.*, 2010).

Apart from a loss in natural area, expansion of cities also fragments natural habitat. One key effect of this fragmentation is the increase in isolation of natural habitat patches, as the average distance between them increases. Increased isolation tends to reduce population and gene flow among patches, and in the limit, may break a large regional population into several discrete subpopulations. Migration, both seasonal and intergenerational, is also restricted. In general, most mobile taxa like birds are less affected by isolation than less mobile taxa like amphibians, although some apparently mobile species have significant aversion to migrate across urban land cover (Saunders *et al.*, 1991).

Another key effect of fragmentation is to increase the amount of habitat that is near a habitat/nonhabitat edge (Murcia, 1995). This causes a systematic alteration of conditions near the edge, affecting the species and processes found there (Fagan *et al.*, 1999). At forest/nonforest edges, for example, temperature is significantly increased at the edge during the growing season due to greater solar insolation. This can increase average temperatures for tens of meters into the forest interior, changing the climate more than movement of hundreds of kilometers distance in latitude (Smithwick *et al.*, 2003). Roads create a particular type of edge, with particular ecological effects (Forman, 2000). Road noise is another commonly studied edge effect, and has been shown to significantly alter when and how bird species sing (Rheindt, 2003). Finally, biotic interactions may change near edges. For instance, many birds' nests are more parasitized by cowbirds when the nest is near an edge (Lloyd *et al.*, 2005).

Edge effects and the process of urbanization in general, increase the number and extent of nonnative invasive species (e.g., McDonald and Urban, 2006). Indeed, many authors have commented that there is a suite of "cosmopolitan" species that are present in most cities around the world (Kuhn and Klotz, 2006; McKinney, 2006). Meanwhile, the combined effects of habitat loss, increased isolation, and increased edge effects, as well as competition from invasive species, often leads to the loss of "sensitive" species dependent on larger, more natural blocks of habitat. The net result is often termed "biotic homogenization." Species richness in cities may actually be higher than that in rural areas, depending on the richness of the suite of cosmopolitan species relative to that in

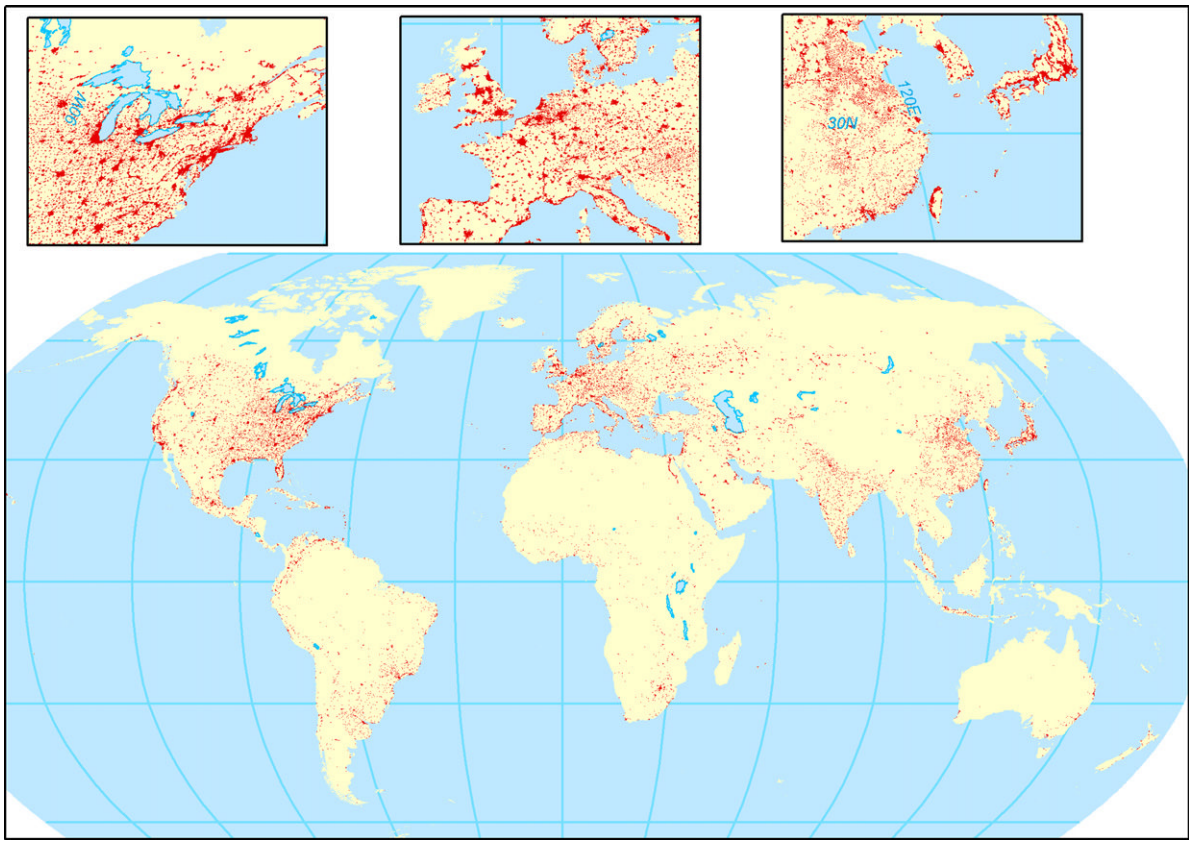


Figure 3 Distribution of urban areas, taken from the Global Rural/Urban Population (GRUMP) of the World database (CIESIN *et al.*, 2004). Note that the GRUMP database defines urban area based on nighttime light imagery, and includes suburban or exurban areas within the urban extent.

natural habitat, but global species richness declines. The flora and fauna of the world's cities thus becomes similar and homogeneous (Pysek *et al.*, 2004; Hobbs *et al.*, 2006; Grimm *et al.*, 2008).

Ecosystem Service Provision

Just as cities impact biodiversity, they also impact a myriad of ecosystem processes in sometimes complex ways (McDonald *et al.*, 2009). For example, fire ignitions tend to increase near urban land uses, increasing the risk of wildfires starting (Matos *et al.*, 2002). However, widespread suppression of fires soon after they start, designed to minimize the risk of damage to property, often decreases the magnitude of fires that do occur. The exact effect on the natural ecosystem varies, but in almost every case proximity to urban areas alters the fire regime. Occasionally, anthropogenic alteration of the fire regime may have unintended consequences, as when fire suppression leads to fuel buildup, increasing the risk of catastrophic fire.

Altered ecosystem processes of course lead to altered ecosystem service provision as well (McDonald and Marcotullio, 2011). An example might be the capacity of the atmosphere and biosphere to absorb atmospheric pollutants or mitigate their effect. Removal of forest cover as cities grow reduces standing carbon stock and decreases carbon sequestration. Without natural land cover, the natural ability of vegetation to moderate temperature swings is lost and the urban heat island effect (the tendency of cities to be warmer than their surroundings) is magnified (cf. Kalnay and Cai, 2003). In some cases, the loss vegetation also decreases the environment's ability to absorb or filter low-level air pollutants (Escobedo and Nowak, 2009). However, for other low-level air pollutants like VOC, natural vegetation can actually enhance formation of the pollutant, with negative effects to human health (Leung *et al.*).

Another example of an ecosystem service affected by urbanization would be the water quality (Carpenter *et al.*, 1998). Once urbanization passes a certain threshold in a watershed, water quality is significantly degraded. Urbanization, and particularly new construction activities that disturb the soil, increases sedimentation into streams. Excess nutrients, often from sewage or lawn fertilizer, end up running into streams as well. Finally, in hot climates water running into streams after flowing across hot pavement may thermally shock freshwater ecosystems. All of these factors negatively impact freshwater provision for people downstream (McDonald, 2009).

Cities and Protected Areas

Cities of course affect the biota in protected areas (**Figure 4**), just as they would affect that of any natural habitat, but city/park relationships also pose a few unique challenges to protected area management (McDonald *et al.*, 2009). At a small scale, many urban parks have vegetation damaged by

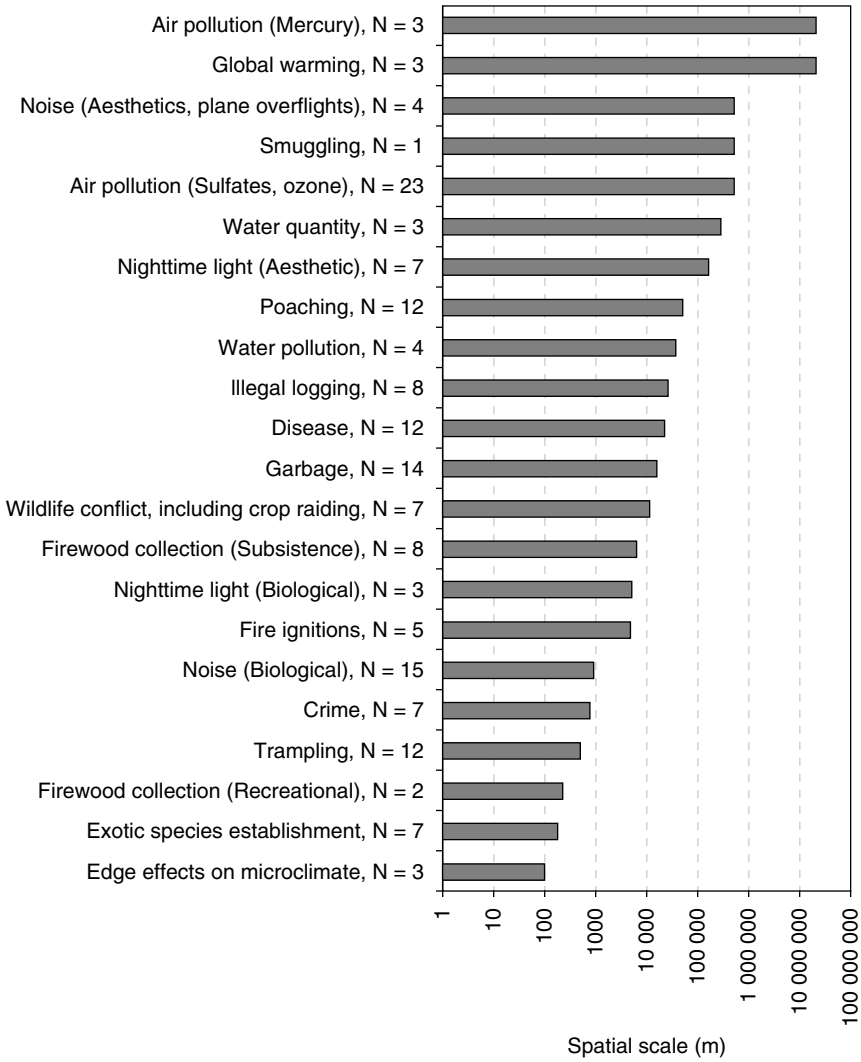


Figure 4 Maximum spatial scale at which negative and positive effects from urban areas propagate out and have been observed to affect protected areas. Each line represents an average of the reported distances in the literature. Note that the spatial scale is graphed using a logarithmic axis, and varies from local to global. The number of studies in each category is given (total $N=189$). Adapted from McDonald RI, Forman RTT, Kareiva P, Neugarten R, Salzer D, and Fisher J (2009) Urban effects, distance, and protected areas in an urbanizing world. *Landscape and Urban Planning* 93: 63–75.

trampling or are used for illegal activities (Liddle, 1991). At a medium scale, in many developing countries, illegal logging (Yonariza and Webb, 2007) and hunting within protected areas are more problematic near human settlements. At a broader scale (Cinzano *et al.*, 2001), light from human settlements can affect many species' movement, and can detract from people's wilderness experience in a park (Longcore and Rich, 2004). Finally, urban emissions of greenhouse gases will affect essentially all protected areas on Earth through enhanced global warming.

Perversely from the perspective of a conservationist, protected areas also tend to attract development to their edges (McDonald *et al.*, 2007). In developed economies, there is a large literature that uses hedonic analysis to quantify the amenity value of being near a protected area. Increased habitat conversion for new residential development is also observed near the border of protected areas (e.g., Mansfield *et al.*, 2005). In developing economies, there are some indications that a similar phenomenon is at work, although the paucity of good global data make broad-scale analyses difficult (Joppa *et al.*, 2009). Certainly there are many cases of tourism spending attracting new people near a park to earn a livelihood (Kruger, 2005).

Positive Impacts

After all this litany of negative effects, it is worth remembering that there are positive things about having cities close to protected areas, or natural areas more generally (McDonald, 2009). Parks near cities are of more value for recreation, simply because people can visit them more easily, an issue of some importance given the epidemic of obesity in many developed economies (Earnhart, 2004; Giles-Corti *et al.*, 2005;

Cohen *et al.*, 2006). Natural areas also offer a place for urbanites to experience nature, relax from their hectic lives, and perhaps draw some spiritual solace from the experience. These noble goals were the idea behind the creation of many large urban parks (Rosenzweig and Blackmar, 1998). Moreover, the role that places like Central Park in Manhattan played in connecting urbanites with nature is important, given that the majority of political activism and fundraising for conservation comes from urbanites (Nash, 2001).

Another benefit of cities arguably is the avoidance of a worse environmental impact under the baseline scenario of no rural-to-urban migration. In a world without such migration, billions more people would be in rural areas, increasing impacts there. In some extreme cases, population migrations (either to cities or elsewhere) have led to agricultural land being abandoned and a return to a more natural state. For instance, New England was once >80% cleared for farming, but was largely abandoned when farming in the US shifted elsewhere. The region is now >90% reforested, and many large mammals like moose have returned to the landscape (Foster and Aber, 2004).

Dependence of cities on Biodiversity and Ecosystem Services

Now to turn from the negative to the positive, cities not only impact biodiversity and ecosystem service provision, but they also depend on them for their continued functioning and the well-being of their citizens. As opposed to the impact of cities on the environment, often a negative economic externality, cities' dependence on ecosystem services is often an example of a positive externality. For many ecosystem services that cities depend on, whose value is not fully captured in economic decisions, there are theoretical reasons to assume that those positive externalities are unlikely to be fully supplied by the free market (Kolstad, 2000).

Dependence on Biodiversity

Despite the relative paucity of biodiversity in urban habitats, or at least taxa that are of global rarity, cities are culturally often the center of demand for biodiversity conservation. Urban dwellers are historically advocates for conservation, often more than rural dwellers. Indeed, in the last 50 years, the processes of urbanization and economic development are correlated with the creation of protected areas, one of the principal manifestations of the conservation movement. To some extent, this correlation supports the idea of land protection as an economic amenity – people value the existence of biodiversity, and its option and bequest values, more after they have achieved some base level of economic development. But a recent analysis of country-level data over time suggests that the degree of education of a populace, or the degree of freedom in a political system, is as important as the level of economic development (McDonald and Boucher, 2010).

US cities make a good case study, clearly showing the links between cities and advocacy for conservation. Teddy Roosevelt, born in wealth in Manhattan, and others in Eastern cities, played a key role in the creation of what is now Yellowstone National Park (cf. Schullery and Whittlesey, 2003). Without citizens in San Francisco organizing into the Sierra Club, it is unlikely that Yosemite would have been protected (e.g., Cohen, 1988). Similarly, residents of New York City played a key role in the protection of the Adirondacks (Rosenzweig and Blackmar, 1998).

Dependence on Ecosystem Services

Just as cities are the center of demand for biodiversity conservation, so they are the center of demand for ecosystem services. Indeed, some of the desire for protected areas was not due to a biodiversity value, but rather a desire to protect the recreational amenity or scenic beauty of a site. Most people are in cities, so in one sense it is not surprising that aggregate cities are the center of demand for ecosystem services. However, there is a debate in the literature about whether the average urban dweller is more or less dependent on ecosystem services than the average rural dweller, and whether conservation projects that aim to maintain ecosystem services truly benefit the poor (Boerner *et al.*, 2007; Corbera *et al.*, 2007; Ferraro and Hanauer, 2011). On the one hand, most of the very poor globally are located in urban areas, and it is hard to imagine someone more dependent on ecosystem services than a subsistance farmer. On the other hand, cities are very dependent on a flow of ecosystem services to survive; most urban dwellers have little capacity to raise their own food, for example, and are dependent on someone else growing the food.

To be a true ecosystem service, a desirable ecosystem process has to occur near consumers of that service (McDonald, 2009). The degree to which proximity is essential – the transportability of an ecosystem service – varies from service to service. Furthermore, most cities have a sharp gradient in population density with a relatively dense city core and then a decline in density as one heads away from the core. These two factors lead to a gradient in ecosystem service demand gradient that varies in steepness depending on the ecosystem service. The extent to which some ecosystem process is important to urban dwellers increases with more people nearby and with greater transportability of an ecosystem service.

An example of an ecosystem service that must be supplied extremely close to people is the shade from trees. Centuries of urban planning have traditionally created tree-lined boulevards in many cities. Shade has taken on greater quantitative importance in terms of a city's energy consumption with the rise of electromechanical heating and cooling of homes. One tree's shade, in a hot climate, can reduce annual electricity consumption by 3.6–4.8 kWh day^{-1} if it properly shades a residence (Akbari *et al.*, 1997). Such benefits, of course, only occur if trees are near buildings, and there's the rub. In many cases, those who benefit from the shade, whether individual homeowners or electric utilities, do not control what is planted on the streetscape, the domain of municipal authorities.

A slightly less local ecosystem is open space for use in recreation. Many studies looking at people's use of parks show that for day-to-day recreational opportunities, parks need to be within about a kilometer of the home (Giles-Corti *et al.*, 2005). Less frequent trips on weekends may be made within a short trip of 30 min to an hour in duration. In many cities, there is an expectation that municipal authorities will supply

recreational opportunities as a public good. In the US, this has led many cities to pass municipal bonds worth millions of dollars to protect open space.

Urban residents, of course, need water for their daily activities, and essential services like disposal of human wastes is impossible without some water. Typically the urban water source is relatively nearby the city, in the same watershed, although where necessary, cities have moved water hundreds of kilometers (e.g., San Francisco Public Utilities Commission, 2005). In many cases, natural vegetation filters water on the way to a reservoir, increasing water quality, or promotes groundwater infiltration, increasing water quantity (cf. Dudgeon *et al.*, 2006). The value of these ecosystem services may be partially expressed through the protection or restoration of such vegetation by municipal water planners. In other cases, however, planners may be unaware of their cities' dependence on functioning ecosystem processes.

At an even broader regional or continental scale, cities are dependent on how vegetation mitigates the effect of air pollution and ensures urban residents access to clean air to breathe. The "ecosystem service" shed, or area of importance for ecosystem service provision, varies depending on microclimatology, particularly the prevailing winds and vertical patterns of atmospheric circulation. It also depends on the pollutant; for some, like particulate matter, natural vegetation can act as a filter, while for others like low-level ozone, vegetation can actually increase pollutant formation. In general, very few mechanisms exist for valuing this ecosystem service. Most often, air pollution is controlled by restrictions on where and how it is emitted.

At the end of the spectrum of spatial scale lies climate regulation. Since greenhouse gases are well mixed in the atmosphere, carbon sequestration or avoided emissions services provided by vegetation are of service to urban dwellers regardless of where that vegetation is located. Right now, the lack of a coherent international agreement to replace the Kyoto Protocol has depressed the economic incentives for carbon sequestration or avoided emissions. Nevertheless, some cities have set ambitious targets like New York City's goal to plant a million new trees.

Land protection and other conservation activities can restore or enhance ecosystem services for urban residents. The "what" defines the "where": because of different transportability, what ecosystem service is to be conserved affects where conservation actions can be contemplated. Strictly from the perspective of demand for ecosystem services, it is usually better to be near the urban core, where there are lots of people (**Figure 5**). But stewardship, maintaining ecosystem processes on a preserved parcel, may be more difficult with lots of people nearby. All of the negative impacts of urban areas on ecosystem service provision noted above make the life of land stewards harder (McDonald, 2009).

It is also very expensive to purchase land or otherwise do conservation near a city center (McDonald, 2009). Cities have a sharp urban rent gradient (**Figure 5**), with land in the center being many times more valuable than land in the countryside. Thus, based on the ecosystem service demand gradient and the urban rent gradient, optimum placed for conservation of a particular ecosystem service varies depending on its transportability. Generally, conservationists should go as far out from a city as possible while doing ecosystem service protection, but no farther transportability is allowed. For instance, supply shade trees next to people's houses, and plant trees for carbon sequestration on very cheap land far from the city center.

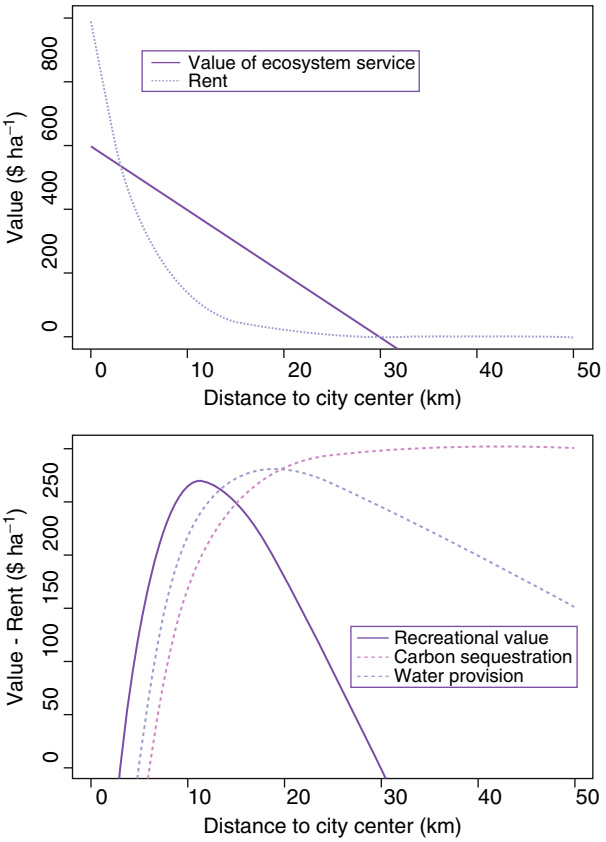


Figure 5 Framework for thinking about ecosystem service conservation and cities. Top panel: Hypothetical decline in the value of an ecosystem service that has some transport cost with distance from city center, compared with an urban rent gradient. Note that the maximum value for a conservation action to preserve the ecosystem service will occur when the tangent to the rent curve parallels the tangent to the ecosystem service curve. Bottom panel: The net value of a conservation action to preserve an ecosystem service as a function of distance from city center, for three hypothetical ecosystem services. Note that the curve for recreational value is derived from the curves in the top panel. Adapted from McDonald RI, Forman RTT, Kareiva P, Neugarten R, Salzer D, and Fisher J (2009) Urban effects, distance, and protected areas in an urbanizing world. *Landscape and Urban Planning* 93: 63–75.

Solutions

In this last section, the author discusses potential solutions to the issues discussed above. Some of these solutions aim to reduce the impact of cities on the environment, preventing or mitigating damage. Other solutions seek to enhance or at least maintain the flow of ecosystem services for urban residents. In the first subsection, we look at how urban form can make a city more or less sustainable. Next, we look at efforts to "mainstream" ecosystem services, putting them into everyday

decision-making. Finally, we discuss how urbanization can lead to lower population growth and rates of habitat conversion in the countryside.

Shaping More Sustainable Cities

Above, we discussed how the shape of a city controls its impact on natural habitat. Reductions in urban areas will generally reduce the amount of natural habitat lost during development, while more spatial contiguity of urban development will limit edge and isolation effects from habitat fragmentation (Seto *et al.*, 2010). The urban form of a city also mediates how people live and work, affecting the overall environmental footprint of a city. The urban form thus controls a city's environmental impact, but over the long term is also shaped by the environment, both in terms of the physical landscape (e.g., topography) and in terms of the city's dependence on ecosystem services. Other factors important in determining urban form include mode of transit, public policy, and zoning. Most importantly, urban form also has cultural dimensions, as esthetics and a people's sense of priorities can shape urban form profoundly.

One of the most commonly advocated ways to make cities more sustainable is to make them more compact. Compact city theory postulates that cities that are more dense and less fragmented have lower per capita use of many natural resources, less per capita pollution, and promote a psychologically and socially healthier environment for urbanites (Alberti, 2005; Leichenko and Solecki, 2005). For instance, average vehicle kilometers traveled (VKT) is well predicted by average population in cities in the developed world (Kenworthy *et al.*, 1992; Newman *et al.*, 1992; Kenworthy and Laube, 1999), wherein cities with densities lower than 30 people ha^{-1} have significantly more VKT per capita (**Figure 6**). More dense cities tend to use less energy for heating, cooling, and lighting than less dense cities, primarily because there are significant economies of scale by sharing space in one building rather than having multiple buildings (Steemers, 2003; Holden and Norland, 2005). Of course, the smaller amount of urban area implied by a compact city has obvious habitat conservation benefits, and indeed it is suburban or exurban

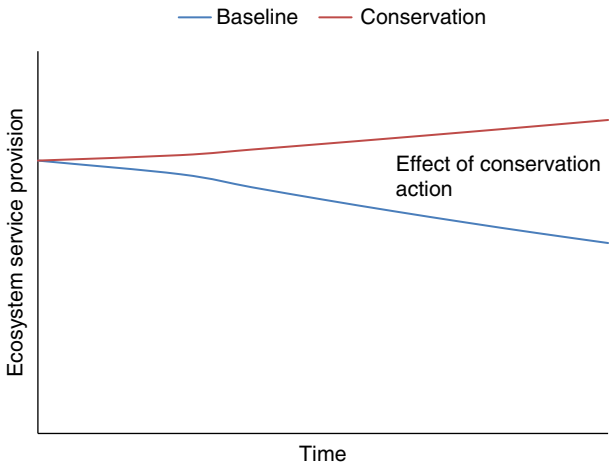


Figure 7 Conceptual graph of the additional provision of an ecosystem service after a conservation action, measured above a baseline (i.e., what would have occurred without the conservation action).

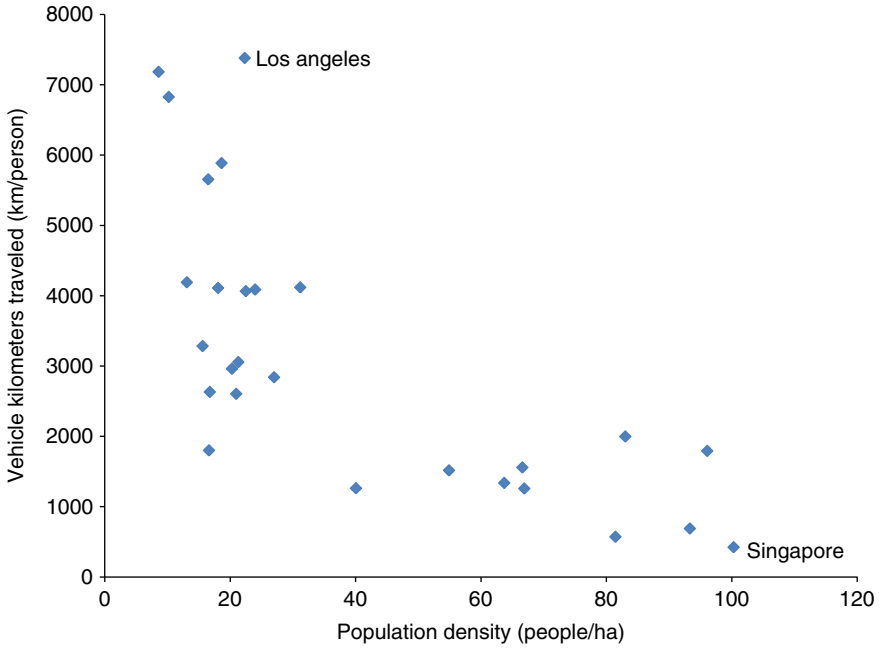
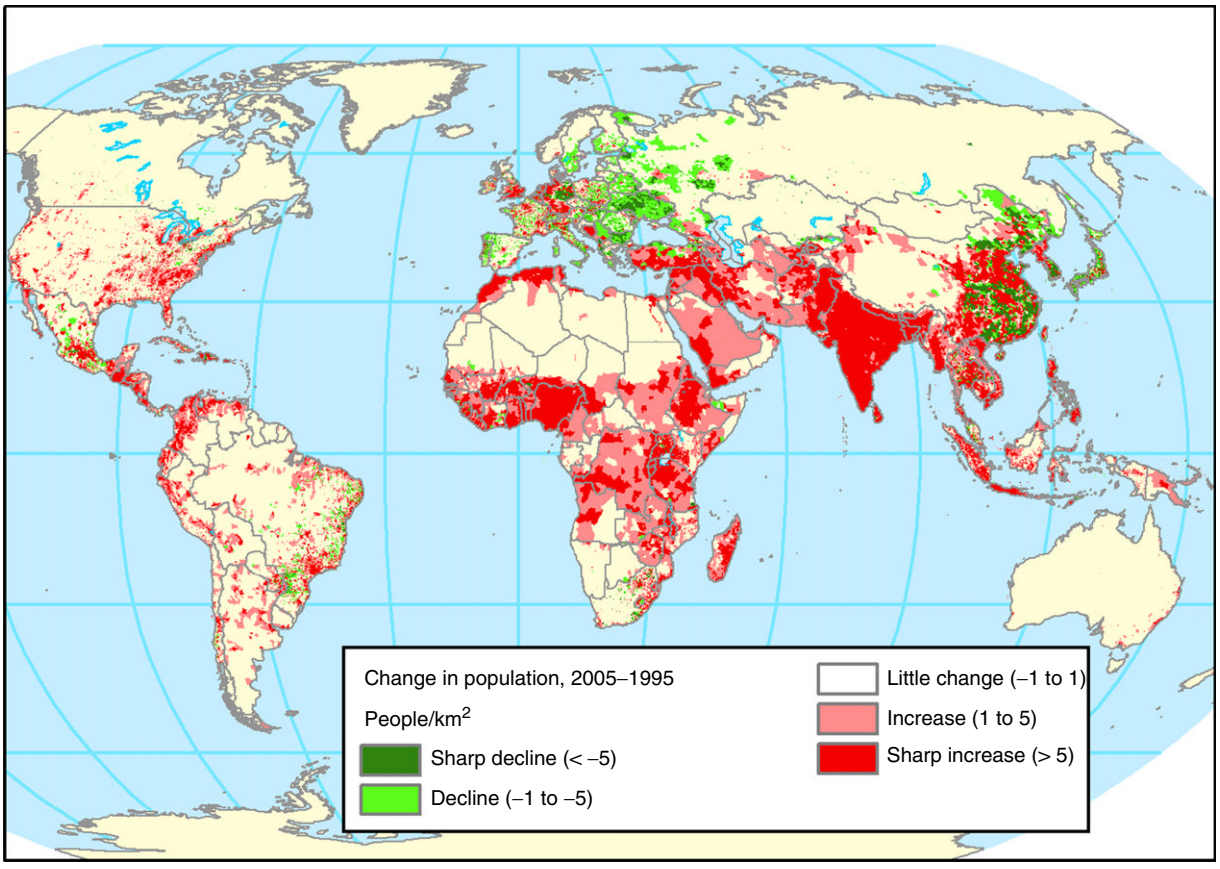
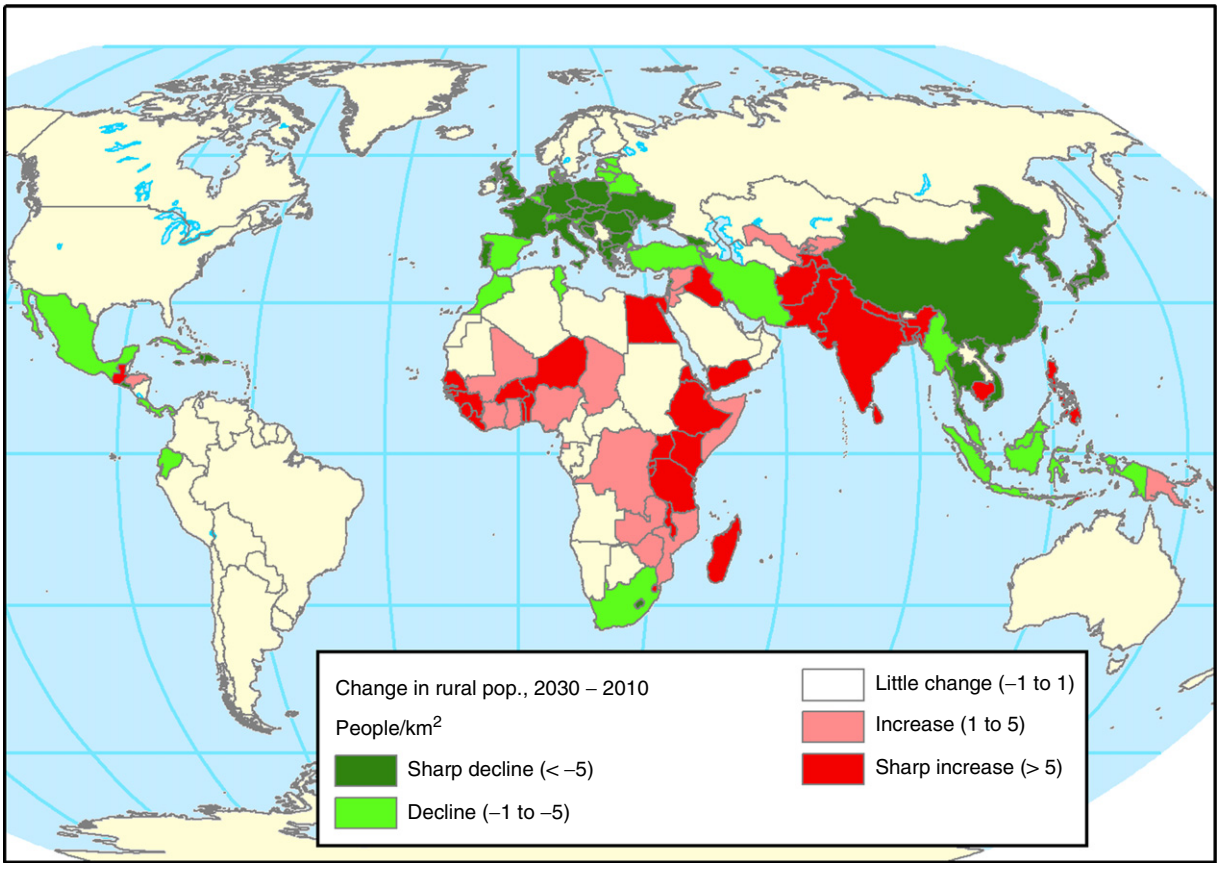


Figure 6 Correlation between population density and vehicle kilometers traveled (VKT) for a sample of cities in developed countries. For reference, the two most extreme cities in this sample are labeled on the graph. Data taken from Kenworthy J and Laube F (2000) *An International Sourcebook of Automobile Dependence in Cities, 1960–1990*. Boulder: University of Colorado Press.


Change in population, 2005–1995
People/km2
Sharp decline (< –5)
Decline (–1 to –5)
Little change (–1 to 1)
Increase (1 to 5)
Sharp increase (> 5)


Change in rural pop., 2030 – 2010
People/km2
Sharp decline (< –5)
Decline (–1 to –5)
Little change (–1 to 1)
Increase (1 to 5)
Sharp increase (> 5)

sprawl (settlements at the edge of cities at very low density) that is the biggest habitat conversion threat in many cities (Theobald, 2001). And it is often argued, particularly by those of the "New Urbanist" school of planning, that more compact cities are better for creating walkable neighborhoods and a sense of social cohesion (e.g., Duany *et al.*, 2000).

In all modern cities, zoning plays a significant role in shaping urban form (Jackson, 1985). Originally, zoning was about segregating industrial land-uses from the residential ones for public health reasons. However, zoning has become a complex set of rules that perform a huge variety of functions, from preventing development at high densities (so-called exclusionary zoning) to controlling signs that landowners can post. There has been a movement away from strictly separating land uses toward multiuse zoning that is more compatible with the compact city theory discussed above. Still, the overall effect of zoning in the US is to guarantee that new developments have relatively low population density, as neighborhoods seek to restrain developers and maintain their current character (Pendall *et al.*, 2006). Current residents benefit from the maintenance of the status quo, while future potential residents are kept out. Zoning has thus arguably led to greater urban sprawl, at least in the US, and hence greater environmental impact.

Cities (or neighborhoods in cities) that developed after the widespread availability of the automobile have a very different urban form than those that were developed before (Jackson, 1985). Automobile-dependent areas are typically at a lower density, and have less of a central business district and more dispersed business parks. However, public policy toward automobiles and mass transit varies from one jurisdiction to another, with discernable effects on urban form. Note that all countries subsidize both road construction and mass transit infrastructure, often from general tax revenue but sometimes from gas taxes or user fees, but some countries spend a greater proportion of subsidies on automobiles (e.g., US), while other spend a greater proportion on mass transit (e.g., Japan). Generally, investment in mass transit leads to a more compact city, with all of its environmental benefits, while investment in roads leads to a more dispersed city.

Mainstreaming Ecosystem Services

As discussed above (*see* Effects of Cities on the Environment), most urban decision-makers face a profoundly dysfunctional incentive structure, where their dependence on ecosystem services is ignored and their impact on ecosystem services for others little considered. One hope of environmentalists is that by incorporating ecosystem services into urban decision-making, better environmental outcomes will be reached. This could be through a direct payment for the ecosystem services from beneficiary to provider, where according to economic theory the market transaction will increase overall welfare for the two participants (Kolstad, 2000). More commonly, public policy incentives are used to create actions that maintain or enhance ecosystem services. For instance, city ordinances that encourage using drought tolerant plants in gardens are, in a sense, protecting freshwater resources by minimizing human withdrawals for landscape irrigation.

The spatial scale at which a market or incentives must operate depends on the transportability of the ecosystem service (McDonald, 2009). It also depends on administrative or political boundaries. While these boundaries are usually drawn without regard to ecological boundaries such as watersheds, they nevertheless often define the administrative area that might pay for a service or help provide. Finally, any ecosystem service market or incentive must exist in a specific cultural context, which controls what forms the market or incentive can practically take. For instance, in much of the Western world the concept of private property is widely accepted, and so it is a small intellectual step to think about private water-use rights or grazing rights. However, in many traditional pastoral systems, such as in Mongolia, most land is considered the common property of all, making notions of an ecosystem service market in the strict sense more problematic.

As an example of how an ecosystem service market or incentive might work in a city, consider shade trees and the important climate regulation they provide, significantly reducing heating and cooling costs. As discussed above (*see* Effects of Cities on the Environment), there is a disconnect between the beneficiaries of this service (homeowners and especially electric utilities) and those who plant and maintain landscaping (often building managers or municipal governments). Picture an electric utility offering to pay a set fee to municipal government for each tree it plants that provides useful shade, perhaps with restrictions on the orientation of the site, proximity to a building, and type of tree species involved so as to ensure the needed shade is truly provided. The utility is expressing how much they value the service by the price they offer to pay, and the municipality is expressing how much it costs to plant and maintain trees by the number of trees they plant at that price.

One can imagine a set of problems that might (and do!) occur when trying to set up an ecosystem service market or incentive like this (Brown *et al.*, 2007). There are very real transaction costs in negotiating and monitoring a deal between ecosystem service beneficiaries and providers. This is particularly true when there are multiple individual actors in either group, which poses a coordination and group action problem. Another potential hurdle ecosystem services schemes must jump is how to accurately, yet tractably, measure the quantity of an ecosystem service provided by an action. For instance, how much does any particular shade tree reduce cooling bills, and how can one cheaply yet relatively and accurately estimate that quantity? Specifically, one must measure the quantity of an ecosystem service provided relative to some baseline (**Figure 7**), the amount of ecosystem service provision

Figure 8 Geography of population decline and population increase. Top panel: Increase in population from 1995 to 2005, from the Gridded Population of the World database (CIESIN & CIAT, 2005). This dataset incorporates information from population censuses at the subnational scale, and includes information on both rural and urban populations, but because of the spatial aggregation of urban populations the observed pattern is largely due to changes in rural population. Bottom panel: National-level forecasts for the increase in rural population from 2010 to 2030, from the World Urbanization Prospects database (UNPD, 2009).

that would have occurred without any additional action. Related to this, one must also ensure that any increase in ecosystem service provision in one place does not cause actions at another site that reduces ecosystem service provision, a phenomenon termed "leakage" (e.g., Kindermann *et al.*, 2008).

A more regional example of an ecosystem service scheme might be water funds (Tallis *et al.*, 2009). In a water fund, users of water (often a water utility) set aside money that funds conservation actions upstream of the reservoir that increases water quantity or quality. For example, in Quito, the main water utility has set aside money for land protection and the construction of riparian buffer strips upstream in the paramo. These water funds have generally been easiest to set up when there is a single beneficiary and only a few landowners or land management agencies control a large area upstream. Note that the conservation strategies required to improve water quality, generally reduction in erosion of excess N and P run-off, are quite different than those required to increase water quantity, generally making agricultural irrigation more efficient or changing land cover to one that has lower evapotranspiration.

Finally, climate change furnishes a good global example of an ecosystem service market. Any project that reduces emissions or increases carbon sequestration results in a global benefit, but the cost of performing the project is borne at the site. The idea behind things like the Clean Development Mechanism and the Voluntary Carbon Standard is to pay local providers for providing a global benefit. However, a qualifying "project" could be interpreted much more broadly for cities than is currently done in the market for carbon offsets, leading in the limit to what you could call payment for sustainable urbanization (McDonald, 2008). For instance, a project to create a new subway system in a poor, rapidly growing city arguably might have global carbon value if, after accounting for baseline and leakage, a city is polluting less because fewer people drive.

New Opportunities in the Countryside

Rural-to-urban migration may reduce population growth in some rural areas significantly, and in some regions may lead to an actual decline in population (**Figure 8**). This potentially provides new opportunities for conservation. The reduced development pressure in rural areas means remaining natural areas can be more easily protected, and if agricultural land is abandoned, restoration may be a possibility. Areas with an absolute decline in population over the period 1995–2005 include much of Eastern Europe and the former Soviet Union, as well as portions of Mexico and China. Forecast patterns for the next 20 years look similar, with other countries in Southeast Asia beginning to lose rural population as well. However, in practice, most developing countries with substantial rural-to-urban migration still have large population growth rate in rural areas, limiting the importance of this "rewilding" opportunity (Boerner *et al.*, 2007; Corbera *et al.*, 2007; Ferraro and Hanauer, 2011). This is particularly true for Sub-Saharan Africa and South-Central Asia, which see quite striking increases in rural population over the same time period.

Appendix

Potential Courses for this Article

Landscape Ecology
Landscape Architecture
Urban Planning
Conservation Planning

See also: Energy Use, Human. Habitat Loss and Fragmentation. Human Impact on Biodiversity, Overview. Land-Use Issues. Urban–Suburban Biodiversity. Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy

References

Akbari H, Kurn DM, Bretz SE, and Hanford JW (1997) Peak power and cooling energy savings of shade trees. *Energy and Buildings* 25: 139–148.

Alberti M (2005) The effects of urban patterns on ecosystem function. *International Regional Science Review* 28: 168–192.

Anas A, Arnott R, and Small KA (1998) Urban spatial structure. *Journal of Economic Literature* 36: 1426–1464.

Angel S, Sheppard S, Civco D, *et al.* (2005) The dynamics of global urban expansion. *Transport and Urban Development Department*. Washington, DC: The World Bank.

Bairoch P (1988) *Cities and Economic Development: From the Dawn of History to the Present*. Chicago: University of Chicago Press.

Boerner J, Mendoza A, and Vosti SA (2007) Ecosystem services, agriculture, and rural poverty in the Eastern Brazilian Amazon: Interrelationships and policy prescriptions. *Ecological Economics* 64: 356–373.

Brown TC, Bergstrom JC, and Loomis JB (2007) Defining, valuing, and providing ecosystem goods and services. *Natural Resources Journal* 47: 329–376.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, and Smith VH (1998) Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications* 8: 559–568.

CIESIN & CIAT (2005) *Gridded Population of the World Version 3 (GPWV3): Population Density Grids*. Socioeconomic Data and Applications Center (SEDAC), Columbia University, Palisades, NY. Available at http://sedac.ciesin.columbia.edu/gpw.

CIESIN, Columbia University, IFPRI, World Bank and CIAT (2004) *Global Rural–Urban Mapping Project (GRUMP): Urban Extents*. Center for International Earth Science Information Network, Palisades, NY. Available at: http://sedac.ciesin.columbia.edu/gpw/.

Cinzano P, Falchi F, and Elvidge CD (2001) The first World Atlas of the artificial night sky brightness. *Monthly Notices of the Royal Astronomical Society* 328: 689–707.

Cohen DA, Ashwood JS, Scott MM, *et al.* (2006) Public parks and physical activity among adolescent girls. *Pediatrics* 118: E1381–E1389.

Cohen M (1988) The History of the Sierra Club, *1892–1970*. San Francisco: Sierra Club Books.

Corbera E, Kosoy N, and Tuna MM (2007) Equity implications of marketing ecosystem services in protected areas and rural communities: Case studies from Meso-America. *Global Environmental Change – Human and Policy Dimensions* 17: 365–380.

Dietz T and Rosa EA (1997) Effects of population and affluence on CO_2 emissions. *Proceedings of the National Academy of Sciences of the United States of America* 94: 175–179.

Duany A, Plater-Zyberk E, and Speck J (2000) *Suburban Nation: The Rise of Sprawl and the Decline of the American Dream*. New York: North Point Press.

Dudgeon D, Arthington AH, Gessner MO, *et al.* (2006) Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews* 81: 163–182.

Earnhart D (2004) Time is money: Improved valuation of time and transportation costs. *Environmental & Resource Economics* 29: 159–190.

Escobedo FJ and Nowak DJ (2009) Spatial heterogeneity and air pollution removal by an urban forest. *Landscape and Urban Planning* 90: 102–110.
Fagan WE, Cantrell RS, and Cosner C (1999) How habitat edges change species interactions. *American Naturalist* 153: 165–182.
Ferraro PJ and Hanauer MM (2011) Protecting ecosystems and alleviating poverty with parks and reserves: 'Win–Win' or tradeoffs? *Environmental & Resource Economics* 48: 269–286.
Florida R (2002) The economic geography of talent. *Annals of the Association of American Geographers* 92: 743–755.
Forman R (2000) Estimate of the area affected ecologically by the road system in the United States. *Conservation Biology* 14: 31–35.
Foster D and Aber J (2004) *Forests in Time: The Environmental Consequences of 1000 years of Change in New England*. New Haven: Yale University Press.
Giles-Corti B, Broomhall MH, Knuiman M, *et al.* (2005) Increasing walking – How important is distance to, attractiveness, and size of public open space? *American Journal of Preventive Medicine* 28: 169–176.
Grimm NB, Faeth SH, Golubiewski NE, *et al.* (2008) Global change and the ecology of cities. *Science* 319: 756–760.
Hobbs RJ, Arico S, Aronson J, *et al.* (2006) Novel ecosystems: Theoretical and management aspects of the new ecological world order. *Global Ecology and Biogeography* 15: 1–7.
Holden E and Norland IT (2005) Three challenges for the compact city as a sustainable urban form: Household consumption of energy and transport in eight residential areas in the greater Oslo region. *Urban Studies* 42: 2145–2166.
Jackson K (1985) *Crabgrass Frontier*. New York: Oxford University Press.
Joppa LN, Loarie SR, and Pimm SL (2009) On population growth near protected areas. *Plos One* 4.
Kalnay E and Cai M (2003) Impact of urbanization and land-use change on climate. *Nature* 423: 528–531.
Kenworthy J and Laube F (1999) Patterns of automobile dependence in cities: An international overview of key physical and economic dimensions with some implications for urban policy. *Transportation Research Part A – Policy and Practice* 33: 691–723.
Kenworthy J and Laube F (2000) *An International Sourcebook of Automobile Dependence in Cities, 1960–1990*. Boulder: University of Colorado Press.
Kenworthy JR, Newman PWG, and Lyons TJ (1992) The ecology of urban driving. 1. Methodology. *Transportation Research Part A – Policy and Practice* 26: 263–272.
Kindermann G, Obersteiner M, Sohngen B, *et al.* (2008) Global cost estimates of reducing carbon emissions through avoided deforestation. *Proceedings of the National Academy of Sciences of the United States of America* 105: 10302–10307.
Kolstad CD (2000) *Environmental Economics*. New York: Oxford University Press.
Kruger O (2005) The role of ecotourism in conservation: Panacea or Pandora's box? *Biodiversity and Conservation* 14: 579–600.
Krugman P (1998) What's new about the new economic geography? *Oxford Review of Economics and Policy* 14: 7–17.
Kuhn I and Klotz S (2006) Urbanization and homogenization – comparing the floras of urban and rural areas in Germany. *Biological Conservation* 127: 292–300.
Leichenko RM and Solecki WD (2005) Exporting the American dream: The globalization of suburban consumption landscapes. *Regional Studies* 39: 241–253.
Leung D, Tsui JKY, Chen F, Yip WK, Vrijmoed LLP, and Liu CH (2011) Effects of urban vegetation on urban air quality. *Landscape Research* 36, 173–188.
Liddle MJ (1991) Recreation ecology – effects of trampling on plants and corals. *Trends in Ecology & Evolution* 6: 13–17.
Lloyd P, Martin TE, Redmond RL, Langner U, and Hart MM (2005) Linking demographic effects of habitat fragmentation across landscapes to continental source-sink dynamics. *Ecological Applications* 15: 1504–1514.
Longcore T and Rich C (2004) Ecological light pollution. *Frontiers in Ecology and the Environment* 2: 191–198.
Mace R (2008) Reproducing in cities. *Science* 319: 764–766.
Maddison A (2001) *The World Economy: A Millennial Perspective*. Paris: Organization for Economic Cooperation and Development.
Mansfield C, Pattanayak S, McDow W, McDonald RI, and Halpin PN (2005) Shades of green: Measuring the value of urban forests in the housing market. *Journal of Forest Economics* 11: 177–199.
Matos D, Santos C, and Chevalier D (2002) Fire and restoration of the largest urban forest of the world in Rio de Janeiro City, Brazil. *Urban Ecosystems* 6: 151–161.
McDonald RI (2008) Global urbanization: Can ecologists identify a sustainable way forward? *Frontiers in Ecology and the Environment* 6: 99–104.
McDonald RI (2009) Ecosystem service demand and supply along the urban-to-rural gradient. *Journal of Conservation Planning* 5: 1–14.
McDonald RI and Boucher T (2010) Global development and the future of the protected area strategy. *Biological Conservation* 144: 383–392.
McDonald RI, Forman R, and Kareiva P (2010) Open space loss and land inequality in United States' cities, 1990–2000. *PLoS One* 5: e9509.
McDonald RI, Forman RTT, Kareiva P, Neugarten R, Salzer D, and Fisher J (2009) Urban effects, distance, and protected areas in an urbanizing world. *Landscape and Urban Planning* 93: 63–75.
McDonald RI, Kareiva P, and Forman R (2008) The implications of urban growth for global protected areas and biodiversity conservation. *Biological Conservation* 141: 1695–1703.
McDonald RI and Marcotullio P (2011) Global effects of urbanization on ecosystem services. In: Niemelä J (ed.) *Handbook of Urban Ecology*. Oxford, UK: Oxford University Press.
McDonald RI and Urban DL (2006) Edge effects on species composition and exotic species abundance in the North Carolina Piedmont. *Biological Invasions* 8: 1049–1060.
McDonald RI, Yuan-Farrell C, Fievet C, *et al.* (2007) Estimating the effect of protected lands on the development and conservation of their surroundings. *Conservation Biology* 21: 1526–1536.
McKinney ML (2006) Urbanization as a major cause of biotic homogenization. *Biological Conservation* 127: 247–260.
MEA (2005) *Millennium Ecosystem Assessment Synthesis Report*. Washington, DC: Island Press.
Montgomery M (2008) The urban transformation of the developing world. *Science* 319: 761–764.
Montgomery M, Stren R, Cohen B, and Reed HE (2003) *Cities Transformed: Demographic Change and its Implications in the Developing World*. Washington, DC: National Academies Press.
Murcia C (1995) Edge effects in fragmented forests – implications for conservation. *Trends in Ecology & Evolution* 10: 58–62.
Nash R (2001) *Wilderness and the American Mind*, 4th edn. New Haven: Yale University Press.
Newman PWG, Kenworthy JR, and Lyons TJ (1992) The ecology of urban driving. 2. Driving cycles across a city – their validation and implications. *Transportation Research Part A – Policy and Practice* 26: 273–290.
Pendall R, Puentes R, and Martin J (2006) *From Traditional to Reformed: A Review of the Land Use Regulations in the Nation's 50 Largest Metropolitan Areas*. Washington, DC: The Brookings Institution.
Pysek P, Chocholouskova Z, Pysek A, Jarosik V, Chytry M, and Tichy L (2004) Trends in species diversity and composition of urban vegetation over three decades. *Journal of Vegetation Science* 15: 781–788.
Rheindt FE (2003) The impact of roads on birds: Does song frequency play a role in determining susceptibility to noise pollution? *Journal Fur Ornithologie* 144: 295–306.
Rosenzweig R and Blackmar E (1998) *The Park and the People: A History of Central Park*. Ithaca, NY: Cornell University Press.
San Francisco Public Utilities Commission (2005) *A History of the Municipal Water Department and Hetch Hetchy System*. San Francisco Public Utilities Commission, San Francisco.
Sassen S (2001) Cities in the global economy. In: Paddison R (ed.) *Handbook of Urban Studies*. London: Sage Publications.
Saunders DA, Hobbs RJ, and Margules CR (1991) Biological consequences of ecosystem fragmentation: A review. *Conservation Biology* 5: 18–32.
Schneider A, Friedl MA, and Potere D (2009) A new map of global urban extent from MODIS satellite data. *Environment Research Letters* 4: 0044003.
Schullery P and Whittlesey L (2003) *Myth and History in the Creation of Yellowstone National Park*. Lincoln: University of Nebraska Press.
Seto KC, Sanchez-Rodriguez R, and Fragkias M (2010) The new geography of contemporary urbanization and the environment. *Annual Review of Environment and Resources* 35: 167–194.
Smithwick E, Harmon M, and Domingo J (2003) Modeling multiscale effects of light limitations and edge-induced mortality on carbon stores in forest landscapes. *Landscape Ecology* 18: 701–721.
Steemers K (2003) Energy and the city: Density, buildings and transport. *Energy and Buildings* 35: 3–14.
Tallis H, Goldman R, Uhl M, and Brosi B (2009) Integrating conservation and development in the field: Implementing ecosystem service projects. *Frontiers in Ecology and the Environment* 7: 12–20.
Theobald DM (2001) Land-use dynamics beyond the American urban fringes. *Geographical Review* 91: 544–564.

UNFPA (2007) *State of the world population 2007: Unleashing the potential urban growth*. New York: United Nations Population Fund.

UNPD (2009) *World Urbanization Prospects: The 2009 Revision*. New York: United Nations Population Division.

Van Ommeren J and Rietveld P (2005) The commuting time paradox. *Journal of Urban Economics* 58: 437–454.

von Thünen JH (1875) *Der isolirte Staat (The Isolated State)*. Berlin: Wiegandt, Hempel, and Parey.

Yonariza and Webb EL (2007) Rural household participation in illegal timber felling in a protected area of West Sumatra, Indonesia. *Environmental Conservation* 34: 73–82.

Inbreeding and Outbreeding

Katherine Ralls, Smithsonian Conservation Biology Institute, Washington, DC, USA
Richard Frankham, Macquarie University, North Ryde, NSW, Australia
Jonathan D Ballou, Smithsonian Conservation Biology Institute, Washington, DC, USA

Published by Elsevier Inc.

Glossary

Allele One of two or more different forms of a gene.
Coadapted gene complexes Chromosomes, loci, or genes that are adapted to function well together.
Effective population size (N_e) Size of the ideal population used in population genetics theory that would have the same rate of increase in inbreeding or decrease in genetic diversity as the actual population under study.
Genetic drift Changes in the frequency of alleles within a population over time due to the random process of passing alleles from parents to offspring.
Lethal equivalent Group of mutant alleles that would cause an average of one death if homozygous; for example, one lethal equivalent might represent two mutant alleles, each with a 50% probability of causing death, or any other combination of mutant alleles that would produce an average of one death.
Outcrossing A mating system in which individuals breed with other individuals that are not related to them.
Selection differential(s) A measure of the intensity of selection on a phenotypic character.

Inbreeding

Inbreeding is the production of offspring from the mating of individuals related by ancestry. This includes self-fertilization, brother–sister, parent–offspring, and cousin matings, as well as matings between more distant relatives. Inbred offspring are more likely to inherit recent copies of the same allele from both parents, that is, alleles that are identical by descent (derived from a common ancestor of both the parents). Two alleles that are identical by descent are homozygous, but not all homozygous alleles are identical by descent. A homozygous individual has two alleles at a locus that are functionally similar. However, these two alleles may or may not be identical by descent.

The inbreeding coefficient (usually symbolized by F) of an individual is the probability that the individual has two alleles at a locus that are identical by descent. Because F is a probability, it ranges from 0 for noninbred individuals to 1 for completely inbred individuals. For example, the inbreeding coefficient of an individual resulting from self-fertilization is 1/2 and that for an individual resulting from a parent–offspring or brother–sister mating is 1/4.

Because the process of inbreeding increases the likelihood that two alleles in a genotype are identical by descent, inbreeding increases the frequency of homozygotes, and reduces the frequency of heterozygotes in direct proportion to the inbreeding coefficient. A population with an average inbreeding coefficient of 10% will have 10% less heterozygosity than a similar noninbred population.

Inbreeding Depression

Natural populations contain low frequencies of deleterious recessive mutations (due to the balance between their occurrence by mutation and removal by selection) that are normally found as heterozygotes and are thus not expressed. Inbreeding increases the probability of individuals inheriting these recessive alleles as homozygotes, thus enabling the expression of their deleterious effects. Consequently, in most populations of animals and plants, inbreeding results in a decline in reproduction and survival (reproductive fitness), which is called inbreeding depression.

Another cause of inbreeding depression is overdominance. Different genotypes have different effects on fitness. In some traits or loci, even though alleles are not recessive, the homozygous genotypes (e.g., AA and aa) are, on average, less fit than the heterozygous genotype (Aa). These traits or loci are referred to as overdominant or showing heterozygote advantage. As inbreeding decreases the proportion of the fitter heterozygotes in a population, the average fitness (e.g., reproduction and survival) of the population declines, resulting in inbreeding depression.

Studies show that most inbreeding depression is due to the accumulation of deleterious recessive mutations, rather than overdominance.

Evidence for Inbreeding Depression

The deleterious effects of inbreeding were known long before the discovery of the underlying Mendelian mechanisms that cause it. In the nineteenth century, Charles Darwin clearly documented inbreeding depression based on studies in 57 species of plants, as well as the experience of livestock breeders. These early observations were amply confirmed by subsequent studies. There is now extensive evidence for inbreeding depression in laboratory and domestic animals and plants, as well as considerable evidence for inbreeding depression in wild animals and plants. Many studies on inbreeding were massive, involving large numbers of animals over many years, and the literature is extensive. For example, Sewall Wright's classic experiments on inbreeding in guinea pigs resulted in the production of 29,310 inbred and 5105 control young from 1906 to 1924. Only the five most vigorous inbred lines survived to the

end of the experiments; 30 other inbred lines went extinct or declined so severely that Wright discontinued breeding them before the end of the experiment. Nevertheless, the surviving inbred guinea pigs were consistently inferior to the controls in number of young born, percentage of young born alive and raised to 33 days, and weight at 33 days.

Inbreeding and selection have been used to fix desirable traits in modern breeds of livestock. However, reduced fertility was a major problem during the early periods of inbreeding and the inbred lines were repeatedly outcrossed to restore vigor and fertility. The North Central regional dairy cattle-breeding project (involving Iowa, Michigan, Minnesota, Missouri, Ohio, South Dakota, and Wisconsin), begun in 1947, is an example of the many extensive studies on inbreeding in livestock. This study found that inbreeding usually increased juvenile mortality and decreased milk yield, fat yield, growth, and reproductive performance.

Inbreeding depression is also well known in zoo animals. For example, in the early 1980s, scientists at the National Zoo found that juvenile mortality was higher in inbred than in noninbred offspring in 41 of 44 populations of mammals in zoos, including many primates, antelopes, and deer, as well as a variety of smaller mammal species (**Figure 1**). On average, the progeny of father–daughter and brother–sister matings suffered a 33% reduction in juvenile survival compared with outbred offspring. However, the severity of inbreeding depression varied widely across species. Although inbreeding in zoo animals usually results in less vigorous and fertile individuals that appear phenotypically normal, it is also responsible for some genetic diseases in captive populations, including blindness in wolves and dwarfism in California condors. Zoo populations are now routinely managed to avoid inbreeding.

Evidence for inbreeding depression in natural environments has been reported in many species (Keller and Waller, 2002). A review of the literature on inbreeding effects in wild vertebrates and plants found that 90% of 157 data sets showed inbreeding depression, indicating that species frequently exhibit moderate to high levels of inbreeding depression under natural conditions (Crnokrak and Roff, 1999). Species showing inbreeding depression include mammals (golden lion tamarins, lions, native mice, shrews, and Soay sheep), birds (greater prairie chicken, Mexican jay, song sparrow, American kestrel, and reed warbler), fish (Atlantic salmon, desert topminnow, and rainbow trout), a reptile, a snail, an insect (butterfly), and many species of plants. Some studies have failed to find inbreeding depression, but these have typically been small studies or ones where paternity has not been verified. For example, no inbreeding depression was found in superb fairy wrens in Western Australia, but subsequent paternity studies using molecular genetic methods revealed that 64% of progeny were not fathered by the males to whom paternity had previously been attributed. Because inbreeding depression is so common, management of an unstudied outbreeding species should be based on the assumption that it will suffer reductions in reproductive fitness if it is inbred.

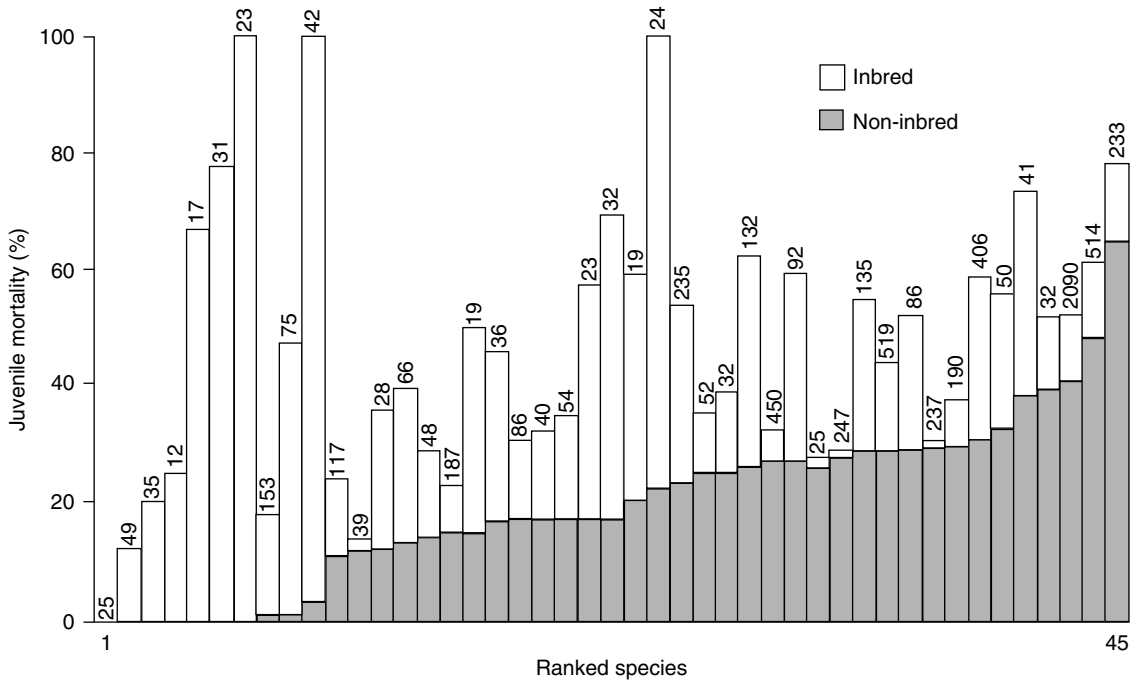


Figure 1 Juvenile mortality in inbred and noninbred young in 45 populations of mammals in captivity. Inbreeding levels were calculated with respect to the founders of the population for which the pedigree data were available. Most populations were founded with wild-caught animals but some were founded with animals from other zoos or of unknown origin. The noninbred category includes all animals with an inbreeding coefficient of 0 and the inbred category included all of those with an inbreeding coefficient greater than 0. For the larger species, all young surviving up to 6 months or more were considered to have survived. For the small species, one-half of the age at sexual maturity was used as the criterion age. Numbers above the bars indicate sample sizes for each species. Reproduced from Ralls K and Ballou J (1986) Captive breeding programs for populations with a small number of founders. *Trends in Ecology and Evolution* 1: 19–22.

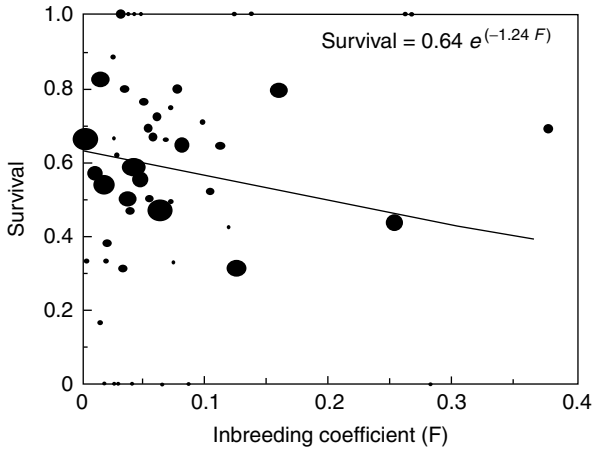


Figure 2 Decreasing juvenile survival with increasing inbreeding in golden lion tamarins (*Leontopithecus rosalia*). The proportion of offspring with a given inbreeding coefficient surviving to 7 days of age is plotted against the inbreeding coefficient. The sizes of the ovals represent the number of offspring available to estimate survival for each level of inbreeding. Large ovals represent samples of more than 50 offspring, medium ovals from 30 to 50, small ovals from 10 to 30, and tiny ovals from 1 to 10.

Factors Affecting the Severity of Inbreeding Depression

The degree of inbreeding depression in a population depends on the extent of inbreeding, the original frequency of deleterious recessives, the environment, and chance. Continued inbreeding results in greater inbreeding depression. For example, average levels of various components of reproductive fitness, such as juvenile survival in mammals (**Figure 2**) or grain yield in maize, theoretically decrease in a linear fashion as the inbreeding coefficient increases. The more deleterious alleles that were formerly masked by heterozygosity, the more severe the effects of inbreeding will be.

Severity of inbreeding depression also depends on the environmental conditions. Under benign conditions, many inbred plants and animals may survive. However, inbreeding depression is greater under more stressful conditions. Selfed progeny of one species of plant showed 75% inbreeding depression in the field but only 59% in garden plots and 53% in a greenhouse (Dudash and Fenster, 2000). A meta-analysis of 34 studies concluded that inbreeding depression was, on average, 69% more severe in stressful than in benign environments (Crnokrak and Roff, 1999).

Measuring Inbreeding Depression as Lethal Equivalents

The extent of inbreeding depression in survival of animals can be measured in terms of lethal equivalents. The number of lethal equivalents per gamete or individual can be calculated from the rate that survival decreases with increasing inbreeding. A study of 40 captive populations of mammals found an average of 4.6 lethal equivalents per individual for juvenile survival (Ralls *et al.*, 1988). Thus, each individual contained enough deleterious mutations to have an effect equivalent to slightly <5 single lethal mutations if they were homozygous. This figure is similar to estimates done in humans and birds. However, the number of lethal equivalents varied widely across captive populations. Estimates of lethal equivalents based on survival to progressively older ages tend to increase. For example, wild song sparrows had 2.9 lethal equivalents based on survival from egg to independence but 5.4 lethal equivalents based on survival from egg to age at first breeding (Keller, 1998).

Inbreeding Depression in Total Reproductive Fitness

Most studies of inbreeding depression measure only one or a few components of reproductive fitness. However, since all traits have the potential to accumulate deleterious recessive alleles as a result of mutations, it is expected that all components of reproductive fitness will be subject to inbreeding depression. In animals, this includes offspring survival, number of offspring per female, male mating ability, sperm quality, disease resistance, and the quality and quantity of maternal care. Greater inbreeding depression for overall fitness than for its components has been found for old field mice, house mice, chickens, turkeys, Japanese quail, chukar partridges, and song sparrows. A meta-analysis estimated that wild populations have an average of about 12 lethal equivalents across the lifecycle (O'Grady *et al.*, 2006). In song sparrows, eggs with an average inbreeding coefficient of 0.25 have a 79% loss of fitness distributed over various life history stages: egg survival to breeding age was reduced by 49%, adult survival by 24%, and lifetime reproductive success of females by 48%. In the Takahe, an endangered New Zealand rail, inbreeding led to only small reductions in fitness at individual life history stages, the deleterious effects were cumulative across multiple life history stages and significantly reduced long-term fitness. Successful recruitment of second-generation offspring was reduced by 88% in full sib pairs compared to unrelated pairs.

Susceptibility to Inbreeding Depression

Inbreeding depression has a large chance (stochastic) element because of the random sampling of alleles during reproduction. Individuals with the same inbreeding coefficient, that is, the same probability of carrying alleles identical by descent, will differ in actual levels of homozygosity for any given trait, and, thus, will vary in the degree to which they experience inbreeding depression. Families and populations within a species carry different types and numbers of deleterious mutations and differ in their susceptibility to inbreeding depression. Differences in the extent of inbreeding depression among lineages within species are common, including mice, dairy cattle, fruit flies, and flour beetles and are to be expected in all outbreeding species. These differences contribute to the occasional success in establishing inbred lines or wild populations from the progeny of a small number of founding individuals even in species that typically show high levels of inbreeding depression. Differences among populations within species have also been reported. No differences in susceptibility to inbreeding depression among major taxonomic groups are known, but the relevant data are limited.

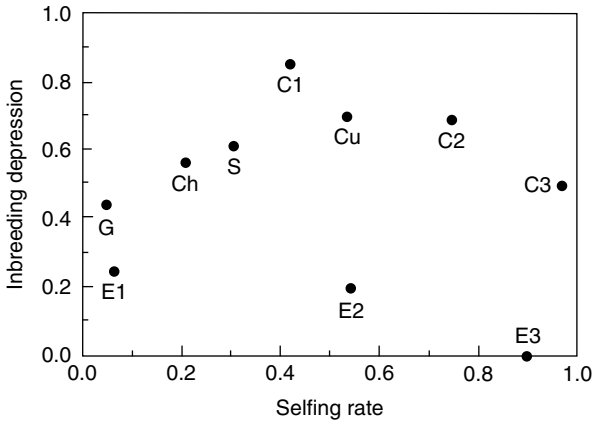


Figure 3 Relationship between inbreeding depression and the selfing rate in several species of herbaceous angiosperms. Species are G, *Gila achilleifolia*; Ch, *Chamaecrista fasciculata*; S, *Sabatia angularis*; Cu, *Cucurbita foetidissima*; C, *Clarkia tembloriensis*; and E, *Eichhornia paniculata*. For *Clarkia* and *Eichhornia*, data from three populations are plotted. Inbreeding depression for these species was measured in the lath house and glasshouse, respectively; all others were measured in the field. Reproduced with permission from original Figure 1.7 in Falk DA and Holsinger KE (1991) *Genetics and Conservation of Rare Plants*, 23 pp. New York: Oxford University Press, Inc.

Species that naturally inbreed (e.g., selfing plants and mole rats) occasionally show lower levels of inbreeding depression than naturally outbreeding species (**Figure 3**). Slow rates of inbreeding over long periods of time have the potential to select against and remove deleterious alleles (i.e., purge, see below) in these inbreeding populations, reducing their inbreeding depression. While these types of species often show less inbreeding depression, they still do exhibit detectable levels of inbreeding depression when inbred further.

Efforts to predict the susceptibility of specific populations or species to inbreeding depression have not been successful owing to the extremely stochastic nature of inbreeding depression.

Inbreeding and Outbreeding in the Wild

Only about 20% of species regularly self-fertilize and these tend to have life histories that strongly favor mating with relatives. For example, many are colonizing species in which the chances of successfully dispersing are much higher if only a single individual has to reach a new habitat. The restricted taxonomic distribution of selfing to some plants, terrestrial slugs, and marine invertebrates suggests that it is an evolutionary dead end.

The majority of species appear to be naturally outbreeding. Many plants that are pollinated by insects have elaborate morphological mechanisms that favor cross-pollination. In other plants, cross-pollination is ensured because the male and female gametes do not mature at the same time. Another mechanism that prevents self-fertilization is self-sterility, in which the pollen either fails to germinate on a stigma of its own flower or germinates, but does not develop sufficiently to fertilize the egg.

Many animals are also thought to avoid close inbreeding. Although a few mammals, notably the naked mole rat, normally mate with close relatives and are highly inbred, most mammals and birds rarely mate with close relatives. Sex differences in dispersal patterns often limit opportunities for inbreeding. In most mammals, males tend to disperse more frequently or farther than females, whereas the reverse is true in birds. Furthermore, many species can recognize close kin, and a variety of species such as jays, woodpeckers, mice, voles, ground squirrels, black-tailed prairie dogs, and chimpanzees are known to actively avoid mating with them.

Inbreeding Depression in Small Populations

Inbreeding is unavoidable in small, closed populations because all individuals eventually become related to each other. Inbreeding in a population of size N_e increases at a rate of $1/(2N_e)$, per generation with random mating. For example, in an effective population of size 10, there is a 5% increase in inbreeding per generation. Consequently, small, isolated populations that have existed for many generations are expected to show inbreeding depression. Small populations of plants, fruit flies, a rock wallaby, Florida panthers, and a snake have been found to suffer from inbreeding depression. However, inbreeding depression may not cause declines in population size. Reduced fecundity and survival will only cause a population to decline to extinction if the reproductive rate drops below the replacement level. Small (and large) populations can also suffer from reduced reproductive fitness because of declines in environmental quality.

Reduction of Inbreeding Depression

Inbreeding depression may be reduced by selection against deleterious alleles, which eliminates, or purges, them from the population. Purging has been documented in plants, mice, birds, and fruit flies, and during the development of inbred lines in a variety of species. Species that naturally inbreed should theoretically show less inbreeding depression than naturally outbreeding species because of the greater opportunity for selection against deleterious recessives.

Purging can reduce inbreeding depression, but it is unlikely to eliminate it. In captive mammals, selection led to a small reduction of inbreeding depression in neonatal survival in 15 of 17 populations. No trends in purging effects were observed in survival to weaning or litter size. The purging effects were not strong enough to be of practical use in captive breeding programs. Lacy and colleagues found that the effectiveness of natural selection in reducing inbreeding depression varied substantially in experiments with three subspecies of captive deer mice, probably due to different histories of inbreeding and selection in the wild prior to capture. Reviews on the effects of purging have pointed to a variety of effects from nil to moderate. In sum, it remains impossible to make strong predictions about the situations under which purging will occur and the probable magnitude of purging effects in individual populations. Therefore, genetic management programs for captive populations currently aim to minimize inbreeding

rather than undertake deliberate inbreeding in an attempt to purge deleterious alleles from populations.

Inbreeding depression can be completely reversed by outcrossing an inbred population to another unrelated population, either an outbred population or an independently inbred population. Crossing two highly inbred, but unrelated, populations results in an outbred population with a high level of heterozygosity and little inbreeding depression, since the high heterozygosity effectively masks the deleterious alleles that were being expressed in the inbred populations. This effect is termed "hybrid vigor." Partial recovery can be achieved by introducing unrelated individuals into an inbred population. Several wild populations showing signs of severe inbreeding depression have benefited dramatically from the introduction of unrelated individuals. Examples of such genetically stimulated recovery are known from a broad range of taxa including mammals (deer mice, gray wolves, Florida panthers), birds (prairie chickens), reptiles (adders), fish, and plants.

Inbreeding, Loss of Genetic Diversity, and Extinction

Many conservation biologists have been concerned that inbreeding and loss of genetic diversity in small populations will increase the risk of extinction. However, some researchers have questioned this view because of the difficulty of obtaining direct evidence that inbreeding contributes to the extinction of wild populations.

Inbreeding and Extinction in Captivity

There is overwhelming evidence that deliberately inbred populations of laboratory and domestic animals and plants suffer elevated extinction rates, as illustrated in **Figure 4**. Extinctions were rare at low levels of inbreeding but sharply increased when inbreeding reached intermediate levels as it generally takes several generations and considerable inbreeding until the growth rate of the population becomes negative.

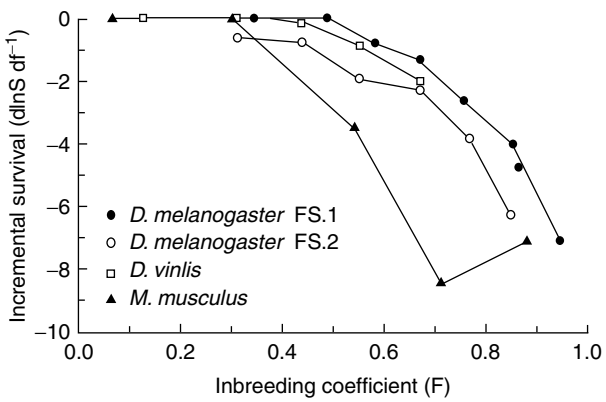


Figure 4 Relationships between survival of populations and inbreeding (F) for *Drosophila melanogaster*, *Drosophila virilis*, and house mice. All populations were inbred using full-sibling mating, except for the use of cousin mating in the first generation in *Mus musculus*. Reproduced from Frankham R (1995) Inbreeding extinction: A threshold effect. *Conservation Biology* 9: 792–799, with permission from Wiley.

Most attempts to develop inbred lines by intense inbreeding are unsuccessful. Of the 20 colonies of laboratory mice maintained by sibling mating in **Figure 4**, 19 became extinct by generation 12, but the remaining colony showed no decline in litter size and became a successful inbred line. Results with fruit flies, guinea pigs, poultry, and Japanese quail are similar.

Slower rates of inbreeding also lead to extinctions. Small populations of fruit flies maintained at effective sizes of 10 and 20 had greatly elevated extinction risks, even though these were slightly lower than those of populations subjected to intense inbreeding (full-sibling mating) when all were compared at the same inbreeding levels. Further, all the larger populations that survived to high levels of inbreeding had fitness levels lower than those of outbred controls, leading these authors to conclude that although slower rates of inbreeding may slightly slow extinction rates, they do not offer a solution to the increased extinction risk associated with inbreeding in small populations. The effects of inbreeding on population viability are complex and will interact with other factors affecting population growth, population fluctuations, or N_e, but they will be deleterious in the long term (**Figure 4**).

Inbreeding and Extinction Risk in Wild Populations

Conservation biologists initially thought that genetic problems played a role in the endangerment and extinction of wild populations. It was clear that inbreeding increased extinction risk in laboratory populations and theoretical work in the 1980s suggested that small populations in the wild should also suffer from increased extinction due to the unavoidable increase in mating between close relatives. However, in an influential paper, Lande argued that random demographic and environmental events drive small wild populations to extinction before genetic factors cause problems. Environmental events, ranging from annual variation in climatic variables (such as rainfall) to catastrophes (such as fires and disease epidemics), do increase the probability of extinction and it is extremely difficult to isolate genetic effects from these other effects because inbreeding typically interacts with demography by reducing fecundity, juvenile survival, and life span. Consequently, other researchers continued to question the significance of genetic factors because of the lack of direct evidence that inbreeding can contribute to the extinction of wild populations. However, more recent work has supported the original view that genetic factors can contribute to the extinction of wild populations.

The hypothesis that nongenetic factors would usually drive species to extinction before genetic factors could impact them was evaluated by comparing heterozygosity levels in threatened taxa with those of taxonomically related nonthreatened taxa (Spielman *et al.*, 2004). In 77% of 170 comparisons, heterozygosity levels were lower in the threatened taxa, which would not be expected under the hypothesis that extinction generally occurs prior to genetic impacts. Thus, the view that extinction would occur prior to genetic impacts was rejected. Direct evidence for an impact of inbreeding on extinction risk was provided by studies on the Glanville fritillary butterfly in Finland (Saccheri *et al.*, 1998). They studied a metapopulation (a population of populations) of this butterfly that consists of

numerous small populations that breed in about 1600 dry meadows of different sizes at varying distances from each other. Caterpillars feed in conspicuous family groups of 50–250 individuals and the smallest populations consist of the offspring of a single pair. Populations in individual meadows often go extinct, but many meadows are eventually recolonized, with an average of 200 extinctions and 114 colonizations per year.

Because small population size results in both inbreeding and loss of genetic variation, the researchers were able to use the degree of genetic variation in each population as a measure of the extent to which it was inbred. They sampled 42 populations and found that populations with less genetic variation were more likely to go extinct. Furthermore, genetic diversity predicted extinction risk after they had accounted for all other known causes of extinction in this well-studied butterfly metapopulation. Inbreeding reduced egg-hatching rate and larval survival, lengthened the duration of the pupal period (so that inbred pupae were more likely to be parasitized), and shortened female life span (so that inbred females tended to lay fewer eggs).

Extensive replicated experiments with the annual plant, *Clarkia pulchella*, also found that decreased genetic effective population size (resulting in both inbreeding and loss of alleles through genetic drift) increased the probability of population extinction substantially above that expected from demographic and environmental stochasticity alone. A Spanish study of the shore campion plant reached similar conclusion.

More indirect evidence that inbreeding depression increases the extinction risk of threatened taxa has been provided by stochastic computer simulation models that include inbreeding effects. For example, O'Grady *et al.* (2006) found that incorporating realistic levels of inbreeding depression in population viability models decreased median times to extinction for 30 species by an average of 37%.

Outbreeding and Outbreeding Depression

Mating between distantly related individuals, such as individuals from different populations or subspecies, is called outbreeding. Crossing populations may increase reproductive fitness by increasing heterozygosity and thus preventing the expression of deleterious recessive alleles ("hybrid vigor,") or may decrease fitness because of various kinds of genetic incompatibilities between the genes from the different populations (outbreeding depression). Outbreeding depression is a reduction in reproductive fitness, reduced ability to mate or pollinate, fertilize, produce offspring, survive, or reproduce, in the first or later generations following the crossing of populations.

Causes of Outbreeding Depression

Three main mechanisms have been proposed as causes of outbreeding depression: chromosomal incompatibilities, adaptive differentiation among populations, and genetic drift. It is well known that fixed chromosomal differences between two populations increase the probability of outbreeding depression if the populations are crossed and many cases have been documented in both animals and plants. Although the risk of outbreeding depression differs according to the type of chromosomal differences between populations, it generally increases as the number of fixed differences increases. The risk is greatest with polyploidy, followed by translocations, centric fusions, and inversions.

There is also theoretical and empirical evidence for outbreeding depression between populations that have adapted to different environments. Natural selection can promote rapid development of reproductive isolation, which results in outbreeding depression, between populations in differing environments. A positive association between ecological divergence in reproductive isolation was found in over 500 species pairs of vertebrates, invertebrates, and plants. Most genes involved in prezygotic reproductive isolation that have been investigated show molecular signals of positive selection. Whether or not all cases of reproductive isolation not due to chromosomal incompatibilities are due to adaptation to differing environments remains controversial but adaptation is clearly involved in many cases.

The ecological mechanism of outbreeding depression requires that populations develop different adaptations in response to different local environments. Crossing individuals from the populations may then produce progeny that are less well suited to either local environment. This kind of outbreeding depression is usually observed in the F1 generation. Populations isolated in different environments may also develop coadapted gene complexes, which are sets of genes that have evolved together to produce a fitter phenotype. Crossing individuals from populations with different coadapted gene complexes could then disrupt these complexes and reduce reproductive fitness in the F2 and subsequent generations, when recombination during meiosis disrupts gene complexes.

Genetic drift in small populations has been proposed as a mechanism leading to reproductive isolation (and hence outbreeding depression). Genetic drift can result in chromosomal differences between populations and this often results in outbreeding depression. However, when the chromosomes in two populations are the same, selection is much more likely to result in reproductive isolation than genetic drift. Both theory and extensive empirical evidence indicate that reproductive isolation results predominantly from selection, rather than drift. For example, phenotypic differences due to natural selection were a better predictor of reproductive isolation than genetic distance due to drift in several groups of organisms, including poison-dart frogs, anole lizards, and cichlid fish.

Evidence for Outbreeding Depression

Evidence for outbreeding depression comes primarily from organisms with extremely limited dispersal, such as some plants, copepods, and scale insects, or from crosses between individuals from vastly different geographic sources, or with significant chromosomal differences. Outbreeding depression appears to be more common in plants than in animals. A recent review listed 35 cases of outbreeding depression in vertebrates, invertebrates, and plants. There is not much evidence for outbreeding depression in larger animals, although

it has been observed in crosses where there are chromosomal differences between populations such as in dik-diks and spider monkeys. Such cases usually indicate the existence of unrecognized species or subspecies. The Arabian oryx suffers simultaneously from inbreeding and outbreeding depression. Outbreeding depression appears to be the larger effect but there are chromosomal differences segregating in the population. The most widely quoted mammalian case concerns two ibex populations in Slovakia that lived in very different environments. Several studies have failed to find outbreeding depression in mammals, including in rhesus macaques and saddle-back tamarins. Studies also found no outbreeding depression in survival in several captive populations, including crosses between Borneo and Sumatran orangutans. Lacy and colleagues conducted numerous crosses among five subspecies of deer mice. These subspecies included closely related populations from similar, contiguous habitats, as well as very divergent subspecies that had long been isolated in dissimilar habitats. In general, the benefits of increased heterozygosity outweighed the costs of disrupted gene complexes.

Predicting the Risk of Outbreeding Depression

The question of whether crossing geographically distinct populations of a species is likely to have beneficial or detrimental effects is an important conservation issue because many species now exist as a series of isolated or semi-isolated populations due to human activities. Inbreeding depression and loss of genetic diversity will ultimately contribute to the extirpation of many isolated populations unless gene flow with other populations is reestablished. The introduction of individuals from other populations to improve reproductive fitness and restore genetic diversity can be an effective management strategy to aid the recovery of small inbred populations with low genetic diversity and has been successful in several natural populations including deer mice, wolves, mountain lions, greater prairie chickens, Swedish adders, desert topminnows, and several species of plants. However, before establishing gene flow between isolated populations, the risk of outbreeding depression should be evaluated.

Inadvertent crossing of populations that belong to separate species can result in outbreeding depression. This problem occurred with small antelopes called dik-diks, as zoos did not originally appreciate that the individuals in captivity belonged to more than one species. Thus, managers should make sure that two populations to be crossed belong to the same species. Fixed chromosomal differences between populations increase the probability of outbreeding depression so karyotyping of populations prior to crossing them is advisable. Predicting the likelihood of outbreeding depression resulting from adaptations to differing environmental conditions is more challenging. Because reproductive fitness is a quantitative character, Frankham and colleagues (Frankham *et al.*, 2011) recently provided theoretical and practical guidance in this area based on an extension of the breeder's equation:

$$\sum AD_t \sim h^2 \sum_{i=1}^{t} S_i(1 - 1/[2N_e])^{t-1}$$

where AD is the adaptive difference between two populations, h^2 the heritability of reproductive fitness, S_i the selection

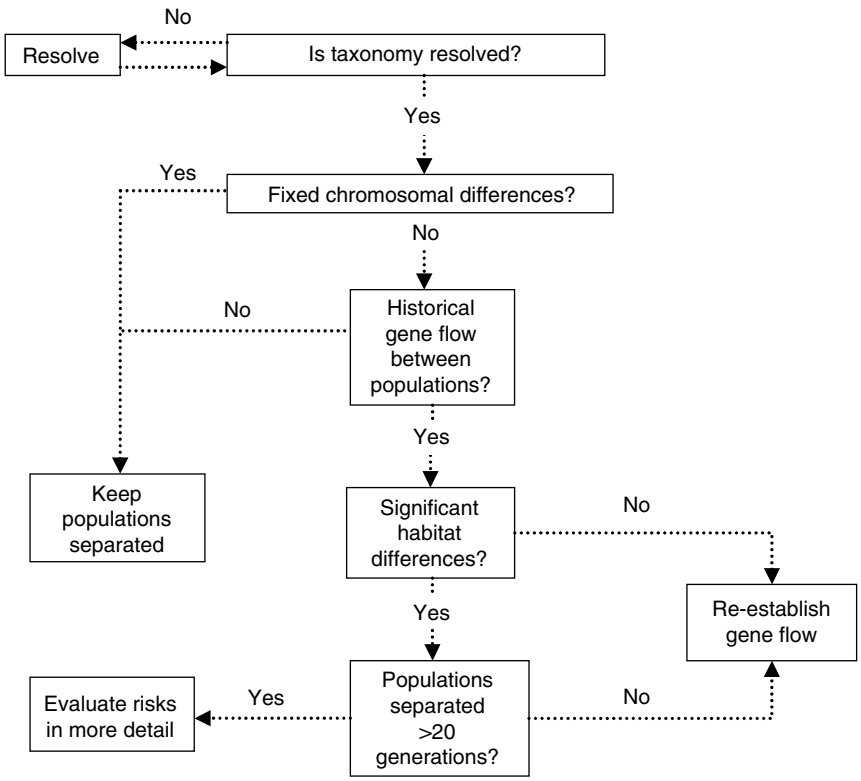


Figure 5 Decision tree for predicting the risk of outbreeding depression when two populations are crossed.

differential in the ith generation, and N_e the effective population size. This work indicated that the degree of adaptive differentiation between two populations, and thus the probability of outbreeding depression if they are crossed, is an increasing function of four factors: (1) the selection differential, which increases with the difference between the environment to which the population was previously adapted and the current environment or environments of the populations to be crossed; (2) the heritability of the traits involved (which depends on the genetic diversity in the populations); (3) the effective population size (which is almost always smaller than the actual population size); and (4) the number of generations the populations have been separated. Calibration of the equation with data from the literature indicated that thousands of generations are required before populations isolated in the same environment begin to develop reproductive isolation but only dozens of generations are required if the populations are adapting to differing environments.

Based on these ideas, Frankham and colleagues developed a decision tree to assess the risk of outbreeding depression if two populations are crossed (**Figure 5**). Applying the tree to known cases successfully predicted the occurrence of outbreeding depression. The questions are based on the findings that small populations whose taxonomy is well-studied and generally accepted, that exhibit no fixed chromosomal differences, have experienced gene flow between the populations within the past 500 years, and inhabit similar environments, or have been in different environments for less than 20 generations have little risk of exhibiting outbreeding depression if crossed. In such circumstances, establishing gene flow between the populations is likely to have beneficial genetic and demographic consequences. Conversely, if the two populations belong to separate species, or exhibit fixed chromosomal differences, or have been isolated for 500 years or longer, or have been living in significantly differing environments for more than 20 generations, there is a modest to high risk of outbreeding depression if the populations are crossed. However, even when crosses of populations result in outbreeding depression, it will not be a long-term phenomenon unless F1 individuals are sterile or have very low fitness, as natural selection will act on the extensive genetic variation in the hybrid population, resulting in better adaptation to its environment.

See also: Ecological Genetics. Genetic Diversity. Population Genetics

References

Crnokrak P and Roff DA (1999) Inbreeding depression in the wild. *Heredity* 83: 260–270.

Dudash MR and Fenster CB (2000) Inbreeding and outbreeding in fragmented populations. In: Young AG and Clarke J (eds.) *Genetics, Demography and Viability of Fragmented Populations*, pp. 35–53. Cambridge: Cambridge University Press.

Frankham R, Ballou JD, and Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Frankham R, Ballou JD, and Briscoe DA (2010) *Introduction to Conservation Genetics*, 2nd edn. Cambridge, UK: Cambridge University Press.

Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR, and Fenster CR (2011) Predicting the risk of outbreeding depression. *Conservation Biology* 25: 465–475.

Keller LF (1998) Inbreeding and its fitness effects in an insular population of song sparrows (*Melospiza melodia*). *Evolution* 52: 240–250.

Keller LF and Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution* 17: 230–241.

Lande R (1988) Genetics and demography in biological conservation. *Science* 241: 1455–1460.

O'Grady JJ, Brook BW, Reed DH, Ballou JD, Tonkyn DW, and Frankham R (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Conservation Biology* 133: 42–51.

Ralls K and Ballou J (1986) Captive breeding programs for populations with a small number of founders. *Trends in Ecology and Evolution* 1: 19–22.

Ralls K, Ballou JD, and Templeton A (1988) Estimates of lethal equivalents and the cost of inbreeding in mammals. *Conservation Biology* 2: 185–193.

Saccheri I, Kuussaari M, Kankare M, Vikmam P, Fortelius W, and Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature* 392: 491–494.

Spielman D, Brook BW, and Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceeding of the National Academy of Sciences, USA* 101: 15261–15264.

Indicator Species

John H Lawton, Imperial College, London, UK
Kevin J Gaston, University of Exeter, Exeter, UK, and University of Sheffield, Sheffield, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 437–450, © 2001, Elsevier Inc.

Glossary

Acid deposition Anthropogenic acidification of terrestrial and freshwater ecosystems by (primarily) sulfuric acid, derived from sulfur dioxide produced by burning oil and coal and deposited in rain and snow (acid rain), directly as particles (dry deposition) and as cloud droplets.

All taxa biodiversity inventory (ATBI) The idea, first suggested by D. H. Janzen, that it might be feasible to produce a complete species list for all the organisms living in one place, a hectare of tropical forest, for example. The goal has so far proved elusive.

Bioassay The use of living cells or organisms to make quantitative or qualitative measurements of the amounts or activity of substances.

Community An assemblage of species populations that occur together in space and time.

Ecotoxicology The use of test organisms (e.g., the water flea, *Daphnia*) to study the toxicity, pathways of accumulation, and breakdown of chemicals, particularly those manufactured by humans (e.g., pesticides).

Endemic species Species confined in their distribution to a particular geographic region. The size of the region is arbitrary (a species can be endemic to North America or to a tiny island).

Hot spots A word with several distinct meanings. Here it is used to denote sites unusually rich in a particular group of species (e.g., birds), compared with average sites in the same geographic region. (The converse is a "cold spot.") Has also been used to denote centers of endemism (*see* Endemic Species), which need not be unusually species rich; it is not used in this context here.

Paleoclimatology The study of past climates from fossils and other traces left in the geological record.

Reserve selection algorithms Mathematical techniques used to maximize efficiency in the selection of protected areas for conservation. The efficiency criteria vary with circumstances but may, for example, be the minimum number of reserves with every species represented, or minimum cost.

Species as Indicators of the State of the Environment

There are three distinct uses of the term "indicator species" in research in ecology and biodiversity. They are a species, or group of species, that do the following:

1. Reflect the biotic or abiotic state of an environment.
2. Reveal evidence for, or the impacts of, environmental change.
3. Indicate the diversity of other species, taxa, or entire communities within an area.

This article explains, provides examples of, and evaluates each of these uses of the term, focusing primarily on terrestrial and freshwater ecosystems; broadly similar conclusions apply to marine ecosystems, but marine examples lie beyond the scope of the article. We pay most attention to the third use of the term "indicator species," because this seems most appropriate for an encyclopedia devoted to biodiversity. The most up-to-date evaluation and review of indicator species in the scientific literature is by McGeoch (1998). She concentrates on terrestrial insects as "bioindicators" (in all three senses of the word) but the general principles that she discusses extend to all ecosystems and organisms, not just to terrestrial insects.

Everybody knows that living organisms are sensitive to the state of their environment. Pollution from human activities kills many species and reduces the abundance of others. These changes in abundance can be used to assay the state of the environment.

An Example: Acid Deposition

Sulfur dioxide, produced by burning fossil fuel, particularly coal, enters the atmosphere and is eventually deposited on terrestrial and freshwater ecosystems via three routes: (a) as tiny solid particles, (b) washed from the air in rain or snow, or (c) as droplets formed in clouds. Deposition often occurs hundreds of kilometers from the source. Dissolved in water, sulfur dioxide forms sulfuric acid, resulting in what is frequently referred to as "acid rain," but because there are three principal routes involved in its transfer to terrestrial and freshwater ecosystems, it is more correctly called "acid deposition" (Erisman and Draaijers, 1995). Sulfur dioxide is not the only source of acidification; oxides of nitrogen, again produced by burning fossil fuel, are also involved, but sulfur dioxide is the main agent of acidification in most ecosystems.

In terrestrial ecosystems, this deposition kills lichens and acidifies the soil, leading to changes in the vegetation. Lakes become progressively more acidic as deposition loads increase, until eventually they may become virtually lifeless. A trained biologist, visiting for the first time an area subject to acid deposition, will often be able to deduce that the habitat is being polluted simply by looking at the species that are present and those that ought to be there but are not. Beautifully

clear Scandinavian lakes, lacking any fish or amphibians, supporting few birds and a species-poor and taxonomically unusual invertebrate fauna, have been reduced to this impoverished state by the transnational export of sulfur dioxide from coal-burning power-stations in the United Kingdom and elsewhere in Europe. Here, living organisms act as powerful indicators of the state of the environment and the damage being done to it by human activities, often performed many hundreds of kilometers away.

Management of European Rivers

Because the species composition and richness of biological communities change as the environment changes, we can use species as indicators of the state of the environment for practical management purposes. The techniques have been particularly well developed to assess organic and inorganic pollution in European rivers, managed for recreation, fisheries, and drinking water. The advantages of using living organisms as indicators of water quality are that they avoid the need for expensive chemical analyses, and, probably more important, organisms integrate the impacts of pollutants over space and time. All chemical traces of a major pollution incident may disappear from a river in a matter of hours as the pollution is flushed from the system. Nonetheless, the biotic community may show evidence of the damage for many months. It is extremely difficult, and prohibitively expensive, for chemists to measure all the organic and inorganic chemical pollutants entering a river, and it is certainly impossible for them to work out what all the combined impacts of such a cocktail might be. But living communities reflect the integrated effects of all the compounds that find their way deliberately and accidentally into watercourses, and hence they act as sensitive indicators of the state of the environment. A valuable source of further information on the use of living organisms to monitor the environmental health of rivers and lakes is provided by Rosenberg and Resh (1993).

Two widely used European examples of this approach are the German Saprobic Index, and RIVPACS in the United Kingdom. RIVPACS is now used by the Environment Agency to manage UK rivers. Both the German and UK approaches require accurately identified samples to be taken of the organisms found along sections of the river. RIVPACS focuses on invertebrates, the German index on invertebrates, microbes, and higher plants. Both rely on the fact that some species are extremely tolerant of pollution (the aquatic larvae of some chironomid midges, for instance), while others are extremely sensitive, particularly to the low oxygen levels produced by organic pollution (for instance, the larvae of many mayflies). The species present in the samples are given scores, depending on their known tolerances, and the data from all species are combined to produce a composite and very sensitive index of pollution levels for any particular section of a river.

Widespread Application

Use of organisms to indicate the state of the environment is widespread, taxonomically and geographically, for a wide range of environmental issues. Use of species as indicators of the state of the environment is not confined to freshwater, or to Europe and North America. A wide variety of organisms has been suggested, or used, as indicators of human impacts. In Europe, suites of plant, fungal, and insect species are only found in, and hence are good indicators of, ancient woodland; they are entirely absent from plantations, even though these may be several hundred years old. The use of lichens as sensitive indicators of air pollution is well known, but organisms as different as mites and geckos (agile, climbing lizards) have been used, or suggested, for similar purposes. Lichens have also been used as indicators of fire history in Brazilian cerado (a type of dry, scrubby forest), tiger beetles as indicators of tropical forest degradation in Venezuela, and day-flying Lepidoptera (butterflies and moths) as indicators of the state of seminatural grasslands for conservation in Europe. Many other similar examples exist.

Interpretation Requires Care

In these and similar cases, considerable care is needed before a species or group of species can be used as reliable indicators of damaging (or beneficial) human impacts on ecosystems. All populations of living organisms fluctuate over time and vary in abundance spatially, because of natural variations in the weather, normal changes in the physical environment, and fluctuations in the abundances of natural enemies, competitors, and essential resources (food and shelter). Just because one or more species is declining does not mean that human impacts are to blame. In the case of lichens and atmospheric pollution or freshwater invertebrates and river quality, the links between anthropogenic pollutants and changes in the distributions and abundances of organisms are thoroughly researched and well understood. But even quite major declines in some species have proved exceptionally difficult to link to damage to the environment caused by people.

Amphibian Decline

The so-called amphibian decline is a particularly dramatic example (Blaustein and Wake, 1995). In many parts of the world, population biologists interested in amphibians (frogs, toads, newts, salamanders, etc.) have recently become alarmed by apparent major declines in the abundance, and the complete disappearance, of many species from areas where formerly they were common, often in regions apparently remote from human impacts. The declines are not happening everywhere, and the magnitude of many of those that have been claimed is difficult to assess because of the lack of long-term data prior to the supposed population collapses; some of them may be perfectly natural. The worrying aspects of the phenomenon are that while it is apparently global in scope, the causal mechanism (or mechanisms) remains obscure. It has been suggested, for example, that the amphibian decline is indicative of rising global levels of damaging ultraviolet light (UV-B) caused by loss of the earth's protective stratospheric ozone layer. Amphibian eggs, exposed in shallow water, and the adults with their thin wet skins may be particularly sensitive to UV-B, as are human sunbathers without sunblock.

Others doubt the explanation. More recently a global pandemic has been implicated. But what should suddenly trigger lethal outbreaks of disease in amphibians is unclear.

Environmental Toxicology

In all the examples so far, the organisms being used as actual or possible indicators of environmental health have been in their natural environment. There is another related but quite separate way in which biologists use the sensitivity of organisms to set environmental standards, namely in the science of environmental toxicology, or ecotoxicology for short. In many areas of human endeavor, the aim is to apply some beneficial technology with minimum environmental damage. Crop spraying with pesticides is a good example, and so is the discharge of treated effluent from a factory. Some environmentalists claim that these types of operations should not lead to any environmental contamination; factories should have zero discharges, and if we must use pesticides, they should be targeted to reach only the crop and the pest and not, for example, the soil, nontarget organisms, or adjacent watercourses.

However, zero discharges or precision pesticides, if they can be achieved at all, can often only be obtained at great economic cost. The more pragmatic solution is to ask whether there are minimal levels of discharge, spray drift into watercourses, and so forth that cause no detectable environmental damage. To provide answers to this admittedly difficult question, environmental toxicologists use a wide variety of laboratory bioassays with standard organisms. Examples from freshwater include the alga, *Chlorella vulgaris*, the water flea, *Daphnia magna*, the amphipod shrimp, *Gammarus pulex*, and the rainbow trout, *Salmo gairdneri*. The fundamental problem is to try and establish acceptable levels of contamination. Defining "acceptable" obviously requires political as well as biological judgment. However, traces of a compound in water, air, or soil that cause no detectable changes in the performance (growth, survival, or reproduction) of the test organisms are clearly more acceptable than doses that kill 50% of the population (so called LD_{50} levels). Basically, the bioassays seek to set environmental standards for levels of potential pollutants in soil, air, and freshwater, using a range of standard laboratory organisms as indicators (Shaw and Chadwick, 1998), but there can be no absolute standards about what is safe or acceptable. The general trend in modern societies is for standards to gradually tighten.

Species as Indicators of Environmental Change

If the amphibian decline (discussed in the previous section) is real, it is an example of a group of organisms acting as indicators not only of the state of the environment, but also as indicators of ongoing changes to the global environment, albeit of an unknown nature. In other words, given that species are sensitive to the condition of their environment, monitoring organisms not only tells you about the current state of an environment, but *repeated* monitoring can tell you about *changes* in that environment. To act as indicators of change rather than current environmental health, it is necessary to have at least two sets of data on the particular indicator species in question, taken in the same way, at the same place(s), on two separate occasions. More frequent sampling allows greater confidence in the direction of apparent trends and the detection of more subtle environmental changes.

Not all Monitoring is about Environmental Degradation

Not all monitoring of species seeks to record environmental degradation. Increasingly after mining operations, for example, mine operators are required to restore spoil heaps and mine pits by sowing or planting native vegetation. Monitoring selected groups of common animals on nearby undisturbed control sites and on the restored land can give a good indication of the recovery of the entire ecosystem and of the success of the restoration project. For instance, when biologists monitored ant assemblages on abandoned, replanted bauxite mines in Australia, they found that the ants provided a good indication of the recovery of these ecosystems. Even after 14 years there were still differences between the ant communities found in the natural *Eucalyptus* forest and the restored land.

Historical Records of Change

Lake Acidification

It may not always be necessary to sample in real time. When anthropogenic acidification of lakes was first discovered, many people doubted that the phenomenon was real. In particular, there was considerable opposition to the notion from the power-generating industry, because solving the problem (by burning low-sulfur coal, adding "scrubbers" to power station chimneys to remove sulfur dioxide, or switching to natural gas) was inevitably going to be expensive. After all, there were few historic data on the state of the acidified lakes. Perhaps they had always been that way?

Resolving the problem required knowledge of the fact that lake phytoplankton (the tiny, unicellular plants that float in the upper layers of lakes) are extremely sensitive environmental indicators, because different species grow best in very different conditions determined by nutrient status and pH (acidity). When algae die, they sink to the bottom where their bodies and characteristic pigments are buried and some are preserved (incipient fossils), particularly the resistant, silicious outer cases of a group called diatoms. An undisturbed core through the sediments records the history of a lake's phytoplankton, with the oldest flora at the bottom. Cores showed unequivocally that many Scandinavian lakes that are acid now were not acid before the Industrial Revolution; the oldest diatoms—species not found in acid lakes—are gradually replaced in the sample column by acid-tolerant species. Diatoms are wonderfully sensitive indicators of environmental change (**Figure 1**).

Plants and Carbon Dioxide

Herbarium specimens (pressed plants collected for taxonomic purposes) and fossil leaves can also be used as indicators of past environmental change. Another consequence of the rapid rise in the burning of fossil fuel since the Industrial Revolution has been an accelerating rise in the concentration

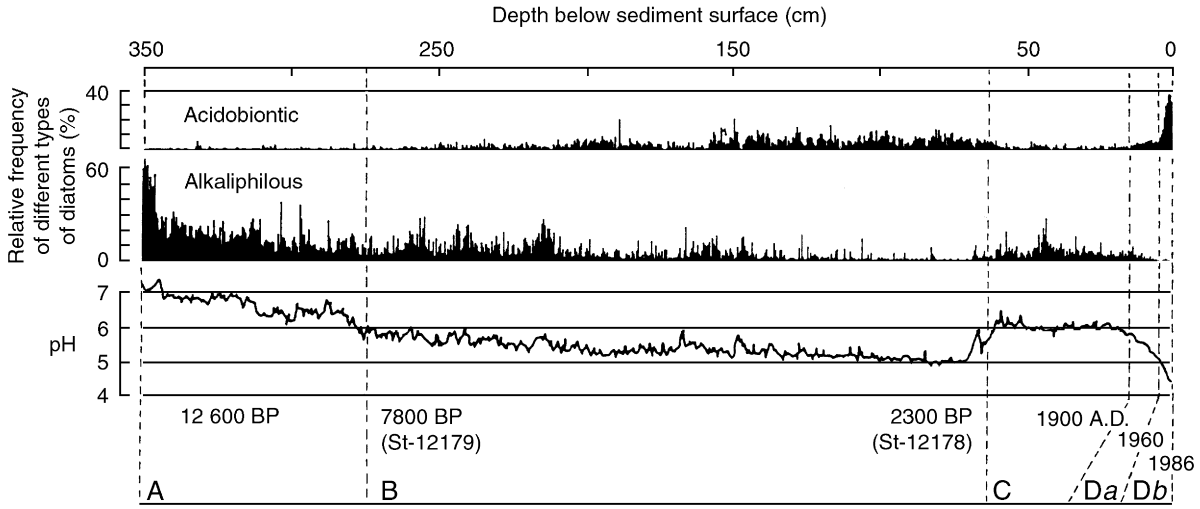


Figure 1 The pH history of Lilla Öresjön, a 0.6 km^2 lake in southwest Sweden. A core of the bottom sediments 3.5 m long records the history of the lake extending back to 12600 BP, using the valves ("shells") of diatoms preserved in the deposits. Different species of diatoms have different pH preferences and can be classified accordingly. Acidobiontic species thrive in acid waters; alkilophilous species prefer more alkaline conditions. Combining data from the remains of all species of diatoms allows the pH history of the lake to be reconstructed. The lake has passed through four pH periods. *A*, an alkaline period after deglaciation. *B*, a naturally more acidic period. *C*, a period with higher pH, which started at the same time as agricultural expansion in the region, and *D*, a rapid, recent period of acidification. The post-1960 phase has no similarity with any of the previous periods. Reproduced from Renberg I (1990) *Phil Trans. R. Soc. London B* 327: 357–361, with permission from Royal Society of Publishing.

of atmospheric carbon dioxide, one of the main agents of "global warming." We will deal with species as indicators of anthropogenic global climate change (as it is more accurately known) later. Here we want to focus on physiological and developmental responses within single species to rising carbon dioxide.

If plants are grown in a greenhouse under different atmospheric carbon dioxide concentrations, from below the pre–Industrial Revolution levels of about 280 parts per million by volume (ppm), through what are roughly present levels of 350 ppm, to levels that may be reached by the end of the 21st century (700 ppm), several interesting things happen. In particular, in the present context, stomatal densities on the undersides of the leaves decline. Stomata are the tiny pores in the leaf surface through which plants take up carbon dioxide (needed for photosynthesis), and through which they lose water vapor. It has been known for a long time that plants control the opening and closing of stomata to optimize carbon dioxide uptake and reduce water loss. More surprising, we now also know that plants grown in high carbon dioxide have lower densities of stomata; something happens during leaf development to reduce the number of stomata. How and what is currently unclear. Why is simple enough. In a high carbon dioxide world, the plant needs fewer stomata to take up the carbon dioxide it requires and hence can satisfy the needs of photosynthesis and reduce water loss by developing fewer pores in the leaves.

Now back to those herbarium specimens and fossil leaves. If you look at 200-year-old (and very precious) herbarium and modern specimens of the same species, sure enough, stomatal densities decline as global atmospheric carbon dioxide levels increase (**Figure 2A**). The same approach has recently been used to try and trace atmospheric carbon dioxide levels throughout most of the Phanerozoic, from the time when plants first colonized the land. Here the method is more contentious, because different species of truly fossil plants with presumed similar growth forms have to be used in different geological periods. Nevertheless, the pattern of apparent changes in global atmospheric carbon dioxide concentrations over hundreds of millions of years, revealed by this method (**Figure 2B** and **C**), are in reasonable agreement with alternative, independent, and also contentious geochemical methods. Here is a really unusual use of species as indicators of environmental change.

Species as Indicators of Climate Change

The Sensitivity of Species to Climate: Fossils Again

Current, rapidly rising concentrations of atmospheric carbon dioxide are the primary cause of anthropogenic global climate change. However, the earth's climate has always changed, naturally, with no intervention from human beings. One of the ways we know this is through the careful documentation of the types and distributions of organisms in the fossil and subfossil record. The science of paleoclimatology, which seeks to reconstruct the history of earth's climate, relies heavily on changes in fossil and subfossil species assemblages to deduce what the earth's climate was like thousands or even millions of years ago. To take one example, in the modern world many types of corals occur exclusively in tropical marine environments; it is a winning bet that fossil corals of the same type indicate an ancient tropical sea, even though the rocks bearing the fossils may now lie in much colder parts of the world.

In more recent geological time, we can use changes in the distributions and abundances of plants and animals to trace

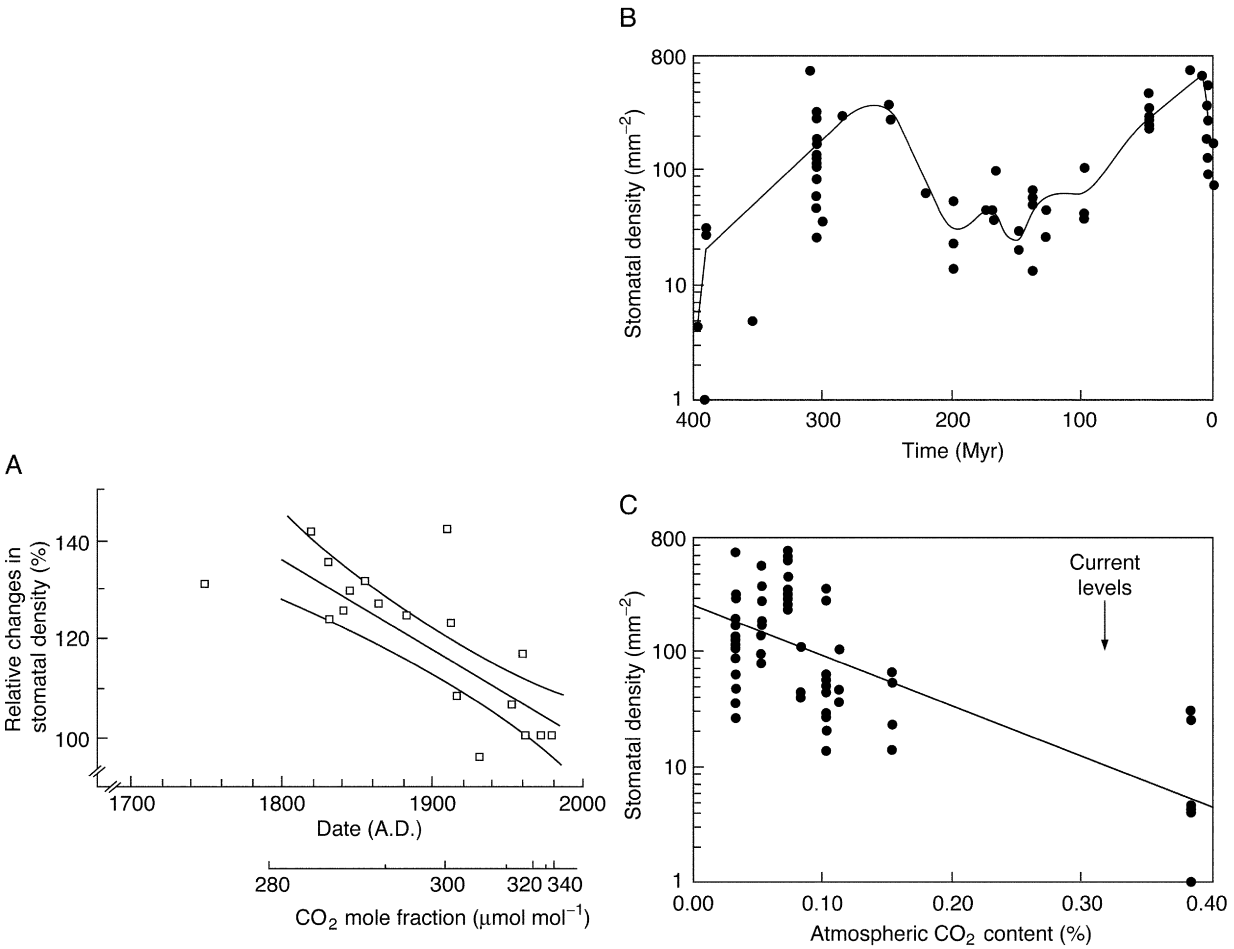


Figure 2 Changes in stomatal densities of leaves as indicators of changes in atmospheric carbon dioxide concentrations. (A) Percentage changes in stomatal densities on the lower surfaces of the leaves from eight species of trees and shrubs collected from the English midlands and preserved as herbarium specimens. The oldest specimen was collected in 1750. As atmospheric CO_2 has risen during and since the industrial revolution, so stomatal densities have fallen. (B, C) Assuming that similar effects occur in all species of plants, fossil leaves can be used as indicators of atmospheric CO_2 concentrations extending back many millions of years. The CO_2 content of the earth's atmosphere appears to have fluctuated markedly and apparently naturally during the past 400 million years. Reproduced from Beerling DJ and Woodward FI (1997) *Bot. J. Linn. Soc.* 124: 137–153, with permission from Wiley.

major changes in the earth's climate during the Holocene (the most recent geological past) and Pleistocene glacials and interglacials. Plant remains preserved in packrat middens in dry air of the southeastern United States attest a much wetter climate only a few thousand years ago. Hippopotamus bones and teeth dug up from under Trafalgar Square provide unequivocal evidence of a much warmer London. Pollen grains preserved in peats and lake sediments record in exquisite detail the march northward of European and North American forests from the end of the last glaciation 12,000 years ago (Huntley and Birks, 1983). The forests spread with remarkable speed (an average of about 200 m per year, but sometimes as fast as 2 km a year) to achieve present distributions in the northern parts of both continents from glacial refugia thousands of kilometres to the south (**Figure 3A** and **B**). The information is not won easily. It requires huge patience and great skill to identify thousands upon thousands of pollen grains extracted onto microscope slides. But once done, the record reads like a speeded-up movie, as spruce, oaks, white pine, hemlock, beech, and chestnut swept north in successive waves through what is now the United States and Canada; in the more species-poor forests of Europe, pines were followed by birch, then oak. These invasions are as dramatic as any in human history, but they were silent and recorded only by pollen grains.

Contemporary Changes in Species Distributions

Historical changes aside, there is now no doubt that the world is currently warming quite rapidly. An upward trend in global annual mean surface temperatures is apparent from about 1920, particularly over the last two decades (from c. 1980); global mean surface temperatures in July 1998 were the highest ever recorded. Do organisms act as indicators of these changes, perhaps, as with the freshwater species discussed earlier, acting subtly to integrate several of the changes humans find difficult to comprehend in the bald statistics? Climate change does not simply involve warming; it involves changes in rainfall, extreme weather events (droughts and

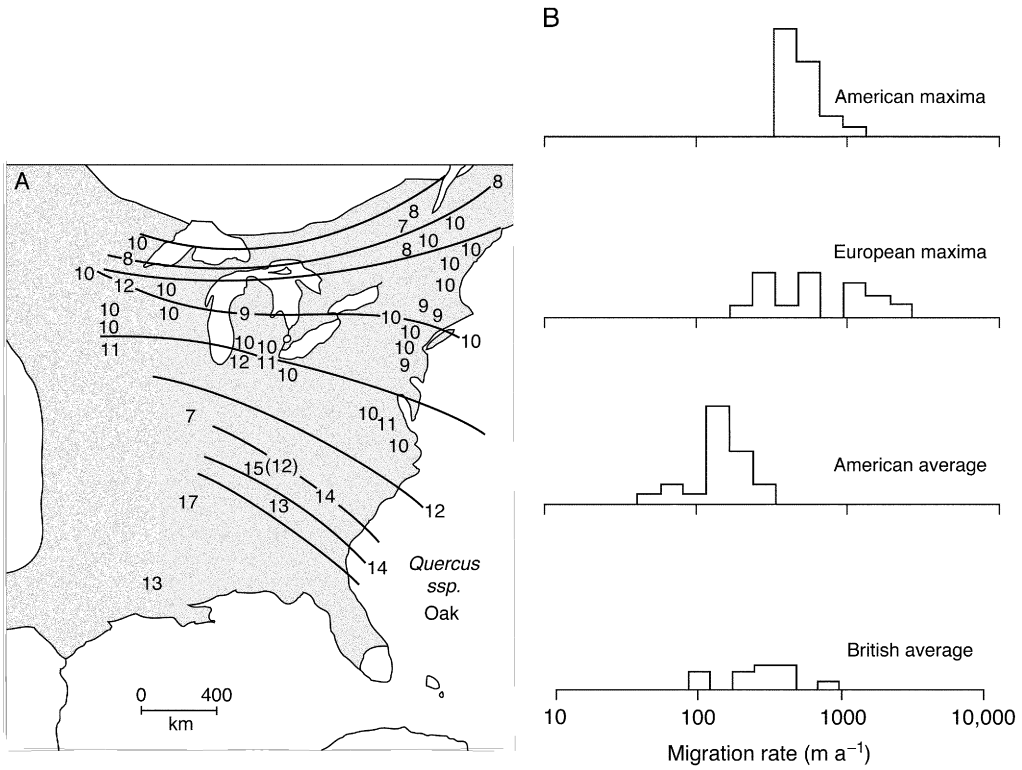


Figure 3 The migration of trees across North America and Europe after the end of the last glaciation, revealed by pollen remains in lake sediments and peat. (A) The spread of oaks in North America, with radiocarbon ages in thousands of years (contours) and the present range of the genus (shaded). (B) Estimates of the overall rates of spread of trees on two continents, based on data of the type shown in part A. Reproduced from Williamson M (1996) *Biological Invasions*. London: Chapman & Hall.

storms), and even locally cooler conditions. All these complex changes should show up in changes in the distributions and abundances of organisms.

They do. Species are proving to be extremely sensitive indicators of contemporary climate change, where historical records allow decent reconstruction of former and current distributions. Populations of Edith's checkerspot butterfly *Euphydryas editha* are disappearing from southern California and northern Mexico, at the current southern end of its distribution, and from more lowland sites; sites where previously recorded populations still exist are on average 2° further north than sites where populations went extinct (**Figure 4**). These are exactly the changes we would expect in a warming world. Twenty years ago in northwest Europe, little egrets *Egretta garzetta* (small white herons) used to be rare visitors from the Mediterranean. Now they are breeding in northern France and southern England in an astonishing expansion of range. Populations of many other European birds, butterflies, and other organisms are spreading north at the present time, as the climate warms.

Of course, none of this tells us whether the climate change that is certainly happening is "natural"—it could have happened anyway and may have nothing to do with anthropogenically produced greenhouse gasses—or whether it is indeed due to human activities. Using species as indicators of climate change tells us unequivocally that the earth's climate is changing, but so does the mercury in the thermometer. What neither tells us is why, and no end of work on species as indicators will solve that dilemma. As we have already seen, this situation is not unique to climate change. It generally holds whenever we use species as indicators of the state of the environment. Indicator species can tell us whether an environment is, or is not, changing. They do not tell us why the changes are taking place. That almost always requires additional detective work, although knowledge of an organism's biology will frequently provide valuable clues. Three examples, using birds as indicators, illustrate the problem in more detail.

Birds as Indicators of Large-Scale Environmental Changes

Birds are widely used indicators, because in Europe, North America, and other parts of the world where there are large armies of amateur bird watchers their populations and distributions have been recorded well enough, for long enough, to reveal major environmental trends.

Peregrine Falcons and DDT

The catastrophic collapse of peregrine falcon *Falco peregrinus* populations throughout the northern hemisphere in the 1950s signaled widespread contamination of the environment by chlorinated hydrocarbon insecticides, first DDT, then other compounds such as aldrin and dieldrin. The total, and rapid, disappearance of these dramatic birds signaled to

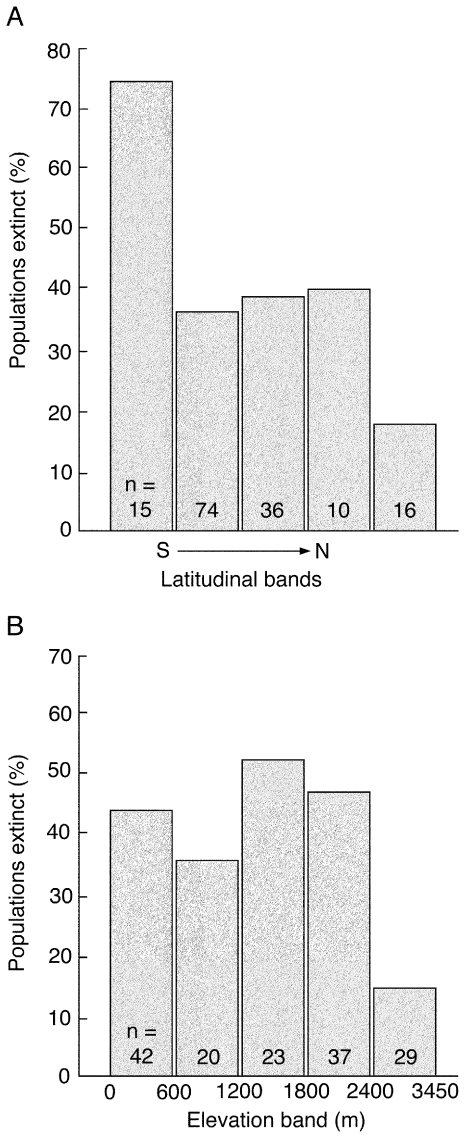


Figure 4 The fate of 151 previously recorded populations of Edith's checkerspot butterfly, *Euphydryas editha*, in western North America. The populations ranged from northern Mexico to southern Canada and were visited by Camille Parmisan and other biologists between 1992 and 1996. Populations that had disappeared because of habitat degradation (e.g., loss of usable host plants) were omitted from the analysis. Dividing the populations into five, evenly spaced latitudinal bands between 30° N and 53° N (A) reveals that significantly more southern populations have gone extinct than northern populations; sites where previously recorded populations still exist were, on average, 2° further north than sites where populations were extinct. Extinctions were also higher at lower altitudes (B) (*n* is the number of populations in each latitudinal or altitudinal band). Both results are consistent with the effects of global climate warming on the butterfly, leading to a northward and upward shift in its geographical range. Reproduced with the permission of McMillan Journals Ltd, from *Nature* **282** (1996), page 766.

ornithologists that something was seriously wrong with the environment, but what? It took a great deal of clever biological detective work (see Ratcliffe, 1980) to link the decline of peregrine populations to the accumulation of these persistent pesticides up the food chain, resulting in eggshell thinning, reproductive failure, and (in extreme cases) direct poisoning of adult birds. Although some populations have now recovered, signaling a recovery in environmental quality, the species is still missing from many parts of its former range—some coastal populations in England, for instance. Nobody knows why.

Migratory Songbird Declines in North America

In North America, considerable concern is currently being expressed over widespread declines in summer migrant birds, particularly warblers. Unlike the so-called amphibian decline, nobody questions the phenomenon; just like the amphibian decline, nobody really knows why it is happening. There is no doubting the data; many species are indeed declining very quickly, in as clear an indication as one wants that something is wrong with the environment, but what? Several possibilities exist, and they are unlikely to be mutually exclusive. One explanation focuses on the destruction of tropical forests in the birds' wintering areas. Another suggestion is that there are other unknown problems there or on the migration routes. A third possibility is extensive habitat fragmentation and urbanization in the breeding forests of the eastern seaboard. This human modification of the northeast forests markedly increases nest losses of migrant songbirds to jays, crows, cowbirds, and racoons, all species that thrive in the slipstream of urban humans.

Declines in Formerly Common Farmland Birds in Northwest Europe

In the intensively agricultural areas of northwest Europe—over the whole of lowland England, for example—a whole raft of formerly "common farmland birds" are also in steep decline (Tucker and Heath, 1994). They include skylarks (*Alauda arvensis*), European tree sparrows (*Passer montanus*), corn buntings (*Milaria calaudra*), gray partridges (*Perdix perdix*), and song thrushes (*Turdus philomelos*). Here the problem is now reasonably well understood, though many details remain unresolved. Modern farming is so efficient and clean that there is little for the birds to eat. Weeds are killed with herbicides, which remove both seeds and rich sources of insects that feed on the weeds. The crop itself is sprayed to remove insects and is harvested so efficiently that few seeds are spilled on the way. Modern farms are biodiversity deserts, an indication of the power of people to squeeze nature to the margins while apparently maintaining a green and pleasant land. If present trends continue, skylarks will be rare birds in Britain in 20 years.

Species as Indicators of Biodiversity

The Nature of the Problem

Common sense suggests that the known losses of plants and birds from European farmland will go hand-in-hand with much more poorly documented declines in many other, less familiar and cryptic taxa, from land snails to glowworms, and hoverflies to harvest spiders. In other words, changes in the distribution and abundance of well-known groups should

serve as broad indicators of the status of, and changes in, a much wider sample of a region's flora and fauna. The assumption here is that birds (or other conspicuous species) might serve as biodiversity indicators—that is, as surrogates of overall biodiversity. But although it seems intuitively reasonable to use familiar, well-studied, and easily censused groups as indicators of what is happening to many other taxa, despite a great deal of research, the idea is actually contentious.

Following (but slightly modifying) the work of McGeoch (1998), we can define a biodiversity indicator as a group of taxa (e.g., genus, tribe, family, or order, or a selected group of species from a range of higher taxa) whose diversity (e.g., overall species richness, number of rare species, levels of endemism) reflects that of other higher taxa in a habitat, group of habitats, or geographic region. The idea is simple enough, and if it can be shown to work, it is important because biologists then have a relatively simple means of assessing overall biodiversity for purely scientific reasons, for setting conservation priorities, or for monitoring the effectiveness of conservation management.

Taxa that have been Suggested as Indicators of Biodiversity

The groups of organisms whose richness has been evaluated most thoroughly in the greatest number of places on earth are also the most familiar. The natural history sections of bookshops are dominated (sometimes exclusively) by volumes on plants and birds. If insects figure at all, butterflies will be on the top of the list, although there are fascinating differences between nations. Japan loves dragonflies. Birds, higher plants, butterflies, and dragonflies are all groups that occur in most places in the world but whose individual species are seldom so widespread. In much of the world they are also groups whose species are, relatively speaking, taxonomically well known and stable, readily identifiable, and have biologies that are well understood. They are easy to find, inventory, and count, and they are reasonably, but not overwhelmingly, diverse in any one place. These are all desirable attributes of groups that might be used as indicators of the diversity of many other, much less well known taxa—that is, as indicators of the overall biodiversity of a region.

Other groups have many of these same attributes but have not gained the same popularity, perhaps because often they are not also large bodied or perceived as being quite so attractive. The list of those that have been advocated as useful biodiversity indicators at one time or another is very long. It includes soil nematodes, moths, beetles galore (tiger, carabid, dung, and buprestid, to name but four), termites, fish, frogs, and snakes. Whatever the group, they must also have one further attribute—namely, that they genuinely indicate levels of biodiversity or at least some of the components of primary interest. The fact that the scientific literature contains suggestions for so many different possible indicators shows that there is little consensus on the matter. Many have been called, but few are chosen. Why? There are two, related, reasons. First, scientific knowledge on the degree of coincidence in patterns of biodiversity between different taxa is surprisingly poor. Second, as knowledge improves, coincidence between many taxa turns out to be much worse than people had imagined, or indeed hoped, would be the case.

Knowledge is Poor Because of the Effort Required

Gathering information on the diversity of different groups of organisms, even in one place, is enormously time-consuming. Two examples illustrate the problem. To map the presence and absence of breeding birds (conspicuous, "easy" to find and to identify) in every 10×10km grid-square in Britain and Ireland (there are 3672 squares) took more than half a million individual record cards, filled in by an army of amateur birdwatchers coordinated by professional ornithologists in the British Trust for Ornithology (BTO). The task took 4 years and about 100,000 hours of fieldwork (Gibbons *et al.*, 1993). Now imagine the effort required to do the same thing for all the other hundreds of different groups of organisms found in this one small corner of Europe. It has been done for a sample of taxa (we will return to what these data show in a moment), but many groups remain unmapped.

At a much smaller spatial scale in a tropical forest in Cameroon, a group of biologists attempted to measure the impacts of forest disturbance on just eight groups (birds, butterflies, flying beetles, canopy beetles, canopy ants, leaf-litter ants, termites, and soil nematodes). The birds and the butterflies took 50 and 150 scientist-hours, respectively, to survey. But the effort required climbed rapidly for smaller-bodied, more cryptic, less well-known groups—1600 hours for the beetles, 2000 for the termites, and 6000 for the nematodes (Lawton *et al.*, 1998). Despite the fact that this work as a whole took about five scientist-years, inventories for most groups that were surveyed were still only partial, and most taxa remained unexamined (fungi, higher plants, spiders, soil mites, collembola, earthworms, lizards, frogs, and mammals, to name some of the most conspicuous gaps).

Given this background, it is hardly surprising that biologists do not have a complete inventory of all the species that occur even in a single, moderately sized area (a field, small wood, or lake)—a so-called ATBI (All Taxa Biodiversity Inventory) (Oliver and Beattie, 1996). A moments thought will also show that to use one or two groups (for the sake of the argument, say birds and butterflies) as indicators of the richness of other taxa in fact requires *several* such areas to be investigated to properly test the hypothesis that high bird diversity (or any other single group) reflects a high diversity of many other groups.

Although progress has been made in this area over the past decade, considerable work remains to be done. Even in otherwise well-studied situations, many groups remain to be examined. Hence, at the present time, and effectively by default, some groups are being used as indicators of biodiversity, even though we cannot show categorically that the richness of one or more groups of organisms truly reflects the overall, or even a major portion of the overall, biodiversity of an area. As a result there is little consensus about what a "good" indicator group, or groups, might be, because there are too few hard data, from a range of habitats and geographic regions round

the world, on which to draw firm conclusions. But as data slowly emerge, they are not encouraging for those who wish to use simple, single-taxon indicators of biodiversity.

Indicator Reliability

Where knowledge exists, it suggests that single or small numbers of taxa will usually be poor indicators of the biodiversity of other groups.

Tropical versus Temperate and Other Major Diversity Gradients

It would be wrong to think that there is no coincidence between patterns of diversity in different groups of organisms. Of course there is. In the broadest terms, it is axiomatic that most major terrestrial and freshwater groups are more species rich in the tropics than in temperate regions, at low elevations than at high ones, in forests than in deserts, and on large land masses than tiny islands. Whether you are a botanist, a bird-watcher, or a bug hunter, to find the most species it is generally advisable to head to hot and humid mainland tropics with lots of trees. It is easy to assume that there must therefore be reasonably good correlations between major diversity gradients for different groups. There can be, but even at this scale often there are not. Penguin diversity peaks in Antarctica, not the tropics, and there are many other examples of similar "reverse diversity gradients" that buck the average trend. On the eastern side of North America, the diversity of breeding warblers *increases* from south to north—suggesting that this conspicuous taxon, which is easy to identify (at least the breeding males!) and to survey, is probably highly unsuitable as an indicator of patterns of biodiversity in most other taxa (in which diversity typically *decreases* from south to north).

In the case of breeding North American warblers, we can spot the problem because we have enough information about the organisms involved. But the whole point about indicator taxa for biodiversity is that typically we will not be armed with, and indeed should not need, information about "other" groups; knowledge of the indicator taxon should suffice and be reliable. The evidence suggests otherwise.

Hot Spots

Major gradients in diversity aside, at similarly large scales an indicator group might be used to identify local geographic hot spots in the species richness of one or more other groups (peaks in the landscape of species richness) or to determine relative levels of richness in those other groups (hot spots versus all spots) (Gaston, 1996b; Reid, 1998). At the continental scale, the procedure has frequently been found to fail on both counts (Gaston, 1996a), with mismatches between the occurrence of peaks in the richness of different groups being commonplace.

Across the United States and southern Canada, hot spots (local areas with unusually high diversity) overlap partially between some pairs of taxa (trees, tiger beetles, amphibians, reptiles, birds, and mammals), but the pattern is not a general one. Numbers of species in different large grid cells for two groups are often significantly positively correlated, for example, birds and tiger beetles or mammals and swallowtail butterflies. But these correlations are frequently weak, of rather limited predictive value, and in some cases explained by latitudinal gradients in diversity. In other words, although such correlations may sometimes enable a very general impression of the patterns in richness of one group to be obtained from the patterns in richness of another, their predictive powers are low.

These conclusions seem to hold at finer resolutions over more constrained areas. Thus, species-rich areas for different taxa in Britain (birds with butterflies, dragonflies, etc.) frequently do not coincide at a scale of 10×10 km squares (Pendergast *et al.*, 1993) (**Figure 5**). Hot spots in this study are not distributed randomly, overlapping more often than expected by chance, but still at a low level. Likewise, different taxa are species poor or species rich in different areas of the Transvaal region of South Africa. At even finer scales, within the Cameroon forest mentioned earlier, disturbance impacted on the diversity of eight taxa in very different ways. All declined drastically in completely cleared areas, but intermediate levels of forest disturbance had very different effects on the diversity of different groups. As a result, changes in the diversity of one taxon could not be used to predict changes in the diversity of any other (Lawton *et al.*, 1998). A summary of these and related studies showing similar results is provided by Gaston (1996a, 1996b) and by Pimm and Lawton (1998).

A Common Sense Explanation

This lack of, or relatively feeble correlation between, species rich-areas for different groups of organisms makes the search for simple, robust, single-taxon indicators of overall biodiversity look increasingly like a lost cause. With hindsight, perhaps this emerging result is obvious (Reid, 1998). Major geographic gradients in biodiversity aside, within particular geographic regions or at smaller habitat scales, the conditions favoring one group of species may be hostile to another. Mollusks like it cool and wet, butterflies like it warm and sunny, and high bird diversity is more likely in tall vegetation than short vegetation, irrespective of weather. Common sense natural history suggests that there is unlikely to be a single indicator taxon able to predict the diversity of all, or even a majority of others.

Rare Species and Endemic Species

Biologists and conservationists are often interested not only in patterns of species richness but also in the distribution of unusually rare species, or of endemic species. Do sites with unusual numbers of rare species frequently coincide across different taxa? Again, the answer seems to be no, or only weakly (Pimm and Lawton, 1998; Prendergast *et al.*, 1993), for the reasons just outlined.

Endemic species may be different (Bibby *et al.*, 1992). There is some evidence that areas rich in endemic birds (e.g., some tropical mountaintops or isolated islands) may also contain unusually large numbers of endemic species in other groups. However, rigorous data and analyses are few, and exceptions are easy to find. Lake Baikal has no endemic birds but supports an exceptionally rich, endemic invertebrate fauna and a unique, endemic freshwater seal.

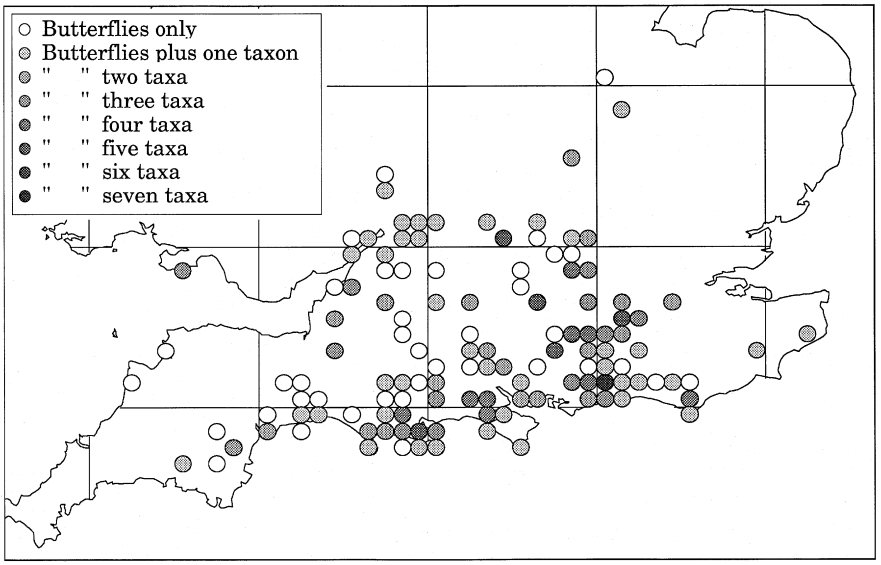


Figure 5 Coincidence between hot spots for butterflies and up to seven other taxa in Britain. Hot spots are unusually species-rich sites (here defined as the top 5 percentile in Britain). All of the most species-rich localities in Britain for butterflies lie in southern England. Increasingly dark shading indicates that butterfly hot spots coincide with hot spots for an increasing number of other taxa. Note that many butterfly hot spots are not unusually rich in any other species (open circles) and that only one locality (in southeast England, just in from the coast) is a hot spot for all eight taxa in this particular survey. The other taxa are breeding birds, dragonflies, moths, mollusks, aquatic higher plants, and liverworts (simple plants). Reproduced from Prendergast JR, Quinn RM, Lawton JH, Eversham BC, and Gibbons DW (1993) Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* 365: 335–337, with additional data and figure kindly provided by John Prendergast.

Selecting Areas to Conserve Biodiversity: Conservation Planning and Reserve Selection Algorithms

Biodiversity Indicators and Conservation

The general lack of reliable indicator groups for biodiversity is undoubtedly unfortunate for scientists wishing to understand how life is distributed across the earth; the road to an atlas of biodiversity seems set to be a long one. In practice, it may actually prove somewhat less of a worry for one of the primary motivations in the search for indicators for biodiversity—namely, conservation planning. Networks of national parks and reserves are central planks in conservation, albeit alone they are insufficient to protect all species. Their establishment is one of the obligations placed on Parties to the Convention on Biological Diversity. A primary argument for using indicators of biodiversity is to determine the effectiveness of these protected area networks in capturing biodiversity, and the best ways in which they might be extended, in the face of stiff competition with alternative forms of land use.

Although hot spots of species richness coincide weakly for different taxa (see the previous section), if we turn the problem around and look at it from a different angle, an interesting picture emerges. Imagine that conservation priorities in Britain have been set by concentrating just on areas rich in birds (the bird hot spots). (Although the general view of many nations is that the British are mad about birds, the country's protected area network is not based solely on birds. We use the example simply to illustrate a point.) What we discover is that the hot spots for this one group tend to embrace a high proportion of the total species in other groups. Thus, the hot spots for breeding birds contain 87% of the breeding bird species in Britain, 100% of the butterflies, 92% of the dragonflies, 92% of the liverworts, and 94% of the aquatic plants. A reserve network established around hot spots for one group does a rather good job of ensuring that most species in other groups find a place in the extended ark of protected sites.

Reserve Selection Algorithms

Although this message is encouraging, it turns out that designing conservation networks simply on the basis of levels of species richness is extremely inefficient, at least if the goal is to capture representative samples of all taxa. The same species may occur repeatedly in different richness hot spots for a group. Conservationists do not have unlimited resources, and this duplication wastes money on purchasing, or managing, unnecessary land. On the other hand, some species, particularly the rare ones of primary conservation interest, may not occur in any richness hot spots at all. What is required is to identify those areas that constitute the greatest complementary species richness; the complementary part of an area's biota consists of those species unrepresented in another biota with which it is being compared. To do this, mathematically minded conservation biologists have developed powerful reserve selection algorithms that help to select sites with maximum efficiency, according to some predetermined criteria (Pressey *et al.*, 1993). The criteria may be to maximize the number of species, rare species, or endemic species in a proposed reserve network, at minimum cost, on a minimum area, closest to existing reserves, or what have you.

Echoing the conclusions of the previous section, the question then arises as to whether the patterns of complementarity of one group of organisms are congruent with those of another. The question is a new one, with few studies available to answer it. Across 50 forests of Uganda, which boasts more species for its size than almost any other country

in Africa, and consistent with our earlier conclusions, there was little spatial congruence in the species richness of woody plants, large moths, butterflies, birds, and small mammals once differences in sampling effort were accounted for. However, sets of forests selected using complementarity determined for single taxa were generally similar to those for all other taxa and hence served to capture well the species richness in all these other groups (Howard *et al.*, 1998).

If these results generalize to other parts of the world, they send an encouraging message to conservation managers struggling to identify the best areas to set aside as reserves and parks. It says that a complementary and therefore efficiently selected chain of reserves based on a single indicator taxon (or perhaps two or three indicator taxa) may efficiently capture complementary sets of many other groups as well. Unfortunately, the Ugandan results are not supported by similar studies in the Transvaal, elsewhere in Africa (van Jaarsveld *et al.*, 1998). It may therefore be too soon to assume that we can find simple indicators for complementary reserve sets embracing many taxa as a means of conserving biodiversity.

Conclusions

The term "indicator species" has three distinct meanings. They are a species, or group of species, that reflect the biotic or abiotic state of an environment; reveal evidence for, or the impacts of, environmental change; or indicate the diversity of other species, taxa, or entire communities within an area. The uses of indicator species in the first two senses of the word are very similar, differing largely in the fact that to indicate change, organisms need to be sampled more than once in the same place and in the same way. Using organisms to indicate the state of, and changes in, the environment has numerous tried and tested applications, from detecting pollution to monitoring recovery of formerly degraded habitats, at many scales, from local to global. The use of indicator species to predict the diversity of other, unstudied taxa for scientific or conservation reasons is much more contentious and may prove to be impossible with any degree of rigor.

See also: Birds, Biodiversity of. Ecotoxicology. Endemism. Environmental Impact, Concept and Measurement of. Greenhouse Effect. Hotspots. Keystone Species. Paleoecology

References

Bibby CJ, Collar NJ, Crosby MJ, *et al.* (1992) *Putting Biodiversity on the Map: Priority Areas for Global Conservation.* Cambridge: International Council for Bird Preservation.

Blaustein AR and Wake DB (1995) The puzzle of declining amphibian populations. *Sci. Am.* 272: 56–61.

Erisman JW and Draaijers GPJ (1995) *Atmospheric Deposition in Relation to Acidification and Eutrophication.* Amsterdam: Elsevier.

Gaston KJ (1996a) Biodiversity—Congruence. *Prog. Phys. Geog.* 20: 105–112.

Gaston KJ (1996b) Spatial covariance in the species richness of higher taxa. In: Hochberg ME, Clobert J, and Barbault R (eds.) *Aspects of the Genesis and Maintenance of Biological Diversity*, pp. 221–242. Oxford: Oxford University Press.

Gibbons DW, Reid JB, and Chapman RA (eds.) (1993) *The New Atlas of Breeding Birds in Britain and Ireland: 1988–1991.* London: Poyser.

Howard PC, Viskanic P, Davenport TRB, *et al.* (1998) Complementarity and the use of indicator groups for reserve selection in Uganda. *Nature* 394: 472–475.

Huntley B and Birks HJB (1983) *An Atlas of Past and Present Pollen Maps for Europe 0–13000 Years Ago.* Cambridge: Cambridge University Press.

van Jaarsveld AS, Fretag S, Chown SL, *et al.* (1998) Biodiversity assessment and conservation strategies. *Science* 279: 2106–2108.

Lawton JH, Bignell DE, Bolton B, *et al.* (1998) Biodiversity inventories, indicator taxa and the effects of habitat modification in tropical forest. *Nature* 391: 72–76.

McGeoch MA (1998) The selection, testing and application of terrestrial insects as bioindicators. *Biol. Rev.* 73: 181–201.

Oliver I and Beattie AJ (1996) Designing a cost-effective invertebrate survey: A test of methods for rapid assessment of biodiversity. *Ecol. Appl.* 6: 594–607.

Pimm SL and Lawton JH (1998) Planning for biodiversity. *Science* 279: 2068–2069.

Prendergast JR, Quinn RM, Lawton JH, Eversham BC, and Gibbons DW (1993) Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* 365: 335–337.

Pressey RL, Humphries CJ, Margules CR, Vane-Wright RI, and Williams PH (1993) Beyond opportunism: Key principles for systematic reserve selection. *TREE* 8: 124–128.

Ratcliffe DA (1980) The Peregrine Falcon. Calton: Poyser.

Reid WV (1998) Biodiversity hotspots. *TREE* 13: 275–280.

Rosenberg DM and Resh VH (eds.) (1993) *Freshwater Biomonitoring and Benthic Macroinvertebrates*. New York: Chapman & Hall.

Shaw IC and Chadwick J (1998) *Principles of Environmental Toxicology.* London: Taylor & Francis.

Tucker GM and Heath MF (1994) *Birds in Europe: Their Conservation Status.* Cambridge: Bird Life International.

Indicator Species: Computation

Pierre Legendre, Université de Montréal, Montréal, QC, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Fidelity The degree to which a species is present at all sites of a group.

Indicator species A biological species that has a high indicator value for a group of sites. In the Indictor value method, species j is an indicator of group of sites k if the indicator value of species j is the highest, among all groups, for that group of sites, and is statistically significant at a preselected significance level.

Indicator value The degree to which a species is indicator of (the conditions found in) a group of sites.

Pseudospecies To model the concept of differential species (i.e., species with clear ecological preferences), which is qualitative, TWINSPAN creates pseudospecies. Each species is recorded into a set of dummy variables, called pseudospecies, corresponding to relative abundance levels (or percentage cover for plants) at each site; these classes are cumulative. If the pseudospecies cutting levels are 1%, 11%, 26%, 51%, and 76%, for instance, a relative abundance of 18% at a site will occupy the first and second dummy pseudospecies vectors with "1" (=presence). Cutting levels are arbitrarily decided by users. A (sites × pseudospecies) data table is created.

Specificity The degree to which a species is found only in a given group of sites.

Ecological Interest for Indicator Species

How to identify characteristic or indicator species is a traditional problem in community ecology and biogeography. It is grounded in the paradigm that species are the best indicators available for particular environmental conditions. Field studies describing habitats or groups of sites often mention one or several species that characterize each habitat. Because indicator species add ecological meaning to site groups discovered by clustering, they provide criteria to compare typologies derived from data analysis, to identify where to stop dividing clusters into subsets, and to point out the main levels in a hierarchical classification of sites. Indicator species differ from species associations in that they are indicative of particular, predetermined groups of sites. Good indicator species should be found mostly in a single group or a limited number of groups of a typology, and be present at most of the sites belonging to that or these groups. This duality is of ecological interest; yet it is not always used in indicator species studies.

There is clearly a need for the identification of characteristic or indicator species in the fields of monitoring, conservation, and management. For example,

- Species may be used as indicators of the state of an ecosystem and of human-induced changes to the environment and biodiversity.
- In long-term environmental follow-up studies for conservation or ecological management, researchers are looking for biological indicators of habitat types or combinations of habitat types they want to preserve or rehabilitate (McGeoch and Chown, 1998).

As for many other ecological concepts, the empirical method of early ecologists who walked through ecosystems and examined data tables to identify indicator species has gradually been replaced by statistical methods that attempt to reproduce and formalize the reasoning of the traditional ecologists, with computer programs facilitating the calculations. In this article, the author will focus on statistical methods for the identification of indicator species, comparing the merits of the Two-way indicator species analysis (TWINSPAN) and Indicator value (IndVal) methods.

Statistical Methods: TWINSPAN and IndVal

TWINSPAN: A Brief View

Two-way indicator species analysis (TWINSPAN) (Hill, 1979) is fundamentally a method for hierarchical divisive classification of communities, based on progressive refinement of a single ordination axis obtained by correspondence analysis (CA) or detrended correspondence analysis (DCA) of a community composition (sites × species) data matrix. The algorithm, which is rather complex, is detailed by Kent and Coker (1992). An attractive feature of the computer output is a two-way table with the sites (columns) sorted according to the splits of the hierarchical classification. The species (rows) are also sorted so as to form blocks corresponding to the groups of sites of the classification. The body of the table contains the highest pseudospecies scores (*see* Glossary) found at the sites.

An additional feature of TWINSPAN is that it computes an indicator values index (I) for the species for every split of the hierarchical classification of the sites. According to Kent and Coker (1992), the index is computed as follows using the pseudospecies data (*see* Glossary):

$$I_j = \frac{n_j^+}{n^+} - \frac{n_j^-}{n^-}$$

where n^+ and n^- are, respectively, the number of sites on the (arbitrarily chosen) positive and negative sides of the split, whereas n_j^+ and n_j^- are the number of sites on the positive and negative sides, respectively, that contain pseudospecies j.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00430-5

A pseudospecies present in every site on the positive side and in none of the sites on the negative side obtains $I_j = 1$, and -1 if it is found in every site on the negative side and in none on the positive side. A pseudospecies that occurs in all sites on both sides of the split obtains $I_j = 0$. In TWINSPAN, only one pseudospecies of a single species is declared an indicator of a split, and that is the pseudospecies that has the highest absolute value of I. n_j^+/n^+ is the measure of fidelity to a group used in the INDVAL method (*see* The INDVAL Method).

The major disadvantage of TWINSPAN is that it can only compute the indicator value of species for the hierarchical classification produced by itself. Another critique is that the importance of a species in the analysis depends on the abundances of the other species because the pseudospecies, which form the data basis of the method, are based on species relative abundances. The method has also been criticized by Belbin and McDonald (1993) because it assumes the existence of a single, strong gradient dominating the data structure, and because the cutting points for the whole group, and then for subgroups, are always chosen to be the centroid of the group to be split instead of a point where a large gap occurs in the data. Because of that, sites that are very similar in species composition may end up in separate groups. The last problem has been alleviated by a modification to the method proposed by Roleček *et al.* (2009).

The INDVAL Method

Dufrêne and Legendre (1997) presented an alternative to TWINSPAN in the search for indicator species and species assemblages characterizing groups of sites. Like TWINSPAN, the indicator value (IndVal) method analyses the species with reference to a prior partition of the sites. The first novelty of IndVal is that it derives indicator species from any hierarchical or nonhierarchical classification of the objects (sampling sites), contrary to TWINSPAN where indicator species can only be derived for a classification obtained by splitting sites along a CA axis. The second novelty lies in the way the indicator value of a species is measured for a group of sites. The indicator value index (INDVAL) is based only on within-species abundance and occurrence comparisons; its value is not affected by the abundances of other species. The significance of the indicator value of each species is assessed by a randomization procedure.

The *IndVal* index is defined as follows. For each species j in each cluster of sites k, one computes the product of two values, A_{kj} and B_{kj}. A_{kj} is a measure of specificity based on abundance values whereas B_{kj} is a measure of fidelity computed from presence data:

$$A_{kj} = N\text{individuals}_{kj}/N\text{individuals}_{+k}$$

$$B_{kj} = N\text{sites}_{kj}/N\text{sites}_{k+}$$

$$INDVAL_{kj} = A_{kj}B_{kj}$$

n the formula for specificity (A_{kj}), N individuals$_{kj}$ is the mean abundance of species j across the sites pertaining to cluster k and N individuals$_{+k}$ is the sum of the mean abundances of species j within the various clusters. The mean number of individuals in each cluster is used, instead of summing the individuals across all sites of a cluster, because this removes any effect of variations in the number of sites belonging to the various clusters. Differences in abundance among sites of a cluster are not taken into account. A_{kj} is maximum when species j is present in cluster k only. In the formula for fidelity (B_{kj}), N sites$_{kj}$ is the number of sites in cluster k where species j is present and N sites$_{k+}$ is the total number of sites in that cluster, as in index I of TWINSPAN. B_{kj} is maximum when species j is present at all sites of cluster k. Quantities A and B must be combined by multiplication because they represent independent information (i.e., specificity and fidelity) about the distribution of species j.

The indicator value of species j for a partition of sites is the largest value of $IndVal_{kj}$ observed over all clusters k of that partition:

$$IndVal_j = \max[IndVal_{kj}]$$

The index is maximum (its value is 1) when the individuals of species j are observed at all sites belonging to a single cluster. A random permutation procedure of the sites among the site groups is used to test the significance of $IndVal_j$. A correction for multiple testing is necessary before reporting the results when multiple tests (for several species) have been conducted. The index can be computed for a given partition of the sites, or for all levels of a hierarchical classification of the sites.

Numerical Example

Table 1 describes the example given by Dufrêne and Legendre (1997), slightly modified, to illustrate the computation of the *IndVal* index. The data represent three species observed at 25 sites, which are divided into five groups. To facilitate comparisons, the sums of the mean group abundances are 20 for all three species. For species 1, $IndVal_{k1}$ has the highest value (0.30) for group 2, so $IndVal_1 = 0.30$. Following similar reasoning, $IndVal_2 = 0.40$ and $IndVal_3 = 0.90$. The permutational p-values computed by functions **IndVal()** of LABDSV and ***multipatt*()** of INDICSPECIES in R are significant in all three cases.

Statistical Method: *INDVAL* Expanded

De Cáceres and Legendre (2009) described other statistics that can be used to assess the indicator value of species. The statistics are divided into (1) correlation indices, which are used for determining the ecological preferences of species among a set of alternative site groups or site group combinations, and (2) indicator value indices, including *IndVal*, which are used for assessing the predictive values of species as indicators of the conditions found in site groups, for example for field determination of community types or ecological monitoring. Each of these families of indices comes in different types: there are indices for presence–absence and for quantitative species data; there are also nonequalized indices that give equal weights to individual sites and group-equalized indices that give equal weights to all groups whatever the number of sites they contain. For studies involving several groups of sites, De Cáceres *et al.* (2010) showed that the interpretation of indicator value analysis could be improved by computing the statistics for all possible

Table 1 Numerical example: abundance of three species at 25 sites divided into five groups

Groups	*Group 1*					*Group 2*					*Group 3*					*Group 4*					*Group 5*				
Sites	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Species 1	4	4	4	4	4	6	6	6	6	6	5	5	5	5	5	3	3	3	3	3	2	2	2	2	2
Species 2	8	8	8	8	8	4	4	4	4	4	6	6	6	6	6	4	4	2	0	0	0	0	0	0	0
Species 3	18	18	18	18	18	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Species 1																									
A_{k1}	4/20 = 0.20					6/20 = 0.30					5/20 = 0.25					3/20 = 0.15					2/20 = 0.10				
B_{k1}	5/5 = 1					5/5 = 1					5/5 = 1					5/5 = 1					5/5 = 1				
$IndVal_{k1}$	0.20					**0.30**					0.25					0.15					0.10				
Species 2																									
A_{k2}	8/20 = 0.40					4/20 = 0.20					6/20 = 0.30					2/20 = 0.10					0/20 = 0.00				
B_{k2}	5/5 = 1					5/5 = 1					5/5 = 1					3/5 = 0.6					0/5 = 0				
$IndVal_{k2}$	**0.40**					0.20					0.30					0.06					0.00				
Species 3																									
A_{k3}	18/20 = 0.90					2/20 = 0.10					0/20 = 0.00					0/20 = 0.00					0/20 = 0.00				
B_{k3}	5/5 = 1					5/5 = 1					0/5 = 0					0/5 = 0					0/5 = 0				
$IndVal_{k3}$	**0.90**					0.10					0.00					0.00					0.00				

Top panel: species abundance data. Bottom: calculation of the specificity (A_{kj}), fidelity (B_{kj}) and $IndVal_{kj}$ index for each species (j) in each group of sites (k). The maximum value of $IndVal_{kj}$ for each species is in bold.

Source: Modified from Dufrêne M and Legendre P (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs* 67, 345–366.

combinations of site groups. Moretti *et al.* (2010) published an application of that method.

A detailed discussion of the limitations of indicator value analysis is found in De Cáceres *et al.* (2010). They point out that more indicator species will be found than expected by chance when the classification of sites was obtained from the same species composition data that are used for IndVal analysis. In that case, p-values must be interpreted with caution: they do not result from a genuine test of significance where the classification of sites has to be independent of the species data used in the test.

Variants of the *IndVal* index have been proposed by Podani and Csányi (2010). Instead of using specificity and fidelity alone, these authors proposed to define the indicator value of a species as the product of two among three quantities: specificity A_{kj} (that they called concentration), specificity (new equation, with allowance for positive or negative species preferences), and fidelity B_{kj}. They provided formulas based on presence–absence or abundance data for each of these three quantities.

According to McGeoch and Chown (1998), IndVal is important for conservation biology because it is conceptually straightforward and allows researchers to identify biological indicators for any combination of habitat types or areas of interest – existing conservation areas, or groups of sites based on the outcome of a classification procedure. The method is sometimes used to identify biological indicators for groups of sites classified using the target taxa, subject to the caveat about p-values mentioned two paragraphs up, or in other instances using nontarget taxa, for example finding insect biological indicators for a classification of sites based on plant community data.

An IndVal index for a species is calculated independently of the other species in the assemblage. Because of that, comparisons of indicator values can be made between taxonomically unrelated taxa, or taxa in different functional groups or in different communities. Comparisons across taxa are robust to differences in abundance that may or may not be due to differences in capturability or sampling methods. The method is also robust to differences in the numbers of sites between site groups, in abundance among sites within a particular group, and in the absolute abundances of very different taxa that may show similar trends.

When a group of sites for which indicator species are sought corresponds to a well-delimited geographic area, superposition of distribution maps for the indicator species of that group should help identify the core conservation areas for these species, even when little other biological information is available.

Taxa described as biological indicators in the literature are often merely the favorite taxa of their proponents; ornithologists prefer birds, lepidopterists butterflies, and coleopterists beetles. IndVal provides an objective method for addressing this problem by enabling assessment of the relative merits of different taxa for a given study area (McGeoch and Chown, 1998). The species that emerge from this procedure as the best indicators for a group of sites should prove useful for monitoring site alterations, natural or man-induced.

Ecological Applications

Carabid Beetles in Belgium

Dufrêne and Legendre (1997) analyzed a large data set of Carabid beetle distributions in open habitats of Belgium

(189 species collected in pitfall traps at 69 sites). Classification of the sites was obtained by distance-based *K*-means partitioning computed as follows: first, a distance matrix (percentage difference, also called Bray-Curtis distance) was computed from the log-transformed species abundance data; this distance matrix was subjected to principal coordinate analysis (PCoA), also called metric multidimensional scaling; all ordination axes (i.e., the principal coordinates) produced by PCoA were used as input data into *K*-means partitioning. Although the clusters produced by *K*-means were not forced to be hierarchically nested, they showed a strong hierarchical structure for $K =$ two to ten groups. This allowed the authors to represent the relationships among partitions as a dendrogram with $K = 10$ groups, which corresponded to the main types of habitat recognized *a priori* by surveying the sites.

Indicator values were computed for each species and partitioning level. Some species were found to be stenotopic (narrow niches) whereas others were eurytopic (species with wide niches, present in a variety of habitats). Other species characterized intermediate levels of the hierarchy. The best indicator species ($IndVal > 0.25$) were assembled into a two-way indicator table; this tabular representation displayed hierarchical relationships among the species.

The IndVal results were compared to TWINSPAN. The partitions of sites used in the two methods were not the same; the TWINSPAN typology was obtained by partitioning CA ordination axes. TWINSPAN identified, as indicators, pseudospecies pertaining to very low cut-off levels. These species were not particularly useful for prediction because they were simply known to be present at all sites of a group. Several species identified by TWINSPAN as indicators also received a high indicator value from the IndVal procedure for the same or a closely related habitat class. IndVal identified several other indicator species, with rather high indicator values, that also contributed to the specificity of the groups of sites but had been missed by TWINSPAN. So, the IndVal method appeared to be more sensitive than TWINSPAN to the fidelity and specificity of species.

More Application Examples

Here are more examples of the many applications of indicator species analysis found in the literature.

- Borcard (1996) and Borcard and Vaucher-von Ballmoos (1997) present applications of the indicator value method to the identification of the Oribatid mite species that characterize well-defined zones in a peat bog of the Swiss Jura.
- The indicator values of beetle species characterizing different types of forests have been studied by Barbalat and Borcard (1997).
- Tuomisto *et al.* (2003) used spatially-constrained clustering to group 86 sampling units, each 500 m in length, forming a 43-km long transect in the Amazonian rain forest in Peru, into spatial clusters on the basis of satellite image pixel values. They also surveyed the ferns and Melastomaceae in the 86 sampling units in the field. Then they used the IndVal method to determine the species of ferns and Melastomaceae that were good indicators of the spatial clusters.
- Legendre *et al.* (2009) used multivariate regression tree analysis (De'ath, 2002) to identify habitat types that were similar in topographic conditions and in species composition in a Chinese permanent forest plot divided in 20 m × 20 m quadrats; then they used the IndVal method to identify, among the 159 tree species, the nine species that were statistically significant indicators of the five main habitat types. In this paper, the species used for IndVal analysis had been used to obtain the classification of the sites; as a consequence, the p-values had to be interpreted with caution.
- De Cáceres *et al.* (2010) carried out indicator species analysis of the vegetation of the Barro Colorado Island (BCI) permanent forest plot in Panama, also divided in 20 m × 20 m quadrats, grouped into seven habitat types identified in the literature. Among 307 tree species, they identified 44 indicator species of individual habitat types and 64 species for habitat type combinations. In this paper, the classification of the sites was independent of the species analyzed for indicator value.

Conclusion

Indicator species are biological indicators of groups of sites representing habitat types or combinations of habitat types; they are of prime interest for ecosystem conservation and management. Statistical methods are now available to identify indicator species in different situations (presence–absence or abundance data, group-equalized or nonequalized indices) and for a variety of purposes: correlation indices are used for determining the ecological preferences of species among a set of alternative site groups or site group combinations, whereas indicator value indices are used for assessing the predictive values of species as indicators of the conditions found at groups of sites. Both types of indices are readily available for computation in R functions. Tests of significance can be computed for these indices, producing p-values that help identify the most interesting indicator species.

Software

In the R statistical language, indicator value indices (IndVal) can be computed by functions ***strassoc()*** and ***multipatt()*** of INDICSPECIES and by function ***indval()*** of LABDSV. The functions in INDICSPECIES offer a choice of the several different indicator statistics described in De Cáceres and Legendre (2009). The IndVal index is also available in the computer package PC-ORD. The TWINSPAN program, in FORTRAN, can be obtained from several Web pages.

See also: Indicator Species

References

Barbalat S and Borcard D (1997) Distribution of four beetle families (Coleoptera: Buprestidae, Cerambycidae, phytophagous Scarabaeidae and Lucanidae) in different forest ecotones in the Areuse Gorges (Neuchâtel, Switzerland). *Écologie* 28: 199–208.

Belbin L and McDonald C (1993) Comparing three classification strategies for use in ecology. *Journal of Vegetation Science* 4: 341–348.

Borcard D (1996) Typologie des assemblages d'espèces d'Oribates (Acari, Oribatei) de la tourbière du Cachot (Jura suisse): espèces indicatrices ou groupements caractéristiques? *Bulletin de la Societe neuchâteloise des Sciences naturelles* 119: 63–73.

Borcard D and Vaucher-von Ballmoos C (1997) Oribatid mites (Acari, Oribatida) of a primary peat bog-pasture transition in the Swiss Jura mountains. *Écoscience* 4: 470–479.

De'ath G (2002) Multivariate regression trees: A new technique for modeling species–environment relationships. *Ecology* 83: 1105–1117.

De Cáceres M and Legendre P (2009) Associations between species and groups of sites: Indices and statistical inference. *Ecology* 90: 3566–3574.

De Cáceres M, Legendre P, and Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119: 1674–1684.

Dufrêne M and Legendre P (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs* 67: 345–366.

Hill MO (1979) *TWINSPAN – A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of the Individuals and Attributes*. Ithaca, New York: Section of Ecology and Systematics, Cornell University.

Kent M and Coker P (1992) *Vegetation Description and Analysis – A Practical Approach*. New York: John Wiley & Sons.

Legendre P, Mi X, Ren H, *et al.* (2009) Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology* 90: 663–674.

McGeoch MA and Chown SL (1998) Scaling up the value of bioindicators. *Trends in Ecology and Evolution* 13: 46–47.

Moretti M, De Cáceres M, Pradella C, *et al.* (2010) Fire-induced taxonomic and functional changes in saproxylic beetle communities in fire sensitive regions. *Ecography* 33: 760–771.

Podani J and Csányi B (2010) Detecting indicator species: some extensions of the IndVal measure. *Ecological Indicators* 10: 1119–1124.

Roleček J, Tichy L, Zeleny D, and Chytry M (2009) Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. *Journal of Vegetation Science* 20: 596–602.

Tuomisto H, Ruokolainen K, Aguilar M, and Sarmiento A (2003) Floristic patterns along a 43-km long transect in an Amazonian rain forest. *Journal of Ecology* 91: 743–756.

Indigenous Peoples and Biodiversity

Víctor M Toledo, National University of México (UNAM), Morelia, México

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bio–cultural axiom Recognition that biological and cultural diversity are mutually dependent and geographically coterminous.

Cultural diversity Variety of human groups distinguished through beliefs, lifeways, dress, food, languages, sexual behavior, forms of productive organization, art, and conceptions of nature.

Endemic languages Languages that are restricted to a single country and, like their species counterparts, hold a high percentage of the unique traits in human language.

Ethnoecology Interdisciplinary study that explores how nature is perceived by human groups through a screen of beliefs and knowledge and how humans, in terms of their images and symbols, use or manage natural resources. The above is synthesized by the complex formed by *kosmos*, *corpus*, and *praxis*.

Indigenous peoples Those who are the original or oldest inhabitants of an area or region, or who have lived in a traditional homeland for many generations, usually many centuries.

Outline

Indigenous people, who number more than 300 million, are inhabitants of practically every main biome of the earth and especially of the least disturbed terrestrial and aquatic ecosystems of the world. Based on an exhaustive review of published data, this article stresses the strategic importance of indigenous peoples in the maintenance and conservation of the world's biodiversity.

Introduction

As a word and concept, biodiversity originated in the field of conservation biology. However, as Alcorn (1994) states "…while proof of conservation success is ultimately biological, conservation itself is a social and political process, not a biological process. An assessment of conservation requires therefore an assessment of social and political institutions that contribute to, or threaten, conservation." One of the main social aspects related to biodiversity is, undoubtedly, the case of the world's indigenous peoples.

Scientific evidence shows that virtually every part of the planet has been inhabited, modified, and manipulated throughout human history. Although they appear untouched, many of the last tracts of wilderness are inhabited and have been so for millenia. Indigenous peoples live in and have special claims to territories that, in many cases, harbor exceptionally high levels of biodiversity. On a global basis, human cultural diversity is associated with the remaining concentrations of biodiversity. Both cultural diversity and biological diversity are endangered.

Given the above, this article offers a review about the multiple importance of indigenous peoples and makes the point that valuable, local-specific views, knowledge, and practices are used by indigenous peoples who have relied for hundreds or thousands of years on the maintenance of biodiversity. Indigenous peoples are often classified as impoverished or treated as invisible. However, in the final analysis, they hold the key to successful biodiversity conservation in most of the biologically richest areas of the world.

Indigenous Peoples

Indigenous people number more than 300 million (**Table 1**). They live in about 75 of the world's 184 countries and are inhabitants of practically every main biome of the earth. Indigenous peoples, also called tribal, aboriginal, or autochthonous peoples, national minorities, or first peoples, are best defined by using several criteria. Indigenous peoples may have all or part of the following criteria: (1) are the descendants of the original inhabitants of a territory which has been overcome by conquest; (2) are "ecosystem peoples," such as shifting or permanent cultivators, herders, hunters and gatherers, fishers, and handicraft makers, who adopt a multiuse strategy of appropriation of nature; (3) practice a small-scale, labor-intensive form of rural production which produce little surplus and has low energy needs; (4) do not have centralized political institutions, organize their life at the level of

Table 1 Estimated numbers of the world's indigenous peoples

Region groups	*Number of cultural group*	*Population*
North America	250	3,500,000
Latin America and the Caribbean	800	43,000,000
Former Soviet Union	135	40,000,000
China and Japan	100	67,000,000
The Pacific	1273	2,000,000
Southeast Asia	900	30,000,000
South Asia	700	100,000,000
Australia and New Zealand	250	550,000
Africa	2010	50,000,000
Total	6418	336,050,000

Source: Reproduced from Burger J (1987) *Report from the Frontier: The State of the World's Indigenous Peoples*. London: Zed Books LTD, and Hitchcok (1994).

community, and make decisions on a consensus basis; (5) share a common language, religion, moral values, beliefs, clothing, and other identifying characteristics as well as a relationship to a particular territory; (6) have a different worldview, consisting of a custodial and nonmaterialist attitude to land and natural resources based on a symbolic interchange with the natural universe; (7) are subjugated by a dominant culture and society; and (8) consist of individuals who subjectively consider themselves to be indigenous.

It is possible to find indigenous peoples carrying out many different activities of use and management of the planet's ecosystems: As forest-dwellers in the tropical lowlands or in the mountains, pastoralists in savannas and other grasslands, or nomadic or semi-nomadic hunters and gatherers in forests, prairies and deserts. Fishing is, in addition, the principal economic activity and source of food for several million coastal and island dwellers, as well as many indigenous peoples inhabiting margins of rivers. Large numbers of indigenous peoples are, however, peasant producers and therefore can be indistinguishable from the nonindigenous peoples living nearby. In the Andean and Mesoamerican countries of Latin America, for instance, indigenous peoples farm like mestizo peasants. Similarly, in India distinctions between scheduled tribes and nontribal peoples cannot be made solely on the basis of productive activities. In these and many other cases, nonindigenous peasants and indigenous peoples produce the same crops with the same farming methods. Since in numberless countries many mestizo peasants are direct descendants of the indigenous peoples and retain most of their cultural traits, it has been pointed out that a broader definition of indigenous peoples might increase the real numbers. Thus, by considering other characteristics than language, it is possible to enlarge the number of indigenous peoples in the contemporary world. Given the above, some authors like Burger think that the number of indigenous people could double that previously estimated. Thus, in the contemporary world there may be as many as 600 million indigenous peoples. However, such figures need qualification.

Based on the percentage of total population identified as belonging to the indigenous peoples, it is possible to recognize a group of selected nations with a strong presence of these peoples: Papua New Guinea (77%), Bolivia (70%), Guatemala (47%), Peru (40%), Ecuador (38%), Mynamar (33%), Laos (30%), Mexico (12%), and New Zealand (12%). However, the absolute number of people recognized as indigenous allows identification of nations with high indigenous population such as India (more than 100 million) and China (between 60 and 80 million).

Biological Diversity and Diversity of Cultures

On a global basis, human cultural diversity is associated with the remaining concentrations of biodiversity. In fact, evidence exists of remarkable overlaps between global mappings of the world's areas of high biological richness and areas of high diversity of languages, the single best indicator of a distinct culture. The above correlation can be certified both on a country-by-country basis as well as using biogeographic criteria.

Measured by spoken language, the entire world's people belong to between 6000 and 7000 cultures. It is reported that 6700 of these are indigenous cultures. Thus, indigenous peoples account for as much as 80–90% of the world's cultural diversity. On the basis of the inventories done by linguists, we can draw up a list of the regions and countries with the greatest degree of cultural diversity in the world. According to Ethnologue, the best existing catalog of the world's languages, there is a total of 6703 languages (mostly oral), 32% of which are found in Asia, 30% in Africa, 19% in the Pacific, 15% in the Americas, and 3% in Europe (Gordon, 2005). Only 12 countries account for 54% of the human languages. These countries are Papua New Guinea, Indonesia, Nigeria, India, Australia, Mexico, Cameroon Brazil, Zaire, Philippines, USA, and Vanuatu (**Table 2**).

However, according to an analysis of biodiversity on a country-by-country basis (Mittermeier and Goettsch-Mittermeier, 1997) there are, similarly, 12 countries which house the highest numbers of species and endemic species (**Table 3**). This assessment was based on the comparative analysis of eight main biological groups: mammals, birds, reptiles, amphibians, freshwater fishes, butterflies, tiger-beetles, and flowering plants. The nations considered as "megadiversity" countries are: Brazil, Indonesia, Colombia, Australia, Mexico, Madagascar, Peru, China, Philippines, India, Ecuador, and Venezuela (**Table 3**).

Thus, the relationship between cultural diversity and biological diversity stands out in global statistics: nine of the 12 main centers of cultural diversity (in terms of number of languages) are also in the roster of biological megadiversity nations and, reciprocally, nine of the countries with the highest species richness and endemism are also in the list of

Table 2 Top 25 countries by number of endemic languages

1. Papua New Guinea (847)[a]
2. Indonesia (655)[a]
3. Nigeria (376)[a]
4. India (309)[a]
5. Australia (261)[a]
6. México (230)[a]
7. Cameroon (201)[a]
8. Brazil (185)[a]
9. Zaire (158)[a]
10. Philippines (153)[a]
11. USA (143)[a]
12. Vanuatu (105)
13. Tanzania (101)
14. Sudan (97)
15. Malaysia (92)[a]
16. Ethiopia (90)
17. China (77)[a]
18. Peru (75)[a]
19. Chad (74)
20. Russia (71)
21. Salomon Islands (69)
22. Nepal (68)
23. Colombia (55)[a]
24. Cote D'Ivoire (51)
25. Canada (47)

[a]Considered "megadiversity" countries by Mittermeier and Goettsch-Mittermeier, 1997.

the 25 nations with the highest number of endemic languages (Tables 2 and 3).

The links between biological and cultural diversity can also be illustrated by using the data of the so-called Global 200, a program that the World Wide Fund for Nature (WWF) developed as a new strategy to identify conservation priorities based on an eco-regional approach. As part of this program, WWF has identified a list of 233 terrestrial, freshwater, and marine biological ecoregions representative of the Earth's richest diversity of species and habitats. A preliminary analysis conducted by the People and Conservation Unit of WWF of the presence of indigenous peoples in the 136 terrestrial ecoregions of Global 200, revealed interesting patterns (Oviedo *et al.*, 2000). As shown in **Table 4**, nearly 80% of the terrestrial ecoregions are inhabited by one or more indigenous peoples and half of the global estimated 3000 indigenous groups are inhabitants of these ecoregions. By geographical region, basically all the regions, except the Palearctic region, have 80% or more of their ecoregions inhabited by indigenous people (**Table 4**).

Table 3 Top 12 countries by number of species richness and endemic (endemism)[a]

	Biological diversity		
	Richness	*Endemism*	*Both*
*Brazil	1	2	1
*Indonesia	3	1	2
*Colombia	2	5	3
*Australia	7	3	4
*México	5	7	5
*Madagascar	12	4	6
*Peru	4	9	7
*China	6	11	8
*Philippines	14	6	9
*India	9	8	10
Ecuador	8	14	11
Venezuela	10	15	12

[a]Calculated for the following biological groups: mammals, birds, reptiles, amphibians, freshwater fishes, butterflies, tiger-beetles and flowering plants. Asterisks indicate countries included in the list of the 25 nations with highest number of endemic languages (See table 3; Harmon, 1996).

Biodiversity and Biomass Appropriation: The Role of Indigenous Peoples

Biodiversity conservation cannot be separated from natural resources use. Human appropriation of nature inflows minerals, water, solar energy, and principally living beings (biomass) from ecosystems. World statistics indicate that almost half the inhabitants of the planet are still people engaged in the appropriation of natural resources. This appropriation is carried out by a myriad of rural or primary producers through the management of terrestrial, marine, and freshwater ecosystems.

Forty percent of the total human population has been recorded by the Food and Agricultural Organization (FAO) as agricultural population (FAO, 2004). It can be estimated that between 60 and 80% of this agricultural population is represented by small-scale, solar-energized productive units based on a multiuse management of nature (Toledo and Barrera-Bassols, 2008). In fact, the statistical record shows that by 1999 around 1500 million rural people were practicing agricultural activities on areas of 5 ha or less. A similar pattern is found in the world's fisheries where more than 90% are small-scale, artesanal operators, acting in a great variety of coastal habitats.

Most of these small-scale farmers and fishers develop their production activities not as socially isolated households but as familial nuclei belonging to specific village communities, many of which, in turn, correspond to cultures that can be considered as indigenous. Moreover, within the core of these community-based producers, those identified as indigenous people are those who carry out the biomass extraction affecting at the lowest level of their local ecosystems. Called "ecosystem people" by some authors such as Dassman and Gadgil, these producers subsist by appropriating a diversity of biological resources from their immediate vicinity. Their quality of life is therefore intimately linked to the maintenance of certain levels of local biodiversity. As a consequence, they are productive actors in little transformed habitats of the world, including the main forest and sea dwellers, slash and burn agriculturalists, some 25–30 million nomadic herders or pastoralists (in East Africa, Sahel, and Arabian peninsula), most of the 15–21 million world fishers, and all the half a million hunters and gatherers who are still recognized as inhabitants of the contemporary world.

Table 4 Indigenous peoples in Global 200 terrestrial ecoregions considered priority areas by World Wildlife Fund for Nature

Region	*Ecoregions*	*Ecoregion with IP*	*%*	*Total IP in ecoregions*	*Number of IP in ecoregions*	*%*
World	136	108	79	3000	1445	48
Africa	32	25	78	983	414	42
Neotropic	31	25	81	470	230	51
Neartic	10	9	90	147	127	86
Asia and Pacific (Indo-Malayan)	24	21	88	298	225	76
Oceania	3	3	100	23	3	13
Paleartic	21	13	62	374	11	30
Australasia	15	12	80	315	335	65

Note: These figures, utilized by WWF International, underestimate the number of indigenous peoples given by several analysts. For example, Harmon (1996) reported a total of 5635 "endemic languages," which can be considered as socially equivalent to indigenous peoples. IP = Indigenous peoples.

In conclusion, indigenous peoples are the fraction of human appropriators of biomass determining the lowest ecological impacts. They generally live in what may be termed as "frontier lands" or "refuge regions;" thus remote areas of great wilderness where the structure, not the components, of original ecosystems remains more or less untouched. In many cases these lands and waters are untamed, unknown, unowned, and unclaimed.

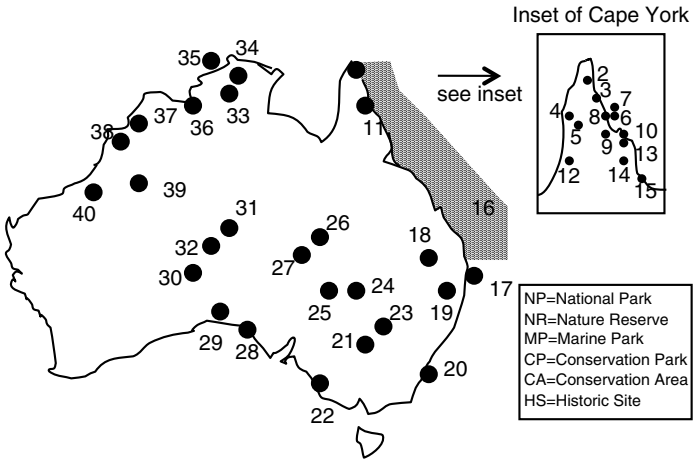
Biodiversity and Indigenous People's Lands and Waters

Indigenous peoples occupy a substantial share of the world's little disturbed tropical and boreal forests, mountains, grasslands, tundra, and desert, along with large areas of the world's coasts and near-shore waters (including mangroves and coral reefs). The importance of indigenous territories to biodiversity conservation is therefore evident. In fact, indigenous peoples control, legally or not, immense areas of natural resources. Among the most remarkable examples are the cases of the Inuit people (formerly known as Eskimo) who govern a region covering one fifth of the territory of Canada (222 million hectares), the indigenous communities of Papua New Guinea whose lands represent 97% of the national territory, and the tribes of Australia with nearly 90 million hectares (**Figure 1**). Although numbering only above 250,000, the Indians of Brazil possess an area of more than 100 million hectares, mainly in the Amazon basin, distributed in 565 territories (**Figure 2** and **Table 5**). Nearly 60% of the priority areas in central and southern Mexico recommended for protection are also inhabited by indigenous peoples (**Figure 3**), and half of the 30,000 rural communities are distributed in the ten most biologically rich states of the Mexican territory. In summary, on the global scale it is estimated that the total area under indigenous control probably covers between 12% and 20% of the earth's land surface.

The best example of notable overlaps between indigenous peoples and biological rich areas is the case of tropical humid forests. In fact, there is a clear correspondence between areas of remaining tropical forests and the presence of indigenous peoples in Latin America, the Congo Basin in Africa, and several countries of tropical Asia such as Philippines, Indonesia, and New Guinea. The strong presence of indigenous peoples in Brazil, Indonesia, and Zaire alone, which accounts for 60% of all the tropical forest of the world, is remarkable.

In Latin America, this geographical relationship has been strikingly verified for the Central American countries by a National Geographic Society map produced by a project headed by Mac Chapin in 1992. The same pattern can be found in the tropical humid areas of Mexico inhabited by 1.6 million indigenous people, and for many regions of the Amazonia basin (see the case of Brazil in **Figure 2**). It has been estimated that in Amazonia, above one million indigenous people of eight countries possess more than 135 million hectares of tropical forest (Davis and Wali, 1994). Many temperate forests of the world also overlap with indigenous territories as, for example, in India (**Figure 4**), Mynamar, Nepal, Guatemala, the Andean countries (Ecuador, Peru, and Bolivia), and Canada.

However, more than two million islanders the South Pacific, most of who are indigenous peoples, continue fishing and harvesting marine resources in high biodiversity areas (such as coral reefs).



1. Jardine River NP
2. Iron Range NP
3. Forbes Islands NP
4. Archer Bend NP
5. Rokeby-Croll NP
6. Flinders Island Group NP
7. Clack Islands
8. Cliff Islands
9. Lakefield NP
10. Starcke NP
11. Cape Melville NP
12. Mitchell and Alice Rivers NP
13. Mt.Webb NP
14. Cedar Bay
15. Mossman Gorge NP
16. Great Barrier Reef MP
17. Fraser Island NP
18. Carnarvon Gorge NP
19. Mt.Yarrowych NP
20. Jervis Bay NR
21. Lake Mungo NP
22. Coorong NP
23. Mt.Grenfell HS
24. Mootwingee NP
25. Gammon Ranges NP
26. Simpson Desert NP
27. Witjara NP
28. Yumbarra CA
29. Nullarbor NP
30. Unnamed CP
31. Watarraka NP
32. Uluru-Kata Tjuta NP
33. Nitmiluk NP
34. Kakadu NP
35. Gurig NP
36. Purnululu NP
37. Proposed Buccaneer Archipelago MP
38. Karlamilyi NP
39. Rudall River NP
40. Karijini NP

Figure 1 Terrestrial and marine protected areas with significant involvement or interest of indigenous peoples in Australia. Elaborated on data from the Australian Federal Government.

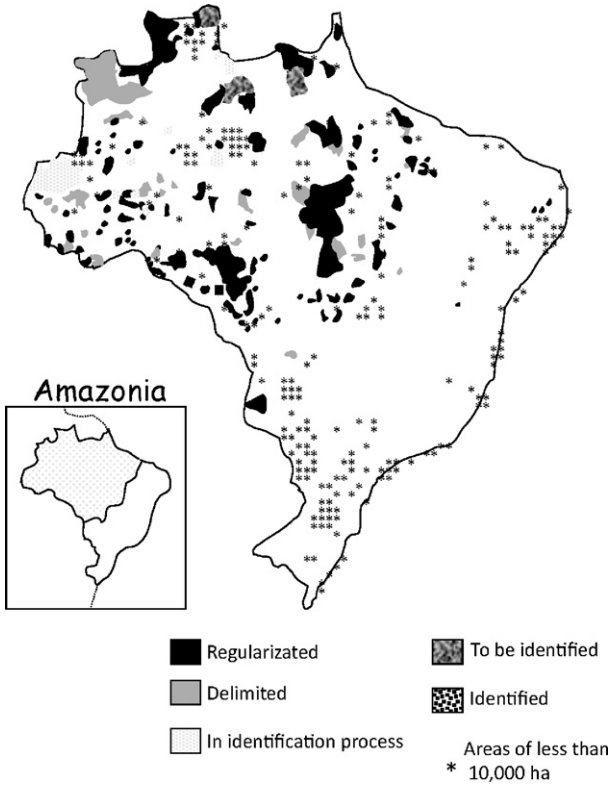


Figure 2 Geographical location of indigenous territories in Brazil, according to their legal situation and size area. Note the large tracts of areas under indigenous control in the Amazonian region, the core of Brazilian biological ichness. Reproduced from the map of Terras Indigenas do Brasil. Instituto Socioambiental, San Paulo, Brasil.

Table 5 Situation of indigenous territory in Brazil (November, 1997)

Legal situation[a]	*Number of indigenous areas*	*Area (ha)*	*%*
Not identified	74	2,749,000	2.60
To be identified	96	4,983,578	4.92
Interdicted	5	8,897,066	8.88
Identified	12	1,998,117	1.97
Delimited	67	19,963,673	19.86
Demarcated and Confirmed	73	14,816,728	14.77
Regularized	238	47,093,429	47.00
Total	565	100,501,591	100.00

[a]According to the National Indian Foundation (FUNAI).

Biodiversity and Ethnoecology: Indigenous Views, Knowledge, and Practices

Biodiversity is a very wide concept that refers to the variety of landscapes, ecosystems, species, and genes, including their different functional processes. Therefore, maintenance and conservation of biodiversity demands efforts on these four levels. While the first level is oriented to preservation of assemblies of "ecosystems," the second one focuses on protection of habitats in which the populations of species live.

At the species level, most knowledge on biodiversity concerns the large plants and animals such as flowering plants and vertebrates. The extent of diversity of smaller plants and animals remains to be inventoried and protected. While most biological diversity is constituted by wild plants and animals, an important subset involves the diversity among domesticated organisms. In this fourth level, the interest focuses on conservation of genetic variation of crops and domesticated animals. This section is dedicated to examining the potential role of indigenous peoples in biodiversity conservation from an ethnoecological perspective. Ethnoecology can be defined as an interdisciplinary approach exploring how nature is seen by human groups through a screen of beliefs and knowledge, and how humans, in terms of their images use or manage natural resources. Thus, by focusing on the *kosmos* (the belief system or cosmovision), the *corpus* (the whole repertory of knowledge or cognitive systems) and the *praxis* (the set of practices), ethnoecology offers an integrative approach to the study of the process of human appropriation of nature (Toledo, 2002). This approach allows recognition of the value of the belief-knowledge- practice complex of indigenous peoples in relation to the conservation of biodiversity.

The Kosmos

For indigenous peoples, land and in general nature, has a sacred quality which is almost absent from Western thinking (Berkes, 2008). Land is revered and respected and its inalienability is reflected in virtually every indigenous cosmovision. Indigenous peoples do not consider the land as merely an economic resource. Under indigenous cosmovisions, nature is the primary source of life that nourishes, supports, and teaches. Nature is, therefore, not only a productive source but the center of the universe, the core of culture, and the origin of ethnic identity. At the heart of this deep bond is the perception that all living and nonliving things and natural and social worlds are intrinsically linked (reciprocity principle). Of particular interest is the research done by several authors (Reichel-Dolmatoff, Boege, Ph. Descola, and van der Hammen) on the role played by the cosmology of several indigenous groups as a mechanism regulating the use and management of natural resources. In the indigenous cosmovision, each act of appropriation of nature must be negotiated with all the existing things (living and nonliving) through different mechanisms as agrarian rituals and shamanic acts (symbolic exchange). Humans are thus seen as a particular form of life participating in a wider community of living beings regulated by a single and totalizing set of rules of conduct.

The Corpus

Indigenous societies house a repertory of ecological knowledge which generally is local, collective, diachronic, and holistic. In fact, since indigenous peoples possess a very long history of resource-use practice, they have generated cognitive systems on their own circumscribed natural resources which

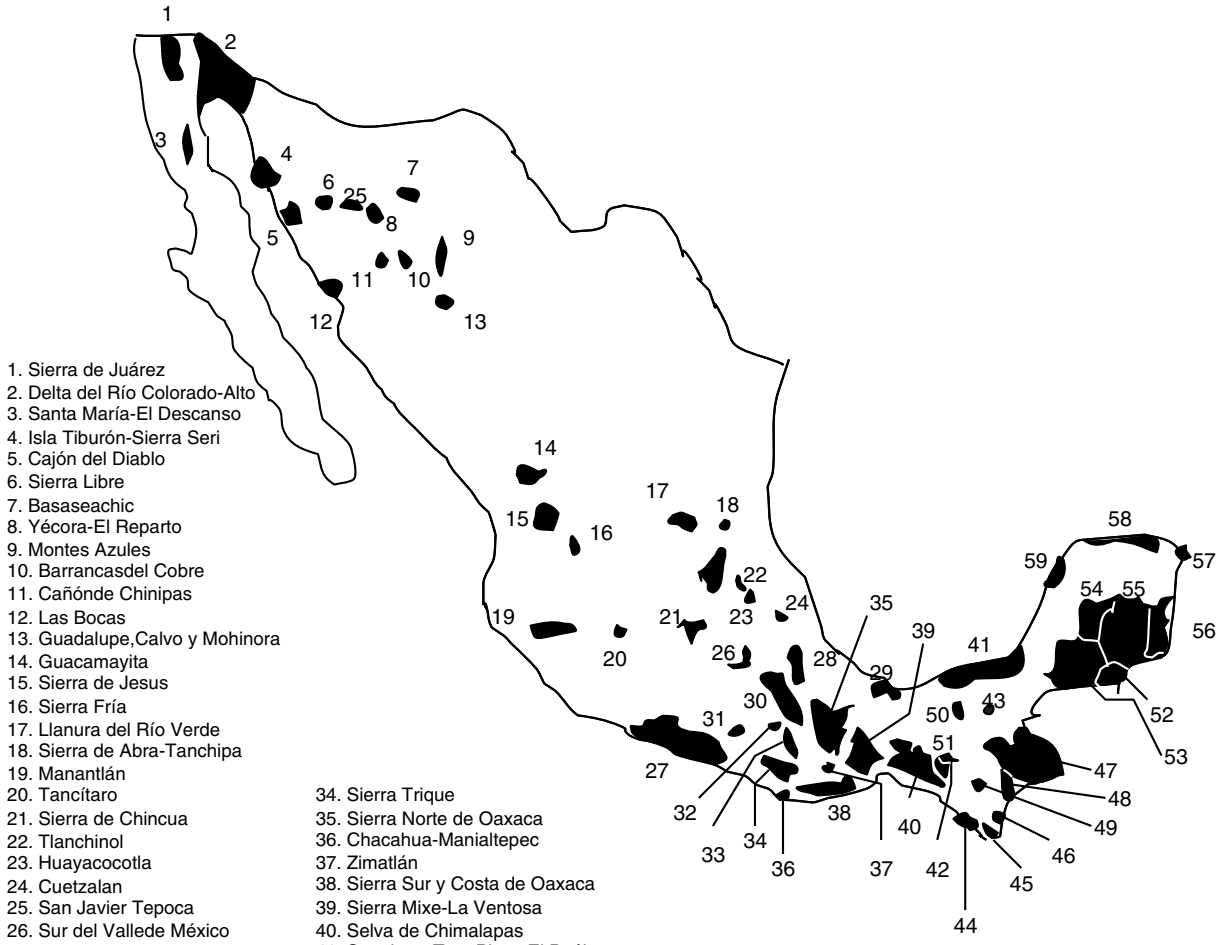


Figure 3 Geographical location of priority areas recommended by the Comision Nacional para el Estudio y Uso de la bodiversidad (CONABIO) of Mexico, overlapping with territories of indigenous communities. Note the high number of overlapping areas in the central and southern portion of Mexico, where most of the biological richness of the country is concentrated. Modified from CONABIO's map on priority areas for conservation, 1996.

are transmitted from generation to generation. The transmission of this knowledge is done through language, hence the corpus is generally an unwritten knowledge. Memory is, therefore, the most important intellectual resource among indigenous cultures.

This body of knowledge is the expression of a certain personal wisdom and, at the same time, of a collective creation, that is to say, a historical and cultural synthesis turned into reality in the mind of an individual producer. For this reason, the corpus contained in a single producer's mind expresses a repertoire that is a synthesis of information from at least four sources: (1) the experience accumulated over historical time and transmitted from generation to generation by a certain cultural group; (2) the experiences socially shared by the members of the same generation or cohort; (3) the experience shared in the household or the domestic group to which the individual belongs; and (4) the personal experience, particular to each individual, achieved through the repetition of the annual cycles (natural and productive), enriched by the perceived variations and unpredictable conditions associated with them.

Thus, indigenous ecological knowledge is normally restricted to the immediate environments and is an intellectual construction resulting from a process of accumulation of experiences over both historical time and social space. These three main features of indigenous ecological knowledge (being local, diachronic, and collective) are complemented with a fourth characteristic, namely holistic. Indigenous knowledge is holistic because it is intrincately linked to the practical needs of use and management of local ecosystems. Although indigenous knowledge is based on observations on a rather restricted geographic scale, it must provide detailed information on the whole scenery represented by the concrete

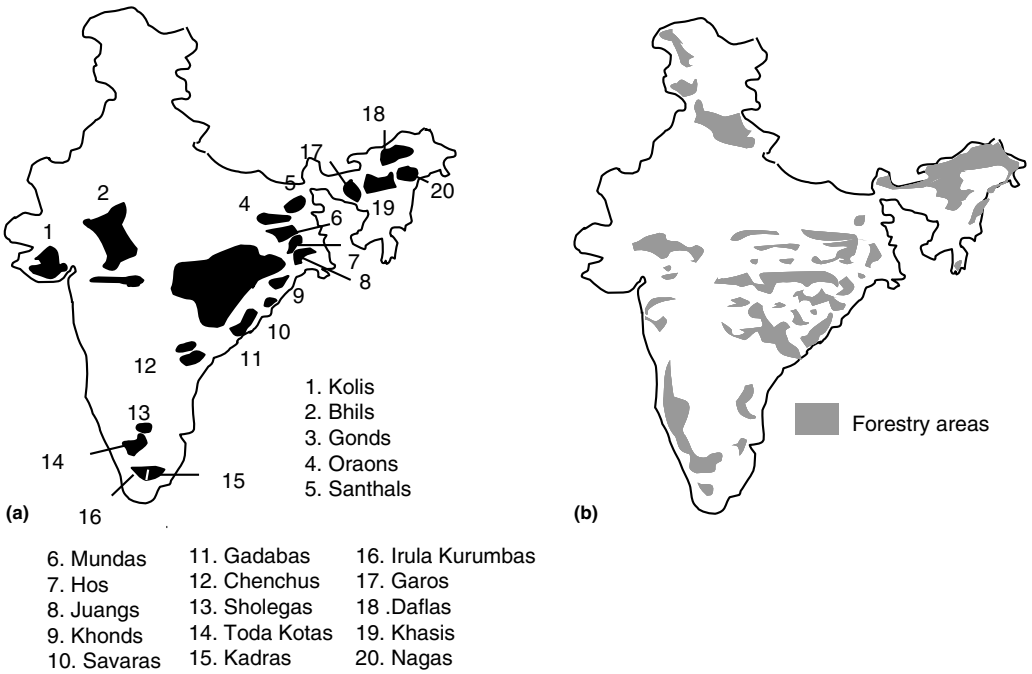


Figure 4 Geographical location of 20 indigenous groups (a), and principal forestry areas (b) of India. Although the long history of migrations of peoples makes it difficult to distinguish indigenous peoples in India, there are about 100 million people considered by the government as "scheduled tribes" speaking over 300 languages. These groups are generally residents of remote hilly or forested areas. Modified from The State of India's Environment 1984–1985.

landscapes where natural resources are used and managed. As a consequence, indigenous minds not only possess detailed information about species of plants, animals, fungi and some microorganisms; they also recognize types of minerals, soils, waters, snows, landforms, vegetations, and landscapes.

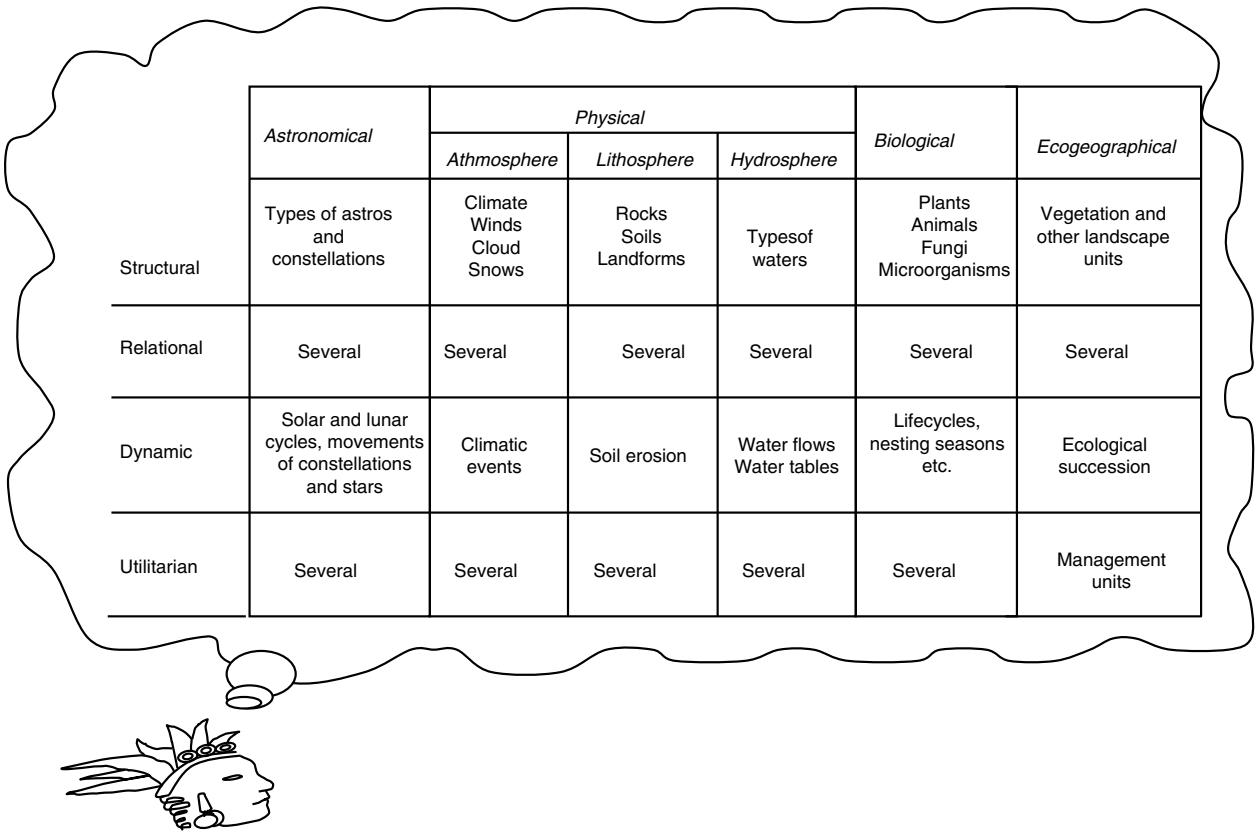
Similarly, indigenous knowledge is not restricted to the structural aspects of nature, which are related to the recognition and classification (ethnotaxonomies) of elements or components of nature, it also refers to dynamics (which refers to patterns and processes), relational (linked to relationships between or among natural elements or events) and utilitarian dimensions of natural resources. As a result, it is possible to integrate a cognitive matrix (**Figure 5**) which certifies the holistic character of indigenous knowledge and serves as a methodological framework to ethnoecological research (Toledo, 2002).

The Praxis

Indigenous societies generally subsist by appropriating a diversity of biological resources from their immediate vicinity. Thus, subsistence of indigenous peoples is based more on ecological exchanges (with nature) than on economic exchanges (with markets). They are therefore forced to adopt survival mechanisms that guarantee an uninterrupted flow of goods, materials, and energy from ecosystems. In this context a predominant use–value economic rationality is adopted, which in practical terms is represented by a multiuse strategy that maximizes the variety of goods produced in order to provide basic household requirements throughout the year (for further details on this strategy see Toledo, 1990). This main feature accounts for the relatively high self-sufficiency of indigenous households and communities.

Indigenous households tend to carry out a nonspecialized production based on the principle of diversity of resources and practices. This mode of subsistence results in the maximum utilization of all the available landscapes of the surrounding environments, the recycling of materials, energy and wastes, the diversification of the products obtained from ecosystems and, especially, the integration of different practices: agriculture, gathering, forest extraction, agroforestry, fishing, hunting, small-scale cattle-raising, and handicrafts. As a result, indigenous subsistence implies the generation of a myriad of products including food, domestic and work instruments, housing materials, medicines, fuelwoods, fibers, animal forage, and others.

Under the multiuse strategy, indigenous producers manipulate the natural landscape in such a way that two main characteristics are maintained and favored: habitat patchiness, heterogeneity, and biological as well as generical variation. In the spatial dimension, indigenous people create a complex landscape mosaic in which agricultural fields, fallow areas, primary and secondary vegetation, household gardens, cattle-raising areas, and water bodies are all segments of the entire production system. This mosaic represents the field on which indigenous producers, as multiuse strategists, play the game of subsistence through the manipulation of ecological components and processes (including forest succession, life cycles, and movement of materials). It has been demonstrated that some natural disturbances can increase biodiversity if they increase habitat heterogeneity, reduce the influence of competitively dominant species, or create opportunities for new species to invade the area.



	Astronomical	*Physical*			*Biological*	*Ecogeographical*
		Athmosphere	*Lithosphere*	*Hydrosphere*		
Structural	Types of astros and constellations	Climate Winds Cloud Snows	Rocks Soils Landforms	Typesof waters	Plants Animals Fungi Microorganisms	Vegetation and other landscape units
Relational	Several	Several	Several	Several	Several	Several
Dynamic	Solar and lunar cycles, movements of constellations and stars	Climatic events	Soil erosion	Water flows Water tables	Lifecycles, nesting seasons etc.	Ecological succession
Utilitarian	Several	Several	Several	Several	Several	Management units

Figure 5 Matrix of indigenous ecological knowledge.

However, the number of species is commonly relatively small in highly disturbed biotic communities, because few populations are able to re-establish themselves before they are reduced by later disturbances. In contrast, a low rate of disturbance provides few opportunities for pioneer species and might allow competitively dominant species to usurp limiting resources. Therefore, biodiversity is often greater at intermediate levels of disturbances than lower or higher rates.

The creation of landscape mosaics under the indigenous multiuse strategy in areas originally covered by only one natural community represents a human-originated mechanism which theoretically tends to maintain (and even increase) biodiversity. Several authors have already stressed the importance of the models of low intensity mosaic usage of the landscape by indigenous peoples and other small-landowner populations for biodiversity conservation. The same diversified arrangement found in indigenous landscapes tends to be reproduced at a microlevel, with multispecies, multistory crops, or agroforests favored over monocultures.

As a consequence, animal and especially plant genetic resources tend to be maintained in indigenous agricultural fields, aquaculture systems, homegardens and agroforests (see several examples in Toledo and Barrera-Bassols, 2008). Polycultural systems managed by indigenous agriculturalists and agroforesters are relatively well known and the recent specialized literature has plenty of case studies illustrating such designs. Especially notable are the homegardens and agroforestry systems of the tropical and humid regions of the world, which operate as human-made refuge areas for many species of plants and animals, notably in areas strongly affected by deforestation (Moguel and Toledo, 1999; McNeely and Schroth, 2006). At farm level, it is broadly recognized that crop populations are more diverse in indigenous farming systems than in agricultural areas dominated by agroindustriality. Therefore, indigenous peoples are recognized as key agents of on-farm preservation of plant genetic resources threatened by agricultural modernization (genetic erosion). The loss of biodiversity is also experienced in farming systems as indigenous cropping polycultural patterns are replaced by fossil-fueled monocrops. Indigenous agricultural systems and landscapes are then acknowledged as designs that preserve not only landraces of crop species, but also semidomesticated and wild crops relatives and even nondomesticated species.

Conserving Biodiversity by Empowering Indigenous Peoples

During the past three decades, as the loss of landscapes, habitats, species, and genes has become an issue of international concern, the protected areas of the world have increased notably both in size and number. However, as protected areas expanded, it became evident that the North originated model of uninhabited national parks could not be applied worldwide. Today, there are nearly 109,000 nationally protected areas (parks and other reserves) in more than 160 countries, which

represent more than 11.5% of earth's land surface (UNEP-WCMC, 2005; World Database on Protected Areas, 2005). Many of the areas that have been established as protected areas and many of those that are suitable for future addition to the protected area network are the homelands of indigenous peoples. In Latin America alone, more than 80% of protected areas are estimated to have indigenous people living within them. However, large tracts of the territories under indigenous control, estimated at between 12% and 20% of the earth's surface, are in the scope of conservationists as future reserves. Moreover, some authors like Nietschmann and Alcorn think that the bulk of the world's biodiversity is embodied within the limits of the indigenous territories of the tropical countries. Given the above, as well as the evidences offered and discussed in the previous sections, the idea that biodiversity conservation is impossible without the participation of indigenous communities is increasingly gaining recognition in national and international conservation circles.

Protected areas based on consultation, co-management, and even indigenous management, are likely to be increasingly important in coming years as the key role of indigenous cultures is being gradually recognized. It is important, however, not to over-idealize indigenous peoples and their resource-management strategies and stewardship skills. Conservationists have been frequently criticized for over-romanticizing indigenous peoples, creating a late-twentieth-century version of "the noble savage." Acknowledgment of the positive links between indigenous peoples and biodiversity has been increasingly tempered by the recognition that under certain circumstances (high population densities, market pressures, unsuitable technologies, local disorganization) indigenous peoples can act as disruptive, not as conservationist, actors.

Biological diversity and sustainable development are today two of the most powerful and central concepts in environmental protection. In recent years, special attention has been paid to the sustainable development of community-based peoples, as a key mechanism for the reinforcement of correct participation of local communities in biodiversity conservation. It is possible to define sustainable community development as an endogenous mechanism that allows a local society to take (or retake) control of the processes that affect it. In other words, self-determination and local empowerment, conceived as a "taking of control," have to be the central objectives in all community development.

Given the demonstrated importance of indigenous peoples for biodiversity conservation, it is essential to recognize the necessity of empowering local communities. That is, to maintain, reinforce, or give control to the indigenous communities of their own territories and natural resources as well as providing sufficient access to information and technology. Important here are legally recognized and enforceable rights to lands and waters, which give the communities both an economic incentive and a legal basis for stewardship. In many countries, national recognition and policy support for existing, community-based property rights systems are crucial. In many Asian and African countries, returning a measure of control over public lands and resources to local communities is also fundamental to slowing biodiversity loss in threatened regions.

Similarly, it is very important to establish new resource-management partnerships between local communities and the state and other society institutions to maintain biodiversity. Local stewardship in conjunction with external governmental and nongovernmental agencies and institutions is perhaps the best way to guarantee effective protection of landscapes, habitats, species, and genes worldwide, and especially in tropical countries.

Concluding Remarks: A Bio–Cultural Axiom

The research accumulated in the last three decades by investigators belonging to the fields of conservation biology, linguism and anthropology of contemporary cultures, ethnobiology, and ethnoecology, have evolved convergently toward a shared principle: that the world's biodiversity will be effectively preserved only by preserving diversity of cultures and vice versa. This common statement, which represents a new bio–cultural axiom, has been nourished by four main sets of evidence: geographical overlap between biological richness and linguistic diversity and between indigenous territories and biologically high-value regions (actual and projected protected areas), recognized importance of indigenous peoples as main managers and dwellers of well-preserved habitats, and certification of a conservationist-oriented behavior among indigenous peoples derived from its premodem belief–knowledge–practices complex.

This bio–cultural axiom, called by Nietschmann the "concept of symbiotic conservation," in which "biological and cultural diversity are mutually dependent and geographically coterminous," constitutes a key principle for conservation theory and applications, and epistemologically is an expression of the new, integrative, and interdisciplinary research gaining recognition in contemporary science.

References

Alcorn J (1993) Indigenous peoples and conservation. *Conservation Biology* 7: 424–426.

Alcorn J (1994) Noble savage or noble state? Northern myths and southern realities in biodiversity conservation. *Etnoecologica* 3: 7–19.

Barrera-Bassols N and Toledo VM (2005) Ethnoecology of the Yucatec Maya: Symbolism, knowledge and management of natural resources. *Journal of Latin American Geography* 4: 9–40.

Berkes F (2008) *Sacred Ecology*, 2nd edn. New York: Routledge.

Chapin M (2004) A challenge to conservationists. *World Watch Magazine* 176: 24–36.

Davis SH and Wali A (1994) Indigenous land tenure and tropical forest management in Latin America. *Ambio* 23: 207–217.

Durning AT (1993) Suporting indigenous peoples. In: Brown L (ed.) *State of the World*, pp. 80–100. Washington, DC: World Watch Institute

Gordon Jr. RG (ed.) (2005) *Ethnologue: Languages of the World*, 15th edn. Dallas, Texas: SIL International. http://www.ethnologue.com/

Hale K (1992) On endangered languages and the safeguarding of diversity. *Language* 68: 1–2.

Maffi L (ed.) (1999) *On Biocultural Diversity: Language, Knowledge and the Enviroment*. Washington, DC: Smithsonian Institution Press.

Maffi L (2005) Linguistic, cultural and biological diversity. *Annual Review of Antrhopology* 29: 599–617.

McNeely JA and Schroth G (2006) Agroforestry and biodiversity conservation: Traditional practices, present dynamics and lesson for the future. *Biodiversity and Conservation* 15: 549–554.

Mittermeier R and Goettsch-Mittermeier C (1997) *Megadiversity: The Biological Richest Countries of the World*. Mexico City: Conservation International/CEMEX/ Sierra Madre.

Moguel P and Toledo VM (1999) Biodiversity conservation in traditional coffee systems in Mexico. *Conservation Biology* 13: 1–12.

Oldfield M and Alcorn J (eds.) (1991) *Biodiversity: Culture Conservation and Ecodevelopment*. Boulder, CO: Westview Press.

Orlove BS and Brush SB (1996) Anthopology and the conservation of biodiversity. *Annual Review of Anthropology* 25: 329–352.

Oviedo G, Maffi L, and Larsen PB (2000) *Indigenous and Traditional Peoples of the World and Ecoregion Conservation: An Integrated Approach to Conserving the World's Biological and Cultural Diversity, and Companion Map Indigenous and Traditional Peoples and the Global 200 Ecoregions*. Gland, Switzerland: WWF-International and Terralingua.

Stevens S (ed.) (1997) *Conservation Through Cultural Survival: Indigenous Peoples and Protected Areas*. Washigton, DC: Island Press.

Toledo VM (1990) The ecological rationality of peasant production. In: Altieri M and Hecht S (eds.) *Agroecology and Small-Farm Development*, pp. 51–58. Boca Raton, Florida: CRC Press.

Toledo VM (2002) Ethnoecology: A conceptual framework for the study of indigenous knowledge of nature. In: Stepp JR, *et al*. (eds.) *Ethnobiology and Biocultural Diversity*, pp. 511–522. International Society of Ethnobiology.

Toledo VM and Barrera-Bassols N (2008) *La Memoria Biocultural*. Barcelona, Spain: Editorial Icaria.

Toledo VM and Barrera-Bassols N (2009) A etnoecologia: Uma ciencia pós-normal que estuda as sabedorias tradicionais. *Desenvolvimento e Meio Ambiente* 20: 7–27.

UNEP-WCMC (2005) *World Database on Protected Areas*. Cambridge: UNEP World Conservation Monitoring Centre. http://sea.unep-wcmc.org/wdbpa

Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon

Giorgini A Venturieri, Universidade Federal de Santa Catarina, Florianópolis, Brazil

© 2013 Elsevier Inc. All rights reserved.

Glossary

Barreais Soils naturally enriched with organic matter and clay carried by the rain and/or a river.
Caboclos Cultural and/or racial crossing between European and indigenous populations.
Macaxeiras Group of varieties of manioc (*Manihot esculenta* Crantz) with low levels of cyanidric acid. These varieties may be consumed without prior treatment to lower their toxicity, as must be done with other varieties of wild manioc. Synonyms: Sweet manioc and aipim.
Marimba Elevated flooring made of wood, used in flooded areas for placing cattle or plants during river flood periods.
Reproductive isolation When a derived population is not able, for some reason, to breed and obtain fertile progeny with the original population.
Wild manioc Manioc with a high level of cyanidric acid.

In this article, concepts are discussed relating to natural biodiversity used to form new crops and for the improvement of already domesticated species. A majority of examples come from the underexploited plants used by Amazonian indigenous people, little known in the international scientific literature, but which are splendid examples of the capacity for breeding and selection of traditional communities. The article also gives curious reports of Amazonian Indian practices of management and interchange of germplasm from useful plants.

What is Domestication?

The evolution of a plant occurs when there is a selection; this can be the result of human activity, which generally values attributes of interest to people in detriment to those necessary for the plant's own survival in the wild. When this process reaches high levels of genetic modification, to the point that human aid is necessary for survival, the plant is considered to have achieved the stage of being domesticated (Harlan, 1975). Useful plants occur at various stages of domestication that require different degrees of management and cultivation. Generally, the more domesticated the plant, the more demanding it is as to the degree of management and cultivation, but the fact of a plant being managed or cultivated does not imply that it is domesticated. However, any plant that has been managed or cultivated for a long time is inevitably undergoing modification towards domestication because it is placed in environmental conditions improved by humans, who reduce or eliminate the usual processes of natural selection such as competition and predation. Domestication is related to the plant's genetic response, management, and cultivation and human activity in dealing with the plant (Harlan, 1975).

This article will deal principally with practices still used by Indians and Caboclos (rural peoples often of mixed ancestry) of Amazonia who may directly (conscious selection) and indirectly (unconscious selection) be inducing modifications in natural plants in the direction of their domestication.

How Does Domestication Happen?

The forces of selection, resulting from management practices or from cultivation *per se* and as induced by farmers, may be consciously or unconsciously provoked (Harlan, 1975). In conscious selection, the farmer sees some interesting attribute in a plant and seeks to propagate it in some way, be it by sexual means (seeds) or asexually (generally by trying to induce a portion of stem to take root), thus forming a new population. If the characteristic selected has good inheritability – that is, it can be transferred from one generation to another – the result is an improved population. In unconscious selection, as the name indicates, this selection does not occur but the populations are modified due to some human intervention, generally in the environment or harvest system, which ends up benefiting some segment of a population to the detriment of the others, generating a selected population. The classic example of unconscious selection is that of rice, which upon being harvested has seeds in its panicle that do not easily fall from the plant. When these seeds are planted, they carry alleles that favor this characteristic, which is selected increasingly more favorably with each harvest. Seeds of domesticated rice today do not fall as easily from the panicle as do those of more primitive varieties (Harlan *et al.*, 1973).

The well-known "genetic drift" of the classic works on evolution is also very much associated with the process of domestication. Genetic drift occurs when a very small portion of a certain population is taken to another area that is isolated and it regenerates there, but does so only with the descendants of the few individuals introduced therein (**Figure 1**). Many times this process does not bring gains in productivity but does induce an increase in more or less stable varieties that end up forming the "local varieties." This process is very common among farmers who have the habit of collecting only a few seeds from a given plant for planting on their farms or villages, generally far from the original population and isolated by some environmental barrier. This may be a large river or, in the case of plants typical of disturbed environments such as the plants domesticated in farm plots, a climax forest in which such species do not occur.

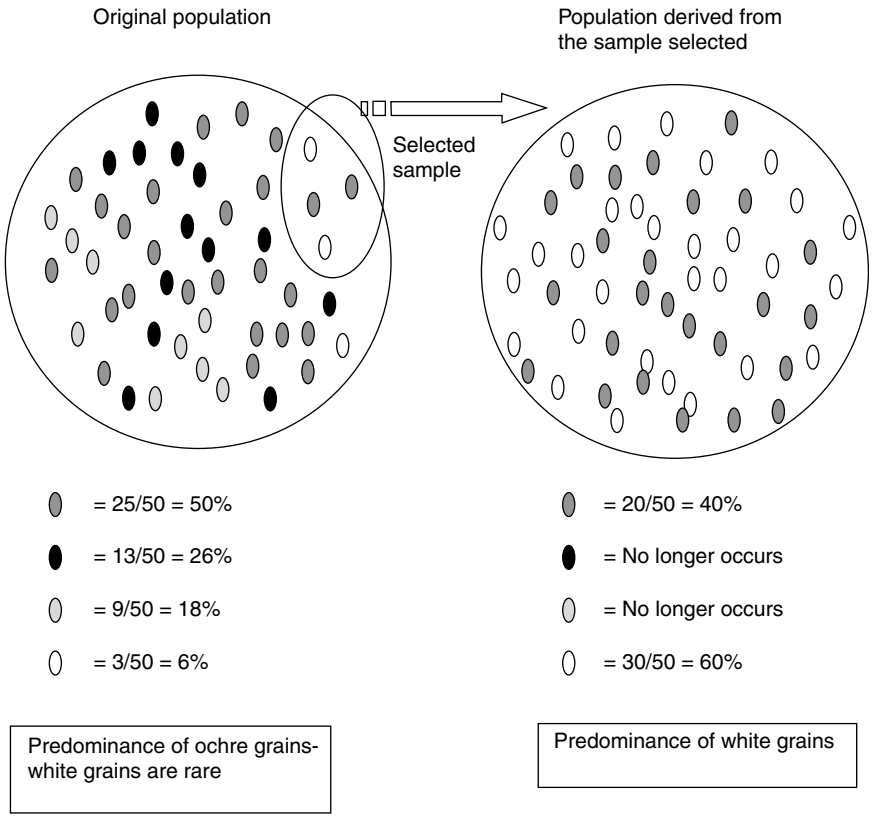


Figure 1 Hypothetical representation of genetic drift for the grain color character, considered not to have adaptive value. Genetic drift occurs when the sample is small and does not represent all the classes of a determined characteristic.

Currently domesticated plants were some time ago wild plants which were in stages used, then managed, and finally cultivated (Clement, 1992a, 1992b). The process of domestication, for the majority of species, took thousands of years. For corn, whose domestication process is the best known, it is estimated that the process took more than 3000 years (Iltis, 1986). If, for an annual plant, so much time was spent, one can imagine how long it took to domesticate the Amazonian fruit trees that are perennials. The appearance of a domesticated plant may also be sudden, due to some mutation that provokes a very radical modification in the plant, such as alterations in the number of chromosomes. For example, Cupuaçu (*Theobroma grandiflorum* Schum) is an Amazonian fruit highly valued for its pulp, but whose seeds constitute 32% of its total weight. A triploid variety that produces only pulp and no seeds has been found and is being propagated by grafting (Venturieri and Mendonça, 1985). Another case of chromosome number modification is observed in Guaraná (*Paullinia cupana* Kunth var. *Sorbilis* (Mart.) Ducke), a stimulant plant due to its high caffeine content, initially of restricted use by Amazon people, has now spread across the world thanks to the energizer and soft drinks industry. The variety *Sorbilis*, domesticated by the Sateré Maué, where it originated, has 210 chromosomes, while other species of the genus have 24 (Freitas *et al.*, 2007). This recent polyploidization is an apparently recent event because the natural genetic variability of domesticated guaraná, assessed by molecular markers, is small (Atroch *et al.*, 2009). Peach palm (*Bactris gasipaes* H.B.K.), considered to be the only palm tree domesticated in the Americas, is probably a hybrid between *B. dahlgreniana* Glassman and *Guilielma insignis* (Clement, 1988). Finally, there is the suppression of some gene or group of genes that made a species poisonous, as is the case with wild manioc, with a high level of cyanidric acid, producing macaxeira with very low levels of cyanidric acid. Nowadays, with the help of advanced selection and biotechnology techniques, the process of domestication can be considerably shortened, but still 99% or more of the species currently domesticated were achieved by farmers before the "green revolution" (Harlan, 1975).

The greater the diversity of these plants, in terms of species as well as of varieties within each of these species, the greater would be the options of choices for humans.

For selection to be efficient there must be reproductive isolation between the population selected and the original group from which it was taken. If there is no isolation, the new population again crosses with the original population and becomes similar to it, without achieving any gain. The more efficient this isolation, the more rapidly a given population may be domesticated (**Figure 2**). In indigenous peoples' systems, this isolation is achieved by environmental barriers or management practices. In summary, to obtain a domesticated plant there must be (1) the chance for choice (genetic diversity), (2) discovery of useful plants and obtainment of interesting varieties, and (3) reproductive isolation of the varieties selected so that they do not return to the wild state. How were these predicates achieved for Amazonian plants?

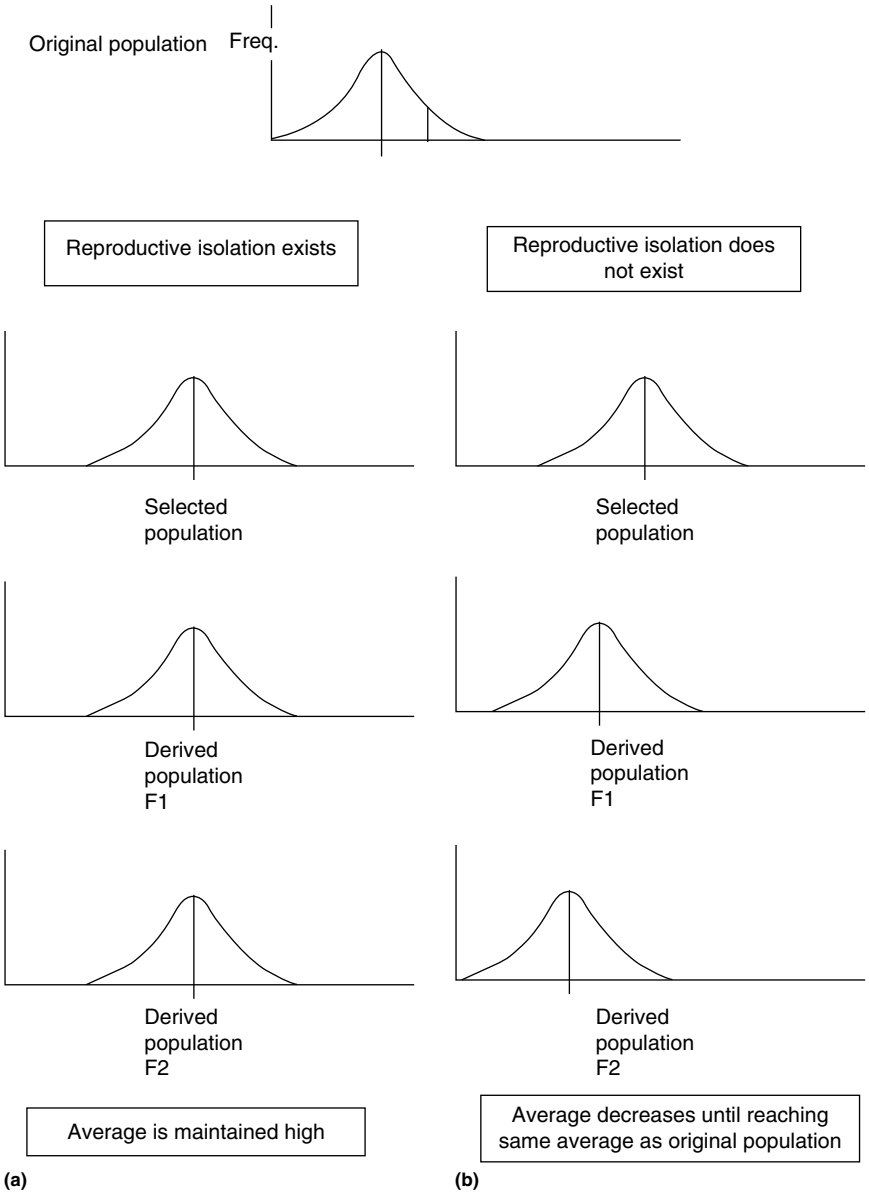


Figure 2 Diagram exemplifying the effect of reproductive isolation. A selected sample with large fruits is removed from an original population. (a) When the isolation is efficient – that is, when the new population is planted in a plot at a distance that avoids gene exchange with the original population – the new population will produce descendants with large fruits. (b) When isolation is not efficient, the new population may cross with individuals from the gene group that gave rise to it and re-establish the characteristics of the original population among its descendants. The more effective the reproductive isolation, the more rapidly genetic gains are achieved.

The Chance for Choice (Genetic Diversity)

The Amazonian forest offers a grandiose biodiversity with an immense number of useful plants for domestication. Considering only the fruiting species (Benza, 1980; Cavalcante, 1991; Clement, 1999; Correa, 1926–1969; Hoehne, 1946; Kerr and Clement, 1980; Kerr and Posey, 1984; Le Coint, 1947; Smith *et al.*, 1992), one may prepare a list with more than 250 typically Amazonian species, some of which are internationally recognized, others with use restricted to Indians who extract them from the native forest, and other "little-known" species in general use among the civilized population that have undergone Indian influence. Certainly, if one considers the vastness of Amazonia with its 4 million km^2 and its incredible variability of plant species, estimated to be between 50,000 and 60,000 for the higher plants alone (Clement *et al.*, 1982), and the current state of economic botany in the region, it may be concluded that this list may still be increased considerably. For example, in two expeditions specifically planned for collecting potential fruiting species in the Amazon region, 10 species were attained that had not yet been cataloged (Instituto Nacional de Pesquisas da Amazônia/Núcleo do Acre, 1984; Venturieri *et al.*, 1986/1987).

The Discovery of Useful Plants and Acquisition of Interesting Varieties

In the eyes of "civilized" society, it is difficult to understand the process of discovering uses for plants. The Indians experiment with almost everything new that they find in the forest or that is brought to them by white people. As presents, they like receiving new plant species and varieties, such as the exotic plants that have recently been introduced in their cultures. The Kayapó Indians are very fond of mango, an Asian tree, and plant the seeds everywhere (Kerr and Posey, 1984), as do the Nambiquara Indians. Some Indians describe methodologies for discovering the usefulness of plants by employing concepts that are difficult to understand; for example, an Indian explained to me that the discovery of papo-de-peru plant (*Aristoloqia* sp.) served as an antidote to snake venom because of the fact that "its flower looked like the skin of a snake." It is common for them to risk their lives for discovery. Among the Hahaintesu group of the Nambiquara the ingestion of the aril of seeds of a *Philodendrum* sp. (**Figure 3**), the latex and fruit of which they recognized as very poisonous and capable of causing death was witnessed. They would split the fruit and with tweezers carefully remove the seeds so that they were not contaminated by the latex and then place them in their mouths to suck the arils, doing so carefully so as not to bite the seeds, which were later discarded. These Indians eat tarantulas (Setz, 1991) after singeing them to remove hairs that can cause irritations to the skin (in this case, in the mouth). They also cook and eat the tuberous stalk of a plant of the genus *Zamia*, which if eaten raw causes a strong headache (**Figure 4**). Apparently, the process of discovering a plant's usefulness is by trial and error since the analogies between aspect and use are not logical. From the amount of information existing on edible and medicinal plants, one may imagine how many lives were lost and pains felt by these intrepid researchers.

Domesticated plants are generally described as having "appeared" through some myth. Among the Hahaintesu group of the Nambiquara there is a myth that describes the origin of man, whose head was the gourd (*Crescentia cujete* L.), whose teeth were the corn (*Zea mays* L.), whose ear was the lima bean (*Phaseolus lunatus* L.), whose bones were the manioc (*Manihot esculenta* Crantz), and whose scrotum was the cará yam (*Dioscorea trifida* L.) (**Figure 5**). In the Kayapó myth, corn was acquired from the stomach of a rat that lived in an enormous tree in the middle of an area that was being prepared for planting; the strength of the entire tribe was needed to fell the tree (Lukesch, 1976). Another Kayapó myth mentions that sweet potato, macaxeira, and banana were brought by the "daughter of the rain" who came down from the sky, became an ordinary Indian, married, had children, and during a time of hunger returned to the sky and brought back these species that were then cultured so that the tribe never again went hungry (Lukesch, 1976). Others describe useful plants as being inherited from conquered people. The Tikuna recognize that many of the obviously domesticated fruits were already present when they arrived, possibly derived from the Omagua Indians, whom they expelled (Ailton Krenak, União das Nações Indígenas, personal communication), or they negotiated with other tribes (among the Hahaintesu group of the Nambiquara there is a variety of manioc called the "manioc of the Wassuso," a tribe from the same linguistic group that became practically extinct – in 1985, there was only one elderly woman of this tribe).



Figure 3 *Philodendrum* sp., a poisonous species whose seeds are removed from the fruit with the aid of tweezers and the aril is sucked carefully so as not to bruise the seeds, which are also toxic.



Figure 4 Indian boy of the Hahaintesu group of the Nambiquara showing a plant of the genus *Zamia*, whose tuberous stalk can only be eaten cooked. When ingested raw, it causes a strong headache.



Figure 5 Species from the Nambiquara myth about the appearance of cultivated plants cultivadas: (a) The head is represented by the calabash (*Crescentia cujete* L.), (b) the teeth by corn (*Zea mays* L.), (c) the ear by the lima bean (*Phaseolus lunatus* L.), (d) the bones by manioc (*Manihot esculenta* Crantz), and (e) the scrotum by the cará yam (*Dioscorea trifida* L.).

The Indians' diet is based on many plants and their varieties. One commonly finds up to 40 varieties of manioc per indigenous tribe (Kerr and Clement, 1980; Boster, 1984; Elias *et al.*, 2000), witnessing the dynamism these people have for renewing these varieties. Manioc is usually propagated by stakes (cloned) which avoids combination between genotypes. However, if this were always the case, how would recruitment of new varieties for selection exist? How would one justify the great number of varieties observed among the Indians? Contrary to what is found in the literature of modern agriculture, manioc is also propagated by seeds (Kerr and Posey, 1984; Pujol *et al.*, 2002). Plantules derived from seeds make possible new combinations resulting from crossing preselected parents propagated by stakes. A curiosity, volunteer seedlings of cassava emerge from a bank of dormant seeds in the soil established during the previous cycle(s) of swidden cultivation, 3–30 years/cycle (average appoximately 10 years) (Elias and McKey, 2000). Dormant seeds need high temperature to germinate and remain ungerminated on lower temperature proportioned by shade through the fallow period;

they germinate synchronously after vegetation is slashed and burned. At the time of harvest, they decide whether to include the plants among those that will be vegetatively propagated and so incorporate the seedling as a new cultivar. As a consequence, the cassava germination process has a short relation with swidden cultivation and so, with its domestication (Elias and McKey, 2000; Pujol *et al.*, 2002). The Indians of the upper Rio Negro recognize plantules of manioc as derived from seed banks of old plantings in areas that have regenerated and been slashed and burned anew. This recognition is easy because the plantules derived from seeds display their cotyledons (Paulo Sodero Martins, Escola Superior de Agricultura "Luíz de Queirós," personal communication). Traditional farmers from Ubatuba – SP, southern Brazil, also recognize cassava seedlings by their pivotal roots, contrasting cuttings that have lateral roots (Sambatti *et al.*, 2000).

The Hahaintesu group of the Nambiquara in 1985 possessed only nine varieties of macaxeira, almost all recently introduced after their areas were demarcated and their extermination by whites was interrupted. The increase in varieties is an indication that peace may be giving the Indians a chance to recompose their genetic collection of plants (Boster, 1984). Generally, as in the case of birth control, the administration of genetic resources is directed by women (Kerr and Posey, 1984). In the case of the Hahaintesu group of the Nambiquara, an elderly Indian woman (as mentioned earlier, the last of the Wassuso, married to the Indian Papai) possessed a great knowledge and love for agricultural plants, and others frequently brought her new plants and varieties. This woman would plant them in different environments and evaluate them. Those that were approved she handed over to the rest of the tribe for planting. This Indian woman and her garden fulfilled what would be the role of "curatorship and experimental station for genetic resources" in our society.

The Amazonian Indians highly value esthetics, as expressed through their clothing, body decorations, and basketry, which are generally exuberant in both form and color. This sentiment is also reflected in the plants and varieties domesticated by them. Examples include peach palm (*Bactris gasipaes* Kunth) (**Figure 6(a)**) and peppers of the genus *Capsicum* (**Figure 6(c)**) and cubiu (*Solanum sessiliflorum* Dunal.) (**Figure 6(e)**), which have a large variety of colors, sizes, and shapes. Among Indians, selection is also made by productivity and taste. Peach palm, one of the best-domesticated plants of the humid



Figure 6 Good examples of the great variety of colors, sizes, and shapes of domesticated plants selected by Amazonian Indians: (a) Peach palm (*Bactris gasipaes* Kunth), (b) Biribá (*Rollinia mucosa* Baill.) var. Tikuna, (c) peppers of the genus *Capsicum*, (d) abiu (*Pouteria caimito* Radl.) var. Tikuna, and (e) cubiu (*Solanum sessiliforum* Dunal). See also color insert (this volume). (b, d) Reproduced with permission of Clement CR.

tropics, can produce up to 50 t ha^{-1}, with varieties selected for flour or oil and others that easily release the skin and have a refined flavor for cooking; there are also varieties with and without spiny trunks (Clement, 1988; Clement and Mora Urpi, 1987). In the state of Roraima, a Brazilian state in southwestern Amazonia, 78 morphotypes of *Capsicum* peppers have been identified (Barbosa *et al.*, 2002). Cubiu (*S. sessiliflorum* Dunal.), a fruiting plant similar to tomato, produces up to 100 t ha^{-1} and has varieties with eight different shapes ranging from fusiform to semirounded, with colors ranging from yellow to purple (Silva-Filho *et al.*, 1993, 1999).

New varieties are also achieved through barter, raiding, and dower. Groups of friendly Indians visiting from other tribes take varieties as presents. Rival Nambiquara trade varieties in a ritual that is difficult to understand. They peep at each other from opposite sides of a river, without exposing themselves so as not to be killed; they sing and call out challenges. Suddenly, they gain confidence and cross the river to leave presents that sometimes are repaid. After the exchanges they go back to being mortal enemies. Among the Alantesu group of the Nambiquara, after marriage is consented to for a woman, some compensation is required from the groom, which may take the form of labor, a bicycle, a radio, or some other article or group of articles. When the debt is paid, the husband takes his wife and with her some plants, which introgresses genotypes from other tribes. Kidnapping is very common among the indigenous tribes of Amazonia (Black *et al.*, 1991). They kidnap children (the chief of the Ureu-au-au Indians, shown in a 1970s documentary about colonization in Amazonia along the BR-364 highway in the state of Rondônia, was a white man without any cultural traces of white civilization). Women are kidnapped, which seems to be a common practice among the Indians that does not annoy the kidnap victim for very long. Among the Hahaintesu group of the Nambiquara, a kidnapped woman was seen, who was imprisoned in a straw hut from which escape would have been easy and was guarded by her kidnapper, who stayed nearby. During the first days, she was aggressive toward those who approached her; two days later, she was seen laughing among the other Indian women. It was later learnt that she returned to the agricultural plots of her original tribe and brought back new varieties of manioc, and thus increased her standing with her new tribe.

Domesticated plants return to their wild state if the selective human pressure is relaxed (Harlan, 1975); therefore, there is a close relation between humans and domesticated plants. This relation is so intimate that gene introgression observed in Indian populations in Amazonia (Black *et al.*, 1991) may be associated with the introgression also found among populations of domesticated plants, such as is observed in peach palm (*B. gasipaes* Kunth) (Picanço *et al.*, 1999) and in caiaué (*Elaeis oleifera* Cortés) (*E. Barcelos*, Embrapa Amazonia Ocidental/Manaus, personal communication, and data partially published in Barcelos *et al.*, 1999).

Among the Tikuna in the region near Leticia, Colombia, there is a tribal custom that reinforces plant selection: "All fruit or other edible product that is especially large or flavorful is divided among the tribe and seed planted." This process, practiced over many generations, led, by massal selection, to an increase in desirable alleles, giving rise to the domestication of fruiting species such as peach palm (*B. gasipaes* Kunth; **Figure 6(a)**), biribá (*Rollinia mucosa* Baill.; **Figure 6(b)**), and abiu (*Pouteria caimito* Radl.; **Figure 6(d)**) (Kerr and Clement, 1980). The fruiting species of that region have very beautiful varieties with giant fruits, showing clear signs of being domesticated; therefore, this is suggested as a center of domestication (Clement, 1989).

The Selection, Fixation, and Status of Amazonian Plant Domestication

Reproductive Isolation

For one population to be different from another, it must fragment off from the original and in some manner maintain itself distinct in genetic terms; in other words, it should have very low or nonexistent chances of crossing with individuals of the population from which it originated. Any reproductive barrier is capable of permitting the fragmentation of a population and its differentiation into subpopulations. Many isolation mechanisms are known from cultivated plants: Geographic and ecological separation, differences in time of blooming, self-fertilization, translocation races, polyploid races, gametophytic factors, cryptic chromosomal differences, and meiotic irregularities (Harlan, 1975). These mechanisms may occur simultaneously and be imperfect, and if the subpopulations meet someday some degree of hybridization may occur. Subpopulations with potential for intercrossing are called varieties. When the isolation is perfect and there are no longer possibilities for a natural cross to occur, then with the generation of fertile progenies a new species has been formed.

The Environment

Practices capable of inducing modifications in the structure of a population are intimately related to the physiognomy of the site and to the ecological restrictions that each environment offers. In Amazonia, although there are a range of ecosystems (for climate, see Salati, 1985; for soils, see Jordan, 1985; and for vegetation, see Pires and Prance, 1985) and a varied number of singular strategies used by the Indians who practice agriculture on lands with very low trophic support – such as the use of barreais near the rivers (Hill and Moran, 1983), in the savannas, with the formation of "managed islands" or apêtê of the Kayapó (Anderson and Posey, 1989), and plaza towns with ditches and roads called "garden cities" managed by the Kuikuros indians (Heckenberger *et al.*, 2008). Recently, earthworks related to water management for irrigation, drainage, and fishponds were discovered in the Amazon, all sophisticated systems of agriculture, constructed before the conquistadors arrived. They are: The raised fields on flooded areas at the Marajó Island in the Amazon delta (Rossetti *et al.*, 2009; Roosevelt, 1991; Schaan, 2008), Coastal Amazon of French Guiana (Rostain, 2010), and Baures region and the larger Llanos de Mojos in the Bolivian Amazon (Erickson, 2010), and ring ditches, also referred to as "geoglyphs," along the Branco, Mamoré, Madeira, Orton, Madre de Dios, Beni, and Guaporé rivers in the Brazilian and Bolivian Amazon (Erickson, 2010, Pärssinen *et al.*, 2009, Saunaluoma, 2010;

Schaan *et al.*, 2008). The oldest earthworks dated back to AD 300, the majority were from between AD 400 and 1400, and the last known to be constructed was from around AD 1700, almost all were abandoned around 1500 (Erickson, 2010; Pärssinen *et al.*, 2009; Roosevelt, 1991; Rostain, 2010; Saunaluoma, 2010; Schaan, 2008). Sophisticated systems of agriculture require genetic material of high productivity, and possibly the people who manned the earthworks generated collection of useful plants and cultivars that we are unlikely to discover how they were achieved. The abandonment of such costly construction as the earthworks was certainly linked to population declines, and with them went many cultivars because they need human intervention to maintain their genetic status (Harlan, 1975). It is speculated that they were adopted by indigenous groups remaining, and much of this material can still be found and collected (Lima *et al.*, 1997). Currently agriculture in indigenous communities is basically practiced on (1) upland areas (terra firme) covered by forests and (2) the river floodplains (várzeas). On terra firme covered by forest, agriculture is practiced in cycles of felling, burning, planting, and fallowing, also called slash-and-burn or swidden fallows. This is the most applied system of agriculture in Amazonia. A site with exuberant forest is located and the understory trees are cut first, followed by the large ones. Three months later, fire is applied. The ashes and organic matter formed by decomposing tree trunks nourish the soil for at most 4 years. When these soils are exhausted, cultivation of annual plants is abandoned and a new cycle must begin. Among Indians and many caboclo communities, this is not a complete abandonment because there is a subsequent recuperation stage of planting with useful perennial plants (Denevan *et al.*, 1984). This system of apparently migratory agriculture has been very efficient in improving plants because each plot receives some seeds of selected plants. Because the plots are relatively isolated from each other, as if islands within the forests, and the majority of plants cultivated in them do not occur inside the forest, the latter serves as a barrier to gene flow (Clement, 1987) (**Figure 1**). On terra firme, the soils are generally poor and need fertilizing, generally with the ashes resulting from the felling and burning of the original forest that covered them. On terra firme there are also areas that are cultivated continuously, although they are rare. These are the *terra-preta-do-índio* (Indian black earth, also called Amazonian dark earth), anthropomorphic soils that are very fertile due to having been enriched by organic detritus and charcoal deposited over countless generations by the Indians who inhabited them (Steiner *et al.*, 2007) (**Figure 7**). They usually have 2–5 ha, but 100 ha or more can be found, and owing to their fertility are preferred even today by white people for planting (Hiraoka *et al.*, 2003; Lehmann *et al.*, 2003; Smith, 1980; Steiner *et al.*, 2007).

Agriculture on periodically flooded lands is practiced on high and low várzeas and on river banks in cycles dictated by seasons of high and low water. Despite this inconvenience, these lands are highly fertile and demand little labor for clearing and planting. On the high várzeas, where flooding rarely occurs, perennial plants that can survive some time with roots submerged and annual plants are cultivated. With some risk, perennial plants typical of terra firme are also cultivated on high várzeas (Smith, 1996).



Figure 7 Anthropogenic soil of the "Indian black earth" (*terra-preta-do índio*). Its great fertility is caused by the accumulation of charcoal, ashes, wastes, and excrement of domestic animals. Many perennial and annual plants are cultivated on these soils. Andirás river, Barreirinha, AM, Brazil.

These two groups of ecosystems have caused a great division between groups of species and varieties of domesticated plants. For the plants most domesticated among Indians, such as manioc (*M. esculenta* Crantz), corn (*Z. mays* L.), and lima bean (*P. lunatus* L.), short-cycle varieties were developed capable of being cultivated on the várzeas before the floods. Although plantain banana (*M. x paradisica* L.) is an exotic and perennial species, the variety known as *pacovi* (Cavalcante, 1991), with small bunches and an extremely short productive cycle, are cultivated. The Indians also developed cultivation systems for it that are adapted to the várzea. For plants propagated by stakes, such as manioc (*M. esculenta* Crantz), when floods arrive the young plants are removed from the soil and placed in a canoe or higher site for planting when the waters recede. In the case of banana plants, marombas (wooden structures elevated above the water level) are made in which the young plants are placed with a little soil and fertilizer so that they are sufficiently grown to be planted when the waters recede. Recent discoveries have provided information about how the floodplain of the Beni river in Bolivia was cultivated by the ancient Indians. They elevated fields of up to 2 m high, known as camellones that are constructed to be above the height of floodwaters, and are surrounded by water channels. The elevated fields are somewhat like the raised beds that some vegetable gardeners construct, though on a much larger scale (from 0.7 up to 176.4 ha). During the rainy season, the surrounding channels fill with floodwater, preventing the crops in the fields being washed away. The water can then be used to irrigate and provide nutrients to the camellones during periods of drought (Erickson, 2010). There is also a similar system used by the ancient Marajoara Indians, who cultivate the "tesos" of Marajó Island, which are elevated fluvial landforms upon which to build their settlements and ceremonial centers of their communities and that eventually increased their height and size (Rossetti *et al.*, 2009) using soil from artificial lakes carved next to serve as fishponds and water reservoirs (Schaan, 2008). On those hillocks Indians could escape from the annual flood. Tesos and campos inundáveis do Marajó are considerably less fertile than várzeas close

to the Andes, the Solimões/Amazon Rivers and their tributaries of white water, and, even today, in contrast to floodplains, closer to the Andes, and it is very hard to grow corn there. Nevertheless, manioc produces well and fruits are profuse. Even today manioc flour is the mainstay in the diet of rural peoples and those living in urban centers consume fruits of palms (mainly açaí and bacaba) and fishes. Many fruits can be collected on the tesos – muruci (*Byrsonima crassifolia* Rich.), mangaba (*Hancornia speciosa* Gomes), bacaba (*Oenocarpus bacaba* Mart.), tucuman or tucum (*Astrocaryum vulgare* Mart.), inajá or najá or anajá (*Attalea maripa* Drude), caju (*Anacardium occidentale* L.), bacuri (*Platonia insignis* Mart.), araçá (*Psidium guianensis* Swartz), mucajá (*Acrocomia aculeata* Lodd), and in the lowlands açaí (*Euterpe oleracea* Mart.), miriti (*Mauritia flexuosa* L.), and taperebá (*Spondias mombin* L.) – so, apparently the Marajo Island mound habitants had a diet similar to what caboclos have today. Várzeas are reputed to have higher potential to produce food than terra firme, and were capable of having maintained a quite dense population, as reported by sixteenth-century chronicles. It is estimated that at the time of discovery, on terra firme there were 0.2 inhabitants km^{-2} and on the várzea 14.6 (Denevan, 1976) – numbers that are higher than those currently observed in Amazonia. Population estimates for the evaluated wetlands are certainly not as high if one takes into account that the foraging areas used by the Indians, in the dry season, were bigger. A clue to how extensive these foraging areas were may be given by the Tupi names of some towns of the Amazon basin beginning with a plant name followed by the suffix "tuba," "duba," "teua," "deua," "tida" and "diva" which means "a land of" or "with a profusion of" (Masucci, 1979; Barbosa, 1955), such as Inajatuba (an allusion to inajá *Attalea maripa* (C. Serra) Drude); Tucumduba, Tucumandeua and Tucumdiva all in allusion to tucum (*Astrocaryum vulgare* Mart.); Marituba an allusion to mari (*Poraqueiba paraensis* Ducke ou *P. sericea* Tul.); Mucajatuba an allusion to mucajá (*Acrocomia aculeata* Lodd), Miriti tuba and Miriti teua allusion to miriti (*Mauritia flexuosa* L.), Timboteua (an allusion to timbó (*Lonchocarpus utilis* Smith), Jabatiteua (possible a corruption of abati, that means maize, *Zea maiz* L.); Ananindeua (an allusion to anani = ananin, *Symphonia globulifera* L.); Caranandeua (an allusion to caranã (*Mauritia carana* Wallace), Açaiteua and Açaitiva an allusion to açaí (*Euterpe oleracea* Mart.); Ajuruteua (in allusion to ajuru (*Chrysobalanus icaco* L., Caratateua in allusion to cará (*Dioscorea trifida* L.), Jutaitiva (in allusion to *Hymenaea courbaril* L.), Araparitiva (in allusion to *Macrolobium acaciaefolium* Benth).

The Management System

The management system causes environmental changes that can influence the demography and genetic structure of a population. For there to be selection and consequent domestication, there must be differentiated reproduction and survival for the genotypes selected; therefore, management and domestication are two intimately linked factors. Forms of management generally influence domestication, and it is through them that explaining this process among Amazonian plants is intended. Generally, the more elaborate the management demands of a plant, the greater its degree of domestication. In the case of Amazonian plants, depending on the degree of management to which they may be submitted, they may be classified as (1) collected, (2) protected or planted with little human interference, and (3) planted with high human interference. There is also another distinct group – those managed unconsciously due to interference in the environment by humans for other purposes or due to dispersion of seeds through feces and/or discarding of nonedible parts.

Plants that are simply collected are left in their natural habitat without selection for any characteristic, and even if the fruits collected from more productive and better plants should be more disseminated, there is no reproductive isolation and presumably the population maintains itself stable. The word "presumably" is used because no matter how subtle, human occupation is capable of introducing some modification in the natural populations of plants. The reduction, through hunting, of frugivorous animals such as agoutis (*Dasyprocta leporina* L., *D. prymnolopha* Wagler, and *D. fuliginosa* Wagler), pacas (*Agouti paca* L.), and bats, which prey on and disperse seeds, may favor the increase in density of a certain species where the more productive plants might have greater chances of increasing the frequency of their alleles in a given location due to the increase in their density. The group of useful plants collected in the forest for medicinal purposes, food, or some other use may possibly comprise more than 1000 species.

The plants protected or planted with low human interference are those that occur naturally in the habitat and that are preserved during felling and later cared for or planted along trails or in clearings but that receive little care (**Table 1**). Examples of plants protected from burning include the tucumã (*Astrocarium aculeatum* G.F.W. Meyer and *A. vulgare* Mart.) and inajá (*Attalea maripa* (C. Serra) Drude) palms. Plants under low human interference are those generally cast along trails and in clearings, as is done by the Kayapó Indians, who collect seeds and plantules of useful trees and plant them along trails on which they usually travel or in clearings opened by tree falls (Kerr and Posey, 1984). They generally belong to the same habitat and although they undergo some selection, reproductive isolation, as in the previous case, is not efficient because these species also occur in forests and can interbreed. Occasional progress in the domestication process may be achieved because the selected plants, when planted together, have greater chances of intercrossing, which facilitates the maintenance of a higher frequency of the alleles selected. Since isolation is not very efficient, progress is very slow. This group is prone to receive greater human influence than is the case with the previous group because their density and fecundity are in most cases favored by the reduction of competition with the other plants. The reproductive isolation of these species, although more efficient than in the previous group, is also not very efficient and there is no directional selection of the plants; they are simply spared.

The dispersion of those planted with high human interference (**Table 2**) is consciously done by humans and during their entire productive cycle they receive care and make up three distinct groups, according to the position they assume in the ecological succession of forest regeneration after burning. This succession is directed by humans toward a final floristic composition with a greater density of useful species than would be found in a forest without any history of human

Table 1 Protected or planted plants atlow levels of human care[a]

Names: English/Portuguese/Spanish	*Latin name*	*Uses*
—/*caruru azedo*/—	*Amaranthus cruentus* L.	Leaves as vegetable
Star nut palm/*tucumã do Amazonas, tucumã- açu/chambira*	*Astrocaryum aculeatum* G. F. W. Meyer	Fruit, fibers from leaves
—/*tucumã-do-pará/aouará, awarra*	*Astrocaryum vulgare* Mart.	Fruit, fibers from leaves
Brazil nut/*castanha-do-pará/nuez del Brasil*	*Bertholletia excelsa* Humb. Et Bonpl.	Nuts, infusion of fruits as a medicinal
Wild cherry/*muruci verdadeiro/Indano, nanci, yoco*	*Byrsonima crassifolia* L. Rich.	Fruit
Pekea/*piquiá verdadeiro/amêndoa-do peru, al-mendro*	*Caryocar villosum* (Aubl.) Pers.	Fruit, and oil from the pulp
—/*castanha-de-porco*/—	*Caryodendron amazonicum* Ducke	Seed as a nut
—/—/—	*Celosia argentea* L.	Leaves as vegetable
—/*castanha-de-cutia*/—	*Couepia edulis* Prance	Fruit
—/*pajurá*/—	*Couepia bracteosa* Benth.	Seeds as a nut
—/*castanah pêndula*/—	*Couepia longipendula* Pilger	Seeds as a nut
—/*sorva, sorvinha*/—	*Couma utilis* Muell. Argov.	Fruit
American oil palm/*caiaué/noli, palma brasilenha, carocito, colorada, palmiche*	*Elaeis oleifera* Cortés	Oil from the pulp and endosperm, seed kernel for crafts
—/*uxi*/—	*Endopleura uchi* Cuatr.	Fruit
Wild coriander/*chicória, coentro do pará/siuca culantro*	*Eryngium foetidum* L.	Condiment
Assai palm/*açaí/chonta, husaí*	*Euterpe oleracea* Mart.	Fruit, trunk for building
—/*mangaba/mango jsú*	*Hancornia speciosa* Gomes	Fruit
Copal/*jatobá/copalhuallo*	*Hymenaea courbaril* L.	Fruit
—/*ingá-açu*/—	*Inga cinnamomea* Benth.	Fruit
Ice cream beam/*ingá-cipó/guamo, guabo*	*Inga edulis* Mart.	Fruit
Pataua/*patauá/batauá, chapil, palma sege*	*Oenocarpus bataua* (Mart.) Burr.	Fruit
Barbasco/*timbó/barbasco*	*Lonchocarpus utilis* Smith	Poison for fishes
—/*maçaranduba, maçaranduba verdadeira*/—	*Manilkara huberi* (Huber) Standl.	Fruit, edible latex
Mauritia, *moriche/buriti, miriti/aguaje*	*Mauritia flexuosa* L.	Fruit, sap as sweet sirup, starch from the trunk
—/*bacaba açu, bacaba verdadeira/ungurauy, camou, manoco*	*Oenocarpus bacaba* Mart.	Fruit, trunk for building, leaves to cover houses
—/*bacaba-de-leque*/—	*Oenocarpus disticus* Mart.	Fruit, trunk for building, leaves to cover houses
—/*bacabi/mapora*	*Oenocarpus mapora* Karsten	Fruit
—/*bacabinha*/—	*Oenocarpus minor* Mart.	Fruit
—/*feijão macuco*/—	*Pachyrrhizum tuberosum* Lam.	Tubers as vegetable
—/*pajurá-da-mata*/—	*Parinari montana* Aubl.	Fruit
Wild passion fruit/*maracujá suspiro, maracujá-de-rato, maracujá-do-mato*/—	*Passiflora nitida* H. B. K.	Fruit
—/*camapu*/—	*Physalis angulata* L.	Fruit as a vegetable, roots as medicine
Backuri/*bacuri/bacuri, matazona*	*Platonia insignis* Mart.	Fruit
—/*umari, mari gordo*/—	*Poraqueiba paraensis* Ducke	Fruit
—/*umari, umari do amazonas*/—	*Poraqueiba sericea* Tul.	Fruit
Canistel/*cutite/caimo, canistel, yema de huevo*	*Pouteria macrophylla* (Lam.) Eyma	Fruit
—/*ucuqui*/—	*Pouteria ucuqui* Pires et Schultes	Fruit
—/*araça-pera*/—	*Psidium acutangulum* DC	Fruit
Guavas/*goiaba/guayaba*	*Psidium guajava* L.	Fruit, medicinal tea from young leaves for diarrhea
—/*araçaí, araça-do-campo*/—	*Psidium guinensis* Swartz	Fruit
—/*bacuripari liso*/—	*Rheedia brasiliensis* (Mart.) Pl. et Tr.	Fruit
—/*bacuripari/pacuriguazu*	*Rheedia macrophylla* Pl. et Tr.	Fruit
Hog plum, plum bush/*taperebá, cajamirim, cajá/mombim*	*Spondias mombin* L.	Fruit, infusion from the bark for leucorrohea
—/*pitomba/carayá-vola*	*Talisia esculenta* (St. Hil.) Radlk.	Fruit
Cocoa, chocolate tree/*cacau/cacao*	*Theobroma cacao* L.	Fruit
Cupuassu/*cupuaçu, cupu, cupuácu verdadeiro/copoazur*	*Theobroma grandiflorum* (Willd. ex Spreng.) Schum	Fruit
—/*cacauí/cacau sacha, chocalatillo*	*Theobroma speciosum* Willd.	Fruit

[a]They receive some influence on the genetic structure of their population. The domestication process is accepted as being in progress but at a very early stage. Some of these species were possibly undergoing a more advanced process of domestication, but a possible retrocession on their status is accepted due to loss of varieties or by the people who used them.

Table 2 Cultivated plants under a high level of human assistance[a]

Names: English/Portuguese/Spanish	*Latin name*	*Uses*
Cashew/*caju/marañón*	*Anacardium occidentale* L.	Fruit, infusion from the bark for leucorrohea
Pineapple/*abacaxi, ananas/piña*	*Anannas comosus* (L.) Merril.	Fruit, fiber from leaves
Soursop/*graviola/guanábana, anona*	*Annona muricata* L.	Fruit, infusion from the leaves for snake bite
Peanut/*amendoin/mani*	*Arachis hypogeae* L.	Seeds
Peach palm/*pupunha/pejibaye, pijuayo, chontaduro*	*Bactris gasipaes* H. B. K.	Fruit, trunk for building
Annatto/*urucum/achiote*	*Bixa orellana* L.	Arielle from seeds as colorant and condiment, largely used for painting the body
Guinea arrow root/*ariá/dale dale*	*Calathea allouia* (Aubl.) Lindl.	Tubercle
—/—/—	*Canna edulis* Ker.	Tubercle
Sweet pepper/*pimenta-doce/aji*	*Capsicum anumm* L. var. minimum	Fruit as a vegetable, condiment, medicinal
—/—/—	*Capsicum chinense* Jacq.	Condiment
—/*pimenta-de-cheiro*/—	C. brasilianum	Condiment
—/*pimenta comari*/—	*C. baccatum* L.	Condiment
Papaya, paw paw/*mamão/lechosa, melon zapote*	*Carica papaya* L.	Fruit
—/*cupa*/—	*Cissus gongyloides* Burch	Liana as vegetable
Calabash gourd/*cuieira/pati*	*Crescentia cujete* L.	Fruit domestic appliances
Yam/*cará/sacha-papa*	*Dioscorea trifida* L.	Tubercle
—/*jaboti*/—	*Duguetia stenantha* R. E. Fries	Fruit
—/*araça-boi/arazá, araza buey*	*Eugenia stipitata* McVaugh	Fruit
Coca/Amazon coca/*ipadu/epadu*	*Erythroxylum coca* Lam. var. *ipadu* Plowman	Leaves, stimulant, hunger and fatigue relief
Genipa/*genipapo/huito*	*Genipara americana* L.	Fruit
Cotton/*algodão brabo/algodón*	*Gossipium barbadense* L.	Fiber
Sweet potato/*batata doce/camote*	*Ipomoea batatas* (L.) Lam.	Tubercle
Manioc/*mandioca/yucca*	*Manihot esclenta* Crantz	Tubercle, leaves as vegetable
Arrowroot/*araruta*/—	*Marantha arundinaceae* L.	Tubercle
Tobacco/*tabaco/tabaco*	*Nicotiana tabacum* L.	Leaves that are smoked or smashed as stimulant
Passion fruit/*maracujá, maracujá, verdadeiro/ granadilla*	*Passiflora edulis* Sims. f. flavicarpa Deg.	Fruit for a refrigerant and as a carminative
Giant granadilla/*maracujá-acú, maracujá mamão/granadilla, granadilla real*	*Passiflora quadrangularis* L.	Fruit
Guarana/*guaraná/cupana*	*Paullinea cupana* H. B. K. var sorbilis (Mart.) Ducke and var. typica	Seeds for infusion used as refrigerant and stimulant
Lima beam/*feijão lima*/—	*Phaseolus lunatus* L.	Beans
—/*cucura, mapati/uvilla*	*Pourouma cecropiaefolia* Mart.	Fruit
—/*abiu/caimito*	*Pouteria caimito* (Ruiz et Parvon) Radlk.	Fruit
Sapote/*sapota-do-peru/zapote, mame colorado*	*Quararibea cordata* (H. B. K.) Vischer	Fruit
—/*biribá*/—	*Rollinia mucosa* (Jacq.) Baill.	Fruit
—/*cubiu/cocona*	*Solanum sessiliflorum* Dunal.	Fruit as a vegetable
Para cress/*jambu/jambu*	*Spilanthes acmella* (L.) Murr.	Leaves as vegetable
—/*cacau-do-peru/macambo*	*Theobroma bicolor* Humb. & Bompl.	Fruit, toasted seeds as a nut
Cocoyam/*taioba/uncucha, malanga*	*Xanthosoma sagittifolium* (L.) Schott.	Tubers and leaves as a vegetable
Corn, maize/*milho/mayz*	*Z. mays* L.	Seed as a grain, estiles as medicinal

[a]The domestication process is accepted as being in an advanced stage.

influence (Denevan *et al.*, 1984; Anderson and Posey, 1989; Balée, 1989). The first group is that of the most domesticated species: These are annual plants that are colonizers and fast growers, with a high productivity of edible biomass. They are manioc (*M. esculenta* Crantz), corn (*Z. mayz* L.), and more recently, beans of the genus *Vigna*. The second group is formed by semiperennial species, sown after the plants in group 1: sweet potato (*Ipomoea batatas* (L.) Lam.), arrowroot (*Marantha arundinaceae* L.), peanut (*Arachis hypogaea* L.), lima bean (*P. lunatus* L.), cará yam (*D. trifida* L.), and pineapple (*Anannas comosus* (L.) Merril). Those of the third group are arboreal species, generally fruit producers with rapid growth and the capacity for growing well in soils already leached by rains and exhausted by annual crops. These plants have a high capacity for production and deposition of their leaf biomass. They usually live for 20 years, but some of them live for 80 years, and cannot bear competition with other trees of the forest. They are peach palm (*Bactris gasipaes* H.B.K.), cucura (*Pourouma cecropiaefolia* Mart.), umari do Amazonas (*Poraqueiba sericea* Tul.), and biribá (*Rollinia mucosa* (Jacq.) Baill). Also in this group are other trees that do not have a great capacity for regenerating the soil but are still able to grow well in poor

soils: Guava (*Psidium guajava* L.), coca (*Erythroxylum coca* Lam. var. *ipadu* Plowman), and abiu (*Pouteria caimito* (Ruiz et Parvon) Radlk). After the planting of these categories, there is the reappearance of the group of plants protected/planted under low human interference. These make up the group in the final phase of regeneration in the direction of a climax forest. Sometimes, species of this group are planted, whereas others are only protected and require more specific ecological conditions, with some shading in the juvenile phase, humidity, and protection from heat. These plants are capable of competing with other trees of the forest and are part of the final climax stand achieved by the system; they are Brazil nut (*Bertholletia excelsa* H & B) and the piquiazeiros (*Caryocar villosum* Pers.), amapá doce (*Brosimum potabile* Ducke), maçaranduba (*Manilkara huberi* (*Huber*) Standl.), bacuri (*P. insignis* Mart.), cacao (*Theobroma cacao* L.), cacauí (*Theobroma speciosum* Willd.), and uxi (*Endopleura uchi* Cuatr.). Because trees of the climax forest and those of this group have the chance to make gene exchanges, diluting eventual gains in selection, they cannot be considered domesticated. It is curious to note that young Brazil nut trees are only seen in disturbed environments, but that in the climax forests they are frequent, which reinforces the hypothesis that the Brazil nut forests, such as those observed in the regions of southern Pará, Acre, and Alenquer, had once been various Indian agricultural plots. These useful forest trees form the so-called maçaranduba islands, Brazil nut islands, amapá-doce islands, bacuri islands, cacao islands, cacauí islands, uxi islands, and so on. When they begin to produce fruits, these islands begin to receive better care. The plants below the canopy are weeded to facilitate fruit collection and observation of game that also come searching for fruits. The Nambiquara Indians of the Hahaintesu have a curious system for transferring information on the localization of these "islands" from one generation to the next until they are ready to be exploited. The sites where they live are eventually abandoned, generally due to over-hunting or disease. The new site chosen is generally an old site that had been used because it had a water supply, offered some strategic defense position, and had trails enriched with useful plants. The former site has a cemetery. The tribe is divided into people: The "people of the curassow," "the jaguar," "the tapir," etc., and the individuals of each "people" have a specific cemetery where they should be buried. When Indians die, their companions take them to be buried in the cemetery of their respective people and thus they are able to visit the old sites, accompany the development of the "islands," and maintain their localization in tribal memory.

The groves of babaçu palms (*Orbignya martiana* Barb. Rod.), which occupy ~200,000 km^2 in Brazil and support ~450,000 subsistence-level households in the north, northeast, and central west regions of Brazil through their multiple products (fibers, thatch, basketry, construction, palm heart and fruits for oil, charcoal, and animal feed) are examples of unconscious management. This palm tree regenerates profusely when the primary forest is burned. The plant survives fire well because it has a unique germination process. Upon germinating, the seed sends its hypocotyl underground to a depth of about 10 cm and then growth of the gemule begins. It remains inside the soil, putting its leaves out above ground. When the forest is burned, the plants that are still in this phase lose only their leaves but are capable of regenerating since the main bud is well protected from the fire. Owing to the reduction in competition with the other plants that were destroyed by fire, the babaçu plants end up dominating the environment (May *et al.*, 1985a, 1985b). Another example of unconscious management is the sowing of plants collected near the huts in piles of trash (**Figure 8**). This process is possibly what happened at the beginning of agriculture and the related sedentarism (Harlan, 1975). The plants thus unconsciously cultivated end up being cared for and recruited within an agricultural system.



Figure 8 Nambiquara Indians of the Alantesu group enjoying fruits (not identified) recently collected from a fruiting plant in the savanna and scattering the seeds near a temporary hut built for support on one of their collecting journeys.

All these systems of management and consequent selection have been applied since before the discovery to the present by the indigenous people, later by the caboclos, and recently by Japanese immigrants to Amazonia (Saragoussi *et al.*, 1988). This has maintained the preservation of already domesticated plants and their varieties and even the appearance or improvement of others. The Japanese farmers of Tomé-Açú and Castanhal and their descendants, using only massal selection, have been able to achieve better cultivars of cupuaçu (*T. grandiflorum* Schum.) and açaí (*Euterpe oleracea* Mart.). Virola softwood from the estuary islands of the Amazon River, which in the past was predated as a source of extractivism, today is wisely managed as a consequence of market forces, and thus is also undergoing domestication by caboclos.

Final Comments

A domesticated plant, it should be remembered, is in this state because humans brought about its domestication. If human-induced selection pressures cease, these plants will not survive, or their descendants will need to recompose their defenses to face competition in a natural environment and in doing so generally lose the attributes selected by humans (Harlan, 1975). Many of the indigenous cultivars were selected not only for flavor or productivity but also for mystical reasons and are thus susceptible to disappearance as a result of changes in habits. In Amazonia, the Indians are being besieged by pressures toward deculturation from the "civilized"

people who are occupying the region. In this process, much of the existing knowledge on plant use is lost. For medicinal plants, this loss is even greater since most of the time their usefulness is not visible, as in the case of fruiting species. With the gradual loss by the Indians of the properties and uses of the plants, everything may simply be considered "jungle." Although Amazonia is considered one of the largest repositories of genetic resources in the world, there are still few efforts in recovery, collection, and evaluation of useful plants among the Indians. If humanity has these marvelous plants, delights in their flavor and their appearance, wears clothes and uses domestic utensils made from them, and uses them to cure pains and diseases, this is because the traditional peoples of Amazonia, very simple people, discovered them, selected them, and maintained them until now. Fortunately their rights were recognized by the Convention on Biological Biodiversity – The Nagoya Protocol on Access and Benefit-Sharing (Secretariat of the Convention on Biological Diversity, 2011), but the way to put it in practice is yet to be found.

See also: Agriculture, Traditional. Biodiversity in Plant Breeding. Domestication of Crop Plants. Edible Plants. Genetic Diversity. Indigenous Peoples and Biodiversity

References

Anderson A and Posey DA (1989) Management of tropical scrub savanna by the Gorotire Kayapó of Brazil. *Advance Economic Botany* 7: 159–173.

Atroch AL, Nascimento-Filho FJ, Ângelo CSP, Freitas DV, and Clement NRCR (2009) Domesticação e melhoramento do guaranazeiro. In: Borém A, Gomes L, Maria T, and Clement CR (eds.) *Domesticação e Melhoramento: Espécies Amazônicas*, pp. 337–365. Viçosa, MG: Editora da Univ. Fed. Viçosa.

Balée W (1989) The culture of Amazonian forest. *Advance Economic Botany* 7: 1–21.

Barbosa AL (1955) *Pequeno vocabulário Tupi-português*. Rio de Janeiro: Livraria São José.

Barbosa RI, Luz FJF, Nascimento-Filho HR, and Maduro CB (2002) Pimentas do gênero Capsicum cultivadas em Roraima, Amazônia brasileira. I. Espécies Domesticadas. *Acta Amazonica* 32(2): 177–192.

Barcelos E, da Cunha RNV, Nouy B, and Sousa NR (1999) Uso de marcadores moleculares RFLP e AFLP no estudo da diversidade genética do dendezeiro (*Elaeis guineesis* Jacq. *E E. oleifera* (Kunth) Cortés). *Genetics and Molecular Biology* 22(3): 289.

Benza JC (1980) *143 Frutales Nativos*. Lima: University Agraria La Molina.

Black FL, Pandey JP, and Santos SEB (1991) Evidências baseadas em HLA e IgG sobre as relações intra e intercontinentais das populações nativas da Amazônia. In: Neves WA (ed.) *Origens, Adaptações e Diversidade Biológica do Homem Nativo da Amazônia*, pp. 55–83. Belém: Museu Paraense Emílio Goeldi (coleção Emilie Snethlage).

Boster JS (1984) Classification, cultivation, and selection of Aguaruna cultivars of *Manihot esculenta* (Euphorbiaceae). In: Prance GT and Kallunki JA (eds.) *Ethnobotany in the Neotropics*, pp. 34–47. New York: New York Botanical Gardens.

Cavalcante PB (1991) *Frutas Comestíveis da Amazônia*, 5th edn. Belém: CEJUP/ Museu Paraense Emílio Goeldi.

Clement CR (1987) A pupunha, uma árvore domesticada. *Ciência Hoje* 5(29): 42–49. (The pejibaye, a domesticated tree. *Ciência Hoje* (Special ed.: Amazonia), 43–47, 1991).

Clement CR (1988) Domestication of pejibaye (*Bactris gasipaes* H. B. K.): Past and present,. *Economic Botany* 6: 155–174.

Clement CR (1989) A center of crop genetic diversity in western Amazonia. *BioScience* 39(9): 624–631.

Clement CR (1992a) Domesticated palms. *Principes* 36(2): 70–78.

Clement CR (1992b) Frutas da Amazônia. *Ciência Hoje* 14(83): 28–37.

Clement CR (1999) 1492 and the loss of Amazonian crop genetic resources. I. The relation between domestication and human population decline. *Economic Botany* 53(2): 188–202.

Clement CR and Mora Urpi J (1987) The pejibaye (*Bactris gasipaes* H. B. K., Arecaceae): Multi-used potential for lowland humid tropics. *Journal Economic Botany* 41: 302–311.

Clement CR, Müller CH, and Flores WC (1982) Recursos genéticos de espécies frutíferas nativas da amazônia brasileira. *Acta Amazonica* 12(4): 677–695.

Correa P (1926–1969) *Dicionário das Plantas Úteis do Brasíl e das Plantas Exóticas Cultivadas* 4 Vols. Rio de Janeiro: IBDF.

Denevan WM (ed.) (1976) *The Native Population of the Americas in 1492*. Madison: University of Wisconsin Press.

Denevan WM, Treacy JM, Alcorn JB, Padoch C, Denslow J, and Paitan SF (1984) Indigenous agroforestry in the Peruvian Amazon: Bora Indian management of swidden fallows. *Interciência* 9(6): 346–357.

Elias M and McKey D (2000) The unmanaged reproductive ecology of domesticated plants in traditional agroecosystems: An example involving cassava and a call for data. *Acta Oecologica-International Journal of Ecology* 21: 223–230.

Elias M, Rival L, and McKey D (2000) Perception and management of cassava (*Manihot esculenta* Crantz) diversity among Makushi Amerindians of Guyana (South America). *Journal of Ethnobiology* 20(2): 239–265.

Erickson CL (2010) The transformation of environment into landscape: The historical ecology of monumental earthwork construction in the Bolivian Amazon. *Diversity* 2: 618–652.

Freitas DV, Carvalho CR, Nascimento-Filho FS, and Astolfi-Filho S (2007) Karyotype with 210 chromosomes in guaraná (*Paullinia cupana*) 'Sorbilis'. *Journal of plant research* 120: 399–404.

Harlan JR (1975) *Crops and Man*. Madison, WI: American Society of Agronomy/ Crop Science Society of Marica.

Harlan JR, de Wet JMJ, and Price EG (1973) Comparative evolution of cereals. *Evolution* 27: 311–325.

Heckenberger MJ, Russell JC, Fausto C, Toney JR, Schmidt MJ, Pereira E, Franchetto B, and Kuikuro Afukaka (2008) Pre-Columbian urbanism, unthropogenic landscapes, and the future of the Amazon. *Science* 29(321): 1214–1217. n. 5893.

Hill J and Moran E (1983) Adaptative strategies of Wakuenai people in the oligotrophic rain forest of the Rio Negro basin. In: Hames R and Vickers W (eds.) *Adaptative Responses of Native Amazonians*, pp. 113–135. New York: Academic Press.

Hiraoka M, Yamamoto S, Matsumoto E, Nakamura S, Falesi IC, and Baena ARC (2003) Contemporary use and management of Amazonian dark earths. In: Lehmann J, Kern DC, Glaser B, and Woods WI (eds.) *Amazonian Dark Earths: Origins, Properties, Management*, pp. 387–406. The Netherlands: Kluwer Academic Publishers.

Hoehne FC (1946) *Phytophysionomia do Estado do Mato-Grosso e Ligeiras Notas a Respeito da Composição e Distribuição da Sua Flora*. São Paulo: Estudo Preliminar, Melhoramentos.

Iltis HHL (1986) Maize evolution and agricultural origins. *International Grass Symposium* 195–220.

Instituto Nacional de Pesquisas da Amazônia/Núcleo do Acre (1984). Levantamento de informações sobre a ocorrncia de frutas nativas no interior do estado do Acre.

Jordan CF (1985) Soils of the Amazon rainforest. In: Prance GT and Lovejoy TE (eds.) *Key Environments: Amazonia*, pp. 83–94. New York: Pergamon.

Kerr WE and Clement CR (1980) Práticas agrícolas com conseqüências genéticas que permitiram os índios da Amazônia uma melhor adaptação as condições regionais. *Acta Amazonica* 10(2): 156–159.

Kerr WE and Posey AD (1984) Informações adicionais sobre a agricultura Kayapó. *Interciência* 9(6): 392–400.

Le Coint P (1947) *Amazônia Brasileira III. Árvores e Plantas Úteis (Indígenas e Aclimatadas)*, 2nd edn. São Paulo: Nacional.

Lehmann J, da Silva Jr. JP, Steiner C, Nehls T, Zech W, and Glaser B (2003) Nutrient availability and leaching in an archaeological Anthrosol and a Ferralsol of the Central Amazon basin: Fertilizer, manure and charcoal amendments. *Plant and Soil* 249: 343–357.

Lima RR and da Costa JPC (1997) *Coleta de Plantas de Cultura Pré-Colombiana na Amazônia Brasileira: Metodologia e Especificações Realizadas Para Coleta de Germoplasma*. Belém: EMBRAPA/CPATU, (Documentos, 99 v.1). 148 pp.

Lukesch A (1976) *Mito e Vida dos Caiapós*. de São Paulo: Editora Pioneira de Estudos Brasileiros/Editora da Univ.

Masucci O (1979) *Dicionário Tupi Português e Vice-versa*. São Paulo: Brasilivros.

May PH, Anderson A, Frazão JMF, and Balick MJ (1985a) Babassu palm in agroforestry systems in Brazil's mid-north region. *Agroforestry Systems* 3: 275–295.

May PH, Anderson AB, Balick MJ, and Frazão JMF (1985b) Subsistence benefits from babassu palm (*Orbignya martiniana*). *Economic Botany* 39(2): 113–129.
Pärssinen M, Schaan D, and Ranzi A (2009) Pre-Columbian geometric earthworks in the upper Purús: A complex society in western Amazonia. *Antiquity* 83: 1084–1095. n. 322.
Picanço DB, Sousa NR, Clement CR, and Nagao EO (1999) Discriminaçao de raças primitivas de pupunha (*Bactris gasipaes*) na Amazônia Brasileira com Marcadores Moleculares (RAPDs). *Genetics and Molecular Biology* 22(3): 290.
Pires JM and Prance GT (1985) The vegetation types of Brazilian Amazon. In: Prance GT and Lovejoy TE (eds.) *Key Environments: Amazonia*, pp. 109–145. New York: Pergamon.
Pujol B, Gigot G, Pinheiro-Kluppel M, Elias M, Hossaert- McKey M, and McKey D (2002) Germination ecology of cassava (*Manihot esculenta* Crantz, Euphorbiaceae) in traditional agroecosystems: Seed and seedling biology of a vegetatively propagated domesticated plant. *Economic Botany* 56(4): 366–379.
Roosevelt AR (1991) *Moundbuilders of Amazon: Geophysical Archeology on Marajo Island*. Brazil, San Diego: Academic Press.
Rossetti D de F, Góes AM, and de Toledo PM (2009) Archaeological mounds in Marajó Island in Northern Brazil: A geological perspective integrating remote sensing and sedimentology. *Geoarchaeology: An International Journal* 24(1): 22–41.
Rostain S (2010) Pre-Columbian earthworks in coastal Amazonia. *Diversity* (2): 331–352.
Salati E (1985) The climatology and hydrology of Amazonia. In: Prance GT and Lovejoy TE (eds.) *Key Environments: Amazonia*, pp. 18–48. New York: Pergamon.
Sambatti JBM, Martins PS, and Ando A (2000) Distribuição da diversidade isoenzimática e morfológica da mandioca na agricultura autóctone de Ubatuba. *Scientia Agricola* 1(57): 75–80.
Saragoussi M, Martel JH, and Ribeiro GA (1988) Comparação na composição de quintais de três localidades de terra firme do estado do Amazonas. In: Posey D and Overal WL (eds.) *Ethnobiology: Implications and Applications. Proceedings of the First International Congress of Ethnobiology*, pp. 295–304. Brasil: PA Belém.
Saunaluoma S (2010) Pre-Columbian earthworks in the Riberalta region of the Bolivian Amazon. *Amazônica* 2(1): 86–115.
Schaan D, Ranzi A, Pärssinen M and (Orgs.) (2008) *Arqueologia da Amazônia Ocidental: Os geoglifos do Acre*. Belém: Editora Universitária UFPA. 192 p.: il.
Schaan DP (2008) The nonagricultural chiefdoms of Marajó Island. In: Silverman H and Isbell WH (eds.) *Handbook of South American Archaeology*, pp. 339–357. New York: Springer.
Secretariat of the Convention on Biological Diversity (2011) *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity: Text and Annex*. Canada: Montreal.
Setz EZ (1991) Animals in the Nambiquara diet: Methods of collection and processing. *J. Ethnobiol.* 11: 1–22.
Silva-Filho DF, Noda H, and Clement CR (1993) Genetic variability of economic characters in 30 accessions of Cubiu (*Solanum sessiliflorum* Dunal) evaluated in central Amazonia. *Revista Brasileira Genét* 16(2): 409–417.
Silva-Filho DF, Noda H, Clement CR, and Machado FM (1999) Cubiu (*Solanum sessiliflorum*): Conservação e melhoramento de etnovariedades na Amazônia. *Genetics and Molecular Biology* 22(3): 708.
Smith NJH (1980) Anthrosols and human carrying capacity in Amazonia. *Annals of the Association of American Geographers* 70: 553–566.
Smith NJH (1996) Home gardens as a springboard for agroforestry development in Amazonia. *International Tree Crops Journal* 9: 11–30.
Smith NJH, Williams JT, Plucknett DL, and Talbot J (1992) *Tropical Forests and their Crops*. Ithaca, NY: Comstock/Cornell University Press.
Steiner C, Teixeira WG, Lehmann J, Nehls T, de Macêdo JLV, Blum WEH, and Zech W (2007) Long term efects of manure, charcoal and mineral fertilization on crop production and fertility on a highly weathered Central Amazonian upland soil. *Plant and Soil* http://dx.doi.org/10.1007/s11104-007-9193-9.
Venturieri GA and Mendonça ML (1985) Cupuaçu sem semente: Histórico do aparecimento da cultivar. *Informativo da Sociedade Brasileira de Fruticultura* 4(4): 12–13.
Venturieri GA, Rodrigues WA, and Menezes JMT (1986/1987) Salikisu e Kwalhakatasu, duas fruteiras dos Índios Nambiquaras com potencial para a domesticação. *Acta Amazonica* 16/17: 1925.

Indirect Land Use and Greenhouse Gas Impacts of Biofuels

Richard J Plevin and Daniel M Kammen, University of California, Berkeley, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Direct land-use change (DLUC) Conversion of land from a noncrop land use to the production of bioenergy crops.

General equilibrium (GE) model Similar to a partial equilibrium (PE) model, but representing all economic sectors, generally in a highly aggregated form. GE models have greater scope but less detail (fewer products in each sector and less detail on each product) than do PE models.

Indirect land-use change (ILUC) Conversion of land from one land-use category to another, induced by the expansion of biofuel production elsewhere.

Life cycle assessment An analytic framework to assess the full environmental impacts of a product or service, from raw material extraction, through production, use, and disposal.

PE model A model of one or a few sectors of the economy that computes an equilibrium state in which supply and demand for each product equilibrate in response to some exogenously defined change, such as a tax, mandate, or loss of supply.

Overview

Biofuels can be produced from virtually any form of biomass, including purpose-grown energy crops, crop and timber residues, and wastes. Commonly used or proposed energy crops include corn, sugarcane, soybeans, rape, wheat, palm, beets, switchgrass, miscanthus, pine, and willow. Just as each crop has different energy content and life cycle consumption of fossil fuel, fertilizers, and other inputs, each crop has a different impact on the overall commodity markets for food, feed, and fiber.

When bioenergy crops are produced on productive cropland, they can displace the production of food, feed, and fiber, driving up the price of the displaced commodities. In response to the price increases, some consumers will switch to lower-priced substitutes and others will simply consume less, while farmers elsewhere will respond by replacing production of the displaced crops. Increasing crop production requires either increasing yields on existing land (intensification) or farming more land (extensification). Intensification involves the use of more inputs, such as fertilizer or water; increasing crop or livestock density; or employing better technology, such as improved genetic varieties. Extensification requires bringing additional land into crop production. If high-carbon value land, such as forest and grassland, is cleared to accommodate the additional production, enough CO_2 and other greenhouse gases (GHGs) may be released from the disturbed soil and biomass to negate the climate benefits of displacing petroleum-based fuels with biofuels. Whether and how much indirect land-use change (ILUC) occurs depends on the relative scale of each of these possible responses.

The production of biofuels from feedstocks that do not compete for land with food, feed, and fiber generally does not induce ILUC. These feedstocks include crop and forestry residues, municipal waste, and crops grown on the so-called "marginal" land that is incapable of supporting economically viable crop production. To the extent that harvesting residues reduces yield, it may result in some land clearing or additional fertilizer use. Although this land clearing and fertilizer use increases GHG emissions, the effect should be much smaller than that induced by purpose-grown energy crops.

Direct and Indirect Land-Use Change

In the context of biofuels, direct land-use change (DLUC) refers to the conversion of land from some other land-use category to the production of bioenergy crops. DLUC can provide environmental costs or benefits. For example, replacing row crops with perennial grasses will often increase soil carbon sequestration, reduce nutrient and pesticide runoff, and improve biodiversity (Davis *et al.*, 2011).

When crops replace pasture or forest, DLUC can result in substantial GHG emissions. For example, Fargione *et al.* (2008) estimate that converting lowland tropical rainforest in Indonesia and Malaysia to palm biodiesel would result in the release of about 700 Mg ha^{-1} of CO_2 for more than 50 years, resulting in a "carbon debt" that would take almost 90 years to repay through substitution of biodiesel for petroleum diesel. Only after this carbon debt is balanced would the biofuel serve as an instrument of carbon reduction. Similarly, these authors estimate that draining tropical peatland rainforest to produce palm would release about 3450 Mg ha^{-1} of CO_2, requiring more than 400 years to repay through fuel substitution. In contrast, production of biofuels from prairie biomass grown on unused or marginal cropland incurs essentially no carbon debt.

In contrast to DLUC, ILUC occurs when bioenergy crops displace other crops, triggering the conversion to cropland of lands, somewhere on the globe, to replace some portion of the displaced crops (Searchinger *et al.*, 2008; Hertel *et al.*, 2010; Al-Riffai *et al.*, 2010). This result occurs because agricultural land is a constrained resource and the demand for food and feed is somewhat insensitive to changes in price (Kløverpris *et al.*, 2008). ILUC is a market-mediated phenomenon: The effects are transmitted through global markets linked by commodity substitutability and the competition for land (Laurance, 2007). Owing to the challenges of modeling global economic behavior, the location and magnitude of

ILUC, and thus the GHG emissions induced by crop-based biofuels, are highly uncertain (Plevin *et al.*, 2010).

Although ILUC emissions are uncertain and likely to remain so due to the assumptions that drive different models, understanding ILUC is critical to understanding whether biofuel production increases or decreases GHG emissions relative to using petroleum-based fuels (Edwards *et al.*, 2010). In most life cycle analyses of the GHG emissions from biofuels, the emission of CO_2 from the combustion of biogenic carbon is considered "carbon neutral" because the biomass was formed from atmospheric CO_2 during photosynthesis (Rabl *et al.*, 2007). However, this carbon neutrality holds only if the biomass production would not have occurred in the absence of biofuel production. If the carbon would have been sequestered and remained so in the absence of biofuel production, releasing this carbon as CO_2 through biofuel combustion is not carbon neutral; it adds to the stock of CO_2 in the atmosphere (Searchinger *et al.*, 2009). To calculate whether biofuel production uses net "additional" biomass requires considering whether feedstock production results in crop displacement and land clearing. So another way of viewing ILUC uncertainty is that if it cannot be proved that ILUC has not occurred, it should not be assumed that a biofuel is carbon neutral (Searchinger, 2010).

Estimating the CO_2 Emissions from ILUC

Although DLUC can be observed and the effects measured, the indirect effects of biofuel expansion are triggered through price changes in global markets and as such are unobservable (Nassar *et al.*, 2011). Estimating ILUC, therefore, requires the use of models. Because the main linkages are economic, economic models have generally been used to estimate ILUC.

Agricultural extensification is recognized as a leading proximate cause of deforestation (Gibbs *et al.*, 2010; Nassar *et al.*, 2011); however, economic effects are only one of the several interacting drivers of land conversion (Geist and Lambin, 2002; Pfaff *et al.*, 2007). Other driving forces include social processes, such as human population growth and migration, and national policies affecting agriculture, land use, and economic development (Geist and Lambin, 2002), as well as cultural, technological, and institutional issues, all interacting in complex relationships (Schaeffer *et al.*, 2005). In this context, deforestation is best understood as an emergent characteristic of a complex system, with a range of proximate and ultimate causes. Given this complexity, the ability to predict ILUC from a single driver such as increases in commodity prices may be quite limited, and thus a core assumption underlying ILUC modeling is called into question, resulting in model uncertainty that is difficult to quantify.

The economic analyses of ILUC performed to date have assumed that a marginal increase in LUC in any region can be estimated from supply and demand functions for commodities and land, despite the presence of other drivers. **Figure 1** illustrates the steps typically used to estimate ILUC emissions from biofuels expansion. When biofuels are produced on former cropland, a cascade of displacements and substitutions may occur. These interactions are estimated using an economic model (step a), which projects changes in economic uses of land, such as cropping, pasture, and forestry. As economic models are not generally spatially explicit, changes in economic uses must be mapped onto specific land cover types (step b). Next, the carbon emissions from conversion of the identified land cover types to its projected new use (step c) are estimated. Finally, for use in biofuel regulations, the estimated ILUC emissions are typically associated with a quantity of biofuel production, resulting in an ILUC emission factor denominated in gram CO_2 per megajoule (step d). Each of these steps is discussed further below.

Estimating Global Market Impacts

There are two main classes of economic models, partial equilibrium (PE) and computable general equilibrium (GE) models, both of which have been applied for estimating the economic and land-use effects of biofuel policies. PE models offer highly detailed representation of one or a few sectors of the economy, such as the agricultural sector; they, however, lack linkages to other sectors of the economy (Kretschmer and Peterson, 2010). GE models, in contrast, are comprehensive in their representation of the economy, reflecting feedback effects among all economic sectors; they, however, offer less detail. For example, a GE model might represent a composite "coarse grains" sector rather than individually representing the characteristics of corn, oats, barley, sorghum, and so on. Computational and data limitations force a tradeoff between detail and scope; combining the scope of GE models with the detail of PE models would result in a model that was computationally intractable. Both approaches have their proponents and detractors; neither is clearly superior for modeling ILUC.

Given the differing modeling approaches, data sets, and parameter choices, models of ILUC emissions have produced widely divergent results. **Figure 2** shows the results from several

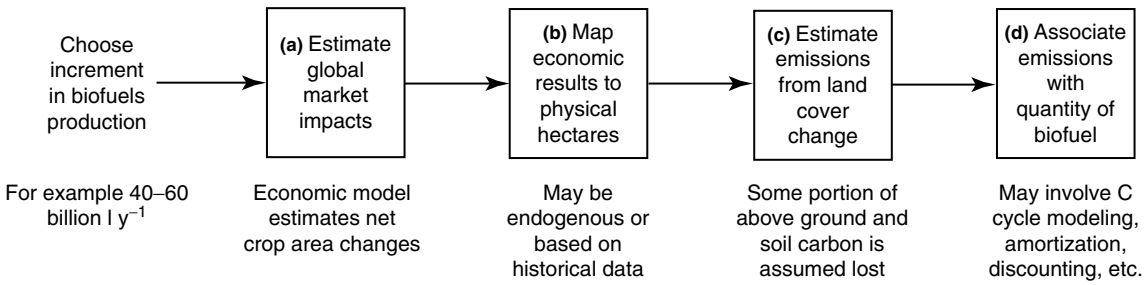


Figure 1 Schematic of indirect land-use change calculation using economic models.

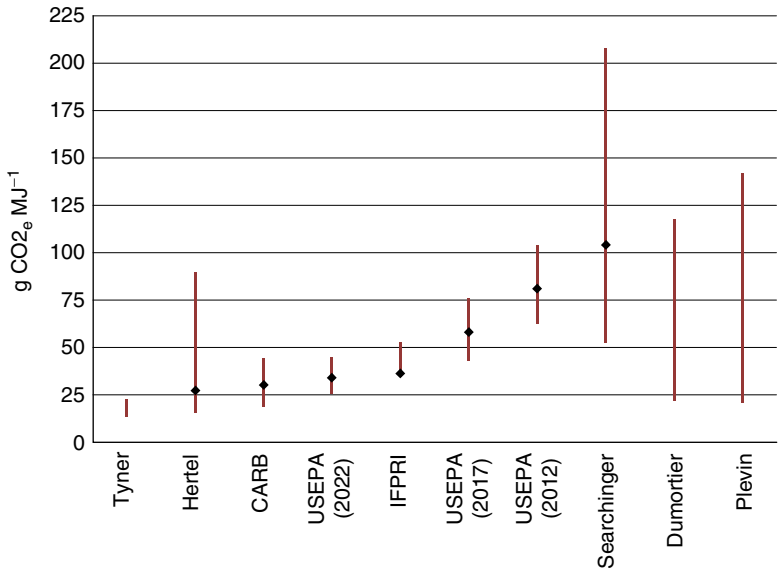


Figure 2 Ranges of results from models of ILUC emissions induced by expansion of corn ethanol production. Point estimates are indicated by the diamond where applicable. USEPA estimated ILUC emissions that would occur to meet the US Renewable Fuel Standard in each of three years: 2012, 2017, and 2022 (Tyner *et al.*, 2010; Hertel *et al.*, 2010; CARB, 2009; USEPA, 2010; IFPRI = Al-Riffai *et al.*, 2010; Searchinger *et al.*, 2008; Dumortier *et al.*, 2009, and Plevin *et al.*, 2010).

studies modeling the ILUC effects of expanded corn ethanol production. The ranges presented reflect on a variety of approaches to explore the sensitivity of ILUC emissions to various model assumptions. All of the studies shown are based on economic equilibrium modeling except for the Plevin *et al.* (2010) study, which is based on a simplified model parameterized from prior economic model results and thus incorporates model uncertainty in addition to parameter uncertainty.

Mapping Economic Results to Physical Hectares

The GHG emissions resulting from conversion to cropping vary with land cover types and across regions (Fargione *et al.*, 2008). Estimating ILUC emissions, therefore, requires assumptions about the specific land cover types converted as a result of biofuels expansion. As most economic models are not spatially explicit, economic model results must be mapped onto land cover types to predict the emissions from land cover conversion.

One mapping approach relies on survey data or remote sensing to identify the agricultural frontier – the regions in which natural land cover has most recently been converted to agriculture. If the economic model predicts, for example, the conversion of forest to cropland in a region, some models assume that the types of forest converted will resemble the types of forest conversion previously detected in that region. That is, it is assumed that prior land conversion patterns are predictive of future patterns.

For example, Winrock International compared moderate-resolution imaging spectroradiometer (MODIS) satellite images from 2001 and 2007 to detect land-use changes (Harris *et al.*, 2009). However, as noted by the authors, comparing already-classified land cover sets from two time periods can involve high error rates as small differences can be magnified when a parcel of land changes just enough to change classification between the two data sets. A more accurate – and much more time-consuming – approach would be to compare the two images before classification and to classify the changes directly.

Another approach considered by the European Union's (EU's) Joint Research Center (JRC) uses a spatial allocation model that incorporates land cover trends detected using remote sensing with predictive information such as the distance from existing cropland and the suitability of the land for producing different crops (Hiederer *et al.*, 2010).

Estimating Emissions from Land Cover Change

The magnitude and time profile of emissions from land cover change depend on the type of land cover and the mode of clearing. Burning releases most biomass carbon immediately, mostly as CO_2, but also as black carbon (soot), organic carbon, carbon monoxide (CO), methane (CH_4), and other volatile organic compounds. Nitrous oxide (N_2O) is also generally released on burning biomass.

Estimates of the total CO_2-equivalent emissions from land conversion usually rely on the well-documented Intergovernmental Panel on Climate Change GHG inventory methods, which include estimates of the CO_2 emissions from above- and below-ground biomass, dead organic matter (dead wood and litter), and soil organic matter, as well as non-CO_2 gases from burning above-ground biomass and dead organic matter (IPCC, 2006). Models of ILUC generally also consider some number of years of foregone sequestration of carbon, that is, sequestration presumed to have occurred if not for land cover conversion.

Total emissions are estimated by multiplying the total area in each land cover type by the associated emission factors. Even for a single set of economic model results, different estimates of total emissions will result from different land cover

allocations. For example, the JRC estimated ILUC emissions induced by the EU Renewable Fuel Directive using the MIRAGE GE model. Applying the spatial allocation model (*see* Mapping Economic Results to Physical Hectares) to the MIRAGE resulted in emissions of 168 Mg CO_2 from above- and below-ground carbon stocks, compared with the prior estimates of 43 Mg CO_2 based on the same MIRAGE results and the Winrock land allocation method (Hiederer *et al.*, 2010; Table 25, p. 106). The increased estimate of emissions results from a greater allocation of land-use changes to areas that have a high-carbon carrying capacity (e.g., forests and shrubland) in the Spatial Allocation Model.

Associating ILUC Emissions with a Quantity of Biofuel

When ILUC emissions are included in biofuel regulations, the total emissions are generally associated with a quantity of fuel production to produce a metric in g CO_2e per megajoule. The simplest approach, used by regulators in USA and EU is straight-line amortization, dividing the emissions over 20 (EU) or 30 (US) years of biofuel production (Note that for any given total for ILUC emissions, amortizing over 20, rather than 30, years increases the per-MJ value by 50%). Another approach compares the total cumulative radiative forcing from biofuels to that of gasoline and measures the difference at an arbitrary time in the future (O'Hare *et al.*, 2009). Others have suggested long time horizons, discounting either emission flows (USEPA, 2009) or damages from emissions (Delucchi, 2011). Others have suggested charging for ILUC emissions when they occur (mostly in the first year) rather than amortizing (Melillo *et al.*, 2009).

Food Price Rationing

As noted earlier, one of the effects of biofuel expansion in the resulting competition for land is an increase in food prices. In response to this increase, economic models of ILUC indicate a reduction in food consumption. Because of reduced consumption, less additional land is required than would be if food demand were unchanged. Thus in models of ILUC emissions, the reduction in food consumption appears as a GHG "benefit" (Searchinger *et al.*, 2008; Hertel *et al.*, 2010). One approach to including this social cost in the ILUC factor would be to hold food consumption constant when estimating ILUC. In one such modeling experiment, holding food consumption constant increased the ILUC factor by approximately 40% (Hertel *et al.*, 2010).

Uncertainty in ILUC Emission Estimates

Despite many uncertainties and a wide range of estimates, the consensus from a suite of different economic modeling efforts is that the biofuels expansion results in net emissions from ILUC (Edwards *et al.*, 2010; Plevin *et al.*, 2010). Substantially narrowing the range will be difficult owing to the challenges of accurately modeling the response of the global economy to changes in production. Economic models tuned to historical performance are better suited in projecting small change to the *status quo* and less well-suited in modeling the effects of the rapid increase in biofuel production witnessed in the recent history and anticipated in the near future. Another challenge is the lack of spatial reasoning in economic models, requiring a variety of imperfect methods for mapping projected economic activities to specific land cover types.

One important uncertainty not considered in current economic models is the effect of climate change on agriculture. Global increases in extreme weather events suggest that existing models will underestimate future yield in some areas and overestimate yields in far more locations. Based on a probabilistic analysis of climate change and its effects on crop yields, Tebaldi and Lobell (2008) wrote: "[We] estimate the chance that global losses from climate change by 2030 will outweigh gains from CO_2 as unlikely for wheat ($<33\%$ chance), likely for barley ($>66\%$ chance) and virtually certain for maize ($>99\%$ chance). In addition, we estimate larger than 80% chance that net losses for maize will exceed 10% over this relatively short time period." Lobell and Burke (2010) noted that "although most cropping systems exhibit a clearly negative yield response to warming, the precise amount of yield loss per degree warming is often not tightly constrained, either from theory or observations." The longer the analytical horizon used to estimate the climate effects of fuels, the more important it becomes to consider the potential effects of climate on agriculture. An important nexus exists here between the economic models and long-range ecological assessments, such as the Millennium Ecosystem Assessment (http://www.maweb.org).

Approaches to Avoiding ILUC

Although estimates of ILUC emissions remain uncertain, ILUC can be reduced, and possibly avoided, by limiting competition between bioenergy feedstocks and other high-demand commodities (Schubert *et al.*, 2010, p. 210). The two main approaches are (1) to ensure that bioenergy crops do not compete with existing food and feed production, and (2) use byproducts of other production systems rather than purpose-grown crops as feedstocks.

Some bioenergy crops can be produced on a so-called "marginal" land that cannot support commercial agriculture. The most beneficial situation from a climate change perspective is to grow perennial grasses, such as switchgrass, miscanthus, or mixed prairie grasses on marginal land, where their deep root systems can enrich soil carbon and improve soil structural properties (Hill *et al.*, 2006; Blanco-Canqui, 2010).

An approach favored by the Roundtable on Sustainable Biofuels (a voluntary certification organization) is to grow bioenergy crops in responsible cultivation areas where the land use from increased bioenergy production is offset locally by intensification, such as increasing cattle stocking density or double cropping, thereby ensuring that the area maintains its prior output while producing energy crops.

The organic portion of municipal solid waste and waste from the food processing and timber industries can also be used to produce biofuels without competing with crops. Estimates of net GHG benefits of using these feedstocks for

biofuels should consider the alternative fate of the biomass. For example, combusting biomass that would have otherwise remained intact in a dry landfill releases additional CO_2 to the atmosphere just as the burning of fossil fuels does. In other cases, the removal of biomass from landfills can reduce methane emissions, resulting in additional GHG reductions.

Crop residues can also be used as a biofuel feedstock. The collection and use of corn stover has been shown to reduce corn yield in subsequent years and to reduce the accumulation of soil carbon (Blanco-Canqui, 2010). Compensating for reduced yield can require extensification, triggering ILUC, or intensification, resulting in additional fertilizer-related N_2O emissions. From a climate change perspective, reducing soil carbon sequestration is equivalent to emitting the carbon. Blanco-Canqui estimates that only 25% of residue can be collected without adversely impacting soil properties, erosion, soil carbon sequestration, and crop production (Blanco-Canqui, 2010). Collecting such a small percentage of stover may not be economically feasible.

See also: Biodiversity in Logged and Managed Forests. Biofuels and Biodiversity: The Implications of Energy Sprawl. Deforestation and Land Clearing. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production. Life Cycle Analysis of Biofuels. Rainforest Loss and Change

References

Al-Riffai P, Dimaranan B, and Laborde D (2010) *Global Trade and Environmental Impact Study of the EU Biofuels Mandate*. Washington, DC: IFPRI. http://trade.ec.europa.eu/doclib/docs/2010/march/tradoc_145954.pdf

Blanco-Canqui H (2010) Energy crops and their implications on soil and environment. *Agronomy Journal* 102: 403–419.

CARB (2009). Proposed regulation to implement the low carbon fuel standard, vol. ii. *Staff Report: Initial Statement of Reasons*. Sacramento, CA: California Air Resources Board.

Davis SC, House JI, Diaz-Chavez RA, *et al.* (2011) How can land-use modelling tools inform bioenergy policies? *Interface Focus* 1: 212–223.

Delucchi M (2011) A conceptual framework for estimating the climate impacts of land-use change due to energy crop programs. *Biomass and Bioenergy* 35: 2337–2360.

Dumortier J, Hayes DJ, Carriquiry M, *et al.* (2009) *Sensitivity of Carbon Emission Estimates From Indirect Land-Use Change*. Ames, IA: Center for Agricultural and Rural Development, Iowa State University.

Edwards R, Mulligan D, and Marelli L (2010) *Indirect Land Use Change From Increased Biofuels Demand: Comparison of Models and Results for Marginal Biofuels Production From Different Feedstocks*. Ispra: EC Joint Research Centre – Institute for Energy.

Fargione J, Hill J, Tilman D, *et al.* (2008) Land clearing and the biofuel carbon debt. *Science* 319: 1235–1238.

Geist HJ and Lambin EF (2002) Proximate causes and underlying driving forces of tropical deforestation. *BioScience* 52: 143–150.

Gibbs HK, Ruesch AS, Achard F, *et al.* (2010) Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proceedings of the National Academy of Sciences*. DOI: 10.1073/pnas.0910275107.

Harris N, Grimland S, and Brown S (2009) Land use change and emission factors: Updates since the RFS proposed rule. Arlington, VA, USA: Winrock International. http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OAR-2005-0161-3163

Hertel TW, Golub A, Jones AD, *et al.* (2010) Global land use and greenhouse gas emissions impacts of U.S. Maize ethanol: Estimating market-mediated responses. *BioScience* 60: 223–231.

Hiederer R, Ramos F, Capitani C, *et al.* (2010) *Biofuels: A New Methodology to Estimate GHG Emissions From Global Land Use Change: A Methodology Involving Spatial Allocation of Agricultural Land Demand and Estimation of CO_2 and N_2O Emissions*. Ispra: EC Joint Research Centre – Institute for Energy.

Hill J, Nelson E, Tilman D, *et al.* (2006) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11206–11210.

IPCC (2006) *2006 IPCC Guidelines for National Greenhouse Gas Inventories*, Vol. 4: Agriculture, Forestry and Other Land Use. Hayama, Japan: Institute for Global Environmental Strategies (IGES), on behalf of the IPCC.

Kløverpris J, Wenzel H, and Nielsen P (2008) Life cycle inventory modelling of land use induced by crop consumption: Part 1: Conceptual analysis and methodological proposal. *The International Journal of Life Cycle Assessment* 13: 13–21.

Kretschmer B and Peterson S (2010) Integrating bioenergy into computable general equilibrium models – A survey. *Energy Economics* 32: 673–686.

Laurance WF (2007) Switch to corn promotes Amazon deforestation. *Science* 318: 1721b.

Lobell D and Burke M (2010) Economic impacts of climate change on agriculture to 2030. In: Reynolds MP (ed.) *Climate Change and Crop Production*. Wallingford, UK: CABI.

Melillo JM, Reilly JM, Kicklighter DW, *et al.* (2009) Indirect emissions from biofuels: How important? *Science* 326: 1397–1399.

Nassar AM, Harfuch L, Bachion LC, and Moreira MR (2011) Biofuels and land-use changes: Searching for the top model. *Interface Focus* 1: 224–232.

O'Hare M, Plevin RJ, Martin JI, *et al.* (2009) Proper accounting for time increases crop-based biofuels' greenhouse gas deficit versus petroleum. *Environmental Research Letters* 4: 024001.

Pfaff A, Robalino J, Walker R, *et al.* (2007) Road investments, spatial spillovers, and deforestation in the Brazilian Amazon. *Journal of Regional Science* 47(1): 109–123.

Plevin RJ, O'Hare M, Jones AD, *et al.* (2010) Greenhouse gas emissions from biofuels: Indirect land use changes are uncertain but may be much greater than previously estimated. *Environmental Science & Technology* 44: 8015–8021.

Rabl A, Benoist A, Dron D, *et al.* (2007) How to account for CO_2 emissions from biomass in an LCA. *The International Journal of Life Cycle Assessment* 12: 281.

Schaeffer R, Vianna R, Leonardo R, *et al.* (2005) Underlying causes of deforestation. *Science* 307: 1046–1047.

Schubert R, Schellnhuber HJ, Buchmann N, *et al.* (2010) *Future Bioenergy and Sustainable Land Use*. London, UK: Earthscan.

Searchinger T, Heimlich R, Houghton RA, *et al.* (2008) Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land use change. *Science* 319: 1238–1240.

Searchinger TD (2010) Biofuels and the need for additional carbon. *Environmental Research Letters* 5: 024007.

Searchinger TD, Hamburg SP, Melillo J, *et al.* (2009) Fixing a critical climate accounting error. *Science* 326: 527–528.

Tebaldi C and Lobell DB (2008) Towards probabilistic projections of climate change impacts on global crop yields. *Geophysical Research Letters* 35: L08705.

Tyner WE, Taheripour F, Zhuang Q, *et al.* (2010) *Land Use Changes and Consequent CO_2 Emissions due to US Corn Ethanol Production: A Comprehensive Analysis*. West Lafayette, IN, USA: Department of Agricultural Economics, Purdue University. http://www.transportation.anl.gov/pdfs/MC/625.PDF

USEPA (2009) *Draft Regulatory Impact Analysis: Changes to Renewable Fuel Standard Program*. Washington, DC: US Environmental Protection Agency.

USEPA (2010) *Renewable Fuel Standard Program (RFS2) Regulatory Impact Analysis*. Washington, DC: US Environmental Protection Agency.

Insecticide Resistance

Ian Denholm, Rothamsted Research, Harpenden, UK
Greg Devine, Queensland Health, Cairns, QLD, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acetylcholinesterase (AChE) The enzyme responsible for breaking down the neurotransmitter acetylcholine (ACh) at nerve synapses, thereby preventing hyperexcitation of cholinergic pathways in the nervous system. The target of organophosphate and carbamate insecticides.
Bioassay (biological assay) A laboratory test for evaluating the response of organisms to a toxin and for diagnosing the presence or absence of resistance.
Cross-resistance The ability of a single gene or mechanism to confer resistance to more than one toxin.
Cytochrome P-450 monooxygenases A ubiquitous group of enzymes involved in the NAPDH-mediated oxidation and metabolism of a broad range of endogenous and exogenous substrates.
GABA receptor Part of the inhibitory ion channel complex gated by GABA (γ-aminobutyric acid) in postsynaptic nerve membranes. The target site of organochlorine insecticides and some other chemical groups.
Glutathione *S*-transferases (GSTs) Enzymes that catalyze the metabolism of a range of substrates following their conjugation with the endogenous tripeptide glutathione.
Multiple resistance The occurrence of more than one resistance mechanism in the same individual or pest population.
Sodium channel A large transmembrane protein that regulates the flow of sodium ions across axonal membranes and mediates the rising phase of action potentials. The target site of pyrethroid insecticides.
Synergist A chemical used at sublethal concentrations to inhibit particular groups of detoxifying enzymes and therefore to implicate these enzymes in resistance.

Introduction

The diversity of organisms and their genetic variation have been forged by evolution. In evolution, undirected mutations generate unstructured diversity, which is then structured by selection. The time scales over which selection and adaptation occur in eukaryotes are usually too large to observe *in situ* (except in a few cases, such as industrial melanism in some moths and ladybirds), but there is one microevolutionary process for which many of the factors driving selection and adaptation are well understood, a process that is being increasingly studied and manipulated by a large number of biologists. This is the phenomenon of pest resistance to insecticides, and its study can reveal much about how biodiversity originates, at the intraspecific level at least. Similar processes underpin the evolution of herbicide resistance by weeds, fungicide resistance by plant pathogens, and drug resistance by disease-causing microbes.

Extent of Resistance

Although a relatively recent phenomenon (resistance to the first synthetic insecticide, dichlorodiphenyltrichloroethane or DDT, was initially reported in the 1940s), insecticide resistance is now very widespread. Available statistics (**Figure 1**) show that reports of resistant arthropod species increased almost exponentially between 1950 and 1980, following the successive introduction of different classes of synthetic insecticides. By 1990, more than 500 species were reported to resist chemicals of at least one insecticide class, and many of these resisted several classes simultaneously. Of the resistant species reported in 1990, 88% were insects (class Insecta) and 12% were mites and ticks (class Arachnida, order Acarina). Four insect orders – Coleoptera (beetles), Diptera (true flies), Hemiptera (aphids, bugs, hoppers, and whiteflies), and Lepidoptera (moths) – accounted for 92% of the resistant insect species; the remainder mostly comprised of cockroaches, thrips, lice, and fleas. Additional resistant species have been reported since 1990, but a change in survey methods makes it difficult to extend the temporal trend evident in **Figure 1**.

Although almost all insecticide classes are now affected by resistance, its extent varies greatly between species. In some insects, resistance only extends to a few closely related compounds in a single class; it may be very weak or restricted to a small part of their geographical range. At the other extreme,

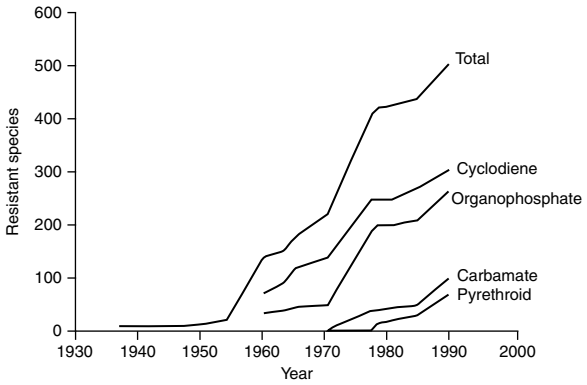


Figure 1 Increase in the number of arthropod species reported to resist insecticides over time, in total, and in response to the four most widely used classes of insecticide (courtesy GP Georghiou).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00104-0



Figure 2 Examples of arthropod pests that have developed resistance to insecticides, (a) The bollworm, *Helicoverpa armigera*, a major pest of cotton and vegetables in the Old World; (b) the tobacco or cotton whitefly, *Bemisia tabaci*, which threatens a wide range of crops through direct-feeding damage and the transmission of virus disease; (c) the peach–potato aphid, *Myzus persicae*, a major pest of vegetable and ornamental crops in temperate countries; and (d) the two-spotted spider mite, *Tetranychus urticae*, a cosmopolitan pest of fruit, vegetable, and ornamental crops.

some widespread pests such as the cotton bollworm (*Helicoverpa armigera*; **Figure 2(a)**), the tobacco whitefly (*Bemisia tabaci*; **Figure 2(b)**), the peach–potato aphid (*Myzus persicae*; **Figure 2(c)**), and the two-spotted spider mite (*Tetranychus urticae*; **Figure 2(d)**) can resist most or all of the insecticides available for their control. The most extensively used insecticide classes – organochlorines, organophosphates, carbamates, and pyrethroids – have generally been the most seriously compromised by resistance, and many principles relating to the origin and evolution of resistance can be demonstrated solely by reference to these fast-acting neurotoxins. In recent years, however, there has also been a worrying increase in resistance to more novel insecticides, including ones attacking the nervous system (e.g., neonicotinoids), developmental pathways (e.g., benzoylphenylureas), respiratory processes (e.g., mitochondrial electron-transport inhibiting (METI) acaricides), and digestive systems (e.g., *Bacillus thuringiensis* (Bt) endotoxins).

Origins of Resistance

As insecticides are not normally considered mutagenic at field application rates, it is assumed that resistance mutations occur independently of insecticide exposure, and are as likely to occur before an insecticide is introduced as they are during its use in the field. Resistance is therefore a preadaptive phenomenon reflecting the selection of individuals possessing heritable genetic traits that promote their survival or reproduction in environments treated with insecticides. Estimates of mutation rates are as imprecise as they are for other adaptive traits, but will depend on the mutational event involved

(*see* Mechanisms of Resistance). The role of enzyme induction in resistance has not been demonstrated satisfactorily but is likely to be only of secondary importance. Increased tolerance due to environmental or biological factors such as diet, age, or climate can sometimes be significant but is outside the scope of this article.

Mechanisms of Resistance

Some of the most significant recent progress in understanding resistance result from applying advances in molecular biology to resistance research. Depending on the mechanism involved, resistance can arise through structural alterations of genes encoding target-site proteins or detoxifying enzymes, or through processes affecting the expression of gene products (e.g., gene amplification or altered transcription rates). Other mechanisms that have been demonstrated or postulated include reduced penetration of insecticides through the insect cuticle, enhanced excretion of insecticides, and behavioral traits enabling pests to reduce or avoid exposure to a toxin. However, the latter are generally considered to be relatively minor in effect or to arise only under very specialized circumstances. **Figure 3** is a schematic representation of some of the resistance mechanisms discussed in the following two sections.

Increased Detoxification of Insecticides

The three major routes of detoxification implicated in resistance are as follows: Enhanced metabolism of insecticides by cytochrome P-450 monooxygenases can potentially confer resistance to most chemical classes. Previously, much of the evidence for this mechanism was indirect, based on the ability of compounds such as piperonyl butoxide, which are known inhibitors of monooxygenases, to reduce the magnitude of resistance when used as synergists in bioassays. However, recent developments in insect genomics are enabling complex families of genes such as those encoding monooxygenases to be cloned and characterized and single enyzymes to be implicated directly in resistance through crossing studies and experiments on insecticide metabolism and binding.

Enhanced activity of glutathione *S*-transferases (GSTs) is considered potentially important in resistance to some classes of insecticide, including organophosphates. However, information on the role of GSTs in resistance is still sketchy but these enzymes, like monooxygenases, exist in numerous molecular forms with distinct properties, and are proving amenable to molecular screening and characterization.

Enhanced hydrolysis or sequestration by esterases (e.g., carboxylesterases) capable of binding to and cleaving carboxylester and phosphotriester bonds undoubtedly plays a significant role in resistance to organophosphates and pyrethroids. Of the three main types of detoxification mechanisms, this has historically been the best characterized biochemically. Resistance due to increased esterase activity can arise through either a qualitative change in an enzyme, improving its hydrolytic capacity, or (as in mosquitoes and aphids) a quantitative change in the titer of a particular enzyme that already exists in susceptible insects.

Alterations to Insecticide Target Sites

Resolving target-site modifications that lead to resistance requires some knowledge of the mode of action of insecticides themselves. At present, this information is most advanced for insecticides binding to enzymes or receptors in the nervous system of arthropods. Three examples of target-site resistance are well understood:

Organophosphates and carbamates exert their toxicity by inhibiting the enzyme acetylcholinesterase (AChE), thereby

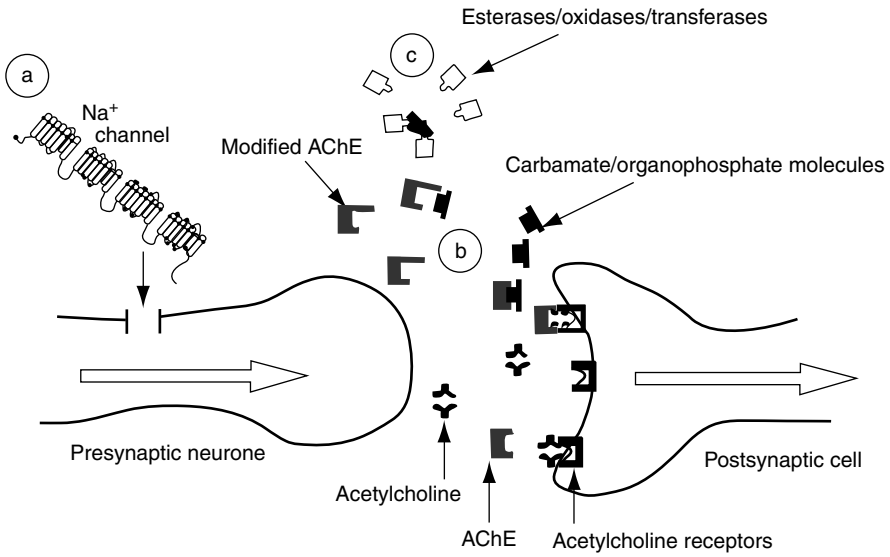


Figure 3 Schematic diagram of a nerve synapse showing examples of insecticide-resistance mechanisms. (a) Changes in the structure of the sodium channel confer knockdown resistance (kdr) to pyrethroids. (b) Modified acetylcholinesterase (AChE) is no longer bound by organophosphates and remains available to breakdown acetylcholine molecules after neurotransmission across the synapse. (c) Detoxifying enzymes degrade or sequester insecticides before they reach their targets in the nervous system.

impairing the transmission of nerve impulses across cholinergic synapses. Mutant forms of AChE showing reduced inhibition by these insecticides have been demonstrated in several insect and mite species. Biochemical and molecular analyses of insecticide-insensitive AChE have shown that pests may possess several different mutant forms of this enzyme with contrasting insensitivity profiles, thereby conferring distinct patterns of resistance to these two large insecticide classes.

Pyrethroids act primarily by binding to and blocking the voltage-gated sodium channel of nerve membranes. A resistance mechanism causing insensitivity of this target site was first identified in houseflies (*Musca domestica*) and termed knockdown resistance (kdr). Kdr is attributable to structural modifications of the sodium channel protein and the genetic mutations responsible have been identified. This mechanism has been found in many pest species. As with insensitive AChE, there can be different forms of kdr resistance (e.g., a more potent "super-kdr" form in which a second mutation potentiates the level of resistance conferred by the original one).

GABA receptors are targets for several insecticide classes, including cyclodienes (a subclass of the organochlorines), avermectins, and fipronils. The primary mechanism of resistance to cyclodienes and fipronils involves modification of a particular GABA receptor subunit, resulting in substantial target-site insensitivity to these insecticides.

Homology of Resistance Genes

Although there are several possible mechanisms, the options for resisting insecticides can also be very limited, especially for mechanisms based on decreased sensitivity of insecticide target sites. The target-site mechanism of cyclodiene resistance involves the same amino acid substitution (alanine-302 to serine) in GABA receptors of several species of diverse taxonomic origin, including beetles, mosquitoes, whiteflies, and aphids. Work on the two other principal target-site mechanisms – altered AChE and kdr – has proved more challenging due to the occurrence of multiple resistance alleles at the same loci. In the case of kdr, however, there is also evidence for the same amino acid substitution (leucine-1014 to phenyl-alanine) in the sodium channel protein conferring a "basal" kdr phenotype in a wide range of species. This phenotype can be enhanced (to "super-kdr" resistance) by further mutations that also recur between species. Despite the structural complexity of the receptors involved, these parallel mutations imply that the opportunities for insects to modify them to avoid or reduce binding of insecticides, while retaining normal functioning of the nervous system, are very limited indeed.

When susceptible individuals of the sheep blowfly (*Lucilia cuprina*) were exposed to the mutagen ethyl methanesulfonate (EMS) and their progeny screened for resistance to dieldrin (a cyclodiene), surviving insects not only exhibited a GABA receptor-based mechanism analogous to that found in nature but also exhibited an identical alanine to serine amino acid substitution in the receptor gene. Similarly, mutagenesis followed by screening with diazinon (an organophosphate) led to the recovery of a resistance mechanism showing identical toxicological, biochemical, genetic, and molecular properties to one that had previously evolved to diazinon under field conditions. These findings reinforce the tight evolutionary constraints on the number of viable resistance mutations, even in the laboratory, where mutations conferring deleterious effects on overall fitness might be expected to survive better than in the open field. Interestingly, they also highlight the potential of using mutagenesis to predict likely resistance mechanisms to novel insecticides in advance of them being used commercially and to tailor resistance-management recommendations accordingly.

In other cases, different types of resistance to the same toxin exist and can account for differences in the toxicological and genetic basis of resistance between species or between different geographical populations of the same species. Insecticidal proteins from the soil bacterium Bt are becoming increasingly important in pest management, especially in relation to insect-tolerant transgenic crops (*see* The Special Case of Transgenic Plants). To date, the only species to have evolved Bt resistance on a large scale in the field is the diamondback moth, *Plutella xylostella*. The majority of Bt-resistant populations examined have exhibited very similar characteristics, including a very consistent pattern of cross-resistance to different Bt toxins and recessive inheritance. However, there are also strains of *P. xylostella* in which the breadth and inheritance of Bt resistance differ markedly from this "mode," implying the existence of distinct resistance genes and mechanisms.

How often do Resistance Genes Arise?

The recurrence of specific resistance mutations within and between taxa begs another question of fundamental significance to the origins of biodiversity: Have such mutations arisen repeatedly within the same species, or appeared on only a limited number of occasions and subsequently spread through migration or human agency? This question is amenable to investigation by sequencing not only the resistance genes themselves but also flanking regions and introns, which would be expected to vary between alleles that have arisen independently. In the mosquito, *Culex pipiens*, organophosphate resistance is primarily conferred by allozymes at two closely linked loci (esterases A and B) coding for insecticide-detoxifying carboxylesterases. Overproduced allozymes (resulting from amplification of A or B genes) tend to recur in geographically disjunct areas. This situation could be explained by recurrent mutation generating each amplification event *de novo* or by a nonrecurrent mutation that has subsequently spread between populations. Restriction mapping of DNA around the esterase genes points to the latter explanation, with large-scale gene flow (even between continents) most likely attributable to passive migration of mosquitoes on ships or airplanes. It is notable that the appearance of a new resistance allele in southern France originated in the vicinity of the international airport and seaport at Marseilles.

Organophosphate resistance in the aphid *M. persicae* is also attributable to the amplification of a gene encoding an insecticide-detoxifying carboxylesterase. Despite the often widespread dispersion of these amplified genes in the aphid genome, restriction analyses have indicated that all copies are in the same immediate genetic background. This suggests that

amplification occurred only once, with the amplified DNA subsequently being moved intact around the genome through chromosomal rearrangements, or perhaps mediated by transposable elements. Similarities in the position and structure of these genes in aphids of diverse geographic origin reinforce the likelihood of a single amplification event that has subsequently become widely dispersed around the world.

However, there is also molecular evidence for some resistance genes having several independent origins in the same species (e.g., for target-site resistance to cyclodienes in the red flour beetle, *Tribolium castaneum*). Results for mosquitoes and aphids, nonetheless, highlight the potential for large-scale inadvertent movement of resistant insects between countries or even continents. For crop pests, these risks are particularly acute due to the increasing international trade in edible and ornamental plants, many of which have been treated with insecticides at their point of origin. In such cases, growers at the receiving end of the trade network face a dual threat: (1) the establishment of new pest species or more aggressive biotypes of existing ones and (2) the possibility that such pests are already strongly resistant to compounds that might otherwise be used to suppress or eradicate them.

Cross-Resistance and Multiple Resistance

Arthropods seldom resist just one toxin. Most commonly, they exhibit differing levels of resistance to a range of related and unrelated insecticides. In its strictest sense, the term "cross-resistance" refers to the ability of a single mechanism to confer resistance to several insecticides simultaneously. A more complex situation is that of "multiple resistance," reflecting the coexistence of two or more resistance mechanisms, each with their own specific cross-resistance characteristics. Disentangling cross-resistance from multiple resistance is a challenging aspect of resistance research. However, knowledge of the mechanisms involved is often essential in order to develop resistance-management recommendations based, for example, on the alternation of insecticides to avoid continuous selection for the same resistance gene or mechanism.

Unfortunately, cross-resistance patterns are inherently difficult to predict in advance, because mechanisms based on both increased detoxification and altered target sites can differ substantially in their specificity. The most commonly encountered patterns of cross-resistance tend to be limited to compounds within the same chemical class (equivalent to the term "side resistance" as used by parasitologists). However, even these patterns can be very idiosyncratic. For example, organophosphate resistance based on increased detoxification or target-site alteration can be broad-ranging across this group or highly specific to a few chemicals with particular structural similarities. The breadth of target-site resistance to pyrethroids also depends on the resistance allele present. The most frequent mutation affects almost all compounds in this class to a similar extent (approximately 10-fold resistance), whereas resistance due to other mutations is highly dependent on the alcohol moiety of pyrethroid molecules, ranging from approximately 10-fold to virtual immunity. Cross-resistance between insecticide classes is even harder to anticipate, especially for broad-spectrum detoxification systems whose specificity depends not only on insecticides having the same mode of action but also on the occurrence of common structural features that bind with detoxifying enzymes.

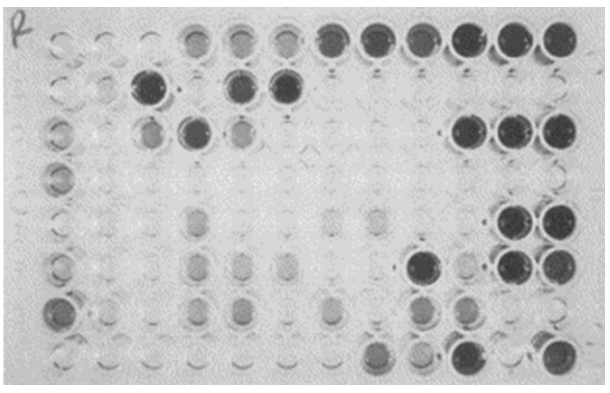
Empirical approaches for distinguishing between cross-resistance and multiple resistance include (1) repeated backcrossing of resistant populations to fully susceptible ones to establish whether resistance to one chemical cosegregates consistently with resistance to another and (2) reciprocal selection experiments whereby populations selected for resistance to one chemical are examined for a correlated change in response to another. If available, biochemical or molecular diagnostics for specific resistance genes can assist considerably with tracking the outcome of genetic crosses or with assigning cross-resistance patterns to particular mechanisms.

Diagnosis of Resistance

A large number of laboratory bioassay methods (e.g., **Figure 4(a)**) have been developed for detecting and characterizing resistance.



(a)



(b)

Figure 4 Examples of resistance diagnostics for the peach–potato aphid, *Myzus persicae*; (a) a bioassay in which insecticide droplets are applied topically to individual aphids; and (b) a biochemical immunoassay to determine levels of an insecticide-detoxifying esterase in individual aphids. The top row shows results for three insects from each of four resistance categories, with resistance increasing from left to right. The remaining insects were drawn randomly from a field sample.

Most of these are limited to defining phenotypes and provide little or no information on the underlying genes or mechanisms. Thus, although bioassays remain the mainstay of most large-scale resistance monitoring programs, much attention has been paid to develop more incisive techniques that not only offer greater precision and throughput but also diagnose the type of mechanism(s) present and, whenever possible, the genotypes of resistant insects. A variety of approaches are being adopted for this purpose, including electrophoretic or immunological detection of resistance-causing enzymes, kinetic and end-point assays for quantifying the activity of enzymes or their inhibition by insecticides, and DNA-based diagnostics for mutant resistance alleles. The sensitivity that these techniques can provide is exemplified well by work on the aphid *M. persicae*, which in northern Europe possesses at least three coexisting resistance mechanisms: (1) an overproduced carboxylesterase conferring resistance to organophosphates, (2) an altered AChE conferring resistance to certain carbamates, and (3) target-site (kdr) resistance to pyrethroids. These mechanisms collectively provide strong resistance to virtually all available aphicides. Fortunately, it is now possible to diagnose all three in individual aphids using an immunoassay for the overproduced esterase (**Figure 4(a)**), a kinetic microplate assay for the mutant AChE, and a DNA-based diagnostic for the kdr allele. The combined use of these techniques against field populations provides up-to-date information on the incidence of the mechanisms and is used to alert growers to potential control problems.

Selection of Resistance Genes

The rate at which resistance genes are selected reflects the combined influence of numerous biotic and abiotic factors. Resistance offers several advantages for research to resolve these factors and their interactions. First, the selecting agent (exposure to insecticides) is well understood; it can usually be carefully documented using treatment histories or manipulated to investigate its effect on selection rates. Second, the selective advantages conferred by resistance genes are often very large, leading to substantial changes in genetic composition over a measurable time frame. Third, most of the major mechanisms of insecticide resistance, unlike those for many stress-related adaptations, are controlled by single genes of major effect (monogenic) rather than many genes, each of small effect (polygenic). This renders resistance more readily amenable to analysis within the theoretical framework of ecological genetics. Finally, the frequent availability of bioassays for quantifying the frequency of resistance phenotypes, or even *in vitro* assays for specific genotypes, enables accurate documentation of responses to selection applied in population cages in the laboratory or under open field conditions.

Factors determining the selection (and hence risk) of resistance to insecticides can, for convenience, be classified into genetic or ecological ones relating to the intrinsic properties of pests and resistance mechanisms and operational ones relating to the chemical itself and how it is applied. Some of the most important factors apparent from the large body of experimental and theoretical research on resistance selection are summarized in the following three sections.

Genetic Influences

For resistance to evolve, resistance genes must confer a selective advantage over their susceptible counterparts. One of the primary challenges for describing resistance is therefore to estimate the relative fitness of different genotypes under exposure to insecticides. There are different ways of achieving this, the most direct being to release individuals of known susceptible and resistance genotypes into insecticide-treated environments and to monitor their survival. This has been done for a variety of pest species and has identified many, often subtle, influences on how resistance genes are expressed in the field. The dominance of resistance genes, which exerts a major influence on selection rates, is a case in point. In laboratory bioassays evaluating the relative survival of susceptible homozygotes (SS), heterozygotes (RS), and resistance homozygotes (RR) over several insecticide concentrations, dominance can be measured precisely, with RS individuals usually responding in an intermediate manner. In the field, dominance is a changeable phenomenon, depending on the concentration of insecticide applied and its uniformity over space and time. Even when the initial concentration is sufficient to kill RS individuals (rendering resistance effectively recessive), the weathering or decay of residues may result in this genotype showing increased survival and resistance becoming functionally dominant in expression. When resistance genes are still rare, and hence mainly present in heterozygous condition, this can have a profound effect in accelerating the selection of resistance genes to economically damaging frequencies.

The diverse mating systems of insects also influence the rate at which resistance evolves. Although most research has focused on outcrossing diploid species (typified by members of the Lepidoptera, Coleoptera, and Diptera), systems based on haplodiploidy and parthenogenesis also occur among key agricultural pests. In haplodiploid systems, males are produced uniparentally from unfertilized, haploid eggs, and females are produced biparentally from fertilized, diploid eggs. The primary consequence of this, exemplified by whiteflies (**Figure 3(b)**), spider mites (**Figure 3(c)**), and phytophagous thrips, is that resistance genes are exposed to selection from the outset in haploid, hemizygous males, irrespective of intrinsic dominance or recessiveness. Whether a resistance gene is dominant, semidominant, or recessive, resistance can develop at a similar rate under haplodiploidy, whereas recessiveness can cause significant delays (initially at least) in diploid populations.

Most species of aphid undergo periods of parthenogenesis, promoting the selection of clones with the highest levels of resistance or the most damaging combination of resistance mechanisms. However, in holocyclic populations (ones that alternate between sexual and asexual reproduction), this effect is at least partially countered by genetic recombination and the subsequent reassortment of mechanisms during sexual reproduction. In fully anholocyclic (asexual) populations, such as those of *M. persicae* in northern Europe, the influence of parthenogenesis is much more severe and has led to strong and persistent associations between resistance mechanisms within clonal lineages exposed to a succession of different selecting agents.

Ecological Influences

Several aspects of pest ecology, including the dynamics, phenology, and dispersal capabilities of pest organisms, act as primary determinants of resistance development. However, their influence on selection rates can be unpredictable without a sound knowledge of how they interact with patterns of insecticide use. As an example, movement of pests between untreated and treated parts of their range may delay the evolution of resistance, due to the diluting effect of susceptible immigrants. Conversely, large-scale movement can also accelerate the spread of resistance by transferring resistance alleles between localities.

For highly polyphagous crop pests, interactions between pest ecology and insecticide treatments play a particularly critical role in determining selection pressures. Key factors to be considered are the seasonality and relative abundance of treated and untreated plant hosts and patterns of migration between hosts at different times of the year. A good example relates to the two major bollworm species (Lepidoptera: Noctuidae) attacking cotton in Australia. Only *H. armigera* (**Figure 3(a)**) has developed a strong resistance; *H. punctigera*, despite being an equally important cotton pest, has remained susceptible to all insecticide classes. The most likely explanation is that *H. punctigera* occurs in greater abundance on a larger range of unsprayed hosts than *H. armigera*, thereby precluding a significant increase in resistance on treated crops.

However, polyphagy can sometimes be deceptive. In the cotton/vegetable/melon production systems of the southwestern USA, the highly polyphagous whitefly *B. tabaci* has become a devastating pest and a primary target of insecticide sprays. The consequences for resistance have varied substantially on a regional basis. Resistance problems have proved much more severe and persistent on cotton in south-central Arizona than in the extreme southwest of the state and the adjacent Imperial Valley of California. This appears attributable to the higher proportion of unsprayed hosts, especially alfalfa, acting as a buffer to resistance in the latter areas and preventing any directional increase in the severity of resistance over successive seasons. In south-central Arizona, these untreated refuges are much less abundant during the time that cotton is treated with insecticides. Thus, a large proportion of the local whitefly population is forced through a selection "bottleneck" on cotton and exposed to intense selection for resistance, despite an abundance of alternative hosts at other times of the year.

Enclosed environments, such as greenhouses and glasshouses, which restrict migration and escape from insecticide exposure under climatic regimes favoring rapid and continuous population growth, provide ideal ecological conditions for selecting resistance genes. Very low or zero damage tolerance thresholds for high-value ornamental or vegetable produce accentuate the problem by promoting overfrequent spraying and hence intensify selection for resistance. Over the years, these environments have proved potent sources of novel resistance mechanisms for a diverse range of control agents and have presented a particular challenge to attempts at resistance management.

Operational Influences

Although closely linked to the aspects of pest genetics and ecology, operational factors are best distinguished as ones which, in principle at least, are at human's discretion and can be manipulated to influence selection rates. Factors exerting a major influence in this respect include the rate, method, and frequency of applications, their biological persistence, and whether insecticides are used singly or as mixtures of active ingredients.

Equating operational factors with selection is often difficult, because without a detailed knowledge of the resistance mechanisms present it is impossible to test many of the assumptions on which genetic models of resistance are based. Anticipating the selection pressure imposed by a particular application dose of insecticide is a case in point. If resistance alleles are present, the only entirely nonselecting doses will be ones sufficiently high to overpower all individuals, regardless of their genetic composition, or ones sufficiently low to kill no insects at all. The latter is obviously a trivial option. Prospects of achieving the former depend critically on the potency and dominance of resistance genes present. A pragmatic solution to this dilemma is to set application doses as far above the tolerance range of SS individuals as economic and environmental constraints permit, in the hope that at least RS genotypes will be effectively controlled. Even this approach can backfire badly if resistance turns out to be more common than suspected (resulting in the presence of RR homozygotes) or resistance alleles exhibit an unexpectedly high degree of dominance. Unless a high proportion of insects escape exposure altogether, the consequence could then be to select very rapidly and effectively for homozygous resistant populations.

In practice, concerns over optimizing dose rates to avoid resistance are secondary to ones regarding the application process itself. Delivery systems or habitats promoting uneven or inadequate coverage will generally be more prone to selection for resistance, because pests are more likely to encounter exposure conditions under which selection is most intense. This was elegantly demonstrated through experiments assessing the relative survival of endosulfan-susceptible and -resistant phenotypes of the coffee berry borer (*Hypothemus hampei*) in coffee plantations treated with this chemical in New Caledonia. The practice of spraying plantations from roadsides with vehicle-mounted mistblowers generated gradients in the concentration of endosulfan that resulted in different selection pressures in different parts of each field. Similarly, underdosing with the fumigant phosphine in inadequately sealed grainstores has been implicated as a primary cause of resistance to this chemical in a range of stored product pests.

The timing of insecticide applications relative to the life cycle of a pest can also be an important determinant of resistance. A good example relates to the selection of pyrethroid resistance in the cotton bollworm, *H. armigera*, in Australia. On cotton foliage freshly treated with the recommended field dose, pyrethroids killed larvae up to 3–4 days old irrespective of whether they were resistant or not by laboratory criteria. As the sensitivity of larvae of all genotypes to pyrethroids was found to decline with increasing larval size, the greatest

discrimination between susceptible and resistant phenotypes occurred only when larvae achieved a threshold age. Targeting of insecticides against newly hatched larvae, as is generally advocated for bollworm control, not only increases the likelihood of contacting larvae at the most exposed stage in their development but also offers the greatest prospect of retarding resistance by overpowering its expression.

In practice, persistent insecticides are often essential to ensure an acceptable period of control, especially when contending with disease vectors or continued invasion of crop pests from alternative host plants. However, persistent applications can accentuate resistance development by exposing a larger number of individuals to the selecting agent. Another problem is that residues of persistent insecticides decay or become diluted through plant growth, so that resistant insects may survive more effectively than they did at the time of application. Empirical studies with a range of pests including mosquitoes, bollworms, and blowflies have demonstrated that aged deposits discriminate more readily between genotypes or phenotypes than ones freshly applied.

In theory, the coapplication of two or more unrelated chemicals as insecticide mixtures offers substantial benefits for delaying the selection of resistance. The underlying principle is one of "redundant killing," whereby any individual already resistant to one insecticide is killed by simultaneous exposure to another, and vice versa. However, achieving this objective requires not only that each type of resistance is still rare but also that the ingredients confer mutual protection throughout the effective life of an application. Failure to ensure that they exhibit similar biological persistence may lead to one compound exerting greater selection pressure than the other, thereby accelerating the selection of doubly resistant phenotypes. Two potentially conflicting challenges of choosing ideal mixture partners are therefore (1) to ensure maximum similarity in efficacy and persistence against the target pest(s) and (2) to ensure maximum dissimilarity in chemical structure and mode of action to minimize the likelihood of cross-resistance. Difficulties with identifying candidate molecules that meet all these criteria have greatly limited the use of mixtures for combating resistance to conventional insecticides, although they have considerable appeal for sustaining the effectiveness of insect-tolerant transgenic crops (*see* The Special Case of Transgenic Plants).

Fitness Costs Associated with Resistance

Despite the advantages they confer under exposure to insecticides, it is often assumed that resistance genes also confer physiological costs that could lead to counterselection when insecticides are not applied. Some of the best examples of such fitness costs come from studies conducted under harsh or stressful environmental conditions, when even slight differences in relative fitness are likely to have major consequences for the survival of genotypes. For example, resistant strains of bollworms, blowflies, and aphids have all been demonstrated to overwinter less successfully than their susceptible counterparts. Possible explanations for these fitness differentials include the reduced viability of certain life stages, a slower reproductive rate rendering resistant insects more vulnerable to adverse climatic conditions or to predation, or a reduced ability to respond to environmental cues promoting survival. In the aphid *M. persicae*, resistant individuals are less inclined to move from senescing to younger leaves and are therefore more vulnerable to isolation and starvation after leaf abscission. Many resistant aphids also appear less able to avoid attack by parasitoids and predators – a more subtle side-effect with interesting implications for combining chemical and biological control strategies.

Fitness costs associated with resistance can be difficult to demonstrate experimentally, since deleterious effects may only be expressed under particular environmental conditions or be conferred by other genes closely linked to the resistance locus. The most convincing examples are ones in which costs have been found consistently in resistant populations of diverse geographic origin or have persisted after several generations of back-crossing to susceptible insects in order to exclude linkage effects. The potential for fitness drawbacks to be overcome by a process of coadaptation, i.e., the integration of resistance genes with other "modifier" loci that ameliorate fitness costs, has also proved challenging to demonstrate. In a few cases, however, repeated back-crossing of resistant insects to susceptible ones has led to seemingly fit resistance phenotypes acquiring a fitness penalty, apparently due to the uncoupling of resistance genes from modifier loci.

Combating Insecticide Resistance

In most studies of evolution, the primary challenge is to identify selective forces and to interpret their effects on the genetic composition of individuals and populations. With insecticide resistance, it is also necessary to intervene in the evolutionary process and find ways of reducing its deleterious impact on pest management. Failure to do so in the past has had many severe consequences, including the economic failure of cropping systems, the resurgence of insect-transmitted pathogens, and damage to the environment by way of increased insecticide applications.

The concept of insecticide-resistance management (IRM) aims to address these concerns through the development of control strategies for overcoming resistance to currently used compounds, or preventing its appearance in the first place. Although drawing extensively on the theoretical and empirical framework that evolutionary biology provides, IRM strategies must also contend with several practical, economic, and political constraints on the choice of possible management tactics and the precision with which they can be applied. The most important of these are as follows:

1. The properties of any resistance genes present will often be unknown, and knowledge of pest ecology may still be rudimentary.
2. It is often necessary to contend with a whole pest complex rather than just a single pest species.
3. There will often be a very limited number of insecticides available for use in management strategies.
4. For highly mobile pests at least, countermeasures may need to be standardized and synchronized over large areas, sometimes whole countries.
5. Resistance is a dynamic phenomenon; i.e., any mechanisms already known to exist may change over time.

Continued monitoring is vital to determine whether management recommendations remain valid or need to be revised in light of changing circumstances or new knowledge gained.

6. To promote compliance with management strategies, the countermeasure adopted should be as unambiguous, rational, and simple as possible.

A strategy first implemented on Australian cotton in 1983 to contend with the bollworm, *H. armigera*, illustrates many features of large-scale attempts at resistance management. It was introduced in response to unexpected, but still localized, outbreaks of pyrethroid resistance in *H. armigera* and was based primarily on the concept of insecticide rotation. The threat of pyrethroid resistance was countered by restricting these chemicals to a maximum of three sprays within a prescribed time period coincident with peak bollworm damage. Farmers were required to use alternative insecticide classes at other stages of the cropping season, in order to diversify the selection pressures being applied.

Compliance with this strategy was excellent, and initially it had the desired effect of preventing a systematic increase in the frequency of pyrethroid-resistant phenotypes. Additional recommendations resulting from work on the ecological genetics of pyrethroid resistance, including the targeting of insecticides against newly hatched larvae (when even resistant insects can be killed) and the plowing-in of cotton stubble to destroy resistant pupae overwintering in the soil, undoubtedly contributed to this success. Unfortunately, the restrictions placed on pyrethroid use were inadequate to prevent a gradual, long-term buildup of pyrethroid resistance. As a result, pyrethroids are no longer considered reliable control agents for *H. armigera*, although they remain highly effective against a coexisting species, *H. punctigera*. The strategy has therefore been revised extensively to place greater emphasis on distinguishing between the two *Helicoverpa* species and on the strategic use of nonpyrethroids against *H. armigera*. Transgenic cotton plants expressing Bt toxins have since become the mainstay for controlling *H. armigera* in Australia, and tactics for deploying these without selecting rapidly for Bt resistance were implemented at the outset.

Another strategy incorporating a wide range of chemical and nonchemical countermeasures was introduced on Israeli cotton in 1987. This had the primary objective of conserving the effectiveness of insecticides against the whitefly, *B. tabaci*. Under recommendations coordinated by the Israeli Cotton Board, important new whitefly insecticides were restricted to a single application per season within an alternation strategy optimized to contend with the entire cotton pest complex and to exploit biological control agents to the greatest extent possible. One major achievement of this strategy has been a dramatic reduction in the number of insecticide applications against the whole range of cotton pests, but especially against *B. tabaci*. Sprays against whiteflies now average less than two per growing season compared with more than 14 per season in 1986. Most importantly of all, the strategy has generated an ideal environment for releasing additional new insecticides onto cotton and for managing them effectively from the outset.

Given the threats posed by resistance to pest control, resistance management has become an increasing preoccupation among crop-protection scientists, and also pesticide manufacturers and bodies that regulate the release and use of insecticides. Major agrochemical companies work jointly to publicize resistance through the global Insecticide Resistance Action Committee (IRAC). In the European Union, resistance risk assessment is an integral part of approving new products and maximizing their effective lifespan.

Resistance in Nonpest Species

Compared with its prevalence in arthropod pests, insecticide resistance is still relatively rare among nonpest species including beneficial organisms. However, it has been well documented in a few species of hymenopteran parasitoids and predatory mites, some of which are being exploited in integrated pest management (IPM) systems. Its rarity among beneficial organisms is probably due in part to difficulties in locating hosts and prey (and hence surviving) under exposure to insecticides. The likelihood of resistance developing in beneficial arthropods may be increased if the insects they depend on as prey are already resistant, although this requires further research. It is also likely that, in comparison with herbivorous species, the enzyme systems of predators and parasites are less well adapted to detoxify xenobiotics.

The propensity for beneficial insects to evolve resistance obviously depends on the degree of selectable variation within their populations. Although laboratory selection has been used to enhance low levels of pesticide tolerance found in field populations, there is a concern that such selection applied at artificially low doses will promote polygenic traits that could fragment and dissipate if released into natural populations. However, when substantial resistance has evolved naturally in the field, its mechanisms have tended to be similar to ones found in pest species. Organophosphate-resistant strains of the green lacewing, *Chrysopa scelestes*, have been shown to exhibit increased activity of acetylcholinesterase (AChE) compared with susceptible insects. A carboxylesterase enzyme, very similar in amino acid sequence to that conferring organophosphate resistance in the aphid, *M. persicae*, has been cloned and sequenced from a malathion-resistant strain of the parasitic wasp *Anisopteromalus calandrae*.

As the development of resistance is dependent on the ecology of systems in which it appears, interactions between beneficial and pest species will greatly affect the epidemiology and dynamics of resistance in both. For example, some parasitoids of stored-grain beetles are resistant to insecticides, and it is thought that this adaptation has been encouraged by the fact that their parasitic larvae are sheltered from insecticides by the grain kernels inhabited by their hosts. This is thought to protect a substantial part of the insect life cycle from insecticide selection and ensures that relatively small shifts in insecticide tolerance by the parasitoid afford significant protection against the decreased insecticide doses that do penetrate their defenses.

Natural enemies may contribute to retarding resistance in pest species by exerting sufficient control to decrease the number of insecticide treatments required. Conversely, there are ways in which natural enemies could promote the adaptation of pests to insecticides. For example, the selection pressure for resistance would be increased if weaker,

sublethally affected individuals were more easily preyed upon or parasitized than their fully resistant and therefore unaffected counterparts.

The Special Case of Transgenic Plants

A major development in crop protection with an important bearing on resistance is the release of crop plants genetically engineered to express genes for insecticidal toxins derived from the microbe Bacillus thuringiensis (Bt). Bt cotton and corn is already being grown commercially on a large scale in the USA, Canada, Australia, Mexico, South America, China, India, and South Africa. In 2006, the total area worldwide planted with Bt crops was estimated to exceed 28 million ha. Existing toxin genes in Bt cotton and corn are active specifically against certain key lepidopteran pests (especially bollworms and corn borers); another engineered into potatoes provides protection against the Colorado beetle, *Leptinotarsa decemlineata*.

Aside from their commercial prospects, insect-tolerant transgenic crops offer numerous potential benefits to agriculture. The incorporation of Bt genes into crops offers constitutive expression of toxins in plant tissues throughout a growing season. This could reduce dramatically the use of conventional broad-spectrum insecticides against insect pests as well as remove the dependence of pest control on extrinsic factors, including climate and the efficiency of traditional application methods. Estimates differ, but one study concluded that conventional insecticide use between 1996 and 2005 fell by more than 35 million kilograms of active ingredient, roughly equivalent to the amount applied per year to arable crops in the European Union. However, the high and persistent level of expression also introduces a considerable risk of pests adapting rapidly to resist genetically engineered toxins. To date, there have been no examples of large-scale control failures due to resistance selected directly by commercial transgenic crops, but resistance to conventional Bt sprays (selected in either the laboratory or the field) has been reported in more than a dozen species of insect. Research on the causes and inheritance of such resistance is providing valuable insights into the threats facing Bt plants and the efficacy of possible countermeasures.

Tactics proposed for sustaining the effectiveness of Bt plants have many parallels with ones considered for managing resistance to conventional insecticides. However, they are more limited in scope due to the long persistence and constitutive expression of engineered toxins and the limited diversity of transgenes currently available. Indeed, for existing "single-gene" plants, the only prudent and readily implementable tactic is to ensure that substantial numbers of pests survive in nontransgenic "refuges," composed either of the crop itself or of alternative host plants. The stacking (pyramiding) of two or more genes in the same cultivar, or possibly mixtures of cultivars expressing different single toxins, are potentially more durable options for resistance management. Whatever measures are adopted, it is essential that Bt plants (and their successors expressing other transgenes) are exploited as components of multitactic strategies rather than as a panacea for existing pest-management problems, including those arising from the development of resistance to conventional insecticides.

Concluding Remarks

Over the past 25 years, few areas of entomology have advanced as rapidly or received such widespread attention as that of insecticide resistance. Research on this topic has provided invaluable insights into the origin and nature of adaptations, and these are in turn proving of much broader significance for understanding genetic responses to manmade change in the environment. In many respects the continuing battle against resistance is analogous to an evolutionary "arms race," in this case pitting human ingenuity in discovering new toxins against the adaptive capacity of pest species. Debates as to who will eventually win this race are of secondary importance to the realization that for many species the race is probably unnecessary. A wider adoption of resistance-management practices, especially through greater exploitation of nonchemical measures, would assist with reducing both the economic impact of resistance and any deleterious effects of pest management on biodiversity in general.

See also: Agricultural Invasions. Agriculture, Industrialized. Diversity, Molecular Level. Ecological Genetics. Ecotoxicology. Genetic Diversity. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Pesticides, Uses and Effects of. Population Genetics

References

Castle SJ, Palumbo JC, Prabhakar N, Horowitz AR, and Denholm I (2010) Ecological determinants of *Bemisia tabaci* resistance to insecticides. In: Stansley PA and Naranyo SE (eds.) *Bemisia: Bionomics and Management of a Global Pest*, pp. 423–466. New York: Springer.

Denholm I, Cahill M, Dennehy TJ, and Horowitz AR (1998) Challenges with managing insecticide resistance in agricultural pests, exemplified by the whitefly *Bemisia Tabaci*. *Philosophical Transactions of the Royal Society Series B* 353: 1757–1767.

Denholm I and Rowland MW (1992) Tactics for managing pesticide resistance in arthropods: Theory and practice. *Annual Review of Entomology* 37: 91–112.

Devonshire AL, Field LM, Foster SP, *et al.* (1998) The evolution of insecticide resistance in the peach–potato aphid, *Myzus persicae*. *Philosophical Transactions of the Royal Society Series B* 353: 1677–1684.

ffrench-Constant RH, Pittendrigh B, Vaughan A, and Anthony N (1998) Why are there so few resistance-associated mutations in insecticide target genes? *Philosophical Transactions of the Royal Society Series B* 353: 1685–1693.

Heckel DG, Gahan LJ, Baxter SW, *et al.* (2007) The diversity of Bt resistance genes in species of Lepidoptera. *Journal of Invertebrate Pathology* 95: 192–197.

Horowitz AR, Forer G, and Ishaaya I (1994) Managing resistance in *Bemisia tabaci* in Israel with emphasis on cotton. *Pesticide Science* 42: 113–122.

McKenzie JA (1996) *Ecological and Evolutionary Aspects of Insecticide Resistance*. Austin, TX: R. G. Landes.

Nauen R and Denholm I (2005) Resistance of insect pests to neonicotinoid insecticides: Current status and future prospects. *Archives of Insect Biochemistry and Physiology* 58: 200–215.

Raymond M, Chevillon C, Guillemaud T, Lenormand T, and Pasteur N (1998) An overview of the evolution of overproduced esterases in the mosquito *Culex pipiens*. *Philosophical Transactions of the Royal Society Series B* 353: 1707–1711.

Roush RT (1989) Designing resistance management programs: How can you choose? *Pesticide Science* 26: 423–441.

Tabashnik BE (1994) Evolution of resistance to *Bacillus thuringiensis*. *Annual Review of Entomology* 39: 47–79.

Tabashnik BE, Gassman AJ, Crowder DW, and Carriere Y (2008) Insect resistance to Bt crops: Evidence versus theory. *Nature Biotechnology* 26: 199–202.

Whalon M, Mota-Sanchez D, and Hollingworth RM (eds.) (2008) *Global Pesticide Resistance in Arthropods*. Wallingford, UK: CABI.

Insects, Overview

Brian V Brown, Natural History Museum of Los Angeles County, Los Angeles, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 479–484, © 2001, Elsevier Inc.

Glossary

Hexapod Group including insects and their primitive relatives.

Instar Stage between molts of immature insects.

Pro-, meso-, and metathorax First, second, and third segments of the thorax.

Introduction

The Class Insecta, or the slightly larger Superclass Hexapoda (which additionally includes the orders Collembola, Protura and Diplura), is the world's most species-rich group of organisms, with about 1 million described species. They are found nearly everywhere on earth, including the terrestrial, aquatic, and, to a much lesser extent, marine ecosystems. They are most diverse in tropical forests, where the undescribed fauna has been estimated to comprise 5, 10, 30, or even 50 million species. Hexapods are a well-established monophyletic group, based on the presence of three major body divisions—head, thorax, abdomen—and a single pair of locomotory appendages on each thoracic segment. Some primitive insects have retained appendages on the abdominal segments, but these are much smaller and less functional than those on the thorax. Most of the more derived groups of insects also have wings, usually one pair on each of the mesothoracic and metathoracic segments, but these have been variously lost or modified in some groups, especially the Diptera (flies).

The fossil record of hexapods extends back to the earliest record of terrestrial life, with Collembola and lower insects recorded from the lower Devonian, almost 400 million years ago, and possibly even earlier traces from the Silurian. Insects have been prominent members of the fossil record ever since, with most prominent major groups having been preserved from late Paleozoic or early Mesozoic formations (200–250 million years ago). A summary of fossil insects is given by Kukalová-Peck (1991).

Major Divisions

Insects and their relatives (together referred to as the Hexapoda) are arthropods, with a chitinous exoskeleton and jointed appendages. They grow by molting, periodically shedding their exoskeleton to allow a new larger body to expand and harden. The closest arthropod relatives to the hexapods are still unknown but probably include some or all of the myriapods (centipedes and millipedes).

The insects formerly included all groups currently classified as Hexapoda, but recently authors have separated the Collembola (springtails) and Protura into a separate group called Ellipura or Parainsecta (**Table 1**). The Diplura are placed either within the Insecta or in a separate order of uncertain affinity. All are soft-bodied hexapods usually found in soil or decaying organic material. Most are small, 1 to 5 mm in length, but some Diplura are much larger. The Collembola is the most species-rich group and the one most commonly encountered, as they are often enormously abundant in soil, compost, under rocks, in damp places, and even on the surface of water or snow.

The "true" insects are defined by characters of the structure and musculature of the antenna, among others. There are two relatively primitive groups, the jumping bristletails (Archaeognatha) and the bristletails or silverfish (Thysanura). Both are elongate, wingless creatures with a long, segmented, median "tail" and two shorter, but still prominent cerci. Both groups are scavengers, feeding on vegetable and sometimes animal debris.

The appearance of wings marks the Pterygota. The origin of wings and flight in insects has been controversial, with a number of theories put forward. Most now agree that wings are formed from a movable part of the leg, with its attendant corollary that the protowings were formed in aquatic insects, possibly used for ventilating gills and even for locomotion in water. Extant insects have wings on only the mesothorax and metathorax (the second and third thoracic segments), but fossils exist for which there are also prothoracic wings or wing precursors.

Also unlike the more primitive insects, pterygote insects have hemimetabolus development. The immature stages, often called nymphs, after molting several times, undergo a final molt that produces fully winged, sexually mature adults (except in the Ephemeroptera, which are unique in having a second adult molt; discussed later). More primitive hexapods are ametabolus, with gradual growth through several molts until the sexually mature adult stage, which differs relatively little from the immature forms, is reached.

The most primitive winged insects are sometimes joined in a group called Paleoptera and include the extant orders Ephemeroptera and Odonata. Unlike their more advanced relatives, the neopterans, paleopterans cannot fold their wings down against the body. Both extant orders of Paleoptera have aquatic larvae, but they differ markedly in other aspects of their life histories. Mayflies are generally herbivores, scraping algae and diatoms from objects in the water. When mature,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00075-7

Table 1 Major divisions of extant insects, simplified

Superclass Hexapoda	
CLASS ELLIPURA (= PARAINSECTA)	
Order Collembola (springtails)	6000
Order Protura	500
CLASS DIPLURA	
Order Diplura	800
CLASS INSECTA	
Primitive insect groups	
Order Archaeognatha (jumping bristletails)	350
Order Thysanura (bristletails, silverfish)	350
Pterygota (winged insects)	
Order Ephemeroptera (mayflies)	2200
Order Odonata (dragonflies and damselflies)	5000
Neoptera (advanced pterygotes)	
Order Plecoptera (stoneflies)	2000
Order Blattodea (cockroaches)	4000
Order Isoptera (termites)	2300
Order Mantodea (mantids)	1800
Order Grylloblattodea (ice crawlers)	25
Order Dermaptera (earwigs)	1800
Order Orthoptera (grasshoppers, crickets, katydids)	20,000
Order Phasmatodea (walkingsticks)	2500
Order Embioptera (webspinners)	200
Order Zoraptera	30
Order Psocoptera (bark and book lice)	3000
Order Phthiraptera (lice)	3000
Order Hemiptera (including Homoptera; true bugs)	98,000
Order Thysanoptera (thrips)	4500
Endopterygota (= Holometabola)	
Order Coleoptera (beetles)	350,000
Order Megaloptera (alder flies, dobsonflies)	300
Order Raphidioptera (snakeflies)	175
Order Neuroptera (lacewings, antlions and others)	5000
Order Mecoptera (scorpionflies)	500
Order Siphonaptera (fleas)	2400
Order Strepsiptera (twisted-winged parasites)	500
Order Diptera (flies)	120,000
Order Lepidoptera (butterflies and moths)	160,000
Order Trichoptera (caddisflies)	7000
Order Hymenoptera (sawflies, ants, bees and wasps)	120,000

Important subdivisions are in bold. A detailed classification can be found in Kristensen (1991, Insects of Australia, Vol. 1). Approximate number of described species from various sources.

they emerge from the water in an intermediate, winged stage called the subimago, which lasts only a short time. They molt once more and emerge as winged adults, ready to mate, and die soon thereafter. Dragonflies and damselflies, in contrast, have highly predaceous larvae. There is no subimago stage, as the adults emerge directly and begin relatively long lives as predators of smaller insects. Dragonflies in particular are strong fliers that can travel great distances from water when in search of prey.

The Neoptera are defined by the ability to swivel and fold the wings back at rest. The functional significance of such a modification is obvious—it allows the insect the opportunity to slip into small places to hide from predators and seek food. It has also opened the path to using the wings as protective structures when not in flight, and the forewings of many groups have become hardened or thickened. This is especially evident in the beetles, whose hardened forewings are largely useless for locomotion and are held out of the way during flight.

The relationships among the more primitive Neoptera are not well resolved. One particularly enigmatic group are the stoneflies (Plecoptera), whose aquatic lifestyle is reminiscent of the paleopterans. The larvae live in clear, cool streams or lakes, where they are herbivores or predators. The adults are weaker fliers than Odonata but longer lived than Ephemeroptera, and they feed on algae, decaying vegetation, or detritus.

The Blattodea, Isoptera, and Mantodea are terrestrial insects sometimes unified in a group called Dictyoptera. Their lifestyles are extremely divergent, however. Cockroaches are flattened, usually cryptic and nocturnal scavengers. They vary in shape and size from small, delicate species a few millimeters in length to large, bulky forms 6 cm long. Most are cryptically colored, but some are mimics of apparently distasteful fireflies (Coleoptera: Lampyridae).

Isoptera are termites, the only eusocial group of insects outside of the Hymenoptera. All species feed on cellulose from wood, leaves, and plant debris that is digested in their gut by a symbiotic microbial fauna of flagellate protozoans or spirochaete bacteria. Some species also culture gardens of fungi in their droppings; the fungi are then consumed and their enzymes aid in digestion. Termites have a number of castes, with most individuals in a colony being soft-bodied workers, along with a lesser number of soldiers and a single queen and king. Nests can be in the ground, in rotting wood or in carton nests built on trees or other structures. Ground-nesting species often build prominent mounds that in some species can reach a few meters in height. Although they appear defenseless, termites have a wide variety of chemical and structural modifications that allow them to resist attack by predators, especially ants.

The praying mantids, Order Mantodea, are voracious predators. They have modified, spiny, raptorial forelegs that grasp and crush their prey, usually other insects but sometimes even small vertebrates. They have no venom or other mechanism for killing their prey other than by eating them. Most mantids are green or brown in color, camouflaging them from attack by predators and detection by their prey. Some are flattened and specifically colored to resemble leaves or petals of flowers.

Unlike the well-known cockroaches, termites, and mantids of the Dictyoptera, the grylloblattids are poorly known and only relatively recently recognized as a separate order. They are found in cool climates, usually at high elevations, in the northern hemisphere living mostly under rocks or in caves. In the spring and summer they can forage on the snow surface, feeding on dead insects and plant material. Ice crawlers lack wings, which they have secondarily lost probably as an adaptation to their cryptic lifestyle.

Earwigs, order Dermaptera, are a distinctive group of insects that have a pair of forceps-like appendages at the posterior apex of the abdomen. Wings are absent in some species, and when present are modified, such that the forewings are shortened and hardened and the hindwings are extensively folded, exposing most of the abdomen. Most species are omnivorous, predatory, or herbivorous and are nocturnal. Usually they are encountered when searching through leaf litter, under rocks, in rotten logs, or in any other hidden crevice.

A much larger group is the Order Orthoptera, the familiar grasshoppers, crickets, and katydids. Most of these have notably enlarged hind legs, used to jump great distances or launch the insects in flight to escape predators. Many katydids are masters of camouflage, with green, leaf-like wings bearing markings that resemble leaf veins, fungal infections, and even insect feeding damage. Other orthopterans are equally cryptically colored, a testimony to the relentless selective pressure exerted by sharp-eyed bird predators. Some grasshoppers and katydids are strong fliers, whereas crickets are usually much more likely to stay on the ground. Most male Orthoptera produce sound to attract mates, and the calls of katydids and crickets are an integral part of the evening chorus in warmer parts (or seasons) of the world. Plants are the food of most species, although there are some predators as well.

Like the Orthoptera, the walkingsticks (Order Phasmatodea) are masters of concealment. With their long, stick-like bodies covered in warts, bumps, and spines, walkingsticks can virtually disappear in the appropriate background. One family, whose species are flattened but still extremely well camouflaged, is more appropriately referred to as the leaf insects. All feed on leaves of plants.

The webspinners of the Order Embioptera (webspinners) are an unusual group of small insects that live in silk galleries of their own production. They are narrow bodied, wingless (except for some males), and have the first tarsal segment swollen and filled with silk glands. The galleries they spin are under bark, beneath stones, or in the open in more humid regions; from them they feed on vegetable debris, rotten wood, moss, or lichens.

The Zoraptera are probably the least well known group of insects. They are nondescript, small insects, a few millimeters long, white or pigmented, usually wingless, and with or without eyes. Their life history is little known, although they may be fungus feeders.

A much more commonly seen, but still relatively obscure group are the Psocoptera, the bark and book lice. These small insects, under 1 cm in length, have large, globular heads, rounded bodies, and may or may not bear two pairs of wings. They live on plants, on bark, in leaf litter, or sometimes in human habitations; their food is plant, fungus, and dead insect debris.

Possibly related to some members of the Psocoptera, the Phthiraptera or lice are specialized ectoparasites of mammals and birds. Some species, such as the chewing lice, feed on skin, hair, or feathers, whereas others, such as the sucking lice, suck blood from their hosts. All are small, wingless, flattened insects found only on their hosts.

Outside of the Endopterygota (discussed later) there is only one truly large order of insects: the Hemiptera. This assemblage, in the broad sense, includes groups sometimes named the Heteroptera, or true bugs, and the Homoptera, the cicadas, leaf-, tree-, and planthoppers, aphids, whiteflies, scales, and others. All have distinctively modified mouthparts that are in the form of a piercing-sucking beak that they use to obtain food. Predatory species pierce their prey, usually other insects, injecting digestive enzymes to kill and begin the process of digestion. Predation is restricted to some heteropterans; a few also feed on vertebrate blood. Most Hemiptera are plant feeders, including the majority of the heteropterans and all the homopterans. Most species are terrestrial, although some heteropterans are aquatic. Some homopterans, especially scales and aphids, produce honeydew, as sweet excretion drips down onto leaves and becomes a major source of food for other insects such as flies and wasps. Some are attended by ants, who harvest honeydew and protect the homopterans from predators and parasitoids.

The thrips, Order Thysanoptera, are another unusual order of insects. They are small, slender-bodied, with or without slender, fringed wings. The last tarsal segment of the legs have an inflatable bladder used to improve the grip on the substrate. The mouthparts are asymmetrical, of the piercing-sucking form, used to feed on debris, fungi, or plants. Thysanopterans also have an unusual form of development in which there are two or three pupa-like stages, similar to the pupal stage of holometabolous insects, but which were independently evolved. At least one tropical species has developed a degree of sociality.

The major group of insects, accounting for about 80% of the species, is the Endopterygota, or Holometabola, two names that refer to different aspects of these insects. Endopterygota refers to the development of the wings from imaginal disks of tissue within the pupa of the immature insect, in contrast to the gradual, external development in less derived insects. Holometabola refers to the development of individual insects through a complete metamorphosis, through egg, larva, pupa, and adult stages. The larva is transformed to an often strikingly different adult through the pupal stage, in which the body tissues are almost completely broken down and then reorganized into adult structures. Imaginal disks, sections of tissue sequestered and formed in the larval stage, are retained in the pupa and are the basis of many of the adult structures, such as gonads, legs, and wings. The almost complete separation of the structure of larval and adult insects allows the holometabolous insects to be extremely specialized in each stage: the larva for feeding and the adult for mating and dispersal. This specialization increases the efficiency of each stage and probably results in the incredible success of holometabolous insects.

Within the Endopterygota, there are four orders that are immensely large and successful: the Coleoptera, Diptera, Lepidoptera, and Hymenoptera. Any one of these megadiversity groups alone is among the largest assemblages of organisms in existence, outnumbering any other group except plants. Together, these four insect orders constitute about 40% of all described species of life on earth.

The most successful order of insects in terms of species number is the Coleoptera or beetles. They are the largest group of organisms in the world, even out-numbering plants. The mesothoracic (first) pair of wings of the beetles are greatly strengthened and hardened, such that they are of little or no use in flight but are superb shields when held over the vulnerable abdomen. These hardened forewings, called elytra, are usually held flush over the back of the beetle, making it slippery and difficult to grasp, as well as hard and difficult to crush. In many groups, the overall body form is solid, flattened, and compact, allowing beetles to hide easily, penetrate cryptic habitats, and even to burrow extensively in soil. Although some beetles have secondarily evolved softened or even greatly reduced elytra, there is no doubt that

this modification has allowed beetles to become the preeminent form of insect life on the planet. Primarily a group that is prevalent in warm climates, beetles become relatively less diverse at higher latitudes, but within their range they perform a number of roles, from scavenging, predation, and herbivory to endo- and ectoparasitism. Many species of beetles can be found under stones, in rotting logs, under bark, in freshwater, in fungi, in leaf litter, and on the foliage of plants. Some beetles found on flowers are mimics of bees and wasps, and some beetles found in army ant colonies resemble their hosts to an amazing degree. There are species of rove beetles found in termite nests that carry a passable model of a termite, formed by extensions of their abdomens, over their heads.

There are three orders of insects that are related to beetles but that do not share their same degree of success: the Megaloptera (alder flies, dobsonflies), Raphidioptera (snakeflies), and Neuroptera (lacewings, antlions, and others). The larvae of megalopterans are predatory and aquatic, usually living in clear, running water. The adults are medium to large-sized insects, some of which have enormously enlarged mandibles. In contrast, the larvae of snakeflies are terrestrial, living on the ground in rotting wood or leaf litter and feeding on smaller insects. Adults have a greatly elongated prothorax, giving them a snakelike appearance. Neuropterans are more diverse, with a number of elaborate or bizarre forms intermingled with the more normal-looking lacewings and the damselfly-like antlions and owlflies. All are relatively soft-bodied, with large wings bearing an elaborately reticulate wing venation. Some adults are known to be predatory, as are most of the immatures. The immatures of antlions dig conical pits into which prey fall and are grabbed by the waiting larva. Lacewing larvae are predatory on aphids and other soft-bodied insects found on foliage.

Another large assemblage of holometabolous orders includes the Mecoptera (scorpionflies), Siphonaptera (fleas), Strepsiptera (twisted-winged parasites), and Diptera (flies). The relationships among these orders is in a state of flux, with molecular data giving some indication that fleas are highly modified scorpionflies and that the Strepsiptera, long associated with the Coleoptera, are actually the closest relatives to the Diptera. Previously, some subdivision of the Mecoptera was thought to be most closely related to Diptera; whatever the situation, the monophyly of the scorpionflies relative to some other groups is seriously in question.

Scorpionflies are a relatively small group of insects, so-named because the male genitalia resemble the stinger of a scorpion. Most are terrestrial; the larvae are scavengers and the adults scavengers or predators on small insects. Adults and larvae of the unusual, brachypterous Boreidae feed on mosses, whereas the adults of one genus of Panorpodidae are herbivorous. One family, Nannochoristidae, has aquatic immatures that feed on larval chironomid midges. Some species have elaborate courtship behaviors involving nuptial gifts of dead insects to the female.

Fleas are highly modified, laterally flattened, wingless blood feeders on birds and mammals. Usually they are found in the nests or other areas frequented by their hosts. The larvae are usually free living, feeding on organic detritus and blood from droppings of the adults, although some are obligate ectoparasites. Adults are extremely laterally flattened, allowing them to travel smoothly between the hairs or feathers of their hosts, and often have backwards-pointing combs of spines that makes them difficult to remove. They have strongly developed jumping legs that allow them leap to and from their hosts.

Strepsiptera are extremely unusual parasitoids of other insects. The adults are highly sexually dimorphic, with the males being free-living, winged insects, whereas the females in all but one family are endoparasitoids—wingless, legless, and with only vestigial eyes and appendages of the head. The female body extrudes from the body wall of the host, emitting pheromones to attract males that copulate with the special openings (external genitalia being absent). The larvae are hypermetamorphic, with active, host-seeking first instars that, upon encountering and infecting a host, molt to a legless, relatively inactive form.

As one of the megadiversity groups described earlier, Diptera are found nearly everywhere. Their distinctive innovation has been the reduction of the metathoracic wings to a pair of knoblike halteres that act as gyroscopes in flight. This modification has increased their maneuverability and allowed Diptera to become unparalleled masters of aerial locomotion. Although some adult flies require extensive protein meals to mature eggs and power flight, most of the feeding is done by larvae, which can be predators, scavengers, herbivores, parasitoids and even true parasites. Free-living Diptera larvae are found in soil, rotting vegetation, feeding on and in plants, and sometimes exposed on vegetation. Aquatic forms are found in the silt or sand underlying the body of water (sometimes interstitially), on the surface of rocks, logs, or vegetation, or in the water column. Parasitoids attack mostly other arthropods, but some endoparasites attack mammals.

The orders Lepidoptera (butterflies and moths) and Trichoptera (caddisflies) are close relatives. The lesser of the two—the caddisflies—have aquatic larvae that are found in nearly every type of freshwater environment. Most construct cases or shelters from plant particles, twigs, stones, or sand grains tied together with silk. Some also construct nets to capture debris for food, whereas others are predatory, attacking other aquatic insects. The adults are slender, mothlike insects, often with long, thin antennae.

The Lepidoptera are among the best-known insects, especially the colorful, diurnal group called butterflies. Most of the diversity of the group, however, is in the nocturnal, often drably colored moths, which constitute about 80% of the species of Lepidoptera. The larvae are usually called caterpillars and are best known as voracious feeders on plants. Larval feeding takes place on the surface of the plant or within it (as in leaf miners and stem borers), and almost every plant part—leaves, stems, roots, flowers, and seeds—can be affected. Some species are also predatory and some feed on animal material, such as wool, but almost all species are phytophagus. Adults of most families have mouthparts that are modified to form a long, coiled tube that is used for taking up liquids, usually nectar from flowers. Many adult and larval Lepidoptera are cryptically colored to avoid detection by predators, whereas others are brightly colored to warn of toxic chemicals sequestered in their bodies. Some adults engage in long-distance migrations.

The final large group is the Hymenoptera, the sawflies, ants, bees, and wasps. Like the Diptera, the Hymenoptera have become extremely adept fliers, but have done so by joining the fore- and hindwings with a row of tiny hooks (hamuli) to produce a single functional pair of wings. The most primitive families are phytophagous, but from within these groups have arisen a great diversity of parasitoids, predators, and plant feeders. The parasitoids include some of the smallest known insects, which attack the eggs of their much larger hosts (such as Lepidoptera). Other parasitoids attack a diversity of immature insects, especially those of other holometabolous groups, and develop as endo- or ectoparasitoids. Some are obligatory hyperparasitoids—parasitoids of parasitoids; others oviposit in plant tissue and induce the formation of plant galls, in which the larvae feed. The predatory Hymenoptera attack a wide range of hosts, especially other arthropods, which they often subdue but do not kill with a venomous sting. Larvae of these species have a supply of fresh food to consume when they hatch from an egg laid on the paralyzed prey. Most species hide their prey in some sort of burrow or nest to prevent its being taken by other insects or scavenging animals. Some of these provisioning wasps have moved on to pollen and nectar as food, as in bees. Sociality has evolved a number of times in the Hymenoptera, with the largest and most complex colonies formed by ants and bees.

Conclusion

As this brief survey shows, the largest, most diverse and arguably most successful groups of insects are those that have undergone three major innovations: (a) the development of wings, (b) the ability to fold wings over the body, and (c) the division of the life history into four major stages. Further outstanding success was found in those species that had hardened forewings (beetles), those that transferred the responsibility of flight to a single pair of wings (Diptera), those that extensively exploited flowering plants (Lepidoptera), and those that combined a single functional wing complex with a number of other traits (such as haplodiploidy) that produced the ants, bees, and wasps (Hymenoptera).

See also: Beetles. Butterflies. Flies, Gnats, and Mosquitoes. Grasshoppers and their Relatives. Hymenoptera. Invertebrates, Terrestrial, Overview. Isoptera. Moths. True Bugs and Their Relatives, Diversity of

References

Boudreaux HB (1979) *Arthropod Phytogeny, with Special Reference to Insects*. New York: John Wiley & Sons.

Hennig W (1981) *Insect Phylogeny*. Chichester: John Wiley & Sons.

Naumann I D (chief ed.) (1991) *The Insects of Australia*, vol 1 and 2, Ithaca, NY: Cornell University Press.

In Situ, *Ex Situ* Conservation

Nigel Maxted, University of Birmingham, Birmingham, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Complementary conservation Application of a range of conservation techniques (including *ex situ* and *in situ*) to conserve the target taxon, one technique acting as a backup to another. The degree of emphasis placed on each technique depends on the conservation aims, the type of species being conserved, the resources available, and whether the species has utilization potential.

Conservation Maintenance of the diversity of living organisms, their habitats, and the interrelationships among organisms and their environment.

Ecogeography Analysis of a species' ecological, geographical, and taxonomic characteristics to assist in the formulation of collection and conservation priorities.

Effective population size (Ne) Number of conserved individuals that would undergo the same amount of random genetic drift as the actual population.

***Ex situ* conservation** Conservation of the components of biological diversity outside their natural habitats.

Gap analysis Formalized analysis of a species' ecological, geographical, and taxonomic range compared with that currently sampled for active conservation to identify 'gaps' and therefore future conservation priorities.

Gene pool The total genetic diversity found within an individual species or species group.

Genetic erosion Loss of genetic diversity from a species, often caused by anthropogenic factors.

***In situ* conservation** Conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species in their natural surroundings and, in the case of domesticates or cultivated species, in the surroundings where they have developed their distinctive properties.

Keystone species Usually the dominant species within a habitat that tend to define it physiognomically and ecologically, for example, by determining nutrient and water cycling.

Introduction

The conservation of biodiversity is of critical importance, first because that diversity is under threat of extinction and erosion, but second because that diversity can be of direct and indirect benefit to humankind and is vital for human well-being. Biodiversity benefits humans through the exploitation of animals and plants in agriculture and horticulture, the development of medicinal drugs, and the pivotal roles played by species in the functioning of all natural ecosystems. Biodiversity is also valuable for ethical, esthetic, and recreational reasons. The fundamental importance of these issues to humankind was paramount at the United Nations Conference on the Environment and Development (UNCED) held in Rio de Janeiro, Brazil, in 1992. The Convention on Biological Diversity (CBD), which has subsequently been ratified by 168 countries (as of March, 2012), has as its objectives

> the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilisation of genetic resources. (Article 1: Objectives; Convention on Biological Diversity, 1992).

More recently the CBD (2010) has established the Strategic Plan for Biodiversity 2011–2020, including the 20 Aichi Biodiversity Targets, that aim to address the underlying causes of biodiversity loss, reduce the direct pressures on biodiversity and promote sustainable use, improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity, enhance the benefits to all from biodiversity and ecosystem services and also enhance implementation through participatory planning, knowledge management, and capacity building.

Why Does Biodiversity Need Conservation?

Estimates of the total number of species in the world vary, but approximately 1.75 million species have been scientifically described out of an estimated 14 million species (Groombridge and Jenkins, 2002). It is difficult to estimate rates of species extinction, but the consensus view summarized by Lugo (1988) was that 15–20% of all species could become extinct by the turn of the century. It is even more difficult, if not impossible, to estimate precise levels of genetic erosion, that is, the loss of genetic diversity from extant species. However, the loss of genetic diversity must always be faster than the loss of species because there will be some genetic erosion from the species that remain extant. Loss of any genetic diversity means that the affected plants or animals will not be able to adapt to changing conditions quite so readily; for example, potato cultivars with a narrow genetic base were unable to withstand the infection of late potato blight (*Phytophthora infestans*) in the late 1840s in Ireland, the crop was devastated, and millions were forced to emigrate or starve. Although genetic erosion cannot be quantified accurately, it seems likely that virtually all species are currently suffering loss of genetic variation to varying degrees. Therefore, using Lugo's figures as a starting point, it may be estimated that 25–35% of plant and

animal genetic diversity could possibly be lost over the 12-year period leading up to the year 2000 (Maxted *et al.*, 1997a).

Of course, loss of biodiversity also occurs at other levels than species, for habitats and ecological communities can also be degraded or destroyed. For example, virtually all of the natural grasslands in the USA have been lost since 1942 (Spellerberg, 1996), over 90% of natural wetlands in New Zealand have been lost since European settlement, and Fernside (1990) estimated that world forest loss was proceeding at 20,000 km^2 per year. In a similar study, FAO and UNEP (1991) found that annual rates of forest loss had increased from 113,000 km^2 to 169,000 km^2 per year between 1980 and 1990 in 76 countries. However, a more recent assessment between 2000 and 2010 found the rate of loss had slowed to 130,000 km^2 per year (FAO, 2011).

It is important to realize, however, that both species extinction and genetic erosion can be natural events, just as species and genetic evolution are natural; nature is and it seems has always been dynamic. Yet the current levels of species extinction and genetic erosion are dramatically higher than the so-called background levels that existed hundreds, thousands, and millions of years ago. Humankind now has the capacity to drastically alter the world environment in ways that were not previously possible, and these anthropogenic changes are undoubtedly increasing the rates of species and genetic extinction.

What Threatens Biodiversity?

The threat to biodiversity as a result of anthropogenic changes is not universal for all species. Some species are in greater danger of genetic erosion (or even of complete extinction) than others. These dangers must be evaluated carefully, so that those exposed to the highest risk can be given higher priority for conservation. However, it must also be borne in mind that levels of threat often change rapidly and unexpectedly. Thus, an area may suddenly come under the threat of industrial development, road-building, or logging. The kind of events that lead to genetic erosion or even extinction may be broadly grouped under the general headings of

- destruction, degradation, and fragmentation of natural habitats;
- overexploitation and incidental take;
- introduction of exotic species that compete with, prey on, or hybridize with native species;
- human socioeconomic change and upheaval;
- changes in agricultural practices and land use;
- calamities, both natural and man-made.

But today the greatest potential threat to biodiversity is climate change, even if its precise quantification is still a matter of debate (IPCC, 2007).

It is valuable to establish a system by which the relative threat of genetic erosion can be assessed objectively. Guarino (1995) proposed a model for estimating the threat of genetic erosion that a taxon (wild or cultivated) faces in a particular region. The model may be used without having to visit the region involved, provided some background data are available on the taxon and the area. The model is based on scoring a number of parameters, such as the abundance of the taxon, the level of agricultural development, and the proximity and intensity of various types of human activity to the populations being studied. The higher the score, the greater the risk of genetic erosion and therefore the higher priority for *in situ* or *ex situ* conservation.

The International Union for Conservation of Nature and Natural Resources (IUCN, now known as the World Conservation Union) has developed a system of categories of conservation status, which is based on detailed knowledge of the population dynamics and characteristics of the species concerned, the so-called IUCN Red Data List Categories (IUCN, 2001) (**Figure 1**). Using these categories, for example, a 'Critically Endangered' species would be assigned higher conservation priority than a 'Vulnerable' species. However, IUCN has tended to focus their attention almost exclusively

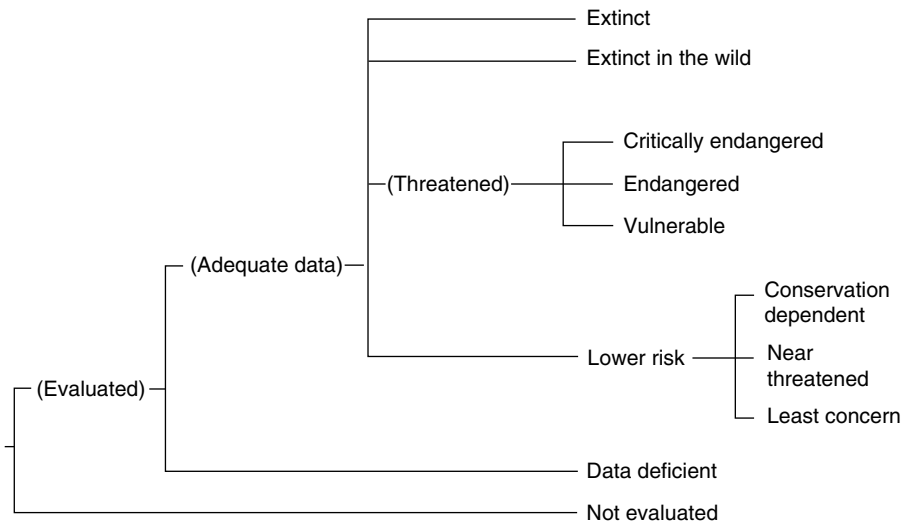


Figure 1 Structure of the IUCN categories of threat. Reproduced from IUCN (2001) *IUCN Red List Categories, Version 3.1.* Gland, Switzerland: IUCN Species Survival Commission.

on species extinction rather than genetic erosion within individual gene pools, and the latter may be of equal importance in terms of overall loss of biodiversity.

What Is Biodiversity Conservation?

Spellerberg and Hardes (1992) stated that biodiversity conservation aims to "maintain the diversity of living organisms, their habitats and the interrelationships between organisms and their environment." But how is this to be achieved?

Conservationists must clearly define and understand the processes involved, and then attempt to develop practical techniques to achieve this objective. When undertaking a particular conservation exercise, a conservationist must use his or her knowledge of genetics, ecology, geography, taxonomy, and many other disciplines to understand and manage the biodiversity that is being conserved.

It is important to stress that conservation is not just about individual plant and animal species, but also includes all aspects of biodiversity from ecosystems (a community of organisms and its abiotic environment), through communities (a collection of species found in a common environment or habitat), species, and populations, to genetic diversity within species. In recent years there has been a differentiation between conservation at the ecosystem and at the genetic levels, and these may be referred to as ecological and genetic conservation, respectively.

Ecological conservation focuses on the conservation of whole communities; although the survival of individuals and the extinction of particular species are a major concern, both are viewed in the larger context of overall community health. This form of whole-community conservation is exemplified by the 'Man and the Biosphere' program (UNESCO, 1996), which established a network of biosphere reserves representing distinct biomes and ecosystems throughout the world. The clear emphasis was on conservation of ecosystems: this program believed that individual species should be conserved as a component of ecosystems rather than on a species-by-species basis.

Genetic conservation focuses more explicitly on particular taxa and attempts to conserve the full range of genetic (allelic) variation within those taxa. The aim of this form of conservation is often utilitarian, for genetic diversity conservation is often linked intimately to human utilization. However, Maxted *et al.* (1997b) stressed the following points: first, species usefulness does not have to be defined in the strictest utilitarian sense, because plants and animals considered of esthetic value are equally worthy of receiving human value as those of immediate use to plant or animal breeders; and second, in many cases individual species cannot be conserved without conserving the communities in which they naturally occur. So the distinction between the two basic forms of conservation is in practice blurred and may be viewed as artificial and of limited semantic importance, because the conservation of ecosystems, species and their genetic diversity are intimately linked. Just as it is difficult to focus conservation effort on the generality of the entire ecosystem, in practice conservationists need to focus on something more tangible that can be monitored and managed, even if only as an exemplar for the ecosystem as a whole.

So to undertake effective conservation, species interactions must be understood as far as possible. Even if the conservation target is a population of a single species, no populations survive in isolation. They exist within a community or ecosystem and interact with other species and the abiotic environment. Obvious examples of interactions include pollinators, seed dispersers, microbial symbionts, herbivores (whether natural or introduced by humans), predators, and pathogens. Thus, even when applying genetic *in situ* conservation, the maintenance of genetic diversity will have to be considered within the context of whole-ecosystem conservation.

The so-called keystone species are important in this context, for these species contribute significantly to the structure of a community or its processes; they are the dominant species. The removal of a keystone species renders other members of a community vulnerable to extinction. Tropical trees that produce a rich food resource in the form of fruit or seeds, for example, can be considered keystone species, as they provide a vital food source for a diverse array of mammals and birds. Generally keystone species play an important part in interactions between different trophic levels, whether they are predators, herbivores, mutualists such as pollinators, or decomposers. So when considering the conservation of any particular species within an ecosystem, one must identify the inherent interactions and ensure their maintenance if the conservation project is to be successful and sustainable.

Methods of Conservation

There is a need to develop appropriate methodologies for biodiversity conservation, particularly in the tropics. The tropical regions of the world have the highest levels of biodiversity, but their fauna and flora are the least well known and are most under threat. Also, tropical nations have few conservationists and often they are insufficiently trained; furthermore, the resources available for conservation activities are relatively limited. To address these issues, the CBD asks nations to:

Promote and encourage research that contributes to the conservation and sustainable use of biological diversity, particularly in developing countries. (Article 12; CBD, 1992)

Maxted *et al.* (1997b) responded to this requirement – to clarify and enhance the methodologies and research programs that currently enable scientists to classify, conserve, manage, and utilize biodiversity – by proposing a model for plant and animal genetic diversity conservation (**Figure 2**). The raw materials of genetic conservation are genes within gene pools, which represent the total diversity of genetic material of the particular taxon being conserved. The product of the gene pool (seeds, eggs, ovules, etc.) is either preserved or utilized as genetic diversity. The processes that link the raw matter and the utilized gene pool represent genetic conservation.

Selection of Target Taxa

Conservation activities will always be limited by the financial, temporal, and technical resources available. Conservation of ecosystems or species has a cost and the effort expended is directly related to how much society values that species and is

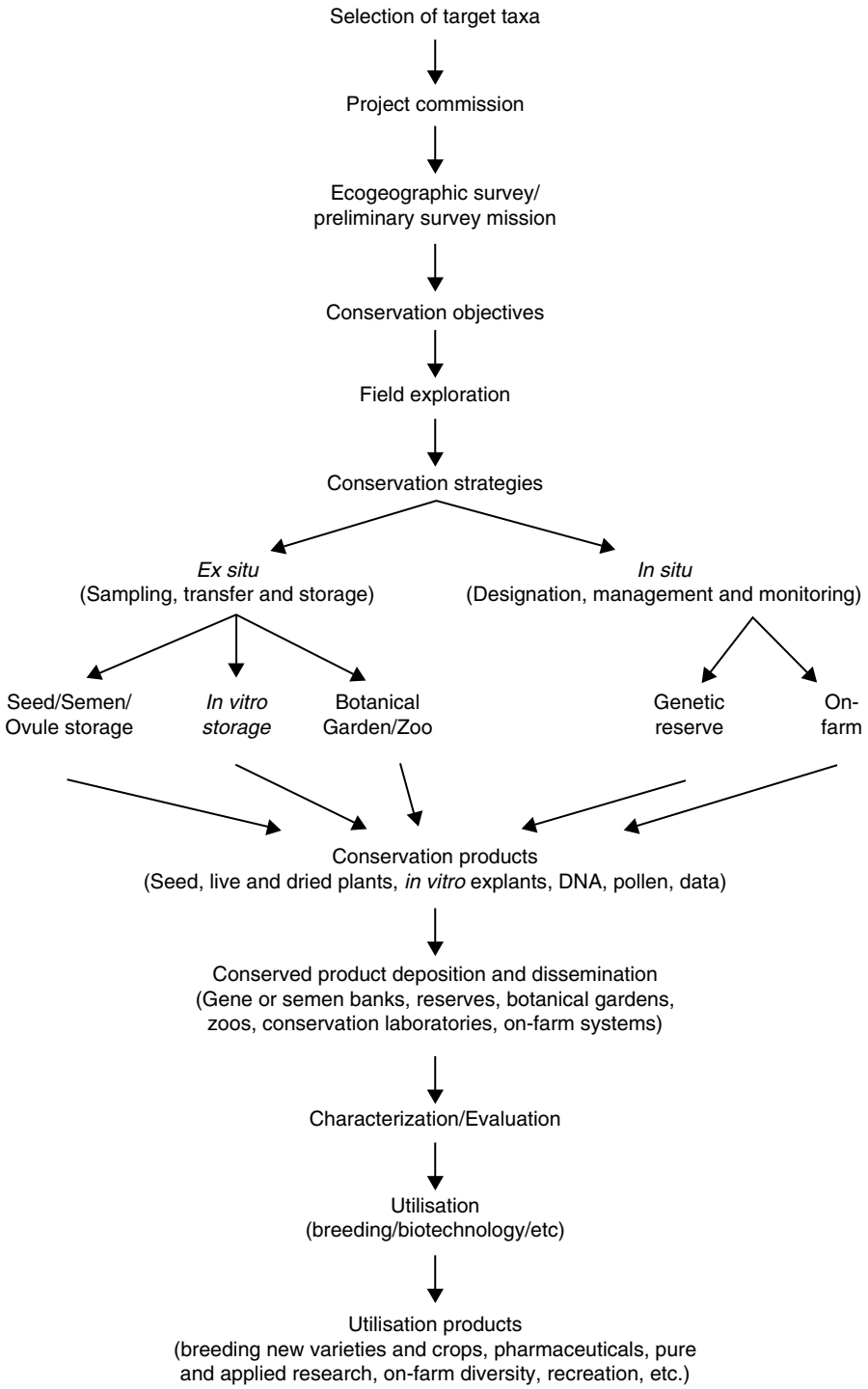


Figure 2 Model of biodiversity conservation. Adapted from Maxted N, Ford-Lloyd BV, and Hawkes JG (1997b) *Plant Genetic Conservation: The In Situ Approach*. London: Chapman & Hall.

therefore willing to pay. It is impossible to actively conserve or monitor all species, so it is important to make the most efficient and effective selection of species on which to focus conservation efforts. This choice should be objective and based on logical, scientific, and economic principles related to the perceived value of the species. Maxted *et al.* (1997a) discussed the sort of factors that provide a species with 'value': current conservation status, potential economic use, threat of genetic erosion, genetic distinction, ecogeographic distinction, national or conservation agency priorities, biological importance, cultural importance, relative cost of conservation, conservation sustainability, and ethical and esthetic considerations.

Rarely will one of these factors on its own lead to a taxon being given conservation priority. More commonly, all or a

range of these factors will be assessed for a particular taxon and then it will be given a certain level of national, regional, or world conservation priority. If the overall score passes a threshold level or is higher than those of competing taxa, the taxon will be conserved; it will then be either collected and conserved *ex situ*, an appropriate reserve will be established, or on-farm conservation will be proposed. Having listed the factors that affect the selection of target taxa in terms of value to society, those related to potential economic use will commonly be given higher comparative value, especially in economically poorer economies where income generation is of the highest priority. This anthropocentric and utilitarian view in the selection of conservation priorities may offend some conservationists, but when financial resources for conservation are, and are likely to remain, limited, and when men, women, and children are still suffering from malnutrition in many parts of the world, there appears to be no practical or ethical alternative to giving those species of most direct use to people the highest value and thus the highest conservation priority.

Project Commission

In practice, once taxa are selected for conservation, the actual conservation activities are necessarily preceded by some form of commission statement. This is likely to establish the objectives of the conservation, specify the target taxa and target areas, state how the material is to be utilized and where the conserved material is to be safely duplicated, and perhaps indicate which conservation techniques are to be employed. A clear, concise commission statement will help to focus subsequent conservation activities.

Who writes the commission? The commission statement may be written by those contracting the conservation or those who actually undertake the conservation work. The commission may vary in taxonomic and geographic coverage, from a systematic collection program for a single species throughout its geographic range to a range of target taxa from a restricted location, for example, onion (*Allium*) species of Central Asia, large cat species worldwide, or chickpeas (*Cicer*) from the Western Tien Shen. In each case, however, a particular group of taxa from a defined geographical area must be considered to be insufficiently conserved (either *in situ* or *ex situ*), of sufficient actual or potential use, or endangered to warrant active conservation.

Ecogeographic Survey, Gap Analysis, and Preliminary Survey Mission

Once the target taxon or group of taxa have been selected and delimited, the conservationist begins to amass and synthesize fundamental biological data to help formulate an appropriate conservation strategy. The synthesis and analysis of these data enable the conservationist to make vital decisions concerning, for example, which taxa to be included in the target group, where to find these taxa, which combination of *ex situ* and *in situ* conservation to use, what sampling strategy to adopt, where to store the germplasm and site the reserve or what captive breeding program would be most successful. If the basic biological data for a particular species, for example, the close lentil relative *Lens culinaris* subsp. *orientalis*, indicate that the species has been previously found on stony slopes at the edge of the Fergana valley in Uzbekistan, then further material of this species is likely to be currently found under similar constraints and is less likely to be found in different habitat types or in far distant regions.

The process of collating and analyzing geographical, ecological, and taxonomic data for use in designing conservation strategies is referred to as ecogeography and was defined by Maxted *et al.* (1995) as:

an ecological, geographical and taxonomic information gathering and synthesis process for a particular taxon. The results are predictive and can be used to assist in the formulation of collection and conservation priorities.

Ecogeographic studies involve the use of large and complex data sets obtained from the literature and from the compilation of passport data associated with herbarium specimens and germplasm accessions. These data are synthesized to produce three basic products: the database, which contains the raw data for each taxon; the conspectus, which summarizes the data for each taxon; and the report, which discusses the contents of the database and conspectus, as well as proposing appropriate conservation strategies. Ecogeographic techniques enhance the efficiency of crop relative and wild species conservation because they enable the conservationist to identify clearly the geographical regions and ecological niches that the taxon inhabits, and so not only identify areas with high numbers of target taxa, but also areas that contain high taxonomic or genotypic diversity of taxa, uniqueness of habitat, economic or breeding importance, and so on.

The ecogeographic study may form part of a more formulized gap analysis, which originally was used in a more limited sense to identify areas in which selected elements of biodiversity are under-represented (indigenous forests, particularly on small islands rich in endemic species) (see Margules, 1989; Balmford, 2003), but the methodology can be adapted for more general species conservation use (Maxted *et al.*, 2008) involving four steps: (1) identification of priority taxa, (2) identification of the ecogeographic breadth and complementary target taxon hotspots using genetic diversity, distribution, and environmental data, (3) match current *in situ* and *ex situ* conservation actions with the identified genetic diversity, ecogeographic breadth and complementary species hotspots to identify the so-called conservation 'gaps,' and so (4) formulate a revised *in situ* and *ex situ* conservation strategy for the target taxon. Gap analysis has rapidly established itself as the methodology for species conservation planning.

If the available ecogeographic data for the target taxon are limited, whether undertaking an ecogeographic survey or gap analysis the conservationist will not have sufficient background biological knowledge to formulate an effective conservation strategy. In this case it would be necessary to undertake an initial survey mission to gather the novel ecogeographic data required on which to base the actual conservation strategy. The survey mission may be in the form of 'coarse grid sampling,' which involves traveling throughout a likely target region and sampling sites at relatively wide intervals over the whole region. The precise size of the interval between sites depends on the level of environmental diversity

across the region, but it may involve sampling every 1–50 km. The population samples and data collected during this mission can then be used to formulate further conservation priorities and to develop an appropriate strategy, thus providing the same result as the ecogeographic survey for groups that are better biologically understood.

Conservation Objectives

The products of the ecogeographic survey, gap analysis, or survey mission provide a basis for the conservationist to formulate future conservation priorities and strategies for the target taxon. Within the target area, zones of particular interest may be identified, for example, areas with high concentrations of diverse taxa, low or very high rainfall, or high frequency of saline soils or extremes of altitude or exposure. In general, areas with very distinctive characteristics are likely to contain plants with distinct genes or genotypes. If a taxon is found throughout a particular region, then the researcher can use the ecogeographic data to positively select a series of diverse habitats to designate as reserves. If a taxon has been found at one location but not at another with similar ecogeographic conditions, then the conservationist may suggest that these similar locations should be searched. Within the target taxon, specific variants may be identified that warrant conservation priority, for example, species that have previously unrecognized utilization potential, populations that are particularly in danger of genetic erosion, or those that had not previously been conserved.

The conservationist must set out a clear, concise statement of the proposed conservation strategy for the target taxon and, if appropriate, prioritize actions. These may have been established in the project commission, but if not the conservationist should undertake the task. This should answer questions such as: Which populations require conservation? Can local farmers play a part in conservation activities? Do population levels require close monitoring? Should a national or international collecting team be directed to collect the priority target taxa? What conservation strategy or strategies are appropriate? What combination of conservation techniques is appropriate or is a more detailed study required before any of these questions can be answered?

Field Exploration

Once the conservation objectives have been clarified, whichever conservation strategy is to be applied, the ecogeographic information is used to locate and identify the general locality of the animal or plant populations that are to be conserved. The ecogeographic data will rarely be sufficiently comprehensive to precisely locate actual populations. Therefore, the preparatory element of conservation activities will be followed by field exploration, during which actual populations are located. Ideally, populations of the target taxon that contain the maximum amount of genetic diversity in the minimum number of populations will be identified, but how is this goal to be achieved? Commonly there will be too much diversity in both crops and wild species to conserve all their alleles, even if these were known then or at some future time. Thus the conservationist must attempt to conserve the range of diversity that best reflects the total genetic diversity of the species. How many animals or plants must be sampled, which specimens and what pattern of sampling is appropriate? To answer these specific questions the conservationist should know the amount of genetic variation within and between populations, local population structure, the breeding system, taxonomy and ecogeographic requirements of the target taxon, and many other biological details. Some of this information will be supplied following the ecogeographic survey, but some will remain unavailable. Therefore, the practice of field exploration will be modified depending on the biological information on the target taxon and target area that is available.

For a botanical project, the field botanist should select populations if they are found on the periphery of the target taxon's distribution or those that contain morphological or ecological variants. Atypical populations or those growing under atypical conditions may possess genes or alleles that are unknown or extremely rare in the target taxon's center of diversity, and this material possibly contains genetic variation that is of special use to breeders (e.g., disease or pest resistance or adaptation to soil or climate that is unknown in the crop itself).

Conservation Strategies

There are two basic conservation strategies, each composed of various techniques, that the conservationist can adopt to conserve genetic diversity once it has been located. The two strategies are *ex situ* and *in situ* conservation. Article 2 of the CBD (1992) provides the following definition of these categories:

In situ conservation means the conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species in their natural surroundings and, in the case of domesticates or cultivated species, in the surroundings where they have developed their distinctive properties.

Ex situ conservation means the conservation of components of biological diversity outside their natural habitats.

There is an obvious fundamental difference between these two strategies: *in situ* conservation involves the location, designation, management, and monitoring of target taxa where they are encountered, whereas *ex situ* conservation involves the location, sampling, transfer, and storage of target taxa from the target area. Because of this fundamental difference, there is little overlap between the two strategies. However, the two techniques should not be seen as alternative, they are complementary and should be applied in tandem to ensure maximum maintenance of diversity. The two basic conservation strategies may be further subdivided into the following specific techniques, which are discussed in sections *Ex Situ* Techniques and *In Situ* Techniques:

Ex Situ Seed/embryo storage *In vitro* storage Semen/ovule/pollen/deoxyribonucleic acid (DNA) storage Field gene bank/livestock parks Botanic/zoological garden, and *In Situ* Genetic reserve On-farm

Ex Situ Techniques

In *ex situ* conservation, genetic variation is maintained away from its original location and samples of a species, subspecies,

or variety are taken and conserved either as living collections of plants or animals in field gene banks, botanic or zoological gardens, and arboreta, or as samples of seed, semen, ovules, tubers, tissue explants, pollen, or DNA maintained under special artificial conditions.

Seed/Embryo Storage Conservation

Ex situ seed/embryo collection and storage is the most convenient and widely used method of genetic conservation (**Figure 3**). Seeds and embryos are the natural dispersal, storage, or generative organs for the majority of species. This technique involves collecting samples from individuals or populations and then transferring them to a gene bank for storage, usually at subzero temperatures. The procedure used for the bulk of orthodox-seeded plant species is to dry the seeds or embryos to a suitable moisture content (5–6%) before freezing at $-20\,^{\circ}C$, but this method is only suitable for species that can be dried and stored at low temperature without losing viability. The advantages of this technique are that it is efficient and reproducible, and feasible for short-, medium-, and long-term secure storage. However, the disadvantages are that there are problems in storing recalcitrant-seeded plant species. The latter species cannot be dried and frozen in the way used for orthodox seeds, because they rarely produce seed or are normally clonally propagated.

Botanical/Zoological Garden Conservation

Historically, botanical or zoological gardens were often associated with physic or medicinal gardens or displays of single specimens of zoological curiosities, and as such they did not attempt to reflect the genetic diversity of the species. These gardens now hold living collections of species that were collected in a particular location and moved to the garden to be conserved. The advantage of this method is that gardens do not have the same constraints as many other conservation agencies; they have the freedom to focus on wild species that may otherwise not be given sufficient priority for conservation. Yet there are two disadvantages to this technique. The first is that the number of species that can be genetically conserved in a botanical or zoological garden will always be limited because of the available space. The majority of these gardens are located in urban areas in temperate countries, and at their present sites most expansion would be prohibitively expensive. The majority of botanical and animal diversity is located in tropical climates, yet because most botanical and zoological gardens are in temperate countries, the collections must be kept in expensive greenhouses or other facilities, which also limits the space available. The second disadvantage is related to the first, namely very few individuals of each species can be held, and this severely restricts the range of genetic diversity found in the wild that is protected. However, if the target species is very near extinction and only one or two specimens remain extant, this objection of course does not hold.



Figure 3 Collecting seed from Tajikistan for *ex situ* conservation.

In Vitro *Conservation*

In vitro conservation involves the maintenance of ex-plants in a sterile, pathogen-free environment, and it is widely used for vegetatively propagated and recalcitrant-seeded species. This method offers an alternative to field gene banks. It involves the establishment of tissue cultures of accessions on nutrient agar and their storage under controlled conditions of either slow or suspended growth. The main advantage is that it offers a solution to the long-term conservation problems of recalcitrant, sterile, or clonally propagated species. The main disadvantages are the risk of somaclonal variation, the need to develop individual maintenance protocols for the majority of species, and the relatively high-level technology and high cost required. The best answer for cheap, long-term *in vitro* conservation has been increasingly shown to be cryopreservation (Hoyt, 1988), that is, the storage of frozen tissue cultures at very low temperatures, for example, in liquid nitrogen at $-196\,^{\circ}C$.

Field Gene Bank/Livestock Park Conservation

The conservation of germplasm in field gene banks or livestock parks involves the collecting of plant or animal specimens from one location and the transfer and conservation at a second site. It has traditionally been the method for recalcitrant- (whose seeds cannot be dried and frozen without loss of viability) or sterile-seeded plant species or for those species for which it is preferable to store clonal material. Field gene banks are commonly used for species such as cocoa, rubber, coconut, mango, coffee, banana, cassava, sweet potato, and yam. Livestock parks or rare breed centers, as distinct from zoos, emphasize captive breeding programs and therefore genetic conservation. The advantages of field gene banks and livestock parks are that the species are easily accessible for utilization and evaluation can be undertaken while the material is being conserved. The disadvantages are that the material is restricted in terms of genetic diversity, is susceptible to pests, disease, and vandalism, and may require large areas of land.

Pollen/Semen/Ovule/DNA Conservation

The storage of pollen grains is possible under appropriate conditions that allow their subsequent use for crossing with living plant material. It may also be possible in the future to regenerate haploid plants routinely from pollen cultures, but no generalized protocols have been developed yet. The development of artificial insemination techniques in recent years has made semen and ovule storage routine, especially for

domesticated animals. The storage of DNA under prescribed conditions can easily and inexpensively be achieved given the appropriate level of technology, but the regeneration of entire organisms from DNA cannot be envisaged at present, although single or small numbers of genes could subsequently be utilized. The advantage of pollen storage is that it is low cost and simple, but the disadvantage is that only paternal material would be conserved, and with DNA storage there are problems with subsequent gene isolation, cloning, and transfer.

In Situ Techniques

In situ techniques involve the maintenance of genetic variation at the location where it is encountered, either in the wild or in traditional farming systems. The majority of existing nature reserves and natural parks were established to conserve megafauna or to protect esthetically beautiful landscapes, but even today few have less glamorous species conservation as their primary goal, let alone the genetic conservation of species (Hoyt, 1988).

Genetic Reserve Conservation

Conservation of wild species in a genetic reserve involves the location, designation, management, and monitoring of genetic diversity in a particular natural location (**Figure 4**). This technique is the most appropriate for the bulk of wild species, because it can, when the management regime is minimal, be relatively inexpensive. Whether dealing with plants or animals, the objective is to contain the minimum number of individuals that can maintain genetic diversity within the species. If too few individuals are protected, genetic diversity will decline over time, and if too many are protected, resources may be wasted in managing the large population. To guide such efforts, conservationists will need to estimate the effective population size (Ne), that is, the number of conserved individuals that would undergo the same amount of random genetic drift as the actual population. Genetic reserves are appropriate for animals as well as for orthodox and nonorthodox seeded plant species, because numerous taxa can be protected in a single reserve that allows the continued evolution of species. However, the disadvantages are that the conserved material is not immediately available for human exploitation and, if the management regime is minimal, little characterization or evaluation data may be available. In the latter case, the reserve manager may even be unaware of the complete specific composition of the reserve that he or she is managing.



Figure 4 Surveying plant populations for an *in situ* reserve in Turkey.

On-Farm Conservation

Farmer-based conservation involves the maintenance of traditional crop or animal breeds or cultivation systems by farmers within traditional agricultural systems. On traditional farms, what are generally known as 'land-races' of plants are sown and harvested, and each season the farmers keep a proportion of harvested seed for re-sowing. Traditional breeds of domestic animal are maintained by interbreeding within and between local village stocks. Thus the land-race or breed is highly adapted to the local environment and is likely to contain locally adapted alleles that may prove useful for specific breeding programs. Home garden plant conservation is a closely related variant of on-farm conservation of land-races but on a smaller scale. It involves the cultivation of more species-diverse material in home, kitchen, backyard, or dooryard gardens. These home gardens focus on medicinal, flavoring, and vegetable species (e.g., tomatoes, peppers, digitalis, mint, thyme, and parsley). The overall advantage of the on-farm technique is that it ensures the maintenance of highly adapted land-races and breeds and those wild species that traditional agriculture often depends on. However, these land-races or traditional breeds may yield less than their modern counterparts, and so traditional farmers may require some subsidy and possibly monitoring to ensure continued farming. It should be noted that contemporary economic forces tend to act against the continued farming of ancient land-races and breeds, which are currently suffering rapid genetic erosion; many face imminent extinction. A back-up system of *ex situ* conservation is therefore essential, as discussed in the section Sustainable and Integrated Biodiversity Conservation.

Community-Based Conservation

When applying all of the conservation techniques discussed here, professional conservationists have too often failed to appreciate the role that local communities have successfully played in conserving animal and plant diversity within their local environment in the past and how they might continue to contribute in the future. It is now generally accepted that the present-day wealth of domesticated and nondomesticated biodiversity would not exist were it not for the conscious effort of local communities over millennia to conserve biodiversity in all its forms. For example, indigenous farmers in the Andes maintain a gene pool of over 3000 varieties of potatoes representing eight cultivated species, and in Papua New Guinea approximately 5000 varieties of sweet potato are cultivated, with a single farmer growing up to 20 varieties in one garden (McNeely *et al.*, 1995). For wild species, Prance *et al.* (1995) showed that four groups of Amazonian Indians use up to 79% of the tree species in their home ranges, and Milliken *et al.* (1992) in a similar study found that 81% of tree and vine

species were utilized; this number rises to 86% when other categories of plants are added from literature sources.

So local communities have had and continue to have an essential role in biodiversity conservation. They not only continue to conserve by using traditional practices for their own future direct and indirect benefit, but also increasingly work in collaboration with professional conservationists to conserve broad-based biodiversity for the benefit of their host countries and humankind as a whole. Specifically, collaboration involving conservationists and local communities increases the overall efficiency of 'professional' conservation, because local communities have a broader local knowledge base concerning the animal and plant species found in their area. Local communities are therefore able to assist in the development of a more practical, focused, and hopefully efficient approach to locally targeted conservation. The employment of a collaborative approach also empowers local people and engenders increased pride in native biodiversity and its conservation. In this way, rather than deferring responsibility to outside science-based experts, they can retain environmental responsibility and take greater pride in maintaining their own environment.

Conservation Products

The products of conservation activities are primarily conserved germplasm (seed, embryos, semen, and ovules), live plants and animals, dried plants, cultures, and conservation data. *Ex situ* conserved orthodox seed or animal semen and ovules are commonly held in gene or semen banks at subzero temperatures and, for seed, low moisture content to prolong their life. Live plants or animals are conserved in genetic reserves, field gene banks, botanical or zoological gardens, or parks and research laboratories. Germplasm that is stored in a suspended form, such as tissue, pollen, or DNA, is kept as cultures in specialist laboratory facilities. Dried voucher plant specimens are held in herbaria and linked to specific samples of germplasm, and are as much as possible made representative of the conserved populations. Conserved material is ideally associated with a range of passport data, which detail the taxonomic, geographical, and ecological provenance of the material. All passport data should be entered into a database and made available for the management of the material, the formulation of future conservation priorities and strategies, and any exploitation. The various conservation products, where they are stored, and where they should be duplicated are presented in **Table 1**.

Conserved Product Dissemination

The conservation products are either maintained in their original environment or deposited in a range of *ex situ* storage facilities. Whether the germplasm, voucher specimens, or passport data are conserved *in situ* or *ex situ*, to ensure their safety they should ideally be duplicated in more than one location. The distribution of duplicate sets of material avoids accidental loss of the material due to fire, economic or political difficulties, warfare, or other unforeseen circumstances. Duplication of the data is relatively easy from the conservation database, and copies should be held by the commissioning agency, relevant host country institutes, and other interested parties.

Biodiversity Utilization

As discussed earlier, if conservation is to be long term and sustainable there should be an intimate linkage between conservation and utilization. The products of conservation, whether they be 'living' or 'suspended,' should be made available for use by humankind for the various forms of utilization. Conservation can be seen as the safekeeping of preserved material, so that the material is available at a future date. In certain cases the material can be used directly, say in the selection of forage accessions or local domesticated animal breeds, where little breeding is undertaken. The conserved material may also be used in reintroduction programs where the traditional breed or land-race has been lost locally owing to civil unrest or the application of perverse government incentives that encourage the alteration of traditional practices.

More commonly, the first stage of utilization will involve the recording of genetically controlled characteristics (characterization) and the plant material may be grown out

Table 1 Conservation products and their storage and duplication sites

Conservation product	*Storage site*	*Duplication site*
Plant germplasm (seed, vegetative organs, etc.)	Gene bank	National, regional, and international gene banks, duplication with other conservation techniques
Animal germplasm (semen, ovules, eggs, embryos, etc.)	Gene bank	National, regional, and international gene banks, duplication with other conservation techniques
Live plants	Field gene bank, botanical garden, genetic reserve, on-farm	Duplication with other conservation techniques, for example, gene bank storage of seed
Live animals	Zoological garden, genetic reserve	Duplication with other conservation techniques, for example, storage of germplasm
Dried plants or preserved animals	Herbarium or museum	National, regional, and international herbaria or museums
Explants or plantlets	Tissue culture	Duplication with other conservation techniques, for example, gene bank storage of seed
DNA and pollen	Various cultures	Duplication with other conservation techniques, for example, gene bank storage of seed
Conservation data	Conservation database	Duplication with other national, regional, and international conservation agencies

under diverse environmental conditions to evaluate and screen for drought or salt tolerance, or be deliberately infected with diseases or pests to screen for particular biotic resistance (evaluation). As well as actual animal or plant characterization, increasingly geographic information systems are being used to help predict which accessions have desired traits, for example, climate change resilience in accession collected from arid areas. The biotechnologist will be screening for single genes, which when located may be transferred into a host organism to generate more rapid growth, for example. The biochemist (bioprospector) will be screening for particular chemical products that may be of use to the pharmaceutical industry. The products of utilization are therefore numerous, including new varieties, new crops, improved breeds, and pharmaceuticals as well as more nebulous but equally valuable products such as a beautiful or more diverse environment for human recreational activities.

Sustainable and Integrated Biodiversity Conservation

Having discussed in detail what constitutes biodiversity and how that biodiversity can be conserved, the point should be made that any biodiversity conservation program should be sustainable and integrated. Each conservation technique has its advantages and disadvantages. The two strategies of *ex situ* and *in situ* conservation should not be seen as alternatives or in opposition to one another, but rather as being complementary, as stated in Article 9 of the CBD (1992). One conservation strategy or technique can act as a backup to another, the degree of emphasis placed on each depending on the conservation aims, the type of species being conserved, the resources available, and whether the species has utilization potential. The efficacy of adopting an integrated approach to conservation, or as some have called it a 'holistic' approach (Withers, 1993), is now well established as the only sustainable option. Therefore, when formulating an overall conservation strategy for a species, conservationists should think in terms of applying a combination of the different techniques available, including both *in situ* as well *as ex situ* techniques, where the different methodologies complement each other. It may be helpful to think of the various techniques as pieces in a jig-saw puzzle that will complete the overall conservation strategy and thus ensure the maintenance of plant or animal genetic diversity.

The adoption of an integrated approach requires the conservationist to consider the characteristics and needs of the particular gene pool being conserved, and then to assess which of the strategies or combination of techniques offers the most appropriate option to maintain genetic diversity within that taxon. To formulate the conservation strategy, the conservationist may also need to address not only biological questions but also the practical and political ones: What are the species' storage characteristics? What do we know about its breeding success in captivity? Do we want to store the germplasm over the short, medium, or long term? How important is the species? Where is the species located and how accessible is it/does it need to be? Are there legal issues relating to access? How good is the infrastructure of the established reserves? What backup is necessary and desirable? How does the species conservation strategy fit within the local community development program? Given answers to these questions, the appropriate combination of techniques to conserve the gene pool can be applied in a pragmatic and balanced manner. The integration of conservation and community development is an important point to stress. We cannot expect local communities to altruistically forgo development for the benefit of a more abstract greater good, and so the practical application of the conservation strategy may need to be a compromise between scientific protocols and meeting the needs and desires of local people. Therefore the actual combination of techniques will be formulated afresh for each species or group of species, demonstrating the flexibility of the integrated approach.

Sustainability in the sense of continuance is a fundamental concept for conservation. Whether seed, semen, or embryos are collected for *ex situ* conservation in a gene bank, animals are incorporated into a captive breeding program, or a habitat is designated as a reserve, each option has a financial cost and it would be a waste of limited conservation funds not to ensure that the conservation project is sustainable at least in the medium term. If the species or genetic material is lost from a reserve, the resources expended on establishing the reserve would have been wasted and the cost of rehabilitating populations using materials stored *ex situ* would have to be considered. The latter option is commonly expensive and may require extensive research to ensure that the reintroduced animals or plants do not likewise go extinct. Unfortunately, many conservation projects are funded on a short-term basis, so it is essential that an effective project exit strategy is developed so the conservation program itself is sustainable.

Not only is it necessary to integrate the different conservation strategies and techniques, and to involve the local community in a sustainable conservation project, but it is also important to integrate the different potential agencies involved. This is particularly true for large *in situ* reserve projects where the project may naturally span national borders and professional disciplines. In these cases, the project team must ensure that the local, provincial, national, regional, and international conservation agencies, as well as professionals from the different disciplines involved, such as environmentalists, foresters, agriculturalists, and politicians, work together to promote the success of the conservation project.

Appendix

List of Courses

1. Conservation Biology
2. Genetic Resources
3. Biodiversity
4. Taxonomy

See also: Endangered Ecosystems. Environmental Ethics. Gene Banks. Genetic Diversity. Keystone Species. Species Interactions

References

Balmford A (2003) Conservation planning in the real world: South Africa shows the way. *Trends in Ecology and Evolution* 18: 435–438.

Convention on Biological Diversity (1992) *Convention on Biological Diversity: Text and Annexes*, pp. 1–34. Montreal: Secretariat of the Convention on Biological Diversity.

Convention on Biological Diversity (2010) *Strategic Plan for Biodiversity 2011–2020, Including Aichi Biodiversity Targets*. Montreal: Secretariat of the Convention on Biological Diversity.

FAO (2011) *Tropical Forest Resources Assessment Project*. Rome: FAO.

FAO and UNEP (1991) *Tropical Forest Resources Assessment Project*. Rome: FAO.

Fernside PM (1990) The rate and extent of deforestation in Brazilian Amazonia. *Environmental Conservation* 17: 213–226.

Groombridge B and Jenkins MD (2002) *World Atlas of Biodiversity: Earth's Living Resources in the 21st Century*. Cambridge: UNEP-WCMC.

Guarino L (1995) Assessing the threat of genetic erosion. In: Guarino L, Ramanatha Rao V, and Reid R (eds.) *Collecting Plant Genetic Diversity: Technical Guidelines*, pp. 67–74. Wallingford, United Kingdom: CAB International.

Hoyt E (1988) *Conserving the Wild Relatives of Crops*. Rome: IBPGR/IUCN/WWF.

Intergovernmental Panel on Climate Change (IPCC) (2007) *Fourth Assessment Report Climate Change 2007: Synthesis Report*. Geneva, Switzerland: Intergovernmental Panel on Climate Change.

IUCN (2001) *IUCN Red List Categories, Version 31*. Gland, Switzerland: IUCN Species Survival Commission.

Lugo AE (1988) Estimating reductions in the diversity of tropical forest species. In: Wilson EO (ed.) *Biodiversity*, pp. 58–70. Washington, DC: National Academy Press.

Margules CR (1989) Introduction to some Australian developments in conservation evaluation. *Biological Conservation* 50: 1–11.

Maxted N, Dulloo E, Ford-Lloyd BV, Iriondo J, and Jarvis A (2008) Genetic gap analysis: A tool for more effective genetic conservation assessment. *Diversity and Distributions* 14: 1018–1030.

Maxted N, Ford-Lloyd BV, and Hawkes JG (1997b) *Plant Genetic Conservation: The In Situ Approach*. London: Chapman & Hall.

Maxted N, Hawkes JG, Guarino L, and Sawkins M (1997a) The selection of taxa for plant genetic conservation. *Genetic Resources and Crop Evolution* 44: 337–348.

Maxted N, Van Slageren MW, and Rihan J (1995) Ecogeographic surveys. In: Guarino L, Ramanatha Rao V, and Reid R (eds.) *Collecting Plant Genetic Diversity: Technical Guidelines*, pp. 255–286. Wallingford, United Kingdom: CAB International.

McNeely JA, Gadgil M, Leveque C, Padoch C, and Redford K (1995) Human influences on biodiversity. In: Heywood VH (ed.) *Global Biodiversity Assessment*, pp. 711–821. Cambridge, United Kingdom: Cambridge University Press.

Milliken W, Miller RP, Pollard SR, and Wandelli EV (1992) *Ethnobotany of the Waimiri Atroari Indians of Brazil*. Kew, United Kingdom: Royal Botanic Gardens.

Prance GT, Balée W, Boom BM, and Carneiro RL (1995) Quantitative ethnobotany and the case for conservation in Amazonia. In: Evans-Schultes R and von Reis S (eds.) *Ethnobotany: Evolution of a Discipline*, pp. 157–174. Portland, Oregon: Dioscorides Press.

Spellerberg IF (1996) *Conservation Biology*. Harlow, United Kingdom: Longman Group Ltd.

Spellerberg IF and Hardes SR (1992) *Biological Conservation*. Cambridge, United Kingdom: Cambridge University Press.

UNESCO (1996) *Biosphere Reserves: The Seville Strategy and the Statutory Framework of the World Network*. Paris: UNESCO.

Withers LA (1993) Conservation methodologies with particular reference to *in vitro* conservation. *Proceedings of the Asian Sweet Potato Germplasm Network Meeting, Guangzhou, China*, pp. 102–109. Manila: CIP.

International Organizations and Biodiversity

Margaret Goud Collins, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

© 2013 Elsevier Inc. All rights reserved.

Introduction

Biodiversity is an inherently international concern. Species distributions and the ecosystem services that they support are not defined by national boundaries. As a reflection of societal interests in those ecological benefits, international organizations have grown up to study, develop, regulate, and conserve the systems on which much of human well-being depends. These organizations, many of which were founded in response to the need for international cooperation in such areas as wildlife and wilderness conservation, scientific research, agriculture, forestry, water resources, aquatic resources, trade, pharmaceutical development, and economic development, have incorporated biodiversity concepts into their missions. Moreover, a number of issues are overlapping concerns of these different groups – invasive species control, for example, is where agriculture, forestry, and wildlife conservation meet up with trade.

As the interest in these issues has increased, and the complexity of, and threats to, global biological systems have become clearer, the diverse scientific communities studying science and policy related to biodiversity have developed stronger international ties. A central tenet of the new field of "sustainability science," which focuses on the complex, dynamic interactions between nature and society, holds that "societies are complex adaptive systems, composed of individual agents who have their own priorities, and who value the macroscopic features of their societies differently. Resolving those competing perspectives is at the core of addressing sustainability" (Levin and Clark, 2010). The emphasis on social–ecological systems makes biodiversity a central element in the global network of institutions researching sustainability. However, the array of international organizations concerned with biodiversity and ecosystem services can themselves be viewed as a complex adaptive system, forming coalitions and adopting interests in response to scientific developments, economic forces, policy initiatives, and environmental changes.

International cooperation is crucial for the scientific investigation and conservation of biodiversity, and for the maintenance of the ecosystem services that biodiversity supports. International organizations foster and coordinate that cooperation, creating a complex network of governmental and nongovernmental international bodies to finance and conduct biodiversity research and conservation activities. Many of them participate in the provision of policy advice through national and international initiatives, undertaken in response to unprecedented ecosystem challenges stemming from economic development and climate change. Other organizations focus on the industries and public services that depend on biodiversity and ecosystems. Increasingly, the interests and mandates of these organizations overlap, requiring collaboration and cooperation among – and even within – the organizations, as it is not unusual for different bureaus within the same association to address different sides of the biodiversity question.

UN Leadership on International Biodiversity Issues

Within this network of organizations, the United Nations (UN) serves as a hub, both for intergovernmental cooperation and for partnerships with governments, nongovernmental organizations (NGOs), and industry. In each of the subject areas discussed here, UN agencies are crucial for both coordinating and funding international activities. Biodiversity-related issues have been among UN concerns since its earliest days, and the role of the United Nations Educational, Scientific and Cultural Organization (UNESCO) in the creation of the International Union for the Conservation of Nature (IUCN) offers one of the earliest examples of UN promotion of international cooperation through NGO partnerships. Since its creation in 1972, the United Nations Environment Programme (UNEP) has been at the heart of the UN environmental activities, and the Global Environment Facility (GEF) was established in 1991 to provide coordinated funding for global environmental issues (GEF Website, 2011). However, environmental concerns are woven into activities throughout the organization. On the UNEP website, a diagram of the "environmental DNA" of the UN system lists 44 UN entities that play explicit roles in environmental issues (UNEP Website, 2011), and most of these include biodiversity and ecosystems among their priorities. The ubiquity of biodiversity interests was summarized in a report prepared for the International Year of Biodiversity (IYB) in 2010 (UN Environment Management Group, 2010). The recognition that the 2010 biodiversity conservation targets had not been met, and the importance of the challenges, led the UN to designate 2011–2020 as the International Decade on Biodiversity, in order to highlight the urgency of achieving the next set of decadal goals (CBD Website, 2011).

The most significant global assessment of biodiversity and ecosystem services to date, the Millennium Ecosystem Assessment (MA), was a joint initiative of four intergovernmental organizations and NGOs – UNEP, UNESCO, World Bank, and the World Resources Institute (WRI) – with support from a total of seven UN entities. It enlisted the skills of more than 1360 experts worldwide, and their findings provide a framework for understanding the importance of biodiversity and ecosystems for human well-being, as well as a record of the state and transformations in ecosystems around the world (Reid *et al.*, 2005). The MA sparked a renewed recognition of the pervasiveness and increasing urgency of biodiversity and ecosystem concerns, as a consequence of which the UN General Assembly approved in 2010 the creation of an Intergovernmental Platform for Biodiversity and Ecosystem Services (IPBES). The IPBES will create a new intergovernmental institution, comparable to the Intergovernmental Program on

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00395-6

Climate Change (IPCC), to provide a mechanism to contribute to global policy deliberations through regular assessments of biodiversity and ecosystem services. These assessments will be performed by the global academic, governmental, industry, and NGO communities with interests in biodiversity (IPBES Website, 2011). The new institution may, in turn, help introduce more coherence and consistency into the conservation and development strategies of governments, industries, and NGOs, if IPBES is able to lead to common, and improved, understandings of the baseline conditions, possible future scenarios, and complexities of ecosystem change.

This article attempts both to catalog institutions and their missions, and to capture the dynamism of their interactions and activities. It outlines the framework of international organizations concerned with biodiversity issues in a variety of contexts, with sections focusing on such topics as biodiversity conservation, biodiversity research, collecting and archiving biodiversity data, and the confluence of biodiversity issues with economic activities such as agriculture, fisheries, forestry, and trade.

International Organizations and Biodiversity Conservation

Wildlife and habitat conservation were the primary concerns of the first international conservation organizations, which served as the kernel of what became the global environmental movement of the late twentieth century. Early conservation organizations, such as the Wildlife Conservation Society (founded in 1895 as the New York Zoological Society) and Flora and Fauna International (founded in 1903 in the United Kingdom as the Society for the Preservation of the Wild Fauna of the Empire), developed early conservation norms. They were then joined by new international organizations as the scope of conservationists' interests broadened. The new and evolving organizations began to apply the groups' lessons and approaches worldwide.

Intergovernmental conservation efforts were among the earliest activities of the UN, principally through UNESCO. Shortly after its establishment in 1945, UNESCO partnered with governments and national organizations to launch the IUCN, initiating a tradition of partnerships between international governmental organizations and NGOs to accomplish conservation ends. The initial focus of IUCN in the 1950s was on endangered species and protected areas, and in 1963 the organization established the Red List of Threatened Species, aimed at drawing attention to the magnitude and importance of threats to biodiversity worldwide (IUCN Website, 2011; International Union for the Protection of Nature, 1948).

IUCN was joined in succeeding decades by an array of international organizations focused on conservation of threatened species and the ecosystems in which they lived. The World Wildlife Fund (WWF) was founded in 1961 to advocate for the protection of wildlife and wild places around the world. Their campaigns used "flagship species," such as the iconic panda of their logo, to raise awareness and stimulate action and funding for their global activities and advocacy (WWF Website, 2011). In the 1980s, the Nature Conservancy went international with its successful US model of purchasing and managing ecologically valuable properties, joining WWF and Conservation International (CI) in pioneering "debt-for-nature" swaps in Latin America and elsewhere (The Nature Conservancy Website, 2011). CI was founded in 1987, and their approach emphasizes partnerships with businesses and governments to provide resources and develop policies for the conservation of biodiversity and natural habitats in concert with economic development (Conservation International Website, 2011). CI has prioritized the preservation of areas they have designated as Biodiversity Hotspots, defined as places with very high numbers of endemic species in immediate need of conservation action in order to survive (CI Website, 2011a). Other international NGOs focus their efforts on narrower swaths of the biodiversity spectrum, such as BirdLife International, Botanic Gardens Conservation International, and the World Association of Zoos and Aquariums. Some concentrate on regional biodiversity conservation concerns, such as the Africa Wildlife Foundation, founded in 1961.

Biodiversity and conservation were among the central concerns in the explosion of environmental interest and activism in the 1970s, and intergovernmental conservation activities saw a concomitant growth. Since 1971, UNESCO has sponsored the Man and the Biosphere (MAB) Program with an emphasis on integrating conservation with human activity, overseeing the creation of a World Network of Biosphere Reserves (WNBR) as a testbed for research into innovative approaches to sustainable relationships between human societies and natural systems. Since the approval of the first WNBR site in 1978, the network has grown to encompass 580 sites in 114 countries. The mandate of the program has evolved from conservation and managed resource use to incorporate research into mitigation and adaptation to climate change in a broad range of ecosystems.

The focus on conservation of wildlife and habitat was a central theme of the UN Conference on Human Environment in Stockholm in 1972, and of the UNEP, which was founded shortly afterward. UNEP emerged as a key player in international biodiversity research and conservation. The 1970s also saw the negotiation and adoption of international legal instruments for biodiversity protection, including the following:

- The Convention on Wetlands of International Importance, or the Ramsar Convention, provided a framework for the conservation and wise use of wetlands, particularly ones of importance for waterfowl; the list of Ramsar sites in 2011 numbered 1952 (Ramsar Website, 2011).
- The Convention Concerning the Protection of the World Cultural and Natural Heritage, or the World Heritage Convention, oversees the preservation of sites that are valuable to the entire world citizenry, whether natural or cultural. Adopted and administered by UNESCO, the list in 2011 contained 936 properties (World Heritage Convention Website, 2011).
- The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) entered into force in 1975, to monitor and regulate the interstate transport of wild animals and plants to ensure sustainability.

- The Convention on Migratory Species of Wild Animals (CMS) was signed in 1979 to regulate the conservation and management of migratory species. Administered by UNEP, CMS develops and promotes regional and global agreements related to specific migratory species, among which are bats, cetaceans, and gorillas.

These piecemeal tactics for protection highlighted the need for a global approach to the protection of species, an issue that was galvanized after the 1987 National Academy of Sciences Conference on BioDiversity highlighted the urgency of a wise policy of conservation and development, and Gro Harlem Brundtland's 1987 report *Our Common Future* called for a Species Convention. UNEP took the lead in drafting what became the Convention on Biological Diversity (CBD). The convention was approved at the 1992 UN Conference on Environment and Development – the Earth Summit – as was the UN Framework Convention on Climate Change (UNFCCC). The three goals of the convention are the conservation of biodiversity, the sustainable use of biodiversity components, and the fair sharing of the benefits of genetic resources (Juma and Henne, 1997). The CBD obligates signatories to develop national plans for biodiversity conservation, and the CBD organization develops programs of work to achieve its conservation goals.

As the understanding of the complexity of environmental systems and their interactions with human society has grown, the agendas of all of these organizations have evolved. All have contributed to the MA and have incorporated its core findings regarding the linkages between ecosystems and human well-being. Each organization maintains a distinct strategy and focus, but their missions and work programs have developed to encompass the importance of ecosystems for the preservation of species, the importance of understanding and harnessing broad economic and political forces for conservation, and the critical role of raising global public awareness of the importance of biodiversity and ecosystem services.

International Organizations and Scientific Research Cooperation in Biodiversity

The concept of biodiversity arose from the scientific community, whose members have a long history of international cooperation in studying and promoting the conservation of natural systems. Scientific research is woven into the activities and strategies of every one of the international governmental and nongovernmental conservation organizations discussed earlier. The scientific community participates in and is supported by the conservation community, in a complex and evolving system of interactions.

The earliest international organizations dedicated to promoting international cooperation in biological research were the International Union of Biological Sciences (IUBS), founded in 1919, and the International Union of Microbiological Sciences (IUMS), founded in 1927, the memberships of which were national scientific societies. IUBS was among the founding members of the International Council for Science (ICSU) in 1931, formed as an umbrella organization to promote international cooperation in science, and IUMS joined in 1982. With its wide-ranging scientific overview, ICSU has promoted interdisciplinary, integrated scientific approaches to global environmental issues, including a number of programs that have made significant contributions to international cooperation in biodiversity research. These include the Scientific Committee on Problems of the Environment (SCOPE; founded in 1969), the Scientific Committee on Ocean Research (SCOR; founded in 1957), the Program on Ecosystem Change and Society (PECS; founded in 2008), and the four programs devoted to global change research: the World Climate Research Program (WCRP; created in 1980), the International Geosphere–Biosphere Program (IGBP; created in 1986), the International Human Dimensions Programme on Global Environmental Change (IHDP; created 1996), and the International Programme of Biodiversity Science (DIVERSITAS; established in 1991 by IUBS, UNESCO, and SCOPE, which were joined by ICSU and IUMS in 1996). In addition to the international coordinating function provided by the ICSU family, some scientific professional societies are international in scope. Most notably, the Society for Conservation Biology has a broad international membership of scientists, social scientists, economists, and professionals such as lawyers and teachers interested in science and policy of biodiversity and ecosystems.

These nongovernmental, scientific coordinating organizations have played a crucial role in defining international scientific priorities, promoting international research collaboration, and providing input to national and international policy deliberations. ICSU is often called upon to represent the international scientific community in UN fora and has served as a partner to UN agencies in achieving their scientific goals. For example, in preparation for the UN Commission on Sustainable Development's "Rio + 20" conference in June 2012, ICSU organized regional science and technology workshops, and participated in intergovernmental preparatory meetings. ICSU's programs are generally organized with input from intergovernmental organizations such as UNEP and UNESCO, and thus reflect global governmental priorities and needs for input.

Two of the ICSU initiatives have an explicit biodiversity focus: DIVERSITAS and PECS. DIVERSITAS was created to integrate the broad range of biological sciences needed to understand observed loss and change in global biodiversity, and its scope evolved in concert with the insights enunciated in the MA to incorporate more policy relevant aspects of biodiversity and ecosystem change. The organization provides an umbrella for international research cooperation on topics from evolution to biodiversity data to sustainability, and cooperates in the biodiversity and ecosystem aspects of the other ICSU programs in the Earth System Science Partnership (ESSP), particularly IGBP and IHDP. DIVERSITAS is also an active partner with intergovernmental organizations, providing scientific support to the CBD and co-chairing the Biodiversity Observation Network of the intergovernmental Group on Earth Observations (GEO BON). DIVERSITAS was a leader in advocating the scientific communities' views as the proposal for IPBES was developed, and represented ICSU at the UN-sponsored negotiations for the establishment of the organization in 2011–2012.

The more recent ICSU initiative related to biodiversity and ecosystems, PECS, aims to complement the ESSP programs by

developing an international program of work on social–ecological systems, with particular focus on assessing the effects of policies and practices on the resilience of critical ecosystem services. Along with DIVERSITAS, it will provide scientific input to the proposed IPBES (ICSU Website, 2011a).

Other international organizations conduct research on biodiversity and ecosystem services, many serving as hubs of regional and global research networks. Among the organizations with significant biodiversity-related activities are the following:

- The WRI, a policy research institute addressing global resource and environmental issues founded in Washington, DC, in 1982, was a key organizer of both the 1995 Global Biodiversity Assessment and the MA. WRI collaborates with UN agencies to produce the World Resources Report biannually.
- The International Institute for Applied Systems Analysis (IIASA), an international nongovernmental institute, founded in 1972 near Vienna, Austria, conducts systems analysis research on global problems, including evolution, land use, and ecosystem services.
- The Inter-American Institute for Global Change Research (IAI) and the Asia-Pacific Network for Global Change Research (APN) are intergovernmental organizations serving as hubs for networks of research institutions working on global change issues in the Americas and the Asia-Pacific region, respectively. Both organizations promote science that cannot be conducted by any one country alone through research grants to international teams, scientific meetings, and capacity building. Biodiversity and ecosystems are key elements of their agendas.

International Organizations and Biodiversity Data

The conduct of research, the development of conservation priorities, and the understanding of changes in ecosystems are all dependent on the analysis of reliable global data sets. The cataloging and preserving of accurate records of life on Earth presents particular challenges. The traditional repositories of biodiversity data are local institutions such as natural history museums and academic collections. They rely on international standards for the names of plants and animals that have long been overseen by the International Association for Plant Taxonomy and International Commission on Zoological Nomenclature. However, recent decades have witnessed an explosion of both remotely sensed data that offer a global context for ecosystem understanding and genetic data that offer a new understanding of the molecular basis for biological diversity. In addition, new information technologies provide previously unimaginable opportunities for global access to biodiversity data. These developments have led to an array of international organizations devoted to collecting and maintaining these data, and to making them available for use by the global science and policy communities.

The Global Biodiversity Information Facility (GBIF), established in 2001 as an intergovernmental initiative to encourage free and open access to biodiversity data via the Internet, had as of 2011 built a global network of 57 countries and 47 organizations. GBIF promotes and facilitates the mobilization, access, discovery, and use of information about the occurrence of organisms over time and across the planet (GBIF Website, 2011). The Consortium for the Barcode of Life (CBOL) is an international initiative established in 2004 devoted to developing DNA barcoding as a global standard for the identification of species. Its International Barcode of Life (iBOL) project is coordinating a 26-nation research alliance to construct a global DNA barcode reference library for that purpose. The Encyclopedia of Life (EOL) is a global partnership, begun in 2007, with a goal of compiling an online, dynamic catalog of every species on Earth. The EOL partnership included 181 content partners as of 2011, making it a link to an array of biodiversity databases on organisms from microbes to whales (EOL Website, 2011).

The UNEP World Conservation Monitoring Centre (UNEP-WCMC) is a collaboration between UNEP and the UK-based charity WCMC that supports international biodiversity conservation initiatives by compiling, storing, and analyzing biodiversity data on matters of relevance to international conventions (such as CBD and CITES), businesses (through its Proteus partnership), and conservation and research organizations. Databases supported by UNEP-WCMC include the World Database on Protected Areas; the UNEP-WCMC Species databases on organisms of conservation importance, including those protected by multilateral environmental agreements; and the CITES database of trade in protected species (UNEP-WCMC Website, 2011).

GEO BON is the biodiversity component of the Global Earth Observation System of Systems (GEOSS). It encompasses a partnership of some 100 governmental organizations and NGOs and is chaired by DIVESITAS, NASA, and the EU Biodiversity Observation Network (E-BONE). The GEO BON goal is to collect and make accessible a comprehensive collection of *in situ* and remotely sensed biodiversity observations, accompanied by tools for analysis that enable decision making in support of conservation and management of resources (GEO Website, 2011).

Among the best known independent databases is the IUCN Red List of Threatened Species. From its beginnings in 1963, the Red List has grown into "the world's most comprehensive inventory of the global conservation status of plant and animal species," utilizing resources from museums, conservation organizations, and academic institutions to provide data on approximately 45,000 species, including both threatened and nonthreatened organisms. Partners such as BirdLife International, which connects avian conservation organizations in more than 100 countries, coordinate specific elements of the database. With a new focus on scientific rigor and transparency in the 1990s, the Red List has evolved into a valuable tool for conservation strategies (IUCN Website, 2013).

International Organizations and Biodiversity for Agriculture and Forestry

Humanitarian, commercial, and conservation concerns converge in global cooperation related to agriculture and forestry. Traditional efforts to develop seeds and efficient production

methods to aid in the fight against hunger or increase the economic viability of agricultural enterprise have been influenced by the recognition that valuable genetic resources were being lost in the process of modernization. Forest management has developed beyond the concept of maximum sustained yield to incorporate an approach that can maximize the value of a broader range of ecosystem services provided by forest ecosystems. Expanded global economic integration has accelerated the introduction of new species into far-flung corners of the globe, some on purpose, others accidentally. Some of these invaders have been so successful that their proliferation threatens the existence of the ecosystems that they have colonized, endangering both economic and cultural services that the ecosystems provide. The global intertwining of ecosystems, ecosystem services, and ecosystem threats has generated a complex international institutional framework aimed at understanding and regulating these ecosystem-dependent economic activities to increase and balance the societal benefits and costs.

Agriculture and Biodiversity

The boom in international cooperation in agricultural research dates back to the 1950s and 1960s, when private foundations and the World Bank began funding the development of "miracle grains" in specialized research institutes around the world. These collaborations were integrated and institutionalized with the creation of the Consultative Group on International Agricultural Research (CGIAR) in 1971, funded by a consortium of donor countries, foundations, and multilateral organizations (CGIAR Website, 2011). During that same period, the erosion of plant genetic resources was attracting the attention of the UN Food and Agriculture Organization (FAO) and was widely discussed at the 1972 UN Conference on the Human Environment. Shortly thereafter, the IUCN added *Threatened Plant Species* to its Red List system, focusing on plants that fell outside the interests of the FAO (Maunder, 2001), and the preservation of plant and animal genetic resources, particularly in the context of agriculture, took a central place on the agendas of international conservation, research, and development organizations.

The FAO began in the 1970s to integrate biodiversity and sustainability into its structure and core activities. The FAO Commission on Genetic Resources for Food and Agriculture (CGRFA), established in 1983, has collaborated with CBD and CGIAR in assessments of plant, animal, and forest genetic resources, and on initiatives for the conservation and sustainable use of soil biodiversity. CGRFA also oversaw negotiations for the International Treaty on Plant Genetic Resources for Food and Agriculture, adopted in 2001 to facilitate access and share benefits, and FAO helps set standards and global frameworks for collection, storage, preservation, and sustainable use of genetic resources.

International efforts to expand seed bank collections as a hedge against species loss have been spearheaded by botanical gardens and the Global Crop Diversity Trust (GCDT), which was jointly established by FAO and CGIAR/Biodiversity International. The Trust has collaborated with the Government of Norway to establish the Svalbard Global Seed Vault as a backup conservator of the world's gene bank collections against strife and natural disaster. The Millennium Seed Bank, sponsored and hosted by Kew Royal Botanic Gardens, works with a network of international partners to conserve wild plant diversity.

As of 2011, the CGIAR comprised a consortium of 15 research centers, and their research agenda had expanded beyond crop genetic improvement to include sustainable pest management and agricultural technologies, as well as research on the efficacy of proposed agricultural policies. In particular, Biodiversity International (formerly International Plant Genetic Resources Institute), located in Rome, focuses on how agricultural biodiversity can be used to improve nutrition and make agriculture more sustainable both economically and environmentally, and the International Food Policy Research Institute (IFPRI) in Washington researches policies and practices that can improve the sustainability of food production. A number of the centers are conducting research in agricultural and agroforestry landscapes, incorporating biodiversity as an element of research programs (CGIAR Website, 2011a).

In an effort to assess the challenges facing the global agricultural enterprise, including its relationship to ecosystems, agricultural livelihoods, agricultural knowledge, and food policies, seven UN agencies launched the International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD). The 3-year effort culminated in a 2008 synthesis report that analyzed issues and global scenarios for agricultural goals related to reducing hunger, improving health and rural livelihoods, and fostering environmental and social sustainability. Five subglobal assessments captured the regional differences in agricultural scenarios (IAASTD Website, 2011).

Conservation organizations and the international conservation science community are also exploring the connections between agriculture and biodiversity conservation. The MA observed a dichotomy between agricultural research and conservation biology, in which, for example, research on related species and landscapes proceeded on parallel tracks and without significant communication. International organizations are involved in efforts to bridge that divide. The DIVERSITAS agroBIODIVERSITY program fosters international research cooperation "to address the trade-offs between food production, biodiversity conservation, ecosystem services, and human well-being in agricultural landscapes" (Jackson *et al.*, 2005). International conservation organizations have also added sustainable agriculture to their activities, advocating alternatives to the slash-and-burn approach that is so destructive of forest ecosystems. In these efforts, conservation organizations are finding institutional partners from the business community. For example, IUCN joined with the World Business Council for Sustainable Development to produce a 2008 report calling attention to challenges related to agriculture and food security, WWF works with partners from local producers to international markets to promote practices and policies for sustainability (World Wildlife Fund Website, 2011), and the Conservation International Center for Environmental Leadership in Business negotiates agreements with international producers such as Starbucks to promote sustainable agricultural practices (CI Website, 2013).

Forestry and Biodiversity

As discussed in the previous section, much of the work by FAO and CGIAR on agriculture and biodiversity has elements related to forestry and forest biodiversity. However, in the context of the UNFCCC, a new international organization has become a major player in forest conservation since 2009. The UN Program for Reducing Emissions from Deforestation and Forest Degradation (UN-REDD) is an intergovernmental agreement, jointly executed by FAO, UNDP, and UNEP, to create financial value for the carbon stored in forests, offering incentives for developing countries to reduce emissions from forested lands and invest in low-carbon paths to sustainable development, including biodiversity conservation. UN-REDD works in close coordination with the World Bank's Forest Carbon Partnership Facility (FCPF), which assists tropical and subtropical forest countries to develop REDD strategies.

Implementation of the REDD program, with its emphasis on sustainability and improved livelihoods, is important for both conservation and the forestry industry. International conservation organizations, such as WWF, IUCN, CI, and the Nature Conservancy, as well as the intergovernmental International Tropical Timber Organization (ITTO), have longstanding programs of research, field work, and policy development related to both forestry and climate, and they are active both in UN-REDD and in advising national REDD programs. Other multilateral development banks, such as the Asian Development Bank and the Inter-American Development Bank, are developing strategies to take advantage of emerging financial incentives for forest conservation under REDD.

Accelerating international trade is leading to increasing risk of introducing destructive invasive species, which can cause severe damage to forest ecosystems and ecosystem services. These invasive organisms range from disease-causing microbes to weeds to destructive insects to imported game animals, and can lead to significant economic harm. For that reason, the international trade negotiators incorporated the right to preserve plant, animal, and human health into the rules governing international trade. The Sanitary and Phytosanitary Agreement in the General Agreement on Tariffs and Trade (GATT) was incorporated into the structure of the World Trade Organization (WTO), giving the WTO a voice in a key area of biodiversity and ecosystem conservation.

International Organizations and Aquatic Biodiversity

Marine and freshwater biodiversity have a long history of international study and regulation, and a concomitantly broad array of organizations with interests in aquatic biodiversity. The reasons are manifold: much of the marine realm is located outside national boundaries; many marine and freshwater organisms are mobile, traveling readily between and among different national and international regimes; aquatic biodiversity provides an important source of food and resources, and the societies and industries that rely on those resources have long recognized the need to protect them from excessive exploitation and damage from pollution. More recently, new imperatives for understanding and protecting marine biodiversity have arisen because of increased understanding of the role of marine microorganisms in critical biogeochemical processes that regulate atmospheric composition and global climate, and the new and poorly understood biological effects of global ocean acidification caused by the ocean's absorption of anthropogenic greenhouse gases.

UN agencies are central to marine biodiversity conservation. The UN Convention on Law of the Sea (UNCLOS) includes a Convention on Fishing and Conservation of Living Resources of the High Seas, which is one among dozens of global and regional instruments of international law designed to govern the exploitation of marine life and protect it from damage through pollution or overexploitation. There are specific agreements on whales, seals, marine mammals, sea turtles, cetaceans, sea birds, Antarctic marine life, highly migratory species, and straddling or anadromous stocks, in addition to the elements of CITES, CBD, and other treaties that apply to marine environments (International Fisheries Agreements, 2008). Some are executed by distinct international organizations such as the International Whaling Commission, set up by the International Convention for the Regulation of Whaling in 1946, which governs the conduct of whaling, maintains whale databases, and funds and publishes whale and whaling research. Another example is the Commission for the Conservation of Antarctic Marine Living Resources, which enforces the Convention on Antarctic Marine Living Resources. Other elements of the UN responsibility for marine life reside in UNDP, UNEP, FAO, CBD, and UNCLOS. Because of the UN's broadly distributed responsibilities for marine resources, a number of coordinating mechanisms have been established. In 1969, the Joint Group of Experts on the Scientific Aspects of Marine Environment Protection (GESAMP) was created to provide scientific guidance on marine issues to any of the nine sponsoring UN agencies, utilizing working groups of international experts; GESAMP prepared a report on marine biodiversity in 1997 (GESAMP, 1997) and has regularly considered biodiversity-related issues. To coordinate UN oversight of marine biodiversity, a task force on Marine Biodiversity Beyond National Jurisdiction was set up by the UN-Oceans interagency organization in 2005. The task force is jointly led by the UNCLOS and CBD.

Marine governance depends on data, and international databases are supported by a network of intergovernmental and independent international organizations. UNESCO's Intergovernmental Oceanographic Commission (IOC) manages the Global Ocean Observing System, which supports a number of remotely sensed databases related to marine biodiversity. Integrated fisheries databases are maintained by the Fishery Resources Monitoring System (FIRMS) in order to provide high-quality information for monitoring and managing fishery marine resources (FIRMS Website, 2011). FIRMS integrates data collected from a network of intergovernmental organizations such as the intergovernmental International Council for the Exploration of the Sea (ICES), which conducts research, maintains marine databases, and provides scientific advice to governments and international regulatory bodies related to the North Atlantic, and the Southeast Asian Fisheries Development Center, which promotes sustainable fisheries development in Southeast Asia, as well as conducting research and training.

International conservation and research organizations are likewise active in marine biodiversity issues. IUCN is leading an international consortium of organizations in a Global Ocean Biodiversity Initiative (GOBI), building on the Census of Marine Life assessment to compile a database to CBD standards that can be used as a scientific basis for conserving marine biodiversity (GOBI Website, 2011). The Nature Conservancy sponsors projects around the world aimed at protecting marine resources while also developing sustainable human communities adjacent to the sea. WWF has a Global Marine Program focusing on ecosystem-based management, "smart fishing," and protection of priority areas of great biological diversity. WWF is also sponsoring roundtables with aquaculture stakeholders to develop standards for sustainable aquaculture.

Biodiversity and International Business Strategies

Biodiversity and ecosystem services are important to global economies. For that reason, international organizations engage in research to understand the role of ecosystems' "natural capital" in economics and work with businesses to ensure that economic development is compatible with the preservation of biodiversity.

The Economics of Ecosystems and Biodiversity (TEEB) is an intergovernmental initiative coordinated by UNEP to evaluate ecosystem services, including costs of biodiversity loss and strategies for governments and businesses to minimize those losses (TEEB Website, 2013).

The potential to use biodiversity as a source for the development of new commercial products, particularly pharmaceuticals and biofuels, has caused conflict between poor countries with rich biodiversity resources and the companies and nations that have the intellectual and financial resources to commercialize them. International organizations play a significant role in developing equitable solutions to those disputes. Moreover, conflicts of this sort combine with information campaigns by environmental advocates to raise public awareness of the importance of biodiversity and the negative effects that global development and trade can have. This awareness plays out in consumer preferences that create incentives for businesses to make their operations more sustainable.

International organizations have developed programs designed to assist enterprises with their sustainability strategies. These programs vary depending on the scale of the business, ranging from households looking to raise their incomes beyond subsistence levels through collection of organic, wild collected products, to global exporters of agricultural products that require a "green" certification for their biodiversity-friendly products and production methods. The International Trade Center (ITC), a joint program of the UN and the World Bank, works with small and medium-sized enterprises to comply with green certification schemes, both to ensure sustainable supply of "biodiversity products" or organic crops and to improve access to markets for small producers. They offer information, advisory services, and training for producers in developing countries (UN-ITC, International Trade Center Website 2011).

The Forest Stewardship Council (FSC) is one example of an international organization that uses a certification program to promote responsible management of the world's biodiversity resources. The FSC develops standards and conducts inspections to assure that forest products are being harvested with due respect to both forest-dependent communities and biodiversity (FSC Website, 2011). To certify the standards of the FSC, it has collaborated with environmental organizations such as WWF, which endorses the FSC on its website. It is also a member of the ISEAL Alliance, the global association for social and environmental standards (ISEAL Alliance Website, 2011). Other biodiversity-related accrediting organizations that are members of ISEAL include the International Federation of Organic Agriculture Movements, International Organic Accreditation Service, the Marine Aquarium Council, the Marine Stewardship Council for sustainable wild seafood, the Rainforest Alliance/Sustainable Agriculture Network, the Roundtable on Sustainable Biofuels, and Union for Ethical Bio Trade for ingredients from native biodiversity.

Civil Society and Biodiversity

A number of international organizations doing research related to international economic development are engaged in work related to biodiversity conservation as an element of sustainable development. Likewise, relief organizations are recognizing that damage to ecosystem services is a consequence of relief activities during humanitarian crises, and mitigating those damages can be a valuable addition to planning. The convergence in interests and activities between these organizations, such as the International Institute for Environment and Development in London, the Center for Global Development in Washington, the Red Cross, and the international science and conservation organizations discussed earlier (*see* International Organizations and Biodiversity Conservation), demonstrates the breadth of the biodiversity agenda and its importance to all elements of human well-being.

Destruction of ecosystems and threats to biodiversity are endemic to human activity, and the international conservation community has recognized that humanitarian crises such as natural disasters, military conflicts, and the refugee emergencies that result can be compounded by the environmental consequences of the international interventions meant to relieve human suffering. In consequence, international conservation organizations have formed Stewardship Council partnerships to help in aid and recovery planning. For example, the UN High Commissioner for Refugees (UNHCR) has worked with international partners to develop sound environmental practices in establishing and operating refugee camps, both to mitigate degradation and to avoid conflicts over resources (UNHCR Website, 2011). UNHCR worked with CARE International, and relief workers in the field, to develop a toolkit to help field workers and managers make rapid environmental assessments to help minimize adverse environmental impacts and conflicts with surrounding communities (UNHCR, 2009). Likewise, WWF has worked with the Red Cross to develop guidelines for environmentally sustainable reconstruction efforts, a partnership that developed in the aftermath of the Indian Ocean tsunami in 2004 (WWF 2011a).

Conclusion

This article offers a framework for understanding how international organizations contribute to the understanding and handling of the questions of biodiversity and human well-being. Although the author has made efforts to include the most important of the groups working in this sphere, the discussion cannot be considered exhaustive and, in addition, is in a constant state of flux, as organizations form, disperse, and change names or focus. Some important organizations and initiatives have undoubtedly been omitted.

See also: Agricultural Invasions. Agriculture, Industrialized. Agriculture, Sustainable. Agriculture, Traditional. Agrobiodiversity. Biodiversity in Logged and Managed Forests. Biodiversity Informatics. Biodiversity-Friendly Farming. Commons, Institutional Diversity of. Conservation Efforts, Contemporary. Conservation Movement, Historical. Conserving Biodiversity Outside Protected Areas. Convention on Biological Diversity. Ecology of Agriculture. Economic Control of Invasive Species. Economics of Agrobiodiversity. Ecosystem Services. Feeding the World and Protecting Biodiversity. Fish Conservation. Framework for Assessment and Monitoring of Biodiversity. Human Impact on Biodiversity, Overview. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Introduced Plants, Negative Effects of. Loss of Biodiversity, Overview. Mammals, Conservation Efforts for. Plant Conservation. Plant Invasions. Priority Setting for Biodiversity and Ecosystem Services. Rainforest Loss and Change. Slash-and-Burn Agriculture, Effects of. The Global Environment Facility: Financing the Stewardship of Global Biodiversity. The Value of Biodiversity. Timber Industry

References

CBD Website (2011) *United Nations Decade on Biodiversity 2011–2020*. http://www.cbd.int/doc/strategic-plan/un-decade-biodiversity.pdf.
CGIAR Website (2011) http://www.cgiar.org/
CGIAR Website (2011a) http://www.cgiar.org/centers/index.html
Conservation International Website (2011) *History*. http://www.conservation.org/about/pages/history.aspx.
Conservation International Website (2011a) *The Biodiversity Hotspots*. http://www.conservation.org/where/priority_areas/hotspots/Pages/hotspots_main.aspx.
Conservation International Website (2013) *Starbucks Partnership*. http://www.conservation.org/campaigns/starbucks/Pages/default.aspx.
Convention on Migratory Species (CMS) Website (2011) *Introduction to Convention on Migratory Species*. http://www.cms.int/about/intro.htm
Encyclopedia of Life Website (2011) http://eol.org
FIRMS Website (2011) *Fishery Resource Monitoring Systems: About FIRMS*. http://firms.fao.org/firms/about/en#Org-Mission.
FSC Website (2011) *Forest Stewardship Council: About FSC*. http://www.fsc.org/about-fsc.html.
GBIF Website (2011) *Global Biodiversity Information Facility*. http://www.gbif.org/
GEF Website (2011) *What is the Global Environment Facility*. http://www.thegef.org/gef/whatisgef
GESAMP (IMO/FAO/UNESCO-IOC/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection) (1997) Marine biodiversity: patterns, threats and conservation needs. *Report Studies* GESAMP (62):24p.
GEO Website (2011) *Group on Earth Observation Biodiversity: Strategic Target*. www.earthobservations.org/geobon.shtml.
Global Ocean Biodiversity Initiative (GOBI) Website (2011) *About GOBI*. http://www.gobi.org/About
ICSU Website (2011a) *Ecosystem Change and Society (PECS)*. http://www.icsu.org/what-we-do/interdisciplinary-bodies/pecs.
International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD) Website (2011) *Frequently Asked Questions*. http://www.agassessment.org/index.cfmPage=FAQs&ItemID=8.
International Fisheries Agreements (2008) *Volume 1 – Multilateral Fisheries Agreements, Table of Contents*. http://www.oceanlaw.net/projects/current/pdf/ifa_toc(draft).pdf.
International Trade Centre (2011) *Trade Support Institutions: Trade, Climate Change and Environment Program*. http://www.intracen.org/projects/tccep/
International Union for the Protection of Nature (1948) *Summary report of the Conference at Fontainebleau: Brussels*. http://data.iucn.org/dbtw-wpd/edocs/1948-001.pdf).
IOC Website (2011) *Intergovernmental Oceanographic Commission: Safeguarding the Health of Ocean Ecosystems*. http://www.unesco.org/new/en/natural-sciences/ioc-oceans/high-level-objectives/ecosystem-health/.
IPBES Website (2011) *About IPBES*. http://ipbes.net/about-ipbes.html
ISEAL Alliance Website (2011) http://www.isealalliance.org/
IUCN Website (2013) IUCN Redlist of Threatened Species: About. http://www.iucnredlist.org/about.
Jackson L, Bawa K, Pascual U, and Perrings C. (2005) agroBIODIVERSITY: A new science agenda for biodiversity in support of sustainable agroecosystems. DIVERSITAS Report N°4. 40 pp.
Juma C and Henne G (1997) Science and technology in the convention on biological diversity. In: Raven PH and Williams T (eds.) *Nature and Human Society: The Quest for a Sustainable World*. pp. 387–397. Washington, DC: National Academy Press.
Maunder M (2001) Plant conservation, overview. *Encyclopedia of Biodiversity* 4: 645–657.
Ramsar Website (2011) *The Ramsar convention and its mission*. http://www.ramsar.org/cda/en/ramsar-about/main/ramsar/1-36_4000_0__.
Reid WV, Mooney HA, and Cropper A (2005) *Ecosystems and Human Well-Being: Synthesis Report of the Millennium Ecosystem Assessment*.
Simon LA and Clark WC (eds.) (May 2010) Toward a Science of Sustainability: Report from Toward a Science of Sustainability Conference, Airlie Center, Warrenton, Virginia, November 29, 2009–December 2, 2009. CID Working Paper No. 196. Center for International Development at Harvard University.
TEEB Website (2013) About TEEB. http://www.teebweb.org/about#
The Nature Conservancy Website (2011) *Our History*. http://www.nature.org/about-us/vision-mission/history/
The Nature Conservancy Website (2011a) *Oceans and Coasts: places we protect*. http://www.nature.org/ourinitiatives/habitats/oceanscoasts/.
UN Environment Management Group (2010) *Advancing the biodiversity agenda: A UN system-wide contribution*.
UN Food and Agriculture Organization (UNFAO) Website (2011) *Biodiversity Instruments*. http://www.fao.org/biodiversity/instruments/en/.
UN-ITC, International Trade Center Website (2011) *Trade Support Insitutions: Biodiversity*, http://www.intracen.org/trade-support/biodiversity/
UNEP-WCMC Website (2011) *About UNEP-WCMC*. http://www.unep-wcmc.org/about-us_17.html.
UNEP Website (2011) *The UN and the Environment*. http://www.unep.org/un-env/.
UNESCO Website (2011) *Ecological Sciences for Sustainable Development: Man and the Biosphere Program*. http://www.unesco.org/new/en/natural-sciences/environment/ecological-sciences/man-and-biosphere-programme/.
UNHCR (2009) Framework for assessing, monitoring and evaluating the environment in refugee-related operations. *Modules* 1–7.
UNHCR Website (2011) *Environment: Looking After the Land*. http://www.unhcr.org/pages/49c3646c10a.html.
Wilson EO (1988) *Biodiversity*. Washington, DC: National Academies Press.
World Business Council for Sustainable Development (WCBSD) and IUCN (2008) *Agricultural Ecosystem: Facts and Trends*.
World Heritage Convention Website (2011) *The World Heritage Convention*. http://whc.unesco.org/en/convention
World Wildlife Fund Website (2011) *Agriculture: Engaging Business*. www.worldwildlife.org/what/globalmarkets/agriculture/engagingbusiness.html.
World Wildlife Fund Website (2011a) *Humanitarian Partnerships*. http://www.worldwildlife.org/what/partners/humanitarian/index.html.
WWF Website (2011b) *Species: Overview, Protecting the Future of Nature*. http://www.worldwildlife.org/species/.

Intertidal Ecosystems

AJ Underwood and MG Chapman, University of Sydney, Sydney, NSW, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Assemblage The collection of animals and plants found together in a patch of habitat.

Diversity index A numerical measure combining information about the number of species present in a sample or habitat and information about their relative abundances.

Ecological functions Attributes or properties of assemblages or habitats that are dependent on the biodiversity of animals and plants present. Examples are production of nitrogen, sequestering of heavy metals, sustained production of harvested resources, maintenance of a diverse food-web, and so on.

Ecological processes Direct or indirect interactions between species – such as grazing, predation, competition, responses to disturbance, and so on – that cause spatial or temporal patterns in the distributions and abundances of species.

Patchiness Variation from place to place (or time to time) in the abundances of animals or plants, caused by the interaction of numerous processes.

Recruitment Arrival of new juvenile animals or plants into a habitat or an older stage of the population. For marine animals, recruitment is at some time after the settlement of many planktonic larvae in the adults' habitat.

Intertidal areas such as rocky shores, mangrove forests, sandy beaches, and mudflats are habitats for very diverse assemblages of plants and animals. These species are usually distributed patchily in space and time because of several important ecological interactions. Local biodiversity is therefore patchy and variable. Measurement of biodiversity is complicated by variability. The functions provided by biodiversity are not well understood in most intertidal habitats. Planning for conservation, management, and restoration of biodiversity is difficult because of variability in space and time.

Introduction

Intertidal habitats range from rocky beaches or platforms to sandy and muddy shores. They occur under a wide range of physical stresses, most notably due to emersion when the tide falls and disturbances by waves. The animals and plants must endure the full range of physical variation, from being covered by water when the tide is in, to being potentially exposed to typical terrestrial conditions when the tide is out. In addition, the amplitude or range of tidal rise and fall can vary between two tides in a day (this is called "semidiurnal inequality") and, everywhere, varies in a fortnightly cycle between spring and neap tides. Spring tides rise much higher and fall much lower than do neap tides. As a consequence, during neap tides, animals and plants at high levels on a shore are not reached by an incoming tide for several days in a row and must endure long periods of inactivity or reduced activity until spring tides start again. At these high levels, moisture from spray (particularly where wave action is great) and during rain is very important for the supply of food and for preventing marine animals from drying out during their long periods of emersion.

In complete contrast, animals and plants at lower levels on the shore are submersed continuously for several days during neap tides because the tides do not fall low enough to expose such areas to aerial conditions. As a result, the animals and plants are, effectively, in a fully marine habitat and can be subject to predation by crabs, fish, starfish, and so on, which are mostly subtidal species. Such predators can forage at will over the lower parts of a shore during neap tides. This variation in periods of emersion versus submersion causes a very stark gradient of physical conditions from the top to the bottom of the shore.

A second strong physical gradient often exists along a shore, due to variation in wave action. As any shore bends away from being fully exposed to the open ocean, the strength of waves hitting the rocks will vary. As a result, there is a general decrease, with decreasing force of waves, in the amount of splash and spray over any area. The two physical forces – the direct force of waves and the ameliorating effects of splash and spray reducing the influences of desiccation during low tide – influence the abundances of animals and plants and the nature of their activities. Thus, grazing and predation are less intense where waves are strong. The diversity of species changes with wave action because of the combined influences of waves themselves, their ameliorating influence on desiccation, and the indirect influences of both physical features on the interactions among species.

These two major processes – tidal rise and fall and wave action – are so well known that their influences were long considered the major, if not the only, influences on local ecological processes and therefore diversity in intertidal habitats. The general features of intertidal regions, particularly the general patterns of distribution and numbers of species on rocky shores, have for a long time been described in terms of these physical forces. Summaries and details are available in Stephenson and Stephenson (1972) and other reviews of particular biogeographical regions or comparisons of various parts of the world.

More modern syntheses, in contrast, have focused on the ecological processes that have more profound effects in terms of creating and maintaining considerable patchiness in the

ways in which animals and plants are distributed in intertidal habitats. The variation often is considerably more conspicuous than the features considered general and easily explained in terms of gradients of physical forces. This forms the central theme of this contribution – the processes that influence the patterns of distribution and abundance of species and therefore the local diversity of species in intertidal habitats. The major focus is on rocky intertidal habitats because most of the experimental work in the last fifty years has been in that habitat. The animals and plants are typically relatively short-lived (although individuals can live for extraordinarily long periods). They are usually small, abundant, and most of them are slow-moving or sessile. As a result, experimental manipulations to test specific hypotheses from competing models that might explain patterns can be carried out without enormous areas of habitat being disturbed and without excessive costs of logistics and infrastructure. There has been a mass of such work and it has revealed and continues to reveal a great deal about the life histories, interactions, and the general and specific ecology of the maintenance of diversity in intertidal habitats (see the review by Underwood, 2000).

In addition to rocky shores, there is also some discussion of other intertidal habitats (sandy beaches, mangrove forests), even though these are discussed in detail elsewhere in this Encyclopedia. Their inclusion here is to demonstrate that the sorts of variability in time and space found on hard, rocky surfaces are more widespread and general and are not confined to rocky shores.

Following an overview of ecological patchiness and some consideration of its causes in intertidal habitats, three consequences of variability in biodiversity are introduced: how to measure and compare diversity from one place to another, how descriptions of biodiversity may relate to functional aspects of diversity, and the sorts of functions that biodiversity may provide in coastal habitats.

Processes Causing Patchiness on Rocky Shores

Numerous ecological processes influence or maintain patterns of local biodiversity in coastal habitats. These include direct and indirect biological processes of recruitment, predation, grazing, and competition. They also include physical disturbances and biological responses to them. These different types of processes are illustrated here.

Recruitment

Variability in the recruitment of coastal marine organisms is well described (Connell, 1985; Thorson, 1950). Many of the animals and plants have a relatively long dispersive phase in their life history, resulting in a large proportion of propagules being killed by predators, failing to develop, or not being able to find a suitable habitat in which to settle and metamorphose. Inevitably, the processes of production and loss of propagules lead to a huge variation in numbers surviving to settle in the adults' habitat and the number actually settling in any area of habitat at any particular time. This has led to the concept of "supply side" ecology (Roughgarden *et al.*, 1987; Underwood and Fairweather, 1989; Underwood and Keough, 2001). This is the idea that several ecological processes and their resulting patterns of diversity in assemblages are, at least partially, dependent on the numbers and types of larvae that arrive to grow in any area. For example, predation cannot be an important process in areas where predators fail to colonize as larvae, that is, fail to maintain populations of adults.

There are two major consequences to issues of biodiversity of this variation in recruitment to adult populations. First, vagaries of types and numbers of species arriving on any shore or in any area on a shore will ensure that the diversity of species continues to be patchy at small spatial scales. Second, across the entire ranges over which species disperse, there can be little large-scale variation in diversity. Thus, along a stretch of coastline over which organisms disperse, the assemblages that develop will, in general, receive recruits from similar mixtures of species. When the organisms can keep recruiting via dispersive larvae, populations will persist. Whatever happens in any area, recruitment from other areas will continue. Thus, assemblages will, on average, persist with the same mix of species. In contrast, at the more local scale of patches of coastline, or individual shores, there will be greater variation in the composition of assemblages.

At larger scales, speciation and changes in diversity from one side of an ocean to another are very much influenced by the recruitment of larvae. Thus, the vast stretches of open, oceanic water in the middle of the Pacific and the Atlantic oceans form barriers to the spread or the continuity of species. Scheltema (1971) sampled planktonic stages of the life history of various gastropod mollusks in samples from the middle of the Atlantic. He found a very good correlation between the number of larvae of a given species and the distribution of adults. Species found on coastlines on both sides of the Atlantic were regularly found as larvae in his samples. Larvae that were found less commonly were from adults that had closely related species or subspecies on the two coasts. This implied that the lack of frequent interchange of larvae was causing differentiation and speciation. Species found on only one side of the Atlantic were rare or absent as larvae in the sample. Thus, where larvae cannot arrive, diversity will differ.

The timing of recruitment by different species can also affect the local diversity of species. Thorson (1950) described the predatory brittle stars (*Amphiura* spp.), which do not feed for two months while breeding. Several species of mollusks with short larval development breed during this period. The brittle stars have a longer period of development. As a consequence, the young mollusks could be released into the plankton, complete their development, recruit, and grow to a size where at least some of them would escape from the voracious predatory recruits of the brittle stars. If the mollusks recruited later, they were unable to avoid their predators. This example needs to be examined experimentally, but indicates that the timing and order of recruitment can have profound influences on which species survive in an area.

Initial patterns of recruitment may be less important in determining local patterns of diversity for mobile than for sessile species. Many species of mobile animals can move over the substratum or move up into the water column to drift across relatively large areas in response to local environmental conditions. This not only affects small-scale patterns of

diversity but also contributes to the temporal variation in diversity in any one place.

A final point about the uncoupling of numbers of adults and recruits in any area of coast is that the mixture of different species and the numbers and ages of each species present in some area will not necessarily predict what will occur at any future time. Recruitment of any or all of the species in any diverse assemblage will not necessarily be related to the types or the sizes of populations present there.

Competition and Facilitation

A second major process that can influence local diversity is competition – negative interactions among individuals of the same or different species caused by absolute or relative shortage of resources needed by all of the individuals. In most coastal habitats, resources are food, space, and, for plants, light. Space is needed to settle and become established (for sessile species) and as an indirect resource for food (e.g., by mobile grazers that need relatively open space over which to feed). Alternatively, sufficient space on the substratum is needed to allow access to the water column for filter feeders.

Generally, competitive interactions are considered to lead to a reduction in abundances or actual elimination of competitively inferior species. Numerous examples demonstrate this, starting with Connell's (1961) classic experimental demonstration of destruction of one species of barnacle (*Chthamalus stellatus*) by another, faster-growing species (*Balanus balanoides*, as described in the original papers). As a result of overgrowth by *B. balanoides*, *C. stellatus* could be eliminated from some lower areas of their vertical distribution on a shore.

Such competitive overgrowth can have dramatic effects on the local richness of species. For example, Paine (1974) has described the major decrease in the number of species living on primary space – the underlying substratum – where mussels are able to overgrow and eventually eliminate most other species (limpets, chitons, algae, barnacles).

These examples do, however, paint a biased picture. For example, where mussels overgrow a primary surface, their shells provide new hard substrata on which the inferior competitors can become established. They also serve as "environmental engineers" (Lawton, 1994), because the packed shells of mussels trap sediments. While feeding on particles in the water, mussels selectively ingest some sizes and types of particles. The rest are wrapped in mucus and eliminated as pseudo-feces, which add substantial substance and nutrients to the surrounding sediments in the matrix of mussels. The result of all these processes is the provision of habitat and food for a very substantial number of species, such that the number of species in mussel-beds is very large (Lohse, 1993). The mix of species among the mussels will, however, be different from that outside a bed of mussels.

On rocky shores, many sessile species can function as ecological engineers, thus providing habitat for a range of species that may not otherwise live on the shore. These engineers are themselves patchily distributed, due to processes such as competition or predation, which, in turn, contribute to patchy distributions of those species that use them as habitat.

In soft sediments, ecological engineering can be very important. For example, large species can alter the habitat considerably by bioturbation, irrigation, changes in nutrient fluxes, and so on. The distributions and abundances of many smaller species are therefore influenced by the activities of a few larger species, which have major impacts on the sediment (i.e., the habitat) itself.

There are also complications due to indirect interactions (reviewed by Wootton, 1993), where the outcome of competition between two (or more) species (say, A and B) influences the outcome of competitive interactions between one of them and yet other species (say, B and C). Indirect interactions have been categorized as being of two distinct types. The first are called "interaction chains," where one species (A) directly influences abundances of a second species (B), which in turn influences numbers of a third species (C). Thus, A has an indirect effect on C, because of A's influence on B. The second type has been called an "interaction modification," where species A directly influences a second species (B). A third species (C) influences the interaction between A and B. The difference between the two types is that, in the latter, species C has no direct influence on either of the other participants themselves (i.e., no direct effect on A or B).

As an example of an interaction chain, Kastendiek (1982) described the interesting case of the turfing red alga, *Halidrys dioica* (consider it to be species B), which outcompetes the alga, *Pterocladia capillacea* (species C), by growing over and denying it access to light (**Figure 1**).

Where there is a canopy of a third species (A), *Eisenia arborea*, however, the inferior competitor (C) survives well under the canopy and is able to "resist" competition from *H. dioica* (B). Thus, competition for light between *E. arborea* and *H. dioica* (i.e., A and B) diminishes the latter's capacity to grow well and to take over space that is occupied by inferior competitors (C), such as *P. capillacea*. As a result, competition by *E. arborea* over *H. dioica* prevents competitive overgrowth of *P. capillacea* by *H. dioica*.

Predation

A major process influencing the diversity of species on rocky shores is predation. Predation has three distinct influences. First is simply that where predators are sufficiently numerous or active, they may eliminate some species from areas of habitat. Thus, toward the bottom of rocky shores on the north-west coast of the US (Connell, 1970), several species of the whelk, *Nucella* (or *Thais*), feed on barnacles. Low on the shore, the whelks had sufficient opportunity (due to prolonged submersion by the tides) to eat all of the barnacles, thus eliminating them from such low areas and thereby reducing the diversity of sessile species there. At higher levels, in contrast, the whelks were less active, because they had a smaller period submersed during low tide in which they could find an item of prey and consume it.

The second influence of predation is indirect. The whelk *Morula marginalba* in southeast Australia consumes a variety of prey, including barnacles and small limpets, *Patelloida latistrigata*. It takes the whelks considerably longer to drill a hole through the shell of a barnacle than to eat a limpet; *M. marginalba* can consume *P. latistrigata* without having to drill a hole through their shells. Fairweather (1985) removed

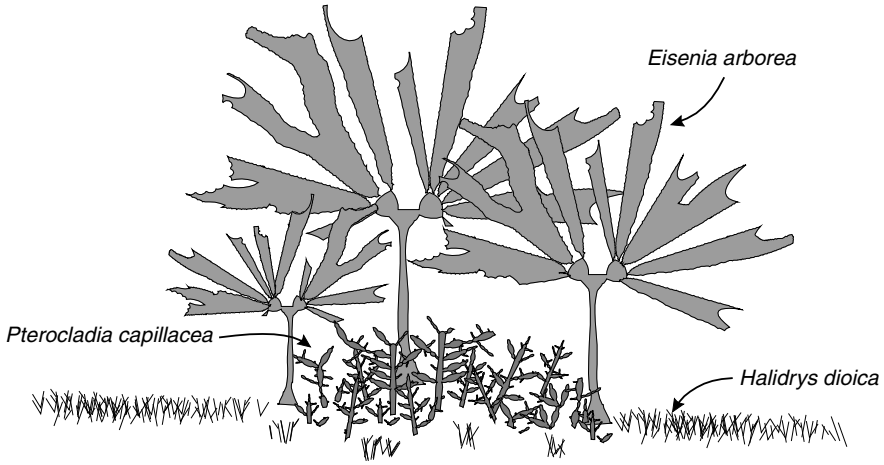


Figure 1 Diagram illustrating indirect interactions involved in competition for space and light among intertidal seaweeds (Kastendiek, 1982). The kelp, *Eisenia arborea*, forms a canopy that provides shade and moisture during low tide. Under the canopy, the red alga, *Pterocladia capillacea*, can thrive. Elsewhere, it is overgrown and outcompeted by the turfing alga, *Halidrys dioica*.

limpets from some areas as they arrived. As a result, predation of small barnacles by whelks increased. This demonstrated an indirect effect of preferences and differences in the time taken to consume prey; the least preferred species (the limpets) took less time to eat than the more highly preferred barnacles. Where there were no limpets, the whelks turned their entire attention to small barnacles, with consequent increased mortality and decreased abundances of the barnacles. Note also that whether or not limpets arrived from the plankton also influenced the eventual outcome in terms of abundances of barnacles and, therefore, the local relative abundances and, thus, diversity of species.

The final major influence of predation is also indirect. Large generalist predators are capable of consuming most other species in any patch of habitat. As a result, they can have direct influences on the abundances of prey and can free space on the shore for species to be able to recruit. Space is often in short supply, causing the various sessile species to compete with each other for space on which to live and causing grazing species to compete for space over which to feed. If the generalist predator disproportionately consumes "winners" of competitive struggles for space, predation will have very important effects on the local diversity of species. The original example of this phenomenon, known as "keystone predation," was investigated by Paine (1974). The large starfish, *Pisaster ochraceus*, eats most of the other sessile and mobile animals on rocky shores on the west coast of the US. In some areas, the starfish preferentially eat mussels (*Mytilus* spp.). Mussels are, however, among the best competitors for space, being able to smother and grow over other species (see earlier). Predatory starfish continuously remove mussels, making space available for inferior competitive species that would otherwise be locally eliminated by overgrowth due to the mussels. In this case, the predators exert an indirect positive influence on a range of species of prey, resulting in there being greater numbers of species (i.e., greater species diversity) where there are predators than where there is none.

Any analysis of the processes maintaining local diversity is made complex by the fact that different types of indirect interactions are not really mutually exclusive categories of ecological influences, as proposed by Wootton (1993; *see* Competition and Facilitation). For example, on shores in southeast Australia, the canopy-forming alga, *Hormosira banksii*, provides shelter from desiccation and heat stress for whelks, *M. marginalba*. The whelks come out from shelter to eat surrounding barnacles, *Chamaesipho tasmanica*. The algae are part of an interaction chain in that they provide habitat for whelks, which, in turn, decrease numbers of the barnacles. At the same time, the size of the algal canopy has been shown experimentally to influence the effectiveness of the whelks as predators (larger canopies are found with greater rates of predation on barnacles). Thus, the algae are also an interaction modifier of the process of predation by whelks on barnacles.

Physical Disturbances

Given that physical factors such as desiccation and wave action can have large and important influences on the abundances and diversity of species in intertidal habitats, it is not surprising that physical disturbances can influence diversity. Where disturbances are large or frequent, it is likely that more delicate species that are unable to withstand the physical forces will be unable to survive. They are either prevented from settling and becoming established in the first place or unable to complete their life cycles before a large disturbance kills them. It is expected under this model that there would be reduced diversity in more physically stressed habitats simply because of the direct loss of species due to disturbances. In contrast, where disturbances are small or rare, most, if not all, species can survive the rigors of the environment. Resources of space and food will then become critical and some species will disappear because of the superior competitive abilities of other species. Under these circumstances, diversity in physically benign areas will also be reduced below what is theoretically possible.

As a consequence of these two processes, it is anticipated that the greatest diversity should be found in areas of intermediate disturbance, such that competition for resources is

not too intense because there is sufficient disturbance to prevent all of the possible species from building up excessive abundances. At the same time, physical stress is not so great that some species are actually eliminated. This is the model of "intermediate disturbance" (Connell, 1978).

From this model, it has been predicted that in areas where physical disturbances can be increased or decreased in intensity or frequency, there will be predictable changes in diversity. For example, where disturbances are small, increasing their severity will cause increases in diversity as superior competitors are prevented from building up sufficient numbers to consume all of the resources needed by other species. If disturbances are increased even more, there will be a decrease in diversity.

Such predictions have been tested in several studies on rocky shores, with mixed success. Sousa (1980) identified patterns of diversity that were fairly consistent with the model, but the mechanisms operating were not as stated. Species were not lost by superior competition for space from other species. Rather, the life histories of the different species and the time taken to be able to recolonize disturbed areas were of greater importance. McGuinness (1987) identified the intermediate disturbance model to be an appropriate explanation for only some of the observed patterns of diversity that have been examined in intertidal boulder fields. Clearly, more work is needed to understand how and when the interactions of competition and disturbance lead to predictable outcomes in terms of the diversity of species.

Boulder Fields

Intertidal boulder fields form a very interesting habitat that forms part of rocky intertidal environments. Boulders have top surfaces that are exposed to waves and sunlight and that may be emersed during low tide. Underneath boulders, however, there is a very different sort of habitat – usually cool, moist, and dark. On shores affected by large waves, boulders are frequently overturned and moved. This alternately exposes each surface of the boulder to the light and causes the boulders to abrade against each other, damaging and killing many animals and plants living on them. These boulders support little biodiversity, mainly encrusting species that are resistant to harsh conditions or ephemeral species that rapidly colonize new habitat when it becomes available.

Boulders on sheltered shores, in contrast, are stable and support great diversity of animal and plant life. Different species are found on the tops of or underneath these boulders. Those on top are similar to those living on nearby rocky shores – after all, they are subjected to similar environmental conditions. Therefore, boulders may be covered with leafy and encrusting seaweeds, have patches of apparently bare rock or support large numbers of snails, limpets, and other common intertidal animals.

The species living under boulders are completely different. There are few leafy seaweeds, but there may be many kinds of encrusting seaweeds, which seem to thrive under dark conditions. There can also be patches of sessile animals – plate-like bryozoans, jelly-like masses of colonial ascidians or sea-squirts, encrusting sponges and the calcareous, and sandy tubes of different types of worms. These sorts of sessile animals are not common on rocky shores, except in rockpools or crevices, or under ledges (**Figure 2**).

Many of the mobile animals living under boulders are rather specific to boulders – they tend to be rare or absent from other intertidal habitats. Many show very specific patterns of behavior that allow them to move quickly from one boulder to another if the boulders are overturned or moved into unsuitable conditions by large waves. Therefore, although chitons on rocky shores tend to be very slow-moving and inactive animals, those under boulders move extremely rapidly over the rock-surface or curl up and drop from the boulder as soon as it is overturned. This behavior returns them to the undersurface of the same or nearby boulders.

Intertidal boulder fields are potentially important habitats for the conservation of biodiversity. They often form rather small patches of habitat – especially when compared with sandy beaches, mangrove forests, or rocky shores, which are often continuous for many kilometers. They are also often scattered along the coast, with large stretches of other habitats between them. There are no data about how much interchange there is among populations in different boulder fields or how consistently species would recruit from other boulder fields if locally eliminated. Because many species living in boulder fields are relatively confined to these habitats, they are likely to be vulnerable to human activities that destroy or damage their surroundings, but recent work demonstrates that many species can colonize newly created boulder fields.

Finally, many species are extremely overdispersed. This means that each species tends to be crowded onto relatively few of the available boulders, with most boulders not being occupied by that species. Because different species are often found on different boulders, individual boulders support different suites of species. Patterns of biodiversity are therefore very complex – both within and among different boulder fields.

As yet, the processes that lead to these complex patterns are not well understood, but include such factors as the size and shape of the boulder, what stone it is composed of, where it is (e.g., depth of water, what substratum is beneath it), and the different wave conditions where the boulder field is located. In addition, there are complex interactions among the animals



Figure 2 The undersurfaces of intertidal boulders support numerous species of sessile and mobile animals, many of which are not found in other intertidal habitats. These include encrusting sponges, ascidians and bryozoans, numerous chitons and snails and starfish, sea urchins, and brittle stars.

and plants themselves (Sousa, 1980), causing complex changes to biodiversity depending on which species happen, by chance, to recruit to which boulders.

Mangrove Forests

In contrast to intertidal rocky shores and boulder fields, mangrove forests are dominated by trees. Unlike some other habitats containing many large plants – for example, subtidal kelp beds (Foster and Schiel, 1985), and terrestrial rain forests (Terborgh, 1992) – patterns of diversity in mangrove forests (i.e., variation in the numbers and types of animals and plants found from place to place or time to time) are not well documented (Hutchings and Saenger, 1987). This is particularly true for any measures of small-scale patchiness in diversity within individual forests. This contrasts markedly with other intertidal habitats in which patchiness of animals and seaweeds is relatively well described (see the discussion on rocky shores presented earlier).

This lack of good quantitative measurements of biodiversity in mangrove forests is due to many different factors. First, the plants that make up the forests are, themselves, not very diverse. Typically, they do not develop a complex structure of many species of canopy, understory, and ground cover, which could, in turn, support many different types of animals. Second, because mangrove forests normally grow on very sheltered shores, the vertical gradients of environmental conditions associated with tidal height are very strong (Hutchings and Saenger, 1987). There are also strong latitudinal gradients, causing very large differences between temperate and tropical zones in the types and diversity of trees that make up the forests. These are well documented elsewhere and are not discussed further here. Within any single mangrove forest, however, different species of trees typically live at different levels along the intertidal gradient, so that in any one place, even fewer types of plants are found.

The main factor that has led to a poor description of the distribution and patchiness of biodiversity in mangrove forests is, however, that most diversity is due to small invertebrate animals. This is, of course, true for many habitats where much of the biodiversity is "invisible." In such habitats, appropriate sampling can clearly measure variation in diversity from place to place or time to time, but these sampling designs are inevitably complex and costly because they need to measure diversity at many different spatial and temporal scales.

The animals that make up most of the biodiversity in mangrove forests include many species of crustaceans, such as crabs and amphipods, small snails and bivalve mollusks, worms from many different phyla, insect larvae, and so on. Because these animals are generally very small and live on or under the surface of the mud, they are not readily visible without sampling small patches of mud, sieving the animals out of the sediment, and viewing them under magnification. The arduous work that this entails means that patterns in diversity are not well described for most mangrove forests.

In addition, because the types of plants found in mangrove forests often vary along the intertidal gradient, most studies of the distribution and diversity of animals have also described similar broad-scale patterns. Therefore, changes in the diversity of snails, crabs, and other animals from the seaward to the landward edges of the forest tend to dominate the literature. There have been few descriptions of the small-scale variability or patchiness in this diversity within a shore level, although this may be the most important source of variation in diversity, as is the case on rocky shores. Some studies that have been carried out to measure small-scale patchiness in diversity in mangrove forests show that much of the variability in the diversity and abundances of these animals is at very small spatial scales (i.e., among patches of habitat centimeters or meters apart). It can also change unpredictably through time (**Figure 3**).

This sort of patchiness of diversity is not limited to small invertebrates that live in the mud. Many of the larger animals that spend their juvenile stages in mangrove forests, such as fish, prawns, and large crabs, show similar patterns. A survey of small fishes in patches of mangrove forests in Sydney Harbor showed that most of the variation in abundance and diversity was found at the scale of meters (i.e., from one net to another set in a small patch of mangrove forest). Variation from site to site (hundreds of meters apart in the same bay) and from bay to bay (kilometers part) together accounted for less variability.

In contrast to rocky shores, there is little information about the processes that cause this small-scale patchiness in mangrove forests. This is due to the fact that, with few exceptions, mangrove forests have not been as well studied, particularly using well-designed, controlled field experiments.

The diversity of small animals living in sediments in any aquatic habitat is influenced by the range of grain sizes of the sediment itself. Larger species tend to be more common in coarse-grained sediments; small species are common among finer sediments. The range of sediment in any particular mangrove forest can vary from patch to patch because of such processes as differences in water currents around submerged objects (e.g., the bases of trees) or numerous biological processes, each of which also strongly and independently influence the patterns of diversity of the fauna.

In New South Wales (Australia), when feeding and burrowing, the mangrove crab, *Heloecius cordiformis*, creates mounds of coarser and drier sediments and flat areas of wetter and finer sediments (**Figure 4**; Warren and Underwood, 1986). Although not yet measured, it is likely that this activity directly influences local patterns of diversity, as has been shown for the burrowing behavior of soldier crabs in Tasmania (Warwick *et al.*, 1990). Other small-scale structures – for example, buried shells, the roots of the mangrove plants, worm tubes – are equally likely to alter patterns of diversity, although the importance of these small structures and the behavior of other animals as influences on biodiversity are not yet well documented for mangrove forests.

Sandy Beaches and Mudflats

Sandy beaches and mudflats tend to grade into each other along a continuum depending on the relative grain size of the sediment. Sandy beaches have rather coarse-grained sediments and mudflats are, of course, muddy, with fine-grained sediments. Both can occur on very sheltered shores, such as in estuaries, but mudflats do not occur on wave-exposed shores, where sandy beaches may be common (Brown and McLachlan, 1990).

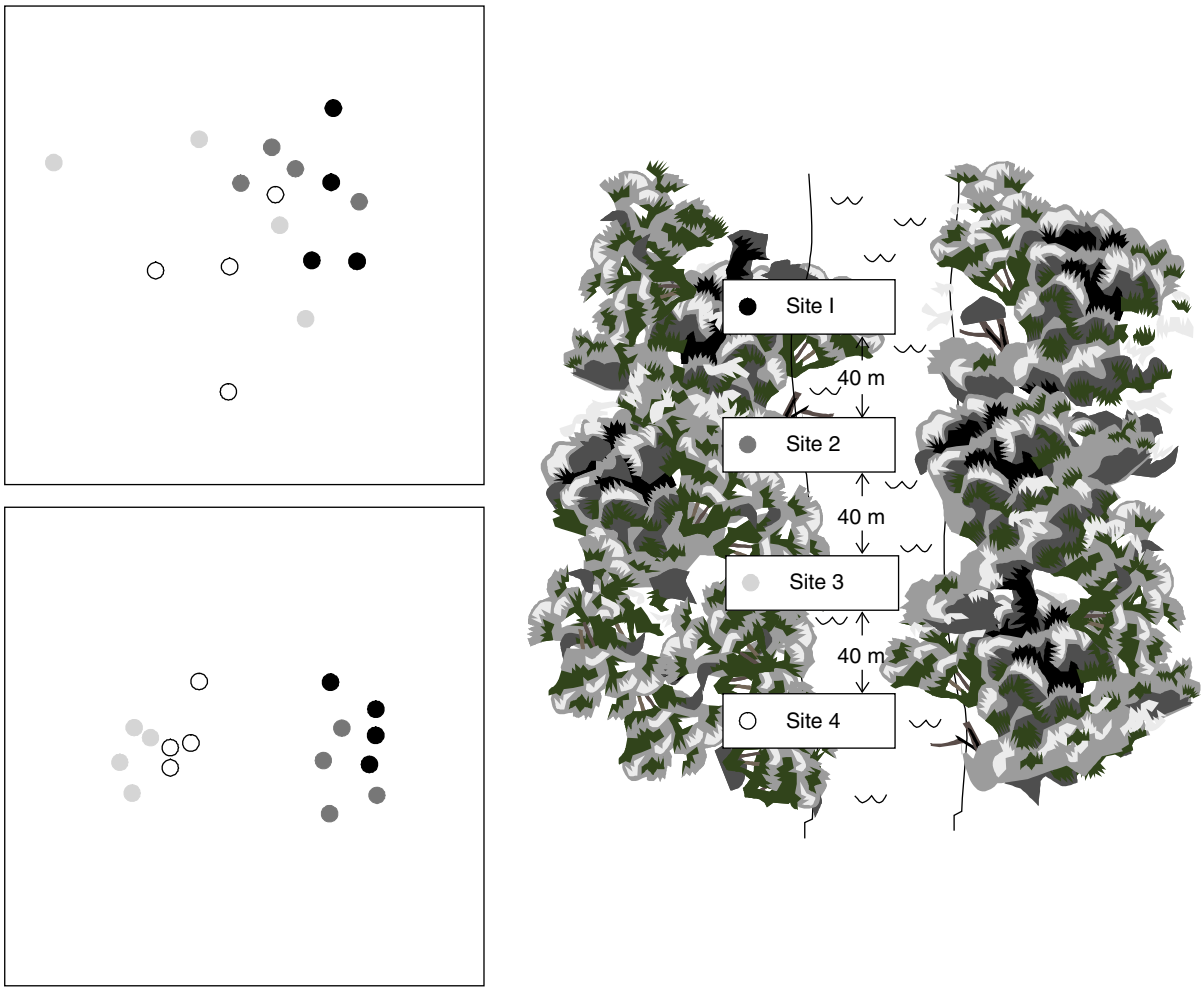


Figure 3 Variability in the numbers and types of animals living in a mangrove forest in New South Wales, Australia. The animals were sampled in four sites along the bank of a river. Each site was sampled using four quadrats, placed 1–2 m apart. Each point on this figure represents the types and numbers of animals in each quadrat. The relative distance between two points indicates how similar the animals were in those two quadrats. Points close together represent quadrats containing very similar types of animals. Points far apart indicate that the quadrats contained very different numbers and types of animals. **Figure 3(a)** shows that when the animals were sampled the first time, the diversity of animals was highly variable from one quadrat to another. Quadrats in the same site (meters apart) showed similar amounts of variability in these measures of diversity as shown by quadrats in different sites. **Figure 3(b)** shows what the patterns looked like only three months later. Animals were more similar at the scale of meters (among quadrats in each site), but the sites appeared to represent a gradient along the river. This relatively large change in the pattern of diversity was not seasonal, predictable, or found in other nearby sites. It was mainly due to changes in the relative numbers of numerous small crustaceans.

Sandy beaches and mudflats have very few large plants, although heaps of decaying plant material, commonly known as beach wrack, can be found on the strandline. This material – decaying seagrasses, seaweeds, and kelps, which have been washed up by the waves – is rapidly decomposed by bacteria, fungi, and small animals, many of which are dependent on wrack for food. This activity ultimately releases nutrients back into the coastal waters (Griffiths *et al.*, 1983).

Sandy beaches and mudflats appear at first glance to resemble deserts – there are often few animals to be seen, other than birds and a few crabs. The birds are generally visitors to these shores, feeding along the shoreline or over the extensive flats during low tide, but spending the rest of their time elsewhere. There are relatively few animals high on the shore, mainly semiterrestrial species, such as ghost crabs that spend the day in burrows above the level of high tide, coming out at night to feed. Nevertheless, despite first appearances, there is diverse life lower on the shore. As is the case in mangrove forests, however, most of the animals live under the sediments.

Very small animals live in the tiny spaces among the grains themselves. These include unicellular and multicellular animals from many different phyla. Many of these very small animals resemble juveniles of larger species. Many are also extremely simple. For example, although multicellular, they may only consist of a few cells and lack structures – such as limbs, kidneys, or a circulatory system – relying on exchange of substances through the body wall to acquire oxygen or eliminate waste (Brown and McLachlan, 1990).

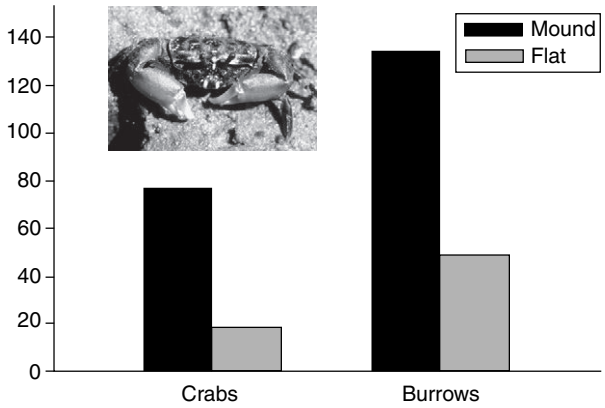


Figure 4 The crab, *Heloecius cordiformis*, and its burrows are more numerous on mounds of sediments than in the flatter patches of habitats in between the mounds. These mounds are created by the crabs themselves.

The larger animals living under the sand or mud either construct and maintain complex tubes or burrows or simply move through the sediment without leaving any structure behind them. On sandy beaches exposed to relatively large waves, the sediment is continually reworked by wave action. Most animals living in these habitats do not construct tubes or burrows because these are likely to be destroyed. Instead, they remain relatively deep in the sediment or tend to have mechanisms whereby they can rapidly burrow back under the sediment if exposed. Mole crabs have paddle-shaped limbs and can burrow extremely quickly during the time available between waves. Similarly, many bivalves have an extendible, mobile foot, which rapidly digs and anchors into the sediment and then contracts, pulling the bivalve under the sediment in only a few seconds.

Because mudflats tend to be very extensive and not steeply sloping, they usually remain quite damp during low tide. Many different species of worms, crustaceans, mollusks, and so on can be found in these habitats. Although some simply burrow through the sediment, feeding on detritus and particles attached to the grains, others live in semipermanent tubes or burrows. These can be very complex. Some burrows of crabs have many entrances and exits and are connected together in a complex underground labyrinth. Some animals living in burrows emerge to feed on the surface of the mud itself during high tide or at night. Other species remain in the burrow, but they may have long tentacles that are spread over the surface of the mud to pick up particles of food. Animals inhabiting burrows that feed on particles in the water often have modified limbs or other appendages, which can create currents of water through the tubes, pulling in the food with the water current. They also have very elaborate filters, which sieve the particles of food out of the water. These filters are modified legs, mouthparts, gills, or other organs.

Despite the fact that sandy beaches and mudflats are very common intertidal habitats, their patterns of biodiversity are not particularly well known. Available data suggest that, like most other intertidal habitats, they are extremely patchy and variable through time. The numbers and types of animals differ from one patch to another only a few meters away. These patterns are caused by the different species responding to subtle changes in the sediments themselves, to local disturbances such as feeding by fish, which creates depressions in the mud, and to other animals living in the sediment via interspecific processes, such as competition or predation.

Some species have rather aggressive behavior that spreads them out, presumably to ensure adequate food for all. Other species tend to aggregate and can create large patches of tubes or burrows, altering the sediments around them. These sorts of processes inevitably alter the small-scale local patterns of diversity, but there is very little information on how such processes may operate. Of course, as described earlier, such processes modify the initial patterns of diversity, which are established by the differential recruitment of species among different patches of habitat.

Measurement of Biodiversity in Coastal Habitats

Methods for and interpretations of measures of biodiversity in coastal habitats have proven complex. As reviewed by Gray (2000), the simplest measure – the number of species in any area – does not describe the structure and variation in assemblages. For this, measures are needed that can collate information on the types of species present (the composition of an assemblage) with information about their relative abundances. Many indices have been used (see Magurran, 1988). One of the most widely used is the Shannon–Wiener index:

$$H' = -\sum_{i=1}^{s} p_i \log_2 p_i$$

where p_i is the proportion of species i in the entire sample. Measurements of diversity in a single sample have been described as "point diversity." Diversity in a set of replicate samples from the same habitat is "alpha diversity." Comparisons from habitat to habitat, for example, along an environmental gradient uses measures of "beta (or between habitat) diversity." This is the combined diversity across a set of alpha diversities. There are also larger scale measures, such as "gamma diversity," which are measured across larger spatial areas, such as an entire coast-line.

There are serious problems of scale in these considerations. For example, a rocky headland could be a relatively large area; thus, diversity measured across it would be gamma diversity. Within the area, there are different habitats such as rock pools, algal beds, boulder fields, and so on. In each of these, samples would provide measures of alpha diversity. Beta diversity would measure the diversity in, say, algal beds along a gradient of increasing exposure to waves.

Suppose, instead, that the study was a series of headlands along the biogeographical gradient from south to north along a coastline. It is now likely that the diversity measured in a set of randomly taken replicates from one headland would be considered to be alpha diversity. That from the set of headlands would be beta diversity. Thus, whether or not a sample allows measurement of alpha or gamma diversity is entirely dependent on the spatial scale being examined. It is clearly important to define the hypotheses being tested very carefully so that only the relevant scales of habitats and sample areas are examined.

Another major difficulty in measures of diversity is that the number of species found is almost always dependent on the size of the sample examined or the length of time spent searching for individuals throughout the habitat. This is inevitable because of two distinct aspects of the patchy distributions of species. First, rare species are not going to occur very often in any sample and will often be missing from small samples. Considerable space needs to be covered to be sure that the rare species are actually encountered. Imagine a comparison of the animals in two habitats, which each have one hundred species in them. If, in one habitat (A), many more of the species are present than is the case in the other habitat (B) and similarly sized small samples are examined from each habitat, there will be apparently fewer species in Habitat A than in B. This is, however, an artifact of sampling.

A second reason for there being inevitable differences in the numbers of species seen in two samples where one is much larger than the other is due to the potential heterogeneity of habitats in the area studied. Suppose that the number of species is recorded in samples totaling 4 square meters of habitat in one area and 8 square meters in the other habitat. The latter sample would normally be expected to contain more species if there are many subhabitats in the area sampled. Thus, on rocky shores, clumps of seaweeds, small rock pools, patches of mussels, and so on may be scattered at scales of tens of centimeters. Each of these subhabitats probably contains several different species, in addition to "cosmopolitan" species found throughout the entire area. As a result, the larger area examined in samples from one habitat will encompass more different patches and therefore contain more species than would be found if a smaller area were examined.

There exist techniques to compensate for differences in the size of samples from one habitat to another. One of the most commonly used is called rarefaction, but this is known to have serious problems of overestimation of numbers of species in small samples (Fager, 1972). A superior alternative is to take random smaller samples from the largest, to match the size of the smaller samples from the other habitat. Such random samples can be used to provide reliable estimates and estimates of error associated with the number of species expected in smaller samples. Comparisons between the two habitats then consist of comparisons of the entire sample from the habitat that was sampled with smaller samples and the randomly determined estimates from the other habitat (reviewed by Gray, 2000).

Some multivariate methods are widely used to compare the diversity of species from one area or habitat to another. These operate on the principle that the abundances or biomasses of every species can be recorded from a series of replicate sample units (cores, grabs, quadrats, etc.). This produces a matrix of species by replicates for a sample from each habitat or area. The difference from any one sample unit to another can then be calculated across all of the species. One popular measure of such difference is the Bray–Curtis dissimilarity index (Clarke, 1993). This is calculated as follows:

$$D_{j,k} = \frac{\sum_{i=1}^{s} |X_{ij} - X_{ik}|}{\left(\sum_{i=1}^{s} X_{ij} + \sum_{i=1}^{s} X_{ik}\right)} \times 100$$

where $D_{j,k}$ is the dissimilarity between sample units j and k, X_{ij} is the abundance of species i in sample j, X_{ik} is the abundance of species i in sample k, and s is the total number of species found in the two sample units. If j and k come from the same sample, this is a measure of variation among units within a sample. If j and k come from two different samples, D is a measure of dissimilarity between the two samples. Because there are replicate measures, there is a sample of measures to estimate variation within each of the two samples and between the two samples. Statistical tests can then be carried out to determine the likelihood of there being more difference between two samples than expected by chance, given the variability within the two samples (Clarke, 1993).

Similar procedures are now being used widely in tests of the effects of environmental impacts and in many ecological studies to measure the diversity of animals in different habitats, where diversity is made up of the composition of species in any habitat and their relative abundances (many are summarized in Anderson *et al.*, 2008). Recent advances in these analytical techniques have increased the range of experimental designs for which they can be used (Anderson, 2001), and there are now many descriptions of complex patterns of natural variability in assemblages across a variety of spatial scales, and from time to time. More intensive sampling can yield large enough samples for independent measures of dissimilarity at different scales, allowing tests of more complex hypotheses about variability in diverse assemblages (Underwood and Chapman, 1998). Although very useful in measuring diversity, these analytical techniques do not, however, provide any clear understanding of the various physical and biological processes that cause such variations in diversity. To understand this, more knowledge is still needed of the ecological processes influencing the individual species that make up diversity in any place and at any time.

Can Intertidal Biodiversity Indicate Ecological Function?

Although it is clear that "biodiversity" covers many levels of variability, from genetic to population diversity within species, from diversity among species to diversity among habitats or ecosystems, the most common perception of biodiversity is still simply the numbers and types of animals and plants in different areas. Therefore, the importance of biodiversity has often focused on terrestrial habitats, the consensus of opinion being that there are more species on land than in the sea (Gray, 1993). Nevertheless, at different levels of organization, marine systems are more diverse – for example, there are more phyla in the sea than found on land. In addition, marine species appear to have more genetic diversity than related terrestrial and freshwater organisms. This suggests that they may be less vulnerable to processes that reduce genetic variation, although there is increasing concern that many nearshore aquacultural practices, including culture of fish and bivalve mollusks, are decreasing the genetic variability of natural populations.

Although many small marine invertebrates still remain to be identified and described, it is generally held that most diversity in the sea is of benthic organisms. There are many more

species of animals and plants living on or in close association with the sea floor, rocky reefs, and others, than living in the water column itself. This is probably related to the diversity of different benthic habitats. There is still controversy about the relative importance of deep-sea habitats and coastal habitats to benthic diversity – controversy arising from the fact that these habitats are so poorly described that the general patterns of biodiversity across large areas of marine habitat, even nearshore and intertidal habitats, are largely guesswork.

At whichever level biodiversity is examined, however, understanding patterns of biodiversity is essential to understanding the role(s) of diversity in the maintenance of different ecological functions. By functions, we mean such processes as maintaining diversity itself, recycling nutrients and chemicals among organisms and their surrounding habitat, and maintaining appropriate habitat for adequate food, shelter, and so on. Because patterns of biodiversity are complex – biodiversity varies at many spatial scales from centimeters to hundreds of kilometers and changes predictably and unpredictably among days, weeks, seasons, years, and eons – the role of diversity in maintaining the functioning of individuals, populations, assemblages of species, and habitats will also be necessarily complex.

Relationships between biodiversity and the persistence of successful ecological functioning may be described by quite different theories (see the review by Hooper *et al.*, 2005). First, there may be considerable redundancy of species, so that many species may be lost without impairment of function because other species simply take on their roles. This may be common in assemblages of very similar and small species, such as the many species of amphipods that live in some algal mats, or the many species of small burrowing worms in sediments. There must obviously be a limit to this; not all species can disappear and the ecology of the assemblage still persists. A second view is that many species may contribute to any ecological function, but the system may withstand loss of these species until a crucial limit is reached, after which there will be a rapid and inevitable degradation of function. The third view, called the idiosyncratic response, has been shown to be most likely in intertidal systems that have been examined. There is loss of function with loss of species, but this is variable and unpredictable because of the complex patterns of abundance and interactions among the various component species.

Most of the experimental tests of the relationships between biodiversity and ecological function have been carried out in terrestrial systems. The results of different experiments are not clear-cut, and there is considerable controversy about their designs and the ways in which they have been interpreted (Huston, 1997). Because of the fundamental differences in the ways in which different processes operate between marine and terrestrial habitats, it is unlikely that the results of such experiments, even if not controversial, could be applied to marine systems. Therefore, it is essential that the importance of biodiversity to the well-being and persistence of coastal habitats be tested experimentally in marine habitats.

Experiments have generally examined the removal of species from assemblages or the creation of artificial assemblages by the addition of different numbers and types of species. Despite differences between marine and terrestrial systems, recent research has indicated some commonality in changes to ecological functioning when the numbers of species in assemblages are altered. Frequently, the number of species matters less to the maintenance of some important ecological functions than does the identity of the species in an assemblage. Some species appear to have large and important effects, whereas many others have little measurable effect.

Nevertheless, although loss of such large and important species can alter many ecological functions, experiments have shown that not all changes were as predicted from current ecological understanding. This shows that there is still much to learn before the effects of changes to the diversity of even ecologically important species can be understood.

To date, the functional role of marine biodiversity at large scales has mostly focused on off-shore processes. For example, people are concerned about the diversity of plankton and its role in large-scale processes, such as commercial fisheries or atmospheric levels of CO_2 and O_2. In coastal habitats, most emphasis has been on the role of diversity in maintaining ecological function and has focused on single species or suites of similar species. Therefore, there is concern about the loss of large filter-feeding bivalves in estuaries and bays because the large numbers of these animals were considered to filter all of the water in the estuary in only a few days. In areas where populations have crashed, the water has become more turbid and is now occupied by different types of animals than was the case in the past. Similarly, certain large predators or grazers on intertidal shores or shallow reefs appear to play a pivotal role in maintaining patterns of biodiversity by selectively removing competitively dominant sessile animals or plants with their associated species, thereby providing space for a host of other organisms that thrive under the altered conditions (see the earlier discussion on keystone predators).

"Ecological engineers" (Lawton, 1994) have also received considerable attention because of their roles in altering the physical environment, creating habitat, and changing the availability of resources to other species. Many coastal animals are ecological engineers – the most obvious of which are corals that may develop into large reefs – a unique and very diverse habitat totally dependent on the corals themselves. Other lesser-known examples include burrowing bivalves and crabs in salt marshes, which can alter drainage, sedimentation, and erosion in addition to having direct impacts on other animals or plants that they may eat or for which they may be an important source of food. Some intertidal animals form habitat for a range of other species. Therefore, mussels of the genus, *Mytilus*, provide habitats for more than 300 other species of animals and plants on intertidal shores in Washington, in the US (Suchanek, 1992).

Mangrove trees stabilize sediments, modify shorelines, prevent erosion, and provide habitat for the juveniles of commercially exploited fish and shellfish. Many studies examining the relationship between structure and function in mangrove forests have focused on changes in the number of trees, especially in areas where they are cut down for firewood.

There has, however, been little consideration of the role of the actual biodiversity of such habitats – the mud-dwelling bacteria and small animals – on the well-being of the trees themselves, or, indeed, on the maintenance of the habitat itself. Unfortunately, as described earlier, little is known of this

Box 1 Despite the fact that there is no unambiguously accepted list of what is meant by the values of biodiversity, many values are cited as being important.

Evolutionary value. This includes genetic diversity that may allow organisms to persist or change in response to localized threats to diversity. This may be very important in intertidal and near-shore coastal habitats, where humans tend to cause most damage to marine habitats.

Ecological value. A naturally functioning habitat with its full complement of biodiversity is necessary for studying and understanding ecological patterns and processes. This understanding is, in turn, considered essential for the persistence of life on earth in the face of humanity's alteration to and degradation of the natural world.

Economic value. There are or may be species of direct economic value. On intertidal shores, this includes plants and animals collected or cultured in fisheries, mangrove trees exploited for firewood, and so on. Habitats that are used for recreation may also have economic value. For example, a wetland with diverse species of birds can have economic value if bird-watchers are prepared to pay to access it.

Aesthetic value. Natural habitats are sometimes perceived to have value simply because they exist. This is also sometimes called "existence value."

Ethical value. There is argument by many sections of society that all biodiversity has ethical value, that is, we, as human beings, are ethically bound to protect biodiversity. This argument comes to the fore when decisions are being made about the rights of humanity to eliminate diseases and other pests from the earth.

biota, other than the fact that it is variable, unpredictable, and complex. Because the species are poorly described and even more poorly understood, it is generally accepted that there is considerable redundancy in the system. Therefore, it has been assumed that, if the larger plants are looked after, the small invertebrate animals and the functioning of the habitat will look after itself.

The same is true of all intertidal habitats. Most ecological functions are thought of in terms of general, widespread, and large-scale processes (e.g., the cycling of nutrients throughout a mangrove forest or the maintenance of food webs across a rocky shore). Yet, the fauna and flora that make up most of the diversity and that maintain these ecological functions are variable and patchy and change unpredictably through time, often at very small and localized scales. The scales at which ecological structure (the species) and ecological functions (many of the processes) are measured are often quite different. It is therefore difficult to relate the stochastic variability that we see in ecological structure with what is often perceived as the predictability in measures of ecological function. It is increasingly important that we measure structure and function at a similar range of scales before we will be able to understand the role of natural patchiness in biodiversity in ecology. This is not yet widely done, but is crucial if we are to be able to conserve functioning habitats in an increasingly altered world.

What Services Does Intertidal Biodiversity Provide?

Consideration of the relationships between the diversity of species and functioning of ecological systems leads very rapidly to consideration of any "values" of biodiversity. Because biodiversity is not well described (i.e., the threats to it are poorly understood and probably variable from place to place and habitat to habitat), there is still considerable controversy about perceived value(s) of biodiversity. This is specifically true for most coastal habitats, especially those on temperate coasts, which have not received as much attention from the media and celebrities as have habitats such as coral reefs or tropical rainforests.

There are many so-called values to biodiversity, some of which are discussed in **Box 1**. Whatever terms are used, however, they tend to fall into two main categories. Ecocentric values are those associated with the well-being of the animals and plants themselves (i.e., they are focused on maintaining the full range of functioning ecological processes). They include such terms as ecological value and evolutionary value (see **Box 1**). Anthropocentric values, in contrast, are centered on humanity and the role of diversity in maintaining or improving human lifestyles. These obviously include a range of values, based on our perceived moral obligations, our ideas of the sort of world that we want to live in, and obvious economic returns from our exploitation of coastal resources.

In contrast to many terrestrial habitats, nearshore coastal habitats, including nearly all intertidal areas, are generally conserved for anthropocentric values, rather than ecocentric values. Following traditions lasting hundreds of years, the sea is still largely regarded as a larder to be plundered rather than a unique set of habitats to be conserved. Apart from certain coastal wetlands, which can support large populations of wading birds, most intertidal habitats do not abound with large charismatic megafauna. Therefore, although they may contain threatened species, these are unlikely to attract media attention.

Although large areas of mangroves can be extensive, diverse, and very interesting forests, they are still generally valued for their economic returns – provision of firewood, habitats for juvenile commercially exploited fishes, sites for aquaculture, and, unfortunately in many countries, prime real estate for reclamation. Similarly, rocky reefs, mudflats, estuaries, and so on are generally considered important for what direct economic value they can provide humanity.

The loss or degradation of nearshore habitats goes largely unseen because the biota that live in them are small, cryptic, and unknown. Yet, they are incredibly diverse and in many parts of the world, make up a large amount of the endemic fauna and flora. Most of these organisms do not have commercial value. As long as these coastal habitats continue to be considered valuable mainly with respect to the direct economic services that they provide to humankind, there will be little initiative to carry out the research and impose the necessary management to conserve this diversity.

Conclusions

The management of biodiversity in intertidal coastal habits needs a different approach from that used in terrestrial habitats. Except for species that are considered to play a key role in local ecological processes, it is unlikely to be useful to attempt

conservation on a species by species level. Exceptions are the so-called keystone species or ecological engineers, which were discussed in some detail earlier. Either these species provide habitat directly or their activities serve to enhance local diversity because of their elimination of competitively dominant species. For the most, however, emphasis on conservation of individual species is unlikely to be profitable.

First, for most species, we know too little about their requirements for habitat, food, and so on to be able to impose sensible managerial options. Second, we know too little about their natural patterns of variability in time or space to be able to evaluate whether any managed populations are persisting with anything like their natural patterns of abundance. Third, we certainly know too little about their interactions with other species to know which are essential to their continued well-being and which are not (i.e., which suite of species to try to conserve). In many cases, it has so far proven impossible to determine the geographical range of habitat occupied by a breeding population of intertidal animals, the extent to which there is interchange from one population to another, or the sources of recruits to any area that must be conserved.

Management of coastal biodiversity will better be focused on habitat. Removing or reducing disturbances to patches of mangrove forest, rocky shores, mudflats, and beaches is one of the more feasible and cheaper options – it is usually easier to prevent access to areas than to try to control people's behavior once they are in them. The fauna and flora will then be left to do the best they can under the circumstances. Fortunately, many species living in shallow-water, marine habitats are able to deal with many disturbances to their habitat, as long as these are neither too frequent nor too extreme. Evidence suggests that many can relatively rapidly recolonize areas once any disturbances are removed; such areas can then develop what appear to be functional ecological systems. Whether all species return to such habitats is not known because it is generally not known what species were there before the disturbance.

Nevertheless, there is obviously a need for more research on interactions among species wherever possible, so that keystone or engineering species can be identified. Where possible, management that will conserve these directly is likely to conserve those species that are dependent on them. Therefore, legislation that protects rocky shores from foragers removing mussels and other large shellfish will not only protect the targeted species but also a diverse range of other animals and plants. There is little chance that these could be protected on a species-by-species basis because too little is known about any individual species to mount a case for its protection.

Finally, with increasing understanding of the landscape ecology of coastal habitats and how "intertidal landscapes" interact, conservation of intertidal biodiversity is probably best within a mosaic of patches of different habitat. There is considerable debate about the value of few, large versus many, small reserves for conservation of terrestrial species. Which may be best is unknown for marine habitats and, in any case, there is no reason why what suits one species or set of species will suit another. The best strategy will probably be to hedge one's bets and try a range of different procedures. Therefore, with the current lack of understanding, conserving patches of habitat of different sizes and shapes, set at different distances apart, and subjected to as wide a range of environmental conditions as possible, will probably be the safest option until we know a lot more about our intertidal biodiversity.

Clearly, our understanding of the processes that influence and change biodiversity is increasing at a satisfactory pace. Our ability to use these insights in the planning and management of conservation is, however, more limited. The natural variability due to numerous processes and the large ranges over which larvae seem to disperse combine to make predictions difficult. The urgent tasks for coastal systems in response to threats due to coastal development, aquaculture, disposal of wastes, and potential rises in sea level require more understanding of the scales of variability in the distributions and abundances of coastal species.

See also: Coastal Beach Ecosystems. Disturbance, Mechanisms of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Function, Principles of. Mangrove Ecosystems. Marine Ecosystems. Measurement and Analysis of Biodiversity

References

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26: 32–46.

Anderson MJ, Gorley RN, and Clarke KR (2008) *Permanova+ for Primer: Guide to Software and Statistical Methods*. Plymouth, UK: Primer-E.

Brown AC and McLachlan A (1990) *Ecology of Sandy Shores*. Amsterdam: Elsevier.

Brown AC and McLachlan A (2002) Sandy shore ecosystems and the threats facing them: Some predictions for the year 2025. *Environmental Conservation* 29: 62–77.

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 18: 117–143.

Connell JH (1961) The influence of interspecific competition and other factors on the distribution of the barnacle Chthamalus stellatus. *Ecology* 42: 710–723.

Connell JH (1970) A predator–prey system in the marine intertidal region. *I. Balanus glandula and several predatory species of Thais. Ecological Monographs* 40: 49–78.

Connell JH (1978) Diversity in tropical rain forests and coral reefs. *Science* 199: 1302–1310.

Connell JH (1985) The consequences of variation in initial settlement versus post-settlement mortality in rocky intertidal communities. *Journal of Experimental Marine Biology and Ecology* 93: 11–46.

Connell SD and Gillanders BM (eds.) (2007) *Marine Ecology*. Oxford, UK: Oxford University Press.

Fager EW (1972) Diversity: A sampling study. *American Naturalist* 106: 293–310.

Fairweather PG (1985) Differential predation on alternative prey and the survival of rocky intertidal organisms in New South Wales. *Journal of Experimental Marine Biology and Ecology* 109: 135–156.

Foster MS and Schiel DR (1985) The ecology of giant kelp forests in California: A community profile. *U. S. Fisheries and Wildlife Service Biological Reports* 85: 1–152.

Gray JS (1993) Marine biodiversity: Patterns, threats conservation needs. *Biodiversity and Conservation* 6: 153–175.

Gray JS (2000) The measurement of marine species diversity, with an application to the benthic fauna of the Norwegian continental shelf. Journal of Experimental Marine Biology and Ecology. *Experimental Marine Biology and Ecology* 250: 23–49.

Griffiths CL, Stenton-Dozey JME, and Koop K (1983) Kelp wrack and the flow of energy through a sandy beach ecosystem. In: McLachlan A and Erasmus T (eds.) *Sandy Beaches as Ecosystems*, pp. 547–556. The Hague: W Junk.

Hooper DU, Chapin FS, Ewel JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75: 3–35.

Huston MA (1997) Hidden treatments in ecological experiments: Re-evaluating the ecosystem function of biodiversity. *Oecologia (Berlin)* 110: 449–460.
Hutchings PA and Saenger P (1987) *Ecology of Mangroves*. St Lucia, Queensland: University of Queensland Press.
Kastendiek J (1982) Competitor-mediated coexistence: Interactions among three species of benthic macroalgae. *Journal of Experimental Marine Biology and Ecology* 62: 201–210.
Lawton JH (1994) What do species do in ecosystems? *Oikos* 71: 367–374.
Lohse DP (1993) The importance of secondary substratum in a rocky intertidal community. *Journal of Experimental Marine Biology and Ecology* 166: 1–17.
Magurran A (1988) *Ecological Diversity and Its Measurement*. London: Croom Helm.
McGuinness KA (1987) Disturbance and organisms on boulders. II. Causes of patterns in diversity and abundance. *Oecologia (Berlin)* 71: 420–430.
Paine RT (1974) Intertidal community structure: Experimental studies on the relationship between a dominant competitor and its principal predator. *Oecologia (Berlin)* 15: 93–120.
Roughgarden J, Gaines SD, and Pacala SW (1987) Supply-side ecology: The role of physical transport processes. In: Gee JHR and Giller PS (eds.) *Organization of Communities Past and Present*, pp. 491–518. Oxford: Blackwells.
Scheltema RS (1971) Larval dispersal as a means of genetic exchange between geographically separated populations of shallow-water benthic marine gastropods. *Biological Bulletin* 140: 284–322.
Sousa WP (1980) The responses of a community to disturbance: The importance of successional age and species life histories. *Oecologia (Berlin)* 45: 72–81.
Stephenson TA and Stephenson A (1972) *Life Between Tidemarks on Rocky Shores*. San Francisco: WH Freeman and Co.
Suchanek TH (1992) Extreme biodiversity in the marine environment: Mussel bed communities of Mytilus californianus. *NorthWest Environmental Journal* 8: 150–152.
Terborgh J (1992) *Diversity and the Tropical Rain Forest*. New York: WH Freeman.
Thorson G (1950) Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews* 25: 1–45.
Underwood AJ (2000) Experimental ecology of rocky intertidal habitats: What are we learning? *Journal of Experimental Marine Biology and Ecology* 250: 51–76.
Underwood AJ and Chapman MG (1998) A method for analysing spatial scales of variation in composition of assemblages. *Oecologia (Berlin)* 117: 570–578.
Underwood AJ and Fairweather PG (1989) Supply-side ecology and benthic marine assemblages. *Trends in Ecology and Evolution* 4: 16–20.
Underwood AJ and Keough MJ (2001) Supply-side ecology: The nature and consequences of variations in recruitment of intertidal organisms. In: Bertness MD, Gaines SD, and Hay ME (eds.) *Marine Community Ecology*, pp. 183–200. Sunderland, Maine: Sinauer Associates.
Warren JH and Underwood AJ (1986) Effects of burrowing crabs on the topography of mangrove swamps in New South Wales. *Journal of Experimental Marine Biology and Ecology* 102: 223–235.
Warwick RM, Clarke KR, and Gee JM (1990) The effect of disturbance by soldier crabs, Mictyris platycheles H Milne Edwards on meiobenthic community structure. *Journal of Experimental Marine Biology and Ecology* 135: 19–33.
Wootton JT (1993) Indirect effects and habitat use in an intertidal community: Interaction chains and interaction modifications. *American Naturalist* 141: 71–89.

Introduced Plants, Negative Effects of

William G Lee, Landcare Research, Hamilton, New Zealand

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 501–515, © 2001, Elsevier Inc.

Glossary

Economic impact Capacity of weeds to limit productive use of terrestrial environments and the costs associated with their control in both managed and natural areas.

Environmental impact Ability of weeds to displace native fauna and flora, alter key ecosystem functions, and change disturbance regimes.

Introduced plants Species transported by humans to regions outside their natural geographic range.

Invasive weed Introduced plant species that establish self-maintaining populations and spread, with and without human assistance, into new areas where they frustrate human intentions in production and natural landscapes.

Introduction

Human migration, settlement, and trade over the past two centuries have produced a rapid globalization of floras with independent evolutionary histories that were previously kept apart by natural barriers. The scale of movement and mixing of floras in recent years is unprecedented in the earth's history, and very few parts of the world currently retain an exclusively native assemblage of plant species. Hundreds of naturalized plants, no longer dependent on humans for their persistence, occur in most regions and countries, where they form a significant and growing proportion of the wild flora. Over large continental areas the contribution of established introduced plants can reach 23% of the total flora (e.g., Canada), but this is greatly exceeded on many islands where the naturalized flora may outnumber the pool of native species (e.g., Bermuda, Ascension, New Zealand). Despite the many benefits provided by introduced plant species, there is growing concern at the environmental damage and economic costs of naturalized invasive plants in both productive and natural ecosystems. Important negative impacts are created by a small subset, probably less than 0.01% of all introduced species growing in a region, but their control, management, and impact consume significant economic resources, especially of developed countries.

The extent, patterns, and consequences of introduced plant invasions are described in another chapter presented in this volume, "Plant Invasions." This chapter highlights the negative impacts of introduced plant species using case studies of notable invasions, particularly those that have spread into natural areas with serious deleterious consequences for native species and communities. Examples of invasive plant species have been selected on the basis of a well-documented, and therefore relatively recent, spread and impact in natural environments. They are not necessarily the most widespread plant species, or the most aggressive, and some are limited to a few regions or countries. However, they are representative of the pattern and scale of intercontinental transfer of introduced plant species and the types of dramatic impacts these species can have when released into new environments. The examples chosen include a range of plant growth forms, altitudinal and bioclimatic zones, regions of origin, and modes of negative impact.

For ten introduced invasive plant species, representing algae (*Undaria pinnatifida*), ferns (*Salvinia molesta*), grasses (*Bromus tectorum, Spartina alterniflora*), small herbs (*Hieracium pilosella*), large herbs (*Fallopia japonica*), cacti (*Opuntia*), shrubs and small trees (*Lantana camara, Mimosa pigra*), and tall trees (*Pinus pinaster*), an account is given of the invasion process; local ecological, cultural, and economic impacts; attempts at control; and likely long-term consequences. Not all invasions are irreversible, and several examples illustrate successful levels of population control, if not eradication.

Notable Examples

Undaria Pinnatifida (Laminariales)

Some of the most pronounced ecological impacts and intractable problems of introduced species invasion involve marine organisms. Both deliberately and inadvertently introduced macroalgae have become invasive in recent decades with the increase in marine farming, transoceanic and coastal shipping, and the discharge of ballast water in and around foreign ports. The Asian laminarian kelp (*Undaria pinnatifida*), native to Japan, Korea, and China, has been introduced and spread since 1970 in other northern hemisphere waters (e.g., England, French Atlantic, and in parts of the Mediterranean near France, Spain, and Italy). More recently (1980s), populations have been discovered in the southern hemisphere (Argentina, Australia, and New Zealand), making it the first large kelp to cross tropical waters successfully. In its native range *Undaria pinnatifida* is widely cultivated as an edible seaweed, and new locations are often viewed positively for their commercial potential. However, the alga is highly invasive, has significant negative conservation impacts on the local native shallow marine fauna and flora, and may be impossible to eradicate.

Native to the north-west Pacific, *Undaria pinnatifida* has an annual, heteromorphic life cycle, with macroscopic sporophytic and microscopic gametophytic stages. The golden

brown, pinnatifid sporophyte, up to 3 m long, grows over autumn and winter, before degenerating in the summer. When growth slows, sporophylls develop as undulating extensions at the base of the stipe and release millions of short-lived (generally less than 2 days) male and female pelagic spores which, following settlement, develop into small gametophytes that may overwinter before producing either eggs or sperm. Fertilization of the female egg cell by male gametes gives rise to the straplike juvenile sporophyte. *Undaria pinnatifida* grows from low water to a depth of 20 m, depending on turbidity and wave exposure, and tolerates sea surface temperatures from 0 to 27 °C. In Japan, it is an important edible seaweed crop and is also cultivated extensively to feed juvenile abalone (*Haliotis spp.*).

Undaria pinnatifida first appeared outside its native range in 1971 in a lagoon on the south coast of France, in the Mediterranean Sea. The source is thought to have been contaminated oyster spat translocated from the Pacific. It has subsequently spread along northern parts of the Mediterranean from Spain to Italy. In 1983 *Undaria pinnatifida* was deliberately introduced to the French Atlantic for experimental aquaculture trials, on the basis that sea temperatures would be too cool (<14 °C) for the release of zoospores. This assumption was proven incorrect when ecological surveys during the 1990s revealed natural establishment at five of the nine cultivation sites, some now abandoned, with several new populations up to 20 km away from nearest putative sources. *Undaria pinnatifida* appeared on the south coast of England in 1994, the population mostly likely deriving from microscopic gametophytes or juvenile sporophytes attached to boats arriving from Brittany.

Spread of *Undaria pinnatifida* into the southern hemisphere has been very recent, beginning in the 1980s, and accidental, presumably from plants attached to ships or in seawater ballast. *Undaria pinnatifida* was first discovered in New Zealand in 1987, and has subsequently spread, mainly via sporophytes and gametophytes attached to hulls of boats, to more than ten localities, up to 1000 km apart (**Figure 1**). Soon after (1988) the kelp was discovered in Tasmania, in and around several major ports, and was recently (1996) reported from coastal waters on the Australian mainland. The first South American population was discovered in Argentina in 1992.



Figure 1 Intertidal kelp forests of *Undaria pinnatifida* recently established along parts of the eastern South Island, New Zealand. Photo: National Institute of Water and Atmospheric Research.

While fouling boats and underwater structures is of nuisance value, the full ecological impact of *Undaria pinnatifida* in these new habitats is hard to predict, as the species is a very recent invader in foreign waters, is mostly confined to highly modified habitats in harbors, and initially preferentially colonizes clean, artificial surfaces. Investigations in the Atlantic suggest that *Undaria pinnatifida* is preempted on denuded rocky substrates by several local native annual and perennial kelps. Because of this competition, it may eventually be restricted along the coast in northern France to the low tide zone. In Tasmania, following disturbance and dieback of the native kelp due to ocean warming, *Undaria pinnatifida* displaces *Macrocystis* kelp forests and dominates subtidal urchin barrens due to its high fecundity and fast growth rate. However, it is less successful against native seaweeds on exposed coasts with slow currents. In New Zealand, *Undaria pinnatifida* reduces habitat for native marine grazers by smothering coralline algal surfaces favored for larval settlement of the animals. The region has a depauperate flora of annual and early successional seaweeds, and *Undaria pinnatifida* has few effective competitors on fresh surfaces created by storms and echinoid grazing. Frequent wave induced substrate disturbances could consolidate the dominance of *Undaria pinnatifida* in many shallow marine habitats.

While the transoceanic and local spread of *Undaria pinnatifida* has been dependent on human activity, the alga has several attributes that make it an aggressive competitor in shallow marine habitats. Sporophytes densely colonize any unvegetated surfaces, and huge populations can establish within a year on boats, wharf piles, mooring buoys, and on breakwater or reclamation debris. These can provide multiple sources for short- and long-distance dispersal. Importantly, *Undaria pinnatifida* occurring outside its native range can, in warmer waters, produce several generations of sporophytes each year, while local kelp species are reproductively dormant for part of the year.

The negative effects of *Undaria pinnatifida* on marine aquiculture and numerous algal communities are potentially enormous, and most countries where the species has established are pursing some form of control. However, the large reproductive potential and cryptic dispersal phases of Asian kelp, together with the partial obscurity provided by the marine environment, combine to make detection and control of *Undaria pinnatifida* almost impossible. Reducing the rate of spread of *Undaria pinnatifida* may be practicable with improved ship hygiene practices in ports and the strict control of aquaculture ventures. Eradication of small populations, using direct removal of sporophytes by divers, or sterilization of mariculture equipment, may also be possible where pioneer plants are restricted to artificial structures (e.g., mooring lines), but this is unlikely to be achieved where *Undaria pinnatifida* has established in harbors and along exposed coasts.

Salvinia Molesta (Salviniaceae)

Ferns are rarely successful weeds in natural ecosystems, but a notable exception is *Salvinia molesta* (kariba weed, African

pyle, Australian azolla, water fern, giant azolla), a free-floating perennial aquatic fern native to tropical and subtropical areas of South America. In the later part of the 20th century it has been spread worldwide, becoming invasive in fresh waters in warm-temperate to tropical areas, and causing major disruptions to the utilization of important water resources.

Only formally described in 1972, *Salvina molesta* is a sterile pentaploid, probably of natural hybrid origin. Native to parts of Brazil and Argentina, it grows there in lagoons, swamps, and river margins with still or slow-moving freshwater, in a diverse community of floating and emergent plant species. In these situations *Salvinia molesta* occurs at low densities, does not form extensive mats, and is rarely invasive.

Morphologically the species is quite plastic, exhibiting several distinctive forms. At each node *Salvinia molesta* supports submerged (which function as roots) and floating leaves. The size, shape, and density of the foliage changes depending on environmental conditions (temperature and nutrient concentrations) and growth phase. When *Salvinia molesta* first establishes, usually in calm, sheltered waters, the plants are small (20 mm), with oval, flat floating leaves. With time, leaves become rounded and larger, and in the final mat-forming stage are folded, crowded, and between 50 and 60 mm across. *Salvinia molesta* is efficient at accumulating and reutilizing nutrients such as nitrogen.

Salvinia molesta has become widely distributed through aquarium and horticultural industries and is now naturalized in Australia (**Figure 2**), Africa, Madagascar, India, Sri Lanka, Southeast Asia, Philippines, Indonesia, Papua New Guinea, Fiji, and New Zealand. Dispersal between water systems is also invariably by human activities (e.g., boats, aquaria), but spread within and between interconnected water bodies is assisted by wind and water movement of vegetative fragments. Outside its native environment, *Salvinia molesta* shows phenomenal growth rates, via the expansion of auxiliary buds at stem nodes. Under optimum conditions (30 °C and high nitrogen levels), plant area can double every 8 days and biomass every 2.2 days. In Australia total floating biomass values of 1600 g^{-1} m^{-2} dry weight, or 400 t ha^{-1} fresh weight, have been recorded. Frequently *Salvinia molesta* forms mats 1m deep across extensive waterways.



Figure 2 Attempted mechanical removal of extensive mats of *Salvinia molesta* blocking large waterways in Queensland, Australia. Photo: Queensland Department of Natural Resources.

During the 1970s and 1980s, *Salvinia molesta* was seen as potentially one of the worst invasive weeds globally, with major ecological and socioeconomic impacts. The basis for this concern is well illustrated by accounts of the species invasion in Papua New Guinea. In the early 1970s *Salvinia molesta* became established in the Sepik River flood plain, and within a decade covered 250 km^2 of water surface. The weed impeded water transport, preventing fishing, trade between villages, and access to medical and educational facilities. Fresh water for humans, stock, and wildlife became restricted, and *Salvinia molesta* mats proved suitable habitats for insect vectors of serious human diseases such as malaria. Similarly, in Zimbabwe, *Salvinia* threatened utilization of Lake Kariba, a massive impoundment on the Zambezi River created in the 1950s for hydroelectricity generation, commercial fishing, recreation, and transport. Within 15 years of the scheme being completed, mats of *Salvinia molesta* were limiting human access, decreasing fish populations, and modifying nutrient regimes. Elsewhere, *Salvinia molesta* has also invaded rice fields, blocked irrigation ditches, eutrophied small water bodies, and smothered emergent, floating, and submerged indigenous aquatic plants. A severe reduction in light and dissolved oxygen, together with an increase in carbon dioxide and hydrogen sulphide, have had a drastic impact on most benthic biota.

The management and eradication of *Salvinia molesta* focused initially on mechanical removal and the use of chemical herbicides, and these were occasionally effective, especially where original populations were small or isolated. However, repeated treatment was required to prevent the rapid recovery of the plants, and the techniques used were expensive and occasionally environmentally damaging. Since 1980, biological control of *Salvinia molesta*, using a foliage- and stem-feeding weevil (*Cyrtobagus salviniae*—Curculionidae) from its native habitat, has been outstandingly successful in suppressing populations and achieving control in many countries (e.g., parts of Australia, Papua New Guinea, Fiji, South Africa). For example, the weevil had reduced the cover of *Salvinia molesta* on two 16 ha dams in Zimbabwe from over 90% to under 2% in less than two years. Similarly impressive results were achieved at a much larger scale in the Sepik River. Following the weevils' introduction, large mats of *Salvinia molesta* are typically reduced to sparse plants around the margins of water ways, interspersed with other aquatic plant species of the region. After the collapse of the *Salvinia* plants, weevil numbers also decline but appear to remain at densities sufficient to keep the weed in check. The global benefit of this successful example of biological control has been estimated to surpass $A200 million.

Bromus Tectorum (Poaceae)

The development of rangelands for pastoral use, through stock grazing, fire, and the deliberate and incidental introduction of new plant pasture species, has involved some of the most extensive modifications to natural ecosystems ever undertaken by humans. In the semi-arid intermontane west of North America more than 400,000 km^2 of native sagebush-steppe communities have been penetrated or displaced this century by *Bromus tectorum* (cheatgrass, downy brome, downy chess, drooping brome), a winter annual grass native to arid Europe and Central Asia. The dominance of this species has resulted in changes in land use, increases in erosion potential, reductions in native biodiversity, and alterations to the major disturbance regimes over large areas. *Bromus tectorum* is considered to be the most significant plant invasion in North America.

The native habitat of *Bromus tectorum* is centered on dry continental climates. Commonly associated with human activities, it occurs in grazed and ungrazed grasslands, among crops, and along roadsides. Flowers are hermaphroditic and self-fertile, and populations survive the dry summer period as seed. It characteristically germinates in autumn, estivates during the coldest parts of winter, and can grow rapidly, to around 50 cm tall in spring, when it flowers. The plants die back during the hot, largely rainless, summer. Seeds can be widely dispersed via animals, as they have long awns that are frequently embedded in fur or wool. Transfer by humans is usually accidental as a seed contaminant and among farm machinery. *Bromus tectorum* is now found in temperate East Asia, North and South America, and Australasia.

Prior to European settlement, the natural vegetation of the Great Basin area contained perennial cespitose grass-dominated steppes, the grasses (*Agropyron, Festuca, Stipa, Poa*) and forbs on mesic sites, giving way to shrubs (*Artemisia, Chrysothamnus*) in drier areas. The accidental introduction of *Bromus tectorum* into the intermontane region of the Pacific Northwest late in the 19th century, and several deliberate attempts to establish it as an alternative pasture species, coincided with large-scale domestic grazing by cattle, sheep, and horses. Because of the low abundance of large ungulates (bison and deer) during the Holocene, the original grasses and shrubs were generally poorly adapted to grazing ruminants. The flora also lacked aggressive colonizing annuals that could reclaim sites following disturbance. Burning by early settlers to remove shrubs and improve stock access further depleted the vegetation. Initially a weed of cultivated fields and roadsides, *Bromus tectorum* rapidly spread via wind and animal dispersal, and in seed lots, into overgrazed and vulnerable ungrazed native communities, becoming a widespread and dominant weed in the 1930s and reaching its current distribution in North America by the 1950s. Spread since then continues but at a slower rate.

The expansion of *Bromus tectorum* has caused multiple changes in natural ecosystems, largely through its impact on the grass/fire cycle. Accumulation of its highly flammable dry litter following spring growth, has resulted in an increase in the frequency of natural lightning-induced summer fires from 2 to more than 20 per century. Frequent burning has killed native shrubs and depleted the recruitment of perennial grasses. In contrast, *Bromus tectorum* benefits from fire, due to its large seed bank, prodigious and prolonged germination, rapid seedling growth, and heavy annual flowering. It is therefore able to preempt the recolonization of burned areas by native species already weakened by regular fires. Greater intensity and frequency of burning following *Bromus* invasion reduce plant cover, giving rise to erosion-prone landscapes. *Bromus tectorum* provides poor-quality herbage for stock, and its dominance has resulted in a decline in rangeland quality (**Figure 3**).

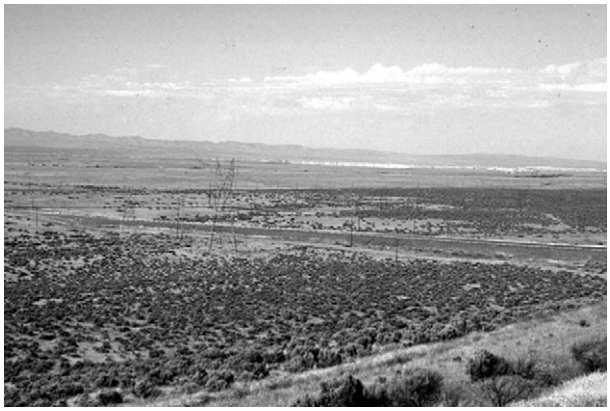


Figure 3 Extensive arid steppe in Washington State, USA, dominated by *Bromus tectorum* which, following fire and grazing, displaces native shrubs (foreground) and grasses. Photo: Richard Mack.

The outcome for conservation and pastoralism of the transformation, in less than half a century, of natural semiarid steppe dominated by shrubs and perennial grasslands, to communities predominantly of introduced annual grasses, remains uncertain. *Bromus tectorum* appears uncontrollable under extensive management systems. It is still spreading into more isolated areas of the Great Basin and into the Great Plains region of the United States.

Hieracium Pilosella (Asteraceae)

Numerous flat-weeds (rosette species) have accompanied European colonization of temperate regions around the world. Most of these plants are small, shortlived (2 years), dependent on human disturbance for local persistence, and generally inconsequential as weeds outside intensively managed systems. *Hieracium pilosella* (mouse-ear hawkweed), a perennial, stoloniferous, mat-forming daisy native to Europe, has a worldwide distribution in temperate regions and was generally considered to be relatively benign until it started to invade grasslands in New Zealand in the 1950s. Currently it occupies about 1 million ha of montane-subalpine short tussock (*Festuca novae-zelandiae*) grassland in subhumid and humid zones and is seen as the greatest threat to the conservation and long-term pastoral use of native grasslands. Key ecological factors responsible for the spread of *Hieracium pilosella*, and management options for its control, are controversial, and highlight the problems of large-scale weed management in nonforest ecosystems.

Native to Europe, *Hieracium pilosella* is a prostrate rosette-forming herb that reproduces vegetatively via stolons, which are initiated following flowering. Seed can be produced by both sexual and apomictic processes. In addition, hybridization and polyploidy are common and give rise to complicated population structures. Rosettes produce an erect short scape, which supports a solitary lemon-yellow flower that produces abundant small wind-dispersed achenes. In its native range the species is characteristic of disturbed dry habitats, most often grazed pastures, wastelands, and outcrops, particularly on calcareous soils below 1000 m above sea level.

Introduced into New Zealand in the middle of the 19th century, most likely as a seed contaminant, *Hieracium pilosella* remained locally restricted and largely insignificant for more than a century. The invasive phase of *Hieracium pilosella* became obvious during the 1960s, initially in the semiarid intermontane basins in the South Island, where the species became dominant in grazed *Festuca-Poa* grassland induced by pastoralism. Following burning of the original native forest and woodland during early Polynesian settlement, approximately 600 to 1000 years ago, tall tussock *Chionochloa* grassland and shrub land extended onto these areas. Early European pastoralists, in the first few decades of the 19th century, increased fire frequency and brought in sheep and cattle which, assisted by periodic eruptions of introduced rabbit populations, destroyed dominance by the tall cespitose grass species. The decline in plant cover and increasing bare ground provided numerous sites for invasive grazing-tolerant plant species. Foremost amongst these has been *Hieracium pilosella*, which slowly occupied the most depleted areas (those with 500 to 1200 mm rainfall) over about 50 years, becoming widespread and dominant in the 1970s when it began to extend into subalpine tall tussock grassland. In the 1990s *Hieracium pilosella* became the most widespread and significant introduced weed in native grasslands (**Figure 4**), accompanied in parts of its range by the introduced congeners *Hieracium praealtum* and *Hieracium lepidulum*.

During its colonizing phase *Hieracium pilosella* forms distinctive circular patches that gradually coalesce to cover large areas. In eastern parts of the South Island, instantaneous rates of increase of approximately 8% per year have been recorded, with local cover values of *Hieracium pilosella* exceeding 50%. Much of this local expansion is attributable to vegetative reproduction, with seedlings being rare. *Hieracium pilosella* smothers and displaces small native herbs and grasses and restricts the survival of productive pasture grasses and legumes. Soil changes beneath *Hieracium pilosella*, especially increasing rates of acidification, also limit the growth of introduced legumes.



Figure 4 Mass flowering of *Hieracium pilosella* in *Festuca-Poa* short tussock grassland, eastern South Island, New Zealand. Photo: Stan van Uden.

There is little doubt that human activities, especially repeated fire, and the grazing of domestic and feral animals, have increased the vulnerability of indigenous grasslands to invasion by introduced plant species. However, some relatively unmodified native habitats, especially in subalpine areas, on braided river beds, in semiarid environments, and on unstable hillslopes, may also be vulnerable due to tectonically induced natural erosion maintaining bare ground. *Hieracium pilosella* possesses several attributes that enable it to outcompete many low-growing native grass and herb plant species, especially in drier regions. These include rapid growth, vigorous vegetative reproduction, drought-tolerance, and grazing resistance. Furthermore, *Hieracium pilosella* is more efficient than native plant species in utilizing seasonal pulses of nutrients and water. The widespread aerial application of superphosphate fertilizer in upland areas may also have contributed to the spread of *Hieracium pilosella* in recent decades. New Zealand plants of *Hieracium pilosella* are entirely polyploid with high levels of genetic variation enabling the rapid evolution of innovative genotypes for colonizing new habitats.

The future of *Hieracium pilosella* in indigenous grasslands probably depends on land-use goals. In tall-tussock grassland, in the absence of mammalian grazing, natural vegetation cover and successional processes eventually leave *Hieracium pilosella* restricted to marginal habitats. However, the outlook for induced short-tussock grassland, with or without stock or feral animal grazers, is less clear, and *Hieracium pilosella* may well monopolize these communities in the longterm, or at least until succeeded by shrubs. Active management in areas under extensive pastoralism involves grazing to reduce flowering and applications of fertilizer to improve the competitive ability of associated grasses and legumes. A biological control program in New Zealand is currently underway, and two agents collected from Europe have been utilized: a pathogenic rust fungus (*Puccinia piloselloidarum* var. *hieracii*) and a gall wasp (*Aulacidea subterminalis*), which forms galls on stolons.

Spartina Alterniflora (Poaceae)

Estuarine habitats, characterized by waterlogged soils and regular saltwater incursions, provide a severe environment for plants and usually support a sparse halophytic flora. One kind of plant that is able to form dense monospecific stands in these environments is the grass *Spartina* (cordgrass), especially *Spartina alterniflora* and its derivatives, which have spread in estuaries in many subtropical and temperate regions of the world. The invasion of *Spartina alterniflora* is significant because of the rapid evolutionary process involved and the growing recognition of the biodiversity values of estuarine areas.

Spartina species are perennial, deep-rooted (30 cm), rhizomatous, sward-forming erect grasses that mostly occupy low- to midtidal mudflats. They are well adapted physiologically to tolerate saline conditions with special salt excretion, dilution, and restriction mechanisms. A high water-use efficiency, due to their C_4 photosynthetic system, and the unusual ability to maintain higher rates of photosynthesis than other C_4 and many C_3 species under cool temperate conditions (5–10 °C), increases their environmental range.

The native geographic distribution of the genus centers on the east coast of North and South America, with fewer species on the west coast of North America, Europe, and north Africa. *Spartina alterniflora*, native to the Atlantic coast of North America, was accidentally introduced to the United Kingdom early in the 19th century, and produced rare hybrids with the local congener *Spartina maritima*, which is restricted to western Europe and Africa. The male sterile F_1 hybrid *Spartina* x *townsendii* subsequently produced, via doubling of chromosomes, a new fertile species *Spartina anglica*. It quickly became apparent that *Spartina townsendii*, and particularly *Spartina anglica*, were markedly more invasive than either of the parent species. Estuarine habitats in southern England were transformed from exposed mudflats into tall-grass meadows. In less than a century *Spartina anglica*, via seed dispersal, clonal spread, and human plantings, covered approximately 10,000 ha of intertidal salt marsh along the coast of Britain.

Early this century, estuarine stabilization to maintain commercial waterways and reclamation for industrial and agricultural development were seen as important goals of coastal management. Experience in the United Kingdom resulted in *Spartina alterniflora* and its derivatives (*Spartina* x *townsendii*, *Spartina anglica*) being widely planted in parts of North America, in regions occupied by native *Spartina* species, and in areas well outside its native range (e.g., Australia, and New Zealand) where they were similarly successful, especially *Spartina anglica*. The heightened competitive vigor and environmental range of *Spartina anglica* appears to derive from a greater genetic heterozygosity, due to hybridization and polyploidy, and expanded phenotypic plasticity. Spreading *Spartina* species appear to have fewer insect and avian herbivores in their new ranges.

Invasive *Spartina* species induce major geomorpholgical changes in tidal habitats by enhancing rates of sediment accumulation. Sediment accretion rates of up to 17 cm yr^{-1} (average 5–10 cm) have been recorded within the tall grass sward, and in southern England *Spartina* caused elevation of mudflat surfaces by 1.8 m in 37 years. Declining tidal influence, and the accumulation of biomass (dry weight up to 7500 kg ha^{-1}), transforms salt marsh areas into nonestuarine ecosystems. Impacts on biodiversity are no less dramatic. At the pioneer phase, establishment via seed or planting produces circular clumps that expand at rates of between 3 and 7 m yr^{-1}, eventually coalescing to form extensive meadows on sand and mud substrates in the low- to midtidal range (**Figure 5**). *Spartina anglica* physically displaces many native indigenous low-growing halophytes (e.g., *Selliera*, *Salicornia*, *Schoenus*, *Puccinellia*) and modifies habitats formerly occupied by a range of bird, fish, and invertebrate species. In southern New Zealand, the loss of fish and other food sources in estuaries colonized by *Spartina anglica* threatens the persistence of 34 species of foraging shorebirds in these habitats.



Figure 5 Colonizing patches of *Spartina alterniflora* on mudflats at Wilapa Bay, Washington, USA. Photo: Glen Miller.

The hybridization of native and introduced *Spartina* species during the 19th century in the United Kingdom is presently occurring in western North America, where the recently established *Spartina alterniflora* (introduced) is genetically assimilating the common *Spartina foliosa* (native) through introgressive hybridization. Abundant *Spartina alterniflora* pollen improves seed set in *Spartina foliosa*, producing fertile hybrids that also cross with the native species. Locally, the introduced *Spartina* species and progeny could readily displace the native species and colonize lower reaches of the mudflats, currently beyond the environmental tolerance of the *Spartina foliosa*.

Because of growing appreciation of the biological values of estuaries, new plantings of *Spartina alterniflora* have largely ceased, and the species is now considered a serious weed. Extirpation of existing *Spartina* meadows has, however, been difficult, but containment and the removal of outliers has been achieved using systemic herbicides.

Opuntia (Cactaceae)

Cacti are prized as ornamentals for their novel form, spectacular flowers, edible parts, and medicinal uses. *Opuntia* (prickly pear) a genus of about 90 species native to North and South America, has been distributed globally over the past two centuries. Sixty species have become naturalized, and at least 15 are considered major weeds. *Opuntia* species provide some of the earliest and most spectacular examples of invasive plants in dry environments and the effectiveness I of insects as biological control agents. In the major centers of *Opuntia* invasion (Australia, South Africa, and India), several species have displaced horticultural crops, reduced the grazing potential of rangelands, and threatened the conservation values of native grasslands, scrublands, and woodlands. Eradication may be unattainable, but some invasive *Opuntia* species are now at low population densities, while others are continuing to spread.

Opuntia species are fast-growing perennial succulents with thickened, often flattened, segmented cladodes, usually supporting spines. Most *Opuntia* will regenerate from seed, cladode fragments, and underground tubers. Ranging in size from

low-growing shrubs to small trees, the species characteristically occupy dry habitats with seasonal water deficits. As with all cacti, *Opuntia* species have a distinctive photosynthetic system (crassulacean acid metabolism or CAM), which enables them to fix carbon at night when evaporative stress is reduced, as well as during the day, if adequate moisture is available. A high water-use efficiency, coupled with a large internal water-holding capacity and the ability to restrict water loss by tightly shutting off stomata, enhance drought tolerance. In North and South America *Opuntia* species are usually sparse and restricted regionally, and several are considered threatened in the wild. The major invasive weed species come from a range of native regions: *Opuntia stricta* (common or erect prickly pear) eastern North America, West Indies and adjacent South America; *Opuntia aurantiaca* (tiger pear, jointed cactus) temperate South America; *Opuntia ficus-indica* (sweet prickly pear, Indian fig) Central America; and *Opuntia vulgaris* (drooping prickly pear, Barbary fig) eastern South America.

Initially *Opuntia* species were distributed outside the New World as ornamental plants, but a range of uses rapidly emerged, especially as stock fodder, hedges, and as fruits or vegetables. A novel use was the production of carminic acid, a commercially important red dye, derived from the crushing of dried cochineal beetles fed a diet of cacti. *Opuntia* can provide an important source of food and income for local communities, and attempted extirpation is not universally accepted.

Introduced into Australia in the early 19th century, *Opuntia stricta* had by 1925 infested more than 25 million ha of eastern Australia and was spreading at the rate of 100 ha per hour, aided by the dissemination of vegetative and seed material by floods, humans, and feral and wild animals. Areas most susceptible were pastoral and arable land cleared amongst *Acacia* and *Casuarina* woodland, although the cacti also penetrated the woodland understorey. The Australian flora has few native succulents, and none with the environmental tolerance of *Opuntia*. Approximately 10 years after initial establishment at a site, ground cover became dominated by impenetrable thickets of *Opuntia* reaching densities of 16,000 plants and a biomass of 250,000 kg, per hectare. The weed frustrated the farming ambitions of European settlers and caused the ruin of early rural economies.

Opuntia ficus-indica, a tall shrub, was established in southern Africa at least 250 years ago. By the early 20th century, infestations occupied 900,000 ha, mainly in the Eastern Cape region, occupying grassland, succulentkaroo, and the savanna biome (**Figure 6**). *Opuntia aurantiaca*, a low-growing species, which may be of hybrid origin, was introduced much later (early 19th century), spreading originally from garden plantings around Cape Town. A total infested area of about 400 ha in the 1890s has grown this century to around 830,000 ha, mainly in eastern parts of South Africa. It is replacing important pastoral plants in grasslands and savanna, injuring domestic and feral animals, and degrading natural rangelands. A single cladode can produce up to 145 new cladodes over a 200-day growing period, and the potential for vegetative spread is enormous. Currently *Opuntia aurantiaca* is considered to be South Africa's most expensive weed. *Opuntia vulgare* has also been an important weed in the latter part of the twentieth century, mainly in western coastal areas.



Figure 6 *Opuntia ficus-indica* invading rangeland in the Eastern Cape Province of South Africa. Photo: John Hoffmann.

In Australia and South Africa, the control of *Opuntia* assumed top priority for land management agencies early in the 20th century, when various herbicidal and mechanical methods were attempted. Poisoning, using a combination of arsenic pentoxide and sulfuric acid, and mechanical cutting were successfully used to control small infestations of *Opuntia* in open lands but were very expensive and achieved little at a regional scale. However, spectacular success has been achieved using plant-sucking cochineal insects *Dactylopius* species and the cladode-eating larva of the moth *Cactoblastis cactorum*. Indications of the potential of insects as biocontrol agents first appeared in 1795, when the accidental introduction of *Dactylopius ceylonicus* resulted in widespread death of *Opuntia vulgaris* in India. Subsequently, deliberate introductions of several *Dactylopius* species have drastically reduced *Opuntia* weed infestations in southern Africa, and *Cactoblastis* has been similarly used to diminish population densities of *Opuntia stricta* in Australia. Overall, biological control agents have lowered population densities of *Opuntia* by 90%, especially in drier climates. *Opuntia* remains a widespread and invasive weed in these countries, but the economic impact on agricultural land uses has diminished hugely, especially in Australia.

Fallopia Japonica (Polygonaceae)

The human passion for gardens increases biodiversity in urban areas, often creating large source populations, which initiate the effective dispersal of introduced plants into adjoining landscapes. One garden ornamental that has become an invasive weed is *Fallopia japonica* (Asiatic knotweed), a large perennial herb native to the Far East. In the 20th century the species has infiltrated much of central and western Europe, North America, and several southern temperate countries, where it excludes native plant species in artificial and highly disturbed habitats, riparian areas, and in open woodlands. In the United Kingdom, *Fallopia japonica* is presently the tallest and most aggressive common herbaceous species.

Fallopia japonica is a rhizomatous, clump-forming, perennial native to China, Japan, Korea, and Taiwan. In these countries it is an early successional species, growing to less

than 2 m tall, colonizing volcanic debris and disturbed riparian habitats, under a range of edaphic conditions, from lowland to subalpine. *Fallopia japonica* establishes on new sites by seed in its native range and persists to develop a large biomass (12 tonnes ha^{-1}), deep litter layer, and a rhizome system that extends for up to 20 m beyond the shoots.

Taken from Japan to the United Kingdom in 1825, for planting in gardens, *Fallopia japonica* was subsequently established in other northern (e.g., North America) and southern hemisphere (e.g., New Zealand) countries late in the 19th century. Although mostly found on heavily modified and disturbed sites within urban environments, it has also spread in natural habitats, especially along waterways. In the United Kingdom it is presently found in over half the 10-km grid squares, being most abundant in southern and western regions (**Figure 7**). In Wales, for example, *Fallopia japonica* occupies 84% of river systems with average flow rates greater than 2.3 $m^3 s^{-1}$. In North America it has become naturalized in most mesic western coastal regions, north to Alaska, all of the northeastern United States, and parts of central and southern United States. In western Pennsylvania, stands extending for hundreds of hectares cover wetlands, stream sides, and adjoining hill slopes.

Around waterways, *Fallopia japonica* is capable of forming dense thickets up to 500 m across, overtopping and supplanting native herbs, shrubs, and small trees; impeding water runoff and increasing flooding, and modifying nutrient availability and cycling patterns through sequestration in a large standing biomass. Early emergence and fast growth of new shoots and flowering stems, usually completed before midsummer, create a dense overstory, and this, together with prodigious annual litter production, restricts the presence of other species, except where regular flooding removes surface litter, and the shade from tall trees reduces its vigor.

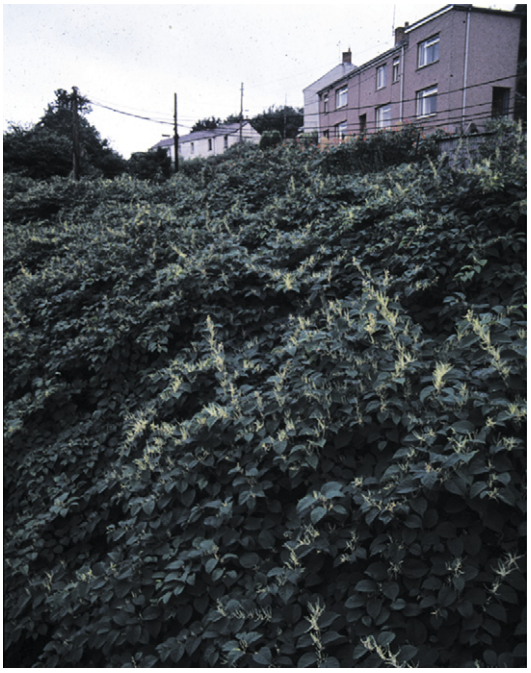


Figure 7 Complete cover of *Fallopia japonica* in urban wasteland, Swansea Valley, United Kingdom. Photo: Simon Fowler.

In its native range, *Fallopia japonica* bears female and hermaphrodite flowers on separate individuals. Elsewhere the species is male-infertile and, although seeds are produced from hybridization with other introduced congeners, spread via successful sexual reproduction is rare. However, spread of rhizome fragments is extremely potent, and viable plants can establish from pieces as small as 7 g, and from burial at depths of up to 1 m. Stem fragments are also able to produce new shoots. In most instances vegetative dispersal is achieved by water currents and the movement by humans of soil and gravel contaminated with rhizome pieces.

The current geographic range in Europe extends north to 63°N and is broadly correlated with degree-days, absolute minimum temperatures (> -30.2 °C), and annual rainfall (>500 mm). The shoots are vulnerable to late frost and summer droughts, and these factors may constrain the large-scale distribution of *Fallopia japonica*, especially in continental climates. Tolerant of a broad range of soil conditions, *Fallopia japonica* is strongly light-demanding and rarely enters forests.

Once established, *Fallopia japonica* is difficult to remove. In Europe and North America young shoots of *Fallopia japonica* are occasionally eaten by stock, and this may reduce the development of clumps and slow the rate of spread. Where accessible, regular cutting or spraying with herbicides can achieve control, but follow-up treatments over several years are needed to weaken and eventually kill the rhizomes.

Lantana Camara (Verbanaceae)

International trading in plants and the intensive search for novel cultivars for horticulture has produced new hybrids with high invasive capacity and given them near-global distributions and access to a great range of habitats. *Lantana camara* (lantana, white sage, wild sage, tick berry), a perennial shrub with brightly colored flowers, was taken from Brazil to Europe in the middle of the 17th century as a hothouse plant. During the next several hundred years, extensive hybridization and propagation of different varieties, augmented by new collections from the native region, produced a proliferation of forms. Commonly grown in gardens, *Lantana camara* has now become a major weed worldwide in tropical, subtropical, and warm temperate regions. It invades pastures, plantation crops, and a range of disturbed-open natural and modified scrubland, woodland, and forest communities, displacing the indigenous biota and limiting public access and use. Because of its broad distribution, invasive ability in both agricultural and natural ecosystems, and local persistence, *Lantana camara* is considered one of the world's top ten weeds.

Lantana camara is an aromatic shrub with distinctive four-angled stems, often armed with recurved prickles. It has multicolored, insect pollinated, bisexual, potentially self-fertile flowers that produce purplish-black drupes by assorted sexual, semisexual, and apomictic mechanisms. Native to parts of Central and South America, *Lantana camara* is actually a complex of species and varietal forms of mixed hybrid origin. It is a genetically diverse group, with ploidy levels ranging from diploid to hexaploid. The most weedy variety

(*Lantana camara* var. *aculeata*) is tetraploid. The growth form (e.g., ground creeper, prostrate shrub, tall shrub, liane) of *Lantana camara* is equally plastic, varying with light and soil conditions, but with support plants can reach 15 m into the canopy.

The species currently grows in many countries between 45° N and 45° S, but appears to be limited to minimum mean monthly temperatures above 5 °C, rainfall in excess of 650 mm per annum, and nonsaline soils. It is a serious weed in the Caribbean, in eastern Africa, South Africa, India, Madagascar, Australia, and the Pacific Islands.

Lantana camara is an important weed of plantation crops. In India, it has taken over land used for tea and sugarcane production, causing the displacement of whole villages, and in other countries it is a major nuisance in cotton, coffee, coconut, oil palm, bananas, pineapple, rubber, and rice crops. In continental countries such as Australia, India, and South Africa, it is natural grasslands that have been most extensively modified by *Lantana camara*. In eastern Australia alone approximately 4 million ha of grassland have become covered, eliminating both indigenous species and pastoral agriculture. Remnants of semideciduous forest are also invaded via edges and natural and human disturbances of the canopy. In South Africa, 44% of 1-km grid squares were infected with *Lantana camara* in a survey of 14 major forest reserves.

The spread of *Lantana camara* on remote volcanic islands with distinctive biotas has been of international concern and has usually been associated with European settlement and vegetation clearance via fire, roads, and grazing mammals (**Figure 8**). On the Galapagos Islands, famous for their biodiveristy and role in the origin of evolutionary theory, *Lantana camara* endangers both rare plants and animals. It is considered to have caused the extinction of one plant species and threatens the demise of at least eight others. The weed is also closing in on the last remaining colony of the darkrumped petrel (*Pterodroma phaeopyia*), and will, if not managed, seal off bird access to burrows. In the Indian Ocean, *Lantana camara* is one of several shrub-creepers invading and depleting native forests on La Réunion Island. Initially colonizing natural gaps, it subsequently ascends to smother and weaken canopy trees, making them more susceptible to damage during cyclones. It is predicted that *Lantana camara* will increase dominance with successive cyclones and eventually displace the forest cover. Thirteen years after its introduction to the Hawaiian Islands in 1858, *Lantana camara* had become naturalized on all the islands and is considered to be the number one problem plant. It currently occupies more than 160,000 ha, mainly in coastal and lowland areas, covering cropland, shrub land, savanna, and destroying opened up forest.



Figure 8 Flowering *Lantana camara*, a major threat to indigenous biodiversity on oceanic islands. Photo: Simon Fowler.

Apart from a broad environmental tolerance, the species has numerous traits that facilitate its dispersal, penetration, dominance, and persistence in different ecosystems. In tropical climates flowers and seeds may be produced all year, and fruits are eaten and widely dispersed by a range of birds and mammals, including some similarly invasive introduced species (e.g., Indian mynah, *Acridotheres tristis*). Most disturbed vegetation types are vulnerable to invasion by *Lantana camara*. Grassland, shrub land, and woodlands, from coastal to mountain areas, are transformed into dense stands of *Lantana camara* as it smothers herbs and other shrubs. Continuous closed forest is less susceptible as *Lantana* has greatly reduced vigor in deep shade. However, plants can penetrate forest from around the margins, readily establish in gaps created by natural treefall, and prevent the regeneration of canopy trees. Plants are flammable, even when green, and facilitate the spread of forest fires. Foliage and seeds are toxic for many grazing mammals and the thick, prickly stands of *Lantana camara* are virtually impenetrable. Thickets also provide habitats for vermin and disease vectors (e.g., tsetse fly).

Techniques for effectively controlling *Lantana camara* are slowly emerging, although most depend on the rapid establishment of vigorous crops or pastures following initial suppression of the weed. Mechanical methods are generally less successful, because *Lantana camara* generally resprouts from plant fragments, basal shoots, and roots after fire, cutting, or digging. Herbicides, especially those based on phenoxy or benzoic acid or pyridine, have variable success depending on climate, season, and growth form. Biological control appears to be sometimes effective, especially in drier areas, utilizing a mixture of seed- and foliage-feeding insects.

Mimosa Pigra (Fabaceae)

Nitrogen-fixing shrub legumes provide some of the most aggressive and disruptive invasive plant species in many parts of the world. One of the most spectacular plant invasions this century has been the spread of *Mimosa pigra* (giant sensitive plant, zaraz, dormilona) in tropical wetlands in parts of Asia, Africa, and Australia. In less than a century, 450 sq km of natural habitat associated with flood plain areas and rivers around Darwin, northern Australia, have been transformed into dense stands of *Mimosa* (**Figure 9**). The species endangers the conservation and use of natural wetland ecosystems in tropical regions worldwide.

Mimosa pigra is a prickly shrub native to Central and South America where it forms shrub lands up to 5 m tall in areas with seasonally high humidity. Its current pantropical distribution reflects human movement of plants since the 16th century, most likely because of fascination with the touch-sensitive rapid folding of the foliage. *Mimosa pigra* has



Figure 9 Dense stands of *Mimosa pigra* during the wet season, northern Australia. Photo: CSIRO Entomology.

naturalized beyond its native range in Asia (e.g., Thailand, Malaysia), America (e.g., Costa Rica, Brazil), Africa (e.g., Namibia, South Africa), and Oceania (e.g., Australia, New Guinea). Climatically, *Mimosa pigra* favors the seasonally dry tropical zone, with an annual rainfall of between 750 and 2250 mm, mean annual temperature above 17 °C, and mild winters.

Characteristics that make *Mimosa* such an aggressive invasive plant species include rapid growth rate (10 mm per day), rapid maturation (germination to first flowering within 6 months), potentially autogamous, abundant seed production (9000 seeds per m^2 annually), a large, long-lived (>10 years) seed bank in soil, and an effective dispersal system (flotation of clusters of capsules and via attachment to animals). Spines on the stem and leaf rachis deter most mammalian grazers and, outside its native range, there appear to be few invertebrate or fungal attackers. In many habitats, *Mimosa pigra* has the ability to completely dominate a site, forming impenetrable, monospecific shrub stands up to 6 m tall.

Rates and pattern of invasion have been best documented for northern Australia. Movement away from the entry point in Darwin has taken more than 50 years, but spread since the 1970s has been rapid. Within a river system stands expand on average at the rate of 76 m yr^{-1}, resulting in a doubling time of 1.4 years. Peripheral expansion is fastest following higher than average rainfall during the wet season, suggesting the importance of water for both the dispersal of seeds (in the wet season) and the survival of seedlings (through the dry season). An example of this spectacular rate of local spread has been given for a site of known history where a stand grew from a few individuals to one covering an area of 60,000 ha in approximately 10 years. At a larger scale, the doubling time for new infestations is slower (6.7 years), due to habitat heterogeneity.

The role of introduced grazing animals in the invasion of *Mimosa pigra* is debated. In northern Australia, the feral asiatic water buffalo (*Bubalus bubalis*) was thought by pastoralists to limit the spread of *Mimosa*, because of its conspicuous expansion since the lowering of buffalo densities as part of a disease (brucellosis and tuberculosis) eradication program in the mid-1980s. However, in the two decades prior to this, both buffalo numbers and *Mimosa pigra* infestations increased in parallel, and recent evidence shows that the most readily invaded habitats (wetland margins) are those most disturbed by large grazers, lacking tall trees, and with long periods of inundation.

Mimosa pigra drastically alters the composition and use of natural ecosystems. In Australia, it has supplanted native sedge land and grassland communities on flood plains and has invaded and displaced adjoining *Melaleuca*, *Eucalyptus*, and *Pandanus* woodland. The tall cover of *Mimosa* decreases native biodiversity and threatens the rich wildlife associated with open habitats. Ecosystem modifications caused by *Mimosa pigra* have reduced native resources accessible for traditional aboriginal use, pastoral grazing, and ecotourism ventures. The enormous scale and potential level of impact of *Mimosa pigra* is highlighted by estimates provided for Kakadu National Park (13,000 km^2), a World Heritage Area containing a full range of natural habitats within 150 km of the coast. Over 80% of the park is vulnerable to invasion (29% seasonally flooded areas susceptible to complete displacement; 54% largely open forest and woodland, which could have *Mimosa pigra* as a common element). Only 17% (extreme edaphic sites) of the area appears beyond the ecological tolerance of *Mimosa*.

The management and control of *Mimosa pigra* in natural ecosystems present a major challenge. It is susceptible to mechanical removal, herbicides, fire, and competition from grasses, and all are being used, often in combination, to eradicate new infestations and control the expansion in Australia. However, to be effective, treatments need to be repeated at regular intervals for at least a decade due to resprouting from damaged stumps and recolonization via the seed bank (10^4 seeds m^2). Currently an extensive biological control program is being undertaken testing several insects and pathogens.

Pinus Pinaster (Pinaceae)

Members of the Pinaceae, notably *Abies*, *Larix*, *Picea*, and *Pinus*, are the most widely planted exotic tree species in the world and form the basis of production forestry in many countries. The evolutionary history of the family is centered on the northern hemisphere, but in the 20th century the geographic range of northern conifer species has been greatly extended to many southern hemisphere regions where they have been evaluated and used for forestry, amenity planting, reforestation, and erosion control. Translocation of *Pinus* species in particular has occurred on a grand scale, initiated by European settlers. In South Africa, for example, more than 80 of the 111 species worldwide have been planted since the late 17th century. Outside their native range, at least 20 species have established away from plantations into adjoining grassland, shrub land, and woodland communities. Because of their size, growth rate, and fecundity, pines have been responsible for generating widespread concern internationally about the potential impacts of introduced plant species. *Pinus pinaster* (cluster pine, maritime pine), native to the Mediterranean region, typifies the invasion history and types of impacts that pines can have, particularly in nonforest ecosystems.

Naturally occurring in the Mediterranean Basin from Portugal to Algeria, *Pinus pinaster* is a two-needled, tall (40 m),

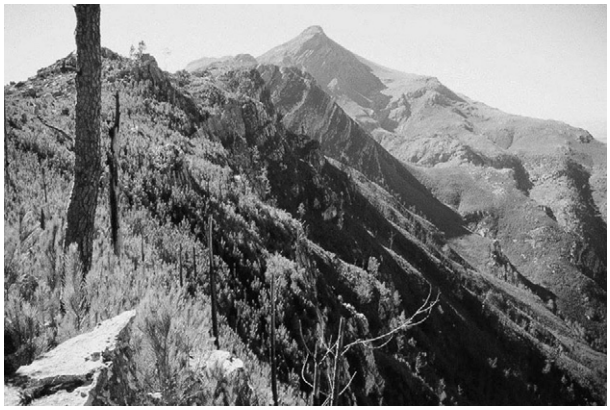


Figure 10 Thousands of hectares of *Pinus pinaster*, after 100 years of spread in the Riviersonderend Mountains near Genadendal, Western Cape Province, South Africa. Photo: David Richardson.

open-canopied pine species frequently growing on sandy soils and dunes among oaks and heathland shrubs at low to moderate altitudes, but extending up to 1500 m. Valued for resin potential and for erosion control, *Pinus pinaster* has been planted widely in South America, Australasia, and South Africa, where it has invaded a range of natural and modified communities. Plantations have also been established in Europe, within its native range, and these are similarly spreading.

The ecology and spread of *Pinus pinaster* has been best documented in South Africa, where it was introduced in the late 17th century and subsequently widely planted for timber and drift sand stabilization. Two hundred years later *Pinus pinaster* has invaded 3256 km^2 of natural vegetation (**Figure 10**) and is one of several pine species (*Pinus halepensis, Pinus radiata*) that threatens the conservation of one of world's biodiversity gems, the fynbos biome. Extending for 71,337 km^2 in the Cape Floristic region, the fynbos biome comprises complex, species-rich (7300 vascular plants), fire-adapted, mainly shrub-dominated ecosystems with a high proportion (80%) of local endemics. The pines have spread peripherally around source plantations and also via long-distance (up 5 km) wind dispersed seed. These outliers gradually coalesce to form extensive stands in which *Pinus pinaster* is more than 25% of the cover. On the Cape Peninsula near Cape Town, pines are the most common introduced plant species in 10 of the 15 vegetation types represented at the Cape, with *Pinus pinaster* dominant over 560 ha or 1.1% of the total area.

The success of *Pinus pinaster* and other invading pines is due to a combination of factors including particular species traits, natural features of the fynbos, and various human activities. Weedy pines, in this ecological context, are distinguished from nonspreading species by their resilience to fire (e.g., short period to reproductive maturity, above-average levels of serotiny) and dispersibility (small seeds with low seed-wing loading). The fynbos vegetation is fire prone, and natural fires at intervals of a decade or more maintain a mosaic of shrub land and gully forest, depending on the timing and intensity of the burn. Fire frequencies have increased with greater human activity in the region, creating ideal invasion opportunities for fire-adapted, fast-growing, and early-maturing introduced species. Plants derived from abundant seed sources blown in from established plantations have penetrated the fynbos after fire, increasing (by up to 300%) the available flammable biomass, and gradually consolidating dominance with successive fire cycles.

In the absence of management, *Pinus pinaster* and the other pines would eventually transform much of the diverse fynbos vegetation into an introduced conifer forest. The shift in life-form dominance threatens some rare native plant species and will suppress many other light-demanding plants, although they may provide new habitats for arboreal native birds. The large increase in stand biomass caused by pine invasion has increased interception and transpirational losses resulting in decrease stream flow (by between 30 and 70%) for human consumption from pine-dominated catchments. The pines have induced a new vegetation structure and system, mainly as a result of increasing the intensity of burning, and few native species can cope with the new disturbance regime.

Management for the conservation of the fynbos biome must include controlling the invasive pines. Where plantations adjoin watersheds, massive clearing operations are financially justifiable in view of the major impact of invading pines on water delivery from catchments.

Options currently being undertaken involve prescribed burning at intervals of 12 to 15 years, in conjunction with mechanical clearing of pine stands and outliers. Pines are felled 12 to 18 months before burning to allow time for seeds to be released from serotinous cones, and seedlings to establish. These are subsequently killed by the fire. Biological control approaches using seed-attacking insects are being explored for *Pinus pinaster* and *Pinus halepensis* as long-term solutions for effectively managing pines in fynbos landscapes.

Predicting and Evaluating Impacts

Most countries want to limit the influx of potentially invasive plant species and to identify and contain those that already exist within the large reservoir of introduced plants confined to cultivated areas. The commonest approach to improving biosecurity at the border and enabling recognition of new weeds at a very early stage of invasion when control is both cost-effective and practicable is to apply some form of weed risk assessment analyses. These invariably include evaluations of potential impact on human activities (e.g., agricultural production, animal health) and natural environments. Information is commonly derived from weed databases on the behavior of the plant species elsewhere in the world or from assumptions about the relation between plant traits and species impact. In general, features that enable a species to either capture or produce an important resource will have the greatest effect on plant community composition and structure and will initiate major changes in ecosystem processes and functions. Key traits vary depending on the local environment and the type of land management, but those that either assist species to acquire an unequal share of key resources (e.g., relative growth rate, height), foster new disturbance regimes (e.g., fire), enhance local persistence (e.g., long-lived soil seed bank), or create a major increase in resource availability (e.g., nitrogen-fixing rhizobia) are likely to be the most important

in any risk analysis. However, stochastic processes (e.g., change in land use), complexity of multiple interactions in natural ecosystems (e.g., presence of seed dispersers), and the increasing prevalence of introduced plant species lacking a weed history elsewhere collectively contribute to making weed impact prediction a major challenge. Methods for evaluating comparative benefits (e.g., food source for humans and native animals, erosion control) and costs (e.g., loss of crop production, reduced native biodiversity) of introduced invasive weed species are only slowly emerging due to difficulties in assessing qualitative and quantitative elements of these categories in simple economic terms.

Acknowledgments

This article benefited from comments from Stephen Brown, Garry Cook, Curt Daehler, Penny Edwards, Barrie Forrest, Peter Grubb, Simon Fowler, Mic Julien, Richard Mack, Wendy Nelson, Trevor Partridge, Ian Payton, Dave Richardson, Alan Rose, Leslie Seiger, and Peter Williams.

See also: Fires, Ecological Effects of. Introduced Species, Impacts and Distribution of. Migration. Plant Biodiversity, Overview. Plant Invasions

References

Brock JH, Wade P, Pysek P, and Green D (eds.) (1997) *Plant Invasions: Studies from North America and Europe*. Leiden: Backhuys.

Cowling RM, Richardson DM, and Pierce SM (eds.) (1997) *Vegetation of Southern Africa*. Cambridge: Cambridge University Press.

Cronk QCB and Fuller JL (1995) *Plant Invaders: The Threat to Natural Ecosystems*. London: Chapman and Hall.

Groves RH, Shepherd RCH, and Richardson RG (eds.) (1995, 1998) *The Biology of Australian Weeds*. RG Frankton and FJ Richardson.

Luken JO and Thieret JW (eds.) (1997) *Assessment and Management of Plant Invasions*. New York: Springer.

Pysek P, Prach K, Rejmanek M, and Wade M (eds.) (1995) *Plant Invasions: General Aspects and Special Problems*. Amsterdam: SPB Academic Publishing.

Richardson DM (ed.) (1998) *Ecology and Biogeography of Pinus*. Cambridge: Cambridge University Press.

Simberloff D, Schmitz DC, and Brown TC (eds.) (1997) *Strangers in Paradise: Impact and Management of Non-Indigenous Species in Florida*. Washington, D.C: Island Press.

Starfinger U, Edwards K, Kowarik I, and Williamson M (eds.) (1998) *Plant Invasions: Ecological Mechanisms and Human Responses*. Leiden: Backhuys.

Vitousek PM, D'Antonio CM, Loope LL, Rejmanek M, and Westbrooks R (1997) Introduced species: A significant component of human-caused global change. *New Zealand Journal of Ecology* 21: 1–16.

Walker DI and Kendrick GA (1998) Threats to macroalgal diversity: Marine habitat destruction and fragmentation, pollution and introduced species. *Botanica Marina* 41: 105–112.

Introduced Species, Impacts and Distribution of

Daniel Simberloff, University of Tennessee, Knoxville, TN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biological control Introduction of a natural enemy (parasite, predator, herbivore, or pathogen) of an undesirable species, usually itself introduced.

Coevolution Evolved mutual adaptations between species, in which each species influences the evolution of the other.

Hybridization Mating between individuals of two different populations, usually classified as separate species. Mixing of their genomes is not required.

Introgression Gene flow between populations whose individuals hybridize, achieved when hybrids backcross to one or both parental populations.

Island biogeography, theory of A theory that the number of species in the biota of each island (or island-like habitat) is in dynamic equilibrium, with species frequently going extinct on the island and new species frequently arriving.

Introduction

What constitutes an impact of an introduced species has never been formally defined. One common use of the term is for interactions with native species (e.g., predation), as well as consequences of these interactions for populations, communities, and structure and function of the ecosystem to which the species was introduced. Frequently this term refers to the economic cost generated by an introduced species (including loss of goods or services and costs of control). A range of less frequent meanings of "impact" will be clear as impacts are described. Nor is the meaning of "introduced species" clearcut. It is widely used for any species introduced to a new region by humans, whether deliberately or accidentally. Some authorities wish to restrict "introduced" to species deliberately imported by humans and describe species that arrive with human assistance but are not deliberately imported as "immigrants." In this restrictive definition, introduced species plus immigrants comprise the category "nonindigenous species." A few authors also include as introduced species those that arrive on their own in a location far from their original homes – for example, the Old World cattle egret (*Bubulcus ibis*) in Florida. Finally, an introduced species is often defined legally as one not native to a particular nation; a species carried by humans from one location to another within a nation would not be introduced. Many other common terms – alien, exotic species, bioinvaders – are less precisely used. In this article, introduced species will be those that arrive in a new region with human assistance, deliberately or inadvertently, whether or not the new region is within the national borders of the original range.

The impacts of introduced species are numerous, often subtle, and idiosyncratic. Because intensive study of invasions by introduced species is quite recent, many statements about their impacts are based on intuition rather than empirical data, and intuition is sometimes incorrect. Thus, the decline of the native otter (*Lutra lutra*) in the UK was long attributed to the introduction of the American mink (*Mustela vison*), whereas subsequent research exonerated the mink and implicated organochlorine pesticides (Chanin and Jefferies, 1978). Conversely, the destruction of the forest birds of Guam was at first attributed to pesticide poisoning before the introduced brown tree snake (*Boiga irregularis*) was proven to be the cause (Perry and Rodda, 2011).

Another complication in predicting and determining impacts of introduced species is that, after introduction, there is frequently a time lag before they spread and begin to exert an impact. For example, Brazilian pepper (*Schinus terebinthifolius*), which now infests *ca.* 300,000 ha in south Florida, was an innocuous, restricted species for over 50 years (Ewel, 1986). Sometimes the causes of these lags seem evident (e.g., delay in the arrival of an obligatory pollinator for a plant); other times they are mysterious. The opposite phenomenon, dramatic increase in numbers and impact by invaders, followed by a decline independent of directed human actions, is also occasionally observed. For instance, on several Pacific islands massively invaded by the terrestrial giant African snail (*Lissachatina fulica*), a rather abrupt decline in density was observed (Mead, 1961). Similarly, the Pacific alga *Caulerpa taxifolia*, after spreading rapidly for two decades in the Mediterranean, recently began to decline in striking fashion. Again, causes sometimes suggest themselves (e.g., a pathogenic disease), other times they do not.

Despite uncertainties surrounding impacts of introduced species, a recent burst of research shows that they are often dramatic and, in sum, constitute a major, ongoing global change. They are widely recognized as the second greatest cause of species endangerment and extinction (after habitat change, with which they often interact). A recent estimate by D. Pimentel and others suggests an annual global cost of invasive species greater than US$1.4 trillion (Pimentel *et al.*, 2002). This is not to say that most invaders have dramatic impacts. They do not; most do not even survive. M. Williamson's widely cited tens rule (about 10% of introduced species establish without human assistance, about 10% of those have impacts large enough to be seen as pests) is simplistic, but it accurately depicts the lack of substantial impact by most invaders.

The Geography and Magnitude of Invasion by Introduced Species

How Many Are There, Where Do They Come From, and Where Do They Go?

No comprehensive list of introduced species exists for most taxa in most locations; indeed, there is often no list of native species. However, for well-studied groups, some figures are impressive. In Florida, for example, 27% of established plant species, 8% of insects, 29% of land snails, 24% of freshwater fishes, 22% of reptiles, 5% of birds, and 24% of land mammals are introduced (Simberloff *et al*., 1997). For the Hawaiian Islands, as for many islands, some of the analogous figures are even greater: almost half the plants, 25% of insects, most freshwater fishes, 40% of birds (Staples and Cowie, 2001). In some areas (e.g., Alaska) there are far fewer introduced species, but almost no regions are immune.

Although the absence of adequate quantitative data makes it difficult to describe this pattern fully, it is widely believed that Eurasian species are more likely to invade other regions than vice-versa and more likely to have large impacts. For instance, most major human pathogens and most plant pathogens that have had global impacts originated in Eurasia. Similarly, more Eurasian insects, vertebrates, and plants have invaded other regions than *vice-versa*. The reasons for this imbalance are obscure. Some authors have argued that Eurasian species have an innate superiority, generated either by the larger numbers of species evolving greater competitive ability on the larger landmass, or by the happenstance that Eurasian species were highly coevolved – plants, pathogens, and animals (including especially grazers and humans) – and overwhelmed native species by their joint action. For example, the structure and behavior of Eurasian hooved livestock was devastating to native tussock grasses in the North American prairie but favorable for Eurasian turfgrasses, which now dominate vast regions (Mack, 1986).

However, even if Eurasian species, singly or in groups, were not innately superior, one might have expected a preponderance of them among introductions, because the opportunities for such species to reach other regions were greater throughout most of recent history. For instance, the majority of introduced insects in the US through the eighteenth century came from Europe in soil ballast, loaded in Europe and exchanged in North America for various raw materials (Lindroth, 1957). Similarly, as Europeans colonized other regions, they formed acclimatization societies to introduce the birds of their homelands (Lever, 1992). Except for gamebirds, there was little analogous movement of species in the opposite direction. Nowadays, as travel and trade are burgeoning worldwide, opportunities for introduction from any region to any other one are greatly enhanced, and one might expect habitat and climatic matching to become more important as limiting factors.

When Did They Get There?

The timing of introductions has depended heavily on available means of transport and patterns of travel and trade, and it has tended to increase strongly from the late eighteenth century through the present. For example, introductions of aquatic plants and animals into the Great Lakes rose steadily from one species between 1810 and 1839 to over 40 between 1960 and 1990, a surge associated with the opening of the Saint Lawrence Seaway (Mills *et al*., 1993). The advent of rapid steamship transport across oceans coincided with a dramatic increase in introductions of many taxa, as hitchhikers that could not have stayed alive over a month or more in transit were able to survive a voyage of two weeks across the Atlantic, for example. Air travel decreased the need to survive a long transit period still further. Overlaying this dominant pattern of increasing rates of introduction with increased transport volume and decreased transit time are idiosyncrasies associated with particular taxa and regions. For example, beginning around 1920, there was a decrease in the rate at which foreign insect species, especially herbivorous species, were introduced to the US; this downturn coincided with the enforcement of plant quarantine laws. At the same time, a dramatic increase of introduction of wasp species reflected greatly increased biological control efforts, especially the use of parasitic wasps (Simberloff, 1986).

Distribution of Introduced Species among Habitats

Habitats modified or routinely disturbed by humans generally house more introduced species and larger populations of introduced species than do pristine habitats, though even the latter are occasionally invaded. The reason for this pattern has been hotly debated. In general, the very fact of human disturbance renders a habitat less suitable to the native species that had evolved in the original habitat. However, some species, often associated with humans, are superbly adapted to habitats that humans create. Thus, certain introduced plants are routinely found on lists of serious weeds from many parts of the world. A common claim that disturbed habitats are invasible because they are species-poor is incorrect. Many habitats relatively undisturbed by humans, such as saltmarshes and mountaintops, have few native species, but they also have few introduced species, and for the same reason – they are biologically difficult environments and few species of any kind have physiological adaptations that permit them to thrive there. Conversely, enormously diverse tropical communities often have as many invaders (with as large impacts) as less species-rich temperate analogs. Nor does periodic disturbance per se automatically conduce to invasiveness. For example, the intact upland pine forests of north Florida are far less stricken by invaders than many other habitats in the state, yet they frequently burn naturally because of lightning strikes. If anything, this disturbance favors the many native species that are adapted to a fire regime over a plethora of potential invaders that are not. However, the increasing invasion of this region by fire-adapted cogon grass (*Imperata cylindrica*) poses a threat to the entire forest ecosystem (Schmitz *et al*., 1997).

Most well known examples of introduced species with enormous impacts come from either terrestrial or freshwater habitats. Marine habitats are less represented in the invasion literature. Freshwater habitats – lakes and rivers – can often be seen as habitat islands, surrounded by land, and thus their communities are believed to be inherently invasible for the same reasons that islands appear to be particularly invasible,

as will be discussed in the following section. Certainly, many lake and river communities have evolved without certain types of organisms, such as large predatory or herbivorous fishes, so their species are particularly prone to damage from introduced species of these particular sorts. Whether marine habitats have suffered less impact from introduced species is highly questionable. Terrestrial habitats are immediately visible, so the effects of introduced species are surely more likely to be detected there than in the sea, just as the general ecology of most marine habitats is poorly studied and understood relative to that of most terrestrial habitats. Additionally, it is now widely recognized that many marine species that had been viewed as native over wide regions are in fact of unknown provenance – they are "cryptogenic." In most places they may well have been accidentally introduced by humans. Finally, a number of marine invasions have been observed lately that vie in apparent impact with the most heralded terrestrial and freshwater invasions. The tropical alga *Caulerpa taxifolia*, purged from the public aquarium in Monaco in the early 1980s, spread to cover over 10,000 ha of nearshore habitats off the coasts of Spain, France, Italy, Monaco, and Croatia. It particularly overgrew and replaced meadows of marine grasses such as *Posidonia oceanica*, the nursery and home of many fishes and invertebrates (Meinesz, 2001). As noted previously, populations of *C. taxifolia* have recently begun to recede, but the related Pacific alga *C. racemosa*, which arrived in the Mediterranean later, has spread widely. In the Black Sea, the western Atlantic ctenophore *Mnemiopsis leidyi* was first recorded in 1982. By 1988, its biomass exceeded that of all other zooplankton combined, and, as this comb jelly is a predator, its impact on other animals, including young fishes, was enormous. It caused a dramatic crash of several commercial fisheries. About a decade later, another comb jelly from the western Atlantic, *Beroe ovata*, arrived in ballast water, preyed on and greatly reduced populations of *M. leidyi*, and the Black Sea ecosystem began to recover (Shiganova, 2011).

Island Vulnerability

Islands are notoriously prone to ecological damage by introduced species. Many island bird species and subspecies have been eliminated by introduced rats, feral cats, and other introduced carnivores. In the West Indies, the small Indian mongoose (*Herpestes auropunctatus*) has extinguished several species and subspecies of endemic snakes and lizards (Hays, 2011). Introduced grazers have caused many plant extinctions on islands. The introduced carnivorous rosy wolf snail, *Euglandina rosea*, has eliminated at least seven species of terrestrial snails on the Pacific island of Moorea alone, as well as many populations of endemic tree snails in the Hawaiian Islands (Cowie, 2001). Worldwide, about 1.6 times as many mammal species and three times as many bird species have been successfully introduced to islands as to mainland areas (Groombridge, 1992).

The reasons for the great number of introduced species on many islands, and their apparently great impacts on many of them, are probably threefold. First, there are generally more deliberate attempts to introduce species to islands. As mainland inhabitants colonized islands, they saw them as biotically impoverished and attempted to introduce species – often those from home – to fill the perceived gap. Second, many of the islands that have been most assaulted by introduced species were massively modified largely independently of introductions. For example, almost all the lowland areas of the Hawaiian Islands, Mascarenes, and many of the West Indies were cleared of native species for agricultural and housing purposes. That the introduced species largely replaced them does not necessarily mean that the native species were in some sense weaker than the invaders; rather, the habitats to which the native species were adapted were simply removed. Because of the smaller size of islands than mainland, habitat destruction on islands was generally much less likely to leave refugia in which the native species persisted. Third, the biotae of oceanic islands often evolved without entire groups of organisms, and so lacked adaptations to survive their impacts once they were introduced. Thus, most oceanic islands lack native predatory mammals, so birds evolved to nest on the ground and were highly susceptible to the invasion of such species. Similarly, native browsing and grazing herbivorous mammals are lacking from islands, so the native plants have not evolved traits to survive herbivory when mammals such as goats and reindeer are introduced.

Direct Effects

Habitat Modification

The most widespread impacts of invaders are usually due to modification of the habitat, which affects many species together. Almost all introduced species believed to have had major impacts on entire ecosystems or landscapes have done so by modifying habitats.

Grazing and Rooting

Goats were introduced to St Helena in 1513 and built up enormous herds. These are believed to have extinguished at least half of some 100 endemic plant species before botanists had even visited the island (Cronk, 1989). On the island of South Georgia, grazing by introduced reindeer (*Rangifer tarandus*) has heavily degraded tussock grassland, mesic meadow, dwarf-shrub sward, and dry meadow communities (Leader-Williams, 1988). These changes have facilitated erosion and invasion by exotic plants. In many places, these changes resulted in a totally new plant community. Recovery of certain plant communities, and the insects that inhabit them, may take a long time, but exclosure experiments show that few of the changes are irreversible. European wild boar (*Sus scrofa*), feral domestic hogs, and hybrids between them have wreaked havoc in both bottomland and upland forests in the US. For example, in the Great Smoky Mountains National Park, they "root" primarily in high-elevation deciduous forests in the summer, reducing understory cover and number of species. By selective feeding, they can locally extinguish plant species with starchy bulbs, tubers, and rhizomes. They greatly modify soil characteristics by thinning the forest litter, mixing organic and mineral layers, and accelerating mineral leaching (Singer *et al.*, 1984).

Shading and Overgrowth

South American water hyacinth (*Eichhornia crassipes*) blankets many nearshore areas of Lake Victoria, some lakes and rivers of the US Southeast, and many water bodies in other regions. It blocks light and smothers beds of native submersed vegetation. Decaying water hyacinth can deposit over 1000 metric tons (wet weight) of detritus per hectare per year, which in turn can heavily modify water chemistry, such as nutrient concentrations. Dissolved oxygen drops, affecting many animals. The entire biotic community changes (Hill *et al.*, 2011). Terrestrial plants such as Asian kudzu (*Pueraria montana*) in the southeastern US can similarly cover existing vegetation and eliminate it by shading, as can marine plants such as the tropical algae *Caulerpa taxifolia* and *C. racemosa* in the Mediterranean.

Modified Fire Regime

In much of the American West and Hawaii, Old World grasses such as cheatgrass (*Bromus tectorum*) increase the frequency and intensity of fires, greatly harming native plants and the animals that use them (Mack, 2011). Similarly, the Australian paperbark tree (*Melaleuca quinquenervia*) has a spongy outer bark and highly flammable foliage, and it also produces a more abundant litter than native herbaceous communities in south Florida. These features have produced a changed fire regime that has facilitated the invasion of 100,000 ha. Both tree and ground fires are now more intense and frequent, to the detriment of the native plants (Serbesoff-King, 2003). An added threat to the region is the recent arrival in the area of Old World climbing fern (*Lygodium microphyllum*), which transmits fires to the top of the tree canopy (Volin, 2004).

Modified Hydrology

Mediterranean salt cedars (*Tamarix* spp.) have invaded the US Southwest. They are deeply rooted and transpire rapidly; once established, they can survive on water deep in the soil, and their transpiration is a significant pathway of water loss in arid areas. For example, at Eagle Borax Spring in Death Valley, California, within 25 years of invasion by salt cedar, the surface water of what had been a large marsh had disappeared completely, along with the majority of its associated biota (McDaniel *et al.*, 2005). In Israel, Australian *Eucalyptus* trees have been used deliberately to drain swamps and bogs, thus eliminating the original vegetation and animals dependent on them (Biger and Liphschitz, 1995). In addition to changed evapotranspiration rates, plants can also affect hydrological regimes by changing soil elevation. Off the southern coast of Great Britain and parts of the coast of the state of Washington in the US, the cordgrass *Spartina anglica* has elevated intertidal areas and transformed gently sloping, productive mudflats into poorly drained marshes with enormous impact on resident animal species. *Spartina anglica* is a species formed in Great Britain in the late nineteenth century by the hybridization of introduced North American cordgrass (*S. alterniflora*) with native European *S. maritima* (Simberloff, 2011).

Modified Nutrient Regime

Many plant invaders fix nitrogen, and, in nitrogen-poor areas, this added increment can favor other invaders and be detrimental to native species, which have evolved to thrive in a low-nitrogen environment. The Atlantic nitrogen-fixing shrub *Morella faya* has invaded young, nitrogen-deficient volcanic regions of Hawaii. There are no native nitrogen-fixers, and the invader thus alters productivity, nutrient cycling, and ecological succession. Because many nonindigenous plant species in Hawaii are more successful on more fertile sites, *M. faya* may enhance the likelihood of other invasions (Vitousek, 1986).

Competition

Some introduced species use a particular resource so effectively that they deprive native species of it. In addition, an introduced species can depress a native one not by eliminating some resource but by direct interference.

Resource Competition

Shading and water depletion, already discussed, are means by which some introduced plants outcompete native species for a crucial resource. Introduced animals have similar effects. For instance, the house gecko, *Hemidactylus frenatus*, has invaded many Pacific islands, and it depresses the insect food base locally so that some native lizard populations decline (Petren and Case, 1998). In Great Britain, the greater foraging efficiency of the introduced American gray squirrel (*Sciurus carolinensis*) has led to the decline of the native red squirrel (*S. vulgaris*). The impact of the invader is enhanced by the fact that it carries a pathogen to which it is resistant but that devastates the native squirrel (Gurnell *et al.*, 2004).

Interference Competition

The South American fire ant, *Solenopsis invicta*, which has spread throughout much of the southeastern US, attacks individuals of many native ant species and is replacing some species, such as the "native" fire ants (perhaps themselves pre-Columbian introductions) *S. geminata* and *S. xyloni* in several habitats (Sanders, 2011). In a plant analog of aggression, the African crystalline ice plant, *Mesembryanthemum crystallinum*, accumulates salt (Vivrette and Muller, 1977), and the salt remains in the soil when the plant decomposes. In California, this ice plant excludes native plants that cannot tolerate the salt (Vivrette and Muller, 1977). In both of these examples, the invader does not actually render a resource in short supply for native species; rather, it directly inhibits the native, preventing its access to the resource.

Predation

Some predaceous invaders have devastated particular native species or groups of them. The Nile perch (*Lates nilotica*), introduced to Lake Victoria, eliminated many species of endemic cichlid fishes by eating them (Pringle, 2011). Introduced rats (*Rattus* spp.) on over 80% of the world's archipelagoes have destroyed at least 37 species and subspecies of island birds worldwide (Pascal, 2011). The brown tree snake and the rosy wolf snail, mentioned previously, are predators that have extinguished many native prey species. These are all dramatic examples of how an introduced predator can affect a native fauna that has not coevolved with it. However, even prosaic

introduced predators may have a substantial impact on native prey. For instance, there are up to 30 million feral housecats in the US, and in Wisconsin plus Virginia alone they are believed to kill more than six million birds annually (Winter, 1999).

Many predators have been introduced deliberately for biological control of previously introduced species, especially when the latter inflict losses on agriculture. The introduction in 1889 of the Australian vedalia beetle (*Rodolia cardinalis*), which controlled cottony-cushion scale (*Icerya purchasi*) on California citrus, is an early example, and there are many other successes (Simberloff, 2012). More recently, biological control has been employed for conservation purposes. On the island of St Helena, a tropical American scale insect (*Orthezia insignis*) that had threatened the native gumwood tree (*Commidendrum robustum*) was controlled by introducing a South American lady beetle, *Hyperaspis pantherina* (Booth *et al*., 2001). However, some predators introduced for biological control have devastated nontarget native species while exerting little if any control on populations of the target. The rosy wolf snail (*Euglandina rosea*) discussed previously is a prime example; it was introduced to many islands to control the giant African snail but failed to do so because the latter snail is too large. Similarly, the small Indian mongoose, previously mentioned as having extinguished native reptiles in the West Indies, was introduced there and on the Hawaiian Islands, Fiji, Mauritius, Okinawa, and other islands for the biological control of rats. On all these islands, it preyed on native, nontarget species, sometimes to the point of endangerment or extinction. The mosquitofish (*Gambusia affinis*) from Mexico and Central America, introduced in Europe, Asia, Africa, Australia, and many islands for mosquito control, is widely seen in many regions as being at best no better than native predators at controlling mosquitoes. It often preys on other small fishes as well as invertebrates, some of which prey on mosquito larvae. For instance, in Australia, it is implicated in the extinction of several endemic fish species through predation and both interference and resource competition (Pyke, 2008).

Herbivory

The extinction of half of the native flora of St Helena by goats, discussed earlier, shows that herbivory by introduced species on native ones can have major impacts. Other examples abound in a variety of habitats. Chinese grass carp (*Ctenopharyngodon idella*) have been introduced to Europe, Africa, and North America to control aquatic plants. Although the targeted plants are generally introduced species, the grass carp is not highly selective and often prefers native plants. It is so voracious that it can completely eradicate nontarget species from particular water bodies before foraging on the target species (Bain, 1993). Herbivory by introduced European rabbits (*Oryctolagus cuniculus*) has caused enormous damage on islands worldwide. Their impact in Australia is legendary. The most serious damage there is by the stripping and killing of seedling trees and perennial shrubs. In addition, rabbits often cause extensive erosion (Thompson and King, 1994).

Innumerable crop plants, most of them introduced, have been devastated by introduced insects. Damage in the US from the Russian wheat aphid (*Diuraphis noxia*) alone exceeded $600 million between 1987 and 1989 (US Congress Office of Technology Assessment, 1993). The South American cassava mealybug (*Phenacoccus manihoti*) has invaded most cassava-growing regions of Africa and causes yield losses of up to 84% (Norgaard, 1988). Similarly, introduced insects attack both native and introduced trees, often with staggering impact. Losses in eastern US forests to the European gypsy moth (*Lymantria dispar*) were estimated at $764 million in one year alone (US Congress Office of Technology Assessment, 1993). The Asian balsam woolly adelgid (*Adelges piceae*) has eliminated the Fraser fir (*Abies fraseri*) in many areas of the southern Appalachians, while the Asian hemlock woolly adelgid (*A. tsugae*) is currently devastating hemlock (*Tsuga*) forests through the eastern US, and the Asian emerald ash borer (*Agrilus planipennis*) threatens to eliminate extensive forests of ash (*Fraxinus*) in the US East and Midwest (Simberloff, 2012).

Herbivorous insects have been employed in biological control projects against many introduced aquatic and terrestrial weeds. Among striking successes are the use of the South American cactus moth, *Cactoblastis cactorum*, against pest prickly pear (*Opuntia* spp.) in Australia and the South American alligatorweed flea beetle (*Agasicles hygrophila*) on aquatic alligatorweed (*Alternanthera philoxeroides*) in Florida. In each instance, the weed had covered vast areas before the insect introduction, which subsequently limited it to small, ephemeral infestations (Simberloff, 2012). Such introductions of insects for control of introduced weeds have occasionally led to threats to nontarget native species. For example, the cactus moth, introduced to control pest prickly pear on the island of Nevis in the West Indies, reached the Florida Keys, where it attacked every individual of the native semaphore cactus, *Opuntia coricolla*, and it has now spread north from Florida to South Carolina and west to Louisiana; it is believed to threaten 79 species of *Opuntia* in the US and Mexico (Simberloff, 2012). The Eurasian weevil *Rhinocyllus conicus*, introduced to North America to control Eurasian pest thistles, especially musk thistle (*Carduus nutans*), attacks several native thistles, including the endangered Suisun thistle (*Cirsium hydrophilum* var. *hydrophilum*) (Simberloff, 2012). The key to these unplanned threats to native species, in each instance, is that the introduced insect feeds on several species in addition to the target pest plant and is able to maintain high numbers on alternative hosts (generally the target species), so that decline of the native species did not induce a decline in populations of the herbivore.

Parasitism and Disease

Introduced pathogens of various sorts have had enormous impacts. The most damaging to entire ecosystems are plant pathogens that affect dominant native plants; in so doing, they affect entire communities that interact with these plants. The Asian chestnut blight fungus (*Cryphonectria parasitica*) reached New York City in nursery stock in the late nineteenth century. In less than 50 years, it spread over about 100 million hectares of the eastern US, destroying the aerial parts of almost all mature chestnut (*Castanea dentata*) trees. As chestnut had

been the most common tree in many forests, the ecosystem impacts of this invasion were almost certainly enormous, although few crucial data were gathered. Several insect species host-specific on chestnut went extinct, and nutrient cycling rates were probably greatly modified (Freinkel, 2007).

A pathogen of animals that affected many entire ecosystems was the virus rinderpest, native to India. Introduced to sub-Saharan Africa in cattle in the 1890s, it infected many native ungulate species, with mortality in species such as the wildebeest (*Connochaetes* spp.) and buffalo (*Syncerus caffer*) surpassing 90%; the distribution of some species remains affected a century later. When ungulate populations plummeted, populations of carnivores such as the lion (*Panthera leo*) crashed as well. Because of the crucial role played by ungulates in aspects of vegetation structure and dynamics, the impact of rinderpest far surpassed the death of individuals it infected (Plowright, 1982). A long campaign to rid Africa from rinderpest finally succeeded in 2011 (McNeil Jr, 2011), but it is uncertain how long it will take for native ecosystems to recover, if they ever do.

Many introduced diseases have heavily affected individual species or small groups of them without apparent major impact on entire ecosystems. Ranges of Hawaiian native birds have been drastically circumscribed by habitat destruction, but avian malaria, caused by *Plasmodium relictum capristranoae* and vectored by introduced mosquitoes, afflicts the remaining populations and helps restrict them to upper elevations. The plasmodium was introduced with Asian songbirds (van Riper *et al*., 1986). The European fish parasite *Myxosoma cerebralis* causes whirling disease in salmonid fishes. Following World War II, live North American rainbow trout (*Oncorhynchus mykiss*) were transferred freely among European sites. There they acquired the parasite from the native brown trout, *S. trutta*, which has significant resistance. Eventually, frozen European rainbow trout were widely exported. The parasite probably reached North America in imported frozen fishes. Either these were fed accidentally to fish in a trout hatchery in Pennsylvania, or the viscera may have been discarded in streams near the hatchery. In any event, hatchery fishes became infected, and these were shipped to many other states. In large parts of Montana and Colorado, the great majority of rainbow trout contract the disease, and sport fisheries have collapsed in both states (Bartholomew and Reno, 2002).

Introduced macroparasites also often have major impacts. The parasitic plant witchweed (*Striga asiatica*) probably invaded North Carolina from Africa with military equipment after World War II. It attacks primarily grasses, including corn, and is a sufficient agricultural pest that it has been the target of a long eradication campaign and associated quarantine, which have led over 50 years to a great reduction in its introduced range (Eplee, 2001). The trematode *Cyathocotyle bushiensis*, which causes heavy mortality in ducks, has advanced along the St Lawrence River recently concurrently with its invasive introduced intermediate host, the Eurasian faucet snail *Bithynia tentaculata* (Sauer *et al*., 2007).

Many parasitic wasps and flies, as well as some disease pathogens, have been introduced as biological control agents for introduced pests. For example, the yellow clover aphid (*Therioaphis trifolii*) is controlled in California by three parasitic wasps, *Praon palitans*, *Trioxys utilis*, and *Aphelinus semiflavus* (van den Bosch *et al*., 1964). Populations of the cassava mealybug in Africa, discussed earlier, have been greatly reduced by the introduced South American parasitic wasp *Epidinocarsis lopezi*. Perhaps the best known biological control pathogen is the New World myxoma virus, introduced to mainland Europe (where the European rabbit is native) and Great Britain and Australia (where it is introduced). Initial epizootics caused over 99% mortality in Great Britain and Australia, and over 90% in France. The initial virulent strains in all three countries largely evolved to more benign strains, and, at least in Great Britain and Australia, rabbits evolved a degree of resistance. Thus, successive epizootics caused decreasing mortality, until an equilibrium appeared to have been reached (Bartrip, 2008). More recently, rabbit calicivirus from Asia was introduced to Australia and New Zealand, again intitiating epizootics that will almost certainly attenuate in time as the virus becomes more benign and rabbits more resistant. A number of pathogenic fungi have been introduced to control weeds, with varying degrees of success. For instance, African bridal creeper (*Asparagus asparagoides*) has been substantially controlled in parts of Australia by the African rust fungus *Puccinia myrsiphylli* (Morin and Edwards, 2006).

Hybridization and Introgression

Introduced species can eliminate native species by mating with them, a particularly strong threat when the native species is not as numerous as the introduced one. Both the New Zealand grey duck *(Anas superciliosa superciliosa*) and the Hawaiian duck (*A. wyvilliana*) are thus threatened by extensive hybridization and introgression with the North American mallard, (*A. platyrhynchos*), introduced as a game bird (Rhymer and Simberloff, 1996). Similarly, the white-headed duck (*Oxyura leucocephala*), now restricted in Europe to Spain, is threatened there by hybrization and introgression with North American ruddy ducks (*O. jamaicensis*), which were introduced to Great Britain as an amenity, escaped, and eventually reached Spain (Muñoz-Fuentes *et al*., 2007).

Both plants and animals are threatened by such introgression, and its extent is increasingly demonstrated by molecular techniques. This problem is much more common in regions that exchange closely related species, such as Europe and North America, than in those with species so distantly related that they are unlikely to be able to mate and exchange genes, such as Australia and either Europe or North America. Exchange of genes is not even necessary for hybridization with an introduced species to affect a native species inimically. Many females of the endangered European mink (*Mustela lutreola*) hybridize with male introduced American mink (*M. vison*), which become sexually mature earlier than the male European mink. The embryos are all aborted, but the loss of reproduction by the European mink exacerbates their population decline (Maran and Henttonen, 1995).

Indirect Effects

The foregoing impacts of introduced species are direct effects of various sorts. The actions of individuals of the introduced

species are directly on individuals (and often, ultimately, on populations) of one or more native species; they may attack them, eat them, poison them, infect them, and so forth. However, an introduced species can also affect other species indirectly in many ways.

Classical Indirect Effects

A strictly defined indirect effect occurs when one species affects the interaction between two others. For example, the chestnut blight led to the replacement of chestnut in much of the eastern US by oak species. Red oak (*Quercus rubra*) increased greatly, and this species is particularly susceptible to oak wilt disease (*Ceratocystis fagacearum*). The increase in oak wilt disease on red oak, in turn, raised the frequency of oak wilt disease on many less susceptible native oak species. Thus, the chestnut blight fungus indirectly affected the interaction of oaks and oak wilt disease, in addition to its direct effect on chestnut (Ellison *et al.*, 2005).

A common sort of indirect effect, called "apparent competition," occurs when the population of an introduced species becomes so large that it sustains a large population of a predator, parasite, or herbivore. If the predator or parasite does not greatly affect that population of the introduced species, it may nonetheless attack some native species and cause its population to decline. The introduced species is not actually competing with the native species, but the impact "appears" to be the same. A case in point is the Eurasian weevil *Rhinocyllus conicus*, discussed previously, which was introduced to North America to control invasive Eurasian musk thistle (*Carduus nutans*). Enough musk thistle remains that weevils remain numerous, and they attack several native North American thistle species, bringing one (the Suisun thistle, *Cirsium hydrophilum* var. *hydrophilum*) to the brink of extinction. Another common indirect effect is the "trophic cascade," in which an introduced species in a higher trophic level, by feeding on a species in the trophic level just below it, reduces the population of the latter sufficiently that the population of species in a still lower trophic level, previously controlled by the latter, increases. A major trophic cascade was precipitated by the introduction of the sea lamprey (*Petromyzon marinus*) to the North American Great Lakes in the nineteenth century. Lampreys devastated populations of all large native fishes in the lakes, and the declines of the latter rippled through the food web, causing increases in populations of several smaller fish species (Sorensen and Bergstedt, 2011). Successful biological control of plant-eating insects relies on trophic cascades. For instance, the introduced lady beetle that rescued the St Helena gumwood tree, discussed earlier, did so by eating the scale insect that was eating gumwood leaves.

Classical indirect effects can be complex and surprising. For example, two gall flies (*Urophora affinis* and *U. quadrifasciata*) introduced to control spotted knapweed (*Centaurea maculosa*) in western North America failed to do so. However, they have established enormous populations and have generated a substantial indirect effect by subsidized populations of the deer mouse *Peromyscus maniculatus*, comprising 85% of deer mouse winter diet in knapweed-infected areas and elevating deer mouse populations two to threefold. The full impact of increased populations of deer mice, a generalist feeder, are unknown but could include eating seeds of native plants, competition with other small mammals, predation on and competition with insects, and even increased transmission of the hantavirus to humans (Pearson and Callaway, 2003).

Chain Reactions

A number of the introductions discussed earlier affected one species directly, but this direct effect generated subsequent impacts in chain-like fashion. Thus, the direct effect of chestnut blight on American chestnut led indirectly to the extinction of several insect species host-specific on chestnut. The modification of nutrient cycles as a chestnut-dominated litter was replaced by litter dominated by leaves of oaks and other species almost certainly affected populations of litter inhabitants as well as plants, but these effects were never investigated. Rabbits in Australia directly affected certain plant species, as noted earlier, and also probably directly contributed to the elimination of two small burrowing marsupials, the boodie rat (*Bettongia lesueur*) and the lesser bilby (*Macrotis leucura*) by competition for burrows. Furthermore, rabbits indirectly led to the elimination of the common wombat (*Vombatus ursinus*) from part of its range by modifying succession and locally eliminating native perennial plants. On coastal islands, the erosion and vegetation changes caused by rabbits have been highly detrimental to seabirds that use these islands for resting and breeding.

Some chain reactions induced by introduced species are so complicated that predicting them would have been difficult. Caterpillars of the large blue butterfly (*Maculina arion*) in Great Britain required development in underground nests of the ant *Myrmica sabuleti*. The ant cannot nest in overgrown areas. Changing land use patterns and reduced livestock grazing had left introduced rabbits as the main grazer maintaining the habitat. The biological control introduction of myxoma virus, discussed previously, reduced rabbit populations sufficiently that ant populations declined, and the butterfly was extirpated (Ratcliffe, 1979).

Chain reactions can generate effects far from the site of the initial introduction. For example, landlocked kokanee salmon (*Oncorhyncus nerka*) were introduced to Flathead Lake, in western Montana, in 1916 and largely replaced native cutthroat trout (*O. clarki*) as the dominant sportfish. By 1931, the kokanee were spawning in McDonald Creek (Glacier National Park), some 100 km upstream from Flathead Lake, and the spawning population was so large that it soon attracted large populations of bald eagles, grizzly bears, and many other predators. The catch of kokanee rose to over 100,000 fishes per year through 1985. Between 1968 and 1975, the opossum shrimp (*Mysis relicta*), native to several large deep lakes in North America and Sweden, was introduced to three lakes in the upper portion of the Flathead catchment in a misguided attempt to enhance kokanee productivity. The shrimp drifted downstream, reaching Flathead Lake by 1981. There was an immediate, dramatic decline in densities of copepods and especially cladocerans preyed on by the shrimp. This competition with the shrimp for prey caused the kokanee population to decline rapidly; there was no catch at all in 1988 and

1989. Bald eagle numbers fell precipitously, as did those of grizzly bears and several other predators (Spencer *et al*., 1991).

Vectoring of Pathogens

In a number of instances discussed earlier, an introduced species carried a pathogen that greatly affected one or more native species – rinderpest, whirling disease, avian malaria. In some cases, these could be classified as classical indirect effects. For instance, in Hawaii the avian malaria plasmodium changed the interaction between introduced mosquitoes and native forest birds, to the detriment of the latter. The impact of an introduced pathogen can be more complex – for example, the introduced species that carries it to a new region need never be found together with a native species ultimately affected. Consider the introduction of the Chinese grass carp to the US. The carp was introduced to Arkansas in 1968; it spread to the Mississippi River, carrying with it a parasitic tapeworm from Asia, *Bothriocephalus acheilognathi*. The tapeworm quickly infested other fishes, including the red shiner (*Notropis lutrensis*), a popular bait fish. Fishermen or bait dealers introduced infected red shiners to the Colorado River, and by 1984 they had reached the Virgin River, a Utah tributary. There they infected the woundfin (*Plagopterus argentissimus*), a native minnow already threatened by dams and water diversions. Tapeworm infections caused woundfin numbers to decline precipitously, possibly because they are less able to compete with the red shiner, which have some resistance to the tapeworm (Moyle, 1993).

Invasional Meltdown

Some ecological theories suggest that introduced species should interfere with one another and thereby lessen one another's impact. The theory of island biogeography, for example, implies that each successive species in a series introduced to an island (or a habitat island) has a lower probability of surviving. Biological control practitioners have argued about whether releasing several species of potential natural enemies of a single pest might lessen the overall impact on that pest because of competition among them. However, positive interactions between introduced species are often detected. Sometimes one introduced species facilitates the existence and enhances the impact of another; other times the combined impacts of two or more introduced species exceed the sum of what the same species might have accomplished individually. Such situations, in which different introduced species enhance one another's impacts, are collectively termed "invasional meltdown." The existence of many such cases suggests that an accelerating wave of invasions with ever-increasing impacts may characterize the future in many regions (Simberloff and Von Holle, 1999).

Introduced Animals Pollinating and Dispersing Introduced Plants

Introduced fig (*Ficus*) species are frequently planted as ornamentals in south Florida. If the wasp species that pollinates a fig species in its native range is absent, the fig cannot reproduce. Thus, figs in Florida long remained where planted. However, within the past 25 years, breeding populations of three host-specific pollinating wasps were introduced to Florida, and the species they pollinate now regularly produce seeds. *Ficus microcarpa* is spreading most quickly and has become an invasive weed, the small fruits of which are dispersed by birds and ants. The spread of the same complex of *F. microcarpa* trees and its wasp pollinator *Parapristina verticillata* is also occurring in Bermuda, Mexico, and Central America.

Some introduced animals disperse introduced plants that disrupt native plant communities. The red-whiskered bulbul (*Pycnonotus jocosus*) has dispersed a number of alien species, such as *Rubus alceifolius*, *Cordia interrupta*, and *Ligustrum robustum* on the Indian Ocean island of La Réunion. *Cordia* was not viewed as a problem until it was widely distributed by the bulbul; the bird is the primary dispersal agent for *Ligustrum*. In the Hawaiian Islands, introduced pigs selectively eat and thus disperse several invasive introduced plant species, and their rooting and defecation favor several introduced invertebrates. Further, the pigs grow larger because of the presence of introduced, protein-rich European earthworms in the soil. The Asian myna bird (*Acridotheres tristis*), introduced to control pasture insects, has dispersed the New World weed *Lantana camara* widely in the lowlands of the Hawaiian islands, including into native forest areas.

Introduced Plants that Modify Habitat

As noted in a previous section, the Atlantic nitrogen-fixing shrub *Morella faya* in Hawaii may facilitate invasion by other introduced plants that are currently limited by the nitrogen-poor volcanic soil and absence of native nitrogen-fixers. This plant also enhances populations of introduced earthworms, which in turn increase the rate of nitrogen burial and thus exacerbate the impact of the plant on nitrogen cycling. The soil drying by introduced salt cedars already described favors the introduction of non-native grasses in both the US Southwest and Australia. The increased soil salinity produced by the African crystalline ice plant, described earlier, does more than kill native plants. When wind or other disturbances create holes in the carpet of ice plant, they are colonized not by native plants but by the ice plant itself or by weedy introduced plants such as *Malva parviflora* or *Erodium cicutarium*.

Many invasive introduced plants increase fire frequency or intensity to the detriment of native species but to their own advantage and that of other invaders. Old World grasses have come to dominate many New World grasslands through this facilitation. For instance, in Hawaii introduced perennial tufted beardgrass, *Schizachyrium condensatum*, invaded seasonal submontane shrub-dominated woodland. It fostered much more frequent fires over larger areas, killing most native trees and shrubs. But *S. condensatum* recovered quickly and the even more flammable introduced perennial molasses grass, *Melinis minutiflora*, also invaded (D'Antonio and Vitousek, 1992).

Introduced Animals that Modify Habitat

The disturbance wrought by large, congregating, introduced herbivores in North America favored the establishment of

Eurasian grasses adapted to such animals and helped them to replace native tussock grasses. Similarly, the Asian water buffalo (*Bubalus bubalis*), introduced to northeastern Australia as a beast of burden and for meat, spread throughout the flood plain of the Adelaide River by the late nineteenth century. It devastated native plant communities, compacted the soil, eroded creek banks, and altered hydrology. A Central American shrub, *Mimosa pigra*, had been a minor, noninvasive weed in the vicinity of Darwin for a century. This legume produces large numbers of small seeds that are readily dispersed by water. The sundering of the flood plains by the water buffalo created ideal germination habitat for *M. pigra* seedlings. In many areas, native sedgelands were converted to a monoculture of *M. pigra*.

The Caspian zebra mussel, *Dreissena polymorpha*, arrived in the Great Lakes in the 1980s. Its huge numbers and great filtering ability have greatly affected native freshwater communities over much of the eastern and midwestern US. In particular, they convert large amounts of seston into excreted feces, creating a soft substrate of rich organic material. Among species whose populations are increased by the zebra mussel is the invasive Eurasian faucet snail discussed previously. The filtering also increases water clarity, and this change has favored certain invasive macrophytes, such as Eurasian watermilfoil (*Myriophyllum spicatum*).

Mutualisms

A number of instances in which one introduced species facilitates the existence, spread, and population growth of another introduced species constitute mutualisms, in that the latter species also aids the former. For example, introduced macrophytes like Eurasian watermilfoil provide additional settling substrate for the zebra mussel, and they can also help the mussel disperse between water bodies. *Mimosa pigra* aided water buffalo by protecting them from aerial hunters trying to eradicate them. Fig wasps and their associated fig species obviously form coevolved mutualistic introduced species pairs. In the Hawaiian islands, introduced African big-headed ants (*Pheidole megacephala*) tend an introduced scale insect, *Coccus viridis*, a classic mutualism. The scale is on the introduced plant *Pluchea indica*; among other impacts, the ant hinders introduced predatory lady beetles and parasitic wasps. In this one case, a range of interactions are taking place among an entire assemblage of introduced species. Sometimes, as with *Mimosa pigra* and water buffalo, the mutualists cannot have coevolved, as they originate on different contintents. The African big-headed ant in Hawaii also tends the South American gray pineapple mealybug (*Dysmicoccus neobrevipes*), a particular pest because it helps spread a wilt disease of pineapple.

Time Lags and Evolution

Impacts of introduced species may change dramatically, as, for example, when an invader is quiescent during a time lag, after which it rather abruptly spreads and increases in number. As observed in the Introduction section, such time lags complicate efforts to predict the impacts of an introduced species and are sometimes mysterious. It is frequently suggested that a mutation in an invader could account for a particular time lag, and that evolution in general could greatly change the impacts of an introduced species. In no instance can a sudden increase in population size and impact of an introduced species be clearly attributed to a mutation, but there is ample evidence that such species evolve in their new homes. How frequently and to what extent does such evolution change their impacts?

Morphological evolution of introduced species is often apparent. The small Indian mongoose, whose depredations of native island species were discussed earlier, has become larger and more sexually dimorphic on all islands to which it has been introduced. This change is probably selected for in the absence on the islands of the slightly larger gray mongoose (*Herpestes edwardsii*), with which the small Indian mongoose coexists in the region of origin for the island populations (Simberloff *et al.*, 2000). The North American muskrat (*Ondatra zibethicus*), introduced as a fur bearer to Europe in 1905, has spread throughout much of Europe in less than a century. The muskrat is well known in its native range as a species with substantial variation in morphology (especially size), to the extent that many subspecies have been named. In Europe, it has already evolved approximately the same degree of variation (Pankakoski, 1983). Similarly, the European house sparrow (*Passer domesticus*) was introduced in New York City in 1853 and quickly spread to become one of the most common birds in North America, covering much of the continent. It has evolved so much that distinct "races" are now easily identified in different parts of its introduced range (Johnston and Selander, 1971). Many plant species are known to differ substantially between their native and introduced ranges, though there is generally far less evidence than in the vertebrate cases just cited that the differences are genetic. What is rarely known in these demonstrated instances of evolution after introduction is how these changes affect the impacts of these species on native communities and ecosystems.

In agroecosystems, however, good evidence of increased impacts generated by gradual evolution within introduced species is found among several insect pests of agriculture. For instance, the southern cowpea weevil, *Callosobruchus maculatus*, is native to sub-Saharan west Africa but has become a worldwide pest of legume seeds. The original host was the wild progenitor of black-eyed peas in Africa, but the weevil has now evolved biotypes that thrive on at least 12 species of crop beans, with increased survival on the new host plants (Wasserman, 1986). The western corn rootworm, *Diabrotica virgifera*, is native to Central America and invaded North America by the early twentieth century. In response, farmers adopted the tactic of rotating corn with soy. However, recently a biotype has evolved that prefers to oviposit and feed on soy, and genotypes have also evolved with extended periods of diapause, so they can survive the rotation period with soy and attack the next corn crop (Levine *et al.*, 2002).

The hybridization of an introduced species with a native one can bring in new genes that can increase impact and can even lead to the formation of a new, more invasive species. For example, North American cordgrass (*Spartina alterniflora*), introduced in shipping ballast to southern England, occasionally hybridized with a noninvasive native congener, *S. maritima*. These hybrids were sterile, but eventually one hybrid

individual underwent a doubling of chromosome number to produce a fertile new species, *S. anglica*, which turned out to be highly invasive (Thompson, 1991).

The coevolution of the introduced rabbit and myxoma virus in Australia and Great Britain certainly affected the impact of the rabbit on native systems. The fact that the rabbit became somewhat more resistant to the virus, and that the virus evolved to be somewhat more benign, lowered the rabbit mortality in each successive epizootic, thus the degree of various impacts outlined earlier. Because pathogens and their hosts often coevolve, changes of impact might be expected. They are not always seen, however. Chestnut blight has not yet become perceptibly less devastating to American chestnut, and only one native Hawaiian bird has evolved any traits conferring resistance to avian malaria.

Many introduced pest insects have evolved resistance to chemical pesticides and thus generated much greater impacts, generally on agricultural or silvicultural plants. Just as strains of human pathogens have evolved resistance to whole sequences of antibiotics, some pest insects have evolved resistance to several insecticides. This resistance can arise in three ways: insects can evolve to tolerate greater amounts of the chemical, they can evolve physiological means of detoxifying the chemical, or they can evolve behavioral traits (such as going to the bottom of a leaf instead of the top) that help to avoid contact with the chemical. Introduced weedy plants also evolve resistance to herbicides, though this phenomenon has not been as widely studied as insect resistance.

The evolution of resistance to chemicals by introduced pests gives impetus to the desire to control them biologically, through the introduction of natural enemies. However, this approach has fostered concerns about subsequent evolution. For pathogens, one concern is that the target organisms will evolve resistance to them. Much of the market for insect pathogens today consists of products involving two organisms, the bacterium *Bacillus thuringiensis* and heterorhabditid nematodes. Both types have no natural association with insect pests on plants and thus must be routinely reapplied. It is possible that target pest insects could become resistant to them, simply through repeated exposure plus selection of resistant strains. This possibility is heightened for the bacterium by the fact that *B. thuringiensis* toxins are now being genetically engineered into crop plants, thus increasing exposure of pest insects to them and increasing the rate of natural selection for resistance. However, introduced biocontrol insects can evolve greater adaptation to a potential host and thus become more rather than less lethal. In the US, the Egyptian alfalfa weevil, *Hypera brunneipennis*, was originally quite immune to the European ichneumonid parasitic wasp *Bathyplectes curculionis*, encapsulating 35–40% of the eggs and larvae in an immune response. Fifteen years later, only 5% of the eggs were encapsulated (Messenger and van den Bosch, 1971).

A second concern is that introduced biological control agents could evolve to switch hosts. For introduced insects, the fact that host range can be controlled by a single gene only heightens this concern, as does the fact that certain features of most biological control introductions favor fast evolution. These features include initially small population size, rapid population growth in new environments, and novel, abundant resources. In light of the possibility, it might seem surprising that few examples of host-switching are known. Because many of the pathogens used or contemplated as biological control agents have broader host ranges than parasitic insects, the probability of host-switching accompanied or followed by evolutionary adaptation to the new host is enhanced.

The introduction of new individuals of a native or previously introduced species can bring new genes that enhance the impact of that species, sometimes dramatically. For instance, common reed (*Phragmites australis*) has been present in the US for thousands of years. In the last 150 years it suddenly became highly invasive in wetlands, with its geographic range and abundance greatly increasing. This expansion arose from the arrival, probably in soil ballast of ships, of a non-native genotype that has both displaced native genotypes and spread to areas not formerly occupied by the species (Saltonstall, 2002). Reed canary grass (*Phalaris arundinacea*) similarly spread rapidly in North America after multiple introductions from Europe bearing new genotypes (Lavergne and Molofsky, 2007). The introduced European green crab (*Carcinus maenas*) spread further north along the eastern coast of North America after new genotypes arrived in the 1980s, probably in ballast water, and proved more cold-tolerant than the previously introduced ones (Roman, 2006).

Quantifying Impact

"Impact" and "effect" have never been formally defined in invasion biology, so that rankings of introduced species in terms of greatest impact are impressionistic. Area of occupancy is a frequently used index of effect, although the number of individual invaders per unit area and what each individual actually does would be important (though more difficult to measure). Impacts on ecosystem processes, such as nutrient cycling and fire frequency, are seen by some biologists as the best measures of effect of introduced species, on the grounds that such alternation of processes could affect a large number of species simultaneously. Both areal coverage and impact on processes often envision outcome for native species, communities, and ecosystems as the crucial measure of impact. The real impact, in these terms, would require accurate knowledge of the range size of the species, the distribution of abundances of species within this range, and the impact on the native species, both on-site and elsewhere, per individual invader (or per unit biomass of the invader) (Parker *et al.*, 1999). The first two factors can be measured or at least estimated in straightforward fashion, although the effort for some species would be enormous. The per capita impact, however, would probably entail not only detailed observation but also experiment, perhaps long-term, at individual, population, community, and ecosystem levels. As one example of the subtleties that might be involved, the fact that an introduced predator is observed to eat individuals of a native species does not necessarily mean that this predation affects the population size, behavior, or ecology of the prey species. Ecological methods are well enough developed that these sorts of questions can be answered, but they generally require detailed research, not quick observation.

Economic costs of an invasion are often proposed as estimates of the impact of introduced species, but these costs are more easily tallied from a human standpoint than from that of native species, communities, and ecosystems. It is a fairly straightforward matter to tabulate the costs expended in controlling a particular weed or insect pest, in terms of materiel and personnel used. And one could relatively easily add the monetary cost of many services or goods lost because of an introduced species – say, a percentage of some crop. Costs of other lost services are not so easily tabulated, even in human terms. For example, how expensive is the loss in enjoyment suffered by citizens who can no longer walk through Fraser fir or American chestnut forests in the eastern US because these were eliminated by introduced species? Economists have attempted to measure these costs. For example, travel-costs of citizens to a particular natural site, as evidence of how much they are willing to pay for such an experience, are sometimes used to estimate the "value" of a site, at least to humans. But all of these methods are controversial. And costs to nonhuman species are even harder to weigh in monetary terms. What was the cost, in dollars, to the insect species extinguished because of the chestnut blight?

In sum, even though a bewildering variety of impacts of introduced species can be demonstrated, measuring and comparing these impacts for different species remains as much a conceptual and practical challenge as predicting them does.

See also: Coevolution. Fires, Ecological Effects of. Grazing, Effects of. Human Impact on Biodiversity, Overview. Introduced Plants, Negative Effects of. Island Biogeography. Species Coexistence

References

Bain MB (1993) Assessing impacts of introduced grass carp in large systems. *Environmental Management* 17: 211–224.

Bartholomew JL and Reno PW (2002) The history and dissemination of whirling disease. *American Fisheries Society Symposium* 29: 3–24.

Bartrip PWJ (2008) *Myxomatosis: A History of Pest Control and the Rabbit*. London: Macmillan.

Biger G and Liphschitz N (1995) Australian trees in the land of Israel. *Journal of Israeli History* 16: 235–244.

Booth RG, Cross AE, Fowler SV, and Shaw RH (2001) Case study 5.24. Biological control of an insect to save an endemic tree on St. Helena. In: Wittenberg R and Cock MJW (eds.) *Invasive Alien Species: A Toolkit of Best Prevention and Management Practices*, p. 162. Wallingford, UK: CAB International.

Chanin PRF and Jefferies DJ (1978) The decline of the otter *Lutra lutra* L. in Britain: An analysis of hunting records and discussion of causes. *Biological Journal of the Linnean Society* 10: 305–328.

Cowie RH (2001) Invertebrate invasions on Pacific islands and the replacement of unique native faunas: A synthesis of the land and freshwater snails. *Biological Invasions* 3: 119–136.

Cronk QCB (1989) The past and present flora of St. Helena. *Journal of Biogeography* 16: 47–64.

D'Antonio CM and Vitousek PM (1992) Biological invasions by exotic grasses, the grass/fire cycle, and global change. *Annual Review of Ecology and Systematics* 23: 63–87.

Ellison AM, Bank MS, Clinton BD, *et al.* (2005) Loss of foundation species: Consequences for the structure and dynamics of forested ecosystems. *Frontiers in Ecology and the Environment* 9: 479–486.

Eplee RE (2001) Case study 2.10. Co-ordination of witchweed eradication in the U.S.A. In: Wittenberg R and Cock MJW (eds.) *Invasive Alien Species: A Toolkit of Best Prevention and Management Practices*, p.36. Wallingford, UK: CAB International.

Ewel JJ (1986) Invasibility: Lessons from south Florida. In: Mooney HA and Drake JA (eds.) *Ecology of Biological Invasions of North America and Hawaii*, pp. 214–230. New York: Springer.

Freinkel S (2007) *American Chestnut The Life, Death, and Rebirth of a Perfect Tree*. Berkeley, CA: University of California Press.

Groombridge B (1992) *Global Biodiversity Status of the Earth's Living Resources*. London: Chapman & Hall.

Gurnell J, Wauters LA, Lurz PWW, and Tosi G (2004) Alien species and interspecific competition: Effects of introduced eastern grey squirrels on red squirrel population dynamics. *Journal of Animal Ecology* 73: 26–35.

Hays W (2011) Small Indian mongoose. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 631–634. Berkeley: University of California Press.

Hill M, Coetzee J, Julien M, and Center T (2011) Water hyacinth. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 689–692. Berkeley, CA: University of California Press.

Johnston RF and Selander RK (1971) Evolution in the house sparrow. II. Adaptive differentiation in North American populations. *Evolution* 25: 1–28.

Lavergne S and Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences (USA)* 104: 3883–3888.

Leader-Williams N (1988) *Reindeer on South Georgia*. Cambridge, UK: Cambridge University Press.

Lever C (1992) *They Dined on Eland The Story of the Acclimatisation Societies*. London, UK: Quiller.

Levine E, Spencer JL, Isaid SA, Ornstad DW, and Gray ME (2002) Adaptation of the western corn rootworm to crop rotation: Evolution of a new strain in response to a management practice. *American Entomologist* 48: 94–107.

Lindroth CH (1957) *The Faunal Connections between Europe and North America*. New York: Wiley.

Mack RN (1986) Alien plant invasion to the Intermountain West: A case history. In: Mooney HA and Drake JA (eds.) *Ecology of Biological Invasions of North America and Hawaii*, pp. 191–210. New York: Springer.

Mack RN (2011) Cheatgrass. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 108–112. Berkeley, CA: University of California Press.

Maran T and Henttonen H (1995) Why is the European mink (*Mustela lutreola*) disappearing? A review of the process and hypotheses. *Annales Zoologici Fennici* 32: 47–54.

McDaniel KC, DiTomaso JM, and Duncan CA (2005) Tamarisk or saltcedar, *Tamarix* spp. In: Clark JK and Duncan CL (eds.) *Assessing the Economic, Environmental and Societal Losses from Invasive Plants on Rangeland and Wildlands*, pp. 198–222. Champaign, Illinois: Weed Science Society of America.

McNeil Jr. DG (2011) Rinderpest, scourge of cattle, is vanquished. *New York Times*, June 28: D1.

Mead AR (1961) *The Giant African Snail. A Problem in Economic Malacology*. Chicago: University of Chicago Press.

Meinesz A (2001) *Killer Algae. The True Tale of a Biological Invasion*. Chicago: University of Chicago Press.

Messenger P and van den Bosch R (1971) The adaptability of introduced biological control agents. In: Huffaker CB (ed.) *Biological Control*, pp. 68–92. New York: Plenum.

Mills EL, Leach JH, Carlton JT, and Secor CL (1993) Exotic species in the Great Lakes: A history of biotic crises and anthropogenic introductions. *Journal of Great Lakes Research* 19: 1–54.

Morin L and Edwards PB (2006) Selection of biological control agents for bridal creeper: A retrospective review. *Australian Journal of Entomology* 45: 287–291.

Moyle PB (1993) *Fish: An Enthusiast's Guide*. Berkeley, CA: University of California Press.

Muñoz-Fuentes V, Vilá C, Green AJ, Negro J, and Sorenson MD (2007) Hybridization between white-headed ducks and introduced ruddy ducks in Spain. *Molecular Ecology* 16: 629–638.

Norgaard RB (1988) The biological control of cassava mealybug in Africa. *American Journal of Agricultural Economics* 70: 366–371.

Pankakoski E (1983) Morphological variation and population structure of Finnish muskrats, *Ondatra zibethica* (L.). *Annales Zoologici Fennici* 20: 207–222.

Parker IM, Simberloff D, Lonsdale WM, *et al.* (1999) Impact: Toward a framework for understanding the ecological effect of invaders. *Biological Invasions* 1: 3–19.

Pascal M (2011) Rats. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 571–575. Berkeley, CA: University of California Press.

Pearson DE and Callaway RM (2003) Indirect effects of host-specifc biological control agents. *Trends in Ecology and Evolution* 18: 456–461.
Perry G and Rodda GH (2011) Brown treesnake. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 78–80. Berkeley, CA: University of California Press.
Petren K and Case TJ (1998) Habitat structure determines competition intensity and invasion success in gecko lizards. *Proceedings of the National Academy of Sciences (USA)* 95: 11739–11744.
Pimentel D, McNair S, Janecka J, *et al.* (2002) Economic and environmental threats of alien plant, animal, and microbe invasions. In: Pimentel D (ed.) *Biological Invasions. Economic and Environmental Costs of Alien Plant, Animal, and Microbe Species*, pp. 307–330. Boca Raton, Florida: CRC Press.
Plowright W (1982) The effects of rinderpest and rinderpest control on wildlife in Africa. *Symposia of the Zoological Society of London* 50: 1–28.
Pringle RM (2011) Nile perch. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 484–488. Berkeley, CA: University of California Press.
Pyke GH (2008) Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia* species. *Annual Review of Ecology, Evolution, and Systematics* 39: 171–191.
Ratcliffe D (1979) The end of the large blue butterfly. *New Scientist* 8: 457–458.
Rhymer J and Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* 27: 83–109.
Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society of London B* 273: 2453–2459.
Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences (USA)* 99: 2445–2449.
Sanders NJ (2011) Ants. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 17–24. Berkeley, CA: University of California Press.
Sauer JS, Cole RA, and Nissen JM (2007) *Finding the exotic faucet snail (Bithynia tentaculata):* Investigation of waterbird die-offs on the upper Mississippi River National Wildlife and Fish Refuge. *US Geological Survey Open-File Report 2007–1065.* Washington, DC: US Geological Survey.
Schmitz DC, Simberloff D, Hofstetter RH, Haller W, and Sutton D (1997) The ecological impact of nonindigenous plants. In: Simberloff D, Schmitz DC, and Brown TC (eds.) *Strangers in Paradise. Impact and Management of Nonindigenous Species in Florida*, pp. 39–61. Washington, D.C: Island Press.
Serbesoff-King K (2003) *Melaleuca* in Florida: A literature review on the taxonomy, distribution, biology, ecology, economic importance and control measures. *Journal of Aquatic Plant Management* 41: 98–112.
Shiganova T (2011) Ponto-Caspian invasions. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 549–557. Berkeley, CA: University of California Press.
Simberloff D (1986) Introduced insects: A biogeographic and systematic perspective. In: Mooney HA and Drake JA (eds.) *Ecology of Biological Invasions of North America and Hawaii*, pp. 3–26. New York: Springer.
Simberloff D (2011) How common are invasion-induced ecosystem impacts? *Biological Invasions* 13: 1255–1268.
Simberloff D (2012) Risks of biological control for conservation purposes. *BioControl*. doi: 10.1007/s10526-011-9392-4.
Simberloff D, Dayan T, Jones C, and Ogura G (2000) Character displacement and release in the small Indian mongoose, *Herpestes javanicus*. *Ecology* 81: 2086–2099.
Simberloff D, Schmitz DC, and Brown TC (eds.) (1997) *Strangers in Paradise. Impact and Management of Nonindigenous Species in Florida*. Washington, DC: Island Press.
Simberloff D and Von Holle B (1999) Positive interactions of nonindigenous species: Invasional meltdown? *Biological Invasions* 1: 21–32.
Singer FJ, Swank WT, and EEC. Clebsch (1984) Effects of wild pig rooting in a deciduous forest. *Journal of Wildlife Management* 48: 464–473.
Sorensen PW and Bergstedt RA (2011) Sea lamprey. In: Simberloff D and Rejmánek M (eds.) *Encyclopedia of Biological Invasions*, pp. 619–623. Berkeley, CA: University of California Press.
Spencer CN, McClelland BR, and Stanford JA (1991) Shrimp stocking, salmon collapse, and eagle displacement. *BioScience* 41: 14–21.
Staples GW and Cowie RH (eds.) (2001) *Hawaii's Invasive Species*. Honolulu: Mutual Publishing.
Thompson HV and King CM (eds.) (1994) *The European Rabbit. The History and Biology of a Successful Colonizer*. Oxford, UK: Oxford University Press.
Thompson JD (1991) The biology of an invasive plant: What makes *Spartina anglica* so successful? *BioScience* 41: 393–401.
US Congress Office of Technology Assessment (1993) *Harmful Non-Indigenous Species in the United States*. Washington, DC: US Government Printing Office.
van den Bosch R, Schlinger EI, Dietrick EJ, Hall JC, and Puttler B (1964) Studies on succession, distribution, and phenology of imported parasites of *Therioaphis trifolii* (Monell) in southern California. *Ecology* 45: 602–621.
van Riper C, van Riper SG, Goff ML, and Laird M (1986) The epizootiology and ecological significance of malaria in Hawaiian land birds. *Ecological Monographs* 56: 327–344.
Vitousek PM (1986) Biological invasions and ecosystem properties: Can species make a difference? In: Mooney HA and Drake JA (eds.) *Ecology of Biological Invasions of North America and Hawaii*, pp. 163–176. New York: Springer.
Vivrette NJ and Muller CH (1977) Mechanism of invasion and dominance of coastal grassland by *Mesembryanthemum crystallinum*. *Ecological Monographs* 47: 301–318.
Volin JC, Lott MS, Muss JD, and Owen D (2004) Predicting rapid invasion of the Florida Everglades by Old World climbing fern (*Lygodium microphyllum*). *Diversity and Distributions* 10: 439–446.
Wasserman SS (1986) Genetic variation in adaptation to foodplants among populations of the southern cowpea weevil, *Callosobruchus maculatus*: Evolution of oviposition preference. *Ecologia Experimentalis et Applicata* 42: 201–212.
Winter L (1999) Cats indoors! *Earth Island Journal* (summer) 14(2): 25–26.

Invertebrates, Freshwater, Overview

Margaret A Palmer, University of Maryland, College Park, MD, USA
Holly L Menninger, Cornell University, Ithaca, NY, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Margaret A. Palmer, P. Sam Lake, volume 3, pp 531–542, © 2001, Elsevier Inc.

Glossary

Benthic At or on the bottom of streams, lakes, or riverbeds.
Decomposition The breakdown of organic matter such as dead plants and algae to release carbon and nutrients.
Endemic species A species that is unique to a particular locality.
Exotic species A species that has been introduced to an area outside of its native range.
Groundwater A "reservoir" of water residing below ground in saturated soils and beneath geologic formations.
Larvae Early stages in the development of an organism; for invertebrates, often morphologically quite distinct from the adult.
Local species richness Number of species found at a local site; distinguished from regional or global species richness that "sums" the number of species across a number of individual sites.
Planktonic Organisms that reside in the water column.
Watershed A geographical region in which water drains into common water bodies.

Introduction

Water is the most abundant substance on earth and is essential for all of life. It is critical to the world's climate, to the cycling of nutrients, and is a habitat for much of the earth's biodiversity. Fresh water links the land and oceans *via* groundwater and riverine flow. While invisible to the human eye, groundwaters represent a substantial portion of the fresh waters (approximately 30%) acting not only as the most important reservoir of fresh water on the earth, but also as a home to many unique fauna. Lakes and reservoirs are also important stores of fresh water and harbor several groups of extremely abundant invertebrates. Running waters and wetlands are known to be havens for diverse assemblages of invertebrates, many of whom are important in ecological processes that ensure clean water. The amount and timing of water inputs to many freshwater systems worldwide is changing rapidly with growing human demands for freshwater resources, exacerbated by climate change. Since many invertebrates require specific habitats and flood regimes, their abundance and diversity in many parts of the world are at risk because habitats are being lost and natural flow regimes altered at astounding rates. Also contributing to the recent declines in freshwater invertebrates are, poor water quality and the introduction of exotic species that have led to the extinction of native species, especially bivalves and crayfish.

Major Freshwater Habitats

The major freshwater habitats include running waters (streams and rivers), standing or semistanding waters (reservoirs, lakes, ponds, wetlands), and groundwater (**Figure 1**). Most of the

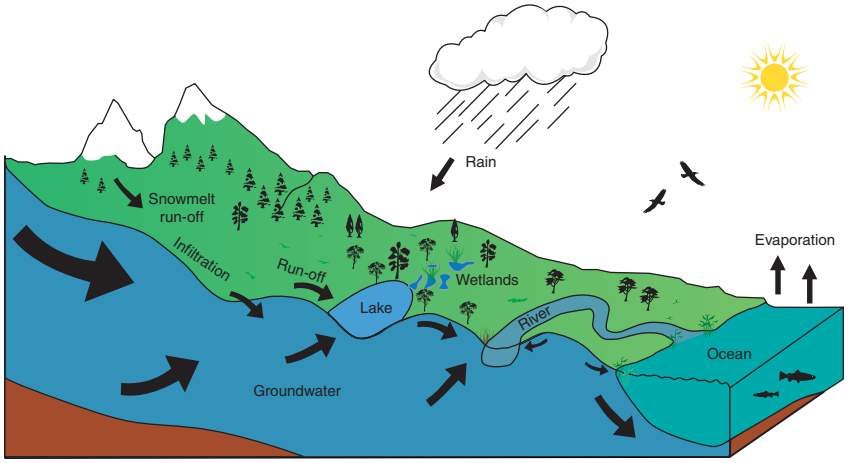


Figure 1 Schematic illustration of major freshwater habitats. Note all water bodies, including lakes, streams and rivers, wetlands, and the oceans, are connected *via* the groundwaters.

35 million cubic kilometers of fresh water on the earth (≈69%) exists in the form of polar ice. Of the remaining fresh water, a huge volume exists out of sight as groundwater (**Table 1**). Lakes, streams, and wetlands represent less total water volume but are biologically rich hot spots for abundant and diverse assemblages of invertebrates. Together, the freshwater bodies form a continuous "pipeline" that links the atmosphere, land, and the oceans. Rainwater is intercepted by vegetation on land, percolates through the soils into groundwaters, and then moves into wetlands, lakes, and streams before it eventually reaches the oceans (**Figure 1**). Since groundwater directly or indirectly "feeds" all other water bodies on earth, maintenance of adequate quantities of high-quality water in this underground reservoir is a high priority worldwide.

Major Groups and Their Lifestyles

A disproportionate number of species live in fresh waters compared to marine systems: oceans comprise 70% of the earth's surface while inland waters make up about 1%. Despite this, about 10% of all animal species alive today live in fresh water. The majority of the freshwater animal species are invertebrates. More than 107,000 freshwater invertebrate species have been described but many more remain undiscovered – new species and genera are found every year (Balian *et al.*, 2008). Freshwater habitats have representatives from most taxonomic groups with the insects being particularly speciose (**Table 2**). The freshwater insects are also quite notable for the large number whose lifestyle changes with age: many have aquatic larval stages but flying adults. This is in contrast to freshwater mollusks and crustaceans that complete their life cycles in water either as members of the plankton or benthos.

Life cycles of freshwater invertebrates can be quite complex with multiple, morphologically distinct stages (Thorp and Covich, 2010) (**Figure 2**). For example, among the freshwater bivalves, some are hermaphroditic while others have separate sexes; some brood their eggs and then release small clams; however, many release larvae into the water. Zebra mussels produce a planktonic larval stage (a veliger larva), while some of the most common freshwater native bivalves have unique parasitic larvae (glochidia larvae) that attach to the skin, scales, or gills of fish. The host fish disperses the attached glochidium larvae, prior to the detachment of larvae and their metamorphosis into adult bivalves. This dispersal phase may be very important to the long-term survival of bivalves particularly in freshwater bodies that are being stressed.

Damselflies and dragonflies provide other examples of complex life cycles typical of many freshwater invertebrates. These insects have flying adults that live weeks to months and lay eggs on aquatic vegetation or other substrates. The eggs hatch within 12–30 days, or they can remain viable in an unhatched state (diapause) if there is a drought. Small larval stages (sometimes called nymphs) hatch from the eggs and then molt multiple times (10–16) before crawling up vegetation and "emerging" as flying adults. The presence of an intact riparian zone with a diversity of native plants promotes population persistence of insects.

Number of Freshwater Invertebrate Species

The number of freshwater invertebrates varies considerably between lakes, running waters, groundwaters, and wetlands. Since the number of species varies vastly from locale to locale as a function of many complex factors, estimates of species richness should be taken as only approximations. Additionally, freshwater invertebrate diversity at any one time may be moderate while accumulated diversity over time may be very high. For example, there may be dramatic declines in invertebrate diversity in many streams during the flood season but high levels of diversity across the entire year. Similarly, temporary wetlands may have much higher total diversity over time than permanent wetlands even though the snapshot diversity in a temporary wetland may be considerably lower.

Many freshwater wetlands are extremely speciose habitats (e.g., up to 2000 invertebrate species at a locale), followed by lakes and streams (80–1400 species of invertebrates typically found at a locale), and then groundwaters (0–150 at a locale). In all these habitats, insects are by far the most speciose group with local species richness levels of 50–500 and global richness levels of nearly 76,000 species. Crustaceans may attain local species richness values of up to 150 and global richness values of 12,000 species; water mites up to 75 species locally and 6000 species globally; annelids up to 50 species locally and nearly 2000 species globally; nematodes and rotifers up to 500 species locally and 6000 species globally; and mollusks up to 100 species locally and 5000 globally.

Some freshwater habitats are hot spots of endemism harboring many unique fauna. For example, ancient lakes such as Lake Baikal in Siberia have rich abundance of endemic fauna, especially for those species that live all of their lives in fresh water and have a poor capacity for dispersal. For one group of crustacean amphipods, the gammarids, this lake has the highest level of endemism in the world (41 genera, of which 38 are endemic). Other lakes that harbor highly endemic invertebrate faunas include the African rift lakes of Lake

Table 1 Freshwater distribution by Continent[a]

	Africa	*Europe*	*Asia*	*Australia*	*North America*	*South America*
Lakes and reservoirs	31,240	2447	29,132	192	26,573	1199
Rivers and streams	195	80	565	25	250	1000
Groundwater	5,500,000	1,600,000	7,800,000	1,200,000	4,300,000	3,000,000
Wetlands	341,000	'Eurasia' 925,000		4000	180,000	1,232,000

[a]Modified from the World Conservation Monitoring Center's Freshwater Biodiversity (1998) *A Preliminary Global Assessment*. Data are volume of water in km^3, except for wetlands, which refer to area in km^2. Wetlands includes marshes, swamps, lagoons, and floodplains.

Table 2 A list of some of the most common freshwater invertebrates with estimates of the number of described species globally

	Common Name	*Aquatic Stage*	*Existence*	*Number of Described Species*
Phylum Porifera	Sponges	RW, SW	B	219
Phylum Cnidaria	Hydras, jellyfish	RW, SW	B, P	13–23
Phylum Platyhelminthes				
Class Turbellaria	Flatworms, Planaria	GW, RW, SW	B	1297
Phylum Rotifera	Rotifers	RW, SW	B, P	1948
Phylum Nemertea	Ribbon worms	SW	B	22
Phylum Nematoda	Free-living roundworms	GW, RW, SW	B	1800
Phylum Nematomorpha	Horsehair worms	RW, SW	B	326
Phylum Gastrotricha	Gastrotrichs	RW, SW	B	318
Phylum Bryozoa	Bryozoans	RW, SW	B	88
Phylum Tardigrada	Water bears	RW, SW	B	62
Phylum Annelida				~1800
Class Polychaeta	Bristleworms	GW, RW, SW	B	168
Class Oligochaeta (Clitellata)	Earthworms	GW, RW, SW	B	1100
Class Hirudinea (Clitellata)	Leeches	RW, SW	B	482
Phyllum Mollusca				5000
Class Bivalvia	Clams, mussels	RW, SW	B	1026
Class Gastropoda	Snails	GW, RW, SW	B	3795–3972
Phylum Arthropoda				
Class Crustacea				11,990
O. Branchiopoda	Clam, tadpole and fairy shrimp	RW, SW	B, P	508
O. Cladocera	Water fleas	RW, SW	B, P	620
O. Ostracoda	Seed shrimp	GW, RW, SW	B	1936
O. Copepoda	Copepods	GW, RW, SW	B, P	2814
O. Mysida	Opossum shrimps	GW, RW, SW	B, P	72
O. Isopoda	Sow bugs	GW, RW, SW	B	942
O. Amphipoda	Scuds	GW, RW, SW	B	1866
O. Decapoda	Crabs, shrimp, crayfish	GW, RW, SW	B, P	2832
Class Arachnida				
Subclass Acari: Hydrachnidia	Water mites	GW, RW, SW	B, P	6007
Class Insecta				75,000
O. Ephemeroptera	Mayfly larvae	RW, SW	B	3046
O. Odonata	Dragonfly and damselfly larvae	RW, SW	B	5680
O. Plecoptera	Stonefly larvae	RW, SW	B	3497
O. Heteroptera	True bugs	RW, SW	B, P	4801
O. Trichoptera	Caddisfly larvae	RW, SW	B	11,532
O. Megaloptera	Dobsonfly larvae, fishflies, hellgrammites	RW, SW	B	328
O. Neuroptera	Spongillaflies	RW	B	118
O. Coleoptera	Water beetles	RW, SW	B	12,604
O. Diptera	Fly larvae (e.g., midges, craneflies, blackflies, mosquitoes)	GW, RW, SW	B, P	32,281
O. Lepidoptera	Butterfly and moth larvae (caterpillars)	RW, SW	B	737

Note: Some groups are more common in running waters (RW), such as streams and rivers while others are more common in standing or semistanding water (SW) bodies (lakes, streams, wetlands) or groundwater (GW). Some are primarily benthic (B), existing in or on the bottom of lake, stream, river sediments, or on aquatic vegetation, while others maintain a water column existence as members of the plankton or pelagic fauna (P). Estimates of diversity are derived from the Freshwater Animal Diversity Assessment.

Tanganyika and Lake Nyasa, Lake Titacaca in Peru, and Lake Lanao in the Phillipines.

Dominant Taxonomic Groups across Freshwater Habitats

In running waters, the dominant invertebrate groups vary considerably depending on the type of bottom substrate (e.g., mud, cobble/boulder), the flow, and available sources of organic matter (Allan, 1995) (**Figure 3**). In rough-bottom creeks and swift-flowing streams, invertebrates live primarily on or beneath pebbles and boulders on the bottom and include many species of insects, crustaceans, and mollusks. In deep riverine areas, fine sediments predominate, the near-bed water flows more slowly, and planktonic invertebrates become common in the water column. In the soft sediments, burrowing insect larvae, bivalves, and worms are common as long as oxygen is present (Palmer *et al.*, 1997).

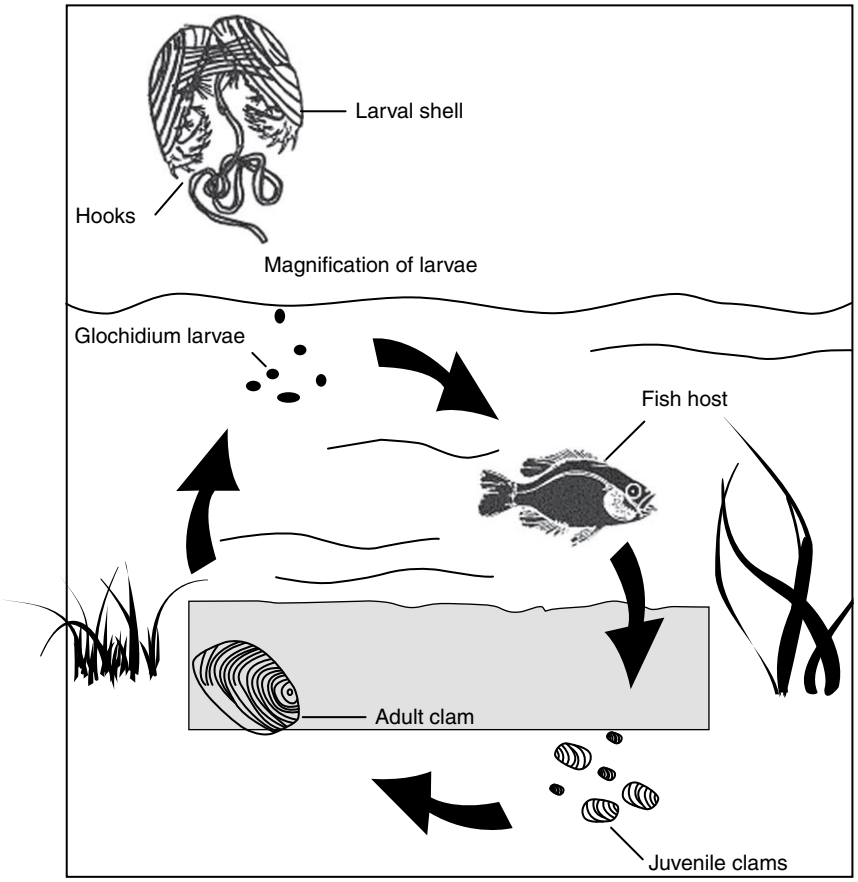


Figure 2 Example of a complex life cycle for a freshwater invertebrate. Many freshwater bivalves are benthic yet produce larvae that are released into the water column. These develop into small glochidium larvae that attach to a fish host. After some time, the larvae metamorphose into small bivalves, detach, and take on a benthic existence.

Benthic species can be classified according to how they feed, also known as functional feeding groups (Merritt and Cummins, 1996). "Shredders" are adapted to feed on decomposing plant material that has fallen into the stream. "Collector-gatherers" filter or gather particulate organic matter or drifting organisms from the water column. In streams where more sunlight reaches the bottom, animals known as "scrapers" feed on the algae that readily grows on rocks and sides of the channel.

In lakes and ponds, the most common invertebrates include crustaceans, rotifers, insects, and oligochaetes; however, taxonomic composition varies dramatically depending on the size of the lake and the position within the lake (Hutchinson, 1967/1993) (**Figure 4**). Benthic invertebrates are particularly abundant and speciose along the lake margins but generally are less numerous in the deeper parts of the lakes where oxygen may be limiting and habitat diversity is much lower. However, a few species that are tolerant of silt and low oxygen such as oligochaetes, midge larvae, and nematodes may be extremely abundant in the deeper part of lakes. In the water column, abundant planktonic invertebrate communities thrive by feeding on phytoplankton, but these invertebrate assemblages typically are dominated by just a few taxonomic groups such as rotifers, copepods, and cladocerans.

In the groundwater, crustaceans, rotifers, and nematodes are very common, and many of these are adapted for adept movement among rocks and particles of sand (Gibert *et al.*, 1994). In some parts of the world, insect larvae are common in the groundwater and may be found great distances away from the nearest surface water. Important crustacean groups in groundwater include copepods and less commonly amphipods, ostracods, isopods, and decapods. Many of these live in caves and sinkholes. Groundwater invertebrates are particularly interesting because of the unique adaptations many of them possess for life in dark, isolated habitats where food may be scarce. For example, the Edwards Aquifer in Texas harbors a unique fauna of 22 species including 10 amphipod species. Many of these species have reduced (or no) eyes, body pigmentation, and body size, low metabolic rates, and an enhanced sensitivity to touch.

In freshwater wetlands, the number of habitats and types of food sources for invertebrates are vast. The soft sediments are filled with burrowing nematodes, oligochaetes, midge larvae (a dipteran insect), and mites. Some invertebrates feed on or around the roots of vegetation, while others burrow through the sediments consuming bacteria, fungi, or other invertebrates. In the overlying water, crustaceans, rotifers, and insect larvae (e.g., mosquito larvae) are common and feed on phytoplankton or invertebrate prey. The stems of emergent and submerged vegetation in wetlands are

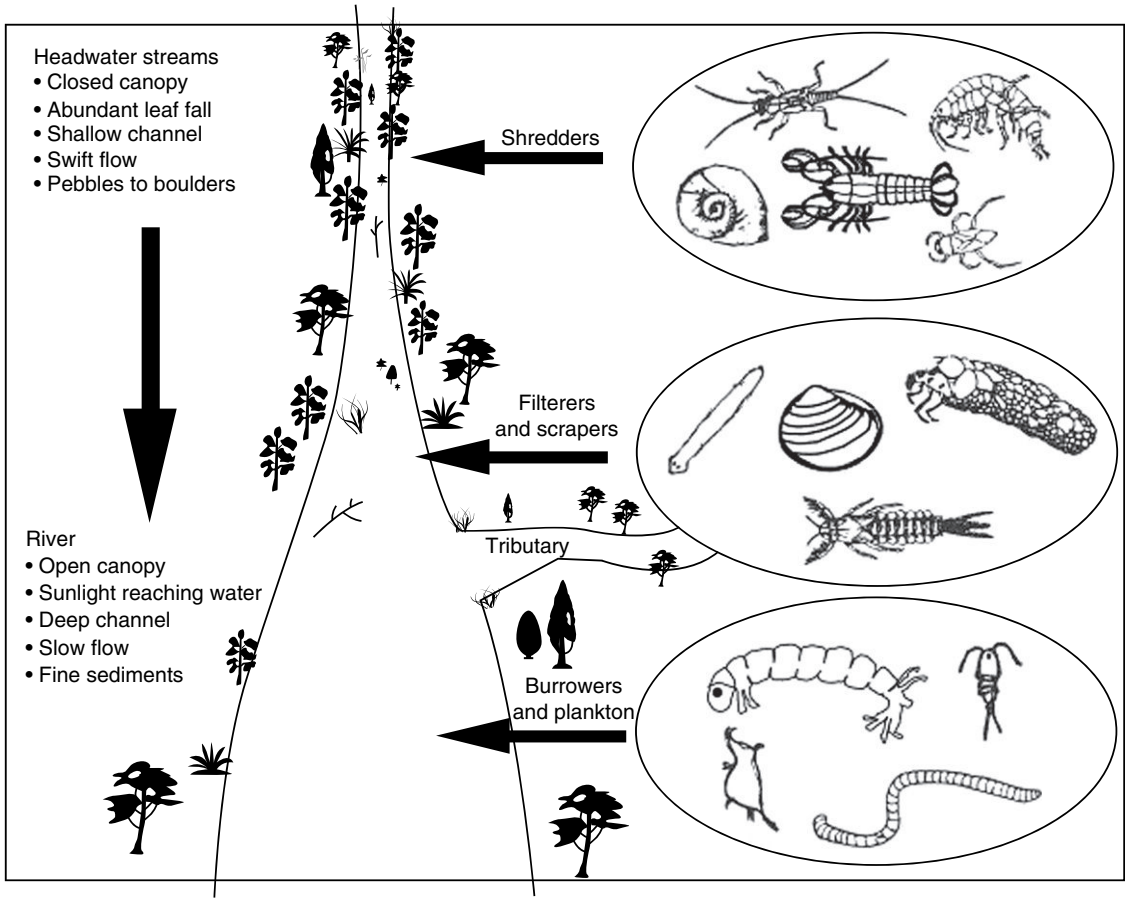


Figure 3 Schematic illustrating the River Continuum Concept (RCC) that proposed a relationship between stream size (small upland streams to large downstream rivers), flow and bottom substrate, and type of macroinvertebrate groups likely to dominate. In the smallest streams, large inputs of leaves and wood along with benthic algae are particularly important sources of food for invertebrates. These invertebrates "shred" the leaf material into finer particles and thus facilitate its further breakdown by fungi and bacteria. Further downstream, where sunlight penetrates the water and flow is reduced, algae are also important sources of food and many invertebrates "scrape" the algae from rock surfaces. In the very open riverine areas, plankton may be an important food source and many of the benthic invertebrates burrow into the fine sediments. Modified from Vannote RL, Minshall GW, Cummins KW, Sedell JR, and Gushing E (1980) The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37: 130–137.

typically crawling with diverse assemblages of crustaceans, mollusks, and insects. Small shrimplike crustaceans (amphipods) may reach very high levels of abundance on aquatic vegetation. Water beetles, odonates (dragonflies and damselflies), and waterbugs (water boatmen, waterstriders, and backswimmers) are commonly the top predators.

Global Patterns in Freshwater Biodiversity

Knowledge of lineage histories is very important because regional patterns of distribution of species are often a reflection of distinctly different evolutionary histories. Adaptations to specific climates and habitat characteristics influence speciation and the ability of an organism that evolved under one set of circumstance to move broadly among different geographic areas. Terrestrial and marine biologists generally have found the highest levels of animal and plant diversity in the tropics with lower levels in temperate and polar regions. This is not necessarily the case for freshwater invertebrates. Species richness for freshwater invertebrates is related in very complex ways to latitude and elevation, and global patterns of diversity are not straightforward for this group.

There have been far fewer surveys of freshwater invertebrates in the tropics than in temperate regions; however, those to date suggest that biodiversity in fresh water is not typically higher in low latitude, tropical habitats. For example, in North America, bivalves and crayfish appear to be the most diverse in temperate zones (about 260 native bivalves and 320 species of crayfish). Temperate zones have a wealth of crayfish species, but tropical fresh waters have a richer shrimp and freshwater crab fauna. Insects, mayflies, and stoneflies are more diverse in temperate latitudes than in the tropics, whereas dragonflies, waterbugs, and water beetles are much more diverse in the tropics.

Factors Influencing Biodiversity in Fresh Waters

Scientists have devoted their lifetimes to the study of what determines the number of species in a given habitat type or

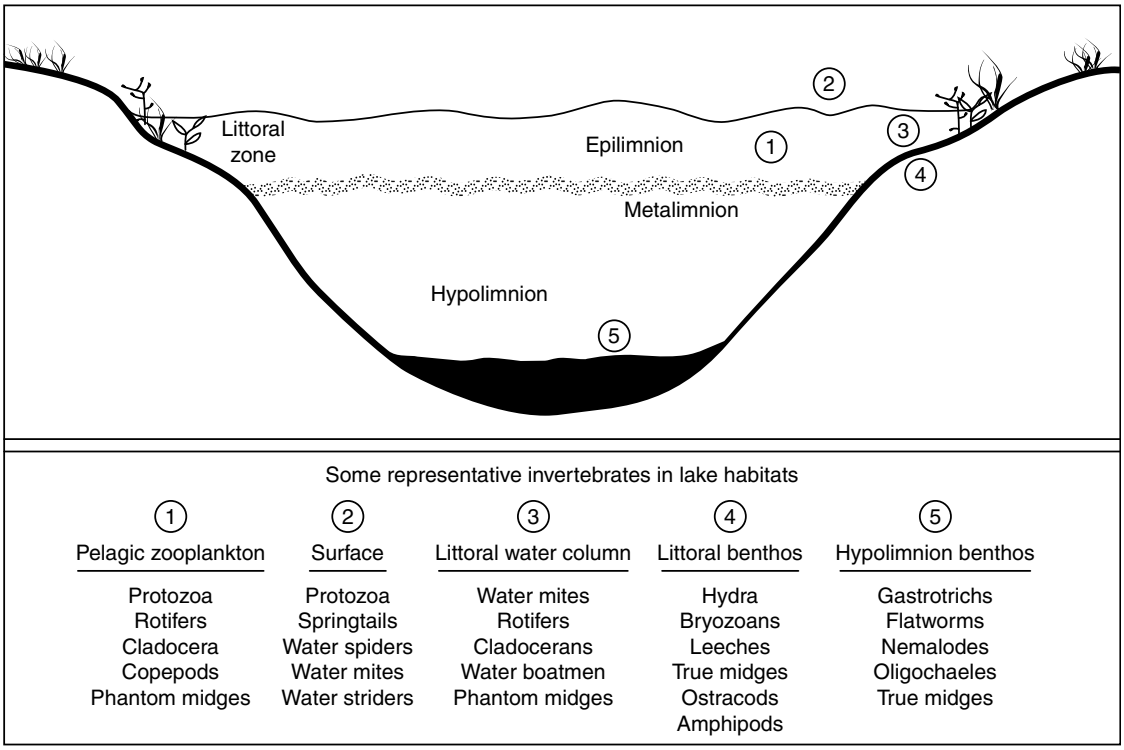


Figure 4 Representative invertebrates from different zones within a typical lake: (1) pelagic zooplankton; (2) surface; (3) littoral water column; (4) littoral benthos; (5) hypolimnion benthos.Modified from Figure 2.12 in Thorp and Covich (2010), *Ecology and Classification of North American Freshwater Invertebrates*, 3rd edn. New York: Academic Press.

geographic region. These studies have involved thoughtful, systematic observations of patterns of distribution dating back many years (e.g., Robert Pennak's work in many freshwater systems), and more recent experimental studies that have cleverly manipulated some factors believed to influence biodiversity. From such studies, we know that biodiversity in fresh waters is determined by both historical (e.g., age of a lineage, evolutionary constraints) and contemporary factors (e.g., habitat heterogeneity, predator–prey interactions). The relative importance of many factors varies geographically as well as between habitats in the same geographic region.

The Origin of Freshwater Invertebrates

Most freshwater invertebrates are believed to be derived from terrestrial or marine ancestors (Bănărescu, 1990). Representatives from many different groups of marine crustacea (shrimp, copepods, cladocera) may have made the transition from salt water to fresh water through many years of evolution in increasingly brackish (low salinity) water. Success in fresh water required that the animals be able to maintain ionic balance in their body fluids (not gain too much body water), reproduce without the aid of complex tidal currents to carry eggs and larvae, and, in some cases, develop life stages that could withstand freezing or drying. Only certain species could make this transition, so at the phylum level, freshwater biodiversity is more limited than in the oceans. For example, echinoderms (e.g., starfish) did not develop the ability to osmoregulate in fresh water.

Many freshwater invertebrates are secondary aquatic forms that moved, in an evolutionary sense, into fresh waters from the land. The most notable group of secondary invaders is the insects, many of which still have terrestrial adults. Some insect orders, such as the Heteroptera, Coleoptera, and Diptera, have both freshwater species and terrestrial species in abundance. Heteroptera and Diptera also have a few highly specialized marine species. Water mites (Acari) coevolved with insects and many species are parasites on the adult insects. With their coevolutionary links, they are also secondary invaders of fresh waters. Freshwater snails consist of two major groups, the prosobranchs and the pulmonates. Pulmonates have invaded freshwater systems from the land; freshwater forms include the two common and ubiquitous families of pond snails, the Lymnaeidae and Physidae.

Habitat Age and Isolation

Some freshwater habitats are extremely old and have been isolated for very long periods of time. These habitats have a high degree of endemism. For example, Lake Baikal in Siberia is extremely old and has many species that are found nowhere else in the world. Such "hot spots" of unique species are considered valuable for conservation purposes and for use as a "natural laboratory" for the study of invertebrate evolution.

For those ancient freshwater bodies that are highly isolated, resident fauna may be at a high risk of extinction. If a freshwater species is highly specialized for conditions at a particular site and conditions change, it may be unable to

survive. Geographic or habitat remoteness is not the only factor promoting isolation of freshwater invertebrates. Some fauna are isolated because they have very limited dispersal abilities. For example, many groundwater and deep-streambed dwelling species are very poor swimmers and lack dispersal stages. Thus local biodiversity may be low even though regional biodiversity is much higher – species from the large-scale regional "pool" of invertebrates (e.g., all rivers in a geographic region) may not be able to move between all watersheds. In general, ecologists assume that local species diversity represents a subset of the regional pool of species and the size of this subset depends greatly on both dispersal abilities of fauna and geographic barriers to dispersal.

Habitat Heterogeneity and Disturbance

A central tenet of ecological theory is that the greater the number and types of habitats, greater is the species diversity. Put simply, different types of organisms with different niche requirements can exist in areas that are heterogeneous over space and time. In freshwater ecosystems, spatial heterogeneity (and species diversity) is enhanced by anything that leads to a great variety of particle sizes in which benthic organisms can burrow and feed. Similarly, spatial heterogeneity and species diversity of planktonic invertebrates is enhanced by anything that promotes the growth of many types of vegetation because this provides the invertebrates with diverse food types and hiding places from predators.

Environmental variation over time (temporal heterogeneity) is also extremely important as a determinant of species diversity. Variable climate and flow regimes will promote seasonal (or longer) changes in species abundances and may determine whether a species can coexist in a location. Some freshwater invertebrates have evolved life cycles to take advantage of or avoid predictable changes in water level, flood flows, or ice scour. However, when these environmental changes become extreme or highly unpredictable (a "disturbance"), species richness may decline as fauna are unable to evolve coping strategies that keep up with such changes. Many ecologists believe that invertebrate species diversity is greatest when natural disturbances like floods are at intermediate levels with respect to severity or predictability. High levels of disturbance will wipe out most species while low levels of disturbance may favor one or two species to the exclusion of other species that are poorer competitors.

Species Interactions

Predation and competition are well known to influence local species diversity in freshwater ecosystems. For example, many fish predators are highly selective in their feeding and may forage on a single species or suite of species. Thus, predation pressure may reduce species diversity locally, or it actually may promote species diversity if the predator is removing a species that is a competitive dominant. Competitive dominants are species that are so good at garnering resources (space, food, etc.) that other species are unable to coexist with them. The process of competitive exclusion also may be prevented by disturbances if the disturbance (e.g., a flood or drought) has a disproportionate effect on the competitive dominant.

In freshwater ecosystems, some of the most dramatic examples of the delicate balance that species interactions play in determining biodiversity come from the introduction of exotic species. Many exotic fish have been introduced into lakes throughout the world and have had devastating effects on their invertebrate prey or on other fish that in turn prey on invertebrates. Introduced carp have reduced invertebrate populations in floodplain wetlands. Nonnative trout introduced to streams have depleted populations of native fish and increased trout predation pressure on native invertebrates.

The Role of Invertebrates in Ecological Processes

Healthy freshwater ecosystems are those in which the ecological processes continue unimpeded to ensure that water is clean and the organic matter in a lake, stream, or wetland is not lost or accumulated in excess. Freshwater invertebrates play key roles in these processes. The most important ecological processes in freshwater systems include decomposition of the organic matter; the uptake and transfer of materials such as sediments, nutrients, and contaminants; and production by plants.

Decomposition of Organic Matter and Recycling of Nutrients

Decomposition releases elemental carbon and nutrients (nitrates, phosphates) to the environment with the result that plants and animals that rely on these nutrients thrive. Additionally, decomposition ensures that wastes and dead organic material do not build up in bodies of water. Terrestrially produced leaf and woody litter and dead aquatic vegetation and plankton sink to the bottom of lakes, streams, and wetlands. Once deposited within fresh waters, this organic matter can be broken down directly by bacteria and fungi. Often however, the material first must be altered by invertebrates, which fragment or shred it before ingesting it. The shredding activities of invertebrates are important because smaller fragments of dead leaves and other organic matter are decomposed by microbes much faster than larger pieces. Additionally, through their feeding activities, invertebrates can enhance the abundance and reproduction of the microbes, which further acts to stimulate the rates of decomposition.

The Uptake and Transfer of Materials

Movement of water, sediment, and organic material within and from freshwater systems has a profound effect on global biological, geological, and chemical processes. Freshwater invertebrates may alter the rate at which water flows through sediments. For example, burrowing oligochaetes, nematodes, midge larvae, and crustaceans can increase the rate at which water moves through lake and stream sediments and to nearby groundwaters. Instead of increasing water flow and percolation through sediments, some freshwater invertebrates bind

sediments by producing mucus "biofilms" or dense mats of tubes or filaments that reduce water flow and the movement of sediments. This may act to improve water clarity so that sunlight penetrates deeper and plant production is enhanced. Similarly, water clarity may be enhanced by dense populations of bivalves that filter water as they feed and remove suspended matter. Freshwater invertebrates also may influence the concentration of contaminants in the water or bottom sediments by accumulating the contaminant in their bodies or *via* direct degradation of toxic materials.

Production by Plants

In freshwater ecosystems, the synthesis of organic matter by plants depends on the presence of sunlight and nutrients. Rates of primary production are extremely variable across freshwater ecosystems, ranging from low in poorly lit groundwaters or turbid lakes to quite high in well-lit shallow wetlands. Aquatic herbivores may directly affect plant production as the mifoil weevil, *Euhrychiopsis lecontei*, does in feeding on both northern (*Myriophyllum sibiricum*) and Eurasian (*M. spicatum*) watermilfoils in North American lakes. However, the indirect effects that invertebrates have on plant production are also significant. The availability of nutrients for aquatic vegetation and phytoplankton depends to a greater extent on the decomposition of organic matter in aquatic sediments and the movement of nutrients. Grazing and movement of invertebrates stir up the bottom sediments and mix water, increasing the availability of nutrients to plants and phytoplankton. Furthermore, planktonic invertebrates that graze on phytoplankton excrete forms of phosphorus and nitrogen that are immediately available to further enhance algal growth. Grazing may be particularly important in the removal of senescent algae, thereby increasing light as well as nutrients.

Threats to Freshwater Fauna

The loss of freshwater invertebrate biodiversity because of damage generated by human activities is an immense global problem (Strayer, 2006). A major difficulty in assessing the dimensions of the problem is the poor state of knowledge of freshwater invertebrate taxonomy and distribution and the lack of reliable monitoring to identify loss of diversity. In spite of these problems, it is clear that invertebrate species in some places already have high levels of population decline and extinction. In fact, it has been estimated that over 10,000 freshwater species are already extinct or greatly endangered as a result of human activities.

The major threats to freshwater biota in general, and freshwater invertebrates in particular, are habitat degradation and loss, pollution and reduced water quality, altered hydrologic regimes, and invasion by exotic species. In many situations, the major threats do not act alone but can act synergistically; moreover, they will interact with and possibly be exacerbated by global climate change (Allan, 2004). Thus, for example, streams in agricultural areas may lose their invertebrate fauna due to habitat loss by sedimentation, pollution by pesticides, altered flow regimes due to irrigation, and the introduction of exotic species, including plants, fish, and invertebrates. Increased variability in the frequency and intensity of rain events as well as rising temperatures may further stress any vulnerable invertebrates remaining. A fifth threat to freshwater fauna is overexploitation. For example, harvesting of the mussel *Margaritifera margaritifera* for pearls has reduced populations by 90%.

Loss and Depletion of Habitat

Humans have damaged water bodies by river channelization and dredging, sand and gravel extraction, wetland drainage, lake and river shore development, dam and barrier construction, water diversion, levee bank construction, and valley fills associated with mountain top mining. Human activities on land can have indirect but significant effects on fresh waters. For example, sedimentation from poor land use, destruction of riparian vegetation, loss of surface runoff, and loss of water by groundwater extraction for irrigation, all can damage fresh waters. Indeed, studies have shown that when just 5–10% of a watershed's area is affected by anthropogenic activities, stream biodiversity and water quality suffer.

Temporary wetlands are common in many regions. They may have surface water only occasionally but when they do, they are very productive and harbor a distinctive invertebrate fauna adapted to survive long dry periods. Unfortunately, in many areas the importance of temporary wetlands as distinctive habitats of high biodiversity value has not been recognized and many of these wetlands have been plowed, drained, and filled in.

Sedimentation of streams and wetlands from erosion can be caused by activities in watersheds such as land clearance, plowing and tillage of soil, road building, and logging. The mass delivery of sand and silt into water bodies can reduce invertebrate diversity and change species composition by filling in pools in streams, wetlands, ponds, and even lakes, burying porous coarse sediments with layers of fine sediments with very low permeability, and covering the aquatic plants. The deep streambed of running waters has a rich and distinctive fauna dependent on the high permeability of gravels and sands. Sedimentation can clog the pores in the streambed reducing water movement and oxygen availability, which leads to the loss of fauna. Many streams that originally had stable, habitat-rich channels, and a rich invertebrate fauna have been damaged when channels were filled by fine shifting sediment that consequently has led to a greatly depleted fauna.

A property vital to the nature of rivers, and essential for maintenance of invertebrate biodiversity, is connectivity – the unimpeded movement of water, longitudinally from source to mouth, and laterally between the channel and its floodplain. Loss of connectivity greatly alters invertebrate species composition and may reduce diversity. Dams on rivers are barriers to longitudinal connectivity because they disrupt the downstream movements of nutrients and organic matter and prevent movement of invertebrates and fish. Levees are barriers to lateral connectivity because they are designed to stop lowland rivers from flooding their floodplains; however, periodic

floodplain inundation is essential for the maintenance of biodiversity of the river system.

Pollution and Reduced Water Quality

Pollutants can enter fresh waters through point sources such as sewage outfalls, or *via* diffuse nonpoint sources such as runoff from agricultural fields. Pollutants in sewage and agricultural runoff include organic matter and nutrients. Organic pollution of rivers usually is released from point sources, and downstream from the point source, dissolved oxygen concentrations may drop greatly. The normally diverse fauna of the river is eliminated and replaced by an abundance of a few species that can tolerate low oxygen levels. As the organic matter is processed, the river may recover further downstream. In lakes contaminated by organic pollution, decomposition of the organic matter and elevation of the nutrient levels also may create low oxygen conditions in the benthic regions so that the normally diverse benthic invertebrate fauna is replaced by species similar, if not identical, to those found in the anoxic sections of organically polluted rivers.

Nutrients such as phosphorus and nitrogen enter fresh waters either from sewage or from diffuse sources, such as agricultural fertilizers. Nutrients may encourage the growth of undesirable algae and may cause algal "blooms" and periods of low oxygen availability, especially in the deeper benthic regions. This process is called cultural eutrophication and is a major problem for lakes in both urbanized and rural areas (Scheffer, 1998). In such lakes, the invertebrate fauna, both planktonic and benthic, is depleted in species diversity, but abundances of a few tolerant species may be very high.

Combustion of fossil fuels can generate sulphur and nitrogen oxides that combine with water to form acids in the atmosphere. Acids dissolved in water vapor can move long distances in the atmosphere before entering water bodies as acid rain. This process is a dramatic example of nonpoint source pollution in which pollutants generated in one region exert their effects on another, often distant, area. Many poorly buffered lakes in Europe and North America have become strongly acidic due to acid rain. In acid lakes ($pH < 4.5$), some invertebrate groups such as crustaceans and mollusks are eliminated along with gill-breathing insect larvae. The most acid-tolerant taxa appear to be predatory waterbugs.

In many parts of the world, irrigation may reduce river flows and create extensive salinization (salt contamination). Salinization occurs when water evaporates from irrigated fields and leaves behind the salts that were dissolved in the water. The increased salt levels in the soil kill terrestrial vegetation. The concentrated salt can enter rivers and lakes where it eliminates most of the freshwater invertebrates, except for a few tolerant midge larvae and crustaceans. The Aral Sea in Russia was once brackish water and fauna rich. Its freshwater inputs were diverted for irrigation, and it now receives very saline waters from irrigated lands. As a consequence, it has shrunk greatly in size to become a saline sea with a very impoverished fauna.

Many pollutants, such as heavy metals and chlorinated hydrocarbons, are not readily broken down or deactivated. Once they enter a freshwater system through point or nonpoint sources, they persist and continue to cause damage. Heavy metals, such as zinc, copper, cadmium, and lead, may enter a water body from mining operations and poison the system, thereby eliminating many invertebrate taxa. Unfortunately the contamination may persist for many years, even long after mining operations cease. In Wales (Britain), Australia, and many parts of the US, rivers have remained contaminated 35–90 years after mining ceased and they continue to have a depauperate invertebrate fauna.

Altered Water Regimes

In the past century with the growth in human populations and the rapidly rising demands for water for human consumption, industry, hydropower, and irrigation, there has been an enormous expansion in the number and scale of dams, barrages, and diversion works. Dams and their large upstream impoundments break the connectivity of rivers. This disruption has been recognized in situations where valuable fish stocks have been threatened. Fishways, porterage schemes, and stocking have been used to reduce threats to fish, but nothing has been done to alleviate the effects on invertebrates. Thus dams may disrupt essential migrations, such as those of shrimps and freshwater crabs in the tropics, and prevent the normal movements of invertebrates down the river.

The waters impounded behind dams may flood valuable habitats – floodplain wetlands, river canyons, even lakes. In Tasmania, Australia, a large dam-created impoundment flooded an entire isolated lake with a remarkably large beach. This lake, Lake Pedder, contained five endemic invertebrate species, four of which appear to have become extinct. Impoundments behind dams may have a greatly depleted littoral (shoreline) fauna due to rapid changes in the water level incurred as a result of dam operations. Drawdown of lake levels to meet hydroelectricity or irrigation can leave many invertebrates stranded. In unimpounded rivers, organic matter and nutrients are transported downstream to be used by downstream invertebrates. This longitudinal transport ceases with impoundment. Water released from dams may come from the deep hypolimnion and consequently may be cold, deoxygenated, and low in organic material – conditions quite unfavorable for riverine invertebrates.

The purposes for which a dam is operated can have direct consequences for downstream invertebrates. Most drastic of all is the simple loss of water. Many dams are water diversion structures that leave the river channel below the dam waterless or with a small minimum flow. In many places, large rivers are left with so little water that they no longer reach the sea. Needless to say, the loss of invertebrates is considerable. Dams may be built and operated to generate hydroelectricity, wherein water levels in the channel below the dam fluctuate greatly and quickly as water is released or suddenly cut off to meet the electricity generation demands. Alternation between flood and drought creates disturbance levels that permit only the hardiest of invertebrates to exist in the channel. Dams operated to store water for irrigation, hold water in the wet season (winter), and release it for irrigation in the dry season. Consequently, the seasonal patterns of flow are reversed, and wet season floods are

greatly diminished. The lack of water in the wet season and high water levels in the dry season disrupt the life cycles of many invertebrates. For many lowland rivers, the effects of greatly diminished floods have been devastating because inundation of the floodplain has been greatly reduced, resulting in a loss of invertebrates that would normally thrive on the floodplain.

Invasion of Exotic Species

Exotic species of plants and animals may reach new localities through deliberate introduction by humans or accidental introduction during transport. The rate of introductions both between and within continents is rising rapidly. The most obvious introductions into freshwater systems have been plants, both aquatic (e.g., water hyacinth) and riparian (e.g., salt cedar), and fish (e.g., carp). Few invertebrates have been introduced deliberately (e.g., opossum shrimp, crayfish), but many, especially crustaceans and mollusks, have been introduced accidentally (e.g., spiny and fishhook water fleas, zebra mussels).

Introduced invertebrates may exert very strong effects on native species. Exotic invertebrates may be vectors for diseases lethal to native species. In Europe, native crayfish stocks have been decimated by the crayfish plague fungus, *Aphanomyces astaci*, which came with the deliberate introduction of North American crayfish. Exotic species may take up habitat space and outcompete local species. Zebra and quagga mussels from the Ponto-Caspian region invaded the Great Lakes, spread rapidly through eastern North America, and have now been detected in reservoirs and lakes in Colorado, Utah, and California. They monopolize habitat space and have greatly reduced native mussels (Strayer *et al.*, 1999). Zebra and quagga mussel populations increase rapidly once introduced to new water bodies and through their immense filtration capacity may greatly reduce planktonic organisms, increasing the clarity of water but reducing the invertebrate diversity.

Introduced aquatic plants, such as water hyacinth and alligator weed, may blanket the surfaces of lakes and rivers. Such blanketing may reduce water quality by causing deoxygenation or by reducing photosynthesis of native plankton and aquatic plants. This, in turn, diminishes food resources for native invertebrates. Introduced riparian plants, such as tamarisk or salt cedar, may cause the loss of habitat for native invertebrates by dominating the stream banks and often by reducing water levels in the channel. Other introduced plants, such as eucalypts in Spain and Portugal, may produce dead plant material that is not readily consumed by stream invertebrates, thereby causing them to starve.

Many fish are predators of the invertebrates. The introduction of exotic fish may increase predation pressure and deplete native invertebrate populations. Introduced carp have reduced invertebrate populations in floodplain wetlands. In New Zealand, introduced trout in invading streams have depleted or eliminated populations of native fish and native invertebrates. Released from the intense grazing pressure of native invertebrates, algae in such invaded streams proliferate and build up to high levels.

See also: Endangered Freshwater Invertebrates. Hotspots. Invertebrates, Marine, Overview. Invertebrates, Terrestrial, Overview. Lake and Pond Ecosystems. River Ecosystems. Wetlands Ecosystems

References

Allan JD (1995) *Stream Ecology: Structure and Function of Running Waters*. New York: Chapman & Hall.

Allan JD (2004) Landscapes and riverscapes: The influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution and Systematics* 35: 257–284.

Balian EV, Lévêque C, Segers H, and Martens K (eds.) (2008) Freshwater animal diversity assessment. *Hydrobiologia* 595, pp. 1–637.

Bănărescu P (1990) *Zoogeography of Fresh Waters*, vols 1–3. Wiesbaden, Germany: Aula-Verlag.

Gibert JA, Danielopol DL, and Stanford JA (eds.) (1994) *Groundwater Ecology*. San Diego, CA: Academic Press.

Hutchinson GE (1967/1993) *A Treatise on Limnology*, vols 2 and 4. New York: Wiley.

Merritt RW and Cummins KW (1996) *An Introduction to the Aquatic Insects of North America*. Dubuque, Iowa: Kendall/Hunt Publising.

Palmer MA, Covich AP, Finlay BJ, *et al.* (1997) Biodiversity and ecosystem processes in freshwater sediments. *AMBIO* 26: 571–577.

Scheffer M (1998) *Ecology of Shallow Lakes*. London: Chapman & Hall.

Strayer DL, Caraco NF, Cole JJ, Findlay S, and Pace ML (1999) Transformations of freshwater ecosystems by bivalves: A case study of zebra mussels in the Hudson River. *BioScience* 49: 19–27.

Strayer DL (2006) Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society* 25: 271–287.

Thorp JH and Covich AP (2010) *Ecology and Classification of North American Freshwater Invertebrates*, 3rd edn. New York: Academic Press.

Vannote RL, Minshall GW, Cummins KW, Sedell JR, and Gushing E (1980) The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37: 130–137.

Invertebrates, Marine, Overview

PJD Lambshead, The Natural History Museum, London, UK
PH Schalk, ETI, Universiteit van Amsterdam, Amsterdam, Netherlands

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 543–559, © 2001, Elsevier Inc.

Glossary

Benthic That connected with the bed of the sea or a freshwater lake, river, or stream.
Deep sea The seabed and immediately overlying water covered by seas at least 200 m deep.
Disturbance A biological or physical factor that impacts a population or community by causing death, reduced reproduction, or increased emigration.
Infauna Animals living within sediments.
Invertebrate All animals that lack a backbone (vertebrae)—that is, most of the animal kingdom.
Pelagic That connected with the water column.
Planktonic Organisms of many different phyla that float in and are carried by water masses in the pelagic of the sea or freshwater.
Production The increase in biomass of an individual, population or community as it grows by converting energy-food into biomass.

Physical Characteristics of The Oceans

The world ocean covers more than two-thirds of the surface of the earth and as much as 90% of the habitable volume. It is the largest habitat on earth and about one billion people depend on the oceans for their primary protein needs. Yearly, well over 90 million tons of fish and shellfish are harvested.

Only 10% of the oceans' area could be classed as shallow water and this is mostly located around the continental margins. Seas deeper than 2000 m cover half the earth's surface. The seashore zone, the area covered and exposed by tides, corresponds to the ecological littoral region. Shallow seas make up the sublittoral zone, also called the continental shelf, which ends at the shelf break, usually at around 200 m. The continental slope marks the boundary between the continental and ocean floor crusts. It extends from the shelf break down to about 1500 m, so it roughly corresponds to the ecological bathyal region, 200 m to 2000 m.

The continental slope falls away into the continental rise, which extends down to a depth of about 4500 m. Beyond this is the abyssal plain at depths of between 5000 to 6000 m. The continental rise and the abyssal plain roughly correspond to the ecological abyssal region, 2000 to 6000 m. Other features of the seafloor are rocky seamounts and ridges, which may be of a considerable size, and deep trenches down to 11,000 m or more. Trenches are usually a feature of subduction zones where the seafloor buckles and deepens beneath a continental crust. Deep trenches correspond to the ecological hadal region. The continental rise and abyssal plain tend to be relatively flat. Geological features such as terraces and submarine canyons, however, sculpture the continental slope. Pressure increases down the water column at the rate of one atmosphere per 10 m.

Much of the seafloor is covered in sediments; bare rock is rather rare. Near the coast these sediments are of terrigenous origin and may be coarse (e.g. clay, sands, or pebbles). Terrigenous material may be found down to the continental rise and vast areas of the seafloor consist of clay sediments or, in productive areas, biogenic oozes mainly from diatoms, radiolarians, and Foraminifera.

The temperature of the water declines rapidly with depth until it stabilizes at about 4 °C, which takes place at approximately 1000 m, depending on latitude. This phenomenon is known as the permanent thermocline. At greater depths, the temperature declines slowly. Water currents are too complex to be discussed here in detail but, broadly speaking, deep water is formed by the sinking of dense, cold saline water in the Antarctic and Arctic. This dense water spreads out across the world ocean, eventually returning to the surface.

In general, marine water is well oxygenated, often at values near saturation. There are areas where rapid seasonal blooms in the upper water column deoxygenate the water leading to oxygen minimum zones. Oxygen concentration falls rapidly with sediment depth, especially in littoral muds and sediments under shallow enclosed seas.

The open ocean or pelagic realm has no obvious boundaries such as the mountain ranges or large water masses (rivers, seas, or oceans) that are so important in driving speciation on land, although to a certain extent the three oceanic basins—Atlantic, Indian, and Pacific—are partially separated by landmasses. The world ocean is essentially a continuum and the most important environmental factors are temperature and salinity, which translate primarily into latitude and depth. However, despite the lack of obvious barriers, speciation has and does occur in the pelagic environment. The geographic species distribution of modern plankton is the product of the geological history of the oceans and continental barriers, limitations of species in adapting to biotic and abiotic environmental factors, and the degree of organism mobility.

The pelagic realm is divided by depth into zones. Light, temperature, and salinity are responsible for the major vertical biodiversity gradients. In the upper 200 m, the epipelagic, sufficient light is available to sustain photosynthetic processes. This is the primary production zone where light and carbon dioxide is converted into carbohydrates, the principle energy (food) source for all open ocean inhabitants. The mesopelagic

Table 1 The division of the pelagic environment into depth zones (measured in meters)

Epipelagic	0	200	The photic zone, primary production
Mesopelagic	200	2000	The realm of diurnally migrating organisms
Bathypelagic	2000	6000	The deep-sea realm, continuous darkness
Abyssopelagic	>	6000	The deep-sea troughs, continuous darkness

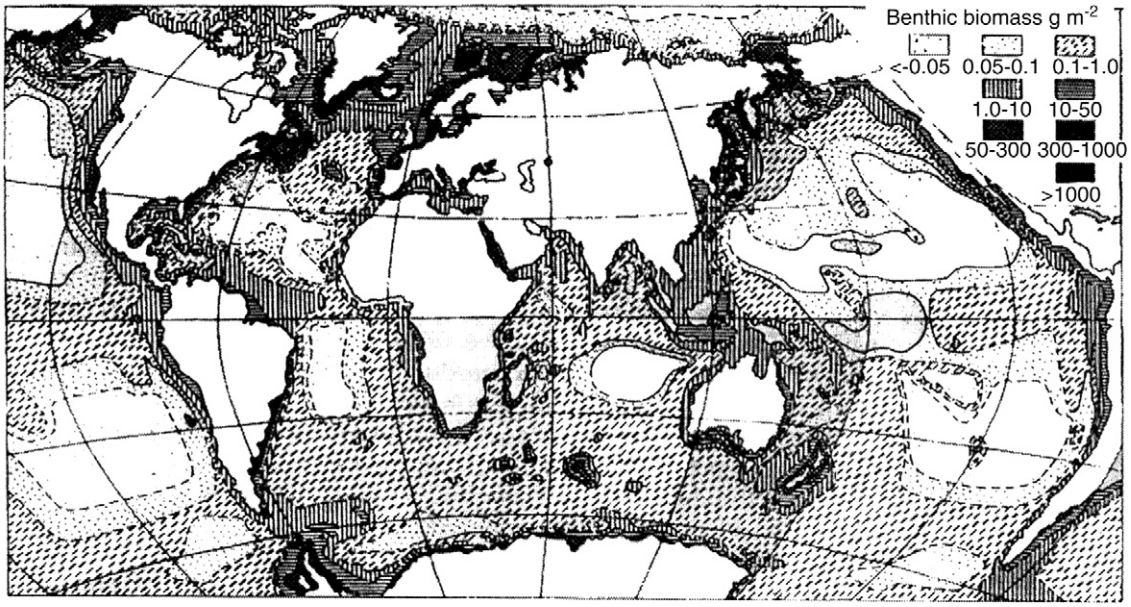


Figure 1 The global distribution of benthic biomass in the world ocean based on the work of the Russian sampling programme of Belyaev and colleagues. This work established that biomass declines in samples proportionally to their distance from continents. Biomass is proportional to food input. (Reproduced with permission from Gage JD and Tyler PA (1991) *Deep-Sea Biology*. Cambridge: Cambridge University Press.)

zone is the realm of the diurnally (daily) migrating organisms and extends from 200 to 2000 m depth. This is the deepest into the water column that light penetrates. Diurnal migration is triggered by light and typically involves moving into shallow waters at night and deeper water layers during the day. Possible reasons for such vertical migration may include hiding in the dark depths from predators during daylight and energy conservation achieved by spending part of the time in deeper, colder water layers, thus allowing organisms to reduce their metabolic rate. The bathypelagic (2000 to 6000 m) and abyssopelagic are the aphotic zones, where there is perpetual darkness (**Table 1**).

The Fundamental Processes Controlling Marine Biodiversity

There are three major differences between marine and terrestrial systems that have an enormous impact on biodiversity. The first is that the oceans are three-dimensional; organisms live, feed, and reproduce at all levels in the water column. The second is that light has low penetration through sea water, so photosynthesis only occurs in the upper water column, the photic zone, so most of the oceanic system has no primary production. The third is that most photosynthesis is carried out by tiny single-celled organisms. Macrophotosynthetic organisms (macroalgae or seaweeds) are rare and tend to be concentrated in a thin zone around the edge of the continents in the littoral and sublittoral regions.

Pelagic organisms in the three-dimensional water column have no structure to "hide" in, either for ambush attacks or defense. Essentially, there are only three possible life strategies. An organism can evolve for speed and agility to capture prey and escape predators, or an organism can evolve as a jelly, offering a poor food return for potential predators. A third strategy is rapid growth and reproduction resulting in huge numbers: schools and swarms offer protection to those within. Many pelagic organisms show a patchy distribution for this reason. The organisms within the sediment, the infauna, exist in a two-dimensional world that has a physical structure within which organisms may shelter. The marine system, with regard to biodiversity, may therefore be divided into two separate domains, the pelagic and the benthic. Some infaunal taxa have a pelagic phase and so have a presence in both domains. We may predict that marine pelagic diversity should be lower than terrestrial diversity given the limited choice of life strategy of pelagic organisms coupled with the active or passive mobility of organisms through the water column. The

benthic domain is structurally similar to the terrestrial environment and may be predicted to have a similar diversity with the caveat that the absence of large photosynthetic organisms removes an entire biotope.

The lack of primary production in most of the oceanic realm results in a dependence on an organic flux that originates on land or in the photic zone. The organic flux becomes weaker further away from the continents or deeper into the water column. The weakening flux gradient governs benthic biomass, which declines with depth and distance from the continental shorelines. The Russian grab-sampling programs of Belyaev and colleagues from the 1950s onward established the global distribution of benthic biomass (**Figure 1**). Productivity has a major influence on biodiversity so the organic flux gradient also will have an impact on benthic biodiversity patterns.

Physical disturbance is also considered to be a key process controlling biodiversity. Physical disturbance was considered to be highest in the littoral region and to decline consistently with water depth down to the abyssal plain. Recent research has caused science to modify this view, as there is now good evidence for hydrodynamic disturbance and seasonal perturbations right down onto the abyssal plain. But in general, the trend is for physical disturbance to decrease with depth. The main physical disturbance processes that show a depth trend are mechanical effects (current energy and wave action), temperature variation (including exposure in the littoral region), and salinity change. Deep trenches are possibly highly disturbed due to slumping of the trench sides.

Sampling and Assessing Marine Invertebrate Biodiversity

Benthos

Marine invertebrates are divided into three size classes on the pragmatic basis of the equipment used to collect them (**Table 2**). Megafauna are large, visible animals that may be seen by eye (or on photographs). Macrofauna are infaunal (sediment dwelling) organisms that are not normally visible but are retained on a 1mm sieve (0.5 mm or less in the case of deep-sea samples). Meiofauna are the infauna too small to be retained on such a sieve. It is not clear whether these arbitrary classifications also have biological significance. Some taxa are almost entirely in one size class for their entire life history; for example, nematodes are always considered to be meiofauna. Others—for example, polychaetes, although predominantly in one size class, in this case macrofauna—may also have meiofaunal and megafaunal representatives for at least part of their life history. Warwick and colleagues have some evidence, most convincingly from coastal sandy sediments, that the division of organisms into macro and meiofauna is ecologically meaningful. It appears that organisms must be either macrofaunal-sized, to actively burrow through the sediments, or meiofaunal-sized, to slip between the sediment particles, but mechanical limitations prevent them being of intermediate size.

Megafauna are numerically dominated by the phylum Echinodermata, although the Crustacea are also important. The most abundant and diverse macrofauna group are the

Table 2 Examples of marine benthic invertebrate taxa classified into the three size groups

Size class	*Megafauna*	*Macrofauna*	*Meiofauna*
Collecting equipment	Agassiz trawl, epibenthic sled, anchor dredge	Grab, corer (notably the USNEL box corer)	Corer (notably the SMBA multiple corer)
Examples of taxa	Ascidiacea (sea squirts)	Arthropoda (amphipods, crustaceans, cumaceans, isopods, mites, tanaids)	Nematoda (thread worms)
	Asteroidea (sea stars)		Copepoda (mainly harpacticoids)
	Brachiopoda (lamp shells)	Mollusca (aplacophorans, bivalves, chitons, gastropods, scaphoda)	Gastrotricha
			Kinorhyncha
	Bryozoa (moss animals)		Loricifera
	Coelenterata (sea firs, jellyfish, sea pens, sea fans, sea anemones, corals)	Oligochaeta (“earth” worms)	Ostracoda
		Phorinda	Tardigrada (water bears)
	Crinoidea (sea lilies, feather stars)	Pogonophora (tube worms)	Turbellaria (flat worms)
	Crustacea (amphipods, barnacles, crabs, prawns shrimps)	Polychaeta (segmented sea worms)	
		Priapulida (worms)	
	Echinoidea (sea urchins)	Nemertea (ribbon worms)	
	Echiura (worms with elongate bifid proboscis)	Sipuncula (peanut worms)	
	Hemichordata (acorn worms)	Turbellaria (flat worms)	
	Holothurioidea (sea cucumbers)		
	Ophiuroidea (brittle stars and basket stars)		
	Pogonophora (including vestimentiferans)		
	Porifera (sponges)		
	Pycnogonids (sea spiders)		

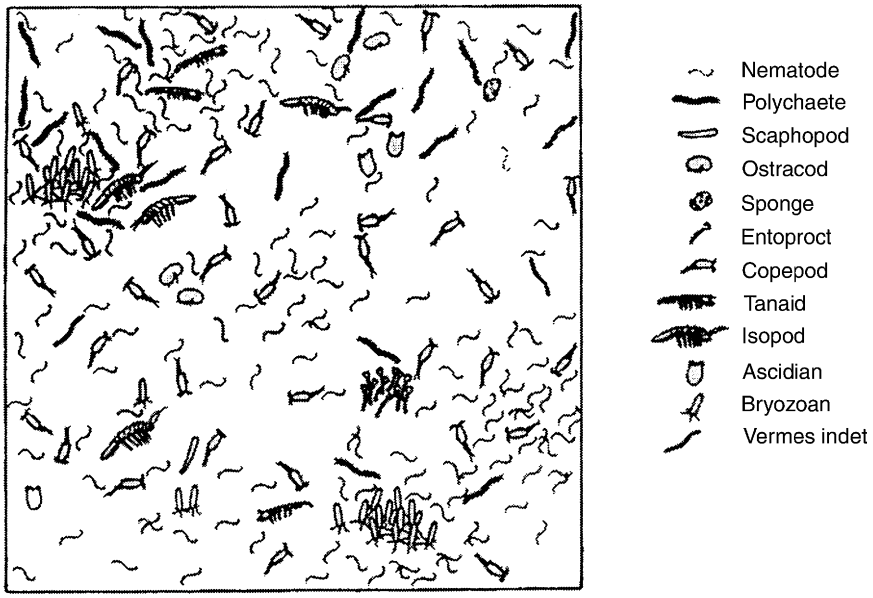


Figure 2 The invertebrate animals found in a 0.25 m^{-2} layer of sea sediment from the central North Pacific by Hessler and Jumars. A worm shape is the most common body form, nematodes being particularly abundant. Note that the size of the animals is exaggerated so that they may be seen. (Reproduced with permission from Gage JD and Tyler PA (1991) *Deep-Sea Biology*. Cambridge: Cambridge University Press.)

polychaete worms (Annelida), while meiofaunal nematodes (Nematoda) dominate the benthos, at least in terms of abundance and diversity of meiofauna. Arthropods are also important in both the macro and meiofaunal size classes. The average marine benthic infaunal invertebrate is a soft-bodied worm (Figure 2), but vermiform animals are less important components of the megafauna.

Littoral benthos from sediments is sampled by inserting a corer into the deposits and washing the removed sediment over an appropriately sized sieve. Sampling offshore is more problematical. Megafauna are collected by equipment such as Agassiz trawls, anchor dredges, or epibenthic sledges towed along the bottom. Such equipment has problems collecting consistent samples and is suitable only for producing qualitative or semiquantitative data. Initially, the smaller infaunal animals were collected with grabs, and these are still used where circumstances dictate but they are inefficient. The development of the box corer by Hessler and Jumars was a critical improvement in the accurate quantitative sampling of macrofauna for biodiversity studies. Development of quantitative meiofauna samplers, notably the Scottish Marine Biological Association's multiple corer, similarly revolutionized investigation into offshore meiofauna biodiversity.

Marine samples taken with modern corers provide data that are more quantitatively accurate than most terrestrial studies, and this has influenced the type of biodiversity questions asked by marine scientists. Marine research has focused on alpha, or ecological, diversity, which is commonly analyzed using diversity indices that incorporate some measure of evenness as well as species richness. The two most common indices employed are the Shannon-Wiener information function (H′) and, in the deep sea, the Sanders rarefaction method expressed as a curve or index that predicts an expected number of species per number of individuals in a sample. Terrestrial biodiversity analysis tends to focus on larger scale measures of diversity based on species richness per area so it can be difficult to compare published marine diversity patterns with those obtained for terrestrial environments.

Pelagos

Functional classification of pelagic organisms is based on locomotion, size, and trophic level or ecosystem function. Plankton are generally passively drifting organisms carried by water movements; examples include bacteria, algae, and small animals. Nekton are actively swimming organisms and so can to a certain extent migrate in the horizontal plane; examples include squid, fish, and some crustaceans. Phytoplankton, marine plants, are primarily unicellular and responsible for primary production. They are the main food source for all life in the open ocean. Zooplankton are animals with a planktonic life style.

Division of pelagic organisms into size classes is related to retention by the different mesh sizes of nets and filters used for sampling (Table 3). Pica and nanoplankton are usually collected with remotely operated opening and closing bottle devices, allowing control of the depth zone sampled. To collect organisms sized from microplankton upward, various types of net sampling gear are used ranging from simple ring nets, which are hauled vertically, to complicated remotely operated opening and closing net systems that can sample horizontal depth layers (e.g., RMT net, Omori net). Nekton are sampled with large, pelagic fish trawls. Often pelagic samplers are used in combination with other equipment that measures physical factors such as depth, temperature, salinity, and light.

Due to the vastness of the pelagic realm relatively little has been studied, although there has been reasonable geographic coverage. It is estimated that less than 5% of its huge volume

Table 3 The size classes of pelagic organisms with typical examples

Planktonic	*Picoplankton*	*< 2 μm*	*Bacteria*
	Nanoplankton	2–20 μm	Bacteria, algae, protists
	Microplankton	20–200 μm	Algae, heterotrophs
	Mesoplankton	0.2–2 mm	Copepods
	Macroplankton	> 2 mm	Shrimps, fish larvae, jellyfish
Nektonic	Micronekton	> 2 cm	Small fish, squid, crustaceans
	Nekton	> 5 cm	Fish, squid, marine mammals

has been sampled once! Deep–sea and very small planktonic species are particularly undersampled; undoubtedly many new species wait to be discovered.

One of the main problems with sampling the benthic and pelagic realm is that scientists are literally grabbing in the dark. In contrary to terrestrial research, where the scientist has a visual concept of the environment and can generally "see" what he or she is doing, marine biologists are completely dependent on remotely operated equipment. This has hampered progress in forming concepts of the marine environment as an ecosystem and estimating its biological diversity. Remotely operated vehicles (ROV's) with cameras and sampling devices have improved our understanding of the marine environment significantly in the past decade. The choice of sampling gear has a large impact on the representativeness of the samples for the community as a whole, hence the frequent discoveries of new species or new information about known species. Giant squids present a nice example of a species that have eluded biologists.

Taxonomy

Estimating, let alone listing, to the number of described species on the earth is a difficult task. For some groups of organisms, usually terrestrial, the actual number of species is reasonably well known with most species named and described; examples include birds, mammals, and higher plants. Other groups have not been comprehensively researched and accurate estimates for species numbers are not available. This category includes beetles, mollusks, and nematode worms. The current best estimate given by the Global Biodiversity Assessment is that there are 1.75 million species known to science. The estimate for the total number of existing species on earth is approximately 12 to 13 million although some scientists think that it may be as high as 20 million.

Classifying and naming the 1.75 million known organisms is an enormous task, but it is vital for comparing and exchanging biodiversity information between different organizations and countries. A profession of approximately 5000 experienced taxonomists worldwide is performing this documentation task. The Kingdom of Plants is the best known and that of Protists least. For the Kingdom Animalia only 20,000 new names are described yearly, including 17,000 species descriptions. The natural history collections stored and maintained in various countries are an important archiving tool for research, and they provide the community at large with a systematic insight into species diversity.

Inventorying marine invertebrates poses exceptional problems. In an area of survey twice as large as the terrestrial environment, marine biology has to deal with problems of the difficult accessibility of the deep sea and the vagueness of the three-dimensional biogeographical borders of the pelagic realm. This makes stock assessments and the definition of distribution ranges extremely problematical.

Species richness is unevenly distributed over the various kingdoms and phyla. The Kingdom Animalia represents the greatest numerical diversity, with well over 1.3 million known species. The terrestrial order of Insecta, in the phylum Arthropoda, takes up about 70% of the known animal diversity.

In both terrestrial and marine domains, the highest diversity is found in the equatorial region. The productivity (in grams carbon per m^2) is more or less equal: 2.2 g C/m^2yr in tropical rain forests and 2.5 g C/m^2yr in coral reefs. In the aquatic environment, current data suggest that coral reefs are the most species-diverse regions and that they are comparable with tropical rain forests. Although reefs only form 0.2% of the seafloor by area, they harbor over 25% of all known marine species. The deep-sea is also believed to be highly diverse. Due to the difficulties in accessing this environment relatively little is known and the discovery of many new species is expected in the coming years (**Table 4**).

It is often surprising for nontaxonomists to discover that most species in the world have never been described and most probably will never be described. This is especially true for the smaller size classes and marine infauna in general. Marine species have received less attention from systematists than terrestrial organisms, less than 15% of described species are marine. It is estimated that there are at least 200,000 known animal species and about 50,000 known plant (algae) species in the oceans. Fish (26,000 species) and marine mammals (250 species) belong to the best known taxa.

Most marine animals live in the deep sea or are tiny and were out of sight and out of mind. Only in the second half of the 20th century has the equipment become available to comprehensively collect deep-sea and small planktonic organisms, and it is still an expensive business limited to a handful of professionals. When this is compared with the long and vigorous tradition of both professional and amateur taxonomic study in the terrestrial domain, the backwardness of marine invertebrate systematics is understandable.

The taxonomically best known marine area in the world is the coastal strip along northwest Europe and to a lesser extent the coasts of the rest of Europe and North America. Outside these zones, less than half the marine invertebrate species are

Table 4 Known and estimated species diversity of various taxa, including Global Biodiversity Assessment (GBA)

	Described species (× 1000)	*Estimates (highest) (× 1000)*	*Estimates (lowest) (× 1000)*	*GBA estimate (× 1000)*	*Accuracy of estimate*
Viruses	4	1.000	50	400	Bad
Eubacteria & Archae	4	3.000	50	1.000	Bad
Fungi	72	2.700	200	1.500	Reasonable
Protozoa	40	200	60	200	Bad
Algae	40	1.000	150	400	Bad
Plants	270	500	300	320	Good
Nematodes	25	100.000	100	400	Bad
Arthropods					
Crustaceans	40	200	75	150	Average
Arachnids	75	1.000	300	750	Average
Insects	950	100.000	2.000	8.000	Average
Mollusks	70	200	100	200	Average
Chordates	45	55	50	50	Good
Others/diverse	115	800	200	250	Average
Total	1.750	111.655	3.635	13.620	Bad

Table 5 The species and family diversity of the major invertebrate macrofauna phyla in box cores from the Bathyal Northwest Atlantic

Taxa	*Number of families*	*Number of species*
Annelids	54	398
Arthropoda	40	185
Mollusca	43	106
Echinodermata	13	39
Nemertina	1	22
Cnidaria	10	19
Sipuncula	3	15
Hemichordata	1	4
Echiura	2	4
Priapulida	1	2
Brachiopoda	1	2
Ectoprocta	1	1

Table 6 Species diversity per m^2 of a mudflat from a British estuary for some of the major meiofauna groups

Taxa	*Number of species*
Polychaetes	5
Nematodes	40
Oligochaetes	2
Copepods	7
Ostracods	2
Hydroids	1
Kinorhynchs	2

known to science. For example, a recent study of 16 box cores taken from bathyal depths off the British coastline revealed 304 polychaete species, of which only 17% could be associated with a known species. A study of 6 liters of coral reef sediment from Hawaii found 158 polychaete species, of which only 30% had been previously described. The largest study of macrofauna ever undertaken in the deep sea in the bathyal, northwest Atlantic counted 798 infaunal species from 233 box cores, but less than half could be named (Table 5). This study illustrates the importance of polychaetes. In general, at least half the macrofauna species of any deep-sea sample will be new to science.

The systematics of meiofauna is even less well known. For example, a nematode study in the Arctic Ocean reported that only 4% of the 92 species found were known to science, while a study in the Venezuela Basin could only name 1.5% of the 136 species found. The taxonomic breakdown of different meiofauna groups in a single region has not been properly explored due to the difficulty of gathering the disparate taxonomic expertise. Some taxa, such as turbellarians, may be important but are commonly ignored in samples because they are so difficult to identify (Table 6).

The identification of marine benthic invertebrates is problematical not just because of lack of taxonomic coverage but because the taxonomic state of the art is inadequate. For many of the soft bodied groups, the species concept (i.e., the criteria used to define a species) is poorly defined. For example, the apparently well-known opportunistic polychaete species *Capitella capitata* and *Streblospio benedecti* have been found to be a complex of genetic sibling species (15 in the case of *C. capitata*).

The lack of species descriptions of small marine animals makes it difficult to assess regional diversity, and this is another factor that encourages marine biologists to consider diversity only over small scales. Only marine megafauna have an adequately censused taxonomy.

Global Diversity

It is impossible accurately to assess the total number of species in the oceans because of limited sampling, the nonrandomness of sampling locations, and inadequate taxonomy. The best estimate has come from Grassle and Maciolek, who sampled benthic macrofauna for 176 km along a bathymetric

gradient 1.2 km to 2.0 km deep in the bathyal zone of the northeast Atlantic. They discovered a rate of species turnover of one new species detected every km. Assuming one new species per km^2, given that there is 10^8 km^2 of deep sea, this implies that there may be up to 10^8 (100,000,000) species in the seas. Such extrapolation techniques are notoriously inaccurate so this figure should be treated with caution but Grassle and Maciolek's calculation is surprisingly similar to the estimate obtained for insect diversity of rain forests by Erwin who used substantially the same technique, suggesting that marine benthic and terrestrial diversity may be similar. A much better knowledge of the geographic ranges of the smaller benthic organisms is essential for an accurate estimation of global marine diversity. This will only be possible when some of the taxonomic problems surrounding these organisms are resolved.

The marine environment certainly demonstrates a greater diversity at higher taxonomic levels than the terrestrial domain. Life evolved in the oceans and many phyla never evolved to survive in the harsher terrestrial environment. It is highly likely that new metazoan phyla remain to be discovered, especially among the small organisms in marine sediments. The free-living phylum Loricifera was only described in 1983 and the commensal phylum, Cycliophora, was first found in 1995 (**Table 7**).

The distribution of benthic invertebrate species richness around the global ocean basins has not been comprehensively assessed. Schopf has shown that both ectoprocts and bivalve mollusks are twice as diverse in the Pacific than the Atlantic. This may reflect the relative sizes of the two oceans, or it may be the product of the relative age of the basins; the Atlantic is younger than the Pacific.

There is still some controversy over whether the highest marine benthic invertebrate diversity is found in coastal regions or the deep sea. Certainly the deep sea is a much larger environment than shallow water, so it might be expected to be home to more species; but it is unclear whether it is more diverse per area sampled. Originally, the low abundance of animals found in the deep sea caused scientists to assume that the highest diversities must occur on the coastal shelf. However, in the 1960s Saunders found the deep sea unexpectedly diverse, indeed with a higher species richness than shallow water. A series of studies appeared to confirm this view but recently an analysis has reopened the debate. Gray carried out a similar study to that of Grassle and Maciolek (noted earlier) but on the Norwegian coastal shelf. He found a similar, or even higher, rate of accumulation of species with sampling on the shelf compared to the deep sea.

Table 7 Comparison of free-living invertebrate diversity of terrestrial and marine environments at a high taxonomic level

Endemic marine phyla	*Endemic terrestrial phyla*	*Cosmopolitan phyla*
Brachiopoda	Onychophora	Annelida
Chaetognatha		Arthropoda
Ctenophora		Chordata
Echinodermata		Cnidaria
Echiura		Ectoprocta
Gnathostomulida		Gastrotricha
Hemichordata		Kamptozoa
Kinoryhncha		Mollusca
Loricifera		Nematoda
Phoronida		Nemertea
Placozoa		Platyhelminthes
Priapula		Porifera
		Rotifera
		Sipuncula
		Tardigrada

Pelagic Diversity Patterns

Van der Spoel and Heijman summarized the general distribution pattern types and biogeographic regions in the open ocean for phyto- and zooplankton species as Arctic, Subarctic, Cool-Temperate, Warm-Temperate, Tropical, Temperate and Subtropical, Subantarctic, and Antarctic. These regions principally reflect seawater temperature and latitude. Contrary to earlier beliefs, relatively few plankton and nekton species have a true cosmopolitan circumglobal distribution. Despite the lack of obvious barriers in the oceans, there is a surprising variety of the types of species distribution. Species specific characteristics and ecosystem relations play an important role in combination with abiotic factors in determining distribution patterns, but many are further modified by regional influences such as ocean basins, currents, and divergence and convergence zones. Typical and easily recognizable biogeographic patterns are belt-shaped patterns related to latitudes (temperature regimes) and neritic distributions (basically around the continental coasts and shallow water areas). The distribution around the Antarctic continent of *Euphausia superba* (the shrimp commonly known as krill) is an example of a neritic distribution (**Table 8**).

Many widely distributed pelagic plankton species show North-South or East-West variation. Without active migration, genetic contacts between populations from one end of the range to the other are limited or even nonexistent, promoting genetic drift. In addition, the differing environmental circumstances (abiotic and biotic) at the ends of each range can give rise to phenotypic differences. This is reflected by both morphological and ecological variation within a species, a first step toward the development of new species. An example of North-South variation is found in the shell shape and size of many species of Pteropoda (pelagic Mollusca). East-West variation can be found for example in *Eucalanus subtenuis* (copepod, Crustacea).

It is expected that over coming decades, knowledge of pelagic biogeography and open ocean distributions will develop exponentially (**Table 9**).

Benthic Large-Scale Diversity Patterns

Latitudinal Gradients

Latitudinal gradients of terrestrial taxa commonly display a decline in species richness from the equator to the poles. Many different processes have been suggested to explain such gradients including competition, predation, mutualism, parasitism, and host diversity, but the most convincing explanation is that the diversity gradient follows the gradient of

Table 8 Pelagic distribution patterns

Type	*Region*	*Example species*
Circum-global	Cosmopolitan	*Stylocheiron maximum*
Belt-shaped patterns	Cosmopolitan	*Rhizosolenia alata*
	(Sub) tropical	*Clausocalanus paupulus*
	Central Waters	*Stylocheiron suhmi*
	Temperate N and S	*Sagita plantonis f zetesios*
	High latitude N	*Calanus glacialis*
	High latitude S	*Thysanoessa macrura*
Central water patterns	Atlantic, Indian, Pacific	*Euphausia brevis*
	Indo-Pacific	*Hydromyles globulosa*
Endemic patterns	Atlantic	*Euphausia americana*
	Indian	*Desmopteris gardineri*
	Pacific	*Sagitta pseudoserratodentata*
Neritic patterns	Cold water north	*Gammarellis homaris*
	Cold water south	*Euphausia crystallorophias*
	Warm water	*Charybdis smithii*
Bathypelagic patterns	Deep water	*Cyclothone pseudopallida*

Table 9 Species diversity in the Atlantic (A), Pacific (P), and Indian (I) Oceans for three example groups: Chaetognatha, Pteropoda, and Euphausiidae (largest diversity is found in the equatorial region)

	Chaetognatha			*Pteropoda*			*Euphausidae*		
	A	*P*	*I*	*A*	*P*	*I*	*A*	*P*	*I*
Arctic	1	1	—	3	4	—	5	5	—
Subarctic	9	6	—	20	17	—	10	10	—
40° N–40° S	25	34	29	120	90	95	25	30	30
Subantarctic	9	8	8	17	17	17	10	10	10
Antarctic	4	4	4	11	11	11	5	5	5
Total	48	53	41	171	139	123	55	60	45

solar energy. There are two potential mechanisms that might cause this relationship. One suggestion is that latitudinal gradients in solar radiation cause concomitant gradients in productivity and hence diversity. The other suggestion is that the relationship between solar energy and diversity is the result of increased evolutionary speed in warmer conditions. The explanations are not mutually exclusive.

The relationship between latitude and diversity for marine benthic coastal fauna could best be described as confusing, with different studies and different taxa yielding conflicting results. For example, if the diversity of core samples is compared using diversity indices, then no significant difference is found between tropical and temperate samples. This experiment has been repeated for both macrofauna by Warwick and colleagues and meiofauna by Boucher and Lambshead.

However, Roy and colleagues were able to test for a latitudinal diversity gradient for prosobranch, gastropod mollusks using data based on number of species per degree of latitude. Unusually for marine studies, this data is similar in scale and format to the terrestrial diversity data used for analyzing large-scale patterns and it is noteworthy that it produces convincing latitudinal diversity gradients (**Figure 3**). Furthermore, the prosobranch diversity gradient follows the solar energy gradient in a similar way to terrestrial latitudinal gradients.

The deep sea is an intriguing location to search for latitudinal gradients because solar energy can have no direct impact on deep-sea diversity. The deep sea is uniformly cold and relatively stable so it can be argued that temperature cannot be a factor in controlling diversity though influencing evolutionary rates. Furthermore, productivity gradients in the deep sea are the result of gradients in the food flux to the seafloor, which vary according to a number of processes.

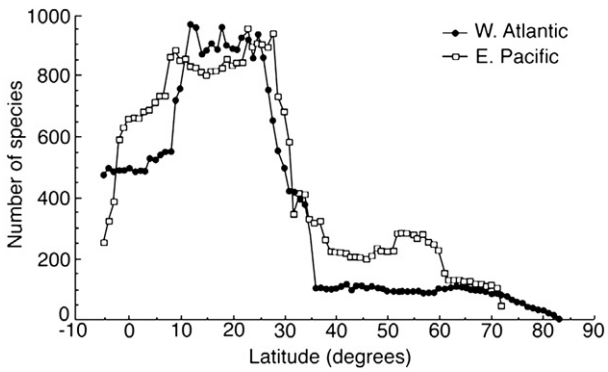


Figure 3 The number of American coastal species of prosobranch gastropods per degree of latitude. On both coastlines the diversity peaks at 10 to 30° N. (Reproduced with permission from Roy K, Jablonski D, Velentine JW, and Rosenberg G (1998) Marine latitudinal diversity gradients; tests of causal hypotheses. *Proc. Natl. Acad. Sci. USA* 3699–3702.)

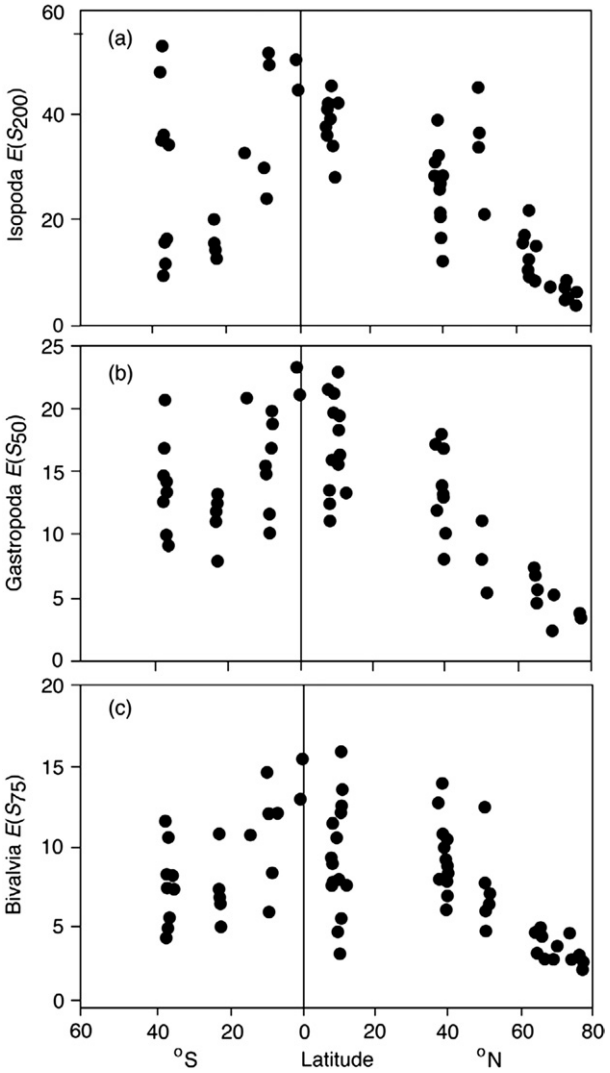


Figure 4 The change in species diversity with latitude for Atlantic deep-sea (a) isopods, (b) gastropods; and (c) bivalves. Rex and colleagues demonstrated that species richness for these taxa falls from the equator northward into the Norwegian Sea. Species richness has been calculated using the Sanders rarefaction index. (Reproduced from Rex, *et al.* 1993, with permission from Nature.)

Rex and colleagues found that deep-sea gastropod mollusks, bivalve mollusks, and isopods showed a decline in diversity from the equator to the Norwegian Sea in the North Atlantic (Figure 4). The explanation is not obvious, as productivity tends to increase northward. Similar gradients have been searched for in the South Atlantic by Brey and colleagues but have not been found, so it is possible that the North Atlantic gradient is due to some individual feature of this ocean, possibly the Quaternary glaciation.

Analysis of the smaller invertebrates by Lambshead and colleagues has shown different patterns. Nematode diversity appears to be correlated with the food flux to the seafloor (i.e., a productivity gradient). In the North Atlantic, there is a weak increase in diversity northward. In the Pacific, nematode diversity declines northward away from the equator, as does the food flux to the seafloor (Figure 5).

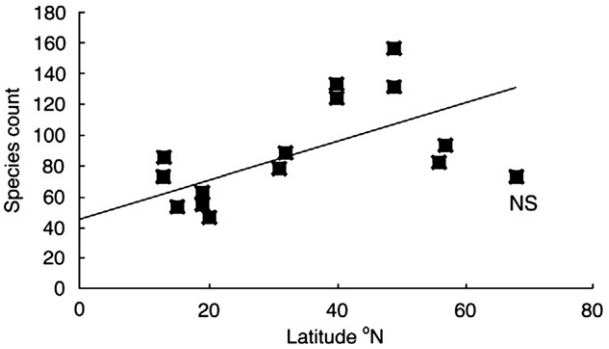


Figure 5 The count of the number of nematode worm species per station in the North Atlantic plotted against latitude. There is a gradual increase in species richness following the increase in food supply from the equator northward. This pattern extends as far north as the Norwegian Sea (marked NS) where there is a marked drop in species richness. The pattern of diversity being positively linked to food input appears widespread for small invertebrates.

Bathymetric Gradients

Bathymetric gradients are probably better understood than any other large-scale diversity pattern for marine benthic invertebrates and for once a consistent pattern is found in all size classes. Rex showed that a number of macrofaunal taxa followed a parabolic diversity curve with depth, diversity peaking at about 2 to 3000 m (Figure 6). A number of studies have since confirmed this finding. A similar pattern was found for megabenthos by Haedrich and colleagues and for nematodes by Boucher and Lambshead (Figure 7).

The exact peak in each bathymetric diversity gradient varies with different taxa and with different locations and, for all we know, there may be temporal variations. However, the basic pattern of diversity peaking in the bathyal or upper abyssal region appears consistent. This consistency suggests that similar processes are responsible for all taxa and locations. Two obvious explanations are that both productivity and disturbance decline with depth. Computer models (e.g. those produced by Huston) have demonstrated how the interaction of productivity and disturbance gradients can theoretically give parabolic diversity patterns.

The exact species composition of the benthic community changes rapidly with depth for some taxa with species having a preferred depth range. Other taxa have wide depth ranges. This has been better studied for the larger organisms than for small organisms due to taxonomic reasons. A wide variety of reasons probably explain the different strategies displayed by different taxa but, in general, the sharpest discontinuity is found at the shelf break with a second boundary at about 1000 and 2000 m. So the basic pattern is of zones of gradual change in species composition separated by areas of rapid change.

Benthic Small-Scale Diversity Patterns

One of the most intriguing observations about marine benthic invertebrates is the high species richness found in individual core samples (Table 10).

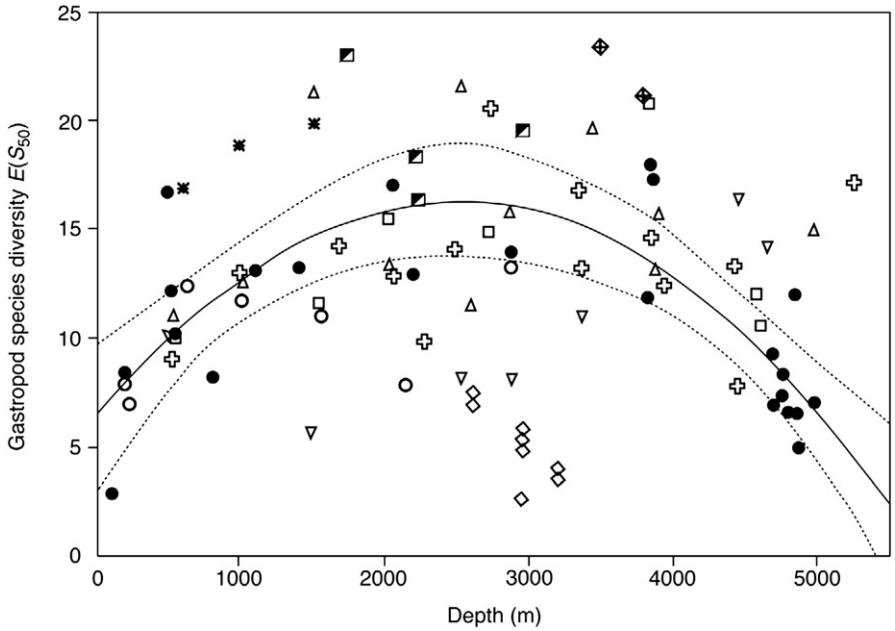


Figure 6 The relationship between gastropod diversity, as measured by the Sanders rarefaction index, and depth in 85 epibenthic sledge samples from various locations in the Atlantic Ocean and Norwegian Sea. Rex and colleagues showed that diversity peaks at bathal depths declining in both shallow and deeper sediments. The pattern is probably caused by the effects of disturbance and food availablity on diversity. Note that the regression lines and 95% confidence limits are for the samples from the North American Basin only. (Reproduced with permission from Ormond RFG, Gage JD, and Angel MV (eds.) (1997) *Marine Biodiversity, Patterns and Processes*. Cambridge: Cambridge University Press.)

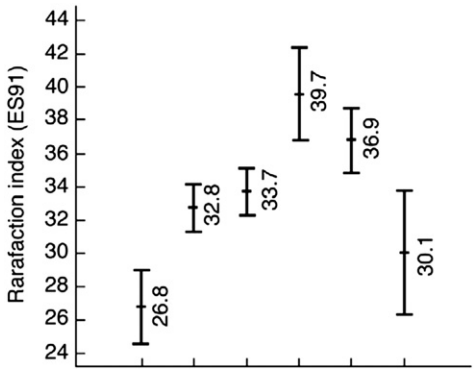


Figure 7 The species richness of marine nematodes in sediments from six biotopes, temperate estuaries, tropical shallow water, temperate shallow water, bathyal depths, abyssal depths, and hadal depths. Species richness is measured using the Sanders rarefaction index. Diversity is highest in bathyal sediments.

Table 10 The species diversity of nematode samples from different depths and habitats from the British Isles westward into the deep sea

Location of sample	*Depth m*	*Species richness*	*Number of individuals studied*
Clyde Estuary	High water, beach	20	272
Clyde Estuary	Mid water, beach	30	1268
Clyde Estuary	Low water, beach	56	549
Irish Sea	56	46	150
Rockall Trough	545	45	104
Rockall Trough	835	50	102
Rockall Trough	1474	56	127
Porcupine Abyssal Plain	4840	71	493

The observation that a large number of similar species can be found together in samples of apparently featureless sediment has always represented a challenge for "equilibrium" explanations for the ecological processes that control diversity. Such explanations depend on the hypothesis that coexisting species must avoid competition by each possessing some unique adaptational specialization. This hypothesis is difficult to accept for such a simple organism as a free-living nematode worm, many of which seem to have similar functional requirements.

An alternative explanation is that large numbers of species can coexist because the sediment is divided into transient patches by biological and physical small-scale disturbance and uneven distribution of the vertical food flux. Grassle and colleagues have coined the phrase *spatio-temporal mosaic* to describe this process (**Figure 8**). A new patch is initiated by a disturbance event, or a food fall, and colonization to exploit the new situation increases diversity. As the patch ages and is exploited, competition increases and diversity drops. But at some point, a new event will recreate a new patch. In this theory, the sediment consists of a heterogeneous mosaic of unique patches of different ages and histories. For any given species, there is a suitable habitat within dispersion distance in

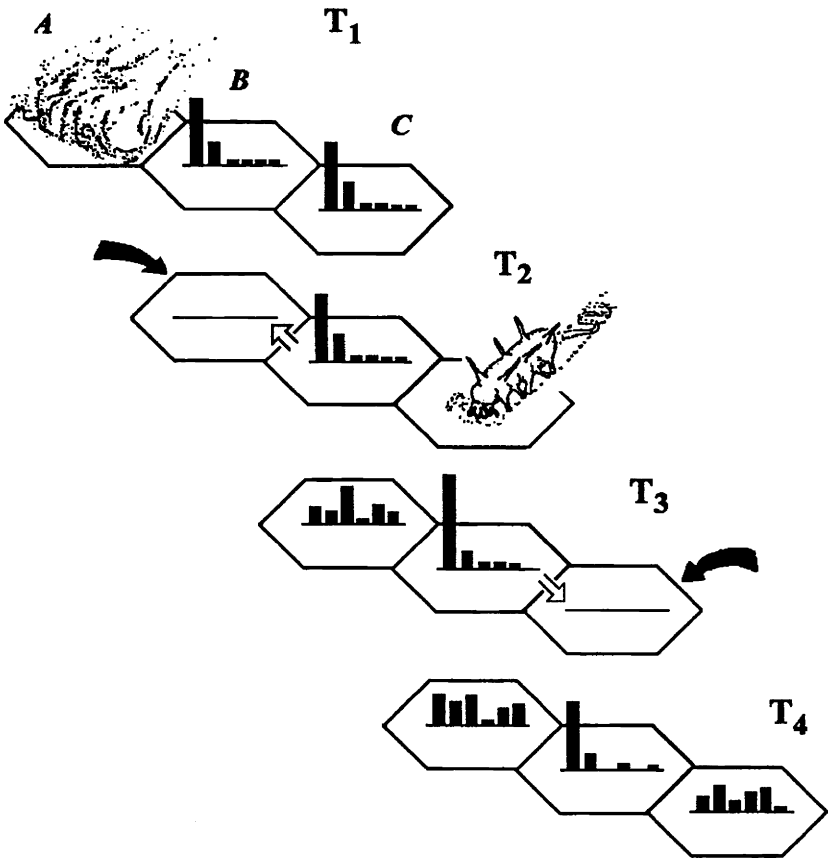


Figure 8 The spatio-temporal mosaic theory of Grassle and colleagues. The concept is that the sediments consist of a series of patches, each of which evolves through an unpredictable cycle. The example shows three patches, A, B, and C through times T_1, T_2, T_3, and T_4. Each vertical bar is a species and the height of the bar represents the abundance of that species. At T_1, patch A is physically disturbed removing all the animals by T_2. The patch is recolonized from the water column (dark arrow) and by migration from adjacent patches (light arrow). By T_3 and T_4, the patch is diverse with six species, five of which are equally abundant. Patch B is undisturbed through time. As the patch matures and is exhausted, the diversity drops to four species through competition. One of these species becomes completely dominant. Initially, patch C is similar to patch A, but it is biologically disturbed by the feeding behavior of a large organism at T_2. This empties the patch allowing recolonization, forming a diverse community. This theory explains how apparently featureless marine sediments can support such a high diversity of life. For any species there is always a suitable new patch available nearby when it exhausts or is outcompeted in its current location. From Paterson.

the mosaic, and when that patch becomes unsuitable, another opens up nearby. The theory depends on limited large-scale disturbance and a lack of barriers for dispersion so it is considered most applicable to the deep sea (and possibly rain forests!).

This spatio-temporal mosaic theory predicts that areas of the deep-sea subject to a flux of phytodetritus would have a more patchy distribution of animals and higher species diversity than other locations. This has been tested for nematode worms by comparing the small-scale spatial distribution and species diversity of nematodes from a station in the phytodetritus-enriched Porcupine Abyssal Plain with a station from the more oligotrophic Madeira Abyssal Plain. Nematodes from the enriched site were more diverse and more species showed aggregation. Interestingly, the species were not aggregated in concordance together around the food or evenly randomly aggregated with respect to each other but showed discordance (i.e., species tended to aggregate in different places). This is consistent with Grassle's theory.

Diversity Over Evolutionary Time

The marine fossil record is more complete than the terrestrial because the ocean floor tends to accumulate sediment. Even so, only certain fauna, notably those with "hard parts," will form an adequate fossil record. The dominant macrofauna group, the polychaetes, presents a limited fossil record, and the dominant meiofauna group, the nematodes, has no record at all. The available fossil record suggests that marine species have an existence of about 4 million years. This indicates on average a 25% species turnover per million years. Assuming that the oceans contain 10^7 (10,000,000) species, we might predict that 2.5 species would become extinct due to natural processes every year. In practice, the fossil record suggests that there are periods of relative quiescence followed by mass extinction and subsequent speciation. The fossil record also implies that 95% of all the marine species that have ever existed are extinct.

The history of marine biodiversity is best tracked at the family rather than the species level to remove some of the

noise associated with an incomplete fossil record. This does tend to make the mass extinction events appear less dramatic. For example, only 54% of marine families were lost in the Permian extinction event, but it is estimated that 77 to 96% of all species became extinct.

The early Cambrian shows a rapid increase in diversity that tends to flatten out in the middle and late Cambrian (**Figure 9**). Most animal phyla appear in the record in this phase. The Cambrian diversity includes a number of "archaic" forms such as trilobites, hyoliths, and inarticulate brachiopods that decline after the Cambrian period. Diversity is not high in the Cambrian and rather unspecialized detritus and low suspension feeding organisms functionally dominate communities, suggesting a simple ecology.

The Ordovician sees a steep rise in diversity but the curve then flattens for some 200 million years. Periods of mass extinction are detectable, notably at the end of the Ordovician and in the late Devonian, but in general diversity stabilized. This new Palaeozoic diversity was associated with an evolutionary radiation of sessile benthic organisms such as crinoids, articulate brachiopods, stenolomate bryozoans, and tabulate and rugose corals.

The Palaeozoic period ended with the catastrophic Permian-Triassic mass extinction. Diversity then apparently rose steadily, with a small extinction at the end of the Cretaceous, past the levels achieved in the Paleozoic until the unique peak of the present day. Current family diversity is apparently twice as high as the Palaeozoic stable level. The Palaeozoic community never seems to have recovered and the modern high diversity is associated with a new fauna of shell-breaking predators and sediment movers. This fauna includes the familiar sea urchins, bivalve and gastropod mollusks, and crustaceans such as shrimps and crabs.

A variety of explanations have been given for this pattern. A number of authors have pointed to the association between the rise in taxonomic diversity and a rise in functional diversity. This almost certainly explains the Ordovician increase in diversity, which is essentially because of increased exploitation of marine sediments. Explanations for the Mesozoic-Cenozoic explosion in diversity are more controversial. One argument is that the breakup of the Pangaea supercontinent caused increased diversity through increased climatic variation leading to more endemicity. It is not clear how this argument would

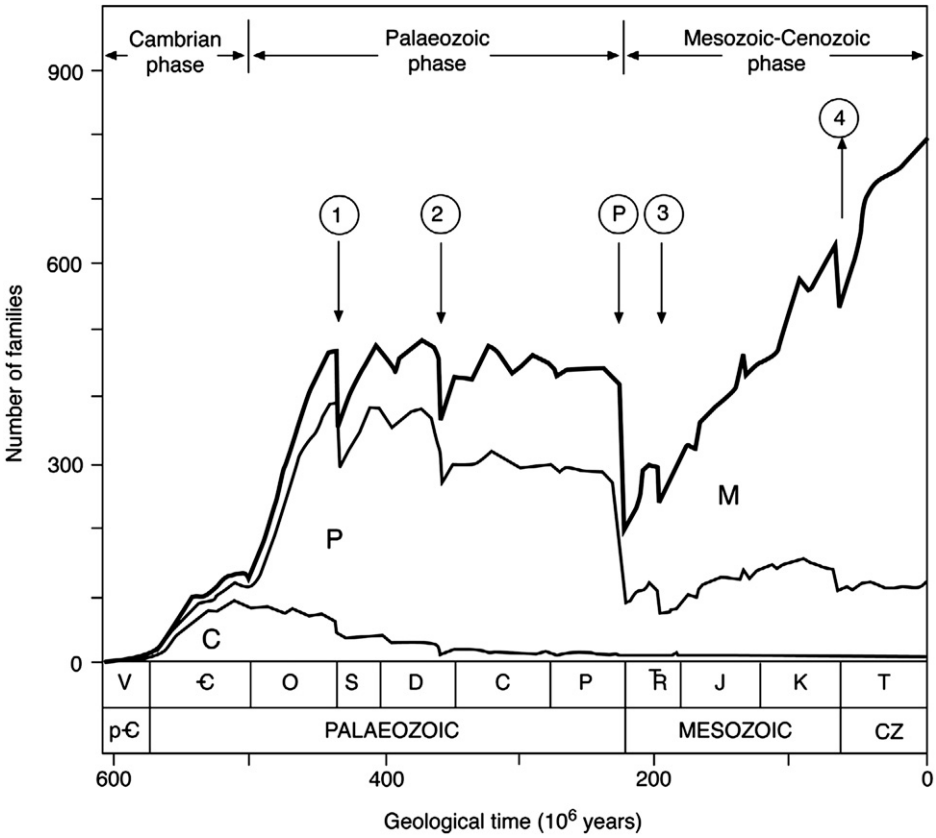


Figure 9 The family diversity of marine animals as measured by their fossil record from 600 M years ago to the present day. Graph C shows the rise and fall of the diversity of the Cambrian fauna of unspecialized detritus and low suspension feeding organisms, P show the rise and fall of the diversity of the Palaeozoic fauna of sessile benthic organisms, and M shows the rise of the diversity of the modern fauna of sediment movers and shell breaking predators. Note that sampling error may have exaggerated modern diversity. The arrows indicate periods of mass extinction, 1 is at the end of the Ordovician, 2 in the late Devonian, 3 at the end of the Triassic, and 4 at the end of the Cretaceous (the dinosaur extinction). P indicates the massive Permian extinction. (Graph by Sepkoski, 1992, reproduced with permission from Heywood VH (ed.) (1995) *Global Biodiversity Assessment*. UNEP, Cambridge: Cambridge University Press.)

apply to the deep sea, where the bulk of marine diversity is to be found. Another suggestion is that evolution in the exploitation of the terrestrial domain has increased nutrient runoff into the oceans concomitantly increasing marine diversity. A similar argument has combined these two explanations suggesting that volcanism associated with continental breakup increased nutrient flux into the oceans.

Finally, Raup has suggested that the high modern marine diversity is an artifact of the fossil record. He observed that across the Mesozoic-Cenozoic time periods there are more younger rock available for study than older rock. If **Figure 9** is adjusted to allow for this, then there is a slower recovery from the Permian mass extinction and modern diversity peaks at much the same level as in the Palaeozoic. If true, it implies that there is a maximum global carrying capacity for marine biodiversity and that the Permian-Triassic extinction was even more devastating than **Figure 9** suggests since the new graph would show that the Palaeozoic benthic fauna has declined consistently since the extinction event. This is possible as the new Mesozoic fauna included many organisms that would tend to disrupt the stable sediment environment on which the sessile Palaeozoic fauna depended.

Hydrothermal Vents and Cold Seeps

The special biological communities around hydrothermal vents were first discovered in 1977 on the Galapogos Rift at a depth of 2500 m. They are now known to be widely distributed in the oceans at tectonically active sites such as subduction and fracture zones, ocean-floor spreading centers, and back-arc basins. Vents are formed where sea water penetrates through fissures in the ocean floor deep into the earth's crust. The water becomes heated and escapes back to the surface through hydrothermal vents. The water temperature at a vent varies from mildly warm, 23 °C, to superheated, 350 °C and can be rich in sulphide and metalliferous ions. The hottest vents are called "smokers" from the precipitation of minerals in the water.

Cold seeps that release sulphide and methane-enriched water have been recorded from a number of areas in the Atlantic and Pacific Oceans. Similar chemical conditions have been found around whale carcasses that are oil rich.

These sulphide-rich habitats are small and ephemeral. The evidence is that many have a life span that measures in decades, although cold seeps may last longer. However, they do appear to be commonplace and may be thought of as chains of islands across the seafloor, separated by distances of 1 to 100 km.

Vent communities have a noticeably different biodiversity from the surrounding ocean floor. The fascinating feature of vent communities is that they derive their energy from chemosynthetic primary production, by reducing compounds such as hydrogen sulphide, rather than photosynthetic sources. It is not clear whether these communities are entirely chemically independent of the photosynthetic world but certainly they are energetically independent.

Biomass is high at hydrothermal vents and primary productivity may be double or triple that of the overlying water. Dense mats of bacteria are found, notably *Thiomicrospira*. A number of invertebrate species feed in dense colonies either directly on these bacteria or by means of symbiotic chemoautotrophic bacteria. Evolutionary selection pressure has been for large, fast-growing species at these productive but ephemeral sites.

Conversely, species diversity is much lower at vents than at other benthic habitats, illustrating once again Rosenzweig's "paradox of enrichment." The dominant species vary from place to place and from depth to depth. For example, the shallow vents off the Palos Verdes Peninsula in California are dominated by the black abalone, *Halioctis cracherodii*; the Mid-Atlantic Ridge vents are characterized by two species of caridean shrimp; and the Eastern Pacific Rise and Galapagos Spreading Centre vents are dominated by the red-plumed, tube dwelling vestimentiferan worms (notably *Riftia pachyptila*) and the large bivalves *Calyptogena magnifica* and *Bathymodiolus thermophilus*. The fauna around the vents are often found to be new to science. Other fauna include a number of polychaetes, crustacea (notably including a number of decapods, and a new primitive genus of barnacle, which is a relic of the Mesozoic).

More than 150 new species, 50 new genera, and 20 new families or subfamilies have been identified with sulphide-rich vents and seeps. Endemism is considered to be high and it is noteworthy that many of the fauna do not appear to be particularly well adapted for high dispersal. Genetic studies on *Bathymodiolus* have shown strong genetic similarity between organisms at sites only 8 km apart but high dissimilarity at sites separated by 2200 km. Vents that are distinctly separated from the chain of sulphide-rich habitats such as the Mariana Trough or Hawaiin volcanic seamounts do show considerable faunistic differences from the norm and the Atlantic sites are less diverse than the Pacific.

See also: Endangered Marine Invertebrates. Invertebrates, Freshwater, Overview. Invertebrates, Terrestrial, Overview. Marine Ecosystems. Pelagic Ecosystems. Plankton, Status and Role of. Vents

References

Butman CA and Carlton JT (eds.) (1995) *Understanding Marine Biodiversity.* Washington, DC: National Academy Press.

Cullen V (ed.) (1995) *Marine Biodiversity I. Oceanus.* Woods Hole: Woods Hole Oceanographic Institution.

Gage JD and Tyler PA (1991) *Deep-Sea Biology.* Cambridge: Cambridge University Press.

Heywood VH (ed.) (1995) *Global Biodiversity Assessment.* UNEP, Cambridge: Cambridge University Press.

Lambshead PJD (1993) Recent developments in marine benthic biodiversity research. *Océanis* 19: 5.

Ormond RFG, Gage JD, and Angel MV (eds.) (1997) *Marine Biodiversity, Patterns and Processes.* Cambridge: Cambridge University Press.

Roy K, Jablonski D, Velentine JW, and Rosenberg G (1998) Marine latitudinal diversity gradients; tests of causal hypotheses. *Proc. Natl. Acad. Sci. USA* 3699–3702.

Spoel, van der S, and Heijman RP (1983) A comparative atlas of zooplankton. Biological Patterns in the Oceans. pp. 1–186. Utrecht: Bunge.

Invertebrates, Terrestrial, Overview

Olof Andrén, Swedish University of Agricultural Sciences, Valsätra-Ultuna, Sweden

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 561–564, © 2001, Elsevier Inc.

Glossary

Annelida Segmented worms, a relatively species-poor phylum.
Arthropoda Joint-legged animals, the most species-rich phylum existing today.
Enchytraeids Segmented worms that are related to but smaller than earthworms.
Invertebrates Animals without a backbone (vertebrae), ranging from unicellular Protozoa to multicellular, complex organisms such as insects.
Phylum Highest level of taxonomic division in the animal kingdom, followed in descending order by class, order, family, genus, and species.

The Terrestrial Invertebrates

Taxonomic Groups, Fully Adapted to Terrestrial Conditions

The truly successful and well-adapted terrestrial invertebrates belong to only two phyla, Arthropoda (joint-legged animals) and Mollusca (snails, etc.). Other taxa, such as Protozoa, Aschelminthes (nematodes, etc.), and Annelida (earthworms, etc.), are limited by their need of an aquatic or at least water-saturated environment, such as in the litter layer and in soil, and a lack of protection against ultraviolet rays. These groups can only move in the true terrestrial environment during nights with favorable weather conditions.

The Insecta are the largest class of animals. About one million species are described, but there are possibly as many as five million species present in the world today. The bulk of the insect species are in the orders Coleoptera (beetles), Lepidoptera (butterflies and moths), Hymenoptera (wasps, ants, bees), and Diptera (flies) (**Table 1** and **Figures 1–3**). Other Arthropoda that successfully cope with the terrestrial environment include the classes Diplopoda (millipedes) and Chilopoda (centipedes) and the similarly built but smaller and less known Symphyla and Pauropoda.

In the class Arachnoidea, scorpions (order Scorpionidea) and spiders (order Araneae) may be the most obvious. However, the smaller mites and ticks (order Acari) are a highly diverse group, containing parasites, predators, fungivores, and detritivores, and are common in almost all litter layers and soils (**Figure 4**).

The crustaceans (phylum Arthropoda, class Crustacea) are mainly aquatic, but the terrestrial isopods (order Isopoda, the "pill-bugs") are quite widespread, even in dry climate zones. However, they lack the water regulation of, for example, insects and spiders and are restricted to moist micro-environments and nocturnal habits.

Among the molluscs (phylum Mollusca) only a few classes have terrestrial capability, and the snails (class Gastropoda) have successfully coped with the dry atmosphere by developing a shell into which they can escape.

Taxonomic Groups with Limited Adaptations

The subkingdom Protozoa contains microscopic animals such as flagellates, ciliates, and amoebae. These are found almost everywhere but are restricted to water films or water bodies.

Table 1 Insect orders, common names of typical members, and approximate numbers of described species

Order	*Typical members*	*Number of species*
Protura	Telsontails	100
Thysanura	Bristletails, silverfish	700
Collembola	Springtails	2000
Ephemeroptera	Mayflies	1500
Odonata	Dragonflies, damselflies	5000
Orthoptera	Grasshoppers, crickets, cockroaches, mantids	23,000
Isoptera	Termites	2000
Plecoptera	Stoneflies	1500
Dermaptera	Earwigs	1100
Embioptera	Webspinners	150
Psocoptera	Psocids	1100
Zoraptera	Zorapterans	20
Mallophaga	Chewing lice	3000
Anoplura	Sucking lice	250
Thysanoptera	Thrips	3200
Hemiptera	Bugs	23,000
Homoptera	Cicadas, aphids, hoppers	32,000
Neuroptera	Dobsonflies, fishflies, snakeflies, antlions	4700
Coleoptera	Beetles	277,000
Strepsiptera	Twisted-wing parasites	300
Mecoptera	Scorpionflies	350
Trichoptera	Caddisflies	4500
Lepidoptera	Butterflies and moths	112,000
Diptera	True flies	85,000
Siphonaptera	Fleas	1100
Hymenoptera	Sawflies, wasps, ants, bees	103,000

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00079-4

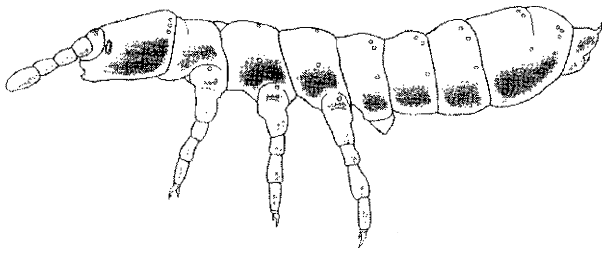


Figure 1 The collembolan (springtail) *Protaphorura* sp. (class Insecta, order Collembola). Reproduced from Coleman DC and Crossley DA Jr. (1996) Fundamentals of Soil Ecology. San Diego: Academic Press.

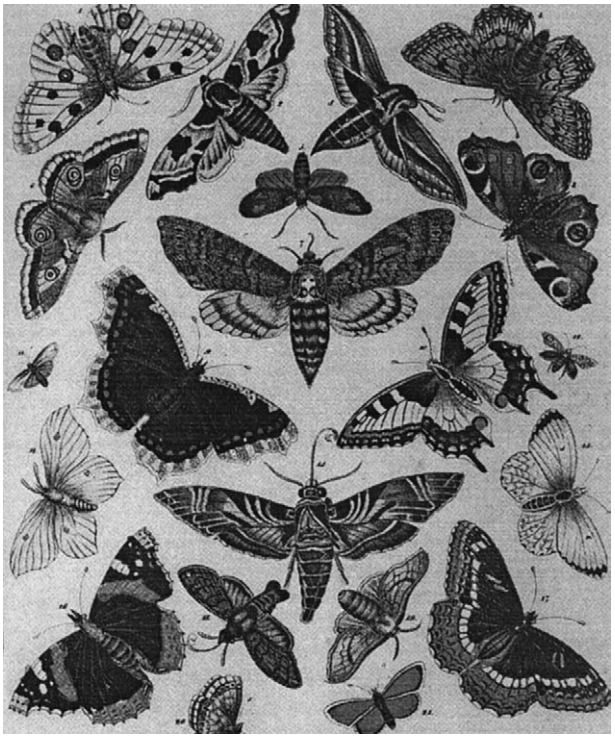


Figure 2 Example of biodiversity among Lepidoptera. (Reproduced from *Bilder ur naturen. En* gåva *till goda barn*. P.A. Norstedt & Söner, 1842, with permission from Dal B (1996) *Sveriges Zoologiska Litteratur*. 1483–1920. Kjuge, Sweden: Orbis Pictus.)

However, several groups have resting stages that can withstand severe desiccation and have the capacity to rapidly reproduce when conditions improve. Thus Protozoa can regulate bacterial biomass and numbers in the soil.

The phylum Platyhelminthes contains primitive flatworms that lack an anus and sometimes even guts. They are mainly parasitic (e.g., tapeworms), but there are free-living species in the class Turbellaria.

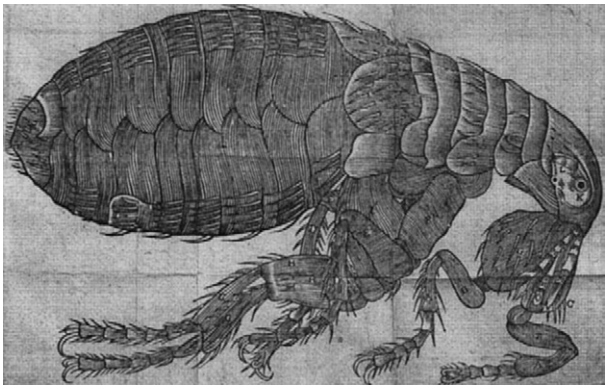


Figure 3 An early micrograph (woodcarving) of a flea (class Insecta, order Siphonaptera). The lettering is not explained in the original. (Reproduced from *Swenska Mercurius*, 1682, with permission from Dal B (1996) *Sveriges Zoologiska Litteratur*. 1483–1920. Kjuge, Sweden: Orbis Pictus.)

Nematodes (phylum Aschelminthes, class Nematoda) are unsegmented, mainly microscopic worms that often have the capacity to form resting stages that can withstand drought and low temperatures. Thus the nematodes have successfully invaded most environments, including arable land (several are pests, but the majority are not) and the inside of other animals (several are internal parasites of mammals, insects, etc.) (**Figure 5**).

Segmented worms are in the phylum Annelida, class Oligochaeta, and contain the well-known "ecological engineers"—the earthworms. Well over 1200 species of earthworms have been described. Lesser known but very abundant in boreal forests and wetlands are the smaller enchytraeids, which are also segmented (**Figure 6**).

The Terrestrial Environment

The terrestrial environment is harsher than marine or freshwater environments. Shortage of water, ultraviolet radiation, rapidly fluctuating temperatures, and a number of obstacles against movement and/or the dissemination of offspring create survival problems, but also opportunities for evolution and speciation. The fact that most of the known animal species are terrestrial is due to the diversity of habitats and a fairly high probability for isolation of populations—a necessary condition for speciation.

Invertebrate biodiversity ranges from the very low in the polar deserts (a few insect and nematode species) to the extremely high diversity found in some rain forest areas, particularly of insects. Note that the high biodiversity in rain forest may be due partly to the high humidity, which reduces the environmental limitations of animal groups that have limited terrestrial adaptations (e.g., slugs). The various ecosystems and habitats that harbor terrestrial invertebrate diversity are described elsewhere in the Encyclopedia.



Figure 4 Ventral and dorsal scanning electron micrographs of the oribatid mite *Carabodes* sp. (class Arachnoidea, order Acari), found in boreal forest litter. (Photo courtesy of T. Persson.)



Figure 5 Micrograph of a free-living rhabditid nematode (class Nematoda, order Rhabditida). (Photo courtesy of J. Lagerlöf.)

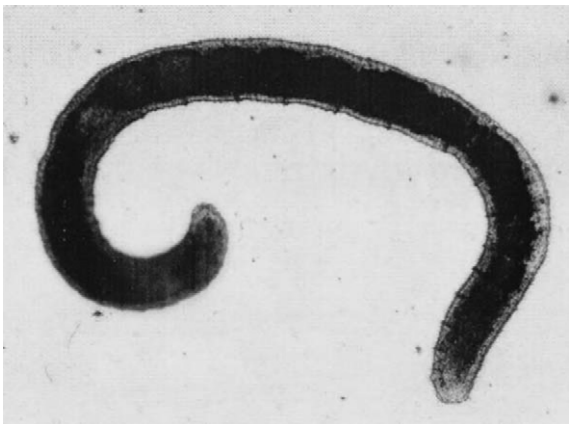


Figure 6 Micrograph of a soil-living enchytraeid worm (phylum Annelida, class Oligochaeta). (Photo courtesy of J. Lagerlöf.)

See also: Invertebrates, Freshwater, Overview. Invertebrates, Marine, Overview

References

Andrén O, Lindberg T, Paustian K, and Rosswall T (eds.) (1990) Ecology of arable land – Organisms, carbon and nitrogen cycling. *Ecol. Bull.* 40. Copenhagen: Munksgaard International.

Coleman DC and Crossley DA Jr. (1996) *Fundamentals of Soil Ecology.* San Diego: Academic Press.

Dal B (1996) *Sveriges Zoologiska Litteratur.* 1483–1920. Kjuge, Sweden: Orbis Pictus.

Dindal D (1990) *Soil Biology Guide.* New York: John Wiley & Sons.

Edwards CA and Lofty JR (1977) *Biology of Earthworms.* London: Chapman & Hall.

Stachowitsch M (1992) *The Invertebrates. An Illustrated Glossary.* New York: John Wiley & Sons.

Island Biogeography

Dieter Mueller-Dombois, University of Hawai'i, Honolulu, HI, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 565–580, © 2001, Elsevier Inc.

Clarification of Concepts

Island Biogeography

Biogeography is a scientific approach to understanding the distribution and abundance of living things, the biota, on our planet. Island biogeographers are primarily interested in isolated areas and the study of fragmented life zones and their relation to the biota. But what living things or biota are included? Nearly all groups are studied: plants, birds, insects, other animals, humans, fungi, fishes, disease organisms, and so on. From this list it is clear that biogeography is not a single discipline. Instead, it is a unifying principle for scientists of different disciplines. The unifying principle is their interest in the distribution and abundance of the organisms with which they have a greater familiarity. Thus, botanists, ornithologists, entomologists, mamologists, mycologists, and anthropologists can all come together and be unified by their interest of biogeography.

Biogeographic Scales

The study of distribution and abundance of biota can be applied at any level of scale in space and time.

Spatial Scales

Traditionally, biogeographers were, and still are, interested in biotic patterns occurring at global and intercontinental scales. Geographically speaking, these are small scales that provide for broad overviews on maps or satellite images. Scientists were concerned with distinguishing and mapping large-area distribution patterns, based on landscape physiognomy, such as tropical rain forests, deserts, savannas, and temperate grasslands, commonly called "biomes." A parallel concern was to distinguish broad patterns of species distribution, which could be mapped as "biotic provinces," areas distinguished as different centers of biodiversity with their "own" floras and faunas. Lately, biogeographers have become concerned with distribution patterns occurring at more detailed spatial scales, such as within-archipelago migration patterns, or the distribution of biota along individual mountain slopes. Moreover, the aspect of abundance versus rareness of species and other taxa has become of concern to biogeographers, an area treated traditionally by ecologists.

Timescales

Present-day distribution and abundance patterns of biota are usually the result of past events and historical processes. Historical processes that led to present-day distribution patterns of biota are often measured at geological timescales in millions of years. Similarly long timescales are considered for biological evolution. But organic evolution by mutations and hybridization can occur at any timescale. Other timescales of biogeographic interest relate to the concepts of succession following disturbances and to phenological change. Successional change in terms of primary succession, which refers to vegetation development on new geological surfaces, can be considered at long timescales involving hundreds and thousands of years. Primary succession can also be related to soil development (= pedogenesis). Secondary succession, defined as following disturbances on already developed soils or previously vegetated substrates, may be considered at plant-demographic timescales, involving a few years, decades, or hundreds of years, if based on the life cycles of certain long-lived tree species. Phenology relates to seasonal changes in biota. The latter two biotic change-with-time concepts—succession and phenology—were developed from ecological research in biogeography. Thus, biogeography as a research approach to discover chronological changes overlaps with those aspects of ecology that deal with distribution and abundance of biota at successional and phenological timescales and also with evolutionary research that seeks to unveil migration patterns and phylogenetic relationships over long and short timescales.

Geobotany

The term *geobotany* is derived from geographical botany, the biogeographical study of plants. In an effort to clarify Central European and Anglo-American terminology in this broad study area, Mueller-Dombois and Ellenberg (1974) synthesized the specializations within the area of geobotany in tabular form, which is reproduced in **Table 1**.

Biogeography was originally understood as consisting primarily of phytogeography and zoogeography. Today, following a review of recent textbooks (e.g., Cox and Moore, 1993; Hengeveld, 1990; Huggett, 1998; and others), one can consider all the ecological disciplines listed under Anglo-American equivalents (**Table 1**) as aspects of plant biogeography.

Three Inspiring Theories

The Theory of Island Biogeography

This theory, originally proposed by MacArthur and Wilson (1963, 1967), is practically synonymous with the concept of "island biogeography." It proposes that species equilibria are formed on islands in relation to the size of land area and its distance from biotic source areas. An equilibrium is suggested where the rate of invasion equals the rate of extinction of island biota. These are the intersection points on **Figure 1**. Curves sloping from the left to the abscissa represent decreasing rates of invasion of biota, from near to far source areas; curves sloping from the zero point on the abscissa up to the right ordinate represent increasing rates of extinction, from large to small islands. The curves are based on a complicated

mathematical model. However, the model is easily understood from **Figure 1**. For example, a large archipelago, such as the Fijian Islands, situated nearer to continental source areas for the dispersion of biota, would have richer species equilibria than a large archipelago, such as the Hawaiian Islands, which is much farther removed from any biotic source area.

The reality of species equilibria is highly questionable, yet the theory has been and continues to be very inspirational. If applied not simply to species per se, but instead to different life forms, such as indigenous trees, ferns, or shrubs, the results may show differing values of species richness and endemism in relation to size of island areas and degrees of isolation. Moreover, these in turn may lead, with additional ecological studies, to a better understanding of the function of biodiversity in different island ecosystems.

At this point, another limitation of this theory should be mentioned. This relates to the fact that size of island area is only a most general predictor of species richness. At least elevational range and substrate type should be added to make the theory more predictable. This brings us to the next theory.

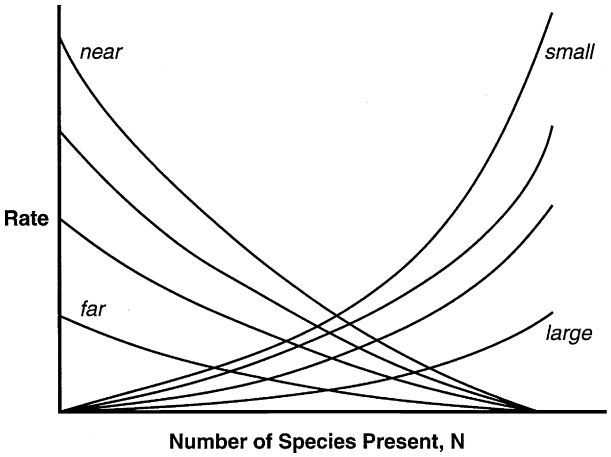


Figure 1 The MacArthur and Wilson (1963, 1967) models of dynamic species equilibria as controlled by distance of islands from biotic source areas (from near to far) and by size of islands (from small to large). Reproduced from MacArthur RH and Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution* 17: 373–387 and MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton, N.J.: Princeton University Press.

The Biome Theory

This theory predicts that there are broad life zones that are indicated by groups of biota of key plant life forms, which are controlled by certain broad-area climatic and edaphic (soil) parameters. For example, the biome theory predicts that in mountainous environments there are altitudinal life zones (Holdridge, 1967) or vegetation zones (Mueller-Dombois and Ellenberg, 1974) that can be distinguished by tree species and other life forms into lowland, upland, and high-altitude zones. Also, such familiar terms as *desert, grassland, deciduous forest, coniferous forest,* and *tundra* depict different latitudinal biome types, which in turn can be defined by climatic parameters as "zonobiomes" (Walter *et al.*, 1975). It is assumed that conditions for life within a biome are more homogenous than life conditions across different biomes (such as grassland versus desert).

Applied to island biogeography, the biome theory lends itself to a more appropriate refinement in the analysis of biodiversity than is offered by the two ecosystem parameters—size and isolation—in the above described theory of island biogeography. For example, comparative biodiversity research within a Pacific-wide biome type, such as the montane rain forest on volcanic high islands of basaltic origin, is scientifically more satisfying than biodiversity research based simply on size of island area. The size approach groups different types of islands into the same category, which thus can be a very

Table 1 Areas of specialization within the field of Geobotany, their synonyms, and Anglo-American equivalents

Area of specialization (and synonyms, European terms)	*Subject matter*	*Anglo-American equivalents (and synonyms)*
Floristic geobotany	Study of geographic distribution of plant taxa and their evolutionary relationships	Plant geography (phytogeography)
Sociological geobotany (vegetation science, plant sociology, phytosociology, phytocoenology)	Study of composition, development, geographic distribution, and environmental relationships of plant communities	Synecology (community ecology, plant ecology in part)
Ecological geobotany (plant ecology)		
Autecology (ecophysiology)	Study of physiological functions of individual organisms in the field environments and communities; life-history studies of species or ecotypes	Autecology (physiological ecology, population ecology in part)
Demecology (population ecology)	Study of structure and function of populations	Population ecology
	Study of genetic variation in populations	Genecology
Synecology (habitat science; ecosystem research)	Study of habitat factors and the physiological response of species and species groups to these factors; study of community functioning, and niche functions of plant populations in an ecosystem context	Ecosystem ecology (community process ecology, functional ecology, systems ecology)
Historical geobotany	Study of historical origins and development of populations and communities	Paleobotany (paleoecology)

Source: Reproduced from Mueller-Dombois D and Ellenberg H (1974) *Aims and Methods of Vegetation Ecology*. New York: John Wiley & Sons.

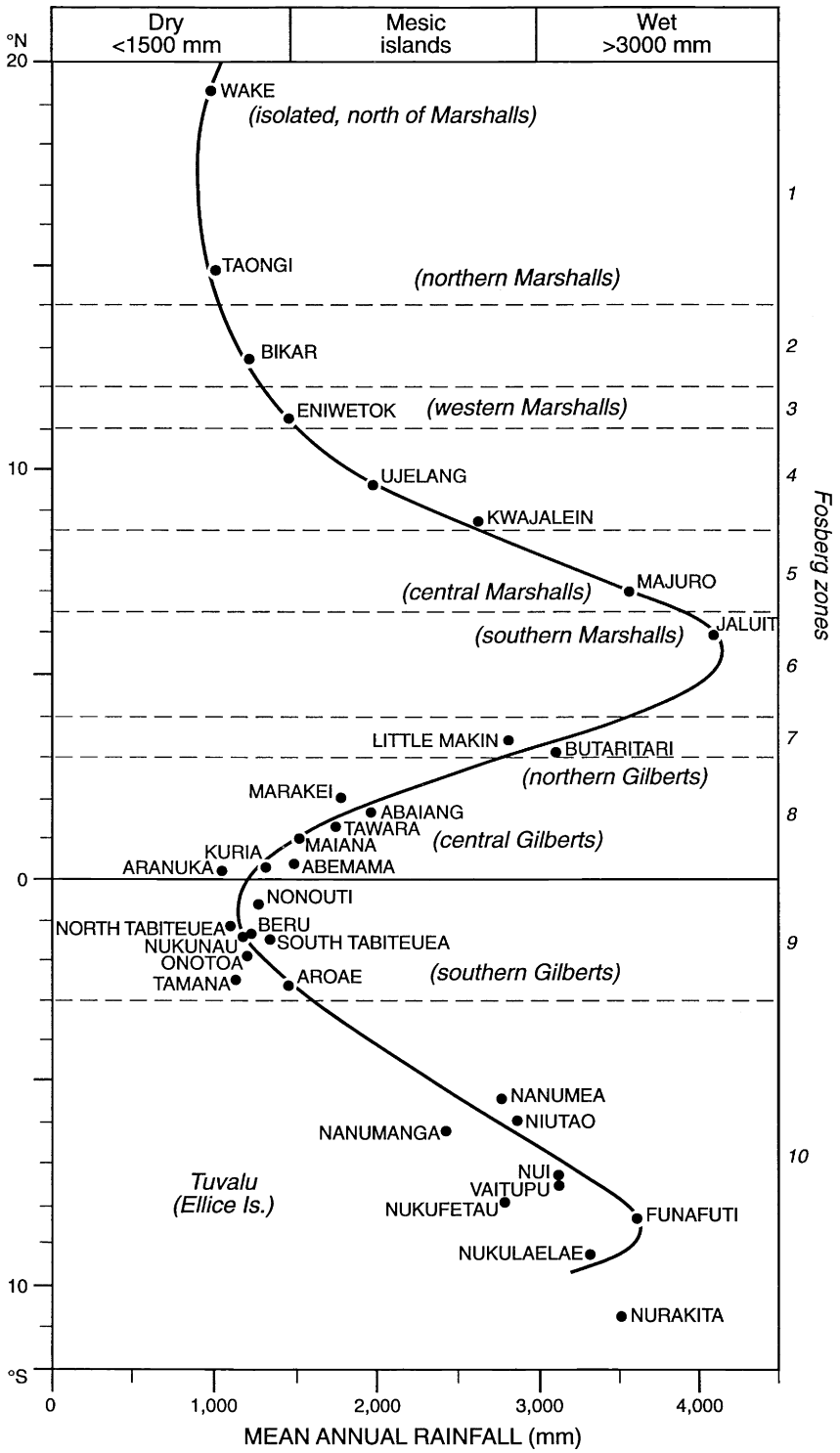


Figure 2 Latitudinal gradients of mean annual rainfall in eastern Micronesia, from the northern Marshalls to the southern Gilberts (Kiribati) and Ellice Islands (Tuvalu), summarized by Fosberg zones. Adapted with modifications from Stoddart DR (1992) Biogeography of the tropical Pacific. *Pac. Sci.* 46: 276–293.

heterogeneous category that includes atolls, raised limestone, and volcanic high islands. Using the biome theory, environmental gradients can be studied among physically similar islands. Examples of environmental gradients for island research of the same biome type are given in **Figure 2**. Atoll and reef islands can be considered as belonging to the same biome, in this case, the same "pedobiome," meaning they have similar, marine-derived, substrates and similar low elevations.

The Theory of Succession

Both the island biogeography and biome theories thus outlined contain elements of succession or community and ecosystem development. MacArthur and Wilson (1967) speak of five fundamental processes as the most difficult to study in biogeography. These are listed as (a) dispersal, (b) invasion, (c) competition, (d) adaptation, and (e) extinction of species. In his concept of succession, the early, influential American ecologist Clements (1916) recognized six processes: (a) nudation, (b) migration, (c) ecesis (establishment by reproduction), (d) competition, (e) reaction (habitat change through organisms), and (f) final stabilization, the climax community. All of these processes are of concern in biogeographic research. For example, the process of "dispersal" and "migration" of species among islands and their "adaptation" in terms of speciation were chosen as the main topics in a recent treatment of Hawaiian biogeography (Wagner and Funk, 1995). Other processes, such as the establishment and regeneration of populations (ecesis) after major disturbances (nudations), their "invasion" relative to "extinction" or their "final" assemblages in communities or ecosystems, also fall into the realm of biogeographical research. Clements proposed a final stabilization, called climax, while MacArthur and Wilson proposed a dynamic (final) species equilibrium as explained earlier. Both, the climax and species equilibrium concepts have been severely criticized. Yet dynamic equilibria remain an area of ecological and biogeographical interest because an understanding of dynamic processes is essential for an improved theory of island biogeography.

In connection with our long-term research on the native Metrosideros forest dieback in Hawaii, I introduced a model of ecosystem development based on the theory of succession (Mueller-Dombois, 1986). This is diagrammatically portrayed in **Figure 3**. The model addresses the concepts of climax as well as those of primary and secondary succession and habitat change with time in a single island biome, the montane tropical rain forest of Hawaii. Here, the process of nudation can be a new pāhoehoe lava flow or a volcanic ash blanket. Both, the volcanic ash and pāhoehoe substrates, may achieve a "climax" in vegetation development in about 1000 to 5000 years in the rain forest climate. After that, a regression phase sets in very slowly, characterized by cation leaching, increasing occlusion of phosphorus (Crews *et al.*, 1995), formation of secondary aluminum (gibbsite) and iron (goethite) minerals, and advanced desilication (Fox *et al.*, 1991). With time, here estimated as 1 million years, the forest undergoes a number of generation turnovers in the form of canopy breakdown or gap formation and recovery. These may be synchronized over larger or smaller areas, depending on the disturbance regime, the cohort structure, and the aging pattern in the forest mosaic. The different kinds of dieback or canopy failure, depicted on the diagram as demographic events, reflect the habitat changes in terms of soil water and nutrient relations that occur over the long timescale. The model also implies that after the plant biomass or biophilic nutrient climax, forest recovery yields successively less tall Metrosideros forests. Associated with the progressive and regressive phases of ecosystem development, different species assemblages also occur. Species equilibria have not been detected among native species. The question of invasive nonnative species is a separate major problem of island biogeography that is discussed in another chapter of this book.

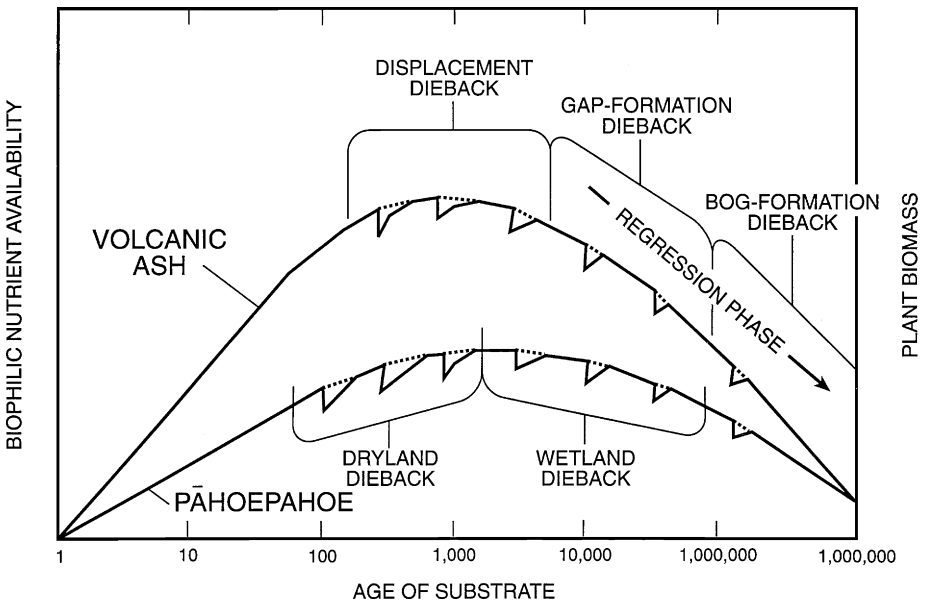


Figure 3 A model of primary succession and regression in the Hawaiian rain forest biome with superimposed secondary successions as caused by Metrosideros canopy failures. These are related to demographic growth cycles and the climatic disturbance regime. A "climax" is shown by the maximum in plant biomass and biophilic nutrient availability. After this climax, indicated mainly by the height of Metrosideros canopy trees, the regression phase is caused by leaching of cations, occlusion of phosphorus, and complete desilication. Reproduced from Mueller-Dombois D (1986) Perspectives for an etiology of stand-level dieback. *Ann. Rev. Ecol. Syst.* 17: 221–243.

Vegetation Analysis Methods

Floristic Checklist Methods

"Floras" are lists of plant species. When more elaborate, they are books with species descriptions, keys, and illustrations for plant identification. Floras usually give little information on vegetation, which is the most visible component in terrestrial ecosystems, there performing the function of primary producer. "Vegetation" is the plant cover of an area in which species play different roles in terms of their abundance or rareness and in terms of their life forms, life histories, and physiology. In nature, plants often interact with one another in communities. "Plant communities" can be distinguished by differences in life-form structure—such as forests versus shrub- or grassland—or by differences in species assemblages. In the study of island biogeography, floristic checklist methods can be applied at the level of biogeographic provinces, which may serve for floristic comparisons of archipelagoes. They can also be applied among individual islands of the same archipelago. At both these broad levels of scale, tests of similarity versus dissimilarity can be applied from floristic records or subregional floras where these exist. However, field research is needed if one wants to compare the floras within the same biome on different archipelagoes, because this information is usually not available. A simple "walk-through method" can provide this information without formalized sampling designs, such as plots or transects. Such broad floristic surveys also help in familiarizing with the territory and, when used together with aerial photographs and topographic maps, they can result in vegetation maps outlining tentative communities. However, floristic comparisons among communities of the same biome in an archipelago, such as the rain forest across the high Hawaiian islands, require a more formalized checklist method in representative sample plots, such as described by the relevé method (Mueller-Dombois and Ellenberg, 1974, discussed later).

Vegetation Sampling Methods

A number of field methods are available. Their proper use depends on the purpose of sampling and the type of vegetation. Here, I will consider plant biodiversity as the general objective of sampling.

The Relevé Method

This is a widely applied and proven floristic checklist method, ideal for many, if not most, plant biodiversity surveys. It is based on plots, called relevés (meaning abstracts), whose sizes are based on the "minimal area concept."

Determining Relevé Size

A minimal area for vegetation sampling can be established by the "nested plot technique." This technique consists of lining out a small quadrat, 0.5×0.5 m, and then enumerating all species encountered in this small quadrat. Next, the quadrat size is doubled into a rectangle of 0.5×1 m, and additional species are noted. The next doubling is a 1×1 m quadrat, followed by 2 m^2, then 4,8,16,32,64 m^2, and so on. At each enlargement, the new species encountered are added. When plotting the cumulative number of species over the area sample, a species/area curve is obtained. In a homogenous vegetation segment, this usually takes the form of a steeply rising curve that levels off with each enlarged quadrat. The result is a curve of "diminishing returns" in terms of number of species encountered with still greater enlargement of area. The area over which the curve levels can be used as the "minimal area." The usual practice is to use a slightly larger area for the size of a relevé. In temperate grasslands, a 10×10 m quadrat is often sufficient to satisfy the minimal area requirement. In temperate forests, 20×20 m quadrats may give a satisfactory relevé size. In multispecies tropical rain forests, 1 ha (10,000 m^2) may not be large enough (see **Figure 4**). But 1ha is now often considered a practical standard for biodiversity research of tropical rain forests.

Recording Species in Layers

A relevé record should contain all plant species found in its boundaries. In forest communities, recording is best done in horizontal layers by vertically defined height strata. For example, one may distinguish two tree layers T1 as trees over 10 m tall, T2 as trees from 5 to 10 m tall (subcanopy trees), then a shrub layer from 2 to 5 m (including smaller trees), then a lower shrub layer from 1 to 2 m high, and further a herb layer from 0 to 1 m high, including smaller woody shrubs and trees (usually seedlings). The vertical stratification is an aid in recording. It also can document if canopy trees are regenerating in the same sample stand (or relevé).

Distributing Relevés in Predefined Strata

Relevés are best laid out in predefined strata. On mountain slopes, these may be predefined elevational strata, such as 200 m contour intervals or the like. In other areas, these are the tentative vegetation segments or communities identified on air photos in conjunction with field reconnaissance and check listing. We have called this process "entitation" (Mueller-Dombois and Ellenberg, 1974), which means defining of entities or map units from air photographs. Depending on the scale and homogeneity of the initial strata or entities, from 2 to 5 relevés are often used to characterize a stratum or entity. Also, depending on the size and complexity of plant biodiversity in an area, from 20 to 140 relevés can amount to a satisfactory biodiversity survey in terms of plant species and plant communities.

Estimating Species Quantities

The relevé method as described here is merely a formalized floristic checklist method, which reports the presence (and absence) of species. For classifying plant communities it is desirable to record how abundant or rare a species is in an area. This can be done simply by adding an abundance symbol to each species on the relevé record sheet. The most widely used rating system in European vegetation surveys is the Braun-Blanquet cover/abundance scale consisting of seven symbols as follows:

5 Any number of individuals, with cover more than 3/4 of the reference area (>75%)
4 Any number, with 1/2 to 3/4 cover (50–75%)

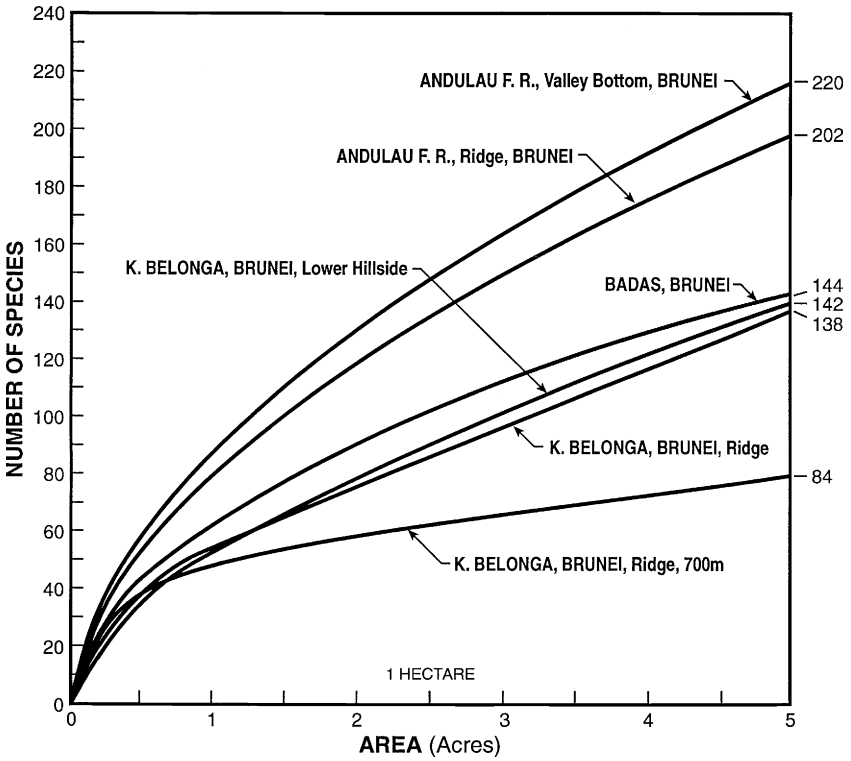


Figure 4 Species/area curves for sites with tropical rain forests in Brunei. Redrafted with modifications from Ashton PS (1964) Ecological studies in mixed dipterocarp forests of Brunei State Oxford Forestry Memoirs, no. 25.

3 Any number, with 1/4 to 1/2 cover (25–50%)
2 Any number, with 1/20 to 1/4 cover (5–25%)
1 Numerous, but less than 1/20 cover, or scattered, with cover up to 1/20 (5%)
+ Pronounced cross) few, with small cover
r Solitary, with small cover

This estimation scale is applied by walking diagonally through the relevé, several times if necessary. It is a crude scale and thus is often criticized, but it takes little time and tells a lot more about the plant composition of a relevé than a mere presence/absence record.

However, estimation methods are often not satisfactory in more complex vegetation. In forest vegetation it may be usefully applied only to the undergrowth species and only in relevés that are small enough to allow such estimation with confidence.

Quantitative Methods

In forest vegetation it is not possible to get a satisfactory estimate of the quantity of tree species by the cover/abundance scaling method as described earlier. In such more complex vegetation, at least the tree component needs to be measured to obtain a satisfactory abundance or cover estimate. The reason for this is primarily that relevé sizes are too small for obtaining an adequate density or cover estimate of trees.

The Count-Plot Method

To get an adequate estimate of tree density per unit area, such as a hectare, one should count trees in many small plots, such as 3×5 m. The small size of such count plots is merely a convenience for an accurate (100%) enumeration per plot. Larger plots may lead to accidental double counts or omissions, thereby introducing serious errors. Count plots are conveniently laid out along transects. They must be repeated until the cumulative mean number of trees does not fluctuate anymore radically with each additional count plot. This may be accomplished by counting at least 30 to 50 trees/species. It is obvious that rare tree species can hardly be adequately counted by this method. Rare tree species can only be reasonable assessed by searching an area of known size for such species, thereby attempting a 100% survey of their presence. In counting trees as a measure of tree density per unit area, it is necessary for plant biodiversity surveys to count trees by species and by measuring some tree-size parameters. Among tree-size parameters, usually the diameter at breast height (Dbh) is measured. From this, tree basal area is obtained by the area formula $r^2\pi$. However, the true basal area is the stem cover at ground level. Where this parameter is desired, one needs to measure also tree diameters at ground level. The Dbh measure is usually preferred in timber volume surveys. Tree volume is established by combining the basal area with a height measure and a tree form factor. More elaborate formulas can be developed for total tree biomass, often sought in productivity studies.

Distance Methods

An alternative to the count-plot method are the distance methods. They often lead more rapidly to an adequate count of trees, since they do not require the layout of plot boundaries. Instead, the average distance between trees is used to determine the mean area per tree and thereby the density per

unit area. The mean area per tree is obtained by squaring the mean distance. This method works well under certain restrictions. Two methods are briefly outlined.

The Point-Centered Quarter Method

A random starting point is established in the forest to be measured. Four distances are measured from this point to the nearest tree in four quarters, forming each a 90° exclusion angle over the point. The sampling point therefore is used like a rectangular cross, and the distance from the point to the nearest tree is measured in each quarter. When repeated at 20 sampling points, established at random or along a compass line, one obtains 80 distances. The mean distance squared then gives an estimate of the mean area per tree. The number of trees per acre ($=4000\ m^2$) or hectare ($10{,}000\ m^2$) is obtained by dividing such reference area by the mean area/tree. The point-centered quarter method is relatively simple to apply, and it compares well to the count-plot method. However, two restrictions apply: the trees should be randomly distributed, and no tree should be measured twice. The second requirement is usually easy to follow as it requires that the sampling points are far enough apart so that no tree is measured twice.

The Wandering Quarter Method

This second recommended distance method is often less affected by the distribution pattern of trees (random versus clumped) as it traverses for the same number of distance measures through a larger area. The method begins with establishing a random sampling point in the forest segment to be sampled. A 90° exclusion angle is established in a certain compass direction, and the nearest tree to the sampling point is measured. That nearest tree then becomes the next sampling point over which a 90° exclusion angle is established in the same compass direction. From this tree the distance is measured to the next tree in that exclusion angle. The procedure is continued for 20 distances in that same direction, moving from tree to tree and always using the same 90° exclusion angle. This results in a wandering movement or zigzag line depending on where the next tree occurs within the 90° exclusion angle. After 20 distances, the direction may be changed following a right angle for another 20 distances, then once more at a right angle in a parallel reverse direction to the first zigzag line, and finally, 20 distances in direction of the starting point. In this case, the sampling also includes 80 distances, but the area traversed is larger than in the point-centered quarter method.

Measures Obtained in Count-Plot and Distance Methods

The two types of methods provide for the same three parameters desired often in quantitative vegetation surveys. These are frequency, density, and cover.

"Frequency" is simply the occurrence of a species in any number of plots in relation to the total number of plots in the sample. In the distance methods, frequency is the occurrence of a species at a sample point out of the total number of sample points. Frequency is a mixed measure of abundance and dispersion or distribution. It is a relative measure as it depends on the plot size used in sampling.

"Density" is the actual count of individuals per species in the total number of count plots. The sample plot area is then adjusted to a standard reference area, such as a hectare. In the distance methods, the number of individuals is obtained from squaring the mean distance of all trees and dividing this measure into the reference area. The number of individuals per species is then obtained from the proportion of species among the total number of individuals.

"Cover" as stem cover per species is obtained from totaling the individual basal area measurements in the sample of count plots or points and by relating this measure to the reference area. For examples, see Mueller-Dombois and Ellenberg (1974).

Another aspect relates to the tree-size stratification in count-plot and distance sampling. In addition to sampling mature trees/species, it is ecologically desirable in biodiversity inventories to also establish the density of seedlings and saplings. First, mature trees have to be defined. This is often done by setting height and/or diameter limits. "Mature trees," for example, may be all individuals 5 m and taller, thereby forming a broad size class. "Saplings" may be defined as individuals from 0.5 to 5 m tall and "seedlings" as those from 0.1 to 0.5 m tall. Smaller individuals may be considered as "germinants" (i.e. not yet well established or ephemeral individuals). Smaller subplots are often required for seedling counts because of their smaller size and frequently greater density. In the distance methods, usually only mature trees are measured because of the random pattern requirement for accuracy. In such cases, seedlings and saplings may be counted in subplots of appropriate sizes at predetermined sampling points.

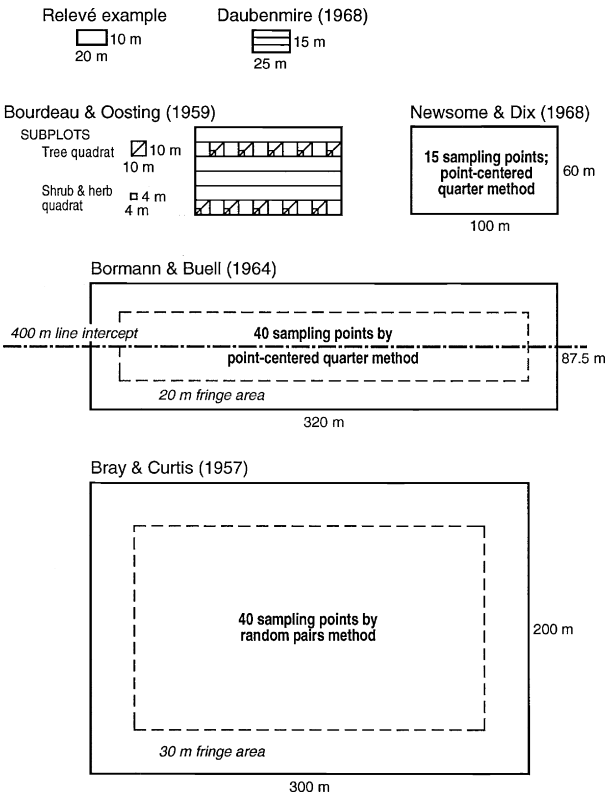


Figure 5 Comparative sizes of sample plots. Reproduced from Mueller-Dombois D and Ellenberg H (1974) *Aims and Methods of Vegetation Ecology*. New York: John Wiley & Sons.

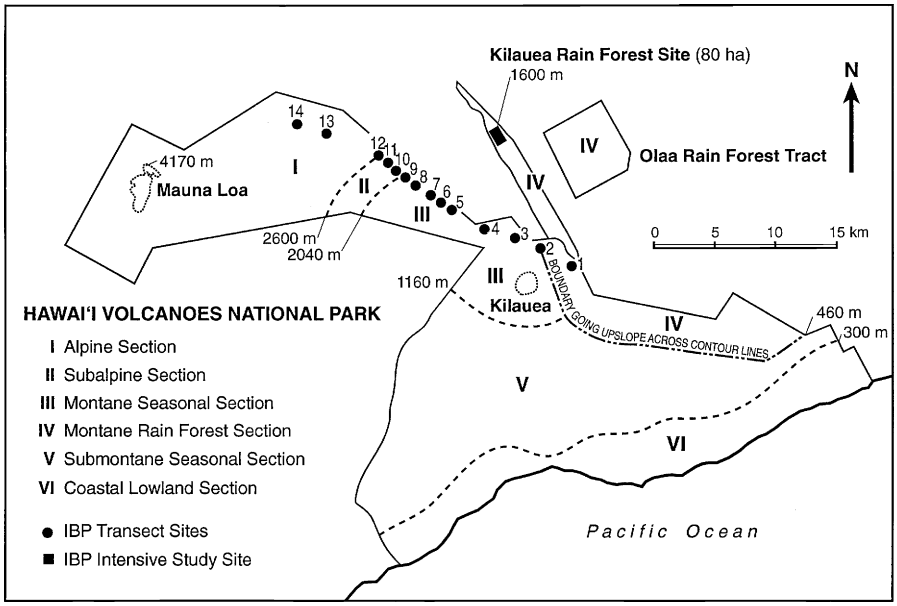


Figure 6 Map of Hawaii Volcanoes National Park with the 14 Mauna Loa transect sites and the 80 ha Kilauea forest site used for integrated biodiversity sampling during the Hawaii International Biological Program (IBP). Reproduced from Mueller-Dombois D, Bridges KW, and Carson HL (eds.) (1981) *Island Ecosystems: Biological Organization in Selected Hawaiian Communities*, U.S./IBP Synthesis Series 15. Stroudsburg, PA: Hutchinson-Ross.

Table 2 Extract of a final two-way table

IBP Focal site no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Relevé no.	001 054 444	001 454 925	001 154 807	001 124 906	001 290 160	011 900 234	011 801 385	011 212 471	011 822 135	011 723 790	011 733 523	000 117 142	000001 466663 434568	000111 667344 790902
Species group 2								*		* *	* *	* * *	* * * * *	* *
PELTER									+	+	+ +	1 1 1	1 + 1 1 1 1	+ 1
ASP ADI							R	+ R		+ +	+	1 1 1	R R R	+ R
ASP TRI								R	+ R	R	+	1 1 +	+ R R R	R
AGR SAN			R							R R	1	+	1 1 + +	+ + 1
								*		* *	* *	* * *	* * * * *	* *
Species group 1		*	* *							* * *	* * *	* * *		
MAC GAH		+	1 1 1							1 1 1	1 1 1	1 + 1		R 1
COP ERN		3	+ + 1						2	1 2 2	1 2 2	2 2 +		
LUZ HAW		1	1 2		+				+ 1	1 1 1	+ 1	R +		
RAI CIL		R	1							+ + 1	+ 2	1 1 R		
COP MON										1 1 2	1 2 2	2 2 2		
PLE THU			+	R						+ +	1 + +	1 +		
POL PEL		+ +	1 2							+ +	1 +			
		*	* *							* * *	* * *	* * *		
Species group 3			* *		* * *	* * *	* * *	* * *	* * *						
HOL LAN		+	+ R	3 2 2	2 1 1	1 2 1	1 1 1	1	1 2 2					
CAR WAH		+	R +		2 + 1	1 1	1 2 2	2 1 1	1 1 1		1	+		
ACA KOA			2	2 2	5 4 3	3 3 3	4 4 3	5 4 3	3 2					
STY TAM		2	3 3 3		2	1 3	2 2 2	1 2 3	3 2					
PAN TEN					1 1 1	2 1 1	1 2 2	2 1 2	2 2					
			* *		* * *	* * *	* * *	* * *	* * *					
Species group 4		*	* * *		* * *	* * *	* * *	* * *	* * *	* * *	* * *	* * *		
DES AUS		1	+ 2		2 1 1	3 1 1	3 2 2	4 2 2	2 2 2	1 1 1	1 1 1	1 1 1		
PTE AQU		1	2 2	2 3 2	2 2 2	1 2 1	1 +	1 + +	1 1 2	1 1	1			
DOD SAN		+	1 2 2			+ 2	1 R +			2 2 2	2 2	2 1 1		
		*	* * *		* * *	* * *	* * *	* * *	* * *	* * *	* * *	* * *		

Note: Rules used: 50% inside, 10% outside.

Application of appropriate sampling methods requires some basic understanding of sample theory and field experience. When the sampling objective is clarified and the nature of the vegetation experienced, it is usually possible to prescribe an appropriate approach from a combination of established field analysis methods.

Size of Sample Plots

Figure 5 displays six sample plot sizes that have been used in proven studies. They are drawn to the same scale. Like the small, 20×10 m grassland relevé, all plots were placed in predetermined entities or tentative map units to obtain representative data. Daubenmire's (1968) preferred plot size of 15×25 m was applied in the mountainous interior forests of the Pacific Northwest. He preferred to assess the undergrowth vegetation within such plots along 25 m transects by determining frequency out of 50 systematically placed 0.1 m^2 frames. Percent cover in terms of shoot cover was also estimated similarly as with the Braun-Blanquet cover/abundance scale. But 50 placements only gave an area sample of 5 m^2, which likely did not fulfill the minimal area requirement as used in the relevé method. The larger plots in **Figure 5** refer to quantitative studies as discussed earlier. Their sizes are based on the principle that tree species need to be counted in numbers that give reliable density and basal area estimates. For obtaining a minimum tree count of 30 to 50 trees/species, larger plot areas are needed, since individual mature trees can occupy mean areas of 25 to 100 m^2 or larger.

The larger sample plots with quantification of tree counts obviously require a much greater time effort than a relevé with a species record supplied with a cover/abundance estimate. Thus, one has to be very clear about the purpose of the field effort. The relevé method is particularly useful in plant biodiversity studies aiming at classifying vegetation by species groups, while quantitative methods are essential if the objective is repeated monitoring of plant biodiversity in permanent plots.

Data Processing and Display Methods

Data Analysis: An Island Example

During the 1970s, we did a multidisciplinary study in Hawaii under the auspices of the International Biological Program (IBP). Our main objective was to study the biological organization in relatively undisturbed natural communities. We focused on two areas, an upper montane rain forest (the Kilauea Rain Forest) and the east-slope of Mauna Loa in Hawaii Volcanoes National Park (**Figure 6**). In the Kilauea Rain Forest, which was dominated by three native keystone

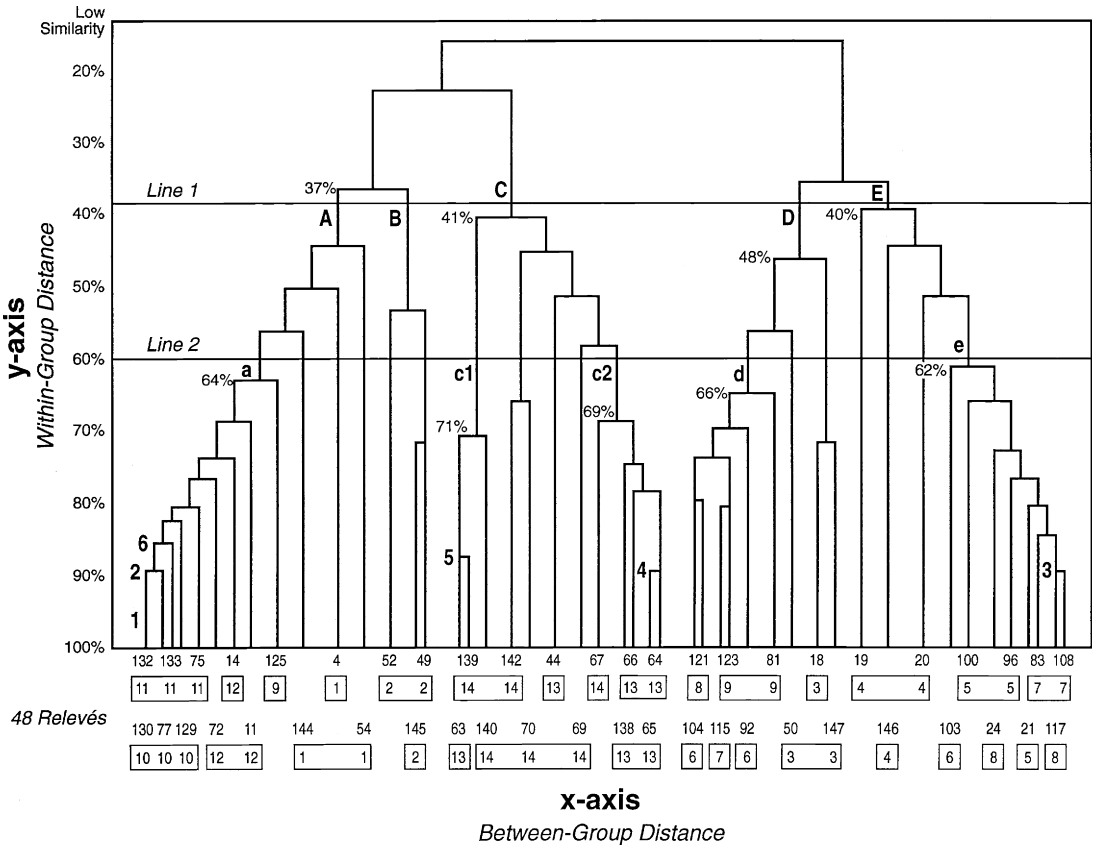


Figure 7 Dendrograph for cluster analysis based on 48 relevés sorted by the quantitative Sørensen index of similarity. Blocked out numbers are IBP transect sites 1–14. Numbers not blocked out are relevé numbers. Reproduced from Mueller-Dombois D, Bridges KW, and Carson HL (eds.) (1981) *Island Ecosystems: Biological Organization in Selected Hawaiian Communities*, U.S./IBP Synthesis Series 15. Stroudsburg, PA: Hutchinson-Ross.

species, Acacia koa, Metrosideros polymorpha, and tree ferns (Cibotium glaucum), we established a large sample plot, 1000 × 800 m = 80 ha greater by 20 ha than the largest plot on **Figure 5**. Our data sets were processed by species in life form and guild categories to get quickly from forest community structure to function by niche differentiation. This aspect is dealt with in part III of our synthesis volume (Mueller-Dombois, Bridges, and Carson, 1981), and will not be elaborated here.

On the east slope of Mauna Loa, we tried to answer a basic question asked by MacArthur (1972, p. 161), "Do different plant species change synchronously, or does each have independent distribution?" This question arose from a dispute over the nature of species distribution within and across neighboring vegetation zones or biomes. To clarify this question, we established 14 transect sampling sites (1–14 on **Figure 6**) in predefined altitudinal intervals from 1100 to 3080 m. These sites were studied for 12 other organism groups

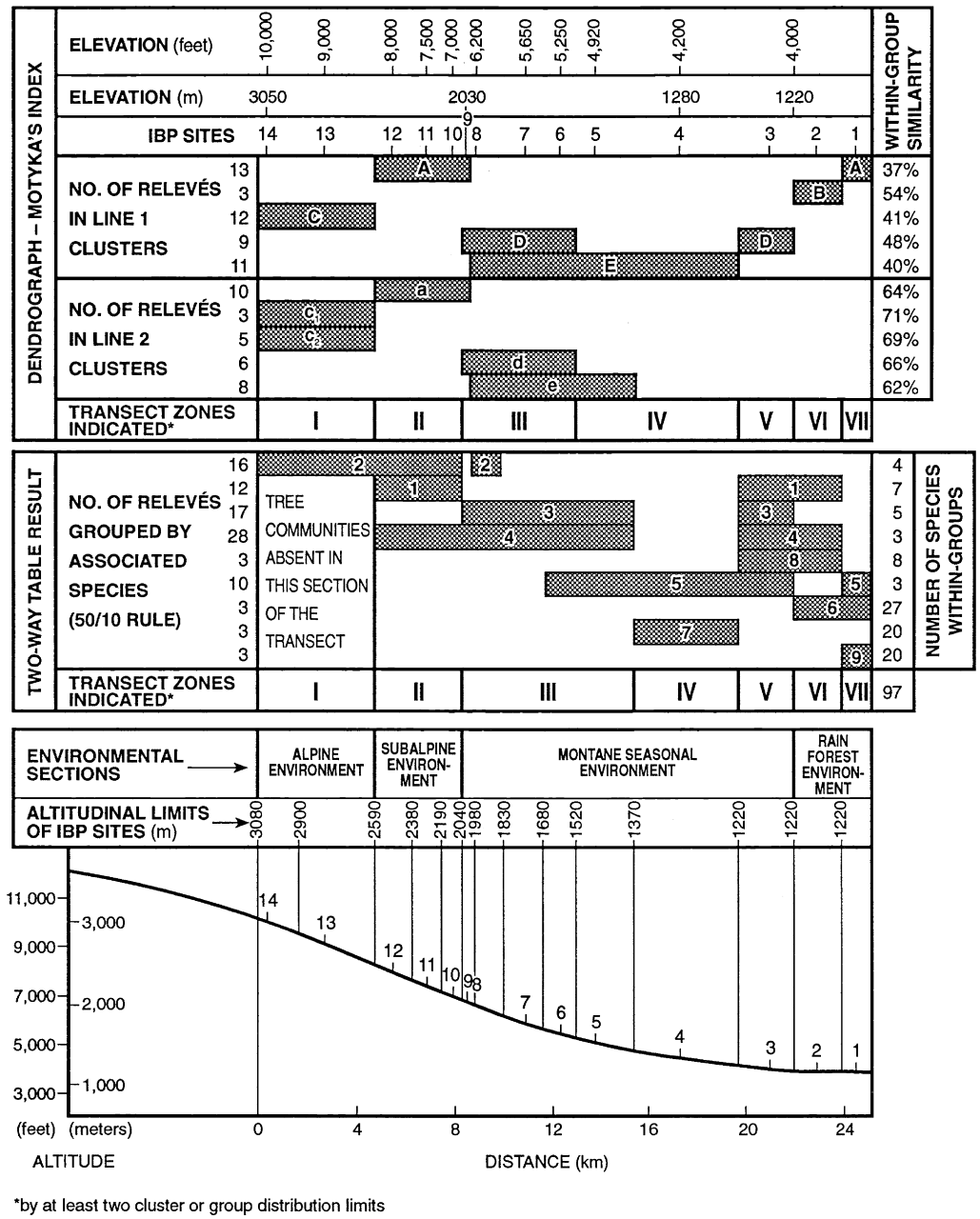


Figure 8 Altitudinal vegetation zones on the east slope of Mauna Loa derived from computer processing of 48 relevés by dendrograph and two-way table techniques. Zone I = Sparse alpine heath scrub, II = Subalpine Metrosideros scrub forest, III = Mountain parkland with Acacia koa colonies, IV = Sapindus-Acacia koa savanna, V = Open Metrosideros dry forest, VI = Open Metrosideros rain forest, VII = Closed Metrosideros rain forest. Reproduced from Mueller-Dombois D, Bridges KW, and Carson HL (eds.) (1981) *Island Ecosystems: Biological Organization in Selected Hawaiian Communities*, U.S./IBP Synthesis Series 15. Stroudsburg, PA: Hutchinson-Ross.

besides plants. For the plant distribution study, we established 48 relevés throughout this mountain transect, approximately three relevé clustered around each of the 14 transect sites.

Two-Way Table and Dendograph Techniques

Following the field study, the 48 relevés were processed by the "two-way table technique." That required entering all relevés into a single "raw table," whereby the relevé number appears at the head of the table and all species names appear at the left side of the table. The species magnitude, or score values, are entered for each species in the respective "relevé columns." When all data are entered in this way, one can see in the "species rows" which species are present (and with what score value) or absent in the "relevé columns." The process then begins of sorting species of similar distribution together by repositioning the "species rows," then also the "relevé columns" are resorted to bring those relevés together in the table that contain species of similar distribution. The "row and column sorting process" is reiterated until the table displays an optimal structure in the sense that one can interpret the clustering of species and relevé group in an ecologically meaningful way. This data processing, formerly done by hand, can now be done mathematically by using one of many multivariate analysis techniques and similarity indices.

An example of a two-way table extract of our 48 Mauna Loa relevés is given in **Table 2**. For corroboration of species distribution and relevé clustering trends one can use various mathematical ordination methods or data classification methods. Ordination methods (not illustrated here) present the same data geometrically, while classification methods present the same data in the form of dendrographs (see **Figure 7**). The three data processing and display techniques are corroborative analytical tools. Particularly the ordination and dendrograph techniques are rather abstract and difficult to interpret by the reader.

Improved Display Methods

It is more "reader friendly" to go a step further in displaying the results of a multivariate data analysis as shown in **Figure 8**. This data display over the Mauna Loa transect with its 14 IBP transect sites clarifies that both multivariate analyses, the two-way table and the dendrograph techniques, gave closely similar results. Both data treatments resulted in the definition of seven altitudinal vegetation zones as named in the legend of **Figure 8**.

The answer to MacArthur's (1972) question is given in **Figure 9**. It displays the distribution of 17 key species, which are primarily responsible for the pattern of the seven vegetation zones identified by multivariate techniques. Several species clearly change synchronously, while others have independent distributions along this island mountain slope. The answer to MacArthur's question is this: There is more than one pattern of plant species distribution along this geologically young tropical mountain slope. Some distributions coincide closely with one another; others do not. However, the species distribution data give clear evidence of the existence of altitudinal vegetation zones as established from mathematically independent multivariate analysis techniques.

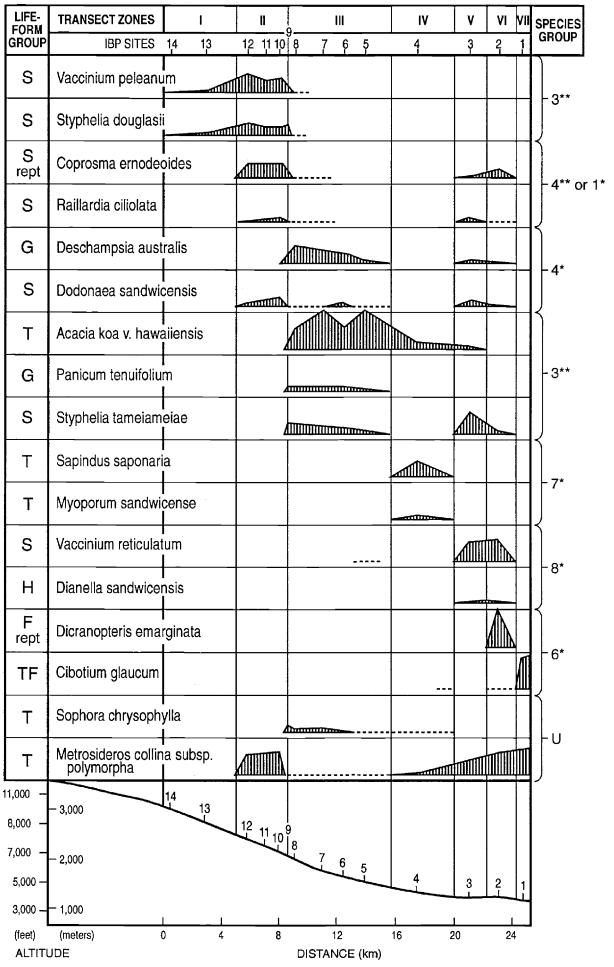


Figure 9 Distribution and abundance of native key species along the Mauna Loa transect. Species groups derived from two-way table analysis. Computer derived group numbers shown at right; U = ungrouped species. Symbols at left are, S = shrub, rept = reptate or creeping, G = grass, T = tree, H = herbaceous plant other than grass, F = fern, TF = tree fern. Reproduced from Mueller-Dombois D, Bridges KW, and Carson HL (eds.) (1981) *Island Ecosystems: Biological Organization in Selected Hawaiian Communities*, U.S./IBP Synthesis Series 15. Stroudsburg, PA: Hutchinson-Ross.

An Experimental Sampling Design For Island Biogeography Research

Transects are generally the most efficient sampling designs in ecological field research. They can often be arranged to cut through a maximum of variation in representative habitats and biodiversity assemblages of an area to be studied, using the shortest distance. A still more important aspect is the use of transects to control environmental variation in its effect on biodiversity. In this way, transects can be used as "experimental sampling designs" by aligning them along changing environmental control factors to study their effects on biodiversity in environmental settings that are relatively uniform or constant. Sampling transects along a mountain slope, on which mean annual temperatures and rainfall change in predictable ways, are good examples. Another example is given

in Figure 2, in which the western Pacific atoll islands are aligned along mean annual rainfall gradients in both the northern and southern hemispheres.

Horizontal within Ecosystem Gradients

In addition to the atoll islands there are several other Pacific-wide biomes. A particularly important one is the upland rain forest on the volcanic islands of the tropical Pacific. This forest is important for two primary reasons: it is still extant on many volcanic high islands as indigenous forest harboring most of the endemic species, and it performs the function as watershed cover in the mountainous interiors of these islands. The volcanic high islands typically are small, isolated land fragments that protrude as mountainous terrain above the vast Pacific ocean surface. Their interior upland forests likewise are fragments that belong to the same larger ecosystem or biome. However, from archipelago to archipelago, these biome fragments are occupied by different sets of species due to their past biogeographic isolation. Connecting these biome fragments horizontally across the ocean by a system of transects provides for a within-biome research design. Here, the broad habitat features are kept uniform, while the indigenous biodiversity sets change from island to island. This truly is a research approach to biodiversity, which Pielou (1979) defined as "geoecology," the study of recurrence of similar communities in similar habitats, which are occupied by different sets of species. The term *geoecology* is an abbreviation of geographical ecology in the conceptual sense used by MacArthur (1972), who contributed substantially to the foundation of island biogeography.

Vertical Between Ecosystem Gradients

A functioning watershed cover is an essential resource component in all Pacific high islands. From here, the fresh water flow begins to be regulated and then influences almost all lower lying island ecosystems, the freshwater wetlands, mangroves, estuaries, fish ponds, fringing reefs, the entire coastal zone, and agriculturally used lowland areas. The ecosystem services of the upland forest watersheds have not been studied in the Pacific islands, in spite of the fact that they have been part of the traditional land-use system in the Pacific high islands, in Hawaii known as the ahupua'a system. This vertically arranged land-use system was recognized in the Hawaiian Islands as comprising four integrated management zones, "wao la'au," upland forest or wilderness area, to be left alone, "kula," the more open foot hill region, "wao kanaka," the agricultural zone in the lowlands, and "kahakai," the coastal zone. This vertically arranged multi-ecosystem human support system proved to be an optimal land-management scheme for the indigenous islanders in the past. Vertically arranged transect sites, using the fresh water flow and hydrology as unifying parameter, may also prove to be a good organizing principle for interdisciplinary research focused on the function of biodiversity at the landscape level.

The PABITRA Initiative

The acronym PABITRA refers to the Pacific-Asia Biodiversity Transect outlined on Figure 10. The transect system connects the high island archipelagoes with indigenous tropical rain forests still extant in their interior uplands. They range

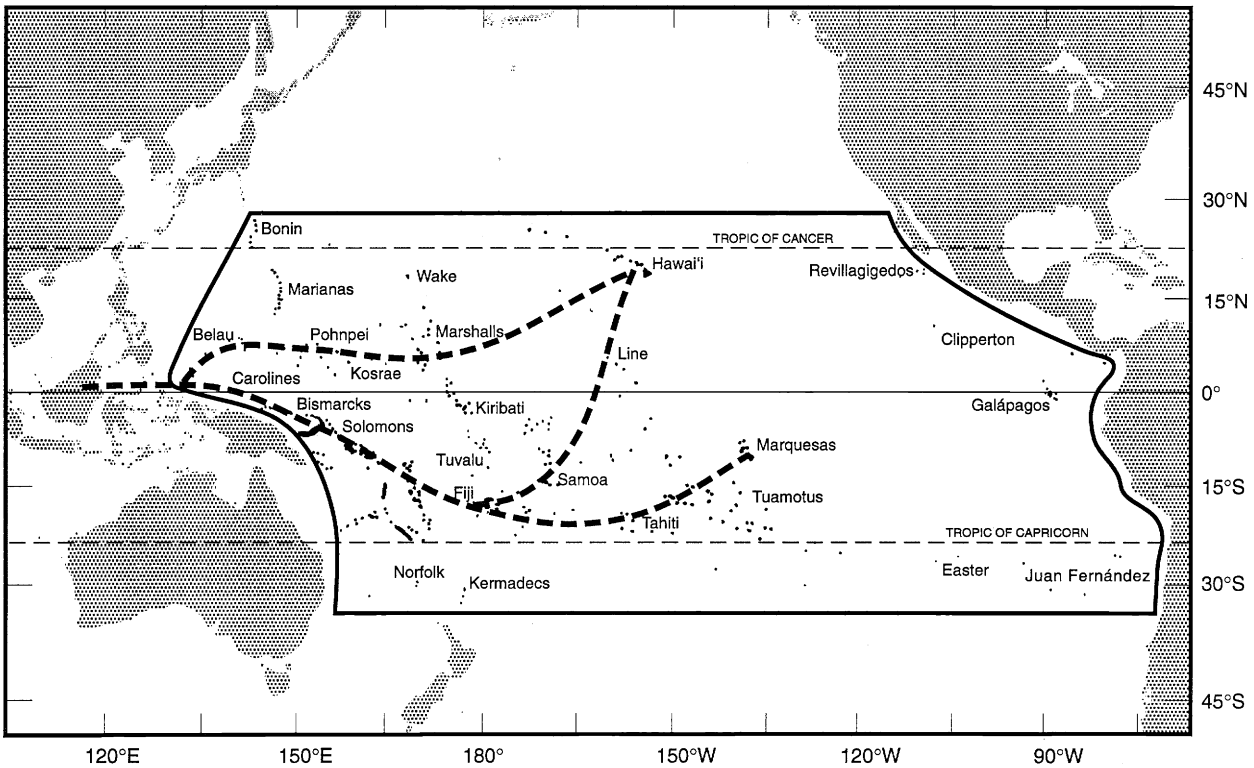


Figure 10 The PABITRA (Pacific-Asia Biodiversity Transect) network as currently being initiated. For more information see text.

from east to west, from the paleotropical outlier islands, the Hawaiian archipelago and the Marquesas, to the biotically rich continental islands of southeast Asia. The PABITRA concept arose from the concluding chapter of a recent book (Mueller-Dombois and Fosberg, 1998) with the chapter heading "The Future of Island Vegetation." Simply put, the indigenous island vegetation only has a future if the scientific community and conservation managers redouble their efforts in learning to understand the function of island biodiversity and thereby take appropriate conservation measures. Another impetus came from a Japanese research initiative, called DIWPA, Diversitas in the Western Pacific and Asia. This program is the Pacific-Asia representative of the new worldwide program "DIVERSITAS," promoted by the International Union of Biological Sciences (IUBS) with support from UNESCO and ICSU (The International Council of Scientific Unions in Paris, France). The promotion of DIVERSITAS started effectively with a week-long forum titled "Biodiversity, Science and Development: Towards a New Partnership" in September 1994 (Younès, 1996). Since then, DIWPA has focused on a "Green Belt," a north-south transect in western Asia for monitoring biodiversity of forests and freshwater lake systems at 42 transect sites. A parallel running "Blue Belt" is also planned with focus on coastal zones and nearshore marine habitats. The PABITRA Initiative is considered a separate, but attached, "Pacific Island Branch of DIWPA," working cooperatively under the Ecology, Conservation, and Environmental Protection (ECEP) Division of the Task Force on Biodiversity in the Pacific Science Association, which is chaired by the writer.

See also: Biogeography, Overview. Dispersal Biogeography. Species–Area Relationships. Succession, Phenomenon of

References

Ashton PS (1964) Ecological studies in mixed dipterocarp forests of Brunei State Oxford Forestry Memoirs, no. 25.

Borman FH and Buell MF (1964) Old-age stand of hemlock-northern hardwood forest in central Vermont. *Bull. Torrey Bot. Club* 91(6): 451–465.

Bourdeau PF and Oosting HJ (1959) The maritime live oak forest in North Carolina. *Ecology* 40(1): 148–152.

Bray JR and Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* 27: 325–349.

Clements FE (1916) Plant succession. An analysis of the development of vegetation. Washington, D.C.: Carnegie Institute.

Cox CB and Moore PD (1993) *Biogeography: An Ecological and Evolutionary Approach*, 5th ed. Oxford: Blackwell Scientific Publications.

Crews TE, Kitayama K, Fownes JH, *et al.* (1995) Changes in soil phosphorus fractions and ecosystem dynamics across a long chronosequence in Hawaii. *Ecology* 76(5): 1403–1424.

Daubenmire RF (1968) *Plant Communities: A Text Book of Plant Synecology.* New York: Evanston & London, Harper & Row.

Fox RL, de la Pena RS, Gavenda RT, *et al.* (1991) Amelioration, revegetation, and subsequent soil formation in denuded bauxitic materials. *Allertonia* 6(2): 128–184.

Hengeveld R (1990) *Dynamic Biogeography.* Cambridge: Cambridge University Press.

Holdridge LR (1967) Classification and characterization of tropical forest vegetation. *Recent Adv. Trop. Ecol.* II: 502–507.

Huggett RJ (1998) *Fundamentals of Biogeography.* London, New York: Routledge.

MacArthur RH (1972) *Geographical Ecology: Patterns in the Distribution of Species.* New York: Harper and Row.

MacArthur RH and Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution* 17: 373–387.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Princeton, N.J.: Princeton University Press.

Mueller-Dombois D (1986) Perspectives for an etiology of stand-level dieback. *Ann. Rev. Ecol. Syst.* 17: 221–243.

Mueller-Dombois D, Bridges KW, and Carson HL (eds.) (1981) *Island Ecosystems: Biological Organization in Selected Hawaiian Communities*, U. S./IBP Synthesis Series 15. Stroudsburg, PA: Hutchinson-Ross.

Mueller-Dombois D and Ellenberg H (1974) *Aims and Methods of Vegetation Ecology.* New York: John Wiley & Sons.

Mueller-Dombois D and Fosberg FR (1998) Vegetation of the Tropical Pacific islands. *Ecol. Studies* Vol. 132. New York: Springer-Verlag.

Newsome RD and Dix RL (1968) The forests of the Cypress Hills, Alberta and Saskatchewan, Canada. *Amer. Midland Naturalist* 80(1): 118–185.

Pielou EC (1979) *Biogeography.* New York: John Wiley & Sons.

Stoddart DR (1992) Biogeography of the tropical Pacific. *Pac. Sci.* 46: 276–293.

Wagner WL and Funk VA (eds.) (1995) *Hawaiian Biogeography: Evolution on a Hot Spot Archipelago.* Washington, D.C.: Smithsonian Institution Press.

Walter H, Harnickell E, and Mueller-Dombois D (1975) Climate diagram maps of the individual continents and the climatic regions of the earth. (Supplement to Vegetation Monographs). 36pp. + 8 maps of world regions at 1:8 mill. And one world climate map at 1:30 mill. Berlin, Heidelberg, New York: Springer-Verlag.

Younés T (1996) Biodiversity Science: Issues and challenges. In: Turner IM, Diong CH, Shirley E, Lim L, and Ng PKL (eds.) *Biodiversity and the Dynamics of Ecosystems*, pp. 1–12. Kyoto, Japan: International Network of DIVERSITAS in Western Pacific and Asia (DIWPA), Center for Ecological Research, Kyoto University. Shimosakamoto 4-1-23, Otsu 52001, Kyoto, Japan. DIWPA Series Vol. 1.

Isoptera

Takuya Abe[†] and Masahiko Higashi[†], Kyoto University, Kyoto, Japan

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 581–611,
© 2001, Elsevier Inc.

Glossary

Alate An imago still possessing its wings.
Caste A group of individuals in a colony that are both morphologically distinct and specialized in behavior.
Imago The adult insects; the final developmental stage when insects possess wings.
Larva An immature individual without an external sign of wing buds or soldier morphology.
Neotenic A secondary reproductive with juvenile morphological characters. Neotenics derive from larvae, nymphs, pseudergates, or workers through at least one special moult.
Nymph An immature individual on the developmental pathway to the imago and which possesses external wing buds.
Presoldier A transitional morph that always precedes the soldiers; an unsclerotized individual whose head shows signs of soldier differentiation.
Primary reproductives Dealate reproductives that founded a new colony after nuptial flight.
Pseudergate (false worker) A temporarily nonreproductive individual serving the colony in nutrition, construction, or brood care, which results from a late, reversible deviation from the pathways to the imago and is characterized by reduced wing buds compared to nymphs of the same stage.
Queens and kings Females and males that actively reproduce in a colony.
Secondary reproductives Reproductives that differentiated in an established colony, whatever their origin and morphology. They may be supplementary reproductives if older reproductives are still present or replacement reproductives if not.
Soldier An individual with a strongly sclerotized head showing defensive adaptations, such as enlarged mandibles, a stopperlike shape, or a frontal gland able to produce a defensive secretion.
Worker An individual resulting from an early, irreversible deviation from the pathway to the imago, and performing helper tasks. Workers are primarily characterized by the loss of the ability to proceed to the winged imago, but they need not be permanently sterile.

Introduction

The Isoptera is a relatively small order of insects that consist of 31 orders, including about 2650 described species. The order name Isoptera originates from the fact that the imago (adult) has wings of equal size (*Isos* in Greek meaning the same and *ptera* meaning wings). The insects of Isoptera are commonly called termites in English. Due to their poor abdominal sclerotization and white appearance, they are often called "white ants" in many temperate regions. However, they are not phylogenetically closely related to ants, but to cockroaches and mantises. Termites are sometimes black, reddish brown, or yellow in body coloration in tropical regions.

Termites are widely distributed in tropical and sub-tropical regions, spreading from humid forests to savannas and even arid areas. Their biomass approaches 10 g wet weight/m^2. Only a few animal groups such as human beings, earthworms, herbivorous mammals in African savanna, and ants claim this order of biomass. The basis of this extraordinary prosperity of termites lies in their feeding habits. They consume cellulose, which is the most abundant organic matter on the earth. The ecological basis of their prosperity is their symbiosis with microorganisms and their highly developed social organization.

The symbiosis produces novel abilities that cannot be attained by either of the partners alone. In particular, the symbiosis between higher organisms and microorganisms often creates capabilities for exploiting new food resources such as dead plant material containing large amounts of cellulose, which is a potential energy source but difficult for most animals including human beings to utilize. On the other hand, sociality enhances the efficiency of resource exploitation. Either of these or both together may cause a rapid growth and spatial expansion of the population, leading to diversification.

Animals in the tropical rain forests where the species diversity is the highest in the world are characterized by the dominance of social insects (termites, ants, social wasps, and social bees), and the dominance of symbiotic associations. Termites possess both of those attributes. Termites and ants occupy about one-third of total animal biomass in Amazonian tropical rain forests. The symbiosis and sociality in termites, together with their super abundance, brings them to play a keystone role of "super decomposer" of dead plant material in tropical terrestrial ecosystems. They consume and transform a large amount of nitrogen-poor dead plant material into nitrogen-rich termite body, which is in turn consumed by a great variety of animals ranging from ants and spiders to chimpanzees and human beings. Thus they form the basis for a

[†]Deceased.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00200-8

Table 1 List of families and principal genera of the world termite fauna, together with distributions of the genera in terms of major zoogeographic regions

Division	*Distribution*
Family Mastotermitidae	
Mastotermes	Australia
Family Kalotermitidae	
Kalotermes	Worldwide
Cryptotermes	Worldwide, except for temperate Eurasia
Neotermes	Worldwide
Rugitermes	Polynesia, New World tropics
Glyptotermes	Worldwide, except for North Temperate Zone
Calcaritermes	New World
Family Termopsidae	
Archotermopsis, Hodotermopsis	Eurasia, Temperate Zone
Zootermopsis	North America
Stolotermes	Africa, Australia, New Zealand
Porotermes	Africa, Australia, New World (Chile only)
Family Hodotermitidae	
Hodotermes, Microhodotermes	Africa
Anacanthotermes	Asia, Africa
Family Rhinotermitidae	
Psammotermes	Eurasia, Africa
Coptotermes	Worldwide, mainly tropical
Heterotermes	Worldwide, mainly tropical
Reticulitermes	Eurasia, North America, temperate only
Prorhinotermes	Worldwide, tropical islands and shores
Termitogeton	Asia
Parrhinotermes	Asia, Australia
Schedorhinotermes	Asia, Africa, Australia
Rhinotermes	New World tropics
Family Serritermitidae	
Serritermes	New World tropics
Family Termitidae	
Subfamily Apicotermitinae	
Speculitermes	Asia, New World tropics
Anoplotermes	Africa, New World
Euhamitermes, Eurytermes, Indotermes, Protohamitermes	Asia
Apicotermes, Hoplognathotermes, Lepidotermes, Trichotermes	Old World tropics
Ahamitermes, Incolitermes	Australia
Microcerotermes	Worldwide except for North America
Amitermes	Worldwide
Drepanotermes	Australia
Gnathamitermes	North America, temperate
Subfamily Termitinae	
Basidentitermes, Crenetermes, Cubitermes, Euchilotermes, Fastigitermes, Foraminitermes, Megagnathotermes, Noditermes, Ophiotermes, Pericapritermes, Procubitermes, Thoracotermes, Promirotermes, Thoracotermes, Unguitermes,	Old World tropics
Cavitermes, Dentispicotermes, Neocapritermes, Orthognathotermes, Spicotermes, Spinitermes	New World tropics
Termes	Worldwide, in tropics
Angulitermes	Eurasia, Africa
Dicuspiditermes, Homallotermes, Microcapritermes, Procapritermes	Asia
Capritermes	Madagascar Africa
Subfamily macrotermitinae	
Acanthotermes, Allodontermes, Ancistrotermes, Protermes, Pseudocanthotermes, Sphaerotermes, Synacanthotermes	Africa
Macrotermes, Microtermes, Odontotermes	Asia, Africa
Subfamily Nasutitermitinae	
Eutermellus, Mimeutermes, Verrucositermes	Africa
Bulbitermes, Hirtitermes, Hospitalitermes, Lacessititermes	Asia, New Guinea
Grallatotermes, Trinervitermes	Asia, New Guinea, Africa
Armitermes; Constrictotermes, Convexitermes, Cornitermes, Curvitermes, Labiotermes, Paracornitermes, Procornitermes, Rhynchotermes, Subulitermes, Syntermes, Velocitermes	New World tropics
Tenuirostritermes	New World
Nasutitermes	Worldwide, in tropics

Source: Reproduced from Wilson EO (1971) *The insect Societies*. Cambridge: Harvard University Press.

large food web in the ecosystems, although they are well-known pests of agriculture and wooden buildings.

Furthermore, termites are prominent ecological engineers, modifying the soil properties by constructing huge mounds and long subterranean galleries and providing many animals and plants with hetrogenous habitats. It is also possible that they are an important player in global change scenarios because they have high biomass and emit a large amount of greenhouse gases such as methane as well as CO_2.

There are several landmarks in the process of global diversification of termites: (a) acquisition and loss of symbiotic flagellates, (b) social evolution, (c) change of life type from one-piece type to separate type, and (d) change of dominant feeding habits from wood feeding to soil feeding. This chapter explains an outline of the diversification of termites in relation to those events as well as their ecosystem functions.

A textbook of termites, *Biology of Termites*, was edited by Krishna and Weesner (1969, 1970), and monographs of termites were published by Grasse (1982, 1984, 1986). A new textbook, *Termites: Their Symbiosis, Sociality and Global Diversification*, was recently edited by Abe *et al.* (2000).

Profile of Termites

Phylogeny and Fossil Records

The Isoptera, which includes 280 genera and more than 2650 described species, is a small order, considering that more than 0.9 millions species of insects have been described. Termites include seven families: Mastotermitidae, Hodotermitidae, Termopsidae, Kalotermitidae, Rhinotermitidae, Serritermitidae, and Termitidae. Principal genera of the world termite fauna and their distribution are given in **Table 1**.

The phylogeny of termites is shown in **Figure 1.** Three families, Mastotermitidae, Hodotermitidae, and Termopsidae, are the oldest living lineages. These three families together with Kalotermitidae, Rhinotermitidae, and Serritermitidae are called lower termites, which are characterized by the presence of symbiotic flagellates (Protozoa) in the gut. A remaining family, the Termitidae, is made up of the so-called higher termites, which are characterized by the absence of symbiotic flagellates, although some higher termites harbor symbiotic amoeba in the gut. The species diversity is the highest in Termitidae (the higher termites, ca. 1900 species), occupying about 70% of all species, while the Kalotermitidae have the highest species diversity (ca. 410 species) among the lower termites.

The fossil record of termites is fragmentary and has not been known from Africa. The number of fossil and living species of each family is shown in **Table 2**. The oldest fossil is the worker of *Meiatermes bertrani* (Hodotermitidae) found in limestone deposits in Spain, dating to 130 million years ago (Cretaceous). A related species, *M. araripena*, was discovered in limestone deposits in Brazil, dating to 110 million years ago. The next oldest fossil species is *Valditermes brenanae* (Mastotermitidae) from England dating to 120 million years ago and *V. acutipennis* from Mongolia. Other Cretaceous species include *Cretatermes carpenteri* (Termopsidae), *Luteitermes prisca* (Hodotermitidae), and *Mastotermes sarthensis* (Mastotermitidae). Those Mesozoic termite fossils are found in Europe, Asia, and North and South America, indicating a broad Pangean distribution in both temperate and tropical habitats. A living termite, Mastotermes darwiniensis, whose distribution is confined to Australia, may well be called a living fossil.

Many fossil species of the Kalotermitidae and Rhinotermitidae, as well as Mastotermitidae, Hodotermitidae, and Termopsidae, have been found in the Tertiary period, while only a few fossil species of the Termitidae (higher termites), which have the highest living species diversity, have been found in the Tertiary period.

Thus, the adaptive radiation of the higher termites seems to have occurred in the late Tertiary or Quarternary (Pleistocene) after the breakdown of Gondwana, yet they show a worldwide distribution. This is an interesting problem remaining to be explained because the flight distance of termite alates is usually limited to some kilometers.

Based on the Cretaceous fossil record and living species diversity, two lower termite groups can be distinguished, primitive and modern. Primitive lower termites characterized

Table 2 Numbers of Known Species of Termites

Groups	*Cretaceous species*	*Tertiary species*	*Pleistocene species*	*Living species*
"Lower" termites				
Mastotermitidae	3	19		1
Kalotermidae		33	0	411
Termopsidae	1	7	0	20
Hodotermitidae	3	14	2	15
Rhinotermitidae		16	1	305
Serritermitidae				1
"Higher" termites				
Termitidae				
Apicotermitinae				186
Termitinae		4		702
Macrotermitinae		3		332
Nasutitermitinae				659
Totals	7	96	3	2632

Source: Reproduced from Kambhampati and Eggleton, 2000, and Thorne *et al.*, 2000.

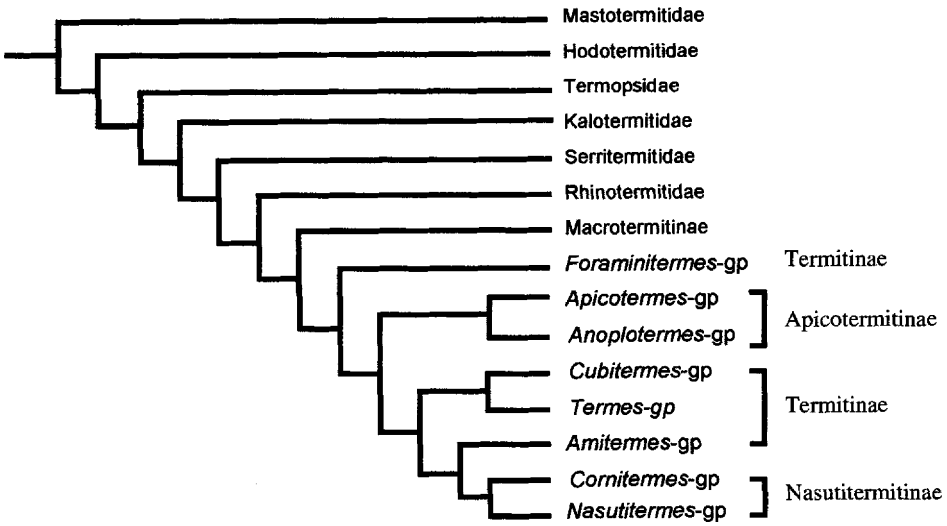


Figure 1 Phylogeny of termites. Reproduced from Kambhampati and Eggleton, 2000.

by Creataceous species and low-living species diversity include Mastotermitidae, Hodotermitidae, and Termopsidae. The three families are the oldest living lineage as shown in **Figure 1**. Modern lower termites characterized by moderate Tertiary and living species diversity include Kalotermitidae, Rhinotermitidae, and probably Serritermitidae.

Distribution

Global generic distribution of termites is shown in **Figure 2**. Termites occur between the latitudes of 30–45N ~ 40–45S and are mainly distributed in tropical regions, especially in western and central Africa, southeast Asia, northeastern Australia, and South America. The species diversity drops from 10 degrees north and south of the equator, but this gradient is not uniform on both sides of the equator (**Figure 3**).

The three groups of termites—primitive lower, modern lower, and higher termites—show different distributions. Primitive lower termites show a fragmental or disjunctive distribution. The Mastotermitidae have only one species, *Mastotermes darwiniensis*, which is distributed in the savanna of northern Australia, although more than 20 species of Tertiary fossils are known from Europe, Australia, and North and South America. The Hodotermitidae including three genera (*Hodotermes*, *Microhodotermes*, and *Anacanthotermes*) and 15 species of grass feeders are distributed in the savanna and arid grassland of Africa and Asia although the family was widely distributed (Spain, Brazil, Europe, Japan, United States) in the Cretaceous and Tertiary periods. *Hodotermes* and *Microhodotermes* occur in Africa, while *Anacanthotermes* occurs in north African deserts, the Middle East, and southern India. The Termopsidae containing 20 living species show an amphitropical distribution: Three species of *Porotermes* are distributed in temperate regions of three southern continents (one species in each of Chile, Australia, and South Africa). Several species of *Stolotermes* are found in Australia, New Zealand, Tasmania, and South Africa, whereas the other three genera are distributed in the temperate and subtropical northern hemisphere (*Archotermopsis* in Kashmir, India, *Zootermopsis in* North America, and *Hodotermopsis* in Japan, Vietnam, and China).

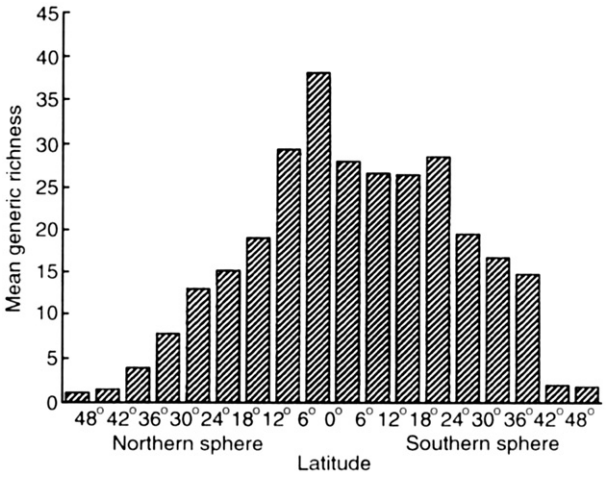


Figure 3 Latitudinal gradients of termite generic richness north and south of the equator. Reproduced from Eggleton, 2000.

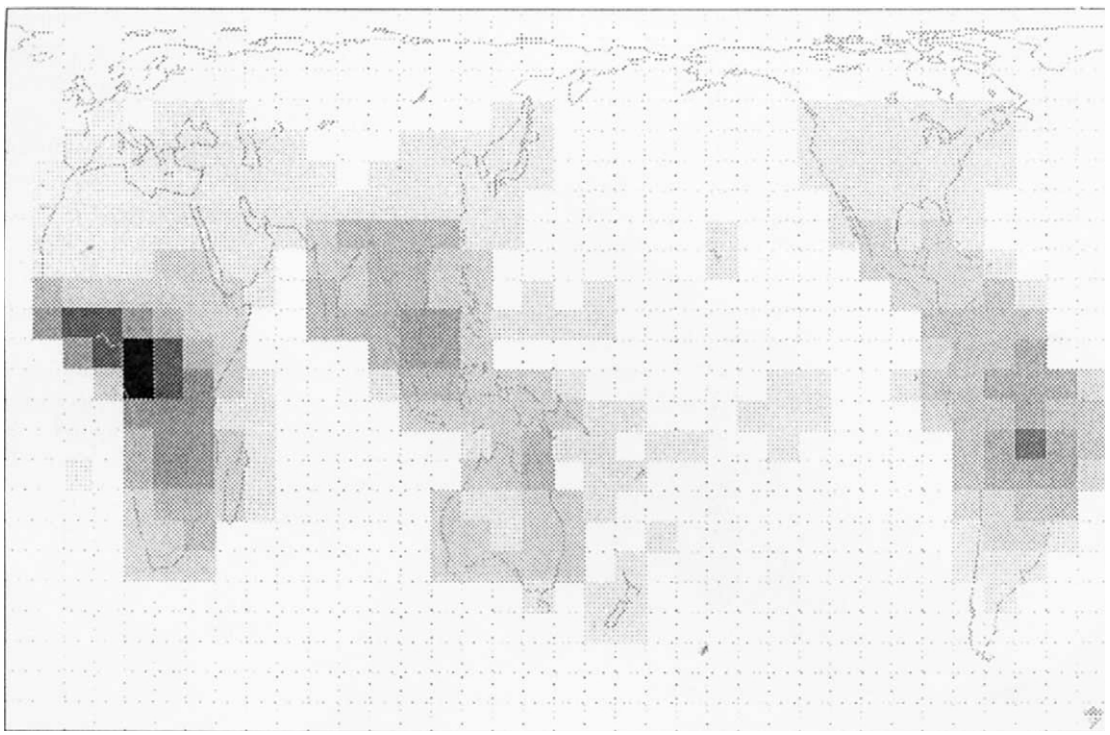


Figure 2 Global generic distribution of termites. The darker the square the higher the generic richness. Maximum generic richness is the southern Cameroon square. Reproduced from Eggleton, 2000.

The modern lower termites except the Serritermitidae show a worldwide distribution. The Serritermitidae includes only one species, *Serritermes serrifer* in Brazil, and its phylogenetic position is under debate. The Kalotermitidae are distributed worldwide but tend to be distributed in coastal forests or islands. Among them, *Glyptotermes*, *Cryptotermes*, and *Neotermes* are pantropical, while *Kalotermes* shows a bipolar distribution. The Rhinotermitidae are widely distributed across tropical, subtropical, and temperate regions, and from wet to dry areas. For example, *Psammotermes* is one of the few insects persisting in deserts of Africa, the Middle East, and northwest India. *Coptotermes*, some species of which are serious pests of wooden buildings, are pantropical, partly because they are carried by humans. The distribution of *Heterotermes* and that of *Reticulitermes* are complimentary: *Heterotermes* is distributed in all tropical regions except for Africa, while *Reticulitermes* shows a Holarctic distribution. *Prorhinotermes* is mainly distributed on tropical oceanic islands.

The higher termites, including the four subfamilies of the Termitidae—Macrotermitinae, Apicotermitinae, Termitinae, and Nasutitermitinae—are characterized by central (tropical) distribution. The Macrotermitinae (fungus-growing termites) are distributed in the central areas of tropical Africa and southeast Asia. Three genera, (*Macrotermes*, *Odontotermes*, and *Microtermes*) are common. The *Apicotermitinae* include the *Apicotermes* and *Anoplotermes* groups. The *Apicotermes* group, most species of which are soil feeders, is distributed in tropical Africa, while the *Anoplotermes* group, which is a soldierless soil feeder, is distributed in tropical Africa, southeast Asia, and South America. The Termitinae are pantropical, including dominant genera of *Cubitermes*, *Termes*, *Amitermes*, and *Microcerotermes*. The Nasutitermitinae are also pantropical, including dominant genera of *Nasutitermes*, *Trinervitermes*, and *Cornitermes*.

It is interesting to note that the higher termites and some lower termites (the Termopsidae, the Kalotermitidae, and *Prorhinotermes* of the Rhinotermitidae) are complementary in geographic distribution: the former show a central distribution (occupying tropical mainland), while the latter show a marginal distribution (localized in temperate regions and tropical coastal areas and islands).

Termites as Eusocial Insects

An excellent introduction to eusocial insects was given by Wilson (1971). Eusociality is generally characterized by two generations of conspecific adults living together (that is,

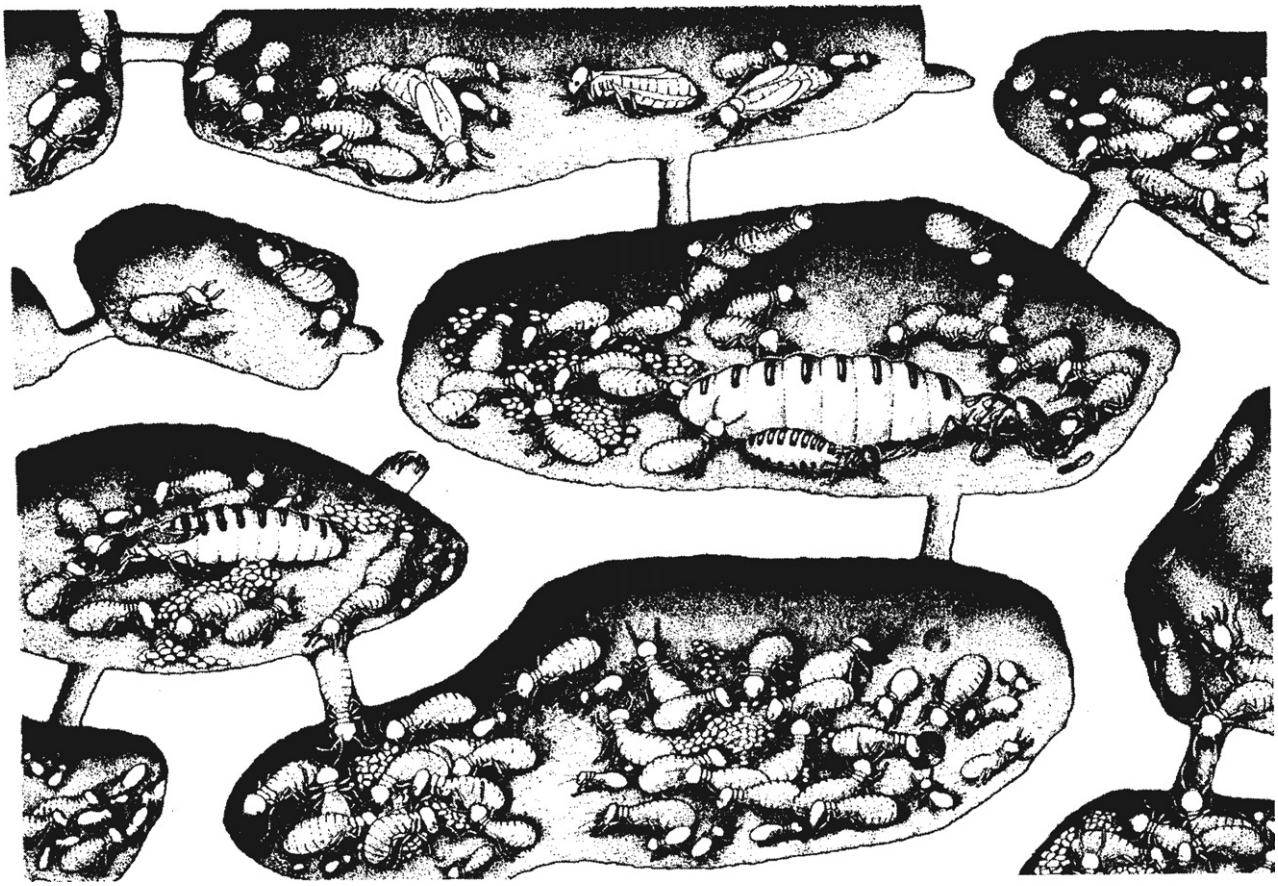


Figure 4 The interior of a nest of *Amitermes hastatus* of South Africa. The primary queen and male sit side by side in the middle cell. To the lower left is a secondary queen, which is also functional in this case. In the chamber at top center are reproductive nymphs, characterized by their partially developed wings. Workers attend the queens and are especially attracted to the heads, to which they offer regurgitated food at frequent intervals. Other workers care for the numerous eggs. A soldier and presoldier (nymphal soldier stage) are seen in the lower right chamber, while worker larvae in various stages of development are found scattered through most of the chambers. Reproduced from Wilson EO (1971) *The Insect Societies*. Cambridge: Harvard University Press.

forming a group called a colony) and cooperation with each other to the extent that a strong reproductive skew is observed—that is, only one or a few adults in the group lay viable eggs (i.e., division of reproduction). The interior of the nest of a higher termite, *Amitermes hastatus*, is shown in **Figure 4**. The primary queen (the largest individual) and king sit in the middle cell. Workers attend to the queen, to which they offer regurgitated food. Developmental pathways of the fungus growing termites are schematically shown in **Figure 5**.

The eusociality of termites has evolved at least four times as sterile castes (once for the soldier caste and three times for the worker caste). Some wasps, some bees, all ants (Hymenoptera), and all termites (Isoptera) are typical eusocial insects, and sterile members have been found also in gall-forming aphids (Hemiptera) (soldiers), gall-forming thrips (Thysanoptera)(soldiers), an ambrosia beetle (Coleoptera) (workers), some mole-rats (workers), and a coral reef shrimp (soldiers), which are included in eusocial animals.

Termites have similar social organization to the Hymenoptera but differ from them in several important aspects (**Table 3**). Termites are characterized by the diversity of soldiers, while the Hymenoptera are characterized by the diversity of workers. Furthermore, in termites, (a) females and males are both diploid (whereas females are diploid and males are haploid in the Hymenoptera), (b) females (queens) and males (kings) both take care of brood and male remains with female through life (whereas males are produced just before the mating season and die soon after mating in the Hymenoptera), (c) workers and soldiers are either male or female or both (whereas they are all females in the Hymenoptera), and (d) termites are hemimetabolous (whereas they are holometabolous in the Hymenoptera).

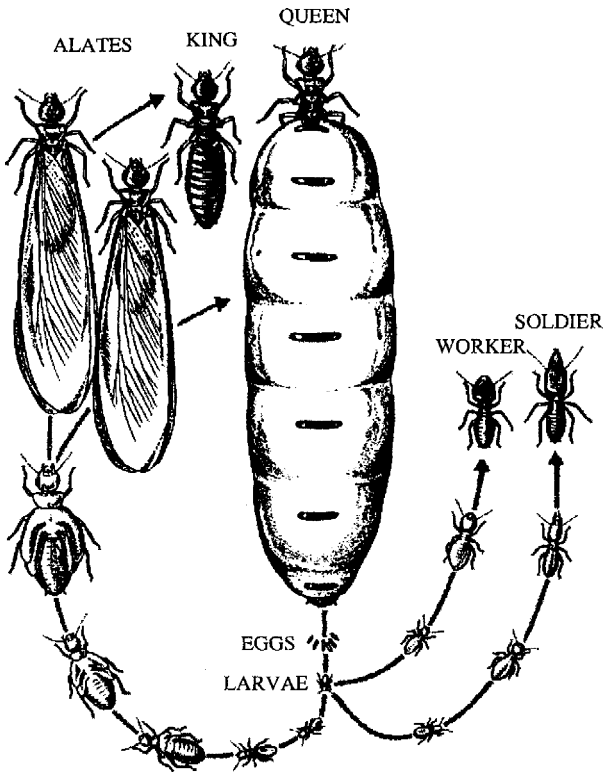


Figure 5 Schematic illustration of the developmental pathway of higher termites. Reproduced from Behnke FL (1977) *A Natural History of Termites*. New York: Charles Scribner's Sons.

Food and Predators

Animals are divided into two large groups in relation to their feeding habits: cytoplasm consumers and cell wall consumers. Most termites are cell wall consumers, whereas the Hymenoptera, including social insects such as bees, wasps, and ants, are cytoplasm consumers. The cell wall of plants consisting of cellulose, hemicellulose, and lignin is very poor in nitrogen but tends to be abundant and clumped, while cytoplasm is rich in nitrogen but tends to be sparsely distributed. Cellulose, a major component of the plant cell wall, is decomposed to glucose, an important source of energy and body construction for animals. However, cellulose is difficult for most animals to decompose, especially when it is combined with lignin, which is extremely resistant to chemical degradation.

Termites as a whole feed on live and dead vegetation, lichens, algae, fungi, dung, soil rich in organic matter (so-called humus), other individuals of the same colony (cannibalism), organic-rich portion of termite nests, and skins of vertebrate corpses (Wood, 1978). Most termites in natural ecosystems do not attack living plants, although *Microcerotermes* sometimes attacks native living trees in Malaysia and some termites consume living native grasses in savannas and grasslands. Lower termites eat dead wood except Hodotermitidae, which are grass feeders, while higher termites are rich in feeding habits. Some higher termites take specialized food. Some genera of the Nasutitermitinae feed on lichens on tree trunks: *Hospitalitermes* in southeast Asia, *Constrictotermes* in South America, and *Grallatotermes* in east Africa, India, the Philippines, and New Guinea. *Ahamitermes* and *Incolitermes* in Australia and *Ophiotermes* and *Tuberculitermes* in Africa are obligate inquilines in the nest of other termite species and feed exclusively on the carton of nest material of their hosts.

Soil feeders, which are recognized in the Termitinae, Apicotermitinae, and Nasutitermitinae of higher termites, include 1100 species, which is approximately 45% of the described termite species. The development of soil feeding is a landmark in the evolution of termites, considering that it is rare among animals. The real food source for soil feeders is unknown.

Termites are included in the diets of a wide range of animals. Ants are the most dominant predators. They are either obligatory or opportunistic predators. Some ponerine ants, such as *Termitopone commutata* in South America, *Leptogenys chinensis* in Sri Lanka, and *Megaponera foetens* and some subterranean *Dorylus* spp. in Africa, are typical obligate predators. The ant with lethal effects on a termite colony changes with the colony's development. For a colony of the fungus growing termite, *Macrotermes michaelseni* in Kenya, it is opportunistic predatory ants (individual foraging ants such as *Pachycondyla* and *Ophthatmopone* and ants with recruitment systems such as *Camponotus* and *Pheidole*) at their early stage, ants with marked recruitment by trail pheromone (e.g., *Megaponera foetens*) at the middle stage, and subterranean group foraging doryline ants (e.g., *Dorylus juvenculus*) at the late stage.

Table 3 Basic similarities and differences in social biology between termites and higher social hymenoptera (wasps, ants, bees)

Similarities	*Differences*	
	Termites	*Eusocial Hymenoptera*
1. The castes are similar in number and kind, especially between termites and ants	1. Caste determination in the lower termites is based primarily on phermones; in some of the higher termites it involves sex, but the other factors remain unidentified	1. Caste determination is based primarily on nutrition, although pheromones play a role in some cases
2. Trophallaxis (exchange of liquid food) occurs and is an important mechanism in social regulation	2. Most species possess soldiers	2. A few species possess soldiers
3. Chemical trails are used in recruitment as in the ants, and the behavior of trail laying and following is closely similar	3. The worker castes consist of both females and males	3. The worker castes consist of females only
4. Inhibitory caste pheromones exist, similar in action to those found in honeybees and ants	4. Larvae and nymphs contribute to colony labor, at least in later instars	4. The immature stages (larvae and pupae) are helpless and almost never contribute to colony labor
5. Grooming between individuals occurs frequently and functions at least partially in the transmission of pheromones	5. There are no dominance hierarchies among individuals in the same colonies	5. Dominance hierarchies are commonplace, but not universal
6. Nest odor and territoriality are of general occurrence	6. Social parasitism between species is almost wholly absent	6. Social parasitism between species is common and widespread
7. Nest structure is of comparable complexity and, in a few members of the termitidae (e.g., *Apicotermes, Macrotermes*), of considerably greater complexity; regulation of temperature and humidity within the nest operates at about the same level of precision	7. Exchange of liquid anal food occurs universally in the lower termites, and trophic eggs are unknown	7. Anal trophallaxis is rare, but trophic eggs are exchanged in many species of bees and ants
8. Cannibalism is widespread in both groups (but not universal, at least not in the Hymenoptera)	8. The primary reproductive male (the "king") stays with the queen after the nuptial flight, helps her construct the first nest, and fertilizes her intermittently as the colony develops; fertilization does not occur during the nuptial flight	8. The male fertilizes the queen during the nuptial flight and dies soon afterward without helping the queen in nest construction

Source: Modified from Wilson EO (1971) *The insect Societies*. Cambridge: Harvard University Press.

Mammals that specialize in digging out termite nests are pangolins in Africa and India and aardvarks in tropical Africa. Other vertebrates specializing in predation on termite brood centers include armadillo in Texas of the United States, anteaters in South America, sloth bear (*Melurus ursinius*) in India, and the numbat and echidna in Australia. Wild chimpanzees in Africa make and use tools for "fishing" termites (**Figure 6**).

Nests and Life Types

Termites are ecosystem engineers. They construct nests of various shapes and sizes, sometimes reaching to 5 m in height (**Figure 7**). The diversity of nests is related to the diversity of social evolution, colony size, and feeding habits as well as the establishment of a microclimate suitable for termites to defend their colony against predators. The inner parts of carton nests of some species contain a large amount of wood fibers, which are eaten by the builders or utilized by inquiline termites. The most remarkable example is fungus gardens of Macrotermitinae.

Nest material has two origins: exogenous material such as particles of soil and wood transported by colony members and fecal. The cohesion of the nest construction is achieved by saliva moistening of exogenous material. Termites usually use both fecal and exogenous materials for nest construction except for the Macrotermitinae, which do not use fecal material. Soil feeders and arboreal nest builders mainly use fecal material. Soil material is important in most epigeous nests (Abe, 1987).

Three life types are distinguishable in relation to nest and feeding sites: (a) a one-piece type, which nests in a piece of wood and consumes only that piece, (b) an intermediate type, which nests in a piece wood and consumes it as food but also constructs galleries to consume other wood pieces, and (c) a separate type, which nests in various sites such as on living tree trunks, in soil, on the ground surface, and so on, constructing galleries to consume various dead plant material outside of the nest.

The one-piece type is supposed to be the most primitive. An evolutionary trend of life type is a progressive separation between food and nest. The first step is the development of a subterranean gallery system, which allows the colonization of new wood pieces (intermediate type), and the second step is the distinction of the nest from food (separate type). This change of life types seems to have occurred independently at least three times (Mastotermitidae, Hodotermitidae, and Rhinotermitidae) in the evolution of termites. The one-piece type is observed in primitive lower termites (Termopsidae) and

modern lower termites (Kalotermitidae and *Prorhinotermes* of Rhinotermitidae). The Termopsidae usually nest in large damp wood and are called damp wood termites, while the Kalotermitidae usually nest in dry wood of living and fallen trees and are called dry wood termites. *Prorhinotermes* of the Rhinotermitidae nest in ordinary wood. In the one-piece type, the size, growth, and longevity of the colony are constrained by the size of a wood piece.

The intermediate type is observed in primitive lower termites (Mastotermitidae) and the modern lower termites (most Rhinotermitidae). The separate type is observed in primitive lower termites (Hodotermitidae), modern lower termites (a few Rhinotermitidae), and the higher termites (Termitidae). Separate type nests are classified into three categories: (a) subterranean nests constructed below ground, (b) epigeous nests protruding above soil surface, and (c) arboreal nests built on a tree trunk or a tree branch but always linked with soil by covered galleries.

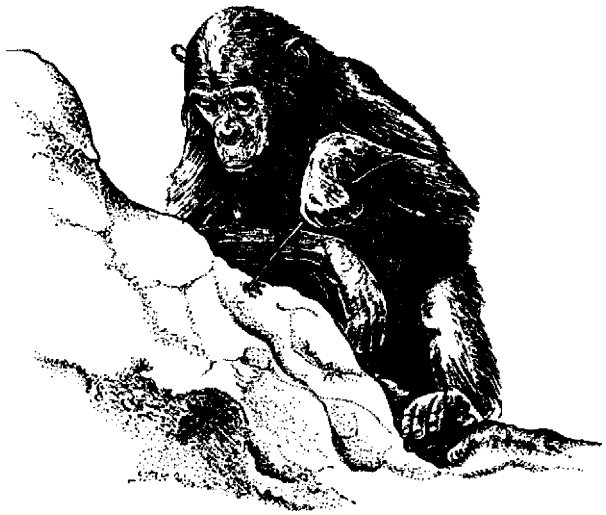


Figure 6 A chimpanzee fishing termites with a stick in a fungus-growing termite mound. Reproduced from Behnke FL (1977) *A Natural History of Termites*. New York: Charles Scribner's Sons.

Symbiosis

Digestive Tube

Termite symbiosis is an obligate nutritional mutualism between the termite and microorganisms in the gut and nest of termites, although symbiosis, in the general sense, means the living together of two or more organisms that are not closely related in phylogeny, without the implication of beneficial exchanges.

The digestive organ of lower termites is schematically shown in **Figure 8**. The digestive tube consists of foregut, midgut tubular (site for secretion of digestive enzymes and for absorption of soluble nutrients), and hindgut (voluminous site for digestion and for absorption of nutrients). The swollen portion of the hindgut is called the paunch. Malpighian tubules, which transport urine and urinary metabolites for excretion, empty at or near the junction of mid- and hindgut. Midgut and salivary glands produce the enzyme for decomposing cellulose into glucose. In the hindgut, cellulose is decomposed into acetic acid by symbiotic flagellates.

The microenvironment becomes more anoxic from foregut to hindgut. The hindgut used to be considered an anaerobic fermentation chamber, but recently this has been found to be false. The hindgut consists of an anoxic lumen surrounded by a microoxic periphery (**Figure 9**). This is consistent with the occurrence of both anaerobic and O_2-dependent microbial metabolism in the hindgut. Bacteria on or near the gut wall constitute an oxygen sink, consuming the inwardly diffusing O_2 and thereby creating anoxic conditions favorable for



Figure 7 Nests of *Nasutitermes triodiae* in Australia.

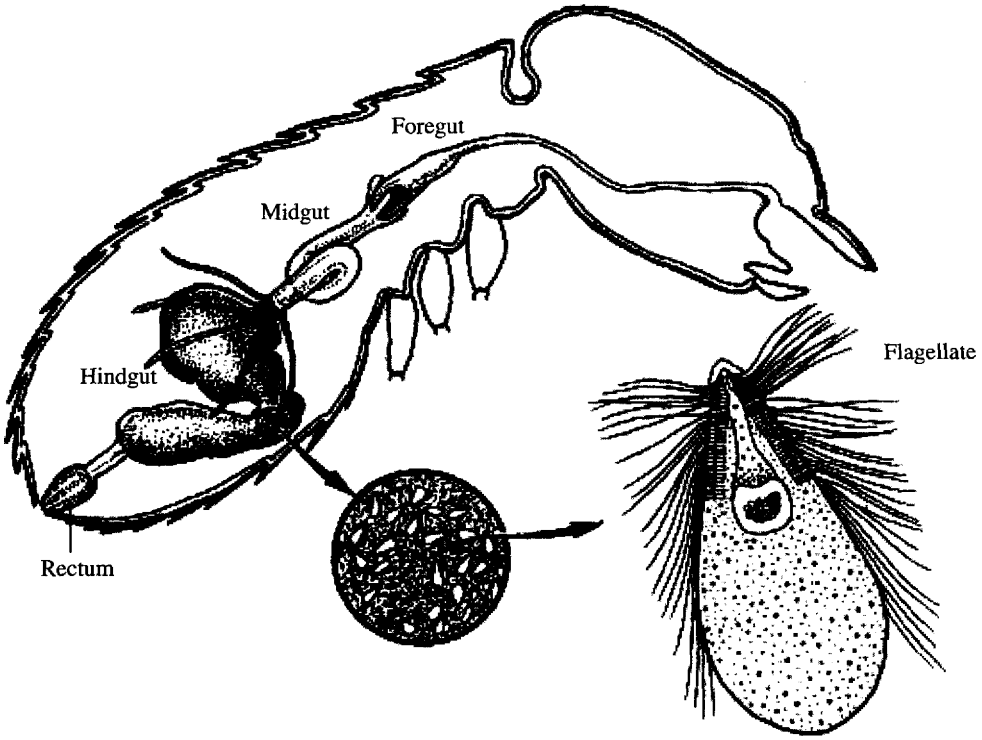


Figure 8 Schematic illustration of digestive system of a lower termite. Reproduced from Behnke FL (1977) *A Natural History of Termites*. New York: Charles Scribner's Sons.

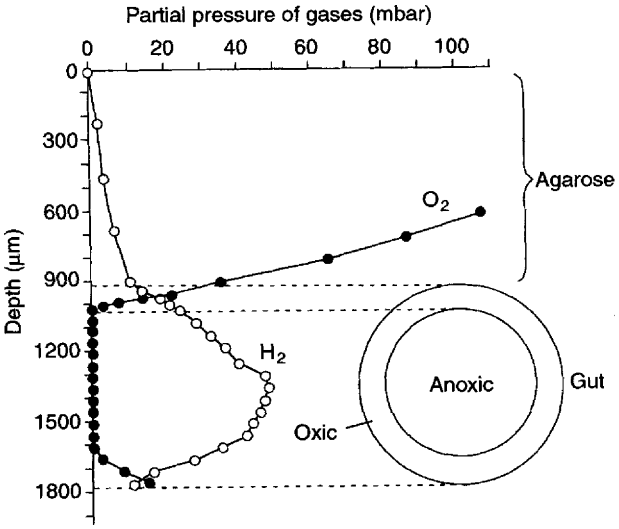


Figure 9 Microenvironment of the hindgut of a termite. Reproduced from Breznak, 2000.

fermentative production of acetate, a major energy source for termites. The hindgut of termites typically has a low redox potential (from -50 to -270 mV) and a pH around neutrality (6.2 to 7.6), although some portions of the hindgut of soil-feeding termites have a pH as high as 11.

Two Types of Symbioses

Woody tissue, a major food of termites, contains only 0.03 to 0.1% nitrogen and their C/N (carbon/nitrogen) ratio is 350 to 500, whereas termite tissues contain 5 to 14% nitrogen and their C/N is 4 to 12. Therefore termites have to solve the "carbon-nitrogen balance" problem in addition to "cellulose decomposition problem" (Higashi *et al.*, 1992). Termites use free-living and symbiotic fungi, gut bacteria, and protozoa for assistance in decomposing cellulose and metabolizing and conserving nitrogenous compounds.

Thus, two symbioses are distinguished in termites: "cellulose decomposition symbiosis" and "C-N balance symbiosis" (**Figure 10**). All the species of termites are, though to different extents, able to produce their own cellulase. This seems reasonable because they have no way to gain glucose, which is used not only for energy production but also for biological synthesis, other than digesting cellulose for themselves. The lower termites, which harbor symbiotic flagellates, need cellulose decomposition symbiosis to increase the efficiency of cellulose decomposition.

On the other hand, C-N balance symbiosis is necessary for all termites, because they are not able to solve the problem alone. Two mechanisms are possible for C-N balance: adding N to their food, referred to in **Figure 10** as route N-1, and selectively outputting C, referred to as routes C-1 and C-2. Termites obtain N either through their food or by fixing atmospheric N. Termites tend to feed selectively on food with above average N context. Some species feed on plant tissue only when it is partially decomposed, and by this means, take advantage of the reduced C/N ratio that results from fungal and microbial decomposition. As selectively eliminating carbon is a "wood-consuming" manner, one-piece type termites, which stay in their wood nest and consume it, are expected to develop mechanisms to add N, such as symbiosis with

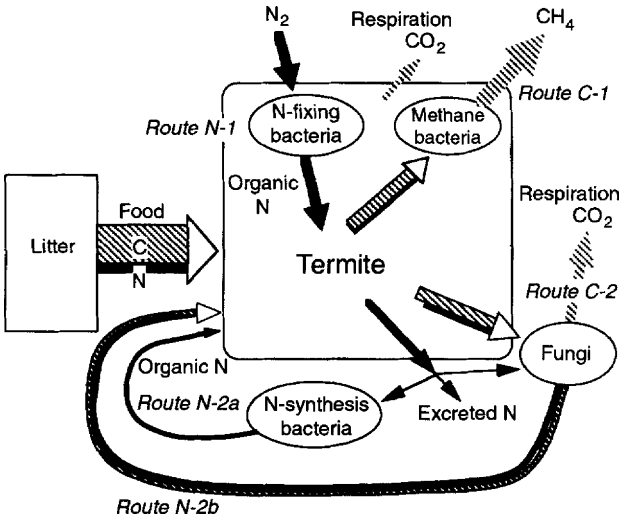


Figure 10 A summary (superimposition) of the possible mechanisms for C-N balance, with the responsible symbionts. Black symbols, nitrogen (N); hatched symbols, carbon (C). Reproduced from Higashi M and Abe T (1997) Global diversification of termites driven by the evolution of symbiosis and sociality. In: Abe T, Levin SA, and Higashi M (eds.) *Biodiversity*, pp. 83–112. Springer.

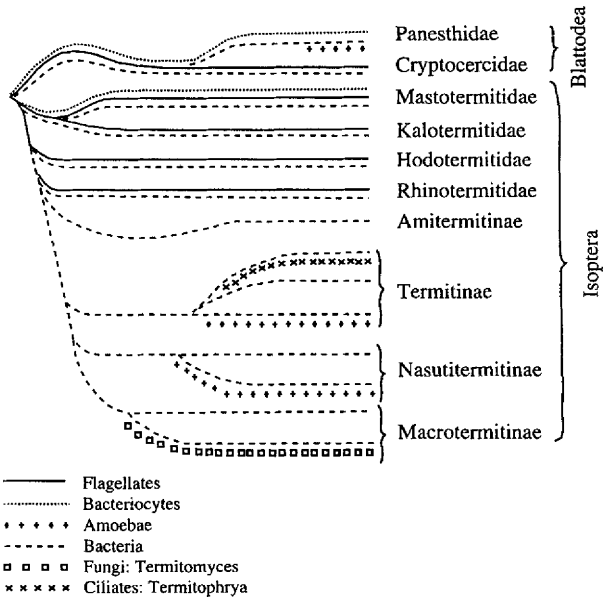


Figure 11 Phylogenetic diagram of the termites, in relation to the evolution of symbiosis with protozoans, bacteria, and fungi. Reproduced from Grasse and Noirot, 1959, with some alterations.

nitrogen-fixing bacteria. On the other hand, separate type termites that are free from the limit in resource utilization are expected to make a full choice among possible mechanisms for C-N balance, including association with methanogenic bacteria. Among the separate type termites, the fungus-growing habit brings additional mechanisms for C-N balance; termites both add N through local N-recycling (route N-2b) and selectively eliminate C through fungal respiration (route C-2).

Cellulose Decomposition Symbiosis

The phylogenetic trend of symbiotic microorganisms in termites and cockroaches is shown in **Figure 11**. Lower termites harbor flagellates in their hindgut. More than 400 species of flagellates of three orders, Oxymonadida, Trichomonadida, and Hypermastigida, have been described from 205 species of lower termites. Each species of lower termites usually harbors more than one species of flagellates. The class Parabasalea, to which Trichomonadida and Hypermastigida belong, is one of the oldest groups among eukaryotes, lacking mitochondria and peroxisomes. Members of this class use hydrogenosomes for energy production and, like the prokaryotes, have a 70S ribosome.

All colony members of lower termites except eggs, newly hatched larvae, and individuals just after molting harbor flagellates. They are transmitted through proctodeal feeding (i.e., transmission of gut contents from anus of a donor to mouth of a receptor). The flagellates of the mother colony are transmitted to offspring colonies by alates. The paunch of the worker is filled with flagellates and bacteria; they occupy 61% by weight of the hindgut contents of *Reticulitermes flavipes* (Rhinotermitidae). The population of flagellates in a gut of *Reticulitermes speratus* often reaches more than 105 individuals.

The association between lower termites and flagellates is well known, since Cleveland (1924) showed that the lower termite died in 3 to 4 if the flagellates were removed, but reinfection by them enabled their host to survive. However, more than 70% of termite species belong to the Termitidae, which lack symbiotic flagellates. Lower termites in addition to higher termites are more or less able to produce cellulases by themselves. They secrete end-β-1, 4-glucanases from the salivary gland, while higher termites secrete it from the midgut. The first cellulase gene for the endo-β-1, 4-glucanas was sequenced from a Japanese lower termite, *Reticulitermes speratus*, in 1998.

A biochemical overview of the major processes involved in the symbiosis between termites and their gut microorganisms is shown in **Figure 12**. The flagellates ingest wood taken by the lower termites and decompose cellulose as follows:

$$C_6H_{12}O_6 + 2H_2O \rightarrow 2CH_3COOH + 2CO_2 + 4H_2$$

The flagellates provide acetate for termite and bacterial metabolism as well as H_2 and CO_2 for bacterial metabolism.

Cellulose in the plant cell wall is combined with lignin to form lignocellulose. Except for fungus-growing termites, the lignin degradation in termite guts remains ambiguous. Some termites such as *Mastotermes darwiniensis* and *Coptotermes acinaciformis* are incapable of digesting lignin, whereas *Nasutitermes exitiosus* seems to be able to degrade lignin at least in part.

C/N Balance Symbiosis

Lower and higher termites harbor a high density and diversity of bacteria in the gut: 109 to 1011 cells per ml of gut fluid. Most bacteria are distributed in the paunch of the hindgut, but no clear evidence to support a significant role for bacteria in

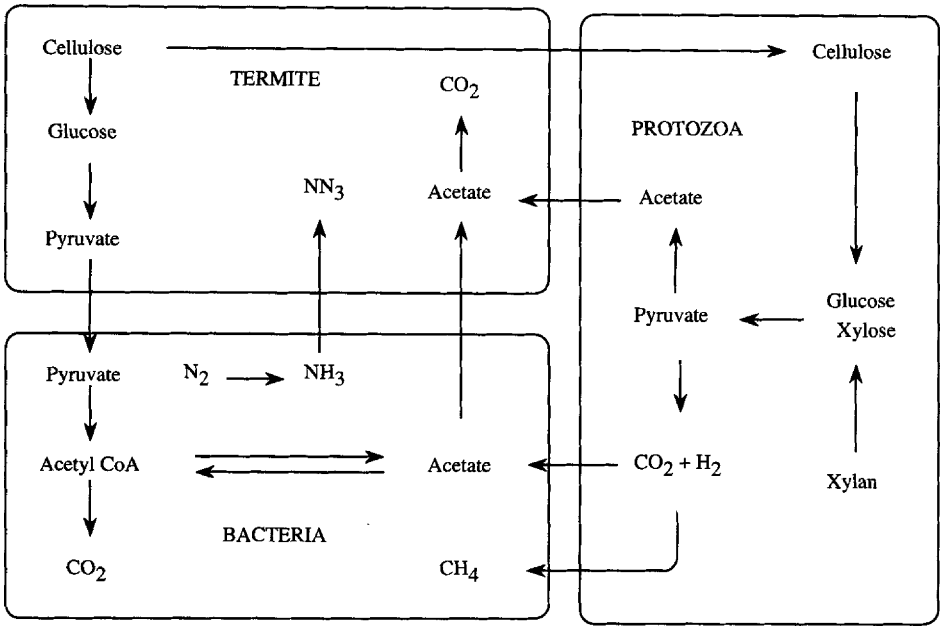


Figure 12 Biochemical overview of the major processes involved in the symbiosis between termites and their gut microbiota. Reproduced from Slaytor, 2000.

cellulose hydrolysis has been given. Intestinal bacteria are important in nitrogen fixation, nitrogen enrichment of the diet, pyruvate metabolism, or the metabolism or acetogenic reduction of CO_2.

Carbon elimination through the activities of microorganisms is the most general mechanism of C/N balance. It entails the release of carbon-rich, nitrogen-poor products of digestion leaving an assimilatable fraction of digesta enriched in nitrogen. Gaseous forms of carbon released are principally carbon dioxide and methane, which is formed through respiration, fermentation, and methanogenesis. Almost all termites emit methane and soil-feeding termites tend to emit more methane than wood-feeding termites. Diverse prokaryotes are capable of participating in such transformations.

As long ago as 1925, Cleaveland suggested that termites use nitrogen fixation to explain their survival on a diet low in nitrogen. Nitrogen fixation, which is energy demanding, is only known from bacteria. Nitrogen fixation in termites was first reported by Benemann (1973) and Breznak *et al.* (1973). Tayasu *et al.* (1994) estimated the extent of nitrogen fixation in termites by comparing the natural abundance of ^{15}N in termite tissues with that of their food and air, showing that 30 to 60% of the nitrogen in the body of *Neotermes koshunensis* (Kalotermitidae) was derived from N_2 fixation. However, the variation in the importance of nitrogen fixation in the termite colony is not clear.

Nitrogen is conserved by internal recycling of nitrogenous waste from termites, by recycling of feces, or by cannibalism and nectophagy. Nitrogenous wastes excreted into the gut through the Malpighian tubules are broken down by gut microorganisms and recycled. Termites have symbiotic, uricolytic bacteria, which use uric acid as an energy source, producing ammonia as the major end product. Ammonia provides a readily available source of nitrogen for protein synthesis in many bacteria.

Symbiosis with Fungi

Higher plants and animals have developed various symbioses with fungi. The Macrotermitinae of higher termites cultivate the basidiomycete fungus, *Termitomyces*, on fungus gardens (fungus combs) in the nest and are called fungus-growing termites. *Termitomyces* is known only from macrotermitine nests. The fruiting body of *Termitomyces* growing on the comb of *Odontotermes formosanus* at the beginning of the rainy season in Okinawa, Japan, is shown in **Figure 13**. Ascomycetes of the genus *Xylaria* grow rapidly on the comb, when the comb is removed from the nest.

Fungus-growing habits are known from ants as well (e.g., *Atta*). It is a curious fact, possibly coincidental, that fungus-growing termites, which consume both fungi and fungus combs, are distributed in Africa and southeast Asia, whereas fungus-growing ants, which consume only fungi growing on fungus gardens, are distributed in America.

In *Macrotermes michaelseni* in the grassland of Kenya, old workers (mainly major workers) go out of the nest and collect dead grass pieces. Young workers in the nest consume the pieces and build fungus comb in the mound chambers by depositing fecal pellets of partially digested grass pieces. The comb, which is shaped into round forms resembling brain or sponge, is inoculated with *Termitomyces*, which develop small round white nodules called conidia. Young workers continually deposit the fecal pellets on the top of the comb and consume the conidia, which contain cellulase as well as glucose. Old workers and soldiers consume the old part of the comb from underneath.

The role of *Termitomyces* has been controversial since Konig and Smeathman discovered the fungus in the 18th century. Suggested roles of the fungus are (a) degradation of lignin in the fungus comb, (b) ingesting and utilization of fungal cellulolytic enzymes in conidia to complement its own digestive

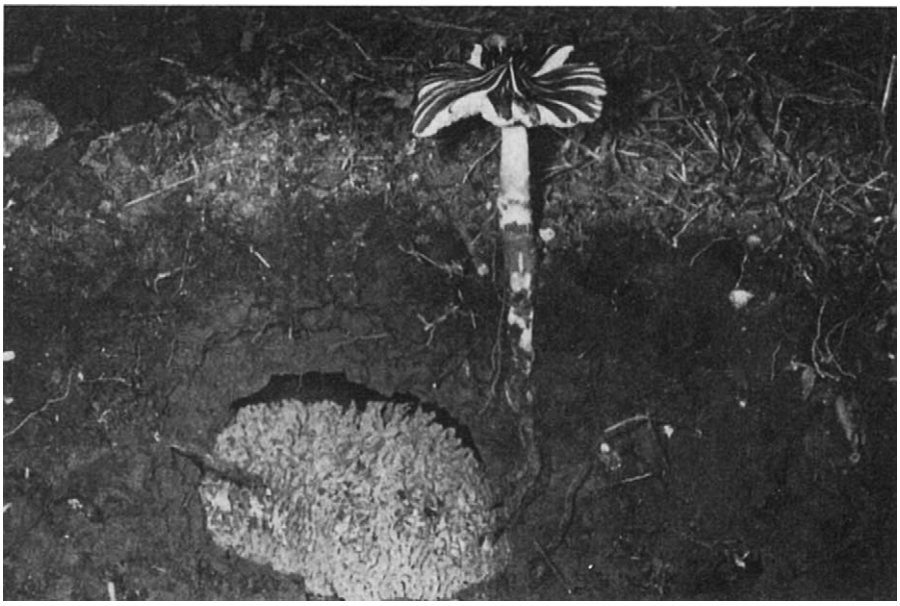


Figure 13 *Termitomyces* (fruit body) growing on the fungus combs of Odontotermes formosanus in Okinawa, Japan. Photo by K. Sugio.

enzymes, and (c) improvement of nitrogen economy in termite nutrition.

Termitomyces is able to break down lignin in addition to cellulose in the comb to simpler compounds, which are reingested by old termites together with the fungi. In *Macrotermes mulleri*, a synergistic interaction between the cellulolytic enzyme of the termite and fungal origin in the gut is known, while in *M. subhyalinus* and *M. michaelseni*, most (90%) of the glucose requirements of workers can be met by the activity of endogenous intestinal cellulases. The extent to which fungal enzymes contribute to cellulose digestion after ingestion by termites is under debate.

The symbiotic fungi seem to be an important nitrogen source for termites, because nitrogen concentration in the fungi is much higher than that in the fungus comb. Thus, all of the three suggested roles of fungus are partly true. The size and morphology of the combs are different among macrotermitine species, and the role of fungi may vary among them. The comb plays a role as reserved food for termites, especially at times when food is scarce out of the nest.

Alates of *Macrotermes bellicosus* and *Microtermes* spp. carry the spores of *Termitomyces* in their rectal chamber, while *Ancistrotermes cavithorax*, *A. guineesis*, *Macrotermes subhyalinus*, *Odontotermes pauperus*, and *O. smeathmani* do not carry the fungal spores. They may depend on the source of spores from fungi growing on the soil surface, which the first workers collect.

Social Life

Caste Differentiation

A termite colony usually consists of a pair of reproductives and their offspring including workers or pseudergates (false workers), soldiers, nymphs, larvae, and eggs. Termites are hemimetabolous, and their developmental pathway consists of several instars with little morphological change until the final molt to alates. The differentiation into castes such as workers, soldiers, and neotenics occurs when immatures deviate from the pathway leading from egg to winged imago (alate). The more evolved the species, the earlier the caste differentiation.

The most important difference in the caste differentiation exists between the one-piece nest type and separate and intermediate nest types. The one-piece type produces pseudergates (false workers), while separate and intermediate types produce workers. The pseudergates, which functionally correspond to workers in separate and intermediate types, retain the capability to be reproductive (alate and neotenic), presoldier, and another pseudergate instar. In this sense, pseudergates are helpers.

The developmental pathways of *Kalotermes flavicollis* of Kalotermitidae (one-piece type) and *Macrotermes michaelseni* of Termitidae (separate type) are shown in **Figure 14**. *K. flavicollis* is characterized by flexible or late caste differentiation. On the way to the imago, which develops from the egg through seven molts, two irreversible deviations to sterile soldiers and neotenics and one reversible deviation to pseudergates occur. There is no sexual dimorphism: the sex ratio among pseudergates and soldiers usually depart little from 1.

Caste differentiation is controlled by pheromones. Reproductives (king and queen) secrete sex-specific pheromones that inhibit the metamorphosis of immatures into reproductives, although the inhibitory pheromones have not been chemically identified (**Figure 15**). Two inhibitory pheromones are produced by the queen and the king and are passed out of their anuses (proctodeal feeding). When the reproductives die, pseudergates or nymphs molt to neotenics and inherit their parents' colony. The presence of soldiers inhibits the metamorphosis of pseudergates into soldiers, but its physiological basis is not yet known.

The flexible caste differentiation in *K. flavicolis* is associated with its nesting habits. The Kalotermitidae to which *K. flavicolis*

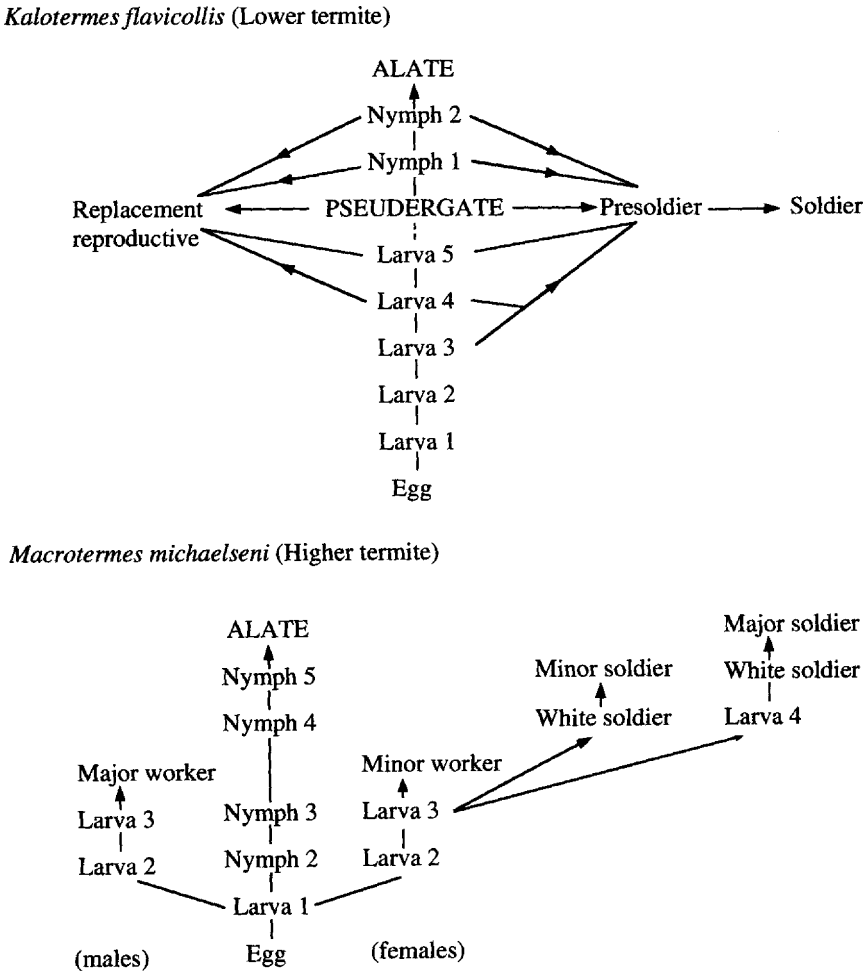


Figure 14 Caste development in lower and higher termites. Reproduced from Roisin, 2000.

belongs nest in the dead branches of living trees, standing dead trees, and rarely fallen trees, consuming the wood. Those nests are unstable, because wood consumption by nesting termites leads to nest destruction.

In *Macrotermes michaelseni* of the Termitidae, the second instar is an irreversible decision point separating two developmental pathways to imago and sterile workers and soldiers. The number of larval instars is three in the Macrotermitinae and two in the other Termitidae. A sexual dimorphism frequently occurs in workers, while soldiers are of one sex. In *Macrotermes*, major workers are males and minor workers are females, while major and minor soldiers are females. In *Trinervitermes* of the Nasutitermitinae, soldiers are males and workers are females. Replacement reproductives of the Termitidae arise either from alates and nymphs or workers.

Evolution of Workers and Soldiers

All species of termites, with some exceptional groups in higher termites, have soldiers. They are adapted to living in the darkness and lack compound eyes. The soldiers of most species are sterile, but fertile soldiers that are able to be neotenics are present in *Zootermopsis*, *Archotermopsis* and *Stolotermes* of Termopsidae, a family of primitive lower termites.

Soldiers, which are specialized in the defense of nest and foraging workers, are rich in the variety of morphology. Two types can be distinguished; mandibulate and nasute types. Mandibulate soldiers occur in all families and rely on large, prominent mandibles to defend the nest and colony members by biting or snapping with mandibles. In nasute soldiers occurring only in Nasutitermitinae of the Termitidae, the front of the head is modified to chemical warfare and the mandibles are in many cases reduced. Defense by nasute soldiers are very effective to ants, major predators of termites (**Figure 16**). Soldiers are of one size in most species, but are dior trimorphic in *Schedorhinotermes* of Rhinotermetidae and Termitidae (e.g., *Macrotermes* and *Acanthotermes* of the Macrotermitinae).

The proportion of soldiers to workers (or false workers) in mature colonies varies considerably. Soldiers occupy up to 30% of colony members in the Nasutitermitinae with nasute soldiers, which are usually smaller than workers, while they occupy about 5% in most termites with mandibulate soldiers, which are usually larger than workers. The proportion of soldiers decreases in soil feeders and some of them (mainly Apicotermitinae of Termitidae) lack the soldier caste.

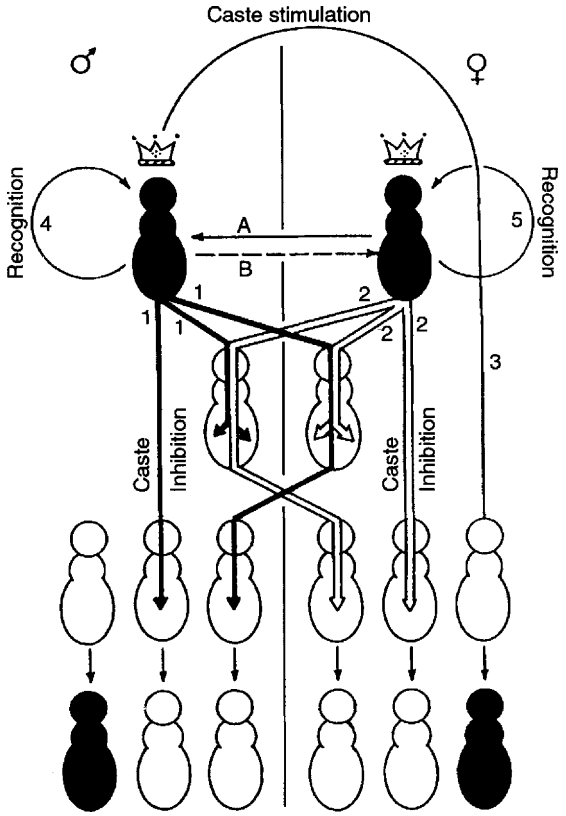


Figure 15 The pathways of pheromone action in the control of reproductive caste formation in the termite Kalotermes flavicollis. In the top row the pair of "crowned" figures represents the reproductives: the functional male (king) to the left and the functional female (queen) to the right. The remaining figures all represent pseudergates. The king and queen produce substances (labeled 1 and 2, respectively) that inhibit transformation of pseudergates into their own royal castes. These inhibitory pheromones are passed directly from the reproductives to the pseudergates and are also circulated indirectly through the digestive tracts of the pseudergates. Another male substance (pheromone 3) stimulates the female pseudergates to change into the reproductive caste, but the reverse relation does not hold. When supernumerary royal males are present, they recognize each other (through pheromone 4) and fight; similarly, supernumerary royal females recognize each other (through pheromone 5) and fight. Finally, royal males stimulate production of pheromone 2 in royal females, and royal females stimulate production of pheromone 1 in royal males; the nature of the stimuli labeled A and B is unknown. Reproduced from Wilson EO (1971) *The Insect Societies*. Cambridge: Harvard University Press.

Workers are numerically the largest caste in the colony and consist of sterile males or females which are wingless, in most cases unpigmented, and lack special external modifications. Compound eyes are usually absent. Workers provide labor for gathering food, feeding dependant castes (larvae, soldiers, nymphs, and reproductives), tending eggs, and repairing and enlarging the nest and gallery system. Worker polyethism is based on age and sex; old workers tend to predominate in outer tasks, while young workers tend to predominate in inner tasks. Workers are of one size in most species, but are dimorphic in some species. Major workers tend to engage in outer tasks, while minor workers tend to engage in inner tasks.

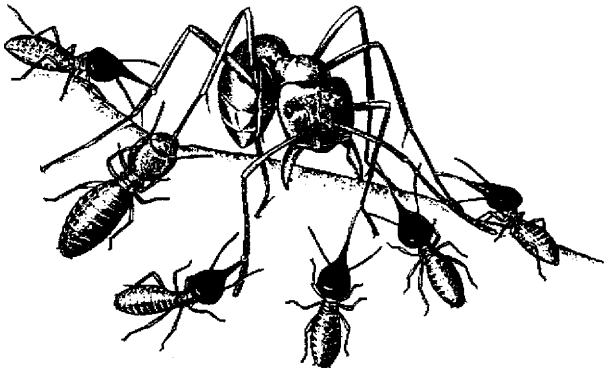


Figure 16 Defensive posture of a worker and nasute soldiers of *Nasutitermes* against an ant worker. Reproduced from Behnke FL (1977) *A Natural History of Termites*. New York: Charles Scribner's Sons.

The presence or absence of sterile castes is shown in relation to life type on the phylogenetic tree in **Figure 17**. Sterile soldiers appeared earlier than sterile workers, evolved once in one-piece type prototermites with false workers, and disappeared secondarily in some soil-feeding termites. On the other hand, the sterile worker caste has evolved polyphyletically. Separate and intermediate type termites (Mastotermitidae, Hodotermitidae, most Rhinotermitidae, and Termitidae) possess a worker caste, whereas one-piece type termites (Termopsidae, Kalotermitidae, and *Prorhinotermes* of Rhinotermitidae) possess a false worker caste. The sterile worker may have evolved from the false worker at least three times independently in accordance with the change from one-piece nest to intermediate and separate types.

Diversity in Social Life

The social organization of termites is similar to that of social Hymenoptera in that the colony is organized based on the division of labor due to caste and age. However, there is an important difference between them. In contrast to Hymenoptera, whose active members are imago (adult, which are suppressed to be reproductive), active members such as workers and soldiers of termites are individuals which are suppressed to be imago.

The degree of eusociality, fecundity of reproductives, and the other life history characteristics are diverse among termite species, and they are summarized in relation to life type in **Table 4**. The reproductive skew, defined as the probability that workers and soldiers forego direct reproduction, is a good indicator of eusociality (0: all colony members reproduce; 1: colony members forego direct reproduction except a single colony breeder). The skew is low in the one-piece type, where pseudergates have the potential to be reproductive, while it is high in the intermediate or separate type, where workers have little chance to become reproductives.

The reproductive skew is negatively correlated with the extent of noetenic production. The percentage of genera producing neotenics is high in the one-piece type (100% in Termopsidae and *Prorhinotermes* of the Rhinotermitidae) except the Kalo termitidae (48%), medium (100% in Mastotermitidae and 71% in Rhinotermitidae) in the intermediate type, and low (67% in Hodotermitidae and 0–20% in Termitidae) in separate types.

No. of species	411	1	20	15	305	1879
Distribution	WW	Au	Amp	Af,As	PT* WW	PT
Nest site	W	W,S	W	S	W	W,S,E,A
Food	W	W	W	DG	W	W,DG,DL,S,Li
Symbionts	P,B	P,B	P,B	P,B	P,B	B,F
Colony size	10^3	10^6	10^4	?	10^4* 10^6	10^6
Life type	O	I	O	S	O* I,S	S
Sterile caste	S	S,W	S	S,W	S* S,W	S(-),W

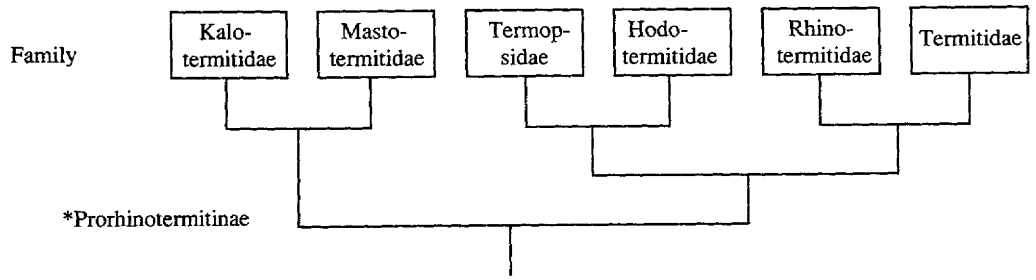


Figure 17 Six families of termites and a summary of their characterization in terms of distribution, nest site, food, symbionts, colony size, life type, and sterile caste. Distribution: Au, Australia; WW, worldwide; Af, Africa; As, Asia; Amp, amphitropical; PT, pantropical. Nest site: W, wood; S, subterranean; E, epigeal mound; A, arboreal. Food: W, wood; DG, dead grasses; DL, dead leaves; S, soil; Li, lichen. Symbionts: P, Protozoa; B, bacteria; F, fungi. Life type: O, one-piece; I, intermediate; S, separate. Sterile caste: S, soldier; W, worker. Reproduced from Higashi M, Abe T, and Burns TP (1992) Carbon-nitrogen balance and termite ecology. *Proc. R. Soc. London B* 249: 303–308.

A remarkable difference among life types is the fecundity of primary reproductives. It is low in the one-piece type, whereas it is sometimes extremely high in the separate type (e.g., a queen of *Macrotermes subhyalinus* lays 36,000 eggs per day).

In comparison with most insects, primary reproductives survive for a long time—10 to 12 years for *Incisitermes minor* (one-piece type) and 20 to 50 years in Macrotermitinae (separate type). It is noteworthy that the longevity of reproductives is similar to that of pseudergates in the one-piece type, whereas it is much longer than that of workers in the separate type.

The age at which termites begin to reproduce is early (beyond the second year in *Kalotermes flavicollis*, beyond the fourth year in *Incisitermes minor*, after the fourth year in *Zootermopsis*, after the sixth year in *Neotermes tectonae*) in the one-piece type, and late (after 5 to 10 years in many species of Rhinotermitidae and Termitidae) in intermediate and separate types. The size of mature colonies producing alates ranges from a few hundred individuals for Kalotermitidae (one-piece type), many thousands or much more in Mastotermitidae, Rhinotermitidae (intermediate type), and Termitidae (separate type).

The colony size is shown in **Table 5**, although the accuracy of estimation is different. The one-piece type such as the Kalotermitidae and Termopsidae is small, usually up to 10,000, while the separate type such as mound builders of the Termitidae, in particular *Macrotermes* and *Nasutitermes*, is very large up to some millions. The intermediate type, including the Rhinotermitidae and Mastotermitidae, is sometimes very large up to some millions.

Organization of Foraging

Foraging of termites has evolved in concert with the change of life type. The one-piece type of termites do not forage out of nest, and the nest site selection by alate reproductives plays a role of food research. Dealate reproductives of *Zootermopsis* (Termopsidae) look for a colony founding site in the wood cambium layer containing nitrogen of relatively high concentration.

In intermediate and separate type nesters, a group of workers and soldiers go out of the nest and search for new food sources. Many species show aggressive behavior toward the other colony members of the same or different species, and the foraging territory, which is intra- and interspecifically defended, is known from many termite species.

The division of labor in foraging activities is based on caste and age. Soldiers usually defend foraging workers, and in some species they explore new areas for food and recruit workers to new food sites. Workers mainly collect and process food, and they also defend foraging territory.

Foraging activities are mediated by trail pheromones emitted from the sternal gland. Several substances are isolated from the gland: n-hexanoic acid in Termopsidae, dodecatrienol in Rhinotermitidae, and neocembrene in

Table 4 Distinguishing social characteristics associated with each nest type

Life type:	*One piece type*	*Intermediate type*	*Separate type*
Termite group	Termopsidae, most Kalotermitidae and *Prorhinotermes* of Rhinotermitidae	Mastotermitidae, few Kalotermitidae, most Rhinotermitidae and some Termitidae	Hodotermitidae, some Rhinotermidae and most Termitidae
Reproductive skew*	0–0.5	0.4–0.75	0.7–1.0
Task specializations	Sterile soldiers, false workers and age polyethism	Sterile soldiers, sterile workers and age polyethism	Sterile soldiers, sterile workers and age polyethism sterile soldiers (except in some soil feeders), sterile workers and age polyethism
Fecundity of primary reproductive (eggs/day)	Small *Cryptotermes havilandi:* 8		Large *Odontotermes obesus*: 26,000–86,000 *Macrotermes subhyalinus*: 36,000 *Nasutitermes surianamensis*: 3,900 *Cubitermes severus*: 50–600
Age of primary reproductives	Field-observed reproductive 4–5 years in *Zootermopsis* spp. and 10–12 years in *Incisitermes minor*	*Mastotermes darwiniensis* lab-reared reproductives up to 17 years; *C. formosanus* at least 9 years; *R. lucifugus* primaries and neotenics up to 7 years in the lab	Primary reproductives 20 years; queens 20–25 years in Macrotermitinae spp. (estimates based on size of physogastric queen)
Age of workers and soldiers	*Zootermopsis* false workers up to 4 years, soldiers 4–5 years in the lab; *Neotermes isularis* soldiers and false workers more than 6 years in the lab	*Reticulitermes lucifugus* soldiers up to 5 years in the lab; *R. hesperus* workers 3–5 years in the field; *Coptotermes acinaciformis* workers and soldiers 4 years in the lab	*Macrotermes* spp. workers and soldiers less than a year in the field and lab; *Macrotermes natalensis* up to 1.5 year in the lab; *Cubitermes ugandensis* worker 196–339 days
Colony foundation	A male and female dealate pair	A male and female dealate pair by a cohort of neotenics and workers	A male and female pair (sometimes multiple reproductives) or sociotomy or secondary dealates reproductives, brachypterous neotenics with workers, rarely ergatoid reproductives
Dispersal distance of alates	Short *Cryptotermes*: 1–45 m *Kalotermes*: 20–50 m *Incisitermes*: 120 m *Zootermopsis*: 300 m	Short *Reticulitermes*: 10–200 m	Long *Odontotermes*: over 0.8 km *Macrotermes*: a few km

Source: Reproduced from Shellman-Reeve, 1997.

Nasutitermitinae of the higher termites. Many chemicals that are not endogenous to termites cause an orientation response. 2- phenoxyethanol from the ink of certain ball-point pens have a trail-following effect on *Coptotermes formosanus*.

Territoriality and elaborate division of labor in foraging were studied well in the Macrotermitinae and the Nasutitermitinae of higher termites, because some species of them forage on the ground. A fungus grower, *Macrotermes michaelseni* in Kenyan savanna, constructs a large mound and has a very large foraging territory based on subterranean galleries that extend up to 50 m from the mound, run horizontally, and lead to foraging holes on the ground surface. The species possesses four sterile castes, major and minor soldiers, and major and minor workers. Major soldiers mainly defend the mound, while minor soldiers mainly defend foraging workers. Major workers are foraging specialists, while minor workers mainly perform intranidal works. The division of labor between major and minor workers is more or less different among species of *Macrotermes*.

During the night, major workers emerge from the foraging hole, cut off dead grasses, fallen twigs and cow dung on the ground, and bring them into the foraging hole. They form foraging columns (up to 1 m) between the foraging hole and the foraging site, both sides of which are guarded by minor soldiers and rarely major soldiers. The column is attacked by various species of ants. The foraging areas of *Macrotermes carbonarius*, a mound builder in southeast Asia, change irregularly but seem to be arranged to avoid redundant foraging in the long term (**Figure 18**).

Black marching termites of *Hospitalitermes* of the Nasutitermitinae in southeast Asia form soldier-led foraging columns reaching up to 100 m in length from the nest at the base of trees to feeding sites on standing tree trunks where workers collect lichens. Soldiers explore the feeding site and recruit workers to the feeding site, defending the workers. Three worker castes show division of labor during food collection; minor workers gnaw at lichens and form a food ball that is passed on medium and major workers, which

Table 5 Characters showing evolutionary change within termites

Characters showing advance		*Characters showing regression*	
Primitive condition	*Derivative condition*	*Primitive condition*	*Derivative condition*
Behavioral-physiological		Imagoes (primary reproductives)	
Small numbers in the nature colony	Large numbers in the mature colony	Y-suture of head present	Y-suture of head reduced or absent
Excavated nests with little construction	Elaborately constructed carton nests	Two well-developed ocelli present	Ocelli reduced or absent
Relatively little care of eggs, nymphs, and adults	Relatively great care of eggs, nymphs, and adults	Antennae with numerous segments	Antennae with fewer segments
Food not store	Food stored	Mandibles with 2 or 3 prominent marginal teeth with sharp basal notches	Mandibles with reduced teeth and notches
Nutritive dependence upon symbiotic intestinal flagellates	Nutritive independence from symbiotic intestinal flagellates	Five tarsal joints present	Second tarsal joint reduced or absent
Damp-wood eaters	Dry-wood eaters	Arolium present between tarsal claws	Arolium absent
Wood eaters	Deaf leaf and grass eaters	Carci with numerous segments (up to 8)	Cerci reduced, usually with 2 segments
Wood eaters	Humus eaters	Styli present	Styli absent
No fungus gardens	Fungus gardens	Egg mass in cluster	Eggs laid separately
Few or no termitophiles	Termitophiles, often species-specific and highly modified	Genitalia similar and clearly homologous to those of cockroaches	Genitalia reduced or absent
Workers forage only within excavated tunnels in wood	Workers forage outside nest, sometimes in covered tunnels and sometimes on exposed odor trails	Hind wing with anal lobe	Hind wing without anal lobe
		Pronotum wide and flat	Pronotum narrow and saddle-shaped
Reproductive castes		Workers	
Small abdomen of queen with small ovaries and glandular tissue	Large abdomen of queen with large ovaries and glandular tissue	Compound eyes faceted and pigmented	Compound eyes reduced or absent
Capacity to produce substitute (neotenic) kings and queens	Reproduction confined to primary reproductive caste (image)	Mandibles as in primitive imagoes	Mandibles as in derivative imagoes
Frontal gland absent	Frontal gland present	Soldiers	
Front coxae smooth	Front coxae with ridge or projection	Antennae with many segments (up to 29)	Antennae with fewer segments(As few as 10)
Workers		Compound eyes pigmented and faced	Compound eyes unpigmented, or nonfaceted or even absent
False worker caste present	True worker caste present	Ocellus spot present	Ocellus spot absent
Soldiers		Mandibles with 2 or 3 large marginal teeth	Mandibles with reduced marginal teeth
Head elongate, large, smooth, somewhat flattened, with large, curved, and toothed biting mandibles	Head round, or extremely flat, or phragmotic; surface rough or with ridges and projections, mandibles elongate and thin, or twised for snapping, or with proejction for frontal gland	Soldiers present	Soldiers absent
Mandibles with smooth cutting edges or with reduced teeth	Mandibles with serrated cutting edges		
Pronotum wide, flatly convex, and with smooth edges	Pronotum narrow, saddle-shaped and sometimes with serrated edges		
Front coxa smooth	Front coxa ridge or projection		

These changes have occurred at various places in termite evolution.
Source: Modified from Wilson EO (1971) *The insect Societies*. Cambridge: Harvard University Press.

carry the food ball to the nest. Medium workers also form food balls.

Two arboreal species of Nasutitermitinae, *Nasutitermes nigriceps*, and *N. nigriceps* form a clear foraging territory in a Panamanian mangrove forest. They build arboreal nests and forage along carton-convered tunnels on branches, which radiate from the nest, and have a nonoverlapping foraging area that is intra- and interspecifically defended.

Life History of Fungus-Growing Termites

A new colony is founded by dealates or budding of the parent colony. In most cases, a pair of dealates found a new colony to be primary reproductives. Multiple alate-derived queens are restricted to the Termitidae (ca. 40 species are known). Colony budding is achieved by active migration of swarms containing all castes or by division of a diffusely organized colony.

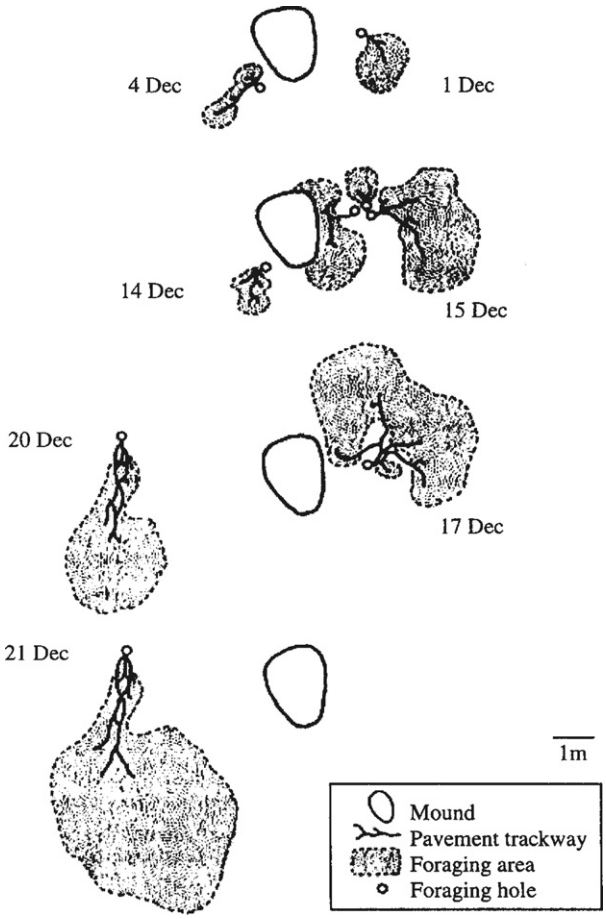


Figure 18 Daily change of foraging areas and pavement trackways of a fungus-growing termites, M. carbonarius in Thailand. Reproduced from Sugio, 1995.

The former, called a "sociotomy," is known only in the Termitidae (e.g., *Trinervitermes bettonianus*, *T. gratiosus*, *Syntermes territus*, and *Anoplotermes* spp.). The latter is known in Mastotermitidae, Kalotermitidae, Termopsidae, Rhinotermitidae, and Termitidae, where the extension of nests is coupled with production of supplementary reproductives.

The life cycle of *Macrotermes michaelseni*, a socially advanced fungus-growing species of the Termitidae in the grassland of Kenya, is shown in **Figure 19**. After the nuptial flight at the beginning of rainy season, a male (king) and female (queen) pair of dealates form a tandem, dig a small cell in the soil, copulate, and initiate a subterranean incipient colony. Occasionally three or more dealates found a new colony together. The primary queen begins to lay eggs within a few days after mating. They hatch after some weeks.

The first brood is tended by the king and queen and is differentiated into soldiers as well as workers. The first workers begin foraging at about 10th week after colony founding in the fungus grower *Microtomes*. The king and queen are similar in body size at colony founding, but the abdomen of the queen becomes much larger than that of the king with time.

Within a few years, *M. michaelseni* constructs a mound that reaches 3 m in height at the mature stage. A large mound builder has a territory based on subterranean galleries, which run horizontally in a shallow layer (10–15 cm). Young subterranean colonies can survive within the territories of large mound builders, because they are located at a deeper layer (20–40 cm) than the subterranean galleries of large mound builders and are able to avoid their attack. However, with colony growth they construct foraging galleries in the shallow layer and are killed by large mound builders. Thus in *Macrotermes michaelseni*, large mounds tend to be uniformly distributed due to territoriality, and small (young) mounds are rare in Kenya.

On the other hand, both large and small mounds of *M. bellicosus* are distributed randomly in Nigerian savanna. The mortality rate of large mound builders of *M. bellicosus* is very high due to predation by doryline ants and aardvarks. Thus, young subterranean colonies are not seriously affected by the proximity of mound builders and survive to construct small mounds, which are randomly distributed.

A marked increase in the number of small mounds following the death of large mound-building colonies due to the severe drought has been reported *M. bellicosus* in Senegal and *M. subhyalinus* in Kenya. Thus, when the mortality rate of large mounds is low, they show uniform distribution due to their territoriality, whereas when their mortality rate is high, young mounds are continuously added and mounds as a whole tend to show random distribution. The duration of mound occupation has been poorly understood: 80 years for *Macrotermes* in Africa, more than 100 years for a large mound of *Nasutitermes triodiae* in Australia, and 20 to 40 years for *Amitermes vitiosus* in northern Australia.

The Global Diversification of Termites

Major Events in Termite History

Three major events of global diversification that Isoptera has experienced are identified (Higashi and Abe, 1997): (a) the evolutionary radiation of original termites, (b) the expansion and diversification of separate type termites and the diversification of one-piece type termites driven by the expanding separate type termites into fragmented habitats, and (c) the radiation of higher termites.

As stated earlier, Cretaceous fossil species belonging to Hodotermitidae, Termopsidae, and Mastotermitidae are known from Europe, Asia, and North and South America. This shows that primitive lower termites had been globally distributed and diversified in the Cretaceous. Therefore, the first event, the evolutionary radiation of original termites, must have occurred at latest in the Jurassic and Cretaceous periods of the Mesozoic. All living species of Hodotermitidae are specialized in grass feeding, while living species of Termopsidae and Mastotermitidae are wood feeders. Grass feeding habits of the Hodotermitidae must have evolved in the Tertiary period, and Cretaceous Hodotermitidae must have consumed wood, because grasses evolved in the Tertiary.

The fact that the oldest fossil record is the worker of Hodotermitidae is interesting, because this suggests that the evolution of the worker caste occurred in the Mesozoic. We do not know if Cretaceous termites of Hodotermitidae nested in wood or soil.

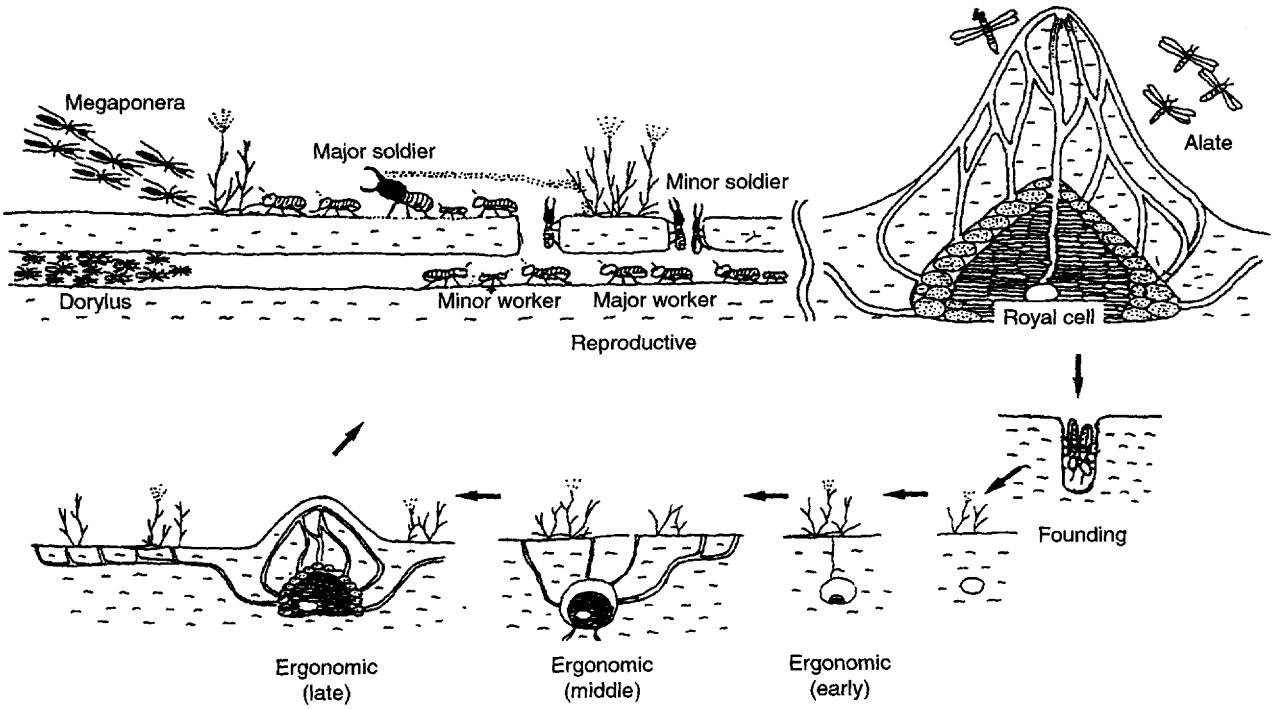


Figure 19 Life cycle of the fungus-growing termite, *Macrotermes michaelseni*, in the grassland of Kenya. Reproduced from Abe, 1995.

Many fossil species of the Kalotermitidae whose living species are mostly of the one-piece type have been found in the Tertiary period. It is probable that the second major event, the expansion and diversification of separate type termites and the diversification of the one-piece type termites driven by the expanding separate type termites into fragmented habitats, occurred in the Tertiary.

Only a few fossil species of the higher termites are known from the Tertiary. The third event, the radiation of higher termites, may have occurred in the late Tertiary or the Quaternary. On the other hand, the worldwide distribution of higher termites suggests their early evolution and dispersal before the breakdown of Gondwanaland in the Cretaceous. As shown later, it is quite difficult for termites to cross the oceans. It is probable that the radiation of higher termites may have occurred in the late Tertiary or the Quarternary, although the origin of higher termites may be before the breakdown of Gondwanaland in the Cretaceous.

Dispersal Ability of Termites

Islands sterilized by volcanic eruption are good for studying the process of colonization. In 1883, Krakatau, located about 60 km from Java and 40 km from Sumatra, erupted violently and more than half of the island disappeared, leaving the remaining Krakatau and two neighboring islands. The entire flora and fauna of the islands were almost destroyed, but they were rapidly colonized by organisms of Java and Sumatra.

In the Krakatau Islands, three, three, and seven species of termite were collected in 1908, 1919–1933, and 1982, respectively. They were characterized by the dominance of the lower termites (Kalotermitidae and Rhinotermitidae), the one-piece type, and the absence of soil feeders, although a wood-feeding arboreal species of *Nasutitermes* of the higher termites was found in the Krakatau Islands. On the other hand, the supposed source area (the tip of West Java) was characterized by higher species diversity (13 species), the dominance of higher termites and separate type, and the presence of soil feeders. In the tip of West Java, the one-piece type termites of the Kalotermitidae are confined to the coastal forest, while soil feeders are confined to inner forests.

Alates of termites are specialized for dispersal within a zone of calm air near the ground. The dispersal distances are short in lower termites (up to 300 m) and a little long in higher termites (a few kilometers). It is not impossible but difficult for termites to disperse widely or to reach remote islands by flight. Therefore, termites usually cross the sea by rafting. The Kalotermitidae, the first colonizers, have some preadaptive attributes for oversea dispersal: (a) they tend to be confined to the coastal area and may be easily swept into the sea, (b) they are more tolerant to sea water than other types, and (c) their caste differentiation is flexible.

On the other hand, it is difficult for soil-feeding termites to cross the sea because they have little opportunity to be in rafting wood. Thus the dominance of the Kalotermitidae (one-piece type) and the Rhinotermitidae (intermediate type) in many islands is explained in terms of the differential dispersal ability of termites.

The Evolutionary Radiation of Original Termites

The first event, the evolutionary radiation of the original termites into a new habitat of wood throughout the world, may have been driven by an efficient utilization of wood as an

abundant and stably supplied food resource and nest substrate, which became possible by the evolution of two symbioses for termites (i.e., the cellulose digestion symbiosis with cellulolic flagellates and the C-N balance symbiosis with nitrogen fixing bacteria). All descendants of the Mesozoic families of Hodotermitidae, Termopsidae, and Mastotermitidae harbor symbiotic flagellates and nitrogen-fixing bacteria in the gut.

Once a termite obtains a means for digestion and C-N balance through association with microorganisms, then wood, which is superabundant but extremely hard to digest and C-N unbalanced, becomes a "well-protected" food resource that can be monopolized by the termite. The diversification of termites into the Hodotermitidae, Termopsidae, and Mastotermitidae together with their worldwide distribution in the Cretaceous may be viewed as an example of niche opening, which often leads to an adaptive radiation.

Conquering the wood habitat should promote the evolution of a false worker caste, because life in a piece of wood causes (a) greater need for concentrating nitrogen, (b) higher possibility of nest succession due to the longer duration time of the nest relative to the longevity of reproductives, and (c) lower possibility of finding an alternative nesting place due to the heterogeneity of resource (wood) distribution. The evolution of this subsociality should further have enhanced the opportunity for eusociality—that is, the production of sterile soldier caste.

Separate Type Expansion and One-Piece Type Diversification

The second major event came with the development of separate type termites, which separate their feeding sites from their nest. They may have expanded, replacing the existing one-piece type termites through a competitive exclusion process, the detail of which was examined by Higashi and Abe (1997). This is a consequence of asymmetrical competition between one-piece and separate types due to the fact that a piece of wood is food and nest for the one-piece type but only food for separate type termites. On the other hand, some groups of one-piece type termites, which were driven out by the expanding separate type termites into fragmented habitats, should have gone through a diversification process.

This process could be reconstructed based on the present peculiar distribution patterns of one-piece type termites and separate type termites. The three groups of one-piece termites show quite different marginal distribution patterns. Damp wood termites (all species of Termopsidae) show a clear amphitropical distribution (**Figure 20**). Ordinary wood termites of *Prorhinotermes* are widely distributed in tropical and subtropical regions but tend to be confined to islands. Dry wood termites (Kalotermitidae) are also widely distributed in the tropical and subtropical regions but tend to be confined to dry dead parts of standing trees.

On the other hand, separate type termites, represented by the Termitidae, are widely distributed in tropical mainlands, the central areas of termite distribution. Intermediate type termites, represented by Rhinotermitidae, are widely distributed from tropical to temperate regions. Thus the distribution of the one-piece type and separate type is complimentary, while that of intermediate type covers the ranges of both types.

Although a full explanation of such distribution patterns requires one to take into consideration other factors such as plate tectonics and global change of climate, ecological mechanisms, specifically interspecific interactions and differential dispersal ability, have been a major force in determining the ranges of distribution.

The asymmetrical competition theory predicts that as the latitude (ocean) gets higher (closer), the resource supply rate of the one-piece type increases relative to that of the separate

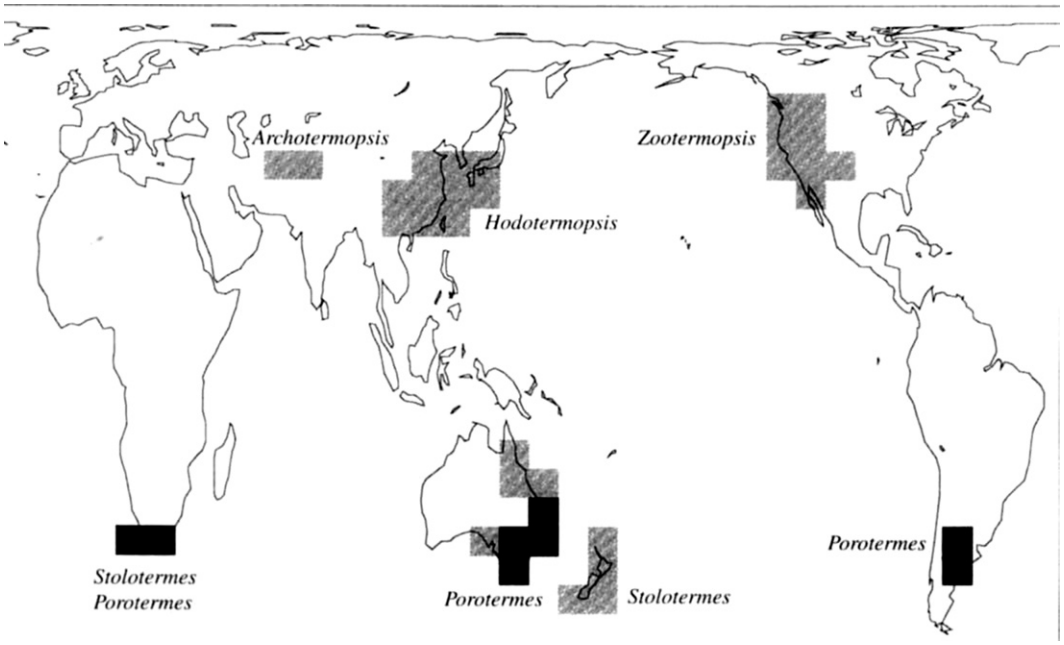


Figure 20 Distribution of the Termopsidae. Reproduced from Eggleton, 2000.

type, because more wood is available to the one-piece type due to the slower decay rate of damp (dry) wood at higher latitude (seacoast) areas under a given predation pressure. Thus, the one-piece type may compensate its disadvantage from its asymmetrical competition with the separate type and becomes dominant at higher latitude (seacoast) areas. This explains the questioned marginal distribution patterns of damp and dry wood termites. The distribution of ordinary wood termites confined to islands may simply be explained as a result of escape from asymmetric competition due to their dispersal ability.

Separation of feeding site from nest (reproductive site) should enhance the C-N balance symbiosis by adding the option of selective carbon elimination and promoting the evolution of a true worker caste. These should further have enhanced the superiority of the separate type in resource utilization, thus driving the expansion and diversification of separate type termites.

Among the one-piece type termites driven out by the expanding separate type termites, those groups driven into fragmented habitats along seacoasts and islands are expected to have gone through a diversification process. In fact, dry wood termites show a high species diversity as a family.

The Radiation of Higher Termites

The third major event, the radiation of higher termites that are free from flagellates, may have been driven by the advantage that a termite would have if it could digest cellulose efficiently enough and eliminate cellulolic flagellates, which became possible by eliminating the lignin cover, the major obstacle in cellulose utilization, perhaps through the shift of their diet from wood to a food with less lignin cover.

Those termites that became higher termites must have obtained their own cellulase activity. The present-day termites have their own cellulase activity. What prevents the elimination of cellulolic flagellates from lower termites may be the inefficiency of the termites in cellulose digestion, which should cause the termites a shortage of (not glucose but) energy supply. Thus, cellulolic flagellates are still symbionts to lower termites, though they are not for cellulose digestion but for energy supply. Higher termites are those that succeeded in increasing the efficiency of cellulose digestion to the extent that they no longer need help from cellulolic flagellates for energy supply either. This may be the driving force for their radiation.

For increasing the efficiency of cellulose digestion, the elimination of the lignin cover is expected to be the most effective. An obvious means to achieve the elimination of the lignin cover is to shift the diet from wood to a food with less lignin cover. The nest-food separation makes possible the selection of food. Thus, it is expected that higher termites may appear only from separate type termites, which is indeed the case.

Three major groups of higher termites are recognized: nasute termites (Nasutitermitinae), soil (humus)-feeding

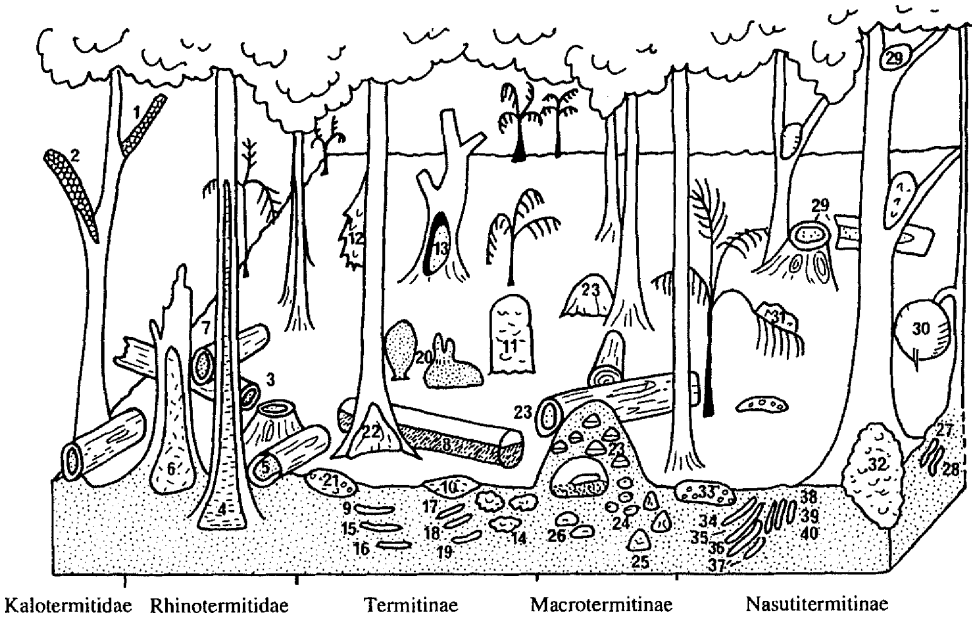


Figure 21 Termites in lowland dipterocarp forest in West Malaysia. Kalotermitidae: 1. *Glyptotermes*; 2. *Neotermes*; Rhinotermitidae: 3. *Termitogeton*; 4. *Coptotermes*; 5. *Parrhinotermes*; 6. *Schedorhinotermes*; 7. *Heterotermes*; Termitinae: 8. *Protohamitermes*; 9. *Prohamitermes*; 10. *Labritermes*; 11. *Globitermes*; 12. *Microcerotermes*; 13. *Amitermes*; 14. *Procapritermes*; 15. *Pericapritermes*; 16. *Coxacapritermes*; 17. *Oriencapritermes*; 18. *Syncapritermes*; 19. *Mirocapritermes*; 20. *Dicuspiditermes*; 21. *Homallotermes*; 22. *Termes* (in *Hospitalitermes* nest); Macrotermitinae: 23. *Macrotermes*; 24. *Microtermes*; 25. *Odontotermes*; 26. *Hypotermes*; Nasutitermitinae: 27. *Hirtitermes*, 28. *Havilanditermes*; 29. *Nasutitermes*; 30. *Bulbitermes*; 31. *Lacessititermes*; 32. *Hospitalitermes*; 33. *Longpeditermes*; 34. *Leucopitermes*; 35. *Aciculitermes*; 36. *Adculioiditermes*; 37. *Subulioditermes*; 38. *Malaysiotermes*; 39. *Praciculitermes*; and 40. *Oriensubulitermes*. Reproduced from Mabberley, 1992.

termites, and fungus growing termites, although there is some overlap among them. They seem to have succeeded in the elimination of the lignin cover.

The Nasutitermitinae show the broadest distribution and the highest species diversity among the higher termites. Their radiation was driven by the evolution of a new type of soldier with a chemical weapon against ants, their major predator, which made it possible to forage out more freely and select food with less lignin cover. The radiation of soil feeders was driven by the shift of diet from wood to humus, which contains less lignin. The radiation of fungus-growing termites was driven by broadening the diet and a new symbiosis with fungi for C-N balance and for lignin reduction.

Termites in Ecosystems

Terrestrial Ecosystems

Terrestrial ecosystems have some characteristics different from aquatic ecosystems—for example, the dominance of the detritus chain among grazing, detritus, and microbial chains; the rare cascading effect of predators on vegetation; the abundance of fungi; and so on. This is partly because vascular plants in terrestrial ecosystems produce lignin, an organic matter extremely resistant to chemical degradation. Plants cover cellulose with lignin to form lignocellulose as a major component of the cell wall. Cellulose is an energy source for many heterotrophs, but its availability decreases markedly when covered with lignin.

Lignin makes it possible for terrestrial plants to support leaves horizontally and transport water from root to leaves and branches, providing various heterotrophs with heterogeneous habitats. Furthermore, lignin regulates the decomposition process of dead plant material, and its remnants play an important role on the formation of humus associated with soil fertility.

Therefore, lignin is a key substance for understanding the architecture and the material flow in terrestrial ecosystems. Considering that microorganisms, especially fungi, are major decomposers of lignocellulose, fungi are reasonably abundant in terrestrial ecosystems, whereas they are poor in aquatic ecosystems that lack lignocellulose.

Brown rot among fungi mainly decomposes cellulose, whereas white rot decomposes both lignin and cellulose. Most soil animals are not able to decompose lignin chemically but decompose it physically and promote the chance of contact between lignocellulose and microorganisms. Termites are distinct among soil animals in that some of them decompose lignocellulose efficiently with the aid of microorganisms in the gut or in the nests.

Many termites live in lowland tropical forests, where 50 to 80 species are found in 1 ha area in **Figure 21**. Their food is rich in variety including dead branches of living trees, fallen logs, fallen leaves, lichens, and soil. Those termites are functionally largely divided into two groups: soil feeders and litter feeders. Litter feeders are furthermore divided into lignocellulose feeders and cellulose feeders.

Soil feeders, confined to three subfamilies of higher termits, are concerned with the formation of humus associated with soil fertility. Lignocellulose feeders (fungus growers of Macrotermitinae) decompose both lignin and cellulose rather completely with the aid of the white rot, *Termitimyces*. As they do not leave many lignin remnants, they may functionally well be called giant moving white rot. Cellulose feeders, including most of the other termites, decompose cellulose in dead plant material, leaving lignin remnants. They may functionally well be called giant moving brown rot. Roughly saying, soil feeders, wood feeders, and lignocellulose feeders occupy about 45%, 45%, and 10% of living termite species, respectively.

We do not know what substrates soil feeders really eat. Soil feeders can generally be distinguished from litter feeders by intestinal morphology, the stable isotope ratios of carbon and nitrogen (**Figure 22**), and by the higher activity of certain gut bacteria, notably methanogens and organisms able to ferment reduced and recalcitrant substrates, including aromatics. An interesting idea that soil feeders eat microorganisms such as bacteria and fungi in ingested soil, and therefore are important members of microbial chain in the tropical forests, is partly supported by the stable isotope analysis and the detection of lysozyme decomposing cell wall of bacteria, but this theory has not been proved.

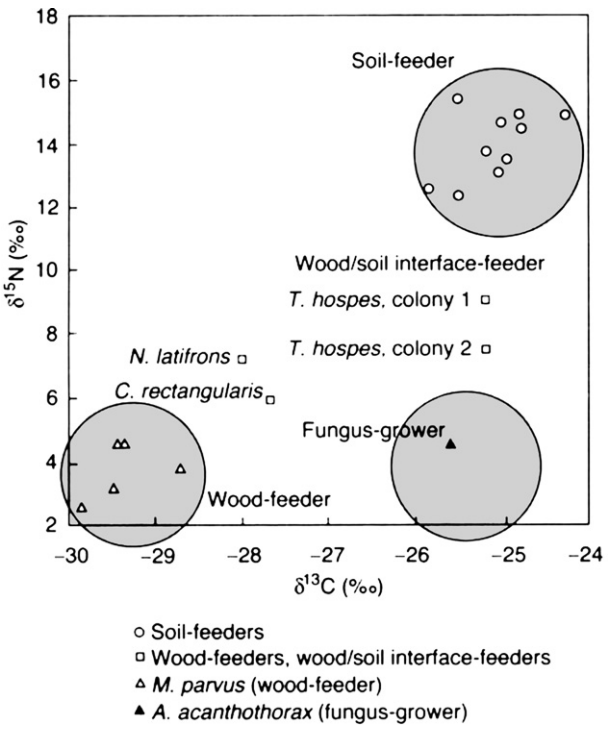


Figure 22 Plot of mean δ^{15}N against mean δ^{13}C for the tissues of worker caste termites from the Mbalmayo Forest Reserve, Cameroon. Soil feeders (colonies of eight species, open circles) cluster at the top right and colonies of the wood-feeding *Microcerotermes parvus* (open triangles) at lower left. Species feeding on highly decayed wood (*Cephalotermes rectangularis*; *Termes hospes*) are intermediate. A single fungus grower, *Acanthotermes acanthothorax*, and the wood-feeding *Nasutitermes latifrons* are also shown. Reproduced from Tayasu *et al.*, 1997.

Table 6 Taxonomic composition and diet structure of termite communities from various African Savannas and Rain Forests

Locality	*Type of vegetation*	*Latitude*	*Annual rainfall (mm)*	*Number of species eating*					*Number of species*					
				Grass and grass litter	*Fresh wood and leaves litter*	*Decomposing litter*	*Humus and soil*	*Kalo-termitidae*	*Rhino-termitidae*	*Termitinae*	*Apico-termitidae*	*Macro-termitidae*	*Nasuti-termitidae*	*Total*
Fétè Olè, Senegal	Wooded steppe (small sand dunes) (reference area)	16° N	375	8	11	0	3	0	3	8	0	3	5	19
Tsavo east, Kenya	Bush and wooded steppe	3.5° S	400	8	17	0	3	4	1	9	0	8	3	25
Cap Vert, Senegal	Woody sub-Guinean savanna	15° N	550	6	17	1	8	1	3	12	2	6	4	28
Mokwa, Nigeria	Primary and secondary woodland	9° N	1175	10	17	1	9	0	1	9	3	11	7	31
Lamto, Ivory Coast	Grassy to woody derived savanna	6° N	1290	10	18	3	13	2	1	10	4	9	10	36
Youhouli, Ivory Coast	Grass savanna	5° N	2000	10	(0)	1	12	0	0	7	5	7	3	22
Edea, Camerooon	Equatorial rain forest	3.5° N	>3000	0	8	4	31	0	1	25	9	5	3	43

Source: Reproduced from Josens, 1985.

Functional Role of Termites

Taxonomic composition and feeding habits of termite assemblages in some African habitats are shown in Table 6. The species diversity is the highest in the tropical forest and decreases from forests to arid ecosystems. The dominant feeding habits gradually change from soil-feeding to wood- and litter feeding (cellulose and lignocellulose feeding). The species diversity of fungus growers (Macrotermitinae) is the highest in slightly dry areas such as savannas and dry forests where soil fungi are poorer than in moist forests. A similar tendency is recognized in tropical forests of southeast Asia.

The change may be explained in relation to the activities of fungi. Soil feeders are abundant and diversified in tropical rain forests where fungi, in particular white rot, may be active and decompose lignocellulose efficiently, decreasing the litter accumulation on the ground. On the other hand, lignocellulose feeders cultivating white rot of *Termitomyces* in the nest chambers with high humidity and temperature are dominant in slightly dry tropical areas where the low humidity may suppress the decomposition of lignocellulose by white rot. Cellulose feeders are widely distributed in areas where fungi may be inactive due to low temperature and inadequate humidity.

Abundance of termites sometimes exceeds 10,000/m^2 and 10 g wet weight/m^2, although it is usually up to 4500/m^2 and 10 g wet weight/m^2. The biomass of 10 g wet height/m^2 is equivalent to 250 persons/km^2 and 40 kg/person in human beings. In the tropical rain forest of Malaysia where fungus growers are dominant, as much as 30% of the leaf litter supply is consumed by termites. In a Guinea savanna in Nigeria where fungus growers are dominant, termites consume up to 55% of surface litter. Termites are responsible for up to 20% of total carbon mineralization in the savannas of Africa. Consumption of dead plant material by termites increases when fungus growers are dominant, because they consume five to six times more food per unit of biomass than other termites.

Termites are consumed by a great variety of animals ranging from ants to human beings, forming a large detritus chain. Furthermore, a large number of alates produced by mature termite colonies are eaten by birds and spiders in the canopy, joining the grazing chain of terrestrial ecosystems.

Termites and Soil

While the distribution and abundance of termites are influenced by climatic conditions (air temperature and humidity), soil properties, and vegetation, termites modify local climate conditions, soil properties, and vegetation through the activities of building mounds and excavating subterranean galleries. The effect of termites on soils has been well examined in the textbook *Termites and Soils* by Lee and Wood (1971).

Termites can modify the soil profile by removing soil from various depths (up to 70 m) and transporting it to the ground surface in the form of runways, sheetings, or mounds, from where it is redistributed by water and wind erosion. In central Africa, termites of *Macroterms* form huge mounds up to 10 m in height and 30 m in diameter. Excavation of subterranean galleries has positive effects on the hydraulic conductivity and infiltration rates.

Large mounds built by species of Macrotermitinae in Africa and Asia often support a vegetation comprising trees and

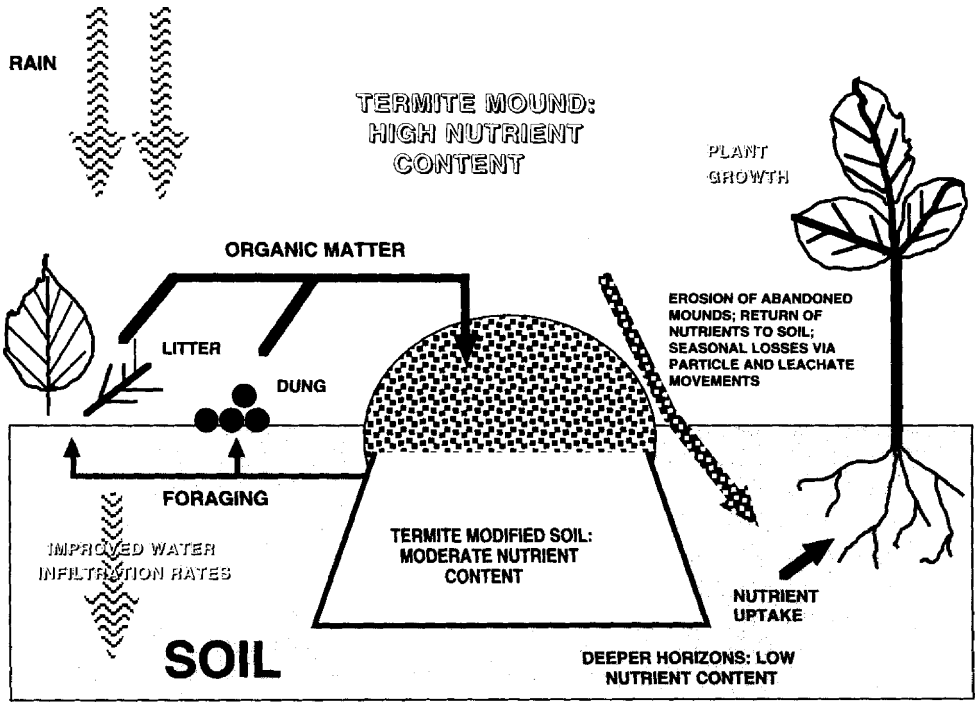


Figure 23 Schematic diagram depicting the influence of termites on plant growth. Reproduced from Conventry *et al.*, 1988.

shrubs that is quite distinct from that of the surrounding soils. This difference is considered due to protection from fire, improved drainage, greater soil depth, higher soil moisture, and improved soil fertility. Cation change capacities (CEC), exchangeable cations (Ca, Mg, Na, K), and base saturation decrease with increasing distance from the mound of *Macrotermes michaelseni* in the grassland of Kenya and available phosphorous and NO_3^- nitrogen are high near the mound. This causes the change of vegetation; maximum standing stock of plants occurs near the mound base and the number of plant species increases with increasing distance from a mound. **Figure 23** shows a schematic diagram depicting the influence of termite mounds on plant growth.

The termite mound supports a higher level of bacterial activity than adjacent top soil, especially of cellulose decomposers, and the activities of these microorganisms result in the release of nutrients into the mound soil. Many species of higher termites have developed to store the fragments of grasses and leaves in the mounds mainly in somewhat dry regions. The stored fragments are heavily colonized by microorganisms. The storage chambers may have the specific function of decomposition chambers and the process may be similar to the fungus gardens of the Macrotermitinae in that it provides the colony with organic material enriched in nitrogen.

Termites and Human Disturbances

The abundance and species diversity of termites tend to decrease when tropical forests and savannas are cleared and modified (**Table 7**). At first soil feeders are replaced by fungus growers and then by wood feeders in Africa and Asia. However, the destruction of natural ecosystems and their replacement by other ecosystems in some cases provides suitable habitats for some groups of termites, leading to an increase in their population. Zimmerman *et al.* (1982) were the first to draw attention to the topic in relation to the production of greenhouse gas by termites.

The global concentration of atmospheric methane, a greenhouse gas, is increasing in recent years at an annual rate of about 1%. This has been mainly attributed to paddy rice cultivation, enteric fermentation (mainly by ruminants), biomass burning, and fossil fuel consumption. Termites have high biomass and emit methane in many tropical ecosystems.

Much research has been done on the global estimates of termites, methane production by various groups of termites, oxidation of methane produced by termites in the soil, and the importance of methane production by termites on global climatic change. The rate of emission of methane by termites differs between species and feeding habits. Soil feeders tend to produce more methane than wood feeders on a

Table 7 Termite abundance and biomass in tropical regions (AFR, African; NEO, Neotropical; ORI, Oriental)

Ecosystems	*Region*	*Abundance nos. m^{-2}*	*Biomass g m^{-1}*
Primary or near-primary forest			
Riverine forest (Congo)	AFR	1000	ca. 11
Riverine forest (Nigeria)	AFR	2646	6.9
Semideciduous forest (Nigeria)	AFR	3163	8
Semideciduous forest (Cameroon)	AFR	2282–6957	8.31–123.2
Rainforest (Puerto Rico)	NEO	87–104	0–1
Rainforest (Amazonia)	NEO	1865	6.0–7.5
Rainforest (Pasoh, W. Malaysia)	ORI	3200–3800	ca.9
Wet evergreen (S. India)	ORI	2011–2449	3.62–3.72
Montane (Gunung Mulu, E. Malaysia)	ORI	38–295	0.01–0.7
Rainforest (dipterocarp, E. Malaysia)	ORI	779–1603	1.1–1.8
Rainforest (kerangas heath forest)	ORI	2271	ca.4
Secondary forest or tree plantation			
Terminalia plantation (1 yr) (Cameroon)	AFR	244–2959	0.43–19.89
Terminalia plantation (5 yr) (Cameroon)	AFR	5170–6703	10.98–35.13
Secondary forest (40 yr) (Cameroon)	AFR	2328–10,488	39.11–114.16
Tropical savanna			
Grass savanna (Central Africa)	AFR	612–701	1.3–1.9
Derived savanna (W. Africa)	AFR	861	1.7
Guinea savanna (W. Africa)	AFR	2966	3.6
Guinea savanna (W. Africa)	AFR	4402	11.1
Acacia savanna (S. Africa)	AFR	732	0.98
Broadleaved parkland (South Africa)	AFR	3300	8.43
Agro-ecosystems			
Grazed pasture (W. Africa)	AFR	2010	2.8
Maize (1 yr, W. Africa)	AFR	1553	1.7
Maize (8–24 yr, W. Africa)	AFR	6825	18.9
Sugar cane (W. Africa)	AFR	4800	5.4
Temperate scrubs			
Mediterranean scrubs (W. Australia)	AUS	27–45	0.97–1.93

biomass-specific basis. An important fact is that a large amount of methane produced by subterranean termites is oxidized by microorganisms in the soil. Thus the annual contribution by termites is estimated to be less than 20 Tg and probably less than 10 Tg (ca. 4% and 2% of global total from all sources, respectively).

Termites usually consume dead plant material and do not attack living plants in natural forests. However, some termites that are able to survive in agricultural ecosystems become harmful to crops (Harris, 1971). This tendency is stranger when the crops are exotic. A few hundred species of termites cause damage and only about 50 species are serious pests (Pearce, 1997).

See also: Amazon Ecosystems. Insects, Overview

References

Abe T (1987) Evolution of life types in termites. In: Kawano S, Connell JH, and Hidaka T (eds.) *Evolution and Coadaptation in Biotic communities*, pp. 125–148. Tokyo: University of Tokyo Press.

Abe T, Higashi M, and Bignell DE (eds.) (2000) *Termites: Their Symbiosis, Sociality and Global Diversification*. Dordrecht, Netherlands: Kluwer.

Behnke FL (1977) *A Natural History of Termites*. New York: Charles Scribner's Sons.

Higashi M and Abe T (1997) Global diversification of termites driven by the evolution of symbiosis and sociality. In: Abe T, Levin SA, and Higashi M (eds.) *Biodiversity*, pp. 83–112. Springer.

Higashi M, Abe T, and Burns TP (1992) Carbon-nitrogen balance and termite ecology. *Proc. R. Soc. London B* 249: 303–308.

Krishna K and Weesner FM (eds.) (1969) *Biology of Termites*, Vol. I, 598. New York: Academic Press.

Krishna K and Weesner FM (eds.) (1970) *Biology of Termites*, Vol. II, 643. New York: Academic Press.

Lee KE and Wood TG (1971) *Termites and Soils*. New York & London: Academic Press.

Wilson EO (1971) *The Insect Societies*. Cambridge: Harvard University Press.

Wood TG (1978) Food and feeding habits of termites. In: Brian MV (ed.) *Production Ecology of Ants and Termites*, pp. 55–80. Cambridge: Cambridge University Press.

Wood TG and Sands WA (1978) The role of termites in ecosystems. In: Brian MV (ed.) *Production Ecology of Ants and Termites*, pp. 245–292. Cambridge: Cambridge University Press.

J

Justice, Equity and Biodiversity

Kai MA Chan and Terre Satterfield, University of British Columbia, Vancouver, BC, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ecosystem services Conditions and processes through which natural ecosystems, and the biodiversity they comprise, sustain and fulfill human life. Examples include provision of clean water, maintenance of livable climates (carbon sequestration), pollination of crops and native vegetation, and cultural, spiritual, and intellectual inspiration.

Environmental justice Balanced distribution of costs and benefits among stakeholders in the design and implementation of environmental policies and projects.

Equity State, quality, or ideal of being just, impartial, and fair.

Justice The appropriate treatment of others; fairness in actions, rules, and outcomes.

Livelihood Possession of income and access to resources sufficient to meet basic life needs.

Restitution Act of restoring to the rightful owner something that has been taken away, lost, or surrendered; the act of compensating for loss, damage, or injury.

Introduction

The struggle to sustain biodiversity is a struggle for justice and equity, a struggle to treat others fairly where the others include not only existing people but also future generations and nonhuman organisms. But in an urgency to protect biodiversity, the interests of existing people – especially those who are physically proximate to protected areas – have often been transgressed. To guard against human injustices while also advancing the fair treatment of nonhuman organisms and future people, we need a coherent and shared understanding of the competing demands of justice.

The authors work toward a common understanding by discussing first the progress made in addressing the interests and justice-demands of existing and future people and nonhuman organisms, and then the current state of knowledge and debates regarding those demands. The authors then discuss how conservation actions have at times involved insufficient attention to the rights and interests of existing people, especially those surrounding or previously within the protected areas. In doing so, the authors draw upon case studies to bring life to the human impacts of the conservation actions.

What is Justice?

Aristotle clearly established that just and equitable treatment is not necessarily equal treatment, as differing circumstances call for differing treatments. Accordingly justice and equity are considerably more difficult to agree upon and to achieve, in part because a common understanding of merit – what a person deserves – has been elusive. To determine 'justice as fairness' in the treatment of people, Rawls (1999) has advanced the notion of the veil of ignorance, by which people are called upon to consider what distribution of circumstances they would prefer given that they imagine the experience of every person who might be affected by the decision. According to Rawls, this approach would better attend to concerns associated with the identities and number of people impacted by an action, which some argue and is ignored by utilitarianism. Like much of economics, utilitarianism seeks the greatest total goods regardless of the specifics of distribution. Utilitarianism may therefore imply preference for alternatives in which the disadvantaged are further disadvantaged, as long as the total gain is more. Utilitarians argue that the amount that people benefit from a given impact depends on their circumstances – for example, the poor benefit more from a dollar than do the rich – such that utilitarianism's aggregation of benefits sufficiently accounts for equity. Other ethicists counter that justice must go beyond an aggregation of what people want for themselves (personal preference) and also attend to what people believe constitutes a just society (principle) (Sagoff, 1998).

Like many other ethicists, Rawls only intended to consider the treatment of people – "rational beings with their own ends and capable of a sense of justice" but efforts to protect biodiversity have been fueled largely by a notion of justice that extends to and beyond the boundaries of the community of existing people. Conservationists argue that the community of beings toward whom we have responsibilities ('moral patients' in philosophy) should include future generations, nonhuman organisms, and perhaps even ecosystems.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00252-5

The authors do not defend a particular theory of justice. For our purposes it is sufficient to note that justice must involve a fair distribution of rights, responsibilities, costs, and benefits. And when actions have impacts with unfair distribution, justice requires appropriate restitution.

The implications of justice for moral acceptability is subjected to debate within ethics: utilitarianism assumes that an action is morally acceptable if injustices are outweighed by other gains, whereas other theories do not allow gains to offset injustices. Regardless, the just treatment of individuals has importance for all ethical theories.

Inclusion of Future People: Benefits of Conservation

The conservation of biodiversity is fueled in part by the recognition that actions taken by existing people for their own direct benefit have diffuse – but cumulative and consequential – impacts on existing and future people through the degradation of ecosystem services (the flow of benefits from nature to humanity). Often the loss of biodiversity and its associated benefits is a vehicle for injustice, as powerful capitalist or authoritarian forces sequester environmental benefits for themselves at the expense of those less empowered. Future people deserve special consideration in the context of biodiversity and its benefits because many losses of biodiversity are irreversible. Accordingly, conservation can and should ameliorate injustice and promote the equitable sharing of biodiversity's benefits across populations and generations. Efforts like the Millennium Ecosystem Assessment, 2005, which recognized the needs of people right from its conception, offer considerable progress for conservation in promoting socioeconomic justice.

Unlike existing people, who can be harmed physically or psychologically in the name of conservation, future people are affected primarily through benefits accrued and denied. Insofar as future people will inherit the harms experienced by their parents and grandparents, the authors consider these other dimensions of justice through the Section on Inclusion of Nonhuman Organisms.

Ecosystem services – as the benefits that people derive from ecosystems – provide a framework for the justice concerns of future people regarding biodiversity. Although the dependence on biodiversity of the supply and stability of individual ecosystem services varies considerably and is generally poorly characterized, ecosystem services nevertheless represent the best expression of the benefits of biodiversity. Ecosystem services can be grouped into five categories: provisioning, regulating, supporting, cultural, and option.

Provisioning

Degradation of ecosystems threatens the ability of existing and future people to provide for their material well-being through 'renewable' natural resources. The collapse and lack of recovery of fish stocks such as the northern cod demonstrate that ecological impacts can prevent some of these resources from being truly renewable. Recent research has revealed that we have removed the majority of the top predators from ocean ecosystems and are now harvesting less-desirable species lower in the food chain ("fishing down marine food webs") (Pauly *et al.*, 1998). Conservation activities attempt to address such impacts through reform of fishery management, new technologies (e.g., turtle excluder devices, low-bycatch longline practices), and marine protected areas (MPAs). Other natural resources are impacted similarly, such as forests, whose wholesale logging may trigger shifts to other forms of dominant vegetation, undermining the productivity of timber resources.

Regulating

Habitat destruction and climate change also diminish the ecological regulation of key biogeochemical processes and thereby the mitigation of damages to existing and future people. Loss of forests and wetlands impedes protection from river flooding; cutting and burning forests exacerbates climate change and its possible negative impacts; clearing of mangroves and loss of coral reefs and other coastal ecosystems undermine protection of coasts from flooding associated with tsunamis and hurricanes. Conservation activities aim to lessen these ecological insults through protected areas, best-management practices on private lands, ecological restoration or afforestation, and alternative land uses and livelihoods. Note that activities intended to mitigate one ecosystem service may have negative impacts on others: for example, afforestation in inappropriate (especially arid) locations may undermine water provision and harm biodiversity. Strategies such as biodiversity accreditation of carbon credits and management for multiple ecosystem services seek to redress these unintended consequences.

Supporting

Ecological degradation also undermines the services that are crucial for realizing value of other services, such as pollination, pest control, and prevention of soil erosion for agriculture. Because the implications of such service losses to future generations are not reflected in the functioning of current economic markets it is especially important that we plan specifically for such indirect, temporally distant effects. Some existing programs do attempt to limit or lessen soil erosion, such as the Conservation Reserve Program in the United States and the Grain-for-Green program in China. We need similar large-scale projects to protect populations of native pollinators and pest enemies.

Cultural

The loss of natural habitat and biodiversity gravely threatens the future contribution by ecosystems to the nonmaterial benefits people derive from human–environment interactions (including place value, cultural identity, and esthetic, artistic, scientific, and spiritual inspiration). Mostly to guard against this, numerous conservation nongovernmental organizations (NGOs) have targeted the protection of beautiful spaces and charismatic megafauna. Similarly, governments across the world have established protected areas that are more effective at protecting iconic but infertile landscapes than biodiversity,

and the World Conservation Union (IUCN) has established a network of World heritage sites for which cultural values are important contributors. To complement these various efforts, many governments have passed legislation protecting threatened and endangered species.

Option

Perhaps more worrisome than the threats to known services are the threats to ecosystems' provision of future values that we do not yet understand: the option values (recognized as a separate category of services by Daily *et al*. (2000) but not the Millennium Ecosystem Assessment (2005)). Plants like the rosy periwinkle and the Pacific yew were considered relatively unimportant until it was discovered that they are especially effective in treating childhood leukemia and other cancers for which medicines are lacking. Only by protecting a large component of biodiversity – including rare and inconspicuous organisms – can we sustain the options of future generations (Mangel *et al*., 1996). Such reasoning was a part of the motivation for the 1992 UN Convention on Biological Diversity, which commits nations to protect biodiversity within their borders and rich nations to assist less-developed countries in this task. Unfortunately, although much has been done, much has not – both within nations and between them.

Major unresolved philosophical issues impede the just treatment of future generations:

1. To what extent do we have responsibilities toward people whose identities depend upon our decisions (the future individuals paradox), we only need to ensure that their lives are worth living or do we have responsibilities beyond that (Parfit, 1984)?
2. How should we deal with shifting baselines (Pauly, 1995)? That is, if future generations grow up with less natural habitats or less-intact ecosystems, they may be satisfied with less than we might be, so how should we calculate our impacts on future generations?
3. To what extent are we responsible for the way that our actions might change the actions of future generations (impacting yet subsequent generations)? When changes are irreversible, most of us would agree that we are responsible for the impacts on all future generations. But if we degrade ecosystems and compromise the ecosystem services of future generations in a manner that is theoretically reversible, human psychology might nevertheless render it unlikely to regain the services, as restoration 'seems' much more onerous, uncertain, and supererogatory than protection. Should we apply a utilitarian approach to calculate the various expected direct implications of our actions or should we apply the logic of Immanuel Kant, according to which we should act only in a way that we would have all generations act? Although utilitarianism might seem to allow some generations to unsustainably harvest natural ecosystems and biodiversity, Kantian logic would seem to prohibit this.
4. What is an appropriate discount rate? To obtain mathematically bounded solutions and to reflect the expected improvement of material well-being, economists discount future costs, and benefits, assuming that foregone benefits will be more easily substituted and that committed costs will be more easily borne (Portney and Weyant, 1999). (Note that this discounting to account for the interest on investments does not entail discounting the well-being of future generations through 'time preference,' which most ethicists rebuke). How is well-being likely to change and how can we appropriately account for our expectations and our uncertainties?
5. What will future people value? What is the relative value of general wealth (converted natural capital) versus (unconverted) natural capital; is it permissible to replace naturally provided ecosystem services with human-engineered services? If so, when and to what extent? The crux of this question seems to be how much future people will appreciate the cultural values associated with biodiversity, some of which seem much more difficult, if not impossible, to replace.
6. What does sustainability entail, and is it necessary in a moral sense, sufficient, or neither? Related to the above, there is a debate regarding the demands of sustainability. Weak sustainability would require us to ensure that future generations enjoy at least the same quality of life (allowing for substitution between natural resources and various other forms of capital), whereas strong sustainability would mean that future generations would enjoy the same level of natural resources. By Rawls' just savings principle, each generation would maintain its just institutions and preserve its material base and also commit savings beyond that, apparently requiring strong sustainability as necessary but not sufficient. However, utilitarianism seems not to require even weak sustainability: if we could benefit many future generations by slightly impoverishing the very next generation, utilitarianism seems to require this.
7. How might we think of justice not only as equity in distribution, but equity in the processes through which such decisions are made?

Inclusion of Nonhuman Organisms

Without necessarily subscribing to a particular theory of justice, many people feel that humanity's collective treatment of nonhuman organisms is deeply unjust and inequitable. Indeed, it is hard to imagine a theory of justice that recognizes the moral standing of any nonhuman organisms but that finds no fault in the pervasive loss of habitat and populations and the impending global extinction of some significant proportion of Earth's species. Widespread agreement on this injustice has contributed greatly to the major progress, the conservation movement has made. In recent decades, the protected areas network has grown considerably to 12% of the global land-base, thanks to the efforts of conservation NGOs and nation states. Many nations have signed the Convention on Biological Diversity and the Convention on International Trade in Endangered Species and numerous other conservation-related international agreements. On the downside, however, the existing protected areas are disproportionately situated on steep slopes with unproductive soils, in low-diversity high-latitude regions. And those that exist on paper are often not adequately enforced or managed,

such that species are routinely extirpated within their boundaries. Furthermore, protection of the marine environment has just begun, and freshwater organisms have been decimated. The imperiled state of biodiversity suggests that more remains to be done to ensure just treatment of nonhuman organisms.

Yet crucial issues remain to be resolved regarding the requirements of justice to nonhuman organisms, despite the considerable attention to environmental ethics (Primack and Cafaro, 2001):

1. There are numerous perspectives on which organisms and entities should have moral standing within a theory of justice. And what is the source of responsibilities we might have toward the nonhuman world? Nonhuman organisms, species, and ecosystems differ from people in a number of ways such that it is unclear that the responsibilities we have toward other people ought to be extended to these organisms. But if responsibilities to others (moral standing) do derive from some notion of sentience or consciousness (a capacity to feel or think), the uncertainty associated with nonhuman organisms (whether they possess that key characteristic) may itself incur moral responsibilities (Chan, in press).
2. What would constitute just treatment of those entities that merit moral standing (1)? What should we aim to protect, to what extent, and at what scale? Much conservation funding gets funneled through large international NGOs, whose actions seem largely aimed to conserve species richness at the global scale. Although both the Nature Conservancy and World Wildlife Fund have organized their conservation efforts by ecoregions, attention is focused on globally threatened species and ecoregions with many such species. Whereas , many governments at various scales have passed laws requiring conservation of species (and in some cases, subspecies) that are threatened nationally or regionally, though not globally. Many of these laws are intended to prevent extinction entirely, but many only require measures that have sufficiently small costs. Does the taxonomic level of species merit such a dominance of attention, or do populations and other levels deserve more concern than they currently receive?
3. What is the relevance of naturalness for the demands of justice? How important is it to protect species in their native surroundings and with their historic biological communities? And what should we do in the face of climate change, which is already shifting species distributions? How should we balance the needs of individual organisms with those of species and populations, and are the justice requirements different for domesticated or invasive species? These are crucial issues for ecological restoration, conservation, invasive species management, and environmental impact mitigation.

These various questions come to the fore in many conservation actions, such as The Nature Conservancy's much-publicized efforts to protect the Santa Cruz Island fox (and numerous endemic plants). This charismatic mammal is globally endangered, threatened most immediately by predation from golden eagles that were attracted by the presence of introduced feral pigs. To the outrage of some animal welfare activists, The Nature Conservancy has been shooting the feral pigs to help dissuade the golden eagles from nesting on the island. This conservation action seems to reflect a valuation of species and naturalness over the welfare of individual animals such that conservation gains outweigh the injustices to the pigs. It is clear that there are still components of society willing to debate the appropriateness of this ethical stance.

Just Attention to Existing People: The Consequences of Conservation

If we accept that justice includes fairness in action, rules, and outcomes, then we must ask the following: are the same parties who benefit from conservation efforts those who bear the costs? And if not, what particular inequities must be attended to in the quest for greater justice in the protection of biodiversity? Currently, as areas are committed to biodiversity protection or are managed ever more strictly, extant communities' rights and access to land are often identified as primary threats to the integrity of the reserves. This threat-identification often occurs without consideration of issues crucial for just outcomes including livelihood needs in the face of a shrinking land-base, the availability of economic alternatives, and prior events such as the evacuation of people from the territories in question in the colonial period. Tension between human and biological priorities is especially strong when affected populations are primarily land- or sea-based (those with pastoralist, subsistence-agricultural, or fishery-based economies) and economic alternatives are nascent or elusive (West and Brechin, 1991; Leitmann, 1998; Salafsky and Wollenberg, 2000). We must also remember that indicators of biodiversity and species health are not solely status reports on physical conditions. They are imbued by values that are the products of a modern industrial world that seeks new markets in the form of wilderness experiences at same moment that it has come to recognize the loss of the 'frontier' and "nature" and now aspires to curtail that loss by restricting the activities of others (Igoe et al., 2010).

Ultimately, most of the human societies value and work to preserve some aspects of biodiversity, and such efforts may be sufficient without outside involvement (Schwartzman *et al.*, 2010). Nonetheless, restrictions imposed by conservation efforts often involve imposing a postindustrial notion of value on those who are largely outside modernism's beneficial reach. We also cannot simply leave the recognition and restitution of injustices to those who raise its specter because the capacity to resist unjust practices is a function of power such that those with a greater ability to articulate their grievances are also those most likely to achieve desirable outcomes.

Distribution of Risks by Class, Gender, and Ethnicity

Our call for justice for human communities in conservation contexts follows the general principle of environmental justice. Poverty advocates argue that the costs of biodiversity protection are unfairly concentrated in the poorest regions of the world and can exacerbate poverty among women and other disadvantaged parties, and that when conservation

benefits do accrue, their local distribution is disproportionately secured by local elites. Thus, distributional injustice expresses itself both within and between communities at a variety of scales. Globally, the overrepresentation of biological hotspots in poor areas (e.g., the tropical Andes, Mesoamerica, Madagascar, and Brazil's Atlantic rainforest) has placed an unprecedented burden of expectation on nation states least equipped to meet conservation goals (Balmford and Whitten, 2003). This inequitable distribution of biodiversity hotspots has the potential for even greater injustice because the beneficiaries of the extractive industries that threaten biodiversity in such places are often not those who suffer its costs, and are often from comparatively wealthy populations in the developed world. Nationally and regionally, it is difficult for poorer nations to resist the inflow of foreign currency promised by an ecotourism industry enabled by protected areas – just as it is difficult to resist monetary benefits promised by extractive industries. This is the very income necessary to build infrastructures of protection and regulation and to reshape a nation's land-base toward conservation objectives, but the sustainability of such infrastructure and reshaping depends upon a base level of well-being among local people that is often lacking. It has been demonstrated too infrequently that the wealth produced by conservation-based development is redistributed to those human communities that are most affected by such land-use change.

The case of Madagascar's Ranomafana National Park offers a particularly alarming example of these distributional injustices (Harper, 2002), making it illustrative but not typical. Home of the golden bamboo and greater bamboo lemurs, the park, was established in the mid-1990s with an ambitious plan to reorient the economic structure of 26 adjacent villages. The intent to establish a park that also bettered the lives of local people was certainly active from the earliest stages. The intended benefits included the promise of health care to compensate for the loss of land and its anticipated accompanying local storehouse of medicinal plants; it also involved the substitutions of 'swidden' or dry rice cultivation with more intensive wet rice cultivation in the hope that this would produce greater yields and so both staple and market-ready grain. In fact, sustained health care was only offered in the form of birth control to minimize population growth and a very limited acute care wherein a traveling physician was assigned to 60,000 residents made occasional visits to dispense limited medicines. Health status among the local population deteriorated more generally both because of diminished land-based income and because the vision of a renaissance of traditional plant-based medicines (in the absence of alternatives) was more of a conservationist fantasy than a contemporary or a historical reality. For the vast majority, the imagined income from crop intensification did not materialize. Problems were exacerbated by the expansion of Betsilo peoples into remaining Tanala land, where the former possessed the necessary resources to take over the irrigated fields of the impoverished Tanala. Further, those who were able to cultivate irrigated rice were sanctified by conservation agents as progressive and modern, whereas long-standing hillside (dry) cultivation was deemed a quaint but destructive tradition or a 'cultural' (namely 'Tanala') predisposition unworthy of economic and social support.

The park itself has also led to an increase in local poverty as a function of pre-existing gender and caste-ascribed positions. The constant and often all-consuming labor owed in exchange for the use of rights for wet rice cultivation is often the basis for indebtedness to better-off landowners. This indebtedness most severely affects older single women who do not have male household members to fulfill these economic obligations and whose status is further compromised by reduced access to swidden (dry-cultivation) fields due to the park. Only a very small set of local elites benefited economically from the park. Not coincidentally, these few were instrumental in the park's genesis and did not have to relinquish land in the formation of the park. Most lost the right to future claims to forested land, but again the elites recovered more easily and fully by absorbing the land belonging to others – as that land was forced into sale – or by using the land of those who had become desperate and so would accept minimal rental sums (e.g., US$12.50 for a three-year period). Justice in the above and other parallel cases requires that we remain diligent about the weighting of conservation's costs and benefits, asking repeatedly, whom does this act benefit and for whom does it cost, particularly at local scales? And, how can sustainable systems of restitution be brought into place quickly?

Distribution of Action, Responsibility, and Power

Anthropologists have also pointed to the problematic nature of the neoliberalisation of conservation. Broadly stated this refers to the conversion of most decisions to those legitimized by market success and measures, thus tending toward a commodification of all of nature's flows or services. But more critically, it also refers to the fact that some of the largest funders of major conservation NGOs, which are responsible for a great and growing proportion of conservation worldwide, are companies implicit in the transformation of ecosystems to intensive use (Chapin, 2004). Although we recognize that good citizenship and hazardous corporate or industrial practices might well co-exist, the danger is that corporate-NGO alliances can also mask or lead to a downplaying of heinous practices on behalf of corporate entities. An illustrative example (Igoe *et al.*, 2010) is McDonalds' Endangered Species Happy Meals campaign. "Although awareness among children of biodiversity loss may well prevail as a result, the environmental contradictions of the fast food industry are far more pervasive and integrated than a Hotspots and Happy Meals partnership can even begin to capture. The rise of the fast food industry in North America is inextricably linked to automobile culture, expanding networks of superhighways, and the industrialization of global food production systems'' (ibid., p. 505).

Such collaborations between those NGOs who act to protect biodiversity and those companies responsible for losses of biodiversity deserve scrutiny. On the one hand, who better to fund biodiversity conservation than those who bear responsibility for the motivating problem? This is the logic of the 'polluter pays' principle. On the other hand, there are dangers that (i) the attention to the voluntary positive actions will distract public attention from behavior and actions that degrade biodiversity (Igoe *et al.*, 2010); (ii) conservation

organizations will have an incentive to turn a blind eye on any such behavior and actions, lest they bite the hand that feeds them (Chapin, 2004); (iii) the governance principles and mindsets that fuel a preoccupation with economic growth – which have largely been adopted by major conservation NGOs – will remain unchallenged despite their possible roles in creating the problems at hand (Corson, 2010); (iv) both the extractive industries and now their conservation – through the restriction of access to wild living resources – may represent competing major threats to the autonomy and sovereignty of people who depend on ecosystems for their livelihoods and well-being.

Reflections on the Distribution of Values: Cultural Sovereignty versus Cultural Naiveté

A justice lens on human communities also requires us to recognize that all conservation projects are also projects of social engineering: We need people to abide by new rules, change what they do, alter their livelihoods, and so on. Although social and biological change is impossible without these engineered consequences, there is a sustained failure among many conservation efforts to comprehend the injustice of the under-recognized clash of values. Biodiversity is not a value-neutral construct despite the widespread desire to avoid wanton destruction and maintain the health of needed ecosystems. By now biodiversity discourse is too deeply embedded in concomitant notions of 'ideal natures,' 'good' and 'bad' cultural practices, and utopian visions (e.g., the hoped for discovery of plant pharmacopias or the protection of primordial nonhuman landscapes). Comprehending the clash of values requires consideration of the implicit prescriptions of ideal behavior that accompany conservation projects and thus become the basis for the apportionment of benefits among the local population. Most extreme is the complete exclusion of local people in the pursuit of idealized 'prehistoric' landscapes as realized not through joint intent but 'fortress conservation' (Brockington, 2002). Subtler examples of culturally naïve expectations are that hunting with spears is acceptable – even culturally authentic – whereas hunting with guns is not, and that harvesting medicinal plants is viable whereas foraging for sustenance is not. Justice can only be served by recognizing our assumptions, their historical roots, and the judgments of desirable behavior that follow. Only then does it become possible to appropriately negotiate values in the prescription of acceptable practice. Goals regarding biodiversity outcomes – which fuel the conservation-derived prescriptions of behavior – are clearly a key component of conservation action; but these too require collaborative negotiation, adaptive management, and some hybridization of local and nonlocal practices, knowledge, and values (Chan *et al.*, 2007).

The Distribution of Rights, Title, and Access

Common property scholars have long attended to the importance of rights, expressed as control over the products and decision making pertaining to local resources. More recently, access has become the central point as this concept highlights the broader social contexts that permit or deny a person the ability to benefit from 'things,' be they material or symbolic resources (Ribot and Peluso, 2003). Further, access to resources is fundamental to one's capacity to achieve and maintain a land- or sea-based livelihood, which is itself embedded in a complex web of cultural and social arrangements. Thus, the distribution of access to resources is central to theorizing justice although the processes underlying this distribution are less visible than is the assignation and 'evidence' of legal title. Use rights are often assigned through oral contract or customary practice, and it may also be mediated through lineage ties, village or religious leaders, or normative pressures. Knowledge of these dynamics is not readily available to conservation agents or even expressed by local people, for whom it has long become an inarticulable common sense. Yet use and access is also tied to systems of power, which may themselves be unjust (Asfaw & Satterfield, 2010). Changing access to achieve conservation is hazardously conducive to blind injustice in that something as seemingly simple as the individualization and legal titling of access rights to property (also a feature of aid and development agendas, and neoliberalism generally) may involve a fundamental shock to the social system.

Even where a system of individualized property rights is firmly established, individuals' rights to take actions that degrade ecosystems and threaten biodiversity are highly contentious. For example, in the United States there is a strong anti-Endangered Species Act property rights movement, which claims that the federal government has no business dictating what a landowner should or should not do with his or her own land. Is it just to force property owners to bear the costs of conservation, just because they were 'unlucky' enough to host endangered species? Such perceived injustice has fueled the common practice of "shoot, shovel, and shut up" for endangered species, because landowners can often escape responsibility when endangered species are not shown to inhabit the property in question. In recognition of these equity concerns, the NGO Environmental Defense and the US Fish and Wildlife Service have instigated several programs that alleviate the burden on landowners (Wilcove and Lee, 2004). In other countries such as Canada, the United States experience has led to a shying away from endangered species legislation pertaining to private lands.

Colonial Legacies and Restitution

Throughout the developing and developed world there exists the undeniable legacy of colonialism in the conception of biodiversity and its conservation. For instance, many national parks were established for the benefit of colonial administrations but have since become important cornerstone landscapes for the benefit of biodiversity. Yet, the populations evacuated from such sites have both living memories and oral records of these injustices. And such injustices are often the basis in the contemporary period for severely degraded relations between the nation state and conservation NGOs, as well as for intertribal or intercommunity conflict. The iconic Serengeti National Park in northern Tanzania offers a case in point (Neumann, 1998). The colonial period itself saw increased disruption of customary land use and practices in the

form of restricted access to hunting and forest resources. By the end of the colonial period, ordinances pertaining to the new network of national parks went so far as to exclude traditional and land-based people entirely. "Under this ordinance, the Tanganyika National Parks become for the first time areas where all human rights must be excluded thus eliminating the biggest problem of the Trustees and the Parks in the past" (Neumann, 2000, p. 125). Serengeti National Park was less a product of the local colonial administration, which recognized the hardship such parks would generate, than of international conservation organizations operating in the 1920s and 1930s, whose vision of what Africa should be was a mythical unspoiled wilderness, rhetorically explained as an effort to save species from extinction. In fact, the displaced (particularly Maasai pastoralists) watched the beneficiaries parade by with prized game whereas poverty and intertribal conflict came to characterize those on the periphery. In a 1957 memo, a Maasai observed, "From time to time we see (white) hunters passing with the trophies of animals that they have shot. It is the same people and their friends who wish to evict us from the National Park, yet we think it is they who are the enemies of game rather than us" (quoted in Neumann, 1998).

In settler nations with significant Aboriginal populations, such as the United States, Canada, New Zealand, Australia, Brazil, Chile, and so on, many heretofore unresolved Aboriginal treaty obligations are closely tied to the establishment of national parks. Some parks were used to justify placing Aboriginal populations on reservations. More recently, some parks have become a means of settling conflicts over traditional territories, but in doing so they also confine Aboriginal populations to limited use of those areas as prescribed by conservation objectives. In South Dakota, the Lakota Sioux have been pushed aside for a series of mining interests and are now faced with the possibility that the land left aside by the extractive industries will be managed for conservation, thus continuing and deepening injustices. The Sioux have not accepted compensation monies, and instead seek the recovery of treaty lands, the recovery of the bison, and the restoration of mountains they deem sacred. The dispute has unfortunately pitted recreational users of the area against the Sioux Nation and has produced such acrimony that the council of the nation "describes the United States as an imperialistic power that has sought to undermine, diminish and abolish not only the Lakota land-base but Lakota rights to self-sufficiency and self-determination" (Halder, 2003, pp. 105–107).

The failure to recognize past injustices often impedes successful community-based conservation, as does the failure to establish meaningful local control over the governance of a conservation area (Timko and Satterfield, 2008). And it is difficult at best to sort out the historical contingencies that invariably complicate any new conservation effort (including which populations or lineages have decision authority over particular parcels or goods derived from within these) (West, 2006). At the very least, justice will require a vastly deeper historical understanding of the local trajectories and consequences of both colonial legacies and associated programs of human evacuation. Such an understanding will enable a vastly improved system of recompense, where the primary goal of negotiation is not solely cash for land, but the recognition of the legitimacy of historical claims and the provision of substantive decision-making control over future uses, regulation, and access to benefits.

Conservation efforts are further complicated by histories of war and genocide. In such postconflict societies, both the land and the people are often traumatized, and the nation states in which they are embedded tend to be comparatively unstable. Central Africa, home to some of the remaining populations of great apes, offers cases in point. Such settings pose a particularly vexing set of problems as concerns justice. In some cases, poverty and sustained human degradation have fueled desperate measures, including the consumption of great apes and other endangered species as "bushmeat," and the further degradation of any available resources. And yet, no one (human or otherwise) will be served in the long run by equally desperate measures to secure conservation by continued force.

Conclusions

Towards the end, the authors agree with Zerner (2000) that biodiversity conservation depends on the development of solutions that are socially just and ecologically informed. As a society, we should strive for sustainable economic development wherein biodiversity conservation is pursued – for its global and future benefits and for its own sake – in harmony with the cultural, social, and economic well-being of local peoples.

See also: Conservation Efforts, Contemporary. Conservation Movement, Historical. Ecosystem Services. Environmental Ethics. Ethical Issues in Biodiversity Protection. Indigenous Peoples and Biodiversity. Poverty and Biodiversity. Property Rights and Biodiversity. Stewardship, Concept of. Sustainability and Blodiversity

References

Asfaw T and Satterfield T (2010) Beyond local justice: gender relations in local-level dispute settlement in Ethiopia's Zehie Peninsula. *Human Ecology Forum* 17: 160–174.

Balmford A and Whitten T (2003) Who should pay for tropical conservation, and how could the costs be met? *Oryx* 37: 238–250.

Brockington D (2002) *Fortress Conservation: The Preservation of the Mkomazi Game Reserve, Tanzania*. Bloomington: Indiana University Press.

Chan KMA (in press) Ethical extensionism under uncertainty of sentience: Duties to nonhuman organisms without drawing a line, environmental values.

Chan KMA, Pringle RM, Ranganathan J, *et al.* (2007) When agendas collide: Human welfare and biological conservation. *Conservation Biology* 21: 59–68. Available online at: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1523-1739.2006.00570.x

Chapin M (2004) A challenge to conservationists. *World Watch* 2004: 17–31. Available online at: http://www.eldis.org/static/DOC18110.htm.

Corson C (2010) Shifting environmental governance in a neoliberal world: US AID for conservation. *Antipode* 42: 576–602. Available online at: http://dx.doi.org/10.1111/j.1467-8330.2010.00764.x

Daily GC, Soderqvist T, Aniyar S, *et al.* (2000) The value of nature and the nature of value. *Science* 289: 395–396. Available online at: http://www.jstor.org/stable/3077353

Halder B (2003) 'Ecocide and genocide': Explorations of environmental justice in Lakota Sioux Country. In: Anderson DG and Berglund E (eds.) *Ethnographies of Conservation: Environmentalism and the Distribution of Privilege*. New York: Berghahn Books.

Harper J (2002) Endangered Species: Health. *Illness and Death among Madagascar's People of the Forest*. Durham, NC: Carolina Academic Press.

Igoe J, Neves K, and Brockington D (2010) A spectacular eco-tour around the historic bloc: Theorising the convergence of biodiversity conservation and capitalist expansion. *Antipode* 42: 486–512. Available online at: http://dx.doi.org/10.1111/j.1467-8330.2010.00761.x

Leitmann J (1998) Options for managing protected areas: Lessons from international experience. *Journal of Environmental Planning and Research* 41: 129–145.

Mangel M, Talbot LM, Meffe GK, *et al*. (1996) Principles for the conservation of wild living resources. *Ecological Applications* 6: 338–362.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington, DC: Island Press.

Neumann RP (1998) *Imposing Wilderness: Struggles Over Livelihood and Nature Preservation in Africa*. Berkeley: University of California Press.

Neumann RP (2000) Land, justice, and the politics of conservation in Tanzania. In: Zerner C (ed.) *People, Plants, and Justice: The Politics of Nature Conservation*. New York: Columbia University Press.

Parfit D (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Pauly D (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology and Evolution* 10: 430.

Pauly D, Christensen V, Dalsgaard J, Froese R, and Torres F (1998) Fishing down marine food webs. *Science* 279: 860–863.

Portney PR and Weyant JP (eds.) (1999) *Discounting and Intergenerational Equity*. Washington, DC: Resources for the Future.

Primack RB and Cafaro PJ (2001) Environmental ethics. In: Levin S, Daily GC, Colwell RK, *et al*. (eds.) *The Encyclopedia of Biodiversity*. San Diego, CA: Academic Press.

Rawls J (1999) *A Theory of Justice*. Cambridge, MA: The Belknap Press of Harvard University Press.

Ribot JC and Peluso NL (2003) A theory of access. *Rural Sociology* 68: 153–181.

Sagoff M (1998) Aggregation and deliberation in valuing environmental public goods: A look beyond contingent pricing. *Ecological Economics* 24: 213–230.

Salafsky N and Wollenberg E (2000) Linking livelihoods and conservation: A conceptual framework and scale for assessing the integration of human needs and biodiversity. *World Development* 28: 1421–1438.

Schwartzman S, Alencar A, Zarin H, *et al*. (2010) Social movements and large-scale tropical forest protection on the Amazon frontier: Conservation from chaos. *The Journal of Environment and Development* 19(3): 274–299. Available online at: http://jed.sagepub.com/content/19/3/274.abstract

Timko J and Satterfield T (2008) Seeking social equity in national parks: Experiments with evaluation in Canada and South Africa. *Conservation and Society* 6: 238–254. Available online at: http://www.conservationandsociety.org/article.asp?issn=0972-4923;year=2008;volume=6;issue=3;spage=238;epage=254;aulast=Timko

West P (2006) *Conservation is Our Government Now: The Politics of Ecology in Papua New Guinea*. Durham, NC: Duke University Press.

West PC and Brechin SR (eds.) (1991) *Resident Peoples and National Parks: Social Dilemmas and Strategies in International Conservation*. Tucson: University of Arizona Press.

Wilcove DS and Lee J (2004) Using economic and regulatory incentives to restore endangered species: Lessons learned from three new programs. *Conservation Biology* 18: 639–645.

Zerner C (ed.) (2000) *People, Plants, and Justice: The Politics of Nature Conservation*. New York: Columbia University Press.

K

Keystone Species

Bruce A Menge and Alison C Iles, Oregon State University, Corvallis, OR, USA
Tess L Freidenburg, California Ocean Science Trust, Oakland, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Bruce A. Menge and Tess L. Freidenburg, volume 3, pp 613–631, © 2001, Elsevier Inc.

Glossary

Dominant species Species that owe their influence to their high abundance. Such organisms account for most of the biomass in a community, and thus are the primary components of community structure. Trees in forests, mussels in rocky intertidal habitats, grasses in grasslands, and kelps or corals in near-shore subtidal habitats are all dominant species.

Interaction webs (functional webs) Subset of species that through their interactions and responses to abiotic factors make up the dynamic core of food webs or communities. These webs include keystone species, dominants, and other strong interactors.

Key-industry species Prey of intermediate trophic status that support a large group of consumers.

Keystone species Consumers that have a large effect, and one that is disproportionately large relative to their abundance, on communities and ecosystems. Uniquely, the strong effects of keystone species on their interacting species exert extensive influence, often indirectly, on the structure and dynamics of communities and ecosystems. They are a distinct subset of a more broadly defined set of "strong interactors" that also include species having strong effects on interacting populations but not necessarily on communities or ecosystems. Keystone species can include predators, parasites, pathogens, herbivores, pollinators, and mutualists of higher trophic status, but generally are not plants, sessile animals, or "resources."

Strong interactors (foundation species) Species that have a large effect on the species (one or a few) with which they interact. Communities and ecosystems may have many strong interactors, and such species may occur at all trophic levels. "Strong interactors" is a more general term that can include keystone species, but not all strong interactors are keystone species. A similar idea is "foundation species," defined as the group of critical species whose effects and interactions define much of the structure of a community.

Weak interactors Species that have little effect on other species, at least under average conditions. Under some circumstances, weak interactors may occupy important roles in ecological communities as a result of changes that lead to temporary increases in their abundance, size, or biomass.

Historical Perspective

Identification of *Pisaster ochraceus* as a Keystone Species

In 1966, Paine reported the results of a study of the role of predatory sea stars in structuring marine intertidal communities on a rocky shore on the outer coast of Washington State. In temperate regions, such communities typically display a striking spatial pattern called "zonation." The hallmark of this pattern is that dominant space-occupying organisms are arranged in vertically stacked horizontal bands. For example, wave-exposed rocky temperate shores can have a band of barnacles on the upper shore, a band of mussels on the middle shore, and a band of macroalgae on the lower shore. On the Washington coast, the macroalga-dominated "low zone" also harbors various mobile and sessile invertebrates, including limpets, chitons, sea urchins, whelks, anemones, and, significantly, the sea star *Pisaster ochraceus*. An intriguing feature of the zonation pattern was the sharp demarcation between the mussel bed and the low algal zone. Observing that the low-zone-occupying sea stars fed regularly on mussels, and the only abundant source of mussels was higher on the shore in the middle zone, Paine postulated that predation by sea stars determined the sharp lower limit to mussels. Specifically, he speculated that at high tide, sea stars moved up from their low-tide resting places in the low zone to remove prey from the lower edge of the mussel bed, then retreated back to the low zone where the prey were consumed.

Paine tested the hypothesis that sea star predation determined the lower limit of the mussels by periodically removing *Pisaster* from a section of a rocky outcropping and comparing this section to an area in which sea stars had been left at natural densities. Within 3 years, the mussel *Mytilus*

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00081-2

californianus had increased in abundance, causing a reduction in local species richness of macroorganisms from 15 to 8 species (Paine, 1966). Ultimately, after 10 years of excluding sea stars, a single-species monoculture of *M. californianus* dominated the shore (**Figure 1**). Ancillary studies suggested that this change depended on two important mechanisms. First, *Pisaster* preferentially fed on mussels. This sea star actually will feed on many invertebrate species, but given a choice, it selects mussels. Second, mussels were dominant competitors for space. Field observations, and later experiments, indicated that mussels could displace all other occupants of space, including macrophytes, sessile invertebrates, and mobile invertebrates, by crowding them out. In doing so, mussels also created habitat for a sharply different set of invertebrate cryptofauna that occupied the byssal "forest" (the fibers that mussels use to attach to the rock) beneath the mussels.

Paine later repeated this experiment on a nearby coastal island (Tatoosh Island) with essentially identical results, further demonstrating that under natural conditions *Pisaster* indirectly enhances the persistence of other space occupiers by preventing mussels from invading the low intertidal zone. The striking difference in species composition and physical appearance of the low zone community wrought by the sea star, a community member that was nowhere close to being "dominant" either numerically or in biomass, was the basis for Paine's identification of *Pisaster* as a keystone predator. Note that contrary to some uses of the concept, the original definition did not comment on keystone species as determinants of species diversity. Although the changes in species diversity that occurred in Paine's (1966) field experiments revealed an important role of such species, Paine and other early workers on the topic (Estes *et al.*, 1978) evidently did not consider these changes to be fundamental to the concept of keystone species.

The formulation of the keystone species concept contributed to a gradual shift in ecological thought. Throughout the 1960s and 1970s, most ecologists believed that competition was the most important process structuring communities. By the late 1960s, advances in ecological theory and modeling, in conjunction with comparative and experimental field studies, demonstrated that processes other than competition could be important. Experiments such as those described above were among the first to show how predation can dramatically affect community structure. Following Paine's pioneering work, ecologists began to identify keystones in other systems.

Are There Others? Keystone Predators Elsewhere

One of the first documented examples of a keystone predator in another system was the sea otter, *Enhydra lutris*. Estes *et al.* (1978) observed subtidal communities in the Aleutian Islands, Alaska, where sea otters were abundant and differed

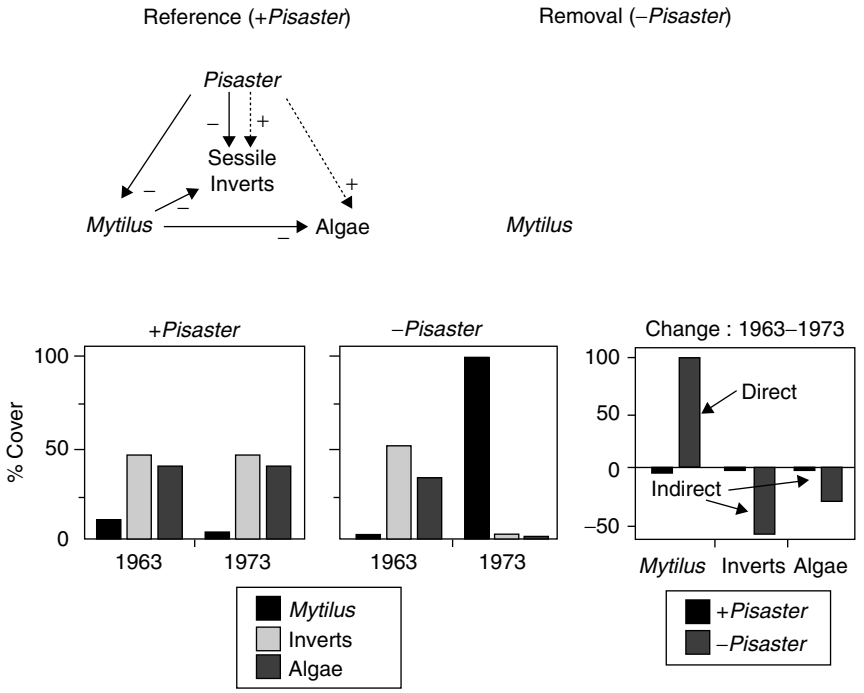


Figure 1 Results of *Pisaster* removal experiments. Upper diagrams: "Reference" shows the natural interaction web in the low zone, and "Removal" shows the interaction "web" (i.e., *Mytilus*) 10 years after removal began. Arrows point to the species or group affected. Solid arrows show direct effects; dotted arrows show indirect effects. "Minus" by an arrow indicates that the effect was negative; "plus" by an arrow indicates that the effect was positive. Lower diagrams: Left and center show covers of *Mytilus*, other sessile invertebrates (barnacles, anemones), and algae in 1963 (beginning of experiment) and 1973 in the presence and absence of *Pisaster*. Right diagram shows difference between covers in 1973 and 1963 for each group of sessile organisms. "Direct" indicates that increase in mussel cover was due to direct predation by sea stars. "Indirect" indicates that decreases in sessile invertebrates and algae were an indirect consequence of sea star predation, through competitive elimination of sessile invertebrates and algae by mussels. Data from Paine, 1974, cited in Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *Bioscience* 46: 609–620.

from those in nearby areas where sea otters had been locally extirpated by hunting. In waters surrounding islands with sea otters, kelp forests at shallow depths dominated nearshore subtidal communities, and within these kelp forests sea urchins, an important consumer of kelp, were relatively small and urchin biomass was low (**Figure 2**). On neighboring islands without sea otters, however, sea urchins were large, urchin biomass was high, and kelps were absent. On the basis of these data and the observation that sea otters preferentially feed on sea urchins, Estes and coworkers postulated that sea otters are a keystone species in this community. They argued that predation by sea otters reduced sea urchin grazing, thereby maintaining kelp forests and associated species. By documenting the "keystone predator" phenomenon in a different ecosystem, this research bolstered the view that such dynamics were potentially widespread.

Subsequent work supported this hypothesis and expanded our understanding of the conditions facilitating a keystone role for sea otters. In a large-scale study published in 1995, Estes and his colleague David Duggins determined the spatial generality of sea otter effects and tested the prediction that after sea otters colonize new areas, sea urchin-dominated subtidal communities would become kelp-dominated (Estes and Duggins, 1995). Surveys in the Aleutians and in southeast Alaska showed that the differences associated with sea otter presence or absence in the earlier studies were general in space – kelps were abundant in the presence of sea otters and scarce in the absence of sea otters. Further, considering all surveyed sites, sea urchin and kelp abundance were strongly inversely correlated (**Figure 3**). Finally, sea urchin abundance declined sharply at sites invaded by sea otters but did not change at sites where sea otter abundance remained constant. An interesting regional difference was that rates of increase in kelp abundance in the presence of invading sea otters were higher in southeast Alaska. Estes and Duggins suggested that higher rates of kelp increase resulted from differences in the recruitment of sea urchins, which was much greater in the Aleutians than in southeast Alaska (1995). Coupled with sea otter preference for larger, adult urchins, higher recruitment rates of

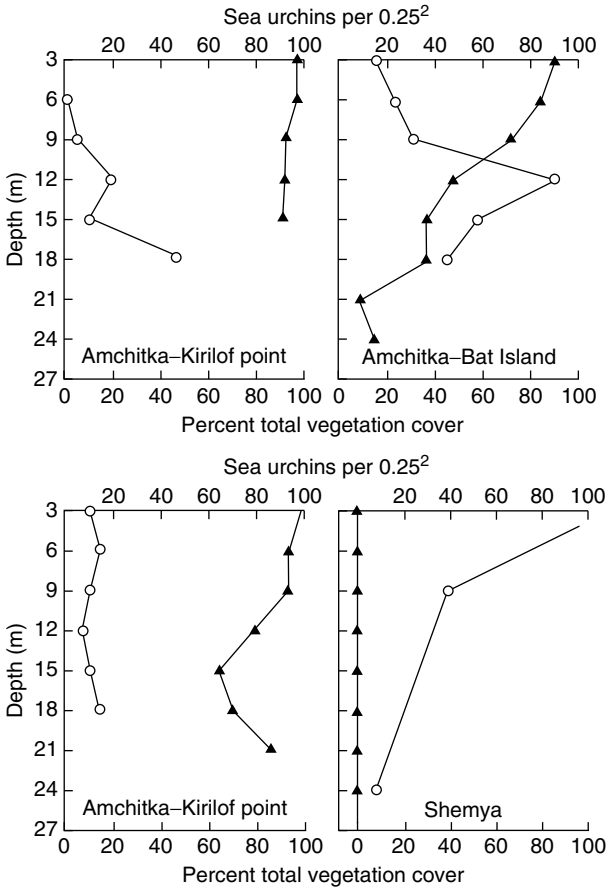


Figure 2 Total cover of macroalgae and sea urchin density by depth at three locations on Amchitka Island – Kirilof Point, Bat Island, and Kirilof Rocks (all with sea otters) – and one at Shemya Island (without sea otters). ▲, vegetation cover; •, sea urchin density. Reproduced from Estes JA, Smith NS, and Palmisano JF (1978) Sea otter predation and community organization in western Aleutian Islands, Alaska. *Ecology* 59: 822–833, with permission from ESA.

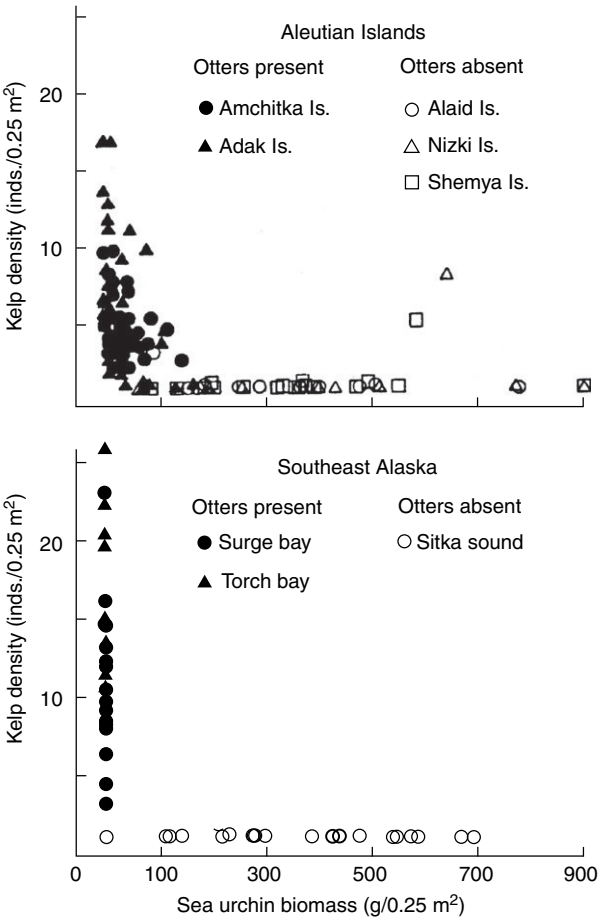


Figure 3 Density of kelp (individuals per 0.25 m^2) vs. estimated sea urchin biomass (gram per 0.25 m^2) for the Aleutian Islands and southeast Alaska. Points show averages for sites within locations. Sea urchin biomass was estimated from samples of population density, size–frequency distribution, and the functional relation between test diameter and wet mass. Reproduced from Estes JA and Duggins DO (1995) Sea otters and kelp forests in Alaska – Generality and variation in a community ecological paradigm. *Ecological Monographs* 65: 75–100, with permission from ESA.

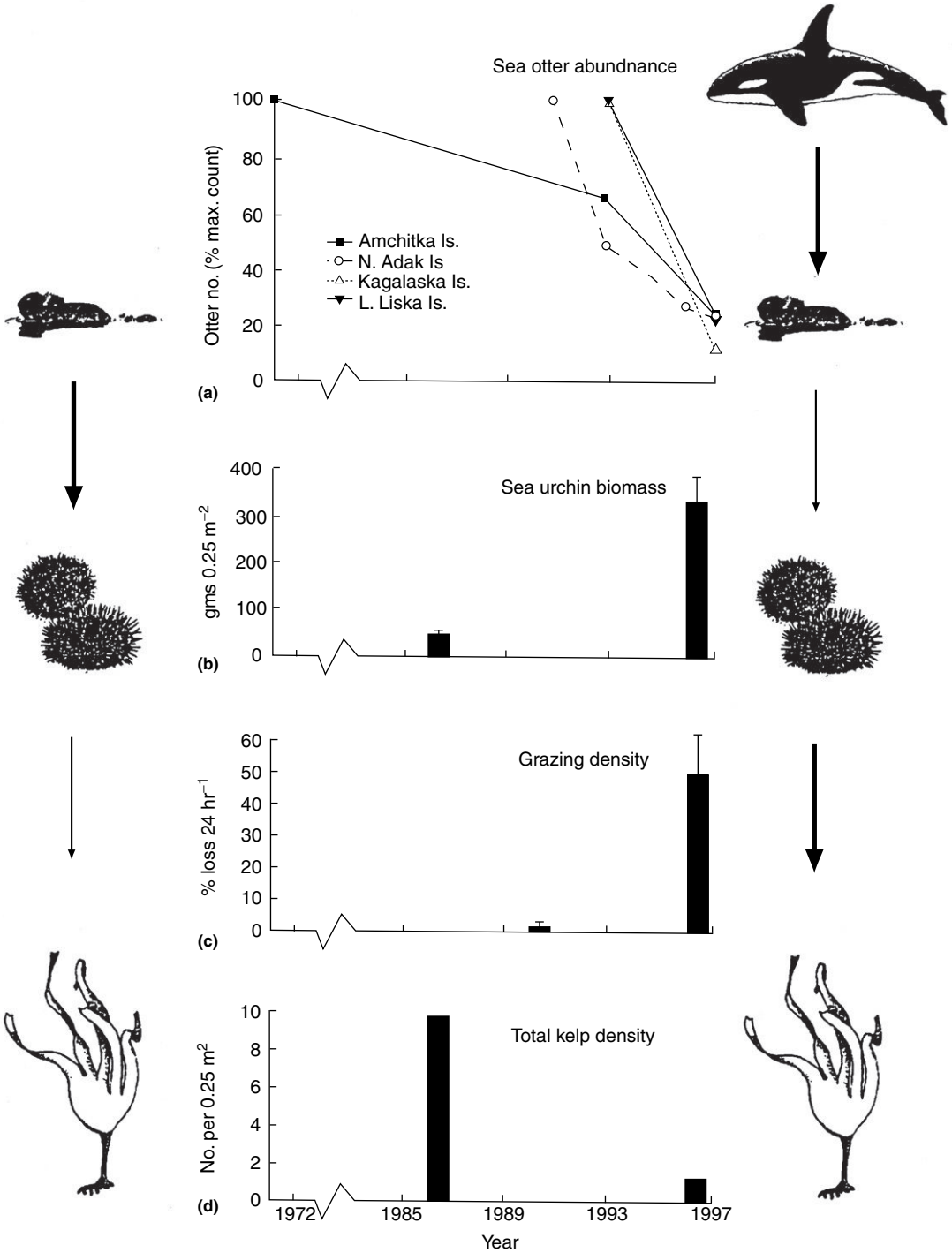


Figure 4 Changes over time in (a) sea otter abundance at several islands, (b) sea urchin abundance, (c) grazing intensity, and (d) kelp density at Adak Island in the Aleutian archipelago. Error bars are 1 SE. Diagrams to the left and right of the data panels suggest the mechanisms leading to the changes. The left diagram shows kelp forest dynamics without orcas and the right diagram shows changes induced by the addition of orcas as the top predator. Thick arrows show strong effects, and thin arrows show weak effects. Reproduced from Estes JA, Tinker MT, Williams TM, and Doak DF (1998) Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science* 282: 473–476, with permission from American Association for the Advancement of Science.

juvenile urchins presumably led to moderate rates of grazing on kelp sporelings and juveniles in the Aleutians, slowing rates of increase in kelp abundance relative to those in southeast Alaska.

A recent complication with troubling conservation implications suggests that killer whale (*Orcinus orca*) predation has begun to eliminate sea otters from their native habitats, with predictable consequences for abundances of sea urchins and

kelp (Figure 4). During the 1990s, sea otter abundance declined sharply at several Aleutian Islands. Once freed from sea otter predation, sea urchin populations increased, leading to severe declines in kelp and associated species. Estes and coworkers suggested that killer whales may have shifted their diet from large marine mammals to sea otters ultimately as a consequence of human activity. Overfishing and increased ocean temperatures (due to global climate change) are associated with declines in forage fish for sea lions and seals, and consequently these marine mammals have declined sharply in number. Historically, these large marine mammals have been the primary prey of killer whales; in their absence, some orcas have shifted their diet to otters. Because sea otters are small relative to pinnipeds, rates of predation loss have evidently been high, with substantial direct and indirect consequences for subtidal communities.

Generalization and Liberalization of the Definition: Other Types of Keystone Species

These earliest examples of keystone species were of consumers that modified the community through predation. However, some investigators suggested that species other than top predators could play keystone roles in communities. Subsequently, a variety of different types of organisms were termed "keystone" species, including herbivores, plants, pollinators, pathogens, habitat modifiers, and mutualists. The remainder of this article presents examples of such alternative keystones, summarizes a critique of these ideas, and offers proposals for further clarification of keystone and other functionally important species.

Herbivores

By consuming primary producers, keystone herbivores can have dramatic impacts on community structure. Among species identified as keystone herbivores are kangaroo rats and bison. In the Chihuahuan Desert, Brown (1998) (see summary) suggested that a guild of kangaroo rats played a keystone role. In plots where kangaroo rats were excluded, grass cover increased threefold. This was accompanied by a change in species composition; species shifted from those typical of desert shrubland to those characteristic of arid grasslands (Figure 5). The mechanisms responsible for this transition were seed predation and soil disturbance by kangaroo rats. Preferentially eating the seeds of competitively dominant grasses, kangaroo rats indirectly released subordinate plant species from competition. Furthermore, kangaroo rat burrowing favored disturbance-tolerant plant species. Hence, the presence or absence of kangaroo rats determined whether the community was desert shrubland or arid grassland.

Note that this example both extends the concept to herbivores and attributes the keystone effect to a multispecies group instead of a single species. Recently Brown (1998) suggested that the kangaroo rat effect was attributable primarily to two species, *Dipodomys merriami* and *Dipodomys spectabilis*. *D. spectabilis* went locally extinct in 1994, so if the system persists unchanged well beyond this extinction, *D. merriami* may be a keystone species. In this study, however, the original interpretation of these herbivores as a keystone "guild" may contribute

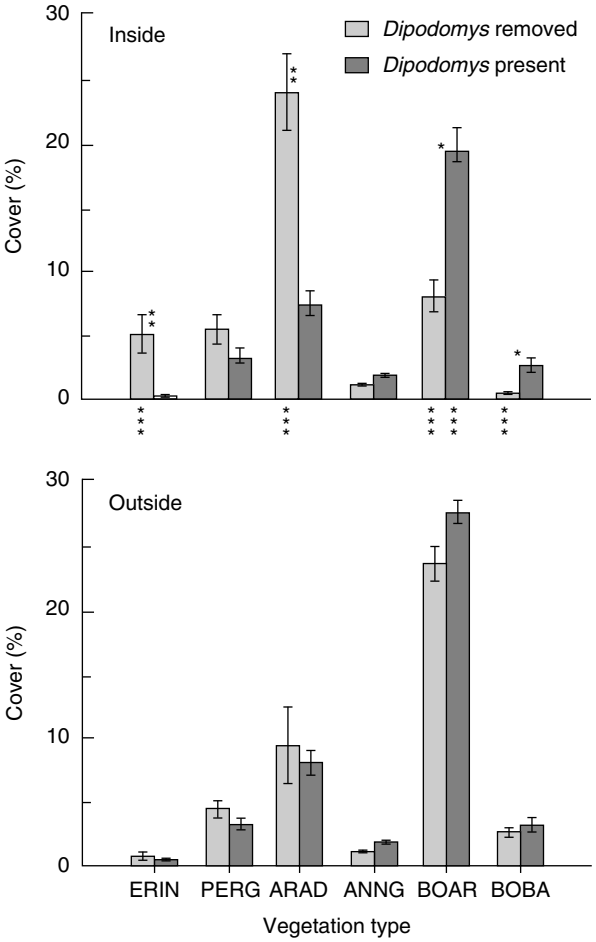


Figure 5 Densities of perennial and annual grasses in the presence and absence of kangaroo rats (*Dipodomys*). Data are mean percent cover (±1 SE) of *Eragrostris lehmanniana* (ERIN), other tall perennial grasses (PERG), *Aristida adscensionis* (ARAD), other tall annual grasses (ANNG), and the short annual grasses *Bouteloua aristidoides* (BOAR) and *Bouteloua barbata* (BOBA). Top: Data from transects inside enclosures from which kangaroo rats had been removed or allowed to remain. Asterisks above bars indicate significant differences between + kangaroo rat and − kangaroo rat enclosures. Bottom: Data from transects outside the respective enclosures. Asterisks between panels indicate significant differences between transects inside and outside plots where kangaroo rats were present or absent. Analyses employed ANOVA with plots as units of replication; *, $p<.05$; **, $p<0.01$; ***, $p<0.001$. Reproduced from Brown JH and Heske EJ (1990) Control of a desert–grassland transition by a keystone rodent guild. *Science* 250: 1705–1707, with permission from American Association for the Advancement of Science.

to confusion in terminology. This study is an excellent example of the strong effects of consumers on community structure, but as explained later, it was not designed to discern between keystone and "diffuse" (multispecies) predation.

Bison have also been identified as keystone herbivores (Knapp *et al.*, 1999). Studies at Konza Prairie have documented strong effects of bison on species composition, diversity, and several physical and chemical aspects of ecosystem function of grassland communities. These patterns occur across scales ranging from the plant level to patches to

landscapes, and are largely a consequence of bison preference for competitively dominant grasses and for persistent habitat features. As will be discussed later, Knapp and coworkers also evaluated the question of whether or not bison are truly keystone herbivores using the operational procedure recommended by Power *et al.* (1996).

Plants (Resources)

Terborgh (1986, cited in Power *et al.*, 1996) suggested that certain plants in Peruvian rain forests can provide keystone "resources." Most tree species in these tropical rain forests fruit in synchrony. However, when these species are out of season, palm nuts, fig trees, and nectar-bearing plants can provide a crucial resource to frugivores. Less than 1% of the plant biomass of tropical forests is made up of these species and yet virtually all frugivorous animals rely on them during the 3 months when other food sources are rare. Consequently, Terborgh argued that these plants maintained high animal diversity in these communities and therefore occupied a keystone role. Although Terborgh's work focused on Peru, he suggested that keystone plant resources may be widespread in tropical rain forests.

In another example, Willson and Halupka (1995) suggested that anadromous fishes swimming upstream to spawn, or their carcasses, were suggested to be keystone resources. They argued that a large number of vertebrate predators and scavengers, especially terrestrial species, rely on the energy input provided by anadromous fishes returning to their natal streams. Such energy sources can be strong interecosystem links and are suggested to have an influence on the diversity of terrestrial mammal assemblages.

The notion that such species or resources can be important to community dynamics is not disputed in this article, but for two reasons, describing these species as keystones is questioned. First, in most such cases, the evidence for keystone effects is weak and conjectural; the investigators have not demonstrated that removal of the resources would lead to wholesale community changes. Second, it is widely believed that the top-down effects (such as keystone predation, keystone herbivory) are qualitatively distinct from bottom-up effects (e.g., primary production and food supply). Therefore, the authors suggest that the examples in which plant-level processes are thought to determine consumer effects, and which thus suggest that the plant is the keystone species, serve to confuse rather than clarify the concept by combining under a single idea a heterogeneous collection of ecological processes. Although the evidence in favor of the role of bottom-up effects as important determinants of community dynamics is strong, these effects commonly lead to variation in top-down effects, often with single species (i.e., keystones) having large effects that are disproportionate to their abundance. It is therefore argued here that retaining the original distinction of keystone species as consumers is simpler and clearer, and hence better serves both understanding and insight into community dynamics and policy decisions regarding ecosystem management and conservation. Studying the role of bottom-up processes as determinants of community structure is an exceedingly important area of research, but the authors believe that the progress both in this area and in the role of keystone species will be enhanced if they are treated as distinct but dynamically linked concepts.

Mutualists (Pollinators, Seed Dispersers)

Flying foxes (*Pteropus* spp.) have been proposed to be keystone "mutualists" in tropical rain forests. Flying foxes are large tropical bats that are important pollinators and seed dispersers in Old World forests and on tropical islands. Especially on isolated islands where other vertebrate pollinators are scarce, these bats can be responsible for dispersing 80–100% of the seeds. Cox and coworkers (cited in Power *et al.*, 1996) argued that, through these pollination and dispersal "services," flying foxes may be responsible for maintaining high plant diversity in the forest communities in which they occur. Comparisons of the proportion of fruits and seeds dispersed by flying foxes on Guam, where flying foxes have been driven nearly to extinction, and on Samoa, with abundant flying foxes, were consistent with this idea (see Rainey *et al.*, 1995, cited in Power *et al.*, 1996). Fruits dispersed by flying foxes on Western Samoa were between 0% and 100%, whereas 0–1% of seeds were dispersed by flying foxes on Guam. Such data are only suggestive; however, further study is needed to demonstrate that the forest community structure reflects these differences in fruit and seed dispersal, and that flying foxes play a keystone role in determining community structure.

Habitat Modifiers

Organisms that influence the availability of resources for other species by modifying the physical environment have also been considered as keystone species. The term "ecosystem engineers" has been coined to identify such organisms (see Lawton and Jones, 1995, cited in Power *et al.*, 1996). Beavers, red-naped sapsuckers, kangaroo rats, and prairie dogs are all examples of species that, through their nonforaging or nontrophic activities, modify and/or create habitat for other species. By altering the hydrology of rivers, beaver dams can have profound effects on sediment retention, nutrient cycling, and the condition of the riparian zone (Naiman *et al.*, 1986). Sapsuckers create habitat for two species of swallows by drilling holes in aspens. In addition, their feeding holes in willows create sap flows, providing a resource for several species of birds, mammals, and insects (see Daily *et al.*, 1993, cited in Power *et al.*, 1996). The activities of kangaroo rats and prairie dogs cause soil disturbance, which affects community structure. These habitat modifiers were all thought to have impacts disproportionate to their abundance and were thus dubbed keystone species.

Critique and Reevaluation

Paine's original definition of keystone species referred to top predators that greatly modified the species composition and physical appearance of the ecosystem by preferentially feeding on the dominant competitors for space. Although not quantitative, the term was clear in concept and interpretation. As documented in the section Habitat Modifiers, however, the term was subsequently applied to a variety of organisms other than top predators and to actions that affected communities in ways other than feeding on competitive dominants. Application to groups of species (e.g., keystone guilds) rather than to a single species blurred the concept even further. These definitional liberalizations, combined with a lack of rigor in determining if candidate species meet the necessary criteria to be

termed "keystone," eventually undermined the usefulness of a potentially powerful concept (Mills *et al.*, 1993). Here, these issues are evaluated and a scheme of classification and definition (slightly modified from that offered by Power *et al.*, 1996) is suggested, which further clarifies and advances the usefulness and application of the concept of keystone species.

Reevaluating the Keystone Species Concept

The seeming lack of a working definition of keystone species prompted Mills *et al.* (1993) to argue that careless usage had sufficiently degraded the value of the concept to justify its removal from usage. They further argued that because the term was poorly defined it was functionally useless and conservation strategies therefore should not be based on protecting keystone species. They advocated focusing on interaction strength rather than on a species' keystone or nonkeystone status as a more useful management strategy.

In response to the criticism by Mills *et al.* (1993), a group of ecologists with expertise in the study of keystone species and strongly interacting species met to evaluate the keystone species concept. The publication that resulted from this meeting was a signal achievement, and much of the present article is patterned after this synthesis (Power *et al.*, 1996). The group agreed with Mills *et al.* (1993) that through misapplication and questionable redefinition, ecologists and conservation biologists had obscured the meaning of the term keystone species. Rather than abandoning the concept, however, the group proposed clarification and adherence to a set of more sharply defined concepts for community dynamics. They defined a keystone species as "one whose impact on its community or ecosystem is large, and disproportionately large relative to its abundance" (Power *et al.*, 1996). This definition retained the essence of Paine's original usage, but expanded it more broadly to include species other than predators.

Power *et al.* (1996) stressed the importance of having a rigorous, quantitative method of assessing the community- or ecosystem-level effects of a species when determining if it is a keystone. To assess a species impact on a community, they proposed a community importance (CI) index, which in practical terms is quantified experimentally as

$$\mathrm{CI}_i = [(t_\mathrm{N} - t_\mathrm{D})/t_\mathrm{N}](1/p_i)$$

where p_i is the proportional abundance of the species i before it was deleted, t_N (for "normal") is a quantitative measure of a community or ecosystem trait under usual conditions (e.g., productivity, nutrient cycling, species richness, and relative abundances of species), and t_D (for "deleted") is the trait in the absence of species i. Keystone species are those whose CI is large relative to that of other species (**Figure 6**). On the basis of these considerations, Power *et al.* suggested that if the total impact of a species is plotted against its proportional abundance, keystone species would cluster toward the upper left of the graph in **Figure 7**. Power *et al.* (1996) defined another group of species, "dominants," as those that had large total impacts but in proportion to their abundance – these would lie to the upper right of the abundance–impact diagram.

Importantly, keystone species are not necessarily simply "strong" interactors, at least as the terms are defined here.

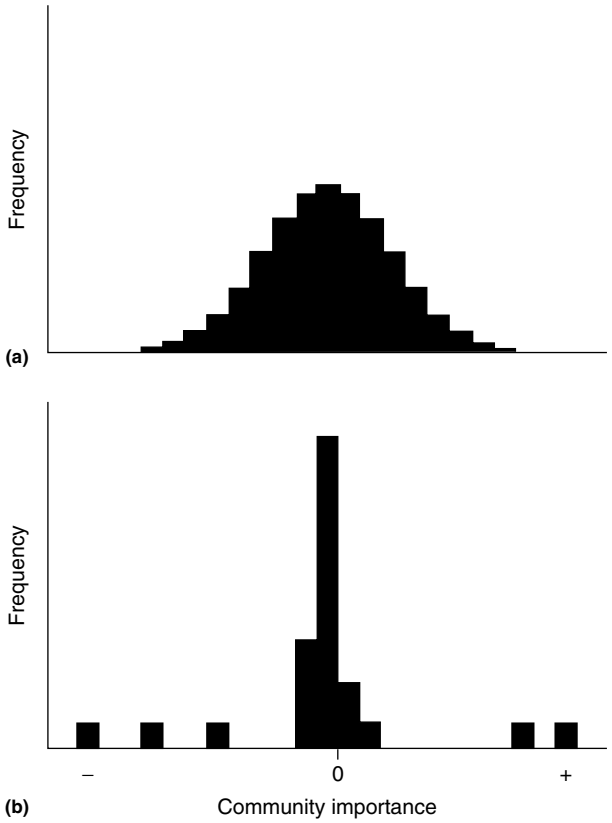


Figure 6 Possible frequency distributions of community importance (CI) values for all species in a community. Positive values occur when a community trait decreases after a species is deleted. For instance, in the absence of a mutualist, the target dominant species would also decrease. Negative values occur when a community trait increases after removal of a species. For instance, in the absence of a consumer, the target dominant species would increase. Community importance values may be normally distributed around zero (a), indicating that most species have small effects and keystones are rare. In some communities (b), CI values may have several modes, with keystone species indicated by values far from zero. Reproduced from Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *Bioscience* 46: 609–620.

Keystone species not only have disproportionately large effects, but also have a community or ecosystem-wide impact through direct and indirect effects cascading through the system (Power *et al.*, 1996; see **Figure 1**). Strong interactors are species that have a large impact on the species with which they interact, but this large impact could affect only a single interacting species. Only when strong interactors have multispecies effects that alter the structure of the community, or the functioning of the ecosystem, are they also keystone species.

Related Concepts

The idea that not all species have equal significance in community dynamics is an old one, and many efforts have been made to assign names or terms to distinguish species or groups of species with important roles. Elton (1927) advanced the concept of "key-industry species" as single species of animals supporting a large number of consumers (e.g., copepods,

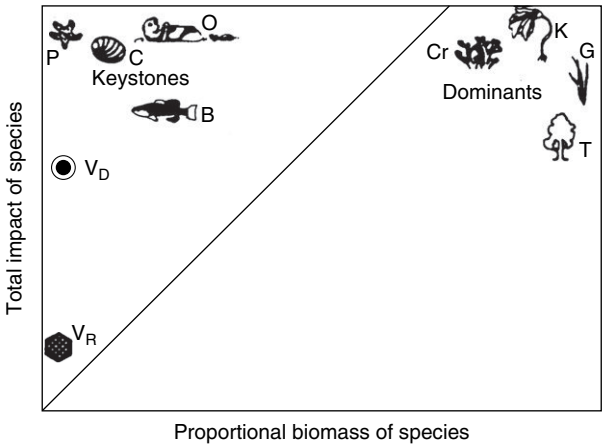


Figure 7 Total impact of a species (absolute value of community impact × proportional abundance of a species) vs. its proportional abundance, p_i. Species whose total impact is proportional to their abundance would fall along the diagonal line $X = Y$. Keystone species have effects that greatly exceed their proportional abundance, both on a per capita (or per biomass) basis and a per population basis, and would cluster toward the upper left region of the diagram. To illustrate, some organisms (e.g., rhinovirus, V_R, which causes colds in wildlife) may have effects that are greater than expected from their biomass, but because the impact on the community is relatively minor, they are not keystone species. Others (e.g., distemper virus, V_D, a killer of lions or wild dogs) may have collective effects on the community that are disproportionately large, and would be keystone species. Examples of keystone species are *Pisaster* (P), sea otters (O), predatory whelks (*Concholepas*, C), and freshwater bass (B). Dominants are high proportional biomass species whose large effects are not disproportional to their biomass, such as trees (T), giant kelp (K), prairie grasses (G), and reef-building corals (Cr). Positions of each species or group represent educated guesses. Reproduced from Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *Bioscience* 46: 609–620.

herring, anchovies, and sea pens). "Foundation species" were defined as the "group of critical species which define much of the structure of a community" (Dayton, 1972). Related ideas include "interaction webs" and "functional webs," which are defined as the subset of strongly interacting species that regulates community structure (Menge and Sutherland, 1987). A similar idea, the "extended keystone hypothesis," states that ecosystems are controlled by a small set of key plants, animals, and abiotic processes. Note, however, that the latter phrase adds environmental stresses ("keystone processes") such as fire, wave-induced damage, substratum movement, and other kinds of disturbances to the biological forces stressed by the other concepts. A contrasting view that biotic and abiotic forces are qualitatively distinct suggests the term "critical processes" to describe abiotic effects that can structure communities.

Recommended Terminology

Although the various generalizations of the keystone concept and the lumping of several related concepts into a broad keystone species definition were well intentioned, this study agrees with Mills *et al.* (1993) and Power *et al.* (1996) that such usage sharply reduces the utility of the keystone species concept. On the basis of current usage and the documented need and desire for clarity in standardized terminology, the adoption of the following terminology is proposed. Species in communities can be defined as:

- *Keystone species*: Consumers having a disproportionately large effect on communities and ecosystems. By this definition, keystone species can include predators, parasites, pathogens, herbivores, pollinators, and mutualists of higher trophic status, but not plants, sessile animals, or "resources" (e.g., salmon carcasses, salt licks, and deep pools). To date, no convincing examples of communities with more than a single keystone species are known. Therefore, the suggested hypothesis is that most communities will have at most a single keystone species. At the ecosystem level, which might include several distinct communities, there may be several keystone species. Evaluation of these predictions awaits future research.
- *Strong interactors (critical species)*: Species having a large effect on the species (one or more) with which they interact. Communities and ecosystems may have many strong interactors, and such species may occur at all trophic levels. Strong interactors would lie toward the upper side of the abundance–impact diagram (**Figure 7**), and therefore include both keystone species and dominants.
- *Weak interactors*: Species having little effect on other species, at least under average conditions. Under some circumstances, weak interactors may occupy important roles in ecological communities as a result of changes that lead to temporary increases in abundance, size, or biomass (Berlow, 1999). Weakly interacting species would all lie toward the lower portion of **Figure 7**.
- *Dominant species*: As noted above, dominant species are those strongly interacting species that owe their influence to their high abundance (**Figure 7**). Such organisms are the species that comprise a large proportion of the biomass in a community, and thus are the dominant components of community structure. Trees in forests, mussels in rocky intertidal habitats, grasses in grasslands, and kelps or corals in nearshore subtidal habitats are all dominants.
- *Key-industry species*: As defined earlier, key-industry species are prey that support a large group of consumers. Following Elton's (1927) usage, key-industry species are therefore animals of intermediate trophic status.

Groups of species in communities or ecosystems can include:

- *Interaction webs*: Interaction webs (= functional webs) are the subset of species that through their interactions and responses to abiotic factors make up the dynamic core of food webs or communities. Interaction webs include keystone species, dominants, and other strong interactors.

Identification of Keystone Species

Lack of experimentation and other rigorous approaches to identify keystones continues to be a pervasive problem. In many cases, species are named keystones based on superficial evidence such as natural history observations. A lack of rigor in identification could result in mislabeling a species as a

keystone and inferring that it is the primary determinant of the structure of its community when in fact it is not. As noted by Mills *et al.* (1993), this can yield a serious loss of credibility with respect to policy and management decisions.

Despite limitations of spatial and temporal scales and other shortcomings, experimentation supplemented by comparison remains the most powerful tool available for revealing the dynamics of communities and ecosystems (Paine, 1994). In many cases, however, practical, legal, and ethical concerns preclude manipulations of suspected keystones (e.g., sea otters, killer whales, or prairie dogs). It is clear, therefore, that the identification of keystone species must necessarily use a variety of approaches. In addition to the comparative-experimental approach, alternative approaches include path analysis, sensitivity analysis, the study of natural or accidental invasions, the study of the consequences of overexploitation, and exhaustive and detailed comparison coupled with natural history and ideally small-scale ancillary experimentation (Power *et al.*, 1996). Inferences based solely on descriptive natural history knowledge (e.g., diet composition and frequencies, behavior, and abundance) are likely to be misleading. For example, under natural conditions a keystone predator may rarely consume the competitive dominant in a system (because it has sharply reduced the prey's availability), and thus be overlooked as a possible regulator of the dominant or the community. In the sea star system studied by Paine (1966), for example, barnacles, not mussels, were the most frequently consumed prey. Thus, although natural history knowledge is fundamental to the understanding of the dynamics of any ecological system, much additional evidence is necessary before the ecological role of a species can be determined.

Experimental Approaches

Keystone species are often identified by removal or exclusion experiments; that is, the presumed keystone is deleted from a portion of the habitat, and the effect of the removal on the community is compared to a separate, control portion of the habitat. One problem with this approach is that if a suspected keystone species is removed but there is no detectable response by the community, it is not possible to conclude that the community lacks a keystone species. In such a case, three alternative interpretations are possible. (1) Predation is weak overall, such that no predator deletion will produce an effect. (2) Predation is strong and there is a keystone species in the community, but it was not the species removed. (3) Predation is strong but diffuse, and multiple predators would need to be removed to produce an effect.

To tease apart these alternatives, an appropriate experimental protocol in testing for keystone predation (Menge *et al.*, 1994; Navarrete and Menge, 1996) should include:

Treatment	*Explanation*	*Tests*
1. + Predators	Intact community, all predators present	Control or reference ("natural" community state)
2. − Predators	Quantifies total predation effect, all predators removed or excluded	Strong vs. weak predation
3. − Single predator	Quantifies single species effects, deletion of each predator species singly, leaving the others present	Keystones vs. diffuse predation

For example, if consumer species are removed both collectively and singly, but predation is weak, the total effects of predation and the effects of each species with respect to their impact on community structure will be small (**Figure 8(a)**). With strong predation regimes, keystone predation would be indicated if total removals and removal of just one of the predator species were similarly large (**Figure 8(d)**). Finally, diffuse predation would be indicated if removal of each predator species singly led to significant but substantially smaller effects than removal of all predator species (**Figure 8(g)**).

This protocol also has important implications for understanding how communities or ecosystems might respond to loss of trophically high species (Allison *et al.*, 1996). In a system where predation is weak, of course, loss of species should have little effect (**Figure 8(a–c)**). With strong diffuse predation, where each consumer has an impact on community structure, species loss may lead to little change in the system as a result of compensation by the remaining consumers, at least until most consumers have been lost (**Figures 8** and **9**). In both cases, the community response to species deletion is relatively predictable (Allison *et al.*, 1996). With strong keystone predation, however, the consequence of species loss is relatively uncertain (**Figures 8** and **9**). By definition, compensation by the weakly interacting predators, while possible, is not likely to fully account for the loss of the keystone species, and if the identity of this species is not known *a priori*, the system response is highly uncertain (see **Figure 9**).

This protocol has rarely been used, but two examples illustrate its efficacy. Menge *et al.* (1986) evaluated the separate and combined impacts of most combinations of four groups of consumers in rocky intertidal communities in Panama. The extremely high diversity of this tropical community imposed an immediate compromise in their design: Simultaneous single-species removals were essentially impossible because >40 consumer species were relatively abundant. The compromise was to remove major consumer groups (i.e., omnivorous crabs, omnivorous fishes, predatory whelks, and grazing molluscs), each consisting of several common species. Their partial factorial design (difficulties in separating crab and fish effects meant only 12 of a possible 16 treatments could be conducted) included treatments assessing the effects of removing consumers in single groups and in total. After 3 years, none of the single-group removals produced an effect similar in magnitude to the exclusion of all the consumers (**Figure 10**). Removal of two and three groups produced changes in prey that were intermediate between single-group and total exclusion treatments. Thus, total predation pressure in this system was high but the effect of individual groups was small. The lack of prey responses from single-group removals compared to the increasingly strong responses from removals of two, three, and all four groups suggests three things: That, up to a point, groups can compensate for reductions in other groups; that there was no keystone species in this system; and that predation was diffuse.

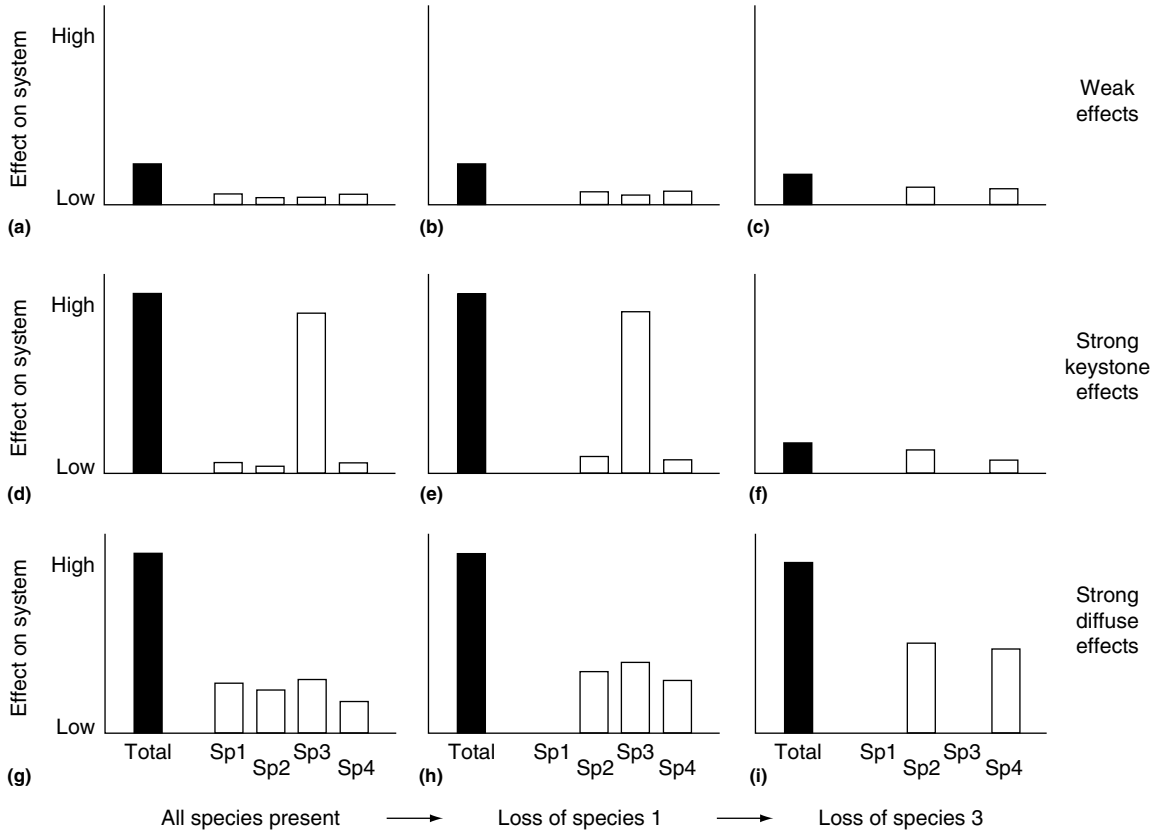


Figure 8 Total effects (solid bars) of a consumer group and effects of single species (open bars) under three regimes of predation effect, with possible responses to sequential loss of single species. Panels (a), (d), and (g) represent the intact assemblage; (b), (e), and (h) represent the changes in the absence of species 1; and (c), (f), and (i) represent the changes if both species 1 and 3 are removed or lost. Panels (a), (b), and (c) show change with weak predation; (d), (e), and (f) show changes with strong keystone predation; and (g), (h), and (i) show changes with strong diffuse predation. In the absence of the keystone species, the remaining species have a small impact on the community (f). In contrast, with diffuse predation, the remaining species compensate strongly for the loss of a consumer so that predation remains strong. Reproduced from Allison GW, Menge BA, Lubchenco J, and Navarrete SA (1996) Predictability and uncertainty in community regulation: consequences of reduced consumer diversity in coastal rocky ecosystems. In: Mooney HA, Cushman JH, Medina E, Sala O, and Schulze E-D (eds.) *Functional Roles of Biodiversity: A Global Perspective*, pp. 371–392. Chichester, UK: John Wiley & Sons.

In another study, Navarrete and Menge (1996) tested the relative intensities of predation on the small mussel *Mytilus trossulus* by the original keystone species, the sea star *P. ochraceus*, and coexisting predatory whelks, *Nucella emarginata* and *N. canaliculata*, on the Oregon coast. In a fully factorial design they quantified rates of predation on mussels in all combinations of presence or absence of each. Their results indicated that, as predicted, *Pisaster* had the strongest effect, by far, on mussel survival (**Figure 11**). Although whelk effects were almost undetectable in the presence of sea stars, they were relatively strong in the absence of the sea stars. Thus, *Pisaster* was the keystone predator in these experiments. Whelks were weak interactors in the presence of sea stars, but, via compensatory increases in density, had moderately strong effects on mussel survival when sea stars were absent.

Nonexperimental Approaches

To be convincing, comparative evidence needs to be extensive in space and time and to incorporate a wide range of approaches. One of the most convincing examples of the determination of keystone species on the basis of description and comparison is the sea otter example described earlier. Another convincing example is that of beavers as keystone species (Naiman *et al.*, 1986). By building dams across streams and through feeding activities, beavers can have profound effects on the structure and dynamics of aquatic ecosystems. Beavers alter hydrology, channel geomorphology, productivity, and biogeochemical cycling at a magnitude far exceeding their proportional biomass in these systems.

Bison are another good example of keystone species designation based on largely nonexperimental evidence (Knapp *et al.*, 1999). Extensive and detailed studies at the levels of individual leaves, individual plants, plant populations, and landscapes showed that selective grazing at the level of species and patch greatly altered patterns of community structure. Bison preferred grasses in burned areas and revisited these sites repeatedly establishing a distinct high-diversity mosaic pattern of vegetation. Their grazing on grasses and avoidance of forbs resulted in high levels of species diversity, high spatial heterogeneity, and high nitrogen availability in tallgrass prairie. Moreover, bison grazing interacted with fire, another

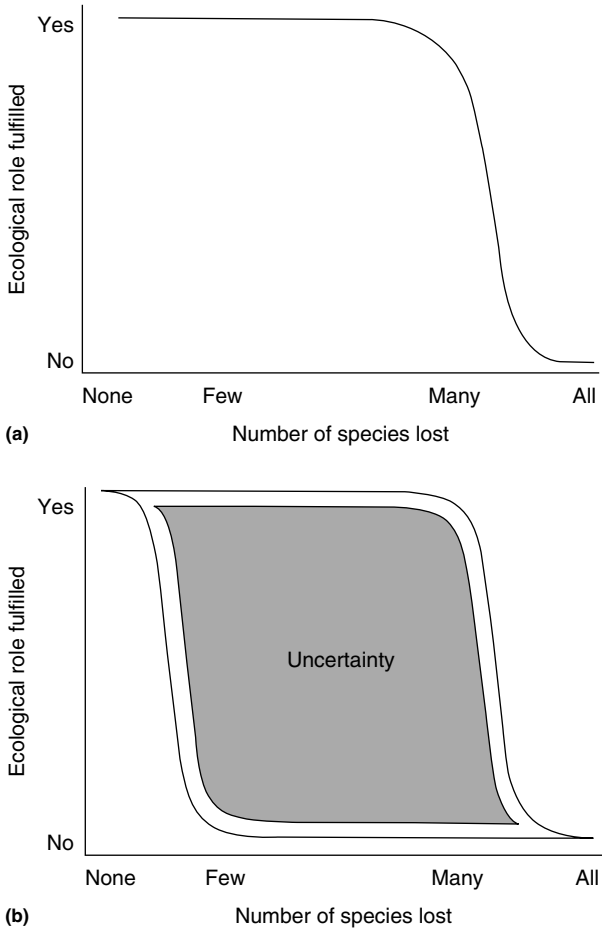


Figure 9 Relationship between the ability of a group of consumers to maintain a system at some persistent state ("ecological role fulfilled") and the regime of predation. (a) Diffuse predation: In moving from left to right on the abscissa, the remaining species compensate for the loss of other species, maintaining similar structural patterns in the community or ecosystem until most are lost. (b) Keystone predation: As species are lost, dramatic changes in structure can occur with each removal, but unless the identity of the keystone species is known, prediction of when such changes occur is uncertain. Reproduced from Allison GW, Menge BA, Lubchenco J, and Navarrete SA (1996) Predictability and uncertainty in community regulation: consequences of reduced consumer diversity in coastal rocky ecosystems. In: Mooney HA, Cushman JH, Medina E, Sala O, and Schulze E-D (eds.) *Functional Roles of Biodiversity: A Global Perspective*, pp. 371–392. Chichester, UK: John Wiley & Sons.

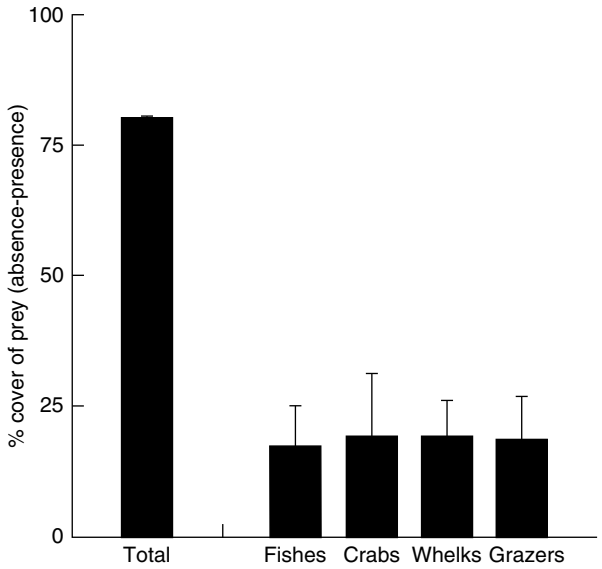


Figure 10 Effect of consumers, collectively and by each of four single groups (fishes, crabs, whelks, and molluscan grazers), on total prey abundance in the rocky intertidal region at Taboguilla Island, Panama. Prey included sessile invertebrates (mostly barnacles and bivalves), colonial invertebrates (hydrozoans), and algae. The estimate of the effect of consumers was based on the percent cover of prey in the absence and presence of all consumers (total) and each group separately. Single-group effects were estimated by determining the difference between treatments that differed in whether or not the particular group was included in the treatment. For example, fish effects were estimated by treatments $+F+C+W+G$ vs. $-F+C+W+G$, $+F+C-W+G$ vs. $-F+C-W+G$, etc. Error bars are 1 SE. Reproduced from Menge BA, Lubchenco J, Ashkenas LR, and Ramsey F (1986) Experimental separation of effects of consumers on sessile prey in the low zone of a rocky shore in the Bay of Panama – direct and indirect consequences of food web complexity. *Journal of Experimental Marine Biology and Ecology* 100: 225–269.

natural element of tallgrass prairies, in ways that enhanced the patchiness, diversity, and nitrogen availability, and generated intermediate levels of net primary productivity. As mentioned earlier, estimates of the CI of bison indicate that these large grazers clearly have a large effect, one disproportional to their biomass, on plant community structure and ecosystem functioning.

Space does not permit further detailed citation, but other convincing examples of keystone species are available from marine, freshwater, and terrestrial habitats, including those cited in Table 1 in Power *et al.* (1996). Many more examples than these are available in the literature that are based on limited evidence and thus remain conjectural. It seems clear, however, that the phenomenon of keystone species is real and widespread, and that a variety of approaches are available to document their existence and importance. What remains unclear from the foregoing is (1) whether or not there are keystone traits that allow *a priori* identification of keystone species, and (2) whether or not there are predictable conditions under which keystone species will evolve and persist. These issues are considered next.

Identification of Keystone Species *a priori*?

Field experimentation is not always feasible. Moreover, even where and when feasible, such studies are often time-consuming and expensive. It seems obvious that it is impossible to study each and every system on the Earth to determine the nature of interactions among species and how these influence community dynamics. For these reasons, an ultimate goal in ecology is to gain the ability to predict the roles of species and the processes underlying the dynamics of communities and ecosystems. Ideally, predictions would be based on observation and measurement of organismal,

population, and community traits and patterns. Evaluation of the reliability of such predictive traits would depend on cycles of study that identified important traits in particular systems and tested their predictive capacity in novel systems.

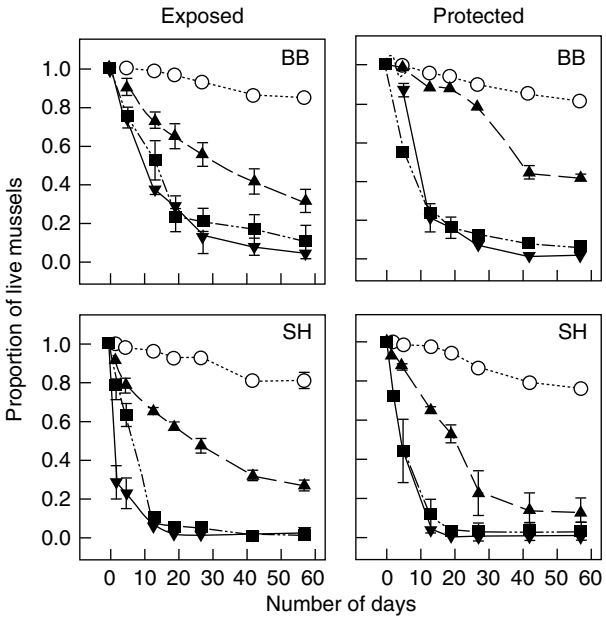


Figure 11 Results of a mussel transplant experiment. Data are the mean proportions of mussels (±1 SE) surviving in each of four treatments at wave-exposed and wave-protected sites at two study sites, Boiler Bay (BB) and Strawberry Hill (SH). Treatments were all combinations of presence and absence of whelks (*Nucella* spp.) and sea stars (*Pisaster*). ○, − *Pisaster* − *Nucella*; ▲, − *Pisaster* + *Nucella*; ■, + *Pisaster* − *Nucella*; ▼, + *Pisaster* + *Nucella*. The experiment was begun 3 July 1993 and completed 30 August 1993. Reproduced from Navarrete SA and Menge BA (1996) Keystone predation and interaction strength: Interactive effects of predators on their main prey. *Ecological Monographs* 66: 409–429.

Traits originally suggested to characterize keystone species were differential impacts on prey species, high consumption rates relative to prey production, and a focus of the keystone's impact on dominant competitors in the system. A survey of well-studied keystone predator examples in marine and freshwater habitats, however, suggested that these, and several other postulated "keystone-identifying" traits, were not predictably associated with keystone species (Menge *et al.*, 1994, see Table 1). In all, 17 studies of strong predation were considered. Eleven were dominated by keystone predation and six were dominated by diffuse predation. As indicated in **Table 1**, differential predation and dominance of the community by a single prey species in the absence of the predator were characteristic of both keystone and diffuse predation systems, and thus do not serve as a distinguishing trait of keystone predation. Although data were limited, prey production rate was the only trait that consistently distinguished keystone (high prey production) and diffuse (low prey production) predation systems. No other potential trait was consistently associated with either type of system. Clearly further research is needed, but based on present knowledge, *a priori* identification of keystone species using traits alone does not seem possible.

Context Dependency

The role of a species as keystone can be context dependent (Menge *et al.*, 1994; Power *et al.*, 1996); that is, a species that serves a keystone role under some set of biotic and/or abiotic conditions may not serve under other conditions. For example, at exposed headlands along the Oregon coast, the sea star *Pisaster ochraceus* was found to occupy a role similar to that documented for this species in comparable habitats on the Washington coast (Menge *et al.*, 1994; Paine, 1966). On more wave-sheltered rocky shores, however, *Pisaster*'s role was weak to absent, despite the presence of populations of this sea star

Table 1 Summary of survey of potential keystone species traits in consumer-dominated aquatic communities

Trait	*Number of interaction webs*					
	Keystone predation			*Diffuse predation*		
	Yes	*No*	*?*	*Yes*	*No*	*?*
1. Differential predation	7	0	4	3	0	3
2. Predation focused on competitive dominant	9	0	2	4	0	2
3. High predator/prey size ratio	10	1	0	6	4	0
4. High functional/numerical response	8	2	1	6	4	0
5. Indeterminate growth of predator(s)	8	3	0	6	4	0
6. High predator mobility relative to prey	10	1	0	6	4	0
7. Spatial refuge for prey	6	4	1	3	2	1
8. Size refuge for prey	1	8	2	2	3	1
9. Space is limiting resource for prey	7	4	0	5	1	0
10. Low community diversity	3	5	3	2	2	2
11. Differential rates of prey recovery	6	1	4	4	0	2
12. High rate of prey production	5	0	6	0	2	4

Question mark means the conclusion is uncertain or unknown. For traits 3–6 the numbers under diffuse predation total >6 because several of the systems had several strongly interacting predators, each of which had a yes or no or ? conclusion.

Source: Reproduced from Table 10 in Menge BA, Berlow EL, Blanchette CA, Navarrete SA, and Yamada SB (1994) The keystone species concept – Variation in interaction strength in a rocky intertidal habitat. *Ecological Monographs* 64: 249–286.

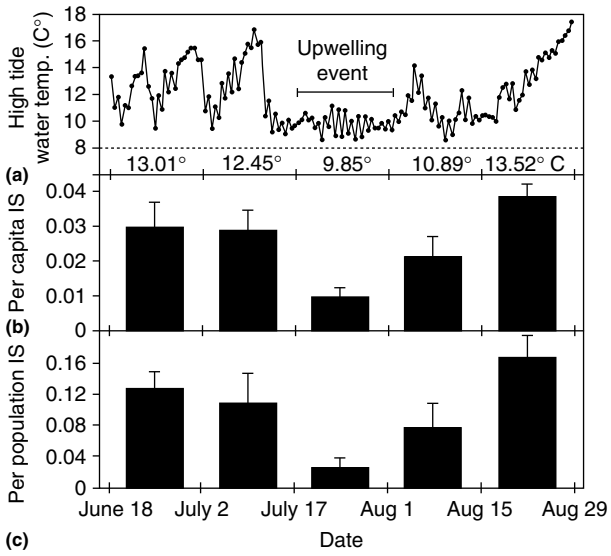


Figure 12 Seawater temperature and sea star–mussel interaction strength (IS) during consecutive 14-day periods (abscissa). Data are means (+1 SE) of values recorded at three sites. (a) High-tide water temperatures (the mean from 2 h before to 2 h after each high tide). Temperatures above the abscissa are the overall means for the 27 high tides of each period. (b) Mean per capita IS of sea stars on transplanted mussels during each 14-day period. (c) Mean per population IS of sea stars during each period. Both per capita and per population IS were lower during upwelling than during the other four periods. Reproduced from Sanford E (1999) Regulation of keystone predation by small changes in ocean temperature. *Science* 283: 2095–2097, with permission from AAAS.

at all locations studied (Menge *et al.*, 1994). The basis for this change was that environmental conditions changed with diminished wave impact. At one sheltered site, prey production rates (a combination of recruitment rates and prey growth) were so low that the sea star population consisted of a few scattered individuals. This and other studies suggest that sea star abundance is highest where food concentrations are high. At the other sheltered site, periodic and unpredictable burial by sand, not sea star predation, eliminated the competitively dominant mussels (*M. californianus*) from the lower shore (Menge *et al.*, 1994). Keystone predation in this system, then, depended on the spatial and environmental context in which the community occurred.

Keystone effects may also be altered by temporal environmental changes. Studies on the Oregon coast showed that *Pisaster* foraging activity is strongly suppressed during periods of cold water temperatures such as those occurring during periodic upwelling (Sanford, 1999; **Figure 12**). Laboratory and field studies in northern California demonstrated that sea star activity and behavior were sensitive to aerial temperature exposure as well, with chronic exposure to warm air temperatures at low tide leading to reduced feeding and growth rates and acute exposure causing increased feeding rates but no effect on growth rates (Pincebourde *et al.*, 2008). Further, the sea star also had the ability to respond to exposure to stress by taking a greater volume of water into its body, evidently to provide a buffer against the possibility of subsequent exposure to hot, desiccating conditions (Pincebourde *et al.*, 2009). Although further study is needed to determine the broader significance of this result, sea and air temperature variation in time or space could also modify the community role of this species. In particular, alteration of upwelling patterns by global climate change could lead to community- and ecosystem-level modifications or shifts in community pattern over a broad geographic region. Thus, investigations conducted over a wide range of environmental conditions indicated that whether or not a species can fill a keystone role is dependent on physical and biological conditions (i.e., context).

In a second example of context-dependent keystone predation, juvenile steelhead were demonstrated to control summer food web structure in California rivers in years experiencing winter flooding. Steelhead were the top predator in a four-level food chain, and when present they controlled invertebrate predators, freeing herbivorous insects from consumer control and allowing them to control the abundance of algae (Power, 1990; cited in Power *et al.*, 1996). When steelhead were excluded, invertebrate predators increased and controlled herbivorous insects, resulting in a high biomass of algae. Winter floods swept these communities away, leading to annual cycles of community redevelopment, with shifts in algal biomass reflecting the development of successive trophic levels through the summer into autumn.

In years without winter flooding, predation-vulnerable herbivorous midge larvae (*Pseudochironomis*) were replaced by predator-invulnerable herbivores (e.g., aquatic moth larvae *Petrophila*, caddisflies *Tinodes* and *Glossosoma*) (Power, 1992, 1995; cited in Power *et al.*, 1996). The invulnerable herbivores sharply suppressed algal abundance and, due to their effective defenses against predators, effectively shortened the food chain from four to two levels. Thus, fishes (juvenile steelhead) were keystone predators in the context of years with winter flooding and were not in the context of years without winter flooding.

A variety of similar examples are available from all major habitats (see Table 2 in Power *et al.*, 1996), suggesting that context dependency is common and perhaps near-universal in ecological communities. Such variation may be a useful tool in investigating the conditions under which keystone species will evolve and persist. Understanding the conditions under which a species plays a keystone role is critical if ecologists are to build a predictive framework to evaluate the effects of global changes in biodiversity.

Interaction Strength

As noted earlier, keystone species are distinguished by both the strength of their interactions with other species and the large indirect consequences of these effects that occur through the structure of the food web. Since determining whether or not a species is a keystone depends on the rigorous application of methods that can quantify these effects, it is important to consider the concept of interaction strength. The patterning and dynamics of species interaction strengths (SIS) and the effects on ecosystem structure and dynamics has been a hot topic for both empirical and theoretical ecologists. Indeed, the variability and context dependency of the strength of species interactions is a great challenge for predictions of

ecosystem dynamics in response to perturbations (Agrawal *et al.*, 2007).

There are many measures of SIS. Both Berlow *et al.* (2004) and Wootton and Emmerson (2005) give comprehensive overviews of these various measures, how they relate to one another, and what their respective strengths and weaknesses are. The various kinds of SIS measures can be categorized into two main groups that measure effects at two different "levels" of a food web (Berlow *et al.*, 2004). The first are measures of SIS that refer to the property of an individual link. Theoretical models, such as the Lotka–Volterra models of interspecific competition, include a coefficient, α_{ij}, which "measures how strong the interactions are" of one species j on another species i (Macarthur, 1972). Maximum consumption rates and more complicated predator functional responses to prey density in predator–prey models are additional examples of interaction strength as a property of an individual link (Berlow *et al.*, 2004). These link-level measures of SIS are independent of the network context that the species are embedded in.

Macarthur (1972) also suggested a second meaning of interaction strength: Interactions can be considered strong if their "removal would produce a dramatic effect." In this case, interaction strength refers to a system-level response of one species on the dynamics of another or on the functioning of the whole ecosystem. In the latter case, by our earlier definition, the strong interactor is a keystone species. Metrics that fall into this category of SIS include the number of secondary extinctions, or changes in one species' abundance resulting from the removal or perturbation of another (Berlow *et al.*, 2004). Most empirical perturbation experiments fall into this category, since species manipulations impact the total direct and indirect effects of one species on another. Paine (1992), for example, quantified interaction strengths of herbivores (sea urchins, limpets, and chitons) grazing on kelp sporelings in field enclosure/exclosure experiments. Specific impacts of individual herbivore species were determined by comparing the densities of kelp sporelings that occurred in the complete absence of all herbivores (a measure of total herbivore pressure) to those occurring in the presence of single herbivore species (a measure of a single species' total effect). Thus, even though Paine's study was focused on specific food web links, the outcome of any perturbation experiment will include all direct and indirect, trophic and nontrophic effects between the species. The interaction remains embedded in its network context and is considered a systems-level response, not a property of an individual link that is independent of the network. This is an important distinction, especially for estimates of weak interactions, which exhibits the greatest variability, likely because when the direct effect is weak, there is a greater influence of indirect interactions (Berlow, 1999).

Despite these issues, system-level measures of SIS remain the most useful tool for distinguishing between systems with keystone versus "diffuse" predation regimes and identifying keystone species for conservation. System-level measures of SIS automatically take into account nontrophic (i.e., nonforaging) interactions, such as ecosystem engineering, interference competition, stress amelioration, and behavioral modifications, which can have important consequences for community structure and dynamics, as it is described above for beavers, sapsuckers, kangaroo rats, and prairie dogs. Unlike trophic interactions, nontrophic interactions have no clear theoretical framework and are unwieldy to include in models of link-level SIS embedded in large ecological networks. However, attempts are being made toward resolving this problem (Arditi *et al.*, 2005; Goudard and Loreau, 2008; Kefi *et al.*, 2012).

These two distinct approaches to measure SIS have created a gap that has long divided theoretical and empirical investigations of food web dynamics (Laska and Wootton, 1998). However, recent work is bridging this gap and leading to exciting new insights. Novak and Wootton (2010) show how reformulating two of the most commonly used system-level measures of SIS, Paine's index and the Dynamic index, can relieve some of their assumptions, such as linear predator functional responses and populations closed to immigration. In addition, Novak and Wootton (2008) describe an observation-based empirical method for estimating interaction strengths of an individual link that is directly comparable to most theoretical estimates. The method estimates species-specific attack rates of consumers expected to exhibit a saturated feeding rate at high-prey density (i.e., a type II functional response). Novak tested this observational approach in the rocky intertidal on the strength of predation of a whelk, *Haustrum scobina*, on its mussel prey, *Xenostrobus pulex* (2010). The observational method produced estimates in high concordance with experimental estimates derived from a best-fit predator–prey model of year-long mussel population dynamics under replicated manipulations of the whelk (Novak, 2010).

Several recent investigations have predicted SIS based on constraints imposed by bioenergetics as described by the Metabolic Theory of Ecology (MTE, Brown *et al.*, 2004). The basis of the MTE is that body size and temperature are critical determinants of metabolic rate, which is the fundamental process affecting higher-level processes in organisms, such as consumption rates. Emmerson and Raffaelli (2004) explored the relationship between predator/prey body size ratios and the strength of species interactions in the Ythan Estuary food web. They found that the patterning of predator and prey body sizes affects the arrangement of interaction strengths and increases the probability of stability in a model food web. Berlow *et al.* (2009) used allometric scaling rules to describe the relationships between body size, metabolism, and food consumption. Putting allometrically scaled parameters into a dynamic food web model led to a realistic method of predicting SIS simply from species body sizes and abundances. Recent laboratory studies have begun to test these theoretical allometries and parameterize food web models by directly measuring the effect of body size and temperature on estimates of link-level SIS (e.g., Rall *et al.*, 2010). Results to date are consistent with the MTE and show great potential for the use of allometric–trophic network models, based on body mass and trophic information, for identifying keystone species and predicting the effects of climate change on food web structure and dynamics.

Despite the issues of reconciling the many theoretical and empirical meanings and measures of SIS, common patterns have emerged over the past two decades. Field evidence for a

strong effect of body size on SIS comes from a study that quantified the relative influences of several biotic and abiotic factors on SIS in rocky intertidal ecosystems using species manipulation experiments (Wood *et al.*, 2010). Another common pattern is that communities repeatedly exhibit only a few strong and many weak interactions (e.g., Bascompte *et al.*, 2005; Paine, 1992; Wood *et al.*, 2010). Furthermore, it appears that the distribution of SIS is nonrandom within a food web. Bascompte *et al.* (2005) analyzed the distribution of SIS in the large pelagic marine food web of the Caribbean Sea. SIS were structured nonrandomly, with the likelihood of two strong interactions connected through two consecutive levels of the food web occurring less frequently than expected by chance. When strong interactors did occur on consecutive levels of the food web, they were accompanied by the stabilizing effect of strong omnivory more often than expected by chance. Using a food web model the authors showed that this patterning of SIS reduced the likelihood of trophic cascades after removing top species from the web, as would occur with fishing.

The discoveries of these patterns have considerably advanced the understanding of how SIS structures communities. Nevertheless, the causes and consequences of variability in species interactions are just beginning to be understood. One interesting example of how context-dependent variability in SIS results in a stable community structure is a study of the strength of predation on juvenile barnacles in small to large patches of cleared rock, which otherwise had an invariant pattern of sparse barnacle cover (Navarrete and Berlow, 2006). Small patches were maintained by predation from whelks and by adult barnacles filtering potential larval settlers, but predation in larger patches was low since the adult barnacle could not filter very far away from themselves and the whelks do not venture far away from the protection of the large barnacles. Thus, the variability in predation strength led to a stable pattern of barnacle abundance.

Keystone species will be those with large *per capita* SIS. Identifying such interactors in this way has shed light on the traits, both individual and community, that generate species with high *per capita* effects. As implied by the alternative meanings for interaction strength offered by Macarthur (1972), however, species impacts will also vary with density. Regardless of how large its *per capita* effect, a single individual will never have the same impact as that of many individuals (Berlow *et al.*, 1999; Navarrete and Menge, 1996). Thus, full understanding of interaction web dynamics, and whether a species is a keystone, a strong interactor, a weak interactor, or a dominant, will depend on knowledge of both *per capita* interaction strength and per population interaction strength. Also note that only a keystone species will have high values of both *per capita* and per population indices of interaction strength, at least at the community or ecosystem level.

Conservation Implications

Biologists are increasingly being asked to inform management decisions concerning the preservation of biological diversity and ecosystem "integrity." Because keystone species can be critical to maintain the species diversity and ecosystem functioning, some have argued that focusing conservation on them should be a priority. A current topic of debate is how best to design reserves, both marine and terrestrial, that will protect species diversity.

Many species have been presumed to be keystones based on observation and conjecture. In many cases no experimental removals were done to determine the actual effect of the species presumed to be a keystone on the rest of the community. As outlined earlier, identification of keystones should involve manipulations to quantify the effect of the suspected keystone on the community. However, from a management standpoint this is not always feasible. It sometimes takes years for field manipulations to yield results, whereas conservation decisions often need to be made in a much shorter period of time. Also, as noted earlier, in addition to potential time constraints, manipulations of some suspected keystones may not be practically feasible or ethical. Removing killer whales, sea otters, lions, polar bears, or other large, charismatic, and often endangered species is not a realistic option. The logistical difficulties in designing controlled exclosure or removal experiments for animals of large size and/or with large ranges are daunting.

Despite the difficulties, and accepting the view that conservation should focus on ecosystems rather than species, the keystone species concept offers several important insights that are relevant to management (Power *et al.*, 1996):

- Seemingly scarce and unimportant species may have unexpectedly large, dramatic effects on communities and ecosystems.
- Conserving species may depend strongly on other species in the community with which the target species has little seeming association, whether as prey, competitor, mutualist, or predator.
- Loss of a species, particularly one high in the food chain, may have surprising and extensive consequences for the remainder of the community or ecosystem.

All these points suggest that great caution is necessary before decisions are made that may result in the loss of a native species or the introduction of exotic species. In this study the authors recognize and agree with the need to make management and policy recommendations on the basis of present knowledge. The authors also suggest that although much has been learned and the pace of new knowledge is increasing, the current state of knowledge is still insufficient to allow making such recommendations with a high level of confidence. In particular, additional study is needed on those putative keystones for which evidence is scant, and continued study is needed on ways to detect keystone species and systems with keystone species. Other urgent issues are to determine the commonness of keystone species and keystone-dominated systems, the patterns of interaction strengths in representative communities, and the environmental and organismal conditions and traits that foster keystone species.

The authors conclude that the concept of keystone species is a powerful one, having broad importance and application to ecological theory, ecosystem dynamics, and conservation.

Many problems remain to be answered, but important advances on these issues are continuously being made.

Acknowledgments

We thank reviewers of the originally published version of this paper for constructive criticism (Kealoha Freidenburg, Heather Leslie, Bob Paine, Mary Power, Eric Sanford, David Skelly, and Michael Webster). This article was written while the authors were supported by grants from the National Science Foundation, the Wayne and Gladys Valley Foundation (B.A.M.), the David and Lucile Packard Foundation (B.A.M.), and an NSERC Graduate Fellowship (A.C.I.). Contribution No. 384 from PISCO, the Partnership for Interdisciplinary Studies of Coastal Oceans: A Long-Term Ecological Consortium, funded by the David and Lucile Packard Foundation.

See also: Competition, Interspecific. Conservation Biology, Discipline of. Ecology, Concept and Theories in. Ecosystem Function, Principles of. Food Webs. Predators, Ecological Role of. Species Coexistence

References

Agrawal AA, Ackerly DD, Adler F, *et al.* (2007) Filling key gaps in population and community ecology. *Frontiers in Ecology and the Environment* 5: 145–152.

Allison GW, Menge BA, Lubchenco J, and Navarrete SA (1996) Predictability and uncertainty in community regulation: Consequences of reduced consumer diversity in coastal rocky ecosystems. In: Mooney HA, Cushman JH, Medina E, Sala O, and Schulze E-D (eds.) *Functional Roles of Biodiversity: A Global Perspective*, pp. 371–392. Chichester, United Kingdom: John Wiley & Sons.

Arditi R, Michalski J, and Hirzel AH (2005) Rheagogies: Modelling nontrophic effects in food webs. *Ecological Complexity* 2: 249–258.

Bascompte J, Melian CJ, and Sala E (2005) Interaction strength combinations and the overfishing of a marine food web. *Proceedings of the National Academy of Sciences of the United States of America* 102: 5443–5447.

Berlow EL (1999) Strong effects of weak interactions in ecological communities. *Nature* 398: 330–334.

Berlow EL, Dunne JA, Martinez ND, Stark PB, Williams RJ, and Brose U (2009) Simple prediction of interaction strengths in complex food webs. *Proceedings of the National Academy of Sciences of the United States of America* 106: 187–191.

Berlow EL, Navarrete SA, Briggs CJ, Power ME, and Menge BA (1999) Quantifying variation in the strengths of species interactions. *Ecology* 80: 2206–2224.

Berlow EL, Neutel AM, Cohen JE, *et al.* (2004) Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology* 73: 585–598.

Brown JH (1998) The desert granivory experiments at Portal. In: Resetarits Jr. WJ and Bernardo J (eds.) *Experimental Ecology: Issues and Perspectives*, pp 71–95. Oxford, United Kingdom: Oxford University Press.

Brown JH, Gillooly JF, Allen AP, Savage VM, and West GB (2004) Toward a metabolic theory of ecology. *Ecology* 85: 1771–1789.

Dayton PK (1972) Toward an understanding of community resilience and the potential effects of enrichments to the benthos at McMurdo Sound, Antarctica. In: Parker BC (ed.) *Proceedings of the Colloquium on Conservation Problems in Antarctica*, pp 81–96. Lawrence, Kansas: Allen Press.

Elton CS (1927) *Animal Ecology.* London: Sidgwick & Jackson.

Emmerson MC and Raffaelli D (2004) Predator–prey body size, interaction strength and the stability of a real food web. *Journal of Animal Ecology* 73: 399–409.

Estes JA and Duggins DO (1995) Sea otters and kelp forests in Alaska – Generality and variation in a community ecological paradigm. *Ecological Monographs* 65: 75–100.

Estes JA, Smith NS, and Palmisano JF (1978) Sea otter predation and community organization in western Aleutian Islands, Alaska. *Ecology* 59: 822–833.

Goudard A and Loreau M (2008) Nontrophic interactions, biodiversity, and ecosystem functioning: An interaction web model. *American Naturalist* 171: 91–106.

Kefi S, Berlow EL, Wieters EA, *et al.* (2012) More than a meal… Integrating non-feeding interactions into food webs. *Ecology Letters* 15: 291–300.

Knapp AK, Blair JM, Briggs JM, *et al.* (1999) The keystone role of bison in north American tallgrass prairie – Bison increase habitat heterogeneity and alter a broad array of plant, community, and ecosystem processes. *Bioscience* 49: 39–50.

Laska MS and Wootton JT (1998) Theoretical concepts and empirical approaches to measuring interaction strength. *Ecology* 79: 461–476.

Macarthur RH (1972) Strong, or weak, interactions? *Transactions of the Connecticut Academy of Arts and Sciences* 44: 177–188.

Menge BA, Berlow EL, Blanchette CA, Navarrete SA, and Yamada SB (1994) The keystone species concept – Variation in interaction strength in a rocky intertidal habitat. *Ecological Monographs* 64: 249–286.

Menge BA, Lubchenco J, Ashkenas LR, and Ramsey F (1986) Experimental separation of effects of consumers on sessile prey in the low zone of a rocky shore in the Bay of Panama – direct and indirect consequences of food web complexity. *Journal of Experimental Marine Biology and Ecology* 100: 225–269.

Menge BA and Sutherland JP (1987) Community regulation: variation in disturbance, competition, and predation in relation to gradients of environmental stress and recruitment. *American Naturalist* 130: 730–757.

Mills LS, Soule ME, and Doak DF (1993) The keystone-species concept in ecology and conservation. *Bioscience* 43: 219–224.

Naiman RJ, Melillo JM, and Hobbie JE (1986) Ecosystem alteration of boreal forest streams by beaver (*Castor canadensis*). *Ecology* 67: 1254–1269.

Navarrete SA and Berlow EL (2006) Variable interaction strengths stabilize marine community pattern. *Ecology Letters* 9: 526–536.

Navarrete SA and Menge BA (1996) Keystone predation and interaction strength: Interactive effects of predators on their main prey. *Ecological Monographs* 66: 409–429.

Novak M (2010) Estimating interaction strengths in nature: Experimental support for an observational approach. *Ecology* 91: 2394–2405.

Novak M and Wootton JT (2008) Estimating nonlinear interaction strengths: An observation-based method for species-rich food webs. *Ecology* 89: 2083–2089.

Novak M and Wootton JT (2010) Using experimental indices to quantify the strength of species interactions. *Oikos* 119: 1057–1063.

Paine RT (1966) Food web complexity and species diversity. *American Naturalist* 100: 65–75.

Paine RT (1992) Food-web analysis through field measurement of *per capita* interaction strength. *Nature* 355: 73–75.

Paine RT (1994) *Marine Rocky Shores and Community Ecology: An Experimentalist's Perspective.* Oldendorf/Luhe, Germany: Ecology Institute.

Pincebourde S, Sanford E, and Helmuth B (2008) Body temperature during low tide alters the feeding performance of a top intertidal predator. *Limnology and Oceanography* 53: 1562–1573.

Pincebourde S, Sanford E, and Helmuth B (2009) An intertidal sea star adjusts thermal inertia to avoid extreme body temperatures. *American Naturalist* 174: 890–897.

Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *Bioscience* 46: 609–620.

Rall BC, Vucic-Pestic O, Ehnes RB, Emmerson M, and Brose U (2010) Temperature, predator–prey interaction strength and population stability. *Global Change Biology* 16: 2145–2157.

Sanford E (1999) Regulation of keystone predation by small changes in ocean temperature. *Science* 283: 2095–2097.

Willson MF and Halupka KC (1995) Anadromous fish as keystone species in vertebrate communities. *Conservation Biology* 9: 489–497.

Wood SA, Lilley SA, Schiel DR, and Shurin JB (2010) Organismal traits are more important than environment for species interactions in the intertidal zone. *Ecology Letters* 13: 1160–1171.

Wootton JT and Emmerson M (2005) Measurement of interaction strength in nature. *Annual Review of Ecology Evolution and Systematics* 36: 419–444.

L

Lake and Pond Ecosystems

Christian Lévêque, Institut de Recherche pour le Développement (IRD), Paris, France

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 633–644, © 2001, Elsevier Inc.

Glossary

Ancient lakes Lakes with a persistence of more than 100,000 years are called long-lived or ancient lakes.

Eutrophication The process of enrichment of a water body due to an increase in nutrient loading.

Species flocks An aggregate of closely related species that share a common ancestor and are endemic to a geographically circumscribed area.

Broad Characteristics of the Biodiversity in Lakes and Ponds

Freshwater habitats are widely considered to be transient in time and space in comparison with both terrestrial and marine habitats. This is true for many lakes and ponds. However, depending on the origin of lakes there are great differences in the nature and the diversity of their biota. Three broad categories of lakes may be recognized:

1. *Lakes and ponds that are permanently or frequently connected to large river systems*. This category includes river-lakes (i.e., Lake Geneva) or lakes that are part of a large floodplain system such as "varzea" lakes. In these lakes, exchanges of flora and fauna occur with the main river system so that their biota is usually greatly similar to the biota of the river system itself with the exception of a few species adapted to still waters. Many endoreic lakes (Lake Chad, Aral Sea) also belong to this group.
2. *Isolated lakes with a limited drainage system*. The biota of the lakes in this category has evolved in isolation from others for a more or less long period of time leading to speciation and endemicity when the period is long enough. The associated ice ages at higher latitudes and altitudes were the phenomena that created most of the lakes in existence today. Therefore the great majority of existing lakes (around 10,000 exceeding 1 km^2 in extent) are geologically very young and occupy basins formed by ice masses or glacial erosion after the retreat of continental ice sheets some 10,000 years ago. All such lakes are expected to fill slowly with sediment and to disappear in the future, along with any isolated biota. Compared to ancient lakes, they acquired their fauna and flora via the rivers that supply them with water as a result of runoff in their basin and from aerial transport by wind or animals. Only a few existing lakes are known to be much older, and most of them occupy basins formed by large-scale subsidence. They may date back at most 20 million (Lake Tanganyika) or 30 million (Lake Baikal) years. These so-called ancient lakes are of particular interest for biodiversity because they exhibit a rich endemic fauna for several major groups of animals. There is also good evidence that some extinct lakes were also very large and long-lived under different climatic and tectonic conditions.
3. *Temporary lakes and ponds whose water budget is controlled by the climate regime*. The fauna and flora in these lakes exhibit special biological adaptations to seasonal drying.

Origin and Peculiarities of Freshwater Biota

It is thought that the early evolution of all the major animal phyla took place in the sea. Most phyla are predominantly marine and benthic: 32 phyla are found in the sea with 11 exclusively marine, whereas 14 are represented in freshwater and only 12 are found on land (May, 1994). Compared to the sea, freshwaters are deficient in major taxa and there are no uniquely freshwater metazoan phyla. The osmotic challenges of life in freshwaters probably discouraged invasion of the habitat by many marine invertebrates. It explains probably the tendency in freshwater invertebrates for larger but fewer eggs than in marine relatives: they must eclose with fully developed osmoregulatory capacities to be at a more advanced stage to cope with the highly dilute surrounding.

Another difference between the species richness of marine and freshwater zooplankton derives from the necessity of diapause or other resting mechanisms as a condition for

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00082-4

persistent successful radiation in freshwaters (Lehman, 1988). Freshwater invertebrates developed anabiotic devices: special resistant eggs, cysts, and other resting stages that are produced to tide the animal over periods of desiccation, extreme cold, heat, anaerobic situations, lack of food, and other adverse conditions. In addition to withstanding unfavorable conditions, resistant stages have the further function of making overland transport and geographical dissemination possible. Without such a function, colonization of freshwater areas would be slow and difficult in the discontinuum of isolated lakes and ponds.

The Latitudinal Gradient

It is usually assumed that species diversity increases from high to low latitudes for most of the major groups of plants and animals and that highest values occur at low latitudes. Indeed, the diversity of marine plankton decreases from low latitudes to high ones, so that tropical and subtropical ocean waters exhibit rich diversity of zooplankton whereas Arctic and Antarctic waters tend to be dominated by copepods and euphausiids. In freshwaters however, the latitudinal trend in species richness is the opposite. Tropical lakes have abbreviated zooplankton faunas compared with temperate locales (Fernando, 1980); they are depauperate in large-bodied species of copepods and Cladocera, and limnetic rotifers are likewise poorly represented.

It could be assumed that the associated ice ages at higher latitudes and altitudes were the phenomena that created most of the lakes in existence today. They are therefore very young compared to ancient lakes, and they acquired their fauna and flora via the rivers that supply them with water via runoff in their basin and via aerial transport by wind or animals.

For fish, the species richness is actually smaller in north temperate lakes of glacial origin than in long lived lakes from tropical areas. At least the endemicity is much lower in temperate lakes than in tropical lakes.

Dumont (1994), in a review of the species richness of the pelagial zooplankton in ancient lakes, provided also evidence that these water bodies have simple pelagial communities. Among 14 pre-Pleistocene lakes across the world, at least one Cyclopoid copepod species is present, often in the genus Cyclops or Mesocyclops, a group of microraptorial species feeding on rotifers, small Crustacea, and immature stage of other copepods.

The number of species regularly found in the pelagic plankton of ancient lakes (pre-Pleistocene) varies from 3 (Lake Tanganyika) to approximately 15 to 20 (up to 5 copepods, 5 cladocerans, 10 rotifers) in Lakes Victoria, Biwa, and Titicaca. In contrast, "young" lakes may have up to 10 species of copepods, 10 of Cladocera, and 10 to 15 species of rotifers occurring together. In the oldest lakes (Baikal, Tanganyika), which also happen to be the deepest, this simplification has gone extreme and the food web reduces to a linear chain.

The question has been raised as to why Cladocera have been almost completely eliminated from some ancient lakes such as Tanganyika, Baikal, and even Malawi (Dumont, 1994). They are able to eat both large items and microplankton and seem all but competitively inferior to other species for food acquisition. Predation had been advocated as a possible cause. In clear-water lakes such as Tanganyika and Baikal, visual predation by fish is more effective than in turbid lakes, and large clumsy swimmers like big Daphnia are likely to be preyed to extinction before the relative small transparent, agile swimmers like the Calanoids. An experimental demonstration of this hypothesis has been the disappearance of all Cladocera from the pelagial of Lake Kivu, within a decade, after the introduction of the zooplanktivore clupeid Limnothrissa miodon, native from Lake Tanganyika (Dumont, 1986).

Vertical Distribution in Lakes

Aquatic organisms are not evenly distributed along depth. Water characteristics are relatively uniform in shallow lakes, which are mixed by winds. However, deeper lakes exhibit patterns of vertical gradients for temperature and light. Briefly speaking, the lake is divided by a thermocline into an upper layer, the epiliminion, and a lower layer, the hypolimnion. Life occurs in the oxygenated upper layer while the lower layer is deoxygenated. In most stratified lakes therefore, biota is very depauperated, except for bacteria, below a few 10 meters in depth. There is an exception, Lake Baikal, which is the deepest lake in the world with a maximum depth of 1620 m. The mechanism of mixing of the deep water zone is still not completely understood, but the entire water column is well oxygenated. Life for fish and invertebrates is therefore possible from surface to maximum depth, which is exceptional for freshwater systems. Lake Baikal is therefore unique among inland systems to the study of bathymetric segregation and includes some of the deepest occurring freshwater animals. Among fishes, the family Abyssocottidae contains 20 species distributed throughout the depths of the lake. Species of the genera *Abyssocottus*, *Cottinella*, and *Neocottus* are adapted to the deep water way of life in that they do not occur above 400 m, the size of the eyes is reduced, and they are physiologically adapted to resisting high pressures (Sideleva, 1994).

The discovery of a deep sea hydrothermal fauna in the 1980s was a surprise for marine biologists. Similarly, hydrothermal vents have been discovered in Lake Baikal, at a depth of 440 m on the sediment floor of Frolikha Bay (Crane *et al.*, 1991), at the foot of an east-west trending fault. Photographs reveal that the center of the vent field is covered by a near-continuous bacterial mat. A white sponge encrusts small cobbles at the periphery of the vent field. Coiled gastropods and whitish translucent amphipods are found among the sponges and on the sediment at the edge of the bacterial mat.

Relationship between Species Richness and Area in Lakes

Community ecologists used to compare isolated freshwater systems to biogeographic islands. The relationship of species number to area containing those species is a well-known empirical observation, and a power function is widely used to describe this pattern mathematically: $S = cAZ$, where S is the number of species, A the area, Z the slope of the regression line, and c a constant. It can also be expressed as $\log S = c + Z\ logA$.

The effect of lake size on species richness of invertebrates has been demonstrated. For crustacean zooplankton, species richness is also significantly correlated with lake surface area (Dodson, 1992). The species area curve for North American lakes is statistically different from and steeper than the corresponding European curve (slopes, respectively, 0,094 and 0,054). The log species richness is also correlated to log of the average photosynthetic flux per cubic meter and log number of lakes within 20 km of the target lake. For 66 North American lakes, the three variables can be combined in a multiple linear regression model, which explains 75% of the variation in log species richness (Dodson, 1992).

Species richness of aquatic birds also increases with lake size. In Swiss lakes, the species number increase steeply with lake size up to 50 km^2 and species richness depends more closely on lake area in fish eaters and diving ducks than in dabbling ducks (Suter, 1994). Actually lake size explains 70 to 85% of the variation of abundance and species richness in fish eaters and diving ducks but only 64% of species richness in dabbling ducks. In Florida lakes, bird species richness was also positively correlated to lake area and to total water column phosphorus concentration value (WCP) for each lake. The multilinear Log (species richness) = 1.12 + 0.56 Log (Lake area) + 0.12 Log (WCP) and accounts for 77% of the variance in species richness (Hoyer and Canfield, 1994).

Morphometry and Species Richness

The species diversity in a lake is a function of the diversity of habitats: the more ecological niches in the lakes, the more species may be expected. The lake's morphometry is basic to its structure: deep, steepsided lakes do not offer as many biotopes as shallow, flat lakes. For the latter, most of the lake bottom may be colonized by plants and animals (the benthic flora and fauna), while in deep lakes, only a small part of the lake bottom is colonized. Generally speaking, deep lakes are dominated by planktonic organisms, which are floating or weakly swimming organisms, usually associated with suspended particles. In shallow lakes, benthic organisms are dominant and the heterogeneity of lake bottom, as well as the development of macrophytes, may increase the diversity of benthic species.

Evaluating Biological Diversity

Despite the efforts of taxonomists, a good estimation of the total number of species occurring in freshwater lakes and ponds does not exist. We shall provide here some recent findings about aquatic biodiversity.

Diversity of Plankton and Microbial Loop

Three major size classes are usually recognized in pelagic plankton: microplankton (20–200 μm), nanoplankton (2–20 μm), and picoplankton (0.2–2.0 μm). In the late 1970s, phototrophic picoplankton was discovered in great abundance in both marine and freshwater ecosystems. However, identifying picoplankton causes considerable taxonomic problems due to the very small sizes of these organisms. We do not know how many bacterial species exists in the world, because bacteria cannot be differentiated under the microscope; we do not even know the right order of magnitude. A new way of classification has been proposed, based on the sequences of ribosomal RNA that led to a phylogenetic classification of bacteria. It is becoming apparent that the genetic diversity among bacteria is much wider than that among the animals and plants.

Most heterotrophic nanoplankters are small (2–5 μm), colorless flagellated protists. They grow at about the same rate as bacteria and are capable of consuming the entire bacterial production. Meanwhile, they regenerate significant amounts of nutrients and serve as prey for micro- and mesoplankton.

The importance of bacteria and protozoa activities in the trophic structure of lacustrine food chains has been largely underestimated in the past. The major role played by microorganisms in controlling energy and nutrient fluxes is now better understood following the discovery of the microbial loop and its role as a source or a sink for carbon and energy flow to higher trophic levels in pelagic systems. We know now that these microorganisms can control major fluxes of energy and nutrients. In some cases, 50% of the photosynthetic production does not pass directly to higher trophic level but is diverted into a microbial loop where nutrients are rapidly remineralized and fed back to the dissolved inorganic pools.

Diversity in Freshwater Sediments

About 175,000 species of organisms associated with freshwater sediments have been described, but the true number is much higher than this (Palmer *et al.*, 1997). The number of species in most taxa can scarcely be estimated and global estimate of microbial diversity remains controversial. For example, some specialists estimate that there are hundred of thousands of aquatic nematodes and only a small percentage of these have been described. Rotifer species diversity is also poorly known for freshwater sediments, but it is estimated that there are thousands of undescribed species.

Most freshwater sediment species are small and concentrated in the upper sediment layers. Availability of light limits the development of plants and photosynthetic bacteria, which are therefore scarce or absent in most sediments. Moreover, oxygen level may influence species richness and the number of species is low in anoxic waters (see **Table 1**).

Diversity in Fish

Presently, 25,000 fish species have been described. Some 10,000 species are found only in freshwaters, a large proportion of which occurs in lakes and ponds. The freshwaters are therefore disproportionately rich in species of fishes on the basis of area when compared to oceans. That could be viewed as the result of the patchy nature of inland waters and the resulting high endemicity of the biota. Fish live in almost every conceivable type of aquatic habitat. They exhibit enormous diversity in size, shape, and biology.

Other vertebrate species occur in freshwaters: a few mammals, several reptiles, and many birds and amphibians. There

Table 1 Species richness of the freshwater sediment biota for many habitat types

Taxon	*Number of species described*	*Probable number of species*	*Range of local species richness*
Bacteria	>10,000	Unknown	>1,000
Algae	14,000	>20,000	0–1000
Fungi	600	1,000–10,000	0–300
Protozoa	<10,000	10–20,000	20–800
Plants	1,000	Unknown	0–100
Invertebrates	70,000	>100,000	10–1,000
Aschelminthes	4,000	>10,000	5–500
Annelida	1,000	>1,500	2–50
Mollusca	4,000	5,000	0–50
Acarii	5,000	>7,500	0–100
Crustacea	8,000	>10,000	5–300
Insecta	45,000	>50,000	0–500
Others	1,400	>2,000	0–100

Numbers are rough estimates and derived from many sources.

Table 2 Number of fish species recorded from several lakes connected to rivers systems

Lake		*Latitude*	*Area*	*Number of fish*
Chad	Africa	13° N	10–20,000	137
Turkana	Africa	3° N	6,750	51
Chilwa	Africa	15° S	675	31
Ngami	Africa	20° S	150	48
George	Africa	0°	270	30
Huron	North America	44° N	59,600	99
Erie	North America	42° N	25,700	113
Michigan	North America	44° N	58,000	114
Superior	North America	47° N	82,400	67
Great Bear	North America	66° N	31,150	12
Great Slave	North America	61° N	27,200	26
Big Trout	North America	54° N	616	24
Chapala	Central America	20° N	1,080	14
Nicaragua	Central America	11° N	8,200	40
Maggiore	Europe	46° N	676	21
Windermere	Europe	54° N	15	9
Ladoga	Europe	61	18,400	48
Aral sea	Europe	45° N	64,500	17

See also Table 3 for ancient lakes.

is no quantitative evaluation of the number of vertebrates whose life cycles include lakes or ponds, but it is far from negligible (see **Table 2**).

Biological Diversity and Ecosystem Functioning

Energy and nutrients in an ecosystem are transferred through successive trophic levels. Photosynthesis provides the basic food for herbivorous animals, which are eaten by the carnivores. Therefore, knowledge of the role of individual species and their relationships in aquatic systems is critical to understand the functioning of the system as a whole. Limnologists pointed out several key issues to the study of the relationships between species diversity and lake functioning.

Food Webs

Food webs are diagrams depicting which species in a community interact in feeding and describing which kinds of organisms in a community eat which other kind. Food webs are thus caricatures of nature, but they give a picture of the processes at work in ecosystems. Connectivity food webs are describing pathways along which feeding interactions occur. These interactions change at least seasonally and not all interactions are equally strong. The interaction web emphasizes connections that appear to have a large effect on the dynamics of the food web structure and function. Food webs occupy a central position in community ecology. Many important interactions (e.g., competition, predation) cannot be isolated from a food web context because the outcome of these interactions can be modified directly and indirectly by other members of the web.

For a long time food webs served principally as heuristic devices, useful in depicting complex ecosystems as diagrams composed of many interactive parts and enhancing our understanding of pathways of energy and material transfer in aquatic ecosystems. However, the recent surge of interest in food webs seems related to the question of the functional role of biodiversity (discussed later).

Few if any of the aquatic food webs are unimpacted by humans both at a local and a global scale. For example, fisheries food webs are complex, involving multiple trophic levels at several spatial and temporal scales. Fish species offer a wide range of body sizes and feeding habits, and thus have a variety of food web roles and interactions with other species. Exploitation of fishes may result in major changes in food webs. However, the consequences of species removal through fisheries are an almost unexplored field of research in most freshwater systems.

The Top-Down Control

In the classical limnological approach, it was usual to consider freshwater ecosystems as operating in a physical-chemical milieu that, largely through nutrient availability, conditions the food chain from primary producers to top predators. In this "bottom-up" control, competition between primary producers for limiting nutrients determines the state of higher trophic levels. A reverse viewpoint, the "top-down control" slowly became prominent. It argues that the effects of top predators cascade down the trophic chain and are responsible for controlling the state of the entire ecosystem. The predators, near or at the top of the trophic pyramid, may be fishes but also may be birds, mammals, and so on, as well as invertebrates. Through grazing, for instance, fish have direct effects on the composition and abundance of phytoplankton, periphyton, and macrophytes, as well as on the dynamics of plankton and benthic communities. Size-selective predation by fish may not only play a major role in the population dynamics of prey species, but also result in shifts in the relative abundance of species.

Predation is now considered to be a major driving force in shaping zooplankton communities. A great number of papers emphasized the size-related alterations in zooplankton communities as a consequence of planktivorous fishes, which select the largest available prey and may rapidly reduce the density of large zooplankters, resulting in a shift of the prey community to small species, predominantly rotifers and small cladocerans. Extinction of large zooplankton has been documented in several habitats, usually following the introduction of new species of planktivorous fish.

Trophic cascades from fishes to water quality in lakes are among the clearest examples of feedbacks from population to ecosystem processes (Carpenter and Kitchell, 1993). A shift in the species composition and size distribution of the fish assemblages alters the community composition and size distribution of the herbivorous zooplankton. The impact of herbivory on phytoplankton depends on the relative abundance of certain herbivores with wide diets, high grazing rates, and rapid population growth rates. Population of these keystone herbivores are sensitive to fish predation. In addition, the size distribution of fishes and zooplankton and their migratory behavior largely determine the rate and spatial pattern of nutrient recycling in pelagic ecosystems. In whole-lake experiments, manipulations of fish community structure have caused significant changes in primary production, algal biomass, nutrient recycling, and sedimentation rates.

Relationships Between Biodiversity and Ecosystem Stability

A major concern for limnologists is to predict response to stress. For a long time, the so-called conventional wisdom in ecology was that increased complexity within a community leads to increased stability. Complexity is used here to mean more species, more interactions between them, and more pathways. The basic assumption is that if the number of pathways increases, any blockage at one point of the network would be compensated for by the opening of another pathway. However, until now this conventional wisdom has not received much support from field or experimental work. Therefore some basic questions remain open and are of particular concern for freshwater lakes:

- How is system stability and resistance affected by species diversity, and to what extent could the integrity and sustainability of ecosystems be maintained in spite of species deletions resulting from degradation of environmental conditions?
- Are rare species an insurance of ecosystem stability? Do these rare species play a role in ecosystem functioning? One hypothesis is that the most stable ecosystems in terms of key functions are those richest in species. However, well documented studies of rare species substituting for declining common species in the maintenance of key freshwater ecosystem functions following disturbance are scant.

Role of Intra- and Interspecies Communication Systems in Ecosystem Dynamics

The structure of the biota is determined by complex interactions between individual organisms acting at different trophic levels. It is now becoming clear that besides energy transfer from one trophic level to another, there is an exchange of information between trophic levels through infochemicals. Moreover, sexual pheromones, but also sounds, electric signals, and visual communication play a role in shaping the structure of the biotic community. This diverse communication network has biological consequences and may modify the behavior of aquatic animals, such as migrations. This is a fairly new field of research.

Biological Productivity and Biodiversity: The Case of Eutrophication

The biological structure and internal biological control mechanisms of freshwater lakes are highly affected by lake water nutrient level and by the extent of nutrial loading. Limnologists distinguish oligotrophic lakes, which are generally deep with steep slope and are characterized by low nutrient levels and clear water. The biomass at all trophic levels is small. On the opposite, eutrophic lakes are often shallow with

gradually sloping edges. The most characteristic features are high nutrients levels, the abundance of plankton, and low water clarity.

One concept of lake succession considers that lakes pass through different trophic states, beginning with low fertility or oligotrophy, gradually moving to a moderately productive or mesotrophic state to reach finally the eutrophic stage. This succession may happen in undisturbed lakes. However, eutrophication (sometimes called cultural eutrophication) is now widespread as a result of human activities. Eutrophication may be defined as the process of enrichment of a water body due to an increase in nutrient loading. The most important nutrients causing eutrophication are phosphates, nitrates, and ammonia. All these chemicals are abundant in waters released from sewage treatment works and from surface and groundwater runoffs in intensive agriculture areas.

The most obvious consequence of eutrophication is the increased aquatic plants and phytoplankton growth, an overall increase in biomass, and a shift in species composition of the lake. For example, at low P-concentrations, north European shallow freshwater lakes are usually in a clearwater stage; submerged macrophytes are abundant, potential piscivores are present in large numbers, and predation pressure on zooplankton is consequently low. At some higher P-concentrations, there is a shift to a turbid stage: submerged macrophytes disappear and the fish stock changes. The fish biomass rises and there is a shift from a system dominated by pike (*Esox lucius*) and perch (*Perca fluviatilis*) to one exclusively dominated by planktivorous-benthivorous fish, mainly bream (*Abramis brama*) and roach (*Rutilus rutilus*).

The Case of Ancient Lakes Species Flocks

About a dozen lakes in the world are up to three orders of magnitude older than most others (**Table 3**) (Martens, 1997). Such lakes have exceptionally high faunal diversity and levels of endemicity.

An important characteristic of ancient lakes biodiversity is the existence of "species-flocks." An aggregate of several species should be identified as a flock only if its members are endemic to the geographically circumscribed area under consideration and are each others' closest living relatives. Briefly speaking, a species flock has to be monophyletic. At present, different rich species flocks for fish and invertebrates have been identified in various ancient lakes, which are therefore exceptional natural sites for the study of speciation patterns.

The processes accounting for these radiations are a matter of debate, but there is more and more evidence that sympatric speciation may occur in isolated water bodies. These species flocks are sometimes considered to be a world heritage that is endangered and has to be preserve from destruction by human activities (Coulter *et al.*, 1986; Nagelkerke *et al.*, 1995) (see **Table 3**).

Fish Species Flocks

The most striking feature of the Great East African Lakes (Victoria, Tanganyika, Malawi) is that each has its own highly endemic lacustrine Cichlid fauna. In Lake Victoria, according to our present knowledge, there is a cichlid species flock of more than 500 endemic haplochromine species. The true species number is almost certainly even higher (Seehausen, 1996). The age of this flock was estimated at 200,000 years, but it is most likely that Lake Victoria had entirely dried up as recently as 12,400 years ago, so that most of the endemic cichlid flock would have evolved during the past 12,400 years (Johnson *et al.*, 1996). In Lake Malawi, the diverse cichlid fauna of this lake could also total much more than 500 species. Species flocks are also reported for the clariid catfish *Dinotopterus* (10 species).

The Lake Tanganyika cichlids are slightly less diverse. However, morphological and electrophoretic data both suggest that several lineages of cichlids from Lake Tanganyika are much older than the Lakes Victoria and Malawi lineages and can be traced back to at least seven distinct ancestral lineages. Species flocks also occur in noncichlid families: seven Mastacembelid species, six species of the bagrid *Chrysichthys*, seven species of *Synodontis*, and four species of the Centropomid *Lates* (De Vos and Snoeks, 1994).

Rates of speciation in cichlids can be astonishingly fast. That has been known since the discovery of five endemic species of cichlids in Lake Nabugabo, a small lake less than 4000 years old and separated by a sandbar from lake Victoria. Still faster speciation rates were suggested by the finding that the southern end of Lake Malawi was dry only two centuries ago, while it is now inhabited by numerous endemic species and "color morphs" that are only found there and are believed to have originated during the past 200 years.

The remarkable diversity of the large barbs (genus *Barbus*) in Lake Tana (Ethiopia) constitutes a potential species flock that has been discovered recently. Nagelkerke *et al.* (1995) hypothesized that intralacustrine speciation has occurred among the barbs of Lake Tana and possibly is still going on.

Table 3 Summary of physical and biological characteristics of some of the larger extant ancient lakes

Lake	*Age (My)*	*Max. depth. (m)*	*Area (km^2)*	*Number of animal species*	*Number of endemic*	*Number of fish species*	*Number of fish endemic*
Baikal	35–30	1,700	31,500	1,825	982	56	27
Tanganyika	9–12	1,470	32,600	1,470	632	330	241
Malawi	4.5–8.6	785	30,800			>600	>600
Victoria	0,6?	70	70,000			>500	>500
Titicaca	3	284	8,448	533	61	29	23
Ohrid	2–3	295	348			17	2
Biwa	4	104	674	600	54	57	11

In South America, the native fish fauna of Lake Titicaca includes the genera *Trichomycterus* and *Orestias*, both endemic to the Andean Altiplano. Twenty-four *Orestias* species are presently recognized in Lake Titicaca (Lauzanne, 1992). However it is probably not a monophyletic group but rather an assemblage that includes, in part, several species flocks.

Lake Baikal (East Siberia) hosts a very diverse fauna, with some 2500 described species (most of which are endemic), which might constitute 50% of the total amount. At present, Lake Baikal comprises 56 species and subspecies of fish, which belong to 14 families (Sideleva, 1994), a group of six species and subspecies, belonging to three families (Thymallidae, Coregonidae, and Acipenseridae), which are relatively endemic; another group of nonendemic fauna includes 21 species and subspecies, which belong to Cyprinidae, Percidae, Cobitidae, Esocidae, Gadidae, Salmonidae, Siluridae, and Eleotridae.

In Asia, Lake Lanao was formed by a lava flow that dammed the streams flowing southwest. The cyprinid fauna presents a widely acknowledged example of adaptive radiation. Of the 23 cyprinid species presently known from Mindanao Island, 15 are reported from Lake Lanao (Kottelat and Whitten, 1996). Unfortunately overexploitation and exotic introductions have decimated the fauna, so that now only few endemic cyprinids are still present.

Lake Biwa is the largest and oldest lake in Japan. The deep basin as seen presently is supposed to have been formed 300,000 years ago. Most endemic fishes exist also since that time (Kawanabe, 1996). Some 500 plant and 600 animal species have been recorded. At present there are 71 species and subspecies of freshwater fishes found in Lake Biwa and its tributaries. There are 13 endemic species and subspecies of fish in Lake Biwa.

The only more or less pristine species flocks left in Asia are to be found in the lakes of the Malili River drainage of the Sulawesi Island (Celebes) in Indonesia (Kottelat & Whitten, 1996). Malili lakes (Lakes Towuti, Matano, Mahalona, Wawontoa, and Masapi) constitute a system of lakes partially isolated from each other and completely isolated from other freshwaters. As a result, most of the animal species known from the lakes are endemic.

Invertebrates

Mollusks focus our attention on parts of the world that seem to be hot spots of endemicity, where the resident clades are remarkably more diverse than in other similar environments. Long-lived lakes are prime examples of these evolutionary theaters. Particular clades, such as the hydrobioid and ceratioidean prosobranchs and the planorbid pulmonates, show repeated patterns of diversification in both extant and fossil long-lived lakes, revealing the common patterns that make them prone to speciate (Michel, 1994). The process of diversification is tied to intrinsic characters shared by many of these clades: reproductive and dispersal strategies (brooders and poor dispersers), genetic structure (tightly constrained genetic systems), morphology (often relatively thick and ornamented shells), substrate specificity (hard bottom stenotopy), and physiology (depth tolerance). The most notable examples of these evolutionary theaters are extant lakes Tanganyika, Baikal, Ohrid, Titicaca, and fossil lakes Idaho, Biwa, and Turkana. However, examples of gastropod radiations are also found in river systems (see **Table 4**).

Table 4 Endemism in ancient mollusk fauna (ancient lakes and rivers)

Lakes	*Baikal*	*Ohrid*	*Tangan.*	*Titicaca*	*Biwa*
Mussels	**2**	**1**	**12**	**0**	
% end.	0	0	75		
Sphaeridae	**10**	**9**	**3**	**4**	
% end.	30	22			
Prosobranchs	**72**	**47**	**52**	**16**	**11**
% end.	93	91	84	93	72
Pulmonates	**61**	**25**	**16**	**3**	**16**
% end.	77	48	6	33	62
Total species	**145**	**82**	**81**	**19**	**27**

Among Crustacea, gammarids have also undergone an enormous evolutionary radiation in Lake Baikal, with a total of 259 species, 98% of which are endemic. There are also several species flocks of Ostracods reported from ancient lakes (Martens *et al.*, 1994).

Major Threats to Biodiversity in Lakes

The concentration of people around freshwater systems has resulted in a much greater degree of degradation to these systems than most open marine or even terrestrial systems.

Competition for Water

Competition for water may result in the total or partial desiccation of lakes and ponds through various diversions and impoundment of tributaries. Water is withdrawn most often from aquatic systems for irrigation, flood control, and urban and industrial consumption. A spectacular example is provided with the Aral Sea, a large saline lake in the terminus of an extensive inland drainage basin in south-central Asia. Water diversion for irrigation purposes of most of the waters in inflowing rivers of the Aral sea, as well as poor agricultural practices, resulted in a marked fall of water level (c. 15 m) and an increase in salinity (from c. 10 to 30 g/l) since the 1960s. These changes have resulted in the degradation of the natural environment. Fish have virtually disappeared from the lake and the diversity of associated bird and wildlife communities has decreased. Many invertebrates also disappeared (Williams and Aladin, 1991).

Habitat Alteration

Siltation from erosion of the lake basin has direct adverse effects on fish by covering spawning sites, destroying benthic food sources, and reducing water clarity to visual feeding animals. However, the increased turbidity may also have indirect effects on biodiversity in lakes. Seehausen *et al.* (1997) provided evidence that increasing turbidity (as the consequence of deforestation and agricultural practices) by curbing the impact of sexual selection on sexual isolation is responsible for the

decline in cichlid diversity in Lake Victoria. Actually, mate choice in these cichlids is determined on the basis of coloration, and strong assortive mating can quickly lead to sexual isolation of color morphs, which is increasing and probably started in the 1920s. By constraining color vision, turbidity interferes with mate choice (Seehausen *et al.*, 1997). The reduced effectiveness of signals causes relaxation of sexual selection for color, with consequent loss of male nuptial coloration and erosion of species diversity due to a breakdown of reproductive barriers. Dull fish coloration, few color morphs, and low species diversity are found in areas that have become turbid as a result of recent eutrophication. This is proof that human activities that increase turbidity destroy the mechanism of diversification and the maintenance of diversity.

Species Introductions

The introduction of alien fish into inland waters has occurred all around the world. The main goals of deliberate introductions by fishery officers were initially to improve sport fisheries and aquaculture, or to develop biological control of aquatic diseases, insects, and plants, or else to fill supposed "vacant niches" and improve wild stocks in old or newly created impoundments.

The introduction of alien species has been considered as the main causes of extinction of endemic species flocks in several ancient lakes. In Lake Lanao, the introduction of the white goby (*Glossogobius giurus*) in the early 1960s resulted in the elimination of numerous species of endemic cyprinid fish. In Lake Titicaca the rainbow trout *Salmo gairdneri* was accused of seriously threatening the endemic *Orestias* fauna and for having been responsible for the disappearance of species such as *Orestias cuvieri*. In Lake Biwa, the recent increases in numbers of the exotic bluegill *Lepomis macrochirus*, black-bass *Micropterus salmoides*, and *Channa maculata*, have been mirrored by serious declines in the native species *Onchorhynchus rhodurus rhodurus* (an endemic), *Hemigrammocypris rasborella*, and *Hymenophysa curta*.

Much has been said about the impact of the introduction of the Nile perch on the hundreds of endemic haplochromines of Lake Victoria (Lévêque, 1997). In the early 1980s this impact was considered an ecological and conservation disaster (Coulter *et al.*, 1986). However, it was later recognized that predation by *Lates* may not be solely responsible for the depletion of haplochromine stocks, and that the haplochromine stock was already affected by fisheries before the establishment of *Lates*, particularly by unregulated fishing or by trawling techniques introduced in the Tanzanian part of the lake. Lake Victoria is now invaded by water hyacinth, and the remaining fish fauna is therefore more and more threatened.

Transport through ballast water is probably one of the most important pathways for alien species invasions in several places, including the North American Great Lakes (Mills *et al.*, 1993). That is the case for the zebra mussel introduced into the Great Lakes, apparently in 1985 or 1986, which spread dramatically throughout the waterways of both Canada and United States, expansion with serious economical and ecological consequences. The recent finding of individual mitten crabs (*Eriocheir sinensis*), a European flounder (*Platichthys flessus*), and the establishment of the alien gastropod *Potamopyrgus antipodarum* in Lake Ontario in 1995 demonstrate that the process of invasion is still going on at a fast rate.

One of the major problems in freshwater species introductions is their irreversibility, at least on the scale of a human's lifetime. Once introduced and established, it is impossible, given current technology, to eradicate a fish, a mollusk, or a plant species from a large natural water body. As a consequence, we are likely to see a continued reduction in native aquatic biodiversity and an increased homogenization of the world's freshwater biotas.

Fisheries Practices

One of the major threats to the unique species flocks of ancient lakes are the fishing practices and particularly overfishing and introduction of new fishing practices. According to Coulter *et al.* (1986), the collective experience in recent years on the African Great Lakes seems to show that large-scale mechanized fishing is incompatible with the continued existence of the highly diverse cichlid communities. Cichlids appear especially vulnerable to unselective fishing because of their particular reproductive characteristics. The structure of endemic cichlid fish communities in the African Great Lakes can change dramatically within a few years when trawlers and other such fishing gear are used. Actually, a number of authors have recorded the effects of over-fishing in Lake Victoria, from the decline of some species to the virtual disappearance of others, and the history of the fishery was briefly reviewed by Witte *et al.* (1992).

It has also been suggested that parks should be developed (Coulter *et al.*, 1986) and that fishing should be rendered impossible in certain areas by placing obstructions on the bottom that would snarl trawls. Lake Malawi National Park will very probably afford protection to widespread species, but no data are at present available to confirm this hypothesis. It is unknown yet whether these reserves can adequately preserve the integrity of populations, but that is probably only possible for stenotopic populations whose distribution coincides with the park area. The size of the reserves, the intensity of fishing in nearby areas, and the possible influence of pollution or introduced alien species should also be taken into account.

Pollution

Pollution can affect aquatic biota through direct mortality at any life stage or by sublethal effects influencing predation, foraging, and reproduction. The eutrophication of Lake Victoria during the past 25 years is quite well documented. Enhanced quantities of nutrients appear to have been entering this lake for many years, both through rivers and from aerosols as a result of human activities in its watersheds. The eutrophication could lead to increased oxygen demand in the lake's deep water and thus decrease the hypolimnetic volume habitable by fish during seasonal stratification. This phenomenon is partly responsible for the threatening or

disappearance of cichlid species belonging to the haplochromine flock.

The release of sulfur and nitrous oxides from the burning of fossil fuels may be transported great distances before being transformed chemically into sulfuric and nitric acids and deposited as rain, snow, or dust. When acid rains occur over areas where waters are poorly buffered, the chemistry and biology of freshwaters can be changed dramatically. Many softwater lakes have been acidified both in North America and Europe, but evidence has accumulated for its occurrence in China, the former Soviet Union, and South America. Monitoring studies indicate a general impoverishment of species numbers in lakes as they become more acidic. Many lakes in the northeastern United States have lost 30% or more of the species in some taxonomic groups. In many northern European countries, acidification strongly modified the fish composition and abundance in lakes.

See also: Endangered Freshwater Invertebrates. Eutrophication and Oligotrophication. Fishes, Biodiversity of. Freshwater Ecosystems. Freshwater Ecosystems, Human Impact on. Invertebrates, Freshwater, Overview. Resource Exploitation, Fisheries

References

Carpenter SR and Kitchell JF (1993) *The Trophic Cascade in Lakes*. Cambridge: Cambridge University Press.

Coulter GW, Allanson BR, Bruton MN, *et al.* (1986) Unique qualities and special problems of the African Great Lakes. *Environmental Biology of Fishes* 17: 161–184.

Crane K, Hecker B, and Golubev V (1991) Hydrothermal vents in Lake Baikal. *Nature* 350: 281.

Dodson SI (1992) Predicting crustacean zooplankton species richness. *Limnology and Oceanography* 37: 848–856.

Dumont H (1994) On the diversity of the Cladocera in the tropics. *Hydrobiologia* 272: 27–38.

Fernando CH (1980) The species and size composition of tropical freshwater zooplankton with special reference to the oriental region (Southeast Asia). *Int. Rev. gesamten Hydrobiol.* 65: 411–426.

Kawanabe H (1996) Asian Great Lakes, especially Lake Biwa. *Environmental Biology of Fishes* 47: 219–234.

Kottelat M and Whitten T (1996) *Freshwater Biodiversity in Asia, with Special Reference to Fish*. World Bank Technical Paper, no. 343.

Lauzanne L (1992) Native species. The *Orestias*. In: Dejoux C and Iltis A (eds.) *Lake Titicaca. A Synthesis of Limnological Knowledge*, pp. 405–419. Dordrecht: Kluwer Academic Publishers.

Lehman JT (1988) Ecological principles affecting community structure and secondary production by zooplankton in marine and freshwater environments. *Limnol. Oceanogr.* 33: 931–945.

Lévêque C (1997) *Biodiversity Dynamics and Conservation: The Freshwater Fish of Tropical Africa*. Cambridge: Cambridge University Press.

Martens K (1997) Speciation in ancient lakes. *Tree* 12(5): 177–182.

Martens K, Goddeeris B, and Coulet G (eds.) (1994) *Speciation in ancient lakes. Advances in Limnology.* Archiv für Hydrobiologie.

May R (1994) Biological diversity: differences between land and sea. *Philosophical Transactions of the Royal Society of London, B* 343: 105–111.

Michel E (1994) Why snails radiate: A review of gastropod evolution in long lived lakes, both recent and fossil. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* 44: 285–317.

Mills EL, Leach JH, Carlton JT, and Secor CL (1993) Exotic species in the Great Lakes: A history of biotic crises and anthropogenic introductions. *Journal of Great Lakes Research* 19: 1–54.

Nagelkerke LAJ, Mina MV, Wudneh T, Sibbing FA, and Osse V (1995) Lake Tana, a unique fish fauna needs protection. *Bioscience* 45(11): 772–775.

Palmer MA, Covich AP, Finlay BJ, *et al.* (1997) Biodiversity and ecosystem processes in freshwater sediments. *Ambio* 26(8): 571–577.

Seehausen O (1996) Lake Victoria Rock Cichlids—Taxonomy, Ecology and Distribution. *Verduyn Cichlids.*

Seehausen O, Jacques JM, and Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277: 1808–1811.

Sideleva VG (1994) Speciation of endemic Cottoidei in Lake Baikal. *Adv. Limnol.* 44: 441–450.

Williams WD and Aladin NV (1991) The Aral sea: Recent limnological changes and their conservation significance. *Aquatic Conservation: Marine and Freshwater Ecosystems* 1: 3–23.

Witte FT, Goldschmidt J, Wanink M, *et al.* (1992) The destruction of an endemic species flock: Quantitative data on the decline of the haplochromine cichlids of Lake Victoria. *Environmental Biology of Fishes* 34: 1–28.

Yu Sherbakov D (1999) Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. *Tree* 14: 92–95.

Landscape Corridors

Ellen I Damschen, University of Wisconsin-Madison, Madison, WI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Connectivity Degree to which the landscape facilitates or impedes movement among habitat patches.
Corridor Linear strip of habitat connecting two otherwise isolated patches of suitable habitat.
Dispersal Movement of organisms, their propagules, or their genes away from their source.
Ecosystem services Resources and processes provided by natural systems that benefit humankind.
Functional traits Characteristics of species that define their ecological roles.
Local extinction When a species ceases to exist in an area that it historically occupied.
Matrix Landscape surrounding suitable habitat patches.
Movement When genes or organisms change location or position.
Patch Areas of suitable habitat that are different from the surrounding habitat (the matrix).
Rescue effect When the likelihood of local extinction in a habitat patch is reduced by emigrants from other habitat patches.

Introduction

Habitat fragmentation is one of the biggest drivers of species extinctions (Wilcove *et al.*, 1998). By severing movement pathways, the loss and fragmentation of habitat has been shown to decrease movement among habitat patches resulting in higher extinction rates (Pimm *et al.*, 1988), loss of gene flow (Lande and Shannon, 1996), decreased species diversity (Leach and Givnish, 1996), and diminished ecosystem services (Daily and Ellison, 2002; Kremen *et al.*, 2002). One proposed conservation solution for ameliorating these negative impacts is to use landscape corridors, or thin strips of habitat, to connect otherwise isolated habitat patches. By reducing local extinctions in habitat patches through increased colonizations (i.e., via "rescue effects" sensu; Brown and Kodric-Brown, 1977), corridors have been theorized to increase gene flow, population sizes, species persistence, and species diversity.

Landscape corridors are also referred to by other names, including "habitat corridors" and "wildlife corridors" (specifically referring to corridors for particular animals). Here, "corridor", "landscape corridor", and "habitat corridor" are used interchangeably to refer to linear habitat linkages that connect otherwise isolated patches of suitable habitat. Corridors have also have been defined in both structural (i.e., the physical presence of a corridor in a landscape) and functional (i.e., serving to promote movement across a landscape) ways, which has created confusion when communicating about corridors (Hess and Fischer, 2001). Here the author discusses how the presence of structural corridors (i.e., linear strips of habitat) affects various ecological functions (e.g., movement, gene flow, population sizes, species diversity). In addition, "landscape connectivity" is a closely related concept in that it is the degree to which the landscape facilitates or impedes movement among resource patches (Taylor *et al.*, 1993; Dunning *et al.*, 1992). Corridors are, therefore, a narrower definition housed within landscape connectivity.

Historical Background

Landscape corridors conceptually emerged from applied island biogeography (Wilson and Willis, 1975) and metapopulation theory (Levins, 1969). This body of theory suggested that larger areas of habitat are better than smaller areas, that habitat patches closer together are better than those farther apart, and that otherwise isolated habitat patches connected with a corridor are better than those without such a connection (**Figure 1**). The intuitive appeal of these proposed designs led to landscape corridors becoming a common feature in the design and implementation of reserves. The Yellowstone to Yukon Conservation Initiative (Y2Y) and the Mesoamerican Biological Corridor from Mexico through Central America are two examples where corridors are being implemented at continental scales. Corridors are also frequently being implemented in regional (e.g., the Talamanca-Caribbean Biological Corridor in Costa Rica) and local landscapes (e.g., Metro Conservation Corridors in the Twin Cities, MN).

Empirical evidence documenting whether and how corridors function, however, has lagged behind implementation. Questions have risen about whether corridors are the most effective conservation strategy in comparison to, for example, preserving additional area alone. And, some have suggested caution about the potential for corridors to have detrimental impacts (e.g., increasing predation risk, spreading disease or fire; Simberloff *et al.*, 1992; Beier and Noss, 1998; Hess, 1994; Orrock and Damschen, 2005). Studies have begun to accumulate over the past decade, allowing for a more informed understanding of whether and how corridors function. Here I do not attempt to conduct an exhaustive literature review (but see Crooks and Sanjayan, 2006; Hilty *et al.*, 2006; Gilbert-Norton *et al.*, 2010; Rosenberg *et al.*, 1997 and the references therein) but instead rely on specific examples to demonstrate key points in the evolution of our understanding of corridor effects.

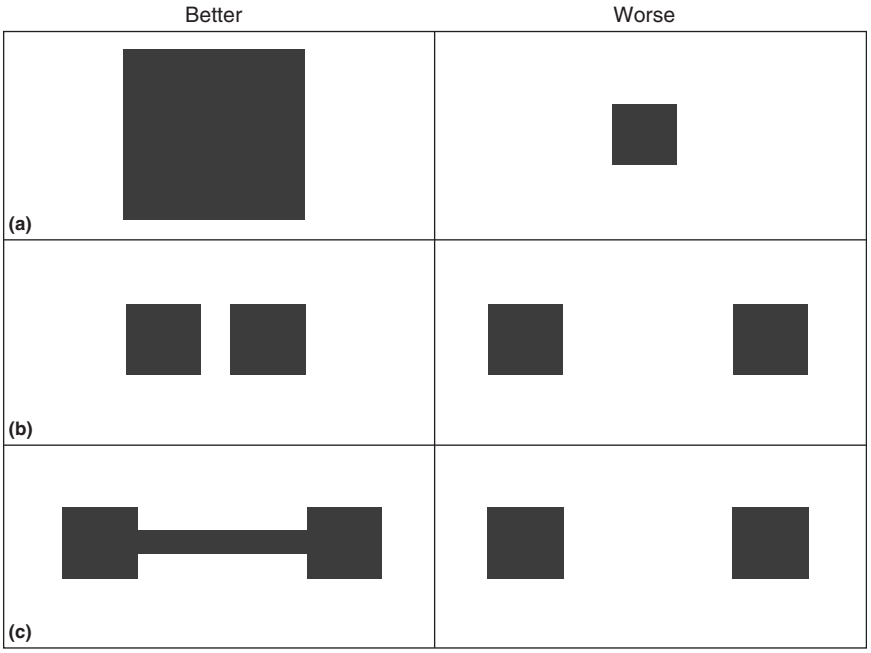


Figure 1 Some of the better and worse reserve designs proposed by Wilson and Willis (1975): (a) larger patches are better than smaller patches, (b) patches closer together are better than those farther apart, and (c) patches connected by corridors are better than isolated patches.

Separating Connectivity Effects from Confounding Factors

Corridor effects can arise from one of three mechanisms: (1) connectivity, whereby corridors increase movement between connected patches; (2) area, because corridors also increase the amount of suitable habitat any increase in movement, population sizes, or diversity could be due to the availability of additional habitat rather than because of a connection per se; and (3) patch shape, because corridors change the shape of a habitat patch by increasing the amount of edge habitat relative to core habitat, responses could be due to patch-shape effects rather than connectivity. Observational studies of existing landscapes are often unable to control for these three separate factors. However, some experiments have explicitly manipulated these factors so that the mechanisms underpinning corridor responses can be deciphered.

Measuring Corridor Responses

The effectiveness of corridors can be evaluated using a variety of metrics. The most commonly measured response is the movement of individuals among habitat patches. This can be done in a variety of ways, including creating mathematical or simulation models, marking and recapturing organisms, conducting real-time follows of organisms as they move, using dyes to trace movement patterns, or determining genetic similarities to infer movement over time. A less commonly measured response is the effect of corridors on population sizes. In large part, this is because tracking populations requires long-term demographic data, which are difficult to obtain at large spatial scales. Similarly, studies of community-level processes (e.g., predator–prey dynamics, species diversity) are also less frequent because they can require the identification of large numbers of organisms often over multiple years. The latter two types of studies (population and community levels) are important to tackle because they are often the ultimate goal of corridor implementation: to prevent species declines and extinctions and to increase biodiversity.

Evidence from Models

Theoretical and simulation models exploring the impact of corridors and connectivity on ecological processes can illuminate new avenues for research, tackle empirically challenging questions, and provide predictive power.

One of the most informative uses of simulation models in corridor research has been to determine how species will respond to various habitat configurations in conjunction with actual conservation decisions. Paul Beier and colleagues have developed such models using geographic information systems (GIS) to illuminate pathways of least resistance or cost to organisms (e.g., Beier *et al.*, 2009; McRae *et al.*, 2008). These models have helped identify habitat connections and prioritize conservation efforts (Beier *et al.*, 2008). For example, the South Coast Missing Linkages project in Southern California utilized this approach for identifying key habitat corridors amid an urban matrix (Beier, 2006).

Simulation models have also been especially useful for evaluating the potential outcomes of negative corridor effects, which may pose special ethical challenges for testing in the field and for considering evolutionary effects that are unable to be tested over short time scales. For example, Hess (1994) determined how corridors might affect the spread of disease

and found that connecting otherwise isolated habitat patches could lead to faster rates of spread (Hess, 1994). By exploring these effects through a model, hypotheses could be initially tested and the study's findings could be used to appropriately guide tests in real landscapes (e.g., Johnson and Haddad, 2011). In another example, Earn *et al.* (2000) determined that landscape connectivity may lead to negative effects by promoting connectivity-mediated synchrony. For evolutionary effects, Orrock (2005) explored how corridors affect rates of adaptive evolution and determined that the effects can be positive or negative, depending on the frequency of disturbance. When disturbance is rare, corridors increase the fixation of beneficial alleles and the loss of harmful alleles (Orrock, 2005). However, when disturbance rates are high, the opposite occurs.

Empirical Evidence

Observational Studies

Observational field studies of corridor effects have provided insight into corridor function by incorporating realistic landscape features such as heterogeneous resources and large spatial scales. Although observational studies provide a very important endpoint on the continuum of understanding corridor effects, they are also limited because they can be confounded by features that co-occur in natural landscapes (e.g., connectivity, area, and patch-shape effects; *see* Separating Connectivity Effects from Confounding Factors) and may lack true replication. Here the author includes what others have referred to as "quasi-experiments" (Rosenberg *et al.*, 1997) or "natural experiments" (Gilbert-Norton *et al.*, 2010) as observational studies since they do not include randomized manipulations or true controls.

One of the earliest studies of corridor effects in real landscapes was reported by MacClintock *et al.* (1977), who demonstrated that bird species richness was higher in a single forest patch that was connected to other forested habitat by a corridor than in another unconnected forest patch. Notably, this early study was conducted at the community level, which remains a challenge for ecologists to study at landscape scales. However, this study contained a single replicate, so although the study's findings were intriguing they were not especially robust.

Several observational studies have made use of existing shelterbelts (wooded habitats planted to provide windbreaks within agricultural landscapes) that connect otherwise isolated woodlots within agricultural settings. Haas (1995) found that shelterbelts in south central North Dakota, USA, promoted the movement of Robins and Thrashers. Harvey (2000) found that shelterbelts in Monteverde, Costa Rica that were connected to forest habitat had higher forest tree-seedling densities and greater numbers of forest tree species. This was likely due to frugivorous birds moving more frequently through the connected shelterbelts because the density and diversity of bird-dispersed tree species were especially high. These studies suggested that corridors (i.e., shelterbelts) promoted movement of birds and bird-dispersed seeds, but because the size and spacing of shelterbelts was not controlled, these studies were unable to determine how corridors functioned.

A quasi-experimental study of corridor effects on diversity was conducted by Schmiegelow *et al.*, (1997) using habitat patches of different sizes that were either connected or unconnected to adjacent existing habitat. However, this study did not control for the area of the corridor. Surprisingly, they found weak effects of corridors and the highest diversity was found in the smallest patch size, which contradicted theoretical predictions. However, area and connectivity effects were confounded in this study because the corridors in their system were wider than the smallest habitat patches. In addition, patches were nonrandomly located so connected patches were always adjacent to riparian areas, resulting in location and connectivity being confounded. This study demonstrated the importance of controlling for the various mechanisms by which corridors can work and for the quality of the surrounding habitat matrix.

Machtans *et al.* (1996) also conducted a quasi-experiment before and after the harvesting of forests that were adjacent to riparian buffer strips. They measured the frequency of bird movements through these landscapes and found that buffer strips enhanced movement of juveniles by promoting connectivity and maintained the movement rates of adults. They also found that corridors were used less frequently when interpatch distances were small, suggesting that corridors may be more important at larger spatial scales where movement is disproportionately enhanced by the presence of corridor.

One of the most robust approaches that can be used in observational studies is to infer landscape connectivity from genetic similarity among individuals because it is a measure that integrates past organism movement across landscapes. For example, Hale *et al.* (2001) measured gene flow in British red squirrels in the UK. They found that recent plantings of conifer forests, which connected habitat fragments in the north of England with the south of Scotland, increased the amount of genetic mixing in squirrel populations. In another example, McRae and Beier (2007) determined genetic similarity for several types of organisms and then correlated this with modeled measures of landscape connectivity. They found an especially effective metric was the amount of landscape "resistance," analogous to electrical resistance, which was strongly correlated with the genetic similarity of the organisms studied. Such tight links between models and data hold promise for predicting the movement of organisms across large landscapes.

Experimental Studies

Experimental studies can overcome some of the inherent challenges of observational systems in that they can control for various factors (e.g., area, connectivity, patch shape) that would otherwise be confounded. They also can use higher levels of replication that can provide robust tests of corridor effects. Controlled and replicated experiments, however, are logistically difficult to conduct at large scales, so most experiments have been conducted using microcosms, mesocosms, or small field plots. Such studies lack the realism of actual landscapes where corridors are being implemented but allow for a more complete understanding of the mechanisms underlying corridor effects. In the following sections,

the author highlights some of the corridor experiments that have been conducted and synthesizes some of the key findings.

Microcosms and Mesocosms

Burkey (1997) provided one of the first experimental tests of corridor function. He used simple communities of bacteria and protozoa in vessels (pits in Plexiglas plates) that were connected with corridors created by removing part of the cylinder wall separating adjacent vessels. The experimental communities contained three trophic levels, and the time to extinction for the top predator was measured in habitats that were different sizes and with different levels of connectivity. Burkey found that fragmented vessels went extinct sooner than unfragmented vessels of the same overall size. However, vessels connected by corridors went extinct significantly sooner than those that were isolated. This study suggested that corridors could have "negative" impacts on some species – in this case, on the top predator in the community – and suggested caution for establishing corridors in fragmented landscapes.

Holyoak (2000) also conducted studies of corridor effects within microcosms (**Figure 2(a)**). He determined how long predator and prey protists persisted in microcosms with different numbers of patches with and without corridors. Holyoak found that metapopulation persistence for the predator could be predicted by the number of patches and corridors, the maximum interpatch distance, and the proportion of patches that provided colonists. However, loops of

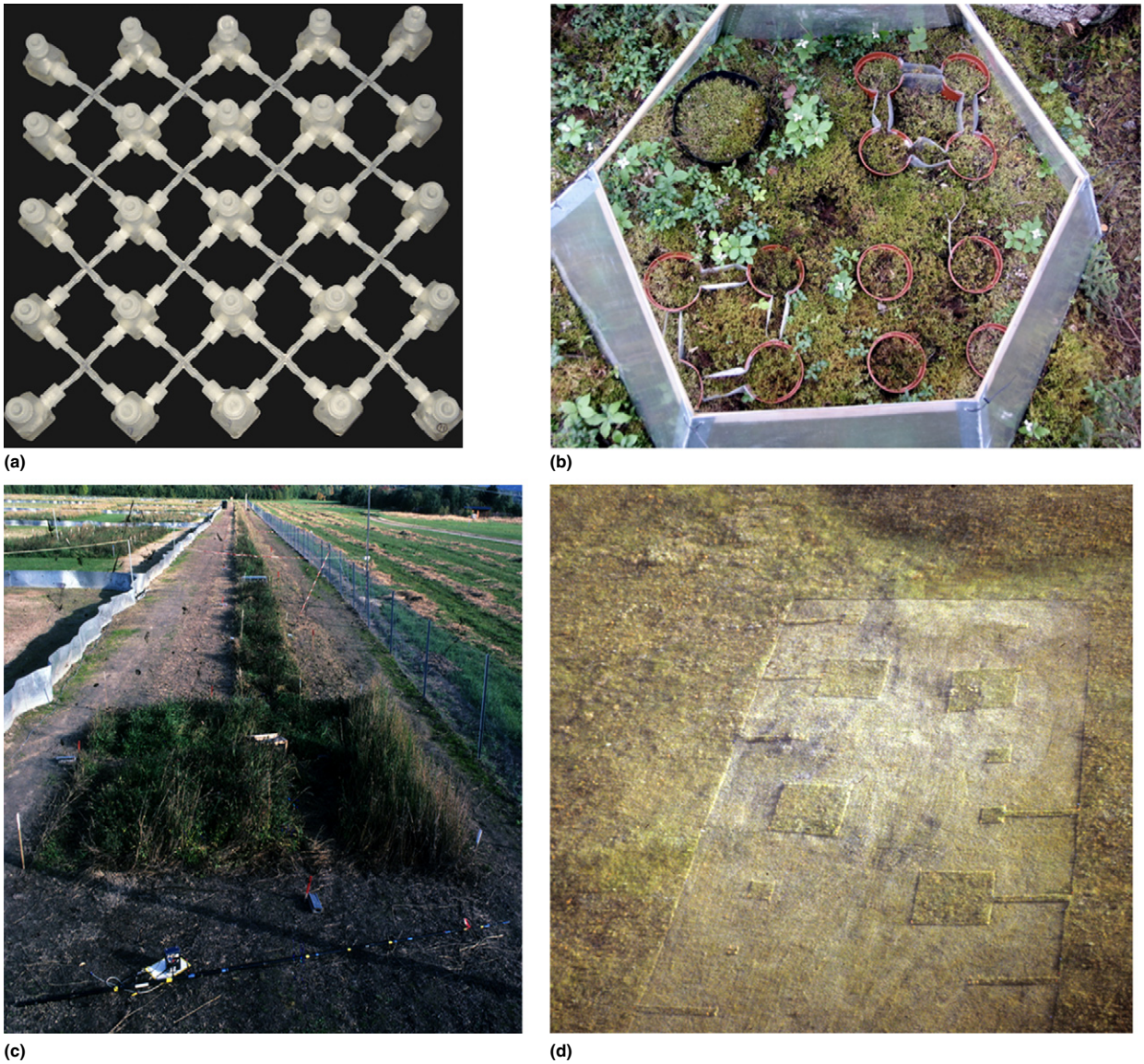


Figure 2 Examples of corridor experiments. (a) A microcosm study by Marcel Holyoak (photo credit: Sharon Collinge). (b) A mescocosm study of arthropods within moss patches by Andrew Gonzalez, Francis Gilbert, and colleagues to test for corridor and climate change effects (photo credit: RolfIms). (c) A plot-based study of corridor presence and width by Harry Andreassen, Rolf Ims, and colleagues (photo credit: Marcel Holyoak). (d) A plot-based study of grassland corridors by Sharon Collinge (photo credit: Andrew Gonzalez). All photos reproduced with permission.

patches generated unexpected results that were not predicted by theory and suggested the need for combining models with experimentation to understand the mechanisms underlying connectivity effects.

A series of mesocosm studies by Andrew Gonzalez, Francis Gilbert, and colleagues have provided robust tests of corridor effects for microarthropod communities within small (~10 cm) patches of moss (Gilbert *et al.*, 1998; Gonzalez and Chaneton, 2002; Gonzalez *et al.*, 1998; Hoyle and Gilbert, 2004). Higher levels of microarthropod diversity and slower extinction rates (contrary to Burkey, 1997) were found in moss patches connected by a corridor (Gilbert *et al.*, 1998; Gonzalez and Chaneton, 2002; Gonzalez *et al.*, 1998). However, a study conducted in the same system several years later found that corridors did not slow the rate of extinction or increase species richness (Hoyle and Gilbert, 2004). The authors concluded that this was likely because of changes in weather and that corridors may be more important for mitigating extinctions during extreme rather than moderate weather conditions. This study revealed that time and environmental conditions might be an important determinant of corridor effectiveness. In new research, Andrew Gonzalez and colleagues are attempting to understand the interaction between connectivity and climate change (i.e., environmental conditions) using microarthropod community mesocosms within open-top warming chambers (**Figure 2(b)**).

Rantalainen *et al.* (2004a, b, 2005) conducted several mesocosm experiments using soil decomposer organisms as a model system. These studies found that corridors increased the species richness of soil fungi in the patches they connected. In addition, corridors sometimes increased the abundance of enchytraeid worms, the species richness of bacteria (dispersed by enchytraeid worms), and the number of microarthropod taxa. One important conclusion from these studies was that corridor effects can depend on particular species interactions and trophic levels.

Plot-based Experiments

One of the first studies to use large field plots to experimentally test for the effects of corridors was conducted by Aars and Ims (1999) who used 20-m × 37.5-m patches of dense meadow vegetation connected by 50-m long corridors within a habitat matrix that was herbicided and fenced (**Figure 2(c)**). This study measured the movement of individuals by tracking levels of genetic heterozygosity among vole demes. Voles within patches connected by a corridor exhibited enhanced interbreeding rates as compared to those within unconnected patches. This has been one of the only experimental studies to examine how corridors affect rates of gene flow and suggests that such connections do enhance the movement of genes. However, the study design did not control for area or patch shape effects, so the underlying mechanism by which corridors operated was not determined.

Coffman *et al.* (2001) conducted another plot-based experiment that tested the impact of corridors on meadow vole populations. Corridors promoted movement and increased survival but did not have any effect on population sizes or recruitment. Populations in better-connected plots, however, had more stable population sizes over time. This study is one of the few that has explicitly measured corridor effects on population sizes and demography, but again the mechanisms responsible for the effects observed were not deciphered.

Another experimental study of corridors was conducted by Sharon Collinge, who used grassland patches within a matrix of mowed vegetation (**Figure 2(d)**) to test for corridor effects on arthropod communities (Collinge, 2000). She found weak effects of corridors on insect diversity and attributed these findings to inadequate contrast between the patch and matrix habitat. In addition, the dispersal ability of the insects within the study system may have been much greater relative to the size and isolation of the patches. Both findings suggested the importance of habitat contrast and the scale of the landscape relative to dispersal ability for determining corridor function.

Landscape Experiments

Experimental tests of corridors at landscape scales offer the promise of providing realistic predictions for real-world landscapes while controlling for factors that are often confounded at such large scales (e.g., area, patch shape, and connectivity). At Savannah River Site, SC, Nick Haddad and the Corridor Research Group created two such experiments to test for the effects of corridors, one of which controlled for area and patch shape. These have been the largest experimental tests of corridors.

In the first experiment, 1.64 ha patches were connected or unconnected by different length corridors (64, 128, 256, and 384 m long). The impacts were subsequently measured on the movement of butterflies, bees, birds, and small mammals (e.g., Haddad *et al.*, 2003; Haddad, 1999a, b; Haddad and Baum, 1999; Danielson and Hubbard, 2000). In general, all 10 of the study species exhibited positive corridor effects (i.e., the percentage increase in movement between connected relative to unconnected patches). Of the 10 species that were examined, five moved significantly more often between connected patches than those that were isolated. For the remaining five species, four appeared to be directed by corridors, but the results were inconclusive because of low sample size. One rodent species (*Sigmodon hispidus*) did not move preferentially between connected patches, likely because the scale of the study was too small for movement of this species across landscapes. In separate studies on the effects of corridor length, movement among patches connected by a corridor was negatively related to corridor length, but corridors increased overall interpatch movement when compared to movement through the surrounding matrix. This suggests that corridors promote movement, which may be especially important ecologically at large spatial scales where infrequent migrants have large impacts on population persistence.

The second experiment was designed to separate the mechanisms by which corridors work – through connectivity, area, or patch-shape effects – by controlling for each factor. Eight experimental landscapes were created, each with a central 100-m × 100-m open habitat patch (**Figure 3**). Four peripheral patches surrounded this center patch with three different patch types. *Connected* patches were 100-m × 100-m patches connected by a 150-m long corridor to the central patch. *Rectangular* patches were 100-m × 137.5-m patches that remained unconnected but controlled for habitat area by adding the area of the corridor to a 100-m × 100-m patch. *Winged* patches were 100-m × 100-m patches with two

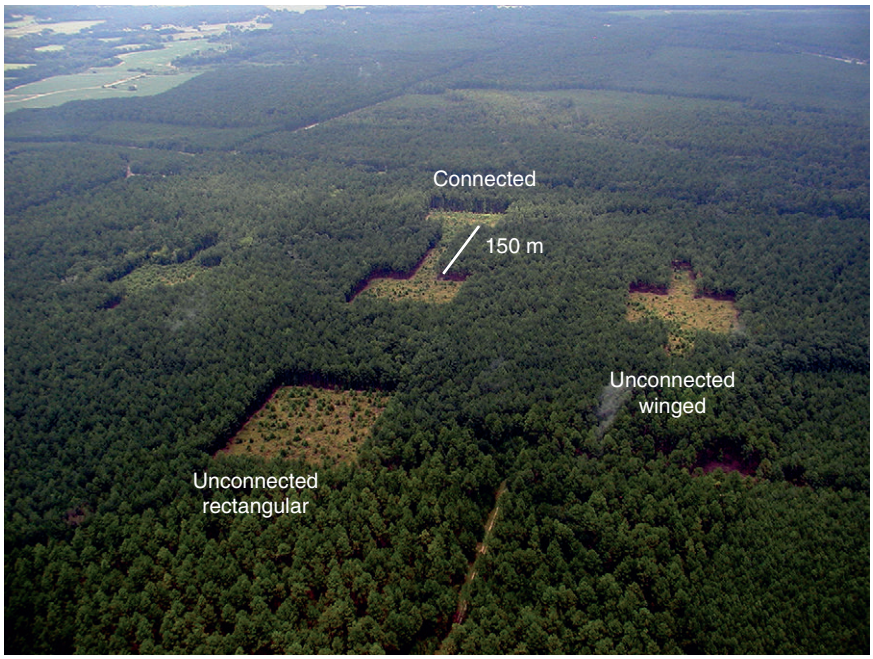


Figure 3 The experimental design used to test the mechanisms by which corridors work at the Savannah River Site, SC, USA. A central 100-m × 100-m patch is surrounded by four peripheral patches that are either connected or unconnected by a corridor. Unconnected patches are of two types: *Rectangular* patches that control for the area of the corridor and *Winged* patches that control for patch shape. Photo credit: Ellen Damschen.

dead-end "wings" projecting from opposite sides of this patch that were each one-half the length of a corridor. These patches had the same area and amount of edge habitat as a connected patch but remained unconnected. By comparing responses in connected, rectangular, and winged patches, it has been possible to determine how corridors are functioning.

Studies from this experiment also found increased movement of birds, plants, and butterflies between patches connected by a corridor (e.g., Levey *et al.*, 2005; Tewksbury *et al.*, 2002; Haddad and Tewksbury, 2005) as well as increased densities for three butterfly species (Haddad and Baum, 1999) and increased gene flow for one butterfly species (*Junonia coenia*) (Wells *et al.*, 2009). In addition, small-scale movement behavior of animals was found to correspond with whether species will use corridors at larger scales (Levey *et al.*, 2005), which suggested that understanding local movement behavior of organisms may allow corridor effects to be predicted.

There was also evidence for negative corridor effects from studies conducted within this experiment. For example, some bird species had decreased survival in connected patches (Weldon, 2006; Weldon and Haddad, 2005). This response was due, however, to patch-shape effects (i.e., because linear landscape features such as corridors necessarily have higher edge-to-area ratios than larger contiguous habitat patches) rather than connectivity effects per se. The birds were preferentially nesting along habitat edges, where they had poor nest success and decreased survival in edgier connected patches. In another example, seeds were more likely to be predated by small mammals in connected patches (Orrock *et al.*, 2003; Orrock and Damschen, 2005). However, seeds consumed by invertebrates exhibited the opposite response, resulting in no difference in total predation (by both types of predators) between connected and unconnected patches (Orrock *et al.*, 2003). This and a subsequent study (Orrock and Damschen, 2005) of two seed species that varied in size demonstrated that corridor effects may differentially affect plant species within the same community. In addition, these studies suggested that corridors may have both positive and negative effects on the same species at different life stages.

Despite these potentially negative impacts of corridors on individual species or life stages, the net effect of corridors on plant community diversity in this system has been positive. Plant species richness was higher in connected versus unconnected patches, and this difference increased over time (Damschen *et al.*, 2006). Various groups of species, however, likely respond uniquely to corridors. For plants, bird-dispersed species responded most strongly and quickly to the connections that corridors provide, whereas wind-dispersed species responded to both patch shape and connectivity effects (Damschen *et al.*, 2008). In another community-level study, there were fewer genera of arthropods and less species turnover in connected patches due to patch-shape effects, whereas evenness was lower in connected patches due to connectivity effects (Orrock *et al.*, 2011). Both of these community level studies of plants and insects suggest that subsets of species within a community respond uniquely to corridors. Understanding species functional traits may provide keys to predicting corridor use.

The Savannah River Site corridor experiments have shown that corridors increase movement for the vast majority of species that have been studied (13 out of 15 species; Haddad *et al.*, 2011). The few cases where corridors have resulted in

negative effects were related to the increased prevalence of edges in connected patches, suggesting that changes in patch shape are more likely responsible for negative impacts than connectivity effects per se (Haddad *et al.*, 2011). Additional studies also found that corridor effects can be predicted by understanding small-scale movements (Haddad, 1999b; Levey *et al.*, 2005) and life-history traits such as the dispersal mode of plants (Damschen *et al.*, 2008).

Conclusions and Future Directions

Synthesizing Current Evidence

Models and observational and experimental studies have demonstrated the utility of corridors in promoting the movement of organisms among the habitat patches they connect. In the majority of cases, corridors have resulted in positive effects for the organisms of interest, including plants, insects, and small mammals (Haddad *et al.*, 2003; Gilbert-Norton *et al.*, 2010). In a recent meta-analytic review of 78 experiments from 35 corridor studies, Gilbert-Norton *et al.* (2010) determined that corridors increased interpatch movement by ~50% when compared to unconnected patches. The cases where corridors have caused neutral or negative effects demonstrate the importance of understanding organism life histories, how the scale of organism movement relates to the landscape, and the underlying mechanisms responsible for corridor use (e.g., connectivity or patch shape effects).

In a small number of cases, corridors have resulted in negative impacts for the species of interest (e.g., Orrock and Damschen, 2005; Hess, 1994; Johnson and Haddad, 2011; Haddad *et al.*, 2011). For example, corridors increased disease transmission in the model by Hess (1994), which was corroborated by a later field study by Johnson and Haddad (2011). Importantly, however, the model was unable to decipher the mechanisms underlying the response to corridors, but the field test was able to do so by experimentally controlling for connectivity, area, and patch-shape effects. The field experiment determined that disease spread was driven by distance from habitat edges (i.e., patch-shape effects due to high edge-to-area ratios in connected patches) rather than connectivity effects per se. In fact, the majority of "negative" corridor effects have been driven by patch-shape and edge effects (Haddad *et al.*, 2011).

For some species, the presence of a corridor has not affected movement or been less effective in promoting movement than the nonhabitat matrix (Haddad *et al.*, 2003). Often, this has been due to the life history traits of the organism. For example, species that are habitat generalists may not respond because quality of the corridor habitat is not different from the surrounding matrix, and these organisms can move through both habitat types equally well (Haddad, 1999b). Or, if the species is migratory and the migratory pathways are in a particular direction, then corridors may not be used unless they are also in the same direction as the migration (Haddad, 1999b). Other reasons for such responses include if the suitability of habitat types is not understood or has been misclassified or if the scale of the study is not appropriate for the study organism (Gilbert-Norton *et al.*, 2010; Haddad, 1999b). These responses demonstrate the importance of understanding the life history of the target species for which a corridor may be designed in relation to landscape features such as matrix quality (Prugh *et al.*, 2008).

Future Directions

The last decade has provided great insight into corridor function, but much yet remains to be understood. In particular, most studies have focused on organism movement alone, whereas the effects of corridors on population size, genetic diversity, and species diversity have received much less attention (Gilbert-Norton *et al.*, 2010). In addition, most studies have been conducted over relatively short time scales, preventing an understanding of long-term population and community dynamics. Future studies are needed to determine whether increased movement results in reduced population extinction rates and increased species diversity over the long term.

Other factors that need to be explored include the role of corridor quality and the habitat matrix in promoting or hindering movement (Prugh *et al.*, 2008). One study found that low-quality corridors promoted the movement of two butterfly species even though they did not establish resident populations and preferred core habitat (Haddad and Tewksbury, 2005). However, for other species that must move through corridors across multiple generations such as plants, habitat quality may be more important. A key conservation question is whether land acquired for corridors is of appropriate quality and contrast to the matrix to achieve management goals.

Corridor studies have also almost exclusively been conducted in terrestrial settings. There are only a handful of examples from marine and freshwater ecosystems. Darcy and Eggleston (2005) conducted a study to determine how seagrass corridors affected the movement of benthic organisms. They found that corridors facilitated the movement of slow-moving polychaete worms, but not for faster-moving species such as grass shrimp and bay scallops. For the faster-moving species, water currents seemed to affect dispersal more than the presence of a habitat corridor. These results from an aquatic system corroborate those from terrestrial settings and similarly suggest that understanding the life history of organisms and how their movement is affected by landscape features is critical for predicting corridor function. In particular, the scale of the corridor relative to the scale at which organisms make movement decisions is a key factor for determining if corridors will promote dispersal. Future studies of corridor effects across ecosystems may help elucidate generalities such as these.

A critical conservation question that has not been adequately addressed is how corridor width affects species. Early studies by La Polla and Barrett (1993) and Andreassen *et al.* (1996) tackled this question by monitoring the movement of voles in a plot-based experiments. La Polla and Barrett (1993) found that corridor presence was more important than width for the movement of male meadow voles within their home

range. Although Andreassen *et al.* (1996) found that a corridor of intermediate width (1 m) promoted movement of root voles better than narrower (0.4 m) or wider (3 m) corridors. In another study of corridor width and quality on Deermice, Ruefenacht and Knight (1995) found that corridor width had no effect but that tree densities were the major driver of movement. Since these studies, there have been no other experimental tests of how corridor width affects movement (let alone population sizes or community diversity) despite the fact that land managers continue to search for answers to this question.

A related question that needs addressing is how the relative scales of organism movement and landscapes might help predict corridor responses. Gilbert-Norton *et al.* (2010) found that corridors worked equally well for all taxa they examined except birds (which still had a positive effect size), which they hypothesized was due to the fact that experiments were more difficult to appropriately scale for birds since they can easily travel large distances over matrix habitat. Further work is needed to explore how scales of organism movement relate to scales of landscape heterogeneity.

Finally, corridors have often been implemented to minimize potential risks of climate change by facilitating the movement of species to newly created suitable habitat from locations where habitat has become unsuitable. However, there has been virtually no research dedicated to this topic. Andrew Gonzalez and colleagues are conducting a new study that addresses whether or not corridors facilitate species persistence in habitats with varying climate (**Figure 2(b)**). Their experimental moss arthropod mesocosm study system is particularly well suited for tackling this challenging question in that manipulating both connectivity and climate at landscape scales would not be possible. By using mesocosms that include connected and unconnected patches embedded within a series of warming open-top chambers (and controls), they will be able to provide a direct test of how corridors function under climatic warming scenarios.

Appendix

List of Courses

1. Community Ecology
2. Conservation Biology
3. Ecology
4. Landscape Ecology
5. Population Ecology

See also: Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. Climate Change and Ecology, Synergism of. Climate Change and Extinctions. Climate Change and Wild Species. Dispersal Biogeography. Extinction, Causes of. Habitat Loss and Fragmentation. Implications of Urbanization for Conservation and Biodiversity Protection. Island Biogeography. Landscape Ecology and Population Dynamics. Metapopulations. Priority Setting for Biodiversity and Ecosystem Services. Valuing Ecosystem Services

References

Aars J and Ims RA (1999) The effect of habitat corridors on rates of transfer and interbreeding between vole demes. *Ecology* 80: 1648–1655.

Andreassen HP, Halle S, and Ims RA (1996) Optimal width of movement corridors for root voles: Not too narrow and not too wide. *Journal of Applied Ecology* 33: 63–70.

Beier P (2006) South coast missing linkages: Restoring connectivity to wildlands in the largest metropolitan area in the USA. In: Crooks KR and Sanjayan M (eds.) *Connectivity Conservation*. Cambridge, England: Cambridge University Press.

Beier P and Noss RF (1998) Do habitat corridors really provide connectivity? *Conservation Biology* 12: 1241–1252.

Beier P, Majka DR, and Newell SL (2009) Uncertainty analysis of least-cost modeling for designing wildlife linkages. *Ecological Applications* 19: 2067–2077.

Beier P, Majka DR, and Spencer WD (2008) Forks in the road: Choices in procedures for designing wildland linkages. *Conservation Biology* 22: 836–851.

Brown JH and Kodric-Brown A (1977) Turnover rates in insular biogeography: Effect of immigration on extinction. *Ecology* 58: 445–449.

Burkey TV (1997) Metapopulation extinction in fragmented landscapes: Using bacteria and protozoa communities as model ecosystems. *American Naturalist* 150: 568–591.

Coffman CJ, Nichols JD, and Pollock KH (2001) Population dynamics of *Microtus pennsylvanicus* in corridor-linked patches. *Oikos* 93: 3–21.

Collinge SK (2000) Effects of grassland fragmentation on insect species loss, colonization, and movement patterns. *Ecology* 81: 2211–2226.

Crooks KR and Sanjayan M (eds.) (2006) *Connectivity Conservation*. Cambridge, United Kingdom: Cambridge University Press.

Daily G and Ellison K (2002) *The New Economy of Nature*. Washington, DC: Island Press.

Damschen EI, Brudvig LA, Haddad NM, Orrock JL, Tewksbury JJ, and Levey DJ (2008) The movement ecology and dynamics of plant communities in fragmented landscapes. *Proceedings of the National Academy of Sciences of the United States of America* 105: 19078–19083.

Damschen EI, Haddad NM, Orrock JL, Tewksbury JJ, and Levey DJ (2006) Corridors increase plant species richness at large scales. *Science* 313: 1284–1286.

Danielson BJ and Hubbard MW (2000) The influence of corridors on the movement behavior of individual *Peromyscus polionotus* in experimental landscapes. *Landscape Ecology* 15: 323–331.

Darcy MC and Eggleston DB (2005) Do habitat corridors influence animal dispersal and colonization in estuarine systems? *Landscape Ecology* 20: 841–855.

Dunning JB, Danielson BJ, and Pulliam HR (1992) Ecological processes that affect populations in complex landscapes. *Oikos* 65: 169–175.

Earn DJD, Levin SA, and Rohani P (2000) Coherence and conservation. *Science* 290: 1360–1364.

Gilbert F, Gonzalez A, and Evans-Freke I (1998) Corridors maintain species richness in the fragmented landscapes of a microecosystem. *Proceedings of the Royal Society of London Series B – Biological Sciences* 265: 577–582.

Gilbert-Norton L, Wilson R, Stevens JR, and Beard KH (2010) A meta-analytic review of corridor effectiveness. *Conservation Biology* 24: 60–668.

Gonzalez A and Chaneton EJ (2002) Heterotroph species extinction, abundance and biomass dynamics in an experimentally fragmented microecosystem. *Journal of Animal Ecology* 71: 594–602.

Gonzalez A, Lawton JH, Gilbert FS, Blackburn TM, and Evans-Freke I (1998) Metapopulation dynamics, abundance, and distribution in a microecosystem. *Science* 281: 2045–2047.

Haas CA (1995) Dispersal and use of corridors by birds in wooded patches on an agricultural landscape. *Conservation Biology* 9: 845–854.

Haddad NM (1999a) Corridor and distance effects on interpatch movements: A landscape experiment with butterflies. *Ecological Applications* 9: 612–622.

Haddad NM (1999b) Corridor use predicted from behaviors at habitat boundaries. *The American Naturalist* 153: 215–227.

Haddad NM and Baum KA (1999) An experimental test of corridor effects on butterfly densities. *Ecological Applications* 9: 623–633.

Haddad NM, Bowne DR, Cunningham A, *et al.* (2003) Corridor use by diverse taxa. *Ecology* 84: 609–615.

Haddad NM, Hudgens B, Damschen EI, *et al.* (2011) Assessing positive and negative ecological effects of corridors. In: Liu J, Hull V, Morzillo AT, and Wiens JA (eds.) *Sources, Sinks, and Sustainability*. Cambridge, UK: Cambridge University Press.

Haddad NM and Tewksbury JJ (2005) Low-quality habitat corridors as movement conduits for two butterfly species. *Ecological Applications* 15: 250–257.
Hale ML, Lurz PWW, Shirley MDF, Rushton S, Fuller RM, and Wolff K (2001) Impact of landscape management on the genetic structure of red squirrel populations. *Science* 293: 2246–2248.
Harvey CA (2000) Colonization of agricultural windbreaks by forest trees: Effects of connectivity and remnant trees. *Ecological Applications* 10: 1762–1773.
Hess GR (1994) Conservation corridors and contagious-disease – a cautionary note. *Conservation Biology* 8: 256–262.
Hess GR and Fischer RA (2001) Communicating clearly about conservation corridors. *Landscape and Urban Planning* 55: 195–208.
Hilty J, Lidicker Jr. WZ, and Merenlender A (2006) *Corridor Ecology: The Science and Practice of Linking Landscapes for Biodiversity Conservation*. Washington, United States: Island Press.
Holyoak M (2000) Habitat patch arrangement and metapopulation persistence of predators and prey. *American Naturalist* 156: 378–389.
Hoyle M and Gilbert F (2004) Species richness of moss landscapes unaffected by short-term fragmentation. *Oikos* 105: 359–367.
Johnson BL and Haddad NM (2011) Edge effects, not connectivity, determine the incidence and development of a foliar fungal plant disease. *Ecology* 92: 1551–1558.
Kremen C, Williams NM, and Thorp RW (2002) Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences of the United States of America* 99: 16812–16816.
Lande R and Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* 50: 434–437.
La Polla VN and Barrett GW (1993) Effects of corridor width and presence on the population-dynamics of the Meadow Vole (*Microtus pennsylvanicus*). *Landscape Ecology* 8: 25–37.
Leach MK and Givnish TJ (1996) Ecological determinants of species loss in remnant prairies. *Science* 273: 1555–1558.
Levey DJ, Bolker BM, Tewksbury JJ, Sargent S, and Haddad NM (2005) Bird foraging behaviors scale up to predict corridor effects on seed dispersal. *Science* 309: 146–148.
Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* 15: 237–240.
MacClintock L, Whitcomb RF, and Whitcomb BL (1977) Evidence for the value of corridors and minimization of isolation in preservation of biotic diversity. *American Birds* 31: 6–13.
Machtans CS, Villard MA, and Hannon SJ (1996) Use of riparian buffer strips as movement corridors by forest birds. *Conservation Biology* 10: 1366–1379.
McRae BH and Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America* 104: 19885–19890.
McRae BH, Dickson BG, Keitt TH, and Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* 89: 2712–2724.
Orrock JL (2005) Conservation corridors affect the fixation of novel alleles. *Conservation Genetics* 6: 623–630.
Orrock JL, Curler GR, Danielson BJ, and Coyle DR (2011) Large-scale experimental landscapes reveal distinctive effects of patch shape and connectivity on arthropod communities. *Landscape Ecology* 26: 1361–1372.
Orrock JL and Damschen EI (2005) Corridors cause differential seed predation. *Ecological Applications* 15: 793–798.
Orrock JL, Danielson BJ, Burns MJ, and Levey DJ (2003) Spatial ecology of predator–prey interactions: corridors and patch shape influence seed predation. *Ecology* 84: 2589–2599.
Pimm SL, Jones HL, and Diamond J (1988) On the risk of extinction. *American Naturalist* 132: 757–785.
Prugh LR, Hodges KE, Sinclair ARE, and Brashares JS (2008) Effect of habitat area and isolation on fragmented animal populations. *Proceedings of the National Academy of Sciences of the United States of America* 105: 20770–20775.
Rantalainen ML, Fritze H, Haimi J, Kiikkila O, Pennanen T, and Setala H (2004a) Do enchytraeid worms and habitat corridors facilitate the colonisation of habitat patches by soil microbes? *Biology and Fertility of Soils* 39: 200–208.
Rantalainen ML, Fritze H, Haimi J, Pennanen T, and Setala H (2005) Colonisation of newly established habitats by soil decomposer organisms: The effect of habitat corridors in relation to colonisation distance and habitat size. *Applied Soil Ecology* 28: 67–77.
Rantalainen ML, Haimi J, and Setala H (2004b) Testing the usefulness of habitat corridors in mitigating the negative effects of fragmentation: the soil faunal community as a model system. *Applied Soil Ecology* 25: 267–274.
Rosenberg DK, Noon BR, and Meslow EC (1997) Biological corridors: Form, function, efficacy. *Bioscience* 47: 677–687.
Ruefenacht B and Knight RL (1995) Influences of corridor continuity and width on survival and movement of deermice peromyscus-maniculatus. *Biological Conservation* 71: 269–274.
Schmiegelow FKA, Machtans CS, and Hannon SJ (1997) Are boreal birds resilient to forest fragmentation? An experimental study of short-term community responses. *Ecology* 78: 1914–1932.
Simberloff D, Farr JA, Cox J, and Mehlman DW (1992) Movement corridors: Conservation bargains or poor investments? *Conservation Biology* 6: 493–504.
Taylor PD, Fahrig L, Henein K, and Merriam G (1993) Connectivity is a vital element of landscape structure. *Oikos* 68: 571–573.
Tewksbury JJ, Levey DJ, Haddad NM, *et al.* (2002) Corridors affect plants, animals, and their interactions in fragmented landscapes. *Proceedings of the National Academy of Sciences USA* 99: 12923–12926.
Weldon AJ (2006) How corridors reduce Indigo Bunting nest success. *Conservation Biology* 20: 1300–1305.
Weldon AJ and Haddad NM (2005) The effects of patch shape on indigo buntings: Evidence for an ecological trap. *Ecology* 86: 1422–1431.
Wells CN, Williams RS, Walker GL, and Haddad NM (2009) Effects of corridors on genetics of a butterfly in a landscape experiment. *Southeastern Naturalist* 8: 709–722.
Wilcove DS, Rothstein D, Dubow J, Phillips A, and Losos E (1998) Quantifying treats to imperiled species in the United States. *Bioscience* 48: 607–615.
Wilson EO and Willis EO (1975) Applied biogeography. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*. Cambridge, Massachusetts: The Belknap Press.

Landscape Diversity

Debra PC Peters, United States Department of Agriculture, Agricultural Research Service, Las Cruces, NM, USA
Sarah C Goslee, USDA-ARS Pasture Systems and Watershed Management Research Unit, University Park, PA, USA
Scott L Collins, University of New Mexico, Albuquerque, NM, USA
James R Gosz, University of Idaho, Moscow, ID, USA

Published by Elsevier Inc.

This article is a revision of the previous edition article by Debra P. C. Peters, Sarah C. Goslee, volume 3, pp 645–658, © 2001, Elsevier Inc.

Glossary

Alpha landscape diversity Number and dominance of patch types within a landscape.
Corridor Strip of land or water that differs from the adjacent landscape on both sides.
Gamma landscape diversity Total number of patch types contained within a geographic region.
Landscape Spatially heterogeneous area composed of a mosaic of interacting components (patches, corridors, and area of matrix).
Matrix Background landform, habitat, or ecosystem in a landscape, characterized by extensive area, high connectivity, and major control over landscape dynamics.
Patch Area that is relatively homogeneous with respect to the characteristics being examined, and that differs from its surroundings.
Region Large geographic area that contains more than one landscape.

Introduction

While biodiversity is usually considered at the species level, maintenance of biodiversity requires management at higher levels of organization, particularly at the landscape scale. It is difficult to manage for each threatened species individually. Alternatively, management can focus on the ecosystems that contain these species and on the landscapes in which ecosystems are found. The relatively new discipline of landscape ecology provides an insight into both landscape diversity and species diversity, and suggests a theoretical and practical basis for conservation planning.

There are three basic characteristics of landscapes that affect their diversity: structure, function, and dynamics. Structure is the most well-understood element of landscapes. It is also the most obvious – nearly any aerial view will show a mixture of different landforms, habitats, or vegetation types. The patch is the basic unit of landscape structure. The characteristics of patches and the spatial relationships among patches are important components of landscapes. The distributions of energy, materials, and species among patches differing in size, shape, abundance, and configuration are particularly important to patterns in diversity at the landscape scale. The other two elements of landscapes go beyond a description of spatial heterogeneity. Function is concerned with interactions among the spatial elements of a landscape, including flows of energy, materials, and species among patches. Landscape dynamics includes characteristics of both structure and function to examine changes in pattern and process over time. The conservation and management of biodiversity require an understanding of all three elements, including the effects of human activities on the system. This article discusses each element in turn, and also considers the underlying determinants of landscape structure, including environmental heterogeneity and disturbance patterns. The authors then discuss classical and current issues in biodiversity management and conclude with a case study of landscape diversity at the Sevilleta National Wildlife Refuge Long-Term Ecological Research site in central New Mexico, US.

It is essential to keep the concept of scale in mind when considering landscape diversity. Spatial scale has two elements: grain and extent. Grain is the minimum resolution sampled, usually the cell size or the quadrat size for ecological studies. The extent is the domain of the study, which is typically the size of the study area. Ecological processes often have characteristic spatial and temporal scales. This means that the grain and extent of sampling in both space and time may strongly affect the results of a study. For example, as quadrat size (grain) increases, species richness may increase, and yet the diversity of patch types within a landscape may decrease since fewer large quadrats can be found within a given area (**Figure 1(a)**). As landscape size (extent) increases, more species and more patches of a constant size may be found that would increase both species and landscape diversity (**Figure 1(b)**).

Description of Landscape Structure

Landscape structure can be most easily described at two hierarchical spatial levels, both of which are relevant to landscape diversity as well as to species diversity. At the lower level, the focus is on the attributes of individual patches, particularly size and shape. Description at the higher spatial level is concerned with the composition and pattern of the entire landscape and its mosaic of patches. The ability to quantify landscape structure at both levels allows the comparison of different landscapes. More importantly, interactions between landscape structure and function have implications for both species and landscape diversity.

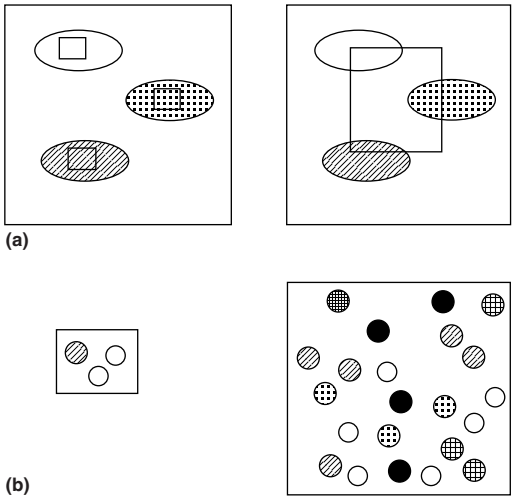


Figure 1 Landscape grain and extent: (a) as the grain or the quadrat size of a landscape increases, yet the extent remains the same, the diversity of patch types decreases since fewer quadrats can be found in the same area; (b) as landscape extent or size increases, more species and more patch types of the same size can be found that result in higher landscape diversity.

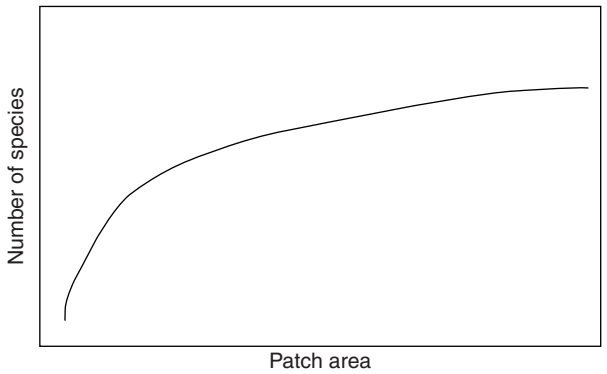


Figure 2 Species–area curve showing the increase in species number with increasing patch size tending toward a regional limit.

Patch Description

A patch is a relatively homogeneous nonlinear area that differs from its surroundings. The definition and identification of individual patches and their boundaries are important steps in characterizing the structure of a landscape. In some systems, boundaries may be easily identified, such as between patches of agricultural field and adjacent woodland in human-dominated systems. In many cases, however, the boundary is not so clear, and patches are more difficult to delineate. Most methods of patch identification combine qualitative and quantitative approaches. A subjective determination of how different two areas must be in order for them to be considered separate patches is often needed. A number of quantitative techniques have been developed to group similar cells into homogeneous patches or to identify repeating patterns across a landscape (Turner and Gardner, 1990). Approaches such as blocking techniques, spectral analysis, and nearest-neighbor analysis are commonly used. Other techniques rely on the detection of edges or boundaries rather than identifying patches directly. These methods include moving window analysis and image analysis to characterize landscapes with sharp transitions (Cornelius and Reynolds, 1991).

Patch identification provides an excellent example of the importance of the spatial scale of the observer. From inside a forest, clumps of trees and grass-dominated openings appear to be separate patches with different vegetation and resource availability. From an aerial view, the entire forest appears to be a single patch. This illustrates the importance of the selection of spatial scale based on study objectives before the determination of patches and their edges.

Once the patches in a landscape have been identified, there are many ways to describe and quantify them (Riitters *et al.*, 1995). Only patch size and shape will be discussed here, since the relevance of these two attributes for species diversity is the most well understood. The relationship between patch size and species richness goes beyond the familiar species–area curve (**Figure 2**). Although the number of species present in a patch tends to increase with patch size up to a certain limit, the kinds of species found also tend to vary with size. Two general types of species can be distinguished. Interior species are found primarily in the interior of large patches. These species often have very specific habitat requirements and are relatively rare. Migratory songbirds that are particularly sensitive to patch size and adversely affected by habitat fragmentation are interior species. In contrast, edge species are found near the edges of large patches and throughout small patches that consist mostly of edge habitats. Edge species are commonly occurring generalists that can use various habitat types, and are often introduced species. Because small patches consist mostly of edge with little interior area, they often have the highest species densities, but contain few or no rare species. Large patches, however, are mostly interior area with lower species densities per unit area, but they contain more rare species and a higher total number of species.

In an important study of tropical deforestation in the Amazon rain forest, species in patches of various sizes were compared to evaluate the importance of patch size to species number (Lovejoy *et al.*, 1984, 1986). Large patches were richest in species and small patches were found to contain only edge conditions. Patch size had important effects on different species, including trees, insects, birds, and mammals, which were noticeable in a short time. This study is one of the few in which patch size was experimentally manipulated to allow comparison with pretreatment conditions as well as control patches.

A simple measure of patch shape is the perimeter: area ratio. This measure is often standardized so that the most compact possible form, either a square or a circle, is equal to 1. More complex shapes have increasingly higher numbers. Another common index of shape complexity is the fractal dimension, which is also derived from the perimeter and area of a patch. The fractal dimension of a patch is between 1 and 2; a simple shape will have a lower fractal dimension than a more complex shape. **Figure 3** illustrates the amount of interior area available in patches of different shapes. Both patches have an area of 25, but the perimeter of shape a is 20, while the perimeter of shape b is 32. Using a scaled perimeter: area ratio, a has a value of 1 and b has a value of 1.6. Assuming that

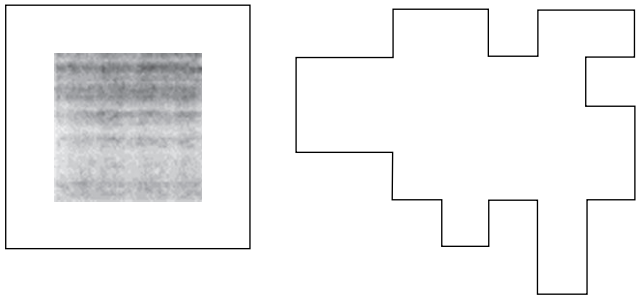


Figure 3 (a) A simple shape and (b) a more complex shape of equal area. The interior area is shaded.

interior area is at least 1 unit from any patch edge, a has an interior area of 9, but b has an interior area of only 2. Thus, a is more compact and less convoluted than b, where more of the area is closer to its edge and can interact with the area surrounding the patch. This suggests that the overall flow of species and resources between *b* and its surroundings is higher than that between *a* and its surroundings. It is also expected that *a* would have higher richness of interior species than *b*, which would have higher richness of edge species.

Landscape Description

At the landscape level, there are two basic components of structure: composition and pattern. Composition refers to the parts (i.e., patch types) that make up the landscape and pattern refers to how these patches are arranged. Although these two components are conceptually different, in practice, they are often related. For example, the pattern of agricultural fields on a landscape is likely to be different from the pattern of undisturbed woodland.

Landscape composition can be measured in ways analogous to measurements of species composition (Romme, 1982). The most straightforward approach is landscape richness or the number of different patch types in a landscape. Another approach includes the relative abundance or dominance of different patch types along with richness. These landscape indices were derived from information theory and are closely related to species diversity measures, such as the Shannon–Wiener and Simpson indices, which are used to describe alpha species diversity (Turner, 1989; Huston, 1994). Using one of these indices, a landscape containing many small patches of different types would have a higher diversity value than a landscape consisting of one large patch and several smaller patches, even if the total number of patches is the same for both landscapes. Landscape measures of richness and evenness were used in a study conducted in different patch types in Yellowstone National Park (Romme, 1982). Changes in landscape diversity through time were related to fire frequency and were hypothesized to have important effects on species diversity as well as wildlife habitat (Romme and Knight, 1982).

Measurements of landscape diversity are analogous to common measurements of species diversity (Whittaker, 1960, 1972). Alpha species diversity is a measurement of species richness (number) and evenness (dominance or distribution) within a patch. Similarly, alpha landscape diversity is a measure of the number of patch types in a region (O'Neill *et al.*, 1988). Large-scale species diversity is called gamma diversity. Gamma landscape diversity of ecosystems is sometimes called ecodiversity (Rowe, 1992; Lapin and Barnes, 1995). The third form of species diversity, beta diversity, describes species turnover along a gradient. Beta diversity has no analog at the landscape level, but is sometimes estimated as gamma diversity/alpha diversity, which yields an average regional beta diversity.

Because different patch types provide different habitats and species compositions, one might expect that the total number of species in a landscape would increase as landscape richness increases. This idea was supported by a study that compared plant species richness in Rhode Island Audubon refuges varying in terrain and soil properties (geomorphological measures) (Nichols *et al.*, 1998). In a related study of one landscape, the diversity of trees and shrubs was higher on plots with the greatest geomorphological heterogeneity, indicating an important link between landscape diversity and species diversity (Burnett *et al.*, 1998). Although this simple relationship between landscape and species diversity is generally true, the interactions between landscape composition and species diversity are more complex, in part because of species preferences to edge or interior types of habitats. The species found in a diverse landscape with many small patches are mostly edge species. Interior species are found only in landscapes with large patches, even though these landscapes have a lower diversity. The total number of interior species increases with the number of large patches on a landscape, similar to the species–area relationship in **Figure 2**. Thus, the type of species that increases with increasing landscape diversity depends on the change in the size and configuration of patches within the landscape.

Landscape pattern, or the spatial arrangement of patches, can be measured in a number of ways, some of which are extensions of the patch-level metrics already discussed. These measures focus on patch abundance without regard to location in the landscape. The distribution of patch sizes can be determined within a landscape and used as information in the management of habitat patches for species that are sensitive to patch size or spatial arrangement, such as the spotted owl in the Pacific Northwest. The effects of forest clear-cutting on changes in patch structure and implications for interior and edge species provide another example of the importance of these measures. Shape complexities and boundaries can also

be scaled up from the patch to the landscape level using the fractal dimension and perimeter: area ratio (Milne, 1988).

The second type of measurement for landscape pattern explicitly considers the location of patches relative to each other and includes patch abundance as well. Dispersion indicates the tendency of patches of one type to be distributed either uniformly, randomly, or aggregated. Contagion describes the tendency of patches of two different types to be near each other. Connectedness can be quantified using nearest-neighbor probabilities that reflect the degree of fragmentation in the landscape. All three of these indices have implications for the flow of species and resources between patches of the same and different types, and thus have important effects on species diversity. The effects of patch characteristics and landscape context on species richness can even be seen in heavily managed agricultural communities (Goslee and Sanderson, 2010).

Two additional structural elements other than patches may be recognized in many landscapes. The second element is the matrix or the background land-form, habitat, or ecosystem in a landscape. The matrix is characterized by extensive cover, high connectivity, and major control over landscape dynamics. Forest patches contained within a matrix of subdivisions are functionally very different from forest patches surrounded by agricultural land. Corridors, strips that differ from the adjacent landscape on both sides, are the third element of landscapes. Corridors are usually linear and always highly connected; stream networks and roadways are common examples. Corridors may also connect larger patches of a similar type, such as a stream flowing between two lakes.

The patch–matrix–corridor model of landscape structure is illustrated in **Figure 4**. Corridors may be particularly important for preserving species diversity by allowing movements of species across diverse landscapes. Corridors can also adversely affect species diversity by allowing nonnative or exotic species to invade and reduce the number of native species in an area. An example is the extensive spread of cheatgrass, an annual introduced to North America in shipments of grain from Asia and Europe in the 1880s (Mack, 1981). Movement of cheatgrass seed along railroad and cattle trail corridors in the early 1900s spread this grass throughout much of the northwestern US, resulting in changes in species composition and dominance as well as losses of diversity.

Controls on Landscape Diversity

Heterogeneity or diversity of landscape structure arises from a number of factors. Patches can be produced through biotic or abiotic causes, including natural- or human-caused disturbance, fragmentation, regeneration, and persistent differences in environmental resources. Once a patch is formed, environmental conditions or interactions among organisms may change through time, leading to successional dynamics on the patch. A landscape consisting of patches in various successional stages is called a "shifting mosaic" (Bormann and Likens, 1979). The spatial pattern of patch formation and the changes within patches are collectively called "patch dynamics" (Pickett and White, 1985). The patch dynamic mosaic is part of the broader landscape transformation that includes changes in corridors and the matrix as well as in the dynamics of species and ecosystem processes. These dynamics are discussed in Section Landscape Dynamics.

Biotic causes of patch generation include the local dispersal of seeds into a landscape, such as by an invasive weed, and the spatial segregation of populations or communities as a result of competition. Spatial structure can also be generated by differences between species in their dispersal abilities and rates of mortality. Naturally occurring and human-created disturbances are common causes of patch formation. A wide variety of natural disturbances are possible, including mud slides, avalanches, windstorms, ice storms, herbivore outbreaks, animal grazing, trampling, and digging, as well as fire. Mounds produced by badger digging activities in tallgrass prairie are an example of patch-producing disturbances that have important influences on patch structure as well as species

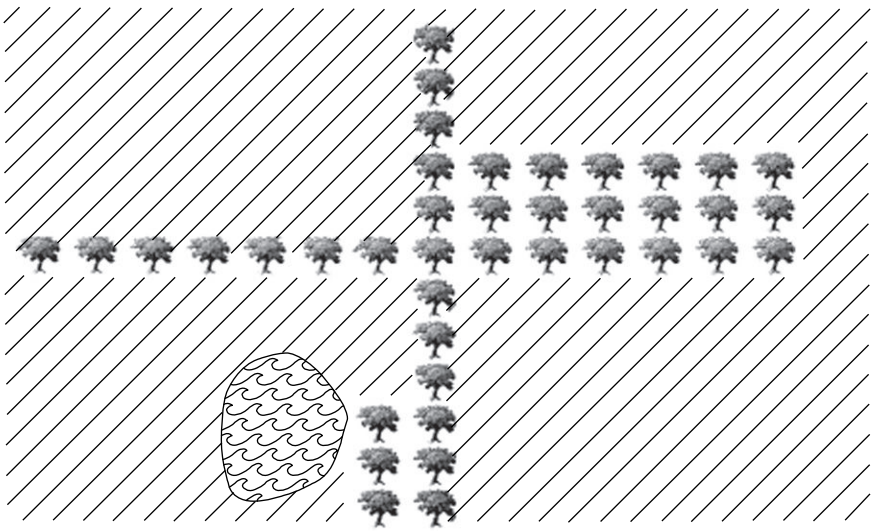


Figure 4 A matrix of agricultural land with patches of two types: woodlots and a farm pond. Windbreaks of trees act as corridors connecting the woodlots. Hedgerows and wooded windbreaks often create travel routes for wildlife.

diversity (Platt, 1975). Human activities, such as forest cutting, altered fire regimes, cultivation, urban development, introduction of pests, and strip mining for surface coal and minerals, also produce disturbance patches. Many landscapes are influenced by both natural- and human-caused disturbances, and distinguishing the separate effects on landscape diversity can be difficult. In a recent study, pollen and charcoal were collected from small lakes in Massachusetts to reconstruct long-term vegetation dynamics as related to disturbance history (Fuller *et al.*, 1998). This reconstruction over the past 1000 years included the period of time before European settlement, when the primary disturbances were fire and wind. Landscape patterns in forest composition following settlement by Europeans were largely influenced by clearing of forests for agricultural purposes and timber. These researchers found that the past history of disturbance as a result of settlement has persistent effects on current landscape patterns.

Landscape fragmentation is closely related to disturbance. Many forms of disturbance effectively break up large patches into smaller pieces. Decreases in patch size, connections between patches, and total interior area as a result of fragmentation have important implications for species and landscape diversity. As landscapes become more fragmented, patch diversity increases, with subsequent increases in edge species, exotic species, and generalists. Richness in interior species tends to decrease. Fragmentation of landscapes by human activities is considered a major threat to biodiversity worldwide (Saunders *et al.*, 1991; Bierregaard *et al.*, 1992). A major focus of the field of conservation biology is the design of nature reserves to maximize the likelihood of species existence and to minimize the loss of species to extinction. These processes are discussed in the Section, Biodiversity Planning at the Landscape Level.

Another cause of patch formation is environmental heterogeneity, which refers to variation in soils, topography, and other landform features. This variation in the physical environmental leads to a heterogeneous or a patchy spatial distribution of resources, including water, nutrients, and light. Plant species found in a resource patch can differ from species in other patches containing different levels of resources. The importance of spatial heterogeneity to species diversity has been well documented, and is most closely related to beta species diversity. These ideas have also been extended to landscape diversity, where studies have linked measures of alpha and beta species diversity with landscape diversity (Romme, 1982; Lapin and Barnes, 1995). Large-scale gradients in landscape diversity can also be related to broad-scale patterns in the environment. For example, spatial variation in climate, topography, and soils was found to be strongly related to latitudinal gradients in the richness of land cover types across the continental US (Wickham *et al.*, 1995).

Landscape Function

Interactions among the spatial elements of a landscape are the major components of landscape function. These flows of energy, materials, and species among patches, or among patches, corridors, and the surrounding matrix, are at least as important to the maintenance of diversity as patch size and configuration. However, these flows have not been as well studied as landscape structure. An example of flows among different patch types is the dispersal of seeds from forest patches into clear-cuts, which has important effects on vegetation dynamics in these open areas. Boundaries or edges between patches or between patches and the mosaic often control the strength of interactions or the amount and kinds of materials that can move between the landscape elements. Because of the importance of edges, boundaries of patches can have very different characteristics than interiors. For example, edges of recently disturbed tropical forest patches have greater tree mortality and increased recruitment of early-successional species compared with interior areas (Bierregaard *et al.*, 1992). Boundaries can also change location through time, with resulting effects on landscape structure. In contrast to boundaries, where movement is generally restricted, corridors linking similar landscape elements tend to improve or enhance flows. Movements of organisms through corridors become increasingly important as the landscape becomes more fragmented (Saunders and Hobbs, 1991).

Studies have also examined the influences of landscape structure on flows of organisms and materials. Patchy environments in Yellowstone National Park were found to be more resistant to large fires than were homogeneous landscapes, and after burning, they had a greater ability to maintain water quality (Knight and Wallace, 1989). Historical migration patterns of wild ungulates, such as wildebeest in the Serengeti of Africa and bison in North America, were mostly related to patterns of rainfall that were spatially variable both locally and regionally (reviewed in Frank *et al.*, 1998). Changes in landscape structure through fencing and urbanization have restricted migration patterns and resulted in animal overabundance and overgrazing in wildlife preserves. Biogeochemical fluxes, such as CO_2 and various forms of nitrogen, can also be affected by patches within a mosaic structure created by human land use. Gene flow and metapopulation dynamics are other examples of processes that respond to spatial structure in a landscape. A population that is spatially subdivided into patches that are connected through dispersal is called a metapopulation. Movement of individuals between subpopulations can reduce the risk of local extinction of species within small isolated patches.

Landscape Dynamics

Landscape structure and function can change for many reasons and in many ways. Changes can occur over very small or very large areas and over short or long time spans. The gap caused by a single tree falling in the forest during a storm is small and temporary, while an entire forest may be leveled by a hurricane and may take decades to centuries to recover.

Vulnerability or sensitivity to change varies from landscape to landscape. This vulnerability (or, conversely, stability) is traditionally divided into two components: resistance and resilience. Resistance is the ability of a patch or a landscape to remain unaffected by a disturbance. A grassland is much more resistant to wind damage than a forest, since grasses can bend with the wind without breaking. Resilience is the ability of a patch or a landscape to recover after a disturbance. Temperate forests recover after clearing much more quickly than tropical

forests (which may never recover) owing to differences in soil depth and fertility.

Change in a patch or a landscape can be caused by any number of factors. Some of these are intrinsic to the population being studied, including recruitment, growth, mortality, and spread or migration, which can lead to invasions or extinctions as well as changes in patch boundaries. Other causes are extrinsic to the ecosystem and are imposed by outside forces, such as climate change and disturbance events. Human transformations of the landscape include deforestation and reforestation, urbanization, corridor construction, and agricultural conversion. The effects of consumers, pathogens, and especially humans can be considered either intrinsic or extrinsic depending on the particular point of view. The potential causes of change may be interrelated in complex ways. A drought may make a forest more vulnerable to pathogens or a new clearing may increase the vulnerability of adjacent trees to windthrow.

Changes in landscape structure can have several spatial and temporal forms. Patches can shrink or expand, or be lost entirely. Successional dynamics on patches can lead to a shifting mosaic of patch types through time. Species interactions with other species and with their environment, as well as dispersal of new species into patches, are primary determinants of the regrowth of plants on these successional patches. Changes in patch size and shape can occur along edges, such as the clearing of forest, to increase the size of a cultivated field (**Figure 5(a)**). A new patch type may spread outward from a corridor (**Figure 5(b)**). For example, housing developments often spread from the course of new roads. Alternatively, a patch type may spread out from a nucleus that could be a remnant of a previous vegetation type or an introduction site for a new patch type (**Figure 5(c)**). Some changes are nearly instantaneous and occur over very short periods of time, such as the effect of fire. Other changes occur slowly and take a longer period of time to develop, such as suburbanization and desertification (Peters *et al.*, 2004).

Patch configuration on a landscape can also change. Patches can become perforated by other patch types, and large patches can be fragmented into several smaller patches. Landscape fragmentation, particularly in the tropics, is having severe effects on species biodiversity. Some of the potential consequences of fragmentation include the loss of patch types and their characteristic species, decreased connectivity with its repercussions for species movements, and decreased interior area. The biggest consequence for species diversity is the associated loss of interior species and the increase of generalist or edge species.

Landscape-level dynamics are often studied with ecological models since the temporal scales of interest are often greater than the human life span, and experiments are difficult to perform at large spatial scales. There are four general classes of models that are used to predict landscape dynamics: transition probability models, individual-based models, ecosystem process models, and biogeographic models (Peters, 2011). Transition process models are useful when the factors causing landscape change are not represented mechanistically. For example, assume that a landscape with three patch types was sampled twice, before and after an event. A table can be constructed showing the percentage of each patch type that remained the same or that was transformed into a different patch type in this hypothetical landscape (**Table 1**). Each row shows the fate of a particular patch type. Over a single time step, 60% of the forest land remained forested, 25% was converted to agricultural uses, and 15% was developed. These transition probabilities can be used to extrapolate into the future by individual time steps. **Figure 6** shows the projected change over 25 time steps if this original landscape had 50 units of forest, 25 units of agriculture, and 10 units of developed land. This landscape will stabilize with a high proportion of developed land and a very small forest area. Transition analyses are very simple to conduct and can be useful for examining the effects of various probabilities and initial conditions. In the simple form presented here, no

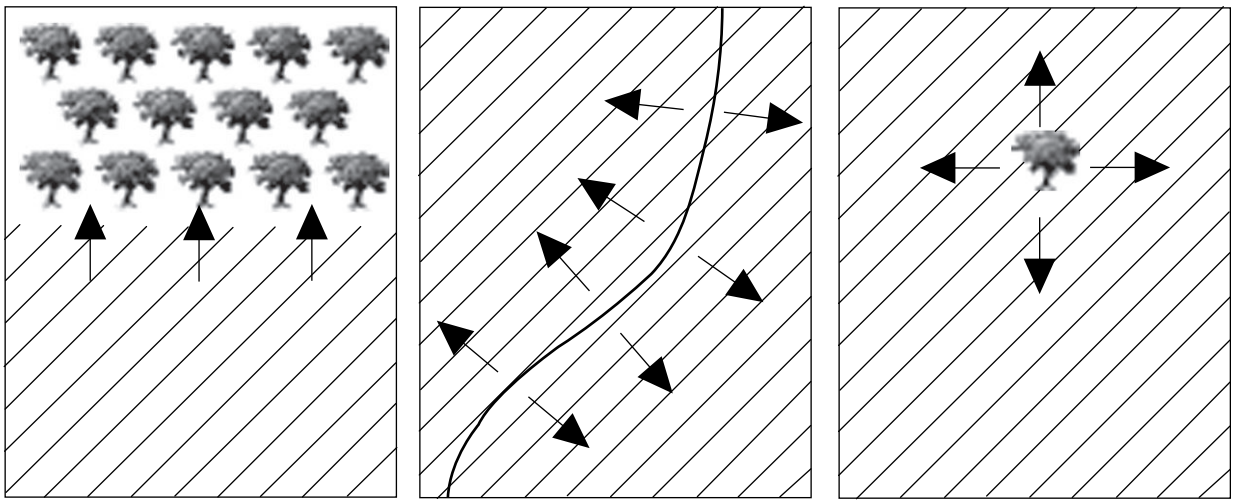


Figure 5 Three common patterns of landscape change. Arrows indicate the direction of spread. (a) An edge where agricultural land encroaches on forest along a field boundary. (b) A corridor where development spreads outward from a new road. (c) A point where reforestation proceeds outward from an isolated forest patch.

allowance was made for variations in the rate of change and no specific spatial component was included.

Each cell shows the percentage of the landscape area that changed from the patch type in that row to the patch type in that column over a single time step. For example, 25% of the original forest land was cleared for agriculture and 15% was developed. **Figure 6** shows this projected change over 25 time steps.

Individual-based simulation models are useful when information is known about the mechanisms underlying changes in landscape structure. These models incorporate life-history traits of individuals and the mechanisms by which they interact with their environment to predict landscape-level dynamics (e.g., Peters, 2002). Landscapes are simulated by linking plots together in a grid or a transect. Plots are spatially interactive through processes such as seed dispersal. Spatially interactive individual-based models can represent a variety of environmental conditions, including differences in soil properties, climate, and disturbance regime (Coffin and Lauenroth, 1994). These models are most commonly used for evaluating changes in the diversity of groups of similar species (i.e., functional types) rather than species diversity (Paruelo *et al.*, 2008).

A third class of models simulates ecological processes, including rates of nutrient cycling, water balance, and primary production (e.g., CENTURY and DayCent: Parton *et al.*, 1987, 1998). These models have been linked with geographic information systems (GIS) to simulate large regions. The effects of climate, soil texture, and management on soil organic carbon dynamics were simulated for the central grasslands of the US (Burke *et al.*, 1991). Across this large region, soil organic carbon increased with precipitation and decreased with temperature and percentage sand content. Biogeographic models are a fourth class of models that can be used to investigate vegetation responses to environmental heterogeneity. These models incorporate large-scale variations in climate and soils, as well as water and energy constraints on plant growth, to simulate continental and global patterns in vegetation. Biogeographic models are most useful for simulating responses of plant functional types at large spatial scales, to either equilibrium or transient environmental conditions (Prentice *et al.*, 1992; Neilson and Drapek, 1998). These models can incorporate landscape-scale processes or connections among patches, such as spread of wildfire (Lenihan *et al.*, 2008).

Although each of these types of models has traditionally been used independently, the linking of different models together has considerable potential for addressing issues related to landscape diversity. Because of important feedbacks between species and rates of ecosystem processes (Schulze and Mooney, 1994), linking a spatially interactive individual-based model with an ecosystem model can simulate the dynamics of landscape structure and function as well as changes in functional group diversity (Epstein *et al.*, 1999). A nonspatial individual-based model linked with a nutrient cycling model was also used to examine the importance of soil heterogeneity to forest responses to global climate change (Pastor and Post, 1988). The incorporation of landscape-scale flows of water, carbon, and nutrients into a spatially interactive individual-based model is an important research area

Table 1 Transition matrix for a hypothetical landscape with three land-use types

	Forest	*Agriculture*	*Developed*
Forest	60	25	15
Agriculture	10	75	15
Developed	0	2	98

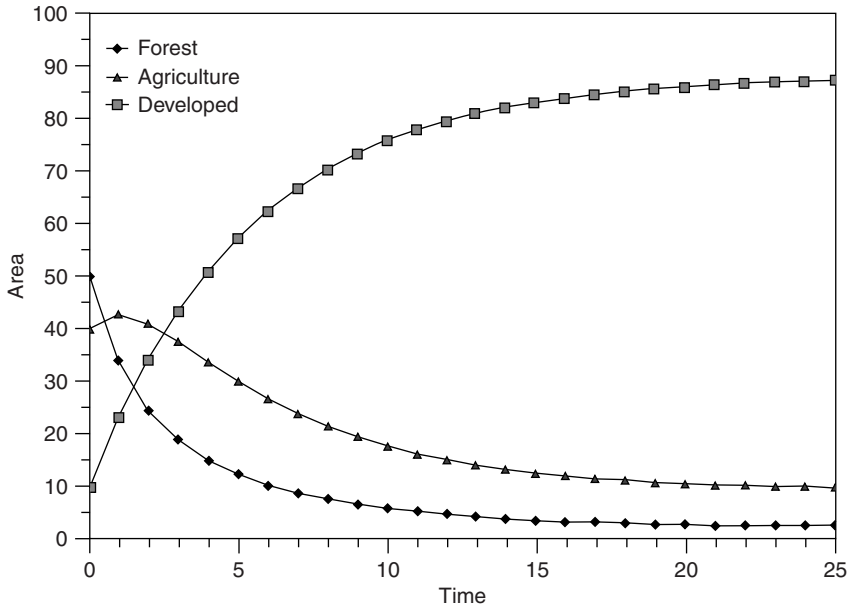


Figure 6 Transition matrix analysis of a hypothetical landscape with three patch types of differing abundance: forest, 50 units; agriculture, 25 units; and developed, 10 units. Transition probabilities are shown in **Table 1**.

for predicting landscape diversity dynamics that includes individual species responses (Peters, 2011). Linkages between biogeochemical and biogeographic models are another area of application to landscape diversity, especially if species or plant functional types are resolved and landscape-scale processes, such as disturbance regime, are included (Lenihan *et al.*, 2008).

Biodiversity Planning at the Landscape Level

To preserve species diversity most effectively, management plans must preserve the habitats and landscape structures needed by the target species, rather than simply preserving the species in isolation from the larger, potentially changing the environment. Management practices aimed directly at a particular species run the risk of losing ecosystem functions that might actually be crucial for the target species, but that were unknown when the management plan was created. Furthermore, maximizing benefits for one species may threaten others. The ideal is to preserve overall ecosystem health, including species diversity. Unfortunately, this is easier said than done. Much of the effort of conservation biologists has been directed toward learning how to manage ecosystems, at both theoretical and practical levels.

One of the classic debates in conservation biology centers around the best reserve design. If limited resources are available to purchase land, is it better to establish one large reserve or a few smaller ones? This has become such a well-known and controversial issue that it has its own acronym: single large or several small (SLOSS). A large reserve provides the most potential habitat for interior species, which are usually the ones most in need of protection. However, a single reserve is vulnerable to all sorts of disasters. If a major hurricane or a pathogen hits that reserve, there are no other reserves to take its place. The establishment of several smaller reserves minimizes the risk of losing everything at the same time. However, a minimum size is needed to sustain populations of interior species as well as to preserve the characteristic species diversity and species composition of the ecosystem. Furthermore, reserves do not operate like isolated islands; thus, connections between reserves and the surrounding habitats are also important.

A related concept involved in determining the optimum size of a nature reserve is the minimum dynamic area (Pickett and Thompson, 1978). Assuming that the disturbance regime of an area is known, the frequency, areal extent, and recovery time can be used to determine the smallest reserve area in which there will always be some mature patch types to provide a species source for the rest of the area as it recovers from disturbance. If a patch is smaller than this minimum dynamic area, it will likely be eliminated through time simply from natural disturbances.

Given the large number of species on the planet, it is impossible, or at best impractical, to manage for every one of them. Instead, conservation biologists are now trying to identify ways to simplify the task of landscape-level management. The most promising methods identify one or a few important species and focus on their management. One tactic is to manage keystone species, those on which important ecosystem functions or other species depend. Another approach is to target umbrella species, those with large ranges or broad habitat requirements. Managing for these species will automatically save many other species with smaller or less inclusive requirements. A similar method identifies a set of focal species, each of which is sensitive to a particular aspect of landscape structure or function. One of the focal species might be especially vulnerable to habitat fragmentation, whereas another might require a high level of connectivity. The protection of this set of sensitive species provides the management goals. When the requirements of the sensitive species are met, other species will likely be provided for as well.

Case Study: Sevilleta National Wildlife Refuge

The Sevilleta National Wildlife Refuge (SNWR; 34.5° N, 106.9° W), located approximately 75 km south of Albuquerque, New Mexico, provides an excellent example of landscape diversity and its relationship with species diversity. This 100,000 ha wildlife refuge was established in 1973 and is currently managed by the US Fish and Wildlife Service. The refuge is also a Long-Term Ecological Research site funded by the U.S. National Science Foundation. The climate at the SNWR is semiarid to arid, with low amounts of precipitation and high temperatures during the April to October growing season. The mean annual precipitation over the past 65 years was 23.4 cm. (SD = 70.4 cm) and the average annual temperature was 14.1 °C (SD = 0.7 °C).

The SNWR is uniquely located at the ecotonal boundary between four major grassland–shrubland biomes found within the continental US (Peters *et al.*, 2006). Two of these biomes, shortgrass steppe ecosystems and Chihuahuan desert grasslands, form transition zones in the eastern part of the refuge (**Figures 7(a)** and **(b)**). Patches of variable size (<10 to >1000 m^2) and shape occur and result in a high landscape diversity (**Figure 7(b)**). These patches can be differentiated into one of two patch types based on the cover of the dominant plant species (Gosz, 1995; Kröel-Dulay *et al.*, 2004). The vegetation of some patches consists mostly of blue grama (*Bouteloua gracilis*), the dominant species in shortgrass steppe ecosystems. A second patch type occurs, where the majority of cover is black grama (*Bouteloua eriopoda*), a dominant grass in Chihuahuan desert ecosystems. Species richness is similar for both patch types, although the canopy cover of plants is higher in black grama compared with blue grama patches (Peters *et al.*, 2006). A transect across the conceptual landscape shown in **Figure 7(b)** goes through each patch type as well as the matrix vegetation where similar cover of both species occurs (**Figure 7(c)**).

Within each patch, a smaller scale of heterogeneity also exists owing to disturbances associated with the burrowing activities of bannertail kangaroo rats (see **Figure 7(b)**). Mounds can be distinguished into one of three types based on plant species diversity as well as composition (**Figure 8(a)**). Active mounds are the site of frequent burrowing; thus, only plant species well adapted to disturbance can survive there (Fields *et al.*, 1999). Typically, this vegetation consists of small plants representing few species (**Figure 8(b)**). After mounds are abandoned and burrowing activities cease, more plant

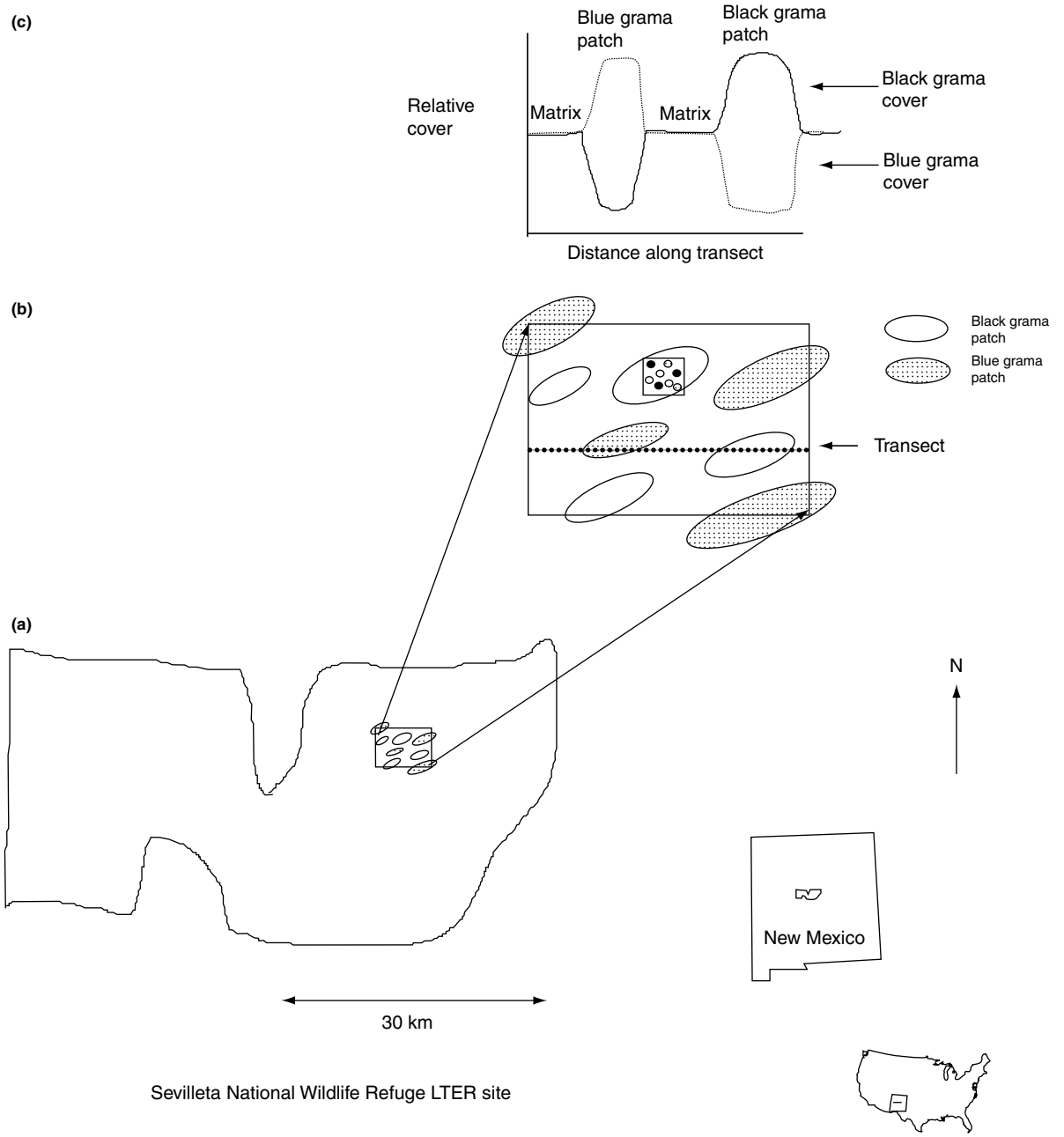


Figure 7 Patch types at the Sevilleta National Wildlife Refuge LTER site (SNWR). (a) Location of the SNWR in central New Mexico, US, and location of patch types in the eastern part of the refuge, (b) location of two patch types on the landscape. The transect across the landscape goes through patches dominated by either blue grama, black grama, or the matrix containing both species, (c) cover of blue grama and black grama along the transect, showing one approach for identifying and delineating patches.

species can survive to larger sizes on these early successional mounds. Through time, competition among plants typically reduces the number of species, although plant sizes can be quite large as one or a few plants come to dominate late-successional mounds. This invasion–abandonment cycle results in a shifting mosaic of mound types through time across the landscape (**Figure 8(a)**). Although the species diversity on mounds changes through time, and the location of mound types varies spatially across the landscape, the total numbers of species and patch types remain constant on the scale of the landscape (**Figure 8(c)**). Therefore, landscape diversity both reflects and determines patterns in diversity at smaller levels of organization, and in particular, species diversity.

Broad-scale drivers such as climate and fire can also affect patterns in diversity within each patch and across the landscape. Black grama is more sensitive to wildfire, livestock grazing, and drought than blue grama (Gosz and Gosz, 1996; Parmenter, 2008). Changes in species production and dominance can result in altered patterns in species diversity, with consequences at the landscape scale (Ryerson and Parmenter,

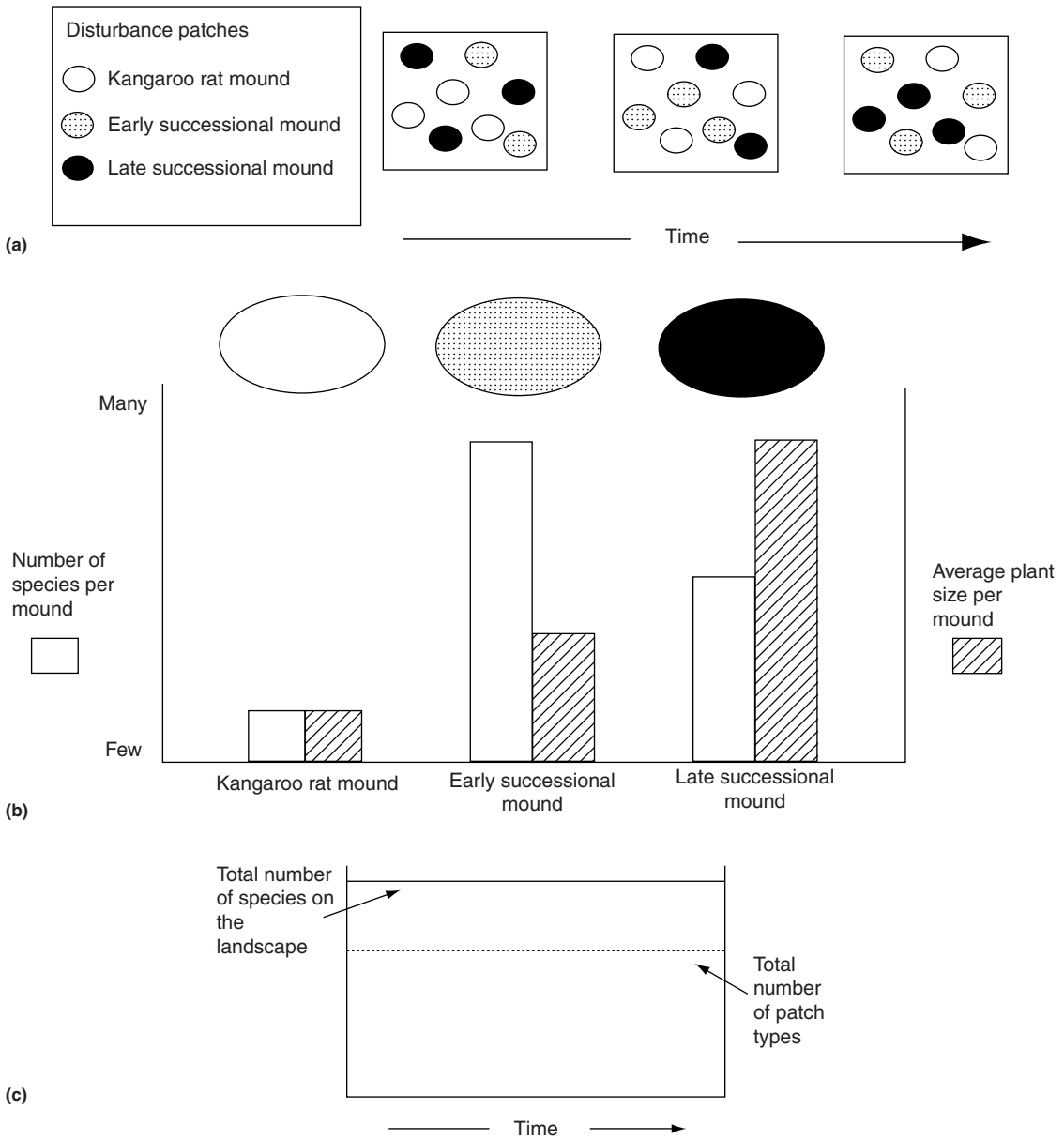


Figure 8 Small-scale heterogeneity within one black grama patch from **Figure 7b**. Three types of mounds can be defined based on plant species composition and burrowing activities of bannertail kangaroo rats. (a) Mounds change from one type to another through time as kangaroo rats abandon some mounds and invade others. (b) Number of plant species and average plant size for each mound type. (c) Because the number of mounds within each of the three types is constant in this hypothetical landscape, the total number of species found on the landscape is also constant.

2001). The current frequency of El Niño Southern Oscillation (ENSO) events that increase winter rainfall does not alter the patterns at the landscape scale; however, productivity patterns may respond on an annual basis (Muldavin *et al.*, 2008; Xia *et al.*, 2010). Climate projections should be considered in modeling how the future ecotone dynamics may change through altered frequency and strength of ENSO events.

Recent experimental manipulations suggest that cover of the dominant grass, black grama, declines drastically (up to 66%) during prolonged drought periods whereas cover of the invasive shrub, creosotebush (*Larrea tridentata*), hardly changes under these same experimental conditions. In addition, cover of black grama in a lightly grazed pasture along the northern boundary of the SNWR declined 50% relative to long-term ungrazed areas on the Refuge. Thus, the abundance of black grama is negatively affected by common disturbances including grazing, fire, and prolonged drought, all of which may combine to further promote shrub encroachment in this region. However, long-term monitoring of plant cover under undisturbed conditions shows that the abundance of black grama is increasing at a greater rate than Great Plains grasses, such as blue grama. Moreover, experimental nighttime warming year-round favors the growth of black grama (Collins *et al.*, 2010) but not blue grama. Instead, blue grama

responded more favorably than black grama to low levels of nitrogen addition (Báez *et al.*, 2007). Thus, the composition and abundance of dominant species across this dynamic ecotone are governed by multiple forces that interact in complex ways to affect the rate and direction of landscape change in this aridland ecosystem.

Conclusions

Although much of the current emphasis on biodiversity has been at the level of species, landscape diversity is also important. The preservation and maintenance of multiple levels of organization, including species, populations, communities, and ecosystems, require an understanding of how these various levels interact with their environment across a range of spatial scales. Maintenance of landscape diversity provides a spatial template for the preservation of these smaller levels of organization, and in particular, for species biodiversity. Changes in landscape structure and function through time have important effects on the distribution of resources, with resulting influences on the survival of species in both natural and managed ecosystems. Because of the overwhelming numbers of species, it may be impractical to attempt to conserve species diversity per se. By focusing on landscape diversity and the perpetuation of dynamic processes across multiple scales, an attempt can be made to preserve entire ecosystems with their full complement of genetic diversity.

See also: Deforestation and Land Clearing. Desertification. Keystone Species. Land-Use Issues. Range Ecology, Global Livestock Influences

References

Báez S, Fargione J, Moore DI, Collins SL, and Gosz JR (2007) Atmospheric nitrogen deposition in the northern Chihuahuan desert: Temporal trends and potential consequences. *Journal of Arid Environments* 68: 640–651.

Bierregaard ROJ, Lovejoy TE, Kapos V, dos Santos AA, and Hutchings RW (1992) The biological dynamics of tropical rainforest fragments. *BioScience* 42: 859–866.

Bormann FH and Likens GE (1979) *Pattern and Process in a Forested Ecosystem: Disturbance, Development and the Steady State Based on the Hubbard Brook Ecosystem*. New York: Springer-Verlag.

Burke IC, Kittel TGF, Lauenroth WK, Snook P, Yonker CM, and Parton WJ (1991) Regional analysis of the Central Great Plains. *BioScience* 41(10): 685–692.

Burnett MR, August PV, Brown Jr JH, and Killingbeck KT (1998) The influence of geomorphological heterogeneity on biodiversity. I. Patch-scale perspective. *Conservation Biology* 12(2): 363–370.

Coffin DP and Lauenroth WK (1994) Successional dynamics of a semiarid grassland: Effects of soil texture and disturbance size. *Vegetation* 110: 67–82.

Collins SL, Fargione JE, Crenshaw CL, *et al.* (2010) Rapid plant community responses during the summer monsoon to nighttime warming in a Northern Chihuahuan Desert grassland. *Journal of Arid Environments* 74: 611–617.

Cornelius JM and Reynolds JF (1991) On determining the statistical significance of discontinuities within ordered ecological data. *Ecology* 72(6): 2057–2070.

Epstein HE, Burke IC, and Lauenroth WK (1999) Responses of the shortgrass steppe to changes in rainfall seasonality. *Ecosystems* 2: 139–150.

Fields MJ, Coffin DP, and Gosz JR (1999) The role of kangaroo rats (*Dipodomys spectabilis) in determining patterns in plant species dominance at an ecotonal boundary. Journal of Vegetation Science* 10: 123–130.

Frank DA, McNaughton SJ, and Tracy BF (1998) The ecology of the earth's grazing ecosystems. *BioScience* 48: 513–521.

Fuller JL, Foster DR, McLachlan JS, and Drake N (1998) Impact of human activity on regional forest composition and dynamics in central New England. *Ecosystems* 1: 76–95.

Goslee SC and Sanderson MA (2010) Landscape context and plant community composition in grazed agricultural systems of the northeastern United States. *Landscape Ecology* 25: 1029–1039.

Gosz JR (1995) Edges and natural resource management: Future directions. *Ecology International* 22: 17–34.

Gosz RJ and Gosz JR (1996) Species interactions on the biome transition zone in New Mexico: Response of blue grama (*Bouteloua gracilis*) and black grama (*Bouteloua eriopoda*) to fire and herbivory. *Journal of Arid Environments* 34: 101–114.

Huston MA (1994) *Biological Diversity: The Coexistence of Species on Changing Landscapes*. Cambridge, UK: Cambridge University Press.

Knight DH and Wallace LL (1989) The Yellowstone fires: Issues in landscape ecology. *BioScience* 39: 700–706.

Kröel-Dulay G, Odor P, Hochstrasser T, and Peters DPC (2004) Compositional comparison of grass-dominated patches at a semiarid–arid ecotone. *Journal of Vegetation Science* 15: 531–538.

Lapin M and Barnes BV (1995) Using the landscape ecosystem approach to assess species and ecosystem diversity. *Conservation Biology* 9(5): 1148–1158.

Lenihan JM, Bachelet D, Neilson RP, and Drapek R (2008) Simulated response of conterminous United States ecosystems to climate change at different levels of fire suppression, CO_2 emission rate, and growth response to CO_2. *Global and Planetary Change* 64: 16–25.

Lovejoy TE, Bierregaard RO, Rylands AB, *et al.* (1986) Edge and other effects of isolation on Amazon forest fragments. In: Soulé M (ed.) Conservation Biology. *The Science of Scarcity and Diversity*, pp. 251–285. Sunderland, MA: Sinauer Associates.

Lovejoy TE, Rankin JM, Bierregaard Jr RO, *et al.* (1984) Ecosystem decay of Amazon forest remnants. In: Niteki MH (ed.) *Extinctions*, pp. 295–325. Boulder, CO: Westview Press.

Mack RN (1981) Invasion of *Bromus tectorum* L. into western North America: An ecological chronicle. *Agro-Ecosystems* 7: 145–165.

Milne BT (1988) Measuring the fractal dimension of landscapes. *Applied Mathematics and* Computation 27: 67–79.

Muldavin EH, Moore DI, Collins SL, Wetherill K, and Lightfoot D (2008) Aboveground net primary production dynamics in a northern Chihuahuan Desert ecosystem. *Oecologia* 155: 123–132.

Neilson RP and Drapek RJ (1998) Potentially complex biosphere responses to transient global warming. *Global Change Biology* 4: 505–521.

Nichols WF, Killingbeck KT, and August PV (1998) The influence of geomorphological hetrogeneity on biodiversity. II. A landscape perspective. *Conservation Biology* 12(2): 371–379.

O'Neill RV, Krummel JR, Gardner RH, *et al.* (1988) Indices of landscape pattern. *Landscape Ecology* 1: 153–162.

Parmenter RR (2008) Long-term effects of a summer fire on desert grassland plant demographies in New Mexico. *Rangeland Ecology and Management* 61: 156–168.

Parton WJ, Hartman M, Ojima D, and Schimel D (1998) DayCent and its land surface submodel-description and testing. *Global and Planetary Change* 19: 35–48.

Paruelo JM, Pütz S, Weber G, *et al.* (2008) Long-term dynamics of a semiarid grass steppe under stochastic climate and different grazing regimes: A simulation analysis. *Journal of Arid Environments* 72: 2211–2231.

Pastor J and Post WM (1988) Response of northern forests to CO_2-induced climate change. *Nature* 344: 55–58.

Peters DPC (2002) Plant species dominance at a grassland-shrubland ecotone: An individual-based gap dynamics model of herbaceous and woody species. *Ecological Modelling* 152: 5–32.

Peters DPC (2011) Grassland simulation models: A synthesis of current models and future challenges. In: Jorgensen SE (ed.) *Handbook of Models used in Ecosystem and Environmental Management*. London: Taylor and Francis Publishing Co.

Peters DPC, Gosz JR, Pockman WT, *et al.* (2006) Integrating patch and boundary dynamics to understand and predict biotic transitions at multiple scales. *Landscape Ecology* 21: 19–33.

Peters DPC, Pielke Sr RA, Bestelmeyer BT, Allen CD, Munson-McGee S, and Havstad KM (2004) Cross scale interactions, nonlinearities, and forecasting catastrophic events. *Proceedings of the National Academy of Sciences* 101: 15130–15135.

Pickett STA and Thompson JN (1978) Patch dynamics and the size of nature reserves. *Biological Conservation* 13: 27–37.
Pickett STA and White PS (eds.) (1985) *The Ecology of Natural Disturbance and Patch Dynamics*. New York: Academic Press.
Platt WJ (1975) The colonization and formation of equilibrium plant species associations on badger disturbances in a tall-grass prairie. *Ecological Monographs* 45: 285–305.
Prentice C, Cramer W, Harrison SP, Leemans R, Monserud RA, and Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19: 117–134.
Riitters KH, O'Neill RV, Hunsaker CT, *et al.* (1995) A factor analysis of landscape pattern and structure metrics. *Landscape Ecology* 10: 23–40.
Romme WH (1982) Fire and landscape diversity in subalpine forests of Yellowstone National Park. *Ecological Monographs* 52: 199–221.
Romme WH and Knight DH (1982) Landscape diversity: The concept applied to Yellowstone Park. *BioScience* 32: 664–670.
Rowe JS (1992) The ecosystem approach to forestland management. *Forest Chronicle* 68: 222–224.
Ryerson DE and Parmenter RR (2001) Vegetation change following removal of keystone herbivores from desert grasslands in New Mexico. *Journal of Vegetation Science* 12: 167–180.
Saunders DA and Hobbs RJ (eds.) (1991) *The Role of Corridors*. Chipping Norton, Australia: Surrey Beaty & Sons.
Saunders DA, Hobbs RJ, and Margules CR (1991) Biological consequences of ecosystem fragmentation: A review. *Conservation Biology* 5: 18–32.
Schulze ED and Mooney HA (eds.) (1994) *Biodiversity and Ecosystem Function*. Heidelberg, Germany: Springer-Verlag.
Turner MG (1989) Landscape ecology: The effect of pattern on process. *Annual Review of Ecology and Systematics* 20: 171–197.
Turner MG and Gardner RH (1990) Quantitative methods in landscape ecology. *The Analysis and Interpretation of Landscape Heterogeneity*. New York: Springer-Verlag.
Whittaker RH (1960) Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs* 30: 279–338.
Whittaker RH (1972) Evolution and measurement of species diversity. *Taxon* 21: 213–251.
Wickham JD, Wade TG, Jones KB, Riitters KH, and O'Neil RV (1995) Diversity of ecological communities of the United States. *Vegetation* 119: 91–100.
Xia Y, Moore DI, Collins SL, and Muldavin EH (2010) Spatial and temporal variation of net primary production in Chihuahuan Desert annual plant communities. *Journal of Arid Environments* 74: 378–385.

Landscape Ecology and Population Dynamics

Scott M Pearson, Mars Hill College, Mars Hill, NC, USA

Published by Elsevier Inc.

Glossary

Core and satellite model Network of habitat patches in which centrally located core patches harbor stable populations that provide emigrants (or colonists) to ephemeral populations in peripheral satellite patches.
Deme Localized breeding group or subpopulation, often confined to a single habitat patch.
Effective population size (N_e) Number of individuals in population that participate in breeding.
Extinction vortex Complex of interacting mechanisms that lead to the extinction of small populations. These mechanisms include loss of genetic diversity as well as environmental and demographic stochasticity.
Habitat fragmentation Process in which habitat loss converts a landscape dominated by large connected patches of habitat to one in which habitat exists in small, disconnected patches or a condition in which habitat is spatially distributed as small, disconnected patches.
Landscape matrix The complex of habitats occupying the space between habitat patches.
Metapopulation Network of individual demes connected by the processes of immigration and emigration.
Population dynamics Change in population size through time and the processes that bring about these changes.
Sink population Deme in which death rates exceed birthrates. Sink populations will go extinct without an input of immigrants.
Source population Deme in which birthrates exceed death rates. Sources produce an excess of individuals that emigrate from that deme.

Introduction

The discipline of landscape ecology emphasizes the causes and consequences of spatial pattern on the functioning of populations, communities, and ecosystems. Nature is generally not homogenous with respect to ecological properties. Those properties include gradients in abiotic factors, such as the temperature, moisture, and the abundance of required resources; and biotic factors, such as the presence of predators, pathogens, and competing species. This spatial heterogeneity exists at multiple scales of space and time. The term "landscape" evokes patterns over a broad scale, such as the view from an airplane or satellite extending over several square kilometers and including the mosaic of forests, grasslands, wetlands, agricultural fields, and urban areas seen below (**Figure 1**). This mosaic and the diversity of habitats therein is the setting where species make their living. The spatial pattern of habitats determines how species and their populations are distributed across this view from above. The influence of spatial pattern is also important at finer scales. Studies in landscape ecology have illustrated how the spatial pattern of resources at the scale of 1 m^2 altered the habitat use by beetles and ants (Wiens and Milne, 1989; Wiens *et al.*, 1993). "Scale" is an important concept in landscape ecology because the influence of spatial patterns often depends on the scale in which they are observed and measured. It is common to measure spatial heterogeneity at multiple scales in order to best discern which scales are important for a given ecological phenomenon (e.g., Price *et al.*, 2005). No matter what scale is studied, the spatial pattern of resources alternatively provides opportunities for population spread and increase, limits and constrains population growth, or threatens the persistence of some species.

Basic Concepts and Terminology about Population Dynamics

A population is a group of interbreeding organisms that produce viable, fertile offspring; under the biological species concept, members of a population belong to a single species. Localized breeding groups are also known as "demes." In a typical landscape, the suitable habitat for most species exists in a patchy pattern (**Figure 1**), so patches of habitat usually define the location and boundaries of demes.

Ecologists are interested in the number of individuals in a population (that is, population size classically denoted by the letter N). The change in N through time (dN/dt) is population dynamics. Population size changes by the demographic processes of birth, immigration, death, and emigration (BIDE). This BIDE model of population growth is

$$dN/dt = \text{Births} + \text{Immigrants} - \text{Deaths} - \text{Emigrants}.$$

Populations in which immigration and/or emigration occurs are considered to be open populations and are connected to other demes by the movement of individuals. The distances between demes influence the frequency of these deme-to-deme movements. Movements may happen regularly for the population but occur infrequently during the life cycle of an individual. Movements often take the form of dispersal from the natal site to a different habitat patch. Dispersal lessens the probability of competition and interbreeding with parents and siblings, especially when the population in the natal patch is small. Analyses of population viability and genetic diversity emphasize effective population size (N_e), which is the number of individuals actually participating in breeding. In most populations, $N_e < N$ because some individuals are unsuccessful in attracting a mate, reproductively immature, or of postreproductive age.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00417-2

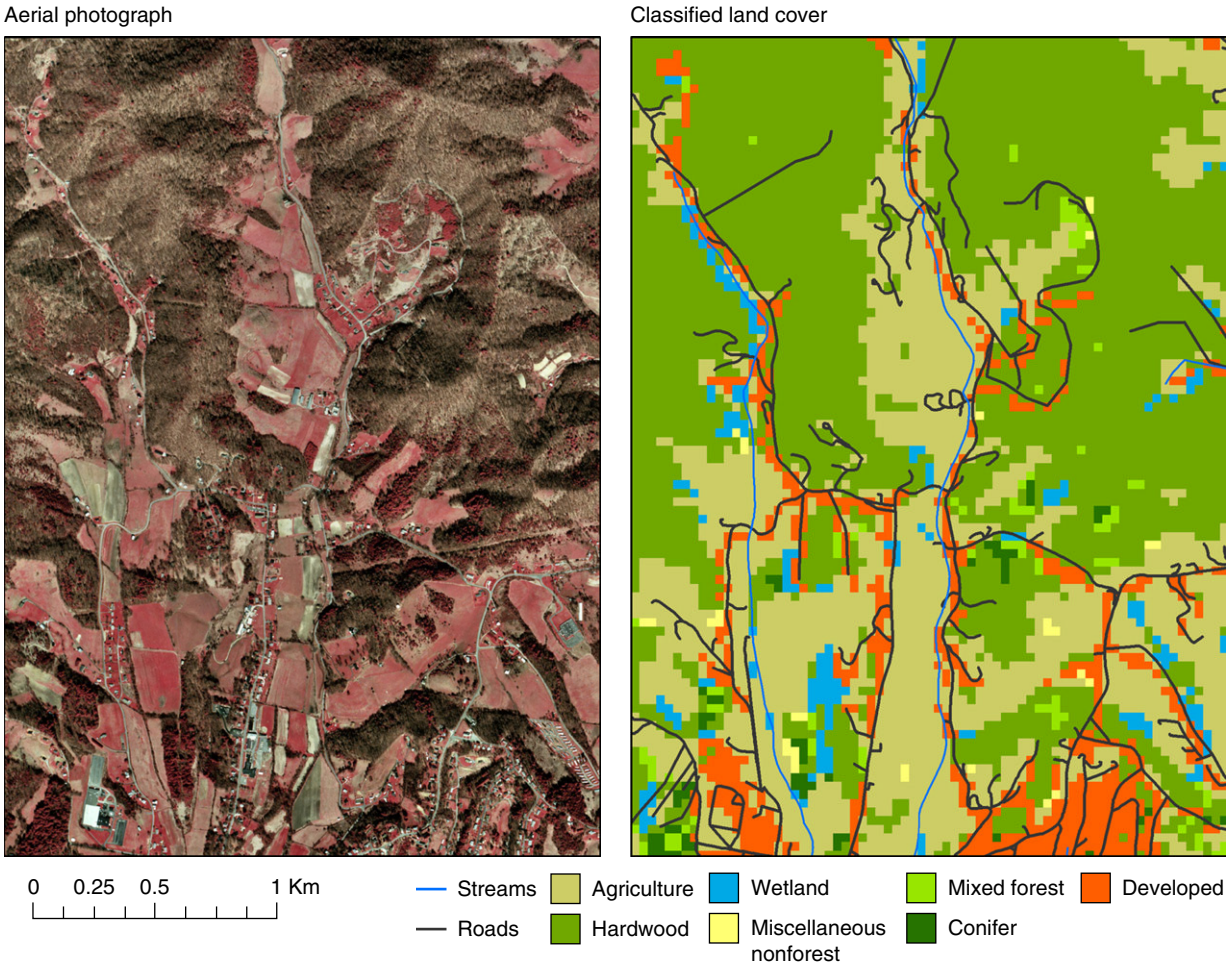


Figure 1 Landscapes are a mosaic of habitat types. Color infrared aerial photograph from 1998 (left) and classification of land cover in 2001 (right) for a region of southern Appalachian Mountains, Madison County, NC, USA.

Population Dynamics in Landscapes

The number of individuals in a population, N, is limited by available resources. The limiting resource(s) may include food or nest sites for animals, or moisture, sunlight, and/or soil nutrients for plants. Suitable habitats are the locations and conditions that provide these resources. Given that resources are finite, N has a maximum value that is sustainable. Ecologists use the term carrying capacity (K) to refer to the maximum N that a given environment can support over the long term. The logistic growth equation is a simple model of population dynamics learned by most students of ecology:

$$\frac{dN}{dt} = rN\left[\frac{K-N}{K}\right]$$

The value dN/dt represents change in N over a unit of time (t), and r is the intrinsic rate of increase, which is a measure of per capita population growth:

$$r = \frac{births - deaths}{N}$$

Assumptions of this model include a closed population (i.e., there is no immigration or emigration) and that r and K are constants. Other ecological forces that could influence population growth, such as predation or competition, are not included. The logistic growth model implies that a population attains equilibrium with resources in its environment. The continuous form of this model produces graphs of N against time as an S-shaped curve in which the population grows until $N=K$ and then N remains constant at that value. The assumptions of the logistic growth model are rather restrictive, and very few if any populations in nature would fulfill all of them. Nevertheless, it expresses a central premise that populations are limited by availability of resources and the amount of habitat in the landscapes where they live. Furthermore, landscape ecologists study how the spatial arrangement of habitat patches can alter population dynamics.

Landscape-level changes in habitat abundance and spatial pattern affect population dynamics. Data on populations of grassland birds in North America provide example of how changes in habitat availability alter population size. At the continental scale, the abundance of bird species that require grassland habitats has been declining since the mid-1900s in

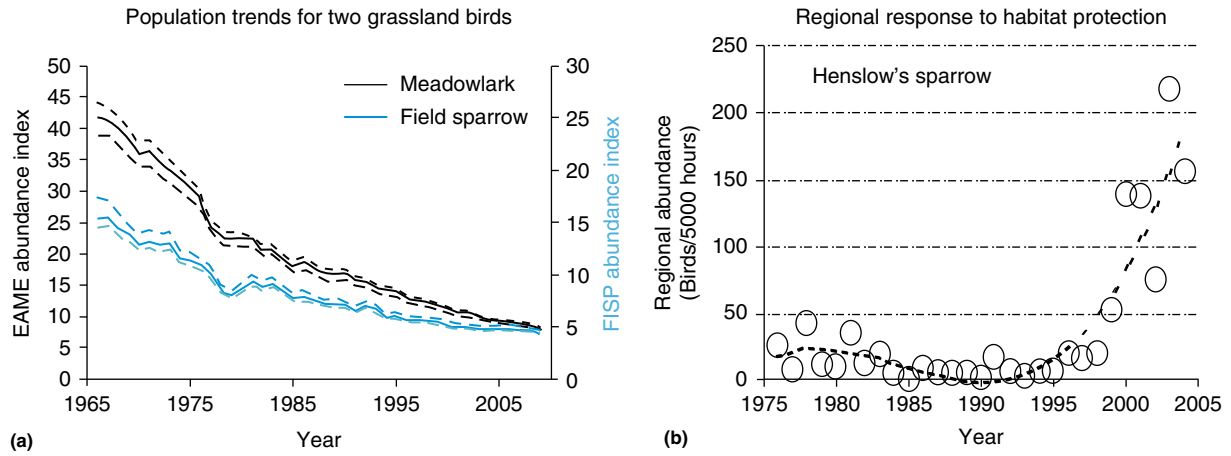


Figure 2 Change in abundance for grassland birds at (a) continental scale for eastern North America and (b) regional scale for Illinois. (a) Data for Eastern meadowlark (*Sturnella magna*) and field sparrow (*Spizella pusilla*) provided by United States Geological Survey (USGS) North American Breeding Bird Survey, http://www.mbr-pwrc.usgs.gov/bbs/bbs.html. Solid line is mean estimate, with dashed lines showing 95% credible interval. (a) and (b) Reprinted from Herkert JR (2007) Evidence for a recent Henslow's Sparrow population increase in Illinois. *Journal of Wildlife Management* 71: 1229–1233, with permission from Wiley.

North America (Peterjohn and Sauer, 1999; **Figure 2**) and parts of Europe (Chamberlain and Fuller, 2001; **Figure 2**). These declines can be explained by the loss of grassland habitats because of changing agricultural practices, expansion of suburban development, conversion of grassland habitats to forests, and long-term disruption of ecological processes (e.g., fire and grazing) that create and maintain grassland habitats (Askins *et al.*, 2007). Landscape-level analyses have revealed the specific habitat features required for foraging and nesting by these birds (Ribic and Sample, 2001; Veech, 2006). These species are sensitive to the fragmentation of their habitats. In response to these population declines, management initiatives at the regional scale have focused on habitats protection and restoration. One such initiative is the US Department of Agriculture's Conservation Reserve Program (CRP). The CRP encourages the protection of large patches of grassland habitat and has the potential to increase habitat availability for these birds. Herkert (2007) documented an increase in Henslow's sparrow (*Ammodramus henslowii*), an obligate grassland species, in response to the establishment of 400,000 ha of CRP lands in Illinois (**Figure 2**) but cautioned that local management practices (such as farmers' mowing schedules) can affect habitat quality. The positive effects of similar habitat restoration efforts have been documented in other regions (e.g., Benson *et al.*, 2006).

Monitoring population dynamics at the landscape scale is challenging. Various survey techniques are available for estimating population size, but measurement of demographic rates is more difficult. It is generally necessary to estimate demographic rates (i.e., births, deaths, and immigration) by sampling a limited number of sites or habitat patches in the landscape because of the large effort required to produce these estimates. Often it is the variation in demographic rates among habitat patches that is of greatest interest. For example, Herkert *et al.* (2003) sampled 39 sites to estimate the effects of patch size on nest survival and parasitism in three grassland bird species. In a study of song sparrow (*Melospiza melodia*) populations, Jewell and Arcese (2008) linked the population growth rates of the sparrows indirectly to landscape-level patterns of land use (**Figure 3**). Rates of brood parasitism by cowbirds (*Molothrus ater*) reduced sparrow reproduction, and cowbird abundance was strongly correlated with land uses such as urban and agricultural lands with livestock grazing. These land uses created areas where parasitism rates were so high that local reproduction in sparrow populations could not replace individuals lost to mortality (i.e., $r<0$). Their results suggest that these populations were maintained by immigration from areas with lower parasitism rates rather than by local reproduction.

Simulation models are often employed to explore the effects of landscape patterns on population dynamics. Spatially explicit models can simulate the occupancy of specific locations as well as the demographic rates characteristic of populations at those locations. Such models allow the researcher to test hypotheses about population dynamics in the complex settings of real and theoretical landscapes. These models are especially useful for studying the impacts of various scenarios of landscape change on populations. Pearson *et al.* (1999) examined how management options related to forest clearing would affect habitats, and therefore the abundances, of a suite of native species. Complex population processes such as survival, reproduction, and movement of individuals can be represented in spatially explicit models. For example, Holdo *et al.* (2011) used a model to study the impact of road construction that would interfere with migratory movements of the Serengeti wildebeest (*Connochaetes* spp.; **Figure 4**). Their simulations suggest that this barrier to movement could reduce wildebeest population by one-third because these animals would be less able to migrate to the most productive foraging sites. Locations of the most productive sites shift in location seasonally, and the road interfered with the animals' ability to track and migrate to those sites. Moreover, the road was predicted to divide a single wildebeest population into two separate subpopulations.

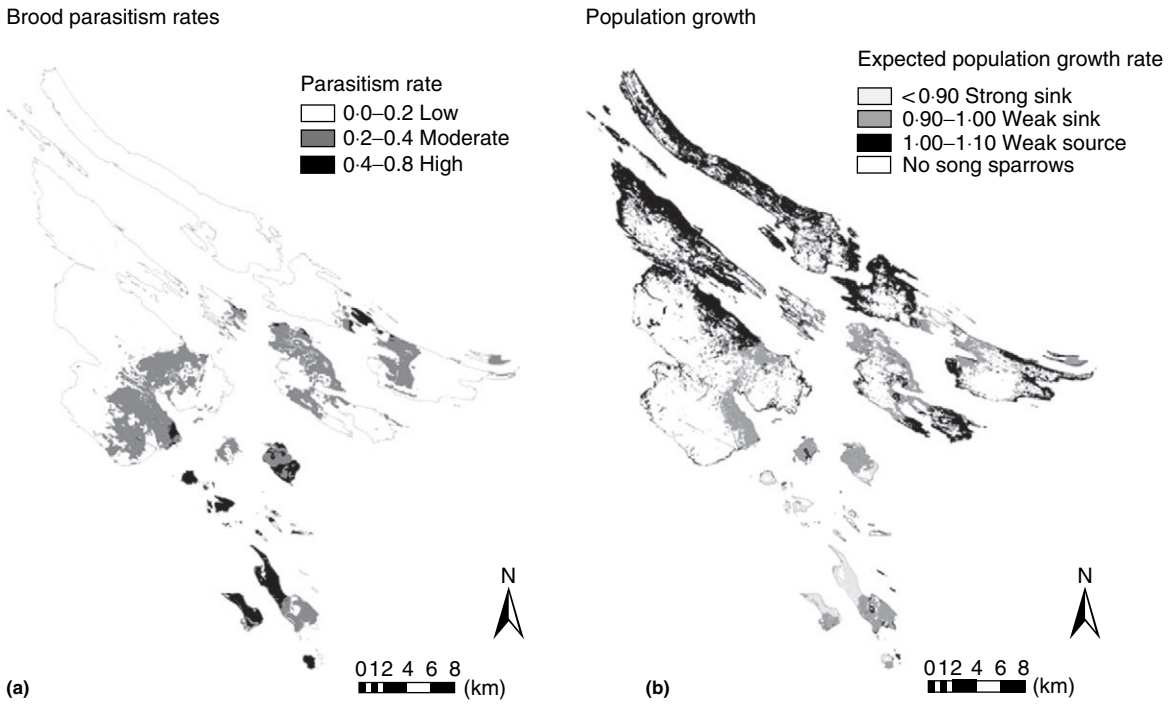


Figure 3 Maps of brood parasitism and population growth of song sparrows (*Melospiza melodia*) in the Southern Gulf Islands of British Columbia, Canada. Rates of parasitism by brown-headed cowbirds (*Molothrus ater*) (a) are affected by land use and results in variation in population growth rates of their sparrow hosts (b). Growth rates are expressed in terms of finite rates of increase. Rates ≥ 1 indicate stable or growing populations (sources), whereas rates <1 represent declining or sink populations. Reprinted from Jewell KJ and Arcese P (2008) Consequences of parasite invasion and land use on the spatial dynamics of host populations. *Journal of Applied Ecology* 45: 1180–1188, with permission from Wiley.

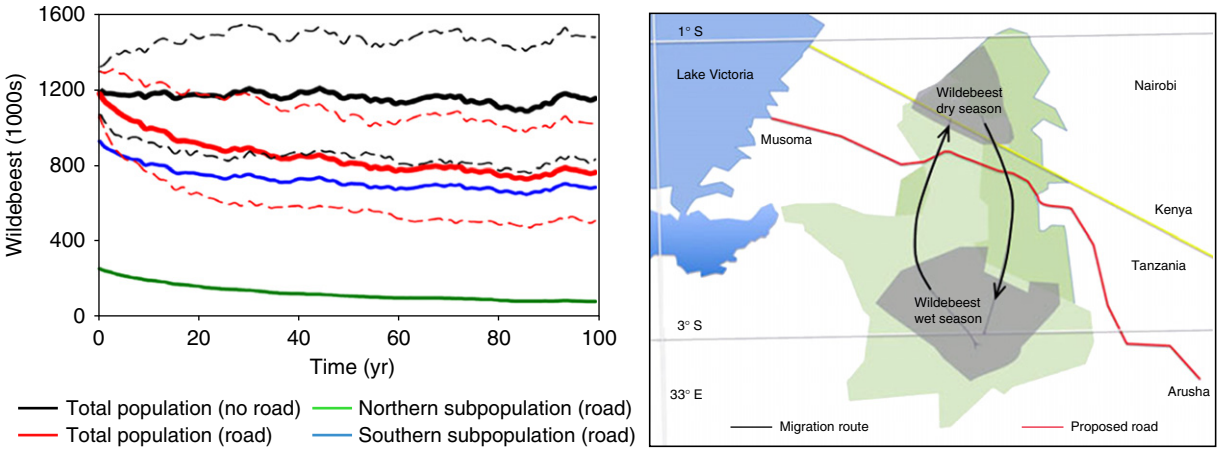


Figure 4 Potential impact of a proposed road on Serengeti wildebeest populations. (a) Results of spatially explicit simulation model on abundance of wildebeest (*Connochaetes* spp.) comparing the landscapes with (red line) and without (black line) proposed road. The road would separate wildebeest into northern and southern populations shown by blue and green lines. Graph reprinted from Holdo RM, Fryxell JM, Sinclair ARE, Dobson A, and Holt RD (2011) Predicted impact of barriers to migration on the Serengeti wildebeest population. *Public Library of Science One* 6: 1–7. Map provided by Stuart Pimm, with permission from PLOS.

Although the overall abundance of required resources is of great importance for populations, the patchiness and spatial arrangement of habitat types have the potential to complicate population dynamics in landscapes. Demographic rates, such as survival and reproduction, may be affected by patch size and shape. Movements between patches can be influenced by the distance between patches and other spatial features. Therefore, an understanding of how population dynamics play out in space is important for research and management.

Landscape Structure and Population Dynamics

Landscape ecologists are interested in spatial patterns. For populations, the overall abundance of habitat and its spatial distribution have a strong influence on the genetic structure and population dynamics of species. The concepts of patch, matrix, and corridor are useful for understanding how populations are influenced by the spatial structure of landscapes (**Figure 5**). For the time being, assume that patches are recognizable as discrete areas of similar habitat and that all habitat patches contain suitable habitat with enough resources to sustain a population. (This assumption of discrete, homogeneous patches will be relaxed later.) The unsuitable habitats found between habitat patches comprise the landscape matrix. Corridors are linear strips of habitat that connect patches and facilitate the movements among patches (**Figure 5**). The number, sizes, shapes, and spatial arrangement of habitat patches, as well as the presence of corridors, impose a spatial structure on the species that occupies them.

Several terms contain the word "structure." The spatial distribution of habitat may be referred to as "patch structure" when applied to a specific habitat type or as "landscape structure" when applied to the spatial pattern of all habitat types in the landscape. The term "population structure" refers to the spatial distribution of breeding groups (i.e., demes) and the rates of exchange of individuals, via emigration and immigration, among these groups. Much of the problem of analyzing population dynamics at the landscape level relates to understanding the influence of landscape (or patch) structure on population structure.

Patches, Matrix, and Corridors

The patch structure of a landscape influences the dynamics of populations living therein because patch structure affects the spatial distribution of individuals and their movements (Saunders *et al.*, 1991). Patch size is one of most important qualities of landscape structure because it is directly related to local population size. Larger patches have a greater abundance of resources and will support a greater N over the long term (i.e., the patch's carrying capacity), whereas smaller patches support fewer individuals. Immigration and emigration are affected by the distance between patches and the intervening landscape matrix. When patches are far from one another, the chance of successful movement between them is reduced. Dispersing animals can suffer mortality from hazards of predation, starvation, or accident while crossing the unsuitable matrix. The dispersal mechanisms of plant seeds may fail to deliver them to distant patches, and dispersing animals may not find a suitable patch.

Corridors connect patches and facilitate interpatch movements. Similarly, small habitat patches (i.e., habitat fragments) scattered throughout the matrix can act as "stepping stones" to promote movement between large patches. Interpatch distances between these fragments are small, and the habitat fragments provide temporary refuge for the dispersing animal before crossing the next gap. Likewise, these fragments may be too small to support a plant population over the long term, but temporary populations can produce seeds capable of colonizing nearby patches. The effectiveness of corridors and stepping stones for movement can be altered by characteristics of the landscape matrix (Baum *et al.*, 2004; Vergara, 2011).

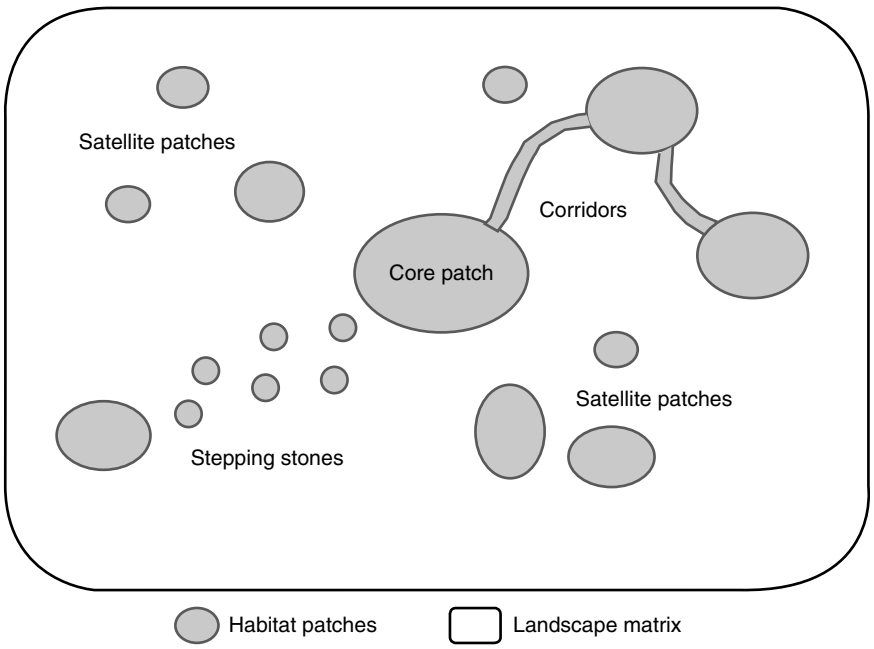


Figure 5 Schematic diagram of patch–matrix–corridor model. In this example, a centrally located core patch is surrounded by smaller satellite patches. Individuals from the population of the core patch disperse to the satellite patches, which support smaller populations that frequently go extinct. Linear corridors or closely spaced habitat fragments (stepping stones) facilitate movement among patches.

Characteristics of the landscape matrix may facilitate or hinder movement. Habitats in the matrix may be harsh or benign relative to the resource needs of individual species. The matrix may contain barriers to movement. The barriers include natural features such as rivers or mountains and human constructed features such as roadways, agricultural fields, and urban areas. The relative ease or difficulty of movement through the landscape matrix is called its permeability. The spaces between habitat patches, filled by the matrix, are called gaps. A species' ability to move across this space is called its gap-crossing ability. Gap-crossing ability and matrix permeability for a given species in a particular landscape depend on the life-history characteristics, habitat needs, and vagility of the species as well as the matrix itself (Chardon *et al.*, 2003; Baguette and Van Dyck, 2007). For example, a bird can readily fly across 100-m wide plowed agricultural field, but that same field is a barrier to a salamander, which risks predation or death by desiccation during an attempt to cross it.

Landscape Context

The suitability and uses of particular habitat patches may be affected by landscape context. The matrix and landscape-level abundance of similar habitats alter the abundance of a species at the patch scale (Radford and Bennett, 2007). When a given habitat type is abundant in the broader landscape, then it is more likely to see occupancy or visits to small patches (Pearson, 1993; Lewis *et al.*, 2011). This effect has been documented in mobile animals, such as mammals, flying insects, and birds, which readily move between patches. Landscape context affects plant population dynamics by altering pollination services and rates of seed predation (Steffan-Dewenter *et al.*, 2001). The abundance and diversity of habitats in the broader matrix affect the probability that pollinators or seed predators will visit a given patch. Some animal species depend on a variety of habitat types during their lives and the ability to move among them. These species may be especially sensitive to landscape context. For example, some amphibians require wetlands for breeding but then disperse to upland habitats after the breeding season (Pillsbury and Miller, 2008). Such species require different habitat types in different seasons or different parts of their life cycles. Therefore, a mix of patch types that provide supplementary or complementary resources (*sensu* Dunning *et al.*, 1992) must be considered when assessing habitat quality and quantity at patch and landscape scales. By altering rates of immigration and emigration, landscape context can affect the genetic diversity and divergence among populations (Blevins *et al.*, 2011).

Quantifying Landscape Structure and Connectivity

Spatial metrics provide a means of quantifying the spatial arrangement of habitat. The most straightforward metrics focus on the statistical distribution of patch sizes. First, the landscape is classified into suitable versus unsuitable habitat (i.e., a binary landscape), and the areas of suitable patches are tallied. Then the "number of patches," "mean patch size," and "size of the largest patch" are useful for comparing landscapes with similar amounts of habitat but having a different spatial arrangement of habitat. Other metrics can measure more complex aspects of spatial pattern such as the shape of patches. These metrics incorporate a ratio of perimeter-to-area for patches. Compact shapes, such as circular or square patches, have a lower perimeter-to-area ratio, whereas elongated patches or those with complex shapes have a greater perimeter-to-area ratio. The metric of "contagion" (e.g., Riitters *et al.*, 1996) measures the intermixing of habitat types, which is correlated with number of patches as well as patch size and shape. There are many spatial metrics available, and many are correlated with each other (Hargis *et al.*, 1998). No single metric will measure all of the important aspects of spatial pattern. The best choices will be those metrics that measure the qualities of landscape structure predicted to influence the population processes under study. One of the most important qualities is habitat connectivity, which measures the ease of moving between habitat patches.

Landscape ecologists have developed a variety of metrics for measuring the connectivity of landscapes (Kindlmann and Burel, 2008). These metrics incorporate the interpatch distances, area of habitat at various distances from the focal habitat, and the presence of stepping stones and corridors (Leidner and Haddad, 2011). A proximity index rates the relative amount and distance to suitable habitat for single patches (Gustafson and Parker, 1994). Deciding which among the available metrics is the best choice for any given species remains challenging. Furthermore, connectivity has been treated as both independent and dependent variables in landscape ecology research (Goodwin, 2003). When considered an independent variable, connectivity was considered a structural aspect of the landscape under study. These studies often compared the performance of populations living in landscapes having different patch structures. When treated as a dependent variable, connectivity was a function of various species responses to pattern in the same landscape. A highly mobile species, such as a bird or large mammal, can easily move between patches, but a sedentary organism or nonflying invertebrate will find the same landscape disconnected (or fragmented) if it cannot readily move among patches. Thus, sensitivity to habitat loss and fragmentation depends on species' habitat needs and life-history traits (Driscoll and Weir, 2005; Sutton and Morgan, 2009).

Landscape Structure, Population Viability, and Conservation

Conservation biology naturally considers the effects of landscape structure on population structure and population viability. Therefore, there is a close and complementary connection between the disciplines of Conservation Biology and Landscape Ecology. The conservation of any species depends on the maintenance of viable populations, and the patch structure of landscapes affects population viability at both the patch and landscape level.

Population Viability in Small Patches

Population viability is strongly related to effective population size (N_e) and genetic diversity, two characteristics that are linked. Small populations suffer from demographic stochasticity and reduced genetic diversity. Populations in landscapes

in which small isolated patches predominate show less genetic diversity overall than populations in unfragmented landscapes (e.g., butterfly (Collier *et al.*, 2010) and toad (Allentoft *et al.*, 2009)). Moreover, population sizes in small patches fluctuate to a greater degree because of demographic stochasticity, which is the fluctuation in N_e, birthrates, and death rates through time. They also fluctuate as a result of environmental stochasticity, which is variation in demographic rates from unpredictable events related to factors such as weather. Such fluctuations create a greater chance of extinction and expose the population to loss of diversity because of genetic drift and population bottlenecks.

The viability of populations in small patches is reduced because of limited genetic diversity. Small populations contain less genetic diversity because there are fewer individuals. In addition, genetic drift and inbreeding further reduce genetic diversity over time. Genetic drift is the random change in allele frequencies because of sampling error in the process of mating. Over the long term, drift ultimately leads to the loss of alleles when their frequencies randomly reach zero. Levels of heterozygosity, a common way of measuring genetic diversity, are positively associated with population growth and persistence (Hanski and Saccheri, 2006). Individuals with higher levels of heterozygosity typically have higher survival and reproductive rates than individuals with less heterozygosity (Coltman *et al.*, 1998). This phenomenon is called heterozygote advantage or heterosis. Heterozygosity declines as alleles are lost and the population becomes fixed for one allele. Similarly, inbreeding reduces heterozygosity among offspring when related individuals breed, and inbred populations have a greater chance of extinction (Saccheri *et al.*, 1998). Inbred offspring have lower levels of heterozygosity and have a greater chance of being homozygous for deleterious recessive alleles. This reduced fitness is termed "inbreeding depression." In small populations, the choice of mates is directly related to N_e, and matings among related individuals will inevitably occur. The primary conclusion is that small populations have lower genetic diversity and stochastic processes of genetic drift, inbreeding, and population bottlenecks will further reduce genetic diversity over time.

The predominance of small patches in fragmented landscapes can affect demography and behaviors. Small populations may suffer reduced population growth rates because of social factors called Allee effects. Allee (1931) suggested that some species have a minimum viable population (MVP) size. Birth and death rates are density-dependent, and recruitment drops to unsustainable levels below a certain N. For example, fertilization rates in flowering plant may depend on visitation rates by pollinating insects, but a minimum density of individual plants is needed to attract the necessary numbers of pollinators (Hackney and McGraw, 2001). When N_e or densities are too low, recruitment rates decline because of low fertilization rates. Likewise, vulnerability to predators can increase in some animal groups. Flock size and herd size are correlated with the ability to detect predators and thereby reduce predation risk for the individual (Lima and Dill, 1990). Allee effects are apparent when survival and reproductive rates are positively correlated with population density and density is below the threshold for density-dependent population regulation (i.e., the carrying capacity). However, care should be taken not to confuse these effects with the negative effects of small population size such as inbreeding (Stephens *et al.*, 1999). In fragmented landscapes, the small patches may support too few individuals needed to accrue the positive benefits of population size and/or density. For animals, patterns of behaviors such as movements can be altered by habitat fragmentation. For example, Mitrovich *et al.* (2009) found differences in the home range overlap and movements and crowding for the coachwhip (*Masticophis flagellum*), a large snake, in small patches relative to large patches. Movements of large highly mobile mammals may be affected. too. Seasonal movement behaviors of white-tail deer (*Odocoileus virginianus*) varied with the mean size and number of patches in a landscape in the midwestern USA (Grovenburg *et al.*, 2011). Deer in landscapes with many large patches tended to be sedentary residents, whereas those in landscapes with small patches were more likely to be migratory.

The genetic and demographic problems of small populations interact synergistically in a phenomenon referred to as an extinction vortex (Gilpin and Soule, 1986). When N_e falls, genetic diversity is lost because of the phenomena of bottlenecks, drift, and inbreeding. Reduced genetic diversity causes drops in population growth, reducing mean fitness because genetic diversity is correlated with the levels heterozygosity in individuals and the population as a whole. In small populations that have lost a significant amount of genetic diversity, low or negative growth rates further reduce N, which results in greater loss of genetic diversity. Moreover, reduced genetic diversity limits the ability of the population to adapt, via natural selection, to changes in its environment. Therefore, small populations are more sensitive to environmental change than large populations having more diversity. This combination of interacting processes eventually threatens the small population with extinction.

What population size qualifies a small population? Although the most accurate answer will depend on the life history of specific species, the "50/500" rule of thumb from conservation biology (Franklin, 1980) provides a useful starting point when detailed life-history information is lacking. This rule states that isolated demes with an effective population size (N_e) greater than 500 will be relatively safe from the dangers of the loss of genetic diversity. These risks progressively increase as N_e declines from 500 toward 50. At $N_e < 50$, all combined forces of the extinction vortex act on the population, and extinction may be inevitable without management intervention. Thus, a $N_e > 50$ is needed for short-term population viability and $N_e > 500$ needed for long-term viability. The "50/500" rule was based on goal of maintaining genetic diversity (Franklin, 1980). Shaffer (1981) outlined the concept known as the minimum viable population. Estimates of a MVP include three parameters: a N_e, a probability of extinction (or persistence), and a specific time period. For example, "An MVP for any given species in any given habitat is the smallest isolated population having a 99% chance of remaining extant for 1000 years despite the foreseeable effects of demographic, environmental, and genetic stochasticity, and natural catastrophes" (Shaffer 1981, p. 132). When incorporating the dangers of demographic and environmental stochasticity, the minimum N_e should be in the range of thousands rather than hundreds (Traill *et al.*, 2010).

Some scientists criticize the notion of general guidelines for MVPs, arguing that the minimum N_e for long-term persistence must be determined on a species-by-species basis (Flather *et al.*, 2011). Population viability and MVP will be affected by the unique combination of species' traits, including its phylogeny, life history, resource needs, habitat availability, population structure and demography, and environmental threats (both natural and human caused) within the context of the landscape where it lives. For landscape-level conservation efforts, it should be more profitable to identify and address the specific causes of population declines (Caughley, 1994) rather than focus on maintaining a minimum number of individuals. Nevertheless, an estimate of MVP can be useful for labeling species as being at risk of extinction, motivating policy makers and managers to take appropriate actions, and setting a target population size for restoration efforts.

Immigration and Rescue Effects

Despite the problems described in the previous section, small populations are frequently observed in nature and seem to persist over the long term. Their persistence can be explained by the fact that these populations are open and connected to other demes in the landscape via the movement of individuals. Population biologists commonly refer to these movements as "immigration" (into) and "emigration" (from) when in reference to a specific deme.

Immigration has several benefits to a small population and some risks. Immigrant individuals can increase genetic diversity if they bring alleles that are absent or rare in the deme. Even infrequent or low levels of immigration are sufficient to restore genetic diversity lost to genetic drift (Slatkin, 1985). Immigrants provide a form of gene flow. The interpatch transport of gametes, such as pollen for flowering plants, is another example of gene flow. Moreover, immigrants provide a form of population growth when local recruitment rates are low (see BIDE model in the section Basic Concepts and Terminology about Population Dynamics). Thus, immigration can periodically rescue a population from the negative effects of demographic stochasticity and loss of genetic diversity, but there are significant risks associated with the movement of individuals. "Outbreeding depression" occurs when immigrants introduce genetic traits that are poorly suited for the local environmental conditions. Offspring expressing these traits will have reduced fitness, and the mean fitness of the population will be lower until natural selection eliminates these traits. If immigration of these traits persists, then the mean fitness and population growth rate will be determined by a balance between immigration and selection. A second risk is the introduction of pathogens and parasites carried by immigrants.

Preserving Habitat at Landscape Level: The SLOSS Debate

Conservation biologists have debated the best strategy for preserving habitat for threatened and endangered species. If the total area of habitat that can be protected is limited, then what is the best scheme for distributing that habitat in space? The essence of this question is whether to protect the same acreage of habitat in "single large or several small" (i.e., the acronym SLOSS) patches. There are advantages and disadvantages for both choices (Diamond *et al.*, 1976). A single large preserve provides habitat to support the largest possible population. If N_e is sufficiently large, then this choice will avoid the problems of small populations discussed previously. However, the danger of a single habitat patch is that some environmental catastrophe such as a hurricane or fire or an introduced virulent pathogen could extirpate the entire species. If the species is spread across several smaller patches sufficiently separated in space, then the chance of a catastrophic event affecting all of them is less than if they are confined to a single patch. For the "several small" strategy to be viable, the demes of respective patches must be large enough to avoid the extinction vortex, or the populations must be open to movement between patches in order to accrue the benefits of immigration. The movement of immigrants may be natural or facilitated by human management in the form of translocations of individuals.

Land management around preserved patches may include the preservation of corridors to connect patches and facilitate movements among them. Corridors likewise have their relative advantages and disadvantages (Haddad *et al.*, 2011), and their effectiveness depends on the natural history of the target species (Gilbert-Norton *et al.*, 2010). Species that are dispersal-limited or have difficulty crossing the landscape matrix receive the greatest benefits corridors. For animals, corridors of low-quality habitat can enhance movement among habitat patches even if populations cannot persist within the corridor itself (Haddad and Tewksbury, 2005).

Landscape Ecology and Metapopulation Dynamics

Most populations in nature have a patchy distribution because of patchiness of their habitat, and population structure is determined by the exchange of individuals (e.g., immigration and emigration) among patches. If the rates of exchange are high and frequent, such as by a highly mobile species, then there is simply one population occupying multiple patches. If exchanges among patches are very infrequent or nonexistent, then each patch has a distinct and separate population. An intermediate level of exchange would include regular movements among patches by dispersing individuals moving from their natal site to another patch. This amount of exchange characterizes metapopulations.

The dynamics of open populations persisting in a patchy landscape often take the form of a metapopulation. A metapopulation is a network of individual demes connected by the processes of immigration and emigration. Metapopulation dynamics exist when species have demes in discrete habitat patches. Each deme has a finite probability of extinction, but extinction events are not synchronized in time across space. Empty patches are common, but none of them are too isolated to be recolonized from an extant deme. Thus, a metapopulation exhibits turnover in the occupancy of patches that is a function of local extinctions and recolonizations (Hanski, 1999). McCullough (1996) provides examples of metapopulation ecology applied to problems of wildlife conservation.

Metapopulation theory has been an important tool for helping ecologists think about population dynamics in fragmented landscapes, but detecting metapopulations and parameterizing metapopulation models from field data is still challenging (Driscoll, 2007). The incidence function model of (Hanski *et al.*, 1996b) has been one of the most useful because it requires only patch occupancy data. By monitoring how many and which patches are occupied over time, ecologists can relate the probability of occupancy to patch size, shape, and isolation. Alternative techniques include analyses of genetic structure to infer the relative isolation of habitat patches and the frequency of immigration among them (Wang *et al.*, 2008).

Sources and Sinks

In the simplest case of population structure, the size of patches and the quality of habitat contained within them is similar, but this case is likely uncommon in nature. In a metapopulation spread over several habitat patches, variation in population growth rates can exist because of patch-specific differences in habitat quality, interspecific interactions, or other ecological factors. Pulliam (1988) described the situation in which some patches have intrinsic rates of increase (r = birthrate − death rate) greater than 0 – that is, local birthrates exceed local death rates. These patches were termed "sources" because local population growth will produce an excess of individuals. The local population growth exceeds the patch's carrying capacity, and these excess individuals must emigrate; if not, then feedbacks from local resource limitation will reduce growth rates. In the same metapopulation, other patches called sinks have $r<0$ because the local birthrate is less than the death rate. If isolated from other populations, extinction is inevitable for sink populations. However, sinks can persist in the context of a metapopulation because excess individuals from source patches can immigrate into sinks. These sinks receive a demographic subsidy and are able to persist because the sum of (births + immigrants − deaths) ≥ 0.

The knowledge of source and sink patches is critical for understanding how a particular metapopulation functions. This understanding is important for landscape-level conservation efforts in which choices are made about the protection (or allowing the destruction) of some fraction of the existing patches. If a critical number of the sources patches are lost, then the entire metapopulation will collapse (Hanski and Ovaskainen, 2002). Unfortunately, patch occupancy alone is not a reliable indicator of the source/sink status of individual patches. The hard work of estimating demographic rates must be undertaken to fully understand this aspect of population dynamics.

Core and Satellite Model

In conservation planning, the "core and satellite model" of habitat protection has been used at the landscape level (Shen *et al.*, 2008; Snep and Ottburg, 2008). This model envisions critical habitat distributed in patches varying in area and quality. Core patches support the largest N_e and populations that persist through time. Core populations may act as demographic sources, and their greatest value to the broader metapopulation is to produce emigrants destined for satellite populations (Runge *et al.*, 2006). Small patches located near the periphery of the core patches support small satellite populations that periodically go extinct. Satellite populations depend on immigrants from the core patches to avoid extinction or for recolonization after local extinction. Satellite populations may be demographic sinks or simply have a small N_e that is vulnerable to the problems of small populations.

Landscape Change, Habitat Fragmentation, Edge Effects, and Extinction Debt

Landscape changes, particularly those caused by human land uses, can alter the abundance and connectivity of suitable habitat. Habitat fragmentation results when habitat loss converts a landscape dominated by large connected patches of habitat to one in which habitat exists in small, disconnected fragments. Fragmentation of native habitats is happening on a global scale (Riiters *et al.*, 2000). This phenomenon affects habitat quality as well as its quantity and spatial distribution. Consequentially, fragmentation affects the spatial structure of populations. The effects of fragmentation can be conceptually separated into the effects of (1) reduced patch size and (2) increased patch isolation. The processes of habitat loss and fragmentation often occur concurrently, but Fahrig (2003) argued that their effects need to be considered independently to best understand the consequences of landscape change on populations and communities.

In fragmented landscapes, the effects of small patch size are confounded with edge effects. Small fragments have more edge, which can alter habitat quality. Small habitat patches have a greater edge-to-interior ratio because of simple geometry. As the area of a circle ($\pi \times \text{radius}^2$) decreases, the perimeter of a circle ($2 \times \pi$ radius) decreases more slowly. Circles and squares are simple, compact shapes in which length and width are approximately equal. Habitat fragments of elongated and complex shapes have a greater amount of perimeter compared to the area. The patch perimeter or "edge" is the boundary between the habitat of the patch and the landscape matrix, and the ecological conditions at the patch edge can be modified by the conditions in the matrix (Matlack, 1993). For example, consider a forest patch in a matrix of agricultural fields. The edges of forest are drier, warmer, windier, and receive more sunlight than locations deep in the interior of the forest patch. These altered conditions are termed "edge effects" and may extend some distance into the forest patch. Edge effects also include differences in the biotic interactions. For example, predation rates may be greater near edges. The "interior habitat" – sites within the patch but some minimum distance from the patch boundary – is considered free of edge effects. Some species are sensitive to the differences between edges and interiors. So-called "edge-sensitive" or "interior-dependent" species avoid edges or experience lower survival and/or reproductive rates in edges. As fragment size decreases, the edge-to-interior ratio increases; edge-sensitive species have less and less interior habitat for use. Below a particular patch size, interior habitat is absent because edge effects penetrate to all regions of the small patch. Therefore, interior-dependent

species are absent from the smallest patches that are dominated by edge-tolerant species (Banks-Leite *et al.*, 2010). Although the example of forest habitat was discussed here, edge effects also exist for grasslands and other nonforest habitats (Renfrew *et al.*, 2005).

The response of populations to the process of habitat loss and fragmentation may lag behind the changes in habitat quantity and quality. For example, some edge-sensitive species may persist in habitat fragments in spite of the loss of interior habitat. This lag occurs because individuals survive in small patches but long-term recruitment rates drop below mortality rates. Thus, small patches are occupied and some reproduction may be occurring, but not enough for long-term population persistence. In addition, the metapopulation dynamics may be disrupted by the overall loss of participating patches or the loss of a critical patch, and the metapopulation collapses (Hanski *et al.*, 1996a). Eventually, these species will go extinct in this landscape, but extinction is delayed by the presence of long-lived individuals and some low levels of reproduction. This delay is termed "extinction debt" (Hanski and Ovaskainen, 2002; Honnay *et al.*, 2005), and it depends on the life history of species affected by habitat fragmentation (Piqueray *et al.*, 2011). Krauss *et al.* (2010) reviewed data for a variety of species living in fragmented grasslands of Europe. They found that long-lived perennial plant species were more likely to exhibit extinction debt than short-lived, mobile species such as butterflies. For species exhibiting extinction debt, their distribution on the landscape is better explained by past habitat patterns than by the distribution and quality of habitat in the contemporary landscape.

Environmental Gradients and Habitat Heterogeneity at Multiple Scales

The concept of source and sink populations (Pulliam, 1988) raised awareness of how variation in demographic performance among patches affects the occupancy of individual patches as well as the landscape-wide distribution of a population. The spatial arrangement of patches is important because sink patches may persist only if they are within the dispersal window of at least one demographic source (Foppen *et al.*, 2000). In the original models of source and sink metapopulations, within-patch conditions were considered homogeneous, but this situation is seldom typical in real landscapes. Spatial heterogeneity within patches may affect demographic rates and population dynamics at fine scales.

For large patches, spatial heterogeneity in habitat quality may create source–sink dynamics among sites within the patch. Environmental heterogeneity can create a mosaic of sites having higher and lower rates of survival, reproduction, and recruitment (Meekins and McCarthy, 2001). Within-patch heterogeneity can be especially high when patches are defined by a coarse-level criterion, such as land-cover type, and fine-scale heterogeneity is ignored. Fine-scale heterogeneity in temperature, moisture, soil nutrients, etc. may be correlated with occurrence of individuals (Palmer, 1990; Fraterrigo *et al.*, 2006) and demographic rates (Meekins and McCarthy, 2001). If habitat quality is strongly correlated with fine-scale edaphic conditions (or some other measure of resource abundance for such food availability), then these conditions may be used to estimate the demographic potential of a given site (Seydack *et al.*, 2000; Ecke *et al.*, 2002; Gonzalez-Megias *et al.*, 2005; Flinn, 2007). Although not separated into discrete patches, source-sink dynamics can occur among the sites within patches. The demographic rates associated with an entire patch will be a function of the average habitat quality of its embedded sites. Landscape ecologists, with their appreciation for issues of scale, are well-equipped to study the relative effects of within-patch, among-patch, and landscape heterogeneity on population dynamics.

Changing patterns of biodiversity in changing landscapes illustrate how spatial patterns at different scales interact with population processes. In portions of Europe and North America, forests are expanding following agricultural abandonment. New forest patches are developing from abandoned farmland via the process of secondary succession; this new habitat is also called postagricultural forest. The spatial arrangement of old forest patches is of great importance for native herbaceous plant species that depend on forest habitats. At the landscape scale, the abundance and proximity of "ancient forests" (not cleared for agriculture) is correlated with the patterns of plant diversity among recovering forest patches (Matlack, 1994a, b). The ancient forests preserved a pool of forest plant species during the time period of heavy agricultural use in the landscape (Pearson *et al.*, 1998). During the period of recovery after the peak of agriculture, these species are able to disperse from the ancient forests to the new forest patches. Moreover, life-history traits influence the establishment and persistence of populations (Vellend, 2003; Flinn and Vellend, 2005) in particular patches. Verheyen *et al.* (2003) studied the distribution of forest herbaceous plant species in these postagricultural forests and their landscapes. They found that recovery after disturbance is a two-stage process in which the dispersal and colonizing ability of a species is initially most important in establishing populations in new forest patches. Next, after a species had colonized a new forest patch, habitat quality and niche specialization of the species influenced abundance and persistence within patches (also Pearson and Fraterrigo, 2011). The distribution of dispersal limited plant species was not correlated with habitat quality at the landscape scale because these species simply could not reach isolated forest patches. In other words, habitat fragmentation negated any effects of varying habitat quality for these postagricultural forests in the initial phase of recovery. Habitat quality became important only if and when a species reached a patch. This process takes many years to play out and will affect forest biodiversity for a century or more of recovery (Vellend *et al.*, 2007).

Continuous Habitat Variability in Landscapes

The patch–matrix–corridor approach to characterizing spatial pattern is useful for landscapes in which patches exist as discrete readily identifiable units, and there are greater differences among patches than within them. However, in many landscapes, habitat variation occurs along environmental gradients that change relatively slowly across space (**Figure 6**). In such a landscape, drawing a definite boundary to denote habitat

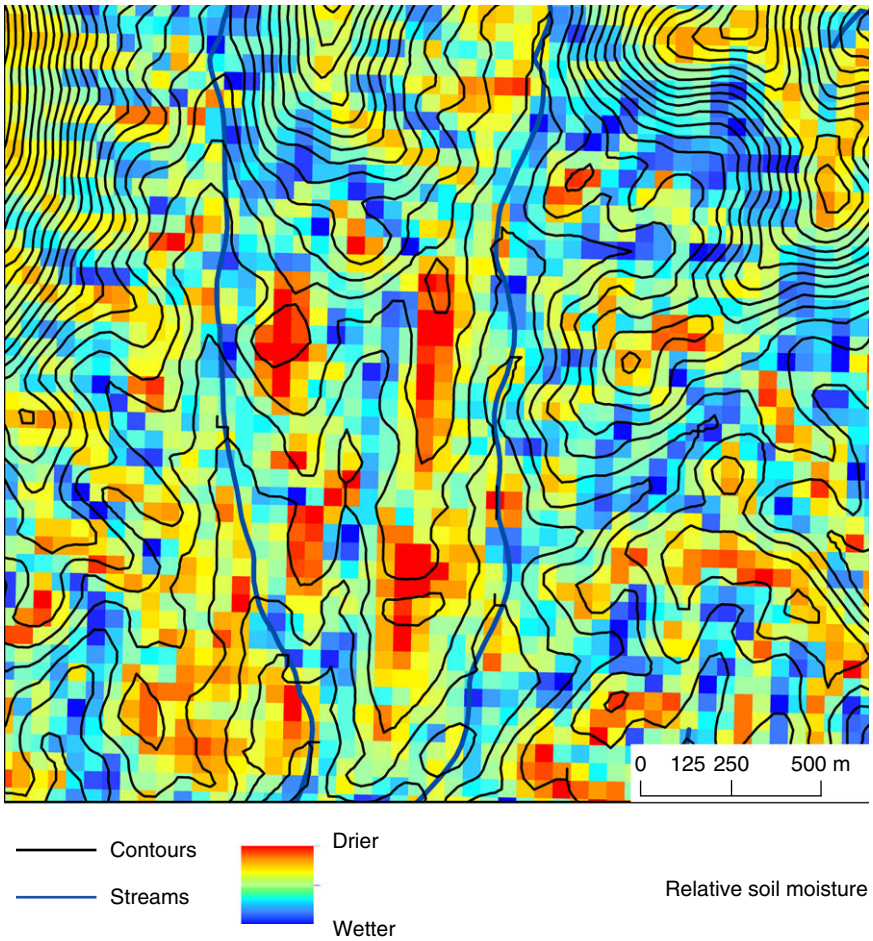


Figure 6 Map of predicted soil moisture as an example of continuous variation in habitat quality. Soil moisture is influenced by topography shown by 10-m contour lines as well as slope, aspect, and surrounding landforms. Drier sites are found on ridges and peaks, whereas wetter sites are located in ravines and valleys. This map was developed using a digital elevation model to calculate a topographic relative moisture index (Parker, 1982). (Reproduced with permission from Simon SA, Collins TK, Kauffman GL, McNab WH, and Ulrey CJ (2005) *Ecological zones in the Southern Appalachians: First Approximation*. Asheville, NC: US Department of Agriculture, Southern Research Station. Forest Service Research Paper SRS-41). Map shows 50-m grid cells.

patches is nearly impossible because suitability varies in a complex and continuous manner across space. In mountainous landscapes, topography creates gradients in temperature, moisture, and light related to elevation, aspect, slope, and terrain shape (**Figure 6**). Instead of maps of discrete patches of suitable and unsuitable habitat, these landscapes are best represented by maps showing continuous variation in habitat quality (Thompson *et al.*, 2006; Franklin, 2009).

Quantifying the spatial pattern of habitats in these landscapes is challenging (Dorner *et al.*, 2002). The statistical characteristics of this continuous variation can be quantified using a range of statistical techniques including blocking, spectral analysis, measures of spatial autocorrelation (Turner *et al.*, 1991), or surface metrics (McGarigal *et al.*, 2009). Patterns of habitat heterogeneity can change between fine and coarse scales, and these analyses can identify the particular spatial scale(s) in which spatial pattern and population processes are linked (Fortin and Dale, 2005). To be useful, the analysis should be approached from an organism-based perspective; that is, landscape structure can be understood by relating environmental gradients to the tolerances and resource needs of individual species (Cushman *et al.*, 2010) and to population processes such as dispersal ability and sizes of home ranges. The influence of environmental gradients in determining species distributions and community composition has long been recognized by the ecologists. However, the theory and analytical techniques for studying populations in these types of landscapes is still under development.

Landscape Structure and Interspecific Interactions

Interspecific interactions are important determinants of population dynamics, and landscape structure can influence these interactions. All species interact with predators, parasites, competitors, and so on as the biotic portion of their environment. In breeding birds of forests, nest predation rates are affected by both landscape-level and within-patch habitat patterns (Lloyd *et al.*, 2005). Predation rates increase as patch sizes decline and the landscape-wide abundance of nonforest

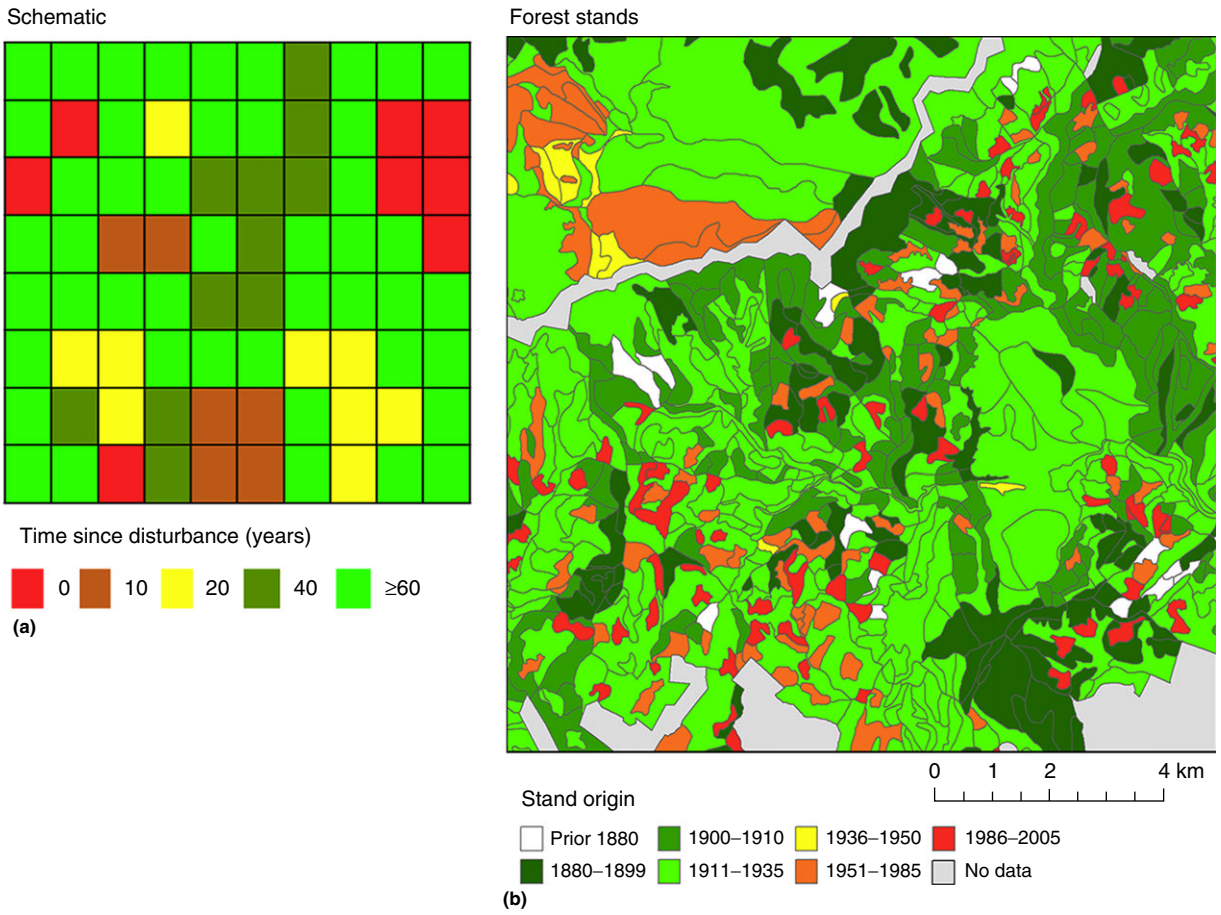


Figure 7 Illustration of a landscape as a shifting mosaic of sites in different stages of recovery from disturbance. In the schematic (a), sites fully recover from disturbance after 60 years of recovery. A landscape in the Pisgah National Forest in North Carolina (b) is a mosaic of forest stands of various ages as measured by time since the year of stand origin. This pattern is a consequence of natural disturbances and human land uses such as timber management.

habitats increases. Moreover, landscape structure influences the movement patterns of different predator species so that the risk of encountering a specific predator species depends on the precise location in the landscape (Bergin *et al.*, 2000). Conversely, spatial complexity can promote coexistence between predators and prey. "Refugia" are habitats (or microhabitats) where preys are relatively safe from predators. When predators are abundant, refugia prevent them from driving their prey to extinction.

Spatial heterogeneity is one of the principal explanations for the coexistence of competing species (Amarasekare, 2003) in community ecology. Coexistence can be achieved by niche partitioning when species specialize on different habitats or resources. Habitat heterogeneity within and among patches promotes coexistence. Spatial structure and heterogeneity through time are also necessary for the coexistence by the competition–colonization trade-off (Tilman, 1994). In this mechanism, competing species differ in their competitive abilities and ability to colonize new habitat patches. The superior competitors have low colonizing abilities, and the better colonizers are poor competitors. Disturbance periodically kills off local populations of the superior competitor, and for a brief time those sites or patches are empty. The superior colonizer (but poor competitor) quickly occupies the empty sites and holds them until the superior competitor arrives. Regular, localized disturbances permit coexistence by continually providing habitat for the poor competitor (e.g., Debout *et al.*, 2009). Coexistence is possible when the landscape is a shifting mosaic of disturbed patches and sites at different stages of recovery from prior disturbances (**Figure 7**). Disturbance and its effects on landscape structure through time are a recurring research theme in landscape ecology.

Summary

Landscape ecology is concerned with the causes and consequences of spatial pattern in nature at multiple spatial scales. Spatial pattern in habitats and resources influences the growth, persistence, and decline of populations. The patch–matrix–corridor model of describing spatial heterogeneity is a useful conceptual model that is directly applicable to populations. Species frequently exhibit a patchy distribution in landscapes as a function of the spatial arrangement of their habitats. This landscape-level pattern (i.e., the landscape structure) determines how populations are organized in

local breeding groups and how frequently individuals move between these groups. This population structure in turn affects population dynamics. Landscape structure, in combination with life-history traits of a species, determines whether that species exists as a single large population, as a metapopulation of demes connected by movement of individuals, or as a collection of separate isolated populations. An understanding of population structure reveals the conservation challenges facing many species. In landscapes in which habitat is highly fragmented, isolated populations in small patches face a number of problems ranging from loss of genetic diversity to edge effects. Thus, the landscape ecology of populations is related to conservation. Some landscapes exhibit continuous variation in habitat quality, and the patch–matrix–corridor model is not directly applicable. Techniques for describing habitat heterogeneity at multiple scales are available, but understanding how this type of variation affects population dynamics is more challenging.

See also: Habitat Loss and Fragmentation. Metapopulations

References

Allee WC (1931) *Animal Aggregations: A Study in General Sociology*. Chicago, IL: University of Chicago Press.

Allentoft M, Siegismund H, Briggs L, and Andersen L (2009) Microsatellite analysis of the natterjack toad (*Bufo calamita*) in Denmark: Populations are islands in a fragmented landscape. *Conservation Genetics* 10: 15–28.

Amarasekare P (2003) Competitive coexistence in spatially structured environments: A synthesis. *Ecology Letters* 6: 1109–1122.

Askins RA, Chavez-Ramirez F, Dale BC, *et al.* (2007) Conservation of grassland birds in North America: Understanding ecological processes in different regions. *Auk/Ornithological Monographs* 124: 1–46.

Baguette M and Van Dyck H (2007) Landscape connectivity and animal behavior: Functional grain as a key determinant for dispersal. *Landscape Ecology* 22: 1117–1129.

Banks-Leite C, Ewers RM, and Metzger J-P (2010) Edge effects as the principal cause of area effects on birds in fragmented secondary forest. *Oikos* 119: 918–926.

Baum KA, Haynes KJ, Dillemuth FP, and Cronin JT (2004) The matrix enhances the effectiveness of corridors and stepping stones. *Ecology* 85: 2671–2676.

Benson TJ, Dinsmore JJ, and Hohman WI (2006) Changes in land cover and breeding bird populations with restoration of riparian habitats in East-central Iowa. *Journal of the Iowa Academy of Science* 113: 10–16.

Bergin TM, Best LB, Freemark KE, and Koehler KJ (2000) Effects of landscape structure on nest predation in roadsides of a midwestern agroecosystem: A multiscale analysis. *Landscape Ecology* 15: 131–143.

Blevins E, Wisely S, and With K (2011) Historical processes and landscape context influence genetic structure in peripheral populations of the collared lizard (*Crotaphytus collaris*). *Landscape Ecology* 26: 1125–1136.

Caughley G (1994) Directions in conservation biology. *Journal of Animal Ecology* 63: 215–244.

Chamberlain DE and Fuller RJ (2001) Contrasting patterns of change in the distribution and abundance of farmland birds in relation to farming systems in lowland Britain. *Global Ecology & Biogeography* 10: 399–409.

Chardon JP, Adriaensen F, and Matthysen E (2003) Incorporating landscape elements into a connectivity measure: A case study for the Speckled wood butterfly (*Pararge aegeria* L.). *Landscape Ecology* 18: 561–573.

Collier N, Gardner M, Adams M, McMahon CR, Benkendorff K, and Mackay DA (2010) Contemporary habitat loss reduces genetic diversity in an ecologically specialized butterfly. *Journal of Biogeography* 37: 1277–1287.

Coltman DW, Bowen WD, and Wright JM (1998) Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. *Proceedings of the Royal Society of London B* 265: 803–809.

Cushman SA, Evans JS, McGarigal K, and Kiesecker JM (2010) *Toward Gleasonian Landscape Ecology: From Communities to Species, From Patches to Pixels*. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. Research Paper RMRS-RP-84.

Debout GDG, Dalecky A, Ngomi AN, and McKey DB (2009) Dynamics of species coexistence: Maintenance of a plant-ant competitive metacommunity. *Oikos* 118: 873–884.

Diamond JM, Terborgh J, Whitcomb RF, *et al.* (1976) Island biogeography and conservation: Strategy and limitations. *Science* 193: 1027–1032.

Dorner B, Lertzman K, and Fall J (2002) Landscape pattern in topographically complex landscapes: Issues and techniques for analysis. *Landscape Ecology* 17: 729–743.

Driscoll DA (2007) How to find a metapopulation. *Canadian Journal of Zoology* 85: 1031–1048.

Driscoll DA and Weir T (2005) Beetle responses to habitat fragmentation depend on ecological traits, habitat condition, and remnant size. *Conservation Biology* 19: 182–194.

Dunning JB, Danielson BJ, and Pulliam HR (1992) Ecological processes that affect populations in complex landscapes. *Oikos* 65: 169–175.

Ecke F, Lofgren O, and Sorlin D (2002) Population dynamics of small mammals in relation to forest age and structural habitat factors in northern Sweden. *Journal of Applied Ecology* 39: 781.

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34: 487–515.

Flather CH, Hayward GD, Beissinger SR, and Stephens PA (2011) Minimum viable populations: Is there a "magic number" for conservation practitioners? *Trends in Ecology and Evolution* 26: 307–316.

Flinn KM (2007) Microsite-limited recruitment controls fern colonization of post-agricultural forests. *Ecology* 88: 3103–3114.

Flinn KM and Vellend M (2005) Recovery of forest plant communities in post-agricultural landscapes. *Frontiers in Ecology and the Environment* 3: 243–250.

Foppen RPB, Chardon JP, and Liefveld W (2000) Understanding the role of sink patches in source-sink metapopulations: Reed warbler in an agricultural landscape. *Conservation Biology* 14: 1881–1892.

Fortin M-J and Dale M (2005) *Spatial Analysis: A Guide for Ecologists*. New York, NY: Cambridge University Press.

Franklin IR (1980) Evolutionary change in small populations. In: Soule ME and Wilcox BA (eds.) *An Evolutionary-Ecological Perspective*, pp. 135–150. Sunderland, MA: Sinauer.

Franklin J (2009) *Mapping Species Distributions: Spatial Inference and Prediction*. New York: Cambridge University Press.

Fraterrigo JM, Turner MG, and Pearson SM (2006) Interactions between past land use, life-history traits and understory spatial heterogeneity. *Landscape Ecology* 21: 777–790.

Gilbert-Norton L, Wilson R, Stevens JR, and Beard KH (2010) A meta-analytic review of corridor effectiveness. *Conservation Biology* 24: 660–668.

Gilpin ME and Soule ME (1986) Minimum viable populations: Processes of species extinction. In: Soule ME (ed.) *The Science of Scarcity and Diversity*, pp. 19–34. Sunderland, MA: Sinauer Associates.

Gonzalez-Megias A, Gomez JM, and Sanchez-Pinero F (2005) Regional dynamics of a patchily distributed herbivore along an altitudinal gradient. *Ecological Entomology* 30: 706–713.

Goodwin BJ (2003) Is landscape connectivity a dependent or independent variable? *Landscape Ecology* 18: 687–699.

Grovenburg TW, Jacques CN, Klaver RW, *et al.* (2011) Influence of landscape characteristics on migration strategies of white-tailed deer. *Journal of Mammalogy* 92: 534–543.

Gustafson EJ and Parker GR (1994) Using an index of habitat patch proximity for landscape design. *Landscape and Urban Planning* 29: 117–130.

Hackney EE and McGraw JB (2001) Experimental demonstration of an Allee effect in American ginseng. *Conservation Biology* 15: 129–136.

Haddad NM, Hudgens B, Damschen EI, *et al.* (2011) Assessing positive and negative ecological effects of corridors. In: Liu J, Hull V, Morzillo AT, and Wiens JA (eds.) *Sources, Sinks, and Sustainability*, pp. 475–503. Cambridge, UK: Cambridge University Press.

Haddad NM and Tewksbury JJ (2005) Low-quality habitat corridors as movement conduits for two butterfly species. *Ecological Applications* 15: 250–257.

Hanski I (1999) *Metapopulation Ecology*. Oxford, UK: Oxford University Press.

Hanski I, Moilanen A, and Gyllenberg M (1996a) Minimum viable metapopulation size. *The American Naturalist* 147: 527–541.

Hanski I, Moilanen A, Pakkala T, and Kuussaari M (1996b) The quantitative incidence function model and persistence of an endangered butterfly metapopulation. *Conservation Biology* 10: 578–590.

Hanski I and Ovaskainen O (2002) Extinction debt at extinction threshold. *Conservation Biology* 16: 666–673.
Hanski I and Saccheri I (2006) Molecular-level variation affects population growth in a butterfly metapopulation. *Public Library of Science Biology* 4: e129.
Hargis CD, Bissonette JA, and David JL (1998) The behavior of landscape metrics commonly used in the study of habitat fragmentation. *Landscape Ecology* 13: 167–186.
Herkert JR (2007) Evidence for a recent Henslow's Sparrow population increase in Illinois. *Journal of Wildlife Management* 71: 1229–1233.
Herkert JR, Reinking DL, Wiedenfeld DA, *et al.* (2003) Effects of prairie fragmentation on the nest success of breeding birds in the Midcontinental United States. *Conservation Biology* 17: 587–594.
Holdo RM, Fryxell JM, Sinclair ARE, Dobson A, and Holt RD (2011) Predicted impact of barriers to migration on the Serengeti wildebeest population. *Public Library of Science One* 6: 1–7.
Honnay O, Jacquemyn H, Bossuyt B, and Hermy M (2005) Forest fragmentation effects on patch occupancy and population viability of herbaceous plant species. *New Phytologist* 166: 723–736.
Jackson HB (2013) Habitat loss and fragmentation. *Encyclopedia of Biodiversity*, 2nd edn. Burlington, MA: Academic Press.
Jewell KJ and Arcese P (2008) Consequences of parasite invasion and land use on the spatial dynamics of host populations. *Journal of Applied Ecology* 45: 1180–1188.
Kindlmann P and Burel F (2008) Connectivity measures: A review. *Landscape Ecology* 23: 879–890.
Krauss J, Bommarco R, Guardiola M, *et al.* (2010) Habitat fragmentation causes immediate and time-delayed biodiversity loss at different trophic levels. *Ecology Letters* 13: 597–605.
Leidner AK and Haddad NM (2011) Combining measures of dispersal to identify conservation strategies in fragmented landscapes. *Conservation Biology* 25: 1022–1031.
Lewis CW, Hodges KE, Koehler GM, and Mills LS (2011) Influence of stand and landscape features on snowshoe hare abundance in fragmented forests. *Journal of Mammalogy* 92: 561–567.
Lima SL and Dill LM (1990) Behavioral decisions made under the risk of predation: A review and prospectus. *Canadian Journal of Zoology* 68: 619–640.
Lloyd P, Martin TE, Roland LR, Langner U, and Hart MM (2005) Linking demographic effects of habitat fragmentation across landscapes to continental source-sink dynamics. *Ecological Applications* 15: 1504–1514.
Matlack GR (1993) Microenvironment variation within and among deciduous forest edge sites in the eastern United States. *Biological Conservation* 66: 185–194.
Matlack GR (1994a) Plant species migration in a mixed-history forest landscape in eastern North America. *Ecology* 75: 1491–1502.
Matlack GR (1994b) Vegetation dynamics of the forest edge – Trends in space and successional time. *The Journal of Ecology* 82: 113–123.
McCullough DR (ed.) (1996) *Metapopulations and Wildlife Conservation*. Washington, DC: Island Press.
McGarigal K, Tagil S, and Cushman S (2009) Surface metrics: An alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology* 24: 433–450.
Meekins JF and McCarthy BC (2001) Effect of environmental variation on the invasive success of a nonindigenous forest herb. *Ecological Applications* 11: 1336–1348.
Mitrovich MJ, Diffendorfer JE, and Fisher RN (2009) Behavioral response of the coachwhip (*Masticophis flagellum*) to habitat fragment size and isolation in an urban landscape. *Journal of Herpetology* 43: 646–656.
Palmer MW (1990) Spatial scale and patterns of vegetation, flora and species richness in hardwood forests of the North Carolina Piedmont. *Coenoses* 5: 89–96.
Parker AJ (1982) The topographic relative moisture index: An approach to soil-moisture assessment in mountain terrain. *Physical Geography* 3: 160–168.
Pearson SM (1993) The spatial extent and relative influence of landscape-level factors on wintering bird populations. *Landscape Ecology* 8: 3–18.
Pearson SM and Fraterrigo JM (2011) Habitat quality, niche breath, temporal stochasticity, and the persistence of populations in heterogeneous landscapes. In: Liu J, Hull V, Morzillo AT, and Wiens JA (eds.) *Sources, Sinks, and Sustainability*, pp. 115–138. Cambridge, UK: Cambridge University Press.
Pearson SM, Smith AB, and Turner MG (1998) Forest patch size, land use and mesic forest herbs in the French Broad River Basin, North Carolina. *Castanea* 63: 382–395.
Pearson SM, Turner MG, and Drake JB (1999) Landscape change and habitat availability in the Southern Appalachian Highlands and Olympic Peninsula. *Ecological Applications* 9: 1288–1304.
Peterjohn BG and Sauer JR (1999) Population status of North American grassland birds from the North American Breeding Bird Survey. In: Vickery PD and Herkert JR (eds.) *Ecology and Conservation of Grassland Birds Of The Western Hemisphere*, pp. 27–44. *Studies in Avian Biology No. 19*.
Pillsbury FC and Miller JR (2008) Habitat and landscape characteristics underlying anuran community structure along an urban-rural gradient. *Ecological Applications* 18: 1107–1118.
Piqueray J, Bisteau E, Cristofoli S, Palm R, Poschlod P, and Mahy G (2011) Plant species extinction debt in a temperate biodiversity hotspot: Community, species and functional traits approaches. *Biological Conservation* 144: 1619–1629.
Price SJ, Marks DR, Howe RW, Hanowski JM, and Niemi GJ (2005) The importance of spatial scale for conservation and assessment of anuran populations in coastal wetlands of the Western Great Lakes, USA. *Landscape Ecology* 20: 441–454.
Pulliam HR (1988) Sources, sinks, and population regulation. *American Naturalist* 132: 652–661.
Radford JQ and Bennett AF (2007) The relative importance of landscape properties for woodland birds in agricultural environments. *Journal of Applied Ecology* 44: 737–747.
Renfrew RB, Ribic CA, Nack JL, and Bollinger EK (2005) Edge avoidance by nesting grassland birds: A futile strategy in a fragmented landscape. *Auk* 122: 618–636.
Ribic CA and Sample DW (2001) Associations of grassland birds with landscape factors in southern Wisconsin. *American Midland Naturalist* 146: 105.
Riiters K, Wickham J, O'Neill R, Jones B, and Smith E (2000) Global-scale patterns of forest fragmentation. *Conservation Ecology* 4: 3.
Riitters KH, O'Neill RV, Wickham JD, and Jones KB (1996) A note on contagion indices for landscape analysis. *Landscape Ecology* 11: 197–202.
Runge Jonathan P, Runge Michael C, and Nichols James D (2006) The role of local populations within a landscape context: Defining and classifying sources and sinks. *The American Naturalist* 167: 925–938.
Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, and Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature* 392: 491.
Saunders DA, Hobbs RJ, and Margules CR (1991) Biological consequences of ecosystem fragmentation: A review. *Conservation Biology* 5: 18–32.
Seydack AHW, Vermeulen C, and Huisamen J (2000) Habitat quality and the decline of an African elephant population: Implications for conservation. *South African Journal of Wildlife Research* 30: 34.
Shaffer ML (1981) Minimum population sizes for species conservation. *BioScience* 31: 131–134.
Shen G, Feng C, Xie Z, Ouyang Z, Li J, and Pascal M (2008) Proposed conservation landscape for giant pandas in the Minshan Mountains, China. *Conservation Biology* 22: 1144–1153.
Simon SA, Collins TK, Kauffman GL, McNab WH, and Ulrey CJ (2005) *Ecological zones in the Southern Appalachians: First Approximation*. Asheville, NC: US Department of Agriculture, Southern Research Station. Forest Service Research Paper SRS-41.
Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology & Systematics* 16: 393–430.
Snep R and Ottburg F (2008) The 'habitat backbone' as strategy to conserve pioneer species in dynamic port habitats: Lessons from the natterjack toad (*Bufo calamita*) in the Port of Antwerp (Belgium). *Landscape Ecology* 23: 1277–1289.
Steffan-Dewenter I, Munzenberg U, and Tscharntke T (2001) Pollination, seed set and seed predation on a landscape scale. *Proceedings of the Royal Society of London: Biological Sciences* 268: 1685–1690.
Stephens PA, Sutherland WJ, and Freckleton RP (1999) What Is the Allee Effect? *Oikos* 87: 185–190.
Sutton FM and Morgan JW (2009) Functional traits and prior abundance explain native plant extirpation in a fragmented woodland landscape. *Journal of Ecology* 97: 718–727.
Thompson LM, van Manen FT, Schlarbaum SE, and DePoy M (2006) A spatial modeling approach to identify potential butternut restoration sites in Mammoth Cave National Park. *Restoration Ecology* 14: 289–296.
Tilman D (1994) Competition and biodiversity in spatially structured habitats. *Ecology* 75: 2–16.
Traill LW, Brook BW, Frankham RR, and Bradshaw CJA (2010) Pragmatic population viability targets in a rapidly changing world. *Biological Conservation* 143: 28–34.
Turner SJ, O'Neill. RV, Conley W, Conley MR, and Humphries HC (1991) Pattern and scale: Statistics for landscape ecology. In: Turner MG and Gardner RH

(eds.) *Quantitative Methods in Landscape Ecology: The Analysis and Interpretation of Landscape Heterogeneity*, pp. 17–49. New York: Springer-Verlag.

Veech JA (2006) A comparison of landscapes occupied by increasing and decreasing populations of grassland birds. *Conservation Biology* 20: 1422–1432.

Vellend M (2003) Habitat loss inhibits recovery of plant diversity as forests regrow. *Ecology* 84: 1158–1164.

Vellend M, Verheyen K, Flinn KM, *et al.* (2007) Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *Journal of Ecology* 95: 565–573.

Vergara P (2011) Matrix-dependent corridor effectiveness and the abundance of forest birds in fragmented landscapes. *Landscape Ecology* 26: 1085–1096.

Verheyen K, Guntenspergen GR, Biesbrouck B, and Hermy M (2003) An integrated analysis of the effects of past land use on forest herb colonization at the landscape scale. *Journal of Ecology* 91: 731–742.

Wang Y-H, Yang K-C, Bridgman C, and Lin L-K (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology* 23: 989–1000.

Wiens JA and Milne BT (1989) Scaling of "landscapes" in landscape ecology, or, landscape ecology from a beetle's perspective. *Landscape Ecology* 3: 87–96.

Wiens JA, Stenseth NC, Horne BV, and Ims RA (1993) Ecological Mechanisms and Landscape Ecology. *Oikos* 66: 369–380.

Landscape Epidemiology

Felicia Keesing, Bard College, Annandale-on-Hudson, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Landscape epidemiology, broadly defined, is the study of spatial variation in disease risk or incidence (Lambin *et al.*, 2010; Ostfeld *et al.*, 2005). The roots of landscape epidemiology stretch back to John Snow's mapping of cholera incidence in London in the 1850s, work which founded the modern science of epidemiology. Snow created maps of the distribution of cholera cases and found that the cases were clustered around a particular water source – the Broad Street pump. When the pump was closed, the number of cholera cases rapidly declined.

Almost 100 years later, a Russian scientist, Evgenii Pavlovsky, coined the term "landscape epidemiology" and extended Snow's concept (Pavlovsky, 1966). Like Snow, Pavlovsky recognized that infectious diseases occur in specific environmental contexts because the distributions of the organisms involved in the transmission of the pathogen are limited to some extent by environmental conditions. Knowledge of these environmental constraints can permit us to characterize, predict, and manage the risk of transmission, just as Snow did. Pavlovsky's innovation was to extend Snow's ideas to the diseases transmitted to humans less directly than cholera.

Cholera is caused by a bacterium, *Vibrio cholerae*, which causes illness when it is ingested by humans. One of the major symptoms of infection with the bacterium is diarrhea, which releases bacteria back into the environment and potentially back to humans through contaminated water. Cholera, then, is a disease directly involving two species – humans and *V. cholerae*. In contrast, "zoonotic diseases" are caused by pathogens that are transmitted to humans from another animal. Hantavirus pulmonary syndrome (HPS) is a notable example. HPS is caused by infection with one of many types of hantaviruses, which are typically shed in the urine or feces of an infected animal, such as a wild rodent; the virus particles can then be inhaled by a human victim. Thus, three species are involved and the disease can only occur where humans, a particular rodent species, and the pathogen coexist in the environment. Some zoonotic diseases are transmitted among hosts by vectors, such as mosquitoes, ticks, or biting flies. These "vector-borne zoonoses" involve at least four species – a human, the pathogen, a wildlife host, and the vector. Pavlovsky was particularly interested in vector-borne zoonotic diseases, how their current and future ranges are shaped by environmental factors, and how the knowledge of these relationships could be used to reduce transmission to humans.

Decades after Pavlovsky, landscape epidemiology is being widely used to ask questions about the spread of pathogens. For example, how does landscape fragmentation influence pathogen spread? At what spatial scales does the presence of resistant species provide a buffer against the spread of pathogens? How can landscapes be designed to control the spread of diseases of humans, wild and domesticated animals, and plants?

The modern study of landscape epidemiology can be divided into two broad research foci. In one approach, investigators study the effects of actual geographic entities on diseases (i.e., map-based approaches); in the other, they utilize abstract geographic entities (i.e., modeling-based approaches). Map-based approaches can be further divided into those that generate an understanding of how spatial patterns of disease change dynamically through time, and those that consider static patterns in disease based on the distribution of hosts, vectors, or disease incidence.

Dynamic Map-Based Approaches: Analyzing Patterns Through Space and Time

Many epidemiologists are interested in understanding the processes that cause spatial patterns of disease risk or incidence to change through time. To address this issue, researchers incorporate dynamic processes (e.g., demographic, genetic, or behavioral changes) into spatiotemporal models to explain or predict epidemiological patterns. Dynamic approaches are exemplified by a recent study of the spatiotemporal patterns of human infection with rotavirus, a pathogen that infects people worldwide (Pitzer *et al.*, 2010). Data collected for 11 years on patients with rotavirus were used to develop alternative models of epidemic patterns in USA. The model with the best fit to the data incorporated differences in birth rates among states, and this model explained geographic shifts in outbreaks over time – states with the highest birth rates had earlier seasonal onset of rotavirus epidemics.

Dynamic spatiotemporal analyses require relatively large amounts of data both in space and in time (but see Tildesley *et al.*, 2010). Data of this quality tend to be available for diseases that are well-known infections of people, such as measles and influenza, or of their domesticated plants and animals. For these types of diseases, sophisticated mathematical analyses can provide useful guidance for the development of prevention strategies. Results of such studies can also reveal patterns applicable to less familiar diseases. For example, an analysis of data from 14 years and 850,000 infections with dengue hemorrhagic fever (DHF) in Thailand showed that cases moved in a traveling wave emanating from Bangkok, the largest city in the country, out to the countryside at 150 km per month (Cummings *et al.*, 2004). This information could guide interventions for DHF, and also for less well-documented infections in the region.

Static Map-Based Approaches: Using Current Distributions to Predict the Future

In contrast to the dynamic approaches, static approaches require fewer data and less sophisticated mathematical models. They are useful for characterizing spatial variation in current risk and for predicting the underlying causes of that variation.

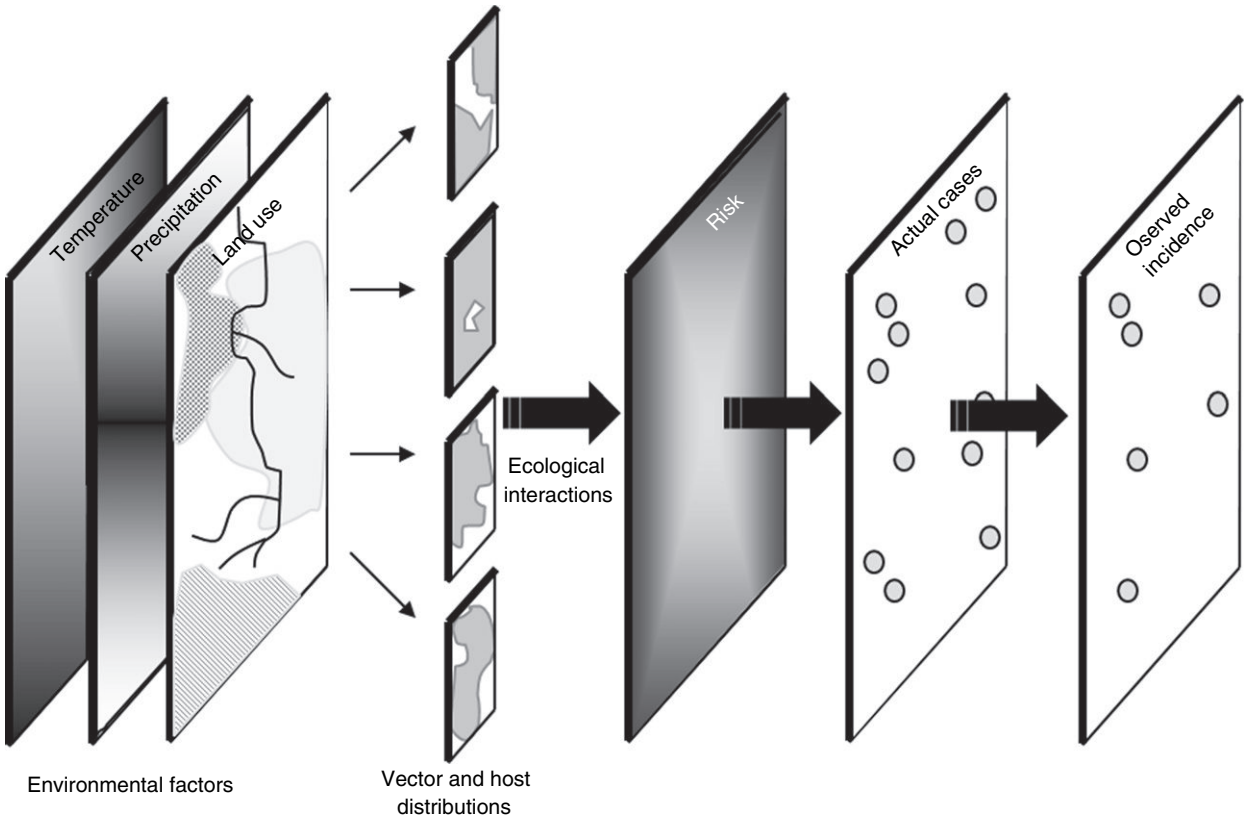


Figure 1 Conceptual model of the relationship between environmental factors that influence disease and observed incidence of that disease in humans. Underlying environmental factors (e.g., temperature, precipitation, land use, and soil type) can influence the distribution of vectors and hosts for a particular disease. These distributions influence the risk to humans of coming into contact with the pathogen that causes the disease. Variation in this risk is one factor that determines the actual number of cases of human disease. Because of differences in surveillance and reporting, the observed incidence in humans is likely to be a subset of actual cases. Modified from Ostfeld RS, Glass GE, and Keesing F (2005) Spatial epidemiology: An emerging (or re-emerging) discipline. *Trends in Ecology & Evolution* 20: 328–336.

The typical approach is to create a map of the distribution of a host or vector and then to use remote-sensing data and geographic information systems to identify underlying geographic variables that correlate with that distribution (**Figure 1**). Once the variables that are most strongly correlated with the distribution are identified, the values of these variables can be projected onto maps outside the spatial limits of the original distribution. Alternatively, the values of these variables can be projected forward in time and used to predict future distributions.

This approach is typified by a study of the distribution of the incidence of West Nile Virus (WNV) encephalitis shortly after its emergence in the New York City area (Brownstein *et al.*, 2002). WNV encephalitis is caused by WNV, which is transmitted to hosts through the bite of any of several species of infected mosquitoes. In this study, human cases of the disease were mapped and found to be correlated with the "normalized difference vegetation index" (NDVI). NDVI is an index of the amount of green vegetation in an area and can be readily obtained from satellite images using remote-sensing technology. The range of NDVI values correlated with human cases was projected onto areas where human cases had not yet been detected. The ability of the variables to predict future incidence in humans was tested by comparing NDVI values to data on the locations of infected mosquitoes; 72% of locations with infected mosquitoes were identified as high risk based on their NDVI values. With this validation, NDVI values could then be used to predict likely areas for future human cases.

Using Data on Incidence

The use of disease incidence data to construct risk maps, as in the WNV example, has several advantages. Incidence data are often more easily obtainable than data on hosts or vectors. Incidence data also integrate a suite of environmental risk factors; if a case occurs, the environmental requirements for host, pathogen, and vector (if relevant) must be met. However, incidence data also present several challenges. First, incidence might not be correlated with risk. This would be the case if, for example, preventative measures (e.g., bednets and repellents) were implemented to reduce the incidence in areas of risk. In such cases, the use of incidence data would lead to an underestimation of areas of potential risk. Incidence data might also be decoupled from risk if the disease system is not in equilibrium, for example, if it is emerging in a new area. In the WNV example, most cases were restricted to a narrow geographic region near the presumed area of emergence of the virus in USA (Brownstein *et al.*, 2002). As a result, only this

area was assessed for fit with predictor variables. If a wider geographic region had been used, the analysis might have identified incorrect predictor variables. Another challenge of using incidence data is that reporting standards are often not uniform across geographic boundaries.

Using Data on Hosts

An alternative to using data on cases of a disease is to use data on the distribution of hosts for the pathogen that causes it. A good example of this type of approach comes from research on appropriate locations for the reintroduction of wild boars (*Sus scrofa*) to Denmark (Fernández *et al.*, 2006). The plan to reintroduce boars is a response to conservation concerns following their extirpation, but the planned reintroduction also raises concerns about possible economic consequences. Boars are reservoirs for the virus that causes classical swine fever, which can be transmitted from wild boars to pigs at commercial pig farms. The researchers predicted boar population sizes at candidate reintroduction sites based on demographic parameters and habitat suitability and then estimated risk to pigs based on the number and types of farms in the area. Based on the underlying model of boar populations, many of the sites that were most suitable for boar reintroduction also created high risk for viral transmission to pigs. Successful reintroduction, therefore, was predicted to require active efforts to minimize contacts between boars and pigs throughout the areas of successful reintroduction.

Host-based approaches have also been used for plant diseases. For example, a recent study assessed the consequences of land-use changes in California forests on the abundance of hosts for *Phytophthora ramorum*, a fungus that causes Sudden Oak Death (Meentemeyer *et al.*, 2011). Aerial images of the study areas from 1942 were compared with images from 2000 to determine changes in suitable habitat for known host plants. Areas that had experienced the greatest increase in woodland (host) habitat over that time period were found to have higher current incidence of disease.

Using Data on Vectors

For vector-borne diseases, data on the distribution of vectors can be used to approximate disease risk. The underlying principle behind this approach is that the presence of the vector is necessary for transmission to take place. However, a major limitation of this approach is that the presence of the vector is not sufficient for transmission – the pathogen must be present as well. Thus, vector mapping can determine areas of no risk (because vectors are absent) and of potential risk (because vectors are present but not necessarily infected), but this method alone cannot determine actual risk.

Many studies of this type are concerned only with the presence or absence of the vector at a given location. A recent investigation of the distribution of blacklegged ticks (*Ixodes scapularis*), which are the vector for a suite of zoonotic diseases, tracked not only the presence of ticks at a suite of locations but also their abundance (Diuk-Wasser *et al.*, 2010). Based on a statistical model parameterized by climatic and habitat characteristics, the best predictors of tick abundance were altitude, vapor pressure deficit, and the presence of nearby sites with abundant ticks. The model was tested by comparing its predictions to sites where tick abundance was known but had not been included in the development of the model. Based on strong concordance between predicted and observed abundance, the model could be used to predict areas that could see growth in tick populations because they contain habitat highly suitable for ticks, but currently do not harbor abundant tick populations.

All three of these approaches – incidence based, host based, and vector based – share a common limitation. All use patterns to determine correlations with predictor variables, but correlations cannot establish cause-and-effect relationships. Nevertheless, these approaches can allow rapid assessment of potential risk.

Mapless Approaches

Abstract spatial models of disease incidence or risk do not rely on explicit spatial placement on maps of entities involved in transmission. These models can be useful for describing patterns, developing theory, and generating predictions that can be tested in real landscapes. There are four main approaches – patch models, distance-transmission models, multigroup models, and network models (Riley, 2007). These models vary in the degree of resolution with which they track transmission events. In patch models, individuals occur in patches (e.g., towns) and patches are connected to other patches to varying degrees. All individuals within a patch are assumed to have the same transmission status – for example, all individuals in the patch are susceptible or all are infected. At the other extreme, in network models, each individual is connected to other individuals to varying degrees, and transmission can only occur between connected individuals.

The choice of the appropriate modeling approach depends largely on the characteristics of the underlying disease system being modeled. For example, a respiratory pathogen that requires close contact between individuals for transmission to occur might best be modeled by a network that takes account of frequency of contact even within a household. In contrast, for a highly contagious pathogen, a patch model might be sufficient because membership in a patch would be likely to result in transmission.

Frontiers in Landscape Epidemiology

Several directions seem particularly promising for the future study of landscape epidemiology. First, there is still much to learn about how the configuration and composition of landscape elements influence diseases. Patch size, shape, and connectivity, for example, have been shown to influence many ecological processes in the study of "landscape ecology," but their effects on disease are less well studied. A major research question in landscape epidemiology is how and when local biotic and abiotic conditions predict disease risk and incidence, and how and when the larger landscape context overrides these local influences (Ostfeld *et al.*, 2005).

Landscape epidemiology is also likely to benefit from "landscape genetics," which takes advantage of burgeoning access to genetic information (Biek and Real, 2010). Knowledge of the genetics of hosts and pathogens through space and time permits a new level of detail in studying disease patterns. For example, amphibian populations are being devastated worldwide by a chytrid fungus, *Batrachochytrium dendrobatidis*. Genetic analyses of *B. dendrobatidis* imply a single genetic origin of the fungus, which has since spread globally. This suggests that human dispersal of the pathogen is a more likely explanation for its spread than factors related solely to amphibian ecology (James *et al.*, 2009; Biek and Real, 2010). Detailed genetic information can also be used to reveal spatial patterns and processes in dynamic analyses. A study utilizing more than 1000 complete genomes of human influenza virus revealed that variants in temperate areas are more closely related to those in other temperate areas in any given year than they are to the variant from the same place the previous year (Rambaut *et al.*, 2009). This suggests a common tropical source that seeds temperate regions each year. A recent study of the spread of chronic wasting disease (CWD) in deer provides a final example of the potential of landscape genetic approaches. Researchers used genetic markers to estimate gene flow among deer populations in Wisconsin and then determined geographic barriers preventing gene flow (Blanchong *et al.*, 2009). The area connected by the least gene flow also had the lowest incidence of CWD, suggesting that (1) the same geographic barrier preventing gene flow (a river) also prevented contact that led to transmission, and (2) host gene flow could be a useful index by which the potential for spatial spread of infections is measured, especially when infection itself is hard to detect.

A final new development comes from "geographic profiling," an approach that was originally developed for use in criminology. As applied to landscape epidemiology, this statistical technique relies on two underlying assumptions: (1) transmission declines with distance from an infected host ("distance-decay"), and (2) if suitable habitats are distributed randomly in the landscape, habitats suitable for transmission are more abundant further from the infected host ("buffer zone") (Le Comber *et al.*, 2011). Using these two assumptions, a geoprofile is created, characterizing possible locations of the infected host or vector within the study area. In a study of malaria in Cairo, Egypt in 2001–2004, a geoprofile was created from human cases. The profile correctly identified six of eight pools with breeding mosquito vectors and was more effective than other techniques that rely on case data to make spatial inferences. Geoprofiling may thus be a productive way to identify important disease agents rapidly from relatively little data.

Summary

Pathogen transmission is an inherently spatial process, requiring contact in space and time between an uninfected host and an infectious source (Killilea *et al.*, 2008). As a result, pathogen transmission can be studied in one direction by identifying where (and perhaps when) the elements necessary for transmission cooccur and then predicting the likely future incidence. In the other direction, data on the incidence of a disease can be used to work backward to predict where (and perhaps when) the elements for transmission coincided.

Because all pathogen transmission has a spatial component, landscape epidemiology could, in principle, encompass virtually all studies of transmission. A more restrictive – and potentially more useful – definition might focus on questions of scale, as has been done for the field of "landscape ecology." At what scale, for example, are the processes affecting transmission determined? Do the processes that affect transmission patterns occur at the same scale at which the patterns are detected, or at a different one? It is in the pursuit of answers to questions like these that the full richness of landscape epidemiology will be achieved.

Appendix

List of Courses

1. Landscape Ecology
2. Disease Ecology
3. Geography

See also: Biodiversity and Human Health. Biogeographical Models. Countryside Biogeography. Habitat Loss and Fragmentation. Landscape Corridors. Landscape Ecology and Population Dynamics. Landscape Genetics. Landscape Modeling. Parasitism. Remote Sensing and Image Processing

References

Biek R and Real LA (2010) The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* 19(17): 3515–3531.

Blanchong JA, Samuel MD, Scribner KT, *et al.* (2009) Landscape genetics and the spatial distribution of chronic wasting disease. *Biology Letters* 4: 130–133.

Brownstein JS, Rosen H, Purdy D, *et al.* (2002) Spatial analysis of West Nile virus: Rapid risk assessment of an introduced vector-borne zoonosis. *Vector Borne Zoonotic Diseases* 2(3): 157–164.

Cummings DAT, Irizarry RA, Huang NE, Endy TP, Nisalak A, and Ungchusak K (2004) Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature* 427: 2–5.

Diuk-Wasser M, Vourc'h G, Cislo P, *et al.* (2010) Field and climate-based model for predicting the density of host-seeking nymphal *Ixodes scapularis*, an important vector of tick-borne disease agents in the eastern United States. *Global Ecology and Biogeography* 19: 504–514.

Fernández N, Kramer-Schadt S, and Thulke H-H (2006) Viability and risk assessment in species restoration: Planning reintroductions for the wild boar, a potential disease reservoir. *Ecology and Society* 11.

James TY, Litvintseva AP, Vilgalys R, *et al.* (2009) Rapid global expansion of the fungal disease chytridiomycosis into declining and healthy amphibian populations. *Public Library of Science Pathogens* 5: e1000458.

Killilea ME, Swei A, Lane RS, Briggs CJ, and Ostfeld RS (2008) Spatial dynamics of lyme disease: A review. *EcoHealth* 5: 167–195.

Lambin EF, Tran A, Vanwambeke SO, Linard C, and Soti V (2010) Pathogenic landscapes: Interactions between land, people, disease vectors, and their animal hosts. *International Journal of Health Geographics* 9: 54.

Le Comber SC, Rossmo DK, Hassan AN, Fuller DO, and Beier JC (2011) Geographic profiling as a novel spatial tool for targeting infectious disease control. *International Journal of Health Geographics* 10: 35.

Meentemeyer RK, Rank NE, Anacker BL, Rizzo DM, Hall J, and Cushman JH (2011) Influence of land-cover change on the spread of an invasive forest pathogen. *Ecological Applications* 18: 159–171.

Ostfeld RS, Glass GE, and Keesing F (2005) Spatial epidemiology: An emerging (or re-emerging) discipline. *Trends in Ecology & Evolution* 20: 328–336.

Pavlovsky E (1966) *Natural Nidality of Transmissible Diseases with Special Reference to the Landscape Ecology of Zooanthroponses*. Urbana, IL: University of Illinois Press.

Pitzer VE, Viboud C, Simonsen L, *et al.* (2010) Demographic Variability, Vaccination, and the Spatiotemporal Dynamics of Rotavirus Epidemics. *Science* 325: 290–294.

Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, and Holmes EC (2009) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453: 615–619.

Riley S (2007) Large-scale spatial-transmission models of infectious disease. *Science* 316: 1298–1301.

Tildesley MJ, House TA, Bruhn MC, *et al.* (2010) Impact of spatial clustering on disease transmission and optimal control. *Proceedings of the National Academy of Sciences of the United States of America* 107: 1041–1046.

Landscape Genetics

Andrew Storfer, Washington State University, Pullman, WA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Analysis of molecular variance Similar to the statistical analysis of variance (ANOVA), but involves partitioning of molecular genetic data in different groups (e.g., within and among populations) based on hierarchical sampling procedures.

Extent Generally used in landscape ecology, referring to the overall size of a study area.

Functional connectivity Describes the processes of dispersal and gene flow of organisms among habitat patches.

Gene flow Refers to the movement of organisms among populations or locales followed by successful breeding event(s), resulting in the transfer of alleles from one population to another.

Grain Generally used in landscape ecology, referring to the spatial scale of individual observations.

Least-cost path A line (vector) of 'least resistance' is generated between two observation points. The resistance refers to a resistance surface, or cost grid in a landscape matrix, which is parameterized according to some understanding of species' dispersal habits. The least-cost distance can be expressed in terms of geographic distance or total resistance along this path.

Mantel test A statistical test of the correlation between two matrices; in landscape genetics, these are a geographic and genetic distance matrix. Significance of the correlation is determined by randomly permuting the values in the rows and columns of the matrices to create a distribution of values to which the observed value can be compared.

Spatial autocorrelation Observations are not independent identically distributed, but are correlated over some distance in space. Positive autocorrelation indicates that nearby values are similar; negative autocorrelation indicates that nearby values are dissimilar.

Structural connectivity Generally refers to the distance between habitat patches and the nature of the intervening landscape matrix.

Tessellation A pattern of nonoverlapping polygons is used to partition the area leaving no gaps. In landscape genetics, tessellation is generally used to create polygons around each point where genotypes are collected.

Introduction

An increasing goal of ecology and evolutionary biology is to understand spatial patterns of genetic variation. However, until recently there have been both technological and analytical gaps between the fields of landscape ecology and population genetics to test the influence of spatial processes on population genetic structure (Manel *et al.*, 2003). That is, whereas landscape ecology improved greatly in its ability to understand *structural connectivity* of landscapes with advances in geographic information systems (GIS) and remote sensing, understanding *functional connectivity* among biological populations remained a challenge because of difficulties associated with accurately estimating migration rates (e.g., low recapture probabilities or estimating breeding without dispersal; Holderegger and Wagner, 2008). However, our ability to obtain high resolution genetic data has dramatically increased, but population genetics had been focused on estimating functional connectivity (i.e., via estimating gene flow, which incorporates dispersal *and* breeding among populations), with little regard to structural connectivity of the intervening habitat matrix between populations. Landscape genetics emerged as an interdisciplinary analytical framework to bridge this gap (Holderegger and Wagner, 2008; Manel *et al.*, 2003) by drawing on analysis methods from landscape ecology, spatial statistics, geography, and population genetics to "…explicitly quantify the effects of landscape composition, configuration and matrix quality on gene flow and spatial genetic variation" (Storfer *et al.*, 2007).

Whereas population genetics studies had previously been limited in spatial inference to tests of isolation-by-distance (IBD), landscape genetics studies test explicitly the relative influence of landscape and environmental features on gene flow, genetic discontinuities (Guillot *et al.*, 2008) and genetic population structure (Manel *et al.*, 2003). This area of inquiry has met with early success, as landscape genetic models generally provide a more complete picture of population genetic structure by typically explaining a significantly higher proportion of variation in gene flow than classic, straight-line Euclidean IBD models (**Figure 1**; Storfer *et al.*, 2010). Further, recent simulation studies have shown that a variety of landscape genetics techniques were able to detect the simulated scenario with relatively high power and accuracy (Balkenhol *et al.*, 2009a; Cushman and Landguth, 2010; Murphy *et al.*, 2008). As a result, understanding landscape effects on functional (genetic) connectivity can provide insight into fundamental biological processes such as: metapopulation dynamics, (ecological) speciation, and ultimately limits to species' geographic distributions.

The integrative approach of landscape genetics has tested both basic and applied research questions, including: identifying barriers to dispersal (Guillot *et al.*, 2008; Latch *et al.*, 2008; Manni *et al.*, 2004), modeling functional connectivity across different spatial scales (Murphy *et al.*, 2010; Rasic and Keyghobadi, 2012), inferring the effect of landscape change

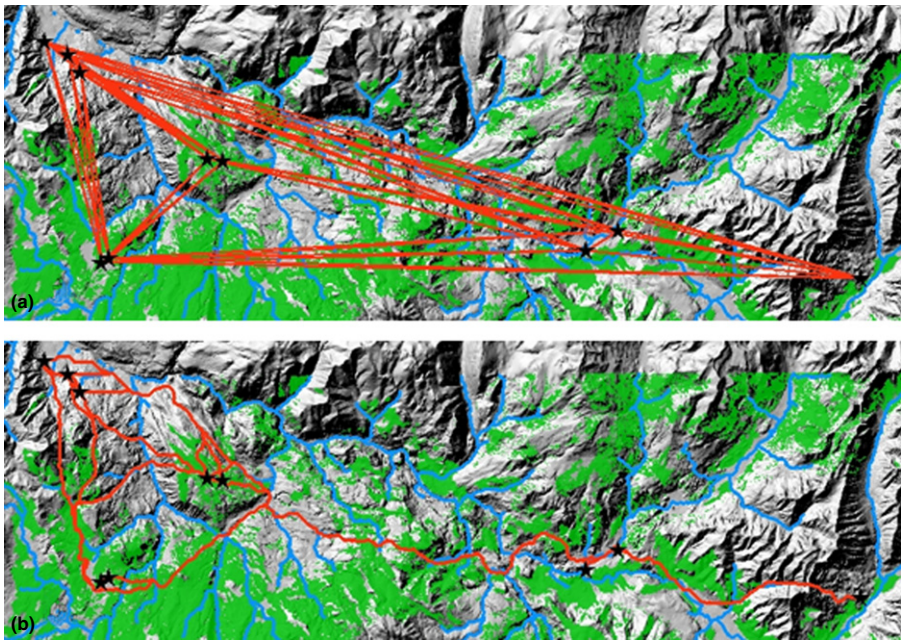


Figure 1 Comparison of straight-line Euclidean distance model (a) with a landscape model that minimizes dispersal costs through wetlands (b). (a) A topographic map of Northern Yellowstone National Park. Black stars represent sampling localities. Green represents forest cover and blue represents steams. Red lines are Euclidean (i.e., straight line map) distances between sites. (b) Same as (a), except red lines are least-cost distances that minimize costs of paths through wetland areas.

(Wagner *et al.*, 2006; Wang *et al.*, 2008), estimating source–sink dynamics (Lowe *et al.*, 2006; Martinez-Solano and Gonzalez, 2008), predicting spread of disease (Biek and Real, 2010), testing whether current spatial genetic patterns are correlated with past landscape change (Spear and Storfer, 2008), and estimating the distribution of adaptive genetic variation across habitat gradients (Joost *et al.*, 2007; Manel *et al.*, 2010a; Segelbacher *et al.*, 2010).

Indeed, due to the wide variety of applications that landscape genetics offers, interest has burgeoned in recent years, and the number of publications is increasing. Researchers new to landscape genetics will quickly note that there are numerous ways to design a study, as well as multiple analytical frameworks and methods. As with any rapidly growing field that integrates previously disparate (sub)-disciplines, landscape genetics has faced challenges with: interdisciplinary communication, the integration of rapidly developing new analytical techniques and software, as well as, integrating and interpreting multiple analysis methods in a single study (Balkenhol *et al.*, 2009b; Holderegger and Wagner, 2008; Storfer *et al.*, 2007). There are several other insightful reviews that have summarized the field as well as its challenges (e.g., Balkenhol *et al.*, 2009a; Holderegger and Wagner, 2006, 2008; see a special issue of *Molecular Ecology* published in September, 2010). As such this review is not meant to be exhaustive, but rather to complement previous reviews by: (1) emphasizing the importance of having a robust study design; (2) discussing some of the most prevalent analytical frameworks; (3) summarizing the major insights that have come from landscape genetics studies to date; and, (4) highlighting some of the current and future challenges facing the discipline.

Study Design and Sampling

A recent review showed that only about one-third of landscape genetics studies had a clear sampling design (Storfer *et al.*, 2010). Perhaps this is not surprising given the relative youth of landscape genetics as a field and the fact that population genetics studies historically tended to rely on opportunistically collected samples (e.g., those obtained along roadsides; Storfer *et al.*, 2007). Because analyses typically tested IBD, sampling schemes tended to be hierarchical, focusing on sampling at localities ranging in (Euclidean) distance from one another. At the time, these sampling designs were sufficient to estimate genetic neighborhood sizes by using, for example, Analysis of Molecular Variance (or AMOVA; Peakall and Smouse, 2006) to partition genetic variance within and among individuals, populations and larger designations. Tests of IBD were often carried out using Mantel tests to determine whether there were correlations between a pairwise genetic distance matrix (e.g., F_{ST} among discrete sampling localities) and a pairwise geographic (Euclidean) distance matrix (Foll and Gaggiotti, 2006).

Landscape genetics studies necessitate a dramatically different sampling approach because the goal is to test explicitly how landscape characteristics influence gene flow (Anderson *et al.*, 2010; Guillot *et al.*, 2009; Storfer *et al.*, 2007). That is, in contrast to traditional population genetics studies, landscape genetics studies are *explicitly spatial* and incorporate landscape data from, for example, GIS, remote sensing and weather stations in addition to genetic data such as allele frequencies (Manel *et al.*, 2003; Spear *et al.*, 2010; Storfer *et al.*, 2007, 2010). As a consequence, sampling designs should ideally be based on knowledge of the dispersal ecology of the study

organism to guide the spatial scale of the study area, as well as the resolution at which GIS and other landscape layer data are sampled (Anderson *et al.*, 2010). Ecological data should include data on home range sizes, migration distances and geographic estimates of population size to guide both the *grain* and *extent* of spatial sampling scales. The grain size at which landscape data are collected should be smaller (e.g., remote sensing data in the USA often is at 30 m resolution) than estimated home range sizes or dispersal distances (Anderson *et al.*, 2010; Fortin and Dale, 2005). Sampling at a grain size that is too large or small may result in failure to capture the range of variation in an environmental (e.g., precipitation) or landscape variable (e.g., elevation), resulting in over- or under-estimation of the influence of that variable on gene flow (Storfer *et al.*, 2007). The extent of the overall study area should be at least as large as the anticipated spatial population size, preferentially with a buffer included (Anderson *et al.*, 2010) to minimize the possibility that gene flow from outside the study area could overwhelm that which occurrs within the study area (Cushman and Landguth, 2010).

Ecological data of habitat preferences and aversions can also be quite useful for guiding which GIS layers will be most relevant for the study organism. For example, habitat suitability modeling can inform parameterization of GIS layers, such as for *least cost path* analysis (**Figure 1**) to assess landscape effects on genetic structure (Wang *et al.*, 2008). Generally, researchers should strive to capture the range of variation in landscape features that are hypothesized to be important, and consequently stratified sampling designs are generally preferred (Storfer *et al.*, 2007). However, whether the most appropriate sampling design is systematic or random may depend on the distribution of the study organism; for example, systematic sampling is most appropriate for species with relatively even and continuous distributions (Harrison and Dunn, 1993; for a detailed discussion on sampling design considerations, see Storfer *et al.* (2007)). Nonetheless, if ecological data on dispersal and habitat use are lacking, data on closely related species can be used to guide sampling grain, extent and landscape variables to be used in statistical models.

Once a sampling strategy has been designed, researchers typically collect geo-referenced (e.g., by using Global Positioning System (GPS)) tissue samples and commonly use hypervariable nuclear deoxyribonucleic acid (DNA) markers (most commonly microsatellites or amplified fragment length polymorphisms (AFLPs); Storfer *et al.*, 2010) across multiple loci to generate a distribution of allele frequencies. Then, usually a genetic distance matrix is computed among sampling localities using analogs of Wright's F-statistics (Wright, 1951) or other distance measures, including chord distance (D_c; Cavalli-Sforza and Edwards, 1967), Nei's distance (D_n; Nei, 1972), or proportion of shared alleles (D_{ps}; Bowcock *et al.*, 1994). These methods are most appropriate for species that have clustered distributions, and for species with relatively continuous distributions, such as many plant species, a growing number of individual-based distance methods are being developed (e.g., Manel *et al.*, 2007; Murphy *et al.*, 2008). Alternatively, allele frequencies themselves are used to cluster individuals among populations, as in assignment tests (Francois and Durand, 2010; Guillot *et al.*, 2008; Pritchard *et al.*, 2000). In turn, researchers test for correlations of hypothesized landscape variables with allele frequency distributions or genetic distances. The highlight here is on the most common analytical frameworks to achieve this goal below, which include use of Mantel (and partial Mantel) tests, barrier detection methods (including use of assignment tests), use of resistance surfaces, and network-based approaches. Other, less commonly used approaches have included ordination and multiple matrix regression, which will be covered briefly.

Analytical Frameworks

Mantel Tests

Mantel or partial Mantel tests remain a predominant way to test for correlations between a matrix of genetic distances and a matrix or matrices of geographic or landscape distances (Storfer *et al.*, 2010). Significant positive correlations with a particular landscape matrix and a genetic distance matrix suggest that a given landscape variable impedes gene flow, whereas significant negative correlations suggest the variable enhances gene flow. Partial Mantel tests, in particular, quantify the correlation between two distance matrices, while controlling for the effects of a third covariant matrix. In landscape genetics, partial Mantel tests are conducted to test the significance of correlations with a genetic distance matrix and usually several different landscape distance matrices. Causal modeling represents a relatively recent application of partial Mantel tests in landscape genetics (Cushman *et al.*, 2006). Causal modeling is used to test a variety of competing hypotheses as to which landscape factors best explain genetic structuring among groups, locales or populations. With a series of partial Mantel tests, competing hypotheses and the statistical predictions of each can be tested, with all statistical predictions met by only the hypothesis with the strongest support (Legendre and Troussellier, 1988). An example is a study of black bears in Northern Idaho, where the influence of a variety of different landscape variables was tested, including land cover, slope, elevation, and roads. Causal modeling of seven competing organizational models showed support for only one, which included only the influence of land cover and elevation on black bear genetic structure (Cushman *et al.*, 2006).

Debate continues about the use of partial Mantel tests in landscape genetics. Some researchers question the appropriateness of permutation procedures for significance testing among multiple matrices with non-independent or multicollinear variables (Castellano and Balletto, 2002; Rousset, 2002). In addition, a comparison of statistical methods (Balkenhol *et al.*, 2009a) showed that partial Mantel tests had lower power and higher type I error than other analyses, such as partial canonical correspondence analysis (CCA), and Bayesian inference of immigration rates (Faubet and Gaggiotti, 2008). Yet, others argue that although caution of high type I error of partial Mantel tests is warranted, this error is not inherent in using the test *per se* (Cushman and Landguth, 2010). Rather, high error rates in Mantel tests can be caused by correlations of expectations from alternative hypotheses (e.g., a particular landscape feature could act as an absolute barrier vs. providing an area of high resistance to dispersal; Cushman

and Landguth, 2010). Together, these studies warrant caution regarding use of partial Mantel tests, careful construction of hypotheses, and the need to use alternative statistical tests to confirm partial Mantel results.

Barrier Detection

As of 2010, approximately half of the landscape genetics studies focused on understanding barriers to gene flow using a variety of approaches (Storfer *et al.*, 2010). Interest in barriers to gene flow is broad, spanning from understanding basic landscape features that generate genetic population structure, to understanding the conservation implications of anthropogenic development (e.g., roads and dams). Early studies sampled on both sides of a putative barrier and compared genetic distance across the barrier with samples taken at similar distances in continuous habitat (Storfer *et al.*, 2007). If significantly decreased gene flow was found across the putative barrier and not the continuous habitat, it was concluded that the landscape feature was, in fact, a barrier. An example was a study that showed that a California highway acted as a barrier to gene flow for both bobcats and coyotes (Riley *et al.*, 2006).

More recently, methods in landscape genetics and spatial statistics have together allowed both more quantitative and spatially explicit ways of investigating potential barriers such as mountains, as well as ways to detect breaks in genetic continuity caused by less obvious or even cryptic landscape features, such as changes in moisture patterns across a study area (Storfer *et al.*, 2007). One approach for detecting changes in allele frequencies in space is Wombling (Barbujani *et al.*, 1989; Womble, 1951). Wombling has been used to identify spatial locations of abrupt changes in allele frequencies by quantifying local variability in allele frequencies (Cercueil *et al.*, 2007). The software WOMBSOFT (Crida and Manel, 2007) utilizes local polynomial regressions to estimate these genetic changes, as well as a binomial test for significance of each genetic break. A recent example is a study of Giant pandas in China, where a Wombling analysis showed genetic breaks between the Xiaoxiangling and Daxiangling Mountains (Zhu *et al.*, 2011). Another approach is using Monmonier's maximum difference algorithm, which creates a break line through the vectors between a network of sampled populations or individuals (Monmonier, 1973). Essentially, the Monmonier algorithm estimates the largest changes in allele frequencies along this break line, and thereby allows inference of barriers in these locations. Noted drawbacks of this method include dependence on the initial sampling design, as well as an *a priori* definition of the number of genetic groups present in the study area (Guillot *et al.*, 2009).

One way to circumvent the necessity of predefining the number of (genetic) populations is by using assignment tests. Assignment tests are most commonly used to determine the most likely number of genetic clusters (or *K*; i.e., populations) based on allele frequencies among samples using a Bayesian approach and Monte Carlo Markov Chain (MCMC) optimization (Pritchard *et al.*, 2000; Guillot *et al.*, 2008). Researchers should be aware that *K* is estimated in a variety of different ways, depending on the assumptions of the method, such as whether there is admixture or not, and the software used (Rowe and Beebee, 2007; Francois and Durand, 2010; Guillot *et al.*, 2009). Barriers or edges are commonly inferred when genetic clustering coincides with one or more geographic features consistent with those that would impede movement of the study organism (**Figure 2**). For example, the program GENELAND (Guillot *et al.*, 2008) was developed as a spatially explicit assignment-based approach to determine the number of genetic clusters (i.e., *K*) present in a dataset. *Tessellation* is used to construct a series of polygons that denotes the boundaries of the genetic clusters in space, independent of the sampling sites (Guillot *et al.*, 2008). Barriers can thus be inferred to coincide with the spatial locations of the polygon edges that delineate the genetic clusters. An individual-based assignment test method provides another approach for detecting genetic discontinuities among samples while avoiding population delineation (Manel *et al.*, 2007). This method generates a spatially referenced probability map for finding the genotype of a particular individual using a moving window approach. A mean slope for all individual probability maps is calculated, and population boundaries are identified by areas of high mean slope (Manel *et al.*, 2007). A comparison of assignment tests and other edge detection methods, including Wombling and Monmonier's algorithm, showed that assignment tests generally perform better on both simulated and empirical data (Safner *et al.*, 2011). Reviews of the applications of, and statistical issues with, assignment tests can be found in Manel *et al.* (2005), Guillot *et al.* (2009), and Francois and Durand (2010).

Resistance Surfaces

Another major way to assess effects of landscape features on the spatial distribution of genetic variation is with the use of resistance surfaces (Spear *et al.*, 2010). Generally, resistance surfaces are estimated by assigning values to each landscape variable or environmental feature contained in a spatial layer of a raster GIS environment (Spear *et al.*, 2010). Higher assigned values translate to greater hypothesized resistance to gene flow by that particular variable, whereas lower values suggest lower resistance. Thus, each grid square (e.g., 30 m or 1 km) in a GIS is assigned a cumulative value based on hypothesized habitat use and movement through each of the landscape variables contained in that grid square. The resistance surface comprised of all grid squares that make up the intervening habitat matrix between sampling localities across a study area (**Figure 3**). Rarely do researchers have direct estimates of dispersal and habitat use to parameterize such surfaces, and therefore assignment of resistance values to each landscape feature remains a challenge for modeling resistance surfaces (Spear *et al.*, 2010). As a result, the majority of resistance surfaces are constructed using subjective methods; a recent review suggested only approximately one-fourth of studies used nongenetic field data, such as radio-telemetry data or direct movement studies, to inform resistance assignment values (Spear *et al.*, 2010). One way to deal with the problem of subjectively assigning resistance values is by assigning a range of values to each landscape variable, thereby generating a variety of surfaces that comprise alternative hypotheses. Then, the multiple surfaces can be statistically

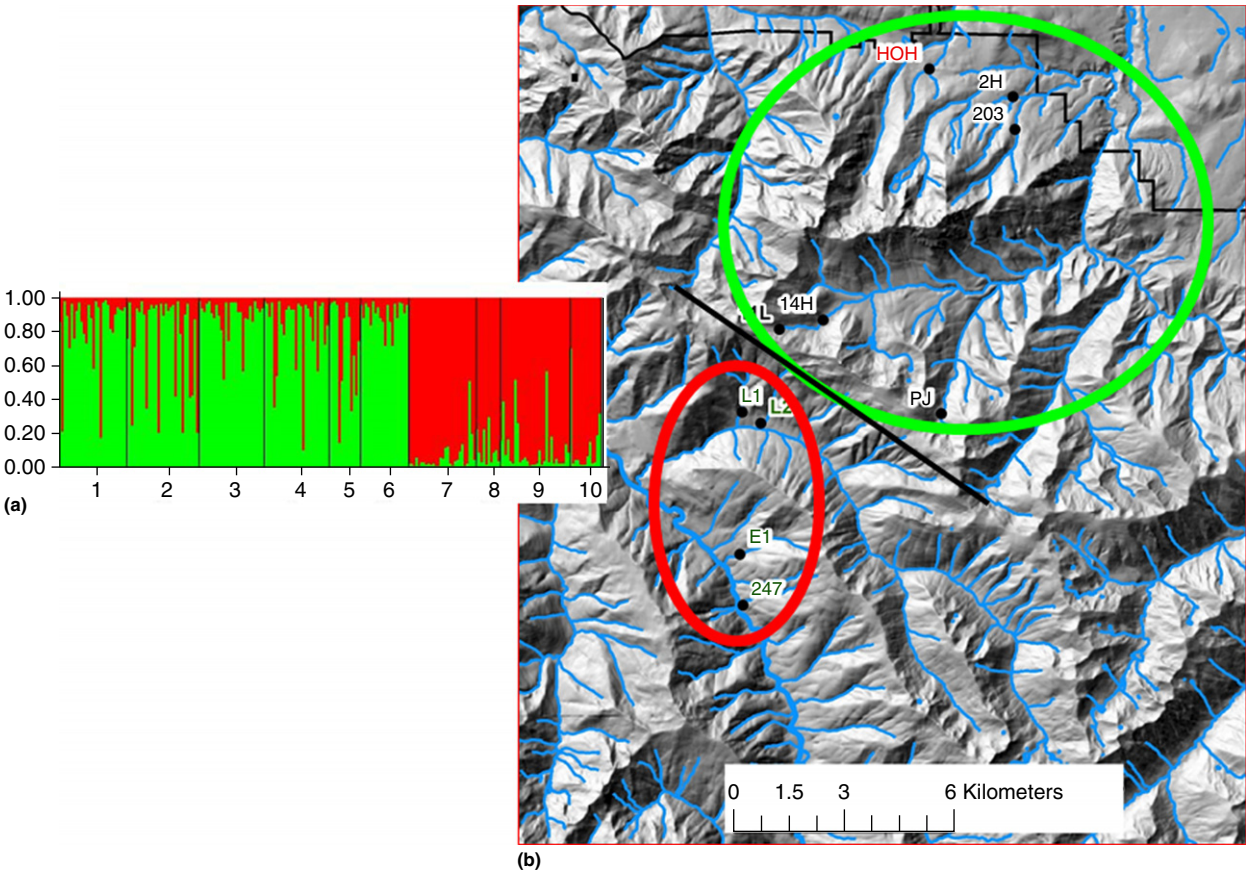


Figure 2 Represents a barrier detection method using a combination of Bayesian clustering analysis and overlay approach in Olympic National Park, USA. (a) Bayesian clustering output using the software STRUCTURE (Pritchard *et al.*, 2000) showing two genetic groups (red and green) identified from multiple sampling locations. Each bar is an individual, and the proportion of its genotype assigned to either population is shown by the proportion of the bar that is red vs. green. (b) Identification of barrier using overlay approach. Letter and number coded dots on map represent sampling localities. Red and green circles correspond to identified populations using Bayesian clustering. Black line is Hurricane Ridge, a visually identified barrier between the green and red genetic clusters.

compared as to which best fits the genetic distance data (Cushman *et al.*, 2006; Cushman and Landguth, 2010). Another potential solution is to incorporate habitat suitability maps into estimation of resistance surfaces by using presence/absence data for the study species (Wang *et al.*, 2008).

Two major modeling frameworks have been used to convert resistance values into measures of population connectivity – least-cost paths and circuit-based analyses. Least-cost paths are based on the underlying assumption that costs, in terms of cumulative resistance values between populations, are minimized such that individuals tend to move along an optimal path (**Figure 1(b)**; Adiransen *et al.*, 2003). As an example, Epps *et al.* (2007) used a least-cost based analysis to find that topographic distance explained a significantly higher proportion of variation in gene flow than Euclidean distance alone for desert bighorn sheep. Least-cost paths have been criticized because they are based on the potentially unrealistic assumption that individuals "know" *a priori*, which is the least costly path to take during dispersal events. One suggestion to overcome this problem is to consider multiple least-cost paths together in order to delineate a zone of probable movement (Rayfield *et al.*, 2010).

Circuit theory-based models simultaneously integrate all possible pathways that connect populations, while still assuming that genetic distance is positively correlated with resistance (McRae, 2006). The software CIRCUITSCAPE is used to estimate overall significance of the resistance surfaces in explaining genetic structuring among samples, along with tests of significance of individual landscape variables (Shah and McRae, 2008). The latter approach was shown to perform better, with a higher proportion of variation in gene flow explained when both least-cost path models and circuit theory-based models were applied to empirical datasets of mahogany and the wolverine (McRae and Beier, 2007). More recently, it has been suggested that least-cost paths may perform better on smaller spatial scales (Anderson *et al.*, 2010). Nonetheless, given the assumptions inherent in least-cost path and circuit theory-based (i.e., that all possible paths are influence movement in some way) analyses, caution should be used when applying either analysis to understand population-level connectivity. Spear *et al.* (2010) provide guidelines and potential solutions for addressing these problems, including use of computer simulations or network-based approaches.

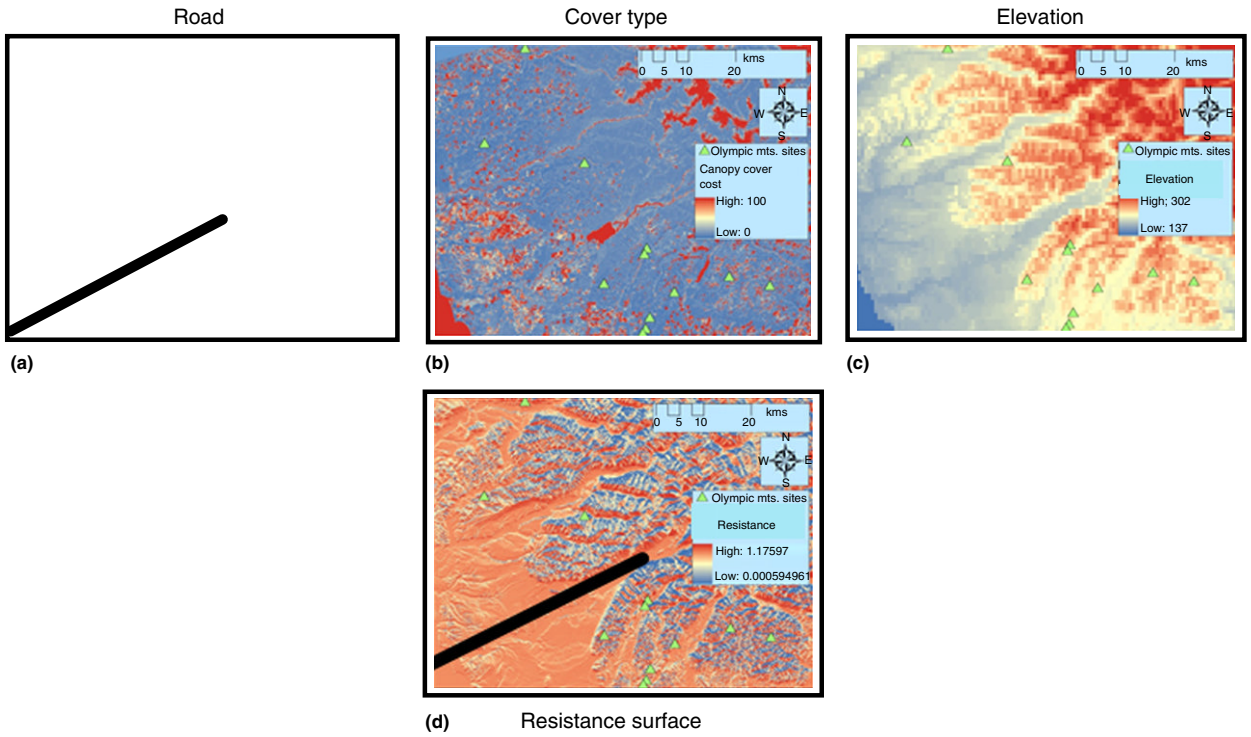


Figure 3 Steps to building a resistance surface using multiple landscape features. (a) Road through the study area. (b) Map of canopy cover in the area; blue indicates low cost areas, red indicates high cost, and green triangles represent sampling locations. (c) Map of elevation in the study area; red indicates high elevation and high cost, yellow shows intermediate cost and blue shows low cost. (d). Multivariate resistance surface indicating cumulative costs from (a)–(c).

Graph Theory and Network Analyses

Graph theory can be applied in population and landscape genetics by treating localities or individuals as nodes and connections between them as edges as in a network (Dyer and Nason, 2004). The strength of each edge is essentially proportional to the rate of gene flow between the two nodes that it connects (as estimated by their genetic covariance), and a complete lack of an edge suggests significant population subdivision (Dyer and Nason, 2004; Dyer, 2007). Originally, this framework was applied using the software POPGRAPH, whereby genetic structure was estimated across all nodes simultaneously, and users could test for population subdivision as indicated by significant deficiencies of edges between user-defined groups of nodes (**Figure 4**; Dyer and Nason, 2004). An empirical study that used this approach showed that, taken together, high elevation sites were significantly genetically differentiated from low elevation sites for the long-toed salamander (Giordano *et al.*, 2007).

Network analyses have been applied in the graph–theoretic approach to model functional connectivity (Garroway *et al.*, 2008; Murphy *et al.*, 2010). In addition to understanding overall connectivity of nodes and substructure across an entire network, analyses can also determine the properties of a specific node. Aside from the degree centrality of a node, which is an estimate of connectivity based on the number of connections it has to other nodes, other commonly evaluated properties include eigenvector centrality and betweenness (Garroway *et al.*, 2008). Eigenvector centrality incorporates both direct and indirect connectivity of particular nodes, thereby weighting well-connected nodes more heavily than those that are less connected. Betweenness is the number of shortest paths that a particular node

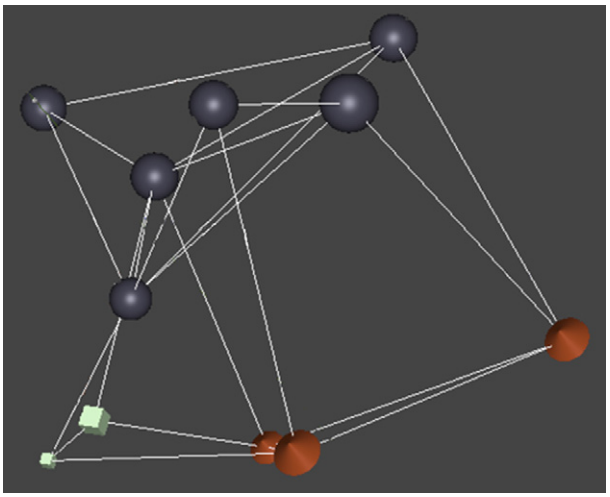


Figure 4 Example of a network using a graph–theoretic approach. Symbols indicate different sampling localities (i.e., nodes); shape size is proportional to the amount of observed genetic structure. Different shapes represent different genetic clusters, and lines between shapes (i.e., edges) represent significant genetic exchange between the two sampling localities that the lines connect.

(or edge) lies on. Assuming interactions within a network generally follow shortest paths, betweenness generally measures the importance of a particular node in terms of how often individuals must path through it (i.e., how much of a bottleneck it creates; Garroway *et al.*, 2008). Estimation of these nodal properties can translate into the understanding of the relative importance of a particular population or locality in sustaining an entire network or metapopulation. An example of a network analysis was conducted with fishers (*Martes pennanti*) in Ontario, Canada, where nodes were highly clustered and generally had short path lengths, suggesting high connectivity and high resiliency to loss of any particular node (Garroway *et al.*, 2008). Analyses also showed animals tended to move from high- to low-quality habitat, as well as evidence of cryptic population structure not indicated with traditional population genetic analyses (Garroway *et al.*, 2008). A later study incorporating use of resistance surfaces with network-based analyses showed that road density, snow depth, and river density all impeded gene flow (Garroway *et al.*, 2011).

Gravity models have also been adapted for use in network analyses of functional connectivity in a landscape genetics study (Murphy *et al.*, 2010). These models are based on Newton's law of gravitation and predict flow from a site within a network; such models have been historically used in modeling economic trade routes, for example (Willig and Bailey, 1979). Essentially, these models incorporate the influence of at-site (node) and between-site (edge) characteristics on number of trips through a network (Anderson, 1979). In a landscape genetics study, some measure of genetic distance is used as the response variable as opposed to number of trips (Murphy *et al.*, 2010). A gravity model-based approach yielded evidence of source–sink dynamics among localities of the Columbia spotted frog (*Rana luteiventris*), whereby presence of predatory fish at sites reduced connectivity but growing season and site productivity were correlated with increased connectivity (Murphy *et al.*, 2010).

Overall, the network analytical framework holds much promise for application in landscape genetics. Typically, landscape genetics studies have focused solely on the influence of between-site (i.e., edge) characteristics on functional connectivity among populations. However, network-based analytical approaches also yield information about the influence of at-site characteristics on gene flow and population genetic structure. These data can be quite useful for conservation and management, for example, by emphasizing protection of sites with high degrees of centrality. These approaches also obviate the need for *a priori* assumptions of the number of populations in a sample (Garroway *et al.*, 2008). In addition, network-based approaches are 'model free' (Dyer and Nason, 2004), in the sense that the significance of observed networks can be obtained by comparisons with fully saturated or randomly generated networks (Garroway *et al.*, 2008).

Ordination and Regression

Ordination (e.g., principal components analysis, or PCA, canonical correspondence analysis or CCA) has also been shown to be a powerful technique for investigating the spatial arrangement of genotypes on a landscape (Balkenhol *et al.*, 2009a). Because space can explicitly be included in CCA (Borcard *et al.*, 1992), the analytical method can be used to identify correlations between allele frequency distributions and gradients in environmental variables (ter Braak and Verdonschot, 1995). Recently, it was shown that partial CCA was effective in accounting for spatial effects through the incorporation of coordinates as covariates (Balkenhol *et al.*, 2009a). A simulation study that compared CCA with several other statistical methods, including Mantel tests and partial Mantel tests, showed the method had higher power to detect landscape effects on spatial genetic structuring (Balkenhol *et al.*, 2009a). Despite its potential, however, ordination has most often been used in empirical studies to cluster the genetic data without including landscape variables (Storfer *et al.*, 2010).

Linear regression or general linear models have also commonly been used in landscape genetics studies. However, these methods should be cautiously used because they are generally inappropriate in situations of substantial autocorrelation (i.e., spatial non-independence of genetic and landscape data), unless significance testing relies on data permutation approaches (Fortin and Dale, 2005). Potential alternatives that attempt to account for non-independence and *spatial autocorrelation* are classification and regression tree (CART) models (Murphy *et al.*, 2010), geographically weighted regression (Spear and Storfer, 2008), and network models (Garroway *et al.*, 2011; Murphy *et al.*, 2010).

Multiple matrix regression techniques also hold promise, and one approach that has been applied in landscape genetics studies is multiple regression of distance matrices (MRDM; Legendre *et al.* (1994)). This procedure involves a stepwise regression of a dependent (genetic) distance matrix and multiple independent pairwise (landscape) matrices. Landscape data are pairwise estimates of spatial separation between sampling points or localities. An MRDM analysis was used to understand landscape influence on genetic structure of the bush cricket, *Metrioptera roeseli*, whereby historic landscape patterns (50 years old) explained genetic structure better than current landscape pattern (Holzhauer *et al.*, 2006). Relative to other methods, including Mantel tests and Bayesian analyses, MRDM had high power and low type I error rates in a simulation study (Balkenhol *et al.*, 2009a).

Research Insights

General Trends

The majority of landscape genetics studies have lacked explicit hypotheses to test the influence of landscape variables on genetic connectivity (Storfer *et al.* (2010); but see Cushman *et al.* (2006) and Wang and Summers (2010)). For example, the most commonly studied feature has been the influence of putative barriers, which have frequently been revealed by visual inspection of genetic clusters for spatial coincidences of cluster boundaries with landscape features (e.g., **Figure 2**). However, given that landscape genetics is a relatively young discipline (Manel *et al.*, 2003), it is not surprising that initial studies have often been exploratory as opposed to hypothesis-

driven. That is, a basic understanding of gene flow and genetic population structure is necessary before testing specific hypotheses related to the influence of landscape variables. This is particularly true for species for which little ecological data regarding dispersal and habitat use are available. Nonetheless, given the multitude of recent theoretical developments and the utility of simulation studies and software (e.g., Cushman and Landguth, 2010; Epperson *et al.*, 2010), landscape genetics should move toward testing specific hypotheses of how environmental and landscape features influence the spatial distribution of genetic variation (Balkenhol *et al.*, 2009a; Manel *et al.*, 2010a; Storfer *et al.*, 2010).

In addition, landscape genetics studies have been largely confirmatory, for example finding support for the influence of variables intuitively thought to be gene flow barriers, such as ridgelines, rivers, or roads. Again, given that the discipline is relatively new, confirmation has largely been positive, demonstrating that spatial models that include landscape features most often explain a significantly higher proportion of variation in gene flow than isolation by (Euclidean) distance models. Nonetheless, some studies showed no improvement in models explaining gene flow despite the inclusion of landscape variables (e.g., Bolliger *et al.*, 2011; Rieux *et al.*, 2011), and several studies showed evidence of counterintuitive effects of landscape features. For example, although rivers were thought as an *a priori* barrier to gene flow among populations of blotched tiger salamanders (*Ambystoma tigrinum melanostictum*) in Yellowstone National Park, they were actually found to facilitate gene flow likely due to the homogenizing influence of recurrent, occasional floods (Spear *et al.*, 2005). Similarly, flooding apparently increased gene flow among populations of the herb, *Primula sieboldii* (Kitamoto *et al.*, 2005). In addition, contrary to predictions, postfire-regenerated shrub habitat in Yellowstone National Park was shown to facilitate gene flow among tiger salamanders (Spear *et al.*, 2005) and boreal toads (*Bufo boreas*; Murphy *et al.*, 2010). Although previously thought of as an impediment to gene flow, shrubs probably facilitate amphibian dispersal by providing better shade and thereby desiccation protection than forested habitat with taller trees.

Other general trends included a focus on temperate versus tropical areas, despite the fact that tropical areas are hotspots of biodiversity (Myers *et al.*, 2000), and increasing targets of conservation efforts (Carnaval *et al.*, 2009). Biases also existed in the use of molecular markers, using primarily microsatellites in animals, with AFLPs used primarily in plants (Holderegger *et al.*, 2010; Storfer *et al.*, 2010). In addition, most studies focused on a single species, and the majority occurred on vertebrate species in terrestrial habitat (Storfer *et al.*, 2010). Overall, there are also fewer plant studies than those of animals, largely because of general assumptions implicit in the most broadly used analytical frameworks that include landscape distance and resistance (Holderegger *et al.*, 2010). Nevertheless, plants are well-suited for studies of barriers to gene flow, as well as for understanding the spatial distribution of adaptive genetic variation (Holderegger *et al.*, 2010). Approximately 80% of landscape genetics studies focused on identification of barriers (50%) and understanding the effects of landscape features on genetic variation (30%), with the effects of anthropogenic development and habitat use also considered by many studies; generalities arising from these three research foci are discussed in the following three sections.

Identifying Barriers

Some of the most frequently considered barriers to gene flow in landscape genetics studies are linear terrestrial features, such as rivers and mountain ridges, roads, and edges created by anthropogenic habitat fragmentation. Rivers have been identified as a barrier to gene flow in several taxonomic groups including small mammals, turtles, deer, and fishers (Coulon *et al.*, 2006; Garroway *et al.*, 2008, 2011; Lugon-Moulin and Hausser, 2002; Mockford *et al.*, 2007), but may actually facilitate gene flow in amphibians (Spear *et al.*, 2005) and transgenic crops (Cureton *et al.*, 2006). Major ridgelines are barriers to gene flow in amphibians (Funk *et al.*, 2005; Giordano *et al.*, 2007) and meso-mammals (Zalewski *et al.*, 2009), but may not limit gene flow in small mammals (Lugon-Moulin and Hausser, 2002). Mountain ranges, as well as low-elevation valleys served as broad scale gene flow barriers to Mojave desert tortoises (Hagerty *et al.*, 2011). A large valley also served as a barrier to gene flow in the African acacia, *Senegalia mellifera* (Guajardo *et al.*, 2011). Although effects of some natural habitat features may be species-specific, roads and other human developments consistently impede gene flow for animals spanning several taxonomic groups, including carnivores (Millions and Swanson, 2007; Riley *et al.*, 2006), ungulates (Epps *et al.*, 2007; Kuehn *et al.*, 2007; Perez-Espona *et al.*, 2008), and amphibians (Manier and Arnold, 2006; Murphy *et al.*, 2010).

Clearly defined linear barriers are also present in aquatic systems. Dams are major barriers to gene flow in aquatic species such as fish, even when ladders and other mitigation features designed to facilitate dispersal across the dam are in place (Deiner *et al.*, 2007; Raeymaekers *et al.*, 2008; Skalski *et al.*, 2008; Taylor *et al.*, 2003; Wofford *et al.*, 2005). Also, natural physical barriers such as waterfalls have been important for shaping genetic structure patterns in yellow perch (*Perca flavescens*; Leclerc *et al.*, 2008), various trout species (Deiner *et al.*, 2007; Taylor *et al.*, 2003; Whiteley *et al.*, 2006; Wofford *et al.*, 2005), and mountain whitefish (*Prosopium williamsoni*; Whiteley *et al.*, 2006). Anthropogenic development adjacent to waterways, such as timber harvest can change water chemistry, temperature, dissolved oxygen levels and light penetrance, thereby acting as a barrier to fish gene flow (Wilcock *et al.*, 2007). Further, fragmentation of breeding habitat may have an additive effect with physical structures such as dams (Leclerc *et al.*, 2008) and roads (Balkenhol and Waits, 2009).

Other, less obvious features can act as barriers to gene flow. For example, unsuitable natural habitat such as savanna acts to fragment lemur populations (Radespiel *et al.*, 2008) and dry grassland habitat increases genetic structuring among salamander populations (Rittenhouse and Semlitsch, 2006). Cryptic barriers to gene flow, such as climate gradients, have been shown to produce unexpected sharp breaks in genetic structure possibly due to species' environmental tolerance limits. An example can be found among salt-water habitat

where cetaceans have abrupt discontinuities in genetic structure (Fontaine *et al.*, 2007). In some cases, the additive effects of several landscape features of low permeability can cause detectable differentiation even among weakly genetically structured populations, such as the combined influence of a river, highways, and canals on roe deer (*Capreolus capreolus*) gene flow in France (Coulon *et al.*, 2006).

Identification of barriers to gene flow can help predict the geographic nature of disease spread or assist with management (Biek and Real, 2010). Barriers to parasite dispersal depend on mode of spread. Directly transmitted parasites are generally dependent on dispersal of the host, and often have smaller genetic neighborhoods than vectored parasites or wind-borne parasites (Biek and Real, 2010). Rivers have commonly been shown to be a barrier to host movement and thereby disease dispersal, from chronic wasting disease in deer (Blanchong *et al.*, 2008), to limiting movement of various primate species and consequently the spread of symian immunodeficiency virus (SIV) (Gagneux *et al.*, 2001). However, a study of two rivers showed that one served as a barrier to raccoon gene flow and consequent rabies spread, whereas rabies persistently spread across a different river (Cullingham *et al.*, 2009). As a result, caution should be used in assuming host genetic structure correlates with disease structure and spread (Biek and Real, 2010). Nonetheless, this type of information facilitated efficient rabies control strategies, with focused deployment of vaccinated raccoon bait where landscape variables were shown to facilitate, rather than act as barriers to, dispersal and gene flow (Real and Biek, 2007).

Influence of Landscape Variables Spatial Genetic Variation

Second to the study of barriers, studies of the influence of landscape variables on spatial genetic variation comprised the largest number of studies (Storfer *et al.*, 2010). Topographic relief was shown to restrict gene flow in many terrestrial species, including most species of amphibians (Funk *et al.*, 2005; Murphy *et al.*, 2010; Spear and Storfer, 2008), small mammals including voles (Berthier *et al.*, 2006), red deer (*Cervus elaphus*; Pérez-Espona *et al.*, 2008), wax palms (*Ceroxylon echinulatum*; Trénel *et al.*, 2008) and golden-brown mouse lemurs (*Microcebus ravelobensis*; Radespiel *et al.*, 2008). Elevational gradients limited gene flow between high- and low-altitude populations of Columbia spotted frogs (*R. luteiventris*; Funk *et al.*, 2005) and long-toed salamanders (*Ambystoma macrodactylum*; Giordano *et al.*, 2007). However, landscape features had little effect on gene flow and population genetic structure of damselflies (Wellenreuther *et al.*, 2011) and several species of birds (Barr *et al.*, 2008; Petren *et al.*, 2005). However, some populations of red-tailed hawks (*Buteo jamaicensis*) genetic structuring occurred across some mountain ranges, but not others (Hull *et al.*, 2008). Instead, habitat preferences were even more important for explaining red-tailed hawk gene flow, with limited dispersal between Mediterranean central California habitats and more xeric inland western and southern habitats (Hull *et al.*, 2008).

Seasonality can also affect gene flow. As an example, flowering time at different altitudes was significantly affected by timing of snowmelt, and flowering segregation, in turn, leading to fine-scale spatial genetic structuring in the alpine snowbed herb, *Primula cunefolia*, (Hirao and Kudo, 2004, 2008). Winter also affects genetic structure in fisher populations, as deep snow restricts gene flow (Garroway *et al.*, 2011).

In river and spring systems, landscape genetic models that included the drainage pattern, direction and speed of water flow best explained patterns of genetic structure in zooplankton (Michels *et al.*, 2001), brook charr (Angers *et al.*, 1999), and aquatic snails (Wilmer *et al.*, 2008). Also, dynamic and sporadic events like flooding were critical for explaining genetic structure (Wilmer *et al.*, 2008). The landscape variables of slope, elevation, and temperature were important predictor variables in multiple studies. For example, fish populations at higher elevations had lower levels of genetic diversity and lower levels of gene flow with other populations (Angers *et al.*, 1999; Castric *et al.*, 2001; Narum *et al.*, 2008), and gene flow was negatively correlated with change in elevation between sampling sites for the stream salamander, (*Gyrinophilus porphyriticus*; Lowe *et al.*, 2006). Precipitation was found to be important for genetic structuring among three Pacific salmon species (Olsen *et al.*, 2011). In contrast, water temperature explained the most variation in genetic structure among Atlantic salmon (*Salmo salar*) along the North American Atlantic coast (Dionne *et al.*, 2008) and steelhead trout in the Pacific Northwest (Narum *et al.*, 2008).

In marine environments, landscape genetic studies have demonstrated that genetic diversity and structure are often best explained by analyses that include ocean currents (blue whiting, *Micromesistius poutassou*; Was *et al.*, 2008, and kelp bass *Paralabrax clathratus*, Selkoe *et al.*, 2010). In other studies, ocean temperatures were the most important predictor variable of genetic structure in the sea urchin, *Centrostephanus rodgersii* (Banks *et al.*, 2007) whereas both temperature and salinity were key variables for explaining genetic structure patterns in Atlantic cod (*Gadus morhua*; Case *et al.*, 2005). In a multispecies analysis along the California coast, cumulative kelp cover, a measure of habitat quality, was the most important variable for explaining genetic diversity and structure in kelp bass, Kellet's whelk (*Kelletia kelletii*) and California spiny lobster (*Panulirus interruptus*; Selkoe *et al.*, 2010).

In sum, a wide variety of landscape variables have been tested for their influence on population genetic structure. In terrestrial systems, elevation, ridgelines, and topographic relief were often found to limit gene flow in terrestrial species. In aquatic species, drainage structure, slope, elevation, and temperature were frequently important explanatory variables. In marine species in particular, ocean currents have also been found to be important.

Tests of Anthropogenic Influence

Effects of anthropogenic land use change have been the focus of a large number of terrestrial and aquatic landscape genetic studies, and results have been mixed. Habitat fragmentation has reduced gene flow in many species, including: formica ants (Mäki-Petäys *et al.*, 2005), common frogs (*Rana temporaria*; Johannson *et al.*, 2005), alpine butterflies (*Parnassius smintheus*; Keyghobadi *et al.*, 2005), and golden cheeked warblers (*Dendroica chrysoparia*). For example, among fragmented populations of the forest herb, *Primula elatior*, lower genetic

variation was found in (developed) habitat patches less than 35 years old relative to those older than 35 years (Jacquemyn *et al.*, 2004). A study in fragments versus closed forest of Andean oak (*Quercus humboldtii*) also showed detrimental effects of anthropogenic habitat use on genetic variability (Fernandez and Sork, 2007). Further, urban areas acted as gene flow barriers for the alpine newt, *Mesotriton alpestris*, but forests acted as dispersal corridors, highlighting the importance of maintaining natural habitat (Emaresi *et al.*, 2011).

For multiple riverine systems, decreased gene flow and genetic diversity were correlated with the presence of anthropogenic landscape features. Dams, tunnels and weirs were found to impede gene flow in yellow perch (Leclerc *et al.*, 2008), several trout species (Costello *et al.*, 2003; Taylor *et al.*, 2003; Wofford *et al.*, 2005), and three-spined stickleback (*Gasterosteus aculeatus*; Raeymaekers *et al.*, 2008). Pulp mill impacts were also found to increase genetic structure in redbreast sunfish (*Lepomis auritus*; Theodorakis *et al.*, 2006) and three-spined stickleback (Raeymaekers *et al.*, 2008). Similarly, copper mine waste outputs restricted gene flow among populations of the saltwater kelp, *Lessonia nigrescens*, in Chile (Faugeron *et al.*, 2005).

However, no effects of fragmentation have been found for several other species, including: Amazon liverworts (*Radula flaccid*; Zartman *et al.*, 2006), the tree species, *Sorbus aucuparia* (Bacles *et al.*, 2004). In a Mexican *Heliconia* species, habitat fragmentation appeared to have little effect on genetic diversity, despite low levels of gene flow (Suarez-Montes *et al.*, 2011). In addition, habitat fragmentation caused by obvious anthropogenic barriers such as roads apparently did not affect gene flow in coyotes (*Canis latrans*); rather gene flow appeared restricted by breaks in habitat types, as well as natal site fidelity (Sacks *et al.*, 2004).

Fragmented landscapes did apparently affect habitat use during dispersal movements for some species highlighting the need to investigate more detailed hypotheses than effects on gene flow or genetic diversity, *sensu stricto*. For example, despite the fact that overall gene flow appeared unaffected in the yellow-footed antechinus, dispersal was reduced through agricultural landscapes (Lada *et al.*, 2008). In addition, gene flow remained high between primary and secondary growth (i.e., postharvest) forest fragments for the forest herb, *Geum urbanum*, but small populations were more genetically diverged than large populations (Vandepitte *et al.*, 2007). Gene flow of Rocky Mountain tailed frogs (*Ascaphus montanus*) more closely followed riparian corridors in deforested landscapes than naturally regenerated postfire landscapes (Spear and Storfer, 2010).

Discrepancies among habitat fragmentation studies may also result from an inability to disentangle the influence of contemporary versus historic landscape features on gene flow (Keyghobadi *et al.*, 2005; Pavlacky *et al.*, 2009; Storfer *et al.*, 2007). On the one hand, historical events could have long-term consequences and potentially confound inferences based on current landscape structure (Landguth *et al.*, 2010). Alternatively, effects of contemporary fragmentation may not be seen due to a lag time necessary for statistical detection. For example, Spear and Storfer (2008) examined correlations of current and historic patterns of timber harvest on Coastal tailed frog (*Ascaphus truei*) gene flow, which correlated strongly with harvest patterns from 20 years ago. This time lag may be due to the speed in which timber harvest can alter landscape structure, as well as the long-generation time of tailed frogs, which could delay the detection of a genetic signature of reduced dispersal. Similarly, genetic structure was explained better by historical natural events, rather than current habitat fragmentation among populations of the Glanville fritillary butterfly (*Melitaea cinxia*; Orsini *et al.*, 2008). In contrast, contemporary deforestation explained more variation in genetic differentiation than a reconstructed pre-European settlement landscape for the rainforest dwelling logrunner (*Orthonyx temminckii*; Pavlacky *et al.*, 2009).

Overall, given the variability in effects of habitat fragmentation on genetic variation and gene flow, studies that test the effects of fragmentation should be conducted before conservation and management decisions are implemented. Anthropogenic features, such as deforestation, agricultural development, damming, and other types of waterway manipulation decreased gene flow or altered movement pathways in some species, but had little to no effect on others. Discrepancies among studies highlight the need for species-specific studies, as well as those that consider current versus historical effects on population genetic structure.

Challenges and Future Directions

Interdisciplinary Communication

It has been noted in several different reviews that because each of the subdisciplines involved in landscape genetics – population genetics, landscape ecology, and spatial statistics – has distinct terminology and analytical approaches, challenges arise associated with cross-disciplinary communication (Balkenhol *et al.*, 2009b; Holderegger and Wagner, 2006; Jacquiery *et al.*, 2011; Storfer *et al.*, 2007, 2010). Recently, however, these challenges are being addressed. First, both a working group and a distributed graduate seminar were funded by the US National Center for Ecological Analysis and Synthesis (NCEAS) to bring diverse groups of researchers from the different disciplines together to collaborate and develop new analytical approaches. A major result of these efforts was a special issue of *Molecular Ecology* (September, 2010) focused on landscape genetics, and papers within the issue were commonly co-authored by collaborators among different disciplines. The distributed graduate seminar involved co-instructors from a number of different universities, as well as different disciplines (Wagner *et al.*, 2012). The goal of this seminar was interdisciplinary training of graduate students through face-to-face and virtual lectures, on-line interactions among students and instructors from remote locations, computer labs, group projects, and then finally a synthesis meeting among a group of students selected from the course as well as course faculty (Wagner *et al.*, 2012). These types of training models have been shown to be successful in helping to break down communication barriers through student immersion in collaborative projects (Reichman, 2004). Second, there have been a number of workshops and symposia associated with national meetings of different professional societies (e.g., Balkenhol *et al.*, 2009b). As we continue into the future, more and more students and postdoctoral researchers will be

trained through interdisciplinary courses, rotations among laboratories and other collaborative activities, thereby eventually making communication in landscape genetics seamless.

The Effects of Spatial Scale

The effects of spatial scale have been investigated in only a few landscape genetics studies, although consideration of multiple scales is important because focusing on one spatial scale may miss the influence of important landscape variables at a finer or broader spatial scale (Anderson *et al.*, 2010). The relative influence of independent variables at different spatial scales can be tested to establish the most appropriate spatial scale to include in a model (Cushman *et al.*, 2006; Murphy *et al.*, 2010). One study showed that different landscape variables influenced boreal toads (*B. boreas*) in Yellowstone National Park at different spatial scales (Murphy *et al.*, 2010). By testing the influence of landscape variables at multiple bandwidths (from 30 m to 960 m) connecting sites, it was shown that habitat permeability and topographic roughness influenced population connectivity at fine spatial scales, ridgelines were important at broad scales, and temperature and moisture had effects across a range of spatial scales (Murphy *et al.*, 2010). Another study showed the effects of spatial scale on genetic connectivity among pitcher plant midge (*Metriocnemus knabi*) populations (Rasic and Keyghobadi, 2012). Specifically, broad-scale landscape features that included bog size and the spatial arrangement and density of pitcher plants influenced fine scale female oviposition behavior, leading to both fine and broad scale patterns of genetic differentiation (Rasic and Keyghobadi, 2012).

These studies highlight the utility of investigations that span different spatial scales as the influence of landscape variables on gene flow at one scale may differ from those at another scale (Murphy *et al.*, 2010), or variables acting at one spatial scale may influence those at another scale (Rasic and Keyghobadi, 2012), possibly leading to incorrect inference about the truly important variables. In addition, studies of spatial scale can inform proposed management or development actions, for example, by simulation testing of the effect of different proposed road widths on species of conservation concern.

Spatial Replication and Multispecies Comparative Studies

Most landscape genetics studies focus specifically on one portion of a species' geographic range (Holderegger and Wagner, 2008; Segelbacher *et al.*, 2010; but see Short-Bull *et al.*, 2011). However, spatial replication is important, as landscape processes in one part of a species' range may affect gene flow differently than in another part of a species' range (Short-Bull *et al.*, 2011). An example is with the black bear, where different landscape variables affected genetic structure differently across the range (Short Bull *et al.*, 2011). For example, elevation was only important in areas that had standard deviations in elevation greater than 300 m (Short Bull *et al.*, 2010). In addition, landscape variables explained significantly more variation in gene flow among 11 out of 12 study areas for the black bear, but classical IBD explained gene flow in the twelfth. As a result, a study that only focused on the twelfth area by change would lead to the erroneous conclusion that landscape effects were not important for gene flow among black bear populations, a logical conclusion for such a highly dispersing species (Short Bull *et al.*, 2010).

Variation in genetic structure could also be due to ecological differences among populations, such as varying interactions with the abiotic environment, species interactions, or sexual selection (Rundle and Nosil, 2005). The strawberry poison-dart frog represents an example, where genetic structuring appears to be correlated with sexual selection based on color pattern variation, rather than landscape features (Wang and Summers, 2010). However, differences in the relationships between gene flow and landscape features across a species' range could simply be due to variation in the landscape itself and not the biology of the species *per se* (Short-Bull *et al.*, 2011).

From an applied research perspective, lack of replication in landscape genetic studies can also impede the conservation of biological diversity. When genetic data are limited to a small portion of a geographic species' range, they may be incorrectly extrapolated to other portions of the range to guide management resulting in conservation measures that are ineffective (Short-Bull *et al.*, 2011; Waples and Gaggiotti, 2006). Thus, replicated landscape genetics studies can also contribute to a better understanding of the evolutionary ecology of a species, and consistency of the influence of landscape variables among regions lends more credibility to their general influence of dispersal and gene flow of a focal study species.

More multispecies comparative landscape genetics studies are also needed to understand whether features of particular landscapes have consistent or contrasting effects across study species. An example was a study of Columbia spotted frogs (*R. luteiventris*) and long-toed salamanders across an agricultural landscape (Goldberg and Waits, 2010). Urban and rural development were associated with the highest landscape resistances for both species, but the effects of specific habitat features differed. That is, a moisture gradient provided the least resistance for long-toed salamanders, but agriculture and shrub/clearcut habitat had the highest permeability values for Columbia spotted frogs (Goldberg and Waits, 2010). Another study that compared landscape genetic structure of three Pacific salmon species, showed similar genetic structuring at broad spatial scales, but at finer spatial scales, subbasin area affected genetic structure only in Chinook salmon (*Oncorhynchus tsawytcha*; Olsen *et al.*, 2011). Life history variation can also affect genetic structuring among closely related species in the same area. For example, bee-pollinated species of the plant *Penstemon* (*Penstemon deustus* and *Penstemon pachyphyllus*) had high genetic structuring and low gene flow relative to the hummingbird pollenated congener *Penstemon rostriflorus* (Kramer *et al.*, 2011). Thus, not only can multispecies studies suggest consistency of effects of some landscape variables, but they can also elucidate how ecological differences among related study organisms can translate to variation in effects of particular landscape features on genetic structure.

Moving Beyond F_{ST}

Although 19 different measurements of genetic distance have been used among landscape genetics studies, F_{ST} and its analogs (including G_{ST}, G_{ST}', R_{ST}), were by far the most commonly used

(Storfer *et al.*, 2010). A variety of concerns have been raised regarding use of F_{ST}, including the fact that direct translation to the number of migrants per generation is often inaccurate (Whitlock and McCauley, 1999). That is, genetic analyses are considered indirect studies of dispersal; a distribution of allele frequencies among groups is used to infer species' dispersal patterns. Several authors have considered a number of problems with the assumption inherent in most landscape genetics studies – that spatial patterns of genotypic distributions reflect *current* dispersal patterns (Jaquiéry *et al.*, 2011; Marko and Hart, 2011). In general, however, F_{ST}-based metrics have equilibrium-based assumptions, and thereby a lag time may be necessary to detect changes in genetic structure that result from land use change (Landguth and Cushman, 2010). This is particularly disconcerting given the abundance of landscape genetics studies that are conducted in anthropogenically altered systems not expected to be in equilibrium. Empirical data support this concern, and studies actually showed that past landscape matrices explained patterns of genetic structure better than current patterns (e.g., Orsini *et al.*, 2008; Pavlacky *et al.*, 2009; Spear and Storfer, 2008). As such, researchers should consider use of other metrics that avoid equilibrium assumptions, such as D_{ps} (Bowcock *et al.*, 1994), or coalescent models that investigate allele genealogies (Marko and Hart, 2011).

Direct studies of animal movement have their own sets of problems, including limitations associated with techniques such as radio-tracking (e.g., cost and size of radiotrackers makes them unfeasible for relatively small study organisms) and mark-recapture studies, which tend to require large sample sizes and relatively high recapture rates (Jaquiéry *et al.*, 2011). In addition, inferring genetic structure from direct dispersal studies is tenuous insomuch as animals that disperse from one habitat to another may not actually breed there due to a variety of reasons that include unfamiliarity with the new habitat, being socially ostracized, or stress associated with the migration event itself. One way of dealing with this issue is combining direct dispersal estimates with indirect genetic estimates, as has been done with studies that integrate habitat suitability modeling and landscape genetic analyses (e.g., Wang *et al.*, 2008).

The predominant use of pairwise distance-based analyses, such as genetic distance matrices and Bayesian assignment methods in landscape genetics relies on the implicit, and sometimes unrealistic assumption that the study species is clustered into groups. However, these methods may be inappropriate for organisms with uniform or non-clumped distributions, such as many plant species or high-dispersing animal species. Use of individual-based genetic distances (Manel *et al.*, 2007) or interpolation methods to create genetic surfaces among individuals (Murphy *et al.*, 2008) avoid assumptions of discrete clustering.

Landscape Genomics

Recent technological advances have led to the availability of vast amounts of molecular data available for landscape genetics studies, and researchers are now able to use large numbers of loci to assess the spatial distribution of adaptive genetic variation (Holderegger *et al.*, 2010; Manel *et al.*, 2010a; Segelbacher *et al.*, 2010). Although the ability of natural selection to shape phenotypic differences among populations has long been recognized, we are now beginning to understand the molecular underpinnings of the spatial distribution of adaptive genetic variation. These studies not only benefit basic evolutionary biology, such as understanding species' range limits or ecological speciation, but it is also becoming increasingly important to assess the adaptive potential of natural populations in the face of global change (Manel *et al.*, 2010a).

Spatial analysis of genomic data in particular requires new analytical methods and approaches given the sheer size of the datasets (Manel *et al.*, 2010a; Segelbacher *et al.*, 2009). Indeed, the high statistical power to conduct tests of correlations of genetic variation and environmental variation using hundreds to thousands of loci is likely to result in a large number of spurious correlations, which may inflate numbers of candidate adaptive loci. Studies with hundreds to thousands of loci are useful for genomic mapping, as well as identifying candidate loci for adaptation to different environmental conditions (Holderegger and Wagner, 2008; Manel *et al.*, 2010a; Stinchcombe and Hoekstra, 2008). The latter approach has been termed "landscape genomics" (Joost *et al.*, 2007).

Four years after the term, "landscape genetics" was coined, Joost *et al.* (2007) developed a spatial analysis method (SAM) for testing the spatial correlation of large numbers of loci with environmental variables to reveal candidate loci for adaptation to environmental conditions. Similar to the advantage landscape genetics has relative to population genetics, SAM analyses add an explicit spatial component to "population genomics" studies, which depend on sampling large numbers of genetic markers dispersed throughout the genome. Population genomics studies reveal spatial patterns of genetic population structure that include "outlier loci", or those with differentiation that is greater than expected by the balance between genetic drift and migration (Luikart *et al.*, 2003). In turn, these loci are assumed to be under selection or physically linked to those under selection. As an example, an assessment of 392 AFLPs in the common frog (*R. temporaria*) among populations at varying altitudes in the Alps revealed eight candidate loci potentially involved in adaptation to altitude (Bonin *et al.*, 2006). Another study investigated whether hundreds of AFLP loci in the alpine plant, *Arabis alpina*, were correlated with environmental variables at both broad and fine spatial scales across an altitudinal gradient throughout Europe (Manel *et al.*, 2010b). Temperature and precipitation consistently influenced the distribution of alleles across all spatial scales, and 43% of the loci strongly correlated with environmental variables were consistent between fine and broad spatial scales. As such, several loci were also thought to be involved with adaptation to environmental variables that were thought to be site-specific (Manel *et al.*, 2010b).

It has been noted, however, that most landscape genomic studies are currently in the exploratory stage of identifying candidate loci (Manel *et al.*, 2010a, 2010b). That is, even hundreds of AFLP loci are not enough to pinpoint locations of candidate loci in the genome and thus, once a candidate locus is identified, a significant amount of downstream work is necessary to confirm its role in adaptation. Indeed, this work often includes fine-scale genetic mapping, along with assessing the fitness value of the genes themselves through

knockout experiments or other selection experiments (Hoffman and Willi, 2008; Manel *et al.*, 2010b). Even for organisms for which complete genome data and fine scale mapping exists (e.g., *Drosophila*, mouse), once loci are identified, the following experiments to confirm the adaptive value of candidate loci are still laborious. Although overcoming bioinformatic and experimental challenges is necessary to integrate genomic and spatial data in landscape genomics studies, it is an exciting time for evolutionary biologists, ecologists and population geneticists as we begin to truly unravel the genetic basis of adaptation.

Acknowledgments

This work was supported by the School of Biological Sciences and Washington State University. Some of the figures were provided by Spear SF.

Appendix

List of Courses

1. Landscape Genetics
2. Molecular Ecology
3. Population Genetics
4. Conservation Genetics

See also: Conservation Genetics. Ecological Genetics. Landscape Ecology and Population Dynamics. Landscape Modeling. Population Genetics

References

Adriansen F, Chardon JP, de Blust G, *et al.* (2003) The application of 'least-cost' modeling as a functional landscape model. *Landscape and Urban Planning* 64: 233–247.

Anderson CD, Epperson BK, Fortin M-J, *et al.* (2010) Considering spatial and temporal scale in landscape genetics studies of gene flow. *Molecular Ecology* 19: 3565–3575.

Anderson JE (1979) A theoretical foundation for the gravity equation. *American Economic Review* 69: 106–116.

Angers B, Magnan P, Plante M, and Bernatchez L (1999) Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology* 8: 1043–1053.

Bacles CFE, Lowe AJ, and Ennos RA (2004) Genetic effects of chronic habitat fragmentation on tree species: The case of *Sorbus aucuparia* in a deforested Scottish landscape. *Molecular Ecology* 13: 573–584.

Balkenhol N, Gugerli F, Cushman S, *et al.* (2009b) Identifying future research needs in landscape genetics: Where to from here? *Landscape Ecology* 24: 455–463.

Balkenhol N and Waits LP (2009) Molecular road ecology: Exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular Ecology* 18: 4151–4164.

Balkenhol N, Waits LP, and Dezzani R (2009a) Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography* 32: 818–830.

Banks SC, Piggott MP, Williamson JE, *et al.* (2007) Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. *Ecology* 88: 3055–3064.

Barbujani G, Oden NL, and Sokal RR (1989) Detecting regions of abrupt change in maps of biological variables. *Systematic Zoology* 38: 376–389.

Barr KR, Lindsay DL, Athrey G, *et al.* (2008) Population structure in an endangered songbird: Maintenance of genetic differentiation despite high vagility and significant population recovery. *Molecular Ecology* 17: 3628–3639.

Berthier K, Galan M, Foltête JC, Charbonnel N, and Cosson JF (2006) Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): Landscape and demographic influences. *Molecular Ecology* 14: 2861–2871.

Biek R and Real LA (2010) The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* 19: 3515–3531.

Blanchong JA, Samuel MD, Scribner K, *et al.* (2008) Landscape genetics and the spatial distribution of chronic wasting disease. *Biology Letters* 4: 130–133.

Bolliger J, Keller D, and Holderegger R (2011) When landscape variables do not explain migration rates: An example from the endangered dragonfly, *Leucorrhinia caudalis* (Odonata: Libellulidae). *European Journal of Entomology* 108: 327–330.

Bonin A, Taberlet P, Miaud C, and Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution* 23: 773–783.

Borcard D, Legendre P, and Dapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73: 1045–1055.

Bowcock JM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, and Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368: 455–457.

ter Braak CJF and Verdonschot PFM (1995) Cononical correspondence analysis and related multivariate methods in aquatic ecology. *Aquatic Sciences* 5/4: 1–35.

Carnaval AC, Hickerson MJ, Haddad C, Rodrigues M, and Moritz C (2009) Stability predicts genetic diversity in the Braxilian Atlantic Forest hotspot. *Science* 323: 785–789.

Case RAJ, Hutchinson WF, Hauser L, *et al.* (2005) Macro- and micro-geographic variation in pantophysin (*Pan*I) allele frequencies in NE Atlantic cod, *Gadus morhua*. *Marine Ecology Progress Series* 301: 267–278.

Castellano S and Balletto E (2002) Is the partial mantel test inadequate? *Evolution* 56: 1871–1873.

Castric V, Bonney F, and Bernachez L (2001) Landscape structure and hierarchical genetic diversity in the brook charr, *Salvelinus fontinalis*. *Evolution* 55: 1016–1028.

Cavalli-Sforza LL and Edwards AWF (1967) Phylogenetic analysis: Models and estimation procedures. *American Journal of Human Genetics* 19: 233–257.

Cercueil A, Francois O, and Manel S (2007) The genetical bandwidth mapping: A spatial and graphical representation of population genetic structure based on the wombling method. *Theoretical Population Biology* 71: 332–341.

Costello AB, Down TE, Pollard SM, *et al.* (2003) The influence of history and contemporary stream hydrology on the evolution of genetic diversity within species: An examination of microsatellite DNA variation in bull trout, *Salvelinus confluentus*. *Evolution* 57: 328–344.

Coulon A, Guillot G, Cosson J-F, *et al.* (2006) Genetic structure is influenced by landscape features: Empirical evidence from a roe deer population. *Molecular Ecology* 15: 1669–1679.

Crida A and Manel S (2007) WOMBSOFT: A R package that implements the wombling method to identify genetic boundary. *Molecular Ecology Notes* 7: 588–591.

Cullingham CI, Kyle CJ, Pond BA, Rees EE, and White BN (2009) Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Molecular Ecology* 18: 43–53.

Cureton AN, Newbury HJ, Raybould AF, and Ford-Lloyd BV (2006) Genetic structure and gene flow in wild beet populations: The potential influence of habitat on transgene spread and risk assessment. *Journal of Applied Ecology* 43: 1203–1212.

Cushman SA and Landguth EL (2010) Spurious correlations and inference in landscape genetics. *Molecular Ecology* 19: 3592–3602.

Cushman SA, McKelvey KS, Hayden J, and Schwartz MK (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist* 168: 486–499.

Deiner K, Garza J, Coey R, and Girman D (2007) Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California. *Conservation Genetics* 8: 437–454.

Dionne M, Caron F, Dodson JJ, and Bernatchez L (2008) Landscape genetics and hierarchical genetic structure in Atlantic salmon: The interaction of gene flow and local adaptation. *Molecular Ecology* 17: 2382–2396.

Dyer RJ (2007) The evolution of genetic topologies. *Theoretical Population Biology* 71: 71–79.
Dyer RJ and Nason JD (2004) Population graphs: The graph theoretic shape of genetic structure. *Molecular Ecology* 13: 1713–1727.
Emaresi G, Pellet L, Dubey S, Hirzel AH, and Fumagalli L (2011) Landscape genetics of the alpine newt (*Mesotriton alpestris*) inferred from a strip-based approach. *Conservation Genetics* 12: 41–50.
Epperson BK, McRae BH, Scribner KT, *et al.* (2010) Utility of computer simulations in landscape genetics. *Molecular Ecology* 19: 3549–3564.
Epps CW, Wehausen JD, Bleich VC, Torres SG, and Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology* 44: 714–724.
Faubet P and Gaggiotti OE (2008) A new Bayesian method to identify the environmental factors that influence recent migration. *Genetics* 178: 1491–1504.
Faugeron S, Martínez EA, Correa JA, and Billot C (2005) Long-term copper mine waste disposal in northern Chile associated with gene flow disruption of the intertidal kelp *Lessonia nigrescens*. *Marine Ecology Progress Series* 288: 129–140.
Fernandez MJF and Sork VL (2007) Genetic variation in fragmented forest stands of the Andean oak *Quercus humboldtii* Bonpl. (Fagaceae). *Biotropica* 39: 72–78.
Foll M and Gaggiotti OE (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics* 174: 875–891.
Fontaine MC, Baird SJE, Piry S, *et al.* (2007) Rise of oceanographic barriers in continuous populations of a cetacean: The genetic structure of harbour porpoises in Old World waters. *BMC Biology* 5: 30.
Fortin M-J and Dale MRT (2005) *Spatial Analysis: A Guide for Ecologists*. Cambridge, UK: Cambridge University Press.
Francois O and Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources* 10: 773–784.
Funk WC, Blouin MS, Corn PS, *et al.* (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* 14: 483–496.
Gagneux P, Gonder MK, Goldberg TL, and Morin PA (2001) Gene flow in wild chimpanzee populations: What genetic data tell us about chimpanzee movement over space and time. *Philosophical Transactions of the Royal Society of London Series B* 356: 889–897.
Garroway CJ, Bowman J, Carr D, and Wilson PJ (2008) Applications of graph theory to landscape genetics. *Evolutionary Applications* 1: 620–630.
Garroway CJ, Bowman J, and Wilson PJ (2011) Using a genetic network to parameterize a resistance surface for fishers, *Martes pannanti*. *Molecular Ecology* 20: 3978–3988.
Giordano AR, Ridenhour BJ, and Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactylum*). *Molecular Ecology* 16: 1625–1637.
Goldberg C and Waits L (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology* 19: 3650–3663.
Guajardo JCR, Schnabel A, Ennos R, *et al.* (2011) Landscape genetics of the key African acacia species *Senegalia mellifera* (Vahl) - the importance of the Kenyan Rift Valley. *Molecular Ecology* 19: 5126–5139.
Guillot G, Leblois R, Coulon A, and Frantz A (2009) Statistical methods in spatial genetics. *Molecular Ecology* 18: 4734–4756.
Guillot G, Santos F, and Estoup A (2008) Analysing georeferenced population genetics data with Geneland: A new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* 24: 1406–1407.
Hagerty BE, Nussear KE, Esque TC, and Tracy TR (2011) Making molehills out of mountains: Landscape genetics of the Mojave desert tortoise. *Landscape Ecology* 26: 267–280.
Harrison A and Dunn R (1993) Problems of sampling the landscape. In: Haines-Young R, Green DR, and Cousins S (eds.) *Landscape Ecology and GIS*, pp. 101–109. London: Taylor and Francis.
Hirao AS and Kudo G (2004) Landscape genetics of alpine-snowbed plants: Comparisons along geographic and snowmelt gradients. *Heredity* 93: 290–298.
Hirao AS and Kudo G (2008) The effect of segregation of flowering time on fine-scale spatial genetic structure in an alpine-snowbed herb *Primula cuneifolia*. *Heredity* 100: 424–430.
Hoffman AA and Willi Y (2008) Detecting genetic responses to environmental change. *Nature Reviews Genetics* 9: 421–432.
Holderegger R and Wagner HH (2006) A brief guide to landscape genetics. *Landscape Ecology* 21: 793–796.
Holderegger R, Buehler D, Gugerli F, and Manel S (2010) Landscape genetics of plants. *Trends in Plant Science* 15: 675–683.
Holderegger R and Wagner HH (2008) Landscape genetics. *BioScience* 58: 199–207.
Holzhauer S, Ekschmitt K, Sander A-C, Dauber J, and Wolters V (2006) Effect of historic landscape change on the genetic structure of the bush-cricket *Metrioptera roeseli*. *Landscape Ecology* 21: 891–899.
Hull JM, Hull AC, Sacks BN, *et al.* (2008) Landscape characteristics influence morphological and genetic differentiation in a widespread raptor (*Buteo jamaicensis*). *Molecular Ecology* 17: 810–824.
Jacquemyn H, Honnay O, Galbusera P, and Roldán-Ruiz I (2004) Genetic structure of the forest herb *Primula elatior* in a changing landscape. *Molecular Ecology* 13: 211–219.
Jaquiéry J, Broquet T, Hirzel AH, Yearsley J, and Perrin N (2011) Inferring landscape effects on dispersal from genetic distances: How far can we go? *Molecular Ecology* 20: 692–705.
Johannson M, Primmer CR, Sahlsten J, and Merilä J (2005) The influence of landscape structure on occurrence, abundance and genetic diversity of the common frog, *Rana temporaria*. *Global Change Biology* 11: 1664–1679.
Joost S, Bonin A, Bruford MW, *et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: Towards a landscape genomics approach to adaptation. *Molecular Ecology* 16: 3955–3969.
Keyghobadi N, Roland J, Matter SJ, and Strobeck C (2005) Among- and within-patch components of genetic diversity respond at different rates to habitat fragmentation: An empirical demonstration. *Proceedings of the Royal Society B: Biological Sciences* 272: 553–560.
Kitamoto N, Honjo M, Ueno S, *et al.* (2005) Spatial genetic structure among and within populations of *Primula sieboldii* growing beside separate streams. *Molecular Ecology* 14: 149–157.
Kramer AT, Fant JB, and Ashley MV (2011) Influences of landscape and pollinators on population genetic structure: Examples from three Penstemon (Plantaginaceae) species in the Great Basin. *American Journal of Botany* 98: 109–121.
Kuehn R, Hindenlang KE, Holzgang O, *et al.* (2007) Genetic effect of transportation infrastructure on Roe deer populations (*Capreolus capreolus*). *Journal of Heredity* 98: 13–22.
Lada H, Thomson JR, Mac Nally R, and Taylor AC (2008) Impacts of massive landscape change on a carnivorous marsupial in south-eastern Australia: Inferences from landscape genetics analysis. *Journal of Applied Ecology* 45: 1732–1741.
Landguth EL and Cushman SA (2010) CDPOP: A spatially explicit cost distance population genetics program. *Molecular Ecology Resources* 10: 156–161.
Landguth EL, Schwartz MK, McKelvey M, Murphy MA, and Luikart G (2010) Quantifying lag time to detect barriers in landscape genetics. *Molecular Ecology* 19: 4179–4191.
Latch EK, Scognamillo DG, Fike JA, Chamberlain MJ, and Rhodes OE (2008) Deciphering ecological barriers to North American river otter (*Lontra canadensis*) gene flow in the Louisiana landscape. *Journal of Heredity* 99: 265–274.
Leclerc E, Mailhot Y, Mingelbier M, and Bernatchez L (2008) The landscape genetics of yellow perch (*Perca flavescens*) in a large fluvial ecosystem. *Molecular Ecology* 17: 1702–1717.
Legendre P, Lapointe F-J, and Casorian P (1994) Modeling brain evolution from behavior: A permutational regression approach. *Evolution* 48: 1487–1499.
Legendre P and Troussellier M (1988) Aquatic heterotrophic bacteria: Modeling in the presence of spatial autocorrelation. *Limnology and Oceanography* 33: 1055–1067.
Lowe WH, Likens GE, McPeek MA, and Buso DC (2006) Linking direct and indirect data on dispersal: Isolation by slope in a headwater stream salamander. *Ecology* 87: 334–339.
Lugon-Moulin N and Hausser J (2002) Phylogeographical structure, postglacial recolonization and barriers to gene flow in the distinctive Valais chromosome race of the common shrew (*Sorex araneus*). *Molecular Ecology* 11: 785–794.
Luikart G, England PR, Tallmon D, Jordan S, and Taberlet P (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics* 4: 981–994.
Mäki-Petäys H, Zakharov A, Viljakainen L, *et al.* (2005) Genetic changes associated to declining populations of *Formica* ants in fragmented forest landscape. *Molecular Ecology* 14: 733–742.
Manel S, Berthoud F, Bellemain E, *et al.* (2007) A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Molecular Ecology* 16: 2031–2043.
Manel S, Gaggiotti O, and Waples R (2005) Assignment methods: Which approaches best address which biological questions? *Trends in Ecology and Evolution* 20: 136–142.
Manel S, Joost S, Epperson BK, *et al.* (2010a) Perspectives on the use of landscape genetics to detect adaptive genetic variation in the field. *Molecular Ecology* 19: 3670–3772.

Manel S, Poncet B, Buehler D, *et al.* (2010b) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Molecular Ecology* 19: 3824–3835.
Manel S, Schwartz MK, Luikart G, and Taberlet P (2003) Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution* 18: 189–197.
Manier MK and Arnold SJ (2006) Ecological correlates of population genetic structure: A comparative approach using a vertebrate metacommunity. *Proceedings of the Royal Society B* 273: 3001–3009.
Manni F, Guerard E, and Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology* 76: 173–190.
Marko PB and Hart MW (2011) The complex analytical landscape of gene flow inference. *Trends in Ecology and Evolution* 26: 448–456.
Martinez-Solano I and Gonzalez EG (2008) Patterns of gene flow and source-sink dynamics in high altitude populations of the common toad *Bufo bufo* (Anura: Bufonidae). *Biological Journal of the Linnean Society* 95: 824–839.
McRae BH (2006) Isolation by resistance. *Evolution* 60: 1551–1561.
McRae BH and Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the USA* 104: 19885–19890.
McRae BH, Dickson BG, Keitt TH, and Shah VB (2008) Using circuit theory to model connectivity in ecology and conservation. *Ecology* 10: 2712–2724.
Michels E, Cottenie K, Neys L, *et al.* (2001) Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: A plea for using GIS modelling of the effective geographical distance. *Molecular Ecology* 10: 1929–1938.
Millions DG and Swanson BJ (2007) Impact of natural and artificial barriers to dispersal on the population structure of bobcats. *Journal of Wildlife Management* 71: 96–102.
Mockford SW, Herman TB, Snyder M, and Wright JM (2007) Conservation genetics of Blanding's turtle and its application in the identification of evolutionarily significant units. *Conservation Genetics* 8: 209–219.
Monmonier MS (1973) Maximum-difference barriers: An alternative numerical regionalization method. *Geographic Analysis* 5: 245–261.
Murphy MA, Dezanni RJ, Pilliod D, and Storfer A (2010) Landscape genetics of high mountain frog metapopulations. *Molecular Ecology* 19: 3634–3649.
Murphy MA, Evans JS, Cushman S, and Storfer A (2008) Representing genetic variation as continuous surfaces: An approach for identifying spatial dependency in landscape genetic studies. *Ecography* 31: 685–697.
Murphy MA, Evans JS, and Storfer A (2010) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91: 252–261.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 409: 853–858.
Narum SR, Zendt JS, Graves D, and Sharp WR (2008) Influence of landscape on resident and anadromous life history types of *Oncorhynchus mykiss*. *Canadian Journal of Fisheries and Aquatic Sciences* 65: 1013–1023.
Nei M (1972) Genetic distance between populations. *American Naturalist* 106: 283–292.
Olsen JB, Templin WD, Crane PA, *et al.* (2011) Comparative landscape genetic analysis of three Pacific salmon species from subarctic North America. *Conservation Genetics* 12: 223–241.
Orsini L, Corander J, Alasentie A, and Hanski I (2008) Genetic spatial structure in a butterfly metapopulation correlates better with past than present demographic structure. *Molecular Ecology* 17: 2629–2642.
Pavlacky DC, Goldzien AW, Prentis PJ, *et al.* (2009) A landscape genetics approach for quantifying the relative influence of historic and contemporary habitat heterogeneity on a rainforest bird. *Molecular Ecology* 18: 2945–2960.
Peakall R and Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.
Pérez-Espona S, Perez-Barberia FJ, McLeod JE, *et al.* (2008) Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). *Molecular Ecology* 17: 981–996.
Petren K, Grant PR, Grant BR, and Keller LF (2005) Comparative landscape genetics and the adaptive radiation of Darwin's finches: The role of peripheral isolation. *Molecular Ecology* 14: 2943–2957.
Pritchard JK, Stephens P, and Donnelly P (2000) Inference of population genetic structure using multilocus genotype data. *Genetics* 155: 945–959.
Radespiel U, Rakotondravony R, and Chikhi L (2008) Natural and anthropogenic determinants of genetic structure in the largest remaining population of the endangered golden-brown mouse lemur, *Microcebus ravelobensis*. *American Journal of Primatology* 70: 860–870.
Raeymaekers JAM, Maes GE, Geldof S, *et al.* (2008) Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evolutionary Applications* 1: 475–488.
Rasic G and Keyghobadi N (2012) From broadscale patterns to fine-scale processes: Habitat structure influences genetic differentiation in the pitcher plant midge across multiple spatial scales. *Molecular Ecology* 21: 223–236.
Rayfield B, Fortin M-J, and Fall A (2010) The sensitivity of least-cost habitat graphs to relative cost surface values. *Landscape Ecology* 25: 519–532.
Real LA and Biek R (2007) Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes. *Journal of the Royal Society Interface* 4: 935–948.
Reichman OJ (2004) NCEAS: Promoting creative collaborations. *Public Library of Science Biology* 2: 311–312.
Rieux A, Halkett F, de Bellaire LD, *et al.* (2011) Inferences on pathogenic fungus population structures from microsatellite data: New insights from spatial genetic approaches. *Molecular Ecology* 20: 1661–1674.
Riley SPD, Pollinger JP, Sauvajot RM, *et al.* (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* 15: 1733–1741.
Rittenhouse TAG and Semlitsch RD (2006) Grasslands as movement barriers for a forest-associated salamander: Migration behavior of adult and juvenile salamanders at a distinct habitat edge. *Biological Conservation* 131: 14–22.
Rousset F (2002) Partial mantel tests: Reply to Castellano and Balletto. *Evolution* 56: 1874–1875.
Rowe G and Beebee TJC (2007) Defining population boundaries: Use of three Bayesian approaches with microsatellite data from British natterjack toads (*Bufo calamita*). *Molecular Ecology* 16: 785–796.
Rundle HD and Nosil P (2005) Ecological speciation. *Ecology Letters* 8: 336–352.
Sacks BN, Mitchell BR, Williams CL, and Ernest HB (2004) Coyote movements and social structure along a cryptic population genetic subdivision. *Molecular Ecology* 14: 1241–1249.
Safner T, Miller MP, McRae BH, Fortin M-J, and Manel S (2011) Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences* 12: 865–889.
Segelbacher G, Cushman SA, Epperson BK, *et al.* (2010) Applications of landscape genetics in conservation biology: Concepts and challenges. *Conservation Genetics* 11: 375–385.
Selkoe K, Watson J, White C, *et al.* (2010) Taking the chaos out of genetic patchiness: Seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology* 19: 3708–3726.
Shah VB and McRae BH (2008) Circuitscape: A tool for landscape ecology. In: Varoquaux G, Vaught T, and Millman J (eds.) *Proceedings of the Seventh Python in Science Conference* (*SciPy* 2008), pp. 62–66. Pasadena, CA, USA.
Short-Bull RA, Cushman SA, *et al.* (2011) Why replication is important in landscape genetics: American black bear in the Rocky Mountains. *Molecular Ecology* 20: 1092–1107.
Skalski GT, Landis JB, Grose MJ, and Hudman SP (2008) Genetic structure of creek chub, a headwater minnow, in an impounded river system. *Transactions of the American Fisheries Society* 137: 962–975.
Spear SF, Balkenhol N, Fortin M-J, McRae B, and Scribner KT (2010) Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology* 19: 3576–3591.
Spear SF, Peterson CR, Matacq M, and Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology* 14: 2553–2564.
Spear SF and Storfer A (2008) Landscape genetic structure of coastal tailed frogs (*Ascaphus truei*) in protected vs. managed forests. *Molecular Ecology* 17: 4642–4656.
Spear SF and Storfer A (2010) Anthropogenic and natural disturbance lead to differing patterns of gene flow in the Rocky Mountain tailed frog, *Ascaphus montanus*. *Biological Conservation* 143: 778–786.
Stinchcombe JR and Hoekstra HE (2008) Combining population genomics and quantitative genetics: Finding the genes underlying ecologically important traits. *Heredity* 100: 158–170.
Storfer A, Murphy MA, Evans JS, *et al.* (2007) Putting the 'landscape' in landscape genetics. *Heredity* 98: 128–142.
Storfer A, Murphy MA, Spear SF, Holderegger R, and Waits LP (2010) Landscape genetics: Where are we now? *Molecular Ecology* 19: 3496–3514.
Suarez-Montes P, Fornoni J, and Nunez-Farfan J (2011) Conservation genetics of the endemic Mexican *Heliconiaux panapensis* in the Los Tuxlas Tropical Rainforest. *Biotopica* 43: 114–121.

Taylor EB, Stamford MD, and Baxter JS (2003) Population subdivision in westslope cutthroat trout (*Oncorhynchus clarki lewisi*) at the northern periphery of its range: Evolutionary inferences and conservation implications. *Molecular Ecology* 12: 2609–2622.

Theodorakis CW, Lee K-L, Adams SM, and Law CB (2006) Evidence of altered gene flow, mutation rate, and genetic diversity in redbreast sunfish from a pulp-mill contaminated river. *Environmental Science and Technology* 40: 377–386.

Trénel P, Hansen MM, Normand S, and Borchsenius F (2008) Landscape genetics, historical isolation and cross-Andean gene flow in the wax palm, *Ceroxylon echinulatum* (Arecaceae). *Molecular Ecology* 17: 3528–3540.

Vandepitte K, Jacquemin H, Roldán-Ruiz I, and Honnay O (2007) Landscape genetics of the self-compatible forest herb *Geum urbanum*: Effects of habitat age, fragmentation and local environment. *Molecular Ecology* 16: 471–479.

Wagner HH, Murphy MA, Holderegger R, and Waits. LP (2012) Developing an interdisciplinary, distributed graduate course for 21st century scientists. *Bioscience* 62: 182–188.

Wagner HH, Werth S, Kalwij J, Bolli J, and Scheidegger C (2006) Modelling forest recolonization by an epiphytic lichen using a landscape genetic approach. *Landscape Ecology* 21: 849–865.

Wang IJ and Summers K (2010) Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* 19: 447–458.

Wang YH, Yang KC, Bridgman CL, and Lin LK (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology* 23: 989–1000.

Waples RS and Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying gene pools an their degree of connectivity. *Molecular Ecology* 15: 1419–1439.

Was A, Gosling E, McCrann K, and Mork J (2008) Evidence for population structuring of blue whiting (*Micromesistius poutassou*) in the Northeast Atlantic. *ICES Journal of Marine Science* 65: 216–225.

Wellenreuther M, Sanchez-Guillen RA, Cordero-Rivera A, Svennson EI, and Hansson B (2011) Environmental and climatic determinants of molecular diversity and genetic population structure in a coenagrionid damselfly. *PLoS ONE* 6(6): e20440.

Whiteley AR, Spruell P, and Allendorf FW (2006) Can common species provide valuable information for conservation? *Molecular Ecology* 15: 2767–2786.

Whitlock MC and McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity* 82: 117–125.

Wilcock HR, Bruford MW, Nichols RA, and Hildrew AG (2007) Landscape, habitat characteristics and the genetic population structure of two caddisflies. *Freshwater Biology* 52: 1907–1929.

Willig RD and Bailey EE (1979) The economic gravity model. *American Economic Review* 69: 96–101.

Wilmer JW, Elkin C, Wilcox C, *et al.* (2008) The influence of multiple dispersal mechanisms and landscape structure on population clustering and connectivity in fragmented artesian spring snail populations. *Molecular Ecology* 17: 3733–3751.

Wofford JEB, Gresswell RE, and Banks MA (2005) Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. *Ecological Applications* 15: 628–637.

Womble WH (1951) Differential systematics. *Science* 114: 315–322.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* 15: 323–354.

Zalewski A, Piertney SB, Zalewska H, and Lambin X (2009) Landscape barriers reduce gene flow in an invasive carnivore: Geographical and local genetic structure of American mink in Scotland. *Molecular Ecology* 18: 1601–1615.

Zartman CE, McDaniel SF, and Shaw AJ (2006) Experimental habitat fragmentation increases linkage disequilibrium but does not affect genetic diversity or population structure in the Amazonian liverwort, *Radula flaccida*. *Molecular Ecology* 15: 2305–2315.

Zhu L, Zhang S, Gu X, and Wei F (2011) Significant genetic boundaries and spatial dynamics of giant pandas occupying fragmented habitat across southwest China. *Molecular Ecology* 20: 1122–1132.

Landscape Legacies

Jennifer M Fraterrigo, University of Illinois, Urbana, IL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Carbon (or nutrient)-accumulation rate The rate at which carbon or nutrients, such as nitrogen and phosphorus, accrue in biomass and soils.
Deforestation The most intense form of timber extraction, wherein forests are converted for another land use either permanently or for a long term.
Landscape legacy The long-term impact of historical events on ecosystem structure, composition, or function.
Natural disturbances Any relatively discrete event in time that is not directly caused by humans; disrupts ecosystem, community, or population structure; and changes resources, substrate availability, or the physical environment (White and Pickett, 1985).
Nitrification The step in the nitrogen cycle wherein ammonia is oxidized and converted to nitrite and then oxidized again to form nitrate.
Plow horizon The homogenous soil layer, typically at 0–20 cm depth, that results from plowing. Designated as the "Ap" horizon in soil morphology.
Selective logging A less intense form of timber extraction involving removal of select trees, often based on species, size, or age class. The forest canopy is typically maintained under this process.
Site history The collection of past events involving natural disturbance or land use conversion at a given site.
Soil organic matter (SOM) depletion The loss of organic matter from the soil. The main constituent of SOM is soil organic carbon. Consequently, this measurement relates directly to soil organic carbon loss and storage.

Introduction

At the start of the twenty-first century, appreciation for the importance of site history was emerging, yet most ecological studies tended to focus on the here and now, and many ecologists were unconcerned with the events that occurred in the distant past. There was a general perception that site history was unimportant for understanding the patterns and processes of modern ecosystems (Foster *et al.*, 2003). However, over the past two decades, ecologists have come to recognize that historical events can have lasting impacts on ecosystem structure, composition, and function. Indeed, many studies now show that site history can explain a substantial amount of the variation in contemporary ecological patterns and processes in both terrestrial and aquatic ecosystems. The strong, lingering effects of historical events on ecosystems are known as landscape legacies.

The duration of such impacts can be surprisingly long, particularly for ecosystem properties that are physically affected by past events. Soil and vegetation communities, as well as the processes they influence (e.g., nutrient cycling and primary production), often retain the imprints of past events for decades to centuries. In several cases, recovery from past events has not yet been observed, thereby leading the scientists to ask whether some ecological changes might persist indefinitely (Peterken and Game, 1984). In areas of northeastern France deforested during the Roman occupation and farmed during AD 50–250, species richness and plant communities of modern forests, as well as the chemical and structural properties of soil, still varied for more than 1500 years later with the intensity of former agriculture (Dupouey *et al.*, 2002). Similarly, modern forest composition, microtopography, soil texture, and the lake and wetland sediment deposits are strongly associated with the Mayan practice of deforestation in AD 700–900 (Turner *et al.*, 2003). Changes brought about by past events can also act indirectly to produce lasting alterations in animal assemblages. In southeastern USA, fish and stream invertebrate communities are less diverse today because historical agriculture has persistently increased the amount of sediments in nearby streams (Harding *et al.*, 1998). Examples such as these have called into question the long-standing assumptions about trajectories of ecosystem recovery and compelled ecologists and managers to rethink of traditional concepts about natural or reference ecosystem states, as well as the controls on species distribution and abundance (Foster *et al.*, 2003; Flinn and Vellend, 2005). The knowledge gained from the studies of landscape legacies could, therefore, be important for directing management policies and more accurately predicting how biodiversity will change in the future.

Drivers of Landscape Legacies

Land Use

Human land use is a primary driver of landscape legacies. Humans have long used the land in ways that would modify the native ecosystems, and a growing body of evidence demonstrates that these changes can persist even long after human activities cease. Agriculture is the quintessential example, having been practiced in the form of cultivation and livestock husbandry since plants and animals were domesticated 10,000 years ago (Diamond, 2002). Although agricultural land use is expanding in some regions, other regions have experienced widespread agricultural abandonment (Ramankutty and Foley, 1999). For example, many modern forests in Western Europe and eastern North America existed as farms or pastures during the nineteenth and twentieth centuries (Rackham,

1980; Peterken and Game, 1984; Foster, 1992; Verheyen *et al.*, 1999).

Despite long periods without additional disturbance, most lands with an agricultural history have characteristics distinct from those of lands without a similar disturbance history. Such agricultural legacies have been observed in a broad range of native ecosystems, including temperate mixed and deciduous forests in Europe and eastern North America (Motzkin *et al.*, 1996; Foster *et al.*, 1998; Compton and Boone, 2000; Richter *et al.*, 2000; Hermy and Verheyen, 2007), grasslands and arid lands in the central and western USA (Burke *et al.*, 1995; Coffin *et al.*, 1996; Lewis *et al.*, 2006), and in tropical forests (McGrath *et al.*, 2001; Thompson *et al.*, 2002; Turner *et al.*, 2003). Although agricultural legacies are pervasive, their longevity and magnitude vary because the intensity and duration of agriculture, and specific management practices attending to agriculture, can differ widely (Coffin *et al.*, 1996; Compton *et al.*, 1998; McLauchlan, 2006). Empirical studies and meta-analyses demonstrate that the effects of cultivation and pasturing on soil carbon and nitrogen pools, for example, can vary markedly depending on farm-residue management, crop rotation, and tillage practices (Fearnside and Barbosa, 1998; Murty *et al.*, 2002). Despite this variation, cultivated lands consistently show more persistent and profound legacies than pastured lands due to the high, sustained intensity of cropping (Compton and Boone, 2000; McGrath *et al.*, 2001; Gerhardt and Foster, 2002; Brown and Boutin, 2009).

Timber extraction can also generate landscape legacies. These legacies tend to be smaller in magnitude and of shorter duration than those attending agriculture, but understanding the legacies of timber extraction is vital due to global importance and long history (spanning for more than two millennia) of this practice (Williams, 2000; FAO, 2006; Hansen *et al.*, 2010). Timber extraction, like agriculture, is practiced with varying intensity. Deforestation (or clearcutting) is the most detrimental of such practices, characterized by wholesale forest clearance and substantial soil disturbance. Selective logging represents the opposite extreme and is characterized by the removal of a few individuals (often economically desirable species) and minimal soil disturbance. The extent of land conversion can be substantial under practice (Asner *et al.*, 2005), but selective logging is generally thought to have less appreciable and short-term effects compared with deforestation (Williams, 1990). Indeed, in studies examining the impacts of historical land use, former selectively logged stands are frequently used to characterize reference ecological conditions when undisturbed stands are unavailable.

Forest loss associated with deforestation has occurred heterogeneously over the past 2000 years. Europe and Asia experienced significant deforestation during both pre-1650 and from 1750 to 1849 but are now undergoing afforestation in many parts (Williams, 1990; FAO, 2006). Extensive forest clearance in eastern North America began in the early nineteenth century and has been followed by rapid afforestation. Today, the regions suffering the largest losses due to deforestation are in South America and Africa, where forests are being cleared for cropping or livestock grazing (FAO, 2006). Tropical forest recovery is of particular concern because of its potential to impact global carbon and nutrient budgets and plant and animal biodiversity.

Detecting land-use legacies is often challenging because of preexisting differences in site conditions. From a statistical standpoint, variation in the initial conditions of sites with similar histories may be large enough to swamp variation associated with previous land use. Legacies may still be present within a given site but may be statistically undetectable because the signal is weak relative to that produced by the physical gradients. Relying solely on mean values to evaluate the effects of land-use history may therefore be misleading, and patterns of spatial heterogeneity or structure may yield a more comprehensive depiction of enduring changes (Fraterrigo *et al.*, 2005, 2006b). Initial site conditions can also confound detection of historical land-use effects by biasing land use. For example, sites with good drainage were favored for cultivation relative to sites with poor drainage, which were avoided. As a result, ecological patterns on historically cultivated and uncultivated lands may not be comparable. Approaches for dealing with this problem include pairing or blocking sites with different histories to control for environmental variation.

Natural Disturbance

Natural disturbances can also leave legacies that persist for decades to centuries. For example, research conducted across a chronosequence of burned lodgepole pine stands in Yellowstone National Park indicates that large, stand-replacing fires contribute to variation in stand density and growth rate for nearly two centuries (Kashian *et al.*, 2005a, 2005b). Previous disturbance history can also influence the severity of future disturbances. In the subalpine forests of western USA, spruce beetle outbreaks appear to reduce the susceptibility of forest stands to low-severity fires by elevating forest floor moisture (Kulakowski *et al.*, 2003), whereas previous blow downs result in more severe burns decades later (Kulakowski and Veblen, 2007).

Despite the potential for natural disturbances to generate considerable variation in ecological patterns and processes, evidence suggests that ensuing dynamics will be constrained by patterns initiated by past land use. In tropical forests of Puerto Rico, successional vegetation trajectories following hurricanes, landslides, and local treefalls correspond more to land use in the nineteenth and early twentieth centuries than to the severity of recent natural disturbances (Thompson *et al.*, 2002). Similarly, patterns of plant and soil variation on New England sand plains closely relate to historic variation in nineteenth-century agriculture, and repeated episodes of fire have had little effect (Motzkin *et al.*, 1996; Eberhardt *et al.*, 2003). These studies reinforce the need to consider historical context when interpreting disturbance effects.

Legacies in Plant Communities

What are the dominant legacies of human land use and natural disturbance on plant communities, and what factors influence them? Research conducted in temperate forest ecosystems demonstrates that previous land use strongly controls the species composition of contemporary forests

(Foster, 1992; Motzkin *et al.*, 1996; Hermy and Verheyen, 2007). Overstory composition of forests in eastern and mid-western USA has shifted has shifted from long-lived and shade-tolerant species to more rapidly growing and short-lived species as a result of historical agriculture and logging (Foster *et al.*, 1998; Rhemtulla *et al.*, 2009). In southern New England, for example, primary forests are dominated by *Acer saccharum*, *Acer rubrum*, *Fagus grandifolia*, and *Tsuga canadensis*, whereas secondary forests abandoned from agriculture a century ago are dominated by *A. rubrum* and *Pinus strobus* (Flinn and Marks, 2007). The species attaining dominance in secondary forests are generally those that sprout effectively or aggressively invade cleared or cultivated lands (Eberhardt *et al.*, 2003; Foster *et al.*, 2003). Studies from tropical and temperate forests demonstrate that non-native invaders, in particular, can quickly become dominant in anthropogenically disturbed stands and persistently preclude re-establishment of native species (Brown and Gurevitch, 2004; DeGasperis and Motzkin, 2007). Selection for species with disturbance-adapted traits has also contributed to the homogenization of overstory species composition at a regional scale (Schulte *et al.*, 2007; Rhemtulla *et al.*, 2009). Such is the case in New England, a region characterized by a history of agriculture, logging, and reforestation (Foster *et al.*, 1998). Paleoecological data indicate that pre-European patterns of forest vegetation that corresponded with differences in climate, substrate, and fire regime have been obscured by European settlement, which has produced a more homogenous regional flora (Fuller *et al.*, 1998). Similarly, the density, age, and size structures of contemporary forests with a history of anthropogenic or intensive natural disturbance often have a more uniform distribution relative to old-growth forests (Goodale and Aber, 2001; Kashian *et al.*, 2005b; Rhemtulla *et al.*, 2009).

Long-term changes in overstory composition and structure can result in lasting changes on the forest floor. Cultivation is generally associated with decreases in foliar litter mass and carbon content due to increased decomposition rates that can accompany the shift to tree species with more degradable litter (i.e., lower foliar C:N) (Compton and Boone, 2000). Working across a chronosequence of formerly farmed sites, Hooker and Compton (2003) estimated that forest-floor carbon content may take over a century to recover after land abandonment. Effects of logging are more equivocal (Yanai *et al.*, 2003), with both increases and decreases in foliar litter carbon content reported (Covington, 1981; Goodale and Aber, 2001; Latty *et al.*, 2004). There is evidence that variation in the magnitude and longevity of logging effects can be explained by differences in logging intensity, which is determined not only by the amount of biomass removed but also by the degree of mechanical disturbance (Yanai *et al.*, 2003). Likewise, logging combined with fire, which often ignited slash piles left on-site after harvests, produce more lasting changes due to increased disturbance intensity (Goodale and Aber, 2001; Latty *et al.*, 2004).

Changes in the amount and size distribution of coarse woody debris are another legacy of land use and natural disturbance. Work conducted at Harvard Forest shows that 73-year-old pine stands planted on abandoned agricultural lands have slightly higher masses and a uniform distribution of smaller size class coarse woody debris compared with deciduous stands without an agricultural history (Currie and Nadelhoffer, 2002). Naturally reforested lands accumulate woody debris more slowly – at a rate of 0.05 mg ha^{-1} $year^{-1}$ – and thus may require more than a century to return to pre-agricultural levels (Hooker and Compton, 2003). Similarly, intensive logging by itself or combined with fire has led to long-term decreases in coarse woody debris, especially in larger size classes, both as a direct result of tree removal and persistent changes in site quality or successional status (McCarthy and Bailey, 1994; Sturtevant *et al.*, 1997; Gough *et al.*, 2007). Fire alone removes considerably less coarse woody debris than logging and, consequently, its effects are smaller in magnitude and shorter in duration (Tinker and Knight, 2000).

Herbaceous plant communities are among the most severely impacted by past land use activity. In European and North American temperate forests, previous agriculture has led to striking changes in understory assemblages, a prominent and consistent feature of which is the reduced abundance or absence of native understory species for a century or more (Peterken and Game, 1984; Flinn and Vellend, 2005; Hermy and Verheyen, 2007). In contrast, most studies indicate that recovery from logging occurs within decades of disturbance (Halpern and Spies, 1995; Roberts and Gilliam, 2003), although long-term effects have been found (Duffy and Meier, 1992). There is also the potential for wide variation in recovery. This is illustrated by semiarid grassland plant communities cultivated half a century ago, where some old fields are compositionally similar to uncultivated grasslands and some are distinct (Coffin *et al.*, 1996). Regardless, these studies have led to important insights about the factors controlling ecosystem response to disturbance. It is clear that life-history traits play a key role in the colonization dynamics of disturbed lands (Verheyen *et al.*, 2003b). Overwhelming evidence supports the hypothesis that species mobility (i.e., dispersal capacity) determines the colonization of disturbed lands (Dzwonko and Loster, 1992; Matlack, 1994; Brunet and von Oheimb, 1998; Bossuyt *et al.*, 1999; Bellemare *et al.*, 2002; Verheyen *et al.*, 2003a). Not surprisingly, the spatial arrangement of disturbed relative to undisturbed lands is important for predicting the nature and rate of species colonization and interacts with dispersal capacity to affect community recovery via its influence on seed availability (Dzwonko and Loster, 1992; Brunet and von Oheimb, 1998; Bossuyt *et al.*, 1999; Butaye *et al.*, 2002). Working in young, secondary stands that varied in isolation from older, species-rich stands, Matlack (1994) demonstrated that diversity increases with proximity to source populations, mainly due to increased representation of poorly dispersing species. The spatial configuration of disturbed habitat relative to the source populations is also important for recovery after large, intense natural disturbances such as fire and flooding (Turner *et al.*, 1998).

Dispersal is the primary factor explaining the colonization of historically altered lands, but site conditions can also affect the distribution of herbaceous plant species by influencing plant recruitment. Indeed, variation in the establishment of species with similar dispersal abilities (Singleton *et al.*, 2001) suggests that site conditions should not be overlooked. Assembly of herbaceous plant communities on disturbed lands likely occurs as a two-stage process in which restricted seed

availability followed by a low recruitment limits colonization (Verheyen *et al.*, 2003a; Baeten *et al.*, 2009a, 2009b). Yet it remains unclear which site conditions affect recruitment. For some species, a lack of microtopographic variation or microsite disturbance hampers establishment (Verheyen and Hermy, 2004; Flinn, 2007; Baeten *et al.*, 2009b), whereas other species appear to be influenced by edaphic characteristics (Verheyen *et al.*, 2003a). Undoubtedly, a species' life-history traits influence its requirements for establishment, which makes it challenging to identify factors that influence all species equivalently. Another complicating factor is that effects of past land use may vary with plant life stage. Several studies demonstrate that previous agriculture has neutral to positive effects on seed germination and adult plant performance (Donohue *et al.*, 2000; Singleton *et al.*, 2001; Endels *et al.*, 2004; Verheyen and Hermy, 2004; Fraterrigo *et al.*, 2006c; Flinn, 2007, but see Vellend, 2005), but negative effects on seedling and juvenile survival and growth, as well as juvenile and adult survival (Endels *et al.*, 2004; Jacquemyn and Brys, 2008). The mechanisms underlying these patterns are poorly understood. One hypothesis is that due to competition from early colonizing, disturbance-adapted species reduces the ability of native species (many of which are slow-migrating perennials) to acquire resources, thereby creating bottlenecks during some stages (De Keersmaeker *et al.*, 2004). Further work is needed to evaluate the role of these and other factors.

Legacies in Soils

Land use, especially agriculture, can impose long-term changes on the physical, chemical, and biological properties of soils. A plow horizon, that is, the depth at which plowing homogenizes the soil, is one of the most apparent legacies of previous cultivation and may persist for centuries after land abandonment (Motzkin *et al.*, 1996). In addition to being visibly distinct, plow horizons are often depleted of carbon (C) and nitrogen (N) due to accelerated decomposition, reduced organic matter inputs, degradation of soil physical structure, and erosion. Labile or reactive soil C and N fractions in the plow layer can see even greater losses than total C and N pools because they are highly degradable and slower to recover (Compton and Boone, 2000). Such conditions can in turn affect soil microbial communities (Buckley and Schmidt, 2001) and contribute to changes in carbon and nutrient cycling (see legacies in ecosystem processes; Compton and Boone, 2000; Fraterrigo *et al.*, 2006a).

However, not all sites subjected to past cultivation show soil C or nutrient depletion. The specific impacts of cultivation vary with soil amendments, such as liming or fertilization, and crop management (Compton *et al.*, 1998; Knops and Tilman, 2000; Richter *et al.*, 2000). Effects on soil organic carbon, for example, depend on the amount and type of organic amendment, method of incorporation, and size and length of time of application (McLauchlan, 2006). As a result, soil C and nutrient content can increase, decrease, or show no change following agricultural land use. The stability of a chemical pool can also influence the longevity of impacts. Phosphorus, for instance, is relatively biogeochemically stable compared with C and N and often remains substantially altered well after other chemical pools have returned to precultivation levels (Honnay *et al.*, 1999; Compton and Boone, 2000; Dupouey *et al.*, 2002; Grossmann and Mladenoff, 2008).

Relative to agriculture, logging and natural disturbance have a more limited impact on soils. Mineral soil C and N content of sites logged approximately 100 years ago is comparable with that of old-growth forests, although logging and fire together can slightly reduce soil C and N (Goodale and Aber, 2001; Latty *et al.*, 2004; Gough *et al.*, 2007). Regional studies of disturbance impacts on soil C and nutrient content suggest that the long-term effects are greatest following agriculture, followed by logging and fire (Grossmann and Mladenoff, 2008).

Despite slow rates of accumulation, depleted soil C and N pools recover eventually in most ecosystems (Knops and Tilman, 2000; Hooker and Compton, 2003; Matlack, 2009). The factors that influence C and N recovery operate via their effects on soil development and C and nutrient inputs and include vegetation recovery, climate, and soil texture and mineralogy (Burke *et al.*, 1995; Ihori *et al.*, 1995; Knops and Tilman, 2000; Post and Kwon, 2000; Springob *et al.*, 2001; McLauchlan, 2006). Because these factors differ across sites, rates of recovery vary widely within and among the ecosystems (McLauchlan, 2006). Empirical and modeling work indicate that active soil organic matter in once-cultivated grasslands returns to precultivation levels within 30–50 years, whereas total soil C may take 100–150 years to recover (Burke *et al.*, 1995; Baer *et al.*, 2002; McLauchlan *et al.*, 2006). Similarly, accrual rates of soil C in formerly cultivated temperate forests range from 1.5 to 5.8 g C m^{-2} $year^{-1}$, which suggests at least a 100-year recovery period (Hooker and Compton, 2003; Paul *et al.*, 2003). Climate and soil texture may ultimately constrain ecosystem C accrual by regulating biomass production, but more work is needed to evaluate how this relates to soil C accumulation (Johnson *et al.*, 2000; Johnson and Curtis, 2001).

Ecosystem Process Legacies

Persistent changes in soil properties can result in altered ecosystem processes. Globally, primary production is controlled by climate and soil texture (Johnson *et al.*, 2000), but within regions, previous land use can influence site quality, thereby reducing the rates of growth and biomass accumulation. These effects are most severe under repeated disturbance or when the duration of previous land use is long. For example, forest stands in the Great Lakes region that were disturbed twice by harvest and fire stored on an average 45% less C annually than those receiving the same disturbance only once due to reduced site quality (Gough *et al.*, 2007). Similarly, pine stands in southeastern USA that were previously farmed for cotton and other crops show acute N deficiency and reduced growth, despite fertilization during cropping and 50–125 years of forest development (Richter *et al.*, 2000).

Previous land use can also result in changes in N cycling and retention. Northern hardwood forests logged and burned 80–110 years ago have lower nitrification rates and export less nitrate to adjacent streams compared with old-growth forests

due to an increase in the strength of the biomass N sink (Goodale *et al.*, 2000; Goodale and Aber, 2001). By contrast, cultivation is associated with increased nitrification rates. The reasons for this are unclear, but one hypothesis is that increased nitrification is related to soil organic matter depletion. As soil microbes become substrate-limited, they may increase the rate at which they degrade soil organic matter and nitrify more N in the process (Compton and Boone, 2002).

Implications

Given the strong influence of historical events on ecosystem structure and function, it is clear that understanding the land use and disturbance legacies is fundamentally important for managing and restoring ecosystems. In many cases, ecosystem responses to current landscape-change drivers cannot be interpreted without knowing the historical context. For example, experimental and observational studies indicate that the N saturation status of forests cannot be accurately predicted without knowing the forest history (Aber *et al.*, 1998; McLauchlan *et al.*, 2007). Rates of C accrued and sequestered in biomass also depend on land-use history, making it an important factor in explaining global C dynamics (Caspersen *et al.*, 2000). Past events can influence the success of ecosystem restoration activities by altering the pool of species available for reintroduction. One result might be that the restoration activities such as prescribed burning have little effect on the patterns imposed by past land use (Motzkin *et al.*, 1996). The composition of disturbed and undisturbed communities may also be so different as to respond to management activities in unique ways (Brudvig and Damschen, 2011). More work is needed to develop strategies for managing landscapes with diverse histories.

Conclusions

Land use and disturbance legacies can be detected in plant communities, soils, and ecosystem processes, resulting in altered ecosystem structure, composition, and function. Recognizing historical legacies aides in interpreting contemporary ecosystem patterns and responses to modern landscape-change drivers. For these reasons, landscape legacies should be viewed as an integral component of ecological science and landscape management.

See also: Deforestation and Land Clearing. Forest Ecology. Human Impact on Biodiversity, Overview. Land-Use Patterns, Historic

References

Aber J, McDowell W, Nadelhoffer K, *et al.* (1998) Nitrogen saturation in temperate forest ecosystems: Hypotheses revisited. *BioScience* 48: 921–934.

Asner GP, Knapp DE, Broadbent EN, Oliveira PJC, Keller M, and Silva JN (2005) Selective logging in the Brazilian Amazon. *Science* 310: 480–482.

Baer SG, Kitchen DJ, Blair JM, and Rice CW (2002) Changes in ecosystem structure and function in a chronosequence of grasslands restored through the Conservation Reserve Program. *Ecological Applications* 12: 1688–1701.

Baeten L, Hermy M, and Verheyen K (2009a) Environmental limitation contributes to the differential colonization capacity of two forest herbs. *Journal of Vegetation Science* 20: 209–223.

Baeten L, Jacquemyn H, Van Calster H, *et al.* (2009b) Low recruitment across life stages partly accounts for the slow colonization of forest herbs. *Journal of Ecology* 97: 109–117.

Bellemare J, Motzkin G, and Foster DR (2002) Legacies of the agricultural past in the forested present: An assessment of historical land-use effects on rich mesic forests. *Journal of Biogeography* 29: 1401–1420.

Bossuyt B, Hermy M, and Deckers J (1999) Migration of herbaceous plant species across ancient-recent forest ecotones in central Belgium. *Journal of Ecology* 87: 628–638.

Brown CD and Boutin C (2009) Linking past land use, recent disturbance, and dispersal mechanism to forest composition. *Biological Conservation* 142: 1647–1656.

Brown KA and Gurevitch J (2004) Long-term impacts of logging on forest diversity in Madagascar. *Proceedings of the National Academy of Sciences of the United States of America* 101: 6045–6049.

Brudvig LA and Damschen EI (2011) Land-use history, historical connectivity, and land management interact to determine longleaf pine woodland understory richness and composition. *Ecography* 34: 257–266.

Brunet J and von Oheimb G (1998) Migration of vascular plants to secondary woodlands in southern Sweden. *Journal of Ecology* 86: 429–438.

Buckley DH and Schmidt TM (2001) The structure of microbial communities in soil and the lasting impact of cultivation. *Microbial Ecology* 42: 11–21.

Burke IC, Lauenroth WK, and Coffin DP (1995) Soil organic matter recovery in semiarid grasslands: Implications for the conservation reserve program. *Ecological Applications* 5: 793–801.

Butaye J, Jacquemyn H, Honnay O, and Hermy M (2002) The species pool concept applied to forests in a fragmented landscape: Dispersal limitation versus habitat limitation. *Journal of Vegetation Science* 13: 27–34.

Caspersen JP, Pacala SW, Jenkins JC, Hurtt GC, Moorcroft PR, and Birdsey RA (2000) Contributions of land-use history to carbon accumulation in US forests. *Science* 290: 1148–1151.

Coffin DP, Lauenroth WK, and Burke IC (1996) Recovery of vegetation in a semiarid grassland 53 years after disturbance. *Ecological Applications* 6: 538–555.

Compton JE and Boone RD (2000) Long-term impacts of agriculture on soil carbon and nitrogen in New England forests. *Ecology* 81: 2314–2330.

Compton JE and Boone RD (2002) Soil nitrogen transformations and the role of light fraction organic matter in forest soils. *Soil Biology & Biochemistry* 34: 933–943.

Compton JE, Boone RD, Motzkin G, and Foster DR (1998) Soil carbon and nitrogen in a pine-oak sand plain in central Massachusetts: Role of vegetation and land-use history. *Oecologia* 116: 536–542.

Covington WW (1981) Changes in forest floor organic matter and nutrient content following clear cutting in northern hardwoods. *Ecology* 62: 41.

Currie WS and Nadelhoffer KJ (2002) The imprint of land-use history: Patterns of carbon and nitrogen in downed woody debris at the Harvard Forest. *Ecosystems* 5: 446–460.

De Keersmaeker L, Martens L, Verheyen K, Hermy M, De Schrijver A, and Lust N (2004) Impact of soil fertility and insolation on diversity of herbaceous woodland species colonizing afforestations in Muizen forest (Belgium). *Forest Ecology and Management* 188: 291–304.

DeGasperis BG and Motzkin G (2007) Windows of opportunity: Historical and ecological controls on Berberis thunbergii invasions. *Ecology* 88: 3115–3125.

Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature* 418: 700–707.

Donohue K, Foster DR, and Motzkin G (2000) Effects of the past and the present on species distribution: Land-use history and demography of wintergreen. *Journal of Ecology* 88: 303–316.

Duffy DC and Meier AJ (1992) Do Appalachian herbaceous understories ever recover from clearcutting? *Conservation Biology* 6: 196–201.

Dupouey JL, Dambrine E, Laffite JD, and Moares C (2002) Irreversible impact of past land use on forest soils and biodiversity. *Ecology* 83: 2978–2984.

Dzwonko Z and Loster S (1992) Species richness and seed dispersal to secondary woods in southern Poland. *Journal of Biogeography* 19: 195–204.

Eberhardt RW, Foster DR, Motzkin G, and Hall B (2003) Conservation of changing landscapes: Vegetation and land-use history of cape cod national seashore. *Ecological Applications* 13: 68–84.

Endels P, Adriaens D, Verheyen K, and Hermy M (2004) Population structure and adult plant performance of forest herbs in three contrasting habitats. *Ecography* 27: 225–241.
FAO (Food and Agriculture Organization) (2006) *The Global Forest Resources Assessment 2005*. Rome, Italy: FAO.
Fearnside PM and Barbosa RI (1998) Soil carbon changes from conversion of forest to pasture in Brazilian Amazonia. *Forest Ecology and Management* 108: 147–166.
Flinn KM (2007) Microsite-limited recruitment controls fern colonization of post-agricultural forests. *Ecology* 88: 3103–3114.
Flinn KM and Marks PL (2007) Agricultural legacies in forest environments: Tree communities, soil properties, and light availability. *Ecological Applications* 17: 452–463.
Flinn KM and Vellend M (2005) Recovery of forest plant communities in post-agricultural landscapes. *Frontiers in Ecology and the Environment* 3: 243–250.
Foster D, Swanson F, Aber J, *et al.* (2003) The importance of land-use legacies to ecology and conservation. *BioScience* 53: 77–88.
Foster DR (1992) Land-use history (1730-1990) and vegetation dynamics in Central New-England, USA. *Journal of Ecology* 80: 753–772.
Foster DR, Motzkin G, and Slater B (1998) Land-use history as long-term broad-scale disturbance: Regional forest dynamics in central New England. *Ecosystems* 1: 96–119.
Fraterrigo JM, Balser TC, and Turner MG (2006a) Microbial community variation and its relationship with nitrogen mineralization in historically altered forests. *Ecology* 87: 570–579.
Fraterrigo JM, Turner MG, and Pearson SM (2006b) Interactions between past land use, life-history traits and understory spatial heterogeneity. *Landscape Ecology* 21: 777–790.
Fraterrigo JM, Turner MG, and Pearson SM (2006c) Previous land use alters plant allocation and growth in forest herbs. *Journal of Ecology* 94: 548–557.
Fraterrigo JM, Turner MG, Pearson SM, and Dixon P (2005) Effects of past land use on spatial heterogeneity of soil nutrients in southern Appalachian forests. *Ecological Monographs* 75: 215–230.
Fuller TL, Foster DR, McLachlan TS, and Drake N (1998) Impact of human activity on regional forest composition and dynamics in central New England. *Ecosystems* 1: 76–95.
Gerhardt F and Foster DR (2002) Physiographical and historical effects on forest vegetation in central New England, USA. *Journal of Biogeography* 29: 1421–1437.
Goodale CL and Aber JD (2001) The long-term effects of land-use history on nitrogen cycling in northern hardwood forests. *Ecological Applications* 11: 253–267.
Goodale CL, Aber JD, and McDowell WH (2000) The long-term effects of disturbance on organic and inorganic nitrogen export in the White Mountains, New Hampshire. *Ecosystems* 3: 433–450.
Gough CM, Vogel CS, Harrold KH, George K, and Curtis PS (2007) The legacy of harvest and fire on ecosystem carbon storage in a north temperate forest. *Global Change Biology* 13: 1935–1949.
Grossmann EB and Mladenoff DJ (2008) Farms, fires, and forestry: Disturbance legacies in the soils of the Northwest Wisconsin (USA) Sand Plain. *Forest Ecology and Management* 256: 827–836.
Halpern CB and Spies TA (1995) Plant-species diversity in natural and managed forests of the Pacific-Northwest. *Ecological Applications* 5: 913–934.
Hansen MC, Stehman SV, and Potapov PV (2010) Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences* 107: 8650–8655.
Harding JS, Benfield EF, Bolstad PV, Helfman GS, and Jones EBD (1998) Stream biodiversity: The ghost of land use past. *Proceedings of the National Academy of Sciences of the United States of America* 95: 14843–14847.
Hermy M and Verheyen K (2007) Legacies of the past in the present-day forest biodiversity: A review of past land-use effects on forest plant species composition and diversity. *Ecological Research* 22: 361–371.
Honnay O, Hermy M, and Coppin P (1999) Impact of habitat quality on forest plant species colonization. *Forest Ecology & Management* 115: 157–170.
Hooker TD and Compton JE (2003) Forest ecosystem carbon and nitrogen accumulation during the first century after agricultural abandonment. *Ecological Applications* 13: 299–313.
Ihori T, Burke IC, Lauenroth WK, and Coffin DP (1995) Effects of cultivation and abandonment on soil organic matter in northeastern Colorado. *Soil Science Society of America Journal* 59: 1112–1119.
Jacquemyn H and Brys R (2008) Effects of stand age on the demography of a temperate forest herb in post-agricultural forests. *Ecology* 89: 3480–3489.
Johnson CM, Zarin DJ, and Johnson AH (2000) Post-disturbance aboveground biomass accumulation in global secondary forests. *Ecology* 81: 1395–1401.
Johnson DW and Curtis PS (2001) Effects of forest management on soil C and N storage: Meta analysis. *Forest Ecology and Management* 140: 227–238.
Kashian DM, Turner MG, and Romme WH (2005a) Variability in leaf area and stemwood increment along a 300-year lodgepole pine chronosequence. *Ecosystems* 8: 48–61.
Kashian DM, Turner MG, Romme WH, and Lorimer CG (2005b) Variability and convergence in stand structural development on a fire-dominated subalpine landscape. *Ecology* 86: 643–654.
Knops JMH and Tilman D (2000) Dynamics of soil nitrogen and carbon accumulation for 61 years after agricultural abandonment. *Ecology* 81: 88–98.
Kulakowski D and Veblen TT (2007) Effect of prior disturbances on the extent and severity of wildfire in Colorado subalpine forests. *Ecology* 88: 759–769.
Kulakowski D, Veblen TT, and Bebi P (2003) Effects of fire and spruce beetle outbreak legacies on the disturbance regime of a subalpine forest in Colorado. *Journal of Biogeography* 30: 1445–1456.
Latty EF, Canham CD, and Marks PL (2004) The effects of land-use history on soil properties and nutrient dynamics in northern hardwood forests of the Adirondack Mountains. *Ecosystems* 7: 193–207.
Lewis DB, Kaye JP, Gries C, Kinzig AP, and Redman Charles L (2006) Agrarian legacy in soil nutrient pools of urbanizing arid lands. *Global Change Biology* 12: 703–709.
Matlack GR (1994) Plant-species migration in a mixed-history forest landscape in eastern North America. *Ecology* 75: 1491–1502.
Matlack GR (2009) Long-term changes in soils of second-growth forest following abandonment from agriculture. *Journal of Biogeography* 36: 2066–2075.
McCarthy BC and Bailey RR (1994) Distribution and abundance of coarse woody debris in a managed forest landscape of the central Appalachians. *Canadian Journal of Forest Research* 24: 1317–1329.
McGrath DA, Smith CK, Gholz HL, and Oliveira FD (2001) Effects of land-use change on soil nutrient dynamics in Amazonia. *Ecosystems* 4: 625–645.
McLauchlan K (2006) The nature and longevity of agricultural impacts on soil carbon and nutrients: A review. *Ecosystems* 9: 1364–1382.
McLauchlan KK, Craine JM, Oswald WW, Leavitt PR, and Likens GE (2007) Changes in nitrogen cycling during the past century in a northern hardwood forest. *Proceedings of the National Academy of Sciences of the United States of America* 104: 7466–7470.
McLauchlan KK, Hobbie SE, and Post WM (2006) Conversion from agriculture to grassland builds soil organic matter on decadal timescales. *Ecological Applications* 16: 143–153.
Motzkin G, Foster D, Allen A, Harrod J, and Boone R (1996) Controlling site to evaluate history: Vegetation patterns of a New England sand plain. *Ecological Monographs* 66: 345–365.
Murty D, Kirschbaum MUF, McMurtrie RE, and McGilvray H (2002) Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. *Global Change Biology* 8: 105–123.
Paul KL, Polglase PJ, and Richards GP (2003) Predicted change in soil carbon following afforestation or reforestation, and analysis of controlling factors by linking a C accounting model (CAMFor) to models of forest growth (3PG), litter decomposition (GENDEC) and soil C turnover (RothC). *Forest Ecology and Management* 177: 485–501.
Peterken GF and Game M (1984) Historical factors affecting the number and distribution of vascular plant-species in the woodlands of central Lincolnshire. *Journal of Ecology* 72: 155–182.
Post WM and Kwon KC (2000) Soil carbon sequestration and land-use change: Processes and potential. *Global Change Biology* 6: 317–327.
Rackham O (1980) *Ancient Woodland: Its History, Vegetation and Uses in England.* London, UK: Edward Arnold.
Ramankutty N and Foley JA (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13: 997–1027.
Rhemtulla JM, Mladenoff DJ, and Clayton MK (2009) Legacies of historical land use on regional forest composition and structure in Wisconsin, USA (mid-1800s–1930s–2000s). *Ecological Applications* 19: 1061–1078.
Richter DD, Markewitz D, Heine PR, *et al.* (2000) Legacies of agriculture and forest regrowth in the nitrogen of old-field soils. *Forest Ecology and Management* 138: 233–248.
Roberts MR and Gilliam FS (2003) Response of the herbaceous layer to disturbance in eastern forests. In: Gilliam FS and Roberts MR (eds.) *The Herbaceous Layer in Forests of Eastern North America*, pp. 302–320. New York, NY, USA: Oxford University Press.

Schulte L, Mladenoff D, Crow T, Merrick L, and Cleland D (2007) Homogenization of northern U.S. Great Lakes forests due to land use. *Landscape Ecology* 22: 1089–1103.

Singleton R, Gardescu S, Marks PL, and Geber MA (2001) Forest herb colonization of postagricultural forests in central New York State, USA. *Journal of Ecology* 89: 325–338.

Springob G, Brinkmann S, Engel N, Kirchmann H, and Bottcher J (2001) Organic C levels of Ap horizons in North German Pleistocene sands as influenced by climate, texture, and history of land-use. *Journal of Plant Nutrition and Soil Science* 164: 681–690.

Sturtevant BR, Bissonette JA, Long JN, and Roberts DW (1997) Coarse woody debris as a function of age, stand structure, and disturbance in boreal Newfoundland. *Ecological Applications* 7: 702–712.

Thompson J, Brokaw N, Zimmerman JK, *et al.* (2002) Land use history, environment, and tree composition in a tropical forest. *Ecological Applications* 12: 1344–1363.

Tinker DB and Knight DH (2000) Coarse woody debris following fire and logging in Wyoming lodgepole pine forests. *Ecosystems* 3: 472–483.

Turner BL, Geoghegan J, and Foster DR (eds.) (2003) *Integrated Land Change Science and Tropical Deforestation in Southern Yucatán: Final Frontiers*. Oxford, UK: Oxford University Press.

Turner MG, Baker WL, Peterson CJ, and Peet RK (1998) Factors influencing succession: Lessons from large, infrequent natural disturbances. *Ecosystems* 1: 511–523.

Vellend M (2005) Land-use history and plant performance in populations of Trillium grandiflorum. *Biological Conservation* 124: 217–224.

Verheyen K, Bossuyt B, Hermy M, and Tack G (1999) The land use history (1278-1990) of a mixed hardwood forest in western Belgium and its relationship with chemical soil characteristics. *Journal of Biogeography* 26: 1115–1128.

Verheyen K, Guntenspergen GR, Biesbrouck B, and Hermy M (2003a) An integrated analysis of the effects of past land use on forest herb colonization at the landscape scale. *Journal of Ecology* 91: 731–742.

Verheyen K and Hermy M (2004) Recruitment and growth of herb-layer species with different colonizing capacities in ancient and recent forests. *Journal of Vegetation Science* 15: 125–134.

Verheyen K, Honnay O, Motzkin G, Hermy M, and Foster DR (2003b) Response of forest plant species to land-use change: A life-history trait-based approach. *Journal of Ecology* 91: 563–577.

White AS and Pickett STA (1985) Natural disturbance and patch dynamics: An introduction. In: Pickett STA and White PS (eds.) *The Ecology of Natural Disturbance and Patch Dynamics*, pp. 3–13. New York: Academic Press.

Williams M (1990) Forests. In: Turner, II B, Clark WC, Kates RW, Richards JF, Mathews JT, and Meyer WB (eds.) *The Earth as Transformed by Human Action: Global and Regional Changes in the Biosphere Over the Past 300 Years*, pp. 179–201. Cambridge, UK: Cambridge University Press.

Williams M (2000) Dark ages and dark areas: Global deforestation in the deep past. *Journal of Historical Geography* 26: 28–46.

Yanai RD, Currie WS, and Goodale CL (2003) Soil carbon dynamics after forest harvest: An ecosystem paradigm reconsidered. *Ecosystems* 6: 197–212.

Landscape Modeling

Robert M Scheller, Portland State University, Portland, OR, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Disturbance A relatively discrete event that results in mortality of plants or animals or an event that is uncharacteristically intense or long-lasting. Wildfire, tornados, and logging are disturbances that are generally stochastic in location and time. Floods or extended dry periods (droughts) are considered disturbances if they are unusually large or intense.

Endogenous process Process that emerges from biotic and abiotic interactions internal to the system of interest. Examples include vegetation growth and succession, dispersal of organisms within a landscape, and topographic influences on microclimate. Some human activities may be considered endogenous – for example, various crop rotation systems – if they are well established and responsive to local needs.

Exogenous process Process external to the system of interest. For landscapes, exogenous processes would include climate change or species migration into the landscape. Human activities may be considered exogenous if they are novel, permanent (e.g., housing development), or driven by external forces (e.g., the world market for timber).

Geographic information system (GIS) Computational system for storing, managing, manipulating, and displaying spatial data. Because landscape models are inherently spatial, strong links to GIS are either built into such models or are required to interface with the model input and output data. Many GIS, both free and commercial, are available for use with landscape models.

Resolution The resolution of information represented within a landscape model has three primary components: spatial, temporal, and taxonomic. *Spatial resolution* refers to the smallest unit size or area represented (synonymous with *grain*). For vector data, resolution is the minimum mapping unit. For raster data, it is the cell size. *Temporal resolution* refers to the frequency at which various social or ecological processes operate. Common examples include hourly, daily, monthly, and annually. *Taxonomic resolution* is the degree of aggregation of species information. The smallest unit possible is species-specific information. *Taxonomic* aggregations typically used in landscape modeling include species, functional groups, potential vegetation types, and successional states.

Scenario A discrete suite of assumptions regarding future drivers of landscape change that are fed into a landscape model or models. The list of assumptions often includes future climate, future disturbance regimes, and future management or land-use policies. Multiple scenarios can be created to form a full factorial matrix that highlights experimentation and statistical rigor. Alternatively, fewer scenarios can be assembled into central "story lines" that highlight key management or policy options.

Stochastic process Process in which one or more simulated events or rates has values that are derived in part from a probability distribution function (e.g., Gaussian, Weibull, exponential). Therefore, such events or rates are not typically "random" (although random is a valid probability distribution function) but are unpredictable. Examples of simulated stochastic processes include the location of lightning strikes, daily maximum temperature, and wind speeds.

Introduction

A sufficient understanding of the threats to biodiversity and the opportunities to preserve or enhance biodiversity requires an acknowledgment of the broader context within which the various components of biodiversity operate. Over the past several decades, managers have realized that consideration of only small watersheds or forest stands is not sufficient to maintain biodiversity (Hansson and Angelstam, 1991; White *et al.*, 1997; Foley *et al.*, 2005; Fischer and Lindenmayer, 2007; Lindenmayer *et al.*, 2008). Insufficient attention to the broader context within which an area of interest (e.g., a national park or wilderness area) resides can lead to population decline, local extirpation, or catastrophic disruption due to large, infrequent disturbance events (den Boer, 1981; Tilman *et al.*, 1994; Huxel and Hastings, 1999; Brashares, 2010; Spencer *et al.*, 2011). This broader context can be considered the *functional landscape*: an area large enough to capture all of those processes that maintain or reduce biodiversity and large enough to provide sufficient habitat for populations of native species. A functional landscape is typically smaller than a biome but larger than a first-order watershed and several times larger than the maximum disturbance size (Baker, 1989a; Dunning *et al.*, 1992; Turner *et al.*, 1993, 1995; Forman, 1995).

Functional landscapes (*landscapes* henceforth) are an intuitive unit for studying and managing biodiversity. By encompassing most of the processes influencing biodiversity, landscapes afford the opportunity to practice inclusionary management that is more holistic and considerate of all relevant processes and elements (the individual components of a landscape – for example, forest patches, stream corridors, and grass or shrub fields). However, landscapes pose distinct challenges for ecologists and managers. Because of their size, they are inherently nonreplicable (Hurlbert, 1984). Also because of size, significant change at the landscape scale typically occurs over many decades: changes to composition, structure, or function emerge more slowly when averaged across large

numbers of each landscape element. Such long durations are generally not commensurate with the available funding for either research or management. For management, these concerns are particularly acute because trends occurring at a landscape scale are often difficult and expensive to reverse through restoration efforts. Borrowing a business aphorism, changing course at the landscape scale is akin to turning an oil tanker around: it happens slowly at best. And no one wants to be held accountable for running an ecosystem into the shoals of lost biodiversity or reduced ecosystem functioning.

Because of these spatial and temporal challenges, both scientists and managers increasingly rely on landscape models to understand and predict change at these spatial and temporal scales. A landscape model is a mathematical and logical representation of a functional landscape that includes the subset of exogenous and endogenous processes expected to explain a majority of the coarse-scale spatial and temporal variability. These processes nearly always include disturbances and vegetative succession. A typical landscape model will include a representation of vegetative change (the author will discuss techniques in the following sections) and one or more stochastic disturbances: wildfire, development and urban expansion, forest management, windthrow, and so on. As with all models, an empirical foundation is necessary to derive the mathematical and logical abstractions of the real system.

Landscape models differ from other kinds of ecological model because they are spatially explicit, depicting the composition, arrangement, and geographic location of the various landscape elements. In practice, this means that the models are often closely linked to a geographic information system (GIS) for managing the spatial information. Results are often provided via maps of biodiversity or other conditions of interest, allowing an intuitive assessment of spatial variation. Maps are particularly valuable for communicating with stakeholders. Unfortunately, accompanying maps of uncertainty are seldom provided.

Landscape models also depict the spatial interactions among constituent elements. Spatial interactions are the lateral transfer of information, matter, or energy generated by ecological processes (Reiners and Driese, 2001). Movement of matter and energy includes the lateral transport of water and dissolved nutrients. The transfer of matter and energy may also be generated through disturbances. Wildfire is an archetypal spatial process because it releases and transfers energy and the size and duration of wildfire is highly dependent on the spatial configuration of fuels and topography (Hargrove *et al.*, 2000; Finney *et al.*, 2007; Schmidt *et al.*, 2008).

The lateral transfer of information is synonymous with species distributions and community composition and is transferred across the landscape through dispersal and migration. Similar to wildfire, dispersal is often highly dependent on landscape configuration and fragmentation (Miller *et al.*, 1997; King and With, 2002; Higgins *et al.*, 2003b; Iverson *et al.*, 2004; Scheller and Mladenoff, 2008). Without spatial interactions – or if spatial interactions are weak – landscape models can be reduced to a collection of individual units (e.g., stands), and these units could simply be multiplied by the area within the landscape that they represent. Further explication of when a spatially dynamic model (one that includes spatial interactions) is necessary is provided below (*see* Landscape Modeling Approaches).

By representing the dominant processes at a scale relevant to understanding and managing biodiversity and by representing the spatial configuration and interactions of a landscape, landscape models allow empirically derived knowledge to be extrapolated to larger spatial and temporal domains. Doing so, we can begin to engage in meaningful *ecological forecasting* (Clark *et al.*, 2001), providing managers and policy makers tools for enhanced decision making that incorporates knowledge about the future, however limited (Pielke, 2003).

Landscape models are frequently paired with scenarios that can be relatively simple narratives about future policies or conditions – for example, the Millennium Ecosystem Assessment (2005). Scenarios can also be framed within an experimental design with multiple factors along two or more gradients (e.g., three levels of fire rotation period × three levels of harvesting in a full factorial design). Each individual scenario is then used to project how a landscape will change given the associated suite of assumptions about the frequency and intensity of the constituent processes (particularly of anthropogenic origin) (Carpenter, 2000). Typically, the results from each scenario are then compared with each other and with empirical data when available (*see* Conclusions). By comparing results among scenarios, the need for absolute accuracy is reduced as the *relative* merits of different policies or assumptions can still be assessed.

Although serving the utilitarian purpose of extrapolating knowledge to larger spatial and temporal domains, landscape models also enable a synthesis that is otherwise impossible. As knowledge about biodiversity expands and the threats to biodiversity multiply, the requisite number of processes that must be integrated and managed to preserve biodiversity grows exponentially. However, the human capacity to synthesize the effects of many processes, each operating at unique and often disparate scales, is limited. Landscape models allow many different processes to operate within and interact across a landscape. Patterns will emerge (*emergent behaviors*) that could not be anticipated using studies of isolated processes or intuition alone. For example, simulations of climate change in Siberia that included wildfire, logging, and insect mortality revealed that the expansion of harvesting in combination with increased insect activity will potentially have a far greater effect on forest composition than the direct effects of climate change alone (Gustafson *et al.*, 2010).

Because landscape models often incorporate multiple stochastic processes, they are also well suited for partitioning the total uncertainty of a landscape projection into its components: parameter, model, and inherent uncertainty (Haag and Kaupenjohann, 2001; Higgins *et al.*, 2003a; Millar *et al.*, 2007). Parameter and model uncertainty are artifacts of the modeling process and represent the current limits of knowledge about the system. Inherent uncertainty is the accumulated effect of all nondeterministic variation. For example, seed dispersal has significant inherent uncertainty as we can never know exactly how far an individual seed may travel. The exact time and location of lightning ignition of fires is similarly unknowable. Inherent uncertainty may seem to deflate ambitions for ever-increasing model accuracy. However, even if the accuracy of a given model projection is weak or unknown (as is the case for all projections of the deeper future), knowledge about the sources of uncertainty

(e.g., Allen *et al.*, 2000; Xu *et al.*, 2009) can be extremely valuable to managers. If a system has relatively low inherent uncertainty, then human agency will be relatively more effective. If the opposite is true, then caution is warranted before investing considerable resources into attempts at guiding or dictating landscape change.

Finally, landscape models also provide a critical heuristic function. The process of formulating, parameterizing, calibrating, and validating a landscape model requires the synthesis of multiple sources of data and models. A landscape model typically requires vast quantities of data that are processed in advance of conducting any simulations. For example, even the simplest landscape model requires remotely sensed data that are classified to a spatial and taxonomic resolution that are congruent with the questions at hand. More detailed or complex landscape models may require additional edaphic or topographic data or more detailed taxonomic data. In addition, formulating how the various processes will interact across space and time requires a deep understanding of the ecology of the system (and therefore modeling efforts often involve teams of scientists). Existing representations of processes may need to be altered or discarded if they were developed for a scale substantially different from the other model components. Validation requires more data still and is difficult to obtain for long-term or projected models. In total, the modeling process is dominated by data synthesis and analysis with a relatively small amount of time devoted to actual simulations.

Landscape Modeling Approaches

When is a landscape model useful or necessary? Landscape models are useful when the representation of exogenous and endogenous processes, represented within a geographic framework, and with spatial interaction among landscape elements is required to project, understand, and manage landscapes for biodiversity. These are also the criteria the author used to select modeling approaches for review. There is an incredible diversity of landscape models – more than can be summarized here. These models span broad gradients of complexity, specificity, usability, and application. Many excellent reviews have been written summarizing the past and current state of landscape models (Baker, 1989b; Mladenoff and Baker, 1999; Keane *et al.*, 2004; Scheller and Mladenoff, 2007; Seidl *et al.*, 2011). More information about the mechanics of developing landscape models is also widely available (Maxwell and Costanza, 1997; Scheller *et al.*, 2010). Rather than an exhaustive and complete list, the author focussed on four of the most common approaches to formulating (or applying) landscape models: state-and-transition models, process models, hybrid models and multimodeling, and neutral models (an approach to applying models, rather than a type). The author will highlight their respective strengths and limitations and the appropriate domain of application for each.

Ideally, a model is no more complex than necessary to answer the given question. Given the enormous additional complexity introduced by the inclusion of spatial interactions (Shugart Jr., 1998; Strayer *et al.*, 2003), it is critical to decide when the representation of spatial processes – and therefore the need for a landscape model at all – is justified. All ecosystems contain spatial interactions. The question is whether the policy, management, or ecological question can be answered with a simpler nonspatial model or whether the additional complexity is required. There are no absolute answers; however, general criteria can be outlined based on the strength of spatial interactions at the chosen model resolution and extent (Strayer *et al.*, 2003). If a spatial process is occurring at a resolution much finer than model resolution, then the process is often considered "noise" that can be safely averaged without loss of important information. For example, when modeling the effects of climate change on terrestrial systems at regional or continental systems (e.g., Bachelet *et al.*, 2003), the inclusion of the effects of competition among neighboring trees or fine-scale propagules dispersal would be unlikely to reduce uncertainty at the intended scale. Inclusion of the effects of tree species migration via dispersal may, however, be required if substantial range shifts are expected over the simulated duration. Similarly, if a spatial process naturally occurs across extents far greater than the focal extent, the process can be represented as a simple rate (potentially changing through time). For regional landscapes (or smaller), climate is often treated in this fashion even though we recognize that weather is a very spatially dynamic process.

All models are compromises. The data available to parameterize and validate models are limited, as are the resources (fiscal, temporal, and computational) available for developing and applying a model. Is it important that the model be able to run on a typical desktop computer? Is an intuitive user interface necessary for the intended user to be able to quickly implement the model? Each model developer must rank the priorities for resolving the question at hand.

Models must also compromise among processes for the allocation of complexity. For example, a landscape model with a fire component that represents hourly growth of a fire and flame lengths at a 1-m resolution necessarily devotes fewer resources to landscape processes that unfold over many decades (e.g., tree species migration). If the model developer devoted most resources to developing a model of individual trees, then the biogeochemical components may be simplified or lacking altogether. Each of the four types (**Table 1**) described in the following sections represent the compromises and such sacrifices. None of these model types is appropriate for all situations and questions.

State-and-Transition Models

State-and-transition models are typically used to address management questions for which distinct vegetative conditions or classes are readily identifiable and their relative abundance is an important consideration. For example, Pastor *et al.* (2005) modeled forest change in northern Minnesota as a function of harvesting using an age-dependent state-and-transition model. They concluded that a narrow range of harvest ages will reduce the amount of the landscape in a regeneration state, a result that is counter to emerging trends toward managing with a broader range of harvest ages.

State-and-transition models place each landscape element into one of a limited number of states (Cattelino *et al.*, 1979;

Table 1 Four general landscape modeling approaches, example questions, and strengths and weaknesses

Model approach	*Example question*	*Strengths*	*Weaknesses*
State-and-transition	How will the relative abundance of discrete vegetation classes change through time?	Relatively straightforward logic; easier to parameterize; simplification facilitates collaboration; reduced computational requirements.	Less flexible responses to novel climates, introduced species, altered disturbance regimes; nonstationary landscape dynamics poorly represented.
Process	How will net biome production change if wildfire increases with climate change?	Highly flexible; ability to represent nonstationary landscape dynamics; can represent changing ecosystem process rates.	Steep parameterization requirements; requires specialized knowledge and therefore is less suitable for stakeholder engagement; high computational requirements.
Neutral	Are patch size or shape important determinants of population migration?	Assess strength of individual processes; location neutral and not context dependent; testing of multiple landscape configurations.	The results are more abstract; artificiality may render results less applicable to actual threats to biodiversity; lack of place-based output maps reduces stakeholder engagement.
Hybrid or multimodeling	Do landscape-scale fire dynamics influence carnivore population viability?	Integration of multiple domains (e.g., forests and animal metapopulations); integration across broader range of spatial and temporal scales; increased ability to assess sources of uncertainty.	High complexity can overwhelm modeler's ability to partition sources of variation; need to validate across multiple domains; communicating system behavior to nontechnical audiences is challenging.

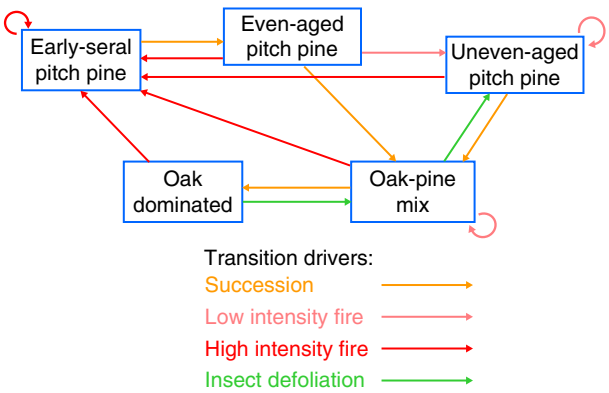


Figure 1 Conceptual diagram of a mid-Atlantic pine barrens state-and-transition model.

Klenner *et al.*, 2000; Acevedo *et al.*, 2001; Monticino *et al.*, 2002; Pastor *et al.*, 2005). These states typically represent the vegetative condition at a given time. For example, in the eastern US, a landscape element could be initially classified as even-aged pitch pine if satellite and field data indicate that the unit is dominated by *Pinus rigida* with an average age between 30 and 100 years (**Figure 1**). The transition to a different predetermined alternative state may occur if either a set amount of time has passed (succession) or a transition is triggered by a disturbance event. The even-aged pitch pine may transition to oak–pine mix 70 years after initiation unless a fire is simulated, changing it to early seral pitch pine or uneven-aged pitch pine, depending on fire severity (**Figure 1**).

Transitions among states may be dictated by disturbance as previously noted or may be determined via a transition probability matrix (Logofet and Lesnaya, 2000; Yemshanov and Perera, 2002). In this case, a state may transition to one of many possible alternate states based on a probability assigned to each transition. The combination of a fixed state and a transition determined in part by a stochastic process is considered a *semi-Markov* model.

The number of states and their taxonomic resolution (e.g., three classes of pitch pine vs. forest, prairie, and savannah), and the number of transitions is dependent on the objectives of the research or the management application. Shorter duration (<50 years) simulations often stratify a landscape into potential vegetation types with a suite of states for each type, whereas longer duration simulations may focus more on transitions among major land types (e.g., agriculture, forest, pasture, developed).

Because of the finite number of states (and therefore transitions), state-and-transition landscape models are conducive to rapid deployment and collaborative model construction. One of the most commonly used state-and-transition models, the Vegetation Dynamics Development Tool (VDDT), has an intuitive graphic interface that allows groups of scientists and managers to quickly visualize a model and focus their efforts on deliberatively constructing the states and transitions that best represent the dynamics of their system (Hemstrom *et al.*, 2001; Shifley *et al.*, 2008). The relative simplicity also maximizes transparency and generates greater "buy-in" from local managers and stakeholders. This process of constructing state-and-transition models to engage a broader group of participants has been widely used by the Nature Conservancy in North America (Shlisky *et al.*, 2005).

However, the finite states also limit the inferential power of these landscape models. Most landscapes around the world today are in part dominated by one or more *nonstationary* processes, meaning that their frequency or magnitude is changing through time. Consequently, many states may cease to be significant components of their respective landscapes or the transitions among states may be substantially altered through time. Novel communities (states) may emerge due to climate change (Williams *et al.*, 2007) or invasive plant species. Invasive (or climatically released) insects may become the dominant process driving vegetative change (Ward and Masters, 2007; Kurz *et al.*, 2008).

Process Models

All landscape models represent one or more processes, and therefore this landscape model type is a misnomer. What distinguishes these models is that all processes, including succession, are explicitly represented and there are no fixed states or transitions. Rather, plant community composition is an emergent property itself. Depending on how succession is represented and on the taxonomic resolution, plant community composition may be a function of disturbance intensity and type, plant competition for light or nutrients, seed dispersal, and other species vital attributes (Mladenoff *et al.*, 1996; Roberts, 1996). Because composition is more nearly continuous, all other processes can be formulated to be more sensitive to the effects of plant community composition and diversity. In particular, ecosystem processes (e.g., aboveground net primary productivity, soil respiration) can more readily be incorporated as these generally require a tight coupling between vegetative composition, structure, climate and soils (Scheller *et al.*, 2011a).

Such landscape process models therefore have much greater flexibility to deal with nonstationary processes, notably climate change. Novel communities can emerge from the interplay between stochastic disturbance processes (themselves responding to climate), species vital attributes, and ecosystem processes that will dictate other physiological constraints (e.g., nitrogen availability, soil moisture, Scheller *et al.*, 2011a) (**Figure 2**). Such an approach is essential for projecting landscape capacity to sequester carbon and provide habitat further into the future (>50 years).

The added flexibility does, however, come at a substantial cost. The data required to parameterize such models is orders of magnitude greater than for state-and-transition models. And the process of constructing and parameterizing these landscape models is at best a specialized skill and at worst hopelessly opaque to anyone outside of the discipline. Policy makers, managers, and stakeholders can feel removed from the modeling process and are generally unable to assess the quality of the research. The important drivers that may emerge – N availability, litter lignin content – are not readily measureable by land managers and often only loosely related back to the task of managing for biodiversity.

Hybrid and Multimodeling

An emerging trend is the combination of one or more modeling approaches to understand and project changing biodiversity. In particular, hybrid models and multimodeling offer the opportunity to include processes or interactions across a broader range of spatial, temporal, or taxonomic resolutions than might be possible with a single modeling approach.

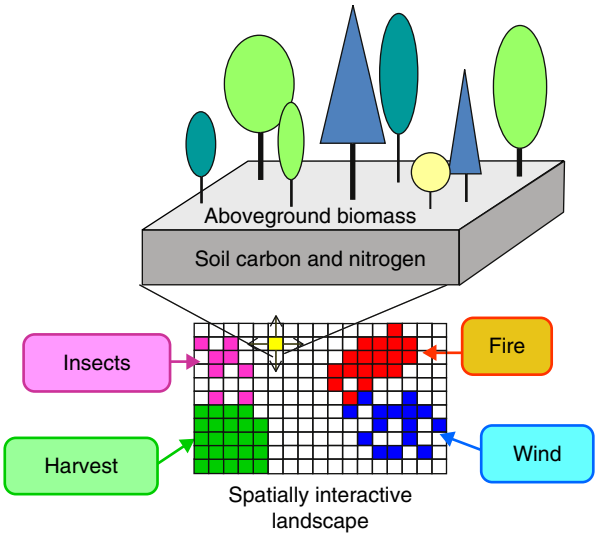


Figure 2 Conceptual diagram of a process landscape model, emphasizing ecosystem processes and spatially dependent disturbances.

To better understand vegetation change, hybrid models have been deployed to represent finer-scale processes using a different modeling paradigm than is used to represent broader landscape process interactions. For example, competition among individual trees and subsequent growth may be simulated using an individual tree model (Robinson and Ek, 2003; Seidl *et al.*, in press) or a gap model (Pastor and Post, 1988; Seidl *et al.*, 2005). These fine-scale models can be stitched together spatially to form a continuous landscape (e.g., Urban *et al.*, 1991; Bragg *et al.*, 2004). Alternatively, individual tree data can be averaged using ecological field theory and scaled up to allow for the efficient calculation of processes that operate at larger spatial or temporal scales (e.g., Seidl *et al.* in press). The advantage of hybrid forest models is the inclusion of individual trees, their growth, and competition with neighbors. The challenge is the additional parameterization and data required. To date, these data are seldom captured by the remote sensing tools available, and hybrid models have been limited to relative small landscapes ($<10{,}000$ ha).

Multimodeling (or *model coupling*) has also been extended to assessing animal populations. The challenges of overlaying stochastic animal populations on top of the existing stochastic disturbance dynamics are considerable. However, advances in software and computer power have accelerated the ability to consider both changing habitats and changing populations simultaneously (Larson *et al.*, 2004; Akcakaya *et al.*, 2005; Shifley *et al.*, 2008). For example, to project changes in the population of fisher (*Martes pennanti*) in the southern Sierra Nevada, California, USA, a landscape change model was coupled with a metapopulation model (Spencer *et al.*, 2011; Syphard *et al.*, 2011). As a result, two of the largest sources of uncertainty were incorporated: wildfire and fisher dispersal.

By doing so, a more accurate estimate of inherent uncertainty was achieved (Scheller *et al.*, 2011b).

Neutral Models

Neutral models are not a model type in and of themselves but rather a unique application of landscape models (Gardner *et al.*, 1987; With and King, 1997). Neutral modeling is the process of applying a landscape model to a strictly hypothetical landscape. In this case, the complexity inherent in parameterizing an actual landscape is jettisoned in favor of a focus on the processes themselves. A neutral modeling approach allows for much more precise testing of model sensitivity to landscape configuration and community composition than is possible on a real landscape. The goal of neutral landscape modeling is a more fundamental understanding of the processes that control landscape change across *all* landscapes rather than the narrower domain of a real landscape. Real landscapes simply never contain the full range of possible landscape configurations and species that may inform longer-range management efforts. As an example, Huxel and Hastings (1999) used neutral models to examine how the placement of restoration efforts (adjacent to existing habitat vs. random placement) could significantly influence habitat restoration success. Attempting to find a suitable suite of similar landscapes (with replication) for such experimentation would have been impossible.

Although further removed from the immediacies of landscape management, neutral models can potentially inform managers about potential solutions or enhancements that are not immediately apparent from an examination of current conditions (Huxel and Hastings, 1999; With *et al.*, 1999; King and With, 2002). The risk is further detachment from management relevancy: maps of actual landscapes and geographically grounded projections of landscape change remain extremely powerful tools for communicating the science of landscape change.

Conclusions

The limitations to landscape models are similar to that of all models, with one important caveat. All models are limited (or strengthened) by the empirical foundation on which they are built: by the data underlying their formulation, the data available to parameterize a model in a novel setting, the computational limits to complex simulations, and the software limitations to managing complex algorithms and architectures. With limited experience, I assume these limitations also apply to models of beating hearts and hurricane trajectories.

However, landscape models, like climate change models, are additionally limited by the ability to validate results. When projecting more than a few years into the future, validation of model results – the cross-check against reality that has made weather forecasting so successful – requires the continual collection of spatially explicit data through time (Rastetter, 1996; Rykiel, 1996; Haag and Kaupenjohann, 2001; Gardner and Urban, 2003). If projecting more than 10 or 20 years, validation would require patience and funding not typically afforded the average scientist or manager. In addition, it is important to remember that because of stochastic events, the actual observed landscape is but one of many possible outcomes that could have occurred. Rarely will a single model simulation match reality. Thus, even assuming that spatial data are collected, validation should be against confidence intervals generated by running the model many times (e.g., model replication).

Complete validation may be the gold standard of modeling (as double-blind experiments are to medicine), but it is only one route to building confidence in the knowledge gained from models. Other avenues for building confidence in landscape model results include the validation of individual processes, model documentation, model transparency (e.g., closed vs. open source), and repeated application across diverse systems. Deploying a rigorous software engineering process during the development stage can maximize model verification (sensu; Aber, 1997) and has become more important as model complexity has increased (Scheller *et al.*, 2010).

It is incumbent on those who build and apply landscape models to follow a few general principles. First, quantify model sensitivity. There is little value obtaining precise leaf nitrogen estimates if wildfire explains 90% of the variation in species composition. Conversely, if leaf nitrogen is the controlling variable, more effort can be placed on obtaining empirical estimates.

Second, the inability to validate the complete landscape does not preclude the validation of individual processes. Often available knowledge about individual processes (tree growth rates, wildfire spread, logging patterns) is large and should be rigorously utilized.

Third, replication is essential. The existence of inherent uncertainty in reality and the incorporation of stochastic processes into models ensures that no single model simulation will be broadly representative of the future.

Fourth, communicate uncertainty. Although maps of projected landscape change are powerful tools for communication, it is essential that the underlying uncertainty be expressed through time series with replicate standard errors, probability surface maps, or similar approaches.

Similarly, it is incumbent on the managers and policy makers to be aware of these principles and to recognize those model applications that do or do not adhere to them. Unfortunately, many managers and policy makers instinctively distrust model results and do not engage in the process. This is perhaps in part due to inadequate education and outreach on the part of those who conduct modeling research (Pielke, 2003). It is also true that model results have often been presented as *predictions*, which imply a high level of accuracy, and the degree of uncertainty has been underrepresented. Modelers need to articulate what managers can learn from models while acknowledging that we live in a ''black swan'' world (Taleb, 2007).

As the collective challenge to preserve biodiversity increases, tools that enable learning about the future are increasingly valuable. Landscape models – in conjunction with well-formulated scenarios – offer a route to understanding and managing biodiversity across large areas and long

durations. The relentless increase in spatial data (particularly from remote sensing), computing power, and the sophistication of landscape models has enabled scientists and managers to look further into the future and with greater accuracy than ever before. Continued success and improvement is possible, dependent on the ability of scientists to provide a transparent, collaborative, and rigorous system that maximizes learning while acknowledging uncertainty.

Appendix

University Courses

1. Landscape Ecology
2. Forest Ecology
3. Simulation Modeling

See also: Biodiversity in Logged and Managed Forests. Climate, Effects of. Computer Systems and Models, Use of. Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems. Dynamic Global Vegetation Models. Forest Ecology. Land-Use Issues. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Modeling Terrestrial Ecosystem Services. Succession, Phenomenon of. Timber Industry

References

Aber JD (1997) Why don't we believe the models. *Bulletin of the Ecological Society of America* 232–233.

Acevedo MF, Ablan M, Urban DL, and Pamarti S (2001) Estimating parameters of forest patch transition models from gap models. *Environmental Modelling & Software* 16: 649–658.

Akcakaya HR, Franklin J, Syphard AD, and Stephenson JR (2005) Viability of Bell's Sage Sparrow (*Amphispiza Belli* ssp. Belli): Altered fire regimes. *Ecological Applications* 15: 521–531.

Allen MR, Stott PA, Mitchell JFB, Schnur R, and Delworth TL (2000) Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature* 407: 617–620.

Bachelet D, Neilson RP, Hickler T, *et al.* (2003) Simulating past and future dynamics of natural ecosystems in the United States. *Global Biogeochemical Cycles* 17: 1–21.

Baker WL (1989a) Landscape ecology and nature reserve design in the Boundary Waters Canoe Area, Minnesota. *Ecology* 70: 23–35.

Baker WL (1989b) A review of models of landscape change. *Landscape Ecology* 2: 111–333.

Bragg DC, Roberts DW, and Crow TR (2004) A hierarchical approach for simulating northern forest dynamics. *Ecological Modelling* 173: 31–94.

Brashares JS (2010) Filtering wildlife. *Science* 329: 402–403.

Carpenter SR (2000) Ecological futures: Building an ecology of the long now. *Ecology* 83: 2069–2083.

Cattelino PJ, Noble IR, Slatyer RO, and Kessell SR (1979) Predicting the multiple pathways of plant succession. *Environmental Management* 3: 41–50.

Clark JS, Carpenter SR, Barber M, *et al.* (2001) Ecological forecasts: an emerging imperative. *Science* 293: 657–660.

den Boer PJ (1981) On the survival of populations in a heterogeneous and variable environment. *Oecologia* 50: 39–53.

Dunning JB, Danielson BJ, and Pulliam HR (1992) Ecological processes that affect populations in complex landscapes. *Oikos* 65: 169–175.

Finney MA, Selia RC, McHugh CW, Ager AA, Bahro B, and Agee JK (2007) Simulation of long-term landscape-level fuel treatment effects on large wildfires. *International Journal of Wildland Fire* 16: 712–727.

Fischer J and Lindenmayer DB (2007) Landscape modification and habitat fragmentation: A synthesis. *Global Ecology and Biogeography* 16: 265–280.

Foley JA, Defries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.

Forman RTT (1995) Some general principles of landscape and regional ecology. *Landscape Ecology* 10: 133–142.

Gardner RH, Milne BT, Turner MG, and O'Neill RV (1987) Neutral models for the analysis of broad-scale landscape pattern. *Landscape Ecology* 1: 19–28.

Gardner RH and Urban DL (2003) Model validation and testing: Past lessons, present concerns, future prospects. In: Canham CD, Cole JJ, and Lauenroth WK (eds.) *Models in Ecosystem Science*, pp. 184–203. Princeton, New Jersey, USA: Princeton University Press.

Gustafson EJ, Shvidenko AZ, Sturtevant BR, and Scheller RM (2010) Predicting global change effects on forest biomass and composition in south-central Siberia. *Ecological Applications* 20: 700–715.

Haag D and Kaupenjohann M (2001) Parameters, prediction, post-normal science and the precautionary principle – a roadmap for modelling for decision-making. *Ecological Modelling* 144: 45–60.

Hansson L and Angelstam P (1991) Landscape ecology as a theoretical basis for nature conservation. *Landscape Ecology* 5: 191–201.

Hargrove WW, Gardner RH, Turner MG, Romme WH, and Despain DG (2000) Simulating fire patterns in heterogeneous landscapes. *Ecological Modelling* 135: 243–263.

Hemstrom MA, Korol JJ, and Hann WJ (2001) Trends in terrestrial plant communities and landscape health indicate the effects of alternative management strategies in the interior Columbia river basin. *Forest Ecology and Management* 5504: 1–21.

Higgins SI, Clark JS, Nathan R, *et al.* (2003a) Forecasting plant migration rates: Managing uncertainty for risk assessment. *Journal of Ecology* 91: 341–347.

Higgins SI, Lavorel S, and Revilla E (2003b) Estimating plant migration rates under habitat loss and fragmentation. *Oikos* 101: 354–366.

Hurlbert SH (1984) Pseudorepublication and the design of ecological field experiments. *Ecological Monographs* 54: 187–211.

Huxel GR and Hastings A (1999) Habitat loss, fragmentation, and restoration. *Restoration Ecology* 7: 309–315.

Iverson LR, Schwartz MW, and Prasad AM (2004) How fast and far might tree species migrate in the eastern United States due to climate change? *Global Ecology and Biogeography* 13: 209–219.

Keane RE, Cary GJ, Davies ID, *et al.* (2004) A classification of landscape fire succession models: Spatial simulations of fire and vegetation dynamics. *Ecological Modelling* 179: 3–27.

King AW and With KA (2002) Dispersal success on spatially structured landscapes: When do spatial pattern and dispersal behavior really matter? *Ecological Modelling* 147: 23–39.

Klenner W, Kurz W, and Beukema S (2000) Habitat patterns in forested landscapes: Management practices and the uncertainty associated with natural disturbances. *Computers and Electronics in Agriculture* 27: 243–262.

Kurz WA, Dymond CC, Stinson G, *et al.* (2008) Mountain pine beetle and forest carbon feedback to climate change. *Nature* 452: 987–990.

Larson MA, Thompson III FR, Millspaugh JJ, Dijak WD, and Shifley SR (2004) Linking population viability, habitat suitability, and landscape simulation models for conservation planning. *Ecological Modelling* 180: 103–118.

Lindenmayer D, Hobbs RJ, Montague-Drake R, *et al.* (2008) A checklist for ecological management of landscapes for conservation. *Ecology Letters* 11: 78–91.

Logofet DO and Lesnaya EV (2000) The mathematics of Markov models: What Markov chains can really predict in forest successions. *Ecological Modelling* 126: 285–298.

Maxwell T and Costanza R (1997) A language for modular spatio-temporal simulation. *Ecological Modelling* 103: 105–113.

Millar CI, Stephenson NL, and Stephens SL (2007) Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* 17: 2145–2151.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well Being: Scenarios,* 596 pp. Washington, DC: Island Press.

Miller GS, Small RJ, and Meslow EC (1997) Habitat selection by Spotted Owls during natal dispersal in western Oregon. *Journal of Wildlife Management* 61: 140–150.

Mladenoff DJ and Baker WL (1999) Development of forest and landscape modeling approaches. In: Mladenoff DJ and Baker WL (eds.) *Spatial Modeling of Forest Landscape Change*, pp. 1–13. Cambridge, UK: Cambridge University Press.

Mladenoff DJ, Host GE, Boeder J, and Crow TR (1996) LANDIS: A spatial model of forest landscape disturbance, succession, and management. In: Goodchild MF,

Steyaert AH, Parks BO, *et al.* (eds.) *GIS and Environmental Modeling: Progress and Research Issues*, pp. 175–179. Colorado, USA: GIS World Books, Fort Collins.

Monticino MG, Cogdill T, Acevedo MF (2002) Cell interaction in semi-Markov forest landscape models. *Integrated Assessment and Decision Support*, Lugano, Switzerland, pp. 227–232.

Pastor J and Post WM (1988) Response of northern forests to CO_2-induced climate change. *Nature* 334: 55–58.

Pastor J, Sharp A, and Wolter P (2005) An application of Markov models to the dynamics of Minnesota's forests. *Canadian Journal of Forest Research – Revue Canadienne De Recherche Forestiere* 35: 3011–3019.

Pielke RAJ (2003) The role of models in prediction for decision. In: Canham CD, Cole JJ, and Lauenroth WK (eds.) *Models in Ecosystem Science*, pp. 111–138. Princeton, New Jersey, USA: Princeton University Press.

Rastetter EB (1996) Validating models of ecosystem response to climate change. *BioScience* 46: 190–198.

Reiners WA and Driese KL (2001) The propagation of ecological influences through heterogeneous environmental space. *Bioscience* 51: 939–950.

Roberts DW (1996) Modelling forest dynamics with vital attributes and fuzzy systems theory. *Ecological Modelling* 90: 161–173.

Robinson AP and Ek AR (2003) Description and validation of a hybrid model of forest growth and stand dynamics for the Great Lakes region. *Ecological Modelling* 170: 73–104.

Rykiel EJJ (1996) Testing ecological models: The meaning of validation. *Ecological Modelling* 90: 229–244.

Scheller RM, Hua D, Bolstad PV, Birdsey RA, and Mladenoff DJ (2011) The effects of forest harvest intensity in combination with wind disturbance on carbon dynamics in Lake States mesic forests. *Ecological Modelling* 222: 144–153.

Scheller RM and Mladenoff DJ (2007) Forest landscape simulation models: Tools and strategies for projecting and understanding spatially extensive forest ecosystems. *Landscape Ecology* 22: 491–505.

Scheller RM and Mladenoff DJ (2008) Simulated effects of climate change, tree species migrations, and forest fragmentation on aboveground carbon storage on a forested landscape. *Climate Research* 36: 191–202.

Scheller RM, Spencer WD, Rustigian-Romsos H, Syphard AD, Ward BC, Strittholt JR (2011) Effects of fire and fuels management on Fishers (*Martes pennanti*) in the southern Sierra Nevada, California. *Landscape Ecology* 26: 1491–1504.

Scheller RM, Sturtevant BR, Gustafson EJ, Mladenoff DJ, and Ward BC (2010) Increasing the reliability of ecological models using modern software engineering techniques. *Frontiers in Ecology and the Environment* 8: 253–260.

Schmidt DA, Taylor AH, and Skinner CN (2008) The influence of fuels treatment and landscape arrangement on simulated fire behavior, Southern Cascade range, California. *Forest Ecology and Management* 255: 3170–3184.

Seidl R, Fernandes PM, Fonseca TF, *et al.* (2011) Modelling natural disturbances in forest ecosystems: A review. *Ecological Modelling* 222: 903–924.

Seidl R, Lexer MJ, Jager D, and Honninger K (2005) Evaluating the accuracy and generality of a hybrid patch model. *Tree Physiology* 25: 939–951.

Seidl R, Rammer W, Scheller RM, Spies T, Lexer MJ (in press) Simulating ecological complexity: A scalable, individual-based process model of forest ecosystem dynamics. *Ecological Modelling*.

Shifley SR, Thompson III FR, Dijak WD, and Fan Z (2008) Forecasting landscape-scale, cumulative effects of forest management on vegetation and wildlife habitat: A case study of issues, limitations, and opportunities. *Forest Ecology and Management* 254: 474–483.

Shlisky AJ, Guyette RP, Ryan KC (2005) Modeling reference conditions to restore altered fire regimes in oak-hickory-pine forests: Validating coarse models with local fire history data. In: *Proceedings of the EastFIRE Conference*, Fairfax, VA.

Shugart Jr. HH (1998) *Terrestrial Ecosystems in Changing Environments*. Cambridge, UK: Cambridge University Press.

Spencer WD, Rustigian-Romsos H, Strittholt J, Scheller RM, Zielinski W, and Truex R (2011) Using occupancy and population models to assess habitat conservation opportunities for an isolated carnivore population. *Biological Conservation* 144: 788–803.

Strayer DL, Ewing HA, and Bigelow S (2003) What kind of spatial and temporal details are required in models of heterogeneous systems? *Oikos* 102: 654–662.

Syphard AD, Scheller RM, Ward BC, Spencer WD, and Strittholt JR (2011) Simulating landscape-scale effects of fuels treatments in the Sierra Nevada, California. *International Journal of Fire Management* 20: 364–383.

Taleb NN (ed.) (2007) *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.

Tilman D, May RM, Lehman CL, and Nowak MA (1994) Habitat destruction and the extinction debt. *Nature* 371: 65–66.

Turner MG, Gardner RH, and O'Neill RV (1995) Ecological dynamics at broad scales. Ecosystems and landscapes. *BioScience* S-29: 443–449.

Turner MG, Romme WH, Gardner RH, Oneill RV, and Kratz TK (1993) A revised concept of landscape equilibrium – disturbance and stability on scaled landscapes. *Landscape Ecology* 8: 213–227.

Urban DL, Bonan GB, Smith TM, and Shugart HH (1991) Spatial applications of gap models. *Forest Ecology and Management* 42: 95–110.

Ward NL and Masters GJ (2007) Linking climate change and species invasion: An illustration using insect herbivores. *Global Change Biology* 13: 1605–1615.

White D, Minotti PG, Barczak MJ, *et al.* (1997) Assessing risks to biodiversity from future landscape change. *Conservation Biology* 11: 349–360.

Williams JW, Jackson ST, and Kutzbach JE (2007) Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences* 104: 5738–5742.

With KA, Cadaret SJ, and Davis C (1999) Movement responses to patch structure in experimental fractal landscapes. *Ecology* 80: 1340–1353.

With KA and King AW (1997) The use and misuse of neutral landscape models in ecology. *Oikos* 79: 219–229.

Xu C, Gertner GZ, and Scheller RM (2009) Uncertainties in the response of a forest landscape to global climatic change. *Global Change Biology* 15: 116–131.

Yemshanov D and Perera AH (2002) A spatially explicit stochastic model to simulate boreal forest cover transitions: General structure and properties. *Ecological Modelling* 150: 189–209.

Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production

Wallace E Tyner and Farzad Taheripour, Purdue University, West Lafayette, IN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biofuels-induced land-use changes Production of corn ethanol causes a series of changes in the markets for feed, food, fuel, and many other commodities. In response to these changes, consumers adjust their demands for goods and services and producers alter their demands for intermediate and primary inputs and adjust their supplies with the new market conditions. In this process, the land-use change occurs through market-mediated effects as higher corn price due to higher demand induces farmers to increase corn production, which can lead to conversion of pasture or forest land to agricultural crop use land. When that land is used for agriculture, carbon stored either in the wood or in the soil is released, causing an increase in greenhouse gases.

BTU British Thermal Unit, a unit of energy equal to 1 055.06 joules.

GREET The Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation Model by Argonne National Laboratory. This is a model used to calculate the direct GHG emissions for different biofuel pathways.

GTAP Global Trade Analysis Project. The name also refers to the computable general equilibrium model and the global data base used with the model. Both are available from the GTAP Center at Purdue University, www.gtap.org

Renewable Fuel Standard (known as RFS or RFS2) A US federal government mandate requiring certain quantities of four categories of biofuels to be used with the total reaching 136 billion liters (36 billion gallons) of ethanol equivalent by 2022.

Introduction

US ethanol production increased from 1.7 billion gallons (BGs) in 2001 to about 10 BGs in 2009. According to the Renewable Fuel Standard (RFS2) in the US Energy Independence and Security Act of 2007 (CRS RL34294, 2007), US corn ethanol production would reach 15 BGs in 2015. This level of ethanol production would affect agricultural activities within USA and around the world. In particular, it can cause land-use changes anywhere in the world, and the implications of land-use changes are complex and controversial. Land-use changes associated with increased corn ethanol production are important because the they can affect the CO_2 emissions associated with ethanol production and consumption.

The Argonne National Laboratory (ANL) (Wang, 1999, 2005) has developed a life cycle model, called "Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET)", which estimates the emissions of greenhouse gases (GHGs, including CO_2, CH_4, and N_2O) of alternative fuels including corn ethanol production. The GREET model classifies GHG emissions into three categories: (1) feedstock production, (2) fuel production – corn to ethanol in this case, and (3) vehicle operation. The total emissions associated with the ethanol supply chain are then compared with the analogous calculations for gasoline. However, prior to this research, there was limited data on GHG emissions from land-use changes due to biofuel production included in the GREET model. Rather, the land-use consequences of biofuel production and their corresponding GHG emissions were highlighted in published articles showing that biofuel production could have extraordinary land-use implications that negated the seeming GHG benefits that had been calculated when land-use changes were not fully considered (Searchinger *et al.*, 2008 (henceforth referred to as SEA); Fargione *et al.*, 2008.)

This article reviews, summarizes, and synthesizes land-use changes and CO_2 emissions induced by US corn ethanol production for several alternative economic conditions and modeling assumptions. By combining land-use-related emissions due to ethanol production with net emissions calculated in GREET, it is possible to estimate total GHG emissions associated with corn ethanol production and use. This total can then be compared with gasoline to determine the net gain/loss for corn ethanol production and use compared with gasoline.

When the use of a crop, such as corn, is diverted from food or animal feed to feedstock for fuel, existing demand for the crop as a food or feed could cause new land to be brought into agricultural production to help meet at least some of the unmet demands. The amount of land so cleared can be estimated by using a computational general equilibrium (CGE) model of the global economy including agricultural commodities. For instance, the Global Trade Analysis Project (GTAP) model (Hertel, 1997) can be modified to assess the economic impacts of ethanol production and its land-use implications for the world under alternative sets of assumptions. (GTAP is a CGE model, which considers production, consumption, and trade of a wide range of goods and services by region and at a global scale).

The second step in determining the GHG impacts of a biofuel is to convert land-use changes, such as estimated in GTAP, into the associated CO_2 emissions. This is done using a module that generates CO_2 emissions factors that are used to convert land-use changes into CO_2 emissions. Such conversions depend on global inventories of carbon in soils and vegetation of different types of ecosystems globally. For instance, the Woods Hole Research Center data set on the soil and land cover carbon profiles divides the whole world into 10 regions and provides data on the soil and land cover carbon profiles for each region.

The third step in determining the GHG net emissions from a particular biofuel is to convert the land-use-related emissions into emissions per gallon of 100% ethanol and add these emissions to those calculated in GREET to get total emissions. This can be done either within the GREET model or by direct calculations.

Rather than using the terms "direct" and "indirect" emissions, as is commonly reported in the literature, this article categorizes the emissions as those calculated in GREET and thus associated with the use of corn for the production and combustion of corn ethanol and those associated with land-use changes that result from diverting the use of corn as a food and feed to its use as biofuel feedstock. By some definitions of the term "indirect," the latter GHG emissions would be labeled "indirect" emissions. For clarity, this article labels them as emissions associated with induced land-use changes.

The Complexity and Global Nature of Land-Use Change

Land-use change is a complicated process. It is driven by many factors and varies through time. There are social as well as economic factors involved in land-use changes. The factors vary by culture, region, and economy (the authors are indebted to Gbadebo Oladosu and Keith Kline of Oak Ridge National Laboratory for providing data and useful perspectives on the land-use change process) and cannot be captured in a given analysis. This article focuses on the impacts of the RFS2 level of US corn-based biofuel production. Since corn is a globally produced and consumed commodity, this US policy of necessity has global impacts driven to a significant degree by changes in global supplies of and demands for food and feed grains. Use of a global general equilibrium model can capture many of these market-mediated effects.

The modified GTAP model is used to illustrate how to analyze the economic and environmental consequences of corn ethanol production. Then the simulations and assumptions are discussed that are the basis for the land-use results predicted by GTAP. Next follows a discussion of the calculation of CO_2 emissions associated with this land use.

The evaluation of the impacts of US corn ethanol production on global land-use requires a model that is global in scope and links global production, consumption, and trade. In addition, the model should properly link energy, biofuel, and agriculture markets. Since biofuel, crop, and livestock industries compete through the land market, the model should link these activities through the land market as well. Furthermore, biofuel byproducts, which can be used in animal feedstuffs, bridge these industries. The connection between these industries through biofuel byproducts affects the nature of competition among these industries for land. All of these features are encompassed in a special-purpose version of the GTAP model and its database. This model will henceforth be referred to as GTAP-BIO-ADV (advanced GTAP-BIO model).

The GTAP model simulates the world economy using a global database that contains input–output tables for almost all countries. These tables provide detailed information on production and consumption of commodities and services along with investment and bilateral trade among regions. This database also includes payments to labor, capital, and land (for details, see Dimaranan (2006)). GTAP data come from a multitude of sources. The database also includes the most updated global land cover and land-use database by region disaggregated into 18 Agro-Ecological Zones (AEZs). These AEZs share common climate, precipitation, and moisture conditions. The land cover and land-use database is based on the Center for Sustainability and Global Environment (SAGE) database (for more information on the land-use database, see Lee *et al.* (2005)). The land-use database provides information on global crop yields as well. Version 6 of the GTAP database, which was used in this analysis, covers 57 groups of commodities and services for 87 countries and regions.

Taheripour *et al.* (2007) have expanded the GTAP standard database version 6 and explicitly introduced three biofuel commodities (ethanol from food grains and from sugarcane

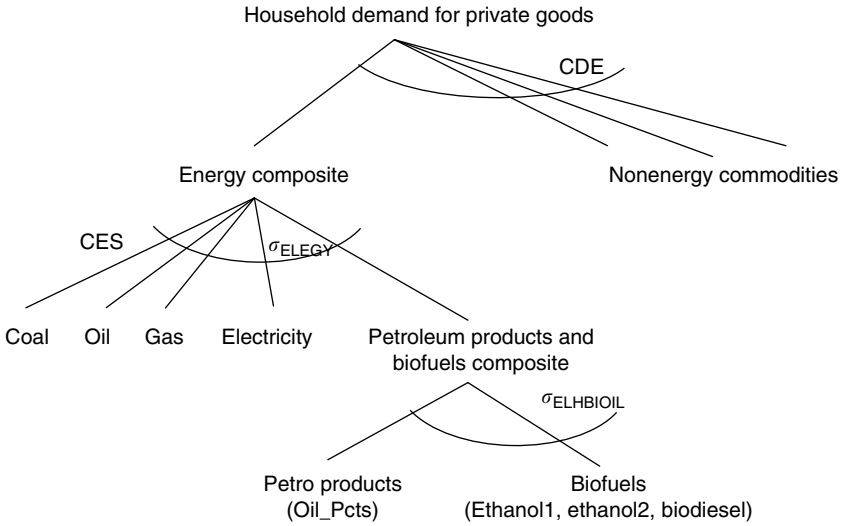


Figure 1 Structure of consumption side of the GTAP-BIO-ADV model.

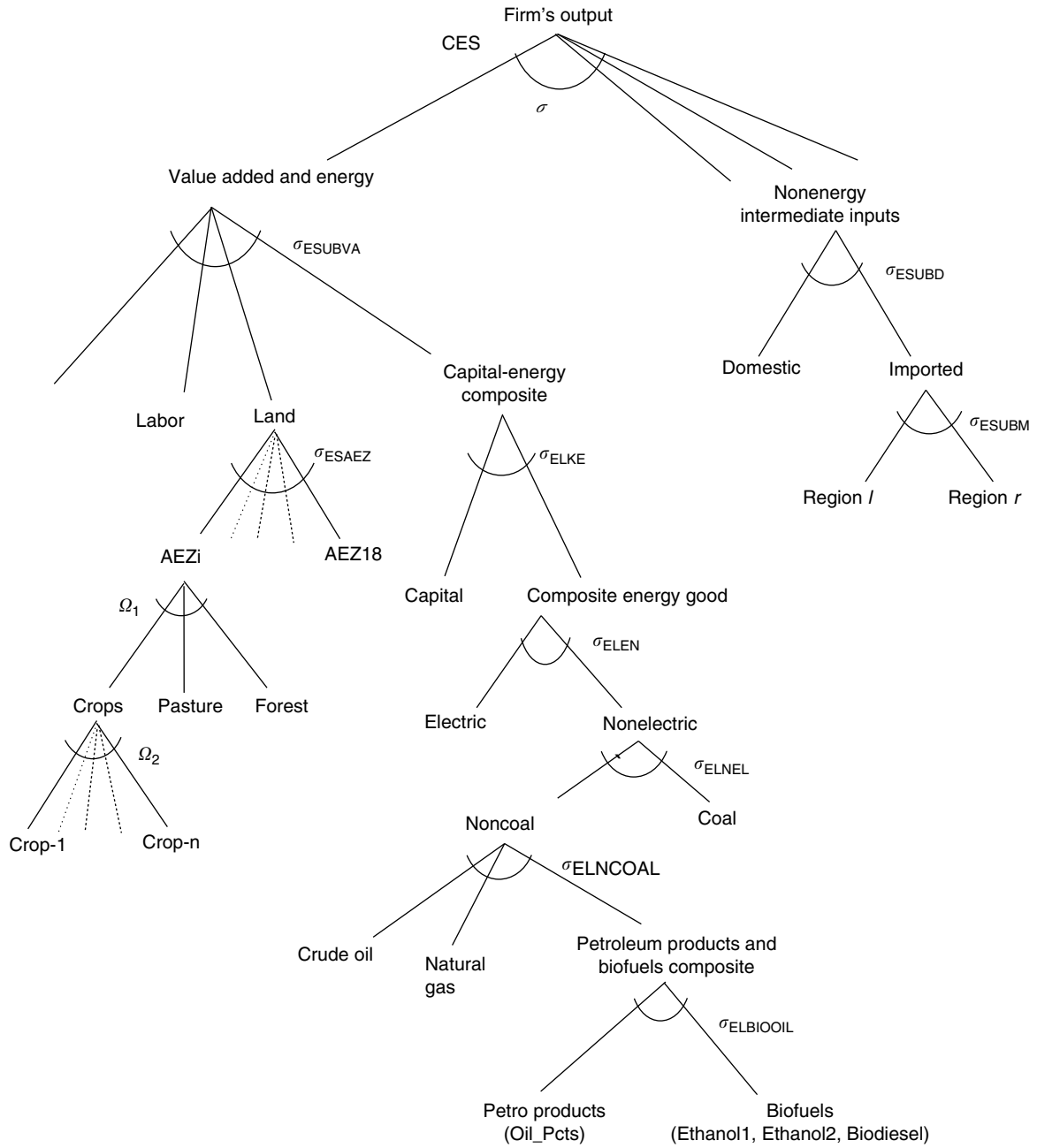


Figure 2 Production structure of GTAP-BIO-ADV.

and biodiesel from oilseeds) into it. Birur *et al.* (2008) have incorporated biofuels into the GTAP-E model. (GTAP-E was originally developed by Burniaux and Truong (2002) to incorporate energy into the GTAP framework, and has recently been modified by McDougall and Golub (2007).) They augment the model by adding the possibility for substitutability between biofuels and petroleum products. **Figures 1** and **2** represent the structure of consumption and production sides of this model. In these figures, CES stands for constant elasticity of substitution, and CDE for constant difference elasticity and is the means of expressing household preferences in GTAP.

Figure 1 indicates that households could use biofuel as a substitute for petroleum products in GTAP-BIO-ADV. On the contrary, **Figure 2** shows that at the bottom-most level of the production side, biofuels are a complement to petroleum products in the production process. It should be noted that in a general equilibrium model like GTAP, all the equations are solved simultaneously, so it is not a stepwise solution process.

Hertel *et al.* (2008) augmented this model with a land-use module to better depict the global competition for land among land-use sectors. The land-use module traces changes in the demand for land across the world at the AEZ level and thereby captures the potential for real competition between alternative land uses. In this module, land does not move across AEZs. However, the distribution of land across its alternative uses can change within each AEZ. Alternative uses

of land are as follows: forest, grassland, and cropland. In this module, livestock producers compete to use grassland, and there is competition among agricultural activities to use croplands. Corn falls in the category of coarse grains along with sorghum, oats, and barley. However, in the USA, this group comprises mostly corn. For example, in 2009, corn constituted 95.4% of the coarse grain production (by weight). Most of the remaining produce was sorghum, which also could be used for biofuels.

Recently, Birur (2010) added two new land categories of cropland pasture and unused cropland (e.g., retired cropland under the US Conservation Reserve Program (CRP)) into supply of land. **Figure 3** represents the new structure of land supply in the modified model.

In the new land supply, tree cropland pasture and unused cropland (mainly CRP) are explicitly defined as components of cropland. CRP land mainly generates environmental benefits. Hence, this type of land is introduced as an input into the sector that provides these services (i.e., Oth_Ind_Se). Cropland pasture is an input into the livestock industry. To facilitate transition of cropland pasture from livestock industry to crop production and vice versa, an industry is added to the model that uses cropland pasture as an input and sells its output (cropland pasture) to the livestock industry. This industry competes in the land market with crops. Finally, the livestock industry combines cropland pasture with pasture land in its production function, as shown in **Figure 4**. This figure indicates that the livestock industry combines pasture land with cropland pasture in the value-added nest and uses feed and nonfeed inputs in its production function.

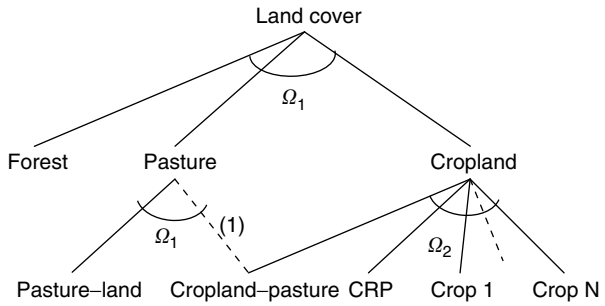


Figure 3 Land cover and land-use activities in the GTAP-BIO-ADV.

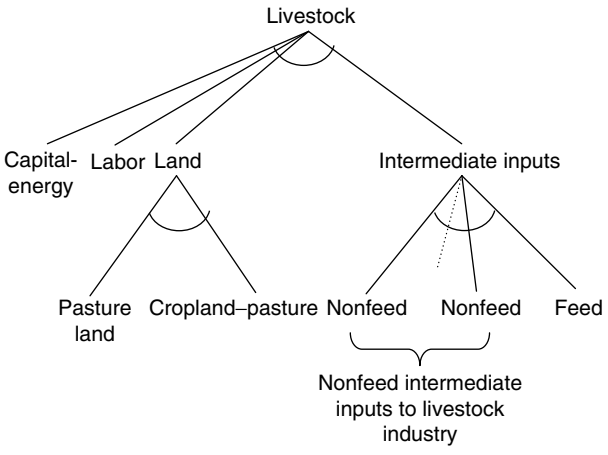


Figure 4 Production structure of the livestock industry.

The land-use module determines expansion of cropland and its distribution among agricultural activities according to two important parameters: price elasticity of yield and ratio of productivities of marginal and average lands. The price elasticity of yield measures changes in crop yield due to the changes in crop price. In the simulations it was assumed that the price elasticity of yield is equal to 0.25. Keeney and Hertel (2008) have provided a detailed discussion on this parameter along with econometric evidence behind it.

The ratio of marginal and average productivities measures the productivity of new cropland versus the productivity of existing cropland. This ratio will henceforth be referred to as ETA (the Greek letter Eta). A set of regional ETAs at the AEZ level is obtained from a bio-process-based biogeochemistry model (terrestrial ecosystem model; Zhuang *et al.*, 2003) along with spatially referenced information on climate, elevation, soils, and vegetation land-use data (for details see Tyner *et al.*, 2010). Regional ETAs vary across the world and among AEZs.

Taheripour *et al.* (2010, 2011) added production, consumption, and trade of biofuel byproducts into the GTAP modeling framework, extending the original GTAP-BIO database (Taheripour *et al.*, 2007) in several directions to better trace the links among biofuel, vegetable oil, food, feed, and livestock industries. Unlike the initial database, these articles distinguish between feedstock of the US and EU ethanol industries. The ethanol industry also produces distiller's dried grains with solubles (DDGS). They also split the "other food products" industry into two distinct industries: processed food and processed feed. In addition, they split the vegetable oil sector into two distinct industries: crude vegetable oil and refined vegetable oil. The crude vegetable oil sector uses oilseeds and produces crude vegetable oil (as the main product) and oilseed meal (as the byproduct). Unlike the original GTAP-BIO database, which directly converts oilseeds to biodiesel, they introduce a biodiesel production technology that uses crude vegetable oil and other inputs to produce biodiesel.

In addition, the latter article uses a three-level nesting structure for the demand for animal feedstuffs in the livestock industry, which brings more flexibility into this part of the model. **Figure 5** depicts this nesting structure. At the lower level of this nesting structure, DDGS and coarse grains are combined to create an energy feed. At this level, oilseeds and oilseed meals are combined to create a protein feed as well for countries that use oilseeds directly as feed. At a higher level, the protein and energy feed ingredients are combined. At this level, other crops also are bundled together. The livestock industry receives some inputs from processed livestock industry as well, and these materials are bundled together at the second level, too. Finally, all feed ingredients are combined to create the feed composite.

Taheripour *et al.* (2011) assigned elasticities of substitution to the different components of the demand for feed to replicate changes in the prices of DDGS and meals in USA and the EU during the time period 2001–2006. In addition, they did several experimental simulations and sensitivity tests to reach displacement ratios between DDGS, grains, oilseeds, and oilseed meals according to the literature in this area. Since oilseeds and oilseed meals are good substitutes in some

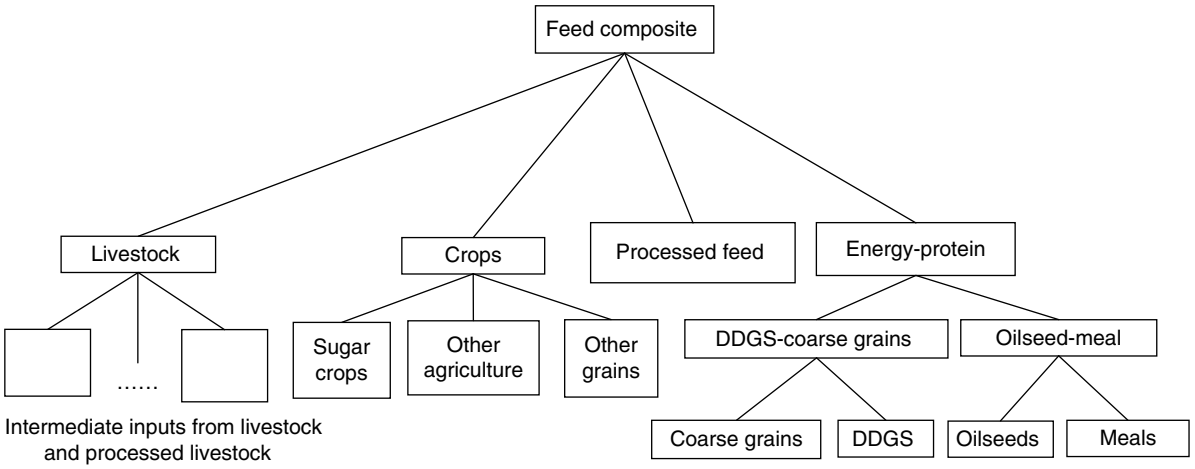


Figure 5 Structure of the nested demand for feed in livestock industry.

Table 1 Cost shares of major feed items in the US livestock industries in 2001 and 2006[a]

Feed items	*2001*			*2006*		
	Dairy	*Meat ruminant*	*Nonruminant*	*Dairy*	*Meat ruminant*	*Nonruminant*
Coarse grains	67.6	68.4	82.9	61.0	57.5	82.5
Other crops	6.4	10.4	2.9	6.0	9.8	2.7
DDGS	5.6	6.4	1.1	12.9	17.6	2.1
Oilseeds meals	20.3	14.9	13.1	20.1	15.0	12.7

[a]Processed feed is dropped from this table to highlight shares of items listed in the table.

regions, they applied a relatively high elasticity of substitution, 20, between these two feed materials for all types of animal species. Following the literature, they used values of 25, 30, and 20 for the elasticities of substitution between coarse grains and DDGS in the dairy farms, other ruminant, and nonruminant feed structure, respectively. They also applied a small value, 0.3, for the elasticity of substitution between the energy and protein feedstuffs because DDGS could displace a portion of meals in some feed rations, as shown in Arora *et al.* (2008) and Fabiosa (2009). In the composite of other crops and composite of processed livestock inputs, they applied elasticities of substitution of 1.5 for all types of livestock industry. Finally, following Keeney and Hertel (2005), they used elasticity of substitution of 0.9 at the higher level of the feed demand nest.

This article uses some GTAP simulation results to show how these elasticities shape the cost structure of the livestock industry. To accomplish this task, it uses the results obtained from the simulations introduced in its next section. In particular, it uses the results of the first simulation of the second group of experiments. This particular simulation replicates transition of the global economy from 2001 to 2006. The results of this simulation predict that the cost shares of coarse grain, other crops, and meals in the US livestock industries declined during the time period 2001–2006, whereas the cost share of DDGS increased. The largest substitution is DDGS for coarse grains, but there is also substitution for other crops and oilseed meals, depending on the livestock species. Note that processed feed was dropped from the list of animal feeds to highlight the changes in the shares of crops, DDGS, and meals during this time period (**Table 1**).

Global markets (including agriculture) are complex and highly linked. This complexity and these linkages make it difficult to accurately estimate how a market perturbation, such as the use of sufficient US corn to produce 15 BGs of ethanol, would impact land clearing and other aspects of land-use change around the world. This discussion indicates how a CGE model of the global agricultural economy has been structured to allow it to make such a prediction. In particular, the GTAP-BIO-ADV model and its database described earlier in this section and in the associated references have the following features:

1. It covers production, consumption, and trade of three types of biofuels: ethanol from crops and from sugarcane, and biodiesel from crude vegetable oil.
2. Byproducts are DDGS and oilseeds meals.
3. The crude vegetable oil industry uses oilseeds and produces crude vegetable oil and oilseed meals.
4. The biodiesel industry uses crude vegetable oil to produce biodiesel.
5. The demand for feedstuffs follows a three-level nesting structure.
6. The land module handles two new land categories of unused cropland and cropland pasture. Although the model could trace changes in these two groups of land

across the world, there are data on cropland pasture for the USA and Brazil and data on CRP only for the USA.
7. ETA has been calibrated for each AEZ and region instead of using the globally fixed ETA parameter as in the past.
8. Energy demand and supply elasticities have been recalibrated for this version.
9. The world economy is divided into 19 regions, 34 groups of commodities and services, 32 industries, and five groups of endowments.
10. When US ethanol is shocked, the production of other biofuels is held constant.

GTAP Simulations and their Results

This article will now consider the results of three different sets of analyses. In the first group, it follows the approach used in Tyner *et al.* (2009). In this approach, the land-use implications of US ethanol production off the 2001 database are calculated. This approach isolates impacts of US ethanol production from other changes that shape the world economy. This method assumes that other factors such as population growth, yield improvement, and economic growth do not affect the land-use implications of producing more ethanol from agricultural resources. Hertel *et al.* (2010) provide more insights into this approach. Even though we use the 2001 starting point, the results are different from the results reported in Tyner *et al.* (2009) because a new model is used that includes all the changes described above.

In the second approach, a baseline that represents changes in the world economy during the time period 2001–2006 was constructed. Then the land-use impact of US ethanol production off the updated 2006 database was calculated, while following the principles of the first group of simulations for the time period 2006–2015. Finally, in the third approach, the updated 2006 database obtained from the second group of simulations was used, but it was assumed that during the time period 2006–2015, population and crop yields will continue to grow. These are two important factors that could alter the land-use impacts of ethanol production in the future. These three groups of simulations and their results are described in the rest of this section.

Group 1: Simulations with No Economic and Yield Growth and 2001 Base

The land-use implications of the US ethanol production for the following six production segments were calculated:

- Ethanol production from 2001 to 2006 level.
- Ethanol production from 2006 level to 7 BGs.
- Ethanol production from 7 to 15 BGs by increments of 2 BGs.

The global biofuel industry has followed a rapid growth path during the time period 2001–2006. The historical observations from this time period have been used to calibrate the biofuel parameters of the model (Hertel *et al.*, 2008). Then gradual increases in the production of US ethanol after 2006 have been considered to evaluate marginal impacts of ethanol production. For this purpose, first the US ethanol production is increased from its 2006 level (4.855 BGs) to 7 BGs. Thereafter, ethanol production is increased by increments of 2 BGs to achieve the goal of 15 BGs in 2015.

Table 2 summarizes the land-use changes from the first set of simulations. These results indicate that producing 13.23 BGs of ethanol (from the 2001 production level to 15 BGs) requires about 2.96 million ha of additional land, of which 1.01 million ha (34%) are expected to be in the USA, with the reminder (1.95 million ha) in other regions (66%). This result suggests that the land-use changes due to US ethanol production will mainly take place outside the USA. Results from this group of simulations also indicate that the size of required land to achieve the 15 BGs of ethanol production is much smaller than the land-use changes suggested by a simple calculation that ignores important factors that could mitigate land-use impacts of ethanol production. (One can determine the land requirement for US ethanol production by multiplying corn yield (370 bushels per hectare of land) by a corn-to-ethanol conversion factor (e.g., 2.7 gallons per bushel of corn) and by the target for ethanol production. Based on this simple calculation, increasing ethanol production from its 2001 level (1.77 BG) to 15 BG needs about 13 million ha of land.) Several factors mitigate the land-use consequences of ethanol production. Among them are less corn consumption in the livestock industry due to use of more DDGS in the livestock industry, reductions in output of the livestock industry, reallocation of croplands across the world among alternative crops, and higher yields in crop production

Table 2 Global land-use changes due to the US ethanol production: off the 2001 database

Changes in US corn ethanol production	*Land-use changes (ha)*			*Distribution of land use changes (%)*			*Hectares per 1000 gallons*
	Within USA	*Other regions*	*World*	*Within USA*	*Other Regions*	*World*	
3.085 BG (2001–2006)	227,982	382,394	610,376	37.4	62.6	100.0	0.20
2.145 BG (2006–7 BG)	162,558	297,766	460,324	35.3	64.7	100.0	0.21
2.000 BG (7–9 BG)	152,990	295,051	448,041	34.1	65.9	100.0	0.22
2.000 BG (9–11 BG)	154,018	310,639	464,657	33.1	66.9	100.0	0.23
2.000 BG (11–13 BG)	154,706	325,639	480,345	32.2	67.8	100.0	0.24
2.000 BG (13–15 BG)	155,000	340,311	495,311	31.3	68.7	100.0	0.25
13.23 BG (2001–15 BG)	1,007,253	1,951,800	2,959,053	34.0	66.0	100.0	0.22

due to higher prices. Hertel *et al.* (2010) provide decomposed contributions of these factors in mitigating the land-use impacts of ethanol production.

The land required to increase the US ethanol production from its 2001 level to 15 BGs obtained from these simulations is smaller than its corresponding value reported in Tyner *et al.* (2009) by about 16.7% (i.e., 2.96 vs. 3.55 million ha). Two major modifications in the GTAP model contribute to this reduction. A portion of this reduction is associated with the land conversion factors. As noted earlier, a new set of regional land conversion factors at the AEZ level is used in this work. The new land conversion factors in several AEZs are higher than the single conversion factor of 0.66, which was used in earlier work (for details see Tyner *et al.*, 2010).

Introducing the new land categories (cropland pasture and unused land) (in these simulations, the area of US CRP land is held constant) into the model also contributes to the reduction in land requirement. In particular, in the USA and Brazil, in the presence of cropland pasture, farmers convert a portion of this type of land to crop production. For example, an increase in US ethanol production from its 2001 level to 15 BGs leads to the conversion of 1.2 million ha from cropland pasture to cropland. Indeed, a portion of this land conversion prevents sharp reductions in production of other crops. It is important to note that the competition between crop and livestock industry prevents full conversion of cropland pasture to crop production. These two modifications not only reduce the land requirement of ethanol production, but also alleviate the adverse impact of ethanol production on the prices and consumption of crops.

Table 2 also indicates that the required land for producing 1000 gallons of ethanol grows with higher levels of ethanol production. For example, for the 2001–2006 simulation, an additional 3.085 BGs of ethanol triggers global land-use changes of roughly 610 ha. This is equal to 0.20 ha per 1000 gallons of ethanol. However, for the 13–15 BGs simulation, an additional 1000 gallons of ethanol requires 0.25 ha of land. To increase ethanol production from the 2001 level to 15 BGs, on average about 0.22 ha of land per 1000 gallons of ethanol are required. The marginal level (0.25) is higher than the average (0.22), which would be expected because as more land comes into production, the yields on the incremental area would be lower.

Table 3 depicts another aspect of the land-use implications of US ethanol production. This table shows the distribution of land-use changes between forest and grassland. About 25% of the required croplands needed to increase ethanol production from its 2001 level to 15 BGs come from forest, and the remaining (75%) come from grasslands. **Table 3** also indicates that with higher levels of ethanol production, the portion of forests in the converted land into crop production increases very slightly (from 23% in 2001 to 25% at the 15 BGs ethanol production).

In the absence of crop yield growth, the increasing global land-use change, given equal increments of US ethanol production, is explained by the differences in the productivity of available lands. Productive lands are employed first before marginal lands, which have lower productivity and lower yields. At low levels of production, more productive lands are available; hence, less land is required to produce additional ethanol. However, at higher levels of ethanol production, most of the productive land is already being used, and only marginal land is available. Given this, more marginal land is required to produce the same increment of US corn ethanol production.

Group 2: Simulations with Updated Baseline for the Time Period 2001–2006

The global economy changed significantly over the 2001–2006 period. Countries followed different economic growth paths, population increased everywhere at different rates, land productivity rapidly increased in many regions (with some exceptions), and technology improved in many areas. These are important factors that could alter the land-use implications of biofuels. In the second group of simulations, these factors are taken into account.

To accomplish this task, a database was developed to include data on crop production, harvested area, forest areas, gross capital formation, labor force (skilled and unskilled), gross domestic product, and population for the whole world at the country level. **Table 4** contains the percentage changes in macroeconomic variables (for details see Tyner *et al.* (2010)). Then this data set was used to generate a baseline that replicates the transition of the global economy from 2001 to 2006, whereas global biofuel production during this time period was targeted in the presence of population, income, and yield growths. In building the baseline, the model was guided to replicate the historical paths of changes in harvested

Table 3 Global land-use changes due to US ethanol production: off the 2001 database

Changes in US corn ethanol output	*Land-use changes (ha)*			*Distribution of land-use changes (%)*		
	Forest	*Grassland*	*Crop*[a]	*Forest*	*Grassland*	*Total*[a]
3.085 BG (2001–2006)	−143,716	−466,652	610,376	23.5	76.5	100.0
2.145 BG (2006–7 BG)	−114,409	−345,912	460,324	24.9	75.1	100.0
2.000 BG (7–9 BG)	−112,330	−335,712	448,041	25.1	74.9	100.0
2.000 BG (9–11 BG)	−116,795	−347,864	464,657	25.1	74.9	100.0
2.000 BG (11–13 BG)	−120,688	−359,650	480,345	25.1	74.9	100.0
2.000 BG (13–15 BG)	−124,151	−371,156	495,311	25.1	74.9	100.0
13.23 BG (2001–15 BG)	−732,089	−2,226,946	2,959,053	24.7	75.3	100.0

[a]The difference between the changes in cropland and the sum of forest and grassland is due to rounding. Cropland pasture is included in cropland.

areas across the world as well. Furthermore, changes in global forest area are traced to match the land use results with the historical changes in forest areas during the time period 2001–2006. Rates of technological improvements were adusted to trace changes in cropland and forest areas. **Table 5** contains the cumulative yield change for each region and crop category over the 5 years.

To generate the 2006 baseline, the major macroeconomic variables are shocked according to the historical observations for the time period 2001–2006. In particular, GDP, gross capital formation, labor force, and population are shocked at the regional level. The global biofuel outputs are also shocked according to actual observations for the same time period. In addition to these shocks, the model is calibrated to replicate observed improvement in yield over the time period 2001–2006 by crop and by region. Finally, technological changes in input–output ratios are introduced to replicate regional changes in the harvested area during this time period. Furthermore, the model is calibrated to trace changes in forest area during the baseline time period. For the list of implemented shocks, see Tyner *et al.* (2010). This experiment provides a new database that represents the world economy in 2006 in the presence of changes in the major derivers of the world economy. To separate out the impacts of the US ethanol program from other derivers of the world economy, this experiment is repeated without the US ethanol shock. The difference between the land-use implications of these two simulations gives us the impact of the US ethanol program for the said time period.

Then the updated 2006 database is used to evaluate the land-use impacts of increasing US ethanol from its 2006 level to 15 BG incrementally. **Table 6** summarizes the global land-use implications obtained from the second group of simulations.

The results obtained from the second group of simulations indicate that about 2.0 million ha of cropland are needed to increase ethanol production from the 2001 level to 15 BGs. This figure is smaller than its corresponding figure obtained from the first group of simulations by 31%. Two main factors contribute to this reduction. During the time period 2001–2006, crop yields were growing faster than the demands for crops in many regions. This reduced the size of land-use changes in this period. In this case, the land-use implications of US ethanol for the time period 2006–2015 is smaller than the corresponding figure of the first cases because crop yields are higher in the updated database.

In the second group of simulations, cropland pasture moves to crop production faster than in the first group of simulations. In the presence of economic growth, about 5.3 million ha of cropland pasture will move to crop production.

Table 4 Percentage changes in macroeconomic variables (2001–2006)

Regions	*Population*	*GDP*	*Skilled labor*	*Unskilled labor*	*Capital*
1 USA	5.2	15.0	5.7	5.2	18.9
2 EU27	1.82	10.2	7.4	− 1.1	13.1
3 Brazil	6.88	17.2	24.4	8.5	11.1
4 CAN	5.31	14.6	9.1	8.3	34.0
5 Japan	0.59	8.8	0.2	− 4.1	0.7
6 CHIHKG	3.59	59.0	17.5	4.7	83.6
7 India	8.51	45.9	27.5	8.7	94.8
8 C_C_Amer	6.41	16.8	33.7	6.8	25.4
9 S_o_Amer	7.19	24.4	50.2	10.1	54.4
10 E_Asia	2.75	25.9	15.1	5.3	21.6
11 Mala_Indo	7.18	29.1	56.5	9.0	30.1
12 R_SE_Asia	7.2	33.7	26.6	9.3	43.0
13 R_S_Asia	10.8	32.5	34.4	15.5	39.0
14 Russia	− 2.38	37.7	2.2	1.2	69.6
15 Oth_CEE_CIS	2.27	25.5	14.9	− 2.2	40.0
16 Oth_Europe	2.27	25.5	14.9	− 2.2	40.0
17 MEAS_NAfr	10.18	26.7	30.7	19.1	47.8
18 S_S_AFR	13.47	27.5	17.3	13.6	45.2
19 Oceania	7.79	17.4	11.1	8.5	54.8

Table 5 Percentage change in yield (accumulation of growth rates 2001–2006)

Region/crop	*Wheat and paddy rice*	*Coarse grains*	*Oilseeds*	*Sugarcane*	*Other agriculture*
1 USA	− 2.3	11.0	11.6	1.8	− 7.3
2 EU27	4.0	7.3	13.5	7.8	− 1.8
3 Brazil	12.4	22.8	3.5	8.1	9.3
4 CAN	10.8	10.2	14.4	33.3	18.1
5 Japan	− 4.1	− 18.4	− 8.6	5.1	-0.5
6 CHIHKG	6.3	17.0	5.6	42.6	5.2
7 India	5.3	16.4	15.6	− 4.1	− 2.4
8 C_C_Amer	4.0	13.2	28.6	13.2	5.4
9 S_o_Amer	10.0	9.0	-0.7	6.4	3.5
10 E_Asia	5.6	48.3	3.6	0.0	5.6
11 Mala_Indo	4.3	19.4	27.4	9.3	19.8
12 R_SE_Asia	10.1	18.1	10.8	− 4.6	15.6
13 R_S_Asia	6.8	37.8	− 5.1	4.4	11.5
14 Russia	20.8	17.2	22.2	48.8	15.0
15 Oth_CEE_CIS	15.1	26.0	16.7	22.6	13.5
16 Oth_Europe	15.1	26.0	16.7	22.6	13.5
17 MEAS_NAfr	20.3	25.3	46.6	4.3	1.7
18 S_S_AFR	6.4	9.8	10.2	− 5.7	3.4
19 Oceania	10.9	− 9.6	0.7	2.2	17.3

Table 6 Simulated global land-use changes due to US ethanol production: off the updated baseline

Changes in US corn ethanol production	*Land-use changes (ha)*			*Distribution of land-use changes (%)*			*Hectares per 1000 gallons*
	Within USA	*Other regions*	*World*	*Within USA*	*Other regions*	*World*	
3.085 BG (2001–2006)	106,870	360,397	467,268	22.9	77.1	100.0	0.15
2.145 BG (2006–7 BG)	77,989	246,464	324,452	24.0	76.0	100.0	0.15
2.000 BG (7–9 BG)	73,308	233,222	306,529	23.9	76.1	100.0	0.15
2.000 BG (9–11 BG)	73,754	233,992	307,746	24.0	76.0	100.0	0.15
2.000 BG (11–13 BG)	74,717	238,378	313,094	23.9	76.1	100.0	0.16
2.000 BG (13–15 BG)	75,731	242,685	318,416	23.8	76.2	100.0	0.16
13.23 BG (2001–15 BG)	482,368	1,555,137	2,037,506	23.7	76.3	100.0	0.15

Table 7 Simulated global land-use changes due to US ethanol production: off the updated baseline

Changes in US corn ethanol output	*Land-use changes (ha)*			*Distribution of land-use changes (%)*		
	Forest	*Grassland*	*Crop*[a]	*Forest*	*Grassland*	*Total*[a]
3.085 BG (2001–2006)	− 151,706	− 315,487	467,268	32.5	67.5	100.0
2.145 BG (2006–7 BG)	− 105,357	− 219,095	324,452	32.5	67.5	100.0
2.000 BG (7–9 BG)	− 99,673	− 206,854	306,529	32.5	67.5	100.0
2.000 BG (9–11 BG)	− 100,005	− 207,740	307,746	32.5	67.5	100.0
2.000 BG (11–13 BG)	− 101,633	− 211,466	313,094	32.5	67.5	100.0
2.000 BG (13–15 BG)	− 103,423	− 214,992	318,416	32.5	67.5	100.0
13.23 BG (2001–15 BG)	− 661,797	− 1,375,633	2,037,506	32.5	67.5	100.0

[a]The difference between the changes in cropland and the sum of forest and grassland is due to rounding.

Table 7 represents distributions of land-use changes between forest and pasture for the second group of simulations. In this group of simulations, on average about 33% of required land for ethanol production comes from forest land. This figure is higher than the corresponding figure of the first group of simulations (25%).

Group 3: Simulations with Crop Yield and Population Growth for the Time Period 2006–2015

Some advocates of the US corn ethanol program argue that crop yields will increase in the future such that this increase could eliminate the land-use implications of ethanol production. This argument neglects the impacts of the future changes in the demand for crops. Demands for crops could increase in the future due to several factors such as changes in population and income, dietary transition as poorer countries consume more meat, or technological progress. In other words, one cannot examine yield (supply) increases alone. One must also include assumptions about increases in crop demand as well. In the third group of simulations, impacts of changes in crop yields and demand are examined as important items that could determine demand and supply for crops and food products.

In the third group of simulations, population growth is introduced into the model as a proxy for food demand increase. It is assumed that population will continue to grow globally during the time period 2006–2015 at the annual growth rate of 2001–2006. It is also assumed that crop yield will increase uniformly at 1% annually after 2006 in all regions and across all types of crops. Although 1% might seem small, it is actually a large number as it is applied in all regions and for all crops. Finally, it is assumed that the regional demands for forest products will increase according to their annual rates of 2001–2006. The latter assumption is made to maintain the long-run pattern in the outputs of forest products. These simulations also include all the changes incorporated in the baseline simulation of the second group of simulations.

To find the land-use impacts of the US ethanol program under these assumptions, simulations were done with and without US ethanol production off the 2006 updated database (obtained in the second group of simulations) for the time period 2006–2015. To understand the land-use implications of the US ethanol program under these assumptions, the land-use implications were first analyzed with no US ethanol production. Table 8 indicates land-use changes due to the yield and population growth for USA, EU, Brazil, and other regions.

This table indicates that after 2006, the cropland areas of USA, EU, Brazil, and other regions would fall due to the simultaneous shocks in yield and population growth. This means that yield growth would dominate the demand growth for crops, and therefore the demand for cropland decreases everywhere. In addition to that, the yield growth contributes to higher levels of food consumption everywhere.

The simulation results indicate that consumption of crops and food products grow faster than population everywhere across the world. This indicates that the yield effect works through two channels: (1) reduction in crop land area needed

Table 8 Simulated global land-use changes due to population and yield growth after 2006 (figures are in 1000 ha)

Period	*Land cover*	*USA*	*EU*	*Brazil*	*Other*	*World*
2006–2007	Forestry	132.2	216.6	419.1	2,168.9	2,936.8
	Cropland	− 163.8	− 241.7	− 106.5	− 2,428.4	− 2,940.5
	Pastureland	31.7	25.1	− 312.6	259.5	3.7
2007–2009	Forestry	357.0	562.3	910.7	6,157.2	7,987.1
	Cropland	− 369.0	− 570.3	− 246.9	− 6,065.2	− 7,251.4
	Pastureland	12.0	8.0	− 663.8	− 91.9	− 735.7
2009–2011	Forestry	522.9	778.8	1083.4	9,211.1	11,596.3
	Cropland	− 410.9	− 693.1	− 279.9	− 7,889.1	− 9,273.0
	Pastureland	− 112.0	− 85.8	− 803.5	− 1,322.0	− 2,323.3
2011–2013	Forestry	712.4	1013.8	1305.5	12,550.1	15,581.9
	Cropland	− 457.8	− 820.4	− 310.0	− 9,859.0	− 11,447.1
	Pastureland	− 254.6	− 193.5	− 995.5	− 2,691.2	− 4,134.7
2013–2015	Forestry	979.3	1348.3	1643.3	15,831.4	19,802.3
	Cropland	− 530.6	− 983.9	− 353.6	− 11,628.8	− 13,496.8
	Pastureland	− 448.7	− 364.4	− 1289.7	− 4,202.6	− 6,305.5

Table 9 Simulated global land-use changes due to US ethanol production: with yield and population growth after 2006

Changes in US corn ethanol production	*Land-use changes (ha)*			*Distribution of Land-use changes (%)*			*Hectares per 1000 gallons*
	Within USA	*Other Regions*	*World*	*Within USA*	*Other Regions*	*World*	
3.085 BG (2001–2006)	106,870	360,397	467,268	22.9	77.1	100.0	0.15
2.145 BG (2006–7 BG)	58,373	175,123	233,496	25.0	75.0	100.0	0.11
2.000 BG (7–9 BG)	57,966	177,186	235,151	24.7	75.3	100.0	0.12
2.000 BG (9–11 BG)	60,830	184,916	245,746	24.8	75.2	100.0	0.12
2.000 BG (11–13 BG)	65,116	199,837	264,953	24.6	75.4	100.0	0.13
2.000 BG (13–15 BG)	70,656	206,057	276,713	25.5	74.5	100.0	0.14
13.23 BG (2001–15 BG)	419,811	1,303,516	1,723,327	24.4	75.6	100.0	0.13

to satisfy demand and (2) higher per capita consumption of food. This means that 1% yield improvement will not end with 1% reduction in cropland, even if there is no population growth.

The released croplands are going to forest to support the long-run growth in forest products. Note that as mentioned earlier, in this group of simulations it is assumed that the global forest sector will continue to grow according to its 2001–2006 growth rate.

After this discussion, impacts of adding biofuel shocks into this picture can be examined. In general, the US ethanol program in this group of simulations generated smaller land-use changes compared to the results of the second group of simulations. Table 9 shows that under the assumptions of this group of simulations, about 1.7 million ha of cropland are needed to increase ethanol production from the 2001 level to 15 BGs. This figure is smaller than the corresponding figure obtained from the second group of simulations by 20%. For the earlier time segments, after 2006 the size of land requirement is significantly smaller than what was observed in the second group of simulations. For example, in this group of simulations, only 0.11 ha of cropland are needed to produce 1000 gallons of ethanol in the time period 2006–2007, whereas the corresponding number obtained from the second group of simulations is about 0.15.

As the year 2015 is approaching, the population growth is dominating the yield growth in some regions and the land requirement is growing. Table 9 shows that the share of USA in land requirement is about 24% on average in this group of simulations.

The distribution of land-use changes between forest and pasture land is similar to that in the second group of simulations. The assumption on the regional demands for forest products drives this result. It is very important to note that adding income growth or changes in other economic factors into this picture may change the geographical distribution of land-use changes or the distribution of the land requirement for ethanol production between forest and grassland (Table 10).

Land-Use CO_2 Emission Factors

Emission factors are used to convert land-use changes into land-use CO_2 emissions. Land conversions of forest and grassland into crop production release CO_2 emissions from two sources: (1) direct CO_2 emissions from land conversion and (2) foregone CO_2 sequestration by forests. The direct CO_2 emissions consist of carbon stored in the vegetation and in the soil, which are released when forests or grasslands are cleared and converted into croplands. The forgone carbon

Table 10 Simulated global land-use changes due to US ethanol production: with yield and population growth after 2006

Changes in US corn ethanol output	*Land-use changes (ha)*			*Distribution of land-use changes (%)*		
	Forest	*Grassland*	*Crop*[a]	*Forest*	*Grassland*	*Total*[a]
3.085 BG (2001–2006)	−151,706	−315,487	467,268	32.5	67.5	100.0
2.145 BG (2006–7 BG)	−75,942	−157,560	233,496	32.5	67.5	100.0
2.000 BG (7–9 BG)	−76,424	−158,735	235,151	32.5	67.5	100.0
2.000 BG (9–11 BG)	−79,870	−165,871	245,746	32.5	67.5	100.0
2.000 BG (11–13 BG)	−86,227	−178,732	264,953	32.5	67.5	100.0
2.000 BG (13–15 BG)	−899,32	−186,782	276,713	32.5	67.5	100.0
13.23 BG (2001–15 BG)	−560,101	−1,163,167	1,723,327	32.5	67.5	100.0

[a]The difference between the changes in cropland and the sum of forest and grassland is due to rounding.

Table 11 Estimated marginal land-use emissions per gallon of E100 for 13–15 billion gallons simulation (30-year payoff method)

Total 30-year emissions from land-use changes (million metric tons)	110.77
Change in ethanol production (million gallons) per year	2,000
Emissions (metric tons per gallon-year of ethanol)	0.0554
Emissions (grams per gallon-year of ethanol)	55,386
One year marginal emissions (grams per gallon of ethanol)	1,846

sequestration accounts for the amount of carbon that could have been stored from annual forest growth, if land had remained forested. This is the opportunity costs of cleared land in terms of its potential to store carbon.

As mentioned earlier (*see* Introduction) the Woods Hole data set (taken from the supporting documents of SEA) is relied upon to develop the land-use emissions factors. This data set divides the world into 10 homogenous regions, determines distributions of forests and grasslands within each region across different types of vegetation cover, and provides detailed information on the carbon stored in the vegetation and in the soil of forests and grasslands within each region.

The Woods Hole data set provides two key carbon figures for each type of land according to its natural vegetation. These figures are carbon stored in the soil and carbon stored in the vegetation. It is assumed that when a natural vegetation area (either forest or grassland) is converted to cropland, about 25% of the carbon stored in its soil will be released into the atmosphere. In addition, it is assumed that 75% of carbon stored in the forest-type vegetation and 100% of carbon stored in the grassland vegetation will be released into the atmosphere at the time of land conversion. (In essence, it is assumed that 25% of the carbon in wood is stored in buildings, furniture, etc.) If more than one type of vegetation is available in an area, the weighted average emissions for that area is calculated, where weights are shares of vegetation areas. The emission factors for forest areas and grasslands are calculated separately. Sensitivity analysis can be conducted on any of the data and assumptions used in this analysis.

Regarding forgone carbon sequestration, it is assumed that when a natural vegetation area is converted to cropland, it loses its carbon sequestration capacity as long as it is under crop production. Again, if more than one type of land is available, weighted average of forgone carbon sequestration is used. The direct and forgone sequestration in each region are simply added. Hence, in each area, there are two groups of emission factors: forest and grassland emission factors. The emission factors are calculated based on the assumption that the converted land to crop production will remain under crop production for 30 years (for more details on emissions factors, see Tyner *et al.* (2010)). We recognize that the 30-year period is somewhat arbitrary. Thirty years roughly equals the life of a biofuel facility, so it seems as reasonable an assumption as any other.

At this point it is important to note that some research indicates that conservation tillage practices and enhanced rotation programs can increase the carbon sequestration ability of croplands. This means that using advanced technologies in corn production can increase carbon stored in soil (West and Post, 2002). This research ignores impacts of advanced tillage methods on the carbon sequestration ability of cropland.

Estimated Land-Use CO_2 Emissions Due to US Ethanol Production

This article now combines simulated land-use changes due to US ethanol production with the CO_2 release emission factors. This is a straightforward process. Suppose ΔLF_{rj} (see Tables 2, 6, and 9) is the size of change in land type j (j=forest and grassland) in region r due to X gallons of increase in the US ethanol production. In addition, suppose that the annual CO_2 emission factor for land type j in region r for a 30-year ethanol production is about F_{rj}. Then the global annual CO_2 emissions due to the production of X gallons of ethanol per year USA will be equal to

$$LUE_W = \sum_r \sum_j \Delta LF_{rj} F_{rj} \quad [1]$$

Using this approach, CO_2 emissions were calculated for all land-use simulation scenarios (three groups of simulations and six time segments) and for all emission factors derived from the Woods Hole data sets. Once emissions have been calculated, the marginal and average land-use emissions due to production of each gallon of pure ethanol (E100) can be calculated for all groups of simulations examined in this article. For example, Table 11 shows how the marginal land-use emissions due to producing each gallon of E100 for the 13–15 BGs were calculated for the first group of simulations.

The value of 110.77 million metric tons of emissions presented in this table is obtained by multiplying regional forest and grassland changes due to an increase in ethanol production from 13 to 15 BGs by their corresponding Woods Hole annual emissions factors and then summed over regions. The result of this calculation is multiplied by 30 to present the magnitude of total emissions more than 30 years. The last row of this table shows the annual marginal land-use emissions of each additional gallon of ethanol when moving from 13 to 15 BGs.

Land Use Emissions for the First Group of Simulations

Table 12 represents marginal and average land-use emissions obtained from simulations off the 2001 database. This table indicates that marginal emissions are increasing in ethanol production. For example, although an increase in ethanol production from 7 to 9 BGs generates 1687 g CO_2 emissions per gallon of ethanol, moving from 9 to 11 BGs generates 1745 g CO_2 per gallon. When ethanol production reaches 15 BGs, each additional gallon of ethanol generates 1846 g of CO_2. **Table 12** indicates that average emissions are increasing in ethanol production as well. This table shows that during the time period 2001–2006, on average each gallon of US ethanol was generating 1477 g CO_2. However, if ethanol production reaches 15 BGs, on average each gallon of ethanol generates 1676 g of emissions. It is important to note that in this group of simulations, about 61% of emissions result from deforestation and 39% from converting grasslands into crop production.

Note that this article ignores impacts of the first 1.77 BGs of ethanol on the average land-use changes per gallon of ethanol production. Incorporating land-use changes due to the first 1.77 BGs of ethanol will moderately reduce the average emissions per gallon of ethanol.

Land-use emissions obtained from this group of simulations are smaller than the authors' earlier estimates of land-use emissions. For example, as shown in **Table 12**, on average each gallon of US ethanol generates 1676 g emissions. The corresponding number reported in Tyner *et al.* (2009) was about 2210 g emissions. This shows about 16.5% reduction in emissions per gallon of ethanol. This is due to the use of new regional ETAs and the incorporation of cropland pasture into the picture.

Land-Use Emissions for the Second Group of Simulations

Table 13 presents the marginal and average emissions for the second group of simulations, where land-use changes are calculated according to the updated baseline for 2001–2006. Emissions obtained from second group of simulations follow the pattern of the first group. However, their magnitudes are smaller than those of the first group.

As shown in **Table 13**, when the US ethanol production reaches 15 BGs, each additional gallon of ethanol generates about 1467 g of emissions. At this level of ethanol production, on average each gallon of ethanol generates 1426 g of CO_2 emissions. These figures are smaller than the corresponding figures of the first group of simulations by 21% and 15%. These reductions are due to yield improvement during the

Table 12 Annual marginal and average estimated land-use emissions due to US ethanol production: obtained from the simulations off the 2001 database

Time segment	*Marginal Emissions (grams CO_2 per gallon of ethanol)*				*Average emissions (grams CO_2 per gallon of ethanol)*			
	Changes in ethanol production	*Forest*	*Grasslands*	*Total*	*Total ethanol production*	*Forests*	*Grasslands*	*Total*
2001–2006	3.085	886	590	1477	3.085	886	590	1477
2006–2007	2.145	990	628	1619	5.23	929	606	1535
2007–2009	2.000	1033	654	1687	7.23	958	619	1577
2009–2011	2.000	1067	677	1745	9.23	982	632	1613
2011–2013	2.000	1097	701	1797	11.23	1002	644	1646
2013–2015	2.000	1122	724	1846	13.23	1020	656	1676

Table 13 Annual marginal and average estimated land-use emissions due to US ethanol production: obtained from the simulations off the updated 2006 database

Time segment	*Marginal emissions (grams CO_2 per gallon of ethanol)*				*Average emissions (grams CO_2 per gallon of ethanol)*			
	Changes in ethanol production	*Forest*	*Grasslands*	*Total*	*Total ethanol production*	*Forests*	*Grasslands*	*Total*
2001–2006	3.085	925	465	1390	3.085	925	465	1390
2006–2007	2.145	1019	399	1418	5.23	963	438	1402
2007–2009	2.000	1020	406	1427	7.23	979	429	1409
2009–2011	2.000	1017	409	1426	9.23	987	425	1412
2011–2013	2.000	1027	419	1446	11.23	994	424	1418
2013–2015	2.000	1040	427	1467	13.23	1001	424	1426

time period 2001–2006. Over this period, yield growth in some regions was faster than demand growth. It is important to note that in this group of simulations, about 70% of emissions come from deforestation and the remainder come from converting grasslands into crop production.

Land-Use Emissions for the Third Group of Simulations

Table 14 shows the marginal and average land-use emission for the third group of simulations, where land-use changes are calculated according to the simulations with the updated 2001–2006 database and population, yield, and forest product growth. As shown in Table 14, in this case during the time period 2006–2015, the marginal emissions grow when the population growth dominates the yield growth. For example, an additional gallon of ethanol produces about 1032 g emissions in the time period 2006–2007, whereas each gallon of additional ethanol produces 1159 g emissions in the time period 2013–2015. In this group of simulations, on average each gallon of ethanol generates about 1167 g of emissions. This figure is smaller than the corresponding figure obtained from the second group of simulations by about 18%.

Final Analysis

This article now compares the land-use emissions obtained from the three groups of simulations with the results of SEA. Table 15 shows lower emissions due to land-use change when all economic and demographic and yield growth are taken into account in the third group of simulations. The average value of the third group of simulations is about 13.6% of the original SEA result. The results of the first and second groups of simulations are approximately 20% and 16.6% of SEA.

Total Emissions from Production and Consumption of Ethanol

Table 18 contains the estimated well-to-wheel ethanol emissions for the marginal and average land-use changes for the three groups of simulations. (In this article, the direct marginal GHG emissions (i.e., nonland emissions) of ethanol for the post-2006 period are taken from 100% dry mill.) For the first group of simulations, the production and consumption of each gallon of ethanol (E100) on average generates about 6800 g of GHG emissions. In this case, about 24.6% of released emissions are related to land-use changes. When changes in population and other factors are incorporated, each gallon of ethanol (E100) on average generates about 6550 g of GHG emissions. In this case, about 21.7% of released emissions are related to land-use changes. Finally, in the third group, when the population and yield growth after 2006 are taken into account, production and consumption of each gallon of ethanol (E100) generates about 6291 g of emissions. In the third case, about 18.5% of released emissions are related to land-use change.

Table 16 indicates well-to-wheel ethanol emissions expressed as grams per gallon of ethanol and as grams per

Table 14 Annual marginal and average estimated land-use emissions due to US ethanol production: obtained from the simulations off the updated 2006 database and with population and yield growth after 2006

Time segment	*Marginal emissions (grams CO_2 per gallon of ethanol)*				*Average emissions (grams CO_2 per gallon of ethanol)*			
	Changes in ethanol production	*Forest*	*Grasslands*	*Total*	*Total ethanol production*	*Forests*	*Grasslands*	*Total*
2001–2006	3.085	925	465	1390	3.085	925	465	1390
2006–2007	2.145	716	317	1032	5.23	839	404	1244
2007–2009	2.000	721	351	1072	7.23	806	390	1196
2009–2011	2.000	698	391	1089	9.23	783	390	1173
2011–2013	2.000	697	452	1149	11.23	768	401	1169
2013–2015	2.000	659	501	1159	13.23	751	416	1167

Table 15 Estimated land-use change emissions due to US ethanol production (comparing GTAP and Searchinger, *et al.* (2008) results)

Searchinger, *et al.* (2008)	Total emissions for 30 years (million metric tons)	3,801
	Change in ethanol production (billion liters of ethanol)	55.92
	Total emissions for 30 years (grams per liter)	67,972
	Liters per gallon	3.785
	Total emissions for 30 years (grams per gallon of ethanol)	257,302
	1-year emissions (grams per gallon of ethanol)	8,577
GTAP results off the 2001 database	1-year average emissions (grams per gallon of ethanol)	1,676
	1-year marginal emissions (grams per gallon of ethanol)	1,846
GTAP results off the 2006 database	1-year average emissions (grams per gallon of ethanol)	1,426
	1-year marginal emissions (grams per gallon of ethanol)	1,467
GTAP results off the 2006 database plus population and yield growth	1-year average emissions (grams per gallon of ethanol)	1,167
	1-year marginal emissions (grams per gallon of ethanol)	1,159

Table 16 Estimated annual well-to-wheel ethanol emissions for marginal and average land-use changes

Description		*Land-use emissions (g gal^{-1})*	*Land-use emissions (g MJ^{-1})*	*Well-to-wheel emissions without land-use emissions (g gal^{-1})*[a]	*Well-to-wheel emissions plus land-use emissions (g gal^{-1})*	*Well-to-wheel emissions plus land-use emissions (g MJ^{-1})*[b]
Simulations off the 2001 database	Marginal	1846	22.9	5100	6946	86.3
	Average	1676	20.8	5124	6800	84.4
Simulations off the 2006 database	Marginal	1467	18.2	5100	6567	81.5
	Average	1426	17.7	5124	6550	81.3
Simulations off the 2006 database plus population and yield growth	Marginal	1159	14.4	5100	6259	77.7
	Average	1167	14.5	5124	6291	78.1

[a]From GREET simulations. The default values were used in GREET version 1.3c for 2015 for the simulations. The marginal and average differ for ethanol direct emissions because the fraction that is wet vs. dry milling decreases over time, yielding slightly lower direct emissions for the marginal case.
[b]Low heating values of gasoline and ethanol are: 116,090 and 76,330 BTU per gallon.

Table 17 Estimated well-to-wheel ethanol and gasoline emissions

Description		*Emissions in grams per gallon of gasoline equivalent*			*Emissions (g MJ^{-1})*		
		Ethanol	*Gasoline*	*Ethanol vs. gasoline (%)*	*Ethanol*	*Gasoline*	*Ethanol vs. gasoline (percentage)*
Simulations off the 2001 database	Marginal	10,564	11,428	92.4	86.3	93.3	92.2
	Average	10,342	11,428	90.5	84.4	93.3	90.5
Simulations off the 2006 database	Marginal	9,988	11,428	87.4	81.5	93.3	87.4
	Average	9,961	11,428	87.2	81.3	93.3	87.2
Simulations off the 2006 database	Marginal	9,520	11,428	83.3	77.7	93.3	83.3
	Average	9,568	11,428	83.7	78.1	93.3	83.7

megajoule (MJ). For the first, second, and third groups of simulations, production and consumption of each gallon of ethanol (E100) on average generates about 84.4, 81.3, and 78.1 g MJ^{-1} emissions, respectively. Land-use emission for the third group of simulation is 14.5 g MJ^{-1}.

Finally, **Table 17** compares total emissions of E100 obtained from the three groups of simulations with the emissions of conventional gasoline. This table indicates that ethanol production induces lower emissions compared to conventional gasoline for all groups of simulations. For example, the total GHG emission due to production and consumption of E100 (including land-use emissions) obtained from the first group of simulations is about 10,342 g per gallon of gasoline equivalent for the average land-use changes. This figure is about 90.5% of the emissions due to production and consumption of conventional gasoline. When the updated 2006 database is used, the total estimated GHG emissions due to production and consumption of E100 is approximately 9961 g per gallon of gasoline equivalent for the average land-use changes. This figure is 87.2% of the emissions due to production and consumption of conventional gasoline. Finally, when the updated database is used, and population and yield increase after 2006 is assumed, the total estimated emission for E100 is 9568 g per gallon of gasoline equivalent for the average land-use changes. In this case, the E100 emission estimate is about 83.7% of emissions associated with conventional gasoline. **Table 17** presents emissions of ethanol and gasoline in grams per gallon of gasoline equivalent and grams per megajoule.

Since the third group of simulations takes into account changes in population, crop yields, economic growth, and growth in primary inputs during the time period 2001–2006 and after that assumes that population and yield growth will continue, the emissions obtained from this group of simulations are lower than in other cases. However, the results are derived from authors' assumptions, in particular for the time period 2006–2015. Any change in these assumptions could alter the results. In other words, this article has assumed 1% global growth in yields for all crops and 2001–2006 population growth through 2015. Changes in these assumptions would alter the numerical results.

Conclusions

The overarching objective of this article has been to show how predictions of the global land-use changes induced by US corn ethanol programs critically depend on processes that have been overlooked in some studies. Although this is a controversial topic, the review and synthesis presented in this article suggest the following. First, with more than a third of the US corn crop today going to ethanol, it is simply not

Table 18 Summary of the different modeling results

Result	*Original January 2009 estimates*	*Model improvements with 2001 database*	*Baseline updated to 2006*	*Updated baseline and growth in demand and yield*
Land needed for ethanol (ha per 1000 gallons)	0.27	0.22	0.15	0.13
Distribution of land-use change between forest and pasture (%forest per %pasture)	23/77	25/75	33/67	33/67
Distribution of land-use change between USA and rest of the world (%US per %Others)	35/65	34/66	24/76	24/76
Average emissions of 15 BG program (grams CO_2 per gallon of ethanol)	1931	1676	1426	1167
% of Searchinger *et al.*	22.5	19.5	16.6	13.6
Emissions per gallon gasoline equivalent (grams CO_2 per gallon)	10,564	10,342	9961	9568
Emissions per MJ (grams CO_2 per MJ)	86.3	84.4	81.3	78.1
Total ethanol emissions as % of gasoline	92.4	90.5	87.2	83.7

credible to argue that there are no global land-use change implications of corn ethanol. The valid question to ask is to what extent ethanol production will induce land-use changes. Second, given the uncertainty of underlying data and parameter estimation and the alternative assumptions that various models might make, one cannot escape the conclusion that predictions of land-use change are uncertain and, of course, are all obtained based on economic predictions. The estimation range depends on what is being simulated, as will be seen in the following paragraphs.

Numerous improvements have been made in the models used for the analysis. These improvements have been spelled out in this article and in other articles. There are better data on land productivity and on cropland pasture and CRP lands, and these data and associated parameters are now in the model. The treatment of livestock and livestock feed sectors has been improved. Similarly, these changes are reflected in the current version of the model. Data on crop yields and many other variables amass for every region of the world, and much of those data are used in this analysis and model calibration. These data and model improvements have significantly improved the analysis and model results.

Table 18 provides a convenient summary of the evolution of some of the results over the different versions of the model and data. The third column replicates the summary results from the January 2009 draft article before all the model changes described were implemented. The January 2009 results are provided only for reference, so the comparisons will be based on the three simulations reported in this article. The fourth column shows all the model improvements and the 2001 database. The fifth column reflects the baseline updated to 2006 as described in this article. The last column depicts the updated baseline to 2006 and the assumed growth in demand and supply as described in this article.

Ethanol emissions as a fraction of gasoline emissions range between 83.7% and 90.5%. In a recent analysis including uncertainty in GHG estimation using an earlier version of GTAP-BIO, Hertel *et al.* (2010) concluded that the corn ethanol-induced emissions from land-use change range between 2 and 51 g MJ^{-1}. The authors' estimate for the last case is 14.5 g MJ^{-1}. This large range taken from another study using similar approaches clearly illustrates the uncertainty inherent in this analysis. It also concludes that zero is not within the error bounds. In other words, emissions induced by land-use change are not zero, but measuring them with high precision is not yet possible.

Limitations of this Research

As indicated in this article, analysis such as that undertaken here is very complex and is limited by data availability, validity of parameters, and other modeling constraints. Economic models, like other models, are abstractions from reality. They can never perfectly depict all the forces and drivers of changes in an economy. However, the basic model used for this analysis, GTAP, has withstood the test of time and peer review. Hundreds of peer-reviewed articles have been published using the GTAP database and analytical framework. In this project, many changes have been made to the model and database to improve their usefulness for evaluating the land-use change impacts of large-scale biofuels programs. Yet uncertainties remain. This article has described the evolution of the modeling and analysis and presents openly the evolution of the results. It is believed quite strongly that analysis of this type must be done with models and databases that are available to others. Replicability and innovation are critical factors for progress in science. They also are important for credibility in policy analysis. This model is available at the GTAP website for GTAP subscribers.

See also: Landscape Modeling. Life Cycle Analysis of Biofuels. Market Economy and Biodiversity. Rainforest Loss and Change

References

Arora S, Wu M, and Wand M (2008) *Updated of Distiller Grains Displacement Rations for Corn Ethanol Life-Cycle Analysis*. Center for Transportation Research, Energy System Division, Argonne National Laboratory.

Birur DK (2010) *Global Impacts of Biofuels on Agriculture, Trade, and Environment: A Computable General Equilibrium Analysis*. PhD Dissertation, Purdue University.

Birur DK, Hertel TW, and Tyner WE (2008) Impact of biofuel production on world agricultural markets: A computable general equilibrium analysis. *GTAP Working Paper No 53*. West Lafayette, IN: Center for Global Trade Analysis, Purdue University.

Burniaux J and Truong T (2002) GTAP-E: An energy-environmental version of the GTAP model. *GTAP Technical Paper No. 16*. West Lafayette, IN: Center for Global Trade Analysis, Purdue University.

CRS RL34294 (2007) Energy independence and security act of 2007: A summary of major provisions. USA: Congressional Research Service.

Dimaranan BV (ed.) (2006) Global Trade, Assistance, and Production The GTAP 6 Data Base. *Center for Global Trade Analysis*. West Lafayette, IN, USA: Purdue University.

Euskirchen ES, McGuire AD, Kicklighter DW, *et al*. (2006) Importance of recent shifts in soil thermal dynamics on growing season length, productivity, and carbon sequestration in terrestrial high-latitude ecosystem. *Global Change Biology* 12: 731–750.

Fabiosa JF (2009) Land-use credits to corn ethanol: Accounting for distillers dried grains with solubles as a feed substitute in swine rations. *Working Paper 09-WP 489*. Iowa, USA: Center for Agricultural and Rural Development, Iowa State University.

Fargione J, Hill J, Tilman D, Polasky S, and Hawthorne P (2008) Land clearing and the biofuel carbon debt. *Science* 319: 1235–1238.

Hertel T, Golub A, Jones A, O'Hare M, Plevin R, and Kammen D (2010) Effects of US maize ethanol on global land use and greenhouse gas emissions: Estimating market-mediated responses. *Bioscience* 60(3).

Hertel T, Tyner W, and Birur D (2010) The global impacts of biofuels mandates. *The Energy Journal* 31(1): 75–100.

Hertel TW (1997) *Global Trade Analysis, Modeling and Applications*. Cambridge: Cambridge University Press.

Keeney R and Hertel T (2005) GTAP-AGR: A Framework for Assessing the Implications of Multilateral Changes in Agricultural Policies. *GTAP Technical Paper No. 24*. West Lafayette, IN, USA: Center for Global Trade Analysis.

Keeney R and Hertel TW (2008) The indirect land use impacts of U.S. biofuel policies: The importance of acreage, yield, and bilateral trade responses. *GTAP Working Paper No. 52*. West Lafayette, IN: Center for Global Trade Analysis, Purdue University.

Lee H-L Hertel TW, Sohngen B, and Ramankutty N (2005) Towards an integrated land use data base for assessing the potential for greenhouse gas mitigation. *GTAP Technical Paper # 25*. Purdue University: Center for Global Trade Analysis.

McDougall R and Golub A (2007) GTAP-E: A Revised Energy-Environmental Version of the GTAP Model. *GTAP Research Memorandum No. 15*. West Lafayette, IN, USA: Center for Global Trade Analysis.

Searchinger T, Heimlich R, Houghton R, *et al*. (2008) Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land use change. *Science* 319: 1238–1240.

Taheripour F, Birur D, Hertel T, and Tyner W (2007) Introducing liquid biofuel into the GTAP database. *GTAP Research Memorandum No. 11*. West Lafayette, IN, USA: Center for Global Trade Analysis, Purdue University.

Taheripour F, Hertel T, and Tyner W (2011) Implications of the biofuels mandates for the global livestock industry: A computable general equilibrium analysis. *Agricultural Economics* 42: 325–342.

Taheripour F, Hertel TW, Tyner WE, Beckman JF, and Birur DK (2010) Biofuels and their by-products: Global economic and environmental implications. *Biomass and Bioenergy* 34: 278–289.

Tyner, WE, Taheripour F, and Baldos U (2009) Land use change carbon emissions due to U.S. ethanol production. *Draft Report to Argonne National Lab*, January 2009. West Lafayette, IN, USA: Department of Agricultural Economics, Purdue University. http://www.agecon.purdue.edu/papers/biofuels/Argonne-GTAP_Revision%204a.pdf

Tyner WE, Taheripour F, Zhuang Q, Birur D, and Baldos U (2010) Land Use Changes and Consequent CO2 Emissions due to US Corn Ethanol Production: A Comprehensive Analysis. *A Report to Argonne National Laboratory*. West Lafayette, IN, USA: Department of Agricultural Economics, Purdue University.

West T and Post W (2002) Soil organic carbon sequestration rates by tillage and crop rotation: A global data analysis. *Soil Science Society American Journal* 66: 1930–1946.

Wang M (1999) *GREET 1.5 – Transportation Fuel-Cycle Model Vol. 1: Methodology, Development, Use, and Results*. Argonne, IL: Center for Transportation Research, Energy Systems Division, Argonne National Laboratory.

Wang M (2005) Updated energy and greenhouse gas emission results of fuel ethanol. *Presented on the 15th International Symposium on Alcohol Fuels*, 26–28 September 2005. San Diego, CA, USA: Center for Transportation Research, Argonne National Laboratory.

Zhuang Q, McGuire AD, Melillo JM, *et al*. (2003) Carbon cycling in extratropical terrestrial ecosystems of the Northern Hemisphere during the 20th century: A modeling analysis of the influences of soil thermal dynamics. *Tellus* 55(B): 751–776.

Land-Use Issues

Kees Klein Goldewijk, The Netherlands Utrecht University, Utrecht, The Netherlands

Published by Elsevier Inc.

This article is a revision of the previous edition article by John M. Marzluff, Nathalie Hamel, volume 3, pp 659–673, © 2001, Elsevier Inc.

Glossary

Afforestation Process of planting trees in areas currently devoid of them. It is done in desert regions and is increasingly common in Europe, where grasslands and pastures are being converted into woodlands and forests.

Agricultural intensification Management of agricultural land to increase yield principally through the use of high-yielding crop varieties, chemical fertilizers and pesticides, irrigation, shorter rotations, larger fields, and mechanization.

Desertification Reduction of the biological potential of land that ultimately leads to the creation of desertlike conditions. It occurs primarily through overuse of marginal lands by humans for agriculture.

Grassland Land that has a vegetation cover dominated by grasses.

Habitat fragmentation Splintering of once contiguous land cover into isolated pieces. Fragmentation happens when habitat is lost from the interior, rather than the edge, of a large block of cover. The resulting habitat patches are sometimes referred to as "habitat islands" and the intervening, converted land is called the "matrix."

Land cover Physical and biotic character of the earth's terrestrial surface. Land cover is typically the vegetation (e.g., tropical forest, marsh, desert, cornfield, and shrubland) or human construct (e.g., road, dwelling, and industrial area) that covers the surface. It may be grossly defined as in the preceding examples, or defined at finer scales (e.g., moist deciduous tropical forest, moist evergreen tropical forest, and dry tropical forest).

Land use The use that humans make of the earth's terrestrial surface, usually to obtain goods or benefits. Examples include agriculture, mineral exploration, settlement, natural reserve, and timber production. Land use refers to the dominant human activity that occurs on the earth's surface, and it often modifies natural characters to change land cover (e.g., replacing native grassland with a cornfield or a road).

Pasture Land used permanently (5 years or more) for herbaceous forage crops, either cultivated or growing wild (wild prairie or grazing land). The dividing line between this category and the category "Forests and woodland" is rather indefinite, especially in the case of shrubs, savannah, etc., which may have been reported under either of these two categories.

Settlement Land used for human habitation, including for the construction of cities, towns, villages, rural settlements, and roads. In a gross sense, settlement is also a land cover when used to refer to the land occupied by the variety of trappings accompanying human living space, including homes, gardens, roads, transportation centers, and industrial areas.

Land use issues concern the processes by which human activities determine land cover. Important issues are agricultural development and intensification, settlement, and extraction of natural resources. In response to human land use, the earth's land cover has changed from a mosaic of native woodlands, forests, and grasslands to an increasingly impacted mixture of degraded and fragmented native habitats, exotic croplands, and impervious urban surfaces. In the last three centuries, studies suggest that primary forests have declined 20%, natural grasslands and savannas have declined nearly 40%, whereas croplands have increased more than 390%, and pastures more than 460% (Klein Goldewijk, 2001; Klein Goldewijk *et al.*, 2011). Almost 39% of earth's land surface is transformed into agricultural land and settlements, and an additional 37% of global land without such use has become embedded within agricultural and settled anthromes (Ellis *et al.*, 2010). This article discusses how land use processes have changed through time and how they have caused the natural pattern of land cover to change. This transformation of the planet's landscape is widely recognized as the primary driver in the current global loss of biodiversity. Several examples of how land use can influence biodiversity are also considered.

Land Use Issues

Humans obtain food, energy, and shelter from the earth by harvesting natural resources and creating new resources. Both of these activities, especially the creation of new resources by agricultural and industrial actions, modify our environment at a scale unknown for other organisms. Our unparalleled recent success as a species (increasing in total global population size by almost three-fold from 1950 to 2010) and our tremendous ability to modify the environment have allowed us to dramatically transform the earth (Turner *et al.*, 1990). It is estimated that we have transformed or degraded large parts of the earth's surface, with 20% remaining seminatural and only a quarter left wild (Ellis *et al.*, 2010) and now command 24–32% of the world's total net primary productivity (Haberl *et al.*, 2007).

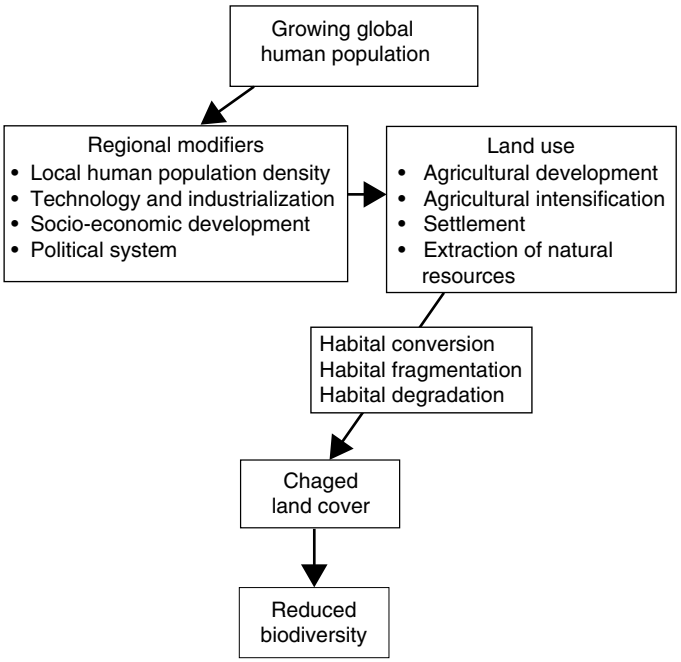


Figure 1 The growing global population of humans reduces biodiversity primarily through land use that converts, fragments, and degrades land cover. Human population growth directly affects land use by demanding more food, living space, and natural resources. The specific effects on regional land use are modified by the local human density, local economy, local technology and industrialization, and local political system.

We can conceptualize our effects on the earth's surface by relating our land and resource use to changes in land cover (**Figure 1**). As our population increases, we garner more of the earth to supply our needs. The aspects of the human population that drive changes in land use and resulting land cover are known only in a general way. Population size is obviously an important determinant of our land use. But its importance relative to global affluence, technological capacity, political organization, and social structuring is hotly contested (Turner II *et al.*, 1990; Steffen *et al.*, 2007). Recent insights suggest that population size may directly determine land use at a global scale, but land use locally is determined by the interaction of regional population size, technology, socioeconomic development, industrialization, and policy (see **Figure 1**). Total biological demand for resources and resulting conversion of land cover are determined by the total world population. Where on the planet changes in land use occur is affected by regional population size, technology, policies, conventions, cultural attitudes, and economic forces that discourage or reward land conversion at the local scale (Turner II *et al.*, 1990; Steffen *et al.*, 2007).

Three main land uses are responsible for changes in the earth's land cover. These are (1) agriculture, which includes cultivation, grazing, irrigation, and drainage; (2) extraction of natural resources, primarily for forestry and energy production; and (3) settlement, which we define loosely to include use of land for human settlements ranging from rural villages to cities. Our domination of the earth is clearly reflected by the earth's current land cover, which is primarily a mixture of fragmented and degraded native habitats, exotic croplands and pastures, and settlements (**Figure 2**). This transformation of land cover *per se* is recognized as the driving force in the loss of biological diversity worldwide. Additionally, land use affects the biosphere indirectly by modifying ecosystem processes and introducing exotic species (Vitousek *et al.*, 1997; Alkemade *et al.*, 2009).

Agriculture

Our agricultural practices have had the greatest influence on the earth's land cover. Before the development of agricultural society, several lines of evidence suggest that the earth was covered with a complex mosaic of forest, woodland, grassland, wetland, tundra, and desert (**Figure 3**). During the last 300 years we have cultivated approximately 11% of the earth's surface (**Figure 4**), and another 24% is currently in use as pasture. This represents an approximate increase of 390–466% since 1700 (Klein Goldewijk *et al.*, 2011). Increases in cultivated land came mostly at the expense of native forests (other than tropical rain forest) and native grasslands (**Figure 5**). We transform the earth by burning and clearing forests, tilling grasslands, draining wetlands, and irrigating arid lands. Agricultural development is responsible for the overwhelming majority of wetland conversion (the global loss of wetlands is not accurately estimated; Meyer and Turner, 1992). The pattern of increasing human population, increasing cropland, and declining forests and grasslands is consistent across the continents except in the USA, Europe, and parts of Asia (Japan, China) where current afforestation is increasing forest and woodland cover (**Figure 6**). This reversal in declining forest area is marked by a first gradual or sudden depletion of the forest resources, later by slow or radical changes in economic structures or organization. These

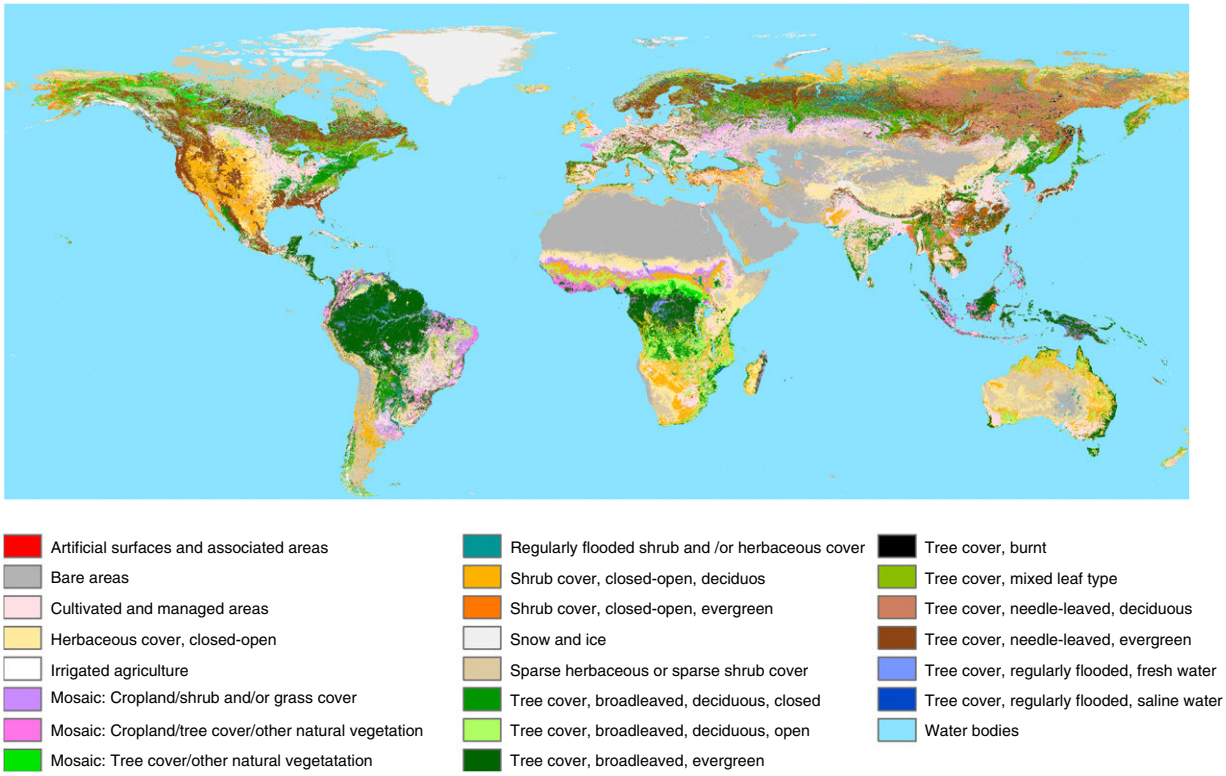


Figure 2 Current land cover of the world. Human domination is evident by the preponderance of agricultural land cover (cultivated, irrigated, and mosaic lands). GLOBCOVER (Reproduced from Arino O, Ranera F, Bicheron P, *et al.* (2008) *Globcover*. Frascati: European Space Agency, (Special Publication) ESA SP).

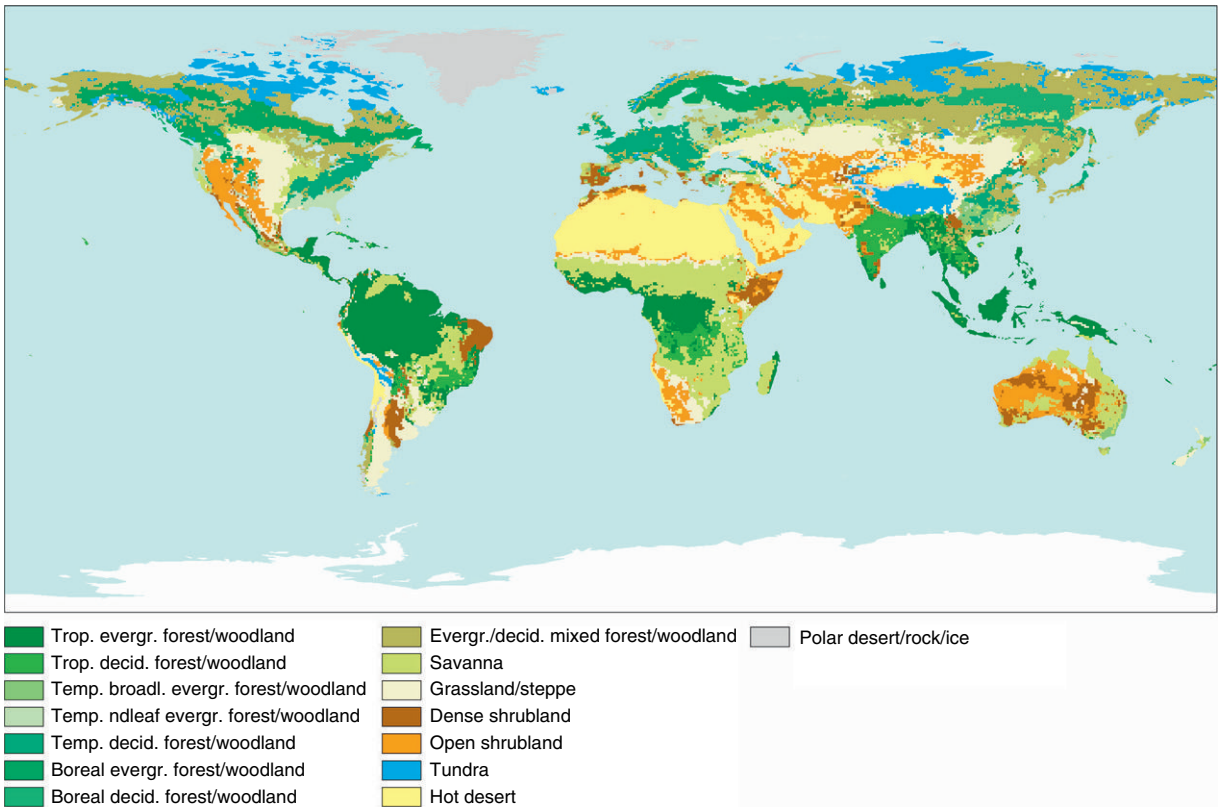


Figure 3 Estimated (modeled) land cover before the development of agriculture. Reproduced from Ramankutty N and Foley JA (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13: 997–1027.

lead to a new appreciation of the resources and eventually in an increase in forest area again. This process is also described as the Forest Transition (Mather, 1990).

Although conversion of native land to exotic croplands has slowed, our use of agricultural land has intensified since "the Green Revolution" began in the 1960s (Matson *et al.*, 1997). Agricultural intensification alters land cover by (1) creating highly productive monocultures, (2) diverting water resources, and (3) fueling technology to produce better seeds and more fertilizers, pesticides, and insecticides.

Increasingly intensive use of grasslands for agriculture has resulted in worldwide overgrazing (**Figure 7**). A direct effect of this is "desertification" or the degradation of land to desertlike conditions. Current estimates indicate that 6% of the earth's land has been converted to deserts by agricultural activity; nearly 25% of the earth is threatened with desertification (Meyer and Turner, 1992).

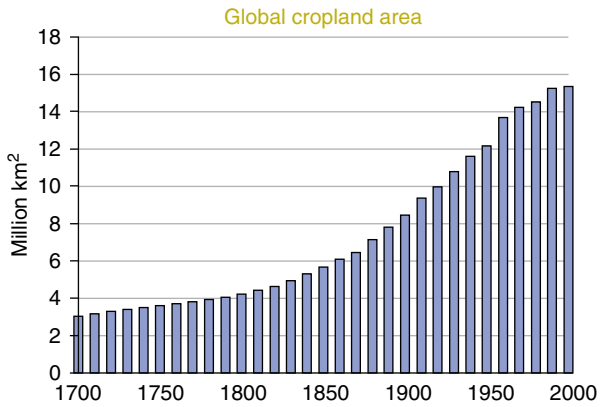


Figure 4 Amount of the world under cultivation (cropland) from 1700 to 2000.

Natural Resource Extraction

We extract a variety of resources from the earth for shelter and energy, but at a global scale timber harvest and management have the largest effect among extractive activities on landscape composition and land cover. In 2010, 31% of the earth was forested (4 billion hectares) and 24% (949 million hectares) of forests are currently maintained as plantations or managed for goods and services (FRA2010, 2010). The rate of deforestation shows signs of decreasing, but remains still alarmingly high according to FAO (globally 13 million hectares per year during 2000–2010 period, compared to 16 million hectares per year for the 1990–2000 period). Large-scale planting of trees is significantly reducing the net loss of forest area globally. The net change in forest area in the period 2000–2010 is estimated at −5.2 million hectares per year (an area about the size of Costa Rica), down from −8.3 million hectares per year in the period 1990–2000. During the last decades, Europe, USA, and especially China have planted large areas with forest, the FRA2010 in 2010 reports 25, 26, and 77 million hectares, respectively.

Humans affect all forests by hunting, gathering, recreating, and suppressing fire. However, only fire suppression has the potential to change forest cover type (if successful, it initially slows forest growth by reducing natural thinning but may allow for greater development of late-successional forests by reducing disturbance). Natural resource extraction has a less direct effect on forest cover than does agriculture. Resource extraction modifies the forest by changing seral stage, degrading forest integrity, and fragmenting continuous forest, but is associated with reduced forest cover in only a few locations (southeastern Asia and parts of Africa, Latin America, and India). Agriculture, in contrast, is the primary agent of forest loss, although it also modifies forest function. Some of the forest loss to agriculture is compensated for by maintenance and planting of forest for timber.

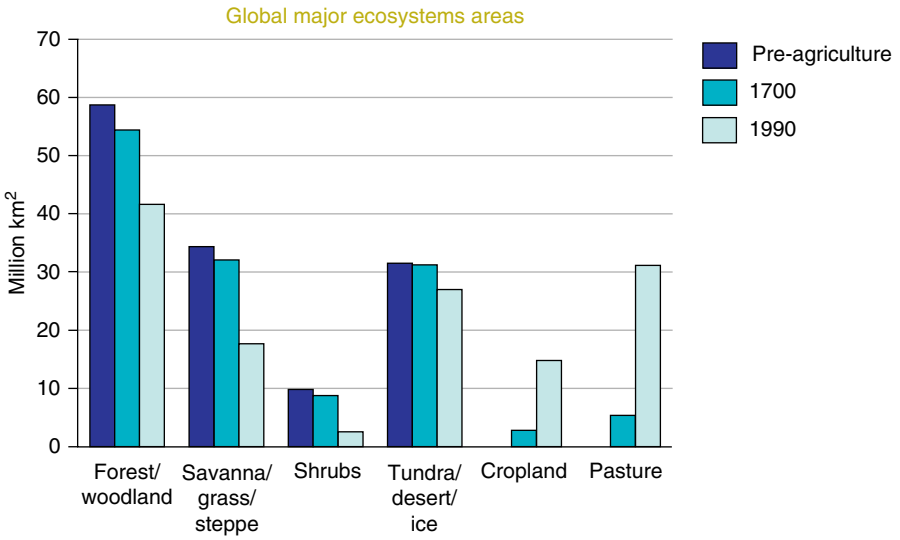


Figure 5 Areas occupied by major ecosystems of the world before agriculture (Klein Goldewijk, 2001). Cultivation has mostly affected forests as well as natural grassland/savanna ecosystems. Reproduced from Klein Goldewijk K, Beusen A, van Drecht G, and de Vos M (2011) The HYDE 3.1 spatially explicit database of human induced land use change over the past 12,000 years. *Global Ecology and Biogeography* 20(1): 73 - 86.doi:10.1111/j.1466-8238.2010.00587.x., with permission from Wiley.

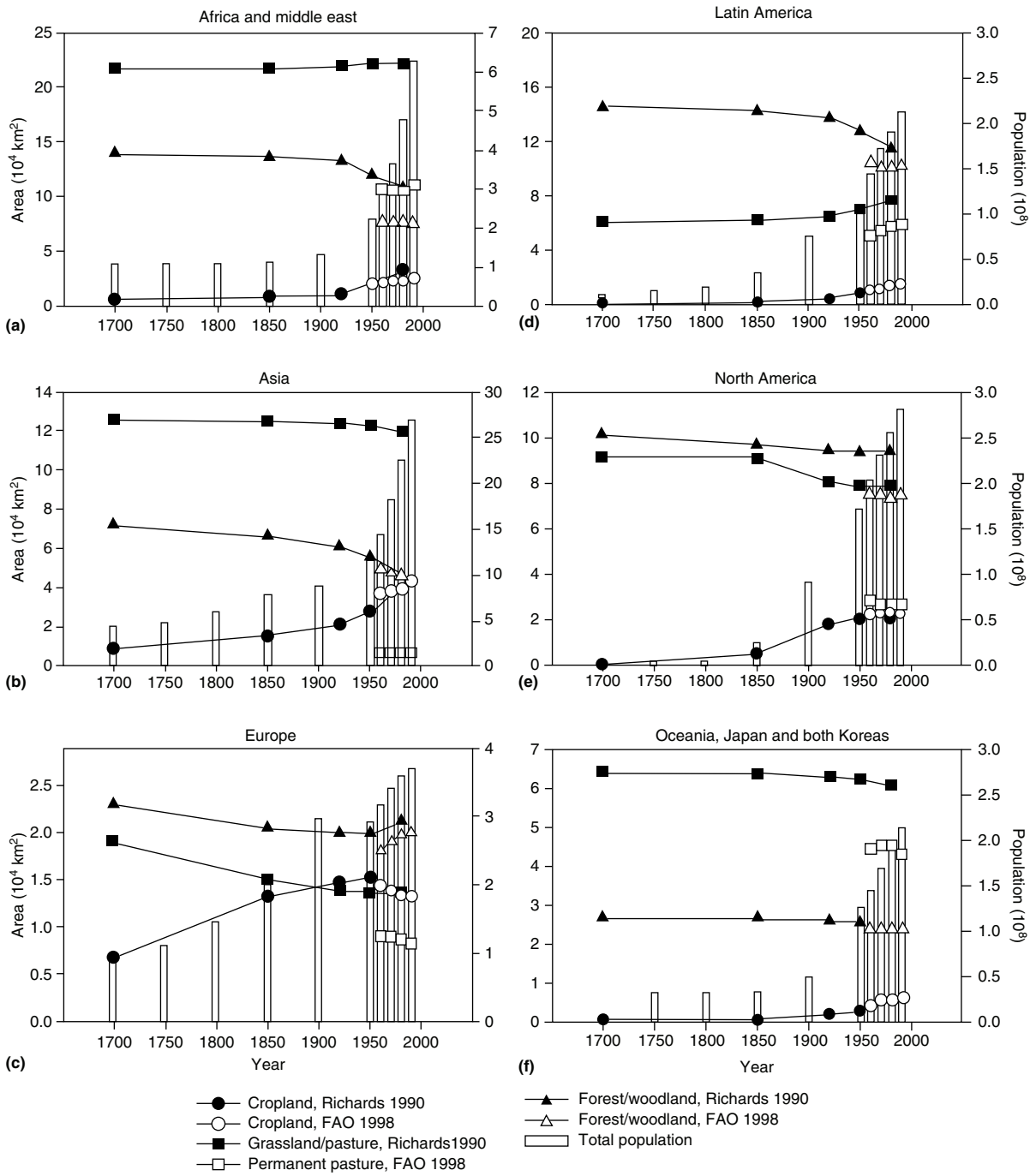


Figure 6 Land use trends and population estimates of major regions of the earth. Land use trends between 1700 and 1980 (closed symbols) are compiled from Richards (1990). After 1950 (open symbols), land use estimates are derived from FAOSTAT (www.fao.org, October 1998). Population estimates in (a) are for Africa only. Panel (f) includes Japan, North and South Korea, and Oceania.

Settlement

The clearing of native land cover and drainage of wetlands for human settlement are the most permanent forms of land conversion. Global population has increased enormously during the last three centuries, from roughly 600 million in AD 1700 to 1.7 billion in AD 1900, 2.5 billion in AD 1950 to 7 billion at present (Klein Goldewijk *et al.*, 2010). This also had of course consequences for the amount and size of settlements as well, since more and more people live in towns and cities. Currently, the urbanization degree globally is more than 50% (UN, 2009). Despite the large increase in population, the absolute area compared to the total global land surface is still very small. Potere and Schneider (2007) compared six global studies of urban areas and found a range between 276,000–727,000 km^2, and one outlier of

3524,000 km^2. That translates to resp. 0.19–0.49% (and outlier 2.4%) of the total land area.

Human settlements typically occupy less than 20% of even the most developed country's area, but this varies widely (**Figure 8**). Each night we advertise our settlements by turning on lights. Viewing this spectacle from the air provides a striking representation of the degree to which settlements (and some large industrial complexes such as oil and gas drilling sites in Siberia for example) occupy the earth (**Figure 9**). Human settlement affects nearly all land in Europe, eastern North America, India, and parts of Asia.

Settlement has increased dramatically over the last 300 years as evidenced by the increasing number of large cities (**Figure 10**). As human populations have grown, we have settled with increasing frequency in urban areas. Today more than 50% of all humans live in urban areas (primarily in the developed countries), compared to only 14% at the start of the twentieth century (UN, 2009). A result of our increasingly urban life is an increased occupation of area by cities and surrounding lower-density housing and industrial areas (urban and suburban sprawl). This expansion of urban areas in developed countries and the increasing density of large urban areas in the developing tropics are responsible for the severe transformation of parts of the earth's surface.

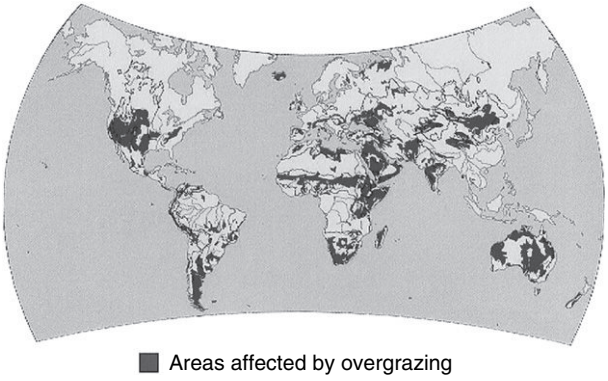


Figure 7 Areas of the world affected by overgrazing. This human-induced activity has been identified as one of the major causes for soil degradation. Overgrazed areas are the result of actual removal of biomass by grazing animals and the effect of trampling. The effects are enhanced where too many animals or the wrong kind of animals are grazed. (Reproduced with permission from UNEP (United Nations Environment Programme)/ISRIC (1997) *World Atlas of Desertification*, p. 44. New York: Arnold/John Wiley & Sons.)

History Repeats Itself: Past and Future Effects of Land Use on Land Cover

Our transformation of the earth during the last five centuries has proceeded by pioneers expanding frontier settlements in a fairly regular way. When humans settle an area, they open a new frontier where natural resources are abundant, cheap, and readily obtained. As resources are consumed and population increases, we expand our "footprint" on the surrounding land in a series of radiating waves. We transform the land at successively greater distances from a settlement's center by clearing and draining land, extracting resources, and converting land cover to suit our needs (Berry, 1991).

The conversion of land cover was fueled on a global scale by Western European nations and their seaborne pioneers from the 1700s to the early 1900s (Richards, 1990; Klein Goldewijk *et al.* 2011; Williams, 2003). Once the European frontiers were settled, the appetite for amenity and subsistence resources propelled the opening of new frontiers around the world. These frontiers then had to satisfy local and European needs, so resources were harvested with frenzied zeal as opportunities seemed boundless on the frontier. Landscapes were quickly degraded and transformed, and as resources dwindled, the frontier closed and pioneers moved on to new frontiers. However, the old frontier was rarely left to return to a natural state. Rather, a more orderly settlement usually persisted, leading to slower, but more permanent, land transformation. Many of these settlements became new centers in a changing world

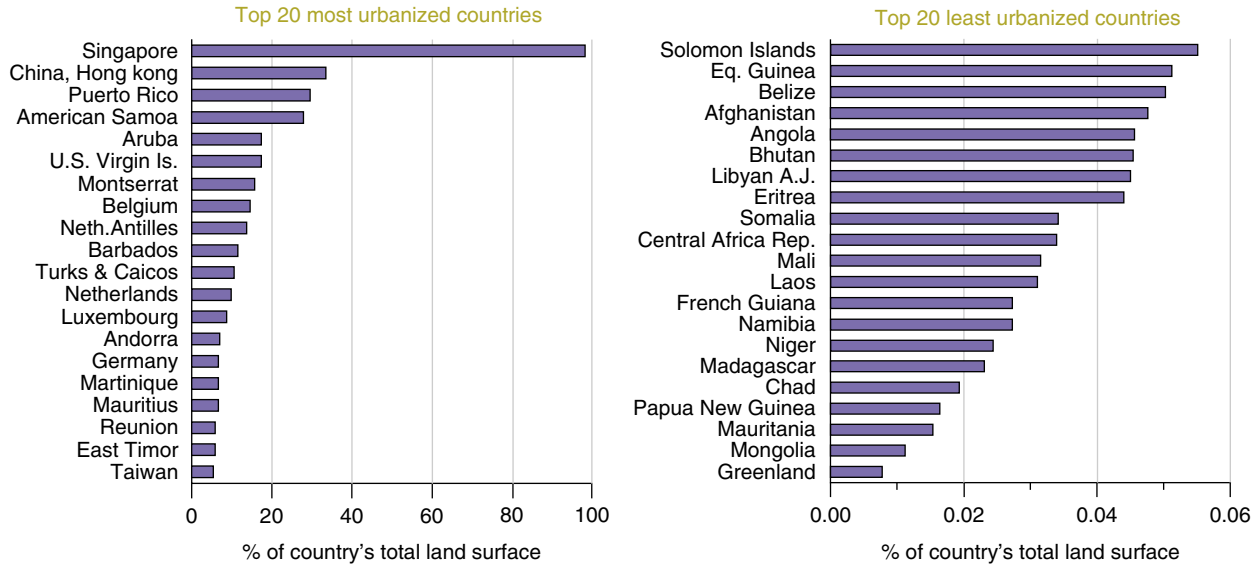


Figure 8 Urban areas as expressed as a percentage of land area (excluding water bodies), in selected countries.



Figure 9 The world at night. Reproduced from NGDC/ NOAA NightLights (2009). The file contains the average of the visible band digital number values of DMSP (Defense Meteorological Satellites Program) F16 satellite with no further filtering. Data values range from 0 to 63. Areas with zero cloud-free observations are represented by the value 255.

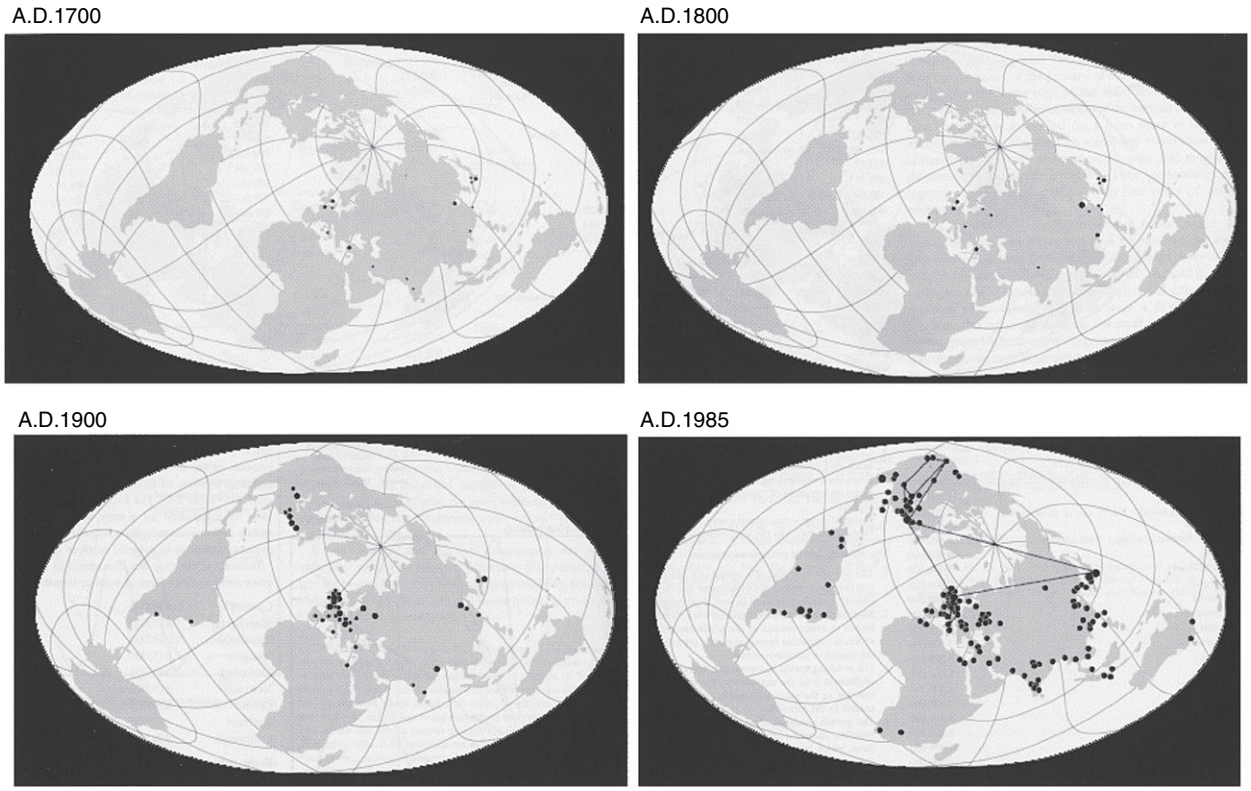


Figure 10 Distribution of the sizes of world cities, between 1700 and 1985. The growth rate of cities in the Americas is more rapid relative to that of cities on the European and East Asian continents. The black dots represent cities of varying sizes (•, 200,000–500,000; ●, 500,000–1,000,000; ●, 1,000,000–10,000,000).

economy, thereby broadening the base from which remaining frontiers were exploited.

The global pattern of urbanization over the last 300 years (see **Figure 10**) illustrates the consequences of our opening and closing the earth's frontiers. New frontiers open in areas increasingly far from Europe, and develop large urban areas as they close. The frontiers of Europe, eastern North America, India, and Asia have closed, leaving dense human settlement

to permanently change land cover. As the remote frontiers of the tropics, mountainous regions, and northern forests and tundras are exploited, settlement will likely increase and land will be transformed to a lasting urban state. Europe can be viewed as a crystal ball for much of the earth's future land use. The forecast is for drastic simplification of a once diverse native land cover to one dominated by settlements, intensive agriculture, and fragmented vestiges of native habitat.

Effects of Land Use on Biodiversity

The current extinction crisis is fueled, at a proximate level, by (1) habitat conversion, (2) habitat fragmentation, and (3) habitat degradation. These are the most common reasons given to explain why modern species are at risk of extinction (IUCN (International Union for the Conservation of Nature), 1996). These three processes are the direct result of our land use and they function as the mechanisms that connect changing land cover to biodiversity (see **Figure 1**). Conversion of native land cover to agriculture and settlement is unquestionably the main cause of habitat loss on a global scale. This conversion typically occurs in a piecemeal fashion and is associated with extensive road building that fragments habitat as it is being lost. Native and converted habitats that are already utilized by humans are further degraded by agricultural intensification, resource extraction, and recreation.

A recent operational tool for policy support on the global to the national scale is the Global Biodiversity model GLOBIO (Alkemade *et al.*, 2009). It can be used to assess (1) impacts of human-induced environmental drivers on land biodiversity in past, present and future, (2) the relative importance of the environmental drivers, (3) trends under future scenarios, and (4) effects of policy response options, such as climate change mitigation, plantation forestry, and protected areas. GLOBIO has been applied in many integrated, global and regional environmental assessments (http://www.globio.info/home). **Figure 11** is an example of an application of GLOBIO from the OECD Environmental Outlook 2030; it depicts the global trend in biodiversity over time.

The effects of land use on biodiversity can be illustrated by considering habitat conversion, fragmentation, and degradation in several regional studies, and we have selected well-documented studies to illustrate the general process. It is important to consider regional changes in land cover because the factors that drive land use and link it to land cover are best understood at 'subglobal' levels (Kummer and Turner, 1994). The following sections examine (1) the complex ways that biophysical and socioeconomic factors affect land use, (2) the linkage between land use and habitat conversion, fragmentation, and degradation, and (3) the effects of land use on biodiversity.

Habitat Conversion

Deforestation in Southeast Asia

The tropical rain forests of Southeast Asia have declined dramatically since the end of World War II (in the late 1940s). Before humans arrived, the Philippines, Indonesia, Malaysia, Vietnam, and Thailand were completely forested (see **Figure 3**). Their rich dipterocarp forests were heavily logged (primarily for export) after the war ended. The Philippines currently retains less than 26% of its original forest cover; Malaysia retains more forest, but it is being extensively harvested (**Figure 12**). The rates of deforestation in the Philippines are well documented and staggering (Collins *et al.*, 1991; Kummer, 1992). From 1948 to 1987, 55% of the forest was lost. From 1969 to 1988, the rate of deforestation was more than 2% per year – or 2 ha every 5 min! The latest FRA2010 reports for the year 2010 – 22% of the total land area in East Asia as being forest, 35% for South and Southeast Asia, 35%

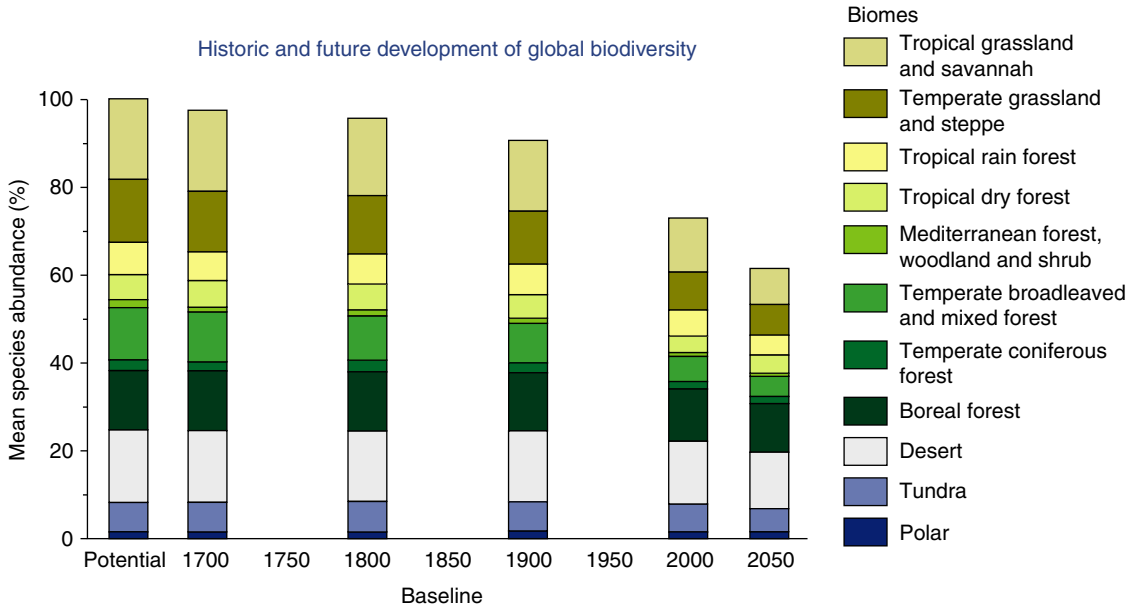


Figure 11 Historical trends in global biodiversity. Reproduced from OECD (2008).

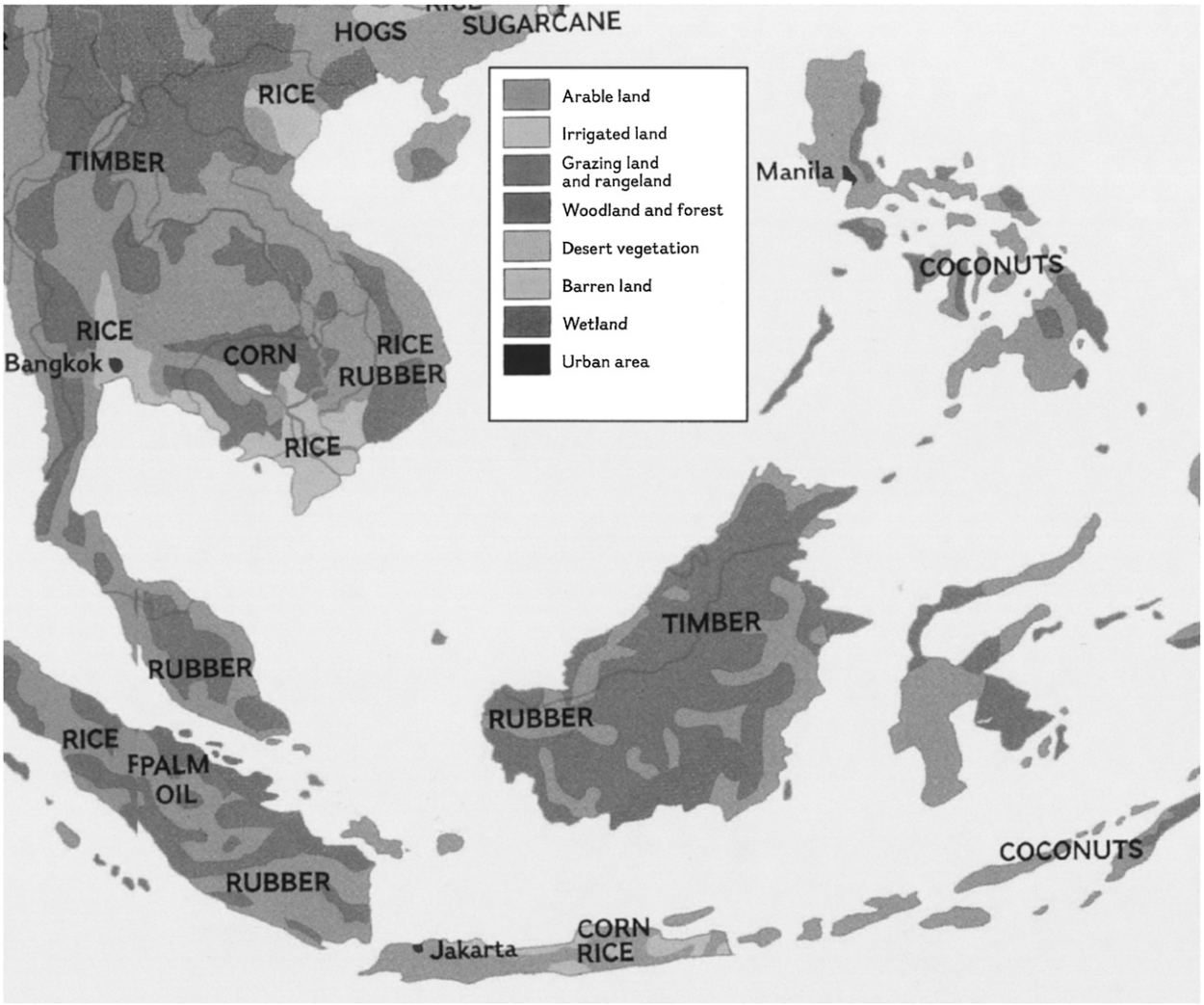


Figure 12 Land cover in the 1990s in Southeast Asia. The tropical rain forests that once covered the area have been now reduced to patches in an agricultural landscape. (Reproduced from National Geographic Society (1992) *The National Geographic Atlas of the World.* Washington, DC: National Geographic Society.)

for Europe, 33% for North and Central America and 49% for South America, and 23% for Africa.

The process of habitat loss (deforestation) in the Philippines serves as a model for all of Southeast Asia (Kummer and Turner, 1994). Two linked land uses, logging and subsistence agriculture, are responsible for converting land cover from tropical rain forest to arable land. Logging dipterocarps converted the primary rain forest to a simpler secondary forest. The resulting partial clearing and road construction allowed poor, landless people to remove the secondary forest and expand agriculture.

This two-step process illustrates how human density, global markets, and local socioeconomic conditions work together to cause changes in land cover. Despite the Philippines' extremely dense (303 people per km^2 in 2010) and still growing population (2.5% in 1993 and 1.8% in 2010), deforestation is not directly driven by the local population. Rather, a global demand for wood, coupled with corrupt and unrestricted logging, cleared land and made it accessible to people in need of agricultural space. A surplus of landless people was needed to drive the forest conversion to agricultural land, but without logging to open a new frontier, urbanization might have increased and agriculture might have remained more centralized and intense rather than dispersed as it is now.

Southeast Asia in general and the Philippines in particular, has tremendous biodiversity. Diversity in the Philippines is poorly known, but there are $>12{,}000$ plants and fungi and >950 terrestrial vertebrates. This incredible diversity is primarily found in a large number (>59) of scattered and poorly protected national parks. Less than 1.3% of the Philippines is managed for conservation, and much of this area is still logged and settled by landless agriculturalists.

The combination of high biodiversity, high endemism, and extreme habitat conversion results in extreme risks of extinction. The magnitude of threat to birds and mammals in Southeast Asia is particularly well known (IUCN (International Union for the Conservation of Nature), 1996). Worldwide, Indonesia has the largest number of birds (104)

and mammals (128) threatened with extinction. The Philippines leads the world in the percentage of its avifauna that is threatened with extinction (15%). These countries are second only to Madagascar in the percentage of mammals threatened with extinction (32%). Vietnam, Thailand, and Malaysia are all in the world's top 20 countries for species endangerment.

To summarize, deforestation in Southeast Asia is an example of how our land use for amenity (dipterocarp logging) and subsistence (agriculture) resources changes land cover and affects biodiversity. Habitat conversion threatens biodiversity, but the greed (exotic wood) and need (agriculture) of a growing global human population fuel the conversion.

Urbanization in Western Washington State

An expanding population and the world's appetite for resources have transformed much of North America from extensive forest and grassland to agricultural and urban land (compare **Figures 2** and **3**). The general story is similar to the case in Southeast Asia. However, in the Pacific Northwest a more dramatic and permanent conversion of land is occurring at a rapid rate. Logging during the last 50 years has converted more than two-thirds of the primary forest in Washington to secondary forest. Rather than continuing to manage these forests for timber production, much of the land is now being converted to human settlements. In western Washington, for example, human populations have doubled in the last 50 years and are expected to double again in the next 50 years. From 1998 to 2000, Washington is expecting a net gain of one person every 5 min! This places a premium value on land for settlement. The predictable result is a rapid conversion of forest to urban and suburban settlements. Indeed, from 1970 to 1997, Washington lost 2.3 million acres of managed forestland. Urban expansion was responsible for about half of this loss. Rights-of-way and agriculture accounted for the rest.

The increasing urbanization in western Washington threatens to reduce biodiversity. Water flows altered by settlement have reduced the spawning and rearing habitat for the region's spectacular salmon diversity. As a result, several salmon runs have been extinguished and many are now listed as endangered. Loss of forests and intensification of resource extraction and recreation on remaining forests have contributed to the endangerment of several birds (e.g., Spotted Owls, *Strix occidentalis*, and Marbled Murrelets, *Brachyramphus marmoratus*), mammals (e.g., Grizzly Bear, *Ursus arctos*, and Gray Wolf, *Canis lupus*), and amphibians (Larch Mountain Salamander, *Plethodon larselli*). Drainage of wetlands and settlement of native woodlands and grasslands have endangered amphibians (Oregon Spotted Frog, *Rana luteiventris*), insects (Oregon Silverspot butterfly, *Speyeria zerene*), mammals (Western Gray Squirrel, *Sciurus griseus*), and reptiles (Western Pond Turtle, *Clemmys marmorata*).

Some species benefit whenever land cover changes. A good example of such a species is the American Crow (*Corvus brachyrhynchos*) in western Washington. Crows are found only in close association with human settlement. As a result, their numbers have increased 10-fold from 1960 to 1995. Increases in human commensals like crows may have a ripple effect through the native biota. Crows prey on the eggs and young of other birds, which may limit their reproduction and reduce overall biodiversity close to human settlements even in remaining forest reserves.

Habitat Fragmentation

The loss of native habitat as lands are converted for human use splinters the remaining habitat into small fragments afloat in a human-dominated matrix. Such "habitat fragmentation" is most evident on a global scale when we compare the original and current distributions of forests (**Figure 13**). Contiguous primary forests have been removed to varying degrees, resulting in a patchwork of forest islands in a sea of settlement, agriculture, and second growth forests.

Forest fragmentation affects biodiversity by (1) outright loss of habitat (recall Southeast Asia), (2) creation of forest edges that differ from interior forest in many physical and biological properties (e.g., wind speed, humidity, temperature, and predator populations), and (3) disruption of movement and dispersal patterns of forest species. The effects of habitat fragmentation depend on the type of land cover that surrounds the fragments (the "matrix"). In general, when the matrix is similar to the native habitat (e.g., secondary forest surrounding native forest), the effects of fragmentation are less than when the matrix is dominated by human settlements or agriculture. The size of the remaining fragment is also important; large fragments conserve more biodiversity than small ones (**Figure 14**).

A variety of recent studies of fragmentation in the tropics suggest that habitat fragmentation reduces global biodiversity but not necessarily local biodiversity (Schelhas and Greenberg, 1996; Laurance and Bierregaard, 1997). This seeming paradox occurs because some species benefit whereas others suffer from fragmentation. The result is that local communities change composition after fragmentation, but the overall number of species in a local area does not necessarily decline. Forest interior specialists and wide-ranging species, notably predators, may go extinct with extensive fragmentation, thus lowering global biodiversity. Generalist and exotic species typically increase in fragmented habitats. Research suggests that many forest species tolerate considerable fragmentation because they can utilize the surrounding matrix. The extent of their ability to sustain viable populations depends on resources being available in the matrix and similarity of the matrix to the remaining native habitat.

Research in Singapore and Hong Kong is beginning to reveal the long-term effects of fragmentation. Forest fragments on these islands have been isolated for 150–350 years. Extinction of birds and mammals has been severe, but many plants are able to persist. Forest interior plants that are sensitive to changes in temperature and humidity (such as orchids and epiphytes) are also extinguished by fragmentation.

Habitat Degradation

Changes in land use can affect biodiversity without changing land cover. Land cover can remain essentially the same whereas the intensity of use increases and degrades habitat. Three forms of land use that degrade land cover rather than drastically changing it are (1) agricultural intensification, (2)

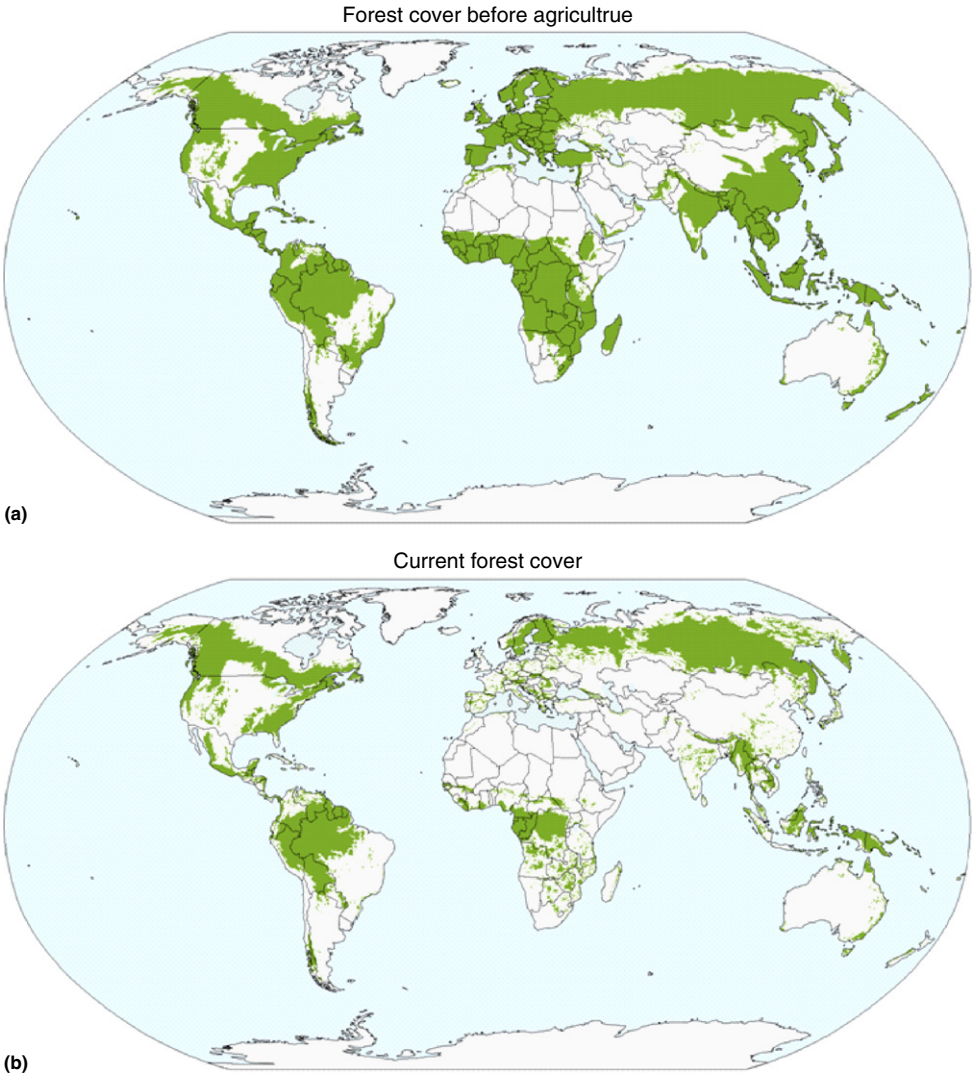


Figure 13 Fragmentation of the world's forests at present (b) compared to in preagricultural time (a). About half of the forest that was present under modern (i.e., post-Pleistocene) climatic conditions, and before the spread of human influence, has disappeared largely through the impact of man's activities. The spread of agriculture and animal husbandry, the harvesting of forests for timber and fuel, and the expansion of populated areas have all taken their toll on forests. The causes and timing of forest loss differ between regions and forest types, as do the current trends in change in forest cover. Until 1995, WCMC's work focused on tropical moist forests because of their high species diversity. GIS data were first assembled for closed moist tropical forests and used to publish the three volumes of the Conservation Atlas of Tropical Forests, covering Asia (1991), Africa (1992), and the Americas (1996). Because digital data were rare at this time, the process of assembling the forest cover data sets involved digitizing manually many paper maps. Continuing on from the tropical moist forest mapping, the next major initiative was to create the first 'World Forest Map'. This was produced in 1996 and was the first digital global forest map showing actual forest extent and protected areas with forested land. Since this achievement, significant work has been carried out to improve data sources and fill in gaps which occurred in this first attempt. This led to the production of the 'Global Overview of Forest Conservation CDROM' in 1997.

recreation, and (3) forest management for sustainable timber production.

Agricultural Intensification in Europe

Agriculture is the dominant form of land use in Europe. Conversion of forests to agricultural lands has resulted in dramatic changes in land cover. Forests once covered more than 80% of Europe, but today they cover 45% of the land (FRA2010). In contrast, slightly more than 40% of the European landmass is now agricultural. Over one-third of this land is permanent grassland maintained for grazing; the remainder is arable land. Conversion of forest to agriculture is no longer a major conservation concern in portions of Europe like Britain that have few remaining primary forests. Instead, the intensification of agricultural practices is degrading the conservation value of lands with grave consequences for biodiversity.

Agricultural intensification in Europe comes in many guises (Newton, 1998). (1) Increased use of chemical pesticides

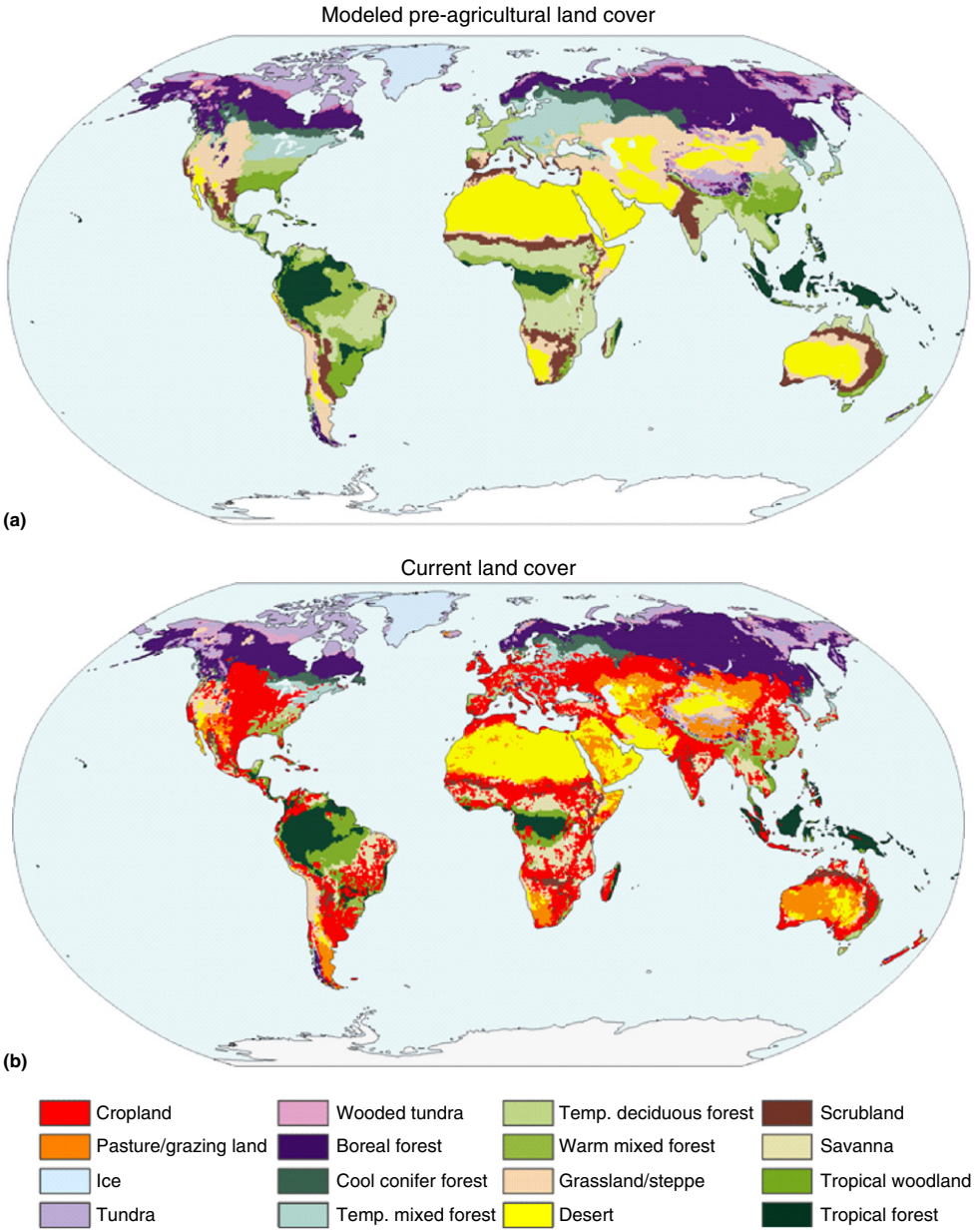


Figure 14 Preagricultural land cover as modeled by the BIOME model (Prentice *et al.*, 1992) and current land cover as modeled by the IMAGE 2.4 model (Bouwman *et al.*, 2006).

and fertilizers. (2) Removal of hedges to increase field size. (3) Plowing in late summer after harvest rather than waiting until the following spring. (4) Draining wetlands to increase arable and pasture lands. (5) Conversion of mixed farms that produced a variety of animal and plant products to monoculture cereal farms. (6) Earlier harvest dates. (7) Intensive grassland management to increase grass growth. (8) Increased stocking density of sheep in hill country.

These changes in agricultural practices cause declines in bird distribution and density for three main reasons (Newton, 1998): (1) chemicals and early harvests reduce breeding success and survivorship of birds; (2) breeding and foraging habitat is reduced by hedge removal, land drainage, late summer plowing, intensive grassland management, and increased grazing; and (3) habitat diversity and resulting bird community diversity are reduced when mixed farms are converted to monocultures. The results of these changes have been dramatic. It is suspected that intensification has affected 42% of bird species of conservation concern in Europe. Nearly 90% of the 26 farmland birds in Britain decreased their range of occurrence from 1970 to 1990, and those species using pastures were negatively affected by increasing grazing intensity. Specific aspects of agricultural intensification affect particular species, but on the whole this form of habitat degradation threatens a significant portion of the common, as well as the rare, European avifauna.

Influence of Recreation on Wildlife Communities

As our landscape becomes more dominated by urban and agricultural land, our psychological need to visit and explore the last remaining wild places increases. We often fulfill this desire by camping, hiking, nature watching, hunting, and fishing in wildlands. These seemingly benign activities can affect biodiversity by harassing animals, trampling plants, and altering competitive, facilitative, and predator–prey relationships (Knight and Gutzwiller, 1995). Although recreation in wildlands rarely changes land cover, its effects are often greatest near settlements, thereby extending the actual area modified by settlement.

How campers and fishermen affect salmon-scavenging birds in the Pacific Northwest of North America is a good example of the subtle ways that our land use affects biodiversity without changing land cover. Three species of birds (American Crow, Common Raven (*Corvus corax*), and Bald Eagle (*Haliaeetus leucocephalus*)) vie for salmon that die after spawning each fall and winter. Bald Eagles are the largest and dominant scavenger, but they are also the least tolerant of human activity. Sport fishermen float and wade rivers in search of live salmon and often flush eagles from their meals. This disrupts eagle foraging and may lead to lower overwinter survival and reduced population viability. The rarity of eagles and their status as a national icon increase the desire of nature watchers to observe them during the winter, thereby disturbing them even further. Recreation effects do not stop here. Camping areas along rivers are common and they provide consistent sources of food that fuel an increase in the wildland crow abundance and distribution. Crows are tolerant of humans and quickly return to carcasses after disturbance. They attempt to eat carcasses dominated by eagles and ravens but rarely succeed until humans displace the eagles (and often ravens), thereby incidentally allowing crows to eat without the check of competitors.

Effects of Intensive Timber Management

Timber harvest directly affects biodiversity by (1) facilitating land cover change through clearing land for agriculture and settlement and (2) fragmenting contiguous primary reserves (*see* Habitat Conversion) discussed earlier. Intensive forest management to produce timber may also affect biodiversity in more subtle ways by (1) reducing the average age of the forest, (2) simplifying the age structure and diversity within forest stands, and (3) increasing the frequency of disturbance in forests.

An intensively managed forest is like an intensively managed agricultural field, with the focus on producing the most marketable timber in the least amount of time. However, the effects of timber management on biodiversity appear to be much less than the effects of agricultural intensification. Animals with large space requirements, especially requirements for the interiors of mature forests, do not fare well in managed forests. But animals and plants that utilize clearings, thickets, and young forests thrive in managed forests, which can actually increase local biodiversity by providing habitat for such species. Nonetheless, the loss of sensitive interior forest species usually leads to lower global biodiversity.

The long-term effects of intensive timber management on biodiversity are poorly known. We do not know if managed forests provide adequate resources for many birds during the nonbreeding season, nor do we know how the shifting nature of forest patches in time and space, so typical of managed forests, affects the ability of plants and animals to maintain viable populations connected by dispersal. It is safe to conclude that maintaining forested land cover is an important contribution to global conservation. The maintenance of managed forests, that perpetuate native forest trees and some of the biodiversity dependent on them certainly helps stem the global loss of native habitat to agriculture and settlement.

Methodological Issues in the Estimation of Land Cover and Land Use

Much of the information presented in this article depends on the estimation of current land use and land cover on a global scale. Our conclusions depend directly on the quality of these estimates. Therefore it is important for the reader to have some insights into how these data are obtained and the degree of uncertainty they contain.

How Are Land Cover and Land Use Estimated?

Land cover is estimated directly by aerial photography at small scales and by satellite imagery at large scales. Aerial photographs resolve fine details in land cover at local or regional scales. Satellite imagery detects features over hundreds of kilometers, but small details are missed because the resolution is coarser.

An important satellite used for imaging the earth is Landsat. Its imaging abilities have improved over the years: Current Landsat TM (thematic mapper) images have better resolution (30 m $\times$ 30 m) that permits the fine-scaled discrimination of land covers. Another satellite that is widely used is the Advanced Very High Resolution Radiometer (AVHRR). Its wider swath and coarser resolution (1.1 km) are good for describing large areas of vegetation.

Although modern land cover can be estimated directly using photography and satellite imagery, land use cannot always be derived from the images. Hence statistics of land use supplement or substitute for imagery. The Food and Agriculture Organization of the United Nations (FAO) has reported tabulated data on land use and cover on an annual basis since the 1950s (these data are used in **Figure 6**). The FAO does not gather data independently but rather compiles national data from returned questionnaires. Globally, the FAO is an important source for time-series estimates of land use and cover from the second half of the twentieth century till present. Recent advances in global satellite products resulted in the GlobCover project. It is an ESA initiative which began in 2005 in partnership with JRC, EEA, FAO, UNEP, GOFC-GOLD, and IGBP. The GlobCover project has developed a service capable of delivering global composites and land cover maps using as input observations from the 300 m MERIS sensor on board the ENVISAT satellite mission (see http://ionia1.esrin.esa.int/).

Limitations in Estimating Land Cover and Land Use

Virtually every determination of land use and land cover is an estimate that includes some level of smoothing, averaging, and guessing. The general conclusions we make about how land use

affects land cover and biodiversity are not affected by these limitations. However, the application of the maps and processes to specific areas on the planet may fail because of estimation.

An important limitation occurs in large-scale mapping of land use and cover. Some minimum mapping unit (the resolution) is used in all compilations of land cover and use. Only one use or cover can be assigned to this minimum unit, and it is the dominant use or cover that is assigned and assumed to occur uniformly over the entire unit. For example, a unit that is simply recorded as forest might actually be covered by 40% forest, 30% lake, 20% grassland, and 10% building. Recording only the dominant cover reduces the occurrence of dispersed (rarely dominant) covers that typify many human-dominated land covers. Therefore, large-scale maps likely underestimate the actual effect that humans have on land cover.

Projected early estimates of land use and cover are by necessity approximations. The map of preagricultural land cover (see **Figure 3**) was derived entirely by modeling the plant growth suspected under varying climatological and soil conditions. Cropland estimates in the 1700s were estimated by assuming that 0.2 ha of land was cleared for each human (Richards, 1990). This number is highly debatable because it varies considerably per region and per time period (Klein Goldewijk *et al.*, 2011) Even the population size at that time was an estimate and has considerable uncertainty in many regions (Klein Goldewijk *et al.*, 2010). Regional estimates of land cover and use are also of inconsistent quality. FAO data vary greatly in quality by country because uniform scales and cover categories are not used. Yet even these large sources of estimation and inaccuracy do not affect the general conclusion that at global scale agriculture has increased tremendously, replaced native forests and grasslands, and caused declines in biodiversity.

Another important bias is an underestimation of habitat degradation. Mapping and imagery capture the basic pattern of habitat conversion and fragmentation, but rarely detect changes in the quality of habitats that retain their basic structure. Selective logging, understory trampling and removal, and disturbance by human presence are not recorded on global scales. Regional, local, and more subjective measures are needed to assess these changes in land use.

See also: Agriculture, Industrialized. Agriculture, Sustainable. Biodiversity in Logged and Managed Forests. Deforestation and Land Clearing. Desertification. Grassland Ecosystems. Landscape Diversity. Range Ecology, Global Livestock Influences. Timber Industry

References

Alkemade R, van Oorschot M, Miles L, Nelleman C, Bakkenes M, and ten Brink B (2009) GLOBIO3: A framework to investigate options for reducing global terrestrial biodiverity loss. *Ecosystems* 12.

Arino O, Ranera F, Bicheron P, *et al.* (2008) *Globcover.* Frascati: European Space Agency, (Special Publication) ESA SP.

Berry BJL (1991) Urbanization. In: Turner II BL, Clark WC, Kates WR, Richards JF, Matthews JT, and Meyer WB, (eds.) *The Earth as Transformed by Human Action,* vol. 1990. pp. 103–119. Cambridge, UK: Cambridge University Press.

Bouwman AF, Kram T, and Klein Goldewijk K (2006) *Integrated modelling of global environmental change. An Overview of IMAGE*, vol. 2.4. p. 228. Bilthoven: Netherlands Environmental Assessment Agency.

Collins NM, Sayer JA, and Whitmore TC (eds.) (1991) *The Conservation Atlas of Tropical Forests: Asia and the Pacific.* New York: Simon and Schuster.

Ellis EC (2010) Land-use and land-cover change, Encyclopedia of the Earth, http://www.eoearth.org/article/Land-use and land-cover change

Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, and Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography* 19: 589–606.

FRA2010 (2010) *Global Forest Resource Assessment 2010.* Rome: FAO.

Haberl H, Erb K-H, and Krausmann F (2007) Global human appropriation of net primary production (HANPP). In: McGinley EM (ed.) *The Encyclopedia of Earth.* http://www.eoearth.org/article/Global_human_appropriation_of_net_primary_production_(HANPP).

IUCN (International Union for the Conservation of Nature) (1996) The 1996 IUCN red list of threatened animals. IUCN, Gland, Switzerland. In: Knight RL and Gutzwiller KJ (eds.) *Wildlife and Recreationists.* Washington, DC: Island Press.

Klein Goldewijk K (2001) Estimating global land use change over the past 300 years: The HYDE database. *Global Biogeochemical Cycles* 15: 417–433.

Klein Goldewijk K, Beusen A, and Janssen P (2010) Long term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. *The Holocene* 20: 565–573.

Klein Goldewijk K, Beusen A, van Drecht G, and de Vos M (2011) The HYDE 3.1 spatially explicit database of human induced land use change over the past 12,000 years. *Global Ecology and Biogeography* 20(1): 73–86.http://dx.doi.org/10.1111/j.1466-8238.2010.00587.x.

Kummer DM (1992) *Deforestation in the Postwar Philippines.* Chicago: University of Chicago Press.

Kummer DM and Turner II BL (1994) The human causes of deforestation in Southeast Asia. *Bioscience* 44: 323–328.

Laurance WF and Bierregaard Jr. RO (eds.) (1997) *Tropical Forest Remnants.* Chicago: University of Chicago Press.

Mather AS (1990) *Global Forest Resources*, 1st edn. London: Belhaven Press.

Matson PA, Parton WJ, Power AG, and Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 277: 504–509.

Meyer WB and Turner II BL (1992) Human population growth and global land-use/cover change. *Annual Review of Ecology and Systematics* 23: 39–61. In: Meyer WB and TurnerII BL (eds.) (1994) *Changes in Land Use and Land Cover: A Global Perspective.* Cambridge, UK: Cambridge University Press.

Newton I (1998) Bird conservation problems resulting from agricultural intensification in Europe. In: Marzluff JM and Sallabanks R (eds.) *Avian Conservation: Research and Management*, pp. 307–322. Washington, DC: Island Press.

Potere D and Schneider A (2007) A critical look at representations of urban areas in global maps. *GeoJournal* 69: 55–80.

Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA, and Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19: 117–134.

Ramankutty N and Foley JA (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13: 997–1027.

Richards JF (1990) Land transformation. In: Turner II BL, Clark WC, Kates WR, Richards JF, Matthews JT, and Meyer WB (eds.) *The Earth as Transformed by Human Action*, pp. 163–178. Cambridge, UK: Cambridge University Press.

Schelhas J and Greenberg R (eds.) (1996) *Forest Patches in Tropical Landscapes.* Washington, DC: Island Press.

Steffen W, Crutzen PJ, and McNeill JR (2007) The anthropocene: Are humans now overwhelming the great forces of nature? *Ambio* 36: 614–621.

Turner II BL, Clark WC, Kates RW, *et al.* (1990) *The Earth as Transformed by Human Action: Global Change and Regional Changes in the Biosphere Over the Past 300 Years*, 1st edn. Cambridge, MA, USA: Cambdrige University Press.

Turner II BL, Clark WC, Kates WR, Richards JF, Matthews JT, and Meyer WB (eds.) (1990) *The Earth as Transformed by Human Action.* Cambridge, UK: Cambridge University Press.

UN (2009) *World Urbanization Prospects, The 2008 Revision.* New York: United Nations Population Division.

Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of earth's ecosystems. *Science* 277: 494–499.

Williams M (2003) *Deforesting the Earth, from Prehistory to Global Crisis.* Chicago, USA: University of Chicago Press.

Land-Use Patterns, Historic

Oliver Rackham, University of Cambridge, Cambridge, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Climax The stable ecosystem that is supposed to result from a sufficiently long period of unchanging environment.
Clone A group of individuals formed by vegetative reproduction from a parent. This term applies particularly to those trees and plants that reproduce by creeping stems or by suckers from roots, forming circular patches of genetically identical individuals.
Coppice Regrowth of a felled tree from the stump.
Cultural ecosystem Ecosystem produced by the long-term interaction of wild plants and animals with human activities.
Endemic Plant or animal species limited to a small area of the world, for example, one island or mountain range.
Heath Open vegetation dominated by dwarf undershrubs usually of the family Ericaceae.
Meadow Grassland maintained by mowing.
Moor Open vegetation, dominated especially by dwarf Ericaceae and *Sphagnum* mosses, on peaty soil.
Pasture Grassland maintained by grazing domestic animals.
Pollard A tree repeatedly cut at about 6–10 ft above ground to produce successive crops of wood.
Savanna Grassland (or other nonforest vegetation) with scattered trees.
Seminatural Vegetation that owes its character to human activity, but in which the plants are wild, not sown or planted. Examples: coppice woods; most heathland; some savannas.
Wildwood Wholly natural forest not affected by sedentary human activities such as agriculture and pasturage, widespread in prehistoric times (including earlier interglacial periods) and still surviving in remote places.
Woodland Forest forming part of a cultural landscape.
Wood-pasture Cultural ecosystems, often savanna-like, combining trees and domestic animals.

Introduction

Human intervention often does not efface all signs of previous ecosystems. The traveler in Texas, despite 170 years of often intense activity by settlers, has no difficulty in recognizing the presettlement vegetation regions (Pineywoods, Blackland Prairie, Post Oak Savanna, etc.).

Even in Britain, after 6000 years of settlement, the landscape is still not wholly artificial. By careful search, one still finds remains of the mid-Holocene vegetation categories (limewoods in lowland England, oakwoods in the north and west, pinewoods in east Scotland). Superimposed on these are the largely artificial ecosystems of arable farmland, sown grassland, and forestry plantations; the seminatural ecosystems produced by the responses of wild plants and animals to continuing land-use (old pasture, hedges, moorland) or to past land-use (abandoned mines, derelict forestry plantations); and the wealth of biodiversity contained in cities, old pits, and industrially derelict land.

Cultural ecosystems are often contrasted with wildwood or "virgin forest," the aboriginal vegetation of much of the world's land surface before the impact of settled human cultures. Wildwood itself, however, may already have been altered by the activities of nonsettled peoples. There is abundant evidence to show that hunting and gathering were not simple, unorganized activities, but could have profound effects on landscape and ecology, even in places remote from the activities themselves. A probable example is the extermination of elephants and other great mammals by Paleolithic peoples. The loss of the only animals capable of breaking down big trees could hardly fail to have profound, if as yet little known, consequences. From the early Holocene, if not earlier, land management by burning affected the ecology where vegetation was combustible, especially in Australia and North America. These activities, as far as we know, created only simple land-use patterns: they would alter, for example, the boundaries between forest and prairie or between combustible and noncombustible vegetation. Some of the Texas ecosystems, such as the Edwards Plateau savannas, were probably produced by Native American land-use and taken over by settlers. The introduction of settlements, domestic animals, and cultivated plants – in European terms, the coming of Neolithic cultures within the past 8000 years – began to create the diversity and complexity of land-use patterns that is familiar today.

Human Activities and Climatic Change

It used to be thought that, given enough time, the plant and animal communities in an area would settle down to a stable *climax* state, determined only by the climate and (maybe) geology and changing only slowly in response to evolution. This may have been so in the Tertiary geological period, but in the Quaternary the dominant factor has been climatic change. Glacial cycles in high latitudes destroyed all vegetation over large areas, including much of Britain. In middle latitudes and

even in the tropics, although there has been little ice cover, climatic changes such as drought have been sufficient to displace or destroy most of the interglacial plant communities.

There have been many glacial cycles over the past 2 million years. In much of the world, forest has been the "normal" vegetation only in interglacial periods like the present, of at most a few tens of thousands of years each. Changes of climate have been too rapid for plants and animals, apart from annuals and other fast-reproducing species, to keep up with them by evolutionary change.

Human activity has recently become the predominant force shaping the world's ecosystems. However, it did not burst into a world of stable environment in which evolution had previously been the predominant force. It came on top of a period of unusual instability: climatic change had been operating on a time scale not much longer than that of human activity itself. In Britain humanized ecosystems have existed for about 6000 years; the "virgin forest" preceding them had existed for at most another 6000 years and was itself affected by the activities of Mesolithic and Paleolithic people.

Climatic change and human activity overlap in time, and their effects can be difficult to separate. In the Mediterranean, the present, rather arid, ecosystems are variously attributed to the strongly seasonal climate with its hot dry summers and to the activities of people "degrading" what would otherwise be a more forested landscape. In reality, the present climate is only about 6000 years old and set in at the same time that prehistoric people were having a profound effect on the landscape.

Properties of Crops and Domestic Animals

Land-use patterns depend on the ecological behavior of a few species of major crop plants and domestic animals, which are usually not indigenous to the region.

Most of the crop plants of European-style agriculture originated in the steppes of southwest and central Asia and have a particular set of ecological requirements. They have modest moisture needs, rather high needs for mineral nutrients, and are very intolerant of waterlogging and shade. As they spread, first to Europe and later to European colonies, farmers strove to convert more and more of the world's ecosystems to an imitation of a Turkish steppe in which wheat and barley are at home. Of the multitude of indigenous American crop plants, only a few – corn, potato, and tobacco – have been widely grown in Europe; these are the few whose ecological requirements are not too different from those of Southwest Asian crops, although they need more moisture. Rice is a Southeast Asian crop with very different requirements, growing in wetlands and making possible the great wetland civilizations of Asia. Hence it comes about that generations of farmers have sought to destroy Europe's fens and convert them into artificial steppe, whereas Southeast Asian farmers, with the utmost ingenuity, have contrived to turn hillsides into artificial fen (**Figure 1**).

Many tropical crops, however, will grow in moderate shade, and some require it. The distinction between forest and farmland, which is so sharp in most European-style agriculture, is not always clear in the tropics.



Figure 1 Cultural landscape of terraced rice fields, carefully leveled so as to be irrigated. Land of the Hani people, Yunnan, SW China.

Keeping domestic animals also involves destroying or modifying natural ecosystems. Cattle, sheep, and goats eat tree leaves but are not forest animals. If kept in the shade of trees, they soon eat all the leafage within reach, creating a *browse-line* beneath which nothing is allowed to grow. To keep them in manageable numbers requires at least some gaps between the trees, in which grasses and low herbage will grow. Pigs, although often thought of as acorn feeders, cannot be kept in forest alone. Acorns are available for only a few months of the year and contain insufficient protein to sustain pigs by themselves. The world's cultural ecosystems would be very different had monkeys been domesticated for human food.

Fields, Hedges, and Terraces

In some parts of the world, arable fields are impermanent, being created out of forest or grassland, cultivated for a few years, and then allowed to revert to natural vegetation. The landscape consists of a mosaic of areas in different stages of reversion. This *shifting cultivation* is typical of regions too infertile to support even primitive agriculture permanently, especially in the tropics. In more fertile regions, such as most of Europe, people avoid having to repeat the immense labor of creating new farmland. Fields are permanent and tend to be private property with fixed boundaries.

In regions not yet reached by weedkillers, arable fields are seminatural ecosystems. Many weeds came from the original homelands of the crops they accompanied, and some annuals have evolved in parallel with particular crops: for example, cow-wheat (*Melampyrum arvense*), whose seeds resemble a grain of wheat, so that they are sown along with the wheat crop. In Japan, rice fields are complex aquatic ecosystems involving cyanophytes, water ferns, amphibians, and insects; some of the plants fix nitrogen and so benefit the rice crop.

Fields tend to be discontinuous. As well as boundaries where they adjoin forest, pasture, or roads, they have boundaries separating ownerships or subdivisions within an ownership (**Figure 2**). These are determined both by practicalities (e.g., erosion control) and by anthropological factors, such as land



Figure 2 Anglo-Welsh cultural landscape of small irregular fields, thick hedges, many hedgerow trees, and scattered farmsteads, of unknown origin but likely to be well over 500 years old. *Llanveynoe, Herefordshire, England.*



Figure 3 Ancient hedge between two fields, managed by coppicing. It has been cut down (leaving a few trees standing) and is beginning to sprout. In 2 years it will be back to its normal state. *Polstead, Suffolk, England.*

tenure, inheritance, or collectivization. The boundaries (hedges, banks, walls, etc.) provide a network of relatively natural ecosystems permeating the farmland. The size of fields varies from country to country. In parts of Canada or Australia, 100 acres is a small field. In Texas, the nineteenth-century allotments of 1 square mile (640 acres) are by now divided into at least 30 compartments in several ownerships. In the hedged parts of medieval England, a typical field would be of about 4 acres. In Crete, fields tend to be of about 1 acre. In Japan, with rice as an immensely high-yielding crop, even this would be a big field.

A *hedge* (in North America, *a fence-row*) is a row or strip of trees or shrubs forming the boundary of a field. In Europe, these are often maintained by cutting in particular styles to form a barrier to keep cattle or sheep from straying into the next field; often trees are left at intervals either uncut or pollarded (**Figure 3**). In America, they are less managed.

In America and to some extent Europe, most hedges have arisen spontaneously as trees have taken root from wind- or bird-dispersed seed at the bases of fences or on balks or walls between fields. In England, hedges are thought of as deliberately planted, but can arise naturally, as many of the older hedges probably did. In Australia, hedges are rare, evidently because most eucalyptuses and other native trees have poor powers of seed dispersal.



Figure 4 Cretan cultural landscape in a semiarid climate. In the bottom of the basin are houses grouped into hamlets, with gardens, orchards, and small fields demarcated by dry-stone walls. Extending up the mountains are remains of terraces, on which grain used to be grown. *Asphéndou, Sphakiá, Crete.*

Hedges presumably go back in some countries to the beginnings of agriculture. In Europe, the earliest evidence for them is about 2000 years old; they probably go back much further, but leave little archeological record. In England, they are well documented for about 1200 years. Many individual hedges are at least 500 years old; these tend to be mixed in their composition, unlike the monotonous hedges of more recent origin.

Hedges are often thought of as linear wood-lots. This is partly true: in England, many hedges were periodically felled as if they were woodland, yielding crops of fuel wood and timber. However, in England and (from what the author has seen of them in America), hedges are distinct vegetation types in themselves, seldom corresponding in detail to the local forest. They usually lack the more exacting woodland plants, apart from those few hedges that are the surviving edges of grubbed-out wood-lots.

Hedges are important as diversifying what might otherwise be featureless farmland. They introduce tree-living (though not forest) birds to open country that would lack nesting sites. They are often regarded as corridors for movement of mammals across farmland.

In many mountainous and densely populated parts of the world, slopes are shaped into terraces to create flattish surfaces on which to grow crops (**Figure 4**). Banks or walls between terraces, like hedges between fields, can be significant sites for seminatural vegetation.

Forest and Woodland

It is often supposed that wildwood consisted of trees, trees, and nothing but trees, and only shade-bearing ground vegetation; gaps arising through the death of trees would promptly be filled by the growth of new trees. Examples can be seen where the dominant trees are of shade-bearing kinds such as *Abies* and *Fagus* species. The argument that wildwood was like this is often based on the general doctrine of climax vegetation, rather than on evidence of the history of specific forests. Pollen analysis can exaggerate the dominance of trees, many of which produce more pollen than herbaceous plants. If this is allowed for, some wildwoods were much more diverse.

Frans Vera has advanced the theory that wildwood was a mosaic of areas of trees and areas of grassland, continually shifting under the influence of deer, wild oxen, and wild horses. Whether this was so is still controversial. It would explain why agricultural technology should have spread into the seemingly hostile environment of north-west Europe, where (in the continuous-forest theory) people would have had to expend immense labor on digging up trees to make fields before they could grow anything. It also explains why wood-pasture is such an important habitat.

In some tropical forests, there is vast biodiversity, owing to the many species of tree and the abundance of vines, epiphytes, and termites, which enable the trees themselves to support complex ecosystems. In other parts of the world, continuous forests, especially of densely shading trees, are relatively poor habitats: many of their plants and animals are concentrated in gaps, cliffs, watercourses, burnt areas, and other breaks in the continuous shade.

Ancient peoples destroyed some areas of forest to create farmland, grassland, and heath; they also affected the remaining forest by various forms of land management. Continuous forest is not very productive for most human purposes. Edible animals and plants occur sparsely, if at all, and the animals are difficult to catch. Tree fruits are out of reach in the high canopy.

Forests were not generally destroyed by people cutting down the trees in order to use them. The author knows of no instance in European history of a forest being destroyed – converted to



Figure 5 A coppice wood. The foreground has just been felled, leaving the stools from which new growth will arise. In the middle is an area of 2 years' growth since last felling. An older part of the wood forms the background. *Bradfield Woods, Suffolk, England.*

nonforest – solely by people using up the trees. Normally they would cut the trees suitable for the purpose in hand and leave the other trees. The result would be a depleted forest, unsuitable for that purpose until a new generation of trees had grown. This is not to be confused with a destroyed forest.

Great trees do not easily furnish wood for fuel and timber for construction. A big tree, when cut down, is a very intractable object. Until the coming of sawmills, vehicles, and railroads capable of dealing with giant trees, people preferred to use the smallest log that would serve the purpose and to manage forests to produce a succession of trees small enough to handle.

Coppicing is an important factor. Most European and North American trees, other than conifers, survive being cut down and sprout from the stump. Clonal trees, like European elms and American beech, sucker from the roots. Coppicing is one of the world's most important practices in historic forest management. By cutting down woodland every 5–30 years and allowing it to grow again from the stumps, a permanent succession of small stems can be assured, of sizes that are easily handled and suitable for light construction and fuel. It was often the practice to leave a scatter of trees of selected species to grow on for three or four cycles to yield constructional timber (**Figure 5**).

Coppicing has been the nearly universal woodland management in Britain, well documented for the past thousand years and known on archeological evidence for some 6000 years. It is the basis of historic forest management in many parts of the broad-leaved and Mediterranean zones of Europe and also in Japan. It was apparently not much practiced by Native Americans but was widely introduced by European settlers in America, where there are large areas of ex-coppiced wood-lots.

Coppicing affects biodiversity by drastically reducing the shade at the start of each cycle of felling and regrowth. Two or three years of relatively open conditions follow, ending as the new growth closes in. This favors various woodland plants and animals. Low-growing herbs such as species of *Viola* and *Primula* flower in abundance in the years of extra light (**Figure 6**). Others such as *Euphorbia* species appear from buried seed produced by their parents at the last felling. Many insects feed on the leaves or nectar of these plants. The middle stages of regrowth, when there is a thicket of young stems, favor warblers and similar small birds – a famous English example is the nightingale (*Luscinia megarhynchos*). The dormouse (*Muscardinus avellanarius*), an English woodland mammal, favors the later stages.

Coppicing is often thought to be artificial, but the ability to coppice is widespread among the world's trees and presumably was an adaptation to some process in wildwood. Sometimes it is a response to fire, but it is not correlated with flammability: few pines (among the world's most flammable trees) will coppice, but fireproof trees such as elms and poplars coppice or sucker. American, European, and Japanese species of *Tilia* (lime, basswood) are self-coppicing and grow naturally in a multistemmed form; so do American and Japanese species of *Magnolia*. Possibly coppicing behavior is an adaptation to tree-breaking mammals: the axes of woodcutters are a replacement of the missing elephants, etc. The author has observed something like a coppicing ecosystem developed in eucalyptus forests in New South Wales after only the third or fourth successive logging (**Figure 7**).

Another historic woodland practice is the creation of permanent edges and open areas. In England, wood-lots have permanent edges (many of them are over a thousand years old) defined by banks and ditches, constructed as a conservation measure (**Figure 8**). They may also have permanent tracks and other open areas in the interior. In a typical wood-lot, well over half the plant species are associated with the boundary, with recently felled areas, or with permanent openings. The species of permanent openings tend to constitute plant communities of their own, distinct from those of



Figure 6 Oxlip *(Primula elatior)* flowering in abundance among hazel (*Corylus*) shoots of 1 year's growth since coppicing. *Hayley Wood, Cambridgeshire, England.*



Figure 7 Eucalyptus forest 2 years after being logged for the third or fourth time. Note the abundant regrowth and ground vegetation. *Merimbula State Forest, New South Wales.*

temporary clearings and of grassland away from woodland. On pollen evidence, something like these permanent open areas already existed in wildwood.

Wood-Pasture and Savanna

Where the environment becomes too dry or cold for forest or where there is too much grazing, forest changes into nonforest in various ways. The trees may suddenly stop, or there may be a zone of trees reduced to the stature of shrubs, or a mosaic of patches of forest and patches of nonforest. Alternatively, there may be a zone, often of great extent, of scattered trees among grassland. This constitutes tropical and subtropical savanna ecosystems and their extensions into northern latitudes. The trees may be single, as in the wood-pasture ecosystems of England, the *dehesa* landscapes of southwestern Spain, and the eucalyptus savannas of Australia. Or they may be in groups, as with the oak and elm *motts* of middle Texas (**Figure 9**). This depends on the properties of the trees: Texan *Quercus fusiformis* and *Ulmus crassifolia* are clonal trees that sucker.

It used to be thought that the "natural" savannas of Africa and Australia were very distinct from the semiartificial treed grasslands and treed heaths of Europe, which were regarded as



Figure 8 Edge of an English wood. The bank with external ditch was constructed about 1000 years ago by the monks of Bury St. Edmunds Abbey to demarcate their wood-lot. The trees on the bank have just been cut; in the interior, they are of 1 year's growth. *Bradfield Woods, Suffolk, England.*



Figure 9 Savanna in middle Texas. The prairie (rich in spring and summer flowers) is dotted with motts – clonal patches – of Texas live oak and Texas elm. This is a cultural landscape of Native American origin, which was already in place when the area was settled in 1846. It needs to be maintained by grazing or burning, without which it is invaded by juniper (*Juniperus ashei*) and turns into forest. *Valley Mills, Texas.*

"degraded forest," meaning forest from which some trees had been removed. Recent research makes the distinction less clear: many tropical savannas are at least partly the result of ancient land management, especially by burning. The pollen record of European wildwood, moreover, often contains nonshade-bearing undershrubs and herbs, showing that it must have had open areas as well as trees.

Cultural treed grasslands are usually differentiated from forest by there being enough browsing animals to prevent trees from occupying all the space. The amount of browsing needed to do this varies. Where the climate is so dry or cold as to be marginal for tree growth, it takes less browsing than in regions, such as England, well within the climatic limit of forest. (In some hotclimate savannas, conversely, the grasses



Figure 10 A complex medieval wood-pasture, combining old grassland, ancient pollard trees, and coppice wood-lots: a particularly important juxtaposition of habitats, reproducing some characteristics of wildwood. *Hatfield Forest, Essex, England.*

are the dominant partners, and too much browsing encourages trees.)

The limits of "natural" treed grassland are vague and controversial. The amount of browsing – the numbers and behavior of the animals – is the most difficult aspect of prehistoric ecosystems to ascertain. Trees tend to be exaggerated, because many nontree plants produce sparse or nondiagnostic pollen grains and also because pollen analysts are trained to interpret tree pollen in terms of forest rather than of single trees.

In glacial times, parts of Europe not covered by ice were evidently dry as well as cold. The natural vegetation consisted of either patches or scatters of trees, which pollen analysts call "forest steppe." This was the situation for at least four-fifths of the past 2 million years, punctuated by interglacials when large areas of forest became possible. In the present interglacial, when the climate became favorable for trees, nonshade-bearing plants became very local – except in the drier parts of southern Europe, where forest was still patchy and savanna-like.

When people took to keeping cattle, sheep, and goats, they replaced an unknown, but probably sparse, scatter of wild herbivores by local concentrations of domestic animals. To feed them, they turned forest into grassland; but they also kept domestic animals in the remaining forest. This involved managing the forest in a different way from where wood was the product. The trees would have to be sparse enough for pasture to grow between them on which the livestock might feed. Trees were also cropped, not by coppicing – for the animals would eat the young shoots and kill the stools – but by the practices of pollarding and shredding, cutting the tree high enough for the shoots to arise out of reach. Sometimes wood was the objective, but in climates where grass did not grow all year, it was the practice to harvest leafy tree branches to store instead of hay for feeding the animals in the cold of winter or drought of summer.

Such seminatural wood-pasture ecosystems came to occupy large areas of Europe and Asia, often associated with communal land management. In England, there was a further development with deer as semidomestic livestock from the eleventh century AD onward. Landowners would keep deer (native and introduced) for meat and as a status symbol in enclosed parks. The king, moreover, had the right to keep deer on certain commonlands, such as Epping Forest and the New Forest, and to kill and eat them. Parks and Forests often contained treed grassland on which deer and other livestock fed. There could also be elaborate systems of temporary enclosure to combine livestock with coppice woods.

Wood-pasture is a most important ecosystem, especially for insects and birds that require both trees (for example as nest sites) and grassland (**Figure 10**). The inland bird fauna of England as a whole falls into this category. Wood-pasture also can preserve ancient trees. A large proportion of the fauna and flora of trees, especially beetles and lichens, is associated with the specific habitats of old trees: old dry bark, heartwood rotted by particular fungi, and cavities of various sizes. These are specialized, poorly dispersed animals and plants and presumably existed in wildwood. Forest trees are not an ideal habitat, for the lives of old trees are often shortened by the competition of neighboring trees. Coppice woods and managed forests are inhospitable to them, unless the ancient bases of coppice stools are a suitable habitat. However, wood-pasture is often very good for old trees and the organisms that live on them. The trees are free from competition and can live to a great age, especially as pollarding prolongs the lives of many species. Much of the information so far comes from England, where the New Forest, Windsor Forest, and Sherwood Forest are major sites for the invertebrates of ancient trees.

Wood-pasture is now much reduced in northern latitudes, but the author has found traces of it as far north as Sweden.

The area surviving in England, though small, contains some supremely complex and important examples. It still flourishes in southern Europe, covering about one-sixth of Portugal and one-eighth of Spain, although few of these contain ancient trees. It extends eastward into the Himalayas. At one time, it was evidently extensive in Japan, where old pollard trees can still be found embedded in forests. In North America, Native Americans and settlers both operated savanna-like landscapes, but pollarding was rare. In Australia, savannas are widespread and sometimes contain trees dating from presettlement times, but often the grassland component has been displaced by introduced grasses and herbs.

Wood-pasture is an important ecosystem for biodiversity: it perpetuates a different set of plants and animals from coppice woods. The most significant examples remaining are those few in which ancient trees and seminatural grassland are both preserved.

Grassland

Many of the world's grasslands are regarded as natural prairies, in regions too dry, or too cold, or too much grazed for trees. Others are part of a cultural landscape in regions that would naturally be forested; if not grazed, they revert to forest. There are large intermediate areas where natural grassland has been extended by burning and other forms of land management.

Much grassland in Europe, America, and Australia is sown pasture, little different from other arable crops. This practice goes back only three centuries. Previously, grassland was permanent, managed as either pasture (grazed) or meadow (mown for hay). Pasturage often goes with communal land use, as in the American Wild West or the chalk downlands of England.

Seminatural pasture grassland in England goes back to the Neolithic period, when the forests that covered the chalklands were converted into open land; prehistoric monuments such as Stonehenge were meant to be visible from a distance. Often there was a phase of farmland before the historic grasslands were established.

Meadow is a more productive kind of grassland, yielding hay which is dried and stored to feed animals in winter when the grass is not growing. This requires iron tools with which to cut the grass. As a large-scale land use, meadow is mainly of the past 1500 years. It was located especially in floodplains, which in medieval England were the most valuable land. In the sixteenth century, elaborate irrigation systems were developed to advance the growth of meadow in spring. Grasslands were often part of a highly integrated pattern of land-uses that transferred nutrients, via the dung of cattle and sheep, to plowland and meadow.

A contrasting type of hay meadow is that on mountains, for example, in northern England or in the Alps. In Italy, in the northern Apennines, the grasslands, originating in the Iron Age, were maintained for centuries to feed local livestock and those of the city of Genoa; recently, they have largely turned into forest and their distinctive plant communities have been lost.

Old grasslands are the classic "seminatural" vegetation. They are artificial in that they were created by human activity, and disappear if that activity is withdrawn. They are natural in that nobody sowed the grasses and other plants of which they are composed. They are of many kinds, depending on whether they were pasture or meadow, on the geology, and on which animals fed on them. Chalk pasture is a dry, relatively infertile form of grassland that is one of the richest plant communities (in terms of plant species per square meter) in Britain. When grazed by sheep or rabbits, it is particularly rich in dwarf herbs and small undershrubs. At the other extreme, wet meadow is rich in tall grasses, sedges, and orchids.

Heathland

Heathland in northwestern Europe is open vegetation other than grassland, dominated by dwarf undershrubs especially of the family Ericaceae. It can be a very rich plant community, with many herbs and also characteristic butterflies and birds. It is a seminatural ecosystem, maintained by grazing and cutting; without these, it quickly turns into woodland (**Figure 11**).

Heathland began in a small way in the Mesolithic period, about 10,000 years ago; people may have encouraged it as part of land management in favor of deer. It expanded in later prehistory, especially as browsing by domestic livestock encroached on forest. Throughout the historic period, it was extensive in England, especially on acid soils in low-rainfall areas.

Heathland should perhaps be seen as a northward extension of the undershrubby plant communities of the Mediterranean, known as *garrigue* and *phrygana*. These, too, are often thought of as the result of destruction of forest: on palynological evidence, they were relatively local and specialized ecosystems until the Neolithic period. However, they are more stable than heath, occurring in places too dry for tree growth. Often they fail to turn into forest if grazing is removed. Phrygana is one of the richest ecosystems in endemic plants.

The heaths of northwestern Europe have fared very badly in the past 200 years. Little remains outside Britain, and even there, heaths, where not made into arable fields or forestry plantations, have been allowed to turn into woodland.

Moorland

Moorland, although often confused with heath, is really a southward extension of the peaty tundra of the Arctic. It is characteristic of northwestern Europe, especially mountains (**Figure 12**); it still covers half of Scotland.

Moorland is dominated usually by ericaceous undershrubs, especially heather, *Calluna vulgaris*. Unlike heath, it occurs in high-rainfall areas and has a peaty substratum. It grades into blanket bog, with pools and abundant *Sphagnum*.

Moorland is more stable than heath and less obviously a cultural ecosystem. It gradually extended during the Holocene, covering areas that once had forest. Sometimes there was an intermediate stage of pasture or even arable land: on the plateau of Dartmoor, southwest England, the boundaries of Bronze Age fields can still be traced. However, it is not clear



Figure 11 Heathland, several hundred years old, now insufficiently grazed: if trees are allowed to take it over, they will destroy its value as a habitat. *Holt, Norfolk, England*.

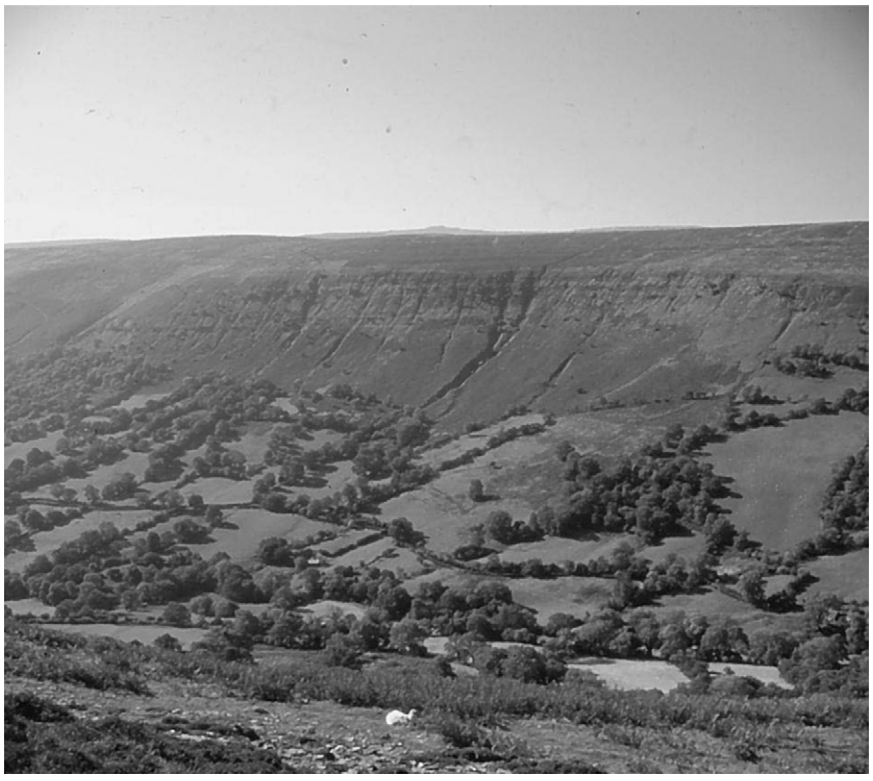


Figure 12 Medieval landscape on the border of Wales, with hedged fields and farmsteads in the valley and moorland on the thin soils of the plateau and on screes below the cliffs. *Cwm Olchon, Herefordshire, England*.

how far the spread of moorland was due to the lapse of time and accumulation of peat, to a change to a wetter climate encouraging peat to grow, or to human activities tipping the balance against forest.

In historic times, moorland has been much used for pasture and for digging peat as fuel. Large supplies of peat have made up for the scarcity of wood fuel in Scotland and Ireland. A skilled tradition of periodic burning has grown up to

maintain the pasture. In the past 200 years, moorland regions have lost human populations, but without forest returning. The main uses of moorland now are for grouse (*Lagopus scoticus*) and deer as game animals.

Buildings and Built-up Areas

Many plants and animals live on buildings, including some quite rare species: for example, the stork in Europe, which nests almost exclusively on buildings and structures, sometimes on platforms provided for it. In England, buildings and abandoned mines are the chief strongholds of bats.

Cities are often the chief concentrations of wildlife in their regions: well-known examples are Cambridge and Glasgow, which contain more wild plant species in a given area than anywhere else in their surroundings. This is because of the diverse habitats provided by gardens (despite the prevalent use of garden chemicals), buildings, and especially ruins and derelict land.

Derelict industrial land can be an important habitat. In England, the quarry that provided the excellent Barnack building stone is now one of the best examples of limestone grassland. Long-standing chemical contamination can produce its own ecosystems, especially where heavy-metal-tolerant plants have evolved.

Sacred sites develop in peculiar ways. In Japan, temples and shrines often own forests, which may preserve ancient trees and evidence of past management practices. In England, many churchyards are excellent habitats (**Figure 13**). They often contain their own peculiar types of grassland. Tombstones and the stonework of the church itself are important sites for rare lichens, especially in areas that lack natural rock outcrops.

The Rise and Fall of Historic Land-Use Patterns

Human intervention creates new ecosystems as well as destroying old ones. Cultural ecosystems are often complex, diverse, and stable. Sometimes they are modifications and extensions of ancient natural ecosystems. Chalk grassland, for example, is in some respects a prolongation in time of the grasslands of the glacial periods. Moorland is in some respects an extension of tundra into southern latitudes. Some, however, seem to be new creations: The author hesitates to guess what the habitat of the plant *Fritillaria meleagris* might have been before hay meadows had been invented.

Some conservationists regard ancient human intervention as destructive, like the modern logging of the Oregon redwoods or the early-modern ravaging of oceanic islands. Sometimes this was so: we would undoubtedly now disapprove of the destruction of almost all the native mammals of Crete by its early human inhabitants. Some interventions were experimental and unsustainable, like the prehistoric attempt to farm Dartmoor. Others may have been almost too gradual to notice, such as the erosion of tree cover on a common pasture.

The ecosystems of cultural landscapes could be preserved for centuries, both by ownership patterns and by deliberate conservation. By AD 1350, some 94% of the forest of England had been converted to nonforest uses. Woodland management had been invented long before, probably for other reasons, but as forest became scarce, conservation practices gradually spread to the remaining forest. It may be no coincidence that the idea of sustained yield is first set down in writing in 1356. Many medieval wood-lots still exist, and are often now nature reserves: "Ancient Woodland" is a scheduled category in modern conservation.

Historic land-use patterns usually worked on a small enough scale to result in useful juxtapositions of habitats. They often perpetuated habitats for non-shade-tolerant plants and animals. Because of the glacial history of the past 2 million years, many species are adapted to nonforest conditions and are threatened by the advance of forests in interglacial periods. They would have fared particularly badly in the present interglacial, because of the widespread extinction of giant tree-breaking mammals. For example, the greatest concentration of endemic plants in Europe is in the Mediterranean, where nearly all of them are nonforest species.

Historic land-use patterns have been in retreat for some 250 years. Usually this was because of outside ideas and interventions, not because they were unsustainable. Commonlands were not destroyed because of the operation of the "Tragedy of the Commons:" the supposed tendency of communal systems to break down because each participant pursues his own short-term advantage regardless of the rights of other participants. That had already been foreseen, and the participants in most commons had drawn up rules of use to prevent it.

In the eighteenth century, multiple and especially communal land-uses became unfashionable. The only proper use of cultivable land (or land thought to be cultivable) was conventional agriculture, carried on by private owners in rectangular plots of not less than 10 acres. For uncultivable land the proper use was timber production organized by the state on the model of either French or German "scientific" forestry (depending on which state). Agricultural writers invented these ideas and bent the ear of governments to get them put into practice; they were not deterred if their schemes failed to work. In this way, the open fields of England were destroyed, and most of the heathland was converted into poor-quality arable. In countries where modern forestry was influential, such as Sweden, Germany, and Italy, most of the wood-pasture ecosystems were destroyed.

In the twentieth century there have been four contrasting processes: extension of cultivation (including plantation forestry) into areas not previously cultivated; mechanization of cultivation in areas already cultivated; urban development; and land abandonment. Often, as in Crete, all four occur almost side by side. They destroy existing seminatural ecosystems without creating an equivalent. A common feature is that the grain of the landscape is coarsened: the small-scale pattern of juxtaposed land uses is replaced by monotonous expanses of the same land use.

One consequence is loss of habitat for creatures that require more than one habitat: birds that feed in the open but nest in tree holes; insects whose larvae feed in rotten wood but whose adults need a nectar source; plants that are weakly competitive when growing in the open but do not flower in

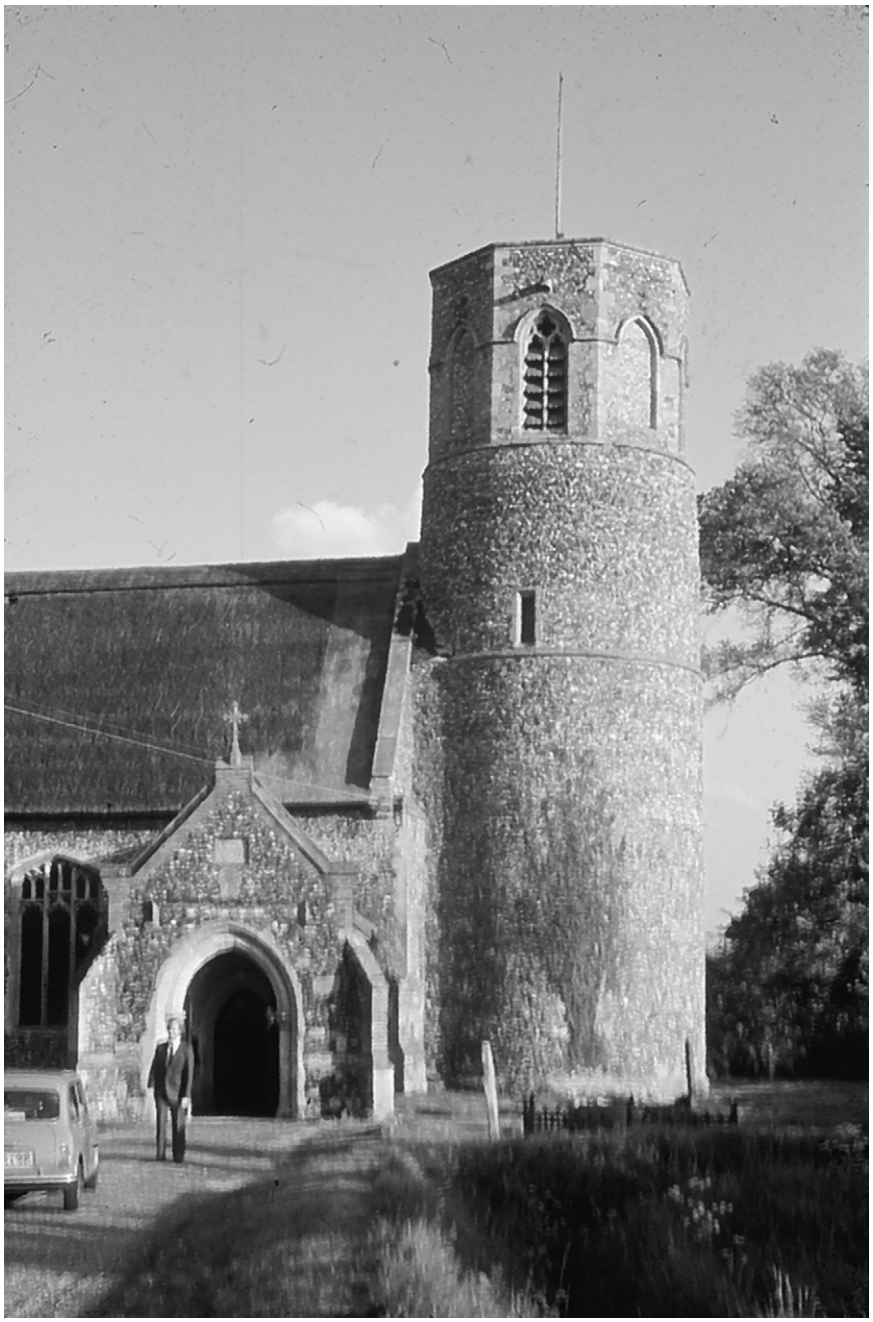


Figure 13 Churchyards have a range of habitats from old grassland (important for reptiles and amphibians) to the stone and brick surfaces of the boundary wall, tombstones, and the church itself (lichens) and the thatched or tiled roof (bryophytes, flowering plants). *Rockland St. Peter, Norfolk.*

shade – all these were favored by the mosaic of historical land-uses, but not by what has replaced them.

Land abandonment does not re-create wildwood (wildwood either as it was before human intervention or as it would be by now had that intervention never happened). It tends to produce uniform expanses of even-aged, densely shading trees: the herbaceous plants of grassland, field edges, etc., disappear without being replaced by woodland herbs. In Mediterranean Europe, where secondary forest often consists of fire-adapted trees, the effect is a recurrent cycle of fires, which destroy whatever escaped the increasing shade. Many conservationists disapprove of goats, but the consequence of removing goats is usually to convert a browsing-dominated landscape into a fire-dominated one (**Figure 14**).

Land-use has become polarized: for example, land is now either forest or pasture but not both. Most of rural Japan, for example, is forest, and the remainder is intensively cultivated rice fields. The forests are now little used, and within them are the remains of many historic ecosystems: coppice woods,



Figure 14 This was a goat-browsed landscape with patches of evergreen oak savanna and forest. It was taken over by foresters, who suppressed goats and planted pines (an ultraflammable tree). The inevitable result was a great fire. *Valor, Alpujarra, Spain.*



Figure 15 An old pollard larch (*Larix leptolepis*) in what is now forest. Its spreading habit proclaims that it was originally a freestanding tree in grassland. *Tateshima, Chino, Japan.*

wood-pastures (**Figure 15**), terraced rice fields, pine savannas, Japanese early-modern forestry, and the growing of the giant grass with which house roofs were thatched.

Forests, on the whole, have become denser as savannas infill, coppice woods grow up, and foresters encourage the growth of timber trees. The older trees in a forest are commonly more spreading than the younger ones, having grown up when they had more room. The author has seen examples in many countries from Canada to New South Wales, often in forests with little obvious human intervention.

An important innovation is the cult of tidiness: the urge among professional planners to destroy old quarries, spoil heaps, ruined buildings, patches of roughland, and even old trees and neglected tombstones, on the grounds that the public might not like them – even though, in practice, derelict land is often a much-loved public amenity.

Implications for Biological Conservation

Some conservationists disdain cultural ecosystems on the grounds that they result from the "degradation" of wilderness ecosystems, especially forest. The only proper business of biological conservation (they say) is to preserve pristine

natural ecosystems where they still exist or to restore them where they do not.

However, this cannot be the whole purpose of conservation. In many parts of the world, such as England, ecosystems have been modified by human activity for much of the Holocene; not enough is known about pristine natural ecosystems to make restoring them a practical objective. Conservation depends on public interest, which it cannot retain if it is limited to remote wildernesses that few of the public ever see. It is unrealistic and impolitic to cut off biological conservation from other kinds of conservation: to separate the archeological interest in an earthwork from the biological interest in the vegetation growing on it. It is essential for students to gain a knowledge of how local ecosystems work before pronouncing on those of distant countries.

Wilderness philosophy often ignores the realities of archeology or vegetation history. In many countries, when a national park is scheduled, the authorities automatically expel the human inhabitants (however long and respectable their history), play down their part in the development of the landscape and ecosystems, and pretend that the park is wilderness. In Yellowstone National Park they even tried to efface all evidence that settlement had ever happened. In reality, few national parks lack human influence altogether; many of them are the last strongholds of precisely those old-fashioned land-uses that biological conservationists ought to support. (The National Trust, biggest landowner in the Lake District National Park in England, goes to much trouble to sustain farming and prevent land abandonment.)

Conservationists are tempted to start from theories of what natural ecosystems ought to be and to recreate them in the image of those theories. They regard large, continuous, mixed-aged forests of tall trees as natural and try to preserve those forests that most nearly approach that ideal. The notion owes more to climax theory and to foresters' notions of an ideal forest than to weighing the evidence of what particular natural forests were like; yet its influence leads to neglect of discontinuous forests, savanna, and forests of short trees, however remarkable.

Biological conservation has many aspects. Priority is rightly given to preventing logging of remote forests that have never been logged, whether or not later research shows that they are really wilderness. However, the real merits of cultural ecosystems should not be overlooked. Wildwood included a great diversity of ecosystems. Those that involved open ground often predominated during glaciations. When giant herbivores were exterminated, this probably made the forests more shady, and the survival of nonshade-bearing plants through interglacials more difficult. Historic land management, to some extent, supplied the place of the missing elephants and restored the habitats of savanna and open ground. Modern changes tend to destroy historic ecosystems altogether or to replace them with uniform, very shady forest. Historic land-use patterns need to be respected and historic cultural ecosystems maintained. Conservationists rightly campaign against destroying trees in tropical rainforests. They often need to campaign against allowing trees to grow in unsuitable places near home.

Acknowledgments

Some of the figures are the result of travels with Mr. A.T. Grove, Dr. D. Lunney, Professor J.A. Moody, Professor Jun-ichi Ogura, Dr. J. Sleath, Professor Kathuyoshi Tsuchida, and the Sphakia (Crete) Survey.

See also: Agriculture, Traditional. Cattle, Sheep, and Goats, Ecological Role of. Ecology of Agriculture. Forest Ecology. Hunter-Gatherer Societies, Ecological Impact of. Indigenous Peoples and Biodiversity. Traditional Conservation Practices

References

Birks HH, Birks HJB, Kaland PE, and Moe D (eds.) (1988) *The Cultural Landscape: Past, Present and Future*. Cambridge, UK: Cambridge University Press.
Buckley GP (1992) *Ecology and Management of Coppice Woodlands*. London: Chapman & Hall.
Cronon W (1983) *Changes in the Land: Indians, Colonists, and the Ecology of New England*. New York: Hill & Wang.
Dargavel J (1995) *Fashioning Australia's Forests*. London: Oxford University Press.
Fairhead J and Leach M (1998) *Reframing Deforestation*. London: Routledge.
Grove AT and Rackham O (2000) *The Nature of Mediterranean Europe*. New Haven, CT: Yale University Press.
Moreno D (1990) *Dal Documento al Terreno*. Bologna: Il Mulino.
Naveh Z and Liebermann A (1994) *Landscape Ecology: Theory and Application*, 2nd edn. New York: Springer.
Parra F (1990) *La Dehesa y el Olivar*. Madrid: Ediciones del Prado.
Rackham O (1986) *The History of the (British and Irish) Countryside*. London: Dent.
Rackham O (1994) *The Illustrated History of the Countryside*. London: Weidenfeld & Nicholson.
Rackham O (2006) *Woodlands*. London: Harper Collins.
Rackham O and Moody J (1996) *The Making of the Cretan Landscape*. Manchester: Manchester University Press.
Vera F (2000) *Grazing Ecology and Forest History*. Wallingford: CABI.
Vos W and Stortelder A (1992) *Vanishing Tuscan Landscapes*. Wageningen: Pudoc.

Large-Scale Biodiversity Experiments

Andrew Hector, University of Zürich, Zurich, Switzerland
Michel Loreau, McGill University, Montreal, QC, Canada

© 2007 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–9, © 2007, Elsevier Inc.

Glossary

Biodiversity A contraction of biological diversity that encompasses all biological variation from the level of genes, through populations, species, and functional groups (and sometimes higher levels such as landscape units).

Ecosystem functioning An umbrella term for the processes operating in an ecosystem.

Ecosystem processes The biogeochemical flows of energy and matter within and between ecosystems, for example, primary production and nutrient cycling.

Ecosystem services An ecosystem process that is beneficial for human beings, for example, the provision of foods and materials, sequestration of carbon dioxide, and stabilization of soils.

Factorial design Involves all possible combinations of the levels of the crossed experimental factors (fully factorial) or an incomplete but informative combination of levels (fractional factorial).

Functional group A group of species thought to have similar impacts on ecosystem functioning (functional effect groups).

Biodiversity Experiments

Biodiversity experiments aim to identify the consequences of changes in diversity for ecosystem functioning and services (as opposed to looking at the covariation of these two factors in comparative analyses of observational datasets). In the early 1990s, ecologists began formulating a set of hypothetical relationships between biodiversity and ecosystem functioning. These hypotheses ranged from those that implied a strong relationship in which the species are complementary and all play an important role, to those that include differing degrees of functional redundancy (species that overlap in functional role) to a null hypothesis of no link. In the early 1990s, these hypotheses were untested and the relationship between biodiversity and ecosystem functioning unknown. The subsequent years have seen a succession of experiments designed to test the effect of biodiversity on ecosystem functioning (Hooper *et al.*, 2005; Kinzig *et al.*, 2002; Loreau *et al.*, 2001, 2002). These experiments have revealed significant effects of biodiversity on ecosystem processes likely to affect the delivery of ecosystem services to human societies.

There are two main ways to manipulate biodiversity. One is by the assembly of synthesized model communities, either in laboratory cultures or in the field using artificial ponds or streams, forest plantations, plots in grasslands and similar experimental systems. An alternative approach is to remove species from natural communities, through weeding and herbiciding of plant species for example (Wardle *et al.*, 1999; Wardle and Zackrisson, 2005). Both approaches have strengths and weaknesses but both approaches attempt to hold other factors as constant as possible while manipulating the diversity of the experimental systems. A third nonmanipulative approach is to infer the relationship between biodiversity and ecosystem functioning by seeing how they are correlated across habitats.

Removal experiments are more familiar as they have a history of use in community ecology to examine interactions between species (principally competition). They are more realistic than assembly experiments in that they are based in real ecosystems. However, increased realism comes at the price of a loss of control through potential unknown confounding factors. The approach also has its own potential limitations through the disturbance involved in species removal and other effects like fertilization through the decay of roots left behind after species removal.

The design and analysis of biodiversity experiments is not straightforward (Schmid *et al.*, 2002). One problem is the substantial levels of diversity present in most systems, which means that the number of possible experimental combinations soon exceeds what is logistically possible and fully factorial designs are only feasible when dealing with very small numbers of species. One way around this is to combine species into a small number of functional groups—species that are expected to have similar effects on ecosystem functioning. However, functional diversity may not fall into discrete groups and objectively identifying these groups is also not straightforward, leading to the development of approaches for looking at continuous measures of functional diversity (Petchey and Gaston, 2006). A second problem is that different aspects of diversity, such as numbers of species and functional groups, are often confounded. The combination of confounded explanatory variables and nonorthogonal designs has made it hard to definitively identify the importance of different aspects of diversity and the mechanisms by which they affect ecosystem functioning.

A further issue for community assembly experiments is how communities should be put together since the diversity gradient formed by the synthesized communities effectively simulates the assumed order of species loss. One approach is to assemble a full community and depauperate versions that simulate a single order of species loss. An alternative approach that provides a more general result is to replicate multiple communities of a given diversity by random selection of

species from the pool available. However, species are not usually lost from ecosystems in random order (although random loss may approximate some real-world cases). Instead, when a given extinction scenario can be predicted, communities can be assembled to simulate it. Alternatively, a fully factorial approach simulates all possible combinations of species but is limited by logistics to small numbers of species (or more often functional groups). Community assembly experiments should give a high level of experimental control but with the usual trade-off with realism. For example, establishment of the assembled communities usually involves removing any existing vegetation and surviving propagules and the disturbance that involves, e.g., soil fumigation. Homogenization of the underlying substrate should reduce the experimental noise, but may also remove environmental heterogeneity that affects the relationship between diversity and functioning.

To date, no one had conducted the same biodiversity manipulation in a given system using both species removal and the assembly of experimental communities and compared the results of the two alternative approaches. Similar results would provide strong support for general effects regardless of experimental methodology. Differences in results could be due to artifacts of the methodology but could also reflect important biological processes. For example, in removal experiments species compensation depends in part on natural levels of colonization of the newly free space, whereas community assembly experiments usually introduce equal numbers of propagules of all species as part of the method.

Ecological experiments can be large scale in different ways as this article illustrates specifically for biodiversity manipulations. Natural experiments may use only a single experimental unit but of a very large size: whole lakes or forest catchments for example. Alternatively, experiments may use a level of replication which is not especially large (a sample size of 30 for example) but where individual experimental units are of a size large for the study system (e.g., fields-scale trials with genetically modified crop species). Other large-scale projects use experimental units of conventional size but in complex and highly replicated designs. Multisite studies can be large if they combine experiments from several locations into a single coordinated project. Long-term studies are relatively unusual but can be of great value, and are large scale in a temporal context.

Mechanisms

Ecological theory identifies two classes of underlying mechanism that can generate impacts on ecosystem processes when diversity changes. Species influence ecosystem processes partly through their intrinsic properties. Whether a species with relatively extreme traits is present or absent from a system is one simple mechanism for impacts of biodiversity change. When species overlap in their ecological traits they are said to be redundant. When redundant species can compensate for the loss of similar species within the same group there will be little if any change in functioning. The redundancy can nevertheless be seen to have value as an ecological insurance. Conversely, when species occupy clearly distinct and complementary niches it will be impossible for full compensation to occur. The degree of niche overlap should control the shape of the relationship between increasing diversity and ecosystem processes; low niche overlap will produce linear patterns while high overlap will lead to saturating curves as the addition of new species increases redundancy.

The interpretation of biodiversity experiments is complicated by these alternative classes of mechanism. Selection effects (Loreau and Hector, 2001) occur when communities become dominated by particular species and if dominant species occur more frequently in high-diversity communities then the traits of these species may drive ecosystem functioning. Alternatively, complementary or positive interactions may allow diverse systems to utilize more resources or to suffer lower levels of pests and diseases. An additive partitioning method (Loreau and Hector, 2001) was developed as part of the BIODEPTH project (see below) to distinguish between these alternative mechanisms. The additive partitioning method (Loreau and Hector, 2001) defines a net effect of biodiversity which is the difference between the observed yield of a mixture and that of the average monoculture yield. This net effect is then partitioned into two additive components: the selection and complementarity effects. The selection effect is the standard statistical measure of covariance applied to the relationship between yield in monoculture and the change in relative yield in mixture. Selection effects will be positive when species with higher-than-average yield dominate communities and negative when species with lower-than-average yield dominate. The complementarity effect uses relative yields to ask whether increases in the abundance of some species exactly cancel with declines in others. When this is the case, resource partitioning is a zero-sum game with some species taking more of a fixed total pool of resources and others taking less. Positive complementarity effects occur when decreases in the abundances of some species do not compensate for the increases in the abundance of other species, which could result from facilitation, resource partitioning, or decreased impact of natural enemies in more diverse communities. The additive partitioning method has revealed widespread complementarity contributing to most published results. However, unanticipated negative selection effects have also emerged as a widespread result. One consequence is that complementarity effects are often masked by the negative selection.

Community Assembly

Historical Precedents

The intellectual linkage of biodiversity with ecosystem functioning can be traced back to the *Origin of Species* where Darwin describes what is arguably one of the first ecological experiments: a large-scale experimental garden comparing the properties of plant monocultures and mixtures (Hector and Hooper, 2002). Interestingly, the results prefigure those of recent experiments in associating higher levels of diversity with higher levels of biomass production (**Figure 1**). However, this early work inevitably lacks features of modern experimental design such as replication (in this experiment diversity was confounded with the origin of the experimental

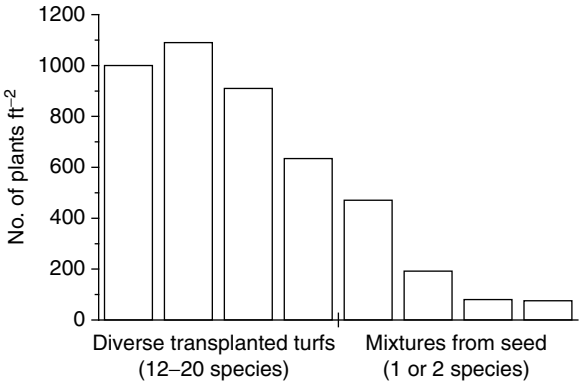


Figure 1 Recently rediscovered data dating to the early nineteenth century which was collected from a large-scale experimental garden at Woburn Abbey, UK has many similarities to modern biodiversity experiments and is arguably some of the earliest experimental work conducted in ecology. Graph by A. Hector.

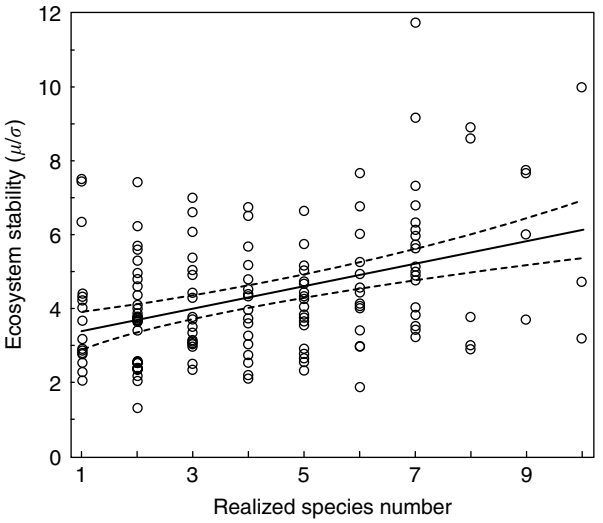


Figure 2 Greater temporal stability of biomass production in higher diversity communities from the first decade of the Cedar Creek prairie grassland experiment. Reproduced from Tilman D, Reich PB, and Knops J (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441: 629–632, with permission from Nature Publishing Group.

communities; that is whether they were transplanted natural turfs or established from seed).

The Ecotron Large-Scale Controlled Environment Facility

The first modern biodiversity experiment was carried out with model communities based on annual plant species grown in the Ecotron large-scale controlled environment facility (Naeem *et al.*, 1994). Most first-generation biodiversity experiments concentrated on a single trophic level. The Ecotron experiment was unusual in taking a single intact community and simultaneously reducing diversity at four trophic levels to produce two increasingly depauperate versions from which species had been omitted at random. The key result of this experiment was that the impoverished communities were progressively less productive. However, because all replicates at each diversity level were identical in composition, the effects of numbers and types of species were confounded and the results may be specific to the particular order of species extinction examined in this experiment and not general across a wider range of possible extinction scenarios.

Large-Scale Field Experiments

The first large-scale biodiversity experiments performed under field conditions were a series of studies by Tilman and colleagues working at Cedar Creek, Minnesota prairie grassland. A pair of biodiversity experiments concentrated on species (Tilman *et al.*, 1996) and functional group effects (Tilman *et al.*, 1997, 2001, 2006), respectively while the Biocon experiment (Reich *et al.*, 2001, 2004) looks at the interactions between biodiversity loss and elevated CO_2 and nitrogen. In contrast to the Ecotron experiment, more extensive diversity gradients were established where each level of diversity (a given number of species) was replicated with different mixtures of species selected at random from the species pool. Biodiversity effects could then be quantified with linear regression and tested against the residual differences between different composition communities within diversity levels. Productivity was positively related to numbers of species and functional groups in the communities with relationships growing stronger over several years. Levels of unconsumed soil nitrate and (potentially leachable) nitrate below the rooting zone were both reduced at higher diversity levels. Detecting how many species contribute to biodiversity relationships has proved one of the most contentious issues in interpreting biodiversity experiments. Tilman *et al.* (2001) used a diversity index approach to see how many of the most productive species in a plot had to be present to best explain their aboveground and total biomass production. For these two ecosystem processes between one quarter and three quarters, respectively, of the species in the high-diversity treatment were needed to best explain the productivity of a plot. Additional analyses of species-specific contributions suggest that much of the biodiversity effects can be explained by legumes and C_4 grasses but with additional species coexisting alongside them and contributing to total productivity (Lambers *et al.*, 2004). In contrast, in the Biocon experiment, which examined a biodiversity gradient under different conditions of elevated CO_2 and N enrichment, species and functional group richness had largely independent effects across the whole range of conditions such that species within groups were not functionally redundant but made separate contributions (Reich *et al.*, 2004).

A decade of monitoring of the Cedar Creek species and functional group diversity experiment confirmed the insurance hypothesis which predicts that biodiversity promotes greater temporal stability of ecosystem functioning (**Figure 2**). The highest diversity plots had a temporal stability measure (mean/standard deviation for a given time period) for biomass production that was 70% greater than the average of the monocultures. Greater temporal stability was generated by the portfolio effect (statistical averaging of asynchronous population fluctuations) and overyielding (where the yield of a mixture exceeds the weighted average of the constituent monocultures). Temporal stability at the ecosystem level was

positively correlated with high root biomass but, interestingly, negatively correlated with the presence of legumes (even though they promote overyielding in this system).

The Jena Biodiversity Experiment (Roscher *et al.*, 2004) simultaneously manipulates numbers of both species and functional groups in a randomized design but in a more balanced way than earlier experiments and in small (3.5×3.5 m) and large (20×20 m) plots (**Figure 3**). The project also contains a dominance experiment comprising plots composed only from a reduced pool of nine species predicted to be dominant (high relative abundance) in the full community.

Initial results from the Jena main experiment show that relationships from small plots can be scaled up to the large plots: results from the small and large plots were statistically indistinguishable. Productivity was positively linearly related to the log of the number of plant species. The slope of the relationship between diversity and productivity in the dominance experiment was statistically parallel to that from the fully randomized experiment but with a higher elevation (**Figure 4**). This results presumably because the average biomass production per species in the dominance experiment is higher than in the main experiment where species were chosen at random from the species pool which contains many species with low maximum aboveground biomass. Comprehensive results on the role of plant functional groups are yet to be published at the time of writing (summer 2006).

Multisite Biodiversity Experiments

The European BIODEPTH project (Hector *et al.*, 1999; Spehn *et al.*, 2005) added two new features to the single-site experiments described above. First, in addition to replicating diversity levels with different species compositions, each composition was itself replicated which allowed the effects of both diversity and composition to be quantified. Second, the same biodiversity experiment was replicated at eight different grassland sites using standardized methodologies. Individual site analyses illustrate the broader range of responses seen in the literature—from null to positive—and the combined multisite analysis revealed an overall log-linear (asymptotic with increasing diversity) relationship that was statistically common to all sites. Reconciling the results of the individual site analyses with the general outcome of the overall analysis has been a point of both confusion and contention. The results of the BIODEPTH project also revealed that productivity could be well explained in terms of the multiple influences of location ($\sim$30% of the total sums of squares), the diversity of both species and functional groups ($\sim$20%), and composition ($\sim$40%). Later results revealed that the effects of biodiversity were not limited only to aboveground biomass production but extended to a wider suite of ecosystem processes and properties including root biomass, levels of intercepted light, nitrogen pools in aboveground vegetation, and available soil nitrogen as well as the decomposition of various substrates. However, the relationship of some of these variables to biodiversity was often more complex and less clear than that of aboveground biomass production (Spehn *et al.*, 2005).

Other multisite experiments have been performed in terrestrial systems (e.g., Van der Putten *et al.*, 2000) but relatively few in aquatic and marine systems. A notable exception is the study by Emmerson *et al.* (2001) who examined the relationship between biodiversity and ecosystem functioning at multiple sites in the United Kingdom, Sweden, and South Australia. They looked at the relationship between the diversity of marine benthic faunal communities (ranging between 1–4 dominant species and 1–3 functional groups) and the production of nitrogen in ammonia form (NH_4–N) as an indicator of ecosystem functioning. Nitrogen production increased with community total biomass but for a given biomass production varied depending on the make up of the community. However, alongside this variation there were positive trends for greater predictability in ecosystem functioning at higher levels of diversity and a greater degree of complementarity between species in more diverse mixtures. In light of



Figure 3 The Jena Biodiversity Experiment manipulates numbers of species and functional groups in small (3.5×3.5 m) and large (20×20 m) plots. Photo by A. Klein.

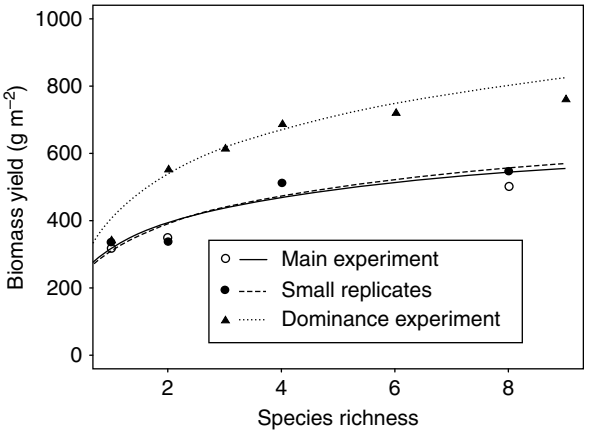


Figure 4 The relationship between diversity and biomass production in the second year of the Jena Biodiversity Experiment. Results from the small and large plots are statistically indistinguishable and the relationship in the dominance experiment has the same slope but with a higher elevation. Graph by J. Schumacher.

this combination of species identity and species number effects Emmerson *et al.* characterized the relationship between biodiversity and ecosystem functioning in these marine systems as positive but idiosyncratic.

Highly Replicated Laboratory Experiments with Microbial Microcosms

In response to the complexities in the analysis and interpretation of biodiversity experiments, Bell *et al.* (2005) used a highly replicated laboratory experiment with a novel design to test whether the relationship between biodiversity and ecosystem functioning found for terrestrial plant communities extended to microbial microcosms. They took a pool of 72 laboratory culturable bacterial species isolated from the communities that develop in tree holes (the pools of water that collect between the exposed roots of trees at the base of their trunks). They took all the factors of 72 as their diversity levels (1, 2, 3, 4, 6, 8, 9, 12, 18, 24, 36, and 72) and replicated each level by drawing species from the pool without replacement. Five independent diversity gradients were formed in this way to give five alternative partitions of the species pool. Finally, each particular partition was repeated twice to provide compositional replication (a feature first introduced by the BIODEPTH experiment; see below). In total, the experiment required 1374 microbial microcosms. This novel approach resulted in an analysis where the log-linear effect of species richness (numbers of species on a log-scale) was orthogonal to the effect of species composition (the collective effects of 72 variables coding for the presence or absence of each species in each microcosm). The main result was a log-linear relationship between diversity and community respiration (**Figure 5**), which was similar to the relationship between diversity and biomass production in experiments with plants. The independence of the log-linear diversity effect from the composition effect, combined with only a small number of strong individual species effects suggests that the result was not produced by dominance of the microbial communities by species with high respiration rates but more likely resulted from complementary or facilitatory relationships or a combination of the two.

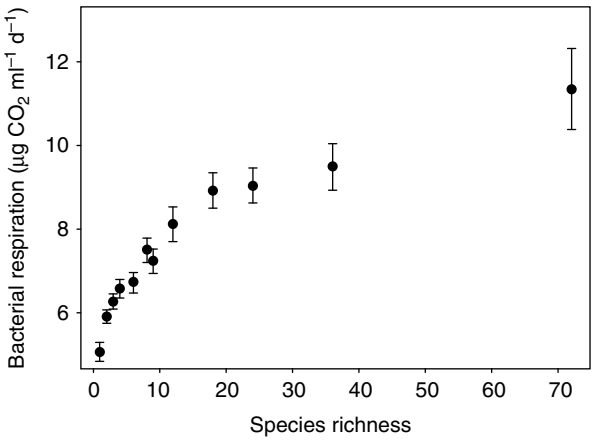


Figure 5 The linear relationship between the diversity of laboratory microbial microcosms and respiration rate (mean ± 1 S.E.M.). Graph by T. Bell and J. Newman.

Species Removal

Wardle and Zackrisson (2005) used a series of islands in a Swedish lake as a study system to examine influences on ecosystem functioning. The island system allowed them to look at patterns when comparing across the habitats (islands) as well as to perform removal experiments within individual systems. These islands show varying diversities and ecosystem process rates due to their size and likelihood of disturbance. Large islands are more prone to fires caused by lightning strike and therefore experience higher levels of disturbance and have lower diversity (species richness is the same on all islands but diversity varies once differences in relative abundance, or "evenness," is taken into account). Note that this relationship is contrary to the predictions of classical island biogeographic theory, which ignores many processes like disturbance. Productivity is greatest on the large, more frequently disturbed islands, and therefore negatively related to diversity (**Figure 6a**). The smaller, low-productivity islands tend to be dominated by "stress-tolerant" species (sensu Grime, 2001) and have lower levels of decomposition and related processes as well as lower productivity. Wardle and Zackrisson (2005) supplemented the across-island comparison by performing manipulative experiments to remove species and functional groups from a subset of the islands of differing sizes. Despite the negative correlation between diversity and process rates across islands, the removal experiments showed significant impacts of the removal of three species of shrub on productivity (**Figure 6b**). Two of the species showed larger removal effects on islands where they were more abundant. This study demonstrates scale-dependence of the relationship between diversity and functioning and how manipulative experiments can reveal diversity effects within systems (positive in this case) which can differ from the larger-scale across-site correlation (which was negative in this study).

In an earlier removal experiment that manipulated the functional group composition of vegetation gaps in a New Zealand pasture grassland, Wardle *et al.* (1999) found some impacts on both above- and belowground processes. However, whether an effect was found and how strong it was depended on the traits of the species that were experimentally excluded. Similarly, in the same Cedar Creek prairie that hosted the community assembly experiments reported above, removal of plant functional groups also negatively affected above- and belowground biomass, soil nitrogen, and drought resistance but where the effects related to which group was removed and to the response of the remaining functional groups (Symstad and Tilman, 2001). Removal of plant functional groups often also facilitated invasion by other species (Symstad, 2000). Many of the results of this and other removal experiments are related to recruitment limitation, a process that is not as evident in community assembly experiments owing to the seed addition used to establish communities.

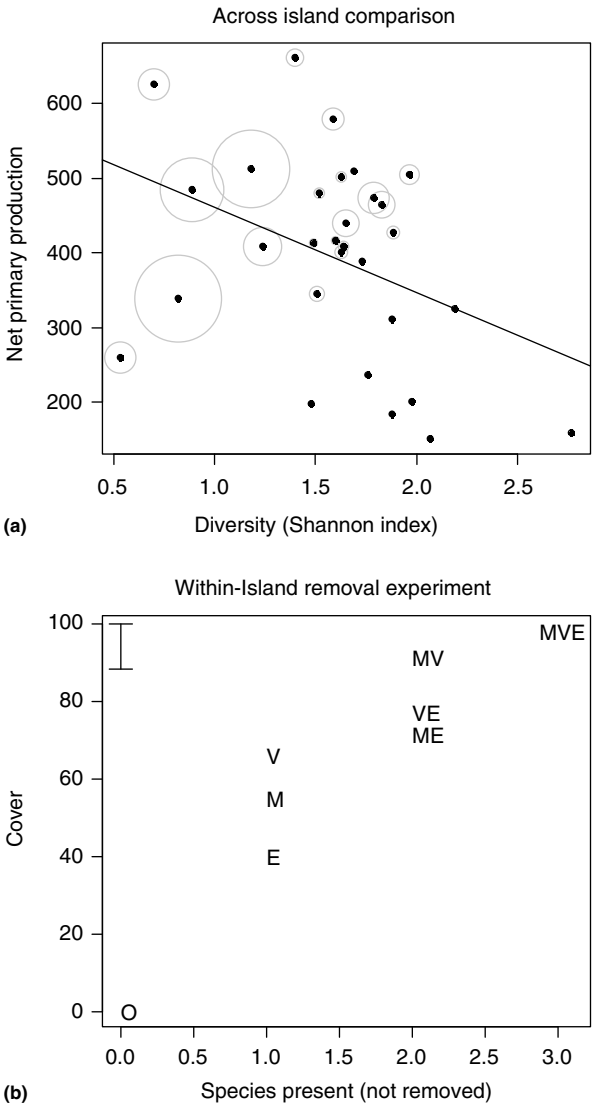


Figure 6 (a) Negative relationship between diversity and productivity across islands of differing sizes (island size is proportional to the size of the gray circles around each symbol). Data provided by D. Wardle. (b) Positive relationship between diversity and productivity (% cover) within islands demonstrated by the experimental removal of combinations of *Vaccinium myrtilis* (M), *V. vitis-idaea* (V), or *Empetrum hermaphroditum* (E) or all vegetation (O). Reproduced from Wardle DA and Zackrisson O (2005) Effects of species and functional group loss on island ecosystem properties. *Nature* 435: 806–881.

Utilizing Biodiversity in Applied Settings

The large-scale biodiversity experiments described above all come from pure ecology and aim to identify generalities in the relationship between biodiversity and ecosystem functioning in various types of ecosystem. However, in the real world, biodiversity is lost due to specific processes (eutrophication, fragmentation, and overharvesting) and in some cases pure studies may not give general predictions which apply to particular real-world cases. A small number of applied experiments have been established to see whether biodiversity can be used in the management of habitats in a way which is beneficial to the provision of ecosystem services to humans.

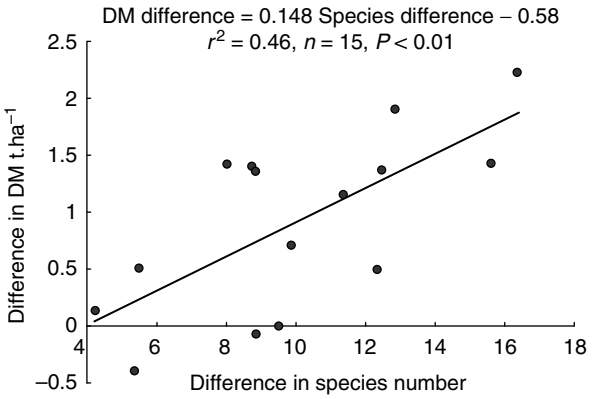


Figure 7 Increases in hay production through the augmentation of agriculturally recommended seed mixtures with additional species. Reproduced from Bullock JM, Pywell RF, Burke MJW, and Walker KJ (2001) Restoration of biodiversity enhances agricultural production. *Ecol. Lett.* 4: 185–189.

Grassland Restoration

Bullock *et al.* (2001) performed a biodiversity experiment on the restoration of hay meadows at seven locations across southern Britain. At each site, plots were sown with "low-diversity" treatments (6–17 species depending on the site) which consisted of the seed mixtures recommended by the (then) UK Ministry for Agriculture Fisheries and Food (MAFF) for the recreation of moderately diverse grassland on land taken out of agricultural production. The recommended seed mixtures were all composed of species from the relevant regional species pool, which were also appropriate for the type of local environment to be restored. The "high-diversity" treatments (25–41 species) were supplemented with species found in diverse hay meadows typical of the region, soils, and hydrology of each site. The high-diversity mixtures therefore contained species appropriate for each site but which had been omitted from the recommended mixture and which were therefore presumably thought to be redundant for restoration purposes. The key result was a linear relationship between hay yield and the number of additional species added to the high-diversity treatments which was consistent across all sites and present for three years after the initial establishment season (**Figure 7**). The selection of a wider diversity of species increased yield by up to 60%. For all the seven sites, the seed mixtures thought to be sufficient for restoration aims led to levels of hay production which were far below those achievable at the same site by more diverse mixtures that contained a wider selection of appropriate species. The increase in yield seemed not to be due to sampling or selection effects for one or a few dominant species although rigorous testing of the underlying mechanisms was not possible without monocultures of all species involved.

Tree Diversity Experiments

Forest ecosystems, particularly in the tropics, are important both for the conservation of biodiversity and the provision

of important ecosystem services but little is known about the relationship between forest biodiversity and ecosystem functioning. Forest ecosystems around the globe that have previously been clear felled or selectively logged are being replanted providing an opportunity to test the importance of restoring levels of biodiversity for the functioning of the rehabilitated ecosystems. There is now an informal network of forest biodiversity experiments that currently ranges from the neotropics (Panama) through Europe (Finland, Germany) to the paleotropics (Malaysian Borneo) as described by Scherer-Lorenzen *et al*. (2005). These experiments signal a new direction in research on biodiversity and ecosystem functioning which moves beyond general tests to large-scale trials in threatened ecosystems that are known to provide important ecosystem services to mankind.

See also: Biodiversity, Definition of. Biodiversity Generation, Overview. Biodiversity, Origin of. Complementarity. Functional Diversity. Functional Groups

References

Bell T, Newman JA, Silverman BW, Turner SL, and Lilley AK (2005) The contribution of species richness and composition to bacterial services. *Science* 436: 1157–1160.

Bullock JM, Pywell RF, Burke MJW, and Walker KJ (2001) Restoration of biodiversity enhances agricultural production. *Ecol. Lett.* 4: 185–189.

Emmerson MC, Solan M, Emes C, Peterson DM, and Raffaelli D (2001) Consistent patterns and the idiosyncratic effects of biodiversity in marine ecosystems. *Nature* 411: 73–77.

Grime JP (2001) *Plant Strategies, Vegetation Processes and Ecosystem Properties*. New York: Wiley.

Hector A and Hooper RE (2002) Darwin and the first ecological experiment, Science 295, 639–640 (table: http://www.sciencemag.org/cgi/content/full/295/5555/639/DC1, accessed November 2006).

Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

Hooper DU, Chapin FSI, Ewel JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge and needs for future research. *Ecol. Monogr.* 75: 3–36.

Kinzig A, Tilman D, and Pacala S (eds.) (2002) *The Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions*. Princeton: Princeton University Press.

Lambers JHR, Harpole WS, Tilman D, Knops J, and Reich PB (2004) Mechanisms responsible for the positive diversity-productivity relationship in Minnesota grasslands. *Ecol. Lett.* 7: 661–668.

Loreau M and Hector A (2001) Partitioning selection and complementarity in biodiversity experiments. *Nature* 412: 72–76. (erratum: 413, 548).

Loreau M, Naeem S, Inchausti P, *et al.* (2001) Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science* 294: 804–809.

Loreau M, Naeem S, and Inchausti P (eds.) (2002) *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*. Oxford: Oxford University Press.

Naeem S, Thompson LJ, Lawler SP, Lawton JH, and Woodfin RM (1994) Declining biodiversity can alter the performance of ecosystems. *Nature* 368: 734–737.

Petchey O and Gaston K (2006) Functional diversity: Back to basics and looking forward. *Ecol. Lett.* 9: 741–758.

Reich PB, Knops JMH, Tilman D, *et al.* (2001) Plant diversity enhances ecosystem responses to elevated CO_2 and nitrogen deposition. *Nature* 410: 809–812.

Reich PB, Tilman D, Naeem S, *et al.* (2004) Species and functional group diversity independently influence biomass accumulation and its response to CO_2 and N. *Proc. Natl. Acad. Sci. USA* 101: 10101–10106.

Roscher C, Schumacher J, Baade J, *et al.* (2004) The role of biodiversity for element cycling and trophic interactions: An experimental approach in a grassland community. *Basic Appl. Ecol.* 5: 107.

Scherer-Lorenzen M, Potvin C, Koricheva J, *et al.* (2005) The design of experimental tree plantations for functional biodiversity research. In: Scherer-Lorenzen M, Körner C, and Schulze E-D (eds.) *The Functional Significance of Forest Diversity (Ecological Studies 174)*. Berlin: Springer.

Schmid B, Hector A, Huston MA, *et al.* (2002) The design and analysis of biodiversity experiments. In: Loreau M, Naeem S, and Inchausti P (eds.) *Biodiversity and Ecosystem Functioning*, pp. 61–78. Oxford: Oxford University Press.

Spehn EM, Hector A, Joshi J, *et al.* (2005) Ecosystem effects of the manipulation of plant diversity in European grasslands. *Ecol. Monogr.* 75: 37–63.

Symstad A (2000) A test of the effects of functional group richness and composition on grassland invasibility. *Ecology* 81: 99–110.

Symstad AJ and Tilman D (2001) Diversity loss, recruitment limitation, and ecosystem functioning: Lessons learned from a removal experiment. *Oikos* 92: 424–435.

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, and Siemann E (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277: 1300–1302.

Tilman D, Reich PB, and Knops J (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441: 629–632.

Tilman D, Reich PB, Knops JMH, Wedin D, Mielke T, and Lehman C (2001) Diversity and productivity in a long-term grassland experiment. *Science* 294: 843–845.

Tilman D, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379: 718–720.

Van der Putten WH, Mortimer SR, Hedlund K, *et al.* (2000) Plant species diversity as a driver of early succession in abandoned fields: A multi-site approach. *Oecologia* 124: 91–99.

Wardle DA, Bonner KI, Barker GM, *et al.* (1999) Plant removals in perennial grassland: Vegetation dynamics, decomposers, soil biodiversity, and ecosystem properties. *Ecol. Monogr.* 69: 535–568.

Wardle DA and Zackrisson O (2005) Effects of species and functional group loss on island ecosystem properties. *Nature* 435: 806–881.

Latent Extinction—The Living Dead

Daniel H Janzen, University of Pennsylvania, Philadelphia, PA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 689–699, © 2001, Elsevier Inc.

Glossary

Agroscape The agricultural, ranching, and plantation countryside, with its roads, irrigation ditches, buildings, and so on. The agroscape stands in contrast to the wildland countryside that is not directly managed by humanity (though it is strongly impacted by it). The agroscape intergrades with wildlands in the form of woodlots, abandoned fields, poor soil sites, hedgerows, and edges of wildlands.

Living dead An individual stripped of the ecological circumstances that allow it to be a reproductive member of its population, but which is living out its physiological life. Living dead are most easily observed as large trees remaining on the agroscape, but they are also present in natural ecosystems.

Megafauna Large mammals that are wolf-sized, deer-sized, and larger. Commonly used in reference to the many species of extinct "Pleistocene megafauna" that 9000 years ago populated the New World. The elimination of this megafauna by hunting (of the herbivores) and starvation (of the herbivore deprived carnivores) was probably the first, and certainly the most dramatically irreversible, of the anthropogenic macroalterations of New World ecosystems. Today, of the extinct Pleistocene megafauna, only the horse remains—evolutionarily invented in the New World but surviving in the Old World until brought back as a gift from the Pleistocene by Spanish soldiers.

Introduction

The idea of the living dead has gradually emerged in my ecological understanding as I have lived past and around the majestic forest giants left standing as the agroscape creeps into Costa Rica's forest ecosystems over the past 4 decades (**Figure 1** and Janzen, 1986a, 1986b). This creep gradually converts the forest to an agroscape of pastures, fields, and roadsides dotted with the occasional adult tree but few or no juveniles. This is an agroscape where a magnificent flower crop now stands bee-less, an agroscape where fruit crops lie rotting below the pasture tree, an agroscape where tree seedlings wither in the dry-season sun or are turned to smoke in the dry-season anthropogenic fires.

I begin this article with a focus on adult large trees and use familiar examples from the Costa Rican countryside. To create breadth, I suggest that you join these verbs with the nouns from the ecosystems you know. This is a conservation biology question, but it applies to more than that, and it applies across the once-forested tropics as well as elsewhere.

Looking across the tropical landscape, the eye is greeted by stately single trees (**Figure 2**), by patches of forest, by the blaze of a colorful flowering episode. Put an inventory to the plant species in a field, in a valley, in an ecosystem. All these species appear in the list. All is more or less well, we conclude, as 96.4% of the species that were here 50 years ago are still present. But are they? How many of them are living dead, part and parcel of latent extinctions?

We live a perceptual lie as we bustle about our agroscapes. That single stately green *Dipteryx* or *Hymenaea* or *Swietenia* or *Enterolobium*, standing in a field, pasture, or roadside, is often just as dead as if it were a log in the litter or the back of a logging truck. That tree was birthed in some favorable circumstance, a circumstance for pollination, seed dispersal, seed germination, and sapling survival.

But one or more of these circumstances is now gone. It was carried away with the forest, put on the hunter's table, pesticided out of existence, or global warmed into oblivion. The long-lived tough adult lives out its physiological life, in the absence of the carpenter with a chain saw, but it is evolutionarily dead. Its pollen no longer flows to other members of the population, its seeds are no longer carried away from seed predators, or its seeds are no longer carried to a favorable site for seedling growth and sapling survival to adulthood.

But because the adult lives on, we are lulled into thinking that the environmental damage really is not all that bad, that extinction has not already occurred. If we can still show the tree to our children, it seems not to be extinct. It is so big and green and strong. Every year we see its flowers, and maybe we even see its fruits on the ground below. And after all, it has clearly weathered all that we have thrown at it. What ever can the matter be?

Humanity's interaction with the world's ecosystems has an enormous perceptual element. We act on what we perceive, be it threat or opportunity. Much of our conservation pragmatics and understanding is based on our knowledge that we really are losing species, losing ecosystems, losing the capacity of the environment to absorb our footprints. But that knowledge comes from what we see and measure. If all members of a tree species were to have the trait that each abruptly falls over dead the moment that it ceases to be a reproductive member of its population in its ecosystem, there would be far stronger alarm cries across the tropics about extinction rates and realities. If trees, the largest organisms on most of our landscapes, were very short lived as compared with humans, there would be less of a perceptual problem—though just as large a conservation problem.

When the terrestrial world was covered with forest ecosystems, the single tree left standing in an aboriginal cornfield

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00085-X



Figure 1 Living dead trees isolated in pasture at the edge of the agroscape (background) as it creeps into old growth forest (foreground). Los Naranjos, Sector Cacao, Area de Conservación Guanacaste, July 29, 1987.



Figure 2 A living dead *Terminalia* tree stands in silhouette, left behind as the rain forest was cleared around it, the natural tree falls in which its seedlings might have survived long since removed. Rincon Rainforest, Area de Conservación Guanacaste, January 6, 2000.

may well have been living dead, but the population from which it was derived was not usually at risk of anthropogenic extinction, unless perhaps dependent on a seed disperser targeted by that aboriginal population (Janzen and Martin, 1982). But when the agroscape is dotted with living dead in the wake of contemporary omnipresent ecosystem alteration, latent extinction is very real. A tree species may be ranked as "common"—meaning visible from a car window along many roads—yet be effectively extinct in a county, state, or region. And since the agroscape now stretches from horizon to horizon, the plant may well be absolutely extinct, since all of its former range may be populated by living dead.

Deforestation and the Living Dead

The forest need not be removed to convert trees to living dead. It is just that when the forest is partly removed, there is a very high chance that this alone will ecologically deprive many individuals of the remaining tree species sufficiently to convert them to living dead status. And, it certainly leaves the living dead very visible.

But even when the forest is left in place, that is no guarantee of a healthy tree population. When the Pleistocene hunters and their carnivorous helpers hunted out the neotropical mastodons and gomphotheres, the glyptodonts and camels, the ground sloths (Janzen, 1983b; Janzen and Martin, 1982), they did not do it by forest clearing. For decades to millennia after this 9000-year-old event, many of the remnant individuals of the tree populations that these big mammals dispersed (**Figure 3**), and for which they created safe sites for seedlings by their browsing and trampling, would have been living dead scattered in the forest.

If some particular species—a pollinator or dispersal agent, for example—in the forest is extinguished, by whatever cause, there will often be surrogates and alternates that will assume, in some form, some portion of the "role" of the extinguished mutualist. The tree species will live on, albeit in some other ecological morph, and therefore in some technical sense will not be extinct. The tree that was "dependent" on the

extinguished species will not, then, be living dead. But the devil is in the details. We need to go case by case. The suite of interactants with a tree species generates a given seed shadow, pollen rain, sapling demography, and microgeographic distribution. Remove one species of interactant. The entire n-dimensional hyperspace shifts in this or that direction. In some places this is toward eventual extinction, in other places it is just a change in demography and microgeographic distribution.

The history of any surviving species is that it must have survived thousands of such handoffs from one mutualist to another, from one moment to the next (e.g., Hallwachs, 1986). What bumps individuals into the category of living dead is the serendipitous event of losing irreplaceable partners. Humanity has a way of removing not only partners, but whole suites of them, as well as altering the physical environment. Our thoroughness and omnipresence creates ecological irreplaceability. Yes, when we lose one ground sloth, a glyptodont picks up some of the slack, though the tree is now a different beast. And at some time, likely as not, some new slothoid arrives by evolution or immigration over the millennia. But lose all these big mammals at once, and the result is guaranteed to be large arrays of living dead.

We have all been nourished by the marvels of evolutionary understanding, leading to the temptation to wonder if rapid evolution will not resuscitate a living dead population, if not many of its individuals. Novel pollinators, dispersal agents, fruit morphology, flowering phenology—all could save the day. In theory yes, but in reality not on the timescales ordained by humanity's charge across the landscape. How long will it take to evolutionarily reinvent a neotropical herbivorous/frugivorous megafauna? Fracture the remaining forest, with its living dead, into small ecological islands (also known as national parks and reserves). Thereby create ideal circumstances for rapid and novel evolution. We still cannot expect natural selection to create a mastodon from a white-tailed deer in anything like the speed required to be an antidote for neotropical rain forest anthropogenic alteration, beginning with the megafauna extinctions.

Certain kinds of habitat destruction are compatible with some tree natural histories. Two common trees, the guanacaste (*Enterolobium cyclocarpum*, Fabaceae) and jicaro (*Crescentia alata*, Bignoniaceae), owe their contemporary prominence on the Mesoamerican Pacific coastal landscape to a particular kind of habitat destruction. For both, large mammals—such as free-ranging horses—swallow the seeds while eating the content of indehiscent fruits fallen below the parent tree (**Figures 3** and **4**), and later defecate them in open sunny habitats (Janzen, 1981, 1982a, 1982b). Forest clearing unto brushy pastures and scraggly roadsides, populated by widely circulating working horses, maintains a healthy population of reproducing guanacaste and jicaro trees in a precarious balance with humanity.

What did these trees do before the Spaniards brought the horse back from its Old World refuge after its neotropical extinction by Pleistocene hunters (Janzen and Martin, 1983)? They probably survived in a peculiar habitat characterized by ample insolated ground yet sufficient rain for there to be large trees and sloppy seed predator rodents (or human fruit and seed harvesters), which offered sufficient seed dispersal. River edges, marsh edges, and the interface between tropical dry forest and desert are such habitats, and the aboriginal village/field edge adds a serendipitous fourth. The Spanish working horse (**Figure 4**) found the fruits abandoned by their extinguished ancestors and spread these two trees so thoroughly that today they are viewed by Mesoamerican societies as native and natural. And, in the case of *Enterolobium cyclocarpum*, cattle are surrogate horses (Janzen, 1982a).

However, as the motorbike and car replace the horse today, and as the cattle industry fades, these two trees are left as very visible living dead scattered across the former ranch lands, their abundant fruits rotting below the parent tree, the newly



Figure 3 A living dead *Crescentia alata* fruit crop presented to earthbound extinct megafauna (**Figure 4**). Sector Poco Sol, Area de Conservación Guanacaste, May 28, 1988.



Figure 4 An earthbound extinct megafauna returned from the Costa Rican Pleistocene by Spanish immigrants, breaking a *Crescentia alata* fruit (**Figure 3**) to eat the molasses and seeds inside. Sector Santa Rosa, Area de Conservación Guanacaste, 1980.

germinated seedlings killed by fungal pathogens nourished by the annually replenished seed crop, and the rare escaped seedling killed by herbicides, grass fires, and cosmetic cleansing.

When is a Tree not Living Dead?

Earlier I noted that if each member of a tree species were to abruptly fall over dead the moment that it ceases to be a reproductive member of its population in its ecosystem, there would be far stronger alarm cries across the tropics about extinction rates and realities.

However, the isolated tree, left an adult in the open as the forest is mined away from around it (**Figure 2**), is not necessarily or automatically a member of the living dead, or at least not necessarily at that moment. At least two circumstances may help to avoid this label. First, the pollinator community and the seed dispersal community for that tree may still be of a structure such that they confer sufficient amounts and patterns of their services and do so with the new reproductive phenology that will be expressed by the tree in its "new" habitat. And males do have fitness. A plant may never set a fruit or never have a surviving seedling from its seed crop, yet it still may be very much a member of the reproducing population (e.g., Aldrich and Hamrick, 1998). Plants contribute pollen "outward" as well as receive it from unseen members of the population. There may be some circumstances where this or that member of the pollinator guild will in fact carry pollen from that isolated tree back into the forest. At least potentially this may remove the living dead label.

Second, the new pattern of seed/seedling/sapling safe sites for that species may be sufficient for population survival, even if different. A novel demography, reproductive phenology, and microgeographic structure will ecologically emerge, reflecting the serendipitous matching of the tree's traits to these new conditions.

For the survivor, such ecological fitting (Janzen, 1985) of an individual (or a population) into the environment newly thrust upon it is the same process as occurs when a tree species is anthropogenically introduced to a new place. Whether introduced by humans or by natural processes, its survival there demonstrates that it has ecologically fit in. Such introduction may occur into a natural ecosystem or one variously anthropogenically perturbed. Sloppy deforestation may create many living dead, only mildly impact some other species, and favor yet new introductions into the region by having removed competitors or consumers.

A population of plants in a newly altered landscape is not necessarily at a given moment either "dead" or "alive." Just as the relationships of an individual to its ecological circumstances may decay slowly, it is also easy to visualize a population being sufficiently anthropogenically impacted that it gradually decays over several decades-to-centuries-long generations. This state of decay is an intermediate between living dead and "normal surviving." The portion of a population of trees at some geographic point may be in a constant state of swinging between being "okay" and living dead, as its associated climate and community of interactors goes through their own changes.

A species' population in its totality may also be waxing or waning in geographic coverage, density, "living deadness," or all three. Living dead are found at the geographic or demographic margins of all populations. It is just that human activity in ecosystem modification (elimination, simplification) simultaneously impacts so many species, and is so omnipresent, that it creates large numbers of living dead in the same place at the same time. These then carry the tragic perceptual load of tricking us into thinking that all is much more well than it actually is.

But ecological neutering, expressed as here in the terms "living dead" or "latent extinctions," is not restricted to the circumstance of the single tree in the field or a single portion of a population. The living dead are an integral part of natural age-structured mortality. Any field biologist can identify a large number of young individuals—seeds, seedlings, saplings—that have a vanishingly small chance of survival as individuals. The forest understory is densely populated with them, as is each squirrel's winter seed cache, as is the patch of seedlings below the healthy parent tree, as is the ground covered with ephiphyte seeds that fell past the branches of the trees above, as is the floor of the cave littered with bat-dispersed seeds. A very large part of the world's herbivore machine is run with this fuel and actually should be labeled "detritivore" rather than herbivore.

The implications for evolutionary biology are huge, given that no matter how much herbivory occurs on these living

dead, there can be no natural selection inflicted on the food populations.

Living dead adult individuals are also a prominent part of many undisturbed habitats and ecosystems. These are the waifs, the strays. Each of these is a plant whose seed arrived, grew to an adult, but found itself in a place lacking whatever is needed to maintain a viable population (Janzen, 1986c). In complex interwoven tropical habitats and ecosystems, the species list in a given place may contain as many as 10 to 20% of these kinds of living dead. For example, if a valley-bottom forest is eliminated, over time a significant number of tree species may disappear from the adjacent ridge, not because of any direct impact on the ridge forest but because the portions of the populations that were there are no longer maintained by seed flow into them from the valley bottom. This phenomenon is particularly visible where a particular soil or slope is thoroughly cleared for a crop, and the natural vegetation is left relatively intact in a neighboring habitat, ostensibly to protect it. Some species disappear because the conserved habitat did not really have its mutualist animals and physical climate conserved, or because it is too small, but others disappear simply because they were naturally occurring living dead.

Not to belabor the obvious, a tree standing dormant in the tropical dry season is not reproducing in the narrow sense, but it is also not necessarily living dead. But this is tricky for the observing human. We are very accustomed to being around trees that are not, at that moment, undergoing anything that appears to be reproduction, yet are members in good standing of quite surviving populations. The living dead tree does not display anything much different at first glance. Recognition of living dead status requires in-depth knowledge of its activities over decades, requires knowing if and where its pollen is going, and requires knowing where its seeds are moving to and what happens to them when they get there. This understanding is not acquired with the casual glance (e.g., Aldrich and Hamrick, 1998; Curran *et al*., 1999; Hallwachs, 1986).

What of Small Plants?

The isolated tree in the pasture has been a convenient illustrative example, but the world to which these ideas apply is far greater than that of large tropical trees. A small herbaceous plant may be a perennial with longevity like that of a tree. When the euglossine bees are extinguished through forest partial clearing, a *Catasetum* orchid they pollinated is left high on the main trunk of a shade tree left behind, a living dead in its own right. The orchid may flower for a century, waiting in vain for its long-distance pollinators (Janzen, 1974). They are long gone, their year-round nectar and pollen sources turned to charcoal. A living dead clump of perennial grass on a landslide scar may for many decades produce its small hard seeds, designed millions of years ago for a trip through a seed-dispersing, now-extinguished, large herbivore to a new disturbed site (Janzen, 1984). It finally succumbs to its individual sterile fate as the landslide scar revegetates to forest. A living dead herbaceous morning glory (Convolvulaceae), sprouting and flowering year after year into the insolated roadside ditch from its underground tuber, may never again see the bees that once moved among its flowers and the flowers of the many other forest-edge species that once sustained them (e.g., Frankie *et al*., 1998).

But as mentioned earlier for a population of trees, even a population of annuals may also be a living dead population. Yes, each year it may flower and seed and disperse and then again germinate with the next rains. But did it make "enough" seeds? Were they set at the "right" time? Did they have the right genetic composition? Did they move to the right safe sites? Were those sites there to be moved to? Does the population do all this and much more to hold its place in the naturally shifting nature of its surroundings? Each year the population may decline a bit. Maybe even in some years it recovers. But overall, gradually it slides into local extinction.

Looking backward at the history of a plant population "going extinct," it may be possible to describe the decay of such a living dead population. Looking forward, however, it is much harder to label than is the living dead tree in a cornfield. After all, all populations have their ups and downs. How to know, other than retroactively, when a down is a downswing versus a slide into extinction? When the habitat destruction is major and obvious, the prediction is much easier, but perhaps more scientifically trivial, than when the habitat destruction is piecemeal, fuzzy, or widespread yet light.

What of Animals, Those Things that Move?

Reproduction—that is, membership in the population—has two components. On the one hand, it is selfevident that the individual needs to be physiologically able to reproduce. On the other hand, if it is ecologically neutered, it is as dead as if sliced off with a chain saw. Selection has not generally favored the ability of a tree to "know" that it has been ecologically neutered by the removal of its pollinators, its dispersal agents, or the safe sites for its juveniles, and then take remedial action. What would the mutant tree have to be able to do? Walk back to the forest? Animals, with their chance to move to a new ecological circumstance, get horny. They search for nesting sites and mates, they may fight harder for their surviving fewer children, or they may migrate or emigrate to other places. But, in the face of the sweeping and omnipresent hand of humanity, busily extending its extended genome to cover the globe with both people and their domesticates (Janzen, 1998), where is the potentially living dead animal to go, and how long does it have to get there? One can search only so long before dying of old age, becoming a road kill, or running out of stored food reserves.

The tropical agroscape, and most wildlands as well, are awash with living dead animals, animal populations, and animal arrays (also known as "communities," whatever those are). Latent extinction is everywhere, but it operates more rapidly on animals with their high turnover rate and their lower capacity for extended lives as dormant seeds, resprouting root stocks, clonal patches, and so on.

Humans contribute in a curious perceptual manner to us being less aware of the animal living dead. At the level of the large animals, "everybody knows" that jaguars and tapirs are still "here" because everyone knows someone who knows someone who saw one once. One sighting of one 10-year-old

jaguar crossing the road at noon 12 years ago will sustain the living dead jaguar in that area for decades, long past its consignment to the litter. It has taken more than three decades for the myth of Costa Rican giant anteaters, which once ranged these forests, to die a natural death.

Collectors and collections do their part as well. There is a snapshot of history present in our museum drawers, each specimen with its neat locality label. These collections continue the illusion of survival long past the reality. Retroactive data capture from museums gives a distribution map not of what is today on the Costa Rican countryside, but rather what once roamed where today sweeps unbroken waves of sugarcane, pasture, plantations, and horticulture. Intellectually every taxonomist knows this, but the orderly march of specimens across the museum drawers that read Panama, Costa Rica, Nicaragua, Guatemala, Veracruz, and San Louis Potosí lull one into thinking "surely over that huge geographic range there are still viable populations." Plants are not immune to these processes. It is just that with the more illusive, the shorter lived, the more mobile, the animal living dead may be more easily manifest in historical collections than on looking out the car window at 70 kmph.

And, when one descends from a field vehicle somewhere, a rare butterfly flutters from the museum drawer and down the roadside ditch, the cruel illusion is reinforced. Highly mobile animals are particularly effective at hiding the living dead from perception. The last living dead Costa Rican green macaws will fly across the countryside for decades. One small viable population of butterflies can create hundreds of living dead individuals searching across the food-plant-free agroscape until dying on windshields, of pesticides, or in the collector's net.

Some animals, like some plants, thrive in the agroscape. Are they living dead as well? The agroscape changes its biotic and its physical traits at the whim of some combination of the market and our technical ability to (re)engineer our domesticates (and produce new ones). Overnight the agroscape can flip from heaven to hell for a particular species. When cotton was the crop of choice on the Costa Rican countryside, the world was an ocean of food for native *Dysdercus* cotton-stainer bugs (as well as for a number of other native cotton herbivores). The local extinction of the bugs' original wild food plants (Malvaceae, Sterculiaceae, Bombacaceae) that accompanied the forest clearing for cotton fields was invisible. But when the downstream shrimp industry decided that it could no longer tolerate the pesticide runoff from the cotton fields, and cotton went the way of history, then so did the populations of cotton stainers. Some remain on as tiny (living dead?) populations on the seeds of local roadside malvaceous and sterculiaceous herbs, but even these may be living dead with their food plants easing their slide into extinction.

Does the ecologically neutered tree try harder, as an animal might? Could there be selection for such behavior? What does the isolated tree in the field perceive? What is perceived by an elephant-dispersed tree in a forest where the elephants have been extinguished? The tree in the field can know that much less pollen of this or that genetic composition now arrives, and may adjust accordingly—it may flower longer, it may set more seeds that are fertilized with its own pollen. It may make more flowers more regularly or it may set more wood or grow a larger crown. All of these things are simple responses to a circumstance that must occur in a natural forest to this or that individual that is not living dead. But the extinction of animal dispersal agents and safe sites for juvenile plants goes unheralded, with not even a potential feedback loop.

And What of the Things that Eat the Living Dead?

All have their predators, their parasites, their mutualists, their scavengers. Many of these are quite dependent on the traits of their hosts. Food is not food is not food. Narrowly host-specific specialists abound.

For every living dead individual, population, or species, there is a large suite of consumers—individuals, and even species—living at the margin of their existence. A seed predator weevil—*Rhinochenus stigma*—passes its larval stages in the pods of guapinol (*Hymenaea courbaril*) on the Costa Rican countryside (Janzen, 1974). It maintains what appears to be a healthy population in the annual to supra-annual fruit crops that are destined to fall and rot below the parent in the absence of both the Pleistocene megafauna and the agouti (*Dasyprocta punctata*), contemporary inheritor of the guapinol (Hallwachs, 1986). But as each of those old guapinol trees dies at the end of its 200 to 500 year life span, the weevil population takes another hit. One day the last living dead guapinol trees will die, and along with them will go what appears today to be a perfectly healthy community of weevils.

The guapinol is also fed on by leaf-eating caterpillars. One, a large saturniid, *Schausiella santarosensis*, eats only guapinol leaves and will go the way of the Rhinochenus weevil. Another, *Dirphia avia*, also a large saturniid, feeds also on the foliage of Spanish cedar (*Cedrela odorata*), mahogany (*Swietenia macrophylla*), oak (*Quercus oleoides*), and guarea (*Guarea excelsa*) (Janzen and Hallwachs, 2000). As the adult guapinol trees dwindle in number, how the *Dirphia avia* population will twist and change will depend in part on how many individuals of the other living dead remain. (You guess: How many Spanish cedar, mahogany and oak trees will be left standing by the Costa Rican roadside?) Perhaps *Guarea excelsa*, its wood of no commercial value, will be the only host plant left. Enough to sustain *Dirphia avia?* Who knows, but it certainly won't be the same moth population that it was before.

The flowers of the living dead *Andira* trees were once a primary food source for tens of thousands of individuals of hundreds of species of bees; today they are visited by only a pale shadow of this bee community (Frankie *et al.*, 1998). But those old adult *Andira* continue to produce their massive flower crops and will do so for many decades to come. Its copious fruits, now largely from pollination by domestic honey bees, lie rotting below their parents in the absence of the masses of frugivorous bats that once dispersed them (Janzen *et al.*, 1976).

As noted earlier, the living dead are a "natural" part of any plant population. They are those individuals that have fallen where they have no chance of survival to reproduction. There are even living dead that have lived past their reproductive age. However, these living dead differ from the tree in the field in a very critical way for those who consume them. These living dead are being continually replenished by the natural

dispersal process. They do not herald an invisible walk to extinction for the consumer.

Are There Living Dead Habitats and Ecosystems?

Even when heavily agroindustrialized, the tropical agroscape often has patches of wildlands (**Figure 5**)—forests along rivers and ravines, broken topography, swamps and marshes, vegetation on bad soil, no-man's land between rival owners, woodlots, hunting preserves, industrial accidents, parks, and parklets. This remaining natural vegetation is a patchwork and a dot map, and it appears to be 1 to 20% of the original vegetation. And it gives one hope.

One says, "aha, there are remnants. There is wild biodiversity on the countryside, in the agroscape. There is hope outside of the reserves" (which are so hard to maintain and seem so expensive in national park status). This is a cruel illusion. Descend to one of these patchlets of forest, so green, so tree-filled. It is a biodiversity desert, lacking 50 to 99% of its original biodiversity that it had when it was once part of a forested landscape. As a package it is a vegetational living dead. Its species list is a mix of actual living dead and a few species that can maintain viable populations under these circumstances. Our major problem is that we visit these patches as tourists. We were not there in 1965 to see their earlier biodiversity, to compare it with its pale shadow in 1999 (but see Frankie *et al.*, 1998).

Why are the survivors living dead, and what happened to those that have gone locally extinct? Part of them went when the area got so small that there were no longer circumstances for a viable population size. Part of them were explicitly mined or hunted. Part of them went when their mutualists, prey, and hosts went. Part of them went when the neighboring habitat, a habitat that spit seeds into the remaining forest and thereby maintained a population there, went to croplands. Part of them went when the seasons got drier, or wetter, or windier, or more fire-rich, or longer, or shorter, or, or, or.

Even those national parks that seem so secure are at major risk from this phenomenon. When the Southeast Asian dipterocarp trees fruit, the wild pigs come from everywhere and the collective seed crop of the preserved forest patch has no chance of satiating these seed predators (e.g., Curran *et al.*, 1999). It may be better to surround a conserved wildland with wild animal-free rice fields than oceans of secondary succession subsidizing waves of animals that then turn the small old-growth forest into yet more secondary succession by defecating seeds all over it (e.g., Janzen, 1983a).

The bottom line is that the complex fabric woven from thousands of interacting species has been ripped to bits. Many of those that seem to have survived are living dead, or the serendipitous few that find this new impoverished habitat to their competitive liking. In short, these patches are only pseudo-remnants, not really smaller pieces of what once was. Even those ecosystems and habitats that have always existed as small units—a marsh, a landslide scar, a volcano top, a patch of serpentine soil—did not live in isolation. Rather, each was maintained by a complex ebb and flow of immigrants, waifs, and influences from the neighbors. When the neighboring natural system is turned to cropland, the integrity of the small

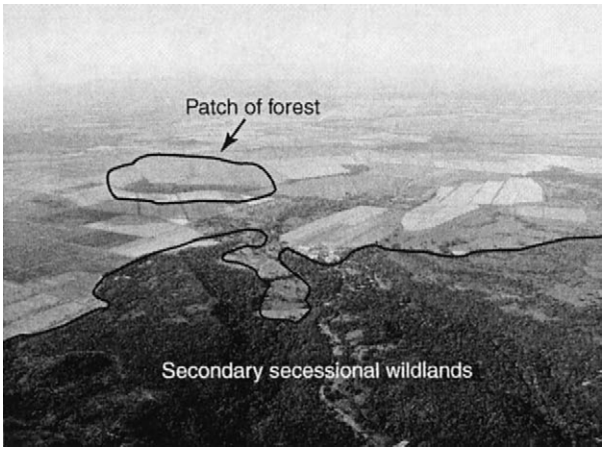


Figure 5 A living dead patch (left center) of natural vegetation, composed primarily of living dead individuals, among rice fields. There is essentially no gene flow between the patch and the secondary successional wildland in the foreground despite the thin connecting strip of riparian vegetation. Southwest of Liberia, Guanacaste Province, Costa Rica, December 14, 1999.

natural patch (e.g., **Figure 5**) is usually trashed almost as badly as if an army of chain saws had run through it. It just takes a bit longer for the living dead to live out their physiological lives.

These impoverished patches are especially deceptive for the bioilliterate. For those to whom a forest is just a batch of large woody plants, for those who cannot or will not read the differences between an advertising ditty and a complex poem, the agroscape with its living dead and pseudo-remnant natural vegetation appears to be not much different from a glade and forest mix in a national park. All seems to be well. But when humanity expects something from that wildland patch, it discovers that almost all of its tropical biodiversity is gone.

These patches have also played a mean trick on the conservation community. A huge portion of the world's conservation policy is based on the understandings of nature held largely intuitively by those who have grown up extra-tropical and learned their lessons from extra-tropical ecosystems. They easily adopt the mantra of trying to save the biodiversity remnants scattered across the agroscape. They are especially prone to do so in the face of the frustration of trying to save very large (and commercially juicy) blocks of intact vegetation. The forest-patchlet-dotted agroscape of Minnesota or Sweden still collectively contains easily more than 80% of the species that were there when the European colonists arrived. However, the same snapshot of a Costa Rican agroscape contains at best 5 to 20% of what once was. And the percent is still falling rapidly because a huge fraction of what remains today is living dead.

The more biodiverse and the more complex an ecosystem, the more likely that human perturbation will create anthropogenic living dead among the species with longer-lived individuals. This is because perturbations strip away mutualists and other biointeractors, leaving behind the physiologically functional individuals to live out their neutered life spans. The more biodiverse and the more complex, the more likely

any given individual is to be dependent on one or more of these interactants to remain a member of the population.

This tropical-to-extra-tropical comparison, derived by spending my life peering closely at both tropical and extra-tropical habitats is a major driver behind the conclusion that in the tropics a triage decision is needed. The living dead are writhing in lethal pain on the battlefield of the tropical agroscape. If we expend our scarce financial, political, and social resources on them instead of saving a few large coherent blocks of multi-ecosystem biophysical units, in the end we will live an even yet more impoverished biodiversity existence.

The future of real conservation in the tropics lies in by-and-large focusing our efforts on the survival of a relatively small number of very large and diverse biophysical units, each complicatedly integrated with local, national, and international societies (Janzen, 1998, 1999). Painful as it may be, resources spent on trying to save individual species and small habitat fragments scattered across the agroscape, often living dead, is bad conservation economics and creates an angry antagonistic *Homo sapiens*.

We have no option in the tropics but to recognize that conserved wildlands are and always will be islands in an ocean of agroscape. Our task is to get on with rendering them into the highest quality islands possible, and not be distracted by, nor lulled by, the living dead individuals and islandlets. Yes, if there remains but just one Rembrandt painting, we of course save it even if it is bullet-holed and faded. However, we must recognize it for what it is and not convince ourselves that by doing so we have preserved our knowledge of European history.

Restoration Biology

The living dead are largely a negative force in the algebra of conservation biology and conservation reality. However, in those few cases where ecosystem restoration is desired or serendipitous, their life span delimits a window of opportunity for the reintegration of their species into the restoring ecosystem. Reintegration is not an unqualified given, however. A single large tree in a pasture being restored to forest may be dropping its seeds and fruits into an early successional old-field community that for decades is still way too unattractive to contain the seed dispersal coterie that will begin to restore the demography of that tree species. Equally, the pollinators of its flowers may already be extinct, or abhor the young secondary succession coming up below the large old parent. And finally, the physical climate of the highly deciduous and dry-season blasted secondary succession may well be a dismal place for a seedling or sapling of that old-growth giant. As every plantation initiator knows, the act of stuffing seeds into the ground does not a plantation make.

Until a very short time ago, the California condor was made up of living dead individuals. They were brought into captivity (e.g., transplanted to a safe field), reproduced (e.g., seeds collected and grown in pots), and have been put back out, hopefully in an agroecosystem with a friendly sociology. This habitat is, however, very seriously impoverished through reduction of marine mammal populations that so kindly generated the cadavers for lunch, and the California condor may always be dependent on human subsidy.

Many species of living dead may be rescued in this manner, if we care enough to spend the resources on them and gather information about them. But before racing out to apply the same technique to the living dead guapinol trees in the centers of Costa Rican pastures, a question very much needs to be addressed. Would not the same money spent on saving large blocks of guapinol-occupied wildlands, complete with their pollinators and dispersal agents, not generate vastly more conservation of guapinol and its hundreds of thousands of compatriot species? Yes, even these large blocks of wildland will contain some living dead. The wildland's biodiversity will attain an equilibrium density at whatever number of species survive the reduction from a continent of wildland to a large island of wildland. Those who are extinguished during this process will suggest the list of who were the living dead.

See also: Central America, Ecosystems of. Conservation Biology, Discipline of. Deforestation and Land Clearing. Forest Ecology. Mammals (Late Quaternary), Extinctions of. Modern Examples of Extinctions. Pollinators, Role of. Range Ecology, Global Livestock Influences. Restoration of Biodiversity, Overview. Tropical Forest Ecosystems

References

Aldrich PR and Hamrick JL (1998) Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science* 281: 103–105.

Curran LM, Caniago I, Paoli GD, Astianti D, Kusneti M, Leighton M, Nirarita CE, and Haeruman H (1999) Impact of El Niño and logging on canopy tree recruitment in Borneo. *Science* 286: 2184–2188.

Frankie GW, Vinson SB, Rizzardi MA, Griswold TL, O'Keefe S, and Snelling RR (1998) Diversity and abundance of bees visiting a mass flowering tree species in disturbed seasonal dry forest, Costa Rica. *Journal of the Kansas Entomological Society* 70: 281–296.

Hallwachs W (1986) Agoutis (*Dasyprocta punctata*): The inheritors of guapinol (Hymenaea courbaril: Leguminosae). In: Estrada A and Fleming T (eds.) *Frugivores and Seed Dispersal*, pp. 285–304. Dordrecht: Dr. W. Junk Publishers.

Janzen DH (1974) The deflowering of Central America. *Natural History* 83: 48–53.

Janzen DH (1981) *Enterolobium cyclocarpum* seed passage rate and survival in horses, Costa Rican Pleistocene seed dispersal agents. *Ecology* 62: 593–601.

Janzen DH (1982a) Differential seed survival and passage rates in cows and horses, surrogate Pleistocene dispersal agents. *Oikos* 38: 150–156.

Janzen DH (1982b) How and why horses open *Crescentia alata* fruits. *Biotropica* 14: 149–152.

Janzen DH (1983a) No park is an island: increase in interference from outside as park size decreases. *Oikos* 41: 402–410.

Janzen DH (1983b) The pleistocene hunters had help. *American Naturalist* 121: 598–599.

Janzen DH (1984) Dispersal of small seeds by big herbivores: Foliage is the fruit. *American Naturalist* 123: 338–353.

Janzen DH (1985) On ecological fitting. *Oikos* 45: 308–310.

Janzen DH (1986a) The eternal external threat. In: Soule ME (ed.) *Conservation Biology: The Science of Scarcity and Diversity*, pp. 286–303. Sunderland, MA: Sinauer Associates.

Janzen DH (1986b) The future of tropical ecology. *Annual Review of Ecology and Systematics* 17: 305–324.

Janzen DH (1986c) Lost plants. *Oikos* 46: 129–131.

Janzen DH (1998) Gardenification of wildland nature and the human footprint. *Science* 279: 1312–1313.

Janzen DH (1999) Gardenification of tropical conserved wildlands: Multitasking, multicropping, and multiusers. *PNAS* 96(11): 5987–5994.

Janzen DH and Hallwachs W (2000) Philosophy, navigation and use of a dynamic database ("ACG Caterpillars SRNP") for an inventory of the macrocaterpillar fauna, and its food plants and parasitoids, of the Area de Conservación Guanacaste (ACG), Northwestern Costa Rica.(http://janzen.sas.upenn.edu).

Janzen DH and Martin PS (1982) Neotropical anachronisms: The fruits the gomphotheres ate. *Science* 215: 19–27.

Janzen DH, Miller GA, Hackforth-Jones J, Pond CM, Hooper K, and Janos DP (1976) Two Costa Rican bat-generated seed shadows of *Andira inermis* (Leguminosae). *Ecology* 56: 1068–1075.

Latitudinal and Elevational Range Shifts under Contemporary Climate Change

Jonathan Lenoir and Jens-Christian Svenning, Aarhus University, Aarhus, Denmark

© 2013 Elsevier Inc. All rights reserved.

Glossary

Front edge Stable margin of a species' range where population size is growing.
Leading edge Dynamic margin of a species' range where it is expanding into neighboring unoccupied areas.
Range core Central part of the distribution of a species.
Range periphery Marginal parts of the distribution of a species.
Rear edge Stable margin of a species' range where population size is declining.
Species' niche The scope of environmental conditions under which a species is able to maintain stable or increasing populations.
Species' optimum The value of a given environmental variable at which a species reaches its maximum probability of occurrence or abundance.
Species' range The spatial distribution of a species, that is, the sum of all geographic locations of its individuals over a given period in time.
Species' response curve The relationship between a species' probability of occurrence or its density (number of individuals per unit area) and the values of a given environmental variable.
Trailing edge Dynamic margin of a species' range where it is becoming locally extinct, causing the range to retract.

The concept of range shifts refers to the dynamic process whereby species shift their distribution over time, tracking geographic shifts in the climate and habitat conditions that they require. Range shifts are also commonly defined as migration events that occur on decadal to millennial or longer time scales, usually beyond the life span of individuals. In contrast, the concept does not include migration events occurring over the life span of individuals, that is, on a seasonal to annual or shorter time scales, as a result of recurring movements of individuals between spatially separate habitats in order to avoid temporarily unfavorable conditions. It also does not include essentially stochastic long-distance dispersal events of a single or few individuals that fail to establish new populations. Recently, species range shifts have attracted a great deal of interest from the scientific community as a direct consequence of the increasing focus on current global environmental changes, notably the changing climate (IPCC, 2007a). There is now a growing body of studies reporting and forecasting species range shifts across the main geographic dimensions (latitude, longitude, and elevation) as a response to both contemporary and future changes in climatic conditions. Here, the focus is on realized range shifts in response to contemporary climate change, a phenomenon that is of key importance for developing well-founded predictive models of range shifts under expected substantial future climate change (IPCC, 2007b). Because dramatic range shifts have also occurred in the geological past, species range shifts, until recently, have been described in textbooks through paleontological examples, but contemporary range shifts are now becoming even more evident.

Species' Range: Geographic Dimensions and Ecological Drivers

As a species' range can be defined by the geographic locations of all its individuals over a given period of time (Brown and Lomolino, 1998), it is often difficult and challenging to capture and represent it. Several abstractions of the real distribution of a single species have been used, with the most frequent approaches being to represent a species' range as outline, dot, or contour maps in a two-dimensional geographic space, most often across the latitudinal and longitudinal dimensions. One could also represent a species' range in a three-dimensional geographic space by adding occurrence along elevational gradients to the representation (**Figure 1**). Another more simplistic abstraction is to project the known or estimated occurrences of the species against one of these three dimensions. Hence, latitude, longitude, and elevation constitute the geographic dimensions over which a species' range is estimated, and a species' range shift is defined as changes in these geographic dimensions over the fourth dimension: time (**Figure 1**).

A species' range is generally understood to be at least partly determined by the species' environmental requirements and tolerances, that is, its ecological niche (Hutchinson, 1957; Soberon and Nakamura, 2009). The ecological niche is typically considered a multidimensional space (or "hypervolume"), in which the different dimensions represent different environmental conditions (Hutchinson, 1957). Especially at large scales, climate is often thought to constitute the niche axes that are most relevant for species ranges (Pearson and Dawson, 2003). Therefore, one should be able to deduce a species' climatic niche from its range. Unfortunately, the geographic distribution of a species is unlikely to fully represent its ecological niche for at least three reasons (Colwell and Rangel, 2009): (1) geographic unavailability of some portions of its ecological niche; (2) geographic inaccessibility due to its limited dispersal abilities (Svenning and Skov, 2004; Blach-Overgaard *et al.*, 2010); and (3) geographic control by constraining species interactions (Case *et al.*, 2005). Throughout this article, therefore, the species' range is solely being considered as an incomplete geographic reflection of the species' niche. Still, it is useful to relate ranges to the niche concept in

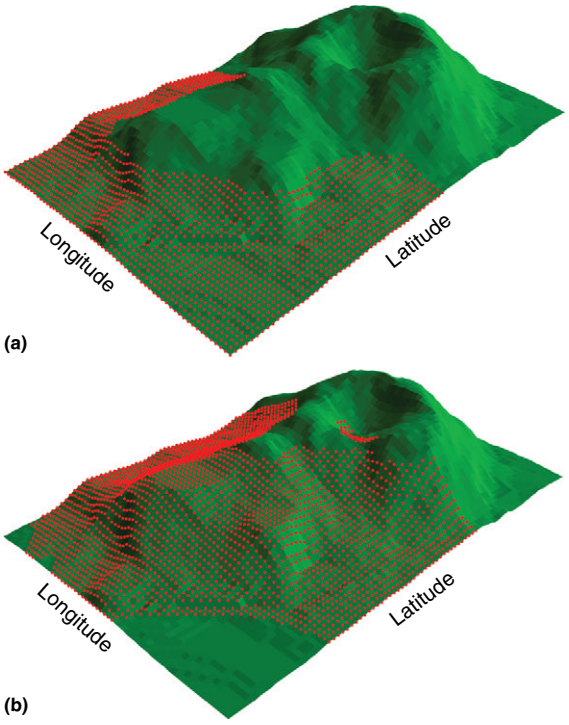


Figure 1 Representation of a species' range (red dots) across the three geographic dimensions (latitude, longitude, and elevation) of a given landscape in the Northern Hemisphere (the "volcano" data set available in R has been used (R Development Core Team, 2010) for illustrative purpose) and at two different time periods: t_1 (a) and t_2 (b). Note that the illustrative species shifted its geographic range upward and poleward between t_1 and t_2.

order to understand the mechanics of range shifts under climate change. Given this conceptual framework, a given species may respond differently to the environmental changes over time. It may (1) relocate toward newly suitable geographic locations, completely or partially tracking its environmental requirements (range shift) (Lenoir *et al.*, 2008); (2) remain within the same locations and acclimate or adapt to its new environment (niche shift) (Davis and Shaw, 2001); (3) go extinct (Svenning, 2003); or some combination of these responses. In short: move, adapt, or die (Maggini *et al.*, 2011).

How Do Species Shift Their Ranges?

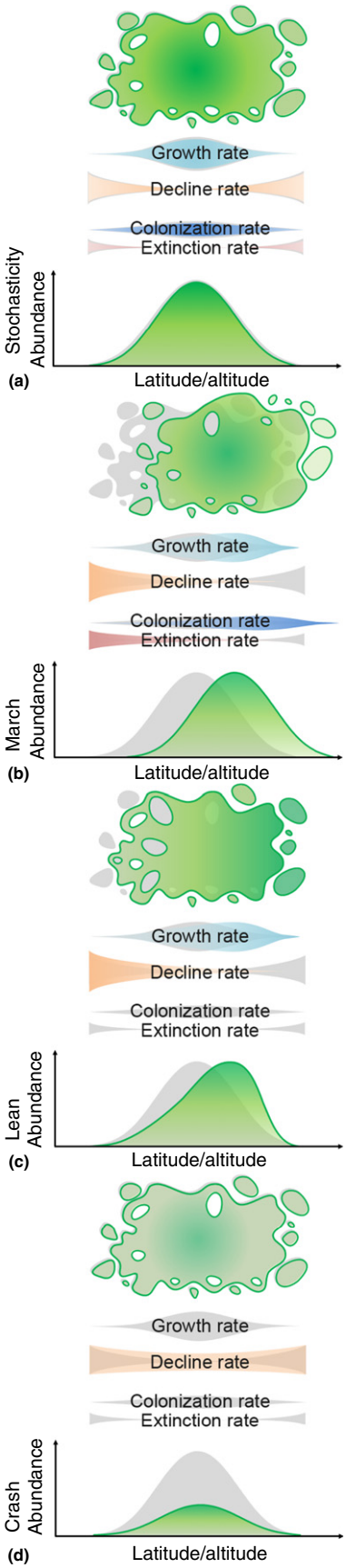
Species' ranges are not static over time, but constantly shift either more or less stochastically by means of population dynamics within a steady-state environment or directionally through population dynamics within a shifting environment that is spatially autocorrelated (**Figure 2**). Although many human-induced changes in the global environment, such as land-use change, climate change, nitrogen deposition, biological invasions, and atmospheric CO_2 increase (Sala *et al.*, 2000) may trigger species range shifts, here the focus is on the effect of contemporary climate change.

The Mechanics of Range Shifts: Distinction between Range Core and Range Periphery

In a steady-state environment, stochastic range shifts are the net result of growth, decline, colonization, and extinction processes that are tuned to different temporal scales (Breshears *et al.*, 2008). Whereas local populations grow and decline over relatively short time periods, local populations colonize and go extinct over much longer time periods, especially for long-lived species. Additionally, growth, decline, colonization, and extinction processes are tuned differently across a species' range (**Figure 2(a)**). Notably, there may be a strong core–periphery distinction leading to lower population densities near the peripheral area in comparison to the core area (Brown and Lomolino, 1998). Here the term "core" instead of "center" is explicitly used when referring to the geographic location of the abundance peak within a species' range. Indeed, the geographic location of the highest population density or abundance of a given species within its range is not necessarily located at the geometric center or centroid of its range (Murphy *et al.*, 2006), even though the core area of a species' range may be represented at the range center for simplicity (**Figure 2(a)**). At the core area of a species' range, sometimes called species' optimum to avoid confusion with the geometric center, local populations are generally thought to experience favorable conditions leading to enhanced fitness and population density. In other words, at the core, birth and emigration rates may exceed death and immigration rates, respectively, leading to sink populations (Pulliam, 2000). Such populations not only are likely to grow within the suitable localities already occupied, but are also likely to colonize and establish populations in suitable localities not previously occupied by the species in the core area of its range and even beyond, if dispersal and biotic interactions are not constraints. Decline and extinction processes might occur there as well, but to a lesser extent, unless there is a drastic shift in environmental conditions. At the peripheral area of a species' range, local populations may experience unfavorable conditions, leading to reduced fitness and reduced population density. In other words, at the periphery, death and immigration rates may exceed birth and emigration rates, respectively, leading to sink populations (Pulliam, 2000). Such range-edge sink populations will tend to decline and eventually go extinct, except when maintained by a "rescue effect" through constant immigration from source populations, typically located close to the core area of the species range. Growth and colonization processes might occur there as well, but to a lesser extent, unless there is a drastic shift in environmental conditions.

Directional Range Shifts under Climate Change

Let us now consider shifting conditions in the environment, such as the temperature and precipitation changes observed during contemporary climate change (IPCC, 2007b). Any shift in climatic conditions in a given location will impact the individuals living there and thus affect population dynamics locally. For instance, species sensitive to temperature may respond to a warmer climate through local changes in growth, decline, colonization, and extinction rates. On the one hand,



locations within the peripheral area of a species' range that were too cold and therefore less suitable might become more suitable in a warmer climate, thus turning sink populations into source populations with higher growth and colonization rates. On the other hand, locations within the core area of a species' range that were highly suitable might become too warm and therefore less suitable, turning source populations into sink populations with higher decline and extinction rates. Across species ranges, the net result of these changes in growth, decline, colonization, and extinction processes will affect their geographic distribution in different ways. However, an increase in temperature is likely to have an overall directional impact on species range shifts, because temperatures are autocorrelated in space, linking warmer conditions at lower latitudes and elevations with cooler conditions at higher latitudes and elevations. Therefore, one expects to observe poleward and upward range shifts as climate warms, even after accounting for dispersal limitations (Engler *et al.*, 2009) or biotic interactions (Araújo and Luoto, 2007). Accordingly, poleward and upward range shifts in the warming climates following the glacial recessions of the past interglacial periods have been widely reported for plants, birds, and mammals (Brown and Lomolino, 1998). For instance, Pleistocene fossils of several species of rodents have been found several thousand kilometers southward of the southern limit of their modern distribution in northern America (Graham, 1986), suggesting strong poleward shifts during the Holocene. Similarly, Pleistocene macrofossils of *Podocarpus* have been found *c.* 1000 m below the lower limit of their contemporary distribution on the Andean flank in western Amazonia (Cárdenas *et al.*, 2011), thus suggesting a large upward shift during the current interglacial period.

Although the expectations given temperature increase alone are relatively straightforward, it is difficult to predict how concurrent changes in other climatic factors, especially precipitation, will affect species ranges. Temperatures are negatively correlated along the latitudinal and elevational gradients and have risen globally over the past decades (IPCC, 2007b). In contrast, precipitation changes are more heterogeneous across space and time (IPCC, 2007b), leading to strong regional differences. Such regional variation in climate change, mainly due to the effect of precipitation changes on the water balance equation (balance between evapotranspiration and precipitation), may affect species ranges as well, leading to unexpected regional range shifts as climate warms globally (Crimmins *et al.*, 2011). Additionally, biotic interactions may also affect the magnitude of species range shifts as climate warms (Araújo and Luoto, 2007) with some suggestions that interactions could explain unexpected range shifts as well (Lenoir *et al.*, 2010a). Therefore, at global to continental scales, poleward and upward range shifts are likely

Figure 2 Conceptual representations of latitudinal and elevational range shifts and their mechanics under a steady-state environment (a) and under climate warming (b)–(d). In each case, the figure shows the relative importance of growth, decline, colonization, and extinction processes across the species range. For the march-, lean-, and crash-range shifts, gray shades represent conditions before a climate warming event whereas overlaying transparent colors represent shifting conditions under a climate warming event.

to be the main responses to the global increase in temperature conditions, albeit counterintuitive shifts should be expected as well at more regional scales due to regional variation in climate change and biotic interactions.

Types of Range Shifts

There are different types of range shifts. Many different terms have been used to refer to the many ways species may shift poleward or upward as climate warms, among others: leading-edge expansion, trailing-edge retraction, optimum shift, and abundance changes. Recently, a catalog of the possible species range shifts observable along a given gradient has been laid out, suggesting that observed species distribution changes are rarely complete shifts, but intermediate states in an ongoing shifting process (Maggini *et al.*, 2011). More generally, three main categories of observed range shifts can be distinguished as climate warms (Breshears *et al.*, 2008): (1) "march"; (2) "lean"; and (3) "crash" (**Figures 2(b)–(d)**).

- March-range shifts represent the most dynamic category of range shifts and involve dynamic edges with establishment from enhanced colonization and growth at the leading edge and mortality from enhanced decline and extinction at the trailing edge. Such dynamics driven by the edges usually lead to expansion toward higher latitudes or elevations (leading-edge expansion) or retractions from lower latitudes or elevations (trailing-edge retraction) (Maggini *et al.*, 2011). Therefore, out of this first category, one can develop several kinds of range shifts involving any combinations of expansion at the leading edge or retraction at the trailing edge, the most dynamic being a combination of both or march (**Figure 2(b)**). Short-lived, small-sized, and mobile species for which colonization and extinction processes occur on a relatively short time scale are likely to exhibit such dynamic range shifts. For instance, short-lived and small-sized butterfly species not strongly limited by their dispersal abilities have been reported to shift northward at both their leading and trailing edges (Parmesan *et al.*, 1999).
- Lean-range shifts involve stable edges with the optimum shifting within the existing range (**Figure 2(c)**). This represents the net result of enhanced growth at the highest latitudes or elevations (front edge) and decline at the lowest latitudes or elevations (rear edge) of the existing range, resulting in a shifting abundance pattern. This kind of shift might be transient and might correspond to early stages of a full shifting process (Maggini *et al.*, 2011). Long-lived, large-sized, and immobile species for which colonization and extinction processes occur on a much longer time scale than the period of observation are likely to exhibit such transient stages (time lag) across their ranges. For instance, long-lived and large-sized tree species, which tend to be strongly limited by their dispersal abilities, have been reported to shift their optimum elevation upward without shifting their range edges: persistence at the rear edge without establishment success beyond the front edge (Kelly and Goulden, 2008).
- Crash-range shifts involve stable edges and a stable optimum, but overall declines across the existing range (**Figure 2(d)**). Rare species having a restricted distribution for which the degree of climate warming far exceeds their ability to shift toward suitable conditions (because of dispersal limitation and natural or anthropogenic habitat fragmentation) are likely to exhibit such changes across their ranges. Similarly, there is no possible escape for endemic species having a distribution restricted to a specific mountain summit and facing contemporary climate change. In such situations, these threatened populations may simply crash. For instance, in 1987, a multipopulation crash of the endemic golden toad of Costa Rica's Cordillera de Tilarán has been reported to have led to the disappearance of this species (Pounds *et al.*, 1997). Additionally, species range shifts at the bottom of a food web may trigger trophic mismatch leading to multipopulation crash within the higher levels of the food web. For instance, fluctuations in plankton have resulted in a general decline in cod biomass and recruitment throughout a bottom-up control in the North Sea (Beaugrand *et al.*, 2003).

Of course, these three categories are not mutually exclusive and can be combined to develop all possible patterns of poleward or upward shifts as climate warms (Breshears *et al.*, 2008). Using the species response curve representation along one of the three geographic dimensions (**Figure 2**), the different combinations may affect the edges, optimum, level of abundance, or skewness of each species' response curve, thus leading to many possible patterns of range shifts (Maggini *et al.*, 2011).

Contemporary Evidence for Latitudinal and Elevational Range Shifts

An overwhelming number of reports of latitudinal and elevational range shifts under contemporary climate change have recently appeared in the scientific literature (IPCC, 2007a). The expected poleward and upward range shifts in warming regions have been reported for many taxonomic groups across the plant and animal kingdoms (Parmesan *et al.*, 1999; Sturm *et al.*, 2001; Lenoir *et al.*, 2008; Moritz *et al.*, 2008). A recent meta-analysis (Chen *et al.*, 2011a) concluded that the distributions of species have recently shifted poleward at a median rate of 16.9 km per decade and upward at a median rate of 11.0 m per decade, which is two to three times faster than reported in a previous meta-analysis (Parmesan and Yohe, 2003). However, these global trends represent averages across considerable variation in range dynamics under climate change (e.g., Chen *et al.*, 2011a). Indeed, most of the studies that have reported the expected range shifts toward higher latitudes or elevations as climate warms globally have also detected, at a smaller frequency and to a lesser extent, species moving toward lower latitudes or elevations (Parmesan *et al.*, 1999; La Sorte and Thompson, 2007; Lenoir *et al.*, 2008, 2010a; Moritz *et al.*, 2008). In most cases, modern latitudinal and elevational range shifts have been documented by using and sometimes resurveying historical occurrence data such as museum collections and field notes (Tingley and Beissinger, 2009). These historically based studies can be divided into two main groups (Parmesan, 2006): (1) studies that directly assess

range shifts by comparing historical and modern data on a species' distribution across its entire ranges or sections of the range; and (2) studies that infer range shifts by comparing historical and modern data on species composition or abundance within local communities. The following sections cover both types of studies.

Latitudinal Range Shifts

We carried out a selective review of historically based studies documenting contemporary latitudinal range shifts. Many species across many organism groups ranging from the most simple such as plankton, algae, and lichens to the most complex such as flowering plants, birds, and mammals, across both marine and terrestrial ecosystems, have shifted their latitudinal ranges over recent decades (**Table 1**). Most of these contemporary latitudinal range shifts are coherent with contemporary climate change, even though some studies reported the absence of latitudinal range shifts or even unexpected latitudinal range shifts, in addition to the expected and empirically predominant poleward range shifts (Thomas and Lennon, 1999; Warren *et al.*, 2001; Hill *et al.*, 2002; Brommer, 2004; Hitch and Leberg, 2007; La Sorte and Thompson, 2007).

Most latitudinal range shifts have been documented within the animal kingdom (invertebrates, fishes, amphibians, reptiles, birds, and mammals), whereas only a handful of cases have so far been documented for plants (**Table 1**). Among animals, there is clear evidence of poleward range shifts through either expansion at the leading edge or retraction at the trailing edge (Parmesan *et al.*, 1999; Hickling *et al.*, 2005), or increasing and decreasing abundance at the front and rear edges, respectively (Sagarin *et al.*, 1999; Myers *et al.*, 2009). For plants, however, reported changes seem to occur principally at the leading/front edge (Smith, 1994; Sturm *et al.*, 2001; Walther *et al.*, 2005b) without much clear evidence for changes at the trailing/rear edge (Lesica and McCune, 2004), in line with conclusions from an earlier review (Jump *et al.*, 2009) on the apparent lack of latitudinal range shifts at the trailing/rear edge of plant species ranges, especially for long-lived plants such as woody species (Jump *et al.*, 2009).

Hitherto, most latitudinal range shifts have been documented in the Northern Hemisphere, with few studies focusing on the Southern Hemisphere (including Antarctica) (**Table 1**), consistent with previously suggested differences in observed climate-induced changes in natural and managed ecosystems between the Northern and Southern Hemispheres (IPCC, 2007a). Additionally, in the Northern Hemisphere, most latitudinal range shifts have been documented within the Arctic, boreal, temperate, and Mediterranean biomes, that is, from middle to high latitudes (**Table 1**). Only few studies of continental extent have covered tropical and subtropical areas below 30° latitude (Beaugrand *et al.*, 2002; Hitch and Leberg, 2007; La Sorte and Thompson, 2007). Hence, there is a striking dearth of evidence for contemporary latitudinal range shifts from areas between −60° and 30° latitude (**Table 1**). The paucity of evidence from the Tropics and the Southern Hemisphere (except Antarctica) may have several explanations (IPCC, 2007a): (1) a marked scarcity of good historical data from developing countries that are suitable for resampling along the latitudinal gradient between −60° and 30° latitude; (2) a lack of taxonomical knowledge, especially in the Tropics, making it difficult to study multispecies latitudinal range shifts; (3) logistic problems, making it even more difficult to collect, access, and manage ecological data in these regions of the world; (4) the fact that mean annual temperature within the Tropics is approximately constant between −21° and 21° latitude, at any given elevation, so that little opportunity for latitudinal escape from warming exists (Colwell *et al.*, 2008); and (5) a huge imbalance of the Earth's land mass distribution with relatively little terrestrial area between −60° and 0° latitude, which probably contributes to the paucity of evidence for contemporary latitudinal range shifts for terrestrial ecosystems at these latitudes.

Elevational Range Shifts

A selective review of historically based studies documenting modern elevational range shifts was also carried out. Although most studies have reported range shifts toward higher elevations for a majority of taxa (**Table 2**), these studies have also reported, to a lesser extent, unexpected range shifts toward lower elevations (Lenoir *et al.*, 2010a). A few studies have even reported a general directional shift toward lower elevations for some taxonomic groups (Hickling *et al.*, 2006; Crimmins *et al.*, 2011). In Northern California, for instance, most plant species shifted downward to track regional changes in climatic water balance rather than temperature (Crimmins *et al.*, 2011). This shows that range shifts may differ not only in extent, but also in direction due to regional variation in climate change.

Contrary to latitudinal range shifts, as many elevational range shifts have been reported for plants as for animals (**Table 2**). Across all taxonomic groups, elevational range shifts toward higher elevations have been documented through changes either at the trailing/rear edge (Wilson *et al.*, 2005; Frei *et al.*, 2010), the leading/front edge (Hickling *et al.*, 2006; Frei *et al.*, 2010), or the core (Konvicka *et al.*, 2003; Lenoir *et al.*, 2008) of species ranges. Therefore, all possible patterns of upward shift between march- and lean-range shifts have been documented along the elevational gradient (Kelly and Goulden, 2008; Moritz *et al.*, 2008; Raxworthy *et al.*, 2008; Bergamini *et al.*, 2009). However, it has been suggested that plants, especially long-lived plants such as woody species, are more likely to lean rather than march their elevational range upward (Breshears *et al.*, 2008).

Just as for latitudinal range shifts, there is a strong imbalance of evidence for observed elevational range shifts under contemporary climate change between the Northern and Southern Hemispheres, most likely due to the higher density of mountainous area in the Northern Hemisphere relative to the Southern Hemisphere. Indeed, most elevational range shifts reports across taxonomic groups have been documented in the Northern Hemisphere, especially between 30° and 60° latitude (**Table 2**). Although there is a strong focus on temperate ecosystems in the Northern Hemisphere, most of the examples have been reported for subalpine, alpine, and nival ecosystems. Therefore, research efforts on elevational range shifts have been focusing on the coldest biomes (Antarctic,

Table 1 Examples of observed latitudinal range shifts under contemporary climate change

Taxonomic group	*No. taxa*	*Latitudinal zone (deg.)*	*Spatial extent*[a]	*No. years*	*T/R*	*C*	*L/F*	*Main changes*	*Reference*
Crustaceans	1	[− 90; −60]	C	75	x	x	x	Decreasing abundance	Atkinson *et al.* (2004)
Chordates	1	[− 90; −60]	C	75	x	x	x	Increasing abundance	Atkinson *et al.* (2004)
Moths	1	[30; 60]	R	30			x	Northward range shift	Battisti *et al.* (2005)
Fishes	2	[30; 60]	R	80			x	Increasing abundance	Beare *et al.* (2004)
Plankton	25	[30; 90]	C	40			x	Increasing richness	Beaugrand *et al.* (2002)
Plankton	11	[30; 90]	C	40	x			Decreasing richness	Beaugrand *et al.* (2002)
Birds	116	[60; 90]	R	10			x	Northward range shift	Brommer (2004)
Birds	34	[60; 90]	R	10	x			Absence of range shift	Brommer (2004)
Butterflies	4	[30; 60]	R	35	x			Northward range shift	Franco *et al.* (2006)
Odonates	24	[30; 60]	R	35			x	Northward range shift	Hickling *et al.* (2005)
Odonates	4	[30; 60]	R	35	x			Northward range shift	Hickling *et al.* (2005)
Arthropods	280	[30; 60]	R	25			x	Northward range shift	Hickling *et al.* (2006)
Herptiles	3	[30; 60]	R	25			x	Southward range shift	Hickling *et al.* (2006)
Fishes	15	[30; 60]	R	25			x	Northward range shift	Hickling *et al.* (2006)
Birds	22	[30; 60]	R	20			x	Absence of range shift	Hickling *et al.* (2006)
Mammals	9	[30; 60]	R	25			x	Northward range shift	Hickling *et al.* (2006)
Butterflies	46	[30; 60]	R	35			x	Absence of range shift	Hill *et al.* (2002)
Butterflies	4	[30; 60]	R	35	x			Absence of range shift	Hill *et al.* (2002)
Birds	27	[30; 60]	C	25			x	Northward range shift	Hitch and Leberg (2007)
Birds	29	[30; 60]	C	25	x			Absence of range shift	Hitch and Leberg (2007)
Fishes	75	[30; 60]	L	20	x	x	x	Decreasing abundance	Holbrook *et al.* (1997)
Birds	254	[30; 60]	C	30		x	x	Northward range shift	La Sorte and Thompson (2007)
Plants	7	[30; 60]	L	15	x			Decreasing abundance	Lesica and McCune (2004)
Algae	13	[30; 60]	R	55			x	Northward range shift	Lima *et al.* (2007)
Algae	26	[30; 60]	R	55	x			Northward range shift	Lima *et al.* (2007)
Acarids	1	[60; 90]	R	15			x	Increasing abundance	Lindgren *et al.* (2000)
Mammals	4	[30; 60]	R	125			x	Increasing abundance	Myers *et al.* (2009)
Mammals	5	[30; 60]	R	125	x			Decreasing abundance	Myers *et al.* (2009)
Butterflies	35	[30; 90]	C	50	x		x	Northward range shift	Parmesan *et al.* (1999)
Fishes	43	[30; 60]	R	25		x		Northward range shift	Perry *et al.* (2005)
Plankton	∼400	[30; 60]	C	45	x	x	x	Northward range shift	Richardson and Schoeman (2004)
Invertebrates	11	[30; 60]	L	65			x	Increasing abundance	Sagarin *et al.* (1999)
Invertebrates	7	[30; 60]	L	65	x			Decreasing abundance	Sagarin *et al.* (1999)
Plants	2	[− 90; −60]	R	25			x	Increasing abundance	Smith (1994)
Plants	4	[60; 90]	R	50			x	Increasing abundance	Sturm *et al.* (2001)
Birds	59	[30; 60]	R	20			x	Northward range shift	Thomas and Lennon (1999)
Birds	42	[30; 60]	R	20	x			Absence of range shift	Thomas and Lennon (1999)
Lichens	12	[30; 60]	R	20			x	Increasing richness	van Herk *et al.* (2002)
Lichens	66	[30; 60]	R	20	x			Decreasing richness	van Herk *et al.* (2002)
Plants	1	[30; 60]	R	50			x	Northward range shift	Walther *et al.* (2005a, b)
Butterflies	46	[30; 60]	R	30			x	Decreasing abundance	Warren *et al.* (2001)
Birds	41	[30; 60]	R	20			x	Absence of range shift	Zuckerberg *et al.* (2009)
Birds	129	[30; 60]	R	20		x		Northward range shift	Zuckerberg *et al.* (2009)
Birds	43	[30; 60]	R	20	x			Northward range shift	Zuckerberg *et al.* (2009)

[a]C: continental; L: local; R: regional.

Examples of recent studies reporting, for both marine and terrestrial ecosystems, modern latitudinal range shifts of organisms at the trailing/rear edge (T/R), the core (C), or the leading/front edge (L/F) of their latitudinal range. We restricted our literature survey to studies that have explicitly used historical occurrence data to assess range shifts under modern climate change. We focused on the following two types of historically based studies (Parmesan, 2006): (1) those that directly assess range shifts by comparing historical and modern data on species distribution across their entire ranges or sections of ranges; and (2) those that infer range shifts by comparing historical and modern data on species composition or abundance within local communities. Shaded cells represent reports for marine ecosystems, whereas nonshaded cells represent reports for terrestrial ecosystems.

Arctic, and alpine) with much less research on elevational range shifts in hotter areas such as Mediterranean and tropical biomes (**Table 2**). However, on the contrary to the latitudinal range shifts, there is relatively many studies documenting elevational range shifts in the Tropics, for example, for plants, insects, amphibians, and reptiles (Pounds *et al.*, 1997; Raxworthy *et al.*, 2008; Chen *et al.*, 2009; Feeley *et al.*, 2011; Juvik *et al.*, 2011). Continental-scale studies of elevational range shifts under contemporary climate change are still lacking (**Table 2**).

Table 2 Examples of observed elevational range shifts under contemporary climate change

Taxonomic group	*No. taxa*	*Latitudinal zone (deg.)*	*Spatial extent*[a]	*No. years*	*T/R*	*C*	*L/F*	*Main changes*	*References*
Birds	41	[30; 60]	L	30		x		Absence of range shift	Archaux (2004)
Moths	1	[30; 60]	L	30			x	Upward range shift	Battisti *et al.* (2005)
Trees	3	[30; 60]	L	40			x	Increasing abundance	Beckage *et al.* (2008)
Trees	3	[30; 60]	L	40	x			Decreasing abundance	Beckage *et al.* (2008)
Bryophytes	61	[30; 60]	R	125	x	x	x	Upward range shift	Bergamini *et al.* (2009)
Moths	102	[0; 30]	L	40		x		Upward range shift	Chen *et al.* (2009)
Moths	28	[0; 30]	L	40			x	Downward range shift	Chen *et al.* (2011b)
Moths	28	[0; 30]	L	40	x			Upward range shift	Chen *et al.* (2011b)
Moths	109	[0; 30]	L	40			x	Upward range shift	Chen *et al.* (2011b)
Moths	109	[0; 30]	L	40	x			Upward range shift	Chen *et al.* (2011b)
Plants	64	[30; 60]	R	80		x		Downward range shift	Crimmins *et al.* (2011)
Epiphyte	1	[30; 60]	R	95			x	Upward range shift	Dobbertin *et al.* (2005)
Trees	38	[− 30; 0]	L	5		x		Upward range shift	Feeley *et al.* (2011)
Butterflies	4	[30; 60]	R	35	x			Upward range shift	Franco *et al.* (2006)
Plants	125	[30; 60]	L	95	x		x	Upward range shift	Frei *et al.* (2010)
Plants	60	[30; 60]	L	95			x	Increasing richness	Grabherr *et al.* (1994)
Arthropods	280	[30; 60]	R	25			x	Upward range shift	Hickling *et al.* (2006)
Herptiles	3	[30; 60]	R	25			x	Downward range shift	Hickling *et al.* (2006)
Fishes	15	[30; 60]	R	25			x	Upward range shift	Hickling *et al.* (2006)
Birds	22	[30; 60]	R	20			x	Absence of range shift	Hickling *et al.* (2006)
Mammals	9	[30; 60]	R	25			x	Upward range shift	Hickling *et al.* (2006)
Butterflies	15	[30; 60]	R	35		x		Upward range shift	Hill *et al.* (2002)
Plants	140	[30; 60]	L	120			x	Upward range shift	Holzinger *et al.* (2008)
Plants	22	[0; 30]	L	50			x	Increasing richness	Juvik *et al.* (2011)
Plants	10	[30; 60]	L	30	x	x	x	Upward range shift	Kelly and Goulden (2008)
Butterflies	119	[30; 60]	R	50		x		Upward range shift	Konvicka *et al.* (2003)
Trees	7	[60; 90]	L	60			x	Upward range shift	Kullman (2002)
Plants	171	[30; 60]	R	100		x		Upward range shift	Lenoir *et al.* (2008)
Plants	60	[30; 60]	R	20	x	x	x	Upward range shift	Lenoir *et al.* (2010b)
Plants	22	[− 90; –60]	L	40	x		x	Upward range shift	le Roux and McGeoch (2008)
Mammals	28	[30; 60]	L	90	x	x	x	Upward range shift	Moritz *et al.* (2008)
Plants	78	[60; 90]	R	40			x	Upward range shift	Odland *et al.* (2010)
Plants	93	[30; 60]	L	50			x	Upward range shift	Parolo and Rossi (2008)
Plants	10	[30; 60]	L	10			x	Increasing abundance	Pauli *et al.* (2007)
Plants	10	[30; 60]	L	10	x			Decreasing abundance	Pauli *et al.* (2007)
Birds	94	[0; 30]		30			x	Upward range shift	Peh (2007)
Birds	94	[0; 30]		30	x			Absence of range shift	(Peh 2007)
Amphibians	50	[0; 30]	L	5	x	x	x	Decreasing abundance	Pounds *et al.* (1997)
Reptiles	30	[− 30; 0]	R	10	x	x	x	Upward range shift	Raxworthy *et al.* (2008)
Birds	1	[30; 60]	R	30	x		x	Upward range shift	Tryjanowski *et al.* (2005)
Plants	31	[30; 60]	L	100			x	Increasing richness	Vittoz *et al.* (2008)
Plants	18	[30; 60]	L	100			x	Increasing richness	Walther *et al.* (2005a)
Butterflies	16	[30; 60]	R	35	x			Upward range shift	Wilson *et al.* (2005)
Birds	41	[30; 60]	R	20			x	Absence of range shift	Zuckerberg *et al.* (2009)
Birds	129	[30; 60]	R	20		x		Downward range shift	Zuckerberg *et al.* (2009)
Birds	43	[30; 60]	R	20	x			Absence of range shift	Zuckerberg *et al.* (2009)

[a]C: continental; L: local; R: regional.

Examples of recent studies reporting, for terrestrial ecosystems, modern elevational range shifts of organisms at the trailing/rear edge (T/R), the core (C), or the leading/front edge (L/F) of their elevational range. We restricted our literature survey to studies that have explicitly used historical occurrence data to assess range shifts under modern climate change (see **Table 1** for further explanations).

Latitudinal and Elevational Range Shifts are Lagging behind Climate Change

It is clear that contemporary climate change is driving latitudinal and elevational range shifts in species distribution worldwide (**Tables 1** and **2**). However, the magnitude of these biotic responses is unlikely to match the magnitude of contemporary climate change due to differences in speed, magnitude, and direction between the climatic changes and the biotic responses as well as differences among species. Time lags between geographical range shifts and past climate changes have already been reported for plants and animals (Davis, 1986; Svenning and Skov, 2007; Araújo *et al.*, 2008). One way to assess the importance of such time lags in the biotic responses to contemporary climate change is to compare the magnitude of observed latitudinal and elevational

range shifts of different taxonomic groups with the magnitude of expected range shifts given climate change alone. In the French mountain forests, for instance, the core of the elevational range for plants ($n=171$ species) shifted toward higher elevations (mean: 29 m per decade $\sim$0.17 °C per decade, considering an adiabatic lapse rate of 0.6 °C per 100 m in that region) only 40% of that expected simply from the observed temperature increase in that region between 1965 and 2005 (0.45 °C per decade $\sim$75 m per decade) (Lenoir *et al.*, 2008). For purposes of comparison, the core of the elevational range for reptiles and amphibians ($n=30$ species) in Madagascar shifted toward higher elevations (mean: 65 m per decade $\sim$ 0.32 °C per decade, considering an adiabatic lapse rate of 0.5 °C per 100 m in that region) approximately 85% of that expected from the observed temperature increase in that region between 1984–1993 and 1994–2003 (0.37 °C per decade $\sim$74 m per decade) (Raxworthy *et al.*, 2008). This suggests that the magnitude of lags in range-shift responses to climate change may vary considerably among taxonomic groups (Davis, 1986).

Many different factors might be responsible for such delayed responses. Life span is obviously one of those factors. Species with shorter life spans such as insects and small mammals should be able to respond relatively fast to climate change, being delayed only a few years after a climatic event, whereas species with longer life spans, including species with slow maturation such as large mammals and many trees, should show much longer time lags, only responding to long-term climatic changes (Davis, 1986). Accordingly, it has been documented under contemporary climate change that species with shorter life cycles have shifted poleward and upward with greater magnitude than species with longer life cycles (Perry *et al.*, 2005; Lenoir *et al.*, 2008) (**Figure 3**). Time lags may also result from low dispersal ability. Species with limited dispersal ability, such as many plants, may not be able to rapidly colonize newly suitable climatic areas, even in the vicinity of their geographic ranges, causing time lags in their colonization of these areas (Davis, 1986). Suggestive of this mechanism, larger elevational range shifts have been reported for species with lighter diaspores as compared to species with heavier diaspores in the Italian Alps (Parolo and Rossi, 2008). Colonization of newly suitable habitats may also be delayed by competition from resident species (Davis, 1986). For instance, recruitment of *Fagus sylvatica* in southern France is delayed in grassland habitats due to its low tolerance to competition from herbaceous species (Kunstler *et al.*, 2007). Time lags in range-shift responses to climate change is furthermore likely to be most pronounced for species with small geographic ranges, as such species are often concentrated in long-term stable areas (Svenning and Skov, 2007; Araújo *et al.*, 2008), and thus likely to be particularly dispersal-limited. Linked to this idea that species with small geographic ranges might have pronounced time lags, range-shifts gaps – that is, lacking overlap between the current range and a future suitable range – are likely to occur and to increase the magnitude of these time lags if the geographic range of a species is smaller than the projected range shift under climate change (Colwell *et al.*, 2008). Crash-range shifts are a likely expectation for such species facing range-shift gaps under contemporary climate change. Finally, physiological adaptation to changing climate conditions through either evolutionary response or phenotypic plasticity may in some cases also cause ranges to shift less than expected from climate change (Davis and Shaw, 2001).

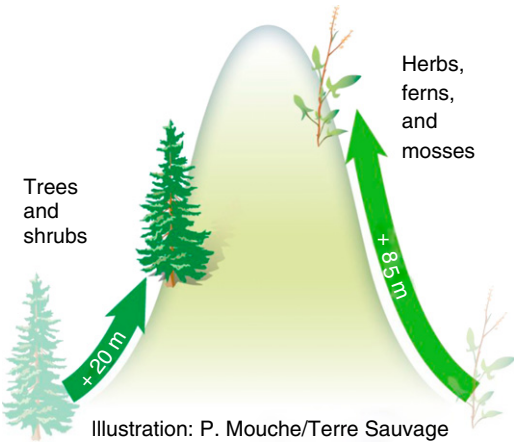


Figure 3 Differential range shifts between short- (herbs, ferns, and mosses) and long-lived (trees and shrubs) plant species along the elevational gradient of the French mountain forests under contemporary climate change (Lenoir *et al.*, 2008). The drawing is from P Mouche/Terre Sauvage. Reproduced from Etienne Hurault.

In addition to these biotic factors, spatial features may also affect time lags in range-shift responses to climate change. Considering contemporary climate warming, closeness among spatial isotherms is probably the most obvious factor. Indeed, time lags in range-shift responses to contemporary climate change are likely to be longer in flatland areas, which offer no short-distance escapes for species facing climate change and where the velocity of climate change is expected to be much stronger, compared to mountainous terrain (Loarie *et al.*, 2009). Therefore, time lags in latitudinal-range-shift responses to climate change should be longer than time lags in elevational-range-shift responses to climate change for a given species, particularly within the tropics, where latitudinal gradients in temperature are virtually absent (Colwell *et al.*, 2008). However, Chen *et al.* (2011a) found greater elevation lags than latitudinal lags in their meta-analysis, arguing that relatively strong elevation lags may arise from topographic and topoclimatic complexity in mountainous areas. Additionally, habitat fragmentation is likely to lengthen time lags in range-shift responses to climate change. In the context of future climate change, it has been simulated that the range of forest plant species in central Belgium will lag behind climate change, especially within the most fragmented landscapes (Honnay *et al.*, 2002).

Ecological Consequences

Ecological consequences of latitudinal and elevational range shifts under contemporary climate change are numerous, affecting biodiversity, communities, and ecosystem functioning. The often mentioned potential polar and mountaintop extinctions constitute an obvious example of the likely biodiversity consequences of climate-change-induced range shifts. Indeed, species restricted to these cold extremes have nowhere

else to go to escape climate warming. Therefore, their potential habitat is shrinking through a general reduction in surface area. Such range-restricted species may experience severe range contractions and may be the first groups in which species have already been driven to extinction by contemporary climate change (Parmesan, 2006). Endemic species restricted to low and isolated mountains in the temperate and tropical zones are especially threatened with extinction as a consequence of range-shift responses to climate warming (Krajick, 2004), for example, alpine endemics restricted to summits of low mountains lacking nival belts or subalpine endemics restricted to summits of low mountains lacking alpine and nival belts. In the long run, these low mountains may act as climatic traps for range-restricted species facing climate warming (Forero-Medina *et al.*, 2011). However, contemporary climate change at these high latitudes (poles) and elevations (mountaintops) also allow the immigration of new species from lower latitudes and elevations, which generally host a higher diversity, with the net result of these contemporary range shifts being a gradual increase in species richness toward the poles and mountain summits (Grabherr *et al.*, 1994; Walther *et al.*, 2005a), though time lags in range-shift responses to climate change are expected and may lead to both extinction debt and immigration credit (Jackson and Sax, 2010). Not only gradual increase in species richness is expected toward the poles and mountain summits as a net result of contemporary range shifts, but also lowland biotic attrition may occur in the Tropics as a net result of range contractions and extinctions without any possibility for compensation, contrary to higher latitudes, where immigrants can come from lower latitudes (Colwell *et al.*, 2008). Similar attrition may also occur in extratropical lowland areas due to migrational lags if source pools are distantly located (Svenning and Condit, 2008).

Communities may also be strongly affected by contemporary climate-driven range shifts, mainly throughout changes in species composition. Indeed, shifting species might be forced to interact with a new set of resident species from the recipient community, thus leading to community reorganization (Walther, 2010). Such new species assemblages may result in the formation of nonanalog communities, that is, communities that are different in species composition from any communities present at a selected reference point in time (Keith *et al.*, 2009). Range-shift-induced community reorganization may have important ecological consequences. For example, plant community homogenization among alpine summits has been reported in the Alps (Jurasinski and Kreyling, 2007) as a consequence of the upward shift of alpine plant species in the same region (Walther *et al.*, 2005a). However, such community reorganization occurs not only "horizontally," that is, within the same taxonomic group, but also "vertically," that is, across taxonomic groups and trophic levels, and, thus, could also strongly affect ecological networks.

Climate change effects on ecological networks could also have consequences for ecosystem functioning (Walther, 2010). For instance, species range shifts within a given trophic level may induce trophic cascades through bottom-up or top-down effects. An empirical example is the impact of sea surface warming on the poleward shift of phytoplankton in the Northeast Atlantic, which has been reported to propagate up the food web throughout bottom-up control on copepod herbivores and zooplankton carnivores (Richardson and Schoeman, 2004). In this case, the tight trophic coupling between phytoplankton, copepods, and zooplankton carnivores allows a preservation of the main functions ensured by this marine food web. However, such ecosystem functioning can be disrupted if the coupling between interacting species is too loose or if interacting species respond differently to climate change. The most striking evidence of such disruptions can be found in reports of pest outbreaks after range shifts of pest species released from the top-down control of their predators. As a typical example, recent range shifts of the pine processionary moth, *Thaumetopoea pityocampa*, has caused major outbreaks in Europe by switching hosts (Battisti *et al.*, 2005). Ecological consequences of these large-scale herbivore outbreaks may involve changes in the carbon sequestration budgets. Other important implications of species range shifts for ecological networks and ecosystem function concern the loss or gain of new functional types. Indeed, some ecosystems are likely to lose or gain functions as a result of species range shifts and community reorganization. In northern Arizona, recent vegetation shifts resulting from extreme mortality of pinyon, *Pinus edulis*, have affected the structure of the pinyon-juniper woodland, causing loss of avian seed dispersers, the disappearance of ectomycorrhizas specific to pinyon, and above- and below-ground competitive release of juvenile trees (Mueller *et al.*, 2005). In parallel, the recent expansion of shrubs in the Arctic (Sturm *et al.*, 2001) is changing the physiognomy of the arctic vegetation by allowing shrubs, a new functional type, to establish, with likely cascading ecological consequences. Indeed, it is suggested that increased vegetation cover in the Arctic may not only provide food resources for invasions of new animal species such as moths, already now causing outbreaks in the Arctic, but also cause ecological changes that may affect the Earth's climate itself, notably via changes in regional land–atmosphere greenhouse gas balances and carbon budgets (Post *et al.*, 2009).

Conservation Implications

Latitudinal and elevational range shifts and their resulting ecological consequences constitute a major challenge for conservation biology. Many recommendations for climate change adaptation strategies for biodiversity conservation have been recently listed in a literature review (Heller and Zavaleta, 2009). Here the focus is on two heavily discussed management recommendations: the traditional habitat-management recommendations and the controversial species-management recommendations. Most management recommendations so far have focused on conserving or restoring habitat quality to enhance species survival. The top-most-cited recommendation for biodiversity conservation under climate change is to increase habitat connectivity by establishing corridors, removing barriers for dispersal, and planning reserves close to each other so as to re-create series of stepping stones of high-quality habitats and increase possibilities for species range shifts (Heller and Zavaleta, 2009). Other habitat-management recommendations include creating and managing buffer zones within and around nature reserves (Heller and Zavaleta, 2009).

For instance, topographic heterogeneity is associated with low climate change velocity (Loarie *et al.*, 2009) and should enhance the long-term conservation capacity of nature reserves (Ackerly *et al.*, 2010) by providing short-distance escapes for species facing climate change (Scherrer and Körner, 2011). However, depending on the magnitude of climate change and the degree of topographic heterogeneity within species ranges, topographic heterogeneity may just temporarily buffer biodiversity before starting to act as a climatic trap by increasing spatial isolation and extinction risks (Forero-Medina *et al.*, 2011).

Nevertheless, it is also increasingly clear that the traditional habitat-management recommendations may be insufficient for safeguarding biodiversity from climate-driven extinctions, and therefore species-management recommendations are being proposed, notably assisted migration (also called assisted colonization or managed translocation) (Thomas, 2011). Among others, dispersal-limited or small-range species might not be able to be saved by more traditional habitat-management recommendations, such as increasing landscape connectivity to enhance species migration (Heller and Zavaleta, 2009; Thomas, 2011). Assisted migration is controversial, especially given the fear that translocated species could increase the risk of native species extinctions in their new recipient habitats (Ricciardi and Simberloff, 2009). However, for climate-threatened, dispersal-limited, or small-range species, it has been suggested that translocation might be the only successful management strategy, and therefore should be implemented provided that risks of extinction to native species in recipient habitats are small (Morueta-Holme *et al.*, 2011; Thomas, 2011). To help managers in making decisions between habitat- and species-management recommendations, a decision framework has already been proposed for assessing when species translocations are possible (Hoegh-Guldberg *et al.*, 2008).

Future Challenges

Despite the increasing number of studies documenting latitudinal and elevational range shifts under contemporary climate change, a better understanding of current range shifts is still needed for developing well-founded predictive models of future range shifts. More research efforts are needed to better understand climate change effects on species ranges at low latitudes and especially within the Tropics, the most species-rich part of the world and the area hosting most threatened taxa (IUCN, 2004). So far, evidence of recent species range shifts in the Tropics have been mostly documented along elevational gradients (Pounds *et al.*, 1997; Raxworthy *et al.*, 2008; Chen *et al.*, 2009; Feeley *et al.*, 2011; Juvik *et al.*, 2011) with at most a handful of reports for plants (Feeley *et al.*, 2011; Juvik *et al.*, 2011). Although the magnitude of climate warming varies geographically, most tropical areas have warmed just as much as temperate areas (IPCC, 2007a). Furthermore, small changes in temperature conditions in the Tropics may have much stronger biotic impact than similar changes in temperate areas. Notably, recent research into the thermal tolerance of terrestrial insects shows that tropical insects are closer to their maximum thermal tolerance than temperate insects (Deutsch *et al.*, 2008), suggesting that tropical species and their ranges may be particularly sensitive to climate warming. Importantly, crash-range shifts are therefore more likely in the Tropics (Pounds *et al.*, 1997) than elsewhere. Even more alarming, there do not exist biodiversity source pools now living in hotter places (as such areas do not exist) that are available to replace climatically displaced or declining species in tropical lowland ecosystems, potentially leading to lowland biotic attrition in the lowland tropical areas (Colwell *et al.*, 2008), with unknown consequences for ecosystem functioning.

Another future challenge is to achieve a better understanding of the mechanisms involved in species range shifts. Notably, it is still unresolved how the different population dynamics involved at both the range edges is driving species range shifts. Are colonization and establishment processes occurring more rapidly at the front edge of a species' range than mortality and extinction processes at the rear edge? For instance, the rear edge of a species' range is likely to host a relatively high genetic diversity due to selection pressure over time (Hampe and Petit, 2005), thus having a high potential for delaying range contractions under contemporary climate change. Therefore, rear edges may be disproportionately important for the survival and evolution of species (Hampe and Petit, 2005). Surprisingly, fewer studies have focused on contemporary range shifts at the trailing edge of species' ranges in comparison to the leading edge. This lack of focus on range shifts at the trailing edge of species' ranges is even more important along the latitudinal gradient for long-lived plants such as woody species (Jump *et al.*, 2009). In short, there is also an urgent need to improve our understanding of species range shifts at the trailing edge of species' ranges. Additionally, the strong variation in the magnitude of contemporary range shifts among species raises questions about the underlying reasons and whether the magnitude of shifts is linked to species traits. A recent study focusing on the relationship between species traits and the magnitude of recent shifts at the leading edge found low explanatory power, suggesting that trait-based range shift forecasts face several challenges (Angert *et al.*, 2011). Finally, the occurrence of unexpected range shifts (such as stasis or even downslope shifts) is an understudied phenomenon of high importance for predicting range responses to future climate change (Lenoir *et al.*, 2010a). This is linked to our current lack of knowledge regarding the effect of biotic interactions and dispersal constraints on species distributions.

Addressing these challenges will allow the development of improved predictive models of future range shifts. Indeed, with exactly this aim, recent discussions have raised the need for improving the field of species distribution modeling (SDM) by developing more mechanistic, hybrid approaches such as coupling stochastic population models with dynamic bioclimatic habitat models (Keith *et al.*, 2008) or by better accounting for nonclimatic factors and ecological complexity, notably biotic interactions and dispersal constraints (Morin and Lechowicz, 2008).

See also: Climate Change and Extinctions. Climate, Effects of. Evolution in Response to Climate Change. Introduced Species,

Impacts and Distribution of. Migration. Phenological Shifts in Animals Under Contemporary Climate Change. Species Distribution Modeling. Translocation as a conservation strategy. Trophic Cascades

References

Ackerly DD, Loarie SR, Cornwell WK, *et al.* (2010) The geography of climate change: Implications for conservation biogeography. *Diversity and Distributions* 16: 476–487.

Angert AL, Crozier LG, Rissler LJ, Gilman SE, Tewksbury JJ, and Chunco AJ (2011) Do species' traits predict recent shifts at expanding range edges? *Ecology Letters* 14: 677–689.

Araújo MB and Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16: 743–753.

Araújo MB, Nogues-Bravo D, Diniz-Filho JAF, Haywood AM, Valdes PJ, and Rahbek C (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography* 31: 8–15.

Archaux F (2004) Breeding upwards when climate is becoming warmer: no bird response in the French Alps. *Ibis* 146: 138–144.

Atkinson A, Siegel V, Pakhomov E, and Rothery P (2004) Long-term decline in krill stock and increase in salps within the Southern Ocean. *Nature* 432: 100–103.

Battisti A, Stastny M, Netherer S, *et al.* (2005) Expansion of geographic range in the pine processionary moth caused by increased winter temperatures. *Ecological Applications* 15: 2084–2096.

Beare D, Burns F, Jones E, *et al.* (2004) An increase in the abundance of anchovies and sardines in the north-western North Sea since 1995. *Global Change Biology* 10: 1209–1213.

Beaugrand G, Brander KM, Lindley JA, Souissi S, and Reid PC (2003) Plankton effect on cod recruitment in the North Sea. *Nature* 426: 661–664.

Beaugrand G, Reid PC, Ibanez F, Lindley JA, and Edwards M (2002) Reorganization of North Atlantic marine copepod biodiversity and climate. *Science* 296: 1692–1694.

Beckage B, Osborne B, Gavin DG, Pucko C, Siccama T, and Perkins T (2008) A rapid upward shift of a forest ecotone during 40 years of warming in the Green Mountains of Vermont. *Proceedings of the National Academy of Sciences of the United States of America* 105: 4197–4202.

Bergamini A, Ungricht S, and Hofmann H (2009) An elevational shift of cryophilous bryophytes in the last century – an effect of climate warming? *Diversity and Distributions* 15: 871–879.

Blach-Overgaard A, Svenning JC, Dransfield J, Greve M, and Balslev H (2010) Determinants of palm species distributions across Africa: The relative roles of climate, nonclimatic environmental factors, and spatial constraints. *Ecography* 33: 380–391.

Breshears DD, Huxman TE, Adams HD, Zou CB, and Davison JE (2008) Vegetation synchronously leans upslope as climate warms. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11591–11592.

Brommer JE (2004) The range margins of northern birds shift poleward. *Annales Zoologici Fennici* 41: 391–397.

Brown JH and Lomolino MV (1998) *Biogeography.* Sunderland: Sinauer Associates.

Cárdenas ML, Gosling WD, Sherlock SC, Poole I, Pennington RT, and Mothes P (2011) The response of vegetation on the Andean flank in western Amazonia to Pleistocene climate change. *Science* 331: 1055–1058.

Case TJ, Holt RD, McPeek MA, and Keitt TH (2005) The community context of species' borders: Ecological and evolutionary perspectives. *Oikos* 108: 28–46.

Chen IC, Hill JK, Ohlemüller R, Roy DB, and Thomas CD (2011a) Rapid range shifts of species associated with high levels of climate warming. *Science* 333: 1024–1026.

Chen IC, Hill JK, Shiu HJ, *et al.* (2011b) Asymmetric boundary shifts of tropical montane Lepidoptera over four decades of climate warming. *Global Ecology and Biogeography* 20: 34–45.

Chen IC, Shiu HJ, Benedick S, *et al.* (2009) Elevation increases in moth assemblages over 42 years on a tropical mountain. *Proceedings of the National Academy of Sciences of the United States of America* 106: 1479–1483.

Colwell RK, Brehm G, Cardelús CL, Gilman AC, and Longina JT (2008) Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. *Science* 322: 258–261.

Colwell RK and Rangel TF (2009) Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences of the United States of America* 106: 19651–19658.

Crimmins SM, Dobrowski SZ, Greenberg JA, Abatzoglou JT, and Mynsberge AR (2011) Changes in climatic water balance drive downhill shifts in plant species' optimum elevations. *Science* 331: 324–327.

Davis MB (1986) Climatic instability, time lags, and community disequilibrium. In: Diamond J and Case TJ (eds.) *Community Ecology*, pp. 269–284. New York: Harper and Row.

Davis MB and Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science* 292: 673–679.

Deutsch CA, Tewksbury JJ, Huey RB, *et al.* (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences of the United States of America* 105: 6668–6672.

Dobbertin M, Hilker N, Rebetez M, Zimmermann NE, Wohlgemuth T, and Rigling A (2005) The upward shift in altitude of pine mistletoe (Viscum album ssp. austriacum) in Switzerland – the result of climate warming? *International Journal of Biometeorology* 50: 40–47.

Engler R, Randin CF, Vittoz P, *et al.* (2009) Predicting future distributions of mountain plants under climate change: Does dispersal capacity matter? *Ecography* 32: 34–45.

Feeley KJ, Silman MR, Bush MB, *et al.* (2011) Upslope migration of Andean trees. *Journal of Biogeography* 38: 783–791.

Forero-Medina G, Joppa L, and Pimm SL (2011) Constraints to species' elevational range shifts as climate changes. *Conservation Biology* 25: 163–171.

Franco AMA, Hill JK, Kitschke C, *et al.* (2006) Impacts of climate warming and habitat loss on extinctions at species' low-latitude range boundaries. *Global Change Biology* 12: 1545–1553.

Frei E, Bodin J, and Walther GR (2010) Plant species' range shifts in mountainous areas-all uphill from here? *Botanica Helvetica* 120: 117–128.

Grabherr G, Gottfried M, and Pauli H (1994) Climate effects on mountain plants. *Nature* 369: 448.

Graham RW (1986) Response of mammalian communities to environmental changes during the late Quaternary. In: Diamond JM and Case TJ (eds.) *Community Ecology*, pp. 300–313. New York: Harper and Row.

Hampe A and Petit RJ (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters* 8: 461–467.

Heller NE and Zavaleta ES (2009) Biodiversity management in the face of climate change: A review of 22 years of recommendations. *Biological Conservation* 142: 14–32.

van Herk CM, Aptroot A, and van Dobben HF (2002) Long-term monitoring in the Netherlands suggests that lichens respond to global warming. *Lichenologist* 34: 141–154.

Hickling R, Roy DB, Hill JK, Fox R, and Thomas CD (2006) The distributions of a wide range of taxonomic groups are expanding poleward. *Global Change Biology* 12: 450–455.

Hickling R, Roy DB, Hill JK, and Thomas CD (2005) A northward shift of range margins in British Odonata. *Global Change Biology* 11: 502–506.

Hill JK, Thomas CD, Fox R, *et al.* (2002) Responses of butterflies to twentieth century climate warming: Implications for future ranges. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 269: 2163–2171.

Hitch AT and Leberg PL (2007) Breeding distributions of North American bird species moving north as a result of climate change. *Conservation Biology* 21: 534–539.

Hoegh-Guldberg O, Hughes L, McIntyre S, *et al.* (2008) Assisted colonization and rapid climate change. *Science* 321: 345–346.

Holbrook SJ, Schmitt RJ, and Stephens JS (1997) Changes in an assemblage of temperate reef fishes associated with a climate shift. *Ecological Applications* 7: 1299–1310.

Holzinger B, Hulber K, Camenisch M, and Grabherr G (2008) Changes in plant species richness over the last century in the eastern Swiss Alps: elevational gradient, bedrock effects and migration rates. *Plant Ecology* 195: 179–196.

Honnay O, Verheyen K, Butaye J, Jacquemyn H, Bossuyt B, and Hermy M (2002) Possible effects of habitat fragmentation and climate change on the range of forest plant species. *Ecology Letters* 5: 525–530.

Hutchinson GE (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22: 415–427.

IPCC (2007a) *Climate Change 2007: Impacts, Adaptation and Vulnerability.* Cambridge, UK: Cambridge University Press.

IPCC (2007b) *Climate Change 2007: The Physical Science Basis.* Cambridge, UK: Cambridge University Press.

IUCN (2004) 2004 IUCN Red List of Threatened Species: A Global Species Assessment. Cambridge and Gland, UK and CH: IUCN Publications Services Unit and the IUCN Species Programme.
Jackson ST and Sax DF (2010) Balancing biodiversity in a changing environment: Extinction debt, immigration credit and species turnover. *Trends in Ecology & Evolution* 25: 153–160.
Jump AS, Matyas C, and Penuelas J (2009) The altitude-for-latitude disparity in the range retractions of woody species. *Trends in Ecology & Evolution* 24: 694–701.
Jurasinski G and Kreyling J (2007) Upward shift of alpine plants increases floristic similarity of mountain summits. *Journal of Vegetation Science* 18: 711–718.
Juvik JO, Rodomsky BT, Price JP, Hansen EW, and Kueffer C (2011) The upper limit of vegetation on Mauna Loa, Hawaï: A 50th-anniversary reassessment. *Ecology* 92: 518–525.
Keith DA, Akcakaya HR, Thuiller W, *et al.* (2008) Predicting extinction risks under climate change: Coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters* 4: 560–563.
Keith SA, Newton AC, Herbert RJH, Morecroft MD, and Bealey CE (2009) Non-analogous community formation in response to climate change. *Journal for Nature Conservation* 17: 228–235.
Kelly AE and Goulden ML (2008) Rapid shifts in plant distribution with recent climate change. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11823–11826.
Konvicka M, Maradova M, Benes J, Fric Z, and Kepka P (2003) Uphill shifts in distribution of butterflies in the Czech Republic: Effects of changing climate detected on a regional scale. *Global Ecology and Biogeography* 12: 403–410.
Krajick K (2004) Climate change: All downhill from here? *Science* 303: 1600–1602.
Kullman L (2002) Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology* 90: 68–77.
Kunstler G, Thuiller W, Curt T, *et al.* (2007) *Fagus sylvatica* L. recruitment across a fragmented Mediterranean Landscape, importance of long distance effective dispersal, abiotic conditions, and biotic interactions. *Diversity and Distributions* 13: 799–807.
La Sorte FA and Thompson FR (2007) Poleward shifts in winter ranges of North American birds. *Ecology* 88: 1803–1812.
Lenoir J, Gegout JC, Dupouey JL, Bert D, and Svenning JC (2010b) Forest plant community changes during 1989–2007 in response to climate warming in the Jura Mountains (France and Switzerland). *Journal of Vegetation Science* 21: 949–964.
Lenoir J, Gegout JC, Guisan A, *et al.* (2010a) Going against the flow: Potential mechanisms for unexpected downslope range shifts in a warming climate. *Ecography* 33: 295–303.
Lenoir J, Gegout JC, Marquet PA, de Ruffray P, and Brisse H (2008) A significant upward shift in plant species optimum elevation during the twentieth century. *Science* 320: 1768–1771.
Lesica P and McCune B (2004) Decline of arctic-alpine plants at the southern margin of their range following a decade of climatic warming. *Journal of Vegetation Science* 15: 679–690.
Lima FP, Ribeiro PA, Queiroz N, Hawkins SJ, and Santos AM (2007) Do distributional shifts of northern and southern species of algae match the warming pattern? *Global Change Biology* 13: 2592–2604.
Lindgren E, Talleklint L, and Polfeldt T (2000) Impact of climatic change on the northern latitude limit and population density of the disease-transmitting European tick Ixodes ricinus. *Environmental Health Perspectives* 108: 119–123.
Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, and Ackerly DD (2009) The velocity of climate change. *Nature* 462: 1052–1111.
Maggini R, Lehmann A, Kery M, *et al.* (2011) Are Swiss birds tracking climate change? Detecting elevational shifts using response curve shapes. *Ecological Modelling* 222: 21–32.
Morin X and Lechowicz MJ (2008) Contemporary perspectives on the niche that can improve models of species range shifts under climate change. *Biology Letters* 4: 573–576.
Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, and Beissinger SR (2008) Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science* 322: 261–264.
Morueta-Holme N, Flojgaard C, and Svenning JC (2011) Climate change risks and conservation implications for a threatened small-range mammal species. *Plos ONE* 5: e10360.
Mueller RC, Scudder CM, Porter ME, Trotter RT, Gehring CA, and Whitham TG (2005) Differential tree mortality in response to severe drought: Evidence for long-term vegetation shifts. *Journal of Ecology* 93: 1085–1093.
Murphy HT, VanDerWal J, and Lovett-Doust J (2006) Distribution of abundance across the range in eastern North American trees. *Global Ecology and Biogeography* 15: 63–71.
Myers P, Lundrigan BL, Hoffman SMG, Haraminac AP, and Seto SH (2009) Climate-induced changes in the small mammal communities of the Northern Great Lakes Region. *Global Change Biology* 15: 1434–1454.
Odland A, Hoitomt T, and Olsen SL (2010) Increasing Vascular Plant Richness on 13 High Mountain Summits in Southern Norway since the Early 1970s. *Arctic Antarctic and Alpine Research* 42: 458–470.
Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics* 37: 637–669.
Parmesan C, Ryrholm N, Stefanescu C, *et al.* (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* 399: 579–583.
Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.
Parolo G and Rossi G (2008) Upward migration of vascular plants following a climate warming trend in the Alps. *Basic and Applied Ecology* 9: 100–107.
Pauli H, Gottfried M, Reiter K, Klettner C, and Grabherr G (2007) Signals of range expansions and contractions of vascular plants in the high Alps: observations (1994–2004) at the GLORIA* master site Schrankogel, Tyrol, Austria. *Global Change Biology* 13: 147–156.
Pearson RG and Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology & Biogeography* 12: 361–371.
Peh KSH (2007) Potential effects of climate change on elevational distributions of tropical birds in Southeast Asia. *The Condor* 109: 437–441.
Perry AL, Low PJ, Ellis JR, and Reynolds JD (2005) Climate change and distribution shifts in marine fishes. *Science* 308: 1912–1915.
Post E, Forchhammer MC, Bret-Harte MS, *et al.* (2009) Ecological dynamics across the Arctic associated with recent climate change. *Science* 325: 1355–1358.
Pounds JA, Fogden MPL, Savage JM, and Gorman GC (1997) Tests of null models for amphibian declines on a tropical mountain. *Conservation Biology* 11: 1307–1322.
Pulliam HR (2000) On the relationship between niche and distribution. *Ecology Letters* 3: 349–361.
R Development Core Team, I (I, 2010) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
Raxworthy CJ, Pearson RG, Rabibisoa N, *et al.* (2008) Extinction vulnerability of tropical montane endemism from warming and upslope displacement: A preliminary appraisal for the highest massif in Madagascar. *Global Change Biology* 14: 1703–1720.
Ricciardi A and Simberloff D (2009) Assisted colonization: Good intentions and dubious risk assessment. *Trends in Ecology & Evolution* 24: 476–477.
Richardson AJ and Schoeman DS (2004) Climate impact on plankton ecosystems in the northeast Atlantic. *Science* 305: 1609–1612.
le Roux PC and McGeoch MA (2008) Rapid range expansion and community reorganization in response to warming. *Global Change Biology* 14: 2950–2962.
Sagarin RD, Barry JP, Gilman SE, and Baxter CH (1999) Climate-related change in an intertidal community over short and long time scales. *Ecological Monographs* 69: 465–490.
Sala OE, Chapin III FS, Armesto JJ, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.
Scherrer D and Körner C (2011) Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. *Journal of Biogeography* 38: 406–416.
Smith RIL (1994) Vascular plants as bioindicators of regional warming in Antarctica. *Oecologia* 99: 322–328.
Soberon J and Nakamura M (2009) Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America* 106: 19644–19650.
Sturm M, Racine C, and Tape K (2001) Climate change – Increasing shrub abundance in the Arctic. *Nature* 411: 546–547.
Svenning JC (2003) Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters* 6: 646–653.
Svenning JC and Condit R (2008) Biodiversity in a warmer world. *Science* 322: 206–207.
Svenning JC and Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters* 7: 565–573.

Svenning JC and Skov F (2007) Ice age legacies in the geographical distribution of tree species richness in Europe. *Global Ecology & Biogeography* 16: 234–245.
Thomas CD (2011) Translocation of species, climate change, and the end of trying to recreate past ecological communities. *Trends in Ecology & Evolution* 26: 216–221.
Thomas CD and Lennon JJ (1999) Birds extend their ranges northwards. *Nature* 399: 213.
Tingley MW and Beissinger SR (2009) Detecting range shifts from historical species occurrences: New perspectives on old data. *Trends in Ecology & Evolution* 24: 625–633.
Tryjanowski P, Sparks TH, and Profus P (2005) Uphill shifts in the distribution of the white stork *Ciconia ciconia* in southern Poland: the importance of nest quality. *Diversity and Distributions* 11: 219–223.
Vittoz P, Bodin J, Ungricht S, Burga CA, and Walther GR (2008) One century of vegetation change on Isla Persa, a nunatak in the Bernina massif in the Swiss Alps. *Journal of Vegetation Science* 6: 671–680.
Walther GR (2010) Community and ecosystem responses to recent climate change. *Philosophical Transactions of the Royal Society B – Biological Sciences* 365: 2019–2024.
Walther GR, Beiβner S, and Burga CA (2005a) Trends in the upward shift of alpine plants. *Journal of Vegetation Science* 16: 541–548.
Walther GR, Berger S, and Sykes MT (2005b) An ecological 'footprint' of climate change. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 272: 1427–1432.
Warren MS, Hill JK, Thomas JA, *et al.* (2001) Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature* 414: 65–68.
Wilson RJ, Gutierrez D, Gutierrez J, Martinez D, Agudo R, and Monserrat VJ (2005) Changes to the elevational limits and extent of species ranges associated with climate change. *Ecology Letters* 8: 1138–1146.
Zuckerberg B, Woods AM, and Porter WF (2009) Poleward shifts in breeding bird distributions in New York State. *Global Change Biology* 15: 1866–1883.

Latitudinal Gradients of Biodiversity

Michael R Willig and Steven J Presley, University of Connecticut, Storrs, CT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alpha (α) diversity The diversity of species, often estimated as species richness, within a local community or site.

Beta (β) diversity The degree of turnover in species (and changes in their abundances) among communities or sites along a gradient or within a larger area.

Biodiversity A multidimensional concept embodying various dimensions (e.g., taxonomic, functional, phylogenetic, and genetic) that characterize variability of the biota in time or space.

Functional biodiversity A characterization of a biota that relates to variation in the importance (e.g., relative abundance, biomass, and relative frequency of occurrence) and functional attributes of taxa, and estimated by components such as guild richness, evenness, dominance, and diversity.

Gamma (γ) diversity The diversity of species, often estimated as species richness, in a larger area as a consequence of both α- and β-diversity.

Phenetic biodiversity A characterization of a biota that relates to variation in phenotypic characteristics (traits) of taxa, and estimated by components that reflect the size of the phenotypic hypervolume defined by constituent taxa, or the dispersion of phenotypes within that hypervolume.

Phylogenetic biodiversity A characterization of a biota that embodies evolutionary history and relates to variation among taxa based on their importance (e.g., relative abundance, biomass, and relative frequency of occurrence) and relationships by descent, including components, such as mean and variance of divergence rates or mean and variance of species ages.

Scale dependence A condition in which either the form or the parameters of a relationship between two variables (e.g., richness and latitude) is contingent on spatial or temporal attributes.

Species diversity A component of taxonomic biodiversity that reflects the variety of organisms in an area and that includes two components, species richness and species evenness (the degree to which all species have the same proportional abundance).

Taxonomic biodiversity A characterization of a biota that relates to variation in the importance and taxonomic identities of taxa, and estimated by components such as species richness, evenness, dominance, diversity, and rarity.

Introduction

Latitudinal gradients of biodiversity are biogeographic patterns that define the way in which components of taxonomic, phylogenetic, functional, genetic, or phenetic dimensions change with latitudinal position on the surface of the earth. By far, most research has considered only taxonomic biodiversity, generally species richness. Consequently, much of this article considers latitudinal gradients in species richness. The general pattern is for species richness to increase from polar to tropical regions (Willig *et al.*, 2003; Hillebrand, 2004), regardless of the taxonomic affiliation of organisms (e.g., mammals, fishes, insects, and plants) or geographic setting in which they occur (e.g., Africa, South America, and the Atlantic Ocean). This is true for extant organisms (**Figure 1**) as well as for those organisms alive during the past 70 million years (**Figure 2**). An increase in species richness with decreasing latitude is the pattern generally observed at three spatial scales, including the level of broad climatic zones (**Figure 3**), assemblages occupying arbitrary geographic subdivisions (i.e., quadrats or bands) of the earth's surface (**Figure 4**), and local ecological communities (**Figures 4** and **5**). Nonetheless, not all taxa increase with decreasing latitude in the same fashion, and a few groups do not even exhibit the general pattern of a latitudinal increase in richness towards the tropics. Moreover, considerable controversy surrounds the mechanisms that affect latitudinal patterns in richness, with ecological, evolutionary, historical, and stochastic processes (**Table 1**) championed as the cause(s) of observed gradients (Rohde, 1992). Indeed, many of the mechanisms are circular or unsubstantiated by empirical data.

Context

Since the voyages of Darwin and Wallace, biologists have been fascinated with the high species richness of tropical regions compared to those in temperate or boreal zones. Indeed, this fascination with tropical species richness catalyzed in many ways the conceptual development of the theory that currently constitutes modern ecology. Moreover, increasing concern about the loss of species, especially in tropical regions, has led to the rapid development of the science of conservation biology. Documenting the way in which biodiversity varies across the globe and understanding the mechanisms that produce such variation are critical steps in the design of global conservation strategies and the implementation of regional management plans.

The mid-1950s to early 1960s saw the emergence of rigorous quantification of broad-scale relationships between species richness and latitude (Fischer, 1960). Within the next 25 years, scientists convincingly had documented the ubiquity of gradients in which species richness increased toward tropical areas. Similar gradients (**Figures 3** and **4**) were

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00086-1

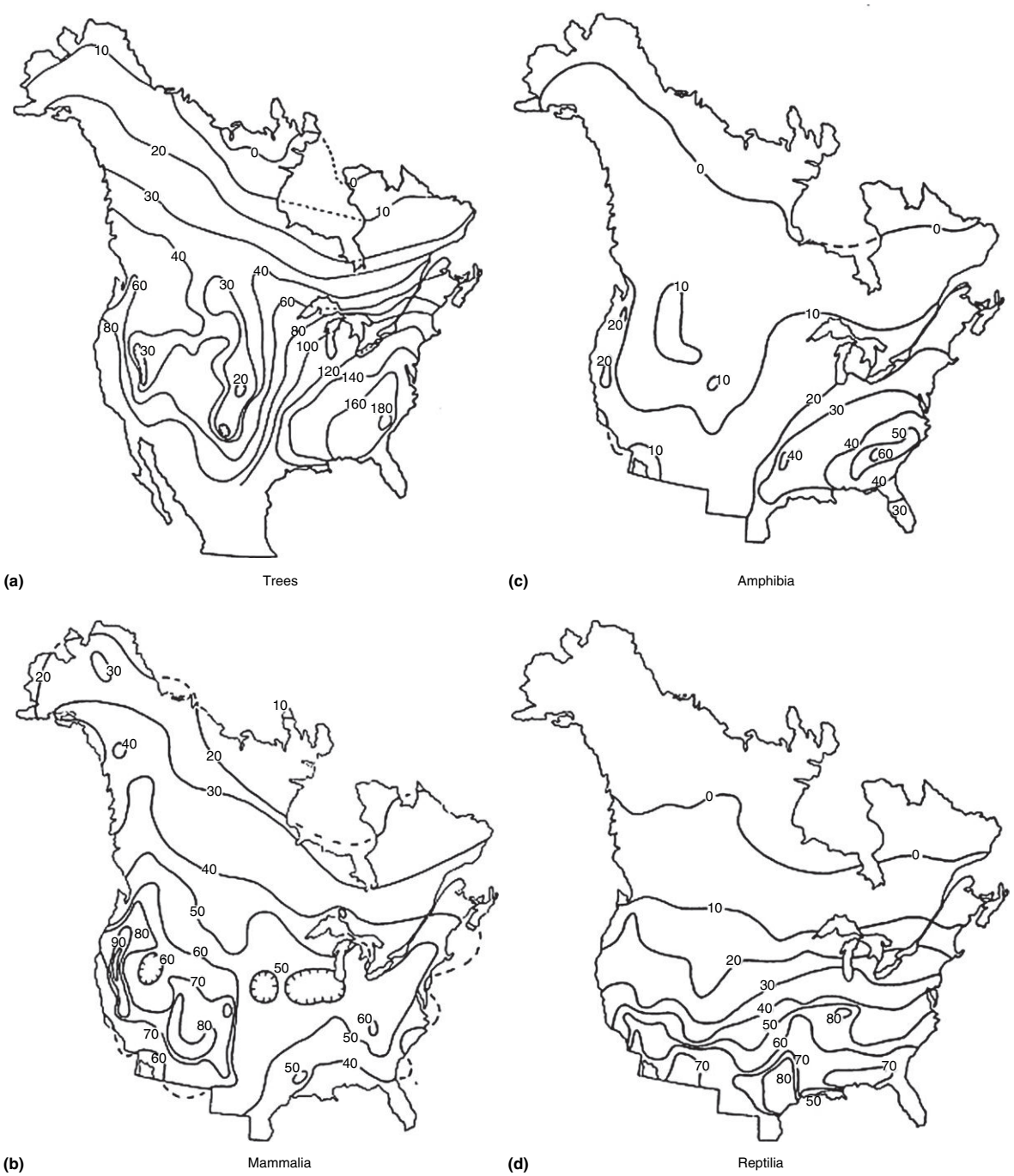


Figure 1 Species richness gradients in Canada and the US for trees (a), mammals (b), amphibians (c), and reptiles (d). Contour lines connect localities with approximately equal species richness. Reproduced from Currie DJ (1991) Energy and large-scale patterns for animal- and plant-species richness. *American Naturalist* 137: 27–49.

documented for the richness of higher taxonomic groups (e.g., genera, families, and orders). Indeed, the increase in species richness for terrestrial and marine environments was quantified successfully for a wide variety of taxonomic groups, such as mammals, birds, reptiles, amphibians, fish, tunicates, crustaceans, mollusks, brachiopods, corals, foraminiferans, and vascular plants. Nonetheless, some taxa representing lower levels in the systematic hierarchy (i.e., orders or families) were notable exceptions in having maximal richness in polar (e.g., seals, penguins, and sandpipers) or temperate zones (e.g., voles, salamanders, ichneumonid wasps, and coniferous trees). Caution must be employed in considering such exceptions because other groups of equivalent rank within the same higher taxon often are restricted to lower latitudes, and the higher taxon exhibits a tropical maximum in species richness.

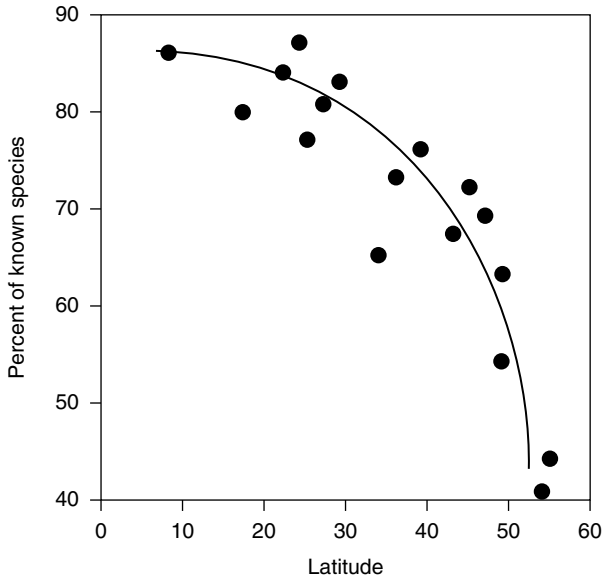


Figure 2 Latitudinal gradient in fossil species richness for marine Foraminifera from approximately 70 million years ago. Modified from Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, MA: Cambridge University Press.

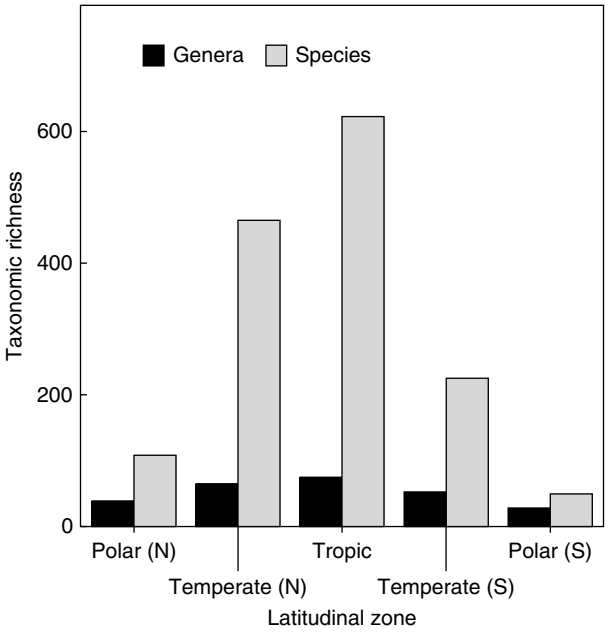


Figure 3 Generic and species richness of tunicates in each of the five major climatic regions defined by latitude and arranged from northern-most latitudes, through the tropics, to southern-most latitudes. Modified from Fischer AG (1960) Latitudinal variations in organic diversity. *Evolution* 14: 64–81.

Latitudinal gradients in other aspects of taxonomic biodiversity (e.g., species or generic evenness, diversity, dominance, and rarity) as well as in other dimensions of biodiversity (e.g., functional, phenetic, and phylogenetic) have received increasing attention in the past decade. This nascent field of investigation is ripe for exploring the form of gradients, the mechanisms that affect them, and the degree to which they are associated with each other.

Patterns

In the context of this article, a gradient implies a gradual change in biodiversity with a gradual change in latitude. In an unambiguous fashion, the form of that pattern is the precise mathematical or statistical relation that describes how biodiversity changes with latitude. As a consequence, three considerations are important in assessing patterns: the general shape of the curve (e.g., symmetry, kurtosis, or linearity), the parameters that characterize the relation, and the degree to which the fit of empirical data to the predicted curve is equivalent to the north and south of the equator. Knowledge of these three aspects of gradients suggests the kinds of causal mechanisms that are in operation. In addition, it facilitates comparison of gradients among taxa within the same geographic domain (**Figure 1**; birds versus mammals versus reptiles versus amphibian in North America) as well as comparisons among different geographic domains for the same taxon (**Figure 4**; North American versus South American for mammalian orders).

Patterns are often scale dependent, with particular mechanisms more likely operating at some spatial scales than at others. Consequently, patterns will be elucidated for each of two foci: biotic assemblages occupying broad areas and ecological communities occupying local sites. These scales are intimately associated with each other. In part, the species richness of regions, biomes, or climatic zones is a consequence of the species richness that is accumulated within constituent local communities. Similarly, the species richness and composition of local communities are affected by the set of taxa that constitute regional species pools.

Assemblages

Most of the empirical research concerning the relationship between biodiversity and latitude considers species richness and has been done using arbitrary sampling units that were based on (1) latitudinal bands, (2) quadrats of fixed area, or (3) quadrats of unequal area defined by lines of longitude (meridians) and latitude (parallels). Alternatively, research has focused on the species richness of biomes or broad latitudinally defined climatic zones. Because the area of any sampling unit may have as large or larger effect on variation in species richness than does its latitude, it is critical to understand how area may affect latitudinal patterns in different ways, depending on the method or approach (see Scheiner *et al.*, 2011).

Generally, analyses of quadrats defined by meridians and parallels are inferior to those based on other sampling units because such quadrats differ in area in a systematic fashion and bias quantitative conclusions. As meridians converge toward the poles, the size of the quadrats becomes smaller. Consequently, any attempt to control for variation in richness among quadrats as a consequence of area will remove at least some of the effects of latitude as well. This confounding effect compromises the detection of pattern across broad latitudinal gradients.

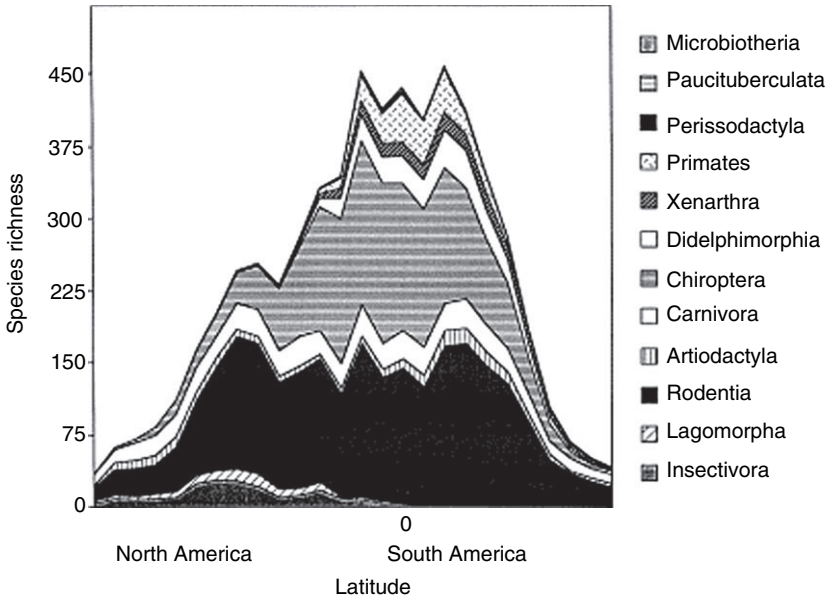


Figure 4 Differential contribution of mammalian orders to the latitudinal gradient of species richness in the New World based on data from 5° latitudinal bands. Reproduced from Kaufman DM (1995) Diversity of New World mammals: Universality of the latitudinal gradients of species and bauplans. *Journal of Mammalogy* 76: 322–334.

Analyses based on latitudinal bands also must control the effect of area because the width of a continent is not constant at all latitudes. A variety of methods have been used to compensate for this problem. Importantly, care must be employed when the area of bands varies in a systematic fashion with latitude because of the shape of the continent (e.g., progressive decreases in area with decreasing latitude in North America versus progressive increases in area with decreasing latitude in South America). In such cases, analyses based on bands may be plagued with the same confounding effects as those for analyses based on quadrats defined by meridians and parallels. Indeed, if classical statistical techniques are used to control for the effect of area in bands defined by 5° meridians in North America, the areal relationship is contrary to both common sense and ecological theory in that species richness increases as area decreases. Subsequent assessment of latitudinal effects may be severely compromised because the width of North America decreases as latitude decreases. Hence, adjusting the latitudinal gradient to account for area also removes an appreciable latitudinal effect. Nonetheless, if continental shape does not confound the effects of latitude and area, then regression techniques hold great promise, especially if appropriate nonlinear approaches are taken to adjust species richness in accord with species area theory (i.e., adjust richness of bands to a common area based on nonlinear regression of nested quadrats within each band).

Quadrats of fixed size also have been used to analyze broad-scale patterns of diversity. Nonetheless, variation in species richness among sampling units could still be a consequence of area, at least partially, when quadrats occupy coastal positions along continental borders. Adjusting for area in these cases may obscure the effects of rapid transition zones in terrestrial communities as they approach land–sea margins. Hence, the consensus is to not consider quadrats unless they are full of land. Subsequent variation among quadrats that is due to latitude can be assessed through a variety of statistical models. However, here too it is important to note that patterns are scale dependent. That is, the pattern detected for quadrats encompassing 100 km^2 could be quite different from those at 10,000 km^2. The importance of scale dependence in latitudinal studies has been emphasized increasingly during the past decade (e.g., Lyons and Willig, 2002).

Much of the early literature on latitudinal gradients in biodiversity was based on the species richness of broadly defined climatic zones (e.g., north polar, north temperate, tropical, south temperate, and south polar) or geopolitical units (e.g., countries, states, and provinces). Taxonomic richness was documented to increase from polar to tropical regions (**Figure 3**). Even when values for richness were not adjusted for the areal extent of geopolitical regions (e.g., snake species in Argentina, ant species in the Americas, and breeding bird species in the Americas), the polar to tropical gradient was obvious. Nonetheless, controversy currently surrounds the interpretation of such data when the focus is on broad climatic zones associated with latitude. Some suggest that the increase in diversity with decreasing latitude primarily is a result of more tropical regions having larger areas than their extratropical counterparts. In contrast, others suggest that extratropical areas are often larger but have fewer species than their more tropical counterparts, suggesting that latitudinal correlates other than area are the driving forces behind the gradient (*see* Geographic Area Hypothesis).

Ecological Communities

The way in which latitudinal variation in biodiversity at broad spatial scales (i.e., γ-scale) is related to patterns of biodiversity at the level of local communities (i.e., α-scale) is unclear. In part, this is because the geographic boundaries of a community are difficult to designate and are ultimately arbitrary

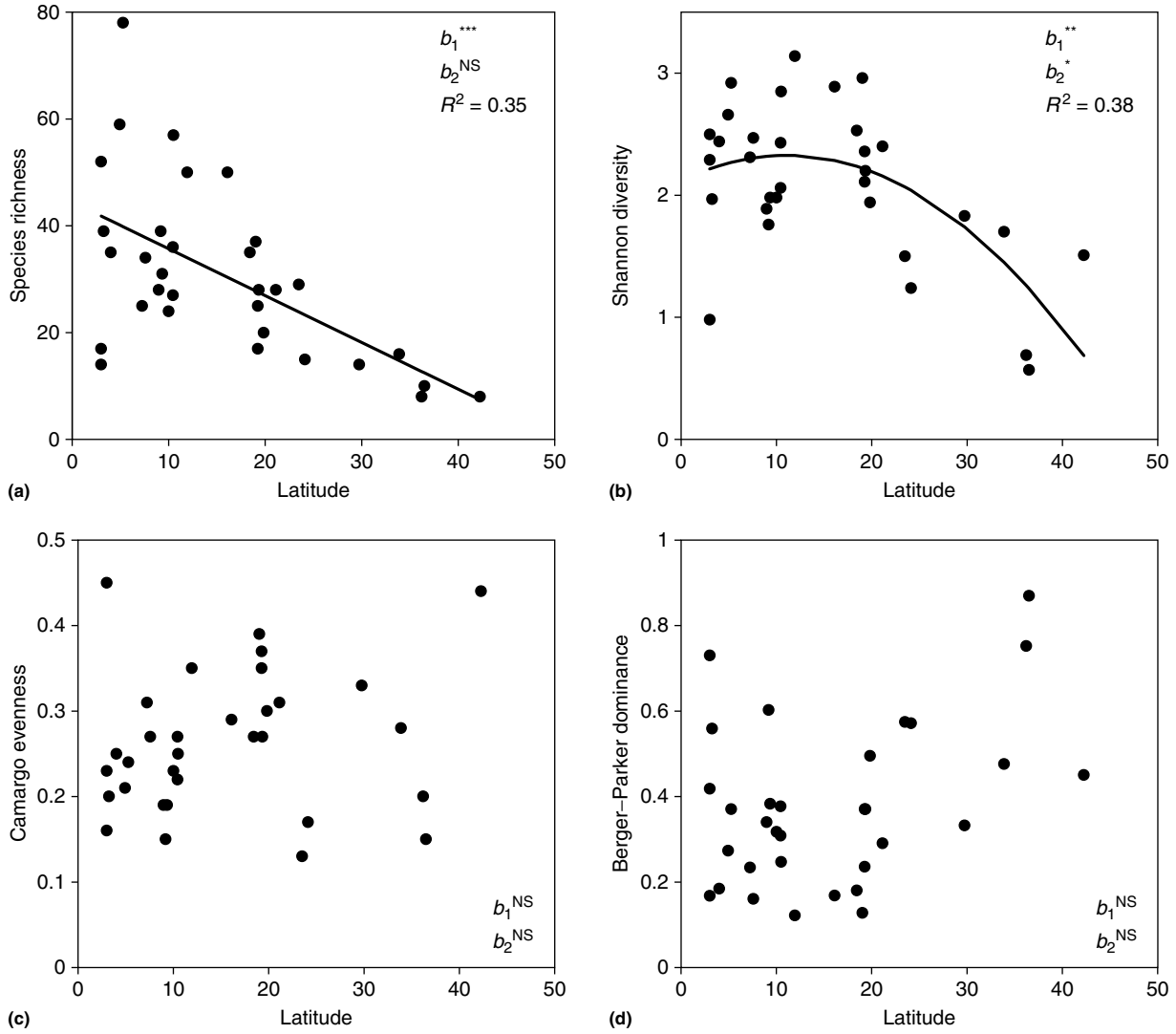


Figure 5 Latitudinal gradients of the taxonomic dimension of biodiversity – species richness (a), Shannon diversity (b), Camargo evenness (c), and Berger–Parker dominance (d) – within local communities for bats from throughout the New World. When regression models are significant, the contribution of linear (b_1) or quadratic (b_2) terms are indicated by superscripts (NS, $p > 0.05$; *, $0.05 \geq p > 0.01$; **, $0.01 \geq p > 0.001$; ***, $p > 0.001$); percent variation in a component of taxonomic biodiversity that is related to latitude is quantified by R^2. Reproduced from Stevens RD and Willig MR (2002) Geographical ecology at the community level: Perspective on the diversity of New World Bats. *Ecology* 83: 545–560.

decisions. Moreover, until recently, little was known about the composition of local communities in tropical regions, making assessment of broad-scale latitudinal patterns of biodiversity a premature endeavor at the community level. Finally, it is unlikely that a single research scientist can gather sufficient data across many sites to assess latitudinal gradients in biodiversity with sufficient power to distinguish pattern from noise. Hence, compositional data must be compiled from the work of many different individuals, who often use different methods, designs, and sampling intensities, to quantify the gradient in a meaningful way. Concern regarding adequate sampling within a community must be tempered by the realization that community composition has a temporal dynamic. Communities represent suites of populations with the potential to interact, and thus must be constrained to some extent by both time and space. Sampling regimes that extend over protracted periods of time (e.g., decades or longer) may have inflated estimates of richness and inaccurate assessments of species composition. Nonetheless, recent progress in this regard allows quantitative evaluation of patterns in a rigorous way, at least for some taxonomic groups.

When care is taken so that a local community is delimited as a geographic area in which constituent species have a high likelihood of interaction and of affecting each others' demographies, it is clear that species richness increases from polar through temperate to tropical regions. For example, data for volant mammals (bats) from 32 local communities (Stevens and Willig, 2002) that met rigorous standards for inclusion in analyses (i.e., well-delimited local areas must have been sampled intensively for at least 1 year to include all seasons in which species are active) clearly exhibit a latitudinal gradient of increasing richness with decreasing latitude (**Figure 5(a)**).

Table 1 Mechanisms potentially affecting the latitudinal gradient in species richness

Circular	*Empirically unsubstantiated*
Competition	Environmental stability
Mutualism	Environmental predictability (contingency)
Predation	Productivity
Epidemics	Abiotic rarefaction
Biotic spatial heterogeneity	Physical heterogeneity
Population size	Angle of the sun above the horizon
Niche width	Area
Population growth rate	Aridity
Patchiness	Seasonality
Epiphyte load	Number of habitats
Host diversity	Rapoport's rule (range size gradient)
Harshness	Ecological time
	Evolutionary time
	Temperature dependence of chemical reactions
	Solar energy
	Evolutionary speed
	Stochastic placement of species ranges

Source: Modified from Rohde K (1992) Latitudinal gradients in species diversity: The search for the primary cause. *Oikos* 65: 514–527.

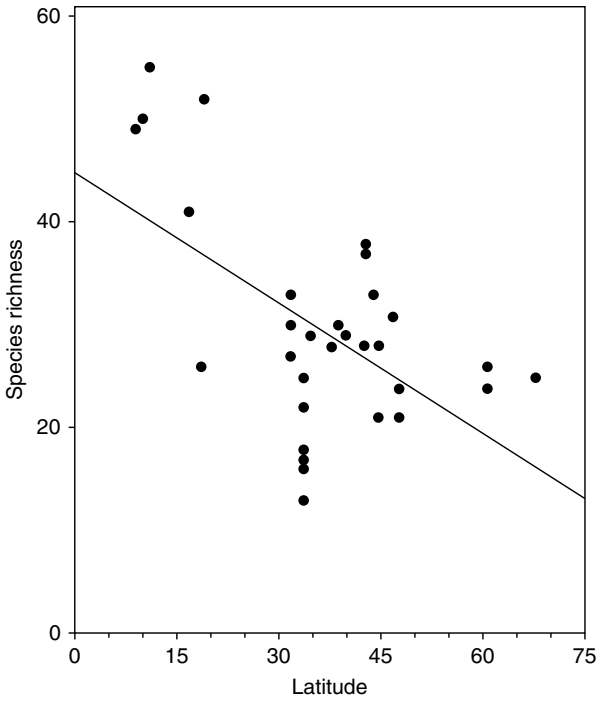


Figure 6 Latitudinal gradient of species richness for mammalian tetrapods in local communities in North America. Reproduced from Kaufman DM (1998) *The Structure of Mammalian Faunas in the New World: From Continents to Communities*. Unpublished PhD Dissertation, University of New Mexico, Albuquerque.

Species richness of nonvolant taxa (tetrapods) in mammalian communities from North America also shows a significant increase from polar to tropical regions (**Figure 6**; Kaufman, 1998).

Beyond species richness, taxonomic biodiversity has a number of other interrelated components such as evenness, diversity, dominance, and rarity. Components other than richness arise by weighting the presence of species by some measure of importance such as relative abundance, relative biomass, or relative frequency of occurrence. Moreover, each component can be estimated by different indexes that differ in the precise mathematical ways in which presence is weighted by importance. Richness is the number of species in a community; it can be equal to the empirical value or can be estimated via a variety of more complicated techniques such as extrapolation or rarefaction. Evenness represents the extent to which the distribution of individuals among species is equable. Diversity is a composite metric, comprising aspects of richness and evenness. Dominance is the extent to which a single taxon represents the largest proportion of individuals in a community. Rarity is a measure of the number of taxa below some threshold level of abundance (e.g., the number of species with less than the average proportional abundance $[1/S]$ in a community of S species). The extent to which each of these components of taxonomic biodiversity exhibits a latitudinal gradient has not been explored for most groups of plants, animals, or microbes. In contrast, the comprehensive analysis of Stevens and Willig (2002) characterized taxonomic biodiversity of New World bat communities based on 14 different indexes representing species richness (3 metrics), evenness (4 metrics), dominance (3 metrics), and diversity (4 metrics). They demonstrated that spatial variation in richness was independent of spatial variation in evenness. Moreover, they documented strong latitudinal gradients in species richness and diversity, but the absence of such gradients in species evenness or dominance (**Figure 5**).

Many factors (e.g., productivity, competition, predation, and disturbance) have been suggested as the dominant forces affecting the composition and structure of local communities. Early theoretical and empirical work stressed the role of deterministic factors such as competition in molding community attributes. Subsequent focus on the distinction between equilibrial and nonequilibrial communities cast doubt on the universality of deterministic mechanisms in general and competition in particular, and raised serious questions about the degree to which local communities were saturated by species. Indeed, variation in the degree to which local communities attain equilibrial richness may contribute strongly to the latitudinal gradient in biodiversity. The latitudinal progression from polar to tropical regions may represent a gradient in the degree to which stochastic density-independent mechanisms or biotic interactions dominate the forces affecting the abundance of local populations and the composition of local communities. Specifically, the species richness of a community may be a consequence of the severity, variability, and predictability of local environmental conditions. Low predictability and harsh conditions predispose communities to be regulated by abiotic parameters and to have low diversity. High predictability and conditions that are clement favor high diversity. To the extent that high solar insolation and warm temperatures represent favorable conditions, and low intra-annual variation in temperature and

rainfall represent predictable conditions, tropical communities should be more species rich than their extratropical counterparts. In essence, many of the factors that affect elevated richness at the local scale likely contribute to enhanced γ-diversity of regions as well.

Hierarchical Configuration of Biodiversity

The hierarchical configuration of biodiversity is an inherently scale-dependent phenomenon that influences the form and parameterization of latitudinal gradients, as well as the relationship between patterns at the local (α) and regional (γ) scale. Indeed, turnover among local sites in species composition contributes to the disparity between biodiversity at local and regional scales, is ostensibly related to habitat heterogeneity, and can be conceptualized via an additive model or a multiplicative model. Unfortunately, considerable confusion and controversy has characterized the discussion of such issues, but a consensus is emerging concerning the meaning and measurement of β-diversity in ecological and biogeographic contexts (see Tuomisto, 2010), with the multiplicative model rather than the additive model gaining ascendancy provided that metrics are adjusted to measure "true diversity" (*sensu* Jost, 2007). Moreover, such hierarchical decomposition can be applied to metrics other than species richness, such as species diversity, evenness, dominance, and rarity. In such a context, β-scale metrics may be considered to estimate the number of distinctive compartments or assemblages along the gradient.

Regardless of model, additive or multiplicative, the increase in species richness from polar to tropical latitudes for assemblages (γ-component) can be a consequence of a gradient in species richness at the α-level, β-level, or both. The β-component of species richness increases toward the equator for a variety of groups of mammals, birds, shallow water bryozoans, trees, and herbaceous plants (summarized in Willig *et al.*, 2003). The exact form of the increase may differ among taxa or between continents. For New World bats, the rate of increase in species richness from extratropical to tropical communities (α-level) is much less than the comparable rate for regional species pools (γ-scale), suggesting that the β-component consistently increases with decreasing latitude (Stevens and Willig, 2002). In contrast, the β-component of species richness remained more or less constant below 30° N latitude for nonvolant mammals in North America (Kaufman, 1998). Moreover, the turnover in species composition between quadrats within elevational bands was independent of latitude in both North and South America (Willig and Gannon, 1997).

Multiple Dimensions of Biodiversity

The past decade has witnessed increasing interest in understanding how functional, phylogenetic, genetic, or phenetic biodiversity responds to spatial, temporal, or environmental variation, and whether such gradients are any different than those produced by parallel gradients in species richness. Research on latitudinal gradients of multiple dimensions of biodiversity is in its infancy, but recent investigations have quantified these gradients, and have shown that they are of multiple forms and are not wholly explicable by gradients in richness. For ease of exposition, the authors illustrate the state of knowledge for latitudinal gradients of functional, phenetic, and phylogenetic biodiversity based on a comprehensive suite of analyses for 32 communities of New World bats (*see* Ecological Communities) for which a strong linear increase in species richness accompanies a decrease in latitude (**Figure 7(a)**).

The functional dimension of biodiversity can be evaluated based on the consideration of functional traits or on the assignment of species to guilds based on considerations of natural history (e.g., foraging mode and diet). Regardless, measures of functional biodiversity evaluate variation in characteristics of a biota that relate to ecosystem processes, species interactions, or functional characteristics of a community. A number of indexes of functional biodiversity can be estimated by weighting guilds by some aspect of importance such as relative species richness, abundance (regardless of species), or biomass (regardless of species), giving rise to guild evenness, diversity, or dominance. Nonetheless, the simplest index is guild richness, the empirical number of recognized guilds in a community. Like the latitudinal gradient of species richness, the latitudinal gradient of guild richness of New World bats (Stevens *et al.*, 2003) is strong, and generally increases from extratropical to tropical latitudes; however, the form of the relationship is markedly nonlinear (**Figure 7(b)**). Of course, species-rich communities should contain more guilds by chance alone (i.e., selection effects). Simulation analyses can evaluate if empirical gradients of guild richness are essentially a consequence of the latitudinal gradient of species richness. Such an approach randomly selects from an appropriate species pool the same number of species as occurs in an empirical community, repeats the random selection process for all empirical communities, estimates parameters of a randomly generated latitudinal gradient in guild richness, and repeats the simulation a sufficient number of times to create a probability density function of parameter values for a gradient created by chance alone. An observed gradient in guild richness is significant if empirical parameter estimates occur in the extreme tails of the generated probability density functions. New World bat communities exhibited a strong and significant latitudinal gradient of guild richness based on such approaches when the suite of bats in the continental New World composed the species pool for each community. In contrast, when each species pool for a community was regional in composition (i.e., each community had a unique species pool defined by the suite of taxa whose geographic distributions overlapped the site), the empirical latitudinal gradient of guild richness was not different than expected given the observed variation in the composition of regional species pools and the species richness of the community. These same approaches were used to assess latitudinal gradients of guild evenness, diversity, and dominance. Guild diversity and dominance significantly, but nonlinearly, increased and decreased toward the tropics, respectively. Like the situation for guild richness, the gradients in guild diversity and dominance were different than those arising from chance alone when considering hemispheric species pools, but no different than expected by chance alone when considering the regional species pool. In contrast, guild evenness evinced

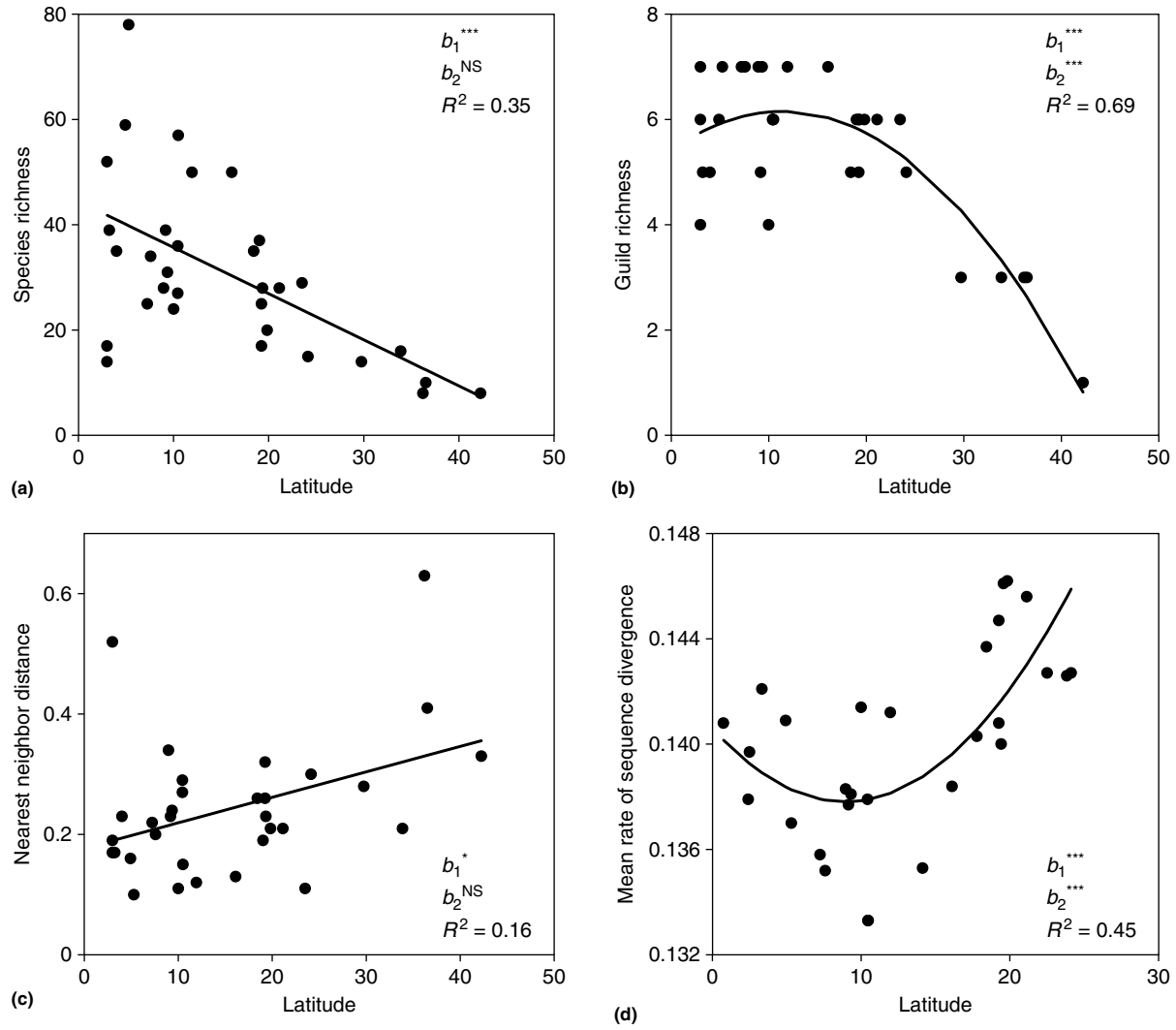


Figure 7 Latitudinal gradient of taxonomic biodiversity (species richness, a), functional biodiversity (guild richness, b), phenetic biodiversity (nearest neighbor distance, c), and phylogenetic biodiversity (mean rate of sequence divergence, d) within local communities for bats from throughout the New World. Reproduced from Stevens RD and Willig MR (2002) Geographical ecology at the community level: Perspective on the diversity of New World Bats. *Ecology* 83: 545–560; Stevens RD, Cox SB, Strauss RE, and Willig MR (2003) Patterns of functional diversity across an extensive environmental gradient: Vertebrate consumers, hidden treatments and latitudinal trends. *Ecology Letters* 6: 1099–1108; Stevens RD, Willig MR, and Strauss RE (2006) Latitudinal gradients in the phenetic diversity of New World bat communities. *Oikos* 112: 41–50, and Stevens RD (2006) Historical processes enhance patterns of diversity along latitudinal gradients. *Proceedings of the Royal Society B* 273: 2283–2289.

random variation with respect to latitude, a pattern expected due to chance, regardless of the nature of the species pool. Taken together, these results suggest that processes, evolutionary or ecological, that determine the composition of regional faunas strongly influence the latitudinal gradient in functional biodiversity.

The phenetic dimension of biodiversity can be evaluated based on considerations of a suite of phenotypic or morphological characteristics. Phenotypic characteristics of species have a profound effect on their ecological attributes, and provide an integrated view of ecological relationships over space or time. Different components of phenotypic biodiversity then relate to aspects of the volume (equivalent to richness) or dispersion (equivalent to evenness) of species within a multidimensional space defined by the phenotypic characteristics of species in a community. If the phenetic characteristics relate to foraging mode or diet, such an analysis could be construed as a trait-based assessment of functional biodiversity. More generally, the selection of phenotypic characteristics to be included in analyses determines the kind of phenetic biodiversity that is reflected in resultant components. By measuring a suite of external and cranial characteristics related to the size and foraging apparatus of bats (e.g., skull measurements related to the size and hardness of prey, or wing measurements related to aerodynamic ability), Stevens *et al*. (2006) characterized a number of aspects of the phenetic dimension of biodiversity. For example, they showed that mean nearest neighbor distances decreased significantly and

linearly toward the tropics (**Figure 7(c)**), and this trend was different from that expected given the richness of communities and the composition of the hemispheric species pool, but no different than expected by chance given the richness of communities and the composition of site-specific regional species pools. Again, historical factors affecting regional species pools have a great impact on gradients of phenetic biodiversity.

The phylogenetic dimension of biodiversity reflects the evolutionary history embodied in a community (e.g., the distribution of taxon ages or the magnitude of relatedness within a community). A variety of metrics then estimate different aspects of that evolutionary history, and each species contribution to those estimates can be weighted by the aspects of importance (e.g., relative abundance and relative biomass). Stevens (2006) estimated a number of different components of the phylogenetic dimension of biodiversity for New World phyllostomid bat communities. He documented a number of distinctive gradients based on the average and variance in the rate of sequence divergence, as well as on the average and variance in the relative ages of taxa. For example, the mean rate of sequence divergence in a community decreased from extratropical to tropical latitudes, but did so in a nonlinear manner (**Figure 7(d)**). Moreover, he showed that the latitudinal gradient in phylogenetic diversity was different from that expected given the latitudinal gradient in species richness and the composition of the hemispheric species pool. These results suggest that significant historical events influence contemporary gradients in biodiversity with the most derived and least variable taxa occurring at the periphery of the geographic range of the family (i.e., in extratropical areas), providing support for the concept of niche conservatism.

Mechanisms

A grand proliferation of hypotheses (**Table 1**), along with subsequent *a posteriori* modifications, is a characteristic feature of the literature concerning the relationship between biodiversity and latitude. A variety of factors contribute to this. Each of these hypotheses represents a conceptual model with only qualitative predictions. As with much of macroecological research, broad-scale data concerning the distribution of species is not available for many taxa. Manipulative experiments designed to disentangle the effects of competing hypotheses are not feasible or ethical. In addition, the inherent geographic factors that might affect richness are often correlated so that efforts to remove the effect of one to assess the other can lead to spurious results due to the confounded nature of the mechanisms.

Indeed, almost all hypotheses provide only a general qualitative prediction that richness should increase toward the tropics. The hypotheses represent conceptual models that lend insight into how nature could operate, but do not generate unique predictions based on direct features of the gradient that allow conclusive elimination of competing hypotheses in the sense of strong inference. Moreover, many hypotheses are circular in nature and the indirect predictions that they make about latitudinal gradients have not been examined comprehensively from an empirical perspective. In addition, some hypotheses only pertain to particular taxa or ecological groups, so that they are not applicable universally.

Rather than elucidating the score of extant hypotheses, which has been done to greater or lesser extents elsewhere (Rohde, 1992; Rosenzweig, 1995), an exposition of selected hypotheses that have generated debate in the contemporary literature follows. These hypotheses represent areas of research that will likely contribute to the complexion of the discipline in the future.

Geographic Area Hypothesis

The latitudinal gradient in which richness peaks in the tropics may be a consequence of the larger landmass of the tropics compared to other geographic zones. This simple idea had its genesis in the work of Terborgh (1973), with considerable development and refinement by Rosenzweig (1995) in subsequent years, during which the effects of productivity and zonal bleeding have been incorporated into a more comprehensive conceptual model. Nonetheless, the geographic area hypothesis has generated considerable controversy and contention (Rohde, 1997). The controversy does not surround whether an areal mechanism operates; it rather focuses on the degree to which variation in area is the dominant factor molding latitudinal gradients in richness.

Two features of the earth's geometry predispose the sizes of tropical regions to be greater than those of their higher latitude counterparts. First, the earth is essentially a sphere. The distance between longitudinal meridians at the equator is greater than that elsewhere on the globe, and intermeridian distance decreases in a regular fashion toward the poles. Second, northern and southern tropical zones are adjacent, whereas the northern and southern variants of other latitudinally defined zones are isolated from each other. Nonetheless, the positions, sizes, and configurations of the earth's continents will affect the proportion of land or water in tropical versus extratropical regions, and this has varied over geological time as a consequence of plate tectonics. In addition, the number and breadth of zones used to subdivide latitude will affect the perception of areal dominance associated with the tropics. For example, a tripartite division (torrid, temperate, and polar) reveals that the tropics rank second in area to north temperate regions at the global scale, with considerable variation in the proportional area represented by the tropics among continents (approximately 38%, 12%, 80%, 41%, and 0% of America, Eurasia, Africa, Australia, and Antarctica, respectively). In contrast, finer resolution of zones to tropical, subtropical, temperate, boreal, and tundra indicates the areal predominance of tropical lands globally.

Most important, the degree of environmental variation within the tropics is less than that in other geographic zones, at least with respect to incident solar radiation and temperature. Specifically, a band of 50° centered on the equator evinces no or little change in mean annual temperature with latitude (constant at approximately 27 °C), whereas mean annual temperature decreases thereafter by approximately 0.75 °C per degree latitude. Hence, regardless of the size of zonal subdivisions chosen to define tropical or extratropical regions, tropical landmasses are larger than any other landmasses with similar variation in temperature.

As a consequence of the areal extent and homogeneity of temperature and solar insolation in the tropics, speciation

rates there should be higher and extinction rates lower than in extratropical regions. Specifically, the larger area of the tropics allows its species to have larger ranges than do their extratropical counterparts. Larger ranges allow species to be represented by more populations or populations of larger size, both diminishing the likelihood of extinction resulting from accidental cause or from environmental perturbation. Simultaneously, larger areas are more likely to contain or experience geological events that produce geographic barriers that enhance the rate of allopathic speciation. The dynamic balance between the rates of speciation and extinction, therefore, yield higher equilibrial richness in the tropics than in extratropical areas (**Figure 8**).

Rosenzweig and colleagues marshal many lines of evidence in support of the geographic area hypothesis by citing two kinds of observations. First, larger biotic provinces, regardless of latitude, have more taxa than do their smaller counterparts (e.g., generic, familial, and ordinal richness of mammals increase with provincial area). Second, diversities from the same biome but from different continents or provinces differ as a function of their areal extent (e.g., rain forest vertebrates and plants as well as tropical freshwater fish increase in richness as their areal extent increases). Situations in which the general pattern does not occur usually include large but unproductive climatic zones with few species – effectively the richness-diminishing effects of low productivity may countermand the dominant role of area in these systems. In contrast, Rohde (1997) considers area not to be the dominant factor that affects high species richness in the tropics. He illustrates the point with Eurasian freshwater fishes, and shows that much smaller tropical regions have much greater species richness than do larger cold-temperate regions. Similarly, the expansive deep-sea biome with more or less constant temperature has far fewer species than its smaller tropical counterparts. Clearly, consensus is elusive concerning the relative importance of area in affecting species richness compared to other mechanisms.

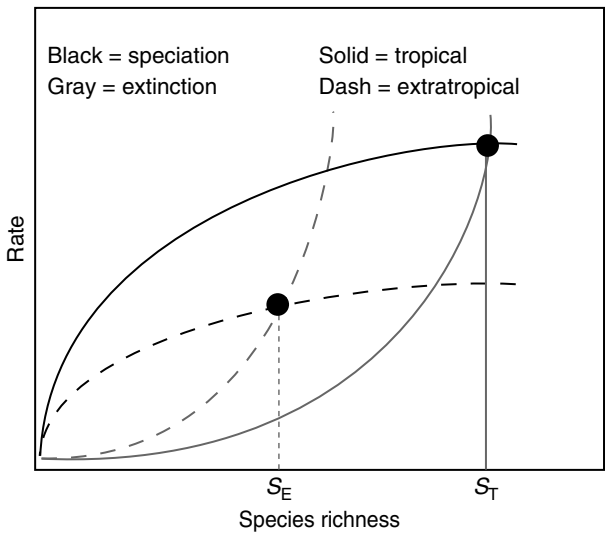


Figure 8 Graphical model illustrating the dynamic equilibrium between rates of speciation and extinction for tropical and extratropical latitudes. Modified from Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, MA: Cambridge University Press.

Because bats exhibit exceptionally strong latitudinal gradients of richness at multiple spatial scales in the New World, they are an appropriate biota with which to test the geographic area hypothesis (see Willig and Bloch, 2006). Importantly, ecogeographic zones of the New World (biome types and their constituent biogeographic provinces, *sensu* Udvardy, 1975) are not larger at tropical versus extratropical latitudes, and the spillover of tropical species into ecogeographic zones within extratropical regions generally does not diminish the association between richness and area. Nonetheless, the latitudinal gradient of species richness is strong and significant at both ecogeographic scales for all bats, as well as for the three most species-rich families (Molossidae, Phyllostomidae, and Vespertilionidae). Area does not drive the latitudinal gradient of bat species richness in the New World. In fact, area represents a source of noise rather than a dominant driver of the pattern.

Evolutionary Speed

After a broad and incisive review of the various mechanisms purported to cause latitudinal gradients in species richness, Rohde (1992) found them all to be lacking. Instead, he suggested that the gradient was a consequence of differential rates of speciation associated with an important latitudinal correlate, temperature, rather than being a product of equilibrium-based ecological processes that presupposed that local communities are saturated with species. His conceptual model is erected on the foundation of three premises. First, tropical environments support shorter generation times for many homeotherms and poikilotherms. Second, mutation rates increase as temperature increases and are highest in the tropics. Third, faster physiological processes occur at higher temperatures; this, coupled with the first two relationships, suggests an accelerated rate of fixation of favorable alleles in tropical populations. This effectively results in greater evolutionary time in the tropics for mechanisms of diversification to attain fruition.

The potential effect of evolutionary speed on latitudinal gradients of species richness has been evaluated for taxa representing terrestrial and aquatic ecosystems as well as vertebrate and invertebrate taxa. Many studies found support for increased diversification rates contributing to latitudinal gradients of species richness; however, evidence that the speed of molecular evolution is an important driver of those gradients is equivocal. For planktonic foraminiferans, temperature affects per individual speciation rates, contributing to the latitudinal gradient of species richness (Escarguel *et al.*, 2008). In contrast, there was no relationship (direct or indirect) between latitude and rates of molecular evolution in birds (Bromham and Cardillo, 2003). Rather than particular environmental conditions, rapid shifts between climatic regimes cause greater species richness for birds in the tropics (Kozak and Wiens, 2010), with sister species being more climatically divergent in the tropics and niche conservatism being the rule in temperate zones (Kozak and Wiens, 2007). The richness gradient for basal bird taxa is indistinguishable from that for all birds,

whereas the richness gradient for derived taxa is much shallower (Hawkins *et al.*, 2006). This suggests that climatic dynamics through evolutionary time and effective evolutionary time, rather than evolutionary speed per se, may contribute to latitudinal gradients of species richness.

Rapoport–Rescue Hypothesis

As its name implies, this hypothesis is a hybrid of two mechanisms operating in tandem: the Rapoport effect and the rescue effect. A geographic pattern in which species range size decreases from high to low latitudes came to the forefront of the macroecological literature a few decades ago (Stevens, 1989), and has been termed Rapoport's rule after the Argentinian scientist who first discussed the pattern in the context of many other areographic principles. Stevens hypothesized that the latitudinal propensity for range size to decrease toward the tropics, when combined with differential movement of individuals from source to sink habitats (rescue or mass effect), can generate the latitudinal gradient of diversity.

Specifically, at any one locale in the temperate zone, an individual must be able to tolerate considerable intra-annual variation in climatic conditions; thus, species that occur in the temperate zone can attain a wide latitudinal distribution because of the broad tolerance of its constituent individuals to varying local conditions. In contrast, an individual in the tropics experiences little seasonal variation in climatic conditions; consequently, species comprising individuals that occur in tropical zones are predisposed to have narrower latitudinal distributions. This creates the Rapoport effect.

The rescue effect is a phenomenon whereby local extinction of a population, often in marginal or sink habitats, is prevented because of immigration of individuals from source or high-quality habitats. Because smaller ranges, which are differentially situated in the tropics as a consequence of the Rapoport effect, have greater perimeter to area ratios, they are predisposed to having greater rescue effect areas relative to range areas. This differentially inflates species richness in tropical areas, generating the latitudinal gradient of diversity.

The generality of Rapoport's rule, as well as the degree to which empirical patterns are generated by the hybrid mechanisms embodied in rescue and Rapoport effects, are controversial. The Rapoport effect has been documented for a diversity of taxa (mammals, reptiles, amphibians, fish, crayfish, amphipods, mollusks, and trees) in aquatic and terrestrial environments, and quickly became engrained as the explanation for species diversity gradients in a variety of ecology textbooks. Additional circumstantial evidence was derived from the observation that taxa, which do not show the general latitudinal gradient in richness, do not adhere to Rapoport's rule, suggesting that both patterns had a shared mechanistic basis (Stevens, 1989). Nonetheless, a growing body of evidence suggests that the Rapoport pattern is far from universal (Rohde *et al.*, 1993; Lyons and Willig, 1997). Moreover, reanalyses of data on marine mollusks that was used to corroborate the Rapoport effect (Stevens, 1989) failed to produce the same patterns in a subsequent study, even though the methods were the same in both studies (Roy *et al.*, 1994). In addition, New World bats and marsupials (Lyons and Willig, 1997; Willig and Lyons, 1998), as well as nonmigratory marine teleosts from surface waters (Rohde *et al.*, 1993), each exhibit strong latitudinal gradients in diversity but do not adhere to Rapoport's rule. Hence, occurrences of latitudinal gradients in diversity do not have one-to-one correspondence with the existence of a Rapoport effect.

Simulation models provide added insight into the phenomenon of Rapoport's rule. The three commonly used methods (i.e., Stevens', midpoint, and most distal point) for assessing a Rapoport effect suffer from serious limitations. Stevens' method is problematic because of a lack of independence associated with counting the same species multiple times in the same analysis (Rohde *et al.*, 1993). Midpoint and most distal point methods suffer from severe mathematical biases – the bounded nature of continents or oceans predisposes correlations between range size and latitude even when ranges are distributed stochastically with respect to latitude (Colwell and Hurtt, 1994; Lyons and Willig, 1997). Finally, a comprehensive set of simulation models (Taylor and Games, 1999) suggests that the Rapoport effect causes a latitudinal pattern in species richness, but the gradient is opposite of the pervasive pattern found in nature in that species richness increases with increasing latitude. Moreover, incorporation of a rescue effect into the model so that it reflects the Rapoport–rescue mechanism still fails to rescue the hypothesis; the predicted pattern remains a decrease in richness toward the tropics. Only the incorporation of competitive effects either to the simulation model based on the Rapoport effect alone or to the combined Rapoport–rescue mechanism produces latitudinal gradients that are consistent with real-world patterns, and in both scenarios, communities must be saturated at equilibrial compositions.

Geometric Constraints Hypothesis

One of the newer ideas to explain latitudinal gradients in biodiversity is the geometric constraints hypothesis (Colwell and Hurtt, 1994; Willig and Lyons, 1998). Geographic constraints may affect patterns of species distribution within a domain, creating modal patterns of biodiversity in the middle of the domain (i.e., the mid-domain effect). Within the context of latitudinal gradients, this hypothesis would consider peak biodiversity in the tropics to be a consequence of the bounded nature of terrestrial and aquatic habitats. Indeed, both simulation (Colwell and Hurtt, 1994) and analytical (Willig and Lyons, 1998) null models suggest that species richness of a biota should increase toward the center of a shared geographic domain in a quasiparabolic or parabolic fashion as a consequence of the random placement of species ranges within the domain. Other mechanisms proposed to account for latitudinal gradients only suggest qualitative increases in richness with decreasing latitude. In contrast, geometric constraint models make quantitative predictions concerning the form of the latitudinal gradient so that expected values for richness occur for each latitude and can be compared to empirical data.

Multiple null models that differ in the number and type of constraints have been developed to evaluate this geometric

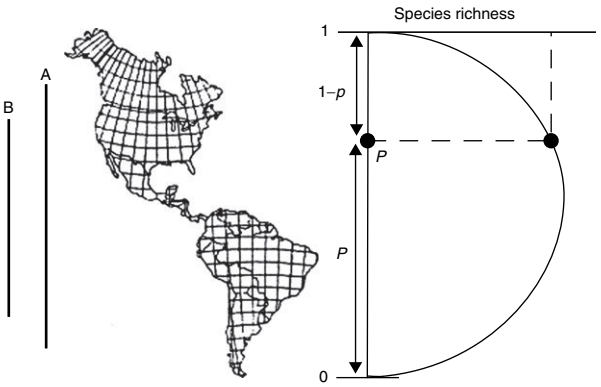


Figure 9 Graphical representation (parabola) of the gradient of species richness that arises from the random placement of species ranges within the latitudinal bounds of the New World, scaled from 0 in the south to 1 in the north. (Modified from Willig MR and Lyons SK (1998) An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. *Oikos* 81: 93–98.) The number of species at any latitude is determined by its proportional distance (p) from the southern boundary and is given by $(2p - 2p^2)S$. The vertical lines labeled A and B represent the latitudinal extents of bats and marsupials, respectively.

hypothesis. The fully neutral model (Colwell and Hurtt, 1994; Willig and Lyons, 1998) defines the expected pattern of species richness while ignoring potential effects of extant environmental gradients, as well as the history of the domain, on species' ranges sizes. Other null models apply additional constraints on randomizations, such as requiring the range size frequency distribution (RSFD) or the distribution of midpoints in the domain to equal the empirical distribution. For the fully neutral model, the number of species at any point within the domain (Sp) is related only to the proportional distance of that point from the boundary (p) and the number of species in the species pool (S), and is given by $Sp = 2p(1-p)S$. This model (**Figure 9**) is an incarnation of both the two-hit broken stick model of MacArthur and the binomial distribution. In essence, if the latitudinal domain of a biota is rescaled to range from 0 in the south to 1 in the north, then the likelihood of a species range overlapping any point P that is exactly p from the southern terminus (and hence $1-p$ from the northern terminus) of the domain is

$$Pr(P) = 1 - p^2 - (1-p)^2 = 2p - 2p^2$$

where p^2 is the proportion of species whose northern and southern boundaries lie to the south of P, and $(1-p)^2$ is the proportion of species whose northern and southern boundaries lie to the north of P. The functional form of the distribution of $Pr(P) = 2p - 2p^2$ is a parabola that peaks at 0.5, and as a consequence, the average range size of species in the biota is half the extent of the domain (0.5).

Distributions of range midpoints and RSFDs are the collective product of responses to environmental gradients, biotic interactions, and historical processes (Hawkins *et al.*, 2005). Consequently, null models that use such constraints incorporate empirical responses to environmental variation and historical processes into "null" patterns that are produced by

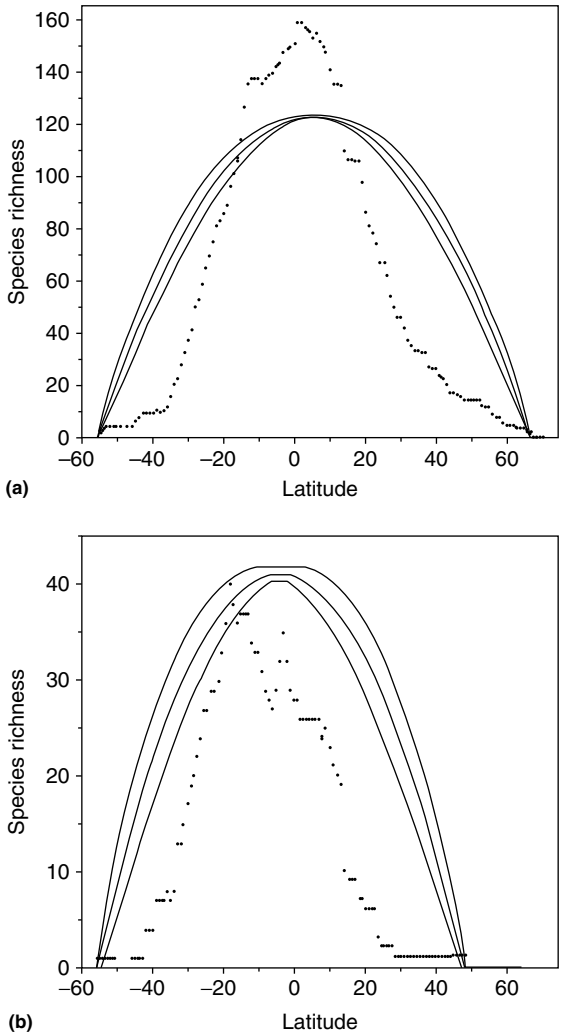


Figure 10 Comparison of the empirical latitudinal gradient of species richness (dots) with the predicted values (outer lines represent 95% confidence bands, and inner line represents predicted value) generated by a null model constrained only by the distributional bounds of the fauna for bats (a) and marsupials (b) in the New World. Modified from Willig MR and Lyons SK (1998) An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. *Oikos* 81: 93–98.

the randomization. Constrained models produce a quasi-parabolic curve that becomes increasingly flattened as mean range size diverges from 0.5 (the mean value in the fully neutral model). If the mean range size is >0.5, species richness in the middle of the domain will be higher and the peak broader than that for a comparable fully neutral model. Alternatively, a RSFD with a mean range size <0.5 will have a richness peak that is lower and broader than that of a RSFD with a mean range size of 0.5 (Colwell *et al.*, 2004).

Considerable debate exists about the application, utility, and interpretation of geometric constraint null models with respect to empirical data (e.g., Colwell *et al.*, 2004, 2005; Zapata *et al.*, 2003, 2005; Hawkins *et al.*, 2005). Importantly, these null models provide baseline expectations for gradients of species

richness in domains in the absence of ecological and evolutionary mechanisms, for the fully neutral model, or given particular distributions of midpoints or species range sizes, for models that incorporate additional biological constraints. Consequently, deviations from null model predictions should be the focus of interpretation for empirical applications of geometric constraint models. For example, a geometrically constrained null model was able to account for 69–95% of the latitudinal variation in species richness for New World bats and marsupials (**Figure 10**; Willig and Lyons, 1998), indicating that latitudinal richness gradients for these taxa approximate null expectations for models that incorporate empirical RSFD, and the ecological and evolutionary mechanisms they represent. Nonetheless, systematic deviations from the null distribution were observed for each taxon, and these deviations represent important foci for further inquiry into mechanisms that mold latitudinal gradients of biodiversity.

Geometric constraint models make predictions (i.e., estimates based on null models) for multiple aspects of latitudinal richness gradients, including the precise shape of the symmetrical curve, as well as the size and location of the single peak in species richness. A comprehensive analysis of geometric constraints on gradients associated with latitude, longitude, ocean depth, or elevation found poor congruence between empirical patterns and those generated by geometric constraint models (Zapata *et al.*, 2003). Nonetheless, this exercise is valuable as systematic deviations from null model predictions can generate more specific hypotheses about mechanisms that shape gradients of species richness. In addition to one-dimensional models, two-dimensional models based on latitudinal and longitudinal constraints have been used to more fully evaluate the potential for geographic constraints to mold gradients of richness. Two-dimensional models fit empirical data more poorly than do one-dimensional models associated with latitude (Zapata *et al.*, 2003; Kerr *et al.*, 2006). Geometric constraints provide the greatest insights when the axis of constraint that defines the domain is associated with important environmental gradients, as is the case for latitude, elevation, and ocean depth, when deviations from predictions suggest possible causal mechanisms.

A geometric constraint approach can enhance the understanding about the relative contributions of evolutionary, environmental, and historical factors to latitudinal gradients of biodiversity. Random processes may predispose particular biotas to produce gradients with peaks in richness at midlatitudes. The challenge for large-scale ecology is to understand the mechanisms that result in deviations from such null models, create peaks in biodiversity, and define the form of the biodiversity–latitude relationship.

Assessment and Synthesis

The ontogeny of theory can be viewed from a variety of perspectives that deal with the detection of patterns, the linkage of patterns to particular mechanisms, and ultimately the integration of those constructs to other theories in the discipline (Scheiner and Willig, 2011). The theory of latitudinal gradients of biodiversity has matured considerably since the first edition of this book. The general patterns of latitudinal increase in species richness remain well documented from an empirical perspective. Recent meta-analyses on more than 500 published latitudinal gradients (Hillebrand, 2004) reaffirmed the generality of the established pattern based on considerations of effect sizes (slopes and coefficients of determination). This analysis established that patterns were stronger and steeper for assemblages (large grains) than for communities or samples at the local level (small grains), but that the latitudinal extent of the domain did not affect the nature of the gradient. The relationships in marine and terrestrial environs were stronger and steeper than those in freshwater environs, and the gradients were different among continents but not between northern and southern hemispheres. Body mass, trophic level, and thermoregulatory capacity of the targeted biota had an effect on nature of gradients as well.

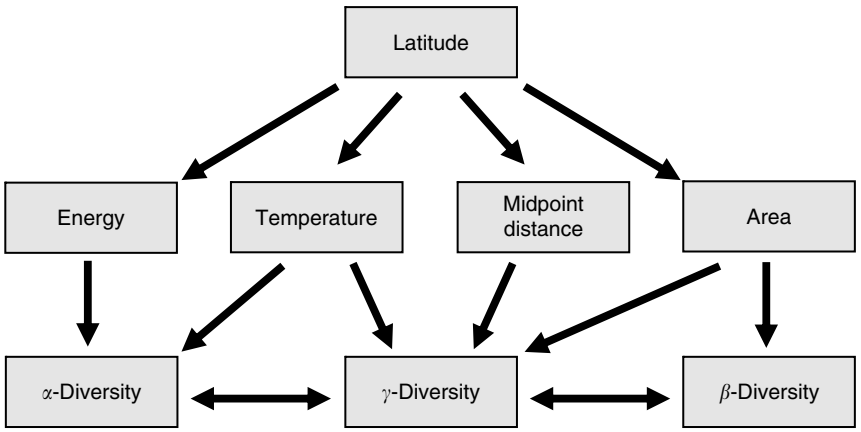


Figure 11 Conceptual model (modified from Willig MR, Kaufman DM, and Stevens RD (2003) Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annual Review of Ecology Evolution and Systematics* 34: 273–309) indicating likely mechanisms whereby latitude may affect alpha- (α), beta- (β), or gamma- (γ) components of biodiversity. Latitude is a surrogate for correlated spatial or environmental characteristics (e.g., energy, temperature, and midpoint distance in the domain, area) with which it covaries. These spatial and environmental covariates figure centrally in dominant hypotheses associated with the latitudinal gradient in species richness.

The manner in which particular mechanisms could affect patterns of species richness has become clearer, and elements of the theory have been used to understand other gradients of biodiversity, such as those related to elevation, depth, or productivity. Moreover, latitudinal patterns of species richness are being integrated with other broad-scale patterns concerning the dominance and turnover of species, as well as to latitudinal patterns of range size and abundance, representing a significant advancement in understanding and integration. The past decade has seen renewed interest in evolutionary and historical mechanisms in addition to those of a more ecological nature.

Although few of the hypotheses postulated to affect the latitudinal gradient in species richness have been eliminated in a conclusive manner, research concerning many of the mechanisms appears more likely to advance theory in the near future. Indeed, recent synthetic works have focused on them to a great extent. Understanding the contexts and degrees to which area, climatic variability or stress, geographic constraints, productivity, temperature, and their interactions mold the latitudinal gradient in diversity remains a challenge for scientists to synthetically address in the coming decade. For example, the synthesis by Willig *et al.* (2003) provides a conceptual model (**Figure 11**) of how latitude as a surrogate for considerations of energy, temperature, geometric constraints, and area (hypothetical mechanisms for the latitudinal gradient) differentially affects species richness at different focal scales as well as the association between them.

In contrast to the situation regarding species richness, the state of theory regarding latitudinal gradients of other aspects of the taxonomic dimension (e.g., evenness, diversity, and dominance) or of other dimensions of biodiversity (e.g., functional and phylogenetic) is in its infancy. General latitudinal patterns for aspects of biodiversity other than species richness have not been well established. Moreover, the mechanistic bases of these more complex gradients have been explored only in a rudimentary manner. Nonetheless, it is clear in the limited number of cases where they have been explored that latitudinal gradients in nontaxonomic dimensions of biodiversity are not wholly explicable by gradients in species richness and that historical factors may play a significantly important role.

Appendix

List of Courses

1. Introductory Biology
2. Ecology
3. Biogeography
4. Evolution
5. Environmental Studies

See also: Biodiversity-Rich Countries. Biogeography, Overview. Hotspots. Species–Area Relationships. Tropical Forest Ecosystems

References

Bromham L and Cardillo M (2003) Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. *Journal of Evolutionary Biology* 16: 200–207.

Colwell RK and Hurtt GC (1994) Nonbiological gradients in species richness and a spurious Rapoport effect. *American Naturalist* 144: 570–595.

Colwell RK, Rahbek C, and Gotelli NJ (2004) The mid-domain effect and species richness patterns: What have we learned so far? *American Naturalist* 163: E1–E23.

Colwell RK, Rahbek C, and Gotelli NJ (2005) The mid-domain effect: There's a baby in the bathwater. *American Naturalist* 166: E149–E154.

Currie DJ (1991) Energy and large-scale patterns for animal- and plant-species richness. *American Naturalist* 137: 27–49.

Escarguel G, Brayard A, and Bucher H (2008) Evolutionary rates do not drive latitudinal diversity gradients. *Journal of Zoological Systematics and Evolutionary Research* 46: 82–86.

Fischer AG (1960) Latitudinal variations in organic diversity. *Evolution* 14: 64–81.

Hawkins BA, Deiniz-Filho JAF, Jaramillo CA, and Soeller SA (2006) Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World Birds. *Journal of Biogeography* 33: 770–780.

Hawkins BA, Deiniz-Filho JAF, and Weis AE (2005) The mid-domain effect and diversity gradients: Is there anything to learn? *American Naturalist* 166: E140–E143.

Hillebrand H (2004) On the generality of the latitudinal diversity gradient. *American Naturalist* 163: 192–211.

Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88: 2427–2439.

Kaufman DM (1995) Diversity of New World mammals: Universality of the latitudinal gradients of species and bauplans. *Journal of Mammalogy* 76: 322–334.

Kaufman DM (1998) *The structure of Mammalian Faunas in the New World: From Continents to Communities*. Unpublished PhD Dissertation, University of New Mexico, Albuquerque.

Kerr JT, Perrign M, and Currie DJ (2006) The missing Madagascan mid-domain effect. *Ecology Letters* 9: 149–159.

Kozak KH and Wiens JJ (2007) Climatic zonation drives latitudinal variation in speciation mechanisms. *Proceedings of the Royal Society B* 274: 2995–3003.

Kozak KH and Wiens JJ (2010) Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters* 13: 1378–1389.

Lyons SK and Willig MR (1997) Latitudinal patterns of range size: Methodological concerns and empirical evaluations for New World bats and marsupials. *Oikos* 79: 568–580.

Lyons SK and Willig MR (2002) Species richness, latitude, and scale-sensitivity. *Ecology* 83: 47–58.

Rohde K (1992) Latitudinal gradients in species diversity: The search for the primary cause. *Oikos* 65: 514–527.

Rohde K (1997) The larger area of the tropics does not explain latitudinal gradients in species diversity. *Oikos* 79: 169–172.

Rohde K, Heap M, and Heap D (1993) Rapoport's rule does not apply to marine teleosts and cannot explain the latitudinal gradient in species richness. *American Naturalist* 142: 1–16.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, MA: Cambridge University Press.

Roy K, Jablonski D, and Valentine JW (1994) Eastern Pacific molluscan provinces and the latitudinal diversity gradient: No evidence for "Rapoport's rule". *Proceedings of the National Academy of Sciences, USA* 91: 8871–8874.

Scheiner SM, Chiarucci A, Fox GA, Helmus MR, McGlinn DJ, and Willig MR (2011) The underpinnings of the relationship of species richness with space and time. *Ecological Monographs* 81: 195–213.

Scheiner SM and Willig MR (eds.) (2011) *Theory of Ecology*. Chicago, Illinois: University of Chicago Press.

Stevens GC (1989) The latitudinal gradient in geographical range: How so many species coexist in the tropics. *American Naturalist* 133: 240–256.

Stevens RD (2006) Historical processes enhance patterns of diversity along latitudinal gradients. *Proceedings of the Royal Society B* 273: 2283–2289.

Stevens RD, Cox SB, Strauss RE, and Willig MR (2003) Patterns of functional diversity across an extensive environmental gradient: Vertebrate consumers, hidden treatments and latitudinal trends. *Ecology Letters* 6: 1099–1108.

Stevens RD and Willig MR (2002) Geographical ecology at the community level: Perspective on the diversity of New World Bats. *Ecology* 83: 545–560.

Stevens RD, Willig MR, and Strauss RE (2006) Latitudinal gradients in the phenetic diversity of New World bat communities. *Oikos* 112: 41–50.

Taylor PH and Games SD (1999) Can Rapoport's rule be rescued? Modeling causes of the latitudinal gradient in species richness. *Ecology* 80: 2474–2482.

Terbrogh J (1973) On the notion of favorableness in plant ecology. *American Naturalist* 107: 481–501.

Tuomisto H (2010) A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33: 2–22.

Udvardy MDF (1975) *A Classification of the Biogeographical Provinces of the World*. IUCN Occassional Paper 18. Morges, Switzerland: International Union for Conservation of Nature and Natural Resources.

Willig MR and Bloch CP (2006) Latitudinal gradients of species richness: A test of the geographic area hypothesis at two ecological scales. *Oikos* 112: 163–173.

Willig MR and Gannon MR (1997) Gradients of species density and turnover in marsupials: A hemispheric perspective. *Journal of Mammalogy* 78: 756–765.

Willig MR, Kaufman DM, and Stevens RD (2003) Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annual Review of Ecology Evolution and Systematics* 34: 273–309.

Willig MR and Lyons SK (1998) An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. *Oikos* 81: 93–98.

Zapata FA, Gaston KJ, and Chown SL (2003) Mid-domain models of species richness gradients: Assumptions, methods and evidence. *Journal of Animal Ecology* 72: 677–690.

Zapata FA, Gaston KJ, and Chown SL (2005) The mid-domain effect revisited. *American Naturalist* 166: E144–E148.

Life Cycle Analysis of Biofuels

Jason Hill, University of Minnesota, St Paul, MN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biofuel Liquid transportation fuels derived from plants or other biological materials.
Carbon footprint Life cycle analysis in which climate change is the impact category of concern.
Functional unit Quantity of reference for comparing two or more products based on their equivalent use.
Life cycle analysis Method of quantifying the suite of environmental effects associated with the production and use of a given product.
Market-mediated effect Changes in the supply and demand of other products due to the production of a given product.
Supply chain Those stages in the production and use of a given product through which materials and energy flow and can be traced.
System boundary The set of processes under consideration in a given life cycle analysis.

Introduction

Life cycle analysis (LCA), also known as *life cycle assessment*, is a method of quantifying the environmental impact of products such as biofuels. In LCA, researchers prepare an inventory of the resources used (e.g., fossil fuels and raw materials) and the substances generated (e.g., greenhouse gases, solid waste, and other pollutants) throughout the full life cycle of the production, transport, and use of the product of interest. From this inventory, researchers prepare an impact assessment that estimates the ultimate effects on human health, ecosystem function, and natural resource depletion. Because LCA considers the entire product life cycle, it is useful both in identifying major sources of environmental impact in a product's supply chain and in mitigating them (Graedel, 1996; Curran, 2008).

The first LCA study is widely held to have been conducted in 1969 by Teasley, who was examining different beverage container options for the Coca-Cola Company (Hunt and Franklin, 1996). To this day, LCA is predominantly used as a comparative tool by corporations, marketers, regulators, and consumers for making informed choices. Liquid transportation biofuels are typically compared against conventional petroleum-derived fuels such as gasoline, diesel, and aviation fuels (Farrell *et al.*, 2006; Stratton *et al.*, 2011). Comparisons are also made between biofuels and alternative uses for biomass such as bioelectricity (Campbell *et al.*, 2009; Ohlrogge *et al.*, 2009).

Types

Various methods for conducting LCA exist, each of which has particular strengths and limitations. One major albeit subtle distinction lies in whether the goal of an LCA study is to assign responsibility for environmental impacts to a given product or whether the goal is to estimate the net change in environmental impacts as a result of a decision to produce a product (Finnveden *et al.*, 2009). The first of these, commonly known as an *attributional approach*, takes the total environmental impact of human activity and seeks to ascribe responsibility for a portion of it to a given product, which may include its entire supply chain or selected parts of it (Weidema, 2003). The second, commonly known as a *consequential approach*, takes the total environmental impact of human activity and seeks to understand how it expands and contracts as a result of producing and using a given product (Earles and Halog, 2011). An attributional LCA is, therefore, arbitrarily defined as a function of how system boundaries are drawn, typically limited by space, time, or stage in a supply chain. It is of primary use to engineers looking to improve process efficiencies or to marketers comparing among purchase options. Alternatively, consequential LCA has infinite system boundaries, which is essentially the same as saying there are no system boundaries. It is strictly defined as the entire suite of direct and indirect impacts, regardless of where or when they occur. It is of primary use to policy makers and regulators who are concerned with the cumulative and comprehensive effects of product decisions.

Attributional LCA is typically conducted using process-based models, input–output models, or a hybrid of the two. Process-based models measure environmental flows along a given supply chain, including processes flowing therein and therefrom. Process-based models offer a fine resolution and product-specific results but typically have narrow system boundaries so as to be tractable and representative of the supply chain under which corporate control may be exerted. By contrast, input–output models utilize data both on trade among sectors and the environmental impact of each of these sectors to allocate environmental impact to a product produced in a given sector (Leontief, 1970; Lave *et al.*, 1995). These models offer much broader system boundaries (e.g., the entire economy of a nation) but present a coarser view, leaving the impact of a product unresolved within a given sector. Hybrid approaches also exist that supplement product specific data from process-based models with data from the expanded system boundaries of input–output models (Suh and Huppes, 2005).

Consequential LCA is methodologically similar to hybrid attributional LCA but with some key differences (Ekvall and

Weidema, 2004). First, consequential LCA takes a dynamic view of the world unlike attributional LCA, which takes a static view. As such, the physical flows in consequential LCA are described using marginal data, whereas attributional LCA uses average data. Second, consequential LCA incorporates economic modeling to describe changes in environmental impact that occur beyond the direct supply chain and use of a given product, including those that occur as a result of market-mediated effects such as economies of scale, substitution effects, and elasticities of supply and demand (Earles and Halog, 2011; Girod *et al.*, 2011). Conversely, attributional LCA is not concerned with economic modeling beyond either allocation of material flows to coproducts or, as is specific to input–output and hybrid approaches, descriptions of existing trade flows. Third, attributional LCA describes a single state of environmental impact (i.e., the world at a given point in time), but consequential analysis by its very nature requires more than one scenario, typically a world in which an item is not produced (i.e., a baseline) and a world in which it is (i.e., the baseline plus the delta). Overall, consequential LCA is inherently more difficult not only because it requires marginal data and economic modeling, but also because it requires analysis of at least one scenario that has not occurred. For retrospective consequential LCA, which is LCA looking at past decisions, this would be the counterfactual (i.e., that which did not occur, whether it be with or without a given product). For prospective consequential LCA, which is LCA used as a predictive tool for estimating future impacts of current or future decisions, this would be at least two scenarios (i.e., the world with a given product and the world without it).

Standardization

Extensive efforts have been made to standardize LCA methodologies. Still, the need to maintain flexibility in its application by different stakeholders is well recognized. Standardization has therefore focused primarily on defining commonly used terminology and ensuring that essential elements of LCA are reported consistently and transparently. The commonly used ISO 14040 series defines LCA as consisting of four phases: goal and scope definition, inventory analysis, impact assessment, and interpretation (ISO, 2006a, b). In the first phase, the goal and scope definition, the system to be studied is identified and described, the system boundary is established, and critical assumptions and limitations are stated. The choice of a functional unit is also explained so as to ensure a proper basis for comparison among different products based on how they are used. For example, a comparison between bar soap and liquid soap should not be made on each package's weight or volume but rather on the basis of an equal number of hand washings. In the case of biofuels, comparisons between ethanol and gasoline are typically made on a per unit energy (e.g., megajoule (MJ) or British thermal unit (Btu)) or per distance driven (e.g., kilometer or mile) basis rather than on a volumetric basis (e.g., liter or gallon) because ethanol has approximately two-thirds the energy density of gasoline (Liska and Cassman, 2008; Cherubini and Strømman, 2011).

The second phase, the inventory analysis, consists of gathering and compiling data on the physical flows of materials and energy throughout the product life cycle as defined by the chosen LCA methodology. In attributional LCA, this includes flows through the supply chain; in consequential LCA, this includes both supply chain and market-mediated flows. In this phase, decisions regarding the appropriate allocation of multiple product streams originating from a single process are also made. In the case of biofuels, this includes the coproduction of ethanol and animal feed from the corn fermentation process or the coproduction of ethanol and electricity from the sugarcane ethanol production process.

The third phase, the impact assessment, seeks to quantify the environmental effects of the changes in physical flows resulting from any supply chain and any market-mediated changes in material and energy as estimated in the inventory analysis. First, impact categories are chosen, which may include climate change, acidification, eutrophication, biodiversity, human health, fossil fuel depletion, and land and water use, among many others (Bare *et al.*, 2003; Curran *et al.*, 2011). Second, characterization models are used to estimate the effect of the inventory analysis on these impact categories (De Schryver *et al.*, 2011). Third, these effects may be compared via normalization, grouping, and weighing. Such comparisons may be useful for assessing trade-offs but may involve value judgments (Gloria *et al.*, 2007).

The fourth phase, interpretation, is an ongoing process throughout all the other three phases. Checks for completeness, sensitivities, and consistencies are to be performed, and any conclusions and recommendations are to be stated.

History and Application

Early application of LCA to biofuels grew from a desire to verify potential reductions in fossil fuel use. The 1970s energy crisis spurred great interest in finding alternatives to petroleum, particularly for use in transportation. Early work such as that of Chambers *et al.* (1979) therefore focused on the fossil energy balance, also known as the *energy return on investment*, of biofuels. Studies also considered the related metric of their petroleum balance, which essentially measures the amount of usable liquid transportation fuel returned from biofuel production. Interest in biofuels as a means of increasing domestic liquid fuel production continues to this day (Hammerschlag, 2006).

More recent application of LCA to biofuels has increasingly focused on environmental impacts beyond fossil resource depletion. Climate change has received particular attention. An LCA in which greenhouse gases inventories are prepared and expressed in CO_2 equivalents is commonly known as *greenhouse gas accounting*, *carbon accounting*, or *carbon footprinting* (Peters, 2010). Estimating net greenhouse gas emissions relative to conventional alternatives has dominated the use of LCA in biofuels research for two main reasons: (1) recent interest in climate change mitigation has been intense, and (2) the fossil fuel inventories (e.g., total use of coal, petroleum, and natural gas) generated when calculating net energy balances of biofuels can readily be translated to greenhouse gas inventories of major climate-forcing species

of interest, including carbon dioxide (CO_2), nitrous oxide (N_2O), and methane (CH_4). Beyond fossil fuel use, additional sources of greenhouse gas emissions that must be tracked when calculating greenhouse gas inventories to include N_2O emissions from nitrogenous fertilizer production and use (Howarth *et al.*, 2002; Hill *et al.*, 2006), CO_2 and CH_4 emissions from any changes in soil carbon storage in biomass production (Johnson *et al.*, 2007; Cherubini, 2010; Karlen *et al.*, 2011), and CO_2 emissions from any changes in land use that may occur (Fargione *et al.*, 2008).

Beyond fossil resource depletion and climate change, other major impact categories explored in biofuels using LCA include water use (Dominguez-Faus *et al.*, 2009; Wu *et al.*, 2009; Chiu *et al.*, 2011; Scown *et al.*, 2011; Yeh *et al.*, 2011) and air quality (Hill *et al.*, 2009; Huo *et al.*, 2009; Tsao *et al.*, 2011). Biofuels examined in these impact categories include first-generation biofuels produced from feed and food crops such as corn and sugarcane, second-generation biofuels (e.g., ethanol produced from lignocellulosic biomass), and advanced biofuels (e.g., hydrocarbons produced from algae) (Wu *et al.*, 2006; Huo *et al.*, 2009; Fargione *et al.*, 2010; Roberts *et al.*, 2010; Brentner *et al.*, 2011; Savage, 2011). Analyses suggest that in general the biofuels with the most beneficial impact will include those produced from crop and forestry wastes and residues, as well as from dedicated perennial crops grown with low agricultural inputs on marginal or abandoned croplands (Tilman *et al.*, 2009; Spatari and MacLean, 2010; Wiens *et al.*, 2011).

Use in Policy and Regulation

Recent years have seen an increase in a desire to apply LCA in biofuel policy and regulation (DeCicco, 2011). The renewable fuel standard of the US Energy Independence and Security Act of 2007, for example, sets greenhouse gas reduction levels for biofuel consumption volumes. Triennial reports to the US Congress describing the effects of increased biofuel production on other impact categories are also mandated. LCA is also employed in the regulation of the California low-carbon fuel standard (Sperling and Yeh, 2009; Boies *et al.*, 2011). Use of LCA in policy making and regulation has engendered a shift away from attributional approaches toward consequential approaches, where the full set of impacts resulting from a decision is taken into account. To date, studies that have taken a full consequential LCA approach to evaluating biofuels are lacking. Many studies utilizing attributional LCA have begun to incorporate elements of consequential LCA so as to approximate environmental changes that may be expected under increased biofuel production. For example, models have been adjusted to include indirect effects commensurate with large increases in land use for biofuels and any resulting indirect land-use change (Wang *et al.*, 2007, 2011). Others have focused on the marginal displaced energy sources (Lemoine *et al.*, 2010; Yeh *et al.*, 2010) and elasticity in transportation fuel markets (De Gorter, 2010). Use of LCA in policy and regulation has also prompted inquiry into means of incorporating variability and uncertainty into the decision-making process (Mullins *et al.*, 2011; Venkatesh *et al.*, 2011).

Appendix

List of Courses

1. Industrial Ecology
2. Environmental Life Cycle Analysis
3. Global Change
4. Environmental Science
5. Landscape Ecology

See also: Agriculture, Industrialized. Agriculture, Nutrient Management, and Water Quality. Air Pollution. Biofuels and Biodiversity: The Implications of Energy Sprawl. Biofuels and Biodiversity, Wildlife Habitat Restoration. Freshwater Ecosystems, Human Impact on. Indirect Land Use and Greenhouse Gas Impacts of Biofuels. Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production

References

Bare JC, Norris GA, Pennington DW, and McKone T (2003) TRACI: The tool for the reduction and assessment of chemical and other environmental impacts. *Journal of Industrial Ecology* 6: 49–78.

Boies AM, McFarlane D, Taff S, Watts WF, and Kittelson DB (2011) Implications of local lifecycle analyses and low carbon fuel standard design on gasohol transportation fuels. *Energy Policy* 39: 7191–7201.

Brentner LB, Eckelman MJ, and Zimmerman JB (2011) Combinatorial life cycle assessment to inform process design of industrial production of algal biodiesel. *Environmental Science & Technology* 45: 7060–7067.

Campbell JE, Lobell DB, and Field CB (2009) Greater transportation energy and GHG offsets from bioelectricity than ethanol. *Science* 324: 1055–1057.

Chambers RS, Herendeen RA, Joyce JJ, and Penner PS (1979) Gasohol: Does it or doesn't it produce positive net energy? *Science* 206: 789–795.

Cherubini F (2010) GHG balances of bioenergy systems – overview of key steps in the production chain and methodological concerns. *Renewable Energy* 35: 1565–1573.

Cherubini F and Strømman AH (2011) Life cycle assessment of bioenergy systems: State of the art and future challenges. *Bioresource Technology* 102: 437–451.

Chiu Y-W, Suh S, Pfister S, Hellweg S, and Koehler A (2011) Measuring ecological impact of water consumption by bioethanol using life cycle impact assessment. *The International Journal of Life Cycle Assessment* 17: 16–24.

Curran MA (2008) Life-cycle assessment. In: Jorgensen SE and Fath BD (eds.) *Encyclopedia of Ecology (4 vols.): Human Ecology*, pp. 2168–2174. Oxford: Elsevier.

Curran M, de Baan L, de Schryver AM, *et al.* (2011) Toward meaningful end points of biodiversity in life cycle assessment. *Environmental Science & Technology* 45: 70–79.

DeCicco JM (2011) Biofuels and carbon management. *Climatic Change.* http://dx.doi.org/10.1007/s10584-011-0164-z.

De Gorter H (2010) Does US corn-ethanol really reduce emissions by 21%? Lessons for Europe. *Biofuels* 1: 671–673.

De Schryver AM, van Zelm R, Humbert S, *et al.* (2011) Value choices in life cycle impact assessment of stressors causing human health damage. *Journal of Industrial Ecology* 15: 796–815.

Dominguez-Faus R, Powers SE, Burken JG, and Alvarez PJ (2009) The water footprint of biofuels: A drink or drive issue? *Environmental Science & Technology* 43: 3005–3010.

Earles JM and Halog A (2011) Consequential life cycle assessment: A review. *The International Journal of Life Cycle Assessment* 16: 445–453.

Ekvall T and Weidema BP (2004) System boundaries and input data in consequential life cycle inventory analysis. *International Journal of Life Cycle Assessment* 9: 161–171.

Fargione J, Hill JD, Tilman D, Polasky S, and Hawthorne P (2008) Land clearing and the biofuel carbon debt. *Science* 319: 1235–1238.

Fargione JE, Plevin RJ, and Hill JD (2010) The ecological impact of biofuels. *Annual Review of Ecology, Evolution, and Systematics* 41: 351–377.
Farrell AE, Plevin RJ, Turner BT, *et al.* (2006) Ethanol can contribute to energy and environmental goals. *Science* 311: 506–508.
Finnveden G, Hauschild MZ, Ekvall T, *et al.* (2009) Recent developments in life cycle assessment. *Journal of Environmental Management* 91: 1–21.
Girod B, de Haan P, and Scholz RW (2011) Consumption-as-usual instead of ceteris paribus assumption for demand. *The International Journal of Life Cycle Assessment* 16: 3–11.
Gloria TP, Lippiatt BC, and Cooper J (2007) Life cycle impact assessment weights to support environmentally preferable purchasing in the United States. *Environmental Science & Technology* 41: 7551–7557.
Graedel T (1996) On the concept of industrial ecology. *Annual Review of Energy and the Environment* 21: 69–98.
Hammerschlag R (2006) Ethanol's energy return on investment: A survey of the literature 1990–present. *Environmental Science & Technology* 40: 1744–1750.
Hill J, Polasky S, Nelson E, *et al.* (2009) Climate change and health costs of air emissions from biofuels and gasoline. *Proceedings of the National Academy of Sciences of the United States of America* 106: 2077–2082.
Hill JD, Nelson E, Tilman D, Polasky S, and Tiffany D (2006) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11206–11210.
Howarth RW, Boyer EW, Pabich WJ, and Galloway JN (2002) Nitrogen use in the United States from 1961–2000 and potential future trends. *Ambio* 31: 88–96.
Hunt RG and Franklin WE (1996) LCA – How it came about – personal reflections on the origin and the development of LCA in the USA. *International Journal of Life Cycle Assessment* 1: 4–7.
Huo H, Wang M, Bloyd C, and Putsche V (2009) Life-cycle assessment of energy use and greenhouse gas emissions of soybean-derived biodiesel and renewable fuels. *Environmental Science & Technology* 43: 750–756.
Huo H, Wu Y, and Wang M (2009) Total versus urban: Well-to-wheels assessment of criteria pollutant emissions from various vehicle/fuel systems. *Atmospheric Environment* 43: 1796–1804.
ISO, International Standard Organization (2006a) *ISO 14040: Environmental management – Life cycle assessment – Principles and framework*. Switzerland: ISO.
ISO, International Standard Organization (2006b) *ISO 14044: Environmental management – Life cycle assessment – Requirements and guidelines*. Switzerland: ISO.
Johnson JMF, Franzluebbers AJ, Lachnicht Weyers S, and Reicosky DC (2007) Agricultural opportunities to mitigate greenhouse gas emissions. *Environmental Pollution* 150: 107–124.
Karlen DL, Varvel GE, Johnson JMF, *et al.* (2011) Monitoring soil quality to assess the sustainability of harvesting corn stover. *Agronomy Journal* 103: 288–295.
Lave LB, Cobas-Flores E, Hendrickson CT, and McMichael FC (1995) Using input–output analysis to estimate economy-wide discharges. *Environmental Science & Technology* 29: 420–426.
Lemoine DM, Plevin RJ, Cohn AS, *et al.* (2010) The climate impacts of bioenergy systems depend on market and regulatory policy contexts. *Environmental Science & Technology* 44: 7347–7350.
Leontief W (1970) Environmental repercussions and the economic structure: An input–output approach. *The Review of Economics and Statistics* 52: 262–271.
Liska AJ and Cassman KG (2008) Towards standardization of life-cycle metrics for biofuels: Greenhouse gas emissions mitigation and net energy yield. *Journal of Biobased Materials and Bioenergy* 2: 187–203.
Mullins KA, Griffin WM, and Matthews HS (2011) Policy implications of uncertainty in modeled life-cycle greenhouse gas emissions of biofuels. *Environmental Science & Technology* 45: 132–138.
Ohlrogge J, Allen D, Berguson B, *et al.* (2009) Driving on biomass. *Science* 324: 1–2.
Peters GP (2010) Carbon footprints and embodied carbon at multiple scales. *Current Opinion in Environmental Sustainability* 2: 245–250.
Roberts KG, Gloy BA, Joseph S, Scott NR, and Lehmann J (2010) Life cycle assessment of biochar systems: Estimating the energetic, economic, and climate change potential. *Environmental Science & Technology* 44: 827–833.
Savage N (2011) The ideal biofuel. *Nature* 474: 9–11.
Scown CD, Horvath A, and McKone TE (2011) Water footprint of U.S. transportation fuels. *Environmental Science & Technology* 45: 2541–2553.
Spatari S and MacLean HL (2010) Characterizing model uncertainties in the life cycle of lignocellulose-based ethanol fuels. *Environmental Science & Technology* 44: 8773–8780.
Sperling D and Yeh S (2009) Low carbon fuel standards. *Issues in Science and Technology* 25: 57–66.
Stratton RW, Min Wong H, and Hileman JI (2011) Quantifying variability in life cycle greenhouse gas inventories of alternative middle distillate transportation fuels. *Environmental Science & Technology* 45: 4637–4644.
Suh S and Huppes G (2005) Methods for life cycle inventory of a product. *Journal of Cleaner Production* 13: 687–697.
Tilman D, Socolow R, Foley JA, *et al.* (2009) Beneficial biofuels – the food, energy, and environment trilema. *Science* 325: 4–5.
Tsao CC, Campbell JE, Mena-Carrasco M, Spak SN, Carmichael GR, and Chen Y (2011) Increased estimates of air-pollution emissions from Brazilian sugar-cane ethanol. *Nature Climate Change* 2: 53–57.
Venkatesh A, Jaramillo P, Griffin WM, and Matthews HS (2011) Uncertainty analysis of life cycle greenhouse gas emissions from petroleum-based fuels and impacts on low carbon fuel policies. *Environmental Science & Technology* 45: 125–131.
Wang M, Wu M, and Huo H (2007) Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. *Environmental Research Letters* 2: 1–13.
Wang MQ, Han J, Haq Z, *et al.* (2011) Energy and greenhouse gas emission effects of corn and cellulosic ethanol with technology improvements and land use changes. *Biomass and Bioenergy* 35: 1885–1896.
Weidema B (2003) *Market Information in Life Cycle Assessment*. Denmark: Danish Environmental Protection Agency. (Environmental Project No. 863 2003).
Wiens J, Fargione J, and Hill J (2011) Biofuels and biodiversity. *Ecological Applications* 21: 1085–1095.
Wu M, Mintz M, Wang M, and Arora S (2009) Water consumption in the production of ethanol and petroleum gasoline. *Environmental Management* 44: 981–997.
Wu M, Wu Y, and Wang M (2006) Energy and emission benefits of alternative transportation liquid fuels derived from switchgrass: A fuel life cycle assessment. *Biotechnology Progress* 22: 1012–1024.
Yeh S, Berndes G, Mishra GS, *et al.* (2011) Evaluation of water use for bioenergy at different scales. *Biofuels, Bioproducts and Biorefining* 5: 361–374.
Yeh S, Jordaan SM, Brandt AR, *et al.* (2010) Land use greenhouse gas emissions from conventional oil production and oil sands. *Environmental Science & Technology* 44: 8766–8772.

Life History, Evolution of

Derek Roff, McGill University, Montreal, QC, Canada

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 715–728, © 2001, Elsevier Inc.

Glossary

Genotype The genetic constitution of an organism.
Heterozygosity The presence of alternate alleles at a given locus in a diploid organism (e.g., A_1A_2).
Homozygosity In a diploid organism, the presence of the same alleles at a given locus (e.g., A_1A_1).
Iteroparity Repeat breeding (see *semelparity*).
Semelparity Breeding once and dying; sometimes called "big bang" reproduction.

Introduction

Plants and animals show profound variation in all aspects of their life histories, which include age at maturity, age-specific fecundity, survival rate, size at birth, and so on. This variation is evident at both the inter- and intraspecific levels. For example, at the interspecific level, species of flatfish range in size from 2 cm-long tropical species that reproduce within their first year of life to behemoths such as the Pacific Halibut (*Hippoglossus hippoglossus*), which exceed 200 cm and take over 10 years to mature. Though the range in variation within a species is not as dramatic as between species, it is still impressive, as illustrated by variation in the flatfish, *Hippoglossoides hippoglossoides*. In this species maturation occurs at age 3 years at a length of 20 cm in populations off the coast of Scotland while the same species requires 15 years to reach maturity at a length of 40 cm on the Grand Banks of Newfoundland. Longevity and maximum size are equally different in the two populations, Scottish fish reaching a maximum length of 25 cm and an age of 6 years, compared to 60 cm and 20 + years on the Grand Banks. Similar observations on variation in life history characteristics could be made in most taxa. But though the diversity of life histories is readily apparent, attempts to understand its origin and maintenance are still in their infancy.

Because fitness varies quantitatively among different life histories, a necessary tool for analysis is mathematical modeling. An influential factor encouraging the use of mathematical investigation into life history variation was Lamont Cole's 1954 paper, "The Population Consequences of Life History Phenomena," which set out the basic mathematical framework by which the consequences of variation in life history traits can by analyzed. Cole's paper ushered in an era of research predicated on the integration of mathematics and biology in the study of the evolution of life history patterns.

In his review, Cole analyzed how changes in demographic attributes, such as the age at first reproduction, influenced the rate of increase of a population. Cole's paper gained widespread notice because of an apparent paradox with respect to the value of semelparity versus iteroparity:

> For an annual species, the absolute gain in intrinsic population growth which could be achieved by changing to the perennial reproductive habit would be exactly equivalent to adding one individual to the average litter size (Cole, 1954, p118, Cole's italics).

Thus, an annual species with a clutch size of 101 would increase in numbers as fast as a perennial that produces 100 young every year forever. There is obviously something amiss with this result, for perennials are common. While there is good evidence that survival and reproduction are negatively correlated it seems highly unlikely that perennial species are committing so much energy into reproduction that they cannot produce one more offspring. Cole's paradox derives from a failure to consider the consequences of juvenile and adult survival rates; specifically Cole implicitly assumed that the survival rate of juveniles is the same as that of the adults. In general, juvenile survival is lower (often by a great amount) than that of the adults and this means that the reproductive output of the annual must be greater than one for the annual to have the same gain in rate of increase as a perennial, the precise amount depending on the mathematical details of the model. Cole's paradox illustrates the need to formulate an analysis of life history variation within a mathematical framework where the biological assumptions are explicitly stated. Since Lamont Cole, the theoretical and experimental analysis of life history evolution has made enormous strides using two different analytical perspectives.

Two Frameworks for Analysis

The evolution of life history variation has been approached via two different modes of analysis, here termed phenotypic models and genetic models.

Phenotypic Models

In this approach no attempt is made to model the genetic basis of traits: it is simply assumed that there exists sufficient genetic variation that evolution is not constrained by genetic architecture. Is this a reasonable assumption? Most life history traits—such as age at first reproduction, fecundity, and survival—are not determined by simple Mendelian mechanisms such as single locus, two allele systems. More generally they

show continuous variation, or if dichotomous can be best interpreted under a threshold model in which there is an underlying continuously varying trait. Nevertheless, polygenic determination is not essentially different from the simpler case of a single locus with two alleles. In the one locus case the phenotype is determined by the additive action of the alleles, or one allele may show dominance. When one allele is dominant, the effect on the phenotype can be statistically decomposed into additive and dominance components. With two or more loci controlling a trait there may be interaction between loci. These interactions, termed epistasis, can produce a wide range of responses but are generally assumed to be absent (because most variation can be subsumed under the additive or dominance portion of genetic variance, this assumption is not as unreasonable as it might appear).

The average similarity between parent and offspring can be measured by the linear regression of mean offspring value (Y) on midparent value (X),

$$Y = (1 - h^2)\mathrm{m} + h^2 X \qquad (1)$$

where m is the population mean of the parents and the slope of the regression, h^2, is termed heritability in the narrow sense. More generally, heritability is defined as follows:

$$h^2 = \frac{V_A}{V_P} = \frac{\text{Additive genetic variance}}{\text{Total phenotypic variance}} \qquad (2)$$

Heritability in the narrow sense, or generally simply heritability, must not be confused with heritability in the broad sense, V_G/V_P, which is a measure of the overall variance in the trait attributable to all genetic influences. Obviously h^2 varies between 0 and 1: at 0 there is no resemblance between parent and offspring due to the additive effects of genes, while at 1 the mean offspring value equals the midparent value. The importance of heritability for evolutionary theory is obvious: the higher the value of h^2, the faster genetic changes in the population can occur. Most important from the point of view of life history theory, if heritability is zero for a particular trait or traits under consideration, then the optimal combination cannot be attained because the effects of selection on parents are not manifested in the offspring. Heritability estimates for different types of traits are shown in **Table 1**. While heritabilities for life history traits are typically lower than those less directly related (in general) to fitness, there is nevertheless sufficient genetic variation to permit rapid evolutionary change even under modest selection pressure.

While the assumption of sufficient additive genetic variation is reasonable with respect to the equilibrium conditions, the genetic architecture can considerably influence evolutionary trajectories and hence the phenotypic approach is contingent on the state having attained an equilibrium. The method seeks to construct the fitness surface and hence the optimal combination of trait values.

Table 1 A summary of heritability estimates for nondomestic animals

Comparison	*Life history*	*Behavior*	*Physiology*	*Morphology*
Drosophila	0.12	0.18	ns	0.32
All animals	0.26	0.30	0.33	0.46
Medians	0.26	0.37	0.31	0.51

Optimality Modeling

The concept of tradeoffs is central to present theories of the evolution of life history traits, for tradeoffs limit the scope of variation. Within the set of possible combinations there will be at least one combination that exceeds all others in fitness. Optimality analysis assumes that natural selection will drive the organism to that particular set. To initiate an analysis using the principle of optimality, we must designate some parameter to be optimized. In the present case we assume that there is some measure of fitness that is maximized by natural selection. The second step is to construct a set of rules that define the life history pattern of the organism, hypothetical or real, under study. Within these rules there will exist a variety of possible life histories; the optimal life history is that which maximizes fitness.

What do we do if the predicted life history does not correspond to that which is observed? The first point to note is that the principle of optimality is not under test. Failure to get a correct prediction is not taken as evidence that fitness is not being maximized, but it is taken to imply that the model is deficient. Having found that the initial model does not work we inquire into the assumptions of the model, namely the rules that define the range and scope of life history variation. These rules are changed either arbitrarily or based on further observation until agreement is gained between prediction and observation. Having found congruence, we are not able to say that therefore the component relationships are correct. The underlying components must be independently verified; a model is only the logical outcome of a set of interactions and the onus is on the person specifying those rules to demonstrate that they are indeed valid. Note again, that the assumption that fitness is being maximized is not under test, except to the extent that the particular measure chosen may be inappropriate. This does not mean that we assume that all traits and trait combinations are the result of adaptive evolution. But we do choose those that we have a priori reason to suppose are under selection.

A primary purpose of optimality modeling is to organize a program of experimentation and data collection. An adequate fit of a model to data gives us reassurance that a sufficient number of factors have been taken into account. Nevertheless, the validity of a model is continually under question and is challenged by the addition of more information. The more tests the model survives the greater the assurance that it is realistically capturing the important elements that determine the set of traits being studied.

Game Theory

Game theory is really a subset of optimality modeling: it is appropriate when interactions are frequency dependent. The approach comprises two essential elements. First, it is assumed that particular patterns of behavior will persist in a population provided no mutant adopting an alternate behavior can invade. Such stable combinations are termed evolutionarily stable strategies (ESS). The concept of the ESS is not unique to game theory: the maximization of fitness measures in

optimality models are all ESSs within the context in which they are appropriate. Second, for each type there must be an assigned gain or loss in fitness when this type interacts with another individual. From this payoff matrix we compute the expected payoff for each behavior. For two behaviors to be evolutionarily stable, their fitnesses must be equal.

Genetic Models

Genetic modeling proceeds by first defining the genetic mechanism determining the phenotypic trait and then cranks the model through the appropriate mathematical machinery to obtain the equilibrium trait values. The critical problems are the correct definition of the genetic architecture and the estimation of the selection gradients.

Simple Mendelian Models

Consider a trait whose expression is governed by a single locus with two alleles, A_1, A_2. A selection coefficient (= fitness) representing their relative contribution to the next generation can be assigned for each of the three genotypes (**Table 2**). The change in the frequency of A_1, p, is then easily obtained as

$$\Delta p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{\overline{w}} \qquad (3)$$

At equilibrium $\Delta p = 0$, and hence it is a trivial matter to predict the frequencies of the three genotypes at equilibrium and hence the mean trait value. Complexity can be added by making the fitnesses a function of density or frequency, but this does not change the basic mathematical approach.

Most traits of ecological interest, fecundity, age at maturity, clutch size, egg size, and so on are continuous in character. Even traits that appear dichotomous, such as liability to disease, wing dimorphism, diapause, and sex ratio, are best understood as being the result of some underlying continuously varying factor exceeding or not attaining a threshold for the expression of the trait. The expression of traits that show continuous variation are not, in general, the result of a single gene, nor two genes, but a large number of genes that acting additively produce a continuous spectrum of phenotypes. The analysis of such traits is the domain of quantitative genetics. This is largely a statistical approach to genetic variation and is founded on a mathematical analysis of variation rather than an understanding of how groups of genes interact to determine a particular trait.

Quantitative Genetic Models

The concept of heritability has already been introduced. For nonzero heritabilities, the rate at which the appropriate combination of trait values can be realized depends in part on the value of the heritability and the intensity of selection, measured as the selection differential, S, the difference between the mean of the population and the mean of the selected parents. For a single trait it can been shown from Equation 1 that the response to selection (the difference between the mean population values in parental and offspring generations), R, is equal to h^2S. Since selection changes the gene frequencies, heritability must change at each generation of selection. This problem is avoided in the application of quantitative genetic theory to natural populations by assuming that selection is weak and population size large, thereby permitting mutation to replace variation eroded by selection. Whether this assumption is reasonable can only by answered empirically and at present there are insufficient data to draw any meaningful conclusion.

Traits are typically not inherited as separate independent units; rather, phenotypic values of several traits are typically controlled in part by a common set of genes. Thus selection on one trait will produce a change not only in the selected trait but also traits that are genetically correlated by virtue of shared genes (genetic correlations can also arise through linkage disequilibrium, but these are transitory and are ignored here). Selection on two traits, X and Y, will produce standardized (in phenotypic standard deviation units) responses R_X and R_Y,

$$R_X = \beta_X h_X^2 + \beta_Y h_X h_Y r_A$$
$$R_Y = \beta_Y h_Y^2 + \beta_X h_X h_Y r_A \qquad (4)$$

where r_A is the genetic correlation between the two traits and β_X, β_Y are the standardized selection differentials on X and Y, respectively. The preceding equation can be more conveniently written in matrix notation, for which the typical notation is $\Delta\bar{z} = G\beta$, where $\bar{z}$ is a vector of trait means, G is the genetic variance-covariance matrix, and β is a vector of (unstandardized) selection gradients for each character.

The evolutionary trajectories for two traits can be easily visualized by constructing a plot of isoclines of equal fitness (defined by the selection gradients) and plotting the changes in trait means each generation as illustrated in **Figure 1**. If there is no genetic correlation between the two traits, the population will move rapidly to its optimal combination. For genetic correlations lying between -1 and $+1$, the trajectory is ''warped'' and if the genetic correlation is exactly ± 1, the optimal combination may be unattainable (**Figure 1**). The reason for this can be understood by considering the regression of the additive genetic values of trait Y on X. According to the quantitative genetic model, we have $Y = c + bX +$ error, where c and b are constants and the error term is normally distributed with mean zero. Provided the variance in the error term is not zero, any combination of Y and X is possible; hence, evolution can always move the traits to such a combination. However, if the variance of the error term is zero, which occurs when the (genetic) correlation between Y and X is ± 1, then the traits are constrained to lie on the regression line. The number of cases in which r_A is exactly equal to unity is very small (*see* The Evolution of Life Histories). With more than two traits evolution can be constrained if one or several eigenvectors of the genetic variance-covariance matrix are zero (singular matrix).

Table 2 The derivation of the optimal trait value in a simple single-locus model

	A_1A_1	A_1A_2	A_2A_2	*Sum*
Ratios before selection	p^2	$2pq$	q^2	1
Fitness	w_{11}	w_{12}	w_{22}	
Ratios after selection	$p^2\ w_{11}$	$2pqw_{12}$	q^2w_{22}	$\bar{w}$

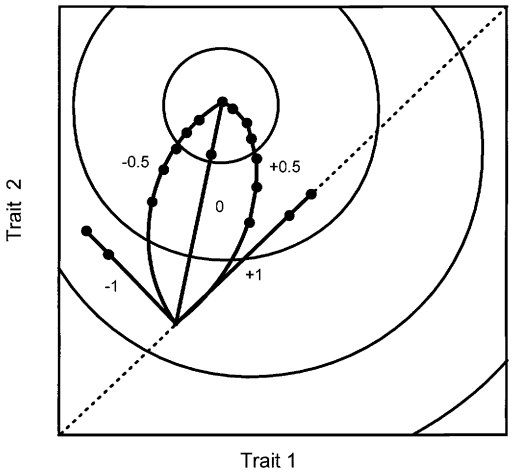


Figure 1 Hypothetical evolutionary trajectories for two traits. The solid lines define contours of equal fitness. Each dot represents the trait combination achieved after some specified number of generations. The values beside each trajectory indicate the genetic correlation between the two traits.

It is possible for there to be no response to selection in spite of both genetic variance for the trait and a phenotypic correlation with another trait under selection. The second observation is particularly important since it implies that a trait may be under selection but show no response. This can be most easily seen by relating the selection differential as a function of the selection gradient and phenotypic variance (V_{PX}) and covariance ($\mathrm{Cov_P}$)

$$S_X = \beta_X V_{PX} + \beta_Y Cov_P \qquad (5)$$

with a similar term for S_Y. Now suppose that (a) there is no additive genetic variance for trait Y, (b) trait Y is phenotypically correlated with X but genetically uncorrelated ($Cov_A = 0$), and (c) there is no direct selection on trait X (i.e., $\beta_X = 0$) but there is selection on trait Y (i.e., $\beta_Y > 0$). From Equation 5 it can be seen that a positive selection differential is generated ($S_X > 0$), but from equation 4 ($R_X = V_{AX}\beta_X + Cov_A/\beta_Y$) it is apparent that there will be no response to this selection, even if there is additive genetic variance for X. A likely case in which this situation can arise is when trait X is a heritable trait such as fecundity and Y is a trait such as nutritional status that might have zero heritability: thus well-nourished individuals have high fecundity, which gives them an apparent selective advantage, but this is not realized because selection is acting only on the environmentally determined component of fecundity. The mechanism described earlier, or ones conceptually similar, have been proposed for the observation of directional selection without evolutionary response in breeding date in birds, clutch size in birds, and tarsus length in birds.

The theory described here assumes that the genetic covariance (i.e., the genetic correlation) remains constant. As with heritability, selection changes allelic frequencies and hence must in principle change the genetic correlations, or more generally the genetic variance-covariance matrix. The G matrix will remain constant provided selection is weak and population size large so that mutation can replace variance lost due to selection and drift.

Tradeoffs: A Central Feature of Life History Analysis

Tradeoffs are essential elements of all the above approaches. In the "phenotypic approach" it is implicitly assumed that the tradeoffs are genetically determined, otherwise there could be no evolutionary response. In quantitative genetic terms, this is interpreted as a negative genetic correlation between two traits. Thus it has been supposed that to demonstrate an evolutionarily important tradeoff it is sufficient to demonstrate that the tradeoff is expressed not only as a phenotypic correlation but also as a negative genetic correlation. However, as discussed earlier, a negative genetic correlation is not, in principle, a barrier to movement anywhere in parameter space. Thus while it is necessary for a tradeoff to be expressed as a genetic correlation for it to be evolutionarily significant, this is not a sufficient demonstration that the optimal combination will be governed by the tradeoff.

Empirical investigations of tradeoffs can be placed into three categories: (a) phenotypic relationships based on field or laboratory observations of unmanipulated situations, (b) experiments in which organisms were manipulated to vary the value of one trait (e.g., manipulation of clutch size to investigate subsequent survival of adults or young), and (c) demonstration of a negative genetic correlation between two traits, obtained by sib analysis or selection. Only the last measure provides definitive proof that the tradeoff is evolutionarily significant but there is, however, merit in the second approach because it can establish the mechanism generating the tradeoff. The first approach is suspect because of the covariation of traits that could produce false conclusions if such covariation is not taken into account. For example, reproduction may be determined by condition, those in poor condition not breeding; consequently it would not be surprising to find the survival rate of nonbreeders to be less than breeders.

There is no reason to suppose that all tradeoffs between life history traits will be negative, but we should expect a significant proportion to be so. In fact, the distribution of genetic correlations between life history traits is very broad (**Figure 2**), with 39% being negative but few estimated to be -1. In contrast, only 23% of genetic correlations between morphological traits are negative. Thus there do appear to be tradeoffs that will at least impede the evolution of life history traits.

The Evolution of Life Histories

The starting point of many analyses of evolutionary change, most particularly that of optimality modeling, is the assumption that there exists some variable maximized by selection. The issue of what is being maximized has been the subject of much discussion, partly because the appropriate measure of fitness changes with circumstance and method of analysis. Broadly, measures of fitness can be separated into two groups: global measures and local measures. A global measure of fitness is one that involves the interaction of all life

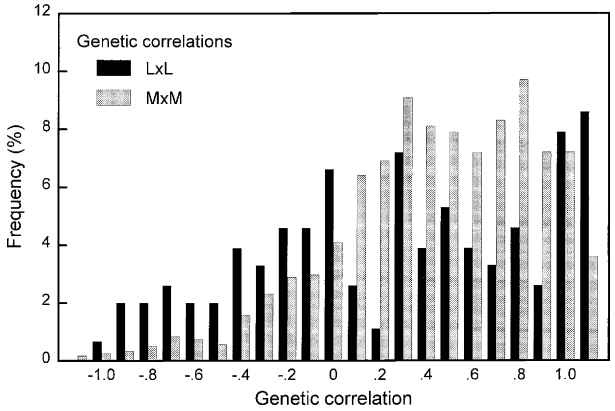


Figure 2 Distributions of the genetic correlation between Morphological traits (M × M) and Life History traits L × L.

history components, the best example being Fisher's Malthusian parameter *r*. Local measures assume that maximization of a fitness component will also maximize the overall fitness of the organism: for example, a common local measure used in foraging theory is the net rate of energy intake. This is an appropriate measure if it can be shown that maximizing this rate does not detrimentally affect other components of fitness, in which case maximizing net rate of energy intake will also increase global fitness.

Local measures are generally tailored for the particular analysis under consideration, but there now exists a general consensus, and more important sound theoretical rationale, of what global measures are likely maximized by natural selection.

Static Environments

A population growing in an unlimited, homogeneous, and constant environment follows the simple exponential growth function

$$\frac{dN(t)}{dt} = rN(t)$$

$$N(t) = N(0)e^{rt} \tag{6}$$

where $N(t)$ is population size at time t and r is the intrinsic rate of increase, comprising the difference between instantaneous rates of birth and death. Equation 6 can also be written as

$$N(t) = N(0)(e^r)^t = N(0)\lambda^t \tag{7}$$

The symbol λ is called the finite rate of increase and is sometimes used instead of r.

Suppose there are two clones with population sizes, $N_1(t)$ and $N_2(t)$, respectively, the first with an intrinsic rate of increase of r_1 and the second with r_2, with $r_1 > r_2$. The ratio of population sizes after some time **t**, given that both clones start with the same population size is,

$$\frac{N_1(t)}{N_2(t)} = e^{(r_1 - r_2)t} \tag{8}$$

It is clear that as time progresses this ratio will increase, clone 1 becoming numerically more and more dominant in the combined population. This conclusion does not depend on the two clones beginning with the same population size: differences in starting condition simply accelerate or retard the rate at which clone 1 increases in frequency relative to clone 2.

For these two clones, an appropriate measure of fitness is r, since the frequency of the clone with the highest value of r will increase toward unity. Thus any mutation in a set of clones that increases r by changing rates of birth or death will increase in frequency in the population. There are no difficulties in assigning r as a measure of fitness in the preceding circumstances. Difficulties arise, however, when sexual reproduction and age structure are introduced. Suppose we have a random mating population in which a mutation arises that increases birthrate or decreases death rate. Since the mutant will initially be rare in the population, its fate can be ascertained by considering the birthrates and death rates of the heterozygote alone. If the heterozygote's rate of increase is enhanced, the mutation will increase in frequency in the population, but its ultimate fate depends on the relative birthrates and death rates of the homozygotes and heterozygotes bearing the mutant allele. If the homozygote carrying both mutant alleles has a higher birthrate or a lower death rate than the heterozygote, the mutant allele will eventually be fixed in the population; otherwise the population will reach a stable polymorphism.

The general assumption, stemming from the work of Fisher, has been that r can be associated with genotypes that follow particular life histories and that selection will favor that genotype with the highest value of r. In an age-structured population, the rate of increase is obtained by solving the characteristic equation

$$\int_0^\infty e^{-rx} l(x)m(x)dx = 1 \tag{9}$$

where $l(x)$ is the probability of surviving to age x and $m(x)$ is the number of female births at age x. The discrete time equivalent of this is

$$\sum_{x=1}^{\infty} e^{-rx} l(x)m(x) = 1 \tag{10}$$

Note that in the discrete version the initial age is subscripted as 1 not 0. The important issue to be considered is the fate of a mutant that increases r. Charlesworth demonstrated that to a rough approximation the rate of progress of a rare gene eventually becomes directly proportional to its heterozygous effect on r. In this case, the probability of survival of a mutant gene of small effect in a near-stationary population is largely determined by its effect on r. Following a more detailed analysis, Charlesworth (1994) concluded

> that for the case of weak selection and random mating with respect to age, the intrinsic rate of increase of a genotype or, more generally, the mean of the male and female intrinsic rates, provides an adequate measure of fitness in a density-independent and constant environment.

Lande tackled the problem of applying quantitative genetic theory to the evolution of r in a population. Assuming weak selection, large population size, and a constant genetic

variance-covariance matrix, Lande showed that in a constant selection regime, life history evolution continually increases the intrinsic rate of increase of a population until an equilibrium is reached. The intuitive appeal of r as a suitable measure of fitness thus receives qualified support from both population genetic and quantitative genetic theory.

If the population is stationary, r is zero and the characteristic equation reduces to

$$R_0 = \int_0^\infty l(x)m(x)dx = 1 \tag{11}$$

R_0 is termed the net reproductive rate and is the expected number of female offspring produced by a female over her life span. The use of R_0 makes analysis easier and can be justified if r is very close to zero or if the density-dependent or other factors that maintain the population at some relatively stable value do not impinge on the traits under consideration. For example, population size might be controlled by density-dependent mortality in the larval stage, while the object of study is the evolution of female age at maturity. In this case we can examine the relationship between the age at maturity and fitness under the working assumption that genotypes do not differ in the characteristics of their larvae. Since population size is stable, the expected lifetime fecundity, R_0, is then the appropriate measure of fitness.

In some analyses, fitness is determined from the reproductive value at age x, $V(x)$,

$$V(x) = \frac{e^{rx}}{l(x)} \int_x^\infty e^{-ry} l(y)m(y)dy \tag{12}$$

The reproductive value of an individual of age x is a measure of the extent to which it contributes to the ancestry of future generations. Williams (1966) postulated that natural selection maximizes r by maximizing reproductive value at every age. A mathematically correct statement of the principle is "reproductive value at each age is maximized relative to reproductive effort at that age, although not necessarily with respect to effort at other ages."

The preceding approach has been very successful in accounting for variation in life history traits. An illustration of this is the analysis of the optimal size at first reproduction in *Drosophila melanogaster*. In *D. melanogaster* body size increases with development time and fecundity increases allometrically with adult body size, which is fixed on eclosion from the pupa. The age schedule of female births is triangular in shape (**Figure 3**) and is described by the equation

$$m(x) = \frac{1}{2} c_1 L^{c_2} (1 - e^{-c_3(x-c_4)}) e^{-c_5 x} \tag{13}$$

where the c_is are constants, L is thorax length, and x is age. Thorax length, L, scales the age schedule of reproduction, larger females producing more eggs, but does not change its position. The constant c_1 is the product of two constants: the coefficient of proportionality within the allometric relationship between length and fecundity, and the proportion of eggs that fail to hatch.

Development time, $d(L)$, scales to body size according to the relationship (**Figure 3**)

$$d(L) = c_6 L^{c_7} + c_8 \tag{14}$$

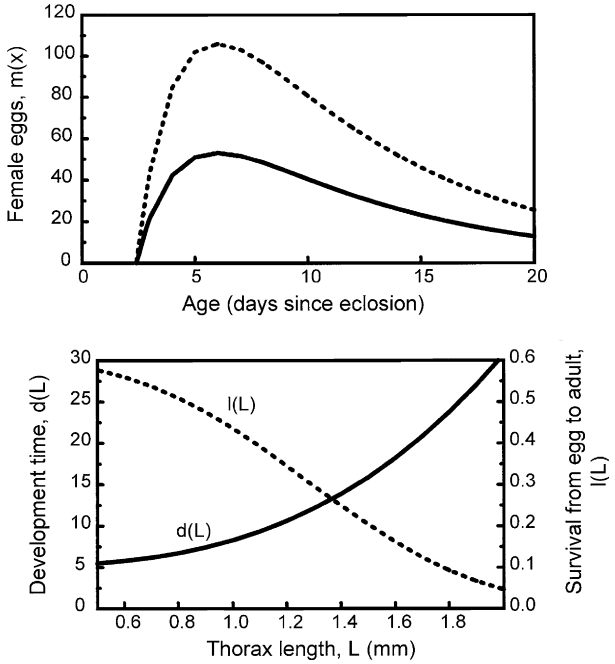


Figure 3 Relationship between body size and life history traits in *Drosophila melanogaster*. Upper panel: Age schedule of reproduction for small (solid line) and large (dashed line) females. Lower panel: Development time and survival from egg to adult in relation to adult thorax length.

Thus development time scales allometrically with size except for the constant c_8, which represents time required for the eggs to hatch and the development within the pupa, both of these components being independent of size.

Because information on mortality rates are so poorly known, it was assumed that instantaneous rates remain constant in the adult and larval phases at M_a and M_l, respectively. Probability of surviving the larval period, $l(L)$, is thus (**Figure 3**)

$$l(L) = e^{-M_l d(L)} \tag{15}$$

Drosophila melanogaster is a colonizing species and hence the appropriate measure of fitness is r. Combining the previous relationships, we obtain the characteristic equation

$$\sum_{x=1}^{\infty} \frac{1}{2} e^{-r[x+d(L)+c_4]-M_l d(L)-M_a(x+c_4)-c_5 c_4} c_1 L^{c_2} (1 - e^{-c_3 x}) e^{-c_5 x} = 1 \tag{16}$$

where, for convenience age has been rescaled to begin at the first day of egg laying. Though this equation is tediously long, its solution presents no great difficulty. Briefly, the method is, first, to evaluate the series making use of the fact that it is a geometric progression and, second, to differentiate implicitly to obtain the optimal length at maturity. The optimum thorax length depends on adult mortality, larval mortality, and the constant initial "mortality" component (p or c_1). **Figure 4** shows the relationship between r and thorax length using values obtained from laboratory stocks and estimates from wild populations. The predicted maximum, 0.95 mm, falls

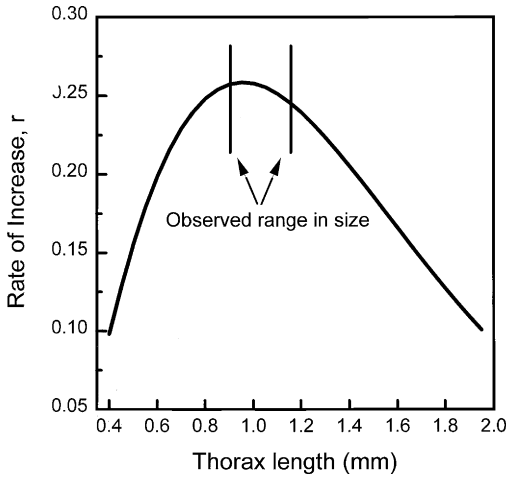


Figure 4 Estimated relationship between fitness, as measured by the rate of increase, and thorax length in *Drosophila melanogaster.*

very nicely within the observed range in thorax length of 0.90 to 1.15 mm.

Stochastic Environments

Environments are typically variable in both time and space. While the assumption of a constant environment may be a reasonable first approximation in many cases, there will be many others in which variation cannot be ignored. The simplest case to consider is where the environment is stochastic and there is no cue as to future conditions.

In a stochastically, temporally fluctuating environment, the correct measure of fitness is the geometric mean of the finite rate of increase. The rationale for this measure is as follows: the size of a population after t time intervals is given by

$$N(t) = N(0)\lambda_1\lambda_{2\ldots}\lambda_t = N(0)\prod_{i=1}^{i=l}\lambda_i \qquad (17)$$

The arithmetic and geometric means are

$$\text{Arithmetic mean} = \frac{1}{t}\sum_{i=1}^{i=t}\lambda_i$$

$$\text{Geometric mean} = \sqrt[t]{\prod_{i=1}^{i=t}\lambda_i} \qquad (18)$$

It is obvious that the geometric mean describes the long-term average per generation change in population size and hence is the parameter that reflects the fitness of different strategies. If r is used in place of the finite rate of increase, the appropriate measure of fitness is the arithmetic mean of r. This can be demonstrated as follows: dividing time intervals into small units we can write

$$N(t) = N(0)e^{r_1}e^{r_2}\ldots e^{r_t} = N(0)e^{\sum_{i=1}^{t} r_t} \qquad (19)$$

from which it is readily apparent that the correct measure of fitness is the arithmetic average of r.

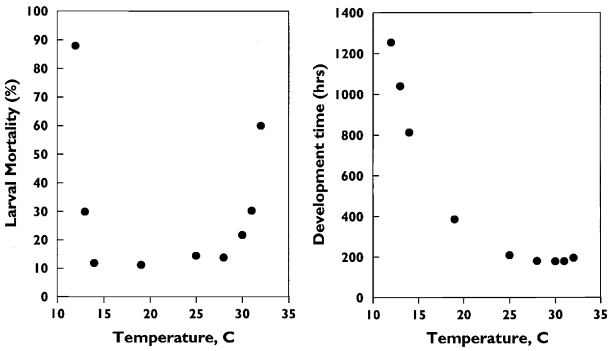


Figure 5 Reaction norms between two life history traits (larval mortality, development time) and temperature in *Drosophila melanogaster*.

Consider two genotypes living in an environment that comprises two types of year, "good" and "bad," each occurring with equal frequency. In "good" years, genotype A has a finite rate of increase of 2 and in a "bad" year a finite rate of increase of 0.5, while genotype B has finite rates of increase of 1 and 1.1, respectively. The arithmetic averages of A and B are 1.25 and 1.05, respectively, but the geometric averages are 1 and 1.1. Thus genotype B has the highest long-term fitness although it has a smaller arithmetic finite rate of increase. Genotype A increases more than genotype B in "good" years but suffers a greater reduction in "bad" years. The relatively high fitness of genotype B resides in the fact that although it has a smaller arithmetic average, it also has a smaller variance in its finite rate of increase. In a stochastic environment the highest fitness may be gained not by the production of a single phenotype but by a genotype producing a range of phenotypes. Such bet-hedging strategies are common, an example being the production of both diapausing and direct-developing offspring in both invertebrates and plants.

Predictable Environments

Most environments show predictable changes either as a result of biotic factors (e.g., overgrazing, community succession) or abiotic factors (e.g., winter, monsoons). Suppose factors that affect one or more life history characters are fully predictable—that is, given some cue E the life history characters will be changed from their present values to values that can be related to E by some function. For each case there is some optimal combination of trait values: given the predictable nature of the cue means that the organism's life history can be altered to accommodate the new conditions. Suppose, for example, that when the environmental conditions are E_i the optimal clutch size is C_i. Upon receipt of the environmental cue, the organism produces the optimal clutch size, C_i. This relationship between trait value and environment is called a norm of reaction, whether or not it is optimal. Reaction norms are ubiquitous. Two examples of reaction norms are shown in Figure 5. A reaction norm may be adaptive or a nonadaptive (even maladaptive) physiological response to the environment. To demonstrate that a reaction norm is adaptive requires that we use the set of tradeoff functions related to the environment to predict the optimal

reaction norm. Consider, for example, the question asked by Stearns and Koella (1986, p. 894): "How should an organism encountering an unavoidable stress that results in slower growth alter its age at maturity to keep fitness as high as possible despite the constraints imposed by slower growth?"

To address this question, Stearns and Koella used the following hypothetical model

1. Fecundity increases with body size, which increases with age:

$$W(x) = c_1(1 - c_2e^{-kx})$$

$$m(x) = c_3W(x) + c_4 \qquad (20)$$

where $W(x)$ is size at age x, k is the parameter determining the rate at which the asymptotic size is approached, and c_1–c_4 are constants.

2. Adult mortality rate, M_a, decreases with growth rate according to the function:

$$M_a = \frac{C_5}{k^{c_6}} \qquad (21)$$

3. Mortality rate of juveniles, M_j, is a function of the component of adult mortality, c_5, plus an amount that decreases with age at maturity, α, and growth rate, k:

$$M_j = c_5 + \frac{c_7}{\alpha^{c_8}k^{c_9}} \qquad (22)$$

Stearns and Koella assumed that fitness is maximized by maximizing r. The predicted norms of reaction between age and size at maturity for four possible scenarios are illustrated in Figure 6:

1. Mortality rate and growth rate independent ($c_6 = 0$, $c_9 = 0$). Decreases in growth rate, k, favors an increase in the age at maturity and a decrease in the size at maturity.
2. Juvenile mortality rate increases as growth rate increases ($c_6 = 0$, $c_9 > 0$). As with the previous case, decreases in k favor an increase in the age at maturity and a decrease in the size at maturity. The actual norm of reaction may have the same shape as case 1 or be sigmoidal.
3. Adult mortality rate increases as growth rate decreases ($c_6 > 0$, $c_9 = 0$). The optimum norm of reaction has the same shape as case 1 but is reversed, size at maturity and age at maturity being positively related.
4. The adult and juvenile rates are variable ($c_6 > 0$, $c_9 > 0$). The relationship between age and size at maturity is "keel-shaped." The portion of the curve in which size at maturity is an increasing function of age at maturity only occurs when adult mortality rates exceed juvenile mortality rates by a factor of 100 or more. Thus, generally the trajectory is determined by mortality rates of juveniles.

The message from the preceding analysis is that reaction norms can be quite complex and dependent not only on the functional relationships determining the tradeoffs but also the parameter values. The situation is made even more complex if the environment is spatially variable: in this case the appropriate fitness measure is the overall r, not the r that is characteristic of a particular environment. This requires modifying the characteristic equation to

$$\int p(h) \int l(x,h)m(x,h)e^{-rx}dxdh = 1 \qquad (23)$$

where $p(h)$ is the probability of habitat h occurring. The optimal combination of trait values can be markedly different from the average set determined by considering each habitat separately.

For reaction norms to evolve to produce the optimal response, there must be genetic variation in how individuals respond to the environment (i.e., there must be genetic variation for phenotypic plasticity). The evolution of phenotypic plasticity can be addressed using two apparently different mathematical perspectives: the character state approach and the reaction norm approach. Both approaches are actually interchangeable and each has advantages and disadvantages.

The character state approach is based on the assumption that the same trait measured in two environments can be considered as two traits that are genetically correlated. This is illustrated in Figure 7, where each line represents a separate genotype. If the lines joining the phenotypic values in the two environments (E_1 and E_2) intersect at a common point between E_1 and E_2, the genetic correlation between the two traits is -1 (Figure 7, panel D). If the lines intersect outside the range E_1 to E_2, the genetic correlation is $+1$ (Figure 7, panel A; note that parallel lines also produce a genetic correlation of $+1$, since mathematically their point of intersection is at the

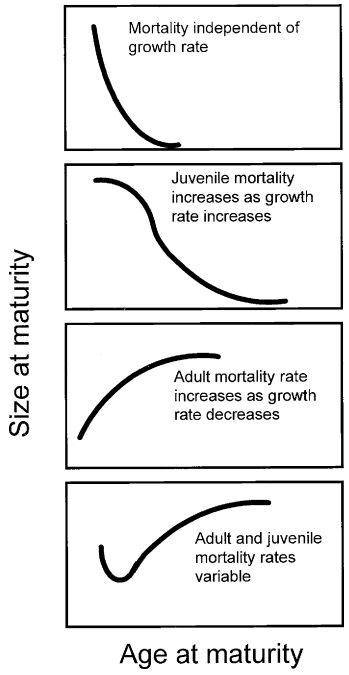


Figure 6 Predicted norms of reaction between size and age at maturity for the life history model described in the text.

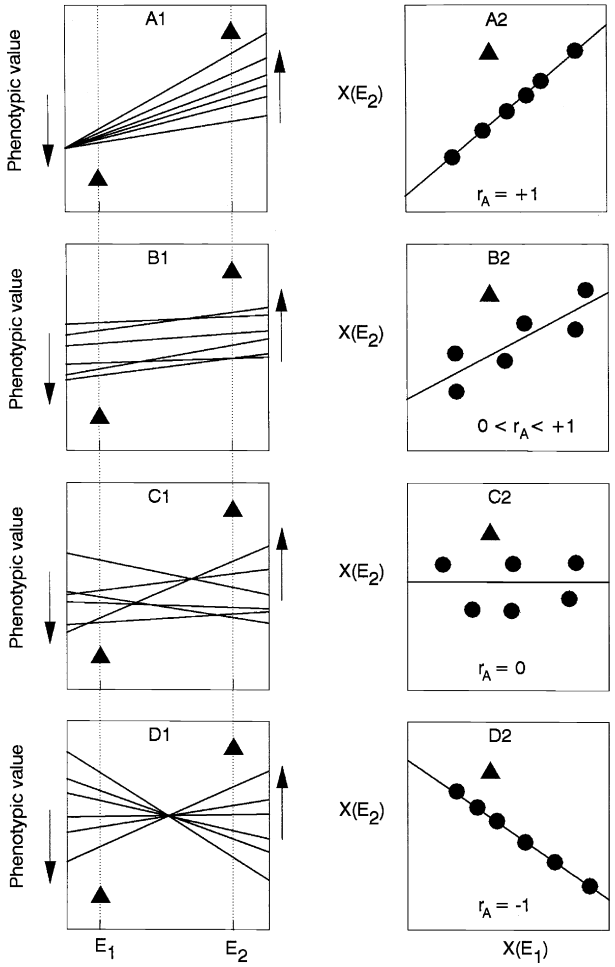


Figure 7 Phenotypic plasticity in two environments as viewed from the character state perspective. The panels on the left show the character states in the two environments E_1 and E_2, each line joining the trait values of a single genotype (the reaction norms of the genotype). The panels on the right show the regression of the trait value from the second environment, $X(E_2)$, on the trait value from the first environment, $X(E_1)$. A: The reaction norms meet at a single point beyond the range of the two environments. The correlation between trait values is $+1$. Note that parallel reaction norms also give a genetic correlation of $+1$ (mathematically this is because the lines meet at infinity). B and C: The reaction norms cross at several points within the range of the two environments. Depending on the distribution of intersections the genetic correlation will be positive but less than $+1$ (B), zero (C), or negative but greater than -1 (not shown). D: The norms of reaction intersect at a single point between the two environments. In this case the genetic correlation is -1. The triangles show a hypothetical optimal combination of trait values. Because in cases A and D all the points lie on a single line ($r = \pm 1$), this combination may not be achievable. In all other cases, because in principle the distribution about the regression line is normal (i.e., no value is excluded), selection can move the population to the joint optimum.

point of infinity point beyond E_1 or E_2). The genetic correlation will differ from ± 1 if there is no common point of intersection (Figure 8, panels B and C).

While the character state approach sees the phenotype as two points in state space, the norm of reaction approach sees a

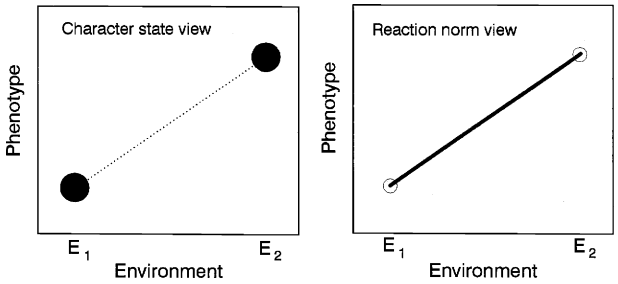


Figure 8 A pictorial representation of the two viewpoints of phenotypic plasticity. The character state approach (upper panel) focuses attention on the trait values in the two environments, while the reaction norm approach (bottom panel) focuses on the line joining the trait values.

line (Figure 8). Thus, for two environments phenotypic variation can be described as $X(E) = c_0 + c_1E + e$, where the two parameters c_0 and c_1 are viewed as traits, and e is an error term that is normally distributed with mean 0. Evolutionary change in X depends on the heritabilities of c_0 and c_1 and the genetic correlation between them. If the genetic correlation is ± 1 then the line that is genetically fixed and corresponds, as it should, to the cases A and D in Figure 7. This description is merely an alternate formulation of the character state approach. The conceptual advantage of this approach is that it extends quite naturally to continuously distributed environmental variables such as temperature. The preceding model can clearly apply to any linear reaction norm; if the relationship between trait and environment is more complex, a more complex function can readily be substituted, for example, the quadratic $X(E) = c_0 + c_1E + c_2E^2 + e$.

The evolution to the optimal reaction norm could be constrained. Suppose, for example, the reaction norm were linear but the optimal reaction norm were quadratic: in this case the optimal reaction norm cannot be attained. If some arbitrarily complex polynomial function is assumed, then many shapes are possible and no constraint will ensue. While there are abundant data indicating that genetic variation for phenotypic plasticity of life history traits is common, there are few data suggesting that the evolution of reaction norms are constrained.

The study by Weis and Gorman (1990) on the evolution of gall size in the insect *Eurosta solidaginis* is one of the few studies that has examined selection on a reaction norm in a natural population. *Eurosta solidaginis* attacks the goldenrod *Solidago altissima*, laying its eggs in the stems and the larvae causing the formation of a gall. From field data, Weis and Gorman determined that survival was a complex function of gall diameter, with an optimum at approximately 24 mm (Figure 9). On the basis of analysis of 16 full-sib families, they obtained a linear relationship between final gall diameter and the time lag between oviposition and gall initiation (Figure 9). Because the time lag showed variation associated with plant sibship but not among insect sibship, Weis and Gorman argued that time lag is a trait of the plants and not the insects. Thus time lag is, from the insect's perspective, an environmental gradient associated with the plant. The heritability of the gall-size/lag time intercept was determined as 0.21 (SE = 0.18) and the slope as 0.54 (SE = 0.25). Although

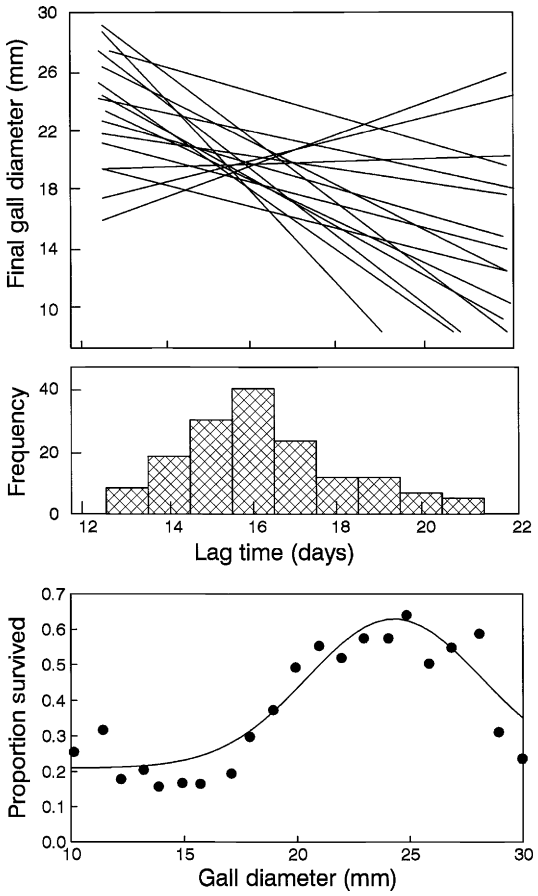


Figure 9 Bottom panel: Proportion of survivors as a function of gall size (x) in *Eurosta solidaginis*. Dots show observed value, the solid line the fitted curve

$$Survival = 0.21 + 0.42e^{-\left(\frac{x-24.3}{3.84}\right)^2}$$

Middle panel: Frequency distribution of lag time between oviposition and gall initiation. Top panel: Reaction norms from the 16 full-sib families.

the former estimate is not significantly different from zero, the ANOVA indicated significant family effects.

Maximum survival occurs when gall diameter is approximately 24.3 mm, and hence we would expect selection to favor reaction norms that produced a gall of this size, regardless of lag time. In this regard it is significant that the greatest intersection of the reaction norms occurs at the modal lag time but gives a gall diameter of approximately 19 mm (**Figure 9**). Thus the fitnesses of the different families are approximately equal in the most frequent environment but are lower than the maximum fitness. Selection should act to shift this area of intersection upward, which could be done by increasing the intercepts or changing the slopes. In the former case an overall increase of approximately 4 mm would achieve the required maximization of fitness within the modal environment, but in the latter case it appears that a much greater range in slopes must be achieved. Therefore, we might expect that selection will act most strongly in a directional sense on the intercept, since the same change in all families is likely to have a greater impact on fitness than a similar change in the slope. This prediction can be tested by estimating the relative strengths of selection on the two parameters. As predicted, directional selection was stronger on the slope than the intercept (Intercept/Slope = 4.4), but stabilizing selection acted approximately equally on the two components (Intercept/Slope = 1.4).

This study reinforces the point that the presence of genetic variation is necessary for reaction norms to evolve to their optima, but at the same time the continued presence of genetic variation remains a phenomenon that itself must be explained.

Environments with Both Stochastic and Predictable Components

The most realistic scenario of the environment is that it has both stochastic and predictable elements. Consider, for example, the problem faced by a organism in which its propagules (eggs or seeds) can survive frost, and so can overwinter, but all other stages are killed by a frost. Now suppose the season available for growth and reproduction is long enough to accommodate on average two generations. Individuals that mature early in the season will lay direct developing eggs that give rise to the second generation. Individuals of the second generation will all lay diapausing eggs that will overwinter. However, individuals of the first generation that emerge late in the season face the problem that there may be insufficient time for a second generation in which case they should lay diapausing eggs. If the season is long enough to accommodate a second generation, diapause eggs have a lower fitness than direct developing eggs, because the former give rise to only a single descendent for that year whereas the latter give rise to many descendants. In an environment in which the day of the first frost always occurred on the same day or could be predicted without error by the first generation females, there would be a sharp switch between the production of diapause and direct developing eggs. On the other hand, in general, the day of the first frost could only be assigned as a probability function: under this circumstance there will exist a period during the season during which the most fit strategy will be to lay a mixture of diapausing and direct developing eggs, the proportion depending on the reliability of the cue.

The phenomenon in which a mixture of types is produced is termed bet hedging or risk spreading. It has been explored theoretically fairly extensively, but there have been few empirical tests. In particular, risk spreading could be achieved by a strategy of avoiding risks (e.g., switching to production of diapausing eggs whenever the probability of not getting in the second generation is less than 1), which has been termed conservative risk spreading or by a single genotype producing a range of phenotypes, which has been termed diversified risk spreading. In an extensive review of insect data Hopper (1999) concluded that there is little evidence that risk spreading has been a major factor in the evolution of insect life histories. The importance of risk spreading in the evolution of life histories remains to be satisfactorily investigated.

See also: Biodiversity, Evolution and. Biodiversity Generation, Overview. Phenotype, a Historical Perspective

References

Boyce MS (1988) *Evolution of Life Histories of Mammals*. New Haven: Yale University Press.

Charlesworth B (1994) *Evolution in Age Structured Populations*. Cambridge: Cambridge University Press.

Clutton-Brock TH (ed.) (1988) *Reproductive Success*. Chicago: University of Chicago Press.

Futuyma DJ (1998) *Evolutionary Biology*. Sunderland, MA: Sinauer Associates.

Hopper KR (1999) Risk spreading and bet-hedging in insect population biology. *Annual Review of Entomology* 44: 535–560.

Roff DA (1992) *The Evolution of Life Histories: Theory and analysis*. New York: Chapman and Hall.

Roff DA (1997) *Evolutionary Quantitative Genetics*. New York: Chapman and Hall.

Stearns SC (1992) *The Evolution of Life Histories*. New York: Oxford University Press.

Limits to Biodiversity (Species Packing)

LB Slobodkin, State University of New York-Stony Brook, Stony Brook, NY, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 729–738, © 2001, Elsevier Inc.

Glossary

Chemostat Apparatus for growing microorganisms in a continually replenished medium.
Community A multispecific aggregation of organisms in a particular location that may interact with each other.
Competition More than one species utilizing one or more common resources.
Ecological niche The set of requirements that must be met if a particular species is to survive. It is sometimes used to mean the place in which those requirements are met.
Ecosystem A region with more or less clear boundaries, which contains a particular set of species and may be characterized by some set of meteorological, climatological, and geochemical properties.
Invasive or alien species Species that have recently colonized some geographic regions different from the one in which they were initially described.
Isocline A line along which some property remains constant.
Multidimensional niche A tempero-spatial region defined by meeting a set of different requirements for viability of a particular kind of organism.
Niche dimensions Ranges of values of environmental measurements in an ecological niche. A range of temperatures, salinities, or oxygen concentrations may be niche dimensions for a population of fish.
Packing The placement of objects in a container.
Reification The assignment of empirical reality to the referenda of a word or theory, regardless of the existence of any such referenda.
Species diversity The number of different species in some area of interest. Global species diversity refers to all species. Local species diversity refers to some geographic region such as Hawaii or New York City.
Species packing The study of how species on the same trophic level coexist in a limited region or container.

Introduction: The Biological Packing Metaphor

Different locations vary conspicuously in the number of species that are present. Some, like coral reefs and tropical rain forests, are famous for their high species diversity, while others, like some Alaskan forests, are notably monotonous in the number of kinds of trees that are present. Differences in diversity have attracted many researchers, although as yet there is no one all-encompassing theory of species diversity. In fact, the overall term "species diversity" seems more useful at the level of scientific policy and public relations than it is at the level of scientific discourse.

There are many theories and models of species diversity, which differ in their foci and assumptions. The theory of species packing is concerned with one aspect of diversity, specifically the coexistence of populations of different species located in a particular space. It carries elements of the concept of "species richness" in that the greater the species richness of a region, the more species have presumably been "packed" into it. It also carries implications for "realized species niches" because it might be inferred that the more species are packed into a region, the narrower the realized niches.

Like so many of the terms of theoretical ecology, "species packing" often is used in a metaphorical sense, which can be misleading. It is best to begin by analyzing the concept in a literal and elementary way.

The word "packing" has at least two meanings, which can cause confusion. It may mean placing objects inside a wrapper or container without altering them, or it may mean squeezing them in to a fixed-size container. Contrast the packing of glass tumblers with that of pillows. Tumblers are packed into a container but if they press hard against each other they shatter. Pillows can be squeezed into a container and fluffed up again when they are taken out.

The possibility of either type of species packing presupposes some sort of container into which things are packed. Is the inferred container equivalent to an environment? Does an environment have fixed geometric dimensions like a wooden box? A pond, a log, a rotting fruit—all can, in fact, be thought of as containers. However, environments can be defined in ways that may have more meaning to investigators than to inhabitants. Are the boundaries of a study area like La Selva, a political unit like Costa Rica, or a geographic area like Central America in some sense meaningful packing containers?

We will introduce the mathematical theory of species packing, examine how the theory has been exemplified in laboratory and field experiments, and, finally, consider the implications of species packing for field observations.

Two Species In One Container

Theory

The problem of species packing was crystallized in the apparently simple question: "Why are there so many kinds of animals?" Thereafter there has been a continual stream of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00088-5

often controversial, difficult, and sometimes acrimonious papers on this subject.

Lotka-Volterra Equations

Species packing theories are concerned with species in an implicitly assumed *container*. The boundaries of the container are usually considered to be permeable to the flow of resources and energy but, unless specified, impervious to the passage of organisms. These are important assumptions.

The formal theory of species packing originated from the equation systems of Lotka, Volterra, and D'Ancona. Lotka considered that in the entire biosphere all individual organisms have similar resource requirements, and no resource is in infinite supply. Therefore, at some level there must ultimately exist a limit to the abundance of organisms and an inverse relation between number of species and number of individuals per species. Further analysis is required before this assertion becomes interesting.

Volterra and D'Ancona considered several species competing with each other for resources in a space through which resources flowed at a constant rate. Competition between species A and B can be considered in terms of pairs of isoclines on a two-dimensional phase space with axes N_A and N_B:

$$\frac{dN_A}{dt} = 0 = \mathrm{K_A} - \mathrm{N_A} - \alpha \mathrm{N_B}$$

$$\frac{dN_B}{dt} = 0 = \mathrm{K_B} - \mathrm{N_B} - \beta \mathrm{N_A}$$

in which the subscripts refer to the two species. The K values refer to the number of organisms present at what was called "population equilibrium" or "population saturation."

It was assumed that any single species population would remain numerically constant after K had been achieved, if the flow of resources and the physical conditions did not change. These simple equations are the starting point for many species packing theories.

The K and N values were initially considered as counts of organisms, but as the theory developed the K values were taken to designate some resource divided up among the organisms. It often helps to think of species abundance as being directly translatable to rate of consumption of a temporally renewable resource.

It was usually assumed that any individual organism will negatively affect the well-being of all the others in the designated space, as measured by their birth and death rates.

Organisms can affect each other by reducing the availability of resources. They can also alter the chemical or physical properties of the local environment. These alterations are referred to as "conditioning" or "crowding." Yeast cells secrete alcohol, which affects other yeast and bacteria. Beetles and mice change the odor of the environment. Mussels and barnacles physically crowd each other.

In elementary packing theory, a constant, α, represents the relative amount of the resources of species A consumed by each individual of species B and, conversely, β represents the crowding of the individuals of species B by one individual of species A. In the multispecies case, α_{ij} is taken as the effect of species j on members of species i.

The Ecological Niche

The possible outcomes for two species were often represented graphically in elementary texts by using a phase diagram with abundance or concentration of each of the two species as axes. The persistence of both species populations is only possible if the two isoclines cross.

For species N_A, the isocline connects the point ($N_A = K_A$, $N_B = 0$) to the point ($N_A = 0$, $N_B = K_B/\alpha$). At each point on this isocline, individuals of the relevant species are crowded enough that births and deaths are equal. The crowding is exercised by either con-specifics, the members of the other species, or some mixture of the two.

If two species are inoculated into a space in which each species can survive if alone, there are basically three possible outcomes:

- If an individual of one species impacts the con-specifics more than hetero-specifics, both species may persist. The two isocline cross at a "knot."
- Another possibility is that individuals of both species are more sensitive to the presence of the other species than to the presence of con-specifics. In this case, the outcome depended on the initial concentrations of the two species. The isoclines cross at a "saddle."
- The final case is one in which one species has a stronger effect on the other, but the second species does not have a very strong effect on the first. The isoclines do not cross.

These three cases are related to simple definitions of the ecological niche. In the first case, the container is considered to contain regions from two ecological niches. In the second case, the container is in the intersection between two niches, and in the third case the container is within the niche of the victorious species.

Laboratory Experiments

While the elementary theory is primarily of pedagogical rather than practical value, it did inspire several laboratory population studies in which populations of two species were actually placed in a container, provided with renewable resources, and permitted to interact.

The early investigations on macroscopic organisms required long periods of observation and repeated tedious counting. Some later laboratory population experiments were done. Current studies of laboratory population dynamics mainly use microorganisms and are primarily of genetic and evolutionary significance. Even in these studies there are significant ecological differences between results derived from liquid culture in test tubes, in chemostats, and those from colonies on semisolid gels.

Without attempting to review all laboratory studies it is possible to summarize insights that have proved relevant in designing field studies and that have suggested theoretical analyses or may be expected to do so in the future.

- Laboratory studies have demonstrated that a single container can be divided in ways that can be more interesting than simply differences in quantitative utilization of resources.

For example, when populations of a moth and of a beetle were placed in a container with intact wheat grains or with flour, the moths were eliminated. Both species survived in cracked wheat. Broken bits of glass capillary tubing added to flour also permitted two species persistence. The moth larvae were being eaten by the beetles unless they were sheltered inside the grains or in the glass tubes. This illustrated that spatial complexity can modify the relation between competing populations. In the absence of physical refuges, the relationship between the moth and the beetle becomes a predator-prey interaction combined with some competition for food. Examples in which the interaction between several species can vary between competition and predation are not uncommon.

No populations, except for those of some microorganisms, grow according to the Verhulst Pearl logistic equation of sigmoid growth. This is interesting because there still are those that will defend the use of this counter-to-fact equation in development of theory.

- As populations become more crowded there are changes in birth and death rates. These changes are always accompanied by other physiological, anatomical, and behavioral changes.

There is abundant evidence of psychological and physiological change caused by individual social history in mammals. While the effects of past history may be most dramatic in mammals, they have been found in essentially all animals and plants. Crowded trees show different shapes than those grown in uncrowded situations.

- The levels of population density that can be reached during population growth experiments are very much in excess of what the organisms would tolerate in nature, as indicated by the ubiquitous tendency of animals to escape from these containers in any way possible.

In one study it was possible to compare responses of house mice in closed and open spaces. When the confined mice became sufficiently numerous to empty their food trays they immediately developed reproductive and behavioral pathologies. Unconfined mice remained in a local area around their food trays until they could empty them after each feeding. They then immediately dispersed.

Flour beetles when crowded move to the surface of their medium and show a strong tendency to fly. This cannot be demonstrated in population vials but becomes apparent if the flour containing a crowded population of beetles is removed from the vials or the vials are unstopped (personal observation). Dispersal when crowded is also found in other grain pests.

In Daphnia, winter eggs (ephippia) are produced at an early stage of food stress. This permits descendants of the population to survive despite whatever occurs in the experimental container. Any pond or lake that contains Daphnia or other Cladocerans will contain abundant ephippia in its sediments. These may hatch after a long dormant period, introducing genetically distinct animals from an animal seed bank for many years, even if the emergent animals of a particular genotype cannot survive in most years.

Even Hydra, which have no locomotory organs, respond to crowding and food shortage by floating free of their substrate, which removes them from a locally crowded situation.

- Laboratory populations can sometimes be used to make some inferences about what may be occurring in nature, but the inferences are not necessarily direct.

Early students of Daphnia assumed that Daphnia *condition* the water in which they lived. It was later demonstrated that the sizes of Daphnia populations in the laboratory, over a crowding range 10 times that ever observed in nature, were linearly dependent on food supply. If conditioning was of importance, this linearity would be impossible and there would have been a curvilinear relation similar to that found in the relation between nutrient flow and population size in chemostat cultures of microorganisms.

This does not imply that Daphnia populations in nature are regulated by their food supply. It does imply that they are not regulated in any significant way by causing chemical deterioration of their environment, because over a range of crowding that greatly exceeds any that has been found in nature, no chemical effect has been found.

Clearly, laboratory experiments and theoretical models cannot adequately imitate nature. However, an experiment can be designed to consider a wider range of conditions than those that occur in nature. In that sense, nature becomes a subset of the experimental world.

In some cases there is a clear logical transition between theory, laboratory experiment, and field experimentation. Gause mathematically considered predation on a two-species competition system. He concluded that the coexistence of two species could, in principle, be stabilized by predation.

In laboratory populations, the coexistence of brown and green species of hydra was facilitated by either removal of some of the animals or by maintaining them in low illumination, demonstrating that coexistence could be stabilized by either a biological or a physical change in the container.

In a field experiment, Paine found that heavy predation of starfish on mussels prevented the mussels from eliminating their competitors. In this case, theory proposed a possibility, laboratory experiments demonstrated its reality, and the field experiment demonstrated its importance in nature. But even in this case there were dangers of misinterpretation. The theory and the laboratory and field experiments demonstrated that under some circumstances a predator might enhance the number of surviving species on the next lower trophic level.

Any loss of organisms imposed on a multispecies system relieves, to some extent, competitive pressure on the survivors, so that weaker competitors may survive. There is, however, an abundance of cases of predators locally eliminating prey species, as is now occurring due to feral cats in Australia and Hawaii.

Expanding the Theory of Packing

Many formal theories of species packing begin by complicating this two-species situation. Among the obvious complications are those in which the isoclines are permitted to be

curved lines, which may be interpreted as social interactions of various sorts. Also, changing the sign of the interactions could crudely mimic predation or symbiosis. Loss of organisms across the "container" boundary could be thought of as mimicking predation.

Expansion of the Theory to Multiple Species

The more general question is: How many species can be introduced into a space? Multiple species packing theory is mathematically more difficult because a multispecies theory cannot be conveniently represented as isoclines on a two-dimensional phase space. If a third species were to be added, the phase space would have to be three-dimensional and isoplanes would be needed instead of isoclines. For coexistence of three competing species, the three isoplanes would have to intersect somewhere in the region below the intersection of any two of the species isoclines.

In the case of multiple species packing, there can be a region in the multidimensional phase space in which resources are divided in such a way that all species can share, but only if each species is relatively strongly selflimiting and relatively mild in its effects on the other species present. The multiple species theory requires clarification of the notion of *ecological niche*. The ecological niche of a species is often considered as the full set of measurements of environmental properties that are relevant to the survival and persistence of that population. Hutchinson and his students developed this into an image of a multidimensional hyperspace.

If two hyperspaces do not contain common dimensions, then the occurrences of the species concerned are independent of each other. As more and more dimensions are identical, or very similar, the stronger are the competitive interactions between the two species. Competition is one partial explanation for replacement of similar species along gradients in such physical features as temperature and salinity.

How Similar can Coexisting Species be?

MacArthur and Levins asked: What might be expected if a third species invades a space in which two species are already resident? They simplified the situation by considering that all species concerned interact through competition on a single niche dimension, or at least a linear representation of a multidimensional niche. They also assumed simple distribution of resources along that dimension. If, for example, the niche dimension is some particular prey, one of the species might eat large specimens while the other eats only small ones.

They considered that the niche of any species would have a shape on any single niche dimension. If the niche can be represented by a straight line, the probability of an individual finding a bit of resource at some point on the line can be represented as an ordinate, generating a two-dimensional shape. If the shapes for several species overlap, then the intensity of competition is measurable in terms of the probability of the several species striving for the same resource—the product of the individual species-specific probabilities.

They could then consider species packing along a niche dimension. Given their assumptions, packing can be closer on any one dimension as the number of niche dimensions increases. If species are arranged in a one-dimensional array along a line, the closest stable packing is admissible when the niche spaces are rectangular and all K values are equal. These conclusions were strongly dependent on their assumptions.

More general conclusions were that closer packing is possible if niche dimensionality is high, niche breadth is small, and the environment is predictable and has high productivity. Centers of adjacent niches are in a multiplicative series of numerical value approximately 1.1.

If the single linear dimension of competition is considered more realistically, so that there is a multiplicity of niche dimensions, MacArthur and Levins concluded that interspecific competition could be less important. In fact, it is sometimes considered that multispecies packing involves as many niche dimensions as there are species. Tilman considers that this is only valid if the organisms are not fixed spatially (i.e., they are free to move about in the container).

Various authors, by slightly modifying the mathematical assumptions, could reach a variety of conclusions about the packing of niches. For example, it was possible to plausibly argue that niches could sometimes be packed infinitely closely.

MacArthur and Levins suggested, as a matter of convenience, that the shapes of niches projected onto a single niche dimension might be thought of as approximating normal distributions. Other shapes are, of course, imaginable.

Species packing became a rich field for construction of theories, which were not necessarily connected to specific biological observations and more. These theories each make slightly modified assumptions or focus on special cases of previous theories. For example, Roughgarden and Feldman considered the significance of shape differences as if the shapes could be empirically demonstrated. Distributions with "thick" tails permitted more species to coexist than those with thinner tails.

There were also studies of multispecies biological situations under natural or artificial circumstances, which used the term "species packing." In these the connection with species packing theory was sometimes explicit but very often the general conclusions of theory were presented in the context of field data without explicit theoretical derivation. This appears to be due to the fact that explicit formal theory production becomes very difficult once more realistic elements of natural complexity are involved. This is typical of mathematical models, at least prior to the recent development of powerful computing machinery.

Sometimes the models were sufficiently complex, combining ecological and evolutionary predictions, that self-contradictory assertions were made within the same theoretical argument and were not noticed for years. For example, one study published in 1983 assumed absence of any genetic heterogeneity and then proceeded to discuss the genetic effects of selection, as noted by Taper and Case.

Niche Separation and Anatomical Difference

The exact ratio of minimum possible distinction between adjacent species can be derived from special, simplified, and

somewhat arbitrary assumptions. Depending on assumptions, it can be asserted that an infinite number of species can coexist, that an infinite number of species could coexist were it not for statistical variance in the parameters, that coexistence is contingent on number of dimensions in the hyperspace, or that species must differ anatomically by a specific ratio in order to coexist.

Hutchinson suggested the possibility that one might find in nature that niche hyper-volumes of the most closely similar, often congeneric, coexistent species would actually differ from each other by a factor whose value might be approximately 1.3. This seemed reasonable based on the common observation that genera are often represented in particular locations by large, small, and medium-sized species. Attention was focused on this by Lack for Darwin's finches, but its occurrence can be easily confirmed by examining intertidal snails of the American northeast coast or the eastern shores of the Mediterranean Sea (personal observation).

"Assembly rules" for competing species on islands were developed in part on the basis of size. These results were criticized on statistical grounds. Were the species distributions really departures from random?

One major step was missing in attempting to demonstrate that a theoretical niche overlap in a simplified mathematical model predicted a morphological ratio between anatomical parts. While Hutchinson expected that trophic structures like jaws might show the appropriate tightly packed ratio, the theoretical connection between niche shape or size and body shape or size seems absent except in very special cases.

It was never really clear why there should be a relation between the overlap of niche hyperspaces and any obvious anatomical differences. There is no rigorous and general way of theoretically deciding which of the infinite possible number of anatomical measurements that might be made is appropriate in any given case. By choosing the measurements correctly or by manipulating the data, almost any ratio might be achieved. Simberloff provided a firm critique of measuring the niche separation by anatomical differences. There is a clear connection between body size and the utilization of preexisting nest holes by snakes, owls, chipmunks, weasels, and others, but any pair of nest hole utilizing species is almost certain to differ in many niche dimensions.

Species Diversity and Species Packing

Does species packing set limits to species diversity in nature? This would require that natural aggregations of organisms are organized into communities and their component guilds, which would act as "containers." Are natural aggregations of organisms in nature strongly mutually interdependent so that "packing in" more species or removing species already present should have discernible effects on those already present? Do cohesive communities of organisms or ecosystems actually exist in nature? Early ecosystem theorists believed in a "holistic" or emergent concept of ecosystems and communities. If these were real, they would constitute containers in which packing theory might be applicable.

There is an enormous literature related to how many species occur in particular locations. Recently there has been a weakening of the basic notion of a natural community as a "container" in any serious sense. There is massive concern about the dangers of invasive species, but recent studies have returned emphasis to individual species and almost discarded the idea of integrated communities. There is a general condemnation of invasive species but they are particularly interesting from the standpoint of species packing. If invasions do not occur, it may be due to lack of opportunity. Large mammals cannot float across an ocean on logs or hide away in ballast tanks. However, if a species can invade a region, then in some sense the region was not packed full, or perhaps successful invasions necessarily result displacements of resident species.

The apparent invader purple loosestrife is sometimes found in dense stands that keep out other species, but other studies find that purple loosestrife had no apparent effect on native plant species and may have benefited native insect diversity.

Ruiz estimated that more than 90% of alien species in estuaries have made no discernible impact on the species diversity or species abundance distribution of the estuary. Levine and D'Antonio report that a "consistent positive relation between exotic species abundance and resident species diversity … [suggesting] that invaders and resident species are more similar than often believed." Apparently some invasive species fit into niches that were in some sense empty.

To what degree is competition, and therefore species packing theory, a serious factor in nature and how could one tell? Attempts to answer this question sometimes hinged on naturalists' insights, sometimes on various statistical analyses, and sometimes on detailed natural history and experiments conducted in more or less natural circumstances. Each of these has its strengths, advocates, and opponents.

In what sense do communities depend for their continuity on particular species? American chestnut trees constituted almost a third of the large trees of the southeastern American forests in the early 19th century. While small specimens still exist, the chestnut trees were essentially wiped out by the Chestnut blight. The forest did not disappear. As noted by Hairston *et al.*, the southeastern forest is as dense with trees as it has ever been. The general appearance of the forests have not changed, but the oaks, which have to a large degree replaced the chestnuts, produce volatile organic compounds that are of biogeochemical significance (M. Lerdau, personal communication, 2000).

Davis, studying the northward expansion of forests after glacial retreat, noted that each individual species seemed to migrate at its own rate—there was no movement of the forest as a community marching together. Whittaker noted that distribution of trees on various environmental gradients was as individual trees, not as communities.

Conclusions

Theory, experiments, and natural history all suggest that communities are not tightly organized, so that species packing may not be strongly relevant in nature. In short, the popular image of an ecological community as an airplane in which each part has a vital role for the integrity of the whole is

dubious. There may be groups of species, in which each one is closely connected to a few others but only loosely connected to other groups.

The term "community" was once extremely useful and is still of pedagogic value if carefully used. As reified objects for research, the concept of communities is now threatening to become what Simberloff and Dayan have referred to as a "panchreston," an idea as likely to generate confusion as enlightenment.

The prospect of a general species packing theory has melted away. Despite the unreality of the models, they did direct attention to what seemed to be a real phenomenon and encouraged experimentalists and field biologists to ask important questions.

Differences in species richness can be partially explained by a multiplicity of factors that do not necessarily relate to each other so as to permit formation of a single coherent theory. Each of the multiplicity of theories of diversity focuses on one or a few of the empirical factors that are known to enhance or diminish the possibility of species coexistence.

Many of these can be seen in laboratory experiments, which have the advantage of clarity but may be of questionable applicability. For example, local geometry complexity influences species diversity in nature as well as in the laboratory.

Certainly in some cases the term "species packing" is used in the sense of species being squeezed into a space. In several cases individuals of particular species in species-rich regions are believed to have a narrower range of activities than individuals of the same species in species-poor situations. Diet or nest sites may be more restricted. In these cases the individual organisms can be imagined to have been constricted by some packing process, like individual pillows in a crate but even this has two possible meanings depending on whether we are concerned with the population level or with individuals.

In comparing different locales, the range of variation among organisms within a unispecific natural population may be reduced when more other species are present. In the case of comparisons of different islands or lakes, this might be tentatively attributed to species packing.

Returning to the initial analogy of packing actual objects, an island is clearly a container. But if packing means filling the ecological space, either by pillows or tumblers, we would not expect islands or speciose lakes to easily admit invasive species. In lakes there may be enormous species richness of fishes, as in the ancient African lake cichlids, but I don't know of any comparisons showing that species-rich lakes show less within species, among individual variation, than species-poor lakes. Is the attempt to crowd multispecific collections of fish together in the same lake equivalent to packing glass tumblers rather than pillows? There is a general impression that species rich systems seem at least as likely to be invaded by exotic species as species-poor systems, violating our sense of what packing might mean.

If the container walls are not apparent but among individuals variation is reduced in one or more populations of a species, is this a sign of species packing? Does the mere fact that among individuals variation is reduced when more species are present imply that there must exist a container wall which may not be obvious?

The overall conclusion is that the theory of species packing does not conveniently predict very much about natural systems but that the images of packing that it generates do informally suggest interesting phenomena to look for.

See also: Competition, Interspecific. Diversity, Community/Regional Level. Habitat and Niche, Concept of. Plant Invasions. Species–Area Relationships. Species Coexistence. Stress, Environmental

References

Abrams P, Nyblade C, and Sheldon S (1986) Resource partitioning and competition for shells in a sub-tidal hermit-crab species assemblage. *Oecologia* 69: 429–445.

Adams DC, di Biteti MS, Janson CH, Slobodkin LB, and Valenzuala N (1997) An "audience effect" for ecological terminology: Use and misuse of jargon. *Oikos* 80: 632–636.

Anderson MG (1995) Interactions between Lythrum salicaria and native organisms: A critical review. *Environmental Management* 19: 225–231.

Ayala F (1971) Competition between species: Frequency dependence. *Science* 171: 820–824.

Barrett GW, Peles JD, and Odum EP (1997) Transcending processes and the levels-of-organization concept. *Bioscience* 47: 531–535.

Bazzaz F, Ceballos G, Davis M, *et al.* (1998) Ecological science and the human predicament. *Science* 282: 879.

Berryman AA (1992) Intuition and the logistic equation. *Trends in Ecology and Evolution* 7: 316.

Boake CRB and Wade MJ (1984) Populations of the Red Flour Beetle Tribolium castaneum (Coleoptera, Tenebrionidae) Differ in Their Sensitivity to Aggregation Pheromones. *Environmental Entomology* 13: 1182–1185.

Burgman M and Lindenmayer D (1998) *Conservation Biology for the Australian Environment*. Chipping Norton, NSW: Surrey Beatty and Sons.

Cornell H and Lawton J (1992) Species interactions, local and regional processes, and limits to the richness of ecological communities: A theoretical perspective. *Journal of Animal Ecology* 61: 1–12.

Crombie A (1945) On competition between different species of graminivorous insects. *Proceedings of the Royal Society of London Biological Sciences* 132: 362–395.

Crombie A (1946) Further experiments on insect competition. *Proceedings of the Royal Philosophical Society of London Series B* 133: 76–109.

Daily GC, Alexander S, Ehrlich P, *et al.* (1997) Ecosystem Services: Benefits Supplied to Human Societies by Natural Ecosystems. Washington, DC: Ecological Society of America.

D'Ancona U (1954) *The Struggle for Existence*. Leiden: E. J. Brill.

Davis M (1986) Climatic instability, time lags, and community disequilibrium. In: Diamond J and Case T (eds.) *Community Ecology*, pp. 269–284. New York: Harper and Row.

Diamond J and Case T (1986) *Community Ecology*. New York: Harper and Row.

Dykhuizen D (1978) Selection for tryptophan auxotrophs of Escherichia-coli in glucose-limited chemostats as a test of energy-conservation hypothesis of evolution. *Evolution* 32: 125–150.

Elton C (1958) *The Ecology of Invasions by Animals and Plants*. London: Methuen and Co.

Feldman MW and Roughgarden J (1975) Populations stationary distribution and chance of extinction in a stochastic environment with remarks on theory of species packing. *Theoretical Population Biology* 7: 197–207.

Ferson S, Stewart S, Downey P, *et al.* (1986) Competing reviews, or why do Connell and Schoener disagree? *American Naturalist* 127: 571–576.

Fryer G and Iles TD (1972) *The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution*. New York: Crown.

Gause GF (1934) *The Struggle for Existence*. Baltimore: Williams and Wilkins.

Glasser JW (1983) Variation in niche breadth with trophic position—On the disparity between expected and observed species packing. *American Naturalist* 122: 542–548.

Griffing TC (1965) *Dynamics and Energetics of Populations of Brown Hydra*. Ph.D. Thesis, University of Michigan.

Hairston N Sr. (1991) *Ecological Experiments: Purpose, Design and Execution*. Cambridge: Cambridge University Press.

Hairston N, Smith F, and Slobodkin L (1960) Community structure, population control and competition. *American Naturalist* 94: 421–425.
Hairston NJ, Lampert W, Caceres C, *et al.* (1999) Rapid evolution revealed by dormant eggs. *Nature* 401: 446.
Harvey P, Colwell R, Silvertown J, and May R (1983) Null models in ecology. *Annual Review of Ecology and Systematics* 14: 189–211.
Hopf FA and Hopf FW (1985) The role of the Allee Effect in species packing. *Theoretical Population Biology* 27: 27–50.
Hutchinson GE (1959) Homage to Santa Rosalia or Why are there so many kinds of animals? *American Naturalist* 93: 145–159.
Hutchinson GE (1978) *An Introduction to Population Biology.* New Haven: Yale.
Kinzig AP, Levin SA, Dushoff J, and Pacala S (1999) Limiting similarity, species packing, and system stability for hierarchical competition-colonization models. *American Naturalist* 153: 371–383.
Lack D (1947) *Darwin's Finches: An Essay on the General Biological Theory of Evolution.* Cambridge: Cambridge University Press.
Leopold A (1966) *A Sand County Almanac: With Other Essays on Conservation from Round River.* New York: Oxford University Press.
Levine J and D'Antonio C (1999) Elton revisited: A review of evidence linking diversity and invasibility. *Oikos* 87: 15–26.
Lomnicki A and Slobodkin LB (1966) Floating in hydra littoralis. *Ecology* 47: 881–889.
Lotka AJ (1925) *Elements of Physical Biology.* Baltimore: Williams and Wilkins.
Lotka AJ (1934) Theorie analytique des associations biologiques. *Actualites Scientifique et Industrielles* 187: 1–45.
MacArthur R (1972) *Geographical Ecology.* New York: Harper and Row.
MacArthur R and Levins R (1967) The limiting similarity, convergence, and divergence of coexisting species. *American Naturalist* 101: 377–386.
Mahmood T, Ahmad MS, and Ahmad H (1996) Dispersion of stored grain insect pests in a wheat-filled silo. *International Journal of Pest Management* 42: 321–324.
May R (1972) *Model Ecosystems.* Princeton: Princeton University Press.
Mullin B (1998) The biology and management of purple loosestrife (Lythrum saliceria). *Weed Technology* 12: 397–401.
Odum EP (1992) Great ideas in ecology for the 1990s. *Bioscience* 42: 542–545.
Paine RT (1966) Food web diversity and species diversity. *American Naturalist* 100: 65–75.
Pantasticocaldas M, Duncan KE, Istock CA, and Bell JA (1992) Population-dynamics of bacteriophage and Bacillus-Subtilis in soil. *Ecology* 73: 1888–1902.
Pianka E (1975) Niche relations of desert lizards. In: Cody M and Diamond J (eds.) *Ecology and Evolution of Communities*, pp. 292–314. Cambridge: Harvard University Press.
Pimm S (1991) The Balance of Nature? *Ecological Issues in the Conservation of Species and Communities.* Chicago: University of Chicago Press.
Ritte U (1969) *Floating and Sexuality in Laboratory Populations of Hydra Littoralis.* Ph.D. thesis, University of Michigan.
Roughgarden J (1974) Species packing and competition function with illustrations from coral-reef fish. *Theoretical Population Biology* 5: 163–186.
Roughgarden J and Feldman M (1975) Species packing and predation pressure. *Ecology* 56: 489–492.
Ruiz GM and Fofonoff P (1999) Non-indigenous species as stressors in estuarine and marine communities: Assessing invasion impacts and interactions. *Limnology and Oceanography* 44: 950–972.
Schoener TW (1985) Some comments on Connell and my reviews of field experiments on interspecific competition. *American Naturalist* 125: 730–740.
Simberloff D and Boecklen W (1981) Santa Rosalia reconsidered: Size ratios and competition. *Evolution* 35: 1206–1228.
Simberloff D and Dayan T (1991) The guild concept and the structure of ecological communities. *Annual Review of Ecology and Systematics* 22: 115–143.
Slobodkin L (1954) Population dynamics in Daphnia obtusa Kurz. *Ecological Monographs* 24: 69–88.
Slobodkin L (1961) *Growth and Regulation of Animal Populations.* New York: Holt, Rinehart and Winston.
Slobodkin L (1964) Experimental populations of Hydrida. *British Ecological Society Jubilee Symposium. Journal of Animal Ecology* 33(Supplement): 131–148.
Slobodkin L (1992) *Simplicity and Complexity in Games of the Intellect.* Cambridge: Harvard University Press.
Slobodkin L (1993) Scientific goals require literal empirical assumptions. *Ecological Applications* 3(4): 571–573.
Strecker R and Emlen J (1953) Regulatory mechanisms in house-mouse populations: The effect of limited food supply on a confined population. *Ecology* 34: 375–385.
Taper ML and Case TJ (1992) Coevolution among competitors. In: Futuyma D and Antonovics J (eds.) *Oxford Surveys in Evolutionary Biology,* vol. 8, pp. 63–109. Oxford: Oxford University Press.
Tilman D (1994) Competition and biodiversity in spatially structured habitats. *Ecology* 75: 2–16.
Treberg M and Husband B (1999) Relationship between the abundance of Lythrum salicaria (purple loosestrife) and plant species richness along the Bar River, Canada. *Wetlands* 19: 118–125.
Volterra V (1926) Variazione e fluttuazione del numero d'individui in specie animali conviventi. *Mem. Accad. Naz. Lincei* 2: 31–113.
Whittaker R (1970) *Communities and Ecosystems.* New York: MacMillan.
Yan GY, Stevens L, Goodnight CJ, and Schall JJ (1998) Effects of a tapeworm parasite on the competition of Tribolium beetles. *Ecology* 79: 1093–1103.
Yoshiyama RM and Roughgarden J (1977) Species packing in 2 dimensions. *American Naturalist* 111: 107–121.

Literary Perspectives on Biodiversity

William Howarth, Princeton University, Princeton, NJ, USA

© 2013 William Howarth. Published by Elsevier Inc. All rights reserved.

Glossary

Contingency Possibility or uncertainty; an event that may occur but is not likely.
Ecocriticism Interdisciplinary study of literature, history, religion, and philosophy with an emphasis on places, evolutionary biology, and environmental problems.
Etiology The study of causes or origins; stories that explain the origins of phenomena.
Land ethic Valuing land as part of a biotic community, not merely as property.
Literature Imaginative and crafted writings, in the form of poetry, prose, fiction, or drama.
Nature The material world and the physical forces or processes that control it.
Pastoralism Agrarian life and work; also literary accounts of rural life, often simplified or idealized.
Places Areas of space with human association; also positions of status in society.

The idea of biodiversity traces back to early concepts of variety and unity, which have their sources in history, philosophy, and literature. These cultural fields have recognized the need for biological variety since the first days of recorded time. Literature comprises many acts of language, from ancient folk chants to poetry on the Internet, often for entertainment and instruction, the *dulce et utile* praised by Cicero in his writings on rhetoric. The main goal of literature is to expand and extend the human imagination. By simulating experiences and performing ideas, literature helps affirm the value of diverse places and species, through a wide range of styles and themes. Not all writings are considered literature, just as all experiments are not science. Like professional science, the ideal of literature implies standards of custom and cultural decree (Turner, 1985)

Literature and Science Relations

Over the centuries, literature has come to mean imaginative writing, the product of disciplined craft that creates drama, fiction, poetry, or essay. Critical theorists disagree as to whether the prime agent in literature is content or form, conveying a message or shaping an attractive verbal structure, like a story or a poem. In the balanced formula of British poet and critic Samuel Coleridge, poetry is "the best words in the best order," which affirms that an ideal of quality guides literary writers. But "quality" is an intrinsic value, one that clashes with the extrinsic aims of science. Science seeks an unbiased view of what nature is and how it operates, while literature observes those conditions through the filters of human belief and emotion.

Although novelist Snow (1959) believed that science and literature are opposing cultures, recent observers find more congruency in their aims. Both disciplines use observational methods to detect patterns and to record them for the benefit of posterity. Biologist Gould (1989) sees science and literature as forms of historical narrative that describe contingency, the recording of uncertain, unexpected events just beyond current knowledge. Literature uses contingency in plot sequences – as events connect and readers turn pages "to find out what happens next." Scientists unravel nature's "plots" by reconstructing them as predictive narratives.

Literary forms represent natural diversity with varying degrees of accuracy. Drama is imprecise because it relies on social situations and voiced sentiments, enacted on a physical stage. Painting and photography may document a place or species, but pictures without words are not sufficiently interpretive. Poetry (whether epic or couplet) and the various forms of prose (novel, story, essay, memoir) are considered more comprehensive because they both describe and evaluate, mingling sensory experience with calculated thought. The meaningful connection of image and word: literature pursues that goal in order to represent biodiversity for the benefit of human understanding.

Language in Science and Literature

Human beings share a common attribute in language, the instrument they use to construct meanings. The path of language moves from variety to unity: many words exist; when combined they form phrases that express thoughts. A thought may be social or personal, as it creates a send-and-receive cycle that communicates. To meet the human need to understand both surface and depth, literature often asserts that natural objects are not alien but connected to human emotions. As Henry David Thoreau wrote in *Walden* (1854), "Shall I not have intelligence with the world? Am I not vegetable mould myself?" Seeking knowledge of nature's variety is thus a mutual goal for literature and science, but sharp differences exist in their methods.

Scientists prefer to use words precisely, to convey strictly limited meanings. Literary writers recognize that words have long histories, in which meanings change across time and space. As philologists learned in the nineteenth century, words behave like organisms (growing from root to stem to branch; cross-breeding; and migrating to other habitats). Some words are viruses, infecting a host, while others lie in dead heaps, like sediment. Through these morphological changes, words shape human values to form and reinforce perceptions, beliefs, and opinions.

Words are also representational, using speech sounds to create signs for rain, snow, leaf, and rock. Combinations with other words produce symbolic and metaphorical overtones, suggesting layers of possible meaning. Ezra Pound's haiku-like poem, "In a Station of the Metro" (1916), compresses a startling comparison of culture and nature into two lines: "The apparition of these faces in the crowd;/Petals on a wet, black bough." Such powers of analogy give words a suggestive and speculative function, which lends itself to stories with a didactic purpose.

Etiological Tales of Nature

The earliest surviving fables and myths are etiological tales, told to explain how the world began, why rain falls, and when crops will bloom or die. Among prehistoric people, nature stories about diverse and complex phenomena became a form of cosmology, providing the basis for ritual ceremony, whether for medical or spiritual healing. Early ideas of disease, as a condition of imbalance or disharmony, encouraged the symbolic belief that the earth was also a vast 'body' that may be cleansed and healed. Some observers (Buell, 1998) note that medicinal rhetoric continues to shape current discourse on toxicity, pollution, plague, and other forms of ecodisaster.

Early tales are efforts to define the myriad natural forms that shape everyday life, from land and species to climate and seasons. At the same time, literature may portray human efforts to resist or control nature, commonly through acts of labor and art. Farmers drain wetlands in order to plant crops on dry land. They weave nets to lift fish from the sea, then bury the fish with seeds to help new life emerge. Stories about these actions form early contributions to understanding biodiversity by recognizing the complexity of nature and the intricacy of its related elements. In a Navajo creation tale, "Changing Woman and Her Hero Twins," the storyteller pours tinted sands on the earth, forming a series of objects and events that are as multiple and shape shifting as the identity of Changing Woman, an earth goddess who has the power to create.

Yet by emphasizing human impact, many early stories also teach that humans and nature are divisible and separate. That idea is distinctly Judeo-Christian and supports the very concept of "nature," meaning all that is born, rising from a source. No equivalent term appears in Asian culture, where the physical world connects humanity to the wider universe (Torrance, 1998). The Old Testament tells two different stories of human creation: man is made separately, in the image of God (Genesis 1:26), but is also formed from dust (Genesis 2:7). The first story appears to justify human dominion over other creatures, while the second suggests a humbler bond with the earth. Belief in a separate, special creation of human beings continues to guide modern fundamentalists, who fiercely resist Charles Darwin's unitary theory of evolution. By including two narratives, the Bible leaves open the problem of literal versus metaphorical interpretation.

Pastoralism in Poetry and Scripture

In the early civilized era, 4000–500 BC, ideas and tales of nature's variety emerged from the extractive activities of hunting, fishing, farming, herding, logging, and mining. By learning to harvest natural materials and process them into goods, humans became experienced observers of seasonal events, of plant and animal lives, and of locations favorable to cultivation. Land use and property rights became the basis for literacy, trade, law, and tribal identity (Crumley, 1994). Nations root themselves in physical places, and when people are forcibly displaced, as in the case of the early Jews, they endure their nomadic years with consoling legends of a Messiah, a warrior-shepherd who will return them to the lost homeland. Such early stories reveal a basic function of literature, to reflect the value of places and how human presence changes them.

In the Old Testament, land is an element made by God, then granted to humans to clear and own. They receive this gift in innocence, but later sin and must accept the burdens of mortality: labor, pain, disease, and death (Genesis 1–3). The status of wilderness is even less benign: uncultivated places, like mountains and deserts, display God's power and offer a test to the faithful. In the story of Noah (Genesis 6–9), God directs Noah to build a great Ark and bring to it at least two of every creature that lives, male and female. Then God floods the earth with rain, until only the Ark and its inhabitants survive. The story illustrates not just divine power, but also how human stewardship may preserve the earth's vast diversity of species. The later vision of Isaiah predicts a savior who will restore the earth to its unfallen state, when all species live together in harmony: "The wolf also shall dwell with the lamb, and the leopard shall lie down with the kid; and the calf and the young lion and the fading together; and a little child shall lead them" (Isaiah 11:6). This scene, oft painted by the nineteenth-century Quaker artist Edward Hicks, represents the differences between a creation marred by original sin and one healed by the world's innate bounty.

In times of constant social change, literature offered comforting accounts of bygone days and places. The very concepts of city and country, or urban and rural land, arose out of literature that invested both realms with a complex relation, possessing zones and boundaries of separation but also connection (Williams, 1973). As urban centers grew into the great city-states of Alexandria, Athens, and Rome, the poets Hesiod, Theocritus, and Virgil wrote pastoral poetry, a literary form known variously as bucolic, georgic, and eclogue; verses that portray rural life as gentle, principled, and close to the soil. Jesus of Nazareth also adapted pastoralism in his parables about fish and seeds to preach transcendence of earthly limitations, a literary device equally effective with audiences in rural Judea and in the fast-growing cities of the Mediterranean basin.

After the Roman empire shifted from paganism to Christianity in the third century AD, early Christians saw their pastoral mission as spreading the Gospels throughout Europe. Monastic enclaves preserved both classical science and literacy in the Dark Ages (400–800), between Rome's fall and the revival of learning in courts and cathedral towns. During the Middle Ages (800–1400), infectious plagues reduced populations and relieved pressure on Mediterranean ecosystems, still recovering from Roman-era deforestation, erosion, and pollution (Hughes, 1994). Medieval literature and science became twinned endeavors, as clerics and scholars sought to

describe the fixed Ptolemaic universe, an effort that culminates in Dante's epic poem, *The Divine Comedy* (1321). Medieval culture's fascination with alchemy likewise fed early studies of natural history and chemistry, described in several of Chaucer's *Canterbury Tales* (1400).

Colonial Expansion and Taxonomy

Chaucer's use of folk tales reflects the popular literacy that spread rapidly after Johannes Gutenberg invented movable type (1455), as printed tracts, manuals, and libraries of books fed a new hunger for learning. Literacy spawned the Renaissance (1500–1700), a period of European expansion into the continents of Africa, Asia, and the Americas. Global exploration refined mapping and navigation, while imperial conquest decimated native populations through warfare and disease. Discoveries of new lands and species also spurred growth in the descriptive sciences, especially geography, geology, and biology. A buoyant verbal fluency stirred the language arts, epitomized in the writings of Shakespeare, Cervantes, and Molière. Colonial empires used their literacy to record events and claim land as property. Native cultures often regarded words as totemic: their oral chants and stories were prayers, repeated to bless earthly cycles. Such radically different uses of language also expressed conflicting views of nature. Europeans sought to own the earth, even as their sciences and arts revealed the immense biodiversity of wild, untouched land.

With the rise of natural science after 1700, literature began to express growing respect for nature, especially those aspects of it that were rare and varied. Optical devices peered into the heavens and the human body, while theorists and engineers created new branches of knowledge, publicized and marketed by print, commerce, and coffeehouse (Jardine, 1999). New forms of technology changed the economic value of natural resources. Land remained a source of food and energy, but wealth increasingly came from products and services. Cities expanded by drawing rural tenants from the countryside; new social classes formed around labor, management, and capital. An expanding, literate middle class called for political reforms, replacing monarchy with representative democracy. Revolutions in the Americas pulled masses of immigrants to the new Western nations, where open territory and cheap land began to transform agrarian life into republican independence. (As the world has grown tamer in recent decades, wilderness has come to have a higher value, as land existing for itself, rather than for human enterprise.)

The study of natural diversity advanced through what Kuhn (1967) has called changing paradigms, theories or conventions that promoted experimentation. Early naturalists described organisms with confusing and inconsistent local vernacular until Carolus Linnaeus compiled *Systerna Naturae* (1734), using Latin nomenclature to assign names by genus and species. Latin was dead and thus fixed, and Linnaeus used social distinctions (kingdom, phylum, class, order, family) to create taxonomy, a systematic language that described the bounty and variety of nature. While literary authors from Jonathan Swift to Fenimore Cooper mocked scientists for obsessive categorizing, taxonomy became a consistent model for geology, archeology, and philology. Those sciences declared that prehistoric changes in the earth and in language follow traceable lines of descent across long stretches of time. Such discoveries replaced mythic accounts of creation with rational studies of the unity and diversity of organic events.

Darwinism and Literature

The new sciences spawned a second age of exploration, 1750–1850, as voyagers from James Cook to Alexander von Humboldt surveyed the continents of Australia and the Americas. In their wake followed a generation of literary naturalists like Gilbert White, William Bartram, and John James Audubon, who drew field sketches and wrote evocative accounts of seasonal journeys through rural or remote lands. The expeditionary narratives of Peter Kalm or Meriwether Lewis and George Rogers Clark are read today as literature and history, for they verify the immense variety of species that led to theories of biodiversity. This tradition of gentleman amateur molded Charles Darwin, who spent a 5-year expedition on the *HMS Beagle*, examining patterns in geology and zoology along the coasts of South America and on Pacific islands. Darwin was one of the last major scientists to be an independent field naturalist, living on his own income. After 1870 science moved to the laboratory bench and university departments, a shift in professionalism that also affected writers, who became teachers and seekers of literary prizes (Jardine, 1996).

In his major works, *Voyage of the HMS Beagle* (1839), *The Origin of Species* (1859), and *The Descent of Man* (1871), Darwin argued that species evolve through the effects of heredity, variation, and natural selection. He had several predecessors in science, from Georges-Louis Leclerc Buffon, who recognized adaptation, to Alfred Russell Wallace, who anticipated and clarified several of Darwin's ideas about selection. The term "evolution" is quite ancient, going back to Latin *evolvere*, to unroll, and may describe inscriptive scrolls, the writings that physically expressed a gradual process of change. The Biblical account of creation (Genesis 1) frames a sequence of originating events, each a division (night from day, land from water, plant from animal) that measures the binary logic of doubling, branching, and splitting, all recurrent patterns in nature. Poetry, tragedy, and biography all depict similar sequences of choice and accident, selection and reproduction. Those lines of cause and effect link the Bible to Homer and Sophocles, and then to Shakespeare or Milton. A constant theme in Western literature is the relation of creator, man, and nature, as chains of interactive events lead to human survival or extinction.

Literary traditions emerged from worship and theology, while Darwin stated that evolution proceeds according to physical laws, without divine intervention, and that human beings did not arise separately from other species. Early confirmation of these views appeared in the research of the Austrian monk Gregor Mendel, whose experiments in hybridization found that reproduction is a process of encoding and transmitting genetic material. Despite the strong controversy these views aroused among people of faith, by the 1860s evolution and genetic science began to alter popular

ideas of nature and produce the first authors who consciously recognized biodiversity.

Two leaders in this movement were Henry David Thoreau (1817–1862) and George Perkins Marsh (1801–1882). Their books (*Walden*, *Cape Cod*, *The Maine Woods*, *Man and Nature*) often examine the physical environment of New England, a region of low economic value but rich biodiversity. A self-trained land surveyor, Thoreau spent several years studying landforms and seasons. After reading *The Origin of Species* in 1860, he wrote the first modern accounts of forest succession and the dispersal of seeds (Howarth, 1982). Marsh, a lawyer and diplomat, published early studies of erosion, blaming land clearance and overgrazing, and he argued vigorously that human beings must learn to control their destructive impact on nature.

Among later writers known as social Darwinists (chiefly Herbert Spencer, Walter Bagehot, and William Graham Sumner), application of evolutionary theory to human behavior produced racist theories of cultural history, favoring white and Nordic peoples over Mediterranean and African. This bias, based on fears of immigration by darker races toward the wealthier nations, slants the novels of Jack London, Frank Norris, and Theodore Dreiser. Later these voices were replaced by modernist writers like T. S. Eliot, Willa Cather, and Ernest Hemingway, who saw racial diversity as healthy and progressive, a human check on industrial monopolies that had lain waste to natural resources and open spaces. Cather explored these issues on the plains of Nebraska, while Hemingway examined waning biodiversity in several lands, from upper Michigan to the savanna of east Africa, where great forests and wild herds are steadily destroyed by human incursion.

Ecology and Industrialism

The concepts of ecology were slow to gain acceptance among biologists and thus entered literary culture at a later date. Ernst Haeckel, a German biologist and advocate of Darwin, coined the term "ecology" in 1869 to describe how organisms form alliances that shape their number and distribution. Based in part on Haeckel's political socialism, his views also echoed rising concern that industrial growth in Germany, France, and England had begun to damage both physical environments and public health (Bowler, 1994).

A vivid picture of decline lay in the western United States, where rapid agrarian settlement after the Civil War destroyed native tribes, bison, and grasslands in a few decades. The conservation movement of 1870–1920 sought to stem those losses by creating public lands and parks for protection of natural resources. Yet in the popular arts, writers of dime novels and outlaw ballads glorified "winning the West," a form of historic triumphalism sanctioned by the theories of Frederick Jackson Turner (1893), who held that the frontier shaped fundamental American values of self-reliance and innovation. Later scholars note that the settlement process was destructive, as pioneers destroyed the very Eden they had sought (Mitchell, 1981).

With the experimental writings of James Joyce, Gertrude Stein, and William Faulkner, a modern critique of mass-industrial civilization began to emerge. Drawing on their perspectives as cultural outcasts (Irish Catholic, Jewish Lesbian, Southern Gentry), they mounted strong attacks on the spiritual vacuity of urban life and, in Faulkner's case, on the economic abuses that had sacked agrarian America. In his 14-novel saga about a fictional county in northern Mississippi, Faulkner assailed profiteering, an ideology that created slavery, miscegenation, one-crop farming, and soil erosion. He was the first modern novelist to comprehend ecological concepts, which gained attention in the 1930s from New Deal efforts to counter the wide-scale failure of capitalism.

The word "ecology" connotes a desire to coexist, if not in symbiosis, at least in a mix of competition and collaboration that marks the dynamics of a healthy, vigorous ecosystem. Darwin's image of nature was "an entangled bank," and throughout the twentieth century, ecologists created an interdisciplinary science that depicted nature as a web of interacting, interdependent forces that sustains itself – not as an unchanging constant, but as a vigorous dynamism sustained by frequent, patchy disturbances (Hagen, 1992).

Three writers who affirmed these principles, in sharply varying styles, were all Californians: Poet Robinson Jeffers, "hard-boiled" novelist Raymond Chandler, and John Steinbeck, a social novelist who was trained in marine biology and in the investigative methods of documentary journalism. In his literary works, especially *The Grapes of Wrath* (1939) and *The Sea of Cortez* (1941), Steinbeck built a vision of interlacing natural and cultural forces, locked in struggles that were shaped by biological conditions of production, consumption, and adaptation (Beegle, 1997).

Land Ethics and Environmentalism

After World War II, two writers emerged who were trained ecologists and willing to apply their influence to public policy. In their writings, Aldo Leopold and Rachel Carson issued firm warnings of environmental destruction by human means. Leopold's *The Sand County Almanac* (1949) celebrates glaciated wetlands in Wisconsin, poor in economic value but rich in its diversity of habitats. His farm lies at the intersection of prairie, forest, and marsh, forming a mosaic of ecosystems that measure the "intricate tangle" of nature. Out of this vision, Leopold sketches the story of how he abandoned early conservationist ideas (a wildlife ranger, he was paid to kill wolves) for a preservationist vision he calls "the land ethic," the sense that land deserves respectful protection as a vital element in the biotic community.

In *Silent Spring* (1962), biologist Rachel Carson intensified that sense of respect by analyzing the destructive force of chemical pesticides. In an incisive passage, she demonstrates how toxicity spreads through ground-water, a system of transport so invisible that it is easy to ignore. For her efforts, Carson was attacked by the petrochemical industry as hysterical and unscientific. But her logic and eloquence impressed the Kennedy administration, brought a ban on DDT and other pesticides, and eventually helped create the Environmental Protection Agency (Howarth, 1999).

Both Leopold and Carson saw that human values are formed by ideas of land, seeing that property is also a shared earth. Their books brought attention to emerging "green" political movements of the late twentieth century, from which two different land ethics developed: landscape, shaped by people

and inflected by aesthetic beauty, and land use, focused on food and energy extraction and marked by utilitarian security. Although often in conflict, the two ethics also reflect agreement that human identity rests on a sense of place, while alienation is feeling displaced, homeless, or unlanded (Spirn, 1999).

Natural Rights and Animal Fables

The perception that land deserves ethical regard soon led to extending rights to nonhuman species, both plant and animal, granting them value and respect while seeking to conserve their populations. Herman Melville's epic novel *Moby-Dick* (1851) anticipates this generosity, expressed in an iconoclastic rhetoric that denounces whaling. The cause of animal rights began with nineteenth-century reformers, who sought protection for religious dissenters, children, women, Indians, slaves, and other powerless groups. These figures came to be seen not as "lower" orders, but creatures having what Darwin called "the mutual affinities of organic beings." That phrase, the title of his penultimate chapter in *Origin of Species*, delivers his belief that humans are descended from all the creatures that have existed and still share the earth.

This theme enters literature in many animal tales written for young readers, from Anna Sewell's *Black Beauty* (1887) to E. B. White's *Stuart Little* (1945) and *Charlotte's Web* (1952). They descend from ancient beast and fairy tales, in which animals have the power to create language, before passing it on to humans. Such stories describe how things came into being, and they still have great attraction for writers drawn to environmental themes (Flynn, 1999). A contemporary example is David Quinn's novel *Ishmael* (1991), a series of dialogues between man and ape, in which the ape explains history as a conflict between Takers and Leavers, identified by their divergent approaches to the natural world. Takers dominate resources and evolve into modern urbanized man; Leavers are indigenous tribes who lead subsistence lives and eventually vanish.

While such accounts are fantastic, they also urge readers to think of animals, water, soil, rocks, and food as more than commodities. The cultural information provided in literary texts clarifies that natural objects result from processes, including the moral and ethical choices made by humans. As such choices recur in history, the value of a creature or a place changes: once the sea was thought to contain monsters, later it became a therapeutic place to bathe and sun; now it is emptying of fish and clogged with sewage (Hendrickson, 1984). Formerly the forest world was seen as profane and lawless, the home of pagans; now it is sacred, a haven for rest and retreat (Harrison, 1992).

Ecocriticism and Biodiversity

Scientists are now warning that a global environmental crisis will be the future outcome of several current trends: human overpopulation, the rise of carbon dioxide emissions, atmospheric warming, and the loss of biodiversity. These dangers transpire across vast frames of time and space that are difficult for nonscientists to imagine. A crisis that occurs in a locality, such as flood or earthquake, is understandable, but one that works over a longer period and through evolutionary selection is invisible to most eyes.

Because its domain is virtual or imaginative, literature helps its readers to envision what they will never directly experience. Yet many variables also affect the ability of literature to depict the threat to biodiversity. A writer's personal involvement in a story will shape its tone, bringing nostalgia (Wallace Stegner) or urgency (Bill McKibben) to the subject matter. The work's form may illustrate its content, as when Peter Matthiessen shapes *The Snow Leopard* (1973) into a mountaineering ascent, using elevation to provoke meditations on higher metaphysical themes. An object described will take on the character of its subject, as when Annie Dillard expresses her notions of wildness through a weasel's alleged indifference to pain.

The work may also have a distinct cultural audience (John McPhee) or veer between standards of elite and popular reception (Edward Abbey). Some writers are sensitive to social and economic conditions (Edward Hoagland), others contrast regional and national concerns (Barry Lopez), still others attend to differences of race and gender (Sue Hubbell), or to the impact of other arts on literature (Anne Matthews). Environmental writing is often a mixed collection of genres; part scholarly research, part sermon-editorial, part whimsy and adventure. Similar manifestations affect the contemporary novels of Thomas Pynchon (*Gravity's Rainbow, Vineland, Mason*, and *Dixon*) and Don DeLillo (*White Noise*).

Because of the great variation in environmental writing, a new mode of reading called "ecocriticism" has emerged to establish interpretive standards. While not agreeing in all respects, ecocritics often seek to translate the specialized language of science, with its jargon and acronyms, into a common vernacular. Ecocritics are interpreters and bridge builders, looking for ways to connect the disciplines that university departments have long separated (Glotfelty and Fromm, 1995). A few scientists, such as Wilson (1998), have written popular accounts of their environmental concerns, but mostly the public has turned to literary writers to gain insight into the causes and effects of waning biodiversity. At times these voices cannot agree about solutions, for the global biosphere is too intricate, too complex, and too unpredictable to describe readily, and some scientists believe that it will not be managed until it returns to preindustrial levels of stable composition (Peters, 1994).

Other ecologists have grave concerns about genetic erosion, through both species loss and also genetic engineering, which in the long term may produce fewer varieties of plants and animals (Karaim, 1999). The role of species biodiversity suggests that its purpose may be to provide redundancy, which has a stabilizing effect on whole communities and ecosystems. In the face of drastic environmental change, diversity may buffer ecosystems against the collapse of ecological function (Nudds, 1999).

If biodiversity works because variety protects a community against stress and disaster, then literary accounts of biodiversity must remain equally diverse and inventive for the sake of textual survival. A review of literary history suggests that the continuance of human life on earth will depend in part on what storytellers make of a changing world. The gift of

language brings insight, but also the wisdom to heed the silence that Annie Dillard calls "nature's one remark." The central question for human beings is not whether they will survive as a species, but whether the survivors will inherit a world worth sharing (Wilson, 1998).

Disaster and Mass-Media

In the first decade of the twenty-first century, several events eroded the optimism that once promoted hopes for sustained biodiversity. Consensus politics in the United States suffered from a series of blows: (1) the disputed election of a conservative president, George W. Bush; (2) attacks on 9/11/2001 by Islamic terrorists in New York and Washington, DC; (3) costly wars in Iraq and Afghanistan; and (4) a long, deep economic recession among developed nations. Hard times often create scapegoats, and in the second millennium, blame has shifted from corporations to government and environmentalists. Since the year 2000, Gallup polls of US public opinion have reflected declining trust in government, loss of economic confidence, and lower regard for environmental advocates (Dunlap, 2010).

At the same time, repeated instances of oil spills, earthquakes, storms, droughts, epidemics, and famine have brought increasing pressure on government and industry to provide fiscal and humanitarian relief from such disasters. A notorious event was the largest accidental marine oil spill in history, known as the *Deepwater Horizon* blowout of April–September, 2010, which caused immense damage to marine and wetland habitat in the Gulf of Mexico. The offending company, BP, acknowledged fault and paid for cleanup and reparations, then continued profitable operations in some 80 countries around the world. Whether the causes are natural or human, citizens expect immediate response, even while fiercely resisting new taxes and programs of assistance (Button, 2010).

This political turmoil has accompanied a profound revolution in mass-media, as the world of publishing and broadcasting has shifted from small nodes of power to a sprawling, intricately linked means of communication, the Internet. Originating in military research in the 1960s, the Internet went commercial in the 1990s and spurred the growth of personal computers and mobile devices, such as smart phones and tablets. Books, magazines, newspapers, radio, and film now circulate instantly in digital form. The effect on cultural production is clear: anyone may create a film, story, or picture and "publish" it on social media sites with little control by gate-keepers or public taste (Dennis and DeFleur, 2009).

For literature, the way ahead is uncertain. Ever since the 1800s, when writers left behind private patrons and embraced a reading public, literature has depended on a well-educated audience for whom reading is a leisure pastime. For two centuries, middle-class nuclear families sustained such popular literary genres as the novel, deriving from fiction a means of virtual experience. That social demography no longer prevails. Families are now two-income or single parent, commuting distances to work and having less time to read popular or serious books. The spread of cable television systems, broadcasting on hundreds of channels, has decimated the once dominant print culture of book stores, reviewers, and readers (Striphas, 2010).

At the turn of the twenty-first century, writers of drama, film, poetry and fiction dealt very differently with questions of biodiversity. Theatrical treatments of the issue have been few, though street and community theater in emerging nations increasingly address issues such as water conservation and public health. In the realm of commercial film, the Australian documentary *Cane Toads* (1987) highlighted the perils of introduced species, as *Microcosmos* (1996), *Winged Migration* (2001), and *March of the Penguins* (2006) celebrated habitat preservation. Two films, in particular, shaped cultural and environmental-policy conversation on a global scale: Albert Gore's *An Inconvenient Truth* (2006), on global climate change, and the pro-biodiversity epic *Avatar* (2009), which earned the highest box-office gross in history.

Between 1990 and 2010, a striking percentage of major literary prizes went to poets inspired by the natural world – Mary Oliver, Robert Hass, Paul Muldoon, W. S. Merwin. In postmodern fiction, by contrast, nature is often cast as a malicious or irrelevant Other. Technoculture, apocalypse, the posthuman, and the social construction of nature dominate the work of literary novelists from Cormac MacCarthy (*The Road*) and Don DeLillo (*White Noise*) to David Foster Wallace (*Infinite Jest*) and Jonathan Franzen (*Freedom*). As tolerance for social and cultural variation increases, hostility toward the natural remains perhaps the last acceptable prejudice in the literary arena. One exception to this trend is the work of Dana Hand, whose prize-winning *Deep Creek* (2010) turns landscape and history into an exploration of the violence and racial hatred at the heart of frontier settlement.

As real human diversity fades under the onslaught of the commercial and the electronic, so does natural biodiversity. Those who work to preserve it heed an early warning, "In wildness is the preservation of the world." Thoreau (2007) saw wildness as the mother of civilization, a rootstock that feeds the tree of life. His hope was that it could not vanish as long as we preserve literacy and write the stories that inform our long history of memory and imagination.

See also: Aesthetic Factors. Darwin, Charles (Darwinism). Historical Awareness of Biodiversity. Religious Traditions and Biodiversity. Social and Cultural Factors. Stewardship, Concept of

References

Beegle S (1997) *Steinbeck and the Environment: Interdisciplinary Approaches*. Tuscaloosa: University of Alabama Press.

Bowler P (1994) *The Norton History of the Environmental Sciences*. New York: W. W. Norton.

Buell L (1998) Toxic discourse. *Critical Inquiry* 24: 639–665.

Button (2010) Gregory Button. *Disaster Culture: Knowledge and UnCertainty in the Wake of Human and Environmental Catastrophe*. Walnut Creek, CA: Left Coast Press.

Crumley CL (1994) *Historical Ecology: Cultural Knowledge and Changing Landscapes*. Santa Fe, NM: American Research Institute.

Dennis EE and DeFleur ML (2009) *Understanding Media in the Digital Age*. Boston: Allyn & Bacon.

Dunlap RE (2010) At 40, Environmental Movement Endures, with less Consensus. *Gallup* (22 April 2010). http://www.gallup.com/
Flynn KM (1999) *Peaceable Kingdoms: Constructions of Animal Life in American Literature*. PhD Dissertation, Princeton University, pp. 1850–1950.
Glotfelty C and Fromm H (1995) *The Ecocriticism Reader: Landmarks in Literary Ecology*. Athens, GA: University of Georgia Press.
Gould SJ (1989) *Wonderful Life: The Burgess Shale and the Nature of History*. New York: W. W. Norton.
Hagen J (1992) *An Entangled Bank: The Origins of Ecosystem Ecology*. New Brunswick, NJ: Rutgers University Press.
Harrison RP (1992) *Forests: The Shadow of Civilization*. Chicago: University of Chicago Press.
Hendrickson R (1984) *The Ocean Almanac*. New York: Doubleday.
Howarth W (1982) *The Book of Concord: Thoreau's Life as a Writer*. New York: Viking Press.
Howarth W (1999) Imagined territory: The writing of wetlands. *New Literary History* 30: 509–539.
Hughes D (1994) *Pan's Travail: Environmental Problems of the Ancient Greeks and Romans*. Baltimore: Johns Hopkins University Press.
Jardine L (1996) *Cultures of Natural History*. Cambridge: Cambridge University Press.
Jardine L (1999) *Ingenious Pursuits: Building the Scientific Revolution*. New York: Little, Brown.
Karaim R (1999) Variety, the vanishing crop. *Washington Post, April* 11: B01.
Kuhn TS (1967) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
Mitchell LC (1981) *Witnesses to a Vanishing America: The Nineteenth-Century Response*. Princeton: Princeton University Press.
Peters RL (1994) *Global Warming and Biological Diversity*. New Haven, CT: Yale University Press.
Snow CP (1959) *The Two Cultures and the Scientific Revolution*. Cambridge: Cambridge University Press.
Spirn AW (1999) *The Language of Landscape*. New Haven, CT: Yale University Press.
Striphas (2010) Ted Striphas. *The Late Age of Print: Everyday Book Culture from Consumerism to Control*. New York: Columbia University Press.
Thoreau Henry D (2007) Walking, or the Wild, Excursions. In: Moldenhauer JJ (ed.) Princeton, NJ: Princeton University Press.
Torrance R (1998) *Encompassing Nature: A Sourcebook*. Washington, DC: Island Press.
Turner F (1985) *Natural Classicism: Essays on Nature and Science*. New York: Paragon House.
Williams R (1973) *The Country and the City*. New York: Oxford University Press.
Wilson EO (1998) *Consilience: The Unity of Knowledge*. New York: W. W. Norton.

Loss of Biodiversity, Overview

Robert Barbault, Museum National d'Histoire Naturelle CNRS, Paris, France

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity The variety of organisms considered at all levels, from genetic variants within the same species to the entire range of species and ecosystems.
Extinction The disappearance of any lineage of organisms, from populations to species and higher taxonomic categories (genera, families, phyla). It can be local or global (total).
Genetic drift The process of random sampling of genes that leads to changes in the genetic composition of a population. The effect of this process is particularly important in small population sizes.
Homozygous Possessing the same gene form (allele) on both chromosomes.
Locus The position on a chromosome of a gene. It is determined by any number of allelic forms.
Metapopulation The set of populations (or subpopulations) of the same species, linked by migration. It is characterized by the processes of local extinction and recolonization.
Pleistocene The geological time that ends with the last glacial period and the appearance of humans. It started 2 million years ago and finished 10,000 years ago with the end of the last Ice Age.
Stochastic The result of chance and random processes.

The extinction of a species, like the death of an individual, is a natural phenomenon – its inevitable destiny. In fact, during the long history of life, earth has undergone several periods of mass extinction. However, the crisis of current extinction differs from the preceding ones in that it is the direct result of human activities. Through our ecological success – have we not invaded the entire planet? – amplified by our industrial and technological revolutions, our species exerts such an impact on the biosphere that today, we are witnessing an acceleration of extinction phenomena without precedent. The objective of this article is to address the patterns, causes, and extent of such losses of biodiversity.

Lessons From The Past

The longest phase in the evolution of life on our planet extends for the 2 billion years from the appearance of the simplest molecules capable of autoreplication to the appearance of the first prokaryotes. Unfortunately, nothing is known about the evolution of biodiversity throughout this very long period, dominated by microorganisms. Our focus concerning extinction must therefore be restricted on two kingdoms, plants and animals, or in other words, on the past 600 million years (My).

Extinction Rates Deduced from the Fossil Record

Table 1 shows the estimates for the average life span of species in various groups of fossils from their origin to their disappearance. Overall, the average life span of the species is about 5–10 My.

Based on this estimate, and keeping in mind that the total interval of time considered is equal to 600 My, one can estimate that the present stock of animals and plants accounts for only approximately 1–2% of all species that have ever lived. However, this evaluation obviously depends on the quality of the fossil record. One should also note that the estimated life spans vary considerably according to the groups considered (Table 1). In particular, according to Raup, the average duration of mammal species is on the order of 1 or 2 My, much lower than the 10 My estimated for (mainly marine) invertebrates. Taking into account this variability, it is estimated that the current species of plants and animals account for approximately 2–4% of those that have ever existed. Rates of extinction tend to vary considerably over time (Figure 1). Thus, four of the five major extinction periods in the fossil record (Figure 1) each eliminated between 65% and 85% of the marine animal species (Ordovician, Devonian, Triassic, and Cretaceous), and one (Permian) eliminated 95% or more.

Mass Extinctions Since the End of the Pleistocene

During the past 50,000 years, several extinction episodes have occurred almost exclusively among terrestrial mammals, some birds, and, to a lesser degree, some reptiles. A comparison of the data gathered on the various continents makes it possible to refine the explanatory assumptions (Figure 2). In Africa,

Table 1 Estimates of species' life spans in various groups from origination to extinction[a]

Group	*Species' average life span in millions of years*
All invertebrates	11
Marine invertebrates	5–10
Marine animals	4–5
Mammals	1–2
Diatoms	8
Dinoflagellates	13
Planktonic Foraminifera	7
Bivalves	10
Echinoderms	6

[a]Based on data from various sources.
Source: Reproduced from May R *et al.* in Lawton JH and May RM (1995) *Extinction Rates*. Oxford: Oxford University Press, simplified.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00298-7

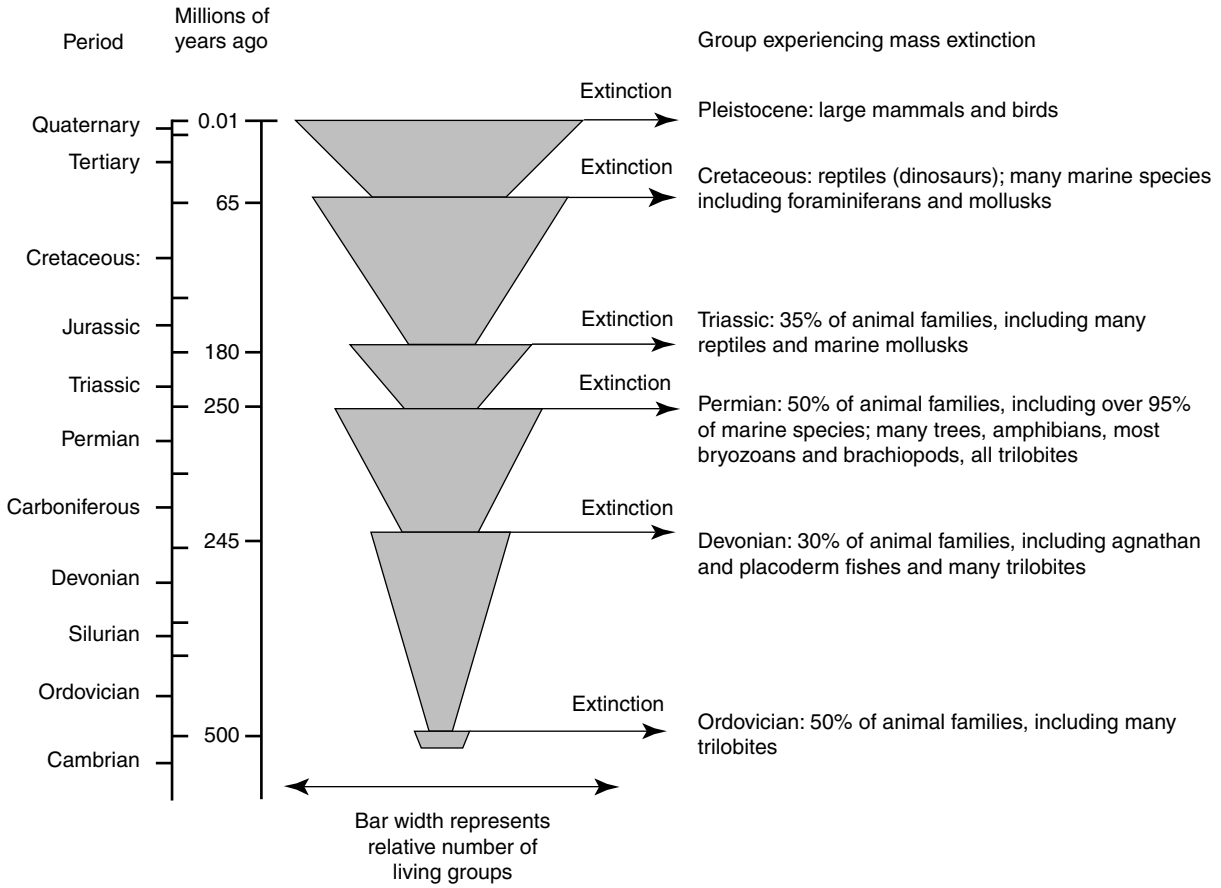


Figure 1 Although the total number of families and species has increased over the eons, during each of five episodes of natural mass extinction a large percentage of these groups disappeared. We are now at the start of a sixth episode, the Pleistocene extinction, as human populations eliminate species through habitat loss and overharvesting. Reproduced from Primack RB (1993) *Essentials of Conservation Biology*. Sunderland: Sinauer Associates.

several genera of mammals had evolved and also become extinct throughout the Pleistocene up to the Holocene, 10,000 years ago, but a primarily intact fauna remained. In Europe, only the largest mammals became extinct (>44 kg). A massive peak of extinctions occurred in Australia, with 86% of the large-bodied animal genera, including the large reptiles and giant birds, becoming extinct at the end of the Pleistocene. The majority of extinctions of the megafauna occurred about 30,000 years ago. Finally, in North America, mass extinction occurred between 10,000 and 12,000 years ago. Of the 37 kinds of mammals that became extinct, 33 weighed more than 44 kg. Similarly, extinctions in South America only affected the genera of animals exceeding 44 kg.

Two kinds of interpretations have been advanced to explain these extinctions, one focusing on human activities as driving extinction and the other focusing on climate changes. Thus, the episode of extinction in North America occurred at the end of the Pleistocene, when temperatures increased, leading to changes in the composition and the structure of plant communities, thus disturbing the interactions between plants and their browsers, and consequently leading to the extinction of many herbivores (particularly the large ones) as well as the predators that depended on them. Similar arguments may be used to explain the extinctions that occurred in Australia, where the climatic change at the end of the Pleistocene was associated with a higher frequency of drought periods. However, the asynchrony of the extinction crises in America and Australia may lead one to lean more toward the role of humans, since the colonization of these continents by humans was also asynchronous.

The well-documented case of Madagascar is particularly interesting. Humans arrived in Madagascar 1500–2000 years ago. At the beginning of the Christian era, 17 species of lemurids inhabited the island. Seven of them disappeared about 1000 years ago; all but one were large species, probably diurnal and mostly terrestrial. The surviving species were, in contrast, either nocturnal or small, arboreal species. Three species of mammals from other families disappeared at about the same time: a pigmy hippopotamus, *Hippopotamus lemeriei*, and two large carnivores, *Cryptoprocta spelea* and *Plesiorycteropus madagascariensis*; furthermore, two species of the giant tortoises, *Geochelone*, and the elephant bird *Aepyornis* (which is, at a weight of 500 kg, the largest known bird), became extinct. While it is not clear whether these extinctions are caused by the direct effects of hunting or by a degradation of the environment, it is certain, however, that they are the result of human activities.

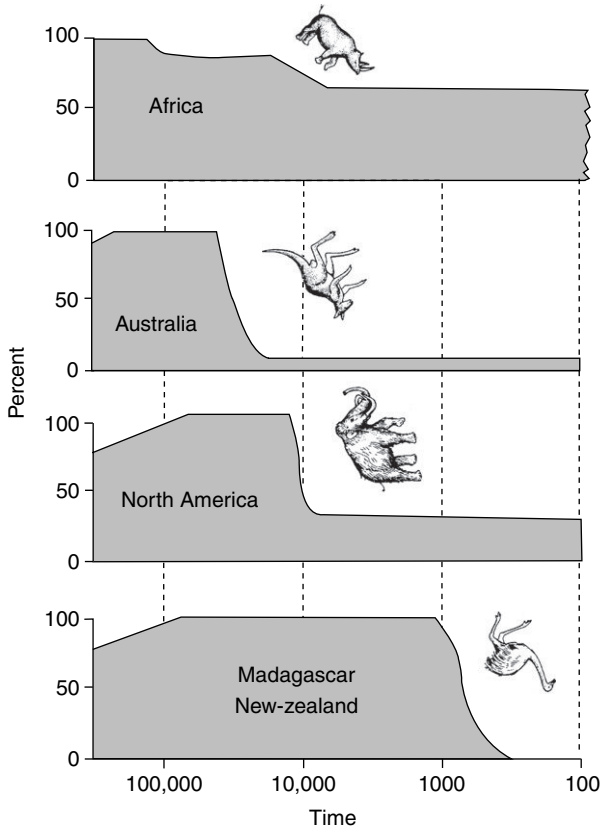


Figure 2 The extinction of large vertebrates coincided with the arrival of humans in Australia, North America, Madagascar, and New Zealand. In Africa, where humans and animals evolved together for millions of years, the damage was less severe. Reproduced from Wilson EO (1992) *The Diversity of Life*. London: Allen Lane, The Penguin Press.

Current Extinction Rates

In 2002, world leaders pledged, through the Convention on Biological Diversity ratified in 1992, to bring about a significant reduction in the rate of biodiversity loss by 2010. This so-called "International year of biodiversity" is gone and the rate of biodiversity loss does not appear to be decreasing (Butchart *et al.*, 2010). However, since the Earth Summit of Rio (June 1992), the scientific information about the current rates of extinction has been strengthened (Lawton and May, 1995; Millennium Ecosystem Assessment, 2005; Leadley *et al.*, 2010).

Recent and Current Extinction Rates

To measure the current erosion in biodiversity, biologists have estimated the current rates of extinction and then compared them with rates that would be expected in the absence of a crisis. The rate of extinction is the percentage of a given systematic group that disappears during a given interval of time. Based on paleontological archives, David Raup has estimated the average life span of animal species to be between 1 and 10 My, depending on the zoological group. Thus, an average life of 5 My per species is usually taken as the normal expectation.

To estimate the observed rates of extinction, biologists have calculated the percentage of all known species that have become extinct in the past 400 years (Smith *et al.*, 1993). However, these fractions largely underestimate the current rates of extinction, because the majority of species were identified only recently, during the nineteenth and twentieth centuries. Only vertebrates and plants were already well inventoried in the eighteenth century. Under these conditions, estimated rates of extinction for the period 1600–2000 are underevaluations of the number of species that have become extinct during the first three centuries of this period.

To avoid this pitfall, the percentage of known species that disappeared during the twentieth century is presented here (**Table 2**). These values represent good estimations for plants and vertebrates, but may still be underestimations for the other groups that were inventoried later because of the scarcity of specialists for the job. As we can see, the reliable values obtained for plants and vertebrate are very high, between 50 and 260 times higher than expected.

Thus, we are really facing a strong acceleration of extinction processes, justifying the title of Richard Leakey and Roger Lewin's book: *The Sixth Extinction* (Leakey and Lewin, 1995).

Estimate of the Future Rates of Extinction from the Area–Species Richness Relationship

Future rates of extinction can be estimated by merging the estimates on the losses of the environment (e.g., deforestation) with the expected relationship between area and species richness. This relationship is an empirical rule, based on a number of studies, that connects the number of species S for a given taxonomic group (coleopterans, birds, vascular plants, etc.) and the area A of the islands (true islands, or ecologically isolated areas such as lakes, mountain peaks, and clearings in a forest) inhabited by the species (**Figure 3**). Generally, the relationship between species richness and the area has the following form:

$$\log(S) = c\log(A)^z$$

where c is a constant and the parameter z is usually between 0.15 and 0.35. Assuming that this equation can also describe the reduction of species richness driven by deforestation or other processes decreasing the area of suitable habitat, and taking into account the current rate of deforestation in the tropics (which varies between 0.8% and 2% per year), one can deduce an annual rate of extinction of between 0.2% and 0.5%. In other words, with a total richness of currently 5 million species inhabiting the planet (our minimal estimate), between 10,000 and 25,000 species will disappear every year.

However, this kind of extrapolation is associated with several problems. It is not known, for example, up to what point the effects of the fragmentation of the environment, particularly in the tropics, can be deduced from the area–species relationship that has been described mainly for insular biogeography. Thus, Simberloff stressed that, during the past two centuries, only three species of birds are known to have become extinct in the forests in the eastern United States, even though the forest has been reduced to small fragments with a total of only 1–2% of their original area.

Table 2 Estimated rates of extinction in the twentieth century

	Number of species described	*Extinctions in the twentieth century*		*Extinction rate for the twentieth century*	*Endangered species*
		Expected	*Observed*	*%*	
Plants	270,000	5.4	270	0.1	30,000
Vertebrates	50,000	1	260	0.6	2300
Molluscs	70,000	1.4	(140)[a]	(0.2)	920
Insects	950,000	19	(50)	(0.005)	537
Crustaceans	40,000	0.8	(5)	(0.02)	407

[a]Likely underestimated.

Source: Reproduced from Teyssèdre *in* Barbault R and Chevassus-au-Louis B (2004) *Biodiversity and Global Change*. Paris: Ministère des affaires étrangères-adpf.

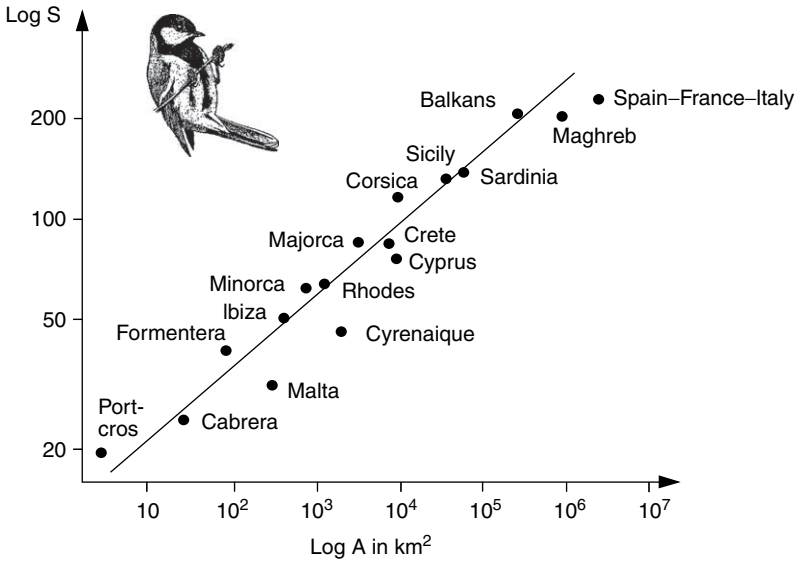


Figure 3 An example of the log–log relationship between the number of breeding-bird species (S) and the area of Islands and continental regions in the Mediterranean (A). Reproduced from Blondel (1995).

Despite such uncertainties regarding our estimates, it is certain that the reduction and the fragmentation of the environment, even if not always resulting in a reduced density of populations, will increase the risk of their becoming extinct. See **Box 1**.

Mechanisms of Extinction

What aspects of the environmental conditions make a species or population become extinct? It is clear that, if the mortality rate continuously exceeds the birth rate, a species will eventually become extinct. However, any prediction (which is based on perhaps a decade of observation) about extinction during the next, say, 3000 or 7000 years assumes that nothing will change. Since we are assuming that the major cause of extinction may be humans, this assumption appears unlikely.

Therefore, we will discuss more short-term risks of extinction, those associated with an environment in which, in principle, the population should persist (where births exceed deaths, at least at low population densities). In doing so, we will distinguish isolated populations (a situation that is rapidly developing with the human transformation of landscapes due to human activities) and interconnected populations constituting so-called metapopulations. The main difference between the two is that (unless due to human intervention) only interconnected populations can benefit from immigration.

Isolated Populations

Low Population Densities: The Allee Effect

At low densities, individuals may be distributed over distances that are much larger than those they usually move in, so that the probability of meeting one another for mating is small. Therefore, the growth rate of the population may decline below unity when the density is below a certain threshold. A well-established example of such an Allee effect leading to extinction is the mottled woodpecker *Dendrocopos medius* in Sweden in 1982. It is likely that an Allee effect is also responsible for the extinction of the migrating pigeon of North America, formerly so spectacularly abundant.

The possibility of extinction due to low population densities is particularly alarming for large animals, which live naturally at a low density, and for species with marked sociality.

Box 1 Future extinction rates deduced from IUCN Red Lists

The International Union for the Conservation of Nature (IUCN) proposed a number of criteria to classify the level of threats imposed on plant and animal species. Seven categories are defined:

Extinct: A taxon that has not been observed in nature for more than 50 years.

Endangered: A taxon close to extinction and whose survival is at risk if no action is taken to eliminate the causes of its disappearance. Included in this category are species whose populations are reduced to a critical level or whose habitats are threatened.

Vulnerable: A taxon at risk of entering the endangered class if detrimental factors continue to exert their effects. This category includes those taxons whose populations decrease as a consequence of overexploitation, large-scale habitat destruction, or any other perturbation of their environment, and those taxons whose populations remain abundant but that are nonetheless threatened by a variety of detrimental factors.

Rare: Taxa whose populations are globally rare even though they are not endangered or vulnerable. Such taxons are generally found in very specific areas or habitats or have widely dispersed small populations.

Indetermined: A taxon that has to be included in one of the above categories, but about which insufficient information is available to assess the level of risk.

Insufficiently known: A taxon thought to belong to one of the preceding categories, but without data to substantiate this.

Threatened: A taxon included in any of the preceding categories.

On this basis it appears (Vié *et al.*, 2009 and November 2009, IUCN synthesis) that from the census of greater than 47,677 species censused in the world, 875 were extinct (809 totally extinct and 66 extinct in the wild) and 17,291 were threatened. The estimates are particularly sound for vertebrates. In the case of the 5490 species of mammals for instance, 79 were extinct, 449 endangered, and 505 vulnerable. However, things are becoming worst particularly for amphibians (30% threatened over 6285 species) and freshwater fishes (37% over 3120 species studied).

Demographic Stochasticity

Changes in population density due to births and deaths necessarily imply the role of chance, as birth and death, but also the sex of the offspring and other parameters are random processes. It follows that there is a positive probability that all of the individuals in a population may die during the same time interval. This probability depends mainly on the population density – the ones with the lowest density being the most vulnerable – and on the mean and variance of the survival rate.

Environmental Stochasticity

In natural populations, the probability of having offspring or of dying is influenced by the environment, and these environmental effects have some degree of correlation among individuals. If the correlation is low, a failure, say, of one individual to produce offspring, may be canceled out, at the level of the population, by a large number of offspring produced by another individual. If, however, the correlation is high, the stochasticity will reduce the variance of the demographic parameters among individuals and therefore increase the risk of extinction. In the extreme, during a total catastrophe the correlation is complete, and all of the individuals die at the same time. Thus, high population densities are by no means a guarantee of population survival.

More formally, environmental stochasticity is evident in the temporal fluctuations of the growth rates of the populations. The increase in the variance of population growth increases the risk of extinction, in particular when the mean growth rate is low, the population density is low, or the carrying capacity of the environment is low.

Genetic Factors

Two widely discussed genetic factors associated with extinction, both of which are important in small populations, are inbreeding depression and the loss of evolutionary adaptability to new environments.

If a population goes through the bottleneck of a low population size, individuals are constrained to mate with close relatives. This is often associated with reduced viability or fertility, a phenomenon called inbreeding depression. The mechanisms leading to inbreeding depression are unknown for the majority of the species. In *Drosophila*, about half of the effect is due to the increased number of homozygous loci carrying recessive detrimental alleles (up to 5000); the other half is due to the accumulation of slightly deleterious and recessive mutations. Inbreeding depression has been established in many domestic species and zoo populations, but unfortunately, information is rare for natural populations.

The Critical Role of Population Size

The central message that emerges from most empirical and theoretical studies on extinction is that extinction risk increases, and thus the life expectancy of the population decreases, as population size decreases (**Figure 4**). Demographic stochasticity considerably increases the extinction risk if population size decreases below about 50 individuals; inbreeding depression is likely to become significant only under circumstances when a previously abundant population goes through a population bottleneck.

Metapopulations: Sets of Several Local Populations

Considering a metapopulation introduces an additional dimension to the risk of extinction, due to the possibility of dispersal from one local population to another. Seeds, spores, or animals may disperse passively or actively from their natal population over the landscape. If they arrive at a favorable site that is not already occupied by members of their species, they may be able to establish a new local population. In the long run, populations of a species may thus be established and disappear repeatedly at the local scale, while at the level of the landscape, one observes the dynamic equilibrium of a species moving among local habitats.

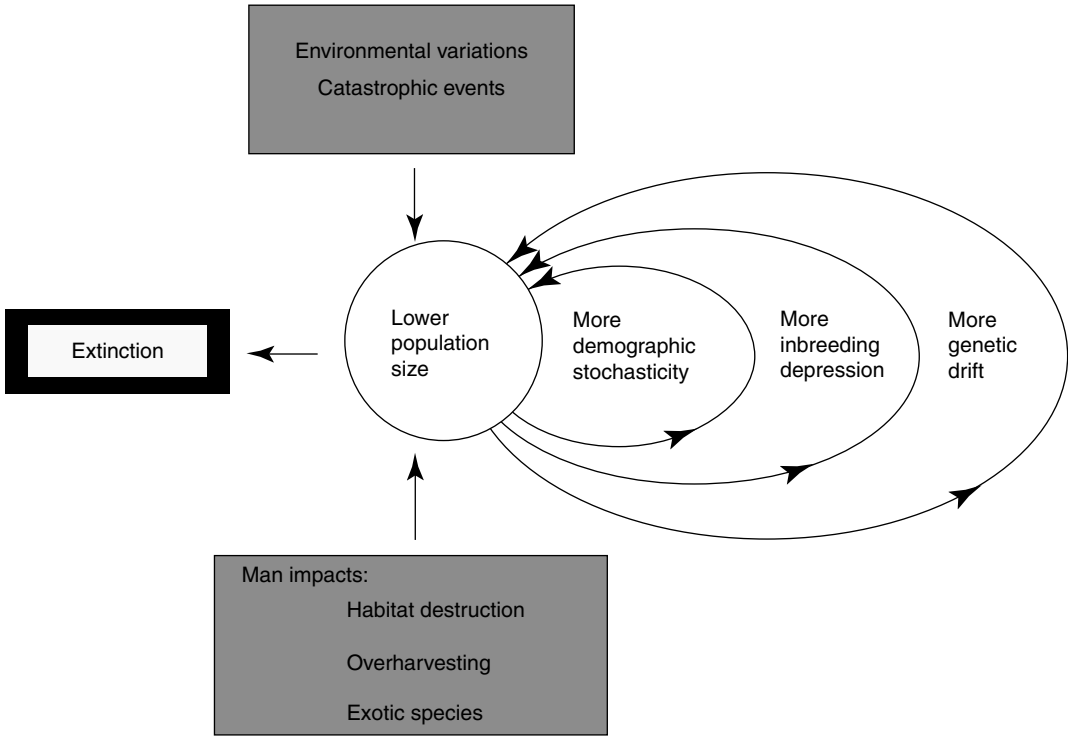


Figure 4 Extinction vortices: the smaller a population becomes, the more vulnerable it is to demographic, genetic, and environmental factors that tend to reduce population size even more and drive the population to extinction. Reproduced from Primack (1993) *Essentials of Conservation Biology*. Sunderland: Sinauer Associates.

The persistence of a metapopulation depends on processes that affect the rates of extinction of local populations, and also on processes that affect the establishment of new local populations. Two conditions must be fulfilled for the establishment of a new population: favorable sites must exist and a sufficient number of individuals must migrate from existing sites to the unoccupied habitats. This makes immigration and dispersal processes important mechanisms for the survival of metapopulations.

The survival of a metapopulation is threatened by two types of stochasticity, which are analogous to the demographic and environmental stochasticities involved in the dynamics of local populations: (1) the stochasticity of local colonization and extinction rates and (2) the regional stochasticity. The first involves the stochastic processes discussed earlier; the second, analogous to environmental stochasticity in isolated populations, applies if stochasticity has effects that are correlated among several local populations within the metapopulation. Since a significant source of environmental stochasticity is related to weather conditions, which are typically strongly correlated over a large spatial scale, one may generally expect a high degree of correlated regional stochasticity in metapopulations and thus high risks of extinction.

Causes of the Loss of Species and Genes

We have argued as if there were no change in the environment experienced by the populations, but there is, of course, hardly any doubt that the current increase in the rates of extinction is primarily caused by drastic changes in the environment. This affects all of the key parameters associated with the risk of extinction: population sizes, the carrying capacity of the environment, the mean and variance of growth rates, the genetic structure of a population, the number and area of suitable habitats, and the number of local habitats forming a metapopulation. From this perspective, we will now consider four types of environmental changes that may initiate the extinction process: the destruction of natural habitats, overexploitation by humans, changes concerning other species, and climate change.

Destruction and Changes of Natural Habitats

All animal and plant populations are adapted to the local conditions of their environment; their persistence depends on the maintenance of these conditions. Large climatic changes of the past, the periods of glaciation, for example, led to local extinctions and a shift of species' geographic ranges. These kinds of changes, however, were very slow, being spread out over periods of millions of years, so that most species were able to gradually adapt to or alter their geographic range according to the novel conditions. Humans introduced a novel dimension to the changes of natural habitats; human activities are more similar to volcanic eruptions than to gradual climatic change and thus could amplify its effects. At the time of the past glaciation, for example, the extinctions of the large mammals during the Pleistocene was partly because of human activities.

The main effects of humans on the environment are as follows:

1. Destruction of habitats (clearing forests, drying lakes, and humid areas, etc.).
2. Degradation of natural habitats (pollution, dams, etc.).
3. Fragmentation of habitats.

Destruction of Habitats

The most severe threat imposed on biological diversity is the loss of suitable habitat – and reducing habitat loss is the most important way of preserving the environment.

Loss of habitat has been the main source of extinction, at least in the recent extinctions of vertebrates. In many areas of the world, particularly on the islands and places where human populations are high, a large portion of the natural habitats has already been destroyed. Thus, in tropical countries, the destruction of the habitat ranges from 29% (Zambia) to 89% (Gambia) in Africa and from 41% (Malaysia) to 97% (Hong Kong) in Asia, and in most countries the number is above 70%.

Fragmentation of Habitats

The fragmentation of habitats is a process by which a large area consisting of a given type of environment is reduced in size and is divided into two or more fragments. These fragments of the original environment are often isolated from one another by a highly modified and degraded landscape. For example, the large stands of vegetation that used to cover vast areas of the world have increasingly divided into separated fragments. Europe as well as densely populated areas of Asia has experienced this process for centuries.

Such isolated fragments differ from the original environment in two critical ways:

1. The ratio of edge to area is increased.
2. The center of each fragment is closer to the edge.

Fragmentation of the habitat threatens the persistence of the species associated with it in various ways.

First, fragmentation may limit the capability of a species to migrate to, and thus to colonize, other habitats, because the landscape between the fragments creates an effective barrier to dispersal. In an undisturbed habitat, individuals can disperse over the entire landscape, which can lead to a stable metapopulation despite local extinctions (discussed earlier). If, however, the habitat is fragmented, the potential for dispersal and for colonization is often reduced. Thus, many birds or mammals, and many insects, are reluctant to leave the cover of their forest habitat and to cross small open spaces, as, if they do, they may be killed by predators prowling forest edges. Cultivated fields, measuring only 100 m in width, can constitute an insurmountable barrier to numerous invertebrate species. If the dispersal ability of, say, some mammals is reduced, plants might also suffer, as they may rely on these species for the dispersal of their seeds. Therefore, isolated fragments of habitat may develop that, although in principle may be suitable for colonization, cannot be reached by dispersing species; they will therefore remain uninhabited.

As some species disappear from a fragmented landscape as a result of natural processes in a metapopulation, while others do not succeed in recolonizing the habitat, the number of species will decline over time. It must be noted here that, worldwide, most natural parks and reserves are too small and too isolated to harbor a large number of species; almost half of the protected sites have an area of less than 100 km^2 and 98% have an area less than 10,000 km^2.

By splitting a large and contiguous population into subpopulations restrained to limited areas, the fragmentation of the environment can precipitate the decline of species and their extinction. As we have seen, small populations are indeed more vulnerable to inbreeding depression, genetic drift, and several other problems. While a large contiguous area may support a single large population, it may well be that none of the fragments can harbor a population large enough to ensure its long-time persistence.

Second, the increased proportion of habitat edges in fragmented environments brings with it changes in microclimate – light, temperature, humidity, wind – but also increased risk of fire and higher vulnerability to novel species of predators and competitors (e.g. invasion of weedy species from disturbed environments; see **Box 2**).

Persecution and Exploitation of Populations

The persecution and exploitation of natural populations generally threatens a certain number of species with extinction, in particular the large-bodied ones. These, and in particular the vertebrates, are also those that attract most attention from the general public, protection agencies, and the scientific community. This is reflected in the considerable amount of scientific literature focusing on rational and controlled exploitation of natural populations, be it hunting or fishing.

It is generally known that large animals require larger reserves than small ones, due to their low population densities. However, although it is generally true that large animals have small population densities, this may be due in many cases to the past persecution or exploitation of these species. It must be emphasized that large-bodied species may survive in disturbed habitats, if they are not persecuted. Thus, in Finland, large birds and mammals, including predatory species, have become more abundant recently in a range of manmade landscapes, most likely because they are now protected.

Changes in the Biotic Environment

A common cause of extinction is the interaction with exotic species that had either been introduced by humans or had arrived naturally (**Box 3**). The introduction of exotic species threatens in particular small isolated islands with specialized endemic species. Large islands, however, are not protected from this effect: After being introduced to Australia, the red fox was responsible for the extinction of several small marsupials.

The isolation of insular habitats favors the evolution of endemic species, but it also makes them particularly vulnerable to invading exotic species. Only a limited number of species can reach islands; plants, birds, and invertebrates are the most common colonizers. However, insular communities generally have few or no predators and browsers (due to the

Box 2 Catastrophic extinctions following deforestation

The looming mass extinction of biodiversity in the tropical forests is a major concern for the future (MEA, 2005).

It is reported here in two enlightening studies supported by broad-based empirical data.

An Amazonian Experiment

The Biological Dynamics of Forest Fragments Project (BDFFP) was initiated in 1979 as a large-scale experiment to assess the effects of fragmentation on Amazonian taxa. Standardized abundance data were collected for trees, birds, amphibians, mammals, and various invertebrate groups prior to experimental isolation of forest fragments of 1, 10, 100, and 1000 ha. These fragments were surveyed 22 years later (Laurance *et al.*, 2002).

The BDFFP results suggest that fragmentation alters population and community dynamics. As expected, they show that species richness is positively correlated with fragment size and that intact forest contains more species per unit area than fragments. Many large mammals, primates, and understory birds, but also some termite, ant, and butterfly species have disappeared from even the largest (100–1000 ha) fragments in the study area. Once isolated, small (1–10 ha) fragments initially lose species at a high rate. For example, dung and carrion beetle assemblages were markedly altered only 2–6 years after isolation. Local extinctions of birds, primates, and butterflies have also occurred more rapidly in small than in large fragments. In contrast, a few taxa like rainforest frogs have remained stable after fragment isolation. Moreover, the BDFFP researchers revealed the remarkable diversity of edge effects in fragmented forest and the importance of matrix effects.

Extinctions in Singapore

Since the arrival of the British in Singapore (a 618 km^2 island) in 1819, some 95% of the 540 km^2 forest was logged. From a census of 3996 species (and those known to have been present around 1819) by Brook *et al.* (2003), 881 were locally extinct 183 years later. Substantial rates of extinctions were found, especially for forest specialists (33%), with the greatest proportion of extinct taxa (34–87%) in butterflies, fish, birds, and mammals. Observed extinctions were fewer for species that prefer or tolerate open and forest-edge habitats (7%) like most amphibians and reptiles.

In conclusion, both experiments show (1) the extent and quickness of extinctions following deforestation; (2) the importance of matrix quality and edge effects for species survival; and (3) the diversity of responses exhibited by different taxa.

Box 3 Examples of extinction cascades due to introduced species

What happened to Guam's avifauna is a good example of the destructive effects of introduced species. Guam is an island belonging to the Marianes Archipelago, situated between Japan and New Guinea. Eighteen species of indigenous birds were known to inhabit the island, as well as seven introduced ones. Then, during the past 20 years, populations declined spectacularly: Today, seven species are considered to be extinct and four others have become so rare that their survival is endangered. Nevertheless, on other islands of the same archipelago, no comparable decline has been observed. Strikingly, the 10 forest species were all affected and in a similar manner: The birds disappeared first from the southerly forests during the 1960s, and then their decline spread progressively toward the north of the island. This decline coincided exactly with the introduction and subsequent spatial expansion of a tree snake, *Boiga irregularis*. Since this snake is absent on neighboring islands, there is good evidence that it was the cause of the observed declines and extinctions. Because of its arboreal and nocturnal habits, this snake is a voracious predator of perching birds and birds sitting on their nests, as well as of eggs and nestlings. Moreover, because it also predates on small mammals and lizards (the latter particularly abundant), it can attain high densities even while exterminating its most vulnerable prey species.

On Santa Catalina, an island close to California, 48 native plant species have been eliminated mainly as a consequence of overgrazing by goats and other introduced mammals.

At Madagascar, where the ichthyofauna is highly endemic, with 14 out of 25 genera unknown elsewhere, a recent inventory of freshwater environments was not able to retrace more than five. Introduced fishes dominate all aquatic environments. The combination of habitat degradation and the introduction of exotic fish seem to be leading the original ichthyofauna toward extinction.

The same has been found in continental aquatic environments, which, for aquatic species, are a kind of island surrounded by inhospitable terrestrial space. Originally, Lake Victoria had more than 350 endemic fish species. Today, many are rare or have become extinct after the introduction of the Nile perch, *Lates nilotica*, in 1960. In fact, the phenomenon is more complicated than that: In 1978, the perch did not yet represent more than 2% of the annual catch; in 1986, it represented 80%. Factors other than predation have played a role (algal blooms creating anaerobic conditions as a consequence of pollution by fertilizers and other pollutants).

difficulties of colonizing the island or of establishing a population on an insufficient area), and indeed, the species representing the highest trophic levels (e.g., the carnivorous mammals) may be missing completely. Since many endemic species on islands have thus evolved in the absence of the selective pressures of predators and mammalian browsers, they have not evolved (or have lost) any means of defense: Birds have lost the ability to fly and lay their nests on open ground and plants do not produce any chemical substances and do not have any protective tissues (spines) that could deter browsers.

Therefore, endemic species that can spread in the absence of these selective pressures may become extinct rapidly when the pressures appear: animals introduced into the islands eliminate them by predation or overgrazing. The introduced plants, however, equipped with a protected or a toxic foliage, have previously evolved to withstand browsing, so that they outcompete the endemic plants, enhancing the selective pressure by the introduced browsers and thus accelerating the extinction of the endemic species.

Moreover, insular species usually do not have an immune defense against foreign diseases; once introduced with the

invading animals, these can spread epidemically and can destroy the native populations. It is thought, for example, that the almost complete destruction of the Hawaiian avifauna – since 1850, 85% of the endemic species have become extinct or have been reduced to very small populations – is due to the spread of variola and bird malaria, introduced together with the mosquito *Culex quinquefasciatus* in 1826.

Extinctions caused by human activities in insular communities are much more frequent than generally recognized. Thus, the rate of extinction on the order of 1% given earlier is a gross underestimate of the true value. Current estimates indicate that 25% of endemic bird species have disappeared from islands as a result of human activities. There are, for example, only three species of winged rails today, while thousands have existed during the past 2000 years. We are justifiably worried by the current wave of mass extinctions threatening the tropical forests, but should also recognize that the oceanic islands have already undergone such mass extinctions.

Climate Change

Climate change over the past decades has produced shifts in the distributions and abundances of many species (Lovejoy and Hannah, 2005; Möller *et al.*, 2004). In some cases, it has been implicated in species extinctions (**Box 4**). Thus, using projections of species distributions for various climate scenarios, extinction risks for sample regions covering 20% of the Earth's terrestrial surface have been estimated by an international team led by Thomas *et al.* (2004). Exploring various approaches in which the estimated probability of extinction shows a power-law relationship with geographical range size, they predict, on the basis of midrange climate-warming scenarios (1.8–2.0 °C, which looks very optimistic!) for 2050, that 15–37% of species in their sample of regions (Europe, Mexico, South Africa, and Queensland) and taxa (terrestrial vertebrates, butterflies and some other invertebrates, and plants) will be committed to extinction. The results of this kind of scenario depend on various hypotheses, such as the different levels of dispersal abilities of the species (Thomas *et al.*, 2004).

In the case of marine species, the impact of climate change is not limited to the effects of warming: there is also a direct effect of the increase of CO_2 in the atmosphere through ocean acidification. It appears as a pervasive stressor that could affect marine organisms and cause profound ecological shifts that may facilitate species extinction (Krocker *et al.*, 2010). Calcifying organisms such as crustaceans or corals are particularly sensitive to such acidification (see **Box 5**).

Ecological Reflection on the Primary Causes of the Current Crisis of Extinction

All organisms modify their environment, and humans are no exception. Thus, in a strategic document, jointly published by the World Resources Institute (WRI), the World Conservation Union (IUCN), and the United Nations Environmental Programme (UNEP), six fundamental causes for the impoverishment of biodiversity are recognized:

1. The rapid and unsustainable growth of human populations and the consumption of natural resources.
2. The continued reduction of the range of products in agriculture, forestry, and fishing.
3. The economic and political systems that do not take into account the environment and its resources.

Box 4 A mass extinction associated with pathogen outbreaks tied to global warming

As the Earth is becoming warmer, many species are likely to disappear, often because of changing disease dynamics (Pounds *et al.*, 2006). For instance, it appears that a recent mass extinction in amphibians is associated with pathogen outbreaks tied to global warming. During the 1990s, in the mountains of Costa Rica, the Monteverde harlequin frog (*Atelopus* sp.) disappeared along with the golden toad (*Bufo periglenes*). An estimated 67% of the 110 or so species of *Atelopus*, which are endemic to the American tropics, have met the same fate, and a pathogenic chytrid waterborne fungus, *Batrachochytrium dendrobatidis* is implicated.

Chytriomycosis is an infectious disease that affects amphibians worldwide. It is capable of causing sporadic deaths in some species and 100% mortality in others. For instance, mass die-offs in populations of the midwife toad *Alytes obstetricans* at high altitude in the western Pyrenees in both France and Spain were observed as early as 2002 and are spreading (http://www.bd-maps.eu; http://www.sosanfibios.org).

Box 5 Marine species face an elevated risk of extinction from global changes

Marine ecosystems are centrally important to the ecology of the planet: covering 71% of the Earth's surface, the ocean nurtured life on our planet and continues to play a dominant role in regulating the climate. Recent studies indicate that rapidly increasing greenhouse gas concentrations are causing a decrease in ocean productivity, altered food web dynamics, shifting species distributions, and elevated risk of extinction in many species (Hoegh-Guldberg and Bruno, 2010). Let us focus here on the case of coral species.

Coral reefs harbor the highest concentration of marine biodiversity and are facing increasing threats at local and global scales. In particular, rapid buildup of carbon dioxide in the atmosphere is leading to both rising sea surface temperatures and water acidification. The conservation status of 704 reef-building coral species was assessed using Red List criteria and measures of population reduction over time. One-third of these are in categories with an elevated risk of extinction. Declines in abundance are associated with bleaching and diseases driven by elevated sea surface temperatures, with the risk of extinction further exacerbated by local-scale anthropogenic disturbances.

A recent investigation of 328 colonies of massive *Porites* corals from 69 reefs of the Great Barrier Reef in Australia shows that calcification has declined by 14.2% since 1990. The data suggest that such a severe and sudden decline in calcification is unprecedented in at least the past 400 years. It could be a response to the declining pH of the upper seawater layers due to the absorption of increasing atmospheric CO_2 (De'Ath *et al.*, 2009).

Table 3 Recent rates of conversion of natural habitat to specialized agriculture (the 10-year rate, up to 1987)

Conversions to cropland (%)		*Conversions to pastureland (%)*	
Paraguay	71.2	Ecuador	61.5
Niger	32.0	Costa Rica	34.1
Mongolia	31.9	Thailand	32.1
Brazil	22.7	Philippines	26.2
Ivory Coast	22.4	Paraguay	26.0
Uganda	21.4	Viet Nam	14.0
Thailand	17.1	Nicaragua	11.8

World Resources Institute and International Institute for Environment and Development.

4. The inequalities in the possession, management, and sharing of the advantages related to the use and the conservation of biological resources.
5. The legislative and institutional systems that favor unsustainable exploitation of resources.
6. The lack of knowledge and its applications.

In fact, the decline in diversity decline results from specific choices during the course of human development (Swanson, 1997). It is clear that, with population growth, demographic changes, and the associated technological development, humans are using an increasing share of the planetary resources. Indeed, it consumes, diverts, or holds approximately 39% of plant productivity, the fundamental source of energy for most living systems (Vitousek *et al.*, 1986). The rate of conversion from natural environments into agricultural systems, high in the developing countries (**Table 3**), remains a major threat for biodiversity. This is alarming not only for the future of many animal and plant species but also for humans.

In addition to the dangerous reduction of renewable resources from the environment, population growth leads to increases in pollution (including the gases responsible for the greenhouse effect), which threaten the balance of ecosystems and the climate. Novel models of development are necessary so that population growth does not exceed the carrying capacity of the planet.

The five other primary causes blame either our economic behaviors or the legal uses of the human societies. For thousands of years, the humans formed a mosaic of relatively autonomous areas. The state of knowledge, strategies of subsistence, and social structures evolved more or less independently in each area. What the populations required of the environment seldom exceeded the capacity of nature.

With the globalization of the economy, starting at the end of the past century, uniformity and interdependence have increased. In agriculture, for example, a small number of cultivated plants have come to dominate the global economy (**Box 5**). In fact, today, human societies mainly depend on four species of plants (maize, wheat, potato, and rice) for their basic needs. Furthermore, these species are used in the form of a decreasing number of high-output varieties. Thus, in Sri Lanka, the 2000 varieties of rice used in 1959 has decreased to 5 main varieties today; in India, 10 varieties out of originally 30,000 represent 75% of production; and about 62% of the varieties in Bangladesh and 74% in Indonesia are estimated to be derived from a single maternal plant (**Box 5**).

The process of conversion, which consists of investing capital in a type of resource that is economically advantageous, is the principal economical force threatening biodiversity: Development and conversion go hand in hand, and conversion is the process by which habitat and its resident species are lost (Swanson, 1997).

The reduction in the number of cultivated species is accompanied by the disappearance of nitrogen-fixing bacteria, mycorrhiza, predators, pollinators, and many other species that have coevolved during centuries with the traditional systems of agriculture. The use of manure, pesticides, and high-output varieties accelerated this erosion.

Similarly, with regard to marine ecosystems, the vast world markets favored the development of blind fishing with, for example, gigantic drifting nets that trap not only enormous quantities of target species, but also a considerable number of other species of fish, mammals, and birds.

In addition, several factors led to an underestimate of the use of environmental resources. First, many resources are directly consumed without appearing on the markets. Second, the benefits of biodiversity are largely public goods. For example, while the protection of wetlands may benefit a population, the benefit is so diffuse that traditional economic theory cannot incorporate mechanisms that would preserve them. Correctly evaluated, the natural systems and their biodiversity are major economic assets. However, as these systems are often underestimated, economic theory considers the conservation of the biodiversity an expenditure rather than an investment.

In response to this challenge, a new discipline is being developed that integrates economics, ecology, and public policy: ecological economics. As Timothy Swanson emphasized, diversity decline is a specific form of institutional failure: the failure to create institutions that internalize the values of biodiversity within the decision-making of states and individuals making conversion decisions.

The reduction in the number of the species and the destruction of habitats are the standard in many countries where a minority of the population controls most of the territory. The rapid profits gained from excessive logging or fishing go to a minority, while the local population, which depends on the sustainable production of the resources, pays the price. In fact, the ownership is often more likely to be granted to those who cut down and colonize forests and other areas covered with natural vegetation than to the inhabitants of the forests living on the sustainable harvest of the natural products. Any uncertainty in the ownership dissuades good management practices and encourages overexploitation.

A second problem stems from the concentration of the control of the resources and the responsibilities for decisions

about environmental policy in the hands of citizens. However, in many societies, it is the women who manage the environment and understand the value of the biodiversity for agriculture and health.

A third problem is that of international trade, the debt, and the policies of technology transfer that support the inequalities and often reinforce the imbalances observed within nations. To preserve biodiversity, the industrialized countries must reverse this flow. If the developing countries continue to be excluded from the markets, to be deprived of access to technologies, and to be blocked by their debts, they will have neither the means nor the incentives necessary to preserve their biological resources.

See also: Extinction, Causes of. Human Impact on Biodiversity, Overview. Mass Extinctions, Notable Examples of. Metapopulations

References

Brook BW, Sodhi NS, and Ng KL (2003) Catastrophic extinctions follow deforestation in Singapore. *Nature* 424: 420–423.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328: 1164–1168.

De'Ath G, Lough JM, and Fabricius KE (2009) Declining coral calcification on the Great Barrier Reef. *Science* 323: 116–119.

Hoegh-Guldberg O and Bruno F (2010) The impact of climate change on the world's marine ecosystems. *Science* 328: 1523–1528.

Krocker KJ, Kordas RL, Crim RN, and Singh GG (2010) Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. *Ecology Letters* 13: 1419–1434.

Laurance WF, Lovejoy TE, Vasconcelos HL, *et al.* (2002) Ecosystems decay of Amazonian forest fragments: A twenty-two year investigation. *Conservation Biology* 16: 605–618.

Lawton JH and May RM (1995) *Extinction Rates*. Oxford: Oxford University Press.

Leadley P, Pereira HM, Alkemade R, *et al.* (2010) *Biodiversity Scenarios: Projections of Twenty-First Century Change in Biodiversity and Associated Ecosystem Services. Technical Series No. 50*. Montreal: Secretariat of the Convention on Biological Diversity.

Leakey R and Lewin R (1995) *The Sixth Extinction. Patterns of Life and the Future of Humankind*. New York: Doubleday Dell.

Lovejoy TE and Hannah L (2005) *Climate Change and Biodiversity*. New Haven/London: Yale University Press.

Millenium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being*. Washington, DC: Island Press.

Möller AP, Fiedler W, and Berthold P (2004) Birds and climate change. *Advances in Ecological Research*, vol. 35. Amsterdam: Elsevier Academic Press.

Pounds JA, Bustamante MP, Coloma LA, *et al.* (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439: 161–167.

Primack RB (1993) *Essentials of Conservation Biology*. Sunderland: Sinauer Associates.

Smith FDM, May RM, Pellew R, Johnson TH, and Walter KS (1993) Estimating extinction rate? *Nature* 364: 494–496.

Swanson T (1997) *Global Action for Biodiversity*. London: Earthscan Publications.

Thomas CD, Cameron A, Green RE, *et al.* (2004) Extinction risk from climate change. *Nature* 427: 145–148.

Vié JC, Hilton-Taylor C, and Stuart SN (2009) *Wildlife in a Changing World: An Analysis of the 2008 IUCN Red List of Threatened Species*. Gland, Switzerland: IUCN.

Vitousek PM, Ehrlich AH, and Matson PA (1986) Human appropriation of the products of photosynthesis. *Bioscience* 36: 368–373.

Macroscopic Patterns in Marine Plankton

Angel López-Urrutia, Instituto Español de Oceanografía, Gijón, Spain

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by William K.W.Li, pp 1–16, © 2007, Elsevier Inc.

Glossary

Comparative analysis Pattern-seeking across natural ecosystems based on the premise that existing gradients and variations generate regularities which can be discerned by statistical analysis of the composite aggregate, of component variables, or of mechanistic connections.

Complex adaptive system A system of multiple interconnected diverse elements whose localized interactions are subject to autonomous selection, leading to continual change and hierarchical structure.

Ecological stoichiometry The quantitative relationship between chemical elements (especially carbon, nitrogen, and phosphorus) arising from ecological processes and interactions such as nutrient acquisition, biomass formation, remineralization, trophic transfer, sedimentation, and dilution.

Macroscopic pattern Regularity observed at a scale higher than that of the interacting units, being the statistical or emergent expression of the ensemble arising from the components.

Multiscale analysis A change in the resolution or range of a measurement to evince a change in pattern.

Plankton functional types A classification of plankton according to commonly shared adaptive features such as physiological and morphological traits which lead in combination to predictable ecosystem or biogeochemical functions.

Plankton size classes A classification of plankton according to size, based on the International System of Units prefixes of the approximate live weight in grams for successive logarithmic intervals of organism width or length: femtoplankton (0.02–0.2 μm), picoplankton (0.2–2 μm), nanoplankton (2–20 μm), microplankton (20–200 μm), mesoplankton (0.2–20 mm), macroplankton (2–20 cm), and megaplankton (20–200 cm).

Plankton trophic types A classification of plankton according to whether they obtain energy from light ("photo") or chemicals ("chemo"), whether they obtain electrons from water ("litho") or organic carbon ("organo"), and whether they obtain carbon from inorganic ("auto") or organic ("hetero") compounds, with the possibility of mixed modes ("mixo"): for example, photolithoautotrophy, chemoorganoheterotrophy, photoorganoautotrophy, photoorganoheterotrophy, and mixotrophy.

Power law A relationship that scales a dependent variable (Y) to an independent variable (X) through a normalization factor (a) and a power exponent (b) in the form $Y = aX^b$.

Scale The distance or time over which a significant change in some quantity of interest can be resolved.

Microscopic Organisms and Macroscopic Patterns

Marine Plankton

Marine plankton are pelagic organisms, motile or nonmotile, that are entrained by the prevailing movement of water. These biological entities range over many orders of magnitude in size from nanometer-scale viruses of the femtoplankton to meter scale gelatinous animals of the megaplankton. By numbers, biomass, and diversity, it is the microscopic forms of plankton that pervade the oceans. The majority of these are prokaryotes and unicellular eukaryotes. Global estimates of prokaryotic marine plankton indicate a stock of 10^{29} cells, a biomass of over 2 Pg carbon, and a genetic richness of perhaps up to 2 million different kinds of taxonomic units. These bacteria and archaea, together with the smaller pool of eukaryotic algae and protists, constitute the foundation of pelagic food webs. In oxygenated waters, most microbial plankton are photolithoautotrophs, chemoorganoheterotrophs, or mixotrophs, but significant number of them also possess capabilities for photoorganoautotrophy or photoorganoheterotrophy. Thus, not only is there great diversity in size and form amongst the microbes, but also there is a wide range of metabolic functions. All other plankton, being metazoan, are chemoorganotrophs and depend ultimately on the microbes. Individually, plankton live short lives and

interact over short distances. However, they have collectively established the conditions for life over the entire Earth since the beginning of evolutionary time, and continue to maintain biospheric integrity through complex adaptations.

Patterns and Processes

In ecology, a pattern can be described as regularities in what we observe in nature; that is, they are widely observable tendencies (Lawton, 1999). Importantly, patterns can exist at various scales in time and space. A macroscopic pattern is one in which the regularity is observed at a scale higher than that of the interacting units. At each scale, the number of biological units is larger than that at the next higher scale. The expression of the ensemble in terms of the component units results in the pattern, which is called macroscopic. The study of macroscopic patterns in ecology is macroecology, and it can be stated as the enterprise of trying to infer laws of nature from the statistical manifestations of the many interacting biotic units of ecological systems (Brown, 1995).

A bulk effect, which is a statistical resultant of aggregating system components, is no less a macroscopic character than an epiphenomenon that emerges from underlying complexity, or one that appears out of constraints imposed from even larger scales. Thus, macroscopic patterns can arise from large number systems in which many independent, essentially identical components interact randomly to give system averages, which are stated in the formalism of statistical mechanics. Macroscopic patterns can also arise from systems in which intermediate numbers of components interact in structured and complex interrelationships. Plankton might be viewed as complex adaptive systems because the essential elements of such systems are in place: natural selection acting on individual components with diversity sustained through localized interactions yielding a subset for replication or enhancement (Levin, 1998). Indeed, macroecology might be considered a perspective on ecological complexity, whether one believes populations and communities to be complex adaptive systems, or simply to be the convenient assemblages of organisms selected for study (Brown, 1995).

In a strict sense, a macroscopic pattern refers only to that which is observed. It is a separate and much more difficult matter to understand the determinant, which is the process. Pattern seeking is an exercise in empirical observation, and the enterprise of large numbers often requires observation over extensive geographic areas and over prolonged durations. Biogeography and macroevolution are therefore sometimes implicated in macroecology, but macroscopic patterns need not necessarily inform about these disciplines. Understanding process through deduction is a part of the macroecological agenda; but it necessarily follows inductive insight gained from pattern detection.

Comparative Ecology and Macroecology

Comparative Analyses of Plankton Ecosystems

The study of macroscopic patterns is a well-established part of biological oceanography, even though the neologistic origin of macroecology is strongly based in terrestrial and avian ecology. At a symposium held in 1956, Alfred Redfield addressed the imbalance of experiment and observation in marine biology and suggested that "to understand the distribution and abundance of life in the sea, the approach must be primarily statistical through the development of significant relationships between large quantities of observations on biological and physical events, occurring often in widely scattered places" (Redfield, 1960). Essentially, this is a prescription for comparative analyses of ecosystems, an approach that has since been extensively used in the study of plankton.

Some of the most common comparative analyses in aquatic ecology are bivariate relationships between plankton trophic groups (Gasol and Duarte, 2000). Plankton are adaptive and complex, so it becomes a simplification when single variables are chosen as surrogates for entire trophic groups. In principle, this simplification is not necessary for comparative analyses, but the demands for large number analyses require measurements that are operationally well defined and easy. For example, the entire photoautotrophic pelagos is often represented by the concentration of chlorophyll *a*. In many cases, this photopigment is not only an adequate and suitable surrogate for the diversity of unicellular algae and cyanobacteria, but is also a useful predictor variable amenable to widespread application. In other cases, however, the surrogate confounds process and obscures patterns. Generality at a high level of description comes at the expense of a loss in detail at a lower level.

Macroecology in Biological Oceanography

The spatial and temporal structures in plankton ecosystems are a fertile area for macroecological study because some aspects are accessible to easy measurement at many scales. For phytoplankton, if we do not restrict the definition of structure only to ecosystem rate parameters, but also admit state variables, then either the surrogate of chlorophyll *a* (and its many proxies such as fluorescence and ocean color) or direct counts of individual cells can provide useful information across many scales. In this way, horizontal distributions of phytoplankton can be studied from turbulent regimes of local environments (Platt, 1972) to biogeochemical regimes of the global ocean (Platt and Sathyendranath, 1999). The premise that phytoplankton distributions are under the broad control of abiotic forcing at regional and global scales allows the pelagos to be viewed in an ecological geography rather than in a traditional biogeography (Longhurst, 1998). In other words, areas of the ocean with common abiotic forcing might be expected to have similar ecologies (Platt *et al.*, 2005). Testing this premise with other plankton is more difficult because of limited data, but some progress can be seen in studies of heterotrophic bacterioplankton, which show marked contrasts at the largest scale of ecological domains (Ducklow, 2003). In the ocean, variations in physical forcing tend to increase with spatial and temporal scales. Red noise, so called, makes a dominant difference in the variance structure between marine and terrestrial systems, and is associated with stronger adaptive responses to long-term change in the ocean (Steele, 1985). Spectral analysis, as a statistical mechanical representation of ecological systems, has deep roots in phytoplankton studies

(Platt and Denman, 1975). Macroecology, it seems, has always been a part of biological oceanography.

Scales of Variability

Multiscale Analysis

The linkage of spatial and temporal variability is a long-standing concept in oceanography. Graphical representations of the variability in a measured quantity in the space–time domain are familiar to physical oceanographers, for example, the Stommel diagram, and to biological oceanographers, for example, the Haury–McGowan–Wiebe diagram (Haury *et al.*, 1978). In logarithmic fashion, time ranges from minutes to hundreds of thousands of years on one axis, space ranges from centimeters to tens of thousands of kilometers on a second axis, and biological variability is indicated on a third axis. These diagrams portray the overlapping hierarchy of patterns along the diagonal, stretching from small space–time scales to large space–time scales. They also convey the reality that what can be perceived is limited by the narrow space–time window through which observations are made. The search for macroscopic patterns is an exercise of increasing the scale of analysis.

In multiscale analysis, a systematic change is made to the scope of a quantity. Scope, which is the dimensionless ratio of extent to grain, expresses space or time as the number of measured units adding up to the length of the investigation. For example, if plankton is collected from a site once every week (the grain) for a total duration of 14 years (the extent), the scope is 728 ($52 \times 14/1$). Various scaling methods exist to change the scope (Schneider, 2001).

Coarse Graining

The method of coarse graining takes the full extent of the weekly measurements and averages them within increasingly larger grain sizes, such as a month, a season, a year, a quadrennial, and so on. An example is shown for time series of bacterioplankton and phytoplankton sampled weekly for 14 and 13 years, respectively (**Figure 1**). Averaging of raw measurements into increasingly larger grains reduces variance. In the example, at the largest grain of a year, all the weekly heterogeneity becomes unresolved, and the between-year heterogeneity decreases (**Figure 1(d)**). It then becomes meaningful to seek a long-term (macroscopic) signal from the annual averages unencumbered by the weekly noise. For the phytoplankton in this example, cell abundance has increased at a rate of 4% per year from 1993 to 2005. As the time series lengthens into the future, it should not be a surprise if annual variance were to increase since a larger extent encompasses more realized situations. Emphatically, this multiscaling analysis demonstrates the difficulty of empirically detecting a

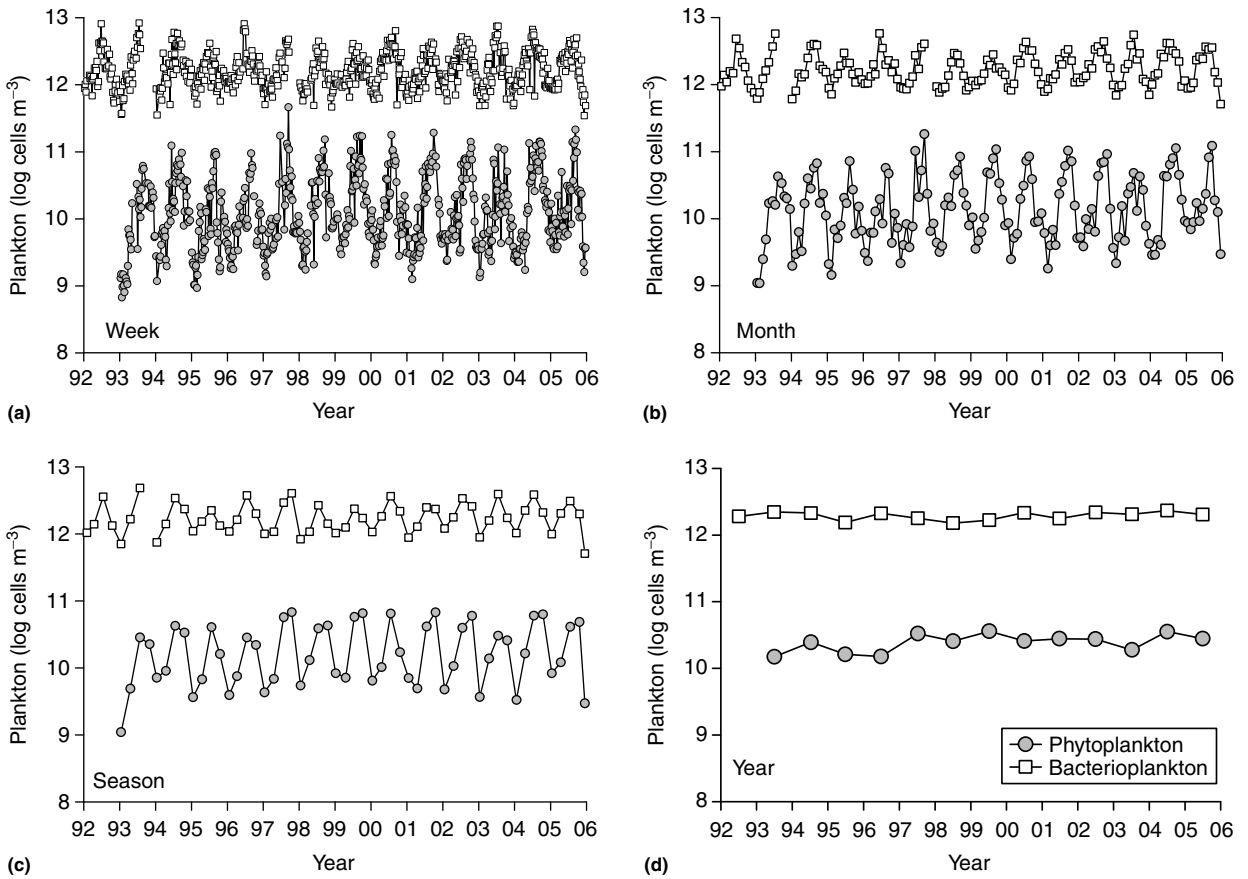


Figure 1 Multiscale analysis by coarse graining. Time series of phytoplankton and heterotrophic bacterioplankton abundance in Bedford Basin, Canada from 1992 to 2005. (a) Weekly measurements, (b) monthly averages, (c) seasonal averages, and (d) annual averages.

multiyear trend if either the grain size of raw observation is not small enough, or if the extent is not large enough.

Accumulating

Another method of changing the scale of analysis is to accumulate an increasing collection of samples. By always using the same operational procedure for collection, each sample may be considered an element in the set of all possible configurations. The more samples accrued, the higher the degree of aggregation. A herculean human endeavor of counting and identifying all 195,983 microphytoplankton cells in 1388 Mediterranean Sea samples, and all 1,111,859 such cells in 1144 Caribbean Sea samples by Margalef (1994) and his colleagues is a majestic exemplar. From these ensembles emerge an elegantly succinct pattern: $S(n) \propto n^{-1}$, which is to say that there is a simple power-law relationship between the number of individuals, n, and the number of species, S, represented by that particular number of individuals (Solé *et al.*, 2002). In other words, only a few species have large populations of individuals, but many species have small populations.

With electronic technology, it is possible to count picophytoplankton and nanophytoplankton with considerable ease. The detection of these cells is based on biooptical characteristics such as light scatter and photopigment autofluorescence. Bivariate diagrams plot each cell as a single dot on an $m \times m$ grid (where m is usually 256 or 1024) representing the position of the cell on the coordinate space of light scatter and pigment fluorescence. A subsequent cell that occupies the same coordinates makes the display dot darker in appearance. The result of analyzing a single water sample is a cytometric diagram showing the density of occupancy of the coordinate space. One such realization taken on a single day in 1996 detected several thousand cells (**Figure 2(a)**). Samples measured week after week (**Figure 1**) accrued over an entire year greatly increase the number of cells. A larger portion of the coordinate space is occupied by the yearly accumulation, and the cell abundance at occupied positions increases. At this annual level, there is an accrual of a few hundred thousand cells (**Figure 2(b)**). The procedure can be carried further to show accumulation of greater than a million cells at the quadrennial scale (**Figure 2(c)**), and finally to the extent of the entire 676-week time series (**Figure 2(d)**) during which 5 million cells have been counted. By pooling all fluorescence measurements at a given light scatter value into a single bin, the cytometric data can be displayed as univariate frequency histograms resembling cell-size spectra. Increasing the scale by accumulation transforms the heterogeneous single-week spectra (**Figure 3(a)**) to more homogeneous annual (**Figure 3(b)**) and quadrennial spectra (**Figure 3(c)**), and finally to the unitary 13-year spectrum (**Figure 3(d)**). As with coarse graining, accrual smooths out variance.

Macroscopic Assembly

Trophic Classification

One of the most enduring constructs in ecology is the trophic pyramid. Organisms are assigned to one, and only one, of a small number of feeding categories, which are then arranged in food chain sequence, showing the importance of each trophic level in relation to those below and above it. Traditionally, the categories are primary producers (autotrophs), decomposers (bacteria), primary consumers (herbivores), secondary consumers (carnivores), and tertiary consumers (apex predator). Importance can be indicated by the number of individuals, by total biomass, or by energy flux. Trophic diagrams of energy are constrained by thermodynamics to have an upward pyramidal shape, but diagrams of numbers and biomass are free to assume other appearances, even inverted pyramid. Plankton biomass pyramids built of autotrophs (phytoplankton) and heterotrophs (bacterioplankton, microzooplankton, and mesozooplankton) have a tendency to be inverted in open ocean waters. This implicates both a fast turnover of phytoplankton and a tight consumer control of primary production (Gasol *et al.*, 1997).

A trophic pyramid might be considered a nonnested hierarchy. In other words, the apex predator in a marine food chain does not nest within it (as reducible components) all the carnivorous zooplankton, the herbivorous zooplankton, and the phytoplankton. Yet, the system is a hierarchy as defined by a system of communication where entities with slow behavior are at the top, while successively faster behaving entities occur lower in the hierarchy (Allen and Starr, 1982). In principle, the behavior of some nested hierarchies may be determinable if observations are sufficiently complete so that the parts can be added up. There are other properties of nested systems that could prevent this realization; but in any case, actual determinability in real ecosystems is likely unachievable. However, nonnested hierarchies are truly indeterminable because the entities are themselves quasi-independent wholes, with nested hierarchies within each (Allen and Starr, 1982).

For several reasons, the pyramid as described is less useful as a construct for plankton ecosystems than it once may have appeared to be. As a system that transfers mass and energy in a linear route from plants to herbivores to carnivores, the traditional pyramid fails to give explicit account to the microbial loop, which diverts primary production to heterotrophic prokaryotes (bacteria and archaea) and thence to phagotrophic protistan eukaryotes (microzooplankton). Microzooplankton is an important prey for mesozooplankton, so trophic linkage remains; but the microbes as a whole are a large respiratory sink, which greatly alters the dynamic functioning of the system. The pyramid also fails to recognize the viral shunt, which by recycling energy within the microbial loop becomes a futile path for any transfer to metazoans. Finally, the idea that an organism must be of a singular trophic type is too restrictive since mixotrophs are common.

Nevertheless, the partition of plankton into nominal trophic groups offers ecological insight that can be particularly robust in statistical terms. The structural relationship between trophic groups is most often represented by a linear model. As more ecosystems are progressively sampled, a greater degree of natural variability is encountered. Patterns formed by a large number of observations have more freedom to reveal nonlinear structures such as thresholds, boundaries, and inflections (Duarte, 1991). For example, the relationship between oceanic bacterioplankton and phytoplankton appears linear in regional ecosystems, and this is consistent

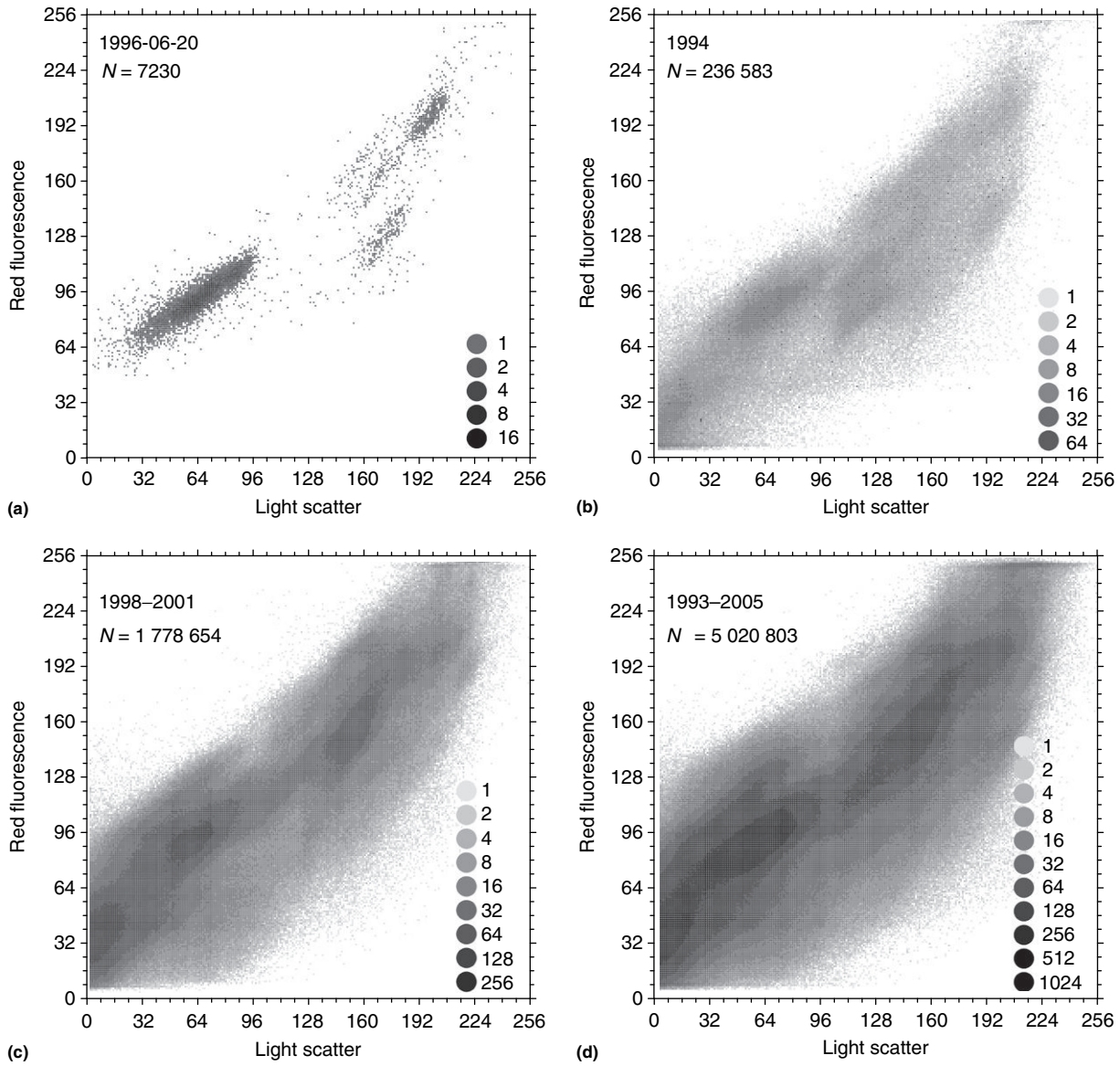


Figure 2 Multiscale analysis by accumulation. Bivariate plots of phytoplankton in Bedford Basin showing the density of particle occurrence in the coordinate space of cellular light scatter (cell size proxy) and cellular red fluorescence (chlorophyll *a* proxy). (a) Measurement on a single day, (b) annual accumulation, (c) quadrennial accumulation, and (d) 13-year accumulation.

with known mechanisms of carbon flux between the interacting trophic groups. However, a global aggregation of these observations reveals a nonlinear constraint on the actualized expression, indicating a need for more descriptive details of other interacting components (Li *et al.*, 2004). Interestingly, with appropriate attention to statistical nuances, it may be possible to invert the problem and use cross-system information to investigate the structural relationship between variables within a single system (Prairie and Marshall, 1995).

Size Classification

Organism size has been a fundamental organizing attribute in plankton ecology since the early days of oceanography when nets of different mesh size were used to selectively retain plankton. The fact that named plankton size classes (see Glossary) are arrayed in logarithmic series is tacit recognition of the importance of power-law scaling. As a means to manage complexity, power-law scaling of plankton size has much appeal because it portends consilience with many other natural systems. Metabolic rate is related to body size of organisms ranging from microbes to the largest vertebrates through a common power exponent whose value is generally (but not universally) taken to be <1. Although new evidence suggests that even suborganismal entities (mitochondria and respiratory complexes) may share the same exponent (West *et al.*, 2002), considerations of nonscalable components within a cell such as a minimum genome and biological membranes of constant thickness may imply a different power exponent for the

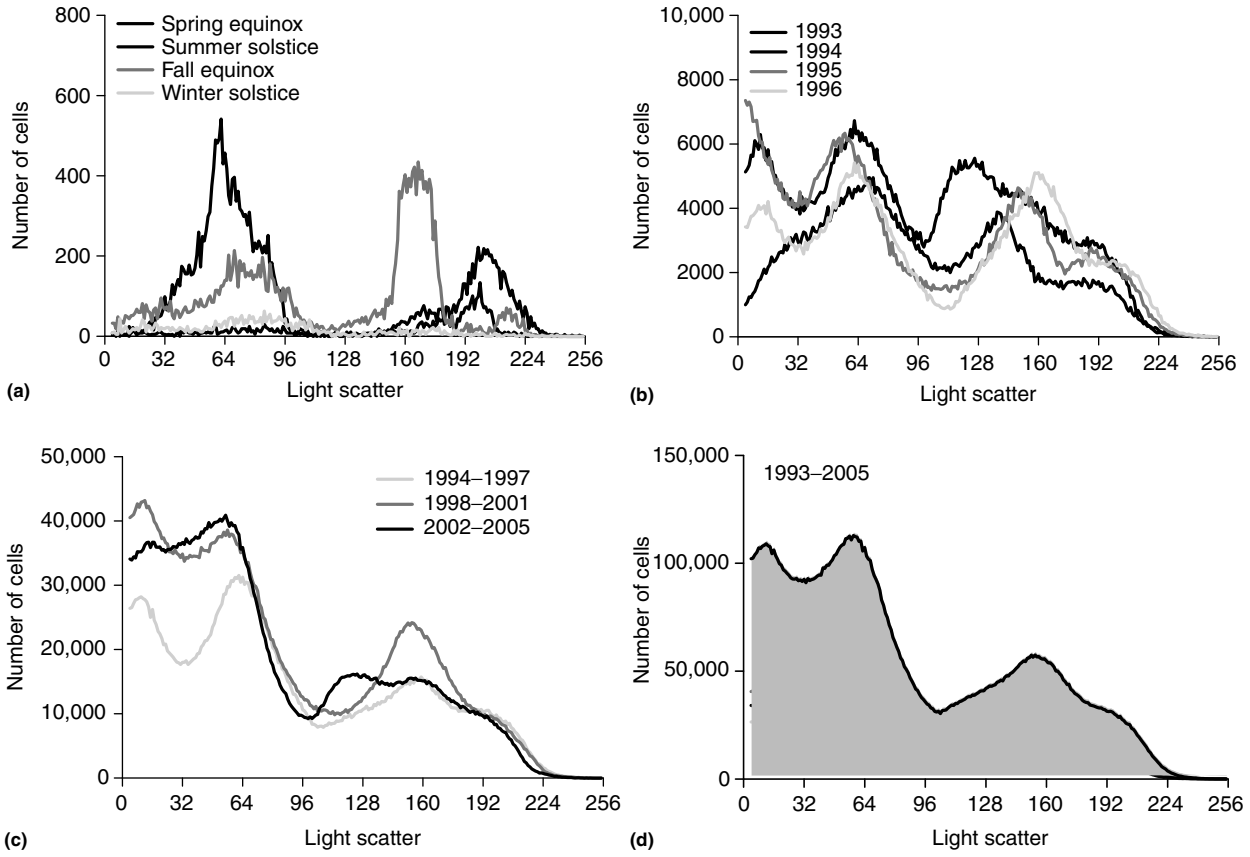


Figure 3 Multiscale analysis by accumulation. Size-frequency histograms of phytoplankton in Bedford Basin. (a) Four weekly measurements, (b) four annual accumulations, (c) three quadrennial accumulations, and (d) 13-year accumulation.

smallest picoplankters (Raven *et al.*, 2005). Notwithstanding the exact value of the exponent, metabolic scaling has given rise to various theoretical models deriving the distribution of organism abundance and biomass according to body size. All such models, implicitly or explicitly, address one of the most striking macroscopic patterns in pelagic ecology. This pattern, first anticipated by pioneering ecologists Charles Elton and Howard Odum, was demonstrated by observations made by Sheldon *et al.* (1972) showing that in the euphotic zone of oceanic regions, "to a first approximation, roughly equal concentrations of material occur at all particle sizes within the range from 1 to 10^6 μm, i.e., from bacteria to whales."

Trophic models of plankton size spectra take apparent universal metabolic scaling within each trophic level, assign a large difference to the body size of predator and prey, and dissipate energy at each feeding step (Kerr, 1974). Biomass invariance across logarithmic size classes can indeed be an outcome of these models (Brown *et al.*, 2004). Microbial ecologists who view the plankton system at a lower hierarchical level may find the simplifications to be unjustified. They see primary producers virtually span the entire size range from picoplanktonic *Prochlorococcus* to macroplanktonic *Trichodesmium*; they see a considerable number of unicells functioning at more than one trophic level; and they see a majority of protistan grazers only slightly larger than their prey, and therefore in a closely adjacent logarithmic size class.

An alternative model of unidirectional biomass flow from small to large organisms (Platt and Denman, 1978), inspired by the cascade of turbulent energy from small to large wave numbers, replaces discrete trophic levels with a size continuum. The model predicts a slight decrease of biomass with size. This has been confirmed in both the North Pacific (Rodríguez and Mullin, 1986) and the North Atlantic (Quiñones *et al.*, 2003) by high-resolution spectra carefully constructed by merging windows of observation derived by different plankton measurement technologies. A strength of this model is that it can be integrated analytically to yield metabolic information for the community as a whole, therefore providing constraints on function based on structure. A difficulty, pointed out by Platt himself, is the need to reconcile the microbial loop, which is a counterflow of material against the general direction.

At an adequately high level of aggregation, such as achieved by accumulating inventories from many different habitats acquired over many years into a "superspectrum," a similitude appears toward a system in which biomass is shared randomly among size classes without preference (Rodríguez, 1994). This type of power-law distribution suggests another way to regard plankton size spectra, namely as a case of scale invariance or self-similarity. In other words, plankton systems may not have any characteristic scale and the components may be organized according to dynamics that cross scales.

Self-similarity thus provides a means to extrapolate between scales.

A notable example of power-law phenomena in plankton can be found in the microbes where, coincidentally, trophic and biomass flow concepts are least clear-cut. In this limited range at the low end of the size spectrum, not only does a power law describe the community (Li, 2002), but also a related power law with the same scaling exponent also describes the relationship between abundance and size for the components: heterotrophic bacteria, *Prochlorococcus*, *Synechococcus*, other picophytoplankton and nanophytoplankton (Rinaldo *et al.*, 2002). The individual distributions are strongly peaked, but partial overlap in size from one to another results in smoothed ensembles. Where one group (e.g., *Prochlorococcus*) might be missing, there might be more of another (e.g., heterotrophic bacteria) or a broadening of the distribution of another (e.g., *Synechococcus*) to fill the gap (Cavender-Bares *et al.*, 2001). Individuals interact with each other locally but become organized into aggregates that appear regular. The size structure of the phytoplankton community may at times be strongly regulated by factors such as competition, grazing, infection, sinking, and physical aggregation, but it also seems that size scaling of cellular nutrient requirements and growth alone can suffice to produce widely observed features such as the power-law relationship, the dominance of small phytoplankton cells in oligotrophic oceans, and the relative increase in large phytoplankton cells under eutrophic conditions (Irwin *et al.*, 2006).

Functional Classification

In terrestrial ecology, plant functional types are sets of species with similar responses to the environment and with similar effects on ecosystem functioning. In essence, this classification seeks functional commonalities in life forms that cut across phylogenetic, autecological, and allometric diversities. Extending this concept to phytoplankton has resulted in detailed assembly rules for freshwater assemblages (Reynolds *et al.*, 2002) and marine dinoflagellates (Smayda and Reynolds, 2001). In these works, the primary emphasis has been on recognizing shared adaptive features that favor a close association of certain organisms well disposed toward certain habitats. These organism associations are mapped against a matrix of habitat gradients (nutrients, irradiance, and water turbulence) to reveal a pattern at a macroscopic level. Functional classification is a polyphyletic approach that arranges organisms according to their biological attributes and sensitivities, which are associated with particular habitat conditions. But even so, proceeding through a hierarchical filter to match the traits of class, genus, and species against habitat conditions, it is still not possible to predict at a low level (even if all traits were completely known) because selection is ultimately stochastic.

The concept of plankton functional types is similar to plant functional types, but oceanographic application emphasizes common effects on ecosystem functioning over common responses to environment. A prospectus for assessing the biogeochemical impact of plankton activities recognizes 10 provisional types: picoheterotrophs, picoautotrophs, phytoplankton nitrogen-fixers, phytoplankton calcifiers, phytoplankton dimethylsulfide producers, phytoplankton silicifiers, mixed phytoplankton, protozooplankton, mesozooplankton, and macrozooplankton (Le Quéré *et al.*, 2005). This is clearly a mixed classification with some elements closely similar to trophic and size classes. Nevertheless, the idea of functional diversity is well accepted. Plankton functional types, appropriately clustered or ordinated, occupy an ecological trait space where the functional distance between types may be quite different from traditional measures of species differences (Weithoff, 2003).

Biogeochemical modeling of plankton functional types at the global scale requires sufficiently well-understood autecologies, a realistic type classification, and robust parameterizations: these are requirements perhaps not well met at present (Anderson, 2005). Nevertheless, it is understood that macroscopic patterns of functional types are a powerful way to study community structure. As an example, phytoplankton silicifiers, which are largely diatoms (but may also include silicoflagellates), can be indicated by their pigment composition or their spectral absorption characteristics. Inferences made from the broad distributions of fucoxanthin (measured by chromatography) and a water-leaving reflectance indicator (derived from satellite data) are consistent with known features of diatom distributions (Platt *et al.*, 2005). Although neither analytical chemistry nor spectral radiometry is incisive enough to provide an irrefutable diagnostic of diatoms, spatial mapping of proxy indicators allows the study of some plankton functional types at the large scale. Other plankton functional types have also been detected by satellites. These include calcifiers that detach high-reflectance coccoliths (Brown and Yoder, 1994), nitrogen-fixing *Trichodesmium* cyanobacteria that have a relatively unique spectral signature of high backscatter from gas vesicles, combined with the absorption and fluorescence of phycoerythrin (Subramaniam *et al.*, 2002), and bioluminescent bacteria which continuously emit light when concentrations of autoinducer are locally high (Miller *et al.*, 2005).

Phylogenetic Classification

It is in the sense of phylogeny that biological diversity is most widely understood. The arrangement of organisms according to their evolutionary relatedness in a nested hierarchy of taxonomic ranks is the basis on which marine autecology, conservation, and resource management is generally conducted. Many general ecological theories are strongly grounded in evolutionary relationships between species; therefore, it seems expedient to explore macroscopic patterns of plankton in that framework. However, to a considerable extent, those theories rely on human perception of morphology and lifestyle as a guide to differentiating between organisms. This guide fails us at the small scales of microbial plankton because organisms with a high ratio of surface area to volume appear to converge on relatively uniform morphological shapes such as spheres and rods. When organisms are very small, adaptive adjustments largely appear within the living volume rather than at the surface. Whereas macrobes tend to be diverse in shape and surface embellishments, microbes tend to be

diverse in metabolism and intracellular functionality. The nanoplankton are poised between the femtoplankton and picoplankton where the concept of morphospecies is conceded to be virtually untenable, and the microplankton, mesoplankton, and macroplankton where the concept is at least of arguable validity.

In microplankton and larger plankton, there is sufficient apparent confidence in taxon recognition of particular forms of diatoms, dinoflagellates, copepods, foraminiferans, and euphausiids that global distributions of these and others can be mapped for biogeographical study. Large-scale distributions of calanoid copepods are particularly well documented. Although taxonomic richness, evenness, and distinctness all decline from the tropics to the poles, they do so in different manners that are best explained by patterns of primary and secondary production consistent with Longhurst ecological geography (Woodd-Walker *et al.*, 2002). From the Continuous Plankton Recorder database, the information on greater than 100 taxa of calanoid copepods is sufficiently extensive in space and intensive in time that diversity can be examined in large contiguous ocean regions at diel, seasonal, and annual scales (Beaugrand and Ibañez, 2002). These multiscale considerations lead to a view that calanoid diversity is mainly regulated by temperature, hydrodynamics, stratification, and seasonal variability of the environment. But even so, a taxon as familiar as *Calanus* still cannot be regarded as taxonomically well described (Longhurst, 1998). In phytoplankton, when the richness of recognized species is examined in relationship to the surface area of the ecosystem, the familiar species–area power law is found, with an invariant exponent that spans many orders of magnitude in spatial extent (Smith *et al.*, 2005). Observations, or experiments, at one scale can therefore be used to inform at other scales. However, as with zooplankton, the taxonomy of even familiar species such as the diatom *Skeletonema costatum* may need revision (Sarno *et al.*, 2005).

In picoplankton (bacteria, archaea, and picoeukaryotes), the concept of species is vague but biological diversity can be inferred from nucleic acid sequences. "The individual is where the gene meets the environment" (Levin *et al.*, 2001); so if gene phylogeny equates to organism phylogeny, then genetically defined organismal units can also be expected to have a macroecology. Plankton phylogenetic trees are most commonly constructed from the small subunit ribosomal RNA gene. The structure and topology of the trees indicate a high degree of novelty and diversity in marine microbes. Were phylogenies better visualized as webs rather than trees, for example, because of prevalent lateral gene transfer, it would not change present consensus that the smallest organisms are also the most diverse.

In fact, the extent of diversity is even greater than that indicated by the dissimilarity of single genes. There is microdiversity. A group of coastal bacterioplankton showing nearly identical ribotype identity to *Vibrio splendidus* was estimated to consist of at least a thousand distinct genotypes (Thompson *et al.*, 2005). Similarly, oceanic bacterioplankton off California revealed greater than a thousand lineages from three clusters alone (Brown and Fuhrman, 2005). Some of this extensive diversity may be neutral with no selective advantage in the ecological context. Microdiverse clusters are thought to arise from periodic selection, but intracluster competition is not strong enough to sweep (purge) members from the group. Yet, in at least one group of marine bacterioplankton, phylogenetically discrete clusters denoted by highly similar gene sequences are represented by ecologically differentiated units. These so-called ecotypes display different cellular characteristics, allowing them to partition environmental gradients in the ocean. In the photoautotrophic cyanobacterium *Prochlorococcus*, ecotypes are distributed in the ocean with respect to temperature, light, and nutrients in accord with functional capabilities conferred by their respective genomes (Johnson *et al.*, 2006). The global basin-scale distribution of *Prochlorococcus* ecotypes at the sea surface is consistent with prevailing environmental conditions resulting from physical oceanographic structuring of the water column (Bouman *et al.*, 2006).

The level of taxonomic resolution at which biogeographic and macroscopic patterns become evident is important in understanding evolutionary and community assembly processes. Genetic and genomic characterization of plankton suggest that some of the most finely resolved taxonomic units may not be distributed in random fashion. It follows that large-scale patterns should be discernible at higher taxonomic levels. At the genus level, the two marine cyanobacteria *Prochlorococcus* and *Synechococcus* have been mapped extensively in the ocean on the operational basis of cytometric characteristics. In the three Southern Subtropical Gyres (Bouman *et al.*, 2006), the surface distributions of the two genera are antiparallel (Pacific and Atlantic Oceans) or parallel (Indian Ocean) to each other (**Figure 4(a)**). In a bivariate plot of cell abundances, the data pairs seem to lie within a triangular polygon (**Figure 4(b)**). The orientation of the polygon indicates a large range of possible densities for *Prochlorococcus* at high densities of *Synechococcus*, but only a restricted high range for *Prochlorococcus* at low densities of *Synechococcus*. At the within-gyre scale, different processes evidently regulate the interaction between the two genera. At the larger between-gyre scale, there is a simple negative correlation between basin-averaged abundances of *Prochlorococcus* and *Synechococcus* in the order of Indian Ocean to Atlantic Ocean to Pacific Ocean (**Figure 4(b)**). This niche complementarity, earlier noted by Chisholm (1992) at a different spatiotemporal scale, is evident at a level of the taxonomic hierarchy higher than that of the interacting units, which are presumably the ecotypes or lower-level genomic variants. A decrease in the numerical abundance of *Prochlorococcus* is not fully offset by an increase in *Synechococcus*, but cells of the former are somewhat smaller on average than those of the latter, so there is a degree of ecological homeostasis in the combined cyanobacterial biomass.

It seems that different taxa can coexist by being sufficiently different or by being sufficiently similar. Taxa repulsion occurs when competition excludes others of moderate similarity. On the other hand, taxa attraction occurs when others are similar enough to avoid competition and results in self-organized clustered niches, a so-called paradox of the clumps (Nee and Colegrave, 2006). As evidence, the frequency distribution of phytoplankton species in Dutch lakes appears to be clustered along a niche axis defined by cell volume (Scheffer and van Nes, 2006). *Prochlorococcus* and *Synechococcus* share similar positions in the food web, physical cell size, dominant metabolic capabilities, and cellular biochemistry. For these

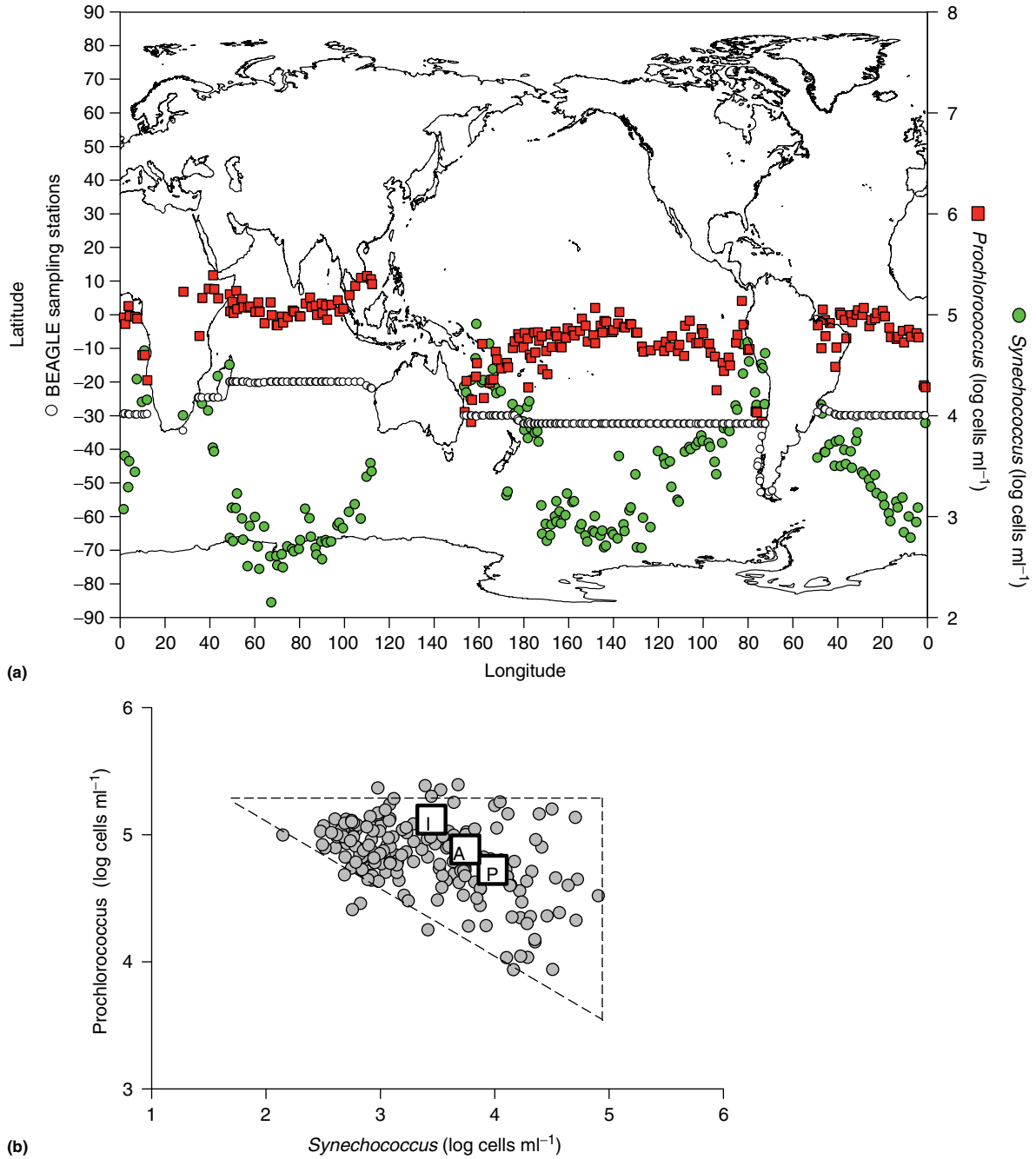


Figure 4 Abundance of *Prochlorococcus* and *Synechococcus* in the Southern Subtropical Gyres. (a) Map shows sampling stations in horizontal transects along 20° and 30° S, and U-shaped (upright or inverted) cell distributions in each ocean basin. (b) bivariate plot of cell abundance for individual stations, and ocean average values: I = Indian Ocean, A = Atlantic Ocean, P = Pacific Ocean. Triangular polygon is drawn by eye to envelop the data cluster.

picoplankton, the phylogenetic perspective becomes intimately enmeshed with those based on trophic linkages, size allotment, functional performances, and indeed also with biotic elemental composition.

Stoichiometric Emergence

All living beings are made of the same few classes of biomolecules (e.g., proteins, carbohydrates, lipids, nucleic acids, nucleotides, and pigments), which are constituted from the same few elements, of which the three major ones are carbon (C), nitrogen (N), and phosphorus (P). A consideration of the C:N:P stoichiometry of plankton is in effect a multiscale analysis from elements and molecules up through the biological hierarchy to populations, communities, ecosystems, and even ultimately to the biogeosphere (Sterner and Elser, 2002). In this sense, ecological stoichiometry is the quintessence of macroscopic emergence from microscopic organisms. Despite perceived

biodiversity and biocomplexity in plankton systems, there is apparent biosimplicity in the stoichiometric equivalence of oceanic plankton biomass and deep dissolved inorganic nutrients in the atomic ratio of 106:16:1, well known as the Redfield ratio. The concept of an elemental composition common to the living and nonliving parts of the Earth is alluring because it lends easily to a teleological metaphor of the biosphere as a superorganism on which selection acts. However, as it is most widely understood, selection acts at a much lower level. It is the genomic instructions within individual organisms that provide controlling directions. Accordingly, the observed stoichiometry of marine seston is a high-level pattern resulting from nested processes leading from fast molecular biology to slow biogeochemistry; it is not a Gaian homeostasis (Levin, 2005).

The concept of a common elemental composition in plankton is useful for biogeochemistry because the laws of matter conservation and stoichiometric proportions allow nutrient inventory to be balanced amongst the major elements and also between the biota and the water. However, because the Redfield ratio is a macroscopic pattern, it may not necessarily be robust at local scales, for example, when used to calculate a nonsteady-state drawdown of dissolved nutrients from observed local biomass accumulation in multiphyletic plankton assemblages (**Figure 5**). The Redfield ratio is an empirical statistical average (**Figure 5(a)**) and not a fundamental biochemical constraint. Stoichiometric plasticity is possible in photolithoautotrophs because the elements are obtained from disparate soluble inorganic forms (e.g., bicarbonate, nitrate, and phosphate) whose supplies are not necessarily coupled. Indeed, phytoplankton has a considerable capacity for intracellular storage of excess nutrients. By contrast, heterotrophs obtain their elements together in preformed food items. If there is stoichiometric similarity between an organism and its resources, then there is some degree of homeostasis (Sterner and Elser, 2002). In this respect, there is a distinction between autotrophs and heterotrophs: Homeostasis is less evident in phytoplankton than in chemoorganotrophic bacterioplankton and zooplankton.

Organisms, especially autotrophs, vary considerably in elemental composition. Systematic phylogenetic differences in C:N:P exist between major superfamilies of phytoplankton, indicative of ancestral phenotypes (Quigg *et al.*, 2003). Furthermore, microalgae exhibit considerable physiological plasticity of C:N:P in response to nutrient and light conditions (Geider and La Roche, 2002). The smallest cyanobacteria (*Prochlorococcus* and *Synechococcus*) are characterized by very high C:P ratios due to low cellular P (Bertilsson *et al.*, 2003). Under P-limitation, most of this element in the cyanobacteria is found in DNA, a nonscalable cell component that constrains both minimum cell size and elemental stoichiometry

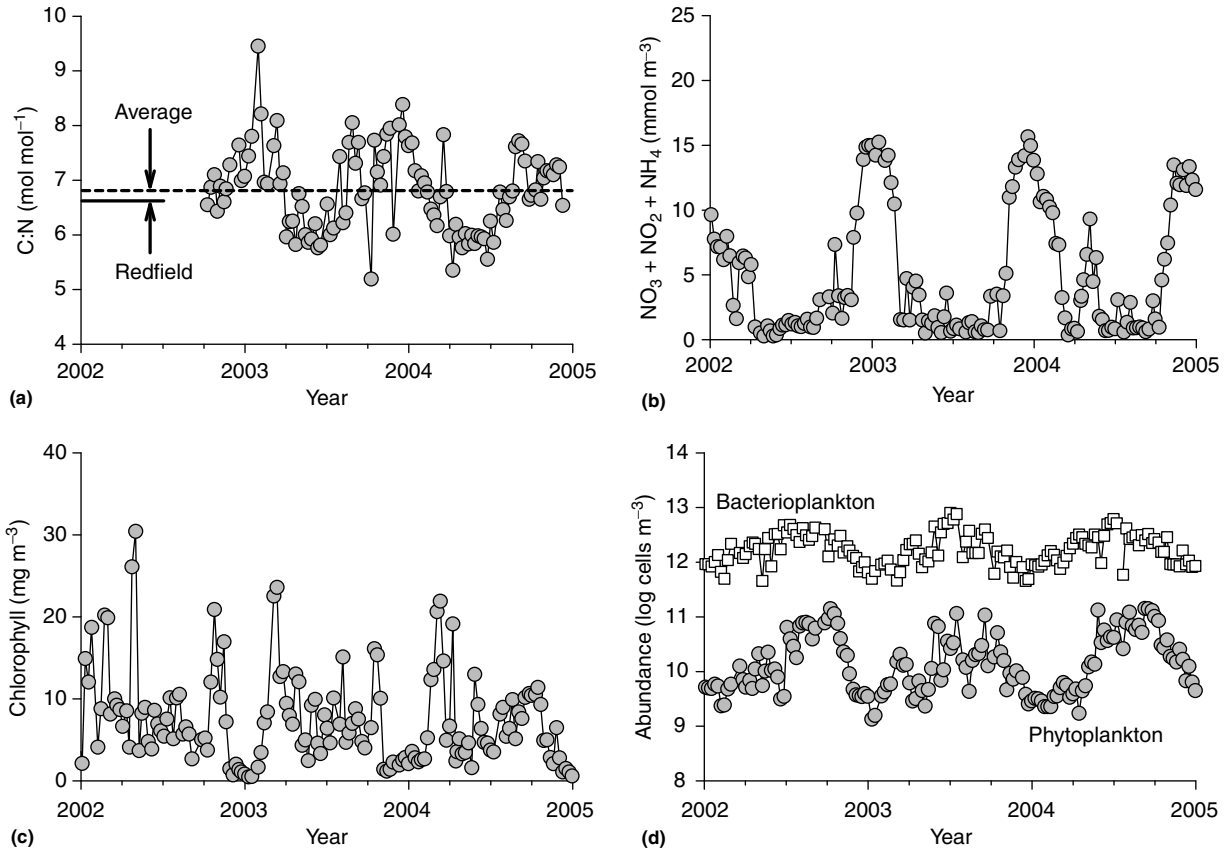


Figure 5 Time series of plankton and nutrients in Bedford Basin. (a) Particulate C:N atomic ratio showing the similarity of the average value to the Redfield ratio, (b) inorganic nitrogenous nutrients, (c) chlorophyll *a* concentration, and (d) numerical abundance of phytoplankton and heterotrophic bacterioplankton.

(Raven, 1994). Exponentially growing phytoplankton depend on intracellular growth machinery such as ribosomes which have a low N:P ratio; conversely, near-equilibrium slow-growth phytoplankton depend on resource-acquisition machinery such as nutrient-uptake proteins and chloroplasts which have a high N:P ratio. The appropriate mix of all these stoichiometric types results in the Redfield ratio (Klausmeier *et al.*, 2004). Oceanographic and atmospheric regimes that shift the relative availability of nutrients to phytoplankton may alter stoichiometric steady states, which might be discernible at regional and global scales.

Temperature Dependence

From Single Molecules to Biochemical Kinetics

Temperature affects the velocity of biochemical reactions which drive organism metabolism. Statistical thermodynamics is in essence a macroscopic rule used to predict the effects of temperature on chemical reactions. In an ideal gas at thermal equilibrium, it is impossible to describe the kinetic energy associated with each individual molecule, but Maxwell–Boltzmann distribution law shows that the probability of a molecule occupying a kinetic energy E is proportional to $e^{-E/kT}$ where k is Boltzmann constant (8.62×10^{-5} eV K^{-1}) and T the system's absolute temperature (K). There is an interesting resemblance between what this law represented (a bridge between microscopic physicists interested in single molecule dynamics and macroscopic physicists interested in whole systems dynamics) and the aim of macroecology as a discipline linking ecophysiology and community ecology.

Implicit in Maxwell–Boltzmann equation is the fact that as temperature increases the proportion of molecules with sufficient kinetic energy to react increases. This ultimately led Svante Arrhenius to formulate that the temperature dependence of reaction kinetics should scale as $e^{-E_a/RT}$ where R is the gas constant (8.31 J mol^{-1} K^{-1}) and E_a is the activation energy of that particular reaction; reactions with higher E_a show stronger temperature dependence. The difference between Boltzmann's factor and Van't Hoff–Arrhenius equation is just that the former uses particle units (E in eV and k in eV K^{-1}) while the latter uses the molar scale (E_a in J mol^{-1} and R in J mol^{-1} K^{-1}). Because k and R are related by $k = R/(N_A \times j)$, where N_A is Avogadro constant ($N_A = 6.022 \times 10^{23}$) and a j conversion factor from electronvolts to Joules ($j = 1.602 \times 10^{-19}$ J eV^{-1}), both expressions have the same value. The more intuitive Q_{10} constant results from simplification of Van't Hoff–Arrhenius equation (Gillooly *et al.*, 2002).

From Biochemical Kinetics to Whole Organism Physiology

A second scaling step is needed to apply Boltzmann's–Arrhenius equation to the response of whole organism physiological rates to temperature. Although Boltzmann's–Arrhenius equation has been widely applied to biological systems, just recently the activation energy measured for whole-organism rates has been related to the activation energy of the main biochemical reactions that drive organism physiology (Gillooly *et al.*, 2001). Metabolic scaling theory proposes that because organism metabolism is mainly driven by mitochondrial respiration, the temperature dependence of adenosine triphosphate (ATP) synthesis should determine the temperature dependence of whole-organism metabolic rate with an average activation energy for both processes close to 0.65 eV (or 62.7 kJ mol^{-1}) (Gillooly *et al.*, 2001). Although whole-organisms greatly depart from the ideal gas conditions for which Boltzmann's–Arrhenius equation was formulated and more complex explanations have emerged, the similar activation energy at both levels of organization is quite remarkable. For planktonic heterotrophs, whose metabolism is mainly driven by the synthesis of ATP, the activation energy for both respiratory (Lopez-Urrutia *et al.*, 2006) and growth rates (Rose and Caron, 2007) is close to the predicted value of 0.65 eV. But the temperature dependence of the metabolic rates of photosynthetic plankton has long been recognized to be lower than that of heterotrophs (Eppley, 1972) with values ranging from 0.29 to 0.39 eV (Lopez-Urrutia *et al.*, 2006; Rose and Caron, 2008). Allen *et al.* (2005) argued that the lower temperature dependence of land-plant photosynthetic rates (0.32 eV) is due to the kinetics of rubisco carboxylation and photorespiration. As temperature increases photorespiration increases relative to carboxylation thus reducing net carbon gain. The applicability of Allen *et al.* (2005) explanation to marine plankton depends on the assumption that CO_2 supply to Rubisco in marine plants, which have different diffusive characteristics and carbon concentration mechanism (Yvon-Durocher *et al.*, 2010), does not modify this theoretical explanation.

From Organisms to Communities, Biogeochemical Cycles, and Biodiversity

Regardless of the theoretical basis for the differential temperature dependence of the metabolic and growth rates of autotrophs and heterotrophs, the implication for to marine plankton community dynamics and biogeochemical cycles are far reaching. Huntley and Boyd (1984) showed that the zooplankton to phytoplankton production ratio would increase as temperature increases. Rose and Caron (2007) argued that phytoplankton blooms might occur more frequently in cold waters because the growth of grazers will be much lower than the growth of phytoplankton as temperature decreases. Laws *et al.* (2000) showed that the proportion of primary production exported to the deep increases with decreasing temperature because, at cold temperatures, the growth rates of heterotrophic decomposers are much lower so most organic matter is exported before it can be decomposed. Lopez-Urrutia *et al.* (2006) scaled the differential temperature dependence of planktonic respiration and photosynthesis to the metabolic balance of whole plankton communities and showed that as temperature increases the ratio of community production to respiration decreases.

Through its effects on organism metabolic rates, temperature also affects community structure. Higher cell division rates with increasing temperature might be responsible for the stronger DNA evolution and speciation rates observed in planktonic foraminifera toward the tropics (Allen and Gillooly, 2006). This kinetic energy hypothesis plays a fundamental role on the observed temperature dependent global patterns of marine biodiversity both in planktonic and other marine communities (Tittensor *et al.*, 2010).

Regional and Global Change

In addressing societal concerns for which plankton ecology is relevant, the dominant patterns are arguably to be sought at the macroscopic level. At this level, we will not always have both pattern and process firmly in place before responding to calls for prediction. Climate change, ecological regime shift, fisheries reallocation, and biodiversity loss occur across multiple scales, but impacts are most consequential at decadal and longer timescales, at large regional and global spatial scales. Modern management of marine resources based on ecosystem considerations requires us to look for evidence and impact of plankton change at the whole system level. Plankton, at least the small ones, live for only days or weeks. Thus, there is a need to link local interactions with regional–global patterns.

Such a linkage can be made when a testable theory is supported by sufficient data. The apparent universal scaling of biological attributes with body size is an important macroscopic pattern. Another is the dependence of metabolic rates on temperature and light (for phototrophs). Together, they underlie a general metabolic theory (Brown *et al.*, 2004), which scales individual rates of photosynthesis and respiration of constituent plankton organisms to ecosystem carbon flux rates of primary production and community respiration (López-Urrutia *et al.*, 2006). Photosynthesis is less dependent on temperature than respiration, and this leads to a prediction that the balance of biogenic carbon flux could be altered by climate change: Less CO_2 would be captured in the surface ocean as temperature increases.

Rules of community assembly prescribe the successional development of populations over a year. In the sea, unlike on land, the scales of variability are such that changing community patterns fit well in the physical context of hydrodynamics (Steele, 1998). Temperature, being an important determinant of both cell metabolism and water column stability, as well as an indicator of seasonal progression and climate change, is an ecosystem driver acting across many scales. Thus, it is perhaps not astounding, but remarkable nevertheless, that over a large region of the temperate and subarctic northwest Atlantic Ocean, there is a coherence in the wax and wane of phytoplankton cells over the annual cycle that is aligned by temperature (Li *et al.*, 2006a). By this rule, phytoplankton are not assembled according to taxonomy, size, or functional attribute, but only by virtue of being individual particulate autotrophic entities. In other words, it might be construed that the molecular basis for trophic mode is ascendant over any particular organismal form under these conditions.

Seasonal cycles recur annually, but the question is whether phase and amplitude remain the same from year to year. There is some evidence that such phenological change is occurring in plankton groups of the North Sea (Edwards and Richardson, 2004), with the implication that a temporal mismatch across trophic levels signals a decoupling of ecosystem components. On the Nova Scotian Shelf and Labrador Sea, there is other evidence that the seasonal means of phytoplankton biomass have been increasing in some areas, but decreasing in others. Indeed, it seems that these multiyear trends have propagated to the heterotrophic bacterioplankton because they also exhibit these trends: a case of coherent sign switching in the direction of long-term change (Li *et al.*, 2006b). Generation times of phytoplankton and bacterioplankton are not greatly dissimilar, but become increasingly disparate from those of primary, secondary, and tertiary consumers at higher trophic levels. Do signals propogate through many levels with fidelity? A continental shelf ecosystem monitored for 40 years appears to show trophic cascades across four levels: large-bodied predators (benthic fish and seals), forage species (pelagic fish and macroinvertebrates), herbivorous zooplankton (calanoid copepods), and phytoplankton (green seston) (Frank *et al.*, 2005). This marine ecosystem is highly complex, and thus the signature of top-down control in the form of simple trophic cascade can be viewed as a macroscopic pattern. The actual details of grazing, forage, and predation are not important in a multidecade trend.

The enterprise of science, like nature, is hierarchical. International consortia work toward a capacity for global ecology; national networks are placed to embrace regional interests; and local alliances exist to address immediate concerns. However, a space–time hierarchy of institutional management of marine resources will not necessarily ensure appropriate normative behavior because human and ecological dynamics are mismatched in scale (Rice, 2001; Ehrlich and Levin, 2005). Moreover, multiscale considerations will inevitably be confronted by ecological fungibility (Steele, 1998), which is the substitutability of one entity by another of equal value or utility. Thus, at some level, the differences in biological patterns caused by historical contingencies and contemporary environmental factors become moot. If fungibility can be expressed in terms of energy flux, allometric linkage, adaptive attributes, functional performance, diversity indices, elemental balance, or other such concepts, then we may be able to see evidence of laws in ecology (Lawton, 1999). Part of the macroecological agenda is to delineate the range of natural variability in given ecosystems so that outliers may be identified, even if the process constraints on natural limits are not yet fully understood. From a management perspective, these limits may suggest points of natural thresholds that trigger precaution toward undesired system states. To describe how a large continental shelf ecosystem has devolved from one regime to another using indicators of ocean climate, plankton, invertebrates, fish, and human influence, and then to predict a potential fishery collapse in an adjacent ecosystem within a decade (Choi *et al.*, 2005) is a bold and consequential test of the empirical approach.

Concluding Remarks

Biological diversity is conceptualized by assigning living entities to different categories. Our perception of meaningful categories is based on what we observe. By and large, we observe whole organisms, namely the units that house the genes. However, it is possible to consider, and in fact construct by technology, an entire plankton community as the sum of all constituent genomes, a metagenome. Sequencing of the community genome yields information on ecological structure, function, and diversity that crosses multiple levels of the conventional biological hierarchy. Although such ecogenomic studies are nascent, comparative ecosystem analyses of metagenomes will undoubtedly yield patterns of macroscopic organization. Macroecology, by its holistic nature, is sometimes seen as an approach to rise above the repeating cycle of

data-intensive reductionism. A multiscale approach enables emergent patterns to consolidate myriad historical contingent events. Seemingly inexhaustible biodiversity can be telescoped to elegant biosimplicity.

See also: Marine Ecosystems. Microorganisms (Microbes), Role of. Ocean Ecosystems. Pelagic Ecosystems. Plankton, Status and Role of. Restoration of Animal, Plant, and Microbial Diversity. Scale, Concept and Effects of. Trophic Levels

References

Allen AP and Gillooly JF (2006) Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecology Letters* 9: 947–954.

Allen AP, Gillooly JF, and Brown JH (2005) Linking the global carbon cycle to individual metabolism. *Functional Ecology* 19: 202–213.

Allen TFH and Starr TB (1982) *Hierarchy: Perspectives for Ecological Complexity.* Chicago: University of Chicago Press.

Anderson TR (2005) Plankton functional type modelling: Running before we can walk? *Journal of Plankton Research* 27: 1073–1081.

Beaugrand G and Ibañez F (2002) Spatial dependence of calanoid copepod diversity in the North Atlantic Ocean. *Marine Ecology Progress Series* 232: 197–211.

Bertilsson S, Berglund O, Karl DM, and Chisholm SW (2003) Elemental composition of marine *Prochlorococcus* and *Synechococcus*: Implications for the ecological stoichiometry of the sea. *Limnology and Oceanography* 48: 1721–1731.

Bouman HA, Ulloa O, Scanlan D, *et al.* (2006) Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. *Science* 312: 918–921.

Brown CW and Yoder JA (1994) Coccolithophorid blooms in the global ocean. *Journal of Geophysical Research* 99: 7467–7482.

Brown JH (1995) *Macroecology.* Chicago: University of Chicago Press.

Brown JH, Gillooly JF, Allen AP, Savage VM, and West GB (2004) Toward a metabolic theory of ecology. *Ecology* 85: 1771–1789.

Brown MV and Fuhrman JA (2005) Marine bacterial microdiversity as revealed by internal transcribed spacer analysis. *Aquatic Microbial Ecology* 41: 15–23.

Cavender-Bares KK, Rinaldo A, and Chisholm SW (2001) Microbial size spectra from natural and nutrient enriched ecosystems. *Limnology and Oceanography* 46: 778–789.

Chisholm SW (1992) Phytoplankton size. In: Falkowski PG and Woodhead AD (eds.) *Primary Productivity and Biogeochemical Cycles in the Sea.* New York: Plenum Press.

Choi JS, Frank KT, Petrie BD, and Leggett WC (2005) Integrated assessment of a large marine ecosystem: A case study of the devolution of the eastern Scotian Shelf, Canada. *Oceanography and Marine Biology: An Annual Review* 43: 47–67.

Duarte CM (1991) Variance and the description of nature. In: Cole J, Lovett G, and Findlay S (eds.) *Comparative Analyses of Ecosystems: Patterns, Mechanisms and Theories*, pp. 301–318. New York: Springer.

Ducklow HW (2003) Biogeochemical provinces: Towards a JGOFS synthesis. In: Fasham MJR (ed.) *Ocean Biogeochemistry: The Role of the Ocean Carbon Cycle in Global Change.* Berlin: Springer.

Edwards M and Richardson AJ (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430: 881–884.

Ehrlich PR and Levin SA (2005) The evolution of norms. *PloS Biology* 3: 943–948.

Eppley R (1972) Temperature and phytoplankton growth in the sea. *Fishery Bulletin(Seattle)* 70: 1063–1085.

Frank KT, Petrie B, Choi JS, and Leggett WC (2005) Trophic cascades in a formerly cod-dominated ecosystem. *Science* 308: 1621–1623.

Gasol JM and Duarte CM (2000) Comparative analyses in aquatic microbial ecology: How far do they go? *FEMS Microbiology Ecology* 31: 99–106.

Gasol JM, del Giorgio PA, and Duarte CM (1997) Biomass distribution in marine planktonic communities. *Limnology and Oceanography* 42: 1353–1363.

Geider RJ and La Roche J (2002) Redfield revisited: Variability of C:N:P in marine microalgae and its biochemical basis. *European Journal of Pharmacology* 37: 1–17.

Gillooly JF, Brown JH, West GB, Savage VM, and Charnov EL (2001) Effects of size and temperature on metabolic rate. *Science* 293: 2248–2251.

Gillooly JF, Charnov EL, West GB, Savage VM, and Brown JH (2002) Effects of size and temperature on developmental time. *Nature* 417: 70–73.

Haury LR, McGowan JA, and Wiebe PH (1978) Patterns and processes in the time-space scales of plankton distributions. In: Steele JH (ed.) Spatial Pattern in Plankton Communities, pp. 277–327. New York: Plenum Press.

Huntley M and Boyd C (1984) Food-limited growth of marine zooplankton. *American Naturalist* 124: 455–478.

Irwin AJ, Finkel ZV, Schofield OME, and Falkowski PG (2006) Scaling-up from nutrient physiology to the size-structure of phytoplankton communitites. *Journal of Plankton Research* 28: 459–471.

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, and Chisholm SW (2006) Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* 311: 1737–1740.

Kerr SR (1974) Theory of size distribution in ecological communities. *Journal of the Fisheries Research Board of Canada* 31: 1859–1862.

Klausmeier CA, Litchman E, Daufresne T, and Levin SA (2004) Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature* 429: 171–174.

Laws EA, Falkowski PG, Smith Jr. WO, Ducklow H, and McCarthy JJ (2000) Temperature effects on export production in the ocean. *Global Biogeochemical Cycles* 14: 1231–1246.

Lawton JH (1999) Are there general laws in ecology? *Oikos* 84: 177–192.

Le Quéré C, Harrison SP, Prentice IC, *et al.* (2005) Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology* 11: 2016–2040.

Levin SA (1998) Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* 1: 431–436.

Levin SA (2005) Self-organization and the emergence of complexity in ecological systems. *BioScience* 55: 1075–1079.

Levin SA, Dushorf J, and Keymer JE (2001) Community assembly and the emergence of ecosystem pattern. *Scientia Marina* 65(supplement 2): 171–179.

Li WKW (2002) Macroecological patterns of phytoplankton in the northwestern North Atlantic Ocean. *Nature* 419: 154–157.

Li WKW, Harrison WG, and Head EJH (2006a) Coherent assembly of phytoplankton communities in diverse temperate ocean ecosystems. *Proceedings of the Royal Society B* 273: 2953–2960.

Li WKW, Harrison WG, and Head EJH (2006b) Coherent sign switching in multiyear trends of microbial plankton. *Science* 311: 1157–1160.

Li WKW, Head EJH, and Harrison WG (2004) Macroecological limits of heterotrophic bacterial abundance in the ocean. *Deep-Sea Research Part I* 51: 1529–1540.

Longhurst A (1998) *Ecological Geography of the Sea.* San Diego: Academic Press.

López-Urrutia Á, San Martin E, Harris RP, and Irigoien X (2006) Scaling the metabolic balance of the oceans. *Proceedings of the National Academy of Sciences of the USA* 103: 8739–8744.

Margalef R (1994) Through the looking glass: How marine phytoplankton appears through the microscope when graded by size and taxnomically sorted. *Scientia Marina* 58: 87–101.

Miller SD, Haddock SHD, Elvidge CD, and Lee TF (2005) Detection of a bioluminescent milky sea from space. *Proceedings of the National Academy of Sciences of the USA* 102: 14181–14184.

Nee S and Colegrave V (2006) Paradox of the clumps. *Nature* 441: 417–418.

Platt T (1972) Local phytoplankton abundance and turbulence. *Deep-Sea Research* 19: 183–187.

Platt T, Bouman H, Devred E, Fuentes-Yaco C, and Sathyendranath S (2005) Physical forcing and phytoplankton distributions. *Scientia Marina* 69(supplement 1): 55–73.

Platt T and Denman KL (1975) Spectral analysis in ecology. *Annual Review of Ecology and Systematics* 6: 189–210.

Platt T and Denman KL (1978) The structure of pelagic marine ecosystems. *Rapports et Procés-Verbaux des Réunions Conseil International pour l'Exploration de la Mer* 173: 60–65.

Platt T and Sathyendranath S (1999) Spatial structure of pelagic ecosystem processes in the global ocean. *Ecosystems* 2: 384–394.

Prairie YT and Marshall CT (1995) On the use of structured time-series to detect and test hypotheses about within-lakes relationships. *Canadian Journal of Fisheries and Aquatic Sciences* 52: 799–803.

Quigg A, Finkel ZV, Irwin AJ, *et al.* (2003) The evolutionary inheritance of elemental stoichiometry in marine phytoplankton. *Nature* 425: 291–294.

Quiñones RA, Platt T, and Rodríguez J (2003) Patterns of biomass-size spectra from oligotrophic waters of the Northwest Atlantic. *Progress in Oceanography* 57: 405–427.

Raven JA (1994) Why are there no picoplanktonic O_2 evolvers with volumes less than 10^{-19} m^3? *Journal of Plankton Research* 16: 565–580.

Raven JA, Finkel ZV, and Irwin AJ (2005) Picophytoplankton: Bottom–up and top–down controls on ecology and evolution. *Vie et Milieu* 55: 209–215.

Redfield AC (1960) The inadequacy of experiment in marine biology. In: Buzzati-Traverso AA (ed.) *Perspectives in Marine Biology*. Berkeley: University of California Press.

Reynolds CS, Huszar V, Kruk C, Naselli-Flores L, and Melo S (2002) Towards a functional classification of the freshwater phytoplankton. *Journal of Plankton Research* 24: 417–428.

Rice J (2001) Implications of variability on many time scales for scientific advice on sustainable management of living marine resources. *Progress in Oceanography* 49: 189–209.

Rinaldo A, Maritan A, Cavender-Bares KK, and Chisholm SW (2002) Cross-scale ecological dynamics and microbial size spectra in marine ecosystems. *Proceedings of the Royal Society of London B* 269: 2051–2059.

Rodríguez J (1994) Some comments on the size-based structural analysis of the pelagic ecosystem. *Scientia Marina* 58: 1–10.

Rodríguez J and Mullin M (1986) Relation between biomass and body weight of plankton in a steady state oceanic ecosystem. *Limnology and Oceanography* 31: 361–370.

Rose J and Caron D (2007) Does low temperature constrain the growth rates of heterotrophic protists? Evidence and implications for algal blooms in cold waters. *Limnology and Oceanography* 52: 886–895.

Sarno D, Kooistra WHCF, Medlin LK, Percopo I, and Zingone A (2005) Diversity in the genus *Skeletonema* (Bacillariophyceae). II. An assessment of the taxonomy of *S. costatum*-like species with the inscription of four new species. *Journal of Phycology* 41: 151–176.

Scheffer M and van Nes EH (2006) Self-organized similarity, the evolutionary emergence of groups of similar species. *Proceedings of the National Academy of Sciences of the USA* 103: 6230–6235.

Schneider DC (2001) Scale, concept and effects of. In: Levin SA (ed.) *Encyclopedia of Biodiversity*, pp. 245–254. San Diego: Academic Press.

Sheldon RW, Prakash A, and Sutcliffe Jr WH (1972) The size distribution of particles in the ocean. *Limnology and Oceanography*. 17: 327–340.

Smayda TJ and Reynolds CS (2001) Community assembly in marine phytoplankton: Application of recent models to harmful dinoflagellate blooms. *Journal of Plankton Research* 23: 447–461.

Smith VH, Foster BL, Grover JP, Holt RD, Leibold MA, and deNoyelles Jr. F (2005) Phytoplankton species richness scales consistently from laboratory microcosms to the world's oceans. *Proceedings of the National Academy of Sciences of the USA* 102: 4393–4396.

Solé RV, Alonso A, and McKane A (2002) Self-organized instability in complex ecosystems. *Philosophical Transactions of the Royal Society of London B* 357: 667–681.

Steele JH (1985) Comparison of marine and terrestrial ecological systems. *Nature* 313: 355–358.

Steele JH (1998) Regime shifts in marine ecosystems. *Ecological Applications* 8(supplement): S33–S36.

Sterner RW and Elser JJ (2002) *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton: Princeton University Press.

Subramaniam A, Brown CW, Hood RR, Carpenter EJ, and Capone DG (2002) Detecting *Trichodesmium* blooms in SeaWiFS imagery. *Deep-Sea Research II* 49: 107–121.

Thompson JR, Pacocha S, Pharino C, *et al.* (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science* 307: 1311–1313.

Tittensor DP, Mora C, Jetz W, *et al.* (2010) Global patterns and predictors of marine biodiversity across taxa. *Nature* 466: 1098–1101.

Weithoff G (2003) The concepts of 'plant functional types' and 'functional diversity' in lake phytoplankton – a new understanding of phytoplankton ecology? *Freshwater Biology* 48: 1669–1675.

West GB, Woodruff WH, and Brown JH (2002) Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proceedings of the National Academy of Sciences* 99: 2473–2478.

Woodd-Walker RS, Ward P, and Clarke A (2002) Large-scale patterns in diversity and community structure of surface water copepods from the Atlantic Ocean. *Marine Ecology Progress Series* 236: 189–203.

Yvon-Durocher G, Allen AP, Montoya JM, Trimmer M, and Woodward G (2010) The temperature dependence of the carbon cycle in aquatic systems. *In: Woodward G (ed.) Advances in Ecological Research, vol. 43, Integrative Ecology: From Molecules to Ecosystems*, pp. 267–313. Academic Press.

Mammals, Biodiversity of

Joshua R Ginsberg, Wildlife Conservation Society, Bronx, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Brachyodont Low-crowned cheek tooth commonly found in omnivorous or burrowing animals.
Cloaca The chamber into which the gut and urinary and reproductive tubes empty before exiting the body.
Diastema A large gap between incisors and premolars that is thought to permit an animal space to manipulate food with the tongue.
Fossorial A burrowing animal.
Macropod A marsupial in the family Macropodidae, the kangaroos and wallabies.
Marsupium Folds of skin that envelope the mammary glands and provide protection to infants: found in metatherian and protherian mammals.
Monophyletic Descriptive of a clade, or branch, of an evolutionary tree that has a single root, indicating that all members of that branch are descended from a common ancestor.
Phylogeography The study of the relationship between and among genetics, morphology, paleontology, and ecology to better understand the spatial aspects of evolution.
Polyphyletic Descriptive of a group of animals that, although grouped together in a phylogeny or evolutionary tree, do not share a common ancestor.
Semelparous Species that breed once and die.

Ranging in size from a shrew, which may weigh only a few grams, to the blue whale, which weighs more than 150 metric tons, mammals are found in the air, on the ground, under the earth, and in the oceans, rivers, and lakes of the world. Modern mammals are represented by 151 families divided into 1226 genera, not an impressively diverse group of animals. Throughout geological history, there have been just more than 400 families of mammals and perhaps as few as 5000 genera. Paleontologically, the structure of the mammalian middle ear is unique, with three small bones that were originally found in the jaw of more primitive vertebrates now forming the structure of the middle ear. Mammals are endothermic; that is, they regulate their internal temperature. To support endothermy, mammals have greater cellular production of energy, lungs with a large capacity, and numerous red blood cells to transport oxygen; an efficient four-chambered heart; and increased digestive efficiency. Soft-tissue characters that are diagnostic of mammals include sweat glands, hair, a folded cerebellum, an epiglottis, and a complex lung structure. Of course, mammary glands are the single most common diagnostic feature for mammals: the production of milk with which to feed and nurture young is a shared feature of all mammals.

Introduction to Mammals

Modern mammals are not an impressively diverse group of animals, at least not numerically. Mammals are an extremely well-described class of animals, with, perhaps, only Aves being better examined. Well described and relatively well studied, modern mammalian diversity is limited to 29 orders, divided into 151 families with 1226 genera and classified, at present, into approximately 4800 species (**Figure 1**). Mammals have about half the diversity of modern birds and pale in significance when compared to the more than 1 million described species of insects and the perhaps 10 million species that remain undescribed. Historical consideration does little to mediate these discrepancies in diversity: throughout all geological history, there have been just more than 400 families of mammals and perhaps 5000 genera.

What mammals lack in numeric diversity is made up in their physical and ecological variation. Ranging in size from just a few grams (e.g., shrews in the family Soricidae) to more than 150 metric tons (the blue whale, *Balaenoptera musculus*), mammals are found in the air, on the ground, under the earth, and in the oceans, rivers, and lakes of the world. Their diversity of body types, diets, and habits has captivated both the general public and scientists, taking an entirely disproportionate share of scientific study and popular examination.

What is a Mammal?

There are literally dozens of diagnostic characteristics of mammals, most of which are related to soft-tissue structure and metabolic function. For paleontological purposes, soft-tissue differentiation of mammals is of little or no value. Soft tissue is rarely preserved and its structure can only occasionally be deduced from the structure of the bones that it covers and from which it is suspended. Paleontologists, therefore, look for skeletal and cranial features that distinguish mammals from other classes. One defining character is found in the structure of the mammalian middle ear. The middle ear contains three small bones, the stapes, incus, and malleus, which were originally found in the jaw in the craniomandibular joint of more primitive vertebrates (the joint where the jaw meets the cranium). The simplification of the craniomandibular is also reflected in further simplification and solidification of cranial bones overall. This simplification and increased ossification are also reflected in the postcranial skeletal structure of mammals. Skeletal growth occurs early in the life history of mammals, with further growth of bones limited to a cartilaginous area at the end of the bone, under the articular

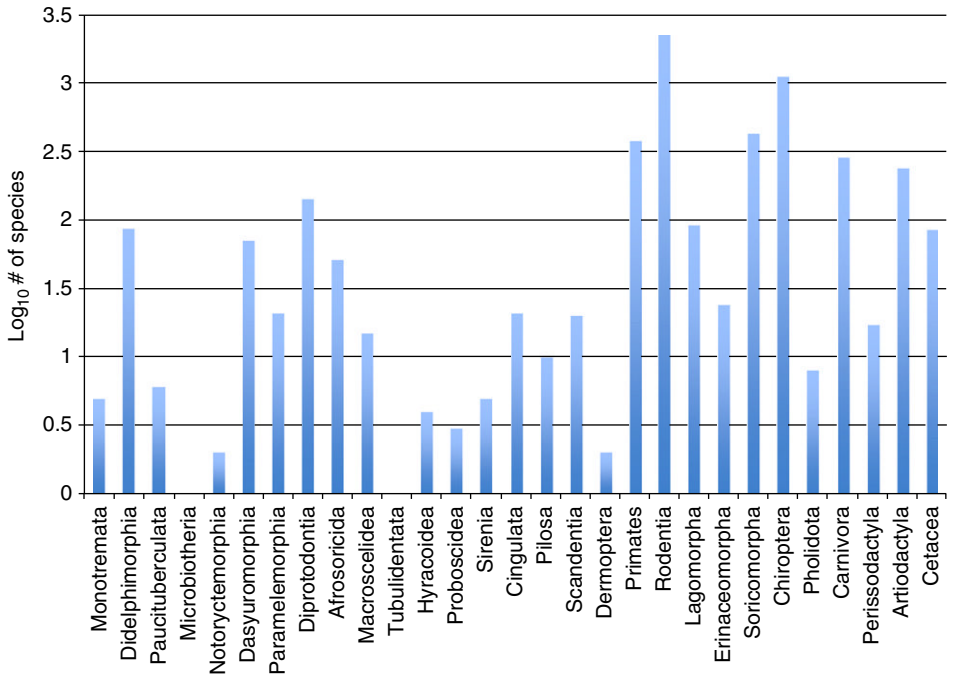


Figure 1 Species diversity of recent mammalian orders. Data derived from Vaughan TA, Ryan JM, and Czaplewski NJ (2000) *Mammalogy*, 4th edn. Philadelphia: Saunders College Publishing. All taxonomy is based on Wilson and Reeder (2005).

surface, called the epiphysis. Eventual closure and ossification of the epiphyses occur after growth is complete.

Physiologically, a critical difference between mammals and reptiles is the development of endothermy, or internal regulation of temperature. While it is unclear whether the mammal-like Therapsids were endothermic, all modern mammals exhibit endothermy, whereas extant reptiles are ectothermic, using behavioral adaptations to regulate temperature in harsh environments or allowing their temperature to fluctuate widely. The cellular and structural differences that endothermy require are all hallmarks of mammalian life: greater cellular production of energy, lungs with a large capacity, and numerous red blood cells to transport oxygen, efficient circulatory systems, and increased digestive efficiency. Soft-tissue characters that are diagnostic of mammals (many of which are related to endothermy) include sweat glands, hair, a four-chambered heart, a folded cerebellum, an epiglottis, and a complex lung structure. Of course, mammary glands are the single most common diagnostic feature for mammals: the production of milk with which to feed and nurture young is a shared feature of all mammals.

Overview of the Class Mammalia

Mammals are usually divided into three subclasses: the Allotheria, which was represented by the extinct multituberculates; the Prototheria, which reached the peak of its diversity in the Mesozoic and which today is represented by a single order, Monotremata, containing the four species of echidna (Tachyglossidae) and the duck-billed platypus (*Orithorhynchus anatinus*); and the Theria, which contains two infraclasses the Metatheria (marsupials) and the Eutheria (placental mammals). Debate continues about the unique origin of the Prototheria.

Monotremata

Viviparity, or live birth of young, is so common among mammals that it is usually, wrongly, considered a defining character of the class. The five species of the order Monotremata all lay eggs. The platypus lays its eggs into a nest, similar to a bird's nest, whereas all four species of the family Tachyglossidae, the echidna, or spiny anteaters, lay their eggs directly into a marsupial-like pouch. Although this curiosity is the root of the common name for the monotremes – egg-laying mammals – the egg is actually a rather insignificant aspect of the monotreme's life history. Incubation is brief, under 2 weeks, following which monotreme development does not differ significantly from that of other mammals. Clearly mammals, they nurture their young with milk that is expressed from mammary glands that lack nipples. Like all mammals, they are endothermic, have hair, possess a single jaw bone, and have the diagnostic three-bone middle ear structure.

Divergence of the monotremes from other mammals occurred approximately 175 million years ago early in mammalian history. Fossil monotremes have only been found from Australasia, and all extant species share this distribution. Monotremes appear to be extremely primitive in their reproductive habits, with not only an egg-laying habit but also a single opening, or cloaca, into which both the excretory and reproductive tracts exit. The cloaca (or single exit) gives the order its name.

Theria

The therian mammals all share the derived character of live birth of their young. The subclass Theria is divided into

two infraclasses: metatherians and eutherians. Metatherians are usually called marsupials because of the pouch, or 'marsupium,' that many members of the subclass possess. A far more diverse order than Monotremata, there are 331 described species of marsupials, which are categorized into 21 families. Metatherians and eutherian mammals probably diverged at about the same time as the monotremes, approximately 100 million years b.p. Since that time, the two major groups have had a divergent, and sometimes convergent, evolutionary history.

Metatheria: Marsupials

In recent times, metatherian mammals have only been found in the Australian subcontinent, the adjacent large islands of New Guinea (the island consisting of Papua New Guinea and the Indonesian province of West Papua) and Tasmania, and the Neotropics. This distribution is more restricted than that given by fossil evidence. Arising in the Mesozoic, about 100 million years ago, marsupials are first found in North and South America. The metatherians made their first appearance in the fossil beds of the Eocene of Europe, 50 million years ago, and it is thought they arrived in Asia, and Australasia, some time thereafter. The fossil evidence for this period from these regions is poor, and some have suggested that this, combined with the present distribution of marsupials, strongly implies an Antarctic origin.

In the New World, the recent marsupial fauna has not developed significant diversity in competition with placentals. The New World family Didelphidae is represented by 87 recent species in 17 genera. Although found across North, Central, and South America, there is relatively little variation in size, with the smallest member of the family the size of a mouse and the largest only the size of a large cat. The most common species in the Americas is the Virginia, or common, opossum (*Didelphis virginiana*), which is widespread and extremely successful in adapting to human presence. The earliest members of this genus are found in South America and it appears that they migrated to North America just more than a million years ago after the establishment of the Central American land bridge. Although marsupials, for the most part, did poorly when they invaded North America, Didelphis is a clear exception to this rule.

In South America during the late Miocene and Pliocene epochs (approximately 5–10 million years b.p.), a diverse radiation of marsupials occurred. Convergent evolution produced a suite of marsupials that resembled contemporaneous placental mammals and included forms similar to large felids (in the family Thylacoleonidae). The most diverse family of carnivorous, South American marsupials was the Borhyaenidae, often called the family of dog-like marsupials because of the general resemblance of some species to canids (e.g., Lycopsis). Borhyaenids persisted only until the late Pliocene. Other families of South American fossil marsupials included rodent-like, mole-like, and rhino-like animals and perhaps the best known, the Thylacosmilidae, or the 'false' saber-tooth tigers (Thylacosmilus).

In the Australian region, until recent times, the metatherians have been free from competition with placental mammals, with the exception of bats and some muroid rodents (rats and mice). Here, the metatheria have undergone a remarkable and fascinating radiation. Many species resemble, in basic form, the eutherian moles, squirrels, and small carnivores. These species tend to be relatively undifferentiated and follow a relatively similar body form within groups. In some cases, however, marsupials in Australia have evolved into forms that little resemble in structure their ecological replacements in areas where placental mammals dominate. Whereas most grazing placental mammals are, to a greater or lesser extent, relatively similar in their body plan, the kangaroos and wallabies (Macropodidae) represent a distinctly different solution to the design of a grazer. The origin of the large ricochetal hind feet and vestigial, but highly flexible, forearms suggests an arboreal origin for the family.

Eutheria: Placental mammals

The greatest diversity of mammals is seen in the eutherians. Both through geological time and in the present, eutherians are the most diverse of the three main branches of mammalian evolution. Representing nearly 95% of the extant species of mammals, eutherians have shown a remarkable variety of forms and adaptations. Fully aquatic forms are represented by the 11 families and 84 species of cetaceans. The 18 families of Chiroptera, representing more than 1000 species, have a near monopoly on the nocturnal skies and are keystone pollinators in tropical forests. But by far the most diverse order is Rodentia, with 33 mammalian families (approximately 22% of the total) and more than 2200 species, nearly half the known mammals alive today. A more thorough discussion of eutherian diversity is given below (*see* Placentals for a description of each Order).

One would expect that the defining character for eutherian mammals is the well-developed placenta, from which the infraclass derives its name. The use of the term is unfortunate as marsupials possess one of four types of more primitive, but nonetheless well-developed, placenta, and placenta-like structures are seen in some fish and reptiles. In all mammals, the placenta serves to provide exchange of nutrients and waste products and acts as a respiratory organ for the developing fetus. Increased complexity of the placenta found in eutherian mammals allows for more efficient transfer and for added functions, including endocrine production, including progesterone, which helps sustain pregnancy.

Differentiating Between Placental and Marsupial Mammals

Cranial and Skeletal Differences

Differentiation of marsupial and placental mammals is easiest in the examination of soft tissues; nonetheless, there are a set of diagnostic characters that distinguish the metatherian and eutherian mammals. Whereas the palates of placental mammals are smooth, those of marsupials have a series of diagnostic holes or vacuities. Marsupials also differ in their dental morphology: in contrast to the serial replacement of deciduous teeth with permanent teeth found in placentals, only a single tooth is replaced in each jaw. With the exception of the third premolar, all premolar, canine, and incisor deciduous teeth are resorbed before they erupt in the jaw. Only a single set of molars forms, erupting sequentially. Recent fossil evidence from the Mesozoic marsupial Alphadon shows that this pattern

of tooth replacement and development is a derived character and suggests that the trait is ancestral to marsupials.

A further primitive character common to the Monotremata and Metatheria, but thought to be lost in the eutherian mammals, is the epipubic bones, which extend forward from the pubic bone in both sexes. Hypotheses about the function of the epipubic bones have been various. Initial speculation was that the structure provided support for the attached young found in marsupials or support for the marsupium, in which they are often contained. Others have suggested it was involved with locomotion, although more recently it has been proposed that epipubic bones provided a place for the attachment of abdominal muscles, thus allowing for greater expansion of the abdominal cavity, increasing the function of the diaphragm and improving efficiency of respiration. Recent paleontological finds show that epipubic bones were present in Cretaceous eutherian mammals. Because the loss of epipubic bones is associated with prolonged gestation, these findings have been interpreted as evidence that the complex suite of adaptations that typify placental mammals were derived later in the evolution of eutherians.

Reproduction

Whereas placental mammals tend to have long gestation periods and relatively short periods of lactation, gestation in the marsupials is always relatively brief and is followed by a much longer period of lactation. Typically, neonatal marsupials are extremely small, only a few grams, and show very incomplete development. The exception to this generality is seen in the jaw and forearm structure of the neonate: the nearly embryonic young need strong arms to navigate up the mother's midline to the teat and well-developed jaws with which to form an immovable attachment to the teat once arriving. A greater proportion of development occurs outside the mother's uterus after the neonate has attached to a teat. In approximately half the marsupial species, the teats are found inside a marsupium. By the time the neonate leaves the pouch, or detaches from the teat, it weighs more or less the same as a placental mammal of an equivalent adult size.

Diagnostic differences in the structure of the reproductive tract of female marsupials and in the development of the embryo may explain many of the life history differences between eutherian and metatherian mammals. While eutherian mammals form a complex placenta that nourishes the embryo, marsupials retain an eggshell membrane with a simple yolk sac to nourish the young. This arrangement limits the nutrients that a mother can transfer to her embryos.

The size of a neonatal marsupial is also limited by the structure of the female reproductive tract. In eutherian mammals, the reproductive tract is arranged linearly (**Figure 2**): the vagina leads directly into the uterus, with two ureters transporting urine from the bladder to the urethra. Growth of the fetus is essentially constrained only by the size and elasticity of the vaginal canal. Given that the spotted hyena gives birth through a vagina that has evolved to look like a pseudopenis, this constraint does not appear to be too great. In contrast, in marsupials, the structure of the vagina and uterus is like a double jug handle (**Figure 2**). Just above the urethra, two lateral vaginas form loops on either side of a pseudovaginal canal. The two vaginal loops rejoin and feed into a bipartite uterus. The ureters pass inside these vaginal loops. Because birth takes place through the lateral vagina, placement of the ureter inside this loop may limit the size of the neonate.

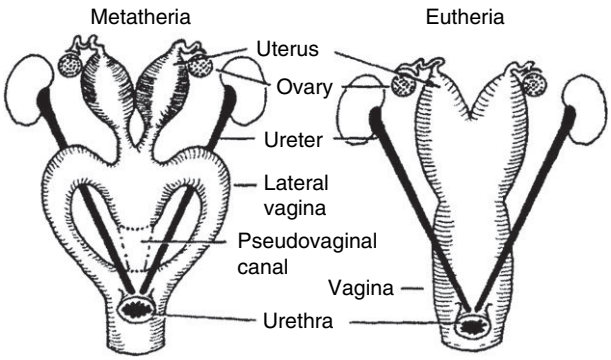


Figure 2 Comparison of eutharian and metatherian reproductive tracts. Reprinted from Sharman GB (1970) Reproductive physiology of marsupials. *Science* 167: 1221–1228, with permission from American Association for the Advancement of Science.

Mammalian Phylogeny

Early History

Mammals have their origins deep in geological history. Just more than 200 million years ago, in the late Triassic period, primitive cynodont mammals evolved from their mammal-like reptile progenitors, the Therapsids. The Therapsids, whose members dominated the terrestrial landscape during the Triassic, faded into insignificance by the end of the Triassic/Jurassic boundary, leaving the stage of the late Mesozoic era open to the radiation of great sauropods, the dinosaurs. For 140 million years, dinosaurs filled most terrestrial (and many aerial) ecological roles, with mammals for the most part relegated to small, terrestrial rodent-like forms.

The earliest evidence of true mammals occurs in the late Triassic, although by this time, mammals were found worldwide, suggesting a somewhat earlier divergence from the mammal-like cynodonts. Poor fossil records from many areas leave gaps, which slowly are being filled. Although the Mesozoic mammals were once thought to be lacking diversity, recent finds suggest this was not the case. Fossil evidence suggests they were never abundant and rarely showed any great size – the first mammal weighing more than 1 kg does not appear in the fossil record until the early Cretaceous. Yet a variety of unusual forms evolved in the Mesozoic, including Symetrodonta, which are characterized by well-developed, triangular molars, the predatory Triconadonta, and the omnivorous and herbivorous members of the Multituberculata.

The great radiation of modern mammals began in the mid-Cenozoic era, in the late Cretaceous period, approximately 100 million years ago. More rapid evolution of larger body sized mammals did not begin until the mid-Cenozoic, beginning about 65 million years ago during the Paleocene epoch of the Tertiary. Following this, mammalian radiation accelerated.

Ordinal Diversification

Structure of the Mammalian Evolutionary Tree

The ordinal diversification of the mammals was more or less completed in the late Eocene, 50 million years ago, and has been relatively stable since that time. Of the 32 major orders that were present 50 million years ago, 20 are extant. Reconstruction of evolutionary paths of ordinal diversification can be made using morphological analysis of modern or fossil taxa, molecular analysis (limited to species from which deoxyribonucleic acid (DNA) can be extracted), or some combination thereof. Extant species have often lost defining skeletal and cranial characters present in earlier fossil forms but have the advantage of soft tissue being available for study. Fossils offer a long time frame, but significant gaps in the fossil record and the absence of soft tissue for comparison may be problematic. Molecular genetic studies have become increasingly sophisticated but suffer from their own inaccuracies and, of course, are relatively useless for independently reconstructing phylogenies from taxa that have gone extinct, leaving no modern (or DNA-bearing fossil or subfossil) forms.

Molecular data have revolutionized our understanding of the phylogenic structure of the major mammalian orders and have brought us closer to resolving significant questions about the exact relationship of several orders. Recent studies group mammals into four different clades: Afrotheria (which includes golden moles, tenrecs, elephants, and dugongs), Xenarthra (which includes armadillos and anteaters), Laurasiatheria (which includes whales, cows, shrews, bats, and dogs), and Eurarchontoglires (which includes rodents, primates, and flying lemurs). These new connections are surprising, given that, with the exception of the Xenarthra clade, the animals within these groups have extremely different morphology. Up until recently, the phylogenies of this 'supertree' was much less clear, with the availability of data being the most limiting factor. However, studies performed in the past decade have been able to analyze more diverse molecular datasets with a higher number of species. A theory currently exists that plate tectonics played an important role in the early evolutionary history of placental mammals, based on the numerous occurrences of convergent evolution found in the new tree's results.

A good example of a problem potentially solved by molecular analysis is the placement of the Hyracoidea, the hyraxes. Both protein sequence data and morphological studies suggest a close affiliation of the hyraxes with the Proboscidea (elephants) and the Sirenia (dugongs and manatees). Yet other morphological analysis suggests that hyraxes and perissodactyls (rhinos, tapirs, and equids) share a common lineage. Gene sequence data, although not without problems of its own, has now comfortably placed hyraxes with the elephants and dugongs in the superordinal clade Afrotheria, separate from the perissodactyls (which are now in the clade Laurasiatheria).

Timing of Ordinal Diversification

While scientists agree that there was a period of rapid diversification of higher orders of mammals, the exact timing of this event is still hotly debated. At the root of this debate are the differences in timing derived from molecular and fossil evidence. The data supporting the timing of events presented above, with the major ordinal diversification completed in the early Tertiary, are derived from the near lack of any fossil evidence for the existence of modern orders of mammals before the Cretaceous/Tertiary boundary 65 million years ago.

Advances in molecular biology have permitted more sophisticated molecular analyses of ordinal diversification of both birds (which are purported to have a similar timing of ordinal diversification) and mammals. These analyses, which employ a number of different genetic 'molecular clocks,' all point to an early ordinal diversification of birds and mammals sometime in the middle to late Cretaceous, approximately 90–100 million years b.p. That the estimated timing of diversification is similar for birds and mammals supports this interpretation. However, of greater interest is the coincidence of the timing of ordinal divergence and the fragmentation of emergent land areas at about the same time.

By the early Cretaceous, the two main land masses formed by the breakup of Pangea, Gondwanaland, and Laurasia had further subdivided into seven continents. By the late Cretaceous, 70 million years ago, 11 continents had appeared, the peak of land mass isolation and diversity. Earlier explanations had suggested that the rapid ordinal diversity resulted as mammals (and birds) radiated into a diversity of empty niches on the new continents. The molecular data suggest, however, that the breakup of land masses itself provided the opportunity for diversification.

If the molecular data are correct, fossil evidence for two dozen orders is 'missing' for upward of 30 million years, and by some estimates 60 million years, of stratigraphic history. The real question asked by many paleontologists, therefore, is, "How likely is it that an increasingly good fossil record could be devoid of evidence for such diversification?" The answer appears to be, not likely. Using standard models to examine the probability of not finding a species in the fossil record, scientists estimate that, even under the most severe conditions, there is a 2% chance of not finding earlier fossils if, indeed, ordinal diversification occurred before the Cretaceous/Tertiary boundary. One possible reconciliation of these two viewpoints would be the Cretaceous splitting of a few lineages, thus giving rise to earlier ordinal diversification, followed by a rapid radiation during the Paleocene.

Whatever the timing of the ordinal diversification of mammals, by the mid-Tertiary, continental drift resulted in a relatively large number of isolated land masses, evolutionary caldrons in which mammalian diversification occurred. Peak continental isolation occurred at the Cretaceous/Tertiary boundary when landmasses, which in modern times are joined, were isolated. This included present-day continents (e.g., South America and Africa) and subcontinents (e.g., India). Even landmasses that were joined, such as Asia and Europe, exhibited greater isolation as a result of landscape features. Diversification accelerated about 10 million years after this isolation was established, but recent work on the 'supertree' structure of mammalian radiations suggests a more complex history.

Island continents, in and of themselves, would have likely produced greater mammalian diversity over time. However, in addition to the isolation of continents, the Tertiary was a time of greater differentiation and seasonality of climates. Much

evidence exists for this, from diversification of Tertiary flora (particularly in the late Tertiary) to developments such as the Antarctic ice cap, which formed in the Miocene. Continental drift, with landmasses moving toward the poles, combined with concomitant changes in ocean currents resulted in increasingly temperate environments. For the homeothermic mammals, this diversification of ecosystems provided a unique opportunity for radiation. In recent times, geologically isolated continents and large islands, such as Madagascar and Australia, continue to show unique assemblages of mammals that are typified by high levels of endemism and diversification of a small number of families.

Phylogeny and Biodiversity: Patterns of Ordinal Diversification

Mammalian generic diversity appears to be rising exponentially throughout the most recent geological period, the Cenozoic. Whereas there is some fluctuation in the levels of diversity, the progressive increase in diversity is notable. This pattern may be driven, in part, by an increasingly heterogeneous environment. However, for the most part, the exponential rise is probably an artifact of an increasingly well-represented fossil record. Mammal diversity probably peaked during the mid-Tertiary, although patterns of diversification (and simplification through extinction) vary from continent to continent. The following sections will briefly describe patterns of diversity in the extant orders of mammals.

Monotremes

The five extant species of the Montremata, found only in Australasia, all represent highly specialized representatives of an order that had an early branching from all other mammalian orders. Living monotremes are toothless, with the exception of young platypuses, a trait that has not made reconstruction of fossil history a simple task. Early fossil remains are known from Australia. A single fossil found in southern Argentina, dated to about 60 million years b.p., represents the only occurrence outside of Australasia.

Monotremes possess what are surely primitive mammalian characters such as egg laying and a cloacal reproductive/excretory structure. Although these primitive characters reflect an early divergence, monotremes are not primitive in all their characters. For instance, all five extant species of montremes share complex and well-developed nervous systems. The most unusual sensory adaptation is the development of electroreceptors across the platypus' bill and in the tip of the echidnas, producing snout or beak. Unique among mammals, electroreception in monotremes is complemented by a complex and large brain.

Monotremes are also the only mammals that have the ability to deliver venom as a defensive measure. Although venom glands are vestigial in the echidnas, in the male platypus a well-developed venom gland is found in the hind legs. More active during the mating season, the venom is assumed to be a defense against intruding males. The venom is toxic to humans and lethal to small mammals.

The sole surviving representative of the family Ornithorhyncidae is the platypus, *O. anatinus*, found only in eastern Australia and Tasmania. The morphology of the species is so distinctive that after early examination of skins sent to London, they were pronounced a fraud. Weighing around a kilogram and with an overall length of 45 cm, the platypus has a broad, duck-shaped bill. The animal is covered in short, dense fur and has webbed feet with claws for digging. The animal is highly aquatic and solitary outside of the breeding season. Females will expand their burrows before breeding and, after mating, will 'gestate' one or two eggs for 10–12 days. After laying, the egg is incubated for a further 12 days. Lactation is protracted, approximately 4 months, but the young becomes independent soon after weaning.

The Tachyglossidae family is represented by two extant genera, Tachyglossus and Zaglossus. Zaglossus is endemic to the mountains of New Guinea and feeds primarily on earthworms and arthropods; it includes three species of long-beaked echidnas, or spiny anteaters: *Zaglossus bruijni*, *Zaglossus attebroughi*, and *Zaglossus bartoni*. Tachyglossus, with *Tachyglossus aculeatus* (short-beaked echidna) as its only species, has a distribution across much of Australasia, found wherever its primary prey, ants and termites, are abundant. Introduction of the dingo 10,000 years ago probably has reduced the range of Tachyglossus, much as it has been thought to have a similar effect on the platypus. The echidnas are larger than the platypus, with Zaglossus weighing up to 10 kg and Tachyglossus reaching about half that weight. The four species have thin protruding snouts and thick coats of dark brown fur interspersed with sharp quill-like spines. Patterns of reproduction differ from those of the platypus. The egg is deposited into a pouchlike structure, where it hatches 10 days later. The young stays within the pouch for 2 months and then nurses outside the pouch for a further 3 months.

Marsupials

Extant marsupials are grouped into seven orders, with 21 families, 92 genera, and 331 species (**Table 1**). In a pattern commonly seen among mammals, some lineages contain disproportionately more families, genera, and species than others. For instance, almost half of all extant marsupial families are in the order Diprodontia, whereas four orders are represented by only a single family. The distributions of genera within families, and species within genera, are also highly skewed. Three orders of marsupials are found only in the New World, whereas the remaining four orders are found in Australasia. The Old World continents of Africa, Asia, and Europe have no extant marsupial representation.

Didelimorphia

The American possums, or didelphid marsupials, represented by a single modern family, are found throughout South America, with the Virginia opossum (*D. virginiana*) found widely over eastern, central, and southern North America. A generalist capable of living on garbage, the Virginia opossum has shown a remarkable range expansion in the past two centuries, expanding its range from the southeastern US north into the central states and the southernmost fringes of eastern

Table 1 Diversity of recent Marsupials

Order	*Common name(s)*	*Number of families*	*Number of genera*	*Number of species*
Didelimorphia	American opossum	1	15	63
Paucituberculata	Rat opossum	1	3	6
Microbiotheria	Minito del monte	1	1	1
Dasyuromorphia	Thylacine, numbat, quoll, antechinus, Tasmanian devil	3	17	63
Peramelemorphia	Bandicoots	2	8	21
Notoryctemorphia	Marsupial mole	1	1	2
Diprodontia	Koala, kangaroo, wombat, wallaby, glider, feathertail, cuscus, phalanger	10	35	117
Totals		21	83	276

Canada and along the west coast up to and just over the Canadian border. In the past 50 years alone, the species has extended its range over an area of two million square kilometers. This expansion has been facilitated by the opossum's extreme potential fecundity. A female may give birth to up to 60 very altricial young, up to a quarter of which may successfully attach to one of her 15 teats.

Despite ranging in size from a small mouse to a large cat (25 g to 5 kg), most didelphids show a clear set of similarities. Arboreal habits are common in the Didelphidae as is reflected in the general morphology of family members. Hands and feet are well developed, with an opposable big toe on a hind foot that has five toes. Many species also have prehensile tails.

The dentition of most didelphids is relatively unspecialized, with 50 teeth found in most species. The family is dominated by generalists, many of which have wide geographic ranges. However, some of the more remarkable specialists are also found in the didelphids. The yapok, or water opossum (*Chironectes minimus*), is the only truly aquatic species, with a fully webbed hind foot and a carnivorous diet consisting of a variety of crustaceans, fish, and amphibians.

Paucituberculata

Paucituberculata is also represented by a single Recent family, the Caenolestidae. All six extant species, placed in three genera of shrewlike marsupials, are found only at higher altitudes in the Andean cordillera, from Ecuador south to Chile. Capable of killing small vertebrates with their sharp teeth, the caenolestids usually subsist on invertebrates, particularly earthworms. Weighing under 20 g, the shrew opossums are nocturnal omnivores, using their long snouts and tactile whiskers, or vibrissae, to locate and kill prey. The extant paucituberculates are represented by a much more diverse fossil fauna. In the Miocene, seven genera of caenolestids were represented by a large number of species with widespread geographic distribution and a relatively large diversity of life forms, including specialized predators.

Microbiotheria

A single recent species of Microbiotheria represents an order whose fossils have been found on three continents: South America, Australia, and Antarctica. With origins in the Paleocene, the microbiotherids are thought to have invaded South America while it was connected to Antarctica and Australia via a southern land bridge. The historical diversity of the group in South America is poorly known. The extant arboreal marsupial mouse inhabits the forests of southern Chile and is nocturnal.

Dasyuromorphia

Dasyuromorphia is known from fossil evidence in Australia in the mid-Miocene (Dasyuridae), but by this time the order already exhibited a relatively large diversity of morphotypes. Extant members of the order are represented by three families and 71 species, more than 20% of all recent marsupials. All but two of these species can be found in the family Dasyuridae.

In historical times, the Thylacinidae were represented by a single species, the thylacine, or *Thylacinus cynocephalus*, also called the Tasmanian wolf or Tasmanian tiger. This species has been considered extinct since the last specimen died in captivity in 1936. Thought to have had a wide distribution across Australia, introduction of the feral domestic dog, or dingo, by Polynesians approximately 10,000 years ago probably led to its extinction in Australia. The cause of extinction of the relictual population in Tasmania is debated, although it was most likely the result of a number of interacting factors common to human predator extinctions: habitat loss, disease, and declining prey supplanted by sheep farmers, who hunted the thylacine because their livestock was perceived to have been threatened by predation. Although it is most likely extinct, tracks and sightings continue to be reported, thus its place in the order remains (Wilson and Reeder, 2005).

The Myrmecobiidae consists of a single, monotypic genus, represented by the numbat, or banded anteater. Weighing approximately 400–500 g, the numbat resembles a small grayish-red squirrel, with a long bushy tail, a half dozen lateral white stripes circling its abdomen, and a prominent protruding snout. As its family name would suggest, however, the numbat is a termite specialist. Like other termite specialists (e.g., the bat-eared fox of Africa), the historical distribution of the species was dictated by the occurrence of its primary prey. Although termites are found across southern and central Australia, today the numbat persists in small fragments of Eucalyptus forest and woodlands (hardly core habitat), its last refuges from agriculture and predation by the introduced red fox (*Vulpes vulpes*).

The 69 species of the family Dasyuridae are often called the marsupial carnivores despite the fact that most members of

the family are insectivores. Their dentition, although similar to that of most marsupials, is similar to the dentition of many carnivores, characterized by bladelike incisors, large, sharp canines, and upper molars modified with large, sharp cusps. Found across Australia and New Guinea, dasyurids are usually small, with most species weighing under 50 g. The family occupies a wide diversity of terrestrial niches, with the smallest species convergent on shrews (family Soricidae) found in Asia, Europe, and North America.

Some of the dasyurids, particularly species in the genus Antechinus, have remarkable life histories. For instance, Stuart's marsupial mouse, *Antechinus stuartii*, has a semelparous life cycle more similar to that of many insects than to that of a mammal. The species is usually arboreal but forages terrestrially for invertebrate prey. Mating is highly seasonal, with males competing aggressively for mates. Females produce an enormous litter for their size (10–12 offspring). The young forage independently at 3 months and are sexually mature by 10 months. What is most remarkable is that after mating, the males die, whereas females rarely survive to breed in a second breeding season. Similar patterns of reproduction are seen in other Antechinus species.

The two largest species of dasyurids rightfully deserve the name marsupial carnivore. The tiger quoll, *Dasyurus maculatus*, may weigh up to 7 kg (males) whereas the largest extant species of Australian carnivore, the Tasmanian devil (*Sarcophilus laniarius*), can reach the size of a medium-sized dog, weighing up to 9 kg. Like the thylacine, the Tasmanian devil once had a widespread distribution across Australia but is now extinct on the continent. It persists in Tasmania, perhaps because it is primarily a scavenger, not a predator, and thus is more catholic in its dietary requirements and less of a perceived threat to farmers. Both quolls and Tasmanian devils are usually solitary; however, aggregations of up to 20 Tasmanian devils have been observed around carcasses.

Peramelemorphia

Found from the deserts of Australia to the rain forests of New Guinea, the 21 species of bandicoots and bilbies are placed in eight genera and three families, the Peramelidae, the Chaeropodidae, and the Thylacomyidae. The bandicoots, found in Permelidae and Chaeropodidae, are insectivorous or omnivorous and for the most part are ecological counterparts of insectivores. Most species forage by digging insects out of the ground. They have small, sharp teeth, and ecological studies show that, despite feeding specialization for insects, they are omnivorous and opportunistic in their diet choice. Thylacomyidae includes two species, the greater and lesser bilby. Similar to bandicoots, bilbies can be distinguished by their longer ears and burrowing abilities that allow them to dig extensive tunnel systems. While the lesser is most likely extinct, the greater is now exclusively found in central Australia (Wilson and Reeder, 2005).

Bandicoots vary in size from the rodentlike mouse bandicoot (*Microperoryctes murina*) at 15–18 cm up to the giant bandicoot, *Peroryctes broadbenti*, at 45–50 cm and up to 5.5 kg. In recent times, the evolution of bandicoots has taken separate paths in Australia and New Guinea. Majority of the five species of echymiperas, or spiny bandicoots, are found on New Guinea, with a few populations on small islands in Indonesia, and one with a relictual distribution in the rain forests of Cape York in northern Australia. Little is known about the fossil history of this family, as New Guinea is all but lacking in fossil record. On New Guinea, species occur at different altitudes, with the most species occurring in mid-altitude regions of 1000–1500 m above sea level.

Like perrisodactyls and artiodactyls, peramelid bandicoots show relatively extreme reduction in their limb morphology. Many peramelids have digitigrade elongated hind limbs, with expanded growth of the fourth digit, and lateral digits showing varying ranges of reduction. Reduction in limb digits suggests strong selection for running in the bandicoots. Whereas the bandicoots are remarkably cursorial, some of the elaboration of the hind limb may also have evolved for improved digging and burrowing.

Notoryctemorphia

The only Australian marsupials adapted for a fossorial existence are the two species of the order Notoryctemorphia, *Notoryctes caurinus* and *Notoryctes typhlops* (northern and southern marsupial mole, respectively). The marsupial mole has a pale, golden yellow coat and a long, tubular aspect about 15 cm long, weighing 35–40 g on average. They have rudimentary, almost vestigial eyes, compressed neck vertebrae, and stubby, almost vestigial forelimbs with two well-developed claws on the third and fourth digits. Other digits have essentially disappeared. The hind limbs are flattened, with only three small claws, and are used to expel dirt from a burrow. Little is known about the behavior or ecology of these species, although they are believed to be solitary and have been observed to eat burrowing insects and their larvae.

Diprodontia

This order includes both the largest and smallest marsupials. The smallest member of the family Burramyidae, or pygmy possums, weighs 6–8 g (*Cercartetus lepidus*). Known only from fossil remains until 1996, the family is now known to have five extant species placed in two genera. At the other extreme of the diprodont scale are the members of the family Macropodidae, the kangaroos and wallabies. Whereas the smallest macropod, the rock wallaby (*Petrogale burbidgei*), weighs 1 kg, the great gray kangaroo may reach 90 kg and stand flat-footed at just more than 2 m. Diprodontia also shows the greatest morphological variation on any extant marsupial order. While it is beyond the scope of this article to review all 143 species in the 11 families that make up the diprodonts, a review of some of the more divergent taxa is informative.

The single species in the family Phascolarctidae, the koala (*Phascolarctos cinereus*), like the kangaroo, is emblematic of the Australian continent. Highly selective in its diet, the koala subsists entirely on the leaves of a few species of smooth-barked eucalyptus trees. Unlike many marsupials, the koala is slow to reproduce, with a single young born to a female each year and the time of dependency between mother and young stretching to up to 1 year. Sexual maturity is not reached until age three or four.

The feathertail glider (*Acrobates pygmaeus*) and the feathertail possum (*Distoechurus pennatus*) are the only two species in the aptly named family Acrobatidae. Found only in New

Guinea and with no known fossil record, both species have long stiff hairs on either side of their tails, from which they derive their name. The gliders, which weigh a mere 10–15 g, have membranes between the elbows and knees, further aiding gliding flight. The membrane is lacking in the possum, which weighs up to 50 g.

Gliding as a strategy appears to have evolved twice in the diprodonts. Six species of lesser glider, three species of triok, and the striped and Leadbetter's possum make up the family Petauridae. Found from Tasmania to New Guinea, the petaurids resemble small squirrels. Weighing from 100 to 700 g, they have long bushy tails and stripes down their backs. The striped possum has black and white dorsoventral stripes along the back, resembling a North American skunk in coloration.

The six gliders bear a remarkable resemblance to flying squirrels of the genus Glaucomys in the Rodentia order. A membrane that stretches between their wrists and ankles provides an almost rectangular, kitelike gliding surface that these animals use to move from tree to tree. Whereas the striped possum is an insectivore, all other members of the family specialize on eating plant exudate (sap and gum) as well as insects. The sugar glider (*Petaurus breviceps*) is particularly specialized, concentrating its efforts on the exudate from a single species of Eucalyptus.

Resembling fossorial rodents, the three extant species of wombat have short, muscular limbs, stocky bodies, and broad long claws. They are large animals, weighing up to 40 kg. Prodigious diggers, they live in complex burrow systems, which caused them to be called marsupial badgers on their first discovery. Unique among marsupials, their teeth are ever-growing like those of rodents and lagomorphs. Wombats are herbivores, concentrating their diet on tussock grasses. Wombats are found only in southern Australia and Tasmania, with the northern hairy-nosed wombat (*Lasiorhinus krefftii*) found in one isolated population of 70 individuals living on 3 km^2 in central Queensland. Wombats have a long fossil history in Australia, with two much larger species occurring throughout the Pleistocene and going extinct about 10,000 years ago, coincident with the arrival of humans and dogs.

The best known family of diprodonts, and probably the most widely recognized family of marsupials, are the 11 genera and 65 recent species of macropods, the kangaroos and wallabies. Ranging in size from 1 to 90 kg, all species share the same basic body plan. Forelimbs are reduced and have five toes, all bearing small claws. The hind limbs are elongated, and, as in the bandicoots, on the hind leg all but the fourth digit has been reduced or lost. The large hind foot allows macropods to take a bipedal stance.

The limbs are highly adapted for ricochetal or hopping locomotion that, because of the limb structure, is highly energy efficient. Each time the kangaroo bounces down, the large elastic tendons in the foot are stretched taut like a rubber band, storing energy from the previous bounce. As the foot is released, the tendons snap back and push the foot off the ground, throwing the kangaroo into the air. On landing, the tendons are stretched taut again, completing the cycle.

In addition to the remarkable limb adaptations, macropods have a digestive system remarkably similar to that of the eutherian ruminant herbivores they ecologically replace. Both groups have evolved segmented stomachs, with a true chambered stomach in the ruminants and a divided, three-part stomach in macropods. This segmentation allows for the acquisition of foregut fermentation, in which symbiotic bacteria digest the tough outer cell walls of the plant materials that the animal eats, thus releasing nutrients from within the cell, providing a digestible form of cellulose (which is broken down by the bacteria), and nutrients from the bacteria themselves as they are absorbed when they pass through the gut.

Macropods have a wide geographic range and habits. Whereas macropods are usually thought of as long-distance, open-plains grazers, in New Guinea tree kangaroos in the genus Dendrolagus have adopted an arboreal habit. Although these animals still spend much of their time on the ground foraging for fallen fruits, they are agile climbers and use their large hind feet to propel them from tree to tree. Many of the smaller wallabies are also habitat specialists, such as the 16 species of rock wallabies (Petrogale sp.).

Placentals

Placentals dominate both the fossil and recent history of the Mammalia. With three times number of orders of marsupials (21 vs 7), placentals have shown remarkable diversification at all lower levels. Recent species are grouped into 21 orders with 130 families, 1134 genera, and approximately 4480 species (**Table 2**). As in the marsupials, some lineages contain disproportionately more families, genera, and species than others. With a fossil record stretching back into the Cretaceous, eutherian mammals have a long evolutionary history derived from marsupials and monotremes. Found on all continents, the placentals show remarkable niche diversification, occupying a wide variety of terrestrial, aerial, and aquatic habitats and subsisting on diets ranging from the true omnivorous habits required to live on the varied detritus of human populations to a single group of grasses (pandas and bamboo) or termites.

Pilosa

The superorder Xenarthra, which includes orders Pilosa (anteaters and sloths) and Cingulata (armadillos), encompasses a group found in recent times only in the New World. Thought to have occurred in Europe, the fossil record is, at best, fragmentary. To appreciate the diversity of the Xenarthra, take a journey back into Miocene South America. The modern-day sloths of Pilosa have quite the ancestral past; giant sloths, commonly called ground sloths or sloth bears, were as huge as their name would indicate. Whereas Tertiary radiations included a variety of species exploiting numerous habitats and niches, the members of the extant families of Pilosa are relatively or highly specialized: all four species of anteaters (Mymecophagidae and Cyclopedidae) are obligate ant and termite eaters, using their long, sticky tongues to collect their food. The sloths (six species in two families) have acquired an odd but apparently effective strategy, shifting along tree branches upside down, their spines hanging in a catenary curve. Sloths are the only green mammals; however, their color is not intrinsic to their fur but is derived from algae and cyanobacteria that grow in grooves in their hair. Sloths come down to the ground only to defecate.

Table 2 Diversity of recent Eutherian mammals

Order	*Common name(s)*	*Number of families*	*Number of genera*	*Number of species*
Xenarthra	Sloths, armadillos, and anteaters	3	13	29
Insectivora	Tenrecs, moles, hedgehog, shrew, mole	7	66	429
Scandentia	Tree shrews	1	5	19
Dermoptera	Colugo	1	1	2
Chiroptera	Bats	18	178	928
Primates	Apes, gibbons, marmosets, lemurs, galagos, Old World and New World monkeys	13	60	236
Carnivora	Canids, felids, mongoose, hyena, seals, weasels, racoons, otters	11	129	271
Cetacea	Whales, dolphins, and porpoises	10	41	78
Sirenia	Sea cows, manatees, dugongs	2	2	5
Proboscidea	Elephants	1	2	2
Perissodactyla	Equids, tapir, rhino	3	6	17
Hyracoidea	Hyrax	1	3	6
Tubulidentata	Aadvark	1	1	1
Artiodactyla	Pigs, peccaries, hippos, deer, antelopes, bovids	10	81	220
Pholidota	Pangolins	1	1	7
Rodentia	Rodents	29	443	2024
Lagomorpha	Rabbits and pikas	2	13	81
Macroscelidea	Elephant shrew	1	4	15
		115	1049	4370

Cingulata

Recently split into its own order from the current superorder Xenarthra, Cingulata contains 21 species of armadillo in its only surviving family (Dasypodidae). Arguably most impressive in Miocene South America (much like its Pilosa counterpart), numerous families of armadillos, including the Glyptodontidae, occupied the landscape with lumbering creatures up to 3 m long. The giant armadillo (*Priodontest maximus*), although half that length, is the largest extant species of armadillo, weighing up to 59 kg. Armadillos are distinguished by their jointed armor and are still endemic to the Americas.

Afrosoricida

Modern insectivores show a wide range of adaptations and are found on every continent except Australia. Taxonomy of the insectivores has been problematic, not only because the order contains some of the most primitive eutherian mammals, as well those with highly derived characters, but also because the order was used as the equivalent of a taxonomic wastebasket in which problematic families were thrown. The six families and 66 genera now classified as insectivores form a monophyletic (single-root clade) with ancient origins in the late Cretaceous or Paleocene.

Afrosoricida, one of the three new orders that split the abandoned order Insectivora, includes golden moles and tenrecs. Tenrecs (family Tenrecidae) show a disjunct distribution, found on islands of the western Indian Ocean (Madagascar, Comorros) and central Africa, and have probably been linked to Africa for all of their evolution. The family, which includes 30 species, can best be described as diverse. Although only ranging in body size from a shrew to a small cat (maximum weight of a kilogram), tenrecs have undergone a remarkable radiation on Madagascar, expanding into vacant niches usually occupied by animals as diverse as otters, shrews, hedgehogs, and moles. Many species are omnivorous. The family Chrysochloridae, the golden moles, is restricted to southern Africa. The 21 species in nine genera are all fossorial and closely resemble both true moles (family Talpidae from another of the split orders, Soricomorpha) and the marsupial mole (family Notoryctidae from the marsupial order Notoryctemorphia).

Soricomorpha

Grouping the shrews and moles into a separate order, Soricomorpha is the second of the three orders that take the place of Insectivora, and is by far the largest in terms of species (428). Two families, Solenodontidae and Nesophontidae, were found in modern times only in the Caribbean. All of the Nesophontidae (also called West Indian shrews) are extinct. The four remaining Solenodon species are found in Hispaniola and, tenuously, in Cuba. Expansion of agriculture and introduction of rats, mongoose, and companion animals (dogs and cats) have led to decline of these species. Solenodon are shrewlike, have a distinctive, highly flexible snout and large hind feet, but are distinguished by an unusually large body size, weighing up to a kilogram. Fossil solenodons are known from North America from Oligocene deposits, approximately 30 million years b.p.

True moles, found in the family Talpidae, are found throughout the Northern Hemisphere and have long evolutionary histories in both Europe and North America. The talpids are more diverse than the Chrysochlorids, with a range of morphotypes including the star-nosed mole (*Condylura cristata*), which has nearly two dozen small, fleshy fingers arranged in a starburst pattern on its nose. This represents the

most bizarre elaboration of a general phenomenon found in the true moles, which have thousands of receptors arranged in a structure called the Eimer's organ on the snout. Cortical development of the brain is linked with the Eimer's organ. The Eimer's organ assists moles in navigating and foraging in their tunnel networks.

The fourth family in Soricomorpha, the Soricidae family, contains the shrews, represented by 26 genera. A very successful mammalian family, shrews are found on every continent except Australia and Antarctica. Soricidae also boasts the world's smallest living terrestrial mammal, the Etruscan shrew (*Suncus etruscus*), which weighs 1.8 g on average.

Erinaceomorphia

The last of the new forms of Insectivora, the order Erinaceomorphia contains hedgehogs and gymnures. The eight species of gymnures are found primarily in the rain forests of Southeast Asia, Gymnures are also called moonrats, and with good reason – they resemble rats much more than they do their much closer relatives, the hedgehogs. The 16 species of hedgehog are all well known for their spines, which are simply stiff hairs hollowed and supported by keratin, and their ability to roll into a ball as a defense mechanism.

Scandentia

The five genera and 20 species of tree shrews (families Tupaiidae and Ptilocercidae) are the only extant members of the Scandentia. The group is dispersed across the rain forests of southern Asia and Southeast Asia and has a relatively long fossil history (first found in the Eocene of Asia). Resembling a small squirrel, weighing 45–50 g, tree shrews have an elongated snout and a long, bushy tail. Like true shrews, many species are known to consume insects; however, the few scientific studies that have been conducted on the family have found that fruit is often the dominant component of the diet. Parental care is unusual, with the female giving birth to a small number of young (one to three) and hiding them in a nest. In a pattern more common in antelopes, the female visits the young to nurse and then leaves them lying in. Maturation is rapid, with independence from the mother at about a month and full sexual maturity at 4 months. Tree shrews are enigmatic in their phylogenetic affiliations. They have variously been assigned to the primates, flying lemurs (Dermoptera), and rabbits (Lagomorpha), with both molecular and morphological analysis producing contradictory results.

Dermoptera

Gliding, as opposed to true flying, is a mode of locomotion that has evolved independently in a number of mammalian orders. The two living species of Dermoptera (family Cynocephalidae), also called colugos or flying lemurs, are a relictual group of gliders found today only in the forests of Southeast Asia. Not closely related to lemurs, the dermopterans first appear in the fossil record in the Eocene of Thailand. Throughout their evolutionary history, only two families of colugos have been described, one North American and the other Asian. Affiliations of the order are unclear, with some paleontologists placing them as relatively close relatives to bats: they are not related to lemurs despite their common name. The gliding membrane, or patagium, stretching over the animal's entire body, connecting along both sides of the tail to the legs, from legs to arms and from the arms to the neck, makes the animal look like a kite. Despite their relatively large size, reaching a maximum size just under 2 kg, the colugos are spectacular gliders, moving more than 100 m from tree to tree in the forest to reach the trees on which they feed. The colugo diet is herbivorous but may include leaves, fruits, and flowers.

Chiroptera

Bats represent the second most diverse order of mammals with 18 families, 202 genera, and 1050 species. Two major suborders of bat have been distinguished, the Megachiroptera, also called fruit bats or flying foxes, and the Microchiroptera. Bats come as small as 1.5 g with a wingspan of 15 cm (Craseonycteridae, Kitt's hog-nosed bat) and as large as the flying foxes in the genus Teropus (Teropidae), which may weigh up to 1.5 kg and have wingspans of 2 m.

The two suborders have long, independent evolutionary histories, although the very nature of their morphology, and in particular their fine, light bone structure, makes preservation unusual. In contrast to modern diversity, sufficient fossil material has been found to describe just more than 30 genera. The lack of fossil evidence has provoked repeated controversy and called into question whether bats are monophyletic or whether the fruit bats are actually more closely related to another group (primates are postulated because of shared evolution with megachiropterans of details of their neural pathways for vision). If this were the case, the remarkable flight structures would have to have evolved twice independently. Molecular data show relatively unambiguously that the two bat groups are monophyletic.

Bats are the only mammals capable of true flight (as opposed to gliding). The wing is formed by a thin membrane that stretches across the arms, elongated fingers, and along the body, forming a diaphanous umbrella-like structure, or patagium. Like all flying animals, bats need wings that are at once light but strong. The wings are greatly reduced in weight with muscles pulled in close to the body and bones reduced in size land volume. Torsional stress is reduced by simplification of joints. The remarkable flying abilities of bats ensure a global distribution, with bats often the only mammal naturally occurring on remote islands.

Whereas birds monopolize diurnal aerial niches, bats are nocturnal specialists. In some areas where birds and bats do not co-occur, bats may become more diurnal, providing thin evidence of competitive exclusion. Flying at night, visual acuity is of relatively little value; nonetheless, all but one species of Megachiroptera rely on vision to navigate. In contrast, microchiropteran bats have evolved navigational tools that are independent of sight. Using their larynxes, microchiropterans produce extremely high frequency sound that they emit through their nose or mouth. The sound produced by bats is referred to as ultrasonic because it is above the range that humans can hear. The ultrasound bounces off both potential prey and obstacles and is received back at the large, elaborated ears of the bat. Because the bat navigates using the sound that bounces back, this form of navigation is called echolocation.

Despite their high specific diversity, patterns of reproduction (in the few species that have been studied) all are

remarkably similar. Females carry their young with them while in flight, and this constraint limits bats to producing one, or occasionally two, offspring. Paternal care of offspring is rare; hence social systems tend to be promiscuous, with few lasting bonds between males and females. Bats tend to be highly gregarious, hanging upside down in roosts containing hundreds, thousands, and in rare cases hundreds of thousands of individuals. Caves, cliffs, and the eves of large houses all provide appropriate roosting areas.

Bats are usually insectivorous, gleaning their prey while airborne, but the diversity of diets is enormous. Diet specialists cover a range of vertebrates and include species that are adapted to fish, much like a small eagle, frog eaters, blood-eating specialists (the vampire bats), nectar feeders, and two groups of specialized fruit eaters. Two families of bat dominate fruit eating: in the Old World, Pteropidae (flying foxes), and in the New World Phyllostomidae (spear-nosed bats). In the forests in which they live, fruit-eating bats are keystone species, dispersing many of the largest forest seeds and fruits. Of course, a frugivorous diet and the large aggregations in which fruit-eating bats can occur make some bats extremely unpopular with fruit farmers around the tropics.

Primates

In recent times, primates have been found on every continent and in every habitat on earth. If one excludes the most abundant and successful primate, humans (*Homo sapiens*), the range of the order is much reduced, with representatives found in the tropics and subtropics of Africa, Asia, and Latin America. One macaque species (*Barbary macaques*, *Macaca sylvanus*) is found in Europe on the island of Gibralter across narrow straits from Africa. Another macaque species, the Japanese macaque, *Macaca fuscata*, survives in a temperate environment by relying on hot springs during the winter.

With the first fossil representation in the late Cretaceous of North America, primates are an ancient order. The modern order is represented by 15 families, including the apes, gibbons, marmosets, lemurs, and galagos. Just less than half of the familial diversity is found in the suborder Strepsirhini, or prosimians, which includes the five families of lemurs, all of which are endemic to the island of Madagascar, the Lorisidae (pottos and lorises), and the Galagidae (galagos), small nocturnal monkeys of Africa often called bushbabies. Some include the Tarsiriidae as well.

The 40 species of lemurs show a remarkable range in size from 50 g (lesser mouse lemur, *Microcebus murinus*) to more than 10 kg (Indri, *Indri indri*) and a diversity of diet from generalized herbivores, insectivores, a larvae specialist (the aye-aye, the only extant species in the family Daubentonidae), and four species of bamboo specialist (Hapalemur species). Many members of this suborder are under threat, victims of the rapid and seemingly irreversible deforestation of Madagascar.

Nineteen species in three genera of galago are recognized, an increase of eight species in the last decade. This change resulted from molecular analyses that showed that a single species of galago there may be contained several cryptic species, virtually indistinguishable morphologically from one another.

The remaining primates are placed in the suborder Haplorhini, which includes the eight families of New and Old World anthropoid primates and the tarsiers. Found across the forests of Southeast Asia, the tarsiers bear a strong physical resemblance to galagos and for many years were grouped with them in the prosimians. Genetic evidence performed in the late 1990s, and a reconsideration of morphological characters, place the tarsiers firmly as distant, but distinct, relatives of the anthropoid primates.

The extant South American primates fall into four families, the Cebidae (56 species, including tamarins, marmosets, capuchins, and squirrel monkeys), the Aotidae (eight species of night monkeys), the Pithecidae (40 species, including tits, uacaris, and saki monkeys), and the Atelidae, (24 species, including howler, spider, muriqui, and wolly monkeys). Marmosets and tamarins, placed in the subfamily Callitrichinae, are unusual in that all but one species regularly produces twins. They have a monogamous or polyandrous mating system (one female, several males), and a diet that is highly omnivorous and may include fruit, seeds, insects, and small vertebrates. Some species, including the diminutive pygmy marmoset (150 g), are exudate specialists, eating the gum of trees.

Many New World primates have been described as monkeys with five legs due to the extensive use they make of their prehensile tails, an adaptation not found in Old World monkeys. Body size ranges from just more than 500 g for squirrel monkeys (Saimiri sp.) up to 12 kg for the wooly spider monkey (*Brachyteles arachnoides*). The diet is predominantly fruits and other vegetation and competition for resources is common among cebids, with larger species dominating smaller ones. As a result, many smaller species have evolved diet or behavioral specialization, including nocturnal foraging (night monkey, Aotus sp.), the ability to eat green fruit (titi monkeys, Callicebus sp.), or habitat specialization (swamp living in uakaris, Cacajao sp.).

Old World monkeys in the family Cercopithecidae dominate the primates numerically, with 132 recent species in 21 genera. Found across Asia and Africa, cercopithid monkeys include animals with divergent life history strategies, sizes, and diets. The guenons, macaques, and baboons (subfamily Cercopithecinae) range in size from 1 to 50 kg and are usually found in large, extended matrilineal groups. Many species are tree-living, but the baboons and some macaques have adopted a terrestrial or semiterrestrial existence. Primarily fruit eaters, most species are omnivorous and will supplement their diet with practically anything that they can capture and eat.

The columbine monkeys, langurs, and leaf monkeys are placed in the subfamily Colobinae. A diverse group of animals, this group tends to be more arboreal and slighter than the cercopithicines. Although the evolutionary history of this group is mainly African, modern representation is dominated by Asian forms. Diet is reflected in an unusual stomach structure, with a large, saclike upper stomach that, like the rumen of ungulates, allows for fermentation of the coarse leaves that constitute much of this group's diet.

The 14 species of gibbons (family Hylobatidae, genera Bunopithecus, Hylobates, Nomascus, and Symphalangus) are commonly called the lesser apes and are often grouped together with the true apes, or Hominidae, in the superfamily

Hominoidea. Found throughout tropical Asia, gibbons are renowned for their brachiating (or swinging) locomotion and their echoing, eerie loud vocalizations. The family is primarily monogamous, and its diet is primarily fruit, supplemented with young leaves and flowers.

The best studied group of primates is the family Hominidae. Formerly broken into two families, Pongidae (orangutans, gorillas, chimps, and bonobos) and Hominidae, data, particularly molecular genetic analyses, have made the isolation of human beings and their ancestors in their own family impossible. Although largely vegetarian (gorillas are almost exclusively folivores), some chimps have been known to actively hunt other monkeys and to kill conspecifics. Most species are extremely social, with the exception of orangutans, which tend to be solitary except when mothers are in association with young or during the mating season. Slow breeders and habitat specialists, great apes are declining in number and distribution as a result of the rampant expansion of their confamilial *H. sapiens*.

Carnivora

A diverse group of animals, the Carnivora share a key diagnostic feature. The carnivore dentition is typified by the evolution of a single pair of slicing teeth, the carnassials. Formed from the upper fourth premolar and the lower first molar, the carnassials are the key to understanding the diversity and plasticity of Carnivora. By isolating the shearing function to a single set of teeth, true carnivores have retained flexibility in their diet while allowing for specialization of shearing of meat. Molars retain their grinding function, whereas premolars in front of the carnassials are retained for grasping, crushing, or puncturing food. The canines are often elaborated into spikes of one size or another, allowing a sharp stabbing of prey. Although carnassials were critical to carnivore evolution, many extant members of the order have reverted to more generalized dentition and lost the carnassials (e.g., pandas grinding bamboo).

Modern carnivores are divided into two major groups, the Feliformia, which includes hyenas, felids, mongooses, and viverrids, and the Caniformia, a diverse group that includes seals, sea lions, walruses, canids, bears, procyonids, and the mustelids.

Pinnipeds are closely related to other carnivores, but there has been a running debate as to the relationship of the three families: the eared seals (Otaridae, 16 species, seven genera of eared seals and sea lions), the earless seals (Phocidae, 19 species and 13 genera of seals), and the walrus (Odobenidae, one species). For many years, these taxa were considered monophyletic, but a suggestion around the 1990s was made that they are diphyletic, with walruses and sea lions branching off from a common ancestor with bears and phocids (seals) being more closely related to the mustelids. More recent molecular and morphological work confirms the monophyly of the pinnipeds.

The Canidae include 35 species of foxes, wolves, jackals, and dogs. Ranging in size from the fennec fox (*Vulpes zerda*, just under a kilogram) to the wolf (Canis lupus, up to 80 kg), canids show a wide range of diet and social organization. The family is best known for its pack-living, social animals (wolves, the Asian wild dog or Cuon alpinus, the African wild dog Lycaon pictus, and the South American bush dog Speothos venaticus), species that raise a single litter of pups communally and that are capable of hunting animals up to 10 times their body weight through coordinated group hunting. Most species, however, live in small family groups and subsist on rodents and small vertebrates. One species, the African social bat-eared fox (Otocyon megaloitis), specializes on termites. Another species, the stilt-legged South American maned wolf (Chrysocyon brachyurus), is a solitary omnivore.

Mustelids, with 59 recent species in 22 genera, are among the most diverse of the modern carnivores. A family that includes weasels, badgers, and otters, Mustelidae are found from above the Arctic Circle to the tropical rain forests and on every continent except Australia. The group includes efficient terrestrial hunters of small mammals (mustelids, badgers) and species that are oceanic and specialize on bivalves for their diet (the sea otter, *Enhydra lutris*). Otters have become the dominant aquatic carnivore in many lakes and rivers across much of the globe. The giant otter, *Pteronura brasiliensis*, of the Amazon basin is the largest otter, weighing 30 kg, and is highly social, 'herding' fish to improve success of prey capture.

Absent from South America, the nine species of badger are mostly nocturnal omnivores of stocky build and with strong, sharp claws. The Eurasian badger (*Meles meles*), a highly social animal that lives in extended burrows or setts, is an earthworm specialist. Weasels and polecats are found from northern Greenland to the tip of Africa. Whereas the European weasel (*Mustela nivalis*) can be as small as 50 g, most weasels follow a remarkably common body form. Predominantly terrestrial and weighing 1–2 kg, with long tubular bodies and short legs, they get their speed from the compression–extension of the body; sharp teeth deliver a killing blow to a number of small vertebrates, such as frogs, rodents, rabbits, or birds. Some species once trapped for their fur are now raised in large commercial farms – e.g., mink (Mustela sp.).

The best known felid is the largest, the lion (*Panthera leo*). In many ways, the lion is an anomalous cat. Although most of the 40 species of cat are solitary hunters, the lion is a highly social, cooperative hunter. Even congeneric large cats such as the leopard (*Panthera pardus*, Asia and Africa), the tiger (*Panthera tigris*, Asia), the snow leopard (*Panthera uncia*, central Asia), and the jaguar (*Panthera onca*, South and Central America) are more similar in their social structure and hunting patterns (solitary, territorial, with male territories overlapping those of females) to the smaller cats in the genera Felis. All cats are highly carnivorous, but diet ranges from the largest bovid species to frogs, fish, and mollusks. Although most are nocturnal stealth predators, the cheetah (*Acinonyx jubatus*) is a true diurnal pursuit predator, capturing prey with bursts of speed of up to 90 km h^{-1}.

Viverrids (civets and genets) and mongooses (Herpestidae) are sometimes classified as a single family, although the most updated phytogenies separate them. Both families are exclusively Old World. Viverrids show a wide range of habits, many are truly arboreal, but there is a semiaquatic form (otter civet, *Cynogale bennettii*) and a terrestrial species from Madagascar that resembles a medium-sized cat (*Cryptoprocta ferox*). Mongooses are either solitary, nocturnal predators that resemble polecats or highly social, diurnal insectivores.

Cetacea

The discovery of Eocene fossil whales confirms an early branching of the Cetacea from the primitive ungulates, closely related to modern artiodactyls. The 11 families of whales, dolphins, and porpoises show a wide diversity of body size from the finless porpoise, *Neophocoena phocoenoides*, at 30 kg, to the world's largest mammal, the blue whale, *B. musculus*, at 150 metric tons. There is an equally large diversity of diet from tiny shrimp or krill to seals, fish, and squid – practically anything that swims. Social structure is variable and ranges from solitary individuals to the communal pods of the Orca, *Orcinus orca*, or killer whale, whose social system more closely resembles that of wolves than other whales. Cetaceans share many adaptations critical to true aquatic living. The skulls are 'telescoped,' with the premaxillary and maxillary bones forming the roof of the skulls and the occipital bones forming the back of the skull. In all whales the tail fin is horizontal.

Whales are divided into two suborders, the Mysticeti, or baleen whales, with 13 recent species in four families, and the Odontoceti, or toothed whales, which includes 71 species in seven families. Baleen whales, which include the aforementioned blue whale, are filter feeders. Incongruously, these leviathans live on tiny prey, from krill and small planktonic copepods, often less than 1 cm in length, to small, schooling fish. Migratory patterns across entire oceans from summer to winter feeding grounds are not uncommon, with many baleen whales being truly cosmopolitan species. Baleen whale populations, severely decreased by aggressive hunting during the nineteenth and twentieth centuries, are still in the process of recovering.

The toothed whales in the family Odontoceti are found in all oceans and all seas. Some of the smaller members of the family are also found in rivers and lakes on all continents except Australia. Like bats, the toothed whales all use echolocation, allowing them to hunt and live in turbid waters where visibility is low. Although echolocation 'pings' or clicks are used for locating prey, longer, more continuous tones are also important for communication among conspecifics. The toothed whales include three species of sperm whale, including the giant sperm whale (Physter catodon, family Phsteridae) made famous by Herman Melville in Moby Dick; six species of Phocoenidae, or porpoises, found throughout the Northern Hemisphere and in near-shore environments of South America; 21 species of the poorly known beaked whales (Ziphiidae), some of whose species have never been seen alive; and two species of highly gregarious Monodontidae, limited to the Arctic oceans, the beluga whale (*Delphinapterus leucas*), or white whale, and the narwhal (*Monodon monoceros*). The narwhal has a long, spirally grooved single upper tooth, used by males in intrasexual competition, which may provide the origin of the unicorn myth.

Sirenia

The five species of the families Dugongidae (with two species: the Dugong, Dugong dugon, and Steller's Sea Cow, *Hydrodamalis gigas*) and Trichechidae (with three species of manatee) are the sole modern representatives of the Sirenia. An anomalous, nearly hairless group of mammals weighing up to 1600 kg, the common name for these animals, 'sea cows,' is remarkably accurate. The only large, fully aquatic mammal that lives on grasses, the sirenids have extremely long intestines. Like perissodactyls, cellulose fermentation takes place in the hindgut, enabling the sea cows to process large volumes of relatively coarse seagrass vegetation. Reproduction is both delayed and slow, with long interbirth intervals and the production of a single offspring that remains with its mother for up to 2 years.

Sirenids have a near-global distribution in tropical and subtropical waters. Manatees are found up the Amazon and in major river systems of West Africa. They suffer little competition for their main resource, aquatic grasses. Yet, despite this, all species are threatened with extinction. These large mammals are docile, represent a large package of meat to human hunters, and are supposed to taste good. This, combined with slow reproduction, has led to overharvesting throughout their range. Pollution, dam building, and injury from the propellers of motorboats all amplify the problems of overhunting.

Proboscidea

Elephants were widely spread across North America, Asia, and Africa until recent times. Throughout the Tertiary, proboscideans showed a relatively wide range of adaptations and a diversity of species. Today, only three species remain: the African bush and forest elephants (*Loxodonta africana* and *Loxodonta cyclotis*, respectively) and the Asian elephant (*Elephas maximus*). These hulking creatures are the largest terrestrial mammals, weighing up to six metric tons. Although the Asian elephant is now limited to the forest of Asia, this probably reflects a historical shift from open grassland and woodland habitat, which has now almost universally been converted to human agricultural use. In Africa, elephants are found throughout the forests, woodlands, and savannas, but at highest densities in savanna woodlands. Some ecologists believe the split of the two African elephants, once defined as subspecies, reflects a more recent fragmentation of populations, with a former, clinal structure of genetic variation now lost.

In both Africa and Asia, elephants are ecological keystone species, one of the few species, other than humans, having a direct impact on the form and structure of the environment in which it lives. Elephants keep savanna ecosystems from reverting to woodland when they knock down trees. 'Catastrophic' destruction of trees may occur in droughts but is clearly part of a long-term natural cycle. In Asia and Africa, the fruits and seeds eaten by elephants are transported and then defecated, seeding the plains and the forests and depositing the seeds in their own fertilizer.

Elephants, fossil and modern, have a highly specialized dentition with six cheek teeth on either side of the jaw. At any one time, one to two teeth are exposed in the jaw. As these teeth wear down, they are replaced by a tooth erupting from behind the current tooth. This pattern of sequential replacement of grinding teeth is unique in mammals. Tusks are formed from elongated upper incisors.

Highly social, female elephants live in extended, matriarchal groups. Adolescence is prolonged, with males remaining in the herd until they are 10–15 years old and females not breeding for the first time until they are 15 years old. Males are solitary, and breeding priority is determined by an

interaction of sheer size (which increases indeterminantly with age) and an endocrinological sexual state called musth. Found in both elephant species, musth is associated with heightened testosterone levels and makes males extremely aggressive. The length of musth increases with age, amplifying the reproductive success of older, larger males.

Elephants have been hunted for their ivory, and in several populations in Uganda and southern Africa, tusklessness in females and reduced tusk size in males have occurred as a result of repeated selection through hunting. In the long term, the greatest threat to elephants is loss of habitat and conflict between man and elephants. But poaching for ivory, which has been cyclical, peaking in the early twentieth century, and again in the late 1980s, showed some decline in the 1990s and the early part of this century. However, particularly in Central and eastern Africa, poaching has again increased sharply and is an immediate threat to elephants.

Perissodactyla

If you are interested in horses, tapirs, and rhinos, the mid-Tertiary was the time to be alive. Although modern perrisodactyls are represented by only eight species of equid in one genus (Equus), four species of tapir in one genus (Tapirus), and five species of rhinoceros in four genera, this represents but the faintest hint of the diversity of this order in the past. Of 14 Cenozoic families, only three families and six genera are represented in the modern mammalian fauna.

Modern equids all look very similar: highly cursorial with a single hoof and consistent compact body, long neck, and highly hypsodont dentition in a large jaw with a distinct gap or diastema. Modern equids are only found in Africa and Asia, although there is evidence of the historical extinction of the tarpan, or European wild horse. Equids went extinct in North America in the Pleistocene, but asses and horses have been reintroduced in many parts of North America, the center of evolution of the Equidae. In contrast, the Pliocene of North America was populated with a few low-crowned, three-toed equids and three-toed horses, hypsodont species of various sizes, from the small, gracile Nannippus species to the larger Hipparion species. The first one-toed horse, Pliohippus, lived side by side with the hippo-like rhino Teleoceras, whereas in Asia the elephant-sized rhinoceros, Sinoiherium, coexisted with the hippo-like rhino Chilotherium.

Only two forms of social organization are seen in modern equids. Those species that live in relatively arid environments, the asses and the Grevy's zebra (*Equus grevyi*), exhibit flexible associations among females and male territoriality. Females move widely in search of food and water, whereas males tend to aggregate around water holes, the limiting resource. Mating is promiscuous. In contrast, horses and the two remaining zebra species, the mountain zebra, *Equus zebra*, and the plains zebra, *Equus burchelli*, form harems of unrelated females and a single male, a social structure also seen in gorillas.

All five species of rhinoceros are threatened with extinction, the victim of the value of their horn. Two uses of rhino horn have been identified: shavings of the horn are used in traditional Asian medicine to bring down fever, whereas in Yemen, the highest value has been placed on dagger handles made from rhino horn. Income from a booming oil economy in the 1970s led to increased demand from Yemen, whereas economic growth in Asia in the 1980s and 1990 s made rhino horn, always expensive, accessible to many more people, thus increasing demand. The black rhino (*Diceros bicornis*) was found widely across eastern and southern Africa until the late 1970s, now all but extinct except in reserves and conservancies. The white rhino (*Ceratotherium simum*), almost extinct in South Africa at the turn of the century, has recovered to where populations number 10,000, but a recent surge in poaching again threatens the species. Outside of South Africa, however, the northern subspecies is probably extinct, and introduced populations in Zimbabwe were decimated in the early 1990s. The Sumatran rhino (*Dicerorhinus sumatrensis*) and the Javan rhino (*Rhinoceros sondaicus*), formerly found across Southeast Asia, are also hanging on by a thread. Of Asian species, only the Indian rhinoceros (*Rhinoceros unicornis*) exists in any numbers, and only as the result of major conservation efforts.

Hyracoidea

Modern hyraxes are classified in a single family with three genera and four species. Limited to Africa, fossil hyraxes are also found in the Middle East and Europe. Although modern forms are all more or less rabbit-sized (2–5 kg) and resemble a very large guinea pig, some extinct forms weighed up to 50–70 kg. Tree hyraxes (Dendrohyrax sp.) are entirely arboreal, whereas both bush hyraxes (*Heterohyrax brucei*) and rock hyraxes (*Procavia capensis*) live on talus slopes, rock outcrops, and cliffs. Affiliations with other animals are uncertain, although it appears that hyraxes represent an early branching from ungulates, perhaps sharing an early evolutionary history with elephants.

Tubulidentata

Found across most of sub-Saharan Africa, the nocturnal and solitary aardvark (*Orycteropus afer*) is the last survivor of the Tubulidentata. Never a very diverse order, the majority of the aardvark's evolutionary history is in Africa and Europe, although aardvarks are known from Asia. Aardvarks weigh about 50 kg and resemble pigs with extremely long snouts, large ears, and long fleshy tails. Like most animals specialized for eating ants or termites, aardvarks have long, sticky tongues that are used to collect their prey. Powerful arms, with large claws, are used to excavate burrows and search out food.

Artiodactyla

The decline of the perissodactyls coincides with the rise of the artiodactyls. Starting in the Miocene, artiodactyls showed a remarkable diversification of families, many of which are still extant. Of 36 families found throughout the Cenozoic, 10 families and 89 genera are still found, with a near-global distribution in North and South America, Europe, Asia, and Africa. Found in all habitats, from the far north of Greenland to the southern tip of Tierra del Fuego, artiodactyls are clearly the most successful of the extant ungulates.

Diagnostic characters of artiodactyls are found in both hard and soft tissues. All members of the order have two- or four-toed hooves, with a plane of symmetry passing through the third and fourth toes. In four families (Suidae, Hippopotamidae, Tragulidae, and Tayassu-idae), four fully functional

digits are present, although peccaries have reduced digits on the hind limbs. All limbs have springing ligaments that capture energy when the leg flexes and return the energy to the foot as the limb pushes off, thus increasing efficiency of locomotion.

Previously categorized in the suborder Ruminantia, six families now stand separate: Tragulidae, mouse deer; Giraffidae, giraffes and okapis; Moschidae, musk deer; Cervidae, deer; Antilocapridae, pronghorns; and Bovidae, cattle, sheep, goats, and antelopes. They were previously grouped together because they are all typified by ruminant digestion, in which the rumen, an expanded first segment of a four-chambered stomach, provides an environment for the bacterial digestion of cellulose. This allows the animal to make use of highly siliceous grasses and vegetation. The cell contents, liberated by digestion in the rumen, and dead bacteria are digested farther along in the intestines. Ruminants also typically lack upper incisors, with lower incisors used to scrape grasses against the maxilla.

The family Tragulidae is represented by eight species in three genera, containing chevrotains, or mouse–deer. These small, primitive ungulates are found in the forests of Africa, southern Asia, and Southeast Asia. Weighing 3–12 kg, the family lacks horns although the upper canines are unusually large and grow continually. Canines are used in intrasexual competition by males and may be used in antipredator defense by both sexes.

Cervids, or deer, are found across the Northern Hemisphere, in most of South America, and in the northern fringes of Africa. The family is absent from sub-Saharan Africa, perhaps excluded by the remarkable radiation of the antelope fauna in that region. Antlers are a defining character and are found in almost all of the 51 extant species. With the exception of caribou, only males carry antlers. Unlike horns, antlers are grown and shed annually. Antlers are found in a variety of forms, from spike or stubs to the elaborate branched and palmate forms seen in the moose (*Alces alces*). Deer are found in nearly every biome, from the snowfields of the Arctic, where they subsist on lichen, to the tropical forests of South America and Asia, where fruit is a common dietary staple.

The seven species of musk deer in the Moschidae family are found in central Asia, the Himalayan plateau, and east through parts of China and Vietnam. A primitive deer, the family was originally subsumed under the Cervidae. Lacking antlers, they have sharp, swordlike upper canines, similar to those found in the chevrotains. Musk is produced by a gland in its abdomen. This waxy substance is used as the base for many expensive perfumes and demand has led to near extirpation of all members of the family across much of its range. Although captive breeding of musk deer by the Chinese might reduce demand on wild populations, prospects are not good for these animals.

Antilocaprids are only found, and have only been found, in North America. Although the family has a long and relatively diverse fossil history, only one species (*Antilocapra americana*) survives. The horns of the modern species are relatively simpler than the much more elaborated fossil forms. Unlike giraffids and all bovids, Antilocapra sheds its horn sheath annually, much like a deer losing its antlers.

The Giraffidae were more speciose during the mid-Tertiary and are represented in modern times by only two species, the giraffe, *Giraffa camelopardalis*, and the okapi, *Okapia johnstoni*. The family, exclusively Old World, has a long fossil history in Europe, where the modern genus Giraffa is first found in the Pliocene. Having evolved in Europe, giraffes dispersed to Africa. Whereas modern forms have two, short stubby horns, fossil giraffes such as the genus Giraffokeryx had four long horns. Camel-like forms, resembling more closely the modern okapi, were also found in Europe.

By far the most diverse of the ruminants is the family Bovidae. With 50 genera and 143 species, the family includes the greatest radiation of modern herbivores. Bovids are more widely dispersed than any other ungulate family, with a near-global distribution, absent only from Australia and the oceanic islands of Asia. But it is in Africa that the family has its greatest diversity. Members of all subfamilies are found in Africa, including 18 species of duiker (Cephalophinae, exclusively found in Africa); 26 of 38 species of 'true' antelope (Antelopinae), all 10 species of hartebeest, wildebeest, and topi (Alcelphinae), all but one of the eight species of grazing antelopes, including oryx (Hippotragus), all eight species of floodplain-dwelling antelopes, including waterbucks and reedbucks (Reduncinae), and the one species of impala (Aepycerotinae). The only subfamilies of the Bovidae that are relatively poorly represented in Africa are the Bovinae (10 of 24 species: African buffalo, eland, and Tragelaphus sp.) and the sheep and goats of Caprinae (three species in North Africa, none south of the Sahara, of the total 35).

The suids, peccaries, and hippos make up the last of Artiodactyls. Despite their global distribution, wild pigs and boars (family Suidae) are an Old World family, with feral populations introduced into North and South America, Australia, Tasmania, and New Guinea. They range in size from the pygmy hog (*Sus salvanicus*) of the Himalayan foothills (9 kg) to the giant forest hog (*Hylochoerus meinertzhageni*) of central and East Africa (275 kg). With 19 species in five genera, there are relatively few modern pig species, all of which closely resemble the domestic hog. Omnivores, pigs usually are forest dwellers, although the warthog (*Phacochoerus aethiopicus*) is found throughout the savanna woodlands and grasslands of Africa. Lacking horns, the upper canines of most pigs are ever-growing and form large, slashing tusks. This adaptation has gone to its extreme in the island endemic, the babirusa (Babyrousa sp.) of Sulawesi and other small islands in Indonesia, whose tusks curl up over the head.

Closely related to the pigs are the three monotypic genera that represent the extant peccaries (family Tayassuidae). Thought to be derived from Old World pigs and with a fossil record in the Old World and New World, the peccaries are South America's pig equivalent. Smaller than pigs (20–40 kg), peccaries are critical frugivores and seed eaters thought to be important for seed dispersal in South America's tropical forests.

The Hippopotamidae family includes the two species of extant hippos, the hippopotamus (*Hippopotamus amphibius*), found in Africa from the Nile basin south throughout West, East, central, and southern Africa, and the pygmy hippo (*Choeropsis liberiensis*), found only in a small section of West Africa. Recent fossil hippos are known from Madagascar, and

in the Miocene, the family was spread throughout Africa and Asia. The hippo is a nocturnal grazer, leaving rivers and lakes at night to forage along their banks.

The first fossil remains of the Camelidae, the family that includes camels and llamas, are in the Old World, although by the Eocene there are camelids in the fossil record of the Americas. Former diversity was much greater than that of the present, where only species remains in each of the Lama and Vicugna genera, both found in South America, and two Camelus are found in the Old World: the dromedary (*Camelidae dromedarius*, southwest Asia and North Africa) and the Bactrian, or two-humped, camel (*Camelidae bactrianus*, Mongolian steppe). Camels are highly adapted for desert life. The feet are splayed to prevent the large (650 kg) animals from sinking into the sand, a reversal of the usual ungulate simplification of foot structure that is evident in the South American camelids and most fossil forms. Water conservation is achieved through well-insulated bodies, a high tolerance for dehydration, dry feces and concentrated urine, and metabolic conversion of fat (stored in the hump).

Pholidota

Pangolins, also called scaly anteaters, are not very diverse – eight species in a single genus, Manis – yet they have a wide geographical range, occurring in a variety of habitats across southern Africa and in much of Southeast Asia. The order is also found in the Tertiary fossil record of all continents except Australia and Antarctica. Resembling a cross between an armadillo and an anteater, pangolins are large (5–35 kg) and both ecologically and culturally important. The scales, which cover the body completely, may constitute up to half the animal's body weight and are prized for their medicinal value across Asia, although no evidence exists that the scales, which are composed, like rhino horn, of agglutinated hair, have any therapeutic value. This has not presented rapid, and extensive, overharvesting across Southeast Asia, with the market now turning to Africa as source populations are depleted. Pangolins have the distinction of being the only extant mammals with no teeth.

Rodentia

The diversity of the order Rodentia deserves a volume of this encyclopedia of its own. With nearly half of all mammalian species (2277), nearly 500 genera, and 33 families, rodents are the most successful group of mammals living today. Two families, Muridae and Cricetidae, previously bound as one before recent analysis of molecular data, represent roughly two-thirds of the living species (hence, one-third of all mammals). Muridae's 730 species, subdivided into five subfamilies, along with Cricetidae's 679 species, subdivided into six subfamilies, includes rats, mice, squirrels, guinea pigs, beavers, kangaroo rats, dormice, jerboas or jumping mice, hamsters, mole rats, porcupines, chinchillas, agoutis, and nutria. Rodents are used by humans for food (e.g., guinea pigs), fur (nutria, beaver, chinchilla), and pets. Most rodents weigh approximately 100 g and have relatively similar body plans: long nose, large eyes, and long tails. Size varies in four orders of magnitude, however, with the smallest member of the order weighing about 5 g (pygmy mouse, Baiomys species of Central America) whereas the largest, the capybara (*Hydrocherus hydrochaeris*), may weigh up to 60 kg and is found near water across the northeast of South America.

Rodents are an ancient order of mammals, found in the Paleocene of both North America and Asia, indicating an earlier first occurrence. Rodents have always been successful, and myriad fossil forms are found in the Eocene of Asia and North America. Modern rodents are usually divided into three major groups: the cavy-like rodents, or Caviomorpha; the mouse-like rodents, the Myomorpha; and the squirrel-like rodents, the Sciuromoprha. Although a convenient division, both paleontological and molecular data suggest that this tri-partite classification may be more useful than it is real.

While supporting a phenomenal diversity in both Recent species and fossil forms, all rodents have adopted a similar jaw structure, diagnostic of the order. The incisors are used for gnawing and clipping and are sharp and ever-growing. Molars are adapted for grinding. The jaw musculature is also modified, with the muscles moving some of their attachments off the zygomatic arch and cranium forward onto the rostrum, with chewing thus providing a forward movement of the jaw.

Rodents have a global distribution and are found in every habitat type from the high Arctic to the driest deserts and the wettest tropical forests. Species like the Norway rat, *Rattus norvegicus*, have hitched rides on ocean-going vessels since people started sailing the seas 10,000 years ago. A crop pest, the Norway rat, is globally distributed and causes for billions of dollars of damage each year. Of greater historical significance, rats have acted as secondary hosts for a number of diseases that plague humans, including the bubonic plague itself, which killed 30 million people in Europe from the fourteenth to seventeenth centuries. Rats have also been one of the main agents of island extinctions.

Lagomorpha

Rabbits (Leporidae) and pikas (Ochotonidae and Prolagidae), although not diverse when compared to the rodents (92 species in 13 recent genera), are nearly worldwide in their distribution. Found in both the New World and Old World, rabbits naturally occur on all continents except Australia, where they have been introduced, as they have been to many larger islands around the world. Fossil remains are first found in Paleocene China, and they occur in a wide variety of habitats, from the Arctic snowshoe hare (*Lepus americanus*) to the tropical species of the genus Nesolagus (the Sumatran and Annamite rabbits). Rabbits are herbivores, with hypsodont molars, and, like rodents, ever-growing incisors are used to clip vegetation. Body size varies over a relatively narrow range (300 g to 5 kg), and body form is consistent: round heads, large eyes, big ears, and extended, ricochetal hind legs, which converge on those of macropod marsupials. Pikas, which are smaller than rabbits (approximately 200 g), closely resemble small caviomorph rodents and are most commonly found on rocky outcrops and talus slopes. Pikas are found in North America along the northwest coast and across the central Asian steppe into Russia.

Macroscelidea

The taxonomic position of the elephant shrews is unclear. Similar to kangaroo rats with long noses, with only 15 living species and a poor representation in the fossil record, the order is usually grouped with the lagomorphs and rodents, although earlier taxonomies have suggested an affiliation with insectivores. Found only in Africa, the order has a narrow range of body size, from the tiny 45-g short-eared elephant shrew (*Macroscelides proboscideusi*) to the rather larger 500-g golden-rumped elephant shrew (*Rhynchocyon chrysopygus*). Absent in West Africa, elephant shrews occupy a diversity of habitats, from the deserts of Namibia to the lowland forests of central Africa. Despite the diversity of habitats occupied, all species are monogamous, terrestrial, and omnivorous, although insects form a large part of their diets where studied.

Evolutionary Trends

Evolution of Brain Size

The defining characters of hominid evolution are an upright gait and an increasing large and complex brain. Humanoids have the largest brain for their body size of any mammal, extant or extinct. Brain size has, not surprisingly, been correlated with intelligence, with other relatively large-brained forms (e.g., dolphins and great apes) being imputed to be more intelligent than the smaller brained mammals. If larger brains do, indeed, confer greater intelligence and survival of an individual is in some way correlated with intelligence, then one would predict strong selection for increasing brain size in mammalian lineages.

At some point in each of their fossil histories, progressive increase in brain size is observed in primates, cetaceans, carnivores, and ungulates. However, increasing brain size has not been a linear effect, but has been punctuated with periods of rapid increase followed by stasis or relatively slow rates of change. Rapid evolution of brain size occurred early in primates, with modern prosimian brain size occurring by the late Eocene. In anthropoids, long held as the best example of progressive evolution of brain size, a rapid increase in brain size in the Oligocene was followed by relative stasis in most lineages, with the exception of the hominid line. Because carnivores, ungulates, primates, and whales have some of the largest brains and are also among the more charismatic species, generalization about the occurrence of progressive brain size evolution in mammals through time may result from a certain large mammal myopathy.

In many lineages of mammal there has been no such progressive increase in brain size – marsupials, edentates, and some lineages of rodents have shown little change in relative brain size since the orders are first seen in the fossil record in the early Tertiary. This stasis is often attributed to differences in predator pressure, with Miocene South American and Australian marsupials not requiring 'higher' brain function and hence there having been no evolutionary pressure on brain size. This explanation is unsatisfactory given the high variation in brain size when evolution occurs in different mammalian lineages.

Because brains tend not to shrink in evolutionary time, the variance in brain size will increase with time, and hence any increase in variance, with size bounded at the lower end of a distribution, will result in a larger average brain size. If selection in evolutionary time for increased brain size in different lineages reflects modern patterns of brain size variation, then examining the ecological and social correlates of brain size variation may give us an insight into the selective forces that may have shaped brain size. For instance, growth of the cerebellum has been correlated with locomotion in three dimensions (flight, swimming) as compared with terrestrial motion. Elaboration of the neocortex has been associated with various aspects of learning. Taxa that have prolonged periods of maternal dependence, and presumably long periods of information transfer, have relatively greater neocortex development than taxa with minimal parental association. Within lineages, in monogamous species males and females have similar brain structure, whereas in those species where males are promiscuous, there is elaboration of the hippocampus, the part of the brain correlated with spatial memory. The hypothesis is that in searching for receptive females, promiscuous males search over large areas, thus requiring greater spatial skills.

Cope's Rule

One of the earliest rules applied to mammalogy is Cope's Rule: species within a lineage will show increasingly large body size through evolutionary time. Cope made his findings in the late nineteenth century working on North American fossil assemblages (some of the best preserved of the Tertiary fossil record), yet most studies since Cope have failed to find support for this generalization. Some have argued that because most lineages originate at small body sizes, it is axiomatic that the only place they have to go is up. Others have stated that the pattern seen in the fossil record is a statistical artifact, the result of passive diversification of species within clades rather than the result of any kind of directed evolution. Recently, it was even suggested that although Cope was an advocate of directed evolution, he did not even do any significant analysis of body size trends and that the attribution of the first observation of this phenomenon is misplaced.

As with all paleontological examination, the detection of such trends relies on the quality of the data set used. Most studies of Cope's law have suffered from either a telescoped time frame or data sets that cover a long time run, but focus on a small number of taxa. But a recent study of Cope's law, which used a data set of 1534 species of North American mammals ranging in age from the late Cretaceous to the Pleistocene, suffers from neither of these faults. Species within a genus were followed through their evolutionary histories, and, on average, new species were 9% larger than the older species within the same genus. Diversification in size was not gradual, but changed rapidly at the Cretaceous/Tertiary (K/T) boundary, coincident with the rapid ordinal diversification of mammals. Average body weight of 29 species in the late Cretaceous was 150 g; by the early Tertiary, the average weight of 33 matched species was just more than 1000 g, an order of magnitude higher. There is an unambiguous directional trend to larger size in North American mammals, but what could possibly explain this trend?

Although data convincingly show that Cope's law is valid, at least for North American taxa, no good explanation has been offered as to why such persistent increases in body size are observed. Not only does body size increase within lineages, but at some point the middle of the size distribution drops out as well, leaving relatively large and relatively small species. This has led some to suggest that there are, perhaps, optimal body sizes for homeothermic mammals. Scientists have noted in modern assemblages of mammals that there are a disproportionately large number of species that weigh about 100 g (coincident with the average weight of Cretaceous mammals) and that this might be a lower end optimal body size for mammals. This hypothesis is supported by data on the relationship between minimum home range area and body size. While home range size scales roughly with body size in mammals, the relationship is U-shaped, with smaller mammals requiring relatively larger areas than expected. An 'optimal' body size is one in which a minimum area is required to support the activities of an individual (the bottom of the U-shaped curve): this minimum occurs between 80 and 200 g.

Upper end optima are not well defined, and gaps in the body size distributions of extant North American mammals are more illusory than real and are easily explained by random statistical variation. If Mesozoic and Cretaceous mammals were competitively excluded by dominant terrestrial vertebrates, the extinction of the dinosaur fauna at the K/T boundary may have opened up the larger end of the body size distribution, allowing mammals to evolve to larger sizes. If larger mammals show higher rates of extinction or lower rates of origination, gaps will develop in the upper end of the mammalian size distribution, perhaps mimicking the random body size gap distribution observed in an ecological time scale and allowing new taxa to continue to evolve to larger sizes.

The Island Rule

When mammals colonize islands, a strange thing happens: small mammals, such as rodents, tend to increase in size, whereas larger mammals, such as carnivores, lagomorphs, and artiodactyls, tend to become smaller. Numerous examples of dwarfism and gigantism have been found in the fossil record and can be documented in extant species, the most notable being the discovery of fossil evidence of pygmy elephants on Mediterranean islands and fossil remains of dwarf mammoth on the Channel Islands of California.

Explanations for this Island Rule have been numerous, and none is completely convincing, but for the most part they focus on a combination of resource limitation, reduced competition as a result of a depauperate fauna, and predator release. For larger species, a limited food supply would favor a smaller body size as individuals of smaller body size require fewer resources, can reproduce more efficiently, and are more likely to leave surviving offspring. Niche partitioning through character displacement will result in selection for the smaller competitor in a feeding guild to become yet smaller. Therefore, it is not surprising that the absence of the larger competitor results in an observed larger body size in the smaller species of a feeding guild. Predation has been hypothesized to have two different effects: in larger mammals, where size is a form of predator defense, release from predators may select for reduced body size. This argument, of course, is confounded with the explanations provided by resource limitation arguments. For smaller mammals, where stealth and crypsis are defense strategies, absence of predators may reduce the adaptive value of small size, thus allowing small mammals to evolve to larger sizes.

A combination of bioenergetics and hypotheses put forth for Cope's Rule may provide a somewhat more synthetic argument for the Island Rule. Studies of mammalian feeding guilds suggest that within a guild of large-bodied mammals, the individuals of smaller species frequently monopolize a greater proportion of the available resources. In guilds dominated by smaller mammals the opposite appears to be true: individuals of larger species within the guild appear to control a greater proportion of the available energy. If there is an optimum size for mammals, arrival on an island in which there was release from predators and guild competitors would allow rapid evolution toward that optimum. Animals that were significantly larger, or smaller, than the optimum would show the greatest divergence between insular and continental body size.

These Legs Are For Walking: Predators and Their Prey

The evolution of ungulate locomotion, as typified by the evolution of the equid leg, is a topic covered by every high-school textbook. Cursorial specialization has evolved independently in a number of mammalian lineages, but there is a general pattern of morphological changes that results from a simple calculation: for an animal to move faster, it must increase either the length of its stride or the number of strides it takes. Hence, shifts in morphology must make the leg longer, faster, or both. The lateral reduction and fusing of bones (especially the hand and foot bones, the metacarpals, and metatarsals), the elongation of limbs, the reduction or loss of the clavicle, and the shifting of muscle mass toward the torso with a concomitant increase in the use of tendons to move the limbs are all part of the suite of changes that improve running speed.

Like their herbivore prey, carnivores have also evolved increasingly long-legged, faster forms through time. Carnivores have gone from being short-legged, small creatures to the long-limbed, slender-bodied forms typified by modern pursuit predators such as the cheetah or wolf. These forms appear to have relatively short evolutionary histories, with repeated evolution of similar body types in different carnivore lineages.

The coevolution of faster herbivores and faster predators has been described as a predator–prey arms race: slight increases in prey speed lead to higher rates of survival of those swift individuals who pass on their genes at disproportionate frequencies. Similarly, faster predators have greater hunting success, produce more offspring, and are better represented in successive generations. There is only one problem with this scenario: the fossil data do not support the hypothesis. Ungulates have tended to evolve ever more cursorial forms throughout the Cenozoic, beginning 55 million years ago in the Eocene. But is was not until the Pliocene, only 5 million years ago, that carnivores 'caught up.' By this time, most of the 'novel' ungulate adaptations for speed were tens of millions of years old.

Long legs confer the ability to move quickly, but what if this ability were a secondary adaptation, the result of selection for another phenomenon? An alternative explanation for cursorial adaptation suggests that the suite of adaptations that has evolved in ungulates has to do with the energetics of movement rather than the energetics of being eaten. All the observed adaptations do one of two things – increase stride length or reduce the costs of moving the leg. Hence, long legs lead to increasing efficiency of movement, whether the animal is walking 10 km to get to a water hole or dashing to get away from a predator. If the opening of the savanna led to increased distance movement in ranging patterns, long, simplified, lightweight legs would confer a great advantage in terms of energy efficiency of movement, whether daily or migratory.

Convergent Evolution

Mammalian evolution has been rich with novelty, whether in the evolution of flight in bats, the suite of adaptations evolved by whales that have brought mammals into the sea to live, or more specifically in one lineage the well-developed venom gland found in the hind legs of the platypus. Throughout mammalian evolutionary history, however, some of the more remarkable evolutionary patterns have involved the repeated evolution of derived characters in widely divergent mammal lineages. Convergent evolution spans a wide array of adaptations, from the suite of carnivorous Miocene South American marsupials that resemble extinct and contemporaneous placental carnivores to the striking similarity of burrowing forms in the marsupial mole (family Notoryctidae), the golden moles, (family Chrysochloridae), and true moles (family Talpidae).

Throughout their evolutionary history, mammals have shown trends to increasing specialization in many morphological and correlated ecological functions. For instance, as grasslands became widespread in the Miocene, ungulate tooth morphology shifted from a dominance of low-crowned, or brachyodont, teeth, which are easily ground down through the animal's lifetime, to a dominance of hypsodonty, or high-crown teeth, which last longer when an animal eats siliceous grasses covered in dust and dirt. This evolution has been complemented by an increased complexity of molar teeth to enable herbivores to more thoroughly grind their food. This pattern is observed in both artiodactyls and perrisodactyls. Rodents have solved the problem differently, evolving ever-growing teeth, whereas proboscideans roll their teeth out of their jaw, each tooth having a limited useful life.

Convergent evolution may address the same problems but find different solutions. The expansion of grasslands and a high-cellulose diet have led to two different solutions to digesting an essentially indigestible substance. Ruminant digestion, widespread in the artiodactyls, uses the fore-gut, or rumen, as a fermentation chamber in which bacteria break down cellulose, making the grass cell contents available for absorption and converting the cellulose into digestible material. Because the passage of materials through the rumen limits intake, artiodactyls tend to be relatively selective, choosing grasses of high quality. In perissodactyls and some kangaroos, a similar solution has evolved – bacterial breakdown of cellulose – but the site of bacterial digestion is in the hindgut and the process is less efficient. Although this means that hindgut fermenters cannot draw out as many nutrients from a given pulse of food, by rapid processing of food, hindgut fermenters can gain sustenance from larger volumes of lower quality forage.

Table 3 lists a few of the more widespread patterns of convergent evolution in mammals.

Geography, Biogeography, and Biodiversity

Gradients of Species Richness: Diversity, Density, Range, and Rapoport's Rule

The tendency for species richness to increase with decreasing latitude was first observed by Alfred Russell Wallace and has been the subject of significant study. Mammals, as well as nearly every taxa studied, appear to show this pattern (Safi *et al.*, 2011). Although nearly a dozen hypotheses have been suggested to explain the phenomenon, no single explanation appears dominant.

In a study of nearly 200 North and Central American mammals, which matched subspecies living at lower and higher latitudes, it was found that those living at lower latitudes had significantly smaller overall latitudinal geographical distributions. The same pattern is observed in a review of 679 North American mammals in which latitude is compared to the range of the species. In this study, in addition to the increase in range size with increasing latitude, a similar pattern is observed in a west-to-east gradient of longitude, with range sizes increasing to the east. Between the tropics and the northern Arctic regions, the average range of a species increases by a factor of 30. Interestingly, this same pattern is not observed in Australian mammals. The continent has a divergent geographic structure, with an arid center fringed by moist mountains, creating habitat diversity (and stability) very different from that observed in northern continents.

The vast majority of North American mammals have a narrow geographical species range that covers but a few habitat types and appear to be habitat specialists. Not surprisingly, patterns of variation are not consistent across orders, with the ranges of carnivores and artiodactyls being larger, on average, and bats and rodents much smaller. Nonetheless, the observed pattern of increasing range with increasing latitude is not an artifact of taxonomy: these patterns are seen both across all mammals and independently within mammalian orders.

Overall, range size shows a log-normal distribution, with very large and very small ranges being exceptional. Species with small ranges would probably result in a greater risk of extinction through severe catastrophic events such as abnormally cold, hot, or dry weather, through disease, or because of predation or competition. Species with unusually large areas may be prone to extinction because of the low densities at which they occur and the high probability that somewhere within their range they will come into conflict with humans.

It has long been noted that species diversity (or more accurately for mammals, species density, the number of species per unit area) decreases with increasing latitude. But

Table 3 Examples of convergent evolution in mammals

Geographic location					
Adaptation	*Holarctic*	*Africa*	*Madagascar*	*South America*	*Australia*
Carnivory	Carnivora Many families	Carnivora Many families	Carnivora Viverrids	Borhyaenidae*	Dasyuroid Phalangeroid
Pursuit predators	Carnivora Dogs Dog bears* Bear dogs*	Carnivora Hyenas Cheetah		Didelphoid Borhyaenid	Dasyuroid *Thylacinus**
Saber-tooth carnivores	Smilodon*	Felids	Viverrids	Borhyaenid*	Phalangeroid
	Felids Nimravids*		Fossa	*Thylacosmilus**	*Thlacoleo**
Semiaquatic	Insectivora Desmans Water voles Pantolestids*	Insectivora Otter shrew	Insectivora Otter tenrec	Didelphoid Water opossum	Monotreme Platypus
	Rodentia Musk rat				Rodentia Water rat
	Carnivora Otters				
Aquatic	Carnivora Sea lions/seals				
	Cetacea Whales				
Herbivores with cellulose fermentation	Ungulates	Ungulates		Ungulates	Phalangeroids
	Artiodactyla Perissodactyla Some rodents	Proboscidians Hyracoids Arsinotheres* Some rodents		Notoungulates* Lipoterns* Pyrotheres*	
Hornlike structures	Perissodactyls Rhinos Brontotheres*	Arsinotheres*		Notoungulates* Toxodontids*	
	Artiodactyls Protoceratids* Oreodonts* Pigs Ruminants				
	Archaic ungulates Uinthatheres*				
	Rodents Mylagaulids*				
Anteaters	Pholidota Pangolins	Pholidota Pangolins		Edentates Anteaters Armadillos	Dayroids Numbat Monotremes Echidna
Gliding	Rodents Flying squirrels	Rodents Scaly-tailed squirrels			Phalangers Flying possums
	Dermopterans Flying lemur Paromomyidae*				

Source: Reproduced from Janis CM and Damuth J (1990) Mammals. In: McNamara KJ (ed.) *Evolutionary Trends*, pp. 301–346. Tucson: University of Arizona Press.
Asterisks indicate extinct taxa.

species density across a continent also has a strong trend, with species density increasing from east to west. The increase in density is greatest in the more southerly latitudes (30–40°) and correlates with increasingly complex topography, and hence greater habitat diversity.

Is there a connection between these two patterns, with narrowly distributed species packed in more closely at southern latitudes and in more complex habitats? Selection for wide tolerances of temperature and moisture extremes is more likely to occur in the temperate zones, where

temperatures can range from freezing or subfreezing in the winter to 35 °C in the summer. Patterns of climatic variation are similar over large areas; hence moving from place to place does not usually expose an individual to temperature or moisture extremes that have not previously been experienced. At more northern latitudes, the range of extremes of moisture, temperature, and light become more extreme, and unless a species is migratory, an individual will be exposed to these extremes year after year throughout its lifetime.

Although wide climatic tolerance might have few costs, it would confer few benefits in the tropics. In the tropics, the range of climatic variation at any given location is low. Microhabitat variation operates over a smaller spatial scale; hence staying in one place does not expose an individual to climatic variation, but moving a few tens of kilometers, or a few hundred meters in elevation, can expose an individual to conditions to which the individual, or its ancestors, have not been exposed. This has prompted one scientist to note that "mountain passes are higher in the tropics." The physical structure of such passes is, of course, no different in the tropics. But the perception of a dispersing animal trying to cross that pass may vary widely, depending in large part on the way adaptation for extremes in temporal variation of ecological variables has equipped the individual to deal with new microclimates encountered on its journey.

The most recent richness measures support the classic latitudinal gradient observation. Climate, however, impacts species diversity and functional diversity differently: for example, tropical climates generally have a lower functional diversity. Assuming stable environmental conditions, this is most likely because each niche experiences higher levels of 'occupancy,' in turn allowing for higher levels of species diversity. If phylogenetic diversity represents the amount of evolutionary time shared by a group of different species, and functional diversity represents the amount of species that share similar traits within a particular habitat or community, then it follows that the comparison of the two will reveal the importance of historical and evolutionary processes relevant to current richness patterns. This is the current line of thinking.

Hot Spots of Species Richness and Endemicity

Broader patterns of species diversity and species density may explain large-scale geographic variation in observed diversity of mammals, but within these patterns there are finer grained anomalies in distribution that result in areas with extremely high diversity, the so-called hot spots. Across all species of terrestrial vertebrates, 35% of all species are confined to a total of 1.4% of the Earth's land surface. Almost by definition, this includes a large proportion of geographic endemics, and not surprisingly, such hot spots tend to be much more common in the tropics than in temperate areas for reasons obvious from the above discussion of geographical patterns of diversity. This pattern has strong implications for priority setting in conservation of mammals in the tropics, with a greater bang for the buck (as measured by species protected for dollar expended) if conservation investments are focused in these hot spot areas.

The geographical and geological precursors that lead to high levels of species endemism are not well understood. Despite the radical difference in life histories and evolutionary histories, patterns of endemism frequently are highly congruent across taxa, indicating necessary and sufficient conditions are probably common to the evolution of endemicity in mammals and other taxa. Indeed, in Africa and its associated islands, endemism of plants and animals converges, with the hottest of the hot spots for all taxa being found in places like Madagascar, the Karoo, the Cape Province of South Africa, and the Ethiopian Highlands.

The diversification of a small number of families on islands is probably as much due to vicariant or random arrival of animals on the island as it is to any set of geological or climatic conditions. The diverse lemur fauna of Madagascar and the rapid radiation of marsupials on the Australian continent owe as much to isolation as they do to anything else. The rapid, and astonishing, diversification of cichlid fishes in the rift lakes of Africa or the convergent pattern seen in the fish genus Semionotus in the Jurassic rift lakes of the East Coast of North America also shows the potential for rapid evolution of endemics in habitat 'islands' within a larger continental setting.

However, isolation *per se* does not explain the evolution of endemism on larger continents, where connections to other source populations of mammals have come and gone over the epochs. Africa, unlike other continents, has retained a relatively stable location on the globe through the past 300 million years. Such stability, relative to the drift of other plates, has provided a relatively constant set of ecological conditions as associated with latitudinal changes. Superimposed on this stability, however, have been the making and breaking of connections to Asia, repeated geological upheavals associated with the repeated formation of an African rift, and a cycle of drier periods that led to expansion and contraction of the extent of the central African forests.

Within this geologically unstable environment, there have been areas of remarkable stability, areas that have been consistent in their patterns of temperature and rainfall, whether hot and dry or hot and wet. This climatic stability has led to the evolution of endemic mammals in places as diverse as the Ethiopian Highlands, the forests of the Congo basin (former Zaire), and the Usambara Mountains (although this area is better known for its avian and floristic endemism).

Endemic hot spots appear to be at greater risk than equivalent areas of similar vegetation type. In tropical forests, for instance, in hot spots, only 12% of the original vegetation persists, whereas 50% of all tropical forests still stand. Furthermore, studies have shown a general pattern of increasing threat with endemism, and that this pattern holds for mammals. This suggests that hot spots are an order of magnitude more threatened than the average tropical forest. One explanation for this ties into the evolutionary explanation for hot spots. Many of the same patterns of geography and climatic stability that lead to the evolution of hot spots are also those that lead to the settlement and cultural evolution of humans. Hence, convergent patterns of human cultural diversification and biotic diversification lead to a situation in which people and hot spots are in closer proximity, competing for the same resources and leading to greater conflict than would be expected if species and people were randomly distributed.

New Discoveries and Descriptions

Mammals may be exceptionally well described, as compared with other taxa, but this does not mean that the taxonomy of Mammalia is either complete, or stagnant. In the decade 1992–2003, more than 340 new mammal species were described, with many of these being the result of taxonomic revision. These are not minor revisions, with 28 of these species belonging to new Genera, and one of them a new, monotypic family. Increasingly sophisticated molecular tools, as well as a better understanding of processes of phylogentic divergence, suggest that this pattern of increasing mammalian diversity will continue, and perhaps in the near term accelerate, as Families are revised and new molecular tools applied. Most of these species are found in the species-rich tropics, and appear to have a restricted range size.

Patterns of Discovery

The discovery of a new mammal, even a large mammal, is not such an unusual event. In the past 60 years, more than 2 dozen species of large mammals have been described and several others are known but await formal description. Mammals in general follow a rate of about 25 new descriptions per year. Consideration of the pattern of description of new species of mammals shows that the curve is more or less bell shaped (**Figure 3**), with the peak of activity in describing mammals coming early in the twentieth century. While new mammals continue to be discovered each year, the rate at which these discoveries have occurred has not changed significantly in recent decades.

If finding new species is not a particularly rare event, can we predict how many species will eventually be discovered? Recent analysis of the discovery and description of large (>2 m) open-water marine mammals suggests perhaps that 47 species await formal scientific description, that we should continue finding species at a rate of one new species every 5 years, and that we are most likely to find new cetaceans. Given the long and intensive history of commercial whaling, such a finding is surprising.

New species can be named in one of several ways. In some cases, specimens collected many decades earlier were neither properly examined nor adequately described. Hence, when a systematist revises a taxon, 'new' species may be found sitting in a museum drawer. Similarly, new forms of evidence from field, laboratory, or genetic studies may force a revision of a species group. Molecular genetics has played a large role in this field, forcing scientists to reassess previously assigned taxonomic names. Finally, previously unknown forms may be found in nature.

The Annamites: The Last Frontier for Large Mammals

The Annamite mountain range, which forms the border between the Lao People's Democratic Republic and Vietnam, has yielded a large number of discoveries of new species and rediscoveries of previously described species of large mammals in the past decade. The Vietnamese warty pig, Sus bucculentus, long thought to be extinct, was found in a food market in Laos in the early 1990s. The field has been particularly rich for discovery of barking deer, primitive deer found throughout the tropics of southern Asia and Southeast Asia. In 1994, the giant muntjac was discovered in the forests of the Annamites. Large in size and with unusual antlers, the species was thought sufficiently morphologically distinct to deserve its own genus, Megamuntiacus. DNA analysis suggested that the species belongs with other muntjacs in the genus Muntiacus (*Muntiacus vuquangensis*) and not in a genus of its own.

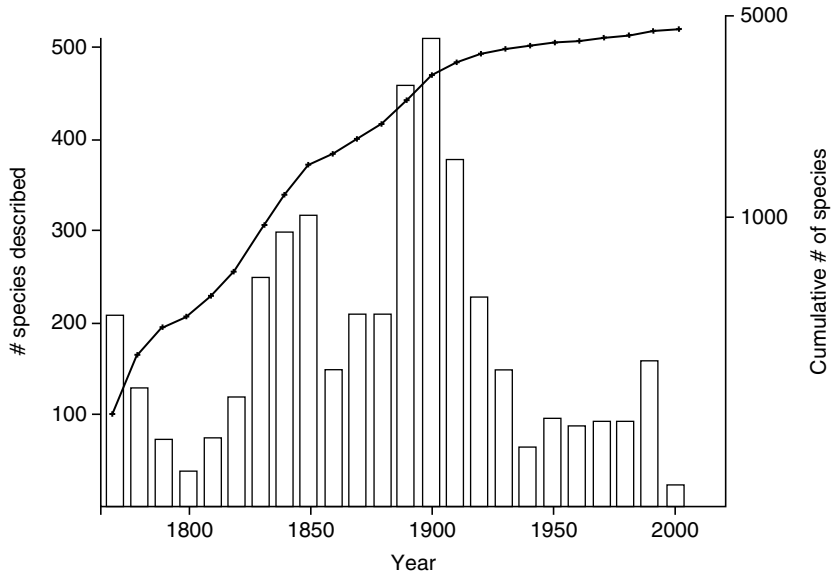


Figure 3 The number of new species described (left-hand scale, histogram) and cumulative number of mammal species (right-hand scale, line). Reproduced from Wilson DE and Reeder DM (1999) *Mammal Species of the World*, 2nd edn. Washington, DC: Smithsonian Institution Press, with permission from Nature.

Two further Annamite Muntiacus species have been found, one a new species from Quang Nam Province, Vietnam (*Muntiacus truongsonensis*), and another from the Lao side of the border, a rediscovery of Roosevelt's muntjac, *Muntiacus rooseveltorum*. In neighboring Myanmar (Burma), a 1997 expedition to the far north of the country, an area not visited by scientists since the 1930s, yielded discovery of the world's smallest deer, another species of muntjac, the leaf deer (*Muntiacus putaoensis*).

The Annamites have also been the source of three discoveries that have both puzzled and captivated scientists. The first is the saola, or *Pseudoryx nghetinhensis*. The species, which weighs approximately 100 kg, derives both its scientific name and Lao common name from the shape of its horns: long and arced, the species' horns resemble both the Africa oryx and the arc of a lyre-like Lao musical instrument. Initially found only in the Vu Quang forest reserve, Vietnam, the species is now known to range widely at higher altitudes in the Annamites. Most closely related to bovids, the saola is clearly unusual and represents a deep branch in the phylogeny of the bovids.

Another, perhaps even more curious, Annamite discovery was that of a new rabbit species (*Nesolagus timminsii*). The rabbit, which has distinct, dark brown stripes running down both its face and back, a reddish rump, and short ears, was first identified in a food market in Ben Lak, Laos, but since has been photographed in Vietnam. The rabbit's closest relative is a critically endangered species found in Sumatra, about a thousand miles away (*Nesolagus netscheri*). Despite extreme morphological similarity, genetic data suggest that the two species may have diverged about 8 million years ago. The two are so divergent genetically that it was debated as to whether the new (as yet unnamed) species should be placed in its own genus.

If these discoveries were not expected, a species described in 2005, the Laotian Rock Rat (*Laonasles aenigmamus*) is perhaps even more remarkable in being not just a species new to science, or Genera, but an entirely new Family, Laonestidae one of only three Families added to the Mammalia since 1900.

Cryptic Species, Phylogeography, and Evolutionary Significant Units (ESU)

New discoveries do not require going to the ends of the earth. The pipistrelle (*Pipistrellus pipistrellus*) is one of the most common bats in Europe, and one of the best studied. In the British Isles it was found that the bat used two frequencies to echolocate, which led scientists to classify these populations as either the 45- or 55-kHz phonic type. Because roosts consisted of bats of a single phonic type, to these scientists it was suggested that the phonic types represented sympatric, but distinct, species. Because of their morphological similarity, they are called cryptic species. Two recent studies have provided further evidence that these two phonic types are, indeed, good species. The two show relatively good separation of diet: whereas both eat mostly the dipteran suborder Nematocera, the dominant prey groups for the 45-kHz phonic type were in the families Psychodidae, Anisopodidae, and Muscidae, while the families Chironomidae and Ceratopogonidae were the main prey groups of bats of the 55-kHz phonic type. Studies of the cytochrome b gene also showed unambiguously that the two phonic types are distinct species. That such cryptic species can be found in such a well-studied group suggests that further genetic studies of mammalian phylogenies will yield some surprises.

Molecular genetics is a powerful tool for disentangling the relationship between the geographical and morphological history of subspecies within a species group, the study of phylogeography. Because morphological, geographic, and genetic data frequently tell different stories (as in the striped rabbit e.g., above), an understanding and reconciliation of these data are critical to understanding both the evolution and conservation of mammals. The recent study of Australia's largest extant carnivore, the marsupial tiger quoll (*D. maculatus*), is instructive. The quoll occurs on Tasmania and on the mainland in two distinct populations, one in the north and one in the south. Previous to genetic analyses, strong body size and morphological convergence led to the southern and Tasmanian populations being grouped as one subspecies (D. m. *maculatus*), whereas the northern population was classified as a separate subspecies (D. m. *gracilis*).

Genetic evidence suggests, however, that despite their size differences, which may be as much as 50%, the two mainland populations are more closely related to one another, while the Tasmanian population is genetically divergent. Why does this matter? If the two mainland populations constitute one ESU while the Tasmanian population constitutes a different ESU, then one management solution would be to manage the mainland and island populations separately. This would ease management concerns as the northern population is classified as endangered, whereas the southern population is relatively more common. Yet, it is likely that if a female northern quoll were placed in a breeding center with a southern male, because of their great size difference, the male would view the female as dinner, not as a potential mate. Similarly, translocation of southern quolls to the north to augment that population could be disastrous. Hence, although phylogeography suggests one management regime, a more complex solution, with each population managed separately, is recommended.

Current Extinction Crisis

How Many Species Have Gone Extinct?

The current rate of extinction of mammals is, by any measure, frightening. In the past 500 years, approximately 88 species of mammals are thought to have gone extinct. This represents approximately 1.9% of the extant species of mammals. Using the 'background' or natural rate of extinction derived from an examination of the fossil record, one would have expected at most one species of mammal to have gone extinct in the same time period. This accelerated rate of extinction can be ascribed, directly or indirectly, to a single cause: humans. While determining what factor is ultimately (as opposed to proximately) involved in a species extinction, habitat loss, introduction of exotic species (including zoonotic diseases), and

overharvesting all have contributed to the high historic rates of mammalian extinction.

One would think that assessing the rate of extinction in such a well-studied group as mammals would be relatively easy. Yet, despite relatively good taxonomy and detailed study of the question, there remains considerable debate among mammalogists and conservationists about the number, the precise identity, and the timing of disappearance of those species that may have gone extinct in historical times. In the 1990s, two efforts to categorize historical extinctions came up with approximately the same number of species going extinct: 85 by one estimate and 88 by another. What is disconcerting, however, is that the two lists contain only 57 species in common, resulting in, overall, 116 different species being listed by one or another of these studies. This suggests that, even among experts, there is some debate about what actually constitutes an extinction.

How much does this matter when assessing patterns of modem loss of mammalian species? On the one hand, documented extinction rates are so far over background rates that definitive listing will not change the implications for conservation of mammals. On the other hand, such lists are playing an increasingly important role both in setting priorities for conservation (the process of keeping species off the list) and in public debate, and an accurate, defensible list is essential.

A clear ability to agree on what is meant by extinction is critical both to conservation planning and to conservation science. While all differences in lists of extinct species need to be reconciled, much can be learned by analyzing why one set of authors excludes species that other authors list as extinct (**Table 4**). The reason(s) why authors disagree can tell us much about the source of such errors, or differences of opinion, and also help us focus our efforts more clearly on those species where changes in status affect conservation action.

Most lists examine the extinction of species over a particular time frame. For instance, one might examine mammalian extinctions since the beginning of the age of exploration starting in 1500. The first question one must ask is, did the species in question really exist at any time during the past 500 years, or was extinction during this time inferred incorrectly either through data in the literature or from misplaced stratigraphy of sub-fossil material? In **Table 5**, clearly a plurality of the 76 species excluded were in this category. There is no dispute whether or not these 31 species went extinct, rather just a question of whether their extinction predated the period under study. While inclusion or exclusion of these species will change the calculation of rates of extinction, such changes have little practical applications to conservation.

Similarly, removing or adding a species from a list for reasons of taxonomic uncertainty or taxonomic revision does not manifestly change conservation status of a species. Take for instance the quagga. Distinct from other zebras, with only vestigial stripes, the quagga, *Equus quagga*, was initially thought to be a distinct species of zebra extirpated from the southern tip of Africa at the turn of the twentieth century. Molecular analysis of museum specimens showed unambiguously that the quagga was a subspecies of the common zebra, *Equus burchelli* – one less extinct species, to be sure – but such revision does not change the way we manage the existing populations of *E. burchelli*. Similarly, if two extinct species are found to be synonymous, as in the case of Johnson's hutia (*Plagiodontia ipaneum*) and *Plagiodontia velozi* (no common name), management plans remain unaffected.

When a species thought to be extinct is determined to still be extant, more likely than not the size and status of the extant population will be either totally unknown or known to be critically endangered. Hence, transition from extinct to extant, unlike other categories, has important conservation implications. In **Table 4**, such cases represent about one-quarter of those species listed. Such a transition should be a red flag indicating that further study, and conservation action, are likely needed.

Table 4 A review of species not accepted as "Extinct"

Species not known to exist in past 500 years	31
Species extant as another taxa	18
Species extinct as another taxa	3
Taxonomy unresolved	3
Species extant	21

Source: Data from McPhee RDEand Flemming C (1999) Requiem: Aeternam: The last five hundred years of mammalian extinctions. In: MacPhee RDE (ed.)*Extinctions in Near Time*, pp. 333–371. New York: Kluwer Academic/Plenum.

Table 5 Geographic patterns of mammalian extinction

Type	*Loss(%)*
Islands	
Caribbean	37.5
Pacific	21.6
Indian Ocean	8.0
All other	6.8
Total	73.9
Continental	
Australia	19.3
Africa	4.5
Eurasia	1.1
Americas	1.1
Total	*26.0*

Source: Data from McPhee RDE and Flemming C (1999) Requiem: Aeternam: The last five hundred years of mammalian extinctions. In: MacPhee RDE (ed.) *Extinctions in Near Time*, pp. 333–371. New York: Kluwer Academic/Plenum.

Patterns of Modern Mammalian Extinctions

Although the rate of extinction is extremely higher than the expected background rate of extinction, the time frame over which historical extinctions have occurred is so short that it is often difficult to discern patterns in data collected on mammalian extinction. Nonetheless, both taxonomic and geographical patterns do emerge, both of which may be informative if we try to project future patterns of extinction in mammals.

Perhaps the most striking pattern seen when one examines data on historical extinctions of mammals relates to the geographic distribution of extinct species. In the past 500 years, the great majority of extinctions have been on islands, and

within island groups, the Caribbean islands have suffered the most extreme loss, accounting for nearly 40% of all recorded extinctions (**Table 5**). Mammalian extinctions are, in this case, no different from those of other well-studied taxa, with more than 90% of modern avian extinctions and 89% of modern molluskan extinctions also occurring on islands. In recent times, continental extinctions have, for the most part, been rather rare, and nearly all continental extinctions have been on the continent of Australia.

A recent sophisticated statistical analysis of the correlates of mammalian extinction in the Holocene (Turvey and Fritz, 2011, p.0955) has shown that although there are correlates of extinction, these have changed more than the past 12,000 years. Extinctions of mammals have been greatest on islands, or on continental areas with long histories of human inhabitation, and relative human domination of ecosystems. Particularly on continents, larger mammals are more susceptible to extinction; as a result of an interaction between human dominance, and body size, the newest waves of extinctions appear to be concentrated in South Asia and the New World, particularly the formerly remote regions of the Amazon. Finally, the threats to mammals appear to have changed over time, with introduced species (of particular importance to island populations) and overexploitation being supplanted by widespread habitat loss and fragmentation as the current most significant threat to mammals. Clearly, the overlay of climate change, and potentially radical shifts in community structure, will amplify and likely become even more important than current threats.

One of the most striking set of extinctions is in the Insectivora, where an entire family (Nesophontidae, the Antillean island shrews) has gone extinct. Eleven percent of all mammalian extinctions recorded in the Modern era have occurred in this family. Insectivora, as an order, appears to have a greater propensity to extinction, with 11 species having gone extinct, representing 2.5% of the ± 450 described species of Insectivora. This is approximately 30% above the overall rate of 1.9% for mammals.

Other orders that have overrepresentation in recent extinctions include two eutherian orders, the most diverse being Rodentia (52% of all extinctions, 44% of described mammals). Similarly, among therian mammals, the most diverse order, Diprodontia, with 117 species, which includes possums, cuscuses, wombats, and the koala, shows overrepresentation in the extinction table (6.8% of extinctions, 2.5% of all mammalian species) as does one of the least diverse, the Perameldia, or bandicoots (3.4%/0.45%).

In contrast, whereas nine species of the order Chiroptera have gone extinct, representing 10% of the known modern extinctions, bats represent nearly 20% of the extant modern mammalian fauna; hence, one could argue bat extinctions have been underrepresented in recent times. Similar arguments could be made for both carnivores (2.2% of extinctions, 5.8% of species) and primates (3.4% of extinctions, 5.1% of species).

Projecting Future Extinctions

Compiling lists of recently extinct species and studying current extinctions tells us something of the recent history of changes in mammalian biodiversity. The strength of such lists is that they are an assessment of the global patterns of extinction and tell us something about the persistence of a species, family, order, or size class. To better understand the immediate future of mammalian extinctions (perhaps the next 200 years) rather than reflect on what has gone extinct, there is greater value in examining which species, families, and orders are under threat and discussing how threats will lead to decline, and ultimately to extinction. Unnervingly, most of the newly described species have been found in increasingly threatened tropical forests in Asia and Latin America, have narrow geographic distributions, and hence are ever more likely to face extinction. These processes are more fully discussed elsewhere in this encyclopedia, as are patterns for particular orders of mammals.

For mammals overall, however, patterns of threat do not correlate well with patterns of recent extinction. A recent assessment of threatened species by the World Conservation Union (IUCN) suggests that 25% of extant mammal species are threatened with extinction. Some of the groups that are overrepresented in analyses of recent extinctions are prominent in the list of threatened species: rodents are somewhat underrepresented, with 'only' 17% of the species in the order threatened with extinction; insectivores, however, look likely to dominate the lists of extinct species in centuries to come, with 36% of the species in this taxon threatened. What is more interesting is that some orders that have shown relatively low rates of extinction such as carnivores and primates have high proportions of their taxa threatened. In carnivores nearly 26% of all species are under threat, whereas nearly half the primates, 46%, may go extinct in the next century.

Analysis of the extinction of subspecies or populations of animals gives some insight into what is generating this pattern. Among both carnivores and primates, rare species have been subject to increasing threats from habitat loss and fragmentation, isolating endemics in smaller and smaller patches. But wide-ranging species are also showing promise for future extinction through similar processes acting at larger scales. The spotted hyena, although still abundant and occurring at relatively high densities for a large carnivore, requires a matrix of resources that are widely dispersed across a landscape. Combined with a relatively low level of dispersal, habitat fragmentation and isolation suggest the species will come under increasing threat. The African wild dog still has a nearly continental distribution, but most populations are small (fewer than 100 individuals) and increasingly isolated. Threats to the species amplify one another: habitat loss and fragmentation, loss of prey, disease introduced by domestic companion animals, active hunting, and road kills all contribute to mortality. The decline of the species in West Africa, where populations became isolated and then went extinct one by one, like candles being snuffed, is probably indicative of patterns that are evolving in central and East Africa. Increasingly, although the species is still extant, and will remain so for decades in southern Africa, it is being lost as a component of the great majority of ecosystems in which it once lived.

While extremely valuable both for planning and for developing an understanding of which species are most likely to go extinct, lists of extinct or threatened mammals only address the concern of evolutionary extinction, the ultimate loss of a species. Increasingly, scientists who study mammals, and

lobby for the preservation of mammalian diversity in particular and biodiversity more widely, will have to address issues related to the local, regional, and global ecological consequences of a species' march toward extinction.

Appendix

List of Courses

1. Biology of Mammals
2. Mammalogy
3. Vertebrate Ecology and Evolution
4. Vertebrate Biology

See also: Birds, Biodiversity of. Endemism. Hotspots. Mammals, Conservation Efforts for. Mammals (Late Quaternary), Extinctions of. Mammals (Pre-Quaternary), Extinctions of. Modern Examples of Extinctions. Vertebrates, Overview

References

Alroy J (1998) Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* 280: 731–734.

Alroy J (1999) The fossil record of North American mammals: Evidence for a paleocene evolutionary radiation. *Systematics Biology* 48: 107–118.

Baillie J and Groombridge B (1996) *IUCNRed List of Threatened Animals*. Gland, Switzerland: IUCN, The World Conservation Union.

Barlow KE (1997) The diets of two phonic types of the bat *pipistrellus pipistrellus* in Britain. *Journal of Zoology* 243: 597–609.

Carroll RL (1988) *Vertebrate Paleontology and Evolution*. New York: W. H. Freeman.

Davies TJ, Buckley L, Grenyer R, and Gittleman JL (2011) The influence of past and present climate on the biogeography of modern mammal diversity. *Philosophysical Transactions of the Royal Society B* 366: 2526–2535.

Dayan T and Simberloff D (1998) Size patterns among competitors: Ecological character displacement and character release in mammals, with special reference to island populations. *Mammal Review* 28: 99–124.

Firestone KB, Elphinstone MS, Sherwin WBV, and Houlden BA (1999) Phylogeographical population structure of tiger quolls *Dasyurus maculatus* (Dasyuridae: Marsupialia), an endangered carnivorous mammal. *Molecular Ecology* 8: 1613–1625.

Janis C (1994) Do legs support the arms race in mammalian predatory/prey relationships? In: Horner J and Ellis L (eds.) *Vertebrate Behavior As Derived from the Fossil Record*. New York: Columbia Univ. Press.

Janis C and Wilhelm PB (1993) Were there mammalian pursuit predators in the Tertiary? Dances with wolf avatars. *Journal of Mammalian Evolution* 1: 103–125.

Jenkins PD, Kilpatrick CW, Robinson MF, and Timmins RJ (2005) Morphological and molecular investigations of a new family, genus and species of rodent (Mammalia: Rodentia: Hystricognatha) from Lao PDR. *Systematics and Biodiversity* 2: 419–454.

Jones G and Vanparijs SM (1993) Bimodal echolocation in pipistrelle bats: Are cryptic species present? *Proceedings of the Royal Society London, Series B: Biological Science* 251: 119–125.

Jones KE and Safi K (2011) Ecology and evolution of mammalian biodiversity. *Philosophical Transactions of Royal Society B* 366: 2451–2461.

Kingdon J (1989) *Island Africa*. Princeton, NJ: Princeton University Press.

Kisel Y, McInnes L, Toomey NH, and Orme CDL (2011) How diversification rates and diversity limits combine to create large-scale species–area relationships. *Philosophical Transactions of Royal Society B* 366: 2514–2525.

Luo Z-X, Cifelli RL, and Kielan-Jaworowska Z (2001) Dual origin of tribosphenic mammals. *Nature* 409: 53–57.

Macdonald DW (ed.) (2001) *Encyclopedia of Mammals*, 2nd edn. Oxford: Oxford University Press.

McKenna MC and Bell SK (1997) *Classification of Mammals above the Species Level*. New York: Columbia University Press.

Morell V (1996) New mammals discovered by new explorers. *Science* 273: 1491.

Novacek MJ (1992) Mammalian phylogeny: Shaking the tree. *Nature* 356: 121–125.

Novacek MJ, Rougier GW, Wible JR, McKenna MC, Dash-zeveg D, and Horovitz I (1997) Epipubic bones in eutherian mammals from the late Cretaceous of Mongolia. *Nature* 389(6650): 483–486.

Pagel MD, May RM, and Collie AR (1991) Ecological aspects of the geographical distribution of mammalian species. *The American Naturalist* 137: 791–815.

Purvis A, Fritz SA, Rodrigues J, Harvey PH, and Grenyer R (2011) The shape of mammalian phylogeny: Patterns, processes and scales. *Philosophical Transactions of Royal Society B* 366: 2462–2477.

Purvis A, Gittleman JL, Cowlishaw G, and Mace GM (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London Series B – Biological Sciences* 267: 1947–1952.

Raven PH and Wilson EO (1992) A fifty-year plan for biodiversity surveys. *Science* 258: 1099–1100.

Reeder DA, Helgen K, and Wilson DE (2006) Global trends and biases in new mammal species discoveries. *Occaisonal Pappers of the Museum of Texas Tech University* 269: 1–35.

Safi K, Cianciaruso MV, Loyola RD, Brito D, Armour-Marshall K, and Diniz-Filho JAF (2011) Understanding global patterns of mammalian functional and phylogenetic diversity. *Philosophical Transactions of the Royal Soceity B* 366: 2536–2544.

Sharman GB (1970) Reproductive physiology of marsupials. *Science* 167: 1221–1228.

Surridge AK, Timmins RJ, Hewitt GM, and Bell DJ (1999) Striped rabbits in Southeast Asia. *Nature* 400: 726.

Turvey ST and Fritz SA (2011) The ghosts of mammals past: biological and geographical patterns of global mammalian extinction across the Holocene. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366: 2564–2576.

Van Valkenburgh B (1999) Major patterns in the history of carnivorous mammals. *Annual Review of the Earth and Planetary Science* 27: 463–493.

Vaughan TA, Ryan JM, Czaplewski NJ, *et al.* (2000) *Mammalogy*, 4th edn. Philadelphia: Saunders College Publishing.

Wible JR, Rougier GW, and Novacek MJ (2005) Anatomical evidence for superordinal/ordinal eutherian taxa in the Cretaceous. In: Rose KD and Archibald JD (eds.) *The Rise of Placental Mammals: Origins and Relationships of the Major Extant Clades*, pp. 15–36. Baltimore and London: Johns Hopkins University Press.

Wilson DE, Reeder DM, *et al.* (1999) *Mammal Species of the World*, 2nd edn. Washington, DC: Smithsonian Institution Press.

Wilson DE and Reeder AM (2005) *Mammal Species of the World: A Taxonomic and Geographic Reference*, (2 vols), 2142 p. Baltimore: Johns Hopkins University Press.

Mammals, Conservation Efforts for

EJ Milner-Gulland, Imperial College London, London, UK
Sarah Durant, Zoological Society of London, London, UK, and Wildlife Conservation Society, Bronx, NY, USA
Rosie Woodroffe, Zoological Society of London, UK
Richard Young, Durrell Wildlife Conservation Trust, Trinity, Jersey

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by
E.J. Milner-Gulland and R. Woodroffe, volume 3, pp 811–824,
© 2001, Elsevier Inc.

Glossary

Conservation efforts Any action that aims to reduce the probability of extinction of a taxon over a specified time period.
***Ex situ* conservation** Conservation activities that involve individuals held outside their native habitats (e.g., in zoos or seed banks).
***In situ* conservation** Activities which aim to conserve wild populations in their native habitats.
Mammal Member of the order Mammalia. A species that provides milk for its young and has fur. They typically (but not exclusively) bear live young.
Persecution Deliberate killing of animals perceived to be a nuisance.
Range state A country in which an animal is found.
Sustainable use Exploitation of wildlife in a manner that avoids depletion of the resource (e.g., limited hunting, ecotourism).
Terrestrial A species that lives on land (as opposed to marine or fresh water) for the majority of its life cycle.

Conservation efforts for mammals must start with a consideration of the particular threats that face these species. Mammals are a disparate group, but similarities in their conservation needs can be discerned, particularly when considering large mammals. Large-bodied species tend to be vulnerable to extinction both because of their biology and their interactions with humans. They typically need large areas for survival and to be relatively slow growing, making their populations less resilient. They also tend to be more likely to be killed by humans for their meat, trophies, or because they are a danger or a nuisance. Because of these varied threats, conservation efforts for mammals are broader than for some other taxa; they encompass both habitat protection and direct protection. Conservation efforts for mammals are also noteworthy because these species are often treated as flagships for the rest of biodiversity conservation, sometimes at the expense of other species. This makes mammals important more broadly for conservation policy.

Introduction

In discussing conservation efforts for mammals, we first need to identify what is particular about their conservation needs. It is hard to generalize about such a broad range of species, from bats and mice through antelopes, manatees, and lynx to elephants and whales, but there are some common characteristics that predispose mammals to threat, at least among larger-bodied species. These characteristics relate both to natural history and to human attitudes.

Large mammals have high energetic demands. They grow slowly, range widely, and occur at low population densities. Together these linked characteristics mean that large mammals are vulnerable. Their population dynamics may be less resilient than those of other species, since they cannot bounce back quickly from population declines. Furthermore, they may be more affected by habitat loss than are other species, because they often have large home ranges, and require larger areas of habitat.

The natural history of mammals is not the only feature that predisposes them to vulnerability – large mammals are often under threat because humans deem them to be worth killing. Being large bodied, most are valuable for meat; those that have trophies such as horns, tusks, or furs can be extremely valuable. The economic value of mammals makes them vulnerable to overexploitation, but it can also be used as a tool for their sustainable use, and ultimately, their conservation. Large mammals can also be prone to human persecution because they represent a threat both to humans themselves and to their livelihoods; they kill livestock, or compete for land and resources. However, some mammals are charismatic, and some are considered iconic (e.g., tigers, elephants, and pandas), which means that people may be prepared to invest heavily in their conservation, and that they can generate substantial sums for conservation through tourism. For this reason, large mammals often act as flagships for the conservation of their ecosystems.

Large mammals can have a significant impact on their environments. This creates an intimate association between the conservation of terrestrial mammals and the conservation of other species in their ecosystems; conservation efforts targeted toward mammals might, therefore, have added value over and above the individual species being conserved. Some mammals have a profound impact on the ecosystems they

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00255-0

inhabit. For example, the reintroduction of wolves to Yellowstone National Park triggered a trophic cascade, in which wolf presence affected elk behavior, and through them the vegetation of the Park, increasing the amount of woody vegetation such as aspen and cottonwood. The reintroduction also reduced populations of the coyote, a smaller predator, which the wolves killed. Scavengers such as grizzly bears benefited from wolves, scavenging at wolf kills. It is difficult to predict in advance the outcome of the interacting effects of removing, or reintroducing, large mammals on their environment, but it is clear that these effects can be profound (Smith *et al.*, 2003).

In this article, we will first discuss in general terms some of the threats facing mammals, and which species are particularly vulnerable to these threats. We will then discuss conservation measures that have been used to address some of the threats. Finally, we will use three in-depth case studies to exemplify a range of species conservation efforts implemented in response to an array of threats operating from localized up to regional scales. We conclude by placing mammal conservation into the broader context of international conservation efforts.

Threats to Mammals

Like other taxa, many terrestrial mammals are at risk from habitat loss. As people become more numerous, they modify their surroundings in a variety of ways; this process creates new environments hostile to many wild species. Direct conversion of natural habitat threatens the high proportion of mammal species that cannot survive in farmed or urban environments. Many terrestrial mammals' large body size and high energy requirements make them particularly sensitive to human activities. Freshwater mammals such as river dolphins are also highly susceptible to habitat degradation and pollution.

Direct killing by people is a serious threat, which perhaps affects mammals more than other groups (with the possible exception of commercial fish species). Some forms of killing are due to exploitation. Subsistence hunting of deer, antelopes, and many primate species can cause population declines, as can commercial hunting to fuel local and international trade. The bushmeat trade (trade in meat from wildlife for local consumption) is predominately based on mammal species; the annual offtake for this trade has been estimated to be 2.2 million tons of meat a year in the Congo basin alone (Fa *et al.*, 2003). There is a strong interaction between hunting and habitat loss; as an area becomes more accessible through the creation of logging roads, for example, hunting rates increase. The extinction of the Yangtze river dolphin has been attributed to a combination of targeted hunting, direct mortality from entanglement in nets, disturbance, collisions with boats, and habitat degradation. Likewise, the presumed extinction of the kouprey, a wild ox, has been linked to intense hunting of large mammals in Indochina. Hunting for the international trade has caused some of the most serious and well-publicized species declines and includes hunting of elephants to make ivory carvings, spotted cats to make fur coats and whales for oil, meat, and baleen. Of particular concern, in recent years, has been the heavy demand for animal body parts used in traditional Asian medicines; this has caused declines of tigers (for their bones and other body parts), rhinoceros (for their horn), and bears (for their gallbladders).

Not all killing of mammals by people constitutes exploitation; several species, especially the large carnivores, have suffered serious declines through persecution. Large-bodied species such as elephants and bears are a risk to human life and are rarely tolerated in areas with high human densities. Much of the problem with human–elephant relationships in recent years is concerned with the fact that elephants are dangerous neighbors. Elephants can flatten a crop in minutes and are responsible for many human injuries and fatalities each year. Likewise, large carnivores are feared as a threat to both people and livestock, and have been the victims of organized eradication campaigns that saw, for example, the extirpation of wolves from most of the United States, and African wild dogs across much of Africa, in the first half of the twentieth century. Humans may also cause inadvertent declines through exploitation or control of relatively common species. Traps set to catch rabbits are a serious threat to endangered Iberian lynx, which are accidentally captured and killed. Extermination campaigns for prairie dogs inadvertently caused the collapse and near-extinction of the black-footed ferret.

Human activities also influence wild mammals' exposure to "natural" threats such as disease and predation. Frequent local extinctions of bighorn sheep in the Western United States, leading to serious species decline, have been largely attributable to infectious diseases contracted through contact with domestic sheep. Small, isolated populations are especially vulnerable to the threat of disease; thus disease may deliver the final *coup de grace* to species already crippled by other factors. Likewise, these populations may be unable to cope with the predation pressures which can be sustained by larger populations. When exotic predators or pathogens are introduced to naive species, the impact can be more severe still. The introduction of red foxes to Australia decimated populations of the smaller marsupial species, and the introduction of rinderpest – a viral disease of Asian cattle – to Africa caused widespread and often irreversible declines of many African ungulates.

The regular updating of the World Conservation Union's (IUCN) Red Lists of threatened species provides an opportunity to analyze the relative importance of different threatening processes. Schipper *et al.* (2008) showed that 25% of mammal species were threatened. Only two other taxonomic groups had then been fully assessed and so are directly comparable, birds (14% threatened) and amphibians (32% threatened). Schipper *et al.* identify six major threat types. For mammals, the top two threat types are habitat loss (37% of the threats) and overexploitation (17%). Other threats, such as introductions (6%) and pollution (4%), are less often serious threats for mammals. Thus any analysis of the threats to mammals and the conservation efforts needed to counteract them is likely to emphasize habitat loss and overexploitation.

Another useful technique for looking at the relationship between threat and extinction risk compares the characteristics of populations that have become extinct with those that survive. These studies have shown that vulnerability correlates with species' natural history, but that this interacts with

anthropogenic factors. For example, carnivores with large home ranges frequently travel beyond the borders of protected areas, where they come into contact with people. Such contact is often fatal, with the result that wide-ranging species require larger parks and reserves for effective protection (Woodroffe and Ginsberg, 1998). The direct and indirect effects of human activity on mammals lead to general associations between high human population density and local extinction of mammal populations. Elephants, primates, and carnivores all demonstrate such associations. This probably reflects complex interactions between direct killing of mammals by people, destruction of mammals' habitat by people, or, on the small scale, simple avoidance of people by mammals. Price and Gittleman (2007) showed that artiodactyls (even-toed ungulates) are more likely to be threatened if they live in areas of high poverty, and that if a species is also hunted, then slower-growing species are more likely to be threatened. The effect of living in poor areas is probably related to the degree of human use of wildlife habitat and habitat clearance that is going on, whereas the effect of the species' biology is probably because slow-growing species are less resilient to high levels of hunting.

In summary, individual species may be threatened by a number of different factors, so that the type of conservation effort needed to counteract these threats can also vary widely between species. However, it is possible to discern some broad patterns that can help to identify species that may be at particular risk of extinction. These patterns relate both to the biology of the species and to the level and type of human pressure on the species and its habitat.

Types of Conservation Effort

The threats described above demand a variety of conservation measures. We discuss these measures briefly, giving more detailed evaluations of their costs and benefits in the case studies.

Legal protection of areas is the most traditional form of conservation; this approach, if effectively enforced, has the advantage of protecting habitats, as well as particular species, against a multitude of threats associated with human activities. While reserves are often targeted at a particular species, and even named as such (e.g., Addo Elephant Park, Gemsbok National Park), well-planned reserves can protect multiple species, habitats, and landscapes. Large mammals are often used as "flagships" to promote such reserves (e.g., India's Project Tiger). Craigie *et al.* (2010) analyzed mammal population trends in African protected areas over the period 1970–2005, and showed an average 59% decline in species' population sizes within protected areas, which suggests that overall, protected areas are not enough to conserve Africa's mammal populations. However, they uncovered major regional differences; in southern Africa populations were generally stable, while severe declines were recorded in West Africa. This suggests that a better understanding of what makes protected areas successful in some areas and not in others could markedly improve conservation effectiveness.

Legal protection of species or populations may also be used as a conservation measure. For example, stringent legislation in the United Kingdom protects the European badger from killing and disturbance. International treaties have effectively protected many cetacean species from intense overhunting, and led to population increases in previously highly threatened species such as the bowhead whale (George *et al.*, 2004). Difficulties of enforcement mean that such legislation may have limited value as a conservation tool in countries without effective governance, however; for example, a wide range of mammals is still openly traded for meat in the Democratic Republic of Congo despite having full legal protection (Rowcliffe *et al.*, 2004). Species threatened by exploitation may be protected by *legislation governing trade* rather than preventing killing. For example, it is not illegal to kill leopards in some countries, such as the Gambia, but international trade in their skins is prohibited from or to member countries of the Convention on International Trade in Endangered Species (CITES).

Problems surrounding the enforcement – and ethics – of legislation prohibiting the exploitation of wild populations or use of protected areas have led to initiatives aimed at conservation through *sustainable use*. The rationale behind such initiatives is that sustainable use should ensure that local people value and protect natural resources, if the benefits accrue to local communities. Like the establishment of reserves, this approach has attracted controversy; the array of political, legislative, and logistical concerns that it entails is different from, but no less complex than, that surrounding more traditional legal protection. At the international level, sustainable use and equitable sharing of the benefits from biodiversity are at the heart of the Convention on Biological Diversity.

Threatened populations may also be protected through direct management, particularly when the population has become very small and unlikely to survive in the wild without intervention. *Habitat management* involves measures such as provision of waterholes or den sites, or control of predators (both native and invasive). *Population management* describes a variety of interventions such as vaccination against infectious diseases and supplementation from other populations, wild or captive. Species conservation may also involve reestablishment of extinct populations through reintroduction of captive-bred or wild-caught mammals. Such measures have been important in the recovery of some endangered species (e.g., golden lion tamarins and red wolves), as well as for the reestablishment of populations of species that remain widespread but have suffered local extinctions (e.g., gray wolves).

Finally, the conservation of some endangered species may benefit from *captive breeding*, maintaining a population in captivity that may or may not be intended for release into the wild. Such efforts have proven valuable for critically endangered species reduced to their last few individuals (e.g., Arabian oryx, black-footed ferret, and island fox), in which managed breeding can minimize the loss of genetic variation. However, captive breeding is a choice of last resort, as it can lead to the loss of wild behaviors. Animals that are habituated to people are particularly vulnerable to persecution or exploitation once reintroduced. Przewalski's horse, however, provides an example to show that all is not lost when species are reduced just to a few individuals in captivity. In this case, the individuals were not even genetically pure, with their genepool having been contaminated by domestic horses.

However, the species has now been reintroduced to the wild and is displaying natural behavior (Boyd and Bandi, 2002).

Case Studies

Mammals cover such a range of habitats, biology, and interactions with humans that it is very difficult, in three short case studies, to give a flavor of the status and conservation efforts for all mammal species. However, we have chosen three case studies that exemplify many of the issues, covering a variety of species and continents; the saiga antelope in Central Asia, hutias and solenodons (small mammals found on the islands of the Caribbean), and the cheetah and wild dog across their range in Africa. Major omissions include marine mammals and the bats, but we hope that the insights from our case studies shed light on some of the issues facing other mammals as well.

Saiga Antelopes

The saiga antelope (*Saiga tatarica*) has had a torrid time in the past two decades. Buffeted by geopolitical forces – the breakup of the Soviet Union and high Chinese demand for its horn – its population crashed in the first half of the 2000s, leading to the dramatic uplisting of the species on the IUCN Red List from Near-threatened to Critically Endangered, the highest threat level, in one step. Since the mid-2000s, however, substantial conservation efforts have been mobilized on every scale, and the species now represents a conservation success story, despite the continued threats which it faces.

Background

The saiga antelope (**Figure 1**) is a nomadic herding species found in the semiarid rangelands of Kazakhstan, Russia, and Mongolia (**Figure 2**). It has two subspecies, *S. t. tatarica* and *Saiga t. mongolica*; the latter is found only in Mongolia. The Kazakhstan populations of the species migrate over very long distances, up to 1000 km, moving between winter pastures in southern desert regions and summer pastures in the steppe zones of the north. This enables them to track green vegetation and to escape from the heaviest snows of the winter (Bekenov *et al.*, 1998). Within its summer and winter ranges, the species is nomadic, a response to the harsh climatic conditions of the continental ecosystem in which it lives, allowing herds to move rapidly away from areas where there is bad weather or where other threats exist.



Figure 1 Picture of a saiga antelope. The saiga is about the size of a domestic goat, with a sandy-colored upper body and a creamy underside. This picture shows an adult male, which has horns of an unusual translucent amber color. The species' most striking feature is a protuberant nose, which swells further in rutting males. Picture by Navinder Singh.

The species stops to give birth in huge aggregations in the spring, which is thought to be a predator-swamping adaptation, protecting the calves from wolf predation. Females twin consistently once they reach maturity. This is unusual among ungulates and means that the population can increase very rapidly in good years, allowing populations to recover quickly from overhunting, harsh winters or disease outbreaks. It has a harem breeding system, in which each adult male controls and mates with a group of 12–30 adult females, making the population resilient to male-biased hunting, which is important because only adult males bear horns, which are highly prized in Chinese medicine. Overall, therefore, the saiga antelope is a resilient species that is capable of withstanding relatively heavy hunting pressure and of recovering quickly from episodes of high mortality.

The saiga has been hunted for its meat, horns and hide since prehistoric times. In the eighteenth and nineteenth centuries, it was hunted in large numbers by the St Petersburg Imperial court. By the early twentieth century, hunting had reduced it to near-extinction. Horn prices were very high, with horns exported for use in Chinese medicine. During the Soviet period, up to the mid-1990s, the population was well managed, with legal protection and regulated commercial hunting, and grew to relative stability. However, the situation changed dramatically with the breakup of the Soviet Union. Rural poverty and unemployment, combined with the opening of the border with China and high prices for horn, as well as a collapse in law enforcement and management, led to massive poaching of the species, causing a 95% population decline in just 10 years and triggering its listing as Critically Endangered on the IUCN Red List in 2001 (**Figure 3**). Since conservation efforts started, around 2003/2004, the global population of the species has started to rise steadily, although it is still well below its 1980s level.

Threats

The major threat that has affected populations of *S. t.tatarica* over the past 20 years has been massive illegal hunting for both horns to supply international markets and meat for local consumption, causing the population decline shown in **Figure 3**. However, limiting factors in the past have included habitat degradation from overgrazing, infrastructure development, particularly canals in which large numbers drowned, and epidemics of disease, probably contracted from the livestock that share the saiga's range (Bekenov *et al.*, 1998). It is important not to forget these other factors when focusing on the main threat of poaching, as was demonstrated in May 2010, when 12,000 saigas were found dead in the Ural population after several years without any sign of disease. The dead saigas were adult females which had just given birth, and although the ultimate cause of death is still not clear, this was



Figure 2 Map of the range area of the saiga antelope, from Milner-Gulland *et al.* (2001). There are four distinct populations of *Saiga t. tatarica*, shown as stippled areas. Three of these are in Kazakhstan (2: Ural, 3: Ustiurt, 4: Betpak dala), the fourth in the pre-Caspian area of Russia (1). The other subspecies, *Saiga t. mongolica*, has only a few thousand individuals and is found in Mongolia (5).

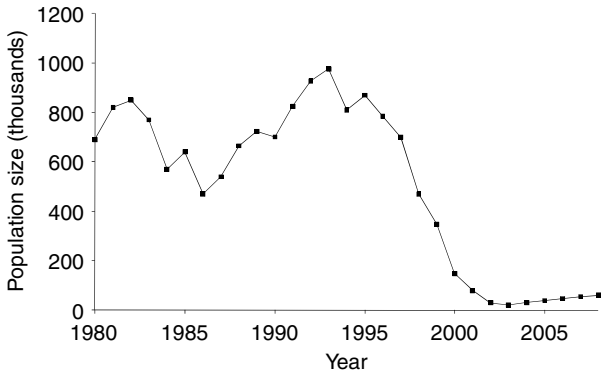


Figure 3 The population trajectory of the saiga antelope in Kazakhstan, 1980–2010. Data are from the Institute of Zoology, Kazakhstan, and are predominately from aerial censuses. In the 1980s, the population was heavily hunted, and population size also fluctuated due to environmental changes from year to year.

a warning that disease is a potentially potent threat; this single week of mortality killed one-third of the estimated Ural population. Looming future threats include the effects of climate change on vegetation productivity and temperatures, and development of the extensive steppe regions where the species lives, particularly for gas and mineral extraction.

Conservation Efforts

The fact that the main threat to saigas over the past two decades has been hunting means that engaging with people is the only way to address this urgent threat. Conservation efforts have therefore focused on strengthening law enforcement, building public awareness of the problem and changing attitudes among hunters and consumers, locally and in China. A few projects have also offered alternative livelihood opportunities to local people, but these have been small-scale and it is hard for a conservation project to counteract the momentum of massive social and economic changes at the national and regional levels. Law enforcement has generally been carried out by governments, with financial and logistical support from large international NGOs, whereas the public engagement activities have been carried out by in-country conservationists in collaboration with smaller, more species-specific NGOs.

The saiga antelope is a good example of how international treaties can guide and integrate conservation efforts toward effective action. In 2006, saiga conservationists from government, NGOs and international bodies met under the Convention on Migratory Species (CMS) to sign a Memorandum of Understanding (MOU) on saiga conservation, and to endorse a Medium Term International Work Programme, which prioritizes conservation actions according to their urgency and their importance for preventing extinction, both for the species as a whole and for individual populations. In 2010, the parties met again to evaluate progress and to agree a new work programme for the next 5 years. Often international meetings can be "talking shops" with little relevance to on-the-ground conservation, but the evidence suggests that the CMS saiga MOU has been very effective in channeling conservation activities into agreed high priority areas, holding all those who

carry out saiga conservation to account, and promoting communication, cooperation, and transparency (CMS, 2010).

Conservationists are continuing to work with local people in the saiga's range states, and to support law enforcement activities. However, the emphasis is shifting, and it is now important to evaluate the lessons learnt from the successes and failures of conservation interventions since 2003, for example, whether local people are actually changing their attitudes in response to public awareness campaigns (Howe *et al.*, 2011). Looking to the future, Kazakhstan is developing a landscape-scale system of protected and multiple-use areas for one of its saiga populations, in order to enable people and saigas to coexist in the steppe ecosystem sustainably, in the face of development pressures and climate change. There is also an increasing emphasis on engaging with China, both at the governmental level and through NGOs and consumer groups, toward a long-term solution to meeting the demand for saiga horn, which otherwise will continue to act as an irresistable force incentivising poaching.

Urgent conservation intervention and a lot of publicity for the saiga's plight were needed in order to ensure that the saiga did not crash to extinction as a result of the collapse of the Soviet Union. Relatively small-scale conservation interventions by dedicated individuals and organizations on the ground were able to support the species through the dark times; now that the range states' circumstances are more stable and they are keen to conserve their biodiversity, it is possible to plan for the future of this species.

The Hutias and Solenodons of the Caribbean

Before humans colonized the Caribbean islands some 6000 years ago, a diverse mammalian fauna of around 120 endemic species of primate, rodent, sloth and soricomorph (shrews and relatives) occurred in the region. The vast majority are now extinct, representing perhaps the highest rate of extinction of any mammalian fauna in the past few thousand years. Only 13 species of hutia (a family of the order Rodentia) and two species of solenodon (a family of the order Soricomorpha) survive today on the islands of Hispaniola, Cuba, Jamaica and the Bahamas. They are an extremely threatened group; all but two are globally threatened, with 10 classified as Endangered or Critically Endangered on the IUCN Red List of threatened species.

Background

The only surviving Caribbean soricomorphs are the Hispaniolan solenodon *Solenodon paradoxus* (**Figure 4**) and the Cuban solenodon *Atopogale cubana*, the last two representatives of the Solenodontidae family. These species have traditionally been placed in the same genus (*Solenodon*), but recent research suggests they diverged from each other around 25 million years ago and should be placed in separate genera. The original solenodon lineage separated from the ancestors of all other living mammals an astonishing 76 million years ago (Roca *et al.*, 2004), before the extinction of the dinosaurs. This ancient mammalian heritage is reflected in some unique characteristics, most notably the solenodon's deep grooves in its incisors, which act as channels to inject venomous saliva into their prey (from the Greek 'soleno' meaning channel or pipe and "don" meaning tooth).



Figure 4 Picture of the Hispaniolan solenodon *Solenodon paradoxus*. Around 1 kg in weight, this adult solenodon resembles a giant shrew with tiny eyes, an elongated and mobile snout and large feet which it uses for digging for invertebrate prey and for burrowing. Picture by Jose Nunez-Mino, Durrell Wildlife Conservation Trust.

The Hispaniolan solenodon lives in the dry, moist, and pine forests of south-western, central and eastern Dominican Republic and is also known from the Massif de la Hotte in Haiti (**Figure 5**). Adult Hispaniolan solenodons weigh around 1 kg, and resemble large, well-built shrews with an unusual side-to-side gait. They are mainly nocturnal, sheltering during the day in rock crevices or burrows. Solenodons can often be found in family groups usually made up of between two and four individuals, consisting of an adult pair and one or two offspring. The Cuban solenodon, once widely distributed across Cuba, is now restricted to the eastern end of the island and is extremely hard to find (**Figure 6**). It has been recorded from the Sierra del Cristal and Alexander Humboldt National Parks (Fa *et al.*, 2002). It shares many characteristics with the Hispaniolan solenodon, but has a darker coat and a striking white mane of fur. Its exact conservation status is not clear and research is urgently needed to better understand its ecology, main threats, and conservation requirements.

The hutias (*Capromyidae*) have an ancient heritage in the Caribbean with the oldest hutia fossils from around 18 million years ago. Many hutia species were lost following Amerindian and European arrival in the Caribbean, including some which have disappeared within the past century. With one species surviving in each of Jamaica (*Geocapromys brownii*), Hispaniola (*Plagiodontia aedium*, **Figure 7**), and the Bahamas *Geocapromys ingrahami*, most hutia species are found on the main island of Cuba and its surrounding islands (**Figures 5** and **6**).

The hutias resemble the cavies of mainland South America and vary in size from the rat-sized dwarf hutias (four *Mesocapromys* spp.) to the Desmarest's hutia *Capromys pilorides* which can weigh up to 7 kg. Some species have prehensile tails to facilitate their almost entirely arboreal lifestyle, whilst other species, such as Ingraham's hutia from the Bahamas, are essentially terrestrial. Hutias are herbivorous and can be found in a wide range of habitats, from small and arid coastal cays to high elevation moist forest. In comparison with other rodents,

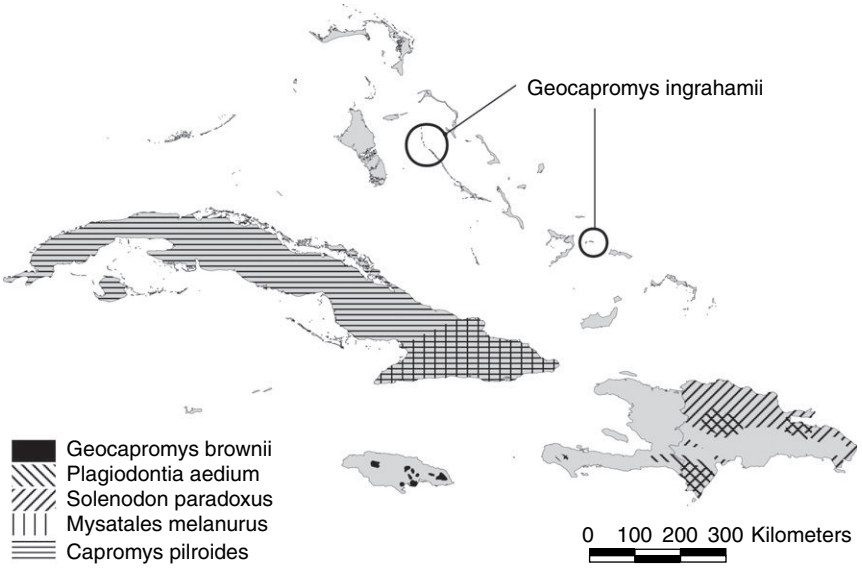


Figure 5 Map of the geographic ranges of five of the hutias and the Hispaniolan solenodon *Solenodon paradoxus* within the Caribbean. The species' ranges are based on data from the IUCN Red List of Threatened Species (IUCN, 2010).

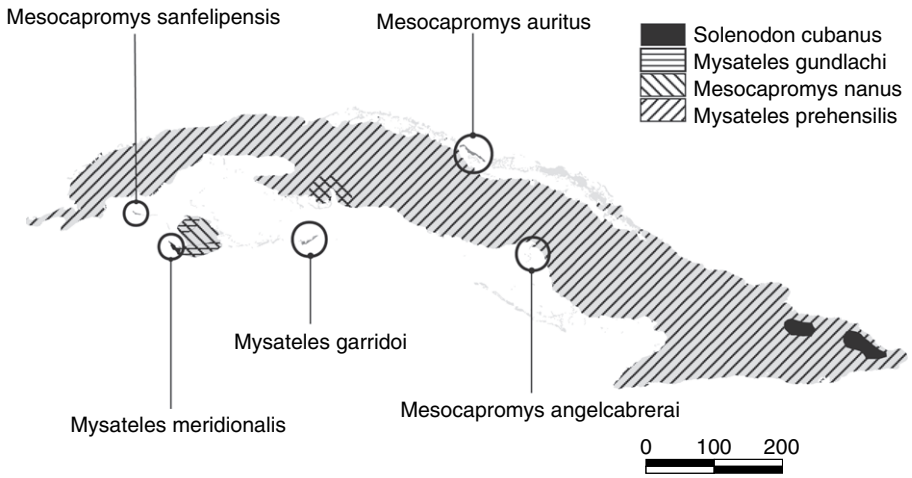


Figure 6 Map of Cuba, showing the geographic ranges of *Solenodon cubanus* and the eight Cuban hutias. The species' ranges are based on data from the IUCN Red List of Threatened Species (IUCN, 2010).

hutias have very low reproductive rates and typically give birth to one or two young per litter; nevertheless in well protected areas some species can reach very high densities. For example, in Guatanamo Bay in Cuba, Desmarest's hutia, the least threatened of all hutias, reaches densities of more than 1000 animals per square kilometer and its population is now managed to reduce impacts on human infrastructure. With high density populations and good tasting meat, hutias were a very important source of food for Amerindians as they colonized the region, and they are thought to have been kept in semidomestication and transported between islands as a food source. The first meat that Columbus tasted on his arrival to the New World was probably hutia.

Compared to many mammals, the ecology and status of solenodons and hutias is very poorly known. This constrains the prioritization of conservation resources and the design of targeted species-specific conservation actions and habitat protection measures. At various times during the past century, a number of the solenodon and hutia species have been declared extinct and then periodically "rediscovered" when animals were searched for or opportunistically found dead by scientists. However, three of the species now seem very likely to be extinct, and surveys are urgently needed to confirm their status.

Threats

The earliest major human impact on the Caribbean land mammals was hunting for food by Amerindians, which is likely to have caused a number of species extinctions. Hunting hutia for food has carried on until today, although the extent and population impacts aren't precisely known. In Haiti, there is some evidence that Hispaniolan solenodons are also



Figure 7 Picture of the Hispaniolan hutia *Plagiodontia aedium* showing off its arboreal skills. This species can be found in a range of habitats from mangroves and coastal dry scrubland right up to montane cloud forest. Picture by Jorge Brocca, Sociedad Ornitológica de la Hispaniola.

opportunistically hunted (Turvey *et al.*, 2008). Other major threats to the solenodons and hutias since European colonization of the Caribbean are habitat loss and predation or competition by invasive mammals. In some countries, habitat loss over the past few centuries has been severe. For example, only 3% of Haiti's forests now remain, which has undoubtedly resulted in massive range contraction of the Hispaniolan solenodon and hutia within this country. Although much more forest remains in the Dominican Republic, Cuba, and Jamaica, it is highly fragmented, isolating small populations of both hutias and solenodons and making them more vulnerable to other threats. Since Columbus arrived in the Caribbean in the 1490s, a number of nonnative mammals such as brown and black rat, feral cat, and dog, and small Indian mongoose, have been introduced into the region with devastating consequences for biodiversity in general, and Caribbean land mammals in particular. The threat of invasive species is particularly pressing for those hutia species which are restricted to small islands around Cuba. This is highlighted by the extinction of the Little Swan hutia in the 1950s following the introduction of cats onto Little Swan Island, Honduras. However, even for more wide-ranging species, invasive mammals may be having a long-term detrimental impact which is typically difficult to demonstrate. For example, there is evidence of widespread predation of Hispaniolan solenodon by domestic dogs in the increasingly fragmented forests of the Dominican Republic, possibly leading to population extinctions in small habitat patches near human habitation.

Conservation Efforts

Conservation of the solenodons and hutias requires a combination of habitat protection, invasive species control and species-specific measures, such as raising awareness of the impacts of domestic dogs on these species in order to change husbandry practices and reduce predation rates. In the Dominican Republic, Cuba, and Jamaica many of the key habitats for solenodons and hutias fall within protected areas. However, management effectiveness of these areas is patchy with some high-profile national parks suffering from extensive illegal timber extraction, agricultural encroachment, and fires. There are also likely to be many undocumented populations of solenodons and hutias occupying smaller patches of habitat that occur outside of protected areas, which also require protection.

The control of invasive species, including their eradication from smaller islands, is a conservation priority in the Caribbean and it is urgently needed to save the most threatened hutia species in the islands off the coast of Cuba. For example, the Isla De La Juventud Tree hutia (*Mysateles meridionalis*), restricted to the Isle of Pines in Cuba, has declined dramatically through hunting and the impacts of rats. This species has not been seen by scientists in over 25 years and is likely to be on the brink of extinction. With possibly four hutia species on Cuban islets numbering fewer than 250 individuals, captive breeding represents an important tool for saving these species. Solenodons are difficult to keep and breed in captivity but hutias have proved more amenable, with Desmarest's hutia and the Jamaican hutia widely kept by zoos. Ingraham's hutia in the Bahamas has been captively bred and successfully translocated from its last remaining site of East Plana Cay to other cays in the archipelago in order to reduce its extinction risk.

Finally, another major threat to the Caribbean land mammals is a lack of awareness regionally and globally of their diversity, importance, and conservation requirements. They are a classic example of species which have been largely overlooked by the international conservation community and, if they are to be saved, it is vital that they are pushed up the conservation agenda.

Cheetah and African Wild Dogs

Background

The cheetah (*Acinonyx jubatus*) is Africa's most threatened large cat. The species has declined dramatically over the past century. Historically the cheetah was widely distributed throughout Africa, except for the central and western forest belts, and east through to India. Today the species has completely disappeared from most of its Asian range, only tiny populations remain in north and west Africa, and its main strongholds are concentrated in southern and eastern Africa (IUCN/SSC, 2007; IUCN, 2010). The species is classified as Vulnerable by the IUCN, with the Saharan and Asiatic subspecies listed as Critically Endangered, the latter persisting in a small remnant population in Iran. Cheetah are also listed on Appendix 1 in

CITES and are classified as an endangered Appendix 1 species by CMS.

African wild dogs (*Lycaon pictus*), though not closely related to cheetah, show a similar geographic distribution and face similar threats. Historically wild dogs were found throughout Africa south of the Sahara, absent only from the lowland forests of the Congo basin and the driest deserts. Today, the species resides in a small fraction of this historic range, mostly in southern and eastern Africa – only a handful of tiny populations remains in West Africa. The species is considered Endangered by the IUCN, and is listed on Appendix 2 of CMS. Both cheetah and wild dogs are now the subject of a range-wide conservation planning strategy, launched in 2007, which is an excellent example of conservationists addressing conservation challenges across a large scale.

Cheetah and wild dogs depend on small to medium sized antelopes for food. However, they are unusual among large carnivores in that natural populations are not limited by prey but by other large carnivores, particularly lions and spotted hyaenas (Durant *et al*., 2010; Creel and Creel, 1996). These predators can kill adults and young, and also steal kills. Up to two-thirds of cheetah mortality in the Serengeti has been attributed to predation by larger carnivores, and this is also the most important cause of natural mortality among wild dogs. Perhaps as a consequence of avoiding larger carnivores, both cheetah and wild dogs range very widely, with home ranges up to 3000 km^2 having been documented (Marker *et al*., 2008; Fuller *et al*., 1992). Despite ecological similarities between cheetah and wild dogs, the two species have profoundly different social organizations. Female cheetah are solitary with overlapping home ranges, while male cheetah are often social, living in egalitarian coalitions of usually two to three males. Male cheetah can also hold territories, which are usually substantially smaller than female home ranges; in the Serengeti, male territories are on average less than 10% of the size of female home ranges. This combination of semisociality and ranging patterns has not been documented in any other mammal, and may be unique. It is thought that this unusual natural history has evolved as it enables cheetah, particularly females with vulnerable cubs, to access nomadic prey, whilst avoiding larger and more aggressive competitors such as lions and spotted hyaenas (Durant *et al*., 2010). In contrast with cheetah, wild dogs are intensely social. Pack members cooperate to hunt, to defend their group territory, and to raise pups. Although reproduction is usually restricted to a single socially dominant pair, all pack members contribute to care of the pups and also feed the mother when she is confined to the den. Wild dogs have large litters, and the energetic demands of reproduction are consequently high; females are rarely able to raise pups alone, and larger packs experience higher reproductive success. Most young wild dogs leave the packs where they are born to seek breeding opportunities elsewhere, travelling as small groups of either males or females. These dispersal groups can travel hundreds of kilometers in search of mates and territories, and are sometimes sighted in regions where no resident populations have occurred for decades.

These features of cheetah and wild dog ecology place constraints on their abundance, with the result that natural densities of the two species in good habitat are usually substantially lower than those of other large carnivores such as lions and hyenas. This has serious implications for the conservation of the two species. It is well understood that huge areas are required to conserve large carnivores, such as lions, tigers, and wolves; however, cheetah and wild dogs require conservation planning on an even larger scale. This means that protected area systems are unlikely ever to be sufficient for cheetah and wild dog survival. Indeed, today, only around a quarter of the world's remaining cheetah and 40% of the remaining wild dogs occur inside protected areas (**Figure 8**).

Threats

The main threats to cheetah and wild dogs today are habitat loss and fragmentation; habitat degradation; and retaliatory killing due to conflict with people and their livestock (IUCN/SSC, 2007). Historically, the captive trade was likely to have paid a strong role in the decline of cheetah. Cheetah were captured from the wild and used for coursing across much of Asia and the middle East. Akbar the Great reportedly kept a stable of 3000 captive cheetah in India. However, cheetah rarely bred in captivity, and the demand for cheetah is likely to have created a constant drain on wild populations. This was probably responsible for the massive contraction in cheetah range across Asia and northern Africa. Cheetah are rarely used for hunting today, but there are concerns that trade in cheetah may be increasing again, as the species is increasingly in demand for private collections (IUCN/SSC, 2007). Wild dogs are seriously affected by accidental snaring; they readily become captured in snares set for ungulates, and the resulting mortality has caused decline of several populations. Infectious disease – especially rabies – has also attracted attention as a threat to wild dogs, largely due to its role in the loss of wild dogs from the Serengeti in 1991.

Conservation Efforts

Cheetah and wild dog survival depends on ensuring the availability of large areas of suitable habitat, including protected areas as well as private or community areas where land use practices are compatible with the species' presence, such as livestock or game ranching. However, because cheetah and wild dog densities are low, it is critical that connectivity is maintained between smaller patches ($<10{,}000\ km^2$) of suitable habitat. Therefore transboundary cooperation, and the incorporation of the requirements of cheetah and wild dogs in national land use planning, are critically important. In order to address these challenges, a range-wide conservation planning process for cheetah and wild dogs was launched in 2007, followed by a conservation implementation phase launched in 2008. The planning process was designed to address the large area requirements of both species, and engaged national governments in a regional planning process to develop conservation strategies.

The cheetah and wild dog conservation planning process has two key themes: (1) an assessment of the current state of knowledge for both species; and (2) developing conservation strategies for the species. These themes are addressed in regional workshops in a participatory and consensual process which involves all key stakeholders, including national wildlife authorities, NGOs and species biologists. Participants in regional workshops are able to draw on at least some information on the effectiveness of particular tools for conserving



(a)



(b)

Figure 8 (a) Picture of a female cheetah with cubs. The cheetah is a large cat, built for speed, with long legs, small head, and flexible spine. Cheetah litters usually number from three to four cubs at birth, but can number up to seven. Because of high mortality, large litters of older cubs are unusual. Picture by SD. (b) Picture of African wild dogs interacting. Wild dogs are coursing predators which may chase their prey for several kilometers. This picture shows three members of a pack engaged in a "rally," an enthusiastic greeting ceremony which precedes every hunt and illustrates the intense social bond which exists among pack members.

the two species, such as evaluations of methods for monitoring cheetah, and for addressing threats to wild dogs. To collate information on distributions a mapping process is used, based on one first used for jaguars by the Wildlife Conservation Society (Sanderson *et al.*, 2002). Workshop participants map current knowledge of the extent of range for cheetah and wild dogs. This includes mapping areas where the species is known to occur and areas where it is known to be extirpated, as well as areas where information is unavailable (**Figure 9**). This then leads to discussion and agreement of a regional plan of action.

In both eastern and southern Africa, strategies identified the need to conserve cheetah and wild dogs outside protected areas, including finding ways to increase tolerance of these carnivores by local communities. Fortunately, unlike some other large carnivores, such as lions and leopard, cheetah and wild dogs do not pose a direct threat to people. However, both species can take livestock, and this causes conflict with livestock farmers. Moreover, the increasing popularity of game farms, particularly in southern Africa, can result in high levels of conflict as in these systems, farmers are ranching the natural prey of cheetah and wild dogs. In some areas, particularly those bordering protected areas, the appeal of the two species can be used to attract wildlife tourism, providing economic benefits to local communities. However, where this is not possible it is critical to reduce livestock loss and raise awareness of the importance of the survival of cheetah and wild dogs outside protected areas. Where meat from local livestock is exported, then accreditation schemes, such as cheetah friendly meat, may play an important role. This involves meat from farms which support cheetah commanding a premium from environmentally conscious consumers.

Regional and national workshops raise awareness of the conservation needs of the species, and enable the agreement of a plan of action by a range of stakeholders. However, there is a danger that planning exercises remain academic exercises; good on paper but never implemented. The cheetah and wild dog planning process is unusual because the planning is followed by a post-strategy implementation phase. This enables national governments to develop their own national conservation action plans for cheetah and wild dogs in line with a regional strategy, facilitating transboundary cooperation. Regional coordinators have been appointed to ensure that national action plans are established across cheetah and wild dog range following on from the regional strategies, and that they are implemented.

By the end of 2010, national plans had been developed in 10 of the 13 known cheetah and wild dog range states in Eastern and Southern Africa, covering 72% of current known cheetah range. Regional offices had been established in eastern and southern Africa, helping overstretched wildlife authorities to prioritize cheetah and wild dog conservation over a plethora of competing interests. The project has focused on improving capacity for cheetah and wild dog conservation within the range states, and acts as a model for the conservation of other wide-ranging terrestrial species.

Themes of the Case Studies

One of the main overarching themes from all three case studies is the paramount need to consider human needs and incentives (financial or otherwise) when designing conservation interventions. Human activities, whether direct exploitation or habitat conversion, are at the root of the

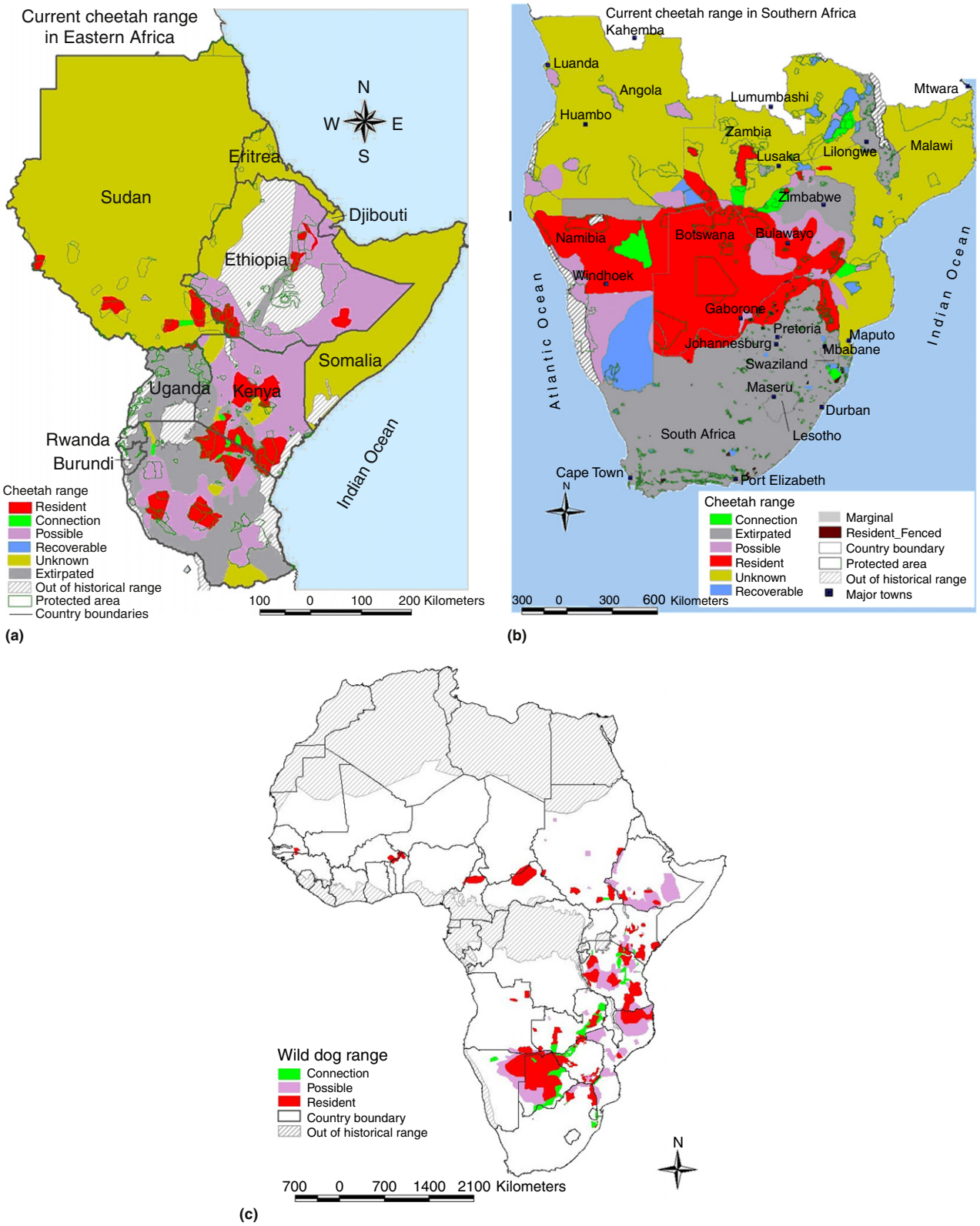


Figure 9 Map of the range areas of (a and b) cheetah and (c) wild dog in eastern and southern Africa in 2008, based primarily on the information provided at the planning workshops (reproduced from IUCN/SSC (2007) *Regional Conservation Strategy for the Cheetah and African Wild Dog in Eastern Africa.* Gland, Switzerland: IUCN Species Survival Commission).

problems that the case study species face. Changes in attitudes and improved knowledge of the needs of the species can lead to changes in behavior by people living in the area, which is required if these species are to have a sustainable future in the long term. However, change in the behavior of individual people is not enough. External events, such as the break-up of the Soviet Union for saigas, or the ongoing poverty and disasters faced by Haiti, are not easily addressed by projects targeting particular species. Similarly, the future of wide-ranging species such as the cheetah and wild dog can only be secured if conservation is planned and implemented at a landscape scale, involving the participation of multiple range states. In the case of the saiga, relatively small-scale conservation efforts at the local level saw the species through the dark times until the economic and social situation stabilized and larger-scale conservation efforts were able to step in.

In the case studies which we have discussed, the focus of conservation efforts has been on single species. However, in the case of both saigas and cheetah and wild dogs, the species are being used as flagships for wider landscape-scale conservation. This leads on to the necessity for an ecosystem approach to conservation, in which the interlinkages between species and their habitat are at the forefront. Perhaps the conservation of mammals has retained its single-species focus for longer than in other taxonomic groups because of their charismatic nature, and the potential that they have as flagships. Even the small and cryptic hutias and solenodons are potentially flagships for the conservation of their habitat due to their ancient lineage, fascinating biology, and odd appearance.

The case studies also highlight the importance of understanding the biology of the system and of the species themselves in order effectively to plan conservation actions. In the case of the Caribbean mammals, one of the main impediments to conservation is that the species are very poorly known; there are a number of these species which may or may not be extinct, their distributions are not well understood and very little is known about their ecology. In these circumstances, conservation actions must include basic research. The saiga is much better known, but even for this species, distributions, movement ecology, and population sizes are still unclear, because the species is nomadic, ranging widely, and crossing national boundaries in their seasonal migrations. The hutias and solenodons also highlight the importance of captive breeding as a safety net for species that are restricted to single sites, exist in very low numbers and for which the *in situ* threats are extreme. Saigas and cheetah breed poorly in captivity, and remain at relatively high numbers in the wild, demonstrating that captive breeding is not an appropriate solution for all species. In all cases, understanding species' biology is not enough; conservation scientists also need to carry out research to understand the social and economic dynamics of the system in order to be able to target conservation action at the right people. For example, cheetah and wild dog conservation involves gaining the support of livestock farmers and game ranchers, because protected areas alone cannot support adequate populations of the species.

As information accumulates about the needs and distribution of species and the attitudes and circumstances of the people living with them, it is important to reassess priorities and approaches. The cheetah and wild dog planning workshops are an excellent example of taking a strategic approach to conservation, at an appropriate scale. Similarly, the periodic reassessment of conservation progress and priorities that is facilitated by the CMS's saiga conservation MOU is very useful in ensuring that conservation is tailored to the circumstances that the species faces. In the case of the saiga, conservation is moving from emergency mode into a stage when conservationists are able to think ahead and plan strategically for the future, including ensuring that protected area networks are robust to climate change and that the cost-effectiveness and impact of their interventions are properly evaluated.

All three case studies involve species that are found in relatively low-income areas and in places where the capacity to carry out research and conservation is limited. This includes the technical capacity to carry out tasks such as GIS mapping, the financial capacity to fund the research and conservation that is required, and the administrative capacity to implement conservation action. In each case, the conservation process has included supporting these countries to build their capacity, with the goal of enabling countries to manage their own biodiversity in collaboration with others, rather than relying on outsiders. Ensuring that countries feel ownership over their science and conservation is the only way to ensure sustainable conservation in the long run. International processes in which stakeholders meet to agree on priorities and activities and can coordinate their efforts across boundaries are crucial. This may apply less to the Caribbean mammals, although Hispaniola is an example in which the Haitian and Dominican Republic governments share conservation problems which require a joint solution.

Conclusions

Terrestrial mammals are often used as flagship species for conservation. The classic example of this is the panda used by WWF, but tigers, elephants, whales, great apes, and other charismatic species are also featured in publicity material about conservation because they are appealing to the public. Many people, particularly in wealthier countries, place very high intrinsic values on mammals and are extremely keen to preserve them and to ensure the welfare of individual representatives of the species. Other people, particularly those who live alongside wild mammals, may be very keen to use them, or to get rid of them. Many mammals are economically valuable because of their meat, furs, or trophies, or are destructive neighbors that invite persecution. At the same time, because of their biology, these species can be very susceptible to persecution or overexploitation.

These factors combine to make many mammals high-profile species among both the general public and conservationists. They often provide a battleground on which conservation battles of more general significance are fought. An example of this is the highly charged and public debate about the listing of the African elephant on the CITES Appendices. This began in 1989 with a successful proposal to transfer the species from Appendix II (controlled international trade) to Appendix I (no international commercial trade), on the basis that the elephant was threatened with extinction

from an uncontrolled, and largely illegal, international ivory trade. The ensuing public debate aired issues ranging from the principles of sustainable use, national sovereignty over resources, animal welfare and animal rights, bringing relatively sophisticated arguments to a wider audience. In March 1999, CITES agreed to a limited resumption of ivory sales, several of which have taken place since then. Much investment has gone into large scale monitoring of the ivory trade and of illegal killing of elephants, in an attempt to understand whether CITES actions have had an effect on elephant poaching. Although this is a hugely complex task, due to the many confounding factors, the recognition that conservationists need scientifically to assess the effects of conservation interventions is an important one (Burn, 2007).

Mammal conservation is high profile and an area in which much general conservation policy is debated and eventually made. The danger is that this policy may not be at all relevant to other components of biodiversity, with their own specific conservation requirements and biological peculiarities. Policies made with respect to one high-profile mammal species may not even be positive for other mammals. Flagship species have to be chosen wisely, since experience shows that species' conservation needs cannot be predicted from their body size or trophic position. For example, in the 1980s, India designated a network of reserves designed to protect tigers, under the banner of "Project Tiger." The project was expected to have the added benefit of protecting other species, including the dhole, a species of wild dog. This expectation was not met; dhole require much larger areas to persist than do tigers and have disappeared from several Project Tiger reserves (Woodroffe and Ginsberg, 1998).

Conservation efforts for mammals are particularly noteworthy because of their impacts on biodiversity conservation in general. This is partly because mammals can be powerful forces structuring ecosystems. But it is also because they are often flagship species for conservation policy; this can be a mixed blessing for the conservation of biodiversity as a whole, as policy made with one particular group of species in mind will not necessarily be appropriate for others. Though this calls for caution in extrapolating the needs of mammals to conservation policy on the broader scale, it is also important to remember that mammals are particularly in need of conservation efforts in their own right, and that some mammal species, particularly small and cryptic ones, are still largely ignored both by conservationists and the general public.

See also: Captive Breeding and Reintroduction. Conservation Efforts, Contemporary. *In Situ*, *Ex Situ* Conservation. Mammals, Biodiversity of. Mammals (Late Quaternary), Extinctions of. Marine Mammals, Extinctions of. Predators, Ecological Role of

References

Bekenov AB, Grachev Iu A, and Milner-Gulland EJ (1998) The ecology and management of the saiga antelope in Kazakhstan. *Mammal Review* 28: 1–52.

Boyd L and Bandi N (2002) Reintroduction of takhi, *Equus ferus przewalskii*, to Hustai National Park, Mongolia: Time budget and synchrony of activity pre- and post-release. *Applied Animal Behaviour Science* 78: 87–102.

Burn RW (2007) Elephants and ivory. *Significance* 4: 118–122.

Convention on Migratory Species (2010) Progress Towards the Fulfilment of the CMS Medium Term Work Programme for the Saiga Antelope for the Period October 2006–September 2010. http://www.cms.int/species/saiga/2ndMtg_Mongolia/Mtg_docs/Doc_05_Rev3_MTW_Progress_Report_Oct2006_Sept2010_E.pdf

Craigie ID, Baillie JEM, Balmford A, *et al.* (2010) Large mammal population declines in Africa's protected areas. *Biological Conservation* 143: 2221–2228.

Creel SR and Creel NM (1996) Limitation of African wild dogs by competition with larger carnivores. *Conservation Biology* 10: 1–15.

Durant SM, Dickman AJ, Maddox T, Waweru M, and Pettorelli N (2010) Past, present and future of cheetah in Tanzania: From long term study to conservation strategy. In: Macdonald DW and Loveridge AJ (eds.) *Biology and Conservation of Wild Felids*, pp. 373–382. Oxford: Oxford University Press.

Fa JE, Currie D, and Meeuwig J (2003) Bushmeat and food security in the Congo Basin: Linkages between wildlife and people's future. *Environmental Conservation* 30: 71–78.

Fa JE, Soy JP, and Capote R (2002) Biodiversity of Sierra del Cristal, Cuba: First insights. *Oryx* 36: 389–395.

Fuller TK, Kat PW, Bulger JB, *et al.* (1992) Population dynamics of African wild dogs. In: McCullough DR and Barrett H (eds.) *Wildlife 2001: Populations*. London: Elsevier.

George JC, Zeh J, Suydam R, and Clark C (2004) Abundance and population trend (1978—2001) of Western Arctic Bowhead whales surveyed near Barrow, Alaska. *Marine Mammal Science* 20: 755–773.

Howe C, Medzhidov R, and Milner-Gulland EJ (2011) Evaluating the relative effectiveness of alternative conservation interventions in influencing stated behavioural intentions: A case study of the saiga antelope in Kalmykia. *Environmental Conservation* 38: 37–44.

IUCN (2010) IUCN Red List of Threatened Species. Version 2010.4. http://www.iucnredlist.org. Downloaded on 5 December 2010.

IUCN/SSC (2007) *Regional Conservation Strategy for the Cheetah and African Wild Dog in Eastern Africa*. Gland, Switzerland: IUCN Species Survival Commission.

Mace GM and Balmford A (2000) Patterns and processes in contemporary mammalian extinction. In: Entwhistle A and Dunstone N (eds.) *Future Priorities for the Conservation of Mammalian Diversity*. Cambridge: Cambridge University Press.

Marker LL, Dickman AJ, Mills MGL, Jeo RM, and Macdonald DW (2008) Spatial ecology of cheetahs on north-central Namibian farmlands. *Journal of Zoology* 274: 226–238.

Milner-Gulland EJ, Kholodova MV, Bekenov AB, *et al.* (2001) Dramatic declines in saiga antelope populations. *Oryx* 35: 340–345.

Price SA and Gittleman JL (2007) Hunting to extinction: Biology and regional economy influence extinction risk and the impact of hunting in artiodactyls. *Proceedings of the Royal Society B* 274: 1845–1851. doi:10.1098/rspb.2007.0505

Roca AL, Bar-Gal GK, Eizirik E, *et al.* (2004) Mezozoic origin for West Indian insectivores. *Nature* 429: 649–651.

Rowcliffe JM, de Merode E, and Cowlishaw G (2004) Do wildlife laws work? Species protection and the application of a prey choice model to poaching decisions. *Proceedings of the Royal Society of London B* 271: 2631–2636.

Sanderson EW, Redford KH, Chetkiewicz CLB, *et al.* (2002) Planning to save a species: The jaguar as a model. *Conservation Biology* 16(1): 58–72. doi:10.1046/j.1523-1739.2002.00352.x

Schipper J, Chanson JS, Chiozza F, *et al.* (2008) The status of the world's land and marine mammals: Diversity, threat, and knowledge. *Science* 322: 225–230.

Smith DW, Peterson RO, and Houston DB (2003) Yellowstone after wolves. *BioScience* 53: 330–340.

Turvey ST, Meredith HMR, and Scofield RP (2008) Continued survival of Hispaniolan solenodon (*Solenodon paradoxus*) in Haiti. *Oryx* 42: 611–614.

Woodroffe R and Ginsberg JR (1998) Edge effects and the extinction of populations inside protected areas. *Science* 280: 2126–2128.

Mammals (Late Quaternary), Extinctions of

Paul S Martin, University of Arizona, Tucson, AZ, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 3, pp 825–839, © 2001, Elsevier Inc.

Glossary

Extinction spasm or pulse A catastrophic burst of extinctions, peaking in less than a millennium.

First contact Initial human arrival on a landmass followed by human colonization.

Late Quaternary extinction event Selective prehistoric extinctions, typically catastrophic, eliminating within the past 40,000 years two-thirds or more of large land mammals of America, Australia, and Madagascar and at least half the species of land birds on remote islands of the Pacific. Humans are present or suspected to be present in virtually all cases.

Megafauna Large terrestrial vertebrates variously defined as >1, >10, and >44 kg adult body weight. Equivalent to 100 pounds and similar to the average weight of adult humans, a body size exceeding 44 kg is easily visualized and is adopted here.

Quaternary The ice age of at least the past 1.81 million years, including the Pleistocene and the Holocene, the latter representing the past 10,000 years.

Radiocarbon dating An isotopic or nuclear decay method for inferring age of organic materials. Carbon 14 is produced in the upper atmosphere by cosmic ray bombardment and oxidized to form $C^{14}O_2$. Distributed through the earth's atmosphere and oceans, a small percentage is incorporated into a variety of organic materials to decay with a half-life of 5700 years. By dating tree rings of known age, and other methods, radiocarbon determinations can be calibrated. Although routine application of radiocarbon dating is usually limited to dates of less than 40,000 years, ages up to 75,000 years have been measured.

Introduction: Prehistoric Extinctions

The late Quaternary extinction event (LQEE) is best known by the loss of large mammals (megafauna) from certain continents. Familiar faunal examples of the LQEE include woolly mammoths of the Northern Hemisphere, woolly rhinos in northern Eurasia, ground sloths and glyptodonts in the Americas, and diprotodonts and giant kangaroos in Australia. Comparisons between landmasses are revealing. The continents differ significantly in both the magnitude and in the timing of their extinctions. First Australia, then North and South America, and lastly Madagascar (the Island Continent) rapidly lost two-thirds or more of their large mammals (**Figure 1**). In contrast, losses were comparatively gradual and much less severe in Eurasia and Africa. Although small animals of the continents seldom suffered extinctions, thousands of endemic mammals, birds, and reptiles vanished from oceanic islands.

At least 85 genera of large (>44 kg) mammals, half of the terrestrial megafauna of the planet, disappeared from the continents in the past 100,000 years. Now "out of sight, out of mind," the extinct mammoths and other large animals of the Quaternary coevolved with modern plant species, thereby shaping certain features of vascular plant anatomy and biochemistry. Undoubtedly, the extinct species were just as influential as living megaherbivores in determining the structure of natural communities (Owen-Smith, 1988). Before 13,000 years ago the diversity of large mammals in the New World would have resembled that currently found in African or Asian game parks. Megafaunal extinctions have made it difficult for ecologists to envision the multiple stable states prevailing in evolutionary time. In the absence of their megafauna, the savannas, grasslands, riparian habitats, and other biotic communities of the New World, Australia, and Madagascar are not functionally "natural."

Prehistoric settlement on deep-water islands in the remote Pacific began 4000 years ago. Extinctions followed, eliminating many endemic birds, reptiles, and land snails along with depletion or size reduction brought by overharvesting of nearshore fish, shellfish, and marine turtles. The relatively small number of oceanic islands that escaped prehistoric discovery suffered losses during historic time of such animals as the dodo, Steller's sea cow, giant rats, and parrots (**Figure 2**). No late Quaternary extinction pulse is seen in the fossil record of the whales, other marine fauna, or in Antarctica.

An intense and often controversial search during the past 40 years for an explanation for the cause or causes of the LQEE has failed to yield any widespread consensus among paleontologists, archeologists, paleoecologists, geographers, and other interested parties, many of them convinced of the efficacy of climatic change and of climatic extinction models. In recent years, as the global pattern of the LQEE has become better known, much more attention has been focused on anthropogenic models (compare MacPhee (1999) with Martin and Klein (1984)). Always on the scene, humans can no longer be ignored in the search for the cause or causes of "extinctions in near time."

The Comparative Approach

In the view of most geologists the Quaternary or Pleistocene ice age of multiple glaciations embraces at least the past 1.81

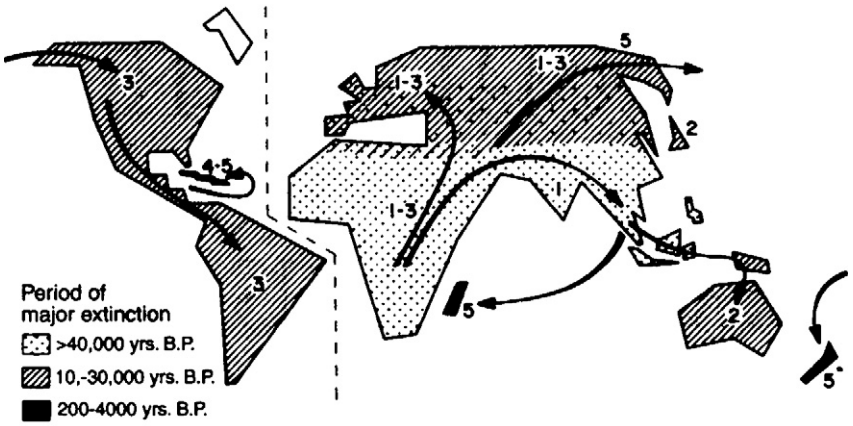


Figure 1 Late Quaternary extinction and prehistoric human dispersal; for numerical sequence and estimates of extinction timing, see the legend to **Figure 2** (Reproduced from Martin PS and Steadman DW (1999) Prehistoric extinctions on islands and continents. In: MacPhee R (ed.) *Extinctions in Near Time*, pp. 17–55. New York: Kluwer/Plenum, with permission from Springer.).

million years (Ma), with the late Quaternary representing the past 200,000 years. In North America the late Quaternary includes the Rancholabrean mammalian age stage, marked by the appearance of the genus *Bison*.

With the possible exception of Australasia, most late Quaternary extinctions fall within the reach of radiocarbon assay, essentially the past 40,000 years. Fossils uncontaminated by groundwater and yielding bone collagen or gelatin are suitable for dating. Environments favorable for the preservation of bone collagen include dry caves of arid regions, the frozen ground of high latitudes, lacustrine deposits, and saline or petrolic sediments such as those at Rancho la Brea, California. Burned bone and associated charcoal are ideal for dating, although charcoal may prove intrusive. A variety of geochemical dating methods, especially radiocarbon dating, allow biogeographers to compare both the rate of extinction within a landmass and the timing of extinctions between landmasses. A robust chronology allows intercontinental and insular comparisons of considerable value in inferring cause of the extinctions.

Although details regarding just how the extinctions were triggered remain speculative, it is increasingly clear that the two major contending explanations of what forced LQEE—that is, climatic changes and cultural impacts—can now be evaluated globally. When approached on a comparative basis, human expansion and climatic change track a punctuated extinction episode in North America and perhaps in South America, whereas climatic change is less clearly involved in Australia or Madagascar and appears to be of no significance in New Zealand and other remote Pacific islands.

North America

Focusing on the past 4 million years only (the Pliocene and the Quaternary), it is possible to evaluate the LQEE using age-stage divisions of biostratigraphers (**Table 1**). The overwhelming importance of extinctions of North American large mammals at the end of the Quaternary, shown in **Figures 3** and **4** and in **Table 1**, has been evident for at least the past 50 years.

The widespread adoption of screen washing of fossil deposits has vastly enriched the fossil record of small vertebrates. The harvest of identified small mammals is vital in supporting the conclusion that no appreciable extinctions of small mammals occurred (**Figure 3**)—certainly not matching the extinction spasm of large mammals approximately 11,000 radiocarbon years ago (**Figure 4**).

The fossil record is based on a much larger sample of late Pleistocene than early Pleistocene and Pliocene faunas (**Table 1** and **Figures 3** and **4**). Presumably this accounts at least in part for the increase in number of genera of both large and small mammals in the later part of the record. Although *Mammuthus*, *Equus*, and extinct *Bison* included multiple species, most of the large genera listed in **Table 1** are monotypic (represented by one species only). Although a few survived on other continents, most large genera totally vanished (**Table 1**).

In North America north of Mexico, 33 genera of large mammals disappeared in the late Quaternary, which constitutes two-thirds of the total late Rancholabrean fauna of 48 genera and more than all the generic extinctions recorded during the previous 4 million years prior to LQEE (**Table 1**). Events that might trigger regional or global extinctions in the early part of the past 4 million years include intercontinental suturing of the Americas with faunal intermingling, extraterrestrial accidents such as the impact of the Eltanan bolide, and the initiation of continental glaciation. However, only at the end of the Quaternary is there a major extinction spasm—one that impoverished America's large mammalian faunas (**Figure 4**).

Unlike the case of the large mammals, a much larger number of small genera (48) vanished in the past 4 million years. Within the minifauna (rat, mouse, or shrew-size mammals), which make up the majority of the 75 living genera of small mammals, no generic extinctions are known in the late Quaternary. Among the medium-sized mammals, three genera were lost: the antilocaprid *Stockoceros*, the skunk *Brachypotoma*, and the rabbit *Aztlanolagus*. The difference deserves close inspection. Although small mammals escaped virtually unscathed, the LQEE (late Rancholabrean losses) blighted all terrestrial orders of large land mammals and eliminated one, the Proboscidea (**Table 1**).

Since the vast majority of Cenozoic (last 65 Ma) extinctions occurred long before radiocarbon time or FC, it is essential to know something of the Cenozoic pattern. Do earlier bursts of extinction match the LQEE pattern at the very end of the Cenozoic? As a result of a relatively rich fossil record of the Cenozoic, the question is readily approached in North America.

Paleontologist John Alroy (cited in MacPhee 1999) recently analyzed 65 Ma of change in mammalian faunas through 1.0 Ma sampling bins. His data consist of extinction time rate series that are computed from a multivariate ordination of 4015 faunal lists from 2415 publications that span the Late Cretaceous through Sangamonian (last interglacial). The lists can be viewed on the World Wide Web at the North American Mammalian Paleofaunal Database: http://homebrew.si.edu/nampfd.html.

According to Alroy (as quoted in MacPhee, 1999, p. 120), standing diversity is defined by

> the number of species that cross each boundary between sampling bins, e.g., the time planes at 65.0, 64.0 … 1.0 Ma. Extinctions are computed by counting the number of species in each cohort that fail to survive until the next time plane. Note that this excludes single-interval species (those that appear and disappear in the same bin), which makes the analysis less vulnerable to sampling artifacts.

Alroy's results indicate that within the past 55 Ma there were more extinctions of large species in the LQEE than there were during any earlier time. The difference in the mean body masses of the victims and the survivors peaks at the end of the Pleistocene (**Figure 5**).

Although there is a general increase in mean mass of mammals through the Cenozoic, the mean mass of victims (extinct species) is highest in the past 1 million year bin within the Cenozoic (**Figure 6**), which includes the LQEE. In other words, when compared with 64 turnovers earlier in the Cenozoic, the 65th has a unique property—excessive extinction of large mammals.

With the exception of bison (*Bison*) and deer (*Odocoileus*), the common fossils of large herbivores found in Quaternary deposits before the LQEE are not the bones or teeth of species that survived the extinctions, such as *Alces*, *Antilocapra*, *Cervus*, and *Rangifer*, but rather those that did not, especially *Camelops*, *Cervalces*, *Equus*, *Mammut*, *Mammuthus*, and *Platygonus*

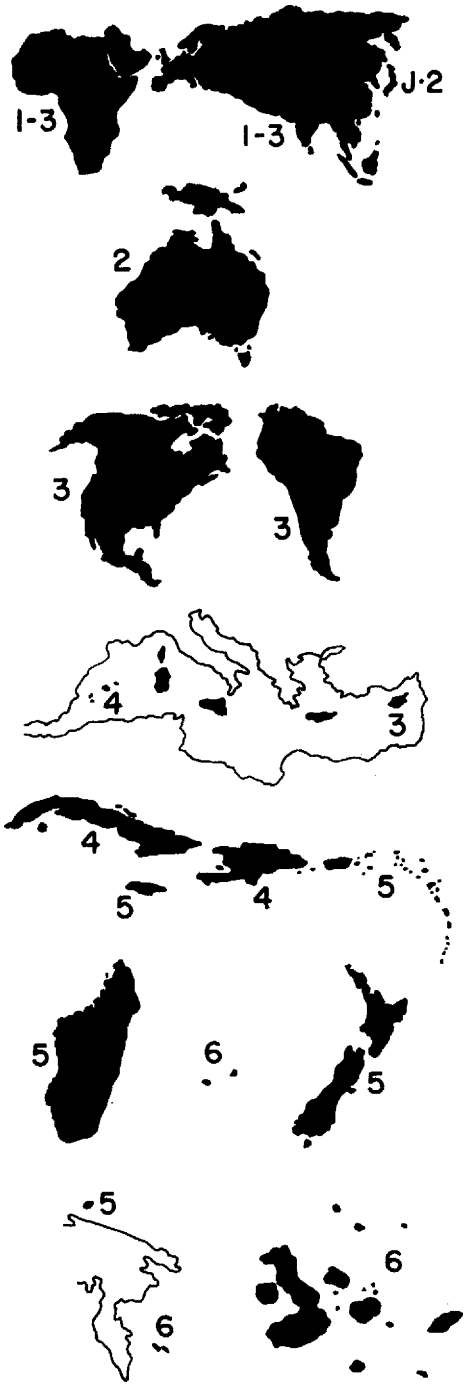


Figure 2 Map silhouettes of continents and islands showing time-transgressive sequence of late Quaternary extinctions and FC. *Africa and Eurasia* (1–3): Sequential extinctions of large mammals during the past 100,000 years including Nauman's elephant and giant deer in Japan [J-2] ca. 30,000 years ago. *Meganesia* (2): Humans arrive and major extinctions occur 30,000–40,000 yr bp; neither event is well-dated or constrained. *Americas* (3): Well-dated evidence for human arrival and for megafaunal extinctions center on 11,000 yr bp (13,000 calendar years). *Mediterranean islands*: Epipaleothic arrival and dwarf hippo and dwarf elephant extinction on Cyprus (3) 10,500 yr bp; mid-Holocene colonization and extinction of goat antelope (*Myotragus*) on the Balearic Islands (4); FC and LQEE chronology elsewhere are uncertain. *Antilles*: Humans arrive in Cuba and Hispainola in early or mid-Holocene (4) and in Jamaica and the Lesser Antilles in the late Holocene (5); few radiocarbon dates on extinctions. *Madagascar* (5): Humans arrive 2000 yr bp; major episode of extinction terminates 1500 AD. *Mascarenes*, east of Madagascar (6): Humans arrive 1600 AD, followed by extinctions of dodo, solitaire, other birds, and giant tortoises. *New Zealand* (5): Humans colonize by 1200 AD; the Polynesian rat perhaps colonized much earlier. Giant flightless birds (moas) are extinct by 1500 AD. *Wrangel Island* (5): Last woolly mammoths dated 4000 yr bp, 1000 years older than oldest cultural material. *Commander Islands* (6): Humans arrive 1741 AD; Steller's sea cow extinct within 30 years. *Galapagos Islands* (6): Bishop of Panama arrives 1535 AD; extinction rates increase over background by two orders of magnitude (Reproduced from Martin PS and Steadman DW (1999) Prehistoric extinctions on islands and continents. In: MacPhee R (ed.) *Extinctions in Near Time*, pp. 17–55. New York: Kluwer/Plenum, with permission from Springer.).

Table 1 Large (>44 kg) Plio-Pleistocene terrestrial mammals of North America north of Mexico[a]

Stage duration (Ma)	*BLANCAN*				*IRV*			*RLB*		*H*
	1,2	*3*	*4*	*5*	*E*	*M*	*L*	*E*	*L*	
	1.0	*0.5*	*0.5*	*0.2*	*0.9*	*0.4*	*0.2*	*0.2*	*0.1*	*0.01*
Xenarthra										
+ *Glyptotherium*, glyptodont			X	X	X	X	X	X	X	
+ *Holmesina*, northern pampathere			X	X	X	X	X	X	X	
+ *Pachyarmatherium*, ground sloth							X			
+ *Eremotherium*, giant ground sloth					X	X	X	X	X	
+ *Nothrotheriops*, Shasta ground sloth					X	X	X	X	X	
+ *Megalonyx*, Jefferson's ground sloth	X	X	X	X	X	X	X	X	X	
+ *Paramylodon*, big-tongued ground sloth		X	X	X	X	X	X	X	X	
Carnivora										
+ *Borophagus* plundering dog	X	X	X	X	X					
Canis, dire wolf, gray wolf				X	X	X	X	X	X	X
+ *Protocyon* Troxell's dog					X					
Ursus, bears		X	X	X	X	X	X	X	X	X
[+] *Tremarctos*, Florida cave bear		X	X	X	X	X	X	X	X	
+ *Arctodus*, short-faced bear						X	X	X	X	
+ *Chasmaporthetes*, hunting hyena	X	X	X	X	X					
+ *Meganteron*, western dirktooth	X	X	X	X						
+ *Smilodon*, sabertooth					X	X	X	X	X	
+ *Ischyrosmilus*, Idaho sabertooth	X									
+ *Homotherium*, scimitar cat					X	X	X	X	X	
+ *Dinofelis*, false sabertooth		X	X	X						
[*] *Panthera*, American lion, jaguar						X	X	X	X	X
+ *Miracinonyx*, American cheetah				X	X	X	X	X	X	
Felis, cougar, puma					X	X	X	X	X	X
Rodentia										
+ *Procastoroides*, large beaver	X	X	X							
+ *Castoroides*, giant beaver						X	X	X	X	
+ *Neochoerus*, giant capybara					X	X	X	X	X	
[+] *Hydrochoerus*, Holmes's capybara							X	X	X	
Proboscidea										
+ *Mammut*, American mastodon	X	X	X	X	X	X	X	X	X	
+ *Stegomastodon*, stegomastodon		X	X	X	X					
+ *Rhyncotherium*, rhynchothere	X	X	X	X						
+ *Cuvieronius*, gomphothere			X	X	X	X	X	X	X	
+ *Mammuthus*, extinct mammoths					X	X	X	X	X	
Sirenia										
+ *Hydrodamalis*, Stellar's sea cow	X	X	X	X	X	X	X	X	X	X
Trichechus, manatee	X	X	X	X	X	X	X	X	X	X
Perissodactyla										
+ *Cormohipparion*, extinct equid	X	X								
+ *Nannipus*, gazelle–horse	X	X	X	X						
+ *Plesippus*, extinct equid	X	X	X	X						
[+] *Equus*, horse species		X	X	X	X	X	X	X	X	
[+] *Tapirus*, tapir species					X	X	X	X	X	
Artiodactyla										
+ *Megalotylopus*, large camelid	X	X	X							
+ *Blancocamelus*, giraffe—camels				X						
+ *Titanotylopus*, gaint camelid	X	X	X	X	X					
+ *Camelops*, camel species				X	X	X	X	X	X	
+ *Hemiauchenia*, llama species	X	X	X	X	X	X	X	X	X	
+ *Palaeolama*, stout-legged llama							X	X	X	
+ *Mylohyus*, long-nosed peccari			X	X	X	X	X	X	X	
+ *Platygonus*, flat-headed peccari	X	X	X	X	X	X	X	X	X	
+ *Bretzia*, false elk		X								

(*Continued*)

Table 1 Continued

Stage duration (Ma)	*BLANCAN*				*IRV*			*RLB*		*H*
	1,2	*3*	*4*	*5*	*E*	*M*	*L*	*E*	*L*	
	1.0	*0.5*	*0.5*	*0.2*	*0.9*	*0.4*	*0.2*	*0.2*	*0.1*	*0.01*
Odocoileus, deer species	X	X	X	X	X	X	X	X	X	X
+ *Torontoceros*, extinct large cervid									X	
+ *Navahoceros*, mountain deer									X	
Rangifer, caribou				X	X	X	X	X	X	X
**Alces*, moose, broad-fronted moose									X	X
+ *Cervalces*, stag-moose									X	
Cervus, wapiti (elk)				X	X	X	X	X	X	X
+ *Tetrameryx*, four-horned pronghorns					X	X	X	X	X	
+ *Hayoceros*, Hay's pronghorn							X			
+ *Slockoceros*, four-horned pronghorns							X	X	X	
Antilocapra, pronghorn									X	X
[+]*Saiga*, saiga									X	
**Oreamnos*, mountain goats					X	X	X	X	X	X
Ovis, bighorn or mountain sheep						X	X	X	X	X
+ *Euceratherium*, shrub ox					X	X	X	X	X	
+ *Soergelia*, Soergel's ox					X	X	X			
+ *Bootherium*, bonnet-headed musk ox							X	X	X	
+ *Praeovibos*, extinct musk ox							X			
Ovibos, musk ox									X	X
[*]*Bison*, bison species								X	X	X
+ *Platycerabos*, flat-horned ox								X		
Primates										
Homo, *modern H. sapiens*									X	X
Originations	—	7	4	6	13	4	7	2	8	0
Extinctions	1	2	2	6	5	0	4	1	33	0
Total genera	18	24	26	30	37	36	43	41	48	15

[a]After Kurtén and Anderson (1980), Martin and Steadman (1999), and Anderson (personnel communication). Abbreviations used: IRV, Irvingtonian; RLB, Rancholabrean; H, Holocene; E, early; M, middle; L, late; +, extinct genus.

*extinct species; brackets indicate generic survival on other continents. Living genera including extinct taxa of large size that vanished with late RLB extinction include *Dasypus belli* (giant armadillo), *Canis dirus* (dire wolf), *Panthera leo atrox* (American lion), *Alces latifrons* (broad-fronted moose), and *Oreanos harringtoni* (Harrington's mountain goat). Age estimates in millions of years (Ma): Blancan 1,4.0–3.5 Ma; Blancan 2,3.5–3.0 Ma; Blancan 3, 3.0–2.5 Ma; Blancan 4, 2.5–2.0 Ma; Blancan 5, 2.0–1.8 Ma (Blancan–Irvingtonian Boundary); early Irvingtonian, 1.8–0.9 Ma; middle Irvingtonian, 0.9–0.5 Ma; late Irvingtonian, 0.5–0.3 Ma; early Rancholabrean 0.3–0.1 Ma; late Rancholabrean, 0.1–0.01 Ma; Holocene, last 10,000 years. There may not be specimen records for all cells within the temporal range of a genus as plotted here. Checklist and age estimates are courtesy of Elaine Anderson (1996), Denver Museum of Natural History, Denver, Colorado, Adapted from Kurtén B and Anderson E (1980) *Pleistocene Mammals in North America*. New York: Columbia Univ. Press.

(Graham and Lundelius, 1994), All have terminal radiocarbon records at approximately 11,000 years ago (Stuart, 1991). One might imagine that if the LQEE were a natural catastrophe the ranges of the Holocene survivors (15 species listed in **Table 1**) would have expanded rapidly as niche space opened following extinctions. For example, if some lethal climate condition at the end of the last glaciation eliminated their North American relatives, one might expect that afterward early Holocene warming would see the surviving populations of bison, elk, and moose expanding or at least maintaining their range. In fact, bison range shrank away from both coasts (Graham and Lundelius, 1994). Only in recent years, with the help of local introductions by game departments, has the local range and numbers of moose, elk, and black bear expanded.

The historic fauna of bison, deer, moose, elk, pronghorn, bears, wolves, etc. is traditionally regarded as the "natural" fauna of North America. However, the view from the past shows that they are an aberrant, extinction-pruned remnant, unrepresentative of the diversity or the ecological amplitude of native large mammals that the continent supported during the latter half of the Cenozoic. To view free-ranging bison, deer, moose, elk, pronghorn, bears, and wolves, the protected fauna of Yellowstone National Park in Wyoming or Denali National Park in Alaska, as representative of an American Serengeti is to seriously underestimate the diversity, productivity, and evolutionary potential of the continent.

South America, Australia, and Madagascar

The continents of South America, Australia, and Madagascar share a common property with North America—heavy extinction of large mammals late in the Quaternary. South America lost all large mammals more massive than a tapir (300 kg); Australia lost all mammals larger than a gray kangaroo (60 kg); and, excepting the bush pig which may have been introduced, Madagascar lost all terrestrial vertebrates

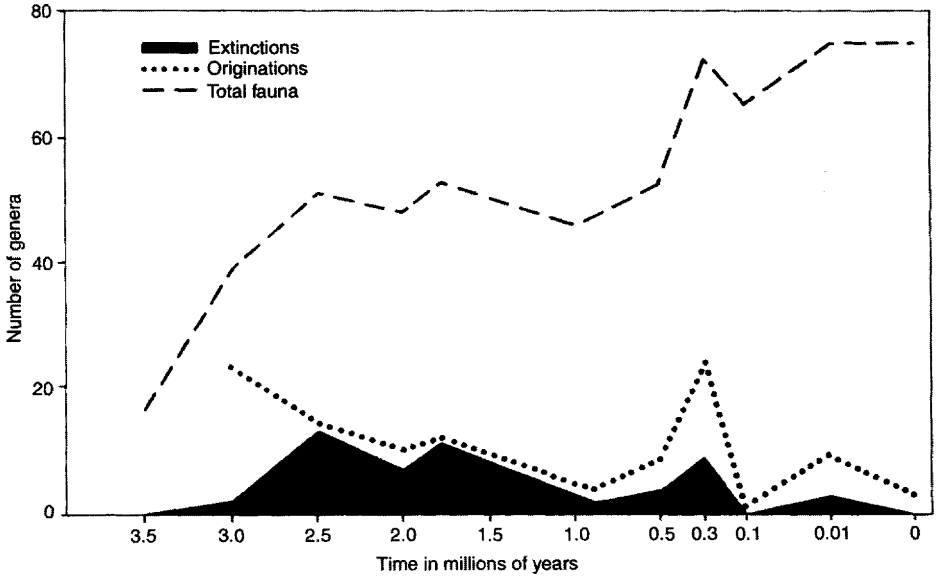


Figure 3 North American small mammal originations (dotted line), extinctions (solid line), and total fauna (dashed line) since the Miocene from data shown in **Table 1**. Note the change of scale in the past 0.5 Ma (Reproduced from Martin PS and Steadman DW (1999) Prehistoric extinctions on islands and continents. In: MacPhee R (ed.) *Extinctions in Near Time*, pp. 17–55. New York: Kluwer/Plenum, with permission from Springer.).

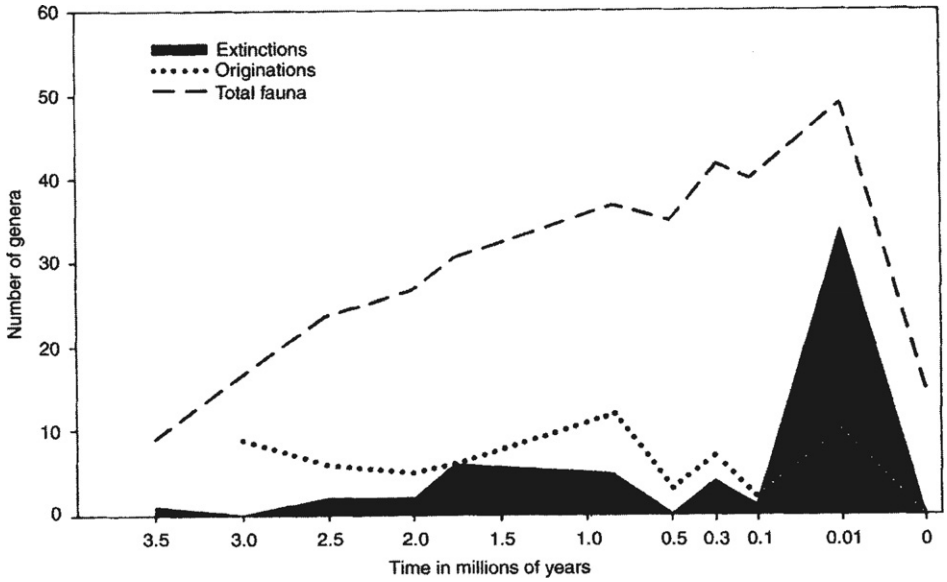


Figure 4 North American large mammal originations (dotted line), extinctions (solid line), and total fauna (dashed line) since the Miocene from data shown in **Table 1**. Note the change of scale in the past 0.5 Ma (Reproduced from Martin PS and Steadman DW (1999) Prehistoric extinctions on islands and continents. In: MacPhee R (ed.) *Extinctions in Near Time*, pp. 17–55. New York: Kluwer/Plenum, with permission from Springer.).

larger than the living *Indri*, a lemur which weighs 7–10 kg. Timing of the extinctions varies; it was earliest in Australia, later in South (and North) America, and latest in Madagascar. Although Australian small mammals (0.5–2.0 kg) suffered historic losses, there is no indication that severe late Quaternary extinction of small mammals accompanied the loss of large ones on the continents.

South America

The Quaternary of South America includes the Ensenadan and Lujanian land mammal ages dated at 1.5 to 0.5 and 0.5 to 0.01 Ma, respectively. The preceding Uquian land mammal age from 2.5 to 1.5 Ma is of great interest because it immediately follows the intercontinental suturing of North and South America. The Uquian shows the effects of the famous Great American Faunal Interchange, a natural experiment in which two continents exchanged long-isolated terrestrial faunas across a new land bridge. The land bridge connecting the continents at the end of the Pliocene brought Northern Hemisphere carnivores, gomphotheres, artiodactyls, and perissodactyls into contact with South American endemic orders, the xenarthrans, notoungulates,

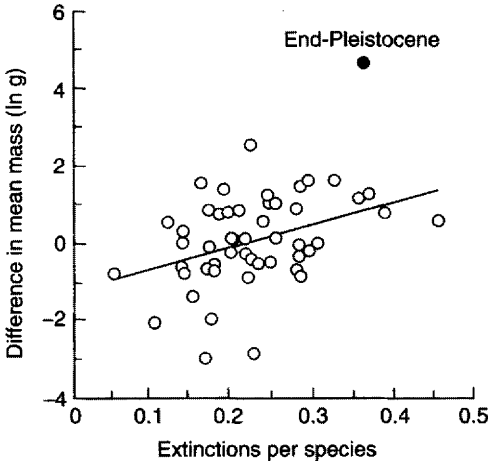


Figure 5 Diversity and extinction in the Cenozoic Data are prepared by multivariate ordination and randomized subsampling. The figure plots the correlation among 1.0-Ma intervals between extinction intensity and the difference in the mean body masses of the victims and the survivors. The end Pleistocene value (●) is among the highest during the past 55 Ma (Reproduced from MacPhee R (ed.) (1999) *Extinctions in Near Time*. New York: Kluwer/Plenum, with permission from Springer.).

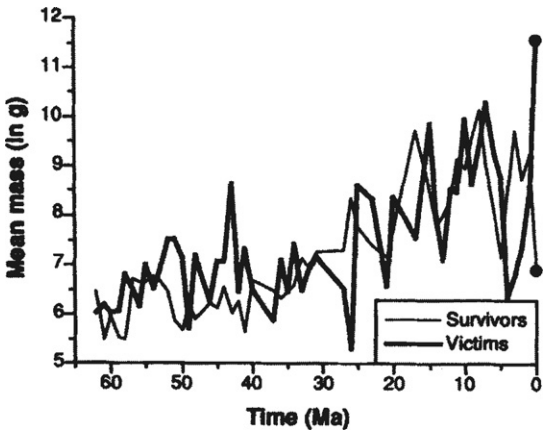


Figure 6 Trend through the Cenozoic in the mean body mass of victims and survivors in each 1.0-Ma interval (see legend to **Figure 5**) (Reproduced from MacPhee R (ed.) (1999) *Extinctions in Near Time*. New York: Kluwer/Plenum, with permission from Springer.).

and litopterns. Interestingly, millions of years before the intercontinental conjunction, megalonychid ground sloths managed to colonize both the Greater Antilles and North America.

Occasionally, biogeographers have sought to invoke intercontinental exchange as a cause of the LQEE. A detailed summary of the South American fossil faunas by Marshall and Cifelli (1990) shows that although small-mammal extinctions occurred in the Uquian, heavy megafaunal extinction did not occur until the Lujanian, approximately 2 million years after the exchange had begun and apparently coeval with the LQEE in North America.

South American extinctions of the late Quaternary were even heavier than those in North America north of Mexico,

Table 2 Large (>44 kg) extinct mammals of the Lujanian (Rancholabrean equivalent) in South America

Endemics
Order Xenarthra
Family Dasypodidae
Eutatus
Holmesina
Pampatherium
Propraopus
Family Glyptodontidae
Chlamydotherium
Doedicurus
Glyptodon
Hoplophorus
Lomaphorus
Neothoracophorus
Neosclerocalyptus
Panochthus
Parapanochthus
Plaxhaplous
Sclerocalyptus
Family Megalonychidae
Nothropus
Nothrotherium
Ocnopus
Valgipes
Family Megatheriidae
Eremotherium
Megatherium
Paramegatherium
Family Mylodontidae
Glossotherium
Lestodon
Mylodon
Scelidodon
Scelidotherium
Order Perissodactyla
Family Equidae
Equus
Hippidion
Onohippidium
Order Artiodactyla
Family Camelidae
Eulamaops
Hemiauchenia
Palaeolama
Family Cervidae
Agalmaceros
Charitoceros
Morenelaphus
Paraceros
Family Tayassuidae
Platygonus
Order Litopterna
Family Macraucheniidae
Macrauchenia
Windhausenia
Order Notoungulata
Family Toxodontidae
Mixotoxodon
Toxodon
Pre-land bridge invaders
Order Rodentia
Family Hydrochoeridae
Neochoerus
Family Octodontidae
Dicolpomys
Post-land bridge invaders
Order Carnivora
Family Canidae
Theriodictis
Family Felidae
Smilodon
Family Ursidae
Arctodus
Order Proboscidae
Family Gomphotheriidae
Cuvieronius
Haplomastodon
Notiomastodon
Stegomastodon

Source: Adapted from Marshall LG and Cifelli RL (1990) Analysis of changing diversity patterns in Cenozoic land mammal age faunas, South America. *Paleovertebrata* 19: 169–210.

involving 50 genera of which 35 belong to extinct families. The losses were differential and involved all megaherbivores such as giant ground sloths, glyptodonts, gomphtheres, and two endemic orders, Litopterna and Notoungulata (**Table 2**). It is not obvious that the interchange had anything to do with the LQEE.

Defensible radiocarbon dates on most of the South American fauna remain to be assembled and may be difficult to obtain given the poor preservation of bone collagen in the open sites where Lujanian faunas are often found. One source of ideal material for radiocarbon dating—dung, hair, and perishable tissue—has been obtained on mylodontid ground sloths from Cueva del Mylodon in southern Chile. Fifteen dates from 13,000 radiocarbon years before present (yr bp) to 10,600 yr bp support the extinction chronology for ground sloths in North America.

Australia

Flannery and Roberts (cited in MacPhee, 1999) report that near the time of first contact (FC) Meganesia (Australasia) lost approximately 28 genera and 55 species of large vertebrates (body weights exceeding 10 kg). The more important extinct genera of mammals in the Australasian (Australia and New Guinea) extinct megafauna are all marsupials. They include the extinct "lion" *Thylacoleo*; the wombats *Phascolomys, Phascolonus*, and *Ramsaya*; a Palorchestid, *Palorchestes*; four diprotodontids—*Diprotodon*, *Euwenia*, *Nototherium*, and *Zygomaturus*; a potoroid, *Propleopus*; and five giant kangaroos—*Fissiuridon*, *Procoptodon*, *Protemnodon*, *Sthenurus* and *Troposodon*. In addition, there was an extinct giant flightless bird, the mihirung (*Genyornis*); a giant lizard, *Megalania*; an extinct horned turtle, *Meiolania*; and a giant python, *Wonambi*. Although older, the pattern of extinction in Australia resembles that in the Americas and Madagascar. In all three landmasses kill sites or processing sites are few, in dispute, or absent. As elsewhere, the lack of kill or butchering sites has deterred many Australian archeologists and paleontologists from invoking an overkill. Although no extinctions of Australian mammals less than 10 kg are known at FC, a sizable number of medium-size mammals are threatened or have vanished in historic time, presumably a side effect of European settlement in Australia (Flannery, 1994).

Despite the problem in direct radiocarbon dating on extinct fauna, the absence of extinct species from well-dated Tasmanian archeological sites up to 25,000 years in age supports the view that the Australian LQEE was over by then. The Franklin River region in Tasmania serves as an example. As in other parts of eastern Australia, Tasmania once harbored a variety of extinct species of giant marsupials. Although 21 dated late Pleistocene sites in caves or rock shelters in southwestern Tasmania ranging in age from 10,000 to at least 25,000 yr bp have yielded hundreds of thousands of animal bones, all are of living species and mainly of Bennett's wallaby, *Macropus rufogriseus* (J. Allen as cited in Kirsch and Hunt, 1997). If the extinct mihirung, giant macropods, diprotodonts, etc. were still alive 25,000 years ago or later, it is difficult to believe that they would not be represented in such rich zooarcheological material. This negative evidence supports the view that extinctions occurred earlier.

Until more fossil faunas are dated, there will be unavoidable uncertainty about when extinction occurred. In this regard, new environmental information from stable isotopes from samples of *Genyornis* eggshell is especially promising. The record to date indicates extinction of *Genyornis* by 45,000 yr bp in several parts of its wide range and in more than one climatic province; these findings argue against extinctions driven by some climatic bottleneck (Miller *et al.*, 1999). In Australia extinctions long predate any hypothetical late glacial climatic forcing.

Finally, in terms of human origins, a comparison between Australia and America is instructive. In the New World archeologists have seen a variety of sites claimed to be 13,000–22,000 years or older fail to be verified or locally replicated, despite regular assertions of their proponents. In contrast, in the past two decades dozens of Australian sites have repeatedly yielded geochemical dates indicating dozens of sites older than Clovis in North America.

Given the smaller area of Australia compared with that of the New World, its much less productive soils and much more variable precipitation (Flannery, 1994), and the much smaller number of archeologists, geologists, paleontologists, and amateurs searching for artifacts and fossils, the abundance of sites 10,000–40,000 years old and older in Australia, compared with the half a dozen claims of pre-Clovis sites in the New World and their debatable status, is a red flag to environmentalists. Unless proposed early sites in the Americas are critically replicated by the geoarcheological community at large, as in Australia, the claims for a pre-Clovis occupation and a pre-Clovis culture, no matter how detailed or how often repeated, remain in limbo (Martin and Steadman, 1999).

Madagascar

Within the past two millennia, at a time when no other continents suffered appreciable extinctions of large mammals (Martin and Klein, 1984), Madagascar lost many large mammals, including 20 species of lemurs up to the size of a gorilla. All are larger than the largest living lemurs. In addition, Madagascar lost two hippos; the highly endemic anteater known as bibymalagasy (*Plesiorycteropus*), recently assigned to its own order; a large (bobcat-size) species of carnivore (*Cryptoprocta spelea*); a rabbit-size rodent (*Hypogeomys australis*); two genera of giant flightless birds; and two giant tortoises, one exceeding 100 kg in mass (MacPhee and Marx as cited in Goodmans and Patterson, 1997). Although most if not all of these extinctions in Madagascar coincide with human colonization, none have been found in well-defined kill sites, leading MacPhee and Marx to model hyperdisease. They generalize their model to include all lands of prehistoric human colonization. Despite its implausibility, the model invites testing since the fossil record has yet to reveal how LQEE occurred. In the Americas the lack of many Clovis kill sites may be overcome by modeling "blitzkrieg," a rapid elimination of preferred prey, leaving minimal field evidence, by a potent human predator (Martin, 1990).

If aridification contributed to extinctions in Madagascar, it was evidently a potential problem only in the southwest. Sudden overkill by a "blitz," a sweep of the vulnerable megafauna, seems unlikely if the start-up of colonization took more than 1000 years. Robert Dewar (as cited in Goodman and Patterson, 1997) suggested that the impact of wild cattle on the native fauna rather than human hunting forced the extinctions.

All this may be resolved eventually by enrichment of the stratigraphic and geochronological record, including a major effort at extending the radiocarbon chronology of the extinct Malagasy megafauna into glacial times to determine if late glacial climatic change forced extinctions long in advance of human presence. Radiocarbon dates of 13,000 and 26,000 yr bp on a large ($\sim$70- kg) extinct lemur, *Megaladapis* (Simons as cited in Goodman and Patterson, 1997), indicate that glacial-age faunas exist. Apart from *Megaladapis*, virtually nothing is known of the glacial-age faunas of Madagascar. Like other extinct genera, *Megaladapis* survived until the end of the Holocene and may have been known historically.

Unlike Madagascar, rich glacial-age faunas are known in New Zealand, Tonga, and Hawaii. To date, no glacial-age extinctions have been detected in these islands, strengthening

the case for an anthropogenic agency as the unique cause of LQEE (Martin and Steadman, 1999).

Perhaps the most interesting case of comparative extinctions is the global decline of giant tortoises. Giant tortoises can be regarded as extremely vulnerable to human foragers. They survived historically only on remote warm-water oceanic islands that escaped prehistoric colonization, such as the Mascarenes, Seychelles, and Galapagos. In the late Holocene in Madagascar, giant tortoises vanished with the giant lemurs. In Africa they disappeared much earlier, perhaps more than 2 million years ago, an event that may represent the earliest example of an anthropogenic forcing of megafaunal extinctions.

Eurasia and Africa

Compared with the sweeping, catastrophic loss of late Quaternary megafauna in the New World, Madagascar, and arguably Australia, the losses in Eurasia were relatively minor and they occurred gradually (**Figure 7**; Martin and Steadman, 1999; Stuart, 1991). Outside the tropics, the extinct Eurasian fauna can be divided into two parts. One was a "warm" or interstadial fauna with straight-tusked elephants (*Elephas* (*Palaeoloxodon*) *antiquus*), temperate rhinoceroses (*Stenorhinus hemitoechus* and. *S kirchbergensis*), hippo (*Hippopotamus amphibius*), and a cave bear (*Ursus spelaeus*). The other was a "cold" or stadial fauna with woolly rhinoceros (*Coelodonta*

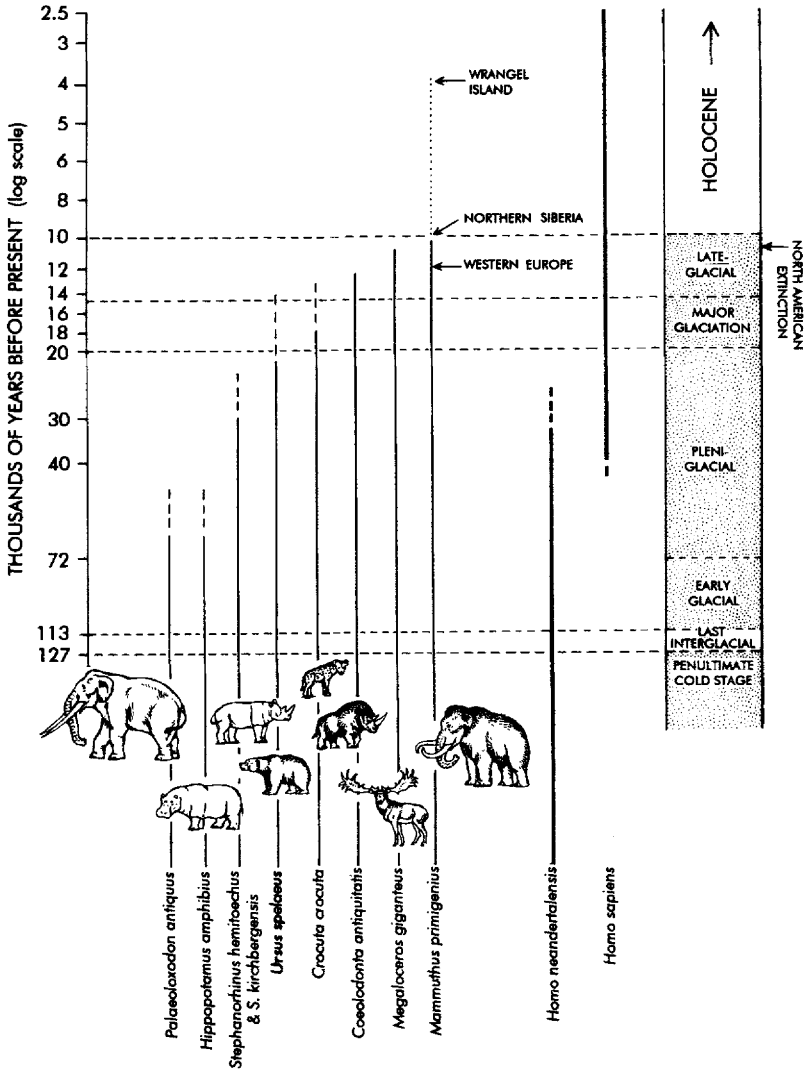


Figure 7 Extinction or extirpation chronology of large mammals in northern Eurasia. Note how the inferred extinction dates are very staggered compared with the strong pulse of extinctions in North America (right). During at least 50,000 yr bp the loss of Eurasian large mammals was sequential; the loss in North America was sudden. NE, northern and central Europe; It, Italy; Ja, Japan; SF, southern France; Sp, Spain; Lv, Levant; E, England; Sc, Scandinavia; Ir, Ireland; NS, north-central Siberia; *Elephas (Palaeoloxodon) antiquus*, straight-tusked elephant; *Hippopotamus amphibius*, hippopotamus; *Stenorhinus hemitoechus* and/or *S. kirchbergensis*, temperate rhinoceroses; *Ursus spelaeus*, cave bear; *Crocuta crocuta*, spotted hyena; *Coelodonta antiquitatis*, woolly rhinoceros; *Megaloceros giganteus, giant deer; Mammuthus primigenius, woolly mammoth* (Reproduced from MacPhee R (ed.) (1999) *Extinctions in Near Time*. New York: Kluwer/Plenum, and Stuart AJ (1991) Mammalian extinctions in the Late Pleistocene of Northern Eurasia and North America. *Biol. Rev.* 66: 453–562.).

antiquitatis), woolly mammoth (*Mammuthus primigenius*), and musk ox (*Ovibos moschatus*). This stepwise sequence of change in temperate and boreal faunas of Eurasia is gradual compared with the sudden extinction of both cold- (woolly mammoths and woodland musk oxen) and warm-adapted species (tapir and jaguar) in North America.

Intriguing differences in the pattern of Old and New World extinction are best illustrated in the case of proboscideans. Extinction of straight-tusked elephants did not begin until early in the last cold stage of the late Quaternary. Extinct throughout their European range by 60,000 years ago (Stuart, 1991), *Elephas* persisted in Japan until 30,000 years ago or less. The largest Holocene survivors in the megafauna of Japan were Sika deer (*Cervus*), wild boar (*Sus*), and black bear (*Selenarctos*). Although dwarf relatives of *Elephas* evidently did not survive on Mediterranean islands as late as the Holocene, as was once thought (Martin and Klein, 1984), mid-Holocene woolly mammoths have been found on Wrangel Island off Siberia, 6000 years younger than any other known mammoths—contemporary with the Pharaohs. The Akrotiri *Aetokremnos* bone cave, an epipaleolithic site on Cyprus, yielded dwarf straight-tusked elephants along with abundant bones of dwarf hippo. The deposit is radiocarbon dated at 10,500 years old. Just as Wrangel Island served as a refugium for the last of the boreal woolly mammoths, thousands of years after they were extinct throughout the rest of their range, Japan and Cyprus apparently sheltered the last populations of temperate Eurasian *Elephas* long after its extinction on the continent (Martin and Steadman, 1999).

Unlike the slow pace of extinction of the Proboscidea in northern parts of Eurasia, in North America prior to an extinction spasm 10,000–11,000 yr bp the animals suffered no loss of taxa, no clear-cut reduction in range, and no apparent decline in numbers (see maps in Graham and Lundelius, 1994). Had American Proboscidea followed the Old World extinction pattern for elephants, the temperate elephant of midlatitudes, the Columbian mammoth, would have predeceased high-latitude extinction of woolly mammoths by tens of thousands of years while the order would endure in tropical America, represented by a living species.

Both ivory carvings and cave paintings depicted by Stone Age artists reveal their intimate knowledge of the anatomy and behavior of Old World woolly mammoths and many other animals. Although field evidence of mammoth kills is scarce, low rates of increase of modern megaherbivores means that modern elephants, rhinoceros, and hippo would have been highly vulnerable to human predation (Owen-Smith, 1988). Even with governmental protection in this century, the future of surviving megaherbivores is by no means ensured. The gradual decline of Eurasian mammoths could readily have resulted from a minimal amount of human hunting or interference, especially if younger age classes were targeted.

Although the nature of the association and its meaning in modeling insular faunal extinctions in the Mediterranean are uncertain, the record at Akrotiri *Aetokremnos* suggests a Paleolithic "commando raid" in which Stone Age hunters or foragers overran an island, slaughtered preferred prey (dwarf hippo, dwarf elephant, and endemic deer), did not find sufficient resources for a sustainable economy, and soon left with no other evidence of their passage besides the remarkably rich contents of one cave. Possibly Crete, Ireland, Wrangel Island, and the Channel Islands of California also lost megafauna to commando raids that left scant evidence or at least none discovered to date. Given the scarcity of convincing kill or processing sites on continents, the rarity of such features on offshore islands is not surprising. Only in the remote Pacific such as on the low limestone islands in Tonga are extinct animals (bones of megapodes) abundantly associated with artifacts of the first colonists. In other cases (Hawaii and Madagascar), such associations are rare or unknown (MacPhee, 1999).

Africa has been held up to modelers of overkill as a conundrum. If people are such effective exterminators, how did the African megafauna survive? However, the argument cuts two ways. If late Pleistocene climatic change played a major role in triggering megafaunal extinctions elsewhere, how did so many large animals manage to survive in Africa and Asia, continents no more immune to Pleistocene climatic change than other corners of the globe?

One reply to the conundrum is that in Africa and Asia there was sufficient time to evolve a balanced predator-prey relationship, perhaps with human populations locally suppressed by sleeping sickness and other endemic human diseases and on occasion by intertribal buffer zones (Martin and Klein, 1984). The extinction record in both Africa and Asia is so unlike that found elsewhere on the planet that important differences in historical biogeography and human ecology may be postulated. The record of gradual human evolution and radiation confined to the Afro-Asian landmass, ending in the late Quaternary with explosive expansion of human colonists onto all other temperate landmasses, fits the model of a time transgressive overkill.

Oceanic Islands

Late Quaternary losses were limited neither to continents nor to mammals. They swept through oceanic islands. For example, on the larger oceanic islands the extinction of large terrestrial vertebrates included flightless moas in New Zealand; gorilla-size extinct lemurs, extinct hippo, giant tortoises, and giant flightless birds such as *Aepyornis* in Madagascar; house cat to bear-size ground sloths in the West Indies; dwarf elephants and dwarf hippo on islands of the Mediterranean; and dwarf elephants and/or giant tortoises on oceanic islands beyond the continental shelf in southeastern Asia such as Timor, Flores, and Sulawesi In one unusual case, prehistoric extinction only 4000 years ago eliminated elephants (woolly mammoths) not from a deep-water island but from a shelf island, Wrangel, in the Arctic Ocean (**Figure 2**).

With their discovery 25 years ago on Hawaii of a flightless goose (*Thambetochen*) and a flightless ibis (*Apteribis*), ornithologists Alexander Wetmore and Storrs Olson triggered the search for unknown extinctions on Pacific archipelagos. In the following years, archeologists and paleoecologists began to uncover rich fossil faunas reflecting prehistoric (Holocene) extinctions. During the past two decades their efforts on the Cooks, the Marquesas, the Kingdom of Tonga, and the Line Islands, to name a few examples from the South Pacific, have yielded many extinct species of small birds, especially flightless

rails, pigeons, and parrots (Steadman, 1995; Steadman as cited in Kirch and Hunt, 1997).

On the basis of fossil deposits from a sample of the 800 islands in the Pacific that are more than 1 km^2 in area, Steadman (1995) estimates the loss of 10 species or populations from each—a total loss of 8000 species or indigenous populations of land birds (including an estimated 2000 endemic species of taxonomically distinctive flightless rails). In addition to human harvesting and possible surplus killing, the introduction of Pacific rats may have eliminated many flightless rails and other birds. In some cases, the rats may even have colonized in advance of their human vectors (Holdaway as cited in MacPhee, 1999).

Extensive colonies of petrels, shearwaters, and other seabirds vanished from many Pacific islands, dimming the avian beacon which helped guide the prehistoric explorers to remote uninhabited islands (Steadman as cited in Kirch and Hunt, 1997). The fossil records of the relatively small number of islands that experienced FC and heavy extinction within historic time, such as the Galapagos, the Mascarenes, and remote Atlantic Islands such as St. Helena, have yet to reveal any pulse of prehistoric (Holocene) extinctions to match those found on islands that were colonized prehistorically.

New Zealand provides the best fossil record of island extinctions, starting with the loss of flightless wrens and small petrels, which were eliminated by accidental Polynesian introduction of Pacific rats (*R. exulans*), and beginning perhaps 1000 years in advance of Polynesian settlement. After Captain Cook's arrival in 1770, more species of mammalian predators reached New Zealand, including the dog, Norway rat, feral pig, feral cat, house mouse, black rat, ferret, and stoat (Holdaway as cited in MacPhee, 1999).

To bring prehistoric extinctions into focus, it is necessary to scan the globe and to probe the magnitude, timing, pattern, and natural history of all prehistoric losses on all landmasses, continents, and islands. This simple comparative approach in paleontology opens new vistas—ones that the popular MacArthur–Wilson model of island biogeography, based on modern distributions, will not reveal. High rates of vertebrate extinction on oceanic islands in the past five centuries, the focus of attention by conservation biologists (MacPhee and Flemming as cited in MacPhee, 1999), pale in magnitude to the extinction spasm of the previous four millennia.

Can extinctions of the LQEE be reversed? They certainly can if we do not insist on replacing lost species with taxonomically identical populations. On oceanic islands restoration ecologists need to consider restarting evolution of lost lineages of rails, using Guam rails on islands once occupied by flightless rails. In North America, a Pleistocene park should minimally include proboscideans, camelids, and equids as well as bison and other historic megafauna that survived the LQEE.

Pattern and Cause

Through an understanding of the pattern and timing of late Quaternary extinctions it may be possible to establish their cause. As the chronology, stratigraphy, paleoecology, and global pattern of the extinctions become better known, the ultimate questions of rate are clarified. For example, Darwin's view in *Origin of Species* that prehistoric extinctions of large mammals occurred gradually and sequentially is supported by the LQEE in Eurasia (**Figure 7**). Although an extraterrestrial accident such as an asteroid or cometary impact is often linked to some earlier mass extinctions, especially at the end of the Cretaceous, the space rock scenario can be ruled out in the late Quaternary. Over radiocarbon time the global pattern of loss was sequential or time transgressive, from more than 30,000 years ago in Australia to 1000 years ago in New Zealand (**Figures 1** and **2**), virtually eliminating any possibility of a one-shot global climatic catastrophe, such as a unique lethal cold snap, hyperdrought, or a great flood.

The prehistoric extinctions reviewed here are increasingly suspected of being the preamble of a vastly larger number to follow, a true mass extinction event in the making. Its cause would be anthropogenic.

Anthropogenic Models

The two major competing hypothesis or models to LQEE involve either anthropogenic or climatic forcing or both in combination (Burney, 1993; MacPhee, 1999; Martin and Klein, 1984). The anthropogenic model includes a variety of possible human impacts: direct overkill, surplus killing, a predator pit (doomed but still functioning predators, such as saber cats, reinforcing human predation), introduction of pandemic disease ("hyperdisease"), and human-initiated changes in habitat. The anthropogenic model is based on the close timing of extinctions to the global spread of anatomically and behaviorally modern people. In addition, some propose that Neanderthals or earlier hominids also had a role both in shaping the evolution of large mammal communities in the Old World and in the extinctions of the most vulnerable prey species providing rich resource packages such as giant tortoises, animals whose antipredator strategies were ineffective against humans. The kinds of animals lost in the LQEE appear to represent species known to be or likely to have been preferred prey.

A new player among the possible contenders for anthropogenically driven LQEE pulses is infectious disease, brought in at FC by human colonists themselves or, more likely, everything and anything that came along with them, such as domesticated, commensal, and synanthropic species. Especially on oceanic islands, the impact of human colonists would be amplified by the introduction of aliens, such as the Pacific rat *R. exulans*, chickens, domestic pigs, and dogs, along with their exotic diseases. The logic of hyperdisease is similar to that of classic overkill, but it contemplates humans as passive rather than active agents in causing faunal crashes. The basic idea is that "emerging" diseases, introduced into species without any natural immunity to them, would have induced incredible levels of mortality in susceptible populations, conceivably leading to their extinction. Once a "new" infectious disease process got started in a susceptible population, it would run its course quite independently of the rate or direction of human expansion.

Critics of anthropogenic models note that kill sites indicating human predation on and processing of allegedly preferred prey are rare, with few in the Americas and none in Australia or Madagascar and few in the Americas. The case of Madagascar is especially curious because its losses occurred only within the past 2000 years. Skeptics challenge the chronology of extinction, denying that within North America or Australia all extinct large mammals vanished abruptly, on the heels of human colonization, as implied by overkill. Some contend that humans arrived in both continents significantly before extinctions occurred. If human arrivals triggered megafaunal extinctions in the New World and Australia, critics ask, how do we explain the coexistence of *Homo sapiens* and large mammals in Africa and tropical parts of Eurasia? Vertebrate paleontologists have traditionally appealed to climatic or environmental changes as the main forcing function for many mammalian extinctions. During the past 65 million years the vast majority of extinctions occurred before any possibility of human involvement.

Climatic Models

Climatic models are based on the highly variable nature of late Quaternary environments with rapid switches from cold or cold and dry to warm and wet accompanied by changes in CO_2 pressure. In some cases, such as the Allerod–Younger Dryas shift within the late glacial, the switch appears to coincide with megafaunal extinctions, especially in North America. At least one habitat, the steppe–tundra of polar regions in the Northern Hemisphere, has been identified as an extinct biome (Guthrie as cited in Martin and Klein, 1984) whose end doomed woolly mammoths and other subarctic megafauna. Some LQEE models invoke a switch from less extreme to more extreme seasonality with out-of-step breeding cycles eliminating ruminants whose life cycles could not accommodate the climatic changes (Kilte and Graham and Lundelius as cited in Martin and Klein, 1984).

Critics discount late Quaternary climatic change as a forcing function since the paleoclimatic record of the Quaternary is rich in rapid and severe changes long before as well as during episodes of extinction. Fewer megafaunal extinctions occurred in the 3.5 million years combined, before human arrival, than in the late Quaternary, suggesting that species in Quaternary biotas were buffered against environmental switches typical of the last ice age. In any case, a late Quaternary extinction spasm is not evident in the oceans or in freshwater habitats and involves small animals only as parasites or on oceanic islands on which the reduced area would have amplified human impacts. Finally, the character of LQEE universally points to species that would either be preferred prey for human foragers and predators or be vulnerable (as in the case of minute endemic island snails) to aliens introduced from the continent.

Although details of how prehistoric humans might have triggered prehistoric extinctions are not easy to interpret from the fossil record, on a global scale the LQEE strongly reflects a deadly syncopation between extinction and human arrival. Viewed from the Cenozoic, Alroy (as quoted in MacPhee, 1999, p. 105) observed,

> The event's timing, rapidity, selectivity, and geographic pattern all make good sense according to the anthropogenic model and no sense at all otherwise. I believe that the overkill hypothesis, at least in general terms, already has been "proven" as thoroughly as any historical hypothesis can be.

See also: Extinction, Causes of. Mammals, Biodiversity of. Mammals (Pre-Quaternary), Extinctions of. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of

References

Burney DA (1993) Recent animal extinctions: Recipes for disaster. *Am. Sci.* 81: 530–541.
Flannery TF (1994) *The Future Eaters*. Melbourne: Reed Books.
Goodman S and Patterson B (eds.) (1997) *Natural Change and Human Impact in Madagascar*. Washington, DC: Smithsonian Institution Press.
Graham RW and Lundelius EL (1994) *Faunmap: A Database Documenting Late Quaternary Distributions of Mammal Species in the United States*, Scientific Papers, Vol. 25. Nos. 1 and 2. Springfield: Illinois State Museum.
Kirch PV and Hunt TL (eds.) (1997) *Historical Ecology in the Pacific Islands: Prehistoric Environmental and Landscape Change*. New Haven, CT: Yale Univ. Press.
Kurtén B and Anderson E (1980) *Pleistocene Mammals in North America*. New York: Columbia Univ. Press.
MacPhee R (ed.) (1999) *Extinctions in Near Time*. New York: Kluwer/Plenum.
Marshall LG and Cifelli RL (1990) Analysis of changing diversity patterns in Cenozoic land mammal age faunas, South America. *Paleovertebrata* 19: 169–210.
Martin PS (1990) 40,000 years of extinctions on the "planet of doom." *Palaeogeogr. Palaeoclimatol. Palaeoecol. (Global and Planetary Change Section)* 82: 187–201.
Martin PS and Klein RG (eds.) (1984) *Quaternary Extinctions: A Prehistoric Revolution*. Tucson: Univ. of Arizona Press.
Martin PS and Steadman DW (1999) Prehistoric extinctions on islands and continents. In: MacPhee R (ed.) *Extinctions in Near Time*, pp. 17–55. New York: Kluwer/Plenum.
Miller GH, Magee JW, Johnson BJ, *et al.* (1999) Pleistocene extinction of *Genyornis Newtoni*: Human impact on Australian megafauna. *Science* 283: 205–208.
Owen-Smith RN (1988) *Megaherbivores: The Influence of Very Large Body Size on Ecology*. Cambridge, UK: Cambridge Univ. Press.
Steadman DW (1995) Prehistoric extinctions of Pacific Island birds: Biodiversity meets zooarchaeology. *Science* 267: 1123–1131.
Stuart AJ (1991) Mammalian extinctions in the Late Pleistocene of Northern Eurasia and North America. *Biol. Rev.* 66: 453–562.

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF
Simon A Levin
Princeton University, Princeton, NJ, USA

VOLUME 5



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xv

Contributors to Volume 5 xxi

Foreword to the First Edition xxv

Preface to the First Edition xxvii

Preface to the Second Edition xxix

How to use of the Encyclopedia xxxi

VOLUME 5

Mammals (Pre-Quaternary), Extinctions of 1
William A Clemens

Mangrove Ecosystems 10
Peter J Hogarth

Marine and Aquatic Communities, Stress from Eutrophication 23
Jonathan H Sharp

Marine Conservation in a Changing Climate 32
Malin L Pinsky, Kristy J Kroeker, Daniel J Barshis, and Cheryl A Logan

Marine Ecosystems 45
J Frederick Grassle

Marine Ecosystems, Human Impacts on 56
Juan Carlos Castilla

Marine Mammals, Extinctions of 64
Glenn R Van Blaricom, Leah R Gerber, and Robert L Brownell Jr.

Marine Protected Areas: Static Boundaries in a Changing World 94
Elizabeth Mcleod

Marine Sediments 105
Paul VR Snelgrove

Marine Symbioses: Metazoans and Microbes 116
Elisha M Wood-Charlson

Marine Viruses 127
Grieg F Steward, Alexander I Culley, and Elisha M Wood-Charlson

Market Economy and Biodiversity 145
Heather Klemick and R David Simpson

Mass Extinctions, Concept of 156
J John Sepkoski Jr.

Mass Extinctions, Notable Examples of 167
Richard J Twitchett

Measurement and Analysis of Biodiversity 178
Wade Leitner and Will R Turner

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data 195
Nicholas J Gotelli and Anne Chao

Mediterranean-Climate Ecosystems 212
Philip W Rundel and Richard M Cowling

Megaherbivores 223
Norman Owen-Smith

Metapopulations 240
Peter Chesson

Microbial Biodiversity 252
Kristin L Matulich, Devon J Mohamed, and Jennifer BH Martiny

Microbial Biodiversity, Measurement of 259
Kate M Scow, Egbert Schwartz, Mara J Johnson, and Jennifer L Macalady

Microbial Biogeography 271
China A Hanson, M Claire Horner-Devine, Jennifer BH Martiny, and Jed A Fuhrman

Microbial Diversity 280
Paul V Dunlap

Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program 292
Kenneth H Nealson and Shibu Yooseph

Microorganisms (Microbes), Role of 299
Tom Fenchel

Migration 309
Daniel I Rubenstein and Mace A Hack

Modeling Biodiversity Dynamics in Countryside and Native Habitats 321
Henrique M Pereira and Luís Borda-de-Água

Modeling Marine Ecosystem Services 329
Anne D Guerry, Mary H Ruckelshaus, Mark L Plummer, and Dan Holland

Modeling Terrestrial Ecosystem Services 347
Erik Nelson, Nirmal Bhagabati, Driss Ennaanay, Eric Lonsdorf, Derric Pennington, and Manu Sharma

Modern Examples of Extinctions 362
Gábor L Lövei

Molluscs 373
David R Lindberg

Moths 384
David L Wagner

Museums and Institutions, Role of 404
Michael J Novacek and Suzann L Goldberg

Myriapods 421
Alessandro Minelli and Sergei I Golovatch

N

Natural Extinctions (not Human Influenced) 433
Christopher N Johnson

Natural Reserves and Preserves 442
Alexander N Glazer

Nature-Based Coastal Defenses: Can Biodiversity Help? 451
Bregje K van Wesenbeeck, John N Griffin, Mark van Koningsveld, Keryn B Gedan, Michael W McCoy, and Brian R Silliman

Near East Ecosystems, Animal Diversity 459
Joseph Heller

Near East Ecosystems, Plant Diversity 478
Avinoam Danin

Neotropical Seasonally Dry Forests 488
Julio Calvo-Alvarado, Arturo Sánchez-Azofeifa, and Carlos Portillo-Quintero

Nest Parasitism 501
Scott K Robinson, Stephen I Rothstein, and Brian D Peer

Neutral Theory and Beyond 510
James O'Dwyer and Ryan Chisholm

Nitrogen Deposition and Terrestrial Biodiversity 519
Christopher M Clark, Yongfei Bai, William D Bowman, Jane M Cowles, Mark E Fenn, Frank S Gilliam, Gareth K Phoenix, Ilyas Siddique, Carly J Stevens, Harald U Sverdrup, and Heather L Throop

Nitrogen, Nitrogen Cycle 537
Sandy L Tartowski and Robert W Howarth

Nomenclature, Systems of 547
David L Hawksworth

North America, Patterns of Biodiversity in 560
Martin J Lechowicz

Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms 568
Laura F Landweber, Tamara L Horton, and Jonatha M Gott

O

Ocean Ecosystems 581
Richard T Barber

Oceanic Islands: Models of Diversity 590
Rosemary G Gillespie

Oil-Palm Plantations in the Context of Biodiversity Conservation 600
Erik Meijaard and Douglas Sheil

Origin and Diversification of Angiosperms 613
William L Crepet

Origin of Life, Theories of 628
Susanne Brakmann

Origins, Evolution, and Breakdown of Bacterial Symbiosis 637
Joel L Sachs

P

Paleoecology 645
Thompson Webb

Parasitism 656
Klaus Rohde

Parasitoids 674
HCJ Godfray

Pelagic Ecosystems 683
Andrea Belgrano, Sonia D Batten, and Philip C Reid

Pesticides, Uses and Effects of 692
Paul C Jepson

Pharmacology, Biodiversity and 703
Paul Alan Cox

Phenological Shifts in Animals Under Contemporary Climate Change 716
Marcel E Visser

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 5

Yongfei Bai
Chinese Academy of Sciences, Beijing, China

Richard T Barber
Duke University, NC, USA

Daniel J Barshis
Hopkins Marine Station of Stanford University, Pacific Grove, CA, USA

Sonia D Batten
Sir Alister Hardy Foundation for Ocean Science, Plymouth, UK

Andrea Belgrano
Institute of Marine Research, Lysekil, Sweden

Nirmal Bhagabati
World Wildlife Fund – US, Washington, DC, USA

Luís Borda-de-Água
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

William D Bowman
INSTAAR University of Colorado, Boulder, CO, USA

Susanne Brakmann
Evotec Biosystems AG, Germany

Robert L Brownell Jr.
National Marine Fisheries Service, Pacific Grove, CA, USA

Julio Calvo-Alvarado
Instituto Tecnológico de Costa Rica, Cartago, Costa Rica

Juan Carlos Castilla
Pontificia Universidad Católica de Chile, Santiago, Chile

Anne Chao
National Tsing Hua University, Hsin-Chu, Taiwan

Peter Chesson
The University of Arizona, Tucson, AZ, USA

Ryan Chisholm
Smithsonian Tropical Research Institute, Panama, Republic of Panama

M Claire Horner-Devine
University of Washington, Seattle, WA, USA

Christopher M Clark
US Environmental Protection Agency, Washington, DC, USA

William A Clemens
University of California, Berkeley, CA, USA

Jane M Cowles
University of Minnesota, Saint Paul, MN, USA

Richard M Cowling
Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

Paul Alan Cox
Institute for Ethnomedicine, Jackson, WY, USA

William L Crepet
Cornell University, NY, USA

Alexander I Culley
University of Hawaii at Manoa, Honolulu, HI, USA

Avinoam Danin
The Hebrew University of Jerusalem, Jerusalem, Israel

Paul V Dunlap
University of Maryland Biotechnology Institute, Baltimore, MD, USA

Driss Ennaanay
Riverside Technology, Inc., Fort Collins, CO, USA

Tom Fenchel
University of Copenhagen, Copenhagen, Denmark

Mark E Fenn
Pacific Southwest Research Station, Riverside, CA, USA

J Frederick Grassle
Rutgers University, New Brunswick, NJ, USA

Jed A Fuhrman
University of Southern California, Los Angeles, CA, USA

Keryn B Gedan
Smithsonian Environmental Research Center, Edgewater, MD, USA

Leah R Gerber
Arizona State University, Tempe, AZ, USA

Rosemary G Gillespie
University of California, Berkeley, CA, USA

Frank S Gilliam
Marshall University, Huntington, WV, USA

Alexander N Glazer
University of California, Berkeley, CA, USA

HCJ Godfray
University of Oxford, UK

Suzann L Goldberg
American Museum of Natural History, New York, NY, USA

Sergei I Golovatch
Russian Academy of Sciences, Moscow, Russia

Nicholas J Gotelli
University of Vermont, Burlington, VT, USA

Jonatha M Gott
Case Western Reserve University, Cleveland, OH, USA

John N Griffin
University of Florida, Gainesville, Florida, USA

Anne D Guerry
Stanford University, Stanford, CA, USA

Mace A Hack
Nature Conservancy, NE, USA

China A Hanson
University of California, Irvine, CA, USA

David L Hawksworth
IUBS/IUMS International Committee on Bionomenclature, Ciudad Universitaria, Spain, and Natural History Museum, London, UK

Joseph Heller
The Hebrew University of Jerusalem, Jerusalem, Israel

Peter J Hogarth
University of York, York, UK

Dan Holland
National Oceanic and Atmospheric Administration, Seattle, WA, USA

Tamara L Horton
Princeton University, Princeton, NJ, USA

Robert W Howarth
Cornell University, Ithaca, NY, USA

Paul C Jepson
Oregon State University, OR, USA

Christopher N Johnson
James Cook University, Townsville, QLD, Australia

Mara J Johnson
University of California, Davis, CA, USA

Heather Klemick
National Center for Environmental Economics, Washington, DC, USA

Mark van Koningsveld
Van Oord Dredging and Marine Contractors, Rotterdam, The Netherlands, and Delft University of Technology, Delft, The Netherlands

Kristy J Kroeker
Hopkins Marine Station of Stanford University, Pacific Grove, CA, USA

Laura F Landweber
Princeton University, Princeton, NJ, USA

Martin J Lechowicz
McGill University, Montreal, QC, Canada

Wade Leitner
The University of Arizona, Tucson, AZ, USA

David R Lindberg
University of California, Berkeley, CA, USA

Cheryl A Logan
California State University Monterey Bay, Seaside, CA, USA

Eric Lonsdorf
Chicago Botanic Garden, Glencoe, IL, USA

Gábor L Lövei
Aarhus University, Slagelse, Denmark

Jennifer L Macalady
University of California, Davis, CA, USA

Jennifer BH Martiny
University of California, Irvine, CA, USA

Kristin L Matulich
University of California, Irvine, CA, USA

Michael W McCoy
University of Florida, Gainesville, FL, USA

Elizabeth Mcleod
The Nature Conservancy, Austin, TX, USA

Erik Meijaard
People and Nature Consulting International, Jakarta, Indonesia; University of Queensland, Brisbane, QLD, Australia, and Center for International Forestry Research, Bogor, Indonesia

Alessandro Minelli
University of Padova, Padova, Italy

Devon J Mohamed
Irvine Valley College, Irvine, CA, USA

Kenneth H Nealson
University of Southern California, Los Angeles, CA, USA

Erik Nelson
Bowdoin College, Brunswick, ME, USA

Michael J Novacek
American Museum of Natural History, New York, NY, USA

James O'Dwyer
Santa Fe Institute, Santa Fe, NM, USA

Norman Owen-Smith
University of the Witwatersrand, Wits, South Africa

Brian D Peer
Western Illinois University, Macomb, IL, USA

Derric Pennington
University of Minnesota, St Paul, MN, USA, and World Wildlife Fund – US, Washington, DC, USA

Henrique M Pereira
Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Gareth K Phoenix
University of Sheffield, Sheffield, UK

Malin L Pinsky
Princeton University, Princeton, NJ, USA

Mark L Plummer
National Oceanic and Atmospheric Administration, Seattle, WA, USA

Carlos Portillo-Quintero
Instituto Venezolano de Investigaciones Científicas (IVIC), Maracaibo, Venezuela

Philip C Reid
Sir Alister Hardy Foundation for Ocean Science, Plymouth, UK

Scott K Robinson
University of Florida, Gainesville, FL, USA

Klaus Rohde
University of New England, NSW, Australia

Stephen I Rothstein
University of California, Santa Barbara, CA, USA

Daniel I Rubenstein
Princeton University, Princeton, NJ, USA

Mary H Ruckelshaus
Stanford University, Stanford, CA, USA

Philip W Rundel
University of California, Los Angeles, CA, USA

Joel L Sachs
University of California, Riverside, CA, USA

Arturo Sánchez-Azofeifa
University of Alberta, Edmonton, AB, Canada

Egbert Schwartz
University of California, Davis, CA, USA

Kate M Scow
University of California, Davis, CA, USA

J John Sepkoski Jr.[†]
University of Chicago, IL, USA

Manu Sharma
Stanford University, Stanford, CA, USA

Jonathan H Sharp
University of Delaware, Newark, DE, USA

Douglas Sheil
Mbarara University of Science and Technology, Kabale, Uganda; Center for International Forestry Research, Bogor, Indonesia, and Southern Cross University, Lismore, NSW, Australia

Ilyas Siddique
Universidade Federal de Santa Catarina, Brazil

Brian R Silliman
University of Florida, Gainesville, FL, USA

R David Simpson
National Center for Environmental Economics, Washington, DC, USA

Paul VR Snelgrove
Memorial University of Newfoundland, St. John's, NL, Canada

Carly J Stevens
The Open University, Milton Keynes, UK, and Lancaster University, Lancaster, UK

Grieg F Steward
University of Hawaii at Manoa, Honolulu, HI, USA

[†]Deceased.

Harald U Sverdrup
Lund University, Lund, Sweden

Sandy L Tartowski
US Department of Agriculture, USA

Heather L Throop
New Mexico State University, Las Cruces, NM, USA

Will R Turner
The University of Arizona, Tucson, AZ, USA

Richard J Twitchett
Plymouth University, Plymouth, UK

Glenn R Van Blaricom
US Geological Survey and University of Washington, Seattle, WA, USA

Marcel E Visser
Netherlands Institute of Ecology, The Netherlands

David L Wagner
University of Connecticut, Storrs, CT, USA

Thompson Webb
Brown University, RI, USA

Bregje K van Wesenbeeck
Unit for Marine and Coastal Systems, Deltares, MH, Delft, The Netherlands

Elisha M Wood-Charlson
University of Hawaii at Manoa, Honolulu, HI, USA

Shibu Yooseph
J. Craig Venter Institute, San Diego, CA, USA

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

Mammals (Pre-Quaternary), Extinctions of

William A Clemens, University of California, Berkeley, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Background extinction Rates of extinction that characterized the major part of the evolution of life.
Clade A group of organisms including its common ancestor and all its descendants.
Crown group A clade including a group of modern species, the common ancestor, and all its descendants, including extinct lineages.
Extinction The termination of a lineage of organisms.
Mass extinction A short period of Earth history in which the rate of extinction reached an exceptional high level.
Paraphyletic A group of organisms that includes the common ancestor but not all of its descendants.

Mammals and Other Synapsids

The evolutionary history of synapsids, which includes mammals and their ancestors, began with the origin of the group in the Carboniferous Period approximately 359–299 million years ago (Ma) (Table 1). Prior to this interval, terrestrial vertebrates underwent their first evolutionary radiation. One of these lineages, characterized by a specialized type of egg (the amniote egg), split into at least two clades: reptiles (including birds) and synapsids (including nonmammalian synapsids and crown group mammals; Figure 1). Early members of these clades can be distinguished by different structural patterns of the posterior parts of their skulls. Clearly, part of the ancestry of mammals, early nonmammalian synapsids, lacked specialized features of skull structure that characterize members of the crown group Mammalia – for example, only one pair of bones (the dentaries) making up the lower jaw or the presence of three bones (malleus, incus, and stapes) in the middle ear. Direct evidence is not preserved in the fossil record, but it is probable that these early synapsids lacked hair, mammary glands in the females, and endothermy – traits commonly used to distinguish modern mammals from reptiles.

Unfortunately, in many older and some recent texts, early synapsids are referred to as mammal-like reptiles. These early

Table 1 Geological timescale for the interval of earth history considered in the text[a]

Era	*Period*	*Epoch*	*Age of boundary (Ma)*
	Quaternary	Pleistocene	2.8
		Pliocene	5.3
		Miocene	23.8
Cenozoic	Tertiary	Oligocene	33.9
		Eocene	55.8
		Paleocene	65.5
	Cretaceous		145.5
Mesozoic	Jurassic		201.6
	Triassic		251.0
	Permian		299.0
Paleozoic	Carboniferous		359

[a]Ages of boundaries taken from the 2009 Geological Time Scale, Geological Society of America.

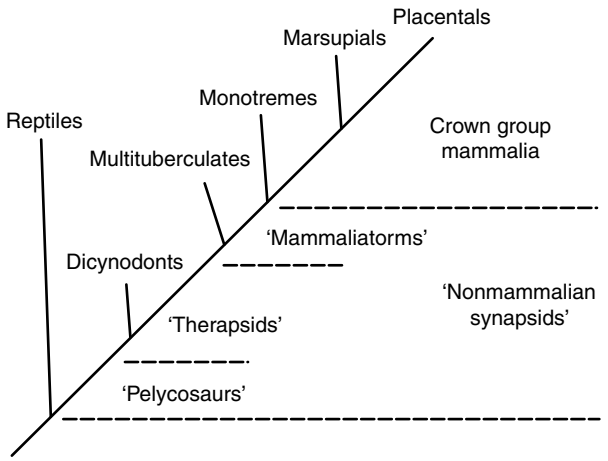


Figure 1 Diagrammatic representation of the phylogenetic relationships of the groups of nonmammalian synapsids and crown group Mammalia. Names of paraphyletic groups are enclosed in quotation marks.

synapsids definitely were not members of the ancestry of reptiles. Morphological similarities to reptiles are the product of inheritance and retention of traits found in their common ancestor. In the following discussion, these early synapsids will be dubbed nonmammalian synapsids, which accurately expresses the fact that they were members of the early part, the stem, of the synapsid clade but not members of the crown group Mammalia (**Figure 1**).

In a broad-brush summary, the evolutionary history of the synapsids can be divided into three phases. Beginning in the Carboniferous, the first evolutionary radiation of nonmammalian synapsids was interrupted by the mass extinction event used to mark the Permian–Triassic boundary. The surviving lineages of nonmammalian synapsids radiated again but did not regain their position as a dominant group of terrestrial vertebrates. Prior to or during the next mass extinction, probably soon before the Triassic–Jurassic boundary (approximately 201.6 Ma), almost all lineages of nonmammalian synapsids became extinct. The few surviving lineages were made up of animals of small body size, sizes equivalent to those of modern mice or shrews. The following Jurassic and Cretaceous Periods, which lasted from approximately 201.6–65.5 Ma, constituted the Age of Dinosaurs. This interval witnessed a major evolutionary radiation that included the origin of the crown group Mammalia. The Cretaceous ended with another mass extinction event; this was the time of extinction of dinosaurs other than birds. During the following Tertiary Period (65.5–2.8 Ma), mammals flourished and their diversity rapidly increased. Biogeographic differentiation, extinction events, and periods of faunal turnover complicated Tertiary evolutionary radiations of mammals.

Patterns and Causes of Extinctions

Although discussed in more detail later, a few issues relating to the nature and causal factors of extinctions need to be reviewed here as the background for analysis of the role of pre-Quaternary extinctions in shaping the course of mammalian evolution.

Extinctions and Pseudoextinctions

In any attempt to assess changes in biodiversity resulting from extinctions, care must be taken to differentiate between true, biological extinctions, and what have been termed as pseudoextinctions. In a biological sense, extinction is the termination of the group or groups of organisms comprising a species. No longer was there successful reproduction and survival of the young to sexual maturity, population size dwindled to zero, and the species, as well as the lineage it represented, was lost from the biota.

Pseudoextinction refers to an artifact produced by zoological nomenclature; a named unit becomes extinct but the lineage it represents survives. The taxonomy (the system of recognition and naming of groups of organisms) of the Equidae, the family that includes modern horses, asses, and zebras, provides many examples of pseudoextinction. The oldest known species of the family are a group of lineages included in the genus *Hyracotherium* (= *Eohippus*). These were diminutive, multitoed inhabitants of tropical forests. The fossil record documents increase in body size, loss of toes, and other significant morphological changes in descendants of *Hyracotherium* – for example, *Orohippus*, *Miohippus*, *Merychippus*, and *Dinohippus*. These extinct genera are the members of the ancestry of modern horses, zebras, asses, and other closely related species, all of which are included in the genus *Equus*.

The genus *Hyracotherium* is composed of the common ancestor of all the species referred to as *Hyracotherium* including the lineage leading to *Orohippus* the next member of the Equid family tree. In the jargon of modern taxonomy, *Hyracotherium* is paraphyletic, a unit that includes its common ancestor and some, but not all, of the common ancestor's descendants. All genera of the family tree ancestral to *Equus* are now extinct. The extinctions of these paraphyletic groups, however, did not mark the termination of the equid lineage; they are products of the system of recognizing and naming organisms. (To highlight their status, the names of paraphyletic groups are often enclosed in quotation marks.)

It also must be stressed that in most studies of pre-Quaternary extinctions, species are not the basic units considered. Usually these studies analyze the differences in numbers of genera or families, a procedure that opens the possibility of significantly muting patterns of change in biodiversity. For example, if a genus consists of only one species, loss of the species also results in the loss of the genus. However, if the genus contains many species, the loss of one species is not reflected at the generic level. Thus, a major extinction event that resulted in the loss of 60% of the species extant in a region might be reflected in the loss of only 30% of the genera or 10% of the families represented.

Background and Mass Extinctions

Since the origin of life approximately 3.5 billion years ago, extinction has been a major process influencing the course of its evolution. Current estimates from the fossil record indicate that more than 90% of the species that once existed are now extinct; some suggest that the number is more than 99%. Modern studies of the patterns of extinction have their roots in the analyses of the times of origin and extinction of families and genera of marine invertebrates, particularly those carried out by Sepkoski, Raup, and their associates (Raup, 1991). These studies and prior research indicated that the rates of extinction fall into two categories, background extinctions and mass extinctions. Background extinction rates appear to characterize the majority of the history of life. Several times during the past 600 Ma, however, rates of extinction increased greatly over short periods of geological time. Taxa that were not particularly closely related and inhabited diverse areas of the oceans became extinct. These events were recognized as mass extinctions. The most severe of the mass extinction events are identified as the "Big Five" mass extinctions.

Although the available database for evolution of terrestrial vertebrates and plants is not as biogeographically extensive and detailed as the marine record, enough is known to show that some of these mass extinctions modifying the marine

biota were correlated in varying degrees with major changes in the evolution of terrestrial life. In particular, the three most recent mass extinctions marking the Permian–Triassic, Triassic–Jurassic, and Cretaceous–Tertiary boundaries in the marine realm were associated with times of significant modifications in the course of evolution of nonmammalian synapsids and early mammals.

Temporal and Biogeographic Scales of Mass Extinctions

Mass extinctions were defined subjectively as short periods of Earth history during which rates of extinction reached exceptionally high levels in widespread areas. In evolutionary terms, how short is short? In studies of recent extinctions in which events are followed on an ecological timescale, short periods of time are measured in terms of years, decades, centuries, or possibly a millennium or two. A recent compilation of recent estimated durations of the last three of the "Big Five" mass extinctions revealed values as much as 8.3 Ma to as short as less than a year (Barnosky *et al.*, 2011).

Current radiometric methods for determining ages and durations of pre-Quaternary events lack the degree of resolution found in ecological studies of the modern biota. For example, using the $^{40}Ar/^{39}Ar$ method of age determination, which is one of the most precise methods of radiometric age determination available for pre-Quaternary deposits, the age of the Cretaceous–Tertiary boundary, as recognized in nonmarine sediments in eastern Montana, is placed at 65.58 ± 0.04 Ma (Wilson, 2005). As data on extinctions of Cretaceous lineages of nonmarine and marine organisms are collected and correlated, all extinction events that occurred within an interval of 80,000 years would necessarily be treated as having occurred simultaneously. Recognition of the role of short-term modifications of the biota on an ecological timescale and their consequences in terms of background and mass extinctions on a geological timescale remains a major challenge.

In comparison to the fossil record of marine invertebrates, which is the foundation for studies of the role of mass extinctions in evolution, the fossil record of terrestrial vertebrates is much more limited, particularly in its biogeographic coverage. For example, detailed records of mammalian evolution across the Cretaceous–Tertiary boundary are limited to sites in the Western Interior of North America (Archibald, 2011). Significant gaps in the fossil records of mammalian evolution in Eurasia and the continents of the Southern Hemisphere characterize this interval. One consequence of these gaps is the uncertainty in determining the origins of mammalian groups that appear soon after a mass extinction event. Were they descendants of a few local survivors or immigrants from another less impacted and yet to be sampled area?

Distal and Proximal Causal Factors

In studies of extinction, a distinction must be made between distal and proximal causal factors. Proximal causal factors are those that impinge directly on individuals and are the immediate causes of their death. Distal causal factors are the regional or global events that result in changes in local (proximal) biological and physical environments affecting individual organisms. In the much longer period of mammalian evolution before the Quaternary, several distal causal factors emerge as major contributory factors to mass extinctions and other significant changes in the course of evolution of the group.

Recognition of some distal causal factors can be traced to the development of our understanding of plate tectonics and the mechanisms involved in movements and changes in configurations of continents and oceanic basins. Among the immediate consequences of modification of the configuration of these basins are changes in their coastlines. Hallam and Wignall (1997) and other workers noted that many of the major mass extinctions were associated with times of withdrawal of shallow continental seas (marine regressions) and reduction of coastal habitats. Additionally, changes in patterns and strengths of oceanic currents coupled with changes in continental topography contributed to modification in global climates. Another product of the changes in configurations of continents during periods of marine regression was the establishment of terrestrial connections between long-isolated continents. For example, the intermittent formation of dry land connections between Siberia and Alaska at various times during the Mesozoic and Tertiary facilitated mammalian dispersal and probably contributed to mammalian extinctions through some combination of competition, predation, and possibly introduction of new diseases.

Volcanic activity has played and continues to play a significant role in modifying the earth's environments. Major eruptions produce large quantities of particulate matter, aerosols, and acids. These clouds can reach magnitudes sufficient to limit or block the sun's radiation from reaching the surface of the earth, causing a decrease in temperature. Later, they would trap heat radiated from the surface to produce a greenhouse warming of the climate. Increased amounts of acid rain also are predictable consequences of volcanic activity. Late in the Cretaceous, high levels of volcanic activity resulting in the formation of the Deccan Traps in India have been hypothesized as a distal causal factor of the mass extinction at the end of the Period.

Although the hypothesis had precursors in earlier decades, the assertion by Alvarez *et al.* (1980) that the impact of an asteroid was the causal factor of the Cretaceous–Tertiary mass extinction focused attention on the potential role of extraterrestrial bombardment in controlling the course of evolution of the earth's biota. Their initial hypothesis argued that the impact of an asteroid was the distal cause of the extinction of nonavian dinosaurs and numerous other kinds of terrestrial and marine organisms. The hypothesized proximal causes were the formation of dust clouds that first produced global refrigeration by blocking solar energy and then a greenhouse effect by trapping heat in the atmosphere. Impact-generated global acid rain and wildfires also were suggested as contributory factors, but the severity and even the occurrence of both has been challenged by recent research.

It has been argued that a mass extinction is an exceptional event, and therefore, requires an exceptional causal factor. Two aspects of this assertion require comment. The level of ecological interdependence of members of the terrestrial biota

has, no doubt, varied through time. Terrestrial ecosystems characterized by relatively low levels of interdependence of their members might well be able to withstand the loss of one or a few species. In contrast, ecosystems with highly interdependent members might not be able to withstand the loss of one or a few species, resulting in the collapse of the whole system and consequent extinction of many other species. Thus, a period of global warming, for example, could have a catastrophic effect on highly interdependent ecosystems, but causes little or no change in those that were characterized by a lower level of interdependence.

Secondly discussions of the causal factors of mass extinctions have too often been simplified into debates favoring one factor, marine regressions of the seas, outbursts of volcanic activity, or impacts of large extraterrestrial bodies, and discounting the effects of the others. Analyses of the geological and fossil records of the Mesozoic and Cenozoic by Arens and West (2008) revealed that the highest frequency of intervals with elevated extinction were found when high levels of volcanic activity producing extensive flood basalts and impacts coincided. Building on the work of community ecologists and other paleontologists, for example, Erwin (2001), they argued that high levels of volcanic activity or climatic change have produced long-term modifications in community structure, the press disturbances. In contrast, pulse disturbances were short-term events that produced abrupt modifications of community structure and may be associated with significant mortality. They hypothesize that episodes of mass extinction probably are not attributable to a single cause, but were the consequences of combinations of press and pulse disturbances.

Finally, the focus of this article is on extinctions of lineages and their effects on patterns of evolution and diversity. Extinction is but one-half of the equation for calculating changes in biodiversity. Variation in the rate of origination of new species can have a marked effect on biodiversity. For example, analysis of the fossil record of nonavian dinosaurs during the Jurassic and Cretaceous shows that they evolved rapidly; a genus of nonavian dinosaur rarely lasted for more than a geological stage (on average 6 Ma in duration) or two. Throughout most of their history, rates of extinction and origination of nonavian dinosaurs were high. In the last two stages prior to their extinction at the end of the Cretaceous, a high rate of extinction of genera appears to have been maintained, but the rate of origination of new genera was depressed.

Origin of Synapsids and the Permian–Triassic Mass Extinction

During the Carboniferous Period (**Table 1**), approximately 359–299 Ma, the reptilian and synapsid clades differentiated. From this beginning, the first phase of synapsid evolution, characterized by their prominent position in terrestrial faunas, would last until near the end of the Triassic. It was interrupted, and the course of synapsid evolution altered, by the massive extinctions that marked the Permian–Triassic boundary.

In the Late Carboniferous and Early Permian, basal nonmammalian synapsids, the pelycosaurs, were the most diverse terrestrial vertebrates but initially limited to the hot, wet environments of the equatorial coal swamps. The pelycosaur-dominated terrestrial faunas of the Late Carboniferous were composed of many more carnivores than herbivores; at first glance, this appears to be an ecologically anomalous situation. It is now thought that the dominance of carnivores is not a product of a bias in preservation or collecting, but reflects an initial stage in the evolution of a terrestrial vertebrate ecosystem. Many of the species of these early nonmammalian synapsids may well have been amphibious, returning to streams and lakes to feed on fish and freshwater invertebrates.

Gradual changes in the climate during the Early Permian resulted in the development of drier, seasonally arid areas. At first, terrestrial herbivorous pelycosaurs became more abundant as a fully terrestrial vertebrate ecosystem evolved. Then, about the beginning of the Late Permian, a rapid radiation of therapsids began (**Figure 1**). This radiation produced a variety of carnivorous and herbivorous forms, including the abundant dicynodonts. It occurred during a time of gradually changing climate and flora. The expansion of the biogeographic range of the therapsids beyond equatorial regions and decline in the diversity of pelycosaurs add to the evidence suggesting that the diversification of therapsids was related to the evolution of fully terrestrial ecosystems well before the end of the Permian (Kemp, 2005).

The mass extinction used to mark the end of the Permian was the most severe yet recorded. It has been estimated that more than 90% of the species and 50% of the families of marine invertebrates were lost during this event. Terrestrial ecosystems of the Late Permian were made up of a variety of therapsids as well as a few lineages of reptiles and amphibians. These too were greatly decimated. Of the approximately 50 families of terrestrial vertebrates present in the last 5 Ma of the Permian, approximately 75% died out. This dramatic loss included large and small herbivores and carnivores. Notable survivors among the nonmammalian synapsids were a few lineages of carnivorous therapsids and the herbivorous dicynodont, *Lystrosaurus*.

The debate continues over the causal factors of the mass extinction marking the end of the Permian (Benton *et al.*, 2003). A variety of lines of evidence suggest that this was a time of global warming, probably related to a marine regression followed by development of anoxic oceans. Extensive volcanic activity, the eruption of the massive flood basalts of western Siberia, was another distal causal factor. This was also a time of elevated concentrations of CO_2 and H_2S in the atmosphere. Currently, students of the Permian–Triassic mass extinction appear to favor some combination of changes in the oceans and volcanic activity as its distal causal factors. No convincing evidence of the impact of an asteroid or other extraterrestrial body has been discovered.

The Second Therapsid Radiation and Late Triassic Faunal Turnover

As a result of the decimation caused by the Permian–Triassic mass extinction, the taxonomic diversity of the earliest Triassic terrestrial faunas was greatly reduced. These faunas were dominated by a dicynodont, *Lystrosaurus*, but included

representatives of a few other lineages of nonmammalian synapsids, which were the basal stocks of another but lesser evolutionary radiation. Later, in the Early Triassic, large and small herbivorous forms, primarily dicynodonts, became prominent elements of terrestrial faunas. In parallel, the surviving carnivorous therapsids radiated, producing a variety of new lineages of large and small carnivores as well as some herbivorous species. In comparison to Permian terrestrial faunas, synapsids did not regain their dominant position in Triassic terrestrial vertebrate faunas.

Analysis of the evolutionary radiation of the carnivorous therapsids during the Triassic reveals trends in modification of skull structure and morphology of the postcranial skeleton that increasingly came to resemble those characterizing modern mammals (Kemp, 2005). For example, the lower jaw was modified through expansion of the dentary bone and the reduction or losses of other bones characteristic of early synapsid or reptilian jaw structure. These were preliminary steps toward acquisition of an articulation of the dentary bone of the jaw with the squamosal bone of the skull, a key skeletal character for distinguishing modern mammals from reptiles. Limb posture began to be modified from a primitive, sprawling stance to a more typically mammalian stance with the elbow and knee tucked close to the body and the limbs brought into an upright posture.

A few million years before the end of the Triassic, some synapsids evolved a functional articulation between the dentary and squamosal. Many textbooks and popular articles present the acquisition of this structure as the hallmark of the Mammalia and identify the end of the Triassic (approximately 251 Ma) as the time of origin of the group. In current phylogenetic classifications, however, the clade Mammalia is restricted to a crown group including the common ancestor of monotremes (the echidna and platypus), marsupials, and placentals and all its descendants. Phylogenetically, some of the very advanced nonmammalian synapsids with a dentary–squamosal jaw articulation and other mammal-like specializations lie outside the crown group; they are sister lineages. In many recent research papers, these very advanced nonmammalian synapsids are referred to as mammaliaforms.

During the Late Triassic, the composition of terrestrial faunas underwent a major change that played a significant role in reshaping the courses of evolution of synapsids and other vertebrates. All the lineages of Triassic synapsids, except for a few lineages of mammaliaforms and one group of herbivorous nonmammalian synapsids, the tritylodonts, became extinct by the end of the Triassic. All the surviving lineages were represented by species of relatively small body size. During this interval two groups of reptiles, the rhynchosaurs and later several lineages of dinosaurs became the dominant terrestrial vertebrates.

Identification of the causal factors behind the Late Triassic decline of therapsids and the contemporaneous increasing diversity of dinosaurs as well as the mass extinctions in both the marine and terrestrial realms marking the end of the Triassic continues to be a matter of debate. It has been argued that the early dinosaurs were in some way competitively superior in their interactions with the therapsids and this was the cause of extinction of these lineages. More likely, the decline in diversity and demise of some therapsid lineages in the Late Triassic is attributable to changes in the environment. Taking advantage of the losses of these lineages, dinosaurs radiated opportunistically. Extensive eruptions of flood basalts in rifts developed by the initial separation of North America from Europe and Africa (the Central Atlantic Magmatic Province) and associated modifications of the atmosphere appear to have been distal causal factors of the Triassic–Jurassic boundary mass extinctions.

Mammalian Evolution During the Age of Dinosaurs

During the Jurassic and Cretaceous, popularly dubbed the Age of Dinosaurs, mammaliaforms and early mammals were not dominant members of terrestrial faunas. Functional interpretations of their dentitions suggest that most were probably carnivorous, feasting on small prey including small terrestrial vertebrates and invertebrates. For the most part, they were very small animals, in the size range of modern mice and rats. A few evolved larger body sizes, rivaling modern opossums or raccoons, but these species were exceptions to the rule.

Recent discoveries reveal that in spite of their relatively small size during the Jurassic, mammaliaforms evolved diverse ecomorphological specializations (Luo, 2007). For example, two lineages of closely related mammaliaforms evolved aquatic specializations, one with a beaver-like flattening of their tails and another with an otter-like transverse compression of their tails. Members of another mammaliaform lineage developed folds of skin linking their neck, limbs, and tail to produce membranes resembling those of modern gliding squirrels.

For the most part, these mammaliaform lineages became extinct by the middle of the Cretaceous. One group that appears in the fossil record in the Jurassic, the multituberculates, evolved a remarkably rodent-like style of dental specialization and is interpreted as having been omnivorous or herbivorous in dietary preferences. This interpretation is strengthened by the persistence of the group into the Paleocene of the Northern Hemisphere. Then, in the late Paleocene as rodents began their evolutionary radiation, the taxonomic diversity of multituberculates dwindled. By the end of the Oligocene, when almost all families of rodents had differentiated, multituberculates had become extinct.

Tracing the course of mammalian evolution during this phase of synapsid evolution is complicated by distinct biogeographic differentiation. Unlike the Permian and Triassic, when most of the terrestrial regions of the world were parts of one supercontinent, Pangaea, through the Jurassic and Cretaceous terrestrial areas and their faunas were fragmented. In part, this was the product of plate tectonic processes that broke up the Pangaean supercontinent and shifted the positions of the resulting continental blocks. Additionally, the Age of Dinosaurs was a time characterized by extensive marine transgressions when shallow seas covered many areas and added to fragmentation of terrestrial areas. By the beginning of the Cretaceous, when the currently sparse fossil record of these faunas gives us our first real picture of biogeographic differentiation, there is evidence of distinct northern (North American–Eurasian) and southern (Australian–Antarctic–South

American) terrestrial faunas. Unfortunately, very little is known of the history of mammaliaform and mammalian evolution in the African continent during this interval.

The Cretaceous also was a time of major changes in the terrestrial biota. Angiosperms made their appearance and began to diversify. First records of most modern families of insects document their diversification. Particularly in the Northern Hemisphere, dinosaurian faunas underwent a major reorganization. Global climates of the Jurassic and Cretaceous were, in general, distinctly warmer and more equable than modern climates. The latitudinal climatic gradient, reflecting the degree of difference between equatorial and polar conditions, was much lower. Throughout most of the interval, there is no evidence of continental glaciations. Finally, to round out this description of the environmental settings of mammalian evolution during the Age of Dinosaurs, it must be noted that evolution of the terrestrial biota was not interrupted by a mass extinction event.

By the middle of the Cretaceous, the mammalian fauna had undergone a major turnover in composition. In the Northern Hemisphere, with the exception of the hardy multituberculates, most of the Jurassic lineages of mammaliaforms had dwindled in taxonomic diversity or became extinct. Currently, the first record of the crown group Mammalia is from the Northern Hemisphere and of Late Jurassic age (Luo *et al.*, 2011). By the Early Cretaceous two of the three surviving lineages of the Mammalia, metatherians (including marsupials) and eutherians (including placentals) had begun to radiate here. Mammalian evolution on a southern continent, composed of the modern Australian, Antarctic, and South America continents, followed a different course. Metatherians and eutherians have yet to be discovered there. The terrestrial fauna included the first records of the third major lineage of the Mammalia, the monotremes. Additionally, at least two groups of mammaliaforms diversified there and survived until at least the end of the Cretaceous. This biogeographic dichotomy strongly suggests that the common ancestor of the crown group Mammalia will be found in the Middle or Early Jurassic, if not earlier.

Cretaceous–Tertiary Mass Extinction and its Consequences

Only the effects of the Permian–Triassic mass extinction overshadow the extent of devastation of marine and terrestrial biotas during the Cretaceous–Tertiary mass extinction. On land, the Cretaceous–Tertiary mass extinction is marked by the demise of nonavian dinosaurs and some other vertebrate lineages (Dingus and Rowe, 1998; Fastovsky and Weishampel, 2005). Similarly, many dominant groups of marine invertebrates and aquatic reptiles died out at this time.

The primary source of information documenting the effects of the Cretaceous–Tertiary mass extinction on mammalian evolution is a series of fossil localities in an area of the Western Interior of North America extending from Colorado northward into Alberta. For decades, a wide range of geological and paleontological research projects have been under way in this area, and the resulting database is extensive. This is both a curse and a blessing. There has been a tendency to regard the results of studies of biotic change in the North American Western Interior as typical of global patterns. Recent discoveries show that this was not the case. However, although geographically limited, the extensively documented patterns of survival and extinction in the North American Western Interior can be used to test the hypotheses concerning the nature and severity of environmental changes that caused the Cretaceous–Tertiary mass extinction.

Wilson (2005), Archibald (2011), and others have analyzed the patterns of survival and extinction of terrestrial vertebrates based on extensive collections from northeastern Montana. During the Cretaceous–Tertiary mass extinction in this area, all the nonavian dinosaurs, approximately 20 genera, became extinct. A small group of freshwater sharks and their relatives shared the fate of the nonavian dinosaurs. These losses, plus extensive extinctions among the Cretaceous lineages of lizards and marsupials, account for almost 75% of the extinctions of the terrestrial vertebrate species in the North American Western Interior. In contrast, no extinctions of lineages of salamanders and frogs have been documented. Only a few species of turtles, crocodilians, and crocodilian-like reptiles died out at this time. Several paleontologists have begun to analyze the very fragmentary fossil record of the evolution of birds in the Late Cretaceous. Their research, including molecular studies of modern forms, shows that the evolutionary diversification of the crown group of birds had begun in the Cretaceous. Although many lineages of more primitive birds died out before or at the Cretaceous–Tertiary boundary, extinction appears to have taken little toll among the basal lineages of the crown group. Among the mammalian lineages present in the North American Western Interior, approximately 60% of the species of multituberculates, which are thought to be the ecological equivalents of rodents, became extinct during the Cretaceous–Tertiary mass extinction. Almost all lineages of eutherian mammals probably survived this event. In stark contrast, all but one or two of the Cretaceous lineages of marsupials appear to have become extinct at this time. The currently available fossil records from other continents in the Northern Hemisphere are characterized by the lack or very limited documentation of their latest Cretaceous and earliest Tertiary vertebrate faunas.

Recent discoveries in Australia and, particularly, South America are beginning to outline the course of mammalian evolution through the Cretaceous–Tertiary mass extinction in the Southern Hemisphere. At this time, Australia, Antarctica, and South America were closely approximated and global climates were much warmer. Direct evidence documents at least some interchange of mammals between Australia and South America. In South America, deposits laid down several million years prior to the end of the Cretaceous contain records of a mammalian fauna dominated by members of several lineages of mammaliaforms accompanied by a lineage of monotremes. A gap in the fossil record separates records of this fauna from deposits documenting the earliest Tertiary faunas known from the area. The change in mammalian faunal composition is striking. Most of the mammaliaform lineages had disappeared. A lineage of monotremes survived into the Tertiary but soon became extinct. The dominant mammals in the earliest Tertiary of South America included a small but diverse group of marsupials, which were

accompanied by a few lineages of placental mammals. These marsupials and placentals appear to be derived from stocks that dispersed southward from North America about the time of the Cretaceous–Tertiary boundary or earlier. A great gap separates the youngest Cretaceous records of mammaliaforms and the oldest Tertiary records of mammals in Africa.

The distal causal factors of the Cretaceous–Tertiary mass extinction are the subject of continuing debate. The pattern of survival and extinction of species of terrestrial vertebrates at the Cretaceous–Tertiary boundary and the role of immigration in their recovery in the earliest Tertiary in the North American Western Interior argues strongly against, if not falsifies, hypotheses calling for global catastrophically destructive changes in the environment as distal causal factors of this mass extinction. Massive volcanic deposits in peninsular India document a period of extensive flood basalt eruptions similar to those that occurred around the Permian–Triassic and Triassic–Jurassic boundaries. These eruptions began before and continued across the Cretaceous–Tertiary boundary. The end of the Cretaceous was a time of major marine regressions linked to changes in continental climates and opening routes of dispersal for terrestrial vertebrates. Discovery of a large crater in Yucatan adds evidence supporting the hypothesis that an asteroid, or other extraterrestrial body, impacted the earth at the close of the Cretaceous. The press–pulse model of extinction (Arens and West, 2008) suggests that extraordinary volcanic eruptions modifying the atmosphere, the consequences of marine regressions, and possibly other factors contributed to changes in global environments stressing Late Cretaceous marine and nonmarine biotas. As a consequence of these stresses, a relatively minor change in the environment might have triggered the mass extinction.

Recovery of Mammalian Faunas During the Early Paleocene

The discovery and collecting biases favoring the fossil record from the Northern Hemisphere continue to limit analyses of the course of mammalian evolution immediately after the Cretaceous–Tertiary extinctions, which marked the end of the second phase of synapsid evolution. Earliest Paleocene faunas are well known only in the North American Western Interior. In northeastern Montana and immediately adjacent areas, a refined fossil record provides the basis for study of the recovery of the mammalian fauna during the first million years of the Paleocene (approximately 65–64 Ma). Radiometric age determinations make it possible to subdivide this period of faunal recovery into a first interval of approximately 400,000 years and a second of approximately 600,000 years (Wilson, 2005).

In the first 400,000 years of the Paleocene, taxonomic diversity of the mammalian fauna of the North American Western Interior was depressed in comparison to that of the latest Cretaceous. The members of a few surviving lineages rapidly increased in abundance, for example, genera of the multituberculates. Only one or two lineages of marsupials survived but their representatives were not abundant. Abundant members of these early faunas were placentals, primarily represented by the paraphyletic condylarths, the so called archaic ungulates. This group is allied to lineages of modern ungulates, hoofed mammals, which would make their appearance later in the Paleocene or in the Eocene. All these earliest Paleocene mammals were characterized by small body size. As far as can be determined from their fragmentary remains, none of them were larger than the largest, latest Cretaceous mammals known from the area, which were about the size of modern opossums or raccoons (Wilson, 2005).

Less than half of the mammalian species in this earliest Paleocene fauna represent lineages present in the preceding latest Cretaceous fauna of the northern Western Interior. The majority is composed of immigrants that appear to have differentiated in other areas and then dispersed into the Western Interior after the Cretaceous–Tertiary mass extinction. This is one line of evidence that suggests the diversification of major placental lineages had begun in other, as yet unsampled, areas prior to the end of the Cretaceous. Support for this view also comes from recent comparative molecular studies of modern placental mammals. Through application of molecular clock techniques, these studies have produced hypotheses suggesting very ancient times, the middle or early part of the Cretaceous, for differentiation of major placental lineages. This conclusion directly contradicts the currently available fossil record, which suggests that the early differentiation of these lineages occurred at most in the latest Cretaceous or was a product of the rapid evolutionary radiation of placentals in the Paleocene (Wible *et al.*, 2007). Probably the history of placental evolution is to be found somewhere between these extremes.

During the next approximately 600,000 years of the Early Paleocene mammalian, particularly placental, diversity in the Western Interior increased greatly. In part, this increase was the result of diversification of lineages present in the area at the beginning of the Paleocene, but immigration of additional groups of placentals and multituberculates from other areas played a significant role. During this interval, body size increased in several lineages. For the first time since the Late Triassic, a few placentals and multituberculates achieved body sizes larger than modern opossums or raccoons and approximating those of modern middle-sized mammals. These trends for increasing taxonomic diversity and appearance of species of larger body size continued throughout the later part of the Paleocene.

Mammalian Evolution During the Tertiary

From the Late Paleocene onward, mammalian faunas of the Northern Hemisphere are increasingly well documented by assemblages of fossils from both Eurasia and North America. Similarly, the available fossil records of Africa, peninsular India, South America, and Australia document the evolution of mammals and other terrestrial vertebrates, particularly later in the Tertiary. After the initial recovery of the mammalian fauna of the Northern Hemisphere from the effects of the Cretaceous–Tertiary mass extinction, taxonomic diversity continued to increase, and many lineages were characterized by the evolution of larger individual body size. Before the end of the Paleocene, some species had attained the bulk of modern rhinos or hippos.

The boundary between the Late Paleocene and Eocene was marked by both a brief period of significant global warming

(the Paleocene–Eocene Thermal Maximum) and significant faunal turnover. In the Northern Hemisphere this was the time of appearance of a number of placental lineages that are represented in modern faunas; for example, it was the time of appearance of the modern orders Perissodactyla (including horses), Artiodactyla (including pigs and deer), and Primates. Additional contributions to mammalian diversity came with the first records of flying mammals (bats) and the appearance of one group of marine mammals (whales). Also, there were instances of decreases in diversity or extinction of major lineages, for example, during the Late Paleocene and Eocene, the diversity of the multituberculates decreased with the appearance and diversification of rodents. In general, however, the taxonomic diversity of terrestrial mammals increased. The sparser fossil records from South America document a radiation of endemic lineages through the Eocene. About the end of the Eocene, immigrant rodents and primates made their appearance. The Eocene faunas of Africa were characterized by both diversification of endemic lineages, for example, the Proboscideans (elephants and their kin) and immigrants from Eurasia (Rose, 2006).

Near the Eocene–Oligocene boundary, approximately 33.7 Ma, many mammalian lineages became extinct. In comparison to other extinctions, this was not an abrupt mass extinction event but an interval of faunal turnover. It occurred during a time of a marked cooling in global climates. Losses were most frequent in groups of placentals thought to have been particularly adapted to widespread tropical forests of the Eocene. In addition to having been cooler, Oligocene climates were characterized by a general increase in aridity and decrease in equability reflected in the expansion of areas of savanna and steppe.

Mammalian diversity in the Northern Hemisphere, measured at the family level, reached a maximum in the Miocene (the Mid-Miocene Climatic Optimum), particularly reflecting a radiation of several lineages of ungulates and rodents. Through the remainder of the Tertiary, the diversity of large ungulates decreased as the climate became cooler and dryer. Although differing in detail, broadly similar patterns in the incidence of extinctions and faunal turnover characterize the evolutionary history of mammalian faunas of the continents of the Southern Hemisphere. Until recently, climatic changes were thought to be the major, if not the only, driving force in mammalian evolution during the later Tertiary. A variety of high-resolution paleontological and geological methods were used to describe and test these hypothesized correlations. The results indicated that times of climatic change and periods of increased diversification or extinction were not precisely correlated. Other environmental or biological factors apparently contributed to controlling the course of later Tertiary mammalian evolution.

Conclusions

The approximately 300-million-year history of synapsid evolution can be subdivided into three major phases. During the first phase, nonmammalian synapsids were among the dominant terrestrial vertebrates. The second phase, which began in the Late Triassic and lasted until the end of the Cretaceous, was characterized by the dominance of the dinosaurs in terrestrial faunas. Mammaliaform synapsids and early mammals were a lesser component of these faunas. The third phase began after the Cretaceous–Tertiary mass extinction and was characterized by a major diversification of mammals returning them to a dominant position.

In addition to the ongoing effects of background extinction, mass extinctions played a significant role in directing the course of synapsid evolution. The Permian–Triassic and, particularly, the Cretaceous–Tertiary mass extinction events greatly reduced synapsid diversity and reset the course of its evolution. Other major turning points in the evolution of the group, which involved extensive extinctions of lineages, were not coincident with mass extinctions – for example, the periods of faunal turnover late in the Triassic or during the transition from the Eocene into the Oligocene.

Attempts to compare patterns and causal factors of pre-Quaternary and Quaternary mammalian extinctions must consider the significant differences in resolution of the timescales at which these events can be studied. In some instances, the fossil record of Quaternary, particularly later Quaternary, extinctions permits analysis of biological change on ecological timescales of hundreds or thousands of years. Resolution of the pre-Quaternary record is much less. The severity of these extinction events may be artificially compounded by our inability to distinguish events that occurred at different times over intervals of hundreds of thousands or a few million years.

Finally, a survey of current research on the distal causal factors of mass extinctions and periods of rapid faunal turnover in synapsid evolution indicates that there probably was not a single grim reaper. Varying combinations of the effects of modification of the configuration of oceans and their patterns of circulation, increases in the intensity of volcanic activity, and the sporadic impacts of extraterrestrial bodies all contributed to sometimes deleterious changes in the physical environment. Similarly, biotic factors played a part. The degree of interdependence of the members of an ecosystem, for example, contributes to the system's resistance to environmental change. At different times in the pre-Quaternary history of mammalian evolution, all these and probably other factors caused extinctions of lineages.

See also: Dinosaurs, Extinction Theories for. Extinction, Causes of. Mammals, Biodiversity of. Mammals (Late Quaternary), Extinctions of. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of

References

Alvarez LW, Alvarez W, Asaro F, and Michel H (1980) Extraterrestrial cause for the Cretaceous–Tertiary extinction. *Science* 208: 1095–1108.

Archibald JD (2011) *Extinction and Radiation: How the Fall of Dinosaurs Led to the Rise of Mammals*. Baltimore: Johns Hopkins University Press.

Arens NC and West ID (2008) Press–pulse: A general theory of mass extinction? *Paleobiology* 34: 455–471.

Barnosky AD, Matzke N, Tomiya S, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57.

Benton MJ and Twitchett RJ (2003) How to kill (almost) all life: The end-Permian extinction event. *Trends in Ecology and Evolution* 18: 358–364.

Dingus L and Rowe T (1998) *The Mistaken Extinction: Dinosaur Evolution and the Origin of Birds*. New York: W. H. Freeman.

Erwin DH (2001) Lessons from the past: Biotic recoveries from mass extinctions. *Proceedings of the National Academy of Sciences* 98: 4503–5399.
Fastovsky DE and Weishampel DB (2005) *Evolution and Extinction of Dinosaurs*, 2nd edn. Cambridge: Cambridge University Press.
Hallam A and Wignall PB (1997) *Mass Extinctions and Their Aftermath*. Oxford: Oxford University Press.
Kemp TS (2005) *The Origin and Evolution of Mammals*. Oxford: Oxford University Press.
Luo Z-X (2007) Transformation and diversification in early mammal evolution. *Nature* 450: 1011–1019.
Luo Z-X, Yuan C-X, Meng Q-J, and Ji Q (2011) A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* 476: 442–445.
Raup DM (1991) *Extinction: Bad Genes or Bad luck?*. New York: Norton.
Rose KD (2006) *The Beginning of the Age of Mammals*. Baltimore: Johns Hopkins University Press.
Wible JR, Rougier GW, Novacek MJ, and Asher RJ (2007) Cretaceous eutherians and Laurasian origin for placental mammals near the K/T boundary. *Nature* 447: 1003–1006.
Wilson GP (2005) Mammalian faunal dynamics during the last 1.8 million years of the Cretaceous in Garfield County, Montana. *Journal of Mammalian Evolution* 12: 53–76.

Mangrove Ecosystems

Peter J Hogarth, University of York, York, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aerenchyma A spongy plant tissue composed largely of air spaces enabling gas exchange to take place by diffusion in underground mangrove roots.

Aerial roots In mangrove species such as *Rhizophora*, Qroots branch out from the stem some distance above the soil surface. Lenticels (pores) in the aerial portion of these roots enable gas exchange to take place, through aerenchyma tissue, with the respiring underground portions of the root.

Mangal A term sometimes used to specify the mangrove habitat as a whole as opposed to "mangrove" applying specifically to the trees themselves. For the most part, however, mangrove is considered to apply to both trees and habitat.

Pneumatophores In some species of mangrove, such as *Avicennia* and *Sonneratia*, underground roots spread laterally from the main stem. Pneumatophores grow vertically from these, typically standing 10–20 cm above the soil surface, enabling gas exchange to take place with the underground roots.

Pseudofecal pellet Fiddler crabs and their relatives collect soil with their mouthparts, separate organic particles from mineral components by a complex flotation process, ingest the former, and discard the latter in the form of compact pellets. These are known as pseudofecal because, although extraction has taken place, the waste material has not passed through the gut.

Mangrove Trees

Currently, the total mangrove area in the world is estimated at 170,000 km^2. They are the principal source of primary productivity in such areas. By their presence, they also provide shelter for other organisms. Mangroves are therefore the energy base, and physical substrate, of an often complex and diverse ecosystem. Mangrove faunas, to a unique extent, comprise organisms of both marine and terrestrial origin (Morrisey *et al.*, 2010; Spalding *et al.*, 1997; Stafford-Deitsch, 1996).

The Mangrove Habitat

The mangrove environment is a demanding one. Typically, mangroves are regularly inundated by tides and are therefore usually in a permanently waterlogged state. The tidal water is saline, so mangrove trees have the problem of coping with salt and acquiring sufficient water against an osmotic gradient. In hot climates, evaporation may make the salinity even greater than that of seawater. In the Indus Delta (Pakistan), for example, the prevailing salinity may be as much as twice that of seawater. Among the vascular plants, only mangroves flourish in such an inhospitable environment (**Figure 1**).



Figure 1 Mangroves (*Avicennia* and *Rhizophora*) fringing a tidal creek in the Indus Delta, Pakistan.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00247-1

Mangroves are defined physiologically as trees that can survive in the mangrove habitat, or mangal. The term is not a taxonomic one, nor does it indicate phylogenetic divergence from a common mangrove ancestor. The approximately 50 species generally recognized as mangroves belong to 20 genera in 16 families, although two families, Avicenniaceae and Rhizophoraceae, dominate in number of species (as they do also in abundance) (**Table 1**). In most cases these genera and families also contain nonmangrove members. Mangrove species have evolved their specialist features as the result of convergent evolution, and mangrove attributes have probably evolved independently at least 15 times.

In addition to true mangrove species, there is also a loosely defined category of mangrove associates. These are species often occurring in mangrove habitats but which also occur elsewhere. Some are found only at the landward margins of the mangal, whereas others, such as creepers and lianes, have their roots above the intertidal zone but invade the mangal by using the mangrove trees purely for support. Other plants associated with mangrove trees are epiphytes, which include ferns and the "ant-house" plants (*see* Insects), and parasitic mangrove mistletoes (Field, 1995; Hogarth, 2007).

Adaptations to the Mangrove Environment

Salinity

Three principal mechanisms enable mangrove trees to survive saline environments. Some species exclude salt at the root surface while continuing to take in water. In *Aegiceras* and *Avicennia*, up to 97% of the salt is excluded, apparently by a physical rather than a metabolic mechanism. This has the effect of locally increasing the salinity of the soil around the roots, with implications for other organisms: mangrove trees modify their environment as well as respond to it. In other instances, trees take in salt but sequester it within cells in such a way that sensitive metabolic processes are protected from contact with excessive salt concentrations. Finally, several mangrove species secrete excess salt, at considerable metabolic cost, from specialized salt glands on their leaves. Many mangrove species use a combination of these mechanisms, as shown in **Table 2**.

Table 1 Distribution of mangrove species by family and genus[a]

Family	*Genus*	*Number of mangrove species*
Avicenniaceae	*Avicennia*	8
Combretaceae	*Laguncularia*	1
	Lumnitzera	2
Palmae	*Nypa*	1
Rhizophoraceae	*Bruguiera*	6
	Ceriops	2
	Kandelia	2
	Rhizophora	8
Sonneratiaceae	*Sonneratia*	5
Bombacaceae	*Camptostemon*	2
Euphorbiaceae	*Excoecaria*	2
Lythraceae	*Pemphis*	1
Meliaceae	*Xylocarpus*	2
Myrsinaceae	*Aegiceras*	2
Myrtaceae	*Osbornia*	1
Pellicieraceae	*Pelliciera*	1
Plumbaginaceae	*Aegialitis*	2
Pteridaceae	*Acrostichum*	3
Rubiaceae	*Scyphiphora*	1
Sterculaceae	*Heritiera*	3
Total		
16	20	55

[a]This follows the classification of Tomlinson (1986); there are althernative views on the status of certain species as true mangroves or mangrove associate species.

Waterlogging

The major problem of waterlogged soils is lack of oxygen. Underground roots, like all tissues, require oxygen for respiration. In a normal soil, gas exchange takes place readily through air-filled spaces between soil particles. In water, the rate of diffusion of oxygen is very low, and in consequence waterlogged soils are generally virtually lacking in free oxygen. One of the most widespread mangrove trees, *Rhizophora*, adapts to such anoxic soils by keeping much of the root mass above the mud surface, surrounded by air. The stretches of these aerial roots (**Figure 2**) close to the soil carry numerous gas-exchange pores, or lenticels, whereas the underground portions are honeycombed with air-filled spaces.

This air-filled tissue, or aerenchyma, is also a feature of *Avicennia* and *Sonneratia*, whose roots are horizontal and close to the surface. These species respire by growing numerous pencil-like pneumatophores, which protrude above the mud surface and allow gas exchange with the underground tissues (**Figure 3**). Pneumatophore growth is facultative: the less waterlogged the soil, the lower the pneumatophore density. In the extreme and atypical case of *Avicennia* growing in sand between the Egyptian Sinai desert and the sea, the soil is so well oxygenated that no pneumatophores develop.

The aerial roots of *Rhizophora* and the intertwined underground horizontal roots of *Avicennia* physically support the trees in what is often a relatively unstable and shifting soil. Aerial roots and pneumatophores provide attachment sites for epibionts and facilitate the accretion of sediment by impeding water movement.

Reproduction

Many mangrove species show some form of vivipary. *Rhizophora* is an example. The ovum is fertilized while still on the

Table 2 Methods of salt tolerance employed by mangrove species

Species	*Exclude*	*Secrete*	*Accumulate*
Acanthus		+	
Aegialitis	+	+	
Aegiceras	+	+	
Avicennia	+	+	+
Bruguiera	+		
Ceriops	+		
Excoecaria	+		
Laguncularia		+	
Osbornia	+		+
Rhizophora	+		+
Sonneratia	+	+	+
Xylocarpus			+

parent tree and grows by a combination of photosynthesis and acquisition of nutrients from the parent until it may reach a length of 50 cm (**Figure 4**). This structure – neither a seed nor a fruit, and hence usually termed a propagule – then falls to the ground. The propagules of some species root almost immediately, but others appear to have an obligatory floating period before they sink and establish themselves. The majority of floating propagules probably settle close to the parent, but long-distance dispersal is also possible. Floating mangrove propagules may remain viable for a month or longer: Depending on current speed and direction, they could travel a considerable distance. It is not uncommon for mangrove seedlings from Mexico, for instance, to be stranded and take root in Texas virtually across the length of the Gulf of Mexico. An even greater dispersal may explain the mangrove species *Rhizophora samoensis*, which is found only in Samoa and adjacent islands, at the opposite extremity of the Pacific from its presumed ancestor, the species *Rhizophora mangle* of Central America. The significance of dispersal ability for the geographical distribution of mangrove species is discussed in section Regional Patterns of Diversity.

Mangrove Animals: Fauna of Terrestrial Origin

Although mangrove roots are periodically immersed, the branches and leaves provide an environment little different from that in adjacent terrestrial forests, with which they consequently share much of their fauna. Mangrove animals of terrestrial, rather than marine, origin include arthropods (particularly insects, but also spiders and myriapods),



Figure 2 Aerial roots of *Rhizophora* in a Malaysian mangrove forest.



Figure 4 Mangrove propagules on a *Rhizophora* tree, Indus Delta, Pakistan.



Figure 3 Mangroves (*Avicennia marina*) on the Red Sea coast of the Sinai Peninsula, Egypt, showing dense pneumatophores.

amphibians, reptiles, birds, and mammals. Virtually none are found exclusively in mangroves.

Insects

Anyone who has worked in mangroves can testify to the abundance of biting insects, particularly mosquitoes and "sand flies" or biting midges (Ceratopogonidae). Mosquito larvae develop in rot holes in mangrove trees, in semi-permanent brackish pools, or in the water retained in crab burrows. In the latter case, one East African species, *Aëdes pembaensis*, ensures a suitable burrow environment for its larvae by laying its eggs directly onto the claws of the crab *Sesarma meinerti*. Prey of adult mosquitoes includes, apart from humans, a variety of mangrove mammals and birds, and in some cases it extends to fish.

Ants are often abundant in mangroves, including the aggressive nest-building weaver ants (*Oecophylla*) of the Indo-Pacific and leaf-cutter ants (*Atta*) of South America. Particularly complex relationships have evolved between ants, epiphytic "ant-house" plants, and mangrove trees (**Figure 5**). Ant-house plants have bulbous stems (which may weigh several kilograms) honeycombed with passages inhabited by ants. One such plant, *Hydnophytum formicarium*, has specialized chambers in which ants deposit the remains of their prey, and from which the plant can absorb nutrients released by fungal action. The situation is further complicated by the presence of butterfly larvae (*Hypochrysops*) which feed on the ant-house plant and which are tended by the ants. The relationship therefore involves interactions between two plant species, two animal species, and one or more fungus.

Most mangrove ants are arboreal and essentially terrestrial animals. In many cases they nest outside the intertidal zone and forage in the mangal only at low tide. One Australian species, *Polyrhachis sokolova*, is truly intertidal, retreating at high tide to nests within the mangrove mud. Nothing is actually known of its physiology: like other intertidal insects, it may retain a surface film of air and therefore avoid the need for any special adaptations to immersion or varying salinity. It does have some ability to swim.

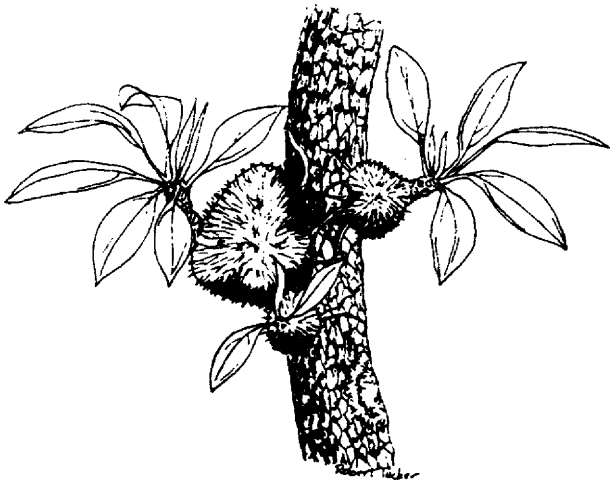


Figure 5 The "ant-house plant" *Myrmecodia*, epiphytic on a mangrove branch.

Probably of greater ecological significance are the various plant-eating insects. Termites play a major role in disposing of dead wood. Some species construct nests of mud on tree trunks several meters above high-tide level, with access galleries snaking down the trunk to the aerial roots and upward to the canopy.

The most important herbivores are those that eat mangrove leaves and seedlings, particularly the larvae of moths and beetles. Typically, only a small proportion of leaf production falls to herbivory. Sometimes, however, it reaches epidemic proportions. Individual trees in an otherwise healthy forest may be completely defoliated, and occasionally areas of many hectares are stripped of leaves. Canopy loss may result in the defoliated trees dying and being replaced by other species that are more tolerant of unshaded conditions. Insect herbivory therefore may alter mangrove community structure.

Other mangrove insects include the spectacular synchronously flashing fireflies of Malaysia (*Pteroptyx*), which occupy the mangrove *Sonneratia* for their displays, and numerous species of butterfly and moth. Hawkmoths, bees, and drosophilidid flies are among the species which are probably of importance in pollinating mangrove flowers.

Amphibia and Reptiles

Amphibia are rare in brackish or salt water, but one species, the crab-eating frog (*Rana cancrivora*), is common in mangrove habitats of Southeast Asia. Tadpoles survive well in salinity up to 50% that of seawater.

Reptiles are more abundant. Numerous species of snake forage within the mangal at low tide, including terrestrial or arboreal species but also some for which the mangal is their primary habitat. Mangrove snakes eat crabs (sometimes reciprocated), insects, and fish. In Southeast Asia, one of the most formidable mangrove predators is the monitor lizard (*Varanus indicus*), which may reach 1 m in length. Crocodiles, caimans, and alligators also occur in mangroves, although these are now rare in many areas due to human activities.

Birds

Birds are highly mobile. Many spend only part of their time in mangroves, migrating seasonally, daily, or tidally. Mangroves provide a feeding area, a nesting site, a refuge from the rising tide, or some combination of these. Waders probe for invertebrates in the mud of the mangal or adjacent mudflats. Kingfishers, egrets, and herons catch fish or invertebrates in the shallow water of mangrove creeks. Larger fish eaters, such as pelicans, ospreys, and cormorants, range further afield and may return to the mangal to roost or breed. In the Caribbean, roosts and nesting colonies of cattle egrets (*Bubulculus ibis*) and scarlet ibis (*Eudocinus ruber*) are so densely packed that the consequent enrichment of the soil with guano leads to significantly enhanced local growth of the mangrove trees.

Mangrove forests typically include numerous passerine species. Nectar feeders such as sunbirds in Malaysia, honeyeaters in Australia, and hummingbirds in South America move

seasonally into mangroves and may be important pollinators. Insectivorous passerines specialize in hawking for insects in the canopy or, among low-lying vegetation, in picking insects off leaves or from bark crevices or from different species of tree. Broadly similar guilds of insectivorous birds, comprising different constituent species, seem to occur in different geographical regions.

Few of the species found within the mangal are mangrove specialists, and those which are restricted to mangroves in one part of the world may occupy different habitats elsewhere. One example is the cosmopolitan Great tit (*Parus major*), distributed from western Europe to China: only in Malaysia is it a mangrove species. The lack of mangrove specialists is probably due to the relative simplicity of the mangrove forest structure compared with typical tropical forest, allowing less scope for niche specialization. Another reason is probably the proximity of a pool of competing species in adjacent tropical rain forest. There are proportionally fewer mangrove specialists in New Guinea, where rain forest usually abuts mangrove habitats, than in Australia, where this juxtaposition is less common. Within Australia, there are fewer specialists in the mangroves of Queensland, which are extensive and contiguous with rain forest, than in northwestern Australia, where this is not the case. Most mangrove birds are probably using the habitat opportunistically.

Mammals

As with birds, many mammal species use the mangal opportunistically. These include small rodents, agoutis, wild pigs, antelopes, deer, and rhinoceroses; the Sundarbans of Bengal are the last major redoubt of the Bengal tiger (*Panthera tigris*). Domestic animals, such as camels and buffalo, are often a major element in the mangrove fauna. Otters may also be abundant, feeding on fish and crabs from the mangrove creeks.

Monkeys are common in mangroves. In Southeast Asia these include macaques (*Macaca*) which forage on the mud for crabs and mollusks. They also uproot large numbers of mangrove seedlings: Because these are seldom eaten or even greatly damaged, the purpose is not clear. Herbivorous monkeys are found in the forest canopy, including leaf monkeys (*Presbytis*) and, in the mangrove forests of Sarawak, the striking proboscis monkey (*Nasalis larvatus*). This is found only in mangroves and riverine forests, and it specializes in eating foliage, which is digested in an elaborate multichambered stomach with the aid of resident bacteria.

Bats are often abundant in mangroves. Resource partitioning in insectivorous bats parallels that of insectivorous birds, with species specializing in different zones of the mangrove vegetation and catching their prey with different flight techniques. A single bat may eat up to one-third of its body weight of insects each night: a 30-g bat might therefore consume 5000 insects nightly. The impact on the insect population of foraging bats must be considerable.

The exclusively Old World fruit bats often occur in mangrove forests in vast numbers: Roosts of an estimated 220,000 individuals have been recorded. Most fruit bats feed on nectar and fruit, and it is this which attracts many species into the mangal. In Malaysia, the long-tongued fruit bat *Macroglossus minimus* is an important pollinator of the mangrove *Sonneratia*: the long tongue is specialized for insertion into the *Sonneratia* flower, which carries large projecting stamens to deposit pollen onto the fur of the feeding bat. *Sonneratia* flowers last for only a single night, possibly because of the wear and tear resulting from visits by such a large pollinator. This species of bat is a true mangrove specialist, and in western Malaysia at least, it has not been recorded from other habitats. Mangrove specialization is possible only because the three species of *Sonneratia* in the area have different flowering patterns so that nectar is available throughout the year. Other fruit bats switch seasonally between mangrove and nonmangrove species.

Fauna of Marine Origin

One of the principal reasons for the high faunal diversity of mangrove ecosystems is their accessibility to occupation by organisms from both terrestrial and marine habitats. Of these, the marine invaders are the more numerous in terms of numbers and diversity of species. These include more or less sessile organisms settling on aerial roots and pneumatophores as well as more mobile species living on and under the mud. Many animal groups are represented in the mangal, the most conspicuous and ecologically most significant being teleost fish, crustacea, and mollusks. As with the land-derived mangrove fauna, the majority of species occur elsewhere and accumulate in mangroves because of the availability of food, shelter, or suitable substrate.

Considering mangrove communities at a scale of, for example, hectares, the diversity of such animals is often high. At smaller scale, however, the anoxic conditions caused by waterlogging, exacerbated by microbial decomposition of detritus, may greatly reduce both species diversity and abundance.

Root Communities

Mangrove roots and pneumatophores provide a hard substrate often covered with a rich and diverse growth of sponges, sea anemones, bryozoans, tunicates, barnacles, tubeworms, and mollusks as well as epiphytic algae. These in turn may attract a more mobile population of browsers or predators. The epibionts are mostly filter feeders, extracting organic particles suspended in the water, or predators of zooplankton, with no direct interaction with their mangrove host. A particularly thick growth, however, can adversely affect the host tree by occluding lenticels and restricting gas exchange with the underground roots. The relationship is sometimes mutually beneficial, as encrusting sponges may transfer nitrogenous nutrients to their host, and encrusting fauna can protect the root from attack by wood borers.

The labyrinthine aerenchyma tissue of the roots is easily penetrated by wood-burrowing organisms. The isopod crustacean *Sphaeroma* is a common root borer and may cause severe damage and even death. *Sphaeroma*-induced damage near the growing tip of a root may induce forking, with a resulting increase in the number of roots entering the soil: This may benefit the tree. The "shipworm" *Teredo* (which is in fact a mollusk)

also bores dead roots and trunks extensively and plays a similar role to that of termites in disposing of woody debris. Like termites, *Teredo* relies on symbiotic microorganisms to digest the more intransigent components of wood.

Fish

Mangrove creeks and inlets are frequently occupied by abundant and diverse fish populations. In Southeast Asia, for instance, records of more than 100 species are by no means unusual. Many of these species spend only part of their time within the mangal, often moving to other habitats seasonally or at different stages of their life cycle. Mullets (*Liza*) eat significant amounts of mangrove detritus, such as shed leaves: most hunt small Crustacea or other invertebrates. Some fish are permanent creek residents, commuting into the forest when it is submerged at high tide and foraging among the mangrove roots.

At low tide, Asian mangroves are occupied by mudskippers, which are relatives of the gobies (**Figure 6**). As their name suggests, they skip across the exposed mud surface using their tails and leg-like pectoral fins, sometimes even climbing up aerial roots or pneumatophores. This amphibious life requires appropriate physiological adaptations, particularly in relation to respiration. Mudskippers are largely air-breathing, with gas exchange taking place not just across the gills but also at highly vascularized areas of the skin. Some store air within their burrows to enable aerial respiration even at high tide.

All mudskippers are probably to some extent omnivorous, but some are predominantly deposit feeders and others carnivores. Prey of the latter include crabs, insects, spiders, shrimps, and snails.

Crustacea

Mangrove habitats, particularly in the Indo-West Pacific, are dominated by crabs belonging to two families, Grapsidae and Ocypodidae. The former are predominantly herbivores or detritus feeders and the latter deposit feeders, extracting fine organic particles from mangrove mud. Predatory crabs, such as the formidable *Scylla*, may also be important components of the mangrove fauna. Shrimps (Penaeoidea) and mud lobsters (*Thalassina anomala*), and smaller Crustacea such as amphipods and isopods, may also be significant as scavengers, in breaking down leaf litter, or as predators of smaller organisms.

Grapsid Crabs

Grapsid crabs of the subfamily Sesarminae, particularly of the genus *Sesarma*, are characteristic of mangroves, although a few species of this genus occur in other habitats (**Figure 7**). More than 40 species of sesarmine have been reported from the mangroves of Malaysia alone, and many species, here and in other regions, undoubtedly remain to be described.

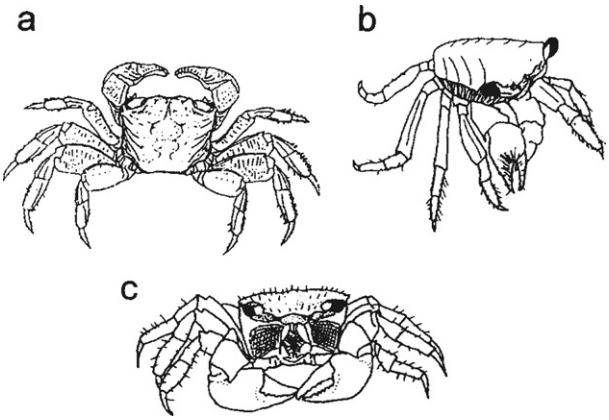


Figure 7 Mangrove sesarmine crabs: (a) *Parasesarma plicata*, (b) *Aratus pisonii*, and (c) *Neosarmatium smithi*. Reprinted from Jones DA (1984) Crabs of the mangal ecosystem. In: Por FD and Dor I (eds.) *Hydrobiology of the Mangal*, pp. 89–109, with permissions from Kluwer Academic Publishers and the author.



Figure 6 Mudskipper on an *Avicennia* pneumatophore.

Sesarma are small (usually less than 3 cm in breadth) and inconspicuously colored. They are amphibious, retreating into burrows at high tide and foraging on the exposed mud at low tide. Respiration in air is achieved partly by recirculating water from the gill chambers over the carapace where it can be reoxygenated: evaporative cooling during this process also serves to reduce the dangers of high air temperature. Water loss can be offset by the acquisition of soil water through tufts of root-like hairs. Sesarmines are euryhaline, although differing degrees of salt tolerance probably contribute to the zonation of crab species along estuaries or with shore level.

In some cases, sesarmine crabs climb trees to feed on fresh leaves or buds. In East Africa, *Sesarma leptosoma* undertakes synchronized mass migrations twice daily from refuges among mangrove roots to forage on the tips of the branches of the trees. The virtually indistinguishable Caribbean species *Aratus pisonii* spends most of its time in trees, only rarely descending onto the mud.

Most sesarmines, however, subsist on fallen leaves or propagules. Mangrove leaves are often rich in tannins and other aversive materials, and several crab species have been shown to select leaves from the more palatable species of tree. Many leaves are collected as soon as they fall and cached in crab burrows. As decomposition proceeds, tannin levels decrease and nitrogen content increases through the accumulation of microbial biomass: storage therefore increases leaf palatability.

Much of the leaf material eaten is not assimilated but redeposited onto the mud as feces, available for microbial decomposition. It has been estimated that processing of leaf material by crabs increases the rate of breakdown of leaf litter 75-fold compared with the rate of decomposition under microbial action alone. Therefore, sesarmine crabs collectively play a very important role in facilitating energy flow through the mangrove ecosystem. By eating propagules, they also affect species distribution and community structure of mangrove trees (*see* Tree Distribution). However, there are geographical differences: in Southeast Asia and Australia, sesarmines are crucial in litter breakdown and selective removal of propagules, whereas in Florida and the Caribbean they are of lesser significance.

Ocypodid Crabs

Some crabs of the family Ocypodidae, such as the Central American hairy land crab *Ucides*, consume mangrove detritus. The majority are deposit feeders. Among these, the most conspicuous are the gaudily colored fiddler crabs (*Uca* spp.), widespread throughout the mangroves of the Old and New World (**Figure 8**).

The common name derives from the one greatly enlarged claw of male fiddlers, which is used in courtship and in deterring rival males. The smaller claw of males and both claws of females are devoted to feeding. Mud is scraped into the buccal cavity in which, by a complicated process of flotation and manipulation by the mouthparts, fine organic particles are separated from the mineral components. The former is ingested and the latter deposited as a ball of sand, or "pseudofecal pellet." The process of separation may be quite selective. In some species, what is extracted consists almost entirely of microbial cells rather than, for example, fragmented leaf material. Others have subtly different extraction techniques and may specialize in the smaller meiofaunal animals. There may be as many as 60 fiddler crabs per square meter, resulting in 500 g of soil being processed daily. The toll on meiofauna is probably considerable, and the effects on soil texture and composition are profound.



Figure 8 Fiddler crab (*Uca*) in a Mozambique mangrove (photograph courtesy of D. Barnes).

Other Mangrove Crustacea

Other crabs found in mangroves are important predators. The most conspicuous is the mud crab *Scylla serrata* of the family of swimming crabs (Portunidae). *Scylla* reaches a carapace width of up to 20 cm, making it the largest invertebrate predator found in mangroves. Equally formidable predators are the mantis shrimps (Stomatopoda), which live in burrows in the mud and lacerate prey by rapidly shooting out their viciously spiked raptorial appendages. Other rarely seen burrowing crustaceans include pistol or snapping shrimps (*Alpheus* spp.) and the mud lobster *Thalassina* (*see* Crustacea as Ecosystem Engineers).

More general mangrove scavengers include hermit crabs, particularly *Clibanarius*, which forage on the mud surface at high tide. Shrimps may also be abundant in mangroves and mangrove creeks. Penaeid shrimps, which in at least some parts of the world depend heavily on mangroves for feeding and breeding, are an important commercial crop. The shrimp *Merguia* apparently lives only in mangroves and has the distinction of being the only semiterrestrial shrimp: it actually climbs trees. Only two species are known. One occurs in the Indo-West Pacific region, from Kenya to Indonesia, and the other occurs in Panama, Brazil, and Nigeria. Indo-West Pacific and Atlantic regions differ in the composition of their mangrove floras, and the separation of the two species of mangrove-associated shrimps may have occurred in parallel with the divergence of the mangroves themselves.

Crustacea as Ecosystem Engineers

All species have an impact on their environment, at the very least exchanging materials in the form of food, waste

materials, and respiratory gases. Some species have effects beyond these simple transactions and alter the nature of their environment in ways that affect species other than their direct competitors, predators, or prey. Such species are often termed "ecosystem engineers."

In a mangrove ecosystem, the trees are the greatest engineers, influencing sedimentation rates and creating a physical environment. Crustacea also, in important ways, transform their surroundings. The topography of mangrove swamps in Southeast Asia is often visibly modified by mud lobsters (see section Other Mangrove Crustacea). While processing mud, *Thalassina* throws up waste material from beneath the surface, forming mounds which may reach 2 m in height. These create patches of dry mud which provide habitats for other species, including the mangrove fern *Acrostichum*, fiddler crabs, and a variety of other burrowing Crustacea and mollusks. Between the mounds the mud surface is lower, and more waterlogged, than it would be otherwise. Burrowing crabs also contour their environment, although less dramatically.

Much of the microbial activity of mangrove mud occurs in the surface layer, to a depth limited by the diffusion and exchange of gases with the atmosphere. As fiddler crabs process surface mud, they continually expose fresh material, facilitating microbial activity, while the active surface of the mud is increased in area by crab burrows. Burrowing activity also oxygenates the deeper soil and creates an underground labyrinth of interconnecting passages, through which significant underground water flow occurs. Experimental evidence suggests that crab activities significantly affect nutrient recycling and enhance growth of mangrove trees. Crustacea therefore alter the state of their environment in ways that significantly affect other species.

Mollusks

Bivalves

The most visible bivalve mollusks of mangroves are the oysters and mussels found attached to roots. Within the mud, however, there is often an abundant population of burrowing species. These, like the oysters and mussels, are largely filter feeders, extracting fine organic particles from suspension. A less typical group of bivalves are the shipworms of the family Teredinidae (*see* Root Communities), including the giant mangrove shipworm *Dicyathifer*, which may reach 2 m in length.

Snails

Gastropod snails are also generally abundant in mangroves. As with the crustacean fauna, these include herbivores, detritus and deposit feeders, and predators. Although a few species are uniquely found in mangroves, the majority of surface-living species also occur on open mudflats.

The principal predatory snails are species of *Thais*, found in mangroves worldwide. These cruise over mud and mangrove roots, feeding on barnacles or smaller gastropods. In the mangroves of Costa Rica, for example, *Thais kiosquiformis* densities may reach more than 200 per square meter, and the species plays a major role in maintaining the function of mangroves by removing encrusting fauna from their roots.

Many gastropod species are deposit feeders, ranging in size from tiny and almost invisible species to the massive *Terebralia* and *Telescopium* of the Indo-Pacific region, which may reach a length of 10 cm (**Figure 9**). One species, *Terebralia palustris*, feeds on small detritus particles when young, but on reaching a length of approximately 3 cm it switches to a diet of fallen leaves. The teeth on the radula (the ribbon-like tongue) of gastropods metamorphose appropriately to a form suitable for the altered diet. In Florida, snails are important consumers of mangrove seedlings, at some locations destroying nearly three-fourths of the seedling population. This is an interesting geographical contrast with other regions, such as Malaysia and Australia, where crabs fulfil this role (*see* Grapsid Crabs).

The most abundant snails on the mangrove trees are often species of *Littoraria*, close relatives of the periwinkles of temperate rocky shores. In Central America, on both sides of the Isthmus of Panama, the common species is *Littoraria angulifera*. In the Indo-Pacific, this species is replaced by many others, which partition between them the slightly different habitats afforded by a tree. In Papua New Guinea, *Littoraria scabra* prefers the bark of trees on the seaward side of a forest,

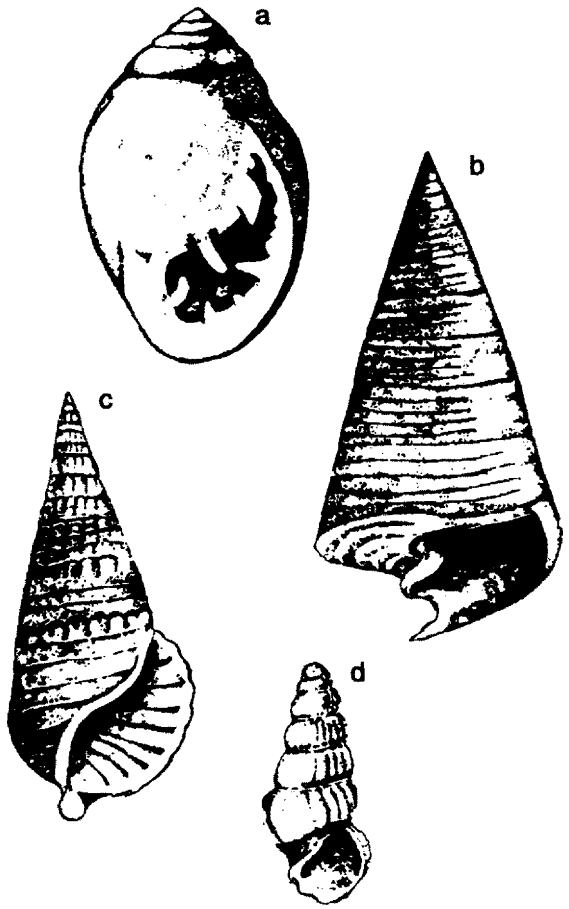


Figure 9 Gastropod mollusks typical of Indo-West Pacific mangroves: (a) *Pythia* (2.5 cm), (b) *Telescopium* (10 cm), (c) *Terebralia* (6 cm), and (d) *Cerithidea* (2 cm). Reprinted from *Advances in Marine Biology* 6, Macnae W, A general account of the fauna and flora of mangrove swamps and forests in the Indo-West Pacific Region, 74–270, 1968, with permission from Academic Press.

Littoraria intermedia prefers trees next to freshwater creeks, whereas the polymorphic species *Littoraria pallescens* is found solely on leaves.

Meiofauna

Within the mangrove mud lies a rich fauna virtually invisible to the naked eye – the meiofauna. Beneath an area of 10 cm^2 of mud there may be many thousands of individuals. Orders of magnitude smaller than the more conspicuous macrofaunal crabs and snails are meiofaunal herbivores, detritivores, and formidable predators, with food chains probably dependent on photosynthetic cyanobacteria ("blue-green algae") and heterotrophic bacteria. Meiofauna colonize fallen leaves, and the stages of leaf breakdown are accompanied by complex interactions and successional shifts in species composition and community structure which parallel, on a microscopic scale, the processes of macroecology.

The numbers of meiofaunal individuals are immense, and their diversity is astonishingly high. Not only are there many species but also the species show a higher level of taxonomic diversity. Among the macrofauna virtually all species belong to just three phyla: arthropods, mollusks, and chordates. The meiofauna from just one mangrove area in Australia, for example, yields turbellarian flatworms, nematodes, copepods, Ciliophora, Foraminifera, bivalve mollusks, oligochaete and polychaete annelids, hydrozoa, archiannelids, kinor-hynchs, tardigrades, and gastrotrichs.

Very little is understood about the meiofauna of mangroves, their interactions, their functional significance in the ecosystem as a whole, and the relationship between the meiofaunal and macrofaunal worlds. Their small size belies their great importance.

Connections

The salient features of typical mangrove ecosystems are relatively high rates of primary productivity, much of the results of which enter decomposition pathways, either directly or after initial breakdown by leaf-eating crabs or mollusks. This is true of leaves and reproductive structures and, on a more protracted timescale, of the woody components of the trees. Particulate organic matter, either small leaf fragments or bacterial cells, is ingested by molluskan and crustacean deposit and filter feeders, enters meiofaunal food chains, or accumulates in the mud.

The ecosystem can be viewed physically as well as in terms of the flow of energy or matter. Mangrove trees supply hard surfaces on which other organisms settle, and they modify (as well as respond to) the physical environment by stabilizing the soil, facilitating accretion of mud, and retarding erosion. The environment is further modified by the physical activities of burrowing Crustacea and other animals.

Mangrove ecosystems cannot be considered in isolation. They interact with adjacent habitats through the trapping of exogenous sediment or export of particulate or soluble organic matter or inorganic nutrients. Animals, by moving between mangroves and other habitats, also contribute to import and export of matter. Commercially important penaeid shrimps use mangroves as nursery areas so that shrimp catches many miles away may depend critically on mangrove productivity. Hard evidence for such connections between mangroves and other ecosystems, however, is sometimes elusive, and the strength of such linkages is almost impossible to quantify.

Mangrove diversity

Mangrove diversity must be considered at a range of spatial scales, from global patterns of species richness to the pattern of distribution, at a particular location, at a scale of a few meters. In considering mangrove fauna, even smaller spatial scales become relevant. At all scales, diversity is affected by the past history of the area, by physical factors, and by biotic interactions, but the importance of each of these and the timescales over which they operate vary with scale (Duke *et al.*, 1998).

Global Patterns

Latitudinal Range and Species Diversity

Mangroves are almost exclusively tropical or subtropical. This distribution is a reflection of a temperature limitation: The global distribution of mangroves correlates very closely with, for example, the winter position of the 20 °C isotherm (**Figure 10**). The number of mangrove species declines with increasing latitude, with the most northerly and southerly mangroves being species of *Avicennia*. In temperate regions, mangroves are replaced by salt marsh vegetation: plants which, like mangroves, are adapted to conditions of salinity and waterlogging but which do not carry the additional burden of being a tree or of producing large propagules.

Longitudinal Differences

Within their temperature and latitudinal constraints, mangroves show interesting patterns of species distribution. The principal biogeographic division is between the Indo-West Pacific (IWP) and Atlantic–Caribbean–east Pacific (ACEP) regions. These two regions have broadly similar areas of mangrove habitat, but the IWP has four times more genera and six times as many species of mangrove: 17 genera compared to 4, and 40 species compared to 7. It is apparent that none of the mangrove genera are very diverse, possibly because of a general limitation on species diversification in harsh intertidal conditions. Genera occurring in the IWP, however, are slightly more speciose than those of the ACEP: 2.35 compared with 1.75 species per genus.

The differences between the IWP and ACEP regions are maintained by major barriers. The most obvious of these is the African continent (**Figure 10**). Less obvious is the barrier represented by the central Pacific. This results principally from dispersal limitations rather than from the absence of suitable habitat. Suitable environments are present on many Pacific islands without natural mangrove populations, as shown by the success of the artificial introduction of mangrove species to Hawaii.

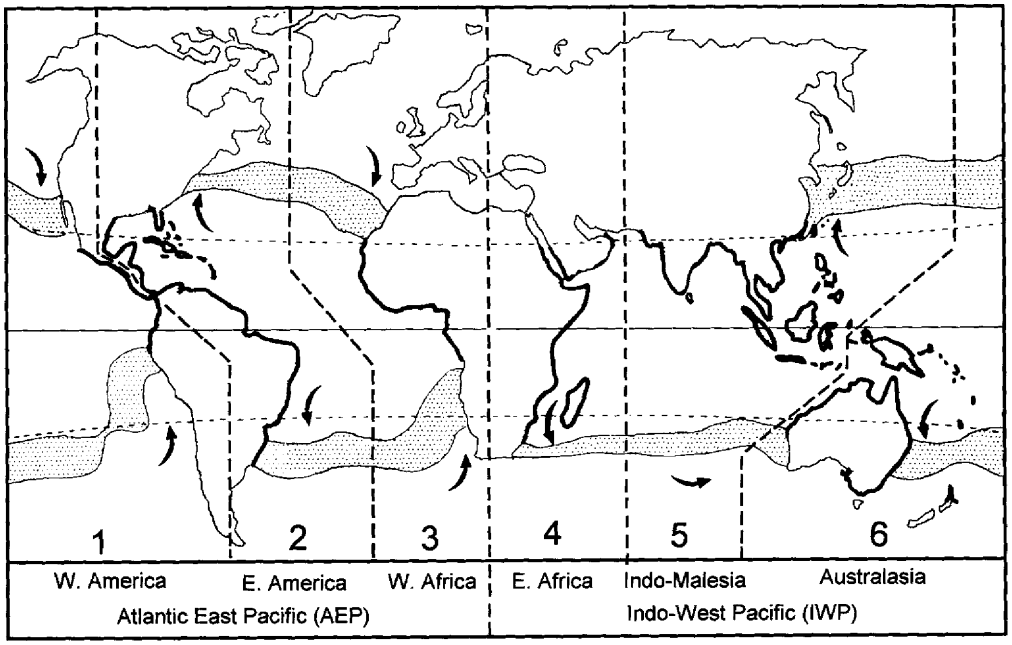


Figure 10 World distribution of mangroves in relation to 20 °C isotherms. Reprinted from Duke NC (1992) Mangrove floristics and biogeography. In: Robertson AI and Alongi DM (eds.) *Tropical Mangrove Ecosystems*, pp. 63–100, with permission from American Geophysical Union and the author.

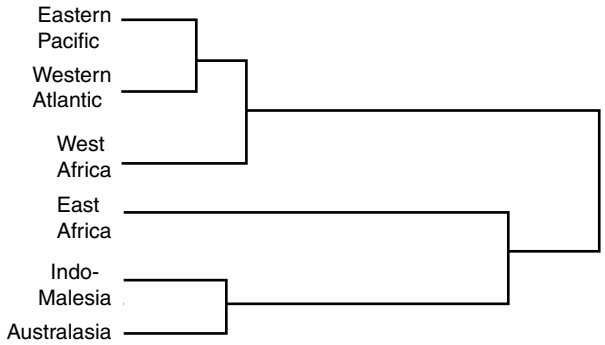


Figure 11 Comparison of the mangrove flora in six geographical subregions. Because of the recent closure of the Isthmus of Panama, the eastern Pacific and western Atlantic (including Caribbean) are most similar in species composition. Note also the separation between Atlantic–Caribbean–eastern Pacific (ACEP) and Indo-West Pacific (IWP) regions.

Further dispersal barriers, including the Isthmus of Panama, open ocean, and arid coasts unsuitable for mangrove occupation, divide the major regions into smaller subregions, each with a more or less distinctive mangrove flora (**Figure 11**). Only one species occurs in all six subregions: the mangrove fern *Acrostichum aureum*. Two genera, *Avicennia* and *Rhizophora*, are common to both IWP and ACEP regions. All other genera are found exclusively in either the IWP or the ACEP, although the close similarity between *Laguncularia* (ACEP) and *Lumnitzera* (IWP) suggests a recent separation of these two genera.

The traditional explanation of mangrove species distribution is of a center of origin and of diversification in Southeast Asia, followed by dispersal restricted by physical barriers. This clearly makes little sense in relation to the current dispersal barriers. Fossil evidence of mangroves is widespread and reveals a much wider distribution during the Eocene and earlier epochs: Fossil *Nypa*, *Avicennia*, and *Rhizophora* pollen and other remains, for instance, have been identified in Eocene and Miocene deposits that now form part of North and South America, Europe, and North Africa as well as south and east Asia. At the time, these locations were connected by the Tethys Sea, continuous through what is now the Mediterranean and Indian Ocean.

Subsequently, this pantropical distribution was partitioned as a consequence of continental movements. Cosmopolitan genera such as *Avicennia* and *Rhizophora* were separated into regional populations by the approach of Africa to Asia 30–35 million years ago which closed the Tethys Sea, and separation of the sister genera *Laguncularia* and *Lumnitzera* followed the widening of the Atlantic barrier. The emergence of modern species ensued within the isolated subregions. Closure of the Isthmus of Panama was geologically very recent (a mere 2 or 3 million years ago) so that differences between eastern Pacific and Caribbean species are slight. One species (*Pelliciera rhizophorae*) is found on both sides of the Isthmus, presumably reflecting a separation into two populations too recently for allopatric speciation to have occurred.

An originally pantropical mangrove distribution was therefore partitioned into regions and subregions, with subsequent evolutionary divergence. Climatic conditions then eliminated mangrove species from areas such as southern Europe and the Mediterranean fringes. The current distribution pattern results from a combination of large-scale geographical factors and more regional climatic ones.

Diversity of Mangrove Fauna

It might be expected that faunal species diversity would follow a similar pattern to that of mangrove tree diversity both because the mangrove fauna has presumably been exposed to the same influences and because of a presumption that faunal diversity should respond to tree diversity.

The IWP region, richer in plant diversity than the ACEP, is also richer in species of mangrove-associated Crustacea and mollusks (**Table 3**). The reverse is true of other taxonomic groups, particularly those that form constituents of the root communities, such as sponges, coelenterates, and echinoderms. This may reflect regional differences in tidal range and availability of roots for settlement. For many groups, unfortunately, little comparable data are available and recorded species numbers reflect taxonomic interest and effort rather than the composition of actual species assemblages.

Regional Patterns of Diversity

Species diversity varies within regions in response to many different factors. The ACEP region, in addition to having fewer mangrove species in total, shows less differentiation between localities within the region, and all the species available in the geographical vicinity are likely to be represented at most locations.

Various factors may result in local variation in species diversity. Mangroves do not grow on rocky shores or in areas where fresh water is completely lacking (which is in part why all tropical shores are not dominated by mangroves). Stretches of inhospitable coastline therefore act as barriers which affect mangrove dispersal and geographical distribution. The arid shores of Somalia, for example, result in the reduction in species number northwards so that *Avicennia marina* is virtually the only mangrove species found in the Red Sea. Separation of mangrove estuaries from each other by arid coastline, and regional-scale variation in physical variables, also affects the species distribution of mangroves around the Australian coasts.

Dispersal ability also affects species distributions within regions. The number of mangrove species on islands of the western Pacific shows clear attenuation with increasing distance from the species-rich areas of Australia and Papua New Guinea. Similarly, among islands off the West African coast there is a clear relationship between the number of mangrove species present and the distance from the nearest landward neighbor (**Figure 12**). Species number also correlates with island size, with larger islands containing more species.

Local Variation in Species Distribution and Diversity

Tree Distribution

At a specific location, the distribution of mangrove species responds to physical variables in the environment. These often vary as gradients: in an estuarine mangal, for instance, salinity and the influence of tidal fluctuations tend to diminish with distance up the river. Sediment composition and nutrient dynamics also alter with distance from the open sea. Mangrove species respond differentially to such upriver/downriver gradients, resulting in zonation of species.

Similarly, in areas dominated by tide rather than river flow, tidal fluctuations establish gradients of physical variables, particularly in salinity and the extent of waterlogging of the soil. Again, mangrove species respond differentially to these physical variables and tend to form distinct zones. Where both river and tidal influences interact, the pattern of species distribution can be extremely complex.

In relation to salinity, species generally grow better at low salinity and differ more in the tolerance range than in their salinity optima. Low salinity, in consequence, tends to be associated with higher species diversity. At higher salinities, tolerance differences result in differing competitive success and translate into zonation of mangrove species along a salinity gradient, with species dominating zones at which they compete best, rather than those corresponding to salinity growth optima.

Although response to physical gradients suggests a gradual transition from one species to another as the determining physical variable gradually alters, this is often not the case. Mangrove species are frequently found in virtually monospecific stands or zones, with a more or less abrupt transition from one dominant species to another. This suggests that interactions between tree species, and mutual exclusion, may play a part in defining zone boundaries. Other physical

Table 3 Number of species recorded from mangroves in various localities in the regions indicated

Taxonomic group	*Atlantic–Caribbean–East Pacific*	*Indo-West Pacific*		
	Caribbean/Western Atlantic	*East Africa*	*Indo-Malesia*	*Australia*
Sponges/bryozoa	36	1	5	7
Coelenterata/Ctenophora	42	12	3	6
Nonpolychaete worms	13	3	13	74
Polychaetes	33	72	11	35
Crustacea	87	163	229	128
Mollusks	124	117	211	145
Echinoderms	29	23	1	10
Ascidians	30	13		8
Fish	212	114	283	156
Reptiles	3		22	3
Birds	138		177	244
Mammals	5		36	7

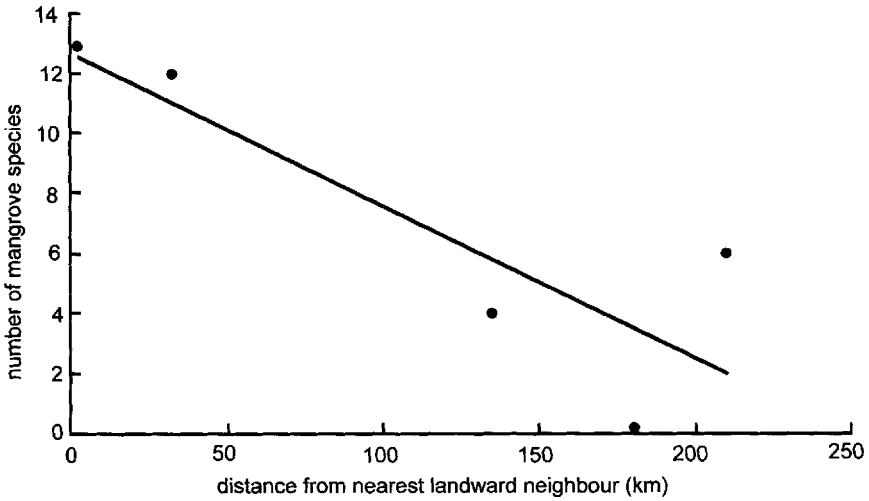


Figure 12 The number of species of mangrove occurring on West African islands in relation to their distance from the nearest landward neighbor.

variables, such as the degree of waterlogging and soil anoxia, nutrient availability, and biotic interactions between species, similarly affect species distribution within the mangal.

Superimposed on the sorting of species under the influence of physical variables are variations resulting from interactions with the mangrove fauna. Of the faunal influences, the most significant is the selective destruction of mangrove propagules by sesarmine crabs (see section Grapsid Crabs). At least in Southeast Asia and Australia, this is a major factor determining mangrove species distribution. Mangrove animals respond to physical gradients of salinity and inundation regime. Sesarmine abundance is often greatest at midshore, and it is therefore here that mangrove propagules are most vulnerable. For reasons related to nutritional value and the levels of aversive tannins, *Avicennia* is generally the preferred food of sesarmines: hence at some locations the distribution of *Avicennia* in the upper and lower shore and their virtual absence from intermediate shore levels.

Random factors can also affect mangrove species distributions. If a gap is created in a mangrove forest because of the death of a tree, it is most rapidly filled by the species that are the best colonizers and best able to flourish in unshaded conditions. In Southeast Asia, the result is often an initial invasion of the mangrove fern *Acrostichum*. This may be succeeded by seedlings of *Bruguiera parviflora*. This species has relatively small and easily dispersed seedlings, whose growth is suppressed by the shade of an intact canopy. These in turn are replaced by slower growing shade-tolerant species such as *Bruguiera gymnorrhiza*. *Avicennia marina* is less tolerant of shade but is less likely to occupy a small gap because of propagule destruction by crabs. If, however, the gap is a large one, *Avicennia* is more likely to establish itself, probably because foraging crabs are vulnerable to predation in large open spaces. The distribution of species within a mangrove forest may therefore be patchy and reflect the stochastic nature of tree death and the subsequent successional history. On a larger scale, extensive death of trees by typhoons, by widespread defoliation by insect attack, or even by oil spills can have profound and long-lasting effects on species composition.

The structure of a mangrove forest is therefore in part explainable in terms of "patch dynamics" – of gaps appearing by chance and being filled by a changing assemblage of species differing in composition (at least for a time) from the surrounding forest. Eventually, something similar to the surrounding forest emerges. With a high incidence of gaps, a mangrove forest could be seen as a mosaic of patches of different successional age: if patches appear relatively rarely, the effect would be transient aberrations in an otherwise homogeneous, or consistently zoned, environment.

Distribution of Mangrove Animals

The species distribution of the mangrove fauna is less well understood since small, cryptic, and often mobile animal species are less easy to describe and analyze than large and immobile trees. A high level of taxonomic confusion compounds the problem. Nevertheless, it seems likely that the same general considerations apply. The distribution of mangrove crabs, for instance, forms zones related to shore level, salinity, and soil texture, whereas mollusks show zonation patterns in relation to shore level and to vertical position on the roots and trunks of mangrove trees.

The distribution of species of mangrove animals may also be related to patch size and the distance between neighboring patches, on a smaller spatial scale than applies to the distribution of mangrove species themselves, corresponding to the more limited dispersal ability of the species in question. This was demonstrated in the classical experiments of Simberloff on the terrestrial arthropod fauna (principally insects) of mangrove islets in the Caribbean. The species richness on a range of mangrove islets increased with the area of the islets and decreased with increasing distance from potential sources of fresh colonists. When the fauna of islets was completely eliminated with pesticides, recolonization soon established an equilibrium species richness similar to that before the elimination. In terms of the representation of different functional groups the previous situation was largely replicated, but the actual species comprising the new assemblages differed. Finally, artificially reducing the area of mangrove in experimental islets reduced species richness, showing that it was causally related to habitat area rather than to habitat diversity.

At an even smaller scale, individual mangrove roots can be regarded as "islands" of habitat suitable for epibiont settlement, surrounded by areas of unsuitable habitat. Here, too, the composition of root epibiont communities appears relatively stable in terms of functional groups. The actual species present are much more unpredictable and particularly affected by physical variables and by the supply of colonizing larvae. These factors are of different significance at different time and spatial scales.

Meiofaunal diversity has scarcely been investigated, although the same considerations apply as in the macrofaunal world. Variation in physical variables, species interactions, patchiness, dispersal, and the other factors relevant to larger organisms must also affect the meiofauna. To date, limited research interest (and the intrinsic difficulty of studying species interactions or measuring, e.g., nutrient gradients at a scale of millimeters) has restricted our knowledge of mangrove meiofaunal diversity and the factors which determine it.

Genetic Diversity of Mangroves

The advent of molecular genetic techniques has made it possible to study diversity at levels lower than the species. To date, few species have been studied, and clear general conclusions cannot be drawn. In some cases, such as the self-pollinating *R. mangle* of Florida and the Caribbean, populations appear to be genetically homogeneous, with slightly more genetic variation toward the northern extremes of the species' range. The extent of intraspecific genetic variation varies with the breeding structure of the population, with dioecious species showing much greater polymorphism. Genetic variation between populations is naturally greater than that within a population at a particular location, although West African mangroves have greater levels of genetic diversity than the same species in the Florida and Caribbean. This confirms the belief that western Atlantic mangroves derive from African populations rather than the reverse. As research proceeds, no doubt many such insights into the causes and consequences of intraspecific diversity will emerge.

Uses and Abuses of Mangroves

Mangroves are of interest not just to biologists. Their diversity and productivity makes them the source, directly and indirectly, of many products of use (and commercial importance) to humans (Bandaranayake, 1998).

Mangrove trees are exploited for timber for construction and firewood. This ranges from the casual collection of fallen wood to major charcoal industries based on the intensively managed mangroves of, for example, western peninsular Malaysia. Foliage may also be grazed directly or harvested for fodder for domestic animals. On a smaller scale, mangrove products are collected for a host of other purposes, including thatching houses, the manufacture of fish traps, for use in medicine, for tanning leather, and for use in various foods and drinks. Indirectly, mangrove productivity supports fisheries, both within the mangal and offshore. Less tangibly, mangroves can be of considerable importance in consolidating shorelines and limiting coastal erosion.

The significance of mangroves to humans varies greatly from place to place, but attempts have been made to achieve an overall economic valuation of the goods and services supplied. One recent estimate indicates that, on average, the annual value of a hectare of mangroves is approximately $10,000, resulting in a worldwide total contribution of $1,648,000,000.

An asset of this magnitude is worth conserving. Unfortunately, sustainable management of mangrove resources is the exception rather than the rule. In almost all parts of the world, mangroves are under pressure from irrigation schemes which divert rivers and prevent fresh water from reaching mangroves and from pollution, overexploitation, or deliberate clearance for construction or for the planting of alternative crops. One of the most destructive processes in many countries of Southeast Asia and Central America has been the clearance of mangroves for the construction of shrimp ponds – an attempt to increase the production of species dependent on mangroves while simultaneously reducing the primary production on which they depend. Not surprisingly, this has not been a success.

During the past few decades, loss of mangrove area in many countries has been dramatic. In the Philippines, for example, 60% of the mangrove area has disappeared, whereas in other countries such as Malaysia, Thailand, and Pakistan, annual losses are of the order of 1–3%. It may be, however, that the tide is turning. The virtual collapse of the shrimp industry in several countries and a greater awareness of the value of mangroves as a natural resource have focused attention on rational management strategies and on the possibility of reversing some of the damage. Much effort is now being put into replanting mangroves in abandoned shrimp ponds and the rehabilitation of denuded areas for coastal protection or in support of local fisheries as well as into developing suitable mangrove areas for ecotourism. The destruction of mangroves has largely been due to human activities: In the future their survival may also depend on mankind.

See also: Coastal Beach Ecosystems. Estuarine Ecosystems. Intertidal Ecosystems. Marine Ecosystems. Tropical Forest Ecosystems. Wetlands Ecosystems

References

Bandaranayake WM (1998) Traditional and medicinal uses of mangroves. *Mangroves and Salt Marshes* 2: 133–148.

Duke NC, Ball MC, and Ellison JC (1998) Factors influencing biodiversity and distributional gradients in mangroves. *Global Ecology and Biogeography Letters* 7: 27–47.

Field C (1995) *Journey Amongst Mangroves*. Okinawa, Japan: International Society for Mangrove Ecosystems.

Hogarth PJ (2007) *The Biology of Mangroves and Seagrasses*. Oxford: Oxford University Press.

Morrisey DJ, Swales A, Dittman S, Morrison MA, Lovelock CE, and Beard CM (2010) The ecology and management of temperate mangroves. *Oceanography and Marine Biology Annual Review* 48: 43–160.

Saenger P, Hegerl EJ, and Davie JDS (1983) Global status of mangrove ecosystems. *Environmentalist* 3(suppl. 3): 1–88.

Spalding M, Blasco F, and Field C (eds.) (1997) *World Mangrove Atlas*. Okinawa, Japan: International Society for Mangrove Ecosystems.

Stafford-Deitsch J (1996) *Mangrove. The Forgotten Habitat*. London: Immel.

Tomlinson PB (1986) *The Botany of Mangroves*. Cambridge: Cambridge University Press.

Marine and Aquatic Communities, Stress from Eutrophication

Jonathan H Sharp, University of Delaware, Newark, DE, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 1–11, © 2001, Elsevier Inc.

Glossary

Eutrophication Overenrichment of aquatic systems with nutrients, often leading to harmful algal blooms and subsurface oxygen depletion.

Harmful algal blooms Development of sufficient numbers of cells of algal species to cause "harmful" effect to ecosystem.

Microbial response Expected responses of microscopic algae to nutrient enrichment is excess production beyond what can be consumed by grazers and species shift to noxious species; this expected response is not necessarily what happens.

Nutrient enrichment Usually excess of nitrogen and phosphorus nutrients to aquatic systems.

Redfield ratio From large-area and time averaging, ratio of carbon, nitrogen, phosphorus, and oxygen for normal ocean plankton and deep-ocean nitrate and phosphate pools.

Subsurface oxygen depletion Due to isolation of waters below the surface, metabolic breakdown of organic matter from the surface waters can cause depletion of dissolved oxygen concentration, leading to very low oxygen (hypoxia) or no oxygen (anoxia).

Introduction

Perhaps the single anthropogenic stress on aquatic and marine environments that is considered to be most ubiquitous is described with the broad term eutrophication. Human activities have mobilized nitrogen and phosphorus through agriculture, urban and suburban sewage, and atmospheric emissions sufficiently to greatly increase fluxes to aquatic environments, especially to lakes, estuaries, and coastal ocean waters. The overly simplistic view of eutrophication is that loading of nitrogen and phosphorus causes increased production and biomass of planktonic algae, with decreased species diversity. Thus, nutrient loading creates a stress to the aquatic community causing adverse community impact. Characteristic symptoms of eutrophication are depletion of oxygen in subsurface waters from excess algal biomass and development of blooms of noxious algal species (Richardson, 1997). Limnological studies of eutrophication have given concepts and observations about the phenomenon which, when applied to estuarine and coastal waters, often lead to incorrect conclusions. Contrary to simple models, nutrient loading in nearshore marine waters often does not support the level of primary production and phytoplankton biomass that would otherwise be expected from the high nutrient concentrations. Depletion of oxygen in subsurface waters is probably more a function of the physics of the specific aquatic system and lack of grazer consumption than it is of nutrient-stimulated excess algal growth. A bloom of a noxious algal species is probably more a function of the response by the entire aquatic community to the overall chemical milieu than it is to nutrient-stimulated growth of the algal species.

This lack of a simple cause-and-effect response to nutrient loading is especially important to evaluate in light of two essentially opposite resource management actions. In nearshore waters, the traditional approach to eutrophication is removal of nutrient inputs to prevent algal production and growth. This approach has been successful in some cases. However, its success may be overestimated and probably, nutrient removal will have little of the intended impact in many estuarine and coastal waters. The second and opposite action with nutrient enrichment comes from recent proposals for large-scale engineering projects to fertilize waters of the open ocean to increase primary production (Cullen and Chisholm, 1999). In the one case, there is the simple goal of decreasing algal production (which is considered bad) by reducing nutrient inputs and in the other case a goal of increasing algal production (which is considered good) by adding nutrients; it's not that simple. A principal reason that we often misinterpret marine eutrophication is that insufficient consideration is given to the hydrodynamic and biogeochemical complexity of the environment as well as the biodiversity of the microbial community (Paerl, 1998).

In discussing biological diversity, Norse (1993) considers hierarchical levels that range from genetic to species to ecosystem. Stress appears to decrease community diversity, often also decreasing the number of species within an individual function, such as primary production. It is critical to consider that increased diversity at one trophic level may decrease diversity at another. An interesting example comes from a recent examination of thermal stress on a planktonic community. Microcosms were studied with examination of several trophic levels (Petchey *et al.*, 1999). It was found that environmental warming caused losses of top predators and herbivores, with increasing dominance by autotrophs and bacterivores. The warming increased extinction of predators with little effect on primary producers and bacterivores. Primary producer and bacterivore biomass increased, bacterial biomass did not change, and there were idiosyncratic impacts on total biomass. Warming directly increased primary production through temperature-dependent physiology and indirectly through changes in trophic structure. Warming also increased decomposition, probably both through physiology and

indirectly through food web structure. The results of this quantitative study are consistent with the general descriptive view of eutrophication where primary production and bacterial decomposition appear to increase while much of the rest of the trophic structure decreases (Vollenweider *et al.*, 1992).

Many anthropogenic inputs to the aquatic environment that may be considered potential stressors have a dual impact, both stimulating growth at low concentrations and decreasing growth at higher concentrations. At a functional level, like primary producers, it is well known that temperature has a stimulating influence on individual species until a threshold is reached, above which it can inhibit. There has been considerable interest recently about the influence of trace metals on primary producers showing this dual stimulating–inhibiting influence (Bruland *et al.*, 1991). The impact of nutrients is clearly similar if one looks at the full ecosystem. Overcoming limiting levels increases primary production and can increase entire ecosystem production, but only to an extent. Then too much nutrient loading may become a negative stress on the ecosystem. It is necessary to discriminate whether the effect is from nutrients or other inputs.

In the next section, the extension to the marine environment of the limnological example of eutrophication is examined. Following that is a section examining relatively low growth in high-nutrient environments and a section on ecosystem stoichiometry. The last section addresses the need for more thorough understanding of multiple stressor effects on the entire community, including grazer and trophic transfer.

Extrapolating from Lakes to the Sea

Lake Eutrophication

The classic example of reducing input of one nutrient to ameliorate eutrophication can be seen with the relatively successful cleanup effort that has taken place with lakes (e.g., Edmondson, 1991). Vollenweider's models of the 1960s have led to the single-element, phosphorus, approach to lake eutrophication. These were successfully applied to the experimental lakes program, where focus on P proved successful (Schlinder, 1981). In these studies, if carbon or nitrogen were in shortage, it could be brought in from atmospheric or sedimentary sources. Schlinder also concluded that micronutrients were not important except in rare cases. The accepted overview for lakes is that phytoplankton production is controlled by annual P-loading and that P-loading is directly related to P concentration and P concentration and chlorophyll are well correlated. Generalizations of this nature may be more valid for lakes than for marine and estuarine waters due to more thorough experimental information and the hierarchical approach used to study lake eutrophication (Hecky and Kilham, 1988). In addition, physical properties in estuarine and marine waters may make these systems more complex. Indeed, Schlinder (1981) has stated, "The control of eutrophication in estuaries has appeared to be much more complex that it is in lakes."

Estuarine and coastal ocean waters have considerable transport circulation driven by both tidal energy and riverine discharge and this circulation is variable on predictable and unpredictable periodicities. These environments also have complicating factors from variations in suspended sediments through specific input sites, circulatory resuspension of bottom sediments, and coagulation along estuarine salinity gradients. Additionally, the salinity gradients of estuaries cause significant physical stress such that populations of organisms will change and moreover the change in ionic strength will influence chemical speciation and solute–solvent interactions. The more complicated and varying physics gives rise to less predictability of populations on a seasonal basis. This pertains not only to primary producers but also to food web structure. There is a large literature on the subject of coastal ocean eutrophication, with recognition of difficulties in applying eutrophication concepts to estuarine and coastal waters, where physics and biogeochemistry complicate the picture (Vollenweider *et al.*, 1992).

Food Web Complexity

In addition to the complicated nature of coastal waters, it is well recognized that food web structure can bring significant variability to all aquatic ecosystems, lakes as well as coastal marine waters. Schlinder *et al.* (1997) demonstrated that the primary production enhancement from nutrient enrichment was less in a piscivore-dominated lake than in a planktivore-dominated lake. The reason for this difference was the suppression of phytoplankton by large zooplankton in the piscivore lake. The study focus included evaluation of drawdown of atmospheric CO_2 from nutrient-enhanced production and confirmed that changes in aquatic CO_2 fugacity could be successfully manipulated in lakes and open-ocean ecosystems.

The recognition of complexity in oceanic microbial ecosystems led to the concept of the "microbial loop," where organic matter from phytoplankton is rapidly consumed by bacteria, which are in turn consumed by protozoans that are consumed by small zooplankton that otherwise would be herbivores exclusively (Azam *et al.*, 1989). This shunt from the phytoplankton–herbivore direct trophic transfer means lower efficiency in the path to higher metazoans and led to debates of bacteria as "source or sink" in oceanic as well as estuarine environments (Sherr *et al.*, 1987). Subsequent work with the microbial loop has led to considering the combined phytoplankton and microbial heterotrophs (bacteria and protozoans) as the primary producer community supporting the metazoans (Sherr and Sherr, 1991). With the primary producer function from a multicompartment ecosystem, it is not surprising to find phytoplankton and heterotrophic bacteria being influenced by different stimuli. For example, Pace (1993) showed bacteria being controlled by phosphorus while phytoplankton were controlled by nitrogen and phosphorus in lake nutrient enrichment experiments. In most cases with any aquatic environment, nutrient addition will bring variable influences on phytoplankton and heterotrophic bacteria and may also influence heterotrophic protozoans and metazoan grazers.

It has been well recognized that phytoplankton production, or the entire primary producer community, is

influenced by removal (top-down control) as well as by resource limitation (bottom-up control). In a recent review of eutrophication in planktonic ecosystems, Glibert (1998) pointed out that grazing and nitrogen recycling are intricately connected in controlling planktonic nitrogen availability. Another important recognition is that top-down control has a major impact on export from the pelagic system (Wassman, 1998). Wassman warned that to view only bottom-up controls (nutrient influence) will not successfully guide biogeochemical studies of marine systems. Thus, it seems obvious that a nutrient influence on phytoplankton should not be considered in the absence of the rest of the beginnings of the ecosystem.

Response of Nearshore Waters to Nutrient Enrichment

The idea that a single nutrient controls primary production comes from classical ecological theory. With a single nutrient, a single phytoplankton species should have an advantage over others and dominate by outcompeting. The fact that multiple species can coexist within an apparent niche was considered a paradox (Hutchinson, 1961) and was the subject of a massive amount of excellent aquatic research. Recognizing the necessity to consider guilds rather than a single species makes understanding of primary production more complicated. Add to this the more recent recognition that multiple compartments of phytoplankton plus bacterial and protozoan heterotrophs may be considered as a primary producer community, and the need for whole-ecosystem experiments is obvious.

There is an oversimplified view that nutrient concentrations (or loading) above those of some "pristine" condition directly cause phytoplankton response, with negative impact in estuarine and coastal waters. The overall impression is that increased nutrients cause increased algal growth with the consequence of excess algal production causing oxygen depletion or the consequence of a bloom of noxious algae. Oxygen depletion from nutrient-enriched phytoplankton growth occurs in environments where summer stratification isolates bottom waters, for example, Chesapeake Bay and mid-Atlantic coastal waters. Oxygen depletion is often quite variable on an innerannual basis and moderated by meteorological forcing. Thus, the occurrence and extent of oxygen depletion are complicated and not simply predictable as a function of nutrient loading.

There is concern that unusual and noxious algal blooms are increasing globally in both geographic extent and intensity, although there is debate on the quantitative significance of such claims (Anderson, 1997). It is important to be careful in defining harmful algal blooms (HABs) as Smayda (1997) has indicated. In most cases, the sign of the "bloom" is the appearance of numbers of cells of a species of a harmful alga sufficient to have a negative environmental impact. Analysis of individual HABs shows that generally the HAB taxa have no unique ecophysiology, including higher affinity for nutrients, and often that the HAB taxa have growth rates lower than those of phytoplankton in general (Smayda, 1997). Many HAB taxa have allelochemically enhanced competition with other algal species and have allelopathic defense against predators as well as against a broad group of other microbial taxa. It appears that noxious algae bloom from their ability to dominate rather than their ability to outcompete other species for nutrients or to grow fast. What actually stimulates these taxa to express their domination is an area in need of more research. Full-ecosystem studies are needed to better understand noxious algal proliferations.

The experimental lakes program mentioned above has provided great empirical evidence to combine with theory in limnology. It is more difficult to manipulate whole parts of estuaries and coastal oceans than it is to do so with small lakes. Controlled mesocosms are a good compromise. The Marine Ecosystems Research Laboratory (MERL) at the University of Rhode Island has been one of the largest and most successful versions of realistic estuarine ecosystems (Oviatt *et al.*, 1986). In these, sufficient volumes of water have been used to overcome many problems of confinement and attempts have been made to simulate estuarine physical and biogeochemical influences. Some excellent research has been done and much learned about the complexity of estuarine responses to nutrients and other stressors.

A much used picture that was developed with information from the MERL experiments and from comparison of phytoplankton biomass and production in various estuaries and coastal waters has been shown by Nixon. The picture indicates phytoplankton increasing proportionately with increasing nutrient concentrations or loading. **Figure 1(a)** shows a generic version of this with phytoplankton production versus nitrogen loading. Note that the nitrogen loading is portrayed on a logarithmic scale (as was done in Nixon and Pilson, 1983), giving the appearance of regularly increasing production as a function of increasing nitrogen over a broad range of nutrient loadings. With transformation to a linear axis for N-loading, it is obvious that production becomes asymptotic after an initial linear increase. There have been several articles in which this conceptual picture has been shown and expanded upon; most of them are extensive evaluations of data from many published works, with the authors indicating that the relationship is complex (e.g., Nixon *et al.*, 1986). As has been cautioned by Nixon and others using it, the relationship is intended to cover a large range of nutrient conditions and to compare a number of different environments. However, a simplistic extension has been made suggesting that there is a simple linear relationship between nutrient loading and adverse phytoplankton production.

The behavior shown in **Figure 1(b)** is probably more correct to indicate that phytoplankton response to increased nutrients is not linear along a very long loading or concentration scale. Relatively small increases in nutrient concentrations and loadings will cause a large increase in primary production, but continued increases do not. In fact, it is likely that with very high nutrient concentrations a decreased phytoplankton response will be encountered as is shown with the theoretical curve in **Figure 1(c)**. Thus, it is not necessarily the case that nutrient enrichment leads to excess algal production. Perhaps, we should be addressing the question of why there is not greater phytoplankton production in estuarine and coastal waters from nutrient enrichment.

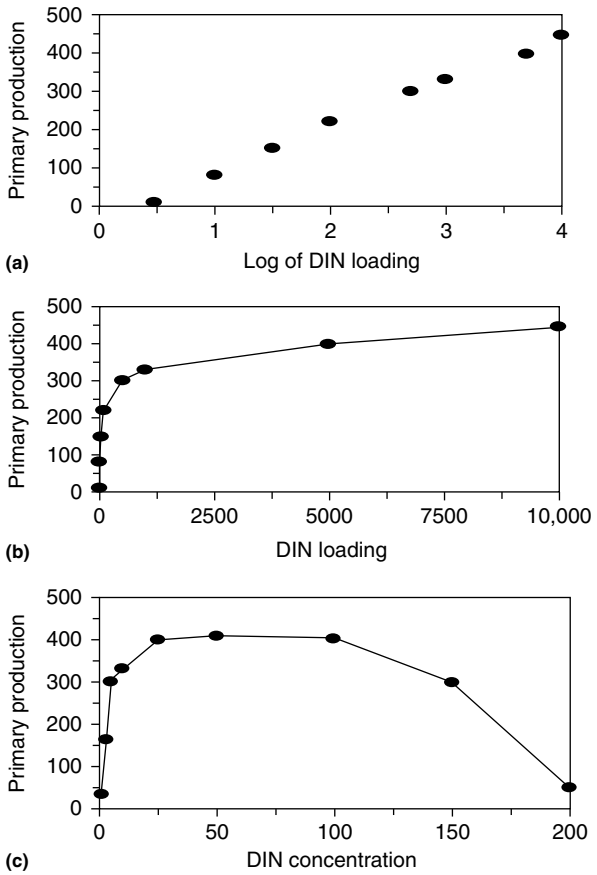


Figure 1 Theoretical relationship of dissolved inorganic nitrogen loading to primary production. Frame (a) shows the often-depicted version with N-loading on a logarithmic scale, while transformation to a linear scale for N-loading is shown in frame (b). Frame (c) shows inhibition at high N concentration (note concentration rather than loading).

Table 1

Characteristics of oceanic HNLC (high nutrient, low chlorophyll) environments

1. Relatively high concentrations of macronutrients
2. Low standing stock of phytoplankton
3. Moderate phytoplankton growth rates

Characteristics of estuarine HNLG (high nutrient, low growth) environments

1. High concentrations and fluxes of macronutrients
2. Moderate to high phytoplankton standing stock
3. Comparatively low phytoplankton growth rates
4. Sometimes, domination of flora by "undesirable" species

Nearshore Ocean Nutrient Response

The Oceanic HNLC Concept

Although it is well known that enhancement of primary production is not simply and directly proportional to enrichment by a single nutrient, there is a tendency to oversimplify this relationship in both nearshore and oceanic waters. I would like to use the oceanic HNLC (high nutrient, low chlorophyll) concept to examine enrichment of nearshore waters. The HNLC concept came from interest in iron limitation in the ocean that has a history going back to at least the 1920s (Martin *et al.*, 1990) but is best recognized in relation to Martin's proposed Antarctic and equatorial Pacific experiments (Coale *et al.*, 1996). Underlying the HNLC is the observation that some areas of the open ocean have relatively high concentrations of nitrogen and phosphorus nutrients but do not support proportionately large phytoplankton biomass; see **Table 1**. In open-ocean experiments, iron as a trace element added to overcome limitation has been shown to increase phytoplankton production with concomitant drawdown of atmospheric CO_2 in both the equatorial Pacific and the Antarctic oceans. The very nature of considering a trace constituent as critical to oceanic production indicates that the ecosystem is more complex than a simple cause-and-effect relationship with a major nutrient. In addition, the details of responses and trophic complexities in these oceanic experiments need more study before simple conclusions should be reached. However, this does not stop engineering plans for commercial fertilization of the ocean that are based on very simple cause-and-effect relationships. The same oversimplification in nearshore waters leads to proposals to solve eutrophication based upon very simple cause-and-effect relationships, i.e., reducing input of a single major nutrient.

An explanation for the observed HNLC conditions is the simultaneous control by grazing and micronutrients, such as iron. Cullen *et al.* (1992) have demonstrated with modeling for the equatorial Pacific upwelling region that grazing is the main control of standing stock but that a trace nutrient (e.g., Fe) might ultimately regulate overall productivity by influencing species composition and food web structure. Frost and Franzen (1992), with a chemostat model using a multiple-step food chain, demonstrate that simultaneous grazing control and trace nutrient limitation (e.g., Fe) could account for the observed conditions. Armstrong (1994) has shown with his multiple-species model that it is possible for each phytoplankton size class to be controlled by herbivores, while at the same time micronutrient limitation (e.g., Fe) may limit the number of size classes that can exist in a community and thus the total phytoplankton biomass that can be supported. A model has been presented (Armstrong, 1999) that shows HNLC conditions controlled by a combination of iron limitation of algal growth rates, ammonium inhibition of nitrate uptake leading to reduced uptake, and dependence of both processes on cell size. This dependence on cell size affects phytoplankton community structure and community uptake of nitrate. Recognition of the combined effect of a bottom-up influence (Fe limitation) and top-down influence (grazer control) would suggest that primary production in any aquatic environment has complex multiple controls.

High Nutrients and Low Growth in Estuaries

I suggest that we use a concept called HNLG (high nutrient, low growth) in estuarine waters that is similar to the oceanic HNLC, but with a different twist; see **Table 1**. A characteristic of the estuarine phenomenon is that phytoplankton show a

comparatively low growth rate and relatively high biomass. Probably a number of factors, individually or in combination, lead to the HNLG phenomenon. Proportions of macronutrients are often inappropriate for sustained high growth and limitation by micronutrients is also likely. Partial light limitation from algal biomass and from nonbiological suspended sediments also is involved in limiting growth. In addition, there is probably a negative influence on phytoplankton by contaminants and also inhibition of grazers by contaminants.

While lakes are generally considered P-limited, the traditional view is that marine waters are N-limited (Ryther and Dunstan, 1971). More recent evaluations have suggested that estuarine waters have alternating controls by nitrogen, phosphorus, and light. It is important to consider that nutrient enrichment is not necessarily an extension of nutrient limitation and that nutrient enrichment will not necessarily cause a direct proportional increase in phytoplankton. In some cases, fairly direct increases can be shown such as the indication that chlorophyll levels in the Chesapeake Bay have increased appreciably, with a doubling in dissolved inorganic nitrogen (DIN) over several decades. In contrast to this are examples in estuaries that do not show predicted increases. For example, Balls *et al.* (1996) show no change in chlorophyll in the Ythan River estuary between 1960 and 1990 when there was a fourfold increase in DIN. Alpine and Cloern (1992) showed a decline in primary production with increasing nutrient enrichment over time in the San Francisco Bay estuary. Reviewing conditions in a number of shallow coastal environments, Cloern (1999) showed a nonlinear response between N-loading and phytoplankton production and suggested that the simple eutrophication model in lakes does not have a current analog in coastal eutrophication.

Examining data on DIN concentration versus measured primary production for summer estuarine transects along the Delaware Estuary, we can show a negative relationship (**Figure 2**). While I will not suggest that the high nitrogen concentrations directly cause a decrease in phytoplankton production, it is clear that there is not a simple continual increase in production proportional to high nitrogen content. Note the similarity of **Figure 2** to the theoretical **Figure 1(c)**, where a small increase in DIN causes a large increase in production at concentrations near limiting, reaching an asymptote followed by a decline at very high DIN concentrations. To normalize measured primary production, the ratio of productivity to chlorophyll biomass (P/B) is often used. With data for the 1980s and 1990s from the Delaware Estuary, average summer P/B from the high-nutrient upper estuary is 61 (grams of C fixed per gram of chlorophyll) contrasted to the lower estuary production maximum region P/B value of 225 (Sharp *et al.*, unpublished data).

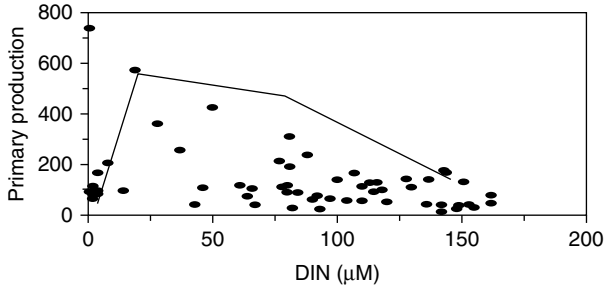


Figure 2 Primary production (integrated areal values as millimoles of C/m^2/day) versus total dissolved inorganic nitrogen (DIN) concentration for samples along the full gradient of the Delaware Estuary. Composite data from summer samplings over 3 years. Highest DIN concentrations are found in the urban freshwater region of the estuary, with slightly lower values immediately upstream in the tidal freshwater region and downstream going into the salinity gradient. Lowest DIN concentrations are found in the lower estuary near the mouth of the Delaware Bay.

In most estuaries, nutrient increases have not been uniform. Often increases in DIN are accompanied by smaller increases or decreases in dissolved inorganic phosphorus (PO_4) such that the N/P ratio has increased. For example, this is the case with the nitrogen increases in the Chesapeake Bay and the Delaware Estuary (Sharp, 1998).

Stoichiometry

Redfield Ratios

From extensive averages of planktonic CNP composition and deep-ocean NO_3 and PO_4 nutrient concentrations, it was noted that a regular and predictable N/P ratio was found (Redfield *et al.*, 1963). This Redfield ratio concept was extended to carbon and oxygen for ecosystem utilization of elements. The concept has been extensively applied to aquatic systems for over a half century and is a valuable guide in understanding biogeochemical fluxes. Analysis of particulate CNP in a number of fresh-water lakes indicates that a variety of conditions exist, ranging from N and P deficiency to sufficiency (Hecky *et al.*, 1993). They concluded that ocean plankton is not as N and P deficient as is lake plankton. From a different viewpoint, Flynn (1990) concluded that only in the presence of excess NH_4 is the cellular response to N-stress fully suppressed. Thus, plankton throughout the ocean show some symptoms of N-stress. He suggested that there are three forms of N status: N-replete (no stress), N-sufficiency (enough stress to depress NO_3 transport and assimilation), and N-deplete (maximum stress, no growth). With this classification, most estuarine, coastal, and open-ocean waters are N-sufficient. To evaluate whether or not there is nutrient stress from too little or too much, it is probably necessary to look over sufficient time to be in essentially steady-state conditions. For instance, it has been suggested that in Southern Ocean waters, Redfield CNP ratios for plankton use were only obtained when averaging over the full vegetative season of the Austral summer (Hoppema and Goeyens, 1999). They suggested that the Redfield ratio was reached only because of nutrient-replete conditions.

Adding Silicon

The concept of the Redfield ratio has been extended to silicon for many studies of aquatic ecology. From laboratory culture studies, the average Si/N ratios for small and large diatom species is close to 1/1 (Brezinski, 1985). This would give an

overall Redfield ratio for CNPSiO of 106/16/1/16/–276 for balanced diatom growth. For a long time, it has been assumed that with deficiency of Si, phytoplankton populations will shift from diatom domination to that of other groups of algae. However, recent research would suggest a more absolute limitation of "healthy" ecosystem production by Si. In oceanic HNLC environments, it has been suggested that "new" production (that which is supported by upwelled NO_3) is reduced by Si limitation and thus export from pelagic primary production is controlled by Si availability. In coastal upwelling regions, it has been demonstrated that iron limitation will cause diatoms to increase the Si/N uptake ratio, depleting the water of Si, leading to secondary Si limitation. Clearly, Si is very important and can influence N response of the primary producers.

Comparative Si availability may be a major feature in the apparent eutrophication response seen in nearshore waters. In a recent 20-year comparison in the Bay of Brest, there was a large decrease in the Si/N ratio, "but, contrary to what has been observed in other coastal ecosystems, phytoplankton stocks have not increased" (LePape *et al.*, 1996). In light of the earlier discussion (*see* High Nutrients and Low Growth in Estuaries), maybe this is less of an exception than LePape *et al.* interpreted. In some cases, an increase in phytoplankton biomass is seen, but not always with a shift from diatoms, and rarely is there an increase in higher trophic level consumption of the primary production. In an extensive study of a very long term record for the Mississippi River outflow into the Gulf of Mexico, Rabalais *et al.* (1996) have shown a large decrease in the Si/N ratio accompanied with an increase in primary production but also an increase in the deposition of biogenic silica in the sediments underlying an increasingly large hypoxia region. The explanation is that with relative Si scarcity, diatoms that are in the plankton are not grazed efficiently, and they fall to subsurface waters and contribute to hypoxia.

Changing Nutrient Ratios

Probably most estuarine waters with impact from human activities show greatly changed N/Si as well as N/P ratios. In **Table 2**, average values for total dissolved inorganic nitrogen (NO_3 plus NH_4), PO_4, and Si are shown for several nutrient-enriched estuaries. All of the examples show large increases in DIN and most have smaller proportional increases in P so that the N/P is usually considerably higher than would be the case without the anthropogenic influence. Since Si is not usually a byproduct of human activity, the Si concentration has not changed much; there is probably a large natural variation depending upon the nature of the land drained for the estuary. A few systems probably have had significant decreases in Si due to decreased natural land erosion (dams, diked river banks); this is definitely the case with the Mississippi (Rabalais *et al.*, 1996). As a result, the N/Si ratio is much different from that prior to human impacts. Inverting this as Si/N, the pristine condition is about 10/1 and most of our nutrient-enriched systems show values of 1/1 or lower. This very likely has a serious negative impact on the primary production community. The importance of Si in relation to eutrophication has been recognized in the past, but usually only in relation to shift from diatom to flagellate flora (e.g., Officer and Ryther, 1980). With more recent information on interactive influences of Si, N, and Fe and on the fate of Si-limited diatom production, it is timely to reinvestigate the role of Si on eutrophication. While species responsible for HABs do not necessarily show greater affinity for nutrients in general, giving them ability to outcompete more "normal" phytoplankton like diatoms, it is probable that changing ratios of N and P to Si do favor some of the HAB flagellates (Smayda, 1990).

The large changes in N/P ratios are often not documented because of lack of complete nutrient records from long-range monitoring. In a data set from the Delaware Estuary, dissolved inorganic nitrogen has been measured regularly along the full axis of the estuary for over 35 years, but parameters for phosphorus measurements have varied over that period. Total P, a composite that includes dissolved organic and particulate phosphorus as well as PO_4, has been measured consistently. A comparison of the N/P ratio change, based on the total P, over a 30-year period is shown in **Figure 3**. Recognizing that the majority of the P in the estuary today is PO_4 and that in the past PO_4 was probably a larger portion of the total, it is possible to view the N/P ratio as indicative of available P. This dramatic N/P ratio change is probably largely due to reduced input of detergent phosphorus and the same change has occurred in many U.S. estuaries (N. A. Jaworski, unpublished data, 1998). In the earlier situation, almost the entire estuary would appear to be replete in relation to P since the N/P was considerably below the Redfield ratio; in the more modern situation, N/P ratios are in the 30–60 range. However, it must be recognized that transport and availability of phosphorus in estuaries is a complex function that also involves geochemical influences. In the past 20 years, the P geochemical reactivity in the Delaware Estuary has changed due to increased pH and

Table 2 Approximate nutrient concentrations at the beginning of the salinity gradient for several nutrient-enriched estuaries[a]

Estuary	*DIN*	*PO_4*	*Si*	*N/Si/P*	*Reference*
Scheldt	550	15	250	37/17/1	Zwolsman, 1994, 1999
Delaware	250	5	125	40/20/1	Sharp, unpublished data
Mississippi	114	7.7	108	15/14/1	Rabalais *et al.*, 1996
Chesapeake	75	1	50	57/50/1	Malone *et al.*, 1996
Northern San Francisco Bay	40	2	200	20/100/1	Peterson *et al.*, 1985
Pristine	10	0.5	100	20/200/1	Fanning and Maynard, 1978; Meybeck, 1982

[a]Average concentrations of nutrients (in μM element) approximated from publications listed. Averages for total dissolved inorganic nitrogen (DIN), dissolved phosphorous (PO_4), and silicate (Si) and ratios normalized to P are listed. Values for pristine estuaries approximated from data for the Zaire and Magdelena River outflow systems.

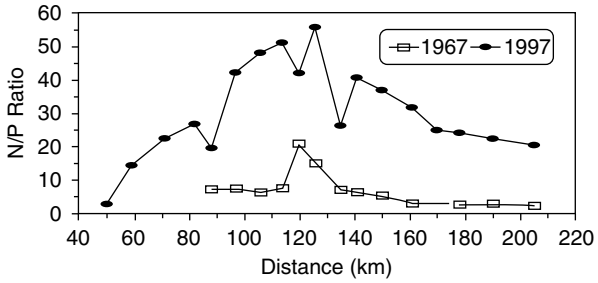


Figure 3 N/P ratios for the Delaware Estuary using total dissolved inorganic nitrogen (DIN) versus total P from summer transcects of the entire estuary in summers of 1967 and 1997 data from Delaware River Basin Commission routine monitoring.

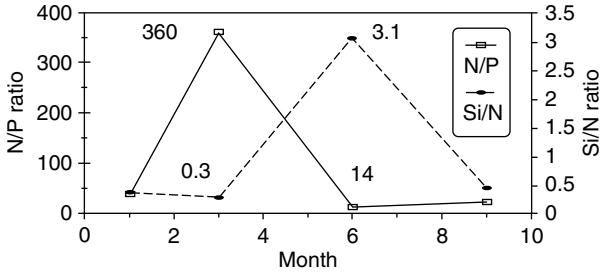


Figure 5 Average ratios of N/P and N/Si for stations in the lower Delaware Bay for winter, spring, summer, and fall. During the spring (month 3), the maximum biomass for the entire estuary for the year is found in the lower bay. The maximum estuarine primary production is found in the summer in the lower bay.

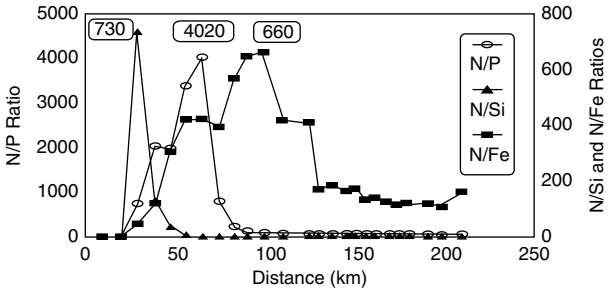


Figure 4 Ratios from sampling along the distance axis of the Delaware Estuary in the spring from measurements of total dissolved inorganic nitrogen (N), P, Si, and Fe. The N/P ratio is generally in the 30–100 range over much of the estuary during much of the year and <16 only near the mouth of the bay. The N/Si ratio is generally between 1 and 2 during most of the year. The N/Fe ratio is in the 100–600 range except near the bay mouth, where it can become <100.

dissolved oxygen. As a result, the N/P ratio based on PO_4 only for average concentrations of the entire salinity gradient of the estuary has decreased in the past 20 years from about 90/1 to 40/1. More thorough analysis of many estuaries may show this dual trend of long-term decrease of N/P loading in an upper estuary but of more available dissolved PO_4 being delivered to the lower estuary. It is important to understand the full biogeochemical picture of estuarine phosphorus before accurate conclusions of nutrient impacts can be made.

In addition to long-term changes in nutrient ratios, there are large spatial changes in estuaries at any single time. **Figure 4** shows nutrient ratios along the full length of the Delaware Estuary from sampling in the spring. In the spring bloom condition in the estuary, NH_4, PO_4, and Si are exhausted from the mouth of the estuary moving up toward a strong-light-limiting turbidity maximum (Pennock and Sharp, 1994). The very high ratios of P and Si to N are due to the large excess NO_3 concentration of the river water as it is advected downstream. It is interesting to see that different regions of the estuary appear to have large differences in the nutrient that could be most limiting. Also, from this picture, it appears that Fe could be more limiting in the upper estuary than near the mouth of the bay. This greater proportional availability of Fe is a year-round occurrence. A noxious algal group that has caused considerable international concern recently is responsible for brown tides. A recent analysis of brown tide occurrence suggests the macronutrient levels are not implicated but that Fe is.

In the Delaware Estuary, the area of the greatest phytoplankton production throughout the year is the lower bay (Pennock and Sharp, 1994). **Figure 5** shows macronutrient ratios for this region averaged on a seasonal basis. The maximum biomass is achieved in the spring bloom, where N/P and Si/N ratios indicate exhaustion of P and Si with residual NO_3; at this time, nutrient regeneration is minimal and there is also little herbivore grazing. Usually, the highest seasonal primary production is found in the summer. At this season, it would appear that grazing controls chlorophyll (Pennock and Sharp, 1994) and grazing is in balance with nutrient regeneration, which is sufficient to allow measurable NH_4, PO_4, and Si levels). At this season, the N/P ratio is close to Redfield and the Si/N ratio is sufficient to support healthy diatom growth although the flora is dominated by small flagellates. Acknowledging the caution of Hecky and Kilham (1988) that nutrient concentration does not equal nutrient utilization, we have demonstrated limitation in several ways (Pennock and Sharp, 1994). Looking at the entire estuary with highly nutrient enriched tidal freshwater region, a light-limited turbidity maximum in the oligohaline region, and clearer and nutrient-diluted lower bay, it would appear that the only time and place that near to Redfield ratio of nutrients is found is in the lower bay in the summer. We have performed preliminary simple mesocosm experiments and find that the CNPO fluxes approach Redfield stoichiometry only in the lower estuary in the summer (Sharp *et al.*, unpublished data). In the lower estuary in the spring bloom, close to Redfield ratios of NPSiO occur but there appears to be an accumulation of C with the biomass accumulation. In the nutrient-rich upper estuary, nothing close to expected stoichiometry is seen. Further research on this anomalous stoichiometry is currently underway in my laboratory.

Contaminants and Stress

The lack of high growth rates of estuarine and coastal phytoplankton in the presence of high nutrient concentrations leads to the suggestion that anthropogenic contaminants,

other than nutrients, may have an influence on ecosystem response. The MERL research facility mentioned earlier has extensive and expensive controls of the large tanks for the studies. A number of excellent experiments have been carried out also in less sophisticated smaller mesocosms in other estuarine areas. One interesting study done in Chesapeake Bay mesocosms over several seasons showed that nutrient enrichment by N and P caused growth of "beneficial" diatom species over flagellates (Sanders *et al.*, 1987). A conclusion of that study was that many other factors probably control community dynamics. As discussed earlier, the lack of predation is often a factor in noxious algal species proliferation. Also discussed earlier is the suggestion that ratios of nutrients may be more important than quantities in causing eutrophication responses. It is possible also that contaminants cause stress to the "desirable" algal species and to grazers that would otherwise consume the primary production.

Most estuaries with large anthropogenic influences have had chronic exposure to many chemicals in addition to nutrients. Chronic exposure to arsenic appeared to cause reductions in phytoplankton cell size with less trophic transfer while chronic exposure to silver resulted in essentially the opposite. Thus, arsenic would appear to be partially responsible for eutrophication response. Many estuaries also have frequent or continuous inputs of chlorination byproducts which must have a selected influence. Sanders (1984) showed that one diatom and a chrysophyte would not grow in aged chlorinated water (with levels below detection for total residual chlorine) where a more resistant flagellate would grow. This result would also favor species other than "normal" with potential to decrease trophic transfer.

Paerl (1998) has illustrated positive and negative interactions and feedback from nutrient loading, emphasizing that negative influences on grazing can increase the impact of primary production. In lakes, it has been shown that combinations of nutrient additions and zooplankton size can have major influence on phytoplankton sizes and thus on trophic transfer. The variability in nutrient and trace metal impacts on phytoplankton, bacterioplankton, heterotrophic protozoa, copepods, fish, and benthos caused variable trophodynamic responses in estuarine mesocosm experiments (Breitburg *et al.*, 1999). These authors concluded that trace elements may mask the response of high nutrient loadings in eutrophic systems. In lake studies, N, P, and C had different controlling effects on zooplankton, phytoplankton, and bacterioplankton, with variable responses in different seasons.

An overall conclusion is that eutrophication as a stress in estuarine and coastal marine environments is not a simple cause-and-effect phenomenon. Nutrient enrichment elicits complex and variable responses from the phytoplankton, bacterioplankton, and protozoa that make up the primary producer community. Increased nutrient concentrations and loadings may cause an overall increase in phytoplankton biomass, but not invariably from higher growth rates. Shifts in dominant species and shunts through the microbial loop may decrease the trophic transfer to higher metozoan levels. The biodiversity of the aquatic community may allow some resilience to the system (Patrick, 1988), but ultimately changes in the diversity are probably more important than a direct response to increased levels of nutrients.

See also: Coastal Beach Ecosystems. Estuarine Ecosystems. Lake and Pond Ecosystems. Plankton, Status and Role of. River Ecosystems

References

Anderson DM (1997) Turning back the harmful red tide. *Nature* 388: 513–514.
Armstrong RA (1999) An optimization-based model of iron–light–ammonium colimitation of nitrate uptake and phytoplankton growth. *Limnol. Oceanogr.* 44: 1436–1446.
Azam F, Fenchel T, Field JG, Ray JS, Meyer-Reil LA, and Thingstad F (1989) The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* 10: 257–263.
Brezinski MA (1985) The Si : C : N ratio of marine diatoms: Interspecific variability and the effect of some environmental variables. *J. Phycol.* 21: 347–357.
Bruland KW, Donat JR, and Hutchins DA (1991) Interactive influences of biactive trace metals on biological production in oceanic waters. *Limnol. Oceanogr.* 36: 1555–1577.
Cloern JE (1999) The relative importance of light and nutrient limitation of phytoplankton growth: A simple index of ecosystem sensitivity to nutrient enrichment. *Aquat. Ecol.* 33: 3–16.
Coale KH, Johnson KS, Fitzwater SE, *et al.* (1996) A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. *Nature* 383: 495–501.
Cullen JJ and Chisholm SW (1999) Commercial ocean fertilization: We know enough to know better. *EOS* 80: OS66–OS67.
Hecky RE and Kilham P (1988) Nutrient limitation of phytoplankton in freshwater and marine environments: A review of recent evidence on effects of enrichment. *Limnol. Oceanogr.* 33: 796–822.
Hutchinson GE (1961) The paradox of the plankton. *Am. Nat.* 95: 137–145.
Martin JH, Gordon RM, and Fitzwater SE (1990) Iron in Antarctic waters. *Nature* 345: 156–158.
Meybeck M (1982) Carbon, nitrogen, and phosphorus transport by world rivers. *Am. J. Sci.* 282: 401–450.
Nixon SW and Pilson MEQ (1983) Nitrogen in estuarine and coastal marine ecosystems. *Nitrogen in the Marine Environment*, pp. 565–648. New York: Academic Press.
Nixon SW, Oviatt CA, Frithsen J, and Sullivan B (1986) Nutrients of the productivity of estuarine and coastal marine ecosystems. *Limnol. Soc. Southern Africa* 12: 43–71.
Norse EA (1993) *Global Marine Biological Diversity.* Island Press.
Officer CB and Ryther JH (1980) The possible importance of silicon in marine eutrophication. *Mar. Ecol. Prog. Ser.* 3: 83–91.
Oviatt CA, Keller AA, Sampou PA, and Beatty LL (1986) Patterns of productivity during eutrophication: A mesocosm experiment. *Mar. Ecol. Prog. Ser.* 28: 69–80.
Pace ML (1993) Heterotrophic microbial processes. *The Trophic Cascade in Lakes*, pp. 252–277. Cambridge, UK: Cambridge Univ. Press.
Paerl HW (1998) Structure and function of anthropogenically altered microbial communities in coastal waters. *Curr. Opin. Microbiol.* 1: 296–302.
Patrick R (1988) Importance of diversity in the functioning and structure of riverine communities. *Limnol. Oceanogr.* 33: 1304–1308.
Pennock JR and Sharp JH (1994) Temporal alternation between light- and nutrient-limitation of phytoplankton production in a coastal plain estuary. *Mar. Ecol. Prog. Ser.* 111: 275–288.
Petchey OL, McPhearson PT, Casey TM, and Morin PJ (1999) Environmental warming alters food-web structure and ecosystem function. *Nature* 402: 69–72.
Rabalais NN, Turner RE, Justic D, Dortch Q, Wiseman WJ, and Sen Gupta B (1996) Nutrient changes in the Mississippi River and system responses on the adjacent continental shelf. *Estuaries* 19: 386–407.
Redfield AC, Ketchum BH, and Richards FA (1963) The influence of organisms on the composition of sea water. *The Sea* 2: 26–77.
Richardson K (1997) Harmful or exceptional phytoplankton blooms in the marine ecosystem. *Adv. Mar. Biol.* 31: 301–385.
Ryther JH and Dunstan WM (1971) Nitrogen, phosphorus, and eutrophication in the coastal ocean environment. *Science* 171: 1008–1013.
Sanders JG, Cibik SJ, D'Elia DF, and Boynton WR (1987) Nutrient enrichment studies in a coastal plain estuary: Changes in phytoplankton species composition. *Can. J. Fish. Aquat. Sci.* 44: 83–90.

Schlinder DE, Carpenter SR, Cole JJ, Kitchell JF, and Pace ML (1997) Influence of food web structure and carbon exchange between lakes and the atmosphere. *Science* 277: 248–251.

Schlinder DW (1981) Studies of eutrophication in lakes and their relevance to the estuarine environment. In: *Estuaries and Nutrients*, pp. 71–82. Clifton, NJ: Humana Press.

Sharp JH (1998) Trends in nutrient concentrations in the Delaware Estuary. In: Majumdar SK, Miller EW, and Sage LE, (eds.) *Ecology and Restoration of the Delaware River Basin*, pp. 78–192 . Pennsylvania Academy of Sciences.

Sherr BF, Sherr EB, and Albright LJ (1987) Bacteria: Link or sink? *Science* 235: 88.

Sherr EB and Sherr BF (1991) Planktonic microbes: Tiny cells at the base of the ocean's food web. *Trends Ecol. Evol.* 6: 50–54.

Smayda TJ (1990) Novel and nuisance phytoplankton blooms in the sea: Evidence for a global epidemic. In: Graneli E, Sundstrom B, Edler L, and Anderson DM (eds.) *Toxic Marine Phytoplankton*, pp. 29–40. Amsterdam: Elsevier.

Smayda TJ (1997) What is a bloom? A commentary. *Limnol. Oceanogr.* 42: 1132–1136.

Vollenweider RA, Marchetti R, and Viviani R (1992) *Marine Coastal Eutrophication.* Amsterdam: Elsevier.

Marine Conservation in a Changing Climate

Malin L Pinsky, Princeton University, Princeton, NJ, USA
Kristy J Kroeker and Daniel J Barshis, Hopkins Marine Station of Stanford University, Pacific Grove, CA, USA
Cheryl A Logan, California State University Monterey Bay, Seaside, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acclimation Acclimatization that occurs in a controlled, experimental setting such as a laboratory (*see* Acclimatization).
Acclimatization Ability of an organism to adjust its phenotype to new environmental conditions over the course of its lifetime given its genotype (*see* Phenotypic plasticity).
Adaptation Beneficial modifications to an organism's phenotype that are caused by heritable changes in its genotype.
Eurythermal Ability to tolerate a wide range of temperatures.
Extirpation Local extinction whereby a group of organisms ceases to exist in a particular location but may still exist in other locations.
Local adaptation Process by which populations of organisms become adapted to their local environment or ecological niche. Populations that are locally adapted generally exhibit the highest fitness in their native environment and reduced fitness in other, contrasting locations or conditions.
Net primary productivity Rate of production of biomass by plants and other primary producers such as algae. This forms the basis for all food webs.
Ocean acidification Ongoing changes in carbonate chemistry of the Earth's oceans (e.g., decrease in pH and decrease in carbonate ions) caused by their uptake of anthropogenic carbon dioxide from the atmosphere.
Phenotypic plasticity Ability of an organism to change its phenotype in response to changes in the environment.
Resilience Ability of a system to maintain function in the face of disturbance, including recovery from disturbance.

Introduction

Although climate change has quickly become a central concern in the conservation of biodiversity, marine systems have been largely overlooked (Richardson and Poloczanska, 2008). The effects and rates of climate change, however, are just as dramatic in the ocean as on land (Burrows *et al.*, 2011). Marine species and ecosystems have adapted to climate variability throughout geologic history, but the rate of change in climates both now and through the rest of the century is expected to reach or exceed their abilities to respond and persist into the future. In addition, marine communities are faced not only with climate change but also with fishing, habitat destruction, pollution, invasive species, new pathogens, and other anthropogenic impacts, all of which compound the detrimental effects of climate change.

Rising atmospheric carbon dioxide (CO_2) is expected to result in many abiotic changes to the marine environment. At ocean basin scales, numerous long-term changes have already been observed. Global average air and ocean temperatures are increasing, dissolved oxygen levels are decreasing, and global average sea level is rising (IPCC, 2007). In a process called "ocean acidification," elevated CO_2 in the atmosphere is dissolving into the ocean and causing a decrease in pH and carbonate ions. Changes in precipitation, ocean salinity, wind patterns, and aspects of extreme weather, including heavy precipitation and tropical cyclone intensity, have also been attributed to climate change. Besides the ongoing changes that have been detected, other changes are predicted to occur in ocean currents, vertical structure of the ocean, flow of nutrients, and general patterns of variability.

Marine organisms are directly affected by this rapidly changing environment. Potential responses for species to climate change can be broadly grouped into four general categories (sensu Fuller *et al.*, 2010; Parmesan, 2006): (1) they move in pace with the changing climate (e.g., range expansions, colonization of new niches, migration into novel habitats), (2) they adjust to new environmental conditions with their existing abilities (acclimatize), (3) they evolve new abilities to tolerate the climate, or (4) they face extinction or extirpation. Species may also have multiple and simultaneous responses (O'Connor *et al.*, 2011).

In this chapter, we explore the state of knowledge regarding both the impacts of climate change on marine species and the range of strategies proposed for marine conservation. For the sake of organization, we present impacts separately in terms of physiology, population abundance and distribution, evolution, and communities, though in reality there are important interactions among these impacts. We then discuss conservation strategies as integrated responses to help marine ecosystems cope with and adapt to the coming challenges of climate change. For further information, we direct readers to a number of well-crafted reviews over the past two decades (Fuller *et al.*, 2010; Brierley and Kingsford, 2009; Fields *et al.*, 1993; Hoffmann and Sgro, 2011; Harley *et al.*, 2006).

Physiological Impacts

How will coming abiotic changes affect the physiology of organisms living in the sea? In this section, the focus will be on

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00336-1

how two commonly studied aspects of climate change affect the physiology of marine animals: rising seawater temperatures and ocean acidification. By examining how marine organisms respond to these two climate stressors, we gain a better understanding of the types of physiological traits that are most important. Determining the capacity of organisms to acclimatize to environmental change is also key to understanding how populations and communities will respond to climate change.

Rising Temperatures

The vast majority of marine fish and invertebrates are "ectotherms," meaning that their body temperature fluctuates with that of the surrounding environment. Ectotherms are limited in their ability to internally regulate temperature, thus making them highly vulnerable to global warming. Marine ectothermic fishes and invertebrates have certain temperature ranges in which their performance is maximal. For example, Atlantic cod (*Gadus morhua*), an important fisheries species, grows and reproduces best within a temperature window close to $\sim 10\ ^{\circ}C$ (Pörtner *et al.*, 2008). In general, as temperatures warm, growth and metabolic rates increase. This may be initially beneficial, but further increases in temperature will push fish outside of their optimal window and lead to excessive metabolic demands. Beyond optimal temperature thresholds (termed "pejus temperatures"), cod struggle to maintain cardiac function and respiration. This leads to decreased swimming performance and less energy available for growth and reproduction (Pörtner and Knust, 2007). As temperatures increase even further, enzymes and other proteins within tissues and cells lose their conformational shape and fail to catalyze reactions (termed "denaturation temperatures"). Exposure to temperatures beyond this denaturation threshold can only last a short amount of time before mortality ensues. During this time, animals can turn to cellular defenses (including heat shock proteins and antioxidants) to temporarily prevent molecular malfunction until they return to cooler waters. The generalized scheme of thermal windows defined by optimal, pejus, and denaturation thresholds applies to both fish and invertebrates (Pörtner and Farrell, 2008).

To a limited extent, organisms can also adjust their thermal windows up or down in response to the temperatures they have experienced over longer time periods (for example, weeks or months). This process is called "acclimatization." It remains uncertain whether animals can adjust their thermal window far enough and fast enough upward to keep pace with climate change without negative impacts on other aspects of their physiology (e.g., Donelson *et al.*, 2010).

A species' thermal history can often help scientists determine how susceptible species will be to climate change. For example, animals already living near their thermal limits (the edges of these windows) may be most at risk (Somero, 2010). A recent study on a marine in-shore fish called the red moki (*Cheilodactylus spectabilis*) in the Tasman Sea showed that temperatures may have already reached levels associated with negative metabolic effects (Neuheimer *et al.*, 2011). Increasing temperatures were correlated with increased growth for populations in the middle of the species range, but warming appeared to cause reduced growth rates for populations at the warm northern edge of the species' distribution. This pattern indicates that temperatures may have already reached levels associated with high metabolic costs at this northern edge (Neuheimer *et al.*, 2011). If warming continues, a range contraction may be inevitable for this species. A suite of studies on intertidal porcelain crabs (*Petrolisthes* spp.) and marine snails (*Chlorostoma* spp.) showed that the most heat-tolerant species have little ability to increase their lethal and sublethal thermal tolerance windows through acclimatization (reviewed in Somero, 2010). In contrast, an in-depth examination of acclimatory and adaptive capacity of thermal tolerance limits in European diving beetles (Calosi *et al.*, 2008) found the opposite trend. Organisms with the highest upper thermal tolerance limits demonstrated the greatest capacity for acclimation. Contrasting findings such as these mean that broad generalizations regarding thermal tolerance and responses to future rapid climate change should be treated with caution for the moment.

Animals living in narrow, constant thermal habitats are also thought to have relatively low capacity to shift their thermal windows upward. For example, antarctic notothenioid fish (*Trematomus bernacchii)* have lived in a cold, constant environment for more than 15 million years and have very limited tolerance for higher temperatures. Some antarctic invertebrates have likewise shown limited ability to acclimatize to warmer temperatures (Peck *et al.*, 2004). These fish and invertebrates do not have the option of moving to a cooler location. Like the crabs and fishes described previously, these fish may also be the losers in a warmer world.

In contrast, animals living in fluctuating environments may be climate change winners. For example, a set of studies of reef-building corals from American Samoa found that corals from a highly variable environment may be acclimatized (and locally adapted) to tolerate greater levels of environmental stress when compared to corals from more stable neighboring areas (Barshis *et al.*, 2010; Smith *et al.*, 2008). Several physiological mechanisms have been described to explain this, including changes in the symbiotic algae that live within corals and a greater production of stress proteins and antioxidants during extreme heat stress events (Barshis *et al.*, 2010; Smith *et al.*, 2008). Whether marine organisms living in fluctuating environments are generally more tolerant to temperature stress has yet to be determined.

Ocean Acidification

Compared to temperature, less is known about the complex effects of ocean acidification on marine animals, not to mention the combined effects of temperature and changing ocean chemistry. Research to date suggests that ocean acidification has three primary physiological effects on marine animals. First, decreased carbonate ion concentrations make the calcification process more difficult for calcifying organisms such as corals and shellfish. These organisms require carbonate ions to build their shells and skeletons. Second, lowered pH and a greater amount of dissolved CO_2 in the water could affect acid–base regulation and metabolic rates as well as other physiological processes. Third, greater amounts of

dissolved CO_2 could alter the ability of primary producers to photosynthesize.

Most of the research so far has focused on negative calcification effects, and this research generally suggests that many calcifiers will be "losers." A decrease in the concentration of carbonate ions has been correlated with reduced calcification in several organisms, including abundant planktonic species (e.g., coccolithophores, pteropods, and foraminifera) and benthic invertebrates (e.g., coral, calcifying algae, mollusks, echinoderms) (Erez *et al.*, 2011; Guinotte and Fabry, 2008; Raven *et al.*, 2005; Fabry *et al.*, 2008). Marine calcifiers are thought to be highly sensitive to acidification not only because fewer carbonate ions in the water make calcium carbonate formation more difficult, but also because many calcifiers have low capacity for acid–base regulation (Hofmann and Todgham, 2010; Pörtner *et al.*, 2005). Calcification is energetically costly (Palmer, 1992) and appears to be one of the first physiological processes to be slowed down under conditions of acidification (Todgham and Hofmann, 2009). However, some species are able to maintain their calcification rates when supplied with enough energy (i.e., food and nutrients) (Cohen and Holcomb, 2009). Because the cellular mechanism of calcification in most of these animals is still unknown and because responses can vary across different life stages, this story is far from clear. Some marine calcifers are affected much less than others, and even closely related species can exhibit quite different responses (Dupont *et al.*, 2010).

The role of pH is diverse in physiological processes, ranging from maintaining the proton gradients that allow organisms to produce energy from food (respiration) and fix carbon (photosynthesis) to affecting the functions of whole organisms. As CO_2 levels in the ocean increase and seawater pH decreases, the ability of organisms to maintain extracellular pH homeostasis (acid–base balance) may be compromised. In sensitive animals, increases in the amount of CO_2 may result in depressed metabolic rates, reduced ion exchange, and lower protein synthesis rates (Pörtner *et al.*, 2005). Overall, this could shift a species' metabolic equilibrium and slow its growth. Sensitivity to elevated CO_2 levels is lower in fish than in most invertebrates, especially those fish with higher metabolic rates. Fish may be less sensitive because they have a more advanced oxygen transport system and because vertebrates have a greater ability to regulate ion exchange than invertebrates (Pörtner *et al.*, 2005). For example, even moderately higher amounts of CO_2 in the water can cause abnormal increases in the acidity of body tissue and fluids (acidosis) of some invertebrates, including some squid and mussel species, and it may contribute to their low resistance and relatively high mortality (Pörtner *et al.*, 2005). In contrast, with the same levels of CO_2 in the water, most fish can return the pH of their fluids to normal using compensatory mechanisms. Thus, the vast majority of acidification studies focus on invertebrates or on early developmental stages in fish before they gain these compensatory mechanisms.

In contrast, some physiological processes such as photosynthesis may benefit from increased levels of CO_2 in the water. For example, certain species of sea grass have increased photosynthetic rates and reduced light requirements when CO_2 levels are high (Guinotte and Fabry, 2008; Zimmerman *et al.*, 1997). However, most marine producers have evolved carbon-concentrating mechanisms that allow them to cope with the relatively low levels of available CO_2 in seawater (Reinfelder, 2011). Although CO_2 does not generally limit photosynthesis in marine primary producers now, it is possible that higher CO_2 concentrations in the future will reduce the energetic and nutrient costs of carbon-concentrating mechanisms and that this could then modulate primary production and species composition of phytoplankton communities in the future (Reinfelder, 2011)

Finally, the interaction and synergism of climate change effects will undoubtedly produce new responses that may prove to be more deleterious than either stressor alone. For example, synergistic stressors such as ocean acidification and hypoxia may narrow thermal windows, depending on species-specific sensitivities (Pörtner and Farrell, 2008). In other cases, multiple stressors could cancel out the negative effects of each other (Crain *et al.*, 2008). Research into the effects of multiple stressors is only beginning but promises to be a fruitful area for a better understanding of the physiology of climate change.

Impacts on Distribution and Abundance

Many conservation and management efforts, including fisheries management programs, are concerned with maintaining the abundance and distribution of populations. A population responds to climate by integrating both the physiological impacts described previously and the changing species interactions and evolutionary forces described later in this chapter. Responses can include changes in rates of growth, reproduction, mortality, dispersal, or timing of major events. Together, these responses determine population abundance and distribution.

Impacts on Abundance

As a starting point for understanding how populations will respond to climate change, we can attempt to scale up from the physiological effects of temperature, acidification, oxygen, and other predicted changes. For example, laboratory studies have shown that eelpouts (*Zoarces viviparus*) are physiologically limited and that the growth rates of individual fish decline above 17 °C (Pörtner and Knust, 2007). These effects can also be detected at the population level. In the Wadden Sea where these fish live, mortality increases and population abundance declines above 17 °C (Pörtner and Knust, 2007), suggesting that further warming will drive this population toward extirpation.

Larvae of marine organisms appear particularly sensitive to environmental conditions, and replenishment may be the first population-level process that declines when conditions worsen. Larval survival is tied to temperature (Pepin, 1991) and CO_2 concentration (Munday *et al.*, 2010), and successful recruitment back to suitable adult habitat often depends on oceanographic conditions (Roughgarden *et al.*, 1988). In coastal upwelling zones of the ocean, including the coast of California, recruitment of species like barnacles is tied to relaxation of the winds that otherwise drive currents away from

the coast (Roughgarden *et al.*, 1988). Changes to upwelling, which are a likely consequence of climate change, are therefore likely to have important impacts on population replenishment by larvae (Harley *et al.*, 2006). In spite of the sensitivity of larvae to climate, however, declines in population abundance will lag further behind because adults can continue to survive in a given location for many years, particularly among long-lived species.

It is also important to realize that changes in growth, mortality, and reproduction can compensate for each other, thereby mitigating the immediate effects of climate change. For example, higher temperatures can negatively affect survival and recruitment, but abundance can nevertheless stay the same if individuals grow to reproductive age more quickly (Doak and Morris, 2010). Care must therefore be taken before directly projecting from physiological studies to population-level impacts without carefully considering the full range of processes that affect populations.

Other population-level impacts difficult to predict from physiological studies are related to phenology, or the seasonal timing of ecological events. The life cycles of fishes are often synchronized to the seasonal production of plankton so that offspring have sufficient food for growth and survival. Research in the North Sea, however, has shown that fish larvae appear earlier in warm years, whereas many lower trophic level species (e.g., some phytoplankton) do not change their timing, likely because they are cued to light levels (Edwards and Richardson, 2004). Mismatches between reproduction of fish and the availability of their planktonic food can cause dramatic decreases in abundance, as has been documented for Atlantic cod (*Gadus morhua*) (Beaugrand *et al.*, 2003). More broadly, populations are affected indirectly by changes to the productivity, structure, and composition of ecosystems that provide food and shelter necessary for growth and survival (see the section "Community Impacts").

One of the most comprehensive efforts to consider the range of climate factors that impinges on population abundance has been undertaken for tuna. A valuable fisheries species, these predators undergo oceanwide migrations every year in search of food and suitable reproductive habitats. By building a population dynamics model on top of global projections for temperature, oxygen, and ocean biogeochemistry, Lehodey and colleagues (2010) predicted that spawning habitats for bigeye tuna (*Thunnus obesus*) would move poleward and toward the eastern tropical Pacific, that access to food would improve in some areas, and that adult survival would decline. Overall, the model predicted stable populations for decades but a declining abundance at the end of the 21st century.

Impacts on Distribution

Even if populations decline in some places, they may flourish in others or colonize new, previously unoccupied habitats as conditions become more favorable. Extinction within "trailing edge" populations or colonization of "leading edge" populations – or a combination of the two – creates range shifts. Despite little attention to range shifts in the ocean during the beginning of climate change research, range shifts appear to be a particularly common and dramatic occurrence in the marine realm (Sorte *et al.*, 2010). Historical surveys have revealed species shifting an average of 19 km yr^{-1} (Sorte *et al.*, 2010), which is an order of magnitude faster than rates observed on land (Parmesan and Yohe, 2003). Species also move deeper as temperatures warm, with rates of 3.6 m per decade measured among North Sea fishes (Perry *et al.*, 2005; Dulvy *et al.*, 2008). Of 129 geographic shifts in marine species that are known so far, 75% are in poleward directions, as would be predicted from widespread warming (Sorte *et al.*, 2010).

On the west coast of North America, however, species such as blue mussels (*Mytilus* spp.) have instead moved toward the equator in recent years (Hilbish *et al.*, 2010). Although at first counterintuitive, this shift is entirely consistent with species ranges that respond to climate. This shift appears to be driven by temporary regional cooling in the California Current ecosystem caused by decadal-scale climate oscillations. This and other examples showing rapid responses of marine species ranges to climatic conditions (e.g., Fisher *et al.*, 2008) suggest that marine species distributions respond quickly and perhaps predictably to climate. Taking this hypothesis to its logical end would predict a massive redistribution of marine species toward the poles and away from the tropics over the coming decades (Cheung *et al.*, 2009). Not all marine species can move quickly in response to climate, however. Some fishes and invertebrates spend very little time dispersing as larvae and move little as adults (Kinlan and Gaines, 2003; Shanks, 2009). Whether these and other species will keep up with climate change remains an important question.

Other researchers have also pointed out a number of factors that make predicting future geographic shifts in species ranges more complicated than simply mapping future temperatures. At one level, local microclimates can diverge substantially from regional averages, so there are no simple latitudinal gradients in temperature for species ranges to follow. For example, intertidal mussels experience hot midday air temperatures only when the tide is low, and the timing of tides on the west coast of North America is such that sites in Washington and Oregon experience more extreme temperatures than many in California (Helmuth *et al.*, 2002). Because adult mussels are sedentary, these microclimates are crucial for determining growth and survival. Local microclimates may be less important in deeper waters where oceanic mixing ensures greater homogeneity, but finding appropriate methods to predict local climates in the ocean remains an area of important research (Stock *et al.*, 2011).

Range shifts for species are also unlikely to be smooth transitions over time that track closely with mean climate conditions. Instead, punctuated transitions occur when extinctions and colonizations are driven by stochastic extreme events of large effect. Although air temperatures have warmed gradually over the past 30 years at Tatoosh Island, the upper distributional limit of an intertidal red alga moved downward in large, punctuated steps (Harley and Paine, 2009). These extirpations at the species' upper limit only occurred during summers when temperatures were high and seas were calm. Without both factors at the same time, no change in distribution occurred. Colonization of new habitat is also likely to be highly stochastic (Clark *et al.*, 2001). This is especially true in the ocean, where organisms rely on highly variable

oceanographic currents to deliver their larvae to new habitats (Siegel *et al.*, 2008).

In addition, species ranges are not determined solely by temperature. Coastal mosaics of ocean acidification are likely to shape the distribution of calcifying species in particular (Feely *et al.*, 2008), whereas changes to the extent of oxygen minimum zones will affect many fishes (Cheung *et al.*, 2011). Changes to runoff from land (driven in part by changes to precipitation), storm frequency, and sea-level rise will also be important for intertidal and near-coastal species.

Finally, ecological processes beyond simple tolerance for climatic conditions will be critical considerations when trying to project future species ranges. From theory and empirical studies, we know that species ranges can be set by currents (Gaylord and Gaines, 2000), habitat availability (Gilman, 2006), or species interactions (Harley, 2003) in addition to climate conditions. Research, however, is needed to understand how important these processes are for predicting future changes to species distributions.

Evolutionary Impacts

The ocean is as varied in environmental conditions as it is vast. In response, marine species have evolved complex mechanisms of survival, physiological tolerance, communication, navigation, and food production and capture. A central evolutionary concern is that the current pace of anthropogenic climate change is likely to be faster than the evolutionary rate of most species. In other words, environmental shifts will likely outpace many species' abilities to adapt evolutionarily to the new conditions. Whether some species in some locations will be able to adapt, however, remains a fascinating question. Adaptations in response to climate change induced shifts in environment have already been found in a handful of terrestrial species (Bradshaw and Holzapfel, 2006; Hoffmann and Sgro, 2011), but similar examples from the marine realm have yet to be uncovered.

Adaptation in a Changing Climate

Traditionally, the theory of natural selection posits that adaptation occurs in slow, successive steps. An increasing body of evidence, however, has shown that rapid, "contemporary evolution" may be more common than previously thought and perhaps can even occur at a pace similar to expected rates of climate change (e.g., Barrett *et al.*, 2010; Marko, 2004; Hoffmann and Sgro, 2011; Hendry and Kinnison, 1999). For example, Barrett *et al.* (2010) found that cold tolerance in marine sticklebacks underwent a rapid evolutionary shift of 2.5 °C over three generations (~27 months) after transplantation to freshwater, one of the most rapid rates of phenotypic evolutionary change ever observed in a natural population.

In contrast, populations with long generation times, low effective population sizes, and low genetic diversity may not possess the flexibility to adapt over such rapid timescales (e.g., Harley *et al.*, 2006). In the case of the sticklebacks, there have been repeated historical transitions between marine and freshwater habitats over past glacial cycles (McKinnon and Rundle, 2002) that are likely to have created a high level of standing genetic diversity in these species (P. Shulte, personal communication). Hence, although natural selection may be acting quickly in this system, the evolutionary material that is under selection may have evolved over many thousands of years. Additional studies of contemporary evolution via experimental manipulations (sensu Barrett *et al.*, 2010), further inference based on historical records (sensu Marko, 2004), and additional contemporary analysis across a variety of taxa will be critical in assessing the capacity of marine organisms to adapt to future climate change more generally.

Local Adaptation

Understanding local adaptations of populations to specific environmental conditions and ecological roles is also an important process to consider when predicting how organisms will respond to future climate change. Equatorial populations of reef-building corals, for example, have higher upper thermal tolerance limits and are genetically differentiated from more southern populations (Smith-Keune and van Oppen, 2006), suggesting that thermal tolerance limits have evolved to match surrounding water temperatures. Local adaptation of upper thermal tolerance limits has also been found in populations of an intertidal snail along the West Coast of the US (Kuo and Sanford, 2009), although in a somewhat different pattern. In the snails, those with greater thermal tolerance limits were found at higher latitudes due to the higher frequency of extreme thermal exposures in northern intertidal locations. This difference between the two studies serves to illustrate the importance of understanding local microclimate and habitat differences and how they relate to environmental selection pressures.

Local adaptation can also have important effects on how genes and phenotypes are distributed in response to climatic changes. If local environmental conditions become more extreme, then populations of organisms locally adapted to historically extreme conditions (e.g., Barshis *et al.*, 2010) may serve as sources of genetic material for locations that only recently shifted to more stressful environmental conditions. Thus, flow of locally adapted genes could provide susceptible populations with the genetic material necessary to adapt to climate change impacts. Alternatively, local adaptation may restrict the ability of certain populations to withstand climate change because genotypes specifically tuned to local environmental conditions may not be capable of surviving additional perturbation caused by global climate change. It also remains difficult to assess whether physiological differences across species ranges are the result of local adaptation or phenotypic plasticity (Sanford and Kelly, 2011), and conservation strategies may differ, depending on the mechanism involved.

Genetic Diversity and Climate Tolerance

Evolutionary theory predicts that peripheral populations in a species' range are likely to contain lower genetic diversity and higher genetic differentiation due to greater distance and smaller effective population size relative to more central

populations (Eckert *et al.*, 2008; Wulff, 1950). This suggests that the populations at the edges of ranges should be less able to adapt to rapid changes in environmental conditions. Indeed, Pearson *et al.* (2009) found that populations at the southern edge of the range of *Fucus serratus*, an intertidal brown alga, showed a greater susceptibility to desiccation and increased heat shock response than their geographically central conspecifics. This was interpreted as increased cellular stress when exposed to heat shock and decreased resilience to environmental stressors. However, Sagarin and Somero (2006) found that levels of heat shock proteins in intertidal mussels and snails showed no correlation with position along the species range, but they found a strong positive correlation with local areas expected to have long daytime exposure during summer months. In this case, the extreme environmental heterogeneity in temperature due to the timing of midday tides may explain why edges of populations did not exhibit higher stress responses as compared to more centrally located populations (e.g., Harley and Helmuth, 2003). The general idea that "edge" populations may be the first to face local extinction must be carefully considered in light of environmental heterogeneity.

Community Impacts

All species are embedded in communities governed by complex webs of species interactions. Although much research about climate change has occurred on a species-by-species basis, it is important to consider how the impacts of climate change on individual species and populations will translate to the broader community and ecosystem. The effects of climate change on a species or population can indirectly affect many other organisms through species interactions, causing cascading impacts on the community structure and function. In addition, interactions among species can either worsen or ameliorate the effects of climate change on a sensitive population. An integrated understanding of the concepts covered in previous sections is needed to predict how marine communities and ecosystems will be impacted by climate change. Differences in physiological tolerance and adaptive capacity among species can cause a restructuring of ecological communities through altered composition and interactions among species. At a global scale, these changes can cause shifts in the productivity, trophic structure, and resilience of marine ecosystems.

One of the clearest impacts of climate change on a community will occur if habitat-forming species are affected. Habitat loss caused by climate change will also result in the loss of the many species reliant on these habitats. For instance, corals are the biogenic habitat formers of coral reef ecosystems, and repeated bleaching events associated with warm temperatures cause both coral mortality and local extirpations of species that rely on live corals (e.g., Graham *et al.*, 2006; Knowlton, 1992). A sharp decline in corallivorous and planktivorous fish that rely on live coral, for example, has been reported shortly after bleaching events (Graham *et al.*, 2006).

Furthermore, climate change is predicted to cause a shift from branching corals with high structural complexity to massive and encrusting corals due to these species' differential susceptibility to climatic stressors (Loya *et al.*, 2001). A reduction in structural complexity increases competition for remaining habitat and increases predation on small coral reef fishes that use the reef for protection, both of which lead to a dramatic decline in fish abundances (Graham *et al.*, 2006). In addition to coral reefs, many marine communities are also built around biogenic habitat formers that have shown sensitivity to increased temperature or acidification, including kelp forests (Schiel *et al.*, 2004), mangroves (Alongi, 2008), and oyster reefs (Talmage and Gobler, 2011). Sea ice is also an important habitat for numerous species, from zooplankton to polar bears (Atkinson *et al.*, 2004). Much like the response of sensitive biogenic habitat formers, the melting of sea ice will result in a loss of the species reliant on this habitat (Atkinson *et al.*, 2004).

Climate change can also have cascading impacts on the community via its effects on other important species, besides those important to habitat formation. For example, many marine communities have strong trophic cascades where a single species plays a disproportionately strong role in structuring the abundance, distribution, and behavior of other species (Shurin *et al.*, 2002; Paine, 1966). A keystone predator in rocky intertidal communities, the seastar *Pisaster ochraceus*, is known to set the lower distributional limit of mussels (Paine, 1966). Ocean acidification and increased temperatures increase the growth and feeding rate of this species, which in turn reduces the abundance of mussels and the many associated species that grow within mussel beds (Sanford, 1999; Gooding *et al.*, 2009).

In contrast, climate change could in some cases shift the competitive balance away from the keystone or competitive dominant species. For example, ocean acidification enriches macroalgal growth at the expense of a competitively dominant coral species *Acropora* sp. (Diaz-Pulido *et al.*, 2011). This could drive a shift from coral-dominated to algal-dominated reef ecosystems, which would have cascading impacts on a multitude of coral reliant species (Mumby *et al.*, 2007; Anthony *et al.*, 2011).

Much attention has also been given to environmentally driven changes in oceanic primary productivity (the production of biomass through photosynthesis) and how these changes will travel through food webs with climate change (Brander, 2007). Net primary productivity is affected by the availability of light and nutrients, which are in turn affected by dust, runoff, ocean mixing, cloud cover, and the solar cycle. Sea surface warming causes increased stratification of the ocean with warmer waters on top and cooler water below. This process prevents the nutrient-rich deep waters from mixing into the upper, light-rich zone where photosynthesis can occur (Sarmiento *et al.*, 2004). The decrease in global productivity from 1997 to 2006 was associated with warmer surface waters, particularly in low-latitude, stratified oceans (Behrenfeld *et al.*, 2006), suggesting there will be lower productivity with climate change. However, patterns in productivity differ geographically (Sarmiento *et al.*, 2004; Richardson and Schoeman, 2004). For example, phytoplankton abundance increased from 1958 to 2002 as waters warmed in the northeast Atlantic where nutrients were abundant but decreased in southern regions where nutrients were more rare (Richardson and Schoeman, 2004), suggesting

productivity hot spots may shift with climate change. Reductions in productivity and shifts in the location of productivity hot spots have important implications for communities and ecosystems, including important fishery species. For example, reductions in the recruitment of Atlantic cod (*Gadus morhua*) and abundance of Atlantic salmon (*Salmo salar*) have been correlated with climate-associated changes in the plankton community (Beaugrand *et al.*, 2003; Beaugrand and Reid, 2003).

In addition to changes in productivity, the trophic structure of marine food webs might shift toward stronger control by consumers in a warmer ocean (Moran *et al.*, 2010; O'Connor *et al.*, 2009). The metabolic theory of ecology predicts that as temperatures increase, metabolism and feeding rates of consumers will increase. A consequence may be that plant–herbivore interaction strengths will also increase. Plant–herbivore interaction strengths are stronger in warmer conditions (Pennings and Silliman, 2005), suggesting there may be stronger consumer control of marine ecosystems under conditions of climate change (López-Urrutia *et al.*, 2006; O'Connor *et al.*, 2009; Moran *et al.*, 2010). Stronger consumption could offset any increased primary productivity associated with warmer, well-mixed waters in nearshore ecosystems.

In addition to the strong effects of changes to primary productivity or keystone species, numerous weak interactions among the hundreds of other species can influence the community (Berlow, 1999). Because of the complexity of these interactions, it can be difficult to predict the community and ecosystem impacts of climate change from single species experiments. Large mesocosm and natural experiments are especially insightful in these instances. For example, volcanic carbon dioxide vents off the coast of Italy cause changes in the carbonate chemistry of the seawater similar to those predicted with climate change and ocean acidification (Hall-Spencer *et al.*, 2008). The benthic rocky reef community surrounding the vents suggests that, in extremely acidified conditions, the diverse assemblage of calcifying species (e.g., urchins, coralline algae, and mollusks) could be replaced by low-diversity communities that lack important calcifying species and that are dominated by fleshy algae and a few invertebrate species (Hall-Spencer *et al.*, 2008). Decreased abundance of calcifying species has also been reported in mesocosm investigations of acidification on coral reef ecosystems (Kuffner *et al.*, 2008).

Range shifts of marine populations will also influence community composition through species introductions and local extirpations (Barry *et al.*, 1995; Southward *et al.*, 1995). If the invading species are strong competitors or predators, species introductions can significantly alter community and ecosystem function (Zeidberg and Robison, 2007; O'Connor and Crowe, 2005). For example, the sea urchin *Centrostephanus rodgersii* (Diadematidae) has recently expanded its range poleward from mainland Australia to Tasmania. *C. rodgersii* is known for causing large barrens and sharp reductions in biodiversity in kelp forests. The increasing rate of species introductions associated with climate-driven range shifts will create novel communities, and the consequences remain difficult to predict.

At a broader level, ecological communities function as complex adaptive systems that have a certain degree of inherent resilience (Levin and Lubchenco, 2008), but climate change may overcome this resilience. Generally, resilience is the ability of an ecosystem to maintain its function despite disturbance events, whereas loss of resilience is often associated with abrupt shifts between dramatically different states (Gunderson, 2000; Scheffer *et al.*, 2001; Levin and Lubchenco, 2008). Resilience is often considered to have two components: resistance to change and ability to recover (Levin and Lubchenco, 2008). In ecosystems, these abilities are provided in part by a combination of diversity and redundancy among the components that make up that ecosystem. For example, a diversity of herbivorous fishes and sea urchins graze on algae and prevent coral reefs from being overgrown. Overfishing in parts of the Caribbean, however, created reef ecosystems with low redundancy because, without herbivorous fishes, only urchins were left to feed on algae. The result, as shown dramatically in Jamaica in the 1980s, was that the coral reefs were transformed into algal fields when a disease swept through urchin populations and removed the last effective herbivore (Hughes, 1994). More recent work has shown that a lack of herbivorous fishes allows algae to overgrow reefs and prevent corals from regrowing after catastrophic bleaching driven by high temperatures, whereas the presence of fishes allows the reef to recover (Hughes *et al.*, 2007).

A critical question is whether or where marine communities will be resilient to the impacts of climate change. In the face of a sustained stressor such as climate change, the individual components that make up marine ecosystems will change, but will this alter the overall ecosystem function? For example, many species will be extirpated while new species will colonize local ecosystems as their distributions generally move poleward. If replacement species fill similar ecological roles (e.g., herbivores), the entire ecosystem may prove generally resilient and be able to continue providing the types of goods and services that humans want (Levin and Lubchenco, 2008). Where and under which conditions this is likely to be true remains an open area of investigation. In general, however, ecosystems with a larger number of species appear to be more resilient to disturbance, perhaps because species respond uniquely to disturbances and, in systems where many species fill similar ecological roles, key roles can be maintained even as some species decline (Worm *et al.*, 2006).

Conservation Strategies

The ultimate conservation strategy for marine ecosystems faced by climate change is a rapid and dramatic reduction in greenhouse gas emissions. However, at this point, the ocean will become warmer and more acidic and sea level will rise due to climatic inertia alone, even if action is taken now to reduce greenhouse gas emissions. Therefore, this section focuses on what can be done in terms of local or regional mitigation and adaptation to the effects of climate change.

Conservation of marine ecosystems requires considering the range of factors that species and communities need to persist, from growth and survival to dispersal, reproduction, and preservation of ecosystem function. Given the effects of climate change we expect, conservation strategies can broadly

aim to protect species and assemblages as they currently exist, plan ahead for likely shifts in location and community composition, or attempt to do both. It is important to remember, however, that climate conditions are likely to become unfavorable for certain species and populations in certain locations, and conservation of the current status quo will not be possible everywhere. When and where to focus on mitigating the impacts of climate change on existing species versus planning for and facilitating a transition to new communities of species remains an important question for conservation planning.

One approach is to focus conservation on broad ecosystem properties such as diversity and resilience rather than particular species. In a recent review, Worm and Lotze (2009) argue that high biodiversity can provide resilience against climate change impacts and needs to be an important consideration in management actions. In systems with higher biodiversity, species can exhibit a wider range of responses to climate change. This range both increases the probability that at least some species will continue performing important ecosystem functions and reduces the overall variability of the ecosystem by averaging the responses across many different species (the portfolio effect). Alleviation of other, more targeted stressors such as overfishing, land-based pollution, and destructive resource uses is one strategy to manage for resilience because it protects the redundancy of ecosystem functions and enhances the ability of a given ecosystem to tolerate or bounce back from the impacts of climate change. Flexibility and active experimentation with conservation strategies, termed "adaptive management," will also be key to ensuring effective marine conservation as new recommendations are generated through research.

Prioritizing Species for Research and the Role of Predictive Modeling

Can we predict which fishes and invertebrates will be the climate change "winners" and "losers"? If so, how can marine managers incorporate this information into policies to protect marine biodiversity? Given limited resources, identification of priority species for climate change research is key to making informed management decisions. Efforts can focus on species that are already known to be vulnerable (e.g., species living close to their environmental limits), species that play key ecological roles such as keystone or foundation species (e.g., corals), and key indicator species that have easily studied and monitored traits relevant to climate change.

Species' capacity for physiological plasticity and genetic adaptation in response to climate change can also be incorporated into predictive population or community-level models (Chevin *et al.*, 2010). These models can help predict which conservation strategies might be most effective for a particular species or group of species. For example, it has been proposed that reefs with more diverse coral communities should be protected because high diversity is an indicator of ecosystem health and high functional redundancy. In contrast, it has also been proposed that less-diverse communities with a high prevalence of thermally tolerant coral species and symbiotic algae should be protected because these may have greater tolerance of climate change (West and Salm, 2003). Baskett *et al.* (2010) used a theoretical community-level model that included adaptation and acclimation to show that protection of more diverse coral communities would be more effective for long-term maintenance of coral cover than protection of communities with high abundances of thermally tolerant corals. These types of ecophysiological models can be combined with empirical and observational assessments to help conservation managers decide how and where resources can best be allocated (Dawson *et al.*, 2011).

Marine Protected Areas

One common conservation approach is the use of marine protected areas (MPAs). MPAs are areas of the ocean set aside with limited or no fishing and other uses. Although MPAs do not directly address the threats from climate change, they promote biodiversity and ecosystem resilience in the face of climate change because they reduce some of the other stressors on marine populations (McLeod *et al.*, 2009). For example, herbivorous fish maintain coral-dominated ecosystems by grazing competitive macroalgae. Maintenance of healthy herbivorous fish communities in MPAs can increase the resilience of the coral reef ecosystem by grazing macroalgae and thereby indirectly promoting coral recruitment and growth after an extreme climate event (Mumby *et al.*, 2007).

Planning MPAs for climate change is similar to planning for general conservation, and it is important to consider the principles of representation, replication, and spacing (McLeod *et al.*, 2009). Representation refers to protecting the full range of habitat types within an ecosystem, whereas replication ensures that a catastrophe will not destroy the only protected example of that habitat. Together, representation and replication accomplish the goal of spreading the risk of unexpected events, including extreme climatic events (McLeod *et al.*, 2009). In terms of climate change, it is important that replicated MPAs exist in different climate regimes such that they will not all be hit by the same extreme climate events. After such events, the unaffected MPAs can be important sources of larvae to allow recovery in the affected areas. MPA spacing ensures that protected populations are close enough together so as to exchange larvae, maintain viable metapopulations through larval dispersal, and allow for recolonization after disturbances. General guidelines of 15–20 km spacing or less have been suggested to allow sufficient larval exchange for many fishes (McLeod *et al.*, 2009). Ensuring that MPAs are close together is likely to become increasingly important with climate change because increasing temperatures cause larvae to develop more quickly and to disperse shorter distances (O'Connor *et al.*, 2007).

Although the preding guidelines and considerations are important for MPAs everywhere, a fundamental question remains: where should MPAs be located to promote population persistence in the face of climate change? For this question, it is important to consider not only which populations are likely to persist if protected, but also where those populations will be in the future. This requires an understanding of physiological tolerance limits, population dynamics, community structure, and evolutionary history. Coral reefs in particular

have received substantial attention given the near-term threat posed by climate. Mumby and colleagues (2011) present an integrative modeling approach that selects MPA sites across the Bahamas based on both thermal stress resistance and likely evolutionary scenarios. They find that although it is difficult to pick MPAs that hedge against all potential evolutionary outcomes, 15% of sites were selected in all scenarios, making them high-priority targets for early protection (Mumby *et al.*, 2011). If resources are initially focused on these high-priority sites, effort will be maximized until future research and monitoring sheds light on which additional populations are likely to survive.

Given the rates of geographic range shifts already observed (*see* Impacts on Distribution, above), future shifts should also be considered in MPA design. For short-term shifts, placing MPAs across relevant environmental gradients will allow populations to move while still staying within an MPA (M. Carr, personal communication). For example, species are likely to move deeper and poleward, and so arranging MPAs to encompass a range of depths and along latitudinal gradients will allow for future shifts. Gradients of wave exposure and upwelling may also be important given projected changes in storms and upwelling intensity and the strong role these factors play in shaping coastal marine communities. Finally, positioning MPAs near undeveloped coastlines will allow intertidal areas to move upward and inland as sea levels rise.

Ecosystem-Based Management

Whereas the implementation of MPAs involves designating specific areas as protected, ecosystem-based management represents an effort to consider and manage all the human activities affecting an ecosystem (McLeod and Leslie, 2009). Such an approach is especially important for mitigating threats that spread easily throughout and between ecosystems, including land runoff and pollution. For example, local ocean acidification can be reduced by preventing the runoff of acidic fertilizers and excess nutrients from land and by reducing the discharge of sulfur dioxide into the air (Kelly *et al.*, 2011).

One form of ecosystem-based management is called "marine spatial planning," and it involves designation of ocean zones for full development, limited use, conservation, or full protection (Crowder and Norse, 2008). From a conservation perspective, the principles just described for MPAs still apply to marine spatial planning, including replication, representation, size, and spacing of conservation zones and general ecosystem-based management of broader threats (Crowder and Norse, 2008).

Marine spatial planning also introduces new considerations, however, because some full development zones may involve building semipermanent structures. These include oil and gas drilling, mining operations, and wind and wave energy. Once built, many of these structures will last for decades, which is sufficient time for climate change to induce substantial alterations in the marine ecosystem. For example, prime habitats for species could move 100 km north in only a few decades (see the preceding description of range shifts), which would move the ideal locations for both conservation and fishing zones and potentially conflict with existing development. Considering these future shifts proactively within the framework of marine spatial planning remains an area for further attention.

Fisheries Management

A sustainable fishery requires a persistent and viable population in the wild. The level of fishing that can be maintained depends on the productivity of the population, which is a function of growth rate, reproduction, and natural mortality. These are also the population factors affected by climate, and the level of fishing that is sustainable is therefore likely to change with climate change (Brander, 2007). Reductions in fishing mortality will be necessary for some populations, whereas increases will be possible in others. Determining where, when, and how much to alter fisheries regulations, however, currently presents a major challenge for fisheries science (Peterman, 2009).

As one example of what may happen to future fisheries, a coupled climate-population model for Atlantic croaker (*Micropogonias undulatus*) in the northeastern US suggested that yield from the species could increase 30–100% over the coming century (Hare *et al.*, 2010). This model was developed for a region near the species' northern range limit, where such effects might be expected. In contrast, one would expect declining fisheries toward the species' southern range limit. One remaining uncertainty, however, is how much the production of plankton (the source of food ultimately available to fish) will change or even in which direction at a regional level (Sherman *et al.*, 2009; Cheung *et al.*, 2011). No matter what the direction of change, the lesson is clear: past experience will no longer be a good guide to sustainable levels of fishing in the future (Peterman, 2009).

Climate-driven shifts to the geographic ranges of fished species also pose new challenges for regional and international fisheries management (Cheung *et al.*, 2009). As populations move poleward, they are likely to become inaccessible to some fishing ports but move within reach of others. When these ports are in different management regimes, fisheries management may not be ready for the change. The so-called Mackerel Wars of 2010 underscored the fact that such problems are not a distant prediction (Anonymous, 2010). Warming climate has shifted the mackerel fishery north from Britain and into Icelandic waters, sparking an international debate about the fisheries quota that should be allocated to each country, and, in the meantime, allowing more fish to be taken from the population than is likely to be sustainable.

This example demonstrates an important fact: fisheries management is about managing people, not fish. Ensuring a sustainable fishery requires transparent and inclusive institutions that can enforce limits on exploitation, engage stakeholders, and design incentives that align public and private benefits (Grafton *et al.*, 2008). In a changing climate, the need for these institutions is even greater given uncertainty about the magnitude and timing of climatic impacts and the likelihood of difficult transitions within the fishing industry as the available species change. Under traditional fisheries management, climate is rarely considered (Hare *et al.*, 2010; Hilborn and Walters, 1992), but this will have to change rapidly so that

institutions and mechanisms for adjusting fisheries management can be in place before the impacts of climate change become more pronounced. Fisheries science has tried to develop management strategies to deal with climate variability, including reference points, fixed decision rules, and climate regime–specific harvest rates, but implementing them effectively requires close monitoring of both fish populations and the ecosystems they inhabit (King and McFarlane, 2006; Walters and Parma, 1996). Even when impacts cannot be predicted ahead of time, fisheries management can be ready to both detect climate-driven changes and respond with new fisheries management when new conditions require doing so.

Assisted Colonization and Restoration

For some marine species – particularly those most sensitive to climate and those with poor dispersal abilities, such as corals – climate conditions may change much faster than the their ability to adapt or disperse to more favorable conditions. In these situations, actively moving the species to more favorable locations may be warranted but should only be undertaken after careful consideration of the costs and benefits. "Assisted colonization," as it is called, may be a reasonable solution when the risk of extinction due to climate change is high, the translocation of the species is technically feasible, and the probability of success outweighs the ecological and socioeconomic costs (Hoegh-Guldberg *et al.*, 2008). The history of marine invasive species provide cautionary tales about needlessly introducing species into new ecosystems (Ruiz *et al.*, 1997). However, introducing species beyond their geographic range but within the same biogeographic province is likely to have fewer consequences because many of the same community members will still be present in the new location (Hoegh-Guldberg *et al.*, 2008). In cases where either assisted or natural colonization is hampered by lack of habitat (e.g., coral reefs for many reef species), construction of artificial habitats may be a temporary solution in the near term (Hoegh-Guldberg *et al.*, 2008).

Research Needs

There is little doubt that effective conservation strategies aimed at mitigating climate change impacts need to be based on sound, robust research. Yet many gaps remain as to how specific species and assemblages will react to climatic shifts, whether multiple environmental stressors will have synergistic effects, how quickly populations can adapt and acclimatize to rapid environmental change, and how best to allocate resources given these uncertainties. Future research needs to focus on filling in these gaps across a wide variety of species in order for conservation to be effective.

For instance, it is likely that increased ocean acidification will have significant impacts, particularly on calcifying organisms, yet long-term studies of response to individual climate change impacts (e.g., ocean acidification, global warming), as well as response to multiple stressors combined, are lacking in the literature (Hofmann *et al.*, 2010; Brown and Cossins, 2011). Similarly, very few studies have attempted to examine long-term capacity for acclimation and rates of adaptation in the face of climate change (Dawson *et al.*, 2011). In addition, although the need to move beyond predictions is very apparent, experimental evolution research is impractical for nearly all slow-growing, long-lived species. In these cases, careful parameterization of models that consider evolution and phenotypic plasticity can provide valuable insight into the potential responses of marine species in a changing climate (Baskett *et al.*, 2010). Lastly, further efforts to integrate our knowledge of physiology, population biology, evolution, and community interactions into conservation strategies are needed to ensure that present-day efforts remain effective in a rapidly changing climate.

Colonization

List of Courses

1. Climate: Past, Present, and Future
2. Climate Change: Scientific Basis
3. Marine Biology
4. Conservation Biology
5. Ocean, Atmosphere, and Climate
6. Sustainability Science
7. Ecosystems and Global Change

See also: Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. Climate Change and Ecology, Synergism of. Climate Change and Extinctions. Climate Change and Wild Species. Climate, Effects of. Conservation Efforts, Contemporary. Conservation Genetics. Evolution in Response to Climate Change. Extinction, Causes of. Fish Conservation. Human Impact on Biodiversity, Overview. Latitudinal and Elevational Range Shifts under Contemporary Climate Change. Loss of Biodiversity, Overview. Marine Ecosystems. Marine Protected Areas: Static Boundaries in a Changing World. Resource Exploitation, Fisheries. Role and Trends of Protected Areas in Conservation. Species Distribution Modeling. Translocation as a Conservation Strategy

References

Alongi DM (2008) Mangrove forests: Resilience, protection from tsunamis, and responses to global climate change. *Estuarine Coastal and Shelf Science* 76: 1–13.http://dx.doi.org/10.1016/j.ecss.2007.08.024.

Anonymous (2010) Mackerel wars: Overfished and over there. *The Economist*, 4 September 2010.

Anthony KRN, Maynard JA, Diaz-Pulido G, *et al.* (2011) Ocean acidification and warming will lower coral reef resilience. *Global Change Biology* 17: 1798–1808.http://dx.doi.org/10.1111/j.1365-2486.2010.02364.x.

Atkinson A, Siegel V, Pakhomov E, and Rothery P (2004) Long-term decline in krill stock and increase in salps within the Southern Ocean. *Nature* 432: 100–103.http://dx.doi.org/10.1038/nature02950.

Barrett RDH, Paccard A, Healy TM, *et al.* (2010) Rapid evolution of cold tolerance in stickleback. *Proceedings of the Royal Society B: Biological Sciences* 233–238.

Barry JP, Baxter CH, Sagarin RD, and Gilman SE (1995) Climate-related, long-term faunal changes in a California rocky intertidal community. *Science* 267: 672–675.

Barshis DJ, Stillman JH, Gates RD, Toonen RJ, Smith LW, and Birkeland C (2010) Protein expression and genetic structure of the coral *Porites lobata* in an environmentally extreme Samoan back reef: does host genotype limit phenotypic plasticity? *Molecular Ecology* 19: 1705–1720.http://dx.doi.org/10.1111/j.1365-294X.2010.04574.x.

Baskett ML, Nisbet RM, Kappel CV, Mumby PJ, and Gaines SD (2010) Conservation management approaches to protecting the capacity for corals to respond to climate change: A theoretical comparison. *Global Change Biology* 16: 1229–1246.http://dx.doi.org/10.1111/j.1365-2486.2009.02062.x.

Beaugrand G, Brander KM, Alistair Lindley J, Souissi S, and Reid PC (2003) Plankton effect on cod recruitment in the North Sea. *Nature* 426: 661–664.http://dx.doi.org/10.1038/nature02164.

Beaugrand G and Reid PC (2003) Long-term changes in phytoplankton, zooplankton and salmon related to climate. *Global Change Biology* 9: 801–817.

Behrenfeld MJ, Malley RTO, Siegel DA, *et al.* (2006) Climate-driven trends in contemporary ocean productivity. *Nature* 444: 752–755.

Berlow EL (1999) Strong effects of weak interactions in ecological communities. *Nature* 398: 330–334.

Bradshaw WE and Holzapfel CM (2006) Evolutionary response to rapid climate change. *Science* 312: 1477–1478.http://dx.doi.org/10.1126/science.1127000.

Brander KM (2007) Global fish production and climate change. *Proceedings of the National Academies of Science* 104: 19709–19714.

Brierley AS and Kingsford MJ (2009) Impacts of climate change on marine organisms and ecosystems. *Current Biology* 19: R602–R614.

Brown BE and Cossins AR (2011) The potential for temperature acclimatisation of reef corals in the face of climate change. In: Dubinsky Z and Stambler N (eds.) *Coral Reefs: An Ecosystem in Transition*, pp. 421–433. Berlin: Springer.

Burrows MT, Schoeman DS, Buckley LB, *et al.* (2011) The pace of shifting climate in marine and terrestrial ecosystems. *Science* 334: 652–655.http://dx.doi.org/10.1126/science.1210288.

Calosi P, Bilton DT, and Spicer JI (2008) Thermal tolerance, acclimatory capcity and vulnerability to global climate change. *Biology Letters* 4: 99–102.

Cheung WWL, Dunne J, Sarmiento JL, and Pauly D (2011) Integrating ecophysiology and plankton dynamics into projected maximum fisheries catch potential under climate change in the Northeast Atlantic. *ICES Journal of Marine Science* 1–11.http://dx.doi.org/10.1093/icesjms/fsr012.

Cheung WWL, Lam VWY, Sarmiento JL, Kearney K, Watson R, and Pauly D (2009) Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries* 10: 235–251.

Chevin L-M, Lande R, and Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biology* 8: e1000357.

Clark JS, Lewis MA, and Horvath L (2001) Invasion by extremes: Population spread with variation in dispersal and reproduction. *The American Naturalist* 157: 537–554.

Cohen AL and Holcomb M (2009) Why corals care about ocean acidification: Uncovering the mechanism. *Oceanography* 22: 118–127.

Crain CM, Kroeker K, and Halpern BS (2008) Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology Letters* 11: 1304–1315.http://dx.doi.org/10.1111/j.1461-0248.2008.01253.x.

Crowder L and Norse E (2008) Essential ecological insights for marine ecosystem-based management and marine spatial planning. *Marine Policy* 32: 772–778.http://dx.doi.org/10.1016/j.marpol.2008.03.012.

Dawson TP, Jackson ST, House JI, Prentice JC, and Mace GM (2011) Beyond predictions: Biodiversity conservation in a changing climate. *Science* 332: 53.http://dx.doi.org/10.1126/science.1200303.

Diaz-Pulido G, Gouezo M, Tilbrook B, Dove S, and Anthony KRN (2011) High CO_2 enhances the competitive strength of seaweeds over corals. *Ecology Letters* 14: 156–162.http://dx.doi.org/10.1111/j.1461-0248.2010.01565.x.

Doak DF and Morris WF (2010) Demographic compensation and tipping points in climate-induced range shifts. *Nature* 467: 959–962.

Donelson JM, Munday PL, Mccormick MI, and Nilsson GE (2010) Acclimation to predicted ocean warming through developmental plasticity in a tropical reef fish. *Global Change Biology* 17: 1712–1719.

Dulvy NK, Rogers SI, Jennings S, Stelzenmller V, Dye SR, and Skjoldal HR (2008) Climate change and deepening of the North Sea fish assemblage: A biotic indicator of warming seas. *Journal of Applied Ecology* 45: 1029–1039.http://dx.doi.org/10.1111/j.1365-2664.2008.01488.x.

Dupont S, Ortega-Martinez O, and Thorndyke M (2010) Impact of near-future ocean acidification on echinoderms. *Ecotoxicology* 19: 449–462.

Eckert CG, Samis KE, and Lougheed SC (2008) Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology* 17: 1170–1188.

Edwards M and Richardson AJ (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430: 881–884.http://dx.doi.org/10.1038/nature02808.

Erez J, Reynaud S, Silverman J, Schneider K, and Allemand D (2011) Coral calcification under ocean acidification and global change. In: Dubinsky Z and Stambler N (eds.) *Coral Reefs: An Ecosystem in Transition*, pp. 151–176. Berlin: Springer.

Fabry VJ, Seibel BA, Feely RA, and Orr JC (2008) Impacts of ocean acidification on marine fauna and ecosystem processes. *ICES Journal of Marine Science* 65: 414–432.

Feely RA, Sabine CL, Hernandez-Ayon JM, Ianson D, and Hales B (2008) Evidence for upwelling of corrosive "acidified' water onto the continental shelf. *Science* 320: 1490–1492.http://dx.doi.org/10.1126/science.1155676.

Fields PA, Graham JB, Rosenblatt RH, and Somero GN (1993) Effects of expected global climate change on marine faunas. *Trends in Ecology and Evolution* 8: 361–367.

Fisher JAD, Frank KT, Petrie B, Leggett WC, and Shackell NL (2008) Temporal dynamics within a contemporary latitudinal diversity gradient. *Ecology Letters* 11: 883–897.

Fuller A, Dawson T, Helmuth B, Hetem RS, Mitchell D, and Maloney SK (2010) Physiological mechanisms in coping with climate change. *Physiological and Biochemical Zoology* 83: 713–720.

Gaylord B and Gaines SD (2000) Temperature or transport? Range limits in marine species mediated solely by flow. *The American Naturalist* 155: 769–789.

Gilman SE (2006) The northern geographic range limit of the intertidal limpet *Collisella scabra*: A test of performance, recruitment, and temperature hypotheses. *Ecography* 29: 709–720.

Gooding RA, Harley CDG, and Tang E (2009) Elevated water temperature and carbon dioxide concentration increase the growth of a keystone echinoderm. *Proceedings of the National Academy of Sciences of the United States of America* 106: 9316–9321.http://dx.doi.org/10.1073/pnas.0811143106.

Grafton RQ, Hilborn R, Ridgeway L, *et al.* (2008) Positioning fisheries in a changing world. *Marine Policy* 32: 630–634.http://dx.doi.org/10.1016/j.marpol.2007.11.003.

Graham NAJ, Wilson SK, Jennings S, Polunin NVC, Bijoux JP, and Robinson J (2006) Dynamic fragility of oceanic coral reef ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 103: 8425–8429.

Guinotte JM and Fabry VJ (2008) Ocean acidification and its potential effects on marine ecosystems. *Annals of the New York Academy of Sciences* 1134: 320–342.

Gunderson LH (2000) Ecological resilience – in theory and application. *Annual Review of Ecology and Systematics* 31: 425–439.

Hall-Spencer JM, Rodolfo-Metalpa R, Martin S, *et al.* (2008) Volcanic carbon dioxide vents show ecosystem effects of ocean acidification. *Nature* 454: 96–99.http://dx.doi.org/10.1038/nature07051.

Hare JA, Alexander MA, Fogarty MJ, Williams EH, and Scott JD (2010) Forecasting the dynamics of a coastal fishery species using a coupled climate-population model. *Ecological Applications* 20: 452–464.

Harley CDG (2003) Abiotic stress and herbivory interact to set range limits across a two-dimensional stress gradient. *Ecology* 84: 1477–1488.

Harley CDG and Helmuth BST (2003) Local- and regional-scale effects of wave exposure, thermal stress, and absolute versus effective shore level on patterns of intertidal zonation. *Limnology and Oceanography* 48: 1498–1508.

Harley CDG, Hughes AR, Hultgren KM, *et al.* (2006) The impacts of climate change in coastal marine systems. *Ecology Letters* 9: 228–241.

Harley CDG and Paine RT (2009) Contingencies and compounded rare perturbations dictate sudden distributional shifts during periods of gradual climate change. *Proceedings of the National Academy of Sciences of the United States of America* 106: 11172–11176.http://dx.doi.org/10.1073/pnas.0904946106.

Helmuth B, Harley CDG, Halpin PM, O'Donnell M, Hofmann GE, and Blanchette CA (2002) Climate change and latitudinal patterns of intertidal thermal stress. *Science* 298: 1015–1017.

Hendry AP and Kinnison MT (1999) The pace of modern life: Measuring rates of contemporary microevolution. *Evolution* 53: 1637–1653.

Hilbish TJ, Brannock PM, Jones KR, Smith AB, Bullock BN, and Wethey DS (2010) Historical changes in the distributions of invasive and endemic marine invertebrates are contrary to global warming predictions: The effects of decadal climate oscillations. *Journal of Biogeography* 37: 423–431.

Hilborn R and Walters CJ (1992) *Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty.* Boston: Kluwer Academic Publishers.

Hoegh-Guldberg O, Hughes L, McIntyre S, *et al.* (2008) Assisted colonization and rapid climate change. *Science* 321: 246–345.

Hoffmann AA and Sgro CM (2011) Climate change and evolutionary adaptation. *Nature* 470: 479–485.http://dx.doi.org/10.1038/nature09670.

Hofmann GE, Barry JP, Edmunds PJ, *et al.* (2010) The effect of ocean acidification on calcifying organisms in marine ecosystems: An organism-to-ecosystem perspective. *Annual Review of Ecology, Evolution, and Systematics* 41: 127–147.

Hofmann GE and Todgham AE (2010) Living in the now: Physiological mechanisms to tolerate a rapidly changing environment. *Annual Review of Physiology* 72: 127–145.

Hughes TP (1994) Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science* 265: 1547–1551.

Hughes TP, Rodrigues MJ, Bellwood DR, *et al.* (2007) Phase shifts, herbivory, and the resilience of coral reefs to climate change. *Current Biology* 17: 360–365.http://dx.doi.org/10.1016/j.cub.2006.12.049.

IPCC (2007) Climate change 2007: Synthesis report. In: Pachauri RK and Reisinger A (eds.) *Contribution of Working Groups I, II, and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.

King JR and McFarlane GA (2006) A framework for incorporating climate regime shifts into the management of marine resources. *Fisheries Management and Ecology* 13: 93–102.

Kelly RP, Foley MM, Fisher WS, *et al.* (2011) Mitigating local causes of ocean acidification with existing laws. *Science* 332: 1036–1037.

Kinlan BP and Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: A community perspective. *Ecology* 84: 2007–2020.

Knowlton N (1992) Thresholds and multiple stable states in coral-reef community dynamics. *American Zoologist* 32: 674–682.

Kuffner IB, Andersson AJ, Jokiel PL, Rodgers KS, and Mackenzie FT (2008) Decreased abundance of crustose coralline algae due to ocean acidification. *Nature Geoscience* 1: 114–117.http://dx.doi.org/10.1038/ngeo100.

Kuo E and Sanford E (2009) Geographic variation in the upper thermal limits of an intertidal snail: Implications for climate envelope models. *Marine Ecology Progress Series* 388: 137–146.

Lehodey P, Senina I, Sibert JR, *et al.* (2010) Preliminary forecasts of Pacific bigeye tuna population trends under the A2 IPCC scenario. *Progress in Oceanography* 86: 302–315.

Levin SA and Lubchenco J (2008) Resilience, robustness, and marine ecosystem-based management. *Bioscience* 58: 27–32.http://dx.doi.org/10.1641/B580107.

López-Urrutia Á., San Martin E, Harris RP, and Irigoien X (2006) Scaling the metabolic balance of the oceans. *Proceedings of the National Academy of Sciences of the United States of America* 103: 8739–8744.

Loya Y, Sakai K, Yamazato K, Nakano Y, Sambali H, and van Woesik R (2001) Coral bleaching: The winners and the losers. *Ecology Letters* 4: 122–131.

Marko PB (2004) 'What's larvae got to do with it?' Disparate patterns of post-glacial population structure in two benthic marine gastropods with identical dispersal potential. *Molecular Ecology* 13: 597–611.

McKinnon JS and Rundle HD (2002) Speciation in nature: The threespine stickleback model systems. *Trends in Ecology & Evolution* 17: 480–488.

McLeod E, Salm R, Green A, and Almany J (2009) Designing marine protected area networks to address the impacts of climate change. *Frontiers in Ecology and the Environment* 7: 362–370.

McLeod KL and Leslie HM (2009) *Ecosystem-Based Management for the Oceans*. Washington, DC: Island Press.

Moran ER, Reynolds PL, Ladwig LM, O'Connor MI, Long ZT, and Bruno JF (2010) Predation intensity is negatively related to plant species richness in a benthic marine community. *Marine Ecology Progress Series* 400.http://dx.doi.org/10.3354/meps08406.

Mumby PJ, Elliott IA, Eakin CM, *et al.* (2011) Reserve design for uncertain responses of coral reefs to climate change. *Ecology Letters* 14: 132–140.http://dx.doi.org/10.1111/j.1461-0248.2010.01562.x.

Mumby PJ, Hastings A, and Edwards HJ (2007) Thresholds and the resilience of Caribbean coral reefs. *Nature* 450: 98–101.http://dx.doi.org/10.1038/nature06252.

Munday PL, Dixson DL, McCormick MI, Meekan M, Ferrari MCO, and Chivers DP (2010) Replenishment of fish populations is threatened by ocean acidification. *Proceedings of the National Academy of Sciences of the United States of America* 107: 12930–12934.www.pnas.org/cgi/doi/10.1073/pnas.1004519107.http://dx.doi.org/10.1073/pnas.1004519107/-/DCSupplemental. www.pnas.org/cgi/doi/10.1073/pnas.1004519107.

Neuheimer AB, Thresher RE, Lyle JM, and Semmens JM (2011) Tolerance limit for fish growth exceeded by warming water. *Nature Climate Change*.http://dx.doi.org/10.1038/NCLIMATE1084.

O'Connor MI, Bruno JF, Gaines SD, *et al.* (2007) Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences of the United States of America* 104: 1266–1271.

O'Connor MI, Piehler MF, Leech DM, Anton A, and Bruno JF (2009) Warming and resource availability shift food web structure and metabolism. *PLoS Biology* 7: e1000178.

O'Connor MI, Selig ER, Pinsky ML, and Altermatt F (2011) Toward a conceptual synthesis for climate change responses. *Global Ecology and Biogeography* http://dx.doi.org/10.1111/j.1466-8238.2011.00713.x.

O'Connor NE and Crowe TP (2005) Biodiversity loss and ecosystem functioning: Distinguishing between number and identity of species. *Ecology* 86: 1783–1796.

Paine RT (1966) Food web complexity and species diversity. *The American Naturalist* 100: 65–75.

Palmer AR (1992) Calcification in marine mollusks: How costly is it? *Proceedings of the National Academy of Sciences of the United States of America* 89: 1379–1382.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics* 37: 637–671.

Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Pearson GA, Lago-Leston A, and Mota C (2009) Frayed at the edges: Selective pressure and adaptive response to abiotic stressors are mismatched in low diversity edge populations. *Journal of Ecology* 97: 450–462.

Peck LS, Webb KE, and Bailey DM (2004) Extreme sensitivity of biological function to temperature in Antarctic marine species. *Functional Ecology* 18: 625–630.

Pennings SC and Silliman BR (2005) Linking biogeography and community ecology: Latitudinal variation in plant–herbivore interaction strength. *Ecology* 86: 2310–2319.

Pepin P (1991) Effect of temperature and size on development, mortality, and survival rates of the pelagic early life history stages of marine fish. *Canadian Journal of Fisheries and Aquatic Science* 48: 503–518.

Perry AL, Low PJ, Ellis JR, and Reynolds JD (2005) Climate change and distribution shifts in marine fishes. *Science* 308: 1912–1915.

Peterman RM (2009) Fisheries Science in the Future. In: Beamish RJ and Rothschild BJ (eds.) *The Future of Fisheries Science in North America*, pp. 167–184. New York: Springer Science + Business Media.

Pörtner HO, Bock C, Knust R, *et al.* (2008) Cod and climate in a latitudinal cline: Physiological analyses of climate effects in marine fishes. *Climate Research* 37: 253–270.

Pörtner HO and Farrell AP (2008) Physiology and climate change. *Science* 322: 690–692.

Pörtner HO and Knust R (2007) Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science* 315: 95–97.

Pörtner HO, Langenbuch M, and Michaelidis B (2005) Synergistic effects of temperature extremes, hypoxia, and increases in CO_2 on marine animals: From Earth history to global change. *Journal of Geophysical Research* 110: C09S10.http://dx.doi.org/10.1029/2004JC002561.

Raven JA, Caldeira K, Elderfield H, *et al.* (2005) *Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide, Policy Doc. 12/05*. London: The Royal Society.

Reinfelder JR (2011) Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Annual Review of Marine Science* 3: 291–315.

Richardson AJ and Poloczanska ES (2008) Under-resourced, under threat. *Science* 320: 1294–1295.

Richardson AJ and Schoeman DS (2004) Climate impact on plankton ecosystems in the Northeast Atlantic. *Science* 305: 1609–1612.

Roughgarden J, Gaines SD, and Possingham H (1988) Recruitment dynamics in complex life cycles. *Science* 241: 1460–1466.

Ruiz GM, Carlton JT, Grosholz ED, and Hines AH (1997) Global invasions of marine and estuarine habitats by non-indigenous species: Mechanisms, extent, and consequences. *American Zoologist* 37: 621–632.

Sagarin RD and Somero GN (2006) Complex patterns of expression of heat-shock protein 70 across the southern biogeographical ranges of the intertidal mussel *Mytilus californianus* and snail *Nucella ostrina*. *Journal of Biogeography* 33: 622–630.

Sanford E (1999) Regulation of keystone predation by small changes in ocean temperature. *Science* 283: 2095–2097.

Sanford E and Kelly MW (2011) Local adaptation in marine invertebrates. *Annual Review of Marine Science* 3: 509–535.

Sarmiento JL, Gruber N, Brzezinski MA, and Dunne JP (2004) High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature* 427: 56–60.http://dx.doi.org/10.1038/nature02127.

Scheffer M, Carpenter S, Foley JA, Folke C, and Walker B (2001) Catastrophic shifts in ecosystems. *Nature* 413: 591–596.

Schiel DR, Steinbeck JR, and Foster MS (2004) Ten years of induced ocean warming causes comprehensive changes in marine benthic communities. *Ecology* 85: 1833–1839.

Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin* 216: 373–385.

Sherman K, Belkin IM, Friedland KD, O'Reilly J, and Hyde K (2009) Accelerated warming and emergent trends in fisheries biomass yields of the world's Large Marine Ecosystems. *Ambio* 38: 215–224.

Shurin JB, Borer ET, Seabloom EW, *et al.* (2002) A cross-ecosystem comparison of the strength of trophic cascades. *Ecology Letters* 5: 785–791.

Siegel DA, Mitarai S, Costello CJ, *et al.* (2008) The stochastic nature of larval connectivity among nearshore marine populations. *Proceedings of the National Academy of Sciences of the United States of America* 105: 8974–8979.

Smith LW, Wirshing H, Baker AC, and Birkeland C (2008) Environmental versus genetic influences on growth rates of the corals *Pocillopora eydouxi* and *Porites lobata* (Anthozoa: Scleractinia). *Pacific Science* 62: 57–69.

Smith-Keune C and van Oppen M (2006) Genetic structure of a reef-building coral from thermally distinct environments on the Great Barrier Reef. *Coral Reefs* 25: 493–502.

Somero GN (2010) The physiology of climate change: How potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. *Journal of Experimental Biology* 213: 912–992.

Sorte CJB, Williams SL, and Carlton JT (2010) Marine range shifts and species introductions: Comparative spread rates and community impacts. *Global Ecology and Biogeography* 19: 303–316.

Southward AJ, Hawkins SJ, and T BM (1995) Seventy years' observations of changes in distribution and abundance of zooplankton and intertidal organisms in the western English Channel in relation to rising sea temperature. *Journal of Thermal Biology* 20: 127–155.http://dx.doi.org/10.1016/0306-4565(94)00043-I.

Stock CA, Alexander MA, Bond NA, *et al.* (2011) On the use of IPCC-class models to assess the impact of climate on living marine resources. *Progress in Oceanography* 88: 1–27.

Talmage SC and Gobler CJ (2011) Effects of past, present, and future ocean carbon dioxide concentrations on the growth and survival of larval shellfish. *Proceedings of the National Academy of Sciences of the United States of America* 107: 17246–17251.http://dx.doi.org/10.1073/pnas.0913804107.

Todgham AE and Hofmann GE (2009) Transcriptomic response of sea urchin larvae *Strongylocentrotus purpuratus* to CO_2-driven seawater acidification. *Journal of Experimental Biology* 212: 2579–2594.http://dx.doi.org/10.1242/jeb.032540.

Walters C and Parma AM (1996) Fixed exploitation rate strategies for coping with effects of climate change. *Canadian Journal of Fisheries and Aquatic Sciences* 53: 148–158.

West JM and Salm R (2003) Resistance and resilience to coral bleaching: Implications for coral reef conservation and management. *Conservation Biology* 17: 956–967.

Worm B, Barbier EB, Beaumont N, *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.

Worm B and Lotze HK (2009) Changes in marine biodiversity as an indicator of climate change. In: Lechter T (ed.) *Climate Change: Observed Impacts on Planet Earth*, pp. 263–279. Amsterdam: Elsevier B.V.

Wulff EV (1950) *An Introduction to Historical Plant Geography.* Waltham, MA: Chronica Botanica Company.

Zeidberg LD and Robison BH (2007) Invasive range expansion by the Humboldt squid, *Dosidicus gigas*, in the eastern North Pacific. *Proceedings of the National Academy of Sciences of the United States of America* 104: 12948–12950.

Zimmerman RC, Kohrs DGLSD, and Alberte RS (1997) Impacts of CO_2-enrichment on productivity and light requirements of eelgrass. *Plant Physiology* 115.

Marine Ecosystems

J Frederick Grassle, Rutgers University, New Brunswick, NJ, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Benthic Pertaining to the bottom of the sea or other aquatic environment.
Benthos Organisms living on, in, or near the seabed or at the bottom of some other aquatic environment.
Coastal Estuaries, semienclosed seas, and shallower regions of the ocean, including areas influenced by rivers and runoff from land.
Community A group of species cooccurring in an area and interacting through trophic and spatial relationships.
Coral reef Benthic environments characterized by reef-building corals with symbiotic dinoflagellates.
Deep sea Volumes of water or areas of ocean bottom at depths greater than 200 m.
Mangrove Environments characterized by mangrove trees.
Nekton Actively swimming pelagic organisms.
Pelagic Pertaining to the water column in aquatic environments.
Plankton Organisms that float freely in the water column and do not maintain their position independent of water movements. Phytoplankton (literally plant plankton) is plankton with photosynthetic pigments and zooplankton is animals of the plankton.

Marine Ecosystems

Ecosystem Units

On land, ecosystems are separated into two-dimensional biomes, land areas defined by characteristic primary producing plants such as trees, grasses, and shrubs. Most shallow lakes and streams are similarly two dimensional; however, a few freshwater deep, ancient lakes, such as Lake Baikal in Siberia, and large rivers such as the Amazon have spatial complexity comparable to many coastal marine ecosystems. The ocean biosphere has an average depth of 4 km and comprises 99.5% of the biosphere. The dense seawater medium allows at least part of the life cycle of almost all marine organisms to be transported and dispersed by ocean currents. One ocean phylum is entirely pelagic, and about a third of the ocean phyla have representatives that spend their entire life cycle in near-surface waters as plankton. The boundaries that define ocean habitats and communities may involve a variety of overlapping criteria such as depth, distance from land, separation by landmasses, ocean currents, water masses of characteristic salinity and temperature, depth, and sea bottom characteristics such as sediment texture, composition, and surface topography. In addition, interactions with land and rivers and patterns of ocean circulation, light, nutrients, hydrology, and physical energy of water movements can strongly influence the distribution of species.

Descriptions of species boundaries are few and biogeographical classification depends heavily on the groups of organisms considered and how well they have been sampled. The ocean generally lacks the obvious barriers to dispersal characteristic of terrestrial environments. There may be multiple criteria for defining biogeographical provinces or marine ecosystems. For example, 64 large marine ecosystems (LMEs) are defined as oceanic areas of 200,000 km^2 or greater that extend from a coast to the margins of continental shelves or the seaward extent of predominant coastal currents (Sherman and Duda, 1999). The United Nations Environment Programme (UNEP) Regional Seas Programme has developed legally binding conventions for countries bordering 18 LMEs to protect the marine environment. At present, Marine Protected Areas occupy 1.17% of the global ocean surface or 4.32% of continental shelf areas, well below the 10% target set by the Convention on Biological Diversity in 2004 (United Nations Environment Programme, 2010).

Vulnerable Marine Ecosystems are defined by "geographic boundaries or by types of organisms and ecological processes (e.g., seamounts, coral reefs, continental slopes, abyssal plains) and ecosystems" (FAO, 2007). Ecosystem vulnerability may depend on the sensitivity of individual species, the ecological roles of threatened species, local types of hydrothermal vent and seep habitats, deep-sea sediment transport, and unique major bottom features such as seamounts, deep canyons, and slopes of volcanoes (Snelgrove, 2010). For example, in the Southern Ocean large areas of relatively undisturbed deep-sea bottom sediments are of special interest as habitat with a high diversity of species. Near the end of the first Census of Marine Life in 2010, ANDEEP (ANtarctic DEEP-sea biodiversity: colonization history and recent community patterns) had identified 1400 species of invertebrates and more than 700 species appear to be new to science. Of the 674 species of isopods, 87% are believed to be newly discovered species (Ramirez-Llodra *et al.*, 2010). To date, the Ocean Biogeographic Information System (OBIS), a legacy of the Census of Marine Life, has 30 million records of 118,000 species in 31,000 genera (http://www.iobis.org/node/299).

Major estuaries, where freshwater from rivers mixes with ocean water, are among the smallest individual ecosystem units in area. The largest units are regions defined by major boundary currents: features such as the Gulf Stream, Kuroshio, and Brazil currents, and the north and south subtropical ocean gyres (the Sargasso Sea and South Atlantic Gyre in the Atlantic and the North Pacific Subtropical and South Pacific Subtropical Gyres in the Pacific). In the far north, the Arctic Ocean ecosystem is a distinct ocean basin covered by ice, and the Southern Ocean around Antarctica is separated from the circulation of the Atlantic, Indian, and Pacific Oceans by the cyclonic circulation of the Antarctic Circumpolar Current.

As with terrestrial environments, marine ecosystems may be classified by their characteristic primary producers, i.e., single-celled phytoplankton that float in the surface layers of the ocean, marsh grasses, seagrasses, mangrove trees, seaweeds such as those forming kelp beds, the single-celled plants called zooxanthellae that live symbiotically with corals, and the chemosynthetic bacteria living in water, sediments, or symbiotically with other organisms at hydrothermal vents or other seep environments rich in chemically reduced compounds such as sulfides or methane.

Using combinations of coastline, coastal bathymetry, ocean current systems, surface winds, and biota, the near-surface pelagic layer of the ocean where primary productivity occurs has been classified into 51 provinces (**Figure 1**) by Longhurst (1998). Similar criteria have been used to classify coastal areas (Briggs, 1974). Marine sediments cover almost the entire surface of the ocean floor, yet a consistent global biogeographic classification of these benthic ecosystems has yet to be developed (Snelgrove *et al.*, 1997). In 2007, a classification of Marine Ecoregions of the World (MEOW) was proposed by major international conservation organizations (Spalding *et al.*, 2007). Twelve realms were divided into 62 provinces which were further divided into 232 ecoregions.

Comparison of Marine Environments with Land

The ocean occupies 71% of the surface area of the globe and the deep sea at depths below 200 m occupies 63.5% of the earth's surface. Seawater is 830 times more dense than air and supports most of the biomass in the ocean. The volume of seawater in the ocean provides 99.5% of the livable volume of the earth (Cohen, 1994).

Concentrations of near-surface chlorophyll in the ocean are measured according to wavelengths of light reflected from the surface of the ocean, which are sensed by earth-orbiting satellites. Extensive studies of the relationship between near-surface chlorophyll and primary production allow satellite-derived information on chlorophyll to be converted to maps of primary productivity. Until very recently, overall primary production was thought to be approximately half that on land. Using distribution of chlorophyll in satellite photographs and models, primary productivity of the oceans has been shown to be about the same as that on land ($\sim$45–50 Pg C per annum in the ocean and $\sim$55 Pg C per annum on land; Falkowski *et al.*, 1998). For regions without ice cover, average net primary productivity (NPP) per area in the ocean is a third of that on land (ocean: 140 g C m^{-2} yr^{-1}, and land: 426 g C m^{-2} yr^{-1}) Only about 1.7% of the ocean surface area has NPP greater than 500 g C m^{-2} yr^{-1} compared to 25% for land. Most productivity in the marine environment is from phytoplankton. Attached, multicellular algae contribute only about 2%. The highest productivity occurs in estuaries and upwelling areas – these highly productive areas contribute approximately 18% to net ocean primary productivity. In the open ocean, the greatest primary productivity is near the equator and at mid-temperate latitudes in the Northern Hemisphere where there are regional maxima in terrestrial productivity. A smaller peak in productivity occurs in the Southern Subtropical Convergence where physical processes supply high concentrations of nutrients to surface waters (Falkowski *et al.*, 1998; Field *et al.*, 1998).

Marine primary producers are small and mobile whereas terrestrial primary producers are mostly large and rooted in the ground – trees account for approximately 80% of the primary production in terrestrial systems. By contrast, in central ocean gyres, photoautotrophic bacteria less than 2 μm in diameter and short generation times account for most of the primary production. Oceanic biomass is extremely dilute and filtering of organic particles is an important mode of feeding in marine environments.

Oceanic food webs have an average food chain length of nearly six trophic links as opposed to four trophic links in terrestrial systems (Cohen, 1994). The number of species of smallest marine organisms, such as the various groups of

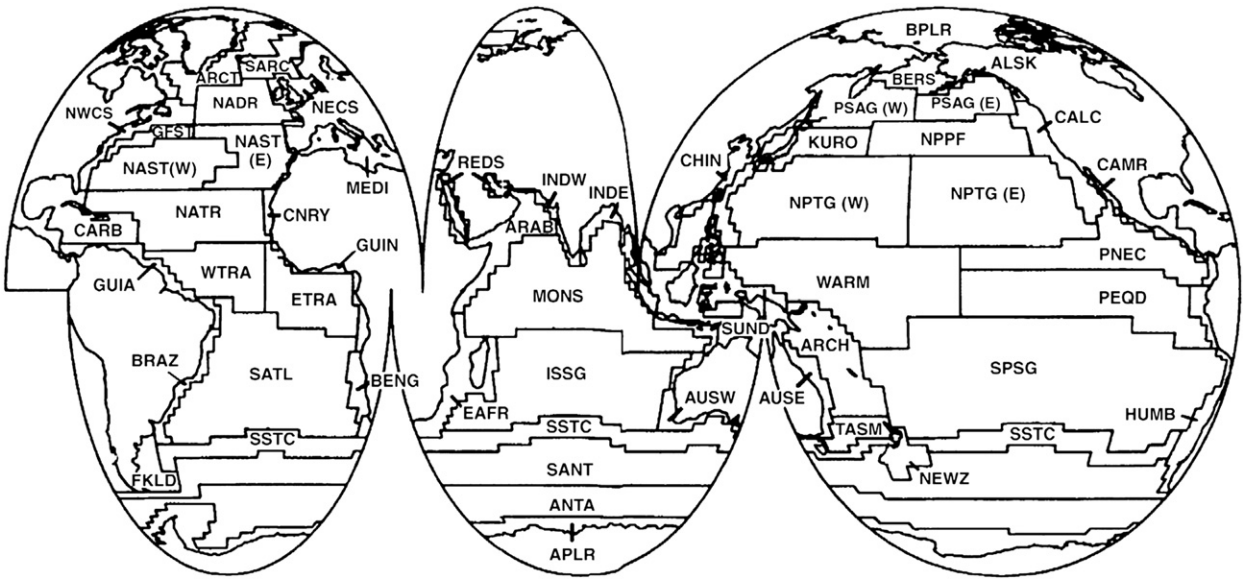


Figure 1 Pelagic biomes. Reproduced from Longhurst A (1998) *Ecological Geography of the Sea*. San Diego, CA: with permission from Academic Press.

one-celled marine organisms, are extremely poorly known. The relationship between the spectrum of individual body size and the spectrum of rates of population growth differs in marine and terrestrial systems (**Figure 2**). In open ocean food webs, the hierarchy of size is not apparent at the lower trophic levels because of the broad overlap in size of consumers and primary producers (**Figures 3** and **4**, Karl, 1999).

The pattern of temporal variability of the physical environment differs between oceans and land. Marine ecosystems are characterized by about the same environmental variation over weeks and years as over days – variability is constant at frequencies ranging from days to decades. In terrestrial environments the variance of environmental parameters (e.g., temperature) increases steadily from frequencies of hours to millennia. Beyond 50 years the variance increases with increasing frequency as it does over the entire time spectrum on land (Steele, 1985).

On land, individual organisms have a high probability of surviving the relatively predicable patterns of environmental variation that occur over time periods up to decades. For example, individual trees and many vertebrate animals resist adverse effects of variation at all frequencies up to several decades because of their large size and long generation time. In the open ocean, time series measurements at a single station and satellite measurements of ocean color from space show that net primary production varies significantly on periods from days to decades (Karl, 1999; Behrenfeld *et al*., 2006). Both seasonal and daily differences in cloud cover may result in three-fold variation in light at the surface. Vertical displacements of phytoplankton by internal waves further increase the amount of light absorbed by seawater before it reaches the photosynthetic organisms, creating a further source of variability. Small bacterial and flagellate primary producers have reduced the adverse effect of this variation in light by supplementing their diet from the pool of dissolved organic matter excreted by other organisms.

Other distinctive features of marine populations are outlined by Cohen (1994) and in a US National Academy of Sciences book on marine biological diversity (National Academy of Sciences, 1995). Plant and animal populations in marine ecosystems generally spend part of their life cycle as floating or swimming stages in the plankton. Unlike most terrestrial systems, the connections between benthic and planktonic life-history stages assume great significance and there is an unusually broad range of dispersal abilities, reproductive rates, and generation times. Almost all species have the ability to disperse in the water column as larval stages produced by some form of sexual reproduction. As a consequence, marine ecosystems are largely open and distant marine habitats can be linked by dispersing larvae. Terrestrial systems are more localized functionally and localized extinction of species occurs more frequently. Invertebrate predators and grazers generally have very high reproductive output, which makes population fluctuations more likely. Fluctuations at the highest trophic levels affect interactions among species at successively lower trophic levels. This cascading effect often has unpredictable consequences, and even the lowest trophic level of primary producers may be controlled from the top down. Bottom-up control of food webs is exerted through the effects of nutrients and physical processes on primary productivity.

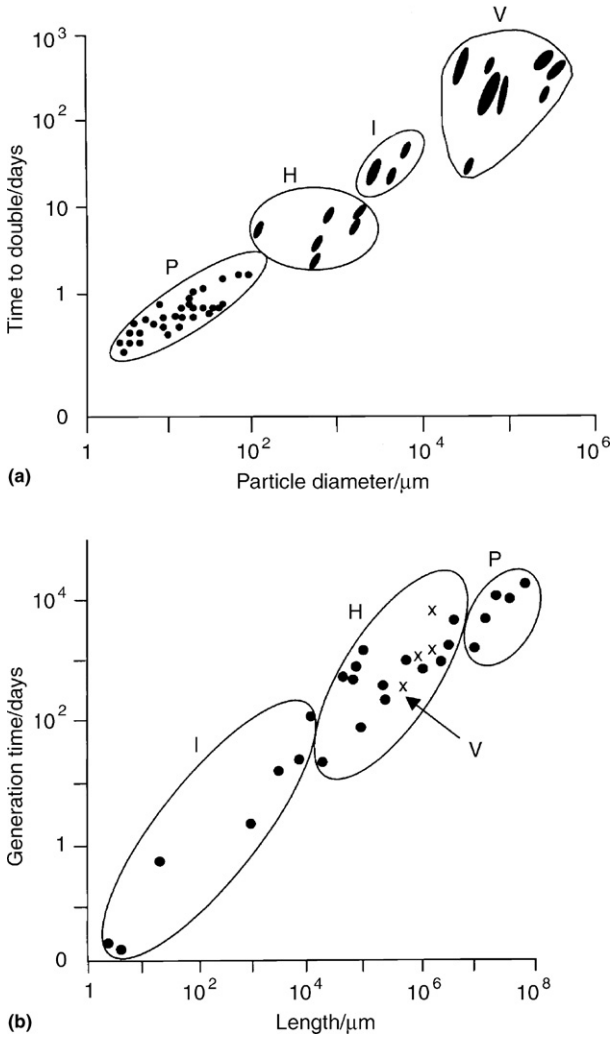


Figure 2 Relation of size to growth for plants (P), herbivores (H), other invertebrates (I), and vertebrates (V). (a) Reproduced from Sheldon RW, Prakash A, and Sutcliffe Jr. WH (1972) The size distribution of particles in the ocean. *Limnology and Oceanography* 17: 327–340, for pelagic marine ecosystems. (b) Reprinted from Bonner (1965) using only the terrestrial species. Derived from Cohen (1994, p. 60), with permission from Princenton University Press.

Biodiversity of Marine Ecosystems

Higher Taxa

The three main biological lineages are the Bacteria, Archaea, and Eukarya (includes plants, fungi, protists, and animals). Recent advances in molecular-biological techniques permit the first measurements of highly diverse oceanic assemblages of bacteria and archaea that cannot presently be cultured in the laboratory. Bacteria are more abundant in the photic zone and Archaea are more abundant in deeper water.

The Eukarya (all taxa except the Bacteria and Archaea) are divided into 71 well-defined monophyletic groups with no

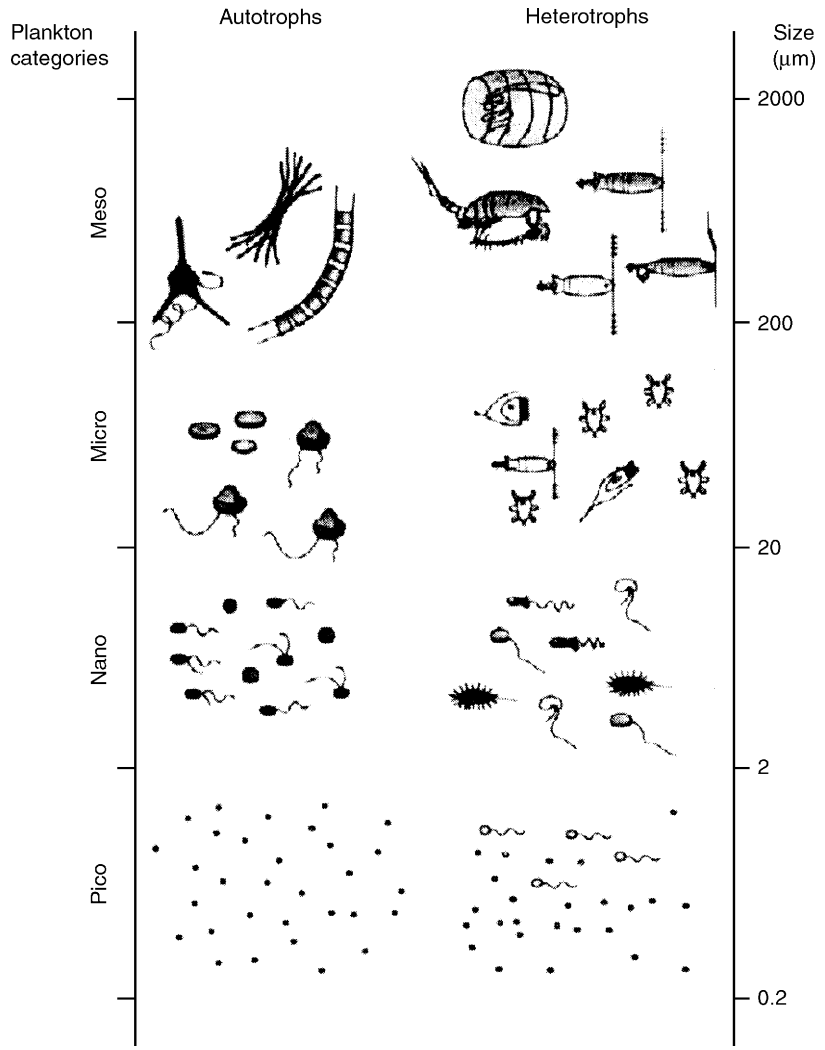


Figure 3 Representative classification of planktonic organisms by size showing the diversity of various autotrophic and heterotrophic groups. Size, *per se*, cannot be used to separate autotrophs from heterotrophs in North Pacific Subtropical Gyre plankton assemblages. Courtesy of Albert Calbet in Karl DM (1999) A sea of change: Biogeochemical variability in the North Pacific Subtopical Gyre. *Ecosystems* 2: 181–214.

apparent taxonomic affinity with one another on the basis of cell organization (Patterson, 1999). Each of these groups includes taxa formerly assigned to the protists. By this classification animals and their relatives the choanoflagellates, and fungi and their relatives the chytrids, are defined as a single group. Plants are in another group altogether with 11 categories (~7000 species) of green algae.

Important groups of primary producers have affinities with several other monophyletic groups. The red algae are a distinct group with about 4000 known species; the ~1000 species of dinoflagellates are related to the ciliates. The ~10,000 species of diatoms are in a highly diverse lineage that includes kelps and other brown algae. The conspicuous red, green, and brown seaweeds of rocky shores are divided among three separate lineages. The two most important primary producers in the open ocean were formerly called blue–green algae. They are actually prokaryotic bacteria in two groups: the *Synechococcus* with three lineages, and the *Prochlorococcus* group with two lineages. These organisms account for most of the phototrophic standing stock and primary production in the open ocean (Andersen *et al.*, 1996).

Among the many non-photosynthetic unicellular marine organisms, the ubiquitous Foraminifera are common both on the bottom at all depths and as pelagic organisms. Two abundant, poorly described benthic groups, the Komokiacea and the Xenophyophora (~40,000 known species), are big enough to be seen on the surface of deep-sea sediments. A leaflike form of Xenophyophora may be as large as 25 cm in diameter. These groups are separate lineages with no obvious relatives.

In the classification of marine, free-living, multicellular animals there are 31 phyla. **Table 1** (modified from May and Godfrey, 1994) compares the described diversity and abundance among marine benthic, marine pelagic, freshwater, and terrestrial environments. Of the 31 known phyla, all are known to have lived in the ocean and 14, or about half, are known only from the ocean. Living representatives of the Phylum Onychophora are presently found only on land in

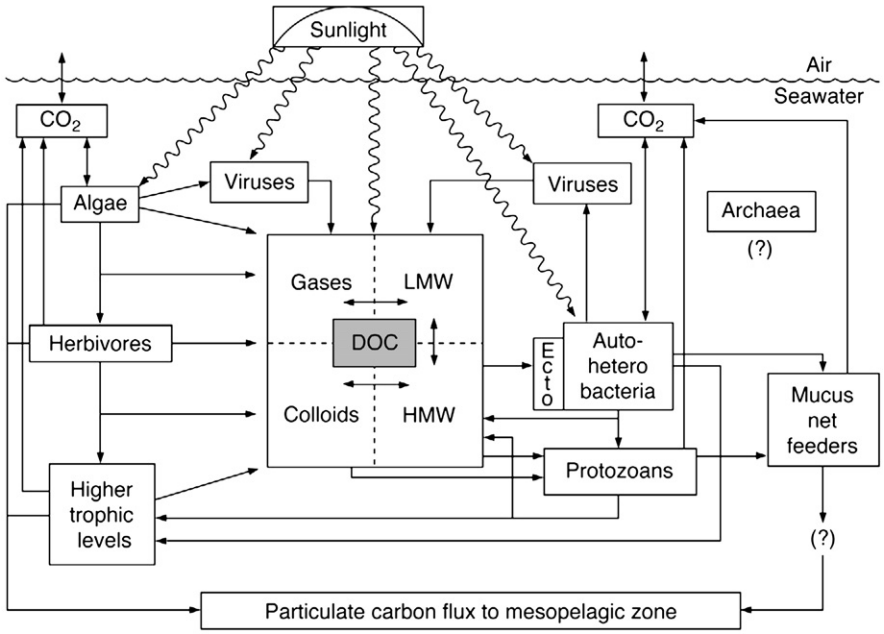


Figure 4 Schematic representation of the oceanic food web showing, on left, the classic pathway of carbon and energy flow through the photosynthetic eukarya to herbivores and on to higher trophic levels. Depicted on the right is the microbial food web, which uses energy stored in the nonliving, detrital carbon pool to produce microbial biomass that can re-enter the classic pathway of carbon and energy flow. Cell-associated ectoenzymes (ECTO) enable bacteria to use high molecular weight (HMW) dissolved organic carbon DOC in addition to the more traditional low molecular weight (LMW) and gaseous carbon substrates. Also shown in the microbial food web are viral particles and Archaea. At the present time, there is only rudimentary knowledge of the role of Archaea in the oceanic food web. Shown at the bottom of this diagram is the downward flux of particulate carbon (and energy), which is now thought to fuel most subeuphotic zone processes. The classic algae–herbivore grazer pathway is most important in this regard. Reproduced from Karl DM (1999) A sea of change: Biogeochemical variability in the North Pacific Subtopical Gyre. *Ecosystems* 2: 181–214, with permission from Springer.

the Southern Hemisphere, but are also known from fossil organisms that lived in the ocean more than 300 million years ago. Species diversity on land is dominated by insects and trees, groups that play a significant role only at the margins of the marine environment. Only about 15% of described species are found in the marine environment, but this may reflect the much greater cumulative effort devoted to species descriptions on land, rather than an actual difference in the number of species (May and Godfrey, 1994).

Species

Species are the units of evolution and both within-species and between-species genetic variation provided the bases for evolutionary change. For whole collections, species diversity is measured as the number of species and their relative abundances within and between habitats, regions, or other ecological or geographical units. Species richness is measured by collecting enough samples to represent very large numbers of individuals over very large areas. Ideally, communities should be sampled until the rate at which new species are found declines, and a plot of species versus area approaches a constant number of species. This level of sampling effort is achieved for groups with few rare species (e.g., larger animals including most vertebrates, planktonic organisms, and macrophytic plants). For species-rich taxa of bottom-dwelling invertebrates from coral reef or deep-sea habitats, this level of sampling has not been attained. Where habitats are patchy and the vast majority of species are rare, it is seldom possible to collect and process enough samples to estimate species richness accurately.

For individual samples, indices based on the absolute number of species and the relative abundance of species are used to study species diversity. The most commonly used index is the Shannon–Wiener information function, H′, which equals the frequency of each species $p_i = s_i/\Sigma s_i$ multiplied by $\log_2 p_i$ summed over the number of species (n) collected (e.g., $\Sigma p_i \log_2 p_i$). Another measure, Hurlburt rarefaction, calculates a species versus individuals curve for each sample based on the expected number of species in successively smaller subsamples drawn from an actual sample. These species diversity curves are especially useful in comparing samples of unequal size.

There are approximately 200,000 described species of animals in the marine environment (**Table 1**). The most species-rich and least well known areas are coral reefs and the sediments of the deep-sea floor. There are no precise estimates for these environments but estimates for coral reefs alone exceed 600,000 species (Reaka-Kudla, 1997). Based on quantitative analysis of 233 box core samples from the Atlantic Ocean continental slope and rise off the east coast of North America, Grassle and Maciolek (1992) estimated 1 to 10 million macrofaunal species in the deep sea (Gage and Tyler, 1991). May and Godfrey (1994) estimated 0.5 million based on the portion of species previously undescribed in the Grassle and Maciolek study. Poore and Wilson (1993)

Table 1 Free-living animal phyla and their relative numbers of described species (5 = >10^5, 4 = >10^4, 3 = >10^3, 2 = >10^2, 1 = present)

Phylum	*Marine*			
	Benthic	*Pelagic*	*Freshwater*	*Terrestrial*
Annelida	4	1	2	3
Arthropoda	4	3	3	5
Brachiopoda	2	0	0	0
Bryozoa	3	0	1	0
Chaetognatha	1	1	0	0
Chordata	3	3	3	3
Cnidaria	3	2	1	0
Ctenophora	0	1	0	0
Dicyemida	0	0	0	0
Echinodermata	3	1	0	0
Echiura	2	0	0	0
Gastrotricha	2	0	2	0
Gnathostomulida	2	0	0	0
Hemichordata	1	0	0	0
Kamptozoa	1	0	1	0
Kinorhyncha	2	0	0	0
Loricifera	1	0	0	0
Mollusca	4	2	3	4
Nematoda	3	0	3	3
Nematomorpha	0	0	0	0
Nemertea	2	1	1	1
Onychophora	0	0	0	1
Phoronida	1	0	0	0
Placozoa	1	0	0	0
Platyhelminthes	3	1	3	2
Pogonophora	2	0	0	0
Porifera	3	0	1	0
Priapula	1	0	0	0
Rotifera	1	1	2	1
Sipuncula	2	0	0	0
Tardigrada	1	0	2	1
Total (31)	27	11	14	11
Endemic	11	0	1	

Source: Modified from May RM and Godfrey J (1994) Biological diversity: Differences between land and sea. *Philosophical Transactions of Royal Society of London Series B* 343: 105–111.

analyzed samples from the Southern Pacific Ocean off Australia and, on the same basis, estimated that there are 5 million species in deep-sea sediments. Multicellular animals small enough to pass through a 1-mm sieve (meiofauna), such as nematode worms, are even less well known and Lambshead has argued that there may be 100 million species if nematodes are included (Lambshead *et al.*, 1994). Reasons for high diversity of species in the oceans include their long evolutionary history, the vast area of deep-sea floor (3×10^8 km^2) with relatively few barriers to dispersal, and the episodic nature of patch formation within and between habitats on a variety of spatial and temporal scales.

Genes

Genetic diversity is the heritable variation among individuals measured as allelic diversity at a broad sampling of genetic loci or as genetic sequence information at the molecular level within populations. Genetic variation occurs among subpopulations as well as within populations. Differentiation among subpopulations results from natural selection for genetic variants adapted to local patterns of environmental variation or random loss of genetic variants in small isolated subpopulations. Species with relatively high rates of dispersal are less likely to form subpopulations and species with very poor dispersal ability are more likely to diverge from parent populations as a result of random processes. In coastal areas, genetic divergence is related to the length of life of dispersal stages and barriers to current flow from one place to another along a coastline. For some shallow-water species, genetic isolation of island populations is related to distances among islands. The archipelagos in the central Indopacific in the vicinity of Indonesia and Papua New Guinea have a high richness of species, which then declines eastward to relatively isolated peripheral island archipelogos (Stehli, 1965). In the same region, in a study of population differentiation in four species of sea urchins, Palumbi (1997) found high genetic diversity (mitochondrial DNA sequence diversity) in the central area (1.6% variation among individuals) and much lower genetic diversity (0.5% variation among individuals) in peripheral island localities to the east. For these species, genetic diversity and species diversity co-vary across gradients suggesting a similarity in the processes maintaining gradients in diversity despite different mechanisms for the origin of the variation. Fluctuations in population size in relatively isolated populations could result in both loss of genetic variants and reductions in number of species (Palumbi, 1997).

In the deep ocean, hydrothermal vents are analogous to islands in the sense that the fluid flows support widely separated biological communities, linearly aligned along the Mid-Ocean Ridge. The patterns of deep-sea ocean currents that transport dispersal stages of species restricted to hydrothermal vents are poorly understood, but it is possible to make estimates of gene flow from the extent of genetic differentiation among populations of individual species. The flow of hydrothermal fluids, containing energy-rich reduced compounds such as hydrogen sulfide, supports chemosynthetic primary productivity. At East Pacific Rise vents, the flow of hydrothermal fluid may last only a decade or two at any one site and all populations are maintained by dispersal over considerable distances. Species can be divided into three categories: those that show no geographic pattern of genetic differentiation, those that are isolated by distance, and species without a free-living larval dispersal stage, which apparently have good dispersal to sites along a single ridge segment but poor dispersal between separated ridge segments (Vrijenhoek, 1997). The latest methods for measuring genetic diversity are being applied to increase numbers of marine species and rapid advances can be expected (DeSalle *et al.*, 2005).

Ecosystem Function

It is useful to classify members of species assemblages according to their feeding relationships with other species in the ecosystem. A trophic unit includes all species that eat the same kinds of foods or are consumed by the same kinds of consumers. Within a food chain, there is a hierarchy of consumers

from primary producers to primary consumers followed by a further sequence of consumers. Each step in a food chain results in a reduction in biomass, and simple food chains are often described as a pyramid with plants at the base and apex predators at the top. In the water column, unicellular phytoplankton form the first trophic level of marine food chains. The second level is formed by herbivores and detritivores and subsequent levels are formed by successive levels of predators. Species at the highest trophic levels can affect the food web relationships among species at lower levels. For example, removal of a top predator can have cascading effects on herbivores and ultimately on primary producers.

Because of the dilute seawater medium, a great many marine species have developed both active and passive means for filtering or trapping food particles from the water. Copepods, the most common animals in the water column, have filtering appendages and gelatinous zooplankton cast mucous nets to feed on phytoplankton. Baleen whales filter zooplankton (krill) from the water column. On the sea bottom, clams and sea cucumbers pump water past internal filters and many animals in sediments pump water through their burrows in order to feed. Other bottom animals have appendages protruding above the sediments that trap or filter food particles. In many marine organisms, the distinction between producers and consumers is blurred. Reef-building corals use their tentacles to trap zooplankton yet may take most of their sustenance from photosynthetic dinoflagellates living symbiotically in their tissues. Other animal–plant relationships of this sort are found in tropical clams and one-celled radiolarians and foraminifera.

Some marine species play another important functional role by providing habitat for other species, either on a large spatial scale – as with coral or coralline algae reefs, polychaete worm reefs, oyster reefs, seagrasses, kelps, marsh grasses, and mangrove trees. On a smaller scale, biogenic sediment structures (tubes, burrows, mounds, fecal aggregations) and more persistent structures made by tube builders, sponges, or shell-bearing animals may serve as habitat for other species.

Some species significantly affect the ecosystem by regenerating nutrients that limit primary production. Burrowing animals release nutrients into the water column which stimulate phytoplankton growth. In chemically reduced sediments, animals pump water into sediments for respiration or feeding and supply oxygen to chemosynthetic primary producers living in the burrow. The role of single species is often not obvious, and several different criteria may be used to assign species to functional groups within an ecosystem (Braeckman *et al.*, 2010). In general, redundancy of ecosystem function within a functional group has the potential to stabilize ecosystem processes despite fluctuations in the environment. Loss of functional groups implies drastic changes in ecosystem function.

Ecosystem Diversity

Edge of the Ocean

Intertidal Beaches

Beaches can be classified according to topography, organic content of sediments, and wave action. Reflective beaches are dominated by low wave energy, low organic content, and coarse sand. Reflective beaches have waves 1 m high or less and are generally found on open coasts with deep embayments, on tropical coasts, and on the coasts of polar seas. Surging wave action filters and drains large volumes of water through the interstices of the sediments, resulting in well-flushed and highly oxygenated coarse sand deposits (Alongi, 1998). Dissipative beaches, at the other extreme of a continuum, are produced by a combination of high waves (>2.5 m) and fine sand deposits with higher amounts of organic matter. These are common on the west coasts of Australia and Southern Africa and seasonally on the west coast of North America where high wave swells and fine sands are abundant. Extensive intertidal sand and mudflats are common on dissipative beaches.

Many beaches have adjacent seagrass beds, kelp beds, or other sources of macro-detritus, which are deposited as thick layers of wrack on the beach. These accumulations support communities that include both marine and terrestrial invertebrates (e.g., beach hoppers, beetles, and kelp-fly larvae). Other beaches are characterized by growth of diatoms in the sediments and inputs of small, filterable organic particles. Many animals live in the sediments, and in some high energy situations animals such as mole crabs and small bivalves move up and down the beach with the tides filtering particles from the waves. Large sand flats, such as in the Wadden Sea in the Netherlands, may be especially productive and support high standing stocks of grazing invertebrates.

Kelp Beds

Kelps attach to the bottom and form a surface canopy at depths up to ~ 20 m. Under the most favorable conditions these large marine plants form subtidal forests and attain rates of primary production in excess of $1000\ \mathrm{g\ C\ m^{-2}\ d^{-1}}$. These forests provide protection and food for a rich community of fish and invertebrates. The biomass and abundance of kelps may be regulated by sea urchin consumers. Sea otters play an important role in maintaining kelp forests by controlling the abundance of sea urchins. In the absence of sea otters, kelp forests are reduced by urchins to a pavement of encrusting algae and sea urchins. Kelp forests are important nursery areas for many species of fish and their detrital production enhances the abundance of benthic populations (United Nations Environmental Programme, 1995). Kelp populations are influenced over large scales by oceanographic climate. Nutrient-rich conditions during La Niña years result in increased growth and reproduction of the competitively dominant, canopy kelp species, *Macrocystis pyrifera*. Inter-decadal-scale shifts in community composition result from fluctuations in kelp density (Dayton *et al.*, 1999).

Rocky Shores

Rocky coasts exposed to the open ocean are characterized by wave action resulting in communities of attached seaweeds and filter-feeding invertebrates such as bivalve mollusks (e.g., mussels), barnacles, and tunicates that provide physical structure for other species. Wave energy enhances the productivity of these ecosystems by continually renewing nutrients and food. The shore face and the organisms that reside on the shore can be divided into zones according to tidal height and

length of exposure to air and the interactions of the dominant species with herbivores such as snails (gastropod mollusks) and predators (particularly snails, starfish, and birds). The large-scale pattern of rocky-shore communities depends on the distribution of rocky outcrops and sporadic changes in climate resulting in unusually heavy waves, ice cover, or sedimentation from rivers. The interaction of physical change and biological relationships among species at a variety of spatial scales (from local to regional) and temporal scales (from annual storm events to inter-decadal climatic change) are most thoroughly understood for rocky intertidal ecosystems.

Coral Reefs

Coral reef ecosystems occur where conditions are favorable for growth of reef-forming corals with dinoflagellate primary producers living symbiotically in their tissues. Growth of corals over many generations in geologic time results in major limestone structures such as coral atolls or the Great Barrier Reef off Australia. Dense growths of coral can sometimes occur in the deep sea, but these species lack photosynthetic symbionts, grow relatively slowly, and do not form major reef structures.

Reefs grow in strong light and clear water at temperatures from 18 to 30 °C at latitudes between 30° N and 30° S. Coral reefs are adversely affected by high nutrient concentrations, runoff of sediments from land, direct removal, and overfishing. The average primary production of corals in combination with their symbiotic dinoflagellates is about 25 g C m^{-2} d^{-1} and varies greatly from species to species. Over large areas, net primary productivity of the most actively growing reef crests and slopes ranges from 1 to 5 g C m^{-2} d^{-1}.

Reefs support an enormous species richness and complexity of interactions among species. Conspicuous large animals include enormous coral heads and large fish such as groupers, stingrays, and manta rays. Many of the colorful reef fish do not move far and develop complex behavioral relationships both within and between species. Some live symbiotically with other species, for example, individual anemone fish live in close association with patches of anemones. Cleaner fish set up cleaning stations where they feed on the ectoparasites attached to the gills of other fish. Some species mimic the cleaner fish and take bites out of the fish expecting to be cleaned of parasites.

Continental Shelves

Continental shelf coastal areas, on the order of 10,000 km^2 or more, have been called "Large Marine Ecosystems" (Sherman, 1993). These are separated from other areas of the ocean by continental shelf depth and ocean currents, and the shapes of coastlines form major seas, bays, or gulfs. Examples include the Baltic, North, Mediterranean, Black, Caspian, Red, Arabian, Barents, Bering, Okhotsk, Japan, Yellow, East China, Sulu, Celebes, and Caribbean Seas; the Bay of Bengal and Walvis Bay; and the Gulfs of Alaska, California, and Mexico. Primary productivity in these systems ranges from less than 35 g C m^{-2} yr^{-1} in the low latitude, warm waters of the Red Sea and high latitude, cold waters of the Beaufort Sea (10–20 g C m^{-2} yr^{-1}) to the very high primary productivity of Eastern Boundary Current upwelling areas in the Southern Hemisphere (1000–2000 g C m^{-2} yr^{-1}) in the Peru Current and Walvis Bay (Walsh, 1988). Most of the world's major fisheries are on continental shelves in mid-latitudes.

The Open Ocean and Deep Sea

Pelagic

The largest ecosystems in the ocean are the central gyres of the Atlantic, Pacific, and Indian Oceans. Ecosystem processes in the North Pacific Subtropical Gyre (NPSG) have been summarized by Karl (1999). This ecosystem is the largest circulation feature on the planet (2×10^7 km^2) and one of the most persistent, its boundaries having remained approximately the same for the past 10^7 years. The NPSG has a clockwise circulation of less than 4 cm s^{-1} and forms a circumscribed, stable, and relatively homogenous habitat. The surface mixed layer varies from 40 to 100 m depth and is characterized by surface temperatures of 24 °C or higher, low nitrate concentrations but relatively high dissolved organic nitrogen, and low standing stocks of organisms. The zone of primary productivity can be divided into two layers: an upper layer where chlorophyll increases in the winter and decreases in the summer, and lower layer (100–175 m), where chlorophyll increases in the spring and declines in the fall. Recharge of nutrients is from deeper water as a result of vertical eddy diffusion and episodic mixing events. This leads to considerable spatial variability in mixing processes and nutrient concentrations that vary by as much as three orders of magnitude. Phytoplankton primary production was once thought to be mostly by eukaryotes (diatoms and flagellates), but is now known to be more than 90% from the small bacterial taxa *Synechococcus* and *Prochlorococcus*. The standing stocks of these autotrophic bacterial groups comprise 80% of chlorophyll *a* and feed a microbial loop that internally regenerates nutrients and maintains a pool of dissolved organic matter, which supports them (**Figure 4**). The abundance of these autoheterotrophs is controlled by light, nutrients, and predation by bacteria and a mixed assemblage of protists. Viral infection may also be an important source of mortality for these organisms. Archaea are abundant but it is not clear whether they are significant chemosynthetic primary producers because little is presently known about these organisms.

Very little organic matter escapes remineralization and the microbial loop provides negligible subsidy to the rest of the food web. The classic food chain pathway of eucaryote phytoplankton to copepod herbivores and on to higher trophic-level fish is ephemeral and occurs more frequently in surface waters during the summer. Organic matter produced by the eucaryotic phytoplankton food chain constitutes most of the exportable carbon during aperiodic, pulsed events.

Falkowski *et al.* (1998) have summarized the biogeochemical processes controlling primary production in the open ocean. The central ocean gyres in the Atlantic, Pacific, and Indian Oceans have been considered analogous to deserts on land with low primary productivity and contain only ~0.2 mg m^{-3} of chlorophyll. Coastal upwelling regions,

seasonally mixed regions of temperate and boreal seas, divergent subpolar gyres, and meso-scale features with eddy-induced pumping have sufficient vertical flux of nutrients to support 5 mg m^{-3} of chlorophyll. Throughout most of the coastal and open ocean, primary production is limited by the availability of inorganic fixed nitrogen. In some instances, the cyanobacteria that fix nitrogen in the open ocean are limited by iron and an important source of iron to the ocean is dust carried from land by winds. Limitation of primary production by lack of iron is especially notable in the South Pacific (Falkowski *et al.*, 1998). There has been a lively debate about using iron fertilization on a large scale as a geo-engineering solution to increasing levels of atmospheric CO_2. A series of short-term, small-scale open ocean iron fertilization experiments have provided useful information on open ocean ecosystems but have not supported this as a CO_2-mitigation strategy (Strong *et al.*, 2009)

Benthic

The deep-sea floor is divided into major ocean basins by continents and the Mid-Ocean Ridge. Communities within ocean basins may be further divided according to depth, sediment type, and level of energy of deep-sea currents. The deep ocean floor is the least-known part of the planet but, through use of manned and unmanned submersibles, distinct ecosystem processes at hydrothermal vents, continental margin seeps, sea-mounts, ocean trenches, and areas of strong bottom currents are being explored and described.

The largest ocean basins and deep ocean trenches each have some species that live only in that basin and nowhere else. Hydrothermal processes along the Mid-Ocean Ridge mix seawater through porous rock at high temperatures yielding an energy-rich fluid containing reduced compounds. These compounds support chemosynthetic microorganisms that provide primary production for a discrete ecosystem clustered around each hydrothermal vent. Flows of subsurface fluid seep out of sediments deposited along some ocean margins providing similar energy-rich fluid to chemosynthetic organisms.

The food supply for most of the deep sea comes from the productivity of surface waters. When diatoms bloom, or gelatinous animals such as salps multiply rapidly, they die and sink, so that organic material accumulates in low areas of the uneven surface of the sea floor and in burrows and depressions left by the larger inhabitants. Even in the central ocean gyres where export production is low, the dead remains of fish, marine mammals, or terrestrial plant material carried seaward sink and form widely separated organic patches on the sea floor. Species respond to these patches at different rates and the probability that two species reach the same patch at the same time is low. This reduces the likelihood of species competing and of one species eliminating another. Most deep-sea species are small and many, including most fish species, are relatively slow growing, long lived, and late in maturation. Attempts to sustain deep-water fisheries have proven unsuccessful because low rates of population growth cannot keep up with rates of removal.

Species that grow relatively fast characteristically respond to patchy but concentrated sources of food from the ocean surface, such as wood from rivers, or the bodies of pelagic animals that settle to the bottom (Baco and Smith, 2003). For example, wood-boring bivalves rapidly colonize pieces of wood, grow to maturity in a few months feeding on their wood habitat, and produce thousands of eggs and larvae to colonize the next piece of wood that settles to the sea floor (Turner, 1973). Other species of bivalves grow very slowly in relatively homogeneous sediments, take several decades to reach maturity, and may produce only one egg at a time (Turekian *et al.*, 1975) in contrast to the rapid maturation and production of millions of eggs by most shallow-water bivalves.

Submarine canyons form conduits for sediment from continental shelves into the deep ocean. Unpredictable events of sediment erosion or scouring by intense currents result in relatively few species in the soft sediments at the bottom and sides of canyons. Seamounts are under sea mountains formed by the same processes at the hot spots on the ocean floor that form volcanic islands. Seamounts often support large populations of fish, and more than 70 species of commercially important fish have been reported. Interactions of currents with the steep topography of seamounts result in areas of enhanced primary productivity and high abundances of zooplankton that provide food for fish and dense concentrations of suspension-feeding benthic invertebrates (Rogers, 1994).

Mid-Ocean Ridges and Hydrothermal Vents

The 40,000 nautical mile Mid-Ocean Ridge system is the largest feature on the deep-sea floor. In 1977 a unique ecosystem was discovered at sites where a plume of high-temperature fluid rich in reduced compounds pours out into the water column. It is now known that sulfur oxidizers are among the most numerous bacteria and form a major base of the food chain. Other energy sources include hydrogen sulfide, methane, and reduced metals such as iron and manganese. In the Pacific, large, red-plumed worms up to 2 m long and large clams and mussels dominate the vents. These animals feed on organic compounds produced by symbiotic sulfur bacteria living in their tissues. Vents in the Atlantic have some of the same kinds of animals, but the most conspicuous are shrimp, which swarm over the surface of vent chimneys. Vents usually have a restricted distribution on any given ridge segment and persist for about 10–20 years, until there is local extinction of the vent community. Animals colonize new vents rapidly, grow fast, and produce enough offspring to colonize the next vent. In comparison with the rest of the deep sea, few species have adapted to the extreme thermal (4 °C up to temperatures in excess of 150 °C), chemical (high concentrations of cadmium, lead, cobalt, and arsenic) conditions at hydrothermal vents (Grassle, 1986). Most species found at hydrothermal vents live exclusively in this environment. Of the 443 species found at hydrothermal vents, 15 have been found in other sulfide-rich environments and only 30 species are known from elsewhere in the deep sea (Tunnicliffe *et al.*, 1998).

Potential Consequences of Anthropogenic Change

Eutrophication

Eutrophication is the increase in the rate of supply of organic matter to an ecosystem. Increases in global inputs of

nitrogenous fertilizers and the mining of phosphate rock have generated increased concern about the effects of eutrophication on enclosed marine ecosystems (Nixon, 1995). Eutrophic ecosystems have algal production in excess of 300 g C m^{-2} yr^{-1}, which results in areas of anoxia and loss of habitat for fish and other organisms. Relatively high rates of denitrification on continental shelves remove excess nitrogen originating from land sources and, in concert with dilution, help prevent adverse eutrophication effects in open coastal areas (Soetaert and Middelburg, 2009).

Overfishing

Globally, about 30% of commercial fish stocks are over-fished and another 44% are being fished at or near the maximum potential long-term catch rate. Atlantic halibut, cod, orange roughly, and many species of salmon are now severely depleted. Significant changes in community structure as a result of overfishing have occurred in ecosystem structure in the Bering, Barents, and Baltic Seas (National Academy of Sciences, Committee on Ecosystem Management for Sustainable Marine Fisheries, 1999). Bottom-fishing has been shown to result in physical destruction of some bottom habitats. Efforts to reverse the effects of over-fishing and habitat destruction may be having some positive effects in certain ecosystems (Worm *et al.*, 2009).

Overfishing has resulted in major changes in coral reef ecosystems. Normally, herbivorous fish heavily graze the attached algae, ensuring enough open reef surface for corals to settle and grow. This is especially true following major storms when wave action reduces coral coverage and circumstances are favorable for rapid algal growth. In the Caribbean, under normal circumstances, sea urchin grazing may compensate for reductions in fish grazing. A combination of overfishing and the decimation of sea urchin grazers by disease favored algal growth following a hurricane, which resulted in reefs dominated by algae (National Academy of Sciences, 1995). Recent research on coral-algal competition, in combination with the effects of no-take marine protected areas, have provided many insights into the positive and negative interactions among members of coral reef food webs (Hughes *et al.*, 2007; McCook *et al.*, 2010).

Invasive Species

Unwanted, non-indigenous species are often introduced to new geographic regions, both deliberately to start new fisheries and accidentally through release from aquaria or ballast water carried by ships, sometimes with disastrous consequences. The Asian clam became established in the San Francisco Bay in 1986 and quickly displaced other species from large areas of the seabed and altered the water chemistry of the bay (National Academy of Sciences, 1995). The introduction of predatory green crabs to coastal environments on the east coast resulted in major reductions in shellfish beds and the green crab has become invasive in many other areas. In short, invasive species have become a significant problem in many marine coastal environments and considerable effort is needed to curb this severe problem (Ruiz *et al.*, 2000).

In summary, the oceans encompass a broad array of habitats that differ in their diversity, function, and vulnerability. Much of the vast area of the oceans is poorly described, but we have some understanding of a variety of globally essential ecosystem processes and predictors of biodiversity (Tittensor *et al.*, 2010). Species loss may threaten not only the organisms themselves but also the many ecological processes that serve the rest of the planet and its human populations.

See also: Coastal Beach Ecosystems. Corals and Coral Reefs. Endangered Marine Invertebrates. Estuarine Ecosystems. Human Impact on Biodiversity, Overview. Intertidal Ecosystems. Invertebrates, Marine, Overview. Mangrove Ecosystems. Pelagic Ecosystems. Vents

References

Alongi DM (1998) *Coastal Ecosystem Processes*. Boca Raton: CRC Press.
Andersen RA, Bidigare RR, Keller MD, and Latasa M (1996) A comparison of HPLC pigment signatures and electronic microscopic observations for oligotrophic waters of the North Atlantic and Pacific Oceans. *Deep-Sea Research II* 43: 517–537.
Baco A and Smith C (2003) High species richness in deep-sea chemoautotrophic whale skeleton communities. *Marine Ecological Progress Series* 260: 109–114.
Behrenfeld MJ, O'Malley RT, Siegel DA, *et al.* (2006) Climate-driven trends in contemporary ocean productivity. *Nature* 444: 752–755.
Braeckman U, Provoost P, Gribsholt B, Van Gansbeke D, and Middelburg (2010) Role of macrofauna functional traits and density in biogeochemical fluxes and bioturbation. *Marine Ecological Progress Series* 399: 173–186.
Briggs JC (1974) *Marine Zoogeography*. New York: McGraw Hill.
Cohen JE (1994) Marine and continental food webs: Three paradoxes? *Philosophical Transactions of Royal Society of London B* 343: 57–69.
Dayton PK, Tegner MJ, Edwards PB, and Riser KL (1999) Temporal and spatial scales of kelp demography: The role of oceanographic climate. *Ecological Monographs* 69: 219–250.
DeSalle R, Egan MG, and Siddall M (2005) The unholy trinity: Taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of Royal Society of London B* 360: 1905–1916.
Falkowski PG, Barber RT, and Smetacek V (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* 281: 200–206.
FAO (2007) Report on the FAO Workshop on Vulnerable Ecosystems and Destructive Fishing in Deep-Sea Fisheries. *FAO Fisheries Report No. 829*, Rome, 26–29 June 2007.
Field CB, Behrenfeld MJ, Randerson JT, and Falkowski P (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281: 237–240.
Gage J and Tyler PA (1991) *Deep-Sea Biology: A Natural History of Organisms at the Deep-Sea Floor*. Cambridge: Cambridge University Press.
Grassle JF (1986) The ecology of deep-sea hydrothermal vent communities. *Advances in Marine Biology* 23: 443–452.
Grassle JF and Maciolek NJ (1992) Deep-sea species richness: Regional and local diversity estimates from quantitative bottom samples. *American Naturalist* 139: 313–341.
Hughes TP, Rodrigues MJ, Bellwood DR, *et al.* (2007) Phase shifts, herbivory, and the resilience of coral reefs to climate change. *Current Biology* 17: 360–365.
Karl DM (1999) A sea of change: Biogeochemical variability in the North Pacific Subtopical Gyre. *Ecosystems* 2: 181–214.
Lambshead PJD, Elge BM, Thistle E, Eckman JE, and Barnett PRO (1994) A comparison of the biodiversity of deep-sea marine nematodes from three stations in the Rockall Trough, Northeast Atlantic, and one station in the San Diego Trough, Northeast Pacific. *Biodiversity Research* 2(95): 107.
Longhurst A (1998) *Ecological Geography of the Sea*. San Diego, CA: Academic Press.
May RM and Godfrey J (1994) Biological diversity: Differences between land and sea. *Philosophical Transactions of Royal Society of London Series B* 343: 105–111.
McCook LJ, Ayling T, Cappo M, *et al.* (2010) Adaptive management of the Great Barrier Reef: A globally significant demonstration of the benefits of networks of marine reserves. *Proceedings of the National Academy of Sciences* 107: 18278–18285.

National Academy of Sciences (1995) Committee on Biological Diversity in Marine Systems. *Understanding Marine Biodiversity*. Washington, DC: National Academy Press.

National Academy of Sciences (1999) Committee on Ecosystem Management for Sustainable Marine Fisheries. *Sustaining Marine Fisheries*. Washington, DC: National Academy Press.

Nixon SW (1995) Coastal marine eutrophication. A definition, social causes, and future concerns. *Ophelia* 41: 199–219.

Palumbi SR (1997) Molecular biogeography of the Pacific. *Coral Reefs* 16(suppl.): S47–S52.

Patterson DJ (1999) The diversity of eukaryotes. *American Naturalist* 154(suppl.): S96–S124.

Poore GCB and Wilson GDF (1993) Scientific correspondence: Marine species richness (with reply to RM May). *Nature* 361: 597–598.

Ramirez-Llodra E, Brandt A, Danovaro R, *et al.* (2010) Deep, diverse, and definitely different: Unique attributes of the world's largest ecosystem. *Biogeosciences* 7: 2851–2899.

Reaka-Kudla ML (1997) The global biodiversity of coral reefs: A comparison with rain forests. In: Reaka-Kudla ML, Wilson DE, and Wilson EO (eds.) *Biodiversity II*. Washington, DC: Joseph Henry Press.

Rogers AD (1994) The biology of seamounts. *Advances in Marine Biology* 30: 305–350.

Ruiz GM, Fofonoff PW, Carlton JJ, Wonham MJ, and Hines AH (2000) Invasion of coastal marine communities in North America: Apparent patterns, processes, and biases. *Annual Review of Ecological System* 31: 481–531.

Sheldon RW, Prakash A, and Sutcliffe Jr. WH (1972) The size distribution of particles in the ocean. *Limnology and Oceanography* 17: 327–340.

Sherman K (1993) Large marine ecosystems as global units for marine resources management – An ecological perspective. In: Sherman K, Alexander LM, and Gold BD (eds.) *Large Marine Ecosystems*. Washington, DC: AAAS Press.

Sherman K and Duda AM (1999) An ecosystem approach to global assessment and management of coastal waters. *Marine Ecology Progress Series* 190: 271–287.

Snelgrove PVR (2010) Discoveries of the census of marine life. *Making Ocean Life Count*. Cambridge: Cambridge University Press.

Snelgrove PVR, Blackburn TH, Hutchings PA, *et al.* (1997) The importance of marine sediment biodiversity in ecosystem processes. *Ambio* 26: 578–583.

Soetaert K and Middelburg JJ (2009) Modeling eutrophication and oligotrophication of shallow-water marine systems: The importance of sediments under stratified and well-mixed conditions. *Hydrobiologia* 629: 239–254.

Spalding MD, Fox HE, Halpern BS, *et al.* (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* 57: 573–583.

Steele JH (1985) A comparison of terrestrial and marine ecological systems. *Nature* 313: 355–358.

Stehli FG (1965) Taxonomic diversity gradients in pole location: The recent model. In: Drake ET (ed.) *Evolution and Environment*. New Haven, CT: Yale University Press.

Strong AL, Cullen JJ, and Chisholm SW (2009) Ocean fertilization. Science, policy, and commerce. *Oceanography* 22: 236–261.

Tittensor DP, Mora C, Jetz W, *et al.* (2010) Global patterns and predictors of marine biodiversity across taxa. *Nature* 466: 1098–1101.

Tunnicliffe V, McArthur AG, and McHugh D (1998) A biogeographical perspective of the deep-sea hydrothermal vent fauna. *Advances in Marine Biology* 34: 353–442.

Turekian KK, Cochran JJ, Kharkar DP, *et al.* (1975) Slow growth rate of a deep-sea clam determined by ^{228}Ra chronology. *Proceedings of the National Academy of Sciences* 72: 2829–2832.

Turner RD (1973) Wood-boring bivalves, opportunistic species in the deep sea. *Science* 189: 1377–1379.

United Nations Environmental Programme (1995) In: Heywood VH and Watson RT (eds.) *Global Biodiversity Assessment*, pp. 370–399. Cambridge: Cambridge University Press.

United Nations Environment Programme (2010) *Regional Seas, Marine Biodiversity and Outloook Series*. Global Synthesis. A Report from the Regional Seas Conventions and Action Plans.

Vrijenhoek RC (1997) Gene flow and genetic diversity in naturally fragmented metapopulations of deep-sea hydrothermal vent animals. *Journal of Heredity* 88: 285–293.

Walsh JJ (1988) *On the Nature of Continental Shelves*. San Diego: Academic Press.

Worm B, Hilborn R, Baum JK, *et al.* (2009) Rebuilding global fisheries. *Science* 325: 578–585.

Marine Ecosystems, Human Impacts on

Juan Carlos Castilla, Pontificia Universidad Católica de Chile, Santiago, Chile

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 27–36, © 2001, Elsevier Inc.

Glossary

Alien: introduced, exotic, nonindigenous, nonnative, invasive species A species that has been transported by human activity (i.e., mariculture), intentionally or accidentally, to a site at which it does not naturally occur.
Ballast Water Water carried by a vessel to improve stability.
Benthic Organism An organism pertaining to the seabed; bottom-dwelling.
Biodiversity The variability among living organisms from all sources and the ecological systems of which they are a part.
Disturbance Any relatively discrete event in time that disrupts ecosystem, community, or population structure and changes resources, substrate availability, or the physical environment.
Ecosystem A complex nonlinear community of organisms in their physical environment.
Ecosystem Engineer Species Species that directly or indirectly modulate the availability of resources (other than themselves to other species) by causing physical state changes in biotic or abiotic material and in so doing modify, maintain, and/or create habitats.
Eutrophication Enrichment of a body of water with nutrients causing excessive growth of phytoplankton, seaweed, or vascular plants and often accompanied by a depletion of oxygen.
Food web, Trophic web A network of interconnected trophic chains in a community. A network of consumer–resource interactions among a group of organisms, populations, or aggregate trophic units.
Guild A group of species having similar functional roles in the community (e.g. herbivores).
Keystone species A group of species whose effects on the structure, dynamics, and functioning of the community is disproportionately large relative to its abundance.
Pelagic organism A free-swimming (nekton) or floating (plankton) organism that lives exclusively in the water column.
Resilience The resistance to a disturbance of a system and the speed of return to an equilibrium point, or the disturbance that can be absorbed before the system changes in structure by the change of variables and processes that control system behavior.
Species diversity The number of species in a given community (= species richness) and the way the species' abundances (i.e., number, biomass, and cover) are distributed among species (= species evenness).
Trophic level Feeding level in a food chain or pyramid (e.g., carnivores).

Introduction

The main threats to marine ecosystems are the human alteration of habitats, the excessive extraction of resources, pollution (Castilla, 1996), invasive species (i.e., introduction through mariculture and ballast water; Cohen and Carlton, 1998), eutrophication, and nonanthropogenic environmental changes (National Research Council (NRC), 1999; Castilla and Camus, 1992). Furthermore, multiple and compounded perturbations related to physically and biological based disturbances are resulting in communities entering new domains or "ecological surprises" (Paine *et al.*, 1998), with important modifications in their structure (i.e., species composition) and dynamics (i.e., alternative states).

Single, multiple, or compounded impacts on ecosystems may directly or indirectly affect their structure, including species diversity and functioning. Ecosystems are complexly linked nonlinear systems and their dynamics may be sensitive to past conditions and subjected to shifts when exposed to anthropogenic and nonanthropogenic environmental stress (NRC, 1999).

The concept of biological diversity (biodiversity; Heywood, 1995) is defined as: the variability among living organisms from all sources and the ecological system to which they are part. The analysis of biodiversity considers four levels: genetic, species, community, and ecosystems. This article focuses on the species diversity (richness, the number of species in a given community; evenness, species abundance), community resilience, and ecosystem functioning. One of the best avenues to integrate species diversity functioning and community resilience (Holling, 1973) is to study their dynamics through long-term manipulations. The article reviews long-term experiments and impacts on marine communities and ecosystems in which humans are one of the key ecological factors (Castilla, 1999).

Human Impacts on Marine Communities and the Effects on Species Diversity and Functioning

Rocky Intertidal Communities

Castilla (1999), based on a 16-year intertidal human exclusion experiment in central Chile (Las Cruces fenced Marine Coastal Preserve; ECIM), summarized the ecological roles played by humans as top predators on rocky mid-intertidal marine communities. The functional intertidal food web, without

humans (inside the ECIM preserve) and with humans (outside ECIM), differed substantially. On these rocky shores the impact of intertidal food gatherers is significant (Durán *et al.*, 1986). The collectors target mainly the keystone muricid snail *Concholepas concholepas*, locally known as "loco" (Castilla *et al.*, 1998). The high density of locos inside ECIM, following its closure to collectors in 1982, resulted in strong loco predation on the competitive dominant mussel *Perumytilus purpuratus*, which cannot "escape in size" from its predator. Therefore, a few years after the fencing of ECIM, the original dense mid-intertidal mussel beds inside ECIM were almost completely eliminated by the locos (Castilla, 1999). The primary space, so liberated, was readily invaded by two species of barnacles, *Jehlius cirratus* and *Notochthamalus scabrosus*, and several species of algae. Despite the fact that the loco also consumes barnacles, they have persisted for several years since they have a "weed recruitment strategy" (Castilla, 1988): After removal they keep reinvading the shore. This is not the case for *P. purpuratus*, which requires special substratum conditions to reinvade the shore (Navarrete and Castilla, 1990). Following the closure of the rocky shore at ECIM, species richness and evenness of sessile organisms using primary substrata increased inside ECIM. Outside ECIM (control), under reduced loco density due to food gathering, primary space is still dominated almost exclusively by the competitive dominant mussel *P. purpuratus*, and the biological diversity of the sessile primary substrata users is reduced since the mussels are long-term winners and appropriate the rock resource (**Figure 1**). Castilla (1999) provided a detailed account of direct and indirect human impacts on these communities and discussed differences in their functioning. For instance, it was noted that the settlement of keyhole limpets, *Fissurella* spp., was indirectly negatively impacted inside ECIM since their recruitment substratum, the beds of the mussels *P. purpuratus*, were absent due to loco's direct predatory impacts (**Figure 1**).

Nevertheless, in the papers previously noted, no mention was made that rocky intertidal species diversity should be viewed in a more comprehensive way so as to include the secondary substrata generated by *P. purpuratus*, an ecosystem engineer species (Jones *et al.*, 1994). Mussel matrices allow for the establishment of a rich macroinvertebrate and algal community composed of dozens of species (Paredes and Tarazona, 1980; Lohse, 1993) which live inside the matrices and on mussel shells. Although in central Chile this effect has not been evaluated, the *P. purpuratus* matrices enhance species richness (for southern Chile, see López and Osorio, 1977) in sites impacted by humans (outside ECIM) compared to those not impacted (inside ECIM, J. Castilla, unpublished results).

Similar ecological direct and indirect human impacts and drastic modification in rocky intertidal species evenness and intertidal community functioning (**Figure 2**) have been reported at Mehuín's southern Chile coastal preserve (Moreno *et al.*, 1984). Lindberg *et al.* (1998), through manipulative and "natural" experiments, described a three-trophic-level interaction among the American black oystercatcher (*Haematopus bachmani*), limpets (*Lottia spp.*), and erect fleshy algae in rocky intertidal bench communities of central and southern California. Human disturbances, such as the selective collection of large-size limpets and the reduction of shorebirds (in shores frequented by humans), drive the communities to a state dominated by small limpets and high cover of fleshy algae. Intertidal benches in relatively isolated islands (e.g., San Nicolas in central California) with large densities of oystercatchers and an absence of limpet human collection present communities in a different alternative state, which is characterized by large-size limpet populations and comparatively reduced fleshy algal cover.

Rocky Subtidal Communities

The Cape rock lobster *Jasus lalandii*, commercially the most important lobster species in South Africa, causes profound direct and indirect effects on subtidal competitive dominant mussel species, such as *Choromytilus meridionalis* and *Aulacomya ater* (Griffiths and Seiderer, 1980), severely modifying species diversity and community functioning. Barkai and Branch (1988a, b) compared the nearshore benthic communities of two adjacent islands on the west coast of South Africa: Malgas and Marcus Islands (33° S, 18′ E), which are approximately 4 km apart. The biotas of both islands have been protected from human exploitation since 1929. In the 1960s both islands supported populations of rock lobsters, but later, due to overfishing, a management plan was established which included a catch quota. Currently, Malgas still supports an unusually dense population of *J. lalandii* (probably partly due to the management plan) with densities of up to 10 individuals per square meter, whereas Marcus has a very reduced adult population of lobster. The benthic communities of both islands have only 34% of species in common. The biota of Malgas is dominated by numerous species of algae, whereas that of Marcus consists of thick beds of the black mussel *C. meridionalis*, an autogenic ecosystem engineer species that has a rich and diverse associated fauna (Barkai and Branch, 1988a). At Malgas, the predatory lobsters have eliminated a large proportion of spatial competitors, including mussels and barnacles, and sea urchins are absent. As a consequence, macroalgae proliferated. At Marcus, due to the absence of lobsters, the competitive dominant *C. meridionalis* formed dense beds, outcompeting other species of mussels, such as *A. ater* and algae; sea urchins are common (Castilla *et al.*, 1994). Barkai and Branch (1988a, b) discussed this ecological situation and argued for the existence of alternative stable states on the contrasting islands. **Figure 3** provides a summary of the main species involved, relative biomass, and direct, indirect, positive, and negative interactions between organisms on both islands.

The ecological impact of the Cape rock lobster at Malgas was experimentally demonstrated by Barkai and McQuaid (1988). The experiments showed that the drastic community differences between the islands were due to the dense population of lobster at Malgas and its absence at Marcus. In fact, the introduction of 1000 lobsters at Marcus ended amazingly: The lobsters were attacked by thousand of snails, *Burnupena* sp., which exist at Marcus in densities of up to 250 per square meter, and the lobsters perished within 30 min. This may explain their absence at Marcus, supporting the existence of an alternative ecological state.

In South African waters, it is unknown to what extent the commercial exploitation of rock lobsters or conservation

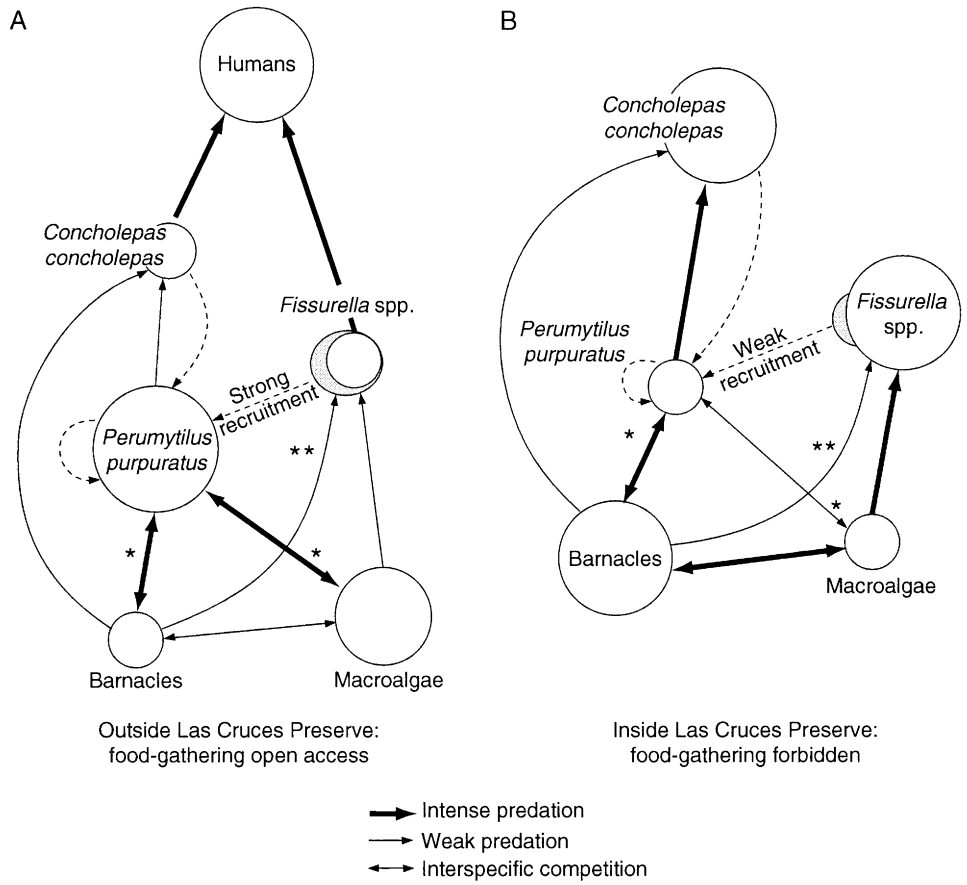


Figure 1 Diagrammatic representation of rocky intertidal food webs and human impacts outside (A) and inside (B) the Las Cruces (ECIM) Marine Preserve, central Chile. The size of the circles represents the approximate density of populations. An arrow with a singe asterisk indicates predation. The point of the arrow shows the flow of energy and the width indicates strong (wide) or weak (narrow) interactions. A double-asterisk arrow represents interspecific competition and the point width indicates the long-term competitive dominant (wider) or subordinate (narrower) species (intraspecific interactions are not considered). Settlement is shown by dashed lines and the arrows on these lines show the settler facilitator. One asterisk indicates that barnacles and macroalgae, apart from their ability to settle directly on rock, settle on top of mussel shells. A double asterisk indicates keyhole limpet browsing on young barnacles. *Concholepas concholepas* is a carnivore muricid. *Fissurella* spp. are herbivore gastropods (reprinted from Castilla, Rocky intertidal food webs and human impacts © 1999, p. 281), with permission from Elsevier.

measures (i.e., coastal closures) have impacted the nearshore rocky subtidal communities or in how many cases (other than Marcus and Malgas Islands) alternative stable states have been reached. This is a classical example in which both extreme attitudes—overexploitation and total conservation (no-take areas)—can result in drastically different species diversity and community functioning, mediated by the role of a high-trophic-level predator.

Humans and Linkages between Coastal and Oceanic Waters

Enhydra lutris, the northern sea otter, is found in near-shore environments ranging across the Pacific rim from Hokkaido (Japan) to Baja California (Mexico). The exploitation of their pelts led to the near extinction of otter populations in approximately 1911, when unregulated hunting was ended. Since then, the recovery of otter populations has occurred, particularly in the Aleutian Island chain, where by the 1970s the populations reached near maximum densities in some areas, were growing rapidly in others, and remained absent from others. Otters as keystone species (Power *et al.*, 1996) control the local biomass and the abundance of sea urchins, which regulate benthic algae biomass and productivity. Aleutian interisland comparisons (Estes *et al.*, 1998) have shown that kelp deforestation occurred in islands with low sea otter densities due to the increased density of sea urchins, whereas islands with high sea otter densities showed high kelp biomass. Estes *et al*. reported the complete transformation of a subtidal kelp forest in islands of the Aleutian Archipelago from three to four trophic-level systems and the release of sea urchin populations from the limiting influence of their predator, *E. lutris*. In the original circumstances, in the absence of sea otters, sea urchin populations increased rapidly and overgrazed the kelp forest, setting in motion a suite of different ecological impacts which drastically transformed the coastal ecosystems. These transformations had implications on the functioning of the communities and affected species diversity. Humans are highly involved in Estes *et al*.'s. findings. In recent years in western Alaska, declines of *E. lutris* populations have been observed. The authors have argued that this is probably due to the recent increased predation on sea otters by killer whales, *Orcinus orca*. *Orcinus* may have initiated

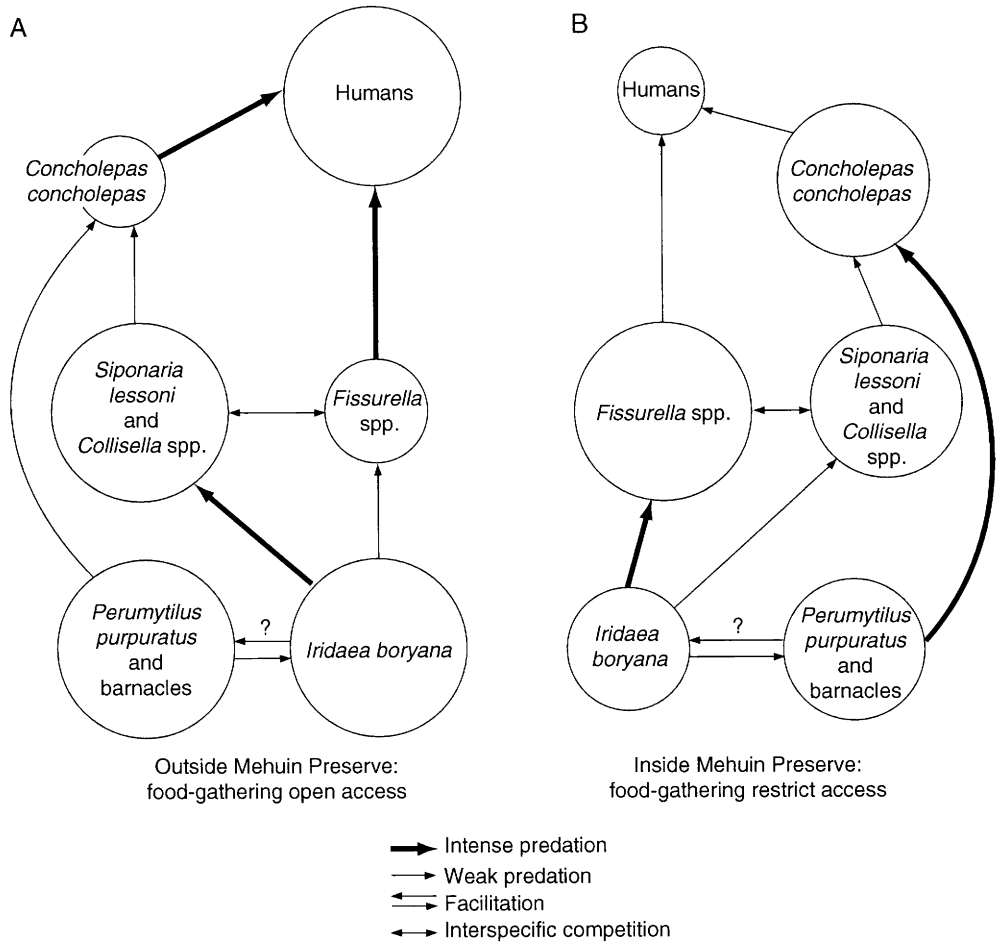


Figure 2 Diagrammatic representation of rocky intertidal food webs and human impacts outside (A) and inside (B) the Mehuín's Marine Preserve, southern Chile. Symbols are as described in the legend to **Figure 1** (reproduced from Moreno CA (1986) Un resumen de las consecuencias ecológicas de la exclusión del hombre en la zona intermareal de Mehuín-Chile. *Estud. Oceanol.* 5: 59–66), with permission from universidad de Antofagasta.

predatory influences that cascaded down successively lower trophic levels, first through the reduction of densities of sea otters, which triggered the increase of sea urchin populations, and ultimately the depletion of kelp biomass due to overgrazing. Estes *et al.*'s paper includes documented information on declines of sea otter populations and increases in the density and intensity of grazing of sea urchins on the kelp beds. Sea otters and killer whales have coinhabited the Aleutian Archipelago for millennia and Estes *et al.* attributed the sudden change of behavior of killer whales to a shift in their prey resource base. This has probably resulted from the collapse of pinniped populations, such as the Stellar sea lion and harbor seals, which were among the killer whale's main prey items. It has been suggested that the pinniped populations may have collapsed due to the Northwest Pacific midwater-trawl overfishing of walleye pollock (*Theragra chalcogramma*) (Alverson *et al.*, 1994) and/or increases in the ocean temperature. The authors recognized that some of their arguments contained speculations and that the critical one refers to the direct/indirect impacts of humans on marine ecosystems. In fact, sea otters, pinnipeds, and whales are under national and international protection in the Aleutians through different treaties and agreements signed dozens of years ago, but it also has to be recognized that their food resources have been depleted independently through overfishing. For instance, there is evidence that in the case of the pinnipeds a reduction (population collapses in some cases) has occurred mostly due to overfishing of pinnipeds, or of their fish resources, and also to climate changes. Overfishing is directly linked to human activities, and in Estes *et al.*'s scenario, humans and not killer whales may be considered as the apex predator. Humans have redirected the functioning of oceanic and coastal marine ecosystems in these localities and modified trophic linkages.

These examples indicate that there are at least two aspects of human ecological influences on marine communities that are difficult to evaluate and hence demonstrate an indisputable cause–effect situation. First, in many cases, the functioning of the marine communities is affected indirectly by anthropogenic activities—for example, human overfishing of pinniped's fish resources, collapse of pinniped populations, a shift in the prey item of killer whales, predation on the sea otter, population explosion of sea urchins, and overgrazing of kelp beds. The cascading down to successively lower trophic levels is complex and requires long-term observation and experiments

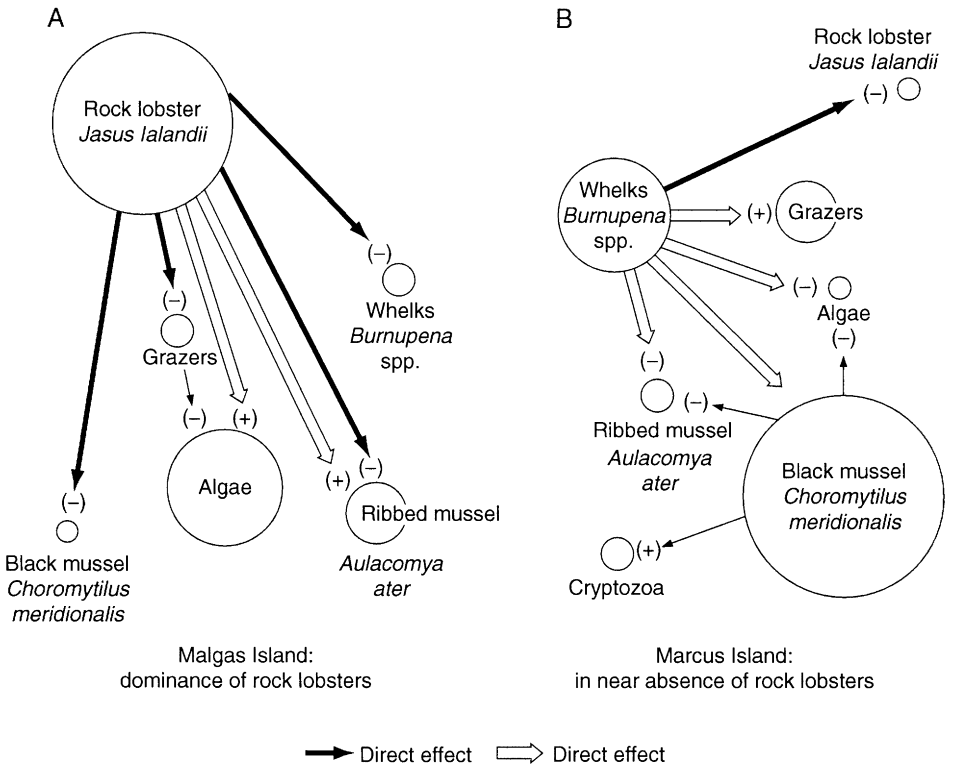


Figure 3 Rock lobster direct (+) and indirect (–) effects on mussel, welk, and grazer (sea urchin) preys in two South African islands. (A) Malgas, with a high density of adult lobsters. (B) Marcus, with a virtual absence of lobsters. The circles indicate relative biomasses (reproduced from Castilla JC, Branch GM, and Barkai A (1994) Exploitation of two critical predators: The gastropod *Concholepas concholepas* and the rock lobster *Jasus lalandii*. In: Siegfried WR (ed.) *Rocky Shores: Exploitation in Chile and South Africa*. Ecological Studies, vol. 103. pp. 101–130, with permission from Berlin: Springer-Verlag).

to be understood. Furthermore, nonanthropogenic impacts also need to be considered. Second, limited knowledge exists on the resilience properties of marine communities and ecological conclusions on linkages between marine ecosystems are based on preliminary data.

Humans and Ecosystem Engineer and Invasive Species

Ecosystem engineer species are species that directly or indirectly modulate the availability of resources (other than themselves) to other species by causing physical state changes in biotic or abiotic materials, and in so doing they modify, maintain, and/or create habitats (Jones *et al.*, 1994). Jones *et al.* distinguished (i) autogenic engineers, when the changes in the environment occurred via their own physical structure, living or dead tissues (e.g., coral reefs), and (ii) allogenic engineers, when they produced changes in the environment through the transformation of living or nonliving materials from one physical state to another via mechanical means (e.g., rabbits and burrows). In marine coastal communities, there are numerous autogenetic engineer species playing roles in the functioning of the community and ecosystem and creating the physical conditions for other species to exist (e.g., mussels, Lohse, 1993). In the Southern Hemisphere, rocky littoral zone tunicates of the genus *Pyura* play such a role (see Fielding *et al.*, 1994, for *P. stolonifera* in South Africa). These tunicates are also important as species extracted for food and/or bait by recreational fishers, divers, and inter tidal food gatherers (for *Pyura praeputialis* in Australia, see Fairweather, 1991; for *P. praeputialis* in Antofagasta, northern Chile, see Castilla, 1998). The tunicates form dense intertidal and subtidal belt monocultures and attain collective cemented beds, creating microhabitats used by several dozen macroinvertebrates and algae. Fielding *et al.* identified 83 taxa of macroinvertebrates in intertidal and subtidal *Pyura stolonifera* beds around Durban, South Africa, whereas more than 100 taxa of macroinvertebrates and algae have been found in intertidal *P. praeputialis* beds in Antofagasta.

The *P. praeputialis* beds in Chile present a very restricted geographical distribution of only 60–70 km around Antofagasta Bay (Clarke *et al.*, 1999). According to a working hypothesis (J. Castilla, work in progress), the species might have been introduced recently to Antofagasta by ships or arrived on floating objects from Australia. In Antofagasta, a contrasting situation concerning species richness is found in sites with *P. praeputialis*, with more than 100 taxa in the *Pyura* beds, as opposed to sites without the tunicate, which have about one-third to one-fourth of the species. It is unknown how much ecological damage, if any, human extraction causes on the dynamics of *Pyura* populations or on species diversity at a local scale. However, preliminary information at Antofagasta indicates that following *Pyura* removals by waves, predators, or humans, the species reinvades intertidal sites (the center of its distribution) within one year (J. Castilla, work in progress).

A higher rate of anthropogenic or nonanthropogenic removal of engineer species than the rate of recovery may be key to local species diversity.

Invasive species are displacing native species throughout the world. They are altering the physical nature of habitats (e.g., the effects of the Asian clam *Potamocorbula amurensis* in the San Francisco Bay) and causing changes in food webs of economically important species (NRC, 1999). The best reported case is that of the Bay of San Francisco, in which ship activities (i.e., the elimination of ballast waters) have increased drastically the number of exotic species in the bay's benthic communities (Carlton, 1996). At the pelagic level the introduction in the bay of the zooplanktonic mysid *Acanthomysis* sp., which displaced another species of mysid, *Neomysis mercedis*, a major food item of the striped bass *Morone saxitilis*, is partly responsible for a severe decline in the bay's bass population (NRC, 1999). Furthermore, there are recent reports showing that the predator green crab *Carcinus maenas* has invaded the San Francisco Bay and is spreading through the coastal waters of California (Cohen and Carlton, 1998).

Mariculture

The intensive and extensive marine farming of fish, shellfish, and algae has a long history and is a controversial issue. For instance, mariculture production expectations have not been achieved (NRC, 1999) and adverse environmental effects, such as contamination of surface waters by fish wastes, eutrophication, spread of diseases, introduction of unwanted species, and deterioration of coastal habitats (e.g., mangroves in connection with shrimp farming in Asia and Latin America), have occurred (Chamberlain, 1997; Anderson, 1997). The introduction of exotic cultured species may be a serious and irreversible event to native ecosystems which merits careful consideration. For instance, oysters have been transported by man from country to country and there are several cases of the concomitant spread of pests (unwanted species) and diseases, even under strict import controls. The introduction of the American oyster *Cassostrea virginica* into English waters (late 1800s and 1939) brought in several exotic species, the worst being the American oyster drill *Urosalpinx cinerea* and the gastropod competitor *Crepidula fornicata* (Edwards, 1990). Critical epizootic disease events in the Gulf of St. Lawrence that caused serious oyster stock depletions were ascribed to the transplant of oysters in 1914 from New England to Canada (Edwards, 1990). No comprehensive ecological reports on the ecological effects of these species introductions and diseases on local species diversity or community functioning have been published.

The intensive farm-raising of high-value species, such as shrimp and salmon, is far from trouble-free. There are concerns about the increase in the deposition of particulates and accumulation of organic matter under salmon cages in intensive mariculture installations due to unwanted effects, such as anoxic conditions and the production of toxic gases (Beveridge, 1996). Coastal ecosystem destruction, nutrient loading, antibiotics wastes, accidental release of alien or genetically altered organisms, and disease spreading to native species are some of the threats to community and ecosystem functioning.

Human Overfishing, Diseases, and Trophic Cascades

Hughes (1994) and Jackson (1997) reported major ecological effects on coral reef communities as a consequence of the overexploitation of herbivorous fishes and a disease killing sea urchins. In Caribbean coral reefs, a chain of effects, starting with the overfishing of herbivorous fishes, appeared following category 5 hurricane Allen in 1980. Allen severely damaged coral reefs in Jamaica, but by 1983 there was evidence of their recuperation. Nevertheless, at that time a disease devastated the herbivorous populations of the sea urchin *Diadema antillarum*. The elimination of the herbivore guild caused dramatic food cascading effects, resulting in reefs overgrown by algae and the detention of their recuperation. Species diversity and community functioning severely changed: The coral cover was reduced from approximately 52% in 1977 to 3% in the early 1990s, and cover of macroalgae increased from approximately 3 to 92% (Hughes, 1994).

Pollution and Artificial Reefs

The cases exemplified are among the best known ecological situations in which human impacts and the function of communities or ecosystems, combined with changes in species diversity, have been observed or studied. However, there are additional examples showing anthropogenic negative, as well as positive, impacts on marine communities and ecosystems. Among negative impacts on marine species diversity and community functioning, the most conspicuous (not discussed here) is pollution (Castilla, 1996). Among positive impacts is the building of marine reefs for fishing enhancement and recreational purposes. Artificial habitats may locally enhance species diversity and resources and drive community structure toward alternative states (Buckley, 1982).

Nonanthropogenic Environmental Changes and Variability

Nonanthropogenic environmental changes and impacts on marine populations and communities have been well documented. For instance, Soutar and Isaacs (1974) reported large fluctuations in the density of scales of hake, anchovy, and sardines in sediment cores during the past 2000 years, well before fishing was a factor. Large-scale ocean climate changes, such as El Niño Southern Oscillation (ENSO) events, have dramatic negative (Arntz and Fahrbach, 1996) or positive (Castilla and Camus, 1992) impacts on fish, shellfish, and algae populations in the Southeastern Pacific. ENSO also causes multiple positive and negative oceanic, freshwater, and terrestrial impacts throughout the world.

Barry *et al.* (1995) reported changes between 1931 and 1933 and between 1994 and 1995 in species richness and evenness of intertidal invertebrates at a rocky intertidal transect in the Hopkins Marine Station, Monterey, California. They reported species' latitudinal range shifting northward, suggesting a consistency with predictions associated with anthropogenic-linked climate warming (but see alternative explanation by Denny and Paine, 1998). Nevertheless, it is debatable whether the current global warming trend, due partly to the build-up of several

greenhouse gases, is part of a long-term climatic trend. In any case, marine species with different geographical origins would have different responses to water temperature alterations (Castilla and Camus, 1992). Moreover, in the case of the oceans, water temperature modifications would be just one of the potential factors affecting the distribution of species. For instance, temperature effects on the turbulence of the ocean waters, and their association with wind stress, may have major implications for plankton dispersal. Also, the predicted north–south interhemispheric asymmetry, due to the thermal inertia in the south, must be considered before drawing firm conclusions on marine species latitudinal shifts (Bernal, 1994). Furthermore, since the ocean is affected simultaneously by several climate forces (including anthropogenic greenhouse effects), it is difficult to determine the real cause of any observed change, such as that in surface seawater temperature. Shifts in marine populations, community structure, and their functioning represent the integrated response of species assemblages to nonanthropogenic long-term climate changes superimposed on the effects of numerous short-term factors, including anthropogenic forcing.

Conclusions

This article discussed several marine examples in which direct anthropogenic and nonanthropogenic impacts (or combinations), such as species extraction and oceanic water temperature modifications, caused drastic ecological shifts on marine benthic intertidal, subtidal, and coastal–oceanic communities, and thereby modified species diversity and the functioning of associated communities. Interestingly, extreme conservation measures (e.g., the establishment of no-take areas) to protect species, habitat, community, or ecosystem may also cause drastic modifications in the functioning of marine communities and drive communities into alternative ecological states (Castilla *et al.*, 1994; Estes *et al.*, 1998; Castilla, 1999). This article highlighted that anthropogenic activities (e.g., mariculture) and impacts (e.g., overfishing) on different ecological categories of species (predator, keystone, engineer, invasive, and competitive dominant) translate into differential responses and functioning at the species diversity and community level. The unique ecological role played by humans and their apex keystone position in trophic webs were discussed.

Acknowledgments

The author acknowledges financial support received through the Cátedra Presidencial en Ciencias, Chile (awarded in 1997) and the Pew Charitable Fund as a Pew Fellow in Marine Conservation (1997). Minera Escondida Ltd. supported fieldwork in Antofagasta, Chile. Comments and suggestions on a preliminary manuscript made by a referee are acknowledged.

See also: Aquaculture. Intertidal Ecosystems. Marine and Aquatic Communities, Stress from Eutrophication. Marine Ecosystems. Resource Exploitation, Fisheries

References

Alverson DL, Freeberg MH, Murawski SA, and Pope JG (1994) A Global Assessment of Fisheries Bycatch and Discards. *FAO Fisheries Technical Paper No. 339*. Rome: Food and Agriculture Organization.

Anderson JL (1997) The growth of salmon aquaculture and the emerging new world order of the salmon industry. In: Pikitch E, Huppert DD, and Sissenwine M (eds.) *Global Trends in Fisheries Management*, American Fisheries Society Symposium No. 20., pp. 175–184. Bethesda, MD: American Fisheries Society.

Arntz WE and Fahrbach E (1996) *El Niño: Experimento Climático de la Naturaleza*. México D. F., México: Fondo de Cultura Económica.

Barkai A and Branch GM (1988a) Contrast between the benthic communities of subtidal hard substrata at Marcus and Malgas Islands: A case of alternative stable states? *S. Afr. J. Mar. Sci.* 7: 117–137.

Barkai A and Branch GM (1988b) The influence of predation and substrata complexity on recruitment to settlement plates: A test of the theory of alternative states. *J. Exp. Mar. Biol. Ecol.* 124: 215–237.

Barkai A and McQuaid CD (1988) Predator–prey role reversal in a marine benthic ecosystem. *Science* 242: 62–64.

Barry JP, Baxter RD, Sagarin RD, and Gilman SE (1995) Climate-related long-term faunal changes in a California rocky intertidal community. *Science* 267: 672–675.

Bernal PA (1994) Global climate change in the ocean: A review. In: Mooney HA, Fuentes ER, and Kronberg B (eds.) *Earth System Response to Global Change: Contrast between North and South America*, pp. 1–15. San Diego: Academic Press.

Beveridge MCM (1996) *Cage Aquaculture*. Oxford: Fishing News Book.

Buckley RM (1982) Marine habitat enhancement and urban recreational fishing in Washington. *Mar. Fish. Rev.* 44: 28–37.

Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology* 77: 1653–1655.

Castilla JC (1988) Earthquake-caused coastal uplift and its effects on rocky intertidal kelp communities. *Science* 242: 440–443.

Castilla JC (1996) Copper mine tailing disposal in northern Chile rocky shores: *Enteromorpha compressa* (Chlorophyta) as a sentinel species. *Environ. Monit. Asess.* 40: 171–184.

Castilla JC (1998) Las comunidades intermareales de la Bahía San Jorge: Estudios de Línea Base y el Programa Ambiental de Minera Escondida Ltda. en Punta Coloso. In: Arcos D (ed.) *Minería del Cobre, Ecología y Ambiente Costero*, pp. 221–224. Talcahuano, Chile: Editora Anibal Pinto S. A.

Castilla JC (1999) Coastal marine communities: Trends and perspectives from human-exclusion experiments. *TREE* 14: 28–83.

Castilla JC and Camus PA (1992) The Humboldt-El Niño scenario: Coastal benthic resources and anthropogenic influences, with particular reference to the 1982/82 ENSO. *S. Afr. J. Mar. Sci.* 12: 703–712.

Castilla JC, Branch GM, and Barkai A (1994) Exploitation of two critical predators: The gastropod *Concholepas concholepas* and the rock lobster *Jasus lalandii*. In: Siegfried WR (ed.) *Rocky Shores: Exploitation in Chile and South Africa. Ecological Studies*, vol. 103. pp. 101–130. Berlin: Springer-Verlag.

Castilla JC, Manríquez P, Alvarado J, Rosson A, *et al.* (1998) Artisanal "Caletas" as units of production and co-managers of benthic invertebrates in Chile. In: Jameison GS and Campbell A (eds.) *Proceedings of the North Pacific Symposium on Invertebrates Stock Assessment and Management* Can. Spec. Publ. Fish. Aquat. Sci. 407–413.

Chamberlain G (1997) Sustainability of world shrimp farming. In: Pikitch E, Huppert DD, and Sissenwine M (eds.) *Global Trends in Fishery Management*, pp. 195–209. American Fisheries Society Symposium No. 20. Bethesda, MD: American Fisheries Society.

Clarke M, Ortiz V, and Castilla JC (1999) Does early development of the Chilean tunicate *Pyura praeputialis* (Heller, 1878) explain the restricted distribution on the species? *Mar. Sci. Bull.* 65: 745–754.

Cohen A and Carlton JT (1998) Accelerating invasions rate in a highly invaded estuary. *Science* 279: 555–558.

Denny MW and Paine RT (1998) Celestial mechanics, sea level changes, and intertidal ecology. *Biol. Bull.* 194: 108–115.

Durán LR, Castilla JC, and Oliva D (1986) Intensity of human predation on rocky shores at Las Cruces, central Chile. *Environ. Conserv.* 14: 143–149.

Edwards E (1990) Be careful with introductions. *Fish Farming Int.* 17.

Estes JA, Tinker MT, Williams TM, and Doak DF (1998) Killer whale predation on sea otters linking oceanic and nearshore ecosystems. *Science* 282: 473–476.

Fairweather PG (1991) A conceptual framework for ecological studies of coastal resources: An example of a tunicate collected for bait on Australian seashores. *Ocean Shoreline Manag.* 15: 5–142.
Fielding PJ, Weeters KA, and Forbes AT (1994) Macroinvertebrate communities associated with intertidal and subtidal beds of *Pyura stolonifera* (Heller) (Tunicata: Ascidiacea) on the Natal coast. *S.-Afr. Tydskr. Dierk.* 29: 46–51.
Griffiths CL and Seiderer JL (1980) Rock lobsters and mussels limitations and preferences in a predator–prey interaction. *J. Exp. Mar. Biol. Ecol.* 44: 95–109.
Heywood VH (1995) *Global Biodiversity Assessment.* United Nations Environmental Programme. Cambridge, UK: Cambridge Univ. Press.
Holling CS (1973) Resilience and stability of ecological systems. *Annu. Rev. Ecol. Syst.* 4: 1–24.
Hughes TP (1994) Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science* 265: 1547–1551.
Jackson JJ (1997) Reefs since Columbus. *Coral Reefs* 16 (Suppl.): S23–S32.
Jones CG, Lawton JH, and Schachak M (1994) Organisms as ecosystems engineers. *Oikos* 69: 373–386.
Lindberg DR, Estes JA, and Warheit KI (1998) Human influences on trophic cascades along rocky shores. *Ecol. Appl.* 8: 880–890.
Lohse DP (1993) The importance of secondary substratum in a rocky intertidal community. *J. Exp. Mar. Biol. Ecol.* 166: 1–17.
López MT and Osorio C (1977) Diversidad biológica en una comunidad intermareal de Putemún, Chiloé. *Bol. Soc. Biol. Concepción* 51: 123–127.
Moreno CA (1986) Un resumen de las consecuencias ecológicas de la exclusión del hombre en la zona intermareal de Mehuín-Chile. *Estud. Oceanol.* 5: 59–66.
Moreno CA, Sutherland JP, and Jara FH (1984) Man as a predator in the intertidal zone of southern Chile. *Oikos* 42: 155–160.
National Research Council (1999) *Sustaining Marine Fisheries.* Washington, D.C: National Academy Press.
Navarrete SA and Castilla JC (1990) Barnacle walls as mediators of intertidal mussel recruitment: Effects of patch size on the utilization of space. *Mar. Ecol. Prog. Ser.* 68: 113–119.
Paine RT, Tegner MJ, and Johnson EA (1998) Compounded perturbations yield ecological surprises. *Ecosystems* 1: 535–545.
Paredes C and Tarazona J (1980) Las comunidades de mitílidos del mediolitoral rocoso del Departamento de Lima. *Rev. Peruana Biol.* 2: 59–73.
Power ME, Tilman D, Estes JA, *et al.* (1996) Challenges in the quest for keystones. *BioScience* 46: 609–620.
Soutar A and Isaacs JD (1974) Abundance of pelagic fish during the 19th and 20th centuries as recorded in anaerobic sediments off California. *Fish. Bull. U. S. Fish Wildlife Service* 72: 257–275.
White PS and Pickett STA (1985) Natural disturbance and patch dynamics: An introduction. In: Pickett STA and White PS (eds.) *The Ecology of Natural Disturbance and Patch Dynamics*, pp. 3–13. New York: Academic Press.

Marine Mammals, Extinctions of

Glenn R Van Blaricom, US Geological Survey and University of Washington, Seattle, WA, USA
Leah R Gerber, Arizona State University, Tempe, AZ, USA
Robert L Brownell Jr., National Marine Fisheries Service, Pacific Grove, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Baiji The species of marine mammal also known as the Yangtze River dolphin or Chinese river dolphin, *Lipotes vexillifer* Miller, 1918, the sole member of the odontocete cetacean family Lipotidae. The last living baiji was seen in 2002, and the species was reported by experts to be extinct in 2007. This was the first known extinction of a cetacean species in recorded human history, and the fourth known extinction of a marine mammal species since the eighteenth century AD.

Baleen Modified, hardened epithelial structures suspended from the upper jaws of the mysticete whales. Baleen is in the form of multiple elongate plates within mysticete whale mouths, with a fibrous fringe on the inside edge that allows whales to filter zooplanktonic or fish prey from sea water.

Cetacea The order of mammals that includes the whales, dolphins, porpoises, and allies.

Delphinidae The family of odontocete cetaceans that includes the dolphins. Delphinids are the most diverse family of marine mammals with 36 species currently recognized, approximately 28% of the known modern species of marine mammals.

Desmostylia The only known fully extinct order of marine mammals, closely related to the sirenians. Desmostylians were large herbivorous quadripedal animals living in shallow waters during Oligocene and Miocene times, possibly resembling the modern hippopotamus superficially in ecology, behavior, and appearance. Desmostylians are known in the fossil record only from coastal regions of Washington and Oregon, USA.

Mysticeti The suborder of cetacean that includes the four families of baleen whales. Most mysticetes are quite large, including the blue whale which reaches 33 m in length and is recognized as the largest animal in the earth's history. Approximately 9% of the known modern species of marine mammals are mysticetes.

Odontoceti The suborder of the cetacean that includes the 10 families of toothed whales, including sperm whales, beaked whales, dolphins, porpoises, belugas, narwhals, and river dolphins. Approximately 56% of the known modern species of marine mammals are odontocetes.

Pinnipedia An obsolete but frequently used taxonomic term for the three families that comprise the seals, sea lions, fur seals, and walrus. Approximately 28% of the known modern species of marine mammals are pinnipeds. All pinnipeds are in the mammalian order Carnivora.

Scientific whaling Whaling activity currently undertaken by the Institute of Cetacean Research (Japan) with the stated goal of improved understanding of relationships of large whales and marine ecosystems in the world's oceans. The activity is permitted under a procedural exception to the moratorium on commercial whaling as managed by the International Whaling Commission. Scientific whaling is highly controversial and under ongoing intensive debate among scientists, management organizations, conservation advocacy groups, government agencies, politicians, and the public at large, and there is widespread concern that the activity may facilitate development of illegal trade and export of whale products, possibly leading to uncontrolled and damaging illegal harvests of whales.

Sirenia The order of mammals that includes the manatees, dugongs, and sea cows.

Introduction

Taxonomic Definition of ''Marine Mammals''

The ''marine mammals'' include one extinct order and three major extant taxa that were or are fully aquatic, in most cases occurring entirely in the marine habitats of the major ocean basins and associated coastal seas and estuaries. In addition, a few species of largely terrestrial taxa are currently regarded as marine mammals. We consider 127 Holocene mammal species in total to be marine mammals for purposes of this review (Rice 1998; see also Reeves *et al.*, 2002, and Jefferson *et al.*, 2008, for excellent non-technical accounts). We acknowledge that species numbers within any taxon are subject to virtually continuous revision as new systematic methods and philosophies emerge. Our primary bases for defining our list of marine mammal species are the protocols of the US federal government, defined by the US Marine Mammal Protection Act (MMPA) of 1972 (16 USC §§1361–62, 1371–84, and 1401–07 (Supp. IV 1974)) as amended (MMPA) and managed by two US federal agencies: the National Marine Fisheries Service (NMFS) and the Fish and Wildlife Service (FWS). Our principal source for taxonomic nomenclature, including common names, is the review of Rice (1998).

The order cetacea includes the whales, dolphins, porpoises, and their allies (**Table 1**). The ''pinnipedia'' is a group of species in three families in the mammalian order carnivora (**Table 1**). The pinnipeds include the seals, fur seals, sea lions, and allies. The order sirenia includes the extant manatees and dugong, and the now-extinct Steller's sea cow (**Table 1**). The order desmostylia is the only recognized order of marine mammals to become entirely extinct.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00187-8

Two largely terrestrial families of the order carnivora include species recognized as marine mammals (**Table 1**). The sea otter and chungungo (family mustelidae) live entirely or primarily in marine habitats. The polar bear (family ursidae) spends a significant proportion of time at sea as well.

Many other species of mammal utilize aquatic or marine habitats, including monotremes, ursids, mustelids, canids, primates, rodents, bats, and ungulates, among others. Ultimately, the distinction among aquatic, marine, and terrestrial taxa is arbitrary. Thus, our reliance on definitions and protocols of MMPA, NMFS, and FWS is subjective, although largely consistent with common practice among most specialists.

We use the term "marine" to refer to large, contiguous aqueous habitats with significant dissolved salt content in ambient waters, including the world's oceans, seas, and estuaries. We apply the term "aquatic" to aqueous habitats without significant measurable dissolved salt concentrations in ambient waters, such as lakes and rivers above the point of significant mixing with marine waters, and to inland saline lakes that lack outlet streams connecting to marine habitats. "Terrestrial" habitats are those lacking standing water under normal circumstances.

Table 1 Major taxa and species numbers of marine mammals, following the conventions of Rice (1998)

Taxon	*Number of species*
Cetacea: whales	83
Mysticeti: baleen whales	12
Balaenidae: right whales	2
Neobalaenidae: pygmy right whale	1
Eschrichtidae: gray whale	1
Balaenopteridae: rorquals	8
Odontoceti: toothed whales	71
Physeteridae: sperm whales	1
Kogiidae: pygmy sperm whales	2
Ziphiidae: beaked whales	20
Platanistidae: Indian river dolphin	1
Iniidae: Amazon river dolphin	1
Lipotidae: Chinese river dolphin	1[a]
Pontoporiidae: La Plata dolphin	1
Monodontidae: beluga and narwhal	2
Delphinidae: dolphins	36
Phocoenidae: porpoises	6
Carnivora, "pinnipedia"	36
Otariidae: sea lions and fur seals	16[b]
Odobenidae: walrus	1
Phocidae: seals	19[c]
Carnivora, other marine taxa	3
Mustelidae: marine otters	2
Ursidae: polar bear	1
Sirenia: manatees, dugongs, and sea cows	5
Trichechidae: manatees	3
Dugongidae: dugong and sea cow	2[d]
Total species	127

[a]The indicated species was last observed in natural habitat in 2002 and is considered extinct.
[b]One of the indicated species, the Japanese sea lion, was last observed in natural habitat in 1951 and is considered extinct.
[c]One of the indicated species, the Caribbean monk seal, was last observed in natural habitat in 1952 and is considered extinct.
[d]One of the indicated species, the Steller's sea cow, was last observed in natural habitat in 1767 and is considered extinct.

General Features and Habitat Boundaries

Marine mammals are characterized by a number of striking modifications, as compared with terrestrial mammals, in anatomy, physiology, and ecology (**Table 2**). In some cases, the modifications are sufficiently extreme that phylogenetic linkages to terrestrial ancestry are obscured, difficult to resolve, and matters of controversy among scientists. The degree of modification is correlated with the duration of the evolutionary history of the major marine mammal taxa as marine organisms.

Although marine mammals are largely defined by marked departures from the terrestrial mammalian model, it is instructive to consider the major features of terrestrial mammals retained in marine mammals. In the context of extinction processes in general, and anthropogenic extinctions in particular, two retained features are of particular importance. First, although most marine mammals spend most of their lives immersed at sea, they retain a largely terrestrial respiratory architecture and must surface and breathe in order to allow for exchange of respiratory gases. Second, marine mammals are homeothermic, with core body temperatures typically near 38 °C, like their terrestrial relatives. The need to breathe at the surface and the need for major anatomical

Table 2 Distinguishing characteristics of the major marine mammal taxa

Characteristic	*Cetacea*	*Sirenia*	*Pinnipedia*
Body streamlined	X	X	X
Limbs modified	X	X	X
Rear limbs modified as flippers			X
Rear limbs and pelvic girdle absent	X	X	
Propulsion by caudal spine and flukes	X	X	
Loss of pelage	X	X	
Subcutaneous blubber layer	X	X	X
Simplification of dentition	X	X	X
Expansion of anterior skull	X		
Development of acoustic capability for communication and echolocation	X[a]		
Amphibious capability			X

[a]Echolocation capability is known only for the odontoceti.

adjustment to minimize rates of heat loss are constraints that foster vulnerability to unsustainable rates of exploitation and to certain types of pollution. The significance of these constraints is developed in the case studies we present in this article.

The diving capabilities of marine mammals define the three-dimensional nature of their habitats at sea. Nearly all extant marine mammals dive to forage, although the ranges of diving capability and pattern are broad. Most marine mammals also spend significant time submerged while traveling, socializing, or breeding.

Among the cetaceans, sperm whales, beaked whales, and narwhals likely dive deeper and longer than other species on average. Sperm whales can dive to 3000 m, remaining submerged for an hour or more. The diving behavior of beaked whales is poorly known, but there is emerging evidence that beaked whales also may routinely make repetitive dives of long duration to great depth. Narwhals are known to be capable of dives of 1800 m depth and may dive to depths in excess of 1000 m 20–30 times per day. Most baleen whales and many of the smaller cetaceans commonly dive for less than 10 min at a time, and to depths no greater than a few hundred meters.

Among the pinnipeds, the elephant seals (in the family phocidae) have maximum diving capabilities to nearly 2000 m, and are known to make remarkably long sequences of repetitive deep (400–600 m), long (20 min or more) dives with surface intervals of only 2–3 min. These sequences may be maintained day and night for tens of days at a time. Many other phocid seals are thought to have similar capabilities. The sea lions and fur seals (otariidae), in contrast, generally dive for only a few minutes at a time, and generally to maximum depths of a few hundred meters, although many otariids are known to be capable of continuous sequences of repetitive shallow dives of 10–12 h or more.

In contrast to cetaceans and pinnipeds, sirenians are weak divers, generally remaining in shallow water (<20 m) and diving only for 2–3 min when active. Deeper dives (to 70 m) may occur on occasion, and dive duration can be quite long (>1 h) when animals are resting at the bottom. Sea otters are also relatively weak divers, reaching depths of 100 m and remaining submerged for a maximum of approximately 5 min, although most dives are to 30 m or less and last only for 1–2 min. There are no data available on the diving capabilities of the chungungo.

Few field observations of Steller's sea cow were made before extinction, but morphological analysis suggests that sea cows were unable to dive below the sea surface, surviving instead by foraging on kelp forest canopies and other macroalgae floating on the sea surface. Polar bears are able to make shallow dives, but do not typically engage in the extended repetitive dive sequences typical of many marine mammals, and apparently do not forage while diving. Polar bears instead use stealth, quickness, and great strength to capture phocid seals, their primary prey, at seal breathing holes on the ice surface. Thus, the extent to which the at-sea habitat of marine mammals is truly three-dimensional varies widely among the major taxa as well as the individual species. Within species, there is marked ontogenetic variation in diving capability and pattern as well.

The marine mammals are geographically ubiquitous in the world's oceans, seas, and estuaries (e.g., Perrin *et al.*, 2009). Cetaceans occur in marine environments at all latitudes. For example, killer whales may have the largest natural geographic range of the earth's mammals. Most of the mysticetes and some of the larger odontocetes have global ranges or are distributed antitropically. Smaller cetaceans are widely dispersed as well, although individual populations typically concentrate in regions of predictably high local biological productivity. Several species of small cetaceans, including delphinids, phocoenids, and the three extant monotypic families of river dolphins, are found in major river systems in South America and Asia. Pinnipeds occur in all the world's major marine habitats, but most species are concentrated in middle or high latitudes, in close association with regions of high productivity (e.g., Kovacs *et al.*, 2011). In addition, there are several pinniped species or populations confined to isolated large lakes in Europe, Asia, and North America. Most sirenians are confined to tropical or subtropical latitudes, in shallow seas that provide adequate macrophytic food and refuge from predation, and are thermally tolerable. The sea otter is confined to the coastal North Pacific Rim, and the chungungo to the temperate coastal southeastern Pacific. Polar bears occur only in the Arctic and sub-Arctic, rarely traveling south of 60° N latitude except in the relatively frigid northwestern Atlantic and Hudson Bay.

Cetacea

The distinguishing anatomical and functional features of the cetaceans are summarized in **Table 2**. The two major taxonomic subdivisions of cetaceans are the suborders odontoceti, or "toothed whales", and mysticeti, or "baleen whales". The odontocetes are the most diverse of the major Holocene marine mammal taxa, with 10 families and 71 species. The best-known families are the delphinidae and phocoenidae (**Table 1**). In ecological terms the family ziphiidae is among the most sparsely studied groups of mammals on earth. Other odontocetes include the sperm whales, beluga whales, narwhals, and river dolphins. The mysticetes (**Table 1**) include species that are the largest animals in the earth's history, with the largest of all, the Antarctic blue whale, reaching 33 m in length and 150,000 kg in mass.

Most cetaceans dwell in the open sea, or in the seas and estuaries of the continental margins. The exceptions are two delphinids, a phocoenid, and three odontocete species known as river dolphins (**Table 1**). Of the river dolphins, one species, the Franciscana or La Plata dolphin, may be found in coastal marine waters of Brazil, Uruguay, and Argentina. The other two river dolphins are exclusively aquatic. One of these, the boto, is known to leave river channels and travel within the adjacent flooded forests of the Amazon basin during the wet season. The Irrawaddy dolphin, a delphinid, occupies estuarine habitats and major river systems in southeastern Asia. The tucuxi, another delphinid, occupies the Amazon watershed and coastal marine habitats of tropical Atlantic South America. The finless porpoise, a phocoenid, occurs in the Yangtze River watershed and other large southern Asian rivers, and coastal marine habitats from the Persian Gulf to Japan.

Mysticetes often segregate feeding and breeding activities, both in time and space, connecting the two categories of

activity with extensive seasonal migration. Feeding is done primarily at high latitude during summer, and breeding and parturition primarily at low latitude during winter. Thus, a significant poleward migration is required during spring, and an equatorward return trip is necessary in autumn. The segregation of feeding and breeding, and the associated migratory behavior, is best developed and understood in the largest cetaceans. However, some large mysticetes, such as the bowhead whale of the Arctic region and the Bryde's whale of tropical and subtropical latitudes, undertake only modest seasonal migrations in contrast to species such as the humpback whale or gray whale. The smaller cetaceans, including most of the odontocetes, appear to disperse primarily on the basis of food availability and productivity at sea. Odontocetes typically do not engage in the dramatic seasonal migrations known for some of the mysticetes, although there are exceptions among the larger species.

Healthy mysticetes never leave the water, nor do they attempt to leave the water, in favor of land. A few odontocetes occasionally strand intentionally on beaches or river banks while in pursuit of prey. Perhaps the most familiar example is the brief intentional stranding of individual killer whales (*Orcinus orca* (Linnaeus, 1758)) while foraging on pinnipeds in the surf zone of Argentine beaches. In general, the cetaceans are not functionally amphibious. Aside from the noted exceptions, strandings of cetaceans can be considered abnormal, even pathological behavior typically resulting in death or serious injury.

Pinnipedia

The pinnipeds have a shorter evolutionary history as marine mammals and are derived from different ancestral taxa than the cetacea. Thus, although pinnipeds and cetacea often use similar aqueous habitats, the pinnipeds have many obvious differences from the cetacea in form and function. Features shared by most pinnipeds, including the phocids, otariids, and odobenids, are summarized in **Table 2**. The largest pinnipeds are adult males of the highly dimorphic southern elephant seal, exceeding 5 m in length and reaching 3700 kg in mass. Adult male northern elephant seals are only slightly smaller. Walruses are also quite large, reaching 3.5 m and 1500 kg.

Despite many anatomical, physiological, and ecological features obviously associated with life at sea, pinnipeds are best regarded as amphibious. All species utilize solid substrata for breeding and for postbreeding maternal care. Solid substrata are used also as short-term resting sites, and for protracted periods during molting of the skin and pelage. Although the proportion of time spent on land ("hauled out") over the long term varies significantly among age and sex categories, generally, the pinnipeds spend a major portion of their lives on land near shore, or on pack or shorefast sea ice.

Otariids use terrestrial habitats near shore for breeding, *post partum* maternal care, molting, and resting. Preferred hauling sites are those near areas of high oceanic productivity and free of large terrestrial predators. Thus, hauling grounds for otariids typically are islands or mainland locations protected by cliffs or rough terrain from land predators. Often the hauling grounds are localized at high latitude or the upwelling zones of midlatitude eastern boundary currents, such as the Humboldt, California, and Benguela currents, where production of preferred food species is high and temporally predictable. Optimal hauling grounds for breeding otariids are few in number, and often limited in size. Breeding activities of otariids are highly synchronous, occurring during a narrow time window when food availability for lactating females and newly produced juveniles is seasonally optimal. As a consequence of the various spatial and temporal constraints, pinniped breeding often involves conditions of extreme crowding on haulout sites.

Phocids and walruses haul out for the same purposes as do the otariids. Unlike otariids, the phocids and walruses use two different kinds of substrata. About half the phocid species use coastal land for hauling grounds, selecting sites for largely the same reasons as described for the otariids. Thus, timing and location of breeding for land-breeding phocids and otariids are generally similar. As a consequence, reproductive activities for many land-breeding phocids also occur under conditions of extreme crowding. Known exceptions include some populations of harbor seals with spatially dispersed, largely aqueous breeding systems, and the two extant species of monk seal with temporally asynchronous breeding systems at low latitude. The walrus and the remaining phocids breed, care for young, rest, and molt on ice at high latitudes rather than on land. Ice as a substratum varies widely over time and among locations in stability, vulnerability to predators, and provision of access to the surrounding sea. Thus, among ice-breeding pinnipeds, there are significant resultant variations in social and breeding strategy, and the degree of crowding at hauling sites. A major predator of ice-hauling phocids, the polar bear, is present in ice-covered marine habitats of the Arctic region, but not the Antarctic. This pattern has a number of interesting consequences for interhemispheric differences in the ecology of ice-hauling pinnipeds.

Some phocid species or subspecies occur only in aquatic habitats. Two subspecies of ringed seal occur only in Lake Saimaa, Finland, and Lake Ladoga, Russia, respectively. The Caspian seal is found only in the Caspian Sea, and the Baikal seal only in Lake Baikal, both in Russia. A population of harbor seals occurs year-round in Lake Iliamna, Alaska, but the degree of exchange, via river connection to populations of harbor seals in the nearby Bering Sea, is unknown.

Sirenia

Several of the defining anatomical and functional features of the sirenians (**Table 2**) are convergent with those of the cetaceans. However, in contrast to the cetaceans, there are few Holocene sirenians (**Table 1**). The sirenians are the only herbivorous marine mammals, sharing common ancestry with the desmostylians and the terrestrial subungulates (e.g., elephants and hyraxes).

Sirenians are large in body mass and linear dimension compared with most terrestrial mammal species. The Steller's sea cow was the largest of the Holocene species, reaching 7 m in length. Body mass of the Steller's sea cow was never directly measured, but the estimated maximum is 10,000 kg. Maximum adult lengths and masses of the three manatee species range from 3 to 4 m and 450 to 1600 kg, respectively. Adult dugongs reach maxima of 3.3 m in length and 400 kg in mass.

All modern sirenians are fully aquatic or marine, and are incapable of leaving the water. Sirenians feed on large plants

growing on the sea bottom, in midwater, at the surface, or closely overhanging the surface. They forage exclusively in shallow habitats. Manatees utilize freshwater, estuarine, and fully marine habitats, often interchangeably. They generally concentrate in habitats that are relatively warm and physically protected from extremes of weather and sea. Dependence on relatively warm water temperatures may result from the combination of obligate homeothermy and a relatively low basal metabolic rate, possibly a consequence of herbivory, compared with other marine mammals. The limited tolerance of low water temperature likely contributes to seasonal migration, and a tendency to concentrate at high density in warm water refugia during the winter. Manatees may also congregate near sources of fresh water, although fresh water is not a physiological requirement. The smallest of the modern sirenians, the Amazonian manatee, occurs only in the freshwater habitats of the Amazon River watershed of South America. Dugongs are fully marine and forage primarily on benthic seagrasses in shallow coastal tropical marine habitats. Steller's sea cow, known only from the sub-Arctic Northwest Pacific, was the most aberrant of the Holocene sirenians. Sea cows likely foraged exclusively on canopy-forming kelps and other large algae along exposed shores.

Desmostylia

Desmostylians are the only known extinct order of marine mammals. The small numbers (<10) of recognizable species in the fossil record are of Oligocene and Miocene age, and are confined geographically to the North Pacific region. Desmostylians were quadrupedal amphibians sharing common evolutionary ancestry with the sirenians. Habitats of desmostylians probably were shallow waters supporting productive populations of algae and aquatic vascular plants, their primary food, in latitudes ranging from subtropical to cool temperate. In habits and superficial morphology, desmostylians are often characterized as being similar to the modern hippopotamus. Some have argued that at least some of the desmostylians fed on clams and other benthic invertebrate prey, but the consensus view is that they were primarily, if not strictly, herbivores.

Marine Otters

There are 13 recognized extant species of otter worldwide, comprising the mustelid subfamily Lutrinae. Here we consider two species: the sea otter (*Enhydra lutris* (Linnaeus, 1758)) of the North Pacific Rim and the chungungo (*Lontra felina* (Molina, 1782)) of Peru, Chile, and southernmost Argentina. Both are marine species with amphibious characteristics. Other otter species may utilize marine environments, but they also have obligatory associations with aquatic and terrestrial habitats. The sea otter and chungungo do not appear to utilize freshwater habitats, except occasionally and facultatively.

The sea otter is arguably the most derived of the lutrines. It is the largest of the mustelids in mass, with some adult males reaching 45 kg in mass and 1.6 m in total length, but among the smallest of the marine mammals. Sea otters are relatively weak divers compared with most marine mammals, and feed almost entirely on large-bodied sessile or slow-moving benthic invertebrates. Sea otters often haul out on coastal beaches and reefs to rest and conserve heat, especially in the northern portions of their geographic range during periods of harsh weather or reduced sea surface temperature. Sea otters are not known to utilize freshwater habitats for any purpose.

The smallest of the marine mammals, the chungungos reach maxima of 6 kg in mass and 1.1 m in length as adults. Chungungos are morphologically similar to the other seven species in the genus *Lontra*. The ecological characteristics of chungungos are not well known. They feed primarily on small crustaceans, molluscs, and fish taken during dives in nearshore marine habitats along open coasts. They may also forage in fresh water, taking small crustaceans. They haul out between foraging periods on exposed rocky shores, and maintain shoreline dens that are the focal areas for social and reproductive behavior. Chungungos produce litters of two to four pups with a mean of two, and are one of only two marine mammal species (polar bears are the other) with a litter size typically greater than one.

Polar Bears

Polar bears are one of seven recognized bear species. Several of the other bear species utilize marine and aquatic habitats for foraging, but polar bears are more dependent on marine habitats for food than other bears. Along with chungungos, polar bears are perhaps the least modified morphologically, compared with terrestrial mammals, of the world's recognized marine mammal species. Although generally similar to other bears morphologically, polar bears have several distinguishing features that reflect their associations with frigid terrestrial and sea ice habitats, and with Arctic marine ecosystems. Polar bears are small compared with many marine mammals, but large compared with most other bears. Adult males reach 2.5 m in length and 800 kg in mass. Although polar bears do not dive repetitively in the manner typical of many marine mammals, they are efficient swimmers, able to traverse large expanses of open water. Polar bears also cover large distances at sea by walking or running across sea ice. Although primary prey species are pinnipeds taken from the surface on sea ice, polar bears are capable of supplementing the diet from other sources. They are the only recognized species of marine mammal that travels extensively on land away from the shoreline, and the only species that consumes both plant and animal species as a regular part of the diet. In addition, polar bears are unique among marine mammals in producing altricial young. Litter size ranges from one to three, with a mean of two, and may be linked to maternal body condition and the availability of food.

Synopsis of Evolutionary Histories of Major Marine Mammal Taxa

Cetacea

The oldest recognized cetaceans are Eocene fossils of the cetacean suborder archaeoceti. Archaeocete fossils are found primarily in rocks of present-day Egypt, Pakistan, and India, in strata thought to be associated with the Tethyan Sea of ancient times. Thus, it is thought that cetaceans originated in the Old World Tethyan environment and share common ancestry with an extinct terrestrial ungulate taxon known as the mesonychians. The earliest recognizable cetacean fossils date to

approximately 50 mybp. The archaeocetes included a number of "missing link" species, displaying intermediate forms indicating progressive reduction and loss of the hind limbs, development of the caudal spine and flukes for propulsion, elongation of the anterior skull, and simplification of dentition.

The archaeocetes were largely extinct at the beginning of the Oligocene, approximately 38 mybp. The first precursors to the modern suborders odontoceti and mysticeti appear in the fossil record during the Oligocene, but the first fossils linked unambiguously to modern cetacean families appear primarily during the Miocene. For example, the earliest sperm whales appear in early Miocene strata. Beaked whales appeared first in the middle Miocene, and the earliest dolphins and porpoises in the late Miocene of approximately 11 mybp. In the mysticetes, the earliest rorquals and right whales are also in Miocene strata. The oldest identified gray whale fossils are from the Pleistocene, and there are no known fossils providing insight into the early evolution of the modern gray whales.

Although evolutionary linkages across the history of the cetacea are uncertain at best and often entirely unclear, it is apparent from the fossil record that the modern taxa of cetaceans have been preceded by many extinct species, likely outnumbering extant species by a considerable number. For example, an early mysticete group, the family cetotheriidae, contains approximately 60 known species dating from the Oligocene. Such extinct large taxa imply significant episodes of diversification and subsequent extinction well before the modern families of cetaceans appeared. The record also suggests a dynamic pattern of biogeographic variation, such as the occurrence of ancestral monodont (the family that now includes only the belugas and narwhals of high northern latitudes) fossils in Miocene strata of Mexico. The dynamic evolutionary record almost entirely predates homonid evolution and the emergence of anthropogenic influences on patterns of extinction.

Pinnipeds

The earliest known pinnipeds are represented by fossils of the late Oligocene or early Miocene, approximately 25 mybp. The two major recognized lineages culminate in the modern phocidae and the modern superfamily otarioidea, the latter including the otariidae and the odobenidae. The earliest fossils of both lineages are of similar late Oligocene or early Miocene age. In contrast to the cetaceans, there are recognized linkages of good quality between the earliest pinniped fossils and the major modern pinniped taxa.

The traditional view of pinniped evolution is that the phocids and otarioids evolved independently. Phocids are said to have emerged in the North Atlantic region from ancestral forms linked to modern mustelids. The appearance of phocid species in the Pacific likely occurred much more recently, possibly in events related to the extinction of early Pacific otariids and odobenids. Otarioids are thought to have evolved from ancestral ursids in the North Pacific. The diphyletic model is supported by traditional analyses of cranial morphology and by the absence of early fossil otarioids from strata of the North Atlantic region. Recent analyses of postcranial material and of molecular data support an alternative model in which the pinnipeds are a monophyletic lineage sharing common ancestry with the modern mustelids. Current consensus favors the monophyletic model.

Recently published evidence indicates controversy over the affinities of the odobenidae. Some analyses indicate that the odobenids are in fact more closely related to the phocids than the otariids, whereas others favor the more traditional view, with odobenids closely allied to otariids in the otarioidea. We follow the traditional view here, but acknowledge the evidence in support of the alternative scenario.

Ancestral taxa of otariids and odobenids show a high level of diversity compared with modern forms. The recognized modern genera of odobenids and otariids appeared primarily during the late Pliocene or the Pleistocene. The fossil record for phocids is not well developed, especially in some North Atlantic regions thought to be important in understanding early evolution in the group. Miocene fossils from the southern hemisphere include significant numbers of monachine (phocid subfamily monachinae) seals, ancestral to the Holocene monk seals, elephant seals, and Antarctic ice seals. Ancestry and relationships of the more derived forms, including the modern phocine (phocid subfamily phocinae) seals of northern temperate and polar latitudes, have not been resolved definitively. As with the cetaceans, the pinniped fossil record indicates significant episodes of diversification and extinction before the emergence of the Holocene forms, and clearly predating anthropogenic influences.

Sirenians

The earliest fossil sirenians are from early Eocene strata, approximately 55 mybp. Significant radiation into at least three families had occurred at the time the earliest cetacean fossils appear in the middle Eocene. Thus the sirenians appear to be the oldest order of living marine mammals on earth. The fossil record does not provide clear evidence of the ancestral groups that gave rise to the sirenians, although studies of modern forms indicate common ancestry with the other subungulates. Although the oldest sirenian fossils are from Jamaica, the sirenians are thought to have emerged first in the Old World Tethyan environment. The early radiations of the sirenians appear to share common ancestry with the extinct family protosirenidae. The protosirenids appear to have given rise to the modern families trichechidae (manatees) and dugongidae (dugongs and sea cows). Dugongids probably first appeared in Mediterranean waters during the Eocene, whereas trichechids apparently first evolved along the South American coast during the Miocene. Sea cows first appeared in the southeastern Pacific during the early Miocene, later radiating throughout the Pacific basin. Some sea cows were unusual among sirenians in their great body size and use of relatively exposed cold-water habitats. They occurred along the coasts of California, Japan, and the sub-Arctic North Pacific Rim during the Pliocene and Pleistocene. Although the predominant sirenian family in the Holocene, the trichechids have a relatively poor fossil record. The allopatric distribution of the three modern species resulted from temporary geological isolation of the Amazonian watershed, and from a chance colonization of West African coastal waters from an ancestral Caribbean population.

Desmostylians

As noted above, the desmostylians apparently first appeared in the oligocene, approximately 35 mybp. The oldest known fossils are from the coasts of Washington and Oregon in the Northeastern Pacific, and all desmostylian fossils come from the North Pacific. The desmostylians share common ancestry with the sirenians and terrestrial subungulates, although many details of these relationships are unknown.

Marine Otters and the Polar Bear

The sea otter, chungungo, and polar bear represent three separate, relatively recent entries into marine environments by largely terrestrial or aquatic taxa. Chungungos still resemble other otters so closely that a meaningful fossil record of their evolution as marine mammals does not exist. There are some Pleistocene fossils of polar bears, but neither fossil nor modern forms represent significant departures from the ancestral ursid morphology. Thus, for both chungungos and polar bears, the history of adaptation to marine life is best inferred from modern biological data.

The sea otter appears to have a somewhat more extensive fossil ancestry, in the marine environment, than the polar bear or chungungo. There are two extinct genera of sea otters in the fossil record. Seven fossil species of *Enhydriodon* have been found in Africa and Europe, in late Miocene and Pliocene strata. *Enhydritherium lluecai* (Villalta Comella and Crusafont Pairó, 1945) is known from the late Miocene of Europe, and *Enhydritherium terraenovae* (Berta and Morgan, 1985; Berta *et al.*, 2005) from the late Miocene through the middle Pliocene in Florida and California. *Enhydritherium* is thought to be a direct ancestor of *Enhydra*. *Enhydra* is confined to the North Pacific region. The extinct *Enhydra macrodonta* (Kilmer, 1972) is known only from the late Pleistocene. The single surviving species of sea otter, *E. lutris*, dates from the early Pleistocene.

General Factors Contributing to the Vulnerability of Marine Mammals to Extinction

Obligatory Dependence on the Sea Surface for Respiration

Marine mammals must exchange respiratory gasses directly with the atmosphere, in the same manner as their terrestrial relatives. Thus, unlike marine fishes and invertebrates, marine mammals at sea are never entirely free of their association with the sea surface and must return periodically to the surface to breathe. The process of breathing at the surface is often associated with conspicuous activities such as splashing, exhalations audible over long distances, and the production of clearly visible clouds of condensed water vapor associated with exhaled gasses. Marine mammals are often physiologically obligated to remain at the surface for several minutes, allowing multiple exchanges of gas volumes contained in lungs in order to set the biochemical stage for successful subsequent dives. Human travel in watercraft at the sea surface is relatively efficient and advanced. By understanding the respiratory behavior of marine mammals and watching for signs of exhalation, humans in surface watercraft can position themselves to facilitate close contact with surfacing animals. The result is high vulnerability of marine mammals at sea to human hunters.

Large Body Mass and Linear Dimensions

The marine mammals are, on average, large compared with most other animals. The large body size probably evolved in response to certain constraints associated with life at sea, most notably those associated with thermoregulatory and hydrodynamic efficiency, foraging ecology, reproductive ecology and physiology, and habitat preference. The return to the consumer per unit of hunting effort will increase with the mean size of the prey, all other factors being equal. Among the marine mammals, all the mysticetes and many of the odontocetes, pinnipeds, and sirenians are large enough in body mass to be highly desirable as targets by human hunters. With twentieth-century refinements to the technology of marine mammal hunting at sea, pursuit of even the most mobile and dangerous of the marine mammals produced highly desired rates of economic return as long as stocks were not depleted. Thus, large body size, *per se*, increases the vulnerability of marine mammal populations to extinction simply by improving the economic return on investment in hunting activity.

Relatively High Predictability of Spatial and Temporal Distributions in Association with Regions of High Biological Productivity at Sea

With certain exceptions, marine mammals are rarely far removed from locations in which they can forage efficiently. The metabolic generation of heat, fueled by food consumption at high rates, is the only option available to marine mammals at sea for replacing heat lost continuously in a heat-consumptive environment. In species with significant annual migrations to food-poor areas for breeding, high rates of intake during the feeding season are vital for survival of extended travel and fasting as well as for maximizing reproductive fitness. In species whose breeding systems include extensive seasonal fasts, seasonal hyperphagy, requiring proximity to abundant food, may be crucial to reproductive fitness and to long-term survival of both sexes.

Most marine mammals feed on planktonic invertebrates and small schooling fishes and squids. Over the long term, seasonal and spatial patterns of production of zooplankton, forage fish, and squid are relatively predictable. Successful tracking of resources that vary predictability in space and time is vital to survival and reproductive success. Accumulated ecological data for marine mammals indicate that most populations successfully track food resource productivity most of the time. The result is an array of stereotypical seasonal and spatial movements by marine mammals that, in many cases, are readily identified. Clearly, an understanding of movements of marine mammal stocks over time reduces the investment risk in developing strategies for efficient hunting of marine mammals. Thus spatial and temporal predictabilities in marine mammal foraging facilitate efficient hunting by humans and add to the risk of anthropogenic extinction.

Impaired Mobility, Contagious Dispersion, and Temporal and Spatial Predictabilities when Hauled Out on Land

Pinnipeds, the two marine otters, and polar bears are functionally amphibious. Hauling behavior of pinnipeds is particularly synchronous and predictable, producing seasonally dense hauled-out aggregations that can be anticipated readily

in space and time. These patterns are most extreme for some of the land-breeding pinnipeds that dwell at middle or subpolar latitudes, but are prevalent in many other pinniped species as well. Two primary factors contribute to the pattern. First, good hauling sites that are near seasonally predictable foraging locations and are free of the disruptive effects of terrestrial predators such as bears or wolves are typically few in number and small in size. Pinnipeds depending on such hauling sites have evolved high site fidelity and strong navigational capabilities in order to minimize the risk that good hauling sites will be overlooked when needed for breeding, molting, and resting. Second, in the case of pinnipeds at middle and high latitudes, highly productive foraging locations near good hauling sites tend to be strongly seasonal in food availability.

Pinnipeds are awkward on land, and can be captured easily by human hunters on hauling sites, even if methods are primitive. Seasonal hauling patterns are easily recognized in space and time. As a result, human hunters can plan highly efficient hunting of hauled pinnipeds with a low risk of poor return on invested effort. Human activity on preferred hauling grounds can cause significant unintended disruption of breeding activity and social interaction among exploited pinnipeds in addition to directed harvest. The stampeding of panicked animals can lead to premature births, trampling of small pups, permanent separation of mothers from pups, disrupted dominance hierarchies, permanent abandonment of haulouts by adults, and other forms of disturbance. The net added result of hunter-associated disruptions is increased mortality and reduced birth rate in the short term. Repeated disruptions associated with human activity can lead to increased long-term risk of local extinction for populations at particular hauling sites.

Subsistence and Commercial Market Demand for Oil, Blubber, Meat, Baleen, Pelts, and Other Body Parts

By virtue of their frequently large size, morphology, physiology, and chemical composition, the harvested carcasses of marine mammals provide a number of products significant in human subsistence and commercial contexts. Harvested marine mammals provide large quantities of meat, organs, and blubber per unit of hunting effort invested. Meat and organs are used directly for human consumption or to feed domestic animals, such as sled dogs, on which human enterprise may depend. Meat may also be marketed commercially for human consumption or as pet food. Blubber is also consumed directly, but traditionally the majority of blubber is rendered and refined to products such as oil or fuel, used both for subsistence and as commercial commodities. In the case of the odontocetes, oil taken from the organs of acoustic transmission may be refinable to high-quality lubricants that, until recent decades, could not be duplicated synthetically.

Skeletal material from hunted marine mammals has been used traditionally by subsistence cultures for tools, boat construction, dwelling construction, and as raw material for handicrafts and objects of ceremonial significance. Handcrafted articles made from marine mammal skeletons often have significant commercial value as well. Mysticete baleen and the vibrissae of walruses and other pinnipeds have, in the past, been marketed commercially as components of clothing or personal toilet articles, although such uses are largely in the past. Hides and pelts may be used for clothing and for construction of boats or dwellings. The pelts of sea otters and fur seals have been sought for centuries as commodities of high commercial value for clothing, or as adornments for ceremonial robes and artifacts. Teeth and other body parts from sea otters are also known to have had ceremonial significance to indigenous peoples. Polar bears and walruses have been hunted for meat, oil, hides, and pelts by subsistence cultures for millennia, and have been targets of trophy hunters since the nineteenth century, particularly since the development of modern firearms as tools for hunting.

Much of the demand for marine mammal products centers on tissues involved in thermoregulation. The blubber of cetaceans, pinnipeds, and sirenians, and the pelts of fur seals, sea otters, and polar bears are the respective primary organs of heat retention, allowing the preservation of homeothermy in a chilling environment. Thus the homeothermic physiology of marine mammals underlies their desirability as commodities, and is a significant contributor to vulnerability of marine mammals to anthropogenic extinctions.

Low Demographic Potential for Rapid Recovery from Disturbance or Overexploitation

The marine mammals are significantly convergent in many aspects of life history. Mean litter size is one for all species, except polar bears and chungungos, and multiple births are quite rare in cetaceans, pinnipeds, sirenians, and sea otters. The age of first reproduction is often relatively high, especially in sirenians and the larger odontocetes. None of the extant marine mammals has a birth interval of less than one year, and for many species the birth interval is at least several years. In all species, parental care is entirely maternal, and the energetic costs of lactation and other forms of care are extensive for the adult female. Reproductive success of newly mature females is often low in marine mammals, increasing only with experience. Survival rates of weaned offspring may be low during the first few years of independence.

The combined effects of the above characteristics are low potential rates of growth in marine mammal populations, even when constraints of food limitation, competition, predation, or natural disturbance are alleviated. Maximum potential annual rates of population growth typically are 2–4% for the cetaceans and sirenians and 10–12% for the pinnipeds and sea otters. Exceptional cases, both higher and lower than mean rates, are known for both groups. Realized rates of growth, affected by variations in food supply and the effects of disturbance, predation, competition, and possibly other factors, often are much closer to zero and at times may reach negative values. Given these patterns, recovery from depletion associated with excessive exploitation or disturbance may require many years, and intervening additional harvest or disturbance may raise the risk of extinction to irreversible levels.

Morphological, Physiological, and Ecological Predisposition for Bioaccumulation of Lipophilic Anthropogenic Contaminants and Toxins

Most marine mammals utilize a well-developed subcutaneous blubber layer as the primary means of thermoregulation. The lipid-based blubber layer is also a primary organ for energy storage, as are fat deposits in polar bears. Many species of

cetaceans and pinnipeds, the Amazonian manatee, and some polar bears have an extensive seasonal fast each year, relating to migration away from primary feeding areas, an extended haul-out or denning period in association with reproductive activity, or shifts in habitat structure. During such fasts a significant proportion of the blubber layer or fat deposits is metabolically mobilized to meet energy and water demands during the fasting period. Following the fast, animals return to the feeding grounds and forage intensively to reconstitute the reduced blubber layer or fat deposits. In the case of the odontocete cetaceans, the acoustic melon is a second concentration of lipid-based tissue.

Odontocetes, pinnipeds, the two marine otters, and polar bears occupy high trophic levels in marine food webs. Many stable lipophilic contaminants are transmitted through food webs, such that top-level consumers may be exposed to high levels of contaminants. Such patterns are particularly well known for environmental contaminants such as organochlorines, a group including the polychlorinated biphenyls (PCBs) and the various derivatives of dichlorodiphenyltrichloroethane (DDT), and the polybrominated diphenyl ethers (PBDEs). In this context, many marine mammals face the double jeopardy of a high position in their respective food webs and, thus, the risk of high exposure to stable lipophilic contaminants, and extensive, metabolically active lipid-based tissues vital to survival, but vulnerable as sites for accumulation of contaminants. Some scientists question the demographic significance of contaminants to marine mammals, noting the relatively small number of cases of quantitative linkage of contaminant levels to alterations of population dynamics. However, in several cases, lipophilic contaminants have been correlated with reduced immune competence, disease outbreaks, and significant mass mortalities in marine mammal populations. There is also evidence that contaminants may cause endocrine disruption and reproductive malfunctions such as premature deliveries of pups. Thus, the combination of lipophilic contaminants in the marine environment, the pattern of foraging at high trophic levels by marine mammals, and the metabolically active lipid-based tissues present in many marine mammals result in increased vulnerability of marine mammals to anthropogenic extinction.

Overlap of Diet or Habitat with Commercial or Recreational Fisheries

Marine mammals often feed preferentially on prey species also sought by commercial, recreational, or subsistence fisheries, or forage in habitats in which significant commercial fishing activity occurs. These patterns create two types of problems that may enhance the vulnerability of marine mammals to anthropogenic extinction. In the first case, marine mammal populations are viewed by commercial, recreational, or subsistence fishers as competitors for a common resource. As a consequence, legal recourse may be sought to actively reduce the range or numbers of marine mammals by killing, translocation, or harassment in order to reduce the intensity of competition in favor of harvesting interests. Illegal activities may also result, including unauthorized killing or harassment intended to reduce the intensity of competition between marine mammal populations and fisheries. Such circumstances can lead to conflicting management goals by interested parties, particularly if the involved marine mammal populations are small. From the perspective of the involved marine mammal species, competition from harvesting interests may alter the quantity or distribution of food availability and may produce significant consequences at the population level.

The second type of problem is inadvertent or incidental take of marine mammals by entanglement in fishing gear. Such taking may include injury or death of individual animals, possibly producing significant effects at the population level. Potential solutions to such problems include tolerance of taking by management authority, changes to fishing techniques or effort to reduce rates of taking, or displacement of fishing effort to other locations. Such interactions may still result in population-level effects if reduction in rates of taking is inadequate or if parties affected by displacement of fishing effort resort to illegal taking of involved marine mammals as a form of retribution. Thus both types of problems may contribute to increased risks of extinction.

General Factors Hindering Effective Identification and Monitoring of Marine Mammal Populations Vulnerable to Extinction

Availability Bias

Marine mammals at sea spend most of their time below the sea surface. Depending on water clarity, typical depth of dive, angle of observation, and platform of observation (i.e., surface vessel or aircraft), a varying proportion of individual marine mammals in a field of view cannot be seen, and thus cannot be enumerated, in a population survey (e.g., Garner *et al.*, 1999). The proportion of animals not visible because of submergence is the availability bias of the survey. Availability bias reduces both accuracy and precision of population estimates, and reduces the statistical power of a survey effort to detect population trends correctly. Availability bias can be estimated with detailed information on water clarity in the survey area, diving characteristics of target species, and detection characteristics of survey observers. Estimates of bias allow the effects of the bias to be incorporated into calculations of coefficients of variation (CVs) for population estimates. Elimination of availability bias generally is not possible for surveys at sea.

Availability bias may also be a problem for surveys directed to hauled animals on shore. For example, topographic irregularities such as rocky overhangs may obscure animals that are present in the defined survey area, and animals at high density may obscure one another. Judicious timing and modified survey angles may sometimes reduce availability bias in surveys of hauled animals to nearly zero.

Observer Bias

Observer bias, also known as detection bias, results from the inability of survey observers to correctly enumerate the number of animals visible in a field of view. Like availability bias, observer bias typically is a low bias. That is, observers typically fail to see and count all animals that are present in a field of view. Observer bias has the same implications for population estimation as summarized above for availability bias. Observer

bias can be a significant source of error in both surveys at sea and surveys of hauled animals on shore. Observer bias can be reduced with increased experience of individual observers, and can be estimated using double-counting techniques with paired independent observers or by comparing observer counts in the field with counts from aerial photographs taken at the same time and place.

An important form of observer bias involves errors in estimates of group size in marine mammal surveys. Typically, marine mammal surveys involve counting of groups in the survey area. The group count is then multiplied by the mean group size, often based on a separate survey effort, as the first step in population estimation for the survey area. However, estimates of group size are themselves subject to observer bias, contributing to an increase in the CV for population estimates. Observer bias in group size estimates can be assessed under good conditions by comparing group size counts by observers with group size counts based on aerial photographs of the same group of animals taken during the surveys.

Low Statistical Power of Population Survey Data to Detect Trends in Population Size

In many cases the single most important type of information for assessing the status of a marine mammal population is the trend in population size over time. A trend is simply a time series of counts in which the slope of a fitted line is significantly different from zero. In a statistical context, the ability to detect a trend correctly is influenced by four factors. The first is the strength of the trend. Strong trends are those in which the absolute value of the slope of the fitted line is large. When other factors are constant, strong trends are more likely to be detected correctly than weak trends. The second factor is estimation error. Other factors being equal, trends in survey-based estimates of population size over time are more likely to be detected correctly if the associated CV is small than if it is large. The third factor is the number of replicate surveys available for a given time period, to be used to calculate a single estimate of population size. The probability of correctly detecting a trend increases with the number of replicate surveys used to calculate each point in the population time series, other factors being equal. The fourth factor is the number of years in which surveys are done. The probability of correctly detecting a trend, with other factors equal, is increased with an increasing number of survey years.

Many marine mammal surveys have characteristics that reduce the probability of correctly detecting trends. Weak trends may portend significant conservation concerns for marine mammals, but are inherently difficult to detect if CVs are large and replication minimal. Large CVs are common in all types of marine mammal population surveys, although CVs are gradually being reduced by the efforts of involved researchers and managers. The level of replication generally is directly dependent on levels of funding. Funding for field surveys in marine mammal science is often compromised by the challenge of limited legislative appropriations and competing priorities. Well-executed marine mammal population survey programs generally detect strong population trends successfully. However, many surveys lack the statistical power to detect weak trends that may nevertheless be important in the context of avoiding eventual extinction of marine mammal populations. The only solution is to extend survey effort over a number of years, thus improving the odds of recognizing a trend. Such an approach carries the obvious risk of potentially delaying the recognition of a significant conservation problem for the target population, and inevitably increases the overall monetary cost of the monitoring effort. In many cases there is no methodological alternative to extending the time scale of the monitoring effort.

Inadequate Understanding of Vital Demographic Parameters

Determination of effective measures to eliminate a negative trend in a marine mammal population often depends on a reasonable understanding of the demographic characteristics of the subject population. Such an understanding improves the odds that limited resources for conservation work will be applied where the greatest benefits will accrue. The demographic parameters of marine mammal populations often are known only with poor levels of accuracy or precision. In such cases, conservation effort can be readily misdirected. For example, measures to reduce preweaning mortality in a marine mammal population will be ineffective if the larger problem is a high rate of adult female mortality obscured by imprecise or inaccurate estimates of adult female survival rate. Misdirection of limited resources for conservation has obvious effects on extinction probabilities for populations in jeopardy.

Accurate, precise measurements of population parameters in marine mammals are difficult to obtain. In all marine mammal species, the time line necessary to obtain good parameter estimates is lengthy, and the best data come from studies that extend beyond a decade in duration. One of the important results of long-term demographic research on marine mammals is evidence, in some cases, of marked interannual variability in demographic parameters, further emphasizing the critical need for a long time scale in such studies. For many species good estimates of demographic parameters require tagging of individuals, an invasive process that can impose risks of reduced survival for the tagged individual and risks of disturbance to groups of animals such as pinniped breeding colonies on haulouts. In some cases, tagging has the potential to bias the demographic data recorded from the tagged individual. Finally, because of the lengthy duration and labor intensity involved, demographic studies may be quite costly and therefore difficult to implement.

Inadequate Understanding of Effects of Environmental Uncertainty on Dynamics of Populations

There is limited evidence that apparently stochastic environmental fluctuations may have important effects on the dynamics of marine mammal populations. The best-known cases involve the pinnipeds, primarily because pinnipeds breed on solid substrata and therefore have more readily observed and better-known population dynamics than most other marine mammal taxa. For example, during 1982–1983 and 1997–1998, food supplies for many temperate pinniped populations in the North Pacific were disrupted by global-scale oceanographic disturbances generally known as "El Niño – Southern Oscillation" (ENSO) events. ENSO events involve a suite of changes in global wind and ocean current patterns, with major large-scale effects on patterns of biological

productivity. Although often drastic, local changes in productivity typically are temporary, returning to normal levels over time periods from a few months to a few years. ENSO events have stochastic characteristics regarding both the frequency and the intensity of occurrence, and may be stochastic in their effects on survival rate and population trajectory of marine mammals. Both of the referenced ENSO events caused significant reduction in some pinniped population sizes, and increased rates of mortality in pups of the year. In the context of extinction, the major problem in incorporating naturally occurring stochastic events into conservation planning is the lack of good-quality data from an adequate number of events. Thus, it is not possible to generalize about effects of stochastic events, nor is it possible to reasonably consider mitigation measures to minimize increased extinction risks associated with stochastic events.

High Cost of Survey Efforts

Despite their obvious conservation value, survey efforts for marine mammal populations often are compromised or eliminated by funding constraints. The central problem is the high cost per unit effort of a good-quality survey for marine mammals at sea. Most surveys at sea use either aircraft or large surface vessels as platforms for observation. Both types of platforms are costly to operate at the level of rigor and safety necessary to obtain statistically defendable estimates of target population size. The monetary costs of good at-sea surveys may be a significant proportion of the research and management budgets of resource-oriented government agencies, even in wealthy countries such as the US. In most cases, low-cost alternatives simply do not exist. As a consequence, only the most serious issues of population trend in marine mammal conservation can reasonably attract a quantitatively rigorous level of survey effort. Thus, funding realities improve the risks of extinction by improving the odds that trends of concern will be overlooked for lack of the necessary survey effort.

Lack of Consensus and Consistency in Definition of Objective Criteria for Determination that Particular Populations are Vulnerable

There are many forms of institutional or organizational protocols designed to provide protection for depleted populations of wildlife, including marine mammals. Many such protocols originate from and have the backing of government agencies. However, few protocols for the protection of species in peril include explicit, objective criteria for determining the level of jeopardy, or for specifying recovery from jeopardy status. Moreover, those cases in which objective criteria are specified rarely accommodate the quantitative uncertainties in population estimation, demographic parameterization, characteristics of ecological disturbance, and related issues that are universal problems in population data for depleted wildlife. Surveys of criteria for status determination of wildlife in peril reveal little consistency, and often only minimal consideration of fundamental biological patterns, even within particular political jurisdictions.

Resolution of the problem of objective criteria is difficult. Different taxa of organisms have different population characteristics, and may warrant different approaches to the development of objective criteria for jeopardy and recovery. Differences in culture, values, and political traditions among jurisdictions also complicate any effort to establish common objective criteria. However, the separation of jeopardy criteria from arbitrary judgments, often intertwined with political considerations, is crucial to ensure that extinction risk is measured by biological criteria. A failure to achieve such separation may increase the odds of extinction for some populations in peril.

Problems in Distinguishing ''Natural'' from ''Unnatural'' Extinction

As indicated above, marine mammals have been present on earth since the early Eocene. The majority of the marine mammals that have evolved on earth are now extinct, and virtually all extinctions occurred before the evolution of the homonid primates, and particularly *Homo sapiens* (Linnaeus, 1758). In the context of conserving the earth's biodiversity, however, we are primarily concerned about anthropogenic loss of species. Thus, excepting material in our section Extinction or Near Extinction of Major Taxa Over Evolutionary Time (below), our time window is narrowed to the Pliocene, Pleistocene, and Holocene, and our conceptual focus to extinctions that result from conscious human action rather than from other, ''natural'' processes such as changing habitat, disturbance and catastrophe, or displacement of one species by another. We suggest that not all anthropogenic extinctions are necessarily ''unnatural''. However, separating the ''natural'' from the more philosophically distasteful ''unnatural'' is a great challenge, much burdened by the complicating considerations of values, culture, politics, and economics. Thus, the distinction, however important, is well beyond the scope of our contribution.

Patterns and Case Studies of Extinction in Marine Mammals

In this section, and particularly in sections Modern Anthropogenic Extinctions of Species, Subspecies, and Major Populations through Species, Subspecies, or Populations Once Thought to be Near Extinction, Now Showing Evidence of Increasing Numbers, we review case studies of taxa in varying degrees of peril with regard to extinction. It is not our intent to be 100% comprehensive with regard to taxa in jeopardy. Rather, our intent is to provide examples in sufficient number to illustrate the full range of factors that influence extinction risk on a global scale. In addition, geographic bias will be apparent in the cases we present. The bias results primarily from the geographic tendencies of our respective research experience and from our relatively greater familiarity with marine mammal conservation issues in the northern hemisphere, and in particular with cases from coastal regions of North America.

Extinction or Near Extinction of Major Taxa Over Evolutionary Time

Odobenidae

The odobenids have a fossil record of striking diversity, including some species resembling modern walruses and others

superficially similar to the modern sea lions and fur seals. All known odobenid fossils have been found in the northern hemisphere. The odobenids seem to have diverged from the otariids in the early Miocene. Peak fossil diversity is in strata from the late Miocene and early Pliocene. Diversity declined abruptly during the late Pliocene and Pleistocene. The single surviving species, the modern walrus, appeared in the fossil record during the Pleistocene. Extinct odobenids were also broadly distributed across latitude. The modern walrus is limited to high northern latitudes and is distributed aberrantly relative to the extinct odobenids.

The fossil record indicates that modern walruses are the single relict of a largely lost taxon. Most extinctions of odobenids seem to be associated with global-scale cooling and related large-scale habitat change during the Pliocene. However, ecological characteristics of the extinct odobenids are in many cases difficult to understand because the surviving contemporary model seems aberrant. Thus, it is difficult to develop meaningful functional models of extinction in the odobenids.

Sirenia

Sirenian diversity clearly peaked during the Miocene, a period of warm global climate coincident with extensive warm shallow coastal marine habitats. The dugongids, now represented by a single surviving species, were the most diverse family of sirenians through the fossil record. Sirenian diversity declined sharply during the relatively cool Pliocene and Pleistocene, and modern forms are relicts of a largely extinct order. Two of four recognized sirenian families are now fully extinct.

Unlike the odobenids, the surviving sirenians seem to provide good ecological models for the extinct forms. Modern sirenians typically consume macrophytes in shallow waters and clearly prefer warm protected waters. Thus, it is likely that Pliocene cooling and the coincident widespread loss of warm shallow seas were major factors in the decline of the sirenia. Surviving species were those able to retreat to low-latitude refugia or, in the case of the Steller's sea cow, a subpolar refuge with abundant food and apparently with no significant predators.

Desmostylia

Desmostylians did not diversify to nearly the extent of the other major marine mammal taxa, and did not survive beyond the end of the Miocene. Lacking modern ecological models, we prefer not to speculate on specific ecological mechanisms of extinction in the desmostylians. However, loss of the taxon coincided in time with the decline of sirenians, with which desmostylians share common ancestry. Thus habitat needs and associations may have been similar between late Miocene sirenians and desmostylians, and loss of optimal habitats during Pliocene cooling could have had similar effects on both groups.

Modern Anthropogenic Extinctions of Species, Subspecies, and Major Populations

Species Level

Following the taxonomic format of Rice (1998), four species of Holocene marine mammals are known to be extinct. In all cases the probable causes were anthropogenic. In addition, Rice (1998) reports anecdotal evidence from non-technical sources that a recent species of pinniped, undescribed and entirely unknown to science, once occurred in the Chagos Archipelago and Seychelles Islands of the tropical southwestern Indian Ocean. If such a species occurred it is now extinct as a result of unknown factors.

Baiji (Yangtze River Dolphin): Lipotes vexillifer *Miller, 1918*

The baiji, also known as the Yangtze River dolphin or Chinese River dolphin, was the sole known member of the family lipotidae. The baiji dwelled exclusively in fresh water habitats of the middle and lower Yangtze River (from Yichang eastward to the river mouth at Shanghai) and the adjacent Qiantang River in southeastern China. Accounts of population surveys are summarized by Turvey *et al.* (2007). The baiji had been recognized as imperiled for a number of decades. Baiji were not seen in the Qiantang River after the 1950s. Surveys of baiji habitat in the Yangtze River in the late 1970s suggested a population of 400 individuals. By the middle 1990s population estimates were less than 100 animals. The last authenticated observation of a live free-ranging baiji was in 2002. Turvey *et al.* reported the results of a combined intensive survey of the entire known range of the species in 2006 using multiple vessels and visual and acoustic survey methods. No animals were located and the baiji appears to be extinct in all known native habitats. There are no living baiji in captivity. Likely causes for the loss of the baiji include high rates of bycatch in fishing gear, erosion of environmental quality due to dam construction and industrial pollution, and shipstrikes. Turvey *et al.* suggest that fishery bycatch was the predominant risk factor contributing to apparent extinction. Bycatch intensity probably correlated directly with the extreme density of the human populations, and associated high demand for dietary protein, in the Yangtze watershed. The region was home to roughly 650 million people at the time of the baiji survey in 2006, approximately 10% of the earth's human population. The intensive industrialization of the Yangtze region, and the obviously diminished air and water quality characteristic of the area, likely contributed to loss of the baiji as well.

Caribbean Monk Seal: Monachus tropicalis *(Gray, 1850)*

The Caribbean monk seal was one of three recent species of *Monachus*. All occur or occurred in tropical or subtropical latitudes. The pre-exploitation range of *M. tropicalis* included the islands of the Caribbean Sea, the coastal regions of Venezuela and Caribbean Colombia, southern Florida, the east coast of Mexico south of the Bay of Campeche, and the Caribbean coasts of Belize, Guatemala, Honduras, Nicaragua, Costa Rica, and Panama. Estimates of pre-exploitation population sizes are not available. Intensive hunting of seals for meat and oil began soon after arrival of European explorers in the late fifteenth century. Seal populations were reported as depleted as early as the seventeenth century, but a few animals survived into the middle twentieth century. The last confirmed sighting of a Caribbean monk seal was at Seranilla Bank, west of Jamaica, in 1952. Directed surveys for seals in the later 1950s and 1960s found no living animals.

Japanese Sea Lion: Zalophus japonicus *(Peters, 1866)*

The Japanese sea lion is sometimes considered a subspecies (*Zalophus californianus japonicus*) of the California sea lion. Here we follow the convention of Rice (1998) in regarding the sea lion as a separate species. The Japanese sea lion originally ranged along the shores of Japan and Korea, and the southern Pacific shores of Russia. The species was subject to a long history of hunting for meat and oil. No population surveys were done. The sea lion was thought to be extinct at the end of the nineteenth century, but a group of animals was reported from the Takeshima Islands in 1951. There have been no subsequent sightings and the species is now considered extinct.

Steller's Sea Cow: Hydrodamalis gigas *(Zimmerman, 1780)*

Steller's sea cow was first observed by a scientist in 1741, at which time it occurred only along the shorelines of the Komandorskiye Ostrova, east of the Kamchatka Peninsula of Russia. At the time of discovery, sea cows probably numbered no more than a few thousand individuals in total, but population surveys were not done. There are reports that the sea cow had previously occurred along the eastern shore of mainland Kamchatka and in the Near Islands of the Aleutian Archipelago, but the primary evidence is material from stranded carcasses that could have resulted from long-distance drift at sea. Sea cows were subject to intensive hunting, soon after discovery, by crews of sea otter hunters working in the Komandorskiye and Aleutian Islands. Sea cows provided high-quality meat and blubber in great quantity and facilitated extended harvesting expeditions by otter hunters. The last sea cow observation was recorded in the Komandorskiye Ostrova in 1767. The only scientific observations of living Steller's sea cows were made by G.W. Steller at Bering Island in the Komandorskiye Ostrova. Steller recorded his observations under extremely difficult conditions while stranded during the winter of 1741–1742, following the shipwreck of the *Saint Peter*, one of two vessels comprising the Bering Expedition from Kamchatka to southern Alaska.

Three factors probably contributed to the extinction of Steller's sea cow. First, directed hunting was intensive and unmanaged. Second, at the time living sea cows were observed by Steller, they occurred only on islands lacking aboriginal human populations. The pattern suggests that aboriginal hunters may have previously reduced the range and numbers of sea cows, predisposing them to extinction when hunting intensified. Third, sea cows foraged on near-shore benthic kelps and other macroalgae. The rapid depletion of sea otter populations by otter hunters in the Komandorskiye Ostrova may have allowed sea urchins, the primary prey of otters in the region, to overgraze their preferred food, the same kelps utilized by sea cows. Thus, catastrophic loss of food supply could have facilitated the rapid extinction of sea cows.

Subspecies or Population Level

North Atlantic Gray Whale: Eschrichtius robustus *(Lilljeborg, 1861)*

Whaling records and subfossil remains indicate that a population or populations of gray whales (*E. robustus*), apparently not taxonomically distinct from North Pacific populations, once occurred along both coasts of the North Atlantic. Available evidence suggests that gray whales occurred along the Atlantic coast of North America into the seventeenth century but probably not beyond. There are no available data on population size, foraging or breeding habitats, or migratory corridors, nor is there evidence bearing on the question of genetic relatedness between gray whales off western Europe and those off eastern North America. Subfossil specimens have been found in Europe from the central Baltic coast of Sweden to Cornwall in the UK, and in North America along the US east coast from New York to South Carolina. Confirmed sightings of a single gray whale in the Mediterranean Sea in spring 2010, off Israel and Spain, are considered to be the result of navigation errors by a whale from the eastern North Pacific population, facilitated by opening of new travel corridors due to climate-based loss of sea ice cover in high northern latitudes. The most likely explanation for extinction of North Atlantic gray whale populations is prolonged excessive harvest by the whaling industry.

Species, Subspecies, and Populations in Imminent Peril of Extinction

Synopsis

Here we summarize the status of 26 marine mammal taxa or populations that in our view face a substantial probability of extinction during the twenty-first century (**Table 3**).

We provide more detailed information for five case studies of populations in this category. In all but one of the 26 cases, available data suggest total abundances of less than 1000 individuals. Most of the taxa or populations considered in this category have several significant past or current anthropogenic sources of mortality that facilitate extinction risk. In nearly all cases, the scope of operational resource investment and human societal adjustment necessary to avoid extinction is high. The recently reported extinction of the Yangtze River dolphin (*see* Baiji (Yangtze river dolphin): *Lipotes vexillifer* Miller, 1918) represents a case in point.

North Pacific and North Atlantic Populations of the Right Whale: Balaena Glacialis *Müller, 1776*

Right whales occur in three major regions, the North Pacific, the North Atlantic, and the Antarctic Ocean, together with adjoining regions of the South Pacific, South Atlantic, and South Indian Oceans. Because of intervening land masses and the antitropical distribution of the species, rates of migration and genetic exchange among the three regions are low. The primary conservation problem for the northern populations is small population sizes, imposed largely by centuries of unregulated commercial and subsistence whaling, from which recovery has not occurred. Rice (1998) considers all populations of right whales to be a single species, *B. glacialis*. Although we follow Rice's systematic treatise in this article, we note that most specialists divide right whales into more than one species. The more commonly supported protocol classifies North Pacific right whale populations as *Eubalaena japonica* (Lacépède, 1818), North Atlantic right whales as *Eubalaena glacialis*, and southern hemisphere right whales as *Eubalaena australis* (Desmoulins, 1822).

Right whales were identified by the earliest whalers as targets of choice because of their abundance, large size, high

Table 3 Species, subspecies, or populations thought to be in imminent peril of extinction

Taxon	*Population*	*Estimated population size*[a]	*Current and historical risk factors*
North Pacific right whale: *B. glacialis* Müller, 1776[b]	Eastern North Pacific	30	CH, ISW, ITF, LRN, VC
North Pacific right whale: *B. glacialis* Müller, 1776[b]	Western North Pacific	Unknown but small	CH, ITF, LRN, VC
North Atlantic right whale: *B. glacialis* Müller, 1776[b]	Eastern North Atlantic	A few animals, possibly extinct	CH, ITF, LRN, VC
North Atlantic right whale: *B. glacialis* Müller, 1776[b]	Western North Atlantic	360	CH, ITF, LRN, VC
Southern right whale: *B. glacialis* Müller, 1776[c]	Chile and Peru	A few animals	CH, ITF, LRN, VC
Bowhead whale: *B. mysticetus* Linnaeus, 1758	Davis Strait and Baffin Bay	450	CH, LRN, OW
Bowhead whale: *B. mysticetus* Linnaeus, 1758	Svalbard, Barents Sea	A few animals, possibly extinct	CH, LRN
Bowhead whale: *B. mysticetus* Linnaeus, 1758	Sea of Okhotsk	A few hundred	CH, LRN, OW
Gray whale: *E. robustus* (Lilljeborg, 1861)	Western North Pacific	130	BSR, CH, ITF, LRN, MO, VN, VC
Humpback whale: *M. novaeangliae* (Borowski, 1781)	Arabian Sea	A few hundred	ISW, ITF, LRN
Vaquita: *P. sinus* Norris and McFarland, 1958	Northern Golfo de California	500–600	ITF, LRN
North Island Hector's dolphin: *Cephalorhynchus hectori maui* (Baker *et al*., 2002)	North Island, New Zealand	120	CP, DIS, HLP, ITF, LRN VC, VS
Pacific humpback dolphin: *Sousa chinensis* (Osbeck, 1765)	Eastern Taiwan Strait	100	CP, ITF, LR, LRN
Killer whale: *O. orca* (Linneaus, 1758)	Southern resident population (Washington, USA, and British Columbia, Canada)	90–100	CFP, CP, DT, LCR, LRN, VN
Beluga whale: *D. leucas* (Pallas, 1776)	Cook Inlet & Northern Gulf of Alaska	340	CH, CP, LRN, OW, SH
Beluga whale: *D. leucas* (Pallas, 1776)	Ungava Bay, Canada	A few animals	CH, CP, LRN, OW, SH
Indian river dolphin: *Platanista gangetica* (Roxburgh, 1801)	Ganges, Indus, Brahmaputra, and Karnaphuli Rivers	Unknown	CP, DD, HLP, ITF, LRN, SH, TAM
La Plata dolphin: *Pontoporia blainvillei* (Gervais and d'Orbigny, 1844)	Uruguay and Rio Grande del Sul, Brazil	300	CFP, CH, HLP, LRN, ITF, MDI
Harbor porpoise: *Phocoena phocoena* (Linnaeus, 1758)	Black Sea and Aegean Sea	Unknown	CP, CH, DIS, LRN, MO, TAM
Harbor porpoise *Phoecoena phocoena* (Linneaus, 1758)	Northern Baltic Sea	600	CH, CP, ITF, LRN
Northern bottlenose whale: *Hyperoodon ampullatus* (Forster, 1770)	Scotian Shelf, particularly the "Gully" submarine canyon	160	CH, CP, LRN, VN
Mediterranean monk seal: *M. monachus* (Hermann, 1779)	Eastern Mediterranean Sea, northwestern Africa, and Madeira	350–450	CFP, CP, DHD, LRN, PO, SH
Hawaiian monk seal: *Monachus schauinslandi* Matschie, 1905	Hawaiian Islands west to Kure Atoll	Approximately 1000	FLI, HOF, ITF, LRN, PR
Lake Saimaa ringed seal: *P. hispida saimensis* (Nordquist, 1899)	Lake Saimaa, Finland	270	ITF, LRN, OW
Ungava seal: *Phoca vitulina mellonae* Doutt, 1942	Ungava Peninsula, Quebéc, Canada	100–600	LRN
Chungungo: *Lutra felina* (Molina, 1782)	All	Less than 1000	CH, CP, IHC, ITF, LCR, PO

[a]Indicated estimates of population size reflect a wide range of quantitative rigor. Primary sources for estimates shown are online resources associated with NMFS, FWS, and RLTS. The responsibility for errors rests exclusively with the authors of this article.

[b]Some taxonomists consider the North Pacific right whale a separate species, *Balaena japonica* (Lacépède, 1818). Here we follow the convention of Rice (1998), regarding the North Pacific and North Atlantic right whales as one species.

[c]Some taxonomists consider the southern right whale a separate species, *Balaena australis* (Desmoulins, 1822). Here we follow the convention of Rice (1998), regarding the North Pacific and North Atlantic right whales as one species.

Codes for risk factors are as follows (codes also apply in Tables 4 and 5): BSR, biased sex ratio; CFP, competition with fisheries for prey; CH, commercial harvest; CP, contaminants and pollution; DD, dams and diversions; DHD, deliberate habitat destruction; DIS, diseases; DT, disturbance by tourism activities; FLI, food limitation from oceanographic anomalies; HLP, habitat loss for prey; HOF, haulout flooding from sea level rise; IHC, illegal harvest to reduce competition for prey with fisheries; ISW, illegal Soviet whaling; ITF, incidental take in fisheries; LCR, live-capture harvest; LR, land reclamation; LRN, limited range or low numbers; MDI, marine debris ingestion or entanglement; MO, marine oil/gas exploration, production, or transportation; OW, ocean warming and sea ice reduction; PO, poaching; PR, predators; SH, subsistence harvest; TAM, terrestrial agriculture, mining, deforestation, and urbanization; VC, vessel collisions; VN, vessel noise.

yields of meat and blubber, relatively docile behavior, and relative buoyancy *post mortem*. Historical records suggest that significant exploitation of the North Atlantic populations began along the European coast early in the second millennium AD. Changes in hunting effort over time and space followed the stereotypical pattern of overexploitation. As coastal European stocks of whales were depleted, whalers expanded efforts westward to Greenland, Newfoundland, and Labrador beginning in the sixteenth century. Subsequent stock depletion led to further expansion of harvest effort southward in the seventeenth century to the waters of what are now Nova Scotia, and the US right whale populations off the northeastern US were depleted after the middle eighteenth century and may have numbered as few as 20 individuals. The historical record of right whale harvest is less lengthy for the North Pacific, but almost certainly followed the same general pattern. Commercial whaling was most intensive in the nineteenth century. Recovery of North Pacific right whales during the twentieth century was damaged by substantial illegal harvest by whalers from the USSR during the 1960s. North Atlantic, North Pacific, and southern hemisphere right whales are listed as "endangered" throughout their respective geographic ranges, pursuant to the US Endangered Species Act (ESA) of 1973 (16 USC §§1531–44) as amended (ESA). The Red List of Threatened Species (RLTS), managed by the International Union for the Conservation of Nature, shows North Atlantic and western North Pacific right whale populations as "endangered" and the eastern North Pacific population as "critically endangered".

Between 360 and 400 individual right whales are thought to be present currently in the North Atlantic. The eastern population, off the coast of Europe, probably contains only a few individuals and is in extreme jeopardy of extinction. It is possible that the eastern population is effectively extinct, and that right whales occasionally seen in the region are strays from the western population. The western population ranges along the coast of Canada and the US and is thought to contain at least 360 individuals. During the twentieth century, estimated annual growth rates of the western population never exceeded 2.5%. The two populations recognized in the North Pacific are also quite small. The western population, ranging from the western Bering Sea to southern Japan, is thought to be somewhat larger than the eastern population, which ranges from the eastern Bering Sea to southern Baja California, Mexico. The eastern population numbers approximately 30 individuals and may have suffered irreparable damage from the illegal Soviet whaling noted above. An estimate of trend in numbers in the eastern population has not been made. The eastern North Pacific right whale population is almost certainly the most critically endangered of the world's large whale populations for which population size estimates are available.

Failure of northern right whale populations to recover from the cessation of commercial whaling has been difficult to understand, except the obvious damage from illegal whaling in the North Pacific in the twentieth century. Southern right whale populations protected from exploitation have grown at rates estimated as high as 6–7% per year, and there is no evidence of major differences in genetically defined vital rates between northern and southern populations. Thus, other factors must be retarding population growth in the north. Northern populations of right whales tend to be concentrated in regions that support productive and highly capitalized fisheries, facilitating damaging rates of incidental entanglement. Right whales are also concentrated in areas frequently transited by large ships, facilitating lethal ship strikes. Vulnerability to shipstrikes is enhanced by the apparent tendency of right whales to rest quietly at the surface for long periods, in a manner minimizing the probability of detection by vessel operators. Survival of northern right whale populations beyond the twenty-first century requires continued prohibition of all forms of directed harvest and the largest possible reductions in rates of incidental taking and shipstrikes. Eastern populations in the North Atlantic and North Pacific may be destined for extinction within the century regardless of actions taken.

Western North Pacific Gray Whale: E. robustus *(Lilljeborg, 1861)*

The western North Pacific population of gray whales probably originally summered in the Sea of Okhotsk, and migrated south to winter habitats in southern China. Focused ecological studies of the population have been done only within the last two decades. The western gray whale population is among the smallest of large whale populations in the world and, as with the eastern North Pacific right whale, is highly vulnerable to extinction in the twenty-first century. The principal conservation concerns for western gray whales are the small population size resulting from commercial whaling in previous decades, incidental take in Japanese net fisheries during migration, a significant male bias in population sex ratio, a lack of knowledge regarding migratory corridors and destinations, and risks posed by offshore oil development in the current summer feeding range of the population. The western gray whale population is listed as "endangered" as per ESA, and "critically endangered" as per RLTS.

Western gray whales forage on benthic invertebrates during summer months in shallow coastal waters off northeastern Sakhalin Island, Russia. Foraging activity is particularly intensive in a small near-shore area off the entrance channel to a large coastal lagoon, Zaliv Pil'tun (52°50′ N, 143°20′ E). Data from photoidentification studies indicate high fidelity of foraging gray whales to the area off northeastern Sakhalin, particularly near Zaliv Pil'tun, within and among years since 1995. Adult females with calves show particularly high fidelity to the feeding area near Zaliv Pil'tun, suggesting that the area is important to calf rearing. In recent years, small numbers of western gray whales have been observed during summer at near-shore locations in the southwestern Bering Sea.

Previous estimates of the western gray whale population, based on minimal data, were in the range of 200–300 individuals. Intensive recent studies have focused on photographic identification of individual whales, exploiting individual-specific markings on the skin. Application of population models to the photoidentification data suggests a current population size of approximately 130 individuals. The photoidentification data coupled with other direct field observations also indicate a significant male bias in sex ratio (male:female ratio ~ 0.65) for unknown causes. The bias in sex ratio influences effective population size and could reduce the growth potential of the population.

Six gray whales have been reported killed by entanglement in Japanese fishing gear, and a seventh in Chinese fishing gear, with five of the events occurring since 1995. Most entanglements involve a type of set net known among Japanese fishers as the "fixed shore net." It is likely that the whales died during migration, as all events occurred well south of known summer foraging areas. Given the small population size and the recognized male bias in western gray whales, loss of even a few animals to bycatch can cause significant negative effects on population growth potential and the risk of population-scale extinction. It was recognized in the early 2000s that Japanese legal protocols may have been contributing to bycatch risk by permitting fishers to sell meat or other products from bycaught whales. In 2008, new regulations were implemented that prohibit sale of products from bycaught whales and encourage fishers to release entangled whales alive when feasible. Sufficient time has not yet passed to determine the effectiveness of the new protocols.

Efforts are underway to determine migratory corridors and calving grounds for western gray whales, a reflection of intensive interest by the conservation community to identify the unknown portions of western gray whale distribution for purposes of eventual additional protection for the whales, in habitats critical to life history and survival. Satellite telemetry tags were attached to whales beginning in 2010, with continued effort in 2011. A surprising early development was the movement of an adult male from the Sakhalin Island region to the coastline of North America in autumn 2010, where it apparently joined the typical southbound autumn migration of eastern gray whales. Although the animal's transmitter failed in 2010 while it was southbound off the coast of Oregon, photoidentification surveys identified the same whale back at Sakhalin in summer 2011. Individuals in the western gray whale population are known to be genetically distinguishable from eastern gray whales, suggesting that exchange of breeding individuals between populations is uncommon. However, recent analyses involving geographic records for photographically identified individual whales indicate that other cases of movement between North Pacific populations have occurred in recent years. Additional tagging data will place the observed movements into perspective as well as serving the original purpose of locating migratory pathways and wintering habitats.

The summer feeding area near Zaliv Pil'tun has been subject to intensive exploration and development of marine oil reserves for the past two decades. Offshore petroleum exploration involves frequent use of "seismic surveys" that generate intensive low-frequency sounds, and development of located petroleum resources involves increased shipping activity, placement of structures, modification of adjacent sediments and benthic communities, and the risk of unintended spills of drilling fluids, vessel and machinery fuels, and extracted crude oil. The area now supports two large offshore production platforms linked to onshore infrastructure with subsea pipelines, and intensive shipping activity in support of petroleum extraction and transport. The possible addition of a third offshore production platform near Zaliv Pil'tun was announced by the industry in 2011.

Recovery planning for the western North Pacific gray whale population will benefit from acquisition of significant additional data. There is a clear need for determination of winter range, calving grounds, migratory corridors, and improved insight into rates of genetic exchange with eastern North Pacific gray whales. An understanding of causes for the bias in population sex ratio will inform more effective and pragmatic population modeling. An improved understanding of the affinity of summering animals, and particularly females with dependent calves, for the feeding grounds off Zaliv Pil'tun will also support more effective conservation planning, as will ongoing critical assessments of activities associated with development of petroleum resources in the coastal zone of northeastern Sakhalin Island. Risks of extinction will be reduced significantly if observed rates of bycatch in fisheries along apparent migratory corridors can be eliminated.

Gulf of Alaska Beluga Whale: Delphinapterus leucas *(Pallas, 1776)*

The Gulf of Alaska population of beluga whales is normally concentrated in upper (northern) Cook Inlet on the southern mainland coast of Alaska. Individuals from this population have been seen on occasion in coastal waters of the Kodiak Archipelago, in Prince William Sound and Yakutat Bay, and along the outer coastal waters of the central Gulf of Alaska. Historically, belugas have been widely distributed within Cook Inlet, a turbid estuary subject to extreme tidal mixing and substantial freshwater input from several large river systems. The whales concentrate seasonally near and in the river mouths to forage on migrating salmon.

Belugas in Cook Inlet were hunted in previous decades by commercial whalers, although commercial harvest no longer occurs. Passage and implementation of MMPA in 1972 assured the rights of native peoples in Alaska to pursue the traditional practice of hunting beluga whales, including those in Cook Inlet. Hunting of belugas in Cook Inlet by natives increased in the 1980s and 1990s compared with earlier decades. During the 1990s, it was recognized that the geographic range of beluga whales, based on opportunistic observations from research vessels, was much smaller than it had been in the 1970s. No sightings of belugas in other locations of the Gulf of Alaska have been made in the recent years. Surveys in the 1990s confirmed that the population was declining. Recent survey data suggest a population size of approximately 340 individuals with no apparent trend in numbers over time. Based on known whale kills by native hunters and evaluation of survey data, it was determined that native harvest of belugas may have been excessive, possibly placing the Cook Inlet beluga population at risk of extinction. The population was listed as "endangered" in 2008 as per ESA, and is designated "critically endangered" as per RLTS. Native tribal organizations and US federal agencies are now planning for recovery actions and a more carefully regulated harvest under a comanagement framework that includes threshold population numbers, below which harvests will be prohibited.

The Cook Inlet beluga population is the first known modern case in which excessive harvest by indigenous North American peoples has placed a population of marine mammals in jeopardy. Other populations of belugas are exploited regularly by natives in Alaska, but none are in peril. The tenuous status of Cook Inlet belugas, compared with other Alaskan populations, probably results from three factors. First, the Cook Inlet

population probably has always been the smallest of the Alaskan beluga populations. Second, Cook Inlet belugas were subject to commercial whaling in previous decades. Third, the cultural characteristics of native harvest of belugas in Cook Inlet differ from all other locations in Alaska where native whaling occurs. In all cases except Cook Inlet, native whaling is done from single coastal villages, using practices consistent with lengthy tradition, subject to stringent oversight by village elders, and based on the subsistence needs of the village. In contrast, most native whalers working in Cook Inlet have moved from their home villages to the relatively large and secular city of Anchorage. Thus it appears that cultural norms and limits on whaling activity characteristic of native Alaskan villages have been lost in the case of Cook Inlet. A cautious and conservative comanagement approach to future harvests will best serve the sustainability of the Cook Inlet beluga population. Otherwise extinction is likely within the century.

Vaquita: Phocoena sinus *Norris and McFarland, 1958*

The vaquita is a small porpoise limited to the northern part of Golfo de California, Mexico. The vaquita has the smallest natural geographic range of any marine cetacean, and the single population probably has never contained a large number of individuals. Numbers are estimated at 500–600 individuals (Table 3). The population biology of the vaquita is poorly known. Even the most rigorous surveys of the vaquita population have low precision. Thus, trends in the population are difficult to discern with confidence. Available data suggest that the population may be declining at rates up to 15% per year. The vaquita is listed as "endangered" as per ESA and "critically endangered" as per RLTS.

In recent decades the primary concerns for vaquita conservation have been high rates of incidental take in gillnet and shrimp trawl fisheries, effects of contaminants, effects of reduced genetic diversity, and effects of diversion of the Colorado River away from the northern Golfo de California. Losses due to incidental take are currently viewed as the primary conservation problem. Concentrations of organochlorines, including PCBs and DDT congeners, are known to be low in vaquita tissues, low in other consumer species within the vaquita range, and low in the vaquita habitat. The levels of reduced genetic variability in vaquitas do not necessarily result in genetically based reduction of reproductive rates, particularly in the context of populations such as vaquitas that probably have always been small. Biological productivity in the upper Golfo de California is high compared with other coastal marine ecosystems despite diversion of the Colorado River.

Continued incidental take of vaquitas will likely cause extinction of the species during the twenty-first century, given even the most conservative estimates of the current rate of take. Elimination of the risk of anthropogenic extinction requires significant reduction in the level of fishing effort, or changes in gear design or deployment strategy to reduce take rates, in fisheries responsible for incidental take.

Mediterranean Monk Seal: Monachus monachus *(Hermann, 1779)*

The Mediterranean monk seal was found originally in the western Black Sea, throughout the Mediterranean Sea, and along the coast of northwestern Africa from the Strait of Gibralter to approximately 34° N latitude. At present, total numbers are estimated at approximately 600 seals, occurring primarily in two populations. One occurs in the eastern Atlantic on the Island of Madeira and along the coasts of Western Sahara and Mauritania in northwestern Africa. The second population occurs in the eastern Mediterranean, primarily off coastal Turkey and Greece. The largest concentration of seals currently known dwells in the coastal waters of Greece. Mediterranean monk seals are listed as "endangered" as per ESA and "critically endangered" as per RLTS.

Mediterranean monk seals probably have been subject to directed subsistence harvest for meat, oil, and hides for several millennia. The precarious status of modern populations seems to result from a number of factors associated with the large, multicultural human populations of southern and eastern Europe and northern Africa. For some years, monk seals have been perceived as direct competitors of fisheries, and have been harassed and killed in substantial numbers, often illegally, as a result. Harassment has included directed destruction of caves and other shoreline locations favored by seals for breeding and resting. Monk seals probably also have been affected by loss of prey due to overfishing and to various forms of contamination of habitats and food webs. In 1997 a mass mortality event was observed at a previously significant seal concentration at Cabo Blanco, near the border of Western Sahara and Mauritania, from an unknown cause. Although comprehensive demographic and population survey data are lacking, the consensus view is that the total number of Mediterranean monk seals is probably declining over time.

In addition to the small size of the two known populations, two factors add great difficulty to the prospects for implementation of a successful recovery strategy for Mediterranean monk seals. First, the habitat of the monk seal is bounded by a large number of culturally disparate political jurisdictions. Historically, political and cultural diversity in the region has interfered with cooperation among jurisdictions. Thus the attainment of consistent broadly supported conservation priorities for monk seals may be an unrealistic political objective. Second, ongoing damage to the monk seal populations apparently results from a number of factors acting in concert rather than one clearly pre-eminent problem. Thus, agreement on conservation priorities and actions may be difficult even within jurisdictions.

Mediterranean monk seals appear to be destined for extinction, possibly within the twenty-first century, unless marine conservation authorities in countries bordering seal habitat can agree in two areas. First, risk factors for the seals must be evaluated dispassionately and placed in order of significance. Second, involved authorities must agree on a plan for recovery of seal populations based on the assessment of risk factors and convince the human populations of their respective jurisdictions that seal conservation is a worthwhile objective.

Species, Subspecies, and Populations of Significant Concern with Regard to Extinction

Synopsis

Here we consider taxa or populations that are of significant concern with regard to extinction (Table 4). In contrast to the

Table 4 Species, subspecies, and populations of significant concern with regard to extinction

Taxon	*Population*	*Estimated population size*[a]	*Current and historical risk factors*
Southern right whale: *B. glacialis* Müller, 1776[b]	All	7000	CH, ITF, VC
Antarctic blue whale: *B. m. intermedia* Burmeister, 1871	Southern Ocean	Unknown	CH
Fin whale: *Balaenoptera physalus* (Linnaeus, 1758)	Northern hemisphere	40,000	CH
Fin whale: *B. physalus* (Linnaeus, 1758)	Southern hemisphere	20,000	CH
Sei whale: *Balaenoptera borealis* Lesson, 1828	All	65,000	CH
Humpback whale: *M. novaeangliae* (Borowski, 1781)	Oceania (Eastern Australia, Tonga, New Caledonia, Cook Islands, French Polynesia)	11,000	CH, ISW, SH
Hector's dolphin: *Cephalorhynchus hectori* (van Bénéden, 1881)	All	7500	CP, DIS, ITF, LRN, VC, VS
Atlantic humpback dolphin: *Sousa teuszi* (Kükenthal, 1892)	All	A few thousand	CP, HLP, ITF, LRN, SH,VC
Black Sea bottlenose dolphin: *Tursiops truncatus* (Montagu, 1821)	All	A few thousand	CP, CFP, CH, DIS, ITF, LCH, LRN
Irriawaddy dolphin: *Orcaella brevirostris* Owen *in* Gray, 1866	All	Unknown	CP, DD, ITF, LCR, LRN, TAM
Yangtze River finless porpoise: *Neophocoena phocoenoides asiaorientalis* (Pilleri and Gihr, 1972)	All	Unknown	CP, CFP, ITF, HLP, LRN
Baltic Sea ringed seal: *P. hispida botnica* (Gmelin, 1788)	All	Unknown	CP, DIS, ITF, LRN, MO, OW, SH, VN
Caspian seal: *Pusa caspica* (Gmelin, 1788)	All	111,000	CFP, CH, CP, DIS, ITF, OW, PO, PR
Lake Ladoga ringed seal: *P. hispida ladogensis* (Nordquist, 1899)	All	Unknown	CP, LRN, OW, SH, VN
Western North Pacific harbor seal: *P. vitulina stejnegeri* Allen, 1902	All	6500–7500	ICF, ITF
Steller's sea lion: *E. jubatus* (Schreber, 1776)	Western North Pacific	58,000–66,000	CFP, FLI, IHC, ITF, PR
Australian sea lion: *Neophoca cinerea* (Péron, 1816)	All	13,800	CH, DT, ITF, SH
Hooker's sea lion: *Phocarctos hookeri* (Gray, 1844)	All	12,000	CFP, CH, DIS, DT, ITF, LRN, SH
Guadalupe fur seal: *Arctocephalus townsendi* Merriam, 1897	All	7000	CH, LRN
Juan Fernandez fur seal: *Arctocephalus philippii* (Peters, 1866)	All	12,000	CH, ITF, LRN, MDI, PO
Atlantic walrus: *O. rosmarus rosmarus* (Linnaeus, 1758)	All	20,000	CH, CP, DIS, OW, PO, SH
Laptev Sea walrus: *O. rosmarus laptevi* Chapskii, 1940	All	4000–5000	CH, CP, OW
Amazonian manatee: *Trichechus inunguis* (Natterer, 1883)	All	Unknown	CH, CP, DD, ITF, LRN, SH, TAM
West African manatee: *Trichechus senegalensis* Link, 1795	All	Less than 10,000	CH, PO, SH, TAM
West Indian manatee: *T. manatus* Linnaeus, 1758	All	6400–9400	CH, CP, DD, ITF, MDI, SH, TAM, VC, VN
California sea otter: *E. l. nereis* (Merriam, 1904)	All	2800	CP, CH. DIS, ITF, MO, PR
Northern sea otter *E. l. kenyoni* Wilson, 1991	All	78,000	CP, CH, ITF, MO, PR

[a]Indicated estimates of population size reflect a wide range of quantitative rigor. Primary sources for estimates shown are online resources associated with NMFS, FWS, and RLTS. The responsibility for errors rests exclusively with the authors of this article.

[b]Some taxonomists consider the southern right whale a separate species, *B.australis* (Desmoulins, 1822). Here, we follow the convention of Rice (1998), regarding the northern and southern right whales as one species.

Codes for risk factors are as shown in Table 3.

section Species, Subspecies, and Populations in Imminent Peril of Extinction, we do not regard entries in this group to be in imminent peril of extinction during the twenty-first century. In most cases, population sizes are large enough and conservation issues tractable enough that less dire predictions seem more reasonable. Proactive management will be required to prevent taxa and populations in this group from reaching a more precarious status. We list 27 taxa or populations (**Table 4**), providing more detailed summaries for five arbitrarily selected examples.

Blue Whale: Balaenoptera musculus *Linnaeus, 1758*

Blue whales occur in coastal and pelagic marine habitats worldwide. Currently three subspecies are known. The pygmy blue whale (*B. musculus brevicauda* (Ichihara, 1966)) occurs in southern cool temperate and subpolar latitudes. The Antarctic blue whale (*B. musculus intermedia* (Burmeister, 1871)) summers in the Antarctic Ocean, and the northern blue whale (*B. musculus musculus* (Linnaeus, 1758)) is found in the North Pacific and North Atlantic. Eastern and western populations are known in the North Atlantic, and at least five populations have been described in the North Pacific. The population structure in the southern hemisphere is unclear. All blue whale populations and subspecies are listed as "endangered" pursuant to ESA and "endangered" as per RLTS, throughout their respective geographic ranges. Available data suggest that blue whales migrate seasonally, utilizing higher latitude habitats in summer for feeding primarily on euphausiid crustaceans ("krill"), and lower-latitude habitats during winter for courtship, breeding, and parturition. The spatial scope of blue whale migrations appears less than those of iconic migrating species such as gray and humpback whales. In addition, there is evidence that individual blue whales may transit the full length of their migratory corridors asynchronously several times annually in response to spatial and temporal variations in prey availability.

Blue whales are an obvious target of choice for commercial hunting because of their great body size. However, blue whales are swift swimmers and are negatively buoyant *post mortem*. Before exploitation, blue whales were most abundant in the Antarctic Ocean and other marine habitats distant from human population centers. Thus, they were beyond the technological capabilities of commercial whalers before the twentieth century. Blue whales became priority targets of whalers only after development of the steam engine, factory ships, explosive harpoons, and air compressors to inflate carcasses after killing. Thus the harvest and depletion of blue whales occurred primarily during the twentieth century. At least 360,000 blue whales were killed in the Antarctic region before commercial whaling declined in the 1960s, and other populations were exploited as well. Illegal harvests by Soviet whalers occurred after a moratorium was imposed on blue whale catches by the International Whaling Commission (IWC) in 1965. At least 8000 additional pygmy blue whales were taken.

The consensus view at present is that blue whale populations are now but a fraction of pre-exploitation size. Recent data indicate that the combined North Pacific populations number approximately 3300 individuals. The combined North Atlantic populations have been estimated at 100–600 individuals. The estimated numbers of blue whales in southern hemisphere habitats range from 400 to 1400 (all subspecies combined). There are published arguments that blue whale populations off southern Japan, in the eastern Gulf of Alaska, and off northern Norway are locally extinct or quite small. In contrast, blue whale numbers seem to have increased rapidly off northern California since the 1970s, suggesting that redistribution may be a contributing factor to regional trends. Because of small population sizes and large CV associated with surveys, it is often not possible to identify trends in blue whale populations. The survey effort required to collect data sufficiently powerful for trend detection is prohibitively costly. Thus, the status of the world's blue whale populations is generally poorly known, and prospects for confident understanding of status of populations in the foreseeable future are limited.

The current status of blue whale populations seems to be largely the result of excessive past commercial harvests, including illegal Soviet whaling. There is evidence that shipstrikes and entanglement in fishing gear may cause some mortality of blue whales in some coastal habitats, but current anthropogenic mortalities are probably minimal. Recovery of small blue whale populations may require indefinite suspension of all forms of harvest, and prompt detection and elimination of emerging sources of anthropogenic mortality.

Western North Pacific Steller's Sea Lion: Eumetopius jubatus *(Schreber, 1776)*

Steller's sea lions occur in coastal waters of the North Pacific Rim from southern California to northern Japan, and in the Bering Sea. In body mass they are the largest of the world's otariids, with males averaging 1000 kg. Genetic data have been used as the basis for dividing the species into two populations, western and eastern, with the boundary at Cape Suckling (144° W longitude), Alaska. Both populations are designated "endangered" as per RLTS. The eastern population is dispersed along the west coast of North America, numbers approximately 52,000 individuals, is increasing slowly overall, and is listed as "threatened" as per ESA. The western population numbered approximately 150,000 animals in the 1950s, but has since declined precipitously, with current total numbers, including US and Russian habitats, estimated at between 58,000 and 66,000 individuals. The rate of decline has varied over time, with the highest rates (approximately 15% per year) from 1985 until 1990. The western population is listed as "endangered" as per ESA. The population is currently thought to be stable overall, but with significant variation in trends among locations within the range of the population. The same local-scale variance in trend is known for the eastern population.

The cause or causes of long-term decline in the western population of Steller's sea lions are not understood. Possible risk factors include incidental take in fishing gear, competition with fisheries for prey in common, hunting by indigenous peoples, illegal hunting or harassment, inadvertent rookery disturbance, consumption by killer whales, disease or parasitism, contaminants, and changes in the structure and productivity in the marine ecosystems of which Steller's sea lions are part. Based on extensive research since the decline was first recognized, the consensus view at present is that ecosystem

change and competition with fisheries are the factors most likely driving the decline. Resolution of the question of cause has become the focus of intensive political interest because of the potential economic consequences of resulting recovery actions by management authorities. The groundfish fisheries of the Gulf of Alaska and the Bering Sea are the most valuable and highly capitalized of the fisheries in the coastal waters of the US. Steller's sea lions feed extensively on groundfish species, such as walleye pollock, targeted by fisheries. Determination that competition with fisheries is contributing to the decline could result in forced reduction of fishing effort, with great economic loss and political discord.

Several lines of evidence favor the argument that ecosystem change has contributed to the decline in sea lion numbers. Prey species composition has changed and dietary diversity has declined, and there have been measurable shifts in the spatial distributions of preferred prey over the past three decades. Numbers of other piscivores in the habitat, including seabirds and harbor seals, have declined over a similar time scale in many parts of the sea lion range. However, fishing activity may be intensive near important breeding locations, and it has not been possible to eliminate competition with fisheries as a potential cause of the decline. Recent rulings in US federal courts have restricted fishing activity in some locations and seasons for the purpose of reducing the rate of decline, intensifying the associated political debate. Because natural oceanographic changes can neither be predicted nor controlled, management authorities may have no choice but to focus on understanding and minimizing anthropogenic risk factors, despite the political consequences, in order to reduce the probability of eventual extinction.

West Indian Manatee: Trichechus manatus *Linnaeus, 1758*

The West Indian manatee occurs in coastal habitats and the lower reaches of rivers in the southeastern US, the Islands of the Caribbean Sea, and the mainland shores of the Gulf of Mexico, the Caribbean coast of Central America, and northeastern South America. Two subspecies are recognized. The Florida manatee (*T. manatus latirostris* (Harlan, 1824)) is found in US coastal waters, especially in Florida. The Antillean manatee (*T. manatus manatus* (Linnaeus, 1758)) occupies the remainder of the range of the species. The current population in Florida is the largest of the species, numbering approximately 3800 individuals, with annual rates of increase in the range of 2–4%. Population structure, numbers, and trends in other locations are not well known, with estimates of total numbers of Antillean manatees ranging from 2600 to 5600 individuals. The species is listed as "endangered" as per ESA and "vulnerable" as per RLTS throughout its geographic range.

Manatees have been hunted by indigenous peoples for meat and other products for centuries. Commercial hunting probably contributed to a reduction of population sizes, but there are few data available to assess the rates or the significance of commercial harvest. Manatees have been protected from all forms of directed harvest in Florida since the 1960s, but subsistence harvest continues in other populations, for meat and oil and for ceremonial purposes. During the twentieth century a number of risk factors for manatees emerged, all in association with an expanding human population in immediate proximity to manatee habitat. The primary problems are increased levels of contaminants and a variety of modifications and ongoing disturbances of habitat. The latter include watercourse diversions and impoundments in aquatic manatee habitat, and disturbance and collision risk associated with recreational boating.

Long-term studies of stranded carcasses indicate three major sources of mortality in Florida manatees. Perinatal mortalities involve newborn animals with the proximate cause of death uncertain, but possibly linked to contaminants. Significant mortality rates are also associated with entrapment in dam floodgates, entanglements in fishing gear, and collisions with recreational powerboats. Manatees clearly prefer areas without significant powerboat traffic, and the ongoing expansion of human populations and associated demand for recreational opportunities inevitably lead to continuing reduction in the size of available manatee habitat.

Manatees are particularly sensitive to water temperature compared with other marine mammals and typically congregate in warm water refugia during winter, especially when ambient sea surface temperatures drop below approximately 18 °C. Refugia include natural warm aquatic springs, and lagoons associated with thermal effluent from electrical generating plants. Natural refugia for manatees are few in number and generally small in size, such that their value to manatee conservation may be highly sensitive to threats from encroaching effects of human development and activity. Use of power plant lagoons may be risky if plant operations dictate precipitous shutdown, resulting in cutoff of thermal effluent and rapid chilling of lagoons. Manatees in Florida have experienced several significant mass mortalities in the recent decades, possibly resulting from an interaction of contaminants, compromised immune systems, disease, and natural toxins associated with harmful phytoplankton blooms.

Habitat loss and disturbance are the primary conservation problems for all populations of West Indian manatees, and subsistence or illegal harvest remain significant risk factors for populations outside of Florida. Over the long term, avoidance of extinction may require cessation of all forms of human harvest and an effective, broadly supported strategy for balancing the habitat needs of manatees with the consequences of human population growth, especially in Florida.

California Sea Otter: E. lutris nereis *(Merriam, 1904)*

Sea otters originally ranged throughout the coastal waters of California, including San Francisco Bay and the southern California islands, and probably numbered 16,000–20,000 individuals before the onset of commercial-scale harvests for pelts in the eighteenth century. The population originally extended southward into mainland Baja California and Isla Guadalupe, Mexico, and was contiguous with other otter populations ranging through the North Pacific Rim to northern Japan. Sea otters in California comprise one of three subspecies currently recognized.

Sea otters probably have been hunted by the indigenous peoples of the North Pacific Rim for several millennia, for meat and pelts and for ceremonial purposes. Observations of the Bering Expedition of 1741–1742 and other voyages of exploration in the North Pacific found sea otters to be abundant throughout their range, and commercial harvest of sea

otters for pelts began soon after. In California the harvest was pursued by hunters from Russia, often utilizing enslaved Aleut hunters, and from Spain, Mexico, and the US. All commercial hunts for sea otters were terminated with approval of the Treaty for the Preservation and Protection of Fur Seals (37 Stat. 1542, T.S. No. 542) by Japan, Russia, Great Britain (on behalf of Canada), and the US. The Treaty was proclaimed in 1911 and ratified in 1912. Article V of the Treaty afforded protection to sea otters from further hunting. Protection of sea otters from hunting in US waters was extended by Presidential Proclamation in 1913 by US President Wilson, and by Executive Order in 1929 by President Coolidge. The State of California enacted regulations to prohibit hunting of sea otters in 1913. However, in California and Mexico the sea otter populations had been depleted commercially by the 1860s. By 1900 only two small populations survived. One was along the Big Sur coast south of the Monterey Peninsula, numbering approximately 50–100 animals. The second, of unknown size off the Islas San Benitos, Mexico, was extinct by the 1920s. The Big Sur population grew at a rate of approximately 5% per year through much of the twentieth century and now numbers approximately 2800 animals, ranging from near Half Moon Bay in San Mateo County southward to the vicinity of Cojo Anchorage in Santa Barbara County. The observed rate of growth in California is currently approximately 3% per year, much lower than rates commonly observed in more northerly populations with protection from harvest and available adjacent vacant habitat. In the late 1980s a new colony was established by translocation at San Nicolas Island (SNI) off southern California. The SNI colony now numbers approximately 50 animals, shows consistent local pup production, and is growing slowly in number, fostering increasing optimism for long-term persistence, Reeves *et al.* (2002) to the contrary notwithstanding. The California sea otter population is listed as "threatened" as per ESA and is designated "endangered" by RLTS.

Several risk factors are known for sea otters in California. Near-shore set-net fisheries were responsible for significant rates of incidental take in the late twentieth century. Changes in fishery regulations have reduced but not eliminated incidental take. Sea otters are known to compete with near-shore marine shellfish fisheries in California, particularly for abalones, sea urchins, clams, and crabs. Concern about interactions of sea otters and shellfisheries is significant, often leading to controversy among scientists, managers, and stakeholders. There is long-standing concern about illegal killing of sea otters to protect shellfisheries, but little solid evidence of a significant problem. Since the 1980s shark attack has been recognized as a significant source of sea otter mortality, particularly in the northern portion of the current range, where great white sharks are relatively common. In the recent decades new concerns have emerged regarding a range of diseases and parasitic infestations that may be highly significant sources of sea otter mortality. In some cases, diseases may have been introduced to sea otters by way of anthropogenic infrastructure such as domestic sewage outfalls, whereas others may have dispersed naturally from terrestrial mammal populations. Increased occurrence of exotic intestinal parasites from unknown sources, particularly a species of acanthocephalan not previously known in sea otters, has been associated with significant rates of fatal peritonitis in the population. The California sea otter population is also regarded to be at high risk of potential damage from oil spills, associated with significant tankship traffic and marine petroleum development, both implemented and proposed, in coastal marine habitats either occupied or suitable for occupation by sea otters. Oiling mats the pelage and eliminates the thermoregulatory function of sea otters, leading rapidly to chilling and death. Oiling is also known to cause other pathologies. The *Exxon Valdez* oil spill of 1989 in Prince William Sound and adjacent waters, Alaska, killed hundreds of sea otters and provided extensive opportunities for detailed insights into pathologies resulting from oiling of animals. The event validated concerns about potential population-scale damage to sea otter populations wherever they overlap in space with marine petroleum extraction, transportation, or processing activities.

A clear consensus on the cause or causes of relatively low growth rates of the California sea otter population has not emerged. Primary risk factors are thought to be ongoing incidental taking in fishing nets, effects of contaminants, mortalities from diseases and exotic parasites, and shark attacks. There is a perception that sea otters in California are more susceptible to disease impacts as compared with other sea otter populations. The long-term survival of the sea otter population in California will likely require ongoing research to more effectively characterize current risk factors, and the development of strategies to minimize associated mortalities. The large and growing human population in California is a major underlying cause of the jeopardy status of sea otters, and broad popular support, including economic compromise by shellfishery and marine oil interests, likely will be required if recovery is to be successful.

Northern Sea Otter: E. Lutris Kenyoni *Wilson, 1991*

The original geographic range of the northern sea otter extended from Attu Island, at the western extreme of the Aleutian Archipelago, Alaska, eastward and southward along the coasts of Alaska, British Columbia, Washington, and Oregon. Sea otters currently dwelling in waters off Washington, British Columbia, and southeastern Alaska are descended from animals reintroduced in the 1960s and 1970s. Source populations in all cases were northern sea otters moved from either Amchitka Island or Prince William Sound. As noted in Section "California Sea Otter: *E. lutris nereis* (Merriam, 1904)", large-scale commercial hunting of sea otters for pelts began in 1741. It is estimated that between 500,000 and 900,000 northern sea otter pelts were taken during the commercial harvest period. When legal hunting ended in 1911, the number of surviving northern sea otters was probably between 1000 and 2000 individuals, scattered among seven isolated remnant populations from the Queen Charlotte Islands, Canada, to the Rat Island group of the Aleutian Archipelago. The Queen Charlotte population was extinct by 1920. The northern sea otter has been subject to ongoing subsistence harvest by native peoples for meat and pelts, probably for many centuries.

Populations from Prince William Sound westward largely recovered from effects of the fur harvest period without assistance, other than prohibition of harvest, during the twentieth century. Observed annual population growth rates were as high as 10–15% in some cases. The translocation projects

done in the 1960s and 1970s restored northern sea otters to habitats vacated by excessive fur harvests in southeastern Alaska, British Columbia, and Washington.

Legal subsistence harvest of northern sea otters by native villages in Alaska has averaged approximately 500 animals per year since the mid 1980s, pursuant to MMPA. The harvest is concentrated primarily in Prince William Sound and southeastern Alaska. The *Exxon Valdez* oil spill of 1989 killed between 500 and 2500 sea otters in Prince William Sound and nearby coastal areas. Despite the intensive public interest and media coverage that fed into apocalyptic scenarios, it appears that sea otter numbers in the Sound did not experience long-term reduction, with the possible exception of a few local areas most heavily affected by the oil spill. Some effects on sea otter prey populations and habitats may persist even to the present, but sea otter numbers are large and growing in most areas affected by the spill. The most recent surveys of sea otters in Prince William Sound, done in 2003, produced a population estimate of approximately 12,000 animals.

Northern sea otter populations number approximately 78,000 individuals, summed across all habitats in Alaska, British Columbia, and Washington. For management purposes, northern sea otter populations in Alaskan waters have been divided into three geographic stocks: Southeast (east of Cape Yakataga through southeastern Alaska to Dixon Entrance), Southcentral (Cape Yakataga west–Cook Inlet including Prince William Sound and Kachemak Bay), and Southwest (west of Cook Inlet including the Alaska Peninsula, Kodiak Archipelago, Bristol Bay, and Aleutian Archipelago). The Southwest Stock is listed as "threatened" as per ESA. The Southcentral and Southeast Stocks and the Washington population are not listed as per ESA. All northern sea otter stocks and populations are designated "endangered" as per RLTS. The British Columbia population is classified as "of special concern" by the Committee on the Status of Endangered Wildlife in Canada.

The most recent northern sea otter population estimate for the US Southwest Stock is approximately 47,700 animals and suggests an increasing trend over the previous several years. However, the recent data do not capture extraordinary fluctuations in the region in the recent decades. It is estimated that the Stock declined in numbers by over 50% from the mid 1980s to the beginning of the twenty-first century. The decline in numbers continued into the early 2000s, with numbers decreasing by 63% in 2 years in the central Aleutian Islands and 48% over the same period in the eastern Aleutians. The declines apparently did not occur in the Kodiak Archipelago or along portions of the Alaska Peninsula, and growth in the latter areas apparently has resulted in an overall recent increase in numbers in the stock despite the population-scale collapses in the Aleutian Islands.

The most broadly accepted explanation for the extreme declines is an abrupt increase in predation by killer whales of the mammal-eating ("transient") ecotype on sea otters beginning in the late 1980s and the early 1990s. The principal lines of evidence for areas showing the most precipitous declines include absence of beach-cast sea otter carcasses, despite high search effort and the knowledge that prior periods of food stress produced substantial numbers of beach-cast carcasses; excellent body condition and lack of disease or other pathologies in surviving sea otters; high abundance of preferred prey; higher observed survival in sea otters occupying habitats with physiographic barriers to killer whale access as compared with areas with whale access; the presence of foraging killer whales; and occasional observations of attacks by killer whales on sea otters. The killer whale predation hypothesis has also engendered skepticism, primarily because of the small number of direct observations of attacks by killer whales on sea otters and because of the minimal value of an ingested sea otter to the metabolic needs of a killer whale.

The credibility of the killer whale hypothesis has been damaged by a second published hypothesis to explain the inferred shift in killer whale diet from larger marine mammals to sea otters. The "serial depletion" hypothesis argues that commercial whaling in past decades in the North Pacific region reduced availabilities of preferred prey of mammal-eating killer whales. The postulated result was a series of prey population depletions, with killer whales shifting sequentially to smaller and less rewarding prey items following overexploitation of larger prey. The hypothesis suggests that sea otters became vital prey when populations of larger marine mammals, particularly pinnipeds, became too rare to sustain the whales metabolically. Publication of the hypothesis drew a series of published rebuttals and remains controversial. However, regardless of the underlying cause, we find it difficult to argue with evidence that the collapse of some northern sea otter populations has been caused primarily by killer whale predation. We view the evidence to be compelling and inconsistent with any plausible alternative explanation.

The northern sea otter clearly recovered from the excessive harvests of the seventeenth, eighteenth, and nineteenth centuries. Current sources of concern include negative interactions with fishing gear and vessels, proper management of legal harvests by native hunters, illegal harvests, and possible damage from chemical contamination, including oil spills. The primary concern, however, is the threat of extreme consumption rates by killer whales. There is no feasible and politically tractable management intervention that can reduce killer whale predation on free-ranging sea otter populations. However, predators rarely exploit prey to the point of extinction and it seems more likely that the whales will switch to different prey when faced with a diminished supply of sea otters. Contrary to the proponents of the serial depletion hypothesis, we suggest that the whales have a number of alternative prey, all larger and more valuable metabolically than sea otters, to which they will likely switch long before whale predation by itself places the northern sea otter in peril of extinction. Beyond the issue of killer whale predation, northern sea otters will likely remain out of jeopardy, as long as legal harvests are conservatively regulated, and all feasible efforts are made to reduce or eliminate incidental takes by fishing activity, illegal harvests, and the effects of contaminants and oil spills.

As noted above and in Section "California Sea Otter: *Enhydra lutris nereis* (Merriam, 1904)", translocations have made significant contributions to the restoration of sea otter populations in several regions. Here we explore the role and significance of translocation projects in greater detail. Sea otters are unique among the marine mammals in the tractability of population-scale translocation projects as a conservation tool.

Compared with other marine mammals, sea otters are small in body mass, can be captured with relative ease and with some control over age and gender of targeted animals, and tolerate temporary captivity and transportation reasonably well. The utility of population-scale relocation efforts as a conservation tool for sea otters has fallen out of favor among management authorities, in part because the implementation of such projects is costly, politically challenging, and because completion of requisite associated logistical, administrative, and legal tasks may require up to a decade in advance of the movement of animals. In addition, individual animals involved in relocation may be placed at risk. Six major sea otter translocation projects were initiated between 1965 and 1987 in the US and Canada. Each had the intent of restoring populations to regions from which sea otters had been hunted to extinction during the eighteenth, nineteenth, and early twentieth centuries. With one exception, animals to be relocated were captured from northern sea otter populations at Amchitka Island and Prince William Sound, Alaska. The recipient regions for translocation were vacant sea otter habitats in southeastern Alaska (initiated in 1965), the Pribilof Islands (1968), British Columbia (1969), Washington (1969), Oregon (1970), and SNI (1987). The source population for the SNI project was the California sea otter population along the mainland California coast. In each case, initial postrelease losses of sea otters, either by documented death or by dispersal to unknown locations, were typically in the range of 70–90%, causing justifiable alarm. In the Pribilof Island and Oregon cases, otter numbers dwindled to near zero and the projects were deemed unsuccessful, although single animals can be observed occasionally in both locations to the present. Three of the four other efforts (southeastern Alaska, British Columbia, and Washington) ultimately emerged as highly successful, producing populations with a summed total currently estimated at approximately 15,000 individuals, with numbers either stable or growing in all cases. The re-established population in southeastern Alaska now appears to be stable at 9000–10,000 individuals and supports sustained subsistence harvests by native hunters of up to several hundred animals per year.

The fourth and most recent project (SNI) also included high initial postrelease loss rates. Many of the relocated animals likely died following misdirected dispersal from SNI, but the project revealed for the first time the strong large-scale site fidelity of sea otters, with nearly one-third of relocated animals (all of which were tagged with visual color-coded markers on the flippers) returning across minimum distances of up to 400 km to original capture sites on the mainland California coast. Thus, instinctive homing, in many cases spatially misdirected, may have driven early losses of released animals observed in all translocation projects. Although the ultimate success level of the SNI project remains uncertain 25 years after implementation, the recent data are increasingly encouraging as noted. With the exception of prohibition of unregulated commercial harvest, implementation of translocation projects has produced conservation dividends that far exceed any other management action taken on behalf of sea otters since development of the International Fur Seal Treaty in 1911. Translocation projects have substantial conservation value if the requisite and daunting prior investment in time, effort, and funds can be managed.

Species, Subspecies, and Populations Once Thought to be Near Extinction, Now Showing Evidence of Recovery

Synopsis

Here we summarize data for 10 species, subspecies, or populations of marine mammals that have been near extinction in the recent past, but are now either recovered or apparently enroute to recovery (Table 5). We regard the taxa or populations in this group as unlikely to become extinct in the foreseeable future as long as risk factors presently unknown do not appear. In this section we present case summaries for four taxa, selected arbitrarily, to illustrate patterns typical of the group.

Western Arctic Population of the Bowhead Whale: Balaena mysticetus *Linnaeus, 1758*

Bowhead whales occur in at least five populations in northern habitats characterized by frequent sea ice. Four of the five populations are small and at risk of extinction (Table 3). The western Arctic population is by far the largest, numbering approximately 10,500 individuals, with annual growth rates averaging 3.5%. The population occurs in the Bering, Chukchi, and Beaufort Seas, migrating northeasterly to the Beaufort Sea in spring, and returning to the Bering and Chukchi Seas in autumn each year. The population is listed as "endangered" as per ESA, and is designated as "lower concern – conservation dependent" by RLTS.

The population probably numbered between 10,000 and 23,000 individuals when first exploited by US whalers during the late 1840s. By the demise of the commercial harvest in 1919 approximately 20,000 whales had been taken, and the population was thought to number between 1000 and 3000 individuals.

Bowhead whales have been hunted for subsistence and ceremonial purposes by the indigenous peoples of Alaska, western Canada, and Russia for many centuries. Alaskan native villages continue to harvest whales from the western Arctic population based on a quota approved by IWC and managed jointly by representatives of native villages and US federal agencies. Management policies have been successful, allowing continued growth of the population despite annual hunts for whales by several villages. In the recent years the annual harvest from the western Arctic population by native villages in Russia, the US, and Canada has averaged 41 whales. For a number of years there has been concern about effects of marine petroleum exploration and development activities in the habitat of bowhead whales, especially in the Beaufort Sea. Offshore oil activity is known to influence movement patterns of whales during migration, but demographic effects have not been demonstrated. A recent increase in active and proposed marine oil development along the Arctic coast of Alaska has renewed concerns regarding conservation implications for bowhead whales.

The western Arctic bowhead whale population appears to be recovering while other bowhead whale populations are not. The likely primary reason is that other populations were pushed much closer to extinction by commercial whaling, possibly imposing depensatory effects and limiting the capacity for recovery. It appears that the western Arctic population will continue to grow, perhaps approaching carrying capacity

Table 5 Species, subspecies, or populations once thought to be near extinction, now showing evidence of increasing numbers

Taxon	*Population*	*Estimated population size*[a]	*Current and historical risk factors*
Bowhead whale: *B. mysticetus* Linnaeus, 1758	Western Arctic (Bering, Chukchi, and Beaufort Seas)	10,500	CH, MO, SH
Humpback whale: *M. novaeangliae* (Borowski, 1781)	All except Oceania (IWC breeding stock "E") and Arabian Sea	50,000–60,000	CH, ITF, VC, VN
Gray whale: *E. robustus* (Lilljeborg, 1861)	Eastern North Pacific	20,000	CH, ITF, VC, VN
Eastern spinner dolphin *Stenella longirostris orientalis* (Perrin, 1990)	All	613,000	ITF
Northern elephant seal: *M. angustirostris* (Gill, 1866)	All	125,000	CH
California sea lion *Z. californianus* (Lesson, 1828)	All	250,000	CH, CP, FLI, IHC, ITF, MDI, SH
Galapagos fur seal: *Arctocephalus galapagoensis* Heller, 1904	All	30,000–40,000	CH, ITF, FLI
Subantarctic fur seal: *Arctocephalus tropicalis* (Gray, 1872)	All	270,000–360,000	CH
Antarctic fur seal: *Arctocephalus gazella* (Peters, 1875)	All	Approaching 4,000,000	CH
Russian sea otter: *E. lutris lutris* (Linnaeus, 1758)	All	20,000	CH, CP, MO, PO

[a]Indicated estimates of population size reflect a wide range of quantitative rigor. Primary sources for estimates shown are online resources associated with NMFS, FWS, and RLTS. The responsibility for errors rests exclusively with the authors of this article.
Codes for risk factors are as shown in Table 3.

within the current century, as long as harvests by native villages are regulated conservatively, and as long as other significant risk factors do not emerge.

Humpback Whale: Megaptera novaeangliae *(Borowski, 1781)*

Humpback whales are distributed globally in marine habitats. At least seven populations are recognized in the southern hemisphere: three in the North Pacific, one in the North Atlantic, and an unusual nonmigratory population in the Arabian Sea. Humpback whales often occur near shore. Lengthy seasonal migrations between high-latitude summer feeding areas and low-latitude winter breeding areas are well documented for all populations except for the Arabian Sea whales.

Humpback whales were among the earliest targeted species of commercial whalers because of their abundance, large size, and coastal distribution. During twentieth-century commercial hunts in the southern ocean, several hundred thousand individuals were taken. Illegal harvest by whalers from the USSR was extensive and potentially damaging to populations following the decline of large-scale legal commercial whaling. Subsistence harvest by indigenous peoples was common in the previous centuries. Additional modern risk factors include shipstrikes, incidental take in fishing nets, and habitat disturbance by large recreational cruise ships. All humpback whales populations are listed as "endangered" as per ESA. RLTS designates populations in the Arabian Sea and Oceania (the latter including eastern Australia, Tonga, New Caledonia, the Cook Islands, and French Polynesia) as "endangered". Other populations are not listed as per RLTS.

Although trends in most populations have not been confidently documented, there is a broad perception that humpback whale numbers are increasing on a global scale. Recent estimates are 11,500 individuals in the North Atlantic, and a combined total of 21,000 individuals in the North Pacific. Population growth rate estimates range from 3% to 7% per year. Data from the southern hemisphere and the Arabian Sea are not sufficient to develop confident population estimates or characterize trends.

The population wintering near Tonga, in the western South Pacific, experienced excessive subsistence harvests during the twentieth century. A moratorium was placed on the harvest of humpback whales by the Tongan government in 1978, when it was recognized that the number of whales breeding in Tongan waters was seriously depleted and that further subsistence harvest was not sustainable. As noted above, the Tongan breeding population of humpback whales is lumped with breeding whales at other locations as a single breeding stock by IWC. However, IWC recognizes the Tonga case as an identifiable substock, facilitating improved scrutiny. Humpback whales in the Arabian Sea are found from approximately the Gulf of Aden eastward to Sri Lanka. The population was reduced over the long term by commercial whaling, and more recently by illegal Soviet whaling and entanglement in fishing gear. It appears there may be a few hundred animals in the population.

The eastern and western North Pacific populations of humpback whales remain small enough to be of ongoing concern in a conservation context, and the relatively small Arabian Sea population may not be able to sustain current levels of incidental taking by local fisheries. We suggest that prospects for survival of humpback whales, as a species, are good as long as commercial and subsistence harvests remain suspended, incidental taking is reduced, and habitat disturbance by human activities continues to be restricted and carefully monitored.

Eastern North Pacific Gray Whale: E. robustus *(Lilljeborg, 1861)*

The eastern North Pacific population of gray whales ranges from the Bering, Chukchi, and Beaufort Seas of the Arctic region during summer to the shallow protected coastal lagoons of Baja California Sur, Mexico, during winter months. Eastern gray whales were hunted intensively by commercial whalers from the eighteenth through the early twentieth centuries. The population may have reached a minimum of approximately 4000 individuals in the late 1890s. The population numbered approximately 13,000 when regular quantitative surveys were begun in 1967. Based on standardized surveys done from 1967 through 2007, the current population is thought to be 20,000 individuals, with a mean annual growth rate of 3.3%. The time series includes a dramatic mortality event in 1999 and 2000 apparently linked to fluctuating prey availability. Patterns in population trend, observed adult mortality rates, and the number of whales in unusual locations during summer are all consistent with the perception that the population is approaching carrying capacity. In 1994 the eastern gray whale was the first marine mammal to be removed from the List of Endangered and Threatened Wildlife as defined by ESA. The eastern gray whale is not listed by RLTS.

Several risk factors are of current concern for eastern gray whales. Modest rates of incidental take in fishing nets are reported. Disturbances by boats supporting "whale watching" activities are of concern, both in migratory corridors along the heavily populated west coast of North America and especially in the breeding and calving lagoons in Mexico. As with all large cetaceans, mortalities from shipstrikes are a chronic risk factor for gray whales.

Eastern gray whales are subject to ongoing annual subsistence harvests by indigenous peoples in Russia, Alaska, and Washington State. The subsistence harvests are authorized by IWC quota and are comanaged by village elders and government agencies in Russia and the US. Subsistence hunts have taken approximately 100 whales per year for several decades, with most of the harvest occurring along the Russian coast of the Bering and Chukchi Seas during summer months. In 1998 the Makah tribe of northwestern Washington State was allocated a harvest quota of five whales, in exchange for a reduction of the quota by five whales for Russian and Alaskan hunts. The allocation of harvest quota to the Makah people was consistent with the Treaty of Neah Bay of 1855 (12 Stat. 939), between the Makah tribe and the US government. One whale was taken in 1999 by Makah hunters. The extraordinary public interest, political discord, and coverage by news media were vastly disproportionate to the demographic significance of the hunt. Owing to a range of US federal legal proceedings and court rulings, Makah tribal hunters have harvested no additional gray whales since 1999 pursuant to the IWC quota. One gray whale was killed illegally by five Makah tribal members in 2007.

In our view the eastern gray whale population is no longer at risk of extinction in the foreseeable future. The status of the population should remain stable and sustainable as long as subsistence harvests are conservative and carefully managed. Rigorous monitoring and appropriate management responses to increasing problems with incidental taking, disturbance by boats, and shipstrikes will increase the probability that the population will remain out of jeopardy.

Northern Elephant Seal: Mirounga angustirostris *(Gill, 1866)*

Northern elephant seals breed and molt on coastal islands and isolated mainland sites of Baja California, Mexico, and southern and central California, US. When not hauled out, seals forage in pelagic habitats of the temperate and subpolar North Pacific. Two complete migrations are made each year between hauling sites and foraging habitats. Seals pup, nurse, and wean young, and breed from December through middle March at the hauling sites, then swim north to forage. They return to haulouts to molt during spring and early summer, then return again to foraging areas. Schedules for migration vary among age and sex categories. When hauled out, elephant seals are concentrated at high density. Hauled seals are easily approached by humans, substantially more so than other land-breeding pinnipeds of North America.

Northern elephant seals have been utilized for food and oil over the centuries by native peoples of North America. Intensive commercial harvests for oil began early in the nineteenth century. Most colonies of seals were severely depleted by 1850, but commercial harvests continued until 1884, at which point the species was considered extinct. Subsequent discoveries of small numbers led to continued harvests by scientific collectors working for museums of natural history. Scientific collecting continued until at least 1911. The total number of surviving seals is thought to have been as low as 20–100 individuals in 1890. At the time harvests finally ended, elephant seals survived only at Isla Guadalupe, west of Baja California, Mexico.

Northern elephant seals numbered approximately 125,000 individuals in 2005. They occupy at least 16 major hauling sites in Mexico and California. In the recent decades at least four new mainland breeding and molting colonies have developed along the central California coast. Population growth has been approximately 6% per year through most of the century. Recent studies indicate a marked lack of genetic diversity at assessed loci, almost certainly a result of the severe population "bottleneck" associated with commercial and museum harvests. Northern elephant seals are not listed as per ESA or RLTS.

The elimination of commercial harvest has allowed northern elephant seals to recover fully from near extinction despite a striking loss of genetic diversity. Most current management concerns for the species involve problems of overabundance rather than rarity. Indefinite survival should be assured as long as harvests and disturbances to habitat can be monitored and controlled.

Scientific Whaling and Marine Mammal Extinction

Whaling in international waters is regulated currently by the IWC, pursuant to the International Convention for the Regulation of Whaling, approved by original member nations in 1946. Although often referenced as "international law," the regulatory authority of IWC applies only to member nations and relies entirely on their voluntary compliance. IWC protocols also include options for member nations to file exceptions to adopted regulations, allowing pursuit of activities contrary to IWC protocols. Several nonmember nations are engaged currently in active whaling.

In 1986 IWC established a moratorium on whaling based on the argument that information on the status and dynamics of whale populations was inadequate for development and application of an effective and informed management strategy. The moratorium was a reflection of the tendency of IWC in the recent decades to adopt more precautionary, conservation-oriented approaches to the regulation of whaling, but was opposed by some member nations. The moratorium remains in effect to the present. In 1988 Japan, an IWC member nation, was granted a "special permit" to allow resumption of whaling, ostensibly for scientific research purposes. The initial request was for an annual quota of 800 whales with an emphasis on minke whales in the southern ocean. The magnitude of the initial quota request drew sharp criticism from a broad range of interested parties and was soon pared to 300 whales. The stated primary intentions of harvest, implemented by the Institute of Cetacean Research in Japan (ICR), were to improve the understanding of interactions of whales with harvested fish stocks and other elements of marine ecosystems. ICR indicated an intention to market meat and other products from whales taken under the special permit once data were obtained from animals killed.

ICR harvested 6778 Antarctic minke whales (*Balaenoptera bonaerensis* Burmeister, 1867) over an 18-year span following receipt of the special permit (Baker *et al.*, 2010). Harvesting also occurred in the western North Pacific region. To date, ICR harvests are known to have included North Pacific common minke whales (*Balaenoptera acutorostrata scammoni* Deméré, 1986), sei whales (*Balaenoptera borealis borealis* Lesson, 1828), Bryde's whales (*Balaenoptera brydei* Olsen, 1913), sperm whales (*Physeter macrocephalus* Linneaus, 1758), and Antarctic fin whales (*Balaenoptera physalus quoyi* (Fischer, 1829)), in addition to Antarctic minke whales. The permit remains in effect to the present and the ICR harvest is continuing.

There is no indication to date that whaling by ICR pursuant to the special permit from IWC has directly increased the jeopardy status of any whale population from which animals have been taken. Nevertheless, the ICR harvest has been subject to criticism in several categories. First, there have been persistent arguments that use of lethal take to obtain targeted scientific data is unnecessary, and cannot be justified given the range of alternative methods available to investigate many of the scientific questions identified by ICR. Second, published presentations and interpretations of data from harvested animals have been criticized as inadequate, erroneous, substandard, and redundant relative to information already available in literature. Third, characterization of the harvest as a source of scientific insight has been intensively criticized as a ruse, with the argument that the true purpose is to supply whale products in response to high demand in Japanese markets. Fourth, it has been argued that the harvest should not be permitted for the same reasons that were used to justify the original implementation of the moratorium on whaling by IWC. Fifth, the IWC's internal process for evaluating results of work under the special permit has been criticized as lacking rigor and compromised by conflicts of interest within the reviewing panels. Finally, there have been concerns that entry of whale meat into open markets will undercut the ability of regulatory protocols, most especially the Convention on International Trade in Endangered Species (CITES), to restrict trade across international borders for products obtained from species or populations in peril.

International criticism of ICR whale harvests was further stimulated by results of a study published by Baker *et al.* in 2010. Servings of uncooked whale meat were purchased at "sushi" restaurants in California and South Korea and were subject to detailed genetic evaluation. The data were compared with genetic information obtained previously from whales known to have been harvested by ICR pursuant to the special permit. Baker *et al.* concluded that whale products purchased in the restaurants were from animals harvested as per the special permit, with an extremely low statistically based probability of error. The restaurant samples were found to include products from North Pacific common minke whales, Antarctic minke whales, sei whales, and fin whales, as well as material from Risso's dolphin (*Grampus griseus* (G. Cuvier, 1812)). Baker *et al.* noted that the appearance of whale products in California and South Korea represented evidence of clear violations of CITES.

The published allegations of a CITES violation raise grave concerns regarding the conservation of large whales and the prevention of population- or species-level extinction. The principal concern is that illegal trade across international borders will create an expanded "black market" for whale products and will provide economic incentives for expanded illegal whaling. With the cooperation of Russian scientists, the scope and impacts of illegal harvests in past decades by whalers from the USSR have been well documented and provide a clear indication of the dire conservation consequences of unregulated whaling. For example, as noted previously, illegal Soviet harvests of eastern North Pacific right whales probably had a significant role in driving the population to the threshold of extinction. Illegal harvests of terrestrial wildlife to service the demands of black markets are well known, have placed a number of species into extreme jeopardy, and are extraordinarily difficult to eliminate. It follows that any economic stimulus for covert illegal whaling is a matter of highest concern in the context of marine conservation and the potential for extinctions of the large whales. Thus, continuation of authorized whaling under the special permit should receive the most intensive possible critical scrutiny by IWC. To the maximum feasible extent, political considerations in maintaining the special permit should be superseded by legal and conservation perspectives, most especially the implications of alleged CITES violations.

Global Climatic Trends and Marine Mammal Extinction

Climate change, including ocean warming and acidification, will likely produce substantial negative ecological consequences for marine mammals, although timelines and spatial patterns for impacts are generally unknown and difficult to predict. For marine mammals, most recent discussions linking climate change to extinction risks focus on high-latitude species, especially those that have obligatory ecological links to sea ice (e.g., Moore and Huntington, 2008). Large-scale reductions in high-latitude sea ice cover have been recognized for at least two decades, and there is no evidence to suggest changes in the pattern in the foreseeable future.

We suggest that fifteen Holocene marine mammal species, all known to have obligate ecological relationships with sea ice, are at significant risk of population reduction, possibly including extinction, as global warming continues. Although low- and medium-latitude marine mammals may be able to survive long-term trends in climate by adjusting latitudinal ranges and habitat use patterns, ice-dependent species have no alternative habitat options once projected losses of sea ice reach critical ecological thresholds. Following are ice-dependent species at significant near-term risk as a result of global warming patterns, with an indication of the type of ice typically associated with each species:

Arctic and sub-Arctic pinnipeds (all are phocids except the walrus):

Bearded seal: *Erignathus barbatus* (Erxleben, 1777): pack ice, shorefast ice;
Harp seal: *Pagophilus groenlandicus* (Erxleben, 1777): pack ice;
Hooded seal: *Cystophora cristata* (Erxleben, 1777): heavy pack ice;
Ribbon seal: *Histriophoca fasciata* (Zimmerman, 1783): pack ice;
Ringed seal: *Pusa hispida* (Schreber, 1775): pack ice, shorefast ice;
Spotted seal: *Phoca largha* Pallas, 1811: pack ice;
Walrus: *Odobenus rosmarus* (Linnaeus, 1758): pack ice.

Other Arctic species:

Beluga: *D. leucas* (Pallas, 1776): pack ice;
Bowhead whale: *B. mysticetus* Linnaeus, 1758: pack ice;
Narwhal: *Monodon monoceras* Linnaeus, 1758: heavy pack ice, especially in winter;
Polar bear: *Ursus maritimus* Phipps, 1774: pack ice.

Antarctic species (all phocid pinnipeds):

Crabeater seal: *Lobodon carcinophaga* (Hombron and Jacquinot, 1842): pack ice;
Leopard seal: *Hydrurga leptonyx* (Blainville, 1820): pack ice;
Ross seal: *Ommatophoca rossii* Gray, 1844: medium and heavy pack ice;
Weddell seal: *Leptonychotes weddellii* (Lesson, 1826): shorefast ice.

The list of vulnerable species is longer and more diverse for the Arctic region compared with the Antarctic, correlating with two patterns that may serve as explanations. First, the Arctic marine habitat is proximate to the continental land masses of North America and Asia, possibly influencing the range of options for evolutionary transitions from terrestrial to marine living on an evolutionary scale. In contrast, Antarctic marine ecosystems are well removed from ice-free continental land masses. Second, the Arctic marine environment is an oceanic water mass largely surrounded by land whereas the Antarctic is a land mass fully surrounded by oceanic waters.

Antarctic marine ecosystems are used intensively by blue, fin, and minke whales along with killer whales. However, although reduced ice presence will likely produce negative consequences, fin, blue, minke, and killer whales have more flexible habitat options and higher probabilities of surviving catastrophic reduction of sea ice than the three Arctic cetacean species with obligatory associations to sea ice. An informative example is provided by geographic changes in foraging patterns of the eastern North Pacific gray whale population in the recent decades. In apparent response to recently documented changes in distribution and productivity of gray whale prey in the Bering Sea, gray whales have expanded their foraging range farther north, with increased presence and feeding in the Chukchi and Beaufort Seas, and farther south, with foraging groups commonly using a number of locations, including the Kodiak Archipelago in Alaska, and sites along the coasts of British Columbia, Washington, Oregon, and northern California. Factors other than climatic shifts may partially explain changes in foraging habitat, including the postulated approach of the population to carrying capacity (see Section "Eastern North Pacific Gray Whale: *E. robustus* (Lilljeborg, 1861)"). However, it is clear that the species can adapt successfully to a shifting prey landscape.

We suggest that polar bears and the southern North Pacific population of spotted seals, the latter dwelling in the Sea of Japan and the northern Yellow Sea in the far western North Pacific region, are likely to be among the earliest of the obligate ice-associated marine mammals to suffer major population reductions as global warming proceeds. The southern spotted seal population is at a lower latitude than other ice seals and may be among the first to experience profound loss of ice habitat. Recent observations indicate that pupping in the population is in transition from pack ice to land, and that dependent pups have been seen leaving haulouts and entering the water before weaning. Both behaviors are abnormal for the species and may indicate the onset of population-scale stresses linked to climatic trends. The southern spotted seal population is listed as "threatened" as per ESA, but is not listed in the RLTS. Pack ice loss rates are also known to be high in Davis Strait (between Greenland and Baffin Island, Canada) and along the eastern coast of Greenland), facilitating vulnerability of resident populations of harp seals and hooded seals.

Polar bears may already be experiencing population-scale stress associated with negative trends in sea ice cover in the Arctic. Polar bears occur in Alaska, Canada, Greenland, Norway, and Russia, with summed populations numbering between 20,000 and 25,000 individuals. Population trend characterizations are difficult to produce for polar bears (*see* General Factors Hindering Effective Identification and Monitoring of Marine Mammal Populations Vulnerable to Extinction), but there is a general perception that populations are declining in a pattern that roughly tracks observed declines in ice cover. Polar bears in Alaska are listed as "threatened" as per ESA, and global populations are listed as "vulnerable" as per RLTS. Reduction in ice extent poses at least three major ecological problems for polar bears. First, pack ice is the primary habitat for ringed seals, the most important prey species for polar bears in a nutritional context. Thus, reduced ice cover will result in diminished foraging habitat for polar bears, and reduction in prey abundance to the extent that lost ice cover negatively influences ice-dwelling seals. Second, as ice cover dwindles, polar bears will be constrained to increase the amount of swimming required to reach appropriate habitat for foraging. Polar bears must be on the ice surface, rather than in the water, to hunt ringed seals efficiently. Swimming is more costly metabolically than

walking or running on the ice surface for travel. It follows that with declining ice cover, polar bears will be required to expend greater energy per unit of time to reach productive feeding areas, finding smaller and more scattered feeding areas in which to forage on arrival. Finally, significant loss of sea ice will impose increased durations of onshore fasting periods in warmer seasons, possibly beyond metabolic tolerances, causing negative impacts to vital population parameters (e.g., Molnár et al., 2011).

We suggest that other obligate ice-associated species listed above are also likely to experience serious reductions in numbers as global warming proceeds. For these species, sea ice has crucial effects on ecosystem function and prey productivity, on which obligate ice-associated cetaceans and pinnipeds are highly dependent. For the pinnipeds, sea ice is vital as breeding and resting habitat as well. The physical configuration of sea ice has significant effects on key attributes of breeding systems used by ice-dwelling seal species on evolutionary scales, including the intensity of male–male competition, behavioral patterns in territorial defense by males, the level of sexual dimorphism in size, and the degree of polygyny or polyandry. We suspect that drastic reduction or loss of sea ice will overextend the adaptability of the listed species and greatly increase the probability of a collective reduction, perhaps leading to extinctions in some or all cases.

In addition to expanded marine oil exploration and development, already underway, the loss of sea ice in the Arctic region is projected to facilitate a dramatic increase in other forms of human activity, including the introduction of shipping traffic, fishing vessel activity, and other forms of human industrial development in habitats with a largely pristine history. With such actions, a predictable suite of conservation risks will be introduced to ice-dependent Arctic marine mammals, including shipstrikes, fishery bycatch, competition with fisheries for prey in common, chemical contamination, including oil spills, and noise pollution. Disruption to Antarctic ecosystems, with consequent increased risks of extinctions in marine mammal populations, will also emerge with the reduction or loss of sea ice, although likely consequences of the introduction of human industrial presence in the Antarctic are more difficult to predict. Recognizing that projections of ecological effects of climatic trends in middle and low latitudes are unclear, we refrain from specific projections of effects on marine mammals in ice-free regions. We acknowledge that consequences of climate change at lower latitudes could be substantial for marine mammals and likely will test the adaptability of lower-latitude species.

Discussion

Synopsis of Factors and Processes Known to be Facilitating Modern Anthropogenic Extinctions of Marine Mammals

In our review of examples of taxa and populations of marine mammals currently in jeopardy of anthropogenic extinction, we see two major, recurring categories of vulnerability. The first is prolonged excessive harvest, usually for commercial purposes, reducing numbers to a small fraction of pre-exploitation status. In this category we find taxa or populations from all marine locations on earth, including many that are distant from human population centers. The second is a combination of risk factors strongly linked to a proximate and encroaching human population. The factors include habitat loss or chronic habitat disturbance, incidental taking in fishing gear, and contaminants, as well as directed harvests. The second category primarily includes taxa or populations restricted ecologically to a limited geographic range in near-shore marine or aquatic habitats. Here we consider the vulnerabilities of taxa or populations in the two categories.

The first category primarily involves excessive directed harvest. All species of mysticetes, the larger odontocetes, nearly all pinnipeds, the marine otters, and the polar bear have been hunted extensively at least over the last few centuries, in most cases in order to obtain articles of commerce. Most exploited taxa lack the necessary demographic features to sustain viable populations at the level of harvest experienced. There are two important results. First, such exploitation has often reduced populations to small sizes. Second, populations thus affected require decades or even centuries to recover to levels free of the risks of extinction. Directed, commercial-scale harvests of marine mammals ended for most species during the twentieth century, and many species now are subject to rigorous protection. However, the risk of extinction persists for some reduced populations despite relaxation of harvest activity.

Small populations are vulnerable to any factor that reduces survival or collective reproductive success. Survival and reproduction can be impaired by anthropogenic factors such as contaminants or disturbance to critically important breeding locations, or by natural fluctuations in the biological habitat. Anthropogenic factors should be controllable in principle by the appropriate management actions, but in reality, effective risk management is difficult. Natural fluctuations are effectively stochastic in timing, duration, and intensity, and cannot be anticipated or controlled by any form of management authority. Once marine mammal populations are reduced, they recover slowly and are, therefore, at risk of the damaging effects of anthropogenic damage or natural disturbances over an extended period even under the most rigorous protection. For example, North Atlantic right whales have not been harvested for decades, but prior harvest reduced them to low numbers. Now even low rates of shipstrike mortality or incidental take in fishing gear are adequate to hold the population at a dangerously small size.

The second category involves a group of factors associated with growing human populations in coastal regions, interacting with taxa or populations constrained to life in the coastal zone. Species in this category include the sirenians, river dolphins, and a number of coastal odontocetes and pinnipeds. The essential problem here is that growing human populations produce a suite of effects, each damaging to proximate marine mammal populations. Reduction of the effects often requires a deliberate curtailment of economic enterprise such as fishing, or of institutional infrastructure such as waste disposal, flood control, or the provision of drinking water. The emergent dilemma is the perception by political institutions that there must be a choice between human welfare and the welfare of nearby marine mammals. Our case studies suggest that, given the choice, human cultures of major population centers act in favor of human needs.

Thus, for example, recreational boating activities continue to crowd needed habitat for manatees in Florida despite decades of documentation that manatees do not tolerate powerboat activities, and fishing interests continue to set nets in vaquita habitat despite widespread recognition that incidental take is driving the population to extinction.

Species in our first category have reasonable probabilities of survival. Northern elephant seals escaped the period of vulnerability associated with small population size, and other taxa seem well on their way. Some taxa or populations probably will not persist. Western gray whales and northern right whales, particularly the eastern North Pacific population, will survive the new century only with the most rigorous imaginable protection. We are less optimistic about species in our second category, because their ultimate survival depends on conscious economic restraint by human cultures, and a possible reevaluation of values regarding the survival of marine mammals and other species in habitats also used or coveted by people. Under this premise we anticipate that the vaquita will be the next marine mammal extinction at the species level.

Excessive subsistence harvests, anthropogenic noise, contaminants, oil spills, and depletion of genetic diversity are issues that have at least occasionally been invoked as risk factors for extinction of marine mammals. We find that there are relatively few taxa or populations clearly falling toward extinction as the direct result of any one of these factors. Subsistence harvest by native peoples is without question a serious risk factor for Cook Inlet beluga whales, and may have been a crucial precursor to the extinction of Steller's sea cow. However, the western Arctic bowhead whale population has been growing steadily for years despite regular annual subsistence harvest. Thus, subsistence harvests are manageable risk factors and need not be regarded universally as unacceptable practice. None of the other listed factors are alone causing widespread extinction risk, although there are isolated examples for each. Of greater concern here is the problem of significant effects arising from the interaction of multiple factors. The best-known cases involve mass mortalities that result from disease outbreaks. Such outbreaks often result from immune suppression, which in turn may result from contaminants or from natural disturbances such as the toxic byproducts of certain phytoplankton. Interacting factors often are a problem near human population centers, and are difficult, if not impossible, to manage. Thus, in our view, the danger of many risk factors discussed here is not the direct effect of a single factor, but the synergistic effects of multiple factors that may be less damaging when separated from one another. The inexorable risks posed to marine mammals by human population expansion are further problematic because of inherent uncertainties in population data for many species, confounding efforts to identify and track the status of taxa jeopardized by contact with human populations.

Geographic Regions of Greatest Concern with Regard to Anthropogenic Extinctions of Marine Mammals.

As suggested above, marine mammal taxa and populations constrained to life near human population centers are in general most vulnerable to anthropogenic extinction. Extinction risks will be greatest where human cultures have the fewest economic options when confronted with the need for restraint in order to solve conservation problems. Such circumstances are most likely in "developing" countries at lower and middle latitudes, where large concentrations of people face ongoing economic shortfalls. Over the next century we anticipate the greatest extinction risks for coastal marine or aquatic marine mammals in southern and southeastern Asia, eastern Europe, central America, and central Africa.

Marine mammals in the southern hemisphere generally should be at lower risk of extinction over the next century than those in the north. Most major human population centers and most cases of coastal habitat degradation occur in the northern hemisphere. However, many southern populations remain small because of past excessive harvests. Thus, distantly located southern taxa and populations will be removed from the risk of extinction only to the extent that all forms of harvest are regulated with extreme caution and conservatism. Moreover, small distant populations are at risk of incorrect conclusions about number of individuals, trend in numbers, or demographic characteristics because of the statistical limitations associated with survey and demographic data that can be obtained only infrequently and at high cost.

General Approaches Toward Minimizing the Rate of Anthropogenic Extinctions of Marine Mammals

Excessive directed harvests have caused more cases of jeopardized marine mammal taxa and populations than any other single factor. Thus the most direct and straightforward approach to the control of extinction risk for marine mammals is a precautionary approach to the concept of marine mammal harvest on a global scale. Fortunately, this is the approach currently adopted by many governments, and by international regulatory cooperatives such as IWC, for some species. In this context we offer three points of caution. First, the protections provided by international treaties and conventions such as the IWC, and by individual governments, do not necessarily extend to all marine mammals. There are high-profile protective protocols, with active ongoing oversight, for the larger cetaceans, some of the pinnipeds, sea otters, and polar bears. Many small cetaceans, some pinnipeds, and some sirenians are not actively or explicitly protected at the national or international level. Second, many of the populations subject to active protection are small in number, and as a consequence will be subject to risks associated with stochastic events, both natural and anthropogenic, for decades if not centuries. Thus, some extinctions are possible because of unequal conservation effort and because the most aggressive protection cannot eliminate all risk factors. Third, the crucial process of detecting trends in small distant populations is so costly that errors are likely in determining which populations are most seriously jeopardized. Thus, despite the best of human intentions, protective effort may be misdirected or inappropriately withheld.

The group of risk factors associated with human population growth will almost certainly cause some extinctions of coastal marine mammals in the current century. Here the outcome is in our view more certain, and the methods of prevention more intractable, than in the simpler cases of small

populations distant from major concentrations of people. Effective protection of imperiled species requires that human cultures forego economic benefits for the good of jeopardized marine mammals. Such sacrifices must extend indefinitely to be effective, given the demographic limitations of most marine mammals. The acceptance of foregone benefits is most needed in cultures least able to accommodate the loss. Cultural acceptance of economic loss motivated by a conservation ethic will require education and reorientation at a level that we find difficult to imagine.

Both evolution and extinction of species have been characteristic features of the history of marine mammal taxa on earth since the early Eocene. Although the anthropogenic loss of species is both regrettable and inevitable in our view, there is no reason to deny that new species will evolve as well. For example, killer whales off Washington and British Columbia occur in three distinctive social configurations, termed "ecotypes", that correlate with subtle but reliable morphological differences. The different ecotypes have entirely different diets, different acoustic repertoires, and rarely co-occur in space. Some have characterized these patterns as a step in the process of speciation, although there are alternative viewpoints. The human mind can easily perceive extinction but often not speciation as a finite event, because of differing time scales and imprecise definitions of the latter. Thus, we are biased toward greater conscious awareness of extinction than of speciation. Although anthropogenic extinctions should not be accepted without exhaustive attempts at recovery, it is perhaps more pragmatic to view such events as one part of a dynamic process of evolution, rather than as catastrophic failures of human behavior.

See also: Antarctic Ecosystems. Arctic Terrestrial Ecosystems. Census of Marine Life. Climate Change and Ecology, Synergism of. Climate Change and Extinctions. Climate Change and Wild Species. Comparing Extinction Rates: Past, Present, and Future. Conservation Biology, Discipline of. Conservation Genetics. El Niño and Biodiversity. Endangered Ecosystems. Endangered Mammals. Estuarine Ecosystems. Extinction, Causes of. Extinction in the Fossil Record. Government Legislation and Regulations in the United States. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Impact of Past Global Warming on Biodiversity. Indigenous Peoples and Biodiversity. Mammals, Conservation Efforts for. Mammals (Pre-Quaternary), Extinctions of. Marine Conservation in a Changing Climate. Marine Ecosystems. Modern Examples of Extinctions. Natural Extinctions (not Human Influenced). Ocean Ecosystems. Pelagic Ecosystems. Population Dynamics. Population Viability Analysis. Social and Cultural Factors. Speciation, Process of. Speciation, Theories of. Stress, Environmental. Threatened Species: Classification Systems and Their Applications. Translocation as a Conservation Strategy

References

Baker CS, Steel D, Choi Y, *et al.* (2010) Genetic evidence of illegal trade in protected whales links Japan with the US and South Korea. *Biology Letters* 6: 647–650.

Berta A, Sumich JL, and Kovacs KM (2005) *Marine Mammals: Evolutionary Biology*, 2nd edn. San Diego: Academic Press.

Garner GW, Amstrup SC, Laake JL, Manly BFJ, McDonald LL, and Robertson DG (eds.) (1999) *Marine Mammal Survey and Assessment Methods*. Rotterdam: A.A. Balkema.

Jefferson TA, Webber MA, and Pitman RL (2008) Marine Mammals of the World. A Comprehensive Guide to Their Identification. Amsterdam: Elsevier.

Kovacs, K.M., Aguilar A., D. Aurioles, *et al.* Global threats to pinnipeds. *Marine Mammal Science*, in press. doi:10.1111/j.1748-7692.2011.00479.x

Molnár PK, Durocher AE, Klanjscek T, and Lewis MA (2011) Predicting climate change impacts on polar bear litter size. *Nature Communications* 2: 186. doi: 10.1038/ncomms 1183.

Moore SE and Huntington HP (2008) Arctic mammals and climate change: Impacts and resilience. *Ecological Applications* 18(Suppl.): S157–S165.

Perrin WF, Würsig B, and Thewissen JGM (2009) *Encyclopedia of Marine Mammals*, 2nd edn. San Diego: Academic Press.

Reeves RR, Stewart BS, Clapham PJ, Powell JA, and Folkens P (2002) *Guide to Marine Mammals of the World*. New York: Knopf.

Rice DW (1998) Marine Mammals of the World. *Systematics and Distribution*. Lawrence, Kansas: The Society for Marine Mammalogy. Allen Press. Special Publication Number 4.

Turvey ST, Pitman RL, Taylor BL, *et al.* (2007) First human-caused extinction of a cetacean species? *Biology Letters* 3: 537–540.

Marine Protected Areas: Static Boundaries in a Changing World

Elizabeth Mcleod, The Nature Conservancy, Austin, TX, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptive management Integration of design, management, and monitoring to systematically test assumptions to adapt and learn.

Connectivity Natural linkage between marine habitats which occur via larval dispersal and the movements of adults and juveniles.

Ecosystem-based management (EBM) Environmental management approach that recognizes the full array of interactions within an ecosystem, including humans, rather than considering single issues, species, or ecosystem services in isolation. EBM focuses on cumulative impacts; multiple objectives; embracing change; linkages between species, ecosystems, societies, economies, and institutions; and learning and adaptation.

Ecosystem function Physical, chemical, and biological processes or attributes that contribute to the self-maintenance of an ecosystem (e.g., nutrient cycling, primary productivity).

Ecosystem resilience Ability of an ecosystem to maintain key functions and processes in the face of stresses or pressures, either by resisting or adapting to change; resilient systems are characterized as adaptable, flexible, and able to deal with change and uncertainty.

Ecosystem services Benefits people obtain from ecosystems; include the provision of food, water, timber, fiber, and other resources; the regulation of floods, disease, wastes, and water quality; the support of cultural practices, including recreation, religion, and art; and the maintenance of biological processes through such phenomena as soil formation, photosynthesis, nutrient cycling, and so on. Ecosystem "goods" include food, medicinal plants, construction materials, tourism and recreation, and wild genes for domestic plants and animals.

Marine protected area (MPA) Clearly defined geographical space recognized, dedicated, and managed through legal or other effective means to achieve the long-term conservation of nature with associated ecosystem services and cultural values.

Marine reserve Subset of an MPA and an area of ocean completely protected from all extractive and destructive activities.

MPA network Collection of individual MPAs operating cooperatively and synergistically – at various spatial scales and with a range of protection levels – to fulfill ecological aims more effectively and comprehensively than individual sites could alone.

No-take area Marine area that is permanently or temporarily completely closed for any form of extraction, and where no disturbance of any kind is allowed.

Introduction

Humans depend on the oceans for food security, shoreline protection, recreational opportunities, cultural heritage, climate regulation, and other services. Despite their tremendous value, the health of the world's oceans has continued to decline due to human activities such as overfishing, pollution, and climate change (Jackson *et al.*, 2001; Worm *et al.*, 2006; Halpern *et al.*, 2008a). These impacts are leading to ecosystem collapses in all the major coastal and ocean regions of the world (Wilkinson, 2004; Hughes *et al.*, 2005; Jackson, 2008).

Over the last several decades, about one-third of coastal and marine habitats, such as mangroves, seagrasses, coral reefs, and salt marshes, have been lost due to human activities (Valiela *et al.*, 2001; Wilkinson, 2004; Duarte *et al.*, 2008; Waycott *et al.*, 2009; Spalding *et al.*, 2010). More than half of the world's fisheries stocks are fully exploited and producing catches at or close to their maximum sustainable limits, and more than 25% are overexploited, depleted, or recovering from depletion (FAO, 2007). Fundamental changes to ecosystem structure, such as changes in species diversity, population abundance, size structure, sex ratios, habitat structure, trophic dynamics, biogeochemistry, and biological interactions, are occurring worldwide (Lubchenco *et al.*, 2003). These changes affect marine ecosystem function and have critical implications for people that depend on these ecosystems for goods and services (Lubchenco *et al.*, 1995).

Marine protected areas (MPAs) have been identified as one of the most effective tools for conserving marine ecosystems (Kelleher, 1999; Palumbi, 2003). A number of terms are used, often interchangeably, to refer to marine areas that are protected by spatially explicit restrictions, including MPAs, marine reserves, closed areas, harvest refuges, and sanctuaries (Agardi, 2000). In this chapter, an MPA is a "clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values" (Dudley, 2008). A marine reserve is a subset of an MPA and is defined here as "an area of the ocean completely protected from all extractive and destructive activities" (Lubchenco *et al.*, 2003).

MPAs may include areas with multiple uses (e.g., fishing, tourism), no-take areas and reserves, or restriction of certain areas to a specific use (e.g., local fishing). MPAs range in size from small marine parks designed to protect endangered or threatened species, unique habitat, or cultural or historical sites to large reserves designed to achieve a range of conservation,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00347-6

social, and economic objectives encompassing different types of protection (Agardi, 2000). Ecological objectives include protection of critical habitats (spawning aggregations, nursery grounds, areas of high biodiversity, and migration routes), maintenance of ecosystem function, and species protection. Socioeconomic objectives include the protection of commercially valuable species, cultural and historic sites, recreation and tourism sites, and sites important for education or research (Salm *et al.*, 2000). If MPAs are well designed and managed, they have the potential to protect and in some cases restore coastal and marine ecosystems and support communities that depend on these ecosystems.

MPAs are most effective when combined with other management tools such as integrated coastal management, marine spatial planning (MSP), and fisheries management (Salm *et al.*, 2006; Dudley, 2008). MPAs are vulnerable to activities outside their boundaries (e.g., pollution and unsustainable fishing) that can affect species and ecosystem functions within protected areas (Kaiser, 2005). Therefore, integrated coastal management to control land-based threats such as pollution and sedimentation and other forms of resource management such as fishery management tools (e.g., catch limits, gear restrictions, regulations regarding fishing grounds, fishing seasons; Kaiser, 2005; Keller *et al.*, 2009) are necessary to support the effectiveness of MPAs. MPAs may also be more effective when combined with traditional marine management approaches (McClanahan *et al.*, 2006).

Systems Approach to Marine Conservation

Conservation and fisheries management efforts have evolved over the last decade toward managing systems as opposed to species or specific habitats and managing cumulative impacts. These approaches have been referred to as *ecosystem-based management* (EBM) and ecosystem approach to fisheries (Rosenberg and McLeod, 2005; Levin and Lubchenco, 2008; Palumbi *et al.*, 2008; FAO, 2010). In these approaches, humans are recognized as critical parts of dynamic ecosystems, and the inclusion of the human dimension is now seen as an essential component of effective conservation. Fisheries management has shifted from a focus on maximum sustainable yield of individual species at a single scale to multispecies stock assessments at multiple scales (Pikitch, 2004). EBM supports ecological processes that maintain resources, recognizing the diverse ecological roles of species and habitats at multiple scales (Graham *et al.*, 2003). MPAs protect geographical areas, species, and their biophysical environments and thus can offer an ecosystem-based approach to conservation or fisheries management (Lubchenco *et al.*, 2003).

Evolution of MPA Definitions

The shift toward a multiscale, integrated human–ecological system, and process-oriented perspective (Hughes *et al.*, 2005) is evident in the changing views of MPAs and the call for networks of MPAs worldwide. In 1999, an MPA was defined as "any area of the intertidal or subtidal terrain, together with its overlying water and associated flora, fauna, historical, and cultural features, which has been reserved by law or other effective means to protect part or all of the enclosed environment" (Kelleher, 1999). In 2008, the World Conservation Union defined an MPA as a "clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values" (Dudley, 2008). The explicit reference to conservation for the benefit of people (ecosystem services and cultural values) is highlighted in this latest definition.

Evolution of MPA Objectives

Conservation managers have been called on to consider protection of ecosystem function, structure, and integrity in addition to species and habitat protection (Agardy and Staub, 2006). There has been a shift from the conservation of commercially important species to management of functional groups (i.e., collections of species that perform a similar function, regardless of their taxonomic affinities) supporting processes and maintenance of ecosystem services (e.g., fisheries) (Hughes *et al.*, 2005). The view of no-take areas as primarily fisheries management tools has evolved to include other conservation objectives, including managing biodiversity, trophic structure and function, and ecosystem resilience (Hughes *et al.*, 2005). This shift toward managing ecosystem structure, function, and services emphasizes the importance of ecological roles and species interactions (including humans) for maintaining ecosystem resilience. The effectiveness of MPAs is now evaluated based on their impacts on local communities, in addition to ecological impacts. In addition, integrated studies are developing that assess how ecological performance of reserves is related to both socioeconomic characteristics in coastal communities and reserve design (Pollnac *et al.*, 2010). The results of such studies are useful for highlighting the complexities around human dimensions of marine reserves and informing MPA design and management.

Ecosystem resilience refers to the ability of an ecosystem to maintain key functions and processes in the face of stresses or pressures, either by resisting or adapting to change (Holling, 1973; Nyström and Folke, 2001). Resilient systems are characterized as adaptable, flexible, and able to deal with change and uncertainty (Hughes *et al.*, 2005); thus *resilience* has been identified as a critical component of MPA network design and management. Designing and managing networks for resilience provides MPAs with the best chance to recover from or withstand environmental fluctuations or unexpected catastrophes caused by climate change and other human impacts (West and Salm, 2003; Mcleod *et al.*, 2009).

Evolution of MPA Networks

The conservation community and government agencies have called for the establishment of networks of MPAs worldwide. An MPA network is defined as a "collection of individual MPAs operating cooperatively and synergistically, at various spatial scales, and with a range of protection levels, in order to fulfill ecological aims more effectively and comprehensively

than individual sites could alone'' (WCPA/IUCN, 2007). MPA networks have also been defined as a ''network of people managing the components of individual MPAs and promoting the network's viability and longevity'' (Dudley, 2008).

Individual small MPAs may not be effective at conserving biodiversity, fish and invertebrate populations, and the communities that depend on them. Single MPAs large enough to sustain populations and habitats are often impractical due to economic, social, and political constraints (Dudley, 2008). Networks of MPAs have been proposed to help reduce socioeconomic impacts while maintaining conservation and fisheries benefits (PISCO, 2007). Networks are also important for maintaining ecosystem processes and connectivity and supporting ecosystem resilience by spreading the risk of reduced viability of a habitat or community type following a large-scale disturbance or management failures (Keller *et al.*, 2009). The commitment to the establishment of global networks of MPAs has been demonstrated at international meetings such as the World Summit on Sustainable Development in 2002, the World Parks Congress in 2003, and the Convention on Biological Diversity in 2004. The scaling up of individual MPAs to networks demonstrates a systems approach to marine conservation because it allows for the protection of species and habitats in addition to ecological processes, structure, and function.

Evolution of MSP

Over the last decade, marine spatial planning (MSP) has been increasingly recognized as a critical tool to achieve EBM (Douvere, 2008). MSP provides an integrated planning framework that moves away from sectoral management to address multiple objectives related to achieving economic and ecological sustainability and the need to reduce conflicts in marine environment (Agardi *et al.*, 2011). MSP has been defined as a ''process of analyzing and allocating parts of the three-dimensional marine spaces to specific uses, to achieve ecological, economic and social objectives that are usually specified through the political process; the MSP process usually results in a comprehensive plan or vision for a marine region'' (Ehler and Douvere, 2007). More broadly, the purpose of MSP is to balance demands for development with the need to protect the environment (Douvere, 2008).

Potential benefits of MSP include a holistic approach that addresses social, cultural, economic, and environmental objectives and thus achieve sustainable development; better integration of marine objectives (both between policies and between different planning levels); improved site selection for development or conservation; a more strategic and proactive approach that delivers long-term benefits; management coordination at the scale of ecosystems as well as political jurisdictions; reduced conflicts among uses in the marine area; and reduced risk of marine activities damaging marine ecosystems, including improved consideration of cumulative effects (Gilliland and Laffoley, 2008; Foley *et al.*, 2010).

MSP has been applied to help manage the multiple uses of marine space, particularly in areas where conflicts exist among users and the environment. MSP is central to the management strategy of the Great Barrier Reef in Australia and has also been applied in other marine areas such as the Florida Keys, Channel Islands, Wadden Sea, North Sea, Irish Sea, and Baltic Sea, among others. MSP is not intended to replace MPAs. MPAs are still recognized as an important tool for managing the marine environment, but they should be considered in the wider context of an MSP strategy that balances the MPAs with economic, social, and biodiversity objectives. By integrating MPA planning in broader MSP and ocean zoning efforts, MSP can help to utilize the benefits of MPAs while avoiding their potential shortcomings (Agardi *et al.*, 2011).

Benefits of MPAs

MPAs have the potential to provide a number of benefits to local communities, fisheries, and the marine environment, including (1) conserving biological diversity and ecosystems; (2) protecting critical spawning and nursery habitats; (3) protecting sites with limited human impact to help them recover from stresses; (4) protecting settlement and growth areas for marine species and spillover benefits to adjacent areas; (5) protecting sites for educating the public about marine ecosystems and threats to them; (6) supporting nature-based recreation and tourism; (7) providing control sites as baselines for scientific research; and (8) reducing poverty and increasing the quality of life of adjacent communities (IUCN-WCPA, 2008). Benefits of MPAs – specifically, marine reserves – have been demonstrated through empirical studies for mollusks, crustaceans, and fishes in habitats ranging from coral reefs, kelp forests, temperate continental shelves, estuaries, seagrass beds, and mangroves (Gell and Roberts, 2003). The following four sections outline the ecological and socioeconomic benefits of marine reserves based on global and regional meta-analyses and site-based studies.

Increases in Size, Abundance, Biomass, Diversity, and Increased Reproductive Potential

Benefits to marine reserves include increases in abundance, biomass, and diversity of many species within reserve boundaries (**Table 1**), yet the range of responses to reserve establishment is very large (Lester *et al.*, 2009; Gaines *et al.*, 2010a). Kenchington (1990) identifies several classes of species for which marine reserves may not be effective such as species with planktonic larvae and planktonic or pelagic adults (e.g., most phytoplankton and zooplankton species to pelagic fishes with large home ranges). However, these species may have a stage that depends on a nursery area or spawning site, and they could be protected by a reserve, assuming that other life stages outside the reserve are not overexploited (Allison *et al.*, 1998).

Within reserves, individuals can grow larger, live longer, and develop increased reproductive potential and populations increase in size (Bohnsack, 1998). Enhanced production of eggs and larvae within reserves are predicted to result in greater export and settlement of juveniles outside boundaries (Gell and Roberts, 2003). In addition, reserves have helped to restore ecosystem structure and function (Sobel and Dahlgren, 2004; Mumby *et al.*, 2006). However, these benefits do not

Table 1 Global and regional meta-analyses of ecological benefits of marine reserves within reserve boundaries

Indicator	*Results*	*Taxonomic group*	*# Studies/marine reserves analyzed*	*Source*
Biomass	446% increase	Algae, invertebrates, and fishes	124 reserves (global)	Lester *et al.* (2009)
Density	166% increase			
Individual size	28% increase			
Species richness	21% increase			
Biomass	352% increase	Algae, invertebrates, and fishes	89 studies; 70 reserves (global)	Halpern (2003)
Density	151% increase			
Individual size	29% increase			
Species richness	25% increase			
Biomass	1.9 times higher	Algae, invertebrates, and fishes	30 reserves (global; temperate only)	Stewart *et al.* (2009)
Density	1.7 times higher			
Species richness	1.5 times higher			
Abundance	25% increase	Fishes	19 reserves (global)	Côté *et al.* (2001)
Species richness	11% increase			
Abundance	3.7 times higher (for target species) No change (nontarget species)	Fishes	12 studies (global)	Mosquera *et al.* (2000)
Density	66% increase	Fishes	32 reserves (global)	Molloy *et al.* (2009)
Density	2.46 times larger	Fishes	12 reserves (regional – European)	Claudet *et al.* (2008)
Species richness	No effect			
Biomass	2.1 times greater	Fishes	12 reserves (regional – Mediterranean)	Guidetti and Sala (2007)
Density	1.2 times greater			

always occur due to fishers' behavior in response to reserves, fishing regulations outside the reserve, and the regulations regarding activities within and outside the reserve (Gaines *et al.*, 2010b.)

Benefits to Nontarget Species

Recent reviews on the benefits of marine reserves typically focus on benefits to target species as opposed to nontargeted groups such as fish, invertebrates, or algae or corals (Babcock *et al.*, 2010). However, it is important to understand the impacts of reserves on nontarget groups if the goal of protection is to maintain ecosystem structure and function. Studies have documented that nontarget species either do not respond to protection (Jennings *et al.*, 1995; Rakitin and Kramer, 1996) or respond negatively (i.e., reduced abundances in response to increased predation within reserves; McClanahan *et al.*, 1999). Other studies have shown that nontarget habitats improve following protection (Mumby *et al.*, 2005; Shears and Babcock, 2003), where changes in ecosystem structure have been documented due to the restoration of predator populations. For example, in tropical systems, enhanced coral recruitment has occurred following a recovery in herbivores that graze down macroalgae and thus encourage coral settlement (Mumby *et al.*, 2005). In temperate systems in New Zealand, the recovery of lobsters and large fishes led to predation and declines of sea urchin populations that led to a reduction in grazing and subsequent recovery of kelp forests (Shears and Babcock, 2003). Reserves have the potential to provide useful insights into the indirect effects of overfishing on ecosystem structure and function (Babcock *et al.*, 2010).

Benefits to Adjacent Fisheries

The ability of reserves to provide conservation or fisheries benefits to adjacent waters is highly controversial (Gell and Roberts, 2003; Hilborn *et al.*, 2004; Halpern *et al.*, 2010), yet recent research suggests that higher abundances within reserves can lead to spillover of adults to adjacent waters (Roberts *et al.*, 2001; Abesamis and Russ, 2005; Kellner *et al.*, 2008; Perez-Ruzafa *et al.*, 2008; Halpern *et al.*, 2010). Spillover occurs through the net export of adults and juveniles (spillover effect) and propagules (recruitment effect) (Russ, 2002). The spillover effect operates on local scales (hundreds of meters to kilometers for reef fish), whereas the recruitment effect operates at scales of tens of kilometers (scales of dispersal of pelagic larvae; Palumbi, 2001; Russ *et al.*, 2004). Spillover from reserves may result in economic benefits from enhanced fisheries and tourism (White *et al.*, 2008) yet may take decades to develop fully (Roberts *et al.*, 2001; Russ *et al.*, 2004).

Long-Term Ecological Benefits

Research has shown that fisheries and conservation benefits of marine reserves increase with greater years of protection (Russ *et al.*, 2004; Claudet *et al.*, 2008; Molloy *et al.*, 2009; Selig and Bruno, 2010). To measure the long-term benefits of marine reserves, time-series data are needed to describe ecological changes due to protection and the stability of such changes. Babcock *et al.* (2010) analyzed data from temperate and tropical marine reserves collected on decadal time scales and found that even though most target species showed initial direct effects (e.g., change in abundance, size of individuals, biomass), their trajectories over time were highly variable. The abundance of some target species continued to increase after protection, whereas some leveled off and others decreased over time. Decreases in abundance were likely due to natural fluctuations, fishing impacts from outside reserves, and increases in predation within reserves. Despite these differences, populations of targeted species were more stable in reserves than fished areas, indicating increased ecological resilience (Babcock *et al.*, 2010). Although some benefits are evident shortly after protection (individuals live longer, mortality rates are lower), other benefits take longer to fully develop, such as increases in reproductive output, biodiversity, and stabilization of communities and ecosystem structure and function. Therefore, understanding that indirect effects (e.g., ecosystem-wide recovery) from removal of fishing pressure take time (e.g., >13 years; Babcock *et al.*, 2010) is essential for managers, policy makers, and communities to have realistic expectations of the benefits of marine reserves.

Socioeconomic Benefits

Although the ecological benefits of MPAs, particularly marine reserves, are well established, there has been less emphasis on the social and economic benefits to human communities. Further, rigorous policy analyses are lacking that consider the full range of economic costs and benefits of MPAs (Rudd *et al.*, 2003; Pelletier *et al.*, 2005). A limited number of recent assessments review the economic impacts of MPAs but are concentrated in North America, Australia, and Europe (e.g., Carlsen and Wood, 2004; Carter, 2003; KPMG, 2000; Leeworthy and Wiley, 2002; Roncin *et al.*, 2008), thus leaving out areas with the highest tropical marine biodiversity worldwide, such as Southeast Asia and the Pacific.

Socioeconomic assessments of the benefits of MPAs typically differentiate between extractive (e.g., fishers) and nonextractive users (e.g., recreational users such as divers, snorkelers, bathers, ecotourists, and sightseers) because these groups are likely to be impacted differently by the MPA (see **Table 2** for a summary of the potential social and economic costs and benefits of MPAs for these user groups). For extractive users, adverse impacts from MPA establishment may include loss of access rights to fishing grounds or increased risk due to traveling farther to access alternative fishing grounds. For a reserve to provide fisheries benefits to human communities, the reserve must lead to a net increase in yield (i.e., increases in harvest must be large enough to compensate for the area removed from fishing). The establishment of a marine reserve may actually reduce fishing opportunity and yield if the fisheries are already sustainably managed (Hastings and Botsford, 1999; Sladek-Nowlis and Roberts, 1999; Ralston and O'Farrell, 2008). The implementation of an MPA can increase costs if fewer fish are available due to harvest restrictions; fuel or labor costs are higher due to traveling farther to fishing grounds, and congestion increases in fishing grounds (Rudd *et al.*, 2003). A number of case studies suggest that fishers perceive the costs of MPAs (in terms of lost harvest) as greater than the benefits provided (e.g., from spillover) (Wolfenden *et al.*, 1994; Sant, 1996; Suman *et al.*, 1999).

Table 2 Summary of potential social and economic benefits and costs of MPAs

Categories	*Benefits*	*Costs*
Extractive users (e.g., commercial and recreational fishers)	Increase in catch (and associated income) Enhanced catch variety (greater species variety, greater frequency of older/larger fish)	Decrease in catch (and foregone fishing income) Crowding of displaced effort User conflicts Higher costs associated with choice of fishing location Increase in safety risks
Nonextractive users (e.g., divers, tourists)	Maintain species diversity Greater habitat complexity and diversity Higher density levels of marine species Enhanced recreational opportunities (e.g., scuba, snorkeling) Research opportunities Protection of other ecosystem services (e.g., coastal protection from erosion and storm surge by healthy reefs)	Damage to marine ecosystem Loss of traditional fishing community
Management	Savings in enforcement costs over nonspatial management Revenues derived from charging users of the MPA Scientific knowledge Hedge against uncertain stock assessments Educational opportunities	Increase in monitoring and enforcement costs Direct costs of setting up MPA Costs of compensatory measures for displaced activities Foregone income from resource extraction (oil, gas, and mineral exploration, and bio-prospecting) Increased congestion and possibly degraded ecosystem if MPA is not well managed due to increased use

By contrast, costs may be lower for fishers due to steady and reliable spillover to adjacent fishing grounds and enhanced catch variety (Sumaila, 1998; Roncin *et al.*, 2008). The fishing benefits of MPAs are challenging to assess because fish mobility between reserves and open areas to fishing are often poorly documented and because MPA benefits to fishers are highly dependent on the level of fishing in open areas (Roncin *et al.*, 2008). Thus, the economic value of spillover depends more on fishers' behavior and the cost of fishing as opposed to biological factors (Rudd *et al.*, 2003).

For nonextractive users, MPAs are likely to improve the quality of the marine ecosystem within the MPA that may be valuable to visitors (Rudd and Tupper, 2002). Because marine reserves support increases in the size and abundance of many species within reserve boundaries (Halpern, 2003) and recreational users such as snorkelers and divers prefer viewing larger and more abundant species (Williams and Polunin, 2000; Rudd and Tupper, 2002), the profitability of recreational and tourism providers may be increased by MPA establishment (Rudd *et al.*, 2002). Although an increase in visitors may increase revenues, too many visitors could adversely affect marine ecosystems within MPAs, particularly when their activities are not well managed. Further, increases in congestion may result in a decline in people's willingness to pay for wildlife viewing (Rudd and Tupper, 2002).

Economic valuations of MPAs provide valuable cost–benefit analyses and also highlight how the benefits of an MPA may be distributed. For example, the economic value of a healthy Great Barrier Reef to Australia is currently estimated to be around $5.5 billion Australian dollars annually and is increasing (McCook *et al.*, 2010). This estimate includes only use values (e.g., jobs, tourism, and fishing) and underestimates the total economic value. The costs associated with zoning and management of the Great Barrier Reef Marine Park are significantly less than the estimated economic value of the Great Barrier Reef; management costs are consistently less than 1% of the economic returns (McCook *et al.*, 2010). Similarly, in the Florida Keys National Marine Sanctuary, the management costs of the conservation program represented only 2% of the total benefits derived from the MPA (Bhat, 2003). In a recent economic analysis of 12 marine reserves in Europe, results suggest that the amount of income generated by fishing and diving in the MPA represents 2.3 times the management costs of the MPA (Roncin *et al.*, 2008). An economic assessment conducted for an MPA in Kenya also demonstrated that the income generated from the MPA was substantially higher than the management and opportunity costs for the park; income from the MPA was $1.6 million annually from tourism and $39,000 from fisheries compared to management and opportunity costs of less than $200,000 (Emerton and Tessema, 2001). Economic benefits from the MPA, such as shoreline protection, marine productivity, wildlife habitat and nursery, and cultural and aesthetic values, were not included in this assessment (Emerton and Tessema, 2001) but would have made the estimate of benefits even greater. The valuation demonstrated that some groups (commercial tourism operators) received the main economic benefits from the MPA, whereas others such as the local fishing communities (which had reduced fishing opportunities) and the park office (responsible for managing the MPA) bore the cost. Such analyses provide valuable indications of the equity of the distribution of MPA benefits that can directly affect compliance and overall protection.

In addition to assessing the economic benefits of MPAs, social benefits should also be addressed. Studies have documented the importance of assessing the perception of people affected by MPAs, as their perceptions affect the degree of support or opposition to the MPA, and consequently the effectiveness of protection (Pelletier *et al.*, 2005). A critical social benefit of MPAs is reducing and anticipating conflicts between different user groups. Much of the world's coastal areas are characterized by conflict between user groups or jurisdictional agencies (Agardi, 2000). For example, recreational use may conflict with shipping and mineral extraction, and commercial and subsistence fishing may conflict with scuba diving and nature-based tourism. In such cases, zoning can be used to accommodate a wide variety of uses and can be used as a tool to settle disputes when they occur (Reynard, 1994; Agardi, 2000). Other social benefits include improving visitors' satisfaction and increasing public knowledge about marine ecosystems and biodiversity (Pelletier *et al.*, 2005). The potential of MPAs to help alleviate poverty in coastal communities dependent on coral reefs has also been acknowledged (Leisher *et al.*, 2007). It is important to note that although MPAs may achieve their biological objectives, they may fail at achieving their social objectives. Therefore, assessments of the social effects of MPAs are critical to determine their long-term benefits (Christie, 2004).

Global Commitment to MPA Establishment

Governments around the world have demonstrated their commitment to conserving coastal and marine ecosystems for the benefit of the communities that depend on them. National leaders have formed regional initiatives to establish networks of MPAs to support fisheries and food security, sustainable tourism, ecosystem services, livelihoods, and cultural heritage (e.g., the Micronesia Challenge, the Caribbean Challenge, the Coral Triangle Initiative, and the Western Indian Ocean Challenge). These Initiatives are critical to building political will to support marine conservation efforts, improved integration with development priorities, and the development of sustainable financing mechanisms (Toropova *et al.*, 2010).

The number and areal extent of MPAs has increased dramatically over the last decade; the current global coverage of MPAs has increased 60% over the last three years and more than 150% since 2003 (Chape *et al.*, 2008). Currently, the total number of MPAs worldwide is about 5878 and covers more than 4.2 million km^2 of ocean ($\sim$1.2% of the global ocean; Toropova *et al.*, 2010). The World Parks Congress in 2003 set a target for conserving 20–30% of the world's oceans, yet the costs of running a global MPA network have been estimated at $5–19 billion annually (Balmford *et al.*, 2004). Such an effort would require an increase in current areal and financial investment in marine conservation by two orders of magnitude. A recent assessment of the progress toward global marine protection targets identified a mismatch between the resources available and those required to implement and monitor a global network of protected areas (Wood *et al.*, 2008). The authors (Wood *et al.*, 2008) suggest that once a

global network is developed, it is likely to be a compromise between quantity (how closely the targets are met) and quality (how well designed and effectively managed the protected areas are).

MPA Effectiveness

Despite the increases in the total number of MPAs worldwide, it is essential to assess how well these MPAs are meeting their objectives. According to a recent global analysis, 27% of the world's coral reefs are located within MPAs, yet only 6% of these are effectively managed (Burke *et al.*, 2011). Further, nearly half of the MPAs worldwide are ineffective at reducing the threat of overfishing (Burke *et al.*, 2011).

MPA performance is highly variable (Kelleher *et al.*, 1995; Halpern, 2003), and a number of factors have been identified that limit the effectiveness of MPAs in conserving biodiversity, maintaining fisheries, or providing additional ecosystem services to human communities. Some have suggested that social and political factors, as opposed to biological factors, are the primary determinants of MPA success or failure (Kelleher and Recchia, 1998; McClanahan, 1999; Mascia, 2003; Leisher, 2008). For example, some MPAs are ineffective because the management framework is ignored or not enforced (these are referred to as "paper parks"), or they have regulations that are fully and effectively implemented but are insufficient to address the threats within the MPA. Other factors affecting MPA effectiveness include differences in reserve design (e.g., size of no-take and buffer zones; Claudet *et al.*, 2008) or reserve shape (Kramer and Chapman, 1999). The location of MPAs also affects effectiveness; MPAs are often placed in areas where threats are lowest (e.g., large MPAs in remote areas such as the northwest Hawaiian Islands; Burke *et al.*, 2011) and thus may do little to mitigate local threats.

MPAs may function more effectively when devolution of authority for MPA development and management occurs (e.g., from national government to local governments, nongovernmental organizations, and resource users; White *et al.*, 2002). Other factors supporting effectiveness of MPAs include adaptive and participatory decision-making arrangements; clearly defined MPA boundaries; clear, easily understood, and easily enforceable rules; legitimacy of rules and regulations; political commitment and leadership; and collaborative MPA management structures linking resources with local interests and knowledge (Mascia, 2003).

The result of establishing a reserve in one location is that fishing effort simply moves elsewhere, and the reallocation of fishing efforts can have adverse impacts on species and habitats outside the reserve (Hilborn *et al.*, 2004). In places where fishing systems are effective at protecting stock (e.g., through catch, size, and area limits), it is not clear whether establishing MPAs will provide additional benefits (Hilborn *et al.*, 2004). The effectiveness of reserves also varies due to differential responses of species to protection (Micheli *et al.*, 2004; Molloy *et al.*, 2009). Researchers have noted significantly large increases in abundance of some large-bodied commercially important species following reserve establishment (e.g., Russ and Alcala, 1996; Claudet *et al.*, 2008), yet many species respond less predictably to protection (Mosqueira *et al.*, 2000; Molloy *et al.*, 2009). Although commonly fished predatory species are likely to benefit from reserve establishment, prey species may decline due to trophic cascades (Micheli *et al.*, 2004; Molloy *et al.*, 2009). Research suggests that MPAs are effective at protecting sedentary species, but they are less effective at preserving highly mobile species that may spend considerable time outside MPA boundaries (Kaiser, 2005), although exceptions have been documented (McCook *et al.*, 2010). Therefore, the scales of adult movement and propagule dispersal can be critical to MPA effectiveness. Empirical studies are needed to clarify the benefits of MPAs to highly mobile species, and innovative management approaches are needed to complement strategically placed MPAs to support them (Gaines *et al.*, 2010b).

Role of MPAs in a Changing World: Rising to the Challenge

Climate change impacts are already occurring in coastal and marine ecosystems worldwide and include shifts in ocean current patterns, ecosystem changes (e.g., widespread coral loss from mass bleaching), changes in larval development and transport, and species range shifts and interactions (Wilkinson, 1998; Parmesan and Yohe, 2003; IPCC, 2007; Rosenzweig *et al.*, 2008). MPAs are a core strategy in marine conservation, yet they are geographically fixed and thus poorly suited to accommodate shifts in species ranges and habitats. In addition, most existing MPAs are designed based on current climate conditions. The recognition of their limitations in a changing world has led some to question their relevance as a conservation response in an era of rapid climate change (Araújo *et al.*, 2004; Hannah *et al.*, 2007).

The ability of MPAs to protect ecosystems and species in the face of climate change and other changes (e.g., increasing global population) is debated (Mora *et al.*, 2006; Graham *et al.*, 2007; McClanahan, 2008; Selig and Bruno, 2010). Some researchers suggest that habitat loss (e.g., coral reefs) in response to climate change, storms, and diseases are unlikely to be mitigated by MPAs (Jameson *et al.*, 2002; Aronson and Precht, 2006; Graham *et al.*, 2008). Site-specific studies have suggested that MPAs do not always protect biodiversity better than unmanaged areas in response to climate impacts (Jones *et al.*, 2004; Graham *et al.*, 2007; McClanahan, 2008). Further, some studies suggest that thermal stress can cause proportionally greater coral mortality of protected than unprotected corals (McClanahan *et al.*, 2007; Graham *et al.*, 2007; Graham *et al.*, 2008; Darling *et al.*, 2010). This may be due to the different coral species composition between protected and unprotected sites (e.g., higher abundance of thermally sensitive corals such as *Acropora* and *Montipora* within reserves) (Côté and Darling, 2010; Darling *et al.*, 2010).

Recent global analyses, however, have confirmed that MPAs can be effective in preventing coral loss (coral cover remained constant in MPAs over 38 years, whereas coral cover on unprotected reefs declined; Selig and Bruno, 2010). Surveys in the Bahamas showed significantly higher increases in coral cover in reserve boundaries compared to outside the reserve (Mumby and Harborne, 2010). Mumby and Harborne (2010) suggest that reserves play an important role in increasing coral

reef recovery rates provided that macroalgae have been depleted by more abundant communities of grazers benefiting from reduced fishing pressure, particularly in the Caribbean.

Empirical evidence both supports (Lafferty and Behrens, 2005; Mumby *et al.*, 2006; Babcock *et al.*, 2010) and refutes (McClanahan, 2008) the idea that species and habitats within reserves are more resilient than those outside reserve boundaries (Gaines *et al.*, 2010b). Therefore, additional studies are urgently needed to establish the ability of MPAs to support resilience in a variety of habitats and geographic locations and in response to diverse threats.

To address the challenge of climate change, conservation practitioners and researchers are applying new tools such as ecological forecasting and climate envelope models to identify sites most likely to protect biodiversity in the future and to assess the ability of reserves and networks to protect species under different climate change scenarios (Araújo *et al.*, 2004; Hannah *et al.*, 2007; Hannah, 2008). Researchers have cautioned that the use of these tools is limited by the lack of existing data needed to support the models and uncertainties inherent in climate change projections and most ecological forecasting approaches (Thuiller, 2004; Lawler *et al.*, 2006; Lawler, 2009).

Complementing the new tools to support MPA design, major advances in recommendations for MPA and network design have also occurred over the last decade (Roberts *et al.*, 2003; Halpern *et al.*, 2006; Gaines *et al.*, 2010a), particularly recommendations specifically designed to address climate change impacts (Lawler, 2009; Mcleod *et al.*, 2009). Such principles may include the identification and protection of refuges (e.g., sites resistant to climate change impacts; such sites can provide the larvae needed to reseed areas that succumb to coral bleaching), pathways of connectivity that link these refuges with damaged areas, and measures to build redundancy into networks, thereby ameliorating the risk that climate change impacts will result in irrevocable biodiversity loss (West and Salm, 2003; Mcleod *et al.*, 2009).

Researchers have suggested that more and larger MPAs will be needed in the future to address climate change impacts (Lawler, 2009). Specific recommendations include increasing the size of existing reserves, adding buffers around existing reserves, and adding larger reserves to reserve networks (Halpin, 1997; Noss, 2001). Establishing more and larger reserves may be insufficient to protect biodiversity and maintain ecosystem services if the reserves are not located in the right places; a more strategic approach involves locating reserves so that they capture the most potential for habitat heterogeneity under a variety of climate scenarios (Lawler, 2009), or in places predicted to escape the brunt of climate change (West and Salm, 2003; Mumby and Steneck, 2008; Mcleod *et al.*, 2009; Côté and Darling, 2010). In addition, whereas larger MPAs may provide protection for increased and functional groups, they may not be politically, socially, or economically feasible. Research also suggests that large MPAs may be less effective than other traditional management approaches. For example, traditional management regimes in Indonesia and Papua New Guinea involving periodic closures were significantly more effective than national parks with permanent closures – a more than 40% increase in targeted fish biomass within reserves in traditional management regimes compared to less than 2% increase in national parks (McClanahan *et al.*, 2006). McClanahan *et al.* (2006) note that whereas large MPAs may provide the best protection for species susceptible to overfishing, traditional management approaches may provide the best solution for meeting conservation and community goals and reversing the degradation of reef ecosystems.

MPA and zone boundaries should be designed to be flexible in space and time so that they can be expanded or contracted, have seasonal or other fixed time limits, or be moved to different levels of protection to help them meet their objectives in response to future changes. Where habitat shifts are predicted, managers should proactively plan for landward migration, particularly in areas where habitats have the potential to expand (e.g., mangrove migration landward in response to sea-level rise). It is important to identify and protect areas likely to serve as refuges in the future (i.e., predictive protected areas; Herr and Galland, 2009) and also areas that have demonstrated resilience to climate change impacts.

Finally, to help MPAs continue to achieve their social, ecological, and economic objectives, adaptive management is essential. Adaptive management refers to the integration of design, management, and monitoring to systematically test assumptions in order to adapt and learn (Salafsky *et al.*, 2001). In the context of MPAs, adaptive management involves the integration of the best available science into MPA strategies and monitoring to systematically test the effectiveness of management methods and refine them over time. Conservation managers should develop management approaches that are flexible and able to incorporate future species and habitat migrations, and they need to apply risk-spreading strategies to ensure the protection of key larvae, species, and habitats. Monitoring should go beyond simply assessing whether current policies are effective (e.g., is biodiversity declining?) and should focus on resolving the underlying causes (e.g., how can we reverse the decline?) (Hughes *et al.*, 2007). Monitoring programs must address thresholds, regime shifts and feedbacks, and the capacity of ecosystems to maintain ecosystem services in response to future changes (Hughes *et al.*, 2007). Understanding how ecosystem services will be affected by climate change is necessary for setting conservation priorities and designing and managing restoration projects (Lawler, 2009).

MPAs have a critical role to play in protecting marine ecosystems and the benefits derived from these systems and in securing the communities that depend on them. There may be trade-offs between MPAs designed for biodiversity, sustainable use, and climate change. Therefore, it is important to include risk assessments, scenario planning, and adaptive management approaches that incorporate these potential trade-offs (Secretariat of the Convention on Biological Diversity, 2009). To be successful in a changing world, MPAs must strive to achieve the complementary goals of maintaining biodiversity, promoting ecosystem values, and enhancing resilience.

Appendix

List of Courses

- Applied Ecology and Environmental Management
- Conservation Biology

- Marine Ecosystem Management
- Marine Protected Areas
- Fishery Management

See also: Coastal Beach Ecosystems. Conservation Efforts, Contemporary. Corals and Coral Reefs. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Services. Ethical Issues in Biodiversity Protection. Fish Conservation. Identifying Conservation Priorities Using a Return on Investment Analysis. Mangrove Ecosystems. Marine and Aquatic Communities, Stress from Eutrophication. Marine Conservation in a Changing Climate. Marine Ecosystems. Marine Ecosystems, Human Impacts on. Modeling Marine Ecosystem Services. Natural Reserves and Preserves. Ocean Ecosystems. Pelagic Ecosystems. Resource Exploitation, Fisheries. Role and Trends of Protected Areas in Conservation. Seagrasses. Wetlands Ecosystems. Wetland Creation and Restoration

References

Abesamis RA and Russ GR (2005) Density-dependent spillover from a marine reserve: Long term evidence. *Ecological Applications* 15: 1798–1812.

Agardi T (2000) Information needs for marine protected areas: Scientific and societal. *Bulletin of Marine Science* 66: 875–888.

Agardy T, di Sciara GN, and Christie P (2011) Mind the gap: Addressing the shortcomings of marine protected areas through large scale marine spatial planning. *Marine Policy* 35: 226–232.

Agardy T, and Staub, F (2006) Marine protected areas and MPA networks. *Network for Conservation Educators and Practitioners*, American Museum of Natural History. CD-ROM. Available electronically at: http://research.amnh.org/biodiversity/ncep

Allison GW, Lubchenco J, and Carr MH (1998) Marine reserves are necessary but not sufficient for marine conservation. *Ecological Applications* 8: S79–S92.

Araújo MB, Cabeza M, Thuiller W, Hannah L, and Williams PH (2004) Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology* 10: 1618–1626.

Aronson RB and Precht WF (2006) Conservation, precaution, and Caribbean reefs. *Coral Reefs* 25: 441–450.

Babcock RC, Shears NT, Alcala AC, *et al.* (2010) Decadal trends in marine reserves reveal differential rates of change in direct and indirect effects. *Proceedings of the National Academy of Sciences of the United States of America* 107: 18256–18261.

Balmford A, Gravestock P, Hockley N, McClean N, and Roberts CM (2004) The worldwide costs of marine protected areas. *Proceedings of the National Academy of Sciences of the United States of America* 101: 9694–9697.

Bhat MG (2003) Application of non-market valuation to the Florida Keys marine reserve management. *Journal of Environmental Management* 67: 315–325.

Bohnsack JA (1998) Application of marine reserves to reef fisheries management. *Australian Journal of Ecology* 23: 298–304.

Burke L, Reytar K, Spalding M, and Perry AL (2011) *Reefs at Risk Revisited.* Washington, DC: World Resources Institute, The Nature Conservancy, World Fish Center, International Coral Reef Action Network, UNEP World Conservation Monitoring Centre and Global Coral Reef Monitoring Network.

Carlsen J and Wood D (2004) *Assessment of the Economic Value of Recreation and Tourism in Western Australia's National Parks, Marine Parks and Forests.* Townsville: Sustainable Tourism CRC.

Carter DW (2003) Protected areas in marine resource management: Another look at the economics and research issues. *Ocean and Coastal Management* 46: 439–456.

Chape S, Spalding M, and Jenkins M (2008) The World's Protected Areas. *Status, Values, and Prospects in the Twenty-First Century.* Berkeley, CA: University of California Press.

Christie P (2004) MPAs as biological successes and social failures in Southeast Asia. In: Shipley JB (ed.) *Aquatic Protected Areas as Fisheries Management Tools: Design, Use, and Evaluation of these Fully Protected Areas*, pp. 155–164. Bethesda, MD: American Fisheries Society.

Claudet J, Osenberg CW, Benedetti-Cecchi L, *et al.* (2008) Marine reserves: Size and age do matter. *Ecology Letters* 11: 481–489.

Côté IM and Darling ES (2010) Rethinking ecosystem resilience in the face of climate change. *PLoS Biology* 8: e1000438.http://dx.doi.org/10.1371/journal.pbio.1000438.

Côté IM, Mosqueira I, and Reynolds JD (2001) Effects of marine reserve characteristics on the protection of fish populations: A meta-analysis. *Journal of Fish Biology* 59: 178–189.

Darling ES, McClanahan TR, and Côté IM (2010) Combined effects of two stressors on Kenyan coral reefs are additive or antagonistic, not synergistic. *Conservation Letters* 3: 122–130.

Douvere F (2008) The importance of marine spatial planning in advancing ecosystem-based sea use management. *Marine Policy* 32: 762–771.

Duarte CM, Dennison WC, Orth RJW, and Carruthers TJB (2008) The charisma of coastal ecosystems: Addressing the imbalance. *Estuarine Coast* 31: 233–238.

Dudley N (ed.) (2008) *Guidelines for Applying Protected Area Management Categories.* Gland, Switzerland: IUCN.

Ehler C and Douvere F (2007) Visions for a sea change. *Report of the First International Workshop on Marine Spatial Planning. Intergovernmental Oceanographic Commission and Man and the Biosphere Programme. IOC Manual and Guides No.48. IOCAM Dossier No. 4.* Paris: UNESCO.

Emerton L and Tessema Y (2001) *Economic Constraints to the Management of Marine Protected Areas: The Case of Kisite Marine National Park and Mpunguti Marine National Reserve, Kenya.* Kenya: IUCN Eastern Africa Programme.

Foley MM, Halpern BS, and Micheli F (2010) Guiding ecological principles for marine spatial planning. *Marine Policy* 34: 955–966.

Food and Agriculture Organization of the United States (FAO) (2007) *Report: The State of World Fisheries and Aquaculture 2006.* Rome: FAO Fisheries and Aquaculture Department.

Food and Agriculture Organization of the United States (FAO) (2010) *Report: The State of World Fisheries and Aquaculture 2010.* Rome: FAO Fisheries and Aquaculture Department.

Gaines SD, Lester SE, Grorud-Colvert K, Costello C, and Pollnac R (2010b) The evolving science of marine reserves: New developments and emerging research frontiers. *Proceedings of the National Academy of Sciences of the United States of America* 107: 18251–18255.

Gaines SD, White C, Carr M, and Palumbi S (2010a) Designing marine reserve networks for both conservation and fisheries management. *Proceedings of the National Academy of Sciences of the United States of America* 107: 18286–18293.

Gell FR and Roberts CM (2003) Benefits beyond boundaries: The fishery effects of marine reserves. *Trends in Ecology and Evolution* 18: 448–455.

Gilliland PM and Laffoley D (2008) Key elements and steps in the process of developing ecosystem-based marine spatial planning. *Marine Policy* 32: 787–796.

Graham NAJ, Evans RD, and Russ GR (2003) The effects of marine reserve protection on the trophic relationships of reef fishes on the Great Barrier Reef. *Environmental Conservation* 30: 200–208.

Graham NAJ, McClanahan TR, MacNeil MA, *et al.* (2008) Climate warming, marine protected areas and the oceanscale integrity of coral reef ecosystems. *PLoS One* 3: e3039.

Graham NAJ, Wilson SK, Jennings S, *et al.* (2007) Lag effects in the impacts of mass coral bleaching on coral reef fish, fisheries, and ecosystems. *Conservation Biology* 21: 1291–1300.

Guidetti P and Sala E (2007) Community-wide effects of marine reserves in the Mediterranean Sea. *Marine Ecology Progress Series* 335: 43–56.

Halpern BS (2003) The impact of marine reserves: Do reserves work and does reserve size matter? *Ecological Applications* 13: 117–137.

Halpern BS, Lester SE, and Kellner JB (2010) Spillover from marine reserves and the replenishment of fished stocks. *Environmental Conservation* 36: 268–276.

Halpern BS, Regan HM, Possingham HP, and McCarthy MA (2006) Accounting for uncertainty in marine reserve design. *Ecology Letters* 9: 2–11.

Halpern BS, Walbridge S, Selkoe KA, *et al.* (2008a) A global map of human impact on marine ecosystems. *Science* 319: 948–952.

Halpin PN (1997) Global climate change and natural area protection: Management responses and research directions. *Ecological Applications* 7: 828–843.

Hannah L (2008) Protected areas and climate change. *Annals of the New York Academy of Sciences* 1134: 201–212.

Hannah L, Midgley G, Andelman S, *et al.* (2007) Protected area needs in a changing climate. *Frontiers in Ecology and the Environment* 5: 131–138.

Hastings A and Botsford LW (1999) Equivalence in yield from marine reserves and traditional fisheries management. *Science* 284: 1537–1538.
Herr D and Galland GR (2009) *The Ocean and Climate Change. Tools and Guidelines for Action*. Gland, Switzerland: IUCN.
Hilborn R, Stokes K, Maguire JJ, *et al.* (2004) When can marine reserves improve fisheries management? *Ocean and Coastal Management* 47: 197–205.
Holling CS (1973) Resilience and stability of ecological systems. *Annual Review of Ecology, Evolution, and Systematics* 4: 1–23.
Hughes TP, Bellwood DR, Folke C, Steneck RS, and Wilson J (2005) New paradigms for supporting the resilience of marine ecosystems. *Trends in Ecology and Evolution* 20: 380–386.
Hughes TP, Gunderson LH, Folke C, *et al.* (2007) Adaptive management of the Great Barrier Reef and the Grand Canyon World Heritage. *Ambio* 36: 586–592.
Intergovernmental Panel on Climate Change (IPCC) (2007) Climate change 2007: Impacts, adaptation and vulnerability. *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
IUCN World Commission on Protected Areas (IUCN-WCPA) (2008) *Establishing Marine protected area networks – Making it Happen*. Washington, DC: IUCN-WCPA, National Oceanic and Atmospheric Administration and the Nature Conservancy.
Jackson JBC (2008) Ecological extinction and evolution in the brave new ocean. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11458–11465.
Jackson JBC, Kirby MX, Berger WH, *et al.* (2001) Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293: 629–637.
Jameson SC, Tupper MH, and Ridley JM (2002) The three screen doors: Can marine "protected" areas be effective? *Marine Pollution Bulletin* 44: 1177–1183.
Jennings S, Grandcourt EM, and Polunin NVC (1995) The effects of fishing on the diversity, biomass and trophic structure of Seychelles' reef fish communities. *Coral Reefs* 14: 225–235.
Jones GP, McCormick MI, Srinivasan M, and Eagle JV (2004) Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Sciences of the United States of America* 101: 8251–8253.
Kaiser MJ (2005) Are marine protected areas a red herring or fisheries panacea? *Canadian Journal of Fisheries and Aquatic Sciences* 62: 1194–1199.
Kelleher G (1999) *Guidelines for Marine Protected Areas*. Gland, Switzerland: IUCN – The World Conservation Union.
Kelleher G, Bleakley C, and Wells S (1995) *Global Representative System of Marine Protected Areas*. Washington, DC: The World Bank.
Kelleher G and Recchia C (1998) Lessons from marine protected areas around the world. *Parks* 8: 1–4.
Keller BD, Gleason DF, Mcleod E, *et al.* (2009) Climate change, coral reef ecosystems, and management options for marine protected areas. *Environmental Management* 44: 1069–1088.
Kellner JB, Nisbet RM, and Gaines SD (2008) Spillover from marine reserves related to mechanisms of population regulation. *Theoretical Ecology* 1: 117–127.
Kenchington RA (1990) *Managing Marine Environments*. New York, USA: Taylor and Francis.
KPMG Consulting (2000) *Economic and Financial Values of the Great Barrier Reef Marine Park*. Townsville: Great Barrier Reef Marine Park Authority. Research Publication 63.
Kramer DL and Chapman MR (1999) Implications of fish home range size and relocation for marine reserve function. *Environmental Biology of Fishes* 55: 65–79.
Lafferty KD and Behrens MD (2005) Temporal variation in the state of rocky reefs: Does fishing increase the vulnerability of kelp forests to disturbance? In: Garcelon DK and Schwemm CA (eds.) *Proceedings of the Sixth California Islands Symposium*, pp. 511–520. Ventura, CA: Institute for Wildlife Studies.
Lawler JJ (2009) Climate Change Adaptation Strategies for Resource Management and Conservation Planning. The Year in Ecology and Conservation Biology. *Annals of the New York Academy of Sciences* 1162: 79–98.
Lawler JJ, White D, Neilson RP, and Blaustein AR (2006) Predicting climate induced range shifts: Model differences and model reliability. *Global Change Biology* 12: 1568–1584.
Leeworthy VR and Wiley P (2002) *Socioeconomic Impact Analysis of Marine Reserve Alternatives for the Channel Islands National Marine Sanctuary*. Silver Spring, MD: NOAA.
Leisher C (2008) What Rachel Carson knew about marine protected areas. *BioScience* 58: 478–479.
Leisher C, vanBeukering P, and Scherl LM (2007) *Nature's Investment Bank: How Marine Protected Areas Contribute to Poverty Reduction*, 52 pp. Arlington, VA: The Nature Conservancy, Available from: http://conserveonline.org.
Lester SE, Halpern BS, Grorud-Colvert K, *et al.* (2009) Biological effects within no-take marine reserves: A global synthesis. *Marine Ecology Progress Series* 384: 33–46.
Levin SA and Lubchenco J (2008) Resilience, robustness, and marine ecosystem-based management. *BioScience* 58: 27–32.
Lubchenco J, Allison GW, Navarrete SA, *et al.* (1995) Biodiversity and ecosystem functioning: Coastal systems. In: United Nations Environmental Programme (ed.), *Global Diversity Assessment*, pp. 370–381. Cambridge, UK: Cambridge University Press.
Lubchenco J, Palumbi SR, Gaines SD, and Andelman S (2003) Plugging a hole in the ocean: The emerging science of marine reserves. *Ecological Applications* 13: S3–S7.
Mascia M (2003) The human dimension of coral reef marine protected areas: Recent social science research and its policy implications. *Conservation Biology* 17: 630–632.
McClanahan TR (1999) Is there a future for coral reef parks in poor tropical countries? *Coral Reefs* 18: 321–325.
McClanahan TR (2008) Response of the coral reef benthos and herbivory to fishery closure. *Oecologia* 155: 169–177.
McClanahan TR, Ateweberhan M, Muhando CA, Maina J, and Mohammed MS (2007) Effects of climate and seawater temperature variation on coral bleaching and mortality. *Ecological Monographs* 77: 503–525.
McClanahan TR, Marnane MJ, Cinner JE, and Kiene WE (2006) A comparison of marine protected areas and alternative approaches to coral-reef management. *Current Biology* 16: 1408–1413.
McClanahan TR, Muthiga NA, Kumukuru AT, Machano H, and Kiambo RW (1999) The effects of marine parks and fishing on coral reefs of northern Tanzania. *Biological Conservation* 89: 161–182.
McCook LJ, Ayling T, Cappo M, *et al.* (2010) Adaptive management of the Great Barrier Reef: A globally significant demonstration of the benefits of networks of marine reserves. *Proceedings of the National Academy of Sciences of the United States of America* 107: 18278–18285.
Mcleod E, Salm RV, Green A, and Almany J (2009) Designing marine protected area networks to address the impacts of climate change. *Frontiers in Ecology and the Environment* 7: 362–370.
Micheli F, Halpern BS, Botsford LW, and Warner RR (2004) Trajectories and correlates of community change in no-take marine reserves. *Ecological Applications* 14: 1709–1723.
Molloy PP, McLean IB, and Cote IM (2009) Effects of marine reserve age on fish populations: A global meta-analysis. *Journal of Applied Ecology* 46: 743–751.
Mora C, Andrefouet S, Costello MJ, *et al.* (2006) Coral reefs and the global network of marine protected areas. *Science* 312: 1750–1751.
Mosqueira I, Côté IM, Jennings S, and Reynolds JD (2000) Conservation benefits of marine reserves for fish populations. *Animal Conservation* 4: 321–332.
Mumby PJ, Dahlgren CP, Harborne AR, *et al.* (2006) Fishing, trophic cascades, and the process of grazing on coral reefs. *Science* 311: 98–101.
Mumby PJ, Foster NL, and Glynn Fahy EA (2005) Patch dynamics of coral reef macroalgae under chronic and acute disturbance. *Coral Reefs* 24: 681–692.
Mumby PJ and Harborne AR (2010) Marine reserves enhance the recovery of corals on Caribbean reefs. *PLoS ONE* 5: e8657.
Mumby PJ and Steneck RS (2008) Coral reef management and conservation in light of rapidly evolving ecological paradigms. *Trends in Ecology and Evolution* 10: 555–563.
Noss RF (2001) Beyond Kyoto: Forest management in a time of rapid climate change. *Conservation Biology* 15: 578–590.
Nyström M and Folke C (2001) Spatial resilience of coral reefs. *Ecosystems* 4: 406–417.
Palumbi SR (2001) The ecology of marine protected areas. In: Bertness MD, Gaines SD, and Hay ME (eds.) *Marine Community Ecology*, pp. 509–530. Sunderland, MD: Sinauer Press.
Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications* 13: S146–S158.
Palumbi SR, McLeod KL, and Grunbaum D (2008) Ecosystems in action: lessons from marine ecology about recovery, resistance, and reversibility. *BioScience* 58: 33–42.
Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.
Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO) (2007) *The Science of Marine Reserves*. (2nd Edition, United States Version) www.piscoweb.org. 22 pp.

Pelletier D, García-Charton JC, Ferraris J, *et al.* (2005) Designing indicators for assessing the effects of marine protected areas on coral reef ecosystems: A multidisciplinary standpoint. *Aquatic Living Resources* 18: 15–33.
Perez-Ruzafa A, Martin E, Marcos C, *et al.* (2008) Modeling spatial and temporal scales for spill-over and biomass exportation from MPAs and their potential for fisheries enhancement. *Journal of Nature Conservation* 16: 234–255.
Pikitch EK, Santora C, Babcock EA, *et al.* (2004) Ecosystem-based fishery management. *Science* 305: 346–347.
Pollnac R, Christie P, Cinner J, *et al.* (2010) Marine reserves as linked social–ecological systems. *Proceedings of the National Academy of Sciences of the United States of America* 107: 18262–18265.
Rakitin A and Kramer DL (1996) Effect of a marine reserve on the distribution of coral reef fishes in Barbados. *Marine Ecology Progress Series* 131: 97–113.
Ralston S and O'Farrell MR (2008) Spatial variation in fishing intensity and its effect on yield. *Canadian Journal of Fisheries and Aquatic Sciences* 65: 588–599.
Reynard Y (1994) Resolving conflicts for integrated coastal management: The case of Soufrière, St. Lucia, Caribbean. *Parks and Protected Areas Bulletin* 5: 5–7.
Roberts CM, Andelman S, Branch G, *et al.* (2003) Ecological criteria for evaluating candidate sites for marine reserves. *Ecological Applications* 13: S199–S214.
Roberts CM, Bohnsack JA, Gell F, Hawkins JP, and Goodridge R (2001) Effects of marine reserves on adjacent fisheries. *Science* 294: 1920–1923.
Roncin N, Alban F, Charbonnel E, *et al.* (2008) Uses of ecosystem services provided by MPAs: How much do they impact the local economy? A southern Europe perspective. *Journal of Nature Conservation* 16: 256–270.
Rosenberg AA and McLeod KL (2005) Implementing ecosystem-based approaches to management for the conservation of ecosystem services. *Marine Ecology Progress Series* 300: 241–296.
Rosenzweig C, Karoly D, Vicarelli M, *et al.* (2008) Attributing physical and biological impacts to anthropogenic climate change. *Nature* 453: 353–357.
Rudd MA, Folmer H, and van Kooten GC (2002) Economic evaluation of recreational fishery policies. In: Pitcher TJ and Hollingworth C (eds.) *Evaluating Recreational Fisheries: An Ecological, Economic, Social Balance Sheet*, pp. 35–52. Oxford: Blackwell Science.
Rudd MA and Tupper MH (2002) The impact of Nassau grouper size and abundance on scuba diver site selection. *Coastal Management* 30: 133–151.
Rudd MA, Tupper MH, Folmer H, and van Kooten GC (2003) Policy analysis for tropical marine reserves: Challenges and directions. *Fish and Fisheries* 4: 65–85.
Russ GR (2002) Yet another review of marine reserves as reef fisheries management tools. In: Sale PF (ed.) *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem*, pp. 421–443. San Diego, CA: Academic Press.
Russ GR and Alcala AC (1996) Do marine reserves export adult fish biomass? Evidence from Apo Island, central Philippines. *Marine Ecology Progress Series* 132: 1–9.
Russ GR, Alcala AC, Maypa AP, Calumpong HP, and White AT (2004) Marine reserve benefits local fisheries. *Ecological Applications* 14: 597–606.
Salafsky N, Margoluis R, and Redford K (2001) *Adaptive Management: A Tool for Conservation Practitioners*. Washington, DC: Biodiversity Support Program.
Salm RV, Clark JR, and Siirila E (2000) *Marine and Coastal Protected Areas: A Guide for Planners and Managers*. Washington, DC, USA: IUCN.
Salm RV, Done T, and Mcleod E (2006) Marine protected area planning in a changing climate. In: Phinney JT, Hoegh- Guldberg O, Kleypas J, Skirving W, and Strong A (eds.) *Coral Reefs and Climate Change: Science and Management*. Washington, DC: American Geophysical Union.
Sant M (1996) Environmental sustainability and the public: Responses to a proposed marine reserve at Jervis Bay, New South Wales, Australia. *Ocean and Coastal Management* 32: 1–16.
Secretariat of the Convention on Biological Diversity (2009). Connecting biodiversity and climate change mitigation and adaptation. *Report of the Second Ad Hoc Technical Expert Group on Biodiversity and Climate Change*, 126 pp. Montreal: Secretariat of the Convention on Biological Diversity.
Selig ER and Bruno JF (2010) A global analysis of the effectiveness of marine protected areas in preventing coral loss. *PLoS ONE* 5: e9278.
Shears NT and Babcock RC (2003) Continuing trophic cascade effects after 25 years of no-take marine reserve protection. *Marine Ecology Progress Series* 246: 1–16.
Sladek Nowlis J and Roberts CM (1999) Fisheries benefits and optimal design of marine reserves. *Fisheries Bulletin* 97: 604–616.
Sobel J and Dahlgren CP (2004) *Marine Reserves: A Guide to Science, Design and Use*. Washington, DC: Island Press.
Spalding M, Kainuma M, and Collins L (2010) *World Atlas of Mangroves*. London: Earthscan, with International Society for Mangrove Ecosystems, Food and Agriculture Organization of the United Nations, The Nature Conservancy, UNEP World Conservation Monitoring Centre, United Nations Scientific and Cultural Organisation, and United Nations University.
Stewart GB, Kaiser MJ, Côté IM, *et al.* (2009) Temperate marine reserves: Global ecological effects and guidelines for future networks. *Conservation Letters* 2: 243–253.
Sumaila UR (1998) Protected marine reserves as fisheries management tools: A bioeconomic analysis. *Fisheries Research* 37: 287–296.
Suman D, Shivlani M, and Milon JW (1999) Perceptions and attitudes regarding marine reserves: A comparison of stakeholder groups in the Florida Keys National Marine Sanctuary. *Ocean and Coastal Management* 42: 1019–1040.
Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology* 10: 2020–2027.
Toropova C, Meliane I, Laffoley D, Matthews E, and Spalding M. (eds.) (2010) *Global Ocean Protection: Present Status and Future Possibilities*. Brest, France: Agence des aires marines protégées; Gland, Switzerland, Washington, DC and New York, USA: IUCN WCPA; Cambridge, UK: UNEP-WCMC; Arlington, USA: TNC; Tokyo, Japan: UNU; New York, USA: WCS.
Valiela I, Bowen JL, and York JK (2001) Mangrove forests: One of the world's threatened major tropical environments. *BioScience* 51: 807–815.
Waycott M, Duarte CM, Carruthers TJB, *et al.* (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences of the United States of America* 106: 12377–12381.
WCPA/IUCN (2007) Establishing networks of marine protected areas: A guide for developing national and regional capacity for building MPA networks. *Non-Technical Summary Report*. Washington, DC: IUCN-WCPA, National Oceanic and Atmospheric Administration and The Nature Conservancy.
West JM and Salm RV (2003) Resistance and resilience to coral bleaching: Implications for coral reef conservation and management. *Conservation Biology* 17: 956–967.
White AT, Salamanca A, and Courtney CA (2002) Experience with marine protected area planning and management in the Philippines. *Coastal Management* 30: 1–26.
White C, Kendall BE, Gaines S, Siegel DA, and Costello C (2008) Marine reserve effects on fishery profit. *Ecology Letters* 11: 370–379.
Wilkinson C (ed.) (1998) *Status of Coral Reefs of the World: 2004, Vol. 1. Global Coral Reef Monitoring Network*. Townsville, Australia: Australian Institute of Marine Science.
Wilkinson C (ed.) (2004) *Status of Coral Reefs of the World: 2004, Vol. 1. Global Coral Reef Monitoring Network*. Townsville, Australia: Australian Institute of Marine Science.
Williams ID and Polunin NVC (2000) Differences between protected and unprotected Caribbean reefs in attributes preferred by dive tourists. *Environmental Conservation* 27: 382–391.
Wolfenden J, Cram F, and Kirkwood B (1994) Marine reserves in New Zealand: A survey of community reactions. *Ocean and Coastal Management* 25: 31–51.
Wood LJ, Fish L, Laughren J, and Pauly D (2008) Assessing progress towards global marine protection targets: Shortfalls in information and action. *Oryx* 42: 1–12.
Worm B, Barbier EB, Beaumont N, *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.

Marine Sediments

Paul VR Snelgrove, Memorial University of Newfoundland, St. John's, NL, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Abyssal plains Relatively flat areas of the ocean bottom below ~4000 m depth.

Benthos Bottom-living organisms, including those that reside on hard and soft bottom surfaces and others that reside between sediment grains.

Continental rise An area at the base of the continental slope between 3000 and 4000 m where the bottom slope is slight and sediments often accumulate.

Continental shelf A region of ocean bottom extending from the low water mark at the edge of continents to a depth (~200 m) at which the incline increases markedly and the continental slope begins.

Continental slope Ocean bottom extending from the edge of the continental shelf at an ~4° incline to a depth (3000–10,000 m) at which the slope decreases and the continental rise begins.

Deposit feeders Organisms that feed primarily by ingesting organic material occurring on or between sediment grains.

Macrofauna Animals large enough to be retained on a 300- or 500-μm, sieve.

Megafauna Animals those are sufficiently large to be identified from bottom photographs.

Meiofauna Animals small enough to pass through a 500-μm sieve but large enough to be retained on a 44- or 63-μm sieve.

Suspension feeders Organisms that feed primarily on particles of organic material suspended in the water above the bottom.

Benthos and the Benthic Environment

Depending on the specific habitat and species, different benthic (bottom-dwelling) organisms may reside just above the bottom but closely associated with it (hyperbenthos), on the sediment surface (epifauna), or among the sediment grains (infauna). Because sediments have limited permeability, very little oxygenation (penetration of millimeters) is typically achieved through diffusion. Greater oxygen penetration is achieved only when bottom currents mix sediments or, more commonly, when bioturbation (biological mixing of pore water and sediments by benthos) facilitates transfers of oxygenated water from the water column into the sediment pore water. Bioturbated sediments are usually oxygenated within the top few centimeters, although active bioturbation can lead to deeper pockets of penetration. Light does not penetrate well through sediments and most food is provided by phytoplankton detritus sinking from surface waters above, macrodetritus transported from coastal areas (e.g., kelps, seagrasses, and terrestrial plants), or from primary producers living attached to the sediment surface in shallow nearshore areas where light penetrates all the way to the bottom. Thus, most of the primary production or detritus in benthic systems concentrates near the sediment surface rather than within the sediments. This combination of limited food and oxygen penetration confines the vast majority of organisms to the upper few centimeters of sediment near the sediment–water interface. Smaller organisms that can tolerate low (hypoxia) or even no (anoxia) oxygen and larger organisms that maintain a burrow or appendage to the surface can live deeper, but even then distributions are generally limited to 10–20 cm or less (see Gray and Elliot, 2009).

As in other environments, feeding modes include herbivores, predators, scavengers, parasites, and omnivores. However, most organisms that reside in marine sediments are classified either as deposit feeders or as suspension feeders, and some species are capable of either feeding mode. Deposit feeders ingest particles associated with sediments or, in many cases, they ingest the sediment particles themselves and strip off nutrition in the form of detritus associated with the sediment grains and also associated microbes. In shallow areas, benthic algae (diatoms) may provide a food source for some deposit-feeding species. Deposit feeders are important to sediment geochemistry because, as they move through the sediments, they facilitate pore water and sediment grain movement. Suspension feeders remove particles such as phytoplankton and detritus from the water column. Because they rely on suspended particles, suspension feeders tend to be most abundant in energetic environments and absent where weak currents reduce horizontal flux of food particles. Differences in the physical energetics of an environment also influence the stability of the benthic habitat and the organisms that live there. Quiescent areas usually have fine-grained sediments that are relatively stable except where storms pass through. High-energy environments are more dynamic in terms of sediment movements; fine-grained sediments are usually absent and ambient flow conditions often move around the coarse-grained sands that remain. Not surprisingly, shallow coastal environments are the most dynamic, whereas much of the deep sea is relatively quiescent because wave energy rarely penetrates beyond 50–60 m with sufficient strength to move sediment grains.

For reasons of sampling and taxonomic practicality, benthic researchers routinely focus on just one of several size groupings. Megafauna are defined as organisms sufficiently large to be identified in bottom photographs and include flatfish, crabs, and scallops. Macrofauna are defined as those retained on a 300- or 500-μm sieve (standards vary). This size grouping includes polychaete annelids, crustaceans, bivalves, and many other phyla (**Figure 1**). Meiofauna refers to organisms that pass through a 300- or 500-μm sieve but are retained on a 44- or 63-μm sieve and include nematodes, foraminiferans, tiny crustaceans, and many others (**Figure 2**). Microorganisms include living organisms that pass through a 44- or 63-μm sieve

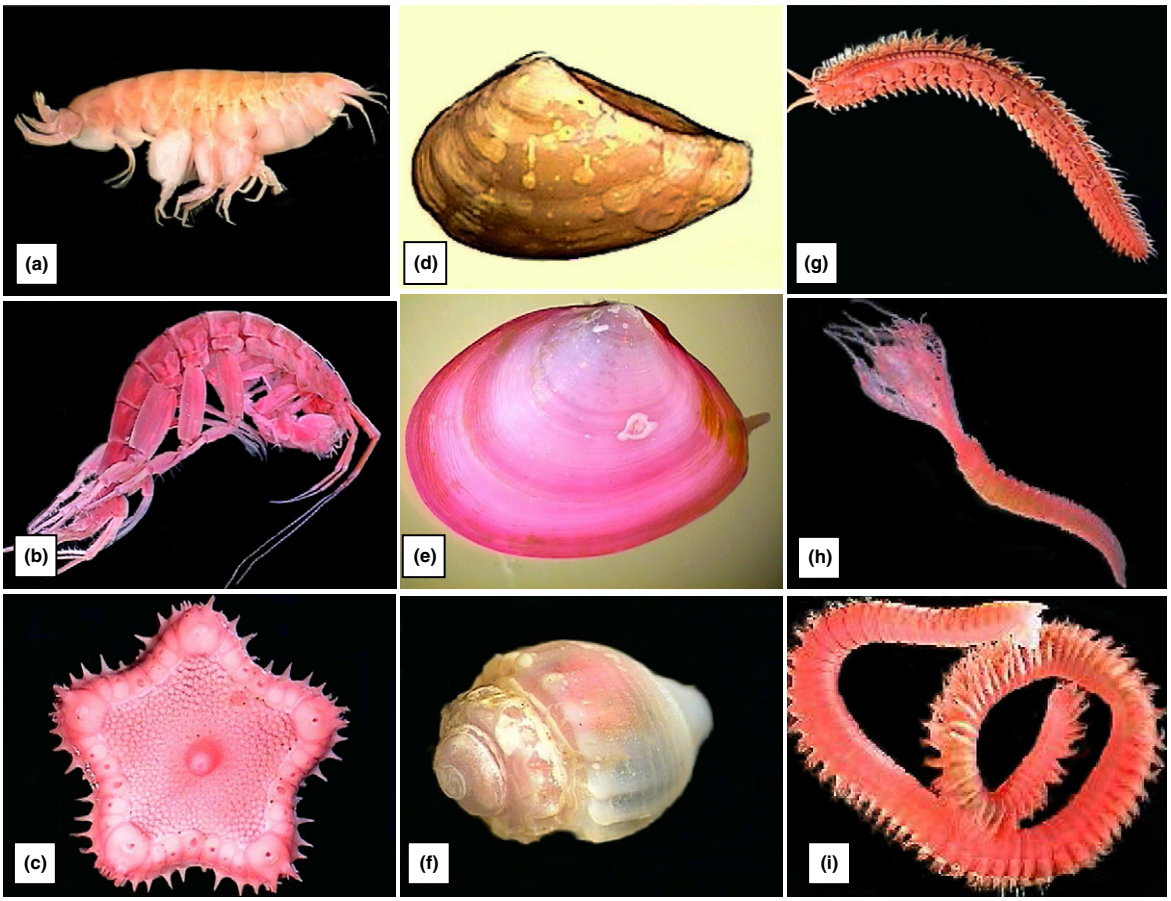


Figure 1 Typical sedimentary macrofauna, including amphipod crustaceans (a and b), echinoderms (c), bivalve (d and e), and gastropod (f) mollusks and annelid polychaetes (g–i). Magnification in the photos varies but individuals range from approximately 1 to 10 mm in length (photographs by P. Ramey).



Figure 2 Typical sedimentary meiofauna, including harpacticoid crustaceans (a), nematodes (b), and foraminiferans (c). Magnification in photos varies but individuals range in size from approximately 100 to 300 μm (photographs by P. Ramey).

and include bacteria and protists. These size groupings are not absolute in that the larval or juvenile stages of one group may be comparable in size with individuals from a smaller group, but this division of organisms is necessary because of sampling logistics. The sample size and equipment appropriate for enumerating megafauna, for example, is inappropriate for sampling meiofauna that are orders of magnitude more abundant and substantially smaller. Moreover, the taxonomic challenges even within each of these size groupings are considerable, and synthesis across groups is therefore rare in a single study.

These groups of organisms do not operate independently of one another. Megafauna prey on infauna, and macrofauna prey on meiofauna. There is also evidence that meiofauna prey on juvenile macrofauna. Microbes form an important dietary component for many meiofaunal and macrofaunal taxa, and macrofaunal bioturbation impacts microbial composition because it oxygenates sediments and thereby regulates the relative importance of aerobic versus anaerobic bacterial activity. Microbes also initiate breakdown of food material that would otherwise be undigestable by metazoans and play a

major role in cycling of organic matter. Thus, there is considerable interaction between these groups of organisms.

Many factors are thought to influence the spatial distribution of benthic sedimentary organisms, including sediment type, productivity, temperature, salinity, pressure, depth, oxygen, sediment stability, air exposure, current speeds and wave action, and biological interaction with other species. A more complete treatment of benthic organisms is given in Gray and Elliott (2009).

Marine Sedimentary Habitats

Oceans range in depth from intertidal habitats to ocean trenches at 10,000 m depth (**Figure 3**). Light can penetrate to 1000 m in the clearest oceanic waters, but in coastal waters penetration may be limited to from 200 m to only a few meters. Like most ecosystems, marine sedimentary communities are fueled by photosynthesis, and because many marine sediments occur below photosynthetic depths they rely on phytoplankton sinking from surface waters above and macrophyte detritus transported from nearshore environments. Not surprisingly, water depth is a major variable in categorizing marine sedimentary habitats.

Intertidal Habitats

The shallowest sedimentary habitats are those that occur at the land–sea interface. They include habitat with emergent vegetation such as mangroves and salt marshes as well as sandflats and mudflats without emergent vegetation that may or may not include diatom mats (**Table 1**). Mangroves and salt marshes occur intertidally, and as such they experience wide ranges of salinity and temperature. They are also extremely productive, with abundant organic matter in the form of decaying vegetation. The productivity of these systems promotes extremely high bacterial respiration and sediments are often hypoxic or even anoxic except at the sediment–water interface. The organic matter produced by vascular plants is also relatively refractory and difficult for most species to digest. In combination with oxygen limitation and tidal exposure, this refractory material contributes to the low diversity of organisms residing in the sediments. Unvegetated sandflats and mudflats tend to be much lower in productivity unless they are immediately adjacent to vegetated areas, and their diversities are slightly elevated. Nonetheless, intertidal areas experience high variation in environmental variables and are characterized by a low-diversity fauna consisting of organisms able to cope with this variability (see Levin *et al.*, 2001).

Subtidal Habitats

At depths below low tide, species diversity increases relative to that of intertidal systems. In the shallowest subtidal habitats, vegetation in the form of seagrass and kelp may occur, and some shallow sediments may be covered in diatom mats. However, all these plant forms are confined to depths of tens of meters or less, and most subtidal sedimentary habitats rely

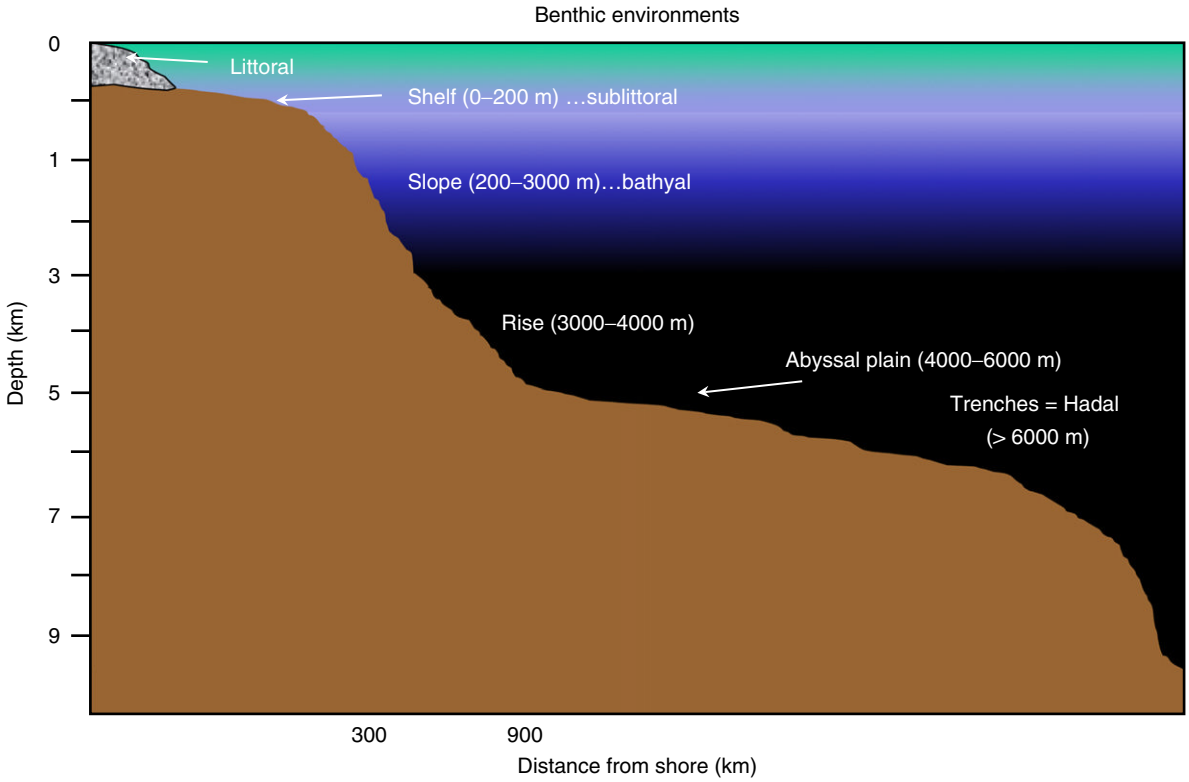


Figure 3 Schematic diagram of the depth zones of the oceans. Dark area indicates the ocean bottom. The distances and depths given are only examples of a range or values observed throughout the world. The horizontal axis is greatly exaggerated in size relative to the vertical axis. See text for more detailed explanation of depth zones.

Table 1 Summary of sedimentary habitats, local diversity, ecosystem services provided, and current threats[a]

Sedimentary habitat	*Diversity*	*Ecosystem services*	*Current threats*
Salt marshes	Low	Nutrient cycling, critical habitat, shoreline stability, filtration, productivity, and pollutant cycling	Habitat destruction, pollution, and global climate change
Mangroves	Low	Nutrient cycling, critical habitat, shoreline stability, filtration, productivity, and pollutant cycling	Habitat destruction, pollution, global climate change, and exotic species
Sandflats	Low to modest	Nutrient cycling, filtration, and productivity	Habitat destruction, pollution, and exotic species
Mudflats	Low to modest	Nutrient cycling, productivity, and pollutant cycling	Habitat destruction, pollution, and exotic species
Estuaries	Low	Nutrient cycling, critical habitat, filtration, productivity, and pollutant cycling	Habitat destruction, pollution, global climate change, and exotic species
Coral reef sediments	Very high	Nutrient cycling and productivity	Habitat destruction, pollution, and fishing
Continental shelf	Modest to very high	Nutrient cycling, critical habitat, filtration, productivity, and pollutant cycling	Habitat destruction, pollution, fishing, and exotic species
Continental slope	Medium to very high	Productivity and nutrient cycling (global)	Habitat destruction and fishing
Abyssal plains	High to very high	Nutrient, element cycling (global)	—
Hydrothermal vent sediments	Low	Nutrient, element cycling (global)	Mineral extraction from adjacent vents
Deep ocean trenches	Low	?	—

[a]The table summarizes generalities for which there are exceptions. In the case of global climate change, the large-scale circulation changes, warming and acidification, predicted in many models would threaten all habitats, and only more immediate impacts are included here.

on sinking plant material as the base of their food chain. Sedimentary environments may range from gravel to fine clays. Fine-grained sediments characterize low-energy environments and coarse-grained sediments characterize many high-energy areas. The latter are often dynamic environments, where bottom flow frequently moves sediments. A wide variety of organisms inhabit the sediments that cover much of the continental shelves. The shallower portions of these environments (up to tens of meters) can experience strong seasonality in productivity and temperature, and salinity may vary considerably where offshore runoff is significant (e.g., estuaries). Within these systems, greater environmental variability and lower diversity go hand in hand. With increased depth, the variability in salinity and temperature is generally reduced, although productivity can still be seasonal. Continental shelf areas are large, productive areas of the oceans and support the majority of marine commercial fisheries.

Deep-Sea Habitats

The deep sea, defined here as habitats beyond the edge of the continental shelf at ~200 m, encompasses continental slopes, continental rises, abyssal plains, midocean ridges, and trenches. Although the sediments that cover the bottom can vary in grain size composition, environmental variability is considerably reduced. Temperatures are largely invariant, salinity does not change, and sediments are rarely moved around other than by bioturbation. Depending on the region, productivity is generally less seasonal than in shallow water, and most deep-sea benthic environments have much less productivity entering the system than their shallow-water counterparts. Because these environments exceed depths of light penetration, productivity consists of whatever material sinks through the thousands of meters of water column to the bottom. As a result of this reduced productivity, numbers of organisms are very low and individuals are relatively small; surprisingly, species diversity is quite high. A very notable exception to low productivity in the deep sea is hydrothermal vents and seeps, for which chemoautotrophic bacteria provide the basis for an extraordinarily productive localized community.

Not all deep-sea communities are diverse. Because hydrothermal vent environments are characterized by fluids rich in compounds such as hydrogen sulfide and heavy metals, most species are unable to tolerate these toxic compounds and species diversity is quite low. Deep-sea trenches are subject to mud slumping and poor circulation. Not surprisingly, species diversity is also low. Upwelling regions are also generally low in diversity, likely because large amounts of organic matter accumulate on the ocean floor and decompose, leading to hypoxia. A few deep-sea areas are subject to intensive "storms," in which currents become intense and resuspend sediments. Evidence suggests that macrofaunal diversity is depressed in such areas, but surprisingly meiofaunal diversity is not. Presumably, the meiofauna are able to cope with the disturbance more effectively than are the macrofauna.

Global Estimates of Species Numbers

Several attempts have been made to estimate the total number of species that live in the ocean. The Danish marine scientist Thorson estimated in the early 1970s that 100,000 species would eventually be described from the marine environment, and until recently this was the ballpark figure most thought appropriate. In the late 1960s, deep-sea biologists Sanders and Hessler documented that the deep sea was not the low-diversity habitat that many had previously thought and could contain a significant portion of global marine species diversity (Hessler and Sanders, 1967; Sanders and Hessler, 1969). The debate on species estimates heated up considerably when

Grassle and Maciolek (1992) published data from a collection of samples along a 176-m-long depth contour along the east coast of the USA, constituting the most extensive quantitative data set ever assembled for a single area of the deep sea. Based on the rate that species were added with increased area sampled, and the total area of the deep sea, they extrapolated to 10 million macrofaunal species. May (1992) pointed out that approximately half of the species in their study were new to science, and one could extrapolate from known to unknown species using a similar ratio. Given that approximately 250,000 marine species have already been described, May's (1992) extrapolation suggests ~500,000 total species living in the oceans. Other work from the Pacific (Poore and Wilson, 1993) suggests that only one in 20 species has been described for that area of the deep sea, in which case May's approach would yield ~5 million species.

Although most of the discussions on species number have centered on macrofauna, the lesser known meiofauna and especially the microbes represent a potentially massive pool of undocumented diversity. But because so little is known about how widely distributed individual species may be, extrapolating numbers is even more tenuous and controversial than for macrofauna (Lambshead, 1993; Lambshead and Boucher, 2003).

Clearly, more data are needed to improve estimates of global species numbers. Current estimates for metazoans range from 500,000 to 100 million species, but vast areas of the ocean and some taxonomic groups remain unsampled or poorly sampled. The South Pacific and tropical latitudes are among the least known areas. Even many shallow environments remain poorly known, and recent evidence suggests that some of these areas may be very species rich. For example, a United States National Research Council (1992) study summarized surveys that found two-thirds of the polychaetes in Hawaiian coral reef sediments were previously undescribed; even Georges Bank, a relatively well-studied area, was found to have only two-thirds of its polychaete species previously described. The situation is even more severe in deep-sea areas and for meiofaunal taxa such as nematodes and copepods, for which proportions of known species can sometimes be less than 5–20% of total species present (Ebbe *et al.*, 2010). Even for taxa thought to be well known, there is increasing evidence based on genetic data that 'cosmopolitan taxa' are often species complexes that are difficult to distinguish based on morphological characters. Regardless of which projection of total species number is appropriate, there remains a huge amount of undescribed species diversity in marine sediments. New approaches such as genetic barcoding should expedite discovery, but the numbers of scientists working on the problem are few relative to the size of the problem.

Global Patterns of Biodiversity

The historical perception of the deep sea as an azoic or species-poor environment changed markedly with a series of papers (Hessler and Sanders, 1967; Sanders and Hessler, 1969) that report on data collected along a transect running from Martha's Vineyard, Massachusetts, to Bermuda, and demonstrated that diversity in deep-sea sediments exceeded that in most shallow areas and rivaled that observed in shallow tropical areas. These findings contradicted a previous notion that diversity would generally decrease with depth. Rex (1983) followed up this work with extensive analysis of large-scale pattern with depth in the North Atlantic and found that diversity is highest at intermediate depths. Diversity in shallow water is low in estuaries and seasonal shallow areas, is somewhat higher on the continental shelf, increases down the continental slope, and peaks on the lower slope. At greater depths it declines again. Other workers have observed a similar pattern for other taxa, although the exact depth of the diversity peak varies among taxa and also among locations (Menot *et al.*, 2010). Moreover, diversity in the abyssal plains of the Pacific appears to exceed that observed at shallower depths, suggesting that a parabolic diversity–depth pattern is not universal. The generality of lower diversity in shallow water was questioned by Gray *et al.* (1997), who presented convincing data that the local diversity of sedimentary fauna of Bass Strait, Australia, (11–51 m deep) may rival or exceed that of the deep sea. The pattern observed by Rex and others has been based largely on data from the North Atlantic, and Gray's findings emphasize the fact that patterns described from one area may not be easily extrapolated globally.

Many studies have documented latitudinal gradients in shallow-water benthic diversity, with decreasing diversity toward the poles. In the deep sea, the existence of such a trend is the subject of debate. The pattern has been proposed for North Atlantic deep-sea macrofauna (Rex *et al.*, 1993), but isopod data from the South Pacific (Poore and Wilson, 1993, 1992) do not indicate such a pattern and nor do some nematode data for the North Atlantic (Lambshead *et al.*, 2003).

The inconsistency of conclusions regarding global patterns in diversity indicates a need for more complete sampling coverage with latitude and depth. Improved coverage would allow more effective comparison of different oceans and reduce the impact of 'noise' added by regional differences in variables such as productivity that are bound to obscure any global patterns. Within this coverage, areas that have been poorly sampled historically, such as tropical sediments and the Southern Hemisphere, must be emphasized to reduce the North Atlantic bias. An additional problem with these large-scale comparisons is that most studies focus on just one of the size groupings of organisms and generally use different sorts of sampling techniques. For example, one must ask whether differences described in large-scale patterns in different taxa and ocean basins reflect differences in biology or sampling coverage. Studies that are more comprehensive in taxonomic and geographic coverage are necessary to resolve these debates.

Regulation of Diversity

Shallow Water

Our understanding of diversity regulation in marine systems is limited. In environments with extremes of temperature (e.g., intertidal sediments), salinity (e.g., estuaries), geochemistry

(e.g., sediments near hydrothermal venting), or oxygen (e.g., eutrophied coastline, mangrove sediments, and fjords with poor circulation), there are obvious physiological limitations for many species that allow only the most robust species to exist. The same may be said for physically disturbed environments such as low-diversity sandflats. Similarly, the reduced diversity in highly seasonal environments presumably results from relatively few organisms able to cope with the variability in the physical environment. However, beyond these extreme habitats, obvious conclusions on diversity regulation are few.

Historical events and modern conditions both play major roles in determining global diversity patterns. Certainly, the evolutionary history of a region must be considered. Jablonski's (1993) analysis of fossil invertebrates suggests that the tropics provide evolutionary novelty, and thus contribute to the higher diversity noted in tropical latitudes. Glaciation effects played a well-known role in establishing current marine patterns. For example, the reduced diversity observed in Arctic relative to Antarctic fauna (a pattern questioned in some recent studies – see Gradinger *et al.*, 2010) likely relates to extinctions during Arctic glaciation and the slow reestablishment of the Arctic community. Thus, historical events set the stage for current ecological processes, which may subsequently play a role. In terms of modern conditions, the high level of total energy input in the tropics has been proposed to explain shallow-water mollusk diversity gradients, and the high degree of seasonality with increasing latitude may be a factor that depresses diversity.

The role that ecological interactions play in sedimentary diversity is less well understood. In shallow-water experiments, the exclusion of predators seems to enhance diversity (see Peterson, 1978). This finding is the direct opposite of Paine's (1966) classic findings in rocky intertidal communities and suggests that it is inappropriate to generalize from relatively well-studied rocky intertidal systems to the less understood sedimentary environment. Whether generalizations regarding predators depressing diversity can be extended to the deep sea remains to be tested. There is a dearth of experiments linking diversity in the water column to sedimentary faunal diversity and little data on how predator diversity might impact benthic diversity. One might predict, for example, that increased diversity in food sources (e.g., phytoplankton) might lead to enhanced diversity in the infauna that feed on those food sources. Increased diversity of predators might also allow tighter species packing in an environment.

Deep Sea

There has been considerable debate on how the deep-sea environment, which appears so physically homogeneous, is able to support such a broad array of species. Sanders, who played a pivotal role in recognizing the high diversity of marine systems, suggested that the high level of stability over evolutionary time in the deep sea has allowed greater specialization and niche diversification. If this were the full explanation, then diversity should be highest on abyssal plains; the depth patterns described in the previous section on patterns do not consistently support this notion though newer data suggest abyssal plains may be more diverse than previously thought. There is also little evidence of niche specialization in deep-sea organisms. Most species appear to be relatively nonselective deposit feeders. In the early 1970s, several contrasting theories were put forward. First, cropping by predators could prevent competitive equilibria from being attained by infauna. If this were the case, however, then fast growth, early reproductive maturity, and an age structure dominated by young individuals might be expected, but this is not a pattern typical for the deep sea. Moreover, evidence from shallow water suggests that predators decrease sedimentary diversity. In any case, the role that predators may play in maintaining deep-sea diversity remains largely unknown. A contrasting theory held that small patches may form disequilibria habitats that promote certain species. Tests of this patch mosaic model have included sampling of natural patches and creation of experimental patches. Both approaches have yielded a similar result: The species that occur in many patch types are usually rare or absent from background sediments, although diversity in patches is usually reduced. Such a pattern is consistent with the patch mosaic model, but experiments so far have demonstrated that patches promote only a modest number of species. Whether it is necessary to sample more patch types, or whether other explanation needs to be invoked, remains unclear.

The large area of the deep sea almost certainly impacts its species richness in that species richness of most environments increases with area. However, the vast rolling plains of the deep sea add less habitat heterogeneity than other species-rich habitats, such as tropical rain forests or coral reefs. Given that abyssal plains are the largest deep-sea habitat in area, the parabolic diversity–depth relationship described earlier is also inconsistent with a simple species–area relationship.

Huston's (1979) intermediate disturbance hypothesis is another explanation for the high diversity of deep-sea systems. The idea is that, at one end of the spectrum, high levels of disturbance depress diversity by eliminating sensitive species. At the other end of the spectrum, very benign habitats allow superior competitors to eliminate weaker species and thus depress diversity. Highest diversity would then be expected in habitats with levels of disturbance that prevent competitive dominants from taking over but not so frequent and harsh that few species can cope. The midslope peak in diversity described earlier (*see* Global Patterns) is consistent with this hypothesis but does not offer an explicit test of the hypothesis. Most disturbed deep-sea environments, such as hydrothermal vents, low-oxygen areas, environments with benthic storms, and slumping areas, have reduced diversity. However, such extreme examples are far from conclusive in drawing comprehensive generalities.

Efforts to understand regulation of biodiversity in the deep sea have been hampered by the considerable logistical challenges of deploying and recovering experiments in deep, oceanic environments. In shallow-water environments, diversity regulation has received considerably less attention, and there is a pressing need to move some of the deep-sea framework and experimental effort into shallow water where logistics are simplified and outcomes are bound to produce interesting comparisons with the deep sea. Our understanding of diversity maintenance in sedimentary systems is generally

very poor, and we are unable to predict, for example, whether loss of pelagic diversity will impact benthic diversity. The problem is that humans impact diversity of marine benthic organisms in many areas through fishing, habitat destruction and modification, pollution, exotic species introductions, and global climate change, but there is often little idea on how different components of the ecosystem will be affected.

Ecosystem Services and Sedimentary Diversity

Although the huge fraction of the Earth's surface covered by marine sediments and the large numbers of species residing within sediments provide strong motivation to understand the pattern and regulation of benthic diversity, there are also good ecological motivators. The oceans provide a variety of ecosystem services (**Figure 4**), and although we understand little about the role that biodiversity plays in maintaining these services there is very clear evidence that benthic species play critical roles. There is a very real chance that ongoing changes in sedimentary biodiversity will result in loss of ecosystem services.

Nutrient Cycling

Benthic marine organisms play a critical role in global cycles of nitrogen, sulfur, and carbon (Giblin *et al.*, 1995; Snelgrove *et al.*, 1997). As organic material sinks to the sediment surface, it carries with it organic carbon and nitrogen. This material may be directly ingested by benthic organisms or microbes may colonize it. In areas with high sedimentation rates or low densities of organisms, there may be substantial burial loss of this material, but most is consumed or decomposed. Even the organic material that passes through organisms undigested will be colonized and broken down by microbes. Organisms that ingest the material respire, are preyed on, or die and decompose in the sediment. Depending on how decomposition and digestion occur, nitrogen may be released as ammonia, nitrite, nitrate, or nitrogen gas that diffuses or is physically mixed into the water column above. In coastal sediments, this cycle is a critical part of regenerating the dissolved nitrogenous compounds that are essential for primary producers; without this regeneration the cycle would end and primary production would cease. Carbon is cycled in a similar way. Material sinks to the bottom and may eventually be buried (e.g., fossil fuels) or it may be decomposed into carbon dioxide or other carbon compounds. The relative rates at which carbon is buried, tied up in living organisms, or respired to carbon dioxide are closely tied to microbial activity and thus to the animals that influence microbial activity through feeding and oxidation of sediments. As a result, benthic organisms play a key role in global carbon budgets. Sulfur cycling through marine sediments is also dependent on sediment oxygenation and therefore bacterial activity. Whether sulfur is buried and stored in sediments or recycled is determined in part by metazoan and microbial activity.

Pollutant Cycling

Because benthic organisms move around and ingest particles, and may themselves be ingested by predators, they can greatly impact the burial fate and mobility of pollutants. As sediment particles are bioturbated, so are any pollutants linked to them. As a result, benthic organisms can dilute pollutants at the sediment–water interface by mixing them downward. By the same process, the continual remixing of sediments by infauna

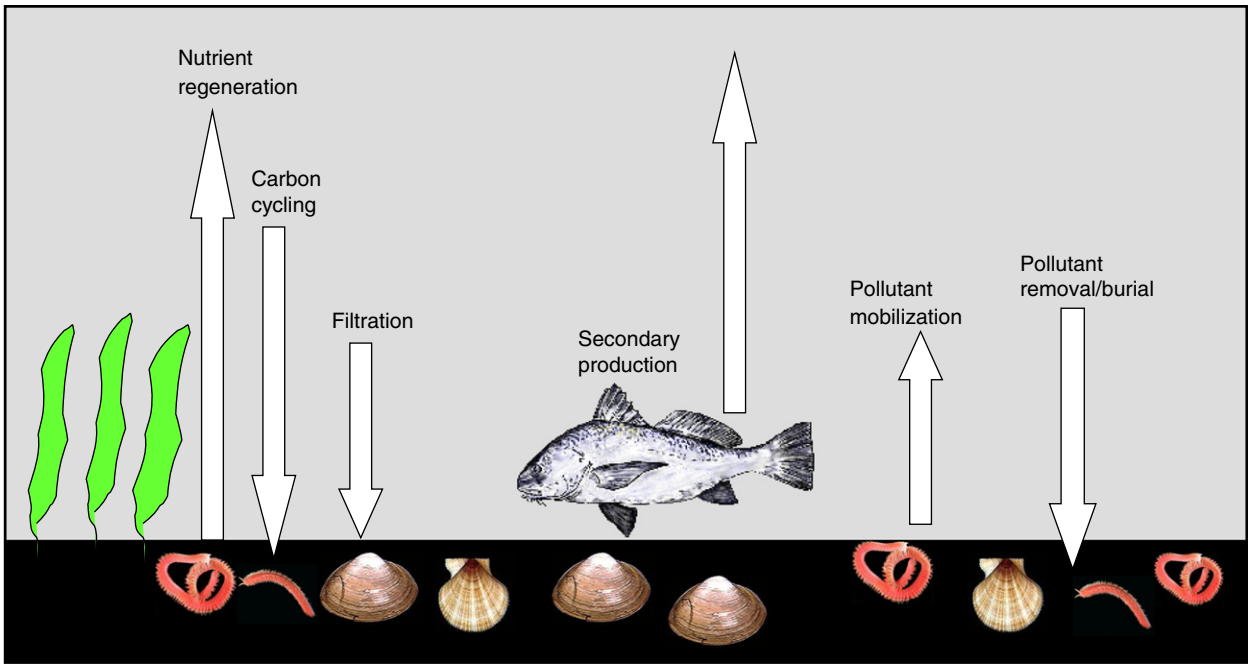


Figure 4 Schematic representation of key ecosystem services associated with sedimentary organisms. Arrows indicate transfer of material between the water column of the ocean and the sediment.

may increase burial time of pollutants that might otherwise be buried more quickly by sedimentation. Bioturbation also tends to destabilize sediments and increase the likelihood that they will be resuspended. Some sedimentary organisms can also stabilize sediments by excreting mucous (e.g., bacterial mats) or creating tubes that they inhabit. Benthos may also increase effective grain size by repackaging grains as fecal pellets. All these processes impact whether sediments are likely to be resuspended or not. Infauna that ingest pollutants may also provide a conduit up the food chain if they concentrate the pollutant in their tissues. For example, fish that feed on polychaetes with high levels of a heavy metal may also have high concentrations of that metal. Finally, benthic organisms have some capacity to break down pollutants. Microbes in particular may have the capacity to break down some toxic compounds and render them harmless.

Sediment and Shoreline Stability

Because benthos can destabilize (bioturbation) and stabilize (microbial mats, mucous excretion, and tube building) sediments, they can have a major impact on sediment transport and coastal geology. Mucous excretion, for example, binds sediment grains together and therefore increases the amount of energy required to move sediments. However, as some organisms produce products that bind sediment grains together, other organisms move among the sediment grains, breaking apart binding structures, increasing the water content, and increasing the likelihood of resuspension.

The vegetation in seagrass beds, mangroves, and salt marshes acts as flow baffles and therefore traps fine sediments, a process that is further aided by root structures that add further stability. The presence of these plants protects the shoreline and results in accretion rather than erosion. Indeed, rapid shoreline growth has been documented in areas with mangroves and marshes, where land is extended seaward on scales of kilometers more than timescales of hundreds to thousands of years.

Filtration

The benthos can also impact water clarity in coastal areas. The sediment-trapping capacity of emergent vegetation reduces sedimentation to the coastal waters beyond the marsh, mangrove, or seagrass bed. In tropical areas, for example, mangroves and seagrasses can trap sediments that might otherwise be transported onto coral reefs, where the suspended sediment would reduce water clarity and productivity and potentially smother corals. In temperate latitudes, marshes are well known not only for their ability to trap sediments but also for their capacity to take up excess nutrients. Agricultural runoff to the coastal environment can result in eutrophication, but the presence of salt marshes can significantly offset this problem and provide a natural water filtration system. Mangroves have a similar capacity.

Suspension-feeding benthos can also impact water clarity by filtering out suspended particles. The reduction of oysters in estuaries along the east coast of the US through overfishing, destruction of oyster reef habitat, and disease has resulted in a marked decrease in water clarity. Impacts of this magnitude are generally confined to shallow nearshore habitats with very high densities of suspension feeders, but the effect is nonetheless important.

Secondary Production

Benthic megafauna and macrofauna provide an important source of secondary production. Some of this production, such as clams, scallops, shrimp, and crab, is directly consumed by humans and forms the basis of important commercial fisheries. However, even groups such as the polychaetes, which have no obvious economic importance, form an important food source for benthic-feeding fishes and crustaceans. Because sinking organic material collects on the bottom, the secondary production in the benthos can often be considerably higher than that in the water column above.

Linking Biodiversity and Ecosystem Services

Although benthic organisms inarguably provide key ecosystem services (Snelgrove, 1998), there is little evidence that species diversity *per se* is important in maintaining these services. This deficiency does not necessarily mean that species diversity is unimportant but rather that it has not been studied. In virtually all the categories listed, there is a particular functional group that is critical to the service, but in most cases there are more than one species of a particular functional group in a given habitat. For example, deposit feeders are the major bioturbators in marine sediments, and bioturbation has major impacts on nutrient cycling, pollutant burial and mobilization, and sediment stability. However, most sediments that contain deposit-feeding organisms will contain multiple species. At least superficially, it appears that different deposit feeders often do the same sorts of things, although there are species that feed on particular grain sizes and at particular depth horizons in the sediment. What is not known is whether those species that remain after others are eliminated can provide the same ecosystem services as the more diverse community. Will the removal of a given species result in fewer total numbers of deposit feeders or will other species increase in number or activity to compensate? Elmgren and Hill (1997) provide evidence from the Baltic Sea that complete removal of a functional group (in this case suspension feeders) will indeed fundamentally impact trophic linkages. Similarly Danovaro *et al*. (2008) found a positive relationship between diversity of different functional groups of meiofaunal nematodes and carbon production in deep-sea sediments. Further experiments to test these ideas and that span a wide range of habitats are sorely needed. It is clear that changes in functional groups will alter the way in which ecosystem services are carried out, but any linkage between regional or local biodiversity and ecosystem services in marine systems remains to be demonstrated.

In some instances, there are likely to be keystone species that provide services disproportionate to their abundance. Clearly, the loss of these species will have a major impact. In other instances, there may be only a single species providing a service so that loss of that species will have an easily predicted

outcome. Seagrasses and mangroves, for example, are often composed of only one or two species of emergent plant in a given area, and the loss of that species will have a clear effect. The loss of seagrasses to disease in Europe earlier this century had very clear effects on coastal sediments and shoreline erosion. Similarly, the loss of oysters (*see* Ecosystem Services and Sedimentary Diversity) had an obvious impact. Unfortunately, we are not always able to predict keystone species and we are even less able to predict the cumulative impacts of multiple species that belong to a given functional group.

Threats to Sedimentary Diversity

Documented extinctions in the marine environment are relatively few, and many of these are birds or mammals. However, there are certainly biodiversity losses occurring in the sedimentary fauna. Particularly in coastal areas, where human impacts have been most severe, large areas of habitat are being damaged or lost (Norse, 1993; National Research Council, 1995). At the very least, genetic diversity is disappearing with these habitats. Species that live in coastal habitats are often adapted to local conditions and may therefore have a specific tolerance to variables such as salinity and temperature. Once this gene pool has been lost, there is no guarantee that conspecifics from other areas will be able to invade that specialized environment. There is also increasing evidence that a significant portion of sedimentary diversity is undescribed and may subsequently be lost without our realizing it even existed. With these losses, changes in ecosystem services could also be missed or misinterpreted.

It is possible to generalize that with increasing proximity to the shoreline, and therefore to human populations, the number and magnitude of threats to sedimentary diversity increase (**Figure 5**). Thus, salt marsh and mangrove habitats are being reduced at a rapid rate, continental shelf habitat is being damaged in many although not all areas, and deep-sea habitats are currently being impacted in mostly isolated instances. The types of impacts are far from uniform between systems.

Fishing Impacts

Most of the major marine fisheries in the world occur on continental shelf areas, and they can impact sedimentary habitats in several different ways (see review of Dayton *et al.*, 1995). Among the most destructive forms of fishing are dredges and trawls that are dragged across the bottom to collect benthic fishes and invertebrates (Watling and Norse, 1998). These types of fishing gear are often heavy and specifically designed to dig into sediments, and in doing so they alter the sediment composition and geochemistry, damage infaunal organisms, and destroy any benthic epifauna and plants and the habitat they create for other species, including some commercial fish. Many fisheries are also notoriously nonselective, and subsequently remove (as by-catch) large numbers of species other than those intended. By-catch, which is often 50–80% of the total catch, is often thrown overboard; most organisms discarded in this manner will be unable to recover, and pockets of decaying organisms then create localized areas that are similar to eutrophied habitat.

Fisheries also tend to remove top predators, and because these are large and abundant taxa they are species that have a major impact on their ecosystem. There are well-documented cases of shifts in species composition and trophic mode that occur as dominant predators are fished out (Botsford *et al.*, 1979).

In general, fisheries tend to reduce benthic diversity and homogenize habitat, although the specific means by which diversity loss occurs is only well documented in the case of

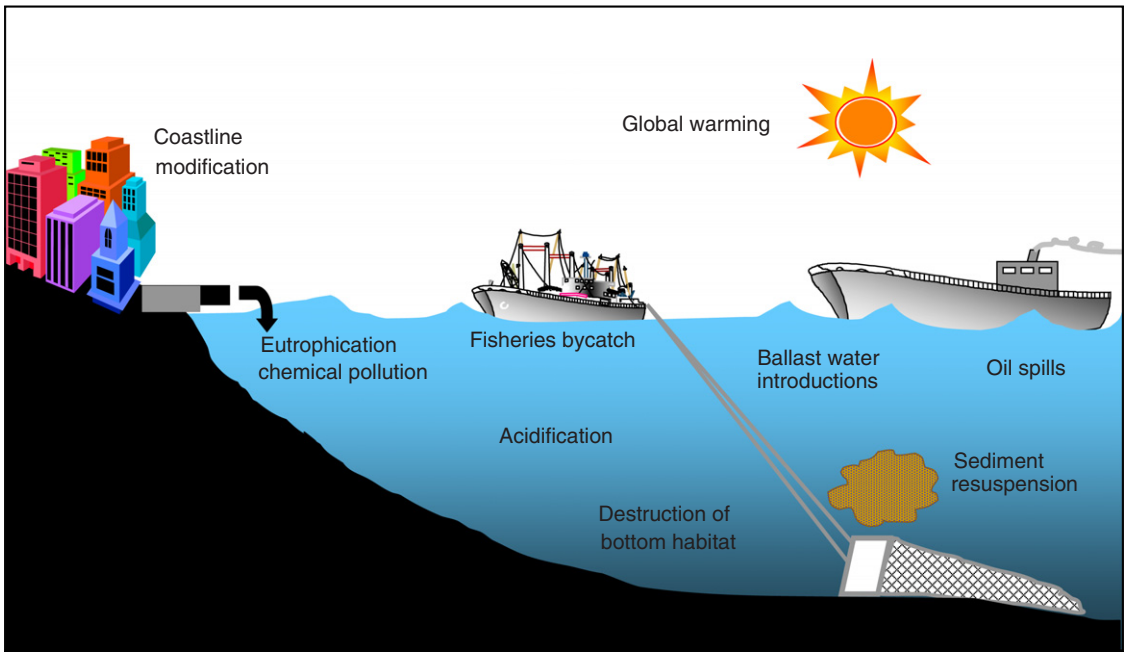


Figure 5 Schematic representation of major threats to sedimentary organisms and the services they provide.

physical destruction of habitat. Indirect effects, such as predator removal, require additional study.

Habitat Destruction, Degradation, and Shoreline Modification

Fishing activities are certainly the greatest source of degradation of bottom habitat beyond the immediate coastline, but coastal development has resulted in considerable reclamation of intertidal habitat ranging from salt marshes to mangroves and mudflats. A significant portion of The Netherlands and Boston's Logan Airport are just two examples in which shoreline has been engineered by filling or erecting dykes to allow development at the land–sea interface. Globally, large areas of marsh habitat have been modified to allow building, and large areas of tropical mangroves have been eliminated to allow development of shrimp aquaculture farms. Clearly, the loss of large areas of habitat changes local species composition and likely reduces genetic diversity of individual species. Extinctions are likely occurring in some areas.

Pollution and Eutrophication

The large human populations and industries that occur along coastlines throughout much of the world create problems of eutrophication, pollution, and habitat degradation (Pearson and Rosenberg, 1978). Eutrophication occurs when excess nutrients are supplied to coastal water from agricultural runoff and sewage. Phytoplankton and shallow vegetation respond to increased nutrients by growing, and as they die and sink to the bottom microbial decomposition occurs. Where large amounts of organic matter are available, microbial respiration exceeds photosynthetic production of oxygen and sediments, and bottom waters may become hypoxic or anoxic.

Pollutants are also problematic in coastal environments. Industrial and storm drain runoff deliver heavy metals, hydrocarbons, synthetic compounds (polychlorinated biphenyls (PCBs), dichlorodiphenyltrichloroethane (DDT), dioxin, etc.) and countless other toxic compounds to sedimentary communities. Some of these compounds result in increased respiration, reduced reproductive output, and countless other physiological responses that may ultimately lead to death. Alternatively, the physiological stress caused by contaminants may weaken an organism and increase its susceptibility to disease. Transfer of contaminants may occur by diffusion, ingestion of contaminated particles, or feeding on contaminated organisms.

Problems of eutrophication and pollution are most intense in industrialized and densely populated coastal areas where circulation with the open ocean is limited. Thus, partially enclosed estuaries and harbors provide no mechanism for rapid dilution and removal of contaminants. Not surprisingly, the deep ocean is largely buffered from most of these impacts by the size of the continental shelf, given that most contaminants arrive as runoff of some sort from land or rivers. Nitrogenous compounds may also be delivered over broader distances by atmospheric deposition. For the coastal areas that are impacted by eutrophication and pollution, many benthic species are unable to survive under these conditions, and bottom habitats become low in diversity and dominated by weedy species that thrive under such conditions.

Exotic Species Introductions

Within recent years, there has been increasing concern regarding the impact that introduction of nonnative (exotic) species can have on benthic communities (Carlton, 1985). Although many people assume that the oceans are open systems and lack barriers to dispersal, there are actually many discontinuities in temperature, land barriers, depth, and ocean currents that limit the dispersal of reproductive propagules from different areas of the world. Historically, ships have provided a mechanism by which exotic species are transported to nonnative habitat. In the case of hard-substrate fauna, fouling of ship hulls allows adults to cross deep ocean basins that they could not otherwise cross because of a lack of suitable shallow habitat. However, for sedimentary invaders, ship ballast water is the main culprit. A tanker may take on large volumes of ballast water in one port, sail to a new port, and dump the ballast water before loading goods. The dumped ballast water may contain adults and reproductive propagules that have been carried through unfavorable open-ocean habitat.

In recent years, the increasing interest in developing marine aquaculture has also helped spread nonnative species. Nonnative species are transported into new environments in which adults or reproductive propagules can escape into the wild, invade nonnative habitat, and displace native forms. An additional problem is that parasites and pathogens can be transported along with species imported for aquaculture, creating additional problems. Another transfer mechanism that is less common, but embarrassing, is the transport of organisms throughout the world for scientific study; although scientists are usually conscientious about preventing individuals or reproductive propagules from escaping into the natural environment, there are instances of accidental release.

In many instances, nonnative individuals will not successfully colonize the new environment, but particularly when conditions are stressful for native organisms (e.g., pollution and unusually low salinity or high temperatures), the exotic species may not only invade but also displace native organisms. When displacement occurs, even after conditions return to normal, the invasive species may remain the dominant taxon. San Francisco Bay and Long Island Sound are two good examples in which it is well documented that many of the dominant species today are nonnative. In some instances, the exotic and native species may be able to coexist, resulting in slightly higher biodiversity, but the complexity of species interactions can sometimes result in elimination of native species. In other words, the introduction of nonnative species is a very dangerous and unpredictable disturbance whose potential impacts are only beginning to be understood.

Global Climate Change

Perhaps the least easily predicted impact of human activity on sedimentary biodiversity is the long-term effects of global climate change (Smith *et al.*, 2000). In a simplistic sense,

species distributions might be expected to shift toward the poles to compensate for warming of the oceans. However, community response is likely to be far more complicated. Because the habitat of a given species is more complex than just temperature, it will often not be possible to simply shift distributions because suitable sediment and geomorphic features (e.g., productive banks) may be unavailable. As sea level rises, intertidal communities will not simply encroach landward because in populated areas such encroachment will be prevented to preserve expensive coastal development. Moreover, models of global warming predict potential circulation changes in the oceans, and if this indeed happens then a whole suite of ecological processes ranging from primary productivity to dispersal of reproductive propagules will be impacted. It is very likely that coastal habitats and species will be lost if significant ocean warming occurs, and it is possible that deep areas will also be impacted if these changes alter large-scale circulation patterns.

Summary

Although it is likely that ocean sediments contain in excess of 500,000 species, little is known of how this diversity is maintained and how loss of biodiversity may impact the ecosystem services that benthos provide. It is critical that benthic ecologists working in sedimentary habitats ride the current wave of interest in biodiversity to improve on these deficiencies in our knowledge. Marine sedimentary systems are widespread, rich in biodiversity, and ecologically important in terms of the processes that occur within them. It is therefore reasonable to expect exciting discoveries regarding biodiversity in these habitats in the coming decade.

See also: Coastal Beach Ecosystems. Estuarine Ecosystems. Eutrophication and Oligotrophication. Intertidal Ecosystems. Marine Ecosystems. Resource Exploitation, Fisheries

References

Botsford LW, Castilla JC, and Peterson CH (1997) The management of fisheries and marine ecosystems. *Science* 277: 509–514.

Carlton J (1985) Transoceanic and interoceanic dispersal of coastal marine organisms: The biology of ballast water. *Oceanography and Marine Biology: An Annual Review* 23: 313–371.

Danovaro R, Gambi C, Dell'Anno A, *et al.* (2008) Exponential decline of deep-sea ecosystem functioning linked to benthic biodiversity loss. *Current Biology* 18: 1–8.

Dayton PK, Thrush SF, Agardy MT, and Hofman RJ (1995) Environmental effects of marine fishing. *Aquatic Conservation: Marine and Freshwater Ecosystems* 5: 205–232

Ebbe B, Billett DSM, Brandt A, *et al.* (2010) Diversity of abyssal marine life. In: McIntyre AD (ed.) *Life in the World's Oceans: Diversity, Distribution and Abundance*. Oxford: Wiley-Blackwell.

Elmgren R and Hill C (1997) Ecosystem function at low biodiversity—The Baltic example. In: Ormond RFG, Gage JD, and Angel MV (eds.) *Marine Biodiversity: Patterns and Processes*, pp. 319–336. Cambridge: Cambridge University Press.

Giblin AE, Foreman KH, and Banta GT (1995) Biogeochemical processes and marine benthic community structure. In: Jones CG and Lawton JH (eds.) *Linking Species and Ecosystems*, pp. 29–36. New York: Chapman and Hall.

Gradinger R, Bluhm BA, Hopcroft RR, *et al.* (2010) Marine life in the Arctic. In: McIntyre A (ed.) *Life in the World's Ocean: Diversity, Distribution, and Abundance*, pp. 183–202. Oxford: Wiley-Blackwell.

Grassle JF and Maciolek NJ (1992) Deep-sea species richness: Regional and local diversity estimates from quantitative bottom samples. *American Naturalist* 139: 313–341.

Gray J, Poore G, Ugland K, Wilson R, Olsgard F, and Johannessen Ø. (1997) Coastal and deep-sea benthic diversities compared. *Marine Ecology Progress Series* 159: 97–103.

Gray JS and Elliot M (2009) *Ecology of Marine Sediments*, 2nd edn. Oxford: Oxford University Press.

Hessler RR and Sanders HL (1967) Faunal diversity in the deep sea. *Deep-Sea Research* 14: 65–78.

Huston M (1979) General hypothesis of species-diversity. *American Naturalist* 113: 81–101.

Jablonski D (1993) The tropics as a source of evolutionary novelty through geological time. *Nature* 364: 142–144.

Lambshead PJD (1993) Recent developments in marine benthic biodiversity research. *Oceanis* 19: 5–24.

Lambshead PJD and Boucher G (2003) Marine nematode deep-sea biodiversity - hyperdiverse or hype? *Journal of Biogeography* 30: 475–485.

Lambshead PJD, Tietjen J, Ferrero T, and Jensen P (2000) Latitudinal diversity gradients in the deep-sea with special reference to North Atlantic nematodes. *Marine Ecology Progress Series* 194: 159–167.

Levin LA, Boesch D, Covich A, *et al.* (2001) The role of biodiversity in the function of coastal transition zones. *Ecosystems* 4: 430–451.

May R (1992) Bottoms up for the oceans. *Nature* 357: 278–279.

Menot L, Sibuet M, Carney RS, *et al.* (2010) New Perceptions of Continental Margin Biodiversity. In: McIntyre AD (ed.) *Life in the World's Oceans: Diversity, Distribution, and Abundance*, pp. 79–102. Oxford: Wiley-Blackwell.

National Research Council (1995) *Understanding Marine Biodiversity*. Washington, DC: National Academy Press.

Norse E (ed.) (1993) *Global Marine Biodiversity Strategy: Building Conservation into Decision Making*. Washington, DC: Center for Marine Conservation.

Pearson TH and Rosenberg R (1978) Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology Annual Reviews* 16: 229–311.

Peterson CH (1978) Predation, competitive exclusion, and diversity in the soft-sediment benthic communities of estuaries and lagoons. In: Livingston RJ (ed.) *Ecological Processes in Coastal and Marine Ecosystems*, pp. 223–264. New York: Plenum.

Poore GBC and Wilson GDF (1993) Marine species richness. *Nature* 361: 597–598.

Rex MA (1983) Geographic patterns of species diversity in the deep-sea benthos. In: Rowe GT (ed.) *The Sea*, pp. 453–472. New York: Wiley.

Rex MA, Stuart CT, Hessler RR, Allen JA, Sanders HL, and Wilson GDF (1993) Global-scale latitudinal patterns of species diversity in the deep-sea benthos. *Nature* 365: 636–639.

Sanders HL and Hessler RR (1969) Ecology of the deep-sea benthos. *Science* 163: 1419–1424.

Smith CR, Austen MC, Boucher G, *et al.* (2000) Global change and biodiversity linkages across the sediment-water interface. *BioScience* 50: 1108–1120.

Snelgrove PVR (1998) Getting to the bottom of marine biodiversity: Sedimentary habitats. *BioScience* 49: 128–138.

Snelgrove PVR, Blackburn TH, Hutchings PA, *et al.* (1997) The importance of marine sediment biodiversity in ecosystem processes. *Ambio* 26: 578–583.

Watling L and Norse E (1998) Physical disturbance of the seabed by mobile fishing gear: A comparison to forest clear-cutting. *Conservative Biology* 12: 1180–1197.

Marine Symbioses: Metazoans and Microbes

Elisha M Wood-Charlson, University of Hawaii at Manoa, Honolulu, HI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Clade Group of organisms with a common ancestor; not a taxonomic identification.
Co-evolution Mutual evolutionary influence as the result of organisms being involved in symbiosis.
Commensalism A type of symbiosis; symbiont obtains benefit with no impact on the host.
Holobiont Collective term that includes all organisms engaged in a symbiotic relationship.
Kleptoplastic The ability to harvest chloroplasts from algae; used by the secondary host to make photosynthetic products.
Mutualism Type of symbiosis; host and symbiont benefit from the association.
Parasitism Type of symbiosis; symbiont obtains benefit at the expense of the host.
Primary mutualism The initial association between a symbiont and a host.
Secondary mutualism The incorporation of a primary mutualism into a secondary host; the primary host either becomes the symbiont or is lost.
Symbiosis Intimate, protracted relationship between two or more dissimilar organisms.

The scope of biodiversity covered in this volume hints at the complexity of life on our planet. However, no organism exists in isolation, and the evolution of organisms must be considered in the context of their interactions with each other. Symbiotic associations are intimate interactions involving two or more organisms that live in a close, protracted relationship. These intimate interactions are a strong driver for the evolution of the organisms involved in the association. Although *symbiosis* can be used to describe many diverse interactions between a variety of organisms in almost every environment on Earth, this article focuses on the diversity of mutualistic partnerships involving metazoans and microbes in the marine environment. Mutualisms can be found in every marine ecosystem across the globe: from the coast to the open ocean and from the surface waters to the deep sea (**Figure 1**). These associations have influenced the evolution and diversification of most, if not all, extant marine organisms. In fact, several marine ecosystems, such as coral reefs and hydrothermal vents, are only possible because of mutualistic symbioses between metazoans and microbes.

Introduction

The study of symbiosis in nature began in the late 1600s with the first observation of microorganisms by Antonie van Leeuwenhoek (1632–1723). The term *symbiosis*, derived from the Greek for "living together," was proposed 200 years later by Heinrich Anton de Bary (1831–1888) to describe phenomena such as mutualism and parasitism. However, the defined outcome of the relationship (positive or negative, depending on the perspective of the host) may become blurry as environmental conditions change (reviewed in Karst *et al.*, 2008; Sachs and Simms, 2006; Thrall *et al.*, 2007). This becomes an important consideration for many mutualisms in the coastal surface oceans where anthropogenic impacts such as eutrophication and increased sea surface temperatures are the most pronounced.

Structure in Symbiosis

Symbiotic associations come in a variety of forms and functions. The structural characteristics that form the relationship can be described in several categories: the relative location of partners to each other, how associations are transmitted between generations, and the obligacy and specificity of the relationship. These structural characteristics can be used to summarize any symbiotic association regardless of where the mutualism occurs or the overall function of the partnership.

Location is usually described as the position of the smaller "symbiont" in relation to the larger "host" (**Figure 2(a)**). Ectosymbionts (from *ecto*, the Greek for "out") occur on the surface of the host, whereas *endosymbionts* (from *endo*, the Greek for "within") occur internally, either within a body cavity but outside the cells (extracellular) or within cells (intracellular). Because of the intimate and protracted nature of mutualistic symbioses, by definition they must be maintained across generations, either through a closed/vertical or an open/horizontal transmission of symbionts to a new host (**Figure 2(b)**). In vertical transmission, hosts directly supply their offspring with symbionts – for example, by packaging them into egg cells – which ensures that all offspring inherit symbionts. If this is the only mechanism operating, then the system is closed to partner exchange and the association diverges together over time, resulting in coevolution of the host and symbiont lineages. In horizontal transmission, hosts produce nonsymbiotic offspring that must acquire symbionts from environmental sources. For these less intimate mutualisms, coevolution between partners typically breaks down because of the potential for partner exchange through environmental acquisition. However, the interaction between host and symbiont still drives their evolution. For example, extracellular, horizontally transmitted associations may not demonstrate coevolution, but the association still results in the evolution of specialized cells or tissues in the host that serve as accommodation for their symbionts. These evolutionary changes are not found in nonsymbiotic hosts.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00409-3

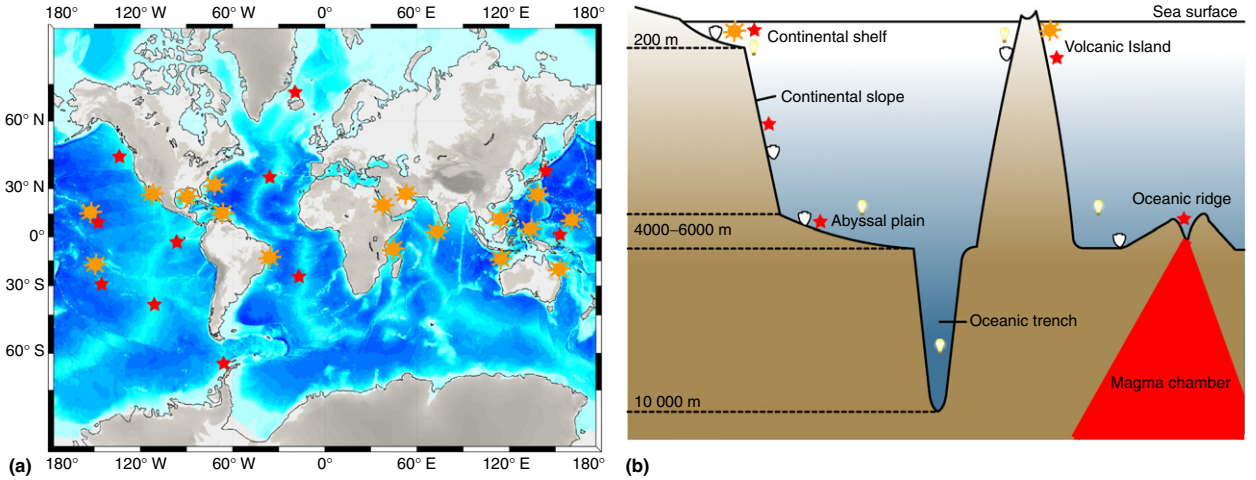


Figure 1 Distribution of symbioses. (a) Global distribution of photosynthetic symbioses (sun, between 30° N and 30° S) and confirmed locations of vent sites with chemosynthetic symbioses (star). (b) Example distribution of photosynthetic (sun), chemosynthetic (star), light-producing (light-bulb), and protective (shield) associations in discrete regions of the ocean. Map courtesy of S Soares (Oceanography Department, University of Hawai'i at Manoa).

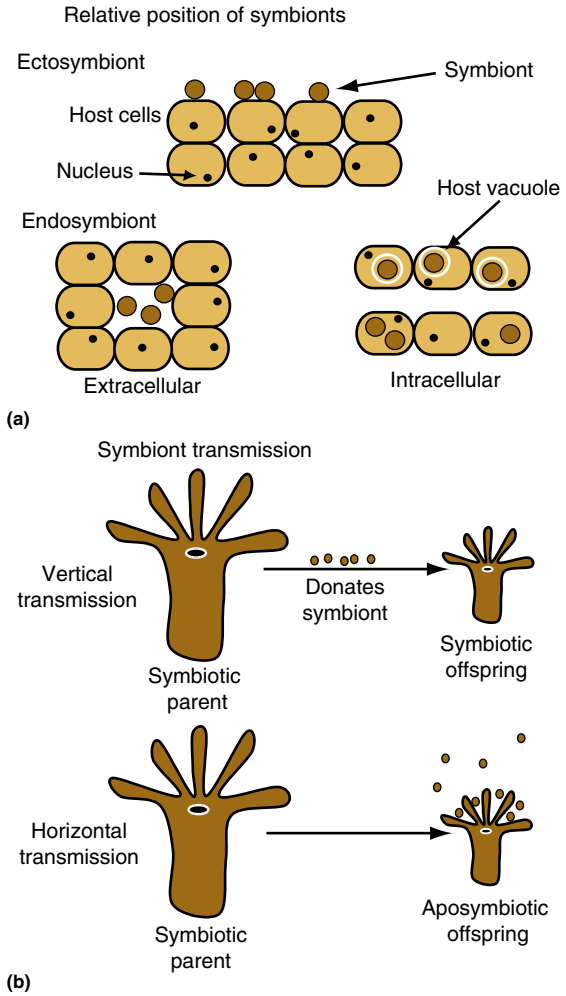


Figure 2 Structure of symbiosis. (a) Relatively position of symbionts (brown circles) to host cells (orange ovals with black nuclei). (b) Transmission of symbionts from host parent to offspring.

In many cases, the mode of transmission may indicate the necessity of the interaction, especially for partnerships with vertical symbiont transmission. However, necessity may not be a two-way street; one may be an obligate partner, requiring the association for survival, and the other may be facultative, with individuals able to live with or without the association. One final layer of complexity to consider when describing a symbiotic association is the specificity between partners. Does the association occur between two partners defined at the species, clade, or subclade level, or does it lie under some higher level of classification? Can one partner be found in association with several potential partners, either occurring in a range of hosts or as hosting a range of symbionts? Marine symbiotic associations occur in a variety of structural forms, with different combinations achieving the same symbiotic function.

Functions of Symbiosis

The diversity of marine symbioses can be categorized according to the function provided to the host by the symbiont. For example, imagine hosting small plants that live inside your skin and feed carbohydrates directly to your body. Or how about a small compartment just under your skin that glows in the dark because it houses light-producing bacteria? This article describes marine mutualisms grouped by three main functions: nutrient exchange, production of light, and production of protective compounds. Nutrient-exchange symbioses bring together organisms with complementary metabolic capabilities. Light-production symbioses typically occur in the deep, dark oceans of the world or in association with nighttime activities in which they provide hosts with a mechanism for capturing prey, avoiding predators, or communicating. Symbioses that produce protective compounds function to deter grazers, competitors, predators, or pathogenic microbes.

Humans also host numerous microbial symbionts known to function in nutrient exchange and the production of

protective compounds (Turnbaugh *et al.*, 2007), so it is important to remember that metazoans utilize symbiotic associations as a way to expand our physiological and metabolic capabilities. In general, metazoans are quite limited in their capacity to process or obtain nutrients, and they rely on the diverse metabolisms of microbes to supplement their native capabilities. The symbioses discussed in this article highlight some of those traits and describe how microbes have been coopted to function in metazoan hosts, creating a unique niche for both organisms in the marine environment.

Nutrient-Exchange Mutualisms

Photosynthetic Symbioses

The phenomenon of photosynthesis evolved once, originating in cyanobacteria (Cavalier-Smith, 2006). Today, a significant amount of global photosynthesis is still performed by cyanobacteria in the oceans, with land plants and algae contributing to the rest of the photosynthesis on this planet. If photosynthesis evolved only once in cyanobacteria, where did plant and algal chloroplasts originate? In 1905, Constantin Mereschkowsky first presented the concept that chloroplasts are derived from an ancient endosymbiotic event that combined cyanobacteria with a eukaryotic host (translated in Martin and Kowallik, 1999). This theory was later strengthened and expanded by Lynn Margulis in 1967 (as L. Sagan), with her presentation of cytological, biochemical, and paleontological evidence for a generalized symbiosis theory (Sagan, 1967) and by the discovery that chloroplasts contain DNA related to cyanobacteria (Bonen and Doolittle, 1975; Ris and Plaut, 1962; Zablen *et al.*, 1975). Aside from land plants and lichens (fungi in association with photosymbionts), all multicellular organisms with photosynthetic symbionts are found in aquatic environments.

The marine environment contains a variety of "photosynthetic animals," or metazoans in association with photosynthetic microbes such as cyanobacteria. However, most of the photosynthetic animal hosts in the marine environment are actually secondary mutualisms – situations in which a primary mutualism such as a eukaryotic alga containing chloroplast organelles is incorporated into a secondary host. These secondarily photosynthetic associations are often found thriving under low nutrient conditions that typically limit food abundance (Johnson, 2011; Vaughn *et al.*, 2010). The photosynthetic symbiont uses sunlight to fix carbon dioxide (CO_2) into organic carbon, much of which is directly translocated to the host to supplement feeding. In return, the host provides a stable, high light environment with a constant supply of CO_2 and nutrients such as nitrogen and phosphorus. Several host examples highlighted in this article are cnidarians, such as corals and anemones, bivalves like the giant clam, and the sea slug (*Elysia* spp.), which has a unique kleptoplastic nature (**Figure 3**).

Cnidarians have a relatively simple body plan comprised of two cell layers: an outer epidermis, which is a central acellular lay of mesoglea, and an inner gastrodermis that lines the digestive cavity and houses the intracellular endosymbionts (**Figure 3(a)**). Symbiont transmission between generations can be vertical or horizontal, depending on the host species. Hosts that employ vertical transmission typically transmit symbionts in egg cells that become fertilized to produce offspring or, in the case of brooding species, may transfer symbionts to offspring while they are still within the parent colony (**Figure 2(b)**). Horizontal transmission is established through feeding: the host's gastrodermal cells phagocytose the symbionts, and the symbionts survive by resisting digestion through mechanisms not yet understood. For reef-building corals, the symbiosis is considered obligate. However, some hosts, such as the sea anemone *Aiptasia* (**Figure 3(b)**), and some strains of the symbiotic algae, can be found free-living. Regardless of lifestyle, symbiotic cnidarians are only found in association with dinoflagellates in the genus *Symbiodinium*, which is described later in this section.

Corals (animal host and algal symbiont) living in warm, shallow waters have a unique role in the marine environment. They create the biological and structural foundation for the most diverse marine ecosystem on the planet: coral reefs. The reef structure is built by coral polyps, clonal individuals resembling anemones, that deposit a thin layer of calcium carbonate, layer by layer, generation after generation, for decades or even centuries (**Figure 3(c)**). Since corals typically live in low-nutrient waters, the nutritional supplement provided by their photosynthetic symbionts significantly enhances calcification and therefore growth of the colony. The coral structure provides habitat for many of the diverse organisms that live on coral reefs. Some corals with branching morphology show rapid growth, laying down 10–20 cm of skeleton per year, whereas corals with a more uniform, massive structure are estimated to add only 0.5–1 cm of growth per year (**Figure 3(d)**). Reef building is a dynamic process. Even as corals build the reef, wave motion and biological perturbation continuously degrade it. The reefs we are familiar with today have been around for thousands of years; however, changing environmental conditions on a global scale (increasing sea surface temperatures and ocean acidification) and on local scales (sedimentation and eutrophication from coastal runoff) are undermining the corals efforts to maintain the structure of the reef, and these important yet fragile ecosystems may eventually be lost in many places around the globe.

Giant clams are the largest bivalves in the world, reaching up to 120 cm in length and weighing more than 200 kg (**Figure 3(e)**). They are found in association with coral reefs in the South Pacific and Indian oceans. They also host *Symbiodinium* dinoflagellates as symbionts, with symbionts being acquired by host larvae during feeding. Symbionts that successfully avoid digestion are housed extracellularly in a tubular extension of the stomach that branches out into the mantle cavity. This structure is designed to provide the symbionts with ample light for photosynthesis and also creates the unique patterns observed in the clam's mantle tissue. For the clam, this is an obligate association, with up to 95% of the photosynthetically fixed carbon being translocated to meet the host's metabolic requirements.

Symbiodinium, the genus of dinoflagellate symbionts found in cnidarians and giant clams, was originally thought to be a single species, *Symbiodinium microadriaticum* Freudenthal (Freudenthal, 1962). Advances in genetic techniques have revealed at least nine different lineages within *Symbiodinium*

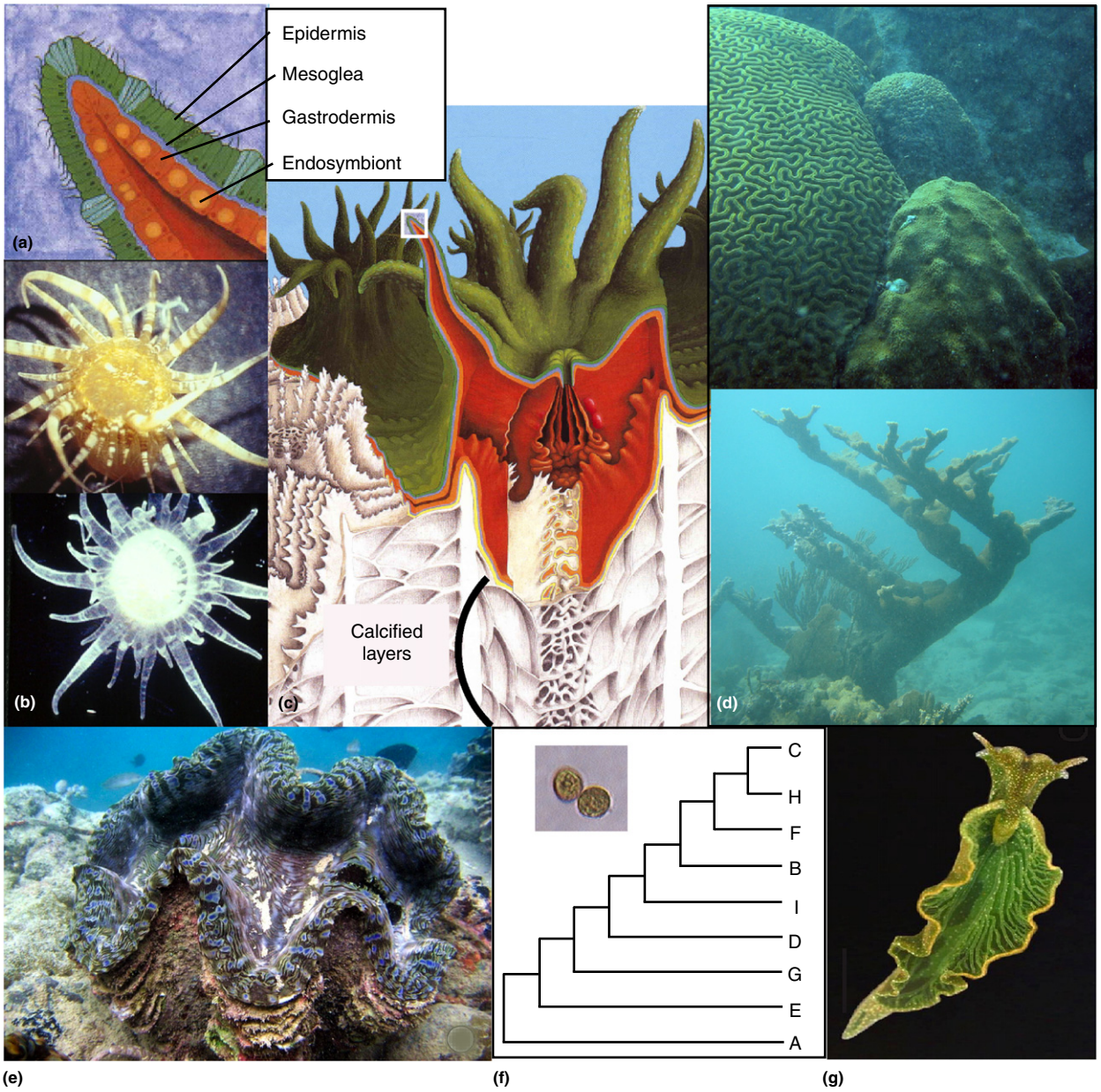


Figure 3 Photosynthetic symbioses. (a) Cross-section of a cnidarian tentacle depicting the tissue layers and location of symbionts (orange circles). (b) *Aiptasia pallida* in symbiosis with *Symbiodinium* (top) or aposymbiotic (bottom). (c) Deposition of calcium carbonate layers by coral polyps builds the reef structure. Reproduced with permission from Veron JEN (1993) *Corals of Australia and the Indo-Pacific*. Honolulu, HI: University of Hawai'i Press. (d) Two types of coral morphology: massive (top) and branching (bottom). (e) *Tridacna squamosa*. Reproduced from Hernawan UE (2008) *Biodiversitas* 9(1): 54, with permission from CMORE. (f) General phylogeny for eight subclades of *Symbiodinium* with photo insert. (g) *Elysia chlorotica* adult with chloroplasts. Reproduced with permission from Rumpho ME (2008) *Proceedings of the National Academy of Sciences of the United States of America* 105(46): 17868.

so far (**Figure 3(f)**). Although the taxonomy of the symbiont has not been fully resolved, host specificity within *Symbiodinium* can range from highly specific (associating with single subclade of *Symbiodinium* through time and space and between generations) to relatively flexible. For example, the giant clam associations appear to be relatively flexible, harboring heterogeneous populations of *Symbiodinium* in a single host. Outside of these metazoan associations, *Symbiodinium* is also found in symbiosis with single-celled protists such as foraminiferans.

Green sea slugs are one of the more bizarre examples of a nutrient-exchange symbiosis creating a photosynthetic animal. This association has a different interpretation of an animal–algal mutualism because the host sea slug steals chloroplasts from the algae, a process also known as *kleptoplasty*. The green sea slug feeds on algae, but instead of acting as a secondary host to the algal-chloroplast association, the sea slug bypasses the primary host by harvesting the chloroplasts directly from the algal cells (**Figure 3(g)**). To maintain a functioning

chloroplastic organelle without the primary algal host, the animal genome has acquired genes, typically found in plant or algal genomes, that support photosynthetic function. The majority of studies on kleptoplasty have been conducted on sacoglossan sea slugs in the genus *Elysia*, but this phenomenon also occurs in marine ciliates and foraminifera (reviewed in Wise *et al.*, 2006). The record for kleptoplastic living is held by *Elysia chlorotica*, surviving for 10 months exclusively on plastid-generated photosynthetic products (Rumpho *et al.*, 2011).

Chemosynthetic Symbioses

Although most of Earth's environments are supported by photosynthetic activity, several unique systems rely on chemosynthesis to fix carbon. Like the process of photosynthetic light harvesting, the oxidation of reduced sulfur and methane also releases high-energy electrons that can be converted to cellular energy and used to fix CO_2 into organic compounds. The most familiar chemosynthetic habitats, the hydrothermal vent systems, are found in the deep sea, but chemosynthetic symbioses are present at a range of depths and across many ocean environments (**Figure 1**). The more we look, the more chemosynthetic diversity is discovered in terms of habitat (vent, cold seeps, etc.), metabolic pathways, and symbiont–host partnerships (for a comprehensive review see Dubilier *et al.*, 2008). As with coral reefs, the deep ocean chemosynthetic symbioses provide the foundation, the source of fixed organic carbon, for the entire vent ecosystem. To highlight nutrient-exchange symbioses characterized by chemosynthetically fixed carbon, the author will focus on the classic hydrothermal vent tubeworms and bivalves (**Figure 4**).

Hydrothermal Vents

Formations of hydrothermal vents occur above pockets of magma under the seafloor, usually along tectonically active oceanic ridges like the East Pacific Rise and the Mid-Atlantic Ridge. Cold seawater seeps in through the surrounding seafloor, where the pocket of magma heats the water causing it to react with surrounding rocks, changing the chemistry of the water until it becomes highly acidic and rich in sulfides and other reduced metals. The heated water rises back to the seafloor and is released through vent sites at temperatures up to 400° C.

Biological communities were discovered along hydrothermal vents during the exploration of the Galapagos rift (1977–1979) (Karl *et al.*, 1980). Compared to other regions of the deep ocean, the vent community contained substantially more biomass than expected, since photosynthesis is not a significant source of organic carbon to the deep ocean. Originally, these communities were thought to be supported by free-living chemosynthetic bacteria, but further analyses revealed that mutualistic associations between animal hosts, such as giant tubeworms and bivalves, like clams and mussels, and chemosynthetic bacterial symbionts were responsible for generating the organic carbon that supports these vent communities.

Riftia vent tubeworms were the first animals found to host chemosynthetic bacteria (Cavanaugh *et al.*, 1981), and in spite



Figure 4 Giant tube worms and bivalves from hydrothermal vents at the East Pacific Rise at 2500 m depth. Each individual in the photo exceeds one meter in length. Photo credit: reproduced with permission from MBARI (2003).

of the challenges of studying organisms that live miles below the surface of the ocean, significant progress has been made in understanding these interactions. Sulfur-oxidizing bacterial symbionts are intracellular within the trophosome, a highly vascularized symbiosis organ inside the *Riftia* tubeworm. The symbionts provide chemosynthetically fixed carbon to the host, and the host provides the symbionts with the necessary substrates: sulfides and oxygen (O_2) from the environment and CO_2 from host respiration. Symbiont transmission between generations is horizontal, as host larvae are symbiont-free and acquire symbionts from the environment. As the larvae develop, their digestive structures are lost. Adult tubeworms have no digestive system, being entirely dependent on their symbionts for nutritional support. The *Riftia* tubeworms host a specific subclade of γ-proteobacteria, a large group of bacteria that contain free-living and pathogenic strains, as well as symbionts found in many other marine symbioses (Thornhill *et al.*, 2008).

One of the most interesting features of this association is the modified physiology of the host (Stewart and Cavanaugh, 2006). For example, the host must supply sulfides and O_2 to the trophosome. The transport of O_2 in the blood of the host is performed by hemoglobins, which are easily poisoned by sulfides that compete with O_2 for binding sites on the protein. In addition, reduced metals begin to react with dissolved O_2 immediately on release from the vent, oxidizing the metals and decreasing their energetic yield for chemosynthesis. To solve these problems, the host tubeworm grows at the interface between the sulfide-rich vent flow and the surrounding O_2-rich seawater, allowing access to both resources. To eliminate the toxicity of sulfides in the blood, the host tubeworm has developed a modified hemoglobin structure adapted to reversibly bind sulfide (for release in the trophosomes), independent of binding O_2 for host respiration.

Clams and mussels (bivalves) are found farther from the vent water source and typically house their symbionts in or on gill tissue, depending on the species (Dubilier *et al.*, 2008). The symbionts are again represented by a unique subclade of γ-proteobacteria, although symbiont transmission depends

on the association. Vesicomyid clam symbioses show cospeciation between host and symbiont, which is reflected in their symbiotic form: an intracellular, vertically transmitted association in which the symbiont is carried within the eggs. For other bivalve partnerships, less is known about their modes of transmission; however, there is no clear cospeciation, implying horizontal or mixed modes of transmission.

As for *Riftia* tubeworms, the bivalve host must supply their symbionts with the materials necessary for chemosynthesis and therefore face the same problems. Harboring the symbionts on the gill tissue provides an adequate supply of O_2 and CO_2 to the symbionts, but the concentration of reduced metals in the seawater decreases with distance from the source. In response, bivalves have developed several mechanisms for concentrating sulfides. For example, clams can extend their foot deep into sulfide-rich vent cracks that fissure out from the main vent site, so O_2 and sulfide acquisition is spatially separated. Another mechanism for concentrating reduced metals is the temporal separation of substrate acquisition. Some bivalves move up and down in the sediment. They collect O_2 from the sediment–seawater interface and then burrow deep into the sediment where sulfides are enriched. As a mechanism to deal with sulfide toxicity in the circulatory system, the host has developed a separate carrier protein that binds sulfides, leaving hemoglobin to carry O_2.

Vent sites are patchy in space and time, occurring where and when tectonic conditions are appropriate. Once they occur, they are ephemeral, typically lasting on the scale of years to decades. Therefore, establishment and community production at vent sites occurs rapidly, reaching maturity within a few years. Mature communities produce a vast amount of motile larvae that move along the ocean currents in search of new vent sites.

Nutrient Conversion

Not all nutrient-exchange symbioses involve autotrophic symbionts. The diverse capacity of microbial metabolism has resulted in symbiotic associations with varied functional roles, including the conversion of complex carbon compounds and the fixation of atmospheric nitrogen into forms that can be used by the host. Many symbioses, especially those occurring in nutrient poor regions of the ocean where nitrogen is limiting, such as cnidarian–algal mutualisms (*see* Photosynthetic Symbioses) have secondary associations with nitrogen-fixing microbes, typically marine cyanobacteria. The fixation of atmospheric nitrogen is covered in depth elsewhere in this volume and a recent review by Fiore *et al.* (2010) highlights nitrogen fixation in marine symbioses, including interesting associations between cyanobacteria and diatoms. Instead this section highlights two marine symbioses involved in the conversion of complex carbon compounds. The first association is unique within the field of symbioses as a whole and so far found nowhere else but in the oceans; the second is very common on land but unique in the marine environment.

Worms and Whales

The death and sinking of a large marine vertebrate, such as a whale, delivers a rich carbon source to the deep ocean, where large scavengers such as sharks, crabs, and fish recycle the majority of the fallen biomass into living tissue. However, when only the bones remain, a unique marine microbial symbiosis is there to break them down.

During an exploration of the ocean floor in Monterey Bay, California, in 2002, the bones of a gray whale were found covered in red feathery plumes of worms later determined to be a new genus named *Osedax* (Latin *bone devourer*) (Rouse *et al.*, 2004) (**Figure 5**). Like their hydrothermal vent relatives, the adult worms have no mouth or digestive system, and the larvae appear to be without symbionts, so the association is likely horizontally transmitted. Their symbionts are also γ-proteobacteria; however, members of this unique *Osedax*-associated subgroup of γ-proteobacteria are heterotrophic. They utilize the lipid-rich hydrocarbons inside the bone marrow as a carbon source for their host, making this nutrient-exchange association unique in the world of symbiosis. Also, like their vent counterparts, this symbiosis relies on ephemeral nutrient sources, so they establish quickly, occur in large populations, and reproduce in vast numbers. Since their original discovery, 17 species of *Osedax* worms have been identified throughout the world's oceans.

''Worms'' and Ships

Although these organisms are called *shipworms*, or "termites of the sea," they are actually bivalve mollusks that bore into and consume wood submerged in the ocean. Historically, they have caused extensive damage to human-made structures, including the demise of several ships during Christopher Columbus's fourth voyage (early 1500s), and they continue to feed on the wooden bridges and piers in coastal cities today. Like termites on land, shipworms subsist on wood, which is very nitrogen-poor and difficult to digest, so the host, once again, relies on γ-proteobacterial symbionts to convert nutrients into a form usable by metazoans (Luyten *et al.*, 2006). Shipworms host a consortium of intracellular, endosymbiotic cellulose-digesting, nitrogen-fixing microbes.



Figure 5 *Osedax frankpressi* worms on a bone fragment from a gray whale in Monterey Canyon.

Production of Light

Bioluminescence, the endogenous production of light through biological activity, involves a chemical reaction in which the oxidation of luciferin (substrate) by luciferase and a photoprotein (enzyme) produces oxyluciferin and light (products) (**Figure 6(a)**). Bioluminescence is relatively rare on land and found in arthropods such as fireflies, a snail, a few earthworms, and some fungi. Bioluminescence in the marine environment, however, is quite common and found in a diverse array of organisms ranging from bacteria to fishes; it is especially common in nocturnal creatures or those living in the deep ocean where little sunlight penetrates (reviewed in Haddock *et al.*, 2010). The luminescence reaction in the marine environment produces light in the blue-green spectrum (~470 nm), probably because blue light transmits the farthest in water. One easy-to-see example sometimes occurs at night along the coast and appears as waves that glow as they break on the shore. This light is produced by dinoflagellates during bloom events (**Figure 6(b)**).

Although many marine organisms produce their own light, others have become hosts to light-producing marine microbes. For example, many squid and bony fish harbor packets of luminescent bacteria that glow within the body of the animal. These light-producing symbioses appear to have several independent origins – hosting different species of luminescent bacteria in different locations around the host and for different purposes such as communication, prey attraction, and camouflage. The hosts in these associations are highly specific for their luminescent symbionts even though the associations are extracellular and most if not all are horizontally transmitted. The bacterial symbionts and their luminescent free-living relatives are members of the γ-proteobacteria. Symbiotic strains include at least one cultured species in the genus *Vibrio*

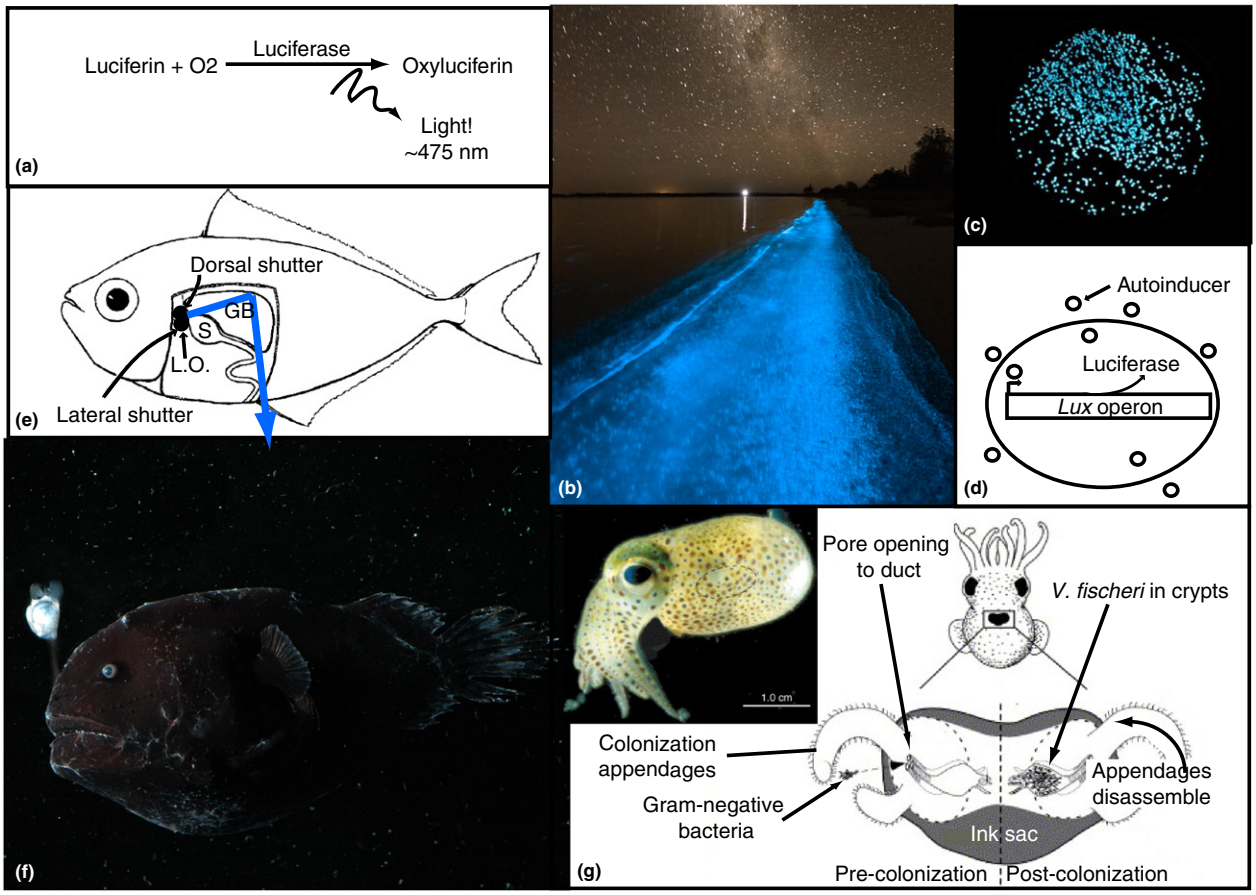


Figure 6 Light-producing symbioses. (a) Chemical reaction responsible for producing light in biological systems. Luciferase catalyzes a reaction between luciferin and O_2, releasing light. (b) Bioluminescence on the shore of Gippsland Lakes, Victoria, Australia. Reproduced from Hart P (2009), with permission from CMORE, http://www.philhart.com (c) *Vibrio fischeri* colonies glowing on a petri dish. (d) Diagram of quorum sensing by a bacterial cell that results in light production. Increased concentrations of the autoinducer allows for diffusion into the bacterial cell, which induces production of the luciferase enzyme. (e) Diagram of light production and emission in ponyfish. The light organ (LO) emits light that is controlled by the dorsal and lateral shutters. The arrow demonstrates the path of light released by the dorsal shutter. Reproduced from Mcfall-Ngai MJ and Dunlap PV (1984) External and internal sexual dimorphism in leiognathid fishes: Morphological evidence for sex-specific bioluminescent signaling. *Journal of Morphology* 182: 71–83, with permission from Wiley. (f) Live image of the anglerfish *Chaenophryne longiceps* with a bioluminescent lure. Reproduced from Haddock S (2011), with permission from Steven Haddock of the Monterey Bay Aquarium Research Institute. (g) *Euprymna scolopes*. The diagram shows pre- and postcolonization mechanisms involved in selecting *V. fischeri* from other Gram-negative bacteria in the seawater. Reproduced from Visick KL and Ruby EG (2006) *Current Opinion in Microbiology* 9: 1, with permission from Elsevier.

(*V. fischeri*) (**Figure 6(c)**) and several species of *Photobacterium* (*P. leiognathi* and *P. phosphoreum*). The necessity of the association for host survival is less well defined than the nutrient-exchange symbioses. However, hosts lacking the ability to produce light for prey attraction or predator avoidance will likely not be successful over time. From a symbiont perspective, the culturable strains are considered facultative and free-living under nutrient-rich conditions, but several strains remain resistant to culturing, suggesting that they are highly specialized for survival inside the host light organ. However, most bacterial species are currently unculturable, so this feature does not guarantee that they are in an obligate association with their host. As a mutualism, the hosts gain a variety of functions that will be highlighted shortly, and the bacteria are provided with a nutrient-enriched environment for growth, allowing populations to survive in much greater densities than in the open ocean.

Density becomes an important consideration for light production by bacteria, since a bacterium is only several microns long. The amount of light produced by a single cell quickly becomes a waste of valuable energetic resources. However, when enough cells are present in a small area such as a host's light organ, the light output becomes substantial. But how do bacteria know when enough cells are present? The regulation of light production includes a mechanism for *quorum sensing*, allowing the cells to chemically detect density (also important in pathogenesis). Quorum sensing involves the continual release of a chemical into the environment. When the chemical, an autoinducer, reaches a critical concentration, it triggers the induction of a metabolic process. In the case of light production, the autoinducer turns on the production of the luciferase enzyme, thereby turning on luminescence (**Figure 6(d)**).

Fish

Light production in bony fish is often autogenic, but those that host bacterial symbionts have developed a variety of symbiosis structures (Haygood, 1993). Several of these host organs occur as outpockets of the gut, which suggests that colonization occurs during feeding from free-living populations (confirmed by culturing). Others house the bacteria in highly specialized superficial structures that are open to the environment. However, their symbionts have not been seen free-living (nor cultured), so colonization mechanisms are unknown. The location of the light organ and direction of light emission provides some information about function, with a few representative examples described in the following paragraphs.

Ponyfish, or Leiognathids, reside in shallower, near-shore waters in the Indo-Pacific Ocean and are characterized by a protrusible mouth and a circumesophageal light organ (**Figure 6(e)**). The light organ is colonized by *P. leiognathi*, and emission of light by the light organ is controlled by highly modified structures in the host fish. The light organ is positioned next to the gas bladder, with a dorsal shutter that controls the entrance of light into the gas bladder. The gas bladder is silvered over to allow the light to be reflected to the posterior end and out through the ventral side of the fish body near a transparent anal fin (Mcfall-Ngai and Dunlap, 1984). Other lateral shutters are present around the light organ to control lateral light emission. The position of light emission appears to correlate with function; ventral emission near the anal fin acts as camouflage, and lateral emission plays a role in communication.

Anglerfish or Ceratioids, occur in the deep, dark abyssal ocean as solitary, relatively rare fishes. They are characterized by a prominent bioluminescent lure at the end of a modified dorsal spine that extends out over the head (**Figure 6(f)**). The superficial light organ is colonized by a strain of unculturable *Vibrio*-like species of symbiont. The function of the lure, easily identified by the location, is used for prey attraction. Symbiont transmission is unknown, but the rarity and solidarity of these fishes and the fact that their symbiont has not been cultured suggest a mechanism for vertical transmission.

These few examples highlight the major functions of luminescence in marine fishes but are not the only examples. Similar phenomena are seen in more than 20 families of bony fishes and are present in all marine environments around the globe.

Squid

Bioluminescence is very common in squids, with at least 70 luminous genera. Most produce endogenous light, but only a few families are known to utilize bioluminescent bacteria: the *Sepiolidae* and *Loliginidae*. One species, *Euprymna scolopes*, the Hawaiian bobtail squid, has become the model organism for studying light-producing symbioses, as well as a model for host symbiont–pathogen interactions and metazoan immunity in general (Mcfall-Ngai, 2008; Mcfall-Ngai *et al.*, 2010). This particular association has become a well-studied scientific model because the organisms are easy to catch in the wild, the host can be reared from egg to adult in aquaria, and the *V. fischeri* symbionts can be cultured independently of the host (**Figure 6(c)**).

Hawaiian bobtail squid are small, nocturnal creatures, about 4 cm in length, that burrow in the sand during the day and come out at night to hunt. The light organ lies ventral to the ink sac, which reflects light emitting from the light organ. This reflected light creates down-welling, counter illumination that masks any shadow caused by overhead moon or starlight. The light organ consists of two enclosed sacs called *crypts* which are connected to the surrounding seawater through ducts that open up as pores in the mantel cavity (**Figure 6(g)**). In spite of the continuous connection with the environment, the light organ houses a dense monoculture of *V. fischeri*, which are acquired from the environment by juvenile squids, typically within hours after hatching. Adults regulate symbiont population by venting up to 95% of their symbionts into the environment at the end of every night. Then they bury themselves in the sand for the day, allowing the bacterial culture to repopulate; by nightfall, the light organ is back to full capacity.

The initiation, recognition, and establishment stages of this symbiosis are very well understood. They are presented in some depth here to illustrate the intricate regulation mechanisms involved in setting up a highly specific, horizontally transmitted symbiosis (Nyholm and Mcfall-Ngai, 2004). Most

if not all horizontally transmitted associations likely require several selection stages as well, but for most associations the mechanisms remain unknown.

Initiation begins at contact between potential partners. Typically, there is an initial mechanism of recognition that allows for the two partners to remain in contact, especially in the marine environment where water motion can easily disrupt the interaction. In the *Euprymna–Vibrio* association, juvenile squid ventilate water into the mantle cavity, drawing in the surrounding microbial community. The presence of bacteria in the seawater induces the host to generate mucus on specialized colonization appendages near the ducts to the light organ. Even though any bacteria can induce mucus shedding, only some (the Gram-negative bacteria) are capable of adhering to the mucus; out of those that can attach, *V. fischeri* is preferentially recruited, even though it only represents $<0.1\%$ of the population. This is the first stage of the "winnowing" process (Nyholm and Mcfall-Ngai, 2004).

Recognition continues as potential symbionts migrate toward the pores and down the ducts to the light organ. The environment within the ducts is highly discriminatory, and only *V. fischeri* is able to reach the crypts for colonization.

Establishment within the crypts of the light organs is not just a matter of getting there. Once *V. fischeri* has entered the crypt, it begins to divide and populate the chamber. However, *V. fischeri* strains that are unable to produce light are unable to maintain residence (Lupp *et al.*, 2003); therefore, the host has some mechanism of sensing light production, ensuring that its symbiont culture is doing its part. Once a culture of symbionts has been successfully established, molecular signals are sent to the host to halt mucus production in the mantle cavity, disassemble the colonization appendages near the entrance to the ducts, and resist subsequent infections. The mature symbiosis persists for the life of the squid.

Production of Protective Compounds

So far, the mutualistic marine microbes presented in this chapter have been shown to fix carbon using light or chemical energy, convert cellulose and atmospheric nitrogen into a biological usable form, and generate light. This final section introduces the production of protective compounds (which function in host defense) by microbial symbionts. Many of these protective compounds are metabolite derivatives, byproducts of metabolic pathways not found in metazoans. A diverse array of marine microbes are capable of producing an assortment of unique protective compounds; as marine metazoans go, sponges are the experts at hosting microbial symbionts (reviewed in Taylor *et al.*, 2007). In fact, one facet of research on marine sponges targets the isolation of bioactive natural products from sponge symbiotic communities for use in pharmaceuticals.

Sponges

Sponges represent the base of the metazoan tree of life. They are sessile filter feeders with a relatively simple body structure, minimal coordination or differentiation of tissues, and a well-adapted innate immune system (**Figure 7**). Their body structure is designed to filter large volumes of water for food particles, with estimates of up to 20,000 l of water per kilogram of sponge per day (Vogel, 1977). This high-throughput lifestyle exposes the sponge to a vast array of microbes from the environment, which may end up as food, microbial mutualists, or even pathogens. Regardless, sponges continue to be very successful, representing a significant component of the benthic marine community worldwide. Sponges, as well as other sessile filter feeders such as corals, rely on the production of protective compounds as a defense mechanism against natural threats such as pathogens, parasites, predators, biofoulers, and competitors. So far, many of these compounds appear to be secondary metabolites produced by microbial symbionts, and sponges host an amazing diversity of microbes from all domains of life – Bacteria, Archaea, and Eukarya. The sponge microbial community can contribute up to half of the host's biomass, with microbial densities 100 times greater than the surrounding seawater. The sponge holobiont is therefore capable of producing a vast array of bioactive metabolites, with several hundred new metabolites characterized each year.



Figure 7 Sponge biodiversity. Included are the yellow tube sponge (*Aplysina fistularis*), the purple vase sponge (*Niphates digitalis*), the red encrusting sponge (*Spiratrella coccinea*), and the gray rope sponge (*Callyspongia* sp.). Photo credit: reproduced from Twilight Zone Expedition Team (2007) NOAA-OE.

With all of this sponge-associated microbial diversity, several fundamental questions arise before the system can be considered a symbiosis, such as "How different are the microbial assemblages from the surrounding environment?" and "Do the associations persist across space and time?" With recent advances in molecular biology and genetics, it seems clear that sponges do host specific microbial communities, relatives of free-living bacteria found only in sponges that seem to be relatively stable across space and time. Some symbionts are intracellular endosymbionts, whereas others are extracellular endosymbionts within the sponge body (**Figure 2**). Some symbionts are vertically transmitted in the egg, sperm, or brooded larvae, and others are horizontally transmitted through mechanisms that are not well understood (Webster *et al.*, 2010). The complex diversity within these associations masks a clear determination for the facultative–obligate nature of the relationships. However, from the perspective of a

mutualism, the sponge seems to act as a protective, high light (in shallow water habitats), nutrient-enriched ecosystem that harbors a complex symbiont community. In addition to producing various protective compounds, the symbionts may also participate in nutrient-exchange associations similar to those previously described such as fixing carbon via photosynthesis in high light environments and chemosynthesis near vent systems and sponge-specific strains of cyanobacteria that fix atmospheric nitrogen for hosts living under low nutrient conditions. However, as with several other mutualisms described here, many sponge communities occurring near coastal areas are under threat from changing environmental conditions.

Sponge symbioses demonstrate the elegant complexity of marine symbioses, but they also highlight the difficulties inherent in trying to study these complex associations. When considering their important role in the ecosystem and their potential benefit to the field of human medicine, it is important that science continues to explore the role of hosts and symbionts in complex partnerships. The sponge's symbiotic community also highlights the importance of considering individual organisms in the context of their association with each other.

Symbioses in the Future

The symbioses presented in this chapter not only represent the diversity of mutualisms between metazoans and microbes in the marine environment but also illustrate the range of complexity in symbiotic associations – from light production involving a single host species and a monoculture of microbial symbionts to sponge associations, where one host may associate with many symbionts of different forms and functions – with a level of microbial diversity that rivals an ecosystem. In either case, neither the host nor the symbiont can be viewed exclusively as an isolated organism. Their life histories are interconnected, interdependent, and interactive on all levels of biology, from cell to ecosystem, across evolutionary time. From a diversity perspective, each organism involved in a symbiosis should be considered as an individual and as a collective organism. If one considers parasitism and viral infection, then every organism on Earth becomes part of a collective. This is especially important when trying to understand an organism's past, present, and future evolutionary trajectory.

From creating the most diverse marine ecosystems in the ocean to potentially providing cures for disease, the importance of marine symbioses cannot be overstated. However, like much of the diversity presented in this volume, these symbioses face huge challenges in the future. The ocean is undergoing rapid change, with increases in sea surface temperature and acidification occurring on a global scale, and eutrophication and invasive species impacting local ecosystems. Although some associations may be relatively protected from these human impacts by their remote locations (e.g., hydrothermal vent communities), those that thrive in the surface oceans, such as cnidarian–algal symbioses and complex sponge communities, are under severe stress from these rapid changes. In addition to the acute impacts, these stressors also affect the physiological health of the holobiont, making the collective more susceptible to infection and disease. For the few marine mutualisms in which the impacts of changing ocean conditions have been tested, the outcome is generally negative, causing disruption or dissolution of the symbiosis (Dubinsky and Stambler, 1996; Ove Hoegh-Guldberg, 1999; O. Hoegh-Guldberg *et al.*, 2007; Weis, 2010). However, we are still decades away from really understanding what our future oceans will look like and whether these globally important marine symbioses have the potential to adapt and survive.

See also: Coevolution. Corals and Coral Reefs. Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea. Eukaryotes, Origin of. Fungi. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Introduced Species, Impacts and Distribution of. Invertebrates, Marine, Overview. Marine and Aquatic Communities, Stress from Eutrophication. Marine Conservation in a Changing Climate. Marine Ecosystems, Human Impacts on. Marine Ecosystems. Ocean Ecosystems. Vents

References

Bonen L and Doolittle WF (1975) On the prokaryotic nature of red algal chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* 72: 2310–2314.

Cavalier-Smith T (2006) Cell evolution and Earth history: Stasis and revolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361: 969–1006.

Cavanaugh CM, Gardiner SL, Jones ML, Jannasch HW, and Waterbury JB (1981) Prokaryotic cells in the hydrothermal vent tube worm *Riftia pachyptila* Jones: Possible chemoautotrophic symbionts. *Science* 213: 340–342.

Dubilier N, Bergin C, and Lott C (2008) Symbiotic diversity in marine animals: The art of harnessing chemosynthesis. *Nature Reviews Microbiology* 6: 725–740.

Dubinsky ZVY and Stambler N (1996) Marine pollution and coral reefs. *Global Change Biology* 2: 511–526.

Fiore CL, Jarett JK, Olson ND, and Lesser MP (2010) Nitrogen fixation and nitrogen transformations in marine symbioses. *Trends in Microbiology* 18: 455–463.

Freudenthal HD (1962) *Symbiodinium* gen. nov. and *Symbiodinium microadriaticum* sp. nov., a Zooxanthella: Taxonomy, life cycle, and morphology. *Journal of Eukaryotic Microbiology* 9: 45–52.

Haddock SHD, Moline MA, and Case JF (2010) Bioluminescence in the sea. *Annual Review of Marine Science* 2: 443–493.

Haygood MG (1993) Light organ symbioses in fishes. *Critical Reviews in Microbiology* 19: 191–216.

Hernawan UE (2008) *Biodiversitas* 9(1): 54.

Hoegh-Guldberg O (1999) Climate change, coral bleaching, and the future of the world's coral reefs. *Marine and Freshwater Research* 50: 839–866.

Hoegh-Guldberg O, Mumby PJ, Hooten AJ, *et al.* (2007) Coral reefs under rapid climate change and ocean acidification. *Science* 318: 1737–1742.

Johnson M (2011) The acquisition of phototrophy: Adaptive strategies of hosting endosymbionts and organelles. *Photosynthesis Research* 107: 117–132.

Karl DM, Wirsen CO, and Jannasch HW (1980) Deep-sea primary production at the Galapagos hydrothermal vents. *Science* 207: 1345–1347.

Karst J, Marczak L, Jones MD, and Turkington R (2008) The mutualism–parasitism continuum in ectomycorrhizas: A quantitative assessment using meta-analysis. *Ecology* 89: 1032–1042.

Lupp C, Urbanowski M, Greenberg EP, and Ruby EG (2003) The *Vibrio fischeri* quorum-sensing systems *ain* and *lux* sequentially induce luminescence gene expression and are important for persistence in the squid host. *Molecular Microbiology* 50: 319–331.

Luyten YA, Thompson JR, Morrill W, Polz MF, and Distel DL (2006) Extensive variation in intracellular symbiont community composition among members of a

single population of the wood-boring bivalve *Lyrodus pedicellatus* (Bivalvia: Teredinidae). *Applied Environmental Microbiology* 72: 412–417.
Martin W and Kowallik KV (1999) Annotated English translation of Mereschkowsky's 1905 paper. *European Journal of Phycology* 34: 287–295.
Mcfall-Ngai M (2008) Host–microbe symbiosis: The squid–*Vibrio* association – A naturally occurring, experimental model of animal/bacterial partnerships. In: Huffnagle GB and Noverr MC (eds.) *GI Microbiota and Regulation of the Immune System*, pp. 102–112. New York: Springer.
Mcfall-Ngai M, Nyholm SV, and Castillo MG (2010) The role of the immune system in the initiation and persistence of the *Euprymna scolopes–Vibrio fischeri* symbiosis. *Seminars in Immunology* 22: 48–53.
Mcfall-Ngai MJ and Dunlap PV (1984) External and internal sexual dimorphism in leiognathid fishes: Morphological evidence for sex-specific bioluminescent signaling. *Journal of Morphology* 182: 71–83.
Nyholm SV and Mcfall-Ngai M (2004) The winnowing: Establishing the squid–*Vibrio* symbiosis. *Nature Reviews Microbiology* 2: 632–642.
Ris H and Plaut W (1962) Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas. The Journal of Cell Biology* 13: 383–391.
Rouse GW, Goffredi SK, and Vrijenhoek RC (2004) Osedax: Bone-eating marine worms with dwarf males. *Science* 305: 668–671.
Rumpho ME (2008) *Proceedings of the National Academy of Sciences of the United States of America* 105(46): 17868.
Rumpho ME, Pelletreau KN, Moustafa A, and Bhattacharya D (2011) The making of a photosynthetic animal. *Journal of Experimental Biology* 214: 303–311.
Sachs JL and Simms EL (2006) Pathways to mutualism breakdown. *Trends in Ecology & Evolution* 21: 585–592.
Sagan L (1967) On the origin of mitosing cells. *Journal of Theoretical Biology* 14(225–274): IN1–IN6.
Stewart FJ and Cavanaugh CM (2006) Symbiosis of thioautotrophic bacteria with *Riftia pachyptila*. In: Overmann J (ed.) *Molecular Basis of Symbiosis*, pp. 197–225. Heidelberg, Berlin: Springer.
Taylor MW, Radax R, Steger D, and Wagner M (2007) Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiology and Molecular Biology Reviews* 71: 295–347.
Thornhill DJ, Wiley AA, Campbell AL, Bartol FF, Teske A, and Halanych KM (2008) Endosymbionts of *Siboglinum fiordicum* and the phylogeny of bacterial endosymbionts in Siboglinidae (Annelida). *Biology Bulletin* 214: 135–144.
Thrall PH, Hochberg ME, Burdon JJ, and Bever JD (2007) Coevolution of symbiotic mutualists and parasites in a community context. *Trends in Ecology & Evolution* 22: 120–126.
Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, and Gordon JI (2007) The human microbiome project. *Nature* 449: 804–810.
Vaughn JM, Balch WM, Novotny JF, *et al.* (2010) Isolation of *Emiliania huxleyi* viruses from the Gulf of Maine. *Aquatic Microbial Ecology* 58: 109–116.
Veron JEN (1993) *Corals of Australia and the Indo-Pacific.* Honolulu, HI: University of Hawai'i Press.
Visick KL and Ruby EG (2006) *Current Opinion in Microbiology* 9: 1.
Vogel S (1977) Current-induced flow through living sponges in nature. *Proceedings of the National Academy of Science USA* 74: 2069–2071.
Webster NS, Taylor MW, Behnam F, *et al.* (2010) Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environmental Microbiology* 12: 2070–2082.
Weis VM (2010) The susceptibility and resilience of corals to thermal stress: Adaptation, acclimatization or both? *Molecular Ecology* 19: 1515–1517.
Wise RR, Hoober JK, Rumpho ME, Dastoor FP, Manhart JR, and Lee J (2006) The kleptoplast. In: Govindjee, Amesz J, Barber J, *et al.* (eds.) *The Structure and Function of Plastids*, pp. 451–473. The Netherlands: Springer.
Zablen LB, Kissil MS, Woese CR, and Buetow DE (1975) Phylogenetic origin of the chloroplast and prokaryotic nature of its ribosomal RNA. *Proceedings of the National Academy of Sciences of the United States of America* 72: 2418–2422.

Marine Viruses

Grieg F Steward, Alexander I Culley, and Elisha M Wood-Charlson, University of Hawaii at Manoa, Honolulu, HI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bacteriophage A virus that infects a bacterium.
Cyanophage A virus that infects a cyanobacterium.
Lysogen A bacterium which has a temperate virus integrated into its genome. So called because it may lyse at some later time as a result of the virus excising from the genome and initiating replication (induction). This occurs spontaneously with low probability or with high probability in response to a stress to the host cell.
Lytic virus A virus for which the usual or only mode of replication is through synthesis of new virions and lysis of the host cell.
Plasmid A piece of DNA (usually circular) that can replicate independently of the chromosome of a cell and which codes for supplementary (noncore) functions.
Temperate virus A virus for which there is a high probability that the viral genome will reside passively in the host cell, sometimes by integrating its genome into the host genome, rather than initiating a lytic infection, thereby creating a lysogen or a persistent infection.
Transposon A DNA sequence that can transpose itself, or a copy of itself, from one location to another within a larger segment of DNA or from one molecule of DNA to another. Also referred to as "jumping genes."
-virad A suffix denoting viruses that belong to an order of viruses (e.g., viruses in the order *Picornavirales* are picornavirads).
-virid A suffix denoting viruses that belong to a family (e.g., viruses belonging to the family *Phycodnaviridae* are picornavirids).
Virion An individual viral particle.

The Discovery of Viruses

The agents of viral diseases were being studied by some eminent microbiologists such as Louis Pasteur, by at least the 1880s, but the truly unique character of these infectious agents was not fully appreciated for some time. The first clue that viruses were not like other known pathogens came from experiments conducted independently in the 1890s by the Russian scientist Dmitri Ivanovsky and the Dutch scientist Martinus Beijerinck on a disease causing agent in tobacco plants in the 1890s. These scientists discovered that sap from infected plants retained its infectivity even after passage through filters with pores small enough to remove bacteria. This filtration test became the basis for assigning other pathogens to this new operationally defined class referred to as "filterable agents" (Levine and Enquist, 2007). In the subsequent 114 years, the study of viruses has not only led to a detailed understanding of their molecular nature, diversity, and ecology, it has resulted in the elucidation of some of the fundamental processes of cells in a new field of research (molecular biology), and led to relief from some of the most virulent diseases known to humankind. Some landmarks in virology at the beginning of the twentieth century included the discovery and identification of the viral agents of diseases such as yellow fever, influenza, rabies, and polio, and the realization that viruses were involved in some types of cancer. The discovery that bacteria are also susceptible to viral infections was discovered independently by Frederick Twort in 1915 and Felix d'Herelle in 1917. The first clear proof that viruses are particles and not a fluid was provided by d'Herelle's experiments in which he showed that dilutions of the lytic agent resulted in discrete zones of lysis on a lawn of bacteria. This was later confirmed with the application of electron microscopy (1939) and X-ray crystallography (1941) to the analysis of viruses, which provided the first detailed views of the structure of viruses. The relative simplicity of bacteriophages, and the ease with which they could be propagated in the lab under controlled conditions, made them an attractive model system for the nascent field of molecular biology. Work with bacteriophages resulted in the demonstration that nucleic acids, not proteins, were the repository for genetic information (1950s), the elucidation of the mechanisms of gene regulation and of the triplet code of genes (1960s), and the development of tools that led to recombinant DNA technology (1970s). The first genome to be sequenced was that of the RNA-containing bacteriophage, MS2, in 1976, followed by the DNA genome of bacteriophage, øX174, in 1977. These were the inaugural events in the genomics revolution in biology. The number of viral genomes completely sequenced is now in the thousands and rising rapidly.

Viruses and Biodiversity

In the context of biodiversity, viruses occupy a curious position. Since they are not organisms *per se*, they might be considered exempt from the discussion. However, they are extraordinarily diverse biological entities, both genotypically and phenotypically, and have been instrumental in promoting biodiversity among all cellular life forms. Viruses are clearly distinct from the cellular life on which they depend for replication in several important respects (Lwoff, 1957) the most salient being that they replicate by assembly rather than division (**Figure 1**). They are, in essence, molecular symbionts of cells, but unlike other molecular symbionts, such as plasmids and transposons, they package their genome in a protein shell (capsid) for efficient dispersal outside of a cell. Encapsidation serves to protect the viral nucleic acid from degradation in the

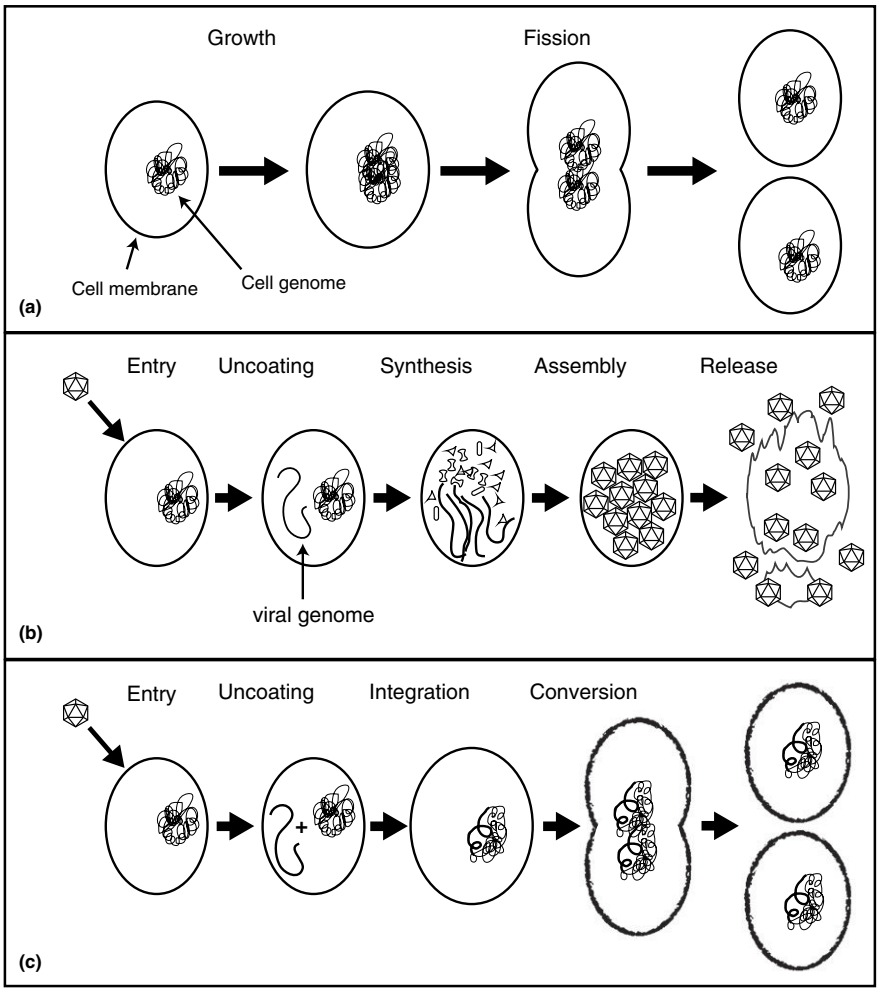


Figure 1 Comparison of the modes of replication of cells and viruses. Cells acquire energy and material from the environment to grow in size and to increase the population by binary fission (a). Viruses neither grow nor divide, but replicate by directing the machinery of a host cell to synthesize constituent parts, which then self-assemble into new virions (b). In the mode of replication shown in (b), the viruses act as parasites (or parasitoids) and the exit of the virions from the host may or may not result in lysis of the cell. Some viruses are capable of a different mode of replication wherein the genome merges with the host genome (c). The virus genome is then replicated along with the host genome and the viral DNA is propagated into each daughter cell. Genes encoded by the integrated viral DNA can alter (or convert) the phenotype of the host. Such conversions are often advantageous for the host, and in such cases the virus-host hybrid can be viewed as a mutualism.

environment and provides the mechanism by which a virion can recognize and infect another appropriate host cell. Because of the explicitly extracellular phase in the viral replication cycle – a phase in which the virions have a proteome and a distinct morphological phenotype – viruses are often casually gathered under the umbrella term "microorganisms." This usage, unfortunately, blurs the fundamental difference between viruses and cells, but perhaps explains why they, alone among the molecular symbionts, have a well-developed taxonomy modeled on that of other organisms.

Viruses are assigned to orders, families, genera, and species, but they are not explicitly represented in the canonical "Tree of Life," because they do not share common ancestry with cells. Rather, different types of viruses appear to have arisen from cells in multiple, independent events over the history of life on earth (Morse, 1994). The diversity among viral genomes, which may be single- or double-stranded RNA (ssRNA or dsRNA) or DNA (ssDNA or dsDNA) depending on the virus, is suggestive of multiple origins. This diversity in the nature of the genetic material is in stark contrast to all extant cell-based organisms. Despite the impressive morphological and metabolic diversity seen in the domains *Archaea*, *Bacteria*, and *Eukarya*, all have genomes composed of dsDNA. The term "virus," then, does not refer to an evolutionarily coherent group, but is better seen as a convenient name for a collection of replicating, evolving, biological entities that share some distinctive features in their lifestyle. Even among the viruses, however, there is variation in the mechanism of replication and the relationship between virus and host cell.

Because some viruses kill their host cell, they are sometimes referred to as predators. Given the nature of virus–host relationships, however, they are more appropriately classified along the spectrum of symbioses that has been defined for cellular organisms (Kuris and Lafferty, 2000). In one mode of infection, viruses replicate in a cell, releasing multiple progeny at the expense of the host. In some cases, the virus kills the

host in the manner of a parasitoid (**Figure 1(b)**). In other cases, viruses are produced, but are shed from a cell without killing it, making the virus a parasite. In another mode, the viral genome merges with the genome of the host cell that it has infected (**Figure 1(c)**). By becoming part of the genome, the virus is replicated passively whenever the cell divides and each daughter cell obtains a copy of the hybrid genome. The presence of viral DNA in the genome increases the replication burden of the host. If the presence of the viral DNA provides no compensatory benefits, the integrated viral genome is a (molecular) parasite. In many cases, however, genes encoded by the integrated virus are expressed, which alters the phenotype of the host cell in ways that are beneficial. This phenomenon, which is particularly well documented in pathogenic bacteria, is known as lysogenic conversion and represents a mutualism.

In life, symbiosis is the rule rather than the exception and it seems most probable that molecular parasites and virus-like agents were present even at the earliest stages of the origin of life (Koonin, 2009). Even some viruses, known as satellite viruses, have obligate symbiotic relationships with other viruses (Krupovič and Cvirkaite-Krupovič, 2011). Although viruses are frequently associated with negative consequences (disease and death), there is ample evidence that they are every bit as much a creative force in the evolution of cellular life. By virtue of their obligate, intracellular, symbiotic relationships with cells, viruses serve as a mechanism by which genetic diversity is introduced to a cell's genome. Viruses may alter a cell directly via lysogenization, but they also mediate gene exchange between organisms and gene duplication within organisms, thereby providing raw material for natural selection. Remnants of viral genes litter the genomes of cells and hint at the profound contribution of viruses to the expansion of biodiversity (Villarreal, 2004). Therefore, although viruses may not appear explicitly, their presence is implicit in the shape and branching patterns of the tree of life (Villareal and Witzany, 2010). In this article, viral diversity *per se* will be the focus, rather than the mechanisms by which viruses promote biodiversity in cellular organisms.

Classification of Viruses

Various classification schemes have been devised in attempts to impose some logical order on the extraordinary diversity of viruses. Earlier schemes include one in which viruses were all placed in the order *Virales* and were classified based strictly on the type of organism they infect (Holmes, 1948) or on the diseases they cause, but these criteria resulted in arbitrary and artificial distinctions and a taxonomy fundamentally at odds with evolutionary relationships. More satisfying is to distinguish viruses based on the properties of the virions themselves. One such system is that proposed by Lwoff *et al.* (1962) and referred to, accordingly, as the LHT system. A strictly sequence-based system, molecular taxonomy, was proposed as the most rational approach for classifying viruses (Lanni, 1964), but was criticized for its lack of practicality (Lwoff and Tournier, 1966), because sequencing and bioinformatics were poorly developed at the time. An alternative nucleic acid-centric system (Baltimore, 1971) classifies viruses based only

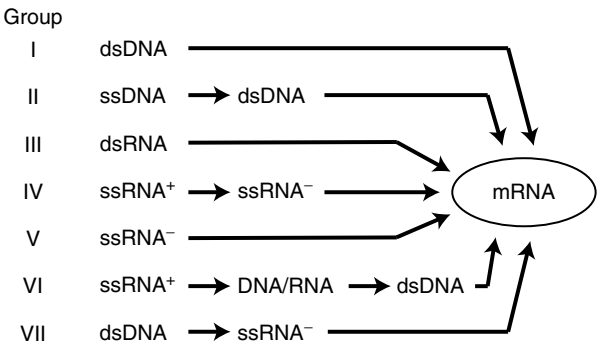


Figure 2 The Baltimore Classification scheme. Depending on the nature of a viral genome, it must go through different steps to be converted into messenger RNA (mRNA), which is required for the information to be translated by the protein-making machinery of the host cell. The nature of the genome and the conversions it goes through to result in mRNA is the basis for discriminating seven different types of viruses that are known as Baltimore Groups I–VII.

on the type of nucleic acid (DNA or RNA, single- or double-stranded) that makes up the viral genome and by the steps required to decode and replicate the genome. Regardless of nucleic acid type, all viruses need to produce mRNA that can be translated into protein. The steps by which this is achieved vary according to the nucleic acid type and, for single-stranded genomes, the sense in which the information is encoded. The modern Baltimore classification scheme distinguishes seven groups of viruses that account for these differences (**Figure 2**). This scheme is partially congruent with some of the higher order viral taxonomy as presented by the International Committee on Virus Taxonomy (ICTV), but with only seven categories it is clearly insufficient for detailed classifications. The LHT system provided finer level discrimination and, with modifications, served as the early basis for a comprehensive viral taxonomy. As sequencing has become more commonplace, molecular taxonomy as envisioned by Lanni has become the accepted basis for delineating taxonomic groups at the finest level. In this article the presentation of viral types is organized according to Baltimore Group, followed by the ICTV-approved orders, families, and genera.

The Scope of Virion Phenotypic Diversity

The collection of biological entities referred to as viruses is defined by certain shared characteristics as noted above, but the physical characteristics of virions vary with respect to their genome (e.g., chemical composition, size, and architecture) and the packaging of the genome (e.g., capsid shape, symmetry, surface projections, and presence or absence of an envelope). To provide a sense of the scope of this variability, some of the information drawn from the Ninth Report of the International Committee on Taxonomy of Viruses (King *et al.*, 2012) is summarized here. This summary includes those viruses classified to the genus level, and specifically excludes those with no known encapsidated stage.

In addition to the four physical configurations already mentioned (dsDNA, ssDNA, ssRNA, or dsRNA), viral genomes may also be distinguished based on their length, whether they

are circular or linear, and whether they are mono or multipartite. If multipartite, the segments may be packaged together in a single virion or they may be individually encapsidated. Monopartite genomes range from 1775 nucleotides (nt) of ssDNA for a member of the family *Circoviridae* to 1,259,197 base pairs (bp) of dsDNA for a member of the family *Mimiviridae*. The total genome length for multipartite viruses (sum of the length of all segments) ranges from 3100 bp dsRNA (in the family *Partitiviridae*) to $\geq$600,000 bp dsDNA (in the family *Polydnaviridae*). The smallest individual segment of a multipartite genome in which the segments are packaged together is reported to be only 200 bp dsRNA (family *Reoviridae*), but is more typically $>$600 nt or bp. For multipartite genomes in which the segments are packaged individually, the segments range from 923 nt ssDNA to 9800 nt ssRNA.

Describing all of the known viral morphologies in detail is beyond the scope of this article, but some examples are provided to illustrate the range of morphological types one may encounter (**Figure 3**). These include the familiar polyhedral (often icosahedral) capsids, which may be isometric or elongated, and may be enveloped or not. A number of viruses appear to be spherical, but have polyhedral symmetry in their construction from subunits. Among the viruses infecting *Bacteria* and *Archaea*, are those that have a polyhedral capsid

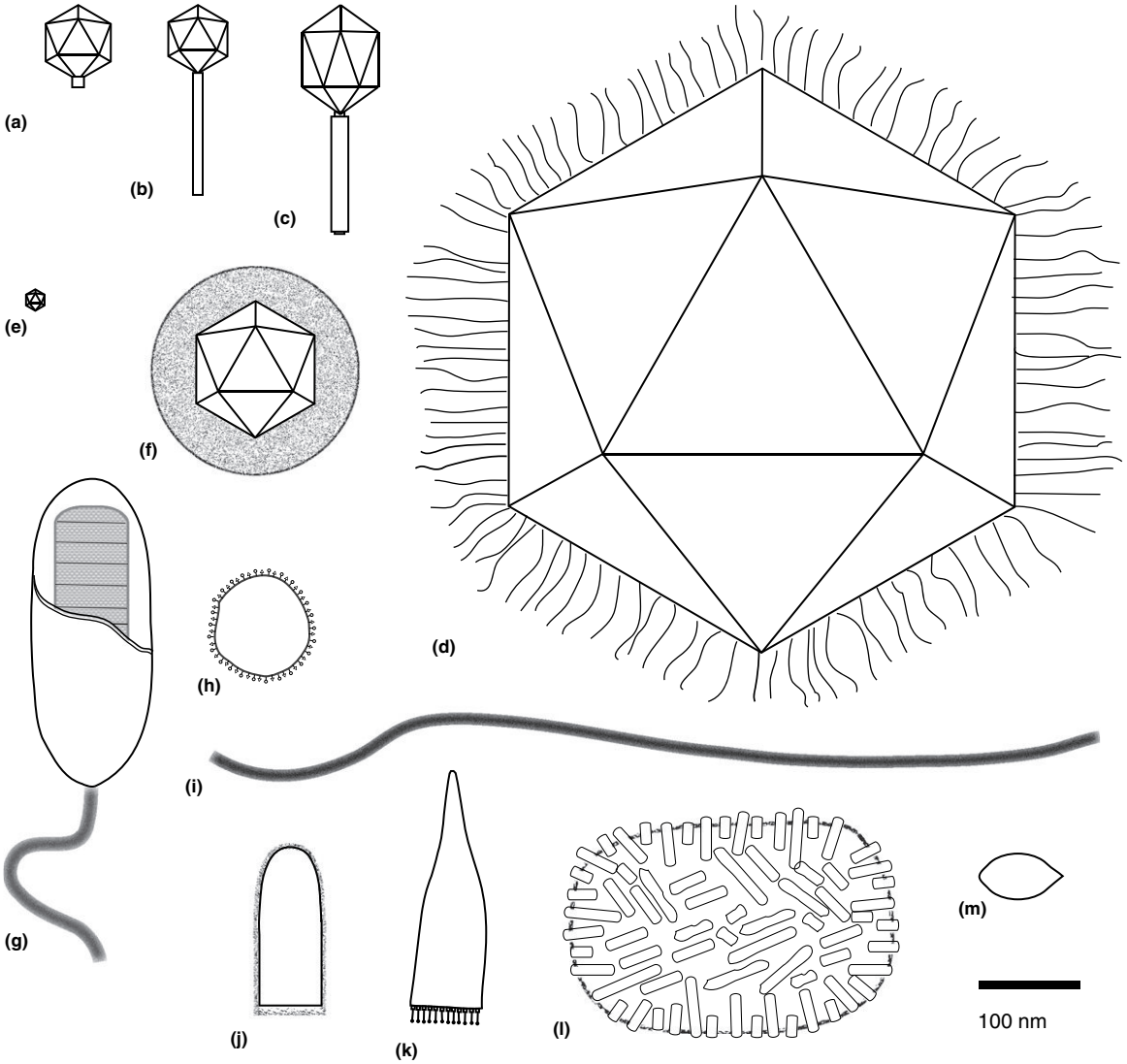


Figure 3 Examples of some the morphological diversity among viruses. (a)–(f) Show virions with capsids having icosahedral symmetry. (a)–(c) Illustrate the types of tails that are observed among virions in the order *Caudovirales* and (c) also illustrates a prolate icosahedron. (d) and (e) Illustrate the range in size observed among virions with isometric capsids and represent the largest and smallest viruses recorded. (f)–(h) and (j)–(m) represent viruses in which the capsid is enclosed in a membrane. (g) Shown in cut away to reveal the rod-shaped capsid within and (h) illustrates a pleomorphic virion. (h) Illustrates a filamentous virus with helical symmetry. (j)–(m) Illustrate some of the other unusual morphologies of virions. The taxa in which virions with these morphologies may be found are: (a) *Podoviridae*, (b) *Siphoviridae*, (c) *Myoviridae*, (d) *Mimiviridae*, (e) *Parvoviridae*, (f) *Herpesvirales*, (g) *Nimoviridae*, (h) *Paramyxoviridae* (also *Bunyaviridae*, *Arenaviridae*, and *Coronaviridae*), (i) *Inoviridae* (also *Closteroviridae*, *Alphaflexiviridae*, and *Betaflexiviridae*), (j) *Rhabdoviridae*, (k) *Ampullaviridae*, (l) *Poxviridae*, and (m) *Fuselloviridae*.

joined to a proteinaceous tail. The tail may be short and stout, long, and flexible, or long and contractile. Other viruses with helical symmetry take the form of filaments or rods. The presence of an external envelope (usually derived from membranes of the host cell) can result in a virion morphology that differs from the nucleocapsid within. These virions may be spheres, prolate spheres, or pleomorphic. Other viruses have shapes reminiscent of bullets, bricks, bottles, and lemons. In addition to the overall morphology, viruses can vary in the number and position of peplomers, or glycoprotein spikes, on the surface, which are involved in recognition of, and entry into, an appropriate host cell.

One hallmark of viruses is their exceptionally small size. An electron microscope is required to resolve the morphological details of viruses, but some (e.g., Acanthamoeba polyphaga mimivirus) are just large enough ($\geq$0.5 mm) to be discernible with a light microscope. Among those viruses having isometric capsids, diameters of individual virions range from 17 to 500 nm. The thinnest filamentous viruses (in the family *Ophioviridae*) have a diameter of approximately 3.5 nm, but can be up to 3200 nm long. Because of the disparate morphologies of viruses, it can be useful for comparative purposes to calculate equivalent spherical diameters from the estimated viral volumes. In the case of multipartite genomes,

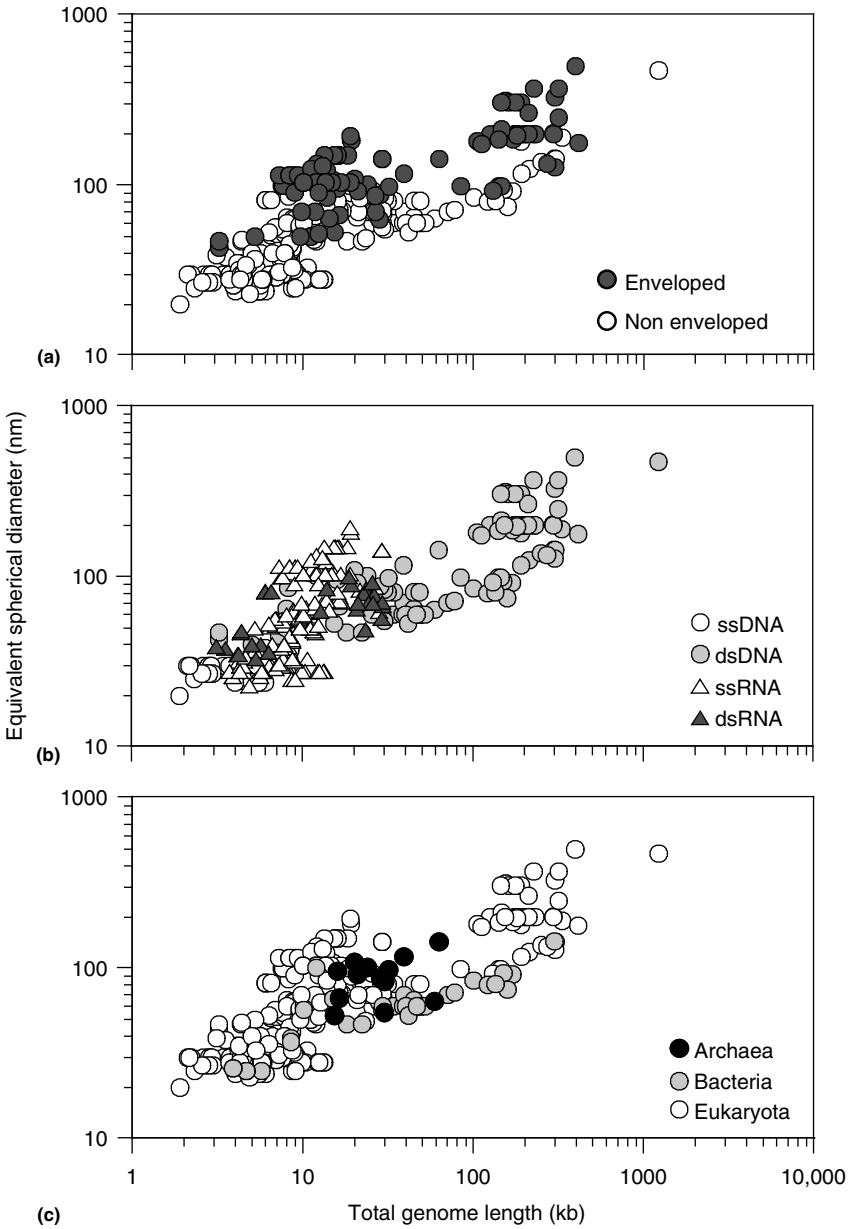


Figure 4 The relationship between genome length and capsid size. Average values of total genome length and equivalent spherical diameter were calculated for currently classified virus families using data reported in the Ninth Report of the International Committee on Virus Taxonomy (King *et al.*, 2012) and other literature data. The data are coded by whether the virions are enveloped or not (a), the type of nucleic acid contained in the virion (b) or the domain of life that serves as the host for the virus (c).

the total genome length is reported as the sum of the length of all segments. For viruses in which the segments of a multipartite genome are packaged separately, total virus volume to be the sum of the volumes of the set of virions that contain one complete set of genome segments has been considered. Equivalent spherical diameter (and thus virion volume) increases as a function of genome length regardless of the morphotype, but there is considerable scatter in the relationship (**Figure 4**). Envelopes, by adding an extra layer, tend to increase diameter of a virus for a given genome size (**Figure 4(a)**). Unenveloped, isometric viruses are noteworthy in that they represent both the smallest and the largest classified viruses both in terms of genome size and equivalent spherical diameter. The genome and capsid size distributions differ among viruses containing different types of nucleic acids, with ssDNA-containing viruses being the smallest, and dsDNA-containing viruses being the largest (**Figure 4(b)**). The size range also varies by host type (**Figure 4(c)**). Viruses infecting members of the domain *Archaea* have the narrowest range of capsid and genome size, while those infecting members of the domain *Eukaryota* have the widest range.

Viral infections have been reported in organisms representing the entire spectrum of life on the planet. In some cases, viral infections are only known from the observations of virus-like particles inside cells using an electron microscope, but viruses have also been isolated and characterized from members of all the domains and kingdoms of life. The number of different viruses confirmed to infect any given organism typically increases in proportion to the effort put into the search. Hundreds of human-infecting viruses have already been identified, for example, and scientists continue to discover more for our species. It thus appears that every species on earth is susceptible to infection by many distinct viruses. The phylogenetic breadth represented by the different types of organisms a virus can infect is referred to as its host range. Some viruses are able to infect many different hosts (broad host range), but often viruses appear to be rather specific (narrow host range). Host range is always construed in a relative sense, since the number and types of organisms tested is not standardized and a definitive host range determination is not empirically achievable. Despite the uncertainties stemming from variable and incompletely defined host ranges, the existing host range studies suggest that viruses are probably much more diverse than the entire spectrum of cellular life.

Viruses in the Sea

Viruses can be found wherever life is found and in seawater they outnumber *Bacteria*, *Archaea*, and *Eukarya* combined by an order of magnitude or more (Wommack and Colwell, 2000; Suttle, 2007). The concentration of virions in surface seawater in the open ocean is on the order of 5–10 billion per liter (**Figure 5**). Their concentration declines with depth, but viruses are present in the deepest waters sampled (>4000 m). The concentration of virions in surface marine sediments is typically even higher than in the immediate overlying waters. As in the water column, viruses become less abundant with depth in the sediment, but remain detectable to depths

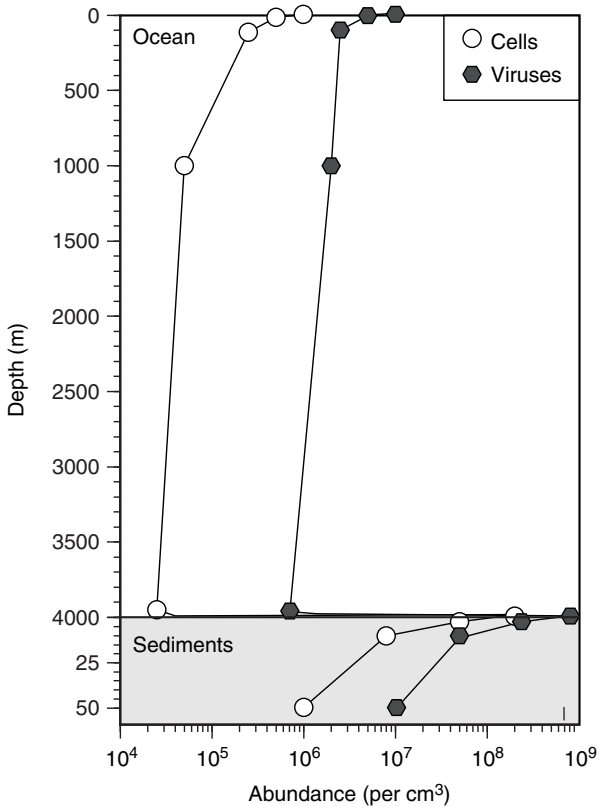


Figure 5 Idealized plot of the concentrations of viruses and cells in open ocean seawater and sediments. Sediments are indicated by the shaded area. Sediment depth (in bold) starts over from zero at the sediment–water interface and is plotted at a different scale for clarity. "Cells" refers to all cells, but these numbers are dominated by bacteria and archaea. The numbers shown are approximate and concentrations at a given depth can vary by an order of magnitude depending on location. Concentrations tend to be higher in areas of high productivity such as the coastal zone.

>100 m below the seafloor (Bird *et al.*, 2001) and in sediments with estimated ages of hundreds of thousands of years or more (Middelboe *et al.*, 2011).

Virions suspended in seawater are a component of the marine plankton and are therefore often referred to as virioplankton. Those in sediments are sometimes referred to as viriobenthos (Danovaro *et al.*, 2008). Plankton ecologists find it convenient at times to refer to plankton according to which of seven different size categories they belong. Because of their small size, most of the viruses in seawater are members of the smallest named size category, the femtoplankton (0.02–0.2 μm), but larger viruses are also part of the picoplankton (0.2–2 μm), a size class that also includes almost all of the bacteria and archaea and the smallest protists (**Figure 6**).

Viral infections have been reported for a broad range of marine organisms (Munn, 2006) and these infections have wide-ranging effects on everything from evolution, ecology, and ocean biogeochemistry (Fuhrman, 1999; Rohwer and Vega Thurber, 2007) to human health and economics. Viral infections of phytoplankton, the foundation of the marine

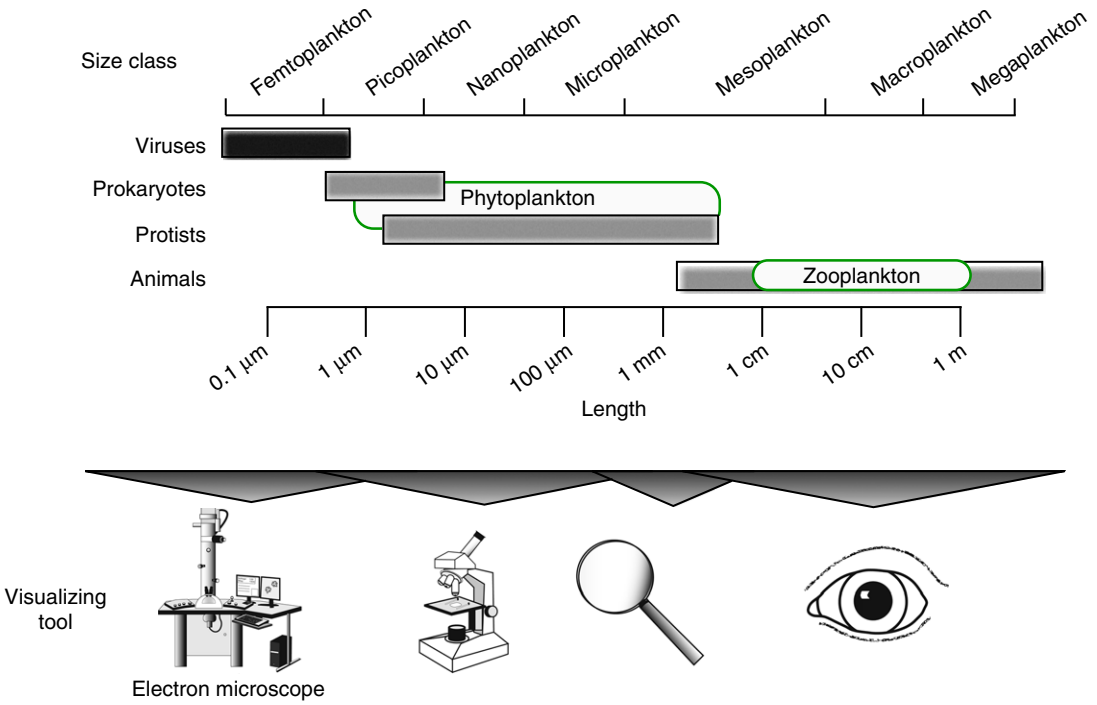


Figure 6 Viruses as members of the marine plankton. As a group, viruses are the smallest and most abundant members of the plankton and an electron microscope is required to reveal their details. Plankton are often named by the size category to which they belong, but may also be referred to in functional categories such as virioplankton, phytoplankton, and zooplankton.

food web, affect how much of the ocean's primary production is transferred to higher trophic levels (Brussaard, 2004). Large increases in phytoplankton concentrations (blooms) are a common phenomenon in the ocean and these periodic increases in primary production serve as a critical food supply for life in the sea. Viruses also spread more easily through populations at high density, however, and phytoplankton blooms can be decimated by the rapid spread of viral infection. Lysis of phytoplankton decreases the efficiency with which they are consumed by grazers and much of the primary production is instead decomposed by bacteria (Bratbak *et al.*, 1998). Bacteria are themselves susceptible to viral infections and viral lysis of bacteria likewise diverts material and energy from protistan grazers, which favors further growth and respiration by other uninfected bacteria (Middelboe *et al.*, 1996). It thus appears that the net effect of lytic viral activity in the marine food web is to decrease the amount of carbon that reaches higher trophic levels (e.g., fish) while increasing the amount of carbon, which is respired by the bacteria (Fuhrman, 1992; Murray and Eldridge, 1994).

From the perspective of humans, viruses could be lamented as reducing the efficiency of wild fisheries, but by how much is unknown. In any case, there is little to be done about the persistent, ubiquitous influence of viruses on the ocean's food web; it is an inescapable feature of the natural cycle of life in the sea. However, viruses are a legitimate human concern with tangible economic consequences when they cause death and disease among farmed marine species such as shrimp and fish. As wild fisheries become increasingly threatened by overexploitation, humans are relying more heavily on aqua- and mariculture as a source of protein. Intensive mariculture involves growth of single species at high densities, which, unfortunately, is a condition under which viral and other diseases can easily spread within the farm and from the farm to wild populations. The economic losses resulting from viral diseases in aquaculture are estimated to be in the billions of dollars annually (Walker and Winton, 2010) so there is a strong incentive to find solutions to combat viral diseases in commercially exploited marine species.

Life in the sea is very diverse, with representation of most of the major types of life on the planet. As one might then expect, the diversity of viruses in the sea is also broad (Breitbart *et al.*, 2007). Among viruses infecting the marine organisms, members of all Baltimore groups except for Group VII have been found so far (**Table 1**). The next section is a synopsis of the viruses belonging to these groups that have been isolated and characterized.

Taxonomic Diversity among Marine Viruses

This section is organized first by Baltimore Group, then by order (where one has been defined), then by family. Specific genera within a family are noted in some cases where it seems of particular interest. The taxonomic groups represented are, for the most part, restricted to those explicitly mentioned in the most recent report of the ICTV (King *et al.*, 2012), but some information is included from other sources and proposals that have been accepted since the report was completed. For the purposes of this presentation, a virus was

Table 1 Families and subfamilies of viruses for which marine isolates have been described

Group	*Genome*		*Taxonomy*			*Marine hosts*	
	Type	*Segments*	*Order*	*Family*	*Subfamily*	*Types*	*Names*
I	dsDNA	Mono	*Caudovirales*	*Myoviridae*	–	B	Cyanobacteria and proteobacteria
				Podoviridae	–	B	Cyanobacteria and proteobacteria
				Siphoviridae	–	B	Cyanobacteria and proteobacteria
			Herpesvirales	*Alloherpesviridae*	–	V	Fish
				Herpesviridae	*Alphaherpesvirinae*	V	Turtle
					Gammaherpesvirinae	V	Seal
				Malacoherpesviridae	–	I	Oyster
			Unassigned	*Adenoviridae*	–	V	Fish
				Corticoviridae	–	B	Bacteria
				Iridoviridae	–	V	Fish
				Mimiviridae	–	Pr	Heterotrophic protist
				Nimaviridae	–	I	Shrimp and crab
				Papillomaviridae	–	V	Porpoise and dolphin
				Phycodnaviridae	–	Pr	Micro- and macroalga
				Poxviridae	–	V	Marine mammals and birds
				Unassigned	–	Pr	Microalga
II	ssDNA	Mono	Unassigned	*Anelloviridae*	–	V	Sea lion
				Circoviridae	–	V	Herring gull
				Inoviridae	–	B	Proteobacteria
				Microviridae	*Gokushovirinae*	B	Proteobacteria
				Parvoviridae	–	I	Shrimp
				Unassigned	–	Pr	Microalga
III	dsRNA	Bi	Unassigned	*Birnaviridae*	–	I, V	Molluscs, crustaceans, and fish
		Multi	Unassigned	*Reoviridae*	*Spinareovirinae*	I, V	Bivalves and fish
					Sedoreovirinae	Pr, I	Microalga and crab
IV	ssRNA (+)	Mono	*Nidovirales*	*Roniviridae*	–	I	Shrimp
		Mono	*Picornavirales*	*Dicistroviridae*	–	I	Shrimp
				Marnaviridae	–	Pr	Microalga
				Unassigned	–	Pr	Labrynthulid and microalga
		Mono	Unassigned	*Astroviridae*	–	V	Sea lion and dolphin
				Caliciviridae	–	V	Seals and whales
		Bi		*Nodaviridae*	–	V	Fish
		Mono		*Togaviridae*	–	V	Fish and mammals
V	ssRNA (−)	Mono	*Mononegavirales*	*Paramyxoviridae*	*Paramyxovirinae*	V	Seals and whales
				Rhabdoviridae	–	V	Fish
		Multi	Unassigned	*Orthomyxoviridae*	–	V	Fish, sea birds, seals, and whales
VI	ssRNA (+)	Mono	Unassigned	*Metaviridae*	–	I, V	Sea urchin and fish

Viruses are presented by Baltimore Group, genome type, and taxonomy to the subfamily level along with the type (B, bacteria; Pr, protist; I, invertebrate; V, vertebrate) and common names of marine hosts.

considered marine if it infects an organism commonly referred to as marine. This includes organisms for which the entire life cycle is completed in the ocean, but also sea turtles (marine reptile), a variety of birds such as gulls, terns, cormorants, and pelicans (marine birds) as well as seals and sea lions (marine mammals). The general characteristics of the virions in each taxon and the types of marine organisms they infect are noted.

Group I: dsDNA Viruses

The dsDNA virus isolates known to infect marine organisms are assigned to 13 of the 30 families in Baltimore Group I. Three families are members of the order *Caudovirales*, another three are members of the order *Herpesvirales*, and six families are not assigned to an order. A number of viruses infecting marine plankton have also not yet been assigned to a family.

Order Caudovirales

The viral order *Caudovirales* contains three related families of tailed viruses, all of which infect bacteria or archaea. Members of this order have nonenveloped capsids with icosahedral symmetry, but these may be isometric or prolate with diameters ranging from around 40–145 nm in diameter. Helical tails (10–355 nm long and 5–18 nm wide) aid in attachment to the host cell and serve as a conduit through which the viral genome is delivered to the interior of the cell. The tail structure serves as a taxonomic character that divides the order into the families *Myoviridae*, *Podoviridae*, and *Siphoviridae* (**Figure 3(a)–3(c)**; Ackermann, 2006). Marine representatives can be found in all three of these families, but marine isolates reported to date are restricted to those infecting bacteria, probably because isolates of marine archaea remain so rare. Members of all three families are very common in seawater.

Family Myoviridae

Viruses in this family are characterized by having a long (50 to more than 200 nm) hollow tail with a neck and a contractile sheath. Because of the sheath, the tails of these viruses are thicker (16–18 nm) than those of other caudovirads (**Figure 3(c)**). Myovirids that infect marine bacteria, particularly bacteria in the phylum *Gammaproteobacteria*, have been classified, but many more marine isolates are known. These include a large number of isolates infecting cyanobacteria in the genera *Prochlorococcus*, and *Synechococcus* (Sullivan *et al.*, 2010). Although there is considerable overlap, the genomes and capsids of myovirids tend to be larger than those of the other two families in this order.

Family Podoviridae

Viruses in this family are characterized by the presence of a short, noncontractile tail that is 10–48 nm long and about 8 nm wide (**Figure 3(a)**). Podovirids appear to be the most numerous of all tailed phages in the marine environment. Podovirids infecting marine bacteria in the phyla *Alpha-* and *Gammaprotoebacteria* have been officially classified by the ICTV, but many more isolates infecting a range of bacteria exist, including podovirids infecting *Prochlorococcus* and *Synechococcus* (Sullivan *et al.*, 2003).

Family Siphoviridae

Viruses in this family have a long, noncontractile tail that is around 100 to >300 nm long, and 5–10 nm wide (**Figure 3(b)**). These long, unsheathed tails often appear to be quite flexible, displaying bends and curves in electron micrographs.

Order Herpesvirales

Members of the order *Herpesvirales* have a large icosahedral capsid enclosed in a tegument and a lipid bilayer membrane (**Figure 3(f)**). The genome ranges from 125,000 to 295,000 bp. Capsids are around 125 nm, but the entire virion is closer to 200 nm in diameter. These viruses are very widespread in the vertebrates. Examples of herpesvirads infecting marine vertebrates are known in all three families of the order (Davison *et al.*, 2009).

Family Alloherpesviridae

Members of this family include those that infect fish and amphibians. Marine and estuarine representatives are in the genus *Ictaluvirus* and include those that infect catfish, sturgeon, salmon, and cod (Hanson *et al.*, 2011).

Family Herpesviridae

This family consists of genetically related herpesvirids that infect reptiles, birds, and mammals. Marine representatives have been assigned to two of the three subfamilies (*Alpha-* and *Gammaherepesvirinae*), but are not yet assigned to a genus. Marine hosts include sea turtles, seals, sea lions, dolphins, and whales (Maness *et al.*, 2011).

Family Malacoherpesviridae

This new family consists of a single genus (*Ostreavirus*) and species (*Ostreid herpesvirus 1*). This virus infects oysters and is the only herpesvirad identified so far that infects an invertebrate.

Families Not Assigned to an Order

Family Adenoviridae

Members of this family are nonenveloped and have a capsid 70–90 nm in diameter with icosahedral symmetry and a genome between 25,000 and 50,000 bp. Adenovirids infect a wide variety of vertebrates. There are a number of reports of infections by adenovirids in marine mammals (pinnipeds and cetaceans) and in gulls, but none have been isolated and officially classified. Sea lions may be a host for canine adenovirus. The only reported fish adenovirus is in the genus *Ichtadenovirus* with the sole representative being *Sturgeon adenovirus A*, which infects white sturgeon. Although not truly marine, the sturgeon habitat includes brackish coastal waters.

Family Corticoviridae

This family is represented by a single genus (*Corticovirus*) and a single species (*Pseudoalteromonas phage PM2*) that infects a marine bacterium (Espejo and Canelo, 1968). The virion has an icosahedral capsid 57 nm in diameter, no tail, and a lipid bilayer membrane interior to the capsid (Harrison *et al.*, 1971). The genome in the virion is supercoiled with a length of 10,079 bp (Männistö *et al.*, 1999). Although only one species has been classified at present, related sequences have been detected within the genomes of a variety of bacteria suggesting that they could be prevalent in the marine environment (Krupovič and Bamford, 2007).

Family Iridoviridae

Members of this family have a large icosahedral capsid (120–200 nm diameter) and a large genome (around 100,000 to 200,000 bp). They infect invertebrates (arthropods) and vertebrates (amphibians, reptiles, and fish). Those infecting invertebrates are nonenveloped while those infecting vertebrates acquire a lipid envelope that derives from the host cell membrane during virion budding. Members of two of the five genera in this family (*Lymphocystivirus* and *Megalocytivirus*) infect marine fish.

Family Mimiviridae

This is a relatively new family of viruses, the members of which have unusually large genomes and icosahedral capsids. The first representative to be fully sequenced, *Acanthamoeba polyphaga mimivirus* (APMV) infects a freshwater amoeba (Scola *et al.*, 2003). At the time of its discovery, it was the largest known virus with a genome of 1,181,549 bp and a capsid diameter of 500 nm (**Figure 3(d)**). The virus is densely covered with hair-like fibrils that extend its effective diameter to 750 nm. Recently, another virus related to APMV with a larger genome (1,259,157 bp), but smaller capsid (440 nm) was isolated from a marine sample using several freshwater *Acanthamoeba* species as hosts (Arslan *et al.*, 2011). The natural host of this new isolate, named *Megavirus chilensis*, is unknown, but is presumed to be a marine or estuarine protist. This is the largest viral genome known at the present time.

Family Nimaviridae

This family consists at present of a single genus (*Whispovirus*) and species (*White spot syndrome virus*). Virions consist of a rod-shaped nucleocapsid (approximately 65–70 × 300–350 nm) enclosed in a tegument and a trilaminar envelope. The overall

virion dimensions are 120–150 nm in diameter and 270–290 nm in length. The virion is unusual in having a thread-like appendage that gives it the superficial appearance of a bacterium with a polar flagellum (**Figure 3(g)**).

Family Phycodnaviridae

The members of this family that have been officially classified so far have large icosahedral capsids (around 100–200 nm diameter) with no outer envelope and genome sizes ranging from approximately 100,000 to 500,000 bp. These viruses infect photosynthetic protists (micro and macroalgae) in both freshwater (genus *Chlorovirus*) and the marine environment (genera *Coccolithovirus*, *Phaeovirus*, *Prasinoviruses*, *Prymnesioviruses*, and *Raphidoviruses*). Members of the genus *Phaeovirus* infect filamentous brown algae, but all others infect unicellular plankton.

Family Poxviridae

The capsids of poxvirids are enveloped and brick shaped or ovoid with tubular surface processes (**Figure 3(l)**). They vary from 200 to 450 nm in length, 140–280 nm in width, and 140–280 nm in thickness. The genomes are linear dsDNA ranging from 130,000 to 380,000 bp. Members of this family infect a wide range of hosts. The poxvirids infecting vertebrates are in the subfamily *Chordopoxvirinae*, which contains nine genera. The subfamily *Entomopoxivirinae* comprises three genera, the members of which infect various insects. The family is well known for the serious human disease caused by one of its notorious members, small pox (now eradicated). In the marine realm, poxviruses have been identified in penguins, seals, sea lions, dolphins, and whales, but none of these have been officially classified.

Unassigned

Several new viruses infecting marine heterotrophic and phototrophic protists have been isolated and described, for which the taxonomic positions remain unclear. A large (300 nm) virus, CroV, that infects *Cafeteria roenbergensis* (Garza and Suttle, 1995) is the first isolate that infects a phagotrophic marine protist. Based on the large genome size (750,000 bp) and specific sequence of the DNA polymerase of CroV, it may belong to the family *Mimiviridae* (Fischer *et al.*, 2010). A number of other virus isolates infecting photosynthetic marine protists (*Chrysochromulina ericina* virus 01B, *Phaeocystis pouchetii* virus 01, and *Pyramimonas orientalis* virus 01B) have capsids closer in size to phycodnavirids (140–220 nm), but their DNA polymerase genes place them closer to APMV and CroV, suggesting they may also be members of the *Mimiviridae*.

A large virus (200 nm diameter, 350,000 bp genome) was isolated that infects the bivalve-killing dinoflagellate, *Heterocapsa circularisquama* (Nagasaki *et al.*, 2003). The DNA polymerase sequence of this isolate is very similar to that of African swine fever viruses in the genus *Asfivirus*, which is the sole genus in the family *Asfarviridae*. It is, therefore, highly divergent from other dsDNA viruses infecting phytoplankton and a new genus name, *Dinodnavirus*, has been proposed for this virus.

Group II: ssDNA Viruses

Families Not Assigned to an Order

Family Anelloviridae

Members of this family have a nonenveloped icosahedral capsid approximately 30 nm in size and a circular genome ranging in size from 2000 to 3900 nt. They appear to be widespread in mammals, but their involvement in specific diseases remains unproven. Novel anelloviridis have recently been detected by metagenomic analyses of lung tissue from seals and sea lions (Ng *et al.*, 2009, 2011), but these have yet to be approved as species.

Family Circoviridae

This family is comprised of two genera (*Circovirus* and *Gyrovirus*) known to infect birds and swine, although distantly related viruses have been described in primates. The virions are nonenveloped and appear spherical with icosahedral symmetry. The capsid diameter is 20.5–25 nm and the genome ranges from 1775 to 2300 nt. The only classified member so far that might be considered marine is the gull circovirus in the genus *Circovirus*.

Family Inoviridae

Virions in this family are filaments (**Figure 3(i)**) or rods with helical symmetry with a circular genome ranging from 4500 to 12,400 nt in length. Those in the genus *Inovirus* are long (700–900 nm) and thin (7 nm) and infect gammaproteobacteria, while those in the genus *Plectrovirus* (10–15 nm wide and 70–280 nm long) infect Gram-positive bacteria. Among the known hosts of inovirids are members of the bacterial genus *Vibrio*, which are common in coastal estuarine and marine waters. Inovirids cause persistent infections and are responsible for lysogenic conversion. The inovirid, Vibrio cholera CTXø, for example, codes for an enterotoxin that converts *Vibrio cholerae* into a human pathogen.

Family Microviridae

Virions in this family have a small, nonenveloped capsid approximately 25–27 nm in diameter and a circular, positive-sense genome ranging in size from 5300 to 6100 nt (genus Microvirus) or from 4400 to 4900 nt (genera *Chlamydiamicrovirus*, *Bdellomicrovirus*, and *Spiromicrovirus* in the subfamily *Gokushovirinae*). The hosts for members of the genus *Bdellomicrovirus*, are bacteria in the genus *Bdellovibrio*. *Bdellovibrio* spp. are predatory bacteria found in a range of freshwater, estuarine, and marine habitats. All strains of bdellomicroviruses isolated so far have been from freshwater. However, there is evidence from metagenomics that additional members of the *Gokushovirinae* are widespread in the ocean (Angly *et al.*, 2006; Tucker *et al.*, 2011). Sequence comparisons indicate that these viruses are distantly related to known chlamydiamicroviruses and bdellomicroviruses, but their hosts have not yet been identified.

Unassigned

Several genetically similar ssDNA viruses infecting diatoms in the genus *Chaetoceros* have been isolated. The virions are nonenveloped and have an icosahedral capsid ranging in size from 30 to 38 nm and a genome around 6000 nt in length.

Although these viruses do not yet belong to an official family, a new genus, *Bacilladnavirus*, has been approved to accommodate this group of viruses.

Group III: dsRNA Viruses

Families Not Assigned to an Order

Family Birnaviridae

There are four genera in the family *Birnaviridae*. The icosahedral particles of birnavirids are 65 nm in diameter, nonenveloped and are composed of a single capsid protein. The genome consists of two linear segments of dsRNA approximately 5900–6900 bp in total length that encodes a polymerase and a large structural polyprotein that is cleaved post-translationally. Birnavirids infect rotifers, insects, fish, and birds and are globally distributed. In the marine environment, aquabirnaviruses cause disease in salmonids, amberjack (yellowtail), bream, and cod among other fishes, and even molluscs and crustaceans.

Family Reoviridae

The family *Reoviridae* has fifteen genera divided among two subfamilies, the *Spinareovirinae* that include viruses that have turret-like proteins projecting from their capsid, and the *Sedoreovirinae*, the reovirids without turrets. The virions in this family are nonenveloped icosahedrons that range from 49 to 100 nm in diameter. Reovirids are distinctive in that their capsid can be composed of as many as three concentric protein shells that surround from 9 to 12 segments of dsRNA, that together range between 18,500 and 33,300 bp in total length. The viruses in this family infect a wide range of multicellular organisms including protists, fungi, plants, insects, reptiles, and humans. The dsRNA virus Micromonas pusilla reovirus (MpRV) from the genus *Mimireovirus* was one of the first marine RNA viruses isolated and it infects a marine protist (Brussaard *et al.*, 2004). The host, *M. pusilla*, is a small ($<2\,\mu m$), autotrophic flagellate considered the most abundant picoeukaryote in the global oceans (Not *et al.*, 2004). Species of bivalves and finfish are subject to infection from reovirids in the genus *Aquareovirus*, while orbiviruses have been found in seabirds (e.g., murres, puffins, and razorbills) and cardoreoviruses in crabs.

Group IV: Positive-Sense ssRNA Viruses

Order Nidovirales

The order *Nidovirales* is divided into three families, two subfamilies and seven genera. Virions in this order range widely in size from 50 to 200 nm, but all are enveloped. The genome of a nidovirad is a single, linear molecule of RNA that ranges in size from 12,700 to 29,100 bp and has a 5′ cap and a poly(A) tail that extends from the 3′ end. The gene order is highly conserved, with open reading frames (ORFs) nearest the 5′ end encoding the replicase, followed by 3–12 structural proteins. The genomes of viruses in the *Nidovirales* contain a characteristic -1 frameshift that is involved in the regulation of viral protein production during replication. Members of this order are the largest RNA viruses known to date, both in genome and particle size.

Family Roniviridae

The genus *Okavirus* is the only genus in the family *Roniviridae*. The virions of ronivirids are enveloped, bacilliform in shape, vary in size from 150 to 200 nm, and are studded with spike-like glycoprotein projections that have bulbous distal ends. The genome of these viruses has features typical of the order *Nidovirales* and can be quite large for an RNA virus, exceeding 26,000 nt in length. The genome configuration, particularly in the structural genes, as well as the absence of a discrete structural membrane protein gene, distinguishes ronivirads from other viruses in the order. Viruses in the family *Roniviridae* infect the giant tiger shrimp and the blue shrimp.

Order Picornavirales

The sprawling order of the *Picornavirales* contains five families, two subfamilies, 24 genera, and approximately 120 species of viruses. Picornavirads are "picorna-like" in that they have similar features to viruses in the family *Picornaviridae*. Members of the order have a nonenveloped icosahedral capsid that ranges in diameter from 22 to 30 nm. The linear, ssRNA genome is 7000–15,500 nt in length and encodes a polyprotein that is post-translationally processed. Picornavirads also have in common a conserved nonstructural gene order and a small protein covalently linked to the 5′ end and a 3′ end poly(A) tail. These viruses infect a wide range of eukaryotes, including humans. *Picornavirads* of note include human enterovirus C, which causes poliomyelitis, human rhinovirus, responsible for the common cold, and Israeli acute paralysis virus, which has been linked to honey bee colony collapse disorder.

Family Dicistroviridae

The family *Dicistroviridae* has two recognized genera, the cripaviruses and the aparaviruses. The particles of dicistrovirids are nonenveloped icosahedra approximately 30 nm in size. The genome is a single molecule of linear RNA that can be from 8500 to 10,000 nt in length and encodes two polyproteins, or cistrons (*ergo* the name of the family). Dicistrovirids infect only invertebrates, including aphids, bees, and butterflies. In the marine environment, Taura syndrome virus (TSV), an aparavirus, infects penaeid shrimp. TSV epidemics have been responsible for high rates of mortality in shrimp farms.

Family Marnaviridae

Heterosigma akashiwo RNA virus (HaRNAV) is the only member of the *Marnaviridae*. *Heterosigma akashiwo*, the host of HaRNAV, is a small, motile photosynthetic protist that is a common member of the temperate coastal phytoplankton community. During times when surface waters are stratified as a result of higher sea surface temperatures and low winds, *Heterosigma akashiwo* is capable of forming immense blooms that can cause extensive fish kills. These toxic blooms affect the fish farming industry in the fjords of western Canada, and Scandanavia in particular (Smayda, 1998). In 2003, Tai and colleagues isolated HaRNAV, the first RNA virus reported to infect a protist (Tai *et al.*, 2003). HaRNAV has a nonenveloped, icosahedral capsid that is approximately 25 nm in diameter. The burst size of HaRNAV is estimated to be 20,000 viruses per cell (Lawrence et al., 2006). The linear ssRNA genome of HaRNAV is 8600 nt in length and is monocistronic – unlike

AuRNAV (*see* Proposed Family *Labyrnaviridae*), the bacillarnaviruses, TSV (*see* Family *Dicistroviridae*), and HcRNAV (*see* Proposed Family *Alvernaviridae*). However, the HaRNAV genome has the same gene order as the other marine ssRNA viruses with the nonstructural genes nearest the 5′ end of the genome followed by the structural genes. Phylogenies based on the conserved domain sequences from HaRNAV and the corresponding regions of other viruses place HaRNAV outside previously characterized virus taxa within the order *Picornavirales*.

Proposed genus Bacillarnavirus

The genus *Bacillarnavirus* has not been assigned to a family within the order *Picornavirales*. It has three recognized members thus far, Rhizosolenia setigera RNA virus (RsetRNAV), Chaetoceros tenuissimus Meunier RNA virus (CtenRNAV), and Chaetoceros socialis f. radians RNA virus (CsfrRNAV) all of which infect marine diatoms. Diatoms are widely distributed, highly diverse, and are among the most abundant eukaryotes in the ocean. They form a large fraction of the total marine primary producers and are thus significant contributors to the cycling of energy in the ocean (Graham and Wilcox, 2000). The hosts of CtenRNAV and CsfrRNAV are members of the genus *Chaetoceros*, a group of chain-forming diatoms that can form blooms associated with fish kills. The two viruses have icosahedral symmetry and have ssRNA genomes that are dicistronic and approximately 9500 nt in length. CtenRNAV has an estimated capsid diameter of 31 nm (Shirai *et al.*, 2008), while that of CsfrRNAV is only 22 nm (Tomaru *et al.*, 2009). A striking difference between the two viruses is that CtenRNAV has an estimated burst size of 10,000 virions while CsfrRNAV produces approximately 66 per lytic event. RsetRNAV was the first virus brought into culture that infects a diatom. Like CtenRNAV and CsfrRNAV, RsetRNAV has an icosahedral capsid (approximately 32 nm in diameter), replicates in the cytoplasm and has a positive-sense ssRNA dicistronic genome approximately 8900 long (Nagasaki *et al.*, 2004). The genomes of all three RNA viruses have a similar gene order, with the structural genes nearest the 5′ end, followed by the helicase, protease, and replicase and terminating in a poly(A) tail. Alignments of the replicase genes of these viruses clearly demonstrate that they are closely related and form a well-supported monophyletic clade (Tomaru *et al.*, 2009).

Proposed Family Labyrnaviridae

Aurantiochytrium RNA virus (AuRNAV) is the only member of the family *Labyrnaviridae*. AuRNAV infects the thraustochytrid *Aurantiochytrium* sp. Thraustochytrids are single-celled marine organisms that were once thought to be fungi, but are now recognized as a class of protists. These organism are capable of feeding directly through absorption and are involved in the degradation of detritus (Raghukumar, 2002). Like HcRNAV (*see* Proposed Family *Alvernaviridae*) and the bacillarnaviruses, AuRNAV has a small (roughly 25 nm), nonenveloped icosahedral capsid. AuRNAV replicates in the cytoplasm and has a phenomenal burst size of approximately 60,000 particles per lytic event (Takao *et al.*, 2005). The genome of AuRNAV is a single molecule of RNA 9000 nt in length with a poly(A) tail. The three ORFs of the AuRNAV genome encode the capsid proteins, nonstructural genes and a protein of unknown function (Takao *et al.*, 2006). Interestingly, AuRNAV produces subgenomic RNA during replication. Although the specific mechanism is unknown, this suggests that the AuRNAV genome does not possess an internal ribosome entry site structure that is characteristic of other picornavirads.

Families Not Assigned to an Order

Proposed Family Alvernaviridae

The family *Alvernaviridae* has a single representative at present, Heterocapsa circularisquama RNA virus (HcRNAV), which infects the eponymous dinoflagellate. The dinoflagellates are a diverse group of unicellular flagellates that are capable of deriving energy in a number of different ways, including through parasitism, photosynthesis, and predation. Dinoflagellates are found in marine and freshwater, sediments, snow and ice, and even inside other organisms. They are an important food source for plankton, a critical component of coral reefs, and can periodically form potent harmful algal blooms (Graham and Wilcox, 2000). *Heterocapsa circularisquama* is an armored, free-living protist that can form extensive blooms and these blooms have resulted in severe financial losses to the aquaculture industry (Matsuyama, 1999). HcRNAV particles do not have an envelope and are approximately 30 nm in diameter. The linear RNA genome of this virus is only 4400 nt long (Tomaru *et al.*, 2004). The two ORFs of the HcRNAV genome encode a protease, a replicase, and a single capsid protein (Nagasaki *et al.*, 2005). During the final stages of HcRNAV replication, the virus forms tightly packed crystalline arrays within the cell that result in the release of an estimated 20,000 viral particles during lysis (Tomaru *et al.*, 2004).

Family Astroviridae

The family *Astroviridae* includes two genera. The capsids of astrovirids have nonenveloped, spherical particles that range in size from 28 to 30 nm and are covered in mushroom-like projections involved in virus attachment and entry. The genome of these viruses is a single linear molecule of RNA that is generally from 6400 to 7700 nt long and terminates in a poly(A) tail. Like the caliciviridis, the astrovirids produce a subgenomic RNA that encodes the structural proteins. Astrovirid infections cause gastroenteritis in a variety of birds and mammals, including humans. In the marine environment, members of the *Mamastrovirus* genus infect Bottlenose dolphins, and California and Steller sea lions.

Family Caliciviridae

There are five genera of viruses within the family *Caliciviridae*. Viruses in this family produce nonenveloped, icosahedral virions 27–40 nm in diameter that have distinctive cup-shaped depressions. The genome of these viruses is linear, ranges from 7300 to 8300 nt long and has a poly(A) tail. Unlike some other taxa of group IV viruses, the caliciviridis make a subgenomic RNA that encodes the structural genes during replication. Humans, pigs, rodents, and cats, among other vertebrates are all infected by calcivirids. Viruses in the genus *Vesivirus* have been found to infect diverse marine mammals, including sea lions, cetaceans, and walruses. Vesiviruses exhibit a broad cell tropism and a broad host range. For example, vesiviruses isolated from marine mammals are capable of infecting pigs and humans (Green, 2007).

Family Nodaviridae

There are two genera in the family *Nodaviridae*. Viruses in this taxa are generally spherical in shape, (from 25 to 33 nm in diameter) and lack an envelope and discernible surface structure. The genome of nodavirids consists of two molecules of RNA (approximately 4500 nt in total length) that have 5′ caps and lack poly(A) tails. Nodavirids are capable of infecting a variety of insect species including two species of *Aedes* mosquitos, in which they cause paralysis and death. Viruses in the genus *Betanodavirus* infect and cause disease in marine fish. Examples include flounder, grouper, and pufferfish, among others. In addition, several nodavirid-like viruses not yet approved by the ICTV infect prawns.

Family Togaviridae

This family comprises two genera, the alphaviruses and rubiviruses. While the *Alphavirus* genus has approximately 40 recognized members, the Rubella virus is the only member of the *Rubivirus* genus. The particles of these viruses are enveloped, spherical in shape and have an approximate diameter of 70 nm. The genome of a togavirid is a single strand of RNA that is from 7700 to 11,800 nt in length and has a 5′ cap and a poly(A) tail. While most togavirids are transmitted via an arthropod intermediary, transmission of the alphaviruses that infect marine mammals and fishes do not appear to require a vector. Alphaviruses have been isolated that infect elephant seals, Atlantic salmon, and rainbow trout.

Group V: Negative-Sense ssRNA Viruses

Order Mononegavirales

The *Mononegavirales* is a large order of negative-sense single-stranded viruses that includes four families, two subfamilies, and 16 genera. Although the morphology of these viruses can vary widely (ranging from 80 to 805 nm in size), all virions are enveloped and most possess spike-like structures that are involved in attachment to and fusion with the host cell. The genomes of mononegavirads are linear and nonsegmented, range from 8900 to 19,100 kb in size, lack a 5′ cap and a poly(A) tail, have a characteristic gene order (wherein from the 3′ end, core protein genes are followed by envelope genes and then the polymerase), and encode an RNA-dependent RNA polymerase whose sequence is highly conserved. During replication, each cistron is transcribed utilizing a stop-start mechanism starting at a single promotor region located on the 3′ end. The *Mononegavirales* is a taxon of particular relevance to human health, because it includes virulent human pathogens such as the Ebola virus, rabies virus, and measles and mumps viruses, but also includes many representatives infecting marine organisms.

Family Paramyxoviridae

The paramyxovirids are viruses of vertebrates that are responsible for notorious diseases in humans and animals such as measles, mumps, pneumonia, rinderpest, and Newcastle disease. The family is organized into two subfamilies comprised of seven genera. The virions of paramyxovirids are spherical in shape and approximately 150 nm in diameter. The genome of these viruses is a single molecule of negative-sense RNA that can be from 13,300 to 18,200 kb in length. All genomes of paramyxovirids encode a protein, creatively named "F," which causes the fusion of the virus and cell membranes at neutral pH. An interesting characteristic of the replication of paramyxovirids is the "rule of six." This rule alludes to the fact that the efficiency of RNA replication in these viruses is severely diminished if the number of nucleotides in the RNA template being replicated is not exactly a multiple of six. It is believed that the precise packaging requirements of the virus drive this surprising requirement (Lamb and Parks, 2007). The Cetacean morbillivirus and the Phocine distemper virus are species in the genus *Morbillivirus* that infect marine mammals, while the Newcastle disease virus has been isolated from the common murre (Lang *et al.*, 2009).

Family Rhabdoviridae

More than 200 different strains of rhabdovirids have been isolated from plants, invertebrates and vertebrates. The most distinctive feature of this family of viruses is that they all possess a bullet or rod-like morphology (Rhabdo from Latin, *rhabdos*, "rod") that can be from 100 to 430 nm in size (**Figure 3(j)**). The genome of a rhabdovirid is a single molecule of negative-sense RNA that can be from 11,000 to 14,900 kb in length. Viruses in the genus *Novirhabdovirus* infect diverse fishes, including salmon, flounder, and eels. Two novirhabdoviruses of note, infectious hematopoietic necrosis virus and viral hemorrhagic septicemia virus, are responsible for the mortality of significant numbers of wild and farm-raised salmon each year (Lang *et al.*, 2009). The rhabdovirid Dolphin rhabdovirus-like virus was isolated from a white-beaked whale in the North Sea (Bressem *et al.*, 1999).

Families Not Assigned to an Order

Family Orthomyxoviridae

The virions of orthomyxovirids are enveloped, spherical in shape, and range in diameter from 80 to 120 nm (e.g., **Figure 3(h)**). They contain a negative-sense RNA genome composed of six to eight segments of different length that are in total from 10,000 to 14,700 kb in size. Because of their ubiquity and virulence, orthomyxovirids have been intensely studied since the early twentieth century. This research has resulted in an unprecedented understanding of the mechanisms of zoonosis and evolutionary history of these viruses. Analyses of genomic data suggest that wild aquatic birds may be the primary reservoir of the orthomyxovirids that now infect a wide range of vertebrates, including the human-infecting H1N1 strain responsible for the worldwide human pandemic of 1918 (Palese and Shaw, 2007). Of the five genera in the family, the genera *Isavirus* and *Influenzavirus A* have viruses that infect fishes and marine mammals.

Family Bunyaviridae

The family *Bunyaviridae* comprises five genera whose viruses infect vertebrates and plants primarily via an arthropod vector. Among other distinguishing features, bunyavirids have a genome comprised of three unique molecules of negative or ambisense ssRNA, that range in total size from 11,400 to 19,200 kb, that encodes four structural proteins. The particles of these viruses are pleomorphic and range in size from 80 to

120 nm (e.g., **Figure 3(h)**). The terminal sequences of the three genome segments are highly conserved among bunyavirids and play an important role during viral replication. Most bunyavirids are capable of replicating in both arthropods, where little cytopathogenicity occurs, and vertebrates, where they are cytolytic (Plyusnin *et al.*, 2012). Mourilyan virus is a bunya-like virus that infects two penaied prawns. Additional bunyavirids have been isolated that infect seabirds, including murres and puffins.

Group VI: Positive-Sense ssRNA Viruses that Replicate through a DNA Intermediate

Family Metaviridae

Metavirids blur the line between virus and retrotransposon in that only one of the species of virus in this family produces infectious extracellular particles (these virions are irregular in shape and range in size from 50 to 100 nm). Oddly, encapsidated particles are produced within the cytoplasm of the host cell and in most cases do not exit. Nevertheless, viruses in this family encode the hallmark genes (major capsid protein, reverse transcriptase, and integrase), share the same gene order, and have a similar genome length to retroviruses (5200–7600 kb). There are two known marine metavirids that infect the same species of pufferfish, the Takifugu rubripes Sushi virus (TruSusV) and Fugu rubripes Suzu virus (FruSuzV), classified in the genera *Metavirus* and *Semotivirus*, respectively.

Phenotypic and Genotypic Diversity of Marine Virioplankton

Detailed characterizations of cultivated viruses have been essential for developing a coherent viral taxonomy. This structure has brought order to the bewildering complexity of the viral world, but viral diversity as viewed through the lens of taxonomy has its limitations. The stringent requirements for establishing official taxonomic classifications mean that the process is slow and deliberate. The rate at which novel viruses can be identified far outstrips the rate at which they can be properly classified. Taxonomy also does not concern itself with relative abundance or biogeography, which is an issue of interest to ecologists. Microbial ecologists investigating viruses in the ocean have therefore been somewhat less concerned with taxonomic classifications than with practical measures of diversity. Some studies rely on cultivation (e.g., establishing host ranges), but many others have focused on quantifying abundance and diversity (phenotypic and genotypic) using cultivation-independent methods.

Host Range

The low-throughput of cultivation is a bottleneck in the characterization of viral communities, but the ability to propagate a virus in the lab allows one to collect ecologically relevant information that cannot be obtained any other way. One can determine, for example, in how many different types of cells the virus can replicate. Investigations of viruses with different host ranges can provide insight into the fitness costs that must necessarily be incurred to offset the benefits of being able to infect a broader range of cell types. Although viruses are often classified as having a broad or narrow host range, such a semiquantitative conclusion is necessarily conditional, since not every possible host can be screened. Despite that limitation, host range analysis has been very enlightening, because it can be exquisitely sensitive in discriminating between strains and it does not require sophisticated or expensive technology. To perform this analysis, each viral isolate is tested for its ability to lyse a suite of organisms. Since viruses tend to be more or less host-specific, the suite of organisms tested should primarily consist of different strains of the same species or different species in the same genus, but testing more distantly related organism allows one to identify viruses with particularly broad host range. The more organisms one tests, the greater one's chance of discriminating between closely related viruses. Use of this technique with marine bacteria (Moebus and Nattkemper, 1981; Wichels *et al.*, 1998) has shown that the number of viruses infecting any given genus of marine bacterium is extraordinarily large.

Morphology

The subnanometer resolving power of the transmission electron microscope (TEM) allows one to discriminate fine structural details of individual virions. This power has been exploited to conduct qualitative surveys of marine viral diversity (e.g., Frank and Moebus, 1987) that hint at an extraordinary richness. Electron microscopy has been used for quantitative, comparative assessments of diversity in natural communities as well, but the resolving power cannot be fully exploited in this context for technical and economic reasons. Negative staining, which reveals the greatest detail of virions, is difficult to control and best suited to qualitative rather than quantitative assessments of diversity. Even with the greater uniformity provided by positive staining, it is difficult to confidently discriminate similar viruses in a mixed assemblage. A single virus may appear different depending on its orientation or as a result of damage or distortion during preparation. As a consequence of these limitations, quantitative surveys of morphological diversity have relied on classification into a handful of broad categories (e.g., <30, 30 to <60, 60 to <80, 80 to <100, and ≥ 100 nm). The results of TEM surveys indicate virus-like particles in the ocean are usually numerically dominated by those with capsid diameters in the range of 30–60 nm, but range in size from around 20 to 750 nm (Gowing, 1993; Weinbauer, 2004; Børsheim, 1993). Viruses in the ocean thus span a similar size range as that observed among all classified viruses (18–500 nm). A significant, but variable, proportion of the viruses in seawater enumerated by TEM have tails, suggesting that they belong to the order *Caudovirales* and infect bacteria or archaea (Weinbauer, 2004).

Genome Size

Genomes of viruses vary over orders of magnitude in length. Among marine viruses, for example they range from 3000 nt to 1,290,000 bp. Quantitative analysis of the genome length

distribution in a sample can therefore provide some sense of diversity (Klieve and Swain, 1993; Steward *et al.*, 2000; Wommack *et al.*, 1999). Because conventional electrophoresis can only resolve DNA if it is less than about 20,000–30,000 bp in length, pulsed-field gel electrophoresis (PFGE) has been used to analyze natural assemblages of viruses. The size range of virus genomes detected in seawater with this technique ranges from around 20,000 to >500,000 bp. Under the right conditions, complex, fine-scale banding patterns can be seen, indicating that natural viral communities are quite diverse, but not all genome sizes are equally represented. In a wide variety of marine environments, the distribution of genome sizes appears to be multimodal, with peaks of genome copy numbers frequently occurring for sizes in the vicinity of 30, 55, and 60 kb. The peaks are broad, however, and there are bands detectable throughout the range. Bands >100 kb are usually detected, meaning they contain a significant amount of DNA, but because there is more DNA per large genome, the copy number is low. The mean genome size of dsDNA viruses in seawater is around 50,000 bp, but the mode is closer to 30,000 bp. Because much of the DNA is clustered into several narrow size ranges, many genomes are not adequately resolved by PFGE. This method is therefore of limited value in quantifying viral diversity, especially richness, but remains a convenient means to track changes in community composition (Sandaa *et al.*, 2010).

Single-Gene Surveys

Investigation of the diversity of microbial communities was revolutionized with the development of methods for cultivation-independent molecular surveys. Development of the polymerase chain reaction was transformative because it provided a means to amplify only specific genes of interest from a complex background of DNA (Britschgi and Giovannoni, 1991). Analysis of the amplified material by sequencing or molecular fingerprinting allows one to determine how many variants of the targeted gene were present and thus evaluate genetic diversity. The small subunit ribosomal RNA gene has been the primary target for analyzing the diversity of cellular life, because of its many desirable properties including the fact that it is present in all cellular organisms (Olsen and Woese, 1993). Viruses, however, do not code for ribosomal proteins. Viruses are so diverse, in fact, that there is no single gene that can be used as a universal viral target. Nevertheless, one can identify genes that are present and sufficiently conserved within subgroups of viruses that they can be used to survey diversity for a portion of the viral community. For example, structural genes of myovirids (Fuller *et al.*, 1998; Filée *et al.*, 2005) and phycodnavirids (Larsen *et al.*, 2008), DNA polymerase genes of podovirids (Chen *et al.*, 2009; Labonté *et al.*, 2009) and phycodnavirids (Chen and Suttle, 1995), and the RNA polymerase gene of picornavirads (Culley *et al.*, 2003) have been used to assess the diversity of these various groups in the marine environment. Results from these surveys indicate that members of those viral groups are persistent and globally distributed, but within each group the specific sequence types detected and their relative abundance varies as a function of location, time, and depth. Comparing the viral gene sequences recovered by this cultivation-independent technique with homologous genes from viruses in culture has been very revealing. We now know that most of the viral diversity present in seawater is not represented in culture collections (**Figure 7**).

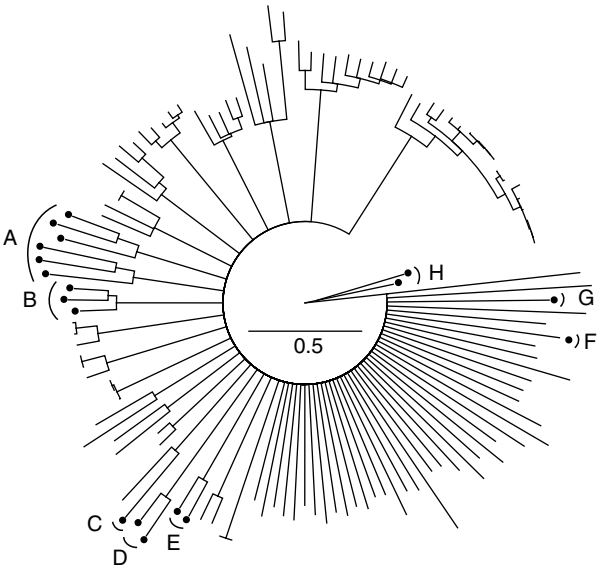


Figure 7 Comparison of the diversity among cultivated and uncultivated viruses. The extensive diversity of viruses revealed by single-gene surveys and how it compares to the diversity among viruses that have been classified is illustrated using marine picorna-like viruses as an example. A phylogenetic tree was created using partial sequences of the RNA-dependent RNA polymerase gene for picornavirad strains in culture (branches marked with a black circle at the tip) or those identified from single-gene surveys of seawater samples (unmarked branches). Groups of sequences identified by an arc and a letter contain representative members of established groups of picornavirads as follows: group A, *Secoviridae*; group B, Bacillarnavirus; group C, *Labyrnaviridae*; group D, *Picornaviridae*; group E, *Dicistroviridae*; group F, *Marnaviridae*; group G, *Iflaviridae*; and the outgroup H, *Potyviridae.*

Metagenomics

There are a number of limitations implicit in the targeted gene approach. First, one recovers information only about variants of a single gene (usually just a portion of the gene). Second, those genes come from only a subset of the viral community, and third, creating a meaningful target in the first place requires at least some prior knowledge about the viruses being sought. Metagenomic analysis, in contrast, allows for a relatively unbiased sampling of all of the viral genomes present in a sample, with no prior knowledge of their sequences (Edwards and Rohwer, 2005). To analyze a viral metagenome, viruses are first harvested from a seawater sample, separated from cells and dissolved nucleic acids, then the viral nucleic acids are extracted. The extracted genomic material is then broken into random, small pieces for sequence analysis. For technical reasons, metagenomic analyses have so far tended to focus on either the DNA viruses or the RNA viruses present in a sample, but it is possible to analyze both at once.

After the analysis, one has a collection of sequence reads that derive from random fragments of the various viral genomes that were present in the original sample. Because one is sampling diversity across entire genomes, rather than a single gene, metagenomic analyses require much greater numbers of sequencing reads (thousands to millions) than are typical of single-gene surveys (tens to hundreds) to draw meaningful conclusions about diversity within a sample and how it compares to other samples. A first-order analysis one can perform is to compare each of the sequences that make up a metagenome to all of the sequences archived in various repositories to see if they are similar to genes that are already known. One can then classify each sequence according to its likely function and the type of organisms from which it might derive. In addition to analyzing each read independently, it is also informative to determine how many of the sequence reads derive from overlapping regions of the same genome, and can thus be reassembled into a single longer read. The higher the diversity in a sample, the lower is the probability that any two or more reads in the randomly selected pool of sequences that comprise the metagenome will overlap. By making a number of assumptions, one can estimate the diversity of the original sample from the assembly statistics for the metagenome (Angly *et al.*, 2005). Although metagenomics has its limitations (Steward and Rappé, 2007), it is providing many new insights into the genetic diversity of viruses in the environment (Kristensen *et al.*, 2010).

DNA-Containing Viruses

In metagenome analyses of marine DNA-containing viruses, most (around 75%) of the viral genes do not resemble any gene with a known function or which derive from a known organism. Among those sequences that are recognizable, the largest proportion (90–95%) of the recognizable reads is most similar to genes of bacteriophages in the order *Caudovirales*. A much smaller proportion resembles ssDNA-containing phages and sequences from eukaryotic viruses, primarily those in the families *Phycodnaviridae* and *Mimiviridae*. Since these data are presumed to represent a random sampling, this suggests that the most common DNA-containing viruses in the ocean are viruses that infect bacteria, followed by viruses that infect protists. DNA-containing animal viruses are very rare in the virioplankton, but metagenomic analyses of samples taken directly from marine animals do reveal sequences resembling known animal viruses (Ng *et al.*, 2009; Ng *et al.*, 2011).

The percentage of viral DNA metagenomic sequences that can be reassembled into larger fragments is typically low. Statistical analysis of assembly data suggest that there are many thousands of distinct virus types in a cubic meter of coastal seawater and tens of thousands in a kg of coastal sediment. The shape of the rank-abundance curves for virus sequence types appears to be best described by a power-law function (Hoffmann *et al.*, 2007). Even the most abundant sequence type in these communities is estimated to make up only around 1% (water) or $<0.01\%$ (sediment) of the total which illustrates the exceptional diversity of marine viral communities (Edwards and Rohwer, 2005).

RNA-Containing Viruses

The results from analyses of metagenomes prepared from viral RNA have been quite different than those prepared with viral DNA. A larger proportion of the reads (around 50%) are similar to known genes and almost all of them resemble eukaryote-infecting viruses. A small percentage of reads appear to derive from ribosomal RNA from cellular contamination, but none of the sequences resemble any of the known RNA bacteriophages. Most of the reads recognizable as viruses are most similar to sequences deriving from ssRNA-containing viruses in the order *Picornavirales*, but sequences resembling ssRNA viruses in other families outside of that order such as the *Tombusviridae* and dsRNA-containing reoviruses have also been detected (Culley *et al.*, 2006).

Larger percentages of the sequences in RNA viral metagenomes have been assembled than has been the case for DNA metagenomes. In fact, entire genomes have been readily assembled from random reads (Culley *et al.*, 2006, Culley *et al.*, 2007). Although RNA viruses are smaller than DNA genomes on average, the results also suggest that RNA viral diversity in the sea may be lower than that of the DNA viruses. However, there are no reliable estimates yet of just how diverse the RNA virus community might be and few environments have been sampled so far.

Conclusion

Viruses are central to the issue of biodiversity. Whether or not one considers them to be alive, they are nevertheless principal contributors to the diversification of life on the planet through their interactions with every organism in every habitat, and they are exceptionally diverse entities in their own right. As rich and complex as the known virus world is, recent molecular probing of the ocean tells us that we are only skimming the surface of the viral diversity.

Appendix

List of Courses

1. Virology
2. Environmental Microbiology
3. Marine Microbiology
4. Marine Plankton Ecology
5. Diseases of Marine Organisms

See also: Aquaculture. Bacterial Genetics. Biodiversity, Evolution and. Census of Marine Life. Diseases, Conservation and. Ecology, Concept and Theories in. Eukaryotes, Origin of. Evolution, Theory of. Food Webs. Macroscopic Patterns in Marine Plankton. Marine Ecosystems. Marine Sediments. Marine Symbioses: Metazoans and Microbes. Microbial Biodiversity. Microbial Biogeography. Microorganisms (Microbes), Role of. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms. Ocean Ecosystems. Origin of Life, Theories of. Origins, Evolution, and Breakdown of Bacterial Symbiosis. Parasitoids. Plankton, Status and Role of. Protozoa. Speciation, Process of. Wildlife Management

References

Ackermann H-W (2006) Classification of bacteriophages. In: Calendar R (ed.) *The Bacteriophages*, 2nd edn. Oxford: Oxford University Press.

Angly F, Rodriguez-Brito B, Bangor D, *et al.* (2005) PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information. *BMC Bioinformatics* 6: 41.

Angly FE, Felts B, Breitbart M, *et al.* (2006) The marine viromes of four oceanic regions. *PLoS Biology* 4: 2121–2131.

Arslan D, Legendre M, Seltzer V, Abergel C, and Claverie J-M (2011) Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proceedings of the National Academy of Sciences of the United States of America* 108: 17486–17491.

Baltimore D (1971) Expression of animal virus genomes. *Bacteriology Reviews* 35: 235–241.

Bird DF, Juniper SK, Ricciardi-Rigault M, Martineu P, Prairie YT, and Calvert SE (2001) Subsurface viruses and bacteria in holocene/late pleistocene sediments of Saanich Inlet, BC: ODP Holes 1033B and 1034B, Leg 169 S. *Marine Geology* 174: 227–239.

Børsheim KY (1993) Native marine bacteriophages. *FEMS Microbiology Ecology* 102: 141–159.

Bratbak G, Jacobsen A, and Heldal M (1998) Viral lysis of *Phaeocystis pouchetii* and bacterial secondary production. *Aquatic Microbial Ecology* 16: 11–16.

Breitbart M, Thompson LR, Suttle CA, and Sullivan MB (2007) Exploring the vast diversity of marine viruses. *Oceanography* 20: 135–139.

Bressem M-FV, Waerebeek KV, and Raga JA (1999) A review of virus infections of cetaceans and the potential impact of morbilliviruses, poxviruses and papillomaviruses on host population dynamics. *Diseases of Aquatic Organisms* 38: 53–65.

Britschgi TB and Giovannoni SJ (1991) Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. *Applied and Environmental Microbiology* 57: 1707–1713.

Brussaard CPD (2004) Viral control of phytoplankton populations – A review. *Journal of Eukaryotic Microbiology* 51(125-38):

Brussaard CPD, Short SM, Frederickson CM, and Suttle CA (2004) Isolation and phylogenetic analysis of novel viruses infecting the phytoplankton *Phaeocystis globosa* (Prymnesiophyceae). *Applied and Environmental Microbiology* 70(3700-5).

Chen F and Suttle CA (1995) Amplification of DNA polymerase gene fragments from viruses infecting microalgae. *Applied and Environmental Microbiology* 61: 1274–1278.

Chen F, Wang K, Huang S, *et al.* (2009) Diverse and dynamic populations of cyanobacterial podoviruses in the Chesapeake Bay unveiled through DNA polymerase gene sequences. *Environmental Microbiology* 11: 2884–2892.

Culley AI, Lang AS, and Suttle CA (2003) High diversity of unknown picorna-like viruses in the sea. *Nature* 424: 1054–1057.

Culley AI, Lang AS, and Suttle CA (2006) Metagenomic analysis of coastal RNA virus communities. *Science* 312: 1795–1798.

Culley AI, Lang AS, and Suttle CA (2007) The complete genomes of three viruses assembled from shotgun libraries of marine RNA virus communities. *Virology Journal* 4: 69.

Danovaro R, Corinaldesi C, Filippini M, *et al.* (2008) Viriobenthos in freshwater and marine sediments: A review. *Freshwater Biology* 53: 1186–1213.

Davison AJ, Eberle R, Ehlers B, *et al.* (2009) The order *Herpesvirales*. *Archives of Virology* 154: 171–177.

Edwards RA and Rohwer F (2005) Viral metagenomics. *Nature Reviews Microbiology* 3: 504–510.

Espejo RT and Canelo ES (1968) Properties of bacteriophage PM2: A lipid-containing bacterial virus. *Virology* 34: 738–747.

Filée J, Tétart F, Suttle CA, and Krisch HM (2005) Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12471–12476.

Fischer MG, Allen MJ, Wilson WH, and Suttle CA (2010) Giant virus with a remarkable complement of genes infects marine zooplankton. *Proceedings of the National Academy of Sciences of the United States of America* 107: 19508–19513.

Frank H and Moebus K (1987) An electron microscopic study of bacteriophages from marine waters. *Helgoländer Meeresuntersuchungen* 41: 385–414.

Fuhrman J (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* 399: 541–548.

Fuhrman JA (1992) Bacterioplankton roles in cycling of organic matter: The microbial food web. In: Falkowski PG and Woodhead AD (eds.) *Primary Productivity and Biogeochemical Cycles in the Sea*. New York: Plenum Press.

Fuller NJ, Wilson WH, Joint IR, and Mann NH (1998) Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. *Applied and Environmental Microbiology* 64: 2051–2060.

Garza DR and Suttle CA (1995) Large double-stranded DNA viruses which cause the lysis of a marine heterotrophic nanoflagellate (*Bodo* sp.) occur in natural marine viral communities. *Aquatic Microbial Ecology* 9: 203–210.

Gowing MM (1993) Large virus-like particles from vacuoles of phaeodarian radiolarians and from other marine samples. *Marine Ecology Progress Series* 101: 33–43.

Graham LE and Wilcox LW (2000) *Algae*. New Jersey: Prentice-Hall, Inc.

Green KY (2007) *Caliciviridae*: The noroviruses. In: Knipe DM and Howley PM (eds.) *Fields Virology*, 5th edn. Philadelphia: Wolters Kluwer Lippincott Williams & Wilkins.

Hanson L, Dishon A, and Kotler M (2011) Herpesviruses that infect fish. *Viruses* 3: 2160–2191.

Harrison SC, Caspar DLD, Camerini-Otero RD, and Franklin RM (1971) Lipid and protein arrangement in bacteriophage PM2. *Nature New Biology* 229: 197–201.

Hoffmann KH, Rodriguez-Brito B, Breitbart M, *et al.* (2007) Power law rank-abundance models for marine phage communities. *FEMS Microbiology Letters* 273: 224–228.

Holmes F (1948) Order *Virales*, the filterable viruses. In: Breed RS, Murray EGD, and Parker Hitchins A (eds.) *Bergey's Manual of Determinative Bacteriology*, 6th edn. Baltimore: Williams and Wilkins.

King AMQ, Adams MJ, Carstens EB, and Lefkowitz EJ (eds.) (2012) *Virus Taxonomy. Ninth report of the International Committee on Taxonomy of Viruses*. New York: Elsevier.

Klieve AV and Swain RA (1993) Estimation of ruminal bacteriophage numbers by pulsed-field gel electrophoresis and laser densitometry. *Applied and Environmental Microbiology* 59: 2299–2903.

Koonin EV (2009) On the origin of cells and viruses: Primordial virus world scenario. *Annals of the New York Academy of Sciences* 1178: 47–64.

Kristensen DM, Mushegian AR, Dolja VV, and Koonin EV (2010) New dimensions of the virus world discovered through metagenomics. *Trends in Microbiology* 18: 11–19.

Krupovič M and Bamford DH (2007) Putative prophages related to lytic tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic bacteria. *BMC Genomics* 8: 236.

Krupovič M and Cvirkaite-Krupovič V (2011) Virophages or satellite viruses? *Nature Reviews Microbiology* 9: 762–763.

Kuris AM and Lafferty KD (2000) Parasite–host modeling meets reality: Adaptive peaks and their ecological attributes. In: Poulin R, Morand S, and Skorping A (eds.) *Evolutionary Biology of Host–Parasite Relationships: Theory Meets Reality*. Elsevier Science

Labonté JM, Reid KE, and Suttle CA (2009) Phylogenetic analysis indicates evolutionary diversity and environmental segregation of marine podovirus DNA polymerase gene sequences. *Applied and Environmental Microbiology* 75: 3634–3640.

Lamb RA and Parks GD (2007) *Paramyxoviridae*: The viruses and their replication. In: Knipe DM and Howley PM (eds.) *Fields Virology*, 5th edn. Philadelphia: Wolters Kluwer Lippincott Williams & Wilkins.

Lang AS, Rise ML, Culley AI, and Steward GF (2009) RNA viruses in the sea. *FEMS Microbiology Reviews* 33: 295–323.

Lanni F (1964) *Plant Viruses*. Gainesville, FL: University of Florida Press.

Larsen JB, Larsen A, Bratbak G, and Sandaa R-A (2008) Phylogenetic analysis of members of the *Phycodnaviridae* virus family using amplified fragments of the major capsid protein gene. *Applied and Environmental Microbiology* 74: 3048–3057.

Lawrence JE, Brussaard CPD, and Suttle CA (2006) Virus-specific responses of *Heterosigma akashiwo* to infection. *Applied and Environmental Microbiology* 72: 7829–7834.

Levine AJ and Enquist LW (2007) History of virology. In: Knipe DM, Howley PM, Griffin DE, *et al.* (eds.) *Fields Virology*, 5th edn. Philadelphia: Lippincott, Williams and Wilkins.

Lwoff A (1957) The concept of virus. *Journal of General Microbiology* 17: 239–253.

Lwoff A, Horne RW, and Tournier P (1962) A system of viruses. *Cold Spring Harbor Symposia on Quantitative Biology* 27: 51–55.

Lwoff A and Tournier P (1966) The classification of viruses. *Annual Review of Microbiology* 20: 45–74.

Maness HTD, Nollens HH, Jensen ED, *et al.* (2011) Phylogenetic analysis of marine mammal herpesviruses. *Veterinary Microbiology* 149: 23–29.

Männistö RH, Kivelä HM, Paulin L, Bamford DH, and Bamford JKH (1999) The complete genome sequence of PM2, the first lipid-containing bacterial virus to be isolated. *Virology* 262: 355–363.

Matsuyama Y (1999) Harmful effect of dinoflagellate *Heterocapsa circularisquama* on shellfish aquaculture in Japan. *Japan Agricultural Research Quarterly* 33: 283–294.

Middelboe M, Glud R, and Filippini M (2011) Viral abundance and activity in the deep sub-seafloor biosphere. *Aquatic Microbial Ecology* 63: 1–8.

Middelboe M, Jørgensen NOG, and Kroer N (1996) Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. *Applied and Environmental Microbiology* 62: 1991–1997.

Moebus K and Nattkemper H (1981) Bacteriophage sensitivity patterns among bacteria isolated from marine waters. *Helgoländer Meeresuntersuchungen* 34: 375–385.

Morse SS (1994) Toward an evolutionary biology of viruses. In: Morse SS (ed.) *The Evolutionary Biology of Viruses*. New York: Raven Press.

Munn CB (2006) Viruses as pathogens of marine organisms – From bacteria to whales. *Journal of the Marine Biological Association of the United Kingdom* 86: 453–467.

Murray AG and Eldridge PM (1994) Marine viral ecology – Incorporation of bacteriophage into the microbial planktonic food web paradigm. *Journal of Plankton Research* 16: 627–641.

Nagasaki K, Shirai Y, Takao Y, Mizumoto H, Nishida K, and Tomaru Y (2005) Comparison of genome sequences of single-stranded RNA viruses infecting the bivalve-killing dinoflagellate *Heterocapsa circularisquama*. *Applied and Environmental Microbiology* 71: 8888–8894.

Nagasaki K, Tomaru Y, Katanozaka N, *et al.* (2004) Isolation and characterization of a novel single-stranded RNA virus infecting the bloom-forming diatom Rhizosolenia setigera. *Applied and Environmental Microbiology* 70: 704–711.

Nagasaki K, Tomaru Y, Tarutani K, *et al.* (2003) Growth characteristics and intraspecies host specificity of a large virus infecting the dinoflagellate *Heterocapsa circularisquama*. *Applied and Environmental Microbiology* 69: 2580–2586.

Ng TFF, Manire C, Borrowman K, Langer T, Ehrhart L, and Breitbart M (2009) Discovery of a novel single-stranded DNA virus from a sea turtle fibropapilloma by using viral metagenomics. *Journal of Virology* 83: 2500–2509.

Ng TFF, Wheeler E, Greig D, Waltzek TB, Gulland F, and Breitbart M (2011) Metagenomic identification of a novel anellovirus in Pacific harbor seal (Phoca vitulina richardsii) lung samples and its detection in samples from multiple years. *Journal of General Virology* 92: 1318–1323.

Not F, Latasa M, Marie D, Cariou T, Vaulot D, and Simon N (2004) A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Applied and Environmental Microbiology* 70: 4064–4072.

Olsen GJ and Woese CR (1993) Ribosomal RNA: A key to phylogeny. *FASEB Journal* 7: 113–123.

Palese P and Shaw ML (2007) *Orthomyxoviridae*: The viruses and their replication. In: Knipe DM and Howley PM (eds.) *Fields Virology*, 5th edn. Philadelphia: Wolters Kluwer Lippincott Williams & Wilkins.

Plyusnin A, Beaty BJ, Elliott RM, *et al.* (2012) Bunyaviridae. In: King AMQ, Adams MJ, Carstens E, and Lefkowitz EJ (eds.) *Virus Taxonomy Ninth Report of the International Committee on Taxonomy of Viruses*. London: Elsevier Academic Press.

Raghukumar S (2002) Ecology of the marine protists, the Labyrinthulomycetes (Thraustochytrids and Labyrinthulids). *European Journal of Protistology* 38: 127–145.

Rohwer F and Vega Thurber R (2007) Viruses manipulate the marine environment. *Nature* 459: 207–212.

Sandaa RA, Short SM, and Schroeder D (2010) Fingerprinting aquatic virus communities. In: Suttle CA, Wilhelm SW, and Weinbauer MG (eds.) *Methods in Aquatic Viral Ecology*. Waco, TX: American Society of Limnology and Oceanography.

Scola LA, Audic B, Robert S, *et al.* (2003) A giant virus in amoebae. *Science* 299: 2033.

Shirai Y, Tomaru Y, Takao Y, Suzuki H, Nagumo T, and Nagasaki K (2008) Isolation and characterization of a single-stranded RNA virus infecting the marine planktonic diatom *Chaetoceros tenuissimus* Meunier. *Applied and Environmental Microbiology* 74: 4022–4027.

Smayda TJ (1998) Ecophysiology and bloom dynamics of *Heterosigma akashiwo* (Raphidophyceae). In: Anderson DM, Cembella AD, and Hallegraeff GM (eds.) *Physiological Ecology of Harmful Algal Blooms*, pp. 113–131. NATO ASI Series. Berlin: Springer.

Steward GF, Montiel JL, and Azam F (2000) Genome size distributions indicate variability and similarities among marine viral assemblages from diverse environments. *Limnology and Oceanography* 45: 1697–1706.

Steward GF and Rappé MS (2007) What's the "meta" with metagenomics? *The ISME Journal* 1: 100–102.

Sullivan MB, Huang KH, Ignacio-Espinoza JC, *et al.* (2010) Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environmental Microbiology* 12: 3035–3056.

Sullivan MB, Waterbury JB, and Chisholm SW (2003) Cyanophage infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* 424: 1047–1051.

Suttle CA (2007) Marine viruses – Major players in the global ecosystem. *Nature Reviews Microbiology* 5: 801–812.

Tai V, Lawrence JE, Lang AS, Chan AM, Culley AI, and Suttle CA (2003) Characterization of HaRNAV, a single-stranded RNA virus causing lysis of *Heterosigma akashiwo* (Raphidophyceae). *Journal of Phycology* 39: 343–352.

Takao Y, Nagasaki K, Mise K, Okuno T, and Honda T (2005) Isolation and characterization of a novel single-stranded RNA virus infectious to a marine fungoid protist, *Schizochytrium* sp. (Thraustochytriaceae, Labyrinthulea). *Applied and Environmental Microbiology* 71: 4516–4522.

Tomaru Y, Katanozaka N, Nishida K, *et al.* (2004) Isolation and characterization of two distinct types of HcRNAV, a single-stranded RNA virus infecting the bivalve-killing microalga *Heterocapsa circularisquama*. *Aquatic Microbial Ecology* 34: 207–218.

Tomaru Y, Takao Y, Suzuki H, Nagumo T, and Nagasaki K (2009) Isolation and characterization of a single-stranded RNA virus infecting the bloom-forming diatom *Chaetoceros socialis*. *Applied Environmental Microbiology* 75: 2375–2381.

Tucker KP, Parsons R, Symonds EM, and Breitbart M (2011) Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. *The ISME Journal* 5: 822–830.

Villarreal LP (2004) *Viruses and the Evolution of Life*. Washington, DC: ASM Press.

Villareal LP and Witzany G (2010) Viruses are essential agents in the roots and stem of the tree of life. *Journal of Theoretical Biology* 262: 698–710.

Walker PJ and Winton JR (2010) Emerging viral diseases of fish and shrimp. *Veterinary Research* 41: 51.

Weinbauer MG (2004) Ecology of prokaryotic viruses. *FEMS Microbiology Reviews* 28: 127–181.

Wichels A, Biel SS, Gelderblom HR, Brinkhoff T, Muyzer G, and Schütt C (1998) Bacteriophage diversity in the North Sea. *Applied and Environmental Microbiology* 64: 4128–4133.

Wommack KE and Colwell RR (2000) Virioplankton: Viruses in aquatic ecosystems. *Microbiology and Molecular Biology Reviews* 64: 69–114.

Wommack KE, Ravel J, Hill RT, Chun JS, and Colwell RR (1999) Population dynamics of Chesapeake bay virioplankton: Total-community analysis by pulsed-field gel electrophoresis. *Applied and Environmental Microbiology* 65: 231–240.

Market Economy and Biodiversity

Heather Klemick and R David Simpson, National Center for Environmental Economics, Washington, DC, USA

Published by Elsevier Inc.

The opinions expressed herein are those of the authors, and do not necessarily represent those of the US Environmental Protection Agency.

Glossary

Cap and trade A type of "market-based incentive" that sets a regulatory requirement that firms or other entities covered under the program achieve an aggregate environmental target (such as a maximum amount of pollution or minimum quantity of habitat preservation), and allows the compliance obligation to be traded across firms to meet the objective at lowest cost.

Certification Procedure of determining whether a particular process is conducted in an ecologically benign fashion (also referred to as ecolabeling).

Command and control Policy instrument in which regulated firms must meet individual environmental targets, often requiring the use of specific technologies. Command and control policies contrast with "market-based incentives," which allow firms more flexibility in how to meet aggregate environmental targets.

Conservation concession Type of "payment for ecosystem services" arrangement in which a buyer purchases the right to manage a large tract of natural habitat for conservation purposes instead of commercial exploitation.

Individual transferable quota (ITQ) Type of "market-based incentive" in which regulators set an aggregate limit on harvests of an open access resource (typically fish), assign "property rights" for the allowable catch to individual participants, and permit participants to trade those property rights.

Market economy Organization of economic activity in which private individuals buy and sell goods at prices that balance supply and demand.

Market-based incentives Policy designed to change behavior by altering the price paid, or cost borne, by a person engaged in an activity that affects biodiversity.

Payments for ecosystem services (PES) A type of "market-based incentive" that compensates communities, individuals, or other entities whose actions increase the supply of "public goods" from natural ecosystems.

Introduction

The topic, "Market Economy and Biodiversity," could be interpreted in several ways.

This article focuses on three interpretations and their consequences:

1. A "market economy" might be defined as a *laissez faire*, private ownership system for the production and distribution of goods. A private market economy tends to underprovide public goods. Because biodiversity gives rise to public goods, such a system will not preserve as much biodiversity as would be socially desirable.
2. The above observation notwithstanding, some of the benefits of biodiversity use can be privately appropriated. We could then think in terms of markets for certain aspects of biodiversity that might be embodied in private goods. A popular strategy for biodiversity protection in recent decades has been to capitalize on these privately appropriable values by encouraging markets in sustainably produced goods and services. The authors review some considerations in the development of such markets.
3. Although there are "missing markets" for the public goods provided by biodiversity, public action such as regulation can create artificial markets for some of these goods. Creating markets for biodiversity services can move a market economy closer to the ideal of an efficient, perfectly competitive economy, in which all goods are priced to reflect their value to society. The authors consider some examples and limitations of such markets.

Biodiversity in a Private Ownership Market Economy

In a private, unregulated market economy, private agents guided by self-interest decide for themselves what to make, buy, and consume. A celebrated result dating back to Adam Smith's *Wealth of Nations* (1904 (1776)) holds that the "invisible hand" of the market will generate a socially desirable outcome even though it is driven by the selfish interests of myriad independent actors. This result was later formalized in what is known in economics as the "First Welfare Theorem": A perfectly competitive market economy is efficient. (Efficiency has a specific and limited meaning in economics: An outcome is efficient if no one can be made better off without making another worse off.) Efficiency does not address another social objective: equity. However, another fundamental result, known as the Second Welfare Theorem, holds that any efficient outcome – including any that are both efficient and equitable – can be achieved by first reallocating the initial distribution of wealth and then allowing trading in a market economy. There is, however, controversy as to whether it is ever possible to reallocate wealth without altering the

incentives to amass wealth, which would violate the assumptions under which the welfare theorems are derived.

Even introductory economics courses quickly teach their students that the perfectly competitive market economy as envisioned by Smith and mathematically formalized in the middle of the twentieth century by Nobel laureates such as Kenneth Arrow, Gerard Debreu, and John Nash rests on a number of very restrictive assumptions. For readers interested in environmental policy, the most problematic of these assumptions is that there are complete markets; that is, all goods in the economy are, in fact, owned by someone, and, consequently, if one person could realize greater satisfaction from a good than could its current owner, the former will purchase the good from the latter. This sounds like a reasonable depiction of the circumstances of ownership of, say, bread, cars, and cell phones. For what types of goods is this description not adequate?

The answer, in brief, is "public goods." It may be best to define a public good by contrasting it with a private good. You and I cannot both consume all of the same loaf of bread, because bread is a private good. It is by definition, a good from whose enjoyment I can exclude you – if I have a loaf of bread, I can prevent you from simply taking it from me without compensating me a mutually agreed-on amount (Of course in most societies, preventing you from appropriating my belongings involves more than personal defense; police and courts play a role as well. Such mechanisms are no less necessary in the allocation of public goods than they are in the defense of private ones, but the authors will abstract from further consideration of them here.) – and whose enjoyment is rival – if I eat my loaf of bread, you cannot eat the same loaf of bread. A public good such as environmental quality, in contrast, is neither excludable nor rival. You cannot be excluded from enjoying the benefits of, say, cleaner air, even if it is my actions that have resulted in the air being cleaner. Nor does my consumption of clean air diminish in any way the benefits you enjoy from cleaner air.

Biodiversity gives rise to many public goods. For example, preserved natural ecosystems may both harbor a diverse collection of species and provide downstream communities with protection against flooding by retaining water. The flood protection, however, is necessarily available to all members of the downstream communities if it is available to any. There are numerous other examples. Consider, for instance, the simple moral or esthetic satisfaction one realizes from knowing that such unique and fascinating creatures as giant pandas, California condors, and tigers continue to exist. This is a quintessential public good. My satisfaction in knowing that there are still tigers in the wild in no way reduces the satisfaction you may take in the same knowledge. Furthermore, villagers in India may "pay for" preserving tigers in that tigers they refrain from killing may attack their livestock or themselves. These villagers are providing a benefit to people elsewhere who care about the continued existence of tigers for purely moral or esthetic reasons.

The types of markets envisioned by Smith are incapable of allocating public goods efficiently. The reason is simple. Goods are efficiently produced when the benefits accruing to all members of society collectively from one additional unit of their production are exactly offset by the cost of that production. This balancing occurs naturally when a private purchaser compares the benefit she alone receives from the purchase of a good with price she must pay for it, and that price is exactly the same as the incremental cost the good's provider incurs to offer.

When one provides a public good, however, the benefits accrue to many people. Unless one accounts for the benefits to others when making the choice of whether to provide a public good, the public good will be undersupplied. This would not occur if everyone who benefited could be compelled to pay for the benefits she receives, but this is where the problems of nonexcludability and nonrivalry arise. People will not choose voluntarily to pay their fair share of the cost of providing public goods because each would be better off to let others pay their shares and to "free ride" on the benefits the others provide.

Consider again the tiger example. Those of us in the wealthier countries who care about the survival of tigers could pay villagers not to kill the tigers, or perhaps pay to relocate the villagers to areas away from tigers. If I were willing to pay for this outcome, though, why would you want to "buy" a good that I was providing to you anyway? Perhaps we could devise a system under which we both contributed to the payments necessary to compensate Indian villagers for tiger preservation. When you consider that such a scheme would have to involve millions of people to cover all the beneficiaries of tiger preservation, however, you appreciate that what we are then contemplating is really a public program for the provision of a public good – an alternative to a private free market.

The organization of economic activity into more-or-less private markets is, by and large, a phenomenon that began several hundred years ago in the West and has expanded worldwide in more recent decades (while the world's major economies have increasingly been organized along market lines, virtually all remain "mixed" economies, in which economic activity is apportioned in varying degrees between private and public sectors). (For an interesting perspective on changes in social views concerning private self-interest over the centuries, see Heilbroner, 1999.) During this period over which market economies became predominant, the pace of population and economic growth was historically unprecedented. When Columbus arrived in the New World fewer than half a billion people inhabited the planet. Population did not increase to a billion until the early years of the nineteenth century. It took more than a century to reach two billion, but then rose to 3 billion by 1960, and doubled again between then and the end of the millennium (De Long, 1998). Even as the numbers of people on the planet skyrocketed, their economic activity increased still more rapidly. J Bradford DeLong estimates the level of income at which the vast majority of the world's population barely subsisted in 1500 as 115 current US dollars per capita. That figure increased steadily to $210 in 1800, $679 in 1900, $1622 in 1950, and something greater than $6500 today. Multiplying population and per capita income figures, the scale of economic activity increased by a factor of 1000 in the course of the last 500 years.

The scale of economic activity neither tracks exactly the degradation of the environment in general nor the decline in

biodiversity. Technological improvements may result in the production of both more valuable and less environmentally damaging goods. The empirical fact is, however, that biodiversity has declined with the appearance and expansion of modern market economies. It is easy to link the causes of biodiversity loss with the hallmarks of economic growth. Overharvesting results when growing demands for fish, timber, and other biological resources interact with emerging technologies for their extraction and exploitation. Modern market economies are not conducive to the types of social norms and local institutions that have, in many cases, led to sustainable resource extraction from common-pool resources in small-scale preindustrial communities (e.g., Ostrom, 1990). International trade and travel are leading causes of the introduction of exotic diseases, pests, and predators that have eliminated native populations, particularly in isolated habitats. (It is worth noting, however, that prehistoric human migrations also had devastating effects on native biota. Paleontological evidence suggests that the extinction of American megafauna were at least suspiciously contemporary with the migration of humans across the Bering land bridge, even if experts disagree as to the culpability of humans. The extinction of Pacific island fauna, such as the giant Moa of New Zealand, has been more definitively linked to the arrival of Polynesian voyagers and, in some instances more importantly, the rats and pigs they brought with them.) In the early nineteenth century, William Blake wrote that the industrial revolution had brought "dark satanic mills," to "England's green and pleasant land," and by the end of the twentieth century the industrial air and water pollution that had transformed landscapes in the worlds' wealthier nations was also to be found, often in greater quantities and concentrations, in less-developed countries. Perhaps most importantly, the sheer scale of human activity has resulted in the destruction of natural habitats to provide more area for industry, residences, and agriculture.

While it is accurate to say that the world's most prosperous economies are now those that follow a more-or-less "market" paradigm, blame for ecological damage may not be laid exclusively at the doorstep of market economies. The Soviet Union wreaked incredible havoc on its native ecosystems (Feshbach and Friendly, 1993). Other economies with legacies of central planning have also often shown disregard for environmental degradation (see, e.g., World Bank and the State Environmental Protection Agency of China (World Bank), 2007, on China; or Kirby, 2004, on North Korea). (It is somewhat ironic that states whose governments purport to represent the interests of their citizenry as a whole have sometimes had abysmal records in providing for the common welfare through environmental protection. We argue that public action is needed to provide public goods, but it has often been the case that the economies whose public sectors were most dominant provided the least of the public good of environmental protection.)

One could argue whether the emergence of market economies and their associated technologies and mores itself drove the decline of biodiversity, or if other more fundamental factors drove both; consider for example, Max Weber's celebrated treatise on *The Protestant Ethic and the Spirit of Capitalism* (1958 (1904–1905)). Diamond (1999) advances the fascinating thesis that the distribution of biodiversity underlay the global dominance that Western Europe eventually achieved: Western Europeans – and their far-flung progeny – came to control much of the world because they were endowed with the richest biodiversity to exploit. If this is true, one might say that biodiversity sowed the seeds of its own destruction through a circuitous process encompassing ever more productive agriculture, the domestication of animals, coevolution with pathogens, the emergence of larger and larger polities, accelerating technological progress, and the breakneck rate of growth the authors have summarized above.

Whether or not the emergence of the market economy may be blamed in whole or in part for the biodiversity crisis, there is increasing interest in harnessing markets to halt biodiversity loss. The ways in which this might be done comprise the topic of the remainder of this article. Let us make one more observation before turning to that topic, however. It is commonly observed in the economic literature that environmental improvement is a normal good – that people manifest a greater concern for environmental protection when their incomes increase (see, e.g., Kolstad, 2000). There may yet be some hope, then, that even if the emergence of market economies is to blame for the decline of biodiversity, the wealth such economies have produced may also be channeled into preserving what is left of the natural world.

Private Markets for Biodiversity-Related Services

The private market economy will not efficiently provide public goods, such as those associated with biodiversity. Not all of the services associated with biodiversity are public goods, however. Consider, for example, experiences associated with travel and tourism. In some nature reserves, admission is restricted and entrance fees are charged. If it is, in fact, feasible to exclude those who do not pay for a good from enjoying it, the good is essentially rendered private and it may be effectively preserved in a private market economy.

In other cases, it may not be feasible to restrict entry *per se*, but if tourists visiting an area require overnight accommodation, guides, or specialized forms of transport, it may be possible to restrict access to these and charge a price for, say, hotel rooms, that includes some compensation for the unique biodiversity to which guests have access.

Such "ecotourism" ventures are examples of a broader strategy of making sustainable use of biodiversity. A sustainable use strategy was advocated in policy statements of major conservation groups in the 1980s and 1990s (IUCN *et al.*, 1981, 1991) and has since often been a major item in international conservation policy and, more to the point, international conservation and development strategy. The development component has been particularly important, as there is a rough inverse correlation between the economic performance and the biodiversity endowments of nations, with the greatest diversity of species often found in the some of the least prosperous regions.

It was then hoped that by making sustainable use of biodiversity, incentives would be created for its conservation. Besides ecotourism, these sustainable use strategies include collection and sale of products from natural areas (often

classified in forest areas as "nontimber forest products"), "bioprospecting" (the search among naturally occurring biological products for chemical entities of value in agricultural, industrial, and, particularly, pharmaceutical applications), and sustainable hunting, farming, and timber harvesting.

There can be little dispute that consumer willingness-to-pay for the sustainable exploitation of natural products creates incentives for the maintenance of the sources of such products. What has proved to be considerably more controversial, however, is whether promotion of such markets is in fact a viable and desirable conservation strategy. Some interrelated issues arise in the consideration of that question:

- Do such markets provide sufficient incentives for conservation?
- Do such markets provide incentives for the conservation of biodiversity *per se*, or rather for selected components of biodiversity, with adverse consequences for others?
- Is the promotion of markets for natural products a cost-effective conservation strategy, or would limited conservation resources be better devoted to other approaches?

These questions have been addressed in a lively debate on the efficacy of Integrated Conservation and Development Projects (ICDPs), ventures intended to achieve the twin goals of biodiversity conservation and economic development by promoting the sustainable use of natural systems. While ICDPs received considerable attention and funding in the 1990s, they have attracted criticism from both conservation practitioners (Terborgh, 1999; Oates, 1999; Terborgh *et al.*, 2002) and social science commentators (Wells *et al.*, 1992, 1999; Southgate, 1998; Ferarro and Simpson, 2002), who suggest that confusion between conservation and development objectives sometimes results in neither being met.

Let us briefly consider the questions posed above in turn. The answer to the first must certainly be "no;" the global community cares about biodiversity for a suite of both instrumental and intangible reasons. Since the demand for particular products does not reflect less tangible concerns for the existence of other life forms, such a demand will never result in the appropriate level of conservation. While this argument is easily demonstrated (see, e.g., Pearce and Moran, 1994), it does not necessarily imply that sustainable use is not a useful strategy to conservation in a world willing to devote only limited resources to conservation. While one might argue that sustainable use strategies do not motivate enough conservation, they may have the virtue of motivating at least some conservation.

Expectations may need to be tempered, however. A fundamental result in economics holds that economic value increases with scarcity. In economic parlance, a payment that arises solely as compensation for parting with a scarce good is known as a rent (or sometimes more explicitly as a "scarcity rent"). Put in another way, a rent is a payment over and above the cost of providing a good. (Economic nomenclature can become somewhat balky. If one is a stickler for exactitude, she might insist instead that an economic rent be defined as the opportunity cost of selling a scarce good to one prospective purchaser rather than to another). While society's concern for the decline of biodiversity arises from the growing appreciation that the benefits of nature are growing increasingly scarce *vis a vis* the scale of the artificial world, it does not necessarily follow that all of the goods and services to which nature might give rise – and especially those that might be appropriated in private markets – are growing desperately scarce. One must be careful to distinguish between compensation paid for tangible investments and the rents commanded by access to nature *per se*. One might, for example, build a luxury hotel in the midst of stunning natural beauty, but much of the price of a night's stay there may reflect compensation for investment in the hotel rather than payment for access to the surrounding nature. There remain many, many very attractive places in the world in which to construct hotels, and very few of them are sufficiently unique as to command extremely high rents to benefit their indigenous settings.

Similar arguments have been made concerning the value of biodiversity for use in new product development. Bioprospecting agreements, which have typically involved pharmaceutical and agricultural biotechnology firms, allow companies access to undeveloped habitat that could harbor organisms with genetic information useful for developing new products in exchange for monetary compensation to host communities or countries. There was considerable enthusiasm in 1991 when Merck, one of the world's largest pharmaceutical companies, and Costa Rica's National Biodiversity Institute entered into the first international biodiversity contract. A number of other companies and ventures followed. While it is true that nature has provided a vast collection of blueprints for new product development, many advocates have been disappointed by the lack of success to date in bioprospecting. There have been no spectacular discoveries for which developing countries have been compensated, and many of the pioneering bioprospecting ventures have lapsed or been abandoned over time (MacIlwain, 1998; Eisner, 2003; Firn, 2003). While it is not yet clear what the ultimate commercial success of bioprospecting will be, neither its commercial nor ecological goals have yet been reached. (Of course, the lack of resulting commercial products would not necessarily be unexpected, even if prospects were generally favorable. The pharmaceutical research industry is notoriously risky, and for any single project, failure is far more likely than success.)

Some authors have argued that this record arises not because biodiversity does not afford frequent enough sources of new products, but, paradoxically, because potential sources of new products are too abundant. The value of an untested species as a lead in the development of a new product is not determined by simply multiplying the value of the product being sought times the probability that the species under consideration will lead to its discovery. Rather, the value of the untested species is the value of the product being sought times the probability that the species under consideration will lead to its discovery times the probability that the new product will not be discovered among any other available natural or artificial sources (Simpson *et al.*, 1996; see also Rausser and Small, 2000; Costello and Ward, 2006). If nature or synthetic chemistry can provide a multitude of alternatives, bioprospectors will not be willing to pay much for any particular natural lead.

There are, of course, any number of other reasons for which natural sites that could be used for tourism or bioprospecting should be preserved; and, in fact, the development of new products may confer substantial benefits over and above those that may be appropriated by private developers (Craft and Simpson, 2001). These observations underscore our main point; however, markets for the privately appropriable benefits of biodiversity may serve only to preserve a small fraction of the resources about which society cares.

Let us now turn to the second question, which concerns the character of the conservation motivated by sustainable use strategies. Commercial ventures in ecotourism or nontimber forest products are often focused on a subset of attributes or products of an area, and the focus on one element of a system may come at the cost of another. It is sometimes alleged that ecotourism ventures may "love nature to death" by allowing overexploitation and excessive use (see, e.g., Honey, 1999). (The authors consider in this discussion whether, in focusing on one set of assets, ecoentrepreneurs hasten the degradation of another. A related issue is whether, in a rush to exploit natural assets, ecoentrepreneurs hasten the degradation of the very assets on which their enterprises depend. The authors will consider this second issue below in their discussion of property rights.) In some instances, biodiversity-rich areas may attract tourists not so much for their biological as for their topographical wonders – waterfalls such as Angel and Victoria Falls come to mind – and tourism promoters might see elements of biodiversity such as poisonous animals or large predators as impediments to tourist access rather than enhancements to the overall experience (Terborgh, 1999).

Similar phenomena arise in the collection of nontimber forest products. Harvesting edible products from the wild necessarily means that nutrients are removed from the ecosystem that would otherwise have been available to indigenous animals (Peters, 1994). Moreover, the history of agriculture is essentially a record of increased concentration and specialization. Those plants that humans found more attractive were selected and cultivated, while those that competed with crops for space, sunlight, and nutrients were selectively eliminated. Ironically, conservation initiatives intended to promote the collection of "natural products" have, on occasion, led to the establishment of *de facto* plantations of things such as tagua nut trees (Chomitz and Kumari, 1996). Similar phenomena may be encountered in situations in which the preservation of some "desirable" animals may mean the culling of those that prey on or compete with them.

Whether one regards such "managed natural systems" as viable conservation approaches may depend on one's objectives and sense of possibilities. To some biologists such as John Terborgh (1999), the conservation of biodiversity means the preservation of systems large enough to admit the continuation of natural processes of evolution, including predation by large and dangerous animals. Others would be satisfied by the adoption of agricultural production processes that are simply less intensive than the more industrial alternatives now available (see, e.g., Daily and Ellison, 2002).

Sustainable alternatives to conventional products have gained an additional foothold in the marketplace through the certification of "green" products. Many people are reluctant to buy things such as lumber, ivory, fish, or coffee without some assurance that they have not caused undue environmental harm. By certifying such products, organizations such as Green Seal or the Forest Stewardship Council enhance their value to environmentally conscious consumers. Such labeling programs help lower the information-gathering costs of determining product-specific biodiversity impacts for consumers who are willing to pay a premium for goods and services produced in a more sustainable way.

However, the potential of certification programs is limited by the public good characteristics of biodiversity conservation already mentioned: Each consumer who pays more for a product or service than it would cost to purchase an alternative, less ecofriendly item is, in effect, making a contribution for conservation. Other consumers, even though they may have similar environmental sentiments, will be tempted to free-ride on such contributions. (See Sedjo and Swallow (2002) or Mason (2006) for formal analyses of some of these issues.)

Advocacy of a sustainable use approach begs a question. If sustainable use is, in fact, sustainable – i.e., if it both conduces to the survival of the biological assets on which it depends and provides a livelihood for the people conducting it – why would international conservation donors need to underwrite it? Why would the local people who stand to profit from the exploitation of their biodiversity not decide to do so without the assistance of foreign donors (Simpson and Sedjo, 1996)?

Several answers might be proposed. First and least satisfactorily, one might suppose that local people do not know what is best for them. As this strains credulity and smacks of condescension, let us proceed to a second potential explanation: Local people may face constraints in technology transfer or capital investment that prevent them from exploiting otherwise profitable opportunities. There may well be merit to this argument, and a cogent model of such constraints and their implications for conservation policy has been proposed (Groom and Palmer, 2010). It should, however, be noted that such arguments may cut both ways. Constraints on capital and credit could prevent local people from investing in ecotourism hotels or bioprospecting laboratories, but such constraints might also prevent them from acquiring chainsaws or building hydroelectric dams (see Rowthorn and Brown (1999) for a formal model of the sometimes counterintuitive effects of lower interest rates on conservation). (One of the authors has heard a revealing, if possibly apocryphal, tale of a nontimber forest product collection ICDP in Madagascar that proved so successful in enhancing local incomes that villagers were able to purchase chainsaws.)

A third possibility is that sustainable exploitation of natural products would generally not be profitable absent subsidies from conservation donors. This possibility seems the most parsimonious explanation of the two facts that, first, sustainable use projects often do not rise indigenously, and second, that foreign donations for sustainable use projects are often made in the form of grants rather than loans.

If, then, subsidies are required to motivate the establishment of markets, the next question to ask is whether such subsidies are the most cost-effective way in which to achieve conservation objectives. Ferraro and Simpson (2002) present a

simple model arguing that they are not. Heuristically, if conservation donors care about conservation, they can generate more effect for each conservation dollar by paying for conservation directly rather than by paying for the purchase of assets (ecotourism facilities, bioprospecting laboratories, and investment in human capital) that only indirectly advance the conservation objective. It should be noted, however, that this result arises from a highly schematic representation of the economic incentives facing both local people and conservation donors, and makes no allowance for political and social factors that could be very important in real-world conservation decisions. Other authors have suggested a richer range of possibilities (see, e.g., Groom and Palmer, 2010), and it may not be surprising that recent research indicates that World Bank-funded development projects that incorporated biodiversity conservation goals were no less successful, but were less costly, than development-only projects (Kareiva *et al.*, 2008).

Creating New Markets for Biodiversity Services Using Public Policy

The innovations to engage the private sector in conservation just discussed focused on preserving biodiversity as a joint output used in the production of other goods and services, and were not always effective tools in stemming species loss. While these initiatives can work in certain circumstances, their limited successes suggests that mechanisms to harness conservation financing should target as directly as possible the desired outcomes – either preservation of specific species or habitat more generally.

Integrating market economics and biodiversity conservation provides other promising approaches to achieve this objective. As a result, enthusiasm for involving the private sector in conservation activities has shifted rather than waned. In recent decades, market-based policies have been adopted to address a variety of environmental and conservation priorities, with cap-and-trade programs such as the US Acid Rain Program for sulfur dioxide and the E.U. Emissions Trading System for greenhouse gases offering two widely known examples.

Market-based approaches like cap-and-trade programs create a regulatory requirement that firms or other entities covered under the program not exceed an aggregate environmental target (such as the amount of pollution emitted) and then allow the compliance obligation to be traded across emitters to meet the objective at lowest cost. While principles of market economics are used to design such systems, public policies (and possibly public financing) are needed to give them force. The approach can be applied to biodiversity conservation by setting a species- or habitat-based target, such as maximum quantity of fish harvested or minimum forested area preserved. Market-based mechanisms for conservation include not only quantity-based instruments (which rely on caps or quotas), but also price-based instruments such as pollution taxes or conservation subsidies.

Fishery management is one of the most prominent examples in which regulatory institutions have created market mechanisms to address species loss resulting from the "tragedy of the commons" (Hardin, 1968) under which open access to a common pool resource results in its overexploitation. "Catch share" systems create a maximum allowable harvesting limit and then assign harvesting rights to a limited number of participants.

Recall that the fundamental problem with protecting biodiversity in a market economy is that no one owns all of the attributes of biodiversity. Thus, it is impossible to structure transactions in which payments are made from one party to another to preserve biodiversity. This problem would, of course, be obviated if ownership rights could be assigned in the attributes of biodiversity. The most prevalent approach to fishery privatization is the Individual Transferable Quota (ITQ), a system in which a regulator sets a total allowable catch and allocates tradable catch shares to private owners in the market – essentially a cap-and-trade program for fisheries. ITQs and other catch share programs started to gain traction in the 1980s after decades of overharvesting and ineffective policy approaches led to severe declines in commercially valuable fish stocks. Prior to catch share systems, many countries sought to reverse fishery collapse via command-and-control policies restricting fishery inputs such as season length, number of vessels, and technology. Such approaches often had the unintended consequence of spurring a "race to fish" in which actual harvests were not limited, while fishers competed ever more aggressively for a dwindling supply. Based on the trajectory of overharvesting seen in the twentieth century, a widely cited study predicted global fishery collapse by 2048 (Worm *et al.*, 2006).

In contrast to command-and-control policies, ITQs employ a market-based approach in which regulators create and allocate property rights to fish harvests and allow those rights to be traded. Overall catch limits are set by regulators with input from ecologists, and programs can cover single or multiple species. Secure property rights give fishers an incentive to harvest sustainably, instead of out-competing other vessels to catch as much as possible before stocks disappear. Other similar systems include territorial use rights programs (essentially marine concession areas), which can be coupled with catch limits, and cooperatives, which allocate property rights to groups rather than individuals. The success of any particular catch share mechanism depends on ecological and socioeconomic factors. However, a firm, enforceable catch limit set at a sustainable level is an element critical to success that is common across these different mechanisms. Unfortunately, ecological advice is not always heeded in setting limits, as demonstrated by one study of the EU over 1987–2006, which found that policymakers routinely exceeded scientific recommendations when setting catch limits, often by large margins (Piet *et al.*, 2010).

ITQs represented more than 10% of ocean fish harvests as of 2008 (Chu, 2008). EDF (2010) has documented 275 catch share programs, of which 80% use tradable shares to create a fully market-based system (most nontradable catch share programs allocate rights to co-operatives or traditional groups). In total, such programs cover more than 500 fish species in 35 countries (EDF, 2010). ITQs and other catch share mechanisms remain concentrated in New Zealand, Australia, Iceland, Canada, and the US, but they have begun to gain traction in South America, Africa, and Asia more recently.

Using fishery regulation to create market-based incentives for conservation has had promising results to date. Costello

et al. (2008) found that commercial fisheries managed under ITQs were only half as likely to collapse as those without ITQs, based on fishery landings data from 1950 to 2003. Fishers under ITQs experience higher profits as a result of rebounding harvests and lower harvesting costs.

The Alaskan halibut and sablefish ITQ program provides an example of these successes (Bonzon *et al.*, 2010). In the 1980s, the Alaskan halibut fishery had fallen prey to an extreme "race for fish" situation in which overcapitalization of the vessel fleet led regulators to restrict the season to only a few days annually. The frenzy of activity during this curtailed season prompted concerns about worker safety conditions and inability to provide the market with fresh fish throughout the year, while catch limits were still exceeded. Since an ITQ system was introduced in 1995 (later expanded to include sablefish), profits have increased due to lower costs and a longer season that allows for more fresh fish sales. In addition, worker safety has improved, sablefish bycatch has decreased, and stocks are not overfished. The US National Marine Fishery Service sets the total catch limit annually and also coordinates with the even more comprehensive Integrated Groundfish ITQ across the border in British Columbia.

While catch share programs have arisen in recent decades in the context of commercial fisheries hard hit by overexploitation, they echo similar mechanisms that evolved among traditional societies. In fact, some contemporary catch share programs are the direct descendants of such systems, including customary fishery rights mechanisms in Papua New Guinea, Fiji, the Solomon Islands, and Vanuatu, as well as fishing co-operatives in Japan (Bonzon *et al.*, 2010). Although the property rights conferred by these traditional systems are not always tradable, they embody a core principle of market economics by effectively privatizing an open-access resource.

Fishery catch shares demonstrate how a market-based regulatory program can protect biodiversity in the case of a commercial product where the buyer side of the market – seafood consumers – is well established. However, as already noted, biodiversity provides not only commercially salable products, but also many other services of value to society for which there is no obvious buyer, including a suite of ecosystem services such as flood protection and maintenance of a robust genetic library. Moreover, these ecosystem services are the products of complex biological systems, whereas ITQs and related approaches typically focus on a few commercially salient elements of diverse systems. Market-based mechanisms can also be designed to capture and market the benefits of broader systems in order to finance biodiversity preservation, however. In the past decade, the paradigm of payments for ecosystem services (PES) has arisen in part to address this issue.

PES is an umbrella term that can apply to any number of arrangements between buyers of conservation services and sellers able to provide these services or to manage land to generate them. Such programs include tradable permits, conservation easements, direct land purchases, and other approaches. Nongovernmental organizations (NGOs) are frequently involved, providing technical, institutional, or financial support. The buyers who participate in PES programs can include governments themselves, or they can be private firms fulfilling a regulatory requirement to offset destructive activities or voluntarily opting in to display corporate social responsibility. Because the types of ecosystem services targeted by most PES initiatives are pure public goods such as flood protection and carbon sequestration, the potential for strictly voluntary contributions to support PES at a large scale seems unlikely, however. Government funding or regulatory requirements for industry participation are probably needed to realize the full potential of these programs.

PES offers a flexible but direct approach to conservation, allowing buyers to seek out biodiversity services where they can be provided at least cost. As already noted, direct payments for land uses that foster ecosystem services – and to an even greater extent, performance-based biodiversity payments – hold out promise to reach conservation targets more efficiently than previous mechanisms. As Ferraro and Kiss (2002) put it, "The basic principle is that the cheapest way to get something you want is to pay for what you want (e.g., protected rain forest), rather than pay for something indirectly related to it (e.g., capital for improving ecotourism)."

Several governments have spearheaded PES programs financed by public (or in some cases, international) funds. Wunder *et al.* (2008) documented eight government-financed PES programs from China to Europe to Latin America. Most of these programs transfer subsidy payments to farmers or other individual landowners in exchange for undertaking specific land use or agricultural practices that generate a variety of ecosystem services. The US Conservation Reserve Program, established in 1985, is the longest-running such scheme. It pays farmers to retire land in order to reduce soil erosion, improve water quality, harbor wildlife, and, more recently, sequester carbon. Mexico's Payment for Hydrological Environmental Services program, started in 2003, is similar in concept but focuses on preservation of forested areas for watershed and aquifer protection (though carbon sequestration cobenefits are also anticipated).

A still more recent class of initiatives that have attempted to link conservation sellers with private sector buyers includes conservation concessions. Conservation concessions have emerged in the past decade as a much-touted incentive-based biodiversity conservation tool primarily reliant on international financing. Essentially a hybrid of a conservation easement and a commercial concession, the mechanism allows investors (which to date have predominantly been international NGOs) to finance conservation by compensating would-be users for the gains they would have realized by exploiting their resources. Conservation International (CI) pioneered the idea, helping to establish initial projects in Peru (2001) and Guyana (2002). Since then, similar programs have been proposed or established in Mexico, Cameroon, Sierra Leone, Bolivia, Indonesia, and elsewhere by CI and other NGOs.

Conservation concessions and similar agreements offer an approach for conserving large tracts of publicly or communally owned land where host countries or communities are reluctant to establish new protected areas without compensation. Thus, their scale and supporting institutional arrangements are distinct from PES schemes that target individual landholders who control relatively small areas. Like easements, concessions restrict land uses, effectively transferring use rights without a transfer of land ownership. In theory, they

provide a mechanism for conservation to compete directly with timber, fishing, mining, and other extractive interests for access to desirable land, though in practice, they have been largely used as a preventive strategy to protect pristine habitats that have not yet come under direct extractive pressures.

While they offer the potential for large-scale habitat preservation at low transaction costs, results remain modest when compared to the extent of resource extraction in commercial concessions (Pearce, 2004). For instance, timber concessions in Central Africa frequently extend to millions of hectares (Karsenty, 2007), while conservation concessions are typically approximately or less than 100,000 ha each. However, the size of concessions does compare favorably with the land area conserved under national PES programs; Costa Rica's PES program had enrolled 270,000 ha of forest as of 2005, while Mexico's had targeted 600,000 ha by 2006 (Wunder *et al.*, 2008). A recent survey finds that momentum for biodiversity markets has been sustained despite the recent downturn in economic activity more generally (Madsen *et al.*, 2011).

Concessions can vary substantially in their implementation. Programs can permit different land uses, with some barring commercial activities and others allowing sustainable harvesting, ecotourism, and research. They might involve payments to local residents, community associations, regional or national governments, or some combination of these groups. They can also target diverse biomes from grasslands to temperate forests to aquatic habitats, though most projects are located in tropical forests. This diversity makes concessions potentially relevant in a wide variety of future applications and provides flexibility to be adapted if significant public or mandatory financing becomes available to supplement the voluntary private efforts that have supported concessions to date.

Proponents of market-based biodiversity conservation efforts typically tout their effectiveness and low costs relative to other policy approaches, but another advantage seen in both the government-financed PES and privately financed conservation concession examples is that tracts of undeveloped land can provide many ecosystem services besides harboring biodiversity. Marketing these "bundled" services could offer a means to leverage multiple sources of funding to preserve highly valued sites. Tropical forests in particular have the potential to supply carbon sequestration and hydrological regulation, both classified as regulating services by the Millennium Ecosystem Assessment, in tandem with biodiversity protection. Combining biodiversity preservation with provisioning or cultural services, like sustainable harvesting or ecotourism activities, might also be feasible. For example, some participants in Chile's area-based catch share program for abalone and other benthic organisms have combined fish harvesting with dive tourism (Bonzon *et al.*, 2010). However, the latter approach risks degrading pristine sites, making these activities incompatible with environmental goals (Terborgh, 1999).

Bundling biodiversity conservation with carbon sequestration offers particular promise as a means to access additional funding for conservation with the growing interest in REDD (Reducing Emissions from Deforestation and Forest Degradation in Developing Countries), which is being discussed under the United Nations Framework Convention on Climate Change (UNFCCC) (see OECD, 2008). This synergy would allow biodiversity conservation to take advantage of new markets created to address other externalities – in this case, climate change from greenhouse gas emissions. If REDD is adopted as a climate mitigation strategy under a post-Kyoto climate agreement, large sums of money could become available for reducing deforestation and degradation (e.g., Kindermann *et al.*, 2008).

Challenges to Creating Markets for Biodiversity

Tradable quota systems, conservation concessions, and other market mechanisms reduce costs relative to command-and-control regulation by creating incentives for those entities with the lowest costs of achieving biodiversity benefits to take the actual physical measures required to reduce environmental impact. Despite these advantages, market-based biodiversity conservation approaches are not invulnerable to criticism. The authors discuss three broad challenges associated with setting up such programs (though they are not all unique to market mechanisms and could also apply to traditional conservation approaches). These include implementation issues; equity and development considerations; and the high costs of conservation relative to the benefits – or at least, those benefits likely to be reflected in market mechanisms.

Applying market-based mechanisms to preserve biodiversity raises a number of nontrivial implementation challenges. In particular, biodiversity-rich habitat and associated ecosystem services are extremely heterogeneous; preserving land in a temperate forest gives rise to very different ecosystem services than a similarly sized tract in a tropical forest, and even within the same biome, habitat can vary substantially in the number and type of species supported. Although habitat area is a convenient metric, delineating tradable units in terms of area alone ignores these quality differences – a particular problem if poor quality habitat also costs the least to preserve. This issue has been recognized in the environmental economics literature for many years; the possibility that high-abatement-cost polluters might be concentrated in "hotspots" may necessitate further restrictions on trade in such areas or the imposition of different trading ratios depending on the location at which pollution is emitted (Tietenberg, 1985). Even in the case of catch share programs where compliance is measured in terms of species-specific harvest limits, complementary policies like marine reserves could be warranted to protect particularly sensitive aquatic habitats such as spawning grounds.

According to Peterson *et al.* (2008), "the more specifically ecosystems are categorized to match ecology, the more barriers to trade these ecological differences erect" (see also Chomitz *et al.*, 2004). A program in which local developers can trade obligations to maintain habitat for a particular endangered species might afford only limited cost savings if there is little flexibility in the choice of feasible habitat configurations. The cost of conserving a certain area of habitat, or even of sheltering a certain number of species, might be greatly reduced if we relaxed any requirements regarding the character of the habitat or the identity of the species. Alternatively, costs will likely be higher if restrictions are placed on trades to preserve particular locales or populations.

Besides the complications related to measuring and trading units of highly heterogenous biodiversity services, market-based programs for habitat conservation also raise concerns about additionality, leakage, permanence, and verification – four different issues that must be addressed to achieve real, long-term conservation gains relative to business-as-usual conditions (Wunder, 2007). Additionality relates to projects targeting land that would likely remain undeveloped even without the conservation program; in such a situation, the program has yielded no "additional" environmental gains. Many projects have focused on remote, sparsely settled areas to date because of limited funds and reluctance to displace land uses contributing to economic development (Barnes *et al.*, 2006). Because the opportunity cost of land is equal to the net present value of the stream of future earnings it is expected to generate, land that is likely to remain undeveloped well into the future has a low net present value, making it less costly to preserve – but also less threatened. Projects targeting these areas divert resources away from sites where they might have more of an impact. Market-based programs can be designed with selection criteria that reflect the likelihood of land conversion to increase additionality.

Even if conservation initiatives successfully preserve natural ecosystems on sites that would likely be converted to other uses, leakage – the displacement of extractive activities to other undeveloped land – can weaken aggregate conservation outcomes (Chomitz, 2002; Wunder, 2007; ICMM, 2005). It would do little good to secure the preservation of one tract of pristine ecosystem if, by doing so, development pressure were simply diverted to the next-most-advantageously located tract of pristine ecosystem.

Preventing leakage (also sometimes called "slippage") is particularly difficult when considering habitats that produce commodities such as timber and minerals that are exchanged in international trade. Any reduction in supply from the market could induce an increase in the price the product commands in international trade, which will, in turn, attract more supply from other sources. The elasticity of demand for harvested products is an important determinant of leakage; less leakage will occur when demand is elastic, while more is expected when demand is relatively inelastic. Fortunately, empirical evidence to date suggests that hectare-for-hectare leakage is unlikely (though biodiversity loss could still be high without complete leakage if activities are displaced from less- to more-diverse land). For example, an empirical study of the Conservation Reserve Program in the US found that for every 100 acres of land enrolled in the program, 20 acres of non-cropland was converted to agricultural production (Wu, 2000).

Permanence is another critical feature of conservation programs that might be in conflict with a market-based approach. The very flexibility of market-based conservation programs could become a drawback if such programs simply delay species extinction for one or two decades because land is eventually converted. Setting firm aggregate conservation targets and structuring contracts to apply long-term and allow the option to be renewed can help address this issue.

Monitoring and verification are also key implementation issues, given the remoteness of many conservation activities and the difficulty in measuring them. Aggregate conservation targets will have little meaning if land use practices and species survival outcomes cannot be credibly assessed. Technologies such as satellite-based monitoring can be helpful for verification in remote areas, and programs can be structured to provide payments only after certain commitments have been fulfilled by the sellers of conservation services.

Equity and development concerns present a different set of challenges from applying market-based approaches to biodiversity conservation. As is true in transactions for pure private goods, market signals lead to the most efficient outcome – the greatest overall monetary benefit at the lowest cost, given the preexisting distribution of resources. But such outcomes may not ensure the well-being of the worst-off members of society or foster a level playing field. For instance, fishery ITQs have confronted whether to auction off rights to the highest bidder, freely allocate them to fishers based on historical harvests, or distribute them using some other approach. This initial distribution of property rights has large implications for who reaps the benefits and bears the costs of the program, even though it does not affect the total monetary benefits and costs to society of instituting a catch share system.

In cases where conservation projects help local people establish secure property rights (whether to fish harvests, land, or other valuable resources), they can facilitate economic gains for host communities while improving environmental outcomes by shifting land or oceans away from an open access regime. For instance, devolving control over wildlife habitat to local communities encouraged economic development through sustainable resource management and ecotourism in Zimbabwe's CAMPFIRE program (Frost and Bond, 2008). Similarly, control over use rights to tracts of Guatemala's Maya Biosphere Reserve has allowed local communities to reap gains through sustainable forest management (Avery and Southgate, 2002).

When property rights are unaffected by PES or other market mechanisms, economic benefits experienced by communities are not expected to be appreciably different under conservation than other resource uses if payments cover but do not exceed true opportunity costs. Indeed, community welfare could even decrease if outsiders pursue property rights to key resources in anticipation of payments.

Even if the communities directly involved benefit, conservation projects could have regressive outcomes by taking land out of production and raising the costs of commodities on which the poor spend disproportionately more of their incomes. If conservation induces substantial land use changes, some of the changes in the prices of agricultural products, wages, and other affected commodities that result could exacerbate inequality (Zilberman *et al.*, 2008; Ross *et al.*, forthcoming).

Recent surges in global food prices in response to higher fuel prices, biofuel production, and demand from expanding economies suggest that potential food and resource price effects of large-scale conservation programs should be scrutinized. Food prices are expected to remain high over the coming decade (von Braun *et al.*, 2008). Food security is, of course, an absolutely vital concern of many people in developing countries. This is particularly true of those who barely sustain themselves at the margins of imperiled ecosystems. Conversely, fishery catch shares that increase the amount of biological resources available for direct human use through more

sustainable management could have beneficial effects on food availability. The general conclusion to be drawn from these examples is that conservation programs on a large enough scale – regardless of the mechanisms – are likely to affect the production and consumption of other goods, triggering a variety of positive or negative impacts on individuals within a market economy.

Even when market-based programs identify equitable and relatively low-cost approaches to conserve biodiversity, their costs can still be high – perhaps too high to achieve desired conservation outcomes. Once institutions and property rights have been established, biodiversity conservation enters the market economy as another service, and individuals must perceive the private benefits they will reap from supplying it to outweigh the opportunity costs.

Undeveloped lands are often arrayed in a haphazard patchwork, and the cost of their conservation can be high, as many areas are now accessible to transportation and services that make them more attractive for development. While the costs of acquiring and managing threatened habitat in some areas of the developing world have been estimated to be relative bargains at US $10 per hectare or less (Pimm *et al.*, 2001), it is these low-opportunity cost areas that raise concerns about additionality. Much land remains in areas that are not yet subject to high development pressure, but studies examining the potential opportunity costs of conservation have shown them to be substantial in many cases. Sandker *et al.* (2007) showed that logging a forest tract and planting oil palm would provide much greater economic benefits to both local communities and potential migrant laborers in Malinau district, Kalminantan, Indonesia over a 40-year timeframe. The authors' simulation results indicated that PES of US $25–50 million annually would be required to make up the opportunity cost of economic gains derived from timber and oil palm harvests.

Whether resources are sufficient to pay the costs of conservation will depend, at least in part, on how many of the benefits of conservation have been internalized by the program. Just as consumer products and services like sustainable coffee and ecotourism only reflect a portion of the ecosystem services provided by natural habitats, even market-based conservation programs mandated by governments could be insufficient to motivate sufficient biodiversity preservation if they do not reflect the full suite of services from which people around the globe derive fulfilment. For example, an economic assessment of Madagascar's Masoala National Park found that although conservation would generate greater social benefits globally, the financial returns to logging exceeded those from an ICDP (Kremen *et al.*, 2000). In the case of fishery catch share programs, it is possible at least in theory that it could be rational to drive low-growth fisheries to extinction under private management (Clark *et al.*, 2010). This result would not be socially optimal if the fish species provide other benefits to society besides food consumption.

Conclusion

Market-driven economic growth and biodiversity conservation have long been seen as a fundamental conflict. In recent decades, conservationists have attempted to change this perception by harnessing private markets to protect biodiversity. This convergence has followed two basic approaches. The first relies on private markets to capture individuals' willingness to pay for those aspects of biodiversity that are reflected in consumer goods and services. The second uses public policy tools to create market-based mechanisms to motivate the conservation of both open access resources and ecosystem services. Both strategies have seen some limited successes to date, and the second approach holds particular promise for conservation efforts that are both large-scale and cost-effective. Market-based approaches do not necessarily resolve all of the dilemmas that have plagued past species conservation efforts, however, from implementing credible programs to covering high costs to achieving equitable outcomes for the local communities affected by conservation. Any conclusions that might now be offered concerning the relationship of "market economy and biodiversity" must remain provisional, awaiting both greater experience with the application of market incentives and greater appreciation of their potential limitations.

See also: Biodiversity as a Commodity. Biodiversity, Human Well-Being, and Markets. Economic Value of Biodiversity, Measurements of. Economics of Agrobiodiversity. Economics of the Regulating Services. Ecosystem Services. Property Rights and Biodiversity. The Value of Biodiversity. Valuing Ecosystem Services

References

Avery D and Southgate D (2002) Endangering tropical forests with good intentions. *Knight Ridder Tribune*, http://www.cgfi.org/2002/05/30/endangering-tropical-forests-with-good-intentions/ (accessed 30 May 2002).

Barnes A, Ebright M, Gaskin E, and Strain W (2006) *Conservation Concession: Protecting Forest Ecosystems with Payments for Ecosystem Services in Guyana.* Columbia University, School for International and Public Affairs. http://www.translinks.org/DesktopModules/Bring2mind/DMX/Download.aspxTabId=409&language=en-US&Command=Core_Download&EntryId=140&PortalId=11

Bonzon K, McIlwain K, Strauss CK, and Van Leuvan T (2010) Catch share design manual: A guide for managers and fishermen. *Environmental Defense Fund,* p. 105.

von Braun J, Ahmed A, Asenso-Okyere K, *et al.* (2008) High food prices: The what, who, and how of proposed policy actions. *IPFRI Policy Brief.* Available from: http://ifpri.org/PUBS/ib/FoodPricesPolicyAction.pdf

Chomitz K (2002) Baseline, leakage and measurement issues: How do forestry and energy projects compare? *Climate Policy* 2(1): 35–49.

Chomitz K and Kumari K (1996) The domestic benefits of tropical forests: A critical review emphasizing hydrological functions. *World Bank Policy Research Working Paper,* No. 1601.

Chomitz K, Thomas K, and Brandao AS (2004) Creating markets for habitat conservation when habitats are heterogeneous. *World Bank Policy Research Working Paper,* No. 3429.

Chu C (2008) Thirty years later: The global growth of ITQs and their influence on stock status in marine fisheries. *Fish and Fisheries* 10(2): 217–230.

Clark CW, Munro GR, and Sumaila UR (2010) Limits to the privatization of fishery resources. *Land Economics* 86(2): 209–218.

Costello C, Gaines SD, and Lynham J (2008) Can catch shares prevent fisheries collapse? *Science* 321: 1678–1681.

Costello C and Ward M (2006) Search, information, and biodiversity prospecting. *Journal of Environmental Economics and Management* 52(3): 615–626.

Craft AB and Simpson RD (2001) The social value of biodiversity in new pharmaceutical product research. *Environmental and Resource Economics* 18(1): 1–17. (January).

Daily G and Ellison K (2002) *The New Economy of Nature: The Quest to Make Conservation Profitable*. Washington, DC: Island Press.
De Long and Bradford J (1998) Estimating World GDP, One Million B.C. – Present. Available from: http://econ161.berkeley.edu/TCEH/1998_Draft/World_GDP/Estimating_World_GDP.html
Diamond J (1999) *Guns, Germs, and Steel*. New York: W. W. Norton & Company.
EDF (2010) Catch Shares Around the World. Available from: http://www.edf.org/content_images/catch-shares-world-map-hi.jpg
Eisner T (2003) Hard times for chemical prospecting. *Issues in Science and Technology* 19(4): 47–48.
Ferraro PJ and Kiss A (2002) Direct payments to conserve biodiversity. *Science* 298(5599): 1718–1719.
Ferraro PJ and Simpson RD (2002) The cost-effectiveness of conservation payments. *Land Economics* 78(3): 339–353.
Feshbach M and Friendly A (1993) *Ecocide in the USSR: Health and Nature Under Siege*. New York: Basic Books.
Firn RD (2003) Bioprospecting – Why is it so unrewarding? *Biodiversity and Conservation* 12: 207–216.
Frost PGH and Bond I (2008) The CAMPFIRE programme in Zimbabwe: Payments for wildlife services. *Ecological Economics* 65(4): 776–787.
Groom B and Palmer C (2010) Cost-effective provision of environmental services: The role of relaxing market constraints. *Environment and Development Economics* 15(2): 219–240.
Hardin G (1968) The tragedy of the commons. *Science* 162(3859): 1243–1248.
Heilbroner Robert L (1999) *The Wordly Philosophers: The Lives, Times and Ideas of the Great Economic Thinkers*. New York: Touchstone.
Honey M (1999) *Ecotourism and Sustainable Development: Who Owns Paradise?* Washington, DC: Island Press.
International Council on Mining and Metals (ICMM) (2005) *Biodiversity Offsets: A Proposition Paper*. Available from: http://www.forest-trends.org/biodiversityoffsetprogram/BBop%20library%202/International/Not%20Printed/BioDiv%20Offsets%20Proposition%20Paper.pdf
International Union for the Conservation of Nature, United Nations Environment Program, and World Wildlife Fund (IUCN, *et al.*) (1981) *World Conservation Strategy: Living Resource Conservation for Sustainable Development*. Switzerland: Gland.
International Union for the Conservation of Nature, United Nations Environment Program, and World Wildlife Fund (IUCN, *et al.*) (1991) *Caring for the Earth: A Strategy for Sustainable Living*. Switzerland: Gland.
Kareiva P, Chang A, and Marvier M (2008) Development and conservation goals in World Bank projects. *Science* 321: 1638–1639.
Karsenty A (2007) *Overview of Industrial Forest Concessions and Concession-based Industry in Central and West Africa and Considerations of Alternatives*. CIRAD.
Kindermann G, Obersteiner M, Sohngen B, *et al.* (2008) Global cost estimates of reducing carbon emissions through avoided deforestation. *Proceedings of the National Academy of Sciences* 105(30): 10302–10307.
Kirby B (2004) *North Korea's Environmental Crisis*. Available from: http://news.bbc.co.uk/2/hi/science/nature/3598966.stm
Kolstad CD (2000) *Environmental Economics*. New York: Oxford University Press.
Kremen C, Niles JO, Dalton MG, *et al.* (2000) Economic incentives for Rain Forest conservation across scales. *Science* 288: 1828–1832.
Madsen B, Carroll N, Kandy D, and Bennett G (2011) *Update: State of Biodiversity Markets*. Washington, DC: Forest Trends. Available from: http://www.ecosystemmarketplace.com/reports/2011_update_sbdm
Mason CF (2006) An economic model of eco-labeling. *Environmental Modeling and Assessment* 11(2): 131–143.
McIlwain C (1998) When rhetoric meets reality in debate on bioprospecting. *Nature* 392(6676): 535–540.
Oates JF (1999) *Myth and Reality in the Rain Forest: How Conservation Strategies are Failing in West Africa*. Berkeley, CA: University of California Press.
OECD (2008) Chair's summary on the OECD workshop on capturing carbon and biodiversity benefits to reduce deforestation. *A Joint Workshop of the Working Group on the Economic Aspects of Biodiversity and the Annex I Expert Group on the UNFCCC*. Available from: http://www.oecd.org/dataoecd/42/61/41241295.pdf
Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.
Pearce D (2004) Environmental creation: Saviour or oversell? *Portugese Economic Journal* 3(2): 115–144.
Pearce D and Moran D (1994) *The Economic Value of Biodiversity*. London: Earthscan.
Peters C (1994) *Sustainable Harvest of Non-timber Plant Resources in Tropical Moist Forest: An Ecological Primer*. Biodiversity Support Program, Washington.
Peterson AL, Hill C, and Gallagher LA (2008) *Balancing Biodiversity: A Global Instrument for Meeting the 2010 Biodiversity Target. Nicholas School of the Environment and Earth Sciences*. Duke University. Available from: http://dukespace.lib.duke.edu/dspace/handle/10161/551 (accessed 29 September 2008).
Piet GJ, van Overzee HMJ, and Pastoors MA (2010) The necessity for response indicators in fisheries management. *ICES Journal of Marine Science* 67: 559–566.
Pimm SL, Ayres M, Balmford A, *et al.* (2001) Can we defy nature's end? *Science* 293(5538): 2207–2208.
Rausser GC and Small AA (2000) Valuing research leads: Bioprospecting and the conservation of genetic resources. *Journal of Political Economy* 108(1): 173–206.
Ross M, Depro B, and Pattanayak SK (2009) Assessing the economy-wide effects of the PSA program. In: Platais G and Pagiola S (eds) *Ecomarkets: Costa Rica's Experience with Payments for Environmental Services*. Washington, DC: The World Bank.
Rowthorn B and Brown Jr. G (1999) When a high discount rate encourages biodiversity. *International Economic Review* 40(2): 315–332.
Sandker M, Suwarno A, and Campbell BM (2007) Will forests remain in the face of oil palm expansion? Simulating change in Malinau, Indonesia. *Ecology and Society* 12(2): 37. URL: http://www.ecologyandsociety.org/vol12/iss2/art37/
Sedjo RA and Swallow S (2002) Voluntary eco-labeling and the price premium. *Land Economics* 78(2): 272–284.
Simpson RD and Sedjo R (1996) Paying for the conservation of endangered ecosystems: A comparison of direct and indirect approaches. *Environment and Development Economics* 1(2): 241–257.
Simpson RD, Sedjo RA, and Reid JW (1996) Valuing biodiversity for use in pharmaceutical research. *Journal of Political Economy* 104(February): 163–185.
Smith Adam (1904) [1776]. *Wealth of Nations*. London: Methuen & Co.
Southgate D (1998) *Tropical Forest Conservation: An Economic Assessment of the Alternatives for Latin America*. Oxford: Oxford University Press.
Terborgh J (1999) *Requiem for Nature*. Washington: Island Press.
Terborgh J, van Schaik C, Davenport L, and Rao M (eds.) (2002) *Making Parks Work: Strategies for Preserving Tropical Nature*. Washington: Island Press.
The World Bank and the State Environmental Protection Agency of China (World Bank) (2007) *Cost of Pollution in China: Economic Estimates of Physical Damages*. Washington: The World Bank.
Tietenberg T (1985) *Emissions Trading: An Exercise in Reforming Pollution Policy*. Washington, DC: Resources for the Future.
Weber M (1958) 1904. *The Protestant Ethic and the Spirit of Capitalism*. New York: Charles Scribner's Sons.
Wells M, Brandon K, and Hannah D (1992) *People and Parks: Linking Protected Area Management with Local Communities*. Washington: World Bank.
Wells MS, Guggeheim P, Jepsen A, Khan, and Wardojo W (1999) *Investing in Biodiversity: A Review of Indonesia's Integrated Conservation and Development Projects*. Washington: The World Bank.
Worm B, Barbier EB, Beaumont N, *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.
Wu JJ (2000) Slippage effects of the conservation reserve program. *American Journal of Agricultural Economics* 82(4): 979–992.
Wunder S (2007) The efficiency of payments for environmental services in tropical conservation. *Conservation Biology* 21(1): 48–58.
Wunder S, Engel S, and Pagiola S (2008) Taking stock: A comparative analysis of payments for environmental services programs in developed and developing countries. *Ecological Economics* 65(4): 834–852.
Zilberman D, Lipper L, and McCarthy N (2008) When could payments for environmental services benefit the poor? *Environment and Development Economics* 13(3): 255–278. (June).

Mass Extinctions, Concept of

J John Sepkoski Jr.[†], University of Chicago, IL, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 97–110, © 2001, Elsevier Inc.

Glossary

Benthos Organisms living on or in sediment, below the water.

Extinction The disappearance of a species upon death of its last surviving individual. In the fossil record, this is treated as the last fossil occurrence of individuals of a species.

Foraminifera An order of animal-like protists, many of which secrete calcareous skeletons ("tests").

Mass extinction The simultaneous extinction of a disproportionate number of species over timescales of 10^0 to 10^6 years resulting in loss of biodiversity.

Phanerozoic The geological interval of abundant animal fossils, beginning approximately 545 million years ago.

Stratigraphic section An outcrop of rock (with fossils in this case) or a drilled core. The term can also refer to a composite for a region in which fossil ranges and stratigraphic events have been summarized into a synthesized rock column.

Tetrapod vertebrates Vertebrate animals with four limbs (or vertebrates that have evolved from such animals, such as snakes).

History of the Concept

The concept of extinction of species goes back at least several centuries. The extirpation of aurochs (wild relatives of cattle) and disappearance of lions from Europe were well known in the era of enlightenment and ascribed to human interference. The fact that species could become extinct from nonhuman causes was promoted by Cuvier at the end of the 18th century through his exquisitely detailed studies of mammalian fossils of the Paris Basin. His arguments were not accepted by all intellectuals at the time, and, in fact, Thomas Jefferson, the third president of the United States, doubted species could disappear before humans; he assigned Lewis and Clark a secondary mission in their explorations of the American northwest to search for living mammoths and mastodons.

The division of sequences of sedimentary rock into geologic systems by British geologists and paleontologists in the first half of the 19th century reflected a concept of major changes in marine faunas between these still-used time periods. But the first quantitative depiction of mass extinctions—major declines in biodiversity followed by recovery—appears to be Phillips's (1860) count of known numbers of fossil species and interpretive graphing of massive drops in diversity between the Paleozoic and Mesozoic eras and the Mesozoic and Cenozoic eras (terms he coined; see **Figure 1**).

The study of mass extinctions rested largely in limbo from Phillips's pioneering work into the mid-20th century. This was perhaps because of emphasis on documenting evolutionary continuity in the fossil record and an assumption of substantive uniformitarianism, inherited from Lyell. However, with accumulation of paleontological data, the greatest of all Phanerozoic mass extinctions—the end-Permian, or "Permo-Triassic," event—could not remain unnoticed. Schindewolf (1963) wrote a seminal paper discussing this event and invoking lethal radiation from an extraterrestrial catastrophe of a nearby supernova explosion. In response, Newell (1967) carefully counted fossil taxa (mostly described families) and argued that there were at least five events of mass extinction in addition to the Permo-Triassic. These papers set the stage for modern studies: examining detailed biostratigraphic data on local species disappearance and global compilations of taxonomic ranges.

Despite the contributions of Schindewolf and Newell, work on mass extinctions remained largely a "cottage industry" among paleontologists until 1980. Workers would examine one of Newell's events (usually in isolation of others) and posit some associated physical event as the cause, such as fall of sea level, or invent ad hoc hypotheses, such as heavy metal poisoning in the oceans as a result of mountain building.

Maturation of the study of mass extinction came with the bold hypothesis of Alvarez *et al.* (1980) that the Mesozoic-Cenozoic event, recognized 120 years before by Phillips, was caused by impact of a 10-km meteorite. The initial evidence of Alvarez and coworkers was concentration of the rare terrestrial element, iridium, at solar-system abundances in a clay layer at the Cretaceous-Tertiary boundary. But the hypothesis implied other testable questions: (a) Is there additional physical evidence of an impact at the Cretaceous-Tertiary boundary and (b) Is the abruptness of biological extinctions at the event consistent with a catastrophe induced by meteorite impact? Impressive evidence affirming the first question has been assembled, including global identification of the iridium-rich clay layer (in both marine and terrestrial stratigraphic sections); presence of shocked mineral grains (e.g., quartz), microspherules of shock-melted minerals, and abundant soot in the global clay layer; and the presence at 65Ma of the largest

[†]Deceased.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00091-5

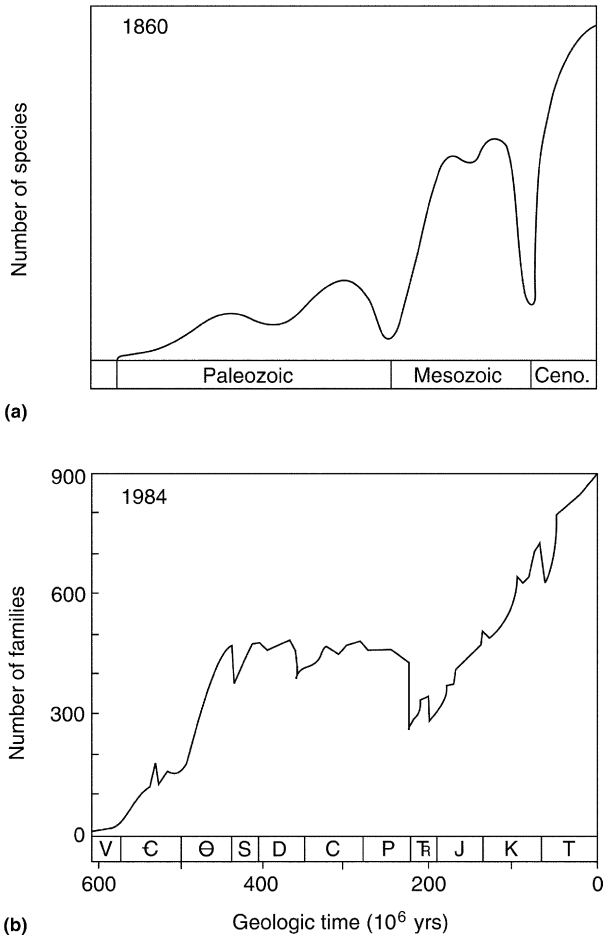


Figure 1 The first diversity curve for fossil organisms (a) compiled by Phillips (1860) compared to a late-20th century diversity curve for fossil marine animals (b). While Phillips' curve was generalized and based on comparatively few described taxa, the end-Permian mass extinction (between the Paleozoic and Mesozoic eras) and end-Cretaceous mass extinction (between the Mesozoic and Cenozoic ("Ceno") eras) were recognized. The lower graph illustrates three additional mass extinctions among those recognized by Newell (1967): the end-Ordovician, Late Devonian, and end-Triassic. Abbreviations along the abscissa for geologic time in (b) and in other figures are V = Vendian; C = Cambrian; O = Ordovician; S = Silurian; D = Devonian; C = Carboniferous; P = Permian; Tr = Triassic; J = Jurassic; K = Cretaceous; T = Tertiary. Reproduced with permission from Sepkoski Jr. JJ and Schopf JW (1992) The Proterozoic fossil record: Special problems in analyzing diversity patterns. In: Schopf JW and Klein C (eds.) *The Proterozoic Biosphere: A Multi-Disciplinary Study*, pp. 525–527. Cambridge: Cambridge University Press.

impact crater known from the Phanerozoic in the Yucatan, Mexico. Answers to the second question have proved more difficult and will be considered later.

The recognition of extinction events in the geologic past demands comparison to modern extinctions, from the auroch to the myriad other species documented to have disappeared from human interference. Questions include not only how modern rates of extinction compare to those in the fossil record but also the consequences of extinction: What kinds of organisms are most susceptible to extinction and what are the patterns of biotic recovery after the pressure of extinction has been removed? These questions will be considered further later in this discussion.

Models of Mass Extinction

A variety of patterns of extinction have been observed in the fossil record around events, ranging from abrupt termination of species at an extinction level to gradual disappearance up to it, and perhaps beyond, the extinction level. Three scenarios for mass extinction have been proposed based on empirical observations: abrupt extinction, gradual extinction, and stepwise extinction.

Abrupt Extinction

This is the pattern hypothesized by Alverez *et al.* (1980) of species disappearing in a geologic instant (which could in fact be 10^1 to 10^3 years). Observed declines in diversity before the event (such as seen in detailed records of foraminifera or broader records of dinosaurs before the Cretaceous-Tertiary) are a result of the Signor-Lipps sampling effect (discussed later).

Gradual Extinction

The fossil record is taken on face value, especially if there has been extensive sampling around the horizon of extinction. Slow attrition of species up to the end of a mass extinction has been claimed for extensively sampled foraminifers around the Cretaceous-Tertiary event, where several large, but rare foraminifers seem to disappear below the stratigraphic boundary and some small, generalized foraminifers occur above the boundary.

Stepwise Extinction

The fossil record is again taken at face value but exhibits a series of pulses of species terminations. This pattern has been hypothesized for situations such as the Late Devonian mass extinction ("upper Kellwasser Event") where intensively sampled taxa of different groups appear to disappear in small pulses separated by 10^4 to 10^5 years. The model can be expanded to intervals such as the end-Ordovician (the second largest marine mass extinction of the Phanerozoic) when many trilobites and other marine animals of tropical areas appear to become extinct at the onset of major glaciation, and then, perhaps 10^6 years later, surviving deep-shelf benthos disappear as normal conditions of low oxygen return with the end of glaciation. (In this case, each pulse of extinction could be dissected to determine if it had been abrupt, gradual, or stepwise at finer timescales.)

These scenarios need to be distinguished from the low levels of extinction that are observed in all geologic intervals. These levels are normally termed "background extinction" to distinguish them from events of mass extinction. Background extinction for marine animals appears to decline through Phanerozoic time. Thus, smaller extinction events are more obvious in Newell-type data over the Mesozoic and Cenozoic eras than during the early Paleozoic when background rates were high.

Interpreting Data from the Fossil Record

The fossil record provides direct evidence of previous mass extinction but only an incomplete accounting because of differences in preservability of organisms (e.g., bivalve mollusks versus polychaete worms) and in scientific sampling (e.g., Europe and North America versus Antarctica). With an incomplete record, observed last occurrences of fossil species are only a minimum estimate of actual times of extinction. This consideration was formalized for mass extinction by Signor and Lipps (1982) who modeled how observed terminations of species would appear around an abrupt extinction (**Figure 2**). With less intensive sampling or less complete preservation, the expectation is a pattern of gradual disappearance of fossil species up to a boundary of abrupt mass extinction. This sampling pattern holds true whether one is examining detailed stratigraphic sections of fossils or analyzing compilations of fossil taxa, like those of Newell.

Raup presented a very intuitive example of the Signor-Lipps effect (**Figure 3**). He used data collected for ammonites below the Cretaceous-Tertiary boundary to investigate how the fossil record of extinction would have appeared if the mass extinction had occurred at an interval 100 m lower in the stratigraphic section. (A meteorite impact or other unpredictable catastrophe potentially could occur anywhere in a stratigraphic section.) The decline in observed diversity up to the artificial event does not appear substantially different from the gradual decline of diversity actually observed below the true Cretaceous-Tertiary boundary. This suggests that problems of variable preservation and incomplete sampling of fossils can indeed influence empirical patterns of species disappearance around extinction events, hindering discrimination among the three models of mass extinction.

More powerful statistical methods have been developed to investigate how apparent last occurrences of fossil species may relate to actual times of extinction. These methods involve calculating "confidence intervals" on the time of last sampled occurrence of a species (**Figure 4**). The basic concept is that the last fossil occurrence of a species that is rarely sampled is a poorer predictor of actual time of extinction than the last occurrence of a species that is densely sampled. Using various models of the density distribution of sampling of a fossil species, probability statements can be made about how far true extinction lies above the last observation of a species.

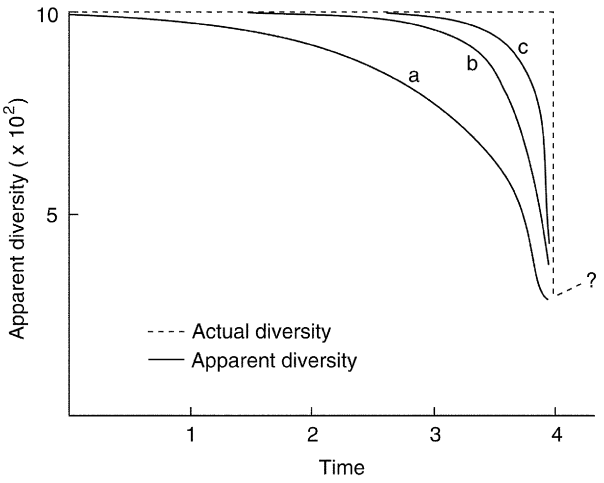


Figure 2 Hypothetical and calculated diversity curves reflecting the "Signor-Lipps effect"—that is, imperfect sampling around a catastrophic mass extinction. The dashed curve represents true diversity, which is treated as constant until time unit 4 at which there is a catastrophic mass extinction reducing diversity by three-quarters. If fossil taxa are not sampled up to the times of their true extinctions, declines in diversity will appear more or less gradual, depending on the intensity of sampling. Curve a represents the poorest sampling, and curve c the best (three times better than a). Reproduced with permission from Signor PW and Lipps JH (1982) Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. *Geological Society of America Special Paper* 190: 291–296.

Magnitudes of Mass Extinction

Questions of abrupt, gradual, or stepwise extinction involve patterns in the fossil record resolved over 10^3 to 10^5 years (encompassing the range from the late Pleistocene extinctions of large mammals to the historical extirpation of species). On larger timescales, general magnitudes of mass extinction can be measured from global fossil data. These data are best for the marine record of animals from continental shelves and seas and fall into roughly three classes of magnitude (**Figure 5**).

The End-Permian Mass Extinction

This class of magnitude stands alone in its effects on the biota (Erwin, 1993). Compilations of taxa lost indicate that more than 50% of animal families and 80% of genera in the oceans became extinct. Extrapolations of species loss have been attempted, using ecological rarefaction (how many species would be lost given measured declines in genera or families, assuming some distribution of species within higher taxa); results range from 90 to 96% loss of marine species. This loss of marine biodiversity at the end-Permian is unprecedented. Recent work suggests that on land important groups, including insects, tetrapod vertebrates, and plants, also experienced substantial declines in diversity.

Four Other Events of Marine Mass Extinction

This class of magnitude eliminated substantial proportions of marine animals and seem to have had nearly equal magnitudes: the end-Ordovician, Late Devonian, end-Triassic, and end-Cretaceous events. (The occurrence of these events at or near the end of geologic periods reflects the use of faunal change to define intervals of geologic time.) The four events have measured family extinction in the oceans of 15 to 25% and extrapolated species extinctions of 64 to 85%.

Other Intervals of Unusual Amounts of Extinction

This class of magnitude is now termed "extinction events" (**Figure 6**). These appear in more detailed compilations of the kind Newell made as well as in precise biostratigraphic and paleogeographic analyses. Paleogeographic analyses suggest that many of these third-order events were not global in extent

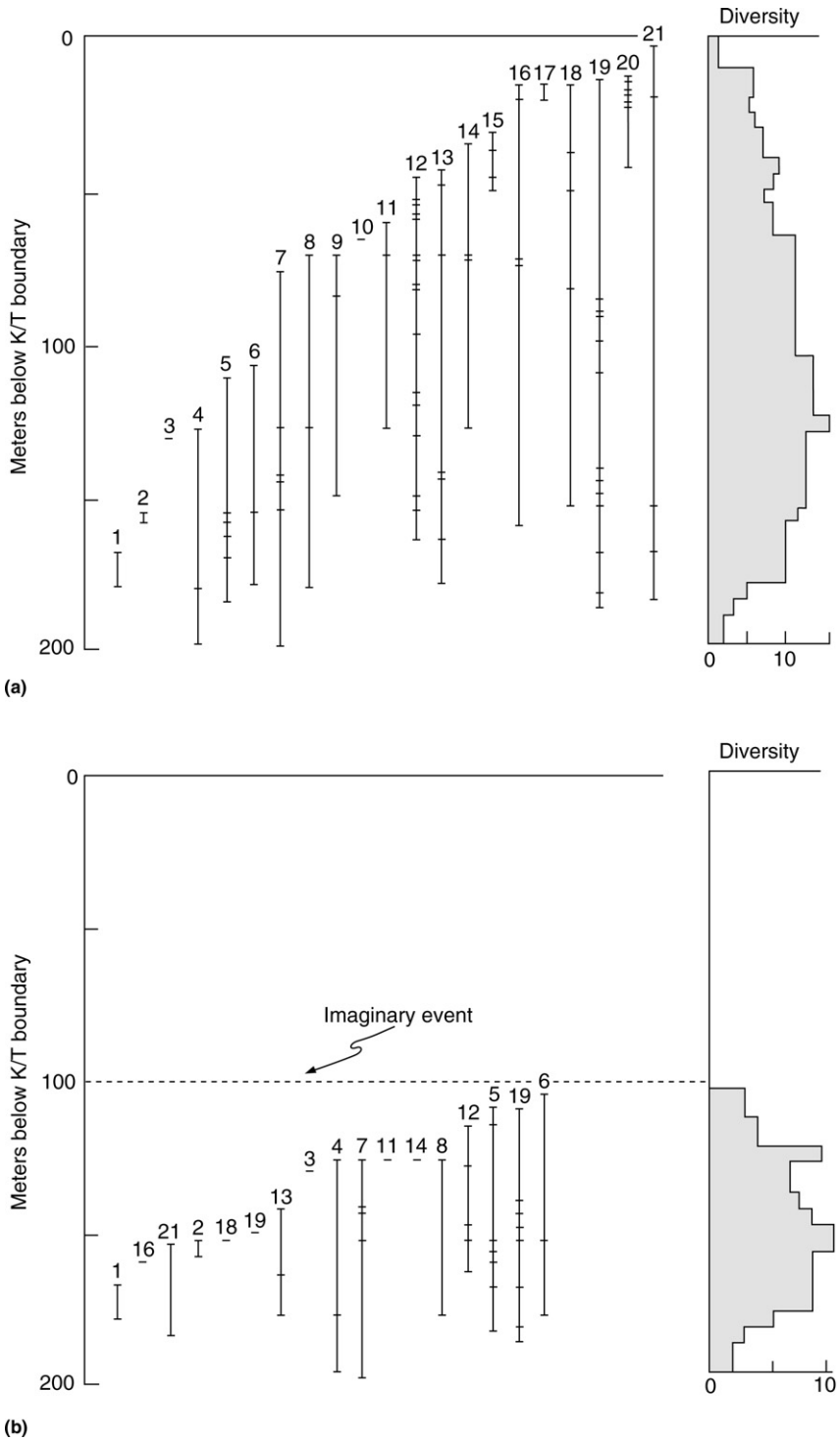


Figure 3 Apparent gradual extinction at an imaginary sudden extinction event. (a) Actual observed geologic ranges of 20 fossil species. The *y*-axis is stratigraphic position, in meters, below the Cretaceous/Tertiary boundary. Vertical lines show the total observed ranges of the species and horizontal ticks indicate horizons at which they were actually sampled. The curve at the right, labeled ''Diversity,'' is the sum of the ranges of the species. (b) Resultant ranges if an imaginary catastrophic extinction were imposed at 100 m. Apparent last occurrences of species again are graded below the mass extinction and diversity appears to decline, both as a result of species being irregularly sampled through the geologic interval. (Note that the species in b have been reordered based on their highest geologic occurrence.) Reproduced with permission from Sepkoski Jr. JJ and Koch CF (1996) Evaluating paleontologic data relating to bio-events. In: Walliser OH (ed.) *Global Events and Event Stratigraphy*, pp. 21–34. Berlin: Springer.

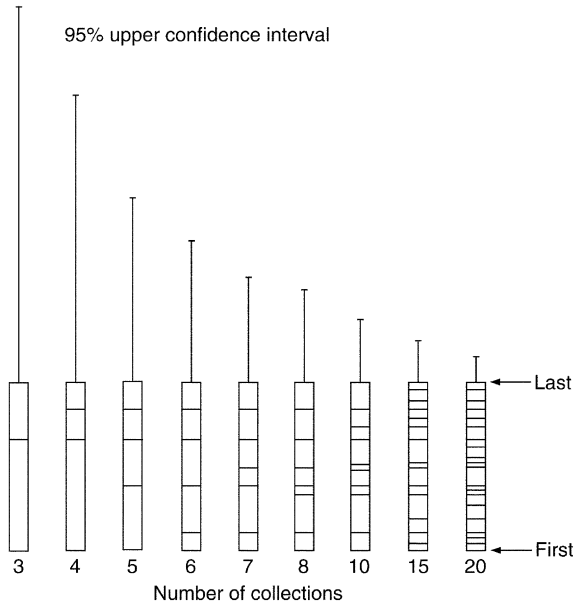


Figure 4 Confidence intervals on times of extinction based on the density of sampling of fossil species. The bars represent observed ranges of fossil species with horizontal ticks (including bottom and top of the bars) representing samples of the species. Vertical lines capped by ticks are calculated relative intervals during which there is a 95% probability that extinction actually occurred (thus allowing a 5% probability that real extinction actually occurred higher). The probable position of true extinction is closest to the position of last observation when species are densely sampled, but true extinction can still be expected to be somewhere above the position of last observation. Reproduced from Sepkoski Jr. JJ and Koch CF (1996) Evaluating paleontologic data relating to bio-events. In: Walliser OH (ed.) *Global Events and Event Stratigraphy*, pp. 21–34. Berlin: Springer.

or taxonomic effects, unlike the two previous categories. Examples include the Cambrian biomere events, marked by nearly complete extinction of North American trilobites but with uncertain effects in other parts of the world or on other taxa (although some events are recognized in Australia, China, Siberia, and Europe); the lower Toarcian event affecting mollusks in Europe but not around the Pacific basin; and the Pliocene event, again documented for mollusks, in the south temperate Northern Atlantic Ocean and Mediterranean Sea but not evident in the tropics or in the Pacific.

There are extinction events that could be classified at lower ranks. For example, it has been demonstrated that marine animal communities persist on order of 10^6 years during the middle Paleozoic, punctuated by rapid changes in faunal composition (Miller, 1997). This raises the question of Raup (1991a) as to whether extinction events represent a continuum between rare major events and frequent small events (discussed later).

Hypothesis of Periodicity

Given the varying magnitudes and geographic extents of extinction events, one would expect differing forcing agents. With a plethora of forcings acting independently over geologic time, it would be expected that extinction events would be distributed at random through the fossil record. Thus, it came as a surprise when Fischer and Arthur (1977) and Raup and Sepkoski (1984) observed that extinction events of first through third rank appeared regularly distributed through time. Raup and Sepkoski performed extensive statistical analyses of Newell-type data and concluded there was a strong periodicity of 26 myr for events during the Mesozoic and Cenozoic eras (**Figure 7**). This suggested some sort of clocklike mechanism behind mass extinction with a periodicity unknown in terrestrial processes. Because one of the periodic events was the end-Cretaceous mass extinction, Raup and Sepkoski speculated that the clocklike mechanism might be extraterrestrial.

This speculation engendered both intriguing hypotheses from astronomers and geologists and critical scrutiny of data and statistical methods from paleontologists, geologists, and statisticians (Raup, 1991b; Sepkoski, 1989). The best-known hypothesis is Nemesis, sometimes called "Shiva;" this is an hypothesized small binary companion to the sun with a large and eccentric orbit that brings it through the Oort Cloud of comets beyond the planets every 26 myr. Nemesis' gravitational perturbation scatters up to 10^9 comets of which 10^1 to 10^2 might impact the earth, disrupting climate and causing mass extinction.

Nemesis has not been observed, and astronomical models cast doubt on the possibility that a small star could maintain a stable orbit at large distance from the sun through the 4.6 billion-year history of the solar system. Also, there is direct evidence of extraterrestrial impact for only a few of the periodic extinctions, despite intensive investigation: Shocked quartz has been found at several end-Triassic localities, and concentrated iridium, microtektites (annealed globules of impact melt), and a crater have been found in the Late Eocene, although their temporal correspondence to extinctions remains unclear. Searches of this nature are difficult and time consuming, and one should not be surprised if more discoveries are made.

Questions about data and analyses with respect to the hypothesis of periodicity have been largely technical (Sepkoski, 1989). Questions of data concern the veracity of fossil families and genera as indicators of extinction in the geological past and the accuracy of geochronological dating of intervals of extinction. Statistical questions involve problems of analyzing data that are not sampled evenly through time (geologic intervals vary in duration; **Figure 7**) and problems of significance when performing multiple tests when searching for periodicities of best fit. Some of these questions entered new analytic ground, leading to new kinds of analyses, but not all analytical issues have yet been resolved.

Thus, a nonrandom, periodic distribution of extinction events remains a hypothesis on the table, with neither a clear forcing mechanism nor a definitive analytical test.

The Kill Curve and Self-Organized Criticality

Raup (1991a) treated extinction in the fossil past not as a hierarchy of events but as a continuum, in which the obvious events of large magnitude are rare relative to less obvious

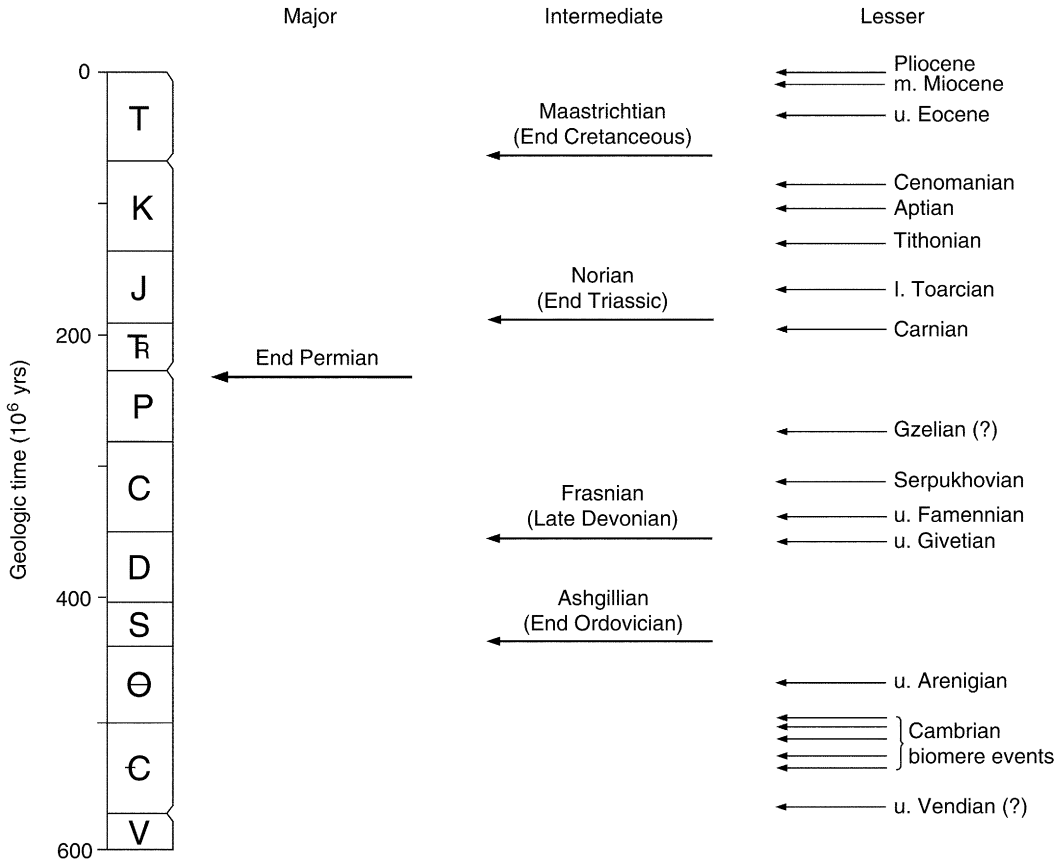


Figure 5 Time distribution of notable extinction events in the marine fossil record, classified as major ("first order"), intermediate ("second order"), and lesser ("third order") events. Only major and intermediate events are normally referred to as "mass extinctions." Abbreviations are mostly as in **Figure 1**; l = lower, u = upper. Names refer to stratigraphic intervals (mostly form stages). (?) indicates an event that has not been well documented. Reproduced with permission from Sepkoski Jr. JJ (1992) Phylogenetic and ecologic patterns in the Phanerozoic history of marine biodiversity. In: Eldredge N (ed.) *Systematics, Ecology, and the Biodiversity Crisis*, pp. 77–100. New York: Columbia University Press.

events of small magnitude. An analogy is to the historical record of river floods, in which the 1000-year event stands out but is really part of a spectrum with smaller decadal or even yearly events. The 1000-year event does not occur regularly but rather has a chance of 0.001 of occurring during any given year. Raup used a standard function for these kinds of events to analyze probabilities of species extinction of a given magnitude during some duration of geologic time (**Figure 8**): the longer the duration, or "waiting time," the larger a pulse of extinction that can be expected. The limitation of Raup's analysis is that it was based on generalized geological intervals, the same as illustrated in **Figures 6** and **7**, and cannot distinguish between sudden events, such as observed at the end of the Cretaceous, and cumulative small events, such as may also have occurred during the last stratigraphic stage of the Cretaceous during the 9 million years prior to the impact-induced mass extinction.

Raup's analysis was posited on an assumption that external perturbations of varying magnitude caused most extinction through geologic time, and he in fact demonstrated a linear relationship between his empirically based kill curve and the similar curve established for the flux of meteorites of varying magnitude (Raup, 1992). Other approaches have posited extinction to be a result of internal dynamics of the biota over time. Solé *et al.* (1997), for example, analyzed Phanerozoic time series of extinction and diversity fluctuation to determine if there was evidence of "self-organized criticality." The prediction was that a power series from fourier transforms of the data would exhibit a linear $1/f^x$ relationship when power was plotted logarithmically against the logarithm of frequency, f. The exponent, x, is expected to be between 1 and 2, which is indeed what was found.

Self-organized criticality refers to systems of interactive components that grow to a degree of complexity that leads to cascades into chaos or collapse. The standard analogy is to a pile of sand onto which one grain is added at a time: As the pile approaches critical size, one more grain can cause a small shuffle of sand while, far less frequently, another grain can cause a major avalanche. The sizes of the avalanches follow the $1/f$ power law. For species, one more new population packed into an interacting community might lead to extinction of a variable number of species if the community were self-critical; usually, extinction numbers would be small, but occasionally there would be a major avalanche of extinction.

There are three criticisms of this argument. First, some mass extinctions are clearly associated with physical disturbances, such as the meteorite impact at the end of the Cretaceous. Second, many other processes can produce $1/f$ power

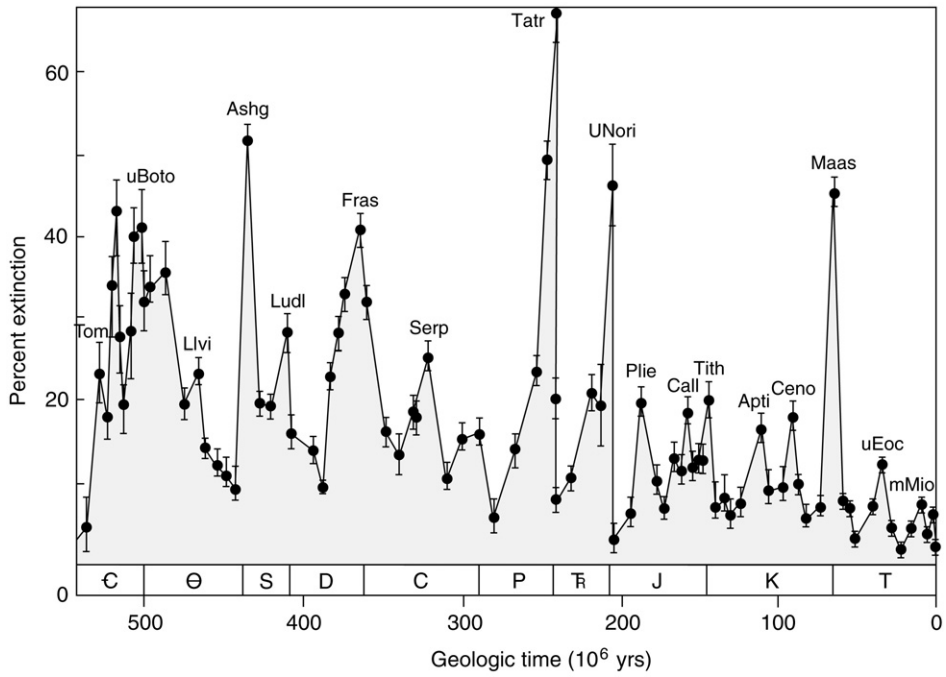


Figure 6 Extinction intensities for marine animal genera through the Phanerozoic. Intensity is measured as percent extinction ($=100\times$ number of extinctions/observed diversity) in geologic intervals averaging about 5.5 myr in duration. Estimated errors of estimate are indicated by error bars (± 1 standard deviation). Significant events are labeled: Tom = Tommotian; uBoto = upper Botomian; Llvi = Llanvirnian; Ashg = Ashgillian (end-Ordovician mass extinction); Ludl = Ludlovian; Fras = Frasnian (Late Devonian mass extinction); Serp = Serpukhovian; Tatr = Tatarian (end-Permian mass extinction); uNori = upper Norian (= end-Triassic mass extinction); Plie = Pliensbachian; Call = Callovian; Tith = Tithonian; Apti = Aptian; Ceno = Cenomanian; Maas = Maastrichtian (end-Cretaceous mass extinction); uEoc = upper Eocene; mMio = middle Miocene. Reproduced with permission from Sepkoski Jr. JJ (1996) Patterns of Phanerozoic extinctions: A perspective from global databases. In: Walliser OH (ed.) *Global Events and Event Stratigraphy*, pp. 35–52. Berlin: Springer.

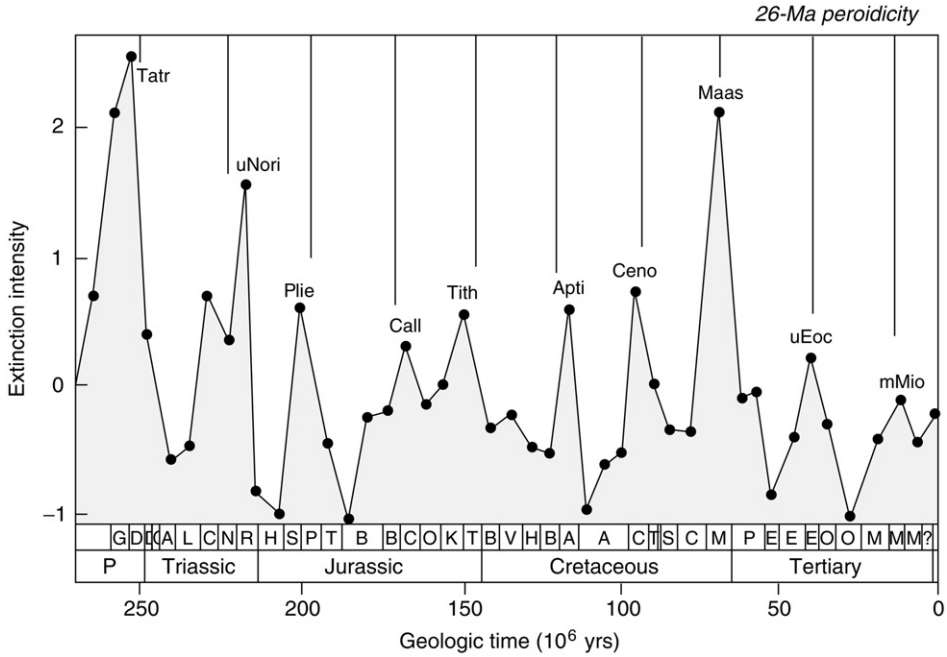


Figure 7 Time series for extinction intensity of marine animal genera through the Mesozoic and Cenozoic showing the putative 26 myr periodicity of mass extinction (vertical bars). The abbreviations are as in **Figure 6**. Extinction intensity here is measured as an average of percent extinction in major taxa that have been standardized to zero means and unit standard deviations. Reproduced with permission from Sepkoski Jr. JJ (1990) The taxonomic structure of periodic extinction. *Geological Society of America Special Paper* 247: 33–44.

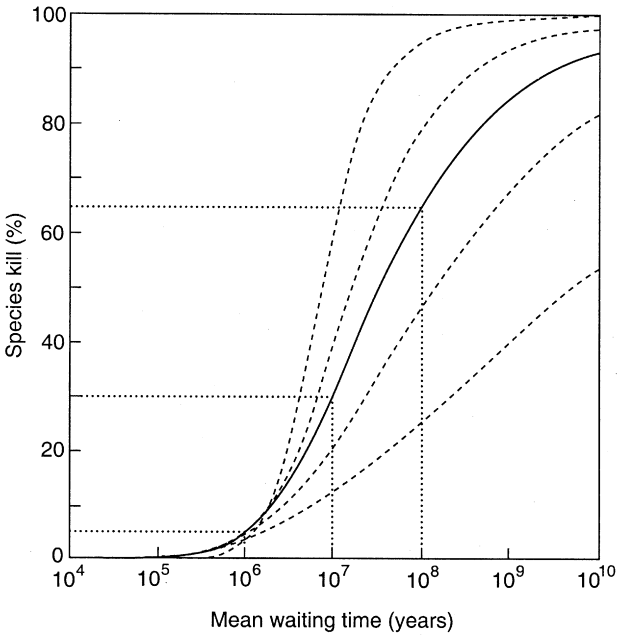


Figure 8 Raup's "kill curve" showing the expected extinction of marine species, as a percentage of standing diversity, over intervals of time ("mean waiting time"). During intervals of less than 10^6 years, percent of diversity expected to become extinct is low, but over longer intervals, especially greater than 10^8 years, extinction events of large magnitude become quite probable. The solid line is the best fit curve to fossil data; the dotted curves fall at bounds of uncertainty for empirical data. Reproduced from Raup DM (1991a) A kill curve for Phanerozoic marine species. *Paleobiology* 17: 37–48, with permission from Geoscience World.

laws when analyzed in the way organized self-criticality is; for example, the secular decline in background extinction over time, perhaps induced by changing taxa with different characteristic rates of extinction (Sepkoski, 1984), can cause power series to decline relative to increasing frequency. Finally, there is no independent evidence that species within marine communities, for which data have normally been analyzed, are as interactive as posited.

There still may be some interesting avenues of inquiry to pursue. The state of the biota is not constant, given variations in the earth's climates, changing positions of continents, and varying tectonic events and their effects on environments and the time over which the biota evolves to these varying conditions. A perturbation at one stage of the earth's conditions and the biota's development could have very different effects than an identical perturbation at a different time (Sepkoski, 1989). Thus, an expectation of some chaotic element in the record of extinction cannot be ruled out.

Selectivity of Mass Extinction: Victims and Survivors

An obvious effect of the end-Cretaceous mass extinction is that all (nonavian) dinosaurs disappeared whereas some mammals survived. At other mass extinctions there are also cases of major taxonomic groups disappearing and others surviving. This observation has led to a search for rules as to what makes some kinds of organisms more vulnerable to mass extinction than others. Patterns that have been found are largely probabilistic, and all seem to have unexplained exceptions:

1. Taxa that have high rates of extinction during times of background extinction suffer disproportionately at mass extinctions. Rates of extinction measured for higher taxa (e.g., classes) during background times tend to be good predictors of magnitudes of extinction at major events. Thus, trilobites, with higher rates of background extinction than articulate brachiopods, suffered more loss at the end-Ordovician, and brachiopods, with higher characteristic rates of extinction than bivalves, suffered greater proportional declines at the end-Permian and even the end-Cretaceous, when brachiopods were relatively marginalized (Sepkoski, 1984). It remains a major question as to why major taxonomic groups have different characteristic rates of extinction during background times and why these rates are conserved over vast stretches of macroevolutionary history and species turnover.
2. Animals, particularly terrestrial vertebrates, with large body size seem to be particularly prone to extinction. Dinosaurs perished at the end of the Cretaceous whereas many of the small mammals survived; many large mammals and birds that had evolved by the Eocene disappeared during the Eocene-Oligocene transition; and the late to end-Pleistocene extinctions that afflicted all continents other than Africa (and Antarctica) are often referred to as the "Pleistocene large mammal extinctions." Explanations for this selectivity involve (a) lower absolute population numbers of large animals, exposing them to greater chance of all individuals perishing during perturbations and (b) longer gestation times of large mammals, lengthening the time to restoration of pre-perturbation population numbers. Size-selectivity may not apply to marine invertebrates. Jablonski (1996) examined marine mollusks (gastropods and bivalves) for size-selectivity at the end-Cretaceous and found no differential extinction among large-sized species. On the other hand, Fischer and Arthur (1977) argued that marine vertebrate "megacarnivores" were very vulnerable to smaller extinction events.
3. It is often claimed that tropical plants and animals in the fossil past have been more susceptible to extinction than organisms living at higher latitudes. This is based in part on the observation that tropical reef communities disappear at several of the major mass extinctions, including the Late Devonian, end-Permian, and end-Cretaceous events, and millions to tens of millions of years intervene before diverse reefs re-evolve. The pattern for level bottom marine communities is not so clear: Jablonski and Raup (1995) analyzed the biogeography of bivalve extinction at the end of the Cretaceous and found no difference in extinction between tropical and temperate genera that were not associated with reefs.
4. Taxa with restricted geographic ranges appear to have been more prone to extinction than widespread taxa. Jablonski (1995) documented this for Late Cretaceous molluskan species during normal times of background extinction. At the end of the Cretaceous, however, the geographic range of individual species did not seem to be a factor, with

widespread species suffering as much extinction as restricted species. At the rank of the genus, however, genera with geographically dispersed species tended to survive preferentially relative to genera with species restricted to one or a few biogeographic provinces. Similar results have been obtained for Cambrian trilobites at various extinction events.

5. Environment seems to have played only a minor role in selectivity of extinction among marine animals at mass extinctions. Rates of extinction among Paleozoic marine animals tended to increase toward shallower waters during background times but to be equal across continental shelves at mass extinctions. There is some suggestion, however, that deep-water animals fared better: Marine shelves seem to have been repopulated by descendants of deep-water trilobites after Cambrian extinction events; shelf communities contained disproportionate numbers of deep-water sponges and corals after the Late Devonian mass extinction (McGhee, 1996); and descendants of mollusks typical of low-oxygen deep-water communities of the late Paleozoic are common after the end-Permian mass extinction (Erwin, 1993).

These reports of taxonomic selectivity at mass extinctions are largely confined to analyses around a few events of extinction, and much larger comparative studies need to be conducted across all mass extinctions for both marine and terrestrial organisms. However, if there are consistent probabilistic biases in terms of the properties of species and higher taxa that survive mass extinctions, then extinction events on frequencies of 2.6×10^7 years, eliminating 30 to 90% of the biota, could be a major factor in the history of life.

Recoveries from Mass Extinctions

Global data on mass extinctions exhibit not only the geologically rapid declines of the biota at mass extinctions but also subsequent recoveries of biodiversity in the aftermaths. With modern data, these recoveries are seen to encompass longer intervals of time than the extinction events, ranging up to about 5 myr for third-order extinction events to around 10 to 15 myr for second-order mass extinctions; recovery from the great end-Permian mass extinction took even longer but was interrupted by the end-Triassic event, some 30 myr later.

On long geologic timescales, recoveries from mass extinctions follow simple patterns expected for a diversifying biota. At more detailed scales, there is great complexity, with some patterns repeated after every mass extinction.

Large-Scale Rebound of Diversity

To a first approximation, diversification in the oceans can be described as a system with equilibrium constraints imposed by limitations of resources and their utilization (Sepkoski, 1984, 1992). This is most obvious in the long plateau of diversity during the Paleozoic era, spanning approximately one-quarter billion years. This interval witnessed two second-order mass extinctions, the end-Ordovician and Late Devonian. Significantly, after each, animal diversity rebounded to previous, unperturbed levels and then continued the Paleozoic plateau (**Figure 1b**). Also, the rate of rediversification after these mass extinctions was approximately the same as during the Ordovician radiations that established the Paleozoic plateau of global marine diversity.

Such a system can be modeled if it is assumed that Paleozoic animals were constrained by an equilibrium diversity, conditioned on the environment and the way these animals subdivided resources (**Figure 9**). The simplest mathematical description of diversification is a logistic model, assuming decreasing rates of origination as diversity increases. A mass extinction is a perturbation to the environment that increases extinction rates (**Figure 9b**). Following the perturbation, when extinction rates return to background levels, taxa rediversify to the normative level (**Figure 9c**).

Patterns are more complex through the quarter-billion years after the Paleozoic era, as diversity in the oceans generally increased. The increase reflects expansion of the modern marine fauna, which may have subdivided resources in ways different from the fauna of the Paleozoic era. Nonetheless, rebounds after mass extinctions during the Mesozoic and Cenozoic eras were much more rapid than the long-term increase in marine diversity when measured over millions of years and were congruent with initial rates of diversification of the taxonomic groups involved (cf. Miller and Sepkoski, 1989).

Delayed Diversification

When rebounds from mass extinction are analyzed at finer timescales, many complications become apparent (Erwin, 1998). One is delayed recovery: rediversification does not commence immediately after perturbations. For example, there are only slow rates of rediversification for 10^5 years after the end-Cretaceous event among planktonic foraminifers (actually, fast at first and then slow for some four myr; D'Hondt *et al.*, 1996), benthic mollusks (Jablonski, 1998), and terrestrial mammals (Maas and Krause, 1994). Recovery seems to be delayed for nearly 5 myr after the much larger end-Permian event, with only depauperate faunas in the oceans and on land (Erwin, 1998).

Explanations for this pattern have varied. Some workers have suggested that the extended post-extinction intervals of low diversity reflect lingering effects of the external perturbation—that is, continued environmental stress or instability. Others have hypothesized that substantial amounts of time are required for the evolution of new species that can reestablish normal ecosystem function which, in turn, can support high diversity. An example is the evolution of planktonic microherbivores after the end-Cretaceous mass extinction (D'Hondt *et al.*, 1998). The delayed interval of recovery after the severe end-Permian event could even be conceived as a long lag in a logistic system after diversity had been reduced by an order of magnitude.

"Disaster Species" and "Lazarus Taxa"

A feature that appears consistent during early phases of recovery from mass extinction is the appearance of "disaster

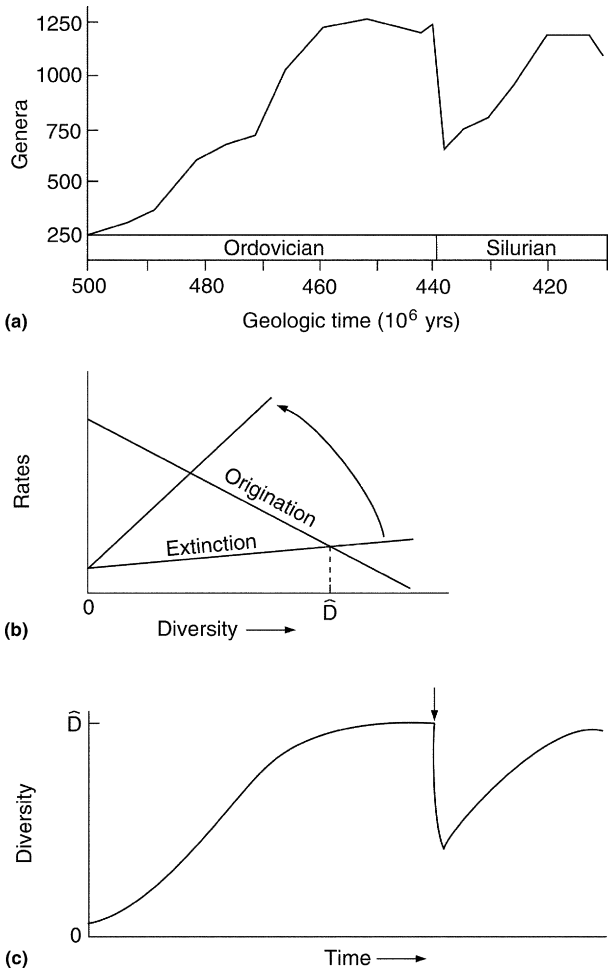


Figure 9 Observed and modeled loss and recovery of diversity around a mass extinction. (a) Observed numbers of genera through the Ordovician and Silurian periods, illustrating the major decline in diversity at the end of the Ordovician and subsequent recovery during the Silurian. Note that the rate of recovery is not very different from the rate of radiation during the Ordovician (cf. **Figure 1**) and that diversity during the Silurian recovers to the same level as during the late Ordovician. (b) Structure of a logistic model to describe diversification, mass extinction, and recovery. It is assumed that diversity increases multiplicatively and that per capita rates of origination decrease with increasing diversity, leading to sigmoidal growth of diversification. A mass extinction is a temporary steepening of the extinction function (or, alternatively, increasing of its elevation). $\hat{D}$ signifies the diversity at which origination rate equals extinction rate during normal times and diversity no longer grows. (c) A solution to the logistic model using rates measured from Ordovician marine animals, with a perturbation (arrow) applied at a model time corresponding to the observed end-Ordovician mass extinction. Diversity declines precipitously when extinction rate is increased, but then rebounds when extinction is returned to normal. Diversity in the model solution rebounds at the same rate as the Ordovician radiations and equilibrates at the same level as previously. Reproduced with permission from Sepkoski Jr. JJ (1992) Phylogenetic and ecologic patterns in the Phanerozoic history of marine biodiversity. In: Eldredge N (ed.) *Systematics, Ecology, and the Biodiversity Crisis*, pp. 77–100. New York: Columbia University Press.

species." These are remarkably abundant and geographically widespread species that appear in the waning stages or aftermaths of mass extinctions. Examples include the small foraminifer, *Guembelitria*, that spread out from marginal environments to form nearly monospecific assemblages immediately after the end-Cretaceous extinction and the terrestrial "fern spike" observed at the same time on land. Another example is the terrestrial synapsid ("mammal-like reptile"), *Lystrosaurus*, that produced monospecific assemblages in Gondwanaland following the end-Permian mass extinction.

Disaster species are characterized by not only great abundance in fossil deposits but also short geologic durations. Marine disaster species seem to flourish for around 10^5 years and then disappear, often to be replaced by another "disaster species" with similarly short duration. This appears like an evolutionary "boom and bust" cycle. Thus, the earliest Tertiary foraminifer, *Guembelitria*, declined and was replaced in dominance by a succession of bursts of *Paravularugoglobigerina*, "*Eoglobigerina*," and finally *Woodringina* over timescales of 10^5 years as the planktonic ecosystem seemed to regain stability.

Perhaps the antithesis of disaster species are "Lazarus taxa." These are lineages that disappear around mass extinctions—seemingly to have died—only to reappear in the fossil record some 10^6 to 10^7 years later (Erwin, 1998; Jablonski, 1986). These can be conspicuous but usually never dominant members of the biota prior to a mass extinction. Presumably, these taxa survived in some environmental or geographical refugium until the ecosystem regained sufficient function so that their dispersal—and recovery of sufficient abundance to be encountered by paleontologists—was possible.

The Modern Biodiversity Crisis

The collapses of ecosystems in the fossil past can inform thinking and actions with respect to the contemporary loss of biodiversity. The fossil record is incompletely known, but the data seem little worse than estimates of living biodiversity and current rates of species extinction. Estimates of modern global biodiversity range over more than an order of magnitude, from 5 to 100×10^6 species. Estimates of present-day loss of species range over nearly two orders of magnitude, from 5 to 150 species extinction *per day*. Most of these extinctions are terrestrial animals (largely insects), whereas the best data from the fossil record are for marine species. Still, in the geologic past, mass extinctions appear to have occurred largely contemporaneously both on land and in the seas (except for the end-Pleistocene event), so some simple calculations can be made.

If a median contemporary extinction rate of 41 species per day is assumed and species attrition is treated as a negative exponential, then it would take only about 16,000 years for 96% of the modern biota to become extinct (Sepkoski, 1997). This is a long time by human standards, but it is beyond the limits of geologic resolution at 250 Ma (the end-Permian), the only other time when 96% of Earth's biota disappeared.

This discussion has been largely in terms of fossil diversity. Another variable that can be measured in the geologic record is ratios of stable carbon isotopes. Because organisms tend to utilize ^{12}C in slight bias over the heavier ^{13}C, the ratio of the two isotopes preserved in the rock record can indicate, among other things, the activity of primary producers. For example, after the end of the Cretaceous, there was no difference in isotopic ratios of carbon in the skeletons of surface planktonic

foraminifers and deep-sea benthic foraminifers. This indicates that organic carbon stopped sinking to the deep ocean. This change must have resulted from either (a) a major decline in oceanic productivity or (b) a major collapse of community structure in marine producers or consumers (D'Hondt *et al.*, 1998).

Summary

Exploration of the fossil record has demonstrated that the earth's biota is fragile at timescales of 10^7 years, suffering numerous declines in diversity. The magnitudes of these declines have been variable, but at least five times in the last 500 myr animal diversity was reduced by more than 50%, with the most severe event, at the end of the Permian, eliminating more than 90% of animal diversity. Recoveries from these events have been slower than the mass extinctions, often taking 5 to 10 myr or more. Although the long-term rebounds of biodiversity are predictable, the detailed patterns of recovery are complex, involving outbreaks of disaster species and considerable ecological instability over timescales of 10^5 to 10^6 years. Current rates of species extinction could eliminate as many species as seen in past mass extinctions in a geologically short interval, and biotic recovery could be long and unpredictable.

See also: Dinosaurs, Extinction Theories for. Extinction, Causes of. Fossil Record. Marine Mammals, Extinctions of. Mass Extinctions, Notable Examples of. Modern Examples of Extinctions

References

Alvarez LW, Alvarez W, Asaro F, and Michel HV (1980) Extraterrestrial cause for the Cretaceous-Tertiary extinction. *Science* 208: 1095–1108.

D'Hondt S, Donaghy P, Zachos JC, Luttenberg D, and Lindinger M (1998) Organic carbon fluxes and ecological recovery from the Cretaceous-Tertiary mass extinction. *Science* 282: 276–279.

D'Hondt S, Herbert TD, King J, and Gibson C (1996) Planktic foraminifera, asteroids, and marine production: Death and recovery at the Cretaceous-Tertiary boundary. *Geological Society of America Special Paper* 307: 303–317.

Erwin DH (1993) *The Great Paleozoic Crisis*. New York: Columbia University Press.

Erwin DH (1998) The end and the beginning: recoveries from mass extinctions. *Trends in Ecology and Evolution* 13: 344–349.

Fischer AG and Arthur MA (1977) Secular variation in the pelagic realm. In: Cook HE and Enos P (eds.) *Deep-Water Carbonate Environments. Society of Economic Paleontologists and Mineralogists Special Publication* 25, pp. 19–50.

Jablonski D (1986) Causes and consequences of mass extinctions: A comparative approach. In: Elliot DK (ed.) *Dynamics of Evolution*, pp. 183–229. New York: Wiley Interscience.

Jablonski D (1995) Extinctions in the fossil record. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 25–44. Oxford: Oxford University Press.

Jablonski D (1996) Body size and macroevolution. In: Jablonski D, Erwin DH, and Lipps J (eds.) *Evolutionary Paleobiology*, pp. 256–289. Chicago: University of Chicago Press.

Jablonski D (1998) Geographic variation in the molluscan recovery from the end-Cretaceous extinction. *Science* 279: 1327–1330.

Jablonski D and Raup DM (1995) Selectivity of end-Cretaceous marine bivalve extinctions. *Science* 268: 389–391.

Maas MC and Krause DK (1994) Mammalian turnover and community structure in the Paleocene of North America. *Historical Biology* 8: 91–128.

McGhee Jr. GR (1996) *The Late Devonian Mass Extinction*. New York: Columbia University Press.

Miller AI (1997) Coordinated stasis or coincident relative stability? *Paleobiology* 23: 155–164.

Miller AI and Sepkoski Jr. JJ (1989) Modeling bivalve diversification: The effect of interaction on a macroevolutionary system. *Paleobiology* 14: 364–369.

Newell ND (1967) Revolutions in the history of life. *Geological Society of America Special Paper* 89: 63–91.

Phillips J (1860) *Life on Earth: Its Origin and Succession*. Cambridge: Macmillan.

Raup DM (1989) The case for extraterrestrial causes of extinction. *Philosophical Transactions of the Royal Society of London B* 325: 421–435.

Raup DM (1991a) A kill curve for Phanerozoic marine species. *Paleobiology* 17: 37–48.

Raup DM (1991b) Periodicity of extinction: A review. In: Muller DW, McKenzie JA, and Weisert W (eds.) *Controversies in Modern Geology*, pp. 193–208. London: Academic Press.

Raup DM (1992) Large-body impact and extinction in the Phanerozoic. *Paleobiology* 18: 80–88.

Raup DM and Sepkoski Jr. JJ (1984) Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences, USA* 81: 801–805.

Schindewolf OH (1962) Neokatastrophismus? *Geologische Geselschaft, Zeitschrift Jahrg* 114: 430–445.

Sepkoski Jr. JJ (1984) A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* 10: 246–267.

Sepkoski Jr. JJ (1989) Periodicity in extinction and the problem of catastrophism in the history of life. *Journal of the Geological Society of London* 146: 7–19.

Sepkoski Jr. JJ (1990) The taxonomic structure of periodic extinction. *Geological Society of America Special Paper* 247: 33–44.

Sepkoski Jr. JJ (1992) Phylogenetic and ecologic patterns in the Phanerozoic history of marine biodiversity. In: Eldredge N (ed.) *Systematics, Ecology, and the Biodiversity Crisis*, pp. 77–100. New York: Columbia University Press.

Sepkoski Jr. JJ (1996) Patterns of Phanerozoic extinctions: A perspective from global data bases. In: Walliser OH (ed.) *Global Events and Event Stratigraphy*, pp. 35–52. Berlin: Springer.

Sepkoski Jr. JJ (1997) Biodiversity: past, present, and future. *Journal of Paleontology* 71: 533–539.

Sepkoski Jr. JJ and Koch CF (1996) Evaluating paleontologic data relating to bio-events. In: Walliser OH (ed.) *Global Events and Event Stratigraphy*, pp. 21–34. Berlin: Springer.

Sepkoski Jr. JJ and Schopf JW (1992) The Proterozoic fossil record: Special problems in analyzing diversity patterns. In: Schopf JW and Klein C (eds.) *The Proterozoic Biosphere: A Multi-Disciplinary Study*, pp. 525–527. Cambridge: Cambridge University Press.

Signor PW and Lipps JH (1982) Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. *Geological Society of America Special Paper* 190: 291–296.

Solé RV, Manrubia SC, Benton M, and Bak P (1997) Self-similarity of extinction statistics in the fossil record. *Nature* 388: 764–767.

Mass Extinctions, Notable Examples of

Richard J Twitchett, Plymouth University, Plymouth, UK

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Douglas H. Erwin, volume 4, pp 111–122, © 2001, Elsevier Inc.

Glossary

Mass extinction A rapid loss of a large fraction of biodiversity on timescales of 10^0–10^6 years, generally involving a variety of unrelated groups.

Supercontinent The amalgamation of many continental masses into a single mass through continental drift. The supercontinent of Pangaea (from 300 to approximately 190 million years ago (Ma)) included most continental areas other than those that comprise east Asia.

Introduction

The 3.6-billion-year history of life on Earth has been punctuated by numerous extinction events, but much of that history is poorly known due to a patchy and incomplete rock and fossil record. Only the record of the past 600 million years is sufficiently well preserved to provide an adequate archive of these biotic crises. Five great mass extinctions punctuate this recent slice of Earth history, which spans most of the evolution of animals and multicellular plants. Additionally, there are many smaller biotic crises, but not all of these are well studied.

Great Mass Extinctions

The five most severe intervals of biotic crisis occurred during the Late Ordovician Period, the Late Devonian, the Late Permian, the Late Triassic, and at the end of the Cretaceous (**Figure 1**). A variety of different causes have been invoked, from climate change (both warming and cooling) to the impact of extraterrestrial objects, but despite numerous attempts to find a common mechanism for these events no single factor appears to have caused all events. One of the challenges of studying these ancient events is that the rock and fossil records are both imperfect archives, with gaps and missing information. This is one reason that estimates of past biodiversity are often compiled at the taxonomic resolution of families (**Figure 1**) or genera (**Figure 2**), rather than species. As new fossiliferous locations are discovered or studied in more detail, and as new methods of coping with biases in the rock and fossil records are developed, our knowledge of these events improves and estimates of extinction rates are revised. Technological advances, especially in the fields of geochemistry, geochronology, and microscopy are also providing important new information concerning the timing of these events and their associated environmental changes.

Late Ordovician

The first of the five great mass extinctions occurred during the Late Ordovician, approximately 445 million years ago (Ma). At that time there were few, if any, organisms on land, and this crisis was confined to marine groups. Detailed studies have revealed the presence of two linked, but clearly separate, extinction events: one at the very end of the Ordovician and another 1–2 million years earlier. Overall, approximately 25% of all families disappeared, with extinction of approximately 49% of all genera in the first event and 44% in the second event. Despite the apparent magnitude of the extinction, it had few lasting ecological effects: with some minor exceptions, Silurian faunas look much like Ordovician ones, with similar groups of organisms occupying similar niches.

Graptolites, a group of colonial, floating (planktonic) marine organisms, were commonly found on the outer shelves and were almost completely eliminated by this extinction, with only a few species surviving. Other planktonic groups and swimming (nektonic) groups, such as conodonts and nautiloids, all appear to have suffered considerable extinction. On the seafloor, the articulate brachiopods were one of the major components of Ordovician seafloor (benthic) ecosystems, and approximately 83% of brachiopod genera became extinct, primarily during the first of the two events. A brachiopod genus, *Hirnantia*, and its associated fauna, appeared in many areas in the aftermath of the first extinction event. At any particular locality, fewer than 10 species are typically present in total and all appear to be adapted to cold water. This 'Hirnantian fauna' reached equatorial latitudes during this interval, but was wiped out by the second extinction event. Trilobites display a similar pattern of extinction followed by the spread of a low diversity, broadly distributed group between the extinctions; however, the second extinction appears to be more significant than the first. Trilobites living in deeper water suffered greater extinctions than those occupying shallow-water habitats. Bivalves and echinoderms also experienced considerable extinction, whereas bryozoans, a colonial group of filter feeders, were not affected as severely. Approximately 70% of rugose corals disappeared, but they recovered quickly so the extinction did not have a major impact. In fact, this is the only major mass extinction that did not trigger an extensive turnover in reef communities.

The causes of the Late Ordovician mass extinction appear to be linked to global climate change and its associated environmental impacts. Warm temperatures characterized the

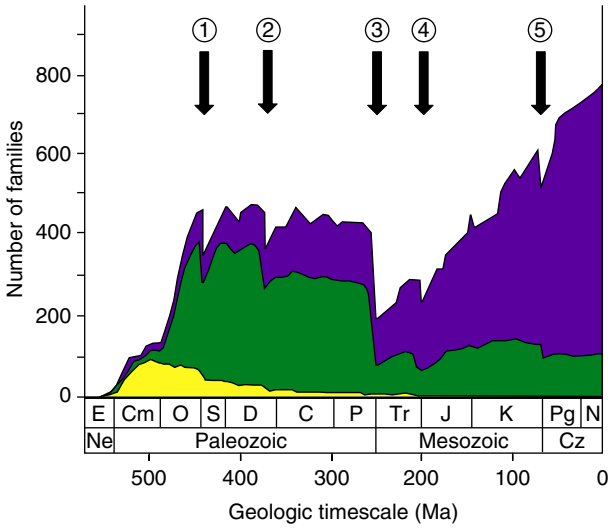


Figure 1 Past diversity of marine families. The uppermost line shows the total number of marine families recorded in the fossil record through the geologic periods of the Phanerozoic and uppermost Neoproterozoic. The circles correspond to the five major mass extinctions: (1) Late Ordovician; (2) Late Devonian; (3) Late Permian; (4) Late Triassic; and (5) end-Cretaceous. The lower, yellow shading shows the families assignable to the Cambrian evolutionary fauna; the green shading shows the families assignable to the Paleozoic evolutionary fauna, and the upper, purple shading shows the families assignable to the modern evolutionary fauna. E, Ediacaran; Ne, Neoproterozoic; Cm, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; Pg. Paleogene; N, Neogene; Cz, Cenozoic; Ma, million years ago. Modified from Sepkoski Jr. JJ (1984) A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinction. *Paleobiology* 10: 246–267, with permission from Palcantological Society.

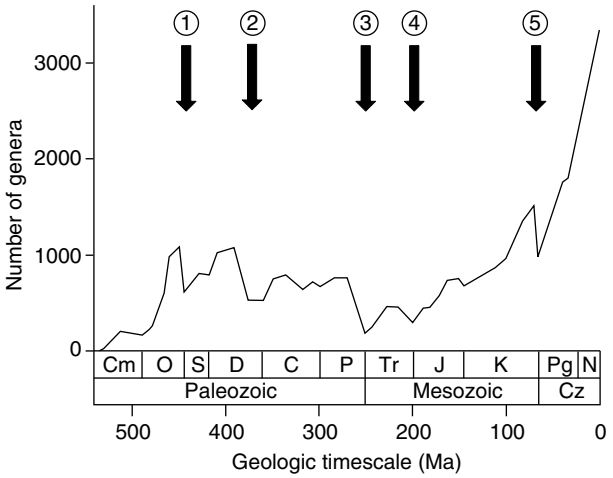


Figure 2 The past diversity of marine genera. The line shows the total number of marine genera recorded in the fossil record through the geologic periods of the Phanerozoic. The circles correspond to the five major mass extinctions: (1) Late Ordovician; (2) Late Devonian; (3) Late Permian; (4) Late Triassic; and (5) end-Cretaceous. Cm, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; Pg. Paleogene; N, Neogene; Cz, Cenozoic; Ma, million years ago. Modified from Alroy J, Aberhan M, Bottjer DJ, *et al.* (2008) Phanerozoic trends in the global diversity of marine invertebrates. *Science* 321: 97–100, with permission from AAAS.

Early Ordovician as the Earth was experiencing a greenhouse climate, with carbon dioxide levels and temperatures much higher than those of the present day. Through the Early Ordovician there is evidence that climate gradually cooled, and from the Middle Ordovician tropical sea surface temperatures were much more similar to those of the present day (Trotter *et al.*, 2008). Global climate remained fairly equable through the Middle and into the Late Ordovician, but sometime in the Katian (the penultimate stage of the Ordovician, about 450 Ma) temperatures begin to fluctuate. It appears that a tipping point was reached toward the end of the Katian and global temperatures plummeted by some 8 °C through the early part of the Hirnantian Stage, plunging the world into the grips of a global icehouse (Finnegan *et al.*, 2011). As climate cooled ice sheets would have expanded, causing sea level to drop, eventually exposing the continental shelves. The first extinction event, which eliminated nearly half of all genera, is intimately linked to cooling and sea level fall at this time, as organisms succumbed to pressures such as lowered temperatures, changes in ocean circulation and reduced space on shallow marine shelves.

The reason for the climatic instability and onset of cooling are still unresolved. Plate tectonics might have had an influence as a large continental landmass moved close to the South Pole during the Ordovician. Elsewhere, the chemical weathering of mountain belts formed by plate collisions may have helped reduce the level of carbon dioxide in the atmosphere. Some evidence suggests that ice sheets may have been present, at least transiently, as early as the Middle Ordovician, implying a lengthy and complex interval of climate change preceding the onset of catastrophic cooling (Finnegan *et al.*, 2011). In the wake of the first extinction event, as global temperatures and sea level fell, the low-diversity, cool-water Hirnantian fauna became widespread. Around the middle of the Hirnantian stage, global cooling ceased and temperatures began to rise once more through the latter part of the Hirnantian. As the Earth emerged from global icehouse conditions and ice-sheets retreated once more, sea level and ocean temperatures would have increased. The warm, rising seas, associated with the spread of low oxygen, or anoxic, conditions, led to the second episode of extinction at the end of the Hirnantian.

The sea-level fall of the Late Ordovician glaciation prevented the deposition of shallow marine rocks in many areas, as the shelves became exposed. This complicates the scientific analysis of biodiversity changes during this interval. Do certain species disappear because of real biological extinction, or simply because the types of rocks that preserve their fossils are no longer present?

This is a problem that affects the analysis of all past extinction events to a greater or lesser extent, especially further back in time when the global rock record is generally less complete (Smith, 2007). Although rocks recording the Late Ordovician crisis are common in Europe and China, for example, deposits of that age in North America are limited to a few places in central Nevada, USA, and Anticosti Island,

Canada. The Nevada sections record a diversity of environments within a small region, however, which is rarely the case elsewhere and which allows the extinctions to be tracked between different environments (Finney *et al.*, 1999).

The potential problems of biases in the rock record and inadequate or patchy sampling have led some scientists to question whether the trends in Phanerozoic diversity identified by Sepkoski (**Figure 1**) reflect a true picture of past biodiversity, or merely reflect the vagaries of the fossil record. A National Science Foundation initiative to document fossil diversity more completely and thoroughly than previous attempts have been able to manage – known as the Paleobiology Database project – has resulted in a revised view of the history of life on Earth that attempts to take into account biases in sampling. While the 'Big 5' extinctions still coincide with reductions in global diversity (**Figure 3**), some of the details of the timing, magnitude, and duration of the various peaks and troughs are different. This alternative view is not accepted by all scientists, and the Paleobiology Database is far from complete, but the approach demonstrates how biases that might be affecting our understanding of the history of life on Earth could potentially be filtered out.

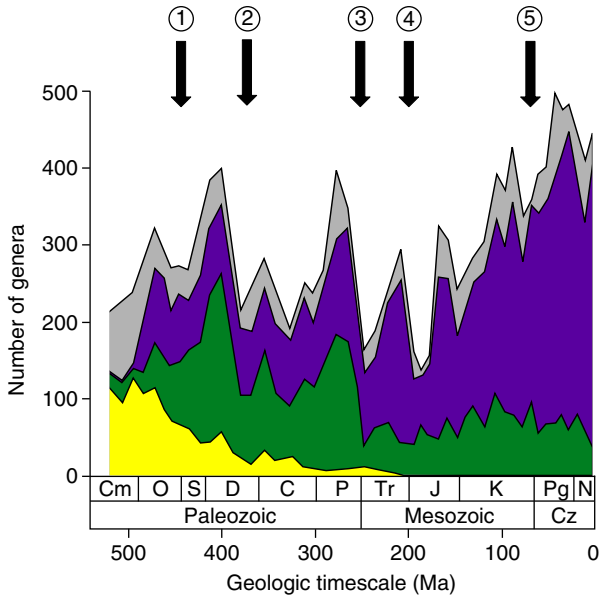


Figure 3 The past diversity of marine genera, corrected for sampling bias. The uppermost line shows the number of marine genera recorded in the fossil record through the geologic periods of the Phanerozoic and uppermost Neoproterozoic. The circles correspond to the five major mass extinctions: (1) Late Ordovician; (2) Late Devonian; (3) Late Permian; (4) Late Triassic; and (5) end-Cretaceous. The lower, yellow shading shows the genera assignable to the Cambrian evolutionary fauna; the green shading shows the genera assignable to the Paleozoic evolutionary fauna, the purple shading shows the genera assignable to the modern evolutionary fauna. E, Ediacaran; Ne, Neoproterozoic; Cm, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; Pg, Paleogene; N, Neogene; Cz, Cenozoic; Ma, million years ago. Modified from Alroy J (2010) The shifting balance of diversity among major marine animal groups. *Science* 329: 1191–1194, with permission from AAAS.

Late Devonian

The Late Devonian Period records a number of closely-spaced, apparent biotic crises of varying magnitude, which together create a substantial decrease in diversity at this time (**Figure 1**). These biotic crises have been recognized since the 1870s, but the rapidity of these events remains unclear. For the Givetian and Frasnian Stages of the Late Devonian, Sepkoski (1996) records a 57% extinction of marine genera (47% for filtered data) and approximately 22% extinction of marine families. The most significant events of the Devonian appear to have occurred at the end of the Givetian Stage (the 'Taghanic event', some 385 Ma), at the end of the Frasnian (the 'Kellwasser event', 375 Ma), and finally at the end of the Famennian (the 'Hangenberg event', 360 Ma). Terrestrial plants had evolved and experienced their first major biotic crisis during this interval as well, although the available data are somewhat patchy (McGhee, 1996; Hallam and Wignall, 1997). Although these events show some similarities with each other, they evidently represent separate events separated by intervals of biotic recovery.

Middle Devonian shallow seas were dominated by extensive reef ecosystems, extending from 45° to 50° south of the equator to perhaps 60° north. Major Devonian reef complexes are known from Australia, western North America, Russia, and Europe, with areal extents often many times that of the present day Great Barrier Reef. These reefs were dominated by rugose and tabulate corals (the major reef builders of the Paleozoic), stromatoporoids (a type of calcified sponge), in addition to reef dwellers such as brachiopods, bryozoans, and large foraminifera. The extinctions suffered by these groups during the Late Devonian resulted in collapse and near elimination of reef ecosystems. The end-Givetian Taghanic event witnessed the most severe extinction for the corals and reef dwellers. At least six families of brachiopods, 47% of stromatoporoid genera, and 50% of coral genera disappeared at this time, and the bryozoans suffered their largest biotic crisis. The few surviving tabulates, colonial rugosans, and stromatoporoids regrouped during the early Frasnian, forming low diversity, but regionally extensive, reefs that reached peak sizes and areal extents by the Middle Frasnian. Reefs then declined through the latest Frasnian becoming very isolated and geographically restricted just before their final extinction immediately preceding the Frasnian/Famennian boundary (Copper, 2002). Alongside the decline in coral reefs, McGhee (1996) indicates that 75% of brachiopod genera disappeared during the Frasnian. Brachiopods recovered quickly during the Famennian, but corals and stromatoporoids did not. No coral reefs are known from the Famennian, and just a few isolated examples of stromatoporoid patch reef are all that remain of the former dominance of these metazoan reef builders. Most Famennian reefs were constructed by microbial organisms, and stromatoporoids were completely extinct by the end of the Devonian (Copper, 2002).

Trilobites, although much reduced in diversity from the Cambrian and Ordovician, were still important during the Devonian. Eight orders and 13 families are known from the Givetian, but only a single family and eight genera survived to the Famennian. Although there is elevated extinction from the Givetian through the Frasnian, the only discrete crisis

occurred during the end-Frasnian Kellwasser event when numerous important groups disappeared. Deep-water and some shallow-water trilobites diversified immediately after the extinction during the Famennian, but trilobites suffered considerable extinction again at the end-Famennian and never regained their previous diversity. Two other arthropod groups, the eurypterids (sea scorpions) and the ostracods, were also members of Devonian marine communities. Eurypterids were relatively rare, large predators and their diversity appears to be declining throughout the interval, although several families became extinct during the last two stages of the Devonian. Ostracods were more common, comprising both benthic and pelagic forms, and suffered considerable extinction during the Kellwasser event. Some 75% of benthic ostracods disappeared then, with genera tolerant of low-oxygen conditions showing greater survival.

Ammonoids, the coiled cephalopods that were distant cousins of the modern chambered *Nautilus*, almost disappeared, with an estimated 88% species extinction. Ammonoids display a boom-and-bust pattern throughout their history, and this includes massive extinctions during the Taghanic event, in which only two genera survived, and the Kellwasser event, with only a few surviving genera, mostly from deeper water habitats. Famennian ammonoids diversified considerably, but only a single family, and perhaps a single genus, survived the Hangenberg event at the end of the Devonian. Nautiloids, the other major group of fossilizable cephalopods, suffered minor perturbation during the Kellwasser event, but suffered perhaps their greatest crisis during the Hangenberg event when 11 of 19 families disappeared. In contrast to the cephalopods, other molluscs such as bivalves and gastropods suffered only minor losses in diversity.

The Devonian is sometimes known as "the Age of Fishes" due to the enormous diversity of fish during that period. Most fish groups did poorly during the Late Devonian, which marked the complete extinction of the jawless, armored forms as all nine of their families disappeared. More advanced fish and other vertebrate groups suffered as well during both the Kellwasser and Hangenberg crises, with generic losses of 19% and 32% at each event, respectively. The latter crisis led to a major restructuring of vertebrate communities with the diversification of previously less common groups such as sharks and tetrapods that dominate marine and nonmarine habitats at the present day (Sallan and Coates, 2010).

What patterns of ecological selectivity are apparent from the variations in extinction intensity among different groups? Often, patterns of differential extinction are among the best clues to the causes of mass extinctions and other biotic crises. The patterns of change are complex, with long-term decline in many groups as well as during the discrete Taghanic, Kellwasser, and Hangenberg events. The first of the three biotic crises primarily affected shallow-water benthic groups. The Kellwasser is the best known of the three events (McGhee, 1996). Warm-water groups, particularly brachiopods and reef forms, suffered the greatest extinction, whereas groups adapted to high latitudes and colder waters seem to have fared better. Most swimming groups either became extinct or declined greatly in diversity. Whether the Hangenberg crisis qualifies as a mass extinction is unclear. Many swimming groups, including placoderm fish, nautiloids, and ammonoids, as well as trilobites and stromotoporoids experienced considerable extinction (Hallam and Wignall, 1997).

Extraterrestrial impact, sea-level fluctuations with attendant spread of anoxic waters, and climatic changes such as global cooling are among the more prominent suggested causes of the biotic crisis (McGhee, 1996). The evidence for impact would include demonstration that extinctions are rapid and virtually simultaneous; a coincident enrichment of iridium, which is rare on Earth but common in extraterrestrial objects; mechanical indications of impact including shocked quartz; fused glass from the impact site; and a crater of the correct age. Although evidence of impact is present in Late Devonian rocks, including the 52-km Siljan crater in Sweden and heightened levels of iridium, these are not closely associated with extinction horizons.

Evidence in support of global cooling includes the heightened survival of cold-water groups, the presence of glacial sediments in Brazil and the eastern USA, and the rapid sea-level fluctuations, reminiscent of Pleistocene glacially induced fluctuations (Copper, 1998). While enhanced survival of cold-water taxa can also be interpreted as survival of deep-water forms, McGhee (1996) concludes that global cooling best explains the pattern of extinction and survival, although he argues that the rapidity of the extinctions perhaps indicates that glaciation is not an adequate explanation.

A number of alternative scenarios have been proposed, most of which invoke sea level, shifts in the carbon cycle, and spread of oxygen-poor waters (Hallam and Wignall, 1997). The scenarios differ in the triggering mechanism, which ranges from rapid glaciations to oceanic volcanism and the effects of increased flux of nutrients to the oceans caused by the first development of land plants. Evidence for warming and anoxia has been questioned, with Copper (1998) noting that many of the atrypid brachiopods adapted to low-oxygen environments became extinct while other brachiopods from the same communities survived. Investigations of these events continue, but currently there does not appear to be sufficient evidence to provide definitive tests of the various competing models for the causes of this extinction. The most reasonable explanation, however, appears to involve a multitude of interacting causes rather than a single critical cause.

Late Permian

The most severe mass extinction event of the Phanerozoic occurred just prior to the very end of the Permian Period, in the late Changhsingian (252 Ma), and caused the most pervasive reorganization of marine communities since the beginning of the Paleozoic (Erwin, 1993). All subsequent marine communities, right down to the present day, reflect the heritage of this event. Extinctions on land were no less severe, and include widespread disruption of plant communities, extinction of most vertebrate taxa, and the only mass extinction suffered by insects. This unparalleled upheaval in the evolutionary history of animals and plants was preceded by a smaller magnitude episode of marine extinction at the end of the Guadalupian, some 8 million years earlier. In some former studies this end-Guadalupian crisis was treated as part of the

'Late Permian' mass extinction event, although detailed analyses from the 1990s clearly show them to be separate and distinct.

A number of major groups of organisms that had dominated the oceans for the previous 250 million years or more disappeared completely during the Late Permian. Some of those, such as the trilobites, had already dwindled to a few species and had long since passed their peak. Others, such as the rugose and tabulate corals, were common and widespread yet still vanished completely. Reef-building organisms, such as sponges, corals, and algae, are common in Changhsingian rocks in many areas, for example, Sichuan Province, China, but after the extinction event metazoan reefs are not recorded again until the Middle Triassic. Scleractinian corals – major reef-builders during the Triassic as they are at the present day – diversified from some unknown, perhaps soft-bodied, extinction survivor to fill the niches that were vacated by the extinction of the Permian reef builders. Brachiopods are perhaps the most diagnostic group of Paleozoic marine organisms, but lost some 90% of their genera during the Late Permian event. Two echinoderm classes that were diverse and abundant during the Paleozoic, the echinoids and crinoids, were saved from complete extinction by the survival of just one or two genera. The morphology of all modern sea urchins is largely inherited from the one surviving genus *Miocidaris*. Our understanding of the impact of the Late Permian extinction on the other echinoderm classes, namely the brittle stars, sea stars, and sea cucumbers, is impeded by a very poor quality fossil record for those groups during the crucial interval. Initial diversification of modern starfish appears to have taken place during the Triassic, perhaps indicating that a single genus survived the extinction, whereas it appears that no major groups of sea cucumbers were lost during the Late Permian, implying the survival of a number of genera. Molluscs dominate most post-Paleozoic marine shelf communities, from the earliest Triassic to the present day, and it is not surprising that they suffered lower extinction rates than many other organisms, although they still lost significant numbers of taxa. Gastropods, for example, experienced approximately 75% extinction at the generic level, and bivalves lost 60% of their genera. Within the cephalopods, which lost 75% of genera overall, the ammonoids, as usual, fared particularly badly. Marine communities in the immediate aftermath were dominated by a handful of bivalve genera, with a few gastropods and inarticulate brachiopods, and show remarkably little variation worldwide.

Extensive biotic crises occurred on land as well. Clearly, any proposed cause that affects the oceans to the exclusion of land is ruled out. Twenty-two orders of insects had appeared by the end of the Permian, and nine became extinct and another 10 suffered serious diversity declines. This is the only mass extinction ever suffered by insects. Many of the groups that disappeared could not fold their wings back over their body (as do flies and butterflies) but rather held them straight out to the side; dragonflies are a relic of these Permian groups. In contrast to studies in the 1980s, which suggested that plants experienced relatively little extinction, recent work has revealed massive perturbations, including the loss of high-latitude gymnosperm forests. Terrestrial ecosystems supported a great variety of vertebrates, including large herbivores and saber-toothed carnivores, but approximately 78% of tetrapod families became extinct at this time. Fossil spores and pollen preserved in shallow marine rocks show that ecological disruption on land occurred at the same time as extinction in the seas. Efforts to accurately date the extinctions have been ongoing since the late 1990s (Bowring *et al.*, 1998) and it seems likely that most extinctions occurred in a period of some 20,000 years or less (Twitchett *et al.*, 2001).

Almost every conceivable cause has, at one time or another and with varying degrees of support, been suggested for the Late Permian mass extinction event. Certainly, the cause was terrestrial in origin, as despite decades of investigation and numerous published studies since the early 1980s, none of the evidence put forward for extraterrestrial impact has ever held up to detailed scrutiny or analysis. Some hypotheses have become obsolete as more precise measurements of the timing and duration of the crisis have been made. Thus, the once-favored explanation that the extinctions were the direct result of the assembly of the supercontinent Pangaea can be rejected because the assembly was complete at least 20 million years before the Late Permian, and the crisis itself was far too rapid to be explained by the slow dance of the continents. Another view, held by several generations of workers, was that sea level fall was responsible. This hypothesis was based on the observation that in many places, especially in North America, marine strata of latest Permian and earliest Triassic age are absent, and the Permian/Triassic boundary is marked by an unconformity. High-resolution studies in the 1990s that focused on areas where the rock record is much more complete, such as northern Italy and Svalbard, demonstrated that although a significant sea-level fall did indeed occur in the Changhsingian stage of the latest Permian, it was associated with few apparent extinctions, and the mass extinction interval actually occurs during the subsequent sea level rise.

The hypothesis that global cooling and widespread glaciations were the ultimate cause still has low levels of support, despite the facts that no glacial deposits of the correct age have been identified anywhere on the planet and, as noted, the extinctions are demonstrably not associated with falling sea level. There is much more evidence to support the alternative view that the Late Permian extinction event was associated with an episode of catastrophic global warming.

The overall climatic trend through the Permian is one of increasing warmth. During the late Carboniferous and Early Permian, the Earth was in the grip of a major icehouse, with continental scale ice-sheets covering much of the southern hemisphere land surface. Gradually, as atmospheric carbon dioxide levels rose through the Permian, the icesheets disappeared so that by the Late Permian all evidence of glacial deposits has vanished, replaced by widespread evidence for hot desert conditions at lower latitudes and lush, dense forests at higher latitudes. Computer-generated climate simulations support the rock record evidence that the Late Permian was a warm interval in the planet's history. Then, in the final stage of the Permian, there is a further increase in atmospheric carbon dioxide levels, recorded in changes in the density of pores (stomata) on fossil leaves and in the chemistry of fossil soils. This sudden spike in greenhouse gases would have fueled additional global warming, which would have led to increased temperatures, sea level rise, and changes to atmospheric and

ocean circulation with catastrophic consequences for the terrestrial and marine biota. Evidence of rising temperature at this time was first gathered in the late 1980s from chemical analyses of shallow water, tropical limestones, which suggested an increase of 5–6 °C through the marine extinction event. More accurate measurements 20 years later on better preserved, tropical, shallow marine brachiopod shells gave estimates that were much higher: a temperature increase through the extinction of perhaps 8–10 °C (Kearsey *et al.*, 2009).

What caused this catastrophic Late Permian greenhouse? It is probably no coincidence that the Late Permian also witnessed the formation of the largest volcanic province to have erupted on land during the Phanerozoic. The eruptions covered an area of some 5 million square kilometers of present-day western Siberia, Russia, with layer on layer of basaltic lava to a depth of more than 3 km in some places. The total volume of extrusive and shallow intrusive igneous rocks is estimated at between 2 and 4 million cubic kilometers (Saunders and Reichow, 2009). Within error, the ages of most lava flows span the extinction event. These eruptions would have released large amounts of carbon dioxide, which gradually built up in the atmosphere, possibly further enhanced by the presence of underlying coal-rich rocks that would have been cooked as the molten magma passed through on its way to the surface.

Rising temperatures alone would probably have caused extinction of some groups of animals and plants. Warming would also cause other environmental changes that would have disrupted the food supply, nutrient supply or life cycle of organisms. Extreme seasonality, changes in rainfall patterns and expansion of climatic belts would have affected terrestrial plants and animals. In many areas it appears that lush gymnosperm forests gave way to sparse shrubby vegetation. Seasonal rainfall and reduced vegetation cover resulted in a change from meandering to braided river systems, as shown by the sudden appearance in the rock record of thick conglomerates and pebble beds in most localities. Erosion of the landscape was enhanced by elevated CO_2 levels, and the rivers must have discharged vast plumes of brown, muddy water into the coastal seas causing an increased flux of land-derived nutrients into shallow marine ecosystems. Marine rocks in the immediate aftermath of the extinction record this increased sedimentation rate in thick, muddy rock sequences (Algeo and Twitchett, 2010). Aquatic ecosystems would have suffered increased turbidity, harmful to many organisms, and marine shelves would have experienced frequent anoxic episodes as the enhanced nutrient flux would have fueled blooms of phytoplankton.

Beyond the shallow shelves, the rest of the ocean was also severely impacted. As global warming tends to cause greater temperature increase in polar regions, the thermal gradient between the tropics and the poles would become much reduced. Normally, this gradient drives the thermohaline circulation that provides a constant supply of cooler, oxygen- and nutrient-rich waters to the deep ocean. As warming proceeded, circulation would become less vigorous and many parts of the ocean would become starved of oxygen and nutrients. Global ocean productivity levels would fall as upwelling zones became less productive. Evidence of anoxic deep oceans has been documented from the rock record of Japan. Fossilized molecules derived from green sulfur bacteria, which need both light and hydrogen sulfide, are present in many marine rocks in the immediate aftermath of the extinction, showing that at times the oceans were anoxic from the seafloor up to the photic zone.

Some regions evidently fared better than others, and patterns of recovery vary between regions. At least some seamounts and offshore highs, such as those in Oman documented by Krystyn *et al.* (2003), were not affected by increased sedimentation or anoxic conditions and record a relative rapid ecological recovery. Although much work still needs to be done, it appears that the Siberian Trap eruptions were relatively short lived and within a million years or so atmospheric CO_2 levels declined to pre-extinction levels. Most tropical ecosystems on land and in the oceans did not fully recover until late in the Early Triassic, some 5 million years after the extinction.

Late Triassic

The Late Triassic biotic crisis records the second smallest extinction magnitude of the 'Big 5' events, but is ranked third in terms of its ecological impact (McGhee *et al.*, 2004). Although the least studied of the major events, it has attracted an increasing amount of scientific attention since the start of the twenty-first century due to present-day climate concerns: the Late Triassic event is associated with clear evidence of global warming and is the most severe crisis to have affected groups that make up the modern marine fauna (**Figures 1** and **3**).

Sepkoski's (1996) data show that 53% of all marine genera became extinct (40% when the data are filtered) and 22% of marine families. Among marine groups, Sepkoski's database shows significant extinctions among ammonoids, bivalves, gastropods, and brachiopods. Ammonoids experienced a boom-and-bust pattern throughout their history, and this event marks one of their largest crises. The group was quite diverse during the latest Triassic, but at least six superfamilies became extinct near the boundary, followed by a rapid rediversification in the Early Jurassic. Some estimates suggest that only a single ammonite genus survived the crisis. Although no family-level extinctions occurred among bivalves, approximately half the genera disappeared and turnover at the species level was particularly high. Endemic genera and those in deep offshore facies or associated with reefs appear to have experienced greatest extinction. Conodonts, a group of phosphatic microfossils formed by simple chordates, had survived all previous mass extinctions, but finally became extinct at the end of the Triassic. Extinction of foraminifera and ostracods was relatively minor. Reefs dominated by scleractinian corals first appear in the Middle Triassic and become prominent in the Carnian. These corals, associated sponges, and other reef taxa suffer a sharp reduction in diversity during this episode, followed by an interval during which reefs are absent.

Some terrestrial groups experienced extinction at this time as well, although disruption was less severe than during the Late Permian or end-Cretaceous events (McGhee *et al.*, 2004). Pollen records show that many seed ferns became extinct, as did some other plant groups, but a lack of a detailed floral record for the interval has hampered analysis in Europe. In eastern North

America the boundary is well preserved in the Newark Supergroup and approximately 60% of the pollen species disappeared within approximately 500,000 years. Similar patterns have been found in Arctic Canada but not in Australia or Siberia, possibly indicating considerable regional variation in extinction intensity. Of particular relevance is a sharp increase in the abundance of fern spores at the extinction, reminiscent of the fern spike associated with the K/T extinction. Whether or not any vertebrate extinctions occurred remains highly contentious. One analysis of the Newark group vertebrates indicated extensive extinctions, but this conclusion has been challenged by many vertebrate paleontologists who have argued that the Newark Supergroup record is sparse and there are no other data supporting an extinction of terrestrial vertebrates.

Debates concerning the causes of the Late Triassic extinction event are similar to those of the Late Permian. The hypothesis of an extraterrestrial cause was effectively rejected when accurate dating of the leading candidate crater, the Manicouagan impact crater in Quebec, showed that it formed some 13–14 million years before the extinction event, in the middle of the Norian stage. Other Late Triassic craters and horizons of impact ejecta likewise represent pre-Rhaetian impacts that appear to have caused no measureable rise in extinction rates (Walkden *et al.*, 2002). Most attention is now focused on sea-level changes and the effects of global warming triggered by flood basalt volcanism.

Evidence from fossil soils and from fossil leaves of plants related to modern *Gingko* has shown that atmospheric carbon dioxide levels rose from less than 1000 ppm to between 2000 and 5000 ppm across the Triassic/Jurassic boundary, with values returning to pre-extinction values within the Hettangian (Schaller *et al.*, 2011). This rise in carbon dioxide levels is hypothesized to have resulted in global warming, with corresponding effects on marine and terrestrial systems as those recorded for the Late Permian event. The similarity with the Late Permian crisis also extends to the involvement of a major igneous province as a source of the carbon dioxide. Recent age dates reveal that the Central Atlantic Magmatic Province, an extensive region of flood basalts extending across eastern North America, Europe, northwest Africa, and northeastern South America that marks the rifting of Pangaea and the initial opening of the Atlantic, erupted coincident with the Late Triassic mass extinction (Marzoli *et al.*, 1999). The eruptions would have released volcanic gases including carbon dioxide into the atmosphere, with serious consequences for climate, ocean circulation, and the biosphere.

End-Cretaceous

What caused the demise of the dinosaurs 65 Ma? This has been both a popular and a scientific question almost since the fossils of these remarkable beasts were first discovered in the early half of the nineteenth century. When the leading explanation became the impact of an extraterrestrial object, both popular and scientific interest exploded. Along with the dinosaurs, other large vertebrates disappeared as well as a wide variety of marine organisms, but overall extinction magnitudes during the end-Cretaceous event were much lower than those of the other four major intervals of biotic crisis (McGhee *et al.*, 2004). The end-Cretaceous event is the most recent, most studied and has the best fossil record of the five great mass extinctions.

Among marine organisms, 16% of the families and 47% of the genera disappeared. In common with other extinctions, some groups suffered more than others. Ammonites were wiped out completely, and the final extinction of this iconic group has been intensively studied. Problems of preservation and sampling mean that an instantaneous mass extinction can produce an apparently gradual pattern of disappearance leading up to the extinction horizon (MacLeod *et al.*, 1997). One way to address this problem is to undertake statistical analyses of fossil occurrences. In an analysis of 28 Maastrichtian ammonite species from the coast of northern Spain, Marshall and Ward (1996) demonstrated that 25% probably became extinct gradually during the several million years leading up to the Cretaceous/Paleogene (formally Cretaceous/Tertiary, or K/T) boundary; the disappearance of up to 35% is probably linked to an episode of sea level fall just prior to the boundary; and the remaining species probably went extinct right at the boundary (**Figure 4**; Marshall and Ward, 1996). Other groups that suffered catastrophic extinction include the planktonic foraminifera, a group of very abundant, floating single-celled organisms that produce calcareous skeletons. In contrast, benthic foraminifera, which live on and in the seafloor sediment, suffered ecological disruption but little apparent extinction.

Broadly speaking, benthic organisms suffered lower extinction rates than pelagic or planktonic ones. Even amongst the benthos, some groups were evidently more extinction prone. A detailed study of sea urchins by Smith and Jeffrey (1998) has shown that feeding strategy strongly correlates with survival, hinting that the extinctions may have been driven by changes in food supply. In contrast, amongst bivalves, the lineages that tended to survive were those with broader geographic ranges (Jablonski and Raup, 1995). The factors controlling survival during the end-Cretaceous mass extinction thus appear to have varied between groups and were also different to factors that appear to control evolution during normal, background times.

Unlike many invertebrates, the problem with vertebrates is that the scarcity of their bones makes them largely unsatisfactory for documenting the pace of the extinction. North America contains one of the best records of vertebrate diversity, and studies have shown that different groups experienced very different extinction patterns. Sharks, marsupial mammals, lizards, and dinosaurs suffered very high extinction, whereas frogs, salamanders, turtles, placental mammals and crocodiles suffered virtually none. Evidence is emerging that Mesozoic birds also suffered major extinction at this time. Both large and small vertebrates are part of the list of extinct forms, countering earlier claims that only large vertebrates became extinct. Animals in freshwater settings suffered much less extinction than fully terrestrial forms: approximately 90% of freshwater forms survived but only 12% of the land species survived. Most scientists accept that the last dinosaurs disappeared about a meter or so below the extinction horizon in North America, but there is still some uncertainty whether dinosaurs were in decline prior to the impact. A few remains are known from rocks that postdate the extinction, and while

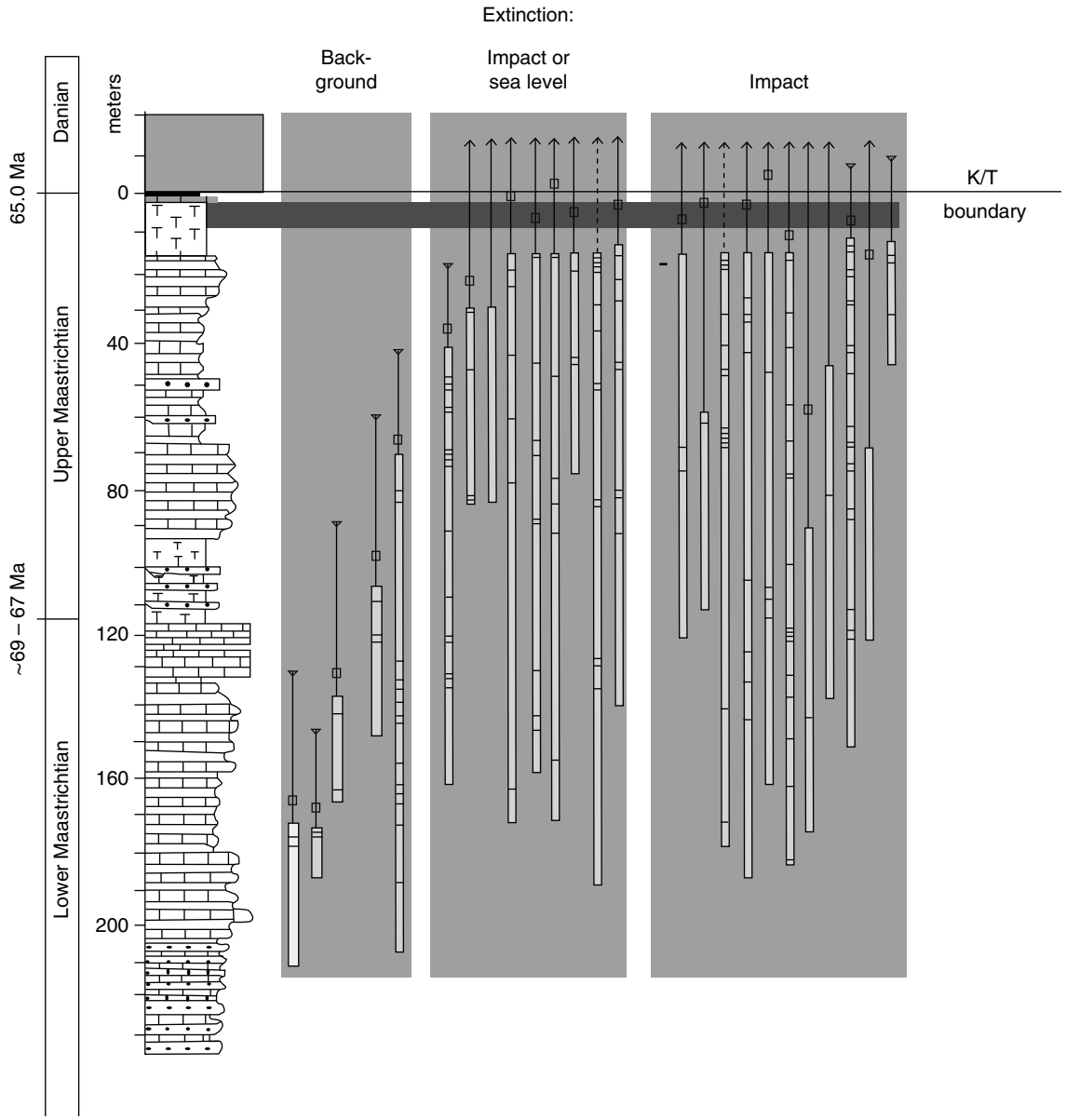


Figure 4 Graph showing the observed temporal ranges and occurrences (short horizontal lines) of 28 species of ammonites at a K/T (K/Pg) boundary section in Spain. Since the observed ranges underestimate true ranges, they can be corrected by statistical analysis. The 50% and 95% confidence intervals on the final occurrences are represented by the small squares and inverted triangles, respectively. Species that extend beyond the boundary are indicated by arrows on their ranges. The species are classified into three categories: (1) those that became extinct prior to the boundary, (2) those that may have become extinct due to either impact or a decrease in sea level shortly before the boundary, and (3) those that became extinct at the boundary. Discussion and further details are provided in Marshall and Ward (1996). Reproduced with permission from Pope KO, d'Hondt SL, and Marshall CR (1999) Meteorite impact and the mass extinction of species at the Cretaceous/Tertiary boundary. *Proceedings of National Academy of Science USA* 95: 11028.

claims for dinosaurs persisting after the end of the Cretaceous have generally been dismissed, the possibility remains that some did survive for a short time in those regions where the terrestrial extinction was much reduced.

Amongst terrestrial plant communities, a remarkable increase in the abundance of fern spores appears at the boundary in sites throughout North America and coincides with the iridium spike and the occurrence of shocked quartz. The interpretation is that these ferns represent the pioneering community that blanketed the bare ground following the extinction of the other plants. In western North America, an extinction of almost 80% of plant species has been documented. Further afield, however, the crisis is less evident so that in Antarctica and New Zealand the extinction is almost unrecognizable.

In 1980, Luis Alvarez and colleagues published one of the most influential geological papers of the twentieth century, in which they hypothesized that an elevated iridium

concentration in a thin clay layer at the Cretaceous/Palaeogene boundary section at Gubbio, Italy, was the result of the impact of a 10-km wide asteroid and that the impact caused the extinctions. Iridium is extremely rare in crustal rocks on Earth, but much more abundant in meteorites, which are the source of nearly all of the iridium present on the surface of the Earth. Any spike in iridium concentrations in sedimentary rocks suggests either a significant depositional hiatus or an increase in the rate of extraterrestrial material reaching the planet's surface. A spectacular way to enhance the delivery rate of iridium is through the impact of a large asteroid. As predicted by Alvarez *et al.* (1980), the iridium anomaly has since been discovered at many other K/T boundary sections, both marine and terrestrial, and other indicators of an extraterrestrial impact have also been recovered, including shocked quartz. The discovery of an impact crater of the correct age and approximately 200 km in diameter at Chicxulub in the Yucatan Peninsula, Mexico, was the final convincing piece of evidence that Alvarez *et al.* (1980) were correct and a large impact had indeed occurred. But, did this impact cause the extinctions?

The Chicxulub bolide is believed to have come from the southeast and North America probably felt the greatest effect of the blast, which might account for the higher terrestrial extinction rates there. Many of the supposed extinction-causing consequences of the impact have, however, proved controversial. The suggestion that terrestrial biotas were decimated by continental-scale wildfires is not supported by measurements of the amount of charcoal in rocks (Belcher *et al.*, 2003). The theory that dust from the impact would have plunged the world into a lengthy period of photosynthesis-sapping darkness is not supported by detailed modeling (Pope, 2002). Whether the amount of water vapor and sulfur dioxide thrown into the atmosphere by the impact would have significantly cooled the Earth for a long period and produced extensive acid rain is open to debate. The second hypothesis of the Alvarez *et al.* (1980) paper, that an extraterrestrial impact caused the end-Cretaceous extinction, is not universally accepted by all scientists.

The latest Cretaceous was a time of considerable environmental change and other causal mechanisms have been proposed. Although the Cretaceous was a greenhouse period, with elevated temperatures and carbon dioxide levels, the final 100,000 years were much cooler overall, with, as already noted, a major episode of global sea level fall happening just prior to the boundary. Some ammonite species, at least, disappeared then. In common with the Late Permian and Late Triassic mass extinctions, the end-Cretaceous event also coincides with the eruption of a large flood basalt province: the Deccan Traps in India. If the eruption of flood basalt provinces led to extinction-causing environmental changes at earlier times in Earth history, it seems unlikely that the Deccan eruptions would have had no effect on the Cretaceous biota. According to one study, ecological disruption of the plankton through the latest Cretaceous and earliest Palaeogene at a number of sites across the globe seems to correspond to intervals of eruptive activity (Keller, 2005). While the existence of an impact coincident with the extinction seems indisputable (e.g., Norris *et al.*, 1999), the various proposed kill mechanisms are disputed and the alternative hypotheses that climatic or environmental changes were responsible for at least some of the extinctions has yet to be rejected.

Lesser Mass Extinctions

Within the fossil record a number of smaller extinction events have been identified (e.g., Hallam and Wignall, 1997; for postextinction recoveries, see Erwin, 1998). Many of these events are not well studied, and a few only appear on large-scale, synoptic analyses (compare, e.g., some of the diversity troughs in **Figures 1** and **3**), while others may be statistical artefacts. Several of these events are significant however, and they provide a useful perspective on the larger mass extinctions. They also emphasize the continuum in biotic disturbances between the great Late Permian mass extinction and the less extreme but more common biotic crises, which do not qualify as mass extinctions. With current scientific endeavor seeking to predict future consequences of current climate change and biodiversity loss, these extinction events are highly relevant to the present day as they contain archives of global-scale experiments in global warming.

Neoproterozoic and Paleozoic

What happened to the enigmatic, soft-bodied Ediacaran biota of the late Neoproterozoic? This diverse assemblage of metazoans is found worldwide and includes a diverse assortment of cnidarian-grade organisms and a few probable ancestors to the clear metazoans of the Cambrian. Virtually all the Ediacaran forms disappear from the record prior to the base of the Cambrian. Whether this represents a mass extinction, a change in preservational style so that Ediacaran fossil were no longer preserved, or a gradual replacement of these forms by diversifying animals is unclear.

The Early Cambrian is characterized by a diverse assemblage of small shells, spicules and tubes, trilobites, and the sponge-like archaeocyathids. Two pulses of extinction near the end of the Early Cambrian each eliminated approximately 40–50% of marine genera. Reefs, small shelly fossils, and some brachiopods were particularly affected. These extinctions mark the end of this characteristic fauna and paved the way for the trilobite-dominated marine communities that dominate the remainder of the Cambrian. Metazoan reefs did not reappear again until the Ordovician, when the origination of the Paleozoic corals (tabulates and rugosans), bryozoans, and some sponges heralded a new suite of reef-builders that lasted until the Late Permian. The causes of these events are unclear, but marine anoxia, a decrease in sea level, and possibly changes in atmospheric composition have been suggested. Equally puzzling are a series of crises affecting near-shore trilobites and other marine benthos during the Middle and Late Cambrian in North America. Each event begins with a radiation of new trilobite families across marine shelves, evidently from deeper water ancestors. These new forms are initially widespread, with broad environmental tolerances. The new forms rapidly produce a diverse, endemic trilobite assemblage, in which individual taxa have relatively narrow environmental tolerances, that is then wiped out by a sharp extinction event that cuts across different marine environments. Brachiopods, conodonts, and other marine taxa are also affected. Although the extinctions were initially linked to decreases in sea level, with the diversifications occurring as sea

level increased, other paleontologists have suggested temperature decreases or incursions of low-oxygen waters. At least four of these events occurred from the Middle Cambrian into the earliest Ordovician, with the interval between extinctions being about 4 million years.

Mesozoic

A number of Mesozoic extinction events have been proposed. Two of the most prominent occurred in the earliest Toarcian and at the Cenomanian/Turonian boundary (Hallam and Wignall, 1997). If the Paleobiology Database analyses are correct, then the Jurassic/Cretaceous boundary also may be a significant event in the Mesozoic (**Figure 3**).

The Cenomanian/Turonian (C/T) mass extinction during the Late Cretaceous (93.5 Ma) eliminated approximately 8% of marine families and 26% of marine genera. Statistical analysis of extinction patterns from western USA suggests the extinction occurred in many discrete steps. Benthic organisms were particularly affected, whereas cephalopods and other groups living high in the water column suffered relatively little extinction. The crisis occurred during warm, equable climates, relatively high sea level, and is associated with marine anoxia and the formation of black shales. In England, for example, the C/T boundary forms a well-known "black band" in quarries. The greenhouse climate evidently led to sluggish oceanic circulation and warm, saline bottom waters, which stratified the oceans and aided in the formation of anoxic conditions (Harries and Little, 1999).

There are some remarkable similarities between the C/T event and an earlier mass extinction during the early Toarcian (Harries and Little, 1999). The extinction intensities were essentially identical, as were the environmental changes recorded, especially in northwestern Europe, which was particularly affected by this event. As with the C/T event, the early Toarcian crisis primarily eliminated benthic taxa, particularly bivalves, gastropods, and brachiopods, as well as many ammonites. In both extinctions diverse faunas were replaced by a much less diverse, stressed fauna dominated by epifaunal bivalves and ammonites. The early Toarcian extinction is clearly associated with global warming and the expansion of oxygen-poor waters in many ocean basins. As with the Late Permian event, however, anoxic waters were not present everywhere; and yet in regions with no record of oxygen stress, such as Spain, the extinction can still be recognized.

Cenozoic

The Cenozoic experienced many smaller biotic crises, each associated with climatic shifts, both cooling and warming. The most recent extinction – that of large mammals during the late Pleistocene – is also associated with human hunting, and whether climate or hunting was the primary cause is the subject of continuing debate.

The late Eocene to early Oligocene was an interval of heightened extinction both on land and in the oceans, although correlation between the two realms remains difficult. Approximately 80% of the foraminifera disappeared, and molluscs and echinoids (sea urchins) display a prolonged period of increased extinction, evidently due to a period of global cooling, which appears to have been the primary factor in the marine extinctions. The same is true on land, where the latest Eocene or earliest Oligocene marks the disappearance of many primitive mammal groups and the rise of modern mammals. Primates and rodents suffered considerable extinction, as did aquatic crocodiles and turtles (Hallam and Wignall, 1997).

The most recent mass extinction occurred within the past 100,000 years, and some claim it continues today. Many large mammal species disappeared from Europe, North and South America, and Australia; only in Africa were they unaffected. In North America, mammoths, mastodons, camels, saber-tooth cats, ground sloths, and a host of other animals disappeared approximately 10,000–12,000 years ago. A similar pattern is found in northern Europe. The leading explanation for this extinction is the overkill hypothesis, with human hunting the most important factor. Proponents of this view note that humans became important ecological factors in each region about the time that the pace of extinction increased. Opponents have favored climatically driven extinction, although the nature of the climatic change has been continuously modified to accommodate new information. The overkill hypothesis has also been extensively criticized.

Appendix

List of Courses

1. Courses Related to the Evolution of Life on Earth

See also: Carbon Cycle. Dinosaurs, Extinction Theories for. Extinction, Causes of. Extinction in the Fossil Record. Mammals (Late Quaternary), Extinctions of. Mass Extinctions, Concept of

References

Algeo TJ and Twitchett RJ (2010) Anomalous Early Triassic sediment fluxes due to elevated weathering rates and their biological consequences. *Geology* 38: 1023–1026.

Alvarez LW, Alvarez W, Asaro F, and Michel HV (1980) Extraterrestrial causes of the Cretaceous–Tertiary extinction. *Science* 208: 1095–1108.

Belcher CM, Collinson ME, Sweet AR, Hildebrand AR, and Scott AC (2003) Fireball passes and nothing burns – the role of thermal radiation in the Cretaceous–Tertiary event: Evidence from the charcoal record of North America. *Geology* 31: 1061–1064.

Bowring SA, Erwin DH, Jin YG, Martin MW, Davidek K, and Wang W (1998) U/Pb zircon geochronology and tempo of the end-Permian mass extinction. *Science* 280: 1039.

Copper P (1998) Evaluating the Frasnian–Famennian mass extinction: Comparing brachiopod faunas. *Acta Palaeontologica Polonica* 43: 137.

Copper P (2002) Reef development at the Frasnian/Famennian mass extinction boundary. *Palaeogeography, Palaeoclimatology, Palaeoecology* 181: 27–65.

Erwin DH (1993) *The Great Paleozoic Crisis*. New York: Columbia University Press.

Erwin DH (1998) The end and the beginning: Recoveries from mass extinctions. *Trends in Ecology and Evolution* 13: 344.

Finnegan S, Bergmann K, Eiler JM, *et al.* (2011) The magnitude and duration of late Ordovician-early Silurian glaciation. *Science* 331: 903–906.

Finney SC, Berry WBN, Cooper JD, *et al.* (1999) Late Ordovician mass extinction: A new perspective from stratigraphic sections in central Nevada. *Geology* 27: 215.

Hallam A and Wignall PB (1997) *Mass Extinctions and Their Aftermath*. Oxford: Oxford University Press.

Harries PJ and Little CTS (1999) The early Toarcian (early Jurassic) and the Cenomanian–Turonian (late Cretaceous) mass extinctions: Similarities and contrasts. *Palaeogeography, Palaeoclimatology, Palaeoecology* 154: 39–66.

Jablonski D and Raup DM (1995) Selectivity of end-Cretaceous marine bivalve extinctions. *Science* 268: 389.

Kearsey T, Twitchett RJ, Price GD, *et al.* (2009) Isotope excursions and palaeotemperature estimates from the Permian/Triassic Boundary in the Italian Dolomites. *Palaeogeography, Palaeoclimatology, Palaeoecology* 279: 29–40.

Keller G (2005) Biotic effects of late Maastrichtian mantle plume volcanism: Implications for impacts and mass extinctions. *Lithos* 79: 317–341.

Krystyn L, Baud A, Richoz S, *et al.* (2003) A unique Permian-Triassic boundary section from Oman. *Palaeogeography, Palaeoclimatology, Palaeoecology* 191: 329–344.

MacLeod N, Rawson PF, Forey PL, *et al.* (1997) The Cretaceous–Tertiary Biotic transition. *Journal of the Geological Society London* 154: 265.

Marshall CR and Ward PD (1996) Sudden and gradual molluscan extinctions in the latest Cretaceous of western European Tethys. *Science* 274: 1360.

Marzoli A, Renne PR, Piccirillo EM, Ernesto M, Bellieni G, and De Min A (1999) Extensive 200-million-year-old continental flood basalts of the central Atlantic magmatic province. *Science* 284: 616.

McGhee Jr. GR (1996) *The Late Devonian Mass Extinction*. New York: Columbia University Press.

McGhee GR, Sheehan PM, Bottjer DJ, *et al.* (2004) Ecological ranking of Phanerozoic biodiversity crises: Ecological and taxonomic severities are decoupled. *Palaeogeography, Palaeoclimatology, Palaeoecology* 211: 289–297.

Norris RD, Huber BT, and Self-Trail J (1999) Synchroneity of the K–T oceanic mass extinction and meteoritic impact: Blake nose, western north Atlantic. *Geology* 27: 419.

Pope KO (2002) Impact dust not the cause of the cretaceous–tertiary mass extinction. *Geology* 30: 99–102.

Sallan LC and Coates MI (2010) End-Devonian extinction and a bottleneck in the early evolution of modern jawed vertebrates. *Proceedings of the National Academy of Science* 107: 10131–10135.

Saunders A and Reichow M (2009) The Siberian Traps and the End-Permian mass extinction: A critical review. *Chinese Science Bulletin* 54: 20–37.

Schaller MF, Wright JD, and Kent DV (2011) Atmospheric pCO_2 perturbations associated with the central Atlantic magmatic province. *Science* 331: 1404–1409.

Sepkoski Jr. JJ (1984) A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinction. *Paleobiology* 10: 246–267.

Sepkoski Jr. JJ (1996) Patterns of Phanerozoic extinction: A perspective from global data bases. In: Walliser OH (ed.) *Global Events and Event Stratigraphy*, pp. 35–51. Berlin: Springer-Verlag.

Smith AB (2007) The shape of the Phanerozoic marine palaeodiversity curve: How much can be predicted from the sedimentary rock record of Western Europe? *Palaeontology* 50: 765–774.

Smith AB and Jeffrey CH (1998) Selectivity of extinction among sea urchins at the end of the Cretaceous period. *Nature* 392: 69–71.

Trotter JA, Williams IS, Barnes CR, Lecuyer C, and Nicoll RS (2008) Did cooling oceans trigger Ordovician biodiversification? Evidence from conodont thermometry. *Science* 321: 550–554.

Twitchett RJ, Looy CV, Morante R, *et al.* (2001) Rapid and synchronous collapse of marine and terrestrial ecosystems during the end-Permian mass extinction event. *Geology* 29: 351–354.

Walkden G, Parker J, and Kelley S (2002) A late Triassic impact ejecta layer in southwestern Britain. *Science* 298: 2185–2188.

Measurement and Analysis of Biodiversity

Wade Leitner and Will R Turner, The University of Arizona, Tucson, AZ, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 123–144, © 2001, Elsevier Inc.

Glossary

Abundance The number of individuals of a given species in a region.
Estimator A statistic calculated from data to estimate the value of a parameter.
Incidence The number of samples containing at least one of a given species in a census.
Sampling scope The area, interval of time, and taxonomic grouping over which sampling takes place.
Spatial heterogeneity In community sampling, variation in species capture probabilities across space; may be due to habitat variation, intraspecific clumping, interspecific association, or other factors.
Species abundance distribution The number of species found in each interval of abundance in a community.
Species accumulation curve A plot of the total number of species observed in a census against some measure of cumulative sampling effort.
Species diversity Any of many measures concerning the number of species in an area and/or the distribution of their abundances.
Species richness The number of species present in a region.
Temporal heterogeneity In community sampling, variation in species capture probabilities over time; may arise from temporal environmental variation, migration, speciation, or other factors.

Introduction

Why Measure Diversity?

Biological diversity pervades every aspect of community ecology. As the key distinguishing feature of communities, it forms the basis of many ecological studies. Many studies have used a wide range of techniques to quantify diversity. These have included studies focused on patterns of species number in both time and space. Projects with more mechanistic aims have employed diversity measures to draw inferences about processes influencing diversity (e.g., migration, habital selection, and competition at ecological timescales; and speciation, colonization, and extinction over evolutionary scales). The measurement of diversity applies to studies of the potential consequences of diversity (e.g., stability and resistance to invasion). In addition, diversity measurement plays a critical role in the study of human impacts on biological systems. Its uses in conservation include estimation of extinction rates due to habitat loss, development of conservation strategies, indication of effect due to disturbance, and use as a barometer of ecosystem status. These and other fields of inquiry require adequate methods for quantifying species diversity.

Measures of Diversity

Diversity measures are classified into two broad categories: species richness (SR) and shape of the species abundance distribution. The first simply reflects the number of species present in an area. The second effectively measures the probability distribution of population sizes of the species in an area. Evenness and equitability measures belong to the second category. Evenness is highest when all abundances have equal probability in a certain range, whereas equitability is highest when all species have the same population size. Depending on the question under investigation, both categories of diversity measures may be used to document patterns. Many hypotheses concerning patterns in diversity use the relatively simple concept of SR to make concise statements connecting pattern and process. Although richness simply refers to the number of species, it encompasses many sophisticated issues in sampling design and estimation. For these reasons, most studies use species richness. We restrict our focus to species richness as well.

The Problem

Species richness, the number of species present, is conceptually the most straightforward of diversity measures. Measuring richness, however, is not so simple. The number of species observed in a sample, S_{obs}, will always tend to underestimate the true species richness because we lack the resources for exhaustive sampling of communities. Even if we had such resources, properties of community structure may change during the course of sampling. This problem is exacerbated by the fact that many biological communities have many rare species, which are unlikely to be detected by sampling efforts. These issues are not unique to ecology.

The problem of determining the number of classes of objects in a collection has a long history. Many different disciplines have sought ways to estimate the number of undiscovered classes from the classes already observed. For example, how many dies were used to mint a collection of ancient coins? How many undiscovered bugs remain in a large computer program? No one method has been successful with all such problems because we often find that we have the most information where it is the least useful. That is, a few classes account for most of the observations, whereas a few

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00093-9

observations are scattered over most of the classes. Therefore, we wind up with small sample sizes for the rare observations where we need to characterize variation the most. In the most extreme case, imagine a community in which every species is represented by just one individual. Then, for any sample size below the true number of species we shall always observe fewer than the true number of species. Of course, in this case we could observe that the number of species was a linear function of the number of individuals. This might suggest to us a way to estimate the number of undetected species if we knew how many individuals were present in the study area. This example illustrates the strategy followed by richness estimators: Model the regularity in the behavior of the number of species detected as a function of sample size. Then, use this model to predict the number of species when the sample size becomes large. The regularity could involve a clear relationship between the average sample size and the average number of species observed; this is the basis of extrapolation-based estimators. On the other hand, the regularity could mean that we can replace a complicated sampling process with a more tractable model; estimators based on sampling theory follow this approach.

Article Overview

An exhaustive review of all methods used for estimations of species richness is beyond the scope of this article. Therefore, this article reviews the most widely used methods for the estimation of species richness. *Definition of Symbols* defines the symbols used in this article. *Theoretical Properties of Richness Estimators* lists some theoretical properties of richness estimators. *Practical Sampling Considerations* discusses practical considerations of using community sampling to estimate species richness. The next two sections present the estimators: richness estimators based on fine-scale, theoretical models of sampling are detailed in *Estimators Based on Sampling Theory*, whereas those based on coarse-scale, global modeling of species accumulation (extrapolation techniques) are discussed in *Estimators Based on Extrapolation*. *Which Estimation Method to Use?* addresses the practical problem of evaluating and selecting estimators for use on new data sets.

Definition of Symbols

We adopt the convention that random variables are written as uppercase symbols, whereas lowercase is reserved for deterministic variables or fixed constants. The definitions we shall need are given in **Table 1**.

Table 1 Symbols Used

s	Number of species in a community
n	Number of individuals in a community
x_i	Number of individuals of species i in a community
r_i	Number of species with i individuals in a community
S_{obs}	Number of species in a census
S_{est}	Estimate of the number of species in a community using given method
N	Number of individuals in a census
C	The sample coverage of the census
t	Number of samples in a census
X_i	Number of individuals of species i in a census
$X_{i,j}$	Number of individuals of species i in sample j
Y_i	Number of samples containing species i in a census
$Y_{i,j}$	The presence (1) or absence (0) of species i in sample j
R_i	Number of species containing exactly i individuals in a census
F_i	Number of species occurring in exactly i samples in a census
P_i	Probability of detecting species i in a sample
q_i	Probability that an individual is of species i
$I(\omega \in A)$	1 if $\omega \in A$ and 0 otherwise
$E[X]$	The expectation of the random variable X
VAR[X]	The variance of the random variable X
COV[X, Y]	The covariance of the random variables X and Y
CORR[X, Y]	The correlation of the random variables X and Y
$\gamma(Z)$	Coefficient of variation of the random variable Z

Theoretical Properties of Richness Estimators

Salient features of richness estimators include the following:

Type of data: Estimators differ in whether they use incidence or abundance data. The incidence of species i is the number of samples in which i occurs. Formally, $Y_i = \sum_{j=1}^{t} Y_{i,j}$. The abundance of species i is the number of individuals of i contained in the census. Formally, $X_i = \sum_{j=1}^{t} X_{i,j}$.

Sampling assumptions: Many estimators make assumptions about the sampling process (e.g., invariance of capture probabilities across all samples in a census).

Parametric/nonparametric: The estimator $S_{est} = \psi(\mathrm{X})$ is nonparametric if the function ψ does not depend on a given distribution of X.

Bias: Bias measures average deviation from the true value. We define bias $\equiv E[\psi(\mathrm{X})] - s$. An estimator of s is unbiased in the strict statistical sense if $E[\psi(\mathrm{X})] - s = 0$ for every s.

Variance: An estimator is a random variable. Therefore, we can use variance as a measure of the uncertainty in an estimate.

Sufficiency: Let $w = \psi(\mathrm{x})$ be an estimator of s. The statistic $W = \psi(\mathrm{X})$ is sufficient for s if $P\{\mathrm{X} \in A | W,\ s\} = P\{\mathrm{X} \in A | W\}$. Suppose $S_{est} = \psi(\mathrm{X})$ is sufficient for s. Then there are no other estimators (other than functions of ψ) that we could compute using X that could increase or decrease our confidence about our estimate of s.

Practical Sampling Considerations

Limiting the Influence of Sampling Biases

We lack the resources for complete censusing of communities. Thus, sampling is the window through which we view the ecological world. The species we observe when we sample are determined both by underlying ecological processes and by biases associated with sampling. The simplest of these is bias due to sample size. If we plot the number of species observed

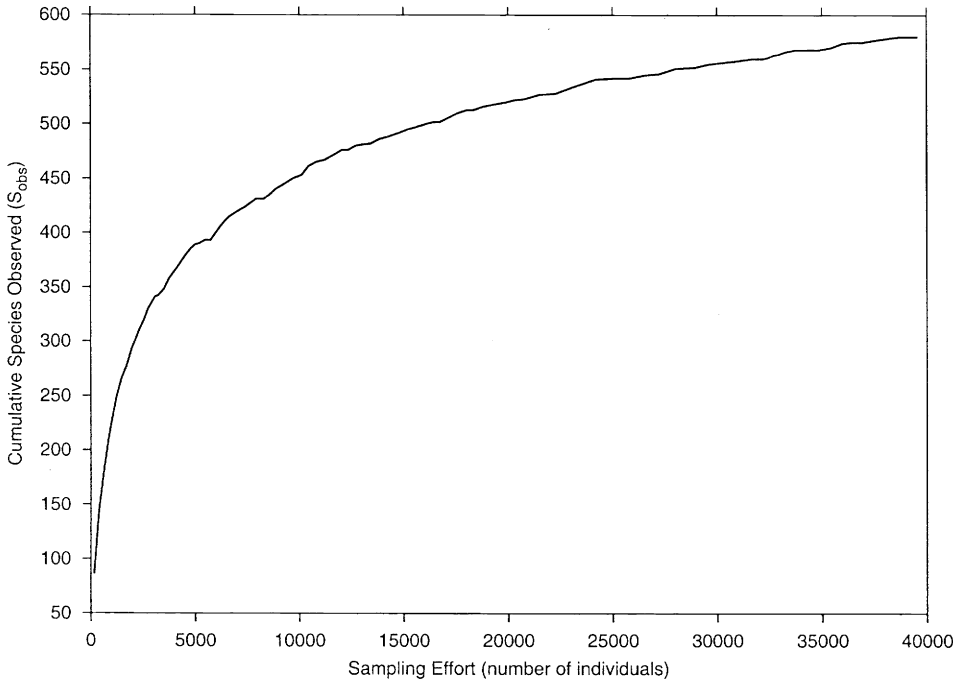


Figure 1 The species accumulation curve for a simulated data set. Note the nonlinear trend in the increase of species observed with increasing sample size.

in a census against some measure of sampling effort, such as individuals counted, we get a species accumulation curve as shown in **Figure 1**. Sample size bias accounts for much of the increase in a species accumulation curve with increasing sampling effort. This bias may be reduced by increasing sample size or by estimating the number of species actually present from incomplete census data. An increase in the accumulation curve also arises from sampling over different habitats and at different times. Spatial and temporal heterogeneity are difficult to quantify, as are their effects on the performance of richness estimators. It may be possible, however, to reduce such effects through careful sampling design. As a general rule, one should minimize sources of variation in capture probabilities over sampling. Where possible, the spatial and temporal scale of a census (see *Practical Sampling Considerations: Sampling Scope*) should be small relative to the scale of the processes generating diversity.

Sampling Scope

Any sampling effort can be described as taking place within some area, over an interval of time, and involving a particular taxon or group of taxa. Even within the area sampled, some habitats may not be sampled. Likewise, within an interval of time, some periods are often not considered (e.g., sampling that takes place over a week but never at night). We define sampling scope as the area (and habitats) sampled, the interval of time (and periods within that interval) over which sampling occurred, and the group of taxa sampled. The consideration of sampling scope is important because it means that the sampling design must be tailored to the question of interest.

The number of species predicted by richness estimators reflects the scope of sampling. In general, these methods estimate only the number of species in the taxa, area, and time interval sampled.[1] For example, estimates from ground-level plant surveys do not predict epiphyte diversity. Similarly, daytime sampling does not predict nocturnal animal species, estimates based on beetle abundance do not predict ant diversity, summer bird diversity does not inform us of winter migrants, and so on.

Abundance and Incidence

With quadrat-based sampling, the area to be sampled is partitioned into many subsamples (quadrats) of equal size. A subset of t of these quadrats is then sampled using capture or observation techniques appropriate for the taxon of interest. The data recorded may take the form of either species incidences $Y_{i,j}$ or abundances $X_{i,j}$.

Abundance data may not be possible to collect in all cases. For some taxa (e.g., some fungi), individuals are not readily recognizable as such. Additionally, some sampling methods may not allow abundances to be recorded. For example, point counts based on bird vocalizations are a practical method of recording species in bird communities but do not allow accurate counting of individuals. Even when abundances can be observed, obtaining abundances may require more effort than obtaining incidences alone. Nonetheless, a wider range of estimators may be used if abundance information is available.

[1]There are exceptions. Methods based on extrapolation of accumulation curves, in particular, might be used to estimate species richness at larger spatial (and temporal) scales than that sampled. There are many scaling considerations to be made that are unique to this type of extrapolation.

Thus, despite the greater effort involved in collecting abundance data, it may be useful to record species abundances so that abundance-based estimators may be used as well.

Other Sampling Issues: Surrogates for Quadrats and Species

Two other sampling issues bear mention here. First, it is important to note that the use of areal units (quadrats) for sampling units is not required. Indeed, although quadrats are often appropriate for plants and other sessile organisms, they are less so for mobile organisms. For mobile taxa, it is useful to substitute a more practical measure—such as observer-hours, trap-days, or volume of substrate—as a sampling unit. Additionally, all is not lost if sampling units are of unequal sizes. Equal-sized samples will meet the assumptions for a larger number of estimators. However, some estimators exist which may be applied when it is not possible to keep sample sizes approximately constant over a census.

Second, techniques for estimating species richness are potentially applicable to taxonomic levels other than species. For any of a variety of reasons, it is often impractical or impossible to key all organisms censused to the species level. For example, in studies of paleontological data, it may only be possible to place an individual to the generic level. In other cases, the presence of undescribed species or other lack of knowledge of species status may necessitate the use of a morphospecies concept. In any case, it is important to note that richness estimators will estimate the number of classes (genera and morphospecies in the previous examples) at whatever taxonomic level was used in sampling.

Estimators Based on Sampling Theory

Species richness estimators are all alike in the obvious way: They all estimate species richness. However, SR estimators take different approaches to estimation. This section presents approaches involving fine-scale, detailed models of the sampling process. *Estimators Based on Extrapolation* discusses extrapolative methods based on coarse-scale, global models of species accumulation.

Models of the Sampling Process: General Concepts

For the fixed collection from which we sample, $(r_0, \ldots, r_n)$ specifies the distribution of abundances among the species. Sampling introduces randomness into our data so that when we census the area, we record $(X_0, \ldots, X_{S_{obs}})$ from which we compute $(R_0, \ldots, R_N)$. When we sample a set of t quadrats, we may keep track of the abundance of the ith species within the j th quadrat, $X_{i,j}$. Then, we compute the abundance of species i by $X_i = \sum_{j=1}^{t} X_{i,j}$. However, in many cases, we may be unable to distinguish individuals. In these cases, we record $Y_{i,j}$ from which we compute the incidence of species i by $Y_i = \sum_{j=1}^{t} Y_{i,j}$. The distribution of $X_{i,j}$ depends on how we sample individuals.

Suppose we sample from a region (an urn) containing n individuals (marbles) distributed among s species (colors). Write x_i for the number of individuals contained in species i. We collect N individuals out of n and S_{obs} species out of s. If we sample with replacement, we find that the number of observations of species i, written X_i, has probability

$$P\{X_i = a\} = \binom{N}{a}\left(\frac{x_i}{n}\right)^a\left(1-\frac{x_i}{n}\right)^{N-a} \tag{1}$$

If we sample without replacement

$$P\{X_i = a\} = \frac{\binom{x_i}{a}\binom{n-x_i}{N-a}}{\binom{n}{N}} \tag{2}$$

Under certain conditions, both of these distributions can be approximated by the Poisson distribution according to

$$P\{X_i = a\} = \frac{(N\lambda_i)^a}{a!}e^{-N\lambda_i} \tag{3}$$

where $\lambda_i = x_i/n$. Strictly speaking, the approximation of Eq. (1) is exact only for communities containing infinite species richness and infinite abundance. Practically speaking, the Poisson approximation should work best on large, species-rich communities. A similar argument holds for the approximation of Eq. (2). The requirement that n be considered infinite is sufficient to get to Eq. (3), but it is not necessary. For example, if the successive times between discoveries of an individual of species i are exponentially distributed and the number of individuals found in disjoint samples are independent, then we will get Eq. (3). There is a natural way in which this can happen. First, detections of individuals must be independent events. Second, once we know the current value of X_i, the distribution of future values X_i must be completely determined. Then, we can expect that X_i can be modeled with a Poisson process.

Species Observed (S_{obs})

As a starting point, we consider the number of observed species, S_{obs}, as our first estimate of s. Our presentation of S_{obs} serves as the template for the more complicated estimators to follow. First, we consider the type of data we record. Then, we state the assumptions made in deriving the estimator and its properties. We next turn to theoretical properties of the estimator, for example, its bias and variance. Finally, we discuss issues related to the application of the estimator to real data.

Estimators may be applied to two types of data: species abundance and species incidence. If we record abundance, denoted $X_{i,j}$, our data set will be a matrix of the number of individuals of species i captured on sampling occasion j.

We define the abundance statistics, R_i, to be the number of species containing exactly i individuals. Then,

$$R_i = \sum_{j=1}^{s} I(X_j = i) \tag{4}$$

We define the incidence statistics, F_i, to be the number of species occurring in exactly i samples in a census. Then,

$$F_i = \sum_{k=1}^{s} I(Y_k = i) \tag{5}$$

When we wish to refer to the vector of R_i or F_i values for a given data set we will write **R** and **F**, respectively. The number

of species observed in a census S_{obs} may be written in terms of either the abundance or the incidence statistics: $S_{obs} = \sum_{i=1}^{N} R_i = \sum_{i=1}^{t} F_i$. In some cases, we use only the lower order statistics to correct estimator bias. We can visualize the difficulties inherent in estimator design by considering how information accumulates over the R_i. Let the number of species having fewer than k individuals be S_k. Similarly, N_k is the number of individuals representing these S_k species. Therefore,

$$S_k = \sum_{i=1}^{k} R_i \tag{6}$$

$$N_k = \sum_{i=1}^{k} iR_i \tag{7}$$

Figure 2 shows a plot of N_k and S_k vs k for real data collected on mist-netted birds. In **Figure 3**, we plot N_k as a function of S_k. This plot reveals an apparently exponential relationship between N_k and S_k. These figures show that information may be distributed very unevenly in the data set. This means that we shall have few data points (e.g., N_1 and N_2) where we wish to concentrate our estimation effort (S_1 and S_2). This is the worst possible distribution of information if we want to estimate the number of rare species from a simple transformation of the data. For example, an estimator using only $E[N]$ and $E[S]$ will probably do poorly for many real data sets.

Suppose we model sampling in terms of the probability of detection of individuals of each species, $\mathbf{q} = (q_1, \ldots, q_s)$, and the number of individuals observed, N. We assume that each individual of a species has the same capture probability throughout the entire census. Thus,

$$E[S_{obs} | (q_1, \ldots, q_s] = s - \sum_{i=1}^{s} (1 - q_i)^N \tag{8}$$

We record the incidence, $Y_{i,\,j}$, of species i during sample j giving a data matrix whose entries are either 0, for no detection, or 1 when any individuals of i get detected on occasion j. For an entire census, the number of detections of species i is $Y_i = \sum_{j=1}^{t} Y_{j,\mathrm{j}}$. Then we may define $\mathbf{p} = (p_1, \ldots, p_s)$ to be the vector of species detection probabilities. After t samples, we compute

$$E[S_{obs} | (p_1, \ldots, p_s)] = s - \sum_{i=1}^{s} (1 - p_i)^t \tag{9}$$

Equations (8) and (9) show that S_{obs} is biased over finite sampling. In the following, consideration of incidence and abundance versions of S_{obs} parallel on another. Therefore, we focus on the incidence data.

As we increase sample size we have for any i, $\lim_{t \to \infty} P\{Y_i > 0 | \mathbf{q}\} = 1$. However, in practical terms, we would like to know just how fast S_{obs} gets close to s. Suppose that the p_i are independent and identically distributed (iid) with distribution $H(p)$ and probability density $h(p)$ and consider $E[s - S_{obs}]/s$. The relative rate of convergence of the mean error is given by this expectation taken over all $\mathbf{p}$:

$$\frac{s - E_i[S_{obs}]}{s} = \int_0^1 (1 - p)^t h(p) dp \leq (1 - E[p])^t \tag{10}$$

When we have few rare species, E [p] will be close to 1 and S_{obs}/s will converge quickly to 1. Note that in absolute terms if s is large we may still have many rare, undetected species.

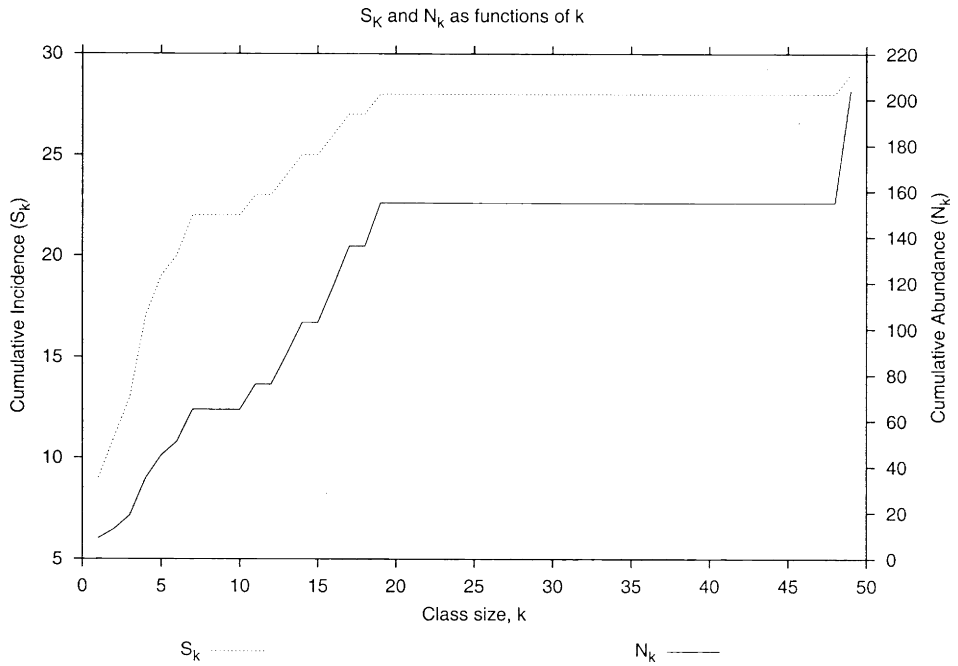


Figure 2 The accumulation of species and individuals with increasing abundance class size for a bird community in southeastern Arizona (W. Leitner, unpublished data). Note that N_k changes relatively slowly where S_k changes quickly and vice versa.

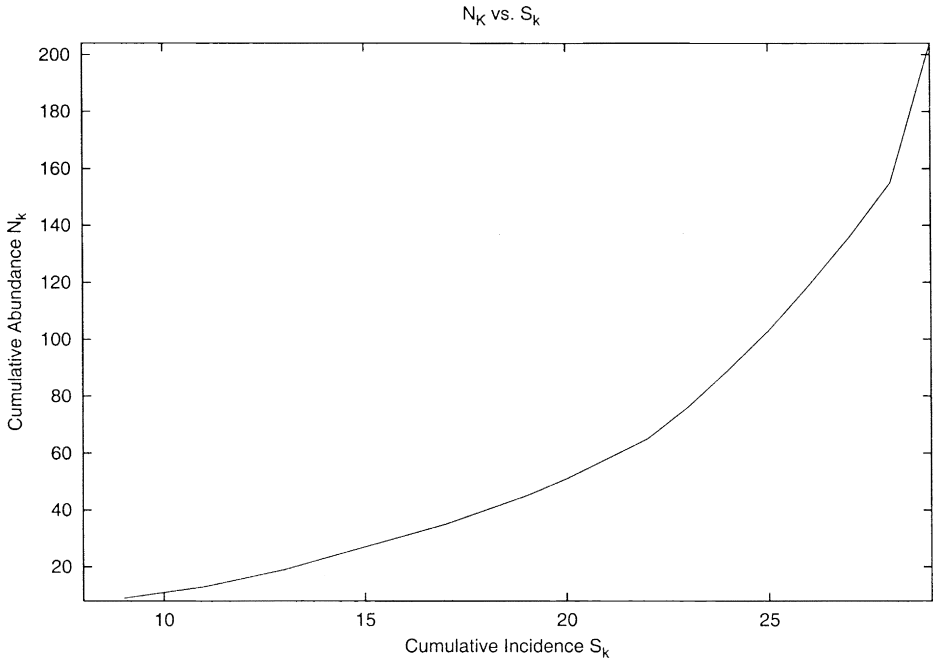


Figure 3 The accumulation of individuals as a function of accumulated incidence for a bird community in southeastern Arizona (W. Leitner, unpublished data). Note that N_k is nearly exponential with S_k. We shall have few data points (e.g., N_1 and N_2) at which we wish to concentrate our estimation effort (S_1 and S_2).

When we have a large proportion of rare species, the convergence will of course be much slower. Now, S_{obs} is a random variable and the variance in S_{obs} is

$$\text{VAR}(S_{\text{obs}}) = s\int_0^1 ((1-p)^t - (1-p)^{2t})h(p)dp \qquad (11)$$

The root mean square error will then decrease approximately like the square root of the error in the mean. Perhaps one of the indices of diversity that measure evenness could be used to sharpen confidence intervals in some way. Before undertaking the design of such a method, it would be beneficial to know if any function of **F** could be used to increase our confidence in our estimate of $E[S_{\text{obs}}]$. Burnham and Overton (1978) show that $(F_1, \ldots, F_t)$ are sufficient for **p** under the assumption of iid p_i. Now, the distribution of S_{obs} is determined by the distribution of **p** because $S_{\text{obs}} = \sum_{i=1}^{t} F_i$. Therefore, to increase our confidence in our estimator we should focus directly on reducing its variance. Thus, our goals in estimating SR should include getting estimators that decrease the bias faster than $(1-E[p])^t$ and whose standard errors decrease faster than $\sqrt{(1-E[p])^t}$. Often, these goals will conflict with one another. However, knowing that the bias decreases like $(1-E[p])^t$ tells us that for large enough t, the bias decreases faster than t^{-k}; $k = 1, 2, \ldots$. Resampling methods can use this fact to accelerate the rate of bias reduction.

Resampling Methods (S_{Jk}, S_{B})

We can very rarely expect

$$E[f(X)] = f(E[X]) \qquad (12)$$

When $f(x)$ is linear in x then Eq. (12) holds. Suppose that $f(x)$ is nonlinear. Then as x gets spread away from $x = E[X]$ the function f can over- or underweigh the contribution of X to the mean. Thus, both the nature of f and the distribution of X contribute to bias. For this reason, outliers (points far from $E[X]$) can severely alter the approximation $\text{E}[f(X)] \approx f(E[X])$. Most estimators derive from just such an approximation. Therefore, when we have outliers, we can sometimes improve our estimate by deleting them from the data set. A version of this approach applied to data not so clearly identified as extreme is the intuitive basis for the jackknife and bootstrap estimators. If the order in which data are recorded does not matter, then we should expect that the relationship between the distribution of t and $t-1$ data points should inform us about the relationship between the distribution of t and $t+1$ data points. Thus, the common property that underlies both techniques is the exchangeability of the data. There are two resampling approaches we consider: jackknifing and bootstrapping. In each case, we describe the resampling strategy that gives rise to the estimator. However, in each case it turns out that actual resampling need not be done. The **F** statistics contain the same information we would obtain from resampling.

Jackknife Estimator (S_{Jk})

Jackknifing involves computing the average value of a statistic on a reduced data set. One removes each combination of k data points in a data set and computes the statistic of interest to give a set of new pseudostatistics. Then, by taking a suitably weighted average of these statistics we get a reduced-bias version of the original statistic. The number of data points removed at a time gives the order of the jackknife. The most obvious statistic to which we can apply the jackknife is the

number of observed species, S_{obs}. As discussed previously, S_{obs} is a biased estimator of s. To apply the jackknife technique to S_{obs} we need three assumptions:

1. The capture probabilities may vary across species,
2. The capture probabilities do not change during the census,
3. The bias in S_{obs} decreases at least as fast as $1/t$.

Under these assumptions, Burnham and Overton (1978) propose a generalized jackknife estimator for species richness, simplified by Smith and Van Belle (1984):

$$S_{Jk} = S_{obs} - \frac{1}{k!}\sum_{j=1}^{k} F_j \sum_{i=j}^{k} (-1)^i \binom{k}{i} \frac{\binom{t-j}{i-j}}{\binom{t}{i}} (t-i)^k \quad (13)$$

where k gives the order of the estimator. S_{Jk} does not require that we actually do the resampling. To better understand the properties of S_{Jk}, we consider how the first-order jackknife estimate of S_{obs}, written S_{J1}, would be produced.

Suppose we compute a statistic, $D(X_1, \ldots, X_t) = D_t$ from t observations of X. Now, we want to consider the effect of each observation in turn by removing it from the data set. However, we do not want the order in which we do the removal to matter. Within a given sampling occasion, there may be time of day or other local differences. For example, during bird banding operations we may consistently catch flycatchers later in the day than warblers because it may take time for flying prey to become active. Then, resampling within a trapping occasion may introduce heterogeneity. Thus, we shall conduct jackknife resampling by adding or removing entire quadrats.

Let us remove the ith quadrat and compute D without this quadrat. Call this value $D_{(i)}$. Do this for each i. The mean of these $D_{(i)}$,

$$D_{t-1} = \frac{\sum_{i=1}^{t} D_{(i)}}{t} \quad (14)$$

gives rise to the first-order jackknife estimate

$$D_{J1} = tD_t - (t-1)D_{t-1} \quad (15)$$

Now, D_{J1} will be less biased than D_t if

$$E[D_t] = s + \sum_{i=1}^{\infty} \frac{b_i}{t^i} \quad (16)$$

where the b_i do not depend on t. If the bias of D_t decreases like $1/t$—that is, $b_1 \neq 0$—then the bias in D_{J1} decreases like $1/t^2$.

Suppose $D = S_{obs}$. If assumption 3 holds, then we can easily get

$$S_{J1} = S_{obs} + \frac{t-1}{t} F_1 \quad (17)$$

The higher order estimates are considerably more complicated but depend on the assumptions in the same way. As we increase the order of the jackknife, we get greater bias reduction. Unfortunately, we can expect that removing data to reduce bias might increase estimator variance. We see from Eq. (13) that S_{Jk} is a linear combination of the F_i with constant coefficients, $a_{i,\,k}$ given t and k. Burnham and Overton (1978) give the unconditional variance as

$$\mathrm{VAR}(S_{Jk}) = \sum_{i=1}^{t} a_{i,k}^2 E[F_i] - \frac{E[S_{Jk}]^2}{s} \quad (18)$$

Note that as we increase the order from k to $k+1$ the coefficient of F_{k+1} becomes nonzero and therefore increases the variance. Furthermore, a quick check of the lower order coefficients reveals that they also increase with k. Therefore, as we expected, the variance increases with k. Thus, we tradeoff bias reduction against increased noise. Two methods have been proposed for choosing the order of the jackknife. The first method tests the hypothesis that $E[S_{Jk} - S_{J,k+1}] = 0$ using the test statistic

$$T_i = \frac{S_{J,k+1} - S_{Jk}}{\sqrt{\mathrm{VAR}(S_{J,k+1} - S_{Jk} | D)}} \quad (19)$$

If we can reject this hypothesis for $k = 1$ then we proceed to test successive values of k until we fail to reject. The difficulty with this test is that quadrat information comes in discontinuous jumps so that the order used in estimation may jump around during the course of a census. Burnham and Overton (1978) suggest an interpolation technique that smoothes this behavior. The power of the jackknife technique is that it requires very little knowledge about the distribution of the data. It is thus nonparametric. An argument like the one presented previously on species observed tells us that assumption 3 is reasonable for finite collections of species. The first two assumptions, however, require biological information. For example, suppose we sample a site containing migratory animals. Then, when we sample determines, to a large extent the distribution of number of individuals observed. If we sample across a productivity gradient then we would not expect the distribution of a given species' abundance, $(X_{1,1}, \ldots, X_{1,t})$, to be symmetric. That is, the quadrats would not be exchangeable. This changes the importance of each quadrat on removal, making some have a deterministically larger effect. Consequently, the effective rate of bias reduction may become much smaller than $1/t^2$. Nonetheless, with attention to census design, such issues can be minimized. If in addition we assume that the p_i are iid random variables, then Burnham and Overton (1978) show that $\mathbf{F}$ is sufficient for the distribution of the capture histories. This result does not depend on the resampling protocol. This assertion will apply to some of the other estimators we examine.

Bootstrap Estimator (S_B)

Bootstrap estimation begins by constructing a surrogate data set by sampling the original data set with replacement. We shall bootstrap to remove bias in S_{obs}, so we shall once again focus on incidence-based data. Smith and van Belle (1984) analyzed the bootstrapping process to derive

$$S_B = S_{obs} + \sum_{i=1}^{S_{obs}} \left(1 - \frac{Y_i}{t}\right)^t \quad (20)$$

Just as in the jackknife estimator, we analyze the resampling process to get the estimate that would result if we actually did the resampling. We would like to think of resampling as a surrogate for doing multiple replicate

censuses. Let $s = \text{bias} + E[S_{obs}]$. If we knew the bias and $E[S_{obs}]$ we would be done. To estimate the bias over all possible sampling outcomes, we would need to know the distribution of the X_i. Let S_{obs} play the role of s, and let $(X_1, \ldots, X_{S_{obs}})$ approximate $(x_1, \ldots, x_s)$. Then, if we resample the X_i multiple times, we can estimate S^*_{obs}, the mean number of species found in the replicates. Then $\text{bias}_r = S_{obs} - S^*_{obs}$ is an estimate of $s - E[S_{obs}]$, the true bias. Now, we have a single point estimate of $E[S_{obs}]$, namely, S_{obs}. Therefore, we estimate s by $S_B = S_{obs} + \text{bias}_r$. We just need to compute bias_r. We assume that

1. Detection probabilities, p_i, may vary between species,
2. The p_i remain fixed during sampling.

Then, on resampling we find that

$$\text{bias}_r = \sum_{i=1}^{S_{obs}} \left(1 - \frac{Y_i}{t}\right)^t \tag{21}$$

which leads us to Eq. (20).

The variance in S_B given by Smith and Van Belle (1984) is

$$\text{VAR}(S_B) = \sum_{i=1}^{S_{obs}} \left(1 - \frac{Y_i}{t}\right)^t \left(1 - \left(1 - \frac{Y_i}{t}\right)^t\right)$$
$$+2\sum_{i=1}^{t}\sum_{i=j}^{t} \left(\left(\frac{Z_{ij}}{t}\right)^t - \left(1 - \frac{Y_i}{t}\right)^t \left(1 - \frac{Y_j}{t}\right)^t\right) \tag{22}$$

where Z_{ij} are the number of quadrats that jointly lack species i and j. For large t we can replace $(1 - Y_i/t)t$ by e_i^{-Y}. Then, as t gets large the variance in the estimator goes to zero as long as the tendency of species to co-occur decreases sufficiently fast as t increases. Therefore, species that interact strongly, for example, core members of a mixed foraging flock, will tend to inflate estimator variance.

Moment Estimators (α, S_{AM}, S_{IM})

These methods all derive relationships between the moments of various distributions. The moments of a distribution, defined by $E[X^r]$, where r is a positive integer, are used to simplify the analysis. At a minimum, the relationship between various moments reflects assumptions about the sampling process. In other cases, we require special assumptions about the abundance distribution.

Fisher's α

Fisher's α is intermediate between species richness and species diversity. It measures species richness but does so in terms of the parameters of the abundance distribution and the sampling process. It is a parametric estimator of richness when applied as described later. Unfortunately, α involves some rather technical arguments. This has led to two different ways to define α. The first follows from Fisher's use of the gamma distribution to model the abundance distribution. The second results from a limit taken on the functional relationship between the mean sample size and the mean number of observed species in a sample. As we shall see, properties that are exact for one definition of α are approximations for the other. We shall focus on two key properties of α: (i) sample size dependence of α and (ii) estimating species richness from α. We begin by defining α, as Fisher did, in terms of the sampling process and the abundance distribution under study. Then, we derive the second version of α from the first by considering what happens when the distribution under study represents an arbitrarily large number of species. Finally, we consider the practical problem of calculating and interpreting α when we know very little about the true nature of the abundance distribution.

Derivation of α

Fisher's α was first introduced (Fisher *et al.*, 1943) as a means for estimating the parameters of empirical abundance distributions that had a negative binomial shape. Fisher made three key technical assumptions in his derivation of α that show up in its properties. First, Fisher assumed that the sample was drawn from a community of infinitely many individuals. As we shall see, this allowed Fisher to use the Poisson process to model the sampling of individuals. Second, he assumed that the community was composed of species whose mean abundances were distributed according to a gamma distribution. Finally, Fisher considered the limiting case of a community containing an infinite number of species.

If we assume that the order in which we discover species is unimportant (*i.e.*, individuals are exchangeable) then $P\{N_1 = x\}$ specifies the probability that we observe a species represented by x individuals. With $P\{N_1 = x\}$ we can compute $E[S_{obs}]$ and $E[N]$ because

$$E[S_{obs}] = s\sum_{x=0}^{n} P\{N_1 = x\} \tag{23}$$

$$E[N] = s\sum_{x=0}^{n} xP\{N_1 = x\} \tag{24}$$

Fisher begins with the assumption that $P\{N_i = x | m_i\} = (m_i^x/x!)e^{-m_i}$, where m_i is the number of individuals of species i expected in the sample. The number of individuals in a sample, N, or the sample size, is usually a random variable because trapping and other forms of detection are random processes. Suppose that detection is a Poisson process, or very nearly so, and that our sampling region has extent A. Then, we expect that $m_i/A =$ where λ_i is the density of species i.

Fisher's second assumption was that the λ_i possess the gamma distribution with parameters $1/p$ and k. The density of a gamma random variable, $f(\lambda, p, k)$, is given by

$$f(\lambda, p, k) = \frac{1}{\Gamma(k)}\frac{1}{p}\left(\frac{\lambda}{p}\right)^{k-1} e^{-\lambda p}$$

where $\Gamma(k) = \int_0^\infty x^{k-1}e^{-x}dx$. The gamma distribution has three features which made it an obvious first choice. First, gamma random variables must be nonnegative just like species' densities. Second, the gamma distribution can take on a wide range of shapes. Lastly, computations using the gamma distribution are often tractable. To spread out densities evenly over a community, k must be close to zero.

Suppose we sample a region by censusing areas of increasing size. The larger areas should contain more

individuals. Let $\mu_A = E[N|A,s]$ and $v_A = E[S_{obs}|A, s]$ give the expected number of individuals and species, respectively, in a sample of size A drawn from a community containing s species. Then,

$$P\{N = x|A,s\} = \int_0^\infty \frac{(\lambda A)^x}{x!} e^{-\lambda A} f(\lambda) d\lambda \tag{25}$$

$$= (1+pA)^{-k} \frac{\Gamma(x+k)}{\Gamma(k)x!} \left(\frac{pA}{1+pA}\right)^x \tag{26}$$

Using Eq. (26) in Eqs. (23) and (24) it can be shown that

$$v_A = s(1-(1+pA)^{-k}) \tag{27}$$

$$\mu_A = skpA \tag{28}$$

Now, if the relative abundance of species is the same at all sampling scales it must be that the distribution's shape and thus k are fixed. Although we expect smaller sample sizes in smaller A, the number of species from which we sample remains fixed at s. This means that μ_A can only change as a function because of A or p. Therefore, we collect all the constants from Eq. (28) into α and write $\mu_A = \alpha\, pA$ and set $\alpha = sk$. This is Fisher's definition of α (with the effect of sampling extent made explicit). When defined this way, α depends on only the shape of the distribution and not on the expected sample size. Using this definition for α, we can rearrange Eq. (27) and substitute Eq. (28) to get

$$\left(1-\frac{v_A}{s}\right)^{-s\alpha} = 1+\frac{\mu_A}{\alpha} \tag{29}$$

Equation (29) defines μ_A as a function of v_A with parameters α and s. In fact, Eq. (27) is a species–area relationship. At $A = 0$ we expect no species and for $A \to \infty$ we expect s species. A little calculus shows that

$$\frac{dv_A}{d\ln(A)} = \frac{\alpha}{1+\frac{\alpha}{\mu_A}} \frac{s-v_A}{s} \tag{30}$$

Consequently, if we plot $\ln(s-v_A)$ vs $\ln(A)$ we get a line of slope $(\alpha/s)(1/(1+(\alpha/\mu_A)))$. Now, μ_A quickly exceeds α in many cases at very small areas. Thus, for most sample sizes we have constant slope α/s. Therefore, we may estimate s if we know α. Unfortunately, this will give only a lower bound on s because the sensitivity of the slope to changes in s decreases as s increases. Nonetheless, we need a method for estimating α that does not involve s. To do this, we will want to consider communities that contain many species. Again, calculus tells us that if we let $s \to \infty$ in Eq. (29) we get

$$e^{-v_A/\alpha} = 1+\frac{\mu_A}{\alpha} \tag{31}$$

The solution to Eq. (29) gives us the second way of defining α. In order to avoid confusion, we write α^* for the solution of Eq. (31). We do not need to assume anything about the sampling procedure or the abundance distribution, except that μ_A and v_A exist, in order to compute α^*. Note that $\alpha \to \alpha^*$ as $s \to \infty$. Unlike α, α^* may depend on sample size. This can be seen from

$$\frac{\partial \alpha}{\partial A} = \frac{-\alpha^{*2}}{\mu_A v_A + \alpha^* v_A - \alpha^* \mu_A} \left(\frac{\partial \mu_A}{\partial A} - \frac{\mu_A + \alpha^*}{\alpha^*} \frac{\partial v_A}{\partial A}\right) \tag{32}$$

because, in general,

$$\frac{\partial \mu_A}{\partial A} - \frac{\mu_A + \alpha^*}{\alpha} \frac{\partial v_A}{\partial A} \neq 0 \tag{33}$$

If we apply the assumptions Fisher made about sampling from a gamma abundance distribution, then

$$\frac{\partial \mu_A}{\partial A} - \frac{\mu_A + \alpha^*}{\alpha^*} \frac{\partial v_A}{\partial A} = p\alpha$$
$$\left(1 - \frac{\alpha(\mu_A+\alpha^*)}{\alpha^*(\mu_A+\alpha)} \left(1+\frac{\mu_A}{\alpha}\right)^{-\alpha/s}\right) \to 0 \text{ as } s \to \infty \tag{34}$$

Therefore, the independence of α from A carries over to α^*, once again, in the limit of large s. Most often, however, computations of α do not in fact use v_A or μ_A. Rather than average values, the raw numbers of species and individuals observed serve as estimates of the respective expectations. This leads to $\hat{\alpha}$ an estimate of α^*:

$$N = \hat{\alpha}(1 - e^{-S_{obs}/\hat{\alpha}}) \tag{35}$$

Unfortunately, Eq. (35) cannot be solved directly for $\hat{\alpha}$ in terms of S_{obs} and N.

We can get an approximate solution to Eq. (35) by expanding the function $\hat{\alpha}(N, S_{obs})$ about the point (μ_A, v_A) in a second-order Taylor series using the fact that $\hat{\alpha}(\mu_A, v_A) = \alpha$. Then, we find the bias is approximately

$$E[\hat{\alpha}(N, S_{obs}) - \alpha] \approx \frac{\alpha^2(\mu_A+\alpha)}{(\mu_A v_A + \alpha v_A - \alpha\mu_A)^3}$$
$$(v_A^2 \mathrm{VAR}(N) + \mu_A^2 \mathrm{VAR}(S_{obs}) - \mu_A v_A \mathrm{COV}(N, S_{obs})) \tag{36}$$

and the sample variance of $\hat{\alpha}$ is

$$\mathrm{VAR}(\hat{\alpha}) \approx \frac{\alpha^4}{(\mu_A v_A + \alpha v_A - \alpha\mu_A)^2}$$
$$\left(\mathrm{VAR}(N) + \left(\frac{\mu_A+\alpha}{\alpha}\right)^2 (\mathrm{VAR}(S_{obs}) + 2\mathrm{COV}(N, S_{obs}))\right) \tag{37}$$

To this point, we have needed no information regarding abundance distributions or sampling methods. Therefore, any conclusions we can draw from Eqs. (36) and (37) apply to any study system. Let $\gamma(X)$ be the coefficient of variation of X. Then, even for moderate sample sizes,

$$E[\hat{\alpha}(N, S) - \alpha] \approx \frac{\alpha^2}{v_A} (\gamma(N)^2 + \gamma(s)^2$$
$$-\gamma(N)\gamma(S_{obs})\mathrm{CORR}(N, S_{obs})) \tag{38}$$

For moderate A the sampling process will be nearly Poisson. Therefore, N will be Poisson because it is a sum of Poissons. Thus, $\gamma(N)$ will decline with sample size. As we increase sampling effort, A, it must be that v_A is nondecreasing because we cannot

undiscover species. For most collections, the probability of finding new species decreases as we increase sampling effort because we generally find common species first and rare species after prolonged sampling. Thus, $\gamma(S_{obs})$ must also decline with increases in A. From $-1 \leq \mathrm{CORR}(N, S_{obs}) \leq 1$ we can see that bias decreases as fast as α^2/v_A. This means that $\hat{a}$ should depend weakly on sample size. Once a sample large enough to get good estimates of the various coefficients of variation has been gathered, $\hat{\alpha}$ will show little further change. This holds for any abundance distribution for which the variances and means of N and S_{obs} exist. The fluctuations of $\hat{a}$ can be seen from Eq. (37) to decrease as α/v_A as well so that $\hat{\alpha}$ closely approximates α^* for large A. However, we should remember that α expresses a relationship between the means of N and S_{obs}. Therefore, we will need more information on the abundance distribution if we are to interpret the number of undiscovered species remaining in terms of remaining in terms of $\hat{\alpha}$.

Chao's Abundance-Based Estimator (S_{AM})

Abundance-based methods apply to data sets in which individuals can be readily identified and counted. From our data we compute the statistics $(R_1, \ldots, R_N)$ and seek an estimate of $E[R_0]$, the expected number of unobserved species. Chao (1984) proposed

$$S_{AM} = S_{obs} + \frac{R_1^2}{2R_2} \tag{39}$$

as a lower bound for s in the limit of infinite sample size.

S_{AM} makes the following two assumptions:

1. Capture probabilities, q_j, may vary among species
2. Capture probabilities remain fixed during sampling

where q_j gives the probability of capture of an individual of species j. The starting point for the derivation of S_{AM} begins with a model of the sampling process.

Using the Poisson approximation as described previously, we get

$$E[R_i] \approx \sum_{j=1}^{s} \frac{(Nq_j)^i}{i!} e^{-Nq_j} \tag{40}$$

From Eq. (40) we see that this approximation works best when i is small relative to N and q_i is small. Thus, this method should perform best when applied to collections that have relatively few common or abundant species but many rare species. The exponential form of the Poisson approximation permitted Chao (1984) to do some clever analysis using a judiciously chosen distribution function, $F_{AM}(x)$. The key consequence of the Poisson approximation is that the ith moment of $F_{AM}(x)$, written μ_i, can be used to estimate $E[R_{i+1}]$ because

$$\mu_i \approx (i+1)! \frac{E[R_{i+1}]}{E[R_1]}. \tag{41}$$

Next, Chao (1984) connects the seen with the unseen species through

$$E[R_0] \approx E[R_1] \int_0^N \frac{1}{x} F_{AM}(dx) \tag{42}$$

Once again, we have used the Poisson approximation to estimate the expected number of species. This integral representation reveals two pathways to further analysis: approximate the integrator or approximate the integrand. Chao (1984) approximated the integrator rather than the integrand. By constructing a suitable distribution function that possesses the same first and second moments as $F_{AM}(x)$, Chao shows that the smallest value taken by Eq. (42) must be $E[R_1]/\mu_1$. Using Eq. (41) and taking the R_i as estimates of $E[R_i]$, we find

$$S_{AM} = S_{obs} + \frac{R_1^2}{2R_2} \tag{43}$$

The derivation incorporates the following assumptions into S_{AM}:

1. Sampling is a Poisson process,
2. Third-order and higher moments of F_{AM} may be neglected,
3. The R_i are good estimates of $E[R_i]$

These assumptions determine the statistical properties of S_{AM}. The first assumption could be used to get the sufficiency of $(R_0, \ldots, R_N)$ if the q_i were iid random variables with distribution $H(q)$. Then,

$$P(r_0, \ldots, r_N) = \binom{s}{r_0, \ldots, r_N} \prod_{i=0}^{N} [\theta_i(H)]^{r_i} \tag{44}$$

where $\theta_i(H) = \int_0^1 \binom{N}{i} q^i (1-q)^{N-i} h(q)$. Paralleling Burnham and Overton (1978), Eq. (44) can be rewritten to permit use of a factorization theorem to assert the sufficiency. If we relax the assumption that the q_i be identically distributed, then we need a model for selecting $H_i(q)$, the distribution function for q_i. For example, if the H_i corresponded to habitat heterogeneity, then the sufficient statistics would presumably incorporate information on variation in habitat type during sampling. The first and the last assumptions allow us to avoid assuming a form for the abundance distribution. S_{AM} is therefore nonparametric. Assumption 2 makes S_{AM} a lower bound for s that increases as sample size increases, indicating that S_{AM} is a biased estimator. Assuming that we have a large enough sample size, we expand the function S_{AM} as a first-order Taylor polynomial about $E[S_{obs}]$, $E[R_1]$, and $E[R_2]$. Then we can use

$$\begin{aligned} \mathrm{VAR}(aX + bY) = a^2\mathrm{VAR}(X) + b^2\mathrm{VAR}(Y) \\ +2ab\mathrm{COV}(X, Y) \end{aligned} \tag{45}$$

to estimate the variance. Working through the analysis and assuming the covariance terms are small gives

$$\begin{aligned} \mathrm{VAR}(S_{AM}) \approx \left(1 + \frac{R_1}{R_2}\right)^2 \mathrm{VAR}(R_1) \\ + \left(1 - \frac{R_1^2}{2R_2^2}\right)^2 \mathrm{VAR}(R_2) + \sum_{i=3}^{N} \mathrm{VAR}(R_i) \end{aligned} \tag{46}$$

Even when we have the abundance distribution, the indicated variances can be challenging to compute. Chao (1984) used bootstrapping techniques coupled with simulations and tests on well-characterized data sets to evaluate the performance of the S_{AM}. When applied to data sets with known s, S_{AM}

generally had smaller bias and narrower confidence intervals than S_{J1} or S_{J2}.

Chao's Incidence-Based Estimator (S_{IM})

Incidence-based methods apply to data sets in which individuals are not readily identified or counted. We compute $(F_1, \ldots, F_N)$ from the data matrix $X_{i,j}$ and seek an estimate of $E[F_0]$, the expected number of unobserved species. Chao (1987) used the same approach as used for S_{AM} to derive

$$S_{IM} = S_{obs} + \frac{F_1^2}{2F_2} \tag{47}$$

for data sets of species incidence. Like the jackknife estimator, the original application of this method was to estimate population sizes. It produces a lower bound for s. The starting point for the derivation of S_{IM} begins as S_{AM} did with a model of the sampling process. Let p_i be the probability of detection of species i. In this case, sampling is as in the jackknife estimator. Therefore, we assume

1. Detection probabilities, p_i, may vary between species.
2. The p_i remain fixed during sampling.
3. The p_i are iid random variables with distribution $H(p)$.

As we saw for S_{JK}, if these assumptions are met then

$$(F_0, \ldots, F_t)$$

will be sufficient for H. The last assumption allows us to write

$$E[F_i] = \int_0^1 \binom{t}{i} p^i (1-p)^{t-i} h(p)\,dp \tag{48}$$

From here we can once again use the Poisson approximation to simplify Eq. (48). In this case, the focus on data sets that concentrate probability over F_1 and F_2 is less critical. The Poisson approximation will still work well if the probability of large p_i is small. From here, Chao (1987) parallels the derivation of S_{AM} using an absolutely continuous version of $F_{AM}(x)$. This distribution function, $F_{IM}(x)$ has moments that satisfy

$$\mu_i \approx (i+1)!\ \frac{E[F_{i+1}]}{E[F_1]} \tag{49}$$

An integral relationship between F_0 and F_1 analogous to Eq. (42) once again leads us to approximate either integrand or integrator. While S_{IM} arises from approximating the integrator as before, we gain some insight into the relationship between S_{JK} and S_{IM} by considering the other alternative first. If we approximate the integrand, $1/x$, with a kth order polynomial then we find

$$E[F_0] \approx \sum_{i=1}^{k} (-1)^{i+1} \binom{k}{i} E[F_i] \tag{50}$$

Now, the bias should decrease as the expected number of undiscovered species decreases. This expectation, given by

$$E[F_0] = \int_0^1 (1-p)^t h(p)\,dp \tag{51}$$

must decrease at least as fast as $1/t$ because $(1-p)^t \le e^{-tp}$ for $0 \le p \le 1$. Equation (50) has, as Chao (1984) points out, the same form of the kth-order jackknife estimator. The connection between S_{JK} and S_{IM} reflects the assumption made in any jackknife estimator that the bias can be expressed as a power series in $1/t$. Chao (1987) obtains S_{IM} by approximating the integrator rather than the integrand. Then, using an analog to Eq. (41) and taking the F_i as estimates of $E[F_i]$, Chao proposes

$$S_{IM} = S_{obs} + \frac{F_1^2}{2F_2} \tag{52}$$

As with S_{AM}, we have several assumptions built into the derivation of S_{IM}:

1. The p_i are small enough to justify the Poisson approximation to the binomial probabilities of Eq. (48).
2. Third-order and higher moments of F may be neglected.
3. The F_i are good estimates of $E[F_i]$,

First, we note that S_{IM} is nonparametric. Assumption 2 makes S_{IM} a lower bound for s that increases as sample size increases, indicating that S_{IM} is a biased estimator. Chao (1987) estimates the variance of S_{IM} as

$$\mathrm{VAR}(S_{IM}) = \frac{F_2}{4}\left(\left(\frac{F_1}{F_2}\right)^4 + 4\left(\frac{F_1}{F_2}\right)^3 + 2\left(\frac{F_1}{F_2}\right)^2\right) \tag{53}$$

Equation (53) can be used to get approximate confidence intervals for S_{IM} provided that $p_i t$ is not too large. As we increase t, we expect the lower order $F_i \to 0$ if we are sampling a finite collection of individuals. This would seem to imply that the variance goes to zero with increasing t. However, large t means large $p_i t$ and so makes Eq. (53) a poor approximation. Chao (1987) used real data sets with known s to evaluate S_{IM}, finding that it performed well relative to S_{JK}. For low t, S_{IM} displayed smaller standard error and less negative bias than S_{J2}. However, as t was increased, S_{JK} outperformed S_{IM}. Notably, the data sets tended to have more recaptures with increasing t. Chazdon *et al.* (1998) found similar results.

Estimates Derived from Sample Coverage (S_{AC}, S_{IC})

Sampling of a finite community must eventually capture each individual multiple times. Thus, we can expect that exhaustive sampling would lead to $F_i = 0$ and $R_i = 0$ for the lower order i. An alternative that measures representation of species in the data set while retaining analytic tractability is sample coverage. The sample coverage, C, is defined by

$$C = \sum_{i=1}^{s} q_i I(X_i > 0) \text{ for abundance data or}$$

$$C = \sum_{i=1}^{S} p_i I(Y_i > 0) \text{ for incidence data}$$

Chao's Abundance-Based Coverage Estimator (S_{AC})

Chao and Lee (1992) propose a nonparametric estimate of s using estimates of $E[S_{obs}]$, $E[C]$, and $\gamma(q)$. They define the abundance coverage-based estimator S_{AC} as follows:

$$S_{AC} = \frac{E[S_{obs}]}{E[C]} + \frac{E[R_1]}{E[C]}\gamma(q)^2 \tag{54}$$

Intuitively, we see that S_{obs} is inflated by C to adjust for the fraction of the distribution left uncovered. Additional bias correction reflecting the shape of the distribution comes through R_1 and the γ of the distribution. Thus, we are using information from more than just the lower order abundances and may expect S_{AC} to perform better than S_{AM} and S_{IM} when higher order abundances support significant probability. We start with the same two basic assumptions that apply to S_{AM}:

1. Individual detection probabilities, q_i, may vary between species.
2. The q_i remain fixed during sampling.

Thus, the q_i can be regarded as a sequence of random variables with mean $\overline{q}$, variance σ_q, and coefficient of variation $\gamma(q)$.

Then, we find that

$$E[S_{obs}] = s - \sum_{i=1}^{s} (1 - q_i)^N$$

$$E[C] = 1 - \frac{\sum_{i=1}^{s} (1 - q_i)^N}{\sum_{i=1}^{s} q_i}$$

Of course, if we knew the true values of the q_i we could compute s directly. Instead, we get estimates of $E[S_{obs}]$ and $E[C]$ from our data. Now, $E[S_{obs}]$ and $E[C]$ are functions of s and the q_i. Therefore, we would like to eliminate the dependence on the q_i in order to estimate s. If our estimates of $E[S_{obs}]$ and $E[C]$ are close to the truth, we can use Taylor's theorem to expand about the point $(q_1, \ldots, q_s)$ to get

$$s \approx \frac{E[S_{obs}]}{E[C]} + \frac{E[R_1]}{E[C]} \gamma(q)^2 \tag{55}$$

Now, all we need are estimates of the expectations and the coefficient of variation for the abundance distribution. Chao and Lee (1992) suggest

$$E[S_{obs}] \approx S_{obs} \tag{56}$$

$$E[R_1] \approx R_1 \tag{57}$$

$$E[C] \approx \hat{C} = 1 - \frac{R_1}{N} \tag{58}$$

for the various expectations. The coefficient of variation is somewhat more complicated. First, it must always be positive. Second, when the true γ is large, we can expect to need more data to estimate it. It can be easily shown that

$$\gamma^2 = \frac{1}{s} \frac{\sum_{i=1}^{s} q_i^2}{\overline{q}^2} - 1 \tag{59}$$

and

$$E\left[\sum_{i=1}^{N} i(i-1)R_i\right] = n(n-1)\sum_{i=1}^{s} q_i^2. \tag{60}$$

Chao and Lee (1992) then suggest that

$$\hat{\gamma}^2 = \max\left\{\hat{N}_0 \frac{\sum_{i=1}^{N} i(i-1)R_i}{\left(\sum_{i=1}^{N} iR_i\right)^2} - 1, 0\right\} \tag{61}$$

where $\hat{N}_0 = S_{obs}/\hat{C}$ when $\hat{\gamma}^2 \le 0.5$. Then we have

$$S_{AC} = \frac{S_{obs}}{\hat{C}} + \frac{R_1}{\hat{C}} \hat{\gamma}^2 \tag{62}$$

In Eq. (61) the term $\hat{N}_0$ is an initial estimate of s. If $\hat{\gamma}^2 > 0.5$ the bias can be reduced by using S_{AC} in place of $S_{obs}/\hat{C}$ to yield

$$\tilde{\gamma}^2 = \max\left\{S_{AC} \frac{\sum_{i=1}^{N} i(i-1)R_i}{\left(\sum_{i=1}^{N} iR_i\right)^2} - 1, 0\right\} \tag{63}$$

Then, for larger γ we have

$$\tilde{S}_{AC} = \frac{S_{obs}}{\hat{C}} + \frac{R_1}{\hat{C}} \tilde{\gamma}^2 \tag{64}$$

It turns out that $S_{obs}/\hat{C}$ is biased low when the q_i differ from one another (when sample sizes are large enough to permit estimation γ (q)). Thus, S_{AC} should be biased low. The variance of S_{AC} is approximately a linear combination of $COV(R_i, R_j)$. It can be shown that

$$VAR(R_i) = E[R_i]\left(1 - \frac{E[R_i]}{s}\right) \tag{65}$$

$$COV(R_i, R_j) = -\frac{E[R_i]E[R_j]}{s} \tag{66}$$

Now $R_i \le s$ so the variance is bounded and decreases with increasing richness. However, for a fixed s, variance depends on how the R_i spread out as we increase t. This in turn depends on the sampling model we use. Here, we have assumed that the q_i remain fixed so that we sample with replacement. Therefore,

$$E[R_i] = \sum_{j=1}^{s} \binom{N}{i} p_j^i (1 - p_j)^{N-i} \tag{67}$$

Evidently, as N increases, $R_i \to 0$ with probability 1. Now, the R_i represent N different abundance classes. However, we have only s species. Therefore, the probability that $R_i > 0$ will be spread out over an increasing set of the R_i as N increases. Thus, $E[R_i]$ should become small relative to s as N increases. The simulations performed by Chao and Lee (1992) indicate that at high sample coverage S_{AC} is nearly unbiased. However, Colwell and Coddington (1994) found that S_{AC} tends to overestimate s when sample sizes are small. The difficulty appears to occur for large values of γ. In this case, many of the individuals in a sample represent common species, so there are a few dominant q_i in the collection. In this case, $\hat{C}$ is a poor estimate of sample coverage. However, we can easily ascertain the number of common species in a collection. Chao *et al.* (1993) suggest truncation of the data set to species with k or fewer representatives. The new estimates for $E[C]$ and $E[S_{obs}]$

that result are

$$S_{\text{obs},k} = \sum_{i=1}^{k} R_i \tag{68}$$

$$\hat{C}_k = 1 - \frac{R_1}{\sum_{i=1}^{k} iR_i} \tag{69}$$

Similarly, we have

$$\hat{\gamma}_k = \max\left\{\frac{S_{\text{obs},k}}{\hat{C}_k}\frac{\sum_{i=1}^{k} i(i-1)R_i}{\left(\sum_{i=1}^{k} iR_1\right)^2} - 1, 0\right\} \tag{70}$$

so that

$$S_{\text{AC}k} = S_{\text{obs}} - S_{\text{obs},k} + \frac{S_{\text{obs},k}}{\hat{C}_k} + \frac{R_1}{\hat{C}_k}\ \hat{\gamma}_k^2 \tag{71}$$

Chao's Incidence-Based Coverage Estimator (S_{IC})

The incidence-based coverage estimator, S_{IC}, applies to data sets that have the quadrat structure outlined for S_{IM}. Briefly, when we have the presenceabsence of data then incidence-based estimators apply. Lee and Chao (1994) found ways to relax many of the assumptions present in S_{AM}, S_{IM}, S_{JK}, and S_{AC}. They follow Pollock (1976) by partitioning variability into between-species, within-species, and temporal components. The results involve more complex notation and more complicated proofs but use the same conceptual framework as S_{AC}. For brevity, we present only the case of unequal detection probabilities among species. The estimator for incidence-based coverage is

$$S_{\text{IC}} = \frac{E[S_{\text{obs}}]}{E[C]} + \frac{E[F_1]}{E[C]}\gamma(p)^2 \tag{72}$$

We define p_i as the probability of detection of species i. For S_{IC} we assume that

1. Detection probabilities, p_i, may vary between species.
2. The p_i remain fixed during sampling.

Thus, the p_i can be regarded as a sequence of random variables with mean $\bar{p}$, variance σ_p, and coefficient of variation $\gamma(p)$.

Then, we find that

$$E[S_{\text{obs}}] = s - \sum_{i-1}^{s}(1-p_i)^t$$

$$E[C] = 1 - \frac{\sum_{i=1}^{s}(1-p_i)^t}{\sum_{i=1}^{s} p_i}$$

$E[D]$ and $E[C]$ are functions of s and the p_i. Therefore, we would like to eliminate the dependence on the p_i in order to estimate s. We can use Taylor's theorem to expand about the point $(p_1, \ldots, p_s)$ to get

$$s \approx \frac{E[S_{\text{obs}}]}{E[C]} + \frac{E[F_1]}{E[C]}\gamma(p)^2 \tag{73}$$

Using the approximations

$$E[S_{\text{obs}}] \approx S_{\text{obs}} \tag{74}$$

$$E[F_1] \approx F_1 \tag{75}$$

$$\hat{C} = 1 - \frac{F_1}{\sum_{i=1}^{t} iF_i} \tag{76}$$

$$\hat{C} = 1 - \frac{(t-1)F_1 - 2F_2}{(t-1)\sum_{i=1}^{t} iF_i} \tag{77}$$

and

$$\hat{\gamma}_2 = \max\left\{\hat{N}_0\frac{\sum_{i=1}^{t} i(i-1)F_i}{\left(\sum_{i=1}^{t} iF_i\right)^2} - 1, 0\right\} \tag{78}$$

where $\hat{N}_0 = S_{\text{obs}}/\hat{C}$ when $\hat{\gamma}^2 \le 0.5$, we get

$$S_{\text{IC}} = \frac{S_{\text{obs}}}{\hat{C}} + \frac{F_1}{\hat{C}}\hat{\gamma}^2 \tag{79}$$

If $\hat{\gamma}^2$ is large, the bias can be reduced by using S_{IC} in place of $S_{obs}/\hat{C}$ to compute $\tilde{\gamma}^2$. Then, using $\tilde{C}$ for $\hat{C}$, we have

$$\hat{S}_{\text{IC}} = \frac{S_{\text{obs}}}{\hat{C}} + \frac{F_1}{\hat{C}}\tilde{\gamma}^2 \tag{80}$$

As $\hat{C}$ increases, the bias must decrease. Therefore, we can expect that the assumptions of fixed p_i will lead to S_{IC} underestimating s in most cases. The variance of S_{AC} is approximately a linear combination of COV(F_i, F_j). It can be shown that

$$\text{VAR}(F_i) = E[F_i]\left(1 - \frac{E[F_i]}{s}\right) \tag{81}$$

$$\text{COV}(F_i, F_j) = -\frac{E[F_i]E[F_j]}{s} \tag{82}$$

The variance is thus bounded and decreases with increasing richness because $F_i \le s$. Now, as we increase t each of the species will be detected a different number of times with probability 1 (even if they have the same p_i). Therefore, there will be at most s nonzero F_i among $(F_1, \ldots, F_t)$. Thus, $E[F_i] \to 0$ as $t \to \infty$. As with S_{AC}, the effects of a few dominant p_i can be adjusted for using the methods of Chao *et al.* (1993). If we truncate the data set to species with k or fewer representatives. The new estimates for $E[C]$ and $E[D]$ that result are

$$S_{\text{obs},k} = \sum_{i=1}^{k} F_i \tag{83}$$

$$\hat{C}_k = 1 - \frac{F_1}{\sum_{i=1}^{t} iF_i} \tag{84}$$

Similarly, we have

$$\hat{\gamma}_k{}^2 = \max\left\{\frac{S_{\text{obs},k}}{\hat{C}_k}\frac{\sum_{i=1}^{k} i(i-1)F_i}{\left(\sum_{i=1}^{k} iF_i\right)^2} - 1, 0\right\} \tag{85}$$

so that

$$S_{\mathrm{IC}k} = S_{\mathrm{obs}} - S_{\mathrm{obs},k} + \frac{S_{\mathrm{obs},k}}{\hat{C}_k} + \frac{F_1}{\hat{C}_k}\hat{\gamma}_k{}^2 \qquad (86)$$

Estimation on Heterogeneous Samples

Increasing sample size often improves estimator performance through the operation of the laws of large numbers. We get better estimates because the means of many summary statistics become closer to the truth and the variation in these estimates decreases with increasing sample size. Of course, increasing sample sizes increases the cost of performing the census. However, a potentially more serious problem accompanies increases in sample size. In most cases, larger samples take more time or cover larger areas. Inevitably, this means increasing heterogeneity in the sampling process. For example, seasonal variation or habitat heterogeneity will cause detection probabilities to change within a species as sampling progresses. A similar problem occurs in estimating the population size of a single species. Pollock *et al.* (1984) and Lee and Chao (1994) modeled the contribution of heterogeneity in individual capture probabilities to population size estimates. Just as most of the estimators presented so far have been adapted from population size estimators, Nichols *et al.* (1998) applied Pollock's robust estimation method to species richness estimation. These methods incorporate many parameters to characterize the effect of heterogeneity on capture probabilities. Consequently, the estimators are substantially more complicated. Although we may better adjust an estimator to the needs of a given census, the added complexity tends to obscure the overall relationship between sample size and species richness.

Estimators Based on Extrapolation

In order to approximate exhaustive sampling or to elucidate large-scale patterns, many studies extrapolate beyond the extent of a given census. These studies focus less on the sampling process and more on biological processes that operate over scales that are too large to survey in detail. As noted previously, studies that increase the scope of sampling usually encounter a wider range of heterogeneity. Heterogeneity increases due to the responses of individual species to the range of environmental and biological conditions as we increase the study area or the time span of the sample. As we increase sample size, we expect that the correlation in the abundance of arbitrarily selected species should decrease. This decreasing correlation permits us to average over the fine-scale, detailed distribution of the components of S_{obs} and focus on the coarse-scale, global behavior of S_{obs}. When most species in large communities interact weakly with most other species, our experience with the laws of large numbers suggests that there should be a trend relating species richness to area or time span. If the interactions are weak enough relative to sample size, the distribution of fluctuations about the trend may display regular, perhaps even Gaussian, behavior. Our models of species accumulation can then be built with two pieces of information. First, we shall need the average rate of increase in richness as a function of sample size. Second, we need the variability about the mean. The first piece gives us a trend we can extrapolate. The second allows us to fit the model to the data. Most extrapolation methods focus on the trend model and assume a form for the noise that makes curve fitting convenient.

Once we have a suitable model, we can then extrapolate beyond the areas in which we actually sample. We must keep in mind that this extension confounds the contribution of ecological process and sampling effort to diversity. At smaller scales, sampling effort dominates the accumulation of diversity. However, at larger scales, we encounter heterogeneity effects that may scale differently than the effects of sampling effort. Often, we have no general guidelines for selecting one of the several functional forms for species accumulation. Most of the difficulty in applying these models stems from fitting the nonlinear models to data. For brevity, we omit the technical details of curve fitting. We briefly consider a few representative forms here. These are species–area curves, birth chains, and the Michaelis–Menten equation.

To evaluate extrapolation models, we need at least three pieces of information. First, is the trend nondecreasing? Even though it may seem counterintuitive, there are models whose rate of species accumulation increases at small sample sizes (Leitner and Rosenzweig, 1997; Soberón and Llorente, 1993). Second, does the trend have an asymptote? Third, what curvature does the functional form of the model predict? The fitting of the model to data depends on the overall shape of the function. Less obviously, the model transforms the noise that gives rise to the residuals used to evaluate the goodness of fit.

Species–Area Curves

Species–area curves may use samples from disjoint areas or nested areas. This approach has a long history due primarily to its flexibility. There are several functions that may be fit to the data in order to extrapolate to a large area. Chief among these are $\log(S_{\mathrm{obs}})$ vs $\log(N)$ and S_{obs} vs $\log(N)$. Note that these curves have no asymptote. Furthermore, their curvature is often too shallow to give the best fits to the data. The key parameters of interest are the slope (the z-value) and intercept of the plot. The slope in particular has drawn much attention in both theoretical and applied circles. Given the abundance distribution, one may use methods similar to Fisher's approach in deriving α to make theoretical predictions about the magnitude of the slope. Preston (1962) and May (1975) both claim theoretical predictions for z-values built solely on the assumption of the lognormal distribution abundances. This work implied that the species–area curve, at least for lognormal distributions, reflected nothing more than sampling. However, this conclusion rests on the tacit assumption that the abundance distribution be lognormal at all scales. This assumption cannot be achieved without invoking ecological process. Leitner and Rosenzweig (1997) demonstrate that realistic species–area curves must include information about spatial heterogeneity. The curves obtained in this work did have finite asymptotes and variable curvatures. This work does

not, however, provide insights into the selection of log–log or log–linear models.

Birth Chains

Many different functions can exhibit similar asymptotic behavior and curvature. Suppose we select the candidate functions that fit the data best. Functions that give equally good fits to the data can make very different asymptotic predictions. One way to select among these candidate models would be to choose the model that best captures the mechanisms of species accumulation. Soberón and Llorente (1993) use the framework of birth chains for this purpose. Let $\{S(t)\colon t \geq 0\}$ be the number of species observed by time t. Then $S(t)$ is a birth chain if

$$P\{S(t+dt) - S(t) = 1 \,|\, S(t)\} = \lambda(S(t), t)dt \tag{87}$$

$$P\{S(t+dt) - S(t) > 1\} = 0 \tag{88}$$

$$P\{S(0) = 0\} = 1 \tag{89}$$

This set of conditions means that the time between discoveries of new species is exponentially distributed. However, the rate at which new species occur depends on the time spent sampling and the number of species already found. The rate function, $\lambda(S, t)$, captures the details of the mechanism of species discovery. For example, suppose we assume that the rate of discovery declines linearly with species discovered. Then $\lambda(S, t) = a - bS(t)$. Soberón and Llorente (1993) show that

$$E[S(t)] = \frac{a}{b}(1 - e^{-bt}) \tag{90}$$

This method facilitates curve fitting because it allows us to derive expressions for the variance of $S(t)$. Existing methods that lack mechanistic foundations may be justified using this approach. For example, Soberón and Llorente (1993) provide insights into the popular Michaelis–Menten equation using this birth chain model.

Michaelis–Menten

The Michaelis–Menten equation has often been fitted to species accumulation data. It is a widely known function that has a finite asymptote, is nondecreasing, and accommodates a wide range of curvatures. The equation states that

$$S_{\text{obs}}(t) = \frac{at}{1 + bt} \tag{91}$$

Soberón and Llorente (1993) use a birth chain model to suggest a mechanistic basis for the Michaelis–Menten equation. The birth chain model assumes that the number of new species found is an inhomogeneous Poisson process with rate

$$\lambda(S, t) = a + \frac{b^2}{a}\frac{at}{1 + bt} - 2\frac{b}{a}S \tag{92}$$

where j is the number of species found by time t and a and b are constants to be fit to the data. Soberón and Llorente (1993) argue that if the chances of adding a new species improve with effort, to some upper limit, then a form such as Eq. (92) obtains. The key drawback with this and other curve-fitting models is that we do not have a distribution for the changes in S_{obs}. This means that we cannot rationally assign different weights to the data during fitting.

Which Estimation Method to Use?

Many estimators have been put forth, leaving the investigator with the practical question of which estimators are appropriate for use with a given new data set. Two facts may give us confidence a priori in an estimator's applicability to a new data set. As mentioned previously, each estimator makes assumptions about properties of the community being sampled (e.g., variation in capture probabilities across species or samples) and the sampling strategy used. If the community sampled and the sampling procedures used satisfy the assumptions of the estimator, then we can have some confidence in applying that estimator to a new data set.

However, meeting a method's explicit assumptions is not always sufficient. Another difficulty with relying solely on the theoretical basis of estimators is the fact that one or more assumptions are often violated in the sampling of real communities. Thus, theoretical support alone may not justify the use of a particular estimation method.

A second way to have confidence a priori in an estimator is to know its performance on similar communities sampled using similar methods over similar scales. Evaluating estimator performance on data allows testing implicit assumptions. Also, many estimation methods (e.g., extrapolation) have little biology built in: They exist because they seem to work in a few cases, and our confidence in them may come only from verification with relatively complete data sets. One should take care, however, not to conclude too much about an estimator based on its performance on single data sets. Rather, only through comprehensive testing on real data sets from a variety of taxa, locations, and spatial/temporal scales, as well as testing on simulated data, will we gain insight into estimator performance.

Estimator Evaluation

Evaluation on Real Data and Simulated Data

The testing of estimators has used data from real communities and from computer simulation. Each type of data has advantages and drawbacks. Real data take considerable effort to obtain. As a result, replication of real data sets is difficult. Simulated data sets, on the other hand, can be turned out in comparatively large numbers. This allows replication of data and testing of various sampling strategies on the same data.

To test estimators effectively with data, we must have some idea of the true parameters underlying our data (otherwise, how do we know what estimators should predict?). With simulated data, each parameter's true value is known because it is either specified by the investigator or is directly measurable from the simulated community. With real data, the truth may be approximated by testing estimators on smaller sets of

data from communities in which the parameters are reasonably well-known from separate, intensive sampling.

Finally, there are many properties of ecological communities which are likely to influence estimator performance. These include factors such as species richness s, various properties of the species abundance distribution, intraspecific clumping, and interspecific associations. Each of these can vary widely. Simulated data allow us to explore large parts of this parameter space. More important, with simulated data we can cover the range corresponding to anticipated field conditions. This allows us to test and select estimators before taking them into the field. Nevertheless, only real data can tell us which parts of the vast parameter space are relevant. Thus, we cannot ignore either approach: The biological reality of actual data is complemented by the versatility of simulated data.

Some progress has been made in the evaluation of estimation methods with data. Every such study cannot be detailed here, and not enough work has been done to allow generalizations to be made. Chazdon *et al*. (1998) provide one useful example of an approach combining both real and simulated data. They examined the performance of eight estimators using young woody regeneration in six tropical forest sites. Noticing that patchiness (intraspecific clumping) varied across sites, they then resampled the data sets to create simulated data sets with a range of patchiness levels. Testing estimators on these simulated data sets revealed the effect of patchiness on the performance of various estimators.

Estimator Evaluation Criteria

Clearly, one comprehensive study of the performance of all estimators over all possible data is impossible. Evaluation of estimators, then, will necessarily be done by many investigators with different communities and over different ranges of simulated data. To facilitate this cooperative approach, we must identify a common set of evaluation criteria that can be used to measure the performance of an estimator on a data set. These criteria are slightly different from theoretical properties of estimators. Numerous such criteria have been put forth; although they differ in calculation methods, they are often variants of a few common themes, which are listed as follows:

Bias: As defined previously, bias measures deviation from the true value. Common measures of bias include $E[S_{est}] - s$, or deviation from s, and $(E[S_{est}] - s)/s$, or relative deviation from s. In the evaluation of estimators with data, bias can be calculated as a mean across replicate data sets or can be assessed with increasing sample size within a data set.

Variance: As discussed previously, variance measures uncertainty in an estimate. If we have high variance, we can have little confidence that a single observation is a good indication of the mean. Variance may be calculated numerically over replicates, or an analytical estimator of variance may be used. These measures can also be expressed as confidence intervals and may be used in hypothesis testing when comparing estimated richnesses. Note that this measure does not depend on s, which is considered by other criteria.

Sample size independence: This measures the rate of convergence of an estimator's mean to its asymptotic value. An estimator that rapidly approaches its asymptotic value requires less sampling effort to obtain an equally good estimate. Such an estimator is relatively sample size independent. This measure, too, does not depend on s.

Note that the previous three criteria may be calculated over multiple realizations of the same underlying parameters (e.g., the capture probabilities q_i). The following two may only be considered among multiple data sets in which these parameters vary.

Correlation with S: Since s is a fixed parameter, we introduce S as a random variable representing the true richness of any of a number of data sets, between which S and capture probabilities q_i may vary. If the correlation between an estimator S_{est} and true richness S across these data sets is high, S_{est} may be useful in comparing SR between sites or census occasions (Palmer, 1990). Note that correlation with S does not require that an estimator have low bias. Rather, correlation reflects a relative deviation of S_{est} from S that is somewhat constant across data sets.

Robustness: A robust estimator is one for which performance (as measured by any of the previous criteria) changes little across a range of data sets.

It is important to note that the relative importance of each of these criteria depends on the problem of interest. In measuring absolute species richness, minimizing bias may be a priority. In comparisons across time or space, variance and correlation with S become more important, whereas bias may become more tolerable. A final criterion represents a minimum requirement of sorts for estimator performance:

Beating S_{obs}: An estimator should perform better than S_{obs}, the number of species observed. Most estimators should have lower bias than S_{obs} most of the time. However, there are trade-offs. For example, the variance of some estimators may be higher than that of S_{obs}.

Selecting an Estimator

Although it is tempting to think that an estimator may exist which is robust to all census conditions, such an estimator is unlikely given the difficulty of the problem. Therefore, estimators should be chosen based on the context of the problem of interest. To select the estimators that are most likely to perform well on a new data set, one might use the following approach. First, consider the anticipated properties of a data set—based on knowledge of the system's biology—in relation to the modeling assumptions behind estimators. Select those estimators whose assumptions are best met by the expected data. If possible, select or modify the sampling strategy so that assumptions will be better met and/or more estimators may be used. Then, test estimators using data (simulated or previously existing) similar to that anticipated. The previous approach should indicate (i) which estimators might perform best on new data and (ii) whether these estimators perform well enough to meet the needs of the problem at hand. Although evaluating a number of estimators on multiple replicate data sets may seem a daunting task, computer programs (e.g., EstimateS, R. K. Colwell, *http://viceroy.eeb.uconn.edu/estimates*; WS2M, W. R. Turner *et al.*, unpublished) may be used to automate many of these calculations. However, use of such programs should complement, not replace, thoughtful

consideration of modeling assumptions, sampling design, and the biology of the community under study.

See also: Economic Value of Biodiversity, Measurements of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Function Measurement, Terrestrial Communities. Framework for Assessment and Monitoring of Biodiversity. Microbial Biodiversity

References

Burnham KP and Overton WS (1978) Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65: 625–633.

Chao A (1984) Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.* 11: 265–270.

Chao A (1987) Estimating the population size for capture–recapture data with unequal catchability. *Biometrics* 43: 783–791.

Chao A and Lee SM (1992) Estimating the number of classes via sample coverage. *J. Am. Stat. Assoc.* 82: 210–217.

Chao A, Lee SM, and Jeng SL (1992) Estimating population size for capture–recapture data when capture probabilities vary by time and individual animal. *Biometrics* 48: 201–216.

Chao A, Ma MC, and Yang MCK (1993) Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* 80: 193–201.

Chazdon RL, Colwell RK, Denslow JS, and Guariguata MR (1998) Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of northeastern Costa Rica. In: Dallmeier F and Comiskey JA (eds.) *Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies*. Paris: Parthenon.

Colwell RK and Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philos. Trans. R. Soc. London B* 345: 101–118.

Fisher RA, Corbet AS, and Williams CB (1943) The relations between the number of species and the number of individuals in a random sample of an animal population, *J. Anim. Ecol.* 12: 42–58.

Lee SM and Chao A (1994) Estimating population size via sample coverage for closed capture–recapture models. *Biometrics* 50: 88–97.

Leitner WA and Rosenzweig ML (1997) Nested species–area curves and stochastic sampling: A new theory. *Oikos* 79: 503–512.

May RM (1975) Patterns of species abundance and diversity. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*, pp. 81–120. Cambridge, MA: Belknap/Harvard Univ. Press.

Palmer MW (1990) The estimation of species richness by extrapolation. *Ecology* 71: 1195–1198.

Pollock KH (1976) Building models of capture–recapture experiments. *Statistician* 25: 253–260.

Pollock KH, Hines JE, and Nichols JD (1984) The use of auxiliary variables in capture–recapture and removal experiments. *Biometrics* 40: 329–340.

Preston FW (1962) The canonical distribution of commonness and rarity. *Ecology* 43: 185–215.

Smith EP and van Belle G (1984) Nonparametric estimation of species richness. *Biometrics* 40: 119–129.

Soberón J and Llorente J (1993) The use of species accumulation functions for the prediction of species richness. *Conserv. Biol.* 7: 480–488.

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data

Nicholas J Gotelli, University of Vermont, Burlington, VT, USA
Anne Chao, National Tsing Hua University, Hsin-Chu, Taiwan

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biotic similarity A measure of the degree to which two or more samples or assemblages are similar in species composition. Familiar biotic similarity indices include Sørensen's, Jaccard's, Horn's, and Morisita's indices.

Hill numbers A family of diversity measures developed by Mark Hill. Hill numbers quantify diversity in units of equivalent numbers of equally abundant species.

Individual-based (abundance) data A common form of data in biodiversity surveys. The data set consists of a vector of the abundances of different species. This data structure is used when an investigator randomly samples individual organisms in a biodiversity survey.

Nonparametric asymptotic estimators Estimators of total species richness (including Chao1, Chao2, abundance-based coverage estimator (ACE), incidence-based coverage estimator (ICE), and the jackknife) that do not assume a particular form of the species abundance distribution (such as a log-series or log-normal distribution). Instead, these methods use information on the frequency of rare species in a sample to estimate the number of undetected species in an assemblage.

Phylogenetic diversity Adjusted diversity measures that take into account the degree of relatedness among a set of species in an assemblage. Other things being equal, an assemblage of closely related species is less phylogenetically diverse than a set of distantly related species.

Rarefaction A statistical interpolation method of rarefying or thinning a reference sample by drawing random subsets of individuals (or samples) in order to standardize the comparison of biological diversity on the basis of a common number of individuals or samples.

Sample-based (incidence) data A common form of data in biodiversity surveys. The data set consists of a set of sampling units (such as plots, quadrats, traps, and transect lines). The incidence or presence of each species is recorded for each sampling unit.

Species accumulation curve A curve of rising biodiversity in which the x-axis is the number of sampling units (individuals or samples) from an assemblage and the y-axis is the observed species richness. The species accumulation curve rises monotonically to an asymptotic maximum number of species.

Species diversity A measure of diversity that incorporates both the number of species in an assemblage and some measure of their relative abundances. Many species diversity indices can be converted by an algebraic transformation to Hill numbers.

Species richness The total number of species in an assemblage or a sample. Species richness in an assemblage is difficult to estimate reliably from sample data because it is very sensitive to the number of individuals and the number of samples collected. Species richness is a diversity of order 0 (which means it is completely insensitive to species abundances).

Introduction

Measuring Biological Diversity

The notion of biological diversity is pervasive at levels of organization ranging from the expression of heat-shock proteins in a single fruit fly to the production of ecosystem services by a terrestrial ecosystem that is threatened by climate change. How can one quantify diversity in meaningful units across such different levels of organization? This article describes a basic statistical framework for quantifying diversity and making meaningful inferences from samples of diversity data.

In very general terms, a collection of "elements" are considered, each of which can be uniquely assigned to one of several distinct "types" or categories. In community ecology, the elements typically represent the individual organisms, and the types represent the distinct species. These definitions are generic, and typically are modified for different kinds of diversity studies. For example, paleontologists often cannot identify fossils to the species level, so they might study diversity at higher taxonomic levels, such as genera or families. Population geneticists and molecular biologists might be interested in more fine-scale "omics" classifications of biological materials on the basis of unique DNA sequences (genomics), expressed mRNA molecules (transcriptomics), proteins (proteomics), or metabolic products (metabolomics). Ecosystem ecologists might be concerned not with individual molecules, genotypes, or species, but with broad functional groups (producers, predators, and decomposers) or specialized ecological or evolutionary life forms (understory forest herbs and filter-feeding molluscs). However, to keep things simple, this article will refer throughout to "species" as the distinct categories of biological classification.

Although the sampling unit is often thought of as the individual organism, many species, such as clonal plants or colonial invertebrates, do not occur as distinct individuals that can be counted. In other cases, the individual organisms, such as aquatic invertebrate larvae, marine phytoplankton, or soil microbes are so abundant that they cannot be practically counted. In these cases, the elements of biodiversity will

correspond not to individual organisms, but to the sampling units (traps, quadrats, and sighting records) that ecologists use to record the presence or absence of a species.

Species Richness and Traditional Species Diversity Metrics

The number of species in an assemblage is the most basic and natural measure of diversity. Many important theories in community ecology, including island biogeography, intermediate disturbance, keystone and foundational species effects, neutral theory, and metacommunity dynamics make quantitative predictions about species number that can be tested with field observations and experiments in community ecology. From the applied perspective, species richness is the ultimate "score card" in efforts to preserve biodiversity in the face of increasing environmental pressures and climate change resulting from human activity. Species losses can occur from extinction, and species increases can reflect deliberate and accidental introductions or range shifts driven by climate change.

Although species richness is a key metric, it is not the only component of species diversity. Consider two woodlands, each with 20 species of trees. In the first woodland, the 20 species are equally abundant, and each species comprises 5% of the total abundance. In the second woodland, one dominant species comprises 81% of the total abundance, and each of the remaining 19 species contributes only 1% to the total. Although both woodlands contain 20 species, a visitor to the first woodland would encounter most of the different tree species in a brief visit, whereas a visitor to the second woodland might encounter mostly the single dominant species (**Figure 1**).

Thus, a comprehensive measure of species diversity should include components of both species richness and the relative abundances of the species that are present. Such measures are referred to in this article as "traditional" diversity measures. Ecologists have used dozens of different traditional diversity measures, all of which assume (1) individuals within a species are equivalent, (2) all species are "equally different" from one another and receive equal weighting, and (3) diversity is measured in appropriate units (individuals, biomass, and percentage cover are most commonly used).

Phylogenetic, Taxonomic, and Functional Diversity

As noted in the previous section the first assumption of traditional diversity metrics (individuals within a species are equivalent) can be relaxed by changing the operational definition of "species" to other categories of interest. The second assumption (all species are equally different from one another) ignores aspects of phylogenetic or functional diversity, but can also be incorporated through a generalization of traditional diversity metrics.

For example, consider two woodlands with identical tree species richness and evenness but with no shared species. The species in the first woodland are all closely related oaks in the same genus (*Quercus*). The species in the second woodland are a diverse mix of oaks (*Quercus*) and maples (*Acer*), as well as more distantly related pines (*Pinus*). Traditional species diversity metrics would be identical for both woodlands, but it is intuitive that the second woodland is more diverse (see **Figure 2** for another example).

The concept of traditional diversity can therefore be extended to consider differences among species. All else being equal, an assemblage of phylogenetically or functionally divergent species is more diverse than an assemblage of closely related or functionally similar species. Differences among species can be based directly on their evolutionary histories, either in the form of taxonomic classification (referred to as taxonomic diversity) or phylogeny (referred to as phylogenetic

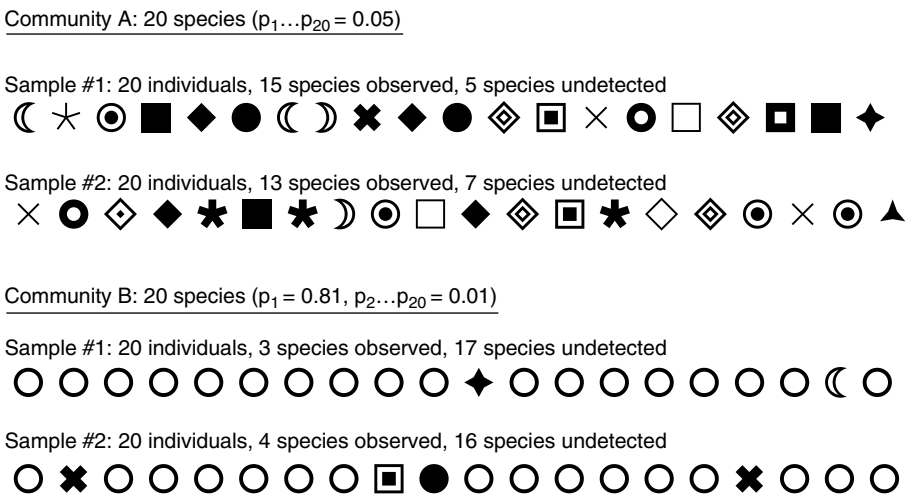


Figure 1 Species richness sampling in a hypothetical walk through the woods. Each different symbol represents one of 20 distinct species, and each row contains 20 characters, representing the first 20 individual trees that might be encountered in a random sample. Community A is maximally even, with each of the 20 species comprising 5% of the total abundance. In this assemblage, the two samples of 20 individual trees yielded 15 and 13 species, respectively. Community B is highly uneven, with one species (the open circle) representing 81% of the total abundance, and the remaining 19 species contributing only 1% each. In this assemblage, the two samples of 20 individual trees yielded only three and four species, respectively.

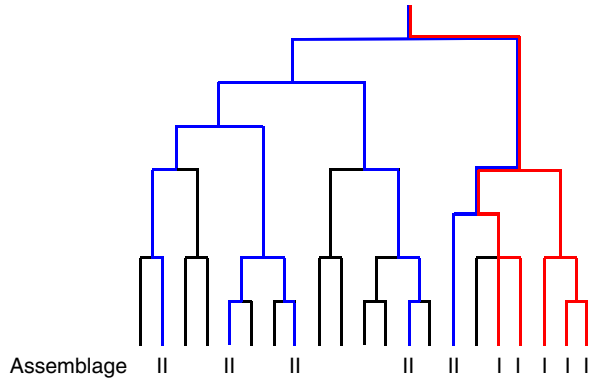


Figure 2 Phylogenetic diversity in species composition. The branching diagram is a hypothetical phylogenetic tree. The ancestor of the entire assemblage is the "root" at the top, with time progressing toward the branch tips at the bottom. Each node (branching point) represents a speciation or divergence event, and the 21 branch tips illustrate the 21 extant species. Extinct species or lineages are not illustrated. The five species in Assemblage I represent an assemblage of five closely related species (they all share a quite recent common ancestor). The five species in Assemblage II represent an assemblage of five distantly related species (they all share a much older common ancestor). All other things being equal, the community of distantly related species would be considered more phylogenetically diverse than the community of closely related species.

diversity (PD)) or indirectly, based on their function (referred to as functional diversity). These metrics relax the second assumption discussed in the section Species Richness and Traditional Species Diversity Metrics (all species are "equally different" from one another) by weighting each species by a measure of its taxonomic classification, phylogeny, or function.

Biotic Similarity

These concepts of species diversity apply to metrics that are used to quantify the diversity of single assemblages. However, the concept of diversity can also be applied to the comparison of multiple assemblages. Suppose again that a person visits two woodlands, both of which have 10 trees species, each species contributing 10% to the abundance of individual trees within the woodland. Thus, in terms of species richness and species diversity, the two woodlands are identical. However, the two woodlands may differ in their species composition. At one extreme, they may have no species in common, so they are biologically distinct, in spite of having equal species richness and species diversity. At the other extreme, if the list of tree species in the two woodlands is the same, they are identical in all aspects of diversity (including taxonomic, phylogenetic, and functional diversity). More typically, the two woodlands might have a certain number of species found in both woodlands and a certain number that are found in only one.

Biotic similarity quantifies the extent to which two or more sites are similar in their species composition and relative abundance distribution. The concept of biotic similarity is important at large spatial scales for the designation of biogeographic provinces that harbor distinctive species assemblages with both endemic and shared elements. Biotic similarity is also a key concept underlying the measurement of beta diversity, the turnover in species composition among a set of sites. In an applied context, biotic similarity indices can quantify the extent to which distinct biotas in different regions have become homogenized through losses of endemic species and the introduction and spread of nonnative species. Differences among species in evolutionary histories and functional trait values can also be incorporated in similarity measures.

Bias in the Estimation of Diversity

The true species richness and relative abundances in an assemblage are unknown in most applications. Thus species richness, species diversity, and biotic similarity must be estimated from samples taken from the assemblage. If the sample relative abundances are used directly in the formulas for traditional diversity and similarity measures, the maximum likelihood estimator (MLE) of the true diversity or similarity measure is obtained. However, the MLEs of most species diversity measures are biased when sample sizes are small. When sample size is not sufficiently large to observe all species, the unobserved species are undersampled, and – as a consequence – the relative abundance of observed species, on average, is overestimated.

Because biotic diversity at all levels of organization is often high, and biodiversity sampling is usually labor intensive, these biases are usually substantial. Even the simplest comparison of species richness between two samples is complicated unless the number of individuals is identical in the two samples (which it never is) or the two samples represent the same degree of coverage (completeness) in sampling. Ignoring the sampling effects may obscure the influence of overall abundance or sampling intensity on species richness. Attempts to adjust for sampling differences by algebraic rescaling (such as dividing S by n or by sampling effort) lead to serious distortions and gross overestimates of species richness for small samples. Thus, an important general objective in diversity analysis is to reduce the undersampling bias and to adjust for the effect of undersampled species on the estimation of diversity and similarity measures. Because sampling variation is an inevitable component of biodiversity studies, it is equally important to assess the variance (or standard error) of an estimator and provide a confidence interval that will reflect sampling uncertainty.

The Organization of Biodiversity Sampling Data

This article introduces a common set of notation for describing biodiversity data (Colwell *et al.*, 2012). Consider an assemblage consisting of N^* total individuals, each belonging to one of S distinct species. Species i has N_i individuals, so that $\sum_{i=1}^{S} N_i = N^*$. The relative frequency p_i of species i is N_i/N^*, so that $\sum_{i=1}^{S} p_i = 1$. Note here that N^*, S, N_i, and p_i represent the "true" underlying abundance, species richness, and relative frequencies of species. These quantities are unknowns, but can be estimated, and one can make statistical inferences by taking

random samples of data from such an assemblage. This article distinguishes between two sampling structures.

Individual-Based (Abundance) Data

The *reference sample* is a collection of n individuals, drawn at random from the assemblage with N^* total individuals. In the reference sample, a total of S_{obs} species are observed, with X_i individuals observed for species i, so that $\sum_{i=1}^{S} X_i = n$ (only species with $X_i > 0$ contributes to the sum). Thus, the data consist of a single vector of length S, whose elements are the observed abundances of the individual species X_i. In this vector, there are S_{obs} nonzero elements.

The *abundance frequency count* f_k is defined as the number of species each represented by exactly k individuals in the reference sample. Thus, f_1 is the number of species represented by exactly one individual ("singletons") in the reference sample, and f_2 is the number of species represented by exactly two individuals ("doubletons"). In this terminology, f_0 is the number of *undetected species*: species that are present in the assemblage of N^* individuals and S species, but were not detected in the reference sample of n individuals and S_{obs} species. Therefore, $S_{obs} + f_0 = S$.

Sample-Based (Incidence) Data

The reference sample for incidence data consists of a set of R replicate *sampling units* (traps, plots, quadrats, search routes, etc.). In a typical study, these sampling units are deployed randomly in space within the area encompassing the assemblage. However, in a temporal study of diversity, the R sampling units would be deployed in one place at different independent points in time (such as an annual breeding bird census at a single site). Within each sampling unit, the presence or absence of each species is recorded, but abundances or counts of the species present are not needed. The data are organized as a species-by-sampling-unit *incidence matrix*, in which there are $i = 1$ to S rows (species), $j = 1$ to R columns (sampling units), and the matrix entry $W_{ij} = 1$ if species i is detected in sampling unit j, and $W_{ij} = 0$ otherwise. If sampling is replicated in both time and space, the data would be organized as a three-dimensional matrix (species × sites × times). However, most biodiversity data sets are two dimensional, with either spatial or temporal replication, but not both.

The row sum of the incidence matrix $Y_i = \sum_{j=1}^{R} W_{ij}$ denotes the incidence-based frequency of species i for $i = 1$ to S. Y_i is analogous to X_i in the individual-based abundance vector. Species present in the assemblage but not detected in any sampling unit have $Y_i = 0$. The total number of species observed in the reference sample is S_{obs} (only species with $Y_i > 0$ contribute to S_{obs}).

The *incidence frequency count* Q_k is the number of species each represented exactly $Y_i = k$ times in the incidence matrix sample, $0 \le k \le R$. For the incidence matrix, $\sum_{k=1}^{R} kQ_k = \sum_{i=1}^{S} Y_i$ and $S_{obs} = \sum_{k=1}^{R} Q_k$. Thus, Q_1 represents the number of "unique" species (those that are each detected in only one sample) and Q_2 represents the number of "duplicate" species (those that are each detected in exactly two samples). The zero frequency Q_0 denotes the number of species among the S species in the assemblage that were not detected in any of the R sampling units.

Species Richness Estimation

A simple count of the number of species in a sample is usually a biased underestimate of the true number of species, simply because increasing the sampling effort (through counting more individuals, examining more sampling units, or sampling a larger area) inevitably will increase the number of species observed. The effect is best illustrated in a *species accumulation curve*, in which the x-axis is the number of individuals sampled or sampling units examined and the y-axis is the number of species observed (**Figure 3**). The first individual sampled always yields one species, so the origin of an abundance-based species accumulation curve is the point [1,1]. If the next individual sampled is the same species, the curve stays flat with a slope of zero. If the next individual sampled is a different species, the curve rises to two species, with an initial slope of 1.0. Samples from the real world fall between these two idealized extremes, and the slope of the curve measured at any abundance level is the probability that the next individual sampled represents a previously unsampled species. The curve is steepest in the early part of the collecting, as the common species in the assemblage are detected relatively quickly. The curve continues to rise as

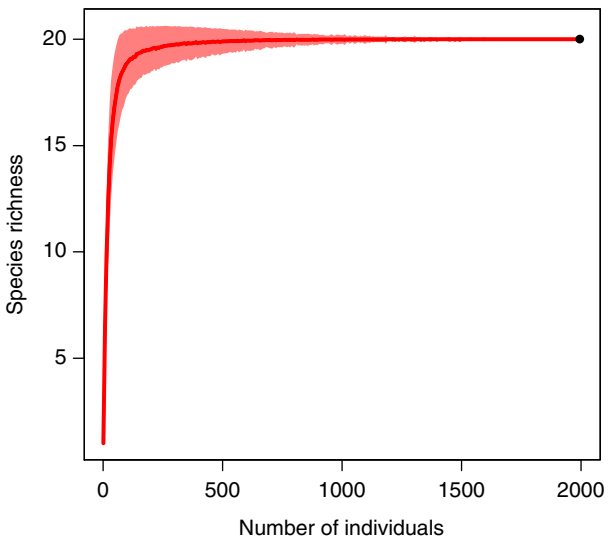


Figure 3 Species accumulation curve. The curve was generated by assuming an assemblage of 20 species whose relative abundances were created from a broken stick distribution (Tokeshi, 1999). The x-axis is the number of individuals sampled and the y-axis is the number of species observed. The species accumulation curve is the smooth red line, which represents the average of 1000 random draws, sampling with replacement, at each level of abundance. The shaded envelope represents a symmetric 95% bootstrap confidence interval, calculated from the estimated variance of the random draws. The shape of this species accumulation curve is typical: it rises rapidly at first as the common species are initially encountered, and then continues to rise very slowly, as much more sampling is needed to encounter all of the rare species. For random samples of 500 or more individuals, it is almost always the case that all 20 species are encountered.

more individuals are sampled, but the slope becomes shallower because progressively more sampling is required to detect the rare species. As long as the sampling area is sufficiently homogeneous, all of the species will eventually be sampled and the curve will flatten out at an asymptote that represents the true species richness for the assemblage. For incidence data, a similar accumulation curve can be drawn in which the x-axis represents the number of sampling units and the y-axis is the number of species recorded.

Interpolating Species Richness with Rarefaction

A single empirical sample of individuals or a pooled set of sampling units represents one point on the species accumulation curve, but the investigator has no way of directly determining where on the curve this point lies. To compare the richness of two different samples, they must be standardized to a common number of individuals, for abundance samples (Sanders, 1968; Gotelli and Colwell, 2001, 2011). Rarefaction represents an interpolation of a biodiversity sample to a smaller number of individuals for purposes of comparison among samples. Typically, the abundance of the larger sample is rarefied to the total abundance of the smaller sample to determine if species richness (or any other biodiversity index) differs for a common number of individuals (**Figure 4**). For incidence data, rarefaction interpolates between the reference sample and a smaller number of sampling units.

Let $S_{\text{ind}}(m)$ represent the expected number of species in a random sample of m individuals from the reference sample of

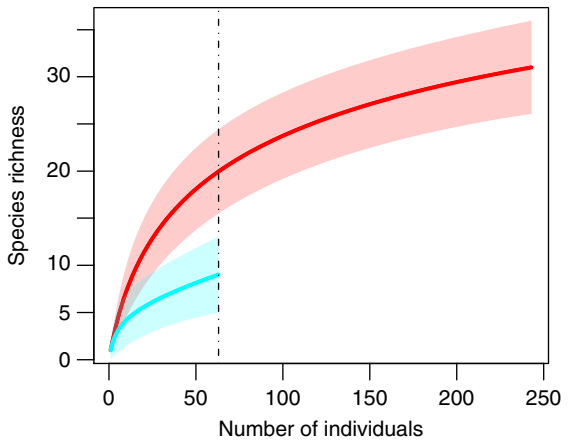


Figure 4 Standardized comparison of species richness for two individual-based rarefaction curves. The data represent summary counts of carabid beetles that were pitfall-trapped from a set of young pine plantations (<20 years old; upper curve) and a set of old pine plantations (20–60 years old; lower curve). The solid lines are the rarefaction curves, calculated from eqn [2], and the shaded polygons are the 95% confidence intervals, calculated from the unconditional variance eqn [5]. The young plantation samples contained 243 individuals representing 31 species, and the old plantation samples contained 63 individuals representing nine species. The dashed and dotted vertical line illustrates a species richness comparison standardized to 63 individuals, which was the observed abundance in the smaller of the two data sets. Data from Niemelä J, Haila Y, Halme E, *et al.* (1988) The distribution of carabid beetles in fragments of old coniferous taiga and adjacent managed forest. *Annales Zoologici Fennici* 25: 107–199.

n individuals ($m<n$). If the true probabilities ($p_1, p_2, \ldots, p_s$) of each of the S species in the assemblage were known, and species frequencies ($X_1, X_2, \ldots, X_S$) follow a multinomial distribution for which the total of all frequencies is n, and cell probabilities ($p_1, p_2, \ldots, p_S$), then

$$S_{\text{ind}}(m) = S - \sum_{i=1}^{S} (1 - p_i)^m \qquad [1]$$

However, the true p_i values are unknown, and there is only the reference sample with observed species abundances X_i. An unbiased estimator for $S_{\text{ind}}(m)$ (Hurlbert, 1971) is

$$\tilde{S}_{\text{ind}}(m) = S_{\text{obs}} - \sum_{X_i>0} \left[\binom{n - X_i}{m} \Big/ \binom{n}{m} \right] \qquad [2]$$

For incidence-based data, the corresponding equation (Shinozaki, 1963) is

$$\tilde{S}_{\text{sample}}(r) = S_{\text{obs}} - \sum_{Y_i>0} \left[\binom{R - Y_i}{r} \Big/ \binom{R}{r} \right] \qquad [3]$$

where $r<R$ is the number of sampling units in the rarefied reference sample. The statistical model for rarefaction is sampling without replacement from the reference sample.

If the area of each of the sample plots has been measured, species richness can also be interpolated from a *Coleman curve*, in which the expected species richness on an island (or sample plot) of area a is based on a Poisson model and is a function of the total area A of the archipelago (or the summed areas of all the sample plots) (Coleman *et al.*, 1982):

$$\tilde{S}_{\text{area}}(a) = S_{\text{obs}} - \sum_{X_i>0} \left(1 - \frac{a}{A}\right)^{X_i} \qquad [4]$$

Although the variance from the hypergeometric distribution has traditionally been used to calculate a confidence interval for a rarefaction curve, this variance is conditional on the observed sample. It therefore has the undesirable property of converging to zero when the abundance level reaches the reference sample size. More realistically, the observed sample is itself drawn from a much larger assemblage, so the confidence intervals generally should widen as the reference sample size is reached. This unconditional variance for abundance data is calculated as follows (Colwell *et al.*, 2012):

$$\sigma^2_{\text{ind}}(m) = \sum_{k=1}^{n} (1 - \alpha_{km})^2 f_k - [\tilde{S}_{\text{ind}}(m)]^2 / S_{\text{est}} \qquad [5]$$

where S_{est} denotes an estimated species richness (such as Chao1, described in the section Species Richness Estimation) and $\alpha_{km} = \binom{n-k}{m} \Big/ \binom{n}{m}$ for $k \le n - m$, $\alpha_{km} = 0$ otherwise. The corresponding unconditional variance for incidence data is (Colwell *et al.*, 2004)

$$\sigma^2_{\text{sample}}(r) = \sum_{k=1}^{R} (1 - \beta_{kr})^2 Q_k - [\tilde{S}_{\text{sample}}(r)]^2 / S_{\text{est}} \qquad [6]$$

where S_{est} denotes a sample-based estimated species richness (such as Chao2, described in the section Species Richness

Estimation) and $\beta_{kr} = \binom{R-k}{r} \Big/ \binom{R}{r}$ for $k \leq R-r$, $\beta_{kr}=0$ otherwise. These variances allow for the calculation of 95% confidence intervals on expected species richness for any abundance level smaller than the observed sample (**Figure 4**).

Because all rarefaction curves converge at small sample sizes toward the point [1,1] (for abundance data) or a small number of species (for incidence data), sufficient sampling is necessary for valid comparisons of curves. Although there are no theoretical guidelines, empirical examples suggest that samples of at least 20–50 individuals per sample (and preferably many more) are necessary for meaningful comparisons of abundance-based rarefaction curves.

Rarefaction curves also require comparable sampling methods (forest samples collected from pitfall traps cannot be validly compared to prairie samples collected from baits), well-defined assemblages of discrete countable individuals (for abundance-based methods), random spatial arrangement of individuals, and random, independent sampling of individuals (or larger sampling units for incidence-based methods). If the spatial distribution of individuals is intraspecifically clumped in space, abundance-based rarefaction will overestimate species richness, but this problem can be effectively countered by increasing the spatial grain of sampling or using incidence-based methods. Perhaps the chief disadvantage of rarefaction is that point comparisons force an investigator to rarefy all samples down to the smallest sample size in the data set, so sufficient sampling is important. However, calculation and comparison of complete rarefaction curves and their extrapolation, with unconditional variances, help to overcome this problem (Colwell *et al.*, 2012).

Nonparametric Asymptotic Species Richness Estimators

Whereas rarefaction is a method for interpolating species diversity data, asymptotic richness estimators are methods for extrapolating species diversity out to the (presumed) asymptote, beyond which additional sampling will not yield any new species. Three strategies have been used to try to estimate the asymptote of the species accumulation curve. *Parametric curve fitting* uses the shape of the species accumulation curve in its early phase to try and predict the asymptote. Asymptotic functions, such as the negative exponential distribution, the Weibull distribution, the logistic equation, and the Michaelis–Menten equation, are fit (usually with nonlinear regression methods) to the species accumulation data, and the asymptote can be estimated as one of the parameters of this kind of model. The chief problem is that this does not work well in comparisons with empirical or simulated data sets, mainly because it does not directly use information on the frequency of common and rare species, but simply tries to forecast the shape of the rising curve. Several different functional forms may fit the same data set equally well, but yield drastically different estimates of the asymptote. Because curve fitting is not based on a statistical sampling model, the variance of the resulting asymptote cannot be evaluated without further assumptions, and theoretical difficulties arise for model selection.

A second strategy is to use the abundance or incidence frequency counts (f_k or Q_k) and fit them to a *species abundance distribution*, such as the log-series or the log-normal distribution. The area under such a fitted curve is an estimate of the total number of species present in the assemblage. The chief weakness of these methods is that simulations show that they work well only when the correct form of the species abundance distribution is already known, but this is never the case for empirical data. It is often not clear that existing statistical models fit empirical data sets very well, which often depart from expected values in the frequencies of the rare species. Moreover, there is no guarantee that two different assemblages follow the same kind of distribution, which complicates the comparison of curves.

The most successful methods so far have been *nonparametric estimators* (Colwell and Coddington, 1994), which use the rare frequency counts to estimate the frequency of the missing species (f_0 or Q_0). For incidence data, these estimators are similar to mark-recapture models that are used in demography to estimate the total population size and are based on statistical theorems developed by Alan Turing and I.J. Good from cryptographic analysis of Wehrmacht coding machines during World War II. The basic concept of their theorem is that abundant species – which are certain to be detected in samples – contain almost no information about the undetected species, whereas rare species – which are likely to be either undetected or infrequently detected – contain almost all the information about the undetected species.

If there are many undetectable or "invisible" species in a hyperdiverse assemblage, it will be impossible to obtain a good estimate of species richness. Therefore, an accurate lower bound for species richness is often of more practical use than an imprecise point estimate. Based on the concept that rare species carry the most information about the number of undetected species, the Chao1 estimator uses only the numbers of singletons and doubletons (and the observed richness) to obtain the following lower bound for the expected asymptotic species richness (Chao, 1984):

$$\hat{S}_{\text{Chao1}} = \begin{cases} S_{\text{obs}} + f_1^2/(2f_2) & \text{if } f_2 > 0 \\ S_{\text{obs}} + f_1(f_1-1)/2 & \text{if } f_2 = 0 \end{cases} \quad [7]$$

with an associated variance estimator of (if $f_2>0$)

$$\hat{var}(\hat{S}_{\text{Chao1}}) = f_2\left[\frac{1}{2}\left(\frac{f_1}{f_2}\right)^2 + \left(\frac{f_1}{f_2}\right)^3 + \frac{1}{4}\left(\frac{f_1}{f_2}\right)^4\right] \quad [8]$$

For incidence data, Chao2 is the corresponding estimator for species richness. It incorporates a sample-size correction factor $(R-1)/R$ (Chao, 1987):

$$\hat{S}_{\text{Chao2}} = \begin{cases} S_{\text{obs}} + [(R-1)/R]Q_1^2/(2Q_2) & \text{if } Q_2 > 0 \\ S_{\text{obs}} + [(R-1)/R]Q_1(Q_1-1)/2 & \text{if } Q_2 = 0 \end{cases} \quad [9]$$

with a variance estimator of (if $Q_2>0$)

$$\hat{var}(\hat{S}_{\text{Chao2}}) = Q_2\left[\frac{A}{2}\left(\frac{Q_1}{Q_2}\right)^2 + A^2\left(\frac{Q_1}{Q_2}\right)^3 + \frac{1}{4}A^2\left(\frac{Q_1}{Q_2}\right)^4\right] \quad [10]$$

where $A=(R-1)/R$. When $f_2=0$ in the Chao1 estimator or $Q_2=0$ in the Chao 2 estimator, the variance formulas in

eqns [9] and [10] need modification; the modified variances are available in Chao and Shen (2010).

The Chao1 estimator may be very useful for data sets in which it is too time consuming to count the frequencies of all abundance classes, but it is relatively easy to count just the number of singleton and doubleton species. Chao1 and Chao2 are intuitive and very easy to calculate, and often perform just as well as more complex asymptotic estimators.

A more general approach is to use information on the frequency of other rare species, not just singletons and doubletons. A cut-off value κ denotes frequencies of rare species (frequency $\leq \kappa$) and abundant species (frequency $> \kappa$). The cut-off $\kappa = 10$ works well with many empirical data sets.

Let the total number of observed species in the abundant species group be $S_{abun} = \sum_{i>\kappa} f_i$ and the number of observed species in the rare species group be $S_{rare} = \sum_{i=1}^{\kappa} f_i$. Define $n_{rare} = \sum_{i=1}^{\kappa} i f_i$ and the coverage estimate $\hat{C}_{rare} = 1 - f_1/n_{rare}$. Coverage is the estimated proportion of the total number of N^* individuals in the assemblage that is represented by the species recorded in the sample. It is a reliable measure of the degree of sample completeness. The *Abundance-based Coverage Estimator* (*ACE*) is (Chao, 2005)

$$\hat{S}_{ACE} = S_{abun} + \frac{S_{rare}}{\hat{C}_{rare}} + \frac{f_1}{\hat{C}_{rare}}\hat{\gamma}^2_{rare} \qquad [11]$$

where $\hat{\gamma}^2_{rare}$ is the square of the estimated coefficient of variation of the species relative abundances:

$$\hat{\gamma}^2_{rare} = \max\left\{\frac{S_{rare}}{\hat{C}_{rare}}\frac{\sum_{i=1}^{\kappa} i(i-1)f_i}{(\sum_{i=1}^{\kappa} i f_i)(\sum_{i=1}^{\kappa} i f_i - 1)} - 1, 0\right\} \qquad [12]$$

An approximate variance for the ACE can be obtained using a standard asymptotic approach.

For incidence data, there is a corresponding *Incidence-based Coverage Estimator* (*ICE*). As with ACE, first a cut-off point κ is selected that partitions the data into an infrequent species group (incidence frequency not larger than κ) and a frequent species group (incidence frequency larger than κ). The cut-off $\kappa = 10$ is recommended. Denote the number of species in the frequent group by $S_{freq} = \sum_{i>\kappa} Q_i$ and the number of species in the infrequent group by $S_{infreq} = \sum_{i=1}^{\kappa} Q_i$. The estimated sample coverage for the infrequent group is $\hat{C}_{infreq} = 1 - Q_1/\sum_{i=1}^{\kappa} i Q_i$. Let the number of sampling units that include at least one infrequent species be R_{infreq}. Then ICE is expressed as

$$\hat{S}_{ICE} = S_{freq} + \frac{S_{infreq}}{\hat{C}_{infreq}} + \frac{Q_1}{\hat{C}_{infreq}}\hat{\gamma}^2_{infreq} \qquad [13]$$

where $\hat{\gamma}^2_{infreq}$ is the squared estimate of the coefficient of variation of the species relative incidences:

$$\hat{\gamma}^2_{infreq} = \max\left\{\frac{S_{infreq}}{\hat{C}_{infreq}}\frac{R_{infreq}}{(R_{infreq}-1)} \times \frac{\sum_{i=1}^{\kappa} i(i-1)Q_i}{(\sum_{i=1}^{\kappa} i Q_i)(\sum_{i=1}^{\kappa} i Q_i - 1)} - 1, 0\right\} \qquad [14]$$

In addition to Chao1, Chao2, ACE, and ICE, the *jackknife* method provides another class of nonparametric estimators of asymptotic species richness. Jackknife techniques were developed as a general method to reduce the bias of a biased estimator. Here the biased estimator is the number of species observed in the sample. The basic idea with the jth-order jackknife method is to consider subdata by successively deleting j individuals from the data. The first-order jackknife turns out to be

$$\hat{S}_{jk1} = S_{obs} + \frac{n-1}{n} f_1 \approx S_{obs} + f_1 \qquad [15a]$$

That is, only the number of singletons is used to estimate the number of unseen species. The second-order jackknife estimator, which uses singletons and doubletons, has the following form:

$$\hat{S}_{jk2} = S_{obs} + \frac{2n-3}{n} f_1 - \frac{(n-2)^2}{n(n-1)} f_2 \approx S_{obs} + 2f_1 - f_2 \qquad [15b]$$

Higher-order jackknife estimators are available, although they give increasingly less weight to the more common species frequencies.

For replicated incidence data, the first-order jackknife for R samples is

$$\hat{S}_{jk1} = S_{obs} + \frac{R-1}{R} Q_1 \qquad [16a]$$

and the second-order jackknife is

$$\hat{S}_{jk2} = S_{obs} + \frac{2R-3}{R} Q_1 - \frac{(R-2)^2}{R(R-1)} Q_2 \qquad [16b]$$

All jackknife estimators can be expressed as linear combinations of frequency counts, and thus approximate variances and confidence intervals can be directly obtained.

Extrapolating Species Richness

Based on a reference sample of n individuals, the extrapolation or prediction problem is to estimate the expected number of species $S_{ind}(n+m^*)$ in an augmented sample of $n+m^*$ individuals from the assemblage ($m^* > 0$). Under a simple multinomial model, Shen *et al.* (2003) derived the following useful predictor with an asymptotic variance:

$$\tilde{S}_{ind}(n+m^*) = S_{obs} + \hat{f}_0\left[1 - \left(1 - \frac{f_1}{n\hat{f}_0}\right)^{m^*}\right] \approx S_{obs} + \hat{f}_0\left[1 - \exp\left(-\frac{m^* f_1}{n\hat{f}_0}\right)\right] \qquad [17]$$

where $\hat{f}_0$ is an estimator for f_0 (the number of undetected species). It is suggested that $\hat{f}_0$ can be obtained by using either the Chao1 estimator ($\hat{f}_0 = \hat{S}_{Chao1} - S_{obs}$) or the ACE estimator ($\hat{f}_0 = \hat{S}_{ACE} - S_{obs}$).

The corresponding extrapolation formula and its asymptotic variance for the Coleman area-based Poisson sampling model were developed by Chao and Shen (2004). An estimator for the expected number of species $S_{area}(A + a^*)$ in an

augmented area $A+a^*(a^*>0)$ based on a reference sample of area A is

$$\tilde{S}_{\text{area}}(A+a^*) = S_{\text{obs}} + \hat{f}_0\left[1-\exp\left(-\frac{a^*}{A}\frac{f_1}{\hat{f}_0}\right)\right] \quad [18]$$

where $\hat{f}_0$ is the same as in the individual-based model.

For sample-based data with R sampling units comprising the reference sample, Chao *et al.* (2009, Appendix A) developed a Bernoulli-product model and derived the following estimator for the expected number of species $S_{\text{sample}}(R+r^*)$ in an augmented set of $R+r^*$ sampling units $(r^*>0)$ from the assemblage:

$$\tilde{S}_{\text{sample}}(R+r^*) = S_{\text{obs}} + \hat{Q}_0\left[1-\left(1-\frac{Q_1}{Q_1+R\hat{Q}_0}\right)^{r^*}\right]$$
$$\approx S_{\text{obs}} + \hat{Q}_0\left[1-\exp\left(\frac{-r^*Q_1}{Q_1+R\hat{Q}_0}\right)\right] \quad [19]$$

Here $\hat{Q}_0$, which is an estimator for Q_0, can be obtained from either the Chao2 estimator $(\hat{Q}_0=\hat{S}_{\text{Chao2}}-S_{\text{obs}})$ or the ICE $(\hat{Q}_0=\hat{S}_{\text{ICE}}-S_{\text{obs}})$.

For each of the these models, Colwell *et al.* (2012) linked the interpolation (rarefaction) curve and the corresponding extrapolation (prediction) curve to yield a single smooth curve meeting at the reference sample (**Figure 5**). They also derived 95% (unconditional) confidence intervals for the interpolated and extrapolated richness estimates. Thus, rigorous statistical comparison can be performed not only for rarefaction but also for extrapolated richness values. This link helps to avoid the problem of discarding data and information from larger samples that is necessary for comparisons using the traditional rarefaction method. However, the extrapolations become highly uncertain if they are extended beyond approximately double the reference sample size. For both individual- and sample-based data, the additional sample size needed, beyond the reference sample, to attain the estimated asymptotic species richness, or to detect a specified proportion of asymptotic richness, is provided in Chao *et al.* (2009) and Colwell *et al.* (2012).

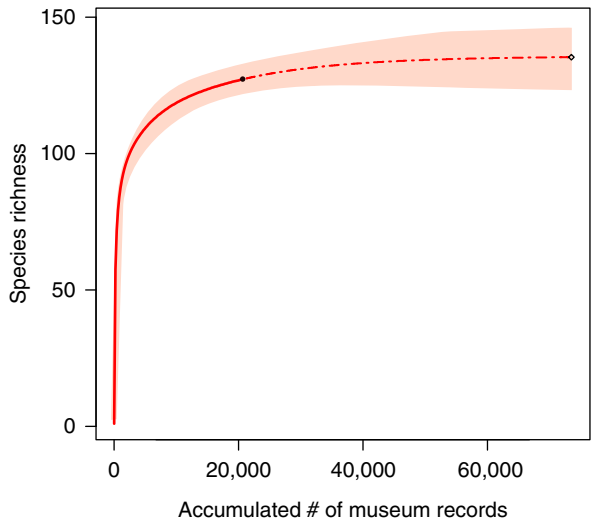


Figure 5 A smoothed rarefaction and extrapolation curve. The *x*-axis is the number of individual, geo-referenced, dated ant specimens in New England, and the *y*-axis is the observed number of species. The total collection (the reference sample, filled circle) included 127 species and 20,225 individual records. The solid curve is the rarefaction curve interpolated from the reference sample. The dashed curve is the extrapolation, which extends to a minimum asymptotic estimator (Chao1) of ~135 species (open diamond). This number accords well with an independent estimate of an additional eight species that occur in suitable habitat in New York and Quebec. These eight species are likely to occur in New England, but so far they have not been collected. However, the extrapolation to reach the Chao1 estimator extends to over 70,000 museum records, and the confidence interval (shaded polygon) is therefore fairly broad. The data set was compiled from museum records and private collections of ants sampled throughout the New England states of the USA (RI, CT, MA, VT, NH, and ME) between 1900 and 2011. Data modified from Ellison AM, Gotelli NJ, Farnsworth EJ, and Alpert GD (2012) *A Field Guide to the Ants of New England*. New Haven, CT: Yale University Press.

Species Diversity

Species Diversity Metrics

Although species richness is the most popular and intuitive measure for characterizing diversity, the section Species Richness Estimation emphasizes that it is a very difficult parameter to estimate reliably from small samples, especially for hyperdiverse assemblages with many rare species. Species richness also does not measure the evenness of the species abundance distribution. Over the span of many decades, ecologists have proposed a plethora of diversity measures that incorporate both species richness and evenness, using both parametric and nonparametric approaches (Magurran, 2004).

For example, parametric approaches assume a particular species abundance distribution (such as the log-normal or gamma) or a species rank abundance distribution (such as the negative binomial or log series), and then estimate parameters from the distribution model that quantify the heterogeneity among species in their relative frequencies. However, as with the estimation of asymptotic richness, these methods often do not perform well unless the "true" species abundance distribution is known, which is never the case (Colwell and Coddington, 1994; Chao, 2005).

Nonparametric methods make no assumptions about the mathematical form of the underlying species abundance distribution, and they have been widely used not only in ecology but also in information science, economics, genetics, and linguistics (see Jost, 2007; Jost *et al.*, 2011; Chao and Jost, in press; Tuomisto, this volume for reviews). The most popular of these measures is the *Shannon entropy*,

$$H_{\text{Sh}} = -\sum_{i=1}^{S} p_i \log p_i \quad [20]$$

where S is the number of species in the assemblage and the ith species relative abundance is p_i. Shannon entropy quantifies the uncertainty in the species identity of a randomly chosen individual in the assemblage. Another measure that has been widely used in economics and genetics, as well as in ecology, is the Gini–Simpson index,

$$H_{\text{GS}} = 1-\sum_{i=1}^{S} p_i^2 \quad [21]$$

which measures the probability that two randomly chosen individuals (selected with replacement) belong to two different species. The measure $1 - H_{GS} = \sum_{i=1}^{S} p_i^2$ is referred to as the Simpson index. With an adjustment for N^*, the total number of individuals in the assemblage, the Gini–Simpson index is closely related to the ecological index *PIE* (Hurlbert, 1971), the probability of an interspecific encounter:

$$\text{PIE} = [N^*/(N^* - 1)]H_{GS} \quad [22]$$

which measures the probability that two randomly chosen individuals (selected *without* replacement) belong to two different species. Both PIE and the Gini–Simpson index have a straightforward interpretation as a probability. When PIE is applied to species abundance data, it is equivalent to the slope of the individual-based rarefaction curve measured at its base.

However, the units of the Gini–Simpson index and *PIE* are probabilities that are bounded between 0 and 1, and the units of Shannon entropy are logarithmic units of information. These popular complexity measures do not behave in the same intuitive way as species richness (Jost, 2007).

The ecologist MacArthur (1965) was the first to show that Shannon entropy (when computed using natural logarithms) can be transformed to its exponential $\exp(H_{Sh})$, and the Gini–Simpson index can be transformed to $1/(1 - H_{GS}) = 1/\sum_{i=1}^{S} p_i^2$, yielding two new indices that measure diversity in units of species richness. In particular, these transformed indices measure diversity in units of "effective number of species" – the equivalent number of equally abundant species that would be needed to give the same value of the diversity measure. When all species are equally abundant, the effective number of species is equal to the richness of the assemblage.

These converted measures, like species richness itself, satisfy an important and intuitive property called the "replication principle" or the "doubling property" (Hill, 1973): if N equally diverse assemblages with no shared species are pooled in equal proportions, then the diversity of the pooled assemblages should be N times the diversity of each single assemblage. Simple examples show that Shannon's entropy and Gini–Simpson measures do not obey the "replication principle." However the transformed values of these indices do obey the replication principle.

Hill Numbers

The ecologist Mark Hill incorporated the transformed Shannon and Gini–Simpson measures, along with species richness, into a family of diversity measures later called "Hill numbers," all of which measure diversity as the effective number of species. Different Hill numbers ${}^qD$ are defined by their "order" q as (Hill, 1973)

$${}^qD = \left(\sum_{i=1}^{S} p_i^q\right)^{1/(1-q)} \quad [23a]$$

This equation is undefined for $q = 1$, but in the limit as q tends to 1:

$${}^1D = \lim_{q \to 1} {}^qD = \exp\left(-\sum_{i=1}^{S} p_i \log p_i\right) = \exp(H_{Sh}) \quad [23b]$$

The parameter q controls the sensitivity of the measure to species relative abundance. When $q = 0$, the species relative abundances do not count at all (no "discounting" for uneven abundances), and ${}^0D$ equals species richness. When $q = 1$, the Hill number ${}^1D$ is the exponential form of Shannon entropy, which weighs species in proportion to their frequency and can be roughly interpreted as the number of "typical species" in the assemblage (Chao *et al.*, 2010; Chao and Jost, in press). When $q = 2$, ${}^2D$ equals $1/(1 - H_{GS})$, which heavily weights the most common species in the assemblage; the contribution from rare species is severely discounted. The measure ${}^2D$ can be roughly interpreted as the number of "very abundant species" in the assemblage. Because all Hill numbers of higher order place increasingly greater weight on the most abundant species, they are much less sensitive to sample size (number of individuals or plots surveyed) than the most popular Hill numbers ($q = 0, 1, 2$). Hill numbers with negative exponents can also be calculated, but they place so much weight on rare species they have poor sampling properties.

Thus, the measure of diversity using Hill numbers can potentially depend on the order q that is chosen. However, because all Hill numbers need not be integers, and all have common units of species richness, they can be portrayed on a single graph as a function of q. This "diversity profile" of effective species richness versus q portrays all of the information about species abundance distribution of an assemblage (**Figure 6**). The diversity profile curve is a decreasing function of q (Hill, 1973). The more uneven the distribution of relative abundances, the more steeply the curve declines. For a perfectly even assemblage, the profile curve is a constant at the level of species richness.

Estimation of Hill Numbers

All of the Hill numbers (including species richness) as well as the untransformed Gini–Simpson index and Shannon entropy are sensitive to the number of individuals and samples collected. The sample-size dependence diminishes as q increases because the higher-order Hill numbers are more heavily weighted by frequencies of common species, and the estimates of those frequencies are not very sensitive to sample size. In contrast, with increasing numbers of individuals or samples collected, rare species continue to be added to the sample, making richness and other low Hill numbers more sample size dependent.

The MLE for the Gini–Simpson index, $\hat{H}_{GS,MLE} = 1 - \sum_{i=1}^{S} (X_i/n)^2$, is biased downward, and the bias in some cases can be substantial. The minimum variance unbiased estimator (MVUE) of the Gini–Simpson index has the following relationship to its MLE:

$$\begin{aligned}\hat{H}_{GS,MVUE} &= 1 - \sum_{i=1}^{S} [X_i(X_i - 1)]/[n(n-1)] \\ &= [n/(n-1)]\hat{H}_{GS,MLE}\end{aligned} \quad [24a]$$

This MVUE is equivalent to the estimator of *PIE* and is relatively invariant to sample size. Thus, a nearly unbiased estimator for Hill number of order 2 is

$${}^2\hat{D} = 1/\sum_{i=1}^{S} [X_i(X_i - 1)]/[n(n-1)] \quad [24b]$$

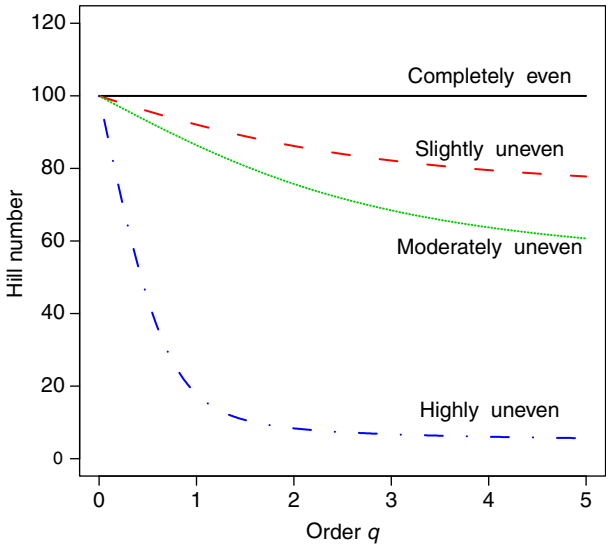


Figure 6 Diversity profile for assemblages of differing evenness. The *x*-axis is the order q in the Hill number (eqn [23a]), and is illustrated for values of q from 0 to 5. The *y*-axis is the calculated Hill number (the equivalent number of equally abundant species). Each of the four assemblages has exactly 100 species and 500 individuals, but they differ in their relative evenness: (1) completely even assemblage (black solid line): each species is represented by five individuals; (2) slightly uneven assemblage (red dashed line): 50 species each represented by seven individuals and 50 species each represented by three individuals (this structure is denoted as $\{50 \times 7, 50 \times 3\}$); (3) moderately uneven assemblage (green dotted line): $\{22 \times 10, 28 \times 5, 40 \times 3, 10 \times 2\}$; (4) highly uneven assemblage (blue dash–dot line): $\{1 \times 120, 1 \times 80, 1 \times 70, 1 \times 50, 3 \times 20, 3 \times 10, 90 \times 1\}$. For $q=0$, the Hill number is species richness, which is equal to 100 for all assemblages. Because Hill numbers represent the equivalent number of equally abundant species, the curve for the perfectly even assemblage (black solid line) does not change as q is increased. Larger values of q place progressively more weight on common species, so the equivalent number of equally abundant species is much lower for the more uneven assemblages than for more even assemblages.

For an integer $q \geq 2$, a similar derivation leads to a nearly unbiased estimator for ${}^{q}D$.

$$ {}^{q}\hat{D} = \left\{ \sum_{i=1}^{S} [X_i(X_i-1)\cdots(X_i-q+1)] \Big/ [n(n-1)\cdots(n-q+1)] \right\}^{1/(1-q)} \qquad [24c] $$

These estimators for $q \geq 2$ are almost independent of sample size, because all the higher-order Hill numbers are mainly dominated by the number of very abundant species. The estimated diversity profile curve is thus generally slowly varying for $q \geq 2$.

The estimation of entropy has been well studied in information science, physics, and statistics. Unfortunately, an unbiased estimator for Shannon entropy does not exist for any fixed sample size of n. As noted earlier, using X_i/n as a simple estimator of the true p_i value yields the MLE of entropy, which is negatively biased. An estimator of Shannon entropy with low bias is the following Horvitz–Thompson-type estimator (Chao and Shen, 2003):

$$ \hat{H}_{\mathrm{Sh}} = -\sum_{i=1}^{S} \frac{\tilde{p}_i \log(\tilde{p}_i)}{1-(1-\tilde{p}_i)^n} \qquad [25a] $$

where $\tilde{p}_i = (X_i/n)(1 - f_1/n)$ is an estimator of the true p_i. Only the detected species contribute to the summation because $\tilde{p}_i = 0$ for any undetected species. The denominator $1-(1-\tilde{p}_i)^n$ is the estimated probability that the ith species is detected in the sample, and the inverse of this probability is used as a weight for the ith species. Thus, the larger the probability of detection, the smaller the weight in the Horvitz-Thompson estimator. The weights adjust the estimator to compensate for missing species. For $q=1$, a low-bias estimator of the Hill number is

$$ {}^{1}\hat{D} = \exp\left(-\sum_{i=1}^{S} \frac{\tilde{p}_i \log(\tilde{p}_i)}{1-(1-\tilde{p}_i)^n} \right) \qquad [25b] $$

In summary, the statistical properties of Hill numbers depend on the order q. Equation [24b] is a nearly unbiased estimator of diversity for $q=2$, and eqn [25b] is a low-bias estimator for $q=1$. As discussed in the section Species richness estimation, total species richness (${}^{0}D=S$) is much more difficult to estimate because it is very sensitive to rare species that are often undetected, even in relatively large samples. Several nonparametric species richness estimators that can be used for estimating ${}^{0}D$ are provided in the section Nonparametric Asymptotic Species Richness Estimators. Then an estimated diversity profile can be constructed by plotting $\{{}^{q}\hat{D}; q = 0,1,2,3,\ldots\}$ with respect to q, based on estimators given in eqns [24b], [24c], and [25b]. The variance of each estimator in the profile can be approximated by a standard asymptotic method, and a 95% confidence interval can thus be constructed as the estimator $\pm$ 1.96 s.e. for each value of q, if the sample size is sufficiently large.

Taxonomic and PD

To quantify taxonomic or PD, species are placed on a branching tree (a cladogram) that describes their evolutionary relationships (**Figure 2**). The base of the tree represents the ancestral taxon, the branching forks (nodes) represent speciation or divergence events, the branch tips represent the contemporary species (not all of which may be represented in any particular assemblage), and time is measured in the vertical axis, increasing from the base of the tree to the branch tips. (For paleontological applications, the tips may be extinct lineages.) All other things being equal, an assemblage in which all the species are closely related and concentrated in one region of the tree should be less diverse than an assemblage in which the same number of species is widely distributed among distant branch tips of the tree.

This article distinguishes two types of phylogenetic trees: ultrametric and nonultrametric trees. A tree is called ultrametric if all branch tips are the same distance from the basal node. For example, if the branch lengths are proportional to divergence time, the tree is ultrametric. A Linnean taxonomic tree, in which species are simply classified into a taxonomic hierarchy (Kingdom, Phylum, Class, Order, Family, Genus, and

Species), can be regarded as a special case of an ultrametric tree. In contrast, if the branch lengths are proportional to the number of base-pair changes in a given gene, or some other measure of genetic or morphological change, some branch tips may be farther in absolute time from the basal node than other branch tips, and such trees are nonultrametric.

Pielou (1975) was the first to notice that the concept of diversity could be broadened to consider differences among species. The earliest taxonomic diversity measure is the *cladistic diversity* (*CD*), which is defined as the total number of taxa or nodes in a taxonomic tree that encompasses all of the species in the assemblage (Vane-Wright *et al.*, 1991). Another pioneering work is Faith's (1992) *PD*, which is defined as the sum of the branch lengths of a phylogeny connecting all species in the target assemblage. In both *CD* and *PD*, species abundances are not considered.

C.R. Rao's *quadratic entropy* was the first diversity measure that accounted for both phylogeny and species abundances (Rao, 1982). It is a generalization of the Gini–Simpson index:

$$Q_{\text{Rao}} = \sum_{i,j} d_{ij} p_i p_j \qquad [26a]$$

where d_{ij} denotes the phylogenetic distance between species i and j, and p_i and p_j denote the relative abundance of species i and j. This index measures the average phylogenetic distance between any two individuals randomly selected from the assemblage. For the special case of no phylogenetic structure (all species are equally related to one another), $d_{ii}=0$ and $d_{ij}=1$ ($i \neq j$), and Q_{Rao} reduces to the Gini–Simpson index.

The *phylogenetic entropy* H_p is defined as a generalization of Shannon's entropy to incorporate phylogenetic distances among species (Allen *et al.*, 2009):

$$H_P = -\sum_i L_i a_i \log a_i \qquad [26b]$$

where the summation is over all branches, L_i is the length of branch i, and a_i denotes the summed abundance of all species descended from branch i. The notation a (abundance) here is not the same as that (for area) in eqn [4].

The replication principle can be generalized to phylogenetic or functional diversity: When N completely distinct trees (no shared nodes during the fixed time interval of interest) with equal diversities are combined, the diversity of the combined tree is N times the diversity of any individual tree. Simple examples can show that neither Q_{Rao} nor H_p satisfies this replication principle. As with traditional diversity measures, Q_{Rao} and H_p can be transformed into measures that do obey the replication principle. In an ultrametric tree with tree height T^* (the time interval between the tree base and tips), the transformed measures are, respectively, $\exp(H_p/T^*)$ and $1/[1-(Q_{\text{Rao}}/T^*)]$.

For ultrametric trees, in addition to the order q, a time parameter T is required to generalize Hill numbers to measure PD in an interval from $-T$ time steps to the present time. The generalized phylogenetic metric is a time-averaged measure of lineage diversity (Hill numbers) at any moment t over the time interval $[-T, 0]$. The lineage diversity ${}^qD(t)$ at any moment t is measured by taking a "cross-section" and finding the lineages that are intersected (**Figure 7**). The relative abundance of each lineage is defined as the sum of the relative abundances of all of the descendants of that lineage in the present-day assemblage. Thus, ${}^qD(t)$ can be quantified by a Hill number. The average of these Hill numbers ${}^qD(t)$ over the interval $[-T, 0]$ gives the *mean PD* of order q over T time steps (Chao *et al.*, 2010):

$$ {}^q\overline{D}(T) = \left\{ \sum_{i \in B_T} \frac{L_i}{T} a_i^q \right\}^{1/(1-q)} \qquad [27]$$

where $\boldsymbol{B}_T$ denotes the set of all branches in the time interval $[-T, 0]$, L_i is the length (duration) of branch i in the set $\boldsymbol{B}_T$, and a_i is the total abundance of extant species descending from branch i; see **Figure 7** for a hypothetic ultrametric tree.

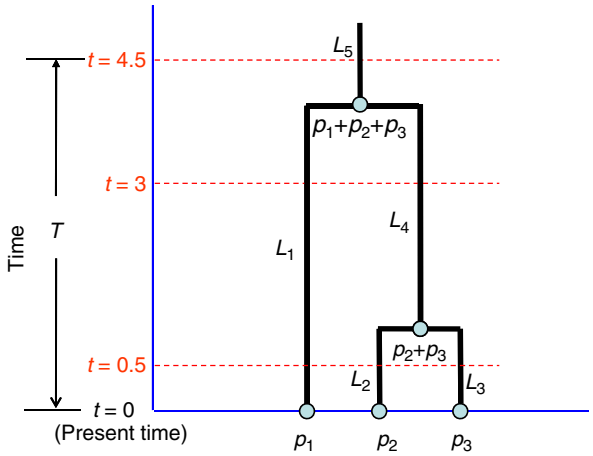


Figure 7 Calculation of mean phylogenetic diversity and branch diversity. In this hypothetical rooted phylogenetic tree, the ancestor to the assemblage is depicted at the top, and there are three extant species living at the present, depicted at the bottom, with relative abundances $(p_1, p_2, p_3) = (0.2, 0.3, 0.5)$. The tree is ultrametric, so the total branch length from the ancestor to any descendant species in the present is the same. To evaluate the phylogenetic diversity at the time $T = 5$ time steps in the past, we first create the set $\boldsymbol{B}_T$ which includes five branches with lengths $(L_1, L_2, L_3, L_4, L_5) = (4, 1, 1, 3, 1)$ and the corresponding abundances $(a_1, a_2, a_3, a_4, a_5) = (p_1, p_2, p_3, p_2+p_3, p_1+p_2+p_3)$. We next measure the lineage diversity ${}^qD(t)$ at any time steps $0 < t < T$. We use three different "sampling times" as examples; these three sampling times correspond to the three distinct assemblages that would be represented by diversity sampling at three points in the past. For the first sampling time, $t=0.5$, lineage diversity ${}^qD(t)$ is measured as the Hill numbers for three lineages (species) with relative abundances (p_1, p_2, p_3); for the second sampling time $t=3$, lineage diversity ${}^qD(t)$ is measured as the Hill numbers for two lineages (species) with relative abundances (p_1, p_2+p_3); for the third sampling time $t=4.5$, lineage diversity ${}^qD(t)$ is measured as the Hill numbers for only one lineage (species) with relative abundance $p_1+p_2+p_3=1$. The average of these Hill numbers ${}^qD(t)$ over the interval $[-T, 0]$ gives the *mean phylogenetic diversity* ${}^q\overline{D}(T)$ of order q over T time steps. The branch diversity is ${}^qPD(T) = T \times {}^q\overline{D}(T)$. These two diversities can be calculated for any fixed $T \geq 0$; their pattern as a function of T is plotted in **Figure 8**. (For example, if T is changed to the tree height $T^*=4$ (distance between tree base and tips), then the branch set includes four branches $(L_1, L_2, L_3, L_4)=(4, 1, 1, 3)$ with the corresponding abundances $(a_1, a_2, a_3, a_4)=(p_1, p_2, p_3, p_2+p_3)$, and the lineage diversity is averaged over $[-4, 0]$ to obtain ${}^q\overline{D}(T^*)$ and ${}^qPD(T^*) = T^* \times {}^q\overline{D}(T^*)$.)

This measure ${}^{q}\overline{D}(T)$ gives the mean effective number of maximally distinct lineages (or species) T time steps in the past. The diversity of a tree with ${}^{q}\overline{D}(T)=z$ in the time period $[-T, 0]$ is the same as the diversity of an assemblage consisting of z equally abundant and maximally distinct species with all branch lengths T.

The branch diversity or phylogenetic diversity ${}^{q}\mathrm{PD}(T)$ of order q through T time steps before present is defined as the product of ${}^{q}\overline{D}(T)$ and T . The measure ${}^{q}\mathrm{PD}(T)$ given below quantifies "the total effective number of lineage-lengths or lineage-time steps" (Chao *et al.*, 2010)

$$ {}^{q}\mathrm{PD}(T) = T \times {}^{q}\overline{D}(T) = \left\{ \sum_{i \in B_T} L_i \left(\frac{a_i}{T} \right)^{q} \right\}^{1/(1-q)} \qquad [28] $$

If $q=0$ and $T=T^*$ (tree height), then ${}^{0}\mathrm{PD}(T)$ reduces to Faith's PD. It also reduces to CD in a taxonomic tree if the branching of each Linnaean taxonomic category is assigned a time step of unit length. A PD profile can be constructed by plotting both ${}^{q}\mathrm{PD}(T)$ and ${}^{q}\overline{D}(T)$ as a function of T for $q=0$, 1, and 2. It is also informative to construct another diversity profile by plotting ${}^{q}\mathrm{PD}(T)$ and ${}^{q}\overline{D}(T)$ as a function of order q for some selected values of temporal perspective T. See **Figure 8** for a numerical example. In most applications, ecologists are interested in the case $T=T^*$ (tree height) or the divergence time between the species group of interest and its nearest outgroup. The divergence time of the most recent common ancestor of all extant taxa is another useful comparison.

For nonultrametric trees, the time parameter T is generalized to $\overline{T}$, where $\overline{T}=\sum_{i\in B_{\overline{T}}} L_i a_i$ represents the abundance-weighted mean base change per species and $\boldsymbol{B}_{\overline{T}}$ denote the set of branches connecting all focal species. The diversity of a nonultrametric tree with mean evolutionary change $\overline{T}$ is the same as that of an ultrametric tree with a time step $\overline{T}$.

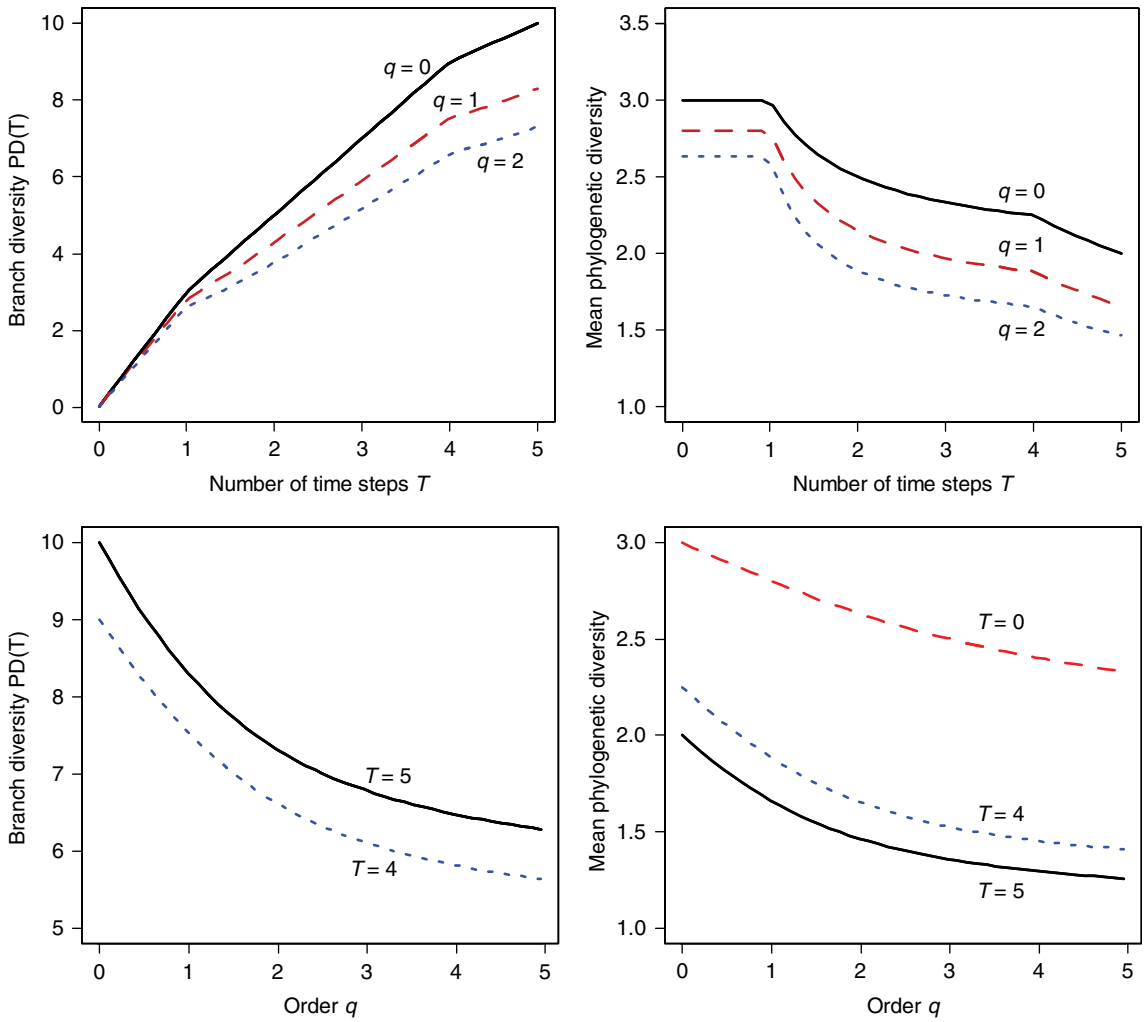


Figure 8 Branch diversity profile and mean phylogenetic diversity profile. (a) Branch diversity profile of ${}^{q}\mathrm{PD}(T)$ (upper left panel) and mean phylogenetic diversity (upper right panel) as a function of the number of time steps (T) in the past ($0<T<5$) for $q=0$, 1, and 2, based on the structure of the phylogenetic tree in **Figure 7**, assuming $(p_1, p_2, p_3)=(0.2, 0.3, 0.5)$ and $(L_1, L_2, L_3, L_4, L_5)=(4, 1, 1, 3, 1)$. (b) Branch diversity profile of ${}^{q}\mathrm{PD}(T)$ (lower left panel) as a function of order q for $T=4$ (tree height) and $T=5$ time steps, and mean phylogenetic diversity profile of (lower right panel) as a function of order q for $T=0$, 4, and 5, assuming $(p_1, p_2, p_3)=(0.2, 0.3, 0.5)$ and $(L_1, L_2, L_3, L_4, L_5)=(4, 1, 1, 3, 1)$. The branch diversity for $T=0$ is 0 as all branch lengths are 0. The mean phylogenetic diversity for $T=0$ are the traditional Hill numbers for $(p_1, p_2, p_3)=(0.2, 0.3, 0.5)$.

Therefore, the diversity formula for a nonultrametric tree is obtained by replacing T in the ${}^{q}\overline{D}(T)$ and ${}^{q}\mathrm{PD}(T)$ with $\overline{T}$. Equation [27] can also describe taxonomic diversity, if the phylogenetic tree is a Linnaean tree with L levels (ranks), and each branch is assigned unit length. It also describes functional diversity, if a dendrogram can be constructed from a trait-based distance matrix using a clustering scheme (Petchey and Gaston, 2002). Thus, Hill numbers can be effectively generalized to incorporate taxonomy, phylogeny, and function and provide a unified framework for measuring biodiversity (Chao and Jost, in press).

Estimation of phylogenetic and functional diversity from small samples has not been well studied. As with the estimation of simple Hill numbers, phylogenetic diversity ${}^{q}\overline{D}(T)$ and ${}^{q}\mathrm{PD}(T)$ can be accurately estimated only for $q=2$. Like the Gini–Simpson index, the MVUE of Rao's quadratic entropy exists under a multinomial model:

$$\hat{Q}_{\mathrm{MVUE}} = \sum_{i,j} d_{ij} X_i X_j / [n(n-1)] \qquad [29]$$

Thus, ${}^{2}\overline{D}(T)$ and ${}^{2}\mathrm{PD}(T)$ can be estimated by a nearly unbiased measure using this estimator based on the transformation ${}^{2}\overline{D}(T) = 1/[1-(Q_{\mathrm{Rao}}/T)]$. With sufficient sampling, these estimators of PD of order $q=2$ are almost independent of sample size. Further research is needed for the development of accurate estimators of PD measures with $q=1$ and 0.

Biotic Similarity

Incidence-Based Similarity Indices

The earliest published incidence-based measure of relative compositional similarity is the classic Jaccard index from 1900. A number of incidence-based similarity measures have been proposed since then (see Jost *et al.*, 2011, for a review). The Jaccard index and the Sørensen index (proposed in 1948) are the most widely used ones, and both were originally developed to compare the similarity of two assemblages. Let S_1 be the number of species in Assemblage 1, S_2 be the number of species in Assemblage 2, and S_{12} be the number of shared species. The Jaccard similarity index$=S_{12}/(S_1+S_2-S_{12})$ and the Sørensen similarity index$=2S_{12}/(S_1+S_2)$. A rearrangement of the Sorensen index$=1/[0.5(S_{12}/S_1)^{-1}+0.5(S_{12}/S_2)^{-1}]$ reveals that it is the harmonic mean of two proportions: S_{12}/S_1 (the proportion of the species in the first assemblage that are shared with the second) and S_{12}/S_2 (the proportion of the species in the second assemblage that are shared with the first). The Jaccard index compares the number of shared species to the total number of species in the combined assemblages, whereas the Sørensen index compares the number of shared species to the mean number of species in a single assemblage. The Jaccard index is thus a comparison based on total diversity, whereas the Sørensen index is a comparison based on local diversity.

When one assemblage is much richer than the other, both Sørensen and Jaccard indices become very small. Although the low similarity value reflects the true difference between the two assemblages, in some applications it can be more informative to normalize a similarity measure so that maximum overlap$=1.0$. Lennon *et al.* (2001) proposed such a modification to the Sørensen index, and it takes the form $S_{12}/\min(S_1, S_2)$; see Jost *et al.* (2011) for details and comparisons.

When more than two assemblages are compared, a typical approach is to use the average of all pairwise similarities as a measure of global similarity. However, the pairwise similarities calculated from data tend to be correlated and are not independent. Most importantly, pairwise similarities cannot fully characterize multiple-assemblage similarity when some species are shared across two, three, or more assemblages (Chao *et al.*, 2008). It is easy to construct numerical examples in which all pairwise similarities are identical in two sets of assemblages, but the global similarities for the two sets are different.

The two-assemblage incidence-based Jaccard and Sørensen indices have been extended to multiple assemblages. Assume that there are N assemblages and there are S_j species in the jth assemblage and S species in the combined assemblage. Let $\overline{S}$ denote the average number of species per assemblage. The multiple-assemblage Jaccard similarity index$=(\overline{S}/S-1/N)/(1-1/N)$. The multiple-assemblage Sørensen similarity index$=(N-S/\overline{S})/(N-1)$. When $N=2$, these two measures reduce to their classical two-assemblage measures. These two measures are decreasing functions of Whittaker's beta diversity for species richness, which is $S/\overline{S}$. When N assemblages are identical, beta diversity ($q=0$) is $S/\overline{S}=1$, and thus both Jaccard and Sørensen similarity indices$=1$. When N assemblages are completely distinct (no shared species), beta diversity ($q=0$) is $S/\overline{S}=N$, and thus both Jaccard and Sørensen similarity indices$=0$.

These incidence-based similarity indices are widely used in ecology and biogeography because of their simplicity and easy interpretation. In most ecological studies, these indices are estimated from observed richness in sample data. The resulting estimates are generally biased downward, and the bias increases when sample sizes are small or species richness is large. They could become biased upward when shared species are common and endemic species are very rare (Chao *et al.*, 2005, p. 149). The classic pairwise Jaccard and Sorensen similarity indices calculated from sample data generally underestimate the true similarity mainly because they do not account for shared species at both sites that were not detected. One strategy could be to use asymptotic species richness estimators (*see* Species Richness Estimation) to estimate species richness in each assemblage and also to estimate species richness in the combined assemblage, and then substitute the estimated values into the similarity formulas. However, this strategy inevitably inflates the variance and often renders the resulting estimate useless. A major statistical concern is that, based on incidence data alone, bias correction and measurements of variances are impossible. Consequently, the interpretation of any incidence-based index based on sample values or estimated values becomes difficult or misleading for comparing two (or more) highly diverse assemblages based on limited data. Only with abundance data can one correct for undersampling bias, as explained in the next section.

Classical incidence-based similarity indices treat abundant and rare species equally, which oversimplifies the relationships between assemblages. If species abundances can be measured, they should be used for a more accurate

representation (and better statistical estimation) of the similarity of assemblages.

Abundance-Based Similarity Indices

Assume that in the combined assemblages, there are S species. Denote the relative abundance vector for the S species in the jth assemblage by $(p_{1j},p_{2j},\ldots,p_{sj})$, some of them may be 0. Thus, for N assemblages, there are N sets of abundances $\{(p_{1j},p_{2j},\ldots,p_{sj});\ j=1,2,\ldots,N\}$. A sample of n_j individuals is taken from the jth assemblage and there are N sets of sample frequencies $\{(X_{1j},X_{2j},\ldots,X_{sj});\ j=1,2,\ldots,N\}$.

For two-assemblage cases, one of the most popular abundance-based similarity metric is the Horn overlap measure (Horn, 1966), which is based on Shannon's entropy:

$$S_H = \frac{1}{\log 2}\sum_{i=1}^{S}\left[\frac{p_{i1}}{2}\log\left(1+\frac{p_{i2}}{p_{i1}}\right)+\frac{p_{i2}}{2}\log\left(1+\frac{p_{i1}}{p_{i2}}\right)\right] \quad [30]$$

Another popular overlap measure is the Morisita–Horn similarity measure (Morisita, 1959), based on the Simpson index:

$$S_{MH} = \frac{\sum_{i=1}^{S} p_{i1}p_{i2}}{\left[\sum_{i=1}^{S} p_{i1}^2+\sum_{i=1}^{S} p_{i2}^2\right]/2} = 1-\frac{\sum_{i=1}^{S}(p_{i1}-p_{i2})^2}{\sum_{i=1}^{S} p_{i1}^2+\sum_{i=1}^{S} p_{i2}^2} \quad [31]$$

Since each of the indices [30] and [31] equals unity if and only if $p_{i1}=p_{i2}$ for all i, these two indices match relative abundances on a species-by-species basis. When the two assemblages are equally diverse and consist entirely of equally common species, the Morisita–Horn index, the Horn index, and the Sørensen index are all equal, and all of these indices give the proportion of shared species in an assemblage.

The first expression in eqn [31] for the Morisita-Horn index has an important probabilistic interpretation. If one individual is selected randomly from each assemblage, then the probability that the two selected individuals belong to *the same shared* species is $\sum p_{i1}p_{i2}$, the numerator in eqn [31]. The denominator in eqn [31] represents a normalizing constant, which is the average of two such probabilities for two individuals drawn from the same assemblages. In this probabilistic interpretation, the abundant species will contribute the most to the probability that two randomly selected individuals belong to the *same* species. As a result, in a hyperdiverse assemblage, the index is dominated by a few abundant species and the relatively rare species (even if there are many of them) have little effect. The index is therefore likely to be resistant to undersampling, because the influential abundant species are always present in samples.

The following MLE of the Morisita–Horn index is always in the range [0, 1], but it has been shown that this MLE systematically underestimates the true similarity.

$$\tilde{S}_{MH,\ MLE} = \frac{2\sum_{i=1}^{S}(X_{i1}/n_1)(X_{i2}/n_2)}{\sum_{i=1}^{S}(X_{i1}/n_1)^2+\sum_{i=1}^{S}(X_{i2}/n_2)^2} \quad [32]$$

A better estimator that is nearly unbiased has the following form, although it may exceed the theoretical maximum value of 1:

$$\tilde{S}_{MH} = \frac{2\sum_{i=1}^{S}(X_{i1}/n_1)(X_{i2}/n_2)}{\sum_{i=1}^{S}(X_{i1}(X_{i1}-1)/[n_1(n_1-1)])+\sum_{i=1}^{S}(X_{i2}(X_{i2}-1)/[n_2(n_2-1)])} \quad [33]$$

It is statistically infeasible to derive analytic or asymptotic variance formulas for these two estimators and the other similarity estimators in this section. A bootstrap method, suggested in Chao *et al.* (2008), is a simple and direct data-resampling method to obtain approximate estimates of variances and confidence intervals especially for complex estimators (Efron and Tibshirani, 1993). This method has found wide applications in various disciplines.

To the extent that ecological processes are often most strongly influenced by abundant species, the Morisita–Horn measure is useful when looking for functional differences between ecosystems. However, when rare species are important, as in many conservation applications, the Horn overlap measure would be more useful. But the MLE of the Horn overlap measure exhibits moderate bias due to undersampling. Until now, only the two-sample jackknife technique has been used to remove part of bias (Jost *et al.*, 2011). More research is required to find a reliable bias-reduced estimator.

For assessing similarity among more than two assemblages, a general multiple-assemblage, abundance-based overlap measure C_{qN} (Chao *et al.*, 2008) is

$$C_{qN} = \frac{[1/(N^q-N)]\sum_{i=1}^{S}\left[(p_{i1}+p_{i2}+\cdots+p_{iN})^q-(p_{i1}^q+p_{i2}^q+\cdots+p_{iN}^q)\right]}{(1/N)\sum_{i=1}^{S}(p_{i1}^q+p_{i2}^q+\cdots+p_{iN}^q)} \quad [34]$$

As with the Hill numbers, here q is a parameter that determines the measure's sensitivity to species' relative abundances, and N is the number of assemblages. The C_{qN} measure includes, as special cases, the classic two-assemblage Sørensen index ($q=0$, $N=2$), the Horn overlap index ($q=1$, $N=2$), the Morisita–Horn similarity index ($q=2$, $N=2$), and their multiple-assemblage generalizations ($N>2$) as follows:

For $q=0$, C_{0N} is the multiple-assemblage Sørensen similarity index:

$$C_{0N} = (N-S/\overline{S})/(N-1) \quad [35a]$$

For $q=1$, C_{1N} is the multiple-assemblage Horn overlap index:

$$C_{1N} = \frac{1}{\log N}\sum_{i=1}^{S}\sum_{j=1}^{N}\left[\frac{p_{ij}}{N}\log\left(1+\frac{\sum_{k\neq j}p_{ik}}{p_{ij}}\right)\right] \quad [35b]$$

For $q=2$, C_{2N} is the multiple-assemblage Morisita–Horn similarity index:

$$C_{2N} = \frac{2\sum_{i=1}^{S}\sum_{j<k}p_{ij}p_{ik}}{(N-1)\sum_{i=1}^{S}\sum_{j=1}^{N}p_{ij}^2} \quad [35c]$$

Jost (2007) was the first to develop a rigorous mathematical formulation of alpha and beta diversities based on Hill numbers of order q. He derived the multiplicative beta diversity ${}^qD_\beta$, which quantifies the effective number of completely distinct assemblages; see also Jost *et al.* (2011), Chao and Jost (in press), and Tuomisto (this volume) for reviews. For

equally weighted assemblages, beta diversity ${}^{q}D_{\beta}$ ranges between a minimum of 1 (when all assemblages are identical) and a maximum of N, the number of assemblages in each region (when all assemblages are completely distinct; i.e., there are no shared species). For example, a set of completely distinct sites in a region of three sites attains the maximum value of 3, whereas another set of completely distinct sites in a region of 10 sites attains the maximum value of 10. Because the maximum depends on the number of assemblages in the region, beta diversities usually cannot be compared directly among multiple regions. Instead, beta diversity should be compared with sample-based rarefaction to a common number of samples or to a common degree of completeness of samples in each region. However, beta diversity can be transformed to the C_{qN} measure in the range [0, 1] by the following nonlinear transform for N equally weighted assemblages:

$$C_{qN} = [(1/{}^{q}D_{\beta})^{q-1} - (1/N)^{q-1}]/[1 - (1/N)^{q-1}] \qquad [36]$$

The transformed measure C_{qN} is unity (when all assemblages are identical) and 0 (when all assemblages are completely distinct).

This nonlinear transformation ensures that C_{qN} preserves an essential property of an overlap index: The transformed index C_{qN} gives the true *overlap* A/S for *all* orders of q if N assemblages each have S equally common species, exactly A species are shared by all of them, and the remaining species are not shared by any assemblages. No linear transformation of beta diversity can achieve this property. See **Figure 9**, Case I, for an example. The measure C_{qN} thus quantifies the *effective average overlap per assemblage*, i.e., average *percentage of overlapped species* (species that are shared by all assemblages, as defined in the above sense) in an assemblage. Because this measure is in a sense of "average" overlap, it can be compared across regions with different number of assemblages. See Chao *et al.* (2008) for more details on the interpretation of C_{qN} based on a simple reference set of assemblages.

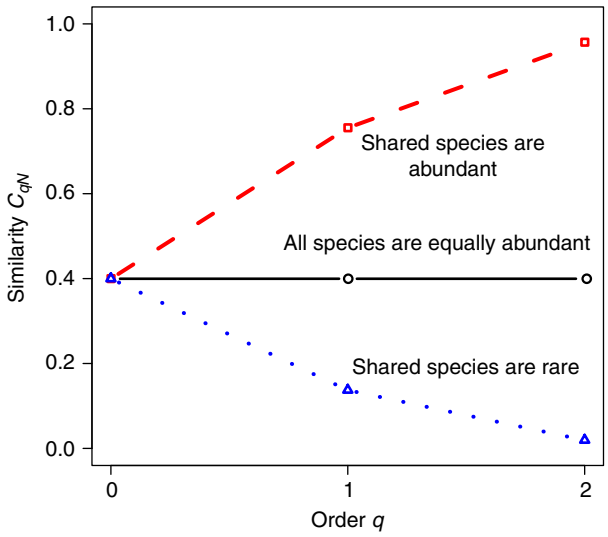


Figure 9 The two-assemblage similarity profile C_{qN} for $q=0$, 1, 2. *Case* I: In each of Assemblages 1 and 2, there are 20 equally abundant species, and eight species are shared. Thus, for all orders of q, the similarity measure C_{qN} equals the percentage of overlap in each assemblage. That is, $C_{qN}=8/20=40\%$ for all q (*black solid line*).
Case II: In each of Assemblages 1 and 2, there are 20 species, and the relative abundances in each assemblage are $p_i \propto K/i$ for $i=1, 2, \ldots, 20$, where K is a normalizing constant such that the total relative abundances is 1. Assume that the shared species are the most abundant eight species. The similarity measure C_{qN} increases as order q is increased (*red dashed line*).
Case III: Same as in Case II, but the shared species are the rarest eight species. The similarity measure C_{qN} decreases as order q is increased (*blue dotted line*).

Just as diversity profiles are used to characterize traditional diversity (**Figure 6**) and PD (**Figure 8**), Chao *et al.* (2008) suggest the use of a similarity profile $\{C_{qN}; q=0, 1, 2, \ldots, N\}$ to describe similarity across N assemblages. It is recommended that investigators calculate at least C_{0N}, C_{1N}, and C_{2N}; see Jost *et al.* (2011, p. 81) and **Figure 9** for examples.

Because the overlap measure C_{qN} is constructed from Hill numbers of order q, the magnitude of the undersampling bias depends on the order q. As with Hill numbers, there exists a nearly unbiased estimator only for $q=2$:

$$\hat{C}_{2N} = \frac{2\sum_{i=1}^{S}\sum_{j<k}(X_{ij}/n_j)(X_{ik}/n_k)}{(N-1)\sum_{i=1}^{S}\sum_{j=1}^{N}[X_{ij}(X_{ij}-1)/n_j(n_j-1)]} \qquad [37]$$

This measure quantifies the similarity of relative frequencies for abundant species. A bootstrap variance estimator and the associated confidence interval are provided in Chao *et al.* (2008). Satisfactory estimators for C_{0N} and C_{1N} that characterize the similarity of relative frequencies for rare species are still lacking.

Similarity Indices Based on Total Abundance of Shared Species

As explained earlier, the Horn overlap measure (eqn [30]) and the Morisita–Horn similarity measure (eqn [31]) match species relative abundances, *species-by-species*. Hence, the typical similarity indices assess a normalized probability that two randomly chosen individuals, one from each assemblage, belong to *the same* species. Another approach by Chao *et al.* (2005) is to consider a (normalized) probability that both individuals, one from each of the two assemblages, both belong to *any shared* species (and not necessarily to the same shared species).

Let U denote the total relative abundances associated with the shared species in Assemblage 1 and let V denote the total relative abundances of the shared species in Assemblage 2. The Jaccard abundance-based similarity index is $UV/(U+V-UV)$ and the Sørensen abundance-based similarity index is $2UV/(U+V)$. These two shared-abundance indices are called the Chao–Jaccard abundance and Chao–Sørensen abundance indices in the literature and in the software package *EstimateS* (Colwell, 2011) and SPADE (Chao and Shen, 2010). This is because, when all species are equally common, $U=S_{12}/S_1$ and $V=S_{12}/S_2$ and the Chao–Jaccard abundance and Chao–Sørensen abundance indices reduce, respectively, to the classic incidence-based Jaccard and Sørensen indices. These two measures yield a maximum value of 1 when all species are shared (i.e., no unique species in both assemblages; $U=V=1$). Also, all indices tend to a minimum value of 0 for completely distinct assemblages (i.e., no shared species in both assemblages).

One advantage of these measures is that the undersampling bias due to unseen, shared species can be evaluated and corrected. Chao *et al.* (2005) used the frequencies of observed rare, shared species to obtain an appropriate adjustment term for U and V to account for the effect of *unseen* shared species and thus remove most undersampling bias. Then the bias-corrected U and V estimators are substituted into the formulas to obtain Chao–Jaccard and Chao–Sørensen estimators. These measures are designed to be sensitive to rare shared species while still taking abundance into account, so they may increase sharply as more shared species are discovered. Because these measures match the total relative abundances of species shared between two assemblages, they are useful if the focus is to construct abundance-based *complementarity* (dissimilarity or distance) measures by subtracting each measure from one. This class of measures can also be extended to replicated incidence data; see Chao *et al.* (2005) for details.

Phylogenetic Similarity Indices

The classic Jaccard, Sørensen, and Morisita-Horn similarity measures all have their own phylogenetic generalizations. Most of the pioneering work was developed by microbial ecologists (Lozupone and Knight, 2005; Faith *et al.*, 2009). The phylogenetic Jaccard and Sørensen measures are based on Faith's total branch lengths and have formulas similar to their classic versions. The phylogenetic Sørensen index can be expressed as $2L_{12}/(L_1+L_2)$, where L_1 and L_2 denote the total branch lengths in Assemblages 1 and 2, respectively, and L_{12} denotes the total length of the shared branches in the same time interval of interest (Lozupone and Knight, 2005). The phylogenetic Jaccard index takes the form of $L_{12}/(L_1+L_2-L_{12})$. When species relatedness is based on a simple Linnean taxonomic classification tree, L_1 and L_2 become the number of taxa in Trees 1 and 2, respectively, and L_{12} becomes the number of shared taxa in the pooled classification tree (Bacaro *et al.*, 2007). In these generalizations, "species" in the traditional indices are replaced by the total branch lengths (or the total number of nodes) in each assemblage. Also, "shared species" in the traditional indices are replaced by the total shared branch length (or the total number of shared nodes). In nearly all applications of these phylogenetic similarity indices, it is assumed that all species are observed; undersampling bias due to undetected species for these measures has not been discussed in this literature.

The classic Morisita–Horn measure has recently been generalized to its phylogenetic version (de Bello *et al.*, 2010). Let $\boldsymbol{B}$ denote the set consisting of all branches in the pooled assemblages in a specific time period of interest, and let the corresponding branch lengths in this set be $\{L_i;\, i\in\boldsymbol{B}\}$. Assume that, in the jth assemblage, a_{ij} denotes the total relative abundance descended from Branch i, $i\in\boldsymbol{B}$, $j=1, 2, \ldots, N$. The phylogenetic Morisita–Horn similarity for N assemblages is a generalization of eqn [35c] based on a normalized Rao's quadratic entropy:

$$S^{*}_{\mathrm{MH}}=\frac{2\sum_{i\in B}\sum_{j<k}L_i a_{ij}a_{ik}}{(N-1)[\sum_{i\in B}\sum_{j=1}^{N}L_i a_{ij}^2]} \qquad [38]$$

A nearly unbiased estimator for this measure is similar to that in eqn [37]. However, the measure in eqn [38] is valid only for ultrametric trees. Extension to nonultrametric trees requires further research.

Appendix

List of Courses

1. Community Ecology
2. Conservation Biology
3. Statistical Ecology

See also: Biodiversity, Definition of. Defining, Measuring, and Partitioning Species Diversity. Diversity, Molecular Level. Diversity, Taxonomic versus Functional. Functional Diversity Measures. Latitudinal Gradients of Biodiversity. Measurement and Analysis of Biodiversity. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms. Species Diversity, Overview

References

Allen B, Kon M, and Bar-Yam Y (2009) A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. *American Naturalist* 174: 236–243.

Bacaro G, Ricotta C, and Mazzoleni S (2007) Measuring beta-diversity from taxonomic similarity. *Journal of Vegetation Science* 18: 793–798.

de Bello F, Lavergne S, Meynard CN, Lepš J, and Thuiller W (2010) The partitioning of diversity: Showing Theseus a way out of the labyrinth. *Journal of Vegetation Science* 21: 1–9.

Chao A (1984) Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11: 265–270.

Chao A (1987) Estimating the population size for capture–recapture data with unequal catchability. *Biometrics* 43: 783–791.

Chao A (2005) Species estimation and applications. In: Kotz S, Balakrishnan N, Read CB, and Vidakovic B (eds.) *Encyclopedia of Statistical Sciences*, 2nd edn, pp. 7907–7916. New York: Wiley.

Chao A, Chazdon RL, Colwell RK, and Shen T-J (2005) A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters* 8: 148–159.

Chao A, Chiu C-H, and Jost L (2010) Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B* 365: 3599–3609.

Chao A, Colwell RK, Lin C-W, and Gotelli NJ (2009) Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* 90: 1125–1133.

Chao A and Jost L (2011) Diversity measures. In: Hastings A and Gross L (eds.) *Encyclopedia of Theoretical Ecology*. Berkeley: University of California Press. In press.

Chao A, Jost L, Chiang SC, Jiang Y-H, and Chazdon RL (2008) A two-stage probabilistic approach to multiple-community similarity indices. *Biometrics* 64: 1178–1186.

Chao A and Shen T-J (2003) Nonparametric estimation of Shannon's index of diversity when there are unseen species. *Environmental and Ecological Statistics* 10: 429–443.

Chao A and Shen TJ (2004) Nonparametric prediction in species sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 9: 253–269.

Chao A and Shen T-J (2010) SPADE: Species Prediction and Diversity Estimation. Program and User's Guide at http://chao.stat.nthu.edu.tw/softwareCE.html.

Coleman BD, Mares MA, Willig MR, and Hsieh Y-H (1982) Randomness, area, and species richness. *Ecology* 63: 1121–1133.

Colwell RK (2011) EstimateS: Statistical estimation of species richness and shared species from samples. Version 9. User's Guide and application published at: http://purl.oclc.org/estimates.

Colwell RK, Chao A, Gotelli NJ, *et al.* (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation, and comparison of assemblages. *Journal of Plant Ecology* 5: 3–21.

Colwell RK and Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society B* 345: 101–118.

Colwell RK, Mao CX, and Chang J (2004) Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology* 85: 2717–2727.

Ellison AM, Gotelli NJ, Farnsworth EJ, and Alpert GD (2012) *A Field Guide to the Ants of New England*. New Haven, CT: Yale University Press.

Efron B and Tibshirani RJ (1993) *An Introduction to the Bootstrap*. London: Chapman and Hall.

Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61: 1–10.

Faith DP, Lozupone CA, Nipperess D, and Knight R (2009) The cladistic basis for the phylogenetic diversity (PD) measure links evolutionary features to environmental gradients and supports broad applications of microbial ecology's "phylogenetic beta diversity" framework. *International Journal of Molecular Sciences* 10: 4723–4741.

Gotelli NJ and Colwell RK (2001) Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4: 379–391.

Gotelli NJ and Colwell RK (2011) Estimating species richness. In: Magurran A and McGill B (eds.) *Biological Diversity: Frontiers in Measurement and Assessment*, pp. 39–54. Oxford: Oxford University Press.

Hill MO (1973) Diversity and evenness: A unifying notation and its consequences. *Ecology* 54: 427–431.

Horn HS (1966) Measurement of "overlap" in comparative ecological studies. *American Naturalist* 100: 419–424.

Hurlbert SH (1971) The nonconcept of species diversity: A critique and alternative parameters. *Ecology* 52: 577–586.

Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88: 2427–2439.

Jost L, Chao A, and Chazdon RL (2011) Compositional similarity and beta diversity. In: Magurran A and McGill B (eds.) *Biological Diversity: Frontiers in Measurement and Assessment*, pp. 66–84. Oxford: Oxford University Press.

Lennon JJ, Koleff P, Greenwood JJD, and Gaston KJ (2001) The geographical structure of British bird distributions: Diversity, spatial turnover and scale. *Journal of Animal Ecology* 70: 966–979.

Lozupone C and Knight R (2005) UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* 71: 8228–8235.

MacArthur RH (1965) Patterns of species diversity. *Biological Reviews* 40: 510–533.

Magurran AE (2004) *Measuring Biological Diversity*. Oxford: Blackwell.

Morisita M (1959) Measuring of interspecific association and similarity between communities. *Memoires of the Faculty of Science, Kyushu University, Series E (Biology)* 3: 65–80.

Niemelä J, Haila Y, Halme E, *et al.* (1988) The distribution of carabid beetles in fragments of old coniferous taiga and adjacent managed forest. *Annales Zoologici Fennici* 25: 107–199.

Petchey OL and Gaston KJ (2002) Functional diversity (FD), Species richness and community composition. *Ecology Letters* 5: 402–411.

Pielou EC (1975) *Ecological Diversity*. New York: John Wiley.

Rao CR (1982) Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology* 21: 24–43.

Sanders H (1968) Marine benthic diversity: A comparative study. *American Naturalist* 102: 243–282.

Shen T-J, Chao A, and Lin C-F (2003) Predicting the number of new species in further taxonomic sampling. *Ecology* 84: 798–804.

Shinozaki K (1963) Notes on the species-area curve. *10th Annual Meeting of the Ecological Society of Japan* (Abstract), p.5.

Tokeshi M (1999) *Species Coexistence: Ecological and Evolutionary Perspectives*. Oxford: Blackwell.

Tuomisto H (this volume) Defining, measuring and partitioning species diversity. *Encyclopedia of Biodiversity*.

Vane-Wright RI, Humphries CJ, and Williams PM (1991) What to protect: Systematics and the agony of choice. *Biological Conservation* 55: 235–254.

Mediterranean-Climate Ecosystems

Philip W Rundel, University of California, Los Angeles, CA, USA
Richard M Cowling, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Philip W. Rundel, pp 1–15, © 2007, Elsevier Inc.

Glossary

Chaparral Evergreen sclerophyllous-leaved shrublands that cover large areas of California.
Exaptation A trait that first evolved to serve one particular function, but subsequently comes to serve another.
Fynbos Evergreen sclerophyll vegetation dominating the Cape Region of South Africa.
Garrigue Low-growing secondzary evergreen shrublands that dominate extensive areas of the Mediterranean Basin.
Gondwanaland Former supercontinent of the Southern Hemisphere from which South America, Africa, Australia, and India are derived.
Kwongan Evergreen heathlands that cover extensive areas of southwestern Australia.
Maquis Vegetation cover of the Mediterranean Basin dominated by evergreen sclerophyllous shrubs and trees.
Matorral Evergreen sclerophyllous-leaved shrublands that dominate large areas of central Chile.
Mediterranean-type ecosystem (MTE) Habitat characterized by mild wet winters and warm dry summers. MTEs occur in California, central Chile, the Mediterranean Basin, the Cape Region of South Africa, and southwestern and South Australia, all at N and S latitudes of 30–35°.
Sclerophyllous Descriptive of leaves with a leathery texture due to the presence of sclerenchyma with large amounts of lignin and cellulose in their tissues.

Introduction

Mediterranean-type ecosystems (MTEs), with their characteristic and unique climatic regimes of mild wet winters and warm and dry summers, occur in just five regions of the world. These regions are California, central Chile, the Mediterranean Basin, the Cape Region of South Africa, and southwestern and South Australia (**Figure 1**, **Table 1**). Under this climatic regime, 90% or more of the annual precipitation typically falls in the 6 months centered on winter. The traditional boundaries of each region have been based not only on climatic regimes, but also on additional factors of floristics, vegetation structure, and historical precedent. Thus, there is not a single defining set of criteria for mapping the MTEs. Because of differing definitions of the boundaries of MTEs and the related biodiversity hotspots, data on species diversity for vertebrate groups differ between references.

Although frosts may occur throughout much of the MTE regions, these are infrequent and relatively mild in lowland areas. Mean annual precipitation is as low as 120–250 mm in coastal areas of the MTEs and reaches to about 900 mm at the upper margins of the classic evergreen shrub zone. Higher montane areas with winter rainfall regimes where the majority of precipitation falls as winter snow are also present within the traditional boundaries of MTEs. Although typically excluded from the boundaries of MTEs, adjacent arid regions that exhibit winter rainfall regimes (e.g., the Mojave Desert in California, Atacama Desert in Chile, Succulent Karoo in South Africa, and northern Sahara Desert) have a Mediterranean-type climate with a regime of winter rainfall.

Biodiversity is particularly notable in the MTE regions for vascular plant species. Although the combined area of these five regions is little more than 2% of the land area of the Earth, MTEs are home to ~50,000 species of vascular plants (**Table 2**), 20% of the world's total of about 250,000 named species. Nowhere outside of lowland tropical rainforests are there ecosystems with higher regional diversities of species, providing a strong justification for each of these regions being designated as a global "hotspot" of evolution. Vertebrate diversity is variable, and less significant overall on a global basis, but highly diverse for many specific groups of reptiles and amphibians. Invertebrate diversity and endemism is high in some MTEs, notably bees and monkey beetles (Hopliini) in the Cape Region, and bees in the Mediterranean Basin.

A characteristic ecological feature of MTEs is the widespread dominance of evergreen shrublands dominated by species with sclerophyllous leaves (di Castri *et al.*, 1981; Cowling *et al.*, 1996; Dallman, 1998). The dense cover and biomass of these shrublands burn readily under dry summer conditions with low humidity, although with differing frequencies characterizing individual MTE regions as described below. Thus, there are morphological, ecophysiological, and phenological traits favoring postfire regeneration of these stands through resprouting and fire-stimulated reseeding. Also widely present in MTE sclerophyll shrub are adaptations (or exaptations) to tolerate low soil nutrient availability and restricted summer water availability.

Despite the general characterization of MTE regions as having dominance by evergreen shrublands, other vegetation forms are also present. Woodlands are widespread in most MTEs, particularly in areas with deeper or richer soils, or as riparian woodlands or gallery forests in wetter sites. Oak woodlands dominated by species of *Quercus* are widespread in California and the Mediterranean Basin, with both evergreen and deciduous species as dominants. These communities can take the form of closed canopy evergreen woodlands, grading

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00245-8

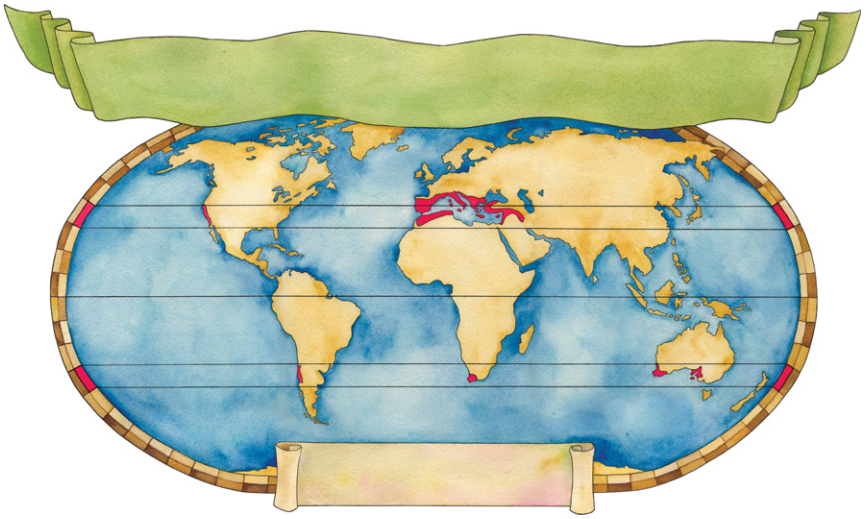


Figure 1 Global distribution of the five Mediterranean-climate regions. The position of 30 and 40° N and S latitudes are indicated by lines.

Table 1 Comparative area, topographic heterogeneity, climatic heterogeneity, and estimated natural fire frequency for the five Mediterranean-type climate regions of the world

Region	*Area (10^6 km^2)*	*Topographic heterogeneity*	*Climatic heterogeneity*	*Natural fire frequency (years)*
California	0.32	High	Very high	40–60
Central Chile	0.14	Very high	Very high	Almost fire-free
Mediterranean Basin	2.30	High	High	25–50
Cape Region	0.09	Moderate	High	10–20
SW Australia	0.36	Low	Moderate	10–15

Source: Adapted from Cowling RM, Rundel PW, Lamont BB, Arroyo MK, and Arianoutsou M (1996) Plant diversity in Mediterranean-climate regions. *Trends in Ecology and Evolution* 11: 352–360.

Table 2 Comparative species richness of native vascular plants and major vertebrate groups for the five Mediterranean-type climate regions of the world. Some caution must be used in interpreting these numbers because of varying contexts of what constitutes core Mediterranean-climate habitats

	Vascular plant species	*Mammal species*	*Bird species*	*Reptile species*	*Amphibian species*	*Fish species*
California[a]	4976	160	350	54	69	73
Chile[b]	3529	64	226	87	43	216
Mediterranean Basin	22,500	224	497	228	77	43
Cape Region	9000	127	324	100	51	34
Southwestern Australia	5571	57	285	177	33	20

[a]Political unit.
[b]"Hotspot" Chile.
Note: Vertebrate data from the Conservation International "hotspot" descriptions of the five MTEs (Mittermeier *et al.*, 2005).

into shrublands as in live oak woodlands of Southern California and the maquis of Europe, or open savannas of deciduous oaks that are widespread in both regions. Central Chile once had widespread dry- and wet-sclerophyll woodlands, but the dominance of these has been dramatically reduced by human activities. Evergreen woodlands dominated by *Eucalyptus* are widespread in Western Australia. Only the Cape Region of South Africa of all of the MTEs is largely lacking in woodlands, with such communities restricted to scattered stands of Afro-montane forest and subtropical thicket along the southern coast in fire-protected areas lacking in strong seasonal drought.

The evergreen, sclerophyll shrublands of California and Chile commonly grade off along drier coastal margins and at their arid interior margin to a plant community dominated by drought deciduous shrubs. This community, which may also dominate early successional disturbance sites or arid microsites in evergreen shrublands, is termed sage scrub in California and coastal matorral in Chile. Structurally similar communities with mixed dominance of low evergreen (and more rarely deciduous shrubs) are called phrygana in Greece and batha in the eastern Mediterranean Basin. Deciduous shrubs are largely lacking, however, from similar habitats in the Cape Region of South Africa and southwestern Australia.

Natural Disturbance Regimes

Although the five Mediterranean-climate regions share many aspects of natural climatic and environmental disturbance regimes, there are significant differences as well (Groves and di Castri, 1991; Rundel *et al.*, 1998; Rundel, 1999). While all five regions experience characteristic summer drought, for example, the magnitude of this drought is particularly severe in California and Chile, where 6–8 months or more may pass without measurable rainfall. Extreme drought such as this is rare in South Africa and southwestern Australia where summer months frequently have light showers and relatively high humidity.

The dense canopies of evergreen sclerophyllous shrubs make fire an important component of natural disturbance regimes. Natural fires, however, are significant ecological events in consuming aboveground vegetation in four of the five MTEs. Chile, where natural fires are a rare event, is the exception among MTEs. The Mediterranean-climate region of central Chile is protected from summer convective storms and lightning moving westward across Argentina by the high Andean Cordillera. The native flora of Chile shows little evidence that fire has been an important ecological disturbance regime in the evolution of life history characteristics.

Natural fire frequencies are quite different among the other four Mediterranean-climate regions (**Table 1**). In South Africa, for example, fynbos vegetation in the Cape Region commonly burns at intervals of 10–15 years, while in California natural frequencies are thought to be 30–50 years or more.

There are other strong environmental differences in addition to potential drought stress, which make the Cape Region of South Africa and southwestern Australia distinct from the other three MTEs. These two areas lie in geologically ancient and stable landscapes, resulting in highly leached and nutrient-poor soils. In contrast, earthquakes, volcanic activity, orogenic uplift, and other dynamic processes create natural disturbance regimes in California, Chile, and the Mediterranean Basin. Unlike South Africa and southwestern Australia, the younger landscapes of these three regions have experienced tremendous changes in climate regime and landscape structure in Quaternary and Holocene times and these changes have had profound impacts on community structure and speciation. The remarkable patterns of speciation in fire-sensitive shrub lineages in the Cape Region of South Africa and southwestern Australia likely have resulted from a combination of relatively mild and stable Cenozoic climatic conditions coupled with high fire frequencies in these nutrient-poor habitats.

Patterns of Speciation

Although there has been widespread speculation for many years on the causal factors responsible for the high levels of vascular plant biodiversity in MTEs, few clear generalizations have emerged. What is known is that there are conditions promoting the coexistence of seemingly ecologically equivalent shrubs, geophytes, and graminoids (Restionaceae) in frequently burned shrublands on nutrient-poor soils of fynbos and kwongan communities. Slow growth rates and diverse strategies of post-fire regeneration and reestablishment appear to promote this coexistence (di Castri and Mooney, 1973; Kruger *et al.*, 1983; Davis and Richardson, 1995; Cowling *et al.*, 1996). MTEs on less-infertile soils and with longer intervals between fires (chaparral, matorral, maquis) have lower diversity as short-lived species are excluded by rapidly growing shrub dominants. Open woodlands in the Mediterranean Basin and California, however, may have very high local diversity of annuals and short-lived perennials, where grazing maintains open habitat for species establishment.

Regional topographic and climatic heterogeneity *per se*, which might be thought to be logical correlates of comparative diversity among the five MTEs, is a poor predictor of diversity (**Table 1**). Instead, natural selection has operated to allow for a fine-scale discrimination of habitats and niches under the selective pressures of stable climates, predictably frequent fires, and periodic drought that promote community turnover and diversification. Thus, fynbos in the southwestern Cape Region and kwongan communities in southwestern Australia have evolved species-rich landscapes in topographically homogeneous areas through successive events of rapid speciation of lineages coupled with low extinction rates. High predictability of interannual rainfall and relatively limited summer drought periods may be an important factor in this low level of extinction (Cowling *et al.*, 2005).

The different speciation histories of the MTEs are reflected in the phylogenetic structure of their floras. In the Cape and southwestern Australia, where environments have been relatively stable throughout large portions of the Tertiary, radiations of endemic lineages have occurred throughout the Neogene; in the other MTEs, most radiations are associated with the onset of Mediterranean climates in the late Miocene and dramatic climate changes beginning in the late Pliocene (Hopper *et al.*, 2009; Linder, 2005; Sauquet *et al.*, 2009).

Biodiversity in California

Introduction

The political boundaries of the state of California cover an area of 411×10^3 km^2, but the area includes more than the core area of Mediterranean-type climate (Rundel and Gustafson, 2005). These political boundaries include winter rainfall portions of the Sonoran Desert and Mojave Desert, as well as areas of cold desert habitats east of the Sierra Nevada. The California floristic province, generally defined as the core MTE area, excludes these desert regions and adds northwestern Baja California and southern Oregon to the floristic province. Under this definition, the California floristic province covers 324×10^3 km^2. Because of the differences between the political and floristic province boundaries of California, some caution must be used in assessing figures on California biodiversity in the literature.

The geomorphic structure of California is complex and the topographic diversity within the floristic region is very high. Thus, this region covers the Coast Ranges extending north and south along the state, the broad Central Valley, the Sierra Nevada range, and the Transverse Ranges of Southern California. The Coast Ranges reach elevations as high as 2700 m,

while Mt. Whitney in the Sierra Nevada is the highest point in the continental United States at 4400 m elevation. The Transverse Ranges in Southern California have a number of peaks reaching above 3000 m. The dynamic geologic history of uplift, faulting, and tectonics has produced complex mosaics of soil structure and parent material, and produced sharp climate shifts over the Quaternary with associated glaciation in the Sierra Nevada.

The foothill regions throughout most of California are typically dominated by mosaics of chaparral shrublands and both evergreen and deciduous woodlands with oak species as the typical dominants. These areas commonly receive 400–800 mm annual rainfall. Rainfall is strongly centered on the winter months, and 6 months without rain is common. Drier areas along the coast and inland at the transition to desert environments support coastal sage scrub dominated by drought deciduous shrubs and a few species of deeply rooted evergreen sclerophylls. Mountain areas above 1500 m in Northern California and 1800 m in Southern California show a transition to montane conifer forests, subalpine forests, and alpine communities with increasing elevation. Higher rainfall areas along the central and northern coast of California support mixes of conifer and hardwood forests, extending into massive coast redwood forests along the northwestern coast. Mean annual rainfall reaches its highest levels above 2500 mm in this region.

Vascular Plant Diversity

California as a political unit contains nearly 5000 native vascular plants species, with about 3500 of these within the California Floristic Province. For the most recent state flora, the total includes 97 ferns and fern relatives, 59 gymnosperms (53 conifers), 833 monocots, and 3987 dicots. The largest family in this flora is the Asteraceae with 627 native species, followed by the Fabaceae with 297 species and the Poaceae with 251 species. The largest five genera make up more than 10% of this total and include *Carex* (131 species, Cyperaceae), *Eriogonum* (112 species, Polygonaceae), *Astragalus* (94 species, Fabaceae), *Phacelia* (93 species, Boraginaceae), and *Lupinus* (71 species, Fabaceae). All of these genera comprise largely of herbaceous perennial and annual species. Notable speciation has also occurred in two shrub lineages, *Arctostaphylos* (Ericaceae) and *Ceanothus* (Rhamnaceae) (Raven and Axelrod, 1978).

Endemism at the species level is relatively high at 61% within the California Floristic Province (Stebbins and Major, 1965). Fifty-two genera are strictly endemic to this province. If another 14 genera that extend only slightly outside of the region into Arizona or Baja California are included, then 8.1% of the genera are endemic. This high level of endemism is heavily influenced by the diversity of annual plants that comprise 27.4% of the vascular plants of the California floristic province. For annual species alone, endemism is 65.3%.

The highest species richness of MTEs in California appears to occur in lightly disturbed grasslands and oak woodlands, where 47–64 species have been reported in 0.1 ha sites. These levels of diversity at this 0.1 ha scale are also matched in postfire stands of chaparral where annual plant diversity is very high. Mature chaparral, however, often exhibits low levels of species diversity.

Vertebrate Diversity

The terrestrial mammal fauna of California includes about 160 species, with rodents making up more than half of this total. Among these, however, are 30 species restricted to the desert regions of the state and thus not part of the strict Mediterranean-climate region. That leaves a total of about 130 terrestrial mammals native to the shrubland, grassland, woodland, and forest regions of California. An additional five species which once occurred in the state have been extirpated in historical times. These include the grizzly bear, wolf, bison, jaguar (only an occasional visitor in the past), and giant deer mouse. Turnover between habitats (beta diversity) accounts for most of the diversity of mammal faunas, with alpha and gamma diversity relatively low.

Resident, breeding, and migrant bird diversity in California is about 350 species. Shore and marine birds make up 39% of this total. Passerines form the largest group of birds with 41% of the total. There are 21 species of hawks, vultures, and eagles, 13 species of owls, and 12 species of woodpeckers and flickers. Two bird species are endemic to California. These are the yellow-billed magpie and endangered California condor. Focusing on passerine birds, the alpha diversity of bird species across landscape gradients peaked in closed woodland and forest habitats, while species turnover between habitats (beta diversity) is greatest in mid-elevation chaparral.

There are 54 species of amphibians and 69 species of reptiles within the political boundaries of California. The salamander fauna is especially notable with 36 species, 24 of which are endemic. The reptiles include two turtles and tortoises, 33 lizards, and 33 snake species. However, the desert areas of the state are the habitats for 38 reptiles and amphibian species, thereby reducing the diversity of the California floristic region.

The native inland fish fauna of California includes 73 native species.

Biodiversity in Chile

Introduction

The political boundaries of Chile provide relatively natural biogeographic boundaries because of the topography of desert, mountain, and ocean boundaries. To the north is the hyperarid Atacama Desert, which reaches its extreme of conditions in the Tacna-Arica region along the Peruvian border. Here, a virtual absence of rainfall separates the Peruvian floristic elements of the desert from the Chilean elements. To the east, the Mediterranean-climate region of central Chile is strongly delineated by the high Andean Cordillera. Many peaks in this range reach well above 6000 m and effectively shield Chile from weather fronts moving westward across Argentina. Although major uplift of the Andes began in the

mid-Tertiary, at least 14 Ma, the range continues to be tectonically active today. The elevation of mountain passes along the Andes in northern and central Chile are too high to allow easy migration of either plants or animals, and thus have helped to isolate the flora and fauna of Chile. Only in southern Chile where the Andes are lower has there been easy migration across this range. However, the severe climatic conditions of cold that characterize Patagonia in southern Chile act to strongly reduce the biological diversity in this area.

Comparisons of species diversity between Chilean organisms and those of other Mediterranean-climate regions deserve some caution in terms of the area included. As with California, the political boundaries of Chile include desert and wet forest ecosystems that are not comparable with core Mediterranean-climate habitats. Most figures for Chilean diversity in the literature are based on political boundaries. An additional issue of political boundaries comes in assessing levels of endemism for the Chilean flora or fauna. Levels of endemism for all groups are much lower if strict adherence to the political boundaries of Chile rather than the more natural boundaries of the Chilean/Patagonian biogeographic province is taken into account.

Much of central Chile, which covers an area of about 140×10^3 km^2, shares a physiographic structure parallel to that of California. Moving inland from the Pacific Ocean, there is typically a coastal range of mountains, a broad central valley, and a high mountain range to the east. The geologically recent Cordillera de la Costa is relatively high west of Santiago at about 33° S latitude, with the major peaks Cerro Campana (1910 m), Campanita (1510 m), and El Roble (2220 m). These peaks are high enough to intercept moisture from humid southwestern winds, producing woodland areas with significant fog interception and thus improved water relations.

The dominant vegetation in central Chile is matorral, an evergreen shrubland similar in general form to chaparral (Rundel, 1981). Along the coast and to the north this community grades into a coastal matorral with a greater dominance by drought deciduous shrubs. At higher elevations and on sites with greater water availability, matorral grades into a sclerophyll woodland community, and to the south into hygrophilous woodlands with many species characteristic of the Valdivian forest region to the south. Much of the central valley of Chile today is dominated by a savanna community termed espinal, with *Acacia caven* as the sole dominant. This community is almost certainly the result of human intervention on landscape processes over the past four centuries. Sclerophyll woodland and matorral would once have covered much of this area.

Unlike California and the Mediterranean Basin, the high mountains of the Andean Cordillera in central Chile do not have a forest zone except in the southern margin of the central region. The young age of these mountains and lack of soil weathering has produced unstable geological conditions on the west-facing slope of the Andes. Matorral communities on the lower foothills of the Andes give way to a low and scrubby montane matorral community at about 2000 m elevation.

The biological diversity of Chile has an ancient origin that dates back to Gondwanaland. Southern Chile in particular exhibits many broad biogeographical linkages with New Zealand and Australia. Central Chile shows other biogeographical connections with southeastern Brazil, a linkage dating back to mid-Tertiary times before the uplift of the Andes. Since the Andean uplift, however, the biota of Chile has evolved largely in isolation from other biogeographic regions.

Biosystematic and biogeographic knowledge of the flora and fauna of Chile has increased greatly in recent decades, and thus the biodiversity of most groups of vascular plants and vertebrates is relatively well known (Simonetti *et al.*, 1995). These syntheses have been applied more generally to the country as a whole, however, rather than to discrete regional areas. Thus, there is a strong need for more regional studies that evaluate patterns of alpha, beta, and gamma diversity in relation to environmental gradients.

Vascular Plant Diversity

The age and evolutionary isolation of the Chilean flora is clearly indicated by the large number of families that are largely endemic to the Chilean floristic region, which includes adjacent areas of austral forest in southern Argentina. Endemic families entirely restricted to Chile are the Aextoxicaceae, Gomortegaceae, and Lactoridaceae and Thyrsopteridaceae (restricted to the Juan Fernandez Islands). Other families that are largely Chilean in distribution but cross political boundaries into desert areas of Peru or austral forests of southern Argentina are the Malesherbiaceae, Mizodendraceae, and Francoaceae. Of these endemic families, the monotypic Aextoxicaceae and Gomortegaceae have distributions centered in the Mediterranean-climate regions of central Chile, while the Malesherbiaceae is centered in the coastal deserts and adjacent arid montane regions of northern Chile and southern Peru.

For continental Chile as a political unit, the vascular plant flora consists of ~4600 native species, divided into about 850 native genera and 180 families. This flora includes 124 species of ferns and fern relatives, 13 species of gymnosperms, and about 4500 species of Angiosperms (Marticorena and Quezada, 1985). The flora of central Chile, excluding the desert areas north of La Serena, the Juan Fernandez Islands, and the forest and moorland areas south of Concepción, is estimated to be about 2400 species. The current concept of the biodiversity "hotspot" of central Chile has been recently expanded well beyond the Mediterranean-climate regions to include the Atacama Desert, Juan Fernandez Islands, and the Valdivian forest region. Thus, under this greatly expanded definition, the flora includes 3539 species.

Endemism is high at the generic level in the Chilean flora. Extrapolating from the literature, 16% of Chilean genera are strict endemics restricted to the political boundaries of the country. Another 17% of the genera are endemic to austral regions of Chile and adjacent areas of Argentina. Together, then a third of the genera are endemic to the Chilean-Patagonian floristic province. There are 22% of the genera with broad South American patterns of distribution and 10% of the genera have Gondwanaland origins with extant species in New Zealand and/or Australia.

The levels of species-level endemism within groups of the Chilean flora are high, but the values reported vary somewhat

depending on whether or not the author is quoting distributions within the strict political boundaries of Chile. A more natural view of endemism would include Andean species or austral forest and moorland species that occur within communities of the Chilean-Patagonian biogeographic province that may extend into Argentina. Using the political boundaries of Chile, one authority estimated that 62% of the vascular plant species were endemic, while others estimate about 50–55%.

A rich flora high in endemism occurs on the oceanic Juan Fernandez Islands, which lie about 500 km off the central Chilean coast and contain floristic elements from both the mainland of Chile and Polynesia. The flora includes 361 species of vascular plants, of which 60% are endemic.

The 20 largest genera of Chilean vascular plants make up about 30% of the flora. This group is led by *Senecio* (Asteraceae) with 218 species, *Adesmia* (Fabaceae) with 140 species, *Oxalis* (Oxalidaceae) with 111 species, *Calceolaria* (Scrophulariaceae) with 86 species and *Calandrinia* sensu latu (Portulacaceae) with 70 species. As might be expected, the degree of endemism within the largest families is higher than that for the flora overall. Eleven of the 20 largest families have 80% or more of their species endemic to Chile. These numbers on species diversity and endemism will no doubt change to some extent as these large and difficult genera become better studied. One of the most charismatic of endemic species is the Chilean wine palm, *Jubaea chilensis*. This large palm was once widespread through central Chile but is much restricted in distribution today.

Vertebrate Diversity

Mammal diversity in Chile, as in other parts of temperate South America, is low. It has been suggested that severe climatic conditions associated with Pleistocene glacial movements in the Andes may have had a strong impact in reducing the diversity of temperate mammal faunas in South America. Major faunal extinctions were also present in South America at the end of the Pleistocene, as in North America, and these extinctions have been associated with the arrival of early man. Many ecological niches in temperate South America appear to be incompletely occupied by mammals in comparison with North America.

Chile has a mammal fauna of 99 native terrestrial species, with 64 species occurring within the hotspot region. Five genera are endemic. The largest single group is the Rodentia, and comprises 60% of this total. Next in abundance are the Carnivora with 14%, and the Chiroptera (bats) with 10%. Large mammal species are particularly low in number. These include species of felids (puma, Geoffrey's cat, and colo colo), three species of canids (all fox species of *Pseudolopex*), four camelids (guanaco, vicuna, and the domesticated llama and alpaca), and three cervids (huemul, northern huemul, and pudu). Mammal faunas of Chile have been placed into five biogeographic groupings: the summer rainfall Altiplano region of northeastern Chile, the Atacama Desert and adjacent winter rainfall Andes of northern Chile, the Andes of central Chile, the Mediterranean-climate region of central Chile, the austral forests of southern Chile, and the Patagonian region. The monito del monte (*Dromiciops gliroides*) forms an endemic family, the Microbiotheriidae. In addition to these native mammal faunas, there are 15 species of terrestrial mammals that have become naturalized in Chile. Five of these occur only on the Juan Fernandez Islands, where they have had an extreme impact on the structure and composition of the native flora.

The total bird diversity of Chile is only moderate. Including the oceanic islands of Pascua and Juan Fernandez and the Antarctic territory claimed by Chile, there are reports of 451 native species and five introduced species for Chile. The hotspot area includes 226 species, with 12 of these being endemic. The biogeographical isolation of the Chilean bird fauna shows parallels with that of the mammal fauna in the manner in which niches have been filled in unusual manners. Corvids, for example, are absent from Chile. Their role as scavengers has been filled by caracaras.

There are 87 native species of terrestrial reptiles in Chile, divided among seven families and 18 genera. These include six snakes (four genera) and 81 lizards (14 genera). For the hotspot boundaries, there are 41 reptile species, with 27 of these being endemic. The diverse lizard fauna of Chile is strongly dominated by the family Tropiuridae, in particular the iguanid genus *Liolaemus*. This diverse genus and its evolution have been the focus of a large number of ecological and evolutionary studies. The occurrence of distinctive populations within individual species has led to the designation of a large number of subspecies. There are two endemic species of snakes.

The highest diversity of reptiles in Chile occurs in the northern desert and Andean areas, outside of the core Mediterranean-climate region. This region is also the site of the greatest degree of endemism.

The amphibian fauna of Chile exhibits an unusually high level of endemism, in comparison with other vertebrate groups. There are 43 native species in the hotspot of Chile, with five of the 12 genera endemic. The highest diversity of these amphibians is located in the forested regions of south-central Chile. Notable for their absence are salamanders, with only frogs and toad species present. More than three quarters of the native frogs and toads are endemic to Chile. Many of these endemic species are quite rare and localized in distribution.

The native fish fauna for the Chilean hotspot region includes 43 species, including two endemic families.

Biodiversity in the Mediterranean Basin

Introduction

The Mediterranean Basin represents the largest area of MTEs in the world, covering a complex landscape with a large amount of topographic and climatic heterogeneity. The total area of the region is $\sim 2.3 \times 10^6$ km^2, nearly 10 times greater in size than any other MTE region. This area includes more than 20 nations arrayed on both sides of the Mediterranean Sea. While coastal areas are extensive because of the large archipelagos and islands within the Mediterranean, much of this area consists of mountainous terrain with many areas above 2000 m elevation and peaks reaching as high as 4500 m.

The geographic position of the Mediterranean Basin is also an important factor in understanding the biodiversity of this region (Blondel *et al.*, 2010). Lying at the juncture of three continental landmasses, it holds a geologic history with dynamic changes associated with plate tectonics, mountain uplift, and active vulcanism. Strong climatic shifts that took place during the Plio-Pleistocene period, most notably major glacial episodes, resulted in a telescoping of many communities into the Mediterranean Basin and provided for opportunities for geographic isolation and speciation. In contrast to other MTEs, the great majority of the Mediterranean Basin is underlain by limestone. Local areas of volcanic or siliceous parent material are also present .

The climatic features of the Mediterranean Basin are often used to define this region, but the range of dominant and widespread woody species such as holm oak (*Quercus ilex*) and olive (*Olea europea*) are also used as bioindicators of the region. To the north, the Mediterranean-climate region grades into more mesic regions with summer or year-round patterns of rainfall. To the south, the Mediterranean region intergrades with the winter rainfall desert of the northern Sahara. Climates of the Mediterranean Basin are notable for their high interannual variation in both rainfall and temperature extremes.

The large area of the Mediterranean Basin, coupled with its topographic and climatic heterogeneity, makes for complex assemblages of vegetation types. There are extensive woodlands dominated by both evergreen and deciduous species of oak and evergreen sclerophyllous shrublands of many forms. These shrublands are often differentiated into types depending on the height of the vegetation. Tall sclerophyllous shrublands that may include small evergreen trees are termed maquis. Several species of Mediterranean pine may be present in this community. A middle height shrubland, generally occurring on calcareous substrates, is termed garrigue. Finally, low semiarid evergreen shrublands in the eastern Mediterranean Basin are commonly termed phrygana in Greece and batha in Israel. A long history of human impacts on the natural landscape has strongly impacted community structure and diversity.

Vascular Plant Diversity

The vascular plant flora of the Mediterranean Basin is estimated to include about 25,000 species (22,500 native species), making this region the richest among MTEs in total plant diversity. By comparison, the remaining portions of Europe with non-Mediterranean-type climate regimes cover four times as much area but have only about 6000 vascular plant species. This large flora in the Mediterranean Basin is a broad mixture of species with disparate evolutionary histories and biogeographic origins (Medail and Quezel, 1997).

One group of species evolved under subtropical conditions that existed in this region prior to the Quaternary. These include such woody plant genera as *Arbutus* and *Calluna* (Ericaceae), *Ceratonia* (Fabaceae), *Chamaerops* (Arecaceae), and *Laurus* (Lauraceae). Another group of taxa represents neo-Mediterranean elements that migrated into the Mediterranean Basin after the establishment of a Mediterranean-type climate. Examples of woody genera in this group include *Amelanchier* (Rosaceae), *Clematis* (Ranunculaceae), and *Cistus*, *Halimium*, and *Helianthemum* (Cistaceae). Three groups of nontropical woody elements that evolved after the onset of Mediterranean-type climates have been identified. These groups are a Mediterranean element evolved *in situ* in mountain areas and high in endemism, a desert and cold steppe group of species entering from Africa and the Middle East, and a Holarctic element of species with Eurasian temperate affinities. Endemism is high at the species level in the Mediterranean Basin, with a level of about 50%.

Species richness at the scale of 0.1 ha is often remarkably high in lightly grazed or disturbed Mediterranean woodlands and shrub grasslands, with as many as 119–179 species in such stands.

Vertebrate Diversity

Vertebrate faunas of the Mediterranean Basin share the characteristic of multiple biogeographic origins with vascular plants. Dramatic climatic oscillations during the Pleistocene led to periodic turnovers of Eurasian and African faunal elements and a resulting isolation of populations.

The present fauna of land mammal species for the Mediterranean Basin number about 224 species of which 25 are endemic. Because of the biogeographic barriers of the Mediterranean Sea and the Saharan Desert, mammal faunas of Mediterranean Europe, the Middle East and North Africa are somewhat distinct. For North Africa, the mammal fauna of the Mediterranean Basin shows its strongest affinities with tropical Africa. A decline in species richness occurred at the end of the Pleistocene and into the Holocene with a combination of sharp climate shifts and human pressures through hunting.

Bird diversity of the Mediterranean Basin includes about 497 regularly occurring species. In contrast to mammals, the affinities of bird faunas are more strongly linked to the Asiatic steppes than to tropical Africa. The evolution of these elements of bird faunas can be linked to Eurasian (153 species) and Eremian semiarid habitats (85 species), where Plio-Pleistocene conditions led to ongoing isolation and speciation. Forest birds of boreal origin are widespread and dominant throughout both Middle Europe and the Mediterranean Basin. Shrubland bird species characteristic of the region represent only about 12% of the total. There are 32 endemic bird species in the Mediterranean Basin.

Reptiles and amphibians of the Mediterranean Basin include 228 and 77 species, respectively, and show distinct Holarctic affinities. Much of the endemism within these groups appears to represent archaic lineages that differentiated during the Middle Tertiary. Reptile diversity is highest in the eastern Mediterranean Basin and drops steadily moving westward. Species diversity on Mediterranean islands is relatively low. Important reptile groups include lizards of the Lacertidae with 63 species (23% of the world total), snakes of the Viviperidae with 19 species (8% of the world total), and tortoises of the Testunididae with four species. Overall, 34% of Mediterranean Basin reptiles are endemic to this region. For amphibian diversity, the pattern is reversed compared with

reptiles as the highest levels of diversity are found in the Euro-Mediterranean areas compared with the North African and Middle Eastern portions of the region. Notable groups of amphibians include the Discoglossidae with 10 species (71% of the world total) and the Salamandridae with 19 species (36% of the world total). Endemism for amphibians in the Mediterranean Basin is 31%.

Fish diversity is relatively high in the Mediterranean Basin compared to the other MTEs, with 216 species. This group includes 63 endemic species, six endemic genera, and one endemic family.

Biodiversity in the Cape Region of South Africa

Introduction

The Cape Region forms a small area on the southwestern tip of the African continent. It is renowned for its showy and diverse flora unlike that of any other area of the world. The landscape of the Cape Region is dominated by steep and strongly folded mountain ranges of quarzitic sandstone that predate the separation of the African continent from Gondwanaland. These once high mountains have been eroded over the past 200 My to form low ranges capped by resistant Table Mountain sandstone. Separating the mountains are gentle valleys and undulating plains that are largely underlain by shales with greater nutrient availability. Relatively young Tertiary and Quaternary limestones and sands mantle extensive areas of the coast.

The characteristic vegetation of the Cape Region, particularly on the nutrient-poor quartzite soils, is fynbos (Cowling, 1992; Cowling and Richardson, 1995). Fynbos is an evergreen shrubland dominated by four major plant morphological groups. These include two shrub groups (the proteoids and ericoids), a sedge-like group (restioids), and geophytes. The proteoids, formed by woody Proteaceae, form the tallest matrix of the fynbos community and commonly reach to 2–4 m in height. The ericoid group gains its name from the Ericaceae, but includes more than 3000 species of small-leaved shrubs representing many families. The restioids are members of the Restionaceae, a family with origins in Gondwanaland but its primary diversification in the fynbos. Finally, the Cape Region contains the highest diversity of bulbs and other geophytes in the world, with more than 1500 species. Many types of fynbos have been described, but a simple classification scheme includes proteoid fynbos, ericaceous fynbos, restioid fynbos, asteraceous fynbos, and grassy fynbos (Manning, 2007).

Another important shrubland vegetation type of the Cape Region is renosterveld, a low shrubland occurring on richer soils with shale parent material. It is floristically differentiated from fynbos by the absence of restioids and minor importance of proteoids. This community once covered more than a quarter of the Cape Region, but has now largely been cleared for agriculture and urban expansion.

Woodland and forest communities are surprisingly rare in the Cape Region. True forests occupy only about 3850 km^2 of moist sites (800–1200 mm annual rainfall) protected from fire along the southern coast. These forests are low in diversity and represent depauperate outliers of Afro-montane forests of tropical East Africa Subtropical thicket, a low, evergreen formation with floristic affinities to Afro-montane forest, occurs throughout the Cape Region in fire-free sites. This community is richer in species and endemics – especially succulents – than the moister Afro-montane forests (Vlok *et al.*, 2003). It appears that fire and not climate is the major determinant of the occurrence of these communities. The quasi-Mediterranean south coastal region has markedly fewer occurrences of high fire danger days.

Vascular Plant Diversity

The Cape Region contains what is arguably the most distinctive and diverse floras of any temperate area of the world (Goldblatt and Manning, 2002). Its status as a distinct floral kingdom is reinforced by the presence of five endemic families – Geissolomataceae, Grubbiaceae, Penaeaceae, Roridulaceae, and Stilbaceae. Moreover, there are 193 endemic genera, which is 19.5% of the total.

Covering an area of only 78,500 km^2, the Cape Region contains ~9000 species of vascular plants. The 10 largest genera of the Cape Region account for over 20% of the flora, led by *Erica* (Ericaceae, 658 species) and *Aspalathus* (Fabaceae, 257 species). Other large genera are *Phylica* (Rhamnaceae), *Agathosma* (Rutaceae), *Oxalis* (Oxalidaceae), *Pelargonium* (Geraniaceae), *Senecio* (Asteraceae), *Cliffortia* (Rosaceae), *Muraltia* (Polygonaceae), and *Ruschia* (Aizoaceae-Mesembryanthema) (Goldblatt and Manning, 2000).

Levels of species endemism in the Cape floristic region are among the highest in the world. For the entire region, endemism at the species level is about 69%. The high levels of endemism present in the Cape Region are largely due to the presence of neoendemics rather than paleoendemic species, although many neoendemic lineages begin diversifying in the mid Cenozoic. This dominance of neoendemics is indicated by the predominance of endemic diversity in large, species-rich genera, the widespread presence of sympatric congeners, and the edaphic specialization of many endemics on geologically young substrates. Rather than being a random ecological or phylogenetic assemblage of species, the great majority of endemics are low shrubs killed by fire and dependent on closely dispersed seeds for regeneration. Four families are notably rich in endemics – the Proteaceae, Ericaceae, Rutaceae, and Polygalaceae. Species richness is greatest in the southwestern Cape Region centered around Cape Town. The Cape Peninsula, for example, supports 2256 species (including 90 endemics) in an area of 471 km^2. Cape Hangklip on the eastern shore of False Bay near Cape Town has 1383 species in 240 km^2.

Smaller regional centers of endemism exist within the Cape Region. Dividing the Cape Region into five floristic zones on the basis of species distributions within seven large families, regional levels of endemism are highest in the southwestern and northwestern Cape (about 50%) and lowest in the Eastern Cape and Inland Mountain regions with nonseasonal rainfall (18–28%). These patterns of regional endemism have been further demonstrated in studies of distribution of the Proteaceae in the Cape Region. For the entire Cape Region, 99.4% of

the 330 species of Proteaceae are endemic. At a regional level, 63% of the Proteaceae in the southwestern region are endemic to that region, compared with only 19% endemism for the Coastal Mountain and southeastern regions. Point endemism is also widespread involving species that are restricted to highly specific edaphic habitats.

Fynbos species diversity is also extremely high at the alpha-diversity level of small stands. Typical fynbos communities support a mean of about 65 vascular plant species in 0.1 ha, with a range of 31–126 species. Renosterveld shrublands have even higher diversities with a mean of 84 species per 0.1 ha and a range of 28–143 species.

Vertebrate Diversity

The Cape Region lacks a distinctive mammal fauna. This region today contains 127 species of native mammals, with 90 being present in the Southwest Cape area. The regional total is less than half of the mammal species occurring within all of South Africa. The largest orders present are the Rodentia and the Carnivora. The rodents are represented by two species of mole rats (Bathyergidae), a porcupine (Hystricidae), two dormice (Muscardinidae), and at least 21 species of Muridae and Cricetidae. There are 27 species of the Carnivora, ranging from mustelids and civets to larger hyenas, jackals, and cats. Large browsers and grazers play an important role in this ecosystem in comparison with other MTEs. There are 20 species of Artiodactyla and five species of Perissodactyla. Very few of these depend on grazing, however, because of the paucity and poor nutritive value of Cape Region grasses. The Chiroptera is a large group with one fruit bat and 14 species of Microchiroptera in the Southwest Cape.

The abundance and diversity of native mammals was probably always relatively low in fynbos shrublands, unlike the rich savanna regions of the continent. At the time of European colonization the highest numbers and diversity of large mammals was present in renosterveld or other open communities with better browse. Large mammals that were once common on the renosterveld plains included bontebok, eland, buffalo, Cape mountain zebra, red hartebeest, and lion.

Endemism, as might be expected, is quite low among mammals in the Cape Floristic Region as most of this fauna extends northward or westward into arid or savanna ecosystems. Only four species in the mammal fauna are endemic. The fossorial Cape dune mole rat and burrowing gerbil among the endemic rodents are associated with sandy soil substrates rather than with any specific vegetation type. The colonial behavior and feeding specialization of mole rats on bulbs may be linked to the remarkably high diversity of geophytes in the Cape Floristic Region.

The regularly occurring bird fauna includes 324 species, with a notable diversity of Falconiiformes with 22 species. Six bird species are endemic. Endemic species are largely dietary specialists such as the Cape sugarbird, orange-breasted sunbird, and protea seedeater that are tied to specific plant resources in the fynbos. The originally direct communication of fynbos habitats with semiarid and savanna shrublands has probably been a factor in limiting the number of endemic fynbos birds. The savanna region of South Africa is far richer in endemic species. At least four of the endemic fynbos birds are characteristic of montane areas or are allied to montane species of East Africa that live in ericaceous shrublands. For fynbos communities specifically, there are only about 10 reported species, with renosterveld habitats generally richer at a local level.

The Cape Region is moderately diverse in reptiles with 100 species present, 22 of which are endemic. Fynbos communities may contain more than 50 species of lizards. The Gekkonidae are the most important group with 18 species. There are 32 species of snakes reported from the Western Cape area. Among snakes, the Colubridae have the highest diversity with 25 species. There are 19 endemic species among the reptiles, 17% of the total. One notable endemic to the southwestern Cape Region is *Psammobates geometricus*, one of the rarest tortoises in the world. The life history of this species seems to have evolved to adapt to fynbos fire cycles, with hatchlings appearing in late autumn after the danger of summer fires is past.

Amphibians are relatively low in diversity in the Cape Region with 51 native species, 16 of these endemic. The largest single group of amphibians is the Ranidae with 13 species. One of the most interesting endemics among amphibians is the arum lily frog, *Hyperolius horstocki*, which lives in the flowers of the common arum lily (*Zantedeschia aethiopica*).

There are 34 native fish species in the Cape Region, 14 of these being endemic.

Biodiversity in Southwestern Australia

Introduction

Core Mediterranean-climate conditions are present in southwestern Western Australia and in South Australia around Adelaide (Hobbs, 1992). The core Southwestern Floristic Province of Western Australia is $\sim 357 \times 10^3$ km^2 in area, similar in size to the California floristic province. A broader region with transitional rainfall patterns with biseasonal distribution is often included as well, almost doubling the core MTE area. While mean annual rainfall is commonly 350–800 mm over much of southwestern Australia, it reaches as high as 1500 mm at the extreme southwestern corner of Western Australia and drops to a low of about 250 mm at the eastern edge of the Mediterranean-climate region in a transition to arid communities. Topographic heterogeneity is very low over this region, with elevations of the few mountains reaching to no more than 1000 m. Much of the region is a low, laterized plateau dissected into broad valleys with deep *in situ* weathering. Soils of southwestern Australia, as in the Cape Region of South Africa, are generally very old, highly weathered, acidic, and low in nutrient availability.

For the true Mediterranean-climate regions, the highest rainfall zones with 800–1200 mm of annual rainfall support evergreen forests and woodlands dominated by *Eucalyptus marginata* (jarrah), *E. calophylla* (marri), *E. diversicolor* (karri), and *E. gomphocephala* (tuart). Low *Banksia* woodlands and coastal heaths are also widespread. At intermediate rainfall regimes of 300–800 mm, the dominant vegetation is a mosaic of shrubby woodlands (mallees) and heathland communities termed

kwongan (Pate and Beard, 1984). The vegetation and dominant species are finely tuned to small changes in edaphic conditions that influence nutrient and water availability. Human impacts on these ecosystems have been severe over the past century.

High levels of species diversity and endemism characterize the vascular plant flora, with only moderate diversity and endemism present in most vertebrate groups (Hopper and Gioia, 2004). The high levels of vascular plant diversity has been related to the development of a complex mosaic of landforms and soils during the Tertiary and Quaternary, the geologic history of oscillating moisture regimes through the Quaternary in the absence of glaciation, isolation of southwestern Australia from the east by the arid Nullarbor Plain, and interactions of gene pools from both paleotropical and temperate assemblages.

Vascular Plant Diversity

The Australian MTE regions including southwestern Australia and South Australia include at least 9000 species of vascular plants. There are about 7400 species present within the Southwest Botanical Province of Western Australia (Hopper and Gioia, 2004). Including species that extend slightly beyond the Southwestern Botanical Province into a transitional interzone, 79% of the flora is endemic. Included are 87 endemic genera and four endemic families – the Ecdeiocoleaceae, Cephalotaceae, Emblingiaceae, and Eremosynaceae. Woody perennials in four families – the Myrtaceae, Proteaceae, Fabaceae, and Ericaceace (Epacridaceae) – dominate the flora (Marchant *et al.*, 1987; Beard *et al.*, 2000). Much of the high species diversity within these families is due to extensive radiation within a few large genera. Large genera for the region include *Acacia* (400 species), *Eucalyptus* (246 species), *Grevillea* (200+ species), *Stylidium* and *Melaleuca* (150+ species each), and *Hakea* and *Caldenia* (100+ species each).

Nodes of unusual species diversity are present along the south coast of Western Australia (Stirling Range, Fizgerald River area) and in the sandplains north of Perth (Mt. Lesueur area). As in the fynbos of South Africa, large numbers of endemics with highly local patterns of distribution also characterize kwongan. Local scale plant diversity reaches almost as high as areas in the western Cape Region of South Africa. Vascular plant diversity in a sample of 0.1 ha stands of heathland in southwest Australia exhibited a range of 43–103 species, while jarrah forests and mallee stands had a lower range of 17–55 species.

At the regional level, southwestern Australia exhibits major differences in centers of highest diversity among the most important woody genera. Some genera – as, for example, *Banksia* (Proteaceae), *Adenanthos* (Proteaceae), *Leucopogon* (Epacridaceae), and *Eucalyptus* (Myrtaceae) – are most speciose near the south coast or in southern kwongan and mallee communities. Other large genera have their highest diversity in northern kwongan – *Grevillea* (Proteaceae), *Conostylis* (Haemodoraceae), and *Lechenaultia* (Goodeniaceae). Finally, a large group of genera show bimodal patterns of diversity reflecting both the nodes of high species diversity in northern and southern kwongan – *Calothamnus* (Myrtaceae), *Melaleuca* (Myrtaceae), *Hakea* (Proteaceae), *Darwinia* (Myrtaceae), and *Dryandra* (Proteaceae). Two other large genera, Acacia (Fabaceae) and Verticordia (Myrtaceae), are most diverse in the inland transition area of rainfall.

Vertebrate Diversity

The vertebrate fauna of Australian MTEs shows no unusual level of high diversity or endemism. The majority of the fauna in these MTEs are populations of more typically arid or mesic habitat species whose ranges extend into southwestern or South Australia. Thus, the majority of the vertebrates appear to have relatively broad ecological niches rather than specialized requirements unique to the MTEs.

For mammals, the Mediterranean-climate fauna includes only 57 species, 12 of these being endemic. Interpreting the patterns of distribution of large mammals is made difficult, however, by the strong impact that both Aboriginal and European populations have had on this fauna. The extinction of a large megafauna in the Late Quaternary left the Australian continent without large grazers or predators.

Bird diversity is also relatively low, with 285 species regularly present. Ten of these are endemic. Most notable among these endemics is the black swan (*Cygnus atratus*), shown on the state emblem for Western Australia.

Diversity is high among reptiles in southwestern Australia, an evolutionary consequence in part of the low mammal diversity. There are 177 native species, with 27 of these being endemic. Amphibians include 33 native species, with 19 endemic species and four endemic genera.

With only small areas of riverine or freshwater habitat, fish diversity in southwestern Australia is low. Only 20 native species are present but half of these are endemic. These include three endemic genera and one endemic family.

Conclusions

Mediterranean-type ecosystems provide many opportunities for comparative studies of the controlling factors in the evolution of biodiversity. These five regions share common characteristics of climatic regime, with an independent evolution of plant and animal species adapted to these conditions within each region. Thus, they provide a natural ecosystem experiment with five independent replications. The value of comparative studies then between these regions lies not only with their similarities, but with subtle differences in climatic conditions, topographic diversity, evolutionary history, and human impacts that have led to the patterns of diversity that we see today.

There has been a long history of comparative ecosystem studies between California and Chile (Cody and Mooney, 1978), and between South Africa and southwestern Australia, and these have traditionally devoted some attention to comparing and contrasting how natural processes operate. The remarkable biodiversity of MTEs together with the large numbers of rare and endangered species in these regions gives a special significance to expanding studies of these regions to better understand the evolution of diversity, particularly by vascular plants. Serious threats of habitat transformation and

degradation today make it critical that there be a better understanding of the conservation biology and sustainable resource management in all five MTEs (Rundel *et al.*, 1998; Rundel, 1999).

See also: Australia, Biodiversity of Ecosystems. Fires, Ecological Effects of. Near East Ecosystems, Animal Diversity. Near East Ecosystems, Plant Diversity

References

Beard JS, Chapman AR, and Goia P (2000) Species richness and endemism in the Western Australian flora. *Journal of Biogeography* 27: 1257–1268.

Blondel J, Bodiou J-Y, Boeuf G, and Aronson J (2010) *The Mediterranean Region: Biological Diversity through Time and Space*. Oxford: Oxford University Press.

di Castri F, Goodall DW, and Specht RL (eds.) (1981) *Mediterranean-Type Shrublands*. Amsterdam: Elsevier.

di Castri F and Mooney HA (eds.) (1973) *Mediterranean-Type Ecosystems: Origin and Structure*. New York: Springer.

Cody ML and Mooney HA (1978) Convergence versus nonconvergence in Mediterranean-climate ecosystems. *Annual Review in Ecology and Systematics* 9: 265–321.

Cowling RF, Ojeda F, Lamont BB, Rundel PW, and Lechmere-Oertel R (2005) Rainfall reliability, a neglected factor in explaining convergence and divergence of plant traits in fire-prone Mediterranean-climate ecosystems. *Global Ecology Biogeography* 14: 509–519.

Cowling RM (ed.) (1992) *The ecology of fynbos: Nutrients. Fire and Diversity*, Cape Town: Oxford University Press.

Cowling RM and Richardson DM (1995) *Fynbos: South Africa's Unique Floral Kingdom*, Cape Town: Fernwood Press.

Cowling RM, Rundel PW, Lamont BB, Arroyo MK, and Arianoutsou M (1996) Plant diversity in Mediterranean-climate regions. *Trends in Ecology and Evolution* 11: 352–360.

Dallman PR (1998) *Plant Life in the World's Mediterranean Climates*. Berkeley: University of California Press.

Davis GW and Richardson DM (eds.) (1995) *Mediterranean-Type Ecosystems: The Function of Biodiversity*. Berlin: Springer.

Goldblatt P and Manning J (2000) *Cape Plants: A Conspectus of the Cape Flora of South Africa*. St. Louis: Missouri Botanical Garden.

Goldblatt P and Manning JC (2002) Plant diversity of the Cape Region of South Africa. *Annals of the Missouri Botanical Garden* 89: 281–302.

Groves RH and di Castri F (eds.) (1991) *Biogeography of Mediterranean Invasions*, 485 pp. Cambridge: Cambridge University Press.

Hobbs RJ (ed.) (1992) *Biodiversity of Mediterranean Ecosystems in Australia*, 245 pp. Australia: Surrey Beatty & Sons, Chipping Norton.

Hopper SD and Gioia P (2004) The southwest Australian floristic region: Evolution and conservation of a global hot spot of biodiversity. *Annual Review in Ecology Evolution Systematics* 35: 623–650.

Hopper SD, Smith RJ, Fay MF, Manning JC, and Chase MW (2009) Molecular phylogenetics of Haemodoraceae in the Greater Cape and Southwest Australian Floristic Regions. *Molecular Phylogenetics and Evolution.*

Kruger FJ, Mitchell DT, and Jarvis JUM (eds.) (1983) *Mediterranean-Type Ecosystems: The Role of Nutrients*. Berlin: Springer.

Linder HP (2005) Evolution of diversity: the Cape flora. *Trends in Plant Science* 10: 536–541.

Manning J (2007) *Field Guide to the Fynbos*. Cape Town: Struik Nature.

Marchant NG, Wheeler JR, Rye BL, Bennett EM, Lander NS, and MacFarlane TD (1987) *Flora of the Perth Region*. Perth: Western Australian Herbarium.

Marticorena C and Quezada M (1985) Catálogo de la flora vascular de Chile. *Gayana* 42: 1–157.

Medail F and Quezel P (1997) Hot-spots analysis for conservation of plant diversity on the Mediterranean Basin. *Annals of the Missouri Botanical Garden* 84: 112–127.

Mittermeier RA, Gil PR, Hoffman M, *et al.* (2005) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Mexico City: CEMEX. (accessed September 2006).

Pate JS and Beard JS (eds.) (1984) *Kwongan: Plant Life of the Sand Plains*. Nedlands: University of Western Australia Press.

Raven PH and Axelrod DI (1978) *Origin and Relationships of the California Flora*, vol. 72, pp. 1–134. University of California Publications in Botany.

Rundel PW (1981) The matorral zone of central Chile. In: di Castri F, Goodall DW, and Specht RL (eds.) *Mediterranean-Type Shrublands*, pp. 175–201. Amsterdam: Elsevier.

Rundel PW (1999) Disturbance in Mediterranean-climate shrublands and woodlands. In: Walker L (ed.) *Ecosystems of Disturbed Ground*, pp. 221–235. Amsterdam: Elsevier.

Rundel PW and Gustafson R (2005) *An Introduction to the Plant Life of Southern California: Coast to Foothills*, p. 300. Berkeley: University of California Press.

Rundel PW, Montenegro G, and Jaksic F (eds.) (1998) *Landscape Disturbance and Biodiversity in Mediterranean-Type Ecosystems*. Berlin: Springer.

Sauquet H, Weston PH, Lisa Anderson C, *et al.* (2009) Contrasted patterns of hyperdiversification in Mediterranean hotspots. *Proceedings of the National Academy of Sciences* 106: 221–225.

Simonetti JA, Arroyo MTK, Sportorno AE, and Lozada E (eds.) (1995) *Diversidad Biologica de Chile*. Santiago: CONICYT.

Stebbins GL and Major J (1965) Endemism and speciation in the California flora. *Ecological Monograph* 35: 1–35.

Vlok JHJ, Euston-Brown DIW, and Cowling RM (2003) Acocks' Valley Bushveld 50 years on: New perspectives on the delimitation, characterisation and origin of thicket vegetation. *South African Journal of Botany* 69: 27–51.

Megaherbivores

Norman Owen-Smith, University of the Witwatersrand, Wits, South Africa

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allometry How a biological feature is related to body mass; for example, basal metabolic rate varies with body mass raised to the power three-quarters rather than in direct proportion with body mass, so that an animal twice as large requires less than twice the amount of food to supply its metabolic needs.

Graminoids Grasses plus grass-like plants such as reeds and sedges.

Holocene The epoch in time extending from 11,700 years ago, approximately when the last ice age ended, to the present.

Home range The area typically traversed by an animal in meeting its life requirements over some period.

Keystone A species or some other entity (such as a habitat) that has a disproportionate influence on maintaining or influencing a structure, for example, the species composition and functioning of an ecosystem.

Pleistocene The epoch or period in time from approximately 2.5 million to 11,700 years ago, which included a series of glacial advancements or ice ages separated by interglacial periods.

Introduction

In this article, the label "megaherbivore" will be restricted to plant-eating animals attaining a body mass exceeding 1000 kg, that is, a metric ton or 10^6 g, as adults. The label "megafauna" has been widely used to encompass species weighing more than 100 lbs (44 kg), including herbivores and carnivores. This label originates from Martin (1966, 1967) and is somewhat arbitrary. In contrast, there are fundamental distinctions in several features of ecology manifested once a body mass of 1000 kg is surpassed. They concern life history attributes, vulnerability to predation, population growth potential, digestive capabilities, biomass densities attained, vegetation impacts, relationships with humans, and conservation issues, defining "the megaherbivore syndrome" (Owen-Smith, 1988).

Megaherbivores were once far more abundant and diverse than they are today, and occurred on all continents. Especially prominent were elephants and elephant-like beasts falling within the order Proboscidea (**Table 1**). Numerous members of the Perissodactyla (odd-toed ungulates), including rhinoceroses and their allies, also attained megaherbivore size, along with some of the largest artiodactyls (even-toed ungulates) and edentates (ground sloths). Anthracotheres, possibly ancestral to hippos, also included species within the megaherbivore size range. Several species exceeded the modern African savanna elephant in size. The massively built brontothere *Paraceratherium* (formerly known as *Baluchitherium*) attained a shoulder height of approximately 5.5 m and body mass of perhaps 20 tons, and was the largest land mammal ever.

Numerous megaherbivores survived into the late Pleistocene, from 100,000 years ago until as recently as 11,000 years ago, when the last ice age ended (**Table 2**). They included several species of mammoths and elephants together with allied gomphotheres. Europe formerly housed three species of rhinoceros. Giant ground sloths in several forms occurred within South America and southern North America, whereas a hippo-like notoungulate existed in South America. Within Australia, a giant wombat-like marsupial attained rhinoceros size.

Extant terrestrial megaherbivores are today confined to Africa and tropical Asia (**Table 3**). They include the African elephants, recently separated into two distinct species (Roholand *et al.*, 2010); the Asian elephant; four rhinoceros species in this size range; the hippopotamus; and, marginally, the giraffe. The West Indian manatee constitutes the sole marine representative. This life form is now absent in the modern faunas of North and South America, Europe, northern Asia, and Australasia.

The author's doctoral study on the behavioral ecology of the white rhinoceros brought to his attention the common ecological features shared among elephants, hippos, and rhinos, as well as the relevance of these attributes to certain contentious scientific issues. These concerned (1) the causes of the wave of large mammal extinctions that reduced megafaunal diversity through most of the world at the end of the Pleistocene and (2) the dilemmas of conserving the megaherbivores that have survived to the present time. Some paleoecologists regret the impoverishment of the North American megafauna and have advocated replicating the ecological role extinct species played by importing modern counterparts, including elephants (Donlan *et al.*, 2005, 2006; Martin, 2005). Similar proposals have been made for Siberia and Europe (Zimov *et al.*, 1995; Svenning, 2007). Are these proposals crazy, or insightful? The reader must judge for his or herself from the information about the ecology of these largest land mammals presented in this article.

Nutritional Physiology and Dietary Tolerance

Except in deserts and polar regions, vegetation accumulates annually as potential food for herbivores to produce the appearance of a "green world." Why is so little eaten? Certain ecologists proposed that this was because predators kept the abundance of herbivores well below the level that this food

Table 1 Megaherbivore genera extant during the earlier part of the Tertiary period

Order	*Family*	*Genus*	*Place*	*Period extant*
Proboscidea	Barytheriidae	*Barytherium*	Egypt	Eocene–Oligocene
Proboscidea	Palaeomastodontidae	*Palaeomastodon*	Africa	Eocene–Oligocene
Proboscidea	Palaeomastodontidae	*Phiomia*	Africa	Eocene-Oligocene
Proboscidea	Deinotheriidae	*Deinotherium*	Africa, Eurasia	Miocene–Pleistocene
Proboscidea	Deinotheriidae	*Prodeinotherium*	Africa	Miocene
Proboscidea	Mammutidae	*Hemimastodon*	Africa	Miocene
Proboscidea	Mammutidae	*Eozygodon*	Africa	Miocene
Proboscidea	Mammutidae	*Zygolophodon*	Africa, Eurasia, North America	Miocene
Proboscidea	Gomphotheriidae	*Gomphotherium*	Africa, Eurasia, North America	Miocene
Proboscidea	Gomphotheriidae	*Platybelodon*	Africa, Eurasia, North America	Miocene
Proboscidea	Gomphotheriidae	*Gnathabelodon*	North America	Miocene
Proboscidea	Gomphotheriidae	*Amebelodon*	Africa, Asia, North America	Miocene
Proboscidea	Gomphotheriidae	*Anancus*	Africa, Eurasia	Miocene–Pleistocene
Proboscidea	Gomphotheriidae	*Zyglolophodon*	Africa, Eurasia, North America	Miocene
Proboscidea	Gomphotheriidae	*Sinomastodon*	Asia	Miocene-Pleistocene
Proboscidea	Gomphotheriidae	*Rhynchotherium*	North America	Miocene
Proboscidea	Gomphotheriidae	*Notiomastodon*	South America	Pleistocene
Proboscidea	Gomphotheriidae	*Tetralophodon*	Africa, Eurasia	Miocene–Pliocene
Proboscidea	Stegodontidae	*Stegolophodon*	Asia	Miocene
Proboscidea	Stegodontidae	*Stegotetrabelodon*	Africa	Miocene
Proboscidea	Stegodontidae	*Stegodibelodon*	Africa	Miocene
Proboscidea	Elephantidae	*Primelephas*	Africa	Miocene–Pliocene
Perissodactyla	Hyracodontidae	*Urtinotherium*	Asia	Oligocene
Perissodactyla	Hyracodontidae	*Paraceratherium*	Eurasia	Eocene–Oligocene
Perissodactyla	Brontotheriidae	*Brontotherium*	North America	Eocene
Perissodactyla	Brontotheriidae	*Megacerops*	North America	Eocene
Perissodactyla	Brontotheriidae	*Dianotitan*	China	Eocene
Perissodactyla	Brontotheriidae	*Notiotitanops*	North America	Eocene
Perissodactyla	Brontotheriidae	*Embolotherium*	Mongolia	Eocene
Perissodactyla	Brontotheriidae	*Aktautitan*	Kazakstan	Eocene
Perissodactyla	Brontotheriidae	*Gnathotitan*	Mongolia	Eocene
Perissodactyla	Brontotheriidae	*Rhinotitan*	Mongolia	Eocene
Perissodactyla	Brontotheriidae	*Diplacodon*	North America	Eocene
Perissodactyla	Brontotheriidae	*Pachytitan*	Mongolia	Eocene
Perissodactyla	Brontotheriidae	*Hyotitan*	Mongolia	Eocene
Perissodactyla	Brontotheriidae	*Brachydiastematherium*	North America, Europe	Eocene
Perissodactyla	Brontotheriidae	*Epimanteoceras*	Mongolia	Eocene
Perissodactyla	Rhinocerotidae	*Teleoceras*	North America	Miocene–Pliocene
Perissodactyla	Rhinocerotidae	*Elasmotherium*	Asia	Pliocene–Pleistocene
Perissodactyla	Rhinocerotidae	*Diceratherium*	North America	Oligocene
Perissodactyla	Rhinocerotidae	*Aceratherium*	Africa, Asia	Oligocene–Pliocene
Perissodactyla	Rhinocerotidae	*Brachydiceratherium*	Europe	Oligocene–Miocene
Perissodactyla	Rhinocerotidae	*Diaceratherium*	Europe	Miocene
Perissodactyla	Rhinocerotidae	*Brachypotherium*	North America, Europe, Africa	Miocene–Pliocene
Perissodactyla	Rhinocerotidae	*Aphelops*	North America	Miocene-Pliocene
Perissodactyla	Rhinocerotidae	*Peraceras*	North America	Miocene
Perissodactyla	Rhinocerotidae	*Chilotherium*	Eurasia	Miocene–Pliocene
Perissodactyla	Rhinocerotidae	*Floridaceras*	North America	Miocene
Artiodactyla	Anthracotherriidae	*Libycosaurus*	Africa	Miocene
Artiodactyla	Anthracotherriidae	*Merycopotamus*	Asia	Miocene–Pliocene
Artiodactyla	Giraffidae	*Libytherium*	Africa	Pliocene

could support (Hairston *et al.*, 1960). However, much of what appears green is loaded with indigestible fiber and nasty chemicals, so that only a tiny fraction is actually edible (Janzen, 1979). Moreover, during the winter or dry season, most plants cease growth and many shed their leaves, so that very little edible material remains available toward the end of the adverse season (Sinclair, 1975). Herbivores must cope with food that appears abundant seasonally, but is mostly of low quality nutritionally, especially in the adverse season.

Larger body size brings one fundamental benefit for herbivores. The metabolic rate of animals increases as a function of body mass raised to the power three-quarters, rather than in direct proportion with body size. This allometric relationship means that the specific metabolic requirement for

Table 2 Megaherbivores becoming extinct during the late Pleistocene

Order	*Family*	*Species*	*Location*	*Last recorded (years BP)*
Proboscidea	Elephantidae	*Elephas iolensis*	Africa	34,000
Proboscidea	Elephantidae	*Palaeoloxodon antiquus*	Europe, Asia	40,000
Proboscidea	Elephantidae	*Mammuthus primigenius*	Eurasia, North America	10,000
Proboscidea	Elephantidae	*Mammuthus columbi*	North America	10,500
Proboscidea	Stegodontidae	*Stegodon* spp	Asia	11,000
Proboscidea	Mammutidae	*Mammut americanum*	North America	10,500
Proboscidea	Gomphotheriidae	*Stegomastodon platensis*	South America	11,000
Proboscidea	Gomphotheriidae	*Stegomastodon waringi*	South America	11,000
Proboscidea	Gomphotheriidae	*Cuvieronius hyodon*	South America	11,900
Perissodactyla	Rhinocerotidae	*Dicerorhinus hemitoechus*	Southern Europe	24,000
Perissodactyla	Rhinocerotidae	*Dicerorhinus kirchbergensis*	Southern Europe	24,000
Perissodactyla	Rhinocerotidae	*Coelodonta antiquitatis*	Eurasia	10,700
Artiodactyla	Hippopotamidae	*Hippopotamus gorgops*	Africa	130,000
Artiodactyla	Giraffidae	*Sivatherium maurusium*	Africa	8000
Notoungulata	Toxodontidae	*Toxodon platensis*	South America	13,000
Edentata	Megatheriidae	*Megatherium americanum*	South America	8500
Edentata	Megatheriidae	*Eremotherium mirabile*	South & North America	13,200
Edentata	Megatheriidae	*Glossotherium harlani*	South and North America	12,500
Edentata	Megatheriidae	*Mylodon robustus*	South America	12,000
Marsupiala	Diprotdontidae	*Diprotodon optatum*	Australia	45,000

Table 3 Megaherbivores still extant

Common name	*Scientific name*	*Adult body mass (kg)*		*Historic distribution*	*Status*
		Female	Male		
African savanna elephant	*Loxodonta africana*	2800–4000	5000–8000	Africa-wide	Still widespread and locally abundant
African forest elephant	*Loxodonta cyclotis*	2500	4500	Forested regions of Africa	Declining under hunting
Asian elephant	*Elephas maximus*	2500–4000	4000–5400	India, Burma, Thailand, Malaysia, Vietnam, Sri Lanka	Locally common but insecure
Hippopotamus	*Hippopotamus amphibius*	1350–2350	1500–2600	Africa-wide	Abundant
White rhinoceros	*Ceratotherium simum*	1600–1800	2200–2400	Southern and northeastern Africa	Secure in south, extinct in north
Great Indian rhinoceros	*Rhinoceros unicornis*	1600	2100	India, Nepal	Restricted to reserves
Black rhinoceros	*Diceros bicornis*	1000–1300	1000–1300	Most of Africa	Uncommon and endangered
Javan rhinoceros	*Rhinoceros sondaicus*	1300	1300	Java	Rare and endangered
Giraffe	*Giraffa camelopardalis*	825–1125	1200–1400	Southern and east-central Africa	Common and secure
West Indian manatee	*Trichechus manatus*	1600	1600	Florida, Caribbean, northern South America	?

energy and nutrients per unit of body mass diminishes with increasing body size. However, the amount of food an animal can consume per day depends on its gut capacity, which is directly proportional to body size and hence mass. As a consequence, larger herbivores can satisfy their nutritional requirements from lower quality food than can smaller animals (Jarman, 1974). Accordingly, a greater fraction of the vegetation is suitable food to support a megaherbivore than is the case for medium-sized herbivores. This basic relationship contributes toward the repeated tendency for mammalian herbivore lineages starting off small to evolve toward larger body mass over time, culminating in the megaherbivore syndrome. The largest birds like ostriches, emus, and the extinct moas (in New Zealand) and elephant birds (in Madagascar) are also herbivores, but remain below the megasize threshold.

Digestive physiology contributes further toward dietary tolerance. The cellulose molecules constituting most of this fiber cannot be digested by mammalian enzymes. Hence, microbial fermentation is employed to degrade the fibrous cell walls into fatty acids, which can be used in place of glucose as an energy-yielding substrate. Ruminants within the order Artiodactyla have a greatly enlarged chamber anterior to the true stomach called the rumen, where this fermentation occurs. The foregut location of this chamber also permits these animals to ruminate (or chew the cud) to increase the surface area available to bacterial enzymes. The largest extant

ruminant is the giraffe (*Giraffa camelopardalis*; body mass typically 1200 kg for males and 825 kg for females), which is entirely a browser on the foliage of tall trees, and among wild cattle the gaur (*Bos gaurus*; males typically 940 kg and females 700 kg). Hence, only males and perhaps exceptional females of these two species enter the megaherbivore size range. Among extinct ruminants, the giant Irish elk (*Megaloceros giganteus*) also came close to megaherbivore size, although its exact weight can only be conjectured. Ruminants gain little digestive benefit from growing larger than the 1000 kg threshold, because cattle and wild bovines are capable of digesting almost all of the plant foliage that is potentially fermentable (Demment and Van Soest, 1985; Clauss *et al.*, 2003). The absolutely indigestible fraction consists of cellulose coupled with lignin, the substance of wood. The limitation of ruminant digestion is that especially fibrous plant tissues do not pass readily through the orifices, retaining material in the rumen, so that ruminants can starve despite having a gut full of potential food. Hence, ruminants seek out the more digestible components of the vegetation, such as green leaves, and avoid stem tissues as much as is possible.

All megaherbivores apart from giraffe are nonruminants, mostly relying on fermentation taking place in the hindgut region, specifically the colon and cecum (Langer, 1988). Larger size means that the forage consumed takes longer to pass through the digestive tract, allowing it to be digested more completely. Nevertheless, the inability to chew material further after ingestion limits the digestive efficiency of nonruminants. Experiments feeding standard diets to zoo animals showed that the extent of fiber digestion achieved by white rhinos was similar to that of medium-sized antelope, but less than exhibited by cattle and wild bovines (Foose cited by Owen-Smith, 1988; Steuer *et al.*, 2010). However, digestive adaptations vary among extant megaherbivores. Elephants show a comparatively rapid passage rate through the gut, similar to that of equids, and digest the forage they eat rather poorly (Clauss *et al.*, 2007a). Hippos ferment forage within an expanded foregut, without the benefit of rumination and rely on prolonged retention to achieve adequate efficiency.

Distinctions in digestive anatomy between grazing and browsing ruminants are underlain by the need to ferment fibrous grass leaves for longer than more succulent tree leaves (Hofmann and Stewart, 1972). Counterpart distinctions in the forage eaten exist among megaherbivores. White rhinos, Indian rhinos, and hippos are mostly grazers. The African savanna elephant and Asian elephant both consume a mixed diet including a substantial proportion of tree leaves, twigs, bark, roots, and fruits as well as grass, thus including some very fibrous material. Black rhinos are mostly browsers and show shorter retention times and hence less efficient digestion than white rhinos. The extinct mammoths appear more specialized than modern elephants for grazing from their dentition, while the extinct woolly rhinoceros was also mainly a grazer. Among gomphotheres and ground sloths, some species concentrated on grass and others on browse (MacFadden, 2000; Prado *et al.*, 2005).

Differences in capacity to detoxify plant secondary chemicals, which are more prevalent in the foliage of woody plants and forbs than in grasses, also exist (Hofmann, 1989). Black rhino diets contain many plant species with potentially toxic chemicals, especially stem succulents in the family Euphorbiaceae with milky latex (Loutit *et al.*, 1987; Luske *et al.*, 2009). African elephants, however, reject many woody species readily eaten by ruminants, suggesting that they are less tolerant of plant secondary chemicals. Their special adaptation is a capacity to utilize very fibrous plant parts, such as bark and roots, during hard times, which they extract by bark-stripping, felling, or uprooting trees and shrubs (Owen-Smith and Chafota, in press).

However, the digestive adaptations of megaherbivores lead to an elevated dietary requirement for sodium and potassium, as has been established for black rhinos compared with horses (Clauss *et al.*, 2007b). As a consequence of lowered digestive efficiency, the fecal bulk produced per unit of dry matter intake is greater, resulting in a greater loss of these minerals in the feces. This finding fits with observations showing that elephants select waterholes with higher sodium contents (Weir, 1972), and ingest sodium-rich soil, in regions where water supplies are deficient in sodium (Holdo *et al.*, 2002). The elevated requirement for sodium, sparsely present in most plants, could restrict the distribution of megaherbivores. For example, white rhinos were historically absent from the Highveld Grassland region of southern Africa, whereas black rhinos occupy especially the drier savannas and shrublands, where evaporation concentrates salts in soils (Owen-Smith, 1988).

On account of their dietary tolerance, megaherbivores would have been widely distributed through landscapes in the past, just as elephants and rhinos had been in Africa before humans restricted their presence. A much larger proportion of the vegetation would have been consumed and digested via herbivore guts, leaving less to be decomposed more slowly by soil organisms.

Vulnerability to Predation

A second fundamental feature of megaherbivores is that they are effectively invulnerable to predation, except by humans, in the adult stage. This removes a second factor restricting the habitat occupation of smaller herbivores: the heightened risk of being killed in places where predators hunt most successfully. Elephants and rhinos brazenly penetrate areas of dense bush and long grass providing ample cover for ambush predators like lions, whereas grazing antelopes usually avoid closed habitats, however good the forage quality offered. These megaherbivores are also relatively more active at night, when ambush predators are more effectively concealed, than most smaller ungulate species (Owen-Smith, 1988). This helps them to meet their greater quantitative food requirements when food is sparse

Apart from a few calves, elephants, rhinos, and hippos hardly featured in predator kills recorded in the Kruger National Park (Owen-Smith and Mills, 2008). Giraffes are frequently killed by lions (*Panthera leo*) in Kruger Park, but rarely in Serengeti National Park (Hopcraft, 2010), indicating the transitional status of their body size. Elephant calves are closely protected by their mothers while very young, but become more vulnerable to being predated during droughts when they lag behind the herd (Loveridge *et al.*, 2006).

In Botswana's Chobe National Park, a large pride of lions specialized in killing young elephants near waterholes at night during the late dry season (Joubert, 1986; Power and Compion, 2009). Young hippos constituted a substantial proportion of lion prey in Zambia and Congo (Bourliere and Verschuren, 1960). Attacks by lions on young or near-adult rhinos of both African species have been recorded (Goddard, 1976; Pienaar, 1970; Joubert and Eloff, 1971; Brain *et al.*, 1999), but seem infrequent. Black rhino calves seem especially vulnerable to being killed by spotted hyenas (*Crocuta crocuta*; Hitchins and Anderson, 1983). A substantial proportion of Indian rhino calves are killed by tigers (*Panthera tigris*; Dinerstein, 2003), and tigers are also a threat to Indian elephant calves (Sukumar, 2003).

The close family groups formed by elephants are basically an adaptation to protect offspring from predator attacks. Adult rhinos of all species are generally solitary, apart from mother–offspring pairs, as also are hippos while foraging on land (Owen-Smith, 1988). However, subadult white rhinos associate in groups numbering from two to three animals up to 12 or more, sometimes including an adult female lacking a calf. When threatened, animals in these groups stand with rumps together and horns facing outward in a defensive formation.

Extinct saber-toothed felids had enlarged upper canine teeth seemingly adapted to slice through the tough hides of very large prey, although only the largest species would have threatened adult megaherbivores (Turner and Anton, 2004). Nevertheless, all animals become increasingly weakened with advancing age. Besides killing young animals, these saber-tooths may also have assisted the death of aging megaherbivores. The dwarfing of elephants and hippos that has occurred repeatedly on islands lacking large predators (Simmons, 1988; Stuart, 1991) testifies to the contribution of predation toward promoting large size.

The arrival of humans organized in groups with projectile weapons breached the basic defense that megaherbivores had evolved against predation, that is, standing their ground to ward off the attack. Spears could be thrown, or arrowheads launched, until the wounds inflicted weakened the animal, leading eventually to a kill. Fleeing is now the response most commonly shown by elephants and rhinos when humans are detected nearby, appropriate only for this threat, backed by attacks that are rarely effective against nonhuman predators. White rhinos today still waver uncertainly between standing defensively and fleeing when approached by humans, unless the identification of the intruder is clearly by scent, at which point they head for the horizon (Owen-Smith, 1988). Alarm calls from associated oxpecker birds (*Buphagus erythrorhynchus*) also confirm the presence of a human that cannot be identified by sight by the rhinos, because of their poor vision. The frightening charges commonly made by black and Indian rhinos serve to deter potential human hunters.

Vegetation Transformation and Ecosystem Processes

Digestive tolerance coupled with invulnerability to predation means that the abundance attained by megaherbivores measured as population biomass density (i.e., mass per unit area) is greater than that manifested by most smaller grazers or browsers. Typically, megaherbivores make up 40–70% of the total herbivore biomass in African savanna ecosystems where their populations have not been decimated by past or current human hunting (**Figure 1**). Within Kruger Park, reintroduced white rhinos have become the third most abundant herbivore in biomass, after elephants and buffalos. The fertile Serengeti ecosystem dominated by migratory wildebeest appears to be a notable exception, but lacks the white rhinos that were once abundant there as documented by fossilized remains in Olduvai Gorge (Leakey, 1965). The high biomass of African elephants and white rhinos does not seem to be at the expense of other large herbivores (Skarpe *et al.*, 2004; Arsenault and Owen-Smith, 2012). This means that their grazing and browsing impacts are added over and above those of smaller herbivores. About half of all the vegetation consumed passes through the digestive systems of megaherbivores, and the remainder through the guts of all of the remaining large herbivores combined.

The high abundance and physical capabilities of megaherbivores mean that they function as habitat engineers altering structural features of the vegetation. Elephants have substantially transformed savanna vegetation structure, and to some extent the composition of woodlands, through their browsing and breakage of trees (Laws, 1970). Quite large sections of regional landscapes have been changed from tree savanna into open grassland or coppice shrubbery, depending on soils and rainfall (Bell, 1981; McShane, 1989; Western, 2007). Intense fires burning the grass layer contribute to the lack of regeneration by trees (Dublin *et al.*, 1990; Holdo,

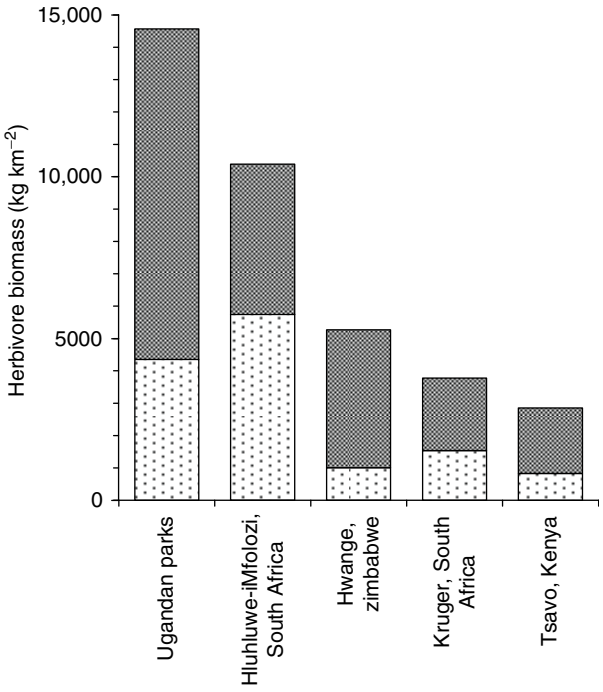


Figure 1 The contribution of megaherbivores (dark shade) to total large herbivore biomass in relatively intact African ungulate assemblages. Uganda is represented by combined data for Queen Elizabeth and Murchison Falls national parks.

2007), but exclosure plots show the overriding effects of elephants both in Uganda (Smart *et al.*, 1985) and in Kruger Park in South Africa (Levick and Rogers, 2008; Levick et al. 2009). Stands of coppicing shrubs cultivated by elephants through breaking leader shoots have been labeled "browsing lawns" (Fornara and du Toit, 2007; Owen-Smith, 2003; Makhabu *et al.*, 2006), by analogy with the grazing lawns described later, implying enhanced food availability for the elephants. Vulnerable plant species may become restricted to steep slopes or places far from water where they are less likely to be encountered by foraging elephants (Edkins *et al.*, 2007). Asian elephants have somewhat less impact on the woodlands and forests to which they are largely confined, but may nevertheless depress the abundance of favored understory trees in deciduous woodlands (Sukumar, 2003). Black rhinos can topple succulent euphorbias (Luske *et al.*, 2009). Heavy browsing or other damage to plants by black rhinos and giraffes can restrict the distribution of sensitive tree species (Bond and Loffell, 2001; Wiseman *et al.*, 2004). Carbon isotope studies indicate that in earlier times, African savanna elephants subsisted largely on grass, suggesting a scarcity of trees through much of their range (Cerling *et al.*, 1999). However, the lack of trees may also have been promoted by the cooler temperatures and lower atmospheric CO_2 levels prevalent during the Pleistocene (Bond *et al.*, 2003).

Grazing impacts by white rhinos have converted areas formerly supporting tall grassland in Hluhluwe-iMfolozi into a mosaic of short grass areas or "grazing lawns" interspersed with taller grass patches (Owen-Smith, 1988; Waldram *et al.*, 2007). Hippos likewise 'cultivate' extended grazing lawns near the pools where they seek day-time refuge (Olivier and Laurie, 1974). Grazing by hippos near rivers and by white rhinos more widely may suppress the spread of fires, facilitating the regeneration of the trees damaged by elephants. Grazing lawns provide more nutritious grasses, attracting concentrations by smaller grazing ungulates (Verweij *et al.*, 2006; Arsenault and Owen-Smith, 2012).

Even in Africa, no area today bears testimony to the ultimate state of the landscapes transformed by the combined impacts of both browsing and grazing megaherbivores at their potential saturation densities. Hippos as well as elephants attained high densities in Queen Elizabeth National Park in Uganda, but their grazing impacts are confined to the vicinity of water rather than affecting the broader landscape (Lock, 1972). Reintroduced white rhinos remain well below their former densities outside the Hluhluwe-iMfolozi Park, and the elephant population reintroduced there is still growing.

During Pleistocene times when megaherbivores of many species were abundant, areas of North America today covered by fairly uniform forests showed a mix of conifers, broadleaf hardwoods, and graminoids, suggesting a parkland mosaic (Wright, 1984; Gill *et al.*, 2009). The prairie grassland that today predominates in western America has no earlier counterpart. The grassy steppe tundra that once stretched from Alaska through Siberia into northern Europe has been replaced by the dense assemblage of shrubs, mosses, and bogs that characterizes the modern tundra (Zimov *et al.*, 1995). Former associations of animal species were quite disparate compared with modern assemblages: Grazing mammoths, equids, and bison coexisted alongside browsing mastodonts, ground sloths, tapirs, and deer, plus mixed feeders like caribou and muskoxen, in the Appalachian region (Guilday, 1984). The mosaic interspersion of vegetation types of the Pleistocene has become transformed into the zonal vegetation bands of today, largely climatically controlled (Guthrie, 1984).

In northwestern Europe, pollen and animals indicative of open vegetation suggest that, even during the last interglacial, grassy meadows were quite common in the landscape, compared with the current predominance of closed forest in areas undisturbed by humans (Svenning, 2002; Vera *et al.*, 2006). Even today, grazing and browsing by cattle and horses suppress tree regeneration locally, thereby maintaining grassy gaps and promoting a cyclic succession pattern (Olff *et al.*, 1999). The complete disappearance of the grassland steppe that prevailed through northern Europe, Siberia, and Alaska during glacial conditions has been related to the loss of megaherbivore impacts maintaining the open soil cover, and thereby restricting permafrost development (Zimov *et al.*, 1995).

Through opening the forest cover, megaherbivores facilitate the penetration of fire, which further modifies these habitats (Svenning, 2002; Gill *et al.*, 2009). Their grazing impacts suppress the spread of fires within grasslands, facilitating tree and shrub establishment. In this way they contributed to vegetation formations prevalent during the Pleistocene that have no modern analog (Wright, 1984).

The vegetation changes induced by megaherbivores also facilitate the coexistence of smaller, more selective grazers and browsers, by bringing more of the vegetation within the height range favored by the latter (Rutina *et al.*, 2005). Moreover, regenerating plant parts are fundamentally more nutritious than the components removed (Makhabu *et al.*, 2006). The overall energy flow through the large herbivore biomass may thus be greatly enhanced by the presence of megaherbivores, at the expense of the amount recycled more slowly through termites and other soil organisms, or incinerated by fire. Hence, megaherbivores may play a 'keystone' role in maintaining both habitat productivity and a diversity of other species (Owen-Smith, 1987, 1989).

Life History Patterns and Population Regulation

Megaherbivores live longer, attain reproductive maturity later, have lengthened gestation periods, and have more prolonged intervals between successive births, than smaller species, because of the allometric dependency of metabolic rates on body size (Owen-Smith, 1988). Specifically, the time taken for these events to occur generally lengthens as a function of body mass raised to the power one-quarter. Elephants of both species can survive up to 65–70 years (Moss, 2001; Sukumar, 2003), whereas the potential longevity of rhinos and hippos is 45–50 years (Owen-Smith, 1988). The highest recorded age for a giraffe is 28 years. Aside from primates and cetaceans (whales and dolphins), other mammals do not survive much longer than 20 years.

Most crucially, gestation periods of megaherbivores generally exceed one year, the exception being the hippo with a gestation time of only 8 months. Consequently, intervals between successive births span two or more years, even for hippos, and hence become partly decoupled from the annual

seasonal cycle. The shortest birth intervals recorded for elephants and rhinos fall close to the allometric trend extrapolated from smaller species, but observed mean birth intervals are somewhat longer (**Figure 2(a)**). Thus, although megaherbivores have the capacity to produce offspring as fast as allometric constraints allow, in practice they tend to reproduce more slowly in response to nutritional shortfalls.

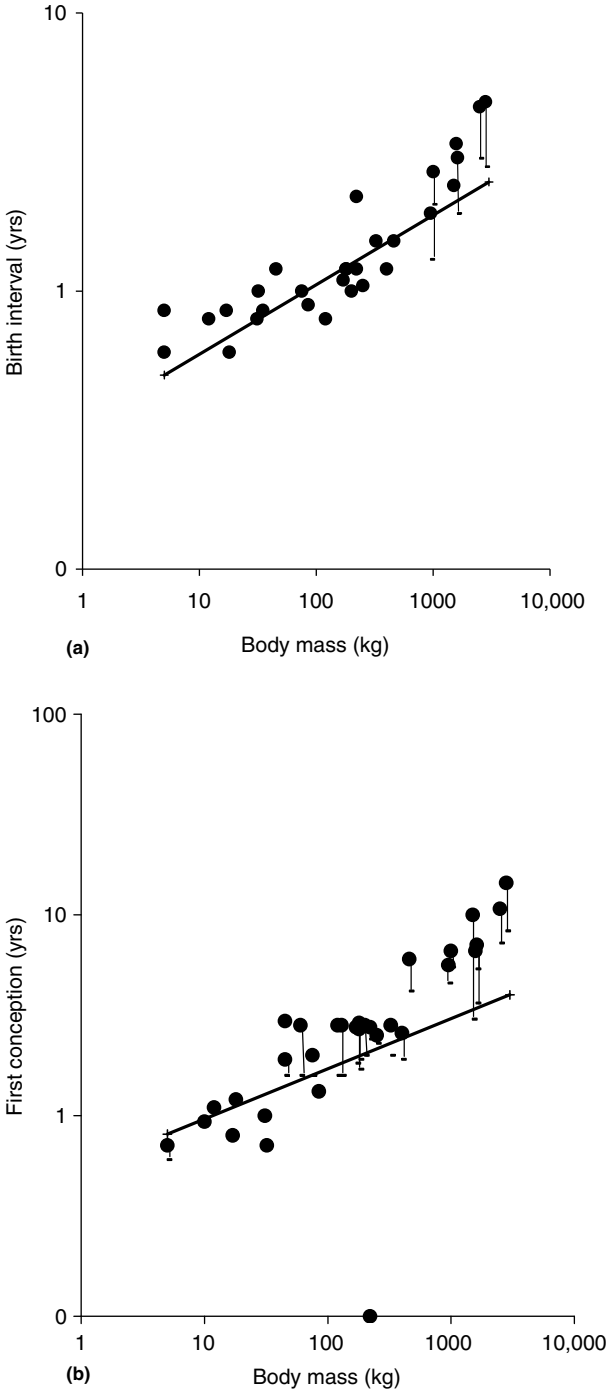


Figure 2 Allometric trends among large mammalian herbivores in (a) interval between births and (b) age at first conception. Circles represent the population averages; lines below indicate minimum records. Trend lines shown have a slope proportional to body mass raised to the power three-quarters.

A similar pattern exists with regard to the age at first conception, but in this case even the youngest ages recorded for elephants and rhinos tend to be greater than would be expected from allometric extrapolation (**Figure 2(b)**). Elephants take an especially long time to reach reproductive maturity, perhaps for social reasons.

Ungulates below megaherbivore size mostly reproduce annually, with offspring born at the time of the year when conditions are most favorable nutritionally for the elevated demands of late gestation and lactation. Among megaherbivores, births are spread through the year, but with a peak in the monthly frequency of conceptions leading to a subsequent peak in births. African elephants, with a 22-month gestation, show a concentration of births early in the wet season, derived from a conception peak 2 months into the wet season (Wittemyer *et al.*, 2007). Asian elephants kept in captivity in India and Myanmar have a birth peak extending into the early dry season, indicating a conception peak during the monsoon rains (Sukumar, 1989). Indian rhinos in Nepal show a similar seasonal concentration of births to these elephants, despite having a shorter gestation (Dinerstein, 2003). However, white rhinos, black rhinos, and giraffes, with gestation times of 15–16 months, exhibit birth peaks during the early part of the dry season, following conceptions concentrated during the rains. Hippos, in contrast, have a conception peak in the early dry season leading to births during the rains. This flexibility in reproductive seasonality is probably an outcome of the spread through the year of critical nutritional demands preceding mating, after birth and following weaning.

Slowed reproduction restricts the proportional rate at which megaherbivore populations can increase. The maximum growth rate for an African elephant population is approximately 6.5% per year, assuming minimal mortality, whereas for a white rhino population, the maximum sustainable rate is 9% per year (Owen-Smith, 1988). Annual population increases may exceed these levels, due to synchronized pulses in reproduction and irregularities in the population structure. Realized population growth rates in the long term are likely to be somewhat slower than the maximum potential rate due to environmental variation and resultant mortality, especially among calves (Moss, 2001).

The pattern of density dependence in the population growth rate shown by megaherbivores may also deviate somewhat from that shown by less large mammals. Mammal populations that are nutritionally limited may experience little reduction in their growth rates until some threshold density has been exceeded (Fowler, 1987). Megaherbivores take this pattern to extremes, as has been documented for whales (Chapman, 1981), and show little response to increasing density until their abundance is quite close to the habitat capacity (**Figure 3**). This explains the lack of signs of density dependence by recovering populations of African elephants and white rhinos that has perplexed some observers (Owen-Smith *et al.*, 2006; Gough and Kerley, 2006). Nevertheless, substantial delays in the age at first reproduction and lengthening of birth intervals have been recorded for African elephants crowded at high densities within protected areas by human settlements and hunting outside (Laws and Parker,

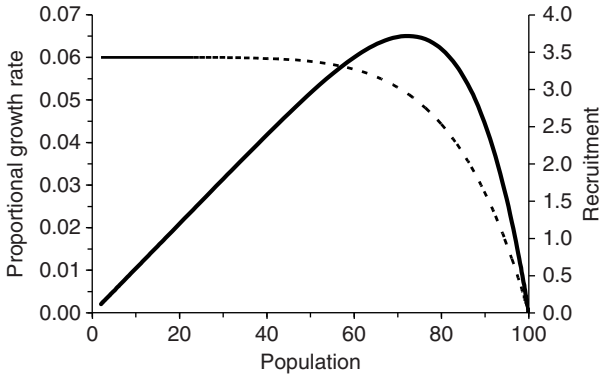


Figure 3 Relationship between the population growth rate and the population size expected for megaherbivores, showing delayed onset of the density-dependent reduction in the relative growth rate (dotted line), and the corresponding right-skewed recruitment curve (i.e., the numerical growth increment; solid line).

1968). Asian elephants, largely confined to forests by human settlements and cultivation, show population growth rates only slightly positive through similar reproductive responses (Sukumar, 1989, 2003). In the Hluhluwe-iMfolozi Park, the shrinking black rhino population showed greatly retarded reproductive rates compared with the expanding white rhino population in the same park, in response to deterioration in its habitat (Hitchins and Anderson, 1983; Owen-Smith, 1988). Indian rhinos in Nepal exhibit substantially longer calving intervals than expanding white rhino and black rhino populations in Africa, although the age at first reproduction differs little (Dinerstein, 2003).

Megaherbivore populations are also less responsive to yearly variation in rainfall conditions than annually breeding ungulates. Elephant population trends are affected little in most drought years, except that conceptions tend to be deferred, leading to fewer calves being produced a year or two after the drought. This response can generate annual peaks and troughs in births (Moss, 2001). Nevertheless, elevated mortality among calves does occur during extreme droughts (Conybeare and Haynes, 1984; Dudley *et al.*, 2001; Foley *et al.*, 2008). Contributing to the heightened calf mortality is the increased daily distance that animals must travel between food and water (Young and van Aarde, 2010), increasing their susceptibility to predation (Loveridge *et al.*, 2006). An exceptional die-off of elephants occurred in Tsavo East National Park in Kenya when severe drought conditions persisted through two successive years in this semiarid region, affecting adult females as well as their offspring (Corfield, 1973). Total mortality amounted to approximately 20% of the population and was concentrated near rivers in the driest region of the park. Predisposing this mortality was the compression of elephants into the park by hunting activities outside (Myers, 1973; Parker, 1983).

Despite the drought conditions prevalent in the Hluhluwe-iMfolozi Park during the late 1960s, the white rhino population continued to increase at its maximum potential rate, perhaps because less rain meant that the grass produced was more nutritious (Owen-Smith, 1988). Hippos are somewhat more susceptible to drought-related mortality than other megaherbivores because of their grazing restriction to the neighborhood of rivers, where forage becomes most severely depleted. Hippos correspondingly show shorter calving intervals and earlier reproduction than other megaherbivores, enabling them to recover faster from population crashes.

Elephants have the capacity to move widely to avoid locally adverse conditions, when not restricted by human settlements or fences. Seasonal home ranges typically encompass 500–1000 km^2 during the dry season, when access to surface water may serve as a restriction, expanding to approximately 2000 km^2 in the wet season, in drier savannas, but tend to be generally smaller in wetter savannas (Young *et al.*, 2009). In semidesert areas, annual ranges traversed by elephants may extend over much as 6000 km^2 for females and 12,000 km^2 for males, although only a fraction of this area is actually utilized (Leggett, 2006). During the wet season, elephant herds concentrate opportunistically in places where local rainfall has resulted in greening of the vegetation (Loarie *et al.*, 2009). Home ranges documented for Asian elephant herds cover 100–800 km^2 (Sukumar, 2003). Within the high-density Hluhluwe-iMfolozi population, female white rhinos moved through overlapping home ranges encompassing 10–20 km^2, whereas males occupied much smaller discrete territories (Owen-Smith, 1975). Larger home ranges of up to 90 km^2 have been recorded for white rhinos elsewhere (van Gysegem, 1984; Pienaar *et al.*, 1993). Black rhinos home ranges are typically under 50 km^2, but may expand to as much as 500 km^2 in desert regions (Owen-Smith, 1988). Home ranges occupied by Indian rhinos are similar in size to those of the African species. Giraffe home ranges typically cover approximately 100 km^2, but may be as large as 500 km^2 in dry regions.

More persistent dispersal movements contribute to the longer term redistribution of elephant populations, responding to changing water as well as food resources (Laws, 1969; Chamaille-Jammes *et al.*, 2008; van Aarde and Jackson, 2007). The regional population of almost 150,000 elephants spread through northern Botswana, western Zimbabwe, eastern Caprivi, and adjoining regions of Angola and Zambia has been increasing in numbers, but not in density, indicating dispersal outward (Junker *et al.*, 2008). The growth of the elephant population within Kruger Park has also slowed as animals moved into areas to the east and west, following the removal of border fences to establish the Greater Limpopo Transfrontier Park. White rhinos showed evidence of dispersal from the high-density core region into outlying areas by subadult animals of both sexes, along with some mature males (Owen-Smith, 1988). Hippos have also been observed transferring between water bodies in response to changing conditions.

Disease has not had much impact on the population dynamics of elephants and rhinos, although white rhinos relocated into East Africa are vulnerable to infection by trypanosome blood parasites causing sleeping sickness in humans. An unusual situation exists in Nambia's Etosha National Park, where periodic mortality from anthrax has stabilized elephant numbers (Lindeque and Turnbull, 1994). However, hippos have been subject to outbreaks of anthrax in the crowded pools where they seek daytime refuge (Wafula *et al.*, 2007).

Stabilization of African elephant populations is likely to come about mainly through reproductive changes, coupled

with episodic mortality mostly among calves during drought years, plus perhaps some predation (Owen-Smith *et al.*, 2006; Wittemyer *et al.*, 2007; Trimble *et al.*, 2009). The robust growth rates shown by most elephant populations not subject to poaching may be a reflection of the exceptionally favorable, low-density conditions that these animals have experienced from birth onward. Cohort effects, whereby animals born under adverse conditions show lower survival and reproductive success throughout their lives (Lindstrom and Kokko, 2002), may come into play at a later stage as populations approach the habitat capacity. Whether regulating processes can operate effectively within the confinements of protected areas is a leading conservation issue.

It is the slow population growth potential of megaherbivores that makes them so susceptible to predation on adults imposed by humans. It also takes longer for their populations to recover after having been reduced. However, this did not make such large animals especially vulnerable to extinction in the past, because of their delayed or threshold density dependence, robust dependence on a wide range of food resources, and capacity to range widely to avert the effects of local droughts or other adverse conditions. Hence, megaherbivores represented an extremely successful life form, until modern humans armed with effective weapons arrived on the scene.

Late Pleistocene Extinctions

The distinctive features of the ecology of megaherbivores are fundamental for explaining the causes of the wave of extinctions that greatly reduced large mammal diversity through most of the world toward the end of the Pleistocene. The crucial features of these extinctions are (1) the strong bias toward larger mammals, which is distinct from all previous extinction episodes, and (2) the differences in timing in different regions of the world, meaning that there is no consistent climatic relationship (Owen-Smith, 1987; Alroy, 1999; Carrasco *et al.*, 2009).

In the late Pleistocene, all megaherbivore species weighing more than 1000 kg, and 80% of macroherbivores in the size range 100–1000 kg, were eliminated throughout all continents except Africa and tropical Asia (**Figure 4**). Around half of the species weighing between 10 and 100 kg disappeared from North America and Australia, but few losses in this size range occurred in Eurasia and South America. Extinctions of carnivores and scavenging birds occurred along with their herbivore prey (Martin, 1967), and in Australia, giant emu-related birds vanished along with the marsupial mammals (Flannery and Roberts, 1999). Africa lost the last of the grazing *Elephas iolensis* in Tunisia approximately 34,000 years ago (Todd, 2005). Four genera of macroherbivores also became extinct in Africa approximately 10,000 years ago: the antlered giraffe *Sivatherium*; the giant buffalo *Pelorovis*; the giant hartebeest *Megalotragus*; and the giant warthog *Metridiochoerus*.

In tropical Asia, the elephant-like *Stegodon* disappeared approximately 11,000 years ago. In Europe, the last recorded date for the straight-tusked elephant *Palaeoloxodon antiquus* was approximately 40,000 years ago, whereas two rhinoceros species vanished approximately 20,000 years ago (Stuart, 1991, 2005). The woolly mammoth and woolly rhino persisted in northern Europe and Siberia until as recently as 11,000–10,000 years ago, and, in dwarfed form, on Wrangel Island until 3700 years ago (Vartanyan *et al.*, 1993; Orlova *et al.*, 2004). In North America, the major pulse of megafaunal extinctions occurred within a narrow time window between approximately 12,500 and 11,500 years ago (adjusting radiocarbon dates reported by Stuart, 1991). Dwarfed mammoths persisted in islands off California, and dwarfed ground sloths in the Caribbean, until somewhat later. In South American the extinction wave apparently intensified slightly earlier, between 14,000 and 11,000 years ago (Barnovsky and Lindsey, 2010). The gomphotheres *Cuvieronius* and *Haplomastodon* were last recorded in Chile and Venezuela, respectively, between 14,000 and 13,000 years ago, whereas the giant ground sloth *Megatherium* persisted in Argentina until perhaps as late as 8500 years ago. In Australia, the extinction wave began shortly after 50,000 years ago (Bowler *et al.*, 2003).

In both North and South America, the timing of most extinctions was closely associated with the Cold Reversal between 11,200 and 13,500 years ago, an abrupt cooling interlude interrupting the warming trend associated with the transition from the Pleistocene into the Holocene (Barnovsky *et al.*, 2004; Barnovsky and Lindsey, 2010). There was no consistent relationship with the timing of climate change elsewhere, and extinctions had not been a feature of previous glacial–interglacial transitions, despite temperature shifts similar to those ending the Pleistocene. Species in the megaherbivore size range should be least vulnerable to climatic influences on habitat conditions, because of their wide food tolerance and broad distributions.

However, a consistent association is apparent between the regional timing of the extinctions and the appearance of modern humans in these places (Burney and Flannery, 2005; Lyons *et al.*, 2004). *Homo sapiens sapiens* moved out of Africa into parts of the Middle East between 120,000 and 80,000 years ago, and dispersed from there along a coastal route to reach Australia by 50,000 years ago (Finlayson, 2005). Extinctions of marsupial mammals and emu-related birds followed shortly after the colonization of Australia by the ancestral Aborigine people (Miller *et al.*, 1999; Roberts *et al.*, 2001; Bowler *et al.*, 2003). Modern humans had spread northward into central Eurasia shortly after 50,000 years ago, but were forced back into more southerly regions during the glacial maximum approximately 20,000 years ago. Thereafter their range expanded through northern Europe and Siberia into North America by 15,000 years ago, to reach the southern tip of South America soon thereafter (Goebel *et al.*, 2008). More primitive humans represented by *Homo erectus* had been present as far north as Georgia and China by 1.7 million years ago, and persisted in the Far East until as recently as 30,000 years ago, without any apparent impact on the large mammal fauna. Neanderthal people (*Homo sapiens neanderthalensis*) occupying Europe between 450,000 and 28,000 years ago were inhabitants of Mediterranean woodlands rather than more northern steppe environments (Finlayson, 2005). Modern humans displacing them had a more gracile morphology adapted for endurance running, which facilitated their expansion into the mammoth steppe.

Megaherbivores that had evolved in the absence of humans outside of Africa and tropical Asia would most likely have

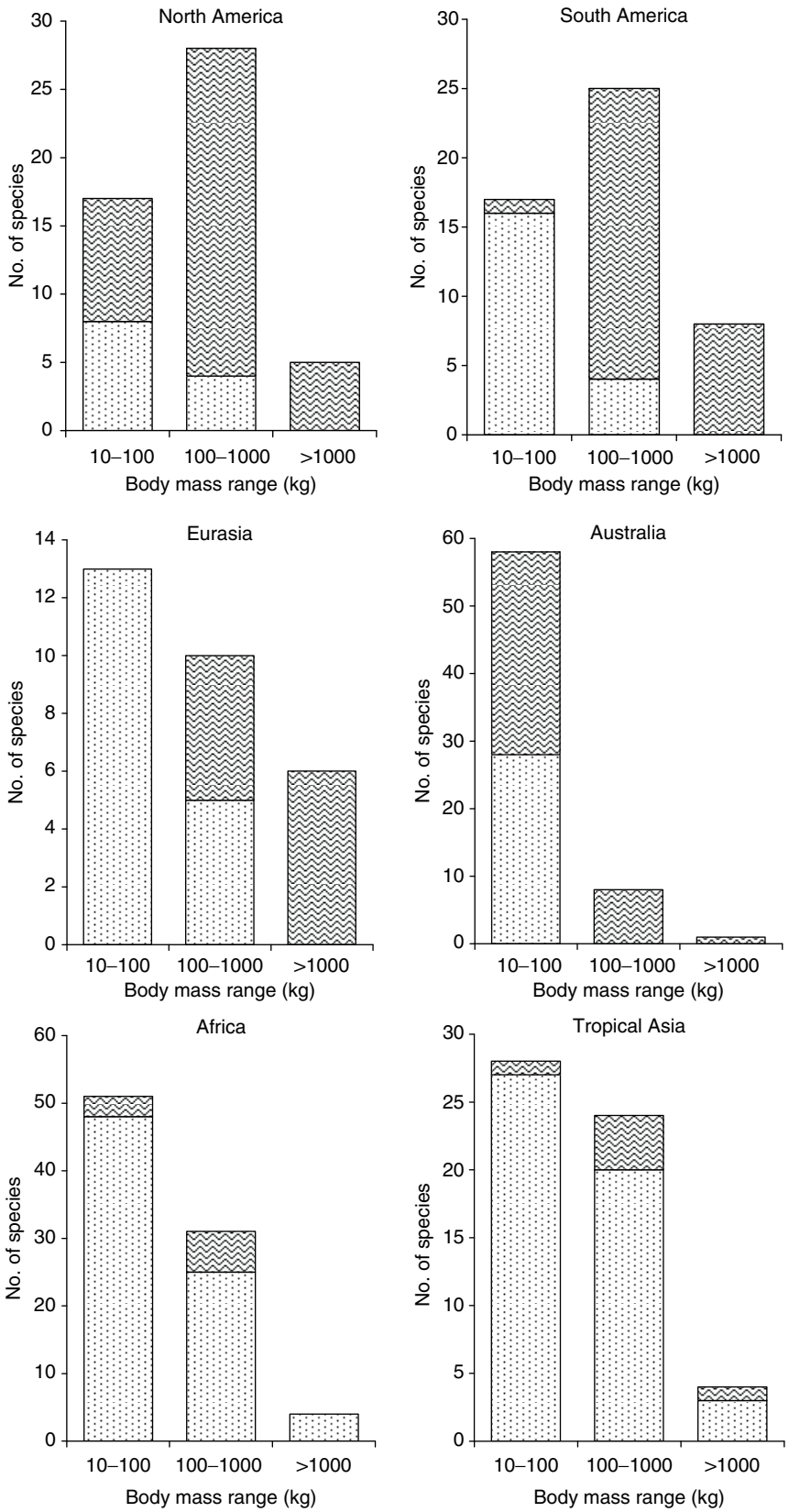


Figure 4 Extent of species extinctions during the late Pleistocene for large mammalian herbivores within different body size ranges among the continents. Stippling at bottom indicates still extant species; wavy line indicates species that went extinct.

responded to threatening humans as they did to other predators, by standing defensively, being too slow moving to escape by fleeing (Berger *et al.*, 2001). They would thus have been sitting targets for skilled groups of human hunters, even if the latter were armed only with stone-tipped spears or arrows. The protein provided, preserved in snow, could have been crucial in enabling human groups to subsist through frigid winters in high northern latitudes in Eurasia and North America. Modern elephants and rhinos have evolved aggressive responses from their long association with humans, deterring hunters until improved weapons reduced the risk.

Megaherbivore populations have limited capacity to sustain predation, especially if imposed across all age classes as would have been the case for human hunters. As the proportional harvest approached the maximum population growth potential, populations would have gone into progressive decline, exacerbated by the right skew in the recruitment curve (**Figure 3**). "Allee effects," manifested through a reduction in the potential growth rate with diminishing density, may have contributed further to a slippery slide toward extinction (Owen-Smith, 1999). Extirpations especially of the megaherbivores became inexorable once humans had built up sufficient numbers (Alroy, 2001). The macromammal types that went extinct tended also to have slow reproductive rates (Johnson, 2002). Nevertheless, predators normally switch to hunting alternative prey species once more favored species become sufficiently rare, thereby averting the extinction of their prey. Humans are unusual in being omnivores as well as effective hunters, and thereby able to subsist on plant food while they continue to hunt large mammals to augment their diet. Climatic shifts may have contributed to the reliance of humans on hunting at a time when crops and wild plants became less dependable.

Around this time the mosaic interspersion of woodland and meadow that typified Pleistocene vegetation patterns became replaced by the broad expanses of uniform forest and extensive grassland persisting into modern times (Guthrie, 1984; Johnson, 2009). Fungal spores typical of large mammal dung fade out rapidly after 14,000 years ago in eastern North America, followed thereafter by a rise in charcoal, indicating extensive fires, together with greater representation of hardwood tree pollen (Robinson *et al.*, 2005; Gill *et al.*, 2009). Last recorded dates of Columbian mammoths after 10,000 years ago fall either in the prairie region along the Canadian border or in the arid region near the Mexican border, where more open conditions remained (Agenbroad, 2003). Some woolly mammoths may also have persisted in Alaska after 10,000 years ago (Haile *et al.*, 2009). Extinctions of numerous large mammals were closely associated with human arrival in the pampas of Argentina and Uruguay and in Patagonia, although elsewhere in South America, humans and large mammals apparently coexisted for several thousand years (Barnovsky and Lindsey, 2010).

The vegetation changes that occurred at the end of the Pleistocene cannot be related solely to climate, because large herbivore impacts and fire also substantially modify vegetation structure and composition (Owen-Smith, 1987; Bond, 2005). In Siberia, there are indications that areas of open steppe-tundra persisted through previous interglacials, but became entirely replaced by boggy tundra after the end of the last ice age (Sher, 1997, reported by Stuart, 2005). The megafaunal extinction by human hunting could have played as great a role as the climate shift in this radical vegetation transformation from graminoid to shrub and moss dominance (Zimov *et al.*, 1995; van der Wal *et al.*, 2004). As noted earlier (*see* Vegetation Transformation and Ecosystem Processes), disturbance by megaherbivores promotes not only structural vegetation diversity but also greater vegetation productivity, at least in the form of regenerating tissues most supportive of smaller herbivores. Hence, deterioration in habitat conditions for macroherbivores may have followed the elimination of megaherbivores by human hunters, making these other species also more vulnerable to hunting pressure, especially at a time when suitable habitats were potentially shifting in location in response to climate change. However, this "keystone herbivore" hypothesis (Owen-Smith, 1987) does not apply in Australia, where a single, rather minor megaherbivore was lost and the use of fire by aboriginal people has been emphasized as playing a major role in the megafaunal extinctions (Bowman, 1998). Within Africa, the disappearance of the grazing *E. iolensis* earlier in the Pleistocene could have been a consequence of widening use of fire by humans, depleting forage during the dry season bottleneck.

Hence, megaherbivores played a central role in the late Pleistocene extinctions, both through being the animal type most susceptible to extinction by human hunters and by contributing through their demise to the vegetation changes that ensued. Their broad dietary and habitat tolerance negates claims that climatic factors were primarily responsible for the extinction wave affecting primarily large mammals. Nevertheless, extreme climatic conditions could have contributed to the timing of the extinctions in North and South America, through increasing human dependency on hunting, rather than directly through vegetation changes (Koch and Barnovsky, 2006).

Recent Human Impacts

Human hunting of the megaherbivores that survived the late Pleistocene extinctions has expanded in modern times. Trade in ivory was initiated by Greeks and Phoenicians around 800 BC, initially exploiting African elephants then still extant in Syria and Asian elephants in India, although later the source shifted to North Africa (Spinage, 1994). The Roman Empire dominated the demand for ivory during the early few centuries AD, but as elephants disappeared from North Africa, the supply became transferred to East Africa and India. By AD 900, Arab traders were importing ivory from East Africa into India and China because Asian elephants had become scarce. During the seventeenth century, Portuguese traders became involved in the ivory traffic, initially in West Africa, extending later to East Africa and southward, closely linked to the slave trade. European hunters played a major role during the nineteenth century, making a living largely from ivory and supplying guns to local people to increase the supply. The amount of ivory leaving Africa was estimated to be approximately 800 tons per year by the late nineteenth century, leading to the elimination of elephants through much of eastern and southern Africa, as reported by Spinage (1994). In South Africa, a few herds survived in forests of the Southern Cape and two or three herds in the region

adjoining Mozambique that became the Kruger Park. Elephants disappeared from Botswana and most of Zimbabwe and Namibia. Elephants are still killed for their ivory today, despite the international ban on trade, so that their numbers north of the Zambezi River continue to shrink. South of the Zambezi, elephant populations have grown to abundance levels widely regarded as excessive. The ivory trade fell less heavily on Asian elephants for the simple reason that females of this species lack tusks. Although male elephants are killed for their ivory, this has less effect on populations than the slaughter of African elephants encompassing both sexes.

Hunting activities during the nineteenth century included rhinos, with horns traded for various purposes (Selous, 1899). White rhinos were abundant throughout the savanna region south of the Zambezi River early in the nineteenth century, but by the end of this century had been extirpated everywhere save for perhaps 50 animals in the newly proclaimed Umfolozi Game Reserve (the original name), and a few survivors in a remote region of Mozambique (Owen-Smith, 1988). The existence of northern subspecifically distinct white rhinos was discovered only in the early twentieth century to the west of the Nile River, but these animals have recently become effectively extinct in the wild. Black rhinos, secluded in bushy habitats, survived somewhat better, and their elimination through most of Africa took place during the twentieth century largely as a result of the expanding demand for rhino horns in the Far East.

A peculiar feature of the distribution of the white rhino is its historical absence from a broad region of central and eastern Africa, where associated grazers and browsers remain abundant. The white rhino is among the most common species found in Pleistocene fossil deposits in Olduvai Gorge and elsewhere (Leakey, 1965). Teeth found in Kenya and rock engravings in Algeria suggest that white rhinos survived in eastern and northern Africa until merely a few thousand years ago (Lang, 1923; Hooijer and Patterson, 1972). Since habitat conditions seem quite suitable through much of the distribution gap (confirmed by successful reintroductions of the species), it seems likely that white rhinos were eliminated from this region through hunting by early Bantu-speaking people moving from central Africa with effective spears due to their Iron Age culture. By the time these people had colonized southern Africa by AD 1000, they had become cultivators and pastoralists and no longer dependent on wildlife for protein.

More recently, the distribution of elephants and rhinos has become restricted by expanding human settlements and continuing, although largely illegal, hunting. The formerly continent-wide distribution of both African savanna and forest elephants is now fragmented, whereas Asian elephants have become narrowly restricted to woodland and forest patches because of the preemption of more open areas by people. Rhinos have become confined almost entirely to protected areas. Hippos, although secure in their aquatic refuges, and giraffes, less of a target for hunters, remain widespread in Africa, although the hippos that once resided in southern Europe have disappeared.

Conservation Dilemmas

The conservation problems associated with megaherbivores are acutely manifested with respect to both elephants and rhinos. Throughout eastern Africa, widening poaching for ivory has greatly reduced elephant numbers, whereas black rhinos have been hunted almost to extinction for their horns. Illegally traded ivory is currently coming largely from forest elephants in central African countries. However, in southern Africa, elephant numbers have grown under effective protection to a situation where progressively worsening consequences of their vegetation impacts are feared (Whyte, 2001; Mosugelo *et al.*, 2002), and their range expansion is exacerbating conflicts with people (O'Connell-Rodwell *et al.*, 2000; Sitati, 2007). Numbers of white rhinos have recovered to levels allowing hunting to be permitted as a legitimate activity outside of protected areas, and even black rhinos are approaching this level in some places. This has led to a conflict in conservation policy within Africa with regard to culling and concerning the international ban on any trade in products. Meanwhile, Asian countries retain only small remnants of their elephants and rhinos, thus fully supporting trading bans.

However, it was in Uganda that culling of elephants for broader conservation reasons was first undertaken in 1965, following the development of high concentrations within Murchison Falls National Park and the consequent woodland disappearance; 2000 elephants were killed (Laws and Parker, 1968). Proposals to reduce elephant numbers in Kenya's Tsavo National Park for similar reasons were blocked by sentiment against lethal intervention. Consequently, nearly 20% of these elephants died during severe drought conditions persisting through two successive years, with numbers reduced further by rampant poaching of the survivors.

In 1966, culling of elephants was initiated in Kruger National Park in South Africa (Pienaar, 1983) and Hwange National Park in Zimbabwe (Cumming, 1981), following observations of worsening vegetation damage as elephant populations increased (van Wyk and Fairall, 1969), and shortly thereafter extended to other parks in Zimbabwe and Luangwa Valley in Zambia (Hanks, 1969). The Kruger population was maintained at a density of one animal per square mile (0.4 per km^2), initially as a precautionary measure but later routinely, through the elimination of approximately 400 elephants annually until 1994, when animal rights groups challenged the scientific justification for this action (Owen-Smith *et al.*, 2006). Culling was then suspended, apart from some live removals, while a national policy review took place. The framework for this assessment concerned broader objectives for biodiversity conservation within protected areas. Meanwhile, elephant numbers within Kruger Park have doubled, and park managers remain worried about the ecological consequences of the continued growth of elephant populations, not only in Kruger Park but also within smaller wildlife reserves in South Africa.

Within Hwange National Park, the ceiling elephant density was set at one animal per square kilometer (Cumming, 1981). After annual culling had failed to hold the population at this level, 10,000 elephants were killed during 4 years between 1983 and 1986. After this intervention was suspended for logistic reasons, the elephant population grew toward a total of approximately 40,000 animals (almost three elephants per square kilometer), at which size it appears to be stabilizing through movement between the national park and adjoining

areas (Chamaille-Jammes *et al.*, 2008). Within Botswana, proposals were made to cull elephants once numbers in Chobe National Park and adjoining regions exceeded 50,000 within an overall range of 100,000 km^2. This policy was not implemented because of logistic, financial, and tourism obstacles. Furthermore, a coordinated set of studies failed to establish any ecological justification for culling along the Chobe riverfront, despite the high elephant concentration there (Skarpe *et al.*, 2004). The elephant population centered on Botswana currently numbers almost 150,000 animals, dispersing increasingly into adjoining regions of Namibia, Zambia, and Angola (Junker *et al.*, 2008). Culling merely to cap this population seems logistically impractical, because this would entail the killing of several thousand animals annually, indefinitely, without alleviating the vegetation impact.

Notable areas regarded as devastated by overabundant elephants include Tsavo East and Amboseli National Parks in Kenya and Chobe National Park in Botswana. In Tsavo, elephants transformed *Commiphora* shrubland into open grassland over a large area, reduced umbrella thorn trees (*Acacia tortilis*) to one-third of their former abundance, and almost eliminated baobab trees (*Adansonia digitata*; Leuthold, 1977). Browsers such as lesser kudu (*Tragelaphus imberbis*) and gerenuk (*Litocranius walleri*) declined in abundance, but grazers like zebra (*Equus quagga*) and oryx (*Oryx gazelle*) increased (Parker, 1983). Following the drastic reduction of the elephant population by poaching, the vegetation has largely recovered (Leuthold, 1996) and no animal populations suffered (Inamdar, 1996). Along the Chobe River, the extent covered by woodland was reduced from 60% to 30%, that covered by shrubland expanded from 5% to 30%, and the tall riparian woodland largely disappeared, following the buildup of high local concentrations of elephants (Mosugelo *et al.*, 2002). Plant diversity was greatly reduced, but the only animal species that suffered much local decline was bushbuck, reduced to a third or less of their former abundance (Skarpe *et al.*, 2004). Savanna woodland away from the river remained little affected (Ben-Shahar, 1993). At Amboseli the vegetation change was drastic, with the extent of fever acacia woodland vastly reduced, except where fenced exclosures kept out elephants (Western, 2007). However, the reduction in browser numbers as a consequence was not drastic, and historical evidence suggested that the development of the fever acacia stands was relatively recent and associated with settlement by pastoral people and their livestock.

As reported earlier, there are reasons to believe that the vegetation impacts of elephants and white rhinos can be beneficial for other large herbivores, through promoting the production of nutritious and accessible plant parts and by generating a mosaic diversity of habitat conditions. However, within local areas the vegetation transformation can reach extremes and can threaten the persistence of many plant species (O'Connor *et al.*, 2007), as well as certain types of animal. For example, the removal of trees is detrimental for the persistence of birds and bats dependent on woodland habitats (Cumming *et al.*, 1997a, 1997b), most notably vultures nesting in tall trees. Issues of spatial scale become central. If the area available is large enough to retain a mosaic of heavily and lightly disturbed areas, the overall biodiversity is likely to be enhanced. However, if the habitat disturbance becomes too extensive, many species may lose out, particularly certain plants. Hence, the most acute concerns arise with regard to protected areas under 1000 km^2 in extent (Boundja and Midgley, 2010). Nevertheless, it remains debatable whether even areas as large as Kruger National Park (19,000 km^2) are big enough to retain sufficient diversity once its elephant population has reached the habitat capacity, especially if surface water is sufficiently widely available for these animals to distribute their impacts throughout the park extent even in the dry season.

Meanwhile, elephant populations north of the Zambezi River remain threatened by illegal hunting supplying the continuing demand for ivory in the Far East, causing Kenya to oppose any change to existing agreements under CITES (the Convention on International Trade in Endangered Species), which block all trade in ivory. Opposing this is pressure from southern range states to allow sales of their mounting stocks of ivory held in vaults. India also strongly opposes any killing of elephants. Conservation issues concerning elephants are comprehensively summarized in the compendium by Scholes and Mennell (2008).

The white rhino population expansion toward 2000 animals within the 900 km^2 Hluhluwe-iMfolozi Park, with local densities exceeding five rhinos per square kilometer, raised similar concerns in a more local context (Owen-Smith, 1981, 1988). Culling was initiated by means of live capture and removal to restock other areas, including sales to private land owners. To simulate the role of dispersal in population regulation, the offtake is restricted to zones designated as dispersal sinks. A total of nearly 20,000 white rhinos have been generated from this source population, around half on private land (Emslie, 2008). The auction price fetched per rhino has reached US $30,000, and legislation allows for males to be shot by hunters, leading some ranchers to breed white rhinos for live sale. However, permits for ownership of hunting trophies have provided a cover for illegal exports of horns from South Africa (Milledge, 2007). Fueled by the extremely high prices paid for rhino horn scrapings used for medicinal purposes in the Far East, rhino poaching has recently increased explosively even in the best protected South African parks.

Black rhinos were formerly widespread through most of Africa, but their numbers have fallen to under 4500 animals, mostly located in South Africa and Namibia. The relatively low density attained by this browser (less than one animal per square kilometer) restricts the population that can be supported in any one area. Hence, a metapopulation management strategy has been introduced, entailing transferring animals from source populations to restock new areas, coupled with movements of breeding males between smaller areas to restrict inbreeding (http://www.rhino.sadc.org).

Indian rhinos remain severely threatened by poaching even in the heavily protected sanctuaries established in India and Nepal. Although some animals have been moved into newer parks, total numbers remain under 3000 animals. The Ujung Kulon Reserve retains no more than 50 Javan rhinos, whereas the few animals that had remained in Vietnam have been extirpated, making survival prospects for this species dismal.

The application of dispersal sink or metapopulation management to control wide-ranging elephants seems impractical even for parks as large as Kruger or Tsavo. The

initiatives most supportive for conserving elephants are transfrontier conservation areas (TFCAs) expanding the scope for elephant movements. Kruger Park has been joined with the Limpopo National Park in Mozambique plus private nature reserves to the west to form the Greater Limpopo Transfrontier Park covering 25,000 km^2, with prospects for further expansion in Mozambique and into Zimbabwe. Even more ambitious is the proposed Kavango-Zambezi (KAZA) TFCA aggregating almost 300,000 km^2 of parks, wildlife and forest reserves, game management areas, and communal land, extending from Botswana through Caprivi, Zimbabwe, Zambia, and Angola (www.peaceparks.org). Ways need to be found to enable elephants to move through settled areas between the protected areas. Studies in northern Kenya have shown how elephants alter their movements to avoid much contact with people (Graham *et al.*, 2009).

Despite these initiatives, a major conservation issue is the conflicts that arise between rural people and elephants with regard to crop damage, in India as well as in Africa (Sukumar, 2003; Dublin and Hoare, 2004). This tension is increasing where elephant populations are growing and expanding their range into settled areas. It is unclear how an eventual resolution will be achieved, except by fencing elephants out or people in and by allowing the killing of elephants transgressing boundaries.

Final Perspectives

Very large mammals falling within the megaherbivore size range (>1000 kg) were once a major feature of terrestrial faunas in all parts of the world. Megaherbivores still vastly dominate large herbivore biomass in part of Africa. Their grazing, browsing, and damaging impacts can transform uniform forests or grasslands into a mosaic interspersion of woodland, meadow, and savanna parkland. The enhanced habitat heterogeneity and productivity generated by such disturbance of vegetation can be favorable for the coexistence of numerous other mammalian herbivores, although plant diversity may be reduced locally.

Ecological conditions and processes operating within regions lacking megaherbivores differ vastly from those that prevailed in past times, before human hunting intruded. Whether the disappearance of these very large beasts is thankful or regrettable raises value judgments around human land-use objectives. The transforming influences of human settlements and cultivation on landscapes have largely replaced the impacts of megaherbivores. The ultimate resolution of the conflicts with humans where elephants are recovering their numbers and distribution ranges is unclear. From an African perspective, hankering for the restoration of mammoth equivalents in North American ecosystems appears crazy, given the problems that would ensue. Nevertheless, at some time in the future, the accumulated fossil fuels that we have depended on to energize our machines will have become irretrievably depleted. Once this situation develops, we may need to resort back to using these great beasts to perform essential services as environmental engineers and cultivators. Hence, we need to ensure their long-term persistence, at least in Africa.

See also: Africa, Ecosystems of. African and Asian Savannas. Climate Change and Extinctions. Endangered Mammals. Extinction, Causes of. Human Impact on Biodiversity, Overview. Keystone Species. Mammals, Conservation Efforts for. Mammals (Late Quaternary), Extinctions of. Mammals (Pre-Quaternary), Extinctions of. Mass Extinctions, Notable Examples of

References

Agenbroad LD (2003) North American proboscideans: Mammoths: The state of knowledge. *Quaternary International* 126–128: 73–92.

Alroy J (1999) Putting North America's end-Pleistocene megafaunal extinction in context. In: MacPhee RDS (ed.) *Extinctions in Near Time*, pp. 105–143. New York: Kluwer/Plenum.

Alroy J (2001) A multispecies overkill simulation of the end-Pleistocene megafaunal mass extinction. *Science* 292: 1893–1896.

Arsenault R and Owen-Smith N (2012) Competition and coexistence among short grass grazers in the Hluhuwe-iMfolozi Park, South Africa. *Canadian Journal of Zoology* 89: 900–907.

Barnovsky AD, Kock PL, Feranec RS, Wing SL, and Shabel AB (2004) Assessing the causes of late Pleistocene extinctions on the continents. *Science* 306: 70–75.

Barnovsky AD and Lindsey EL (2010) Timing of Quaternary megafaunal extinctions in South America in relation to human arrival and climate change. *Quaternary International* 217: 10–29.

Bell RHV (1981) An outline of a management plan for Kasungu National Park, Malawi. In: Jewell PA, Holt S, and Hart D (eds.) *Problems in Management of Locally Abundant Wild Mammals*, pp. 69–90. New York: Academic Press.

Ben-Shahar R (1993) Patterns of elephant damage to vegetation in northern Botswana. *Biological Conservation* 65: 249–256.

Berger J, Senson JE, and Persson I-L (2001) Recolonizing carnivores and naïve prey: Conservation lesson from Pleistocene extinctions. *Science* 291: 1036–1039.

Bond WJ (2005) Large parts of the world are brown or black: A different view on the 'Green World' hypothesis. *Journal of Vegetation Science* 16: 261–266.

Bond WJ and Loffell D (2001) Introduction of giraffe changes Acacia distribution in a South African savanna. *African Journal of Ecology* 39: 286–295.

Bond WJ, Midgley GF, and Woodward FI (2003) The importance of low atmospheric CO_2 and fire in promoting the spread of grasslands and savannas. *Global Change Biology* 9: 973–982.

Boundja RP and Midgley JJ (2010) Patterns of elephant impact on woody plants in the Hluhluwe-Imfolozi park, Kwazulu-Natal, South Africa. *African Journal of Ecology* 48: 206–214.

Bourliere F and Verschuren J (1960) Introduction a l'ecologie des ongules du Parc National Albert. Brussels: Institut des parcs nationaux du Congo belge.

Bowler JM, Johnston H, Olley JM, *et al.* (2003) New ages for human occupation and climate change at Lake Mungo, Australia. *Nature* 421: 837–840.

Bowman DMJS (1998) Tansley Review No. 101. The impact of Aboriginal landscape burning on the Australian biota. *New Phytologist* 140: 385–410.

Brain C, Forge O, and Erb P (1999) Lion predation on black rhinoceros in Etosha National Park. *African Journal of Ecology* 37: 107–109.

Burney DA and Flannery TF (2005) Fifty millennia of catastrophic extinctions after human contact. *Trends in Ecology and Evolution* 20: 395–401.

Carrasco MA, Barnovsky AD, and Graham RW (2009) Quantifying the extent of North American mammal extinction relative to the pre-anthropogenic baseline. *PLoS One* 4(12): e8331.

Cerling TE, Harris JM, and Leakey MG (1999) Browsing and grazing in elephants: the isotope record of modern and fossil proboscideans. *Oecologia* 120: 364–374.

Chamaille-Jammes S, Fritz H, Valeix M, Murindagomo F, and Clobert J (2008) Resource availability, aggregation and direct density dependence in an open context: the local regulation of an African elephant population. *Journal of Animal Ecology* 77: 135–144.

Chapman DG (1981) Evaluation of marine mammal population models. In: Fowler CW and Smith TD (eds.) *Dynamics of Large Mammal Populations*, pp. 277–296. New York: Wiley.

Clauss M, Frey R, Kieffer B, *et al.* (2003) The maximum attainable body size of herbivorous mammals: Morphological constraints on foregut, and adaptations of hindgut fermenters. *Oecologia* 136: 14–27.
Clauss M, Castell JC, Kienzle E, *et al.* (2007a) Mineral absorption in the black rhinoceros as compared with the domestic horse. *Journal of Animal Physiology and Animal Nutrition* 91: 193–204.
Clauss M, Streich WJ, Schwarn A, Ortmann S, and Hummel J (2007b) The relationship of food intake and ingesta passage predicts feeding ecology in two different megaherbivore groups. *Oikos* 116: 209–216.
Conybeare A and Haynes G (1984) Observations on elephant mortality and bones in water holes. *Quaternary Research* 22: 189–200.
Corfield TT (1973) Elephant mortality in Tsavo National Park, Kenya. *East African Wildlife Journal* 11: 339–368.
Cumming DHM (1981) The management of elephant and other large mammals in Zimbabwe. In: Jewell PA, Holt S, and Hart D (eds.) *Problems in Management of Locally Abundant Wild Mammals*, pp. 91–118. New York: Academic Press.
Cumming DHM, Fenton MB, Rautenbach IL, *et al.* (1997a) Elephants, woodlands and biodiversity in southern Africa. *South African Journal of Science* 93: 231–236.
Cumming DHM, Fenton MB, Rautenbach IL, *et al.* (1997b) Elephants, woodlands and biodiversity in southern Africa. *South African Journal of Science* 93: 231–236.
Demment MW and Van Soest PJ (1985) A nutritional explanation for body-size patterns of ruminant and nonruminant herbivores. *American Naturalist* 125: 641–672.
Dinerstein E (2003) *The Return of the Unicorns. The Natural History and Conservation of the Greater One Horned Rhinoceros.* New York: Columbia University Press.
Donlan CJ, Berger J, Bock CE, *et al.* (2006) Pleistocene rewilding: an optimistic agenda for twenty-first century conservation. *The American Naturalist* 168: 660–681.
Donlan J, Green HW, Berger J, *et al.* (2005) Re-wilding North America. *Nature* 436: 913–914.
Dublin HT and Hoare RE (2004) Searching for solutions: the evolution of an integrated approach to understanding and mitigating human–elephant conflicts in Africa. *Human Dimensions of Wildlife* 9: 271–278.
Dublin HT, Sinclair ARE, and McGlade J (1990) Elephants and fire as causes of multiple stable states in the Serengeti-Mara woodlands. *Journal of Animal Ecology* 59: 1147–1164.
Dudley JP, Craig GC, Gibson DSC, Haynes G, and Klimowicz J (2001) Drought mortality of bush elephants in Hwange National Park, Zimbabwe. *African Journal of Ecology* 39: 187–194.
Edkins MT, Kruger LM, Harris K, and Midgley JJ (2007) Baobabs and elephants in Kruger National Park: nowhere to hide. *African Journal of Ecology* 46: 119–125.
Emslie R (2008) Rhino population sizes and trends. *Pachyderm* 44: 89–95.
Finlayson C (2005) Biogeography and evolution of the genus Homo. *Trends in Ecology and Evolution* 20: 457–463.
Flannery TF and Roberts RG (1999) Late Quaternary extinctions in Australasia: An overview. In: MacPhee RDE (ed.) *Extinctions in Near Time: Causes, Contexts and Consequences*, pp. 239–256. New York: Kluwer Academic/Plenum Publishers.
Foley C, Pettorelli N, and Foley L (2008) Severe drought and calf survival in elephants. *Biology Letters* 4: 541–544.
Fornara DA and du Toit JT (2007) Browsing lawns? Responses of *Acacia nigrescens* to ungulate browsing in an African savanna. *Ecology* 88: 200–209.
Fowler CW (1987) A review of density dependence in populations of large mammals. In: Genoways HH (ed.) *Current Mammalogy,* Vol. 1. pp. 401–441. New York: Plenum Press.
Gill JL, Williams JW, Jackson ST, Lininger KB, and Robinson GS (2009) Pleistocene megafaunal collapse, novel plant communities, and enhanced fire regimes in North America. *Science* 326: 1100–1103.
Goddard J (1976) Home range, behaviour and recruitment rates of two black rhinoceros populations. *East African Wildlife Journal* 5: 133–150.
Goebel T, Waters MR, and O'Rourke DH (2008) The late Pleistocene dispersal of modern humans in the Americas. *Science* 319: 1497–1502.
Gough KF and Kerley GIH (2006) Demography and population dynamics in the elephants of Addo National Park, South Africa: Any evidence of density dependent regulation? *Oryx* 40: 434–441.
Graham MD, Douglas-Hamilton I, Adams WM, and Lee PC (2009) The movement of African elephants in a human-dominated landscape. *Animal Conservation* 12: 4445–4455.
Guilday JE (1984) Pleistocene extinctions and environmental change: Case study of the Appalachians. In: Martin PS and Klein RG (eds.) *Quaternary Extinctions*, pp. 250–258. Tucson: University of Arizona Press.
Guthrie RD (1984) Mosaics, allelchemics and nutrients. An ecological theory of late Pleistocene megafaunal extinctions. In: Martin PS and Klein RG (eds.) *Quaternary Extinctions*, pp. 259–298. Tucson: University of Arizona Press.
Haile J, Froese DG, MacPhee RDE, *et al.* (2009) Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences of the USA* 106: 22352–22357.
Hairston NG, Smith FE, and Slobodkin LB (1960) Community structure, population control and competition. *American Naturalist* 94: 421–425.
Hanks J (1969) Growth in weight of the female African elephant in Zambia. *East African Wildlife Journal* 7. 7-1.
Hitchins PM and Anderson JL (1983) Reproduction, population characteristics and management of the black rhinoceros in the Hluhluwe/Corridor/Umfolozi Game Reserve. *South African Journal of Wildlife Research* 13: 78–85.
Hofmann RR (1989) Evolutionary steps of ecophysiological adaptation and diversification of ruminants: A comparative review of their digestive system. *Oecologia* 78: 443–457.
Hofmann RR and Stewart DRM (1972) Grazer or browser: a classification based on the stomach structure and feeding habits of East African ruminant animals. *Mammalia* 36: 226–240.
Holdo RM (2007) Elephants, fire and frost can determine community structure and composition in Kalahari woodlands. *Ecological Applications* 17: 558–568.
Holdo RM, Dudley JP, and McDowell LR (2002) Geophagy in the African elephant in relation to the availability of dietary sodium. *Journal of Mammalogy* 83: 652–664.
Hooijer DA and Patterson B (1972) Rhinoceroses from the Pliocene of north-western Kenya. *Bulletin of the Museum of Comparative Zoology (Harvard University* 144: 1–26.
Hopcraft JGC (2010) Herbivores, resources and risk: Alternating regulation along primary environmental gradients in savannas. *Trends in Ecology and Evolution* 25: 119–128.
Inamdar A (1996) *The ecological consequences of elephant depletion.* Phd thesis, Cambridge University.
Janzen DH (1979) New horizons in the biology of plant defenses. In: Rosenthal GA and Janzen DH (eds.) *Herbivores. Their Interactions with Secondary Plant Metabolites*, pp. 331–350. New York: Academic Press.
Jarman PJ (1974) The social organization of antelope in relation to their ecology. *Behaviour* 48: 215–267.
Johnson CN (2002) Determinants of loss of mammal species during the late Quaternary "megafauna" extinctions: Life history and ecology, but not body size. *Proceedings of the Royal Society of London Series B* 269: 2221–2227.
Johnson CN (2009) Ecological consequences of late Quaternary extinctions of megafauna. *Proceedings of the Royal Society Series B* 276: 2509–2519.
Joubert D (1986) Hunting behaviour of lions on elephants in the Chobe National Park, Botswana. *African Journal of Ecology* 44: 279–281.
Joubert E and Eloff F (1971) Notes on the ecology and behaviour of the black rhinoceros in South West Africa. *Madoqua* 1: 31–36.
Junker J, van Aarde RJ, and Ferreira SM (2008) Temporal trends in elephant numbers and densities in northern Botswana: Is the population really increasing? *Oryx* 42: 58–65.
Koch PL and Barnovsky AD (2006) Late Quaternary extinctions: State of the debate. *Annual Review of Ecology and Systematics* 37: 215–250.
Lang H (1923) Recent and historical notes on the square-lipped rhinoceros. *Journal of Mammalogy* 4: 155–163.
Langer P (1988) The mammalian herbivore stomach. *Comparative Anatomy, Function and Evolution.* Stuttgart: Gustav Fischer.
Laws RM (1969) The Tsavo research project. *Journal of Reproduction and Fertility Supplement* 6: 495–531.
Laws RM (1970) Elephants as agents of habitat and landscape change in East Africa. *Oikos* 21: 1–15.
Laws RM and Parker ISC (1968) Recent studies on elephant populations in East Africa. In: Crawford MA (ed.) *Comparative Nutrition of Wild Animals*, pp. 319–359. Symposium of the Zoological Society of London No. 21
Leakey LB (1965) *Olduvai Gorge 1951–1961, Vol 1. A Preliminary Report on the Geology and Fauna.* Cambridge: Cambridge University Press.
Leggett KEA (2006) Home range and seasonal movement of elephants in the Kunene Region, northwestern Namibia. *African Zoology* 41: 17–36.
Leuthold W (1977) Changes in tree populations in Tsavo East National Park, Kenya. *East African Wildlife Journal* 15: 61–69.
Leuthold W (1996) Recovery of woody vegetation in Tsavo National Park, Kenya, 1970–1994. *African Journal of Ecology* 34: 101–112.

Levick S and Rogers K (2008) Patch and species specific responses of savanna woody vegetation to browser exclusion. *Biological Conservation* 141: 489–498.
Levick SR, Asner GP, Kennedy-Bowdoin T, and Knapp DE (2009) The relative influence of fire and herbivory on savanna three-dimensional vegetation structure. *Biological Conservation* 142: 1693–1700.
Lindeque PM and Turnbull PCB (1994) Ecology and epidemiology of anthrax in the Etosha National Park, Namibia. *Onderstepoort Journal of Veterinary Research* 61: 71–83.
Lindstrom J and Kokko H (2002) Cohort effects and population dynamics. *Ecology Letters* 5: 338–344.
Loarie SR, van Aarde RJ, and Pimm SL (2009) Elephant seasonal vegetation preferences across dry and wet savannas. *Biological Conservation* 142: 3099–3107.
Lock JM (1972) The effect of hippopotamus grazing in grasslands. *Journal of Ecology* 60: 445–467.
Loutit BD, Louw GN, and Seely MK (1987) First approximation of food preferences and the chemical composition of the diet of the desert-dwelling black rhinoceros. *Madoqua* 15: 35–54.
Loveridge AJ, Hunt JE, Murindagomo F, and Macdonald DW (2006) Influence of drought on predation of elephant calves by lions in an African savanna woodland. *Journal of Zoology* 270: 523–530.
Luske BL, Martens T, Lent PC, De Boer WF, and Prins HHT (2009) Impact of the black rhinoceros on a local population of *Euphorbia bothae* in the Great Fish River Reserve, South Africa. *African Journal of Ecology* 47: 509–517.
Lyons SK, Smith FA, and Brown JH (2004) Of mice, mastodons and men: Human-mediated extinction on four continents. *Evolutionary Ecology Research* 6: 339–358.
MacFadden BJ (2000) Cenozoic mammalian herbivores from the Americas. Reconstructing ancient diets and terrestrial communities. *Annual Review of Ecology and Systematics* 31: 33–59.
Makhabu SW, Skarpe C, and Hytteborn H (2006) Elephant impact on shoot distribution on trees and rebrowsing by smaller browsers. *Acta Oecologica* 30: 136–146.
Martin P (2005) *Twilight of the Mammoths. Ice Age Extinctions and the Rewilding of North America*. Berkeley: University of California Press.
Martin PS (1966) Africa and Pleistocene overkill. *Nature* 212: 339–342.
Martin PS (1967) Prehistoric overkill. In: Martin PS and Klein RG (eds.) *Quaternary Extinctions*, pp. 75–120. Tucson: University of Arizona Press.
McShane TO (1989) Some preliminary results of the relationship between soils and tree responses to elephant damage. *Pachyderm* 11: 29–31.
Milledge SAH (2007) Illegal killing of African rhinos and the horn trade. *Pachyderm* 43: 96–107.
Miller GH, Magee JW, Johnson BJ, *et al.* (1999) Pleistocene extinction of *Genyornis newtoni*: Human impact on Australian megafauna. *Science* 283: 205–208.
Moss CJ (2001) The demography of an African elephant population in Amboseli, Kenya. *Journal of Zoology* 255: 145–156.
Mosugelo DK, Noe SR, Ringrose S, and Nellermann C (2002) Vegetation changes during a 36-year period in northern Chobe National Park, Botswana. *African Journal of Ecology* 40: 232–240.
Myers N (1973) Tsavo National Park, Kenya, and its elephants: An interim appraisal. *Biological Conservation* 5: 123–132.
O'Connor TG, Goodman PS, and Clegg B (2007) A functional hypothesis of the threat of local extirpation of woody plant species by elephant in Africa. *Biological Conservation* 136: 329–345.
O'Connell-Rodwell CE, Rodwell T, Rice M, and Hart LA (2000) Living with the modern conservation paradigm: Can agricultural communities coexist with elephants? A five-year case study in East Caprivi, Namibia. *Biological Conservation* 93: 381–391.
Olff H, Vera FWM, Bokdam J, *et al.* (1999) Shifting mosaics in grazed woodlands driven by the alternation of plant facilitation and competition. *Plant Biology* 1: 127–137.
Olivier RCD and Laurie WA (1974) Habitat utilization by hippopotamus in the Mara River. *East African Wildlife Journal* 12: 249–272.
Orlova LA, Kuzmin YV, and Dementiev VN (2004) A review of the evidence for extinction chronologies for five species of Upper Pleistocene megafauna in Siberia. *Radiocarbon* 46: 301–314.
Owen-Smith N (1975) The social ethology of the white rhinoceros. *Zeitschrift fur Tierpsychologie* 38: 337–384.
Owen-Smith N (1981) The white rhinoceros overpopulation problem, and a proposed solution. In: Jewell PA, Holt S, and Hart D (eds.) *Problems in Management of Locally Abundant Wild Mammals*, pp. 129–150. New York: Academic Press.
Owen-Smith N (1987) Late Pleistocene extinctions: The pivotal role of megaherbivores. *Paleobiology* 13: 351–362.
Owen-Smith N (1988) *Megaherbivores. The Influence of Very Large Body Size on Ecology*. Cambridge: Cambridge University Press.
Owen-Smith N (1989) Megafaunal extinctions: The conservation message from 11 000 years BP. *Conservation Biology* 3: 405–412.
Owen-Smith N (1999) The interaction of humans, megaherbivores and habitats in the late Pleistocene extinction event. In: MacPhee RDE (ed.) *Extinctions in Near Time: Causes, Contexts and Consequences*, pp. 57–69. New York: Kluwer Academic/Plenum Publishers.
Owen-Smith N (2003) Elephants and ecosystems. In: Vandewalle M (ed.) *Effects of Fire, Elephants and other Herbivores on the Chobe Riverfront Ecosystem, Proceedings of a Conference organised by the Botswana–Norway Institutional Co-operation and Capacity Building Project*, pp. 16–22. Gaborone: Botswana Department of Wildlife and National Parks.
Owen-Smith N and Chafota J (in press) Selective feeding by a mega-herbivore, the African elephant. *Journal of Mammalogy*.
Owen-Smith N, Kerley GIH, Page B, Slotow R, and van Aarde RJ (2006) A scientific perspective on the management of elephants in the Kruger National Park and elsewhere. *South African Journal of Science* 102: 389–394.
Owen-Smith N and Mills MGL (2008) Predator-prey size relationships in an African large-mammal food web. *Journal of Animal Ecology* 77: 173–183.
Parker ISC (1983) The Tsavo story: An ecological case history. In: Owen-Smith RN (ed.) *Management of Large Mammals in African Conservation Areas*, pp. 37–50. Pretoria: Haum.
Pienaar DJ, Bothma JduP, and Theron GK (1993) Landscape preference of the white rhinoceros in the central and northern Kruger National Park. *Koedoe* 36: 79–86.
Pienaar UdeV (1970) The recolonisation history of the square-lipped (white) rhinoceros in the Kruger National Park (October 1961–November 1969). *Koedoe* 13: 157–169.
Pienaar UdeV (1983) Management of intervention: the pragmatic/economic option. In: Owen-Smith RN (ed.) *Management of Large Mammals in African Conservation Areas*, pp. 23–36. Pretoria: Haum.
Power RJ and Compion RXS (2009) Lion predation on elephants in the Savuti, Chobe National Park, Botswana. *African Zoology* 44: 36–44.
Prado JL, Alberdi MJ, Azanya B, Sanchez B, and Frassinetti D (2005) The Pleistocene Gomphotheriidae (Proboscidea) from South America. *Quaternary International* 126–128: 21–30.
Roberts RG, Flannery TF, Ayliffe LK, *et al.* (2001) New ages for the last Australian megafauna: continent-wide extinction about 46,000 years ago. *Science* 292: 1888–1892.
Robinson GS, Burney LP, and Burney DA (2005) Landscape paleoecology and megafaunal extinction in southeastern New York state. *Ecological Monographs* 75: 295–315.
Rohland N, Reich D, Mallick, *et al.* (2010) Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLoS Biology* 8(12): e1000564. http://dx.doi.org/10.1371/journal.pbio.1000564
Rutina LP, Moe SR, and Swenson JE (2005) Elephant-driven woodland conversion to shrubland improves dry-season browse availability for impalas. *Wildlife Biology* 11: 207–213.
Scholes RJ and Mennell KG (2008) *Elephant Management. A Scientific Assessment for South Africa*. Johannesburg: University of the Witwatersrand Press.
Selous FC (1899) The white or square-lipped rhinoceros. In: Bryden HA (ed.) *Great and Small Game of Africa*, pp. 52–67. London: Rowland Ward.
Simmons AH (1988) Extinct pygmy hippotamus and early man in Cyprus. *Nature* 333: 554–557.
Sinclair ARE (1975) The resource limitation of trophic levels in tropical grassland ecosystems. *Journal of Animal Ecology* 44: 497–520.
Sitati N (2007) Challenges and partnerships in elephant conservation and conflict mitigation. *Oryx* 41: 135–139.
Skarpe C, Aarresstad PA, Andreassen HP, *et al.* (2004) The return of the giants: Ecological effects of an increasing elephant population. *Ambio* 33: 276–282.
Smart NOE, Hatton JC, and Spence DHN (1985) The effect of long-term exclusion of large herbivores on vegetation in Murchison Falls National Park, Uganda. *Biological Conservation* 33: 229–245.
Spinage C (1994) *Elephants*. London: Poyser.
Steuer P, Clauss M, Sudekum K-H, *et al.* (2010) Comparative investigation on digestion in grazing and browsing rhinoceroses. *Comparative Biochemistry and Physiology, Part A* 156: 380–388.
Stuart AJ (1991) Mammalian extinctions in the late Pleistocene of northern Eurasia and North America. *Biological Reviews* 66: 453–562.
Stuart AJ (2005) The extinction of woolly mammoth and straight-tusked elephant in Europe. *Quaternary International* 126–128: 171–177.

Sukumar R (1989) *The Asian Elephant. Ecology and Management*. Cambridge: Cambridge University Press.
Sukumar R (2003) *The Living Elephants. Evolutionary Ecology, Behavior and Conservation*. Oxford: Oxford University Press.
Svenning J-C (2002) A review of natural vegetation openness in north-western Europe. *Biological Conservation* 104: 133–148.
Svenning J-C (2007) 'Pleistocene rewilding' merits serious consideration also outside North America. 〈http://biogeography.blogspot.com/2007/10/pleistocene-re-wilding-merits-serious.html〉.
Todd NE (2005) Reanalysis of African *Elephas recki*: Implications for time, space and taxonomy. *Quaternary International* 126–128: 65–72.
Trimble MJ, Fereira SM, and van Aarde RJ (2009) Drivers of megaherbivore demographic fluctuations: inferences from elephants. *Journal of Zoology* 279: 18–26.
Turner A and Anton M (2004) *Evolving Eden. An Illustrated Guide to the Evolution of African Large Mammals*. New York: Columbia University Press.
van Aarde RJ and Jackson T (2007) Megaparks for megapopulations: Addressing the causes of locally high elephant numbers in South Africa. *Biological Conservation* 134: 289–297.
van der Wal R, Bardgett RD, Harrison KA, and Stien A (2004) Vertebrate herbivores and ecosystem control: cascading effects of faeces on tundra ecosystems. *Ecography* 27: 242–252.
van Gysegem R (1984) Observations on the ecology and behaviour of the northern white rhinoceros. *Zeitschrift fur Saugetierkunde* 49: 348–358.
van Wyk P and Fairall N (1969) The influence of the African elephant on the vegetation of the Kruger National Park. *Koedoe* 12: 57–89.
Vartanyan SL, Garutt VE, and Sher AV (1993) Holocene dwarf mammoths from Wrangel Island in the Siberian Arctic. *Nature* 362: 337–340.
Vera FWM, Bakker ES, and Olff H (2006) Large herbivores: Missing partners of western European light-demanding tree and shrub species? In: Danell K, Bergstrom R, Duncan P, and Pastor J (eds.) *Large Herbivore Ecology, Ecosystem Dynamics and Conservation*, pp. 203–231. Cambridge: Cambridge University Press.
Verweij RJT, Verrelst J, Loth PE, Heitkonig IMA, and Brunsting AMH (2006) Grazing lawns contribute to the subsistence of mesoherbivores on dystrophic grasslands. *Oikos* 114: 108–116.
Wafula MM, Atimnedi P, and Turiwesigye C (2007) Managing the 2004/05 anthrax outbreak in Queen Elizabeth and Lake Mburo National Parks, Uganda. *African Journal of Ecology* 46: 24–31.
Waldram MS, Bond WJ, and Stock WD (2007) Ecological engineering by a mega-grazer: White rhino impacts on a South African savanna. *Ecosystems* 11: 101–112.
Weir JS (1972) Spatial distribution of elephants in an African national park in relation to environmental sodium. *Oikos* 23: 1–13.
Western D (2007) A half-century of habitat change in Amboseli. *African Journal of Ecology* 45: 302–310.
Whyte IJ (2001) Headaches and heartaches: The elephant management dilemma. In: Schmidtz D and Willot E (eds.) *Environmental Ethics: Introductory Readings*, pp. 293–305. New York: Oxford University Press.
Wiseman R, Page BR, and O'Connor TG (2004) Woody vegetation change in response to browsing in Ithala Game Reserve, South Africa. *South African Journal of Wildlife Research* 34: 25–37.
Wittemyer G, Rasmussen HB, and Douglas-Hamilton I (2007) Breeding phenology in relation to NDVI variability in free-ranging African elephant. *Ecography* 30: 42–50.
Wright, Jr. HE (ed.) (1984) *Late Quaternary Environments of the United States. Vol 1. The Late Pleistocene*. Minneapolis: University of Minneapolis Press.
Young KD, Ferreira SM, and van Aarde RJ (2009) Elephant spatial use in wet and dry savannas of southern Africa. *Journal of Zoology* 278: 189–205.
Young KD and van Aarde RJ (2010) Density as an explanatory variable of movements and calf survival in savanna elephants across southern Africa. *Journal of Animal Ecology* 79: 662–673.
Zimov SA, Chuprynin VI, Oreshko AP, *et al.* (1995) Steppe-tundra transition: a herbivore-driven biome shift at the end of the Pleistocene. *American Naturalist* 146: 765–794.

Metapopulations

Peter Chesson, The University of Arizona, Tucson, AZ, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Local population That part of the population of a species found in a particular habitat patch.

Metapopulation The collection of local populations in a region.

Patch An area of suitable habitat for a particular species or particular collection of species, ideally bounded by unsuitable habitat, or habitat with different physical properties. Normally one of many such areas in a region.

Spatially structured population A population whose reproductive and survival rates vary over the region that it inhabits reflecting variation in the physical and biological properties of the localities within the region.

Strict metapopulation (Also called a "classical metapopulation.") A metapopulation satisfying the following conditions:

1. Local populations are partially isolated from one another and are frequently capable of sustaining themselves for several to many generations in the absence of immigration from other local populations.
2. Local population extinction occurs on a time scale of several to many generations.
3. Migration between local populations leads to reestablishment of local populations following local extinction.

Outline

The study of metapopulations is concerned with the patchiness of populations in space, and the role of this patchiness in population dynamics, population stability, coexistence of species, and the maintenance of diversity. Strict metapopulation studies focus on patchiness due to colonization and extinction of local populations in a region. They have important applications in conservation biology, especially concerning the effects of habitat fragmentation. Studies of metapopulations, and their generalization to spatially structured populations, emphasize that patchiness markedly alters population dynamics and the outcomes of species interactions. These findings have implications for the concepts of equilibrium, nonequilibrium dynamics and population regulation, and hence for management practices influenced by these concepts.

Introduction

The concept of a metapopulation has its beginnings in the suggestion of Andrewartha and Birch and others that some, if not most, local populations of organisms in nature, frequently go extinct, and are re-established later by immigration from surrounding areas (Hanski, 1999). The metapopulation – the collection of all local populations – only persists if local extinction is balanced by recolonization. How this balance is achieved is an important focus of metapopulation theory. Because many human activities fragment natural habitats, humans may artificially create metapopulations or may decrease the density of local populations within existing metapopulations, and may put the continued survival of natural populations at risk. Hence metapopulations are a major topic in conservation biology. Regardless of whether a natural population is distributed over easily defined patches of habitat on which distinct local populations may be recognized, essentially all natural populations are patchily distributed in space, and most ecologists believe that patchiness in space and time has functional roles in population dynamics, i.e., in the manner in which population densities change over time. Most important, population dynamics at the regional or metapopulation scale are affected by the patchiness of the populations and their physical environment at lower spatial scales.

One very obvious way that patchiness is important is through mate finding. A sparse population evenly distributed over an area may have a low reproductive rate because males and females encounter each other too infrequently for many eggs to be fertilized, or flowers pollinated. Patchiness can be a solution to this problem. High local concentrations or aggregations of a species remove the problem of low encounter rates of males and females consistent with a low average density. Similar effects occur from predation. Individuals may find safety from predators in a group but be vulnerable when isolated. Such effects of sparse populations are called Allee effects and are potentially one reason why patchiness in space is important. However, disadvantages from clumped distributions in space might occur too, for example, due to competition. Clumps of individuals in space deplete resources locally and reduce population growth rates over what would be experienced were the population more evenly distributed. A similar effect results if predators are attracted to clumps of prey. Rather than prey finding safety in numbers, it is possible that they may be more vulnerable in clumps if predators are able to build up in numbers at clumps of prey.

These explanations for the importance of patchy distributions involve interactions between organisms in the same species, and therefore involve density dependence: dependence of the probability that an individual survives, or dependence of its reproductive rate on population density. The most striking theoretical predictions for the effects of patchiness, however, are for interactions between species, one or more of which is patchily distributed. Major effects of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00172-6

patchiness on population dynamics also arise from spatial and temporal variation in the physical environment. Naturally, if species are concentrated in better localities, they will have higher growth rates in comparison with random distributions, but such environmental variation has its most striking effects by altering the interactions between organisms, locally in space, and hence by altering the collective outcome of those interactions at the level of regional populations. These ideas are all studied within the context of metapopulations, but extend to spatially structured populations generally (Chesson *et al.*, 2005). Metapopulation concepts were originally developed with terrestrial systems in mind (Hanski, 1999), but these ideas have been strongly embraced to understand the ecology of marine systems as well (Kritzer and Sale, 2006).

Strict Metapopulations

The idea of a metapopulation is most commonly invoked when patchiness in space is so extreme that the regional population splits into local populations, each of which is too small to persist indefinitely (Hanski and Gilpin, 1997; Hanski, 1999). Local populations go extinct, but may be recolonized by migration from other local populations. In the strict sense of a metapopulation, local populations are sustained primarily by reproduction of resident individuals but may be subsidized by infrequent immigration from other local populations. The role of immigration, however, is seen mainly as allowing recolonization of a local population after it has become extinct. Local extinction may occur for any of a variety of reasons, as discussed below, but is normally assumed to occur asynchronously in different local populations, or the metapopulation as a whole would be lost. Although nearly all organisms are patchy in space, there is some question of how frequently all the above details of a strict metapopulation are satisfied. Thus, in considering the results of metapopulation theory it is important to distinguish those features applicable only to strict metapopulations, and those applying more generally to spatially patchy populations.

Single Species Considerations

The main interest in metapopulations, strictly defined, is with colonization and extinction dynamics. The variable of prime concern is the proportion, p, of patches (places where a local population potentially could exist) that are occupied by the species. Two parameters are involved with the dynamics of this occupancy fraction: the extinction rate of local populations, ε, and the rate c at which empty patches are recolonized, per fraction occupied. In essence, the fraction of local populations going extinct in one unit of time is ε. The probability of an extinct population being recolonized in one unit of time is cp, because c is measured proportional to the fraction of occupied patches and incorporates the idea that the probability of recolonization should depend on the fraction of other patches that are occupied. Thus, the change in the fraction occupied with time may be expressed as the following differential equation, called the Levin's equation

$$\frac{dp}{dt} = cp(1-p) - \varepsilon p$$

If the rates ε and c remain constant with time, it follows that in a metapopulation with a large number of local populations, the fraction occupied will come to an approximate equilibrium at the value $1-\varepsilon/c$, provided $c>\varepsilon$, because at this value extinction just balances recolonization. If ε is greater than c, local extinction exceeds local colonization and the metapopulation as a whole goes extinct. Conversely c must exceed ε for the metapopulation to persist. This result is often called the threshold condition, and is of major concern for applications of metapopulation theory in conservation. Because actual metapopulations in nature never have infinitely many local populations, the system is never exactly at the equilibrium $1-\varepsilon/c$, but may fluctuate around this value. In metapopulations with only a few local populations, there is a danger that all local populations will simultaneously go extinct, causing the extinction of the metapopulation even though $c>\varepsilon$.

On what biological features do these colonization and extinction rates depend? Consider extinction. Local extinction can occur in many ways, some of which stem directly from the activities of humans in habitat destruction or modification and hunting, but the interest in metapopulation models is the case where local extinction is a natural and repeated phenomenon, whose rates may be influenced by human activities, but whose presence is not. Thus, local populations may become extinct by chance because of small population size. The phenomenon of demographic stochasticity refers to independent chance events in the lives of individual organisms. It is impossible to predict how long any individual will live, and how many offsprings it will have, and these effects summed over individuals in a population cause it to fluctuate. In a small local population, there is always a definite probability of below replacement reproduction in any year; and in a small population, a chance run of such years can lead to extinction. These probabilities drop off dramatically as population size increases, but there are other means that can also lead to local population crashes, including unstable population dynamics and disturbance.

Two different forms of disturbance are commonly considered. An environmental disturbance is caused by an environmental event such as a storm or fire, and a biological disturbance is caused by the invasion of predators or disease. In the strict metapopulation scenario, the interest is in disturbance events that lead to extinction of the local population either coincident with the disturbance or shortly after, following weakening of the population by disturbance. At any time, disturbances must strike patchily in space to cause extinction of only a proportion of local populations, or the metapopulation as a whole would be lost. It is very common for environmental disturbance agents such as fires and storms to distribute mortality very patchily in space even though they may affect a large area in a short period of time.

In a strict metapopulation, to which Levin's equation applies, recolonization of a patch is from immigration from occupied patches. Many organisms may have means of recovering from disturbance by regrowth from seeds, roots, resting eggs, or spores within a recently disturbed patch, but with a strict metapopulation, such populations have not gone extinct. They have merely died back, and regenerate from dormant stages or below ground parts. The strict metapopulation idea that

colonization is from external sources is the reason for the formula cp for the recolonization probability. Regeneration from sources within a patch should not depend on p. Confusing regeneration with recolonization could potentially lead to serious errors in calculations and inferences.

Recolonization can be from mobile individuals, seeds, spores, or other propagules that arrive from external sources in sufficient quantity to re-establish a local population. The value of c will be affected by a variety of factors such as the spacing of local populations. With large distances between local populations, many propagules may perish after leaving their source population before arriving in suitable habitat. The sizes of local populations will also influence c because larger local populations should be expected to send out more propagules. The size of the metapopulation, in terms of the number of local populations, is also important. Because the number of habitat patches is always finite, the fraction of patches occupied is not at the equilibrium value $1-\varepsilon/c$, but fluctuates around this value; the smaller the number of habitat patches, the larger these fluctuations. If the number of habitat patches is too small, there is a danger of simultaneous extinction of all local populations eliminating the metapopulation even if $c>\varepsilon$.

Habitat fragmentation by human activities may lead to a metapopulation structure of local populations even if that were not the pre-existing situation for the natural population of these organisms. With continuing habitat destruction, c will decline due to increased spacing and smaller size of suitable patches, with the danger that c will drop below ε, especially as ε itself may increase as suitable patches decrease in size and support smaller local populations. Thus, the danger exists that even though there is habitat capable of sustaining local populations, local populations may be either small or so vulnerable because of other habitat alterations that the rate of local extinction exceeds local recolonization, which would lead to loss of the metapopulation altogether. Alternatively, if the organisms are not able to disperse between habitat patches, habitat fragmentation is not accompanied by positive colonization rates, and the metapopulation heads to extinction at the rate ε the moment it is created.

The simple destruction of habitat reduces natural population sizes, putting species at risk independently of these metapopulation considerations, but metapopulation theory emphasizes that the viability of populations in the remaining habitat may be seriously affected by its connectedness. A final feature of habitat destruction is that by reducing the areal extent of a population, it increases the probability that extinction events occur simultaneously in different parts of the metapopulation. For example, fire might sweep through the whole area occupied by the metapopulation, or climate fluctuations might have severe effects in all parts of the metapopulation, endangering the metapopulation as a whole.

So far, the rates c and ε have been treated as though they are independent of p, the fraction occupied. This need not be the case. It has been suggested, for example, that extinction rates may decrease as the fraction occupied increases because extant local populations may tend to be higher and less vulnerable to extinction when they are subsidized by immigration. This could be especially important in the recovery of local populations after chance population declines, staving off extinction. For this reason, this phenomenon is called the rescue effect. If the rescue effect is required for c to exceed ε, the metapopulation as a whole may be especially vulnerable to extinction due to chance fluctuations in the fraction occupied, for if this fraction falls too low ε will exceed c promoting further decline and total extinction of the metapopulation.

There are many other details and complications that one can consider in single-species metapopulations such as variation in the sizes of local populations, their distances apart, and variation in habitat quality, features that are discussed in several excellent books on the subject (Hanski and Gilpin, 1997; Hanski, 1999; Hanski and Gaggiotti, 2004; Kritzer and Sale, 2006). These complications, naturally lead to much more complex equations than the Levin's equation given here. The fraction of patches occupied is no longer the key variable in the equations. Instead it is replaced by the probability, unique for each patch occupied. Under certain conditions, this probability is equal to the incidence of occupancy of the patch, which is defined as the fraction of time that the patch is occupied (Hanski, 1999). These sorts of modifications are especially useful when metapopulation theory is applied to a particular natural system when the reality of nature's complexity must be confronted (Hanski and Gaggiotti, 2004).

Multispecies Considerations: Predators and Parasitoids

Extinction and colonization rates may depend on other organisms. Indeed, as mentioned above, local extinction may be a consequence of invasion of the local population by predators or disease. In this case, extinction rates vary with the abundance of predators within the metapopulation, which in turn depend on the abundance of the prey itself. It is quite possible that only the prey has a metapopulation structure, because predators may be more mobile occupying all of the landscape. Their density must be taken into account in determining the extinction rate of the prey but otherwise their inclusion does not lead to major differences in the understanding of metapopulation dynamics, except potentially in terms of the effects of habitat destruction, as discussed below.

If both predator and prey have the same metapopulation structure, a metacommunity results (Holyoak *et al.*, 2005). Simple models of isolated local predator–prey communities are often unstable, causing large fluctuations in their abundances and potentially leading to extinction of the prey or the predator (see **Figure 7**). A metacommunity structure has long been suggested as one means by which predator–prey interactions may be stabilized in nature. The interaction between the two locally in space remains unstable: the prey is driven extinct by the predator, but provided the colonization rate c for the prey exceeds its extinction rate due to predation and other causes, it is able to persist in a metapopulation. Similarly, although the predator depends on the prey whose local extinction leads to local predator extinction, provided the predator can find patches sufficiently well, it can persist in the system also. In this way, it is possible for predator and prey to have a stable equilibrium at the metapopulation level in spite of what might be unstable dynamics on the local scale (Nee *et al.*, 1997). This metacommunity structure may have important conservation implications. Habitat destruction may

more strongly affect the predator than the prey because prey numbers in a region may be buffered at first by reduction in predator numbers, with the effects of habitat destruction on prey being stronger after the predator has been eliminated (Nee *et al.*, 1997).

The potential for metacommunity structure for predator and prey is also of great interest in the study of biological control of insect pests and introduced weeds. Commonly, in biological control, one seeks as a control agent a predator or parasitoid (a parasite that invariably kills parasitized hosts) that is specialized on the pest organism and causes sufficient damage to the pest to reduce its number greatly. Such a control agent, however, may have an unstable interaction with its host, leading to large oscillations in agent and pest densities, an undesirable outcome. However, there are many successful examples of biological control where the pest is maintained at low relatively and stable numbers. One candidate explanation for such control is that local population dynamics are indeed unstable, but regional dynamics are stable because they are metapopulation dynamics (Briggs *et al.*, 1999).

Multispecies Considerations: Competition

One organism may cause another to go locally extinct, not because it is a predator of that other organism, but because it is a competitor. Many years ago, G.E. Hutchinson and J.G. Skellam suggested that inferior competitors may persist regionally by being good colonizers of newly vacant habitat, where they persist only until their competitors find that habitat and drive them extinct. They termed such good colonizing inferior competitors "fugitive species" (Hanski, 1999). An essential component is a tradeoff between dispersal ability and competitive ability: If the better competitors were not poorer colonizers they would arrive too quickly in vacant habitat, giving insufficient opportunity for exploitation of that habitat by fugitives.

The fugitive species idea has been extended in recent years to consider the possibility that a suite of competitors in a system might have a competitive hierarchy, i.e., be strictly ranked in competitive ability with better competitors eliminating poor competitors whenever they occur in the same patch (Hastings, 1980). Such a competitive hierarchy prevents long-term coexistence in any patch, but if all local communities are ephemeral because disturbance takes place locally in space and removes whatever species happen to be there, then every so often any given local community becomes available for colonization by even inferior competitors. Colonization events in the presence of a competitive hierarchy drive succession (a change in species composition in a particular direction) within a local community toward domination by the best competitor. Assuming that inferior species have higher dispersal rates, then they have a higher probability, per unit occupancy (per unit p), of arriving in vacant habitat. Any species could actually arrive first, but inferior competitors have an arrival advantage. Any species arriving at a site that is a superior competitor to the current occupant takes that site. The local community shifts progressively in the direction of the best competitor in the system until disturbance occurs, interrupting the process and starting it all over again.

Because both extinction and colonization occur patchily in space and time, the metacommunity consists of a mosaic of local communities in different stages of succession (**Figure 1**). Each local community may have low species diversity because only the best competitor in the local community persists there for long. However, when summed up over space, the metacommunity has high species diversity due to the fact that different local communities are in different successional states and have different species. It is possible for the metacommunity to be in equilibrium, and maintain this high species diversity regionally. This equilibrium is dependent on the inverse ranking of competitive ability and colonizing ability, but it is not sufficient for an inferior competitor simply to be a

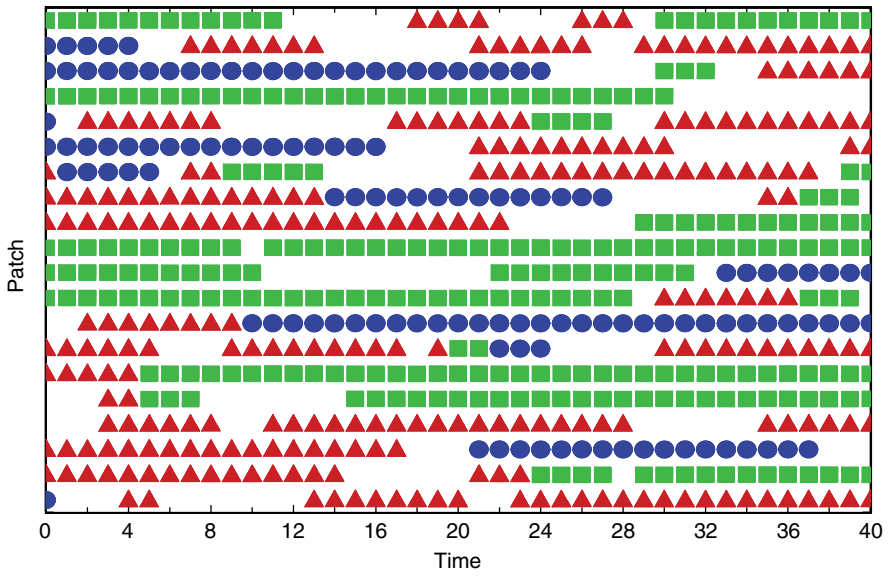


Figure 1 Turnover of species in a subset of the patches of a metacommunity with successional mosaic dynamics. Three different species are indicated by the symbols ●, ■, and ▲ in order of decreasing competitive ability and increasing colonizing ability. Blank means an empty patch.

better colonizer: it has to be better by more than a critical amount that depends on the colonizing abilities of superior competitors, and the disturbance frequency (Hastings, 1980).

The disturbance frequency has a particularly important role in the maintenance of regional diversity. A higher disturbance frequency means fewer patches have the best competitor and more patches are available for other species. At too low a disturbance frequency, lower ranked species do not have sufficient opportunity for colonization and therefore may not persist regionally: their ε values may exceed their c values. However, at higher disturbance frequencies, lower ranked species may persist and coexist regionally with higher ranked species. If the disturbance frequency is too high, the best competitors, which are by assumption poorer colonizers, have too high an extinction rate, ε, relative to their colonization rate, c, and disappear from the system. At extreme levels, only the very best colonizers can persist. It follows that with such competitive hierarchies, both high and low disturbance frequencies lead to low regional diversity. Diversity is maximized at some intermediate value of the disturbance frequency, an idea that is often referred to as the *intermediate disturbance hypothesis*. Disturbances here can be of the physical sort or of the biological sort provided the predators or other agents bringing about disturbance are generalists, and therefore are dependent for their own persistence on all species in the system, not just particular species, which they may control separately from other species. Specialist mortality agents may be important in the maintenance of diversity (Chesson and Huntly, 1997) but this idea is not the intermediate disturbance hypothesis.

The critical feature of the above mechanism of diversity maintenance is a metacommunity that exists as a successional mosaic, i.e., the system is diverse because local communities range over a variety of successional stages having different species that provide colonists of other local communities. Colonization moves succession along or re-establishes local communities following local disturbance (Chesson and Huntly, 1997). This successional mosaic model has been suggested to work with local populations consisting simply of single individuals, for example, single herbaceous plants in a meadow (Lehman and Tilman, 1997). Although this is not exactly what one might think of as a local population, the same principles apply. Death of an individual is then equivalent to local population extinction. Death of an individual may be brought on by the arrival of a superior competitor at the site, disturbance, or simply senescence. Disturbance or senescence open the site to invasion by any species. Otherwise the site can only be invaded by a superior competitor. The idea that many species may coexist in an area by this mechanism operating on a small spatial scale is referred to in the literature as coexistence by competition–colonization tradeoff. This idea is particularly useful in the case of plants or sessile aquatic animals that hold space (Lehman and Tilman, 1997).

These ideas on diversity maintenance in metapopulations have important implications for conservation. If habitat destruction lowers colonization rates, as discussed above, competitively superior species may be most at risk (Tilman and Lehman, 1997) because of their already poor colonizing abilities making them less tolerant to decreases in colonization rates. These models predict that habitat destruction eventually negatively affects all species, however. When the per unit colonization rate, c, of the best colonizer, is reduced below the extinction rate of local populations, all species are doomed to regional extinction even though patches of suitable habitat remain.

Quantitative Effects of Spatial Variation

The discussion so far has made the assumption that the only thing we need to know about a local population is presence or absence: is it extinct or not? This is crude accounting, for surely the number of colonists or propagules sent out by a local population depends on how large that local population is. Patches may be of different sizes and qualities, and therefore support local populations of different sizes and densities, issues that are active areas of research in metapopulation theory (Hanski and Gilpin, 1997). Moreover, local populations vary in size over time in any one patch, and in space from patch to patch even when patches of identical size and quality are compared. Ignoring changes in local abundance over time (i.e., ignoring the dynamics of local abundance) would be justified if local population build up after colonization occurred quickly to some local population equilibrium (a population size at which on average, reproduction and immigration balances deaths and emigration), about which population fluctuations occur until extinction. A broader variety of behaviors of metapopulations can be examined by taking actual local population sizes into account (Chesson, Donahue, Melbourne, and Sears, 2005). In general there is also no reason to restricted consideration to the situation where local populations are defined by discrete habitat patches. Indeed, a variety of techniques exist to study populations that vary continuously in space (Bolker, 2004).

These more general approaches explicitly incorporate the statistical properties of spatial variation of specific relevance to understanding the change in the nature of population dynamics as the spatial scale changes, in other words, the scale transition. These statistical properties are most simply understood in terms of particular variances and covariances in space of population densities, environmental characteristics, and population growth rates. Of particular note is fitness–density covariance (Chesson *et al.*, 2005). This quantity is the covariance in space between the relative density of a local population (local density divided by the spatial average density) and the average fitness of the individuals of a species in that local population. Fitness is defined here as the contribution of an individual through reproduction and survival to the metapopulation as a whole. Fitness–density covariance reflects the extent to which a population is concentrated in favorable locations and defines exactly how much this unevenness of distribution affects the growth rate of the whole metapopulation. Understandably, a population that is more concentrated in favorable locations will grow faster.

Fitness–density covariance, however, is but one aspect defining the change in population dynamics with the change in scale (the scale transition), and indeed, it is just one way of understanding the scale. Fundamental is the interaction between nonlinearities in population growth and spatial

variation. Illustrations of this key concept are given below. It is easier to understand by maintaining the approximation or fiction that space can be divided up discretely, with two relevant spatial scales. The smaller scale is the spatial scale on which interactions between individual organisms occur, i.e., it is the spatial scale of positive and negative density-dependent effects within species, the scale of competition between species and the scale of predation, whichever of these are important in the system of concern. This scale corresponds to the local population scale in the strict metapopulation sense, and shall still be referred to in terms of the patch and the local population. The larger scale is the scale of the whole population, which corresponds to the metapopulation or regional scale in the discussion above.

Single-Species Dynamics

To introduce the fundamental concepts, consider first of all density-dependent population dynamics applicable to organisms with annual life cycles. The dynamics of a local population in the absence of migration can be defined by plotting local population density (numbers per unit area), N_{t+1}, at time $t+1$, as a function of its density, N_t, the previous year, as represented in **Figure 2**. These curves are dynamical relationships, i.e., by applying them repeatedly, one can plot density as a function of time (**Figure 3**). The straight line (relationship I) represents the density-independent case where individual organisms do not interact with one another, and so an individual's contribution to future generations is independent of the number of other individuals. Thus N_{t+1} is simply proportional to N_t. Relationships II and III are two cases of density dependence and this density dependence means they are nonlinear, i.e., they are curved or humped relationships – certainly not straightlines. For relationship II, the resources sustaining a local population are strictly limited, fixing an upper bound on the local population density, which is achieved if the resources are efficiently utilized. The larger the population at time t, the closer the population comes to using all resources and the closer the upper bound on population density is approached. Annual plant populations commonly accord at least approximately with this relationship, which is sometimes referred to as contest competition. Some insect populations may have dynamical relationships more like curve III. Above a certain density, the number of insects in year $t+1$ is a decreasing function of the number of insects in year t. One explanation for a relationship like this is so-called scramble competition: at high densities, a high proportion of the population may starve to death, which leads to loss of the resources that these individuals consumed before death. Thus, a large fraction of the resources that could have been turned into new individuals at time $t+1$ is lost, and the population crashes.

The three different dynamical relationships of **Figure 2** give very different local population dynamics (**Figure 3**). The linear case, relationship I, gives simply geometric increase. Relationship II (contest competition) gives highly stable dynamics: the population quickly converges on an equilibrium population size, which is determined by the point at which the diagonal line in **Figure 2** intersects the dynamical relationship. Relationship III (scramble competition), however, gives irregular fluctuations referred to as deterministic chaos, which result here from the tendency of scramble competition to cause population crashes following population buildup. What happens when local populations with these various dynamics are connected regionally? If there is no variation in population density in space, then regional population dynamics are the same as local population dynamics. However, if there is a large amount of spatial variation in population densities (a common occurrence in nature), regional population dynamics can be very different from local dynamics (**Figure 4**). The density-independent case (relationship I), however, does not show different local and regional dynamics because individuals are not affected by density, and so population dynamics cannot be affected by spatial variation in density. On the other hand, the strong density dependence arising with scramble competition leads to correspondingly striking differences between regional and local dynamics. For the situation depicted, which is defined in detail below, chaotic fluctuations have been replaced

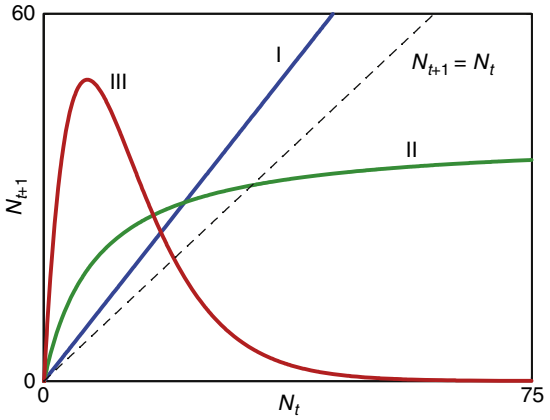


Figure 2 Solid curves: dynamical relationships of single-species local populations, I density-independent dynamics, II contest competition, III scramble competition. The equilibria are the intersections of these curves with the dashed line defining no change in local population size ($N_{t+1} = N_t$).

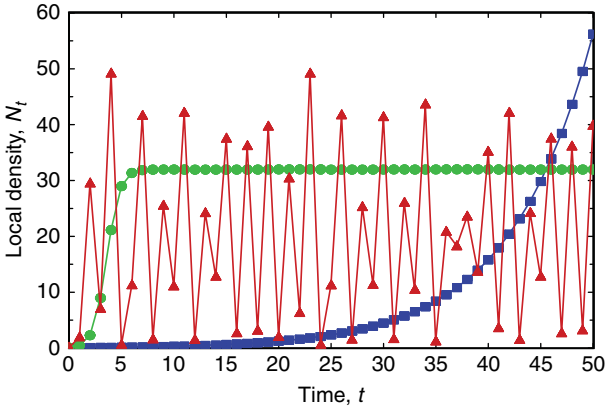


Figure 3 Local population dynamics generated by the dynamical relationships of **Figure 2**. Relationship I (density independent), ■; Relationship II (contest competition), ●; Relationship III, (scramble competition) ▲.

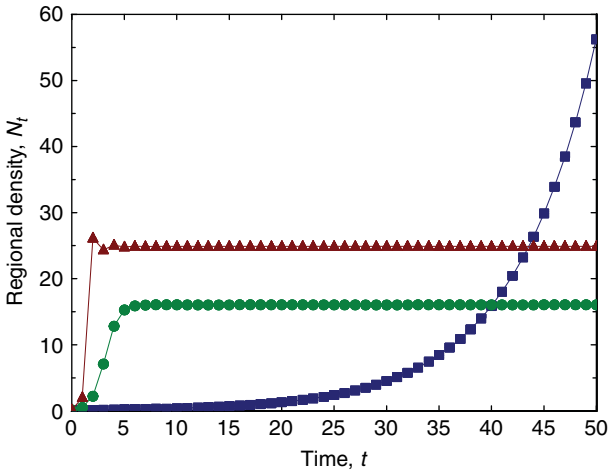


Figure 4 Regional dynamics generated by the interaction of the dynamical relationships of **Figure 2** with spatial variation in population densities. Relationship I, ■; Relationship II, ●; Relationship III, ▲.

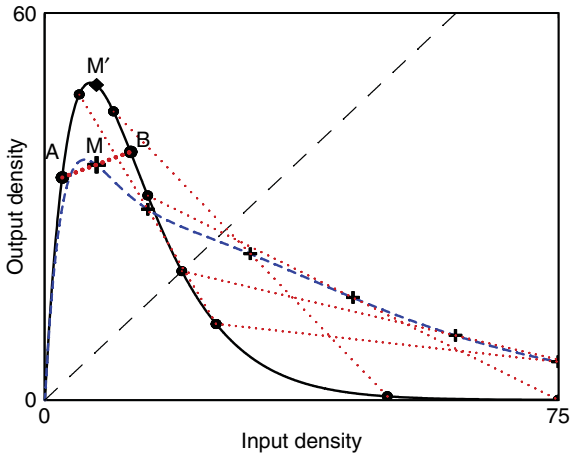


Figure 5 Construction of the regional dynamical relationship from the local dynamical relationship III of **Figure 2** and spatial variation between two densities. Solid curve: local dynamical relationship. Thick dashed curve with + symbols: regional relationship.

by a stable equilibrium. For contest competition (relationship II), the effect of spatial variation on regional population dynamics is quantitative not qualitative: the equilibrium density is lowered.

What is the explanation for these changes in dynamics at the regional level in density-dependent situations? To answer this question, we need to work out the dynamical relationship between $\overline{N}_{t+1}$ and $\overline{N}_t$ defining the dynamics of population density at the regional level. These regional densities can be defined as the averages of the population densities in local populations, weighted if necessary by patch area. Dynamical relationships like those in **Figure 2** continue to define dynamical change within patches, but disturbance and fluctuations due to small population size lead to random deviations about these relationships. Thus, these local dynamical relationships are just mean relationships converting inputs of population density into outputs and some of the output population in any patch may then disperse to other patches.

The essence of the difference between local and regional dynamics can be understood by imagining that dispersal occurs early each year, and for the rest of the year, patches are isolated with mean change governed by dynamical relationships in each patch. Further imagining that after dispersal local populations exist at just two densities, high and low, in equal abundance, then the regional dynamical relationship is easy to derive. This is done in **Figure 5**, where it is assumed that low density patches have inputs 1/3 the regional density, $\overline{N}_t$, and high-density patches have inputs 5/3 of $\overline{N}_t$. The solid curve defines the relationship of local outputs to local inputs (scramble competition) and the dashed curve then defines the relationship between $\overline{N}_{t+1}$ and $\overline{N}_t$, i.e., the regional dynamical relationship. This regional relationship is found by connecting pairs of points on the local relationship, corresponding to low and high-density inputs, and finding the midpoints of the lines joining these pairs of points. For example, the points A and B of **Figure 5** are one such pair of points, and M is their midpoint. The x coordinate of M is thus the regional input, $\overline{N}_t$, which is the average of the x coordinates (local inputs) of A and B. The y coordinate of M is the regional output, $\overline{N}_{t+1}$, which is the average of the y coordinates (local outputs) of the points A and B. The complete regional relationship is found by repeating this procedure for every possible value of the input density $\overline{N}_t$.

Comparison of the points M on the regional relationship with the point M′ on the local relationship reveals the reason for the difference between local and regional relationships. These points both have the same input density, but M′ is the output of the local dynamical relationship, whereas M comes from averaging the outputs of the local dynamical relationship at two different input densities, which give outputs less than M′ due to nonlinearity (curvature) of the local relationship. Thus, nonlinearity in the local dynamical relationship combines with variation in local inputs reducing the hump on the regional dynamical relationship. A similar effect acting on the nonlinearty in the local dynamical relationship after the hump reduces the severity of the decline in the regional dynamical relationship, and it is the combination of these two effects that is responsible for the strong stability of the regional dynamics shown in **Figure 4**. Nonlinearity and spatial variation also affect regional dynamics in the case of contest competition but there the effects are quantitative, not qualitative. The regional relationship for contest competition is shown in **Figure 6**, where it is assumed that 50% of local populations are extinct, and the others have a density twice the regional density. Unsurprisingly, the regional equilibrium is 50% of the local equilibrium, but this reduction, and the regional relationship in total, can be understood using the same averaging approach discussed here for scramble competition.

Having local populations take on just two densities for any given average density is clearly unrealistic. However, the qualitative features of the results above occur with general sorts of variation commonly observed in natural populations. The most common situation in nature is a negative binomial distribution of possible local population densities, which means that densities are somewhat more clumped than random in space. The striking change in dynamics from regional to local found for scramble competition occurs in enhanced

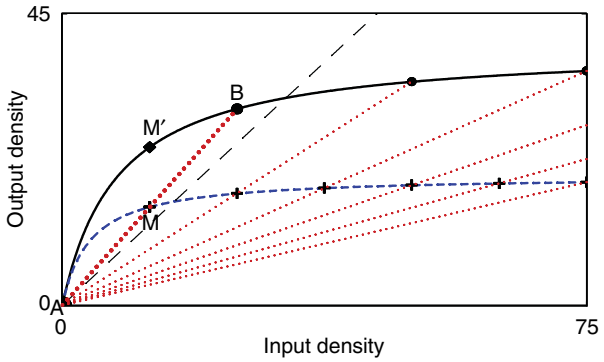


Figure 6 Construction of the regional dynamical relationship from the local dynamical relationship II of **Figure 2** in the presence of spatial variation between two densities. Solid curve: local relationship. Thick dashed curve, regional relationship.

form in that case (Chesson 1998). In the density-independent case (relationship I) the dynamics at the regional level are unaffected by variation from patch to patch: M always equals M′, regardless of the pattern of variation. The nonlinearities, i.e., the deviations from a straight line relationship, present in relationships II and III, are necessary for variation in space to have an effect on the dynamics of regional density, and these nonlinearities arise from density dependence.

There are many causes for spatial variation in local densities of which disturbance and fluctuations due to small population size are mentioned above. Simple random dispersal processes also lead to some variation in density, but spatially varying environmental factors may have large effects on dispersal, and on the ultimate distribution of organisms among patches. For example, in marine organisms, such as reef fishes, local currents may transport larvae away from some places and toward others, and these currents may vary with time of year, storms, and other atmospheric conditions, potentially delivering dispersing larvae to reefs in a very patchy fashion. Moreover, organisms commonly actively seek particular places, and the relative attractiveness of different places may vary greatly. Organisms may also seek out their own species, and so any existing spatial variation in density may be magnified by later dispersal. Unstable population dynamics occurring locally in space can itself be a cause of local variation in density, at least in models. In the illustrations here, spatial variation in density was assumed to have a fixed relationship to regional density (e.g., low-density patches $\overline{N}_t/3$, and high-density patches $5\overline{N}_t/3$). More realistically, local densities will not have fixed relationships to the mean, and therefore the regional relationship will in fact change over time as the relationship between spatial variation and regional density changes (Chesson, 1998). Nevertheless, the way in which nonlinearity and spatial variation change the regional relationship, illustrated in Figs. 5 and 6, remains the same in spite of all these complications. In particular, the stabilizing effect of spatial variation in densities on scramble competition continues to be seen in much more complex circumstances than illustrated here.

These quantitative considerations are very different in spirit from the colonization and extinction issues discussed above for strict metapopulations, but the two approaches can be related. Colonization and extinction dynamics involve implicit density-dependence: to count all occupied patches as equivalent, recently colonized patches must quickly increase in density to some sort of steady state, such as a local equilibrium density, or fluctuation about a local equilibrium. Local dynamical relationship II would serve in this circumstance, with the mean output taking into account the probability of local disturbance. The specific example above with half the patches extinct corresponds to a c to ε ratio of 0.5.

These quantitative considerations apply also to spatially structured populations that are not strict metapopulations, for example, to situations where recovery after disturbance depends to a large extent on regeneration rather than recolonization, and also to situations where only part of the life cycle is spatially confined. For example, it is very common for marine organisms that live on a reef, intertidally, or on the ocean floor, to have widely dispersing larvae, so that only the adult stage has clear spatial structure (Kritzer and Sale, 2006). The adult may be physically attached to the bottom, as in corals, barnacles, and mussels; or if it remains mobile, it may have a territory or at least a home range that is much restricted in extent relative to the distance traveled by dispersing larvae, as in many reef fishes. Such spatial restrictions mean that interactions between individuals are localized, because the individuals are localized and are most likely most strongly affected by individuals living nearby, for example, on the same coral head or part of the reef. Similarly, in terrestrial plants a spatial unit of major importance is the immediate neighborhood of other plants close enough to compete with it. Insect populations often have spatially confined larvae and dispersive adults. Although it is the adults that disperse in this case, the fundamental principles are the same in all of these cases, and the effects of variation in density in space are similar, even though in no case does spatial structure of one part of the life cycle qualify these populations as metapopulations in the strict sense. The effects of nonlinear dynamical relationships that we have explored above depend on having the smaller scale, which corresponds to the local population in a strict metapopulation, correspond to the scale on which those nonlinear relationships apply. As these nonlinear relationships derive from density dependence, the smaller scale is also the scale of density dependence.

How the scale of density dependence is determined depends on how the individuals in a population affect each other. For example, for territorial coral reef fishes, one source of density dependence may simply be competition for space to set up feeding territories. Other fishes may compete for hiding places from predators. The scale on which fishes may move to secure territories or hiding places defines the scale of density dependence. Predators may cause density dependence in their prey populations by responding to the density of the prey, for example, aggregating in areas of high prey density. The scale of density dependence then is the spatial scale on which predators respond to variation in prey density. Predators may also build up in numbers where prey are dense simply because they have more to eat and therefore reproduce faster. The scale on which this effect occurs is bound to be much larger than the scale of aggregation of predators to variation in prey density. The changes in dynamics with a change in scale demonstrated in **Figures 3–6** should occur in all of these

instances when one compares the dynamics defined on the scale of density dependence with the dynamics that emerge at the regional or metapopulation scale (Chesson, 1998).

Predators and Parasitoids

Like the dynamics of colonization and extinction, a rich array of phenomena is uncovered when the quantitative effects of spatial structure are examined in terms of interactions between species. In particular, predator–prey and host–parasitoid dynamics (**Figure 7**) are often stabilized by local interactions and spatial variation (Hassell, 2000). This phenomenon is best understood in host–parasitoid systems where a critical issue is spatial variation in the risk of parasitism experienced by a host, which may result from variation in parasitoid density. **Figure 8** illustrates how this works. The risk of surviving is thought to be an exponentially decreasing function of parasitoid density, as shown by the solid line. Very simple two-level variation in parasitoids is assumed for ease of illustration: no parasitoids in half the patches, and twice the mean parasitoid density in the other half. The effect of averaging

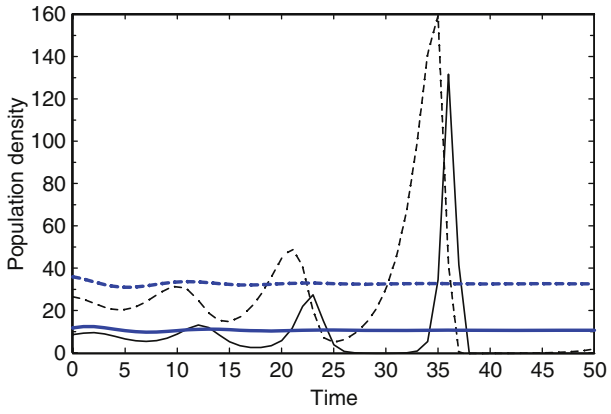


Figure 7 Host–parasitoid dynamics. Thin lines: local dynamics of an isolated host–parasitoid community in the absence of migration. Thick lines: regional dynamics with parasitoids concentrated in half the patches. Hosts: dashed; parasitoids: solid.

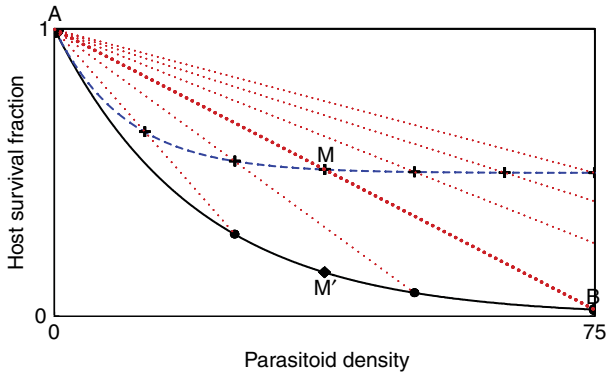


Figure 8 Local host survival rates and regional host survival rates as a function of parasitoid density with half the patches having no parasitoids.

different risk levels, exemplified by **Figure 8** is to moderate the decline in host survival with increasing parasitoid density, which means that host crashes following increases in parasitoid numbers are diminished, tending to stabilize the host–parasitoid interaction regionally (**Figure 7**). Although the reasons for spatial variation in parasitoid density are poorly understood in general, an important feature of predator–prey and host–parasotoid interactions is the potential for the unstable nature of the interactions themselves to generate variation in risk, which may then stabilize population dynamics on the larger spatial scale (Hassell, 2000).

Competitive Interactions

Competition–colonization tradeoffs is one mechanism allowing coexistence of competing species in a patchy environment. Quantitative approaches allow the consideration of other mechanisms and more general spatial variation. Species fail to coexist (i.e., competitive exclusion occurs) when the species best adapted to the environment depletes resources to levels that are too low for other species to persist. Because the environment varies in space, different species may be the best adapted under different environmental conditions. Then many species that depend on the same resources may coexist regionally because each has patches where it is the best competitor, i.e., each has its own spatial niche. If each of these patches were a closed community, i.e., not connected to other places by migration, then just a single species would come to dominate in each locality. However, in the presence of migration, each patch may have many species, because a species that is not the best competitor in that local community has its reproduction supplemented by immigration from places where it is the best competitor.

If environmental characteristics vary over time as well as in space favoring some species sometimes and others at other times, local diversity is even more likely to be maintained at high levels because trends in the local species abundances will be reversed with reverses in the ranking of competitive superiority in a particular locality: species tending toward low relative abundance will be partially restored to higher values. A phenomenon with the opposite effect has been discussed by Rozenzweig (1995), and may be particularly important in animals, especially those with complex behavior. Animals detecting a competitively superior species may leave or keep away from a patch, an effect tending to divide up a landscape into exclusive areas containing only the best species for that area.

These various effects of spatial environmental variation do not require strict metapopulations, but simply spatial structure. This more general theory can be applied, for example, to coexistence of insects, especially various communities of flies, such a *Drosophila* and carrion flies, which lay eggs in ephemeral patches of food, such as fruit, mushrooms, or dead animals. In general, although the larvae of several species are commonly found developing in the same patch of food, it is also a common finding that there is segregation between species in choice of food patches. Although the underlying cause of the choice is not well understood, the effect it has is to promote very strongly stability (Ives, 1988). A second example

of the potential importance of spatial environmental variation is marine communities, especially coral reef fishes, but also various sedentary invertebrates that compete for space, with similar effects applicable to plant species. In marine communities as discussed above, dispersing larvae characteristically arrive at potential settling sites in a highly patchy manner. Such systems are said to be strongly affected by recruitment variation, where recruitment refers to the process of a larva settling at a particular patch (Kritzer and Sale, 2006). Differences in the spatio–temporal patterns of recruitment have the potential to contribute to species coexistence. In this case, the environmental advantage that a species has in a particular patch is proportional to its arrival rate there.

There is a very simple graphical technique for understanding the effects of variation in recruitment rates in space, which serves to illustrate how spatial environmental variation, and its interaction with population density, may have important effects. **Figure 9** assumes a community with two competing species and plots the proportion of new inputs to a local community as a function of the local environment and the regional abundance of the species. The two solid curves correspond to two environment types. The x-axis is the proportion, p_t, of species 1 in the system as a whole. The y-axis is the proportion, s_{t+1}, of the available space at the site taken by new settlers of species 1 during the time interval from t to $t+1$. The upper curve refers to patches favoring settlement of species 1, whereas the lower curve is for patches favoring settlement of species 2. The difference between the different sorts of patch is in the relative arrival rates of larvae of the two species: more of species 1 arrive at upper-curve sites, and more of species 2 arrive at lower-curve sites. The species compete for available space at a site, which is assumed limiting and thus filled by the larvae arriving each year. Note that if there were only one sort of patch in the system, eventually one species would eliminate the other, because it would have higher settlement success everywhere, and would increase in abundance relative to the other species. For example, if all sites favored species 1 (upper curve), then the fraction of species 1 settling would always be higher than p_t. It follows that species 1 would steadily increase until it had taken over altogether, assuming that settled individuals of both species have the same mortality rates (a simple adjustment, however, extends this idea to different mortality rates) (Chesson, 2000).

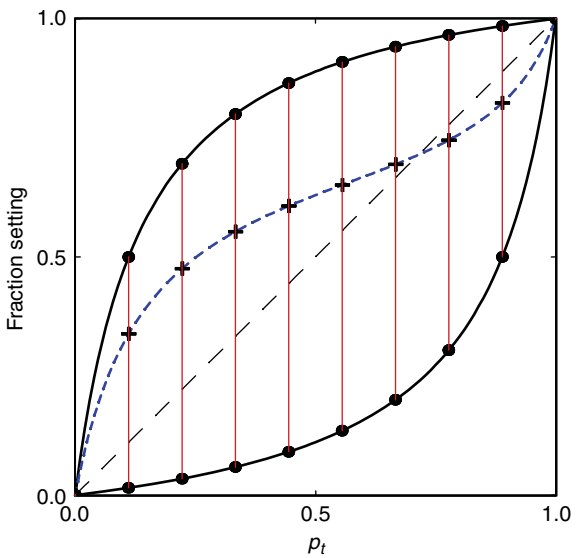


Figure 9 Local and regional settling fractions (proportion species 1) as a function of regional fractional abundance for a marine space-holding community with two species. Upper solid curve: local environment favoring species 1, lower solid curve: local environment favoring species 2, and wavy dashed curve: regional settling fraction with 2/3 of patches favoring species 1. The regional equilibrium is the intersection of the wavy dashed curve with the diagonal dashed line.

In **Figure 9** it is assumed that two thirds of patches favor species 1, and one third favor species 2. This means that the average proportion of settlers, s_{t+1}, that belong to species 1regionally is (2/3) upper solid curve + (1/3) lower solid curve, which is given by the points dividing the vertical lines between the two solid curves in the ratio 1:2. The dynamics of regional settlement are then given by the wavy line between the two curves. Depending on the actual adult mortality rates of settled organisms, the regional proportion, p_{t+1}, of species 1 at time $t+1$ will lie somewhere between p_t and the regional s_{t+1}. Thus, the point where the wavy line crosses the diagonal is the equilibrium point for the system, i.e., the point where p_t does not change over time. It is a stable equilibrium point too: if the value of p_t is less than the equilibrium, then the regional s_{t+1} is always greater than p_t, and so species 1 increases in relative abundance ($p_{t+1} > p_t$). However, if p_t is above the equilibrium, the regional s_{t+1} is less than p_t and so species 1 decreases in the system. Hence, whenever the system is away from equilibrium, it moves back towards equilibrium over time.

These results apply whether a patch continues to favor the same species forever, or favors different species at different times, for example, if environmental conditions responsible for larval transport change over time. These results do have to be modified, however, if instead all patches favor a particular species at a particular time, but the favored species changes over time, for example, if the environmental conditions applicable to a particular year favor one species over the other everywhere. In that case, the environment is varying purely temporally. The way temporal variation affects the dynamics of an organism is greatly influenced by life-history parameters, such as the adult death rate because this determines how long past effects of fluctuations in settlement are reflected by the age-structure of the population. A low adult death rate means that a species can persist over periods when it is at a competitive disadvantage without its population density declining too greatly. It follows that the ability of species to coexist as a result of temporal variation varies inversely with the adult death rates (Chesson, 2003). With irregular temporal fluctuations, the system does not have a traditional equilibrium, but nevertheless can still exhibit stable fluctuations about a mean value as depicted in **Figure 10**.

Temporal variation is an important feature of many systems, and a mechanism like that described here, termed as the storage effect, has been suggested for coexistence in tropical trees where highly variable fruit production occurs, in annual plants where seed germination may vary greatly between years but seeds survive over unfavorable times in the soil seed bank, and also in annual zooplanton in lakes where a resting egg bank takes the place of the seed bank (Chesson, 2003).

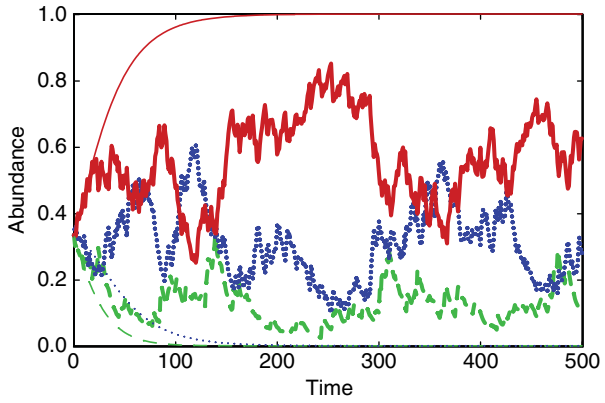


Figure 10 Fluctuations in abundance (proportion of sites held) for three competing space-holding species experiencing independent temporal fluctuations in recruitment rates (thick wiggly curves) and constant recruitment rates (thin smooth curves).

Lessons from Metapopulation Theory

Conservation

Few populations in nature exist without important spatial structure. Although there is debate over how often that structure conforms to a strict metapopulation, the metapopulation concept is important in conservation biology because habitat fragmentation creates distinct local populations even if they could not be defined before human interference. There is no question that the inhabitants of small isolated habitat fragments are at risk of local extinction. Although a collection of such habitat fragments may genuinely be thought of as a metapopulation, recolonization rates may be so low that each local population will go extinct before any are recolonized. Metapopulation theory emphasizes the need, at a very minimum, for the recolonization rates c to exceed the extinction rates ε. Habitat destruction not only puts populations at risk by reducing their total size and spatial extent, but by altering the connectedness of the remaining habitat, it reduces the value of this remaining habitat.

Metapopulation theory also says that even uniform reduction in recolonization rates with habitat destruction may have different impacts on different species in the system. The system therefore may change in character with habitat destruction even under conditions that on the surface may not appear biased in favor of one species or another. Such effects have the potential to disrupt the functioning of ecosystems. This is a general message not specific to strict metapopulations. The metapopulation perspective also teaches that not only inhabitated habitat is a suitable habitat. A species may be only temporally absent or in low abundance in an area, and that habitat may in fact be playing an important role in the system. Empty habitat must be judged in terms of its potential role as part of a larger connected system. Short-term environmental fluctuations may play a role in changing the suitability of habitat patches. Patches that are unsuitable at present may have the potential to become suitable in the future. With the prospect of global climate change, the maintenance of a diversity of patch types, and a landscape structure suitable for migration of organisms between patches, are of major importance in reducing the loss of biodiversity.

Nonequilibrium Dynamics

The metapopulation perspective highlights the role of nonequilibrium conditions in space, i.e., situations in which local populations do not remain near an equilibrium population size. Simple predator–prey and host–parasitoid models commonly do not have both species coexisting stably in isolated patches. Nonequilibrium locally in space, that occurs asynchronously in different patches, has an essential role in stabilizing the spatial extensions of these models. Similarly, in successional mosaic systems, perturbation from equilibrium (domination by one species), locally in space, is critical to maintenance of the system. Such essential roles for nonequilibrium conditions have sometimes been taken as that the system is particularly robust to perturbations, including habitat degradation, by human activities. There is no reason to expect this to be true. Indeed, the message here is that equilibrium occurs on the larger spatial or temporal scales. Interfering with the system through habitat destruction, alteration in the disturbance frequency, alteration in mean environmental states, and alteration of the variance in environmental states all have major consequences, and are therefore of serious concern even though on small scales the system may not be at an equilibrium.

Population Regulation

These metapopulation and structured population perspectives also have the potential to resolve the long-standing debate over the relative importance of environmental fluctuations versus density-dependent interactions within and between species. Andrewartha and Birch were the first to clearly articulate metapopulation structure as an important feature of nature (Hanski, 1999), yet because of the prominent role of stochastic factors (factors with random elements), such as extinction and colonization, they believed that density dependence was of minor importance. The quantitative analysis here shows that far from being of minor importance, density dependence within local populations has a major role in shaping the nature of population dynamics at the metapopulation or regional scale. Density-dependent processes lead to nonlinear dynamical relationships, and we have seen how such relationships interact with fluctuations in time and space to give outcomes at larger temporal and spatial scales that are not predictable on the basis of the separate effects of these factors.

See also: Competition, Interspecific. Conservation Efforts, Contemporary. Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems. Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity. Landscape Corridors. Landscape Modeling. Loss of Biodiversity, Overview. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Modeling Marine Ecosystem Services. Natural Reserves and Preserves. Parasitoids. Predators, Ecological Role of. Population Dynamics. Population Viability Analysis. Scale, Concept and Effects

of. Species Coexistence. Stability, Concept of. Succession, Phenomenon of

References

Bolker B (2004) Continuous-space models for population dynamics. In: Hanski I and Gaggiotti OE (eds.) *Ecology, Genetics, and Evolution of Metapopulations*, pp. 45–69. San Diego: Elsevier Academic Press.

Briggs CJ, Murdoch WW, and Nisbet RM (1999) Recent developments in theory for biological control of insect pests by parasitoids. In: Hawkins BA and Cornell HV (eds.) *Theoretical Approaches to Biological Control*, pp. 22–42. Cambridge: Cambridge University Press.

Chesson P (1998) Making sense of spatial models in ecology. In: Bascompte J and Sole RV (eds.) *Modeling Spatiotemporal Dynamics in Ecology*, pp. 151–166. New York: Springer.

Chesson P (2000) General theory of competitive coexistence in spatially-varying environments. *Theoretical Population Biology* 58: 211–237.

Chesson P (2003) Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. *Theoretical Population Biology* 64: 345–357.

Chesson P, Donahue MJ, Melbourne BA, *et al.* (2005) Scale transition theory for understanding mechanisms in metacommunities. In: Holyoak M, Leibold MA, and Holt RD (eds.) *Metacommunities: Spatial Dynamics and Ecological Communities*, pp. 279–306. Chicago: The University of Chicago Press.

Chesson P and Huntly N (1997) The roles of harsh and fluctuating conditions in the dynamics of ecological communities. *The American Naturalist* 150(5): 519–553.

Hanski I (1999) *Metapopulation Ecology*. Oxford: Oxford University Press.

Hanski I and Gaggiotti OE (eds.) (2004) *Ecology, Genetics, and Evolution of Metapopulations*. San Diego: Elsevier, Academic Press.

Hanski IA and Gilpin ME (1997) *Metapopulation Biology Ecology, Genetics, and Evolution*. San Diego, California: Academic Press.

Hassell MP (2000) *The Spatial and Temporal Dynamics of Host–Parasitoid Interactions*. Oxford: Oxford University Press.

Hastings A (1980) Disturbance, coexistence, history, and competition for space. *Theoretical Population Biology* 18: 363–373.

Holyoak M, Leibold MA, and Holt RD (eds.) (2005) *Metacommunities: Spatial Dynamics and Ecological Communities*. Chicago: The University of Chicago Press.

Ives AR (1988) Covariance, coexistence and the population dynamics of two competitors using a patchy resource. *Journal of Theoretical Biology* 133: 345–361.

Kritzer JP and Sale PF (eds.) (2006) *Marine Metapopulations*. San Diego: Elsevier Academic Press.

Lehman CL and Tilman D (1997) Competition in spatial habitats. In: Tilman D and Kareiva P (eds.) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, pp. 185–203. Princeton, New Jersey: Princeton University Press.

Nee S, May RM, and Hassell MP (1997) Two-species metapopulation models. In: Hanski IA and Gilpin ME (eds.) *Metapopulation Biology: Ecology, Genetics, and Evolution*, pp. 123–147. San Diego: Academic Press.

Rosenzweig (1995) *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.

Tilman D and Lehman CL (1997) Habitat destruction and species extinctions. In: Tilman D and Kareiva P (eds.) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, pp. 233–249. Princeton, New Jersey: Princeton University Press.

Microbial Biodiversity

Kristin L Matulich, University of California, Irvine, CA, USA
Devon J Mohamed, Irvine Valley College, Irvine, CA, USA
Jennifer BH Martiny, University of California, Irvine, CA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Devon J. Bradley, Jennifer B.H. Martiny, pp 1–9, © 2001, Elsevier Inc.

Glossary

Archaea One of the two entirely microbial domains of single-celled organisms, evolutionarily distinct from bacteria. Many archaeal characteristics are more similar to eukaryotes than bacteria.
Bacteria An entirely microbial domain of single-celled organisms, evolutionarily distinct from Archaea.
Eukarya A primarily microbial domain of organisms having a membrane-enclosed nucleus and other organelles; includes animals, plants, fungi, and protists.
Fungi Nonphototrophic, heterotrophic eukaryotic microorganisms that contain rigid cell walls; includes mushrooms, molds, and the fungal part of lichens.
Metagenomics The genomic analysis of microorganisms by direct extraction and cloning of community deoxy ribonucleic acid (DNA) from an environmental sample.
Microorganism Single-celled organisms that can only be observed with a microscope, including bacteria, archaea, small eukaryotes, and viruses.
Operational taxonomic unit (OTU) A group of organisms regarded as being distinct from other groups, based on any clearly defined variables.
Polymerase chain reaction (PCR) A method for copying DNA sequences by repeated cycles of synthesis using specific primers and DNA polymerase.
Virus A genetic element containing either DNA or ribonucleic acid (RNA) that replicates in cells but is characterized by having an extracellular state.

Introduction

Although it might not be immediately obvious, our world is a microbial one. Biodiversity is usually discussed in terms of large organisms, but no organisms are more ubiquitous, abundant, or diverse than microorganisms. Microorganisms were the first cellular life forms and were active more than 3 billion years before the appearance of macroorganisms. The metabolic activities that they carried out during this time were critical for creating the conditions for the evolution of multicellular forms. The universal tree of life (**Figure 1**) emphasizes this point. Multicellular Eukarya (Metazoa, some Fungi, Plantae) are crown groups compared with most microbial Eukarya (e.g., Chromalveolata and Diplomonadida), Bacteria, and Archaea. Microbes force us to stretch our imagination about the limits of metabolic lifestyles, the geography of life, and the roles that organisms play in our lives.

Before the development of molecular methods, it was particularly challenging to understand the evolutionary relationships between organisms. It was generally accepted that there were two basic kinds of organisms, prokaryotes and eukaryotes, distinguished by the absence or presence of a membrane-bound nucleus. In the mid-1970s, Woese and colleagues assembled the universal tree of life based on sequence information of small-subunit ribosomal RNA (SSU rRNA). This breakthrough revealed that life consisted of two very distinct microbial domains, the Bacteria and Archaea, which appear to have diverged very early on in evolutionary history, and a third mostly microbial domain, Eukarya, the exact origins of which remain obscure. Archaea and Bacteria share similar morphological characteristics but Archaea possess several cellular and genetic characteristics that are more similar to eukaryotes and these organisms tend to be more abundant under extreme environments. As a consequence, Bacteria and Archaea often perform quite different roles. For example, photosynthesis is performed by Bacteria, spread across five different clades, and is not found among Archaea, while methanogenesis is restricted to Archaea. Members of these two domains occupy all habitats on Earth.

Characterizing Microbial Diversity

Although delineating species can be challenging for some macroorganisms, defining a microbial species is particularly challenging, in part because of their genetics. For sexual organisms, a species is usually defined as a group of potentially interbreeding organisms. But many microorganisms, bacteria in particular, are asexual. In addition, microorganisms are too small to define by morphological characteristics, as this is usually done by plant and animal taxonomists. Thus, in stark contrast to diversity surveys of large organisms, these microorganisms often focus on variation in their DNA rather than in their phenotype. For instance, bacterial species are defined as a group of strains that share more than 70% of their genome. Operationally, this definition is applied by melting the genomic DNA of two bacteria and measuring the rate at which they reanneal, a rate determined in part by the similarity of the organisms' genomes.

For microbial community studies, the 70% genomic similarity definition is impractical to apply. Instead, researchers

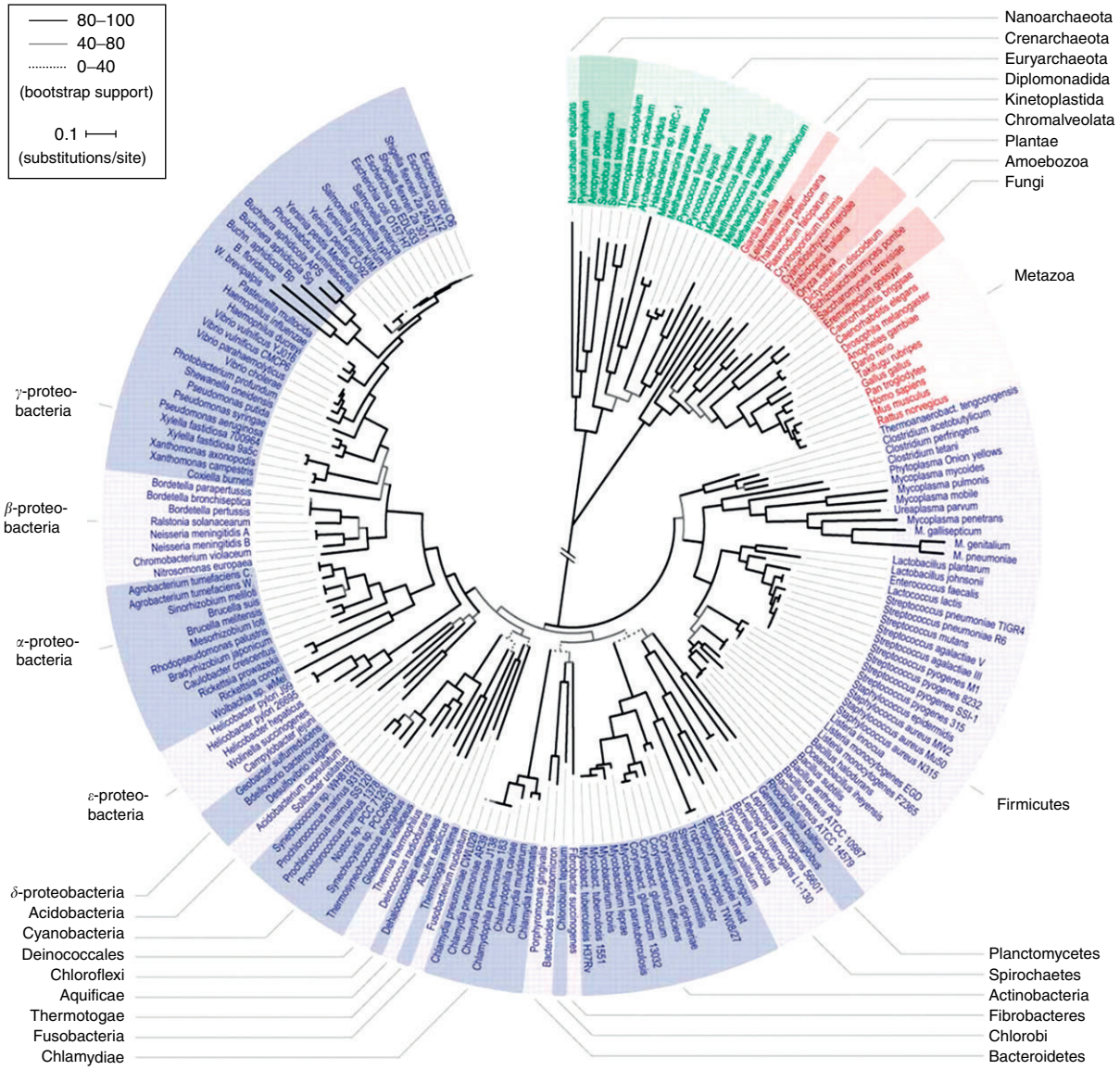


Figure 1 A phylogeny of some fully sequenced organisms showing the relationship between the three Domains. Green section, Archaea; red, Eukaryota; blue, Bacteria. Labels and color shadings indicate various subdivisions. The branch separating Eukaryota and Archaea from Bacteria in this unrooted tree has been shortened for display purposes. Reproduced from Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, and Bork P (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* 311: 1283–1287, with permission from AAAS.

use genetic markers to characterize microbial community composition and define an operational taxonomic unit (OTU). For instance, in studies that sequence the 16S ribozomal DNA (rDNA) of bacteria, a common OTU definition is a group of sequences that are similar by 97%. This definition is based on the rough relationship between 16S rDNA similarity and the 70% genomic similarity definition. Once microbial species, or other taxonomic units, are defined, then the diversity of a microbial community can be measured. In its simplest form, community diversity is the richness (total number of OTUs) and evenness (relative abundance of the OTUs) in a defined area. A full description of microbial diversity, however, includes variation at many levels of biological organization from the alleles at a particular genetic locus to community among habitats.

Culture Techniques

Throughout most of the twentieth century, microbiologists used culturing techniques to characterize the composition of microbial communities. Traditionally, these techniques involve trying to grow microorganisms from environmental samples on artificial media and mimicking natural environmental conditions, such as temperature and light, in a laboratory setting. These traditional cultivation techniques tend to select for fast-growing taxa, and therefore grossly underestimate actual diversity. The techniques are thought to capture generally less than 1% of all bacterial species and therefore provide a biased view of microbial community composition in most environments. In the early twentieth century, however, promising new cultivation techniques

were developed to overcome these deficiencies. For instance, researchers used small diffusion chambers to grow bacterial colonies in a more natural environment and resource regimen. Others have used high-throughput methods to culture novel strains of oceanic bacteria in a low-nutrient media. These newer techniques have allowed microbial biologists to study the ecophysiological properties of organisms that had been previously inaccessible.

Genetic Methodologies

The development of culture-independent methods, based on the genetic characterization of microbial communities, has allowed the detection of nonculturable species and a more detailed description of microbial communities. Because rDNA molecules are found in all organisms and are probably rarely laterally transferred, they have become a standard phylogenetic marker of microorganisms. Pace *et al.* (1986) built on this idea and used the polymerase chain reaction (PCR) to amplify rDNA genes from community DNA in environmental samples. After PCR, the products can then be further assayed to characterize the diversity of the community rDNA genes. Current open-access databases such as GenBank and the Ribosomal Database Project contain thousands of rDNA gene sequences.

Numerous methods can be used to characterize the diversity of PCR-amplified fragments. DNA sequencing provides the most detailed information, but can be expensive. Fingerprinting methods provide an alternative for characterizing microbial community composition. For instance, terminal restriction fragment length polymorphism (T-RFLP), a modification of the conventional RFLP approach, allows the characterization of a community by analyzing the size polymorphisms of PCR-amplified genes that have been cut at precise sequence locations by restriction enzymes. The resulting fragments are separated by gel electrophoresis and produce a banding pattern that can be used as a fingerprint for comparing community composition among samples.

As DNA sequencing has become faster and cheaper, the field of metagenomics has emerged. Metagenomics uses the DNA sequences of fragments of DNA obtained from a natural microbial community to describe the collective diversity of microbial genomes. Shotgun sequencing breaks up an organism's or a community's DNA into a myriad of short fragments that are individually sequenced. These fragments are then computationally ordered based on overlaps in the genetic code and reassembled into larger genomic fragments. For instance, Venter *et al.* (2004) applied this approach to an open ocean environment in the Sargasso Sea. This pilot project detected 150 new species of bacteria and more then 1.2 million new genes. Screening of metagenomic soil libraries has already led to the identification of various novel biomolecules including enzymes and antibiotics. An alternative approach allows the insertion of large continuous sequences of DNA to be incorporated into bacteria to form artificial chromosomes (BACs), which are then sequenced. This latter approach is much more labor-intensive and time consuming, but allows one to see how genes are organized within microbes, which is important information if one is interested in understanding their evolutionary history.

What is Microbial Diversity?

The immense diversity, small size, and clonal nature of most microorganisms explain why quantifying the biodiversity of microorganisms is fundamentally different from quantifying that of macroorganisms. Microorganisms are so abundant and diverse that only a minute fraction of their diversity has been described (**Table 1**). Currently, it is infeasible to completely quantify the microbial diversity in even a gram of soil or a liter of seawater. However, a number of studies have tried to estimate microbial diversity globally or within particular habitat types.

Bacterial and Archaeal Diversity

The sheer abundance of bacteria makes estimating the number of bacterial taxa difficult. It has been calculated that there are more bacteria on Earth than there are stars in the universe and that most microbial communities harbor 10^{10}–10^{17} bacterial cells that compose greater than 107 different taxonomic groups and innumerable functional groups. The magnitude of bacterial diversity is still a matter of debate and perhaps even beyond practical calculation. Dykhuizen (1998) has speculated that there could be 10 billion species of bacteria on Earth. It has only recently been demonstrated that different habitats harbor different amounts of prokaryotic diversity (**Figure 2**). For instance, Torsvik *et al.* (2002) used DNA–DNA reassociation methods to quantify the total genomic diversity in different communities. They estimate that a few grams of soil or sediment contain thousands of bacterial and archaeal "species." In comparison, aquatic environments appear to support fewer bacterial taxa (Torsvik *et al.*, 2002).

Because of their high taxonomic and evolutionary diversity, bacteria are incredibly functionally diverse. Bacteria have a striking variety of modes of energy conversion, many unique metabolic pathways, and can utilize a wide range of substrates. Bacteria can oxidize and reduce inorganic sulfur and can assimilate atmospheric nitrogen. Many bacteria have an anaerobic energy metabolism and are often the only organisms living in anoxic environments. Bacteria can even degrade xenobiotic compounds, such as herbicides and pesticides, which are toxic to other organisms.

Because archaea require particular effort to culture, the diversity of this group is historically less well understood than

Table 1 Comparison of the total number of species currently described and the number of species predicted to exist. The term "species" is used loosely as it must be defined differently across taxa

Group	*Number of "species"*		
	Described	*Estimated*	*% Described*
Eukaryotes	1.8×10^6	12×10^6	15.0
Fungi	$7.2–10 \times 10^4$	1.5×10^6	5.3
Archaea	217		
Bacteria	5007	$4.5–1000 \times 10^6$	0.1
Viruses	4000	1.0×10^8	0.004

Source: Modified from Martiny JBH and Field D (2005) Ecological perspectives on the sequenced genome collection. *Ecology Letters* 8: 1334–1335.

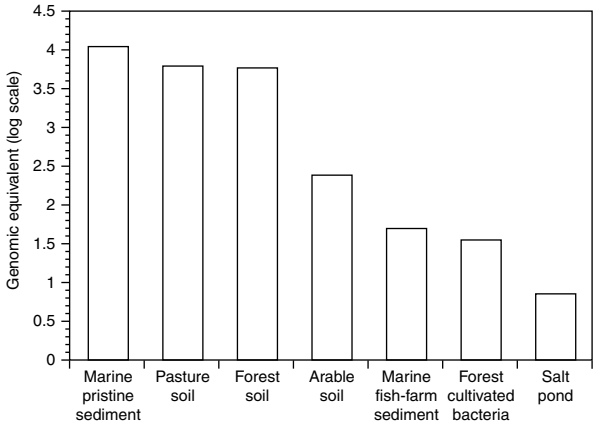


Figure 2 Prokaryotic diversity (genome equivalents) estimated from the reassociation rate of DNA isolated from various habitats. Estimates for pasture soil and arable soil are averages.

bacteria. Molecular phylogenetic surveys, more commonly applied to the study of bacterial communities, are now more frequently being used to characterize the genetic diversity and ecological significance of archaea in a variety of environments including hot springs, deep-sea hydrothermal vents, and coastal waters. A couple hundred archaeal "species" have been described compared with approximately 5000 bacterial "species," and we lack estimates of how diverse these prokaryotes might be.

Archaea have typically been found living at physical or chemical extremes such as high temperature, high salinity, or strictly anaerobic environments and are immensely physiologically and metabolically diverse. Methanogenic archaea, which occupy anaerobic habitats such as sediments, marshes, and hydrothermal vents, produce methane as a by-product of their metabolism, a process restricted to this domain. Recently, archaea have been found in more benign marine, freshwater, and soil environments. For example, it has generally been assumed that specialist bacteria, which convert ammonia into nitrate, a key step in nitrification, are the main ammonia oxidizers in soil. However, it has been recently suggested that archaea may be the most abundant ammonia oxidizers in both pristine and agricultural soils.

Microbial Eukaryotic Diversity

Molecular methodologies that have revolutionized the understanding of bacterial and archaeal diversity have only recently been adapted to study eukaryotic diversity. Perhaps not surprisingly, molecular surveys suggest that morphological characteristics may conceal a large amount of eukaryotic diversity. For example, Slapeta *et al.* (2005) analyzed the rDNA genes of protists in two ponds and found high within-morphotype diversity. In addition, they detected a number of sequences very divergent from known protist species, implying that the description of global protist diversity is far from complete. Others argue that while local microbial eukaryotic diversity is high, global diversity may be relatively low. They document that many free-living protists (as defined by morphology) are cosmopolitan and suggest that nearly all free-living ciliates have already been discovered.

Like protists, fungal species have been traditionally defined morphologically. Hawksworth (1991) suggests that there are at least 1.5 million fungal species globally. Somewhere between 72,000 and 100,000 fungal species have been described, suggesting that as little as 5% of fungal diversity has been discovered. Moreover, with the application of genetic methods such as those applied to bacteria and archaea, estimates of global fungal diversity will likely increase substantially. For example, scientists surveyed the genetic diversity of fungi in the roots of one grass species and found 49 different fungal OTUs, only seven of which were similar to known DNA sequences, and only 6% of which have known ecological roles.

Viral Diversity

Given the remarkable diversity of Bacteria, Archaea, and Microbial eukaryotes, it is perhaps not surprising that the viruses that infect them appear to be similarly abundant and diverse. Viral abundance is estimated to be of the order of 10^{31}, the majority of which infect bacteria and archaea. One kilogram of marine sediment may harbor a million different viral genotypes, and 200 l of seawater contain more than 5000 viral genotypes. In general, viral abundance is well correlated with bacterial abundance, and in fact, bacteriophages are estimated to infect and kill between 4% and 50% of bacteria produced daily in aquatic environments. How, or even whether, viral diversity varies by habitat type is unclear. Researchers have reviewed the extent and distribution of viral diversity in marine habitats and proposed that viruses are highly diverse on a local scale but that this diversity does not vary much over space. In other words, local viral diversity may be representative of global viral diversity.

Maintenance of Microbial Diversity

Ecologists have long been motivated to understand the factors that influence the extraordinary diversity of life on Earth. However, despite the extent of microbial diversity, most of what we know about the maintenance of diversity is based on studies of plants and animals. For example, Hutchinson (1961) observed two species of water-bugs in a homogenous aquatic environment and wondered how they could coexist if they both use similar nutrient resources; the same is true about the thousands of microbial taxa that coexist in all habitats. Although a synthetic view of microbial diversity might not be possible, it is useful to consider whether universal patterns and processes govern the diversity of all organisms, large and small.

From an evolutionary perspective, high microbial diversity is created because the rate of diversification exceeds the rate of extinction. If laboratory studies are indicative, the potential for rapid diversification seems greater in microorganisms than macroorganisms. Bacteria and Archaea, in particular, have several traits that make rapid speciation likely. Because they exist in large populations and have rapid growth rates, they have the capacity to accumulate large numbers of novel mutations, which can generate substantial genetic variation. They can also readily take up DNA and get rid of extraneous DNA. Gene transfer and recombination may also accelerate

diversification further. Less is known about the extinction rates of microbes relative to larger organisms; however, many researchers have posited that extinction rates of microbial taxa may be low because many microorganisms have large population sizes and have an ability to survive in a range of environments and tolerate harsh conditions.

From an ecological perspective, once this diversity is created, the question remains how this incredible evolutionary diversity coexists in contemporary communities. Some have reviewed ecological factors, known to be of importance to plant and animal communities, which may also influence the diversity of microorganisms. Below, the authors discuss four of these factors (environmental heterogeneity, disturbance, productivity, and dispersal), and review recent studies that test whether these factors are important for environmental microbial communities.

Environmental Heterogeneity

Among large organisms, species diversity increases with the heterogeneity of habitat types available. For instance, bird species diversity increases with foliage height diversity and mite diversity increases with the number of habitat types in soil subhorizons. Both laboratory and field studies support the hypothesis that increased environmental heterogeneity also increases bacterial diversity. For example, spatial isolation in the soil matrix, created by low moisture content, plays an important role in determining community composition. Zhou *et al.* (2002) observed that soils saturated with water had fewer bacterial taxa and a less even distribution of taxa than patchy, unsaturated soils. Additionally, researchers found that bacterial diversity increased significantly by adding flow heterogeneity in stream mesocosms. Treves *et al.* (2003) examined the competitive dynamics of two species growing on a single resource in a uniform sand matrix under varied moisture content. One species dominated the community under saturated conditions, which were also highly connected, suggesting that these conditions allow competitive interactions to structure the community. As moisture content and connectivity decreased, the less-competitive species became established in the community. Finally, another study similarly found that litter bags with a mixture of species litter contained a more diverse bacterial and fungal community that litter bags with only a single type of species.

Disturbance

Most communities are subject to different types of disturbances that vary in both frequency and intensity. Events that disrupt the physical environment or alter resource and substrate availability are known to influence plant and animal community diversity. Across a wide range of taxa, including corals, butterflies, plankton, and tropical trees, diversity peaks at intermediate intensities or frequencies of small-scale disturbances. For example, algal diversity on intertidal boulders is highest when waves overturn the boulders, exposing new substrate for colonization.

It is difficult to test whether the intermediate disturbance hypothesis explains patterns of microbial diversity in the wild because disturbance is typically confounded with other factors such as plant cover or soil structure. Field and laboratory studies that have examined the relationship between disturbance and bacterial diversity have generally found that the intermediate disturbance hypothesis holds. For example, Bruce *et al.* (1995) examined the diversity of mercury-resistance genes in sediment-associated bacteria at sites with varying levels of mercury contamination. Consistent with predictions of the intermediate disturbance hypothesis, sites exposed to intermediate levels of mercury had the highest genetic diversity and both pristine and heavily contaminated sites had low diversity. However, other studies have found that disturbance events always result in decreased diversity. For example, Walsh *et al.* (2005) compared archaeal 16S rRNA gene diversity along a salinity gradient prone to high-salinity fluctuations. Heavy-rainfall events impose local high mortality of haloarchaea, and this disturbance regime appears to promote archaeal diversity. Similarly, another study found that the richness of bacterial communities decreased when exposed to low and high doses of the pollutant pentachlorophenol (PCP). Thus, the effect of a disturbance on a microbial community's diversity appears dependent on the type of disturbance as well as the original microbial community.

Productivity

Productivity, the rate at which energy flows through an ecosystem, has long been considered a primary determinant of both plant and animal diversity. A current ecological paradigm is that peak diversity occurs at intermediate levels of productivity, even though this relationship appears to be scale-dependent.

A number of studies consider the relationship between primary productivity and microbial diversity. Productivity appears to be related to microbial diversity; in some instances, these patterns resemble those of plants and animals. For example, Kassen *et al.* (2000) grew cultures of the bacterium *Pseudomonas fluorescens* over a wide range of nutrient concentrations and found that diversity peaked at intermediate productivities. However, another study surveyed two lagoons with different levels of primary productivity and found a greater number of unique bacterial rRNA gene sequences in the more productive lagoon, suggesting that bacterial taxonomic richness increases with productivity. Finally, Horner-Devine *et al.* (2003) estimated bacterial taxonomic richness along a primary productivity gradient in freshwater mesocosms. Although overall bacterial richness was not significantly related to productivity, productivity did influence the richness of separate taxonomic groups. For example, some bacterial groups exhibited a significant hump-shaped relationship with productivity while other groups, such as the α-Proteobacteria, exhibited a U-shaped relationship with primary productivity. Thus, the relationship between ecosystem productivity and microbial diversity appears to be complex and may vary widely across environments.

Dispersal

The small size and high abundance of microorganisms have led microbiologists to generalize that microbial dispersal rates to new habitats is very high, overwhelming spatial differences in taxon abundance. For example, Schauer *et al.*

(2010) examined the bacterial communities in three oceanic basins of the eastern South Atlantic Ocean that were all separated by the Walvis Ridge, a large physical barrier. They found the microbial communities were highly conserved across sites, suggesting a high dispersal capability, despite the physical barrier. However, debate over this view still exists, as other studies demonstrate large biogeographic differences among microbial communities. This evidence suggests that microbial diversity might be influenced by dispersal and isolation patterns. Based on this theory, it has been argued that local bacterial community composition is in part a product of stochastic dispersal events, rather than solely niche differentiation among taxa. The importance of dispersal relative to other factors influencing microbial diversity requires further investigation.

Importance for Ecosystems

Microorganisms are the drivers of key ecosystem processes, including decomposition, nutrient mineralization, and trace gas emission and consumption. Microorganisms are intimately involved in critical pathways of biogeochemical cycles and are the major pathways by which some elemental forms are regenerated to be used by other organisms.

For instance, microbes have a central role in almost all aspects of the nitrogen cycle. Nitrogen fixation is the process by which atmospheric nitrogen is converted into ammonia, which is quickly ionized to ammonium, a form of nitrogen usable by plants. This process, carried out by nitrogen-fixing bacteria, is essential for maintaining soil fertility and agricultural productivity.

If an entire functional group, such as the nitrogen-fixing bacteria, were removed from an ecosystem, the resulting changes would dramatically alter resource supply, processing rates, and plant productivity. However, below this broad functional group classification, it remains unclear whether microbial diversity matters for ecosystem processes. Recently, numerous studies have highlighted the incredible "microdiversity" of bacteria in some habitats. For example, Acinas *et al*. (2004) sampled 16S rDNA sequences from a coastal bacterioplankton community. They found 516 unique sequences, 50% of which fell into clusters containing less than 1% sequence divergence. Thus, one current avenue of research is to understand whether this type of fine-scale diversity or other levels of diversity have consequences for ecosystem processes.

Microorganisms and Global Change

Terrestrial and Marine Ecosystems

The composition of microbial communities is known to vary with land-use type, temperature, agricultural growing practices, nutrient status, pollutants, and other environmental variables associated with global change. Additionally, there is some evidence that certain properties associated with entire microbial communities can be altered by global changes, such as total microbial biomass, rates of respiration, and biogeochemical transformations. Little is known, however, how anthropogenic impacts associated with global change will impact microbial communities, and how those changes will ultimately affect ecosystem functioning.

Several studies have examined the links between microbial composition and terrestrial ecosystem functioning and support the hypothesis that global changes will have important consequences for microbially mediated ecosystem processes. Horz *et al*. (2004) considered the impact of multiple global changes on the community composition of ammonia-oxidizing bacteria (AOB) in a California grassland. The authors observed changes in the community structure and abundance of free-living soil bacteria to simultaneous increases in atmospheric CO_2, precipitation, temperature, and nitrogen deposition. The response of AOB was similar to the observed plant responses: Composition and abundance was altered by nitrogen, total abundance decreased in response to elevated CO_2, and both strongly interacted with temperature and precipitation. In contrast, methane-oxidizing bacteria showed a unique response to increased precipitation and temperature; the relative abundance of common methanotroph clades increased in response to both variables, while novel clades had varied responses. Two of the novel clades did not respond to these environmental changes, while one novel clade increased in relative abundance to increased temperature and precipitation.

Further, studies of feedbacks between microbial and plant communities illustrate the potentially unpredictable nature of ecosystem responses to global change. In a one-year factorial field experiment, researchers tested the influence of decreased plant diversity, elevated CO_2 levels, and increased nitrogen deposition on the severity of disease caused by fungal pathogens in a diverse grassland community. All three treatments substantially increased disease severity in the plant community. Because pathogen loads also affect grassland ecosystem processes, this result suggests a route in which microbes respond to climate change and simultaneously feedback and affect climate change as well. Mychorrizal fungi may also serve as a feedback to the climate system under increased atmospheric CO_2, because they sequester increased amounts of carbon in living, dead, and residual hyphal biomass. In contrast, nitrogen deposition may increase turnover rates of fungal tissue and negate CO_2 effects on fungal biomass.

In marine systems, examination of the impact of global change on microbial communities has only recently begun. Global change is anticipated to impact the physical marine environment in a variety of ways, including temperature increases, shifts in wind and radiation regimes, changes in the frequency of episodic extreme events (such as storms), and acidification. These changes will likely deeply affect the structure and functioning of marine ecosystems, including marine microbial food webs. Recent evidence suggests that global changes, including rising water temperatures, will have dramatic affects on marine microbial systems. For example, Hoppe *et al*. (2008) conducted a mesocosm experiment to examine the effects temperature increases would have on marine microbial communities. When seawater was warmed by 2, 4, and 6 °C, the authors reported an acceleration of bacterial degradation of organic matter derived from a

phytoplankton spring bloom, as well as increased respiration rates. Similarly, others incubated samples in microcosms at different temperatures over short periods of time (24–48 h), and measured the different variables of interest throughout a seasonal cycle in a coastal Mediterranean site. In this study, the warmed samples had on average 20% higher total bacterial carbon demand than unwarmed samples, suggesting a positive feedback loop between warming of coastal waters and CO_2 production.

Overall, predicting the impact of future global change on ecosystems will require a much greater understanding of the response of microorganisms to a variety of environmental variables. Additional research needs to be conducted in both terrestrial and marine ecosystems. Finally, linking the marine and terrestrial components together will require knowledge of the feedbacks that microorganisms will have back on their environment.

Microbial Diseases

The relationship between emerging microbial diseases and global change is a topic of great concern. Many pathogens of terrestrial and marine taxa are sensitive to temperature, rainfall, and humidity. Therefore, global change has the potential to alter the risk of disease through changes in pathogen development and survival rates, disease transmission, and host susceptibility.

So far, evidence suggests that global change will influence a variety of plant, wildlife, and human diseases. For example, several plant diseases are more severe after mild winters or during warmer temperatures, which implies that directional climate warming will alter plant disease severity through longer growing seasons and accelerated pathogen development. Additionally, Kushmaro *et al*. (1998) showed that the bacterium *Vibrio shiloi*, a marine pathogen that infects and causes the bleaching of the coral *Oculina patagonica*, grows well at temperatures close to or exceeding probable host optima.

Research suggests that the most severe and least predictable disease outbreaks might occur if global change alters host or pathogen geographic ranges, resulting in the convergence of formerly separated species and populations. Vector-borne diseases are likely to be the strongest candidates for altered abundance and geographic range shifts because rising temperatures seem to affect vector distribution, parasite development, and transmission rates. Vector-borne diseases of livestock, particularly African horse sickness and bluetongue viruses, recently expanded their ranges. Additionally, vector-borne human pathogens such as malaria, African trypanosomiasis, Lyme disease, tick-borne encephalitis, yellow fever, and dengue have increased in incidence or geographic range in recent decades. Most of these diseases have expanded into regions of higher latitude, in each case accompanied by apparent expansion in the ranges of mosquito, tick, and midge vectors. Research to determine whether these expansions are due primarily to global change or other anthropogenic factors is still ongoing, as well as predicting future distributional changes in disease prevalence.

Although associations between climate and disease do not necessarily imply causation, results from correlational studies and short-term experiments help scientists separate the effects of climate and other components of global change. Overall, further research must be done to improve the ability to predict impacts of global change on microbial diseases.

See also: Archaea, Origin of. Biogeochemical Cycles. Microorganisms (Microbes), Role of

References

Acinas SG, Klepac-Ceraj V, and Hunt DE (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430: 551–554.

Bruce KD, Osborn AM, Pearson AJ, Strike P, and Ritchie DA (1995) Genetic diversity within mer genes directly amplified from communities of noncultivated soil and sediment bacteria. *Microbial Ecology* 4: 605–612.

Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, and Bork P (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* 311: 1283–1287.

Dykhuizen DE (1998) Santa Rosalia revisited: Why are there so many species of bacteria? *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology* 73: 25–33.

Hawksworth DL (1991) The fungal dimension of biodiversity: Magnitude, significance and conservation. *Mycological Research* 95: 641–655.

Hoppe HG, Breithaupt P, Walther K, *et al.* (2008) Climate warming in winter affects the coupling between phytoplankton and bacteria during the spring bloom: A mesocosm study. *Aquatic Microbial Ecology* 51: 105–115.

Horner-Devine MC, Leibold MA, Smith VH, and Bohannan BJM (2003) Bacterial diversity patterns along a gradient of primary productivity. *Ecology Letters* 6: 613–622.

Horz H-P, Barbrook A, Field CB, and Bohannan BJM (2004) Ammonia-oxidizing bacteria respond to multifactorial global change. *Proceedings of the National Academy of Sciences of the USA* 101: 15136–15141.

Hutchinson GE (1961) The paradox of the plankton. *American Naturalist* 95: 137–145.

Kassen R, Buckling A, Bell G, and Rainey PB (2000) Diversity peaks at intermediate productivity in a laboratory microcosm. *Nature* 406: 508–512.

Kushmaro A, Rosenberg E, Fine M, Ben Haim Y, and Loya Y (1998) Effect of temperature on bleaching of the coral *Oculina patagonica* by *Vibrio* AK-1. *Marine Ecology Progress Series* 171: 131–137.

Martiny JBH and Field D (2005) Ecological perspectives on the sequenced genome collection. *Ecology Letters* 8: 1334–1335.

Pace NR, Stahl DA, Lane DJ, and Olsen GJ (1986) The analysis of natural microbial populations by ribosomal RNA sequences. *Advances in Microbial Ecology* 9: 1–55.

Schauer R, Bienhold C, Ramette A, and Harder J (2010) Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. *ISME* 4: 159–170.

Slapeta J, Moreira D, and Lopez-Garcia P (2005) The extent of protist diversity: Insights from molecular ecology of freshwater eukaryotes. *Proceedings of the Royal Society of London B Biological Sciences* 272: 2073–2081.

Torsvik V, Ovreas L, and Thingstad TF (2002) Prokaryotic diversity – magnitude, dynamics, and controlling factors. *Science* 296: 1064–1066.

Treves DS, Xia B, Zhou J, and Tiedje JM (2003) A two-species test of the hypothesis that spatial isolation influences microbial diversity in soil. *Microbial Ecology* 45: 20–28.

Venter JC, Remington K, Heidelberg JF, *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66–74.

Walsh DA, Papke RT, and Doolittle WF (2005) Archaeal diversity along a soil salinity gradient prone to disturbance. *Environmental Microbiology* 7: 1655–1666.

Zhou J, Xia B, Treves DS, *et al.* (2002) Spatial and resource factors influencing high microbial diversity in soil. *Applied and Environmental Microbiology* 68: 326–334.

Microbial Biodiversity, Measurement of

Kate M Scow, Egbert Schwartz, Mara J Johnson, and Jennifer L Macalady, University of California, Davis, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 177–190, © 2001, Elsevier Inc.

Glossary

Clone A population of cells all descended from a single cell; a number of copies of a DNA fragment obtained by allowing the DNA fragment to be replicated by a phage or plasmid.
Fingerprint A pattern produced by DNA fragments or lipids that represents a community or species.
Function A process carried out by an organism or group of organisms.
Genotypic Relating to the genetic composition of an organism.
Oligonucleotide A short, single-stranded nucleic acid molecule either obtained from an organism or synthesized chemically.
Phenotypic Relating to observable characteristics of an organism.
Phylogeny The ordering of species into taxonomic groups based on evolutionary relationships.
Polymerase chain reaction A method for amplifying DNA *in vitro* involving the use of oligonucleotide primers complementary to nucleotide sequences in a target gene and replication of the target sequences by the action of DNA polymerases.
Pure culture An organism growing in the absence of all other organisms.
Species A collection of closely related strains.
Strain A population of cells all descended from a single cell.

Introduction

Microorganisms encompass an extraordinary diversity, in both their taxonomy and their ecological functions. As a group, microorganisms range in size over several orders of magnitude, span the three taxonomic domains (Archaea, Bacteria, and Eukarya), and are uniquely capable of performing all known biochemical transformations. The existence of such an immense variety of organisms, combined with the fact that microorganisms are too small to see without magnification, makes the task of measuring microbial biodiversity challenging and even daunting.

Measuring the biodiversity of microbial communities is important for immediately practical and more fundamental considerations. Understanding microbial biodiversity may translate into benefits for biotechnology; management of agricultural, forest, and natural ecosystems; biodegradation of pollutants; reclamation of damaged lands; waste treatment systems; and biological control systems. On a more basic level, microbial processes underlie many essential processes shared by all ecosystems. However, because of our current minimal understanding of these groups of organisms, microbial communities are treated as black boxes in most studies of community, ecosystem, and landscape ecology. Increased knowledge of microbial biodiversity would clarify the details of many ecosystem processes and help generate the baseline information needed to predict and perhaps ameliorate the effects of global climate change.

This article discusses major issues involved in measuring microbial biodiversity and provides an overview of methods for measuring microbial diversity in natural, managed, and engineered ecosystems. The content is intended for biologists and ecologists who desire to incorporate the often-neglected microbial taxa into surveys of biodiversity.

Biodiversity of Microorganisms

The unit of measurement used in biodiversity studies of macroscopic organisms is usually a taxonomic one: the species. Microbial taxonomy has undergone enormous changes in the past two decades. Traditionally in microbiology, a species was phenotypically characterized with organisms classified into taxonomic groups based on morphology, physiology, and metabolism. Recently, however, phenotypic classification has been replaced or supplemented by the use of genotypic analysis. Nucleic acid sequence information derived from the small subunit of the ribosome (16S ribosomal RNA (rRNA) for prokaryotes or 18S rRNA for eukaryotes) is used to determine the degree of similarity among groups of organisms and the evolutionary relationships of microorganisms and all other life-forms. Certain analytical methods use the DNA sequence that codes for the rRNA rather than the rRNA sequence itself. The term used to describe this DNA is ribosomal DNA (rDNA). Although genetic sequencing for individual microbes is labor-intensive, particularly when characterizing complex microbial systems, some microbial ecologists maintain that an in-depth understanding of recently described systems is not possible until this information is gathered.

Far more species of bacteria, fungi, and other microorganisms exist than have been described to date. For example, between 300,000 and 1 million species of bacteria are estimated to exist based on initial in-depth phylogenetic surveys of several specific ecosystems and on calculations based on the reannealing kinetics of DNA extracted from natural environments. In contrast, only approximately 3800 species have been isolated and described by culture-based analyses. As genetic sequence analyses progress, libraries of small subunit rRNA sequences from specific ecosystems are created and compared to sequences in the existing database of known

species. Often, these new sequences exhibit a low degree of similarity compared with the sequences of known species.

Background on Measuring Microbial Biodiversity

Although measuring the biodiversity of animal and plant communities involves counting and identifying individual species, such an approach is neither practical nor feasible for microorganisms. Though individual microorganisms can be counted using microscopy, their morphologies do not reflect their diversity. Visualization reveals little that is defining about identity beyond the observations that an individual cell is a fungus or bacterium, is a particular shape (rod or coccus), or has certain cell wall properties (gram positive or gram negative). Furthermore, counting is often hindered by interference from the physical environment in which microorganisms live. Some of the traits used as classification criteria for microorganisms are relatively plastic and, given that horizontal gene transfer appears common among microorganisms, using phenotypic information to classify species in microbiology is problematic. Additionally, the end point of many methodologies used to characterize microorganisms is not species identification. Thus, extrapolating the concept of species used in classifying large eukaryotes (plants and animals) to microorganisms is questionable.

Another important issue in exploring microbial diversity is the question of whether taxonomic (e.g., species or family groups present) or functional (e.g., types of processes present) diversity is of interest. Some important microbial processes include the transformation of elements in biogeochemical cycles, decomposition of organic residues, biodegradation of pollutants, fluxes of gases to and from the atmosphere, symbiotic relations with plants, and parasitism of other organisms. In ecological studies, the processes carried out by microorganisms within an ecosystem may be of greater relevance than knowledge of the specific identities of organisms involved. A considerable amount of functional redundancy among microorganisms is thought to exist in most communities based on the fact that there are far more taxa than processes. To date, specific functions have been difficult to measure and interpret, especially by nonmicrobiologists, in studies of ecosystem biodiversity. However, with the development of rapid molecular assays, measuring microbial community compositions has become more feasible, and taxonomic characterization is more widely used. Compiling functional and taxonomic diversity measurements is important given that the ecological roles played by organisms represented by newly discovered ssrDNA sequences are often unknown.

Numerous methods are available for characterizing microbial communities. Determining which approaches to use is a critical step and depends on the objectives of the study. Should the diversity of a community be measured by characterizing its individual members, a task perhaps too large to be feasible, or by directly characterizing the community as a whole? If taxonomic diversity is of concern, what level of resolution is needed? Given that microorganisms are so small, what is the appropriate physical scale on which to sample in a study of biodiversity? These points are discussed later.

Enrichment/Isolation versus Direct Measurement of Microbial Communities

Early studies of microbial diversity involved culturing of bacterial isolates on enrichment media supporting their growth ("culture-based") followed by phenotypic characterization of these isolates based on their morphology, physiology, and specific metabolic capabilities. As the tools of molecular biology began to be applied to microbial ecology, it became evident that culture-based methods were failing to detect large portions of microbial communities (no more than 0.1–5% of bacteria in communities are detected using culture-based techniques). Rapidly growing strains of microorganisms were often favored over other members of the community. A culture-based approach therefore probably offers a skewed perspective of microbial diversity in many environments. Consequently, there has been a shift in studies of microbial diversity toward the use of molecular biology and other biochemical and molecular tools that do not require culturing of organisms. Characterization of taxonomic and functional diversity is conducted using information provided by nucleic acids, lipids, and other cellular constituents. If culture conditions of a microorganism are known, data are gathered after individual strains have been separated from a community using enrichment and isolation techniques. In other cases, especially in soils and sediments, biochemical constituents may be directly extracted from the environmental samples, bypassing the isolation step. Many of these methods are currently under development and new ones are being developed at a rapid pace. Those seriously engaged in studies of microbial diversity have to keep close tabs on emerging trends in the field.

Newer methods may introduce their own biases in the assessment of microbial biodiversity, although these biases are presumably not as severe as the culturing bias. For example, molecular methods often make use of the polymerase chain reaction (PCR), during which original copies of specific DNA sequence are amplified exponentially. Two important considerations are that amplification efficiency varies among DNA sequences and therefore PCR results are rarely considered to be quantitative, and that a DNA sequence with low amplification efficiency could conceivably be prominent in the environment but not appear in PCR-based diversity measurements (*see* Nucleic Acids).

These direct measurement techniques may also be used to amplify sequences of rRNA or messenger RNA (mRNA) in order to shed light on the distributions, relative activities, and gene expression (functions) of specific groups.

Level of Taxonomic Resolution

When measuring microbial taxonomic diversity, a decision must be made about the level of taxonomic resolution (e.g., species, genus, family, order, and domain) needed to sufficiently assess diversity. The very definition of species in microbiology is subject to intense debate. Though counting

numbers of individual species can be relatively straightforward in studies of plants and animals, this approach is not feasible for microorganisms. There is an overwhelming amount of diversity among microorganisms and categorizing all types is beyond the scope and resources of most projects. In some cases, determining families, or even kingdoms, within a previously uncharacterized community or ecosystem provides novel and valuable information.

For instance, the DNA from microbial communities as a whole may be presented in the form of a "fingerprint," which provides a means to classify and describe systems. DNA fingerprints are an abstraction of the original genetic information, but they may be an efficient way to characterize complex microbial habitats in which there is little information linking community structure and community function. They also offer the means to generate hypotheses, test responses of systems to environmental changes and variables, and detect shifts in microbial communities. Efforts may be concentrated on conducting sequence analysis on changing elements of a DNA fingerprint rather than on sequencing all the members of a microbial community.

Environmental Heterogeneity

Obtaining a representative sample of the microbial diversity of a particular environment may be very difficult. It is obvious that most environments are heterogeneous at the spatial and temporal scale of humans. This heterogeneity is even more pronounced at the scale of microorganisms, often in ways that are not immediately evident to humans. For example, microbial communities within the root zone of plants, in which available carbon is high and soil chemistry has been altered, differ substantially from immediately adjacent communities in unrooted soil. These different communities are often only millimeters apart from one another. Also, a community that temporarily forms in a pocket of decaying plant or animal matter may be strikingly different from a community only several centimeters away. Within sediments, communities change substantially with depth as a function of oxygen availability and carbon distribution. This concept is also applicable on the scale of individual soil aggregates. In assessing biodiversity, one approach is to identify important microenvironments within the ecosystem of interest and measure diversity within each major type of microenvironment. With this approach, the number of samples obtained may soon become overwhelming. Another approach is to combine samples from different microenvironments in an attempt to understand the heterogeneity within the entire ecosystem, in which case it may be difficult to interpret factors governing the diversity. An understanding of the underlying physical structure of an environment is a prerequisite to a good sampling design.

Methods for Measuring Microbial Diversity

Far more methods for measuring microbial diversity exist than can be presented here. The methods described here are commonly used and, in composite, represent the breadth of approaches available. Methods can be divided into those that involve the counting of cells, measurement of cellular constituents, or determination of activity. Most methods provide either taxonomic or functional information about communities, whereas only a few (e.g., nucleic acid-based methods) have the potential to integrate both types of information. The following methods were originally developed in pioneering studies of bacteria, but most are also suitable for studies of fungi and other eukaryotes. Commonly, studies of fungal diversity involve isolation and characterization of strains rather than being community-based. Increasingly, community-based molecular methods are being used for eukaryotic microorganisms.

Counting Microorganisms

Since the development of the microscope, microbial population numbers, and to some degree diversity, have been quantified by examining cells under high magnification. Selective growth media, provided either in petri plates or in most probable number (MPN) assays, permit enumeration of culturable organisms that can metabolize specific compounds. Plates and MPN assays thus give potential functional information about a community. Counting organisms by the latter method requires that organisms grow sufficiently to be detectable.

Microscopy

Major microbial groups, such as bacteria and fungi, can be distinguished by their morphologies as observed under a light or fluorescence microscope after staining the cells. The microscope provides a quick but relatively superficial survey of microbial diversity based on the sizes, shapes, and staining properties of microorganisms. The relatively low number of morphotypes among bacteria limits descriptions of diversity. To some extent, fungal diversity can be investigated by examining hyphal and fruiting body structures and spore morphology, if present. Taxonomic resolution can be improved by using dyes tagged with fluorescent monoclonal antibodies specific to groups or strains, which permit selective visualization of members of the group to which the antibody binds. Recently, development of fluorescent *in situ* hybridization (FISH) with specific DNA probes (described later) has made the microscope a powerful and highly quantitative tool for use in studies of diversity. The antibody and DNA tagging methods require initial characterization of the organism or group of interest in order to develop the antibody or probe.

Plate Counts

Plate counts exploit the fact that individual microscopic cells quickly grow into a colony visible to the unaided eye if provided with suitable growth conditions. Microbial diversity has been assessed by culturing microorganisms on solid nutrient media to which substances are added to specifically promote the growth of target organisms and/or inhibit the growth of unwanted groups. For example, media containing complex organic substances combined with a low pH tend to select for fungi instead of bacteria. Colonies that grow on the plates

are counted and may be further differentiated based on their color or other morphological properties. The diversity of culturable microorganisms is estimated by counting the number of different colonies present on agar media inoculated with a dilution series of the microbial community. Diversity measured based on this technique is unlikely to yield comprehensive, ecologically relevant information about the microbial community as a whole. An advantage of using plating techniques is that strains that grow on plates can often be easily isolated, characterized, and identified by traditional methods of microbiology.

MPN Methods

MPNs make use of specifically designed culturing conditions to estimate the number of microorganisms in a community able to carry out specific functions. The MPN medium is designed to select a specific trophic group by providing carbon, energy, nutrients, and environmental conditions needed to support the growth of that group. Thus, iron reducers are selected by providing anaerobic conditions, a carbon source, ferric iron, and other nutrients. A series of dilutions of an environmental sample are inoculated into the MPN medium. Measurements of turbidity, substrate utilization, or product formation confirm cell growth or activity. The number of organisms capable of carrying out the particular function being investigated is estimated from cell counts from the series of dilutions and with standard statistical calculations. MPN estimation of numbers is obviously a culture-dependent method, biased toward those organisms able to grow and compete under laboratory conditions.

Analysis of Cellular Constituents

Nucleic Acids

Methods using information contained in nucleic acids, DNA and RNA, provide the most specific information about an organism and thus potentially the most detailed information about the biodiversity of a microbial community. Figure 1 provides an overview of nucleic acid-based methods used in studies of microbial biodiversity. Extraction of DNA or RNA from environmental samples is the first challenge in using nucleic acid-based methods. One approach is to first extract cells from environmental samples (usually from aquatic environments) and then extract cellular DNA or perform analyses on cells collected on filters (e.g., using *in situ* hybridization). *In situ* approaches are quantitative because taxonomic information can be correlated with actual cell numbers within the community. The other approach, direct extraction of DNA without first extracting cells, is more commonly employed in complex environmental media such as soil and sediment. Numerous approaches have been developed to overcome extraction problems associated with specific types of environments, such as soils with high clay or humic acid contents. Other methods have been optimized for rapid extraction to accommodate large numbers of samples. There is usually a trade-off between speed of extraction and quality of extracted DNA (e.g., with respect to purity of the DNA). The efficiency of DNA extraction associated with either the direct or indirect method is difficult to estimate and may range from 10 to 99%. Extraction efficiencies associated with different groups of microorganisms differ, which could skew conclusions about a community diversity.

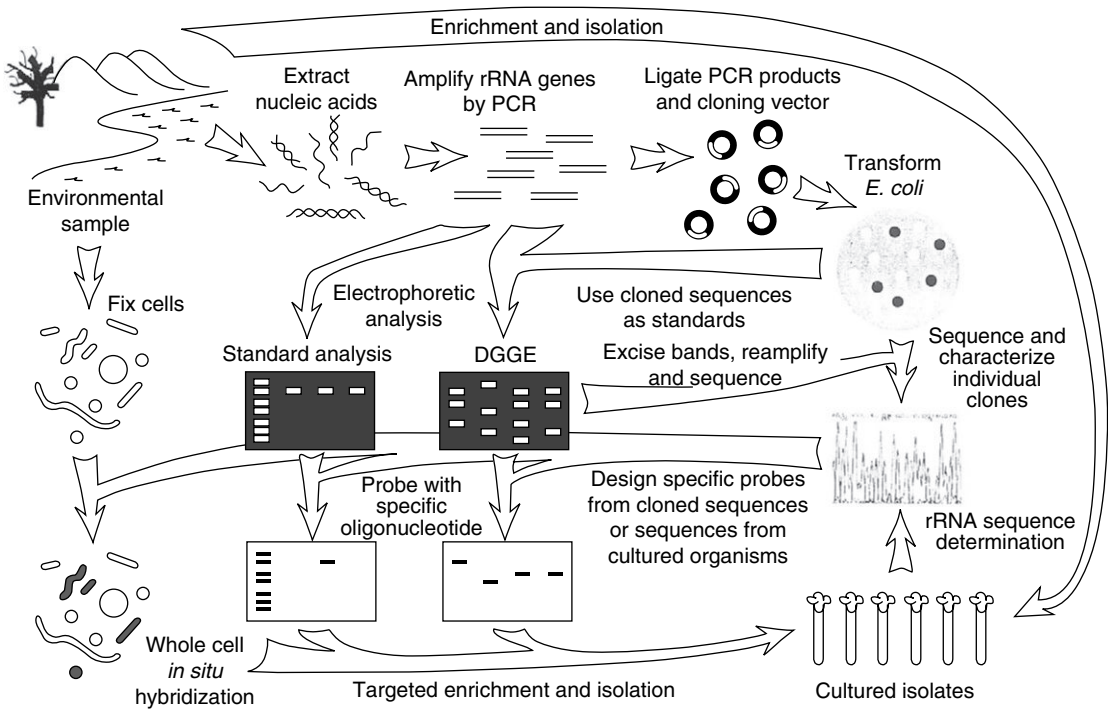


Figure 1 A schematic of commonly used approaches in nucleic acid-based microbial ecology studies (reproduced from Head IM, Saunders JR, and Pickup RW (1998) Microbial evolution, diversity, andecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. *Microbial Ecol* 35: 1, with permission from Springer).

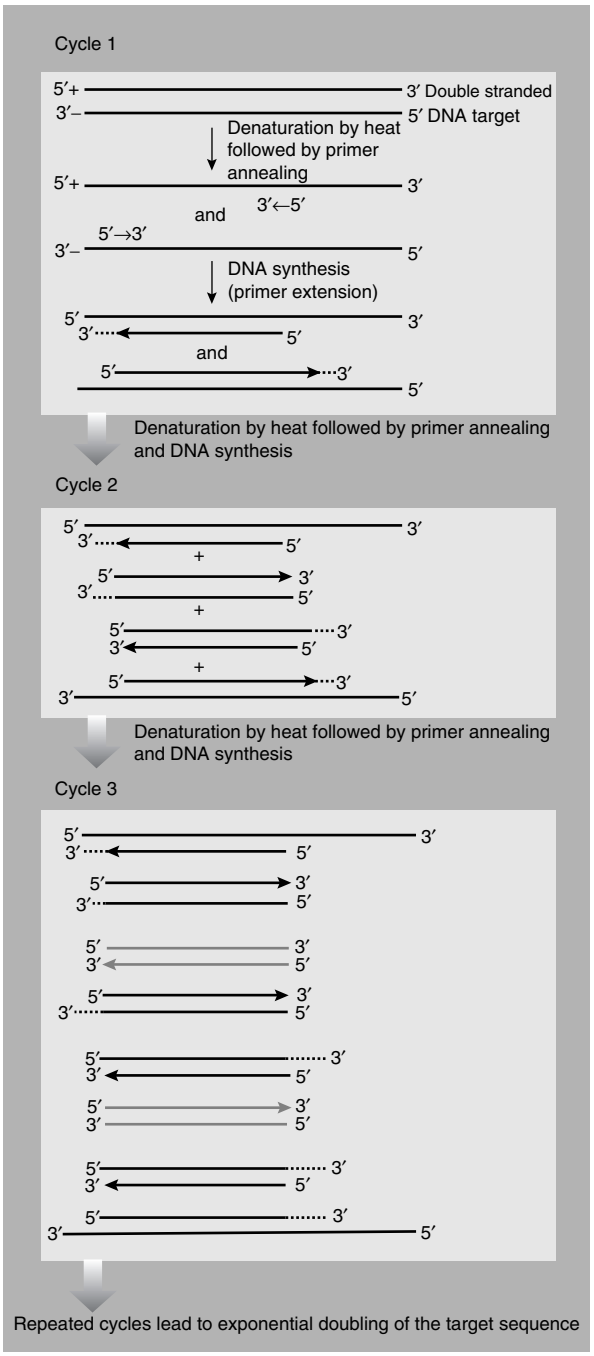


Figure 2 The polymerase chain reaction is used to amplify portions of double-stranded DNA from mixtures of DNA from environmental samples. Typically, 30 cycles are completed to generate thousands of copies of the sequences of interest (reproduced from Old RW and Primrose SB (1994) *Principles of Gene Manipulation*, 5th ed. Oxford: Blackwell, with permission from Wiley).

Nucleic acid-based approaches can be divided into those that employ PCR and those that do not. The PCR makes multiple copies of a specific fragment of RNA or DNA from a mixed pool of nucleic acid fragments and permits detection of sequences originally present at very low densities (**Figure 2**). After PCR amplification, fragments are separated in a polymer gel on the basis of their length or nucleotide composition. Concentrations of the fragments form "bands" in the gel which are visualized by staining. The pattern made by the bands constitutes a "fingerprint" of a community. PCR is also useful for detecting specific strains, species, or phylogenetic groups in environmental samples and, if used this way, requires *a priori* knowledge about sequence variability/conservation in the species or group of interest. The use of primer sets that target evolutionarily conserved regions, such as universal bacterial or fungal primers, circumvents this problem. However, in highly diverse communities, universal primers may amplify so many different nucleic acid sequences that it is not possible to separate them on a gel (e.g., a "smear" may result). Primer sets that target evolutionarily variable regions are useful for differentiating closely related organisms.

PCR amplification efficiencies differ among individual DNA or RNA sequences, introducing PCR bias. This is of particular concern because it is difficult to distinguish artifacts of PCR from true contributions from members of the microbial community. For example, PCR-based diversity measurements may be biased toward populations with sequences most complementary to primers in the PCR reaction mixture. Consequently, the relative intensity of a PCR product band should only be used as a qualitative measure of the relative abundance of that sequence within the community in the absence of carefully constructed studies that address this issue. Recently developed quantitative competitive PCR methods address this issue. Thermocyclers with detectors that quantify PCR products in real time are also now available.

To obtain information about gene expression (protein or enzyme production), microbial ecologists may utilize a variation of PCR called reverse transcriptase-PCR (RT-PCR). Gene expression is associated with enzyme production directed by another type of cellular RNA called mRNA. Messenger RNA constitutes a small percentage of total RNA but contains the genetic code for individual cellular enzymes. With carefully controlled procedures to eliminate the ubiquitous enzymes called ribonucleases that degrade it, mRNA may be isolated from microbes in environmental samples. With the enzyme reverse transcriptase, mRNA is converted to single-stranded DNA. Once the code is in the form of single-stranded DNA, PCR is performed to produce multiple copies of the code for the enzyme for sequence analysis. Future applications of mRNA analysis will provide insight about the functional diversity of microbial communities.

Finally, PCR may be used as a preliminary step in random cloning and sequencing DNA or RNA from environmental samples from which a "library" of DNA sequences (clones) is generated to represent the diversity present within a community.

PCR-independent methods provide the option of obtaining taxonomic or functional information from nucleic acid extracts or even from intact microbial cells. Hybridization methods entail combining complimentary strands of nucleic acids. The complementary strands are genomic DNA in the case of solution hybridization. Other hybridization methods utilize short, single strands of DNA or RNA (oligonucleotide probes) complementary to conserved or variable portions of rRNA or genomic DNA. If the sample contains small amounts of target rRNA or DNA, some of these techniques can be used in conjunction with PCR to reach the threshold of detection.

Table 1 Summary of Nucleic Acid Methods

PCR-Dependent Methods
The polymerase chain reaction amplifies the signal of microbial DNA from environmental samples. PCR enables detection of microbial populations from very small sample sizes and from extreme environments containing unculturable microorganisms.
Denaturing/Thermal Gradient Gel Electrophoresis (D/TGGE)
Small subunit rDNA fragments of the same length are separated based on nucleotide sequence to produce fingerprint patterns. Isolated fragments may be excised and cloned for sequence analysis.
Intergenic Transcribed Spacer (ITS) Analysis
Fragments between the small and large rDNA genes are separated based on length producing fingerprint patterns. Isolated fragments may be excised to differentiate strains by sequence analysis.
Restriction Fragment Length Polymorphism (RFLP)
Small subunit rDNA PCR products are cut into smaller fragments with restriction enzymes and separated based on length to produce patterns that differentiate simple communities or strains.
Cloning and Sequencing
DNA PCR products are inserted in plasmid vectors which are taken up by bacterial cells. The cells are then cloned to generate many copies for sequence analysis and phylogenetic classification.
PCR-Independent Methods
Hybridization techniques have all the specificity of PCR-based techniques, are without PCR bias, and have the additional advantage of a means to quantify microorganisms.
Solution Hybridization
This method takes community double stranded DNA through thermal dissociation and reassociation process to estimate community DNA complexity.
Membrane Hybridization
Group or species-specific probes hybridize with community DNA or RNA immobilized on a membrane to produce estimates of relative abundance.
Fluorescent In-Situ Hybridization (FISH)
Fluorescently labeled group or species-specific probes hybridize with RNA in intact cells. Individual cells may be counted and types of organisms quantified by microscopy.
Oligonucleotide array
Group, species-specific, or functional gene probes are immobilized on a solid support. Hybridization between probes and community DNA is detected by laser technology.

Table 1 provides a summary of nucleic acid methods commonly used in studies of microbial biodiversity.

PCR-Dependent Methods

Denaturing or Thermal Gradient Gel Electrophoresis Gradient gel electrophoresis resolves double-stranded DNA fragments of the same size based on properties defined by their genetic sequence. First, the small subunit rDNA is amplified from microbial community DNA using universal (genetically conserved) or specific (genetically variable) primer sets. PCR products are separated on a gel that contains an increasing chemical gradient of formamide and urea, which differentially denatures or separates the double-stranded DNA fragments. A variant on the method (thermal gradient gel electrophoresis) uses a linearly increasing temperature gradient in the gel, which has an analogous effect on denaturing double-stranded DNA fragments. PCR product fragments of the same size migrate down the gel at a similar speed until a portion of a particular fragment denatures. At that point the migration of the fragment virtually ceases. Individual bands often correspond to individual microbial strains; however, some strains may produce more than one band (if the strain contains more than one ssrDNA sequence).

The result of denaturing or thermal gradient gel electrophoresis (D/TGGE) analysis is a banding pattern that reflects the composition of a microbial community. Statistical analysis involves scoring the bands in the sample as present or absent for a particular migration location on the gel, performing a similarity index based on band information, and then evaluating the relationships of samples with cluster analysis algorithms that create dendograms. Although D/TGGE patterns have been used directly to determine Shannon–Weaver diversity indexes, the potential for PCR bias necessitates that extreme care be taken in using these banding patterns as the basis for traditional diversity measurements. Often, up to 60 dominant bands can be discerned in a DGGE profile, whereas it is known that thousands of different species may be present in a community. Thus, D/TGGE does not reveal information about the entire community but rather provides a survey of the dominant members in the community (at least those sequences that amplify by PCR). D/TGGE patterns are particularly rich in information since bands in the gel can be analyzed by hybridization methods or excised for cloning and sequencing.

Figure 3 depicts a polyacrylamide gel of a DGGE fingerprint obtained from a bacterial mixed culture grown on the pollutant MTBE (lane 6) or on trypticase soy broth (a rich medium) (lane 7). Also shown on this gel are bands obtained from individual bacterial strains isolated from the mixed culture (lanes 2–5). The organism depicted by the bands in lane 5 is able to biodegrade MTBE.

Intergenic Transcribed Spacer Analysis The final portion of the ssrRNA gene and the beginning of the large subunit ribosomal RNA (lsrRNA) gene contain highly conserved

sequences, and primers can be designed to amplify the intergenic transcribed spacer (ITS) region (also called ribosomal intergenic spacer analysis). The lsrRNA gene is located downstream from the small subunit gene. The intergenic spacer region consists of both noncoding and coding regions; coding regions include transfer RNA genes in the case of bacteria (**Figure 4**) and 5.8S rRNA genes in the case of eukaryotes such as fungi. The noncoding regions are poorly conserved and contain insertions, deletions, and mutations that do not have a direct effect on the fitness of the organism. Using ITS primers that are universal to bacteria or eukaryotes in a PCR reaction with microbial community DNA generates fragments of different lengths and sequences. These complex banding patterns, like D/TGGE patterns, can be interpreted as a measure of the diversity of the microbial community. As with D/TGGE patterns, a single species may be represented by more than one band. In addition, a single species may have more than one length of ITS region, and one band may represent more than one species. Because closely related strains may have similar 16S rDNA sequences but dissimilar ITS sequences, ITS analysis is an effective tool to detect diversity in situations in which D/TGGE cannot.

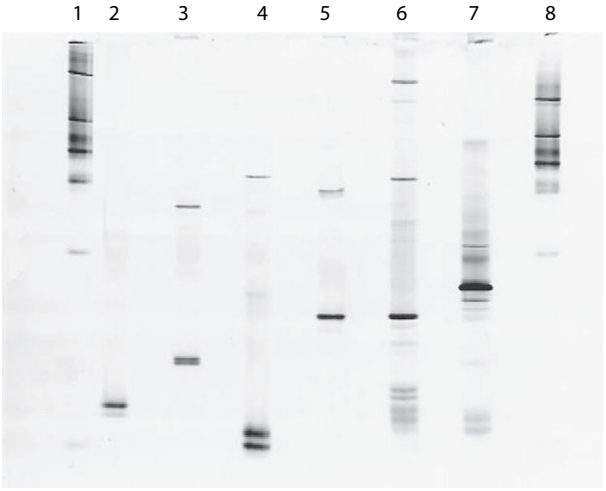


Figure 3 Inverted image of fluorescently stained polyacrylamide gel showing DNA bands in DGGE fingerprints. Lanes 2–4, various bacterial isolates from mixed culture; lane 5, MTBE-degrading isolate from mixed culture; lane 6, mixed culture grown on MTBE; lane 7, mixed culture grown on trypticase soy medium. Lanes 1 and 8 are DNA markers.

Figure 5 shows microbial communities from agricultural soils planted with three different crops discriminated by ITS analysis. The fingerprints were generated from PCR amplification of total soil DNA using universal bacterial ITS primers. Different lanes within each crop category represent different replicates or PCR reactions.

Restriction Fragment Length Polymorphism Small subunit rDNA may also be analyzed by restriction enzymes that recognize and cut double-stranded DNA at precise sequence locations. The fragments derived from the procedure are of different sizes and indicate restriction fragment length polymorphisms (RFLPs). In this method, microbial community ssrDNA is either amplified and then digested directly or individual strains are first cloned and RFLP analysis is performed on each clone separately. Restriction enzyme digestion of the ssrDNA for a single strain typically results in 2–20 DNA fragment lengths, which are separated by gel electrophoresis. The resulting banding pattern is unique and can be used as a fingerprint to distinguish strains or mixtures of organisms from each other. Whole community RFLP fingerprints are extremely complex, resulting in sometimes indecipherable patterns, making this method useful only in systems of low diversity. Cloning the gene first has the advantage of avoiding overly complex community patterns but the disadvantage of limiting investigations due to the additional cost and workload. In terminally labeled RFLP, one of the primers used in amplifying ssrDNA is fluorescently labeled for detection in an automated DNA sequencer. As a result, the fingerprint is less complicated than standard RFLP fingerprints because only one of the fragments formed by the restriction enzyme is detected in the analysis.

Sequencing of ssrDNA Clone Libraries The most direct method for analysis of ssrDNA sequences obtained from environmental samples is to clone and sequence them. In cloning, individual PCR products are integrated into a vector (vehicle for reproduction), such as a bacterial plasmid, so that individual ssrDNA sequences can be reproduced and analyzed. Although this has traditionally been a time-consuming approach, the increasing availability of automated DNA sequencers is making direct sequence analysis more rapid and widespread. The sequence data are used to construct phylogenetic trees, which represent the evolutionary relatedness of microbial species present in a given habitat. **Figure 6** shows

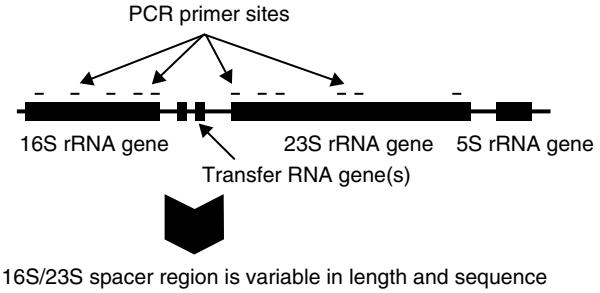


Figure 4 The intergenic transcribed spacer (ITS) region of bacteria contains none, one, two, or more transfer RNA genes between the 16S and 23S ribosomal DNA genes; ITS-PCR products are separated in polyacrylamide gels on the basis of their length.

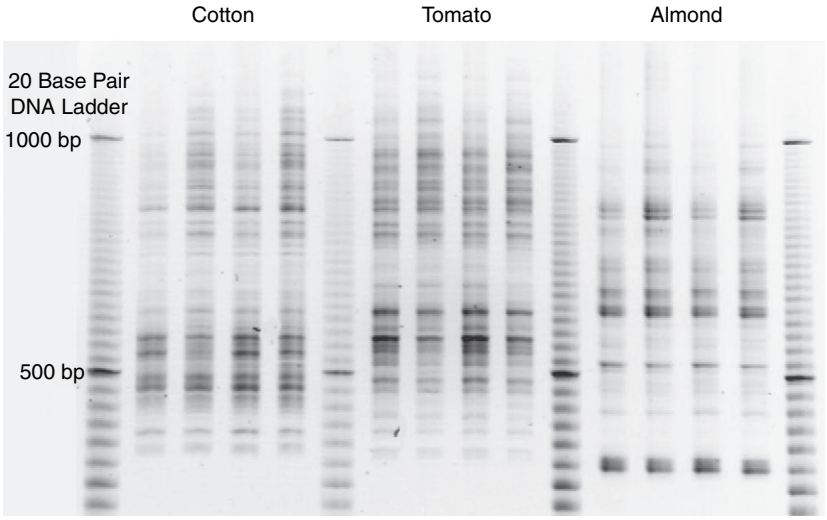


Figure 5 Bacterial ITS analysis of whole soil DNA extracts from the crops cotton, tomato, or almond. Each bacterial DNA fingerprint is represented by a pattern that may be categorized by band location and relative intensity and then compared to other patterns by clustering algorithms or other statistical methods.

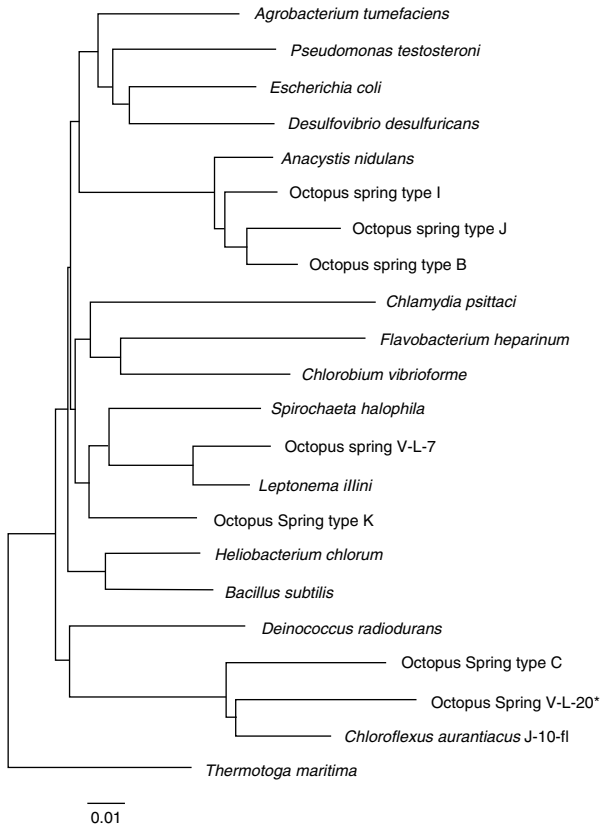


Figure 6 A distance matrix phylogenetic tree constructed to include cloned and sequenced members of a cyanobacterial mat community. The sequences of previously undescribed clones are analyzed and fit with their most closely related relatives in the cluster diagram (reproduced from Weller R, Bateson MM, Heimbuch BK, Kopczynski ED, and Ward DM (1992) Uncultivated cyanobacteria, Chloroflexus-like inhabitants, and Spirochete-likeinhabitants of a hot spring microbial mat. *Appl. Environ. Microbiol.* 58: 3964–3969, with permission from ASM.).

phylogenetic affiliations of 16S rDNA sequences of bacterial clones from a hot spring cyanobacterial mat community.

To date, it has been impossible to exhaustively sample a microbial community using this approach. For instance, in diversity studies of soil bacterial communities, if 100 clones are picked for analysis, each sequence may be unique, suggesting that species richness is far greater than 100. The clones are chosen randomly; thus, the fraction of clones represented by a common sequence can be used to estimate the relative predominance of that population in the community. In addition to PCR amplification bias, artifacts produced by selective cloning of some fragments over others must be considered.

PCR-Independent Methods

Solution Hybridization Microbial diversity of an environmental sample may be estimated using solution hybridization (DNA–DNA reannealing kinetics) approaches similar to those used to determine the genetic similarity of two bacterial strains. In this approach, double-stranded DNA is extracted directly from an environmental sample, denatured, and then allowed to reanneal (return to double-stranded form). Reannealing of DNA can be monitored using spectroscopic techniques. The rate at which the environmental DNA strands reanneal can be compared with the rate at which perfectly matched DNA strands reanneal, and the difference gives a measure of the number of different genomes present in the sample. Results of these assays should be interpreted with two potential artifacts in mind. First, contaminants in the samples (i.e., humic acids in soil DNA extracts) can affect DNA reannealing. Second, since the size of a genome affects reannealing kinetics, necessary assumptions about the "average" size of a genome may skew absolute estimates of the number of genomes present. The genetic similarity between two communities is best estimated by labeling extracted DNA from one community and measuring

the amount of label that hybridizes to DNA from the other community and then vice versa.

Membrane Hybridization Specific taxonomic groups can be detected and quantified by Southern blotting with labeled oligonucleotide probes to complementary ssrDNA sequences. With this method, the sample is fixed to a membrane, and then through a series of steps the probes are allowed to hybridize to complementary portions of microbial nucleic acids. The probes are either radiolabeled or enzyme linked with fluorescent dyes for quantification by densitometry or fluorimetry. The DNA on a single membrane may be stripped of one probe and then hybridized with others to provide estimates of relative abundance of certain taxa or other phylogenetic information. This technique is more rapid than cloning and sequencing, but the appropriate quantification standards must be used. Precise quantification of the number of individuals per species, however, is not possible by this method because the naturally occurring number of copies of the ssrDNA varies among microorganisms. The membrane hybridization technique can also be used with ribosomal RNA, in which case it is called Northern blotting. Northern blots provide quantitative estimates of active taxa because the rRNA concentration is higher in those members of a community that are physiologically active.

Fluorescent In Situ *Hybridization* FISH is a modification of oligonucleotide probing that allows spatial location of specific populations in their native habitat. With this technique, fluorescently labeled oligonucleotides hybridize to ssrD/RNA inside cells and are detected with an epifluorescent or laser confocal microscope. Microorganisms are first pretreated to create gaps in their membranes that permit the labeled oligonucleotides to enter the cells. Group- or species-specific oligonucleotides may be labeled with differently colored fluorescent molecules such that multiple groups of organisms can be observed in a microscopic field. Microscope images can be analyzed automatically to reduce operator errors in identifying and counting fluorescent cells. In soil systems, in which the majority of cells are presumably dormant, FISH results may lead to an underestimation of the microbial diversity due to detection limits. Because this technique allows visualization of individual cells, results are not confounded by variation in the number of ssrDNA gene copies among different microorganisms. This method is problematic in soil systems because soil organic matter fluorescence interferes with the probes' fluorescent signals and because probes may bind to soil constituents.

Oligonucleotide Arrays A promising new variation of hybridization analysis involves attaching arrays of oligonucleotides to a solid surface rather than attaching sample DNA to traditional blotting membranes. In this technique, multiple oligonucleotide sequences are immobilized on a glass or plastic slide, and the sample DNA is radioactively or fluorescently labeled. After labeled PCR fragments or rRNA hybridize to complementary oligonucleotides on the slide, the signal is quantified using laser technology. As many as 10,000 different sequences can be deposited on a small microscope slide. Due to variation in hybridization kinetics, the results are not strictly quantitative but can show trends in population dynamics. An advantage of this technology is that, given sufficient DNA or RNA, it can be used with nucleic acids extracted directly from the environment without an intermediate PCR step. Interpretation of the wealth of data that results from this method remains a formidable challenge. However, it is currently the only technique that can detect thousands of different populations rapidly, and it likely will have a large impact on microbial diversity studies. As more environmental isolates are characterized and sequenced, oligonucleotide arrays will become more useful for studying biodiversity.

Fatty Acid Analysis

In addition to DNA-based identification methods, microorganisms can be classified based on their cellular lipid composition. Lipids can be used to identify individual strains or to characterize whole communities. The most commonly analyzed lipids are the phospholipid fatty acids (PLFAs). PLFAs are the primary component of cell membranes and are present in all living organisms. Fatty acids are distinguished on the basis of chain length, number and location of unsaturations, and location of substituents (e.g., methyls, hydroxyls, and cylopropane rings).

In lipid analysis of individual strains, cells are grown according to a standard protocol and their cellular fatty acids are extracted. Chromatographic analysis of the fatty acids provides a fingerprint, which is compared to a database to identify species. In a modification of the method, PLFAs derived primarily from the cell membrane are analyzed instead of whole cell fatty acids. Analysis of PLFAs in environmental samples is based on an extraction procedure in which polar lipids are extracted and purified using silicic acid chromatography. The ester-linked fatty acids of bacteria, or ether-linked fatty alcohols of Archaea, are analyzed by capillary gas chromatography or gas chromatography with mass spectrometry (White *et al.*, 1979). **Figure 7** shows a correspondence analysis plot of PLFA abundance data for 100 sulfate-reducing bacterial cultures together with uncharacterized sulfate-reducing cultures enriched from sediments in Clear Lake, California. Bacterial strains (represented by points) which lie close together have similar fatty acid compositions.

PLFA analysis is also used to characterize entire microbial communities in environmental samples. There are three ways in which PLFAs can provide information about diversity: (i) The entire PLFA profile can be used as a fingerprint which reflects the composition of the total soil community; (ii) signature fatty acids can be used to detect specific subgroups within the community, e.g., sulfate reducing bacteria, methane oxidizing bacteria, fungi, and actinomycetes; and (iii) certain lipids are physiological indicators of environmental stresses, e.g., the ratio of saturated to unsaturated fatty acids, the ratio of *trans*- to *cis*-monoenoic unsaturated fatty acids, and the proportion of cyclopropyl fatty acids.

Measurement of Process or Function

Substrate Utilization Patterns

The measurement of substrate utilization patterns is commonly used to evaluate functional diversity of bacterial

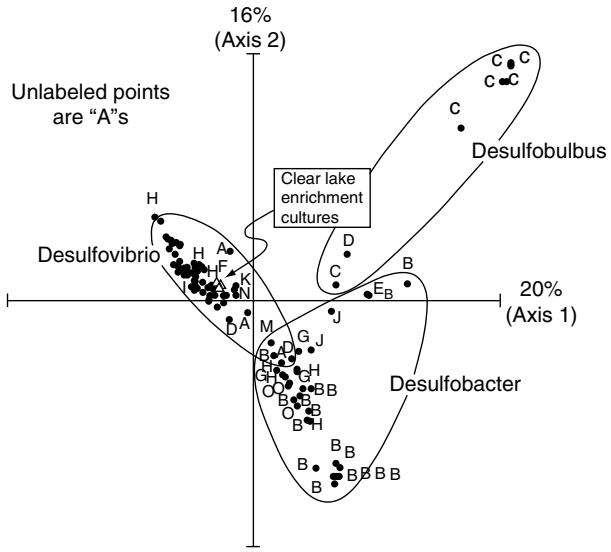


Figure 7 A diagram summarizing PLFA abundance data for pure cultures of sulfate-reducing bacteria and uncharacterized strains enriched from polluted sediments in Clear Lake, California. A, Desulfovibrio; B, Desulfobacter; C, Desulfobulbus; D, Desulfotomaculum; E, Desulfobacterium; F, Desulfomicrobium; G, Desulfuromonas; H, other (see text); I, Desulfomonas; J, Desulfotobacterium; K, Desulfococcus; M, Desulfobotulus; N, Desulfosarcina; O, Desulfomonile.

populations. The commercially available Biolog system, which contains 95 carbon and energy sources on one microtiter plate, is widely used because of its convenience. The plate is inoculated directly with a sample of a microbial community (e.g., a dilution of a soil or sediment suspension) and then incubated under aerobic conditions. Substrate utilization leads to the development of a blue color originating from a redox-sensitive dye. Comparison of well color development patterns among different microbial communities coupled with multivariate statistical analysis may yield information on the functional relatedness of the bacterial populations in the sample.

A major drawback of this method is reliance on the activity of culturable organisms; thus, organisms unable to grow under the incubation conditions are not represented. There is also debate regarding the environmental significance of the carbon sources used in the microtiter plates. These carbon substrates were originally chosen for their ability to distinguish pathogenic bacterial isolates from one another, and therefore bear little relation to substrates encountered by microorganisms in their natural environments. In response to this criticism, plates that contain substrates more common in aquatic and terrestrial environments have been developed and used in some recent applications.

Enzyme Assays

A variety of enzymes have evolved to carry out the diverse metabolic processes in which microorganisms are engaged. Six major classes of enzymes have been defined that, in turn, are divided into subclasses on the basis of the type of reactions they catalyze. **Table 2** provides an overview of common enzymes found in soil. Enzymes found in the environment may be associated with cellular cytoplasm or may be attached to the lipid membranes or surfaces of microbial cells. Other enzymes may be present outside of cells, either in solution or attached to particles in the environment. Enzymes are usually measured based on the type and amount of activity they catalyze under controlled conditions in response to added substrates. "Potential" rather than actual activity is commonly measured.

Some enzymes are present in all organisms (protease and hydrogenase) and give general information about the intensity of biological activity in a sample. Other enzymes catalyze more specific reactions and provide information about the diversity of processes that potentially can be carried out by microbial populations within a given environment. These include enzymes associated with decomposition of polysaccharides (e.g., cellulase, chitinase, amidase, and xylanase), the transformation of phosphorus (e.g., specific kinds of phosphatase) and sulfur (e.g., arylsulfatase), lipid metabolism (e.g., lipase), lignin degradation, and the metabolism of specific environmental pollutants.

Computational Needs

Within the past decade, many large data sets describing microbial communities have been generated using the new biochemical approaches described previously. Much of the information potentially available in these data sets is overlooked because biologists often do not employ the statistical approaches necessary for handling such large data sets. Interesting information may be buried in improperly or unanalyzed data; therefore, serious attention must be given to this problem. One challenge is reducing these large data sets to smaller sets of variables or components. In this way, classification tools can be used and models can be constructed to predict the response of microbial processes to changes brought on by human activity or environmental changes.

Standard multivariate techniques are commonly used to explore the "natural" relationships within a data set (principal component analysis and correspondence analysis) and to determine the correlation of environmental variables to specific components (constrained ordination techniques such as canonical correspondence analysis). In addition, regularized discriminant analysis (RDA) and partial least squares (PLS) are used to test the significance of *a priori* knowledge about the data sets (e.g., the significance of environmental variables or sample category). RDA uses a regularized covariance matrix estimate for the conventional statistical discriminant analysis methods. PLS is a multivariate calibration tool and can also be applied as a modeling method to solve pattern classification problems. PLS finds a set of latent variables in the measurement variable space (e.g., lipids or DNA banding patterns or sequences) that have a maximum covariance with the dependent variable space. Variable selection can then be used to choose which variables are most important for modeling. Artificial neural net analysis can be used to explore nonlinear relationships within the data sets, (e.g., relating community composition to process rates). Using these approaches, it may be possible to address such questions as which traits are shared by all members of certain classes

Table 2 Some Enzymes Extracted from Soils and the Reaction They Catalyze

Recommended name	*Reaction*	*Substrate*
Oxyreductases		
Catalase	$2H_2O_2 \rightarrow O_2 + 2H_2O$	H_2O_2
Peroxidase	Donor $+ H_2O_2 \rightarrow$ oxidized donor $+ 2H_2O$	H_2O_2, pyrogallol, chloroanilines, *o*-dianisidine
Monophenol monooxygenase (polyphenol oxidase)	Tyrosine + dihydroxyphenylalnine $+ O_2 \rightarrow$ dihydroxyphenylalanine + dioxophenylalanine $+ H_2O$	*d*-Catechol, *p*-quinol, *p*-cresol, 3,4-dihydroxyphenylalanine, *p*-phenylenediamine
Hydrolases		
Carboxylesterase	A carboxylic ester $+ H_2O \rightarrow$ an alcohol + a carboxylic acid anion	Malathion, hydroxymethylcoumarin butyrate
Arylesterase	A phenyl acetate $+ H_2O \rightarrow$ a phenol + acetate	Phenyl acetate, phenyl butyrate, naphthyl acetate
Alkaline phosphatase	Orthophosphoric monoester $+ H_2O \rightarrow$ an alcohol + orthophosphate	*p*-Nitrophenyl phosphate
Acid phosphatase	Orthophosphoric monoester $+ H_2O \rightarrow$ an alcohol + orthophosphate	*p*-Nitrophenyl phosphate
Arylsulfatase	A phenol sulfate $+ H_2O \rightarrow$ a phenol + sulfate	*p*-Nitrophenyl sulfate
Cellulase	Endohydrolysis of 1,4-β-glucosidic linkages in cellulose, lichenin, and cereal β-glucose	Cellulose, carboxymethylcellulose
β-Glucosidase	Hydrolysis of terminal, nonreducing β-D-glucose residues with release of β-D-glucose	*p*-Nitrophenyl-β-D-glucoside, cellobiose, *p*-nitrophenyl-β-D-glucopyranoside
β-Fructofuranosidase (Invertase)	Hydrolysis of terminal nonreducing β-D-fructofuranoside residues in β-fructofuranosides	Sucrose
Proteinases	Hydrolysis of proteins to peptides and amino acids	Casein, gelatin, albumin
Urease	Urea $+ H_2O \rightarrow CO_2 + 2NH_3$	Urea
Lyases		
Aromatic-L-amino-acid decarboxylase	L-Tryptophan $\rightarrow$ tryptamine $+ CO_2$	*dl*-3,4-Di-hydroxyphenylalanine, *dl*-tyrosine, *dl*-tryptophan, *dl*-phenylalanine, tryptophan

Source: Reproduced with permission from Paul EA and Clark FE (1996) *Soil Microbiology and Biochemistry*, 2nd ed. Academic Press, New York.

(e.g., communities on a particular land use and vegetation type) and which components of a community are linked to specific processes such as pollutant degradation and nitrogen cycling.

Future Directions

Great care must be taken to not overstate the reliability of microbial diversity measurements currently available. None of the diversity measurements currently available have the ability to (i) exhaustively sample a community, (ii) quantify population densities, and (iii) resolve individual populations. Ecologists who study macroorganisms have the ability to satisfy all three of these criteria and therefore it is difficult to compare the diversity of large organisms with that of microorganisms. It is likely that a large subset of microorganisms exist at low population densities and are dormant. When this fraction of the community encounters environmental conditions suitable for growth, the dormant populations expand, and only then do we have the ability to measure them. Thus, current techniques tend to be biased toward dominant members of a microbial community. Despite this limitation, ecological studies of microorganisms and this limited perspective have already revealed overwhelming biodiversity.

Once the tools are developed to characterize and analyze data describing microbial communities, numerous questions can be considered. Questions of interest may include the following:

1. What environmental gradients (e.g., vegetation, time, soil texture, and topography) are most strongly correlated with the composition of specific communities or patterns of communities in general?
2. What other (e.g., nongradient) factors also regulate community diversity?
3. Do communities undergo succession within a year? How much do communities vary year to year? Are there seasonal patterns common to all communities across biomes?
4. Are there associations between microbial communities (e.g., facilitation, competition, and interaction)?
5. Does human disturbance alter communities to a greater degree and in different ways than do seasonal and vegetation changes, and are there general patterns that are indicative of disturbance?

See also: Measurement and Analysis of Biodiversity. Microbial Biodiversity. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms

References

Akkermans ADL, van Elsas JD, and De Bruijn FJ (eds.) (1995) *Molecular Microbial Ecology Manual*. Dordrecht: Kluwer.

Burlage RS, Atlas R, Stahl D, Geesey G, and Sayler G (eds.) (1998) *Techniques in Microbial Ecology*. New York: Oxford Univ. Press.

Head IM, Saunders JR, and Pickup RW (1998) Microbial evolution, diversity, and ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. *Microbial Ecol* 35: 1.

Old RW and Primrose SB (1994) *Principles of Gene Manipulation*, 5th ed. Oxford: Blackwell.

Paul EA and Clark FE (1996) *Soil Microbiology and Biochemistry*, 2nd ed. New York: Academic Press.

Pickup RW and Saunders JR (eds.) (1996) *Molecular Approaches to Environmental Microbiology.* New York: Horwood.

Van Elsas JD, Trevors JT, and Wellington EMH (eds.) (1997) *Modern Soil Microbiology.* New York: Dekker.

Weller R, Bateson MM, Heimbuch BK, Kopczynski ED, and Ward DM (1992) Uncultivated cyanobacteria, *Chloroflexus*-like inhabitants, and Spirochete-like inhabitants of a hot spring microbial mat. *Appl. Environ. Microbiol.* 58: 3964–3969.

White DC, Davis WM, Nickels JS, King JD, and Bobbie RJ (1979) Determination of the sedimentary microbial biomass by extractable lipid phosphate. *Oecologia* 40: 51.

Microbial Biogeography

China A Hanson, University of California, Irvine, CA, USA
M Claire Horner-Devine, University of Washington, Seattle, WA, USA
Jennifer BH Martiny, University of California, Irvine, CA, USA
Jed A Fuhrman, University of Southern California, Los Angeles, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Community composition It is the identity and relative abundances of all taxa in a community.
Cosmopolitan This refers to being present everywhere, globally distributed.
Dispersal limitation It is the inability to move to and/or colonize/establish at some location.
Drift It is the influence of stochastic demographic variability (such as birth, death, and migration rates) on biotic composition in an area.
Historical processes These are the evolutionary and ecological processes that occur in the recent or distant past including immigration, emigration, speciation, extinction, drift, and/or past selection by the environment.
Microorganism or microbe These are bacteria, archaea, protists, fungi, and small metazoans, typically smaller than 0.5 mm in equivalent diameter and with mass $< 10^{-5}$ g; may also include viruses.
Spatial autocorrelation It is the tendency of geographically close observations to be more similar than geographically distant observations.
Stochastic It means a random effect due to chance events.
Taxon A group into which related organisms are classified based on a designated relatedness cutoff, often using sequence similarity of a marker gene; for example, species-equivalent taxa may be defined as at least 99% similar in their 16S rRNA gene sequence.
Taxonomic resolution The level of genetic variation of the taxa considered.

Introduction

Biogeography is the study of the distribution of biodiversity over space and time. It aims to reveal why organisms live where they do, and at what abundance. The study of biogeography offers insights into the ecological and evolutionary mechanisms that generate and maintain diversity, such as speciation, extinction, dispersal, and species interactions. The results of biogeographic studies also have direct application in the modeling of ecosystem responses as they differ between habitats and respond to global change. Since the eighteenth century, biologists have investigated the geographic distribution of plant and animal diversity, leading to important advances in ecological theory. However, until recently, microorganisms have rarely been considered in this regard.

The Baas-Becking Hypothesis

Although few early microbiologists performed biogeographic studies, a central theme in microbial biogeography today is the so-called "Baas-Becking hypothesis," dating back to almost a century. The hypothesis suggests that for microbial taxa, "everything is everywhere – the environment selects," (Baas-Becking, 1934). The claim "everything is everywhere" implies that microorganisms are so abundant and have such enormous dispersal capabilities that they can rapidly erase any geographic patterns generated by past evolutionary and ecological events. The claim that "the environment selects" implies that different contemporary environments allow for selective growth of distinctive microbial assemblages. If true, and dispersal and selection are effectively instantaneous, the hypothesis implies two expectations: (1) microbial communities will vary among habitats but similar habitats in different locations will harbor similar active microbial communities and (2) the lack of dispersal limitation means that "seed" populations of every kind of microbe can be found everywhere, albeit possibly at very low abundance.

In the sections that follow, some of the recently observed microbial distribution patterns are described. A framework is reviewed for testing the Baas-Becking hypothesis as well as three other alternative hypotheses. It is seen that the current literature generally supports the second expectation of the Baas-Becking hypothesis; the environment clearly has a strong effect on microbial composition, as it does for the composition of larger organisms. However, there is also evidence that the first expectation does not always hold; in other words, that some microorganisms may be dispersal limited.

The Molecular Genetic Revolution

For many decades it was not feasible to evaluate the Baas-Becking hypothesis because suitable methodologies for identifying the majority of microbial taxa in an environmental sample were not available. Until the 1980s classical cultivation techniques were necessary for proper identification of bacteria, archaea, and the smallest protists, and such identification, applied to whole communities, would have been onerous. However, some of the first modern studies to describe the large-scale biogeography of microbes did so by employing morphological identification techniques to group eukaryotic microbes (Protista) to the genus or near species levels. These studies found that virtually all species were detected in a range of widespread and varied habitats, suggesting that such microbes were indeed globally distributed and therefore

corroborating the first expectation of the Baas-Becking dictum (Finlay, 2002; Fenchel and Finlay, 2004). In recent years these claims have been challenged (Foissner, 2006), however, with one argument being that morphologically defined microbial species are "cryptic" in that they are not adequate representations of genetically differentiated species.

The application of newly developed molecular methodologies has now allowed microbial biogeography surveys to characterize microbial taxa genetically, thereby avoiding the caveats associated with cultivation-dependent or morphological-based techniques. Such techniques obtain DNA sequences directly from environmental samples without the need for cultivation. Sequences are then classified into microbial taxa, often referred to as operational taxonomic units (OTUs), which are defined by the nucleotide sequence similarity of one or more genomic regions, usually a marker gene such as 16S rRNA. Sequence similarity cutoffs are often arbitrarily chosen to be coarsely analogous to species or genus in larger organisms. However, taxa can be defined at any level of taxonomic resolution, from unique sequences to OTUs that lump together genetically and phenotypically diverse microbes. Thus, as compared to morphological techniques, sequence-based identification generally allows a greater amount of microbial diversity to be distinguished and taxa to be defined with greater resolution.

Indeed, some of the first studies utilizing culture-independent molecular methods revealed that past culture-based studies missed the vast majority of the microbial diversity present in the environment (e.g., Ward *et al.*, 1990). Additionally because these techniques continue to evolve and improve in coverage and depth of sampling, microbial diversity has been sampled much more deeply and widely than ever before (e.g., Sogin *et al.*, 2006). Microbial biogeography has benefited tremendously from these advances, and in fact most of the work presented in this article relies on this approach.

General Patterns in the Distribution of Microbes

Micro- and macroorganisms are often involved in intimate and specific associations that affect each other's geographic distributions. Studies have recognized for many years that host-associated and pathogenic microorganisms exhibit patterns of genetic, morphological, and functional differentiation related to the distribution of their hosts (Hedlund and Staley, 2003). Thus, there is little debate that most host-associated microbes display biogeographic patterns that parallel those of animals and plants.

Debate remains about the extent to which nonhost-associated microbes display biogeographic patterns. Studies from a diversity of habitats, including soils, sediments, hot springs, marine, and freshwaters, show clearly that community composition differs across space and time, as reviewed extensively in recent years (e.g., Foissner, 2006; Martiny *et al.*, 2006; Lindström and Langenheder, 2011). Not only do the distributions of microorganisms appear to be nonrandom, but these distributions are also often similar to those observed for animals and plants.

Certainly, not every study thus far has found evidence for significant biogeographic patterns among the microbial taxon of interest (e.g., Fenchel and Finlay, 2004). However, the existence of microbial biogeography does not depend on every microbial taxon being nonrandomly distributed. The presence of cosmopolitan microorganisms – the "pigeons" of the microbial world – does not exclude the possibility that other microbial taxa have strikingly uneven geographic distributions. Similarly, microbial communities that contain a large proportion of evenly distributed cosmopolitan taxa can still exhibit biogeographic patterns driven by a minority fraction of taxa that are not evenly abundant everywhere.

The evidence for microbial biogeography is reviewed below by describing some general biogeographic patterns that have been observed for free-living microbes.

Endemism

The existence of endemic microbial taxa would, in theory, be the simplest demonstration that microorganisms are nonrandomly distributed. Endemics are, by definition, restricted to a particular geographic range and therefore are not evenly distributed across the Earth. In practice, however, characterizing the geographic ranges of microbes is difficult. First, because a vast majority of bacterial taxa in a sample are rare (Sogin *et al.*, 2006), many taxa that are present will not be detected. Thus, proving the absence of a microbial taxon within highly diverse natural assemblages is nearly impossible with currently available techniques. Second, it is equally difficult to determine whether observed taxa are actually active in a given sample, and therefore perceiving the location as a suitable habitat. Indeed, many microbes are capable of dormancy stages, and recent evidence suggests that many taxa detected in a location are dormant (Lennon and Jones, 2010), perhaps having arrived there as transient migrants. Thus, without employing a strategy to identify active versus dormant taxa, the actual active habitat range of some microbes might not be distinguishable from a cosmopolitan distribution.

Significant genetic divergence among taxa in different locations is perhaps the strongest evidence for microbial endemism. Many studies employ statistics such as Sewall Wright's F_{ST} or similar metrics, which compare the genetic diversity of a microbial taxon within locations relative to that among all locations combined. Some studies further find that particular microbial clades are restricted to different regions of the world (as demonstrated in **Figure 1**). For instance, Whitaker *et al.* (2003) showed clear genetic differentation among hyperthermophilic archaea from geothermal hot springs on different continents. Significant genetic divergence among geographic locations is not just a demonstration that the genetic diversity of microbes varies in space, but it is also evidence that the sampled microbial communities are not completely mixed by dispersal of individuals between them – that is, the communities are isolated from each other due to dispersal limitation.

Latitudinal Gradients

Since the eighteenth century, it has been widely recognized that there is a broad tendency for animals and plants to

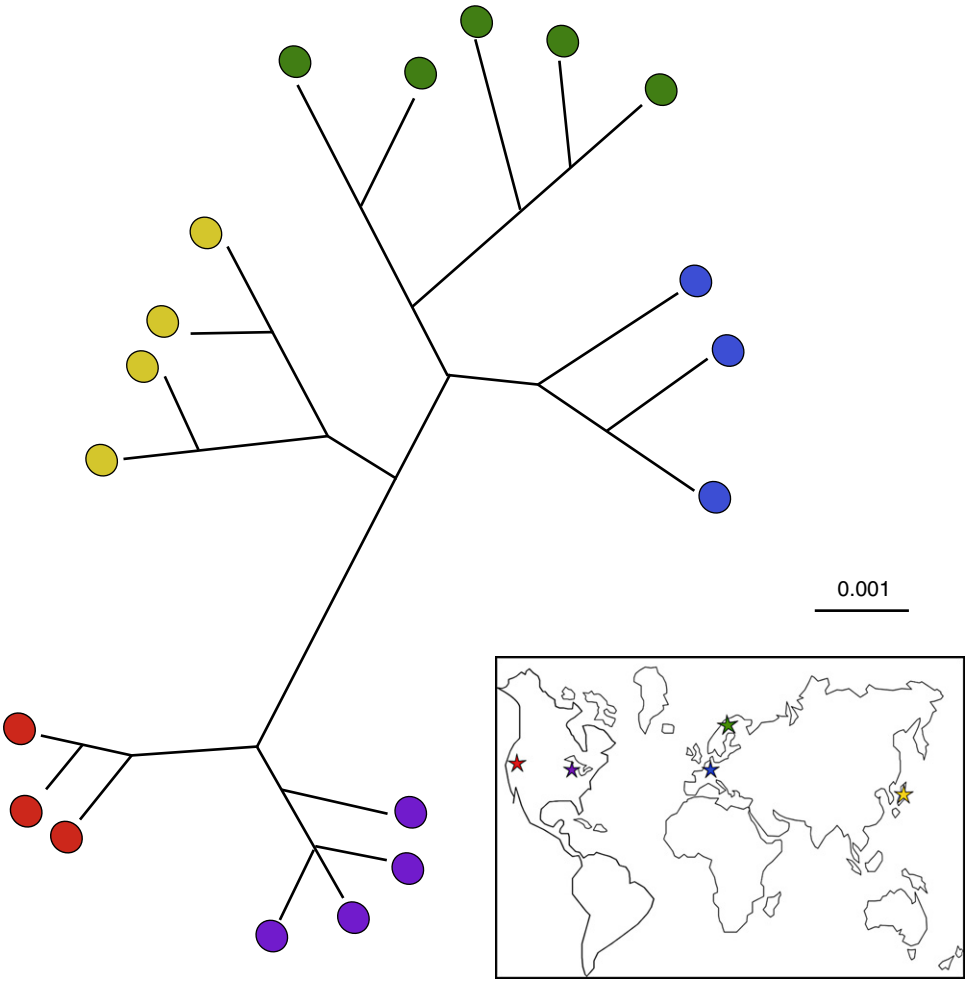


Figure 1 Phylogenetic tree depicting the genetic relatedness of hypothetical isolates (indicated by colored circles) collected from similar habitats around the world (inset). Isolates from geographically nearer locations are more closely related than they are to isolates from more distant locations. Similar patterns have been observed with hot spring Archaea, for example, with isolates from locations in North America being more similar to each other than to those from Asia. Inset: Map showing sampled locations. Different colors correspond to the geographic origin of isolates. Scale bar approximates 1 nucleotide substitution per 1000 bases.

exhibit a gradient of higher species richness (i.e., α-diversity) in tropical regions and lower diversity in polar ones. There are numerous nonexclusive explanations of such patterns, but prominent among them is the suggestion that this relationship is driven by latitudinal changes in temperature and/or productivity (Rosenzweig, 1995; Hillebrand, 2004). Some meta-analyses have indicated that the gradient tends to get weaker as animals get smaller, which might be extrapolated to suggest little or no latitudinal richness gradient for bacteria-sized organisms (Hillebrand, 2004). Recent studies of bacterial diversity have begun to look for such patterns. In a study of 98 soil samples from North and South Americas, Fierer and Jackson (2006) reported no relationship between bacterial richness and latitude or temperature, instead finding that soil type, and especially pH (which varied greatly among samples) was an overwhelmingly important factor, with lower diversity in more acidic soils. In contrast, studies of marine planktonic bacteria, living in the seawater environment that is much more chemically and physically uniform than soils, found evidence of a latitudinal gradient in bacterial richness (Pommier *et al.*, 2007; Fuhrman *et al.*, 2008).

Elevational and Depth Gradients

Scientists have long-recognized elevational gradients in plant and animal community composition. Over relatively short geographic distances, dramatic changes in climate with elevation are often associated with stark biotic turnover in a wide range of taxonomic groups such as birds, trees, and insects. In particular, species richness tends to exhibit either a hump-shaped or monotonically decreasing relationship with elevation. Like with latitudinal gradients, a number of nonexclusive hypotheses have been proposed, but the cause of elevational patterns remains poorly understood. Recent work in the Rocky Mountains showed that bacterial richness decreased monotonically from low to high elevations; while plants from the same system exhibited a unimodal pattern of richness, with richness peaking at midelevations (Bryant *et al.*,

2008). In contrast, bacterial richness in organic soil, mineral soil, and on leaf surfaces along an elevational gradient in the eastern Andes of Peru showed no significant relationship between richness and elevation (Fierer *et al.*, 2011), while bat, bird, and tree richness decreased along the same gradient. Finally, bacterial richness increased monotonically with elevation in a stony stream in China, while diatom richness decreased and macroinvertebrate richness showed a clear unimodal pattern with elevation in the same system (Shen *et al.*, 2011). Overall, these studies suggest that bacteria also vary in composition and richness along elevational gradients, but so far, no consistent patterns have emerged across studies and ecosystems. A similar analysis applies to depth in aquatic and marine systems, which are often physically, chemically, and biologically stratified (layered with depth). Indeed, microbial communities have been known for decades to vary with depth and oceanographic conditions, and even have the potential to be tracers for vertical water movement (e.g., Kriss, 1960). Correspondingly, several studies have demonstrated clear variations of microbial composition with depth (e.g., reviewed by Martiny *et al.*, 2006; Lindström and Langenheder, 2011).

Distance–Decay Relationships

Generally, the similarity between community composition decreases as the distance between two locations increases (Nekola and White, 1999). This so-called distance–decay relationship (**Figure 2**) indicates not only that communities are different in different locations, but also that this variation is spatially autocorrelated. Multiple factors may act alone or in combination to produce this pattern. First, environmental variables tend to be spatially autocorrelated (as in a cline or gradient) and organisms with differing environmental or niche preferences are selected from the available pool of taxa. Second, restricted dispersal of organisms over space can result in geographic isolation or clustering of taxa.

Distance–decay curves have been consistently observed for microorganisms at a range of taxonomic and spatial scales (e.g., Cho and Tiedje, 2000; Green *et al.*, 2004; Horner-Devine *et al.*, 2004; Martiny *et al.*, 2011). For instance, Whitaker *et al.* (2003) showed that not only did archaea diverge among the hot spring locations, but also these differences were spatially autocorrelated. The particular processes underlying these patterns can also vary with taxonomic or spatial scale, as has been shown for bacteria from salt-marsh sediments from around the world (Martiny *et al.*, 2011). Regardless of how the pattern is created, the decline of similarity in communities across space is considered to be powerful evidence of a nonrandom geographic distribution and has been repeatedly observed for microbes in an exhaustive range of habitats and spatial scales.

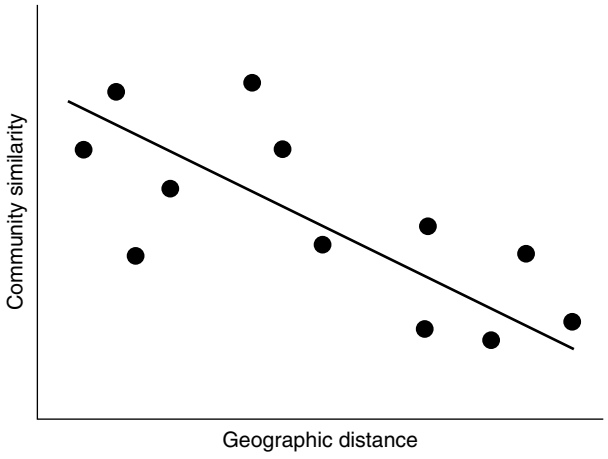


Figure 2 An example of a distance–decay curve. Each point represents a single pairwise comparison of the total number of locations sampled. The solid line represents a negative relationship between community similarity and geographic distance across all points. Using the same hypothetical microbial isolates from **Figure 1**, the pairwise comparisons from North America would be on the upper left side of the figure (more similar to each other and closer together spatially), and comparisons between North America and Europe or Asia would be in the lower right side.

Taxa–Area Relationships

Closely related to distance–decay patterns are species– (or taxa–) area relationships (**Figure 3**). Here species are used to refer to any microbial taxon defined by relatedness. A positive relationship between the number of species observed and the size of the area sampled has been observed repeatedly in plants and animals in a wide range of ecosystems for over a hundred years, often ascribed in part to the fact that as area increases, more types of habitats and microenvironments are included (Rosenzweig, 1995). This empirical species–area relationship usually fits a power-law function, $S = cA^z$, where S is the number of species, A is the area sampled, and c is the intercept in log–log space. The species–area exponent, z, is a measure of the rate of change of the slope with increasing area, that is, the rate of turnover of species across space.

It is rarely, if ever, possible to exhaustively sample the microbial diversity in even a small area, let alone in areas of

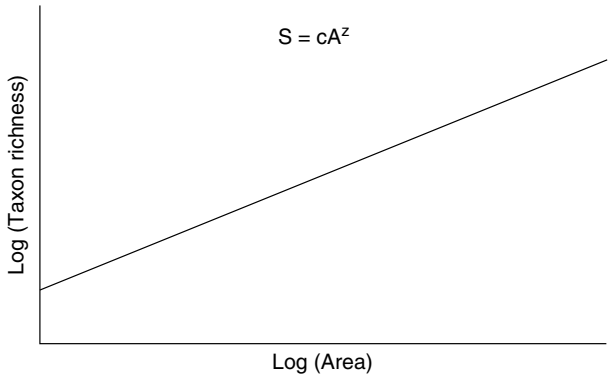


Figure 3 A taxa–area relationship, showing a linear increase in the number of taxa observed with increasing area sampled, on a log–log scale. S is the taxon richness, A is area, and c and z are constants that differ between taxa and habitats, with c the richness at the smallest scales and z the slope of the line reflecting the rate of increasing richness with increasing area or time. Note that on a log–log scale, the taxa–area relationship is equivalent to $\log(S) = \log(c) + z\log(A)$.

increasing size. However, researchers have used distance–decay curves to determine the taxa–area exponent, *z*, and have shown that taxa–area relationships also hold in microbial communities (e.g., Green *et al.*, 2004; Horner-Devine *et al.*, 2004). Currently, much microbial biogeographic research focuses on understanding why "*z*," or the rate of microbial turnover, varies among different taxa, ecosystems, and spatial scales. Such an approach offers an insight into the turnover of community composition across space and how this turnover may vary with the spatial scales sampled. In addition, it is important to consider the spatial scales on which the underlying processes operate and how the relevant scales compare across taxa.

Seasonality and Taxa–Time Relationships

Although most microbial biogeography studies focus on spatial variation, community composition and diversity can also vary nonrandomly over time. For instance, many studies observe very strong temporal and seasonal dynamics in bacterial community composition. Fuhrman *et al.* (2006) showed robust, repeatable temporal patterns in the community composition for bacterioplankton off the coast of Southern California. Similarly, near-surface ocean bacterioplankton in the English Channel exhibited strong repeatable seasonal patterns with peaks in richness during the winter (Gilbert *et al.*, 2009).

As for area, the number of species observed also increases with the time span over which the community is observed (i.e., a positive species–time relationship) due to immigration and also inclusion of more varied conditions. Although less well studied than the taxa–area relationship, organisms from a range of taxonomic groups and ecosystems types exhibit significant and similar taxa–time relationships within and across years (White *et al.*, 2006). Recent work in the Mid-Atlantic Bight showed that time rather than space or measured environmental variables was the primary driver of microbial community composition (Nelson *et al.*, 2008). Bacterial taxa–time relationships have also been observed in terrestrial environments (e.g., Redford and Fierer, 2009).

Interpreting Microbial Biogeographic Patterns

The patterns discussed above (*see* General Patterns in Microbial Distribution) indicate that not all microbial taxa are randomly distributed and show that at least some, if not most, microbes exhibit biogeographic patterns. Indeed, recent advances have been made to describe the existence of such nonrandom patterns in a wide variety of microbial taxa from a broad range of habitat types, as has been reviewed previously (Foissner, 2006; Martiny *et al.*, 2006; Lindström and Langenheder, 2011). However, it is considerably more challenging to interpret these patterns in terms of the ecological and evolutionary processes that may generate them. A community assembly perspective that divides these processes into the effects of current environmental factors and the effects of historical processes is a common framework for evaluating the processes responsible for generating spatial biodiversity patterns in macroorganisms (Ricklefs, 2007) and has been recently applied to microbes (Martiny *et al.*, 2006). Here "history" includes events from the very recent to the geological past that might have influenced present-day distributions, such as past environmental conditions or dispersal limitation. This framework draws many similarities to the metacommunity theory of community assembly, which has also been recently applied toward interpreting observed patterns in the biogeography of microbes (Lindström and Langenheder, 2011).

Following from traditional biogeography, one can consider a framework comparing the importance of the contemporary environment versus the legacies of historical processes for microbial biogeographic patterns. The contemporary environment includes abiotic variables (e.g., temperature, pH, and nutrient availability) as well as biotic factors such as the abundance and composition of other organisms. Historical events include anything from the effects of the past environment on the community, as well as evolutionary or ecological drift and speciation and extinction, which would cause differences in composition among localities. To observe evidence of historical processes, however, there must be dispersal limitation of some microbial taxa in the community, otherwise the effects would be long-since homogenized, and microbial composition would only depend on the current environment. This dispersal limitation allows for the effects of events in the recent or distance past to have a current signature on the microbial taxa present in a location.

Under this framework, one can consider four alternative hypotheses. The null hypothesis (Hypothesis 1 in **Figure 4**) is that microorganisms are not affected by the current environment or historical events. In this case, microbial taxa would be randomly distributed over space. Because there is no selection in the community by any contemporary environmental factor, there is no relationship between community similarity and environmental similarity (dotted line in **Figure 4**). Similarly, because there is no dispersal limitation of the microbial taxa, there are no lingering effects of historical events. Thus, community composition would not be correlated with geographic distance (solid line).

Hypothesis 2 is that microbial biogeography reflects the influence of the current environment, but no historical effects. This is the Baas-Becking hypothesis discussed above. None of the taxa is dispersal limited, thus historical events do not influence present-day assemblages. As a result, community similarity is correlated with environmental similarity, but once this relationship is accounted for, there is no correlation between community similarity and geographic distance. A key point here is that the effect of the current environment must be first statistically removed before testing for a remaining correlation with geographic distance, as environmental variables are usually spatially autocorrelated.

A third alternative (Hypothesis 3) is that all spatial variation in microbial composition is due to dispersal limitation and resulting historical effects. Finally, Hypothesis 4 proposes that, like macroorganisms, microorganisms (or at least some taxa within a community) are both dispersal limited and under current environmental selection. As a result, the distance–decay curve can be partitioned into a correlation

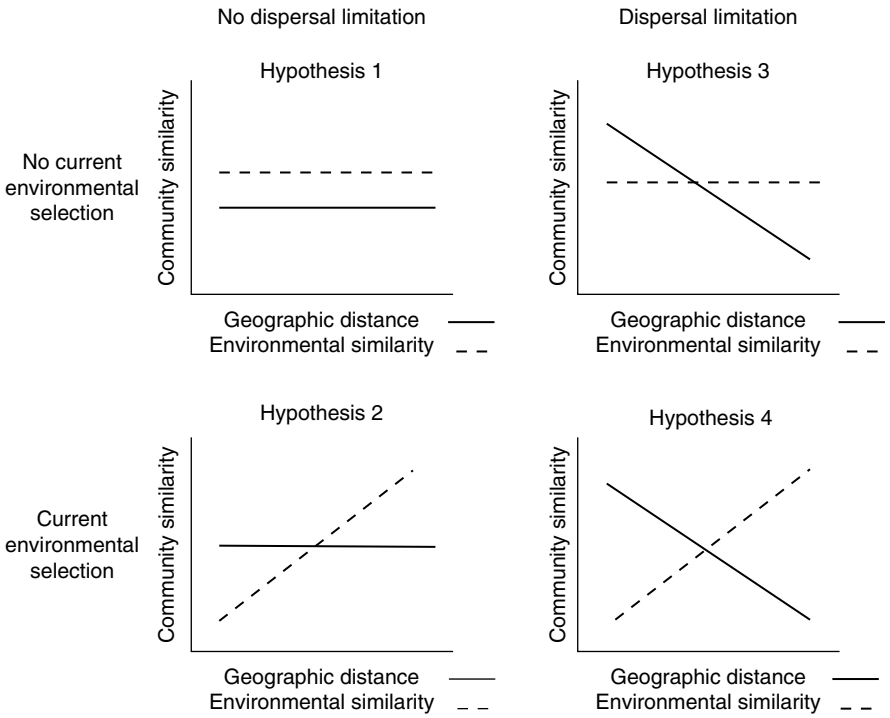


Figure 4 The relationship between community similarity and geographic distance (solid line), or environmental similarity (dotted line) for four hypotheses of microbial biogeography. As described in the text, the relationships depend on the degree of dispersal and the effectiveness of environmental selection.

with the environment as well as a remaining effect of geographic distance.

Evidence for the Importance of Contemporary Environmental Factors

To assess the importance of selection by the current environment on microbial composition (hypotheses 2 and 4; **Figure 4**), biogeography studies examine the correlation between community composition with measured environmental variables. In fact, almost every microbial biogeography study observes some such relationship between similarity in community composition and similarity in environmental characteristics of those samples. For example, numerous studies have linked spatial variation in microbial communities to variation in pH or salinity (e.g., Fierer and Jackson, 2006). Other studies have demonstrated that broadly defined habitat types, which are a proxy measure for a range of associated environmental variables, harbor significantly different microbial communities (e.g., Nemergut *et al.*, 2011). Additionally, seasonal variation in microbial communities tends to be significantly correlated with variation in environmental variables (Fuhrman *et al.*, 2006; Gilbert *et al.*, 2009). Together these studies provide support for the importance of environmental variation in explaining why microbes are nonrandomly distributed in time and space.

Evidence for the Importance of Historical Factors

In contrast, variation in microbial community composition in space may arise because not all taxa can disperse to or establish in all locations equally (hypotheses 2 and 4, **Figure 4**). Assessing the importance of historical factors in shaping biogeographic patterns is not straightforward, however. Because the movement and establishment of microbes is difficult to observe, the degree of dispersal limitation for microbes is often interpreted from the observed correlation between microbial community composition and geographic distance, after removing the influence of contemporary environmental factors. Since environmental variation is often spatially structured, as in a cline or gradient, controlling for this variation, either experimentally or statistically, is key to observing the effect of historical factors.

When microbial biogeography studies use this approach, some find a correlation between microbial composition and geographic distance after removing the effect of the contemporary environment (e.g., Cho and Tiedje, 2000; Whitaker *et al.*, 2003; Martiny *et al.*, 2011). Such studies provide evidence that historical factors, manifested by some dispersal limitation, can indeed be responsible for creating biogeographic patterns in microbes. Moreover, it suggests that not every microbial taxon is cosmopolitan; in other words, not every microbe is everywhere.

At the same time, many other studies find no evidence for dispersal limitation in microorganisms (e.g., Finlay, 2002; Fierer and Jackson, 2006; Van der Gucht *et al.*, 2007; Cermeño and Falkowski, 2009; Zinger *et al.*, 2011). Thus, a current question in the field is to uncover when and where dispersal limitation in microorganisms is more likely. In addition, if there is dispersal limitation, it remains to be investigated what historical processes are creating the signal (e.g., drift, speciation, extinction, or past environmental legacies).

Caveats to Interpreting Biogeographic Patterns

A balance between both contemporary environmental factors and historical processes likely shapes natural microbial communities. However, unraveling the importance of environmental selection, dispersal, and historical factors from observations of biogeographic patterns is not straightforward. First, these processes are likely to be highly intertwined over multiple scales of time and space (Rosenzweig, 1995; Ricklefs, 2007), and pinpointing the singular influence of any particular process within natural communities will be difficult. Second, multiple processes can result in identical patterns, yet many of the underlying processes cannot be directly measured or observed. Therefore, the observation of the existence of a particular biogeographic pattern does not prove the importance of any process.

Additionally, identifying underlying processes from biogeographic patterns is particularly difficult for microbes, as there is a limited understanding of microbial diversity. The physiological requirements and ecological interactions of the majority of microbial taxa in nature are unknown, so it is quite likely that many important selective factors of microbial habitats go unmeasured. For example, species interactions, such as predation, competition, facilitation, or mutualism, are generally under-considered in biogeographic studies of free-living microbes. Nonetheless, these interactions can alter taxonomic occurrences in space and may be strong forces in determining the habitable ranges of some microbes. For instance, Zinger *et al.* (2011) found that the composition of bacteria, archaea, and fungi in alpine soils were significantly correlated with plant species composition, with other abiotic environmental and spatial factors being less important. In general, if any unmeasured, yet selective, habitat features are spatially structured, they can produce a spurious correlation between microbial composition and geographic distance. This would lead to a potentially false conclusion that the biogeographic pattern is shaped by historical factors.

Dispersal is also particularly difficult to quantify for microscopic organisms, and thus geographic distance is used as a proxy indicator for the degree of dispersal between locations. However, this does not adequately represent the nature of the processes involved because successful dispersal is dependent on both the movement of a viable propagule in space and its successful establishment and colonization of a location. The lack of either movement or establishment, therefore, can produce a pattern consistent with dispersal limitation. The difference between these two, while seemingly subtle, could be the result of different ecological processes. For instance, lack of establishment may be caused by unfavorable local habitat conditions or biotic interactions (known as priority effects), whereas degree of movement may be a primarily stochastic process, especially for microbes, which are dispersed mostly by passive means. Therefore, the interpretation of microbial biogeographic patterns would be facilitated by future work on the importance of movement versus establishment. This insight could be attained by use of experimental manipulations, or by measuring the presence of microbial taxa in dispersal vectors such as air and rain.

Future Considerations

Evidence suggests that many microbes exhibit biogeographic patterns, and that these patterns are likely shaped by a combination of ecological and evolutionary processes, similar to animals and plants. However, while at least some microbes are dispersal limited, the importance of the contemporary environment is often more prominent than that of historical effects. Thus the future of microbial biogeography lies in identifying the conditions under which the relative importance of these underlying processes may vary. For example, response to environmental conditions and dispersal abilities may differ by taxonomic groups, taxonomic resolutions, habitat types, and spatial scales. If so, then the ability to detect biogeographic patterns and the importance of particular underlying processes may vary by scope, scale, and focus of the study. Here, some of the conditions under which biogeographic patterns may vary for microbes are highlighted, a consideration of which may facilitate detection of particular underlying biogeographic processes.

Taxonomic Focus and Resolution

Microbial taxa likely vary in traits that affect their dispersal and colonization abilities or their response to the environment, thereby influencing their distributions. Such traits include body/cell size, spore formation, and other dormancy life stages, or the ability to associate with other biological entities that could serve as dispersal vectors. Body size is thought to be negatively correlated with dispersal ability, whereby smaller organisms disperse farther than larger ones (Fenchel and Finlay, 2004). Additionally, spore formation and dormancy stages can increase the likelihood that a dispersing microbe will survive migration, allowing it to potentially travel greater distances until reaching a suitable habitat (Lennon and Jones, 2010). Microbial taxa also vary in traits that could affect the ability of migrants to colonize and establish in new locations through their response to the external environment. Some of these traits could include differential growth rates, "specialist" versus "generalist" physiologies, and susceptibility to pathogens or predators. Generalist taxa, for example, might be more resilient to environmental variation, and thus are able to colonize a larger range. Thus the relative importance of environmental variation and dispersal limitation may vary depending on the microbial taxa examined.

The ability to detect variation in microbial community composition, a necessity of the microbial biogeography study, is highly dependent on the level of taxonomic resolution used to define taxa. Taxonomic resolution is primarily constrained by the sensitivity of the identification method, and can also be arbitrarily defined in sequencing studies via sequence similarity cutoff values. Communities defined by low-resolution definitions or techniques may "lump" together taxa that vary in traits, obscuring differences between the relative importance of environmental selection and dispersal limitation among taxa. As for animals and plants, such broad-level taxonomic groups, akin to genus or family levels, are expected to have larger distributions than a particular species within a genus (Martiny *et al.*, 2006). Thus, biogeographic patterns may be

generally weaker (e.g., Horner-Devine *et al.*, 2004), and the underlying processes more intertwined, when microbial taxonomic definitions are broader.

Habitat Type

Habitat continuity could influence dispersal and therefore its relative importance in shaping biogeographic patterns. Highly connected habitats should be more linked via dispersal, thus harboring more similar communities (Lindström and Langenheder, 2011). Moreover, habitat types vary in degree of connectivity and the mechanisms of dispersal utilized by organisms in those habitats. Aquatic systems, for example, likely facilitate passive dispersal due to the fluidity of the medium. In contrast, movement of microbes may be more restricted within a more solid matrix such as soil. Therefore, microbial communities associated with solid substrates or "patchy" habitat types are expected to be more limited by dispersal than those in highly connected systems, such as the open ocean. Moreover, a biogeographic study across environmental gradients or across habitat types results in large amounts of environmental variation, and thus emphasizes the strength of environmental selection. In this situation, the influence of historical factors could be relatively weak, and possibly undetectable. Thus to target the potential effects of historical processes, habitat variability should be controlled for both via sampling design and statistical methods.

Spatial Scale

The ability to disperse from one location to another should decrease as the distance between those locations increases (Nekola and White, 1999). Accordingly, dispersal limitation should become more apparent as the spatial scale increases. This could result in a greater relative importance of historical factors at larger spatial scales. At the same time, clumping of animal and plant taxa due to restricted dispersal also happens on very short spatial scales. This small-scale clumping could very well be important in some microbial habitats such as biofilms and sediments (Martiny *et al.*, 2011). Thus, the spatial scale at which biogeographic patterns emerge for microbes may provide some insight into the degree of microbial dispersal that leads to certain biogeographic patterns.

Conclusions

Microorganisms exhibit patterns in their distributions across space at multiple scales, but given the immense diversity of microbes and their ecological roles, microbes are probably not uniformly shaped by the same ecological and evolutionary processes. The distributions of free-living microbes do appear to be influenced by both environment and history, with the relative importance of these depending on spatial scale, taxonomic resolution, habitat continuity, and life history traits. While at least some microbes appear to be dispersal limited, it is often secondary to the influence of the contemporary environment. Thus far, these conclusions are broadly similar to macroorganisms, although the relative strength of the underlying processes and particular environmental factors may vary.

The motivation for understanding microbial biogeography extends beyond drawing and interpreting a map of microbial diversity. Understanding microbial biogeographic patterns aids in investigating a range of microbial ecology questions – for example, how much and on what scales to sample microbial diversity, the nature and rate of microbial diversification or speciation, and the role of microbes in ecosystems. In fact, microbial composition drives a number of ecosystem processes, including decomposition, autotrophic and heterotrophic production, and nitrogen cycling. Therefore, even under similar environmental conditions, microbial communities in different regions might function differently. A better understanding of microbial biogeography is essential to predict the effects of microbial diversity on ecosystem functioning.

Appendix

List of Courses

1. Biodiversity
2. Biogeography
3. Environmental Microbiology
4. Microbial Ecology and Evolution

See also: Biogeography, Overview. Dispersal Biogeography. Diversity, Community/Regional Level. Elevational Trends in Biodiversity. Metapopulations. Microbial Biodiversity. Population Genetics

References

Baas-Becking LGM (1934) *Geobiologie of Inleiding Tot de Milieukunde*. The Hague: Van Stockkum & Zoon.

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, and Green JL (2008) Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11505–11511.

Cermeño P and Falkowski PG (2009) Controls on diatom biogeography in the ocean. *Science* 325: 1539–1541.

Cho JC and Tiedje JM (2000) Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology* 66: 5448–5456.

Fenchel T and Finlay BJ (2004) The ubiquity of small species: Patterns of local and global diversity. *Bioscience* 54: 777–784.

Fierer N and Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* 103: 626–631.

Fierer N, McCain CM, Meir P, *et al.* (2011) Microbes do not follow the elevational diversity patterns of plants and animals. *Ecology* 92: 797–804.

Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061–1063.

Foissner W (2006) Biogeography and dispersal of micro-organisms: A review emphasizing protists. *Acta Protozoologica* 45: 111–136.

Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, and Naeem S (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. *Proceedings of the National Academy of Sciences of the United States of America* 103: 13104–13109.

Fuhrman JA, Steele JA, Hewson I, *et al.* (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 105: 7774–7778.

Gilbert JA, Field D, Swift P, *et al.* (2009) The seasonal structure of microbial communities in the Western English Channel. *Environmental Microbiology* 11: 3132–3139.

Green JL, Holmes AJ, Westoby M, *et al.* (2004) Spatial scaling of microbial eukaryote diversity. *Nature* 432: 747–750.

Hedlund BP and Staley JT (2003) Microbial endemism and biogeography. In: Bull AT (ed.) *Microbial Diversity and Bioprospecting*, pp. 225–231. Washington, DC: ASM Press.

Hillebrand H (2004) On the generality of the latitudinal diversity gradient. *American Naturalist* 163: 192–211.

Horner-Devine MC, Lage M, Hughes JB, and Bohannan BJM (2004) A taxa-area relationship for bacteria. *Nature* 432: 750–753.

Kriss AE (1960) Micro-organisms as indicators of hydrological phenomena in seas and oceans. 1. Methods. *Deep-Sea Research* 6: 88–94.

Lennon JT and Jones SE (2010) Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America* 107: 5881–5886.

Lindström ES and Langenheder S (2011) Local and regional factors influencing bacterial community assembly. *Environmental Microbiology Reports* 4: 1–9.

Martiny JBH, Bohannan BJM, Brown JH, *et al.* (2006) Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology* 4: 102–112.

Martiny JBH, Eisen JA, Penn K, Allison SD, and Horner-Devine MC (2011) Drivers of bacterial beta-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences of the United States of America* 108: 7850–7854.

Nekola JC and White PS (1999) The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* 26: 867–878.

Nelson JD, Boehme SE, Reimers CE, Sherrell RM, and Kerkhof LJ (2008) Temporal patterns of microbial community structure in the Mid-Atlantic Bight. *FEMS Microbiology Ecology* 65: 484–493.

Nemergut DR, Costello EK, Hamady M, *et al.* (2011) Global patterns in the biogeography of bacterial taxa. *Environmental Microbiology* 13: 135–144.

Pommier T, Canback B, Riemann L, *et al.* (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology* 16: 867–880.

Redford AJ and Fierer N (2009) Bacterial succession on the leaf surface: A novel system for studying successional dynamics. *Microbial Ecology* 58: 189–198.

Ricklefs RE (2007) History and diversity: Explorations at the intersection of ecology and evolution. *American Naturalist* 170: S56–S70.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.

Shen J, Wang JJ, Soininen J, Zhang Y, Wang BX, and Yang XD (2011) Contrasting patterns in elevational diversity between microorganisms and macroorganisms. *Journal of Biogeography* 38: 595–603.

Sogin ML, Morrison HG, Huber JA, *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America* 103: 12115–12120.

Van der Gucht K, Cottenie K, Muylaert K, *et al.* (2007) The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proceedings of the National Academy of Sciences of the United States of America* 104: 20404–20409.

Ward DM, Weller R, and Bateson MM (1990) 16S ribosomal-RNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* 345: 63–65.

Whitaker RJ, Grogan DW, and Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301: 976–978.

White EP, Adler PB, Lauenroth WK, *et al.* (2006) A comparison of the species–time relationship across ecosystems and taxonomic groups. *Oikos* 112: 185–195.

Zinger L, Lejon DPH, Baptist F, *et al.* (2011) Contrasting diversity patterns of crenarchaeal, bacterial and fungal soil communities in an alpine landscape. *PloS One* 6.

Microbial Diversity

Paul V Dunlap, University of Maryland Biotechnology Institute, Baltimore, MD, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 191–205, © 2001, Elsevier Inc.

Glossary

Aerobe An organism that utilizes or requires the presence of oxygen for growth.
Anaerobe An organism able to grow in the absence of oxygen.
Autotroph An organism able to utilize carbon dioxide as its source of carbon.
Barotolerant and barophilic Able to tolerate high pressures and growing better under high pressure.
Bioluminescence Light production by living organisms.
Chemotroph Organisms that utilize chemicals as sources of energy.
Cryptoendolithic Living within the surface of rocks.
Elective culture The provision of appropriate physical and chemical conditions that elicit the growth of specific metabolic types of microbes.
Extremophile An organism that grows better at, or requires for growth, extremes of temperature, pressure, salinity, or other environmental factors.
Halophile An organism requiring high levels of salts for growth.
Heterotrophy An organism that obtains its carbon from organic carbon compounds.
Microbes Single-celled organisms, such as bacteria, archaea, protists, and unicellular fungi.
Phototroph An organism utilizing the energy of light, as in sunlight, for growth.
Psychrophile An organism that grows better at low temperature or requires low temperature for growth.

"The key to taking the measure of biodiversity lies in a downward adjustment of scale. The smaller the organism, the broader the frontier and the deeper the unmapped terrain" (Wilson, 1994).

Introduction

Rapidly accumulating evidence indicates that microbes most likely account for the vast majority of kinds of organisms on earth. Microbes carry out a stunningly diverse array of metabolic activities, several of which were instrumental in creating conditions for the evolution of other life forms. Through their colonization of diverse and extreme environments, their geochemical cycling of matter, and their biological interactions among themselves and with all other organisms, microbes define the limits of the biosphere and perform functions essential for ecosystem development and health. However, because microorganisms are predominantly unicellular life forms that generally are smaller than can be seen with the unaided eye, they historically have received disproportionately little scientific attention compared to that given to animals and plants. This lack of attention has begun to shift recently as awareness of the diversity of microbes and their biological importance has grown. Of the three presently recognized domains of life, two, the Bacteria and the Archaea, are entirely microbial, and the third, the Eucarya, through its vast array of protists and fungi, is primarily microbial. The essence and full scope of the diversity of microbes is revealed in the dramatic differences among these microorganisms in their phenotypic characteristics of cellular metabolism, physiology, and morphology, in their ecological distributions and activities, and in their genomic structure, expression, and evolution. Appreciation for the true extent of microbial diversity is growing rapidly through the development and use of molecular phylogenetic approaches, which are enabling rigorous analysis of the origins and evolution of microbial life. In combination with classical methods of elective culture, isolation, and phenotypic analysis, the approaches of molecular phylogeny are stimulating the discovery of multitudes of new microorganisms, opening up their biology for study, and providing a clear understanding of the importance of microorganisms as the functional and evolutionary foundation of the biosphere.

The Scope of Microbial Diversity

We live on "a microbial planet" (Woese, 1999) in the "Age of Bacteria" (Gould, 1996). Microorganisms, the first cellular life forms, were active on earth for more than 3.0 billion years before the development of multicellular, macroscopic life forms. During that time and continuing into the present, through the invention of a spectacular array of different metabolic and physiological capabilities, microbes evolved to exploit the multitude of environments and microhabitats presented by the abiotic world. They thereby obtained the cellular building materials and energy necessary for growth and reproduction. In so doing, however, they progressively altered the geochemical conditions of the planet, leading to a continual development of new conditions and habitats, abiotic and biotic. Those new conditions and habitats presented both challenges to survival and opportunities to exploit,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00435-4

leading to continuing evolution of distinct microbial types able to endure or take advantage of the biogeochemical changes taking place on earth. Once cellular life began, it is likely that no place on earth containing the molecules and energy conducive to life remained abiotic for long.

The evolutionary trend toward greater complexity, seen in the relatively recent appearance of multicellular life forms (e.g., plants and animals), however, did not cause microbes to be displaced. The appearance of plants and animals did not shunt the unicellular microbes to forgotten corners of the biosphere to hang on and eke out a marginal existence. Instead, multicellular organisms, which themselves can be viewed as highly evolved, complex assemblages of microorganisms, have provided unicellular microbes with a wide variety of new habitats to colonize and exploit. Consider the various microbes whose growth is favored by the different and changing habitats provided by the growth and senescence of roots, stems, leaves, flowers, and fruits during the life of plants. Consider the multitude of physicochemically distinct habitats of the human skin, of our mucous membranes, and the changing environments of our complex intestinal system. Along with these habitats, colonized often by assemblages of several different kinds of microbes, consider the species-specific developmental and metabolic symbioses certain bacteria have established with plants, such as nitrogen-fixing *Rhizobium* with legumes, and with animals, such as bioluminescent *Vibrio* with marine cephalopods and fishes (**Figure 1**). Instead of passing on the torch of preeminence in life to the developing multicellular organisms and then politely withdrawing from the life center stage, microbes were full participants and driving forces in that development, they have continued to diversify, and they remain fully dominant. Plants and animals provide a highly visible but thin multicellular skin over a biologically rich and complex microbial world. A sense of the biological dominance of microbes is given by estimates of the total number of living bacteria, roughly 5×10^{30} cells, with a collective biomass, despite their small size, possibly equal to that of all other life forms (Whitman *et al.*, 1998). We live in the once-and-always age of microbes.

Microbes of one kind or another, especially bacteria, survive and grow almost everywhere on earth. Whether widely distributed, having been spread globally by winds, water currents, and animals, or occurring only in localized areas where they have adapted to grow under specific environmental conditions, microbes dramatically extend our perception of the limits of the biosphere. Previously, that perception was shaped to a large extent by our notion of the conditions under which plants and animals can live. As aerobic organisms, we require the high levels of oxygen present in air for survival. While many bacteria utilize oxygen as we do, most, however, grow best at lower oxygen levels, and anaerobic microbes of many different types require the strict absence of oxygen to survive, as found, for example, in sediment and gut tracts. Temperatures that from a human perspective are extreme—from just below the freezing point of water in some oceanic waters and sea ice, to several degrees above its boiling point in waters near hydrothermal vents and in hot springs—are not extreme at all to the bacteria that colonize these habitats. Indeed, water temperatures considered cold for humans (i.e., 15 °C) can be lethally hot to true cold-loving, psychrophilic bacteria. Barotolerant and barophilic bacteria, active at and requiring the extremely high pressures of the deep sea, exist at pressures that would crush the human body, while other microbes, the halophiles, found in salt lakes and solar evaporation ponds, require salt concentrations that would quickly pickle our tissues. Add to this the cryptoendolithic microbes living just below the surfaces of sandstone in the Antarctic, the acid-tolerant and acid-requiring bacteria of acid mine drainage and sulfur springs, and the alkalinity-requiring bacteria of desert soils and alkaline lakes that live at pH levels that would be caustic to our skin. This array of

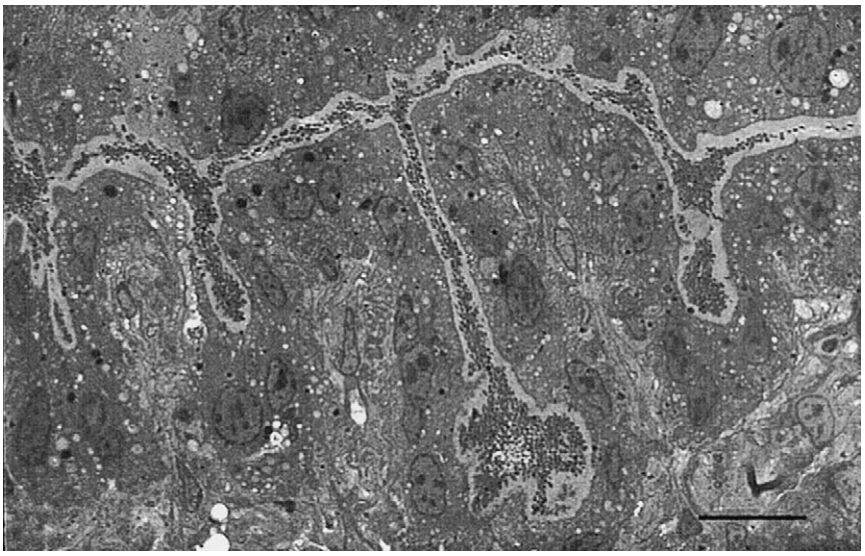


Figure 1 Light-micrograph of a section of the light organ of the sepiolid squid *Euprymna scolopes*. The animal harbors a dense population of the luminous marine bacterium *Vibrio fischeri* extracellularly within a ventral tissue complex, the light organ, and uses the light produced by the bacteria in predator avoidance.

microbial attributes gives a sense of the physical and chemical extremes at which certain bacteria survive and grow (Madigan *et al.*, 1999). These organisms have no "protection" from the environment other than their inherent metabolism and physiology, which, instead of protecting them, exquisitely suit them to living in these extremes, on which they are dependent. Major research initiatives around the world seek to expand knowledge of "extremophile" biology identifying these "exotic" microorganisms and bringing into study the metabolic and physiological attributes that adapt them to life at the physical and chemical limits of the biosphere.

What about microbial diversity in less extreme environments, away from the limits of the biosphere? Many environments, such as garden soil, coastal seawater, and lake sediments, do not exhibit such dramatic extremes of temperature, acidity, pressure, or other factors. Microbial life in these environments is strikingly diverse. While many different types of microbes can be isolated and grown in laboratory culture from these environments, most so far cannot. A typical expectation is that much less than 1 percent to a few percent of the microbial types seen in an environmental sample will grow in culture (e.g., Amann *et al.*, 1995). That means the vast majority of microbes, even from commonly studied environments such as coastal seawater, have not been brought into study or identified. Thus, discoveries of new types of microbes are waiting to be made, and they are being made. Examples of the unexpected include the discovery of magnetotactic bacteria, which synthesize intracellular chains of magnetic granules that orient the cells to magnetic fields (e.g., Amann *et al.*, 1995) and the occurrence of new members of presumably anaerobic Archaea in oxygenated seawater at shallow depth (e.g., DeLong, 1998). These reports demonstrate that microbial life in more accessible and commonly encountered environments is still poorly understood and far more diverse than presently known or expected. Therefore, most habitats that are known to be rich in microbial life and that have been sampled, such as seawater, soil, and animal gut tracts, have not yet yielded for study anywhere near their full complement of microbes. And what about easily accessible habitats likely to be rich in microbial life that have not yet been examined at all or at best have been sampled only minimally? One type of habitat is the gut tracts of the several hundred thousand known insect species, only very few of which, such as the termite, have been examined microbiologically. Yet the gut tract of each species—because of the different foods the animal eats, the gut specific morphology and physiology for digesting that food, and the environmental conditions under which the animal lives—is likely to host its own very different kinds of microbes. Our "microbial planet" is largely unexplored.

The Biological Significance of Microbial Diversity

Microorganisms are "the foundation of the biosphere" (Staley *et al.*, 1997), providing its "essential, stable underpinnings" (Woese, 1999). Microorganisms have played and continue to play fundamental roles in the evolution of higher life forms on earth. They have done so and continue to do so through the essential ecological processes they carry out in obtaining the materials and energy needed for growth and reproduction. A primary example of the evolutionary role microbes have played is oxygenic photosynthesis, invented by phototrophic bacteria more than 2 billion years ago and which releases oxygen as a by-product of energy generation. Over time, that release of oxygen led to a gradual change in the earth atmosphere from reducing to oxidizing. The oxidizing atmosphere, as it developed, allowed energetically more efficient aerobic organisms to evolve and provided a protective shield of ozone against ultraviolet radiation for terrestrial and aquatic organisms. Equally striking examples are the bacterial endosymbiotic origins of chloroplasts, light-harvesting organelles in plants, and of mitochondria, energy-generating organelles, major events in the evolution of plant and animal lineages in the Eucarya. Furthermore, the fixation of atmospheric nitrogen, reducing nitrogen gas to ammonium and converting it into organic nitrogen forms, is an entirely bacterial activity, carried out by various symbiotic and free-living microbes (Madigan *et al.*, 1999).

The ecological processes carried out by microorganisms are equally fundamental. For example, global biogeochemical cycles of major elements, carbon, nitrogen, sulfur, and iron, essential components of all living cells, operate through microbial activity. Specifically, degradation of complex carbohydrates such as chitin, forming the exoskeleton of arthropods, and cellulose, hemicellulose, and lignin, structural polymers in plants, is essential. Without microbial conversion, these polymers would accumulate, removing huge amounts of carbon from the biosphere and blocking a multitude of biological processes that allow micro- and macroorganisms to live. In the absence of these microbial degradative processes, life on earth would soon falter. Besides the microbial degradation of complex organic compounds, consider the range of metabolic diversity in microbes, from oxygenic and anoxygenic photosynthesis, sulfate reduction, methanogenesis, denitrification, iron oxidation, nitrite oxidation and nitrate reduction, hydrogen and methane oxidation, and so on, all ways by which microbes obtain the energy necessary for growth and reproduction. These considerations form in part the basis for a commonly held view that bacteria and other microbes, in carrying out these processes, serve humans and other higher organisms as environmental recyclers and bioremediators. That view, while essentially correct, overlooks an essential point—these activities and processes are the fundamental biology of this planet. Microbes are the biosphere; their activities create and provide the foundation for all other life.

To gain a perspective on the significance of microbial diversity, imagine a biological survey crew tasked with discovering and documenting life forms on a newly encountered planet. Consider that upon landing, the crew found, remarkably, no macroscopic life. However, suppose that an initial sampling of a cubic centimeter of the planet surface revealed the presence of millions of discrete microscopic cellular entities. Imagine that with much additional analysis these entities were found by the crew to represent thousands of different kinds of organisms, distinguishable by their morphology or their dramatically different ways of obtaining the energy and nutrients necessary for metabolism and reproduction. Would the survey crew be surprised? What if upon

analysis of the genetic material from the different types present, these microbial life forms were confirmed to be different and were found in many cases to be dramatically more distinct from each other evolutionarily than are seaweeds and humans, would the crew be impressed? What if the crew continued sampling, spreading out and choosing other locations of the planet surface, and found similar "species richness" wherever they looked, but often with little or no overlap in the types of entities present from one environment to the next. Would that start the crew thinking about, to paraphrase E. O. Wilson (1994), "a strange and vastly complex living world virtually without end"? It would, of course. However, there is no need to invoke new planets. This imaginary scenario describes the reality of microbial diversity on earth. The only difference is the microbially driven evolution of macroscopic life forms on earth, giving rise to the plants, animals, and macroscopic fungi.

A New Era in Biological Sciences

Despite the complexity and richness of microbial life and the essential evolutionary and ecological roles played by microbes, awareness of microorganisms remains limited. The diversity and scientific importance of microbes have been largely passed over in human society, in science, in biology, and even in discussions of biodiversity (Hawksworth, 1991), overshadowed by attention to macroscopic forms. The prevailing and erroneous view for many biologists is that bacteria are "primitive, simple and relatively uniform" (Pace, 1996). This view developed naturally from early technical and scientific limitations for discovering and studying bacteria and other microbes in the 18th and 19th centuries, such as the need for high-resolution microscopy and the need to understand cellular structure, biochemistry, polypeptides, and nucleic acids. Such limitations did not hinder as starkly the beginnings of macrobiology. Later, as limitations to the study of microbial life were overcome in the first three-quarters of the 20th century, the bias against microbes as scientifically important biological systems was nonetheless maintained and reinforced through the lack of a comprehensive phylogeny of microbes. This situation was especially true for bacteria, which generally lack distinctive morphological characters and for which sexual reproduction like that of plants and animals is absent. That bias remains largely extant today. For example, most college and university departments of biology and biological sciences are staffed predominantly with animal and plant biologists, with relatively few if any microbial biologists. Yet "the incongruity (between the scientific perception of microbiology and the preeminence of microorganisms in the real world) is astounding; it is worrisome; it cannot be scientifically justified or tolerated" (Woese, 1999). Fortunately, the bias and incongruity are beginning to be eliminated.

The "Delft School" of General Microbiology

The confluence of two distinct but complementary approaches in biological science, one classical and one more recent, is leading to a shift in awareness about microbial diversity and its scientific importance. The classical approach is that of elective culture, also referred to as enrichment culture, by which new microbial types are brought into culture and isolated for phenotypic analysis. The elective culture approach, through the careful design of growth media and conditions, seeks to provide an appropriate physical and chemical environment that, when inoculated with an environmental sample (mud, for example), will elicit the growth of specific metabolic types of bacteria postulated to be active or present in the sample. For success, the approach requires a thorough knowledge of biochemistry, good observational skills, and sensitivity to potential novelty in microbial metabolism and physiology. A recent example of this approach, demonstrating its central importance and value in microbial research, is the isolation of acetogenic sprirochaete bacteria from the hindguts of termites (Leadbetter *et al.*, 1999). Spirochetes are major members of the diverse microbial consortium resident in the termite hindgut (**Figure 2**), and they have been thought to play key roles in the insect nutrition, which is based on

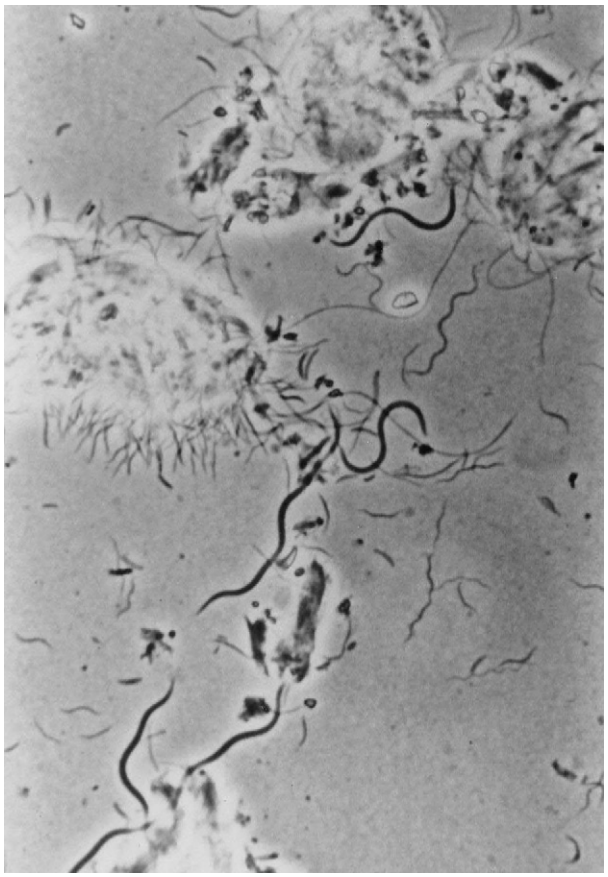


Figure 2 Phase contrast micrograph of the contents (diluted) of the hindgut of the termite *Reticulitermes flavipes* (Kollar). The microbial assemblage of this common eastern subterranean termite includes several types of microbes, including a variety of different kinds of spirochetes, both free and attached to cellulolytic protists. H_2 and CO_2, formed during microbial cellulose digestion, are converted by certain of the spirochetes to acetate (Leadbetter *et al.*, 1999) a major source of nutrition for the animal. The different spirochaetal cells range in length from approximately 5 μm to over 30 μm.

microbial degradation of cellulose and conversion to acetate, a major carbon and energy source for the insect. However, no spirochetes from the termite gut previously had been isolated in pure culture. That inability limited knowledge of the contribution these morphologically distinct and numerically significant bacteria make to host animal nutrition. Leadbetter and coworkers, however, successfully established culture conditions that favored the growth of spirochetes over other bacteria and that simultaneously encouraged the growth of bacteria able to form acetate, from H_2 and CO_2, breakdown products of cellulose degradation. In this way, acetogenic spirochetes from the termite hindgut were brought into pure culture for the first time. H_2/CO_2 acetogenesis, a type of metabolism previously unknown among spirochetes, reveals an important way, formation of acetate, in which these bacteria contribute to termite nutrition. The ability to design culture conditions that elicit the growth of specific bacterial types known or suspected to be present in an environment is central to development of our understanding of microbial diversity and microbial ecology. Once a novel microbial type has been brought into culture, that organism's special or unique cellular metabolism, physiology, and genetics can then be studied in detail.

This now classical elective culture approach developed through the insights of Sergei Winogradsky, a Russian soil microbiologist, and Martinus Beijerinck and later Albert Kluyver, Beijerinck's successor, in Delft, Netherlands. Their work was instrumental in forming the foundations of general microbiology and microbial ecology. When Cornelis van Niel, a student of Kluyver, moved in 1928 from Delft to the Hopkins Marine Laboratory in Pacific Grove, California, he brought with him the "Delft School" tradition in microbiology, which he continued and further developed through his research and through a course he taught in general microbiology and comparative biochemistry (van Niel, 1949). Through his course van Niel trained a generation of microbiologists. Those individuals have gone on to use the Delft School, van Niel approach, in their research, and they in turn have taught other generations of microbiologists, further disseminating the Delft School tradition. Most notably today, the elective culture and isolation approach to the study of metabolically and ecologically diverse bacteria is fostered in young microbiologists and other scientists through the Microbial Diversity summer course of the Marine Biological Laboratory in Woods Hole.

The "Woesean Reformation" of Microbiology

Despite the progress arising from the Delft School tradition in exploring and defining the diversity of bacteria and their metabolic capabilities, a major problem limited the development of general microbiology in the first three-quarters of this century: the lack of a unifying phylogeny of microorganisms. Various efforts at classifying bacteria and systematizing relationships among them, based on phenotypic characters of morphology and biochemical growth attributes, were attempted and abandoned. The frequent lack of characters, the instability or the shared nature of many characters, and the awareness that phenotypic characterization was not grounded necessarily at the genetic level left these efforts with major flaws. The inability to place microbes, especially bacteria, in an evolutionary context caused general microbiology and microbial diversity largely to languish as mainstream sciences at a time when other areas of biology and microbiology (e.g., clinical microbiology and biotechnology) were developing rapidly (Woese, 1999). What was needed was a unifying phylogenetic framework, founded at the genome level, which would allow the true evolutionary relationships among microbes to be analyzed critically and defined.

The use of informational macromolecules, begun more than 30 years ago, is now fulfilling that need for a unifying phylogeny of microbes (Woese, 1987; 1999) and is reforming our view of the evolution and diversity of microbial life. Analysis primarily of ribosomal RNA (rRNA) sequences, especially for the small subunit 16S and 16S-like rRNAs, has created "the first valid microbial phylogenetic systematics" (Jannasch, 1997). The functionally constant 16S and 16S-like molecules, common to all organisms, contain evolutionarily highly conserved regions, suitable for comparing less closely related organisms, and more variable regions, suitable for assessing evolutionary relationships in more closely related organisms. The universality of ribosomal RNA extends the value of this sequence comparison approach to all life forms, but importantly for bacteria it has established a unified phylogenetically based system with which to begin defining bacterial evolutionary relationships, a "first step in microbiology reformation" (Woese, 1999). Along with 16S rRNA, other molecules, such as elongation factor Tu, 23S rRNA, and F_1F_0 ATPase, also provide substantial phylogenetic information and can serve as alternative markers for inferring relationships (Ludwig and Schleifer, 1999). A variety of opportunities exist at various institutions for learning the principles and application of molecular phylogenetic analysis in microbes and higher organisms. Notable among them is the Marine Biological Laboratory course in Molecular Evolution, an intensive 3-week course dedicated to these topics.

A second step in the Woesean reformation of microbiology is the application of rRNA-based molecular phylogeny to microbial ecology (e.g., Hugenholtz *et al.*, 1998; Pace, 1996). Sequence analysis of 16S rRNAs, for example, extracted from natural environments provides direct access to the diversity of bacteria in that environment, bypassing the need for culturing microbes and giving a rapid and potentially comprehensive assessment of microbial community composition. rRNA-based approaches have opened up for study many microbial activities and associations. The result is a rapidly expanding awareness of the diversity and ecology of microbes, both culturable and not-yet-cultured (Amann *et al.*, 1995; Hugenholtz *et al.*, 1998; Pace, 1996, 1999).

The strengths of the elective culture and molecular phylogeny approaches make them naturally complementary. The cultivation of a new microbe leads to acquisition of the organism 16S or 16S-like rRNA sequence in the context of data on its metabolism, physiology and habitat. That sequence provides a defined point of phylogenetic reference and a highly specific tool with which to examine the organism distribution in the environment. Equally exciting is the opportunity to examine and explore the environment for the full, natural microbial diversity present, without concerns about

the bias inherent in and finesse required for culturing. Those explorations identify and define more deeply within a unified phylogenetic framework the diversity of microbes present while also offering potential insights into their metabolism and physiology. That information, in turn, can motivate more refined or novel attempts at cultivation of the sequence-identified microbes. Each approach nurtures and magnifies the strengths of the other. For microbiology in the next century truly to "emerge as the primary biological discipline" (Woese, 1999) the continuing confluence of these approaches must be encouraged.

Major Groups of Microbes

Through analysis primarily of 16S and 16S-like rRNA genes, microorganisms can now be placed in a potentially comprehensive phylogenetic framework, one that includes all living organisms and therefore is universal. Examination of the 16S and 16S-like rRNA-based universal phylogenetic tree (**Figure 3**) shows the three currently proposed domains of life, Bacteria, Archaea, and Eucarya. Future analyses, using expanding sequence data sets and markers of phylogenetic relationships other than 16S and 16S-like rRNA (Ludwig and Schleifer, 1999) will serve to test and refine the validity of this evolutionary grouping. Regardless, the current universal tree reveals in a simple, compelling way the dominance of microbial life forms. The true diversity of life is microbial.

Domain Bacteria

The unit of biological diversity is the species. The classical concept of a biological species, a reproductively isolated interbreeding or potentially interbreeding population of

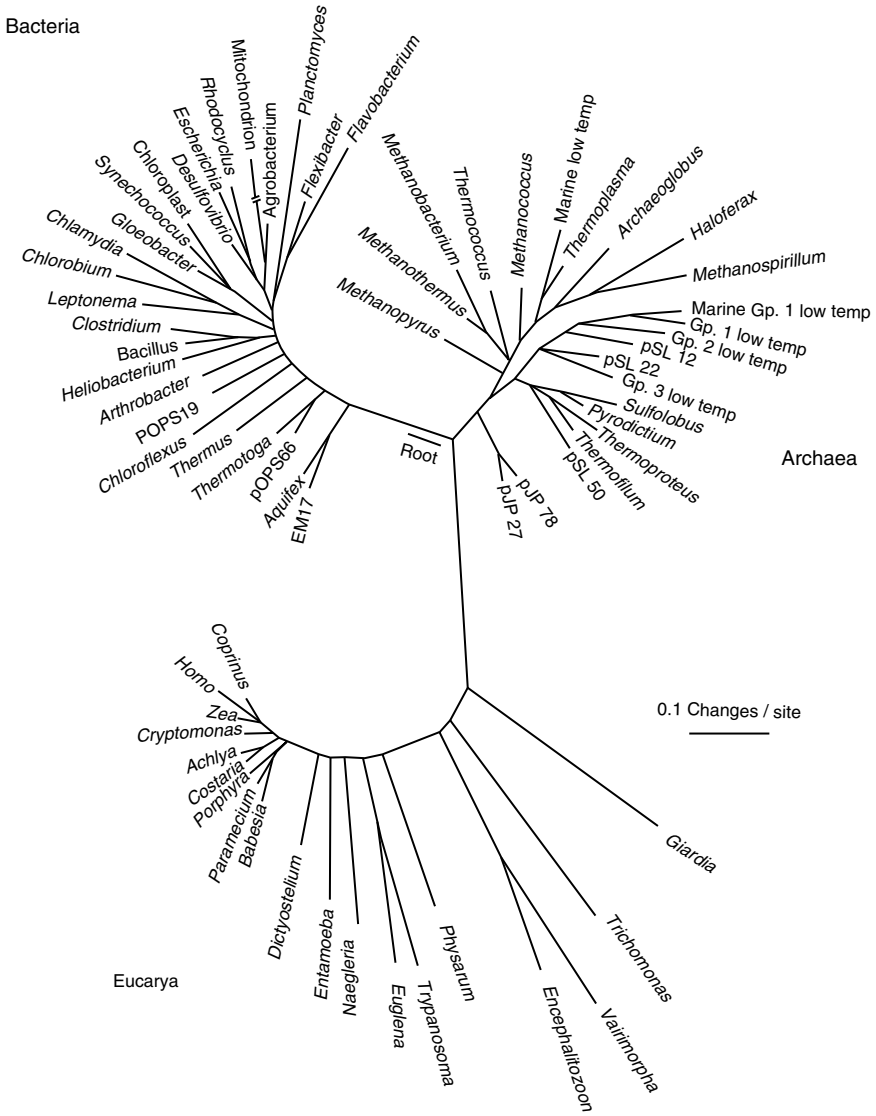


Figure 3 The universal tree of life. The three domains of life, Bacteria, Archaea, and Eucarya, represented by small-subunit rRNA sequences of various organisms within each domain, are shown. The domain Bacteria (Eubacteria) and the domain Archaea (previously the archaebacteria) are entirely microbial, and the domain Eucarya is predominantly microbial.

individuals, however, does not work for bacteria and other microbes that do not interbreed, that have undefined or cosmopolitan distributions, and that generally lack distinguishing morphological characters. So, in discussing bacteria, members of the domains Bacteria and Archaea, how is a species defined? The phylogenetic species concept, a group of individuals for which phylogenetic analysis has demonstrated a shared genealogical relationship (Alexopoulos *et al.*, 1996), seems more workable for bacteria and other asexually reproducing microbes. The question then becomes the extent of genealogical relationship between two individuals necessary to designate them as members of the same species. Operationally for bacteria, if the genomic DNA of two strains is 70% or more similar, as determined by DNA-DNA hybridization analysis, they are considered the same species. With respect to rRNA, 16S rRNA sequence, similarity values below 97% are a strong indication that the two bacteria are different species (Amann *et al.*, 1995; Madigan *et al.*, 1999). The combination of these two approaches, both based at the genomic level, is powerful. When further combined with the identification of distinguishing phenotypic characters, such as cell morphology, motility, and flagellation, response to oxygen, requirement for organic growth substrates, growth temperature range, special metabolic attributes, characteristics of the habitat, and so on, a character-rich, biologically meaningful description of a bacterial species can be obtained.

Conservative estimates place the total number of species of bacteria at 50,000 to 3,000,000 (e.g., Staley *et al.*, 1997). As of 1999, approximately 5000 bacterial species had been described (Pace, 1999), a very small portion of the estimated total number, though many of these descriptions do not yet include rRNA sequence information. Most easily accessible environments, even those commonly sampled for microbial life, are likely to contain a multitude of as-yet-uncultured bacteria. Another important consideration in estimates of the number of bacterial species is the extent to which microhabitats, habitats relevant from the microbe viewpoint, have been sampled. A reasonable assumption is that most biotic and many abiotic surfaces are colonized by bacteria. Each surface presents a different microhabitat and therefore is likely to be colonized by a different individual bacterial type or assemblages of types, with the "many different microenvironments creating an almost infinite variety of selective conditions" (Palleroni, 1994), that is, conditions selecting for the specific metabolic and physiological types. An estimated total number of plant and animal species is approximately 9 million, with insects making up the majority of those species (Staley *et al.*, 1997). Presumably, the external surfaces of these organisms, and also the mouths, gut tracts, and internal tissues of the animals, provide myriad types of microhabitats for bacterial colonization by assemblages of different microbes (Amann *et al.*, 1995) and in many cases by specific bacterial symbionts, as in the various nutritional endosymbioses between bacteria and insects. Consequently, these estimates of bacterial species must be too low. When the approximately 2 million species of fungi and protists, many of which undoubtedly also have largely different assemblages of microbes as well as specific bacterial symbionts, also are factored in, then the total number of bacterial species easily could exceed tens of millions. Future generations of microbiologists will likely find even this estimate to be conservative. Regardless, however, of what the actual total number of bacterial species turns out to be, bacteria are stunningly diverse.

Previously, all bacteria were grouped together taxonomically as prokaryotes. That grouping was based primarily on the common lack of a membrane-bounded nucleus in bacteria. Prokaryotic organisms, however, are now separated phylogenetically into two domains, the Bacteria (or Eubacteria) and the Archaea (formerly the archaebacteria). The Archaea exhibit many similarities to the Eucarya, demonstrating that the prokaryotic body plan is not a phylogenetically definitive character. The bacteria exhibit a wide variety of ways of obtaining for growth the necessary carbon, as in the various kinds of heterotrophic and autotrophic microbes, and energy, as in phototrophic and chemotrophic microbes.

The domain Bacteria presently contains well over two dozen divisions, or kingdoms, of organisms (**Figure 4**). Madigan *et al.* (1999) suggest, however, that the true number of kingdoms in the Bacteria may be 50 or more. As the number of these groups indicates, physiological diversity within the Bacteria is profound. The diversity is especially striking in two of the kingdoms, the Proteobacteria and the Gram-positive bacteria (Madigan *et al.*, 1999), members of which exhibit a wide range of different metabolisms, physiologies, morphologies, and habitats. Examples among the Proteobacteria include the human enteric bacterium *Escherichia coli*, pathogenic pseudomonads, light-producing marine photobacteria, the fever-causing rickettsias, gliding bacteria and fruiting-body formers, stalked and budding bacteria, nitrogen-fixers, sulfate reducers, and so on. Within the Gram-positive kingdom are staphyloccoccal parasites of humans, the milk-sugar fermenting lactobacilli, and spore-forming soil bacteria, among many others. This diversity, however, probably reflects more the relative ease of cultivation and long-term scientific attention members of these two kingdoms have received than the true biological richness of the groups. One can anticipate that much more diversity in other kingdoms of the Bacteria will be revealed with time as bacteria are studied more intensively. Furthermore, as the fusion between elective culture and molecular phylogeny develops, many other types of bacteria, some entirely unexpected, will be discovered and will be found to warrant kingdom status. A recent text (Madigan *et al.*, 1999) highlights the different groups of bacteria, providing information on individual types within a phylogenetic context.

Domain Archaea

The domain Archaea includes the majority of presently known "extremophiles," organisms that live at physical or chemical extremes. Archaea increasingly are being discovered, however, in less extreme types of environments, including the marine plankton, lakes, and sediments (e.g., DeLong, 1998; Vetriani *et al.*, 1999). Diversity within the Archaea is presently less well understood than in the Bacteria and Eucarya because the Archaea often require particular care to culture. Knowledge of how to culture the Archaea has expanded in recent years, and additional habitats supporting their growth are being actively studied, so the disparity in the numbers of different kinds of

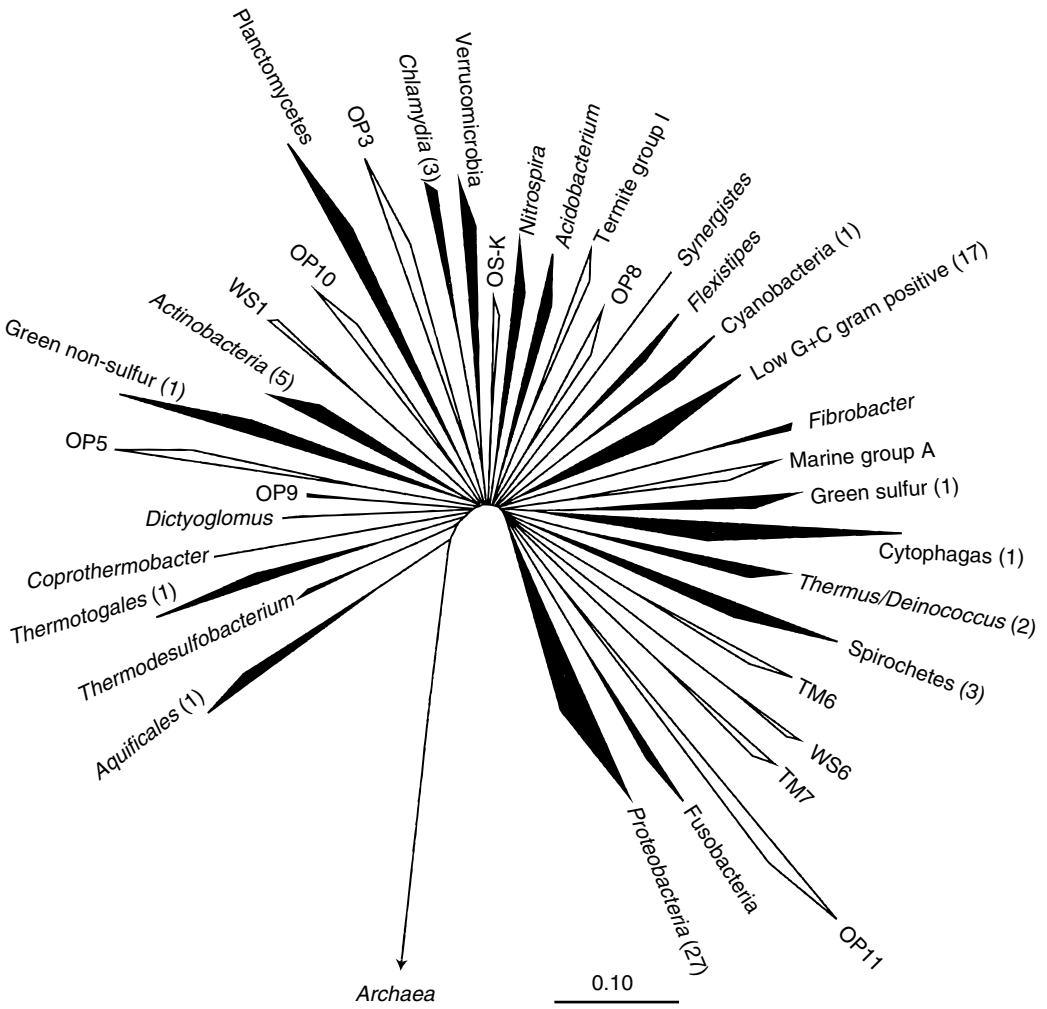


Figure 4 Diversity and phylogenetic relationships among members of the domain Bacteria. The division level grouping are of two types: currently recognized divisions, represented by cultivated bacteria (black), and provisional divisions, represented to date by environmental sequences (outline). The scale bar represents 0.1 changes per nucleotide.

Archaea compared Bacteria should progressively diminish. The Archaea consists at this time of three kingdoms, the Euryarcheaota, the Crenarcheaota, and a provisional kingdom, the Korarchaeaota (**Figure 5**).

Two major types of bacteria, the methanogens and the extreme halophiles, are included within the Euryarchaeota. Methanogens are bacteria that produce methane as an end-product of energy conversion reactions; they occur in a variety of strictly anaerobic habitats, such as sediments, sewage sludge digestors, the rumen of cattle, and the termite hindgut. CO_2 and similar compounds, methanol and other methyl-containing compounds, and acetate are used by different members of this group as substrates for methanogenesis, often with H_2 as an electron donor. Methane, primarily from methanogenic bacteria, is an important greenhouse gas, accounting for possibly as much as a few percent of total primary production. Some methanogens, such as *Methanococcus jannaschii*, are hyperthermophiles. The extreme halophiles, represented by *Halobacterium salinarum*, are bacteria that require for survival and growth the exceptionally high salt concentrations found in salt lakes and solar evaporation ponds. Other groups of Euryarchaeota include a lineage of extremely acidophilic bacteria containing *Thermoplasma* and *Picrophilus*, and two lineages of hyperthermophiles, represented by *Pyrococcus* and *Archaeoglobus* (**Figure 4**).

The Crenarchaeota contains a large and physiologically diverse group of hyperthermophilic, sulfur-metabolizing bacteria from terrestrial and marine hot springs and hydrothermal vents. Recently, several Crenarchaeota (and Euryarchaeota) have been identified at the 16 S rRNA level as members of the plankton in cold oceanic surface and deep waters, coastal sediments, lakes, and in association with animals, indicating that the Archaea are more cosmopolitan in their distribution than believed eariler (e.g., DeLong, 1998; Vetriani *et al.*, 1999).

A third kingdom of Archaea, the Korarchaeota, was established provisionally based on rRNA sequences obtained from samples of the Obsidian Pool hot spring in Yellowstone National Park and distinct from those of other Archaea (Pace, 1996). Attempts to bring these hyperthermophiles into pure culture are underway. The discovery of this third Archaean kingdom and the recent discoveries of Archaea in cold, oxygenated habitats clearly indicate that the true diversity of

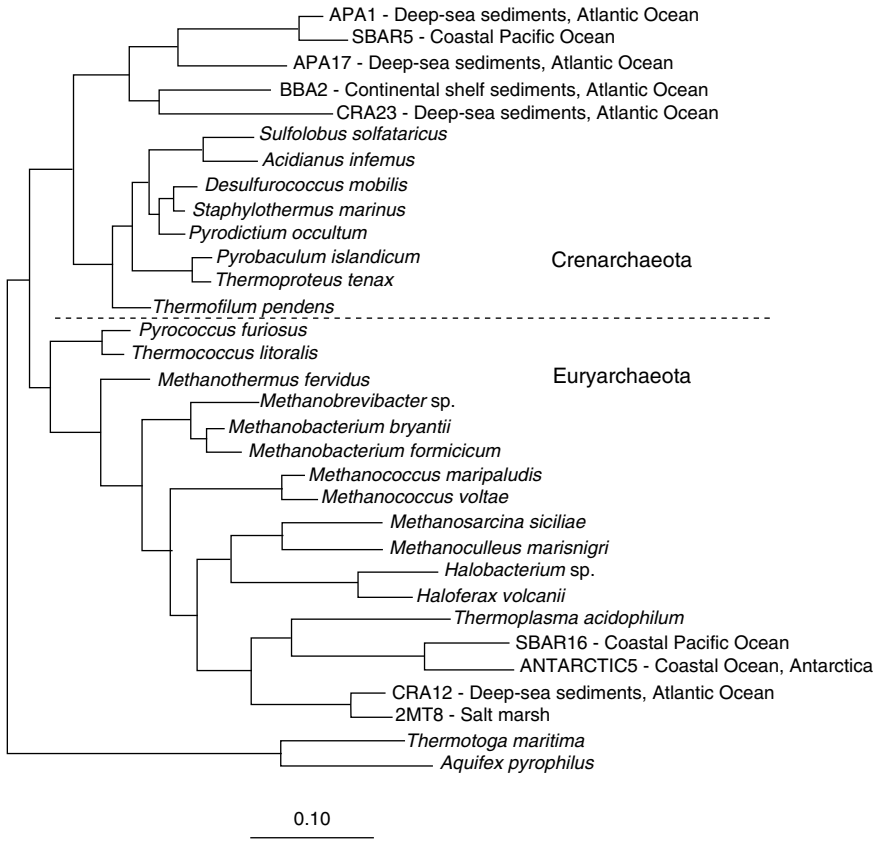


Figure 5 Diversity and phylogenetic relationships among members of the domain Archaea. This maximum likelihood tree, based on small-subunit rRNA, shows representatives of the Crenarchaeota (some extreme thermophiles and some marine environmental phylotypes) and of the Euryarchaeota (some extreme thermophiles, methanogens, halophiles, and marine environmental phylotypes). The environmental phylotypes are indicated with acronyms and numbers, with their environmental origins designated. *T. maritima* and *A. pyrophilus* form the bacterial out-group. The scale represents the expected number of changes per sequence position. See Vetriani *et al.*, 1999 for details and references. Prepared and provided by C. Vetriani, Rutgers University.

Archaea is likely to far exceed that based on presently identified species and sequences obtained from environmental samples. Implicit in this newly evolving view of Archaean phylogenetic diversity are substantially broader metabolic capabilities and wider ecological roles for this group of bacteria (DeLong, 1998; Vetriani *et al.*, 1999).

Domain Eucarya

Microbial groups in the Eucarya are the Protista and the Fungi, organisms that, in contrast to the Bacteria and the Archaea, have a membrane-bounded nucleus. Endosymbiotic events are likely to have been major driving forces in the evolution of eukaryotes. Bacteria are thought to have diverged early from a universal prokaryotic ancestor, followed by the Archaea, both of which retained the prokaryotic body plan. Fusion of an archaean and a bacterium may have led to the nuclear line, which through symbiotic acquisition of phototrophic and nonphototrophic bacteria resulted in chloroplast- and mitochondria-bearing eukaryotic cells. Loss of one or both of these organelles, or failure to acquire them initially, along with secondary symbiotic events, can be seen to account for much of the diversity of modern eukaryotes (Madigan *et al.*, 1999).

Previous groupings placed eukaryotes into four kingdoms: animals, plants, fungi, and protists, with a fifth kingdom, Monera, to contain all the bacteria. Current understanding of phylogeny indicates that this five-kingdom system greatly underemphasized the diversity of bacteria while overemphasizing animals and plants, as described earlier. The five-kingdom system also underrepresented the diversity within the protists. The current, developing view, based on a rapidly increasing database of 16 S-like rRNA sequence information, is that the Eucarya consists predominantly of unicellular microorganisms and that many phylogenetically deep, kingdom-level divisions exist within the protists (Sogin *et al.*, 1996). Diversity within the protists dominates that of the other eukaryotic lineages (**Figure 6**).

Protista

The Protista is a large complex grouping of mostly unicellular eukaryotic organisms. They are morphologically diverse and can be found in most terrestrial, aquatic, and marine habitats as free-living forms and as parasites of other protists, of fungi, and of plants and animals. With their nutritional modes

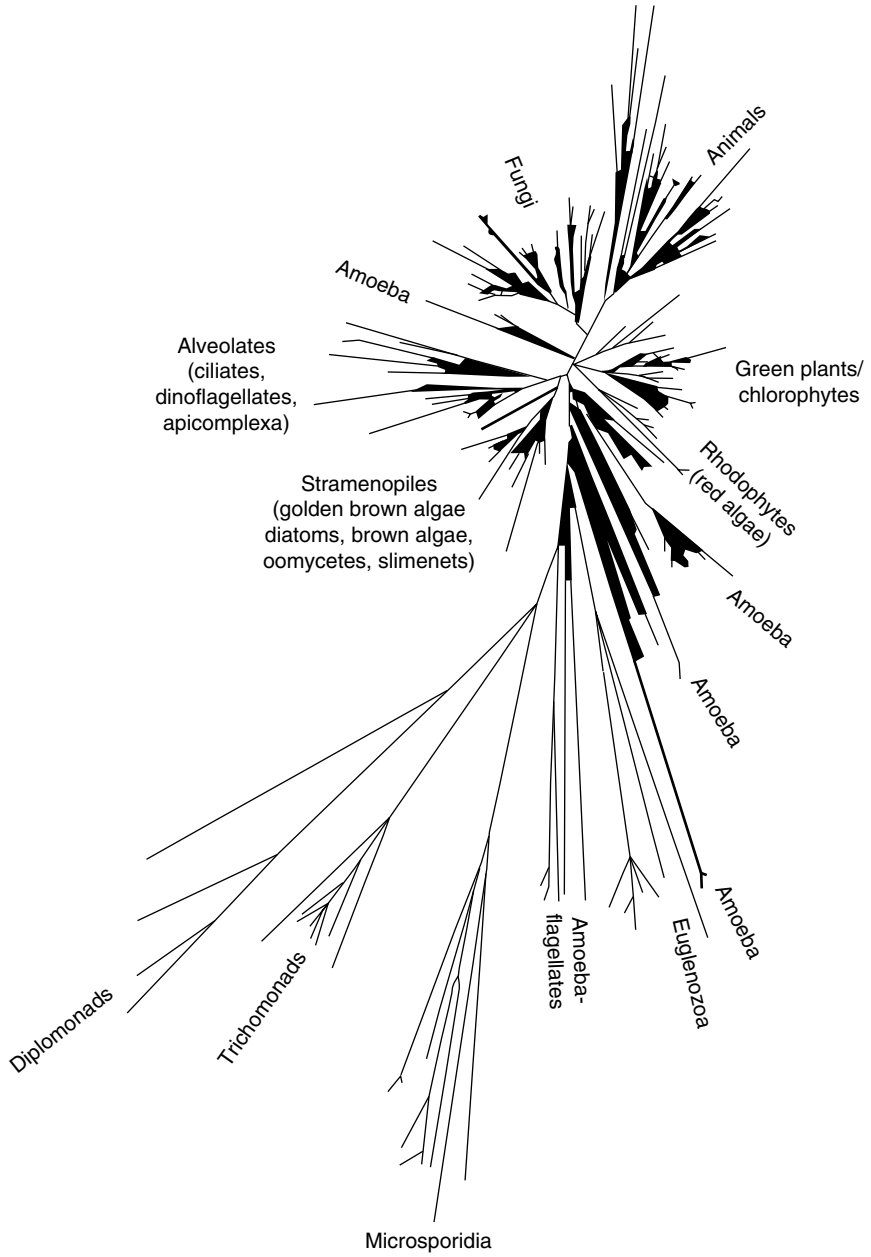


Figure 6 Diversity phylogenetic relationships among members of the domain Eucarya. This evolutionary tree reveals the dominance of the protists over the fungi, animals, and plants.

restricted primarily to osmo- and phago-heterotrophy and phototrophy, protists are metabolically much less diverse than Bacteria and Archaea. Along with various independent amoeboid groups, major groupings include the Alveolates, composed of ciliates (e.g., *Paramecium*), dinoflagellates (e.g., *Alexandrium*), and apicomplexans (e.g., *Plasmodium*), and the Stramenopiles, composed of the brown and golden-brown algae, diatoms, chrysophytes, oomycetes, and distinct groups of slime molds, among other groups. Cryptophytes, Rhodophytes, and Haptophytes are other major groupings of protists. Along with these groups are the diplomonads, trichomonads, microsporidia, amoeba-flagelates, and euglenoids (**Figure 6**).

Fungi

The fungi, sensu strictu, are commonly filamentous, multicellular heterotrophic organisms. Though previously thought to be similar to plants but lacking chlorophyll, fungi phylogenetically are not closely related to plants. Instead, fungi are seen now to have diverged from the animal lineage (Alexopoulos *et al.*, 1996). Many fungi are saprobic, feeding osmotrophically on dead organic matter, and many are also parasites or symbionts of animals. As such, they share the limited range of metabolic capability of animals. Approximately 69,000 species of fungi have been identified, and more than 1,500,000 species are estimated to exist (Hawksworth, 1991).

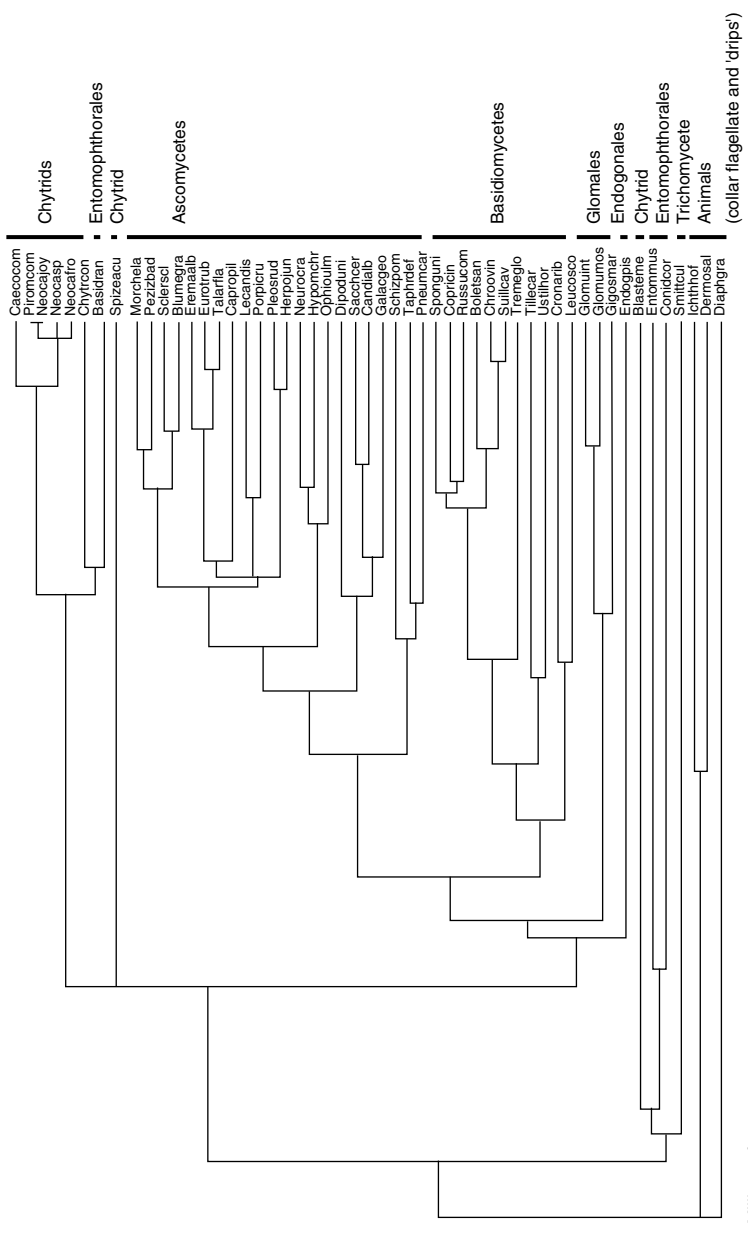


Figure 7 Diversity of fungi and the timing of their divergences. See Berbee and Taylor (1999) for details. Species names are abbreviated with the first five letters of the genus followed by the first three letters of the specific epithet for these fungi (e.g., *Dermocystidium salmonis* = Dermosal).

Four major groups (phyla) of true fungi have been defined (Alexopoulos *et al.*, 1996) (**Figure 7**). The Chytridiomycota contains a single class, Chytridiomycetes, which uniquely among fungi produces motile cells during its life cycle. The motility organelle, a typical eukaryotic flagellum, probably was retained from ancestral protists (Berbee and Taylor, 1999). Chytrids play important ecological roles in decomposing organic materials. The Zygomycota contains two classes, the Zygomycetes, which form thick-walled resting zygospores, and the Trichomycetes, obligate symbionts of arthropods. The Ascomycota, which form ascospore-carrying asci, contains most of the lichen-forming fungi. The Basidiomycota, which produces sexual basidiospores on specialized basidia, contains many of the commonly recognized fungi, such as mushrooms, puffballs, and bracket fungi. Plant pathogens in this phylum include the rust and smut fungi. Other fungi-like microbes commonly studied by mycologists now are grouped with the protists. These include the Stramenopile groups of oomycetes, hyphochytrids, and labryrinthulids, the plasmodiophorids, and the dictyostelid, plasmodial, and acrasid slime molds. The phylogenetic diversity and evolution of the fungi is an area of active research (e.g., Berbee and Taylor, 1999).

Concluding Comments

Understanding the extent and significance of microbial diversity "is the primary goal, the necessary objective, of biology in the next century" (Woese, 1999). Where progress toward that goal will take biology, however, is difficult to even guess, since the unexplored, unrecognized diversity of microbes is immense. However, the fusion of elective culture and molecular phylogeny is creating a revolution in our understanding of that diversity and of the ecological and evolutionary significance of microbes, especially bacteria. Molecular phylogeny also is overturning the previously limited view of protist diversity and is enhancing awareness of fungal phylogeny and diversity. Add to this microbial revolution, just underway in the beginning years of the 21st century, the developing influences of genomic sequencing, differential display and proteomics, and one sees a turbulent, radical, and exciting restructuring of biology underway, one that is being spearheaded by new awareness of the extent and significance of microbial diversity. Important beyond measure in the success of that restructuring will be attracting talented students to professions in microbial research. While biotechnological aspects of microbiology will continue to appeal to capable young researchers, many of the best students will be attracted not by the usefulness of microbes but by their inherent biological beauty and by the potential their study offers for insights into the evolution of life on earth.

Acknowledgements

The author is indebted to M. Sogin, N. Pace, J. Taylor, M. Berbee, K. Suberkropp, J. Breznak, T. Schmidt, J. Fuhrman, C. Vetriani, and E. Leadbetter for providing or suggesting sources of information; to J. Breznak for providing the photograph of termite hindgut microbes; to C. Vetriani for providing the figure of Archaeaphytogeny; and to N. Pace, M. Sogin, M. Berbee, and J. Taylor for permission to use published figures. Research in the author laboratory is supported by a grant from the National Science Foundation.

See also: Archaea, Origin of. Bacterial Biodiversity. Eukaryotes, Origin of. Microbial Biodiversity, Measurement of. Microorganisms (Microbes), Role of

Reference

Alexopoulos CJ, Mims CW, and Blackwell M (1996) *Introductory Mycology*, 4th ed New York: John Wiley & Sons.

Amann RI, Ludwig W, and Schleifer K-H (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev* 59: 143–169.

Berbee ML and Taylor JW (1999) Fungal molecular evolution: Gene trees and geologic time. In: McLaughlin DJ (ed.) *Mycota*, vol. 7, pp. 229–245. New York, N.Y: Springer.

Claes MF and Dunlap PV (2000) Aposymbiotic culture of the sepiolid squid *Euprymna scolopes*: Role of the symbiotic bacterium *Vibrio fischeri* in host animal growth, development and light organ morphogenesis. *J. Exp. Zool* 286: 280–296.

DeLong EF (1998) Everything in moderation: archaea as "non-extremophiles". *Curr. Opin. Genet. Dev* 8: 649–654.

Gould SJ (1996) *Full House*. New York: Three Rivers Press.

Hawksworth DL (1991) The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycolog. Res* 95: 641–655.

Hugenholtz P, Goebel BM, and Pace NR (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol* 180: 4765–4774.

Jannasch HW (1997) Small is powerful: Recollections of a microbiologist and oceanographer. *Annu. Rev. Microbiol* 51: 1–45.

Leadbetter JR, Schmidt TM, Graber JR, and Breznak JA (1999) Acetogenesis from H_2 plus CO_2 by spirochetes from termite guts. *Science* 283: 686–689.

Ludwig W and Schleifer K-H (1999) Phylogeny of *Bacteria* beyond the 16S rRNA standard. *ASM News* 65: 752–757.

Madigan MT, Martinko J, and Parker J (1999) *Brock Biology of Microorganisms*, 9th ed. Upper Saddle River, NJ: Prentice Hall.

Pace NR (1996) New perspective on the natural microbial world: Molecular microbial ecology. *ASM News* 62: 463–470.

Pace NR (1999) Microbial ecology and diversity. *ASM News* 65: 328–333.

Palleroni NJ (1994) Some reflections on bacterial diversity. *ASM News* 60: 537–540.

Sogin ML, Morrison HG, Hinkle G, and Silberman JD (1996) Ancestral relationships of the major eukaryotic lineages. *Microbiologia Sem* 12: 17–28.

Staley JT, Castenholtz RW, Colwell RR, *et al.* (1997) *The Microbial World: Foundation of the Biosphere: A Report of the American Academy of Microbiology*. Washington, DC: American Society for Microbiology.

van Niel CB (1949) The "Delft School" and the rise of general microbiology. *Bacteriol. Rev* 13: 161–174.

Vetriani C, Jannasch HW, MacGregor BJ, Stahl DA, and Reysenbach A-L (1999) Population structure and phylogenetic characterization of marine benthic Archaea in deep-sea sediments. *Appl. Environ. Microbiol* 65: 4375–4384.

Whitman WB, Coleman DC, and Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci. USA* 95: 6578–6583.

Wilson EO (1994) *Naturalist*. Washington, DC: Island Press.

Woese CR (1987) Bacterial evolution. *Microbiol. Rev.* 51: 221–272.

Woese C (1999) The quest for Darwin's grail. *ASM News* 65: 260–263.

Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program

Kenneth H Nealson, University of Southern California, Los Angeles, CA, USA
Shibu Yooseph, J. Craig Venter Institute, San Diego, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Fragment recruitment A computational technique in which reads from a metagenomic data set are mapped to a reference genome. This information allows for assessing the abundance and representation of an organism (given by its genomic sequence) in a microbial community (represented by the metagenomic data).

GOS The Global Ocean Sampling program launched by the J. Craig Venter Institute to study microbial diversity in the world's oceans using metagenomics.

Metagenomic assembly The process of inferring genomic sequences of taxa represented in a metagenomic data set. Reads that belong to the same genomic region are identified using read overlap information, and joined to produce longer contiguous sequences called contigs. When read mate-pair information is available, contigs can be further ordered and oriented to produce scaffolds. An assembly is a collection of contigs and/or scaffolds.

Metagenomics Literally "beyond genomics," this refers to the process of sampling the entire genetic complement of an environment via DNA isolation, sequencing, annotation, and informatic analysis.

MMGSP The Gordon and Betty Moore Foundation funded Marine Microbial Genome Sequencing project (MMGSP) was an international collaborative initiative whose goal was to sequence genomes of ecologically relevant microbes from the marine environment.

Picoplankton The fraction of aquatic biomass in the size fraction that passes through 3.0 and 0.8 μm filters, and is retained on a 0.1 μm filter.

Taxon A term describing a collection of organisms that are sufficiently closely related to be considered to be in the same taxonomic group (species, genus, etc.).

A Historical Perspective: The Need for Genomics and Metagenomics

From the early years of marine microbiology to the late 1970s, the microbial communities of the ocean were easily explained: the dominant groups were familiar to most microbiologists, and included heterotrophs such as *Vibrio*, *Pseudomonas*, *Alteromonas*, and *Photobacterium*, to name a few, as well as other geochemically active aerobes (e.g., familiar species of ammonia and nitrite oxidizers such as *Nitrobacter*; *Nitrosobacter*; sulfur oxidizers, methane oxidizers such as *Methanobacter*), and an impressive array of anaerobic taxa including methanogens and sulfate reducers (Kriss, 1963; Sieburth, 1979; ZoBell, 1946). Much of marine microbiology thus focused on the mechanism(s) and roles of these "key species" in processes such as bioluminescence, symbiosis, fish disease, elemental cycling of carbon, nitrogen, and sulfur, and the production or consumption of important nutrients in the ocean (Sieburth, 1979). At the level of primary productivity, the importance and abundance of the cyanobacterial genera (*Synechococcus* and *Prochlorococcus*) was not yet appreciated, and the view of oceanic carbon cycling was dominated by microalgae, and a few large filamentous cyanobacteria such as *Trichodesmium* (Pace, 1997; Sieburth, 1979).

In the ensuing decades, three key events occurred that forever changed the intellectual landscape of marine microbial diversity and ecology. The first was the "great plate count anomaly" (Staley and Konopka, 1985) and the accompanying realization that most of the microbes seen in the microscope did not form colonies on standard growth media (see Sieburth (1979) and references therein), thus leading to the concept of the uncultivated or "unseen" majority (Whitman *et al.*, 1998). It became clear at this juncture that the "dominant species" represented neither marine microbial diversity nor marine microbial ecology in the broader sense. At this point, it seemed possible that the nongrowing microbes were simply uncultivable members of the "dominant taxa," but since taxa were identifiable only by growth-related criteria, this could not be established.

Second, through the work of Carl Woese and his colleagues (Olsen *et al.*, 1986; Pace *et al.*, 1986; Pace, 1997; Pace *et al.*, 1986; Woese *et al.*, 1990), it became clear that the 16S rRNA gene sequences could be used to identify major groups of microbes, including many that remained uncultivated. Furthermore, through comparative sequence analysis, it became clear that the entire microbial world could be related via sequence similarity, using 16S rRNA gene sequences (Pace, 1997; Pace *et al.*, 1986; Woese *et al.*, 1990). Thus major clades of microbes were seen that were represented only by 16S rRNA gene sequences: i.e., they had no cultivated members. Examples of such organisms were SAR11 (now in cultivation as *Pelagibacter ubique* (Giovannoni *et al.*, 1990)) and SAR 86 (an abundant lineage that has still resisted all attempts at cultivation (Giovannoni *et al.*, 1990)). Molecular analyses of seawater soon revealed the presence of large numbers of some of these clades in seawater, particularly SAR11, and two groups of cyanobacteria – *Synechococcus* and *Prochlorococcus* (Rocap *et al.*, 2002; Worden *et al.*, 2000). With regard to microbial diversity, the notion that the "unseen majority" were simply nongrowing members of more familiar clades could now be dispensed with.

Third, through the work spearheaded by Waterbury (i.e., the cultivation of phototrophic microbes in the groups

Synechococcus and *Prochlorococcus*) (Rocap *et al.*, 2002; Waterbury *et al.*, 1979; Waterbury and Willey, 1988) and Giovannoni (i.e., the cultivation of SAR11 and several other *P. ubique* strains (Rappe *et al.*, 2002; Stingl *et al.*, 2007)), it became possible to do comparative physiology and to integrate the physiology with genomics of these newly cultivated groups.

With the advent of metagenomics (Handelsman, 2004), it became possible to study marine microbial communities using cultivation independent approaches. In a typical metagenomic study of an environment, nucleic acids are extracted from the collected sample and directly sequenced using a random shotgun approach, thus eliminating the need for *in vitro* culture.

Discovery Science: The Global Ocean Sampling (GOS) Program

The GOS Idea

The GOS program was launched by the J. Craig Venter Institute (JCVI) to study microbial diversity in the world's oceans using metagenomics. The program started with a study of samples collected from the Sargasso Sea (Venter *et al.*, 2004), and was followed by sampling of surface seawater from sites around the globe (**Figure 1**). At the very least, a picture of the taxonomic composition of microbes, both known and unknown, could be expected from this study: i.e., a metagenomic glimpse of surface ocean biodiversity could be expected. More optimistically one might hope for insights into the general principles that guide adaptation and evolution of marine surface picoplankton. Seawater samples were collected (Rusch *et al.*, 2007) by pumping water through a 20-μm Nitex® prefilter, and then size fractionating the biological material by serial filtration through 3.0, 0.8, and 0.1 μm membrane filters to obtain three size fractions; the final filtrate was also harvested to analyze for marine viruses. The collected filters were frozen, stored on dry ice, and shipped to the JCVI for sequencing as soon as possible. To this point in time the analysis of these samples has focused primarily on the microbes that pass through the 0.8 μm filter, and are retained on the 0.1 μm filter – commonly referred to as the picoplanktonic size fraction, or the picoplankton. Thus, any coccoid organisms smaller than 0.8 μm in diameter, and any rod-shaped organism with a diameter of 0.8 μm or less would have a good chance of passing through the filter. Accordingly, larger organisms, aggregated cells, and particle associated cells should all be retained on the larger filter(s).

The Sargasso Sea: A ''Low Diversity'' Site

The GOS began with a series of samples from the Sargasso Sea, chosen specifically because of its reported low biomass properties, and the expectation that it would also display low biodiversity. This expedition was regarded as a "test run" to develop and trouble-shoot the methods of sample collection, sequencing and analysis pipelines, and annotation and assembly methods. How would this approach compare to estimates of biodiversity obtained by other methods? (Giovannoni *et al.*, 1990) The Sargasso Sea waters have on the order of 10^5 cells per ml of seawater, about a factor of 10 less than most marine

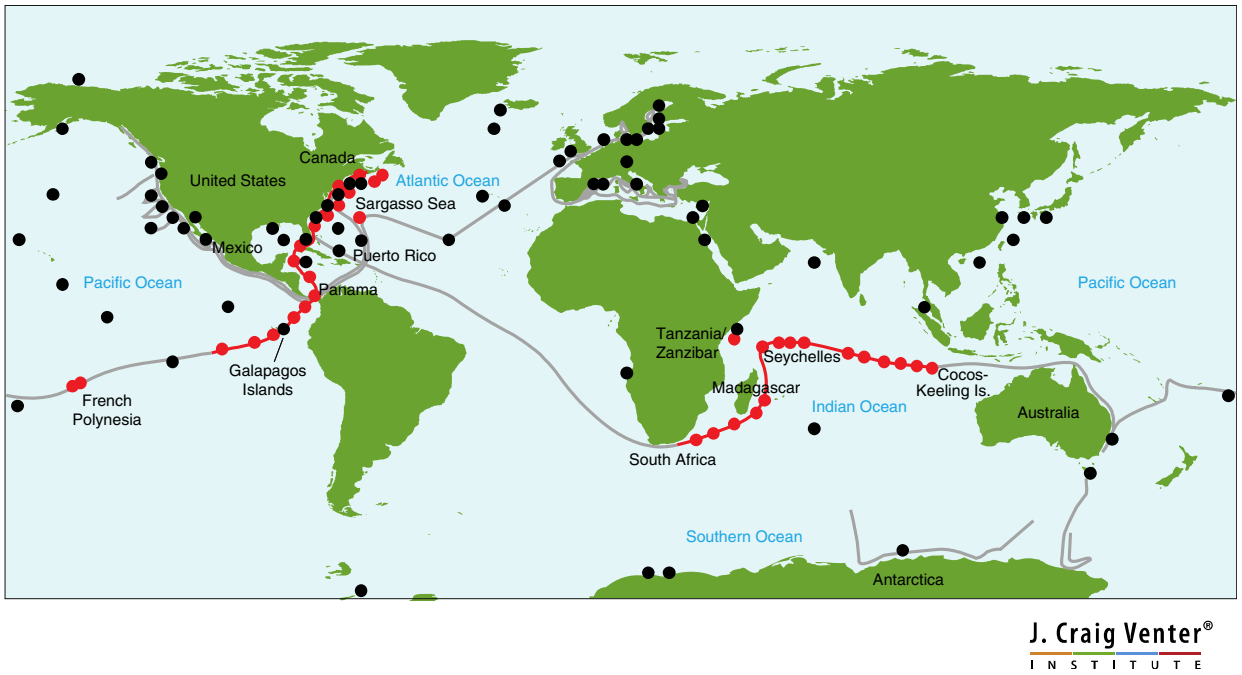


Figure 1 Map showing the GOS route together with locations where marine isolates, sequenced as part of the Marine Microbial Genome Sequencing project, were collected. The red dots are the GOS sample sites that have been published to-date and the black dots are the marine isolates collection locations.

waters (Giovannoni *et al.*, 1990), which are usually at a density closer to 10^6 per ml, and the expectation was that these nutrient-limited conditions would yield a low picoplankton diversity.

While the Sargasso Sea surface samples were less diverse than soil or other complex environments, a substantial number of taxa were seen and it was far more diverse than had been expected. All in all, about 1800 species were estimated to be present based on the sequence relatedness, including nearly 150 previously unknown bacterial phylotypes. In addition, more than 1 million previously unknown genes were seen, including about 800 new bacteriorhodopsin genes (Venter *et al.*, 2004).

Several notable and unusual results from the Sargasso Sea work speak to both the values as well as the hidden traps of the metagenomic approach. First, and perhaps most notable, abundant *Shewanella* and *Burkholderia* were seen in one sample: abundant enough that nearly complete genomes were assembled from the shotgun sequencing data (Venter *et al.*, 2004). While it was not known at the time, these genera would never be seen in abundance again in any GOS sample. The reason for these results remains unresolved. The most convenient, and not unreasonable, explanation is that the sampling system was contaminated (DeLong, 2005). A more cautious view might be that it is due to a stochastic natural event (i.e., "natural contamination"). Given that members of both of these genera are known symbiotic inhabitants of the gut tracts of a wide number of different organisms, including marine and freshwater fish (Kikuchi *et al.*, 2007; Olivier-Espejel *et al.*, 2011; Tanu *et al.*, 2011; Ward *et al.*, 2009), it is also possible that the Sargasso Sea, with its low concentrations of bacteria might be easily overwhelmed by the passing excretion of gut contents into the sampling area by a marine fish: not an uncommon process in the marine environment (Ruby and Nealson, 1978). Finally, one might suppose that this could be the result of the flushing of the waste disposal system on the sampling boat, but this seems unlikely, as no enteric or other commonly encountered human-associated bacteria were seen. For the moment, and probably forever, these results remain enigmatic, and their importance has faded as subsequent sampling has moved them to the "rare and stochastic event" status. What remains important is the observation that several groups of bacteria were found in relative abundance, including *Pelagibacter* and *Prochlorococcus* types, along with several other groups known only by their 16S rRNA sequences. This was the first hint of a pattern that would repeat itself at nearly every marine surface site analyzed by the GOS expedition.

Of equal importance in the Sargasso Sea analyses were things that did not happen. Despite the abundance of *Synechococcus*, *Prochlorococcus*, and *Pelagibacter* types, based on the abundance of 16S rRNA gene sequences, in no case was it possible to assemble even a reasonably long scaffold from any of these three. That is, while the 16S rRNA data seemed to be saying that one species of each was present in abundance, the assembly data were saying that this species was sufficiently diverse that assembly into scaffolds was literally impossible. This was the first example of what was to be many such data sets to be seen for these three groups of microbes.

Finally, one very obvious potential limitation of this approach was noted in the Sargasso Sea work when combined with other genomic data. *Pelagibacter ubique* is according to 16S rRNA taxonomy an alpha proteobacterial phylotype, but contains a genome that is only about 1.2 Mbp, only a fraction of the size of most cultivated alpha proteobacteria (Giovannoni *et al.*, 2005). Thus, the approach to understanding the picoplankton via 16S rRNA is limited in its usefulness in this environment. Until one has an organism in cultivation, or its genome sequenced and sized, there is the distinct possibility that the 16S rRNA gene data provide limited information about the capacity and abilities of the particular taxon. Similar genome streamlining has been shown for two other bacterial groups, the *Flavobacteria* (Woyke *et al.*, 2009) and a gamma *Proteobacteria* group called SAR86 (Dupont *et al.*, 2012) for which both genomes were obtained via the sequencing of single cells.

One of the grand hopes of the GOS program was that function might be linked to identification of new groups of organisms, but in general, with assemblies being so limited, this was not achieved in the initial study. One notable exception to this was the identification of an archaeal 16S rRNA sequence connected via a scaffold to a gene that was similar to an ammonia oxidase (Venter *et al.*, 2004). This microbe has subsequently obtained in pure culture (Konneke *et al.*, 2005) and shown to be widely distributed in many marine environments (Francis *et al.*, 2005). Now called *Nitrosopumilis*, this group is known to be abundant in nearly all GOS samples, and in fact in many other environments around the world, as a major player in ammonia oxidation in many different environments. While the archaea in general are neither dominant nor diverse in the surface oceans, this particular group is both widespread and abundant.

Biodiversity Across Several Biomes and the PLoS Results:

The first attempts at "Global" data analyses were published in 2007, with a compilation of results beginning off the coast of Nova Scotia, and proceeding down the US coast, through the Panama Canal, to the Galapagos Islands, and ending in the equatorial Pacific. All in all, this data set consisted of 44 samples from 41 sites, including a few freshwater and hypersaline sites. The magnitude of this data set was such that patterns began to appear that could not be seen in the Sargasso Sea data (which were also included in this study) (Rusch *et al.*, 2007). For example the appearance of occasional "rare and stochastic events" was again seen, though with a resident population of microbes of 10^6 per ml, the sites were less susceptible to such events. Nevertheless, as a general rule, as more sites were examined, the "occasional" (often singular) appearance of a taxon as a dominant member was still seen, while the nearly continuous appearance of the three major players, and a number of other numerically dominant taxa became a general rule. A site without these phylotypes was unusual indeed.

Fragment Recruitment: Where do the GOS Sequences Come From?

Fragment recruitment (Rusch *et al.*, 2007) is a method whereby an entire genome can be "hybridized" in silico to the

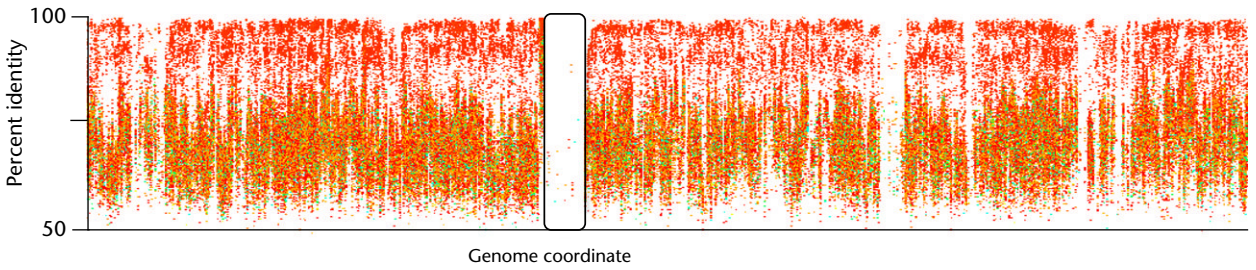


Figure 2 Recruitment of GOS reads to a reference genome. The x-axis denotes the coordinates on the genome (in this case, *Candidatus Pelagibacter ubique* HTCC1062) and the y-axis denotes the percent identity of the match. Reads are colored by GOS samples that they belong to. Region in the box corresponds to a portion of the reference genome that does not recruit any GOS data.

GOS data, and the quantitative contribution of that genome to the surface picoplankton gene pool can be assessed (**Figure 2**). In the initial analysis via fragment recruitment, it was seen that only three genome types provided any significant fragment recruitment, even at the 50% identity level, and these were *Prochlorococcus*, *Synechococcus*, and *Pelagibacter*.

These three groups accounted for only about 15% of the total sequences, leaving a very large proportion of the GOS sequences unaccounted for by any of the genomes of cultivated and sequenced microbes that were tested (Rusch *et al.*, 2007). Thus, it appeared that the oceanic picoplankton were composed of a mixture of (mostly) unknown organisms, with very little genome similarity to any of the cultivated microbes. This overall view of biodiversity has become more solidified as the GOS data have been analyzed, and several hints have been provided by the data set:

First, several taxa were present at nearly every site, suggesting that there are a number of marine taxa, mostly uncultivated, that are very well adapted to life in the surface ocean. While they remain uncultivated, the identification of these "true picoplankters" is now slowly emerging through the abundance studies of the 16S rRNA gene content.

Second, in silico hybridization (fragment recruitment) of the GOS data set (i.e., the complete DNA sequence data with the "stochastic event" DNA samples removed) revealed no significant hybridization by DNA from any cultivated organisms other than *Synechococcus*, *Prochlorococcus*, or *Pelagibacter*. This rather remarkable result suggested that indeed, the picoplankton did not consist of microbes with which we were familiar, and that we really knew very little about the planktonic niche in terms of its taxonomic and functional biodiversity.

Third, as in the Sargasso Sea, assembly of the metagenomes from these environments generated only short contigs and scaffolds (with a majority of reads remaining unassembled), implying again, that there was substantial intraspecies diversity among these three major players. Is this again a property unique to these three genera, or will it be a common mode of adaptation in the marine surface ocean environment?

Functional Diversity of the GOS Sequences

The contribution of the published GOS data to our understanding of protein function and evolution has been significant. More than 6 million proteins were reported in the published GOS study (Yooseph *et al.*, 2007) – in the process more than doubling the number of known proteins. An analysis of these proteins revealed thousands of novel protein families, and also showed that we have a long way to go in cataloging the protein universe, with protein family diversity being driven by a significant contribution from the microbial world. Many of the novel proteins appear to be of viral origin – a likely consequence of under-sampling of viral diversity. An analysis of protein families including RuBisCO, glutamine synthetase, phosphatases, proteases, and kinases, revealed novel subfamilies and substantial sequence diversity added by the GOS data. These data have added insights into the functional capabilities necessary for survival in the nutrient-limited open ocean environment.

Taxonomic and Functional Biodiversity of the Surface Ocean Picoplankton

The Marine Microbial Genome Sequencing Project (MMGSP)

The easiest explanation for the lack of fragment recruitment noted in the GOS study discussed above (*see* Fragment Recruitment: Where do the GOS Sequences Come From?) might be that the microbial genomes used for the hybridization were primarily nonmarine. In fact many of the genomes sampled were either from extremophilic archaea or medically important bacteria, neither of which might be expected to have gene complements abundant in the surface ocean. The Gordon and Betty Moore Foundation funded MMGSP was an international collaborative initiative whose goal was to sequence genomes of ecologically relevant microbes from the marine environment (**Figure 1**). As part of this project, we sequenced and annotated 137 genomes, and analyzed these together with genomes of previously published marine prokaryotes. Our collection included several strains of the "big three" (*Synechococcus*, *Prochlorococcus*, and *Pelagibacter*), as well as microbes from a wide variety of marine biomes.

Fragment recruitment was enlarged to include GOS data from a number of other sites, including the Indian Ocean. Remarkably, when the fragment recruitment exercise was finished, the results remained virtually unchanged. The top recruiters were still the "big three," and virtually no increase in recruitment to other groups was seen with the addition of the MMGSP genomes (Yooseph *et al.*, 2010).

The results from genomic, metagenomic, and 16S rRNA analyses (Yooseph *et al.*, 2010) lend credence to our view of the biodiversity of marine surface picoplankton: there appear

to be a few (i.e., 10–20) dominant and cosmopolitan taxa that represent the true picoplankton in the oceanic surface waters, and of these, only three have been cultivated. And for most of the others, reference genomes are not yet available. Thus we remain "in the dark" with regard to the biodiversity of this group, which we refer to as the "true picoplankton."

Looking at the properties of the GOS data set, some predictions of what the true picoplanktonic population is composed of can be made, and the conclusions reached by this approach provide a picture that is somewhat different from the traditional view: The picoplankton is a group of very small microbes (with reduced genome size) that have given up many of their complex regulatory abilities, and have adapted to the marine surface niche by staying small, growing slowly, and being incapable of adapting to nutrient additions by increasing their growth rate and blooming (Yooseph *et al.*, 2010). These are nonmotile and nonchemotactic aerobes, with major reductions in genome size and metabolic abilities.

Some recent results support this view of the oceanic picoplankton. Woyke *et al.* (2009) recently used single-cell genomics methods to sequence and assemble the genomes of two *Flavobacteria* cells (MS024–2 A and MS024–3C). These genomes showed recruitment to the GOS data supporting the notion that they are cosmopolitan and numerically abundant members of the marine picoplankton. Both genomes showed substantial genome streamlining, with genomes of these two being only 1.9 and 1.5 Mbp, respectively, much less than the 6.1 and 2.9 Mbp genomes of two cultivated *Flavobacteria* isolates (Woyke *et al.*, 2009). In addition, analysis of the gene content suggested that they were adapted to a narrow (focused) metabolic niche: one that might include algal interactions, perhaps as attached forms.

In another study utilizing single-cell genomics and metagenomic assemblies, Dupont *et al.* (in press) showed similar results for four assembled genomes of SAR86-like microbes. Fragment recruitment was significantly higher than seen with the *Flavobacteria* genomes obtained using single-cell genomics, and, as with the *Pelagibacter* genomes, when the required identity level was relaxed from 90% to 50%, fragment recruitment increased dramatically, suggesting to the authors that there is a broad diversity within this group that is yet to be discovered (Dupont *et al.*, in press). The two of the four genomes that are nearly complete (SAR86A and SAR86B) have genomes of 1.25 and 1.7 Mbp, respectively, considerably smaller than other gamma *Proteobacteria* that usually sport genomes of about 5 Mbp. Metabolic mapping of the genome indicates that SAR86 types are aerobes that may be focused on the degradation of lipids and carbohydrates acquired by Ton-B dependent outer membrane receptors. As with both the *Flavobacteria* obtained using single-cell genomics, and *Pelagibacter*, this heterotroph appears to have a narrow metabolic set of preferences, thus avoiding competition via specialization for unique resources.

The GOS approach has allowed the formulation of a particular view of the world's surface picoplanktonic population, which the authors will discuss below. It includes almost exclusively organisms that, until recently, were not known, and many of which are still not cultivated (and many that are not yet sequenced). Based on 16S rRNA analysis and fragment recruitment results, one can speculate that another 10–20 microbial groups of microbes will be required to have a good picture of the diversity of the abundant and cosmopolitan members of the surface picoplankton.

While we now have some sense of the functional adaptation of the picoplankters, the situation remains at best incomplete. With genomes of two photoautotrophs (*Synechococcus* and *Prochlorococcus*) and three chemoheterotrophs (*P. ubique*, *Flavobacteria spp.*, *and* SAR86-like), we can assemble some traits that appear to be common among them with the expectation that this situation will change as more members of the picoplankton are characterized. These traits include: (1) genome reduction; (2) low GC content; (3) lack of or decrease in many functional and regulatory genes, including those involved with transcriptional regulation, energy linked transport systems, efflux systems for toxic metals and drugs, chemotaxis, motility, and quorum sensing, protein secretion, siderophore synthesis and uptake, and anaerobic metabolism. In addition, CRISPRS are notably absent or reduced.

This is a far different picture of the marine picoplankton than anything we imagined before GOS. In this nutrient limited world, the plankters are well adapted, with very little taxonomic diversity, but substantial intrataxon diversity – diversity that challenges the assembly of genomes of these key species.

Given the reduced genome sizes of the picoplankters, the cells are, as expected, quite small. Thus, while they constitute the preponderance of 16S rRNA gene sequences (and thus genomes), they may have far less impact at the level of biomass. For example, some strains of *P. ubique* have cells that are only 1/1000 the biomass of an *E. coli* sized microbe (Giovannoni *et al.*, 2005), so if they are 10^5 ml^{-1} and a *Vibrio* species is at 10^2 ml^{-1}, the relative amount of biomass is approximately the same! Without very deep sequencing, the *Vibrio* cells will not be seen by our shotgun sequencing methods. If one imagines that conditions change to favor growth of the *Vibrios*, then their ability to bloom in response to changes in available nutrients could allow them to rapidly become the dominant biomass in a given oceanic niche, as has been hypothesized and observed by many workers (Hunt *et al.*, 2008; Preheim *et al.*, 2011; Thompson *et al.*, 2005). Given the GOS data set, we conclude that this has not happened in any of our sites, but believe that such phenomena are an integral part of marine biology, and must be part of the ecological strategies of the microbes that are in most of our culture collections. One example of this is the report of a bioluminescent zone (milky sea) off the coast of Somalia, proposed to be a bloom of luminescent *Vibrios* at a concentration of 10^7 ml^{-1} or more (Nealson and Hastings, 2006). Such a bloom could be the natural consequence of the crash of an algal bloom, leading to abundant surface nutrients, and a dominant bloom of light-emitting microbes. Similar nutrient enrichment(s) could stimulate growth of the more familiar cultivated microbes via providing niches such as marine snow (Grossart *et al.*, 2006), stomachs of marine invertebrates and vertebrates (Roeselers *et al.*, 2011; Ruby and Nealson, 1978; Tanu *et al.*, 2011; Ward *et al.*, 2009; Wu *et al.*, 2010), or niches associated with symbiosis or parasitism (Cai *et al.*, 2006). This view is in concert with a recent presentation of marine microbial diversity by Lauro *et al.* (2009) in which it was proposed that the marine picoplankton are essentially a continuum of metabolic types, ranging from ultra-oligotrophs to

"bloom and bust" coprophytes. This is a view that fits well with our own view of biodiversity, although our view is that the true picoplankters (based on abundance and continuity of appearance) do not include the coprophytes, which come and go, but were rarely seen in the GOS sampling program.

Of importance in this discussion is the fact that the GOS data have to this point, examined only those microbes on the smallest filter size – the picoplanktonic populations. Particle associated microbes and large cells may well constitute a major part of the marine biomass, but we maintain that they are a reflection of other niches, and not a part of the "true picoplankton."

The "Faces of Biodiversity"

Putting all of this together, then, the biodiversity of the surface ocean takes on a number of faces, some of which are not yet clearly articulated. The first is that of the true picoplankton, as we call them: the small, reduced genome, reduced function, low biomass microbes that seem very well suited to the nutrient limited surface ocean environment. While we have a sense of the taxonomic diversity of these microbes, we have only a beginning knowledge of their functional diversity. A second area of the unknown is the extent of intragroup diversity, and whether or not such diversity will be also seen in the other true picoplankton.

The second "face" of biodiversity is that of the more familiar cultivated strains of which we now have more than 200 genomes sequenced and in hand. These microbes, for the most part look familiar to us, and do not provide the intragroup diversity seen in the true picoplankton. However, the extent of their functional diversity remains unknown, and their true niches within the marine ecosystem(s) remains mysterious. The niches inhabited by marine luminous *Vibrios* include light organs and stomachs of fishes and squids, marine snow, and particulate matter, fecal pellets, and perhaps more (Nealson and Hastings, 2006), and with the exception of the fish and squid symbioses, whether or not any or all of these environments are taxon-specific is not known. One could easily hypothesize that such niches provide a need for intragroup variation of these types that could be as extreme as that seen in the true picoplankton. The difference, of course, is that microbes in this group are capable of rapid growth and the maintenance of nearly clonal populations, something that would seem difficult among the true picoplankton.

We close this discussion of picoplankton diversity with a brief discussion of a recent paper by Konstantinidis *et al.* (2009) in which populations of deep sea microbes were analyzed for taxonomic and functional diversity. The environment under study was the deep water off Hawaii, where the surface waters look remarkably like the GOS sites discussed. However, the deep waters are substantially different – in fact, they resemble an environment that is populated by species of the type consistent with our "second face" – microbes that we expect to be able to grow, and which might be expected to be found in more nutrient-rich niches. In contrast, this environment is notable because it is dark, nutrient poor, and presumably so depauperate that the surface ocean true picoplankton are not able to survive here!

It is our thesis that the diversity of this deep ocean site is due to the bacteria-laden particles arriving at the site, along with the oceanic eukaryotes, which may be hosts for such microbes, and contribute to their small but continuous population(s) in the deep sea. This view places the eukaryotes squarely in the center of prokaryotic evolution, a place that is not unreasonable, but one that is also only a hypothesis at this time – one that deserves attention and testing.

Conclusions

The GOS program, begun in 2003 and still in progress, has provided a view of the surface ocean microbial communities, with samples collected by the *Sorcerer II* Expedition and Collaborator Cruises from more than 200 sites from around the world. The emerging view of surface ocean microbiology is in harmony with other metagenomic studies of the ocean, and, while not the last word on picoplankton biodiversity, paints a picture of the oceanic plankton that is notably different from the paradigm that was accepted only a few years ago. A few "lessons" serve to highlight the bulk of these differences: (1) the taxonomic diversity of the surface ocean is remarkably low, with just a few (probably 20 or less) numerically dominant taxa; (2) while only three of these taxa have members currently in cultivation, for these three, the intragroup diversity is extremely high, leaving us with a niche that is characterized by a few dominant taxa that are within them, extremely diverse at the genome level; (3) the oceanic DNA sequences reveal a collection of genes that have no strong genomic sequence relationship to the gene content of many cultivated and sequenced bacteria and archaea; and (4) of the few examples of the dominant and cosmopolitan picoplankters that have been sequenced, all appear to have undergone extreme genome reduction, so that while they can be placed in familiar taxonomic clades, their genomic content renders them far different from their phylogenetic brothers and sisters.

However, to date, only a few marine picoplankters have been sequenced, and as these groups are identified and characterized, our appreciation of the metabolic diversity of the surface picoplankton will continue to grow. With regard to this, a few of the abundant and cosmopolitan picoplankters are now being analyzed using a combination of metagenomics and single-cell genomics. The bottom line of these analyses, for the present, suggests that the surface oceanic picoplankton are a group of microbes from different taxa, which are highly adapted to the nutrient limited conditions of the ocean. In contrast, most of the cultivated and sequenced counterparts are capable of "boom and bust" ecology, and occasionally are found in high numbers (as blooms), but were rarely found in abundance in the GOS sites examined. This suggests that the ecology of these readily cultivated forms is distinctly different from the abundant and cosmopolitan picoplankters whose genes dominate the GOS data set.

References

Cai J, Chen H, Thompson KD, and Li C (2006) Isolation and identification of *Shewanella alga* and its pathogenic effects on post-larvae of abalone *Haliotis diversicolor supertexta*. *Journal of Fish Diseases* 29: 505–508.

DeLong EF (2005) Microbial community genomics in the ocean. *Nature Reviews Microbiology* 3(6): 459–469.

Dupont CL, Rusch DB, Yooseph S, *et al.* (2012) Genomic insights into SAR86, an abundant and uncultivated marine bacterial lineage. *The ISME Journal* 6(6): 1186–1199.

Francis CA, Roberts KJ, Beman JM, Santoro AE, and Oakley BB (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proceedings of the National Academy of Sciences of the United States of America* 102(41): 14683–14688.

Giovannoni SJ, Britschgi TB, Moyer CL, and Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* 345(6270): 60–63.

Giovannoni SJ, Tripp HJ, Givan S, *et al.* (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309(5738): 1242–1245.

Grossart HP, Czub G, and Simon M (2006) Algae-bacteria interactions and their effects on aggregation and organic matter flux in the sea. *Environmental Microbiology* 8(6): 1074–1084.

Handelsman J (2004) Metagenomics: Application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews* 68(4): 669–685.

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, and Polz MF (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320(5879): 1081–1085.

Kikuchi Y, Hosokawa T, and Fukatsu T (2007) Insect-microbe mutualism without vertical transmission: A stinkbug acquires a beneficial gut symbiont from the environment every generation. *Applied and Environmental Microbiology* 73: 4308–4316.

Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, and Stahl DA (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437(7058): 543–546.

Konstantinidis KT, Braff J, Karl DM, and DeLong EF (2009) Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 m at station ALOHA in the North Pacific subtropical gyre. *Applied and Environmental Microbiology* 75(16): 5345–5355.

Kriss AE (1963) *Marine Microbiology.* New York: John Wiley & Sons.

Lauro FM, McDougald D, Thomas T, *et al.* (2009) The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 106(37): 15527–15533.

Nealson KH and Hastings JW (2006) Quorum sensing on a global scale: Massive numbers of bioluminescent bacteria make milky seas. *Applied and Environmental Microbiology* 72(4): 2295–2297.

Olivier-Espejel S, Sabree ZL, Noge K, and Becerra JX (2011) Gut microbiota in nymph and adults of the giant mesquite bug (*Thasus neocalifornicus*) is dominated by *Burkholderia* acquired *de novo* every generation. *Environmental Entolmology* 40: 1102–1110.

Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, and Stahl DA (1986) Microbial ecology and evolution: A ribosomal RNA approach. *Annual Review of Microbiology* 40: 337–365.

Pace B, Stahl DA, Lane DJ, and Olsen GJ (1986) The analysis of natural microbial-populations by ribosomal-RNA sequences. *Advances in Microbial Ecology* 9: 1–55.

Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* 276(5313): 734–740.

Pace NR, Olsen GJ, and Woese CR (1986) Ribosomal RNA phylogeny and the primary lines of evolutionary descent. *Cell* 45(3): 325–326.

Preheim SP, Boucher Y, Wildschutte H, *et al.* (2011) Metapopulation structure of *Vibrionaceae* among coastal marine invertebrates. *Environmental Microbiology* 13: 265–275.

Rappe MS, Connon SA, Vergin KL, and Giovannoni SJ (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418(6898): 630–633.

Rocap G, Distel DL, Waterbury JB, and Chisholm SW (2002) Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S–23S ribosomal DNA internal transcribed spacer sequences. *Applied Environmental Microbiology* 68(3): 1180–1191.

Roeselers G, Mittge EK, Stephens WZ, *et al.* (2011) Evidence for a core gut microbiota in the zebrafish. *ISME Journal* 5(10): 1595–1608.

Ruby EG and Nealson KH (1978) Seasonal changes in the species composition of luminous bacteria in nearshore seawater. *Limnology and Oceanography* 23: 530–533.

Rusch DB, Halpern AL, Sutton G, *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biology* 5(3): e77.

Sieburth JM (1979) *Sea Microbes.* New York: Oxford University Press.

Staley JT and Konopka A (1985) Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology* 39: 321–346.

Stingl U, Tripp HJ, and Giovannoni SJ (2007) Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *ISME Journal* 1(4): 361–371.

Tanu DDD, Khandeparker R, Sreepada RA, Sanaye SV, and Pawar HB (2011) A study on bacteria associated with the intestinal tract of farmed yellow seahorse, *Hippocampus kuda*: Characterization and extracellular enzymes. *Aquaculture Research* 42: 1–9.

Thompson JR, Pacocha S, Pharino C, *et al.* (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science* 307(5713): 1311–1313.

Venter JC, Remington K, Heidelberg JF, *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304(5667): 66–74.

Ward NL, Steven B, Penn K, Methe BA, and Detrich 3rd WH (2009) Characterization of the intestinal microbiota of two Antarctic notothenioid fish species. *Extremophiles* 13(4): 679–685.

Waterbury J, Watson SW, Guillard RRL, and Brand LE (1979) Widespread occurrence of a unicellular, marine, planktonic cyanobacterium. *Nature* 277: 293–294.

Waterbury J and Willey JM (1988) Isolation and growth of marine planktonic cyanobacteria. *Methods in Enzymology* 167: 340–343.

Whitman WB, Coleman DC, and Weibe WJ (1998) Prokaryotes – The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America* 95: 6578–6583.

Woese CR, Kandler O, and Wheelis ML (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America* 87(12): 4576–4579.

Worden AZ, Chisholm SW, and Binder BJ (2000) *In situ* hybridization of *Prochlorococcus* and *Synechococcus* (marine cyanobacteria) spp. with RRNA-targeted peptide nucleic acid probes. *Applied Environmental Microbiology* 66(1): 284–289.

Woyke T, Xie G, Copeland A, *et al.* (2009) Assembling the marine metagenome, one cell at a time. *PLoS ONE* 4(4): e5299.

Wu S, Gao T, Zheng Y, Wang W, Cheng Y, and Wang G (2010) Microbial diversity of intestinal contents and mucus in yellow catfish (*Pelteobagrus fulvidraco*). *Aquaculture* 303: 1–7.

Yooseph S, Nealson KH, Rusch DB, *et al.* (2010) Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468(7320): 60–66.

Yooseph S, Sutton G, Rusch DB, *et al.* (2007) The Sorcerer II Global Ocean Sampling Expedition: Expanding the universe of protein families. *PLoS Biology* 5(3): e16.

ZoBell CE (1946) *Marine Microbiology: A Monograph on Hydrobacteriology.* Waltham, MA: Chronica Botanica.

Microorganisms (Microbes), Role of

Tom Fenchel, University of Copenhagen, Copenhagen, Denmark

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bacteria Used as equivalent term for prokaryotes (Eubacteria + Archaebacteria); now frequently used as a synonym for Eubacteria.

Chemoautotrophy (chemotrophy) Energy metabolism exclusively based on the oxidation of inorganic compounds and the energy-requiring reduction of CO_2 for the generation of organic matter.

Cyanobacteria A large group within the Eubacteria with chlorophyll *a* and oxygenic photosynthesis; formerly known as "blue–green algae."

Methanogenic bacteria Anaerobic archaebacteria that convert $CO_2 + H_2$ or acetate into CH_4.

Nitrogen fixation The ability to fix atmospheric N_2 into organic N via ammonia; a widespread ability among several unrelated groups of prokaryotes.

Phagotrophy Feeding on particulate matter occurring exclusively among the eukaryotes.

Phototrophy Energy metabolism based on light energy; photosynthesis also means that the gained energy is also used for the reduction of CO_2 to organic matter.

Prokaryotes (bacteria) Organisms that are not eukaryotes (archaebactreia + eubacteria).

Protists Unicellular eukaryotes (protozoa + protophytes)

Sulfur bacteria Bacteria that depend on the chemotrophic or phototrophic oxidation of reduced sulfur compounds for energy generation.

Definition of Microbes

Microbes are defined mainly by their size; most are unicellular, but some form colonies which may be macroscopic. Some "microorganisms" even exceed the size of the tiniest animals (metazoa). Taxonomically, microbes include the prokaryotes or bacteria (the eubacteria and the archaebacteria) (Dworkin *et al.*, 2006) and members of various unrelated eukaryotic groups, together referred to as protists meaning protozoans (including some fungi-like organisms) and protophytes (unicellular algae) (Hausmann and Hülsmann, 1996). Contemporary understanding of phylogeny (**Figure 1**) reveals that biological diversity is largely microbial: In this representation, multicellular organisms (plants, animals, fungi, and macroscopic algae) appear as minor and evolutionarily recent additions to the "tree of life." This is also reflected perhaps by the fact that except for vascular plants (which are largely responsible for primary production on land) biogeochemical cycling of elements is mainly the result of microbial activity.

Although microbes have been defined in terms of size, it should be emphasized that the size range of microbes is immense. The smallest, free-living bacteria measure < 0.5 μm and the largest protozoan cells measure >1 mm (some foraminiferan species measure >1 cm). The size range of microbes, 1 μm to 1 mm (a factor of 10^3 or a factor of 10^9 in terms of volume), approximately equals that of all vertebrates (guppies to whales).

Even so, some general properties of microbes that relate to their relatively small size can be identified. Among these, the most important is a high "rate of living" – that is, high weight-specific metabolic rates and high growth and reproduction rates. When organisms spanning a wide size range are compared, weight-specific metabolic rates tend to decrease proportionately to the one-fourth power of body weight, and generation times tend to increase proportionately to the three-fourths power of body weight. Under optimal conditions the generation (doubling) time of bacteria may be as short as 20 min and unicellular eukaryotes typically have generation times between 4 and 24 h. Consequently, a relatively small microbial biomass may be responsible for a relatively large part of the flow of energy and materials in ecosystems. Certain microorganisms may in some

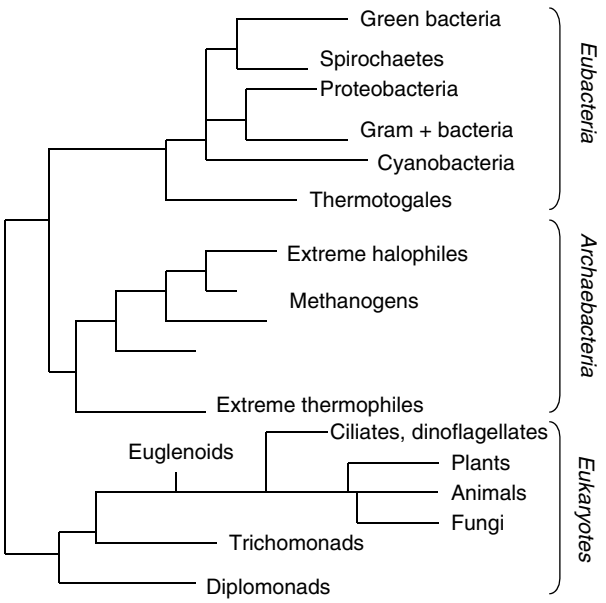


Figure 1 The "universal tree" based on rRNA sequencing. Microbes include, in addition to the prokaryotes (Archaebacteria and Eubacteria), most major eukaryotic groups. For clarity, only a few major groups are included in the tree. Only animals, fungi, plants, and a few (omitted) algal groups represent macroscopic, multicellular organisms; all other organisms are "microbes."

circumstances form conspicuous accumulations of biomass in nature, but usually they are macroscopically invisible constituents of, for example, soils and seawater. However, due to their high weight-specific metabolic rates they dominate as agents of chemical transformations in the biosphere.

There are other characteristics of being small. Uptake of solutes from the surroundings (e.g., dissolved oxygen or organic substrates) is by molecular diffusion, a fact that is important in shaping microbial communities and which imposes constraints on physical transport rates. Due to their small size and low swimming velocities, microorganisms live at "low Reynolds numbers"; this implies that motility depends on viscous rather than on inertial forces. Many properties relating to motility and motile behavior of microorganisms therefore appear counterintuitive (Fenchel *et al.*, 1998).

However, functional diversity of microorganisms is more evident than their similarities. It is therefore expedient to discuss prokaryotic and eukaryotic microorganisms separately.

Prokaryotes (Bacteria)

Principal Properties of Bacteria

The prokaryotes include two major groups of life: the archaebacteria (sometimes referred to as archaea) and the eubacteria (sometimes referred to as bacteria), which are differentiated by numerous genetic and biochemical traits. Structurally and functionally, however, they show so many similarities that it is appropriate to discuss them together and to retain the term bacteria (=prokaryotes) for these organisms that are not eukaryotes. In contrast to eukaryotic cells, bacteria do not have a cytoskeleton and almost all bacteria are enclosed by a rigid cell wall (Madigan, 1997). These features result in certain general properties. The limitation of diffusional solute transport from the surrounding water and within the cell typically constrains bacterial size to 1 or 2 μm. Certain giant bacteria (mainly among cyanobacteria and sulfur bacteria) measure 5–10 μm or more and in one case almost 1 mm, but these giant bacteria usually include a large internal vacuole. Bacteria take up only low molecular-weight solutes from their surroundings. Bacteria that depend on high molecular-weight polymeric compounds as a source of energy and organic carbon must first hydrolyze their substrates extracellularly, using membrane-bound enzymes, before the resulting monomers can be transported into the cell. This transport may be "passive" (facilitated diffusion) or "active" (ATP dependent). The small size of bacteria means that they can exploit and grow in extremely dilute substrate solutions (e.g., approximately 50 nmol l^{-1} of amino acids or glucose).

Two additional properties explain the important role of bacteria in the biosphere; these apply to bacteria as a group rather than to individual species. One is that many bacteria are "extremophiles": Extreme thermophilic bacteria live at temperatures $>80\,^{\circ}C$ (or $>100\,^{\circ}C$ under hyperbaric pressure); extreme halophiles may grow in saturated brine; acidophilic and alkalophilic bacteria occur at a pH <2 and >10, respectively; and some bacteria tolerate toxic metal ions (e.g., copper and zinc) in millimolar concentrations. A variety of habitats (hot springs, brines, subsurface, etc.) are inhabited exclusively by prokaryotes. Many natural habitats lack oxygen (e.g., aquatic sediments, stratified water bodies, water saturated soils, or even the interior of soil aggregates and detrital particles) because diffusive supply of O_2 cannot meet the demand for aerobic microbial respiration. Many bacteria have an anaerobic energy metabolism and they are practically the only inhabitants of anoxic environments except for a few species of anaerobic protozoa. It seems that bacteria have only one absolute requirement for metabolism and growth: liquid water. Active "terrestrial" bacteria (and other microorganisms) are therefore confined to water films surrounding soil particles, leaf litter, plant surfaces, and the like.

Another collective property of bacteria is their metabolic diversity, which far exceeds that found among eukaryotes. Metabolism of bacteria includes processes such as the oxidation and reduction of inorganic sulfur, nitrogen, iron, and manganese compounds and the production and oxidation of methane. Some bacteria can grow using hydrocarbons as a substrate. Some phototrophic bacteria use H_2, H_2S, or Fe^{2+} as electron donors (rather than H_2O as in algae, plants, and cyanobacteria) for the reduction of CO_2 and many bacteria can assimilate atmospheric nitrogen. These are all fundamental (but exclusively bacterial) processes in the biosphere. There is a limited number of types of energy metabolism that seem evolutionary conserved and a few of these have been taken over by eukaryotes. In addition, all natural and many xenobiotic polymers can be hydrolyzed by at least one type of bacterium (Falkowski *et al.*, 2008).

Metabolic Diversity of Bacteria

It is convenient to distinguish between dissimilatory (energy and catabolic) metabolism and assimilatory (anabolic) metabolism, although the distinction is not always sharp. Heterotrophic bacteria typically use a given organic substrate for dissimilation and also as a carbon source for growth, whereas autotrophic bacteria must use energy gained in dissimilatory metabolism for the reduction of CO_2. Energy generation and growth are coupled because most energy generated is used for the synthesis of macromolecules (DNA, RNA, and proteins), for active transport of substrates into the cell, and sometimes for assimilatory reductions. In microorganisms there is therefore typically an almost linear proportionality between energy generation and reproductive rate. Here, the concentration is on energy metabolism because it provides an overview of the functional diversity and roles of bacteria in the biosphere, notwithstanding that the synthesis of bacterial biomass is also an important process in terms of providing a basis for phagotrophic food chains (Fenchel *et al.*, 1998; Zehnder, 1988).

Table 1 presents a (simplified) overview of major types of bacterial energy metabolism. Different types of microbial processes predominate in different habitats according to the chemical environment (availability of substrates and electron acceptors), to thermodynamics (energy yields of the different processes), and to certain physiological constraints. Many of the listed processes are interdependent in nature in that one functional type of bacteria requires the presence of other types of bacteria.

Table 1 Principal types of energy metabolism in bacteria

Fermentation	Anaerobic processes in which (organic) molecules are dismutated. No external electron acceptor, redox equilibrium; low energy yield. Principal end products (of carbohydrate fermentation): acetate and $CO_2 + H_2$, but many fermentation types are incomplete and metabolites then include other low molecular-weight fatty acids and alcohols.
Respiration	Use of an external electron acceptor, electron transport phosphorylation.
Aerobic respiration	O_2 as terminal electron acceptor. A variety of organic substrates are used by different heterotrophs. Chemoautotrophs use O_2 to oxidize various inorganic substrates (reduced S, N, Fe, or Mn compounds and H_2 and CH_4). Aerobic respiration provides the highest energy yield of any known metabolic process.
Anaerobic respiration	Use of external electron acceptors other than O_2. Denitrifiers use NO_3^- (reduced mainly to N_2 or N_2O) to oxidize a variety of organic compounds and reduced S compounds (but not CH_4). A recent discovery is that some bacteria can oxidize ammonia with nitrite with principal end product being N_2. Sulfate reducers produce H_2S while oxidizing especially H_2 or low molecular-weight fatty acids. Iron and manganese reducers use Fe^{3+} and Mn^{4+} for the oxidation of substrates.
Methanogenesis	Methanogenesis is found only among certain anaerobic archaebacteria. Acetoclastic methanogens dismutate acetate into $CO_2 + CH_4$; CO_2/H_2 methanogens produce CH_4 from $H_2 + CO_2$. Some methanogens can also produce methane from reduced C-1 compounds (e.g., methanol).
Phototrophy	Phototrophs use light energy for generation of ATP and (with an external electron donor) for the reduction of CO_2 to organic compounds. Oxygenic phototrophs (cyanobacteria) use H_2O as electron donor and produce O_2 as metabolite. Anoxygenic phototrophs (several groups of bacteria) use H_2S, H_2, or Fe^{2+} as electron donors, thus producing S^0/SO_4^{2-}, H_2O, or Fe^{3+}. Some anoxygenic phototrophs have lost the ability of CO_2 reduction, but can generate ATP in the light; these bacteria have proven to be important in marine plankton. Some other bacteria use the so-called opsins – related to the visual pigments in eukaryotes – for energy conservation in the light.

Under oxic conditions, aerobic bacteria completely dominate in accordance with the fact that oxygen respiration yields more energy than any other metabolic process. Bacterial diversity in aerobic habitats is to a large extent due to differential abilities to hydrolyze different polymers, but most species are capable of complete mineralization of organic monomers to CO_2, H_2O, and mineral N or of the complete oxidation of reduced inorganic compounds.

Anaerobic bacteria, by contrast, are specialists, and a complete mineralization under anaerobic conditions requires several physiological types of bacteria. Fermenting bacteria play a key role in anaerobic habitats because they are (nearly) the only anaerobes capable of hydrolyzing polymers. The complete fermentation of, for example, carbohydrates (to acetate $+ CO_2 + H_2$) is thermodynamically possible only if the ambient H_2 tension is low (approximately $< 10^{-4}$ atm). Low H_2 tension results from the activities of H_2-consuming bacteria, notably sulfate reducers and methanogens. This syntrophic interspecies H_2 transfer is an essential feature of anaerobic habitats. In general, metabolic end products of fermenting bacteria serve as substrates for anaerobic respirers. The type of respiration that will dominate depends on the availability of external energy acceptors and on the energy yield of the different types of oxidation. After aerobic respiration, nitrate respiration is the energetically most favorable process, but its quantitative importance is often limited by the availability of NO_3^-. When nitrate becomes exhausted, Mn^{4+} followed by Fe^{3+} are the electron acceptors of choice; when these are exhausted sulfate respiration prevails (**Figure 2**). In marine sediments, sulfate reduction is typically the dominant process because of the high concentration of SO_4^{2-} in seawater (about 25 mM compared with about 250 μM O_2 at atmospheric saturation; the capacity for oxidation of organic matter through sulfate reduction is approximately 200 times more than through O_2 reduction and so, in contrast to O_2, SO_4^{2-} supply is not rapidly exhausted). Only when sulfate is depleted will methanogenesis be quantitatively important; this only at considerable depth in marine environments, and is more important, for example, in lake sediments. In anaerobic habitats, the microbial processes are therefore spatially or temporally structured; for example, when moving downward from the surface of sediments, O_2, Mn^{4+}, Fe^{3+}, and SO_4^{2-} are depleted sequentially. The reduced end products of anaerobic metabolism diffuse upward and are eventually reoxidized by O_2 (or NO_3^-) by chemoautotrophic bacteria residing at the aerobic–anaerobic interface. The complex interactions between different types of bacteria in anaerobic habitats and at the anaerobic–aerobic interface are often referred to as a "food chain" by microbiologists. It differs fundamentally from phagotrophic food chains, however, in that the organisms do not feed on each other but rather that one functional group of bacteria utilizes the metabolites of another functional group.

Hydrolysis of polymers is often the rate-limiting step in mineralization. Some substrates are inherently difficult to hydrolyze, notably the ligno–cellulose complexes of woody tissue and many other natural plant compounds such as phenols, waxes, tannins, and cork substances (whereas pure carbohydrates, such as cellulose and hemicelluloses, are relatively rapidly degraded under aerobic and under anaerobic conditions). Hydrolysis may also be limited by the availability of mineral nutrients (N and P) that are assimilated ("immobilized") by bacteria during decomposition of mineral-poor substrates (such as most plant tissue and litter). Crude oils (which contain a mixture of normal and branched paraffins, aromatic hydrocarbons, and other compounds) are degraded by a variety of aerobic bacteria and anaerobically by nitrate or sulfate reducers; the rate of crude oil mineralization is generally limited by the access to mineral nutrients, especially N and P. Otherwise, easily degradable monomers (e.g., amino acids) may bind to clay minerals or humic substances to become temporarily or permanently inaccessible to bacterial

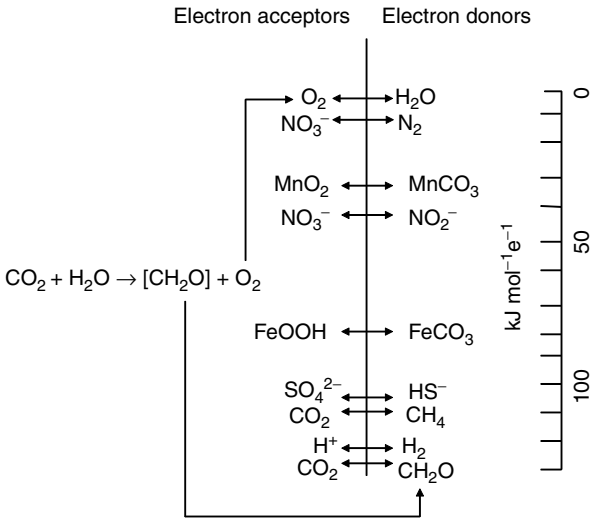


Figure 2 A biosphere model incorporating different microbial respiration processes. Oxygenic phototrophs use light energy to produce O_2 and organic matter ((CH_2O)); this represents the chemical potential that maintains the biosphere. Microorganisms derive energy by redox processes involving oxidants (left) and reductants (substrates) (right). The standard potential of the half-cell processes decreases from above so that the oxidant has to be situated higher than the reductant if the process is to be thermodynamically possible. Oxidations of organic matter may take place stepwise (e.g., first by sulfate reduction and the resulting sulfide can subsequently be oxidized by O_2). A few processes (e.g., N_2 oxidation) are not realized due to kinetic constraints. Data on free energy changes are not precise under physiological conditions because the processes do not necessarily occur under standard conditions and because only a part of the potential energy is conserved as ATP; however, the energetic hierarchy of the processes holds.

attack. Humic substances are complexes consisting mainly of aromatic rings deriving from lignin, quinones, and phenols. Humic substances are very resistant to microbial attack and their rate of turnover in nature must be measured in centuries.

Some polymers are not, or are only very slowly, degraded under anaerobic conditions. This applies to, for example, lignin and some other plant compounds. Anoxia in conjunction with a low pH, such as found in moors and swamps, strongly limits mineralization, and peat, in which the original structure of plant material is often preserved, accumulates. Even in marine sediments, a certain fraction of the organic input is not mineralized but fossilized as kerogen in sedimentary rocks. Peat is eventually (over geological time through abiologic processes) transformed into lignite and coal. Large accumulations of organic matter in marine deposits may eventually transform into crude oils and natural gas.

Roles of Bacteria for Element Cycling in the Biosphere

The single most important role of bacteria in the biosphere is the degradation and mineralization of organic matter produced by (mainly eukaryotic) phototrophs. In some habitats, however, primary production by prokaryotic phototrophs is also significant.

Mineralization processes are often complex due to spatial and temporal heterogeneity, limiting physical transport rates, and kinetic and physiological constraints. This complexity is especially evident where anaerobic processes are important as in sediments. There are considerable differences between major habitat types and these will be discussed separately in the following four sections.

The Water Column of Oceans and Lakes

Prokaryotic oxygenic phototrophs play a specific role in the water column. Thus, the unicellular, 1-μm-large *Synechococcus* cells are ubiquitous in lakes and in seawater, typically occurring at densities of approximately 10^5 ml^{-1}. Their role in primary production (relative to that of eukaryotic phototrophs) is especially important in oligotrophic oceanic waters where it may account for 30–70% of the total production. The tiny *Prochlorococcus* has more recently been shown to be important in the deepest parts of the photic zone (>100 m) in oceanic waters. Blooms of large colonial filamentous cyanobacteria (e.g., *Trichodesmium*) occur periodically in some marine waters; the organism is capable of N_2 fixation and is thus favored by P-rich but N-depleted water masses. Macroscopically conspicuous blooms of various species of colonial cyanobacteria are common phenomena in eutrophic fresh waters (Falkowski and Knoll, 2007; Harris, 1986).

Aerobic heterotrophic bacteria, however, are the most important group of prokaryotes in the water column. They typically occur in numbers ranging between 5×10^5 and 5×10^6 cells ml^{-1} constituting a volume fraction in seawater ($\sim 10^{-6}$) comparable to that of other important biotic components (e.g., phytoplankton). Their relatively constant numbers imply that they are controlled by grazing (mainly by protozoa) and they are also subject to viral attack, and their turnover is relatively rapid (varying from <1 to several days).

Bacteria depend on dissolved organic compounds which again derive from several sources. The most important one seems to be the passive excretion of photosynthate from phytoplankton cells; it is estimated that between 5% and 40% of primary production is lost in this way from algal cells to become mineralized by bacteria. Other sources include allochtonous material (e.g., runoff from land), degradation of macroalgae, and "sloppy feeding" by zooplankton. Dissolved organic matter in natural waters covers a wide range of molecular size from monomers (amino acids and monosaccharides) to colloidal matter. Small monomers have a rapid turnover (sometimes <1 h) and occur at very low concentrations (nM range) due to efficient bacterial utilization. Larger molecules are utilized much more slowly due to their lower diffusion coefficient and because their utilization requires extracellular hydrolysis. Some macromolecular compounds (humic substances) are very recalcitrant; they have a very low turnover rate, but constitute the bulk of dissolved organic matter.

The fact that a relatively large part of primary production is channeled via dissolved organic matter to bacteria, which then enter phagotrophic food chains, has been termed the microbial loop (**Figure 3**). The relative amount of the primary production which is channeled through the microbial loop compared with the "classical plankton food chain" (large phytoplankters → copepods → planktivorous fish → carnivorous fish) depends on circumstances. The classical food chain is

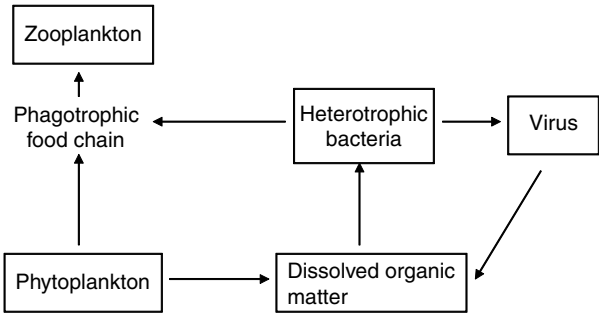


Figure 3 The "microbial loop" in the water column. A fraction of the photosynthate is lost from phytoplankton cells to become incorporated into bacteria, which then enter phagotrophic food chains.

particularly favored by large or periodic influx of mineral nutrients (such as in upwelling zones); otherwise, the microbial loop seems to dominate in terms of carbon flow (Fenchel, 1987).

Special microbial biota are associated with aggregates ("marine snow") which consist of different types of particles (diatom frustules, fecal material from plankton organisms, etc.) held together by colloidal material. They form in the water column where they constitute approximately 10% of nonliving organic matter; they eventually sink to the bottom and become the basis for benthic life. They are, however, rapidly colonized by bacteria (and other microorganisms) and so a substantial part of the organic fraction is mineralized while still suspended in the water column; it has been estimated that 75–80% of the organic matter is lost during a passage of 2000-m depth.

Aquatic Sediments

The most important property of sediments is that vertical solute transport depends on molecular diffusion. Organic material – produced in the water column – sediments and is in part mixed into the sediment through the activity of burrowing animals. The aerobic metabolism of microbes and animals deplete O_2 at a certain depth, depending primarily on the input of organic material. In shallow-water sediments, anoxia typically prevails a few millimeters beneath the surface; at greater water depths the oxic zone typically measures several centimeters. Beneath the oxic–anoxic interface (the "chemocline") mineralization is anaerobic. As previously explained, various other electron acceptors are used by bacteria for the oxidation of their substrates (**Figure 2**). In marine sediments sulfate reduction dominates and, in shallow offshore sediments, may be responsible for more than 50% of the terminal mineralization. The resulting metabolite, hydrogen sulfide, is thus produced in copious amounts. Some of it combines reversibly with iron to form black FeS, but most diffuses upward to become reoxidized by chemoautotrophic sulfur bacteria at the chemocline and therefore a major portion of the O_2 uptake of marine sediments is due to the reoxidation of microbially produced sulfide. Some of the ferrous sulfide reacts (abiotically) with sulfide to form the somewhat more stable pyrite (FeS_2), which appears as a type of fossil fuel in sedimentary rocks.

In sulfate-poor freshwater systems methanogenesis plays a major role as the terminal anaerobic mineralization process. Some of the CH_4 formed is reoxidized aerobically in the sediment surface or in the water column; since CH_4 (in contrast to sulfide) has a low solubility, much escapes to the atmosphere via ebullition.

In productive shallow waters the anaerobic zone may reach the surface of the sediments. This is macroscopically visible as a white cover of chemoautothrophic sulfur bacteria (they owe their color to intracellular elemental sulfur). If the sediments are exposed to light, the surface may instead be dominated by a purple layer of phototrophic sulfur bacteria. In stratified water bodies (such as thermally stratified lakes or where there is a salinity difference between the surface and deep waters), vertical, turbulent transport of dissolved O_2 from the surface layers to the deep layers is prevented. The chemocline may then extend above the sediment surface and into the water column. The phenomenon is common (at least seasonally) in deep, productive lakes and in fjords with a sill; the largest permanent anoxic water body on the Earth is the Black Sea, which is everywhere anoxic and sulfidic below 150- to 200-m depth (Fenchel *et al.*, 1998; Konhauser, 2007).

Terrestrial Soils

The microbial life of soils differs in many ways from that of aquatic sediments. The chemical composition of soil organic matter differs because it mainly derives from vascular plants and therefore includes large amounts of structural polymers (especially lignin and cellulose). Vascular plant tissue is very poor in mineral nutrients; C : N ratios typically exceed 60, which affects the microbial nitrogen cycle of soils. Another important feature is that soils represent a complex matrix that includes air-filled spaces. Since diffusion coefficients are approximately 10^4 times higher in air than in water, soils are not anaerobic throughout, although anaerobic microniches may occur within individual waterlogged soil particles, and therefore anaerobic microbial processes do occur. In general, the biological activity of soils exceeds that of aquatic sediments reflecting the higher input of dead organic matter. Microbial activity in soils affects soil fertility and some processes (e.g., CH_4 oxidation and denitrification) affect element cycling on a global scale.

The overall important, and variable, factor controlling soil microbial activity – qualitatively and quantitatively – is water content. Fungi (ascomycetes and basidiomycetes) are better adapted to water stress and their hyphae can penetrate air-filled spaces; in soils and litter they are therefore important rivals to bacteria as primary decomposers. Certain bacterial groups (including the spore-forming Gram-positive bacteria and the fungi-like actinobacteria and myxobacteria) are better adapted to water stress than other types of bacteria and thus play a relatively larger role in soils. For a given type of soil there is an optimum water content for microbial activity: A very low water content prevents microbial activity; when a large fraction of pores are water filled, anoxic conditions are more widespread, thus inhibiting aerobic processes. Water contents affect, for example, soil nitrogen cycling; low water contents favor ammonia oxidation (nitrification, an aerobic process), whereas high water content favors denitrification (an anerobic process leading to loss of reactive N from the soil).

The highest microbial activities are found within the root zone (rhizosphere). Microbes (in part directly attached to plant roots) exploit dissolved organic material, which is excreted

from the roots. They also form the basis for protozoan grazing. In addition to direct microbe–plant interactions (such as symbiosis with N_2-fixing bacteria), microbial activity has a profound indirect effect on soil fertility through its effects on mineral nutrient cycling (Coleman and Crossley, 1996).

Extreme Environments

The microbiology of extreme environments (here including hyperthermal and subsurface habitats and brines) has recently drawn considerable interest. Studies have expanded our knowledge of microbial diversity (especially of thermophilic archaebacteria) and led to the discovery of novel microbial communities; the potential for isolating microbes producing industrially useful enzymes has also spurred interest in extreme environments.

Microbial mats are communities that develop in environments with no or little animal grazing and disturbance. These are complex biota that include a great variety of interacting physiological types of prokaryotes and in some cases also include eukaryotic microbes. Cyanobacterial mats develop in illuminated thermal springs and brines (salterns); when undisturbed, they develop laminated deposits (stromatolitic mats) over centuries that may attain a thickness of several meters, although almost all biological activity is concentrated in the upper few millimeters. They are a modern analog to Precambrian stromatolites, which seem to have represented the dominating shallow-water biota for approximately 2.8 billion years of the Earth's history; they almost disappeared from the geological record in the past 600 to 700 million years (My), when metazoans made their appearance. Cold or hot seeps of sulfidic water support mats of chemoautotrophic sulfur bacteria. These biota – well known from not only deep-sea hydrothermal vents but also occurring elsewhere – are interesting in that their energy support (sulfide, methane, and reduced Fe) derives from geothermal processes rather than from solar energy via photosynthesis (however, the oxygen, which is necessary for exploitation of these energy sources, derives from oxygenic photosynthesis).

Subsurface microbial biota is a recent discovery. Thus, bacterial life has now been found hundreds of meters underground in shales, limestone, and other rocks, and groundwater ages and other evidence indicate that there has been no contact with the surface for at least 10^4 and up to 10^6 years. The microbial activity seems not only to be hydrogen based (methanogensis, acetogenesis, and sulfate reduction), but organotrophs also occur. Oil reservoirs in deep sediments have also been shown to harbor biota of hydrocarbon-degrading thermophilic sulfate reducers (Fenchel *et al.*, 1998).

Symbiotic Bacteria

There are numerous examples of bacteria living in or on other organisms; in many cases the adaptive significance of these associations is still not understood. Some cases of symbiosis between prokaryotes and eukaryotes, however, have been studied in detail; they are of scientific interest and sometimes even attract substantial economic interests.

Symbiotic polymer degradation is one such example. With few exceptions, animals are not capable of hydrolyzing and utilizing structural plant polymers directly. Animals do not produce the necessary enzymes (cellulase, xylanase, etc.). Moreover, plant tissue generally contains insufficient amounts of N and P to sustain growth, and various secondary plant compounds are toxic to animals. Some herbivores therefore eat copious amounts of plant tissue with low digestion efficiency. They are typically confined to a few (taxonomically related) food plants because each grazer has only developed detoxification mechanisms against specific plant toxins; this applies to many insects. Many animals, however, have solved the problems of herbivory by maintaining consortia of anaerobic microbes in their gut system. Such systems have evolved independently in several groups of mammals and in a few groups of birds (ostriches, ptarmigans, and hoatzins), reptiles, and fish; they are also found in termites and cockroaches, in some sea urchins, and in shipworms.

The best known example is the rumen of cows and sheep. Ruminants exemplify pregastric fermentation. The rumen, which constitutes 10–15% of the volume of the animal, is anatomically a part of the esophagus. Its anaerobic content is a mixture of rumen fluid, ingested plant fragments, and microbes (approximately 10^{11} bacteria and 10^6 protists ml^{-1}) Ruminant saliva does not contain hydrolytic enzymes, but it is strongly buffered. Urea (the principal nitrogenous excretion product of mammals) is also excreted into the saliva, supplying the microbial community with N; it is a mechanism for preserving this essential element for the ruminant through internal recycling. Within the rumen a consortium of different bacteria and some eukaryotic microbes ferment carbohydrates (including cellulose and xylan) principally into acetate, butyrate, propionate, and $CO_2 + H_2$. The fatty acids are absorbed in the intestinal tract, and they constitute the energy and carbon source of the ruminant. Protein is exclusively provided by the digestion of microbial cells in the true stomach; thus, a cow does not profit from eating proteins because these are microbially hydrolyzed, deaminated, and fermented in the rumen prior to the stomach. In the rumen, methanogens convert H_2 into CH_4; the methane represents a (necessary) loss to the cow (approximately 15% of the food intake) and it rids itself of the gas through belching. Ruminants are thus totally dependent on their microbial symbionts; ruminants that have artificially been deprived of their rumen biota cannot survive on a normal diet. In addition to ruminants, pregastric fermentation is found, for example, in kangaroos, sloths, and colubine monkeys (Fenchel *et al.*, 1998).

Postgastric fermentation is more widespread (e.g., horses, rhinos, elephants, lemurs, koalas, lagomorphs, and several others; even humans probably profit slightly from their intestinal microbial biota in this way). Animals with postgastric fermentation possess one or more ceca containing a consortium of fermenting bacteria. Ingested food is first subject to acid digestion in the stomach, with subsequent absorption of monomers, whereas nondigestable plant polymers are eventually fermented in the cecum. In most termites and cockroaches, the principal cellulose decomposers (harbored in the hind gut) are protists (various types of specialized flagellates), although bacteria also play a substantial role in these microbiota.

Nitrogen fixation (the reduction of N_2 to NH_4^+) is a very energy-requiring, anaerobic process, which is known only from bacteria. It is widespread especially among free-living anaerobes

and microaerobes, and some aerobes have special adaptations to protect the nitrogenase complex from O_2 (e.g., some cyanobacteria and *Azotobacter*). Nevertheless, it is estimated that in terrestrial habitats more than 90% of N_2 fixation is symbiotic.

Symbiotic nitrogen fixation is of immense economic importance and since antiquity it has been known that legumes improve soil fertility. The great majority of plants belonging to the Leguminosae form root nodules that harbor the N_2-fixing bacteria (*Rhizobium*). These bacteria normally live in soils; they especially thrive in the rhizosphere of legumes and often adhere to root hairs. Appropriate strains enter root hairs and induce the formation of root nodules. The bacteria transform into nondividing "bacteroids" that fix N_2 on the basis of carbohydrates provided by the plant. A hemoglobin (leghemoglobin) secures a sufficiently low O_2 tension while simultaneously allowing for aerobic metabolism, and the plant assimilates the formed ammonia.

Another type of symbiotic N_2 fixation is found in several unrelated species of trees (e.g., *Alnus*, *Myrica*, and *Hippophae*); it is due to the bacterium *Frankia* (an actinobacterium), which also forms root nodules. The adaptive significance of different examples of symbiotic cyanobacteria in plants seems also to be that of N_2 fixation (e.g., in cycads). The water fern *Azolla* harbors N_2-fixing *Anabaena*; the fern plays a role as green manure in rice cultivation.

There are many other interesting examples of symbiotic relations between bacteria and eukaryotes, including relatively exotic cases such as symbiotic bacteria that confer bioluminescence on certain species of fish and squids. Special mention should be made of various marine invertebrates that farm chemoautotrophic bacteria (especially sulfide oxidizers and, in a few cases, methane oxidizers). First thoroughly studied in the deep-sea hydrothermal vent fauna, the phenomenon has subsequently been found to be widespread in shallow waters as well. Pogonophorans (and the related vestimentiferans) and some bivalves and oligochaetes are gutless and entirely dependent on their symbionts for food; the hosts in turn provide the bacteria with necessary oxygen and sulfide via their circulatory system. Other bivalves maintain sulfur bacteria in their gills, but they are simultaneously capable of filter feeding. Some anaerobic protozoa maintain symbiotic methanogenic or sulfate-reducing bacteria. The protozoa (mainly certain ciliates) have a fermentative metabolism with H_2 production; the symbionts in turn are H_2 scavengers, thus maintaining a low intracellular H_2 tension which again enhances the energy efficiency of the host metabolism.

Unicellular Eukaryotes (Protists)

Protists are known to include many unrelated eukaryotic groups. Eukaryote cells differ from prokaryotes in possessing a cytoskeleton and membrane-covered organelles, among which mitochondria and chloroplasts are recognized as being descendants of endosymbiotic aerobic bacteria belonging to the α group of proteobacteria and to cyanobacteria, respectively. Protists have a much greater size range than bacteria; the smallest free-living species measure approximately 3 μm and the largest may attain sizes of several centimeters.

Phagocytosis is probably a primary property of eukaryotic cells and most extant species depend on particulate food (mainly bacteria or other protists). Many groups have acquired chloroplasts; this has apparently happened independently in different taxa and examples of "chloroplasts," which represent intermediate stages between an endosymbiont, and an organelle are known. Many species have secondarily lost chloroplasts and thus the ability of photosynthesis. Within many groups of phototrophs the ability of phagocytosis has been retained (mixotrophs, found for example among chrysomonads and dinoflagellates); in other groups (e.g., diatoms and green algae) phagocytosis has been irreversibly lost. A few species (some ciliates and foraminiferans) are capable of retaining chloroplasts from their prey cells; the chloroplasts remain functional for some days and this is exploited by the "host." Many heterotrophic protists harbor endosymbiotic phototrophs. Very few protists subsist on dissolved organic matter; in most habitats they would be inferior competitors to bacteria simply due to their larger size. The majority of protists are aerobes; a few specialized protists, however, are obligate anaerobes depending on a fermentative metabolism. Tolerance to extreme conditions is limited relative to that of some bacteria, but otherwise protists are omnipresent; primarily they are the principal consumers of bacteria in all types of environments and phototroph protists are largely responsible for primary production in aquatic habitats. (Fenchel, 1987)

Phototrophs

Macrophytes (macroalgae and vascular plants) are responsible for much of the primary production along shallow coastal waters and lakes. In offshore waters and in the oceans, unicellular phototrophs are solely responsible for light-driven CO_2 reduction and they thus largely constitute the basis for aquatic life. Characteristic annual succession patterns (with respect to species composition) occur in temperate seas and lakes. During winter, mineral nutrients are brought to the photic zone from deeper waters. In early spring, when light conditions allow, a bloom of diatoms develops; the bloom is brought to an end by zooplankton grazing and by nutrient limitation. During summer, phytoplankton is dominated by small forms (chrysomonads and prymnesids), which perform better at low nutrient concentrations; autumn is often characterized by a bloom of large dinoflagellates. In tropical waters which are thermally stratified throughout the year, competition for mineral nutrients is intense and small species prevail. Upwelling areas, such as off the coast of South America, support production of large phytoplankton species and thus represent an exception to this generalization.

Phytoplankton biomass is primarily limited by mineral nutrients – in most marine areas by inorganic N, although P limitation is also known and may be the rule in limnic systems. Diatoms may be limited by available Si and it has been inferred that Fe limits primary production in the Southern Ocean. Increased nutrient additions, especially from fertilizers, via river runoff and groundwater seepage have caused increased phytoplankton production and biomass in offshore waters, bays, and estuaries in addition to lakes in many parts of the world. Such eutrophication may be detrimental to

rooted vegetation (due to competition for light) and in some cases may result in widespread anoxia or hypoxia of bottom waters due to an increased input of organic material (Falkowski and Knoll, 2007).

Large phytoplankton cells are consumed by various zooplankters (copepods, etc.), which again serve as food for fish. Small phytoplankters are to a large extent consumed by protozoa, which then enter planktonic food chains. Many benthic filter feeders (e.g., mussels) also depend mainly on phytoplankton cells.

In shallow aquatic sediments phototrophic protists are important primary producers even though the photic zone extends only 2 or 3 mm beneath the sediment surface. In these biota, diatoms, euglenoids, and dinoflagellates are especially important (together with unicellular and filamentous cyanobacteria).

Phagotrophic Protists

Phagotrophic protists seem to be the principal consumers of bacteria in almost all habitats. In plankton, it is particularly the small (4–10 μm), heterotrophic flagellates that play a role as bacterial consumers, thus closing the microbial loop (**Figure 3**). They occur at densities of approximately 10^3 cells ml^{-1} in most natural waters. Other groups of protists, especially ciliates, mainly consume larger prey such as heterotrophic flagellates and small phytoplankters, whereas heterotrophic or mixotrophic dinoflagellates feed on prey, which often exceeds their own size. In oceanic plankton, the large acantharians, radiolaria, and the planktonic foraminiferan *Globigerina* are mainly phagotrophs feeding on large prey; most of them also harbor endosymbiotic phototrophs. The three (unrelated) groups all produce skeletons (made of strontium sulfate, silica, and calcium carbonate, respectively) and the sedimented skeletal remains of the two latter groups characterize many oceanic sediments.

Sediments, especially from shallow waters, harbor high densities of a wide variety of phagotrophic protozoa, including amoebae, flagellates, and ciliates, filling a variety of niches especially with respect to food requirements and oxygen tension. Foraminiferans are characteristic of marine sediments; their niches are in part taken over by testate amoebae and by heliozoans in lake sediments. Deeper marine sediments harbor large foraminifera and in deep-sea sediments the peculiar xenophyophoreans grow to a size of several centimeters and possess a skeleton made of barite.

In soils, small ameboid protozoa are most important as bacterial consumers; testate amoebae, heterotrophic flagellates, and ciliates are also present. Soil protozoa have various adaptations to thin water films (small size) and desiccation (formation of desiccation resistant cysts and life cycles, which allow exploitation of short periods with high moisture) (Fenchel, 1987).

Symbiotic Protists in Animals

Like prokaryotes, many protists occur as symbionts in animals; probably all animal species (including humans) harbor several protozoan symbionts. Symbiotic polymer degradation by flagellates in termites has already been mentioned. The most important type of symbiosis involving protists is that between animals and intracellular phototrophs. Many aquatic invertebrates harbor such symbionts. In fresh waters the symbionts are usually the green alga *Chlorella* (e.g., in freshwater sponges, the coelenterate *Chlorohydra*, and in the ciliate *Paramecium bursaria*). Most marine cases are based on dinoflagellates (*Symbiodinium*), but other groups (e.g., diatoms, chlamydomonads, and prymnesids) are also represented. In most cases the host combines the nutrition derived from the symbiont (usually in the form of carbohydrates) with particulate food; in some cases, it has been shown that the hosts can subsist entirely on the basis of the symbionts, and in a few cases the ability of phagotrophy has been lost. The most important marine example is that of phototrophic symbionts in reef-building corals; the symbionts are not only responsible for a significant share of primary production of coral reefs but also facilitate carbonate deposition of the host during active photosynthesis. A similar situation applies to giant tropical shallow-water foraminiferans; like corals, they are responsible for the formation of limestone deposits (the Cheops Pyramid is built from limestone consisting of the calcareous remains of the Eocene foraminiferan *Nummulites*). Other marine invertebrates harboring phototrophic symbionts include the giant clam *Tridacna* and various coelenterates (Fenchel, 1987).

Roles of Microbes over Geological Time and in the Contemporary Biosphere

This section briefly explores the role of microbes in the evolution of biogeochemical element cycling on a global scale. The Earth (and the solar system) arose approximately 4.6×10^9 years ago; the earliest sign of life dates to approximately 3.8×10^9 years ago (mid-Archean) in the form of a graphite particle found in sediment on the west coast of Greenland. Their contents of stable C-isotopes indicate that they derive from organic matter that was synthesized via the Calvin cycle presumably by some sort of bacteria. It is generally accepted that life arose on Earth, perhaps 4×10^9 years ago, under anoxic (and presumably chemically reducing) conditions, and that the atmosphere eventually became oxic as a result of oxygenic (cyanobacterial) photosynthesis. There is convincing evidence that atmospheric O_2 slowly increased from very low levels during the Precambrian. The earliest evidence of cyanobacteria (from Warrawoona in Australia and the Fig Tree Formation from southern Africa) is in the form of stromatolites, which are fossil remains of microbial mats. The nature of the oldest of these, approximately 3.5×10^9 years have, however, been contested. Somewhat younger Stromatolites – remains of cyanobacterial mats sometimes with well-preserved fossil bacteria that were very similar to present-day cyanobacteria – are known from throughout the remaining Precambrian. Evidence of an increasing O_2 tension of the atmosphere derives from banded iron formations: laminated (>3 to 2×10^9 years old) deposits containing partly oxidized Fe. The general interpretation is that as O_2 (resulting from photosynthesis) appeared in the atmosphere, dissolved reduced iron, in the otherwise anoxic oceans, was oxidized at the sea surface to become insoluble oxidized Fe that again gave rise to the deposits. Atmospheric O_2 levels thus initially remained low due to oxidation of reduced minerals on the surface of the

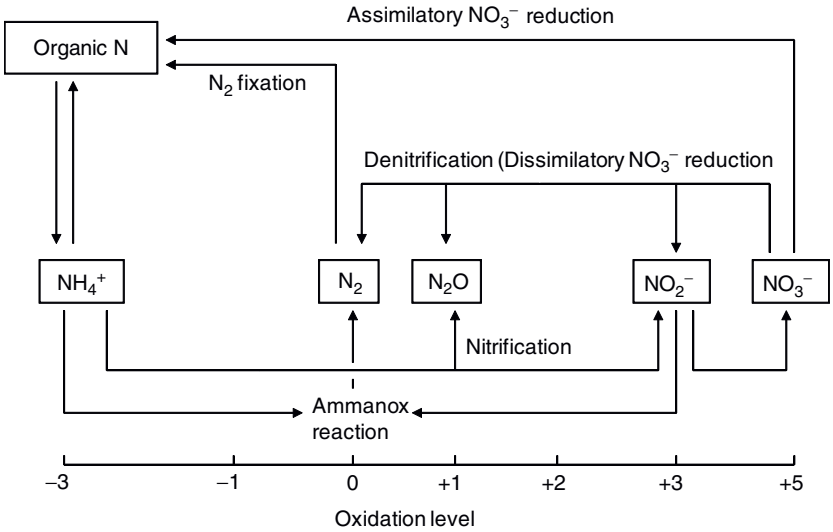


Figure 4 The global N cycle (simplified). Crucial processes (biological N_2 fixation, nitrification, and dissimilatory denitrification) are exclusively bacterial. Denitrification and nitrification also imply production of N_2O and other oxidation levels of N.

Earth (and presumably due to the early origin of biological O_2 respiration). Later, the O_2 level increased, reflecting the accumulation of fossil organic material in sedimentary rocks. There is evidence that the O_2 level had increased to $>1\%$ of the current atmospheric level approximately 2×10^9 years ago (early Proterozoic). When metazoa arose 600 to 700 My ago, atmospheric O_2 must have reached at least 10% of the current level (according to requirements of extant invertebrates), but it could also have been higher.

The conclusion is that oxygenic phototrophs, quite similar to modern cyanobacteria, must have evolved at least 3×10^9 years ago. It seems likely that all basic types of bacterial metabolism had also evolved by then, although this is speculative.

The time for the origin of eukaryotic cells is unknown, although molecular data (**Figure 1**) suggests "deep roots," there is (molecular) evidence suggesting that mitochondria (and thus "modern" eukaryotic microbes) arose approximately 2×10^9 years ago (or slightly earlier than fossils that have been interpreted as remains of protists). In all circumstances, when metazoans and the first macroalgae arose 600 to 700 My ago, almost all major biochemistry and metabolic pathways had evolved in microorganisms and with them the basic biogeochemical cycles as are known today. Colonization of land approximately 100 My later and the evolution of vascular plants and fungi, however, must have had a profound effect on the biosphere and on mineral cycling (Schopf and Klein, 1992).

There is not enough space here to describe the major biogeochemical element cycles (notably the C, N, and S cycles) in detail. A view of the nitrogen cycle (**Figure 4**), however, further illustrates the immense importance of microbes in the biosphere. Atmospheric N_2 is fixed by bacteria, but reactive N is also added to the biota as N oxides formed during electric discharges in the atmosphere and by industrial N_2 fixation. In cells, N occurs in a reduced form in organic matter and it is released as NH_4^+ when organic matter becomes mineralized. Through nitrification (a bacterial process) it is oxidized to NO_3^- which can be utilized (through assimilative reduction) by plants. The only pathway back to N_2, however, is denitrification, an anerobic bacterial process that is necessary for the completion of the global N cycle and ultimately regulates the amount of reactive N available to the biota.

The important mechanisms that control particular microbial processes are largely understood. However, the complexities of the cycles and the many positive and negative feedbacks render it difficult, or perhaps impossible, to make predictions of the effect of environmental changes on an ecosystem or on a global scale. An example is atmospheric CH_4, which is a greenhouse gas. The basic sinks (abiotic atmospheric oxidation and microbial oxidation in soils and other biota) and sources (mainly microbial methanogenesis primarily in wetlands including rice paddies, in lakes and in the sea, and in ruminants and termites; to a lesser extent through biomass burning and natural gas production) have been identified. It is also known that atmospheric levels (currently approximately 1.8 ppm) have increased (by approximately 1% $year^{-1}$) during the past century. However, the effect on CH_4 levels as a function of climate change is in practice impossible to predict. Both CH_4 production and oxidation increase with temperature, and melting of tundras will enhance methanogenesis, but other temperature-correlated factors may affect both processes either way. Determinants of methanogenesis also include input of organic C, competitive interactions with denitrifiers, and sulfate- and iron-reducing bacteria; anthropogenic factors such as the extent of rice paddies and ruminant husbandry should also be included (Falkowski *et al.*, 2008).

Microbes and Man

There are many relationships between humans and microbes and they affect us on many levels of our existence. We are hosts to many prokaryotic and eukaryotic symbionts in addition to being victims of bacterial and protozoan pathogens. Humans have also – long before the existence of microbes was

recognized – utilized microbial processes and learned to prevent some adverse effects of microbial activities.

Thus, microbiological technologies which have been in use since prehistoric times include a variety of fermented foods involving lactic acid bacteria, propionic acid fermenters, and the production of vinegar from ethanol. In addition to their gastronomic qualifications, these techniques have served to prevent or control undesirable or even dangerous microbial spoilage of food, which is also the purpose of salting, smoking, or acidifying food. It has also been suggested that the use of spices and drinking of wine, rather than water, also served to control pathogenic microbes.

A quite different use of microbes is acid mine leaching, which was also in use long before the underlying mechanism was understood. Circulating water through crushed copper ore leads to acid conditions and dissolution of the ore; metallic Cu can then subsequently be recovered from the leachate by chemical methods. The underlying mechanism is that a consortium of acidophilic, chemoautotrophic bacteria (including *Thiobacillus ferrooxidans*) oxidizes both the reduced S and the reduced Fe of pyrite (FeS_2), which is omnipresent in many ores. The resulting sulfuric acid in turn dissolves the ore.

Biological sewage treatment serves primarily to mineralize organic material. Various types of sewage treatment are in use, depending, among other factors, on the scale of the plant; most systems involve aerobic and anaerobic microbial processes. An important aspect is the flocculation of bacteria, a process which is enhanced by the presence of protozoa. Removal of nitrate by microbial denitrification is another important function of biological sewage treatment, whereas phosphate is primarily removed by chemical precipitation. Mainly in smaller plants, anaerobic mineralization can be exploited to produce methane, which can be collected and subsequently used for heating.

In recent times, mass production of certain species of bacteria for the production of enzymes and antibiotics has played an important industrial role. Recently, genetically engineered bacteria or yeasts that express human protein genes (e.g., insulin and other hormones) have been used in the pharmaceutical industry.

Microbial diseases, of which there are many, represent the most direct encounter between humans and microbes. Through recorded history such diseases have played an important role for human populations, most dramatically illustrated, perhaps, by the recurrent plague epidemics in Europe from medieval times to approximately 1700; however, many other bacterial diseases, such as cholera, tuberculosis, leprosy, and typhoid fever, and protozoal diseases such malaria were also important. In North America, Europe, and in some other parts of the world serious bacterial and protozoan diseases have, especially after World War II, largely been brought under control due to the combined effects of hygienic measures, vector control (mosquitoes and rats), immunization programs, and antibiotics and other forms of chemotherapy. However, globally, tuberculosis and malaria remain among the most frequent causes of death. Many bacterial and protozoal diseases of livestock also remain economically significant. Evolving resistance to antibiotics and other types of chemotherapy in agents of disease in humans and animals may represent an increasing problem and indicate that our interactions with pathogenic microorganisms are not a closed chapter in human history.

See also: Ecosystem, Concept of. Eukaryotes, Origin of. Marine Sediments. Microbial Biodiversity. Nitrogen, Nitrogen Cycle. Protozoa. Psychrophiles. Thermophiles, Origin of

References

Coleman DC and Crossley Jr. DA (1996) *Fundamentals of Soil Ecology*. New York: Academic Press.

Dworkin M, Falkow S, Rosenberg F, Schleifer K-H, and Stackebrandt E (eds.) (2006) *The Prokaryotes, A Handbook on the Biology of Bacteria*. New York: Springer.

Falkowski PG and Knoll AH (2007) *Evolution of Primary Production in the Sea*. San Diego: Elsevier.

Falkowski PG, Fenchel T, and DeLong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* 320: 1034–1039.

Fenchel T (1987) *Ecology of Protozoa. The Biology of Free-Living Phagotrophic Protists*. Berlin: Springer-Verlag.

Fenchel T, King GM, and Blackburn TH (1998) *Bacterial Biogeochemistry. The Ecophysiology of Mineral Cycling*, 2nd edn. San Diego: Academic Press.

Harris GP (1986) *Phytoplankton Ecology: Structure Function Fluctuation*. London: Chapman and Hall.

Hausmann K and Hülsmann N (1996) *Protozoology*, 2nd edn. Stuttgart: Thieme Verlag.

Konhauser K (2007) *Introduction to Geomicrobiology*. Malden: Blackwell.

Madigan MT, Martinko JM, and Parker J (1997) *Biology of Microorganisms*, 8th edn. Upper Saddle River, NJ: Prentice Hall.

Schopf JW and Klein C (eds.) (1992) *The Proterozoic Biosphere*. Cambridge, UK: Cambridge University Press.

Zehnder AJB (ed.) (1988) *Biology of Anaerobic Microorganisms*. New York: Wiley.

Migration

Daniel I Rubenstein, Princeton University, Princeton, NJ, USA
Mace A Hack, Nature Conservancy, NE, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Mace A. Hack, Daniel I. Rubenstein, volume 4, pp 221–233, © 2001, Elsevier Inc.

Glossary

Circannual rhythms Endogenous, or internal, rhythmic cycles of one year in duration that govern the onset and cessation of migratory behaviors.

Compass orientation Navigation in a particular direction without reference to landmarks or sites of origin or destination. Migrants are known to use compass information from magnetic fields, chemical gradients, and visual features such as the stars, sun, and planes of light polarization.

Diadromy Migrations that take individuals between fresh and salt-water habitats, a common phenomenon for many migratory fish species.

Partial migration The case where intrapopulational variation in migratory behavior leads some individuals to migrate while others within the same population may only migrate locally or remain sedentary.

Zugenruhe Restlessness exhibited by some migratory species, especially birds, if not allowed to migrate during their usual migratory period. It reflects an underlying physiological transition to a migratory state.

Migration describes the movement of individuals between spatially separate ecological communities, typically on a seasonal or annual schedule. Several characteristics of migrations distinguish them from other forms of animal ranging behavior, including more persistent movements, of greater duration, that follow a more direct path with fewer turnings. Furthermore, migrants do not respond to resources along their path, but show a heightened response to the same resources near the migratory journey's end. This latter feature distinguishes migrations from typical foraging and dispersal movements. Behavioral specializations may include specific activity patterns particular to departure and arrival, and unique patterns of energy allocation to support long-distance movements. The ecological consequences of migration are that they take a species from one community of organisms to another and they partition life histories so that specific phases or events occur in different ecological communities. Given these specific characteristics, human alterations to the landscape and changes to global climate are likely to have dramatic consequences on the nature of migration as well as the migratory species themselves and the communities and habitats they connect.

Introduction

Migrations capture the human imagination like few other animal behaviors. The single-minded struggle of the salmon fighting its way upstream, the barely visible formation of geese piercing the sky overhead, and the thundering line of wildebeest snaking across the open savanna – all speak of ancient rhythms that drive life on our planet. The often great distances moved and the large numbers of individuals involved make animal migrations a conspicuous and essential aspect of many regions' biodiversities. Animals migrate for many reasons but in general do so to avoid temporarily unfavorable conditions or to locate particularly favorable areas that can meet specific biological needs, such as reproduction. However, it is somewhat ironic that migrations might actually increase a species' risk rather than reduce it. Protecting migratory species requires preservation not only of their final destinations but also their migratory routes and stopover points as well. In effect, migration inextricably links the fates of biotas across the length and breadth of the globe. No natural phenomenon makes the point better that biodiversity on local and global scales shares the same actors and the same processes that drive them. In this article we explore the nature, scope, and patterning of animal migrations as a prelude to discussing how human-caused changes in the environment are likely to impact migratory species. We not only address direct effects on migratory species behavior, ecology, population dynamics, and evolutionary potential, but we also consider indirect impacts on the ecological communities and regional biodiversities that disruption of migrations can produce.

What is Migration?

Definitions of "migration" abound in the scientific literature, many of which capture the nature of the behavior for specific taxa but fail to generalize across taxa. Commonly cited elements include long distances traveled and movements from one place to another, then back again. But what is "long"? Dark-eyed juncos (*Junco hyemalis*) and Blackburnian warblers (*Dendroica fusca*) are both small passerine birds that migrate with the seasons; both breed at higher latitudes in

northeastern Canada (among other sites) yet juncos may migrate only a few hundred kilometers to their overwintering range while Blackburnian warblers fly several thousand kilometers to overwinter in the Andean forests of Ecuador. The lengths of these respective trips differ by an order of magnitude, yet in examining the ecology and behavior of each species, we might well consider both to constitute migrations. For example, both involve moving between distinct ecological communities on a consistently timed, seasonal basis. Similarly, "typical" migrants, like most birds, travel the same circuit each year. Others, however, may only complete a single circuit in their lifetime (e.g., salmon) or only part of a circuit before they die, such as insects in which successive generations continue the journey their predecessors began. Should we exclude such species from the "migratory" category despite other aspects of behavior, physiology, and life history held in common with typical migrants?

From a biodiversity perspective, migration drives a species' life history and pattern of resource use, and it ties together ecological communities in different regions. Thus, defining migration in ecological terms of a species' use of space over time is imperative to understanding its functional impacts on biodiversity. However, one also needs to ask, does biodiversity refer to just the specific organisms found within a region, or does it encompass the specialized behaviors such organisms display? Ecological consequences aside, are sedentary populations of Canada geese interchangeable with migratory ones when summing up a region's biodiversity? If migratory and nonmigratory races or subpopulations – whether genetically or phenotypically different – each represent unique components of biodiversity, then for distinguishing them we must also define "migration" in terms of the behavioral mechanisms.

The combined ecological and behavioral dimensions of migration have plagued attempts to come up with a single definition suitable from both perspectives. Ecological definitions have included "the act of moving from one spatial unit to another" by Baker (1978) and "the persisting change that is left over when all other, minor excursions are removed" (Taylor and Taylor, 1983). Both capture the essential idea of migrations extending the ecological space used by individuals, although ambiguity over what constitutes a spatial unit or minor excursion renders each too imprecise to be of much use in distinguishing migratory from nonmigratory species.

Dingle (1996) has proposed a definition of migration that emphasizes the differences between migrations and other forms of animal movement. He uses clear behavioral terms to distinguish migrants from nonmigrants, yet he uses each species as its own reference for differentiating movements with different ecological consequences. Generally put, nonmigratory (ranging) movements are driven by the immediate need to acquire or safeguard resources, and they cease when suitable resources are encountered or are adequately defended. The length and timing of foraging and territorial defense movements will vary daily, or on even shorter time scales, depending on the frequency of encountering resources or potential threats. Commuting daily between refugia and feeding locations, such as the diel movements of zooplankton and other aquatic animals, is similarly directed by resources and is responsive to their location and abundance. Even dispersal to establish a new home range or to leave a natal group usually ceases as soon as a suitable new home range is found. In contrast, once they set out, migrants will ignore many of the resources they encounter, or use them briefly to refuel, and only become responsive again to resource abundance and quality after all or a critical part of their journey has been completed. Migratory movements thus entail specific changes in behavior and physiology that distinguish them from nonmigratory movements, even those that are circular in nature. Defining migrations relative to a species' other forms of movement results in a more generally applicable definition and one that encompasses the diversity of the migratory behaviors animals display.

Specifically, five mechanistic attributes distinguish migratory behavior (Dingle, 1996). The movements are: (1) more persistent and of greater duration than ranging movements and (2) follow a more direct path with fewer turnings. There is (3) an initial suppression of responses to resource-derived stimuli, but often a heightened response to such stimuli near the migratory journey's end. Migrants may also have (4) specific activity patterns particular to departure and arrival and (5) unique patterns of energy allocation to support long-distance movements. Of course, not all migrants will show all five characteristics, but as a group these traits circumscribe the suite of distinct and specialized behaviors entailed in migration. To these behavioral attributes we add two functional hallmarks of migrations: (6) they take a species from one community of organisms to another; and (7) they partition life histories so that specific phases or events occur in different ecological communities.

Two common uses of the term "migration" will not be considered in this article: (1) migration in the paleontological sense of species shifting their historical distributions with climate change or geological events and (2) migration as geneticists use the term to refer to gene flow between populations.

The Evolution of Migration

To move from one area to another and back, organisms must evolve the ability to detect and react to directional cues in the environment. Typically these can range from local cues associated with monitoring and tracking changes in essential resources, to more global cues associated with magnetic, celestial, or odor fields. But evolving these capabilities are not without costs; cue detection is likely to be error prone and to lead to lowering of vigilance levels. Since selection will maximize the difference between benefits and costs associated with enhanced detection ability, an intermediate or modest level of gradient detection ability should be evolved. Moreover, recent models show that because of the costs associated with gradient detection, only a few individuals should evolve gradient detecting abilities in social species (Guttal and Couzin, 2011). Most individuals will be selected to parasitize the gradient detecting abilities of a few leaders. Just as in many other social contexts where some individuals produce resources but others scrounge (Barnard and Sibly, 1981), or some invest in establishing elaborate and costly mating

territories while others hover nearby and surreptitiously steal matings from approaching females (Rubenstein, 1981).

Just as polymorphisms in the detection of directional cues can be maintained in populations, polymorphisms in the synchrony of migration and the fraction of individuals within populations to migrate can evolve. Maintenance of such polymorphisms often occurs in annually reproducing species when subpopulations share a breeding ground and overwinter apart, or share an overwintering ground and breed apart. Such dual strategies evolve when the effects of crowding on survival and reproduction can be reduced by subdividing and separating the population during one of the seasons (Griswold *et al.*, 2010). But partial migration can also evolve when some fraction of the population skips breeding and remains in the nonbreeding habitat while those engaging in breeding migrate to the breeding grounds (Shaw and Levin, 2011). In this case, dual strategies evolve because of tradeoffs in current versus future reproduction. In general, for some in the population postponing breeding and avoiding the costs of migration is favored when mortality associated with migrating, or the benefit of delaying to increase body to enhance future breeding, or the probability of impending harsh conditions will diminish current reproduction is high for a subset of the population. Thus, as environmental conditions become more severe it is likely that collective migratory synchrony will decline and more populations will exhibit dual strategies.

Migratory Patterns: A Taxonomic Survey

The diversity of migratory behavior in animals overwhelms attempts to neatly summarize its character and function. Generally, animals migrate to escape unfavorable conditions or to exploit favorable ones, yet defining "favorable" or "unfavorable" is often specific to the taxon examined and the life-history function at hand. Polar and cold temperate habitats tend to have more migrant species than tropical ones because they vary so strongly in productivity and habitability, although this tendency differs widely among the major taxonomic groups. Migration distance and duration similarly vary to a great degree, even within groups of species that migrate for the same reason. Differences in physiology account in part for this variability, since larger body sizes can store relatively more energy to fuel longer trips and lower the weight specific cost of transport while certain forms of locomotion, such as flying or swimming, are more efficient than others. In addition to physiology, physical forces such as winds and currents act in concert with habitat topography to further shape the migratory route and schedule. Biogeographic history may even play a role in these features of migration since routes may be hard-wired genetically and slow to adapt, or learned and dependent on knowledge within lineages. Indeed, migrations present a truly fascinating mix of evolutionary puzzles in behavior, ecology, physiology, and biogeography (see Dingle, 1996, for an excellent introduction to these). The broad surveys that follow provide an overview of migration prevalence, function, and character for the major taxa with specific emphasis on the biodiversity of migrants per taxonomic group and region.

Birds

Birds epitomize the act of migration for many people. Whether one lives in the temperate zone or tropics, New World or Old, Southern or Northern Hemisphere, migratory species constitute a significant and conspicuous fraction of the avifauna. At the most northern and southern latitudes, close to 100% of bird species migrate out of the region for a part of the year. For example, 135 species breed in the arctic zone, yet all migrate south to spend most of their year elsewhere (Johnson and Herter, 1990). In more temperate regions, a majority of species, and nearly all insectivores, migrate to more tropical latitudes after breeding (Karr, 1980). Approximately 200 species migrate from North America to the West Indies, Central, and South America each year, while many more make shorter distance migrations into the southern US and Mexico. In Europe, 177 species, or 40% of the region's avifauna, migrate from temperate breeding grounds to overwinter in Africa; 104 species from Western, Central, and Eastern Asia join them there. Moreau (1972) has estimated that 5 billion individuals make this migratory journey south to Africa alone; the comparable number for the entire globe is surely several times this. Seasonal fluctuations in temperature ultimately drive food production cycles at higher latitudes, setting the schedule for migratory movements, but in tropical climates distinct wet and dry seasons may function in an identical manner. Substantial numbers of birds follow regular migratory routes that track resources such as fruiting or flowering trees, seeding grasses, or invertebrate flushes brought about by rains. Indeed, this form of short-distance movement may be the most common type of migration in birds. For example, out of 1450 breeding species in Africa, 532 species have been classified as intra-African migrants (Curry-Lindahl, 1981).

Bird migration is clearly a ubiquitous feature of our planet's biodiversity, yet migrant species are not distributed equally around the globe. Particularly, birds that migrate in excess of 1000 km – moderate- and long-distance migrants (Berthold, 1993) – are more prevalent as breeders in the northern hemisphere. This is largely due to the much greater land area at higher latitudes in the Northern than in the Southern Hemisphere. For example, only 20 species among the hundreds that breed in the temperate regions of southern Africa migrate as far north as the equator, while a much larger number and proportion of Palearctic species overwinter in the same equatorial region (noted earlier). Similarly, only 8% of the nearly 600 species breeding in Australia and Tasmania migrate in and out of the temperate zone (Faaborg, 1988). In the New World, 31 species of shorebird (e.g., plovers, sandpipers, and curlews) breed at higher latitudes in the Northern Hemisphere and migrate thousands of kilometers to spend their nonbreeding seasons in the Southern Hemisphere, yet not a single shorebird species that breeds in the south even migrates as far north as the equator during its nonbreeding season (Hayman *et al.*, 1986). Although the actual bases for these fascinating patterns remain relatively unexplored, they suggest a greater sensitivity of northern avifaunas to factors that threaten moderate- and long-distance migrants off their breeding grounds.

Long-distance migrants accomplish truly stunning feats. Blackpoll warblers (*Dendroica striata*), weighing only 10–20 g,

embark from Cape Cod on a nonstop, 86 h, 3500 km flight 2000 m above the waves of the Atlantic to the northeastern coast of South America (Baird, 1999; Williams and Williams, 1978). Tundra-breeding shorebirds such as the Pacific golden plover (*Pluvialis fulva*) make nonstop 5000–7000 km journeys to wintering grounds in the South Pacific (Johnson and McFarlane, 1967). Arctic terns (*Sterna paradisaea*) have the longest migration of any bird, and perhaps any animal, traveling an annual circuit that can exceed 40,000 km (Alerstam, 1985). They breed in the boreal summer along the northern edges of the Old and New World continents, then migrate along the western edges of North America, South America, Europe, and Africa to feed in the rich waters off Antarctica's pack ice during the austral summer. Almost 8 months of each year are spent in transit between these two endpoints.

The Arctic tern's migratory journey seems extraordinary but it typifies why many birds migrate (Gauthreaux, 1982). In doing so, they exploit highly seasonal flushes of food resources, especially to meet the increased demands of breeding. High variability in food supply at higher latitudes, especially in insects and other invertebrates, makes it difficult for species to reside permanently in the habitat. However, the very predictable nature of these fluctuations allows mobile species to rely on them for part of their annual cycle. The absence of resident competitors may make these ephemeral resources even more abundant and accessible to migrants. Thus, the typical migratory movement for birds involves making an annual round trip between seasonally resource-rich breeding sites and nonbreeding areas where resource abundance may be lower but less variable over time.

Seasonal resource variability explains why many birds migrate, but what determines how far they go and where their final destinations lie? Proximal determinants of migration distance and route include weather patterns, history, the distribution of resources along the way, and the character of the landscape (Lövei, 1989; Kerlinger and Moore, 1989; Williams and Williams, 1978; Dingle, 1996; Berthold, 1993). The north–south orientation of mountain ranges in North, Central, and South America tend to funnel migrants along corridors also running north–south, whereas Eurasia mountain ranges run east–west, forcing many Palearctic migrants to make large westerly movements before they are able to fly south. The importance of trade winds or winds generated by weather fronts to migratory journeys can be seen often in the very different routes followed when moving north versus south (e.g., Arctic tern study by Alerstam, 1985). The migratory path of the white-rumped sandpiper (*Calidris fuscicollis*) extends the length of the Western Hemisphere and appears carefully choreographed to coincide with seasonal pulses of invertebrate prey along the route (Harrington, 1999). History too seems to shape migratory paths and destinations as suggested by the very indirect route (via the eastern Mediterranean) taken by red-backed shrikes (*Lanius collurio*) migrating from the Iberian Peninsula to their nonbreeding range in Central Africa (Berthold, 1993).

Habitat suitability is a more ultimate determinant of migration distance and destination. Species with very specific habitat requirements may have limited options for suitable nonbreeding areas. For example, the upland sandpiper (*Bartramia longicauda*) breeds in North American grasslands and must travel 10,000 km to overwinter in the pampas of Argentina – the only similar Southern Hemisphere grassland habitat in the New World. Many tundra-breeding shorebirds migrate to the southern coastlines of South America, Africa, Southeast Asia, and Australia where extensive intertidal mudflats and rocky beaches provide abundant invertebrate prey. Most neotropical migrants – the warblers, flycatchers, vireos, swifts, hummingbirds, swallows, tanager, orioles, and raptors – migrate more moderate distances to forest and scrub habitats in Central America, the West Indies, and northern South America. In addition to the general structure of the habitat, competition with resident forest species and other migrants certainly influences where migrants settle during their nonbreeding seasons (Terborgh, 1989).

Often the question of how far migrants move depends on which population one examines. Many migratory birds conduct "partial" migrations where some individuals migrate while others remain as year-round residents on the breeding grounds or move only short distances (e.g., European blackbird, *Turdus merula*, Berthold, 1993). Different age and sex classes may pursue different migratory strategies, as in the dark-eyed junco of the eastern US (Ketterson and Nolan, 1983). Alternatively, all individuals may migrate but different populations travel different distances. A common pattern in this case is for higher-latitude breeding populations to migrate the longest distances, "leap-frogging" beyond the migratory movements of lower-latitude populations (e.g., fox sparrow, *Passerella iliaca*, Wetmore, 1926). Some sanderling (*Calidris alba*) populations migrate only a short distance from their tundra breeding grounds to overwinter in the northwestern US while others travel 7500 km to nonbreeding areas in Chile. Apparently, the much greater energetic and exposure costs of migrating to South America are offset by much richer food resources and a more hospitable climate, so that the payoffs for each strategy are equivalent, and both short- and long-distance migrants persist (Myers *et al.*, 1985).

The presence of both migratory and nonmigratory strategies in the same species underlines the opportunistic and flexible nature of bird migration. As environmental changes occur or as a species expands its range, migratory behavior can often adapt to fit the new circumstances. The European starling (*Sturnus vulgaris*) is a widespread, permanent resident of Britain today, yet in the eighteenth century starlings regularly migrated out of Britain to overwinter in warmer regions of Europe (Alexander and Lack, 1944). Long-term climate change has presumably made Britain a more hospitable place for starlings to spend their winters. Conversely, several year-round resident species in Europe have extended their breeding ranges north over the past century. These more northerly populations have developed full migrations to southern Europe for overwintering. Migrants may also establish nonmigratory populations along traditional migratory routes as resources become more abundant or habitats are altered, for example, the barn swallow (*Hirundo rustica*) in Argentina (Martinez, 1983) and the Canada goose (*Branta canadensis*) in eastern North America. With industrial parks in North America supporting extensive expanses of grassy lawns some Canada geese cannot only forage sufficiently well at temperate latitudes during the winter, but high levels of vegetative production during the spring and summer enable them to stay and breed.

A fascinating form of bird migration not related directly to food abundance is molt migration. It is particularly common among waterfowl that molt all their flight feathers simultaneously and thereby lose the ability to fly. To escape predation, they may migrate to coastal areas, large lakes, or far offshore until their new feathers have grown in.

Mammals

Small (<5 kg), nonvolant mammals represent a large fraction of our planet's mammalian biodiversity – rodents alone account for 40% of all mammal species – yet very few species are known to migrate. Rather than leave highly seasonal habitats when resource abundances decline, many small mammals hibernate or reduce activity levels and wait for conditions to improve. Species known to migrate include lemmings in the arctic tundra (Stenseth, 1983) and a variety of rats, mice, and shrews living on the Kafue River flats of Zambia (Sheppe, 1972). In both cases, these small mammals move several kilometers to escape flooding conditions as snow melts (lemmings) or rivers flood (Kafue mammals).

Migration occurs more commonly among a second, highly diverse mammal group, the bats (25% of all mammals). The ability to fly and travel large distances more efficiently than via terrestrial forms of locomotion must account in large part for this difference between bats and other small mammals. Fruit-eating bats in West Africa migrate 1500 km annually, following the movement of rains and the consequent pattern of fruit abundance (Thomas, 1983). As in birds, different sex and age classes may migrate different distances; juveniles migrate further, perhaps trading-off reproduction for more abundant food and greater survivorship. Similar migrations tracking fruit, nectar, and insect abundances have also been found in Australia and the New World tropics (Fenton and Thomas, 1985; Richards, 1995).

Insectivorous bats breeding in the temperate regions of North America and Europe commonly migrate, showing two distinct patterns of movement (Griffin, 1970; Nowak, 1994). Like many birds in these regions, species such as the Mexican free-tailed bat (*Tadarida brasiliensis*) and the hoary bat (*Lasiurus cinereus*) travel a maximum of 800–1700 km south to overwinter in warmer sites where their insect food can still be found. These individuals do not hibernate, but other individuals of the same species may make only part of the southward journey to hibernate at sites along the migratory route. The second migratory pattern involves movements of shorter distances (10–200 km), and more variable orientation, from summer ranges to particular winter hibernation roosts. For example, as Griffin discovered, the little brown myotis (*Myotis lucifugus*) hibernates in a few selective caves and mine tunnels in northern New England, yet migrates usually less than 100 km back to summer ranges both north and south of these wintering sites (Griffin, 1970). Seasonality determines the timing of migration in these species but the very specific requirements of hibernation, such as temperature stability, humidity, and protection from predators, drive migration distance and direction. It is not uncommon for 75–95% of the population in a large geographical area, or even an entire species, to migrate and hibernate in only a few caves (Pierson, 1998). Both migratory patterns are particularly common among species that roost in trees during the summer, presumably because trees make very poor, exposed winter hibernation sites.

Among terrestrial mammals, migrations are most common, and most spectacular, in the ungulates, or hoofed mammals. In the Serengeti ecosystem of Kenya and Tanzania, more than 1 million wildebeest (*Connochaetes taurinus*) and several hundred thousand plains zebra (*Equus burchelli*) and Thomson's gazelle (*Gazella thomsonii*) migrate seasonally along a several-hundred-kilometer annual circuit. A spatially varying seasonal rainfall pattern, and the consequent growth in grass that it generates – the primary food of all three species – determines the migration's timing and route (Maddock, 1979), although drinking water availability and salinity may also play important roles (Wolanski *et al.*, 1999). Plains zebra and wildebeest populations elsewhere in Africa make similar migrations between wet and dry season grazing areas, but in smaller numbers. These species are not obligatory migrants, however, since resident, nonmigratory individuals can be found in most populations. Other ungulate migrations driven partly by seasonal food availability include the movements of caribou and reindeer (*Rangifer tarandus*) in the Holarctic region from tundra in the summer to taiga during winter (McCullough, 1985). The more wooded taiga also provides greater protection from harsh winter conditions; eastern North American populations of "woodland" caribou do not migrate at all.

Ungulates in temperate regions often migrate altitudinally with the seasons, moving to lower altitudes in winter to avoid harsh weather, low food abundance, and low food accessibility due to deep snow (McCullough, 1985). Such species include bighorn sheep (*Ovis canadensis*), elk (*Cervus elaphus*), mule deer (*Odocoileus hemionus*), and feral horses (*Equus caballus*) in the Nearctic and chamois (*Rupicapra rupicapra*) in the Palearctic. Migratory distances in these species are usually short. An interesting exception is the mule deer, where each individual or family has a definite summer and winter range yet the migratory route is often not the most direct link between them. Individuals have been observed to travel 150 km in straight-line distance, crossing six mountain ranges in a winding route to do so (Gruell and Papez, 1963). Tradition appears to have a tremendous effect on defining seasonal ranges and migratory routes in this species.

For marine mammals, annual migrations are the rule rather than the exception (Bowen and Siniff, 1999). The baleen whales, or mysticetes, migrate the longest distances, some moving thousands of kilometers from tropical to polar waters and then back again each year. Northern and Southern Hemisphere populations both migrate in this way, but their opposite schedules prevent overlap in the tropics. Food availability and quality at higher latitudes drive the timing of these movements. During the polar summer, long day lengths lead to phytoplankton blooms which, in turn, generate huge abundances of the zooplankton, such as krill, that baleen whales specialize on when feeding. The lipid content of krill and small fish prey, or energy per mouthful from the whale's perspective, also rises during the summer months (Mårtensson *et al.*, 1996). Indeed, polar waters provide such an amazingly rich food source that many species forego feeding the rest of the year, consuming enough during a 3-month summer binge

to not only maintain themselves for the remaining 9 months but also to complete a hemispheric round-trip and breed. In species such as the fin whale (*Balaenoptera physalus*) both mate and give birth 11 months later in warm, low latitude waters (Boyd *et al.*, 1999). Lactation lasts for 6–7 months and the calf is weaned on the summer feeding grounds; fat reserves for lactation just after birth may constitute 50–75% of a female's body mass. Lower thermoregulatory costs in warm water presumably allow the binge-migrate-and-fast strategies of baleen whales to be successful, especially since warm, tropical waters may also provide a thermal environment more conducive to calf growth.

Odontocetes, or toothed whales such as dolphins, pilot whales and beaked whales, may show regular seasonal movements between breeding and nonbreeding grounds, but long-distance migrations rarely occur (Bowen and Siniff, 1999). Most toothed whales feed on single food items like fish and squid and, as a consequence, must feed every day to maintain themselves. Many have large feeding ranges through which they travel 20–60 km per day (Wells *et al.*, 1999). Short-distance migrations typically occur when food abundance in an area changes seasonally, such as when fish schools move off-shore or concentrate in river mouths. Some Arctic odontocetes, such as the narwhal (*Monodon monoceros*) and beluga (*Delphinapterus leucas*), migrate short distances seasonally as the polar pack ice retreats and opens up calving grounds in warmer-water estuaries. When the ice returns, they move back with its advancing edge to deeper, ice-free waters. Only the sperm whale – the largest odontocete – migrates distances comparable to those seen in the baleen whales. Adult males move two- to three-times the distances of females, reaching richer feeding areas at higher latitudes in order to achieve and maintain their much larger body sizes (Best, 1979).

Among pinnipeds, 44% of phocid (true or hair seals) species migrate seasonally while only 14% of otariids (fur seals, sea lions) do so (Bowen and Siniff, 1999). This difference may arise from the phocid's greater reliance on seasonally changing pack ice, rather than solid land, as a substrate for giving birth, raising young, and molting. For example, harp (*Phoca groenlandica*) and hooded seals (*Cystophora cristata*) whelp on pack ice in March near Newfoundland, among other sites, then migrate north with the retreating pack ice to feed. Annual round-trips may cover 4000 km, although there is much variation in migratory distance and route among individuals and populations throughout the North Atlantic. Northern elephant seals hold the current record for the longest migration of any mammal: 18–21,000 km annually as they move twice between Californian islands used for breeding (January–February) and molting (July–August) to higher-latitude feeding areas rich in cephalopod prey (Stewart and DeLong, 1995). Northern fur seals, perhaps the only true migrants among the otariids, migrate in the opposite direction of most phocids, breeding at higher latitudes in the summer months and migrating south to temperate waters to overwinter (Gentry, 1998).

Although clearly not mammals, sea turtles closely resemble pinnipeds in their migratory behavior. Adults migrate to specific island or continental beaches where eggs are laid. Hatchlings venture immediately to the sea where they may spend 30 or more years before maturing and returning to the same beach to complete the cycle (Meylan *et al.*, 1990).

Tropical marine mammals, such as dugongs and manatees, are generally nonmigratory, although the West Indian manatee (*Trichechus manatus*) population in Florida may move several hundred kilometers north in summer to exploit new feeding areas, retreating back to warmer waters in the winter (Reynolds, 1999). Their need for warm water in which to overwinter has in some cases made them dependent on the warm-water discharges of coastal power plants.

Fish and Other Aquatic Species

Fish show a range of migratory patterns, both among different species and different populations of the same species (McKeown, 1984). Migration is relatively common in this group, particularly as a means of linking rich feeding habitats with specialized spawning grounds that provide refuge for eggs and young fish. For diadromous species – the most studied and conspicuous migratory fishes – this journey requires moving from marine to freshwater environments in order to breed (anadromous: e.g., salmonids, shad, sticklebacks, lampreys) or, more rarely, the converse (catadromous: e.g., freshwater eels, southern trout, southern smelt). Anadromous species predominate in cold-temperate and subpolar waters in both hemispheres, while catadromous species more commonly occur in warm-temperate and subtropical waters (McDowall, 1988). The benefits of moving from lower to higher productivity habitats to feed and grow may account for this pattern since marine productivity is higher at high latitudes while freshwater productivity is higher at low latitudes (Gross *et al.*, 1988). Anadromous species constitute a higher proportion of total coastal fish diversity in the Pacific than they do in the Atlantic, largely due to the Atlantic's much higher coastal fish diversity overall.

As with marine mammals (e.g., baleen whales), migrant fishes with separate feeding and spawning areas must feed intensively and store sufficient energy reserves for both migration and spawning. Although many species migrate only a few kilometers between feeding and breeding sites, some fish migrations are truly remarkable in length. Upstream distances in diadromous species include 300–400 km in lampreys, 500 km in shad, and well over 1000 km in some salmons and sturgeons (McDowall, 1988). Sockeye salmon migrating upstream expend 70% of their available energy in reaching the spawning grounds and the rest on spawning itself (Brett, 1986); not surprisingly, both sexes die shortly thereafter. The combined toll of fasting, migrating upstream, and spawning results in semelparity for most diadromous species (McDowall, 1988), but some species, such as the Atlantic salmon (*Salmo salar*) and northern populations of American shad (*Alosa sapidissima*) may migrate and spawn more than once (Wootton, 1990). The dramatic physiological transition required to move between freshwater and marine environments may also be a factor in the prevalence of semelparity among diadromous species.

Salmonid migrations typify anadromous life cycles, but they also illustrate the very flexible migratory strategies of fish (McKeown, 1984). Pacific salmon (*Oncorhynchus* spp.) spawn

in streams at cold-temperate and subpolar latitudes on both sides of the Pacific. After hatching, they either (1) travel immediately downstream to the ocean or, (2) before migrating to the ocean, move downstream to lakes where they spend the next 2–4 years, or (3) remain in their hatching stream for 1 or more years. Once in the ocean, some species migrate the breadth of the Pacific to reach feeding areas while others remain closer to their spawning streams; accordingly, the time spent feeding and growing in the marine environment varies among species but is generally 2 or more years. The same species may have several distinct upstream migrations throughout the year. Further variation on this pattern is shown by individuals (e.g., sockeye salmon: *Oncorhynchus nerka*) that never migrate to the ocean but remain in lakes and move upstream to spawn with ocean-returning migrants. In contrast to other anadromous fish, such as shad and lampreys, salmonid ancestors resided year-round in freshwater and evolved the ability to migrate to marine habitats, perhaps accounting for the great variety of migratory strategies now seen in this group.

The best known catadromous migrants are the freshwater eels (*Anguilla* spp.). They occur throughout the world, including North America and Europe where most studies of their life cycles have been conducted (McCleave and Kleckner, 1985). In North America, adult eels live in rivers and brackish estuaries from the Gulf of Mexico to Labrador, migrating into the Atlantic to spawn. The actual spawning sites still remain a mystery, but small larvae (leptocephali – once thought to be a separate species with no affinity to eels) have been found in the Sargasso Sea, suggesting spawning is here or very nearby. Over the first year, currents carry the leptocephali toward inshore habitats where they metamorphose into glass eels and leave the marine environment. After several years in fresh water, they are ready to migrate, spawn, and complete the cycle, dying after they spawn. Interestingly, adult females are found at greater distances from the spawning grounds – further up rivers, in lakes, and at higher latitudes – than adult males and should thus have greater costs of migration (Helfman *et al.*, 1987). Females also grow larger at maturity suggesting very different life history strategies for the two sexes.

Oceanodromous species, or those that migrate throughout the marine environment alone, often travel complex migratory circuits between sites optimal for different stages of the annual or life cycle. Ocean currents and coastal topography shape these movements, and those of prey items, adding further complexity to the annual migratory pattern. For example, Atlantic herring (*Clupea harengus*) have numerous northern subpopulations, each with distinct spawning and feeding grounds and each following its own migratory route and schedule without intermingling to any great extent (Baker, 1978). Immature fish have separate migrations, moving inshore during warmer months and offshore to deeper waters during winter. On reaching maturity, they follow the migratory circuit of their parents. Other oceanic long-distance migrants include cod (*Gadhus morhua*) and plaice (*Pleuronectes platessa*) at higher latitudes and several tuna species in temperate and tropical waters. In most cases, these migration circuits are of the order of 1000–3000 km in diameter (McKeown, 1984). Littoral migrations also occur commonly among oceanodromous species, but they are generally much shorter. Those species moving to track food resources usually move offshore in winter, while species migrating inshore during winter may do so to spawn in relatively predator-free habitats.

Species living solely in freshwater (i.e., potamodromous) commonly migrate from deep to shallow waters in order to spawn. A large variety of potamodromous fish migrate throughout the world's lakes and rivers, including freshwater rays, sturgeons, suckers, minnows, pikes, sunfishes, darters, perches, and catfishes. Shallower, upstream habitats may exclude certain predators or have water characteristics (e.g., oxygen, temperature, or silt levels) more suitable to the development and growth of eggs and young fish (McKeown, 1984). The South American characin, *Prochilodus mariae*, has both nonmigratory, lake-breeding and migratory, stream-breeding individuals within the same population. The observation that lake breeders expend five-times more energy on egg production than stream breeders suggests that stream habitats provide sufficient benefits in offspring survival to compensate for the high energetic costs of migration (Saldana and Venables, 1983).

Beyond fish and marine mammals, migration in aquatic animals is less common – or perhaps just understudied. Spiny lobsters in the family Palinuridae demonstrate a very curious migratory behavior, queuing up in long lines of up to 50 individuals that snake their way along the ocean bottom toward deeper, more sheltered habitat. These movements, as much as 30–50 km in length over several days, usually occur seasonally in response to a greater incidence of polar storm fronts, which bring colder water temperatures and greater wave disturbance to the shallow water habitats the lobsters feed in during warmer months. Migration thus allows spiny lobsters to exploit a resource-rich, yet seasonally stressful, habitat (Herrnkind, 1985).

Plankton show distinct migrations in the water column driven by the seasonal availability of nutrients (Angel, 1985). During months of high productivity, plankton migrate to surface waters; as productivity declines, they return to lower depths and often enter into a state of diapause. Since changes in light work in concert with endogenous rhythms to set the schedule for these movements, the maximum depth reached during the nonproductive season will depend on how far light penetrates the water. In the Northeast Atlantic, typical maximum depths are 1200 m (Roe, 1983). Seasonal planktonic migrations are most prevalent at higher latitudes and in regions where currents cause seasonal upwellings of nutrients. (Many planktonic species demonstrate diel vertical "migrations" in the water column (Huntley, 1985), but because these constitute daily ranging movements rather than migrations by the definition given earlier, we do not consider them further.)

Amphibians and Reptiles

Amphibian life cycles often require a return to water to reproduce, resulting in migrations from feeding areas or refugia to seasonal ponds, streams, and other water sources (Stebbins and Cohen, 1995). However, salamanders and anurans are not known for their great mobility and consequently most species do not travel far from where they were

born. Red-bellied newts (*Taricha rivularis*) may hold the distance record for their migrations of 1 km from feeding range to breeding site (Twitty *et al.*, 1967). Similar migrations in anurans may extend several kilometers. Still, migrating anurans are remarkable for their fidelity to specific breeding sites and the precision with which individuals are able to return to the same few meters of shoreline each year.

Lizards very rarely migrate, but those that do also move to seek out suitable nesting sites. Green iguanas and a few related species that nest on islands may move as much as several kilometers to lay eggs at specific sites; the scarcity of appropriate soil in their island habitat presumably drives these migrations (Rand, 1968). The red-sided garter snake (*Thamnophis sirtalis parietalus*) migrates to winter hibernacula, such as limestone sinkholes, aggregating in the thousands to buffer the harsh winter conditions of Manitoba (Aleksiuk, 1976).

Insects and Other Terrestrial Invertebrates

Our current state of knowledge regarding insect migrations contains a number of spectacular examples but gives the overall impression that the phenomenon is not common in this group despite its huge contribution to our planet's biodiversity. Well-known examples include the Eurasian milkweed bug (*Lygaeus equestris*) and the North American ladybird beetle (*Coleomegilla maculata*), which perform seasonal movements from summer feeding and breeding sites to more protected areas a few kilometers away (Solbreck *et al.*, 1990). There they diapause and last out the winter surviving on stored fat. Individuals may coalesce at specific sites and form spectacular aggregations. As in other migratory species we have considered, migration allows both species to exploit rich, but seasonally variable, habitats. Because many insects have highly specialized feeding and breeding requirements, even small seasonal variations in temperature, moisture, and light levels may be sufficient to trigger migrations to sites where individuals can wait for local conditions to improve (i.e., diapause). The general migration pattern illustrated by ladybirds and milkweed bugs may thus be widespread among both temperate and tropical species. However, small insects are likely to migrate only short distances and be overlooked and understudied, especially if migrations occur without individuals aggregating into conspicuous groups.

An alternative migration pattern takes individuals from habitats declining in quality to richer sites that allowing feeding and breeding to continue. Among the noctuid, or armyworm, moths – a group found throughout the world – several species follow annual round-trip migratory circuits of several-hundred kilometers that take them to successively higher latitudes or to habitats recently freshened by rain with host vegetation more suitable for breeding (Gatehouse, 1987). In many cases, these species have become significant agricultural pests having a life history ideally suited to quickly exploiting ephemeral but rich food sources. Interestingly, single generations may complete only part of the migratory circuit, raising fascinating questions regarding the evolutionary genetics of migration timing and navigation. Perhaps the most famous migrant of this type is the monarch butterfly, *Danaus plexippus*, of North America (Brower and Malcolm, 1991). Monarch larvae feed on milkweeds and the timing of migration coincides with seasonal milkweed growth. In autumn, adult butterflies from as far north as eastern Canada migrate more than 1000 km south to overwinter in huge aggregations at only a few critical sites in the mountains of central Mexico. In spring, these same adults migrate 200–300 km to the northern edge of the Gulf of Mexico where they breed and die. Successive generations then move north tracking the milkweed growing season.

When large-scale migrations occur in insects, they are truly astounding behaviors if one considers the size of the travelers and the distances covered. Because insects have only limited capacity for fuel storage, they depend to a great extent more than other migrants on winds to propel and direct their movements. North American aphids, for example, fly for 2–24 h and travel from 50 to 1100 km before setting down. Differing wind speeds at different altitudes account for the variation in flight time and distance moved, but aphids probably exert some control over these features by choosing an altitude at which to fly.

In response to deteriorating local conditions, either due to seasonal reductions in food or increased densities, some insects develop long-winged (i.e., macropterous) forms able to travel relatively large distances to find more suitable habitat. Examples include planthoppers and aphids. These irruptive movements also occur in other taxa, such as birds, but they resemble dispersal processes more than the migrations we have so far considered.

Terrestrial crustaceans, such as hermit crabs and ghost crabs, must migrate from nonbreeding, feeding habitats to high-salinity water in order to breed. These movements occur seasonally, especially in response to tidal cycles, and may require movements from 10 to 3000 m each way (Wolcott and Wolcott, 1985).

Consequences of Migration

When animals collectively move in directed ways from one area to another, they connect disparate habitats. When they do so they not only can transfer large amounts of nutrients between ecosystems, but also they can transfer other species as well. And when these hitch hikers are pathogens, the implications for the dynamics of infectious diseases can be profound (Altizer *et al.*, 2011). Cases exist of migrants enhancing the movement of pathogens, including those that may affect human health. Some of the best known are Lyme disease and West Nile virus that moved up and down the east coast of the US from a point of origin in New York City (Rappole *et al.*, 2000; Owen, 2006) and Saiga deer of Kazakhastan that become infected with intestinal worms after moving into areas previously grazed by domestic sheep and then once infected continue to pass the parasites on to other sheep herds along their migratory route (Morgan *et al.*, 2000). But the importance of such transmission to other species may be overestimated. Because pathogens may kill infected migrants before they arrive at their destination or disease agents may be purged from the migrant's system before arrival by virtue of the pathogen's own biology, the actual extent of migrant assisted disease transmission may be limited (Altizer *et al.*, 2011).

Yet migrants are likely to suffer high rates of infection while en route because few species travel nonstop. By stopping to "refuel", at widely spaced and highly restricted sites, migrants create large aggregations that increase disease spread from one migrant to another. And if humans continue to degrade habitats and reduce the number of stop over points, remaining stop over points are likely to become even more effective at spreading disease.

Changes in climate are also likely to alter the way migrants influence disease dynamics. On the one hand, if humans continue to change climate so that migratory ranges expand, then new contacts among previously separated species are likely to have dramatic negative consequences. On the other hand, massive changes in climate or habitats could reduce the benefits of migration so that migration ceases, thus subjecting the transformed sedentary population to higher levels of infection. In monarch butterflies, for example, parasite prevalence and virulence increase with longer residency and more intense habitat use. Thus, milder climates could lead to greater and more devastating infections in emerging nonmigratory populations (Altizer *et al.*, 2000) or if habitat change positions increased parasite reservoirs in areas of species overlap, then greater species to species transmission is likely. Such a process of sedentarization resulting from local habitat enhancement has been implicated in triggering the spread of Nipha virus from flying foxes to pigs to humans (Plowright *et al.*, 2008). Whether or not environmental change enhances or limits the spread of disease, will depend on life history traits of parasites and hosts and contact rate among migrants and other members of the communities they connect.

Attributes of Migrants Affecting Susceptibility to Human Disturbance

As we have seen, migrations appear to serve a few common functions. In general, they enable species to temporarily avoid harsh conditions or to meet important biological needs that are separated by great distance. Not surprisingly, there are certain attributes of migrants that affect their susceptibility to human disturbance.

First, migratory species often use a variety of habitats, leaving them vulnerable to multiple points of disturbance. Often harm is felt mostly at one destination. Bachman's warbler (*Vermivora bachmanii*), for example, was driven to extinction by the destruction of its overwintering habitats in the tropics. For others, however, such as many Neotropical migrant birds, the impacts of human activities are felt at both endpoints of their migrations. Globe-trotting species like the white-rumped sandpiper typify the precarious dependence of migrants on habitat health across a tremendous geographic scale. But harm need not be limited to the endpoints of a migrant's journey. Any diminution in quality of refueling sites along the way could winnow a population and limit its ability to replenish its numbers before the next cycle of migration. And given that migratory trajectories for many species are shaped by the vagaries of prevailing winds or currents, conservation strategies entailing the protection of all stopover areas becomes almost impossible. Even where resting points are protected, unintended consequences associated with the normal nonintrusive behavior of naturalists can put migrants at risk. Steady viewing by bird watchers at refuges along coastal fly routes has been known to force birds to move too far offshore where they can no longer feed on inshore marine invertebrates exposed at low tides.

Sea turtles probably illustrate the best effects of migratory species being vulnerable at many life-history stages to the excesses of human behavior. Not only are their eggs sought after and easily harvested by indigenous peoples, but also the beaches themselves are often degraded by the activities of affluent humans. Entire breeding populations are eliminated when beach habitats are developed or severely compromised in their abilities to launch young when dune buggies destroy nests or excessive night lighting disrupts the water-seeking behavior of newly hatched young. In addition, for those young that mature to subadults and return to the breeding grounds, nets of fishermen, particularly shrimpers, provide the *coup de gras* to the species by drastically reducing the pool of future breeders. In fact, it has been shown for that the most vulnerable period in the life cycle is not the nestling, but rather the subadult, stage (Crouse *et al.*, 1987). Although protecting beaches and increasing the number of functioning beaches gives the species a head start by diversifying recruitment sites, sea turtles are most vulnerable to extinction if the number of subadults is reduced. It is at this point in the species' life history where a long life of breeding commences and reproductive value is highest. Only by insisting that shrimpers insert turtle excluder devices (TEDs) in nets so that the turtles can escape before drowning will there be any hope that those migratory species already endangered will survive.

Second, migratory species often aggregate when traveling. While this behavior might reduce each individual's risk of falling prey to nonhuman predators, both aquatic and terrestrial species in groups are easy targets for the advanced harvesting gear employed by commercial fishers and hunters. In prehistoric times, hunters stampeded herd animals off cliffs and into canyons where massive kills occurred. Such actions have been implicated in the extinction of many of the North American megafauna. Today even the smaller, more elusive, and difficult to capture prey are at risk. Many marine fisheries involving migratory schooling fishes, such as those of cod and haddock, have been overfished. And despite moratoria, many are not recovering.

Third, since many migratory species move according to very precise schedules, any delays caused by human alteration of the habitat or disruption of normal movements could lead to a cascade of effects, jeopardizing a species' ability to replenish its population numbers reduced by ordinary mortality. As noted earlier, migrating shorebirds if forced to linger longer at refueling refuges, could have breeding seasons shortened sufficiently to lose the ability to lay one or more clutches. And since selection favors renesting because nest predation rates are already high, any force constraining such efforts could severely limit a species' recruitment potential.

Despite these attributes that threaten species survival there are some beneficial traits exhibited by migrants. By occupying, at least temporarily, many different habitats, migrants can spread risks and thus can escape catastrophes that befall one location. Unless all individuals in a species simultaneously occupy the same habitat, survivors from habitats not impacted

by the catastrophe can serve as sources for increasing species numbers. Migratory species are also more likely to find and colonize new habitats as they are opened up by climate change or other human-induced and natural changes in the environment.

Ecological Consequences of Human Disturbance on Migrants

As humans fragment landscapes, many migratory species will find themselves in peril. For species that must move from one region to another in order to meet a variety of biological needs, any barrier disrupting these movements could lead to local extinctions. In southern Africa, fences are being built either to prevent mixing of wildlife and livestock and the concomitant transmission of disease (ensuring that exported beef is disease free) or to prevent wildlife from gaining a competitive edge over livestock when both compete for critical resources such as food near water. Without access to these resources at the time of year most critical for developing juveniles, any migratory population's recruitment will be severely curtailed. Many local populations will either go locally extinct or will be transformed from "sources" of new emigrating individuals to "sinks" where excess individuals from healthy populations find refuge. Even for populations that still boast numbers in the hundreds of thousands, such as the plains zebra or even African elephants (*Loxodonta africana*), fragmentation of the large populations is well underway. With human populations expanding into habitats only marginally suited for horticulture, the cry for fencing areas to prevent crops being consumed by migrating herbivores is increasing.

When the movements of these populations are disrupted, a cascade of indirect effects on both the recycling of nutrients in ecosystems and the structuring of animal communities are likely to occur. Since migratory species represent some of the largest aggregations and highest densities of individuals seen in the animal world, any disruption to migration is likely to reduce local density. Thus, it is quite possible that the important impact on the structure of vegetation, or more generally the recycling of nutrients, will be altered (Augustine and McNaughton, 1998). It has even been proposed that the overharvesting of sperm whales and the removal of their large carcasses from the ecosystem have impacted the community stability of the deep-sea communities associated with deep-sea vents. By preventing carcasses from falling to the ocean floor, an important source of renewing "island" resources has been eliminated, thus making it more difficult for species tightly associated with vents to hop from one vent to another (Butman *et al.*, 1995). With respect to structuring communities, if plains zebras are unable to move in large numbers across the landscape from the tops of catenas where they forage during the rains to the wetter valleys where they move when the rains cease, then their facilitative effect on diversifying the herbivore community will be reduced. Typically, plains zebra move to the wetter valleys and graze down the tougher, fibrous forage that dominates these areas making the greener and more nutritious forage available for the ruminants that require such higher quality vegetation. Thus, by disrupting the movements of such a "keystone" species, the diversity of large grazers could be reduced (Rubenstein, 2010). Furthermore, exclosure experiments suggest that the grazing community could shift from one dominated by large ungulates to one populated by small rodents and lagomorphs (Keesing, 1998). Since rodents often come in contact with people and harbor human diseases, the cascade of effects could be quite profound.

Perhaps one of the greatest impacts of disrupting movements will be seen via the effects that elephants have on the landscape. As populations become restricted to reserves if their numbers are allowed to increase, then their ability to transform a mosaic landscape of trees and grasslands into one of mostly grasslands will mean the disappearance of resident forest- and woodland-dwelling species of colobine monkeys and antelopes.

Even for those subdivided populations that can adapt to the nonmigratory habit – and in certain places such as the Ngorongoro Crater in Tanzania, resident and migratory populations of zebra, Thompson's gazelle, and wildebeest coexist – genetic effective population sizes will shrink and the loss of heterozygosity could leave the population vulnerable to the emergence of novel diseases or the devastating effects of inbreeding. The trend to isolate migratory wildlife into reserves that are too small, although serving as an immediate panacea to human population expansion and development, could have adverse long-term consequences for the health and survival of species unless they are designed to be of appropriate size or have appropriate corridors for migration included in their design. The creation of multinational parks in South Africa, Mozambique, and Zimbabwe that include important mixes of habitats required by a wide variety of migratory species is on the drawing board and offers some hope that subdivision and isolation of populations will be reduced.

Migratory species are also likely to be directly affected by global climate change. It has been suggested that the arctic tundra as a habitat might disappear (Peters, 1991). If the upper 2–3 m of permafrost, which helps bind the tundra together, were to disappear, then the tundra as a habitat would vanish. The effects would drastically reduce populations of caribou that traverse large distances across it as they move from summer breeding to more protective wintering feeding sites. Such transformations of animal communities have occurred in the past when climates have dramatically changed and migratory species were excluded from important parts of their ranges. For example, much of the Canadian–Alaskan megafauna vanished 10–20 million years ago when the forests closed and sedge and other patchy but nutrient-rich herbaceous species disappeared. But the scope of these impacts is likely to be greater in the near future because both the magnitude and rate of change associated with the climate effect is increasing and is being coupled with widespread land-use changes that are occurring in areas where the migrants will be forced to move.

Evolutionary Consequences of Human Impacts on Migrants

While it is clear that human behavior disrupting migrations will have ecological consequences, the effects on the evolutionary potential, or even the future characteristics, of

species are also likely to be large. Nowhere is this impact more evident than in North America where the range of the bighorn sheep has been severely reduced and fragmented. Bighorn migrate seasonally from meadows to mountaintops to avoid harsh climate and in search of grass that changes seasonally in quality at each location. As human populations have expanded, populations have been isolated into refuges, numbers are shrinking, and genetic diversity is being lost (Primack, 1993). In addition, hunting for trophy males is removing just those individuals with the best and most diverse genotypes. As a result, the ability of the isolated populations to remain genetically diverse enough to forestall problems associated with inbreeding or to avoid the ravages that would occur if a novel disease appeared has been severely compromised.

Migratory Atlantic salmon (*Salmo salar*) provide perhaps the clearest example of where human activities are causing a species to change under our own eyes. Atlantic salmon typically are born in fresh water and develop for 1–2 years before smolting and heading to sea to grow and fatten by feasting on a rich supply of marine invertebrates. Once they attain a certain size, they become sexually mature and return to rivers. There they travel long distances upstream until they reach clear and cool breeding grounds. Some individuals, however, never head to the sea. Instead, they remain in their natal streams and mature sexually at young ages and at small sizes. Such individuals are called parr and they never go through the smolting process that adapts them to a marine lifestyle. Under pristine environmental conditions, the fraction of the population that becomes parr is small since the reproductive gains of such a strategy are low. When competing for mates with larger more aggressive males, parr fare poorly. What matings they obtain are derived by "sneaking" among a mating pair and releasing milt at just the right time. Such events are rare and the mixing of milt is poor so such sneak matings result in few young being sired by parr. But the survival prospects of parr are high since they do not incur the risks of going to sea and, moreover, they begin breeding at a very early age. Thus, over a lifetime, reproductive success is modest. But as human fishing increases and the netting of the older larger salmon intensifies, the *relative* lifetime reproductive success of parr is improving. Since the costs of migrating to sea and back again have increased dramatically, the long life span of parr gives them a relative advantage. And given that competition with the larger males for mates is also declining, the chances of parr securing matings are also improving. Thus, it is not surprising that the composition of the population is changing as parr increase markedly in abundance. The impact on the long-term stability of the population is unclear. But the long-term impact on the species appearance is clear, just as it is for yet another fishing industry that will go into decline as the large fish disappear.

Appendix

List of Courses

1. Animal Behavior
2. Behavioral Ecology
3. Natural History

See also: Birds, Biodiversity of. Dispersal Biogeography. Moths

References

Aleksiuk M (1976) Reptilian hibernation: Evidence of adaptive strategies in Thamnophis sirtalis parietalis. *Copeia* 170–178.

Alerstam T (1985) Strategies of migratory flight illustrated by Arctic and common terns. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*. Port Aransas, TX: Marine Science Institute.

Alexander WB and Lack D (1944) Changes in status among British breeding birds. *British Birds* 38: 42–88.

Altizer S, Bartel R, and Han BA (2011) Animal migration and infectious disease risk. *Science* 331: 296–302.

Altizer S, Oberhauser K, and Brower LP (2000) Association between host migration and prevalence of a protozoan parasite in natural populations of adult monarch butterflies. *Ecological Entomology* 25: 125–139.

Angel MV (1985) Vertical migrations in the oceanic realm: Possible causes and probable effects. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*. Port Aransas, TX: University of Texas Marine Institute.

Augustine D and McNaughton S (1998) Ungulate effects on the functional species composition of plant communities: Herbivore selectivity and plant tolerance. *Journal of Wildlife Management* 62(4): 1165–1183.

Baird J (1999) Returning to the tropics: The epic autumn flight of the blackpoll warbler. In: Able KP (ed.) *Gathering of Angels: Migrating Birds and Their Ecology*. Ithaca, NY: Cornell University Press.

Baker RR (1978) *The Evolutionary Ecology of Animal Migration*. London: Hodder & Stoughton.

Barnard CJ and Sibly RM (1981) Producers and scroungers: A general model and its application to captive flocks of house sparrows. *Animal Behaviour* 29: 543–550.

Berthold P (1993) *Bird Migration*. New York: Oxford University Press.

Best PB (1979) Social organization of sperm whales, *Physeter macrocephalus*. Behavior of marine animals. In: Winn HE and Olla BL (eds.) *Current Perspectives in Research, Vol. 3: Cetaceans*, pp. 227–289. New York: Plenum.

Bowen WD and Siniff DB (1999) Distribution, population biology, and feeding ecology of marine mammals. In: Reynolds JEI and Rommel SA (eds.) *Biology of Marine Mammals*, pp. 423–484. Washington, DC: Smithsonian Institution Press.

Boyd IL, Lockyer C, and Marsh HD (1999) Reproduction in marine mammals. In: Reynolds JEI and Rommel SA (eds.) *Biology of Marine Mammals*, pp. 218–286. Washington, DC: Smithsonian Institution Press.

Brett JR (1986) Production energetics of a population of sockeye salmon, *Oncorhynchus nerka*. *Canadian Journal of Zoology* 64: 555–564.

Brower LP and Malcolm SB (1991) Animal migrations; endangered phenomena. *American Zoologist* 31: 265–276.

Butman A, Carlton JT, and Paulmbi SR (1995) Whaling effects on deep-sea biodiversity. *Conservation Biology* 9(2): 462–464.

Crouse DT, Crowder LB, and Caswell H (1987) A stage-based population model for loggerhead Sea Turtles and implications for conservation. *Ecology* 68(5): 1412–1423.

Curry-Lindahl K (1981) *Bird Migration in Africa*, vols. 1–2. New York: Academic Press.

Dingle H (1996) *Migration: The Biology of Life on the Move*. New York: Oxford University Press.

Faaborg J (1988) *Ornithology: An Ecological Approach*. Englewoods Cliffs, NJ: Prentice Hall.

Fenton MB and Thomas DW (1985) Migrations and dispersal of bats (Chiroptera). In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*, pp. 409–424. Port Aransas, TX: University of Texas Marine Institute.

Gatehouse AG (1987) Migration and low population density in armyworm (Lepidoptera: Noctuidae) life histories. *Insect Science and its Applications* 8: 573–580.

Gauthreaux Jr. SA (1982) Ecology and evolution of avian migration systems. In: Farner DS, King JR, and Parkes KC (eds.) *Avian Biology* 6. pp. 93–168. New York: Academic Press.

Gentry RL (1998) *Behavior and Ecology of the Northern Fur Seal*. Princeton, NJ: Princeton University.

Griffin DR (1970) Migrations and homing of bats. In: Wimsatt WA (ed.) *Biology of Bats*, vol. 1, pp. 233–264. New York: Academic Press.

Griswold CK, Taylor CM, and Norris DR (2010) The evolution of migration in a seasonal environment. *Proceedings of the Royal Society of London Series B* 277: 2711–2720.
Gross MR, Coleman RM, and McDowall RM (1988) Aquatic productivity and the evolution of diadromous fish migration. *Science* 239: 1291–1293.
Gruell GE and Papez NJ (1963) Movements of mule deer in northeastern Nevada. *Journal of Wildlife Management* 27: 414–422.
Guttal V and Couzin ID (2011) Social interactions, information use and the evolution of collective migration. *Proceedings of the National Academy of Sciences* 107(37): 16172–16177.
Harrington BA (1999) The hemisperic globetrotting of the white-rumped sandpiper. In: Able KP (ed.) *Gatherings of Angels: Migrating Birds and Their Ecologies*. Ithaca, NY: Cornell University Press.
Hayman P, Marchant J, and Prater T (1986) *Shorebirds: An Identification Guide to the Waders of the World*. Boston: Houghton Mifflin.
Helfman GS, Facey DE, Hales LS, and Boyeman EL (1987) Reproductive ecology of the American eel. *American Fisheries Society Symposium* 1: 42–56.
Herrnkind WF (1985) Evolution and mechanisms of mass single-file migration in spiny lobster: Synopsis. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*, pp. 197–211. Port Aransas, TX: University of Texas Marine Institute.
Huntley M (1985) Experimental approaches to the study of vertical migration of zooplankton. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*, pp. 71–90. Port Aransas, TX: University of Texas Marine Institute.
Johnson DW and McFarlane RW (1967) Migration and bioenergetics of flight in the Pacific golden plover. *Condor* 69: 156–168.
Johnson SR and Herter DR (1990) Bird migration in the arctic: A review. In: Gwinner E (ed.) *Bird Migration: The Physiology and Ecophysiology*, pp. 22–43. New York: Springer-Verlag.
Karr JR (1980) Patterns in the migration system between the north temperate zone and the tropics. In: Keast A and Morton ES (eds.) *Migrant Birds in the Neotropics: Ecology, Behavior, Distribution, and Conservation*. Washington, DC: Smithsonian Institute Press.
Keesing F (1998) Impacts of ungulates on the demongraphy and diversity of small mammals in central Kenya. *Oecologia* 116: 381–389.
Kerlinger P and Moore FR (1989) Atmospheric structure and avian migration. *Current Ornithology* 6: 109–142.
Ketterson ED and Nolan Jr. V (1983) The evolution of differential bird migration. *Current Ornithology* 1: 357–402.
Lövei GL (1989) Passerine migration between the Palearctic and Africa. *Current Ornithology* 6: 143–174.
Maddock L (1979) The migration and grazing succession. In: Sinclair ARE and Norton-Griffiths M (eds.) *Serengeti: Dynamics of an Ecosystem*, pp. 104–129. Chicago: University of Chicago.
Mårtensson P-E, Lager Gotaas AR, Nordoy ES, and Blix AS (1996) Seasonal changes in energy density of prey of Northeast Atlantic seals and whales. *Marine Mammal Science* 12: 635–640.
Martinez MM (1983) Nidificación de *Hirundo rustica* erythrogaster (Boddaert) en la Argentina (Aves, Hirundinidae). *Neotropica* 29: 83–86.
McCleave JD and Kleckner RC (1985) Oceanic migrations of Atlantic eels (Anguilla spp.): Adults and their offspring. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*, pp. 316–337. Port Aransas, TX: University of Texas Marine Institute.
McCullough DR (1985) Long range movements of large terrestrial mammals. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*, pp. 444–465. Port Aransas, TX: University of Texas Marine Institute.
McDowall RM (1988) *Diadromy in Fishes*. Portland, OR: Timber.
McKeown BA (1984) *Fish Migration*. Portland, OR: Timber Press.
Meylan AB, Bowen BW, and Avise JC (1990) A genetic test of the natal homing versus social facilitation models of green turtle migration. *Science* 248: 724–727.
Moreau RE (1972) *The Palaearctic-African Bird Migration Systems*. New York: Academic Press.
Morgan ER, Medley GF, Torgerson PR, Shaikenov BS, and Milner-Gulland EJ (2000) Parasite transmission in a migratory multiple host system. *Ecological Modelling* 200: 511–520.
Myers JP, Marion JL, and Sallaberry M (1985) Going to extremes: Why do sanderlings migrate to the Neotropics? *Ornithological Monographs* 36: 520–532.
Nowak RM (1994) *Walker's Bats of the World*. Baltimore: Johns Hopkins University Press.
Owen J, Moore F, Panella N, *et al.* (2006) Migrating birds as dispersal vehicles for West Nile virus. *EcoHealth* 3: 79–85.
Peters R (1991) *Consequences of Global Warming for Biodiversity*. New Haven: Yale University Press.
Pierson ED (1998) Tall trees, deep holes, and scarred landscapes: Conservation biology of North American bats. In: Kunz TH and Racey PA (eds.) *Bat Biology and Conservation*. Washington, DC: Smithsonian Institute Press.
Plowright RK, Field HE, Smith C, *et al.* (2008) Reproductive and nutritional stress are risk factors for Hendra virus infection in little red flying foxes(*Pteropus scapulatus*). *Proceedings of the Royal Society of London Series B* 275: 861–869.
Primack R (1993) *Essentials of Conservation Biology*. Boston, MA: Sinauer Associates.
Rand AS (1968) A nesting aggregation of iguanas. *Copeia* 1968: 552–561.
Rappole JH, Derrickson SR, and Hubalek Z (2000) Migratory birds and spread of West Nile virus in the western hemisphere. *Emerging Infectious Dieseases* 6(4): 320–328.
Reynolds JEI (1999) Efforts to conserve the manatees. In: Twiss JRJ and Reeves RR (eds.) *Conservation and Management of Marine Mammals*, pp. 267–295. Washington, DC: Smithsonian Institution.
Richards GC (1995) A review of ecological interactions of fruit bats in Australian ecosystems. Ecology, Evolution, and Behaviour of Bats. In: Racey PA and Swift SM (eds.) *Symposium of the Zoological Society of London*, vol. 67, pp. 79–96. Oxford: Clarendon Press.
Roe HSJ (1983) Vertical distributions of euphausiids and fish in relation to light intensity in the northeast Atlantic. *Marine Biology* 77: 287–298.
Rubenstein DI (1981) Individual variation and competition in the everglades pygmy sunfish. *Journal of Animal Ecology* 50(2): 337–350.
Rubenstein DI (2010) Ecology, social behavior, and conservation in zebras. In: Macedo R (ed.) *Advances in the Study Behavior: Behavioral Ecology of Tropical Animals,* vol. 42, pp. 231–258. Oxford, UK: Elsevier.
Saldana J and Venables B (1983) Energy compartmentalization in a migratory fish, *Prochilodus mariae* (Prochilodontidae), of the Orinoco River. *Copeia* 1983: 617–625.
Shaw AK and Levin SA (2011) To breed or not to breed: A model of partial migration. *Oikos* 120(12): 1871–8179.
Sheppe WA (1972) The annual cycle of small mammal populations on a Zambian flood plain. *Journal of Mammalogy* 53: 445–460.
Solbreck C, Anderson DB, and Förare J (1990) Migration and the coordination of life cycles as exemplified by Lygaeinae bugs. In: Gilbert F (ed.) *Insect Life Cycles: Genetics, Evolution, and Coordination*, pp. 197–214. London: Springer-Verlag.
Stebbins RC and Cohen NW (1995) *A Natural History of Amphibians*. Princeton, NJ: Princeton University.
Stenseth NC (1983) Causes and consequences of dispersal in small mammals. In: Swingland I and Greenwood PJ (eds.) *The Ecology of Animal Movement*, pp. 63–101. Oxford: Oxford University Press.
Stewart BS and DeLong RL (1995) Double migrations of the northern elephant seal, *Mirounga angustirostris*. *Journal of Mammalogy* 76: 196–205.
Taylor LR and Taylor RAJ (1983) Insect migration as a paradigm for survival by movement. In: Swingland IR and Greenwood PJ (eds.) *The Ecology of Animal Movement*, pp. 181–214. Oxford: Clarendon Press.
Terborgh J (1989) *Where Have All the Birds Gone?* Princeton, NJ: Princeton University Press.
Thomas DW (1983) The annual migrations of three species of west African fruit bats (Chiroptera: Pteropodidae). *Canadian Journal of Zoology* 61: 2266–2272.
Twitty V, Grant D, and Anderson O (1967) Home range in relation to homing in the newt *Taricha rivularis* (Amphibia: Caudata). *Copeia* 1967: 649–653.
Wells RS, Boness DJ, and Rathbun GB (1999) Behavior. In: Reynolds JEI and Rommel SA (eds.) *Biology of Marine Mammals*, pp. 324–422. Washington, DC: Smithsonian Institution Press.
Wetmore A (1926) *The Migration of Birds*. Cambridge, MA: Harvard University Press.
Williams TC and Williams JM (1978) An oceanic mass migration of land birds. *Scientific American* 239(4): 166–176.
Wolanski E, Gereta E, Borner M, and Mduma A (1999) Water, migration and the Serengeti ecosystem: Understanding the mechanisms that control the timing of wildlife migrations may prove vital to successful management. *American Scientist* 87(6): 526–533.
Wolcott TG and Wolcott DL (1985) Factors influencing the limits of migratory movements in terrestrial crustaceans. In: Rankin MA (ed.) *Migration: Mechanisms and Adaptive Significance*, pp. 257–276. Port Aransas, TX: University of Texas Marine Institute.
Wootton RJ (1990) *Ecology of Teleost Fishes*. New York: Chapman & Hall.

Modeling Biodiversity Dynamics in Countryside and Native Habitats

Henrique M Pereira and Luís Borda-de-Água, Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity dynamics models Models describing the dynamics of one or more measures or dimensions of biodiversity (e.g., species diversity or population abundances).

Countryside habitats A landscape made of native and human-dominated habitats.

Human-dominated habitat A habitat that has been significantly altered by humans. Significant habitat alterations include the conversion of natural forest into a pasture or the reclamation of a wetland for agriculture.

Land use The set of human activities within a landscape and their impacts on natural processes, such as physical, chemical, and biological cycles. In practice, land use is often classified in major categories such as cropland, pasture, plantation forest, and natural forest.

Metapopulation model A model of a population restricted to habitat patches immersed in a matrix used only for dispersal. The population in each habitat patch may go extinct through stochastic processes and recolonized by individuals dispersing from other patches.

Native habitat Native habitat is the community of species and the physical environment that occurred in a region before human intervention. In this article, it is equated to the naturally occurring vegetation in a region. In other contexts, it refers to the particular natural habitat of a given species.

Phenomenological models Models that relate observable variables without postulating the underlying (physical, biological, etc.) mechanisms. For example, the species–area relationship is a phenomenological model because it does not explicitly state the mechanisms leading to the relationship between the number of species and area.

Process-based models Models explicitly based on mathematical representations of processes or mechanisms, such as animal or seed dispersal. For example, population dynamics models describe the evolution over time of the population abundance by accounting for processes such as birth, death, and dispersal.

Species–area relationships The relationship between the size of a given habitat or area sampled across habitats, and the number of species (typically of a specific taxonomic group) found in that area.

Source-sink population model A model of a population that occupies multiple habitats in the landscape connected trough dispersal. The habitats where population growth is positive are sources and the habitats where growth is negative are sinks.

Introduction

Despite international efforts and commitments to reduce biodiversity loss, biodiversity continues to decline at high rates compared to those in the background fossil record (Butchart *et al.*, 2010; Barnosky *et al.*, 2011). Land-use change is currently the main driver of biodiversity loss in terrestrial ecosystems and is likely to remain so over the next few decades (Duraiappah *et al.*, 2005; Pereira *et al.*, 2010). It is therefore essential to assess how different patterns of land-use affect biodiversity, in order to compare the biodiversity costs and benefits of different policy and management options.

Much of the recent modeling efforts on biodiversity dynamics of global change have looked at the impacts of climate change (Thuiller *et al.*, 2008; Pereira *et al.*, 2010). However, many of the original projections of global extinction rates were developed from forest loss estimates (Pimm *et al.*, 1995), and there is a renewed interest in models linking biodiversity response to land-use change at different scales (Pereira and Daily, 2006; Wright, 2010; He and Hubbell, 2011). Some of this renewed interest comes from the realization that although the conversion of native habitats to agriculture and other human-dominated habitats is followed by biodiversity loss (Brooks *et al.*, 1999, 2003), countryside habitats can also support important levels of biodiversity, particularly when forest patches or other complex vegetation features remain (Daily *et al.*, 2003; Ranganathan *et al.*, 2008). Another important issue is that global land-use change dynamics are increasingly complex. While forest loss continues to occur in tropical forests and subtropical woodlands, some regions of the world are seeing an expansion of forest trough natural vegetation regeneration or plantations (van Vuuren *et al.*, 2006; UNEP, 2007; FAO, 2010). Therefore, there is a need for models that can capture not only the biodiversity consequences of native forest loss, but also the potential benefits of natural forest expansion or the implications of forest plantations for biodiversity.

Here we review existing models to study the response of biodiversity to land-use change. These models can be grouped in two major categories: phenomenological and process-based (Pereira *et al.*, 2010). Phenomenological models use empirical relationships between habitat characteristics such as area and land-use intensity and species diversity or other biodiversity metrics. The most widely used phenomenological model is the species–area relationship (SAR), and it illustrates the approach of this type of model. There is a well-known relationship between the size of an area and the number of species in that area (Rosenzweig, 1995). The assumption is that the SAR can be applied to project the number of extinctions as a consequence of loss of area of native habitat, even if we do not explicitly model the processes that cause the relationship between the size of an area and its number of species. In contrast, process-based models try to capture the processes that

underlie the relationship between land use and biodiversity. For example source-sink models (Doak, 1995; Pereira *et al.*, 2004) describe the population dynamics of species in habitats of good quality and habitats of poor quality and the movement of individuals between those habitats. Land-use change modifies the proportion and the spatial configuration of the habitats in the landscape. Therefore, in source-sink models there is an explicit link between the land-use change processes and species population dynamics.

Phenomenological Models

Native Habitats: Species–Area Relationship

It has been long known by ecologists that the number of species in a given area increases with the size of the area (Arrhenius, 1921). Several expressions have been proposed to describe the relationship between number of species, S, and the size of an area, A, with the most common one being the power law (Dengler, 2009):

$$S(A) = cA^z \quad [1]$$

where z depends on the sampling regime and scale and c depends on the taxonomic group and region (Rosenzweig, 1995). Therefore, if there is a habitat loss of area a from a total area A, the proportion of species going extinct can be estimated as

$$\frac{S(A) - S(A-a)}{S(A)} = 1 - \left(1 - \frac{a}{A}\right)^z \quad [2]$$

The problem is that, in general, (1) is estimated from patterns that were generated by very different processes from the habitat loss patterns that (2) aims at describing. There are two general types of sampling for the SAR (Drakare *et al.*, 2006; Dengler, 2009): nested sampling and isolate sampling. In nested sampling, progressively larger areas of a continuous ecosystem are sampled, usually nested within each other (**Figure 1(a)**). The SARs generated by this type of sampling are also known as continental SARs (Rosenzweig, 1995), as in most cases they come from studies on continents. In isolate sampling (**Figure 1(b)**), habitat isolates with different area sizes are sampled. Initially, most of the SARs using isolate sampling came from island archipelagos, and therefore this type of SAR has been named the island SAR (Rosenzweig, 1995). A fundamental difference between nested and isolate SARs is that in nested SARs species number must always increase with the area size whereas in isolate SARs species number may increase or decrease with the area size.

The z-values of the SAR, key parameter to calculate extinction projections in [2], are often higher in oceanic islands than in nested areas in a continent. For instance, Rosenzweig (1995) gives a range of 0.25–0.35 for the former and 0.12–0.18 for the latter (**Figure 2**). However, recent studies have shown that the nested SAR exhibits a much broader range of z-values, from about 0.10 to more than 0.40 (Sala *et al.*, 2005; Drakare *et al.*, 2006), and that the z-values are particularly dependent on the scale of the study (Crawley and Harral, 2001; Pereira *et al.*, 2012). Furthermore, the z-value of isolates

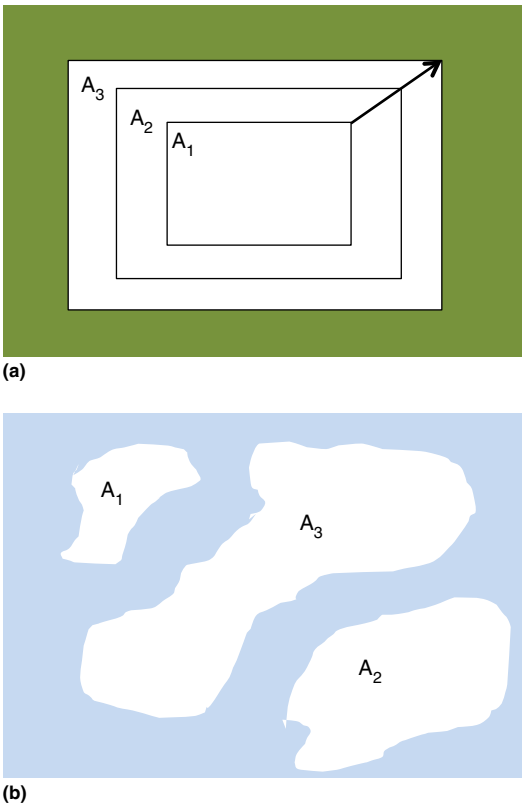


Figure 1 The sampling scheme of the different types of species–area relationship: (a) nested sampling or continental SAR; (b) isolate sampling or island SAR.

on mainlands are often lower than the z-values of true islands in an ocean (Drakare *et al.*, 2006).

The species that are likely to go extinct immediately after habitat loss are the species endemic to the region of destroyed habitat (Kinzig and Harte, 2000; He and Hubbell, 2011). If the loss of the habitat happens in the periphery, then the nested SAR will describe the number of extinctions. This happens because the number of endemics in an outer ring of the habitat of size a is given by [2], assuming that the z-value of the SAR in [1] was estimated as in **Figure 1(a)** (Pereira *et al.*, 2012). If the habitat loss occurs in the center, the z-value from the nested SAR may overestimate the number of endemics. In that case, $z_{\text{isolate}} = 0.25$ the endemics–area relationship (EAR), built by counting the number of endemics in progressively larger areas (as in **Figure 1(a)**), will better describe immediate species extinction. The z-values of the EAR are generally lower than the z-values of the nested continental SAR (He and Hubbell, 2011).

Many species that were not confined to the areas of habitat loss may go extinct in the long term because the remaining area of habitat may not be sufficient to allow for their persistence (i.e., sink species, as their dynamics are dominated by the modified habitats in the landscape which act as sinks). The difference between the species that go immediately extinct and the species that are likely to go extinct in the future has been coined "extinction debt." This second stage of extinctions can take from decades to millennia (Pereira *et al.*, 2010), and it has been proposed that it is better described by the island SAR

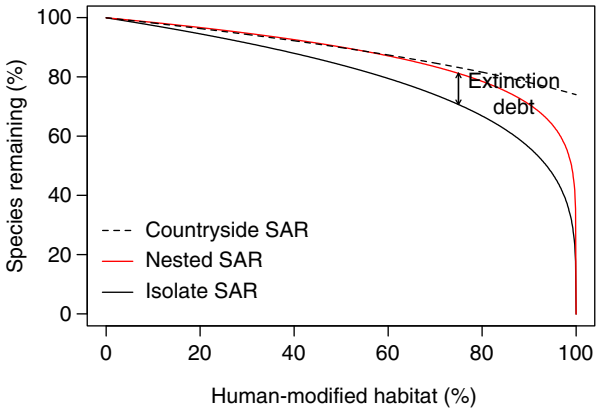


Figure 2 Comparison of the projections of species extinctions for the countryside SAR, the nested SAR and the isolate SAR. The EAR or the nested SAR (here $z_{nested}=0.15$) follow a similar shape, but with lower z-values than the isolate SAR (here $z_{isolate}=0.25$). Notice that when we use the classical SAR or the EAR all the species vanishes when the native habitat is completely lost. In the countryside SAR there are still species surviving when all the native habitat is converted. In this example we assumed that there are two groups of species (group 1 and 2) with the following parameters: $h_{1,Native}=1$, $h_{1,Human\text{-}dominated}=0.25$, $h_{2,Native}=1$, $h_{2,Human\text{-}dominated}=0.5$, $c_1=0.75$, $c_1=0.25$, and $z=0.25$.

(Rosenzweig, 2001). This happens because the island SAR is more the result of colonization/extinction dynamics than of the sampling effects that are intrinsic to the nested SAR and instantaneous extinctions. Note that another process, habitat turn-over, plays a role in both the isolate SAR and the nested SAR. SAR for multiple habitats, discussed in the next section, allows the study of the effects of habitat turn-over on extinction.

One problem with using the island SAR is that a fragment of native habitat surrounded by agricultural areas is ecologically different from an island surrounded by oceans: the agricultural matrix is more inhabitable than an ocean for most species. An alternative is to use the SAR from habitat isolates (e.g., in agricultural land) to estimate extinctions. A problem with this second approach is that, in many cases, time after fragmentation is still relatively short for the communities in sampled habitat isolates to have reached equilibrium. A third problem comes from the difference between what happens when a large region of native habitat is reduced and what are the processes determining species diversity in habitat isolates of different sizes in an agricultural matrix. In the latter, there can be source-sink dynamics between different patches while in the former the only source is the original habitat region; therefore, instead of a spatial source-sink dynamics we have a temporal source-sink dynamics. The problem is further compounded by the fact that the remaining habitat is likely to be itself fragmented.

Countryside Habitats: A Species–Area Relationship for Multiple Habitats

All of the SAR approaches described above assume that no species persist outside native habitats, and that when all habitat is converted to human-dominated habitats (e.g., agricultural lands) all species go extinct (**Figure 2**). In reality many species are able to persist in human-dominated landscapes. For instance, at least half of terrestrial Costa Rican bird species and nonflying mammal species use agricultural habitats (Daily *et al.*, 2001, 2003). Pereira and Daily (2006) proposed a model to describe the use of both human-dominated and native habitats by many species: the countryside species–area relationship (countryside SAR).

The countryside SAR tracks the number of species in groups of species with similar affinities for the habitats in the landscape. The number of species in group i, S_i, is given by

$$S_i = c_i\left(\sum_{j=1}^{n} h_{ij}A_j\right)^{z_i} \quad [3]$$

where h_{ij} is the affinity of group i to habitat j, A_j is the area covered by habitat j, and n is the number of habitats. When the affinity of group i to the preferred habitat k is normalized to unity ($h_{ik}=1$), the habitat affinity h_{ij} becomes a conversion factor from the area of habitat j into an equivalent area of preferred habitat. The total number of species in the landscape is given by

$$S = \sum_{i=1}^{m} S_i \quad [4]$$

where m is the number of species groups considered.

The countryside SAR can be used to project the biodiversity response to scenarios of land-use change (Proença *et al.*, 2009). In contrast with the classic SAR, a proportion of species may remain in the landscape even if all native habitat is converted to human-dominated habitats (**Figure 2**). The countryside SAR suggests that neither the isolate SAR nor the nested SAR describes extinctions. Instead, the equilibrium number of species in a modified landscape is higher than what is predicted by either of these two curves. Extinction debt may occur, but it will be given by the difference between immediate extinctions caused by habitat conversion (the endemic species that cannot survive in the new habitat, which will be fewer than the ones projected by the EAR) and the species that will go extinct in the medium to long-term because they do not have enough source habitats (i.e., sink species). The countryside SAR is likely to describe the short-term extinctions better, but more research is needed.

The countryside SAR can be used to project biodiversity changes, not only from the loss of native habitat but also from an increase in native habitats, such as natural forests (Proença *et al.*, 2009), or other habitats such as forest plantations, as long as those habitats are incorporated into the model and the affinities of species for those habitats are calculable. The countryside SAR [3] has been compared with the classic SAR [1] in mosaic landscapes with multiple habitats, both at the local scale (e.g., $<km^2$, Proença *et al.*, 2009) and at the regional scale (e.g., 10^2–10^4 km^2, Martins *et al.*, 2011). In both cases the countryside SAR performed better that the classic SAR in describing the number of species in a given area, although the improvement was specially marked at the local scale.

There have been other models proposed for a multihabitat SAR (Tjørve, 2002; Triantis *et al.*, 2003). The model from Triantis *et al.* (2003) substitutes area in the SAR by the product of the area with the number of habitats. Tjørve's (2002) model adds the SAR of each habitat in the landscape deducted by the number of species overlapping between habitats to produce an estimate of the overall species diversity. None of the models explicitly tracks groups of species with distinct habitat affinities.

Other Models

Here we discuss briefly three other classes of phenomenological models: direct inferences from geographic ranges, dose–response models and statistical estimators of diversity.

Global species distributions are now available for all terrestrial vertebrate groups (Mace *et al.*, 2005; Hoffmann *et al.*, 2010). Therefore, using the simple assumption that species are restricted to native habitats, the future range of a species can be estimated by obliterating from the species' current range all the areas that are projected to be converted to agriculture, according to spatially explicit scenarios of future land-use change (Jetz *et al.*, 2007). The declines in species ranges can then be classified according to the International Union for Conservation of Nature (IUCN) criteria to produce a list of future-threatened species (Jetz *et al.*, 2007). Projections that account for the persistence of species in human-dominated habitat can be obtained by classifying the suitability of the different land-uses for each species, and only obliterating from the current range of a species the area that will be occupied by habitat with no suitability for that species (Visconti *et al.*, 2011).

Dose–response models are based on correlations between the level of an ecosystem change driver and the biodiversity response. For instance, the GLOBIO model (Alkemade *et al.*, 2009) uses a matrix with estimates of changes in mean species abundances for conversions between any two types of land use, based on 89 empirical studies comparing species abundance between at least one land-use type and primary vegetation (**Figure 3**). Then, based on land-use change scenarios, GLOBIO projects changes in mean species abundance in each grid cell of the land-use maps, which can then be averaged over countries or regions. GLOBIO also uses relationships between the mean species abundance and other drivers such as the level of infrastructure development and the level of fragmentation, and assumes a multiplicative effect of the ecosystem change drivers to estimate the integrated impact on mean species abundance.

When sample plots are taken from a region, the total species diversity of the region can be estimated by using statistical estimators that use information on the number of plots in which each species was found. For instance, if the same species appear in all plots it is likely that all the species in the region were sampled. If most species occur in only one plot it is likely that many other rare species occur in the region and were not sampled, and therefore the species diversity of the region is higher than the number of species collected in the plots. Based on these statistical methods, Hughes *et al.* (2002) estimated the total number of bird species in a countryside region using samples from cattle pastures, coffee plots, mixed agricultural plots, gardens, thin riparian strips of native vegetation, and small forest remnants. Then, in order to test the effect of the disappearance of certain elements from the landscape, they recalculated the diversity of the species in the region without the plots corresponding to those elements.

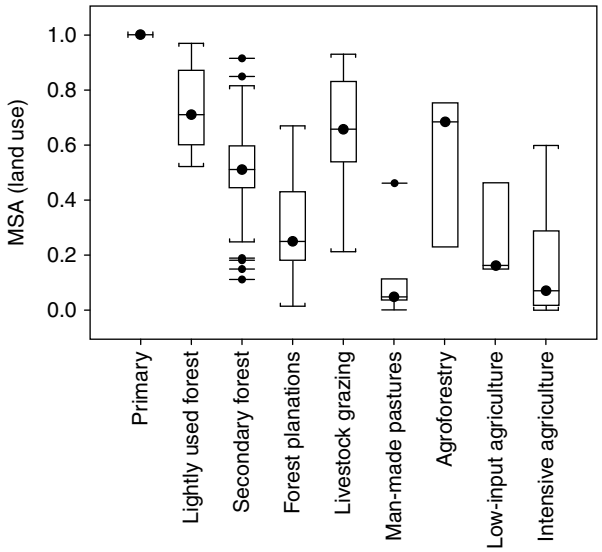


Figure 3 Box and whisker plot of the mean abundance of species (MSA) relative to their abundance in undisturbed/primary ecosystems for several land-use categories. Reproduced from Alkemade R, van Oorschot M, Miles L, Nellemann C, Bakkenes M, and ten Brink B (2009) GLOBIO3: A framework to investigate options for reducing global terrestrial biodiversity loss. *Ecosystems* 12: 374–390, with permission from Springer.

Process-Based Models

Reaction–Diffusion and Integro-Difference Equations

Reaction–diffusion models are spatially explicit models that track the population density of a species at each point in space. They were introduced in population ecology by Skellam (1951) and Kierstead and Slobodkin (1953), and are hereafter referred to as Skellam's equation (but sometimes also called the KISS model, after the name of these three authors). Earlier, reaction–diffusion had been introduced into population genetics by Fisher (1937), to look at the spatial spread of a mutant gene in a population. The main ingredients of Skellam's equation are the dispersal of individuals, which is approximated as a random diffusion process, and the local population growth dynamics (Shigesada and Kawasaki, 1997). For a population inhabiting habitat j, Skellam's equation is

$$\frac{\partial N(x,y,t)}{\partial t} = D_j \nabla^2 N(x,y,t) + F_j(N(x,y,t)) \qquad [5]$$

where the symbol ∇^2 stands for $(\partial^2/\partial x^2 + \partial^2/\partial y^2)$, $N(x,y,t)$ is the density of the population, at location (x,y) at time t, D_j is the diffusion coefficient in habitat j and F_j is the growth rate function of the population in j and it can take distinct forms. If we assume, for example, that the population has logistic growth in habitat j with growth rate r_j and carrying capacity

K_j, then Skellam's equation becomes

$$\frac{\partial N(x,y,t)}{\partial t} = D_j\left(\frac{\partial^2 N(x,y,t)}{\partial x^2} + \frac{\partial^2 N(x,y,t)}{\partial y^2}\right) + r_j N(x,y,t)\left(1 - \frac{N(x,y,t)}{K_j}\right) \quad [6]$$

This equation allows for the simulation of source-sink dynamics (Pulliam, 1988) between different habitats in the landscape (**Figure 4(b)**): the source habitats correspond to habitats where $r_j > 0$ (i.e., the birth rate is higher than the death rate) and the sink habitats correspond to habitats where $r_j < 0$ (i.e., the birth rate is lower than the death rate). The model predicts that there is a minimum area of source habitat required to sustain a species in a human-modified landscape (Pereira *et al.*, 2004).

Skellam's equation is suited to describe the spatial and temporal evolution of a single species, but it can be generalized to communities of several species (Levin, 1974; Ōkubo and Levin, 2001). For instance, Pereira *et al.* (2004) used Skellam's equation to study the response of the avifauna of Costa Rica to land-use change, and Pereira and Daily (2006) applied the same methodology to study the vulnerability of the mammal fauna of Central America in a countryside habitat. They applied Skellam's equation to each species individually using known population growth and diffusion coefficients from the literature or, when necessary, estimated them from allometric relationships. An important result of their work is that it is worth investing conservation resources in countryside habitats, in agreement with the results of the countryside SAR.

One of the advantage of Skellam's equation is its small number of parameters (at a minimum it involves three parameters: population growth rates in the source and sink habitats, and diffusion coefficient). In contrast, most models used in population viability analysis (Possingham *et al.*, 2001) tend to be very parameter rich, turning them difficult to apply to a community of species (but see Keith *et al.*, 2008). Still, Skellam's equation requires (and provides) more information than methods based on the SAR, because the latter assumes all species to be equal (or in the case of the countryside SAR, all the species in a given group), therefore not taking into account their relative vulnerability to land-use change. Skellam's equation lies between the traditional (parameter poor) SAR and (parameter rich) population viability analyses.

There are, however, situations where the approximations implicit in Skellam's equation may not be reasonable, such as, when a species only reproduces and diffuses during specific periods of the year, or when diffusion coefficients are not appropriately modeled by normal distributions, such as when there are long distance dispersal events. In such cases, integro-difference equations are an alternative (Lewis, 1997; Hastings *et al.*, 2005). For a population in habitat j with density or number of individuals N, where r_j is the growth rate function and $k_j(x–x',y–y')$ is the dispersal probability between points (x',y') and (x,y) the integro-difference equation is

$$N_{t+1}(x,y) = \int_A k_j(x-x',y-y')r_j(N_t(x',y'))dx'\,dy' \quad [7]$$

where A is the area of habitat j and we kept t as a subscript to highlight the discrete temporal nature of the equation. Naturally, like Skellam's equation, integro-difference equations can also be applied to entire species communities, but so far we are not aware of such applications to countryside habitats.

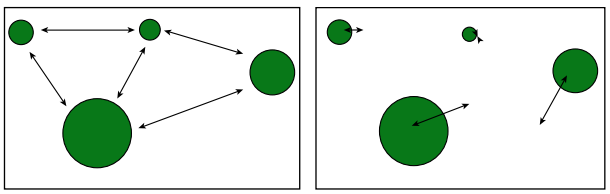


Figure 4 Schematic representation of a metapopulation (a) and of the source-sink dynamics of the Skellam model (b). (a) The population occurs in the green patches, and each patch has a colonization rate (dependent on the distance to other patches) and an extinction rate (dependent on the size of the patch). (b) The population uses the entire landscape, but has a different fitness in each habitat. In a source habitat (green patches) birth rate is greater than mortality rate, while the reverse happens in the sink habitat (white matrix). The former case leads to dispersal ability being a positive factor to the survival of the population because more patches can be colonized. In the latter case dispersal may be detrimental because more individuals reach the sink habitat.

Multispecies Metapopulation Analysis

Like Skellam's equation, metapopulation analysis was originally developed to study only one species. Levins (1969, 1970) developed the basic framework that was later considerably developed by authors such as Ilka Hanski (e.g., Hanski, 1999). The main idea of metapopulation analysis is to consider several populations of the same species inhabiting several patches, which are interconnected through dispersal, and determine how many patches are occupied at one time (**Figure 4(a)**). If the rates of colonization and extinction are c and e, respectively, and the fraction of occupied patches is given by P then the equation governing the time evolution of the fraction of occupied patches is

$$\frac{dP}{dt} = cP(1-P) - eP \quad [8]$$

If $c–e > 0$, the fraction of occupied patches reaches an equilibrium, $P^* = (1 - e/c)$. Naturally, this basic model can be refined. For instance, e and c can depend on the number of occupied patches, one can take into account the size of the patches and their explicit location, or one can consider patches of different "quality."

The metapopulation approach has definitely an appeal for studying countryside landscapes, because habitat loss and disturbance leads to populations inhabiting interconnected fragments, therefore following closely the premises of the metapopulation's theory framework. For example, Kuemmerle *et al.* (2011) used a metapopulation model to assess different strategies to establish (meta) populations of the European bison (*Bison bonasus*) in the Carpathians. However, in contrast with Skellam's model presented in the previous section, it assumes that the matrix (e.g., agricultural land) is only used for dispersal (**Figure 4(a)**). Notice that Skellam's model allows for the matrix to be used as sink habitat, and therefore tracks

the source-sink dynamics of populations across the entire landscape (**Figure 4(b)**).

Another important difference between source-sink models and metapopulation models concerns the impact of dispersal ability. While in metapopulation analyses the role of dispersal is beneficial because individuals can colonize more patches, in reaction–diffusion models dispersal affects the population negatively because individuals tend to disperse away from good patches to sink habitats where population growth rates are negative (e.g., Borda-de-Água *et al.*, 2011). This disparity arises because the source-sink dynamics of Skellam's model are deterministic and the metapopulation models are stochastic. In a stochastic environment it may be advantageous to disperse to spread the risk between different patches, while in a deterministic environment dispersal is costly and tends to be disadvantageous (but see Hamilton and May, 1977 for a competitive benefit of dispersing individuals in a deterministic environment).

How do the relative costs and benefits of dispersal play out in fragmented landscapes or countryside habitats? Unfortunately, most metapopulation models ignore the population costs of dispersal and are unable to answer this question. Casagrandi and Gatto (1999) tackled this issue by developing a stochastic model that takes into account the number of individuals in each patch and, importantly, that takes into account dispersal mortality or inability to colonize. They show that only when dispersal rates are very small does the probability of persistence of the metapopulation increases, otherwise, the probability of extinction increases when dispersal increases (**Figure 5**).

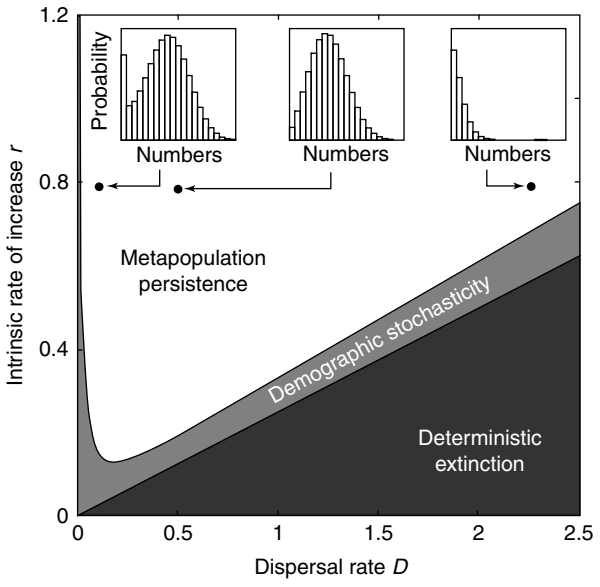


Figure 5 Population persistence (white area) as a function of the intrinsic rate of increase (*y*-axis) and dispersal rate (*x*-axis). When dispersal increases then the region of parameters where deterministic extinction occurs increases as well (dark-gray area). However for very small values of the dispersal rate ($D<0.25$) extinctions caused by demographic stochasticity decrease when the dispersal rate increases (light gray area). Reproduced from Casagrandi R and Gatto M (1999) A mesoscale approach to extinction risk in fragmented habitats. *Nature* 400: 560–562, with permission from Nature.

Most metapopulation studies focus on a single species, but metapopulation models have also been applied to communities of species (Nee and May, 1992; Tilman *et al.*, 1994; Marvier *et al.*, 2004). For example, Tilman *et al.* (1994) used a multispecies model to study the extinction debt in a fragmented landscape, and concluded that the most abundant species in an undisturbed habitat can be the most susceptible to extinction after habitat destruction. This happened because in their model there was a trade-off between being best competitor and best colonizer; hence, if fragmentation affects colonization, the species that had previously been the best competitors, and so the most abundant, became the most affected because of limitations to dispersal due to fragmentation. More recently, Keith *et al.* (2008) have integrated metapopulation models with bioclimatic envelope models to look at the response of a community of 234 plant species of the South African fynbos to climate change. They used age/stage-based matrix models to track the population size in each patch, and incorporated density dependence, environmental, and demographic stochasticity in the model, using the software RAMAS GIS (Akçakaya, 2000). A similar approach could be used to model the biodiversity dynamics of land-use change, where the bioclimatic envelope models and associated climate scenarios would be replaced by habitat suitability models applied to land-use change scenarios.

Neutral Theory

While source-sink and metapopulation models were originally developed to study one species at the time, the neutral theory of biodiversity and biogeography (Hubbell, 2001) has at its core the relative abundance of species within a community. By "neutral" it is meant that all species in the same trophic level have the same probability of dying, reproducing, and speciating. Such an approximation has proved very controversial among ecologists (McGill, 2003; Clark, 2009), but the theory has inspired many recent studies in theoretical and applied ecology (for a review see Rosindell *et al.*, 2011). We should also point out that the classic Theory of Island Biogeography (Mac Arthur and Wilson, 1967) is also neutral because it assumes that all species are equal in their ability of dispersal, colonizing, and becoming extinct. An important difference, however, is that the Theory of Island Biogeography assumes neutrality at the species level, while neutral theory does it at the individual level and, as a consequence, the Theory of Island Biogeography can deal with species richness but not relative abundance.

In the original version laid out by Hubbell (2001), neutral theory considered two different scales: the local community corresponding to the local scale (e.g., the trees of a 50 ha plot of tropical rain forest), and the metacommunity corresponding to the regional scale (such as the Amazon basin). At both scales death and birth are part of the model, but while in the metacommunity speciation is the driving force providing new species (and, in the process, avoiding community dominance by a single species), in the local community that role is provided by migration of individuals from the metacommunity. According to the neutral theory model, at each time step an individual dies and is replaced with a new one. When

dealing with the local community, the new individual can be from a local species or an immigrant from one of the species in the metacommunity (that may or may not be present in the local community). The important characteristic of the theory is that all the processes are function of the species relative abundances; for example, if there is a migration event, the probability of choosing a species from the metacommunity is proportional to its abundance there. The evolution of the meta and local communities leads to the number of species and their relative abundance to reach an equilibrium value, characterized by the speciation and migration rates. This, however, is a dynamic equilibrium because species abundances drift over time; therefore, the species composition may change. It is only the number of species and their relative abundance that remain approximately constant.

Halley and Iwasa (2011) used a simplified version of neutral theory, by ignoring the immigration term, to estimate the time for accruing extinction debt in an isolated fragment. Under their model they derived the following diffusion equation:

$$\frac{\partial p}{\partial t} = \sum_{i=1}^{S-1} \frac{\partial^2}{\partial x_i^2}\left(\frac{x_i(1-x_i)}{N\tau}p\right) - \sum_{i=1}^{S-1}\sum_{j>i}\frac{\partial^2}{\partial x_i \partial x_j}\left(\frac{2x_i x_j}{N\tau}p\right) \quad [9]$$

where S is the number of species in the community, N the number of individuals in the habitat fragment, τ is the mean generation time, x_i is the relative abundance of species i, and $p(x_1, x_2, x_3, \ldots, x_{S-1})$ is the distribution of the relative abundances for up to the last species, whose relative abundance is $x_S = 1 - x_1 - x_2 - x_3 \ldots - x_{S\text{-}1}$. From this equation, they calculated the analytical expression for the time it takes for the number of species to reach half of its original size, S_0, as $t_{50} = \tau N/S_0$. The decay over time of the number of species is given by

$$S(t) = \frac{S_0}{1 + t/t_{50}} \quad [10]$$

Halley and Iwasa found very good agreement between their model and published empirical results for avifauna relaxation in fragments of intermediate size (100–100,000 ha), but noticed that the model broke down either for very small or very large areas, and hypothesized that this breakdown was due to the lack of immigration and speciation in their model.

Other studies have applied neutral models to estimate the number of tree species extinctions in the Amazon Basin for spatially explicit deforestation scenarios (Hubbell *et al.*, 2008), and to predict the rate of extinction of tree species in isolated fragments of the Amazon forest (Gilbert *et al.*, 2006).

Concluding Remarks

There is now a wide range of models available to look at biodiversity dynamics on countryside landscapes. Process-based models require more parameters than phenomenological models, and therefore they have been mostly applied to local scales and to a single species or a small set of species. In contrast, phenomenological models may require as little as one parameter, and they have been applied to projecting changes at global to regional scales for the total number of species. However, phenomenological models have also been used at the landscape scale, and there have been attempts to generalize and upscale the projections of process-based models. The type of projections needed (e.g., individual species vs. species groups, time to extinction vs. number of extinctions) and the data available are the key factors in determining the most appropriate model.

Appendix

List of Courses

1. Applied Ecology
2. Biodiversity and Ecosystem Services
3. Biodiversity and Global Changes
4. Environment and Land Planning
5. Population Resources and the Environment

See also: Biodiversity, Definition of. Conservation Biology, Discipline of. Dispersal Biogeography. Diversity, Community/Regional Level. Extinction in the Fossil Record. Island Biogeography. Land-Use Patterns, Historic. Loss of Biodiversity, Overview. Mammals, Conservation Efforts for. Metapopulations. Plant Conservation. Population Dynamics. Rainforest Loss and Change. Species–Area Relationships

References

Akçakaya HR (2000) Viability analyses with habitat-based metapopulation models. *Researches on Population Ecology* 42: 0045.

Alkemade R, van Oorschot M, Miles L, Nellemann C, Bakkenes M, and ten Brink B (2009) GLOBIO3: A framework to investigate options for reducing global terrestrial biodiversity loss. *Ecosystems* 12: 374–390.

Arrhenius O (1921) Species and area. *Journal of Ecology* 9: 95–99.

Barnosky AD, Matzke N, Tomiya S, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57.

Borda-de-Água L, Navarro L, Gavinhos C, and Pereira HM (2011) Spatio-temporal impacts of roads on the persistence of populations: analytic and numerical approaches. *Landscape Ecology* 26: 253–265.

Brook BW, Sodhi NS, and Ng PKL (2003) Catastrophic extinctions follow deforestation in Singapore. *Nature* 424: 420–423.

Brooks TM, Pimm SL, and Oyugi JO (1999) Time lag between deforestation and bird extinction in tropical forest fragments. *Conservation Biology* 13: 1140–1150.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: indicators of recent declines. *Science* 328: 1164–1168.

Casagrandi R and Gatto M (1999) A mesoscale approach to extinction risk in fragmented habitats. *Nature* 400: 560–562.

Clark JS (2009) Beyond neutral science. *Trends in Ecology and Evolution* 24: 8–15.

Crawley MJ and Harral JE (2001) Scale dependence in plant biodiversity. *Science* 291: 864–868.

Daily GC, Ceballos G, Pacheco J, Suzan G, and Sanchez-Azofeifa A (2003) Countryside biogeography of neotropical mammals: Conservation opportunities in agricultural landscapes of Costa Rica. *Conservation Biology* 17: 1814–1826.

Daily GC, Ehrlich PR, and Sanchez-Azofeifa GA (2001) Countryside biogeography: Use of human-dominated habitats by the avifauna of southern Costa Rica. *Ecological Applications* 11: 1–13.

Dengler J (2009) Which function describes the species–area relationship best? A review and empirical evaluation. *Journal of Biogeography* 36: 728–744.

Doak DF (1995) Source-sink models and the problem of habitat degradation: General models and applications to the Yellowstone grizzly. *Conservation Biology* 9: 1370–1379.

Drakare S, Lennon JJ, and Hillebrand H (2006) The imprint of the geographical, evolutionary and ecological context on species–area relationships. *Ecology Letters* 9: 215–227.

Duraiappah A, Naheem S, Agardy T, *et al.* (2005) *Ecosystems and Human Well-Being: Biodiversity Synthesis*. Washington, DC: World Resources Institute.

FAO (2010) *Global Forest Resources Assessment* 2010. Rome: Food and Agriculture Organization of the United Nations.

Fisher RA (1937) The wave of advance of advantageous genes. *Annals of Human Genetics* 7: 355–369.

Gilbert B, Laurance WF, Leigh EG, and Nascimento HEM (2006) Can neutral theory predict the responses of amazonian tree communities to forest fragmentation? *American Naturalist* 168: 304–317.

Halley JM and Iwasa Y (2011) Neutral theory as a predictor of avifaunal extinctions after habitat loss. *Proceedings of the National Academy of Sciences of the United States of America* 108: 2316–2321.

Hamilton WD and May RM (1977) Dispersal in stable habitats. *Nature* 269: 578–581.

Hanski I (1999) *Metapopulation Ecology*. Oxford, UK: Oxford University Press.

Hastings A, Cuddington K, Davies KF, *et al.* (2005) The spatial spread of invasions: New developments in theory and evidence. *Ecology Letters* 8: 91–101.

He F and Hubbell SP (2011) Species–area relationships always overestimate extinction rates from habitat loss. *Nature* 473: 368–371.

Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the World's vertebrates. *Science* 330: 1503–1509.

Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NJ: Princeton University Press.

Hubbell SP, He FL, Condit R, Borda-de-Agua L, Kellner J, and ter Steege H (2008) How many tree species and how many of them are there in the Amazon will go extinct? *Proceedings of the National Academy of Sciences of the United States of America* 105: 11498–11504.

Hughes JB, Daily GC, and Ehrlich PR (2002) Conservation of tropical forest birds in countryside habitats. *Ecology Letters* 5: 121–129.

Jetz W, Wilcove DS, and Dobson AP (2007) Projected impacts of climate and land-use change on the global diversity of birds. *PLoS Biology* 5: e157.

Keith DA, Akcakaya HR, Thuiller W, *et al.* (2008) Predicting extinction risks under climate change: Coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters* 4: 560–563.

Kierstead H and Slobodkin LB (1953) The size of water masses containing plankton boom. *Journal of Marine Research* 12: 141–147.

Kinzig A and Harte J (2000) Implications of endemics–area relationships for estimates of species extinctions. *Ecology* 81: 3305–3311.

Kuemmerle T, Perzanowski K, Akcakaya HR, *et al.* (2011) Cost-effectiveness of strategies to establish a European bison metapopulation in the Carpathians. *Journal of Applied Ecology* 48: 317–329.

Levin SA (1974) Dispersion and population interactions. *American Naturalist* 207–228.

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of The Entomological Society of America* 15: 237–240.

Levins R (1970) Extinction. *Lectures on Mathematical Analysis of Biological Phenomena* 231: 123–138.

Lewis MA (1997) Variability, patchiness, and jump dispersal in the spread of an invading population. In: Tilman D and Kareiva PM (eds.) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, pp. xiv, 308. Princeton, NJ: Princeton University Press.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Mace GM, Masundire H, Baillie J, *et al.* (2005) Biodiversity. In: Hassan R, Scholes B, and Ash N (eds.) *Ecosystems and Human Well-being: Current State and Trends: Findings of the Condition and Trends Working Group of the Millenium Ecosystem Assessment*, pp. 77–126. Washington, DC: Island Press.

Martins IS (2009) *What Is the Relationship between Habitat Diversity and Species Diversity from the Local to the Regional Scale?* Masters Thesis, Universidade de Lisboa, Portugal.

Marvier M, Kareiva P, and Neubert MG (2004) Habitat destruction, fragmentation, and disturbance promote invasion by habitat generalists in a multispecies metapopulation. *Risk Analysis* 4: 869–878.

McGill B (2003) A test of the unified neutral theory of biodiversity. *Nature* 422: 881–885.

Nee S and May RM (1992) Dynamics of metapopulations – habitat destruction and competitive coexistence. *Journal of Animal Ecology* 61: 37–40.

Ōkubo A and Levin SA (2001) *Diffusion and Ecological Problems: Modern Perspectives*. New York, NY: Springer.

Pereira HM, Borda-de-Agua L, and Martins IS (2012) Geometry and scales in species–area relationships. *Nature* 482: E3–E4.

Pereira HM and Daily GC (2006) Modeling biodiversity dynamics in countryside landscapes. *Ecology* 87: 1877–1885.

Pereira HM, Daily GC, and Roughgarden J (2004) A framework for assessing the relative vulnerability of species to land-use change. *Ecological Applications* 14: 730–742.

Pereira HM, Leadley PW, Proenca V, *et al.* (2010) Scenarios for global biodiversity in the 21st century. *Science* 330: 1496–1502.

Pimm SL, Russell GJ, Gittleman JL, and Brooks TM (1995) The future of biodiversity. *Science* 269: 347–350.

Possingham H, Lindenmayer DB, and McCarthy MA (2001) In: Levin S (ed.) *Population Viability Analysis. Encyclopedia of Biodiversity*, pp. 831–843. San Diego, California, USA: Academic Press.

Proença VM (2009) *Galicio-Portuguese Oak Forest of Quercus Robur and Quercus Pyrenaica: Biodiversity Patterns and Forest Response to Fire*. PhD Thesis, Universidade de Lisboa, Portugal.

Pulliam HR (1988) Sources, sinks, and population regulation. *American Naturalist* 132: 652–661.

Ranganathan J, Daniels R, Chandran M, Ehrlich P, and Daily G (2008) Sustaining biodiversity in ancient tropical countryside. *Proceedings of the National Academy of Sciences of the United States Of* 105: 17852–17854.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.

Rosenzweig ML (2001) Loss of speciation rate will impoverish future diversity. *Proceedings of the National Academy of Sciences of the United States of America* 98: 5404–5410.

Rosindell J, Hubbell SP, and Etienne RS (2011) The unified neutral theory of biodiversity and biogeography at age ten. *Trends in Ecology & Evolution* 26: 340–348.

Sala OE, van Vuuren D, Pereira HM, *et al.* (2005) Biodiversity Across Scenarios. In: Carpenter SR, Prabhu LP, Bennet EM, and Zurek MB (eds.) *Ecosystems and Human Well-Being: Scenarios: Findings of the Scenarios Working Group of the Millennium Ecosystem Assessment*, pp. 375–408. Washington, DC: Island Press.

Shigesada N and Kawasaki K (1997) *Biological Invasions: Theory and Practice*. Oxford: Oxford University Press.

Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* 38: 196–218.

Thuiller W, Albert C, Araujo M, *et al.* (2008) Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology Evolution and Systematics* 9: 137–152.

Tilman D, May R, Lehman CL, and Nowak MA (1994) Habitat destruction and the extinction debt. *Nature* 371: 65–66.

Tjørve E (2002) Habitat size and number in multi-habitat landscapes: A model approach based on species–area curves. *Ecography* 25: 17–24.

Triantis KA, Mylonas M, Lika K, and Vardinoyannis K (2003) A model for the species–area-habitat relationship. *Journal of Biogeography* 30: 19–27.

UNEP (2007) *Global Environment Outlook 4*. Nairobi, Kenya: UNEP.

Visconti P, Pressey RL, Giorgini D, *et al.* (2011) Future hotspots of terrestrial mammal loss. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366: 2693–2702.

van Vuuren D, Sala O, and Pereira HM (2006) The future of vascular plant diversity under four global scenarios. *Ecology and Society* 11: 25.

Wright S (2010) The future of tropical forests. *Year in Ecology and Conservation Biology 2010* 1195: 1–27.

Modeling Marine Ecosystem Services

Anne D Guerry and Mary H Ruckelshaus, Stanford University, Stanford, CA, USA
Mark L Plummer and Dan Holland, National Oceanic and Atmospheric Administration, Seattle, WA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ecosystem services Wide array of benefits that ecosystems and their biodiversity confer on humanity.
Marine Broadly defined to include coastal (on land, within a narrow fringe adjacent to saltwater), intertidal, nearshore, and open ocean.
Production function approach An approach that models ecosystem services as the relationship between ecological and human inputs (e.g., the structure and functions of an ecological system, human labor and capital) and outputs valued by humans.
Valuation Act of estimating or setting the value of something,
Value Relative worth, merit, or importance. Can be measured in various ways, including but not limited to monetary metrics.

Introduction to Marine Ecosystem Services

Humans have always benefited from marine ecosystems, enjoying resources such as seafood and opportunities for activities like recreation and marine transportation. These ecosystems also provide indirect benefits by sequestering carbon and playing key roles in the regulatory processes of other global cycles. Benefits derived from these systems are broadly characterized as the ecosystem services provided by marine ecosystems. As identified by the Millennium Ecosystem Assessment (MA) (2005b), marine ecosystem services span four major categories: provisioning, regulating, cultural, and supporting services (**Table 1**). Broad assessments of marine and coastal ecosystems services based on the MA can be found in Agardy *et al.* (2005) and the United Nations Environment Program (2006), as well as other synthesis documents (e.g., Peterson and Lubchenco, 1997). Costanza (2000), Patterson and Glavovic (2008), and Wilson and Liu (2008) also provide useful overviews of these services, as do descriptions of the particular services provided by fish populations (e.g., Holmlund and Hammer, 1999), coral reef ecosystems (e.g., Moberg and Folke, 1999) and mangroves (e.g., Ronnback, 1999).

Among the four types of marine ecosystem services, provisioning services such as food from capture fisheries, aquaculture, and wild foods are the most obvious and easily valued. About 80 million tons of fish were landed in marine capture fisheries worldwide in 2009, and fish account for approximately 16% of the annual animal protein consumption by humans (FAO Fisheries Department, 2010). Globally, more than 1.5 billion people rely on fish for almost 20% of their animal protein. On average, each person living in 2009 ate 17.2 kg of fish (including fish from aquaculture; the proportion from capture fisheries alone is difficult to calculate given nonfood uses of wild fish) (FAO Fisheries Department, 2010). Other provisioning services include timber and fiber from mangroves and seagrass beds, and biochemicals for cosmetics and food additives. The potential also exists for developing novel natural products from marine species with medical applications (Carte, 1996). For example, researchers recently found that three marine species collected from offshore oil and gas platforms in California's Santa Barbara Channel had potential for biomedical applications, including two that inhibit the division of cancer cells grown in the laboratory (Schmitt *et al.*, 2006). In addition, the ocean may become an important energy source: biofuels from algae and power generation from wave and tidal energy have potential for more widespread use. And finally, the world's oceans provide the highways for the global shipping trade.

Marine systems also provide a wide range of regulating services. As vividly highlighted by the human losses wrought by the 2005 hurricanes on the US Gulf Coast, coastal and estuarine wetlands have value for their ability to reduce storm surge elevations and wave heights (Danielsen *et al.*, 2005; Travis, 2005). Other regulating services provided by marine systems include the transformation, detoxification, and sequestration of wastes (Peterson and Lubchenco, 1997). And "blue carbon" – the role oceans can play as carbon sinks – is a service that is gaining interest (Nellemann *et al.*, 2009).

Human societies have always been drawn to the oceans, which have proven to be a rich source of cultural services. In 1995, an estimated 39% of the world's population lived with 100 km of a coast (Burke *et al.*, 2001). In the US, people love to live near the ocean; one study predicts average increases of 3600 people a day moving to coastal counties through 2015 (Culliton, 1998). Coastal tourism is a key component of many economies around the world and is one of the fastest growing and most profitable sectors of tourism (United Nations Environment Program, 2006). Countless intangible benefits are drawn from the oceans, too. For example, many communities define their identities in relation to the coast, people draw inspiration from oceans in numerous ways, and rich ceremonial traditions are intricately linked to the sea.

Finally, the oceans provide essential supporting services that underpin many of the globe's ecological functions. The oceans hold 96.5% of Earth's water (Gleick, 1996) and are a primary driver of the atmosphere's temperature, moisture content, and stability (Colling, 2001). The global cycles of carbon, nitrogen, oxygen, phosphorus, sulfur, and other key elements flow through the oceans (Peterson and Lubchenco, 1997), and marine ecosystems are responsible for approximately 40% of global net primary productivity (Schlesinger, 1991; Melillo *et al.*, 1993). The oceans are home to vast reservoirs of genetic and ecological diversity, arguably the most fundamental source of supporting services as they are directly

Table 1 Ecosystem services provided by marine systems

Types of service	*Examples*
Provisioning services	
Food production (capture fisheries; aquaculture; and wild foods)	Tuna, crab, lobster; salmon, oysters, shrimp, seaweed; mussels, clams
Fiber production	Mangrove wood, seagrass fiber
Biomass fuel production	Mangrove wood, biofuel from algae
Maintenance of aquatic systems	Shipping, tidal turbines
Generation of genetic resources	Individual salmon stocks, marine diversity for bioprospecting
Production of biochemicals, natural medicines, and pharmaceuticals	Antiviral and anticancer drugs from sponges, carrageenans from seaweed
Regulating services	
Climate regulation	Major role in global CO_2 cycle
Water regulation	Natural stormwater management by coastal wetlands and floodplains
Erosion regulation	Nearshore vegetation stabilizes shorelines
Water purification and waste treatment	Uptake of nutrients from sewage wastewater, detoxification of PAHs by marine microbes, sequestration of heavy metals
Disease regulation	Natural processes may keep harmful algal blooms and waterborne pathogens in check
Pest regulation	Grazing fish help keep algae from overgrowing coral reefs
Pollination/Assistance of external fertilization	Innumerable marine species require seawater to deliver sperm to egg
Natural hazard regulation	Coastal and estuarine wetlands and coral reefs protect coastlines from storms
Cultural services	
Provision of conditions that support or enhance ethical values (nonuse)	Spiritual fulfillment derived from estuaries, coastlines, and marine waters
Provision of conditions that support or enhance existence values (nonuse)	Belief that all species are worth protecting, no matter their direct value to humans
Provision of recreation and ecotourism opportunities (consumptive and nonconsumptive uses)	Scuba diving, beachcombing, whale watching, boating, snorkeling; fishing, clamming
Supporting services	
Nutrient cycling	Major role in carbon, nitrogen, oxygen, phosphorus, and sulfur cycles
Soil formation	Many salt-marsh surfaces vertically accrete; eelgrass slows water and traps sediment
Primary production	Significant portion of global net primary productivity
Water cycling	Most of Earth's water is in oceans; they are central to the global water cycle

Source: Adapted from Guerry A, Plummer M, Ruckelshaus M, and Harvey C (2011) Ecosystem service assessments for marine conservation. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital: Theory and Practice of Mapping Ecosystem Services*. Oxford: Oxford University Press.

linked to the rate of evolution and therefore the ability to adapt to a changing climate (Pergams and Kareiva, 2009).

Since the MA, much work has been done in developing and applying new methodologies to model, map, and value ecosystem services (e.g., UNEP and IOC-UNESCO, 2009; Kareiva *et al.*, 2011; UK National Ecosystem Assessment, 2011). A great deal of the original MA was focused on terrestrial systems, but new work—and new political contexts—have sparked the expansion of an ecosystem services framework to marine systems. For example, the UK recently completed a groundbreaking, countrywide assessment of both terrestrial and marine ecosystem services and their value (Stokstad, 2011; UK National Ecosystem Assessment, 2011). Although agricultural production has increased between 1950 and the present, landings of fish and shellfish from UK waters have declined since the 1960s from almost 900,000 t to slightly more than 500,000 t in 2008 (UK National Ecosystem Assessment, 2011). Protection of UK coastlines from storm-induced flooding and erosion has declined in response to a 10% loss of natural habitats such as dunes, kelps, seagrasses and marshes over the past 60 years (UK National Ecosystem Assessment, 2011).

The MA and subsequent assessments have raised awareness of ecosystem services, the explicit human dependence on them, and the threatened status of many of them. Bringing ecosystem services into the active management of terrestrial and marine ecosystems, however, requires more than just a catalog of services and their total values. More pragmatic is the assessment of the ecological and economic consequences of management activities in particular places. An understanding of how changes in ecosystems are likely to lead to changes in ecosystem services can most clearly provide information to decision makers. Modeling marine ecosystem services can play an important role in providing such insights.

Foundations

Marine ecosystems provide both great opportunities and challenges for the application of the framework of ecosystem

services. The concept of ecosystem services has a long history but has seen a relatively recent resurgence of interest from ecologists, economists, and conservation practitioners. Because ecosystem services are reviewed elsewhere in this volume, we focus here on the conceptual frameworks and methodologies that pertain to marine applications, as well as the challenges facing these applications.

Why Model Marine Ecosystem Services?

Models are important tools for examining the dynamic natures of and interconnections among the biophysical and human elements of marine ecosystems. They provide a way of exploring future scenarios that lie outside the range of past experiences, as well as possible unexpected consequences of policy actions. An ecosystem services framework provides important insights into the challenge of pursuing ecosystem-based management. Ecosystem services are the currencies through which the consequences of ecosystem change flow to people. Using such a framework, ecosystem services and their values can be used as a set of metrics for assessing alternative management interventions and their potential impacts on economic or social well-being.

Models of ecosystems that incorporate both biophysical and human components can present policy makers with an enormous set of potential indicators for gauging the changes brought about by policies. Ecosystem services – the "ecological endpoints" of the system that are directly connected to human well-being (Boyd, 2007) – can provide a guide for winnowing this set down. In the context of ecosystem modeling, such a framework avoids the potential problem of double counting the value of "intermediate" ecological elements (e.g., forage fish) that support other "endpoint" elements with direct value (e.g., fishery harvests) because the value of the intermediate elements are embedded in the value of the endpoints (Boyd and Banzhaf, 2007).

Ecosystem models that focus on ecosystem services also provide the opportunity to overcome problems arising from traditional management in the marine realm, which generally proceeds sector by sector, with distinct bodies making isolated decisions. These single-sector decisions often affect a broad set of ecosystem services, many of which are outside the scope of their authority. As a result, there has been little effort to consider the ways in which single decisions impact the full suite of things people care about and need.

Using an ecosystems services framework can highlight trade-offs among multiple objectives so that decisions to resolve those trade-offs can be made transparently. Increasingly, scientists and managers are pointing to links between diversity, productivity, and resilience attributes of marine systems and their response to human interventions in conserving, harvesting, and regulating marine ecosystem services (Liu *et al.*, 2007a; Levin and Lubchenco, 2008; Murawski *et al.*, 2009; Chan and Ruckelshaus, 2010). Our ability to model and assess trade-offs among objectives of multiple management sectors (e.g., fisheries, wave or wave energy, recreation) and ecosystem services (e.g., value of fishery landings, kilowatt hours of energy generated, revenue from recreational activities) is needed in order to inform more complex cases of ecosystem-based management that can accommodate a broader suite of actors (Foley *et al.*, 2010).

Pioneering efforts to use the framework of ecosystem services to understand the connections between activities in one sector and their impacts on others are underway in the marine environment. Decision makers at various levels of government from around the globe, including the Interagency Ocean Policy Task Force (IOPTF) (2010), the European Commission (2010), and the recently approved UN Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) have recognized that new approaches are needed to ensure the sustainability of benefits people derive from oceans and coasts (UN General Assembly (UNGA), 2010).

Among these parties, there is currently a great deal of emphasis on marine (or "maritime") spatial planning. This approach to comprehensive management of marine and coastal systems analyzes current and anticipated uses, identifies areas most suitable for particular activities, and provides a process to better determine how oceans are used and protected now and for future generations (Douvere and Ehler, 2009; Ehler and Douvere, 2009). Ecosystem services provide a framework and a common language for this type of planning process, allowing agencies and private interests to articulate their goals transparently using common terms and to better understand how their decisions interact with those of other users and activities.

Connecting Social and Ecological Systems

Several related conceptual frameworks provide the context within which ecosystem service information is critical to understanding feedbacks between humans and ecosystem conditions in marine and other environments. One of the most important is the notion that humans are an integral part of ecosystems, and so ecosystem models should encompass their behavior (Holland *et al.*, 2010). Similarly, the theory of complex adaptive systems encompasses humans as parts of coupled systems and presents humans as participating in the dynamics of the system (Levin, 1998; Levin and Lubchenco, 2008). The production of ecosystem services involves a combination of ecological functions and human actions and values (**Figure 1**). In the marine environment, for example, marine ecosystems have fish populations that offer opportunities for commercial and recreational harvests, two of the most valuable services derived from these systems. The human action of harvest, however, is what transforms the potential ecological supply into the actual provision of the ecosystem service (Tallis *et al.*, 2011). Human values in the form of the demand for seafood, the valuation of recreation, and the costs of commercial harvest or recreational angling all determine the value of these services.

Building models to account for the humans who interact with or are affected by an ecological system can therefore provide fundamental insights into the provision and value of ecosystem services. Simple economic models can focus on individual decisions to use more or less of a marine resource, for example. An ecological change that makes it harder to harvest fish is likely to produce a lower level of effort that can be modeled in a simple framework of the demand for that



Figure 1 The production of ecosystem services involves a combination of ecological functions and human actions and values. Recreational fishing as an ecosystem service, for example, depends on the healthy functioning of natural systems that support fish populations. The human action of harvest transforms the potential ecological supply into the actual provision of the ecosystem service.

activity. More complicated models involve complex decisions about where and when to fish, which species to fish for, and so forth (e.g., Sanchirico and Wilen,1999; Wilen *et al.*, 2002). Similarly, human behavior has more latitude to make adjustments over longer periods of time, and so modeling short-run versus long-run behavior of a system can account for such differences (Holland and Brazee, 1996). As they mature, models of marine ecosystem services can include how humans interact with biophysical components of the system and can incorporate realistic mechanisms for changing those interactions based on economic and other incentives. We include some examples of these kinds of truly linked human and biophysical models in the "Production Function Modeling" section.

The coupled social–ecological systems model developed by Ostrom (2009) is one key conceptual framework that includes interactions between biophysical and social systems. Much of the focus of research guided by this conceptual model is to identify relevant attributes for understanding the dynamics of a coupled social–ecological system (e.g., the lobster fishery in Maine and its fishing community) and the effects of different management approaches (e.g., regulation, incentives) on ecological and ecosystem services (e.g., lobster biomass, landings, biodiversity) and social (e.g., efficiency, equity, accountability) outcomes. Under this framework, the features of governance systems (e.g., property rights systems, operational rules), user groups (e.g., social heterogeneity, leadership), and resource units (e.g., economic value, behavior of harvested species) that give rise to different ecological and social outcomes can be quantitatively or qualitatively assessed using well-developed methodologies (Cudney-Bueno and Basurto, 2009; Ostrom, 2009; Basurto and Coleman, 2010).

The closely related conceptual framework of coupled human and natural systems touches on similar issues but with a stronger focus on understanding mechanisms and feedbacks (Liu *et al.*, 2007a, 2007b). A critical question this framework poses is the positive or negative direction of feedback loops between human and natural components of a system and the strength of those interactions (Liu *et al.*, 2007b). This modeling framework encourages exploration of temporal and spatial as well as organizational couplings, thresholds, and resilience in linked systems. For example, whether couplings between human and natural components of a marine ecosystem are direct and local or indirect and global can be a key feature influencing how humans and ecosystem services respond to signs of change in the ecosystem, and it is an increasingly important feature with the globalization of both human and natural processes.

One modeling approach that incorporates these frameworks is the integrated ecosystem assessment (IEA). An IEA outlines both a decision analytical framework and a process containing the logical steps necessary for managing ecosystems, including explicit identification of objectives and the indicators for tracking progress, a risk analysis describing the state of the indicators, an evaluation of management strategies, and an overall assessment of ecosystem response (Levin *et al.*, 2009). IEAs have a successful history of application in fishery-based ecosystem management (Caddy, 1999; Sainsbury *et al.*, 2000; Smith *et al.*, 2007); they are designed to be iterative, expanding in scope and sophistication as

information develops over time (e.g., Dennison *et al.*, 2007; Tallis *et al.*, 2010; Samhouri *et al.*, 2011).

IEAs are one example of current approaches to the management of marine ecosystems that can use ecosystem services in a modeling framework. An IEA uses quantitative analyses and ecosystem modeling to support management decisions that address a range of social, economic, and natural conditions. Ecosystem services can provide a set of metrics for assessing these conditions, and modeled changes in their levels can then provide decision makers with a way to compare alternative policies.

Challenges in Modeling Marine Ecosystem Services

From a scientific perspective, the modeling of marine ecosystem services is sufficiently different from the task on land to warrant its own treatment. Marine systems provide some special scientific challenges in the endeavor to map, model, and value ecosystem services. Most models for terrestrial systems use a land-use or land-cover data layer as a foundation for the initial assessment of services, as well as a baseline for exploring how land conversion and land-use change alter ecosystem services. Fundamentally, the same approach works for marine environments. Marine systems have habitats that are altered or used with consequences, and different management approaches yield different habitat conditions and different delivery of services.

We cannot readily "see" most benthic habitat types using satellite imagery or other remote sensing technology, so maps of habitat type and condition are harder to come by in marine systems than they are on land. Also, in marine systems, associations between species and habitat patches are more difficult to discern. The lack of basic land-use and land-cover data for marine systems underscores the need for a modeling approach to ecosystem and food-web dynamics that is less tightly coupled to detailed habitat maps than the approach often taken in terrestrial systems (e.g., Kareiva *et al.*, 2011).

A second shift in perspective required when considering marine ecosystem services stems from the ways in which humans interact with marine environments. We do not live in the sea, and thus many of our actions that affect the marine environment are indirect. We can think of many terrestrial habitats without considering the effects of marine systems on them, but the converse is not true. Activities on land (e.g., coastal development; nutrient, sediment, and pathogen inputs to freshwater; increases in impervious surfaces) have profound effects on nearshore marine systems (Carpenter *et al.*, 1998; Mallin *et al.*, 2000; Diaz and Rosenberg, 2008).

Understanding and modeling marine ecosystem services is also made more difficult by the fact that we generally have less direct control or knowledge of human actions. In terrestrial areas, property rights are the norm, and zoning often dictates what and where actions can take place (both on public and private property). Human activities and their results are also more observable on land. In marine areas, human activities are not typically managed explicitly at a fine scale, and many areas are effectively open access. Thus, it is often more critical to incorporate endogenous behavioral models of people into models of ecosystem services.

Finally, modeling ecosystem services is not a substitute for all other ways of assessing marine and other ecosystems. Concerns about biodiversity of marine ecosystems, for example, may be poorly addressed by a pure ecosystem services approach. Some assert that protection and restoration of marine systems will improve both biodiversity and the ability of the oceans to provide ecosystem services; others point out that trade-offs can occur (Balvanera *et al.*, 2006; Worm *et al.*, 2006; Naeem *et al.*, 2009; Palumbi *et al.*, 2009; Klein *et al.*, 2010; Brander, 2010). Ecosystem modeling coupled with observations should not ignore this and other approaches to marine conservation, and they should continue to provide needed evidence for the conditions under which biodiversity and ecosystem service provision are positively related and where trade-offs are likely to occur.

Approaches

The modeling of marine ecosystem services can take a variety of approaches. Although not normally considered a "model," a straightforward accounting of ecosystem services and their values has played an important role in past assessments, and so it is included in this section as one approach. Many models of ecosystem services take a production function approach in which the structure and functions of an ecological system are combined with human actions and capital to produce an output valued by humans (National Research Council, 2005). This approach has been used to analyze a variety of individual ecosystem services, as well to provide the framework for developing suites of models that encompass multiple services. Each of these approaches is presented in the following sections.

Accounting Approaches and Basic Valuation Methods

The simplest approach to assessing ecosystem services for marine and terrestrial systems is a straightforward accounting of those services and their values. This approach is not one that explicitly models the ecological or human systems, but it has been used to assess marine ecosystem services (Beaumont *et al.*, 2007, 2008). Generally, the approach begins by defining a particular geographic location and then identifying the set of ecosystem services flowing from that area. The assessment can then be as simple as compiling local information on each of these services for that region – e.g., using revenues from an area's fishery harvests as a measure of the economic value of that type of provisioning service.

A more common method that has been used for terrestrial systems is ecosystem-service mapping (Troy and Wilson, 2006). This method differentiates a particular area by land cover, biome, or some other set of ecologically based landscape or seascape type. Drawing from a standard set of ecosystem service categories (e.g., deGroot *et al.*, 2002), each landscape type (e.g., forestland) is then linked to a set of services (e.g., recreation or carbon sequestration) believed to be provided by that type. The quantity of an ecosystem service produced is then assumed to be linearly related to the area of the landscape type.

From these starting points, local economic data, if they exist, can be used to transform the quantity of ecosystem services into economic values. Such data are not commonly available, however, and so the more typical approach to assigning economic value is to use what is known as *benefit transfer*. Benefit transfer is a method for taking economic data on benefits (or values in general) gathered from one site and applying the data to another site. This method is rarely the "first-best" choice for estimating economic values but the costs of gathering primary, site-specific data have made it a common practice for studies of ecosystem service value (e.g., see Rosenberger and Loomis (2001) and National Research Council (2005) for examples of recreational uses of natural sites). Using the assumption of a linear relation between ecosystem service quantities and landscape area, a "unit value" for a particular ecosystem service and landscape type can then be derived by taking an existing study's estimated value for a particular ecosystem service and dividing by the area of the studied site's landscape (e.g., recreation or carbon sequestration value per acre of forestland). These unit values can then be used to assess the value of ecosystem services at other sites for multiple services and landscape types if they are present.

This accounting approach was pioneered by Costanza and colleagues (1997), who estimated global values for 17 ecosystem services for 16 biomes. Their effort included five marine and coastal biomes: open ocean, estuaries, seagrass or algae beds, coral reefs, and ocean shelf. These five biomes accounted for almost two-thirds of the estimated annual $33 trillion value of Earth's services. Their use of benefit transfer was almost universal; the original data used in the study were gathered in one or more locations and then projected worldwide.

Accounting approaches such as Costanza *et al*. (1997) are simple and play the important role of raising the awareness of the values of traditionally undervalued, nonmarket ecosystem services. They should be used cautiously, however, with important qualifications in mind. First, this approach imposes linearity on the valuation of ecosystem services, which has the potential to produce biased estimates. Human pressures on ecosystems and the services they provide can result in impacts that respond nonlinearly to changes in the scale of the pressure or change discontinuously if a threshold is crossed (Groffman *et al.*, 2006). The valuation of ecosystems services should therefore account for such nonlinearities if the scale under consideration is more than minimal (Barbier *et al.*, 2008; Koch *et al.*, 2009). The human values that apply to ecosystem services can also exhibit nonlinearities as the quantity or quality of the ecosystem service available changes (Bockstael *et al.*, 2000; Toman, 1998). Again, if the scale of a service under consideration is large, projecting a value estimated for a small change has the potential to produce significant bias. And if more than one service is evaluated, the aggregate value of the group of services is likely to be nonlinear with respect to considering more and more services. If estimates of individual ecosystem service values are used in such an exercise, simply adding them together will again produce a biased estimate of the group's value. These problems grow more and more significant as the scale of the ecosystem service valuation exercise increases. Given that this approach is often used to produce national or even global estimates (Anielski and Wilson, 2009; Beaumont *et al.*, 2007, 2008; Ingraham and Foster, 2008; Naidoo *et al.*, 2008), the issue of nonlinearities is an important one.

Second, the method requires rigid association of a particular set of ecosystem services to a particular landscape type; thus it does not allow assessment of how policies might change the flow of ecosystem services and affect their values, short of comparing values under wholesale destruction of the landscape type. Relying on landscape type to assign a fixed presence and quantity of services provides little information to evaluate how a policy will change ecosystem service values in any way other than a change in the landscape type. This "all or nothing" approach may make sense for some cases – for example, replacing forest with open agricultural fields – but not for others, as it is highly unlikely that any policy is capable of converting saltwater estuaries into upland forestland.

A third caution stems from the common use of benefit transfer for ecosystem service mapping exercises and other accounting approaches. The broad categories used in these types of exercises facilitate a comprehensive analysis of ecosystem service values for large geographic areas, but they also subject the analysis to a high likelihood of what is called *generalization error* (Plummer, 2009). As an example, consider the problem of estimating the value of providing "recreation" in a "marine area." *Recreation* covers scores of possible market and nonmarket activities that take place in natural settings, and some may be present in one marine area but not in others. Also, the presence of human-built infrastructure and the accessibility of the marine area are important determinants of the economic value of the recreational activity. Because these factors can vary widely, using estimates of economic value from one site and applying them to another is likely to produce significant errors if the categories of "recreation" and "marine area" are treated at a high level of generalization.

Finally, existing estimates of economic values for ecosystem services are strongly influenced by the methodology employed to make the estimate. As the NRC report on valuing ecosystem services (2005) emphasizes, not all methodologies are equally good. Happily, several valuation approaches are accepted practice by economists, but the interpretation of the resulting values should be informed by the method used.

Revealed Preferences Methods

Standard economic theory is based on the assumption that observable choices made by individuals reveal their expected valuation of a good or activity (but it also allows for some goods with values that are independent of observable behavior) (Slesnick, 1998). The revealed preference approach is a collection of methods for estimating economic values that rely on observable behavior. The most obvious revealed preference method is the use of market data. The production of commercially harvested finfish and shellfish is a leading example of marine ecosystem services with values that can be estimated with market data. Another revealed preference method that uses market data is known as *hedonic analysis* (Palmquist, 1991; Freeman, 1993; Palmquist, 2003); it analyzes goods (e.g., housing) that are sold as a bundle of

characteristics. In some cases, location-specific characteristics may be part of the bundle and include environmental amenities such as air and water quality or proximity to open space. If ecosystem services are a part of these bundled characteristics – for example, shoreline protection provided by nearshore ecological systems – the marginal value of such a service can be estimated with a hedonic analysis.

A final revealed preference method is useful for cultural ecosystem services such as recreation and other environmental experiences that take place outside any formal market. In this approach, the cost of engaging in the activity can be used to derive estimates of its economic value (Clawson, 1959; Knetsch, 1963). Similar to the assumptions for hedonic models, the recreation "good" can be viewed as a bundle of characteristics, some of which are the environmental features important to the recreational experience. If data are available for visits to multiple sites with varying levels of those features, one can then estimate the contribution of a particular feature to the demand for that recreation and from this estimate the feature's value (Morey, 1981).

Stated Preference Methods

Stated preference methods also can provide legitimate estimates of economic value. These methods rely on survey questions that ask individuals to make a choice, describe a behavior, or state directly what they would be willing to pay for specified changes in nonmarket goods or services. These methods are increasingly used in economic studies of environmental quality because they offer the opportunity to estimate the value for anything that can be presented as a credible and consequential choice. If conducted with attention to the many standards of care for its execution (Mitchell and Carson, 1989), this method can provide useful estimates of ecosystem service values.

Other Valuation Methods

Other methods are less reliable and particularly prone to misuse when estimating ecosystem service values. A prominent example is the replacement-cost or cost-of-treatment approach. For example, the cost of municipal water treatment for drinking water can be reduced by the presence of a wetland because the wetland system filters and removes pollutants (Day *et al.*, 2004). Using the cost of a human-built alternative treatment method is an approach sometimes taken to estimate the "value" of the wetland, but in general this cost does not have any necessary relation to the actual services provided by a particular wetland. On the one hand, the wetland may not actually provide filtration services (e.g., because of the absence of contaminants that need filtering or biophysical features of the wetland that prevent filtering). On the other hand, the wetland might simultaneously provide a number of other services (such as carbon storage, nursery habitat support for fisheries) that are not captured by this form of valuation. As a result, the replacement-cost approach to estimating the value of an ecosystem service must be used with great care.

One final word about valuation: much of the dialogue about ecosystem services focuses on economic valuation. Economic valuation, however, is far from the only way to value ecosystem services. Although many services (e.g., provisioning services) are relatively straightforward to value in economic terms, others (e.g., cultural services) often defy economic valuation. The comparison of multiple ecosystem services in a single currency (e.g., dollars) is appealing in some contexts but not appropriate or even desirable in others. Cultural ecosystem services – diverse, nonmaterial benefits that people obtain through their interactions with ecosystems, including spiritual inspiration, cultural identity, and recreation – are difficult but not impossible to value. Various methods (e.g., narrative methods, paired comparisons, structured decision making) can be used to elicit the relative weight that people place on these services (Chan *et al.*, in press).

Production Function Modeling

A production function approach is fundamentally process-based. It represents the relationship between inputs (e.g., the density of mangroves) and outputs (e.g., the degree of protection from storms). Production functions have been used extensively in agriculture, manufacturing, and other sectors of the economy. Ecological production functions can be used to explore how changes in ecosystem structure and function lead to changes in the flows of services (National Research Council, 2005).

Unlike the accounting methods previously discussed (*see* Accounting Approaches and Basic Valuation Methods), a production function approach allows for the comparison of alternative courses of action and their effects on ecosystems and services, and it can thus provide more useful information for decision making. With process-based models, analysts can explore how changes in inputs are likely to lead to changes in outputs. If valuation is of interest, then production function approaches can use both market prices and nonmarket valuation methods to estimate economic value and to show how the monetary value or other value currencies are likely to change under different environmental conditions.

The production function approach has been used to model various ecosystem services on land (Ricketts *et al.*, 2004; Kareiva *et al.*, 2011) and in marine systems (Barbier *et al.*, 2008; Holland *et al.*, 2010). In many cases, the approach is used to model a single ecosystem service, which is appropriate when it is functionally somewhat separate from other services and management actions are focused on that service. In other cases, the services have sufficiently strong connections that a multiple service approach is warranted. Both of these choices are explored in the following sections.

Modeling Single Ecosystem Services: Bioeconomic Models of Fisheries

A leading example of the modeling of a single ecosystem services in marine systems is the modeling of fisheries that incorporates both biological and economic components. For some types of fishery management decisions, a narrow focus on the single ecosystem service of food from fisheries is appropriate. It also provides important lessons for ecosystem service modeling because a critical part of its development has been the evolution of the human behavioral portion of the models. For that reason, an extended discussion

of this type of production function approach is provided in this section.

Of all marine ecosystem service modeling, fisheries models are by far the most sophisticated and have the longest history of use in managing marine systems. Bioeconomic models have been used since the early 1950s to explore how the value generated by fisheries is affected by how they are exploited and regulated. Bioeconomic models have evolved over time from simple models focused on equilibrium outcomes associated with static annual harvests and effort levels (e.g., Gordon, 1954; Schaefer, 1957) to more complex dynamic models that explore the implications of the timing, location, and methods of harvests, the linkages between different species, and the impacts of fisheries on other ecosystem services. From the beginning, these models have emphasized the interconnection of the natural and human components of fisheries and the importance of an approach to modeling that treats human activity as an endogenous part of the system.

It was recognized early on that understanding the optimal exploitation of fisheries would generally require a dynamic modeling approach since fisheries are rarely if ever in equilibrium at some desired state. Scott (1955) introduced the concept of *user cost* to illustrate the trade-offs between harvesting fish today and leaving fish in the water to contribute to future growth. Viewing the fish stock as natural capital, Scott showed that optimal exploitation would require the rate of return from the *in situ* fish stock to equal the rate of return from other uses of that capital (i.e., by extracting and selling fish and thereby transforming them into economic capital).

Although the initial focus of bioeconomic models was on aggregate annual effort and harvests, Gordon's (1954) analysis demonstrated that the value generated by the fishery was also dependent on the spatial distribution of fishing. The exploitation of fisheries over space – and particularly the issue of whether to close some areas to fishing to conserve fish stocks, habitat, and biodiversity – became an important focus of bioeconomic modeling in the late 1990s (e.g., Holland and Brazee, 1996; Sanchirico and Wilen, 1999) and remains so (Costello and Polasky, 2008). It is becoming increasingly clear that spatial heterogeneity of marine ecosystems and of fishing costs provide opportunities to increase the value generated by fishing by managing where and when fishing takes place and which harvest methods are employed. These models show that optimal exploitation depends on relative costs over space in conjunction with spatial variation in productivity (e.g., from differences in suitability and productivity of habitat for species and their prey or from metapopulations with source-sink dynamics).

Another important focus of bioeconomic modeling that has become more prevalent over time is the multispecies nature of fisheries. Many multispecies fishery models focus on the implications of technical interactions that make it difficult and costly to control the relative catch of different species that occupy the same habitat. These models provide insights into how economic incentives or regulations on gear or fishing locations can influence catch composition to increase the value and sustainability of fisheries and reduce undesirable bycatch. Trophic linkages between different fish species are often thought to be important drivers of fishery productivity and stability. However, they have largely been ignored in the assessment of the productivity of fish stocks and when providing management advice (but see Hollowed *et al.*, 2011). This is not surprising: these relationships are typically poorly understood and difficult to quantify.

Nonetheless, there is recognition that understanding and responding to these species interactions can increase value and improve the sustainability of fisheries. Ecosystem models that attempt to depict entire food webs and the impacts of bottom-up forcing (e.g., by climate) and top-down forcing (e.g., predation and fishing) have been developed in recent years; a few such as Atlantis (Fulton *et al.*, 2004a, 2004b) include an endogenous human harvester component to the model. These models offer the possibility of informing a more holistic ecosystem-based approach to managing fisheries and other activities in the marine environment, but their complexity and the lack of data to parameterize them has so far limited their usefulness in providing specific tactical advice for managing individual fisheries. An increasing number of multispecies bioeconomic models focus on small systems of exploited species (see, e.g., the discussion in the section The American Lobster, Atlantic Herring, and Northeast Multispecies Groundfish Fisheries in the Gulf of Maine). For these smaller systems made up of primarily commercial species, there is often more data and research, making it more feasible to parameterize the models accurately enough to quantitatively evaluate multispecies harvest strategies.

Our ability to understand and model fishery systems in the marine environment is hindered by an inability to observe directly the processes that drive outcomes since they occur under the sea and over vast areas. It is often the perturbations caused by human exploitation of the system and the data gathered by fishermen that enable us to develop an understanding of these systems that is sufficient to create empirically based models capable of providing specific insights that can be applied to the management of these systems as we seek sustainable harvest levels. Bioeconometric models, defined by Smith (2008) as structural models that econometrically estimate one or more parameters of the bioeconomic system, have the potential to reveal both ecological and economic parameters of the system simultaneously and to greatly increase our understanding of these systems at relatively low cost (since they exploit data easily collected by fishermen in the course of normal operations).

This field is not new as it includes empirical applications of the Gordon-Schaefer model dating to the 1950s – but until recently was mostly limited to equilibrium models based on aggregate annual catch and effort. These models did not reveal the underlying microecological and economic processes driving the results and required making strong assumptions about system behavior (e.g., that it tends toward equilibria). In recent years, progress has been made in modeling and parameterizing these processes with dynamic models and disaggregated data. This has made it more feasible to model and predict how changes in the underlying ecological system or changes in factors that drive human behavior will change outcomes and the resulting flow of ecosystem services (e.g., Smith *et al.*, 2008).

Efforts to model a single ecosystem service – such as the critical service of food from fisheries – and the ways in which its delivery or value change under different scenarios are of great utility when decisions are being made sector by sector or when a better understanding of a key service sheds light on a previously understudied aspect of a system. But with decision makers increasingly grappling with multisector decisions, modeling of multiple ecosystem services in the same context can allow for more informed decision making.

Modeling Multiple Ecosystem Services and Examining Trade-Offs

Single-sector management that proceeds without the explicit consideration of how single decisions impact the full suite of things people care about and need can lead to misguided assessments of the true costs and benefits to the environment and society of natural resource management options. For example, mangroves are routinely cleared and the resultant open areas used for shrimp aquaculture. A singular focus on aquaculture as an ecosystem service derived from cleared areas often shows a positive bottom line, as the high market price of shrimp provides strong support for this action from a private perspective. Standing mangroves provide other social and ecological benefits that private owners cannot capture, however. More complete accounting using a multiple ecosystem service framework shows that keeping mangroves intact often has higher social benefits once other services such as wood products, support for offshore fisheries, and coastal protection are taken into account (Sathirathai and Barbier, 2001). In other words, a single-sector approach risks ignoring the multitude of connections among components of natural and social systems. These connections are often important for the maintenance of ecosystem health, human well-being, and the sector of interest itself (Millennium Ecosystem Assessment, 2005b).

The explicit recognition of connections between activities and their consequences for multiple ecosystem benefits allows for the exploration of trade-offs and win–wins. Rodríguez *et al.* (2006) review some of the most frequent ecosystem service trade-offs faced by society. In some cases, trade-offs and win–wins can be explored using a common currency (e.g., dollars, Sathirathai and Barbier, 2001), but various metrics can be used (Tallis *et al.*, 2008, 2011; Lester *et al.*, 2010). As one example, the MA explored trade-offs among various categories of ecosystem services (with axes for each service scaled from −1 to 1 to represent positive or negative change from the baseline) for each of four heuristic scenarios (Millennium Ecosystem Assessment, 2005a).

Explorations of multiple ecosystem services and how they are likely to change under various management scenarios has proceeded in two important directions: detailed explorations of particular situations and the development of tools designed to be applicable in various contexts. We focus here on tools. Research teams have taken a number of different approaches to modeling the flow of multiple ecosystem services and examining trade-offs between them. Here we provide a brief introduction to some of the approaches and tools that are most applicable to modeling marine ecosystem services (**Table 2**).

Modeling the flows of ecosystem services can take many forms. Artificial Intelligence for Ecosystem Services (ARIES; www.ariesonline.org) offers a range of approaches including probabilistic Bayesian models, machine learning, and pattern recognition to assess the provision, use, and flow of ecosystem services on a landscape. The tool allows users to evaluate and compare alternative policy and land-use scenarios and their impacts on ecosystem services. The initial marine application of ARIES is in Madagascar in partnership with the United Nations Environment Program's (UNEP's) World Conservation Monitoring Center, where a benefits-transfer approach to ecosystem service valuation is being used (Center for Ocean Solutions, 2011).

More process-based ecosystem service modeling approaches also have emerged. The challenges of applying the land-use or land-cover map-based approach used in terrestrial systems, coupled with the strong foundations of marine ecosystem modeling, has made ecosystem modeling central to the ecosystem services framework for marine systems. Foodweb models such as EcoPath with EcoSim (EwE) (Christensen and Walters, 2004) and ecosystem models such as Atlantis (Fulton *et al.*, 2004a, 2004b) have been used to explore fishery management options in an ecosystem context. These models also have the potential to evaluate a more comprehensive suite of human activities and the ways in which they change the delivery of ecosystem services. Using these models, explorations of the potential outcomes of various human actions can lead to the development of management approaches that consider multiple objectives and trade-offs among objectives.

Integrated Valuation of Ecosystem Services and Trade-offs (InVEST; http://www.naturalcapitalproject.org) helps decision makers visualize the impacts of potential management activities by modeling and mapping the delivery, distribution, and economic value of terrestrial, freshwater, and marine ecosystem services under alternative scenarios (for more information, see Kareiva *et al.* (2011) and Guerry *et al.* (2012)). InVEST can be used to identify trade-offs and compatibilities between environmental, economic, and social benefits. Applicable at a range of scales—from local to global—InVEST was designed to play a key role in real-world decision-making processes. The first phase of the approach involves working with stakeholders to identify critical management decisions and to develop scenarios that project how the provision of services might change in response to those decisions as well as to changing climate, population, and so on. Using these scenarios and basic biophysical and social data as inputs, InVEST quantifies and maps a broad range of ecosystem services. To facilitate the use of InVEST in real decision-making contexts, a key design feature of InVEST is the relative simplicity of data required and the free availability of models on the web. InVEST outputs provide decision makers with information about costs, benefits, trade-offs, and synergies of alternative management strategies.

InVEST uses a primarily production function approach to assess how environmental change, as a result of future policies and management decisions, affects the delivery of ecosystem services. Terrestrial InVEST models have been applied around the world to inform a variety of land-use decisions (Kareiva *et al.*, 2011). InVEST's marine tools (see Guerry *et al.*, 2012) were initially released in 2011 and are being applied in Canada, the US, and Belize. Marine InVEST includes models

Table 2 Decision support tools for use in marine environments, with particular attention to: marine applications, spatial mapping of ecosystem services, capacity to analyze trade-offs between services, and the level of technical expertise needed to use the tool

Decision support tool/developer	*Description*	*Marine applications*	*Spatial mapping of ecosystem services*	*Analysis of ecosystem service trade-offs*	*Level of technical expertise needed*[a]
Artificial Intelligence for Ecosystem Services (ARIES)					
Basque Center for Climate Change (BC3); University of Vermont—Gund Institute for Ecological Economics, Conservation International, and Earth Economics	ARIES is a suite of web-accessible applications—including probabilistic Bayesian models, machine learning, and pattern recognition— used to assess the provision, use, and flow of ecosystem services. ARIES maps both the sources of ecosystem services and their users, along with the flows from ecosystems to users.	ARIES has been used to assess subsistence fisheries and coastal storm regulation in Madagascar.	Yes	Yes	2, 3
Atlantis					
Commonwealth Scientific and Industrial Research Organization (CSIRO) Marine and Atmospheric Research	Atlantis is a three-dimensional, spatially explicit model that incorporates biogeochemical dynamics and fishing behavior. Submodels cover food-web relations, hydrographic processes, and fisheries.	Atlantis has been used to evaluate restructuring of Southeastern Australia fishing fleets, the NOAA Integrated Ecosystem Assessment for the California Current, the Marine Stewardship Council Forage Fish Harvest Guidelines, and groundfish fleet impacts on protected marine mammals in the California Current.	No	Yes	3
Coastal Resilience					
The Nature Conservancy, University of Southern Mississippi, and University of California, Santa Barbara	Coastal Resilience is a spatial planning and action approach that integrates appropriate coastal hazard, ecological, and socioeconomic information within a particular geography. The Coastal Resilience approach is to map sea-level rise and other coastal hazards, natural resources, and human communities at risk and display this information via an internet mapping application that is a data viewer, data discovery tool, and a future scenario mapper.	Coastal Resilience has been used for data exploration with the New York State Emergency Management Office and local towns and villages on Long Island and the Connecticut shores.	No	No	2

Cumulative Impacts					
National Center for Ecological Analysis and Synthesis (NCEAS), University of California, Santa Barbara, and Stanford University	The Cumulative Impacts model uses spatial data and expert opinion to assess the ecological consequences of 17 types of human activities. By overlaying maps of human activities and ecosystem vulnerabilities, this model produces a cumulative mapping of ecological impacts.	Cumulative Impacts has been used in a coarse global analysis and in the California Current and the northwestern Hawaiian Islands.	No	No	2
InVEST					
The Natural Capital Project—Stanford University, World Wildlife Fund, The Nature Conservancy, and the University of Minnesota	InVEST is composed of a number of models for assessing flows of and changes in different ecosystem services including, but not limited to, carbon storage, wave energy, recreation, fishery production, erosion control, habitat quality, water quality, crop pollination, and timber production. InVEST is a toolbox in ArcGIS and runs on both spatial and nonspatial physical, biological, and economic data and information. (An ArcGIS-independent version is forthcoming.)	InVEST has been used for marine applications in Canada (west coast of Vancouver Island) and Belize, and for land–sea connections in Puget Sound, Galveston Bay, and Chesapeake Bay (US). In addition, it is being used in climate adaptation planning and to inform restoration in the Gulf of Mexico and in Monterey Bay (both US).	Yes	Yes	1, 3
Marine Map					
MarineMap Consortium—University of California, Santa Barbara, The Nature Conservancy, and EcoTrust	MarineMap is a web-based decision support toolkit to support marine spatial planning processes. The toolkit includes a spatial data viewer and design tools that allow users to networks of prospective marine protected areas. The application also allows users to share their proposals with others and evaluate their proposals against goals defined in the course of any planning process.	MarineMap has been used for the California MLPA Initiative and the Oregon Territorial Sea Planning process (US).	No	Yes	1
Marxan with zones					
University of Queensland	Marxan is a program for identifying combinations of sites to create conservation networks such as marine protected areas. The program allows the user to set biodiversity and other	Marxan has been applied to Marine Zoning in Saint Kitts and Nevis and to examine four types of protected areas in the context of California's Marine Life Protection Act.	No	Yes	2, 3

(*Continued*)

Table 2 Continued

Decision support tool/developer	*Description*	*Marine applications*	*Spatial mapping of ecosystem services*	*Analysis of ecosystem service trade-offs*	*Level of technical expertise needed*[a]
	targets for one or more zones and then find the set of sites that meets those targets while minimizing the cost of the network.				
Multiscale Integrated Models of Ecosystem Services (MIMES)					
AFORDable futures	MIMES is a multiscale integrated suite of models that incorporate stakeholder input and a variety of data sets to assess trade-offs among multiple ecosystem services. The models simulate ecological and socioeconomic systems and their interactions, and they calculate the values of ecosystem services for different scenarios.	MIMES is being used by the Massachusetts Ocean Partnership to examine the trade-offs between different sectors in spatial planning and to model ecosystem service values at multiple scales.	Yes	Yes	2, 3
Multipurpose Marine Cadastre (MMC)					
Bureau of Ocean Energy Management, Regulation and Enforcement (BOEMRE—formerly Minerals Management Service) and NOAA Coastal Services Center	Originally developed to support the assessment of offshore energy projects, MMC is a web-based geospatial data viewer containing more than 80 data layers from a variety of sources that each can be turned on or off or queried one at a time. The user has the ability to draw lines and other features, create buffers, and perform other geospatial actions.	MMC has been used for reviews of ocean energy projects in Northern California and the outer continental shelf off Massachusetts, and by the Mid-Atlantic Regional Council on the Ocean to support marine spatial planning efforts.	No	No	1

[a]Levels of technical expertise (when more than one level is listed the tool has capabilities that require varying levels of expertise):

1. Minimal training or technical expertise.
2. Minimal training and expertise but process objectives must be set in advance.
3. Expert users.

Source: Adapted from Center for Ocean Solutions (2011) *Decision Guide: Selecting Decision Support Tools for Marine Spatial Planning*. Stanford, CA: The Woods Institute for the Environment, Stanford University.

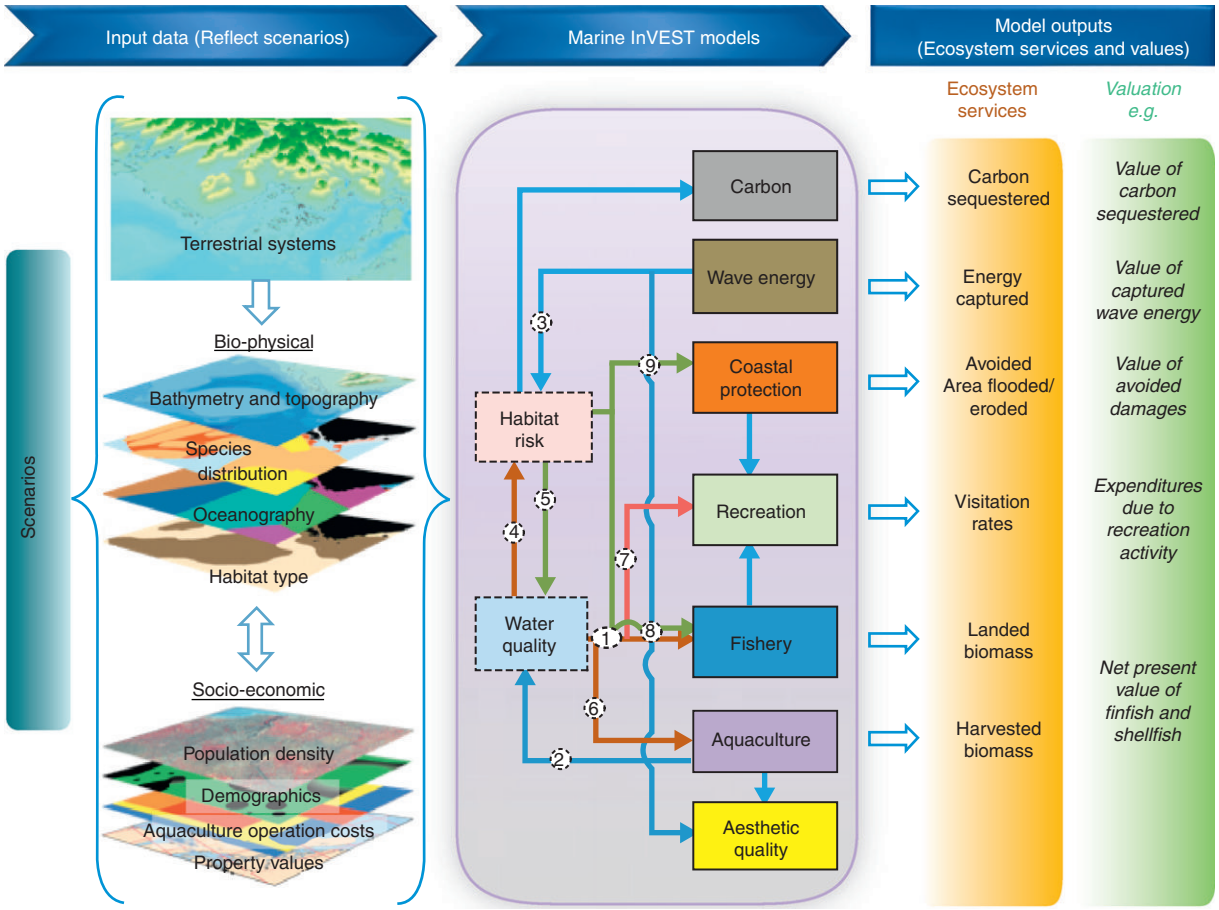


Figure 2 Marine InVEST evaluates how alternative scenarios yield changes in the flow of ecosystem services. First, one translates management or climate scenarios into input data. Inputs include spatially explicit biophysical and socioeconomic information. Next, one feeds input maps into models that predict the delivery of services across the seascape. Intermediate effects of management choices and climate on the flow of services can be evaluated in terms of risks to habitats and changes in water quality. Ecosystem service outputs are expressed in biophysical or socioeconomic units. Some examples of processes or activities that link the models include (numbers correspond to circled numbers on arrows): (1) flow rate and the movement of sediment, nutrients, toxic waste and bacteria; (2) filtration, flows of waste; (3) limiting fishing grounds; (4) light attenuation, sedimentation; (5) water purification; (6) food supply (shellfish); (7) beachgoing; (8) spawning, rearing; and (9) wave attenuation. Adapted from Guerry AD, Ruckelshaus MH, Arkema KK, *et al.* (2012) Modeling benefits from nature: using ecosystem services to inform marine spatial planning. *International Journal of Biodiversity Science, Ecosystem Services and Management*. doi: 10.1080/21513732.2011.647835.

for renewable energy, protection from coastal hazards, fisheries, recreation, aquaculture, aesthetic views, and more (**Figure 2**).

The Multiscale Integrated Models of Ecosystem Services (MIMES) project represents another attempt to build a suite of models that assess the values of ecosystem services to allow managers to understand the dynamics of ecosystem services, how services are linked to human welfare, and how the flow of services might change under various management scenarios in both terrestrial and marine systems (http://www.uvm.edu/giee/mimes/). MIMES involves a relatively complex approach, simulating ecosystems and socioeconomic systems and the interactions between them. It calculates values of ecosystem services using marginal cost pricing. MIMES is currently being used by the Massachusetts Ocean Partnership to evaluate the trade-offs between different sectors in marine spatial planning (Center for Ocean Solutions, 2011).

Applications of Marine Ecosystem Service Modeling to Decision Making

Decision contexts for managing ecosystem services in the sea are varied, so it is to be expected that a mix of quantitative and qualitative modeling approaches are being developed and applied to support biodiversity and ecosystem service management in marine systems. Information needed to support the methodologies we highlight in the following sections ranges from observations and empirical data to modeled ecosystem and social responses, expert opinion and traditional local knowledge (e.g., Kliskey *et al.*, 2009). Typically, the time frames of decisions, local technical capacity, and the nature of information available for each context dictate the approach employed for ecosystem service modeling. We explore four contexts in the following section that display a wide range of the numbers of services modeled, quantitative complexity, and types of decisions being made.

We start where marine modelers and managers are most at home together – with a fisheries-focused example – and then move to efforts to include a number of different types of services.

The American Lobster, Atlantic Herring, and Northeast Multispecies Groundfish Fisheries in the Gulf of Maine

The American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries in the Gulf of Maine provide an excellent example of a system of interlinked fisheries with a value that might be substantially increased by accounting for linkages between them and with the environment. A number of important environmental, ecological, and economic linkages within these fisheries and with the environment were identified as key processes to study and model (**Figure 3**). This system is the subject of an interdisciplinary research project funded by the National Science Foundation aimed at understanding these linkages through empirical research and modeling to provide insights into how the management of each fishery and the system as a whole can be improved. The strategy of this research effort is to link and build on the single-species models that are already used for providing management advice by modeling key linkages between the fisheries and with the environment that might be exploited to increase value and sustainability.

The primary focus for environmental linkages in this system is on impacts of temperature, salinity, and winds on recruitment and juvenile growth. Changes in the abundance of larval food and changes in circulation driven by wind and currents are thought to be key drivers of recruitment variability for all three species. Determining what drives recruitment success requires a combination of empirical fieldwork (determining where and how much spawning and settlement take place) and coupled biophysical models (evaluating how currents, winds, and temperatures affect dispersal and settlement of larvae and consequent recruitment). In addition to circulation and dispersion, model calculations may need to consider spatial patterns of egg production, temporal patterns of hatching, temperature-dependent development, vertical distribution, and mortality (see Incze *et al.*, 2010, and Churchill *et al.*, 2011, for examples that are relevant to this system).

A variety of ecological linkages, natural and artificial, are being studied, including predation of groundfish on herring and juvenile lobsters and predation of herring on groundfish larvae. In addition to predator–prey linkages, the team is studying behavior modification as a result of diminished presence of groundfish predators that may have allowed lobster to expand use of habitat and consequently carrying capacity.

This fishery system is somewhat unusual, though perhaps not unique, in that a key linkage between the fisheries is controlled by humans. The artificial trophic linkage resulting from use of herring bait in quantities (exceeding 50,000 tons

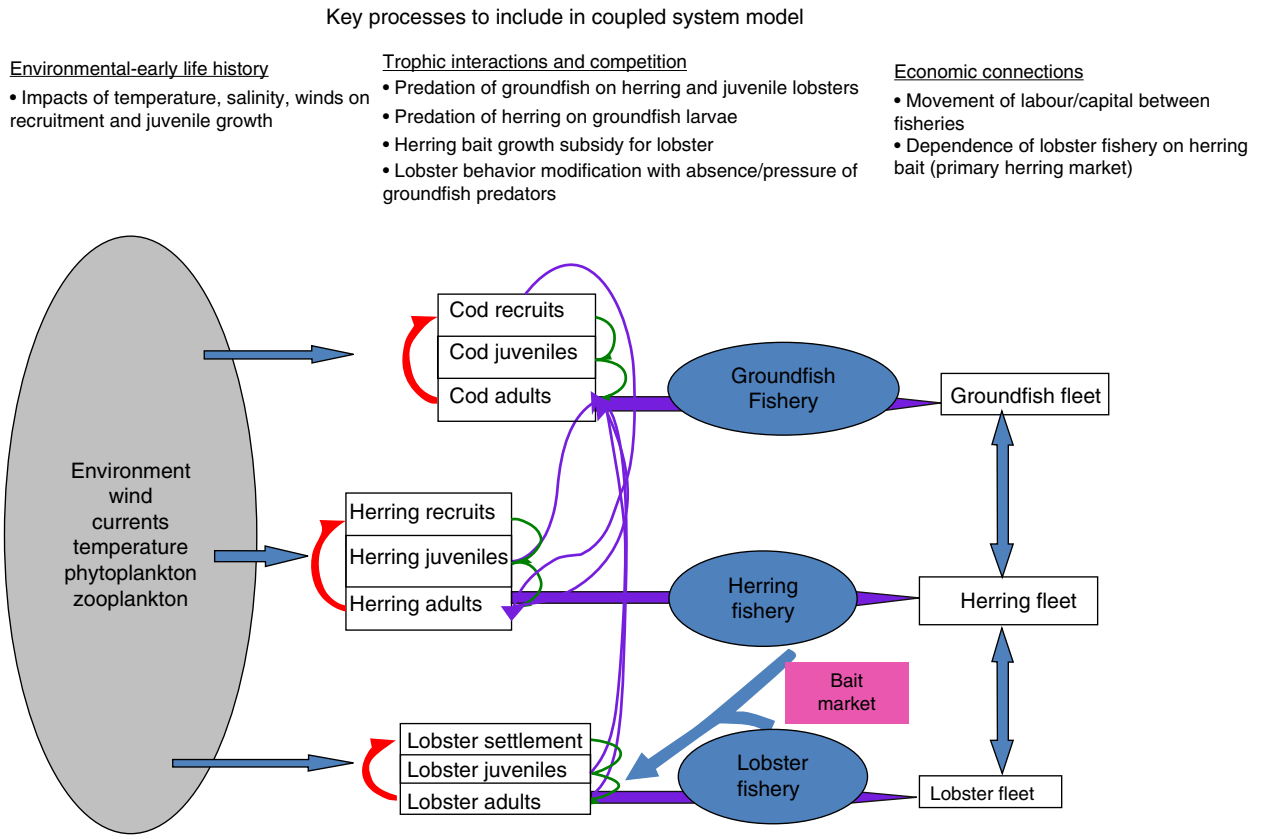


Figure 3 Linkages between lobster, herring, and groundfish fisheries in the Gulf of Maine.

annually in recent years) is thought to be responsible for a substantial increase in productivity of the lobster stock (Grabowski *et al.*, 2010). The way that the lobster fishery is managed substantially alters the amount of herring bait demanded by the fishery, which affects profitability of both fisheries (Holland *et al.*, 2010) and how those profits are distributed (Ryan *et al.*, 2010).

Understanding these linkages among the fisheries and the environment provides useful qualitative insights for resource managers and stakeholders, and it may be possible to use this information to improve single-species models and management advice to account for them as exogenous factors. Ideally, a linked dynamic set of fishery and environmental models will be constructed enabling managers to better understand how changes in one part of the system affect other parts and how best to react or even manipulate the system. In addition to evaluating potential management actions or forecasting future outcomes in the fishery system, models of the system can also be used to determine which variables and processes have the most impact on the system. This knowledge can be used to determine the value of better information and direct research. Ultimately, this information can be used to manage these three key fisheries in a systemwide manner that maximizes the service of provisioning of seafood and the various benefits that flow from that service (e.g., revenues, working waterfronts, community identity).

Coastal Zone Management in Belize

Belize is home to a rich diversity of ocean life, coastal habitats, and the longest barrier reef in the Western hemisphere. Its people are inextricably linked to the marine environment as a source of sustenance, inspiration, economic prosperity, and cultural heritage (**Figure 4**). Yet rapid development, overfishing, and population growth threaten local marine ecosystems. In 1998, the government created the Coastal Zone Management Authority and Institute (CZMAI) with the objective of developing an integrated management plan to guide sustainable economic growth that is consistent with protection of Belize's natural heritage. To date, however, CZMAI has lacked the scientific capacity needed to assess trade-offs among various marine uses and to demonstrate potential win–wins where emphasis on one service yields benefits for another.



Figure 4 Belize is home to a rich diversity of marine life; its people derive sustenance, inspiration, economic prosperity, and cultural heritage from the marine environment. Mangroves and coral reefs not only provide protection from storms and serve as a draw for international tourism, but also are critical habitats for the spiny lobster, a backbone of commercial and artisanal fisheries. The commercial fishermen in this picture are from a fleet of dugout canoes that work with a traditional sail boat (or "mother ship"') to harvest spiny lobster.

CZMAI has partnered with the Natural Capital Project to help create a comprehensive coastal zone management plan for Belize. The plan will identify new marine protected areas, locations suitable for coastal development, and strategies such as payments for ecosystem services (PES) and other market mechanisms (e.g., catch shares) to support effective implementation. The partners are using Marine InVEST to forecast how management strategies and future uses of the marine environment will likely affect the benefits that nature brings to people, such as nursery habitat for fisheries, tourism and recreation, coastal protection, and carbon storage. By providing a platform for stakeholder and agency discussions about trade-offs and win–win situations, this information is moving the process beyond sector-specific issues to the development of a defensible plan.

The first step toward developing a coastal zone plan informed by the science of ecosystem services is to identify and map current and potential future uses of marine and coastal environments. Next is to use modeling approaches (in this case, the InVEST decision support tool) to examine how these activities affect the ecosystem services most important to Belizeans. The team is running marine InVEST models, given future scenarios of marine use, to quantify changes in the ecosystem services that are most compelling to stakeholders and government officials. Scenarios under consideration include explorations of how coastal development and no-take areas – in addition to climate change – might affect services, including (1) provision of commercial and artisanal fisheries for lobster; (2) flooding and erosion protection provided by mangroves, corals, and seagrasses; and (3) maintenance of tourism attractions such as snorkeling and diving. Quantitative outputs in biophysical (e.g., biomass of harvested fish, reduction in land flooded or eroded due to mangroves) and economic (e.g., net present value of harvest, avoided damages to property values) units for each service are informing CZMAI's coastal planning process.

Modeling of ecosystem services in Belize is helping to articulate connections between human activities that are often considered in isolation to align diverse stakeholders around common goals and to make implicit decisions explicit. Ecosystem service modeling results have informed early iterations of the coastal zone plan and will inform the creation of the final plan in 2012.

Protection and Restoration of Reef Habitats in the Coral Triangle

Klein *et al.* (2010) used a combination of quantitative modeling and qualitative information to prioritize strategies for protection and restoration of reef habitats in the coral triangle. The method is designed to allocate a limited budget to management and policy interventions that will abate key threats

affecting marine ecosystem objectives. The approach requires estimates of management costs and opportunity costs of applying alternative actions; and these estimates are user-generated, so they can come from expert judgment, models, or empirical information. The models provide an estimate of the return on investment for different actions so that the relative merits of land- and marine-based interventions on marine ecosystem biodiversity and services can be weighed. This type of practical approach, where a mix of quantitative and expert information is used, allows the users to consider both economic and ecological criteria in a conservation management context. For example, for the ecoregions of the coral triangle, marine-based conservation efforts (e.g., establishment of MPAs, changes in fishing gear types) were often more cost-effective at improving marine ecosystem condition than land-based actions designed to reduce nutrient or sediment runoff.

Putting Offshore Wind in Context in Massachusetts

Like many coastal marine ecosystems, the coast of Massachusetts (US) is crowded with various user groups interested in using a broad range of ecosystem services. Conflicts among sectors in their ability to procure different services from shared, interacting ecosystem resources has prompted calls for the reduction of conflicts through transparent, integrated marine spatial planning. In Massachusetts and elsewhere, offshore wind farms are an emerging, and controversial, ocean use. Numerous wind farms are under litigation or consideration in the state. Large economic gains and "green" energy are purported benefits; impacts on marine mammals and fisheries are some of the potential costs. To explore these issues and inform the dialogue about this new and emerging use of the marine environment, White and colleagues (in press) performed a spatially explicit bioeconomic trade-off analysis among wind energy, fishery, and whale-watching ecosystem services in Massachusetts Bay. They identified optimal planning solutions for wind-energy development that minimize spatial conflicts among sectors and maximize their values. Solutions that considered all three sectors were significantly better than single-sector management outcomes. Solutions were also strongly dependent on the spatial design of the wind farms and the particular fishery examined. This approach of quantifying ecosystem service trade-offs in a crowded, multi-use ecosystem highlights when and how offshore renewable energy may be developed optimally within the complicated contexts of coastal ecosystems.

Conclusions

Marine ecosystems around the globe are under increasing pressure: people rely on them for the delivery of provisioning, regulating, supporting, and cultural ecosystem services. With increasing pressure from increasing and redistributing human populations and from new and emerging uses (e.g., renewable energy), the ability of these systems to sustainably deliver the full range of services that people count on and need is not assured.

Here we have reviewed the services provided by marine environments, discussed some of the potential audiences for information about how changes in marine ecosystems are likely to lead to changes in services, explored critical differences – both in human and scientific contexts – between mapping and modeling ecosystem services on land and at sea, outlined approaches to modeling marine ecosystem services, and described some examples of how these approaches can and are being used in real decision making.

Quantifying, mapping, and valuing marine ecosystem services have the potential to fundamentally change decision making in marine and coastal environments. Making explicit the connections between human activities in one sector and their effects on a broad range of other sectors leads people to think about whole ecosystems and to manage them accordingly. Ultimately, results from marine ecosystem service modeling can help society perceive the critical services oceans and coasts provide, appropriately value marine natural capital, and help human communities make better choices about the use of these life-sustaining environments.

Appendix

List of Courses

1. Marine Ecosystem Services
2. Marine Conservation Biology
3. Marine Ecology
4. Marine Resource Economics

See also: Aquaculture. Biodiversity and Ecosystem Services. Biodiversity, Human Well-Being, and Markets. Carbon Cycle. Census of Marine Life. Computer Systems and Models, Use of. Economic Value of Biodiversity, Measurements of. Economics of the Regulating Services. Ecosystem, Concept of. Ecosystem Function Measurement, Aquatic and Marine Communities. Ecosystem Function, Principles of. Ecosystem Services. Endangered Marine Invertebrates. Estuarine Ecosystems. Evaluation of Ecosystem Service Policies from Biophysical and Social Perspectives: The Case of China. Fish Conservation. Mangrove Ecosystems. Marine Conservation in a Changing Climate. Marine Ecosystems, Human Impacts on. Marine Ecosystems. Marine Protected Areas: Static Boundaries in a Changing World. Modeling Terrestrial Ecosystem Services. Ocean Ecosystems. Priority Setting for Biodiversity and Ecosystem Services. Resource Exploitation, Fisheries. The Value of Biodiversity. Valuing Ecosystem Services

References

Agardy T, Alder J, Dayton P, *et al.* (2005) Coastal Systems. Assessment report. *Millennium Ecosystem Assessment*, ch 19.

Anielski M and Wilson SJ (2009) *Counting Canada's Natural Capital: Assessing the Real Value of Canada's Boreal Ecosystems*. Drayton Valley, Alberta: Canadian Boreal Initiative & The Pembina Institute.

Balvanera P, Pfisterer A, Buchmann N, *et al.* (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9: 1146–1156.

Barbier E, Koch E, Siliman B, *et al.* (2008) Coastal ecosystem-based management with nonlinear ecological functions and values. *Science* 319: 321–323.
Basurto X and Coleman E (2010) Institutional and ecological interplay for successful self-governance of community-based fisheries. *Ecological Economics* 69: 1094–1103.
Beaumont N, Austen M, Atkins J, *et al.* (2007) Identification, definition and quantification of goods and services provided by marine biodiversity: Implications for the ecosystem approach. *Marine Pollution Bulletin* 54: 253–265.
Beaumont NJ, Austen MC, Mangi SC, and Townsend M (2008) Economic valuation for the conservation of marine biodiversity. *Marine Pollution Bulletin* 56: 386–396.
Bockstael NE, Freeman III AM, Kopp RJ, Portney PR, and Smith VK (2000) On measuring economic values for nature. *Environmental Science and Technology* 34: 1384–1389.
Boyd J (2007) The Endpoint Problem. *Resources* 165: 26–28.
Boyd J and Banzhaf S (2007) What are ecosystem services? The need for standardized environmental accounting units. *Ecological Economics* 63: 616–626.
Brander K (2010) Reconciling biodiversity conservation and marine capture fisheries production. *Current Opinions in Environmental Sustainability* 2: 416–421.
Burke L, Kura Y, Kassem K, Revenga C, Spalding M, and McAllister D (2001) *Pilot Analysis of Global Ecosystems: Coastal Ecosystems*. Washington, DC: World Resources Institute.
Caddy JF (1999) Fisheries management in the twenty-first century: Will new paradigms apply? *Reviews in Fish Biology and Fisheries* 9: 1–43.
Carpenter S, Caraco N, Correll D, Howarth R, Sharpley A, and Smith V (1998) Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications* 8: 559–568.
Carte BK (1996) Biomedical potential of marine natural products. *BioScience* 46: 271–286.
Center for Ocean Solutions (2011) *Decision-Guide: Selecting Decision-Suport Tools for Marine Spatial Planning*. Stanford, CA: The Woods Institute for the Environment, Stanford University.
Chan K, Guerry A, Satterfield T, *et al.* (in press) Integrating "cultural" and "social" into ecosystem services: A framework for making decisions about what matters. *BioScience*.
Chan K and Ruckelshaus M (2010) Characterizing changes in marine ecosystem services. *F1000 Biology Reports* 2: 54–60.
Christensen V and Walters C (2004) Ecopath with Ecosim: Methods, capabilities and limitations. *Ecological Modeling* 172: 109–139.
Churchill JH, Runge J, and Chen C (2011) Processes controlling retention of spring-spawned Atlantic cod (*Gadus morhua*) in the western Gulf of Maine and their relationship to an index of recruitment success. *Fisheries Oceanography* 20: 32–46.
Clawson M (1959) *Methods of Measuring the Demand and Value of Outdoor Recreation*. Washington, DC: Resources for the Future.
Colling A (2001) *Ocean Circulation*. Milton Keynes, UK: Butterworth Heineman and The Open University.
Costanza R (2000) The ecological, economic and social importance of the oceans. In: Sheppard CRC (ed.) *Seas at the Millenium: An Environmental Evaluation, vol. 3: Global Issues and Processes*. New York: Pergamon Press.
Costanza R, d'Arge R, de Groot R, *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature* 387: 253–260.
Costello C and Polasky S (2008) Optimal harvesting of stochastic spatial resources. *Journal of Environmental Economics and Management* 56: 1–18.
Cudney-Bueno R and Basurto X (2009) Lack of cross-scale linkages reduces robustness of community-based fisheries management. *PLoS One* 4: e6253.
Culliton T (1998) *Population: Distribution, Density, and Growth. National Oceanic and Atmospheric Administration*. MD: Silver Spring.
Danielsen F, Sorensen M, Olwig M, *et al.* (2005) The asian tsunami: A protective role for coastal vegetation. *Science* 370: 643.
Day Jr. J, Ko J, Rybczyk J, *et al.* (2004) The use of wetlands in the Mississippi Delta for wastewater assimilation: A review. *Ocean and Coastal Management* 47: 671–691.
deGroot R, Wilson M, and Boumans R (2002) Ecosystem functions, goods and services: Classification, description and valuation guidelines. *Ecological Economics* 41: 393–408.
Dennison W, Lookingbill T, Carruthers T, Hawkey J, and Carter S (2007) An eye-opening approach to developing and communicating integrated environmental assessments. *Frontiers in Ecology and the Environment* 5: 307–314.
Diaz R and Rosenberg R (2008) Spreading dead zones and consequences for marine ecosystems. *Science* 321: 926–929.
Douvere F and Ehler C (2009) New perspectives on sea use management: Initial findings from European experience with marine spatial planning. *Journal of Environmental Management* 90: 77–88.
Ehler C and Douvere F (2009) *Mrine spatial planning: A step-by-step approach toward ecosystem-based management*. Paris: UNESCO.
European Commission (2010) *Maritime Spatial Planning in the EU: Achievements and Future Development*. Brussels, Belgium: European Commission.
FAO Fisheries Department (2010) *The State of World Fisheries and Aquaculture. FAO Fisheries and Aquaculture Department*. Rome: Food and Agriculture Organization of the United Nations.
Foley M, Halpern B, Micheli F, *et al.* (2010) Guiding ecological principles for marine spatial planning. *Marine Policy* 34: 955–966.
Freeman III AM (1993) *The Measurement of Environmental and Resource Values: Theory and Methods*. Washington, DC: Resources for the Future.
Fulton EA, Parslow JS, Smith ADM, and Johnson CR (2004a) Biogeochemical marine ecosystem models II: The effect of physiological detail on model performance. *Ecological Modelling* 173: 371–406.
Fulton EA, Smith ADM, and Johnson CR (2004b) Biogeochemical marine ecosystem models. I: IGBEM – a model of marine bay ecosystems. *Ecological Modelling* 174: 267–307.
Gleick P (1996) Water resources. In: Schneider S (ed.) *Encyclopedia of Climate and Weather*, pp. 817–823. New York: Oxford University Press.
Gordon HS (1954) The economic theory of a common-property resource: The fishery. *The Journal of Political Economy* 62: 124–142.
Grabowski JH, Clesceri EJ, Gaudette J, Baukus A, Weber M, and Yund PO (2010) Use of herring bait to farm lobsters in the Gulf of Maine. *PLoS One* 5: e10188.
Groffman P, Baron J, Blett T, *et al.* (2006) Ecological thresholds: The key to successful environmental management or an important concept with no practical application? *Ecosystems* 9: 1–13.
Guerry A, Plummer M, Ruckelshaus M, and Harvey C (2011) Ecosystem service assessments for marine conservation. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital: Theory and Practice of Mapping Ecosystem Services*. Oxford: Oxford University Press.
Guerry AD, Ruckelshaus MH, Arkema KK, *et al.* (2012) Modeling benefits from nature: Using ecosystem services to inform marine spatial planning. *International Journal of Biodiversity Science, Ecosystem Services and Management*. doi: 10.1080/21513732.2011.647835.
Holland DS and Brazee R (1996) Marine reserves for fisheries management. *Marine Resource Economics* 11: 157–171.
Holland DS, Sanchirico J, Johnston RJ, and Joglekar DP (2010) *Economic Analysis for Ecosystem Based Management: Applications to Marine and Coastal Environments*. Washington, DC: Resources for the Future.
Hollowed AB, Aydin KY, Essington T, *et al.* (2011) Experience with quantitative ecosystem assessment tools in the northeast Pacific. *Fish and Fisheries* 12: 189–208.
Holmlund CM and Hammer M (1999) Ecosystem services generated by fish populations. *Ecological Economics* 29: 253–268.
Incze L, Xue H, Wolff N, *et al.* (2010) Connectivity of lobster (*Homarus americanus*) populations in the coastal Gulf of Maine: Part II. Coupled biophysical dynamics. *Fisheries Oceanography* 19: 1–20.
Ingraham MW and Foster SG (2008) The value of ecosystem services provided by the U.S. National Wildlife Refuge System in the contiguous. *U.S Ecological Economics* 67: 608–618.
Inter-agency Ocean Policy Task Force (IOPTF) (2010) *Final Recommendations of the Interagency Ocean Policy Task Force*. Washington, DC: Council on Environmental Quality.
Kareiva P, Tallis H, Ricketts TH, Daily GC, and Polasky S (eds.) (2011) *Natural Capital: Theory and Practice of Mapping Ecosystem Services*. Cambridge: Oxford University Press.
Klein C, Ban N, Halpern B, *et al.* (2010) Prioritizing land and sea conservation investments to protect coral reefs. *PLoS One* 5: e12431.
Kliskey A, Alessa L, and Barr B (2009) Integrating local and traditional ecological knowledge. In: McLeod K and Leslie H (eds.) *Ecosystem-Based Management for the Oceans*, pp. 145–161. Washington, DC: Island Press.
Knetsch JL (1963) Outdoor recreation demands and benefits. *Land Economics* 39: 387–396.
Koch EW, Barbier EB, Silliman BR, *et al.* (2009) Non-linearity in ecosystem services: Temporal and spatial variability in coastal protection. *Frontiers in Ecology and the Environment* 7: 29–37.
Lester SE, McLeod K, Tallis H, *et al.* (2010) Science in support of ecosystem-based management for the US West Coast and beyond. *Biological Conservation* 143: 576–587.

Levin PS, Fogarty MJ, Murawski SA, and Fluharty D (2009) Integrated ecosystem assessments: Developing the scientific basis for ecosystem-based management of the ocean. *PLoS Biology* 7: 23–28.

Levin S (1998) Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* 1.

Levin S and Lubchenco J (2008) Resilience, robustness, and marine ecosystem-based management. *BioScience* 58: 1–6.

Liu J, Dietz T, Carpenter SR, *et al.* (2007a) Complexity of coupled human and natural systems. *Science* 317: 1513–1516.

Liu J, Dietz T, Carpenter SR, *et al.* (2007b) Coupled human and natural systems. *Ambio* 36: 639–648.

Mallin MA, Williams KE, Esham EC, and Lowe RP (2000) Effect of human development on bacteriological water quality in coastal watersheds. *Ecological Applications* 10: 1047–1056.

Melillo JM, Mcguire AD, Kicklighter DW, Moore B, Vorosmarty CJ, and Schloss AL (1993) Global climate-change and terrestrial net primary production. *Nature* 363: 234–240.

Millennium Ecosystem Assessment (2005a) *Ecosystems and Human Well-Being: Scenarios,* vol. 2. Washington, DC: Island Press.

Millennium Ecosystem Assessment (2005b) *Ecosystems and human well-being: Synthesis.* Washington, DC: Island Press.

Mitchell RC and Carson RT (1989) *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington, DC: Resources for the Future.

Moberg F and Folke C (1999) Ecological goods and services of coral reef ecosystems. *Ecological Economics* 29: 215–233.

Morey E (1981) The demand for site-specific recreation activities: A characteristics approach. *Journal of Environmental Economics and Management* 8: 345–371.

Murawski SA, Steele J, Taylor P, *et al.* (2009) Why compare marine ecosystems? *ICES Journal of Marine Science* 67: 1–9.

Naeem S, Bunker D, and Hector A (2009) *Biodiversity, Ecosystem Functioning and Human Wellbeing: An Ecological and Economic Perspective.* New York, NY: Oxford University Press.

Naidoo R, Balmford A, Costanza R, *et al.* (2008) Global mapping of ecosystem services and conservation priorities. *Proceeding of the National Academy of Sciences* 105: 9495–9500.

National Research Council (2005) *Valuing Ecosystem Services: Toward Better Environmental Decision-Making.* Washington, DC: National Academies Press.

Nellemann C, Corcoran E, Duarte CM, *et al.* (2009) *Blue Carbon. A Rapid Response Assessment.* GRID-Arendal: United Nations Environment Programme.

Ostrom E (2009) A general framework for analyzing sustainability of social-ecological systems. *Science* 325: 419–422.

Palmquist RB (1991) Hedonic methods. In: Braden JB and Kolstad CD (eds.) *Measuring the Demand for Environmental Quality*, pp. 77–120. Amsterdam: North Holland.

Palmquist RB (2003) Property value models. In: Mäler K-G and Vincent JR (eds.) *Handbook of Environmental Economics.* North Holland: Elsevier.

Palumbi S, Sandifer P, Allan J, *et al.* (2009) Managing for ocean biodiversity to sustain marine ecosystem services. *Frontiers in Ecology and the Environment* 7: 204–211.

Patterson M and Glavovic B (eds.) (2008) *Ecological Economics of the Oceans and Coasts.* Cheltenham: Edward Elgar.

Pergams O and Kareiva P (2009) Support services: A focus on genetic diversity. In: Levin S (ed.) *The Princeton Guide to Ecology*, pp. 642–651. Princeton, NJ: Princeton University Press.

Peterson C and Lubchenco J (1997) Marine ecosystem services. In: Daily G (ed.) *Nature's Services.* Washington, DC: Island Press.

Plummer ML (2009) Assessing benefit transfer for the valuation of ecosystem services. *Frontiers in Ecology and the Environment* 7: 38–45.

Ricketts T, Daily G, Ehrlich PR, and Michener C (2004) Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences of the United States of America* 101: 12579–12582.

Rodríguez JP, Beard JTD, Bennett EM, *et al.* (2006) Trade-offs across space, time, and ecosystem services. *Ecology and Society* 11: 28.

Ronnback P (1999) The ecological basis for economic value of seafood production supported by mangrove ecosystems. *Ecological Economics* 29: 235–252.

Rosenberger R and Loomis J (2001) *Benefit Transfer of Outdoor Recreation Use Values: A Technical Document Supporting the Forest Service Strategic Plan (2000 revision).* Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Ryan RW, Holland DS, and Herrera G (2010) Bioeconomic equilibrium in a Bait-constrained fishery. *Marine Resource Economics* 25: 281–294.

Sainsbury KJ, Punt AE, and Smith A (2000) Design of operational management strategies for achieving fishery ecosystem objectives. *ICES Journal of Marine Science* 57: 731–741.

Samhouri J, Levin P, James CA, Kershner J, and Williams G (2011) Using existing scientific capacity to set targets for ecosystem-based management: A Puget sound case study. *Marine Policy* 35: 508–518.

Sanchirico JN and Wilen JE (1999) Bioeconomics of spatial exploitation in a patchy environment. *Journal of Environmental Economics and Management* 37: 129–150.

Sathirathai S and Barbier EB (2001) Valuing mangrove conservation in southern Thailand. *Contemporary Economic Policy* 19: 109–122.

Schaefer MB (1957) Some considerations of population dynamics and economics in relation to the management of marine fishes. *Journal of the Fisheries Research Board of Canada* 14: 669–681.

Schlesinger W (1991) *Biogeochemistry: An Analysis of Global Change.* Academic Press

Schmitt RJ, Jacobs RS, Page HM, *et al.* (2006) *Advancing marine biotechnology: Use of OCS oil platforms as sustainable sources of marine natural products.* Santa Barbara, CA: Coastal Research Center, Marine Science Institute, University of California.

Scott A (1955) The fishery: The objectives of sole ownership. *Journal of Political Economy* 63: 116–124.

Slesnick D (1998) Empirical approaches to the measurement of welfare. *Journal of Economic Literature* 36: 2108–2165.

Smith A, Fulton E, Hobday A, Smith D, and Shoulder P (2007) Scientific tools to support the practical implementation of ecosystem-based fisheries management. *ICES Journal of Marine Science* 64: 633–639.

Smith MD (2008) Bioeconometrics: Empirical modeling of bioeconomic systems. *Marine Resource Economics* 23: 1–23.

Smith MD, Zhang J, and Coleman FC (2008) Econometric modeling of fisheries with complex life histories: Avoiding biological management failures. *Journal of Environmental Economics and Management* 55: 265–280.

Stokstad E (2011) Appraising UK's ecosystems, Report envisions greener horizons. *Science* 332: 1139.

Tallis H, Kareiva P, Marvier M, and Chang A (2008) An ecosystem services framework to support both practical conservation and economic development. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9457–9464.

Tallis H, Lester SE, Ruckelshaus M, *et al.* (2011) New metrics for managing and sustaining the ocean's bounty. *Marine Policy.*

Tallis H, Levin PS, Ruckelshaus M, *et al.* (2010) The many faces of ecosystem-based management: Making the process work today in real places. *Marine Policy* 34: 340–348.

Toman M (1998) Why not to calculate the value of the world's ecosystem services and natural capital. *Ecological Economics* 25: 57–60.

Travis J (2005) Hurricane Katrina: Scientists' fears come true as hurricane floods New Orleans. *Science* 309: 1656–1659.

Troy A and Wilson M (2006) Mapping ecosystem services: Practical challenges and opportunities in linking GIS and value transfer. *Ecological Economics* 60.

UK National Ecosystem Assessment (2011) *The UK National Ecosystem Assessment: Synthesis of the Key Findings.* Cambridge: UNEP-WCMC.

UN General Assembly (UNGA) ((2010) Sustainable Development: Report of the Governing Council of the United Nations Environment Programme on its Eleventh Special Session.

UNEP and IOC-UNESCO (2009) An Assessment of Assessments, Findings of the Group of Experts. Start Up Phase of a Regular Process for Global Reporting and Assessment of the State of the Marine Environment including Socio-Economic Aspects.

United Nations Environment Program (2006) Marine and Coastal Ecosystems and Human Wellbeing: A Synthesis Report based on the Findings of the Millennium Ecosystem Assessment. United Nations Environment Program.

White C, Halpern BS, and Kappel CV (in press) Ecosystem service tradeoff analysis reveals the value of marine spatial planning for multiple ocean uses. *Proceedings of the National Academy of Sciences of the United States of America.*

Wilen JE, Smith M, Lockwood D, and Botsford L (2002) Avoiding surprises: Incorporating fisherman behavior into management models. *Bulletin of Marine Science* 70: 553–575.

Wilson MA and Liu S (2008) Evaluating the nonmarket value of ecosystem goods and services provided by coastal and nearshore marine system. In: Patterson MG and Glavovic BC (eds.) *The Ecological Economics of the Oceans and Coasts.* Cheltenham: Edward Elgar.

Worm B, Barbier E, Beaumont N, *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.

Modeling Terrestrial Ecosystem Services

Erik Nelson, Bowdoin College, Brunswick, ME, USA
Nirmal Bhagabati, World Wildlife Fund – US, Washington, DC, USA
Driss Ennaanay, Riverside Technology, Inc., Fort Collins, CO, USA
Eric Lonsdorf, Chicago Botanic Garden, Glencoe, IL, USA
Derric Pennington, University of Minnesota, St Paul, MN, USA, and World Wildlife Fund – US, Washington, DC, USA
Manu Sharma, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Cultural ecosystem services Ecosystems' contributions to the nonmaterial benefits (e.g., capabilities and experiences) that arise from human interactions with ecosystems (Chan *et al.*, 2012).

Ecosystem service An ecological process that creates value for one or more people. For example, the ability of a wetland to remove pollutants from surface and ground water moving through a watershed creates value for society by reducing the amount of money that communities have to spend on water remediation technology and activities.

Ecosystem service value The amount of benefit that an ecosystem service generates for people. The value a service generates will depend on the magnitude of its supply and human demand for the service. For example, animal pollinators can increase the yield and quality of crops. The more nesting habitat and pollinator food the landscape provides the greater the supply of pollinator service on the landscape. The value of increased pollinator abundance will depend on how many crops on the landscape are pollinator dependent and the market or nutritional value of the pollinator-dependent crops grown on the landscape.

Net social benefit When a change in the economy or landscape creates additional societal benefits and costs, the net social benefit of the change is given by the additional societal benefits less the additional societal costs. When the societal benefits generated by a change are greater than societal costs generated by the change, net social benefit is positive and vice versa.

Optimal landscape When land use, land cover, and land management are distributed across a landscape such that it produces the highest level or value of some good, service, or sum of goods or services subject to a series of landscape and economic constraints. For example, we could use a carbon storage and sequestration ecological production function to determine the land use and management pattern across the landscape that maximizes carbon sequestration over time subject to maintaining a certain level of urban and agricultural production value.

Provisioning ecosystem service A consumable good whose production is at least partly dependent on resources in the ecosystem. Generally, the quality or level of a provisioning service will depend on the extent or quality of various supporting ecological conditions and processes on the landscape. For example, crop production, a provisioning service, tends to increase in the organic carbon content of soil. In this case, organic carbon content is a service that supports a provisioning service.

Public good A pure public good is a good or service that is nonrivalrous and nonexcludable. For example, clean air is a pure public good. A good or service that is nonrivalrous but excludable, like a national park that charges for entrance, is a club good. A good or service that is rivalrous but nonexcludable, like fish in a lake open to the public, is a common property resource.

Regulating ecosystem service An ecological process that provides people value indirectly by mollifying or preventing more extreme or destabilizing conditions on the landscape. For example, by sequestering carbon on the landscape we can mollify the pace of climate change and its expected economic damages.

Spatially explicit ecological production function A mathematical mechanism that translates spatially explicit information on conditions, features, and processes in an ecosystem to some ecological outcome or condition in the ecosystem. For example, a pollinator abundance production function uses data on patterns of land cover, land use, land management, local climate, and biophysical features on a landscape to estimate the abundance of pollinators across the landscape.

Supporting ecosystem service Any ecological process or feature that aids in the quality, level, or duration of another ecosystem service.

Introduction

Ecosystem services are ecological processes and conditions that create value for one or more people (MA, 2005; Boyd and Banzhaf, 2006). Terrestrial ecosystem services include all services produced on land or in fresh water bodies. Ecosystem services include processes that support the production of consumable goods, processes that support and regulate life (e.g., storm surge protection, climate regulation), and conditions that enhance life (e.g., recreational, esthetic, and spiritual experiences).

Landowners in market economies are financially rewarded for producing goods and services that are sold in markets. Therefore, profit-maximizing landowners are interested in maintaining ecosystem services to the extent that they cost-effectively enhance the production of the goods and services they produce for markets. However, landowners' self-interested land use and management decisions can affect the

production of their neighbors' marketed goods and a wider range of ecosystem services across the landscape. Because landowners are usually not rewarded (or punished) for the impact of their decisions on the production (or destruction) of marketed goods elsewhere on the landscape and most ecosystem services, they are often largely indifferent to these landscape dynamics (Kroeger and Casey, 2007).

Such indifference may not be socially optimal, however. Given that ecosystem services create value for at least one person, society might realize a net benefit from better management of their provision even if the providers of services do not realize a private net benefit from such actions (e.g., Polasky *et al.*, 2011). For example, if adopting less-intensive cropping practices in the upstream reaches of a watershed generates more in ecosystem service value downstream than the loss in net returns on the impacted croplands, then local policy makers can create a net social benefit by paying upstream farmers just enough to make it in their financial interest to adopt these less-intensive cropping practices (Pattanayak *et al.*, 2010). In order to make calculations like this we need methods for accurately modeling and mapping service provision and value across a landscape, whether for current conditions on the landscape or potential conditions in the future.

Ecosystem Service Supply, Demand, and Value

The amount of value that terrestrial ecosystem services provide to people is a function of the magnitude of service supply, the duration of supply, the number of people that benefit from the supply, and how important or valuable the service is to the people that benefit from the supply (Kareiva *et al.*, 2011). The intersection of supply and demand varies greatly across services and landscapes. For example, land use/land cover (LULC) pattern in a river valley can affect the likelihood of flooding in a town downstream. An ecosystem service is created when upstream landowners make decisions on their lands that reduce the risk of flood in the town downriver. The value of this service is determined by how much flood risk reduction is provided and the demand for flood risk reduction in the town. The owners of town land threatened by flooding would be willing to pay money for the service. The more commercially valuable the land that is at risk for flooding, the more the owners of the land value and are willing to pay for flood mitigation services, all else equal (flood insurance would reduce the willingness of landowners to pay for these services). However, how much value the town landowners receive from flood risk mitigation will be limited by the extent of flood-mitigating behavior by landowners above the town. If landowners upstream of the town engage in activities that minimally reduce the risk of flood downstream, then the expected increase in avoided flood damage will be small regardless of the value of the land threatened in town. As the risk of flooding is reduced more and more by actions upstream, the expected avoided damages in the town will increasingly be determined by the commercial value of the threatened land.

For other services, the suppliers and consumers can be the same people. For example, consider a farm planted with crops whose yield increases in animal-pollinator abundance. If the monetary value of the additional production created by attracting additional pollinators to the farm is greater than the cost of establishing pollinator habitat on the farm, then we would expect a profit-maximizing farmer to establish the necessary habitat. The farmer creates a public good and "external" benefits as well if the pollinators attracted to the farmer's field spillover to neighboring farms with pollinator-dependent crops.

Services and their associated values are comparable on other metrics as well. Some services are rivalrous and excludable whereas other services generate benefits that many people can consume and benefit from. For example, consider the pollination service created by establishing additional pollinator habitat on a farm. The value created by the additional pollinators attracted to the farm by the new habitat is captured by the farm's owner and, in some cases, her immediate neighbors. The benefit is rivalrous (by attracting pollinators to her land the pollinators are not available to other farmers except her immediate neighbors) and excludable (only the farmer and her neighbors capture the value created by the additional habitat). Conversely, when water quality is improved in a public lake it creates benefits for all users of the lake. Further, the benefits are not rivalrous – one person's enjoyment of the clean lake does not prevent another person's enjoyment (it is excludable if there is a fee to enter the park that contains the lake). Because many people benefit from the public good of improved water quality in the lake the total value of this service is likely to be much larger than the total value of the pollinator services that accrue to the farmer and her immediate neighbors.

Further, supply and value provided by a service can decrease or even disappear if service supply reverts back to previous levels or consumers of the service demand less of it. For example, if the commercial value of land in the town downriver decreases, then any reduction in flood risk generated upstream becomes less valuable. Or if the farmer that established pollinator habitat eventually produces a crop that is not animal-pollinator dependent, then the pollinator service previously produced on her farm completely disappears.

Therefore, it is essential that any attempt to model services be able to represent these various dimensions of service supply, demand, and value in general and for the landscape in question in particular. Further, the models need to be sensitive to changes on landscape, whether man-made or natural, so the researcher can predict the ramifications of these changes as mediated through ecosystems and the services they provide. It is the authors' opinion that only spatially explicit production and demand functions exhibit these characteristics. Spatially explicit production functions estimate the supply of ecological processes at all points on the landscape with a mathematical relationship that converts patterns of LULC, land management, biophysical features, geological features, climate, and weather across the landscape into a map of an ecological process, states, or condition. A change in any one of these inputs can change the production of ecological processes, states, or conditions across the landscape. The spatially implicit demand functions determine how valuable the supplies of various ecological processes are. Demand functions are explained by the number of people and pattern of people across the landscape and their willingness to pay for these

services. This production function approach is in contrast to a benefits transfer methodology where the service levels and values provided by a natural feature on a landscape (e.g., a wetland) are based on values the natural feature has been estimated to provide on other landscapes (e.g., Troy and Wilson, 2006). The authors feel that transferred values rarely accurately describe values on the landscape in question given its inevitably unique supply and demand dynamics. Further, it is not clear how transferred values can be accurately modified as conditions on the landscape change.

Reporting Ecosystem Service Supply and Value in Biophysical Terms

To introduce the reader to the concept of ecosystem service modeling with production functions the authors will discuss several terrestrial ecosystem service provision and value models in detail below (space limitations do not allow for an in-depth discussion of all services). However, the authors only briefly discuss the conversion of consumed ecosystem service supply into a monetary value. While it is often very illuminating to convert the consumption of supplied services into a money metric, there may be many instances where biophysical measures of value are sufficient for research or policy questions. For example, a researcher or policy maker interested in determining how much more carbon a landscape can sequester if managed in a certain way will not need to go beyond the models described below. Or a policy maker interested in the trade-off between the volume of wood harvested and the amount of sediment retained on the landscape across various forest management scenarios will find the models discussed here sufficient.

Modeling Ecosystem Service Provision and Value

The authors structure their discussion of ecosystem service modeling using the service taxonomy proposed by the Millennium Ecosystem Assessment (MA) (2005). First, the authors discuss provisioning services. Here the authors highlight the provision of agriculture and other consumable goods on the landscape and ways to measure the contribution supporting ecosystem services make to their production. Second, the authors discuss regulating services such as carbon storage and sequestration and storm peak mitigation. Third, the authors touch on the ecosystem processes that support nature-based recreation and the esthetic and spiritual opportunities provided by ecosystems. Finally, the authors focus on the modeling of services that support the provision of the aforementioned consumable goods, regulating services, and cultural experiences.

Because a service can simultaneously contribute to the production of more than one valuable landscape process, service values can be double-counted if a modeler is not careful (Fisher *et al.*, 2009; Boyd and Krupnick, 2009). For example, pollinators support the production of certain crops. Therefore, it would be inappropriate to aggregate the value of agricultural production and the supportive value of pollination on the landscape; the value of pollinators is already part of agricultural production value. However, it is appropriate to value a supporting service multiple times if it contributes to two or more valuable processes on the landscape. For example, retaining more nitrogen fertilizer on the landscape improves water quality. Therefore, retention of nitrogen on the landscape can decrease the costs of meeting drinking water regulations and improve the recreational fishing, swimming, and boating experiences at local lakes. The avoided water remediation costs and improved recreation experiences can be aggregated and attributed to nutrient retention on the landscape. Proper accounting of the marginal contributions supporting services make to provisioning and regulatory service supply and consumption is particularly important when making claims regarding the monetary value of services produced on the landscape.

Provision of Consumable Goods

Ecological processes contribute to the production of goods that we harvest for consumption. For example, the services that support agricultural production include sediment retention, nutrient cycling in the soil, pollination, and water supply for irrigation. Timber and hydroelectricity are other appropriated goods whose production is supported by ecological processes and like agriculture, tend to be highly managed. Other provisioning services, such as bush meat and medicinal plant production, also rely on various supporting services but tend to be much less managed by humans.

Example: Modeling Crop Production

Crops for food, fiber or fuel are produced with a mix of man-made inputs like seed, fertilizers, pesticides, machinery, and labor, and ecological processes like water supply, nutrient cycling in soil, soil retention, natural pest predators, pollinators, and ecological conditions like soil quality and weather (Swinton *et al.*, 2007). Let Y_{ko} indicate the metric tons of marketable crop k produced in field o. Output Y_{ko} is explained by a vector of man-made inputs applied in field o, σ_o, a vector of ecological processes in field o, θ_o, a vector of ecological conditions in field o, ω_o (e.g., weather), and a vector of past land use in field o, φ_o (e.g., maize production flowing bean production benefits from the nitrogen fixed in the soil by the beans).

$$Y_{ko} = f_{kl(o)}(\sigma_o, \theta_o, \omega_o, \varphi_o) \qquad [1]$$

where $f_{kl(o)}$ is the yield function specific to crop k in the landscape l that contains field o, symbolized by $l(o)$. The vector θ_o includes all supporting ecosystem services other than irrigation water which is included in vector σ_o because irrigation water levels in field o are much more controlled than all other agriculture-related supporting services. The function $f_{kl(o)}$ can be estimated statistically using data on the variables found in [1] (e.g., Foley *et al.*, 2011) or can come from crop simulation models like DSSAT, CropWat, EPIC, and APSIM (IBSNAT, 1989; Smith, 1993; Williams, 1995; Keating *et al.*, 2003). The monetary value of Y_{ko} is given by the revenue it generates less production costs or $R_{ko} = p_{kl}Y_{ko} - C_{ko}$ where R_{ko} is the net revenue from producing k on o, p_{kl} is the per ton market price of crop k on landscape l, and C_{ko} is the cost of producing crop k on field o.

As long as field o is used for marketable crop production it will produce a series of output Y. The frequency and magnitude of output over a given time frame are a function of climate and crop choice. For example, in tropical climates a field may be able to produce two crops in a calendar year while a field in more temperate climates can only produce one crop per calendar year. Or a crop may require several years of establishment before a marketable crop is produced. For example, it typically takes 3–4 years after planting for a coffee bush to produce a marketable crop. The net present value (NPV) of the series of output Y on field o is given by the discounted sum of all R_o values.

Most often we assume a farmer will choose crops k and levels of σ_o over time such that they maximize NPV (in some cases, especially in subsistence farming, the objective may be to minimize risk of crop failure). The choices of k and σ_o that maximize the NPV will change as θ_o, ω_o, and φ_o change. See Ghazoul (2007, 2008), Kremen *et al.* (2008), and Nelson *et al.* (2011) for more on farmer decision making given ecosystem service provision on the landscape.

We can use the yield function in eqn [1] and the assumption of a profit-maximizing farmer to determine the marginal productive value of the ecosystem services that support agricultural production on field o. For example, imagine changes on the landscape increase the field o's animal-pollinator abundance (a change in θ_o) and access to water during the growing season. Suppose the provision of all other supporting services and ecological conditions on field o do not change. Given the change in θ_o and water availability, field o's farmer reassesses her choice of k and levels of σ_o. Assume before the change on the landscape the farmer of o always planted k and k produced one crop per calendar year. Now suppose after the change in θ_o and water availability the farmer of o finds it in her best financial interest to choose a more animal-pollinator- and irrigation-dependent crop. Let $\hat{k}$ and $\hat{\sigma}_o$ indicate the farmer's production and managed input choices (including irrigation water applied) given the change in θ_o, represented by $\hat{\theta}_o$, and water availability. Assuming $\hat{k}$ is also only produced once per calendar year on field o, the annual marginal productive value of the increases in pollinator abundance and water availability is given by

$$\hat{V}_o = f_{\hat{k}l(o)}(\hat{\sigma}_o, \hat{\theta}_o, \omega_o, \varphi_o) - f_{kl(o)}(\sigma_o, \theta_o, \omega_o, \varphi_o) \qquad [2]$$

where $\hat{V}_o$ is measured in metric tons. Ignoring crop and food waste (which is substantial; see Parfitt *et al.*, 2010), all of $\hat{V}_o$ is valuable to society, although the consumers who benefit from the additional production of k on field o may reside far from it. The annual net monetary value of the change in pollinator abundance and water availability in field o is equal to

$$\hat{R}_o = \left(p_{\hat{k}l} f_{\hat{k}l(o)}(\hat{\sigma}_o, \hat{\theta}_o, \omega_o, \varphi_o) - C_{\hat{k}o}\right) - \left(p_{kl} f_{kl(o)}(\sigma_o, \theta_o, \omega_o, \varphi_o) - C_{ko}\right) \qquad [3]$$

All else equal, services that support agriculture are more valuable monetarily on the landscapes that generate crops that are more highly valued in the market and that interact more efficiently with managed inputs. An alternative method for placing a value on the change in the supply of services that support agriculture is to determine the impact of the change on calorie or micronutrient production.

This modeling framework can also be used to estimate how much a farmer would need to be paid to adopt management practices that increased ecosystem service supply on the landscape (i.e., payments for ecosystem services). For example, let $\hat{k}$ and $\hat{\sigma}_o$ indicate the crop and managed input choices that the policy maker would like the farmer of field o to adopt for perpetuity. Assume the climate on o is such that only one crop can be produced on the field per year no matter what the choice of k or managed inputs is. Let s be the annual payment given to the farmer for adopting $\hat{k}$ and $\hat{\sigma}_o$. Accepting the payment of s and its production restrictions is only optimal for the farmer of field o if the value of s is such that

$$\sum_{t=0}^{\infty} \frac{s + \left(p_{\hat{k}lt} f_{\hat{k}l(o)}(\hat{\sigma}_o, \theta_{ot}, \omega_{ot}, \varphi_{ot}) - C_{\hat{k}ot}\right)}{(1+r)^t} \geq \sum_{t=0}^{\infty} \frac{p_{k_t^* lt} f_{k_t^* l(o)}(\sigma_{ot}^*, \theta_{ot}^*, \omega_{ot}, \varphi_{ot}) - C_{k_t^* ot}}{(1+r)^t} \qquad [4]$$

where r is the annual discount rate and k_t^*, σ_{ot}^*, and θ_{ot}^* for all years $t = 0,1,2,\ldots$ indicate the nonsubsidized cropland and management choices, respectively, that maximize the discounted NPV of returns to farming on field o. In words, the value of s needs to be large enough so choosing $\hat{k}$ and $\hat{\sigma}_o$ for perpetuity generates as much expected NPV as not taking s and optimizing the NPV of net revenues on field o.

Other Provisioning Services

The supply of several other provisioning services is heavily dependent on the property rights structure in the ecosystem or resource that provides the good. For example, if a forest is privately held and used primarily for timber production, it is reasonable for a modeler to expect that the owner of the forest manages and extracts timber from the forest in such a way that the NPV of timber harvests over time is maximized (see Nelson *et al.* (2011) for an example of such a model). In this case, NPV is maximized by always keeping a portion of the resource *in situ* for future harvest. This allows the owner to take advantage of growing timber and its associated growth in value. Ecosystem services that support timber production include soil nutrient cycling, pest control, and fire regime management. Fresh water fisheries with private property rights are described by the same harvest dynamics as a private forest; simply substitute the harvest of wood with the harvest of fish. Improved water quality and sediment retention on the landscape support the service of fish production in fresh water bodies.

When resources are privately controlled the entire resource stock at any given time, that which is harvested plus that left *in situ*, provides marketable value to the owner of the resource. Therefore, the ecosystem services that support the growth and health of privately controlled resources tend to be maintained over time. The marginal productive value and NPV of changes in services that support timber and fish production can be determined in a manner similar to that of the services that support agricultural production (see eqn [2]).

If an owner of a resource cannot prevent poaching of its stock or property rights are not well defined or do not exist at

all, then the resource is described by open access (Bromley, 1991). Open access systems are defined by intense harvesting effort because there is no incentive among harvesters to forego harvest of some stock in order to let it and its value grow; someone else will take it in the meantime. See Lopez-Feldman and Wilen (2008), Robinson *et al.* (2008), and Nelson *et al.* (2011) for examples of open access harvest models. Eventually harvest effort for the open access resource will fall as low stock levels dramatically increase the harvester's opportunity cost of resource collection. Much nontimber-forest product (NTFP) harvest in developing countries can be modeled as an open access system (Ticktin and Shackleton, 2011). Again we assume the entire harvest of products from open access systems provide value to harvesters and therefore, society. Unlike the closed access systems, however, *in situ* stock does not provide marketable value to anyone (however, *in situ* stock can provide other service values such as carbon storage). Further, the consumptive value that an open access resource provides over time diminishes quickly as the resource's stock falls. Therefore, the value of services that aid stock production in open access systems will decrease over time. One method for preventing the collapse in resource stocks and value under open access is to establish property rights over the system. This does not necessarily mean privatizing the resource. Ostrom (1999) gives examples of resources under communal property right structures that have maintained sustainable harvests over time.

Another consumptive good whose supply is aided by ecological processes is hydroelectric power. Energy is created when water yielded in a catchment (precipitation less evapotranspiration) is routed through a turbine. The greater the water yield upstream of the turbine and the less that this yield is diverted for other uses, the greater the potential for this water to drive the turbine. Hydrological production at a dam is even more valuable if the reservoir behind the dam does not have to be dredged often to remove deposited sediment. Therefore, the more sediment the landscape can capture upstream of a dam the more net revenue the hydroelectric turbine at the dam provides. Electricity provided when a flowing river or stream passes through a turbine (run of the river systems) is much more variable than electricity provided by dams. Dam managers have the ability to set their preferred production schedule by opening and closing the dam. The better models of hydroelectric power production at dams include this element of manager decision making. Conversely, electricity production at run of the river systems will fall as stream volume falls either due to weather or water use upstream; electricity production at these turbines is relatively management free. However, unlike dam models, which can use seasonal or annual water volume data, run of the river system models requires seasonal or even finer temporal data to simulate undulating flows in the river or stream of interest. See Gibbons (1986), Edwards (2003), and Mendoza *et al.* (2011) for more on hydroelectric power models. The impact of a change in the provision of water yield or sediment retention on hydroelectric power generation and its net revenues provides some information on the value of these supporting services.

Clean fresh water for drinking and other direct human consumption is another vital good whose supply is affected by ecosystems. The supply of fresh water in a region is largely determined by a climate-driven water balance. However, what portion of this supply can be safely consumed by humans without extensive remediation will largely be determined by landscape processes. As water flows across a landscape and through its substrate towards water bodies it will bring pollutants and sediment with it. The more such pollutants are deposited into the landscape's sources of fresh water the more landscape's regulators have to pay to clean the water or in the absence of effective regulation, the greater the likelihood that people will become sick from using the water for drinking, washing, and bathing. Therefore, the more pollutants that the landscape can trap or retain the more value created, either in the form of reduced water purification costs or reduced health damages. The value of the pollutant and sediment retention services is equal to the value of any purification costs or health damages avoided. For example, if pollutant retention on a potential LULC pattern prevents one death and 10 illnesses per year compared to a baseline LULC pattern on the landscape, then the value of additional pollutant retention on the alternative landscape is one death and 10 illnesses. If additional pollutant retention on an alternative landscape does not prevent any deaths or illnesses and does not reduce any regulatory costs, then from a clean fresh water perspective, this supporting service of nutrient retention is valueless.

Regulatory Services

Regulatory services are the processes that provide societal value indirectly by mollifying or preventing more extreme or destabilizing conditions on the landscape. For example, flood control reduces the expected damages after a weather event that introduces an excessive quantity of water on the landscape quickly. Below we highlight several regulatory services and approaches to modeling them. Besides the highlighted services, ecosystems can also regulate the concentration of air pollutants such as nitrous oxides, ground-level ozone, and volatile organic chemicals (Pataki *et al.*, 2011) and help control certain diseases such as malaria (Foley *et al.*, 2005).

Example: Carbon Storage and Sequestration

Over time biomass and soils can capture and store carbon that would otherwise bind with oxygen elements and reside in the atmosphere or ocean as the greenhouse gas (GHG) carbon dioxide (CO_2). By increasing the rate of carbon taken-up by plants and soil and reducing the disturbance of pools of biomass and soil carbon that already exist, even if just temporarily, the process of human-induced climate change and increased ocean acidification can be slowed (Lal, 2004). And because damage to the globe's economies and the environment is expected to increase in climate change, increased carbon sequestration and storage above some baseline expectation of sequestration and storage (additional sequestration and storage) should delay or permanently reduce the path of expected damages. Even just delaying economic damages is valuable because it gives societies more time to adapt and prepare for climate change impacts.

A model of carbon storage and sequestration accounts for the transfer of carbon from one pool to another, including the atmospheric pool, over time. The amount of carbon sequestered in a land parcel over some time frame is the difference in the parcel's storage at the beginning and end of the time frame.

Assume for now that carbon storage levels do not change in parcel x until there is a LULC change on x. Let C_{ixjkt} indicate the amount of carbon stored in pool i on parcel x that is in LULC k but transitioned from LULC j t years ago. The amount of carbon sequestered on parcel x since the transition is given by

$$\Delta C_{xjkt} = C_{axjkt} + C_{bxjkt} + C_{sxjkt} + C_{hxjkt} - C_{axj} - C_{bxj} - C_{sxj} - C_{hxj} \quad [5]$$

where $i = a$, b, s and h indicate the aboveground biomass, belowground biomass, soil, and other carbon storage pools, respectively, and C_{ixj} is the carbon stored in pool i on parcel x in LULC j. In this setting, if LULC does not change on x then after t years sequestration in x is equal to 0. This modeling limitation can be relaxed by making the C variables in eqn [5] a function of land management, climate, and larger landscape dynamics as well. In this case, a change in LULC, land management, climate, and larger landscape dynamics in and around x would engender sequestration. $C_{ixjkt} > C_{ixj}$ if the change in LULC causes pool i to gain more in carbon from other pools (including the atmosphere pool) than it loses to other pools and vice versa. The magnitude of C_{ixjkt} can be determined with a function that is explained by x's biophysical features, LULC j and its features, LULC k and its features, j to k transition dynamics, time, and the climate in x (e.g., Metherell *et al.*, 1993).

A simpler version of eqn [5] assumes that each carbon pool in a parcel that transitions to LULC k will eventually reach a steady-state level of storage where steady-state levels and time to reach steady state are not dependent on any features specific to x or the previous LULC. Let $\hat{C}_{ik}$ be carbon pool i's steady-state level in LULC k and T_{ik} be the years it takes pool i to reach its steady state after a transition to LULC k (note that $\hat{C}_{ik}$ and T_{ik} are not a function of x or LULC j). In this case, eqn [5] becomes

$$\Delta C_{xjkt} = t\left[\frac{\hat{C}_{ak} - C_{axj}}{T_{ak}} + \frac{\hat{C}_{bk} - C_{bxj}}{T_{bk}} + \frac{\hat{C}_{sk} - C_{sxj}}{T_{sk}} + \frac{\hat{C}_{ok} - C_{oxj}}{T_{ok}}\right] \quad [6]$$

If $t > T_{ak}$ then $t[\hat{C}_{ij} - C_{ixk}]/T_{ak}$ is equal to $\hat{C}_{ik} - C_{ixj}$. In eqn [6] the sequestration rate in each pool is constant. This is a modeling simplification as field evidence suggests that terrestrial carbon sequestration follows a nonlinear path over time (e.g., Schlesinger, 1990). Finally, if the modeler lacks information on the level of pool storage in parcel x just before the transition from LULC j to k, then she could assume steady-state storage just before the transition as well. In this case, each storage parameter C_{ixj} in eqn [6] is set equal to its steady-state level, $\hat{C}_{ij}$.

A landscape scenario would indicate the LULC and management transitions and their timing for each parcel on the landscape. For example, suppose a scenario indicated that parcel x would transition from LULC j to LULC k 5 years after the beginning of the modeling time frame and then from LULC k to LULC q 19 years after the beginning of the modeling time frame. Suppose the scenario time frame is 40 years. According to the methodology above, during the first 5 years of the scenario there would be no change in carbon storage on the parcel, the change in storage on x from year 5 to year 19 (where $t = 14$) would be given by eqn [5] (or [6]) where C_{ixj} (or $\hat{C}_{ij}$) is the initial carbon stored in pool i on x given LULC j, and the change in storage on the parcel from year 19 to year 40 (where $t = 21$) would be given by eqn [5] (or [6]) where C_{ixj} (or $\hat{C}_{ik}$) is equal to $C_{ixjk,14}$ (or $14[\hat{C}_{ik} - C_{ixj}]/T_{ik}$ or $14[\hat{C}_{ik} - \hat{C}_{ij}]/T_{ik}$). Over the entire scenario time frame sequestration in parcel x is given by $\Delta C_{xjk,14} + \Delta C_{xkq,21}$.

The portion of sequestration in parcel x that provides value to society is given by the sequestration that is additional compared to some baseline sequestration rate. For example, assume the changes in x described above are motivated by a landscape-wide policy of paying landowners for additional carbon sequestration. Further, suppose in the absence of this policy the expected path of LULC transition on x is LULC j to LULC k 10 years after the beginning of the modeling time frame and no changes thereafter. In the baseline case, sequestration on x over the 40-year modeling time frame is equal to $\Delta C_{xjk,30}$. The term $\Delta C_{xjk,14} + \Delta C_{xkq,21} - \Delta C_{xjk,30}$ gives the additional sequestration on x due to the policy. Because every unit of additional sequestration or storage, even if just temporary, no matter where it occurs, delays expected damages across the globe, all additional sequestration is valuable. If the additional tonnage term is negative then the policy has increased economic damage vis-à-vis the baseline.

The proposed approach to carbon storage and sequestration described above ignores several fundamental carbon storage dynamics present in all managed and unmanaged ecosystems. First, the amount of carbon stored in a parcel is in continual flux due to vegetation growth and decay, organic matter accumulation in the soil, and climatic variability. Several models are well suited to track this continual change. The CENTURY Soil Organic Matter Model (Metherell *et al.*, 1993) continually tracks carbon dynamics in soil and, despite its name, in biomass as well. Dynamic vegetation models simulate vegetation patterns, and associated carbon dynamics, on the landscape as a function of land use and climate (e.g., Bachelet *et al.*, 2001; Sitch *et al.*, 2003). These dynamic vegetation models are particularly useful for predicting the effects of climate change on the terrestrial carbon cycle. A second carbon storage dynamics ignored in the model described above is the long-term storage of carbon in biomass removed from the landscape. For example, wood waste and products made from harvested wood can store captured carbon for up to 100 years (Smith *et al.*, 2006). Such harvested wood product or waste pools could be added to a carbon storage and sequestration model (Conte *et al.*, 2011). Finally, the model proposed above does not account for the annual flux in the GHGs of nitrous oxide (N_2O) and methane (CH_4) from managed land. N_2O is a byproduct of fertilizer use on agricultural fields and traps heat 298 times more effectively than CO_2. CH_4, a byproduct of manure, livestock flatulence, and rice paddy and wetland dynamics, traps heat about 20 times more effectively than CO_2. See Robertson and Grace (2004) for more on the emissions of carbon, N_2O, CH_4, and other GHGs from agricultural fields.

Example: Storm Peak Management

LULC and land management practices in a watershed affect the dynamics of river or stream height downstream after a small to midsize storm event. (A graph that gives the timing of river heights downstream after a storm even is known as a

hydrograph.) For example, LULC and land management changes upstream of a town can reduce the expected peak height of a storm surge in the town. If this reduction in the surge's peak height means less damage in the town after a storm event than it would have been otherwise, then the landscape has provided flood mitigation service. Storm peak and flooding are among the most challenging hydrologic and hydraulic processes to model at the watershed scale. Multiple factors affect how a watershed responds to a specific storm event. A storm peak mitigation model proposed by Ennaanay *et al.* (2011) is composed of three modules. In the first module, Ennaanay *et al.* (2011) use geographic information systems (GIS) software to generate a hydrograph for a downslope point (e.g., a town on a river or stream) given a specific small to medium storm event with uniform depth and duration across the landscape (LULC and land management can do very little to mitigate downriver flooding after very large storm events). The initial module first determines how much of the precipitation runs off each parcel (as opposed to being retained by the parcel's soil and vegetation) upslope of the point of interest. A parcel's excess precipitation or storm runoff is then routed downslope according to an elevation map where some runoff can be retained in wetlands and other depressions with available storage. Any remaining runoff that drains into the watershed's river and stream system will eventually reach the downslope point of interest. The time needed for a parcel's runoff to reach the destination point is found by summing the amount of time the runoff needs to traverse each parcel en route to its drainage point and the subsequent time of travel in the river or stream system to the point of interest. Travel time across the landscape is a function of its roughness and slope where roughness is a LULC and land management characteristic. The first module ends by creating a hydrograph for the point of interest and a map that indicates how long it takes each parcel's runoff to reach the point of interest. In general, the more dissimilar runoff delivery timing is across the parcels upslope of the point of interest, the lower the storm peak is at the point of interest.

The second module of the model proposed by Ennaanay *et al.* (2011) determines how the peak volume will spread beyond the point of interest's channel segment, if at all, after the storm event. For example, if the storm peak of the hydrograph indicates that 100 m^3 of water is expected to overflow a stream's banks at the town of interest, then the module delineates what area and to what depth the excess water will flood. The third module uses this flooding information to calculate flooding damage associated with the storm peak.

Similar to the carbon sequestration and storage model, the supply of storm peak mitigation (the reduction of the hydrograph at its highest point) is determined against a baseline case. First, a baseline landscape is used to create the hydrograph for a particular storm event at the point of interest. Then an alternative landscape is evaluated with the same storm event and the same point of interest. The difference in expected flood damage at the point of interest given the alternative landscape represents how much avoided flood damage the alternative LULC and land management pattern provides. For example, suppose the alternative landscape reduces the height of the storm peak such that 50 m^3 less water will flood the point of interest. Therefore, the alternative landscape has generated 50 m^3 of avoided flood water at the point of interest. The ultimate value of this storm peak mitigation will depend on the monetary value of the land that will be affected by flooding and the reduction in damages, if any, from marginally less flood water volume.

Cultural Ecosystem Services

Ecosystems also provide "nature," a landscape attribute that enhances outdoor recreation, esthetic experiences, and, for many people, provides a larger meaning and purpose to life. While models for the supply and use of nature-based recreation services and some esthetic services are well developed, the ability to model the spiritual values a landscape provides is limited.

Example: Nature-Based Recreation

The first step in nature-based recreation modeling is to collect data on the number of visits a representative population makes to the various nature-based recreation sites on the landscape. If possible, the visit data should be complemented with the representative population's demographics and their costs of traveling to and remaining in the site for a period of time (both out-of-pocket and opportunity costs). Second, the researcher collects information on the amenities found at the visited sites, both natural and man-made. With this dataset the researcher can build a statistical model that predicts the change in the number of visits to the various sites as a function of changes in natural amenities at one or more of the sites (Phaneuf and Smith, 2005). Generally, as the condition of the natural amenities at a site improves (more water, cleaner water, cleaner air, etc.), all else equal, more visits are made to the site and more consumer satisfaction is generated. The structure of the statistical model allows the researcher to isolate visitor response to and the consumer value generated by an improvement in the natural amenity. This value can then be distributed among the ecosystem services that support the natural amenity's improvement.

These recreational models tend to find that the demand for nature-based recreation services increases as people on the landscape become wealthier and can afford more and longer trips. Therefore, nature-based recreation visits have increased around the world as it has gotten wealthier (Pergams and Zaradic, 2006, 2008; Balmford *et al.*, 2009). Further, sites that are easily accessible to many people tend to generate more visits than other sites, all else equal (Bateman *et al.*, 1996). Therefore, services that enhance nature-based recreation at sites that are easily accessible and are attractive to a wealthier cliental are particularly valuable, at least from a monetary perspective, all else equal (e.g., Egan *et al.*, 2009; Vesterinen *et al.*, 2010).

Example: Esthetic Services

Natural vistas provide pleasure to people. Viewing natural vistas may be a primary motivation for a nature-based recreation and the quality of the view at recreation sites should be used to help explain visits to the site (*see* Example:

Nature-Based Recreation). In addition, people may be willing to pay more for homes that have generous viewsheds dominated by natural features such as trees and water bodies (Cavailhès *et al.*, 2009). With the appropriate land and building elevation and land cover maps, GIS software can estimate the land features that compose a home's viewshed. The value a viewshed adds to a home can be estimated with statistical techniques (Hanley and Barbier, 2009). For example, US homes with large natural viewsheds have been estimated to be worth thousands more dollars than comparable homes without such views (Sander and Polasky, 2009). The additional home value due to more extensive natural vistas can then be distributed among the ecosystem services that support the quality of this view. See Tallis *et al.* (2011) and Johnson *et al.* (2010) for examples of viewshed models.

Other Cultural and Spiritual Services

Other cultural and spiritual services include the "nonmaterial benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection…, including, e.g., knowledge systems, social relations…" (MA, 2005, p. 894). Such cultural and spiritual experiences do not readily lend themselves to measurement. However, methods for mapping or ranking cultural and spiritual sites on a landscape have been proposed (Chan *et al.*, 2012). For example, the nutritional and economic value of subsistence animal and plant harvesting in indigenous groups (a cultural experience) can be estimated to obtain a lower bound on the value provided by the ecosystem services that support the growth of harvested natural resources (Chan *et al.*, 2012).

Supporting Services

As exemplified above, most provisioning, regulating, cultural, and spiritual services rely heavily on more fundamental supporting services. For example, in the agricultural model described above we highlighted two services that support agricultural production, pollination, and water supply. Hydropower production is supported by the ecological processes of water yield and sediment retention. The soil's ability to sequester and store carbon increases as animal diversity and abundance in the soil increases (Six *et al.*, 2006). Other ecological processes that support valuable services are nutrient cycling in soils (Giller *et al.*, 1997) and in-stream processes such as water temperature and turbidity. The authors describe the modeling of several supporting services in detail below.

Example: Animal Pollination

Pollinating animals can increase the yield, quality, and stability of the production of many crops, including almond, cacao, canola, coffee, sunflower, tomato, and watermelon. Klein *et al.* (2007) estimate 68% of globally important crops benefit from animal pollination. Gallai *et al.* (2011) find pollinators add approximately $200 billion each year to the value of agricultural production across the world. For pollinators foraging on a landscape, crops may just be one of several sources of food or floral resources. Further, most animal-pollinated crops are better adapted to interact with certain types of pollinators than others. Therefore, how much an animal-pollinated crop will benefit from pollinators on the landscape will be a function of the abundance and spatial location of other floral resources on the landscape, the distance between the field and pollinator nesting sites (pollinators can only fly so far when foraging; Westrich, 1996; Williams and Kremen, 2007), and pollinator-crop compatibility.

Vital to the crop pollinating process is the establishment of pollinator nests on the landscape. Nests are only established on certain types of habitat. Further, pollinators are more likely to establish nests in areas that have close-by floral resources at times they are not hibernating. A model of crop pollination, therefore, can be explained by the relationship between land cover and nest habitat, the location of farms of interest in relation to potential nesting sites, the crops grown on the farms, the temporal and spatial pattern of all floral resources on the landscape, the ability of pollinators to disperse over the landscape, and the dose-response effect of each crop–pollinator combination.

Many of these features are in a pollinator service model proposed by Lonsdorf *et al.* (2009). In the model, the abundance of pollinators at each location in a landscape is determined by the location's suitability for nesting and its accessible floral resources (Kremen *et al.*, 2007). In this model, location x's accessible floral resources include those immediately at x and those on the surrounding landscape, where those resources not at x are discounted by a distance decay function. The model's output is a map of expected pollinator abundance with higher density in locations with better nesting habitat, greater accessibility to floral resources, and that host pollinators that are willing and able to travel longer distances to forage.

The foraging habitat accessible to pollinators nesting at x is given by HF_x,

$$HF_x = \sum_{m=1}^{M} F_m e^{(-D_{mx}/\alpha)} \Big/ \sum_{m=1}^{M} e^{(-D_{mk}/\alpha)} \qquad [7]$$

where F_m represents the relative amount of floral resource available at location m (the series of all locations on the landscape is indexed by $m = 1, \ldots, M$ where location x is part of the series), D_{mx} is the Euclidean distance between location m and x, and α represents the distance a pollinator will travel to forage (Greenleaf *et al.*, 2007; Gathmann and Tscharntke, 2002). The abundance/quality of floral resources at site m, F_m, ranges between 0 and 1, where 1 is the highest possible quality; values are assigned according to m's LULC. In eqn [7] the foraging frequency in location m by pollinators nesting in x declines exponentially with distance between x and m (Cresswell *et al.*, 2000). The model assumes that pollinators forage in all directions from x with equal probability.

Relative pollinator abundance at location x is limited by x's nesting and accessible floral resources; if either one is completely absent at x that pollinator abundance at x, given by P_x, is equal to 0

$$P_x = HF_x HN_x \qquad [8]$$

where HN_x gives nesting habitat suitability at x on a 0–1 scale. In this model HN_x is completely determined by x's LULC. Nesting suitability is assigned to each LULC according to existing data or expert opinion. We can think of P_x as the

pattern of pollinating service supply on the landscape. How useful the supply of pollinators across the landscape will be to humans depends on what types of crops are grown on the landscape and where these crops are grown.

Crop fields surrounded by locations with higher P values should experience higher levels of pollinator visits than crop fields surrounded by locations with lower P values, all else equal. Let $o = 1,\ldots,O$ index crop field locations on the landscape. The fraction of pollinator visits to farm o from location x is given by

$$P_{ox} = P_x e^{(-D_{ox}/\alpha)} \Big/ \sum_{m=1}^{M} e^{(-D_{om}/\alpha)} \quad [9]$$

where D_{ox} (or D_{om}) is the distance between nesting location x (or m) and field o. The numerator of eqn [9] represents the distance-weighted proportion of the pollinators from x that forage on o and the denominator is a scalar that normalizes this contribution by the total area within foraging distance to farm (Winfree *et al.*, 2005). The total pollinator abundance on field o, P_o, is simply the sum over all nesting locations on the landscape, including location x,

$$P_o = \sum_{m=1}^{M} P_{om} \quad [10]$$

P_o from eqn [10] can be part of the θ_o vector of supporting agricultural inputs from eqn [1] where P_o will have no effect on yield if the crops grown on o are not animal pollinated.

For modelers not interested in using the agricultural yield model already introduced above, Lonsdorf *et al.* (2009) propose a simple saturating function to translate the relative abundance of pollinators on farms into an expected yield. Expected crop k yield on field o is given by Y_{ko},

$$Y_{ok} = \overline{Y}_{ko}[(1 - v_k) + v_k(P_o/P_o + \lambda_k)] \quad [11]$$

where $\overline{Y}_{ko}$ is the expected yield of crop k on field o assuming pollination, either animal or wind-aided, is not a limiting factor in crop growth, and v_k represents the proportion of total crop k's yield attributable to wild pollination (i.e., v_k is equal to 1 if k is an obligately outcrossing crop and equal to 0 if k is wind pollinated). The yield $\overline{Y}_{ko}$ can be set equal to the observed yields of k in o's landscape in the cases where researchers know that pollination vectors, either animals or wind, are abundant. The term λ_k in eqn [11] is a half-saturation constant and represents the abundance of pollinators required to reach 50% of pollinator-dependent yield. According to eqn [11] yield for pollinator-dependent crop k ($v_k > 0$) increases as pollinator abundance gets larger (Greenleaf and Kremen, 2006). Equation [11] also indicates that strongly animal-pollinator-dependent crops (k where $v_k \gg 0$) require significant pollinator abundance to yield near their maximum capacity (Allsopp *et al.*, 2008).

The fraction of additional P_o that is valuable to humans is equal to the partial derivative of Y_{ko} with respect to a small change in P_o,

$$\overline{Y}_{ko} v_k \lambda_k / (P_o + \lambda_k)^2 \quad [12]$$

where P_o in (12) is equal to the level of P_o before its change. In words, the positive productive value of additional pollinator abundance in o falls as the pollinator dependence of the crop grown in o, v_k, falls. The net monetary value of a small change in P_o is equal to the marginal value given by eqn [12] times the market price of k (assuming the cost of production does not change given a small change in P_o). The net monetary value of an additional unit of P_o will be higher, all else equal, for more profitable crops. Alternative measures of value, for example, the incremental nutritional value of food produced, will be higher, all else equal, for increases in P_o on fields with more nutritional crops.

The Lonsdorf *et al.* (2009) model has several shortcomings. First, the model does not distinguish among pollinators; all species are represented by a composite pollinator. Second, while the modeler can use F_m to define floral resource preference in the representative pollinator, the pollinators uniformly distribute from the nesting site instead of concentrating at locations on the landscape with preferred resources. See Lonsdorf *et al.* (2011) for a model that addresses some of these limitations.

Example: Nutrient Retention

Pollution produced on a landscape can be transported by surface water runoff into the landscape's water bodies. Such pollution affects the landscape's supply of clean fresh water for human consumption, the abundance of fish and other water life we harvest for food, and the quality of human recreational experiences. Water quality models simulate the transport of pollutants across the landscape to a water outlet. These models typically indicate how land use and management practices at the pixel or hydrologic response unit (HRU; subunits within a watershed) affect water, pollution, and sediment yields at the mouth of a watershed. The most complex water quality models produce time-continuous output of pollutant loadings and require daily weather data along with input data on vegetation, soil properties, and physiography. The Soil and Water Assessment Tool (SWAT) is an example of such a model (Gassman *et al.*, 2007). This tool is remarkable for its ability to model processes such as sediment, nutrient, and pesticide loadings into the stream channel at daily time steps. However, such fine-grain output mandates very detailed sets of data that can be difficult and time consuming to collect. Further, substantial postprocessing time is needed to convert SWAT output to ecosystem service provision and use (Vigerstol and Aukema, 2011). Variable Infiltration Capacity and SPARROW are other hydrological tools that can be used to inform water ecosystem service research and planning.

In a water quality model proposed in Conte *et al.* (2011) nitrogen (N) and phosphorus (P) loadings at a watershed's outlet are calculated on an annual basis instead of a daily or monthly basis. Unlike SWAT and other more complex process models, this rather simple model assigns export coefficients for N and P ($\mathrm{kg\,ha^{-1}\,yr^{-1}}$) and retention efficiencies to each parcel on the landscape according to its LULC type instead of being determined by soil and physiography characteristics specific to the parcel. In the Conte *et al.* (2011) approach, a LULC's annual export values are based on typical or average nutrient transformation processes that take place in the LULC's soil and vegetation. The export coefficient on each

parcel in the watershed is then combined with each parcel's expected annual water yield to determine annual potential nutrient yield from each parcel. Parcel-level nutrient yields or loads are then routed through downslope parcels to the watershed's mouth. As the water and its nutrient load travels across the landscape some of the load will be captured and retained by intermediate parcels. There is no in-stream cycling of nutrients in the model, so the removal and storage of nutrients is limited to the landscape. The model reports total annual loading of each nutrient at the watershed's mouth.

The supporting service modeled by Conte *et al.* (2011) and derivable from SWAT is nutrient retention. The more pollutant retained on the landscape, the cleaner the water, and the more valuable the provisioning and regulating services the landscape's water supports. Further, it reduces the need for costly remediation in water systems and wells in order to meet some regulatory pollution standard, or if standards are not applicable, reduces the human mortality and morbidity associated with excessive nutrient pollution. The supply of nutrient retention on an alternative LULC pattern is defined against a baseline pattern (e.g., current or business as usual conditions). The additional retention created on the alternative landscape represents the alternative scenario's supply of this supporting service. The portion of this additional retention that increases the value of provisioning or regulating services and damages avoided represents the usefulness or value of the additional nutrient retention.

Accounting for Uncertainty in Ecosystem Service Provision and Value

The versions of the models presented above only explain expected results. However, ecosystem service provision and value outcomes are highly uncertain for two reasons. First, provision and value vary because of the inherent randomness of nature, human behavior, and technological progress (Rotmans and van Asselt, 2001). Modelers can attempt to account for this type of uncertainty by systematically varying model input and parameter values (e.g., experimenting with biomass growth rates in a carbon sequestration model) and assumptions regarding human behavior (e.g., farmers are not profit maximizers and instead make decisions according to a goal of minimizing risk of crop failure) and reporting the potential range and likelihood of outcomes. Inexactness, ignorance, and indeterminacy add another degree of uncertainty to modeled results (Rotmans and van Asselt, 2001). For example, several recent studies demonstrate that soil organic carbon levels in an area are determined by a complex mix of ecological variables that are hard to understand let alone model (e.g., Dlugofl *et al.*, 2010; Syswerda *et al.*, 2011; Rumpel and Kogel-Knabner, 2010; Powlson *et al.*, 2011). Our ignorance of ecological processes is particularly problematic when trying to understand and model threshold effects and nonlinearities in ecosystem processes. Evaluating service provision and value on a landscape with disparate modeling approaches is one way to begin confronting indeterminacy issues in ecosystem service modeling. If all or most models predict the same direction and magnitude of change in provision and value, then greater certainty can be attached to the predicted outcomes.

Cost-effective Ecosystem Service Provision and Value

Ecosystem service models are often used to estimate the impact of projected LULC, land management, and climate change on service supply and value. However, we can also use ecosystem service models and conventional market return data to describe the optimal use of a landscape when the objective is to maximize the provision of one or more services subject to some economic constraint. To illustrate this point we consider the market returns to changing LULC and land management in the Seven Mile Creek, Minnesota, USA, watershed in order to increase the rate of P retention and carbon sequestration (the 9531 ha Seven Mile Creek watershed is in the Minnesota River Basin). In this analysis every agricultural parcel in a HRU, typically 2.5 ha, can stay in their current uses (the baseline) or all can convert to switchgrass production, prairie plant production, or timber and pulp production or remain in conventional row crop or pasture production but with lower P application rates. Switchgrass and prairie plants are marketable because they could be used as biomass for electricity generation or the feedstock of liquid fuel (Daniel *et al.*, 2003). SWAT is used to model the effect of LULC and land management pattern across the watershed on annual loading of P (kg ha^{-1}) into the Minnesota River. A version of the carbon storage and sequestration model described above is used to estimate carbon sequestration impacts from LULC change across the watershed. Finally, the annual net return to marketable good production (conventional crops or pasture, switchgrass, prairie grass, and timber and pulp) across the landscape is calculated using crop and other plant yields as predicted by SWAT in conjunction with data on the observed prices and costs of producing conventional crops and timber and the expected prices and costs associated with switchgrass and prairie crop production (Pennington *et al.*, in preparation).

In **Figure 1** we indicate how much more additional P can be retained each year by the landscape compared to the current LULC and land management pattern where each point on a frontier represents a different LULC and land management pattern. There are multiple solutions (LULC and land management patterns) because we vary the minimum allowable change in annual net return to marketed goods and gain in carbon storage produced across the watershed. For example, point A represents the LULC pattern that maximizes P retention *given* the annual net return to marketed goods produced across the landscape has to increase by at least $50,000 compared to the baseline (the current landscape) and steady-state carbon storage levels has to increase by at least 1000 Mg compared to the baseline (at point A changes on the landscape are expected to increase annual P retention by 524 kg). Point B represents the LULC and land management pattern that maximizes change in P retention *given* the annual net return to marketed goods produced across the landscape cannot fall more than $3.5 million per year and steady-state carbon storage levels has to increase by at least 25,000 Mg (at this point changes on the landscape are expected to increase annual P retention by 2074 kg). The next step in this analysis would be to determine how valuable the extent of modeled P retention and increased carbon storage is to people on the landscape

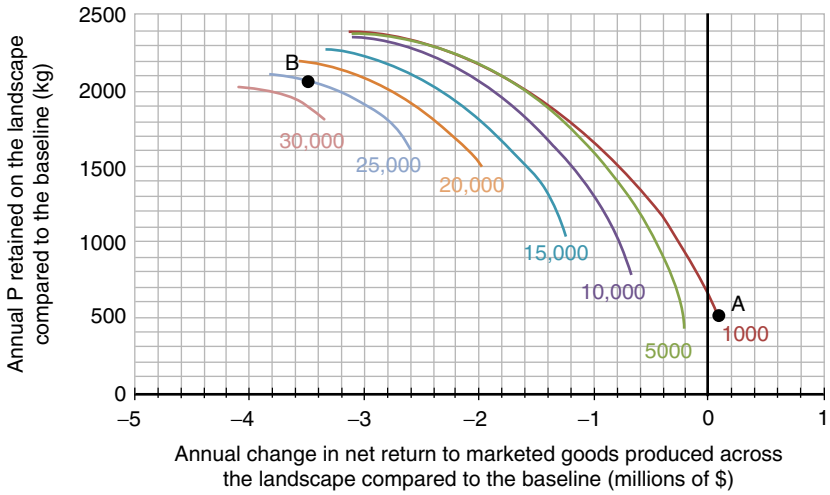


Figure 1 Each point along a frontier gives the maximum gain in annual P retention possible in the Seven Mile Creek watershed compared to baseline retention when gains are constrained by a minimum marketed good value and carbon sequestration production in the watershed. Annual retention of P is increased by changing the LULC and land management pattern in the watershed. Increases in annual retention are constrained by limits on how much of the watershed's annual net return to marketed goods can change (*x*-axis) and minimal necessary gains in steady-state carbon storage on the landscape. Each frontier is associated with a different level of minimal gain in steady-state carbon storage on the landscape. The required gain in metric tons of carbon is indicated by the number at the terminus of a frontier. Baseline output of P retention, carbon storage, and annual return to marketed goods is represented by the origin of the graph. Point A represents the LULC pattern that maximizes P retention given the annual net return to marketed goods produced across the landscape has to increase by at least $50,000 and steady-state carbon storage levels has to increase by at least 1000 Mg compared to the baseline. Point B represents the LULC pattern that maximizes change in P retention given the annual net return to marketed goods produced across the landscape cannot fall more than $3.5 million per year and steady-state carbon storage levels have to increase by at least 25,000 Mg.

and in the case of carbon sequestration, to the people across the globe.

Figure 1 highlights several important points about ecosystem service provision that is useful to communicate to policy makers. First, the more society is willing to sacrifice in marketed net returns and, in some cases, the provision of other services, the more we can gain in a targeted service (in this case, P retention). Second, in this case and as noted in other research (e.g., Polasky *et al.*, 2011), current LULC patterns tend to be inefficient, both from an ecosystem service and marketed returns perspective. In this case the provision of both ecosystem services and marketed net returns could be improved, for example, Point A in **Figure 1**, with smarter land use planning or targeted land use incentive programs.

Market inefficiency is contrary to the fundamental economic assumption that land is always put into its best economic use. These modeled market inefficiencies in **Figure 1** could be due to model limitations (e.g., the costs of LULC transition, not accounted for in our model, are significant, the price and production data we used are erroneous, the modeled yield from SWAT is erroneous) or they could be real (e.g., landowners are choosing not to maximize net returns for some nonmarket-related reason, landowners are unaware of their best options, etc.).

Software That Supports Ecosystem Service Modeling

Up to this point in time, most published ecological process models have focused on one or a collection of closely related services (including the models noted above such as CENTURY, EPIC, SPARROW, etc.). However, there are several inchoate efforts to apply consistent modeling approaches and algorithms across a suite of services (Waage *et al.*, 2011). Artificial Intelligence for Ecosystem Services (ARIES) is one such tool. ARIES first collects information on the user's analysis objective. Next, the user can access data on their landscape of interest from a GeoServer or submit their own maps of the landscape's LULC and biophysical conditions (Villa *et al.*, 2009). ARIES then maps the topology and strength of ecosystem service flow across the landscape for those services that are relevant to the user's analysis objective (Johnson *et al.*, 2010). Flow maps are overlaid with maps of human presence on the landscape to determine the extent to which humans benefit from ecosystem service provision or flow. Default ARIES results are based on observed patterns of service flow on other landscapes that have very similar conditions to the landscape of interest; however, biophysical models can be used to modify these default results as long as the user supplies the necessary information and data (Vigerstol and Aukema, 2011). Service benefits in ARIES can be monetized with a benefits transfer approach or with market, avoided, and replacement valuation techniques discussed above.

Integrated Valuation of Ecosystem Services and Trade-offs (InVEST) is a suite of service models that use production functions to convert maps of LULC, land management, and biophysical conditions into maps of service supply (Tallis *et al.*, 2011). In addition, InVEST models determine the extent to which the supply of services is useful or valuable to people on the landscape. The storm peak mitigation, nutrient retention, and pollination models described above are part of

InVEST. Because data needed to run more complicated service models are often scarce, the current version of InVEST offers relatively simple models with fewer input requirements. The simplicity of the first tier of InVEST models means that results should be viewed with some skepticism. As of now the InVEST is an add-on to ESRI's ArcGIS software. The development team behind InVEST, in cooperation with Google, plans to incorporate several InVEST models into Google's Earth Engine for use with Google Earth. Further, there are plans to publish more complex models and to make InVEST models much more accessible to the open source community.

ENVISION is a GIS-based software system that creates future LULC patterns on a landscape as a function of user-defined economic and environmental policies. The impact of the generated landscape on ecosystem service provision and value can be estimated within ENVISION by accessing any of the ecosystem service models that have been linked to ENVISION.

Other ecosystem service modeling and assessment projects include Ecological Asset Inventory and Management (EcoAIM), EcoMetrix, Ecosystem Services Review (ESR), ESValue, and Natural Assets Information System (NAIS) (Waage *et al.*, 2011).

Using Terrestrial Ecosystem Service Modeling to Affect Decision Making

Ideally, ecosystem service models and analyses will increasingly be used to inform policy-maker decision making and suggest land use policies, regulations, and payment for ecosystem service mechanisms that lead to a net increase in the societal benefits generated on a landscape. As an example of this process, the Sumatra Spatial Planning Forum, a coalition of Indonesian academics and nongovernmental organizations, has been using InVEST to make LULC pattern recommendations that would bolster biodiversity and ecosystem service conservation on the Indonesian island but still provide opportunities for robust economic development. Sumatra's forests have declined in area by a quarter between 2000 and 2010 (Miettinen *et al.*, 2011). Besides reducing the provision of the global public goods of biodiversity conservation and carbon storage, deforestation on Sumatra also degrades the island's water quality and reduces locals' access to valuable forest products.

InVEST has been used to map the distribution of terrestrial carbon stock, water yield, sediment retention, water pollution regulation, and wildlife habitat quality in central Sumatra as of 2008, and under two alternative future LULC scenarios: the Forum's "green vision" of sustainable land use, and the government's land use plan, which corresponds to a business as usual scenario (Barano *et al.*, 2010). These maps are being used to identify areas where conservation-friendly land uses and management practices could maintain or even enhance ecosystem service provision and wildlife habitat at lowest cost where conservation costs are represented by foregone forest plantation and agricultural net returns. The highest priority districts for investment in forest carbon conservation according to the Forum's analysis are those where forest carbon stocks were high as of 2008, the risk of deforestation is high, and monetary returns from foregone plantations and agriculture are low. Similarly, the Forum has identified watersheds where programs to reduce sediment and nutrient pollution could benefit downstream users.

The advice that the Forum is giving the Indonesian government could be improved in several ways. The analysis in Sumatra has been conducted with InVEST's simpler models and coarse global datasets. It is not known how much the targeting maps would change with finer-scale data and more complex models. Some important services, including non-timber forest products and ecotourism, both of which are potentially important for local livelihoods, have yet to be assessed. Further, the rapid rate of deforestation in Sumatra means that targeting maps produced by the Forum may be out of date soon after they are made. Therefore, targeting maps will need to be regularly updated if they are to remain relevant.

In other examples, InVEST has been used to inform a Kamehameha Schools' land use plan (the Kamehameha Schools is the largest private landowner in Hawai'i). The plan was given a 2011 National Planning Excellence Award for Innovation in Sustaining Places by the American Planning Association. InVEST is also being used across Latin America to inform water funds, through which downstream consumers pay upstream land managers for changes in practices to improve water supply. InVEST is used to inform fund size and targeting of payments, in terms of where they are made and the practices they support, in over 30 funds serving major cities. At the largest scale, InVEST is being used in China in a new policy that will designate 25% of the country as "ecosystem function conservation area" status. This policy has twin goals of securing vital natural capital (for flood control, sand storm control, water supply for irrigation and hydropower, and biodiversity) and promoting rural development to alleviate poverty. Decision makers are using InVEST to define the boundaries of the reserves and to evaluate more sustainable livelihoods that could be supported within them.

The Valuing the Arc project, based in Cambridge, UK, is using InVEST models and other modeling techniques and systems to estimate current and future ecosystem service provision and value across the Eastern Arcs Mountain range in Tanzania. The project's goal is to gain a better understanding of the ecosystem services provided by the Mountains and how they contribute to human welfare, and to find solutions to managing these services in a sustainable way. The project hopes to assist policy makers in Tanzania in their efforts to increase their people's welfare.

ARIES has been applied to landscapes in Western Washington, USA; Orange County, CA, USA; San Pedro River, AZ, USA; Vermont, USA; Dominican Republic; Veracruz, Mexico; and Madagascar.

Though the authors have not discussed the approach here in any detail, the benefits transfer approach to creating service value maps has been used to influence policy on several landscapes, including the US state of New Jersey (Liu *et al.*, 2010).

Conclusions

Land use and land management decisions have significant impacts on ecosystems and the goods and services they

provide. By modeling terrestrial ecosystem services we begin to understand the full range of valuable trade-offs associated with land use and land management change, not just those goods and services that are valued by markets. Here we have introduced methods and approaches to estimate how the provision and value of ecosystem services is affected by changes in land use, land management, and ecosystem conditions and human preferences and values on the landscape.

Ecosystem service modeling efforts in the future will need to do a better job of integrating all the spatial and temporal dynamics in service provision and its value. For example, consider the actions of a Minnesota farmer. As described in the detailed model discussion above, her decision to increase the nitrogen fertilization rate on her cropland increases the marketed value of goods produced on her land and may increase the productive value of services that support agriculture but has a detrimental effect on local water quality as nutrient runoff increases. However, if we do not consider the effect of this additional nutrient runoff even further downstream and eventually in the Gulf of Mexico, we have not fully evaluated the impacts of her decision. Further, the timing of impacts (e.g., how long does it take runoff from Minnesota to reach the Gulf) can impact ecosystem service values. Better terrestrial ecosystem service models will consider these complex spatial and temporal dynamics. In addition, the better models will be formally linked to marine ecosystem service models.

Still in its relative infancy, efforts to model trade-offs will become more accurate in the future as we improve our understanding of ecological processes and technical modeling capabilities advance. To improve our understanding of the ecological production functions, more empirical studies designed through an ecosystem service lens are needed to explicitly test the effects of multiple land use and management conditions on ecosystem structure and function. Ideally, such studies would be replicated across multiple ecosystems and biomes to account for spatial or temporal differences. However, no matter how sophisticated our models become, there will be many ecosystem dynamics that will prove difficult or impossible to quantify with certainty, such as threshold responses from changes in global climate. Therefore, responsible ecosystem service modeling will include sensitivity analyses to account for the uncertainty inherent to all modeling efforts.

Appendix

List of Courses

1. Conservation Biology
2. Topics in Ecology
3. Environmental Economics
4. Resource Economics
5. Environmental Studies
6. Sustainability Studies

See also: Agriculture, Sustainable. Biodiversity and Ecosystem Services. Biodiversity-Friendly Farming. Carbon Cycle. Conservation and People. Economic Value of Biodiversity, Measurements of. Economics of the Regulating Services. Ecosystem Services. Feeding the World and Protecting Biodiversity. Identifying Conservation Priorities using a Return on Investment Analysis. Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon. Land-Use Issues. Modeling Marine Ecosystem Services. The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity. The Value of Biodiversity. Tourism, Role of. Trends in Nature Recreation: Causes and Consequences. Valuing Ecosystem Services. Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy

References

Allsopp MH, de Lange WJ, and Veldtman R (2008) Valuing insect pollination services with cost of replacement. *PLoS One* 3: e3128.

Bachelet D, Lenihan JM, Daly C, Neilson RP, Ojima DS, and Parton WJ (2001) MC1: a dynamic vegetation model for estimating the distribution of vegetation and associated carbon, nutrients, and water – technical documentation. Version 1.0. General Technical Report PNW-GTR-508. Portland, OR: United States Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Balmford A, Beresford J, Green J, Naidoo R, Walpole M, and Manica A (2009) A global perspective on trends in nature tourism. *PLoS Biology* 7: e1000144.

Barano T, McKenzie E, Bhagabati N, *et al.* (2010) Integrating ecosystem services into spatial planning in Sumatra, Indonesia. http://www.eea.europa.eu/teeb/integrating-ecosystem-services-into-spatial.pdf.

Bateman IJ, Garrod GD, Brainard JS, and Lovett AA (1996) Measurement issues in the travel cost method: A geographical information systems approach. *Journal of Agricultural Economics* 47: 191–205.

Boyd J and Banzhaf S (2006) *What are Ecosystem Services? The Need for Standardized Environmental Accounting Units*. Discussion Paper 06-02. Washington, DC: Resources for the Future.

Boyd J and Krupnick A (2009) *The Definition and Choice of Environmental Commodities for Nonmarket Valuation*. RFF Discussion Paper 09-35. Washington, DC: Resources for the Future.

Bromley DW (1991) *Environment and Economy. Property Rights and Public Policy*. Cambridge, MA: Blackwell.

Cavailhès J, Brossard T, Foltête J-C, *et al.* (2009) GIS-based hedonic pricing of landscape. *Environmental and Resource Economics* 44: 571–590.

Chan K, Goldstein J, Satterfield T, *et al.* (2011) Cultural services and non-use values. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 206–228. New York: Oxford University Press.

Chan K, Goldstein J, and Satterfield T (2012) Rethinking ecosystem services to better address and navigate cultural values. *Ecological Economics* 74: 8–18.

Conte M, Ennaanay D, Mendoza G, *et al.* (2011) Retention of nutrients and sediment by vegetation. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 89–110. New York: Oxford University Press.

Conte M, Nelson E, Carney K, *et al.* (2011) Terrestrial carbon sequestration and storage. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 111–128. New York: Oxford University Press.

Cresswell JE, Osborne JL, and Goulson D (2000) An economic model of the limits to foraging range in central place foragers with numerical solutions for bumblebees. *Ecological Entomology* 25: 249–255.

Daniel GD, Walsh ME, Shapouri H, and Slinsky SP (2003) The *Economic Impacts of Bioenergy Crop Production on US Agriculture*. USDA, Agricultural Economic Report No. 816.

Dlugofl V, Fiener P, and Schneider K (2010) Layer-specific analysis and spatial prediction of soil organic carbon using terrain attributes and erosion modeling. *Soil Science Society of America Journal* 74: 922–935.

Edwards BK (2003) *The Economics of Hydroelectric Power*. Northampton, MA: Edward Elgar Publishing.

Egan KJK, Herriges JAJ, Kling CLC, and Downing JAJ (2009) Valuing water quality as a function of water quality measures. *American Journal of Agricultural Economics* 91: 106–123.

Ennaanay D, Conte M, Brooks K, *et al.* (2011) Valuing land cover impact on storm peak mitigation. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 73–88. New York: Oxford University Press.

Fisher B, Turner RK, and Morling P (2009) Defining and classifying ecosystem services for decision making. *Ecological Economics* 68: 643–653.

Foley JA, DeFries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.

Foley JA, Ramankutty N, Brauman KA, *et al.* (2011) Solutions for a cultivated planet. *Nature* 478: 337–342.

Gallai N, Vaissière BE, Potts SG, and Salles J (2011) Assessing the monetary value of global crop pollination services. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*. New York: Oxford University Press.

Gassman PW, Reyes MR, Green CH, and Arnold JG (2007) The soil and water assessment tool: Historical development, applications, and future research directions. *Transactions of the ASABE* 50: 1211–1250.

Gathmann A and Tscharntke T (2002) Foraging ranges of solitary bees. *Journal of Animal Ecology* 71: 757–764.

Ghazoul J (2007) Challenges to the uptake of the ecosystem service rationale for conservation. *Conservation Biology* 21: 1651–1652.

Ghazoul J (2008) Debating the ecosystem service rationale for conservation: Response to Kremen *et al. Conservation Biology* 22: 799–801.

Gibbons DC (1986) *The Economic Value of Water*. Washington, DC: Resources for the Future.

Giller KE, Beare MH, Lavelle P, Izac A-MN, and Swift MJ (1997) Agricultural intensification, soil biodiversity and agroecosystem function. *Applied Soil Ecology* 6: 3–16.

Greenleaf S, Williams N, Winfree R, and Kremen C (2007) Bee foraging ranges and their relationships to body size. *Oecologia* 153: 589–596.

Greenleaf SS and Kremen C (2006) Wild bee species increase tomato production and respond differently to surrounding land use in Northern California. *Biological Conservation* 133: 81–87.

Hanley N and Barbier EB (2009) *Pricing Nature: Cost-Benefit Analysis and Environmental Policy-Making*. London: Edward Elgar.

International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) (1989) *Decision Support System for Agrotechnology Transfer Version 2.1 (DSSAT V2. I)*. Honolulu: Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii.

Johnson GW, Bagstad KJ, Snapp R, and Villa F (2010) Service Path Attribution Networks (SPANs): Spatially quantifying the flow of ecosystem services from landscapes to people. *Lecture Notes in Computer Science* 6016: 238–253.

Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) (2011) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*. New York: Oxford University Press.

Keating BA, Carberry PS, Hammer GL, *et al.* (2003) An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* 18: 267–288.

Klein A-M, Vaissière BE, Cane JH, *et al.* (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences* 274: 303–313.

Kremen C, Daily GC, Klein AM, and Scofield D (2008) Inadequate assessment of the ecosystem service rationale for conservation: Reply to Ghazoul. *Conservation Biology* 22: 795–798.

Kremen C, Williams NM, Aizen MA, *et al.* (2007) Pollination and other ecosystem services produced by mobile organisms: A conceptual framework for the effects of land-use change. *Ecology Letters* 10: 299–314.

Kroeger T and Casey F (2007) An assessment of market-based approaches to providing ecosystem services on agricultural lands. *Ecological Economics* 64: 321–332.

Lal R (2004) Soil carbon sequestration impacts on global climate change and food security. *Science* 304: 1623.

Liu S, Costanza R, Troy A, D'Aagostino J, and Mates W (2010) Valuing New Jersey's ecosystem services and natural capital: A spatially explicit benefit transfer approach. *Environmental Management* 45: 1271–1285.

Lonsdorf E, Kremen C, Ricketts T, Winfree R, Williams N, and Greenleaf S (2009) Modelling pollination services across agricultural landscapes. *Annals of Botany* 103: 1589–1600.

Lonsdorf E, Ricketts T, Kremen C, Winfree R, Greenleaf S, and Williams N (2011) Crop pollination services. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 168–187. New York: Oxford University Press.

Lopez-Feldman A and Wilen JE (2008) Poverty and spatial dimensions of non-timber forest extraction. *Environment and Development Economics* 13: 621–642.

Mendoza G, Ennaanay D, Conte M, *et al.* (2011) Water supply as an ecosystem service for hydropower and irrigation. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 53–72. New York: Oxford University Press.

Metherell AK, Harding LA, Cole CV, and Parton WJ (1993) CENTURY: Soil organic matter model environment. *Technical Documentation Agroecosystem Version 4.0. GPSR Technical Report 4*. Fort Collins, CO: United States Department of Agriculture – Agricultural Research Service.

Miettinen J, Shi C, and Liew SC (2011) Deforestation rates in insular Southeast Asia between 2000 and 2010. *Global Change Biology* 17: 2261–2270.

Millennium Ecosystem Assessment (MA) (2005) *Ecosystems and Human Well-being: Current State and Trends* vol. 1. Washington, DC: Island Press.

Nelson E, Montgomery C, Conte M, and Polasky S (2011) The provisioning value of timber and non-timber forest production. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 129–149. New York: Oxford University Press.

Nelson E, Wood S, Koo J, and Polasky S (2011) Provisioning and regulatory ecosystem service values in agriculture. In: Kareiva P, Tallis H, Ricketts T, Daily G, and Polasky S (eds.) *Natural Capital. Theory and Practice of Mapping Ecosystem Services*, pp. 129–149. New York: Oxford University Press.

Ostrom E (1999) *Self-governance and Forest Resources*. Center for International Forestry Research Occasional Paper Number 20.

Parfitt J, Barthel M, and MacNaughton S (2010) Food waste within food supply chains: Quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society, Series B* 365: 3065–3081.

Pataki DE, Carreiro MM, Cherrier J, *et al.* (2011) Coupling biogeochemical cycles in urban environments: Ecosystem services, green solutions, and misconceptions. *Frontiers in Ecology and the Environment* 9: 27–36.

Pattanayak Subhrendu K, Wunder Sven, and Ferraro Paul J (2010) Show me the money: Do payments supply environmental services in developing countries? *Review of Environmental Economics and Policy* 4: 254–274.

Pennington DN, Dalzell B, Mulla D, Polasky S, Nelson E, and Taff S (in preparation) Identifying optimum landscapes for water quality and ecosystem services in an agricultural watershed.

Pergams ORW and Zaradic PA (2006) Is love of nature in the US becoming love of electronic media? 16-year downtrend in national park visits explained by watching movies, playing video games, internet use, and oil prices. *Journal of Environmental Management* 80: 387–393.

Pergams ORW and Zaradic PA (2008) Evidence for a fundamental and pervasive shift away from nature-based recreation. *Proceedings of the National Academy of Sciences US* 105: 2295–2300.

Phaneuf D and Smith VK (2005) Recreation demand models. *Handbook of Environmental Economics* 2: 671–751.

Polasky S, Nelson E, Pennington D, and Johnson KA (2011) The impact of land-use change on ecosystem services, biodiversity and returns to landowners: A case study in the state of Minnesota. *Environmental and Resource Economics* 48: 219–242.

Powlson DS, Whitmore A, and Goulding KWT (2011) Soil carbon sequestration to mitigate climate change: A critical re-examination to identify the true and the false. *European Journal of Soil Science* 62: 42–55.

Robertson G and Grace P (2004) Greenhouse gas fluxes in tropical and temperate agriculture: The need for a full-cost accounting of global warming potentials. *Environment, Development and Sustainability* 6: 51–63.

Robinson EJZ, Albers HJ, and Williams JC (2008) Spatial and temporal modeling of community non-timber forest extraction. *Journal of Environmental Economics and Management* 56: 234–245.

Rotmans J and van Asselt MBA (2001) Uncertainty in integrated assessment modelling: A labyrinthic path. *Integrated Assessment* 2: 43–55.

Rumpel C and Kogel-Knabner I (2010) Deep soil organic matter: Key but poorly understood component of terrestrial C cycle. *Plant and Soil* 338: 1–16.

Sander HA and Polasky S (2009) The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. *Land Use Policy* 26: 837–845.

Schlesinger WH (1990) Evidence from chronosequence studies for a low carbon-storage potential of soils. *Nature* 348: 232–234.

Sitch S, Smith B, Prentice IC, *et al.* (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology* 9: 161–185.

Six JF, Thiet S, and Batten R (2006) Bacterial and fungal contributions to carbon sequestration in agroecosystems. *Soil Science Society of America Journal* 70: 555–568.

Smith JE, Heath LS, Skog KE, and Birdsey RA (2006) Methods for calculating forest ecosystem and harvested carbon with standard estimates for forest types of the United States. *General Technical Report NE-343*. Newtown Square, PA: US Department of Agriculture, Forest Service, Northeastern Research Station.

Smith M (1993) CLIMWAT for CROPWAT, a climatic database for irrigation planning and management. *FAO Irrigation and Drainage Paper No. 49*. Rome: United Nations Food and Agriculture Organization.

Swinton SM, Lupi F, Robertson GP, and Hamilton SK (2007) Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. *Ecological Economics* 64: 245–252.

Syswerda S, Corbin A, Mokma D, Kravchenko A, and Robertson G (2011) Agricultural management and soil carbon storage in surface vs. deep layers. *Soil Science Society of America Journal* 75: 92–101.

Tallis HT, Ricketts T, Guerry AD, *et al.* (2011) *InVEST 2.1 Beta User's Guide*. Stanford, CA: The Natural Capital Project.

Ticktin T and Shackleton C (2011) Harvesting non-timber forest products sustainably: Opportunities and challenges. In: Shackleton S, Shackleton C, and Shanley P (eds.) *Non-Timber Forest Products in the Global Context*, pp. 149–170. Heidleberg: Springer.

Troy A and Wilson MA (2006) Mapping ecosystem services: Practical challenges and opportunities in linking GIS and value transfer. *Ecological Economics* 60: 435–449.

Vesterinen J, Pouta E, Huhtala A, and Neuvonen M (2010) Impacts of changes in water quality on recreation behavior and benefits in Finland. *Journal of Environmental Management* 91: 984–994.

Vigerstol KL and Aukema JE (2011) A comparison of tools for modeling freshwater ecosystem services. *Journal of Environmental Management* 92: 2403–2409.

Villa F, Ceroni M, Bagstad K, Johnson G, and Krivov S (2009) *ARIES (Artificial Intelligence for Ecosystem Services): A New Tool for Ecosystem Services Assessment, Planning, and Valuation*. Burlington, VT: Ecoinformatics Collaboratory, Gund Institute for Ecological Economics, University of Vermont.

Waage S, Armstrong K, Hwang L, and Bagstad K (2011) *New Business Decision-making Aids in an Era of Complexity, Scrutiny, and Uncertainty. Tools for Identifying, Assessing, and Valuing Ecosystem Services*. BSR's Ecosystem Services, Tools & Markets Working Group.

Westrich P (1996) Habitat requirements of central European bees and the problems of partial habitats. In: Matheson A, Buchmann SL, O'Toole C, Westrich P, and Williams IH (eds.) *The Conservation of Bees*, pp. 1–16. London: Academic Press.

Williams JR (1995) The EPIC model. In: Singh VP (ed.) *Computer Models of Watershed Hydrology*, pp. 909–1000. Colorado: Water Resources Publisher.

Williams N and Kremen C (2007) Floral resource distribution among habitats determines productivity of a solitary bee, *Osmia lignaria*, in a mosaic agricultural landscape. *Ecological Applications* 17: 910–921.

Winfree R, Dushoff J, Crone E, *et al.* (2005) Testing simple indices of habitat proximity. *American Naturalist* 165: 707–717.

Modern Examples of Extinctions

Gábor L Lövei, Aarhus University, Slagelse, Denmark

© 2013 Elsevier Inc. All rights reserved.

Glossary

Extinction Disappearance of the last living individual of a species. Extinction can be "local ~", if it concerns a definite population or location; we speak of "~ in the wild" when the only individuals alive of a species are in captivity; and of "global ~" when no living individual remains of a species.

Extinction cascade A chain of extinctions triggered by the extinction of a particular species on which many others depend. Species affected by other species are directly (parasites that live only on that species) or indirectly (predators that rely heavily on the species for food) linked with the extinct species through ecological links. Most species support other ones: a number of specialist herbivores can depend on a plant species for food, or many parasites are host-specific (can only parasitize one species). When these supporting species die out, the dependent species also go extinct. This can trigger a chain of extinctions, the "~". For example, when the passenger pigeon died out, at least two feather lice parasites followed it into extinction. "Co-extinction" occurs when more than one species die out at the same time (this by definition must happen when the host of an obligate parasite dies out).

First-contact extinctions A wave of extinction of species native to a continent or island, following the first arrival of humans to that area.

Living dead A term coined by the American tropical biologist Daniel Janzen, denoting the last living individuals of a species destined to extinction. By definition, extinction happens when the last individual of a species dies. In reality, however, extinction of a species can be certain even earlier. Most species need both males and females to reproduce. If there are no fertile individuals of one sex, the species is doomed to extinction even if several individuals are still alive. Similarly, below a certain population size, a species cannot form a self-sustaining population, and its numbers dwindle. The decline may take many years but its course cannot be easily altered.

Metapopulation A group of populations belonging to the same species that are connected *via* regular migration to each other's habitat patches. An important recent realization is that most species exist as metapopulations and that this is probably the original, 'natural' state of all species.

Pseudo-extinction Extinction of local populations is sometimes erroneously termed as 'pseudo-extinction'. This is misleading because global extinction proceeds through the stepping stones of extinctions of local populations. There is no fundamental distinction between the extinction of a local population and the extinction of a species other than the fact that species becomes extinct when the last local population dies out.

Proximate cause(s) of extinction The actual immediate agent(s) that cause(s) a species to become extinct.

Recolonization The reappearance of a species in an area where it has earlier been present, then became extinct.

Species life span The time between the first appearance of a species in the fossil record and its final disappearance. This time span is typically in the range of millions of years.

Ultimate cause of extinction Being rare (few in numbers) and of limited distribution are precursors to extinction. The causes leading to rarity are the ultimate causes of extinction.

Species Life Spans

The life span of a species can vary widely but no species lives forever. Fossil records indicate that the average life span of an invertebrate species is about 11 million years, while mammal species live for about 1 million years (**Table 1**). Species existing today form only a small fraction of species that have ever lived. If we assume that the average life span of a species is 5–10 million years, and multicellular organisms have been present on Earth for a period of 600 million years, the plant and animal species currently living are not more than 1–2% of all those that have ever lived. For marine invertebrates, an estimated 95% of the species that had ever existed are extinct (Jablonski, 1994). Extinction is thus the natural fate of all living species.

Extinction can and does happen at any time, and extinctions occur continuously. Most of these extinctions are of local populations. For many species, a landscape contains several suitable habitat patches but not all patches are occupied at all

Table 1 Estimates of species' lifespans, from origination to extinction

Group	*Estimated life span, million years*
Dinoflagellates	13
All invertebrates	11
Cenozoic bivalves	10
Diatoms	8
Planktonic foraminifera	7
Echinoderms	6
Marine invertebrates	5–10
Marine animals	4–5
Cenozoic mammals	1–2
Mammals	1
All fossil groups	0.5–5

Source: Reproduced with permission from May RM, Lawton JH, and Stork NE (1995) Assessing extinction rates. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 1–24. Oxford, UK: Oxford University Press.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00215-X

times. Species constantly recolonize unoccupied patches and go extinct in others. The local populations in these patches form a kind of network called a 'metapopulation'. There is constant migration among the habitat patches, some of them (source patches) producing surplus individuals that colonize other patches; others are not so productive (sink patches). When a metapopulation cannot produce enough individuals to compensate for mortality, the species will become regionally extinct. This will result in a range contraction as regular migration between metapopulations does not occur.

Species may become globally rare and subsequently go extinct at any time. The 'background rate of extinction' is estimated at 1–10 species/year through the geological periods (Dirzo and Raven, 2003). However, significant extinctions in Earth's history occurred in clusters. During the last 500 million years, there were five such "mass extinctions", wiping out large proportions of the then-living species. The fossil record related to these has been intensively studied and hotly debated, but without producing an accepted interpretation about their causes. Extinctions, affecting a more restricted group of species have also occurred on a smaller scale. Within those groups the extinctions were significant.

The most recent of such events commenced during the late Quaternary, about 100,000 years before present (yBP), and started to intensify about 40,000 yBP. Since then, several hundreds of land vertebrates, mostly large species (>50 kg body mass) have gone extinct on different continents and islands, at different times. This extinction wave has not yet ended.

Examples, Postglacial

North American Extinctions

In late glacial North America (called the Wisconsin glaciation period, ending about 10,000 yBP), 71% of mid-latitude mammal genera were lost, while in Alaska, the same loss was 56%. This is the opposite that would be expected from environmental conditions – we expect that if climate is the cause of these extinctions, more northerly species would be more severely affected. According to their trophic position, 71% of the herbivores, 67% of the bears, and 50% of the dogs and cats became extinct. Many of these have lived through cycles of glacial and interglacial periods, and extinction was not biased toward either older or newer genera. Environmental changes are unlikely to have caused these extinctions. On the contrary – general conditions were at their worst during the period preceding the extinctions – 20,000–18,000 yBP. Conditions for large mammals have improved afterwards, notably so between 18,000–7000 yBP, when most extinctions occurred.

The postglacial extinctions are generally connected to the appearance of humans in the regions affected. The North American continent has suffered numerous avian extinctions during the end of the last glacial when 19 bird genera became extinct. In spite of taxonomic problems as well as scarce records (10 of these birds are only known from the single area of the Rancho La Brea tar pits in metropolitan Los Angeles, USA), we can generalize that most of these extinct birds were large to very large by avian standards, and the loss of most or all of them can be attributed to ecological dependency on large mammals that also went extinct during the same time. The largest group of these extinct birds were raptors: condors, eagles, accipitrid vultures, and caracaras. Extant hunting birds, including eagles, feed on carrion as well as live prey that they themselves captured; probably the same was true for these birds. The disappearance of large mammals must have resulted in a significant reduction of the available food base, and many of them became extinct. A similar extinction cascade can be observed in the only remaining continent with diverse large ungulate fauna, Africa: where game becomes scarce or disappears, vultures and eagles also disappear. This points to the ecological plausibility of this hypothesis. Two other birds, *Panandris* and *Pyelorhamphus*, related to the North American icterids (Icteridae) of today, are thought to have been in a commensalist association with large herbivores – the 'cowbirds' of the Pleistocene – and followed their hosts into extinction. In Africa, there are several groups of songbirds associated with large mammals, such as oxpeckers and drongos. It is likely that a variety of commensalist relationships also existed in the New World, and these must have been lost with the disappearance of most large mammals in North America.

Australian Extinctions

Australia, until the end of the last glacial, had a fauna of monotremes and marsupials that was as diverse as the placental faunas of other continents. In contrast to those, however, the Australian fauna was rich at the species if not the genus level, and it was not subjected to any significant intercontinental faunal exchange. This lead to a homogeneous fauna that seems to have been unable to withstand ecological stress. During the late Pleistocene, many species went extinct. This loss was comparable, in numbers of species, to extinctions on other continents. For example, while there existed only 15 genera versus the 32 in North America, the number of extinct species is about 60 in Australia and 51 in North America. All 19 species of marsupials heavier than 100 kg, and 22 of the 38 species that are 10–100 kg, have become extinct. Three reptiles and the ostrich-sized *Genyornis newtoni* have met the same fate. A few of the extinct animals are depicted in **Figure 1**. The largest reptile was the varanid lizard *Megalania prisca*, which, at 7 m length, was probably a top predator. Among the extinct monotremes were large echnidas, such as *Zaglossus hacketti*, which was 1 m long and 0.5 m tall. This seems to have been a proportionately large version of the small living echnida. Among the marsupials, there were large carnivores: a large morph of the tiger cat (present on one island until European contact), or the leopard-sized *Thylacoleo carnifex*, a marsupial lion named 'giant killer possum'. One species of the koalas, *Phascolarctos sirtoni*, that was about 30% larger than the living koala, also survived until the very late Pleistocene. The living koala is the only survivor of a diverse family that had its peak in the Tertiary.

The large and varied superfamily of Diprodontoidea has totally disappeared during the late Pleistocene, completing a longer sequence of decline. By the last glacial, only two families were represented. Some species of the Palorchestidae

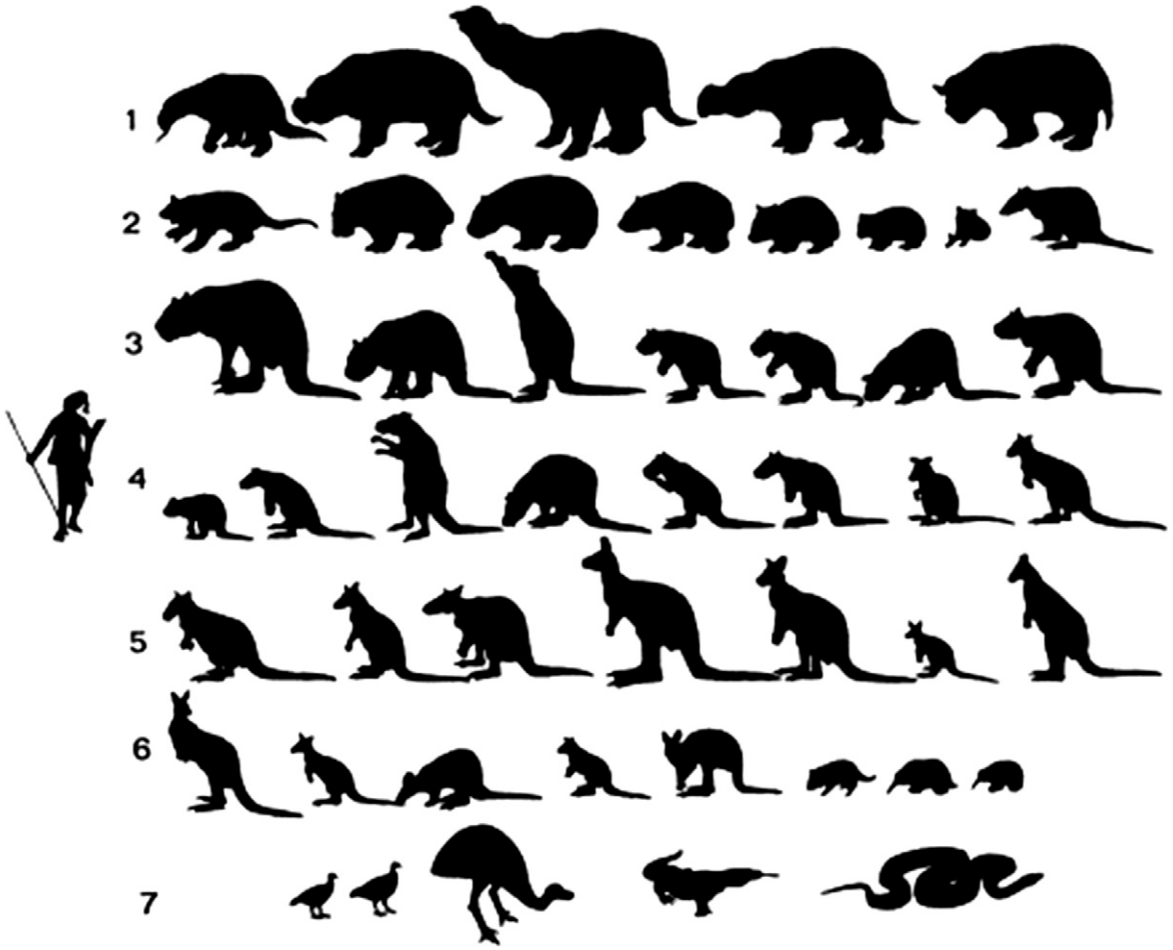


Figure 1 A bestiary of most extinct late Pleistocene Australian vertebrate species. The silhouettes are drawn to scale, with the scale indicated by the human shape. The species, from left to right, are: Row 1: *Palorchestes azcal, Zygomaturus trilobus, Diprotodon optatum, D. minor, Euowenia grata.* Row 2: *Thylacoleo carnifex, Ramsayia curvirostris, Phascolonus gigas, Phascolomys major, P. medius, Vombatus hacketti, Phascolarctos stirtoni, Propleopus oscillans.* Row 3: *Proctoptodon goliah, P. rapha, P. pusio, Sthanturus maddocki, S. brownei, S. occidentalis, S. orientalis.* Row 4: *S. gilli, S. atlas, S. tindalei, S. pales, S. oreas, S. andersoni, Troposodon minor, Wallabaia indra.* Row 5: *Protemnodon roechus, P.anak, P. brehus, Macropus ferragus, M. birdselli, M.siva, M. titan.* Row 6: *M. rama, M. thor, M. piltonensis, M. gouldi, M. stirtoni, Sarcophilus naliarius, Zaglossus hacketti, Z. ramsayi.* Row 7: *Progura naracoortensis, P. gallinacea, Genyornis newtoni, Megalania prisca, Wonambi naracoortensis.* Reproduced with permission from Murray P (1984) Extinctions down under: A bestiary of extinct Australian late Pleistocene monotremes and marsupials. In: Martin PS and Klein RG (Eds) *Quaternary Extinctions: A Prehistoric Revolution*, pp. 600–628, with permission from University of Arizona Press, Tucson.

were beasts that resembled a giant kangaroo but had a tapir-like trunk and huge, curved claws. The cow-like *Zygomaturus trilobus* had a 2 m long body, a huge, broad head, and a narrow, upturned snout. Judging from the frequent fossil remains, it was widely distributed in coastal and mountain Australia. Another browser was the large, slow *Diprotodon optatum* that had a feeding apparatus suggesting that it was browsing tough, succulents and shrubs. The kangaroos (family Macropodidae) today are the largest group of marsupials still living in Australia, although their diversity, too, was seriously reduced by the beginning of the Holocene. From the *Macropus* genus itself, at least eight species died out. Some of these were small, like today's wallabies, but *Macropus titan* and *M. ferragus* were real giants, the latter reaching a height of more than 2.5 m and a body mass of 250 kg. If these species have a living relative, these are about 25% or more smaller.

Although Australia has been inhabited by humans since at least 50,000 yBP, the extinctions were generally not believed to be linked to their presence. Recent evidence based on more exact radiocarbon dating of bid and egg remains casts doubt on this belief, and human predation seems the probable cause also in Australian extinctions.

African Extinctions

When examples of modern extinctions are discussed, Africa is often ignored. Some have called Africa 'the living Pleistocene' because this is the only continent where a diverse and abundant fauna, that bears any resemblance to the Ice Age, remained. However, Africa has had its postglacial extinctions

and as this is the continent with the longest period of human occupation, the analysis of these is potentially very important.

The most impressive examples of such extinctions come from northern Africa. This region, after the dry, hyperarid period between 40,000–12,000 yBP, experienced a moist period during which the fauna included, among other species, the African elephant, white rhinoceros, a zebra, warthog, giraffe, blue wildebeest, hartebeest, eland, roan antelope, and a species of reedbuck. Between 5000–4000 yBP, this moist period changed again, and this fauna disappeared from most of the Sahara, but survived in the Maghreb area of North Africa. Some species were certainly lost, though: the Atlantic gazelle (*Gazella atlantica*), Thomas' camel (*Camelus thomasi*), the giant North African deer (*Megalocerus algericus*), and the long-horned North African buffalo (*Pelorovis antiquus*).

A similar cycle can be observed in southern Africa, although well-dated records are missing from most of this region. In the Cape zone, however, the total disappearance of the long-horned buffalo, giant hartebeest (*Megalotragus priscus*), the giant Cape horse (*Equus capensis*), and the southern springbok (*Antidorcas australis*) happened around 12,000–9500 yBP. Their extinction in southern Africa was, at least partially, related to climate-driven environmental change.

During this period, however, there was a dramatic change in artefact traditions throughout the continent, indicating a very significant shift in human cultures, and with this, probably of hunting techniques and efficiency. Analyses of archeological sites support the hypothesis of increased hunting proficiency: more remains of individuals in their prime age were found as well as relatively more bones of 'dangerous game'. In North Africa, the appearance of domesticated animals may also have contributed to the decline.

The pace and extent of late Pleistocene and Holocene extinctions in Africa parallel that of Eurasia, where there was no sudden and massive extinction wave such as in North America. The only significant difference between these continents and the Americas is the length of human occupation. After Africa, Eurasia has the longest period of human presence, about 700,000 y. North America did not have a previous history of human habitation, and the pattern of extinctions is entirely different. This continent was swept by waves of extinction during the late Pleistocene/early Holocene, coinciding with the migration through the Bering Strait and then southwards of anatomically modern humans. During this period, a very large proportion of the extant fauna disappeared in what is geologically and evolutionarily very short period of time.

Examples, First-Contact Extinctions

On small islands all over the world, many species, especially birds, went extinct during the last 10,000 years (islands have an impoverished mammal fauna to start with, due to dispersal problems). The other common feature of these extinctions is that there was no taxonomic replacement of lost species. These extinctions are so tightly correlated with the arrival of humans, that they were termed 'first-contact extinctions' (FCEs). In the Americas, FCEs occurred between 12,000–10,500 yBP, on the West Indies 7000–5500 yBP, and on Madagascar, 2000–500 yBP (MacPhee and Marx, 1997).

These FCEs can take as little as 1 year on small islands, and up to 1500 y on large islands and continents. On the Commander Islands, east of the Kamchatka Peninsula in the northern Pacific Ocean, humans arrived in 1741. Steller's sea cow (*Hydrodamalis gigas*) was extinct by 1768. On the Mascarene Islands, humans arrived AD1600, and the major extinctions terminated around 1900. In New Zealand, the first human colonists arrived at around AD 1000, and the first major episode of extinctions terminated by AD 1500. In the Mediterranean, humans colonized all the major islands between 10,000–4000 yBP – this also coincides with the extinctions of several endemic species, such as pigmy elephants, rhinoceroses, and hippopotamuses.

Madagascar

The extinction of large mammals and birds on the island of Madagascar during the Holocene was of similar magnitude than the earlier, Quaternary extinctions in North America, Australia, and New Zealand. Today's Madagascar fauna is a pale shadow of a once-diverse assemblage of spectacular species. Humans colonized Madagascar only in historical times. The earliest dated archeological site is from about AD 500. The extinctions started to happen not much later. It is generally agreed that these extinctions were caused by human activities; opinion only differs in what type of activity this was.

Seven of the 17 primate genera have disappeared completely, and two more lost its largest species. The extinct lemuroids were all large, and probably diurnal. The largest of these, *Megadalapis edwardsi*, had males with a body mass between 50 and 100 kg. Members of the smallest extinct genus (*Mesopropithecus*) were about as large as the largest living species, the indri (*Indri indri*). Several of these species had ways of locomotion that are unknown among today's living primates: walking on the ground on four legs (*Hadropithecus*), arm-swinging (*Paleopropithecus*), and vertical climbing similar to that of the koala bear (*Megadalapis*).

The other group that was severely affected includes the large, flightless birds, ratites, commonly known as elephant birds. They are classified into two genera and six to 12 species. The largest of them (*Aepyornis maximus*) had a height of nearly 3 m, and resembled a massive ostrich. The smallest, *Mullerornis betsilei*, was about half this size. These species are thought to have been terrestrial grazers – this ecological group is otherwise only represented by the pigmy hippo (*Hippopotamus lemerlei*). This species, together with a large viverrid (*Cryptoprocta spelea*), and an endemic aadrvark (*Plesiorycterops madagascariensis*) is also extinct. *C. spelea* was the largest known carnivore in Madagascar, and resembled a short-legged puma – so much that earlier it was classified into the cat family (Felidae). The only reptiles that went extinct were giant land tortoises. The two species had carapace lengths of 80 cm and 120 cm, respectively, and were important consumers of ground vegetation.

The Pacific Islands

Humans have gradually colonized the world, and have arrived at several oceanic islands relatively recently. The colonization

history of the Pacific islands is relatively well studied. The human expansion across the Pacific, starting from SE Asia, was accompanied by a wave of bird extinctions on all the islands that humans reached. The time span of this varied due to facts such as distance from neighboring islands, area, and terrain – but the scale of the human-driven extinctions is huge. On all the Hawaiian Islands, we know more endemic species from fossil records only than the number of such species now extant, i.e., more than half the endemic species were lost after human arrival. Every island in Oceania had lost, on average, an estimated 10 species. The total number of islands is about 800, so the loss of species or populations totals 8000. Rails have especially suffered. All Oceanic islands studied so far have had one to four endemic species of rails, and thus an estimated total of about 2000 species, equaling 25% of the global species richness, was lost during human colonization of the Pacific. Most of these species were flightless forest dwellers.

New Zealand

The New Zealand archipelago lies in the southern Pacific Ocean, between the latitudes of 30° and 47°. It is composed of two large and about 300 smaller islands. This land was originally part of the ancient supercontinent of Gondwanaland that also included Antarctica and the other southern hemisphere continents. New Zealand separated from Australia about 75 million yBP and while there are doubts about the Gondwana origin of its species (there being no proof that dry land has been continuous since the separation), the flora and fauna of New Zealand demonstrate a high degree of endemism. Before the arrival of man, the only mammals were marine species and two small bats. There were no large predatory vertebrates, and in their absence, birds prospered. Many species arriving there with the power of flight had become unable to fly, but the best known birds, the ratites, were originally flightless. The New Zealand ratites belong to two orders: the Apterygiformes (kiwis) and the Dinornithidiformes (moas). They have probably been separated from their relatives since the Cretaceous. The kiwis remained small, unobtrusive and nocturnal, and survived into the present.

Moas diversified into several species; current opinion accepts the existence of nine of them. All were ostrich-like, flightless, herbivorous birds with a considerable size range. The largest moa, *Dinornis giganteus*, was about 2 m tall and weighed up to 250 kg , while the smallest, *Megalapteryx didinus,* was about the size of a large turkey, with an estimated mass of 25 kg. In contrast to earlier opinion, most species were probably inhabiting forests, not grasslands. No species of the moas are left and this is one of the best known examples of large-scale human-caused extinctions.

Polynesians have successfully colonized New Zealand about 1000–800 yBP and although we have no reliable record of moa densities before or after this period, nor do we know much about moa evolutionary history in earlier ages, it is well documented that the cause of their demise was that Maori have intensively hunted all species. Archeological sites with large amount of moa bones are found all over New Zealand, some covering many hectares. The most detailed research on them was conducted on the eastern side of the South Island, and these convincingly demonstrate that man was a voracious hunter of moas: their nests were robbed, and their carcasses were probably utilized in a wasteful way. Dogs and rats introduced by man have probably also played a role in the extermination of moas, especially the smaller species. Moa hunting became intensive about one century after the arrival of the Maori, coinciding with a rapid growth in their numbers. Within a few centuries, hunting and forest burning accelerated the decline; by about 400 yBP, moas became so scarce that they were no longer systematically hunted. Continued habitat destruction, sporadic hunting and probably predation by feral dogs continued to destroy birds, and none were left by the time of European settlement.

The extinct bird species that have never been seen by Europeans include not only the moas, but about 20 other bird species (Tennyson, 2010). These were often flightless (79% of all extinct species), ground-nesting (89%), diurnal (96%), and larger than the closest surviving relative (71%). Fifteen of these were endemic to New Zealand, and five were very similar to living Australian relatives. No less than four of the 15 were rails, thus echoing the extinction patterns of the Pacific islands (see earlier). Other birds lost include a flightless goose (*Cnemiornis calcitrans*), a giant rail (*Aptornis otidiformis*), a swan (*Cygnus sumnerensis*) and several flying birds. A co-extinction with the moas was the disappearance of the giant Haast's eagle (*Harpagornis moorei*), the largest known flying bird; it was probably preying on moa. After the extinction of its prey, or possibly even earlier, when the prey became rare, the predator died out.

Examples, Historical Extinctions (1600–)

Since AD1500, during the 'modern era', extinctions were closely correlated with the European expansion, starting with the discovery of America in 1492. The time span of resulting extinctions differ by species and the area affected, but it gradually expanded to include all areas and habitats of the Earth.

New Zealand

During the European period of occupation in New Zealand (although 'discovered' by the Dutch seafarer Abel Tasman in 1642, settlement by Europeans did not start until about 1840), at least five further bird species have become extinct. There is no doubt that the environmental changes brought by Europeans in about 200 years exceed those caused by the Polynesian occupants during the preceding centuries. This difference, however, is not due to intent but due to the difference in technology. The impact of the initial colonization in terms of extinctions is larger and more obvious because the Polynesians arrived to predator-free islands.

One of the recently exterminated species is the Stephen Island wren (*Xenicus lyallii*), the only known flightless songbird. Stephen Island is a small island in the Cook Straight, between the North and the South Islands of New Zealand. The first specimen of this bird was brought to the lightkeeper's house by his cat. Described as a new species to science, it was

exterminated by the same cat within one year (1894). No person has ever seen a live specimen.

The catastrophic impact of predator invasion is exemplified by another New Zealand story, the rat invasion of Big South Cape Island. Big South Cape Island lies south of the South Island, and harbored several endangered species until 1964, when ship rats (*Rattus rattus*) got on shore from a shipwreck. In two years' time, the rats reached very high densities, and four species of birds endemic to New Zealand, one native bat species (greater short-tailed bat, *Mysticina tuberculata robusta*) and numerous invertebrates became extinct. Other species were removed from the island, and thus, for example, the South Island saddleback (*Philesturnus carunculatus*), a thrush-sized bird, survived.

Many more species of birds, reptiles, amphibians, sea mammals, and invertebrates have also suffered a reduction of their former range. Typically, they became extinct on the main islands of New Zealand, surviving only on offshore ones that were typically – but accidentally – free of introduced mammals. For example, the tuatara, *Sphenodon punctatus* (with its sister species *S. guentheri*), the only living relative of the dinosaurs, has been found in early archeological sites on the main islands. Today it only survives on a few offshore islands. It did not survive on any island where Polynesian rats (*Rattus exulans*) are present, but reaches high densities on rat-free islands. Another example of on-islands-only species is the little spotted kiwi (*Apteryx oweni*) that had only one self-sustaining population on Kapiti Island near Wellington, and the recently discovered Middle Island tusked weta (*Motuweta isolata*, a relative of grasshoppers).

Hawaiian Islands

The Hawaiian Islands are a group of volcanic islands, distant from any other land mass, in the middle of the Pacific Ocean. They were reached by Polynesian settlers at around AD 500. These islands have had a very diverse and unique fauna and flora, and especially the vertebrates were seriously decimated. The best documented, again, are the bird extinctions. The extinct species include two species of flightless geese, flightless ibises, rails, a long-legged owl, a sea eagle, a number of honeycreepers, and at least one species of crow (since 2002, a second species, the Hawaiian Crow, *Corvus hawaiiensis* is also extinct in the wild). Further, there is a group of species that have living populations on one island or another, but not on the one where they were found as subfossils.

The patterns of extinction are strikingly similar to those of New Zealand, except that there are not many songbird extinctions reported from New Zealand. Man-induced changes in Hawaii may have been more extreme, or New Zealand originally did not have many songbirds.

Significant paleo-zoological findings are accumulating and it is difficult to draw a reliable and comprehensive picture about the original fauna of the Hawaiian Islands as well as a proper assessment of the extent and nature of extinctions. However, what we know now indicates that the effect of man as exterminator, direct or indirect, of the fauna of this island archipelago is much more significant than earlier thought. Authorities claim that our previous knowledge of the prehuman Hawaiian fauna was so poor and extant species richness is so pale with respect to the original one that ecological and biogeographical studies using recently collected data are critically weakened.

The prehistorically extinct birds of the Hawaiian Islands include one species of petrel, at least 10 mostly flightless species of geese, three flightless ibises, eight of rails, three of long-legged owls, one *Accipiter*, two large crows, one large melliphagid, and 15 species of Hawaiian honeycreepers (relatives of finches).

In the early 1980s, 82 endemic bird species were known from the Hawaiian Islands. Fifty-three of these became extinct prehistorically (before 1778, when a British expedition, lead by James Cook, discovered the islands). Area and elevation show significant positive correlations with the number of fossil and historically recorded bird species: small or low islands lost more species than large or high ones. On Molokai, the smallest of the five largest islands, with 676 km^2 of area and 1515 m a.s.l., there are 21 fossil and nine historically known species. On Hawaii, the largest and highest (10,646 km^2, 4206 m a.s.l.), three fossil and 23 historic species are known – although more fossil species are expected after more excavations are done.

Extinction Paradoxes

Interestingly, current extinction rates seem to be lowest in areas with a long history of human habitation.

Plant extinction rates in areas with Mediterranean climate are low, ranging from 1% of all species in W Australia to 0.15% in the Mediterranean itself. Current threats to plants are one order of magnitude larger: 10.2–15.2% of species are considered threatened. The suspected cause of the current low extinction rate is a ''recording error'': many of the extinctions occurred in the 'prebotanical age'. Indeed, the current extinction rates are lowest where agricultural cultivation has been the longest, 8000–6000 yBP. This is consistent with the view that most vulnerable species will have been lost by the time botanical investigations started.

Similarly, the proportion of bird fauna extinct on the Pacific islands is inversely proportional to the length of human habitation of these islands: 80% of Hawaii's bird fauna is recently extinct or endangered against 10% on Vanuatu. Hawaii has been inhabited for about 1500 years, and Vanuatu for 4000 years. Pimm and co-workers (1995) argue that the sensitive species have been eliminated by the first colonists before record-keeping began, and thus we have no direct evidence of first-contact extinctions in the Pacific.

Examples, Extinction Cascades

A species' ''ecological environment'' almost always includes other organisms that are essential for its survival. Species are linked through trophic links – they eat each other. Other vital ecological links include pollination, dispersal of seeds, and providing habitat. For example, bees, birds and bats pollinate flowers, birds and mammals disperse seeds, trees provide nesting holes for birds. The extinction of a species can have reverberating consequences, affecting other species that are,

directly (such as obligate parasites) or indirectly (such as shared predators) linked with the extinct species through such ecological links. Most species support other ones: a plant species can have a number of specialist herbivores that depend on it for food, or many parasites are host-specific (can only parasitize one species). When these supporting species die out, the dependent species also go extinct. This can trigger a chain of extinctions, termed an 'extinction cascade'.

When the last passenger pigeon, a female named Martha, died in the zoo in Cincinatti, USA, in 1914, at least two species of feather lice that were obligate parasites of this species, must also have perished, although there is no mention of this in any list of extinct species.

All moas, a group of nine species of ratites of different size went extinct not long after Polynesians settled in New Zealand. The largest known raptor, the giant Haast's eagle also followed them into extinction. As there are large middens with thousands of moa bones at several sites in New Zealand but these do not contain bones of the eagle, it is thought that the eagle was not a victim of persecution or hunting, but became extinct after its food base, the formerly very common moa, disappeared.

Likewise, several bird species that went extinct in North America at the end of the last ice age are suspected to have died out in an extinction cascade (see earlier).

Problems in our Understanding of Extinctions

We are aware that our knowledge of the actual extent of even recent extinctions is very fragmented and incomplete. There is, in other words, a huge difference between documented and real extinctions. This is due to a series of reasons. Some of these can be overcome with the development of science, but several of them result from the organisms' biology.

How Many Species are There?

We do not know, even to an order of magnitude, how many species we share the Earth with. Current estimates of global species richness tend to converge to a range from 3 to 15 million species. This has an obvious consequence for the estimation of extinction rates: 1000 species is a different relative share of a global total of 3 million versus 15 million.

Record-keeping is Insufficient/Inadequate

The documentation of extinction is also uneven, both geographically (mostly from islands and the northern temperate region) and taxonomically (higher organisms are better reported).

Since 1500, only 753 animal and 588 plant species are listed as extinct (**Table 2**). We strongly suspect that even among vertebrate groups, documented extinctions are serious underestimates. For example, on the Solomon Islands, where 164 bird species have been recorded, 12 have not been seen this century, but only one is listed as extinct. In Malaysia, a 4-year search for 266 freshwater fish species reported in the last century found only 122, yet few are recorded as extinct.

Table 2 Species in major taxa that have become extinct since 1600 or are threatened with extinction

	Total number of species			*% Extinct*	*% Threatened in 2008*
	Described (thousands)	*Extinct since 1500*	*Listed as threatened in 2008*		
Animals, total	1400	753	8465	0.05	0.60
Invertebrates, total	1, 232.40	256	2499	0.02	0.20
Molluscs	81	191	978	0.24	1.21
Crustaceans	40	4	606	0.01	1.52
Insects	950	61	626	0.01	0.07
Others	161.3	n.a.	289	n.a.	0.18
Vertebrates	61.3	497	5966	0.81	9.73
Fishes	30.7	176	1275	0.57	4.15
Amphibians	6.35	38	1905	0.60	30.00
Reptiles	8.7	23	423	0.26	4.86
Birds	9.9	153	1222	1.55	12.34
Mammals	5.4	107	1141	1.98	21.13
Plants, total	298.5	588	9373	0.20	3.14
Mosses	16	n.a.	82	n.a.	0.51
Ferns and allies	12.8	n.a.	139	n.a.	1.09
Gymnosperms	0.98	2	323	0.20	32.96
Dicotyledons	199.4	120	7122	0.06	3.57
Monocotyledons	59.3	462	782	0.78	1.32
Palms	2.8	4	925	0.14	33.04
Others (lichens, mushrooms, brown algae)	50	n.a.	9	n.a.	0.02
Total, animals + plants	1642.2	1341	17838	0.08	1.09

Combined from May, *et al.* (1995) and IUCN (2010).

The current method of documenting extinct species is not entirely valid biologically. A species becomes 'officially extinct' with the death of the last living individual. A species may be destined to extinction long before this happens. If mortality surpasses reproductive success, a species may get onto an 'extinction trajectory' – numbers will continuously decrease without reversal, but it will take many years or decades until all individuals perish. Likewise, if there are no reproductively successful pairs remaining, the species has no hope of surviving, even though not all individuals are dead yet. These species are termed the 'living dead'.

Sometimes species can go through a 'genetic bottleneck' when populations become so small that genetic variability practically disappears. One such example is the cheetah (*Acinonyx jubatus*), a fast-running predator in the cat family, which must have gone through such a population crisis some thousand years ago. Today all living cheetahs are genetically virtually identical. The black robin (*Petroica traversi*), a small, endemic songbird on the Chatham Islands of New Zealand, had only one fertile female in the 1980s. In one of the success stories of conservation, this species was brought back from the brink of extinction – but genetic variability of the species is much reduced (although apparently without detrimental effects). Without human intervention, this species would have become extinct.

We also have to consider that so many species are recorded from only one location (for example, up to 40% of beetle species kept in the Natural History Museum collection, London, United Kingdom) that it is difficult to assess anything but local extinctions.

Uneven Recording Effort

Documented insect extinctions are 100 times lower than vertebrate extinctions. The difference is even larger if we consider that there certainly are many more insect species than we know today. Part of this is due to a large difference in the effort we devote to the study these groups. There are 100 times more vertebrate than invertebrate taxonomists, and 10 times more than the number of taxonomists of flowering plants. This uneven attention by humankind to different groups is also evident from the rate of describing new species. This is only 0.03–0.05% new species/year for birds. In tropical areas, one of every 100 plant specimens (1%) is new for science; for insects, fungi, and marine macrofauna, this can reach 20–80%.

Taxonomic and Habitat Bias

Our knowledge is also very biased by habitats and taxonomic relationships. We know much more about forests than seas, and while most of the mammal or bird species of the world are known, this cannot be said of other important and species-rich groups of organisms like fungi, nematodes, or arthropods.

The great taxonomic bias in our records is well exemplified in the 61 extinct insect species: 33 of these are butterflies and moths. These groups do not constitute more than half of all insect species (they are more likely to be about 25%); their prevalence merely reflects that butterflies are much better studied than other insect groups. It is perhaps realistic that 51 of these are island species but 42 of them are from Hawaii which is certainly overrepresented. Similarly, of the 10 extinct continental species, nine are from North America. This indicates the distribution of researchers, not the real distribution of threatened or extinct species.

Geographical Bias

Our knowledge is particularly scant in areas of the Earth with the highest biological diversity, the tropics. Any change is much better documented in the northern temperate regions where only a minority of the global biodiversity can be found.

The geographical bias and variability reported in the literature include patterns that are real, while others imaginary. Sixty-one percent of all recorded animal extinctions are from islands – this is probably a real pattern. The numerical preponderance of the Pacific Islands is due to both their large numbers and recent human colonization. However, that 2/3 of recent animal extinctions are from North America and the Caribbean, and 20% is from Australia is certainly an artefact. All 45 plant extinctions in Africa include species from the Cape flora, and 2/3 of continental plant extinctions are from North America and Australia. The rarity of such records from South America, Asia and other regions of Africa is surely an artefact.

Methodological Obstacles

Some of the above artefacts are historical and irreparable. Further inaccuracies result because much of extinction information is gathered by paleozoology. We have never witnessed these extinctions, and only remains of the extinct organisms are found. There are special difficulties in studying and interpreting fossil or subfossil material. Just to mention one, the screens used for sieving soil when excavating animal bones have been, until recently, too coarse to retain bones of small bird species. As a natural consequence, our knowledge of the true extent of bird extinctions is grossly biased because there are many more small than large species of birds (just as in other groups of organisms).

Inherent, Biological Problems

Most of our fossils are from marine organisms, mostly because they often have calcareous body parts that fossilize well. Fossilization on dry land is different: some groups in some climatic regions (for example, insects in tropical climates) simply do not fossilize.

Birds are better known than other organisms because their skeletons fossilize better and their taxonomy, generally, is better known. In contrast, the original vegetation of the Hawaiian Island lowlands is a matter of conjecture as it was largely destroyed before botanists arrived to collect there. Entomologists can only speculate what the effect of this deforestation could have been for the arthropods as very few insects are preserved under Pacific island conditions.

Theoretical Aspects of Extinction

The first step in the extinction process for a species is to become rare. It is conceptually useful to distinguish between ultimate causes of extinction (what causes species to be rare and thus vulnerable to extinction in the first case) and proximate causes of extinction (what is the actual cause of extinction). These latter generally include demographic and environmental stochasticity (random, large fluctuations in density and environmental conditions), genetic deterioration and social dysfunction, although their respective importance is not well understood. Ultimate causes include hunting, habitat destruction, invasion by introduced species, and pollution.

There are two general tendencies that are relevant for the study of extinctions.

1. Species that are widespread tend to be abundant as well, but the causes of this positive correlation are not well understood. This also means that species most at risk from extinction (those that are sensitive to proximate causes) have small geographic ranges, because they will also be locally rare. This double jeopardy may be serious when populations and ranges are artificially reduced by ultimate causes of extinction.
2. The distribution of tropical species is generally more restricted than that of temperate species. Smaller ranges have been documented for tropical than temperate-region trees, molluscs, crustacea (crabs and relatives), fish, amphibians, reptiles, and mammals. A related trend is that average population densities of individual species increase from the Equator toward the poles (proven for invertebrates, mammals, and birds). This fits with trend 1 above, and is also consistent with a decline in range sizes toward the tropics.

As a consequence, disproportionately more tropical than temperate species are threatened with extinction. Of the 1029 threatened bird species, 442 live in tropical forests, more than twice the number of species living in wetlands, the next most threatened habitat category.

Most (direct as well as circumstantial) evidence indicates that most of the recent extinctions were caused by humans. Climate change has been invoked in some cases but the evidence for this is not strong. The actual human impact can be overhunting, habitat destruction, voluntary or accidental introduction of non-native species (mostly predators: cats, dogs, rats or browsing herbivores: pigs, goats, sheep). It was also suggested that humans have spread an extremely virulent pathogen, causing a 'hyperdisease'. In molluscs, birds, and mammals, that went extinct since 1600 and have a known cause, 23% were due to hunting, 36% to habitat destruction, and 39% due to the introduction of exotic organisms. Once again, our knowledge is rather sketchy: in mammals that became extinct since 1600, only 30% have an established proximate cause of extinction.

Introductions as a Threat to Species

Introduced species have often been implicated in the extinction of native species. Many introduced species, however, have had no detectable effect on species in their new environments. The massive spread of organisms by humans to other areas of the globe may increase local diversity, but will result in large losses in global biodiversity. To understand the danger that pan-mixing of the Earth's fauna and flora signify, let us consider a thought experiment in island biogeography. Species richness on an island is largely determined by its area: the larger the area, the more species the island contains. The same applies for continents. For example, mammal species richness is related to the size of the individual continents. The resulting correlation allows us to extrapolate the global species richness. A supercontinent, with an area equal to the total dry land on Earth would support about 2000 mammal species. Currently, there are about 4200 mammal species. Therefore geographical isolation allowed evolution to generate nearly twice the biodiversity that could otherwise, on the basis of habitat area alone, be expected. As today human-assisted invasion is becoming a more and more prevalent biogeographic phenomenon, the individual continents are more and more like one supercontinent. It is not surprising that more extinctions are predicted, with possibly catastrophic consequences for biodiversity.

Often there is more than one cause of extinctions. For example, the kokako (*Callaeas cinerea*), an endemic wattlebird in New Zealand became extinct in most of its former distribution range (and is on the brink of global extinction) due to a combination of factors. These include the contraction and fragmentation of its original forest habitat plus the effects of introduced predators, mainly the European stoat (*Mustela erminea*) and the Australian brushtail possum (*Trichosurus vulpecula*).

Insights from Population Dynamics

A further difficulty in understanding extinction is that the actual process of extinction is also very imperfectly documented. Only a few documented examples exist that link population decline to changes in species distributions. The stepping stones of global extinction are local extinctions, so it is logical to assume that as a species becomes more restricted and rare, its distribution range will become fragmented and gradually smaller. The decline of the European fir tree (*Abies* sp.) was indeed accompanied by population range fragmentation. The skipper butterfly (*Hesperia comma*) in Britain has crashed during the twentieth century. This process left scattered and highly fragmented populations by the 1950s. The same happened with many bird species in New Zealand (kokako, kaka *Nestor meridionalis*, a large parrot or the weka *Gallirallus australis*, a flightless rail).

The sensitivity of fragmented populations is underlined by a trend seen in the success of reintroduction attempts: 76% of 133 documented reintroductions into former "core" areas succeeded while only 48% of 54 translocations to periphery or beyond did so. However, not all species show similar range dynamics in the process of becoming rare. The Kirtland's warbler (*Dendroica kirtlandii*), a small insectivorous bird living in North American forests, withdrew into the historical center of its range during a recent 60% population collapse.

Special Traits Related to Density

While population densities typically fluctuate widely, some species are naturally rare. The study of rarity holds promise to understand processes related to extinction, although only vague clues are available today. It would be important to know, for example, if naturally and anthropogenically rare species are equally sensitive to proximate causes of extinction.

In plants, locally rare and geographically restricted species have lower levels of self-incompatibility, and poorer dispersal abilities. Rare plants are overrepresented in certain families (Scrophulariaceae, Lamiaceae) and under-represented in others (Rosaceae), at least in North America. This may indicate that there are some biological traits and adaptations that are shared by rare species.

Populations of large-bodied species fluctuate less than smaller-bodied taxa (although the measurement of population variability is not as easy as the concept suggests), yet body size is not a useful predictor of extinction risk. In birds, body size was not a useful predictor of rates of population increase or decrease in a global sample of threatened species from 12 families at various trophic levels.

One important but counter-intuitive fact is that trophic position has no consistent effect on extinction. It is difficult to detect a consistent tendency for more frequent extinction of species at higher trophic levels, fossil or extant. This is complicated by the difficulty in separating body size and trophic position (species at higher trophic levels are mostly large). Top predators, in other words, are not more prone to extinction than consumers at other levels.

It seems that large-bodied species are vulnerable to ultimate causes of extinction (hunting, habitat destruction) but less so to proximate causes (their populations fluctuate less).

Time Factor

The above important determinants are thought to vary in ecological time, 10–1000 years. However, as no species lives forever, there may be processes that are operating in evolutionary time. If range and abundance are also species-specific characteristics, some species will be more extinction-prone (i.e., naturally rare and restricted in distribution) than others.

Among songbirds on West Indian islands, older taxa occur on fewer islands, have more restricted habitat distributions and have reduced population densities. However, body size – abundance plots within tribes of birds have more positive relationships than expected in taxonomically ancient tribes, meaning that large-bodied species in old tribes are more common than small-bodied ones. This could be the product of differential extinctions of large-bodied rare species over time.

Bivalves and gastropods in Cretaceous fossils achieved characteristic range sizes early in their history, and this changed little thereafter. In this group, locally endemic species have much higher chances of extinction than more cosmopolitan genera.

The Present: A Full-Fledged Mass Extinction

There are very few documented cases of extinction of lower organisms. This is an inevitable consequence of our ignorance of the degree of biodiversity of those groups. Consequently, it is not surprising that their current rate of extinction can only be roughly guessed. It is extremely probable that the rate of recent extinctions is several magnitudes higher than the 'background extinction rates' and probably surpasses any similar mass extinction events in the Earth's history.

Among the comparatively well-studied birds and mammals, the documented extinctions in the twentieth century numbers about 100 species. There are a total of 14,000 species of these classes, so the documented extinction rate in this century is about 1%. This translates to an expected average life span of a bird or mammal species of around 10,000 years. This is 100–1000 times shorter than the average life span calculated from the fossil record.

Three different methods for predicting impending extinction rates suggest future life spans of birds and mammals of 200–400 years if current trends continue. These impending extinction rates are at least 10,000 times higher than background rates in the fossil record.

All evidence suggests that a sixth mass extinction event in the history of Earth is underway. While the total effect cannot yet be forecast, we know that the sixth mass extinction will be unique in the Earth's history. It will be the first one resulting not from environmental changes but the extraordinary population growth and the activities of one species. Our species now uses an estimated 25–50% of terrestrial net primary productivity. This is without precedent, and will make the coming extinction qualitatively different from all previous mass extinctions. We know enough to realize the gravity of the problem. We need a more elaborate understanding of the phenomenon, how it affects different groups and geographic locations, as our conservation actions will become more and more critical for the future of life on Earth.

See also: Comparing Extinction Rates: Past, Present, and Future. Extinction, Causes of. Extinction in the Fossil Record. Geologic Time, History of Biodiversity in. Human Impact on Biodiversity, Overview. Loss of Biodiversity, Overview. Mammals (Late Quaternary), Extinctions of. Mammals (Pre-Quaternary), Extinctions of. Marine Mammals, Extinctions of. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of. Natural Extinctions (not Human Influenced)

References

Dirzo R and Raven PH (2003) Global state of biodiversity and loss. *Annual Review of Environment and Resources* 28: 137–167.

IUCN (2010) *IUCN Red List of Threatened Species*. Version 2010.4. www.iucnredlist.org.

Jablonski D (1994) Extinctions in the fossil record. *Philosophical Transactions of the Royal Society of London, Series B* 344: 63–76.

MacPhee RDE and Marx PA (1997) The 40,000-year plague: humans, hyperdisease, and first-contact extinctions. In: Goodman S and Patterson B (eds.) *Natural*

Change and Human Impact in Madagascar, pp. 169–217. Washington, DC: Smithsonian Institution Press.

May RM, Lawton JH, and Stork NE (1995) Assessing extinction rates. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 1–24. Oxford, UK: Oxford University Press.

Pimm SL, Moulton MP, and Justice J (1995) Bird extinctions in the central Pacific. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 75–87. Oxford, UK:Oxford University Press.

Tennyson AJD (2010) The origin and history of New Zealand's terrestrial vertebrates. *New Zealand Journal of Ecology* 34: 6–27.

Molluscs

David R Lindberg, University of California, Berkeley, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 235–247, © 2001, Elsevier Inc.

Glossary

Ctenidia (ctenidium *sing.*) The molluscan gill, comprising a characteristic muscular axis from which multiple filaments arise. The relative placement of ciliary bands, nerves, muscles, and blood spaces is similar in all extant taxa.

Mantle An ectodermal tissue layer that secretes calcium carbonate in the form of spicules or shell.

Radula A ribbon-like structure bearing transverse rows of small "teeth," typically used to collect food and move it into the mouth.

Torsion The 180° counterclockwise rotation of the visceral mass and mantle cavity during the larval stage of gastropods.

Introduction

The molluscan body plan typically consists of four body components: (1) a head with tentacles and eyes, (2) a ventral muscular foot, (3) a dorsal visceral mass, and (4) the enveloping mantle that secretes the shell. A space between the covering mantle and the side of the foot forms a mantle (or pallial groove). In most molluscs the groove deepens in the posterior region to form a cavity that contains a pair of gills, or ctenidia, as well as openings of the rectum, paired renal organs, and gonads.

The mouth typically is equipped with jaws and a radula. Numerous glands and sacs are associated with the mouth (or buccal chamber) and esophagus, which opens into a stomach. A digestive gland also opens into the stomach. The hindgut, or intestine, is typically long and highly looped or coiled. The nervous system is concentrated in the head and consists of four ganglia—cerebral, visceral, pedal, and pleural. They are connected by commissures; the paired pedal nerve cords extend ladder-like through the foot. The circulatory system consists of a heart enclosed within the pericardium. The heart has multiple auricles and a single ventricle. The system is typically open except in cephalopods, which have capillaries. The excretory system is paired and connected to the pericardium as well as the gonads in some taxa. The gonads are also paired, but can be fused into a single structure (Polyplacophora) or reduced to a single organ (Gastropoda). Separate gonoducts are present in some taxa, and in other taxa the gonads empty into the kidneys. These connected, mesodermal structures (pericardium, kidneys, and gonads) likely represent the reduced coelom of the Mollusca.

The characters that diagnose the Mollusca are a shell-secreting mantle with a tripartite mantle edge divisible into outer, middle, and inner folds; the secretion of an outer proteinaceous layer, or periostracum; the secretion of multiple crystal types and arrangements (microstructures) within the shell; and a pericardium. Plesiomorphic characters include a mantle groove with flow-through water current patterns, bilateral symmetry, and cephalic tentacles.

This body plan has been substantially modified both among and within groups. Diversification appears to have occurred very early in the history of the Mollusca, but there has been surprisingly little change in some groups. For example, the shells of late Cambrian monoplacophorans are almost identical to those of living taxa despite 450 million years of evolution. Other examples of little change include protobranch bivalves, nautiloids, and scaphopods. There have also been some large and notable extinctions, including those of the bellerophonoids, hyoliths, rostroconchs, and ammonites.

Temporal and Spatial Distributions

The Mollusca include some of the oldest metazoans found in the fossil record. Late Precambrian rocks of southern Australia and the White Sea region in northern Russia contain bilaterally symmetrical, benthic animals with a univalved shell (*Kimberella*) that resembles those of molluscs in many respects. The earliest unequivocal molluscs are found in the early Cambrian, and there is even some support for a Precambrian presence (Coeloscleritophora). Most of the familiar groups, including gastropods, bivalves, monoplacophorans, and rostroconchs, all date from the early Cambrian, whereas cephalopods are first found in the middle Cambrian, polyplacophorans in the late Cambrian, and the Scaphopoda in the middle Ordovician. Most of these taxa tend to be small (<10 mm in length). After their initial appearances, taxonomic diversity tends to remain low until the Ordovician, when gastropods, bivalves, and cephalopods show strong increases in diversity. For bivalves and gastropods this diversification increases throughout the Phanerozoic, with relatively small losses at the end-Permian and end-Cretaceous extinction events. Cephalopod diversity is much more variable through the Phanerozoic, whereas the remaining groups (monoplacophorans, rostroconchs, polyplacophorans, and scaphopods) maintain low diversity over the entire Phanerozoic or became extinct (**Figure 1**).

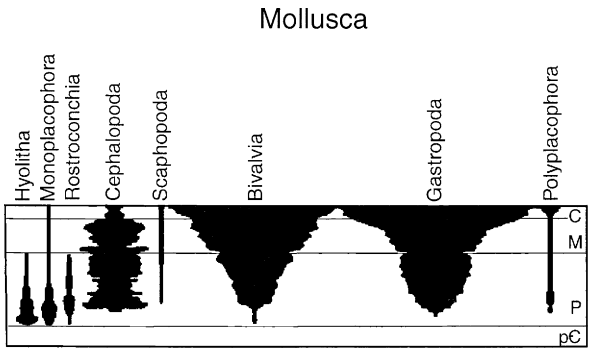


Figure 1 Spindle diagrams of the relative diversity of major molluscan groups through time. (Valentine, James; Phanerozoic Diversity Patterns, Copyright © 1985 by Princeton University Press. Reprinted by permission of Princeton University Press.)

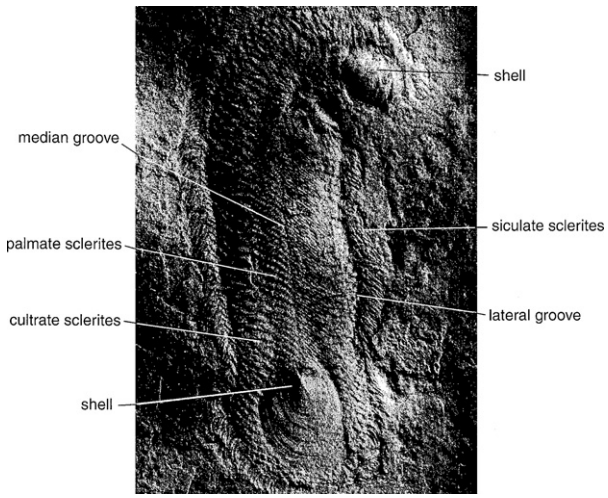


Figure 2 Articulated halkieriid from the early Cambrian of Greenland. (Peel, J. S.; Functional morphology, evolution and systematics of Early Paleozoic univalved molluscs. Bulletin Grønlands Geologiske Undersøgelse 161. Copyright © 1991 by Geological Survey of Greenland. Reprinted by permission of Geological Survey of Denmark and Greenland.)

Molluscs occur in almost every habitat found on Earth. Large concentrations of gastropods and bivalves are found at hydrothermal vents in the deep sea, whereas other gastropods live above tree line in Arctic tundra and on mountain slopes. Molluscs occur in the wettest environments of tropical rain forests and in the driest deserts, where annual activity patterns may be measured in hours. They live below ground in the lightless world of aquifers and caves, and even interstitially in groundwater (stygobionts), but are also familiar organisms of the seashore, rivers, and lakes and even our gardens. The only thing molluscs do not do is fly, unless the passive aerial dispersal of minute land snails among the Pacific Islands by high winds is considered to constitute flying.

Like many other organisms, marine molluscs reach their highest diversity in the tropical western Pacific and decrease in diversity toward the poles. Marine diversity is also highest nearshore and becomes reduced as depth increases beyond the shelf slope. In terrestrial communities gastropods can achieve great diversity and abundance: as many as 60–70 species may coexist in a single habitat and abundance in leaf litter can exceed more than 500 individuals in 4 liters of litter. Abundance and diversity for some groups can also be higher in temperate communities than in tropical settings. In freshwater communities, species diversity tends to be lower but abundances of sessile molluscs such as the zebra mussel *Dreissena polymorpha* can exceed more than 3000 individuals per square meter.

Major Groups

The major groups of living molluscs are clearly dissimilar from one another and have long been recognized as distinct taxa. However, not all were originally recognized as belonging to the Mollusca. For example, the worm-like bodies of the aplacophorans were perplexing to early biologists and required study of their internal anatomy to ultimately recognize their affinities with the other molluscan groups. This problem becomes especially acute with fossil taxa; the extinct groups (indicated with a dagger) discussed below may or may not be molluscs in our current delimitation of the taxon based on living representatives. However, there is little doubt that at some more inclusive grouping these fossil taxa share common ancestors with living molluscan groups.

The converse problem obtains for living taxa. For example, whereas it is possible to relate living taxa to one another using both morphologic and molecular characters, there exists the real possibility that the living taxa do not share a single most recent common ancestor, but may have had multiple, independent derivations from distantly related molluscs or mollusc-like taxa that are now extinct (see below). These alternatives require that both fossils and living taxa be studied and incorporated into evolutionary scenarios and hypotheses of molluscan relationships, especially when the fossil record provides such a wealth of fossils and putative relatives.

Coeloscleritophora†

The worldwide presence of small, hollow, calcareous sclerites in numerous Precambrian and Cambrian sediments (collectively referred to as "small shellies") was an enigmatic component of molluscan evolutionary studies. However, in the early 1990s, an articulated fossil was found in the lower Cambrian of Greenland that was covered with small shellies. It was immediately apparent that what had been thought to be the remains of individual organisms were actually parts of a single larger animal (**Figure 2**). Recent work in the Cambrian of Europe, Asia, and Australia has greatly expanded our knowledge of Coeloscleritophora, and although their relationship to the Mollusca remains uncertain, they likely share a common ancestry with the molluscs as well as with annelids and brachiopods.

Hyolitha†

Hyoliths (**Figure 3**) had cone-shaped shells with a (presumably ventrally) flattened side. The aperture was closed by

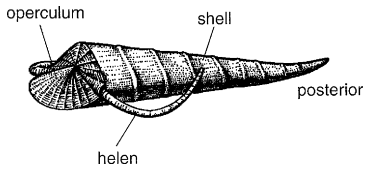


Figure 3 The hyolith, *Haplophrentis carinatus*, middle Cambrian. (Briggs, D. E. G., D. H. Erwin, and F. J. Collier; The Fossils of the Burgess Shale, Copyright © 1991 by Smithsonian Institution. Reprinted by permission of Smithsonian Institution Press.)

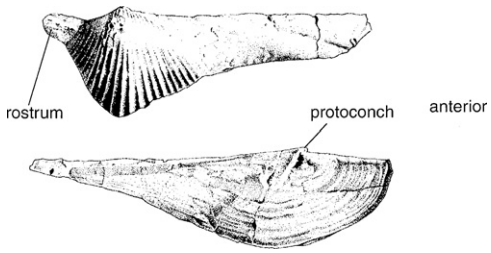


Figure 4 Rostroconchia. Top, *Conocardium* sp., Carboniferous. Bottom, *Pinnocaris* sp. Ordovician.

an operculum, and a single pair of anterior appendages (or Helens) is present in some taxa. The operculum was attached to the shell by pairs of symmetrical muscles. Shells were constructed of crossed lamellar calcium carbonate crystals, and the apical tips bore larval shells. Two forms of larval shells—a smooth globose shell and a point shell with growth lines—perhaps reflected direct and indirect development, respectively. A looped alimentary system connected the ventral mouth to the dorsal anus. Hyoliths were sessile, epifaunal deposit feeders, scrapping sediments from the seafloor. They first appear in the early Cambrian and became extinct at the end of the Paleozoic.

Rostroconchia†

Rostroconchs (**Figure 4**) begin life with a single conch (or valve), but the single shell transforms ontogenetically into a nonhinged bivalved shell that gapes at its margins. Based on muscle scar patterns, the animals appear to have had anterior feeding tentacles that were likely used for deposit or suspension feeding, and both motile and nonmotile forms have been recognized. Rostroconchs first occur in the lower Cambrian and undergo an extensive late Cambrian and early Ordovician radiation. The last known rostroconchs occur in the Permian. Rostroconchs are thought to share common ancestry with the Bivalvia.

Polyplacophora

Polyplacophoran molluscs, or chitons (**Figure 5**), have a dorsal shell that is secreted as eight valves or plates. The plates of individual animals differ in size and shape based on their sequential position; this plate dimorphism is also present in the earliest chitons of the late Cambrian. Although known

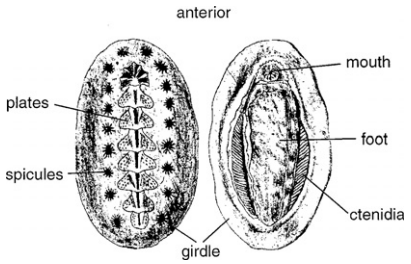


Figure 5 External morphology of the Polyplacophora. Left is dorsal view, right is ventral view. Modified from Gray, Maria; Figures of molluscous animals, selected from various authors. 1857–1859. Reprinted with permission from Gray, Maria.

since the early Paleozoic, chitons do not show a marked increase in diversity until the Cretaceous. Chitons occur worldwide in intertidal habitats and at depths in excess of 7000 m. There are about 850 species and they live on a variety of firm substrates ranging from rocks to algae.

Shell plate morphology has been an important character in determining taxa and relationships. The plates are composed of two distinct shell layers. The outer surface of each plate is high in organic content layer and overlies an inner calcareous layer. Plates may be covered by a thin periostracum. The posterior and anterior edges of each plate have projecting flanges and slits for articulation with other plates, and the outer surfaces are sculptured into distinct regions. The outer plate layer is penetrated by numerous vertical canals that connect to the mantle edge via a canal system that lies between the outer and inner layers. The canals come in two sizes and connect sense organs (including eyes) on the dorsal surface of the valves with the mantle's nervous system.

The plates are surrounded by the girdle, which is formed by a strongly differentiated mantle epidermis. The mantle secretes the plates and surrounds their articulated form. Around the outer plate boundaries the mantle epidermis becomes thicker and secretes a vast array of scales, spines, bristles, and other structures on its dorsal surface. In some species the anterior portion of the mantle epidermis is modified into a large hood that is used to trap food; most species feed by grazing on hard substrates. A poorly defined head and elongated foot that is surrounded by the ventral mantle epidermis of the girdle mark the ventral surface of the animal. The gills are located in the mantle groove that is formed between the girdle and the lateral sides of the foot. The excretory pores and gonoducts also open into this groove.

The visceral mass of the chiton is dominated by a long looped alimentary system. Above the digestive system is a bilobed gonad that appears to result from the fusion of a single pair of gonads. Two separate gonoducts lead from the gonads to opposite sides of the mantle groove. The excretory organs are paired and each connects with the pericardial cavity via a short canal. The center of the pericardial cavity contains an elongated ventricle that is surrounded on either side by enlarged auricles that collect blood from the lateral gills. The polyplacophoran nervous system is ladder-like and almost identical to that of the aplacophorans and monoplacophorans; however, it differs from both by lacking cerebral ganglia. The radular musculature is also similar to that of the monoplacophorans, and the radular tooth configuration is

docoglossate like that of the monoplacophorans and patellogastropod limpets. The typical chiton radula has a weak central tooth that is flanked by strong lateral teeth and weaker marginal teeth, some of which are only plates.

Chiton development has been well studied. The larvae come from relatively large eggs and are nonfeeding. The eggs are covered by a thick hull with elaborate spines and knobs. The chiton embryo hatches as a trochophore, a globular cell mass with an apical tuft and central band of cilia, and subsequently elongates and differentiates into a small chiton. Larval eyes are present but disappear after settlement. The most marked external development is the formation of the foot on the ventral surface of the larva and the formation of transverse bands on the dorsal surface of the larva that mark the sites of valve formation. As the valves form, the apical tuft and prototroch are lost.

Aplacophora

The morphological distance between an aplacophoran mollusc (**Figure 6**) and, for example, a cephalopod mollusc stretches anyone's concept of the rank of phylum. Aplacophora were the last class to be placed in the Mollusca and have until recently been one of the least studied groups. The aplacophoran molluscs include two different morphological groups: the Chaetodermomorpha or Caudofoveata, and the Neomeniomorpha or solenogastres. The Chaetodermomorpha are footless vermiform molluscs that live in sediments. The Neomeniomorpha are more elongate, have a narrow foot, and typically live in association with cnidarians such as hydroids and alcyonaceans. They are found throughout the world and live at depths between 18 and 6000 + m. It is estimated that there are about 250 species of aplacophorans.

The shell of the aplacophoran mollusc is a myriad of calcareous spicules that are secreted by the mantle epidermis. Spicule morphology varies over the body of the aplacophoran, and in some taxa spicules are modified into scales.

The calcareous spicules that cover the bodies of most aplacophorans give the animals a striking sheen. In the Neomeniomorpha there is a pedal groove on the ventral surface of the animal. The pedal groove expands into an anterior pedal pit and contains the narrow foot. The Chaetodermomorpha lack a pedal groove but do have a posterior mantle cavity that contains a gill. In some Chaetodermomorpha there is a slight constriction and distinct change in spicule morphology at the midline of the body.

It is the internal anatomy that provides evidence of the molluscan identity of the aplacophorans. In both groups the anterior end of the alimentary system includes a radula and odontophore. In the Chaetodermomorpha the radula and odontophore are strongly developed, and the alimentary system is more differentiated than in the Neomeniomorpha. Both groups have a dorsal gonad that opens into the pericardium, which contains the heart. From the posterior portion of the pericardium there extends a coelomoduct that loops or bends and ultimately opens into the mantle cavity. In the Neomeniomorpha the posterior portion of the coelomoducts are modified for reproductive functions such as sperm storage or brooding young. The nervous system is ladder-like with a well-developed cerebral ganglion. Radular configurations are quite variable and show a wide range of tooth development and modifications that include jaw-like structures, denticles with cones, and sweepers. This is one of the greatest ranges of radular variation found in a Mollusca, and it stands in marked contrast to the lack of variation in the radular configurations found in both the Monoplacophora and Polyplacophora.

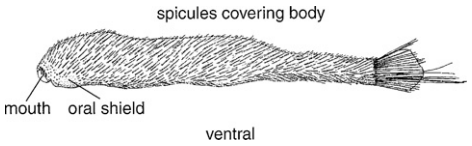


Figure 6 External morphology of a chaetodermomorph Aplacophora. (Redrawn from Scheltema, 1985, Biological Bulletin, 169:496, Figure 3L. Reprinted by permission of the author.)

Development of the aplacophoran mollusc includes a test cell larval stage in which the three tissue types (mesoderm, ectoderm, and endoderm) align and differentiate within an exterior cell layer constructed of large test cells. Aplacophoran eggs are relatively large and free-spawned in the Chaetodermomorpha and fertilized internally in the Neomeniomorpha; some Neomeniomorpha brood their young to various stages of development. After the formation of the test cell larva with an apical tuft and prototroch, the posterior development of the differentiating larva quickly outgrows the constrains of the exterior test and develops directly into the juvenile aplacophoran.

After many years without students, the Aplacophora are now the subject of extensive studies. These studies and interpretations of aplacophoran phylogeny have focused attention on this small group of molluscs. Overall, the aplacophoran body plan is very similar to that of the chiton. Aplacophorans and polyplacophorans differ from the monoplacophorans by having a dorsal gonad rather than a posterior gonad. The pericardium is similar in all three groups as are many of the other organ systems and positions. Major differences are found in the type of shell secreted by the dorsal mantle epidermis.

Bivalvia

Bivalves (**Figure 7**) are laterally compressed molluscs with a hinged, two-part shell. A noncalcified ligament serves as a hinge and connects the two valves along their dorsal surfaces and acts to force the valves apart. The remaining shell surfaces may tightly close or gape with contraction of the musculature. Adult valves can vary greatly in shape and size and certain shell forms appear associated with specific habitats. The shell can be internal and reduced (or even absent) and the bivalve animal worm-like such as in "shipworms" (*Teredo*). Although the first occurrences of taxa referred to the Bivalvia are found in lower Cambrian deposits, it is not until the lower Ordovician that bivalve diversification, both taxonomic and ecological, explodes in the fossil record. This diversification continues unabated through the Phanerozoic, with relatively small losses at the end-Permian and end-Cretaceous extinction events. Today over 10,000 species of bivalves are found in most marine, brackish, and freshwater communities. They

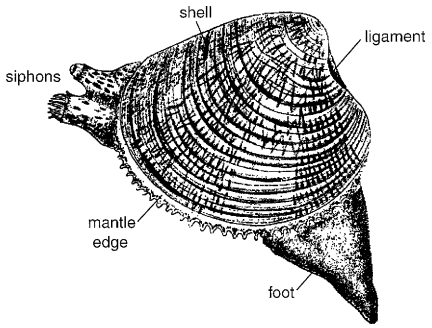


Figure 7 External morphology of the Bivalvia. Modified from Gray, Maria; Figures of molluscous animals, selected from various authors. 1857–1859. Reprinted with permission from Gray, Maria.

may be infaunal or epifaunal, and epifaunal taxa may be either sessile or motile. Bivalves typically display bilateral symmetry both in shell and in anatomy, but there are significant departures from this theme in taxa such as scallops and oysters.

Bivalve shells are constructed of different shell fabrics, including crossed lamellar, nacreous, and foliated microstructures. Most of the variability in shell structure sorts along higher taxon divisions. For example, nacreous structures are present primarily in the basal members of the group (Protobranchia) whereas crown taxa have primarily crossed lamellar shells (Heterobranchia). The outer shell surface is often covered by a periostracum, and the free margins of the shell of some taxa may be uncalcified and flexible. In some taxa such as oysters, one valve is cemented to the substrate. The hinge area often contains an array of sockets and pegs, referred to as "teeth" or hinge dentition. These teeth align the valves as they close and prevent shearing of the valves. The inner surfaces of the valves have scars formed by the adductor shell muscles, typically positioned in the anterior and posterior regions and connected by the pallial line that marks the ventralmost attachment point of the mantle.

The mantle that envelops the bivalve animal forms a large ventral chamber or mantle cavity with paired ctenidia. The mantle edge along the posterior portion of the shell often forms a pair of siphons that facilitate inflow and outflow of water through the mantle cavity. Although the pleisomorphic feeding state for bivalves is likely deposit feeding, the ctendia provide an effective filter-feeding mechanism in most taxa with numerous levels or grades of organization. Bivalves lack a well-defined head and a radula. The foot is often triangular and in some species, such as mussels, the larval byssal gland is retained into the adult and attaches the individual to the substrate with strong fibers.

The visceral mass is primarily situated above the mantle cavity and continues ventrally into the foot. The mouth is often flanked by palps and opens into a stomach. The intestine is irregularly looped and opens dorsally into the excurrent flow near the exhalant siphon. Also opening into this region are the paired kidneys and, when separate from the kidneys, the gonopores of the paired gonads. The heart typically lies below the center of the valves and consists of two auricles and a single ventricle that supplies both anterior and posterior aorta. The nervous system is made up of three pairs of ganglia. These innervate the musculature, mantle, viscera, ctenidia, and siphons. They receive sensory input from statoliths, osphradium, various siphonal sensory structures, and photoreceptors along the mantle margin.

Bivalve development has been well studied, primarily because of the economic importance of bivalves (see below). The eggs are typically small and not very yolkrich. Fertilization is usually external, but in brooding species occurs in the mantle cavity. Cleavage patterns are spiral and both polar lobes and unequal cleavage patterns are present throughout the group. The first larval stage is typically a trochophore that transforms into a veliger. Although morphologically similar to the gastropod veliger stage, phylogenetic analyses suggest that these larval stages are homoplastic rather than homologous. The initial uncalcified shell grows laterally in two distinct lobes to envelop the body. Larval bivalves have a byssal gland that may assist with flotation while planktic but later attaches the juvenile to the substrate.

Monoplacophora

For more than 50 years the Monoplacophora (**Figure 8**) were known only from the fossil record and were recognized by the presence of multiple, symmetrical muscle scars in patelliform shells. The discovery of living monoplacophorans, beginning with *Neopilina galatheae* in 1957, provided anatomical characters for the group. About 20 living species of monoplacophorans have since been discovered worldwide, living at depths between 174 and 6500 m. They are found both on soft bottoms and on hard substrates on the continental shelf and seamounts. Paleozoic taxa are associated with relatively shallow water faunas (>100 m).

In Recent and fossil patelliform monoplacophoran shells, the apex is typically positioned at the anterior end of the shell, and in some species actually overhangs the anterior edge of the shell. Aperture shapes vary from almost circular to pear-shaped. Shell height is also variable and ranges from relatively flat to tall. Shell sculpture varies from rugose concentric ridges to fine growth lines. In some taxa radial sculpture produces a fine reticulate pattern.

The monoplacophoran animal has a poorly defined head with an elaborate mouth structure on the ventral surface. The mouth is typically surrounded by a V-shaped, thickened anterior lip and postoral tentacles; postoral tentacles come in a variety of morphologies and configurations. Below the head

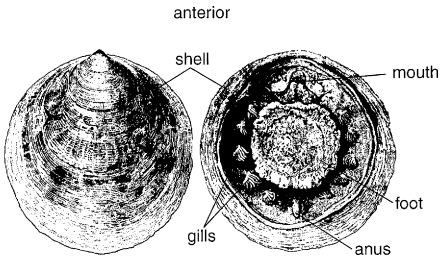


Figure 8 External morphology of the Monoplacophora. Left is dorsal view, right is ventral view. (Reprinted by permission from Nature (179:413–416), copyright 1957 Macmillan Magazines Ltd.)

lies the semicircular foot. In the mantle groove, between the lateral sides of the foot and the ventral mantle edge, are found five or six pairs of gills (less in minute taxa).

Internally, the monoplacophoran is organized with a long, looped alimentary system, two pairs of gonads, and multiple paired excretory organs (four of which also serve as gonoducts). A bilobed ventricle lies on either side of the rectum and is connected via a long aorta to a complex plumbing of multiple paired atria that in turn are connected to the excretory organs. The nervous system is ladder-like and has weakly developed anterior ganglia; paired muscle bundles enclose the visceral mass. The most interesting anatomical structures in the original description of a living monoplacophoran were the paired "dorsal coeloms" that are connected to the anterior excretory organs and are topographically similar to the fused gonads in chitons. Later studies suggest that these structures are more likely extensions of glands associated with the pharynx. The monoplacophoran radula is docoglossate: it has a rachidian tooth, three pairs of lateral teeth, and two pairs of marginal teeth. The lateral teeth are in a stepped configuration: the first and second pairs of lateral teeth are aligned with the rachidian tooth, and the third lateral pair is slightly posterior and lateral to the first two pairs. The cusp of the innermost marginal tooth is frilled.

Developmental studies of monoplacophorans have not been done.

Recent monoplacophorans form a distinct clade, and their similarities and differences with the other extant molluscan groups are easily recognized. There is little question that some Paleozoic taxa also are members of this clade. However, the characters that distinguish some Paleozoic monoplacophorans from torted molluscs (i.e., the Gastropoda) and vice versa are open to alternative interpretations, and the relationships of several major groups of early-shelled molluscs have therefore been the subject of much lively debate.

Scaphopoda

The scaphopod mollusc (**Figure 9**) combines gastropod-like shell morphology with bivalve-like development with cephalopod-like anatomy. Members of the class first appear in the early Paleozoic and the taxon has maintained a slow and steady rate of increase in morphological diversification since then. Our knowledge of the biology of recent species comes primarily from members of a single genus, *Dentalium*. Scaphopods are infaunal organisms and feed on foraminiferans and other interstitial prey. Approximately 350 species occur from the intertidal zone to depths in excess of 7000 m and are present in all the major oceans.

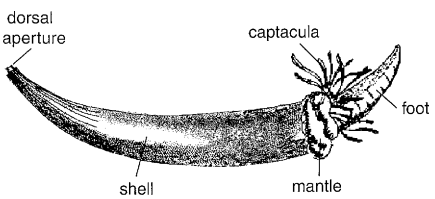


Figure 9 External morphology of the Scaphopoda. Modified from Gray, Maria; Figures of molluscous animals, selected from various authors. 1857–1859. Reprinted with permission from Gray, Maria.

The scaphopod shell is a calcium carbonate tube with equal or unequal apertures; the tube may be either inflated or bowed. The shell microstructure includes prismatic and crossed-lamellar components; the latter is similar in structure to elements seen in members of the Bivalvia.

The head and foot extend through the ventral aperture of the scaphopod shell. The head bears numerous tentacles (captaculae) that are used to capture and manipulate prey. Foot morphology is variable and has been used as a taxonomic character. The pattern of water circulation through the scaphopod mantle cavity is also unique among the molluscs, because water both enters and exits through the small dorsal aperture. Scaphopods also lack gills.

Unlike the previously discussed groups, scaphopods have a U-shaped gut rather than an anterior–posterior configuration of the mouth and anus. The stomach and digestive gland are in juxtaposition, and the intestine loops before passing through the excretory organ and opening into the mantle cavity. The posterior portion of the digestive gland overlies the gonad that connects with the mantle cavity via the excretory organ. There may be a reduced pericardium, but a heart is absent. The radula consists of a rachidian plate, a single first lateral tooth, and a lateral plate.

The ontogeny of several species has been documented. The trochophore larva has an apical tuft and prototroch and develops mantle folds that divide the larva into equal lateral halves. A velum forms from the prototroch on the anterior end and the dorsal and lateral sides of the larvae begin to secrete a shell that fuses along the ventral surface of the larva but leaves an opening at the posterior end. The trilobate foot forms ventrally under the velum, and the larva attains the tubular shell morphology of the adult scaphopod. The loss of the velum and completion of organogensis after settling complete metamorphosis.

Scaphopods have an intriguing set of molluscan characters that have been allied to several scenarios of molluscan evolution and relationships. Their shell structure and development suggest bivalvian affinities, but scaphopods also have a radula. The gross morphology of the scaphopod gut is U-shaped, like that of the gastropods and cephalopods, rather than linear as in monoplacophorans, polyplacophorans, and aplacophorans. However, like monoplacophorans, scaphopods have a gonad that lies under the alimentary system rather than on top of it as in aplacophorans and polyplacophorans. Some workers have suggested that aplacophoran larvae and scaphopod larvae share some features; however, others are unconvinced by these arguments and compare the scaphopod velum to that of higher bivalves.

It has been suggested that scaphopods are descended from ribeirid rostroconchs and they have been grouped with the Bivalvia. Though there is little doubt that scaphopods share some characters with the Bivalvia, the direct derivation of the scaphopod mollusc from a ribeirid rostroconch is contradicted by the U-shaped gut present in scaphopods. Rostroconchs are thought to have had a linear gut. Thus, the transition from the ribeirid condition to a scaphopod morphology would include the construction of a U-shaped alimentary system from a linear one, although the ancestral linear gut would have been an ideal

exaptation for a scaphopod shell morphology open at both ends and with flow-through water circulation. Like many other characters, novelties and confusing patterns are common in molluscan morphology but so are convergent grades (especially in molluscan shell form).

Gastropoda

Externally, gastropods (**Figure 10**) appear to be bilaterally symmetrical; however, they are one of the most successful clades of asymmetric organisms known. The ancestral state of this group is clearly bilateral symmetry (e.g., chitons, cephalopods, bivalves, see above), but gastropod molluscs twist their organ systems into figure-eights, differentially develop or lose organs on either side of their midline, and generate shells that coil to the right or left. And although they have left the more common molluscan state of bilateral symmetry, gastropod molluscs are one of the most diverse groups of animals, both in form and habitat. They occupy habitats ranging from the deepest ocean basins to the highest mountains and from the tropics to high latitudes. Estimates of total extant species range from 40,000 to over 100,000, but may be as high as 150,000, with about 13,000 named genera for both Recent and fossil species. They have figured prominently in paleobiological and biological studies, and have served as study organisms in numerous evolutionary, biomechanical, ecological, physiological, and behavioral investigations. They have a long and rich fossil record that shows periodic extinctions of subclades, followed by diversification of new groups.

The best documented source of gastropod asymmetry is the developmental process known as torsion. Like other molluscs, gastropods pass through a trochophore stage, and then form a characteristic stage of development known as the veliger. During the veliger stage a 180° rotation of the mantle cavity from posterior to anterior places the anus, and renal openings over the head, and twists organ systems that pass through the snail's 'waist' (the area between the foot and visceral mass) into a figure eight. This rotation is accomplished by a combination of differential growth and muscular contraction. In some taxa the contribution of each process is about 50:50, but in other taxa the entire rotation is accomplished by differential growth. Although the results of torsion are the best known asymmetries in gastropods, numerous other asymmetries appear independent of the torsion process. Anopedal flexure, sometimes considered a feature of torsion, is widely distributed in the Mollusca; it is present in the extinct hyoliths as well as in the Scaphopoda and Cephalopoda (and to a lesser extent in the Bivalvia).

Trends in gastropod evolution often feature a reduction in the complexity of many characters. These include reduction of the number of radular teeth, simplification (thought to be due to shell coiling) of the reno-pericardial system (loss of right auricle and renal organ), reduction of ctenidia (loss of the right gill), and associated circulatory and nervous system changes. There is also a reduction of diversity of shell microstructures, simplification of the buccal cartilages and muscles, reduced coiling of hindgut, and simplification of the stomach. Other characters show an increase in complexity, such as life history characters (e.g., internal fertilization with penis and spermatophores and associated reproductive organs). This increase in complexity is correlated with the ability to produce egg capsules and the evolution of planktotrophic larvae and direct development. There is also an increase in chromosome number, and greater complexity of sensory structures (e.g., eyes, osphradium). In the pulmonates (land snails) the pallial cavity is modified into a pulmonary cavity or lung, while in the opisthobranchs (sea slugs) there are secondary gills and elaborate neurosecretory structures.

Generally, gastropods have separate sexes. One major exception is the heterobranch clade (pulmonates & opisthobranchs), which is exclusively hermaphroditic. Basal gastropods typically spawn gametes directly into the water column, where they are fertilized and develop. Brooding of developing embryos is widely distributed throughout the gastropods, as are sporadic occurrences of hermaphrodism. These basal groups also have non-feeding larvae. Eggs are relatively small in most taxa, although eggs of taxa with non-feeding larvae tend to be a little larger than those taxa that have feeding larvae. Egg size is reflected in the initial size of the juvenile shell or protoconch: when preserved in fossil taxa, this feature has been useful in distinguishing feeding and non-feeding taxa in the fossil record. The first gastropod larval stage is typically a trochophore that transforms into a veliger and then settles and undergoes metamorphosis to form a juvenile snail.

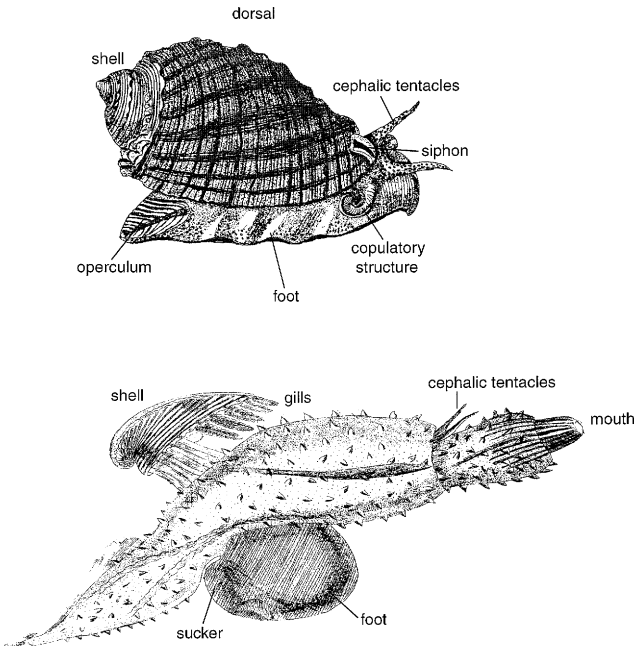


Figure 10 External morphology and diversity of the Gastropoda. Top, typical benthic gastropod mollusc. Bottom, a pelagic heteropod. Modified from Gray, Maria; Figures of molluscous animals, selected from various authors. 1857–1859. Reprinted with permission from Gray, Maria.

Cephalopoda

Cephalopods (**Figure 11**) are dorsoventrally elongated marine molluscs that may or may not have a recognizable external or internal shell. Cephalopods are the most complex and motile

of the nonvertebrate metazoans and show numerous modifications of the general molluscan body plan. The nautiloids first appear in the late Cambrian and undergo a rapid diversification in the Ordovician. Ammonites (Coleoidea) appear later in the Devonian. Cephalopods are much more variable in their diversity through time than other molluscan groups. They are hit by numerous extinctions (e.g., terminal Permian, Triassic, and Cretaceous events) but typically showed rapid replacement by the survivors. The cephalopods include the largest living as well as the largest extinct molluscs: ammonite shells range to over 2 m across and body sizes of living squid range up to 8 m, with tentacles exceeding 20 m in length. Over 10,000 species are known with about 700 living taxa. All are active carnivores in marine benthic and pelagic habitats from nearshore to abyssal depths.

Cephalopods are thought to have evolved from monoplacophoran-like ancestors. Septa formed at the apex as the animal grew and withdrew into a newly formed body chamber. The old chambers are gas-filled and provide buoyancy for the organism. The foot was modified into a funnel and provides jet propulsion for movement. Prehensile arms with suckers surround the mouth on the head (cephalic in origin) and capture prey. In cephalopods with external shells (ammonites and nautiloids), the shells are composed of an inner nacreous layer and an outer prismatic layer. In other cephalopod taxa the shell is typically internal and reduced to a linear stiffener as in cuttlefish and squid or virtually nonexistent as in the octopus, whose pelagic ancestors reinvaded the benthic realm.

Much of the external anatomy of cephalopods appears associated with their highly motile, pelagic habits. These modifications are so strong that it is difficult to readily identify the typically molluscan body axes in a cephalopod. For example, as a squid jets through the water, the most posterior portion of the body, relative to the direction of movement, is the squid's head with its prehensile arms. The most anterior portion is the dorsal surface of the muscular mantle and the posterior mantle cavity is located ventrally. In addition to jets of water expelled through the funnel, squid and some octopi use undulating movements of paired fins at the distal end of the mantle for swimming as well.

The cephalopod mouth contains strong beak-like jaws that are used to deliver a lethal bite to prey items as well as shred them; a radula is also present. Salivary glands can produce highly toxic venoms in some squids and octopi, and ink sacs are present in most taxa except *Nautilus*. The digestive system differs from that of most other molluscs in having a cecum in juxtaposition with the stomach and a relatively short intestinal tract. There are two pairs of ctenidia in living *Nautilus*, but all remaining cephalopods (cuttlefish, squid, and octopus) have a single pair of ctenidia. The circulation system also differs markedly from that of other mollusks in being a closed system capable of maintaining a high blood pressure, rather than open and diffuse as in other taxa. The heart consists of two auricles and a single ventricle that supplies multiple arteries. In all cephalopods except *Nautilus*, a pair of branchial hearts situated at the base of the ctenidia pump blood through the ctenidia. The nervous system is highly concentrated and developed in cephalopods. The three major ganglia are fused and organized into lobes, each with its own specific function. *Nautilus* has fewer lobes than other cephalopods, and lobe size may vary among groups with different lifestyles. Coleoid cephalopods also have two large stellate ganglia on the mantle that control both respiratory and locomotory functions of the mantle. Sensory structures include statocysts, olfactory organs, and eyes. Coleoid eyes are surprisingly convergent with vertebrate systems and are capable of resolving brightness, shape, size, and orientation. Cephalopods are also able to change color patterns using an elaborate chromatophore system that is under nervous control; it provides them with the ability to display incredible camouflaging.

Cephalopods have separate sexes and spermatophores are transferred between males and females by modified tentacles. Eggs are large and yolk-rich, and embryonic development of cephalopods is different from that of all other mollusks. There is no larval form, just direct development into juveniles. Both the eggs and young may be brooded, benthic, or pelagic.

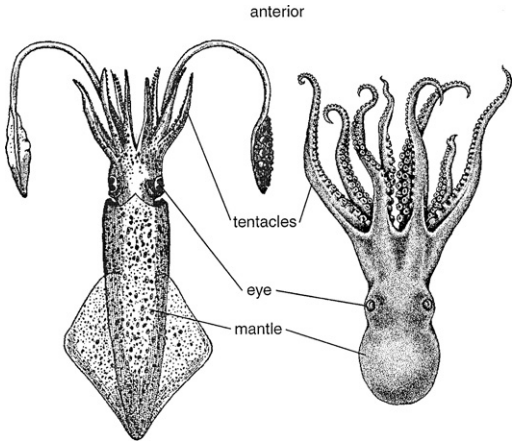


Figure 11 External morphology and diversity of the Cephalopoda. Left, squid; right, octopus. (Octopus redrawn from Beesley, P. L., G. J. B. Ross, and A. Wells, Mollusca: The Southern Synthesis. Part B. Fauna of Australia, Vol. 5. Copyright © 1998 by Commonwealth of Australia. Reprinted by permission of Australian Biological Resource Study.)

Ecology

Molluscs occur on a large variety of substrates, including rocky shores, coral reefs, mudflats, and sandy beaches. Gastropods and chitons are characteristic of these hard substrates; bivalves are commonly associated with softer substrates, where they burrow into the sediment. However, these patterns are not inflexible. The largest living bivalve, *Tridacna gigas*, occurs on coral reefs, whereas microscopic gastropods live interstitially between sand grains. On hard substrates gastropods are often grazers that feed either selectively or indiscriminately on algal, diatom, or blue-green films and mats or animal aggregations. On soft substrates bivalves typically are suspension or deposit feeders.

The adoption of different feeding habits appears to have had a profound influence on molluscan evolution. The change from grazing to other forms of food acquisition is one of the major features in the adaptive radiation of the group. Based on our current understanding of relationships, the earliest molluscs were likely carnivores and detritovores that indiscriminately grazed on encrusting animals and detritus. Truly

herbivorous grazers are relatively rare in the molluscs and are limited to the polyplacophorans and gastropod groups, whereas living aplacophorans, monoplacophorans, scaphopods, and cephalopods are carnivorous. Some protobranch bivalves are also carnivorous; most bivalves are either suspension or deposit feeders that indiscriminately take in particles but then elaborately sort them based on size.

Cephalopods are typically active carnivores specialized on mobile prey such as fish, crustaceans, and other cephalopods. Because they are so abundant in pelagic systems, cephalopods are often important food sources for larger fishes, marine mammals, and even seabirds. In the gastropods several groups such as Janthinidae are planktic pelagic carnivores whereas the heteropods (Caeongastropoda) and the pteropods (Opisthobranchia), like the cephalopods, are active swimmers in search of prey. These taxa spend their entire lives in the water column where they primarily feed on Cnidaria, other molluscs (including small cephalopods), and even fishes.

Bivalve-like suspension feeding has evolved in some gastropods such as the vetigastropod *Umbonium* and in the pelagic gastropod group Thecosomata. Some groups with carnivorous diets have undergone what appear to be true, explosive, adaptive radiations (e.g., the Neogastropoda). Other carnivorous taxa such as the aplacophorans and scaphopods have low diversity and abundance.

In addition to these more typical trophic strategies and interactions, some molluscs (especially among the gastropods) have also become both endo- and ectoparastic in and on other invertebrates—most often echinoderms. Galeommatoid bivalves and eulimid gastropods occur as ectoparasites on holothuroid and asteroid echinoderms, respectively. The reduced, worm-like Entoconchidae occur in the internal body cavity of holothuroids, whereas the larvae of freshwater unionid bivalves parasitize fish and amphibians, although the adults are free-living.

Molluscan groups are ubiquitous and diverse in marine habitats, but only the bivalves and gastropods have invaded freshwater habitats, and only gastropods have invaded terrestrial ones. The major terrestrial clade is the pulmonates, which originated at least by the Carboniferous, but other nonmarine groups such as the Neritopsina (marine, freshwater, and terrestrial) are likely Devonian in origin. Often the terrestrial groups are among the most basal of the extant taxa in the clade. For example, in the Neritopsina terrestrial taxa are thought to be more basal than marine members of the group. These patterns could result from competition among sister taxa and the relegation of one taxon to a unique habitat while the other diversified in the ancestral setting.

Shell morphology is often thought to be correlated with lifestyle and habitat and some substantial changes in body form are clearly associated with major adaptive changes. However, in many cases gross differences in morphology within a group are not readily correlated with habitat. Conversely, similar shell morphologies do not necessarily indicate similar habits or habitats, making paleocological studies with molluscs problematic. For example, limpets occur on wave-swept platforms, on various substrates in the deep sea, in fast-flowing rivers, or in quiet lakes and ponds. It is often suggested that strong wave action selects for limpet morphologies in several groups, but it is obvious from their current distributions that limpets do very well in a wide range of habitats.

Molluscs and Humans

Molluscs and humans are most often associated economically. Molluscs have many important commercial benefits such as fisheries and mariculture but can also be responsible for tremendous economic loss and human suffering. Molluscs are found in some of the earliest human habitation sites in southern Africa over 100,000 years ago, and it is likely that humans have included molluscs in their diet and as material resources for millions of years. More recently, commercial fisheries have focused on bivalve, gastropod, and cephalopod taxa for food (oysters, abalone, scallops, clams, cephalopods, and mussels) as well as pearl and button fisheries.

In freshwater habitats molluscs may serve as intermediate hosts for parasites such as digenetic trematodes and are an integral part of the Schistosomiasis contagion that infects millions in the tropical regions of the world. Nonhuman, molluscan-borne parasites are also responsible for more than $2 billion worth of losses worldwide to the livestock industry each year. Most people are familiar with "garden snails" and the damage they can do to favorite plants, but this local damage is minimal compared to the role of molluscs as pests. Molluscan pest species cause millions of dollars worth of damage to horticultural and food crops throughout the world every year, and as the global economy expands, the probability of potentially damaging introductions increases dramatically. For example, over the past 5 years, almost 5000 molluscs from 100 different countries were intercepted entering the United States.

One of the most notable introductions has been that of the freshwater bivalve *Dreissena polymorpha*, or zebra mussel. Zebra mussels were first detected in Lake St. Clair (Great Lakes chain) in 1988. It is thought that zebra mussels entered the Great Lakes in ballast water of commercial ships from eastern Europe, and without natural enemies to control their numbers, they have thrived and spread to 19 states in less than 10 years. In the United States and Canada, they have caused widespread economic and environmental damage. They typically concentrate where water flows rapidly and therefore can quickly clog intake pipes and cut off water flow to hydroelectric and nuclear power plants, industrial facilities, and public water supplies. One mussel removal project in Lake Michigan removed 400 cubic yards of zebra mussels at a cost of $1.4 million. Zebra mussels can also have tremendous effects on freshwater ecosystems by displacing native species and increasing levels of toxic algae. As zebra mussels have spread across the country, they have devastated native freshwater clam populations (especially the unionid species) that were already endangered and threatened by pollution and habitat modification.

On a more positive note, molluscs are also important in biomedical and biotechnology research. Although not as well known as model organisms such as the fruit fly *Drosophila*, opisthobanch molluscs are commonly used in neurological studies of learning and memory functions, and bivalve molluscs have served as models for environmental and genetic

interactions in tumor research. Extracts of numerous other molluscs have been examined for antibacterial and other properties potentially useful in the medical and health sciences. Molluscs also provide important model organisms for environmental health studies. For example, bivalve molluscs are frequently used in studies on the toxicology and environmental monitoring of heavy metals and other pollutants.

Molluscan conservation issues are primarily centered on freshwater and terrestrial taxa. The largest number of documented extinctions has involved nonmarine molluscs, and nonmarine molluscs are second to the arthropods in numbers of currently threatened species. Terrestrial species are impacted primarily by the destruction of native vegetation and cover, and freshwater species by alter flow, sedimentation rates, and pollution. Although marine taxa are thought to be less impacted by human activities, populations of some taxa have been substantially reduced by overfishing. These include abalone (*Haliotis*), clams (*Tridacna*), and scallops (*Pecten*). Marine taxa, like freshwater and terrestrial taxa, are also threatened by habitat destruction, pollution, changes in sedimentation patterns, and other factors.

As discussed above regarding zebra mussels, the introduction of exotic species can also seriously impact native taxa. This is especially true on islands where the effective population size may be small. One of the best-documented cases was the apparent extinction of the endemic land snail taxon *Partula* from the island of Moorea in French Polynesia. This resulted because of the intentional introduction of the predatory snail *Euglandina rosea* to the island as a biocontrol agent for the Giant African Snail, *Achatina fulica*.

Classification

Given the state of flux in molluscan and metazoan phylogeny and the wealth of new data that is now appearing from molecular, morphological, and paleontological work, any classification proposed here would be rapidly outdated. Several traditional classifications are available in the references cited in the References. However, few are based on hypotheses of relationships, but are instead based on overall similarity and ad hoc scenarios of evolution.

Classifications based solely on morphology have been especially problematic, and much of this confusion has resulted from problematic taxa such as the aplacophorans, scaphopods, and bivalves where possible reduction and loss of organs or other secondary simplification has produced morphologies that may be argued as either primitive or highly derived. Many classifications have also focused exclusively on the morphology of living taxa and have ignored potential, fossil members of the Mollusca. If extinct fossil taxa are included in evolutionary scenarios, they are typically limited to distinctive clades such as the Rostroconchia and Bellerophonta. Other more problematic extinct taxa (e.g., hyoliths) are systematically ignored, arbitrarily excluded from the Mollusca without analysis, or shoehorned into extant groups.

The sister taxa of Mollusca have included the Platyhelminthes, Annelida, Sipuncula, and the Kamptozoa. Within the Mollusca both Polyplacophora and Aplacophora have been argued as the most primitive taxon, and thus the outgroup to all Conchifera.

Most classifications have assumed a single cladogenetic event in the origin of the Conchifera from the supposedly more primitive placophoran groups. Alternative hypotheses have derived the conchiferans in an unresolved polytomy from a hypothetical ancestral mollusc, or HAM. Some workers have interpreted the Cambrian Burgess Shale taxon *Wiwaxia* and other less complete halkieriid-like fossils as molluscan while others have argued *Wiwaxia* to have annelid worm affinities. However, the discovery of an articulated halkieriid from the lower Cambrian and the existence of these and other multishelled placophorans necessitate the reexamination of long-held assumptions of molluscan ancestry and monophyly. The rapidly increasing knowledge of Coeloscieritophora diversity suggests that they may harbor independent ancestors for extant molluscan groups.

Molecular phylogenies for the Mollusca have not fared much better than the morphological studies. Nuclear and mitochondrial DNA sequences have had very limited success in resolving a monophyletic molluscan clade or even producing robust or reasonable groupings within the Mollusca (e.g., the bivalves and gastropods). These problems most likely result because of the deep, Paleozoic divergence of many of the molluscan taxa and the variable rates of change across molluscan genomes.

The following rank-free classification is conservative and only denotes the major clades within the Mollusca.

- Coeloscleritophora
- Sipuncula
- Hyolitha
- Mollusca
  - Chaetodermomorpha
  - Neomeniomorpha
  - Polyplacophora
  - Conchifera
    - Rostroconchia
    - Bivalvia
      - Protobranchia
      - Pteriomorpha
      - Heterodonta
      - Anomalodesmata
    - Monoplacophora
    - Scaphopoda
      - Dentaliida
      - Gadilida
    - Gastropoda + Cephalopoda
      - Gastropoda
        - Patellogastropoda
        - Vetigastropoda
        - Neritopsina
        - Caenogastropoda
        - Heterobranchia
      - Cephalopoda
        - Nautiloidea
        - Coleoidea
          - Aminonoida
          - Decapoda
          - Octopoda

See also: Invertebrates, Freshwater, Overview. Invertebrates, Marine, Overview. Invertebrates, Terrestrial, Overview

References

Beesley PL, Ross GJB, and Wells A (eds.) (1998) *Mollusca: The Southern Synthesis. Fauna of Australia,* Parts A and B, Vol. 5. Melbourne: CSIRO Publishing.

Broadhead TW (ed.) (1985) *Mollusks. Notes for a Short Course* (Studies in Geology 13). Department of Geological Sciences. Knoxville, TN: Univ. of Tennessee.

Davis GM (ed.) (1999) *Interactions between man and molluscs. Malacologia* 41(2): 319–509.

Johnson PA and Haggart JW (eds.) (1998) *Bivalves: An Eon of Evolution—Paleobiological Studies Honoring Norman D. Newell.* Calgary, AB: Univ. Calgary Press.

Moore RC (ed.) (1960) *Treatise of Invertebrate Paleontology. The Mollusca. Part I. Mollusca 1.* Boulder, CO: Univ. Press of Kansas, Lawrence, and Geol. Soc. Am.

Ponder WF (ed.) (1988) *Prosobranch Phylogeny. Malacol. Rev.* Suppl. 4.

Ponder WF and Lindberg DR. Towards a phylogeny of gastropod molluscs—A preliminary analysis using morphological characters. *Zool. J. Linnean Soc.* 119: 83–265.

Taylor J (ed.) (1996) *Origin and Evolutionary Radiation of the Mollusca.* Oxford: Oxford Univ. Press.

Moths

David L Wagner, University of Connecticut, Storrs, CT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aposematic Warningly colored; boldly colored, usually involving reds, oranges, or yellows, as well as black and white; less commonly with only black and white pattern elements.

Batesian mimicry Evolutionary phenomenon whereby a palatable species comes to resemble a distasteful, toxic, or otherwise protected species and thereby gains some protection from predators.

Congener A member of a shared genus.

Crochets Hook-like structures at the ends of abdominal prolegs for engaging silk and other substrates.

Detritivore An animal (larva in this article) that feeds on dead organic debris; although the term covers both plant or animal matter, in moths the term is especially apt to apply to leaf litter feeding.

Diapauses Period of hormonally controlled inactivity generally associated with weather conditions that are unfavorable to survival or reproduction. Often induced by shortening day lengths in late summer and fall, and broken by warm temperatures in spring.

Hemolymph Translucent yellow to green fluid that fills an insect's body, sharing functions of blood and lymphic systems, but differing in that it does not transport oxygen in most insects.

Hypermetamorphic A life cycle that includes two or more larval forms, with each often specialized for a different feeding function (and larval ecology).

Instar A larval stage; the first instar hatches from the egg and on molting enters the second instar. Most moths undergo five or six instars prior to pupation; as few as three and more than a dozen occur in various lineages.

Maxillae The second pair of mouthparts, located between the mandibles and the labium; the third pair of mouthparts.

Monophyletic A group with a single evolutionary origin that includes the common ancestor and all of that ancestor's descendants.

Mullerian mimicry Evolutionary phenomenon whereby distasteful, toxic, or otherwise protected species come to resemble one another. In so doing the members of Mullerian mimicry complex gain by more efficiently educating local predators.

Natural group Monophyletic; group with a common origin (i.e., having a shared common ancestor).

Oviposit To lay an egg or ovum.

Parasitoid Predator that lives internally or externally on its host. Parasitoids are parasite-like in that they are smaller than their hosts, feed from within or on the host's body, and often do so over a period of weeks or months, but functionally they are predators because they nearly always kill their host (prey).

Pharate A "cloaked" or hidden stage, for example, the adult moth prior to its emergence from the pupa (which in some noctuid moths can be delayed by more than 7 months).

Pheromone A chemical released by one individual that elicits a response in a second individual of the same species.

Polyphagous Eating plants from more than two unrelated plant families.

Introduction

Every night at dusk one and a half million Mexican free-tailed bats flood out from beneath the Congress Avenue Bridge in Austin, Texas. Bats from this single colony harvest more than 10,000 pounds of insects every evening, most of which are moths. Moth caterpillars account for much of the above-ground insect biomass in temperate forests; without them one wonders whether there would be songbirds to usher in each spring. Estimates of species-richness for moths have climbed steadily over the past decade, with the more modest extrapolations falling between 200,000 and 300,000 species. Whether measured in terms of biomass, influence on terrestrial ecosystems, or species-richness, moths must be held as one of the most successful lineages of macroscopic organisms on this planet.

Body sizes span two orders of magnitude: nepticulid and heliozelid leaf miners, some with wingspans under 3 mm, are little more than aerial plankton (**Figure 1**). The world's largest moth, a neotropical erebid, has a wingspan that sometimes exceeds 275 mm. All members of the order Lepidoptera can be diagnosed by the presence of scales, which cover the wings and body. A second readily observable structure that is part of the order's ground plan is the epiphysis, a cuticular flap on the inside of the foreleg, which is used by many moths to clean the antenna. A curious lepidopteran attribute is the production of apyrene (anucleate) sperm – in some species 80% or more of the sperm is apyrene and therefore incapable of fertilization. Kristensen (1984) identified nearly two dozen other uniquely derived structures, which, taken as a whole, incontrovertibly established the monophyly of the Lepidoptera and its sister-group relationship with the Trichoptera (caddisflies). Recent molecular studies have corroborated the major findings of Kristensen.

Butterflies and moths make up the order Lepidoptera. Because butterflies derive from within one (or two) moth lineages, it is not possible to identify uniquely evolved features shared by butterflies that are not also found in some moths. Recent phylogenetic reconstructions indicate that butterflies have their closest relatives with the Hedylidae, an odd-looking

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00171-4

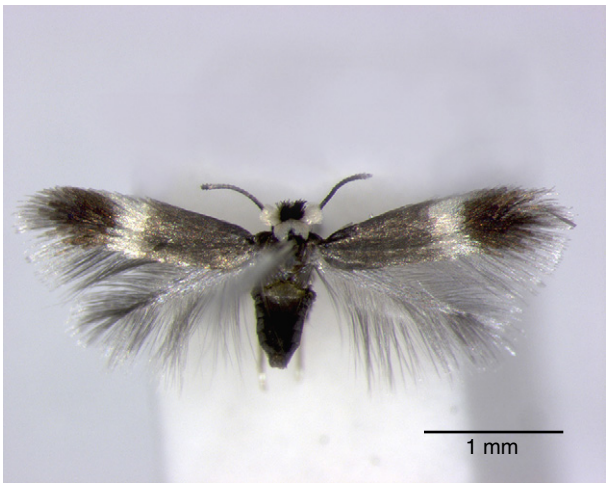


Figure 1 *Stigmella aceris* (Nepticulidae); adults have a wingspan of only 4 mm. Note most of the wing surface area is made up of cilia-like scales. Courtesy Erik Nieukerken.



Figure 2 *Macrosoma nigramacula* (Hedylidae); both morphological and molecular place these moths as a subgrouping within the butterflies (Papilionoidea and Hesperioidea). Courtesy Arthur Anker (UFC, Brazil).

group of neotropical moths that were formerly classified in the Geometroidea (**Figure 2**). Three independent molecular studies place hedylids within the butterflies – stated differently, some butterflies are more closely related to hedylid moths than they are to other butterflies. But regardless of the phylogenetic position within the order, since butterflies are, evolutionarily speaking, nothing more than diurnal moths (and negligible in terms of species richness), "Lepidoptera" and "moths" are used synonymously throughout this article.

Higher Classification

Phylogenetic relationships among the early lineages of Lepidoptera are among the best known of any insect order, so well that they have even been used as a "truth table" for testing the phylogenetic signal in several gene sequencing studies. The basal branches on which the order's extraordinary evolutionary successes have been built are shown in **Figure 3**.

The Four Suborders: The Base of the Tree

Recent higher classifications for the Lepidoptera recognizes four lepidopteran suborders: Zeugloptera, Aglossata, Heterobathmiina, and Glossata (Kristensen, 1998, and authors therein; van Nieukerken *et al.*, 2011; **Figure 3** and **Table 1**). The first and most basal lineage within the order contains a single family, the Micropterigidae, with approximately 250 known species (some 90 of which are still undescribed) distributed across the major continents. Decidedly primitive, this family shares numerous morphological traits with the Trichoptera (caddisflies). Noteworthy are the functional mandibles in the adult (homologous with those of other chewing insects) that are used to grind up pollen and fern spores. The larvae – largely bryophyte feeders – are bizarre creatures with little morphological affinity to other lepidopterans (**Figure 4(a)**); so odd are the caterpillars that at one time these moths were placed in their own order. The spore and bryophyte feeding by adults and larvae, respectively, reflect the family's archaic origins. The Aglossata contains a single *Araucaria*-feeding family the Agathiphagidae with two species, restricted to Australia, New Caledonia, and a few Pacific islands. The third suborder, the Heterobathmiina, is represented by about nine species from temperate areas of Chile and Argentina. So far as is known, all heterobathmiids feed on *Nothofagus* (southern beech), an ancient group of woody angiosperms. Hence, all three of these early lineages have life histories that may agree in detail with those of their earliest Mesozoic ancestors. The remaining diversity occurs in the Suborder Glossata, which possess, among other traits, a coiled proboscis that can be used to siphon up nectar and other liquid substrates. The Glossata includes approximately 41 superfamilies, all but two of which (Hesperioidea and Papilionoidea) are moths.

The Glossatans

Four major clades are recognized within the Glossata: Dacnonypha, Coelolepida, Myoglossata (Neopseusida), and Neolepidoptera. The most basal lineage within this assemblage is the Dacnonypha, with one family, the Eriocraniidae. Eriocraniids are small handsome moths with metallic gold and purple scales known only from the Northern Hemisphere. The two dozen recognized species are leaf miners in new leaves of Fagaceae, Betulaceae, and Rosaceae. Coelolepida include the Acanthopteroctetidae and Australia's lophocoronids. The biology of the Acanthopteroctetidae is poorly known – one is a miner in *Ceanothus* (Rhamnaceae). Likewise, the life history of all six lophocoronids remains unknown: the piercing ovipositor is suggestive that the female lays the eggs within tissues and that the larvae are internal feeders. Neopseustids are peculiar insects that superficially resemble lacewings and other neuopterans. Larval stages are unknown, but again the shape of the female's ovipositor is suggestive that larvae feed internally; the 14 species are found in southern South America, China, and Southeast Asia. The last clade, the

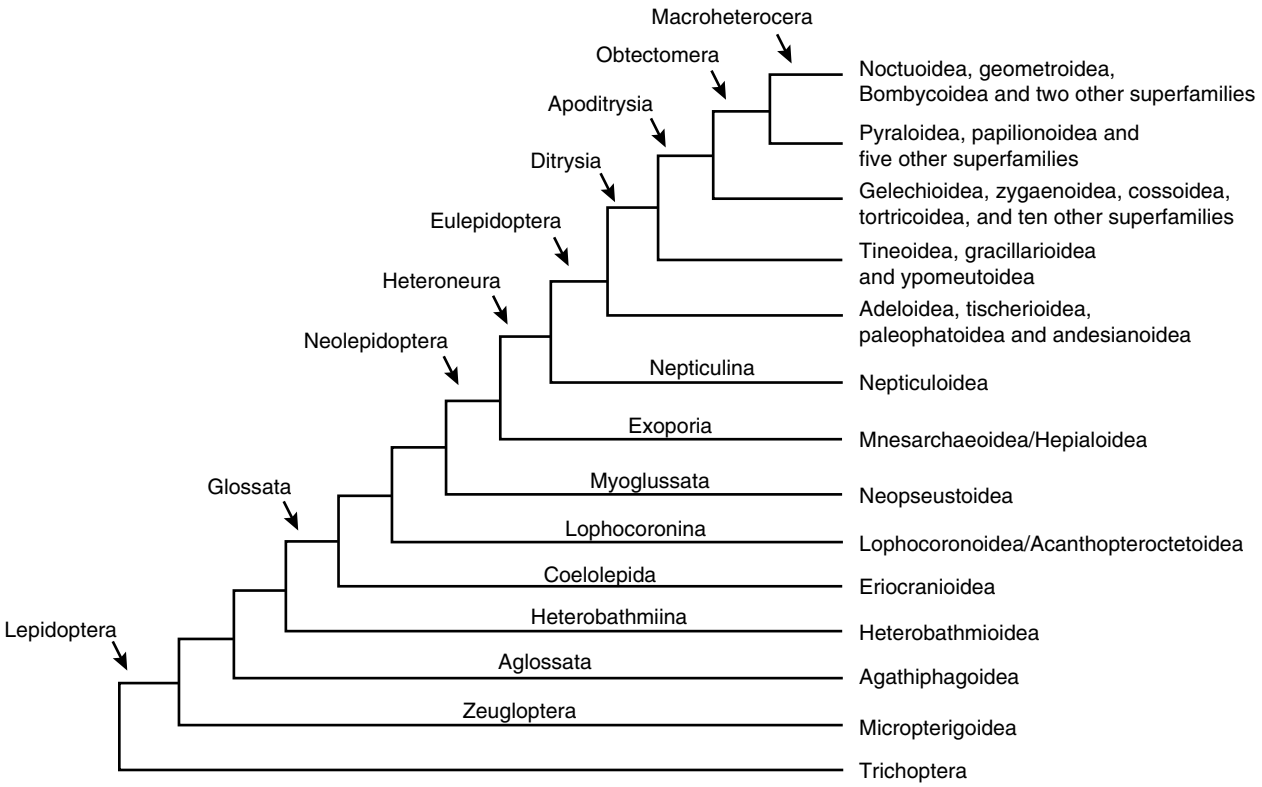


Figure 3 Phylogenetic tree for major lineages of moths. Adapted from Kristensen NP (ed.) (1998) *Lepidoptera, moths and butterflies, Vol. 1. Evolution, systematics, and biogeography Part 35 Arthropoda: Insecta. Vol. IV. Handbook of Zoology,* Berlin: Walter de Gruyter, and with adjustments from van Nieukerken EJ, Kaila L, Kitching IJ, *et al.* (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang Z-Q (ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa* 3148: 212–221.

Neolepidoptera, includes the remainder (>99.8%) of the ordinal species diversity.

The Neolepidopterans

Two infraorders are included: Exoporia and Heteroneura. Exoporia, with six families and more than 636 species, are the only basal moth lineage with appreciable extant diversity. Three of the families are Gondwanan: Mnesarchaeidae (New Zealand, 14 species; seven described), Prototheoridae (South Africa, 12 species), and Anomosetidae (Australia, one species). The largest family, the Hepialidae, with approximately 629 species, is cosmopolitan; although the vast majority of the generic diversity is Gondwanan. The remaining 34 superfamilies are placed into the Heteroneura – the name referring to the venational differences between the fore- and hindwings. Evolutionary relationships among the heteroneurans lineages remain equivocal even in the face of detailed morphological and molecular studies. van Nieukerken *et al.* (2011) recognized two basal (super)clades: the Nepticulina (two included families) and Eulepidoptera (for all remaining Lepidopteran diversity). The latter, in turn, includes four principal clades: Andesianoidea (one family); Adeloidea (five included families); Etimonorysia for Paleaphatoidea and Tischerioidea, each with one family; and Ditrysia. Much of the phylogenetic uncertainty about heteroneuran relationships is believed to be a consequence of rapid diversification within the suborder, with evolutionary events occurring so closely together in time that it has been difficult to accurately reconstruct their sequence. As a group, heteroneurans tend to be internal feeders: most are leaf miners but some also tunnel in stems, buds, and seeds. Among the adeloids are the yucca moths, highly specialized yucca-pollinating mutualists that have been the focus of numerous coevolutionary studies. The most diverse lineage is the (leaf mining) Nepticulidae, with more than 800 named species and perhaps again as many awaiting description. Collectively, all of the aforementioned relictual taxa comprise only about 1.2% described species diversity of Lepidoptera, a percentage that will continue to shrink appreciably as more tropical moth species are described. Ecologically and taxonomically moth diversity is about ditrysians.

The Ditrysia

The remaining 29 heteroneuran superfamilies are placed in the Ditrysia, a group whose monophyly is well supported by the presence of separate copulatory and ovipositional pores, an internal duct that connects the two, and the unique organization of the proboscis musculature. Published phylogenies for ditrysian superfamilies and families are replete with uncertainty. Considerable genetic data will have to be collected before consensus is reached about evolutionary relationships of the families and superfamilies within this remarkably successful group of insects.

Table 1 Diversity of the Lepidoptera

Taxon	*Number of families*	*Number of described (extant) species*
Suborder Zeugloptera		
Micropterigoidea	1	160
Suborder Aglossata		
Agathiphagoidea	1	2
Suborder Heterobathmiina		
Heterobathmioidea	1	3
Suborder Glossata		
Clade and Infra order Dacnonypha		
Eriocraniodea	1	28
Clade Coelolepida		
Acanthopteroctetoidea	1	5
Lophocoronoidea	1	6
Clade Myoglossata		
Neopseustoidea	1	14
Clade Neolepidoptera		
Infraorder Exoporia		
Mnesarchaeoidea	1	7
Hepialoidea	5	627
Infraorder Heteroneura		
Superclade Nepticulina		
Nepticuloidea	2	998
Section Incurvariina		
Superclade Eulepidoptera		
Clade Incurvariina		
Andesianoidea	1	3
Adeloidea	5	582
Clade Etimonotrysia		
Palaephatoidea	1	57
Tischerioidea	1	110
Clade Ditrysia		
unassigned moths	1	104
Tineoidea	3	3823
Gracillarioidea	3	2216
Yponomeutoidea	10	1756
Subclade Apoditrysia	55	
unassigned families	2	30
Simaethistoidea	1	4
Gelechioidea	21	18,489
Elachistidae		3201
Oecophoridae		3308
Gelechiidae		4700
Alucitoidea	2	235
Pterophoroidea	1	1318
Carposinoidea	2	326
Schreckensteinioidea	1	8
Epermenioidea	1	126
Urodoidea	1	66
Immoidea	1	245
Choreutoidea	1	406
Galacticoidea	1	19
Tortricoidea	1	10,387
Cossoidea	7	2881
Zygaenoidea	12	3296
Subclade Obtectomera		
Whalleyanoidea	1	2
Thyridoidea	1	940
Hyblaeoidea	1	18
Calliduloidea	1	49
Papilionoidea	7	18,768
Hedylidae		36
Hesperiidae		4111

(*Continued*)

Table 1 Continued

Taxon	*Number of families*	*Number of described (extant) species*
Lycaenidae		5200
Nymphalidae		6131
Pyraloidea	2	15,576
Pyralidae		5921
Crambidae		9655
Mimallonoidea	1	194
Subclade Macroheterocera		
Drepanoidea	3	672
Lasiocampoidea	1	1952
Bombycoidea	10	4723
Saturniidae		2348
Sphingidae		1461
Geometroidea	4	23,748
Noctuoidea	6	42,407
Erebidae		24,569
Arctiinae		11,000
Noctuidae		11,772
Notodontidae		3800
Totals	124	157,402

Diversity figures were taken from the latter and exclude 86 named fossil taxa. Species-numbers are given for all 44 superfamilies. Families, other than the nominate, that contain more than 2000 described species are given. A few additional families and subfamilies that are mentioned in the text or otherwise deemed to be of special importance are also included. Note: The phyletic sequencing, taxon order, and many of the ranks used here are approximations of evolutionary relationships, for example, the Subclade Macroheterocera is subordinate within the Obtectomera, which in turn is subordinate within the Apodytrisia – taxon (clade) names are best used without their ranks.
Source: Adapted from Kristensen NP (ed.) (1998) *Lepidoptera, moths and butterflies, Vol. 1. Evolution, systematics, and biogeography Part 35 Arthropoda: Insecta. Vol. IV. Handbook of Zoology,* Berlin: Walter de Gruyter, and van Nieukerken EJ, Kaila L, Kitching IJ, *et al.* (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang Z-Q (ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa* 3148: 212–221.

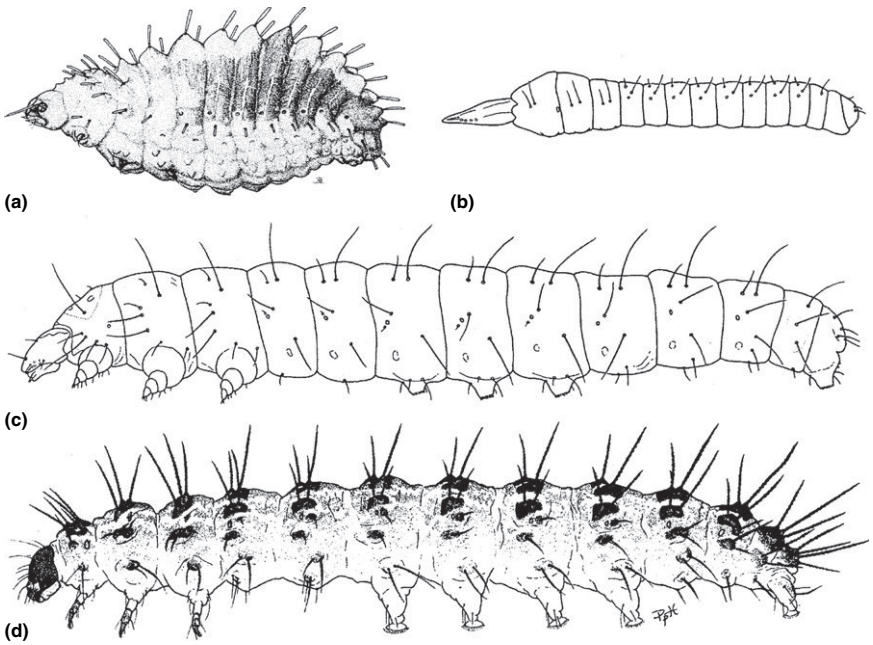


Figure 4 Moth larvae (caterpillars): (a) *Sabatinca* sp. (Micropterigidae). Reproduced with permission from Kristensen NP (ed.) (1998) *Lepidoptera, moths and butterflies, Vol. 1. Evolution, systematics, and biogeography Part 35 Arthropoda: Insecta. Vol. IV. Handbook of Zoology,* Berlin: Walter de Gruyter. (b) *Phyllonorycter blancardella* third instar (Gracillariidae). (c) *Phyllonorycter blancardella* fifth instar (Gracillariidae). (d) *Hypoprepia fucosa* (Arctiinae, Erebidae). (b–d) Reproduced with permission from Stehr FW (ed.) (1987) *Immature insects*, Vol. 1. Dubuque, Iowa: Kendall/Hunt.

Many guides and faunal works divide the Lepidoptera into the so-called Microlepidoptera and Macrolepidoptera. Microlepidoptera roughly corresponds to the basal groups, most of which are small to minute, possess a fused CuP vein in the hindwing, have the larval crochets arranged in a circle, and possess a jugum or the remnants of one (all of which are primitive features that were part of the ground plan for the order). It is a grouping of convenience as the "macros" are clearly derived from the "micros," and therefore some micros are more closely related to the macros than they are to other micros. The Macrolepidoptera, once thought to be a monophyletic grouping that included the butterflies and the last five superfamilies in **Table 1** (most larger moths), has been shown to be a polyphyletic entity – pyraloids would need to be included if the term were intended to reflect a natural group.

Fossil History

Whereas the fossil record of the Trichoptera, the sister group to the Lepidoptera, is adequate and dates into the Permian, the record for lepidopterans is meager, especially for Mesozoic-aged fossils, when the basal lineages were diversifying. There are less than a dozen putative lepidopteran fossils from the Triassic, and most of these cannot be assigned to the order with certainty. The oldest unequivocal moth, *Archeolepis mane*, dates to the Lower Jurassic. By definition, sister taxa are of equal age, and thus it is curious that the oldest caddisfly and moth fossils are of such disparate ages. Skalski has suggested the possibility that some early stem group lepidopteran fossils may be unrecognized because they have been erroneously assigned to the Trichoptera.

More than a dozen moth fossils from the Upper Jurassic and Cretaceous are unassignable to a family. Although most of these clearly belong to preglossatan taxa, this is not true of all. *Protolepis cuprealata* from Upper Jurassic deposits in Kazakhstan appears to have a coiled proboscis, although an alternative interpretation is that the structure in question be an elongate maxillary palpus.

Micropterigids have been identified from several amber and sedimentary deposits. Examples from the lower Cretaceous represent extinct genera, whereas those from the Tertiary are assigned to modern genera. There are no fossil records for either the Aglossata or Heterobathmiina. Glossatans, including a putative lophocoronid, are well substantiated from the Upper Cretaceous. By 102 mya, modern genera of ditrysians are recognizable, suggesting that early evolution of the order occurred on nonseed plants and gymnosperms during the late Jurassic (Labandeira *et al.*, 1994). Not surprisingly, most of the modern superfamilial and species diversity appears after the angiosperms began radiating in the early Cretaceous. The oldest unambiguous macrolepidopteran fossils are Paleocene.

Leaf mining families from the Northern Hemisphere have an extensive fossil record that is believed to reflect their ages of origin as well as their relative abundances. The discoveries of Labandeira *et al.* (1994) of *Ectoedemia*-like and *Stigmella*-like mines (both Nepticulidae) and a phyllocnistine (Gracillariidae) from Albian deposits (102 mya), which agree in detail with those of modern species from related host taxa, are noteworthy in that they demonstrate the antiquity of many moth–plant associations. Another well-represented group of fossil moths are the tineoids, which are richly represented in Dominican amber deposits that date from 40 to 15 mya.

Species Diversity and Biogeography

There is still considerable doubt about the species richness within the order, an alarming observation given that Lepidoptera are among the most well-studied groups of invertebrates. The recent compilation by van Nieukerken *et al.* (2011) lists 157,402 species. Heppner's (1998) earlier figures for global diversity by biogeographic region are shown in **Table 2**. He guessed that half again as many species remain to be described, and that richness for the order would climb to 255,000 species (**Table 2**). Gaston (1991) and Kristensen (2007) have offered higher estimates, in the range of 300,000–500,000 species.

Heppner's tallies indicate that 92% of Palearctic Lepidoptera have received names (**Table 2**). The richest biogeographic realm, the Neotropics, with an estimated 35% of the world's butterfly and moth fauna, has only about 50% of its fauna named. Moreover these figures may be conservative. In recent revisions of neotropical tineids and tortricids 75–90% of the species were regarded to be new. Of the more than 250 species of Gracillariidae that were collected at the La Selva Biological Station in Costa Rica, 98% were believed to be undescribed.

There can be little doubt that microlepidopterans will be a frontier for descriptive taxonomy for decades if not centuries. In general, small cryptically colored taxa have greater proportions of unnamed species. Also, species-rich taxa that are intractable taxonomically have proportionately greater numbers of undescribed species. Microlepidopteran superfamilies with especially high numbers of unnamed species include the Tineoidea, Gracillarioidea, Gelechioidea, and Pyraloidea. But even among the Macrolepidoptera, considerable taxonomic work remains, especially for neotropical taxa. The only two appreciably diverse moth families where the descriptive

Table 2 Worldwide lepidoptera species diversity by biogeographic region (1758–1990)

	Nearctic	*Neotropical*	*Palearctic*	*Ethiopian*	*Oriental*	*Australia*	*Total*
Described species	11,532	44,791	22,465	20,491	27,683	19,603	146,565
% of world fauna	7.9	30.6	15.3	14.0	18.8	13.4	100
Estimated richness	14,000	90,000	25,000	38,000	50,000	38,000	255,000
% of world fauna	5.5	35.3	9.8	14.9	19.6	14.9	100
% of fauna known	82	51	92	51	53	50	57

Source: Reproduced from Heppner JB (1998) Classification of the Lepidoptera. Part 1. Introduction. *Holarctic Lepidoptera*. vol. 5(supplement 1).

taxonomy is approaching completion are the Saturniidae and Sphingidae, veritable behemoths among insects that have long held favor among collectors.

Present tallies suggest that the Noctuoidea is without parallel in richness, with 42,407 described species, most of these falling within the Erebidae and Noctuidae (**Table 1**). It is the most species-rich superfamily on all continents except Australia, where its preeminence is superseded by the Gelechioidea. Moreover, many thousands of noctuoids await description, particularly in the Neotropics. Current tallies of described species place the Geometroidea second in richness with some 23,748 species worldwide, nearly all of which are in the nominate family. Its present ranking could fall to fourth once world microlepidopteran faunas are studied. Presently, there are approximately 18,489 described species of Gelechioidea and 15,576 Pyraloidea, either of which could contain more than 50,000 species and prove to be the most species-rich superfamily within the Lepidoptera. The butterflies, with 18,768 species, will rank no higher than fifth in global importance. In the most well-studied faunas (i.e., those of the north temperate zone and Australia), moths generally comprise about 95% of the lepidopteran species diversity, with butterflies making up the remainder. In terms of biomass and number of individuals, moths and their caterpillars probably have even greater (ecological) importance.

Because tropical faunas are incompletely known, especially for microlepidopterans and smaller Geometroidea and Noctuoidea, a quantitative assessment of global diversity patterns for the entire order is not yet possible. Nevertheless, virtually all lepidopterists are in agreement that greatest species, generic, familial, and superfamilial diversity is found in tropical rain forests. Preliminary estimates of the moth fauna of Costa Rica (51,100 km^2) run between 13,000 and 16,000 species and thus are roughly comparable to the species diversity of the entire North American continent north of Mexico or that of Australia, areas 380 and 150 times greater in size, respectively. Taxa with exceptional tropical richness include the Erebidae, Gelechiidae, Geometridae (especially in middle and higher elevations), Oecophoridae, Notodontidae, and Pyraloidea.

Given that moths are quintessentially phytophagous, their species diversity should parallel that of vascular plants. If this is correct, the highest global richness should occur on the low- to mid-elevational slopes around the Amazon basin. The forests of Southeast Asia and Indo-Australia also would be expected to harbor hyperdiverse moth faunas. In temperate regions, the floras of South Africa and southwestern Australia are among the richest in the world. These latter areas are noteworthy because they support comparatively high percentages of nonditrysian moths and hence can be regarded as two of the world's most important "living museums." Not all moth taxa are hyperdiverse in the tropics. Familial, generic, and species diversity for the nonditrysian moth lineages is highest in extratropical latitudes. In temperate areas of the Northern Hemisphere, the Coleophoridae are numerically important Lepidoptera in grassland, scrub, and desert ecosystems, yet the family appears to be absent in the rain forests of Central America. The Tortricoidea also appear to be less diverse than would be predicted based on the latitudinal richness gradients seen in other Lepidoptera.

On a local scale, Lawton has shown a correlation between herbivore richness and host plant architecture. Trees have more architectural complexity and therefore more herbivores than shrubs, shrubs more than herbs, and so on. An example of the importance of trees is provided by white oak (*Quercus alba*) – across its range in eastern North America it hosts more than 400 species of moths (this number includes both specialist and generalist feeders). It follows that ecosystems with high tree diversity (e.g., lowland tropical rainforests) would have especially rich moth faunas.

Island faunas, especially those of distant islands, are always unbalanced with respect to those of the nearest continents. Macroheterocerans that tend to be well represented on islands include the Erebidae, Geometridae, Noctuidae (especially Heliothinae and Noctuinae), and Sphingidae. Despite their small size, microlepidopterans are present even on remote islands, presumably because they disperse effectively as aerial plankton. Hawaii has been colonized by dozens of different microlepidopteran lineages that can be characterized as weak fliers. A second example: the richness of all leaf mining families (minute to small moths) on Santa Cruz Island, off the coast of California, is nearly that of the mainland, whereas most groups of larger moths and other insects are represented by less than half of the mainland fauna.

Life History

Egg

The ova are variable in shape and surface ornamentation. Microlepidopterans often have flat, smooth eggs, many of which are nearly transparent once laid on a leaf surface; those of macrolepidopterans tend to be raised. The latter may be round, square, disk shaped, or spindle shaped; the chorion (outer surface) varies from smooth to richly ornate. Females generally deposit the egg within or on the surface of the host; in the latter case, secretions released from the accessory glands bond the egg to the substrate. Females of some taxa leave behind abdominal scales, which discourage natural enemies or advertise to other females that the resource already has been found. A few polyphagous species (e.g., arctiines and hepialids) broadcast the eggs as they fly. The latter family also includes the world's most fecund lepidopteran – a single female *Abantiades hyalinatus* laid 30,000 eggs (and died with another 20,000 ova remaining in her abdomen).

Larva

The larva (**Figures 4–9**) is the principal feeding stage, increasing its mass by four (to five) orders of magnitude during its development. There are from three to more than a dozen larval instars, with five being the mode. Dyar was the first to note that there is commonly a linear increase of 1.3–1.4 in head dimensions between successive instars. Dyar's rule holds across many lepidopteran families and often is employed to infer the number of larval instars. In bombycoids and a few other groups the females may pass through an additional instar. Development is remarkably fast in some species, especially those that feed on ephemeral tissues such as fruits and



Figure 5 Moth caterpillar defense (also see **Figures 6–9**). *Isa textula* (Limacodidae). *Isa* caterpillars are armed with hundreds of stinging setae. The setae, positioned about the caterpillar's periphery, may be especially efficient at repelling ants.



Figure 6 *Phobetron pithecium* (Limacodidae). *Phobetron* caterpillars are thought to mimic the cast "skins" of tarantulas, which contain thousands of deciduous urticating hairs – *Phobetron* caterpillars do not sting and are thought to be harmless. Courtesy Noble Proctor.

new leaves and across many desert dwellers. Larger moths of high latitudes may take 2 or 3 years to mature. In *Gynaephora groenlandica* – a lymantriine erebid found well north of the Arctic Circle – caterpillars mature over a 14-year period.

Hypermetamorphosis – where two or more larval forms occur, each specialized for different functions – is rare within the order. The best-known example occurs in the minute to small leaf-mining moths of the family Gracillariidae. In most genera the first two or three instars are flat, legless, possess anteriorly directed mouthparts, lack a spinneret (silk-spinning organ) (**Figure 4(b)**), and feed in a single cell layer, usually the epidermis. The diet is liquid, the larva filtering plant fluids from damaged cells. In later instars the larva takes on a rounded form that is legged, possesses a spinneret, and the mouthparts are directed ventrally (as in other externally feeding caterpillars) (**Figure 4(c)**), and consumes whole cells



Figure 7 *Automeris postalbida* (Saturniidae). *Automeris* caterpillars are fortified with a battery of urticating setae – the effect of which is not unlike that of stinging nettle.



Figure 8 *Euproctis chrysorrhoea* (Lymantriinae, Erebidae). *Euproctis* caterpillars have red middorsal glands on the sixth and seventh abdominal segments that yield a *pot pourri* of defensive secretions – this species purportedly was responsible for two human deaths in Massachusetts.

and tissues. The spinneret is critical in that it is employed by the larva to lay down silk within the mine drawing it into a bubble, thereby creating a space where a rounded body and legs are appropriate. Metamorphosis in some opostegids, *Bucculatrix*, epipyropids, cyclotornids, and others is also hypermetamorphic. Nonfeeding instars are rare, generally occurring in the first or last instars. An example of the former is provided by some zygaenoids. In a number of gracillariids (e.g., *Cameraria* and *Marmara*) there may be as many as two nonfeeding prepupal instars. Whereas a general loss of color is common among prepupae, some turn pink to scarlet prior to pupation (e.g., some Prodoxidae, Gracillariidae, Elachistidae, and Notodontidae).

Pupa

In the most basal lineages the wings, legs, and antennae are weakly fused to the main body of the pupa (**Figure 10**). In some of these, the pupal mandibles, which may be impressively large (**Figure 10(a)**), are used to cut through the cocoon or work the pharate adult free from the pupal



Figure 9 *Hemeroplanes* sp. (Sphingidae). Courtesy of Phil De Vries. *Hemeroplanes* caterpillars are superb snake mimics, despite the fact that mature larvae are but several centimeters in total length. Courtesy of Phil De Vries.

crypt. In heteroneurans the appendages are fused to the body and the mandibles are much reduced and immobile (**Figure 10(c)**). Another phylogenetic trend involves the degree of fusion between adjacent abdominal segments. In the most archaic moths the first seven abdominal segments are movable. Intermediate degrees of fusion occur among microlepidopterans. In the Obtectomera, a clade that includes 12 of the most derived lepidopteran superfamilies (**Figure 3**), the first four abdominal segments are fused and are immobile with respect to one another. Taxa with appreciable abdominal mobility often extrude the anterior end of the pupa from the cocoon or pupal crypt prior to emergence (e.g., most microlepidopterans). Those with a high degree of fusion hatch within the cocoon or pupal crypt and must crawl free before expanding their wings (e.g., all Obtectomera). Male pupae have a single set of paired genital marks on the venter of the ninth abdominal segment; females have an additional paired set of genital "scars" on the eighth segment as well (e.g., **Figure 10(b)**). The pupal stage may be as short as 67 days or may last several years, the latter is not unusual among species that inhabit arid environments.

Most moths spin a cocoon on or within the larval substrate or find their way to the ground and pupate in litter or duff. The cocoon may be a highly elaborate affair that takes days to spin. The most famous example is that of the Oriental Silk moth (*Bombyx mori*); the cocoon is fashioned from a single strand of silk that, when unwound, stretches for more than half a mile (0.8 km). A curiously elaborate cocoon is made by *Marmara* and related genera in the Gracillariidae: the cocoon is ornamented with sometimes more than 100 anally extruded bubbles, each of which is nearly the diameter of the caterpillar's body. Cocoons represent a significant energetic investment, in that silk is entirely protein, and the nitrogen that goes into its production is generally in short supply in plants. In addition to silk, the cocoon may incorporate excrement, scrapings from the larval substrate, or larval setae (especially if these are stiff, poisonous, or urticating). Prepupal limacodids discharge a calcium oxalate solution over the inner surface of the cocoon that hardens the cocoon as it dries. Open, netlike cocoons are often spun by species with aposematic pupae. Taxa that pupate in concealed sites such as below ground or in wood often forego spinning a cocoon. A few moths (e.g., elachistines, pterophorids, sterrhine geometrids, and hedylids) pupate naked, exposed on surfaces, like butterflies.

Diapause and Quiescence

Diapause may occur in any of the four life stages. The largest fraction of temperate moths pass through the winter months in diapause as prepupal larvae or pupae; the second largest fraction overwinters as eggs. Moths that overwinter as adults invariably need to feed over 5–7 months of winter – some of these are among the unwelcome guests in the buckets used to collect maple syrup. Reproductive diapause is the rule among moths that overwinter as adults, with ovarian maturation and mating being delayed until the arrival of spring. Summer diapause and estivation occur in many groups that live in habitats with a pronounced dry season. Enormous aggregations involving tens of thousands of moths (usually noctuids) are known to assemble in caves or on (hill) mountain tops. Moths in these aggregations become important sources of protein for vertebrates, even for grizzly bears. The diapause longevity record belongs to the desert-inhabiting prodoxids – single larval cohorts may yield adults over a period of three decades.

Feeding Biology

Larvae

Given the enormity of the order it is remarkable how little diversity there is in feeding habits. More than 99% of all moths are fundamentally herbivorous. Within this realm, however, they have radiated to exploit virtually every corner of niche space. Terrestrial algae, lichens, fungi, bryophytes, pteridophytes, and seed plants are eaten. All plant tissues and organs are consumed, but especially leaves, meristems, and reproductive structures. A majority require living tissues, sometimes of a very specific age, whereas others subsist on dead and fallen leaves and wood.

The most basal moth lineage (Micropterigidae) with larvae reliant on nonseed plants account for about 1% of extant

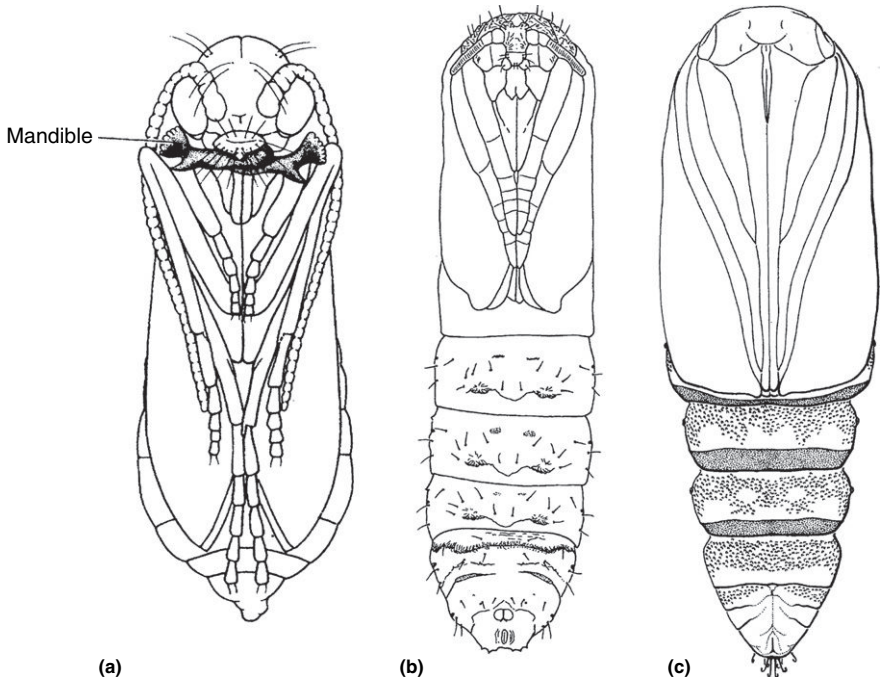


Figure 10 Moth pupae: (a) *Agathiphaga* sp. (Agathiphagidae). (b) *Sthenopis auratus* (Hepialidae). (c) *Asota* (Erebidae). Across this phylogenetic progression there is reduction in the mandibles, increased fusion of appendages, simplification in the external expression of the appendages, and decreased abdominal mobility. (a) and (c) Reproduced with permission from Kristensen NP (ed.) (1998) *Lepidoptera, mothsandbutterflies. Vol. 1. Evolution, systematics, and biogeography Part 35 Arthropoda: Insecta. Vol. IV. Handbook of Zoology*. Berlin: Walter de Gruyter, and (b) reprinted with permission from *Journal of the New York Entomological Society* 97, 7.

species diversity. The order's great species, generic, and even familial radiations were reliant on the diversification of the angiosperms. The first radiation of note, as measured by extant richness, was that of the Hepialidae, with some 620 species worldwide. Their success may be due to their diet, which is arguably the most diverse of any moth family. Larvae eat bryophytes, pteridophytes, and seed plants (including both conifers and hardwoods); leaf and especially woody root and stem tissues; leaf litter and fungi, especially in early instars. Many are cannibalistic, and some even attack and consume soft-bodied arthropods. Another early lineage with remarkably catholic diets are the tineoid families, a high percentage of which are fungivores and detritivores. Members of the Tineidae (e.g., the clothes moths) are noteworthy for their ability to metabolize keratin, the normally undigestible protein that makes up fur and feathers. Others feed in bird and bat guano, dung, and even the dead insects that accumulate below spider webs.

As the Lepidoptera radiated so did their feeding habits and their associations with plants (Powell *et al.*, 1998; Menken *et al.*, 2010). A majority of microlepidopterans are concealed feeders, tunneling into plant tissues, forming leaf mines or shelters, or feeding from within silken cases. Galls are made by some members of more than a dozen families, but none of the lineages can be regarded as particularly successful. Diets tend to be specialized in internal feeders because the host is not only the food but also the environment for some or all of the immature stages. The most intimate host associations occur in taxa where the adults and larvae utilize the same host plant, the most famous example being that of the yucca moths in the genera *Tegeticula* and *Parategeticula* (Prodoxidae). The female yucca moth uses special maxillary "tentacles" to collect yucca pollen and then flies to a second flower where she scrapes the pollen over the stigmatic surface. She then deposits one to a few eggs in the developing ovary, and in so doing, leaves a chemical signature that can be detected by other females. The larvae are seed predators but do not consume all the seeds in a fruit. The mutualism is obligate, as neither moth nor yucca can survive without the other. Two additional moth-pollination mutualisms have come to light over the past two decades. Senita cactus is pollinated principally by a crambid in the Sonoran Desert, and gracillariids in the genus *Epicephala* pollinate members of the Phyllanthaceae genus *Glochidion* in Southeast Asia. As in yucca moths, after actively pollinating a flower the females oviposit in the flowers and the larvae mature on developing seeds.

Leaf miners, minute moths whose larvae feed between the two surfaces of a leaf, account for some early radiations (Menken *et al.*, 2010). The Nepticulidae, whose fossil history traces back more than 102 million years to the Albian, contains more than 1000 species with at least this number remaining to be described. Gracillariidae, with 2200 recognized species and perhaps three times as many unrecognized, also trace their fossil history back to the Albian. The latter family contains members that have been the first shelter-forming Lepidoptera. Early instar gracillariines mine in leaves as do other gracillariids, but middle instars may exit the mine to form a leaf shelter in which they complete their development.

The three largest microlepidopterans superfamilies (the Gelechioidea, Tortricoidea, and Pyraloidea) are concealed

feeders, mostly leaf rollers or shelter formers, but others are borers, leaf miners, gall formers, and detritivores. A handful are predators, inquilines in galls or nests of social insects, and so on. Leaf rollers and shelter formers exhibit intermediate levels of host plant specialization, but polyphagy is not common. An important exception are many tortricids, especially among members of the nominate subfamily that are generalists on woody plants. Several microlepidopteran families have representatives that are external feeders for one or more instars (e.g., the Bucculatricidae, Schreckensteinidae, Pterophoridae, and others), but none accounts for appreciable present-day diversity. The most species-rich lineage of externally feeding microlepidopterans are the zygaenoids, which as a group appear to be well protected – their arsenal includes urticating spines; abundant, long, or dense setae and hairs; and cyanoglucosides (**Figures 5** and **6**).

Roughly 55% of the described Lepidoptera are external feeders. Polyphagy is common among external feeders and is the mode for many macroheteroceran families (e.g., Lasiocampidae, Saturniidae, ennomine Geometridae, Lymantriinae, and some subfamilies of the Noctuidae). It is worth noting that many species that are host plant generalists across their geographic range specialize on just one or a few plants at any one locality (Fox and Morrow, 1981). Regardless of the mode of feeding, internal or external, caterpillars that feed on poisonous plants tend to be host plant specialists.

Fungivory appears throughout the order but is only common in the Hepialidae and Tineoidea. The line between fungivory and feeding on fallen plant material is often an arbitrary one (see Rawlins, 1984) – both of these groups are also detritivores. Other sizable lepidopteran clades, with members specializing on fallen leaves, decaying wood, and other nonliving plant tissues, include the herminiine (Erebidae), Blastobasidae, Lecithoceridae, Oecophoridae, some Australian Tortricidae, and a few pyraloid lineages. In pineland, hardwood, and especially mixed forests of the eastern USA, herminiine erebids may account for more than half of all macrolepidopterans in light trap collections during the month of July. Australia has an exceptionally rich fauna of litter-feeding Oecophoridae and Tortricidae (Common, 1990). There, numbers are so great in eucalyptus forests that the larvae are believed to play a significant role in litter processing.

The most common type of carnivory in the order is cannibalism, often showing up in taxa where the host substrate is limited (e.g., annual plants and architecturally reduced perennials) or where movement between hosts is either impossible or exposes larvae to considerable risks (e.g., internal borers). Essentially phytophagous species that occasionally eat other caterpillars occur sporadically across the order. The Epipyropidae and Cyclotornidae are obligate ectoparasites of homopterans. The latter is only parasitic in the first instar; the second to final instars are brood predators in ant nests. About 200 lepidopterans are known to be predaceous, many of which are lycaenid butterflies. The most common prey are sessile or nearly sessile insects, such as the brood of ants or aphids, scale insects, and related homopterans. Hawaii has spawned the most exceptional predaceous caterpillars – members of the genus *Eupithecia* (Geometridae) are sit-and-wait predators that snatch small insects that are so unwise as to crawl over the caterpillar's body.

Special cases are legion, and worthy of their own review. Crambid acentropines, which number more than 700 species, are aquatic as larvae, with some even adapted for life in rapidly flowing water. The caterpillars feed on submerged aquatic plants tissues or algae. No other Lepidoptera families have yielded speciose lineages that are aquatic or even subaquatic. Adults of three chrysaugine pyralid genera are found on or about sloths, awaiting the once-weekly release of dung, which serves as the substrate for the developing larvae. Other chrysaugines have diversified into a strikingly varied set of niches: North American species include predators of paper wasp brood, cattail borers (Monroe, 1972). *Tallula watsoni*, a phyctine pyralid, has a particularly unusual diet, its larvae feed as inquilines in the webs of a subsocial spider (*Anelosimus studiosus*).

Adults

Micropterigids are mandibulate and use their jaws to grind up pollen and fern spores. An archaic proboscis, formed from the maxillae in Dacnonypha (and other Lepidoptera), is used by adults to collect fluids from bark and other surfaces. These fluids may include dissolved honeydew, but adults do not visit flowers for nectar. Nectar feeding with a coiled proboscis is first seen in the Incurvariina and is retained by virtually all the larger lineages of moths that follow. In addition to nectar and other sources of dissolved sugars, moths use the proboscis to imbibe water, liquids from injured plants and tree flows, dissolved salts (including human perspiration and lacrymal secretions), pus, animal excreta, substances from the crushed bodies of other insects, rotting fruit, dung, carrion, and many other substrates. Calpine erebids use their short, stout proboscis to pierce fruits. A few calpines have taken to collecting lachrymal secretions of birds and mammals; in the most extreme cases, an armored proboscis is used to pierce skin and take blood meals from mammals. At least heliconiine nymphalids and perhaps others have the ability to solubilize (e.g., alkaloids) and even digest solid (e.g., pollen) substrates by pumping oral fluids down through the proboscis.

A favorite collecting technique for moths is "sugaring." A standard bait is made by mixing fruit, beer, and sugar. Whereas overripe bananas are the most universal ingredient, collectors sometimes add watermelon, apricot, grapes, rum, jellies, molasses, and a rather long list of other substances, some of which might best be left unnamed. Once the mixture has been allowed to ferment it is then applied to tree trunks or offered on sponges. Virtually all of the winter-active moths (e.g., Oecophoridae, Tortricidae, and Noctuidae) feed at bait. During summer months, many moths in dry woodlands also come to bait. Presumably, moths that are attracted to bait normally would be feeding at tree wounds, sap flows, rotting fruit, or the honey dew secreted by homopterans.

Plant tissues tend to be low in sodium relative to the metabolic needs of animals. Many lepidopterans reach adulthood with inadequate supplies of this essential element, but they are able to acquire it by imbibing fluids from mud puddles and water courses (**Figure 11**), urine, sweat, and lacrymal secretions. Although "puddling" is well known among butterflies, it also occurs in a diverse array of diurnal and nocturnal moths. In most moth and butterfly species,



Figure 11 *Urania lelius* (Uraniidae). Sunset moths are among the most dazzling of all invertebrates. Males are puddlers that supplement their metabolic needs with the dissolved sodium found in mud and other sources – *see* Adults. Courtesy of Tom Murray.

90–100% of the individuals that puddle (or feed at other sodium-rich substrates) are males. A remarkable case is provided by *Gluphisia septentrionis* (Notodontidae). Puddling males will drink (and forcibly excrete) more than 600 times their own body weight in a single evening. *Gluphisia* males transfer sodium to females with their ejaculate; females in turn pass much of the sodium to their offspring.

Many adult moths lack functional mouthparts or digestive systems, including many (or all) Apatelodidae, Arctiinae, Bombycidae, Hepialidae, Lasiocampidae, Lymantriinae, Notodontidae, Saturniidae, and Zygaenoidea. Adults of late fall and early spring geometrids also tend to be nonfeeding, whereas those that fly during the summer commonly consume nectar.

Natural Enemies and Chemical Defense

Natural Enemies

Moths are fecund invertebrates that produce 100–10,000 times more ova than will ultimately survive. A panoply of biotic and abiotic mortality factors comes into play during every life stage (and larval instar) to limit population growth. Biotic factors include an enormous range of pathogens, parasites, and predators. Population numbers of many pest species, especially in forests and other stable communities, often are controlled by viruses, bacteria, and fungi. Viruses are often highly specific, and one wonders if there might be nearly as many lepidopteran viruses as there are moths and butterflies. There are comparatively few recorded species of bacteria and fungi that attack insects, but the taxonomy of insect pathogens, especially those that cannot be easily cultured, is understudied; although some are highly specific, others have broad host ranges. *Bacillus thuringiensis*, the most widely used pathogen for the control of Lepidoptera in gardens, agriculture, and forests, has a host range that includes hundreds (thousands?) of species across many superfamilies. Fungal pathogens, mostly in the Entomophthorales and Fungi Imperfecti, often require special conditions such as a period of high humidity before appreciable infection will occur. A species of special note is *Entomophaga maimaiga*. This fungal pathogen was introduced into North America from Japan in 1910 and 1911 as a biological control agent to stem the ever-increasing populations of the European gypsy moth (*Lymantria dispar*). The introduction was considered unsuccessful until 1989 when the fungus began showing up in caterpillar cadavers throughout the Northeast. It has since spread throughout much of the North American range of the gypsy moth and continues to be a principal mortality agent for the gypsy moth caterpillar. A diverse assemblage of nematodes and nematomorphans also attacks Lepidoptera, especially the larvae, but their importance is generally regarded as modest among nonsoil dwellers. Microsporidians affect both fertility and longevity in many moths.

Wolbachia are known from across the Obtectomera and Heterocera, but infections likely occur across the order, given that some 25–75% of all insect species are known to be hosts to these endosymbiotic rickettsia-like bacteria (Kozek and Rao, 2007). In many species the cytoplasmic infections are passed vertically from mother to offspring (in the egg), whereas in others transmission occurs through sexual contact. The consequences of infection are both varied and complex, and adequate description is beyond the scope of this treatment. Known effects within the order include male-killing (through selective killing of male embryos); feminization, whereby genetic males are rendered into functional females; and egg–sperm cytoplasmic incompatibility (between sexual partners/mating strains). The nature of the infections is taxon-dependent, and ranges from entirely pathogenic and lethal (e.g., in those where males are killed) to symbiotic (e.g., some *Wolbachia*-infected insect lineages are reported to confer resistance to an array of RNA viruses and some pesticides). In *Phyllonorycter blandcardella* (Gracillariidae) (**Figure 4**) – a leaf miner of apple – *Wolbachia* infections maintain green tissues about the larval mine, even while the remainder of the leaf is yellowing, allowing larvae to complete their development before the host leaf senesces (Kaiser *et al.*, 2010).

Lepidoptera are primary hosts for a multitude of parasitoids. Three of the largest insect families on this planet owe much of their evolutionary success to the Lepidoptera: the Braconidae (>10,000 species) (Hymenoptera), Ichneumonidae (>10,000 species) (Hymenoptera), and Tachinidae (>20,000 species) (Diptera). Another 20 or so families of insect parasitoids attack Lepidoptera, all but two of which (Diptera: Bombyliidae and Sarcophagidae) are parasitic hymenopteran wasps, mostly in Chalcidoidea and Proctotrupoidea. All four life stages are exploited by parasitoids, although most attack larvae. Typically only one or a few instars are attacked by a given wasp species. Most Hymenoptera are endoparasitoids that grow slowly initially, consuming only blood (hemolymph), their own nutritive cells released into the host, or nonvital host cells and tissues, then finish off the host in a burst of feeding and growth, often triggered by the host larva's size or hormonal state. In some cases this may be months after the wasp eggs were laid within the body of the caterpillar (see Wagner *et al.*, 2011 for a brief synopsis).

Some of the smallest insects, trichogrammatid and mymarid hymenopterans, with body lengths of only 0.2 mm, are egg parasites. A great majority of insect parasitoids are specialists that utilize but one or a few related species. Large, long-lived caterpillars support rich guilds of parasitoids: across eastern North America, 29 fly and wasp parasitoids are recorded from *Hyalophora cecropia*, at least three of which are believed to be specialists on this saturniid and its congeners. Large caterpillars may form the base of an entire food web, as evidently there are six hymenopteran hyperparasitoids that parasitize the parasitoids of *Hyalophora*, and one wasp hyperparasitoid that principally attacks the hyperparasitoids. Askew and others have noted that leaf miners and gall-forming caterpillars have an exceptionally rich parasitoid fauna. *Phyllonorycter apparella* (Gracillariidae) has a wingspan of only 7 mm, yet larvae and pupae of this leaf miner are known to host 20 species of hymenopteran parasitoids and hyperparasitoids.

Invertebrate predators of major importance include mites, spiders, predaceous stink bugs, beetles, robber flies, ants, and wasps. Many of the physical, chemical, and behavioral defenses of caterpillars are known to be effective in thwarting the pillages of ants (**Figure 5**). No doubt, these defenses are of special importance in tropical ecosystems where ants make up much of the insect biomass. Mammals, birds, reptiles, and amphibians are the principal vertebrate predators. Birds are such an important selection pressure for externally feeding caterpillars that there is a growing consensus that they have shaped not only what many caterpillars look like, but also what, when, and how they feed (Heinrich, 1993). Their influence on adult color patterns and behavior also seems to be pre-eminent. Bats, too, have been an exceptionally important evolutionary force in the diversification of the order. Hearing organs, sensitive to the frequencies emitted by bats, have evolved multiple times; within one of these groups, the Arctiinae (tiger moths), some chemically protected taxa have evolved the ability to signal back to bats. On detecting ultrasound frequencies, these moths begin clicking, presumably advertising their whereabouts (and toxicity). In addition, there is growing evidence that some can even thwart bats by jamming their echolocation sensory systems. Corcoran *et al.* (2009) have shown that *Bertholida* tiger moths employ a sound-producing tymbal located below the hindwing to generate ultrasonic frequencies that are sufficiently disruptive to cause bats to veer off course and abort their normal attack repertoire.

Defensive Chemistry

The defense strategies of Lepidoptera are legion and make for fascinating reading and study (see Bowers, 1993; Stamp and Casey, 1993; Scoble, 1992, for larvae and adults, respectively). All life stages may be chemically protected from at least a subset of the relevant natural enemies. Protected species (and life stages) tend to be aposematic. Many adult tiger moths (**Figure 12**) and the caterpillar of the monarch butterfly are dramatic examples; but even eggs and pupae of protected species may be rendered in warning colors. Adults of distasteful or otherwise toxic species are often diurnal and may be appreciably slower fliers than their unprotected sister lineage.



Figure 12 *Platyprepia virginalis*; chemically protected moths are often boldly colored and, like this species, diurnal.

Additionally, many tolerate rough handling (e.g., pinching) to a much greater degree than palatable species, a necessity of sorts because some fraction of the protected insect population has to tolerate or least survive the delivery of the noxious chemicals necessary to train the local predator complex (in each generation). Unpalatable species, as first pointed out by Bates (1982), frequently serve as models to one or more Batesian mimics or become entwined in Müllerian mimicry complexes.

Chemical defense against parasitoids and predators likely has played a prominent role in the diversification of the order. The ecological and evolutionary constraints facing unpalatable species differ widely from those faced by palatable taxa (Heinrich, 1993). For example, chemically protected caterpillars are often brightly colored, feed day or night, and rarely take measures to conceal their feeding damage; many are gregarious. Conversely, palatable species are apt to be cryptic in color, to be nocturnal, to move slowly, and to only rarely leave conspicuous signs of feeding. Across the order there is a spectrum of palatability, ranging from those species that can be eaten in large numbers (by birds or indigenous peoples) to those where contact or ingestion can be fatal. There is an enormous literature on this subject – this discussion and examples here are only meant to be illustrative.

Stinging setae, hollow and filled with poison, occur widely in hemileucine (Saturniidae), Limacodidae, Megalopygidae, and Zygaenidae. In rare instances, human fatality has been reported after encounters with the first and third. The body setae of *Euproctis chrysorrhoea* (Lymantriinae) (**Figure 8**) contain histamine, protease, lipase, and phospholipase, and are highly urticating and were perhaps responsible for two human deaths. (These cases may have resulted from severe allergic reactions or exposure and are regarded to be extraordinary. No human fatalities are known in recent times.) Gypsy moth caterpillars and other lymantriines drag the body setae through the dorsal abdominal glands, "arming" them with histamine and a diverse mixture of noxious substances that cause itching and other allergenic responses, which can be severe in some individuals. Other taxa with large numbers of species with urticating setae include the Arctiinae, Lasiocampidae, and thaumetopoeine Notodontidae (see review by Hossler, 2010).

In some lineages, caterpillars and adults synthesize noxious compounds from simple dietary precursors. Zygaenid larvae produce cyanoglucosides from two basic amino acids.

Chemical protection in many Notodontidae is based on simple manufactured compounds (e.g., the late instars of many Heterocampinae eject formic acid and ketones from a cervical gland if molested).

Other moths sequester toxins or precursors from their (poisonous) host plants; toxin concentrations in the hemolymph may greatly exceed those that occur in any of the plant tissues on which the caterpillar has fed. Examples include *Euphorbia*-feeding caterpillars, which retain dietary terpenols, and the milkweed caterpillar (*Euchaetes egle*) (Arctiinae), which, like the monarch butterfly caterpillar, sequesters high concentrations of a cardenolide. Other tiger moths (arctiines) actively seek out and consume host plants high in pyrrolizidine alkaloids (PAs). Experimental studies have demonstrated the effectiveness of PAs in repelling both vertebrate and invertebrate predators as well as their importance in courtship, mating, and paternal investment. If bright coloration can be used as a proxy for chemical protection, there is still a universe of protected moth species whose natural chemistry remains uncharacterized and unappreciated.

In many species the chemicals sequestered or manufactured by the caterpillars are retained by the adult. PAs collected by male *Utetheisa bella* (Arctiinae) caterpillars are used to synthesize the male courtship pheromone. Eisner and his students found that female *Utetheisa* (**Figure 13**) raised on diets devoid of PAs are readily consumed by spiders, but females that have mated with males laden with PAs are cut out and dropped from webs. The eggs and even young caterpillars, resulting from the latter type of pairing, are also chemically protected by the alkaloids collected by the father. Arctiines are not completely dependent on the larval diet because adults of some tiger moths are able to locate and ingest dissolved PAs. They will imbibe fluids from the surfaces or wounds of PA-containing plants (e.g., Asteraceae, Boraginaceae, and Fabaceae) or even the crushed bodies of alkaloid-rich caterpillars or moths. Many tiger moths bubble out a defensive secretion from the prothorax on disturbance. In *Grammia*, the fluid is mostly clear and yellow to green and may represent nothing more than hemolymph (**Figure 14**); in others it is more frothy, sometimes milky, and loaded with acetylcholine, histamine, and possibly pyrazines. Adult female lymantriines collect larval hairs from the cocoon and incorporate these into their egg mass.

Not all pupae or adults retain the protection enjoyed by their larvae (e.g., adult saturniids and limacodids whose caterpillars are armed with stinging setae are generally palatable). Buckeye (Nymphalidae) butterflies whose larvae are protected by sequestered iridoid glycosides are palatable as adults, as much of their chemical arsenal is lost by the pupal stage.

Dispersal and Migration

Long-distance movements occur in many groups, not an insignificant number of which are pest species. Like birds, one of the most common patterns is for populations to cycle through the winter in tropical and subtropical latitudes, then to move north (or south) during the spring and summer. Some migrants move on the leading edge of storm fronts, where both high wind and high humidity facilitate long-distance dispersal. Occasionally moth numbers may be so great that they show up on radar screens. Swallows, swifts, and other insectivorous birds follow these fronts as well, gleaning insects as they go. A second common pattern is for moths to move between wet and dry forests, with adults trickling out of the latter over the course of the dry season then reinvading with the return of the rainy season. The dazzlingly beautiful sunset moths (Uraniidae) are famous for their mass diurnal migrations in Africa and Latin America.

Key Adaptations

One can only speculate as to why there are so many moths. One morphological feature that stands out as universal is the presence of scales over the wings and body. Those covering the



Figure 13 Tiger moths (Arctiidae) (also see **Figures 12** and **14**). *Utetheisa bella*; in addition to sperm, the male transfers alkaloids during mating that protects not only his mate subsequently, but also her offspring.



Figure 14 *Grammia virguncula*; adults of many tiger moths bubble out hemolymph, presumably laden with noxious substances, from their thoraces when disturbed. Courtesy of Duncan Butchart.

wings are easily abraded, so much so that the wings feel slippery. Regardless of the context in which scales evolved, there can be little doubt that they provide considerable protection from natural enemies: adults can fly into a spider web and, if they are lucky, fly out, leaving behind hundreds of scales still attached to the web. Scales allow the wings to slide past one another and often out of the bill or jaws of a would-be predator.

The coilable siphon or proboscis has allowed Lepidoptera to exploit liquid diets (*see* Adults). More than any other feature, this structure has been tied intricately to the coevolutionary history of moths with flowers. Nearly all insect-pollinated flowers with deep corolla tubes (or nectar spurs) are pollinated by Lepidoptera. Charles Darwin knew this, and on being told of a Madagascar flower with a 25-cm corolla, predicted that a sphingid hawkmoth would one day be found on the island with a tongue of at least this length – and it was – and today bears the name *Xanthopan morganii preadicta*. The tongue also allows the collection of dietary constituents that are scarce in plants, such as nitrogen and sodium. No nonfeeding moths are known to migrate, suggesting that the presence of a tongue is a critical attribute for long-distance movement. Similarly all temperate moths that overwinter as an adult have a tongue.

A tympanum or analogous hearing structure, sensitive to the ultrasonic frequencies used by hunting bats, has evolved independently in no less than five nocturnal moth lineages (Scoble, 1992 and **Figure 15**). It may be no coincidence that three of these (the Geometroidea, Noctuoidea, and Pyraloidea) are among the most species-rich moth clades. Fenton and Fullard estimated that 85% of all macrolepidopterans possess some type of hearing organ. Noctuids that detect ultrasound first attempt to fly away from the source, but if the hunting frequencies are sensed at close range, these moths either drop to the ground or initiate a highly erratic flight while spiraling downward. Further testimony as to the importance of the ear is provided by the mites that infest the tympanal cavities of certain noctuids (especially *Leucania*) (**Figure 15**). Treat found that mites placed onto *Leucania* always move into one ear, leaving the other fully operational.

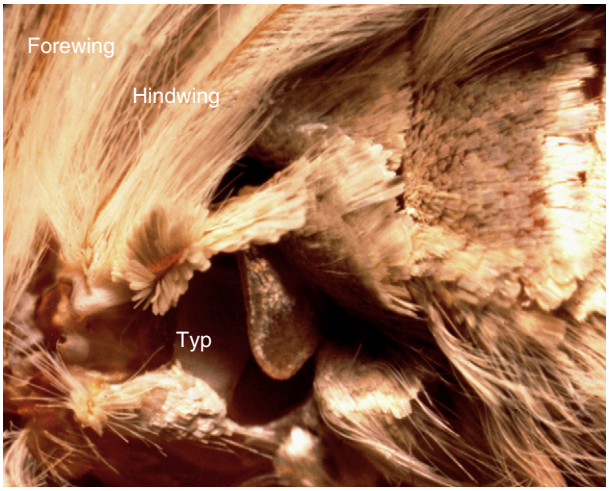


Figure 15 "Ear" of *Leucania* (Noctuidae) located below hindwing (head to right). Courtesy of Asher Treat.

Silk has played a major role in the evolutionary success of the order. Widespread use is seen in the formation of the cocoon or in the attachment of the chrysalis to the pupation substrate. Most microlepidoptera use silk to form a shelter or to line a chamber within the larval feeding substrate. Some caterpillars lay down silk whenever they walk, which presumably allows them to quickly retrace their path, for example, in times of danger. Case making, where the caterpillar constructs a portable silken case, has evolved in no less than a dozen moth groups; two especially successful lineages include the Psychidae and Coleophoridae (both families are appreciably more species rich than their noncase-making sister taxon). Some gregarious caterpillars spin conspicuous nests. Not only do these nests offer protection from natural enemies, but they can also be important in thermoregulation, heating up well above ambient temperatures on cool days (Fitzgerald, 1995). Many species, especially among those taxa with flightless females (e.g., some Psychidae, Geometridae, and Lymantriinae), disperse by "ballooning." Early instars drop from silken lines and wait for winds to blow them about. Among microlepidopterans, it is common to see prepupal larvae dropping from trees on silk lines on their way into leaf litter where pupation will occur. At night many geometrid caterpillars drop from their perches, suspended on a silken thread, and dangle for hours before climbing back onto foliage, always before daylight. In a similar fashion many caterpillars, if disturbed, drop from their perch on a silk belay and crawl back up after the danger has past. Processionary caterpillars (Notodontidae: Thaumatopinae) and tent worms (Lasiocampidae) lay down a silk trail, impregnated with pheromone, that siblings follow. Some microlepidoptera (e.g., Bucculatricidae) form a molting cocoon within which the vulnerable transition between instars takes place. Prior to molting, many species lay down a light sheet of silk into which the crochets are engaged; this allows the new instar to crawl more easily out of its old integument.

One behavioral trait that stands out is the ability of Lepidoptera to exploit the hours of darkness. In terrestrial ecosystems moths are the most diverse and numerically important group of nocturnal flying insects. Perhaps 90% of all Lepidoptera are night active. Associated with their nocturnal habits is the order's remarkable elaboration of morphological, neurophysiological, and behavioral traits that facilitate pheromonal communication (*see* Pheromones).

The extraordinary evolutionary and ecological success of moths may also be a matter of being in the right place at the right time. Because the order was essentially phytophagous and beginning to diversify on nonseed plants and gymnosperms prior to the origin of angiosperms, moths were well poised to ride along with the explosive radiation of flowering plants that was to follow in the Cretaceous. Indeed, the Lepidoptera are the largest single (monophyletic) lineage to have co-diversified with plants (Strong *et al.*, 1984; Menken *et al.*, 2010).

Pheromones

While visual communication is thought to be of modest importance in nocturnal moths, chemical communication is highly developed and quintessentially important to most.



Figure 16 *Hemileuca maia* (Saturniidae); receptive females expose an abdominal gland – here, the small orange structure at the tip of the abdomen, extended nearly perpendicular to the body axis – when "calling" for males. The morphology and location of the sex pheromone gland(s) vary considerably from family to family.



Figure 17 *Antheraea polyphemus* (Saturniidae); male saturniids have strikingly plumose antennae that contain thousands of setae, which capture female sex pheromone molecules.



Figure 18 Male scent-releasing structure of *Trichoplusia ni* (Noctuidae); end view of abdominal brush that encircles the male genitalia. Courtesy of Ken Haines.

The first animal pheromone was identified from the Chinese Silk Moth (*B. mori*) and was appropriately dubbed bombykol. In nearly all nocturnal moths, virgin females release a pheromone that attracts males from distances of up to several hundred meters (reports of up to a mile are likely exaggerated) (**Figure 16**). Males move upwind, in and out of the pheromone plume, until they arrive at its source. The male antenna and associated neurophysiology make up one of the most sensitive chemical detection systems known in nature. Not surprisingly, males often have more elaborate antennae than females, sometimes with secondary or tertiary branches that increase the surface area several fold (e.g., many bombycoids, **Figure 17**). One interesting exception, and one that remains unexplained, are *Arrhenophanes* females (Tineiodea) – the female's antenna is considerably more complex than that of the male.

Female sex pheromones usually are released from glands on the abdomen, often located in the intersegmental regions between abdominal segments seven and eight or eight and nine. The sex pheromone typically contains a "cocktail" of two to several volatile compounds, usually with one or two major components that are responsible for most of the male attraction. The individual components of the pheromone are often odorless. Structurally they tend to have a straight-chain hydrocarbon backbone, one or more double bonds, and an active moiety (e.g., an acetate, aldehyde, epoxide, ketone, or alcohol). Closely related species tend to have different (enantiomeric or isomeric) blends, minor components, or "calling" times.

Once the two sexes are in proximity, a panoply of male courtship pheromones frequently come into play. The structures by which males deploy their pheromones are as varied as any structures in the order. In the simplest case they are merely individual scent scales peppered over the wings or other parts of the body. Male sex scales (or androconia) that are grouped into tufts, brushes, or "pencils" occur over almost any part of the body including the wings, legs, mouthparts, and antennae. They are especially common about the genitalia, where they provide final signals prior to coupling (**Figure 18**).

Remarkably complex scent organs occur in a number of hepialids and various noctuoids (e.g., arctiines and herminiine erebids) (**Figure 19**) and others. A hallmark of androconia is



Figure 19 *Estigmene acrea* (Erebidae: Arctiinae); "calling" male with fully extended coremata (eversible sacs) lined with scent scales. Courtesy of Mark Willis. Reversed calling, with males broadcasting a primary sex pheromone to which females must orient, such as exhibited here by *Estigmene acrea*, is very rare among moths. Courtesy of Mark Willis.

that their abundance and development may differ markedly among closely related species, which suggests that they are under strong sexual selection (and maybe energetically expensive). They tend to be frequent as well as morphologically elaborated in genera where two or more species co-occur, especially if a common host plant is shared.

Male pheromones are thought to serve in species recognition and isolation and to promote female acceptance. In general, they are used at close range in the vicinity of the female. Role reversal, with males calling females from downwind, is very rare but does occur (e.g., see **Figure 19**). Chemically, male pheromones are exceptionally heterogeneous, ranging from simple molecules to long-chain hydrocarbons and ring compounds. Many have a detectable and often pleasant odor. The male pheromone blend of the oriental fruit moth contains both (*Z*)-methyl epijasmonate and (*E*)-ethyl cinnamate; the former is the active constituent in the odor of jasmine, a common ingredient in women's perfumes. Many male scents are based on secondary plant compounds that are sequestered by the larva or collected by the adult. In some arctiines, the female is reluctant to copulate with males that lack appreciable quantities of PAs on their androconia.

Beyond mate location, courtship, and mating, pheromones are also known to serve in a number of lesser roles. Some female moths (e.g., yucca moths) leave behind a marking pheromone at the time of oviposition that alerts other females that the resource is already "in use." The aggregation behavior displayed by some arctiines and noctuids is likely mediated by pheromonal signals. During mating, some male Lepidoptera transfer an "antiaphrodisiac" pheromone to the female that, at least temporarily, repels other would-be suitors.

Coloration, Diurnality, Nocturnality, and Attraction to Light

As a general rule, moths that are active principally at night are subtly colored in earth tones, especially browns, grays, black, and white. The patterning in nocturnal taxa is often variable, presumably because details of the coloration are under relaxed selection pressure. But exceptions to both of these statements abound. Most notably, there are dozens of lineages with brightly colored hindwings that are covered by cryptically colored forewings, with the bright colors and bold patterning being exposed as a flash signal after a moth has been discovered. A principal advantage of such coloration being that the stimulus can be immediately concealed (and therein lost) once the moth alights and again covers the hindwings with the drably rendered forewings. Bright fore- and hindwing coloration is associated with moths that are diurnal or both diurnal and nocturnal.

It is evident that diurnality has evolved repeatedly, perhaps hundreds of times within the order. The most conspicuous example is that of the butterflies, which are, from an evolutionary sense, nothing more than a group (or two) of diurnal moths. Taking the Geometridae as a focal taxon, there are dozens of species and genera that are diurnal; yet there are essentially no species-rich clades (tribes and subfamilies) within the family that are active principally during the day. Diurnality is common among moths that live in alpine or arctic habitats or that fly during the coldest months of the year, as well as that are chemically protected.

From a phylogenetic perspective, diurnality is well distributed across the order. Members of the most basal family, the Micropterigidae, are both diurnal and nocturnal in habit, with some genera exhibiting splendidly metallic golds and reflective purples. These same bright colors are found in a host of other primitive families (e.g., Heterobathmiidae, Eriocraniidae, and Incurvariidae). Like in the Micropterigidae, adults may be active both day and night. Across the Microlepidoptera diurnality is widespread; important diurnal lineages include the Adelidae, Choreutidae, Heliodinidae, Heliozelidae, Scythrididae, and Sesiidae. Many moths are crepuscular, some flying at both dawn and dusk and others limiting their activity to one of these two periods of rapidly changing light intensities.

Nearly all nocturnal moths can be collected at light. Indeed, light traps are routinely employed to conduct moth inventories, to monitor populations of agricultural pests, and to collect gravid females for life history studies. The most commonly asked question about moths is "Why are they attracted to light?" Surprisingly the question has never been answered definitively. One common explanation that has almost no empirical support is the moon-compass explanation. It goes like this: If a moth had a need to fly in a straight line at night it could do so by maintaining a constant angle to any celestial light source. When this same behavior (constant orientation to a distant light source) is applied when a light is proximate, a moth will spiral into a nearby light source (**Figure 20**). And while seemingly logical and compelling, it is rare that a moth spirals into a light – more often the flight path is erratic but directed. An equal mystery is why only a fraction of the moths present in a given area are attracted. Anyone who runs flight interception traps, collects at bait, or knows what caterpillars are feeding in the garden appreciates the fact that the individuals attracted to light represent only a fraction of the whole. Sex ratios in light-collected samples are often heavily biased toward males. Hepialids in the genus *Sthenopis*

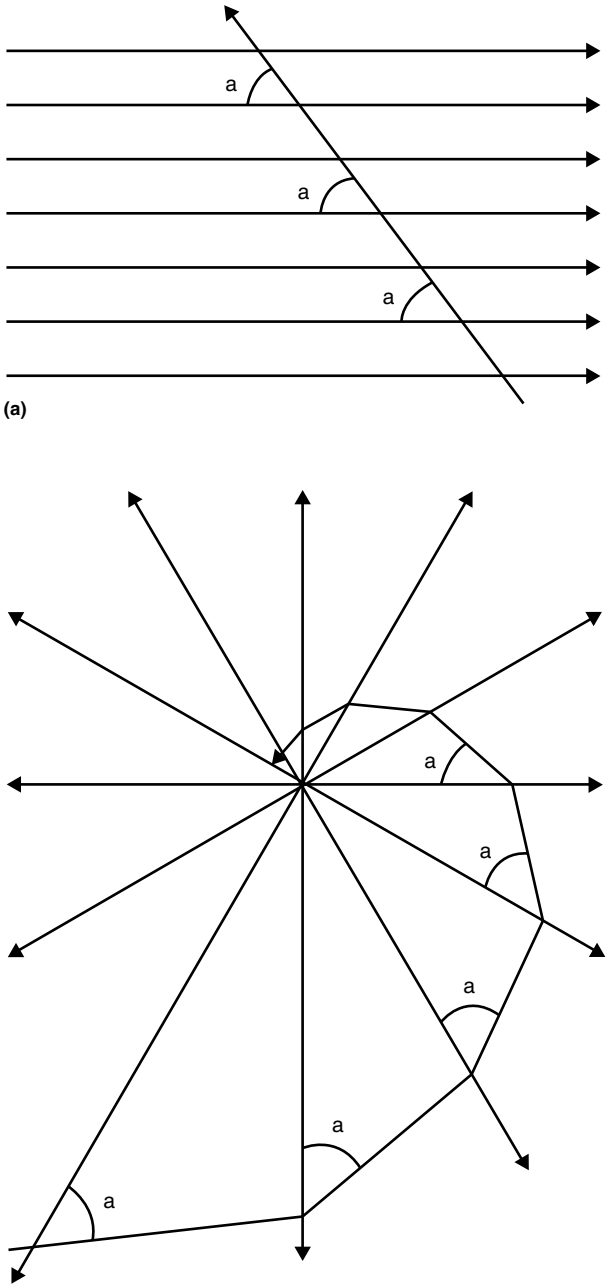


Figure 20 A moth can proceed in a straight line by moving at a constant angle (a) to a distant light source. However, if the light is proximate, (b) this same orientation behavior will cause the moth to spiral into the source. Reproduced from Evans HE (1984) *Insect Biology. A textbook of Entomology*. Addison-Wesley Publishing Company, with permission from Pearson.

provide a counterexample as females predominate in museum collections. There are many moths known primarily from one sex – the sex that is attracted to light. Many of the moths thought to be rare by collectors are simply among those that are not attracted to light with great frequency.

Moths are not entirely dependent on celestial navigation. There is ample evidence that moths also possess a geomagnetic compass that allows them to move in a directed fashion, even on moonless nights and under overcast skies.

Importance

Lepidoptera, in particular moths, are among the most important forest and agricultural pests. There is almost no plant tissue that is not eaten by one or more species of moths. Most of the forest pests are of cyclical occurrence with natural enemies eventually controlling population numbers. Introduced species offer an obvious exception to this generalization – their populations may go unchecked for many years before natural enemies are introduced or recruited from the native fauna. Agricultural pest moths have plagued farmers for millennia. As noted earlier, many are migrants, preadapted to be crop pests (because they are highly dispersive insects that locate resources quickly and thrive in early successional habitats). Additionally, crop pests often have short life cycles that allow them to complete one or more generations before natural enemies can build to appreciable numbers. A species, such as the Indian Meal Moth (*Plodia interpunctella*) (Pyralidae), are important stored product pests that attack grains, cereals, bird seed, dry dog food, and even candy that has sat too long. Clothes moths (*Tinea* and *Tineola* species) (Tineidae) are well known for their ability to damage woolens.

Most moths should be regarded as beneficial principally for their role in terrestrial food webs, especially those that include birds, bats, and rodents. Many songbird nestlings are dependent on caterpillars. With the exception of large, bat-pollinated flowers, most white-flowered plants are pollinated by moths under cover of night. Scents from moth-pollinated flowers may be exceptionally fragrant (e.g., honeysuckle, jasmine, and some orchids). Sphingidae are the primary pollinators for about 6–8% of the tree species in Costa Rica's seasonal dry forests. These large, strong fliers are especially important in tropical forests where individuals of the same species can be widely spaced. Many moth caterpillars feed on leaf litter, decaying wood, and fungi and thus play at least a minor role in litter decomposition and nutrient cycling, such is especially true in the eucalypt woodlands and forests of Australia. The most famous product derived from a moth is silk. This fiber, with a tensile strength greater than that of steel, is still made the same way it has been for more than 4600 years by boiling the cocoons of *B. mori* and unraveling them a single thread at a time. Moths have been used successfully as biological control agents of invasive weeds. Australia had lost 50 million acres of pastureland to introduced prickly pear cactus (*Opuntia*) before the pyralid, *Cactoblastis cactorum*, was introduced from southern South America. Within a decade of its release, most large stands of exotic cactus had been utterly destroyed. The moth also has been effective in controlling invasive *Opuntia* in Hawaii, Mauritius, South Africa, and the West Indies. This story recently has taken an unfortunate twist, as this same moth has found its way into the Florida Keys, where it now threatens stands of three rare (native) species of *Opuntia*. Moreover, the moth has since spread westward to Louisiana and it is feared that it will soon reach the American Southwest and Mexico where *Opuntia* is an ecologically and economically important plant. In Mexico, prickly pear is a

food source for both humans and livestock. In deserts and other xeric communities, dozens of species of vertebrates and invertebrates are reliant on *Opuntia* stands for both food and shelter.

Conservation

Across much of Europe and the USA, lepidopteran faunas are increasingly used as bioindicators by state, provincial, and federal agencies and nongovernmental conservation organizations to evaluate natural areas and management regimes. Moths offer many advantages as bioindicators: most temperate macrolepidopterans are easily identified and their habitat requirements are reasonably well known. By any standard they are hyperdiverse and therefore provide considerable data relative to the amount of effort that goes into their sampling and sorting. Part of their value as bioindicators is that they represent another trophic level – two grasslands that are botanically equivalent would have different moth faunas if one were being managed too intensively (e.g., by fire). One disadvantage of moths as bioindicators is that samples often include many transients. For example, more than one-third of the species of moths that occur in Denmark have been collected in a light trap atop the Zoological Museum on the University of Copenhagen campus at the periphery of the city.

The principal threat to moths is the same as for other terrestrial invertebrates – habitat destruction and fragmentation – this is especially true for tropical forests where moth diversity is thought to be highest. In some ecosystems, reforestation and succession pose threats to moths. Fire suppression in grasslands, chaparral, barrens, scrub forests, and other fire-maintained ecosystems has been detrimental to many local invertebrate populations. And the reverse, ill-timed and too aggressive prescribed burning, can greatly diminish Lepidoptera populations. There has been much discussion about the impacts of artificial lighting on moth populations, much of it focusing on the decline of saturniids and other large moths. Up to this point the published commentary has been anecdotal, but there can be little doubt that street lamps and other high-intensity lamps have had some impact on moth populations – within days, bats learn to forage at them by night, and they are among the first places that birds visit at dawn.

Introduced or alien species can be a threat at many levels. Invasive plants can displace native hosts on which either the larvae or the adults of native moths depend. In exceptional cases, introduced plants can provide strong ovipositional cues for species whose larvae either fail entirely or experience high levels of mortality. Exotic species, be these pathogens or herbivores, that threaten native hosts, can have grave consequences. The loss of American chestnut to an introduced fungal blight is thought to have brought about the extirpation of five native moths. If the emerald ash borer beetle – an introduced cambium feeder from East Asia that kills species of *Fraxinus* – continues unabated, as many as two dozen species of North American butterflies and moths could fall as collateral victims (Wagner and Van Driesche, 2010). Because introduced moths often arrive without natural enemies, their numbers can cause widespread overexploitation of host plants. The widespread defoliation events following introduction of the gypsy moth (*L. dispar*) presumably have had some local impacts on native species. Top-down threats from the release of introduced predators, pathogens, and parasitoids for biological control efforts merit special attention. The tachinid, *Compsilura cocinnata*, was introduced from Europe into eastern North America to control the gypsy moth. The parasitoid's host range includes more than 200 native species. The moth has one generation a year, the fly three to four. It has been estimated that a single acre of a gypsy moth infested forest will yield up to 10,000 adult *Compsilura* in June, which then must seek out alternate (native) hosts.

Although DDT and other pesticides have been implicated in the local extirpation of Lepidoptera, the evidence is invariably circumstantial. Although suppression activities (single applications) would rarely be expected to result in local extinctions, eradication efforts (multiple applications, especially in a single season) certainly impact moths and the insectivorous animals that are dependent on them. The most widely used biological pesticide to control Lepidoptera is *B. thuringiensis* var. *kurstaki* (Btk) – a bacterium that either kills a caterpillar outright or leads to a systemic infection that results in death. Btk is used by home owners to control garden pests, by farmers in agricultural systems, and by government and industry to combat forest defoliators. Product labels indicate that Btk is effective against a taxonomic array of Lepidoptera (and Coleoptera). Susceptibility to Btk ranges from taxa that are not affected to those that are killed weeks after a single application; saturniids and papilionids (swallowtail butterflies) are among the groups reported to be especially susceptible. From one perspective, Btk is a great control agent because it affects few nonlepidopterans and has no mammalian or avian toxicity. Yet from a lepidopterists' perspective it is hard to view Btk in such a positive light. The use of Btk and other pesticides in natural ecosystems is most likely to be a problem over critical habitats that are isolated, restricted, or in some other way imperiled, especially when these occur within a matrix that contains a primary pest. For example, the glades, balds, or barrens communities that occur in an otherwise forested landscape are apt to suffer from the aerial pesticide applications targeted for forest pests. Routine transgenic manipulations in agricultural systems may also prove to have wide-reaching consequences. It appears that no less than three cotton-feeding moths, which originally fed on native cottons, are slipping toward extirpation in North America.

Climate change may be the most significant planet-wide threat to moths, especially to those that dwell in isolated (e.g., islands and mountains tops) or relict (bogs and inland dunes) areas, where successful active or passive movement between suitable habitats patches is unlikely. Significant changes in temperature and precipitation patterns will be a challenge for species worldwide, especially for those taxa that are stenotopic or otherwise occupy specialized niches.

With the exception of the Chinese silk moth, which no longer exists in the wild, there is little evidence that collectors have ever represented a significant risk to moths. Collecting is likely to represent a threat only to highly imperiled and geographically restricted taxa. Where population sizes are very small, the use of light traps that employ a killing agent should be avoided. Paradoxically, a greater threat may be

undercollecting, as it seems probable that only about half of the world's moth fauna has been described. And even in those regions where most of the basic taxonomic work has been completed, it is not until we have extensive and taxonomically reliable occurrence data on a broad geographic scale that conservation biologists will know with certainty which species are increasing, which are stable, and which are slipping toward extinction.

Several British studies report significant declines in moth and butterfly abundance and diversity over the past three decades (e.g., Asher *et al.*, 2001; Conrad *et al.*, 2006). A recent status survey for larger moths in the Netherlands concluded that one-third of the fauna showed signs of decline, even among many common taxa, and suggested that climate change may be at the root of many of the losses (Groenendijk and Ellis, 2011). Across the Northeastern USA and southern Canada moth collectors and taxonomists are reporting declines, especially among Saturniidae, Sphingidae, and Notodontidae. If such declines continue there will be downstream consequences to dependent birds, mammals (especially bats), parasitoids, and sundry others. There is an urgent need for multi-year data sets that can be used to quantitatively document the taxonomic scope, geographical extent, and the ecological magnitude and consequences of moth declines.

Appendix

List of Courses

1. General Entomology
2. Insect Systematics
3. Invertebrate Zoology

See also: Butterflies. Insects, Overview. Invertebrates, Terrestrial, Overview. Parasitoids. Pesticides, Uses and Effects of

References

Asher J, Warren M, Fox R, Harding P, Jeffcoate G, and Jeffcoate S (eds.) (2001) *The Millennium Atlas of Butterflies in Britain and Ireland*. Oxford: Oxford University Press.

Bates HW (1982) Contributions to an insect fauna of the Amazon Valley. Lepidoptera: Heliconidae. *Transactions Linnean Society London* 23: 495–566.

Bowers MD (1993) Aposematic caterpillars: Life-styles of the warningly colored and unpalatable. In: Stamp NE and Casey TM (eds.) *Caterpillars. Ecological and Evolutionary Constraints on Foraging*, pp. 331–371. New York: Chapman and Hall.

Common IFB (1990) *The Moths of Australia*. Melbourne: EJ Brill and Melbourne University Press.

Conrad KF, Warren MS, Fox R, Parsons MS, and Woiwod IP (2006) Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biological Conservation* 132: 279–291.

Corcoran AJ, Barber JR, and Conner WE (2009) Tiger moth jams bat sonar. *Science* 325: 325–327.

Fenton JH, Fullard MB, and Simmons JA (1979) Jamming bat echolocation: The clicks of arctiid moths. *Canadian Journal of Zoology* 57: 647–649.

Fitzgerald TD (1995) *The Tent Caterpillars*. Ithaca, New York: Cornell University Press.

Fox LR and Morrow PA (1981) Specialization: Species property or local phenomenon? *Science* 211: 887–893.

Gaston KJ (1991) The magnitude of global insect species richness. *Conservation Biology* 5: 283–296.

Groenendijk D and Ellis WN (2011) The state of the Dutch larger moth fauna. *Insect Conservation* 15: 95–101.

Heinrich B (1993) How avian predators constrain caterpillar foraging. In: Stamp NE and Casey TM (eds.) *Caterpillars: Ecological and Evolutionary Constraints on Foraging*, pp. 224–247. New York: Chapman and Hall.

Heppner JB (1998) Classification of the Lepidoptera. Part 1. Introduction. *Holarctic Lepidoptera*. Vol. 5(Supplement 1).

Hossler EW (2010) Caterpillars and moths: Parts I & II. Dermatologic manifestations of encounters with Lepidoptera. *Journal of the American Academy of Dermatology* 62: 1–10 13-28.

Kaiser W, Huguet E, Casas J, Commin C, and Giron D (2010) Plant green-island phenotype induced by leaf-miners is mediated by bacterial symbionts. *Proceedings of The Royal Society B* 77: 2311–2319.

Kozek WJ and Rao RU (2007) The discovery of *Wolbachia* in arthropods and nematodes – A historical perspective. In: Hoerauf A and Rao RU, (eds.) *Wolbachia: A Bug's Life in Another Bug (Issues in Infectious Diseases)* Vol. 5. pp. 1–14. Basel, Switzerland: Karger Medical and Scientific Publishers.

Kristensen NP (1984) Studies on the morphology and systematics of primitive Lepidoptera (Insecta). *Steenstrupia* 10: 141–191.

Kristensen NP (ed.) (1998) *Lepidoptera, moths and butterflies. Vol. 1. Evolution, systematics, and biogeography Part 35 Arthropoda: Insecta. Vol. IV. Handbook of Zoology*. Berlin: Walter de Gruyter.

Kristensen NP (2007) Lepidoptera phylogeny and systematics: The state of inventorying moth and butterfly diversity. *Zootaxa* 747: 699–747.

Labandeira CC, Dilcher DL, Davis DR, and Wagner DL (1994) Ninety-seven million years of angiosperm-insect association: Paleobiological insights into the meaning of coevolution. *Proceedings of the National Academy of Sciences* 91: 12278–12282.

Menken SBJ, Boomsma JJ, and van Nieukerken EJ (2010) Large-scale evolutionary patterns of host plant associations in the Lepidoptera. *Evolution* 64: 1098–1119.

Monroe E (1972) Pyralidae, Odontiinae, Glyphrinae. In: Dominick RB, *et al.* (eds.) *The Moths of America North of Mexico. Fascicle 13.1B*. Washington, DC: Wedge Entomological Research Foundation.

van Nieukerken EJ, Kaila L, Kitching, IJ, *et al.* (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang Z-Q (ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa* 3148: 212–221.

Powell JA, Mitter C, and Farrell B (1998) Evolution of larval food preferences in Lepidoptera. In: Kristensen NP (ed.) *Lepidoptera, Moths and Butterflies. Vol. 1. Evolution, Systematics, and Biogeography Part 35 Arthropoda: Insecta. Vol. IV. Handbook of Zoology*, pp. 403–422. Berlin: Walter de Gruyter.

Rawlins JE (1984) Mycophagy in Lepidoptera. In: Wheeler Q and Blackwell M (eds.) *Fungus-insect Relationships. Perspectives in Ecology and Evolution*, pp. 382–483. New York: Columbia University Press.

Scoble MJ (1992) *The Lepidoptera. Form, Function, and Diversity*. Oxford: Oxford University Press.

Stamp NE and Casey TM (eds.) (1993) *Caterpillars: Ecological and Evolutionary Constraints on Foraging*. New York: Chapman Hall.

Stehr FW (ed.) (1987) *Immature Insects*, Vol. 1. Dubuque, Iowa: Kendall/Hunt.

Strong Jr. DR, Lawton JH, and Southwood TRE (1984) *Insects on Plants: Community Patterns and Mechanisms*. Oxford: Blackwell Scientific Publishers.

Wagner DL, Schweitzer DF, Sullivan JB, and Reardon. RC (2011) *Owlet caterpillars of Eastern North America*. Princeton, New Jersey: Princeton University Press.

Wagner DL and Van Driesche RG (2010) Threats posed to rare or endangered insects by invasions of non-native species. *Annual Review of Entomology* 55: 547–568.

Young M (1996) *The Natural History of Moths*. New York: Academic Press.

Museums and Institutions, Role of

Michael J Novacek and Suzann L Goldberg, American Museum of Natural History, New York, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bioinformatics Technology and research aimed at clarifying, coding, organizing, and analyzing biological data through use of computing hardware, software tools, web applications, and theoretical and computational science.
Curation The scholarly study, care, management, documentation, and maintenance of a scientific collection.
Curator A scientist responsible for curation.
Lot A specimen or set of specimens generally stored and studied together, such as individuals of a single species from a single site collected at a single point in time.
Parataxonomist A layperson trained in basic sciences and methods that allow him or her to identify species and higher taxa and conduct biodiversity surveys and inventories.
Phylogeny The evolutionary and kinship relationships of species or the taxa that contain them, usually depicted as branching, tree-like diagrams.
Specimen The whole or a part of an individual organism with a unique data identifier formally added, or accessed, to the institution's collection.
Systematics The study of the evolutionary and kinship relationships of species or the taxa that contain them. Many systematic studies yield results represented by tree diagrams that depict phylogenies (see above).
Type specimen A particular specimen chosen to be the "name bearer" of a species. The characters that diagnose a given species are to be found in the type.
Vouchered specimen A specimen that serves as the basis for extended study and is retained as a reference specimen. A type specimen is a particular vouchered specimen that serves as the basis for taxonomic description and diagnosis of a species. Another example of a vouchered specimen is the one retained as a reference source for tissues used in DNA studies.

Natural history museums (NHMs) and related institutions such as botanical gardens, zoos, and aquaria are devoted to the exploration and study of the biological world, both past and present, and through these efforts provide, with their vast collections, the primary record of life's history and diversity. It has been estimated that some 6700 institutions worldwide serve as the biological repositories for more than 2 billion biological specimens (Duckworth *et al.*, 1993; OECD, 1999; Chapman, 2005; Ariño, 2010). Many of these institutions are also committed to the twin mission of science and education: The research, collecting, curation, and exhibition and educational programs they sponsor serve to enlighten, inspire, and inform a broad public about the richness and wonder of life on earth. In recent decades, these twin missions have taken on a special urgency due to the rapid global degradation and destruction of rain forests, coral reefs, wetlands, and many other natural habitats and the concomitant loss of biological species. That loss, estimated to be at levels 100–1000 times the normal or background rate of extinction recorded in the fossil record (Pimm *et al.*, 1995), constitutes what many biologists call the current biodiversity crisis, an event that possibly could result in loss of anywhere between 30% and 50% of all living species during the twenty-first century, a biological catastrophe approaching the level of the five major mass extinction events in the past 500 million years (Leakey and Lewin, 1996).

Natural history institutions continue to add to their collections and expand on their biological surveys, research, and training. Despite this effort, museum scientists and their colleagues at universities and other institutions still face the enormous challenge of producing a comprehensive census of life's diversity. Between 1.5 and 1.8 million species have been named, described, and classified (May, 1988; Wilson, 2003), but this, by any estimate, is far short of the actual diversity of species living at present. A recognition of our incomplete knowledge of some highly diverse groups such as insects, and our extremely poor knowledge of many groups of invertebrate animals, fungi, algae, plants, and especially diverse bacteria, protists, and other microbial organisms, suggests that the total number of species could range notably upward from at least 10 million species (Wilson, 2003). The challenge of mustering a better accounting of biodiversity is compounded by the current rate of environmental degradation and the loss of species: Many of these species will doubtless be extinct before scientists have a chance to account for them. Thus the work of NHMs and related institutions is critical to our understanding of both the scope and richness of life and the magnitude of the impacts on the natural world resulting from land degradation and loss, climate change, pollution, invasive species, overharvesting, and other drivers caused by human activities.

History: The Emergence of Natural History Institutions

The motivations for establishing museums devoted to natural history are deeply rooted in human culture. The desire to observe, contemplate, and memorialize nature is represented in some of the oldest evidence of human thought and expression, including the elegant wall paintings of animals in the 17,000-year-old paintings in the caves of Lascaux. In the less-distant past, the emergence of a place for scientific collections, study, and enlightenment is certainly apparent. Aristotle's "peripatetic school," formally established in 335 BC at the site of the even more ancient Lyceum, was the base camp from which the philosopher and his students roamed the world collecting animals, plants, and other natural objects (Ross, 1995). Aristotle indeed intended to catalogue the whole of life's diversity, a task whose difficulty, as we know in hindsight, far exceeded his enthusiasm for it.

Yet such motivations in the ancients clearly inspired what followed. History is replete with entities that share many of the qualities and objectives of what we recognize currently as NHMs, botanical gardens, and zoological parks. Although such establishments were not, until more recently, characteristic of eastern societies such as those in India and China, their rise and proliferation in western societies coincided with the cultural fruition spawned by the Renaissance that ushered in the Age of Enlightenment. The transition from the seventeenth century inventions of Galileo, including his improvements on the telescope and his design and fabrication of the geometric compass and the thermoscope, to the eighteenth century refinements of the compound microscope, thermometer, and barometer is vividly represented in original collections of the newly renovated and reopened Museo Galileo in Florence.

The European Model for NHMs

Throughout the eighteenth and nineteenth centuries, the great centers for natural history collections were established to provide a home for the specimens avidly collected from the far corners of the world. In most instances, the passions for collecting by wealthy, influential, and in many cases, aristocratic individuals led to the formation of societies or academies devoted to collections and research. Such developments anticipated, sometimes by many years or decades, the actual acquisition of a facility and the attainment of a scientific program adequate to care for such collections (for a review, see Kohlstadt and Brinkman, 2004). The British Museum, established in 1753, housed (in fact in a private house) the natural history collections that were originally amassed by Sir Hans Sloane, who acquired specimens from many of the British colonies and other parts of the world. Within a few decades, the British Museum could not accommodate the rapidly accumulating collections, a period accompanied by a notorious phase of disarray, where specimens were lost, sold, and even burned in bonfires. Order was imposed by the renowned comparative anatomist and paleontologist Sir Richard Owen, who was appointed superintendant of the natural history departments in 1856. Owen spearheaded the effort to move the collections to a new, appropriately designed building. The resulting grand Romanesque structure, the NHM in South Kensington, opened in 1881. Kew Gardens (or more formally the Royal Botanic Gardens, Kew), established in 1759, grew from the exotic gardens of Lord Capel John of Tewksbury into the world's largest collection of living plants. Some of Kew's most important additions, such as the iron and glass-clad Palm House, came many decades later. Complementing these developments was the establishment of the Oxford University Museum of Natural History in 1860.

Other major museums on the European continent were also established through a prolonged period of gestation. The Muséum national d'Histoire naturelle (MNHN) in Paris was formally founded in 1793 during the French Revolution but was a product of a much earlier, venerable establishment, the Jardin royal des plantes médicinales (Royal Medicinal Plant Garden) created by King Louis XIII in 1635. During much of the eighteenth century the Jardin des plantes was under the directorship of the preeminent naturalist George-Louis LeClerc, Compte de Buffon. The later-established Muséum national d'Histoire naturelle was also home to such influential scientists as Georges Cuvier, Jean-Baptiste Lamarck, and Geoffrey Saint-Hilaire. The largest NHM in Germany, the Museum für Naturkunde, Berlin, associated with the great eighteenth century German explorer and naturalist Alexander Von Humboldt, was established in 1810, but many of its collections date back to the 1700s. More a product of European, rather than Asian, influences is the Zoological Museum of the Zoological Institute of the Russian Academy of Sciences in St. Petersburg. The Zoological Museum is rooted in developments that took place as far back as 1724 and at present functions as one of the world's largest NHMs.

Museums of the Americas

In the western hemisphere, the transition from naturalists' interest to the formation of societies that ultimately sponsored and promoted the establishment of museums was very much in the European mould (Kohlstadt and Brinkman, 2004). Indefatigable collecting and a passion for collections led Charles Wilson Peale and his colleagues to ensure that the comparatively young USA had its own NHM. This effort resulted in Philadelphia's Academy of Natural Science, founded in 1812 but not rehoused in a proper building until 1840 (and later in a building opened in 1876). During the 1850s, massive collections of specimens were stored in the Smithsonian's Castle building in Washington, DC but were not transferred to a newly constructed National Museum of Natural History (NMNH) until 1881. Similarly, the American Museum of Natural History (AMNH) in New York founded in 1869 moved in 1877 from its crowded, temporary home in Central Park to a newly constructed, fireproof building adjoining the Park. Soon to follow were the New York Botanical Garden (NYBG) and the Bronx Zoo (later under the management of the Wildlife Conservation Society) opened in 1891 and 1899, respectively, on land purchased from Fordham University.

Beyond the eastern seaboard, institutions sprouted with American expansionism. The Missouri Botanical Garden, St. Louis, the oldest in the USA, was established in 1859 to encompass its ancestral home, the original estate of Henry Shaw, botanist, and philanthropist. The Field Museum in Chicago founded in 1893 first occupied the Palace of Fine Arts, a magnificent structure built as part of the "White City" for the Chicago Columbia Exposition (1893) but one wholly inadequate for housing museum collections. It was not until 1921 that the Field Museum relocated in its present habitat, the massive neoclassical building on Lake Shore Drive. San Francisco's natural history museum, the California Academy of Sciences (CAS), was founded as a society in 1853 and located in a commercial building whose rental income allowed Academy members to build their own facility in 1891. The latter was destroyed in the great fire of 1906 and the Academy moved to its present location in Golden Gate Park. Its facility there was succeeded by a strikingly modern, environmentally friendly, and energy efficient building that opened in 2008.

These institutions were mainly independent entities, meant to mark the cultural attributes of a city or a particular region. Nonetheless, some of the most important natural

history museums and botanical gardens are actually a part of university campuses, including Harvard's Museum of Comparative Zoology (established, 1859), Harvard's Arnold Arboretum (1872), Yale's Peabody Museum (1866), and The University of California at Berkeley Museum of Vertebrate Zoology (1908).

Outside the USA, the roster of North American museums would not be complete without the mention of Mexico's very important Museo de Historia Natural de la Ciudad de México, Mexico City (MNH). Although the newest instantiation of the MNH dates back to only 1964, when it expanded its active research and education programs in biological subjects, this museum's origins can be traced back to 1790, as one of the Americas' first institutions devoted to curatorial research in natural history. At higher latitudes, Canada has had a long commitment to natural institutions. Among the oldest and most prominent of these are the Royal Ontario Museum (ROM) in Toronto, established in 1912, and the Canadian Museum of Nature in Ottawa, launched with the construction in 1912 of a famous gothic revival structure, the Victoria Memorial Museum Building.

This small sample of established museums and related institutions in North America should not exclude other major museums of the western hemisphere. Some of the latter were indeed established concurrently or prior to those mentioned above. Prominent among South American museums are the Bernardino Rivadavia Natural Sciences Museum, Museo Argentino de Ciencias, Buenos Aires (established in 1826, with the first of its current buildings opened in 1929), Museo del Mar, Mar del Plata (1888), Argentina, and the Museu Nacional, Rio de Janeiro, Brazil (1818).

Museums in Africa, Oceania, and Asia

Outside of the Americas, many museums were originally products of colonialism, subsequently reshaped and renovated with new programs after their home countries became independent. Examples include the Nairobi National Museum of Kenya, anticipated by a modest two-room building opened in 1911, moved to the new Coryndon Museum in 1930, and, following the country's independence, renamed and reconceived as the National Museum in 1963, attracting its famous director Richard Leakey in 1968. South Africa has several important natural history institutions that currently carry on very active research, including the Ditsong NMNH (formerly Transvaal Museum) in Pretoria, South Africa, founded in 1892. Other African natural history museums, in various states of modernization and improvement, can be found in Ethiopia, Egypt, Botswana, Tanzania, and Zimbabwe.

In Oceania, the most expansive museums are located in Australia and New Zealand. Australia's oldest, the Australian Museum in Sydney, founded in 1854, currently supports very active research – including new research centers devoted to biodiversity and conservation – as well as collections growth in keeping with Australia's particular emphasis on studying and sustaining its remarkable, and in many cases unique, biodiversity. But that country can claim a number of important institutions, including the Museum Victoria in Melbourne (founded in 1854), the Queensland Museum in Brisbane (1862) and several others. New Zealand's major institution covering its equally remarkable natural history is the Museum of New Zealand Te Papa Tongarewa, located in Wellington. Its ancestral home was the Colonial Museum founded in 1865, which expanded and moved during subsequent decades culminating in the modern Te Papa that was established in 1992. The Te Papa, which can claim among its arresting exhibits the world's largest specimen of the rare colossal squid (*Mesonychoteuthis hamiltoni*), is particularly notable for integrating themes relating to natural history and biodiversity with the traditions, perspectives, and practices of New Zealand's indigenous peoples.

As noted, Asia historically lacked the longstanding traditions that gave rise to the natural history institutions in Europe, North America, and elsewhere. For example, one of the continent's principal museums, the Beijing Natural History Museum (BMNH), was founded in 1951. Nonetheless, Beijing's museum and many others were direct descendents of China's research institutes established several decades earlier. Indeed, some of Asia's prominent museums, such as the National Museum of Nature and Science in Tokyo, Japan took form as early as 1871. There are a number of other important institutions in Japan as well as other countries that include Israel, India, Malaysia, Pakistan, Philippines, Singapore, South Korea, Taiwan, Thailand, and Jordan. Notable is the growth of relatively new museums as part of the cultural and economic surge in some of the Arab countries. Examples include the Qatar National Museum in Doha and the Sharjah Natural History Museum and Desert Park in the United Arab Emirates.

The Evolution of the Museum Mission

The eighteenth and nineteenth centuries thus marked the emergence and rise of the natural history institutions that still maintain major collections, support active research, and attract tens of millions of annual visitors. Yet the conditions and the premises under which these institutions were originally established were radically different from what they are at present. Although by 1800, the global human population reached 1 billion, the planet still seemed an expansive, mysterious, even intimidating place. Huge areas were poorly known or not known at all. Antarctica, an entire continent larger than Europe and only slightly smaller than South America, was not seen by humans until 1820. Not until well along in the nineteenth century did western scientists encounter, describe, and name the Sumatran rhino, the blue whale, the two living species of tree sloths, the wombat, the pigmy hippo, the spider monkey, and huge numbers of insects and plants. Museums, botanical gardens, and zoos thus sponsored elaborate expeditions to exotic locales, the essence of the so called "Golden Age of Exploration" that marked the 1800s and carried on well into the twentieth century.

Yet even in these early times, museums, explorers, and collectors began to reflect on the diminishing populations of wildlife and the encroachment on natural habitats by people and cities. Several of the great diorama halls, such as the Akeley Gallery of African Mammals in the AMNH in New York that opened in the 1936 (**Figure 1**), were not only a testament to the grandeur of wildlife but were also intended to



Figure 1 A museum's homage to a vanishing world: The Akeley Hall of African Mammals at The AMNH in New York.

memorialize natural settings destined to disappear in the decades ahead. Museums indeed early on fostered a conservation ethic, one, for example, that embodied the concepts and actions of none other than Theodore Roosevelt, the 26th president of the USA, a Harvard graduate with a strong interest in and knowledge of biology and a life long connection to the AMNH and other like institutions. This legacy of environmentalism set the foundation for the expansive research, educational, and exhibition programs in biodiversity and related themes carried out by present-day museums and other collections-based institutions.

Collections

The organization of institutional collections follows several prescriptions. A fundamental unit is of course the specimen, the whole or a part of an individual organism, living or preserved, fossil or recent, with a unique data identifier formally added, or accessed, to the institution's collection. Specimens vary in completeness and mode of storage according to the nature of the material. A specimen may be a single fossil tooth of an individual dinosaur, a herbarium sheet, a sample of frozen tissue, a pinned butterfly, or each tiny fly in a vial holding hundreds of flies. Specimens are often grouped in a lot, a specimen or set of specimens generally stored and studied together, such as individuals of a single species from a single site collected at a single point in time.

The Size and Scope of Natural History Collections

The accumulated collections of the world's natural history institutions and related biological repositories are of course enormous in size but for a variety of reasons it is difficult to reliably estimate this figure. Collection records vary in precision from institution to institution and discipline to discipline. Although formal accession records and actual specimens are interchangeable in many herbaria, this is often not the case with zoological collections. Moreover, the very large collections of highly diverse groups such as insects are customarily catalogued with reference to the published literature rather than by individual specimens. Notwithstanding this variation in curatorial methods, total holdings are reasonably well known for some of the major museums (**Figure 2**). Unfortunately, this is not the case for a large number of other institutions, including those in developing countries, where resources for collections-based institutions can be very limited. A mandatory register of biodiversity-related collections does not yet exist.

Well known estimates for total worldwide collections size, primarily in museums, that account for animal, plant, fossil, and mineral specimens range between 2.5 and 3 billion (Duckworth *et al.*, 1993; OECD, 1999; Chapman, 2005). Newer data sources, such as the Global Biodiversity Information Facility (GBIF) (2003), record lower figures, ranging from 407 to 737 million in holdings, but many data rich institutions are not yet members of GBIF. The application of probabilistic and distribution models and case studies to data derived from GBIF and other sources, such as the Biodiversity Collections Index, yields estimates of total worldwide biological collection size ranging between 1.2 and 2.1 billion units, with a unit equaling a single specimen or a group of specimens registered as a single data record and treated as a whole (Ariño, 2010).

Relative biological collection sizes for 10 major museums, which include among them the largest such institutions (**Figure 2**), represent in aggregate more than 426 million biological specimens (both fossil and extant), and thus a large portion of worldwide holdings. With respect to extant species, all of these collections, and those of many other museums, share some of the same characteristics (**Figure 3**). By far the

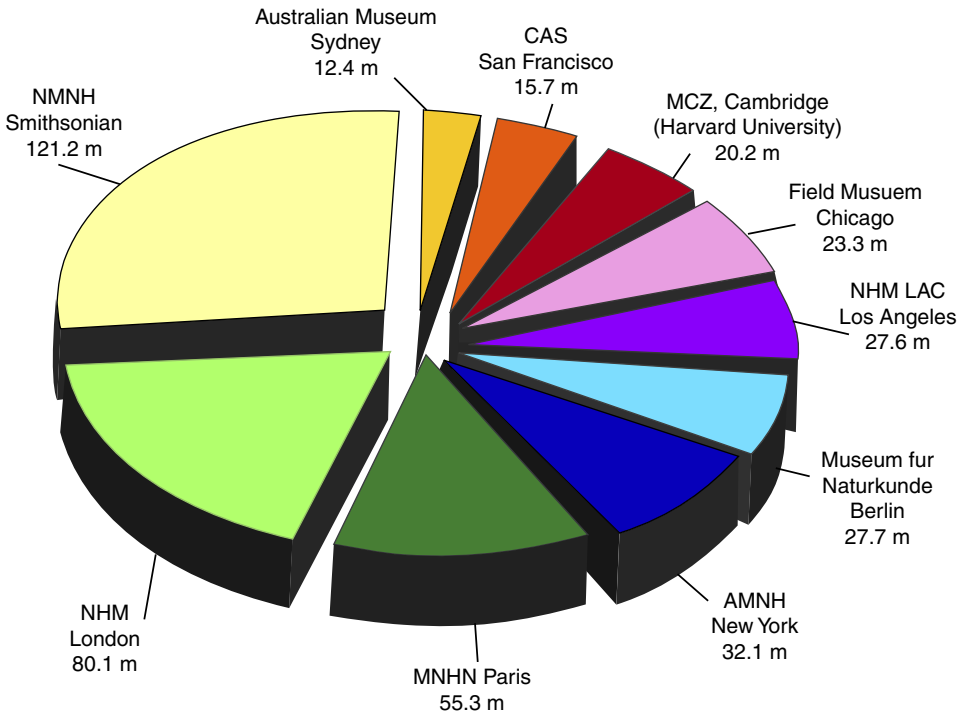


Figure 2 Total size (in millions of specimens) of biological collections (extant and fossil) of several major natural history museums. Abbreviations are: AMNH, American Museum of Natural History; CAS, California Academy of Sciences; MCZ, Museum of Comparative Zoology; MNHN, Muséum national d'Histoire naturelle; NHM, Natural History Museum; NHM LAC, Natural History Museum, Los Angeles County; NMNH, National Museum of Natural History. Data are from institutional websites accessed November and December, 2011 updated through direct communications with collections personnel at several institutions.

largest portions of collections in terms of numbers of specimens contain insects and other invertebrate groups. This fittingly reflects the vastly greater diversity and abundance of these species in nature, although based on this criterion one might judge most museums as having disproportionately large collections of vertebrates.

In other respects, collections in museums can differ markedly. Thus museums such as the Smithsonian, Field, California Academy, and London have substantial collections of plants, whereas others, such as New York's American Museum, do not (**Figure 3**). In some cases, this is due to the presence in the same region of sister institutions, such as the NYBG. The latter and other major botanical gardens, such as Missouri Botanical Gardens, Saint Louis, and Kew Gardens, London, combine preserved plant collections (herbarium sheets of plant parts stored in drawers and cabinets) with elegant collections of living plants more familiar to the public. A comparison of collections for some of the major botanical gardens is shown in **Figure 4**. Zoological parks as well often incorporate notable collections of living plants on their grounds. The San Diego Zoo, for example, is also an officially registered botanical garden, with representatives of more than 6500 plant species.

Most of the figures given above refer to the more traditional collections – namely skins, skeletal elements, shells and other hard parts, whole bodies preserved in alcohol, pinned insects, pressed or living plants, and related specimens. In recent years, a very important new kind of collections has been added to this list. Frozen tissues supercooled in liquid nitrogen represent a significant new component of collections at certain institutions. Such frozen, or cryogenic, facilities provide tissues highly suitable for sampling deoxyribonucleic acid (DNA), ribonucleic acid (RNA) (where tissues are fresh and rapidly frozen), and proteins. With the proliferation of comparative genetics and genomics at many collections-based institutions, these cryogenic collections have become a core resource for modern collections-based research. Such collections are already substantial and growing at several institutions, including the Museum of Vertebrate Zoology at Berkeley, the Australian Museum in Sydney, the Louisiana Museum of Natural History at LSU, and the AMNH in New York. At AMNH, the cryogenic collection as of 2011 contained 71,000 specimens and is rapidly increasing to fill a facility designed to house more than a million specimens. Other institutions, such as the Smithsonian's NMNH, are launching major efforts to build facilities for this purpose.

Another noteworthy category, culture collections, not typically held by museums and gardens are important resources provided by centers and laboratories specially designed to hold them. Notable here are the American Type Culture Collection and the Provasoli–Guillard National Center for the Culture of Marine Phytoplankton, also in the USA along with satellite collections in developing countries (see also Henderson and Chalmers, 2001). Finally, a few institutions have important seed banks, collections extremely relevant to biodiversity related strategies as they maintain storage for live seeds that can be used to propagate endangered or even plant species recently extinct in the wild. It is estimated

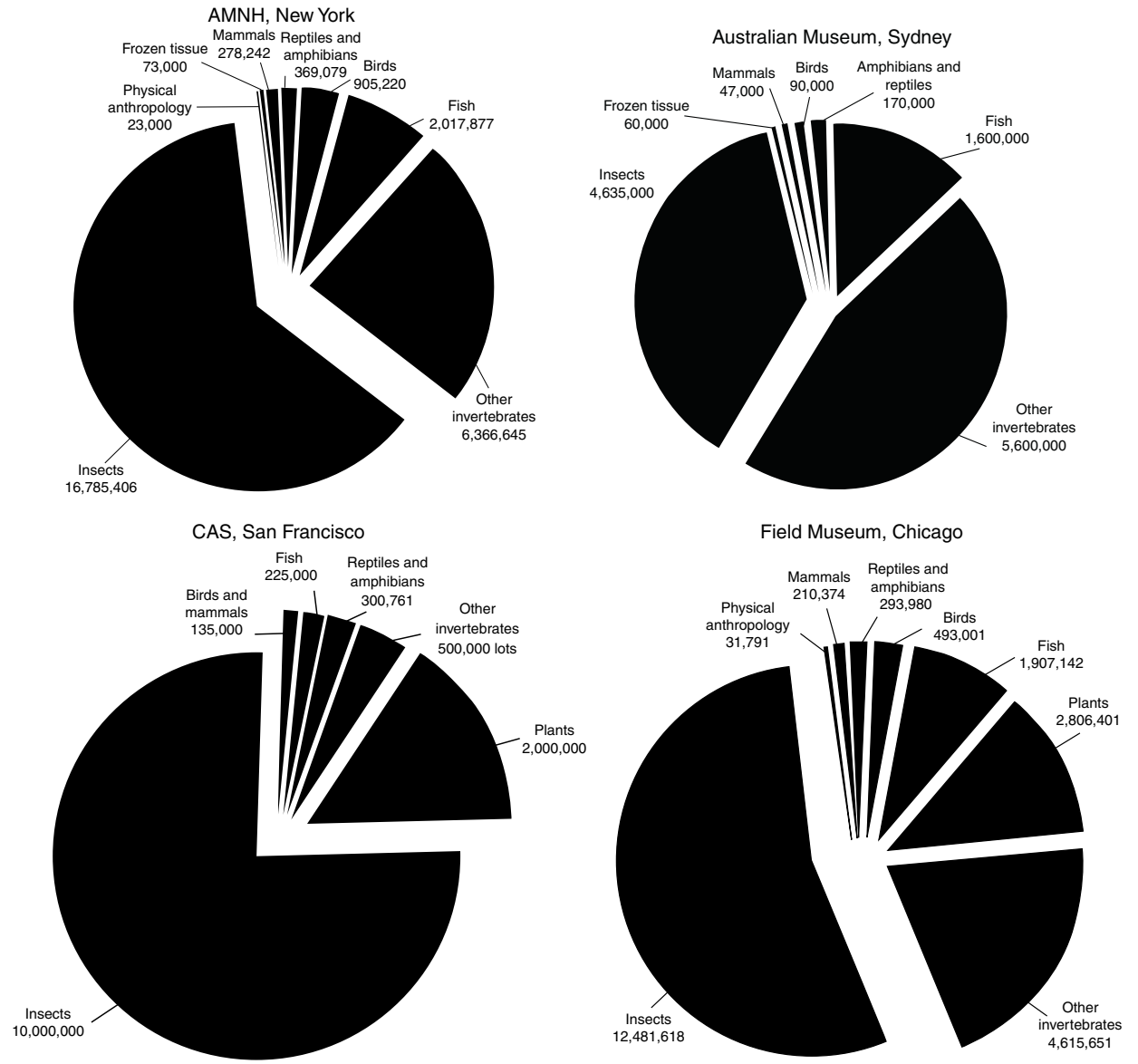


Figure 3 Total size of categories of biological collections (excluding fossils) for several major natural history museums. Abbreviations as in **Figure 2**. Data are from institutional websites accessed November and December, 2011 updated through direct communications with collections personnel at several institutions.

that between 60,000 and 100,000 plant species are threatened or highly endangered (Raven, 1987). Kew Royal Botanic Gardens maintains the world's largest seed bank through The Millennium Seed Bank Partnership (MSBP), a network of partner institutions and researchers across 50 countries. As of 2011, MSBP has successfully banked 10% of the world's wild species with a goal of reaching the 25% mark by 2020. Samples of this collection were used in a striking installation – the "Seed Cathedral" – designed by Heatherwick Studio for the UK Pavilion at the Shanghai World Expo 2010.

Because many museums and gardens are well more than a century old and their collections developed early and rapidly, it might be assumed that current collections growth is eclipsed by that resulting from the initial feverish period of collecting during the "Golden Age of Exploration." Although the more charismatic species of large mammals and other vertebrates, many of which are highly endangered, are not routinely collected at present, museum holdings are still growing at an impressive rate. On average, the AMNH, for example acquires more than 90,000 specimens a year – many of them are insects, spiders, and other arthropods – stemming from the work of more than 125 expeditions annually. Other institutions with comparable expeditionary programs show comparable growth.

Collections Digitization and Databasing

One advantage that comes with new collecting is that the incoming specimens are routinely recorded and catalogued into a highly usable digital database. However, retrospective

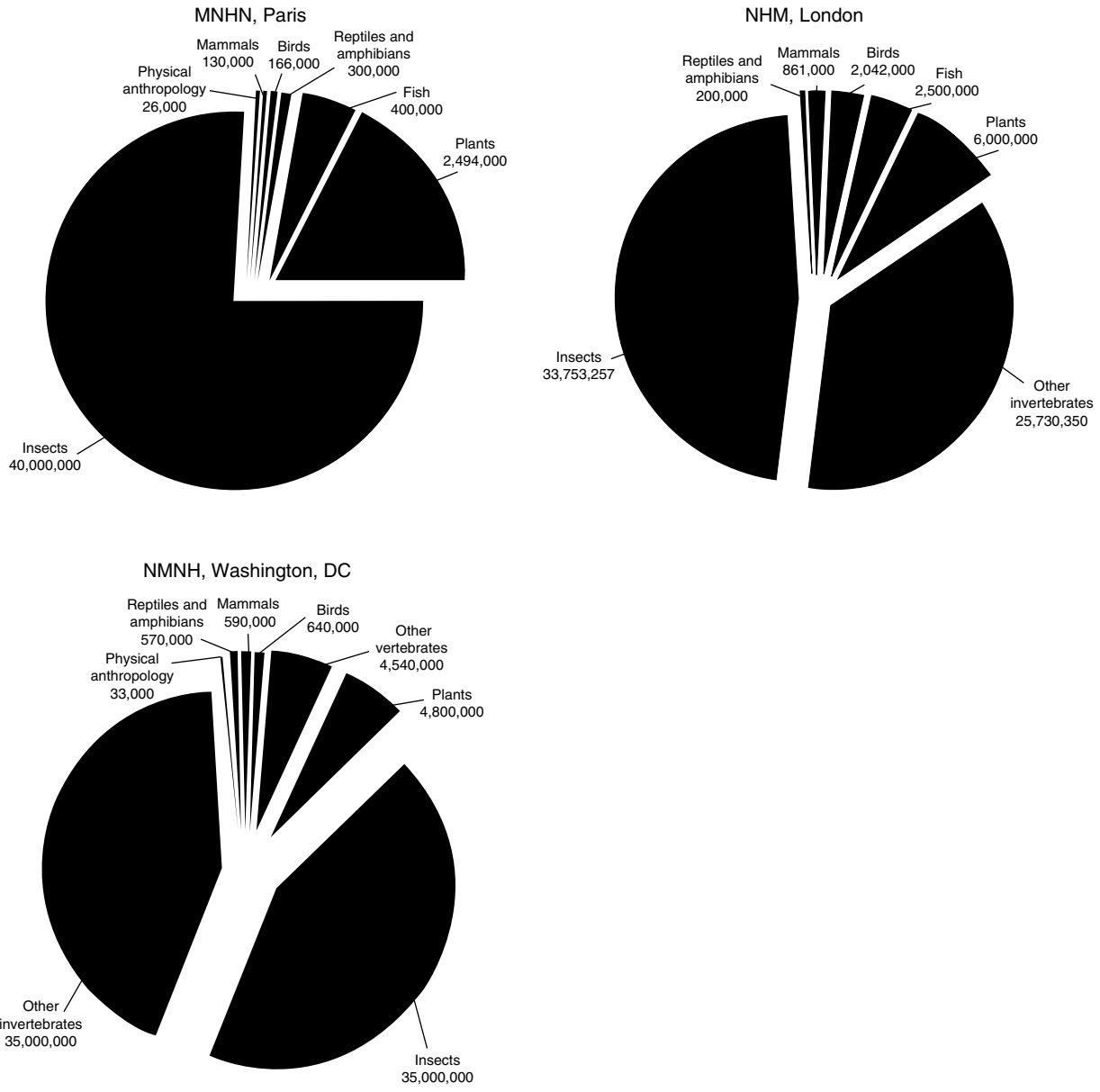


Figure 3 Continued.

cataloguing of older collections is an immense and expensive task, and there is an intimidating backlog of records awaiting digitization. Of the 1.2–2.1 billion units estimated for total collections worldwide, a mere 3% at most are web-accessible through resources like GBIF (Ariño, 2010). Insects represent a particularly monolithic problem as the international community of specialists have traditionally tied records of these huge collections to published literature rather than on a per specimen or per lot basis, and there are as yet no agreed on standards for retrospective data capture. On other fronts there has been some progress. As of 2011, AMNH New York had completed digitization of approximately 86% of its 3.6 million specimen collection of extant vertebrates, but very low percentages for its insects and other invertebrates. This institution as well as others, including the Smithsonian NMNH, are also migrating digital records to a single shared platform (at present the KE EMu system), a move that will enhance content management and data exchange within and among institutions.

The continuing progress toward collections digitization is critical in making these resources more accessible and usable, but it has come with some challenges. Prominent among these is the problem of data storage. Simple text and numerical records are for the most part manageable, but very large amounts of digital images of specimens or parts of specimens, especially the high resolution 3D-images and animations based on computed tomography (CT) scans and related resources, are placing very high demands on server storage (Rowe and Frank, 2011). In fact, the terabyte storage requirements for such image-based data now far outstrip those for

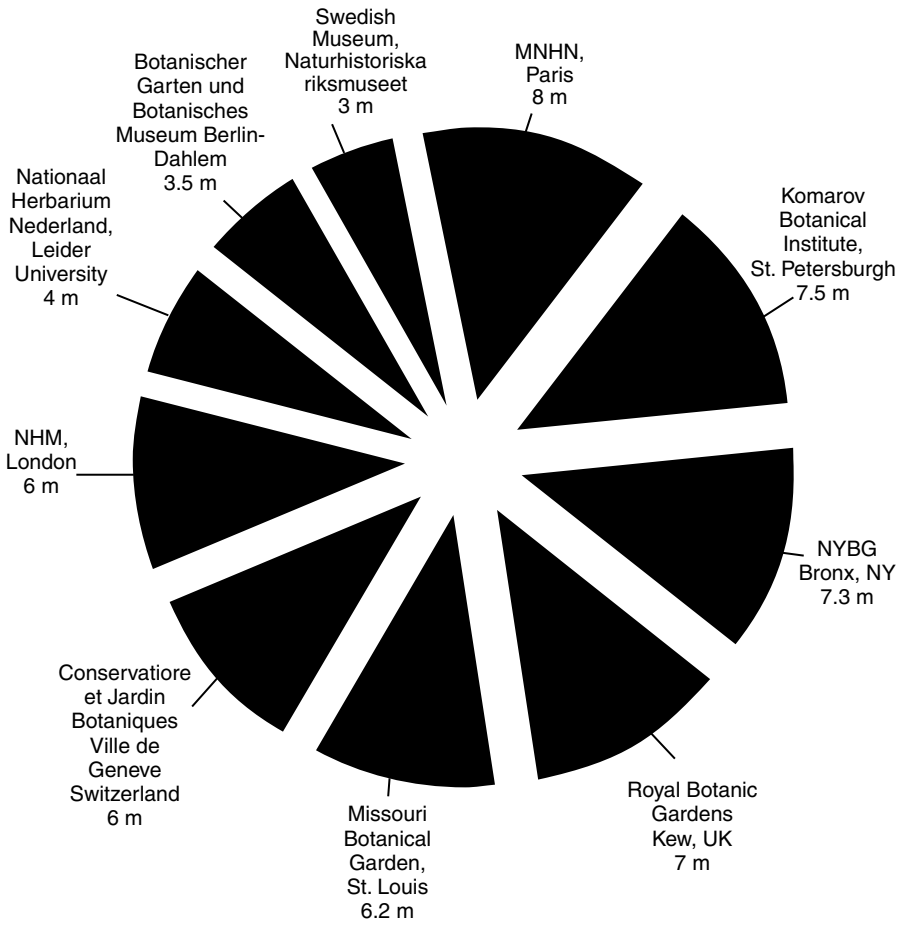


Figure 4 Total size of botanical collections in several major natural history museums and botanical gardens. Abbreviations are: MNHN, Muséum national d'Histoire naturelle; NYBG, New York Botanical Garden. Data are from institutional websites accessed November and December, 2011 updated through direct communications with collections personnel at several institutions.

DNA sequence, amino acid, and other types of molecular data. Digital storage now and going forward is a problem facing the whole museum community and it will require a more concerted effort to find solutions by networked institutions both at national and international levels.

Collections Management and Conservation

Notwithstanding new approaches and technologies, the key factor in ensuring the quality and effectiveness of the natural history collections remains the staff responsible for its management, maintenance, conservation, and use. Here the core positions are the curators and collections managers with expertise relevant to a particular area of the collections. Since collections vary considerably not only in the nature of their taxonomic content but also in the means of their storage preservation and organization, they require a diversity of persons with special expertise and skills. Nonetheless, there is a recent trend toward developing a more hierarchical collections management team within a logically defined area of the collections. For example, there may be one or more curators and one or more collections managers for a particular collection, such as birds or mammals. However, the activities of the collections managers for each of the vertebrate groups can be coordinated by one or more supervising collections managers who report to the various curators responsible for vertebrate collections. Such organization is especially useful where there are shared methodologies and activities, such as the digitization of collections records, the mechanics of loan processing, or the protocol for delivering material for common frozen tissue collections.

Another important function that can be applied across many disciplines and collections is the one involving preparation and conservation of specimens. In recent years, there has been an increase in focus on natural history collections and the special problems they pose for conservation and risk management. This has promoted development of storage facilities that are more temperature- and humidity-controlled and are more effectively sealed from invasion by insect pests. Techniques that call for integrated pest management, involving specimen preservation and optimal cabinetry design are better and safer alternatives to the older and now diminishing use of par-Dichlorobenzine (p-DCBs) and other problematic chemicals to control pests. The First World Congress on the Preservation and Conservation of Natural History Collections held in Madrid in 1993 catalyzed an increase of published studies and activities by such organizations as the Society for the Preservation of Natural History Collections, now in its 27th year.

Museum Research: Its Role in Understanding Biodiversity

Taxonomy and Systematics

Natural history collections of the world, both large and small, comprise the source for the basic science that provides the essential information on present and past biological diversity. Newly discovered species are named and described as part of the discipline of taxonomy, and their evolutionary relationships with other known species are identified in the approach known as systematics. At present, systematic biologists base their studies on genomic data (arrays of genes or whole genomes) and phenomic data (features extrinsic to the genome that include proteins, cell structure, anatomy, and many other characters). Such research leads to the development of phylogenies, tree-like maps that depict the branching relationships of species and higher taxa, and the establishment of biological classifications that array living and fossil species into formally named, more inclusive groups. Natural history museums, botanical gardens, and related institutions represent the greatest concentrations of experts devoted to taxonomy and systematics. Such work and other museum efforts in fields such as biogeography, climate study, paleontology, geology, ecology, molecular biology, and conservation biology provide an important scientific foundation for efforts to conserve and manage biodiversity and contribute to a wide range of applications in animal and human health, agriculture, pharmaceuticals, bioprospecting, resource management, and government policy and practice.

A key practice in taxonomy is the establishment of a type, a particular specimen chosen to be the "name bearer" of a species. The characters that diagnose a given species are to be found in the type. Taxonomy involves a considerable amount of revision, wherein a named species may be found to be indistinguishable from an earlier named species by comparing the types of the two species. Conversely, a new species may be recognized or "split off" from what was once regarded as only a single species because one or more of the specimens referred to the original type show significant differences. The new species will then have its own new type specimen. Such types are extremely important components of any collection. Sometimes even smaller museum collections can be very important resources because they hold a very large number of types relative to the total number of specimens in the collection. Moreover, type collections are often provided special high security locations and are subject to more stringent or even restrictive loan policies.

Another class of specimens important for collections-based research is a vouchered specimen, namely one chosen as the basis for extended study and retained as a reference specimen. A type specimen is a particular vouchered specimen that serves as the basis for taxonomic description and diagnosis of a species. Another example of a vouchered specimen is one retained as a reference source for tissues used in DNA studies. A vouchered specimen may also be one chosen to denote the occurrence of a particular species in a particular geographic locality. Thus, museums provide an important service in housing vouchered specimens for various purposes and are critical to areas such as comparative genetics, biodiversity surveys, and conservation biology.

Genetic, Genomic, and Phenomic Research

Traditionally, specimens and the species they represent have been described and diagnosed on their extrinsic characters – the leaf venation of plants, aspects of internal and external skeletons, hair and feather color, scale number, size and shape measurements, and other features of great variety. For some decades, studies at the molecular level have played an increasing role. Molecular systematics came into play when genetic data could be used to separate groups of individuals or local populations belonging to one species that could not be distinguished by their extrinsic characters. These sibling species, or cryptic species, are thus only distinguishable at the genetic level. A later development is the use of DNA barcoding, wherein slight differences in the order of sequences in a single gene or a few genes are the basis for recognizing different local populations, subspecies, or species (Hebert *et al.*, 2004). DNA barcoding relies on genes with an apparently rapid mutation rate, such as the mitochondrial gene cytochrome *c* oxidase I in animals and chloroplast genes *rbcl* and *matk* in plants. Such genes show high variability even among closely related species. DNA barcoding has been widely adapted and is often applied to the recognition of new species that warrant protection as a conservation priority. Barcoding as well has other conservation uses, such as in identifying the source of illegally transported items such as rhino horn, elephant ivory, or tiger paws, or the true species used as a source for commercial products like lumber, caviar, and tuna. Several museums foster programs that are active in DNA barcoding and its applications.

Genetic approaches, such as DNA barcoding, have been criticized as overly reliant on such a small unit of the entire genome for drawing important taxonomic decisions (Tautz *et al.*, 2003). Earlier phylogenetic studies that relied on DNA sequencing also faced similar criticisms because they accounted for only a few genes out of the tens of thousands of genes in any given genome. The dramatic technological improvements in high throughput DNA sequencing have, however, obviated some of these issues. The cost and time necessary for sequencing large portions of the genome and even whole genomes has decreased dramatically in only the 10 years since the publication of the first sequence of the human genome in 2001 accomplished at costs greater than $100 million (Metzker, 2010). As of June 2011, it was possible to sequence the human genome comprising 25,000 protein-coding genes and 3 billion nucleotide bases for $10,000 in direct costs (Wetterstrand, 2011 National Human Genome Research Institute (NHGRI) website). It is conceivable that in the near future whole genome sequencing will become so automated, efficient, and inexpensive that many museum specimens will be routinely sequenced in this way. In the meantime, most of the current phylogenetic studies on more complex organisms, such as animals and plants, with larger genomes (10–100 billion base pairs) are being carried out at subgenomic levels. For example, a recent study of mammalian phylogeny compared sequences of 26 gene fragments in 164 living mammal species and five vertebrate species outside mammals (out-groups) using newly generated sequence data and the gene libraries available in the international repository GenBank (Meredith *et al.*, 2011).

Another molecular approach allows the sampling of a far greater number of genes by sequencing the RNA that transcribes the sequences of genes that code for proteins (Wang *et al.*, 2009). This transcriptome method can yield patterns that represent thousands of genes, but it does not provide precise maps of the actual base pair sequences of particular genes that one can identify when sequencing whole genomes. Moreover, sampling and sequencing RNA require very fresh, high quality material that exceeds the preservation standards usually found in museum collections and even earlier collected specimens in frozen tissue facilities. This requirement places a premium on careful collecting, rapid freezing and preservation, and a well-curated tissue repository. However, one great advantage of whole genome sequencing, when it becomes even more readily accomplished at lower cost, is that it can also be used on material that has been preserved in collections for decades, such as feathers, hair, skins, and even fossils (Green *et al.*, 2010). Many collections-based institutions, such as NYBG and AMNH, New York: The Field, Chicago; NHM, London; CAS, San Francisco; NMNH, Washington, DC; and MNHN, Paris now have laboratory facilities and active programs in comparative genetics, and some of these (e.g., AMNH) are even sequencing whole genomes of less complex organisms, such as bacteria.

The increase of genetic and genomic research in museums has not diminished the importance of phenomic research at these institutions. Indeed, the museum expertise that combines a deep knowledge of natural history, basic taxonomy, comparative anatomy, and histology with genetic approaches offers powerful insights to major questions that deal with the organization and evolution of biodiversity. A much clearer understanding of the relationship of the genome to the phenome remains a central goal of all biology, and museum scientists have and will continue to play a key role in this quest. Modern comparative phenomics, such as the study of very detailed aspects of anatomy and tissue structure, is greatly aided by powerful imaging and software tools, such as CT-scanners, confocal microscopes, and 3D-digital reconstruction and animation. Such new technology is now readily being used to study fossils, the overwhelming majority of which are inaccessible to DNA sampling. Some museums now support state-of-the-art imaging laboratories to carry on such important research.

Technology and research activity has produced a huge influx of both genomic and phenomic information – a data deluge – which places an enormous burden on the effort to clarify, code, organize, and analyze these data. Thus, museums are committing major resources to the development of on-site and collaborative bioinformatics. Such programs generally comprise three components: (1) computing hardware infrastructure, such as facilities for cluster computing; (2) software tools and web applications for organizing databases, developing data pipelines for processing information, and providing visualizations; and (3) theoretical and computational research reliant on high-end computing, discrete math, and algorithmics. Examples of projects that rely heavily on advanced bioinformatics and involve museum and botanical garden science include a comprehensive study by the New York Plant Genomics Consortium (NYBG, AMNH, Cold Spring Harbor, and New York University) of 150 different species of plants based on 23,000 sets of genes that code for certain proteins. Using powerful databasing and computation tools, these gene data were grouped, ordered, and organized into a phylogenetic tree representing the evolutionary relationships of the plant species (Lee *et al.*, 2011).

On the phenomics side, diverse and complex character information can now be interactively and simultaneously loaded into very large matrices incorporating images, labels, comments, and literature citations through powerful cloud-computing web applications. One such application, MorphoBank, (O'Leary and Kaufman, 2011) is currently the home site for more than 100 published and 383 ongoing projects. MorphoBank is being used for one of the largest phenomics projects ever launched, an investigation supported by the National Science Foundation-sponsored Assembling the Tree of Life project for living and fossil mammals, which has produced a matrix with 4500 characters from the skeleton, dentition, and soft anatomy in a large number of species (Novacek and The Mammal ATOL Team, 2008). The resultant matrix, which incorporates nearly 450,000 labeled images, was developed by 22 international investigators from nine institutions, including four museums.

The development of such bioinformatics capacity will be vital to collections-based institutions going forward. Comparative biological data relating to museum research are among the most complex in any science. The supersized data matrices like those noted above, require sophisticated computational programs for tree construction and other operations. Such complex data require a heuristic approach to achieving resolution among an astronomical number of possible trees (more than 2 million for a data set with only 10 taxa and 4.5×10^{190} for 100 taxa). For this reason some museums have developed research areas that encompass theoretical computational biology in addition to other areas in bioinformatics.

Biogeographic and Environmental Studies

In addition to taxonomy, systematics, genomics, and phenomics, museum research extends into other areas that rely on complex data to address questions important to understanding biodiversity. One of the most notable of these endeavors is the application of Geographic Information Systems, using satellite-based imagery and other means to assess the distribution of species in relation to various environmental parameters, such as climate conditions, topography, vegetation cover, and proximity to cities or other sites of human occupation. Some of these studies incorporate sophisticated computational tools that allow investigators to predict the occurrence of particular species in areas yet unexplored for them by analyzing the multidimensional parameters that are consistently present at localities where these species are known to occur. These analyses then can be applied to assessments of conservation priorities in regions, such as Madagascar, experiencing widespread habitat destruction (Kremen *et al.*, 2008). Other approaches model changes in the range, density, and diversity for groups of species under different likely scenarios for climate change in a given region, such as the critical studies of vulnerable FynBos flora of southern Africa (Midgley and Miller, 2005). The application of such approaches to informing conservation strategies and policies has and will continue to be critical.

Biodiversity: The Need for a Better Census of Life

The aforementioned research areas underscore the importance of museums and like institutions in our effort to understand the vast array of species that occupy the planet. Museum scientists represent a large proportion of working systematists who deal with the description of the basic units of biodiversity, namely species, and place them in an evolutionary context by mapping the relationships of those species and their higher groups within phylogenies. Other aspects of museum research in biodiversity concern the geographic distribution of species as a function of various environmental factors, as well as the study of life cycles and ecological relationships.

Yet we are still confronted with the challenge to more fully account for such diversity, and demonstrate how it is impacted by the global scale environmental changes in play. As noted above, the 1.8 million species named do not, within orders of magnitude, capture the true number, which may be in the order of 10 million or more species yet to be identified (Wilson, 2003). Insects alone represent a staggering amount of known and unknown diversity with 800,000 species named, and doubtless millions more to discover. Approximately 100,000 fungi have been described and named, but this group is thought to number more than 1.5 million species. In terms of sheer biomass, the various worm groups, including segmented annelids and unsegmented nematodes, may represent four out of every five animals on Earth. Approximately 15,000 nematodes species have been named and the existence of millions more is suspected. An even more profound void in the census of life concerns bacteria and related microbial species, the oldest, most ubiquitous, most rapidly evolving, most physiologically diverse, and most poorly known of all organisms. Only approximately 6000 species of bacteria are recognized, but estimates of true diversity in the soil and the water columns using DNA probes and other methods indicate the presence of many millions of species (Gans *et al.*, 2005).

It is clear that the great collections and active research programs in museums and botanical gardens, and other collections-based institutions underrepresent the true diversity of such important groups as terrestrial arthropods, marine invertebrates, fungi, and microbial bacteria and protists. Indeed, specialists on certain of these groups are either nonexistent or a very few of the already insufficient number of some 6000 taxonomists worldwide (Wilson, 2003). There has been a welcome and recent introduction of microbial programs in a few museums, but this area is due much more emphasis as museums consider options and opportunities for future expansion of their collections and research. In addition, there is further need to integrate basic information of species diversity with other biological phenomena, such as population patterns, nutrient cycling, and resource partitioning, in order to demonstrate the important ecological roles of species within biological communities. Such more integrative research requires a scale-up in multidisciplinary approaches involving numbers of museum investigators and colleagues at other institutions.

Biodiversity Science and Conservation

Museums and other collections-based institutions have taken a number of different approaches in responding to the urgency and magnitude of the biodiversity and other environmental crises. Recruitment of curators to leadership scientific positions in some cases has emphasized the need for coverage of poorly known biological groups, or for expertise that relates basic biodiversity knowledge to experience in bioprospecting, surveys, environmental trend analysis, and related approaches. Many institutions have also developed special centers devoted to biodiversity science and conservation biology and programs. These units are meant to initiate programs more connective to the needs of local communities and governments by providing current, accurate, and relevant information on biodiversity and other impacts on the habitats in question. They also help in organizing large scale interdisciplinary projects, such as surveys that enlist teams of experts on diverse taxa and produce results that are of use in setting policy for the management of natural areas and habitats under threat.

It is clear that such scaled-up efforts in surveys and inventories are much needed. Many species now known to science have yet to be sufficiently sampled for their genetic composition. Large areas of the planet, such as the mid ocean, deep sea, and polar regions have barely been explored. There are hundreds of sites and regions throughout the world now considered under high threat or endangerment that contain a large diversity of endemic species, namely, species that are found nowhere else outside of the region in question (Sala *et al.*, 2000). Such highly vulnerable regions include most of Madagascar, the floral communities of southern Africa, Mediterranean coastal habitats, many coral reefs, high-latitude deserts, and montane forests of the Andes, Himalayas, and other ranges (Sala *et al.*, 2005). The identification of such regions and their endemics is the direct result of taxonomic and biogeographic research, much of it museum based. These areas, sometimes referred to as biodiversity hotspots (Myers *et al.*, 2000), help to define the ongoing and future agenda for biodiversity science and conservation.

Museum Advisory Roles and Applications in Biodiversity Conservation

Basic information stemming from surveys and collections is important in informing much broader constituencies about biodiversity loss and thus underscoring the need for mitigation and more effective conservation management. Vouchered specimens from collections serve as keys and references to field experts and parataxonomists in identifying the specimens or in assessing the impact of environmental degradation on animal and plant populations. Web-based publications such as the International Union for the Conservation of Nature's (IUCN's) Red Lists of Threatened Species are to a large part reliant on information from museum-sponsored field surveys and collections (see also Henderson and Chalmers, 2001). The ongoing digitization effort for large collections is an important advance that provides more ready access of information to conservation practitioners and land managers.

Such programs may also link to private or government agencies that manage biodiversity (Henderson and Chalmers, 2001). The latter include the Environmental Resources Information Network in Australia, the National Commission for

the Knowledge and Use of Biodiversity (La Comision Nacional Para El Conocimiento Y Uso de la Biodiversidad) in Mexico, and the National Biodiversity Institute (Instituto Nacional de Biodiversidad; INBio) in Costa Rica. INBio was established as an initiative with a particular emphasis on biodiversity science and conservation across a broad range of activities, from maintaining collections, conducting bioprospecting, and training parataxonomists, to educating the public on the importance of biodiversity resources. Collections- and field-based institutions also serve many other needs relating to biodiversity conservation (Henderson and Chalmers, 2001). They deliver scientific information bearing on the regulation of import and export of biological species, many of which present both environmental and health hazards in areas where they are artificially introduced. They also provide data vital to the development of international treaties and conventions meant to protect species, such as The Convention on International Trade in Endangered Species (CITES). And they play a prominent role in resource management. Accurate identification of species is necessary in developing effective pest management and programs for management of such commercial activities as agriculture, fishing, and domestic breeding. The increased trend toward local farming, green markets, and farmland that embraces a combination of cultivated and natural habitats (Foley *et al.*, 2005) relies on information of local and highly beneficial biodiversity, such as wild species of pollinating bees. Surveys by taxonomic experts disclose biodiversity and population changes before and after high impact industrial projects, such as mining or oil exploration or industrial-induced environmental disasters.

A more formal extension of such activities is the Museum service as a source for recommendations and advice on policy and practice relating to biodiversity loss and other environmental issues. Here the institutions must distinguish their mission as the providers of recommendations based on sound scientific results from that of various advocacy groups that may use such information for various purposes. In this way, the reputation of museums as trusted guides and the impact of the scientific work they foster are assured. Museums and like institutions provide the fundamental data on the realities of habitat degradation and its biological affects due to urban sprawl, deforestation, mining, war, and other human activities. Some countries, such as Australia, require biodiversity surveys and impact studies before any development is approved, and both impact studies and remediation programs are required by many local authorities in many countries.

At the government level, Museums have and will continue to play a critical role in both the development and implementation of policy. The effectiveness of the CITES treaty, for example, requires the continued contributions of accurate identifications and assessments supported by museums and like institutions. Museums are also important participants in developing initiatives prescribed by the Convention on Biological Diversity (CBD), and official delegations of the CBD, the IUCN, and many other international organizations and conventions include museum scientists and staff (Henderson and Chalmers, 2001). Implementation here is aided again by museum-based programs that offer training of in-country expertise in biodiversity sciences.

Applications in Human Health and Other Areas

Museum-based research on biodiversity also has applications that extend to other areas such as health and commerce. Studies by museum scientists and collaborators on the evolution of *Plasmodium*, the microbial parasite responsible for malaria in humans and other species, is contributing to our effort to understand better the nature and effects of the world's deadliest infectious disease (Perkins, 2008). Other insights by comparative biologists are important for forecasting the spread of viral epidemics, managing conditions for the prevention of diseases like cholera and typhus, and finding relationships between environmental degradation and likelihood of disease outbreak promoted by opportunistic species. Importantly, the comparative approach used by most museum scientists to investigate problems in biomedicine and health moves beyond a focus on standard model organisms – such as the fruit fly, mouse, or the bacterium *Escherichia coli* – and instead relies on the 3.6 billion-year-old experiment that is integral to the evolution of all life. Such a comparative approach may lead to insights that allow the mapping of convergent evolution of certain diseases, such as human papillomavirus (Chen *et al.*, 2009) or the variations evolved in blood anticoagulants (Phillips and Siddall, 2005). A classic study by Pershing *et al.* (1990) disclosed the potential for the spirochete *Borrelia burgdorferi* infection before the recognition of Lyme disease as a clinical entity by sampling museum specimens of the deer tick *Ixodes dammini* for the presence of spirochete-specific DNA sequences. The latter sequences were detected in 1940s specimens from Montauk Point and Hither Hills, Long Island, New York. These results suggested that the Lyme disease spirochete was already active in arthropod vectors for at least a generation before the formal recognition of this disease in the USA. Bioprospecting continues to produce discoveries important to pointing the way to new pharmaceuticals or food sources. In fact, the recent burgeoning interest in wild species of plants, fungi, insects, marine invertebrates, and other organisms as sources for healthy, and in some cases gourmet foods, reflects an appreciation for the potential bounty of biodiversity and the scientific information required for utilizing it.

Biodiversity and Culture

Noteworthy also is the knowledge of biodiversity as an inroad to understand the relationship among different cultures and societal traditions. The common formal taxonomy for species names and their classification, which is currently used in virtually all scientific literature, is actually an European development attributed to the work of the great eighteenth century Swedish naturalist Carolus Linnaeus, but some form of "folk" taxonomy has emerged from virtually every region and culture. We know that many people have developed traditional skills and languages for identifying species, and their success at this practice is impressive. Certain tribes in the Amazon are familiar with thousands of species, many of them important as sources for foods and medicines; outsiders have learned much from these great native naturalists. A study by a research team including museum scientists (Fleck *et al.*, 2002) examined the folk taxonomy of bat species used by the Matses Indians in

Amazonian Peru and found it surprisingly comprehensive. Although Matses have only one name for all bats, cuesban, they use descriptive terms to distinguish 43 different bats. Of the 57 species collected by scientists in the area over the years, 34 species had also been hunted by Matses, whose notebooks provided previously unknown information on the roosting habits of several of them. Moreover, at least 14 Matses names closely correspond to standard Linnaean taxonomic names.

This is only one of numerous examples where knowledge of local cultures demonstrates a shared interest and understanding of nature, a common attitude extremely important in developing practices for the sustainability of local species and habitats. Accordingly, staffs at some museums include anthropology curators with specific interests in the ways in which different cultures recognize and use biodiversity as integral to their traditional practices and everyday lives. Biodiversity training programs that take into account local traditions and perspectives are also offered by museum-based biodiversity and conservation centers.

Current Challenges

Despite the clear linkage among collections-based biodiversity science and conservation, human health, and other areas of societal need, there are many challenges to the effort to strengthen this connection. As noted, there is still a deficit in the number of experts, particularly for some of the groups least known and potentially most important for ecosystem services such as microbial organisms, fungi, and pollinating insects. The possible training programs for developing such expertise at universities, museums, and other institutions are limited in number. Even at the broadest levels, the uneven distribution of expertise and collections resources impedes efforts to account for biodiversity on a global scale. Much of the world's biodiversity is located in economically depressed countries very poorly endowed with relevant expertise and scientific infrastructure (Argawal and Redford, 2006). Thus, international programs for in-country training of local people are among some of the most needed and effective programs sponsored by museum-based biodiversity centers and similar organizations.

Results of biological surveys and prescriptions for the implementation of locally- or nationally- directed conservation management are often insensitive to local concerns relating to particular levels of education, economic background, cultural affiliations, and religious beliefs. People in developing countries are, for instance, faced with difficult choices because of their very poor standard of living (Argawal and Redford, 2006). Development of recommendations and programs more sensitive to these factors yields a more effective and feasible strategy for sustaining local natural resources, and is also key to identifying some of the overarching concerns shared by many different communities and cultures.

In the end, the reliable data and results of biodiversity surveys and research may have a less than desired impact in some instances simply because decision-makers and their constituencies fail to recognize biodiversity sustainability and management as a priority need. In many cases, this lack of attention seems rooted in a poor understanding of the fact that biodiversity loss is not a completely isolated issue, but one that bears on the economic, health, and quality of life issues that demand so much attention (Novacek, 2008). Hence, biodiversity research, surveys, and recommendations for conservation must be closely linked with educational efforts at many levels, a point expanded further below.

Training, Education, and Outreach

Graduate and Postdoctoral Training

Museums and the like are in reality major centers not only for research but also for training in the sciences concerned with biodiversity, such as taxonomy, systematics, evolutionary biology, biogeography, and conservation biology. Their increasing role in this respect relates to particular trends in higher education, where much of the training in these fields was displaced at universities by other areas, such as biomedical research, experimental neurobiology, and genomics centered on model organisms. In more recent years, the evolution of these disciplines has come full circle – the flood of technology offered by genomics and computational biology has been readily adopted and expanded on by researchers devoted to field work, collections, and the natural sciences. At the same time, there has been a rebirth at some universities of programs in evolutionary and organismic biology that draw more emphatically on the kind of expertise normally found in museums. A continuing trend in this convergent effort in research and training among universities, museums, and other collections-based institutions offers increased opportunity for the cultivation of biodiversity expertise.

Many museums and botanical gardens for example participate in joint programs with universities in the graduate training of new generations of scientists. Prominent among these are AMNH and NYBG (with Columbia, Cornell, New York University, and the City University of New York), the Smithsonian's NMNH (with George Washington University), the Field Museum (with the University of Chicago), and San Francisco's CAS (with University of California at Berkeley and Stanford University). Many of these arrangements involve a cost share for the graduate program, the advisory role of curators who are also adjunct faculty at the partnering university, and the eventual bestowal on the successful candidate of a PhD degree by the university involved. A more independent role for museums in graduate training is represented by the establishment in 2008 of a PhD program in comparative biology in the Gilder Graduate School at the AMNH. Thus the AMNH joins the Paris MNHN as the only independent museums in the world that offer a PhD degree.

At more advanced training levels, there are numerous postdoctoral programs that support scientists at museums concerned with biodiversity. Some of these incorporate international postdoctoral and graduate level training in order to develop the level of in-country expertise necessary for the global effort. Still, the number of programs and opportunities available in museums and in joint museum-university programs are limited and are far outstripped by the current demand for them by talented, well-educated students.

Student and Teacher Training

Museums of course have a much broader educational agenda than that concerned with graduate and postdoctoral training. These and related institutions intersect daily with vast audiences – families, school children, senior citizens, and other individuals from diverse cultural and economic backgrounds. Much of the investment by museums and related institutions in education involves better ways to serve the millions of students in classes from primary and secondary schools that visit museums each year. Many museums are now tied to school curricula and are timed to intersect with course work in general biology and related fields. After-school programs offer opportunities for field trips and laboratory activities not readily available in classrooms. Much educational learning is greatly aided by digital offerings on the web and software applications that allow students to prepare for museum visits and assemble information after they return to the classroom. Several cities have also developed cooperative educational programs in biodiversity and other subjects that link museums, gardens, zoos, and related institutions in the region as an educational resource for both teachers and students. For example, the Urban Advantage program in New York City is a standards-based partnership designed to improve students' understanding of scientific inquiry through collaborations between urban public schools and science cultural institutions with AMNH as the lead institution, and including, NYBG, the Bronx Zoo, the Brooklyn Botanical Garden, and others. Urban Advantage has now been adapted in other cities, including Boston, Miami, and Denver.

Eliciting Public Understanding and Engagement

Beyond advanced academic training and school programs, museums, botanical gardens, and related institutions have a unique connection with a vast public audience. Even in more economically and technologically advanced countries, there are limited opportunities for the lay public to stay abreast of the rapid rate of scientific discovery (Falk *et al.*, 2007). Outside popular science books, periodicals, films, television specials, and web offerings, the responsibility for providing a lifelong exposure to science falls to museums, botanical gardens, zoos, aquaria, science centers, and similar venues devoted to the public education of science.

These institutions are thus critically important in educating people on biodiversity issues and other environmental problems. However, in doing so they confront some major challenges. A consistent result in surveys of public attitudes, such as the Biodiversity Roadmap Report of 1998, is that the basic message – that the biodiversity enormously important to the sustainability of the environment and the quality of our own lives is at serious risk – is not getting across to many of the target audiences. The most penetrating messages are those that clearly relate scientific insights concerning biodiversity and biodiversity loss to more general environmental problems and in turn to problems rooted in common experience – poor water quality, depletion of fisheries, zebra mussels and other invasive species, forest clearing, open pit mining, urban sprawl, and many others. Basic biodiversity science of course provides the important database for all these arguments. But the public recognition of the importance of this work is elusive without the themes that address more familiar issues (Novacek, 2008).

One bridge that must be crossed in connecting biodiversity science with a diverse public is in inspiring a closer interest in and affinity with nature. The fact that museums and like institutions can offer an encounter with nature that is both vivid and authentic defines their cultural impact (Novacek, 2001). Many people, especially in urban areas, will rarely, if ever, see a relatively unspoiled tract of woodland in their region, let alone a tropical rainforest. For these individuals, an encounter with nature means a visit to a museum or the like. The enthusiastic response of visitors to this opportunity can be appreciated in terms of the huge audiences such institutions attract. Attendance figures for 2010 provided on the websites of just 18 museums, botanical gardens, zoos, and aquaria, including some of those shown in **Figures 1** and **2**, numbered more than 44 million visitors. Another survey claimed that more than 865 million people visited museums, gardens, zoos, nature centers, science centers, and related venues in 1999 in the USA alone (Lake and Perry and Associates, 2001). An additional attribute of museums and institutions as venues for communicating science is the feeling of trust they invoke in the public. Surveys show that natural history and science museums have extremely high credibility ratings (Lake and Perry and Associates, 2001).

The Impact of Exhibits

Notwithstanding the many educational programs offered by museums and like institutions, the primary "flagship" vehicle for connecting with the public has always been and continues to be exhibitions. The accuracy, currency, and inspirational qualities of such exhibits of course reaffirm the museum's role as a powerful source of public education and enlightenment. Yet, on the biodiversity front, there was a period of educational stasis; such institutions had not fully capitalized on their reputation as trusted guides. During the early 1990s, when dimensions of the biodiversity crisis became more evident, very little of the scientific concern was reflected in public exhibits. Many permanent exhibits that included environmental topics had not been revised since the time they first opened decades before, or were not complemented by new halls or temporary exhibits that addressed current themes (Novacek, 2001).

Part of the reason for such arrested development was doubtless financial. Renovating exhibit halls or building new ones are massive and extremely expensive endeavors, and there are often major economic constraints to progress. Despite the challenge, there are more recent notable improvements on this front, and exhibits dealing with current environmental issues, including biodiversity, have been developed by many museums and related institutions. The AMNH in New York was one of the first, if not the first, to open in 1998 a major permanent Hall of Biodiversity. The exhibit is meant not only to inform the public on the stark realities of global species loss and the biodiversity crisis but also to show the beauty, wonder, and diversity of what is at risk. Its features include an electronic "BioBulletin" presenting

regularly updated accounts of events that impact biodiversity and scientific and conservation efforts to mitigate these impacts, in places like Madagascar, Bolivia, the Congo, and Mongolia. The exhibit also has a walkthrough diorama of a section of rain forest of the Central African Republic that shows the impact of human activity on even a relatively pristine environment. The diversity of life in the hall is covered by the "wall of life," that displays more than 1500 specimens and models complemented by videos of organisms in action and interactive touch screens. The wall of life has proven to be one of the museum's most popular attractions. A completely renovated Milstein Hall of Ocean Life incorporating many environmental themes and adjoining the Hall of Biodiversity reopened in 2003.

Elsewhere, exhibitions devoted to more current environmental issues have been developed at various scales. At the ambitious end of the spectrum, the CAS reopened in 2008 in an entirely new structure devoted to collections, research, and exhibition halls, including many exhibits with a strong environmental emphasis. The new museum is itself one of the largest high grade, green buildings in the USA (Barinaga, 2004). London Natural History Museum's new Darwin Center is a bold statement that reveals its massive collections and its laboratories to visitors within an evocative seven-story, concrete "cocoon" encased in a huge glass-walled "cube." The Darwin Center thus brings the public more in contact with the kind of biodiversity- and collections-based research that is usually "behind the walls." Complementing the magnificent Grande Galerie de l'Evolution at the Museum National d'Histoire Naturelle in Paris are a series of temporary exhibits that explore the wonders of biodiversity and the mysteries of oceanic life, such as its recent exhibit "The Deep," which subsequently traveled to other venues. There are other similar installations at the Smithsonian's NMNH, the Field Museum, the Denver Museum of Nature and Science, and others too numerous to describe here. Also, new partnerships among institutions have allowed the sponsorship and nuanced development of timely exhibits on such themes as endangered species, climate change, the importance of evolution in understanding biodiversity, and issues regarding the present and future availability of fresh, clean water. These and other exhibits with similar themes offer clear and consistent messages as they travel to various international destinations.

Public Educational Programs and Citizen Science

Although exhibition remains the principal educational portal and indeed the prime visitor attraction in museums, there has been an increasing emphasis on educational programming either connected to or independent of exhibits. Much of this relates to the programs noted above that link to school programs for teacher and student training and are designed to complement approved curricula. In addition, museums and like institutions are demonstrably very effective locations for major conferences on subjects of high public interest. Several aforementioned institutions sponsor annual symposia on biodiversity issues that include public lectures and panel discussions and the proceedings are now very often webcast in order to reach very large online audiences. Museum scientists are heavily featured in many media offerings including television specials, newspaper articles, and online blogs. For example, a recent feature of the New York Times, "Notes from the Field" provides daily journals from field scientists conducting biodiversity surveys and similar studies in remote regions. Another increasingly common feature of museums, are "Science Cafes," regular evening informal gatherings aimed at young adult audiences, wherein scientists share their adventures, results, and thoughts over cocktails.

These programs and many others are largely built on a strategy of engagement that allows educators to share information and perspectives, and elicit reactions in terms of questions and discussion. What about more active roles on the part of public audiences? A relatively new educational effort aimed at eliciting broader public engagement – dubbed citizen science – involves public–professional partnerships that allow people of all ages an opportunity to participate in real scientific research and to interact with scientists in the process (Cohen, 1997; Brossard *et al.*, 2005). This kind of proactive participation is intended to not only contribute new data on species and habitats but will also increase the participants' understanding of the process and results of the relevant science (Tuss, 1996). Such enlightenment is further hoped to strengthen participants' connections with both science and the environment in ways that cultivate a sense of stewardship.

Citizen science programs require intensive investment of professional time and energy, and, to date, the number of people that actually have the opportunity to become citizen scientists is limited. The problem seems surmountable as more efficient programs, including the use of social networking and web-based technology for linking scientists with science educators and their students, become available. The *Bioblitz* (2007) biodiversity surveys that have been carried out in New York's Central Park, Washington, DC, and many other sites yielded new scientific results that not only further enthused participants and galvanized their activities but also attracted media interest. It seems that programs in citizen science have much potential if they allow more people to participate, their impacts are more thoroughly analyzed, and participants are given a more explicit presentation of the environmental issues that relate to their contribution (Brossard *et al.*, 2005).

Conclusions: Museums and the Biodiversity Challenge

Many major museums and related institutions of the present day have histories extending back well nearly two centuries. Yet these institutions are arguably more important than ever, as we contend with the dramatic environmental transformation of this planet in this century. There is an exciting and unprecedented opportunity for museum researchers to participate in and even lead the great quest to account much better for the great scope of biodiversity and the organization and evolution of life.

There are of course many challenges to the enterprise. On the science front, the lack of taxonomic expertise in many areas, the uneven distribution of scientific personnel and

resources globally, and the huge expenses and investment in time necessary for collecting, organizing, and digitizing collections information are all impediments. On the educational front, the low level of science literacy coupled with public indifference or even resistance are also major obstacles. A very large and diverse public demonstrates a sense of concern about environmental problems (Biodiversity Project, 1998). However, people also often do not recognize the implications of the biodiversity loss in exacerbating many problems more familiar and more important to them, such as health and economics. Thus museums and like institutions devoted to the study of nature are still challenged with the goal defined for the biodiversity agenda more than 20 years ago. They must provide enhanced understanding of biodiversity and its degradation that empowers people to make choices based on sound science and reliable recommendations.

To this end, the momentum in both museum research and education in relation to biodiversity is encouraging. Museums and their kinship institutions support their expeditions, collections, research, exhibitions, and educational programs because they are enlisted in a human quest of paramount importance – to understand biodiversity, the evolution of life, and the critical roles of species in ecosystems, and to share this information with a broad public in ways that elicit enlightenment and engagement with issues important for sustaining the quality of life. The degree to which such institutions are successful in this quest in the coming years will be pivotal in determining how effectively societies will deal with the greatest environmental crisis in the history of the human species.

References

Argawal A and Redford J (2006) Poverty, development and biodiversity: Shooting in the dark? *Wild Life Conservation Society Working Paper 26*, i–iv, 1–58.

Ariño AH (2010) Approaches to estimating the universe of natural history collections data. *Biodiversity Informatics* 7: 81–92.

Barinaga M (2004) California Academy starts on the museum of its dreams. *Science* 304: 669–670.

Bioblitz (2007) National geographic in the field. *National Geographic Society.* http://www.nationalgeographic.com/field/projects/bioblitz-dc-2007.html.

Biodiversity Project (1998) Engaging the public on biodiversity: A roadmap for education and communication strategies. *The Biodiversity Project*, pp. vii + 118. Wisconsin: Madison.

Brossard D, Lewenstein B, and Bonney R (2005) Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education* 27(9): 1099–1121.

Chapman A (2005) *Uses of Primary Species Occurrence Data, Version 1.0*, pp. 106. Copenhagen: Global Biodiversity Information Facility.

Chen Z, van Doorslaer K, DeSalle R, *et al.* (2009) Genomic diversity and interspecies host infection of *α12 Macaca fascicularis* papillomaviruses (MfPVs). *Virology* 393: 304–310.

Cohen KC (ed.) (1997) *Internet Links for Science Education: Student–Science Partnerships*. New York: Plenum Press.

Duckworth WD, Genoways HH, and Rose CL (1993) *Preserving Natural Science Collections: Chronicle of Our Environmental Heritage*. Washington, DC: National Institute for the Conservation of Cultural Property.

Falk JH, Storksdieck M, and Dierking LD (2007) Investigating public science interest and understanding: Evidence for the importance of free-choice learning. *Public Understanding of Science* 16: 455–469.

Fleck DW, Voss RS, and Simmons NB (2002) Undifferentiated taxa and sublexical categorization: An example from Matses classification of bats. *Journal of Ethnobiology* 22: 61–102.

Foley JA, DeFries R, Asner GP, *et al.* (2005) Global consequences of land use. *Science* 309: 570–574.

Gans J, Wolinsky M, and Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* 309: 1387–1390.

Global Biodiversity Information Facility (GBIF) (2003) Global Biodiversity Information Facility Annual Report 2001–2002. Copenhagen.

Green RE, Krause J, Briggs AW, *et al.* (2010) A draft sequence of the Neandertal genome. *Science* 328: 710–722.

Hebert PDN, Penton EH, Burns JM, Janzen DH, and Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator. Proceedings of the National Academy of Sciences USA* 101: 14812–14817.

Henderson P and Chalmers N (2001) Museums, and institutions, role of. In: Levin SA (ed.) *Encyclopedia of Biodiversity*, pp. 271–280. San Diego: Academic Press.

Kohlstadt SG and Brinkman P (2004) Framing nature: The formative years of natural history museum development in the United States. *Proceedings of the California Academy of Sciences* No. 2(supplement 1): 7–33.

Kremen C, Cameron A, Moilanen A, *et al.* (2008) Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science* 320: 222–225.

Lake Snell and Perry and Associates (2001) *Americans Identify a Source of Information they can Really Trust*. Washington, DC: American Association of Museums.

Leakey RE and Lewin R (1996) *The Sixth Extinction: Patterns of Life and the Future of Humankind*. New York: Knopf/Doubleday.

Lee EK, Cibrian-Jaramillo A, Kolokotronis S-O, *et al.* (2011) A functional phylogenomic view of the seed plants. *PLoS Genetics* 7(12): e1002411.

May RM (1988) How many species inhabit the earth? *Science* 241: 1441–1449.

Meredith RW, Janečka JE, Gatesy J, *et al.* (2011) Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334: 521–524.

Metzker M (2010) Sequencing technologies – the next generation. *Nature Genetics* 11: 31–46.

Midgley GF and Miller D (2005) Modeling species range shifts in two biodiversity hotspots. In: Lovejoy TE and Hannah L (eds.) *Climate Change and Biodiversity*, pp. 229–231. New Haven: Yale University Press.

Myers N, Mittermeir RA, Mittermeir CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Novacek MJ (2001) The role of natural history museums in the stewardship of biodiversity. In: Mares M (ed.) *A University Natural History Museum for the New Millennium*, pp. 47–55. Norman, Oklahoma: Sam Noble Oklahoma Museum of Natural History.

Novacek MJ (2008) Engaging the public on biodiversity issues. *Proceedings of the National Academy of Sciences, USA* 105: 11562–11578.

Novacek MJ and The Mammal ATOL Team (2008) A team-based approach yields a new matrix of 4500 morphological characters for mammalian phylogeny. *Journal of Vertebrate Paleontology* 28(supplement 3): 121A.

Organization for economic co-operation and development (OECD) (1999) *Final Report of the OECD Megascience Forum Working Group on Biological Informatics*. Paris: OECD Publications.

O'Leary MA and Kaufman S (2011) MorphoBank: Phylophenomics in the cloud. *Cladistics: the International Journal of the Willi Hennig Society* 27: 1–9.

Perkins SL (2008) Molecular systematics of the three mitochondrial protein coding genes of malaria parasites: Corroborative and new evidence for the origins of human malaria. *Mitochondrial DNA* 19: 471–478.

Pershing DH, Telford 3rd SR, Rys PN, *et al.* (1990) Detection of *Borrelia burgdorferi* DNA in museum specimens of *Ixodes dammini* ticks. *Science* 249: 1420–1423.

Phillips AJ and Siddall ME (2005) Phylogeny of the new world medicinal leech family macrobdellidae (Oligochaeta: Hirudinida: Arhynchobdellida). *Zoologica Scripta* 34: 559–564.

Pimm SL, Russell GJ, Gittleman JL, and Brooks TM (1995) The future of biodiversity. *Science* 269: 347–350.

Raven PH (1987) The scope of the plant conservation problem world-wide. In: Bramwell D, Hamann O, Heywood V, and Synge H (eds.) *Botanic Gardens and the World Conservation Strategy*, pp. 19–29. London: Academic Press.

Ross D (1995) *Aristotle*, 6th edn. London: Rutledge.

Rowe TR and Frank LR (2011) The disappearing third dimension. *Science* 11: 712–714.

Sala OE, Chapin III FS, Armesto JJ, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.

Tautz D, Acrtander P, Minelli A, Thomas RH, and Vogler AP (2003) A plea for DNA taxonomy. *Trends in Ecology and Evolution* 18: 70–74.

Tuss P (1996) From student to scientist, an experiential approach to science education. *Science Communication* 17: 25–44.

Wang Z, Gerstein M, and Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10(1): 57–63.

Wetterstrand KA (2011) DNA Sequencing Costs: Data from the NHGRI Large-Scale Genome Sequencing Program Available at: www.genome.gov/sequencingcosts. (Accessed December 30, 2011).

Wilson EO (2003) The encyclopedia of life. *Trends in Ecology and Evolution* 18: 77–80.

Myriapods

Alessandro Minelli, University of Padova, Padova, Italy
Sergei I Golovatch, Russian Academy of Sciences, Moscow, Russia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Anthropochoric Geographic distribution due to human agency.
Cerotegument Hydrophobic secretion layer covering the body of several arthropods adapted to temporary submersion.
Disharmonic, fauna A fauna in which a strongly reduced number of major lineages is represented, as in oceanic islands, due to their improbable colonization from distant sources.
Laurisilva Warm-temperate woodlands characterized by evergreen broadleaves such as laurel (*Laurus*) trees.
Pupoid Earliest postembryonic stage with incompletely developed appendages, as in insect pupae, hence motionless.
Stenotopic A species (or higher taxon) with a highly restricted geographic range.

Introduction

Myriapods are wingless terrestrial arthropods with at least nine pairs of walking legs in the adult and a trunk not subdivided into thorax and abdomen. Most myriapods begin their postembryonic life with a small number of trunk segments and leg pairs, and obtain the remaining ones through a series of molts. However, no myriapod undergoes a complete metamorphosis as many insects do in their life cycle. These and other features of myriapods have long been regarded as primitive within a large arthropod lineage called the Atelocerata (or Tracheata, Uniramia, or Labrata), traditionally recognized as including myriapods and hexapods, as opposed to the other main arthropodan lineages Chelicerata and Crustacea. However, molecular investigations suggest that insects may be closer indeed to crustaceans than to myriapods, a scenario that invites fresh reinvestigation of the real affinities of the myriapod groups among themselves and to the remaining arthropods. Some studies even suggest that myriapods form with the chelicerates a natural group, the Paradoxopoda or Myriochelata. Irrespective of their eventually closer affinities to the insects and crustaceans or to the chelicerates, myriapods are mostly reported as forming a natural (monophyletic) group (Edgecombe, 2011), supported by morphological characters (e.g., Edgecombe, 2004) as well as by molecular evidence (e.g., Regier *et al.*, 2010).

Four recent and a few fossil classes of myriapods are recognized. Centipedes (class Chilopoda) and millipedes (class Diplopoda) are the main myriapod groups, whereas the two remaining recent classes, the Symphyla and the Pauropoda, lack vernacular names. Of the extinct myriapodous arthropods, Kampecarida are among the most remarkable, because they possibly represented the earliest myriapods ever (Upper Silurian), although it is still debatable whether they were terrestrial or not. The extinct Archipolypoda and Arthropleurida were possibly true Diplopoda; the former were conspicuously spiny and reached 30 cm in length, the latter contained both the smallest (a few millimeters) and the largest myriapods ever (up to 3 m long). Remains of the several other extinct orders are also known from the Paleozoic. Apart from some problematic Silurian remains that have been tentatively interpreted as scutigeromorph legs, the earliest centipedes recorded are the extinct Devonobiomorpha from the Devonian.

Basic Morphology

Chilopoda

Centipedes range in length between a few millimeters and the approximately 30 cm of *Scolopendra gigantea* (**Figure 1**). The centipede body is divided into head and trunk. Compound eyes are only present in Scutigeromorpha, whereas groups of simple eyes (ocelli) are present in most Lithobiomorpha and many Scolopendromorpha; all Geophilomorpha and many Scolopendromorpha are blind, as are several cavernicolous representatives of the Lithobiomorpha. The antennae may be longer than the body, as in the Scutigeromorpha and in some Lithobiomorpha. In these two groups, the number of antennal articles is mostly high and variable, but it is fixed and generally lower in the other groups (14 in all Geophilomorpha and 17 in most Scolopendromorpha). The first trunk segment bears a pair of specialized appendages – the poison claws (or forcipules), each of which contains a poison gland – which are used in the capture of prey and occasionally in defense. Each of the following segments bears a pair of legs, mostly of cursorial type, often with sexual or other specializations in the last pair(s). Curiously, the number of leg-bearing segments is always odd-numbered in the adults, with 15 pairs of legs in all Scutigeromorpha, Lithobiomorpha, and Craterostigmomorpha, but 27 (*Schendylops oligopus*) to 191 (*Gonibregmatus plurimipes*) in the Geophilomorpha. In most geophilomorph centipedes, to the exclusion of the vast majority of the Mecistocephalidae, the number of trunk segments is variable within the species and higher in females. Nearly all species in the Scolopendromorpha have either 21 or 23 pairs of legs, but one species (*Scolopendropsis bahiensis*) includes individuals with 21 as well as individuals with 23 pairs of legs, and its closest relative, the recently described *Scolopendropsis duplicata*, includes individuals with 39 and individuals with 43 pairs of legs (Chagas *et al.*, 2008).

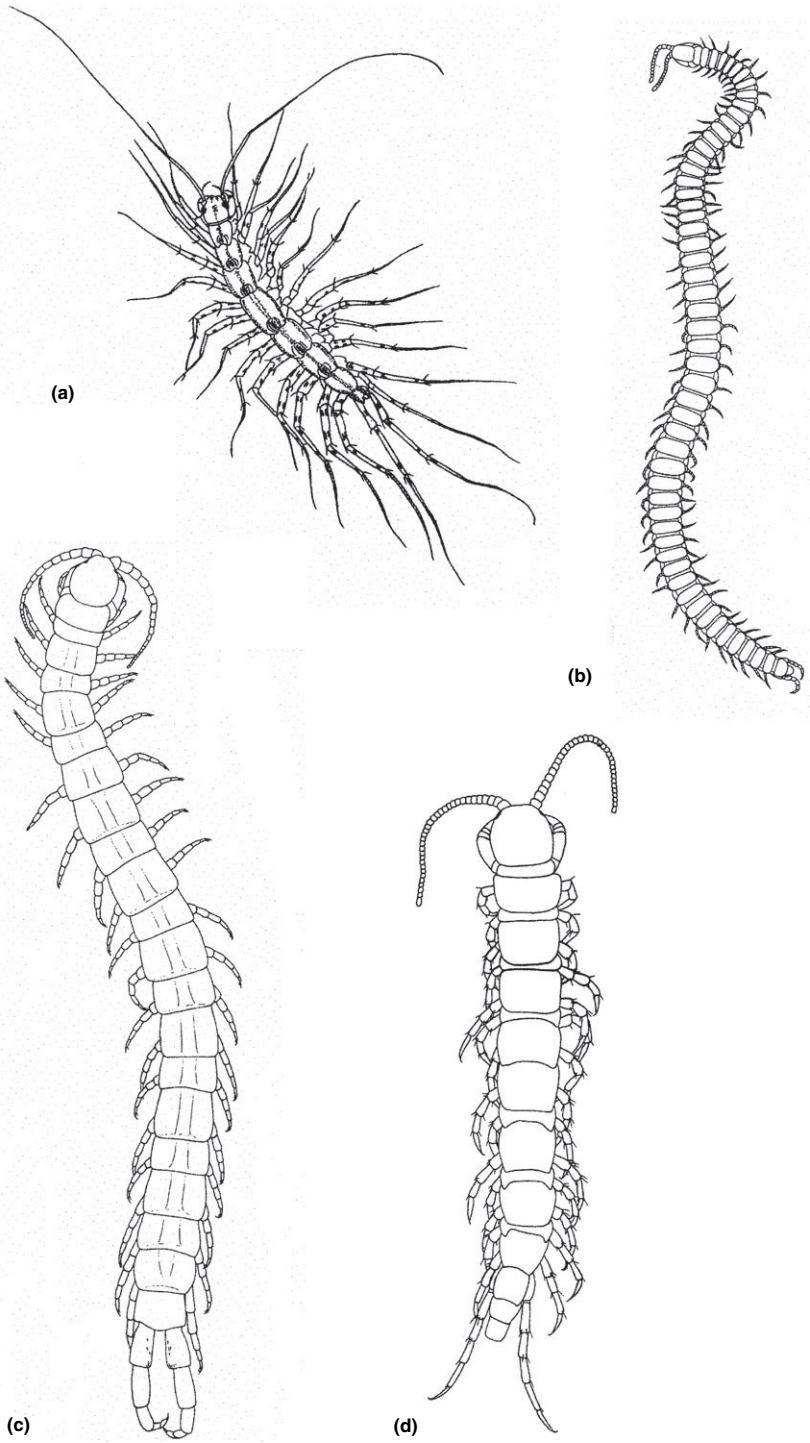


Figure 1 Habitus of representatives of the four main centipede orders: (a) *Scutigera coleoptrata* (Scutigeromorpha); (b) *Geophilus carpophagus* (Geophilomorpha); (c) *Scolopendra cingulata* (Scolopendromorpha); (d) *Lithobius validus* (Lithobiomorpha). Adapted from Centre International de Myriapodologie.

Unique to the Scutigeromorpha is a respiratory system with seven dorsal openings providing oxygen to the dorsal vessel via thick bundles of tiny, short tracheae; the oxygen is further distributed to the whole body with the help of the circulatory apparatus: the circulating liquid (hemolymph) contains an oxygen-binding pigment (hemocyanin) related to the respiratory pigments found in several chelicerates (scorpions and spiders) and crustaceans. In all the remaining centipedes there are tracheae comparable to those of insects, opening on lateral spiracles of most trunk segments (Geophilomorpha

and the scolopendromorph genus *Plutonium*) or on alternate segments (the remaining groups). The genital opening is found on a subterminal segment at the posterior end of the body, near the anus.

Centipede stings are fairly frequent, but usually do not represent a mortal event for humans. In literature there are only a few well-documented cases of death caused by these animals. Nevertheless, several species may cause serious damage in humans, including myocardial ischemia and infarction, hemoglobinuria and hematuria, hemorrhage, and rhabdomyolysis. Also allergy-like complications have been recorded following centipede envenomations. Most centipede envenomations, however, produce transient symptoms that can last a few days but do not require acute medical treatment (Bush *et al.*, 2001).

Information on venom composition is available only for a few species of scolopenders, their venom is a mixture of toxins including myotoxins, cardiotoxins, neurotoxins, cytotoxins, together with several enzymes like esterases, proteases, and hyaluronidases (Undheim and King, 2011).

Some scolopenders have found a place in traditional Chinese medicine alongside scorpions and other arthropods.

Diplopoda

The size of millipedes spans between a few millimeters and 35 cm, one of the largest species being the spirostreptid *Sechelleptus seychellarum*. The body (**Figure 2**) includes a head and a trunk. Groups of simple eyes are present in most orders but are lacking in all Polydesmida and in some other groups as well as in all specialized cavernicolous species. The antennae are generally short and always comprise only eight articles. The mouthparts are generally adapted for cutting and chewing hard matter, such as wood or dead leaves, but some millipedes have evolved adaptations for sucking. The trunk is normally elongate, more or less flattened dorsoventrally, but subcylindrical in several orders. The body wall is rarely soft and flexible (subclass Pselaphognatha with the only order Polyxenida) and the exoskeleton is usually rigid (subclass

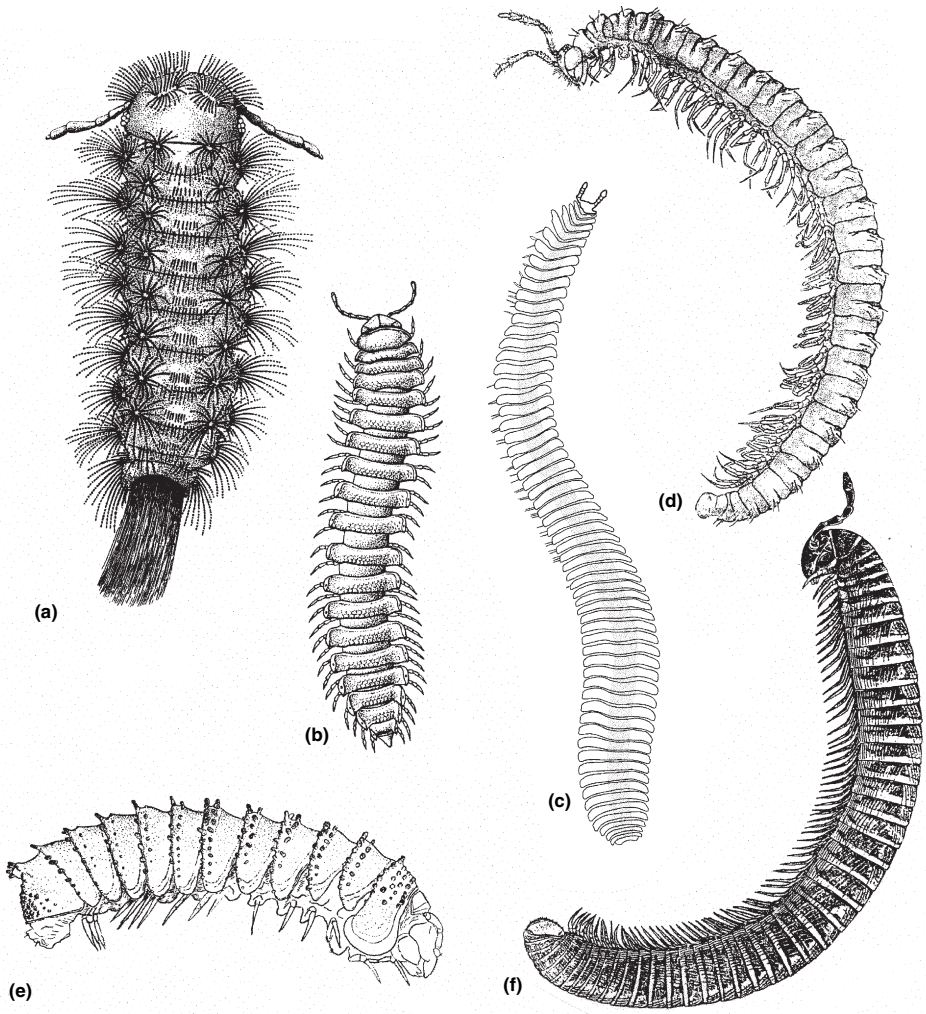


Figure 2 Habitus of representatives of the main millipede orders: (a) *Macroxenus enghoffi* (Polyxenida) (after Nguyen Duy-Jacquemin, 1996); (b) *Polydesmus* sp. (Polydesmida); (c) *Brachycybe* sp. (Platydesmida); (d) *Opisthocheiron canayerensis* (Chordeumatida); (e) *Trachysphaera lobata* (Glomerida); (f) *Stemmiulus adisi* (Stemmiulida). Adapted from Centre International de Myriapodologie.

Chilognatha, with all remaining orders) due to the presence of calcium salts in the internal layers (endocuticle). When the millipede is about to undergo a molt, these salts are dissolved. Accordingly, these arthropods need to obtain from the food sizable amounts of calcium. There are no waxes in their epicuticle, but some protection from evaporation is obtained through the lipoid compounds in their exocuticle.

The first trunk segment (the collum) is legless; it is followed by three (four in the Spirobolida) "thoracic" segments with one pair of legs each and a further number, sometimes very high, of "abdominal" segments with two pairs of legs each. There are 11–17 pairs of legs in the Polyxenida and at least 17 (but usually many more) in the remaining groups, with the highest number being 375 pairs of legs recorded in *Illacme plenipes* (Enghoff, 1990).

Most millipedes are provided with chemical defenses in the form of glands (ozadenes) producing noxious substances which pour out from the series of repugnatorial openings (ozopores), dorsal in the pill millipedes (Glomerida) but lateral in the majority of the diplopods. No chemical defense has been detected in some lineages, that is, the Polyxenida, the Sphaerotheriida, and the Chordeumatida.

A few millipedes spray their secretions up to 40 cm. These secretions are quite effective against both the vertebrates and arthropods. Warning colors are widespread.

The repugnatorial secretions of several large millipedes appear so noxious that some Indian tribes in Central America use them for poisoning arrows prior to hunting. However, some of these tribes cook large millipedes and find them tasty.

A diversity of defense substances have been found in these animals, with distinct classes of chemicals being characteristic of the different millipede groups: The juloids rely on benzoquinones, hydroquinones, phenol, and the acetates of some long-chain carboxylic acids; the polydesmoids rely on several carboxylic acids but especially on benzoic acid, benzaldehyde, hydrogen cyanide, and mandelonitrile; and the pill millipedes (Glomerida) rely on very peculiar and complex heterocyclic compounds (glomerin and homoglomerin) and the Polyzoniida on compounds known as polyzonimine and nitropolyzonimine (Eisner *et al.*, 1978). Several millipedes are adapted to coiling onto themselves (volvation), thus becoming a smooth ball difficult to seize; many others can roll themselves into a spiral flat coil. However, despite such precautions, many millipedes fall victim to insectivorous birds and predatory arthropods.

Special modifications of one or two (more rarely, three or four) leg pairs in the male may involve their transformation into clasping or genital organs. Posterior claspers are characteristic features of the Pentazonia (the pill millipedes and their closer relatives). The male uses these claspers to fix the female during mating. In the males of most other millipedes, one or, more commonly, two leg pairs of body segment 7, or of segments 7 and 8, are modified into gonopods. These are mostly very complex structures directly involved in sperm transfer. The use of the fifth pair of legs of the male *Chelojulus sculpturatus*, which are transformed into functional forceps, is unknown.

The tracheal respiratory system is provided with spiracles normally opening on the sternum near the coxa. The paired gonopores open on the coxae of the second pair of legs or in their proximity. Those of the male are found on a special structure incorrectly termed penis, which has nothing to do with direct sperm transfer.

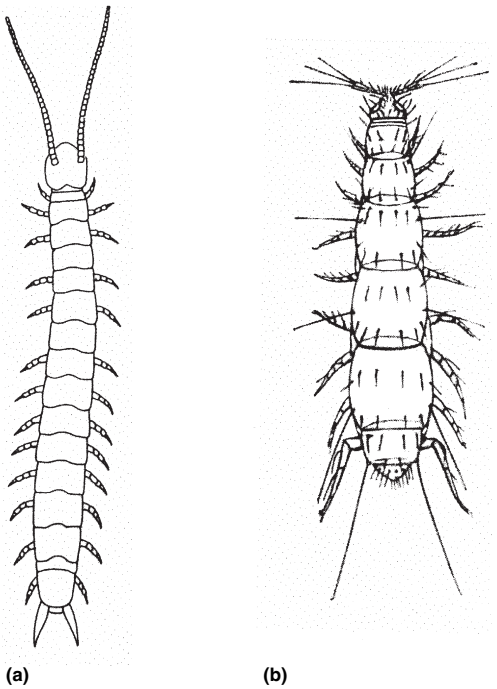


Figure 3 Habitus of (a) a symphylan and (b) a pauropod. Adapted from Centre International de Myriapodologie.

Symphyla

Symphylans are tiny myriapods, mostly 2–9 mm long (but *Hanseniella magna* is 25–30 mm long), generally with a whitish body (**Figure 3(a)**). All species are blind and bear one pair of elongate multisegmented antennae, 12 (rarely 11) pairs of legs, and a pair of posterior uniarticulated appendages called spinnerets. The mouthparts are of masticatory type. The genital opening is found on the fourth trunk segment. Symphylans are the only arthropods with a pair of respiratory openings on the head, but the whole of their very soft and permeable cuticle is important for gas exchange. In these tiny myriapods there is no cuticular protection against evaporation.

Pauropoda

Pauropods are the smallest among myriapods; most species are only 0.5–0.7 mm long, and the largest is just 2 mm long (**Figure 3(b)**). There are no eyes. The antennae are of unique structure, with four (order Tetramerocerata) or six (order Hexamerocerata) basal articles and two apical or subapical branches, each bearing in turn one or two flagella. The mouthparts are comparable to those of millipedes but are generally adapted to the suction of fluid aliment. The trunk bears 9–11 pairs of legs. Most species lack a respiratory system; one pair of tracheae, with spiracles on the coxae of the first pair of legs, is present only in the Hexamerocerata.

Basic Biology

Nearly all myriapods are strictly terrestrial and most species are found in forest leaf litter or in rotten wood, under the bark of dead trees, in the soil, or in caves. Several millipedes also occur in dung, compost, and almost any other kind of plant debris; in nests of ants, termites, and birds; in worm and mammal burrows; near water; on the seashore; on open terrain; under stones; in buildings; etc. That is, they are found in virtually any terrestrial environment.

A few species of diplopods (e.g., *Blaniulus guttulatus*) and symphylans (*Scutigerella immaculata*) are agricultural pests. Owing to their swarming habits, a few millipedes can be of danger for transport when smashed in myriads on roads or railways or for housing, especially due to their smell, but such extremes are exceptional. In temperate regions, migrating swarms of many thousands of millipedes have been recorded, especially swarms of polydesmidans (e.g., *Fontaria virginiensis*, *Parafontaria laminata*, and *Polydesmus collaris*). Causes for and mechanisms of this behavior are little known, but it has been suggested that heavy rainfalls may cause the swarming behavior of a large spirostreptid (*Plusioporus setiger*) in southern Brazil. Larger millipedes have quite long life spans lasting between 2 and 11 years, whereas their life histories usually take 2 or 3 years even in tropical environments. In high-montane habitats, the development is normally retarded by a year or two. However, some tiny tropical species live for less than 1 year and their life cycle takes only about 3 months.

Among myriapods, only some centipedes have successfully colonized eremic environments, with many species of this class (mostly belonging to the Scolopendromorpha) being true deserticoles. This is due to their structural adaptations (waterproof cuticle or complex spiracles) allowing them to reduce water loss. The remaining myriapods are distinctly hygrophilous to mesophilous, often (the bulk of millipedes) calciphilous as well, which explains their difficulty in coping with conditions differing from those prevailing in forest litter.

Tree dwellers are not common among myriapods, being mainly represented by small or slender, if not flat-bodied, millipedes and by some representatives of lithobiomorph and geophilomorph centipedes. Species confined to the tree crowns or living in suspended soil are even fewer. Some scolopendromorphs, however, are fairly common in suspended soils in the tropical forests of America.

Within the Chilopoda, only the Geophilomorpha include halophilous species such as *Tuoba poseidonis*, *Strigamia maritima*, *Hydroschendyla submarina*, and *Geophilus seurati*. *S. maritima* feeds on crustaceans (amphipods and barnacles) and gastropods (the periwinkle *Littorina saxatilis*) (Barber, 2009).

Centipedes are often found in the nest of ground-dwelling mammals, e.g., moles and field mice. Their presence in such specialized environments is likely opportunistic, in the majority of cases, but a distinct preference for mammal nests has been claimed at least for the lithobiomorph *Lithobius crassipes* and the little scolpendromorph *Cryptops hortensis*. Also facultative seems to be the presence of millipedes and centipedes in ant and termite nests, where these arthropods, especially lithobiomorphs, are apparently well tolerated by their hosts.

Myriapod species capable of tolerating sea water are particularly few, although these are often quite common and widespread. Interestingly, no conspicuous morphological types can be discerned among myriapods as obvious adaptations to any of such definitely marginal environments. Again, no common morphological adaptation is shared by the relatively few anthropochoric myriapods that have attained more or less vast, sometimes worldwide, distributions due to human agency. However, several minor modifications of leg structure can be assumed as advantageous for better climbing performance, whereas other adaptations involving the mouthparts and spiracles indicate the species' semiaquatic habits (Adis *et al.*, 1998; Golovatch and Kime, 2009).

In their feeding habits, myriapods are quite diverse. Centipedes, as revealed by their poisonous forcipular fangs, are basically predatory. They are mainly solitary hunters; gregarious habits have rarely been observed, including feeding on barnacles by swarms of specimens of the geophilomorph *S. maritima* in the tidal region along the seashore. Some centipedes, however, are known to occasionally consume plant material. In contrast, most millipedes feed on vegetable matters, particularly dead leaves and wood, but also algae, fungi, lichens, moss, and pollen.

Of recent, somehow unexpected records have noticeably enlarged the prey range of scolopendromorph centipedes (Voigtländer, 2011). This is now known to include scorpions (the European *Scolopendra cingulata* preying on *Euscorpius flavicaudis*), bats (*Sc. gigantea* on a Venezuelan bat species), tree frogs (*Otostigmus tibialis* on *Dendropsophus elegans*), and even sea anemones (*Otostigmus scabricauda* on *Bunodosoma caissarum*).

Millipedes, whose density in forest soil is sometimes higher than 1000 individuals per square meter, are estimated to consume 10–15% of the annual leaf fall in temperate forests. A few species may be qualified as omnivores; very few are carnivores – one of them being *Apfelbeckia insculpta* (order Callipodida), which feeds on earthworms. Coprophagy is also quite widespread in this animal group.

Some suctorial myriapods (all tetramerocerate pauropods and some millipedes), all with mouthparts more or less strongly reduced, are known to live on fungi or semiliquid end products of decaying plant material. Most pauropods feed on fungal hyphae and spores, but *Millitauropus* (Hexamerocerata) feeds on small arthropods (springtails and their eggs). Symphylans are largely vegetarian, rarely saprophagous. Hence, the roles played by various myriapod groups in the terrestrial food chains differ, with millipedes being of particular importance as primary destructors and soil-formation factors and centipedes as active and restless predators. Representatives of only a few millipede orders are capable of manufacturing silk for cocoon production during molting or egg laying.

Reproduction

A few dozen cases of parthenogenesis are known (Enghoff, 1994). One is the lithobiomorph centipede *Lamyctes coeculus*, a synanthropic species known from several continents. Two small European millipedes (*Nemasoma varicorne* and *Polyxenus lagurus*), both mainly living under tree bark, exhibit geographical parthenogenesis, with bisexual populations in less disturbed areas and unisexual ones in areas where the species

appear to be recent colonizers. Similarly, *Poratia obliterata* is bisexual in its native South American habitats, but more or less stable populations recently established in hothouse in Europe are parthenogenetic. There is also evidence of parthenogenesis in pauropods.

The life cycle of some tropical scolopendromorphs is annual, but in temperate regions a centipede life-cycle takes two or more years (Voigtländer, 2011).

All myriapods are oviparous. A lineage of centipedes (the Phylactometria, i.e., Craterostigmomorpha, Scolopendromorpha, and Geophilomorpha) have evolved brood care. In these animals, the female remains coiled (for as long as 3 months in some species) around her brood until hatching. Parental cares are virtually unknown in the other myriapod classes, with two of the rare exceptions being the millipedes *Polyzonium germanicum* (order Polyzoniida) and *Bericostenus humicola* (order Platydesmida), whose habits are comparable to those of the brood-caring centipedes. Egg brooding by the male has been recorded in a few species of platydesmidan millipedes.

A Diversity of Developmental Modes

Newborn millipedes have an incomplete number of segments and legs (**Figure 4**). In most cases, only three pairs of legs are present in the first active or inactive postembryonic stadium. New segments and leg pairs are added at each molt, sometimes inside constructed molting chambers, according to a precise schedule which is different in different orders, families, or genera. Three main developmental modes can be distinguished (Enghoff *et al.*, 1994).

In hemianamorphosis (as in the pill millipede *Glomeris*) the final number of segments and leg pairs is obtained after some molts, but the millipede must undergo further molts without segmental increase before achieving adult size and sexual maturity. In teloanamorphosis (as in polydesmidans) the achievement of the final number of segments and leg pairs coincides with obtaining the adult condition and no further molt occurs. In euanamorphosis (as in the cylindrical juliforms), new segments and leg pairs are added at every molt and the animal continues molting even after reaching sexual maturity.

The males of some western European representatives of the order Julida exhibit the peculiar developmental behavior known as periodomorphosis. Under specific adverse environmental conditions, sexually mature males of these species molt to intercalary males with regressed gonopods and suspended reproductive activity. After one or more additional molts, these animals reach a further mature stage, with newly differentiated gonopods.

Hemianamorphosis is also known in centipedes: Scutigeromorphs, lithobiomorphs, and craterostigmomorphs hatch with 4, 7, and 12 pairs of legs, respectively, and obtain the final number of 15 leg pairs following one (craterostigmomorphs) or more molts but continue molting until they reach maturity. The remaining centipedes (scolopendromorphs and geophilomorphs), however, develop by epimorphosis. That is, they reach the final number of body segments during the embryonic development; no segment or leg pair is thus added during the postembryonic life.

In symphylans, 6 or 7 of the final 11 or 12 pairs of legs are present in the first postembryonic stadium. Quite peculiar is the following developmental schedule in that first-order and second-order molts regularly alternate until maturity is reached. After each first-order molt a new two-segment unit is added, but only one more pair of legs is differentiated, whereas the second-order molts are accompanied by the differentiation of another pair of legs without addition of new segments. Mature symphylans continue to molt, up to 40 times in *S. immaculata*.

In pauropods, the first postembryonic stadium is a motionless pupoid with inarticulated traces of the antennae and of the first two pairs of legs. After a molt it changes into an active stadium with three pairs of legs. The development then proceeds by anamorphosis.

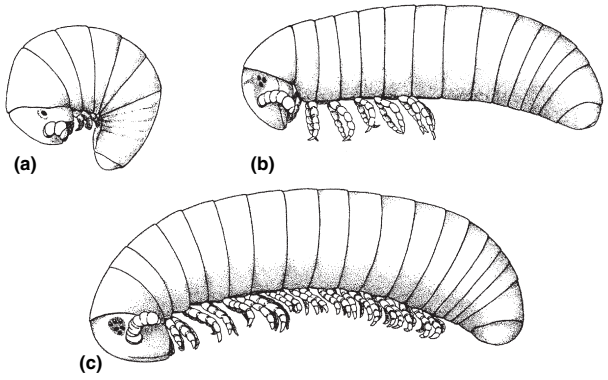


Figure 4 The first postembryonic stages of a millipede (*Anadenobolus leucostigma martinicensis*; Spirobolida): (a) larva I with 3 pairs of legs; (b), larva II with 7 pairs of legs; (c) larva III with 21 pairs of legs (adapted from Nguyen Duy-Jacquemin M (1992) Contribution a 'l'étude du développment postembryonnaire *d'Anadenobolus leucostigma martinicensis* (Chamberlin 1918). Note préliminaire. *Berichte des naturhistorisch-medizinischen Vereins in Innsbruck* 10(supplement): 133–140.

Number of Species Known and Expected

Our knowledge of myriapods is highly fragmentary and incomplete (**Table 1**). Not only does the bulk of existing species diversity remain to be described, but also the number of higher taxa appears to vary considerably as estimated by different taxonomists. In the Diplopoda, by far the largest of the myriapod classes, the situation concerning the number of orders and families, let alone genera and species, is badly equivocal to the point that Hoffman (1980) could equate the state of the art in millipede systematics to that in entomology in the middle of the nineteenth century. Thousands of taxa awaiting description are available in the collections. For example, according to Mesibov's online checklist, (http://www.polydesmida.info/millipedesofaustralia/), the collections of Australian millipedes already kept in Australia's natural history museums contain at least 1000 undescribed species, with only some 426 which have been named. Indeed, it is mostly tropical myriapod faunas that appear to be particularly rich but poorly explored.

Judging from the species described thus far, one millipede order outnumbers all others in terms of species counts –

Table 1 Myriapod diversity (approximate data), known and estimated

Class	Recent orders	Families	Genera	Species	
				Described	Estimated
Chilopoda	5	21	325	3300	6000–10,000
Diplopoda	15–16	130	1800	11,000	50,000–80,000
Pauropoda	1	5	29	700	2000–5000
Symphyla	1	2	13	200	500

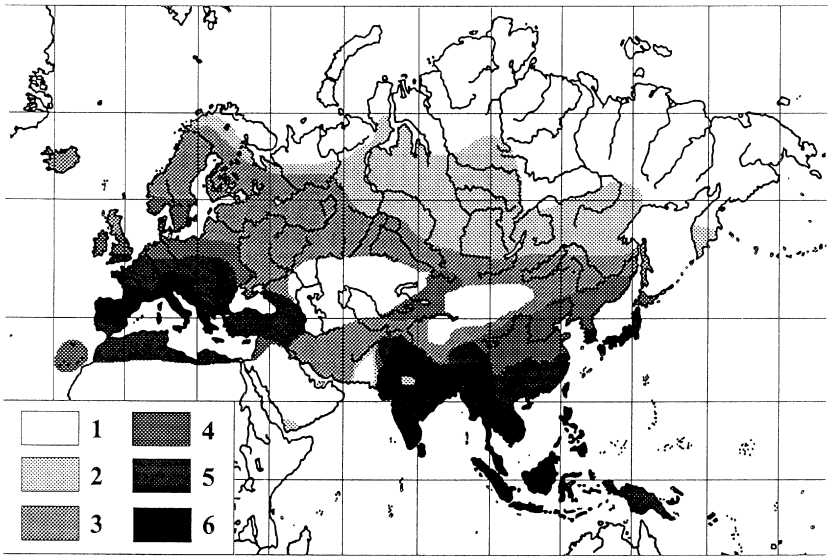


Figure 5 Patterns of millipede generic diversity in Eurasia (reproduced from Golovatch SI (1997) On the main traits of millipede diversity in Eurasia (Diplopoda). *Senckenbergiana Biologica* 76: 101–106. 1, millipede-free territories; 2, 1–4 genera; 3, 5–10 genera; 4, 11–25 genera; 5, 26–50 genera; 6, over 50 genera. The data exclude obvious anthropochores.

Polydesmida (up to one-half of the genus and species richness of the Diplopoda), followed by Spirostreptida, Chordeumatida, and Julida. The remaining orders are minor to marginal. Thus, the pentazonian order Glomeridesmida contains only a single genus with a handful of species in the Indo-Australian and Neotropical realms. Similarly, the order Siphonocryptida is only represented by six species, one each in Macaronesia, Nepal, Taiwan and Sumatra, Indonesia, and further two in mainland Malaysia. The order Siphoniulida is extreme and obviously relict, being known by a single species each in Malaya and Central America.

Probably less than 20% of the world fauna has been described to date; this is particularly true of the largest class Diplopoda (**Table 1**). Indeed, updated counts/checklists are only available for the millipede faunas of some European countries (Kime and Golovatch, 2000). The fact that the numbers of genera and species of Diplopoda currently known in Europe and the Mediterranean are similar to those known from the whole of Asia shows how poorly investigated these are in the Asian, largely tropical fauna. An overview of the geographical distribution of millipedes, as a basis to reconstructing the history of their distribution on earth has been recently provided by Shelley and Golovatch (2011).

Detailed information on the geographical distribution of all centipede species is currently available in a website (Minelli *et al.*, 2006; http://chilobase.unipd.it) and their world distribution has been recently discussed at the level of family and major geographical areas (Bonato *et al.*, 2009; Bonato and Zapparoli, 2011). A world list of Pauropoda is available, but the distributional data concerning Symphyla are still scattered in the primary literature.

Geographical Patterns

The bulk of species richness is definitely confined to tropical countries, mainly woodlands, with certain notable exceptions. Thus, some orders or families appear to be largely temperate and more or less restricted to the Northern Hemisphere, e.g., Lithobiomorpha among Chilopoda or Glomerida, Julida, Callipodida, and Polydesmidae among Diplopoda. For these groups, the hot spots of diversity are mountainous lands at the periphery of the Holarctic Region within the temperate to subtropical areas such as the Atlas, Pyrenees, Alps, Balkans, Caucasus, Tien-Shan, Himalaya, and Appalachians. **Figure 5** shows the patterns of millipede generic diversity in Eurasia.

Harsh environments such as tundra or deserts are populated by myriapods, patchily to marginally if at all, especially by Diplopoda (Golovatch and Kime, 2009) and Symphyla. However, Chilopoda, Pauropoda, and even some Symphyla

occur in permafrost areas, and a few scolopendromorph centipedes are quite characteristic desert dwellers. In contrast, high-montane habitats, however adverse, are regularly populated by myriapods, inclusive of millipedes, and sometimes even by endemic species. The highest altitudes recorded are approximately 4500 m above sea level for millipedes. The highest record for centipedes is one of the *Lithobius* species at 5545 m asl in the Nepal Himalaya. In the Andes, species of *Lamyctes* have been recorded from up to 4400 m, whereas in Africa *Lamyctes emarginatus* occurs at 4000 m on Mt. Kenya and *Lamyctes africanus* at 4200 m on Mt. Ruwenzori.

Given the hundreds to thousands of species in each of the major myriapod groups, one may get the wrong impression that these basically cryptic animals should be highly diverse in most of the local environments, particularly in productive temperate or tropical forests. However, virtually no local fauna is known to include more than two or three dozen species, depending on habitat. Given that myriapods are basically composed of very ancient, widespread, but poorly vagile and largely stenotopic forms, most of the faunules, even among the richest and monotonous-like tropical rain forest, appear different at least to some degree (Golovatch and Kime, 2009). In other terms, ancient and widespread endemism is generally characteristic of Myriapoda.

Cold areas host a very reduced fauna of millipedes. In Europe, only two species (*Polyxenus lagurus* and *Proteroiulus fuscus*) live north of the Polar Circle; only two or no millipedes are known from such areas of Siberia or from Greenland, respectively (Golovatch and Kime, 2009).

Only relatively few nonanthropochoric species display vast distributions; these tend to be particularly large among predators (= Chilopoda), especially soil dwellers. Several geophilomorphs are examples of this pattern of distribution, e.g., *Arctogeophilus glacialis* and *Pachymerium ferrugineum*. Forms populating special but widespread habitats likewise appear widely distributed, e.g., the littoral geophilomorph *S. maritima* or the huge, largely arboricolous, neotropical scolopendromorph *S. gigantea*. Active colonizers of relatively young areas/habitats, such as the trans-Siberian *Angarozonium amurense*, the pan-European *Ommatoiulus sabulosus*, or the subcorticolous pan-European *P. fuscus* among millipedes or the Euro-Siberian *Lithobius curtipes* and some other congeners among lithobiomorph centipedes, reach up to the taiga and, in part, tundra belts in the north, the youngest boreal biomes. The Mediterranean chilopod *S. cingulata* is a notable colonizer of semiarid badlands, whereas the European symphylan *Symphylella vulgaris* is a colonizer of fresh mines.

Worldwide distribution due to human agency is known for several myriapod species, mainly of European origin such as the millipedes *Cylindroiulus latestriatus* and *Ophyiulus pilosus* and the centipede *Lithobius forficatus*. The origin of the synanthropic centipede *L. emarginatus*, also widely distributed, is unknown but it is possibly from the Southern Hemisphere.

Endemism and Speciation

High numbers of endemic species and high percentages of endemics within the total millipede faunas are peculiar to some environments such as caves and some areas, such as islands and some mountain chains in the temperate regions.

The best investigated cases of intense speciation in millipedes are by far those of the genera *Cylindroiulus, Dolichoiulus*, and *Acipes* in the Macaronesian area (Enghoff, 1992). On Madeira, of 60 millipede species known from the island, at least 29 belong to the endemic *Cylindroiulus madeirae* group (Julidae) and another 6 to an endemic lineage of the genus *Acipes* (Blaniulidae). Still more conspicuous is the *Dolichoiulus* (Julidae) species swarm in the Canary Islands, which has 46 species of a total of 79 millipede species thus far recorded from the archipelago. These *Dolichoiulus* species have colonized the most diverse habitats, from the laurisilva to the xeric habitats and the caves, whereas the Madeiran *Cylindroiulus* are all confined to laurisilva forest habitats.

In the Canarian *Dolichoiulus* species swarm, conventional allopatric (interisland) speciation is only marginally responsible for the genesis of the huge diversity of the group. According to Enghoff, speciation occurred mainly within individual islands, with distribution patterns governed mainly by habitat differences between closely related and nearly sympatric species. In several cases, those species which coexist in the same habitat differ conspicuously in size. The same is often seen in Madeiran *Cylindroiulus* and *Acipes* species; this rule, however, is far from universal.

Distribution by Continents

In general, the overall distribution of myriapods is quite consistent with the world's biogeographic regionalization known since the times of Hooker and Wallace. In other words, the traditional division of lands into the Holarctic (= Palearctic + Nearctic), Afrotropical (including Madagascar and the Cape), Neotropical, Oriental (Indian + Malesian), and Australian (including New Guinea and the islands of the southern Pacific) realms is supported by the patterns demonstrated by myriapod higher taxa. Antarctica is the only continent harboring no Myriapoda, whereas, due to the particularly poorly explored tropical faunas, the distribution of genera and species between the remaining continents or biogeographic regions is highly provisional and appears extremely uneven and patchy (**Table 2**).

The Diplopoda (Shelley and Golovatch, 2011) are indigenous to all continents except Antarctica and islands/archipelagos in all temperate and tropical seas and oceans except the Arctic; the group ranges from Kodiak Island and the northern Alaska Panhandle, US, southern Hudson Bay, Canada, and near or north of the Arctic Circle in Iceland, continental Scandinavia, and Siberia to southern Argentina, the southern tips of Africa and Tasmania, and Campbell Island, subantarctic New Zealand.

Glomerida, Platydesmida, Julida, and Callipodida occur exclusively in the Northern continents, whereas Glomeridesmida, Sphaerotheriida, Siphonophorida, Spirobolida, Epinannolenidea, Spirostreptidea, and Stemmiulida are primarily confined to the southern (Gondwanan) areas except for secondary dispersals in Mexico/Central America by all but Sphaerotheriida, which are absent from the New World. Siphoniulida and Siphonocryptida, known from only two and

Table 2 Distribution of centipede and millipede diversity among the major continents/regions[a]

Class	Continent/region				
	Eurasia	*North and Central America*	*Afrotropical region*	*Australia and South Pacific*	*South America*
Chilopoda	1.3	1.5	0.5	0.5	1.1
Diplopoda	7.5	3.5	3.6	0.4	2.8

[a]Approximate number of genera in hundreds.

four areas, respectively, are declining toward extinction. Polyxenida, Polyzoniida, Chordeumatida, and Polydesmida occur on nearly all continents, whereas Cambalidea inhabit North/Central America and southeast Asia with an isolated area in Iran.

Chilopoda are present almost worldwide, but there are no records from Antarctica, most parts of Greenland, the American and Asiatic arctic islands, and from a large part of western Saharan Africa (Bonato and Zapparoli, 2011).

Centipede diversity is highest in temperate and subtropical North America and in southern Europe; to a somehow lesser extent also in southern Africa and in the Japanese archipelago. The Scutigeromorpha have maximum diversity in southern and eastern Africa, the Indian peninsula, south-eastern Asia and Australia, the Lithobiomorpha in all temperate continental areas, the Scolopendromorpha in all tropical and subtropical regions, the Geophilomorpha in North America, most of South America, the Mediterranean basin, the southernmost part of Africa, the Japanese region, and south-eastern Asia. The only two species known of Craterostigmomorpha are restricted to Tasmania and New Zealand.

Hot Spots of Biodiversity and Endemism

The bulk of species richness is confined to the tropical countries, mainly woodlands (Golovatch, 1997), although there are some notable exceptions. The opposite extreme is represented by numerous species endemic, e.g., to a single cave or mountain slope; this is a special characteristic of Diplopoda. Local cavernicolous endemics, notably troglobites, are particularly numerous in southern Europe, especially in the Balkan area, which is rich in karst. The same actually concerns cave millipedes in most of the tropical countries. Similarly, nearly every patch of Amazonian primary rain forest near Manaus, Brazil, contains (much of) its own millipede faunule, with differences being particularly drastic between various kinds of inundation versus nonflooded (= terra firme) woodlands (Golovatch and Kime, 2009). Approximately the same pattern has been observed among Pauropoda.

The Fauna of Islands

Analysis of insular biotas clearly shows the differences in colonizing capability and radiation patterns of centipedes versus millipedes. In most faunas, the ratio of millipede to centipede species is approximately 4:1, but this ratio is much lower in Hawaii, Tahiti, Samoa, and the Fiji Islands. This reveals that centipedes are apparently better dispersers than millipedes. The distribution patterns of millipedes on continental versus oceanic islands are those of poor dispersers. For instance, there are 18 millipede families represented by indigenous species in Sumatra, 17 in Japan, and 14 in Java, but there are only 10 in Australia, 9 in Madagascar, 5 in New Zealand, 3 in New Caledonia, and 1 in the Hawaii Islands. The number of endemic genera is very high in Indonesia and the Antilles, but these island groups do not have endemic families; suprageneric indigenous taxa are only known from Borneo.

The myriapod fauna of most of the oceanic islands is not simply disharmonic but also poor in species. Native myriapods are virtually absent from Kerguelen, Tristan da Cunha, and St. Helena.

Habitats and Adaptations

At least five morphological body plans and a few life-forms can be distinguished among millipedes (Golovatch and Kime, 2009). The pin-cushion millipedes (i.e., the representatives of the worldwide distributed order Polyxenida) are soft bodied, very small (usually <5 mm long), extremely bristly, swift, and capable of withstanding much drier conditions than most other Diplopoda. Many polyxenidans live under loose tree bark or are characteristic inhabitants of microcaverns and small crevices under stones, in the uppermost soil, in litter, and in similar substrates.

The pill millipedes (the largely Holarctic order Glomerida and the Afrotropical and Indo-Australian order Sphaerotheriida) are collectively termed "rollers" because of their ability to roll themselves into mostly glossy balls. Pill millipedes are generally litter dwellers, but some smaller Glomerida are either troglobionts or geobionts.

The colobognathans (i.e., the "sucking millipedes" with reduced mouthparts) and some Chordeumatida, which mostly possess flexible, worm-like, strongly tapered bodies, and shorter legs, are termed "borers."

Many Polydesmida and numerous Chordeumatida, often conspicuously ornamented and relatively shorter, long-legged, and displaying more or less strong paranota (wing-like dorsolateral projections of the diplosegments), are referred to as "wedge types" and are characteristic of forest litter.

The lifestyle, or ecomorphological type, most common and widespread among Diplopoda is that of "bulldozers" or "rammers." Their long, cylindrical, hard body with numerous diplosegments (hence, numerous pushing legs) penetrates the substrate like a bulldozer, using the broad head as a ram. Most juliform millipedes (Julida, Spirobolida, Spirostreptida, Callipodida, and some others) belong to this ecological type. This burrowing habit must have been critical in ensuring diplopods much or even most of their present-day high ecological and

geographical performance. In particular, unlike the remaining ecomorphotypes, only juliforms, among Diplopoda, appear to have colonized virtually any suitable habitat. In fact, in Juliformia belong the few littoral dwellers and deserticoles, but this body plan also provides most of millipede diversity in troglobionts and geobionts, in addition to anthropochores. Indeed, the northernmost record of a millipede belongs to the subcorticolous European species *Proteroiulus fuscus* (Julida) in the forest–tundra belt of Yamal Peninsula, Russia's north, whereas perhaps the most characteristic deserticole among millipedes is *Orthoporus ornatus* (Spirostreptida) in the southern US and adjacent parts of Mexico. Both these species are rare except in their native environments (Golovatch and Kime, 2009).

There are sound reasons to believe that, due to their capability to escape adverse conditions by burrowing in the soil, rotten logs, and similar shelters, juliforms (namely, species of the order Julida) dominate in Europe. This youngest, fully migratory nucleus of the millipede fauna, shows a clear-cut inclination to dwell on open terrain. In contrast, all the remaining millipede orders display very evident geographical trends in diversification. In Europe, even the relatively uniformly distributed Polydesmida have a remarkable center of secondary diversification in Slovenia, whereas the Chordeumatida are particularly species-rich within the Atlantic climatic zone of western Europe (Kime and Golovatch, 2000; Golovatch and Kime, 2009). Furthermore, only a few species of Julida, apparently in response to the strongly adverse conditions that existed in Europe during the Ice Age, appear to have developed a population strategy unique among terrestrial animals – the periodomorphosis strategy mentioned previously, i.e., the extension of male life by means of intercalary stadia.

In summary, most millipede lineages are currently in a phase of rapid evolution and speciation; however, there are a few apparently relict groups (Hoffman, 1980). Polydesmida seem remarkably diverse largely due to their wedge type of burrowing allowing ecological niche partitioning mainly in the litter and at the soil–litter interface. This is particularly obvious in tropical and subtropical faunas. However, examples of troglobionts, geobionts, and myrmecophilous as well as arboricolous species are about as numerous among Polydesmida as among juliforms, whose type of burrowing and lifestyle seem the most characteristic, widespread, and ecologically progressive among all recent Diplopoda.

Burrowing Myriapods

Several major myriapod groups tend to be soil dwellers. Usually, adaptations to geophily involve body elongation due to an increased number of segments. This is quite evident, for instance, in several lineages of geophilomorph centipedes, especially the Himantariidae and the Oryidae, which have a flexible worm-like body and peculiarly short appendages. This is also the case for certain scolopendromorph centipedes and some juliform and colobognathan millipedes, which are active burrowers. Body miniaturization also appears advantageous to dwelling in the crevices or burrows in the soil. This adaptation is observed in most Symphyla, numerous Pauropoda, and some Diplopoda (e.g., some representatives of Glomerida, Julida, and Polydesmida). However, only more or less loose soils appear suitable for such myriapods because harder grounds are apparently too difficult to penetrate, especially by the actively burrowing species.

Open-Country Myriapods

Short-living Polydesmida are dominant in the savanna: The adults are active during the rainy season, whereas the juveniles spend the dry season in resting conditions. Resistance to desiccation varies much between species, with the females being mostly less sensitive than the males, for unknown reasons. Cuticular waxes seem to be present in very few millipede species. One of these species is *O. ornatus* of the American semideserts; when the temperature rises more than 40 °C, it rests in a coiled condition to reduce water loss. Lowest lethal temperatures for millipedes are in the range of −5 to −7 °C, and the highest temperatures are between 36 and 41 °C. Preferred temperatures are clearly lower for hygrophilous forest species (4–18 °C) than for xerophilous ones (e.g., 26–32 °C for *O. sabulosus*, a species common in open environments in southern and eastern Europe).

Special Habitat Myriapods

A few species have semiaquatic habits. One of them is the littoral julid *Thalassisobates litoralis*, which can be found among the stranded decaying matters on the seashore. However, this species does not seem to spend much time under water, as do some cavernicolous species from southern Europe (several julids and the polydesmid *Serradium semiaquaticum*) and some Amazonian species, including *Myrmecodesmus adisi*, whose juveniles spend up to 11 months under water feeding on algae (Golovatch and Kime, 2009).

Both morphological and behavioral adaptations have been observed in myriapods that spend prolonged periods of time under water. Conspicuous are the microtrichia in their spiracles and the cerotegument covering their body which enable flood resistance for weeks and even months compared to the simple flood tolerance (up to a few days) exhibited by non-specialized forms. These morphological adaptations are present in a few millipedes of the seasonally inundated forests of Amazonia; they are also present in the polydesmid *S. semiaquaticum*, which inhabits caves near Verona (northern Italy), whereas its sister species *Serradium hirsutipes* does not have spiracular microtrichia and, as expected, does not have the flood resistance of its congener.

Regarding centipedes, resistance of seashore geophilomorphs to submersion in sea water is not very conspicuous, being 12–24 h in *S. maritima* and up to 36 h in *H. submarina*. Longer resistance to submersion in fresh water has been shown for the lithobiomorph *L. emarginatus* from European floodplains and from populations living along the Amazon River (personal communication from late colleague Adis).

Centipedes living in the floodplains must often face the problem of how to survive high water, basically choosing between running away and resisting somehow on the same spot. Several European lithobiomorphs, for example, *Lithobius microps*, *L. forficatus*, *Lithobius mutabilis*, and *Lithobius curtipes*,

are quite tolerant to inundation: Specimens of *L. curtipes* were found alive after 22 days inundation.

A few Amazonian species of Geophilomorpha show morpho-anatomical modifications in their respiratory opening (spiracles) for surviving submersion as adults; the henicopid lithobiomorph *Lamyctes adisi* has a dormant submerged egg stage; scolopendromorph and other geophilomorph centipedes climb up the trees to live there during the inundation period (Adis *et al.*, 1996; Morais *et al.*, 1997).

Behavioral adaptations to the seasonally flooded forest of Amazonia have also been recorded in the other myriapod groups. Regarding Symphyla, members of the Scolopendrellidae have a dormant stage of 5–7 months duration, which in the case of *Ribautiella amazonica* is spent among roots. Owing to the anoxic conditions of the flooded soil, these arthropods are likely to recur to anaerobic metabolism. A similar situation is presumed for Pauropoda living in the same environment because *Scleropauropus tarumamirimi* has been found to spend 5–7 months under water (Adis and Scheller, personal communication). Myriapods lacking similar morphological or physiological adaptations must rely on vertical migrations along the tree trunks to survive the flood period in the upper trunk or canopy region (e.g., *Cutervodesmus adisi* (Diplopoda), *Hanseniella arborea* (Symphyla), and some Geophilomorpha). The migratory behavior of these species, however, is the most conspicuous aspect of their complex ethological and physiological adaptations. No reproduction occurs during the time they spend on the trunks, and trunk ascent is associated with changes in local climatic conditions ("dry" vs. rainy season) and with macroclimatic influences such as the El Niño Southern Oscillation.

A large fraction of myriapods are represented by cave dwellers. Adaptations to life in caves are often opposite to those to geophily, frequently including larger sizes compared to their epigeic counterparts (Kime and Golovatch, 2000; Golovatch and Kime, 2009), often with strong paranota or hairs, elongated antennae and appendages, and reduced or missing ocelli. Among myriapods, troglobites are known mostly among millipedes and, to a lesser degree, centipedes. Many myriapods have developed adaptations to cave life and can be defined as troglobionts.

Regarding centipedes, approximately 50–60 troglobitic species of lithobiomorphs and a few scolopendromorphs are known from caves of southern Europe: The Pyrenees, Sardinia, Italy, the Balkans, Greece, Turkey, Algeria, Morocco, Central America, and Australia. A troglobitic geophilomorph (*Geophilus persephones*) has recently been described from the cave Gouffre de la Pierre Saint-Martin in the French Pyrenees; it is blind, as expected, but in this case the blindness is not a specific adaptation to cave life because all geophilomorphs lack eyes.

Modified mouthparts have evolved convergently in several unrelated genera of hygrophilous cave-dwelling millipedes from southern Europe and the Caucasus (Julidae: *Leucogeorgia*, *Typhloiulus*, and *Trogloiulus*; Blaniulidae: *Vascoblaniulus*; Polydesmidae: *Serradium*). In these forms, the biting or masticatory part of the mandible is reduced, whereas its pectinate lamellae are hypertrophied, suggesting a function like a filter, used in collecting organic material suspended in water. Investigations on the gut content of *S. semiaquaticum* from Italian caves, however, have failed to substantiate this expectation.

Appendix

List of Courses

1. The M.V. Lomonosov Moscow State University, Invertebrate Zoology and Soil Biology

See also: Arthropods (Terrestrial), Amazonian. Invertebrates, Terrestrial, Overview

References

Adis J, Golovatch SJ, Hoffman RL, Hales DG, and Burrows FG (1998) Morphological adaptations of the semiaquatic millipede Aporodesminus wallacei Silvestri 1904 with notes on the taxonomy, distribution, habitats and ecology of this and a related species (Pyrgodesmidae Polydesmida Diplopoda). *Tropical Zoology* 11: 371–387.

Adis J, Minelli A, de Morais JW, Pereira LA, Barbieri F, and Rodrigues JMG (1996) On abundance and phenology of Geophilomorpha (Chilopoda) from Central Amazonian upland forests. *Ecotropica* 2: 165–175.

Barber AD (2009) Littoral myriapods: A review. *Soil Organisms* 81: 735–760.

Bonato L, Bevilacqua S, and Minelli A (2009) An outline of the geographical distribution of world Chilopoda. *Contributions to Natural History, Bern* 12: 489–503.

Bonato L and Zapparoli M (2011) Chilopoda – Geographical distribution. In: Minelli A (ed.) *Treatise on Zoology – Anatomy, Taxonomy, Biology. The Myriapoda*, vol. 1. pp. 327–337. Leiden: Brill.

Bush SP, King BO, Norris RL, and Stockwell SA (2001) Centipede envenomation. *Wilderness & Environmental Medicine* 12: 93–99.

Chagas Jr. A, Edgecombe GD, and Minelli A (2008) Variability in trunk segmentation in the centipede order Scolopendromorpha: A remarkable new species of *Scolopendropsis* Brandt (Chilopoda: Scolopendridae) from Brazil. *Zootaxa* 1888: 36–46.

Edgecombe GD (2004) Morphological data, extant Myriapoda, and the myriapod stem-group. *Contributions to Zoology* 73: 207–252.

Edgecombe GD (2011) Phylogenetic relationships of Myriapoda. In: Minelli A (ed.) *Treatise on Zoology – Anatomy, Taxonomy, Biology. The Myriapoda*, vol. 1, pp. 1–20. Leiden: Brill.

Eisner T, Alsop D, Hicks K, and Meinwald J (1978) Defensive secretions of millipedes. In: Bettini S (ed.) *Arthropod Venoms (Handbook of Experimental Pharmacology*, vol. 48. pp. 41–72. Berlin: Springer.

Enghoff H (1990) The ground-plan of the chilognathan millipedes. In: Minelli A (ed.) *Proceedings of the 7th International Congress of Myriapodology*, pp. 1–21. Leiden: Brill.

Enghoff H (1992) Macaronesian millipedes (Diplopoda) with emphasis on endemic species swarms on Madeira and the Canary Islands. *Biological Journal of the Linnean Society* 46: 153–161.

Enghoff H (1994) Geographical parthenogenesis in millipedes (Diplopoda). *Biogeographica* 70: 25–31.

Enghoff H, Dohle W, and Blower JG (1994) Anamorphosis in millipedes (Diplopoda) – The present state of knowledge with some developmental and phylogenetic considerations. *Zoological Journal of the Linnean Society* 109: 103–234.

Golovatch SI (1997) On the main traits of millipede diversity in Eurasia (Diplopoda). *Senckenbergiana Biologica* 76: 101–106.

Golovatch SI and Kime RD (2009) Millipede (Diplopoda) distributions: A review. *Soil Organisms* 81: 565–597.

Hoffman RL (1980) *Classification of the Diplopoda*. Genève: Muséum d'Histoire Naturelle.

Kime RD and Golovatch SI (2000) Trends in the ecological strategies and evolution of millipedes (Diplopoda). *Biological Journal of the Linnean Society* 69: 333–349.

Minelli A, Bonato L, Dioguardi R, *et al.* (2006 onwards) *Chilobase: A Web Resource for Chilopoda Taxonomy.* Publicly searchable at http://chilobase.bio.unipd.it.

Morais JW, Adis J, Berti-Filho E, Pereira LA, Minelli A, and Barbieri F (1997) On abundance, phenology and natural history of Geophilomorpha from mixedwater inundation forest in Central Amazonia (Chilopoda). *Entomologica Scandinavica* 51(supplement): 115–119.

Nguyen Duy-Jacquemin M (1992) Contribution à l'étude du développment postembryonnaire d'*Anadenobolus leucostigma martinicensis* (Chamberlin 1918). Note préliminaire. *Berichte des naturhistorisch-medizinischen Vereins in Innsbruck* 10(supplement): 133–140.

Nguyen Duy-Jacquemin M (1996) Systématique et biogéographie des diplopodes pénicillates des Iles Canaries et du Cap Vert. *Mémoires du Muséum National d'Histoire Naturelle, Paris* 169: 113–126.

Regier JC, Shultz JW, Zwick A, *et al.* (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* 463: 1079–1083.

Shelley RM and Golovatch SI (2011) Atlas of myriapod biogeography. I. Indigenous ordinal and supra-ordinal distributions in the Diplopoda: Perspectives on taxon origins and ages, and a hypothesis on the origin and early evolution of the class. *Insecta Mundi* 0158: 1–134.

Undheim EA and King GF (2011) On the venom system of centipedes (Chilopoda), a neglected group of venomous animals. *Toxicon* 57: 512–524.

Voigtländer (2011) Chilopoda – Ecology. In: Minelli A (ed.) *Treatise on Zoology – Anatomy, Taxonomy, Biology. The Myriapoda,* vol. 1. pp. 309–325. Leiden: Brill.

Natural Extinctions (not Human Influenced)

Christopher N Johnson, James Cook University, Townsville, QLD, Australia

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 305–315, © 2001, Elsevier Inc.

Glossary

Conditions Physical features of the environment, such as substrate type, ambient temperature, or salinity, that affect the ability of organisms to survive, grow, and reproduce.
Dispersal The movement of an individual organism away from its place of origin to the place where it breeds (also, movement by an adult from one breeding location to another).
Pseudoextinction The disappearance of a species from the fossil record due not to the death of all its members but to an evolutionary change that results in it being classified as a new species.
Quaternary The past 2 million years (approximately) of Earth history, including the Pleistocene and Holocene (or Recent) epochs, and characterized by the extreme fluctuations in global temperature that produce the ice ages.
Resources Physical and biotic features of the environment, such as shelter sites or foods, that are required by organisms and are consumed such that use by one individual reduces their availability to others.
Secondary extinction Extinction of a species resulting from the extinction of another species on which it relies.

Introduction

Most species have durations that are very short relative to the history of life, and it is likely that 99% of all species that have ever lived are now extinct. Extinction, therefore, is very much a natural process that produces continuous turnover in the membership of biological assemblages as species are steadily lost by extinction and replaced by speciation. Natural extinctions may be divided into two kinds: those that happened during geologically brief periods of crisis when many taxa disappeared (the mass extinction events) and those during the long intervals of "background" time between mass extinctions. Mass extinctions have attracted a great deal of interest because they bit so deeply into standing biodiversity and because the biotas that reformed after mass extinctions were often quite different from those that existed before; they dramatically reshaped world patterns of biodiversity. However, 95% or more of the total number of extinctions have taken place outside mass extinctions (May *et al.*, 1995). These background extinctions have been most important in producing turnover in species assemblages.

Differences in the probability of extinction for different types of organisms interact with rates of origination of new lineages to shape patterns of biodiversity. To understand patterns of biodiversity it is therefore essential that the process of extinction be understood. In particular, we need to know whether extinction rates vary among different types of organisms and to identify the characteristics of taxa that make them more or less vulnerable to extinction. Extinction can be studied at the level of higher taxa—genera, families, and so on—but this article is concerned primarily with extinctions of species.

Natural Causes of Extinction

The causes of extinction can be divided into two types: changes in the environment and interactions with other species.

Changes in the Environment

Populations of all species fluctuate in abundance. A large part of this variation is due to chance variation in environmental conditions, such as variations in weather or events such as cyclones, which can be regarded as environmental "accidents." Species vary in their ability to resist accidents, but for any species there will eventually come an environmental event of sufficient magnitude to wipe it out, or an unlucky succession of smaller blows that drive its abundance down to zero, even when there is no trend in average conditions.

Additionally, species may be driven extinct by directional changes in the environment, such as changes in temperature to levels outside their tolerance or the disappearance of

key habitats. Species can respond to such changes in the environment by (i) acquiring adaptations to the changed conditions ("evolving out of trouble"), (ii) shifting their geographic ranges to remain within a set of suitable conditions ("moving out of trouble"), or (iii) going extinct. If the pace of environmental change is fast relative to the rate of evolution, the first of these responses is much less likely than the other two. The fossil record provides many examples of extinctions apparently driven by environmental changes of many different kinds, especially during mass extinctions events. Environmental changes have been much more extreme and rapid during some periods of Earth history than during others, but environmental conditions have never been truly static. Global temperatures have trended downwards for much of the past 50 million years, and at a finer temporal scale variations in the earth's orbit produce fluctuations in climate with periodicities in the range of 10,000 to 400,000 years. These are called Milankovitch cycles, and during the past 2 million years they have produced the succession of ice ages, but they must have forced rapid swings in climate throughout Earth history (Bennett, 1997).

Interactions with Other Species

The population growth rates of species are constrained by the species predators, parasites, and competitors. The abundance of a species will be reduced if its natural enemies become more abundant or evolve greater efficiency, or if its geographic range is invaded by a new enemy to which it has not evolved defenses.

It is generally difficult to reconstruct such interactions in the fossil record, but some patterns may reflect the impact of predators on prey species. For example, the number of species of endobyssate bivalves gradually declined through the Mesozoic. These species were abundant, immobile, and lay partly exposed on the open seafloor. Their decline coincided with radiations of marine crabs, teleost fish, and carnivorous snails, groups that account for most predation on modern bivalves and to which the endobyssates must have been vulnerable (Stanley, 1979). The few present-day survivors of the group live in conditions with low predator pressure. Recent experience shows that prey species can be rapidly driven extinct by unfamiliar predators that invade their habitat. The carnivorous snail *Euglandina rosea*, for example, has caused the extinction of hundreds of species of endemic snails, including 600 of more than 1000 original species from the Hawaiian Islands, since its introduction to many Pacific islands in the 1970s. Such a rapid course of events would be unresolvable in the fossil record.

Probably, the impacts of environmental changes and of other species interact in subtle ways to cause extinction. For example, a species might experience a slight environmental change that reduces its population without driving it extinct and to which it could readily adapt given sufficient time. However, the same change might favor a competitor or trigger a range expansion by an unfamiliar predator suited to the new conditions. Interactions of this kind could mean that quite small environmental changes could have dramatic consequences leading ultimately to extinction.

Lifetimes of Species

Variation among Taxonomic Groups

The oldest extant species known is the tadpole shrimp, *Triops cancriformis*, a small freshwater crustacean found in temporary pools in arid regions of Eurasia and north Africa that is indistinguishable from 180-million-year-old fossils bearing the same name. Most species do not live to this age, as shown in **Table 1**.

The lifetime of a species begins with its origin in a speciation event and ends either when it dies out, leaving no descendants, or when its characteristics have been sufficiently changed by evolution that it is classified as a new species. The first type of disappearance of species is "real" extinction, and the second is referred to as pseudoextinction. This distinction is important but not easy to draw in practice. Our understanding of species life spans derives from study of the fossil record, and when a recognizable species disappears from the record it can be difficult to determine whether it has died out or evolved into something else. Information in **Table 1** is based on real extinctions when the distinction can be made,

Table 1 Durations of species and extinction rates for major groups of organisms in the fossil record

Taxon	*Estimated mean species duration (millions of years)*	*Extinction rate (extinctions per million species years)*
Single-celled organisms		
Diatoms	8	0.12
Dinoflagellates	13	0.08
Planktonic foraminifera	7–20	0.14–0.05
Benthic foraminifera	25	0.04
Plants		
Early vascular plants	10–14	0.1–0.07
Pteridophytes	7–15	0.14–0.07
Gymnosperms	2–15	0.5–0.07
Monocots	4	0.25
Dicots	3	0.33
All invertebrates	11	0.09
Reef corals	20	0.05
Mollusks:		
Marine gastropods	10	0.1
Marine bivalves	15	0.07
Mesozoic ammonites	1	1
Upper Cambrian trilobites	1.3	0.77
Marine ostracods	8	0.12
Silurian graptolites	2	0.5
Echinoderms		
Echinoids	6	0.17
Crinoids	6.7	0.15
Bryozoans	12	0.08
Freshwater fish	3	0.33
Birds	2.5	0.42
Mammals	1.4	0.71
Cenozoic mammals	1–2	1–0.5
Horses	4	0.25
Primates	1	1

but most estimates include an unknown combination of real and pseudoextinctions.

There is considerable variation in species life spans among different groups of organisms. In general, it seems that species of unicellular organisms last longer than species of multicellular organisms, and invertebrates appear to last longer than vertebrates. The most long-lived animal groups tend to be marine rather than terrestrial, although too little is known about species durations of marine vertebrates or terrestrial invertebrates to judge differences between marine and terrestrial species independent of the difference between vertebrates and invertebrates. Some marine invertebrate groups that are now entirely extinct (the trilobites, ammonites, and graptolites) had short species durations, even though they were successful, abundant, and species rich for long periods of geological time. Among plants, species durations were longer in groups that appeared early in the evolutionary history of plants but have tended to shorten in recently evolved groups.

The inverse of the typical duration of species in a given taxonomic group can be used as a measure of the probability of extinction per unit time for species belonging to that group: Long durations equate to low probabilities of extinction. In **Table 1**, probabilities of extinction have been expressed as the rates of extinction per million species years. Thus, marine gastropods typically last 10 million years, so a sample of 1 million marine gastropod species would be expected to experience 0.1 extinctions per year.

Variation within Taxonomic Groups

The distributions of species lifetimes within higher taxa are typically right-skewed: Most species have short durations, but a small minority are long-lasting (**Figure 1**). There could be two causes for this pattern. First, only a few species might have characteristics that make them intrinsically resistant to extinction, whereas most are sensitive to extinction risk. Second, species might not vary in their susceptibility to extinction, but if many independent factors can cause extinction only a small minority of species will be lucky enough to avoid extinction for long. Probably both factors play a part in shaping distributions of species durations.

Risk of Extinction in Relation to Species Age

Does the likelihood that a species will go extinct depend on its age? One might think that it should because the longer a species stays in existence the more adaptations to its environment it can accumulate; older species should therefore be better at avoiding extinction. This idea can be tested by examining species survivorship curves. A species survivorship curve shows the relationship between age and the proportion of species in a taxon surviving to each age. Such curves are typically log linear; that is, when the proportion surviving is plotted logarithmically its decline with age approximates a straight line. This shows that the proportion of species going extinct is approximately constant with respect to age.

The constancy of extinction probability with species age can be explained by assuming that the environment of any species is continually changing so that even if the species continually adapts, its fitness lags behind the current condition of the environment. Thus, no matter how long a species lasts it can never reach a condition of optimal adaptation. There are two forms of this explanation. The first assumes that the probability of extinction for a species is determined primarily by its interactions with other species and that each species is under constant selection pressure to increase its fitness relative to these other species. However, adaptations in one species that increase abundance at the expense of other species will be met by counteradaptations from them so that, on average, no species improves its fitness with time. This is the Red Queen hypothesis, named after the character in Lewis Carroll's *Alice Through the Looking Glass* in whose world "it takes all the running you can do, just to stay in one place." The Red Queen hypothesis predicts that extinction rates will be

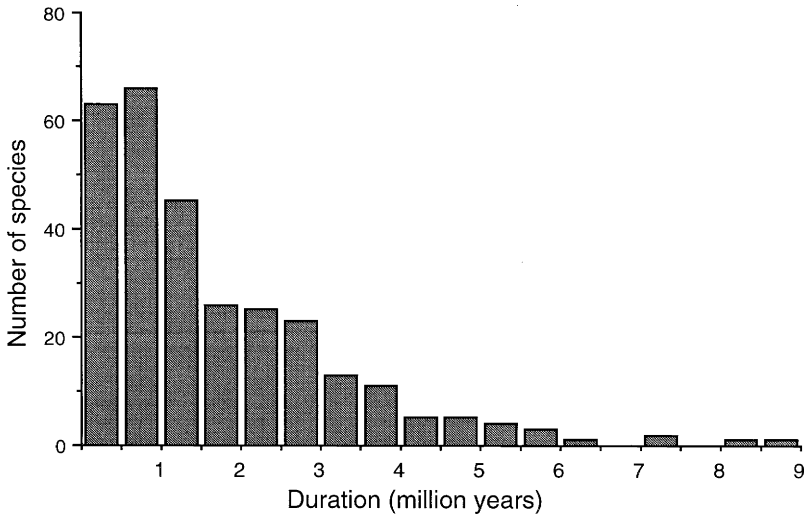


Figure 1 Species durations of Silurian graptolites of the British Isles. (Reproduced from Stanley SM (1979) *Macroevolution: Patterns and Process*. San Francisco: Freeman, with permission from W.H. Freeman and Company.)

approximately constant with respect to absolute time. The second form of the explanation assumes that it is the abiotic environment that is continually changing. Changes in the abiotic environment will tend to be episodic and will therefore produce fluctuations in extinction rates with time, but their effects will still be independent of species age. It is not clear which form of the hypothesis is closer to the truth.

The Selectivity of Extinction

Extinction may be largely a matter of the bad luck of environmental change, exacerbated perhaps by pressure from other species; however, the fact that species durations vary so widely among higher taxa suggests that biological characteristics of species influence their susceptibility to extinction. Three types of studies have helped to identify these characteristics.

First, the attributes of species that have short versus long durations in the fossil record can be compared. A common variant of this technique compares species that persisted through mass extinction events with those that did not. Second, the living faunas of "land-bridge islands" can be compared with their presumed original faunas. A land-bridge island is an island formed as a result of isolation from the mainland by rising sea levels during the Pleistocene. Such islands typically have fewer species than do similar areas of mainland habitat to which they were once connected, reflecting extinctions caused by the pressure of reduced habitat area. Third, extinction of local populations can be observed directly in contemporary ecological studies. The second and third approaches focus on population rather than species extinctions, but the extinction of a species is the end point of a series of population extinctions, so such studies may still reveal traits that correlate with risk of extinction at the species level.

The following traits consistently emerge from many studies of the selectivity of extinction: rarity, dispersal ability, body size, and specialization.

Rarity

The commonness or rarity of a species is a function of two factors: its geographic range and its population density where it occurs. Both components of rarity influence extinction risk.

Geographic Range

There is strong evidence from the fossil record (mainly for marine invertebrates) that species with large ranges have lower extinction rates (Jablonski, 1995; McKinney, 1997). There are probably two causes for this effect. First, at a given scale, a disturbance will affect a smaller proportion of the range of a widespread rather than of a geographically restricted species. In the extreme, a single catastrophic event may wipe out a localized species but have little impact on a widespread species. Second, even if an environmental change affects a very large area, it is likely that some populations of widespread species will persist in isolated refuges that provide some local protection from the change. Species that hang on in such refuges may reinvade their original ranges once conditions improve. This is probably part of the explanation for the existence in the fossil record of "Lazarus species" that vanish during periods of environmental crisis and then reappear much later. The effect of range size on extinction risk tends to be strongest for background extinctions, and it may weaken or disappear in mass extinctions. This seems to have been the case for marine mollusks, for example (Jablonski, 1991), and is probably because mass extinctions were caused by events that affected such large areas and had such profound impacts that even widespread species were susceptible to them.

Local Abundance

Population density is generally not well represented in the fossil record, but patterns of extinction of populations on land-bridge islands and in the present day show that local extinction is more likely for species with low population densities. For example, Foufopoulos and Ives (1999) found that reptile species with low population densities were more likely to go extinct from land-bridge islands in the Mediterranean Sea. The causes of extinction of small populations have been widely discussed in the literature on conservation. Briefly, small populations are more vulnerable than large populations because (i) they may go extinct more quickly when chance environmental variation causes fluctuations in abundance; (ii) they are affected by chance demographic fluctuations, such as occasional production of biased sex ratios of offspring, that would be averaged out in large populations; (iii) they may lose genetic variation and experience high levels of inbreeding and inbreeding depression; and (iv) they may be subject to Allee effects—that is, social or reproductive dysfunction as a direct result of low numbers.

The Relationship between Range and Abundance

The components of rarity—geographic range size and population density—are generally positively related among species: Species with small geographic ranges also tend to have low population densities, and wide-ranging species have high population densities. This pattern has been found in many terrestrial plant and animal taxa, although it remains almost unstudied in marine organisms. Why there should be a positive relationship between range and abundance is not clear. Several ecological mechanisms have been proposed as its cause, but two ideas have been especially influential. First, Brown (1995) argued that niche breadth is positively correlated with both range size and population density because species with broad niches can exist under a wider range of conditions and use more types of resources than can species with narrow niches; variation among species in niche breadth therefore produces a positive relationship between range and density. Second, species that reach high local densities should be both more resistant to extinction on habitat patches because their populations are larger and produce more migrants that are able to recolonize habitat patches after local extinctions (Hanski, 1999). High local abundance therefore results in wide geographic distribution.

To the extent that geographic range and local abundance affect extinction risk independently of one another, the correlation between them should exaggerate differences among species in extinction risk. Rare species face double jeopardy:

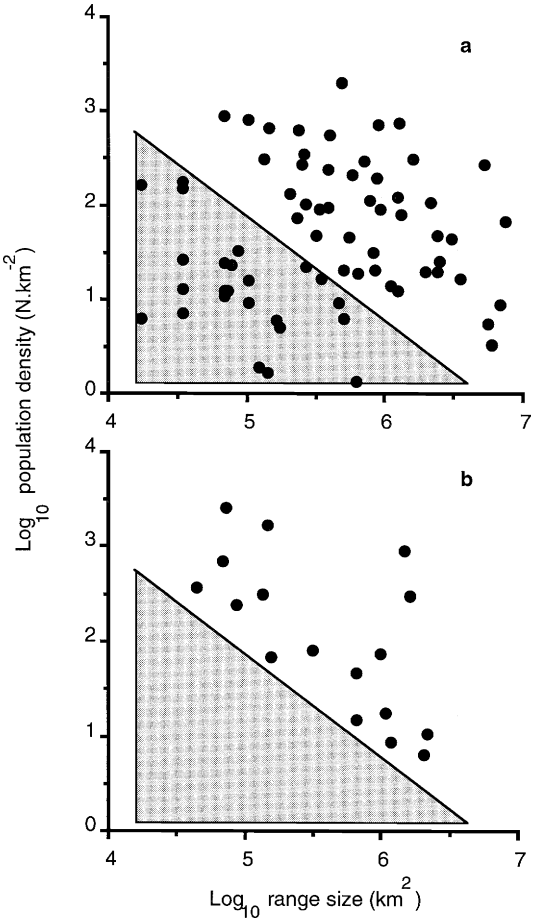


Figure 2 Relationships between size of geographic range and population density for species of Australian marsupials: (a) all species other than those defined as ancient and (b) ancient species only. Ancient species are those that diverged from their closest living relative more than 4 million years ago and have therefore demonstrated resistance to extinction. The shaded triangle defines the region of distribution-abundance space from which species would need to have gone extinct to produce the pattern observed among ancient species. (Reproduced with permission from Johnson CN (1998) Species extinction and the relationship between distribution and abundance. *Nature* 394: 272–274.)

The vulnerability that comes from having a small range is compounded with the vulnerability due to low abundance. Common species, on the other hand, are likely to be highly resistant to extinction because they combine high abundance with large ranges. For example, Johnson (1998) showed that among Australian marsupials, ancient species (those that diverged from their closest relative more than 4 million years ago) are very unlikely to have both small ranges and low population densities, although this combination occurs frequently among young species (**Figure 2**). This suggests that species with low abundance and small ranges tend to be short lived. There are ancient marsupials with small ranges but they have unusually high densities, and conversely there are species with low densities but they have unusually large ranges, implying that a high density can compensate for the vulnerability that comes from having a small range and vice versa.

Spatial Structure of Populations

Another aspect of rarity is ubiquity or patchiness of distribution within the geographic range. The distributions of practically all species are discontinuous at some scale, but the degree of patchiness varies, reflecting specialization for habitats or resources that are themselves patchily distributed. In a species that has a patchy distribution, metapopulation dynamics are likely (Hanski, 1999). That is, the risk of extinction of populations on discrete patches of habitat may be high, but local populations are eventually reestablished by migrants from other populations. Recolonization will involve a time lag that depends on the distance between patches and the dispersal ability of the species.

A general deterioration in the quality of habitat for a species is likely to cause contraction of larger habitat patches and the disappearance of smaller ones, causing local extinctions and simultaneously increasing the distance between surviving patches and thus reducing the probability of migration. Patchily distributed species should therefore be more likely to go extinct as a result of environmental change than should continuously distributed species. The sensitivity of patchy species to environmental change has been demonstrated for insects in Great Britain subject to human-caused changes (Webb and Thomas, 1994), but it is likely to be general to any form of shift in environmental conditions. Patchiness of distribution is loosely correlated with the other components of rarity so that widespread and locally abundant species tend to be continuously distributed within their ranges, whereas rare species are likely to be patchy.

Dispersal Ability

Dispersal ability is positively associated with species longevity for two reasons. First, species that disperse widely tend to have large geographic ranges, and their resistance to extinction can be a direct result of large range. Marine gastropod snails, for example, have two different forms of development. In some, the egg is released into surface waters and develops into a larval form that feeds in the plankton and drifts widely before settling and developing into an adult snail. This form of development promotes wide dispersal. Others have direct development in which eggs and young grow up close to the parent. Species with planktonic development have larger geographic ranges and longer durations in the fossil record than do species with direct development (Jablonski, 1995).

Second, species that disperse widely are better able to recolonize after going extinct from part of the range, and they can more rapidly shift their ranges to track changes in the distribution of their preferred climates and habitats. During the Quartenary in the Northern Hemisphere, the distribution of habitats changed very rapidly as ice sheets repeatedly advanced and receded. These changes produced remarkably few extinctions among Northern Hemisphere beetles. Instead, species shifted their geographic ranges to keep pace with the changing distribution of habitats, an effect that was more dramatic in flighted than in flightless species (Coope, 1995). Probably, dispersal ability and local abundance interact to confer high resistance to extinction because high local abundance means that large numbers of dispersers are produced,

resulting in the potential for rapid shifting of range boundaries as conditions change. This combination of characteristics might protect species that have specialized habitat requirements and small geographic ranges and would otherwise be vulnerable to extinction.

Body Size

It often seems to be the case that large-bodied species are at higher risk of extinction than small-bodied species. Although there are exceptions, this pattern is reasonably consistent in the fossil record both for background extinctions and for mass extinctions (McKinney, 1997), and it also emerges in extinctions on land-bridge islands (Brown, 1995). There are several reasons why large-bodied species might be particularly vulnerable to extinction:

1. Potential rate of population increase declines with body size because large-bodied species generally have longer generation times and lower fecundity than do small-bodied species. This means that large-bodied species will generally recover more slowly from population declines, and because they produce fewer offspring they may be slower to track habitats than will small-bodied species.
2. Individuals of large-bodied species generally need larger areas, and large-bodied species are therefore strongly affected by declines in habitat area.
3. Population density tends to decline with body size so that large species are often naturally rare. This is partly compensated by a tendency for size of geographic range to increase with body size, but this relationship is often weak or absent so that total population size is usually much less for large than small species. The decline in population density with body size is clearest when species of very different size (e.g., mice and elephants) are compared, but in guilds of ecologically similar species it is often the case that density increases with body size, possibly because larger species are better competitors for resources (Cotgreave, 1993). At this finer scale of comparison, therefore, large-bodied species might be more resistant to extinction than small-bodied species.

Specialization

Species can be classed as "specialists" or "generalists" on three criteria: the range of conditions that they are adapted to tolerate, the range of resources that they are able to use, and the degree of their evolved dependence on a small number of other species.

Conditions

Species that are narrowly adapted to environmental conditions are likely to be the first to go extinct when the environment changes. This is difficult to demonstrate in the fossil record, however, because it is not possible to measure directly the environmental tolerances of fossil species. Instead, environmental tolerances are inferred from geographic distributions. Species with small geographic ranges will necessarily occupy a narrow range of climate zones and habitats, but it does not follow from this that their environmental tolerances are narrow. Some species that could potentially be widely distributed have small ranges because of population history, geographic barriers to movement, or poor dispersal ability or because range expansion is prevented by interactions with other species (competitors, predators, and so on). Therefore, although there is abundant evidence from the fossil record that species with small geographic ranges are extinction-prone, there is much less evidence for the commonsense view that the breadth of tolerance of environmental conditions directly affects extinction risk.

Resources

Species that depend on a narrow range of types of resources are likely to be more sensitive to changes in resource abundance than generalists that can easily switch resources. This vulnerability may be partly compensated by the fact that specialization on a particular resource may be more likely to evolve if that resource is abundant and widespread. For example, feeding on grasses by mammals requires extensive specialization of the teeth and digestive system, and increases in body size, to overcome the abrasiveness and poor nutritional quality of grasses. However, because of the abundance of grasses many mammals have evolved these adaptations, and grazing mammals typically have high local abundance and large geographic ranges. Nonetheless, grazing mammals in Africa have suffered higher extinction rates since the Miocene than mixed grazer/browsers, as the extent of grasslands has fluctuated during the Pliocene and Pleistocene (Vrba, 1987).

Interactions

Specialization is taken a step further in species that have evolved a close dependence on one or a small number of other species. For example, many herbivorous insects feed on only one species of plant, and many predators attack only one or a small number among many possible species of prey. Of particular interest are mutualistic interactions, in which species provide benefits to one another. Figs, for example, rely on fig wasps for pollination, and fig wasps in turn lay their eggs only in the flowers of figs. This interaction tends to be highly species specific, with each species of fig visited by only one species of fig wasp and each partner in the interaction completely dependent on the other for reproduction. Such tight species specificity results from coevolution, in which each species in the interaction evolves special characteristics in response to evolutionary changes in the other to increase its benefit from the interaction.

Specialization of this kind is classically regarded as an extinction trap because the specialist will inevitably go extinct if the species that it depends on goes extinct or becomes very rare. This view is probably an oversimplification. Careful study of some species-specific interactions has revealed more flexibility and greater potential for rapid evolutionary response to changes in the abundance of interacting species, including the ability to switch to new partners, than was previously assumed (Thompson, 1994). Also, mutualistic interactions have the general effect of increasing the geographic range and abundance, and stabilizing the population dynamics, of both partners in the interaction. These ecological benefits may at

least partly compensate for the vulnerability caused by dependence on another species.

Because interactions between species are not revealed in detail in the fossil record, there is little direct information on rates of secondary extinction. It is sometimes possible, however, to evaluate the risks of secondary extinction from study of living communities. Many plants cannot set seed without cross-pollination and rely on animals to transfer pollen. Such plants should be vulnerable to reproductive failure, and possibly extinction, if their pollinators decline, and species that have only one or a small number of pollinators are likely to be especially vulnerable. This vulnerability can be reduced by traits such as the ability to propagate vegetatively that reduce demographic dependence on seeds. Bond (1995) showed that in some plant communities there are no species that have both a high dependence on animal pollinators for seed set and a high demographic dependence on seeds, despite wide variation among species in both sets of characteristics (**Figure 3**). One explanation for this pattern is that such species are unusual because their high dependence on other species means that they are very likely to go extinct.

Reproductive and Life History Traits

A wide variety of reproductive and life history traits may influence the likelihood of extinction in some circumstances. Perhaps the most important is sexual reproduction. Sexual lineages appear able to evolve more quickly than parthenogenetic ones because genetic recombination generates more variability among individuals on which natural selection can act. Sexual species may therefore be better able to adapt to changed environmental conditions or to meet new challenges from other species. The taxonomic distribution of parthenogenetic lineages suggests that they are short lived: With few exceptions, no higher taxa (genera, families, etc.) consists solely of parthenogens; almost all parthenogens have close sexual relatives and are likely to be recently derived from sexual species (Maynard Smith, 1978).

Species with resting or resistant stages in their life cycle may be better able to pass through environmental catastrophes than species in which all life stages are vulnerable to environmental changes. In the islands of the Lesser Antilles, for example, butterflies are better able to persist on small islands than are birds or mammals, perhaps because they are able to pass through unfavorable seasons as diapausing eggs and pupae. Similarly, reptiles and amphibians are resistant to extinction, possibly by virtue of their ability to survive for long periods in protected micro environments without having to feed (Ricklefs and Lovette, 1999).

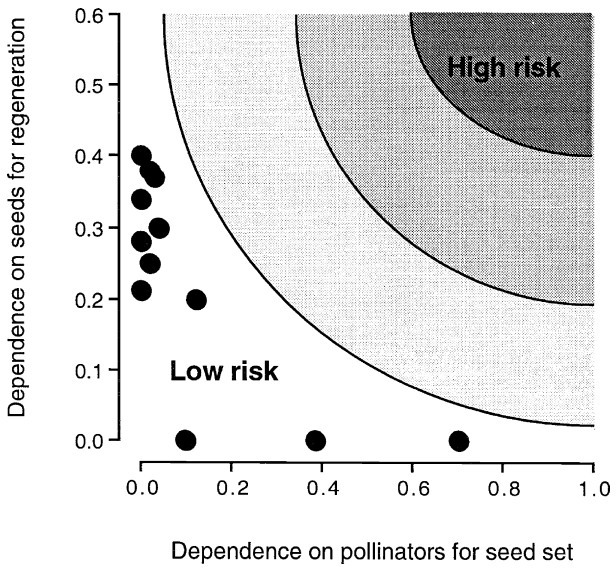


Figure 3 The relationship between an index of dependence on animal pollinators for seed set and dependence on seeds for regeneration for species of spring flowering herbs in a temperate deciduous forest. Any species that was both dependent on pollinators for seed set and dependent on seeds to regenerate would be considered at high risk of extinction due to pollinator failure. (Redrawn with permission from Bond WJ (1995) Assessing the risk of extinction due to pollinator and disperser failure. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 131–147. Oxford: Oxford Univ. Press.)

Geographic Patterns in Extinction

Natural extinction rates vary along two major geographic axes: Extinction rates are lower for marine than terrestrial organisms, and extinction rates may be higher near the equator than at high latitudes. The low extinction rates of marine organisms are presumably due to their large geographic ranges—the oceans are bigger and less subdivided than the continents—and the high dispersal ability of many species that have larval life stages that drift in the plankton. The effects of latitude on extinction rates are less distinct, but where extinction rates have been shown to vary with latitude they tend to be higher in the tropics. Tropical species often have smaller geographic ranges and lower population densities than species at higher latitudes and may thus be more vulnerable to the effects of environmental change (Lawton, 1995). Also, some particularly species-rich tropical environments, such as the low-sediment, low-nutrient, shallow-water platforms that support reef and related communities, are easily disrupted by changes in climate and sea level (Jablonski, 1995).

Natural Extinction and Trends in Biodiversity

The diversity of any lineage of organisms through time reflects the balance of the rates of loss of species by extinction and gains by speciation. Because extinction rates vary among lineages, extinction is an important factor shaping patterns of biodiversity. Much of the explanation for the very high diversity of insects is that, owing presumably to traits such as small size, high abundance, and high mobility, they have low extinction rates. Essentially no insect families have gone extinct during the past 50 million years, and some living genera and species with good fossil records are of Miocene age or older (Labandeira and Sepkoski, 1993). Species diversity of insects has therefore accumulated over long periods of time. Mammals, on the other hand, appear to evolve and speciate rapidly, but they are a minority taxon because their extinction rates are high.

Insects and mammals provide an extreme example of contrasts in the balance of extinction and speciation rates, but in general extinction and speciation rates are positively correlated among different lineages (**Figure 4**). Such a relationship could arise if total species number in a lineage is held at equilibrium by competition between species so that a new species can only establish by occupying a niche left vacant by a species that has gone extinct. However, the relationship was first shown for groups of animals undergoing rapid increase in diversity (Stanley, 1979). There are three possible causes of the correlation between rates of speciation and extinction:

1. The effects of dispersal and gene flow: Dispersal rates should be correlated with resistance to extinction, but the gene flow resulting from high dispersal should also prevent genetic divergence of populations, thus impeding allopatric speciation. This explanation may apply to marine gastropods during the Late Cretaceous, when lineages with planktonic larvae and large ranges not only were less likely to go extinct but also were less likely to produce new species than those with direct development and restricted dispersal (Hansen, 1983).
2. The effects of niche breadth: Ecologically specialized species are likely to be vulnerable to environmental change, but they also tend to have patchy distributions, and this spatial subdivision should promote divergent evolution of local populations and the generation of many species. Generalists may be resistant to extinction, but their lack of spatial population structure also impedes divergent evolution among local populations.
3. The effects of rarity: Species with small populations are prone to extinction. Somewhat controversially, it is believed that speciation may be enhanced by small or fluctuating population sizes (Futuyma, 1998).

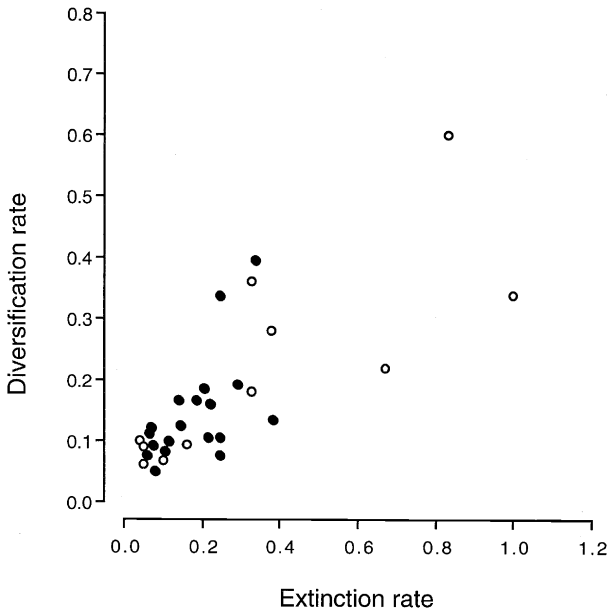


Figure 4 The relationship between extinction rate and diversification rate for lineages of animals (○) and plants (●). Because diversification rate equals speciation rate minus extinction rate, the positive relationship shows that rates of extinction and speciation must be positively correlated (if not, the relationship on the graph would be negative). Extinction rates are measured per million species years and calculated as the reciprocal of species durations.

Because extinction is selective for particular characteristics of species, it can act as a filter through which certain types of species are more likely to pass. Selective extinction therefore shapes the composition of ecological assemblages of species. This filtering effect can be seen most clearly in mass extinctions, which typically leave in their wake "survival faunas" dominated by small numbers of wide-ranging and abundant generalist species of a few simple ecological types. The examples shown in **Figures 2** and **3** illustrate some more subtle effects of the continuous operation of extinction filters, shaping respectively the patterns of distribution and abundance in living marsupials and the combinations of demographic characteristics of living plants.

In some cases, directional trends in evolution and the development of communities can be opposed or curtailed by selective extinction. This effect can produce taxon cycles, which occur if derived species tend to evolve characteristics that make them vulnerable to extinction. For example, species that colonize islands are often generalists that are abundant and competitively dominant and quickly become widespread. Their descendants, however, tend to become more specialized and less mobile, occupy fewer habitats, lose their competitive ability, and sustain smaller and more subdivided populations. These trends eventually lead to extinction and the repetition of the cycle as sets of derived species are replaced by newly invading generalists (Ricklefs and Miller, 1999). Taxon cycles are most visible on island chains, where ecological communities are relatively simple and the patterns of distribution of related species on different islands provide clues to the direction of evolution. However, similar processes probably operate over larger areas on continents and in the oceans.

What can the study of natural extinctions teach us about the coming wave of human-caused extinctions? There are at least two messages. First, species durations among different taxa in the fossil record are positively related to rates of endangerment of living species in the same taxa (McKinney, 1997). This suggests that the characteristics that have made species vulnerable to extinction under natural conditions also make them sensitive to human impacts on the environment, and in the absence of better information these characteristics could be used to identify species likely to be at risk in the near future. Second, the selectivity of extinction means that extinction rates will be higher in certain ecological types and taxonomic groups than in others, and if environmental pressures continue for a sufficient amount of time some groups may be removed entirely. The result will be a much greater impoverishment of biodiversity than would be the case if extinctions were randomly distributed among species.

See also: Extinction, Causes of. Extinction in the Fossil Record. Mammals (Late Quaternary), Extinctions of. Mammals (Pre-Quaternary), Extinctions of. Mass Extinctions, Concept of. Population Density. Species Interactions

References

Bennett KD (1997) *Evolution and Ecology: The Pace of Life*. Cambridge, UK: Cambridge Univ. Press.

Bond WJ (1995) Assessing the risk of extinction due to pollinator and disperser failure. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 131–147. Oxford: Oxford Univ. Press.

Brown JH (1995) *Macroecology*. Chicago: Univ. of Chicago Press.

Coope RG (1995) Insect faunas in ice age environments: Why so little extinction? In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 55–74. Oxford: Oxford Univ. Press.

Cotgreave P (1993) The relationship between body size and population abundance in animals. *Trends Ecol. Evol.* 8: 244–248.

Erwin DH (1998) The end and the beginning: Recoveries from mass extinctions. *Trends Ecol. Evol.* 13: 344–349.

Foufopoulos J and Ives AR (1999) Reptile extinctions on land-bridge islands: Life-history attributes and vulnerability to extinction. *Am. Nat.* 153: 1–25.

Futuyma DJ (1998) *Evolutionary Biology*. Sunderland, MA: Sinauer.

Hansen TA (1983) Modes of larval development and rates of speciation in early Tertiary neogastropods. *Science* 220: 501–502.

Hanski I (1999) *Metapopulation Ecology*. Oxford: Oxford Univ. Press.

Jablonski D (1991) Extinction: A paleontological perspective. *Science* 253: 754–757.

Jablonski D (1995) Extinction in the fossil record. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 25–44. Oxford: Oxford Univ. Press.

Johnson CN (1998) Species extinction and the relationship between distribution and abundance. *Nature* 394: 272–274.

Labandeira CC and Sepkoski JJ (1993) Insect diversity in the fossil record. *Science* 261: 310–315.

Lawton JH (1995) Population dynamic principles. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 147–163. Oxford: Oxford Univ. Press.

May RM, Lawton JH, and Stork NE (1995) Assessing extinction rates. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 1–24. Oxford: Oxford Univ. Press.

Maynard Smith J (1978) *The Evolution of Sex*. Cambridge, UK: Cambridge Univ. Press.

McKinney ML (1997) Extinction vulnerability and selectivity: Combining ecological and paleontological views. *Annu. Rev. Ecol. Syst.* 28: 495–516.

Niklas KJ, Tiffney BH, and Knoll AH (1983) Patterns in vascular land plant diversification. *Nature* 303: 614–616.

Ricklefs RE and Lovette IJ (1999) The roles of island area *per se* and habitat diversity in the species–area relationships of four Lesser Antillean faunal groups. *J. Anim. Ecol.* 68: 1142–1160.

Ricklefs RE and Miller GL (1999) *Ecology*, 4th ed. New York: Freeman.

Stanley SM (1979) *Macroevolution: Patterns and Process*. San Francisco: Freeman.

Thompson JN (1994) *The Coevolutionary Process*. Chicago: Univ. of Chicago Press.

Vrba ES (1987) Ecology in relation to speciation rates: Some case histories of Miocene–Recent mammal clades. *Evol. Ecol.* 1: 283–300.

Webb NR and Thomas JA (1994) Conserving insect habitats in heathland biotopes: A question of scale. In: Edwards PJ, May RM, and Webb NR (eds.) *Large-scale Ecology and Conservation Biology*, pp. 129–152. Oxford: Blackwell.

Natural Reserves and Preserves

Alexander N Glazer, University of California, Berkeley, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Abyssal Applied to the deepest part of the ocean, below about 2000 m.
Benthos (adj. benthic) In freshwater and marine ecosystems, the collection of organisms attached to or resting on the bottom sediments.
Endemism The situation in which a species or other taxonomic group is restricted to a particular geographic region.
Eukaryote An organism whose cells have a distinct nucleus enveloped by a double membrane and other distinctive morphological features.
Hectare A metric measure of surface equal to 10,000 square meters or 2.471 acres (US).
Hypoxia Decreased partial pressure of oxygen.
Nautical mile The unit of length that is used in ocean navigation, equivalent to one minute (1/60°) change in latitude, internationally defined as equal to 1852 m.
Photic zone The layer of water within which organisms are exposed to sunlight.
Prokaryote An organism, in which the cells lack a true nucleus.

Outline

Protected natural areas cover ~12.5% of the Earth's land surface and ~1.2% of the oceans. Collectively, these areas are critical to slowing the loss of biological diversity. Natural biological communities, undisturbed by humans, are also irreplaceable objects for the scientific study of ecological processes. There are many different types of protected areas: national parks, refuges, and sanctuaries, both terrestrial and marine. In recognition of the fact that the purposes for which particular protected areas are managed vary widely, the International Union for Conservation of Nature has classified protected areas into six categories. Examples are given of protected areas that fall into each of these categories. In addition, more detailed descriptions are given for several terrestrial and marine reserves from different biogeographic realms of the Earth to provide a glimpse of their geological diversity and unique biota. It is increasingly recognized that ecosystems, biodiversity, and natural resources are central to sustaining the global economy, societies, and individual wellbeing. Protected natural areas are a key component of assuring the long-term wise management of the Earth. Unfortunately, there are many direct and indirect threats to the future of these areas.

Origins of Protected Areas

"To qualify as conservation, any action or practice must not only prevent or mitigate resource overharvesting or environmental damage, it must also be designed to do so" (Smith and Wishnie, 2000). Preserves, such as sacred groves, established in ancient times, appear to qualify under that definition, because they do protect those areas in their natural state. Royal hunting parks clearly would not.

Preservation of natural biological habitats for their own value originated relatively recently (Adams and Hutton, 2007). The world's first national park, Yellowstone, was established in 1872. It occupies nearly 9000 km^2 mainly in the northwestern corner of Wyoming, and extends into Idaho and Montana. The Yellowstone Act of 1872 reflected the prevailing faith in the inexhaustibility of natural resources. The act prohibited the wanton destruction of fish and game for profit, but it permitted hunting, trapping, and fishing for recreation or to provide food for park residents or visitors. Several other nations established national parks soon after: Australia (Royal National Park, 1879), Canada (Banff National Park, 1885), and New Zealand (Tongariro National Park, 1894). In Africa, and to a lesser extent in the Indian subcontinent, steps toward conservation during the colonial era emphasized large mammals, particularly those that had been intensively hunted. The oldest national park in India, Jim Corbett National Park, a haven for tigers in the Himalayan foothills of Uttar Pradesh, was established in 1936. In contrast, in long-settled parts of Europe, protection was extended to humanized landscapes. Protection has also been given to mangroves and coral reefs. Mangroves are found in the tropics and sub tropics on riverbanks and along coastlines. Mangrove plants include trees, shrubs, ferns, and palms, which have adapted to the largely anaerobic conditions of muddy, aquatic environments by producing stilt roots which project above the water to absorb oxygen. Approximately 900 protected areas include mangroves. Coral reefs, many formed over periods of thousands of years, cover between 300,000 and 600,000 km^2 mainly between the Tropics of Cancer and Capricorn. The Great Barrier Reef Marine Park off the coast of Queensland, Australia, covers 348,000 km^2 of reefs and surrounding waters.

The diverse national parks, reserves, and preserves throughout the world, are now referred to as "protected areas." There are over 100,000 protected areas in the world, representing a total area of 18,764,000 km^2, or over 12% of the world's land surface area. By contrast, as of 2010 only 1.2% of the world's oceans are included in the world's ~5000 Marine Protected Areas. National parks, of which there are currently ~6500 in the world, are among the most popular tourist destinations. In 2009, more than 285 million people visited the national parks in the United States alone.

Table 1 Definitions of the International Union for the Conservation of Nature (IUCN) protected area management categories

Category	*Description*
I	Ia. Strict Nature Reserve: managed mainly for science Area of land or sea possessing some outstanding or representative ecosystems, geological or physiological features and species, available primarily for scientific research and environmental monitoring. Ib. Wilderness Area: managed mainly for wilderness Large area of unmodified or slightly modified land, or sea, retaining its natural character and influence, without permanent or significant habitation, which is protected and managed so as to preserve its natural condition.
II	National Park: managed mainly for ecosystem protection and recreation Natural area of land or sea, designated to (1) protect ecological integrity of one or more ecosystems for present and future generations, (2) exclude exploitation or occupation inimical to the purposes of designation of the area and (3) provide a foundation for spiritual, scientific, educational, recreational and visitor opportunities, all of which must be environmentally and culturally compatible.
III	Natural Monument: managed mainly for conservation of specific natural features Area containing one, or more, specific natural/cultural feature which is of outstanding or unique value because of its inherent rarity, representative or esthetic qualities or cultural significance.
IV	Habitat/Species Management Area: managed mainly for conservation through management intervention Area of land or sea subject to active intervention for management purposes so as to ensure the maintenance of habitats and to meet the requirements of specific species
V	Protected Landscape/Seascape: managed mainly for conservation and recreation Area of land, with coast and sea as appropriate, where the interaction of people and nature over time has produced an area of distinct character with significant esthetic, ecological, or cultural value, and often with high biological diversity. Safeguarding the integrity of this traditional interaction is vital to the protection, maintenance, and evolution of such an area.
VI	Managed Resource Protected Area: managed mainly for the sustainable use of natural ecosystems Area containing predominantly unmodified natural systems, managed to endure long-term protection and maintenance of biological diversity, while providing at the same time a sustainable flow of natural products and services to meet community needs.

Source: Reproduced from Dudley N (ed.) (2008) *Guidelines for Applying Protected Area Management Categories*. Gland, Switzerland: IUCN.

Table 2 Global number and extent of protected areas for 2003 for sites with assigned IUCN Category (excluding noncategorized sites)[a]

Category	*Number of sites*	*% of categorized sites*	*Area covered (km^2)*	*Proportion of total area protected (%)*
Ia	4731	7.0	1,033,888	6.8
Ib	1302	1.9	1,015,512	6.7
II	3881	5.7	4,413,142	29.0
III	19,833	29.1	275,432	1.8
IV	27,641	40.6	3,022,515	19.9
V	6555	9.6	1,056,008	7.0
VI	4123	6.1	4,377,091	28.8

[a]Total number of protected sites in 2003 was 102,102. Of these, 34,036 were noncategorized sites. The noncategorized sites covered an area of 3,569,820 km^2. The total area covered by all 102,102 sites was 18,763,407 km^2.

Source: Reproduced from Tables 1 and 2 of Chape S, Blyth S, Fish L, Fox P, and Spalding M (2003) *2003 United Nations List of Protected Areas*. Gland, Switzerland and Cambridge, UK: IUCN/UNEP-WCMC.

Classification

The International Union for the Conservation of Nature (IUCN) defines a protected area as "a clearly defined geographical space, recognized, dedicated, and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values" (Dudley, 2008). IUCN noted that the purposes for which particular protected areas are managed vary widely and noted specifically (1) scientific research, (2) wilderness protection, (3) preservation of species and genetic diversity, (4) maintenance of environmental services, (5) protection of specific natural and cultural features, (6) tourism and recreation, (7) education, (8) sustainable use of resources from natural ecosystems, and (9) maintenance of cultural and traditional attributes. IUCN classified protected areas into the six categories listed in **Table 1**, reflecting the management purpose. Note that IUCN categories IV–VI encompass areas where limited or extensive use of biological resources is allowed. The following discussion of the characteristics of representative protected areas belonging to each of the categories in **Table 1** clarifies the distinctions. **Table 2**, based on 2003 data, shows the global number of sites within each IUCN category.

Extent of Protected Areas of the World's Major Terrestrial Biomes

The terrestrial extent of protected areas is ~12.5% of the land surface, an area almost equal to that of the entire continent of South America. Summary statistics published in 2003,

Table 3 Extent of protection of the world's major terrestrial biomes

Biome name[a]	*Area (km²)*	*Total number of sites*	*Extent (km²)*	*% Biome protected*
Tropical humid forests	10,513,210	3422	2,450,344	23.31
Subtropical/temperate rainforests/woodlands	3,930,979	6196	665,174	16.92
Temperate needle-leaf forests/woodlands	15,682,817	13,297	1,350,221	8.61
Tropical dry forests/woodlands	17,312,538	5746	2,210,563	12.77
Temperate broad-leaf forests	11,216,659	35,735	856,502	7.64
Evergreen sclerophyllous forests	3,757,144	5334	399,587	10.64
Warm deserts/semi-deserts	24,279,843	2008	2,492,377	10.27
Cold-winter deserts	9,250,252	1235	704,037	7.61
Tundra communities	22,017,390	405	2,606,041	11.84
Tropical grasslands savannahs	4,264,832	318	654,310	15.34
Temperate grasslands	8,976,591	3533	411,839	4.59
Mixed mountains systems	10,633,145	9345	1,735,828	16.32
Mixed island systems	3,252,563	3425	967,129	29.73
Lake systems	517,695	261	7989	1.54
Total	145,605,658	90,260	17,511,641	12.03

[a]Biogeographical classification based on Udvardy MDF (1975) *A Classification of the Biogeographical Provinces of the World* (Occasional Paper no. 18). Gland, Switzerland: International Union for the Conservation of Nature and Natural Resources.

Source: Based on data in Table 4 in Chape S, Blyth S, Fish L, Fox P, and Spalding M (compilers) (2003) *2003 United Nations List of Protected Areas*. Gland, Switzerland and Cambridge, UK: IUCN/UNEP-WCMC.

presented in **Table 3**, are based on a framework that classifies the world into eight biogeographic realms, 203 biogeographic provinces, and 14 biomes (Udvardy, 1975). Protection of temperate grasslands and of lake systems is much lower than that of the other biomes. Recently published results of a draft analysis, based on new, global-level land-cover maps, are in reasonable agreement with those provided in **Table 3**, but show somewhat higher levels of protection for some biomes (see Table 3 in UNEP-WCMC, 2009).

Representative Terrestrial Protected Areas

The protected areas described below exemplify each of the IUCN categories I–VI (see **Table 1**).

Snares Islands/Tini Heke Islands

The subantarctic Snares Islands/Tini Heke Islands, located 209 km southwest of Bluff on South Island of New Zealand, belong to the most diverse and extensive of all subantarctic archipelagos. The islands are made of granite and metamorphic rocks more than 100 million years old. They are classified as category Ia (Strict Nature Reserve) because the vegetation is nearly pristine and the islands are free of introduced mammals. The total area of the Snares Islands Nature Reserve is 328 hectares, covering the islands and their foreshores. The islands are home to millions of breeding seabirds, including the endemic Snares crested penguin, and are strictly protected. Research and monitoring of natural plant and animal communities that does not cause long-term disturbance or damage is permitted. There is no access for tourists.

North East Svalbard Nature Reserve

North East Svalbard Nature Reserve, owned by the Norwegian government, exemplifies category 1b (Wilderness Area). The reserve includes the islands of Spitzbergen, Nordaustlaudet, Edgeøy, Barentsøy, and Prins Karls Forland, and the surrounding territorial waters, a total of 1.9 million hectares, in the northeast of the Svalbard archipelago. The archipelago extends approximately from 74° to 81° N and from 10° to 35° E, an Arctic area about one and a half times the size of Switzerland. This wilderness is covered by high arctic tundra vegetation, with many species near their northernmost distribution in Europe. The large mammals include polar bear, reindeer, and walrus. Access is limited to nonintrusive scientific research and recreation.

Canaima National Park

Canaima National Park (category II, National Park), owned by the Venezuelan government, covers 3 million hectares centered on the Guayana Shield in Bolivar State, south of the Orinoco River. Canaima includes the uplands of the Grand Sabana and the eastern tepuis of the Roraima Range. The tepuis are flat-topped mountains with almost vertical slopes. The world's highest waterfall, Angel Falls, with a fall of 102 m, descends from one of these mountains. The summits of the tepuis have rich flora and fauna with high endemism and form a unique archipelago of isolated but related ecosystems. A small population of traditional occupants of the area, the Pémon, continues swidden (slash-and-burn) agriculture, hunting, and gathering. The park preserves representatives of the geological, biological, and cultural features specific to the Guayana Shield. Hunting and collecting of wildlife is forbidden. Recreational activities, research, and education are encouraged.

Jenolan Karst Conservation Reserve

Jenolan Karst Conservation Reserve (category III, Natural Monument) covers an area of 2694 hectares on the eastern margin of the Oberon Highland Plateau (33°49′14″ S, 150°1′17.2″ E) in New South Wales, Australia. Much of the

area was put aside in 1866 to "protect a source of delight and instruction to succeeding generations and excite admiration of tourists from all parts of the world." This was one of the earliest implementations of conservation for the protection of noncommodity values. The reserve offers outstanding examples of karst landforms and natural phenomena. "Karst" refers to the processes, landforms, and landscapes displaying solutional weathering along surface and subsurface pathways. Karst is usually associated with porous rocks of high solubility, such as limestone or gypsum. An estimated 7–10% of the Earth's ice-free land surface comprises karst terrain. The Jenolan Karst Conservation Reserve terrain was formed by selective dissolution of limestone by acidic natural waters, leading to the formation of a karst landscape of sinkholes, a network of connected cave passages and speleotherms. At over 340 million years old, this cave complex is the world's oldest known and dated open cave system (Osborne, 2010). In addition to its geologic features, the reserve protects rare and endangered flora species and faunal communities totaling over 360 species, including 147 species of troglobitic fauna. Over 250,000 people visit this reserve annually.

Selous Game Reserve

Selous Game Reserve (category IV, Habitat/Species Management Area) is situated in southern Tanzania (7°20′–10°30′ S, 36°00′–38°40′ E). At 5 million hectares, it is one of the largest protected areas in Africa, with altitudes ranging from 100 to 1200 m. Prior to the 1930s numerous villages were in the area currently within the reserve. Periodically, sleeping sickness epidemics would sweep through the area. Control of marauding elephants, which consistently consumed a quarter to a third of the food crop production, was a constant problem. During the period of British colonial rule in Tanzania, between 1919–1961, the authorities would order evacuation of areas affected by sleeping sickness and permanent relocation of the inhabitants. Between 1933 and 1942, some 30,000 inhabitants of an area currently within the Selous Game Reserve either moved voluntarily or were relocated, and the elephants then largely driven into the expanded reserve area (Neumann, 2001). Prior to 1930, the land currently within the Selous reserve included a mosaic of human occupation, much like Dartmoor area today (see below). The current, seemingly pristine wilderness is a relatively recent product of the actions described above. The Selous Game Reserve protects the world's largest area of the deciduous miombo woodland, where the trees grow to heights of 15–20 m, rising over a broadleaf shrub understory with grassland underneath. This woodland requires a particular burning regime for its preservation. The Reserve has a higher density and species diversity than any other miombo woodland area. Selous is famous for its elephant, hippopotamus, and black rhinoceros populations, although only 100–400 of the rhinoceros remain. Sizable populations of other large animals include buffalo, Nyasaland gnu, brindled gnu, hartebeest, Greater Kudu, sable antelope, eland, reedbuck, bushbuck, waterbuck, warthog, zebra, giraffe, and wildebeest. Predators include lion, leopard, spotted hyena, hunting dog, and a small population of cheetah. There are more than 350 species of birds and reptiles, including crocodiles. Ecological monitoring and research provide information for the quotas for regulated sport hunting, the main use of the reserve.

Dartmoor National Park

Dartmoor National Park (category V; Protected Landscape/Seascape) was designated one of the national parks of England and Wales in 1951. It was accepted in Britain that beautiful areas of open space, encompassing mosaics of farms, villages, and small towns, were worthy of protection. National park designations were given to such areas, without requirement that they be pristine. Management of the Park aims to ensure preservation of the natural habitat, the cultural heritage of the area, and continuation of traditional agricultural uses. Dartmoor National Park is a moorland landscape with tors of exposed granite and wooded valleys. It is a rich habitat for wildlife and includes lowland farming areas of meadow, pasture, and woodland, deep valleys, and upland moors. There is a wealth of archeological remains. Approximately 31,000 people live within the 91,300-hectare area of the park. About 10 million visitors come to the Park for recreation each year. Management of the area aims to ensure preservation of the natural habitat, the cultural heritage of the area, and continuation of traditional agricultural uses.

Tamshiyacu-Tahuayo Communal Reserve

Tamshiyacu-Tahuayo Communal Reserve (category VI, Managed Resource Protected Area), established in 1991, is an area of 332,500 hectares of Amazonian rain forest in the state of Loreto in northeast Peru and encompasses a rich flora and fauna. The reserve is divided into three land zones: a totally protected core, a subsistence zone, and an area of permanent settlement. More than 6000 people exploit the resources of the subsistence zone and the bordering areas. These people, broadly characterized as *ribereño* peasants, depend on small-scale agriculture, hunting, fishing, and gathering of nontimber plant products. Peruvian Communal Reserves are a category of protected areas "which secures exclusive use of rights over large, natural areas for the benefit of neighboring rural populations" (Newing, 2009, p. 98). In the long-term, an issue with this approach is that rural communities do not retain a clearly defined membership, because they change in composition and size, cultural characteristics, and consensus. These changes were evident in the makeup of communities on the boundary of Tamshiyacu-Tahuayo Communal Reserve within 20 years of its establishment (Newing, 2009).

Extent of Marine Protected Areas

In comparison to terrestrial biodiversity, marine biodiversity receives much less protection. Biodiversity in the ocean is in decline with huge losses of big fish species, widespread damage to coral reefs, and multiple "dead zones" with low oxygen levels lethal to many marine organisms.

The formal designation of marine protected areas was initiated in the 1970s and the number has grown rapidly in

Table 4 Marine protected areas by IUCN category

IUCN category	*Number of sites*
Ia	419
Ib	49
II	666
III	133
IV	1494
V	571
VI	159

Source: Reproduced from Table 5 in Mulongoy KJ and Chape SP (eds.) (2004) *Protected Areas and Biodiversity: An Overview of Key Issues*. Series No. 21. Cambridge, UK: CBD Secretariat, Montreal, Canada and UNEP-WCMC.

recent years. A tabulation of marine protected areas by IUCN category (**Table 4** in Mulongoy and Chape, 2004) showed that they covered ~0.42% of the global ocean area, whereas protected areas cover ~12% of the earth's terrestrial surface. A 2010 estimate indicates that the coverage by marine protected areas has increased to ~1.2%. Much of the protected area is attributable to two very large sites. The Great Barrier Reef Marine Park and the Northwest Hawaiian Islands Coral Reef Ecosystem Reserve together represent over 680,000 km^2.

There are no marine protected areas entirely within international waters. About 63% of the world's ocean area lies beyond 200 nautical miles (nm) of any coastline and consequently beyond claims of national jurisdiction. A few of the largest marine protected areas extend into the 200 nm zone, the others lie within the 3–12 nm territorial waters of a nation's coastline (Mulongoy and Chape, 2009).

Most protected marine areas are in coastal and continental shelf waters, where they protect the intertidal and subtidal terrain and biota. These areas are also contributing significantly to increase in fish stocks in the protected area and beyond, through the migration of larvae and fish out of the protected area.

Examples of Marine Protected Areas

The four protected areas described below provide a glimpse of the amazing wealth of biodiversity within marine reserves.

Great Barrier Reef Marine Park

Great Barrier Reef Marine Park is a maze of about 2900 individual reefs covering 345,400 km^2 on the northeast continental shelf of Australia and includes 600 continental islands. The Park's northern boundary is from 10°41′ S running east up to the eastern edge of the Great Barrier Reef at 145°19′33″ E. The current Great Barrier Reef structure started growing on top of an old reef platform about 9000 years ago when the sea levels gradually rose at the end of the last Ice Age about 20,000 years ago. It is the world's largest reef system. The Park's relatively pristine area has been poetically described as "the largest living organism in the world." More accurately, the Great Barrier Reef encompasses a tight web of interactions among thousands of different living organisms within a very special environment. The area is a treasure trove of biodiversity with over 360 species of hard, reef-building corals, 4000 mollusk species, and over 1500 species of fish. The islands and keys support over 200 bird species, and over 2000 plant species. There are extensive seagrass beds. The reef provides habitat for endangered species, such as the dugong and the loggerhead turtle, and a breeding area for the humpback whale.

Papahānaumokuākea Marine National Monument

Papahānaumokuākea Marine National Monument (Northwestern Hawaiian Islands Marine National Monument) consists of small islands, atolls, submerged banks, and reefs that extend for more than 2000 km northwest of the main Hawaiian Islands. Partial protection of much of the area now within the Monument dates back to 1909, when as a response to the overharvesting of seabirds, and in recognition of the islands' importance as seabird nesting grounds, President Theodore Roosevelt through an Executive Order created the Hawaiian Island Reservation. The current ~362,000 km^2 Monument was created in 2006 by a Presidential Proclamation by President George Bush. The majority of the islets and shoals are currently uninhabited.

Like the main Hawaiian Islands, the Monument is notable for having flora and fauna with some of the highest marine endemism in the world. Remote and protected, the Monument supports significantly more fish biomass than the heavily fished Main Hawaiian Islands. Large apex predators, reef sharks and jacks, account for over half of the fish biomass. The Monument is home to one of the largest and most important assemblages of seabirds in the world, and also provides important habitat for a number of threatened and endangered species, such as the Hawaiian monk seal, one of the most critically endangered marine mammals in the USA with a reported population in 2005 of 1400, and for over 90% of all Hawaiian green sea turtles. The shallow waters of the Monument have 57 stony coral species, of which 17 endemic species account for 37–53% of the relative abundance of stony corals on each reef within the monument. There are 350 different algal species. The remoteness, limited human presence, protected status and near-pristine condition of the North Hawaiian Islands allow detailed studies of the structure and functioning of coral reef ecosystems. The Monument protects a baseline site for the detection and assessment of human impacts on such systems.

Bowie Seamount Marine Protected Area (Sgaan Kinghlas)

A huge column of upwelling lava, known as a "plume," lies at a fixed location under the Pacific Plate, located beneath the present-day position of the Island of Hawaii. As the ocean floor moves over this "hot spot" at about 12 cm a year, the upwelling lava creates a succession of new volcanoes that migrate along with the plate. Bowie Seamount Marine Protected Area (Sgaan Kinghlas), a part of the Kodiak-Bowie seamount chain, is located at 53°17′58.9″ N/135°39′02.5″ W. It encompasses the Bowie, Hodgkins, and Davidson seamounts, and covers ~6130 km^2. This discontinuous chain of extinct volcanoes reflects the movement of the Pacific plate

over the Bowie hotspot. Bowie, formed 0.075–0.7 Mya, is the youngest volcano in this chain. Bowie is an enormous seamount with a base depth of 2800 m and a peak depth of 34 m. Seamounts are usually extinct volcanoes that have been covered by the sea, and Bowie was once a volcanic island (during the last Ice Age). It is estimated that there may be as many as 200,000 seamounts of 100 m or higher in the oceans. Fewer than 300 seamounts have been thoroughly studied.

Because of their physical characteristics seamounts can generate upwelling of nutrients and currents that trap small migrating organisms. The resulting enhanced primary production attracts birds, mammals, and large pelagic fish, and fisheries. Compared to the surrounding deep-sea, the benthic community is generally more diverse and abundant than that in the surrounding deep-sea. Seamounts can have comparable levels of benthic diversity and endemism to continental margins. The Bowie area was selected for special protection because of high primary productivity. A productive algal community is associated with part of the seamount that lies within the photic zone. The diverse oceanic, coastal, and seamount species include several species at risk, including the killer whale (*Orcinus orca*), the Steller sea lion (*Eumetopias jubatus*), and ancient murrelet (*Synthliborhamphus antiquus*).

The Endeavor Hydrothermal Vents Marine Protected Area

Abyssal plains, underwater plains on the deep ocean floor, cover more than half of the Earth's surface. Abyssal food chains are fueled by organic debris that sediments down to the ocean floor. This limited food source supports a sparse population of abyssal fauna. A very different situation is seen in active seafloor-spreading zones where tectonic plates separate and new oceanic crust is extruded onto the seafloor. Here, cold seawater percolates down through the crust, and comes into contact with the underlying molten lava. The particle rich, superheated fluid emerges through the seafloor as buoyant flumes. Such hydrothermal vent fields support faunal biomass productivity, sometimes a thousand-fold higher than that of the surrounding deep sea, and comparable to that of the most productive marine ecosystems. More than 100 vent fields have been documented along the 60,000-km global mid-ocean ridge system.

In 2003, Canada established the first marine protected area for the conservation of deep ocean hydrothermal vents. The 82 km^2 Endeavour Hydrothermal Vents Marine Protected Area (47° N/129° W), 250 km southwest of Vancouver Island, at a depth of 2250 m, is within the Endeavour Segment of the Juan de Fuca ridge system, and encompasses five vent fields. These fields span hydrothermal venting conditions varying in water temperature, salt content, and mineral chimney morphology. The mineral composition at Endeavour vent sites has been extensively analyzed. Black smokers, with vigorous hydrothermal flow and exit temperatures typically in the range of 250–350 °C, emit black "smoke." The emissions are rich in Fe, S, Ca, Cu, and Zn. The "smoke" consists of particulate and dissolved metal-rich sulfide (anhydrite, chalcopyrite, sphalerite, barite) and sulfate minerals, products of high-temperature mixing between vent fluids and oxygen-containing cooler seawater. The volcanic emissions from black smokers have substantial amounts of dissolved CO_2, H_2S, H_2, and small amounts of methane. These dissolved gases and metals sustain microbial communities that, in turn, serve as the base of the food chain for communities of prokaryotes and eukaryotes of five hundred or more different species of organisms, characteristic of these very special sites. The discovery in 1977 of vent fauna, with its unique suite of organisms and extraordinary metabolic pathways, amazed the scientific community, and gave rise to the speculation that life may have originated in such hydrothermal vent environments.

Threats to Marine Protected Areas

Marine protected areas are vulnerable to the severe, pervasive, and increasing threats to biodiversity in the oceans: the major threats include overfishing, bottom trawling, ocean acidification, and "dead zones," where the water, particularly at the sea floor, is anoxic meaning that the dissolved oxygen concentration is very low or zero.

Persistent oxygen minimum zones, exhibiting severe hypoxia (<0.5 ml O_2 l^{-1}), occur naturally in coastal upwelling zones rich in nutrients at intermediate depths of 200–1000 m, sometimes extending to the ocean floor. In those areas the benthic fauna is adapted to dissolved oxygen levels as low as 0.1 ml O_2 l^{-1}. Such zones occur primarily in the eastern Pacific, in the Atlantic west of Africa, the Arabian Sea, and the Bay of Bengal. Since the 1960s, severe hypoxia gradually developed in many coastal areas with nutrient-rich runoff from population centers and watersheds, giving rise to "dead zones," with mass mortality in the benthic community. There is abundant evidence that the dead zones are produced in response to the delivery of large quantities of nitrogen fertilizers, organic waste and so on. For example, the size of the dead zone in the northern Gulf of Mexico is tightly coupled to the nutrient-rich discharge from the Mississippi River. During years with low river flow, the area of hypoxia is <5000 km^2, and increases to $>15{,}000$ km^2 when the river flow is high. Dead zones have now been reported from over 400 widespread locations and affect a total area of more than 245,000 km^2 (Diaz and Rosenberg, 2008).

Deep-sea areas (below 200 m) cover over 60% of the surface of the Earth. A large fraction of this yet-to-be-adequately-explored area lies beyond the zone of exclusive control of individual nations and is open to fishing and a variety of highly ecosystem-damaging practices. Distinctive deep-sea habitats include seamounts, hydrothermal vents, cold seeps, and cold-water corals. The abyssal muds that cover some 75% of the ocean floor below 3000 m, more than half the surface of the Earth, have a rich biota supported by bacterial degradation of the organic matter that settles on the deep-sea floor. The bacteria, adapted to these great depths were first cultured in 1952, and can only be cultivated under a hydrostatic pressure of $\sim$1000 atmospheres and low temperature. In the deep ocean, these bacteria serve as rungs in a food chain, as food for protozoans or filter-feeders. Bottom trawling destroys large swaths of the abyssal habitat.

As noted above, seamounts attract intensive commercial fisheries, whose trawling activities cause severe ecological damage to seamount biota. Other threats, present and

anticipated, include mining manganese nodules, cobalt-rich ferromanganese crusts, polymetallic sulfide deposits, and precious metals at sites of hydrothermal activity, as well as oil and gas extraction with attendant chemical pollution.

Ocean acidification presents a looming enormous threat to ocean biodiversity (Secretariat of the Convention on Biological Diversity, 2009). The oceans take up nearly half of the carbon dioxide (CO_2) produced by the burning of fossil fuels. As shown below, CO_2 reacts with water to form carbonic acid. As shown in the equation below, the protons produced by the dissociation of the carbonic acid acidify the ocean and the carbonate concentration drops as the equilibrium between carbonate and bicarbonate shifts toward bicarbonate.

$$CO_2 + H_2O \leftrightarrow HCO_3^- + H^+ \leftrightarrow CO_3^{2-} + 2H^+$$

To date the decrease in the ocean pH is 0.1, but is projected to fall another 0.3–0.4 units by the year 2100, assuming that combustion of fossil fuels will continue at the present rate. Such acidification will affect the process of calcification, by which corals and molluscs make shells and plates from calcium carbonate. The tropical and subtropical corals would be among the worst affected, with negative implications for the stability and longevity of the reefs that they build and the organisms that depend on them. Cold-water coral reefs are also likely to be adversely affected (Royal Society, 2005). An acidification event of the projected magnitude that occurred 55 million years ago was accompanied by the extinction of a large proportion of benthic calcifiers.

Values of Natural Reserves and Preserves

Over the past 50 years, the number of natural reserves and preserves around the world has grown to over 100,000 by 2003, and concomitantly the protected area increased exponentially to nearly 13% of the land surface and $\sim$1.2% of the oceans. From the outset, preservation of biodiversity has been the dominant objective. About 1.75 million species of plants, animals, insects, and microorganisms have been identified. Estimates of the actual number of species range from 3 to 13 million. The fraction of the total biodiversity on Earth protected within natural reserves and preserves is not known, but many of the ecosystems that occur in the deserts, forests, wetlands, mountains, lakes, rivers, and oceans, are represented in the protected areas.

The irreplaceable values of natural reserves and preserves include wilderness protection, preservation of species and genetic diversity, and scientific research. Other objectives addressed by protected areas are maintenance of environmental services, sustainable use of resources from natural ecosystems, protection of specific natural and cultural features, maintenance of cultural and traditional attributes, education, tourism, and recreation.

Funding of the Management of Protected Areas

Funding of the protection and management of many natural areas is grossly inadequate. As a result, numerous designated natural reserves are *de facto* paper parks. Arguments for enhanced support have in recent years stressed that the economic benefits provided by natural areas far exceed the required investment. Such arguments are bolstered by examples of ways in which protected areas slow down global climate change, such as the following.

- The 39 Canadian national parks sequester 4 billion tons of carbon dioxide estimated to be worth US$39–87 billion in carbon credits.
- In the Brazilian Amazon, protected lands, including state parks and indigenous reserves, are estimated to prevent the deforestation of 670,000 km^2 by 2050, thereby avoiding the release of 8 billion tons of carbon dioxide into the atmosphere.

However, such impressive benefits do not guarantee that even modest necessary funding will materialize. For example, a 2010 report on the financial sustainability of protected areas in 20 Latin American and Caribbean countries with a combined population of 584 million, documented that the available funds of $314 million per year only addressed basic management activities, whereas more adequate management would require an additional $700 million per year. This funding gap is of particular concern because this region has 4111 designated protected areas covering a surface of around 500 million km^2, of which 466 million are terrestrial and 34 million are marine, and contains almost 40% of the Earth's biodiversity (Bovarnick *et al.*, 2010).

Threats to the Future of Protected Natural Areas

From a biological perspective, the majority of individual terrestrial natural reserves and preserves are isolated islands in a sea of exploited lands. Nominally sanctuaries for wildlife and forests, they are subject to hunting, fishing, the wildlife trade, logging, fuel wood and fodder collection, mining, oil and gas extraction. Depending on the particular area, these activities may be legal, or, more frequently, illegal. Activities, such as poaching for the commercial bushmeat trade, may be pursued by park dwellers, or by outsiders, in extreme cases leading to the disappearance of the species for which the protected area was created. Illegal logging, frequently focused on a few valuable species, threatens forests in many protected areas. Accidentally or deliberately introduced alien invasive species, or illegally cultivated commercial plants, are also widespread threats.

Sustaining Biodiversity in a Rapidly Changing World

The design of protected areas is key to ensuring their long-term value. Even if they are secure from threats such as those described above, they also need to be of the right size and shape and contain all the habitats needed to sustain their biota. If the protected area is isolated and the surrounding natural habitat substantially altered, the populations of many of the species within the protected area may in the long-term be too small to survive. A study of an 86-hectare woodland,

the Bogor Botanical Gardens in west Java, offers a classic example. The woodland became isolated when the surrounding forest was destroyed in the mid-1940s. Out of the 62 bird species breeding in the Gardens during 1932–1952, 20 had disappeared by 1980–1985, four were close to extinction, and populations of five more had declined. Although the wooded habitat within the Gardens was protected, it was too small to retain self-sustaining populations of many woodland bird species without the pre-1940s surrounding habitat (Diamond *et al.*, 1987).

The example of the Bogor Botanical Gardens illustrates the vulnerability of small reserves to changes in the use of the surrounding lands. Is this also a concern for very large protected areas?

The world's first national park, Yellowstone National Park covers 8987 km^2 of which 80% is forest. By 2007, ~37% of the surrounding lands had been converted to intensive human uses including grazing. Since its establishment in 1872, the Park has managed to preserve its wildlife population. The American bison was saved from near extinction through the protection extended in the Park in 1902 to the then remnant population, which reached a low of ~25. The population now fluctuates between 2200 and 4900 depending on season, winter weather, and management. Since at least 1917, the Yellowstone bison have been infected with the bacterium *Brucella abortus*, possibly through contact with cows that were grazed within the Park at that time. Montana and Idaho, the states adjoining the Park, now have brucellosis-free status that provides significant economic benefits to a state's cattle industry. Local livestock officials are concerned that bison could transfer brucellosis to domestic livestock, jeopardizing that status. No such cases of brucellosis transmission in an open range setting have been verified scientifically (Kilpatrick *et al.*, 2009). Since the year 2000, to prevent bison that wander out of the Park from coming near grazing cattle, such bison have either been driven back into the Park or culled. More than 3400 bison have been killed since 2000, and the average bison population within Park is capped at its estimated carrying capacity of 3000. Yellowstone National Park has become a cage for the bison. The inflexible boundaries of protected areas do not conform to the natural ranges of the species they aspire to protect and currently can do little to extend the protection beyond these boundaries.

Climate Change – A Global Challenge to the Protection of Biodiversity

The Intergovernmental Panel on Climate Change projects that by 2050 climate change may be the major cause of biodiversity loss. The biodiversity at highest risk is that of Mediterranean-climate ecosystems, deserts, coral reefs, sea-ice, mountains, and boreal forests. In the Amazon, the climate models indicate significant reductions in precipitation that might lead to the transformation of evergreen tropical forests into grasslands.

The case histories presented above show that for many protected areas, large and small, migration of species to neighboring lands may not currently be possible. Climate change will increase the probability and extent of shifts in the distribution of species, habitats, and ecosystems. Confronting this scenario requires implementation of landscape-level planning to provide secure buffers to expand the high-quality habitats within current protected areas wherever feasible, identifying and securing movement corridors to link protected areas, and, more broadly, improving habitat connectivity in anticipation of species spread in response to the impact of climate change on their current locations. The recently established Great Limpopo Transfrontier Park and the Seamount Marine Management Area show the influence of the above recommendations on the expansion of connectivity and size of existing protected areas.

Great Limpopo Transfrontier Park

The Great Limpopo Transfrontier Park in South Africa is a work in progress. Planning started in 2000, but official opening of the Transfrontier Park is still a few years in the future until full removal of fences allows free movement of animals and people along the length of the international borders within the boundaries of the park. The 35,000 km^2 Transfrontier Park will link the Limpopo National Park in Mozabique, Kruger National Park in South Africa, Gonarezhou National Park, Majinji Pan Sanctuary and Malipati Safari Area in Zimbabwe, and two areas between Kruger and Gonarezhou – the Sengwe communal land in Zimbabwe and the Makuleke region in South Africa.

There is much yet to be learned about the biodiversity within the full area to be included in Great Limpopo Transfrontier Park, but the Park encompasses a wide range of habitats. The species diversity and endemism in Kruger National Park has been surveyed. *Inter alia*, Kruger National Park is home to 18 mammals species on the IUCN Red List of Threatened Species, notably, the wild dog (*Lycaon pictus*) with a population of ~300 and a similar size population of the black rhino (*Diecros bicornis*). The increased range opportunities in the Mozambique and Zimbabwe areas of the Park will contribute to the conservation of such threatened species.

In South Africa, transfrontier conservation activities are encouraged by the belief that as boundaries are dissolved local benefits grow as conservation and development spread regionally. The intent is to achieve benefits for the local people by increasing employment opportunities in service sector activities, such as nature-based tourism, with resulting decrease in dependence on natural resources, that is by bridging conservation and development through 'community-based natural resources management'. The effectiveness of this approach is controversial (Dressler and Büscher, 2008).

Seamounts Marine Management Area

The 2390-hectare Cocos Island, located 530 km southwest of the coast of Costa Rica, is the only major oceanic island of the eastern tropical Pacific with wet rainforest and, above 500 m, a cloud forest. The island also possesses the most diverse and extensive coral reef in the east Pacific, and serves as a reservoir for the larvae of marine species. The waters surrounding the island are famous for the exceptionally high abundance of sharks, including reef sharks, whale sharks, scalloped hammerhead sharks, and of other large ocean predators.

Scalloped hammerhead sharks are on the globally endangered list. These sharks are targeted by fishermen for their fins, which are in high demand, particularly on the Chinese market. Protection of the remaining leatherhead turtles is also an urgent concern. Leatherhead turtles are listed as Critically Endangered on the IUCN Red List of Threatened Species. The Costa Rican population has declined by 90% in the past 20 years, due in large measure to illegal harvesting of eggs in nesting sites. The turtles are also accidentally caught in commercial fishing operations.

In 1978, the island was designated a National Park (IUCN category II), and in 2002 the limit of the protected zone for the surrounding marine ecosystems was extended to 22.2 km from shore. The Las Gemelas Seamounts, a complex underwater mountain system, 35 miles south of Cocos Island, harboring an ecosystem that includes deep-sea corals, sponges, crabs, sea urchins, starfishes, sea cucumbers, and deep-sea fish, lay outside of this protected zone. No trawling has taken place at Las Gamelas seamounts, so that they have retained a relatively intact habitat.

In 2011, the Costa Rican government created a 9640-km^2 marine management area around Cocos Island that encompasses Las Gemelas seamounts. Purse seining is excluded in the expanded management area, but unfortunately longlining will still be allowed in much of this protected area.

See also: Corals and Coral Reefs. Ethical Issues in Biodiversity Protection. Marine Protected Areas: Static Boundaries in a Changing World. Market Economy and Biodiversity

References

Adams WM and Hutton J (2007) People parks and poverty: Political ecology and biodiversity conservation. *Conservation and Society* 5: 147–183.

Bovarnick A, Fernandez-Baca J, Galindo J, and Negret H (2010) Financial sustainability of protected areas in Latin America and the Caribbean: Investment policy guidance. United Nations Development Programme (UNDP) and The Nature Conservancy (TNC).

Diamond JM, Bishop KD, and van Balen S (1987) Bird survival in an isolated Javan woodland: Island or mirror? *Conservation Biology* 1: 132–142.

Diaz RJ and Rosenberg R (2008) Spreading dead zones and consequences for marine ecosystems. *Science* 321: 926–928.

Dressler W and Büscher B (2008) Market triumphalism and the CBNRM 'crises' at the South African section of the Great Limpopo Transfrontier Park. *Geoforum* 39: 452–465.

Dudley N (ed.) (2008) *Guidelines for Applying Protected Area Management Categories*. Gland, Switzerland: IUCN.

Kilpatrick AM, Gillin CM, and Daszak P (2009) Wildlife–livestock conflict: The risk of pathogen transmission from bison to cattle outside Yellowstone National Park. *Journal of Applied Ecology* 46: 476–485.

Mulongoy KJ and Chape SP (eds.) (2004) *Protected Areas and Biodiversity: An Overview of Key Issues*. Cambridge, UK: CBD Secretariat, Montreal, Canada and UNEP-WCMC.

Neumann RP (2001) Africa's 'last wilderness': Reordering space for political and economic control in colonial Tanzania. *Africa* 71: 641–665.

Newing H (2009) Unpicking 'community' in community conservation: Implications of changing settlement patterns and individual mobility for the Tamshiyaku Tahuayo communal reserve, Peru. In: Alexiades MM (ed.) *Mobility and Migration in Indigenous Amazonia. Contemporary Ethnological Perspectives*, pp. 97–116. Oxford, UK: Berghan Books.

Osborne RL (2010) Rethinking Eastern Australian Caves. *Geological Society, London, Special Publications* 346: 289–308.

Royal Society (2005) *Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide. Policy Document 12/05 Royal Society, London*. Cardiff: The Clyvedon Press Ltd.

Secretariat of the Convention on Biological Diversity (2009) Scientific synthesis of the impacts of ocean acidification on marine biodiversity. Montreal: Technical Series No. 46.

Smith EA and Wishnie M (2000) Conservation and subsistence in small-scale societies. *Annual Review of Anthropology* 29: 493–524.

Udvardy MDF (1975) *A Classification of the Biogeographical Provinces of the World (Occasional Paper No. 18)*. Gland, Switzerland: International Union for the Conservation of Nature and Natural Resources.

UNEP-WCMC (2009) *Protected Areas and Biodiversity: An Overview of Key Issues. Biodiversity Series No. 21*. UK: UNEP World Conservation Monitoring Centre.

Nature-Based Coastal Defenses: Can Biodiversity Help?

Bregje K van Wesenbeeck, Unit for Marine and Coastal Systems, Deltares, MH, Delft, The Netherlands
John N Griffin, University of Florida, Gainesville, Florida, USA
Mark van Koningsveld, Van Oord Dredging and Marine Contractors, Rotterdam, The Netherlands, and Delft University of Technology, Delft, The Netherlands
Keryn B Gedan, Smithsonian Environmental Research Center, Edgewater, MD, USA
Michael W McCoy and Brian R Silliman, University of Florida, Gainesville, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity Variety of life forms often expressed in total number of species on a specific spatial scale. Also includes genetic diversity.
Disturbance Events that affect ecosystems and their functioning.
Ecosystem functions Processes of energy exchange in ecosystems such as the production of biomass and the fixation of carbon.
Ecosystem services Ecosystem functions that benefit people.
Facilitation One species benefiting another by mediating stresses.
Niche Position of species in an ecosystem in relation to other species and with respect to resource use, habitat requirement, and so on.
Resilience The potential for recovery of an ecosystem after disturbances.

Outline

The rapid degradation of ecosystems jeopardizes the services they provide. Among the most valuable of these services is protection of coastlines by shoreline ecological communities such as coral reefs, mangroves, and salt marshes. Here we synthesize evidence from multiple sources to evaluate whether aspects of biodiversity may influence the degree of coastal protection afforded by particular coastal ecosystems. This synthesis should empower ecologists, conservation scientists, and practitioners to test for and then harness the unrealized but high yield potential of incorporating biodiversity into coastal defense planning.

Introduction

Ecosystems provide humans with a host of valuable services, including food provisioning, water purification, carbon sequestration, climate regulation, and natural protection (Costanza *et al.*, 1997; Daily, 1997; Barbier *et al.*, 2011). Among the most valuable of these services is protection of coastlines by natural shoreline communities such as coral reefs, mangroves, and salt marshes (MEA, 2005; Barbier *et al.*, 2008). By baffling water movement and dampening potentially destructive waves before they hit the terrestrial shoreline, the biotic structures created by organisms that make up coastal ecosystems reduce flooding and coastal erosion (Das and Vincent, 2009; Koch *et al.*, 2009). These valuable coastal ecosystem services are, however, being threatened because coastal ecosystems are suffering from widespread degradation through habitat destruction and species extinctions and the resultant reduction in biodiversity (Lotze *et al.*, 2006; Byrnes *et al.*, 2007; Worm *et al.*, 2007; Waycott *et al.*, 2009). Meanwhile, we are facing an increasing number of severe storms (Munich RE, 2011) and accelerated rates of sea-level rise are predicted (Mousavi *et al.*, 2011). Of all natural disasters, flooding events are second only to earthquakes as causes of fatalities and financial losses (Munich, 2010), and future costs for moving and protecting human shoreline communities are expected to be massive (Parry *et al.*, 2009). Consequently, understanding, conserving, and harnessing the full benefits of natural ecosystems for coastal protection is a high conservation priority.

Although the value of nature-based coastal storm protection is increasingly recognized, our understanding of how to value this service remains largely based on simple models that do not account for physical or ecological complexities inherent in coastal ecosystems. Recent research, however, has begun to refine these models, yielding important implications for ecosystem service management strategies. For example, although the potential for a given coastal ecosystem to protect the coastline was traditionally thought to increase linearly with its size, recent work has demonstrated that the quantity of ecosystem services generated, including coastal protection, often varies nonlinearly with ecosystem size (Barbier *et al.*, 2008; Koch *et al.*, 2009). Appreciation of this nonlinearity can guide ecosystem-based land-use planning strategies that successfully integrate development and conservation goals. Despite these advances, scientists have yet to explicitly consider the possible role of other key system characteristics, such as biodiversity within ecosystems, in determining the quantity of coastal protection they provide.

The last 20 years have witnessed an explosion in research dedicated to understanding links between various aspects of biodiversity and ecosystem functioning (Loreau *et al.*, 2001; Balvanera *et al.*, 2006; Cardinale *et al.*, 2006). To date, one of the clearest messages to emerge from this large body of work is that the identity of species (i.e., those with key functional traits that enhance focal processes) within an ecosystem can have

strong influences on basic ecosystem functions such as resource uptake and productivity. Although the identity of species can be a dominant determinant of ecosystem functioning, biodiversity–ecosystem functioning research has also shown that the richness (i.e., number) of biodiversity components (usually species, but also alleles or genotypes) can boost ecosystem functions and, therefore, may also boost associated services (Balvanera *et al.*, 2006; Cardinale, 2011; Tilman *et al.*, 2006). Although the precise ecological mechanisms underlying these effects are still being investigated, there is growing evidence that *niche complementarity* (ecological difference among species such as phenology, response to disturbance, resource use, etc.) plays a key role (Dimitrakopolous and Schmid, 2004; Kahmen *et al.*, 2005; Cardinale *et al.*, 2011). Theory and growing empirical evidence suggest that such ecological differences among species will become increasingly important over large spatial and temporal scales and where more than one ecological function is considered (Gamfeldt *et al.*, 2008; Griffin *et al.*, 2009a; Isbell *et al.*, 2011). Differences among species in their responses to environmental conditions can help to maintain a steady level of ecosystem functioning in the face of environmental changes as different species act as ecological "insurance" (Yachi and Loreau, 2009; Griffin and Silliman, 2011). In addition to increasing the breadth of ecological niches occupied and providing ecological insurance, a greater richness of species also increases the probability that particular key species (and combinations of species) will be present in a community.

A key but as yet largely unmet challenge is to translate known effects of biodiversity on basic ecosystem functions (e.g., primary productivity and nutrient cycling) at local scales to ecosystem services that are provisioned over larger scales and have a direct value to human society (Díaz *et al.*, 2006). In this chapter, we take the first step toward this goal, asking how aspects of biodiversity may affect nature-based coastal defenses. Specifically, we first provide an overview of the history of coastal protection and the increasing appeal of nature-based strategies; we then give examples of how aspects of biodiversity can be important at multiple scales, how ecosystem diversity across landscapes may be key for harnessing synergies, and even how diversity of trophic levels can be important. We end with a call for large-scale experiments and replicated comparative studies to test how diversity within different levels of ecological organization (species, genetic, and ecosystem) impacts coastal shoreline protection.

A Short History of Coastal-Protection Strategies

Human populations are becoming increasingly concentrated in flood-prone areas (e.g., along coasts and rivers; Cohen *et al.*, 1997), and losses of life and property from floods are expected to increase. Therefore, innovative methods for flood protection that maintain access to freshwater or aquatic food resources, while not hindering economic development, are urgently needed. Incorporating naturally functioning ecosystems as part of coastal land-development strategies has emerged as one of the most economically viable and effective ways to defend against floods while maintaining other ecosystem services (Barbier *et al.*, 2008; Feagin, 2008; Gedan *et al.*, 2011).

Historically, flood management was a decentralized effort charged to individuals or small communities, and solutions were often focused more on adapting to periodic flooding rather than flood prevention (Van Koningsveld, 2008). For example, many people living in seasonally flooded areas in the Netherlands and in Southeast Asia and Africa have traditionally restricted their domiciles to mounds or to floating rafts and often have economies that are intertwined with flooding events. In much of the world, however, economic development and technological advancements have led to increased centralization of flood management. Consequently, traditional solutions have become ever more displaced by the installation of complex flood-defense systems that require large financial investments, large-scale construction, and long-term management and maintenance strategies (Duxbury and Dickinson, 2007), and these often come at the cost of ecosystem integrity.

Early efforts to develop flood-management systems were singly focused on preventing floods without consideration of ecological impacts. In the 1970s, however, increasing environmental awareness and a response to more aggressive flood-management approaches led to legislation ensuring that environmental impacts were considered in flood-management activities. In the 1980s, flood-management decisions continued to take environmental impacts into consideration but weighed potential socioeconomic impacts of floods more heavily than environmental impacts when making flood-control decisions (van Koningsveld *et al.*, 2008). It was not until the mid 1990s that integrated and multidisciplinary design and planning efforts began to emerge as policy makers started to fully appreciate that protecting socioeconomic interests was not necessarily in opposition to protecting ecosystems. Indeed, it was during this time that integrated coastal zone management (ICZM) (launched in 1994), was developed as a way to guarantee safety against flooding while maintaining other functions of the coastal ecosystems. Over the past 15 years, managers have begun to fully embrace this integrative planning process, and innovative solutions to coastal flooding such as 'living shorelines' (Gedan *et al.*, 2011; Scyphers *et al.*, 2011) and 'building with nature' (Borsje *et al.*, 2011) have emerged that simultaneously enhance the protection of coastal populations, economic development, and ecosystem integrity.

Nature-Based Coastal Defense: Where Are We?

The need for sustainable and cost-effective flood and storm mitigation for coastal human communities has renewed interest in coastal protection provided by natural ecosystems. A variety of coastal and riparian ecosystems, including dunes, mangroves, marshes, seagrass beds, and coral and shellfish reefs, are recognized for their provision of coastal protection (UNEP, 2006; MEA, 2005; Barbier *et al.*, 2008; Barbier *et al.*, 2011). These coastal ecosystems form physical structures that attenuate waves, decrease water velocities, block winds and reduce erosion, and therefore function in a similar way as human-made coastal defense structures such as levees and dams (**Table 1**).

Coastal ecosystems that attenuate waves diminish coastal flooding and increase shoreline stability (Gedan *et al.*, 2011;

Table 1 Listing ecosystems, their coastal defense properties, and additional services

Ecosystem	*Coastal defense properties*	*Additional services*
Coral reef	Wave dampening	Nature, tourism, food provisioning (fish)
Shellfish reef	Wave dampening	Nature, food provisioning (fish and shellfish)
Salt marsh	Wave dampening	Nature, nursery for seafood species
Riverine wetland	Flood defense, water retention	Nature, nursery
Dunes and sandy beaches	Flood defense	Tourism and recreation, water purification
Seagrasses	Mediation of currents and waves	Nature, nursery for fish, clam habitat
Mangroves	Reducing wave impact and current velocities	Nature, provisioning of food and firewood, tourism

Koch *et al.*, 2009; Borsje *et al.*, 2011). The physical structure of coral and shellfish reefs, salt marshes, and mangroves create friction (i.e., drag), which slows water velocities (Christiansen *et al.*, 2000; Mazda *et al.*, 1997) and attenuates waves (*marsh* Coops *et al.*, 1996; Moller *et al.*, 1999; Moller, 2006; *mangrove* Mazda *et al.*, 1997 protection; *shellfish reefs* Borsje *et al.*, 2011; *coral reefs*, Hardy and Young, 1996, Lugo-Fernandez *et al.*, 1998). Mangrove and marsh canopies and dunes block wind and reduce fetch and can further reduce wave heights (Campbell *et al.*, 2009). The structure created by organisms in these coastal ecosystems also reduces the erosion of bed sediments and helps to stabilize shorelines by reducing water velocities and turbulence (*marsh*, Neumeier and Ciavola, 2004; *mangrove*, Mazda *et al.*, 1997 drag; *oyster reefs* Piazza *et al.*, 2005). In addition, biotic structures such as roots, rhizomes, and shell debris provide a physical barrier between sediments and erosive water flow (*mangrove*, Mazda *et al.*, 2006).

A significant feature of all of these coastal ecosystems is that they are biogenic, or built over time by organisms. Modification of the coastal environment by shellfish, corals, sea grasses, dune plants, salt-marsh grasses, and mangroves bears indirectly on sediments and bathymetry in ways that reduce wave heights and shoreline erosion beyond the direct effects of drag reduction. Because wave heights are largely determined by water depth (le Hir *et al.*, 2000), mangrove and marsh plants that affect the elevation of intertidal surfaces by building peat platforms have strong indirect effects on wave heights (Fagherazzi *et al.*, 2006). Biogenic structures can accumulate over decades and centuries, accreting meters in elevation over the underlying substrate (*marsh*, Marani *et al.*, 2007; Kirwan and Murray, 2007). The sediment deposition facilitated by plants and animals can also affect the stability of the coastal system. In salt marshes, the decay of a dense root mass enriches soil organic content, which is associated with slower erosion than more mineral-rich wetland soils (Feagin *et al.*, 2009). Particles bound by mussels and oysters as pseudofaeces increase sediment cohesion and reduce erosion of intertidal flats (van Leeuwen *et al.*, 2010; Borsje *et al.*, 2011 and citations within).

Although, in general, the evidence that biogenic coastal habitats protect shorelines is overwhelming, we still lack important information for predicting when and where these services are effective and can be used for coastal engineering. One major knowledge gap is the scale at which shoreline protection services of different ecosystems are applicable. For example, changes in the shoreline across very large scales along the Thai coastline can be explained by the presence of mangrove forests (Thampanya *et al.*, 2006), whereas oyster reef restorations (Scypher *et al.*, 2011) and salt-marsh vegetation (Feagin *et al.*, 2009) have failed to have a detectable effect on shoreline change at small spatial scales. There are several reasons these discrepancies occur and may impede the generalization and scaling up of shoreline-protection services.

Shoreline-protection services exhibit stark nonlinearities over space and time, due to variation in biotic structures over seasons or years (**Figure 1**, *multiple habitats*, Koch *et al.*, 2009; *coral reefs*, Wilkinson *et al.*, 1999; Sheppard *et al.*, 2005) and due to physical effects that are strongest at the habitat fringe (Kennedy and Mayer, 2002; Möller and Spencer, 2002; Gedan *et al.*, 2011). These nonlinearities have been considered more thoroughly for some coastal ecosystems – for example, seagrass beds (Koch *et al.*, 2009), mangrove forests (Barbier *et al.*, 2008), and marshes (Kennedy and Mayer, 2002; Möller and Spencer, 2002) – than for others. Still, no general rules of thumb have been derived that determine shifts of energy reduction by vegetation. This is mainly due to the lack of controlled measurements on realistic scales (Fonseca and Cahalan, 1992; Mendez and Losada, 2004).

Depending on the biological requirements of the dominant organisms, coastal ecosystems are constrained in their development and self-sustainability by environmental conditions such as elevation, currents, and wave exposure (Piazza *et al.*, 2005; Borsje *et al.*, 2011; Gedan *et al.*, 2011), and better definitions of these constraints are necessary to apply natural shoreline-protection services more broadly. In addition, climatic and tidal-stage conditions and storm track can have a large impact on the effectiveness of coastal ecosystems in shoreline protection (Resio and Westerink, 2008), making it difficult to predict the effectiveness of ecosystems for a given storm or flooding event until after the event occurs. Moreover, there are few scientific studies of wave attenuation or coastal protection during catastrophic wave events such as tsunamis, making it even more difficult to establish, with any level of certainty, the effectiveness of coastal ecosystems in protection from these events, which impedes inclusion of ecosystems in coastal-protection schemes. Nevertheless, damage analyses do often show protection of communities situated adjacent to coastal ecosystems from even these catastrophic events (Laso Bayas *et al.*, 2011; Danielson *et al.*, 2005; Das and Vincent, 2009).

Finally, as we present in this chapter, the roles of species identity and species composition in coastal-protection services have been insufficiently investigated. However, through mechanisms identified in studies of biodiversity and ecosystem function and of natural coastal protection, it is likely that mixtures of species and habitats can increase the effectiveness of natural shoreline-protection services.

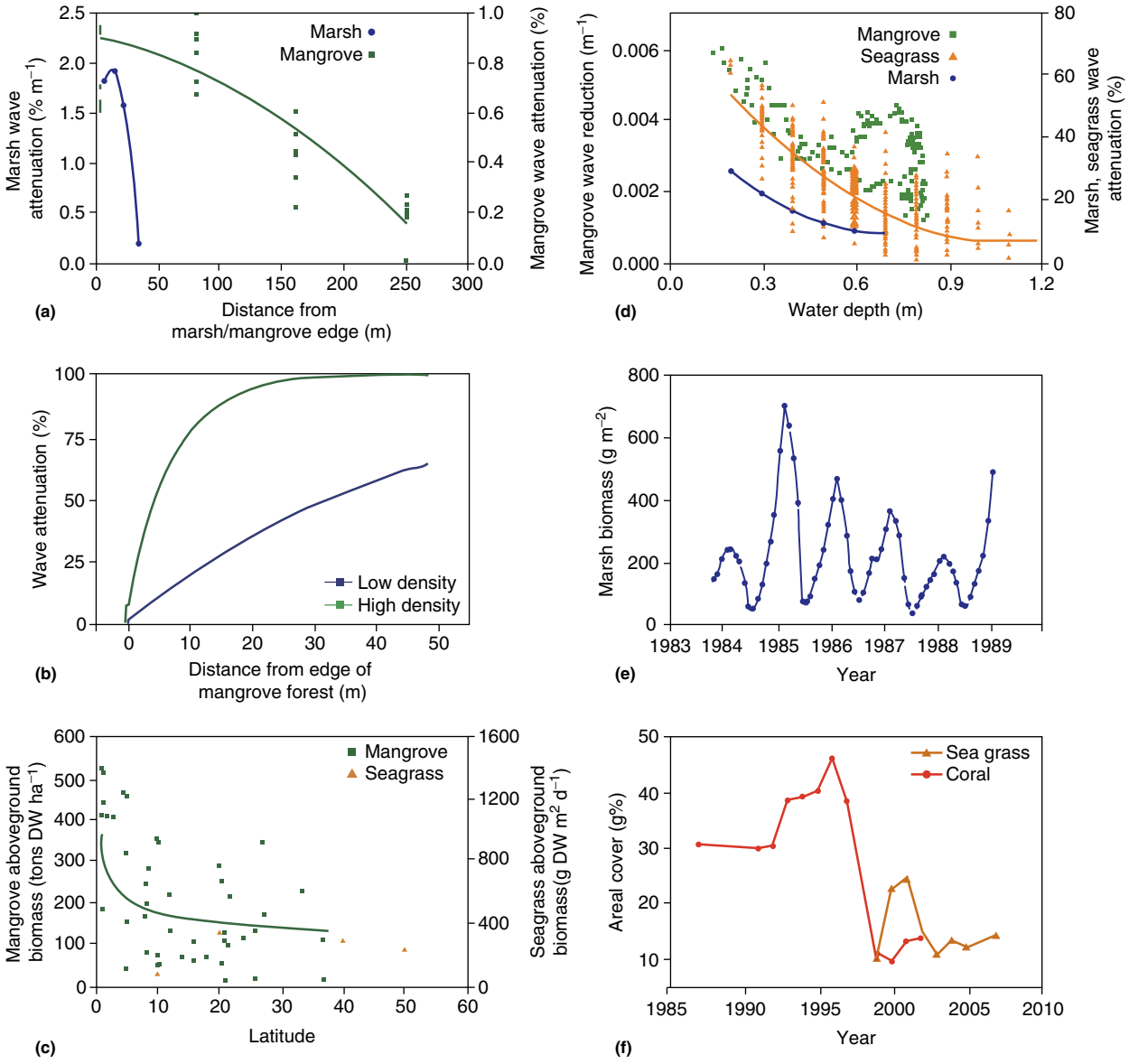


Figure 1 Examples of nonlinearities in wave attenuation for a variety of habitats. Direct measurements of wave attenuation only exist for the smallest spatial (a; Massel *et al.*, 1999; Möller, 2006) and temporal (d; Mazda *et al.*, 2006; Möller *et al.*, 2006; Chen *et al.*, 2007) scales. Wave attenuation for different mangrove densities (b) has been modeled (Massel *et al.*, 1999). Latitudinal wave attenuation (c) is estimated, based on the aboveground biomass of mangroves (Twilley *et al.*, 1992) and sea grasses (Duarte and Chiscano, 1991) – that is, obstructions to water flow. Wave attenuation over different seasons is also assumed to change with marsh aboveground biomass (e; Morris and Haskin, 1990). Interannual variability in coral (modified from McClanahan *et al.*, 2005) and seagrass (Merkel, 2008) area cover (f) is used to estimate long-term trends in wave attenuation. Reprinted from Koch EW, Barbier EB, Silliman BR, *et al.* (2009) Nonlinearity in ecosystem services: Temporal and spatial variability in coastal protection. *Frontiers in Ecology and the Environment* 7: 29–37.

How Biodiversity Can and Does Matter on Different Scales

A range of ecosystem types contribute to coastal protection (e.g., oyster reefs, coral reefs, salt marshes), and although all these systems damp incoming physical stress from water and wind, the type of ecosystem present is likely to be an important determinant of the degree of coastal protection provided. For example, existing studies provisionally suggest that, although shellfish reefs and mangroves both contribute to wave dampening and shoreline stabilization, mangroves are substantially more effective in both (*shellfish reefs*; Folkard and Gascoigne, 2009; van Leeuwen *et al.*, 2010, *mangroves*; Quartel *et al.*, 2007; Ellison, 2000). Within ecosystem types, the efficacy of shoreline protection can also vary with the identity of species. For example, in controlled laboratory settings, differences in effectiveness of wave dampening and sediment trapping have been observed for single-species beds within both marsh plants and sea grasses (Fonseca and Calahan, 1992; Coops *et al.*, 1996; Peralta *et al.*, 2008). These observed differences

between species suggest that the identity of the dominant species may be an important determinant of the degree of coastal protection provisioned by a specific coastal ecosystem, which is consistent with a large body of experimental work examining the effects of biodiversity on ecosystem functions (reviewed by Hooper *et al.*, 2005). The identity of dominant species can change across time and space within any single coastal ecosystem type and should thus be considered in assessments and predictions of coastal-protection potential.

Drawing on theory and past experiments linking biodiversity and ecosystem function (e.g., Balvanera *et al.*, 2006), we predict that the ecosystem services of wave dampening and sediment trapping should be enhanced by species diversity (the number of species present), not simply vary as a function of species identity (i.e., a sampling effect). More precisely, although the elevation of these services with increased diversity could stem directly from the greater probability of including one or more exceptional species, it may also result from emergent effects of diversity per se, through the inclusion of species that inhabit complimentary niches. Although this hypothesis has yet to be explicitly tested for shoreline-protection services, there is evidence that combinations of different species on local scales can have greater impacts on hydrodynamics and sedimentation rates than any species in isolation (e.g., via differences in stoichiometry; Bouma *et al.*, 2005, 2010). Species also clearly show complementarity across space, especially in coastal ecosystems that are often characterized by steep physical gradients and marked zonation of species. It remains unclear how small-scale studies might scale up to explain patterns at larger spatial scales; however, we know from some systems that small-scale interactions can influence landscape patterns over larger scales (e.g., salt marshes; Temmerman *et al.*, 2007; Kirwan and Murray, 2007; van Wesenbeeck *et al.*, 2008).

Niche complementarity is likely to be most important, however, when multiple ecosystem functions or services are considered simultaneously (Gamfeldt *et al.*, 2008; Isbell *et al.*, 2011). This theory may well apply to shoreline protection, as this service involves multiple ecological effects – wave dampening, soil trapping, reduction of water velocity, accretion, and soil effects all contribute to shoreline protection – and even if certain species contribute more to one aspect of coastal protection (e.g., wave dampening by a rigid-bodied species), others contribute to another (e.g., soil stabilization by a soft-bodied species) and no single species (or group of similar species) is sufficient to sustain the multipronged service of coastal protection. For example, Ewel *et al.* (1998) speculate that fringe and riverine mangroves may be more effective at wave attenuation than basin mangrove forests, but basin mangrove forests may be more important for flood storage capacity.

At the landscape level, different kinds of coastal ecosystems have different functional roles in coastal defense, and so cross-ecosystem linkages may be very important for shoreline protection. For example, coral reefs have large implications for wave dampening by functioning as submerged breakwaters (Möller *et al.*, 1999; Kench and Brander, 2006), and mangroves are efficient sediment trappers, thereby stabilizing the shoreline (Ellison, 2000). Often these systems are found co-occurring along coastlines, as mangroves and marshes are generally restricted to low wave-energy environments (Knutson *et al.*, 1981; Ewel *et al.*, 1998), but the presence or placement of a reef at the seaward margin may allow the wetland to persist in a more wave-exposed environment (**Figure 2**). It seems intuitive that waves are dampened more if they meet a shellfish reef and a mangrove forest than if they meet a mudflat and an eelgrass meadow due to the rigidity and the relative height of their biotic structures. However, surprisingly few data support this idea (see discussion in Barbier *et al.*, 2008). Thus, we are limited in our ability to predict if and to what extent different combinations of coastal communities can enhance or hinder coastal shoreline-protection services. Indeed, much additional research is needed in this area.

The extent to which these different coastal ecosystem types can influence coastal defense properties is dependent on having natural, functioning, and stable food webs within each. For example, salt marshes that have experienced significant grazing have distinct features (Bakker, 1981, 1985) that change their effects on wave attenuation. The lower-standing biomass of grazed marshes (**Figure 3**) can reduce wave-dampening



Figure 2 New restorations are pairing multiple habitat types for greater effectiveness in shoreline protection (B. Silliman). Reprinted from Gedan KB, Kirwan ML, Wolanski E, Barbier EB, and Silliman BR (2011) The present and future role of coastal wetland vegetation in protecting shorelines: Answering recent challenges to the paradigm. *Climatic Change* 106: 7–29, with permission from Springer.



Figure 3 Grazed marshes (front) differ considerably from ungrazed marsh systems (back). Evidence suggests that intensive grazing and trampling may have a negative effect on coastal defense properties. Photo published with permission from Martin Stock.

properties of marshes and intensive grazing was shown to reduce the strength of marsh soils (Zhang, 1993), which might make marshes more vulnerable to wave erosion. Further, food-web complexity and structure are also important because they influence ecosystem stability. For example, large diebacks of salt marshes in the southeast US have been linked to declines of a keystone predator, the blue crab (Silliman and Bertness, 2002). Reduction in blue crab numbers releases fungal-farming snails from predation allowing snail numbers to increase to levels high enough to graze down marsh grasses and transform the tidal marsh into mudflats (Silliman and Bertness, 2002). The ecosystem service of coastal protection offered by a healthy salt marsh is considerably greater than that offered by an intertidal mudflat. This example shows that food-web interactions are often complex and that understanding these interactions is crucial for successful implementation of nature-based coastal defenses. Further, it highlights the value of maintaining biotic diversity and food-web structure in ecosystems to maximize their effectiveness in providing coastal protection.

An advantage of natural shoreline protection over human-made coastal defenses is that natural shoreline protection is self-sustaining. Natural shorelines require less maintenance and can naturally adapt to changing environmental conditions such as rising sea levels over time. Therefore, applications of nature-based shoreline-protection strategies should also attempt to maximize ecosystem resilience. Maintaining greater species diversity is expected to maximize ecosystem output (Cardinale *et al.*, 2006) and increase ecosystem stability over time (Cottingham *et al.*, 2001; Griffin *et al.*, 2009b), and more diverse ecosystems are therefore likely to improve the provision of shoreline-protection services over the long term.

Shoreline protection by natural ecosystems represents an elegant compromise for achieving the goals of maintaining biodiversity and ecosystem integrity while also providing shoreline protection as well as many other benefits to human communities. Still, to improve applicability of nature-based flood defenses, more knowledge needs to be collected on design, maintenance, management, and governance. Preferably, this knowledge should be derived from real-scale applications that are accompanied by a thorough monitoring program. Conserving and promoting biodiversity in coastal ecosystems conserved or created for shoreline protection will help to maximize the myriad ecosystem services they provide, including benefits of fisheries and other natural resources, recreational activities, and water and nutrient regulation.

Conclusions

Awareness is rising that functioning, intact coastal ecosystems offer valuable coastal-protection services. By implementing ecosystem management in planning designs for coastal defenses, we can make use of the unique organisms that have evolved in and are well adapted to shoreline stressors. This requires an integrated planning and design process and a solid understanding of natural processes and ecosystem functioning. Making use of general ecological theory is indispensable for adequate restoration and conservation of ecosystems in general.

The effectiveness of coastal ecosystems for wave dampening, reducing current velocities, slowing erosion, and facilitating biotic diversity depends on the types of coastal communities that are present (e.g., oyster reef, coral reef, mangrove, etc.) and on the spatial configuration of these communities.

Degenerated ecosystems that lack natural species assemblages are not likely to be as robust, stable, and resilient as intact ecosystems and therefore are less likely to provide the valuable ecosystem services that many coastal communities depend on. In general, higher diversity is thought to increase ecosystem stability, enhance resilience to perturbations, and increase productivity. Therefore, managing coastal ecosystems to promote diverse and natural species assemblages will enhance the effectiveness of these systems as protection. However, the exact implications for management and conservation depend on whether species-specific effects and possible positive interactions are well understood; if they are, then managing for particular species in order to maximize the yield of certain services may be more efficient than managing for diversity *per se*.

There is a need for large-scale experiments and replicated comparative studies to test how diversity within different levels of ecological organization (species, genetic and ecosystem) impacts coastal shoreline protection. We cannot advance our understanding of biodiversity–ecosystem service relationships unless proper controls are utilized when assessing impacts of proposed nature-based coastal-protection plans.

Incorporating the construction and maintenance of functional coastal ecosystems as a part of coastal defense strategies will require greater collaboration between ecologists, geologists, and engineers. It will also require focused research efforts aimed at documenting and quantifying the flood-defense properties of ecosystems. Such data are needed to develop predictive models that will allow coastal managers to assess the level of protection provided by specific coastal ecosystem configurations and therefore also estimate the potential socioeconomic impacts of ecosystem loss.

References

Bakker JP (1985) The impact of grazing on plant-communities, plant-populations and soil-conditions on salt marshes. *Vegetation* 62: 391–398.

Bakker JP and Ruyter JC (1981) Effects of 5 years of grazing on a salt-marsh vegetation – a study with sequential mapping. *Vegetation* 44: 81–100.

Balvanera P, Pfisterer AB, Buchmann N, *et al.* (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9: 1146–1156.

Barbier EB, Hacker SD, Kennedy C, Koch EW, Stier AC, and Silliman BR (2011) The value of estuarine and coastal ecosystem services. *Ecological Monographs* 81: 169–193.

Barbier EB, Koch EW, Silliman BR, *et al.* (2008) Coastal ecosystem-based management with nonlinear ecological functions and values. *Science* 319: 321–323.

Borsje BW, van Wesenbeeck BK, Dekker F, *et al.* (2011) How ecological engineering can serve in coastal protection. *Ecological Engineering* 37: 113–122.

Bouma TJ, De Vries MB, Low E, *et al.* (2005) Trade-offs related to ecosystem engineering: A case study on stiffness of emerging macrophytes. *Ecology* 86: 2187–2199.

Bouma TJ, de Vries MB, and Herman PMJ (2010) Comparing ecosystem engineering efficiency of two plant species with contrasting growth strategies. *Ecology* 91(9): 2696–2704.
Byrnes JE, Reynolds PL, and Stachowicz JJ (2007) Invasions and extinctions reshape coastal marine food webs. *PLoS One* 2: e295.
Campbell A, Kapos V, Scharlemann JPW, *et al.* (2009) Review of the literature on the links between biodiversity and climate change: Impacts, adaptation and mitigation. Secretariat of the Convention on Biological Diversity. Montreal. *Technical Series* 42: 124.
Cardinale BJ (2011) Biodiversity improves water quality through niche partitioning. *Nature* 472: 86–91.
Cardinale BJ, Matulich KL, Hooper DU, *et al.* (2011) The functional role of producer diversity in ecosystems. *American Journal of Botany* 98: 572–592.
Cardinale BJ, Srivastava DS, Duffy JE, *et al.* (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443: 989–992.
Chen SN, Sanford LP, Koch EW, Shi F, and North EW (2007) A nearshore model to investigate the effects of seagrass bed geometry on wave attenuation and suspended sediment transport. *Estuaries* 30: 296–310.
Christiansen T, Wiberg PL, and Milligan TG (2000) Flow and sediment transport on a tidal salt marsh surface. *Estuarine and Coastal Shelf Science* 50: 315–331.
Cohen JE, Small C, Mellinger A, Gallup J, and Sachs J (1997) Estimates of coastal populations. *Science* 278: 1209.
Coops H, Geilen N, Verheij HJ, Boeters R, and Van Der Velde G (1996) Interactions between waves, bank erosion and emergent vegetation: An experimental study in a wave tank. *Aquatic Botany* 53: 187–198.
Costanza R, d'Arge R, de Groot R, *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature* 387: 253–260.
Cottingham KL, Brown BL, and Lennon JT (2001) Biodiversity may regulate the temporal variability of ecological systems. *Ecology Letters* 4: 72–85.
Daily GC (ed.) (1997) Nature's Services. *Societal Dependence on Natural Ecosystems*. Washington, DC: Island Press.
Danielson F, Sørensen MK, Olwig MF, *et al.* (2005) The Asian Tsunami: A protective role for coastal vegetation. *Science* 310: 643.
Das S and Vincent JR (2009) Mangroves protected villages and reduced death toll during Indian super cyclone. *Proceedings of the National Academy of Sciences* 106: 7357–7360.
Díaz S, Fargione J, Chapin III FS, and Tilman D (2006) Biodiversity loss threatens human well-being. *PLoS Biology* 4: e277.
Dimitrakopoulos PG and Schmid B (2004) Biodiversity effects increase linearly with biotope space. *Ecology Letters* 7: 574–583.
Duarte CM and Chiscano CL (1999) Seagrass biomass and production: A reassessment. *Aquatic Botany* 65: 159–174.
Duxbury J and Dickinson S (2007) Principles for sustainable governance of the coastal zone: In the context of coastal disasters. *Ecological Economics* 63: 319–330.
Ellison AM (2000) Mangrove restoration: Do we know enough? *Restoration Ecology* 8: 219–229.
Ewel KC, Twilley RR, and Eong Ong J (1998) Different kinds of mangrove forests provide different goods and services. *Global Ecology and Biogeography Letters* 7: 83–94.
Fagherazzi S, Carniello L, D'Alpaos L, and Defina A (2006) Critical bifurcation of shallow microtidal landforms in tidal flats and salt marshes. *Proceedings of the National Academy of Sciences* 103: 8337–8341.
Feagin RA (2008) Vegetation's role in coastal protection. *Science* 320: 176–177.
Feagin RA, Lozada-Bernard SM, Ravens TM, Moller I, Yeager KM, and Baird AH (2009) Does vegetation prevent wave erosion of salt marsh edges. *Proceedings of the National Academy of Sciences of the United States of America* 106: 10109–10113.
Folkard AM and Gascoigne JC (2009) Hydrodynamics of discontinuous mussel beds: Laboratory flume simulations. *Journal of Sea Research* 62: 250–257.
Fonseca MS and Cahalan JA (1992) A preliminary evaluation of wave attenuation by four species of seagrass. *Estuarine, Coastal and Shelf Science* 35: 565–576.
Gamfeldt L, Hillebrand H, and Jonsson PR (2008) Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology* 89: 1223–1231.
Gedan KB, Kirwan ML, Wolanski E, Barbier EB, and Silliman BR (2011) The present and future role of coastal wetland vegetation in protecting shorelines: Answering recent challenges to the paradigm. *Climatic Change* 106: 7–29.
Griffin JN, Jenkins SR, Hawkins SJ, Gamfeldt L, and Thompson RC (2009a) Spatial heterogeneity increases the importance of species richness for an ecosystem process. *Oikos* 118: 1335–1342.
Griffin JN and Silliman BR (2011) Predator diversity stabilizes and strengthens trophic control of a keystone grazer. *Biology Letters* 7: 79–82.
Griffin JN, Symstad A, Emmerson M, *et al.* (2009b) Biodiversity and the stability of ecosystem functioning. In: Naeem S, Bunker D, Hector A, Loreau M, and Perrings C (eds.) *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*. Oxford University Press
Hardy TA and Young IR (1996) Field study of wave attenuation on an offshore coral reef. *Journal of Geophysical Research* 101: 14311–14326.
Hooper DU, Chapin FS, Ewel JJ, *et al.* (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75: 3–35.
Isbell F, Calcagno V, Hector A, *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature* 477: 199–202.
Kahmen A, Perner J, and Buchmann N (2005) Diversity-dependent productivity in semi-natural grasslands following climate perturbations. *Functional Ecology* 19: 594–601.
Kench PS and Brander RW (2006) Response of reef island shorelines to seasonal climate oscillations: South Maalhosmadulu atoll, Maldives. *Journal of Geophysical Research* 111: F01001.http://dx.doi.org/10.1029/2005JF000323.
Kennedy G and Mayer T (2002) Natural and constructed wetlands in Canada: An overview. *Water Quality Research Journal of Canada* 37: 295–325.
Kirwan ML and Murray AB (2007) A coupled geomorphic and ecological model of tidal marsh evolution. *Proceedings of the National Academy of Sciences of the United States of America* 104: 6118–6122.
Knutson PL, Ford JC, Inskeep MG, and Oyler J (1981) National survey of planted salt marshes (vegetative stabilization and wave stress). *Wetlands* 1: 129–157.
Koch EW, Barbier EB, Silliman BR, *et al.* (2009) Non-linearity in ecosystem services: Temporal and spatial variability in coastal protection. *Frontiers in Ecology and the Environment* 7: 29–37.
van Koningsveld M, Mulder JPM, Stive MJF, VanDervalk L, and VanDerWeck AW (2008) Living with sea-level rise and climate change: A case study of the Netherlands. *Journal of Coastal Research* 24(2): 367–379.
Laso Bayas JC, Marohna C, Dercona G, *et al.* (2011) Influence of coastal vegetation on the 2004 tsunami wave impact in west Aceh. *Proceedings of the National Academy of Sciences of the United States of America* 108: 18612–18617.
Le Hir P, Roberts W, Cazaillet O, Christie M, Bassoullet P, and Bacher C (2000) Characterization of intertidal flat hydrodynamics. *Continental Shelf Research* 20: 1433–1459.
Loreau M, Naeem S, Inchausti P, *et al.* (2001) Ecology – biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science* 294: 804–808.
Lotze HK, Lenihan HS, Bourque BJ, *et al.* (2006) Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science* 312: 1806–1809.
Lugo-Fernandez A, Roberts HH, and Wiseman WJ (1998) Tide effects on wave attenuation and wave set-up on a Caribbean coral reef. *Estuarine Coastal and Shelf Science* 47: 385–393.
van Leeuwen B, Augustijn DCM, van Wesenbeeck BK, Hulscher SJMH, and de Vries MB (2010) Modeling the influence of a young mussel bed on fine sediment dynamics on an intertidal flat in the Wadden Sea. *Ecological Engineering* 36: 145–153.
Marani M, D'Alpaos A, Lanzoni S, Carniello L, and Rinaldo A (2007) Biologically-controlled multiple equilibria of tidal landforms and the fate of the Venice lagoon. *Geophysical Research Letters* 34.
Massel SR, Furukawa K, and Brinkman RM (1999) Surface wave propagation in mangrove forests. *Fluid Dynamic Research* 24: 219–249.
Mazda Y, Magi M, Ikeda Y, Kurokawa T, and Asano T (2006) Wave reduction in a mangrove forest dominated by Sonneratia sp. *Wetlands Ecology and Management* 14: 365–378.
Mazda Y, Wolanski E, King B, Sase A, Ohtsuka D, and Magi M (1997) Drag force due to vegetation in mangrove swamps. *Mangroves and Salt Marshes* 1: 193–199.
McClanahan TR, Mwaguni S, and Muthiga NA (2005) Management of the Kenyan coast. *Ocean and Coastal Management* 48: 901–931.
MEA (2005) *Ecosystems and Human Well-being: Synthesis*. Washington, DC: Island Press.
Mendez FJ and Losada IJ (2004) An empirical model to estimate the propagation of random breaking and nonbreaking waves over vegetation fields. *Coastal Engineering* 51: 103–118.
Merkel K (2008) Eelgrass habitat surveys for the Emeryville flats and Clipper Cove, Yerba Buena Island (October 1999–2005, and 2007). In: California Department of Transportation, Sacramento, CA.
Möller I (2006) Quantifying saltmarsh vegetation and its effect on wave height dissipation: Results from a UK East coast saltmarsh. *Estuarine Coastal and Shelf Science* 69: 337–351.

Möller I and Spencer T (2002) Wave dissipation over macro-tidal saltmarshes: Effects of marsh edge typology and vegetation change. *Journal of Coastal Research* 36: 506–521.

Möller I, Spencer T, French JR, Leggett DJ, and Dixon M (1999) Wave transformation over salt marshes: A field and numerical modelling study from North Norfolk, England. *Estuarine, Coastal and Shelf Science* 49: 411–426.

Morris JT and Haskin B (1990) A 5-yr record of aerial primary production and stand characteristics of Spartina alterniflora. *Ecology* 71: 2209–2217.

Mousavi M, Irish J, Frey A, Olivera F, and Edge B (2011) Global warming and hurricanes: The potential impact of hurricane intensification and sea level rise on coastal flooding. *Climatic Change* 104: 575–597.

Munich RE (2010) NatCatSERVICE. http://www.munichre.com/app_pages/www/@res/pdf/media_relations/press_releases/2011/2011_01_03_munich_re_NatCatSERVICE_en.pdf

Munich RE (2011) Half-year natural catastrophe review, USA. http://www.munichre.com/app_pages/www/@res/pdf/media_relations/press_dossiers/hurricane/2011-half-year-natural-catastrophe-review-usa_en.pdf

Neumeier U and Ciavola P (2004) Flow resistance and associated sedimentary processes in a Spartina maritima salt-marsh. *Journal of Coastal Research* 20: 435–447.

Parry M (2009) Closing the loop between mitigation, impacts and adaptation. *Climatic Change* 96: 23–27.

Peralta G, van Duren LA, Morris EP, and Bouma TJ (2008) Consequences of shoot density and stiffness for ecosystem engineering by benthic macrophytes in flow dominated areas: A hydrodynamic flume study. *Marine Ecology – Progress Series* 368: 103–115.

Piazza BP, Banks PD, and La Peyre MK (2005) The potential for created oyster shell reefs as a sustainable shoreline protection strategy in Louisiana. *Restoration Ecology* 13: 499–506.

Quartel S, Kroon A, Augustinus PGEF, Van Santen P, and Tri NH (2007) Wave attenuation in coastal mangroves in the Red River delta, Vietnam. *Journal of Asian Earth Sciences* 29(4): 576–584.

Resio DT and Westerink JJ (2008) Hurricanes and the physics of surges. *Physics Today* 61(9): 33–38.

Scyphers SB, Powers SP, Heck Jr. KL, and Byron D (2011) Oyster reefs as natural breakwaters mitigate shoreline loss and facilitate fisheries. *PLoS One* 6(8).

Sheppard C, Dixon DJ, Gourlay M, Sheppard A, and Payet R (2005) Coral mortality increases wave energy reaching shores protected by reef flats: Examples from the Seychelles. *Estuarine, Coastal and Shelf Science* 64: 223–234.

Silliman BR and Bertness MD (2002) Atrophic cascade regulates salt marsh primary production. *Proceedings of the National Academy of Sciences of the United States of America* 99: 10500–10505.

Temmerman S, Bouma TJ, van de Koppel J, van der Wal D, de Vries MB, and Herman PMJ (2007) Vegetation causes channel erosion in a tidal landscape. *Geology* 35(7): 631–634.

Thampanya U, Vermaat JE, Sinsakul S, and Panapitukkul N (2006) Coastal erosion and mangrove progradation of Southern Thailand. *Estuarine, Coastal and Shelf Science* 68: 75–85.

Tilman D, Reich PB, and Knops JMH (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441: 629–632.

Twilley RR, Chen RH, and Hargis T (1992) Carbon sinks in mangroves and their implications to carbon budget of tropical coastal ecosystems. *Water, Air and Soil Pollution* 64: 265–288.

UNEP-WCMC (2006) *In the Front Line: Shoreline Protection and Other Ecosystem Services From Mangroves and Coral Reefs*, 33 pp. Cambridge, UK: UNEP-WCMC.

Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of Earth's ecosystems. *Science* 277: 494–499.

Waycott M, Duarte CM, Carruthers TJB, *et al.* (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences* 106: 12377–12381.

van Wesenbeeck BK, Van De Koppel J, Herman PMJ, and Bouma TJ (2008) Does scale-dependent feedback explain spatial complexity in salt-marsh ecosystems? *Oikos* 117(1): 152–159.

Wilkinson C, Linden O, Cesar H, Hodgson G, Rubens J, and Strong AE (1999) Ecological and socioeconomic impacts of 1998 coral mortality in the Indian Ocean: An ENSO impact and a warning of future change? *Ambio* 28: 188–196.

Worm B, Barbier EB, Beaumont N, *et al.* (2007) Biodiversity loss in the ocean: How bad is it? *Response. Science* 316: 1282–1284.

Yachi S and Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences* 96: 1463–1468.

Zhang HQ (1993) Stabilitätsparameter des Deichvorlandes und deren Beeinflußung durch Bewirtschaftungsverfahren. *Mitteilungen der Deutschen Bodenkundlichen Gesellschaft* 72: 297–300.

Near East Ecosystems, Animal Diversity

Joseph Heller, The Hebrew University of Jerusalem, Jerusalem, Israel

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 329–352, © 2001, Elsevier Inc.

Glossary

Allopatric Describing two taxonomic entities whose geographic ranges do not intersect with each other.
Endemism The fact of being found in a specific location of limited size, rather than being widely distributed.
Eremic Relating to or occurring in desert regions.
Parapatric Describing two taxonomic entities whose geographic ranges are distinct but adjacent to each other.
Xeromorphic Describing a form that has developed in response to highly arid conditions.

The Levant

The Levant is the eastern shoreland of the Mediterranean (Por, 1975), a stretch of land approximately 800 km long and approximately 150 km wide (**Figure 1**). It is wedged in between the Mediterranean Sea in the west and the Arabo-Syrian Desert in the east, stretching from the mouth of the River Orontes in the north to the Isthmus of Suez in the south. It consists of four basic north–south-oriented features: the coastal plain, the western (cis-rift) mountains, the rift valley, and the eastern (trans-rift) mountains. In the east it merges gradually into the Arabo-Syrian Desert.

The 800-km-long Levant has several climatic and floral belts. In its northern (and high-altitude) areas the overall habitat is Mediterranean: Summers are dry and warm, and winters are rainy and mild with occasional snow in the higher mountains. Natural plant associations consist of xerophytic shrubs and trees. In its southern and eastern areas (the Negev, Sinai, and fringes of the Syrian desert) the overall habitat is eremic: Temperatures are high, precipitation is low and very irregular, and the vegetation cover is poor. Between these two major habitats there is an intermediate habitat of steppe character. The combination of the four major north–south relief patterns with the three climatic belts imparts an unusual heterogeneity to this relatively small zoogeographic province and has resulted in a highly dynamic faunal history.

The Extent of Biodiversity

It is difficult to estimate animal diversity throughout the whole Levant because there is better knowledge about the fauna of its southern than of its northern areas. Unfortunately, the term "southern Levant" is loosely applied by various authors to include the area of Israel, and sometimes also the Palestinian Autonomy, and/or Sinai, Jordan, and the Golan Heights. These limitations should be taken in account when considering the data (of representative groups) in **Table 1**. They suggest that animal diversity in the Levant is high and perhaps that this area is one of the richest and most diverse natural regions among temperate regions of the world (relative to its size). For example, the area between the Mediterranean Sea and the Jordan River, together with the Golan Heights, is only approximately 28,000 km^2, but it contains approximately 170 bird and 100 mammal species (breeding). These numbers are not markedly fewer than those in California (210 birds and 110 mammals in an area of

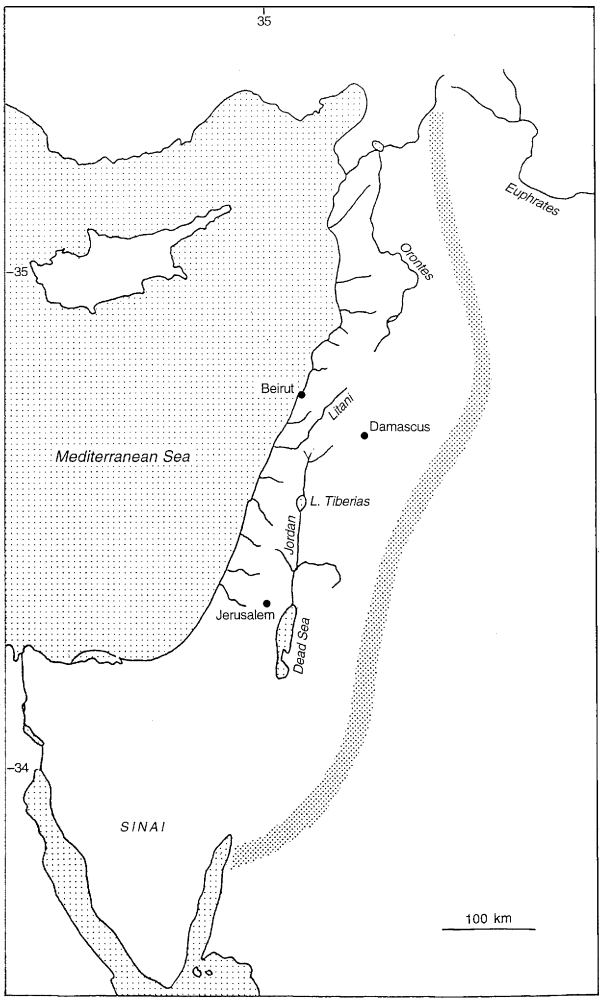


Figure 1 The Levant; its eastern boundary (stippled line) is hazily defined (adapted from Por FD (1975) An outline of the zoogeography of the Levant. *Zool. Scripta* 4: 5). Reprinted with permission from Wiley.

Table 1 Animal diversity in the Levant or in regions of it

Animal group	*No. of species*	*Region*	*Reference*
Crustaceans			
Ostracods (Ostracoda)	53	Israel	Martens and Ortal (1999)
Arachnaeids			
Scorpions (Scorpiones)	16	Southern Levant	Levy and Amitai (1980)
Camel spiders (Solifugae)	54	Israel	Levy and Shulov (1964)
Crab spiders (Thomisidae)	40	Southern Levant	Levy (1985)
Cobweb spiders (Theridiidae)	62	Southern Levant	Levy (1998)
Insects			
Dragonflies (Odonata)	82	All Levant	Dumont (1991)
Termites (Isoptera)	11	Southern Levant	Kugler (1988)
Bushcrickets (Tettigoniidae)	42	Israel	Ayal *et al.* (1999)
Grasshoppers (Acridoidea)	110	Southern Levant	Fishelson (1985)
Flies (Diptera)	3500	Southern Levant	Freidberg (1988)
Asilidae	158	Southern Levant	Theodor (1980)
Tephritidae	85	Southern Levant	Freidberg and Kugler (1989)
Butterflies (Rhopalocera)	106	Israel	Benyamini (1988)
Caddisflies (Trichoptera)	73	All Levant	Botosaneanu (1992)
Wasps (Vespidae)	7	Southern Levant	Kugler (1988)
Ants (Formicidae)	140	Southern Levant	Kugler (1988)
Mollusks			
Land snails (Pulmonata)	100	Southern Levant	Heller (1988)
Vertebrates			
Fish, freshwater (Pisces)	65	All Levant	Krupp (1987)
Amphibia + Reptilia	96	Southern Levant	Werner (1988)
Birds (Aves), breeding	170	Southern Levant	Yom Tov (1988)
Mammals (Mammalia), breeding	100	Southern Levant	Yom Tov (1988)

404,000 km^2), whose southern part lies at the same latitude as Israel.

Several factors contribute to this high diversity. One is the lack of a major disastrous disturbance; the general pattern of the landscape in the Levant has persisted since the early Pliocene. Only minor areas of the Levant (the tops of Mt. Lebanon and Mt. Hermon) were ever glaciated to any considerable extent, and only small regions (e.g., the Golan) were subject to volcanic eruptions that covered large areas with basalt. Consequently, the Levant never suffered the catastrophic wipeouts that hit northern and central Eurasia many times throughout the Pleistocene.

Local, region-scale changes did occur, however, repeatedly separating and then uniting different patches of landscape, thus permitting possible speciation under allopatry. These local geological events probably contributed, through speciation and through invasion, to enrich rather than to impoverish the Levantine fauna. This long, generally eventless but locally eventful history of the Levant is one of the reasons for its rich fauna.

A further reason is the heterogeneous climate. Mean annual rainfall may decrease from 1000 mm to less than 100 mm over a short distance of 150 km.The landscape therefore changes from Mediterranean oak maquis to batha and then to barren deserts, and the fauna gradually changes as one species replaces another. The net result is a fauna richer than would be expected were the landscape more uniform.

In contrast to factors contributing to biodiversity of the Levant, certain environmental factors limit it. The major limiting factor is that southern and eastern areas of the Levant are an extreme desert that, for many animals, is too hostile a habitat for survival.

Historic Zoogeography

Levantine biota constitute one of the most complex ecosystems in the world: Ethiopian and Oriental taxa coexist with Euro-Siberian, central Asiatic, and Mediterranean species, mixed with Saharan and Arabian desert faunal elements. This kaleidoscopic admixture of Palearctic, Paleotropic, and Saharo-Arabian elements has been constantly changing during the Neogene and the Quaternary periods, disposing a new biogeographic configuration during each geological period. Consequently, the Recent fauna is a living monument to a fascinating animal history and ecology. Current zoogeographic patterns in the Levant are the result of four major formative events (Por, 1975; Tchernov, 1988):

1. The breakup of the Tethys Sea during the Miocene and the consequent establishment of the Eurasian–African land bridge: Approximately 18–24 million years ago (late Oligocene–early Miocene) northern regions of today's Levant were still under the large Tethys Sea, whereas southern regions were terrestrial and part of the vast Afro-Arabian continent, with its unique faunal realm (**Figure 2**). Approximately 17 million years ago (Ma) the Tethys contracted, the emerging lands combined the southeastern areas of Asia Minor with the northwestern areas of Arabia, and the Levant was formed (**Figure 3**).

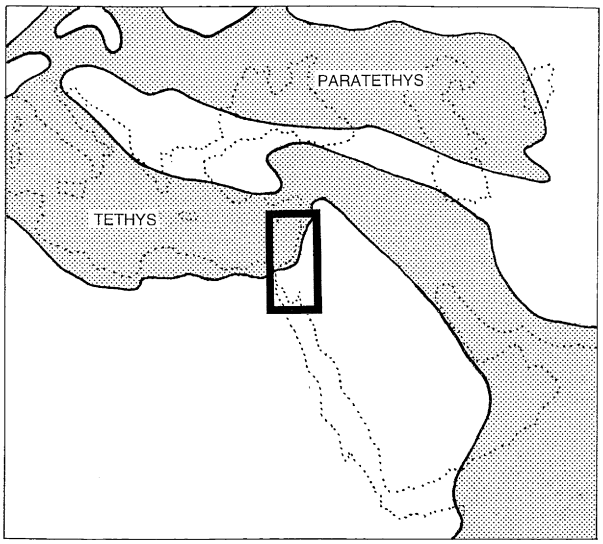


Figure 2 Paleographic map, 20 Ma. The Tethys is connected with the Indian Ocean, thereby separating Afro-Arabia from Eurasia (the Levant is represented by the rectangle). The southern Levant is part of the Afro-Arabian domain, and other parts are submerged (adapted from Tchernov E (1988) The palaeobiogeographical history of the southern Levant. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk). Reprinted with permission from Springer.

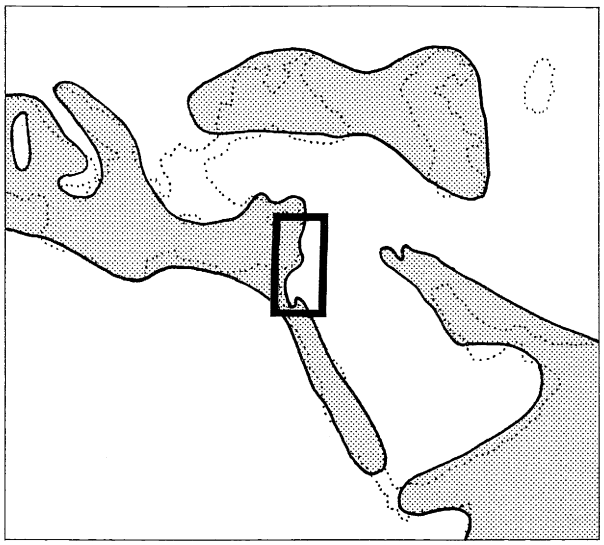


Figure 3 Paleographic map, 10 Ma. The Tethys is no longer connected with the Indian Ocean. In the north, extensive land connections with Eurasia enable the invasion of Palearctic elements into the Levant. In the south, land connections with Africa gradually break up due to the development of the Red Sea (adapted from Tchernov E (1988) The palaeobiogeographical history of the southern Levant. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk). Reprinted with permission from Springer.

Bridging the two continents, the newly formed Levant triggered an interchange of terrestrial biota between Africa and Eurasia. A wide spectrum of habitats were present in the Levant during much of the Miocene, including rivers, lakes, marshy habitats, dense woods, and Paleotropic rain forest, together with open country.

Early Miocene deposits in the southern Levant contain a diverse community of animal taxa, most of which were of African origin (proboscoideans, crocodiles, soft-shelled turtles, and catfish) but with others representing Eurasian elements (viverrines and cricetines). It seems that during this period African forms found it relatively easy to emigrate into the Levant, whereas contemporaneous Eurasian forms found it more difficult.

2. The late Miocene and Pliocene drying out of the Old World subtropics and the formation of the vast belt of the Saharo-Arabo-Syrian deserts, which restrained terrestrial migration between Africa and the Levant: The desertification process was very gradual. At first, approximately 12–15 Ma, open-land biota of savanna and semisteppe began to expand, and in doing so they created an extremely diversified mosaic of biotopes that ranged from the East African highlands through (still connected) Arabia to the Levant and enabled large-scale biotic interchange. Gradually, however, aridity along the Saharo-Arabian latitudes increased (associated with intensification of seasonality) so that by approximately 5.5–6 Ma (late Miocene) a vast arid belt of almost continental size formed which became a severe biogeographic barrier. Eventually, this desert barrier was settled by strongly eremic-adaptive species recruited from its immediate surroundings. The desert fauna of the Levant thus originates from both Eurasia and Africa and is established along adaptive rather than historical lines. It is more between than within faunal realms.
3. Quasi-isolation and biogeographical provincialism of the Levant during the Pliocene: Beginning approximately 5 Ma three geomorphological events had profound effects on the biota of the Levant. To the north, orogeny of the Taurus–Zagros mountain chains restrained many northern Palearctic elements from migrating southwards. To the south, the Red Sea opened up (**Figure 4**) and connected with the Indian Ocean, thereby establishing, for the terrestrial fauna, a marine barrier between Africa and the Levant (in addition to the already existing desert). The third event was the brief flooding of the Mediterranean by oceanic waters through the Gibraltar straight. All along the Levant coastal plains were submerged, and deep marine embayments penetrated into such lowlands of the Levant as (today's) Dead Sea Rift Valley and Plain of Beer-Sheva (**Figure 5**). These marine embayments of the Mediterranean (which gradually retreated later in the Pliocene) formed biogeographic barriers that enabled local speciation.

   Pliocene mammals of the Levant (from a site near Bethlehem) are of open-country species, suggesting a lowland African savanna-like landscape, with some freshwater bodies nearby. The absence of deer, beavers, dormice, and bears (common in the Pliocene of Turkey) is noteworthy.
4. The successive creation of the Jordan–Orontes rift valley in the late Pliocene and early Pleistocene: During the Plio-Pleistocene transition (approximately 1.8 Ma) large-scale tectonism occurred in the southern Levant, which established a longitudinal geomorphological feature that opened initially in the south and gradually northward. The Jordan–Orontes valleys, with their chains of lakes and swamps, were then created, the cis- and trans-rift mountain ranges were then lifted, and basalts were extruded in the

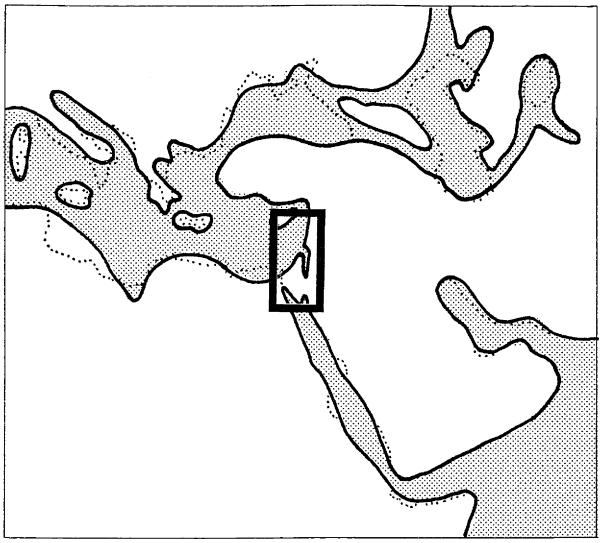


Figure 4 Paleographical map, 5 Ma. The marine connection established between the Red Sea and Indian Ocean results in relative isolation of the Levant from Africa (adapted from Tchernov E (1988) The palaeobiogeographical history of the southern Levant. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk). Reprinted with permission from Springer.

eastern regions of the Levant. It was during the Pliocene–Pleistocene that the contemporary relief structure of the Levant was molded (Por, 1975; Tchernov, 1988).

Levantine faunas during the lower Pleistocene were characterized by a bloom of Palearctic elements and a deterioration of tropical ones, suggesting that the Zagros Mountains filter barrier had by then been overcome and that the Saharo-Arabian barrier, between the Levant and Africa, had already become highly effective (Tchernov, 1988). By 1.4 Ma, two-thirds of the mammal genera (from a site in the Jordan Valley) were of Eurasian origin. The remaining one-third were mainly of tropical origin, and many African mammals entrapped in the Levant underwent local speciation (*Hippopotamus behemoth*, *Praomys minor*, and *Arvicanthis ectos*). By the mid-Pleistocene (800,000–150,000 years ago) additional Palearctic immigrants had invaded the Levant (squirrel, hare, Syrian bear, and wild cat), various tropical elements had become extinct (giraffe, monkeys, the fossil artiodactyls *Pelorovis* and *Kolpochoerus*), and there were no further massive invasions of tropical elements; the fauna was becoming similar to that of today. From 125,000 years ago onwards (upper Pleistocene) there were no drastic changes in faunal turnover rates. In many species of snails, voles, mole-rats, wolves, bear, and gazelle, however, there were considerable changes in body size to adjust to environmental fluctuations. These size alterations suggest that Pleistocene climatic changes in the Levant affected population levels more than species extinctions (Heller, 1988; Tchernov, 1988).

Today, the majority of animal genera in the Levant (more than three-fourths) occur outside the Levant exclusively to the north or northeast. Faunal connections with Eurasia are thus obvious. The extent of the Eurasian element varies from group to group. Almost all species of land snails are of Palearctic origin, as are three-fourths of the fruit flies (Tephritidae), two-thirds of the butterflies, and one-fifth of all chironomids.

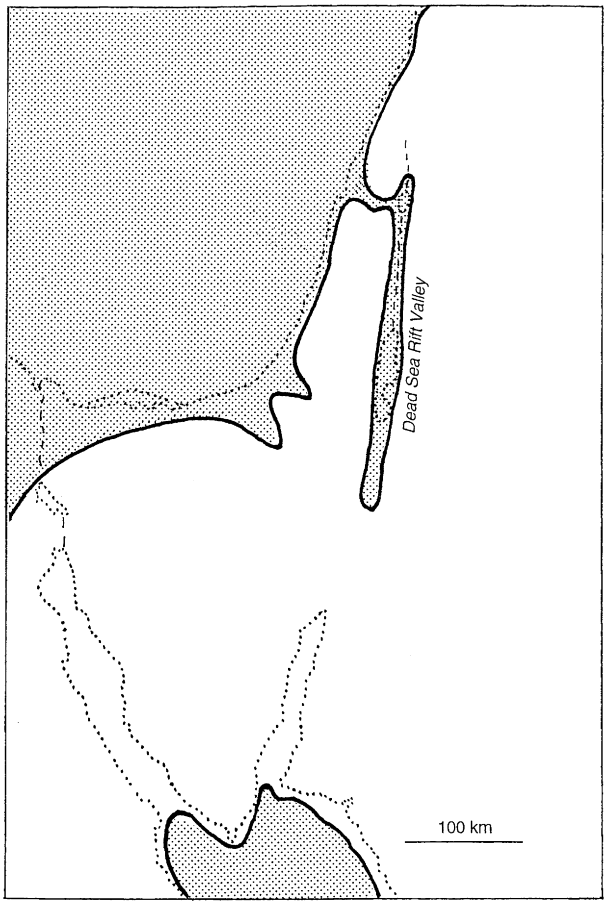


Figure 5 Paleographical map, 4–5 Ma. Marine ingression of the Mediterranean into the newly formed Jordan–Dead Sea Rift Valley brings marine elements into the Jordan Valley (adapted from Tchernov E (1988) The palaeobiogeographical history of the southern Levant. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk, based on Bernor (1987), Gvirtzman and Buchbinder (1977, 1978), and Rögl and Steiner (1983)).

Furthermore, the extent of the Eurasian element within each group may vary from one region of the Levant to another. From Mt. Hermon to Sinai frequencies of Palearctic butterflies decrease from 88 to 43%; within the Palearctic ant genus *Formica*, the number of species decreases from Turkey (12) through Lebanon (1) to Israel (absent; Kugler, 1988; Benyamini, 1988; Heller, 1988).

Tropical elements comprise approximately 5–15% of the fauna, but again this varies from group to group. Approximately 12% of the ant species are tropical or have strong tropical affinities; Diptera (in general) comprise approximately 5–10%, chironomids 40%, land snails 1%, and there are no aquatic oligochaetes. Within the Levant, tropical components in butterflies increase from approximately 8% on Mt. Hermon to 32% in Sinai.

Although most tropical elements are either Ethiopian or Paleotropic, a small Oriental element is also present (**Table 2**). This is noteworthy because the Orient is distant from the Levant and separated from it by a vast desert barrier. The influence of the Oriental species is sporadic and does not reach

Table 2 Oriental representatives in the levant

Group	*Species*	*Reference*
Spiders (Philodromidae)	*Thanatus fornicatus*	Levy (1991)
Ants (Formicidae)	*Polyrachis simplex*	Kugler (1988)
	Monomorium mayri	
	Lophomyrmex quadrispinosus	
Wasps (Vespidae)	*Vespa orientalis*	
Flies (Diptera)		
Chloropidae	*Anatrichus pygmaeus*	Freidberg (1988)
Culicidae	*Culex mimeticus*	
Butterflies (Rhopalocera)	*Madais fausta*	Benyamini (1988)
	Zizeeria karsandra	
	Limenitis reducta	
Caddisflies (Trichoptera)	*Hydroptila adana*	Botosaneanu (1992)
	H. fonsorontina	
	H. hirra	
	H. libanica	
	H. palaestinae	
	Chimarra lejea	
	Setodes alala	
	Stactobia pacatoria	
	Pseudoneureclipsis palmonii	
Fish (Pisces)	*Garra rufa*	(D. Golani personal communication)
	Garra ghorensis	
	Barbus canis	
	Nemachilus insignis	
	N. leontine	
Birds (Aves), breeding	*Ketupa zeylonensis*	Yom Tov (1988)
	Francolinus francolinus	
	Halcyon smyrnensis	
	Merops orientalis	
Mammals (Mammalia)	*Hystrix indica*	Yom Tov (1988)
	Nesokia indica	

sizable percentages in the whole fauna. Freshwater fishes and aquatic insects are exceptions, however.

In conclusion, the Levantine land bridge serves as an important crossroads of biotic exchange. Along the tropical wet and savanna-like landscapes of the rift valley, African animals entered Palearctis. Along the Mediterranean woodland and dry steppe landscapes of the bordering mountains, Eurasian animals invaded Africa. Among the dunes of the shoreland, eremic, Saharan biota advanced northward. The Levant thus functions as a complex biogeographic corridor in which representative species of the Ethiopian, Palearctic, and even Oriental biogeographic realms are encountered side by side. This mixture of faunas is outstanding and perhaps unique.

Endemic versus Widespread Animals

The Levant is characterized by the high frequency of marginal populations of species which range beyond the Levant. However, it is also rich in endemics (**Table 3**). Endemism is a function of a group's low mobility, combined with its tendency to speciate. It is therefore not surprising that within the vertebrates, endemism is higher among (the more sedentary) fish and reptiles of the Levant than among (the more mobile) mammals and birds; within arthropods it is higher among (the more sedentary) ants than among the (highly mobile) butterflies and wasps.

Comparisons within and between groups should not be carried too far, but the very high extent of endemism among land snails is noteworthy. This may be because they are, as a whole, very ineffective in crossing even minor ecological barriers of substratum or aridity. Consequently, many land snail species of the Levant have a very small range, sometimes only a few square kilometers. The endemic enid *Pene galilaea* is confined to a very small area of approximately 1.5×2.5 km in northwestern Upper Galilee, where it is surrounded by another species, *P. sidoniensis* (**Figures 6** and **7**); also, the endemic hygromiine *Trochoidea picardi* is confined to an area of 3×5 km in the coastal plain, where it is completely surrounded by *T. davidiana*. Similarly, the high level of endemism among bushcrickets is probably due to limited dispersal ability due to their loss of ability to fly. Approximately 45% of the Israeli bushcricket species are brachy- or micropterous, and 84% of these flightless species are endemic compared to 13% of the fully winged flying species.

It is tempting to seek "hot spot" regions, in which the extent of endemism is high and cuts across many animal groups. For the Levant, it is difficult to single out such regions. Endemics occur in many different regions.

In contrast to the endemics, several genera of the Levant enjoy an incredibly wide, intercontinental distribution. The

Table 3 Endemism among groups of the levant or regions of it

Animal group	*Endemic species*	*Reference*
Dragonflies (Odonata)	One-third	Dumont (1991)
Termites (Isoptera)	One-third	Kugler (1988)
Bushcrickets (Tettigonildae)	Half	Ayal *et al.* (1999)
Flies; asilids (Asilidae)	Half	Theodor (1980)
Butterflies	None, or negligible	Benyamini (1988)
Caddisflies (Trichoptera)	Two-fifths	Botosaneanu (1992)
Wasps (Vespidae)	None	Kugler (1988)
Ants (Formicidae)	One-fourth	Kugler (1988)
Land snails (Pulmonata)	Two-thirds	Heller (1988)
Fish (Pisces)	One-fourth	D. Golani (personal communication)
Reptiles and Anurans	One-tenth	Werner (1988)
Birds (Aves)	None, or negligible	Yom Tov (1988)
Mammals (Mammalia)	~2%	

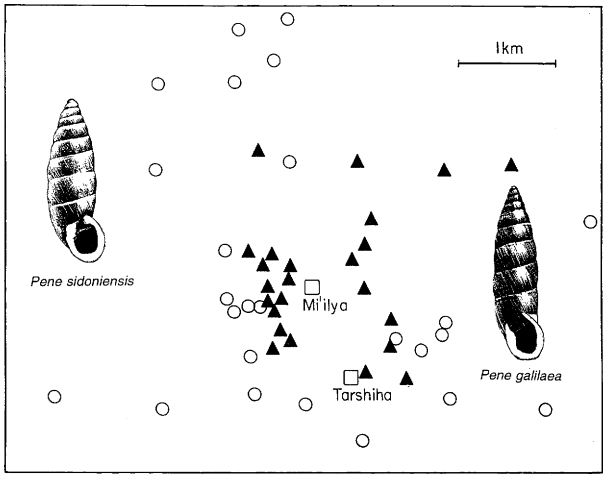


Figure 6 Narrow-ranged endemism in the Levant: total global distribution of the land snail *Pene galilaea* (▲) is only 8 km^2 in the western Galilee (○, sites of *P. sidoniensis*) (adapted from Heller, 1993). Reprinted with permission from Springer.

land snail genera *Truncatellina*, *Lauria*, *Cecilioides*, and *Punctum*, minute snails of the litter or subterranean environment, are found outside the Levant not only throughout Eurasia but also in much of the African continent. *Oxyloma*, an amphibious genus that lives on the vegetation above water, is also included in this list. All these taxa probably reached the Levant by aerial dispersal that, whether by wind or by birds, puts a premium on small size and light weight.

Aerial dispersal separates these widespread animals from many genera of Palearctic origin, which migrated into the Levant by the diffused, gradual movement of populations across hospitable terrain over many generations. It enables constant colonization. In the Levant, for land snails, birds are probably a more important aerial transporter than the wind. They migrate very regularly over considerable distances and are large enough for the accidental transport of minute animals without inconvenience. Birds migrating from Eurasia to Africa via the Levant could thus be responsible for at least some of the minute fauna of the Levant.

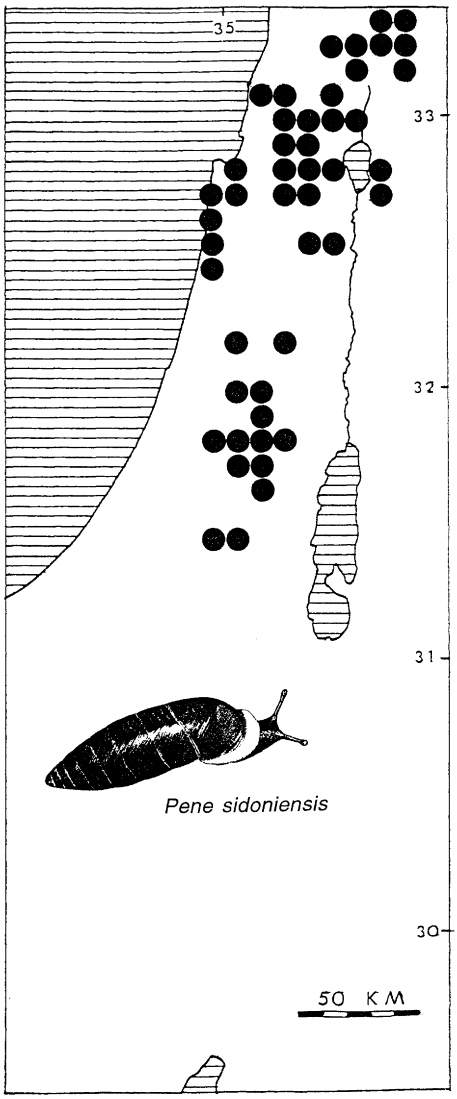


Figure 7 Wide-ranged distribution in the Levant: *P. sidoniensis* ranges over ca. 1200 km^2 in the southern Levant alone; it also occurs in the northern Levant and southern Turkey (adapted from Heller, 1993). Reprinted with permission from Springer.

Factors Determining Diversity

Diversity fluctuates considerably in the different geographic regions of the Levant. The following sections focus on rain and substratum, the major physical environmental forces that, by determining distribution, influence diversity; and on vegetation, an important biotic factor.

Rain

The Levant is situated on the margin of the extreme desert and the mean annual rainfall decreases from 1000 mm to less than 100 mm over the short distance of 150 km.This sharp gradient is reflected in animal ranges, with borderlines coinciding with isohyets so closely that there can be little doubt that rainfall is a major factor and of paramount importance in regulating animal distributions (**Figure 8**).

In rain-regulated distributions the species are classified into two categories: allopatric and parapatric. Allopatric species do not share borders with any other species of their genus. The dependence of such a species' range on rain, directly or indirectly, seems unequivocal. (For example, the Mediterranean-dwelling snails *Paramastus episomus* and *Pene sidoniensis* are the southernmost species of their genera. *Paramastus episomus* does not range below the 500-mm isohyet and *P. sidoniensis* not below the 400-mm isohyet.) Parapatric species pairs, on the other hand, share a common border of distribution. The location of the frontier within such pairs may well be influenced by interspecific competition, and rain could simply be the factor determining the point of equilibrium between the species (*Sphincterochila cariosa* and *S. fimbriata*, on the 500-mm isohyet, exemplify this group).

In many groups, overall biodiversity is a function of the amount of rain. In this context, the 200-mm isohyet, representing the end of the Mediterranean and the beginning of the desert landscape, appears to represent a barrier that many animal species cannot cross. In the Negev and Judean Deserts (areas receiving less than 200 mm of rain annually) only 25 land snail species are found compared to 73 species in the Mediterranean landscape. Of these 25 desert species, 7 occur in both the desert and the Mediterranean habitats. A generally similar pattern occurs at the generic level: Only 10 genera are recorded from the deserts of the Levant compared to 40 genera in the Mediterranean landscape; of these 10 desert genera, 8 occur in both desert and Mediterranean landscape and only 2 occur exclusively in desert areas. These comparative data indicate that adaptations to desert conditions are generally restricted to the species level rather than to higher taxonomic categories.

A similar trend of decrease in diversity also occurs among dipteran families (**Table 4**). In four latitudinal sections of Israel that broadly represent a decreasing gradient in rainfall, Freidberg (1988) noted that of 77 families, 20 were progressively restricted to the more northern sections, with 2 families occurring only in the extreme north. Only 57 families were detected from all four sections.

Among arachnids, the 350-mm isohyet (rather than the 200-mm one) demarcates the end of the Mediterranean and beginning of the desert fauna.

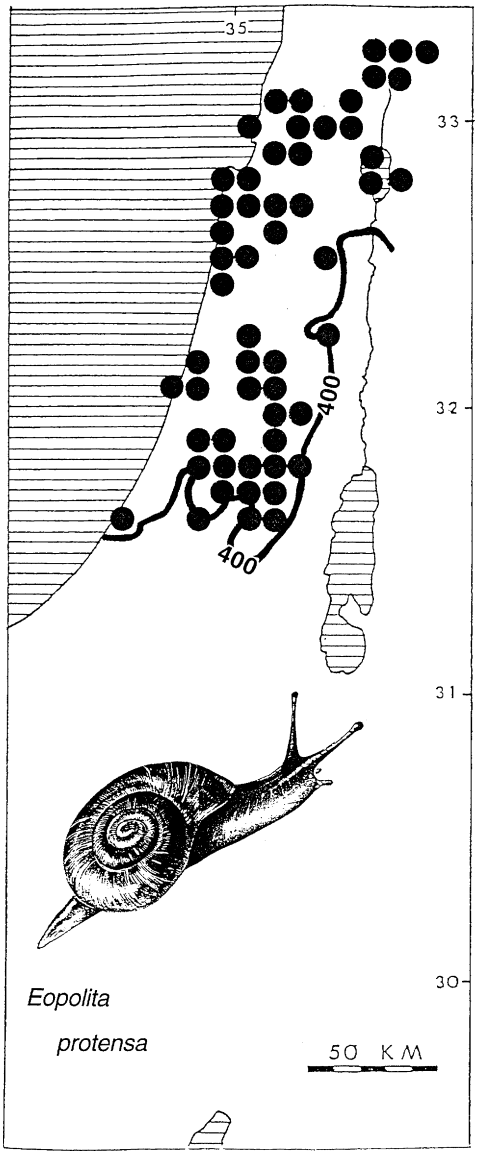


Figure 8 Distribution patterns that correlate with rain: The land snail *Eopolita protensa* is delimited by the 400-mm isohyete. Each dot represents all samples collected within a 5×5 km area (adapted from Heller, 1993). Reprinted with permission from Springer.

Studies of species distributions along climatic gradients may aid in predicting potential faunal responses to global climatic change. However, although a "univariate" approach in which a separate model is constructed for each species is more accurate than a multivariate one, it is not very efficient for analyzing climatic responses of large species assemblages or high taxonomic groups. However, scaling up from the level of individual species to higher taxonomic levels is crucial for evaluating faunal responses to global climatic change. Early suggestions (Heller, 1988) that rainfall is important in determining patterns of land snail diversity in Israel were based on nonquantified generalizations deduced from distribution maps of single species. Recent research (Kadmon and Heller, 1998) analyzes the response of the entire land snail fauna of Israel to gradients in mean annual rainfall by integrating

Table 4 Occurrence of diptera families in latitudinal sections of Israel

Latitudinal section	*No. of families common to various latitudinal sections*				*Total no. of families*
	A–D	*A–C*	*A–B*	*A*	
A (N of 33°)	57	10	8	2	77
B (32–33°)	57	10	8	—	75
C (31–32°)	57	10	—	—	67
D (S of 31°)	57	—	—	—	57

Source: Reproduced from Freidberg A (1988) The zoogeography of the Diptera of Israel. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

geographical information system (GIS) tools and standard multivariate techniques. Faunal variation of land snails was found to be significantly correlated with underlying variation in rainfall (as expected), but the effect of rainfall on the fauna was much greater in drier regions than in more rainy areas. Above 450 mm, no relationships could be detected between the patterns of faunal variation and rainfall.

Adaptations to Dry Conditions

Many desert species of the Levant differ from closely related Mediterranean species, or subspecies or populations, in their physiological adaptations to stress (Shkolnick, 1988). In the harsh environment of the desert, shortage of water and scarcity of adequate food are the major factors that pose a threat to many animals. I briefly discuss the various strategies that land snails and mammals employ to overcome these threats.

Concerning water economy, important adaptive features in determining a snail's ability to inhabit dry deserts include not merely resistance to desiccation but also instant adjustment of the rate of water loss to oncoming desiccation. Within the snail genus *Sphincterochila* (Sphincterochilidae), desert-dwelling *S. zonata* is capable of immediate response to desiccating conditions compared to Mediterranean species that take 3 or 4 days to recruit their water-preserving mechanisms; the steppe-dwelling species *S. fimbriata* is intermediate in its speed of adjustment (Arad *et al.*, 1989). A similar picture emerges among hygromiid and helicid land snails of the Levant. Desert-dwelling *Trochoidea simulata* and steppe-dwelling *Xeropicta vestalis* respond rapidly to desiccation, whereas the Mediterranean *Theba pisana* and *Monacha haifaensis* require up to 4 days to adjust their rate of water loss (Arad *et al.*, 1992).

Barriers to water loss in land snails of Levantine deserts include the shell; consequently, land snails in which the shell is missing, internal, reduced to a small external remnant, or without considerable calcium enforcement do not enter the desert of the Levant (*Limax, Limacus, Milax, Deroceras, Daudebardia, Eopolita, Oxychilus, Vitrea*, and *Oxyloma*). The epiphragm (a mucous-calcareous sheet spread over the shell aperture during seasons of inactivity) is also an important barrier to water loss, and *S. zonata* has the thickest epiphragm and also lower water loss than the Mediterranean species of *Sphinchterochila*. Another adaptation is the extrabody water reservoir that snails carry inside their shell: Species that lose significant amounts of water during desiccation do so almost equally from both body and extrabody compartments. However, in desert species the water content of the body is more closely controlled at the expense of extrabody water, thereby avoiding severe dehydration of soft body tissue.

Superiority in resistance to desiccation, however, is not always unique to desert animals. Mediterranean-dwelling *X. vestalis* enjoy a water loss of only 0.13% per day and cope better with desiccation than do desert-dwelling *T. simulata* (Arad, 1990). However, *X. vestalis* is a semelparous, annual species. It cannot survive even one single rainless year in the desert since all the populations would be wiped out in such an event. *Trochoidea simulata*, with somewhat less efficient mechanisms, reaches a life span of at least 3 years and can survive a rainless year (Heller, 1988; Arad, 1990).

Furthermore, within the genus *Sphincterochila*, desert-dwelling *S. prophetarum* has water regulatory capacities similar to those of the Mediterranean species of this genus. *Sphincterochila prophetarum* dwells underneath stones, where humidity is high compared to other microhabitats of the Levantine desert.

Desert snails of the Levant, to conclude, are adapted to withstand desiccation stress in that they usually have lower and slower rates of water loss and are capable of closer regulation of stable water content of the body compared to Mediterranean snails.

Desert mammals of the Levant differ from Mediterranean mammals in that they have a lower energy metabolism. Among murid rodents of Mediterranean landscapes, *Apodemus sylvaticus* and *A. mystacinus* possess the basic metabolic rate as expected from their body mass (according to the "mouse to elephant curve"). However, in the common spiny mouse *Acomys cahirinus*, a species of both Mediterranean and desert habitats, this rate is only 75% of the expected value, and in the golden spiny mouse *A. russatus*, a species of the extreme desert and regularly active during the day, it is only 55%. Similar comparative patterns have been found among gerbils, hedgehogs, carnivores, and ruminants of the Levant. A low metabolic rate, implying a lower rate of heat generation, saves the water otherwise needed to dissipate heat in a hot environment (Shkolnick, 1988).

In addition to a shortage of water, food in the desert is also at a premium. The frugal food requirements of desert animals may be related not just to a low demand for metabolizable energy. An efficient digestion of the food consumed may also reduce the amount of food required for maintenance. Desert mammals require less food for their maintenance than do nondesert ones. Bedouin goats herded in the southern Levant consume less food, retain it in the gut for a longer time, and digest it more efficiently than do European breeds of goat,

thereby gaining more energy from a given mass of food. In addition, the Bedouin goat has a remarkably spacious rumen, which functions not only as a fermentation vat but also as a voluminous water reservoir, enabling the goat to graze in the water-depleted desert without depending on frequent drinking (Shkolnick, 1988).

Balancing their nitrogen metabolism is as much a challenge for desert mammals as the maintenance of a balanced energy metabolism. The camel, instead of wasting the nitrogenous end products of protein catabolism, first retains them in its kidney and later allows them to be recycled through the gut. Here, the urea and other nitrogenous wastes are used by the microbial symbiotic population for resynthesis of protein, from which the host animal will eventually benefit. The potential capacity for recycling urea is far greater in desert than in Mediterranean species. Recycling of urea, in addition to helping the animal survive on low-protein feed, attenuates the load on the water otherwise required for the elimination of that waste (Shkolnick, 1988).

In conclusion, rain is the major environmental factor determining the diversity of animals within the Levant. The (approximately) 200-mm isohyet marks the arid limit beyond which many species of the Levant cannot exist.

Substratum

Within the boundaries of the relevant isohyets, animal distribution is further limited by lithic factors: The checkered geomorphologic map of the Levant presents a wide range of substrata, including granite, basalt, limestone, chalk, marl, heavy alluvium, calcareous sandstone, and light sand dunes.

Calcium-rich environments are very abundant in the Levant. Concerning one particular animal group, land snails, the abundance of limestone, chalk, or calcium-rich soils derived from them is a prerequisite for the flourishing of any testaceous fauna. Acid environments, which restrict the diversity in many regions of Europe and which in many evergreen forests throughout the world exist due to the gradual addition of acidic elements to the litter, are not known to exist in the Levant to any sizable extent. Granite and basalt, from which the extraction of calcium is difficult, are restricted to small areas. This lithic composition is an important reason for the high diversity of land snails in the Levant.

In some cases, the distribution of land snails varies in close association with geomorphology. The endemic *Cristataria genezarethana* occurs almost exclusively on mid-Eocene formations and is absent from nearby lower Eocene ones. Both formations consist of calcium, but only the mid-Eocene rocks are rich in crevices, which *C. genezarethana* inhabits (**Figure 9**). However, among many groups, only two major categories of substratum-dependent groups can be distinguished: the fauna of the limestone mountains and that of the loose sands. Limestone mountains in the Levant are the substratum that is richest and most diverse in species because they produce a rich mosaic of habitats and provide many shelter niches. Complex systems of rock crevices, moderately dense litter, soil pockets, generous vegetation cover, and a variety of exposed versus shaded sites are all to be found on limestone but not always on other substrata. Accordingly, a 100×100 m of limestone in the Judean Hills (in a Mediterranean landscape that receives approximately 570 mm of rain annually) yields 15 species of snails. A similar nearby area consisting of chalk, marl, or alluvium produces only 4 species, all of which also occur on limestone. Chalk, marl, or alluvium merely contain an impoverished, depauperated element of the rich limestone fauna but without any characteristic component of their own. Crevices provide shelter from the severe heat and drought prevailing throughout the summer. Whereas limestone has a well-developed system of crevices, chalkstone, marl, and alluvium do not and therefore, presumably, crevice dwellers cannot inhabit them. Also, the rich assemblage of litter habits, inhabited mainly by small invertebrates, is often well developed in limestone but nonexistent in nearby chalk or marl. Accordingly, there are no litter dwellers in such substrata. Extensive alluvial plains, such as Yizre'el, Zevulun, or the hinter parts of the Levant's coastal plain, harbor little diversity.

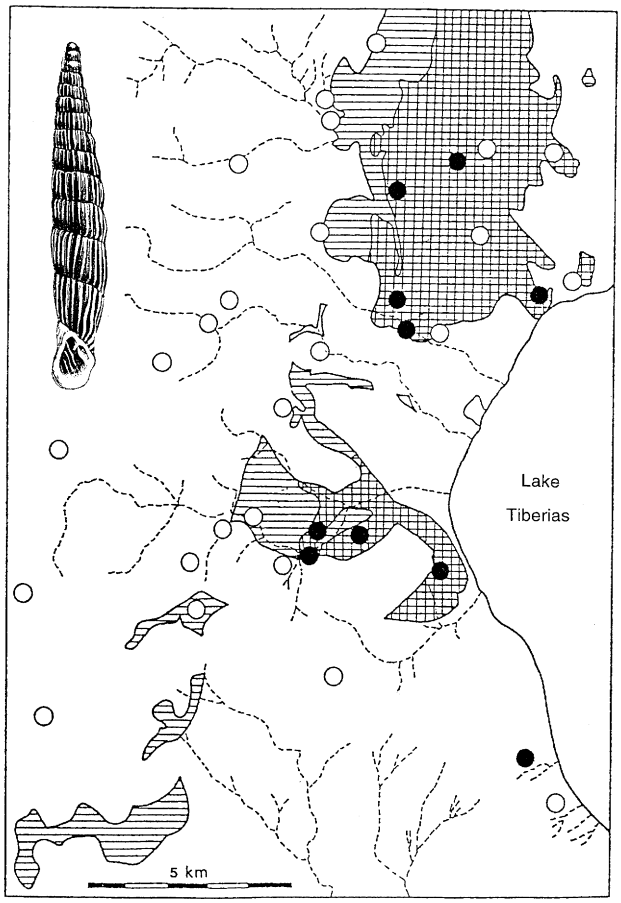


Figure 9 Distribution patterns that correlate with substratum: The land snail *Cristataria genezarethana* is restricted to the mid-Eocene formation (squares) and absent from the lower Eocene formation (lines). •, Sites in which snails were found; ○, sites in which they were not found (adapted from Heller, 1993). Reprinted with permission from Springer.

Sands of the Levant contain a fauna that differs considerably from that of the nearby calacareous mountains and alluvial soils (Kadmon and Heller, 1998). Furthermore, sands of the Mediterranean climatic region may differ from those of the Negev (**Figure 10**), which is arid.

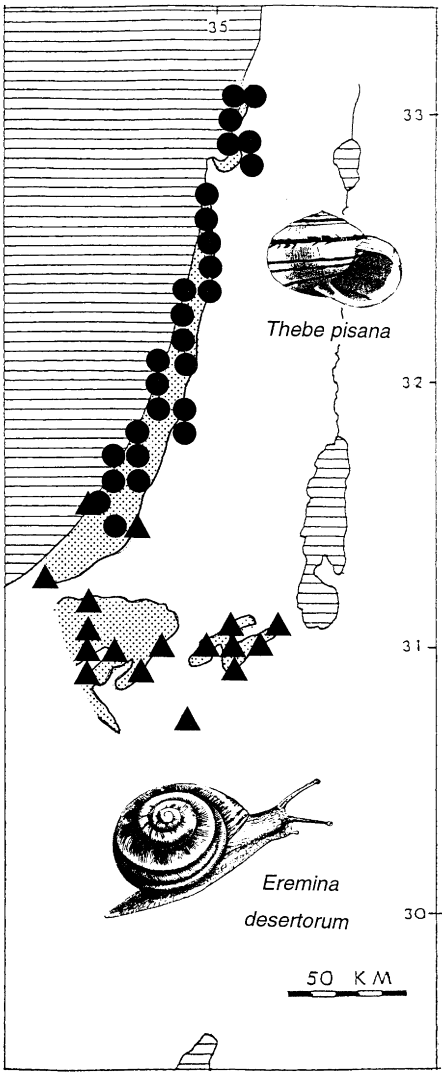


Figure 10 Distribution patterns that correlate with substratum: Land snails restricted to sand. *Theba pisana* (•) occurs in sands of more northern, Mediterranean climatic regions, whereas *Eremina desertorum* (▲) occurs in the more southern sands of the Negev. Each dot represents all samples collected within a 5×5 km area (adapted from Heller, 1993). Reprinted with permission from Springer.

Adaptations to Sand

Sand differs from other substrata in its "near-fluid" texture, which makes locomotion and burrowing more difficult than in other habitats. However, it holds more water than many other substrates so that more water is eventually translated into a more productive biotic community.

Many taxa of the animals restricted to sands converge in their morphological adaptations, thereby signifying the importance of sand as an evolutionary agent. Among reptiles of the Levant, a suite of morphological, physiological, and behavioral adaptations enables certain taxa to live in the sands or endows them with a competitive advantage (Werner, 1985). Locomotion over sand is facilitated in lizards by expansion of the feet. In *Acanthodactylus*, the extent of fringing along the toes varies with the looseness of the substrate intraspecifically and interspecifically (**Figure 11A**). Also, the toes of the gecko *Stenodactylus* ssp. are fringed, and those of the skink *Scincus scincus* are flatly expanded (**Figure 11B**). Sanddwelling viviperids locomote by sidewinding, as do *Cerastes* spp. and *Pseudocerastes fieldi*. *Lytorhynchus diadema* employs a unique variety of serpentine locomotion mechanisms wherein the loops push the sand down rather than sideways. Since burrows tend to collapse, on the one hand, and loose sand enables instantaneous submergence and sand swimming, on the other hand, problems arise from submergence and locomotion within the sand medium. Wedge-shaped snouts, frequent in burrowing reptiles, are prominent in the sand-swimming skinks *S. scincus* (**Figure 11B**) and *Sphenops sepsoides*. In these and in the snake *Lytorhynchus* the mouth opening is protected from sand during submerged progression by its ventral position. *Lytorhynchus* has an expanded rostral, common but not unique to sand-dwelling snakes. *Cerastes* spp. can submerge stationally by shuffling the flattened body sideways. The vertical flanks of *Scincus* (and *Sphenops*) may improve the lateral thrust during sand swimming. The relatively high species density of Gekkonidae on the sand may reflect the preadaptation of spectacled forms to sand. Breathing below sand is problematical not only at the nasal passage level but also at the thoracal: Upon exhalation the sand caves in and stifles inhalation. *Scincus* and *Sphenops* are structured for breathing with the protected ventral surface angularly set off from the lateral ones (**Figure 11B**; Werner, 1985).

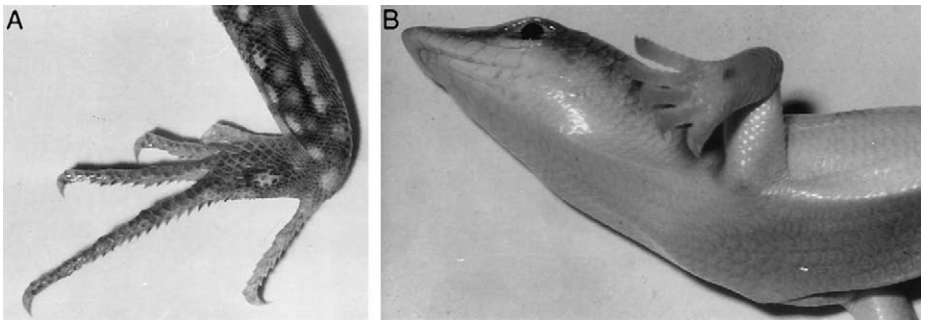


Figure 11 Some adaptations of Levantine psammophilous reptiles to sand. (A) Foot of *Acanthodactylus schreiberi syriacus* showing fringe of spiny scales along third and fourth toes. (B) *Scincus scincus* showing shovel-shaped snout, countersunk jaw, angularly set-off venter, and expanded toes. Photographs printed with permission from Y.L. Werner.

Sand-dwelling scorpions of the Levant can also be distinguished by their morphological adaptations. The most well-known are bristles to aid traction, locomotion, and burrowing in sand. These bristles occur on the tarsomeres and tibiae of all sand-living scorpions. Another leg modification in sand scorpions that acts to increase the effective surface area of the tarsi is long claws (Fet *et al.*, 1998). The bristles and claws create an expanded surface area on the tip of the legs and prevent sinking into the sand. The genus *Buthacus* (represented by *B. arenicola*, *B. yotvatensis*, *B. l. leptochelys*, and *B. l. nitzani*) is a typical sand dweller in the southern Levant, with up to 25 bristles on its walking legs (Levy, 1980).

Sand-dwelling grasshoppers can be distinguished from those of other biotopes by their behavioral adaptations. Sand-dwelling *Hyalorrhipis calcarata*, *Leptopternis* spp., and *Eremogryllus hammadae* dig into the sand with their hindlegs, throwing the sand behind them. They are thus able to cover themselves, leaving only the upper part of the head and the antennae exposed above the surface (species of other substrata usually settle "head to sun," thus casting the smallest shadow; Fishelson, 1985).

In conclusion, substratum is the second major environmental factor determining animal diversity in the Levant. Limestone rocks offer calcium, shelter, and an additional microhabitat—that of litter. Sands form a separate substratum category that usually has a unique fauna.

Vegetation

Both animal and plant distribution are influenced by rainfall and substratum, and consequently their distribution patterns may overlap. Whether vegetation influences animal distribution directly, in that different animal species have food preferences for different higher plants, varies among groups. Among fruit flies (Tephritidae), the larvae of practically all species in the southern Levant are phytophagous. All representatives of the Myopitinae, Oedaspidinae, Tephritinae, and Schistopterinae (65–70 species, 25 genera) develop only in plants of the composites (Compositae), whereas none of the Aciurinae and Dacinae do so. The daciine *Dacus oleae* attacks olives (Oleaceae), whereas other daciines are associated with fruits of milkweed (Asclepiaceae). Many tephritid species are oligophagous; a few are polyphagous, using plant hosts of different families. The parts of plants affected by tephritid larvae also vary considerably among groups. Larvae of Dacinae and Trypetinae develop in fleshy or juicy fruits. However, *Euleia heracleii* is a leaf miner, and *Capparimyia savastani* (**Figure 12**) produces larvae that develop in the flower buds of the caper *Capparis*. Most species of the five subfamilies that attack Asteraceae in Israel develop in fruit heads. The larvae of different species are usually restricted to feeding on certain parts of the flower head, such as flowers, achenes, or the receptaculum. Some species induce the formation of galls (e.g., *Myopites* spp. and *Urophora* spp.). Several species bore into the stems of their hosts, often forming galls (e.g., *Spathulina* spp. and *Oedaspis* spp.). *Orellia falcata*, unlike most members of the Terelliinae, mines along the stem of the goat's beard *Tragopogon* and pupates in the upper part of the root (Freidberg and Kugler, 1989).

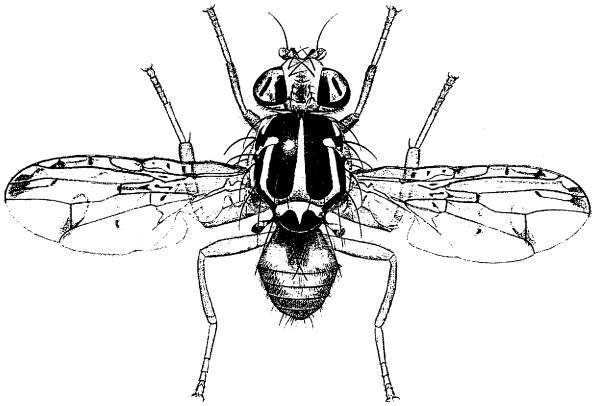


Figure 12 Distribution patterns that correlate with vegetation: *Capparimyia savastani* (Tephritidae), the larvae of which develop in flower buds of *Capparis* (reproduced from Freidberg A and Kugler J (1989) *Fauna Palaestina. Insecta 4—Diptera: Tephritidae*. Jerusalem: Israel Academy of Sciences and Humanities).

Some grasshopper species use plants as a food source, resting source, and hideout. Morphological adaptations to this habitat include large arolia to serve for attachment, streamlined and smooth body structures, heads with an acute or subacute angle, long and delicate colorless wings, and usually green or greenish-gray or brown coloration. Behavioral adaptations include resting along branches or stems and, if flying up, moving from plant to plant. When disturbed, many of these species crawl around the stem or branch, remaining on the opposite side to the intruder; others drop from the plant and dig among the dense branches close to the ground. There is a strong correlation between special types of vegetation and their fauna of plant-dwelling acridoids within the Levantine fauna: The Tropidopola–Ochrilidia association of grasshopper species is found on reed *Phragmites*, reed mace *Typha*, and rush *Juncus*; the *Heteracris–Eyprepocnemis* association shows a great affinity for knotweed *Polygonum* sp. and wormwood *Artemisia monosperma*; and the *Dociostaurus–Morphacris* association occurs mainly in areas with low, usually dry, ephemeral grass (Fishelson, 1985). It is not known if these associations are obligatory.

In contrast to fruit flies, among land snails of the Levant a direct influence of vegetation on diversity is unlikely. Snails feed on a wide spectrum of decaying vegetable matter and saprophytic fungi, occasionally on green tissues, and there is no evidence that food requirements for specific plants determine their wide-scale distribution patterns. Vegetation, however, may play an important role in microgeographic distribution, especially in preventing some snails from occupying certain habitats. Some species of the Mediterranean region (*Buliminus labrosus, Sphincterochila cariosa, Euchondrus septemdentatus, E. saulcyi, Paramastus episomus*, and *Xeropicta vestalis*) are found on the southern slopes of hills, where annual vegetation prevails, and are very uncommon on the northern slopes that are covered with dense oak maquis. In regions in which the vegetational distinction between northern and southern slopes is not obvious, the distinction of the snails between the slopes also becomes obscure, and they also occupy the northern slopes. A preference for the south slopes could perhaps be due to the annual vegetation prevailing on

it. Annual plants do not develop xeromorphic characters and are therefore easily eaten and digested. The oak maquis of the north-facing slope may be cool and damp and offer more shelter, but it consists mainly of perennial vegetation. The development of xeromorphic characters (such as the thick cuticle of the oak leaves) necessary for plant survival in the dry season could perhaps be an obstacle for snail feeding and digestion. Too little information is available on the precise diet of many animal groups. Once this information is gained, I believe that a major breakthrough will occur in our understanding of biodiversity in the Levant.

Freshwater Diversity: Historic Factors

The inland water system of the Levant is dominated by the north–south-oriented topography of the Rift Valley. The three major rivers of the Rift (the Jordan, Litani, and Orontes) run along successive segments of the valley and create a "steeple chase" waterway. Between the Rift basin and the Mesopotamian basin, watershed shifting and headwater capture may well have enabled a faunal transition of transcontinental dimensions.

During the Pliocene the Mediterranean flooded the southern Rift Valley and upon its retreat it left, in the freshwater of the Jordan Valley, several marine relicts. Crustacean species of the Jordan Valley suggested by Por (1975) as marine relicts of the Pliocene invasion include *Loxoconcha galilea*, *Pseudobradya barroisi*, *Nitocra balnearia*, *Monodella relicta*, *Typhlocirolana reichi*, *Typhlocaris galilea*, and *Bogidiella hebraea*. The blind prawn *Typhlocaris galilea* (**Figure 13**) is endemic to a subterranean warm, sulfide-rich spring near the shores of Lake Kinneret, where it feeds on oligochaetes (*Isochaeta israelis*) and snails (*Theodoxus jordani*).

From the Pleistocene onwards there was a continuous existence of freshwater or saline lakes in the Jordan Rift Valley. In this large endorheic system, the smallest climatic change and the changing balance of evaporation and precipitation (in addition to any slight tectonic movement) resulted in considerable fluctuations of water levels, with the creation, fusion, separation, and disappearance of lakes and rivers. This resulted in richly diverse faunas that partly replaced one another throughout the Pleistocene. Some old lacustrine freshwater endemics have survived in the Jordan system since the early Pleistocene (**Table 5**).

The fish fauna of the Levant is a mixture of Palearctic, Oriental, and Ethiopian species. Thirty-four species of freshwater fishes occur in the Orontes, the largest river of the Levant. Of these, approximately two-thirds are Palearctic, one-third are Oriental, and only 3% are African elements. The Orontes shares 18 of its species with the Tigris–Euphrates but only 7 with the Jordan River. The number of species occurring in the Jordan, the second largest river, is only slightly lower, but the species composition is totally different: Of 28 species, the Palearctic, Oriental, and African elements each comprise about one-third. The Jordan River shares only 3 species with the Tigris–Euphrates, but 6 additional species have their closest relatives in Mesopotamia. Apparently, the Jordan and various courses of the Orontes were separately colonized via branches of the Mesopotamian river system, and the Jordan drainage basin became an important center of speciation for several fish lineages with Palearctic affinities. Later, faunal elements of the Jordan reached the Orontes (through the Litani). Some freshwater fish of African origin reached the Jordan Valley by way of freshwater connections to the south. They could have done so up to Pliocene times, and in the Jordan Valley they differentiated specifically or even generically (*Astatotilapia flaviijosephi* and *Tristramella* spp.). During later periods, some African fish may have reached the southern Levant from the Nile via the Mediterranean Sea (the euryhaline cichlid *Tilapia zilli* survives well in brackish and even marine waters) (**Figure 14**). The fish fauna of the Rift's lakes and rivers was thus recruited from the north (Palearctic), east (Euphrates), and south (the Nile; Krupp, 1987).

The headwaters of Jordan constitute a typical temperate Paleoarctic freshwater fauna, with a "normal" level of diversity that includes pulmonate snails, cold-water copepods, and

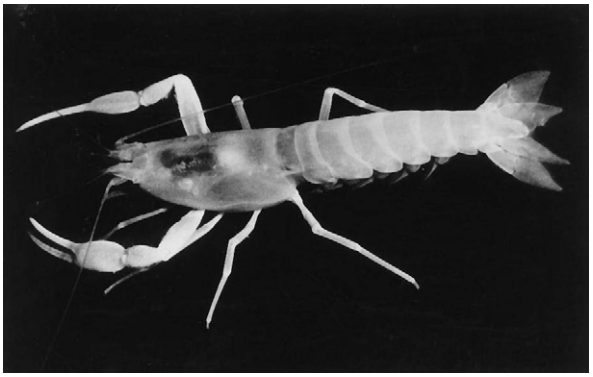


Figure 13 The blind prawn *Typhlocaris galilaea*, a relict of the Pliocene marine ingression into the Jordan Valley (photograph by Dr. M. Tzurnamal).

Table 5 Endemic species of the Jordan Rift Valley

Group	*Species*
Porifera	*Cortispongilla barroisi*
Tricladida	*Dugesia salina*
	D. biblica
Oligochaeta	*Isochaeta israelis*
Halacarida	*Limnohalacarus capernaumi*
	Lohmanella heptapegoni
Crustacea	*Loxoconcha galilea*
	Ilyyocypris harmanni
	I. nitida
	Pseudobradya barroisi
	Nitocra incerta
	Nannopus palustris tiberiadis
	Schizopera taricheana
	Parabathynella calmani
	Monodella relicta
	Typhlocirolana reichi
	T. steinitzi
	Typhlocaris galilea
	Bogidiella hebraea

Source: Reproduced from Por FD (1975) An outline of the zoogeography of the Levant. *Zool. Scripta* 4: 5.

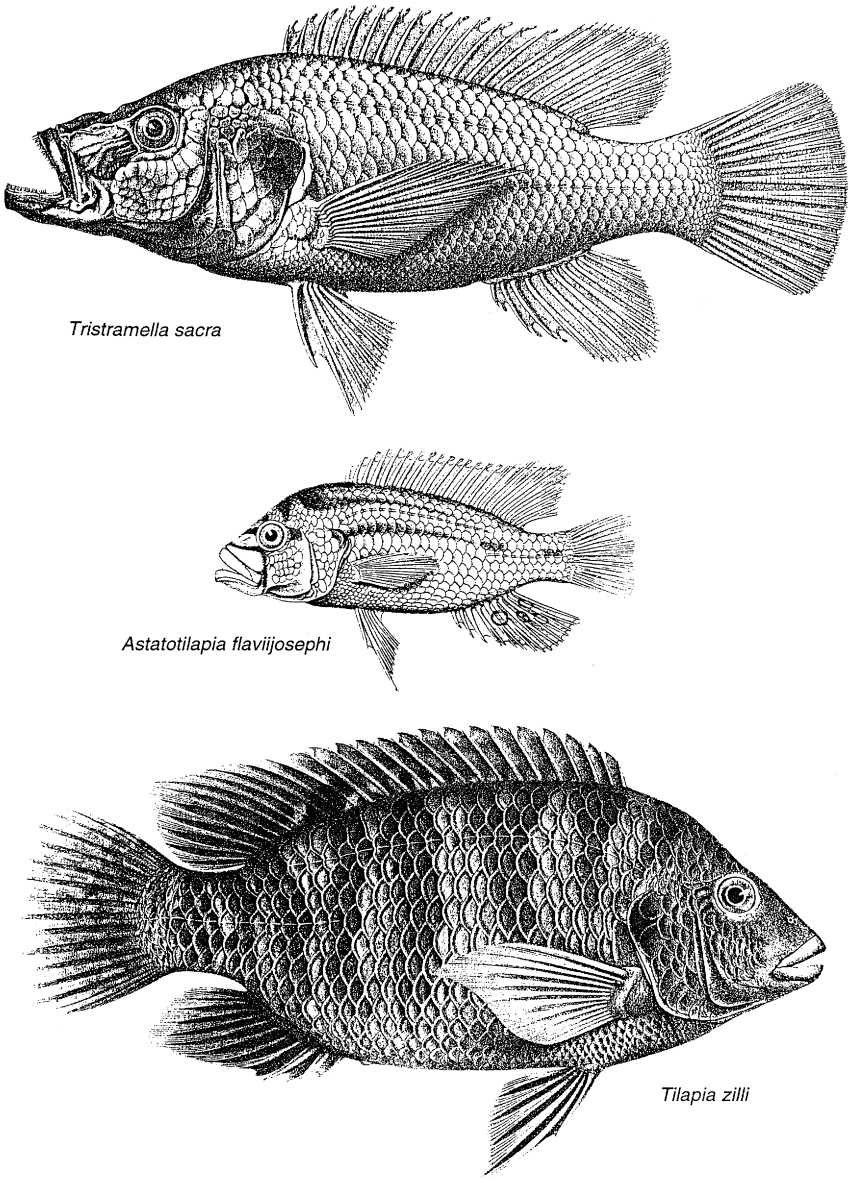


Figure 14 Cichlids of the Levant: *Tristramella sacra*, an endemic genus of African origin; *Astatotilapia flaviijosephi*, an endemic species of African origin; and *Tilapia zilli*, an African species that may have reached the Levant via the Nile and Mediterranean Sea (reproduced from Lortet (1883) Etudes Zoologiques sur la Faune au Lac de Tiberiade, Part I. *Arch. Mus. d'Histoire Nat.* 3, 99–194, Lyon, France).

many stream-living insects such as stone flies (Plecoptera), Elmidae among the Coleoptera, and typically cold-water-like *Rhyacophila* (Trichoptera). There are no less than 12 species of caddis flies (Trichoptera) in the springs of the River Dan; there are 13 species of Ephemeroptera in the River Hatzbani.

Elsewhere in the southern Levant, there is little of this rich freshwater fauna. In all the springs and brooks of Israel, usually only two species of Ephemeroptera can be found, *Cloeon dipterum* and *Caenis macrura*, which are well-known for their broad ecological valency. Only a small number of freshwater invertebrates are found in every spring and brook. This is a monotonous fauna characterized by *Melanopsis buccinoidea* (Mollusca), *Proasellus coaxalis*, *Echinogammarus syriacus* and *Eucyclops serrulatus* (Crustacea), and *Dina lineata* (Hirudinea). However, even this impoverished fauna does not advance much further than the oasis springs of En Avdat and of Qadesh Barnea, on the brink of the extreme desert. In most springs of Sinai, usually only insects or passively transported animals occur (Por, 1975).

Biodiversity: The Human Impact

Since the dawn of the Pleistocene, the Levant has been inhabited by the evolving hominids, starting with the early *Homo erectus* of the Jordan Valley. It was in the Levant that the 10,000-year-old Neolithic revolution of the Old World, with its agriculture and animal domestication, occurred.

The history of the Levant is characterized by tides of highly sophisticated agricultural civilization and ebbs of nomadic destruction. Agricultural apogees saw the development of soil protection, terracing, aqueduct and irrigation works, and drainage of swamps, whereas nomadism brought erosion, overgrazing, destruction of irrigation works, and swamping. Nomads semiconsciously created the deserts, of which they are often called the sons. Throughout history there was a progressive destruction of the woodland habitats by ever-increasing deforestation.

However, the most profound effects of man on biodiversity of the Levant started at the beginning of the twentieth century with the tremendous increase in human populations, use of firearms, extent of cultivated areas, and use of pesticides. These changes have had a pronounced, and in many cases fatal, effect on wildlife in the Levant. The effects of various human activities on animal life in the Levant (mainly in Israel and mainly concerning vertebrates) have been studied in depth by Yom-Tov and Mendelssohn (1988).

Hunting

The proliferation of firearms into the Levant by the end of the nineteenth century was followed by the overhunting and extermination of many game animals, mainly at the beginning of the twentieth century. Roe deer *Capreolus capreolus* and fallow deer *Dama dama mesopotamica* used to occur in the woods of Galilee and of Mt. Carmel; the last roe deer was shot in approximately 1912 and the last fallow deer probably at the beginning of the twentieth century. Oryx *Oryx leucoryx* still lived in Jordan at the beginning of the twentieth century; the last were shot before 1950. The last onagers *Equus hemionus hemippus* survived in the Syrian desert until the early 1930s. The bear *Ursus arctos syriacus*, which used to live near the Sea of Galilee and on the slopes of Mt. Hermon, was shot to extinction during World War I. The last cheetah *Acinonyx jubatus* in Jordan was shot in 1962. Two leopard subspecies existed in the Levant. The northern one, *Panthera pardus tulliana*, used to be found in Turkey, Lebanon, and Syria and was not rare in the Galilee at the beginning of the twentieth century; the last individual was killed in 1965, and today this subspecies is extinct. The other subspecies, P.p. nimr, survives in the Judean desert and Negev. An ostrich (*Struthio camelus syriacus*) was seen in Jordan in 1932; now this species is extinct in the Levant. The Nile crocodile *Crocodylus niloticus* was found in swamps near Mt. Carmel as late as 1912 (Yom-Tov and Mendelssohn, 1988).

Hunting also caused the almost total elimination of the green and loggerhead turtles (*Chelonia mydas* and *Caretta caretta*) which used to nest along the sandy Mediterranean beaches. During 1920–1930 approximately 30,000 turtles were killed along Israel's seashores; in 1985 only 14 nests were found along its whole Mediterranean coast.

Today, effects of hunting on wildlife in the Levant vary in broad correlation to the enforcement of hunting laws (and hence to political borders). In Israel, hunting today has only minor effects. There are only approximately 5000 licensed hunters and the main game species are chuckar partridges, rock pigeons, doves, various wintering waterfowl, hare, porcupine, wild boar, and mountain gazelle where they damage agricultural crops. Many of these may be hunted only during well-defined seasons, and none are seriously affected by hunting. Indeed, increasingly efficient enforcement of hunting laws has enabled several species to recover: the leopard *Panthera pardus nimr*, of which approximately 20 individuals now dwell in the Judean Desert and Negev; the wolf *Canis lupus pallipes*; and the hyena *Hyaena hyaena*. All wild mammals, birds, and mollusks in Israel are completely protected by law, with the exception of designated pest species (rats, house mice, voles, fruit bats, house sparrows, etc).

Habitat Destruction

Most habitats of the Levant have been severely or completely destroyed by urban and agricultural development. Two examples are presented in the following sections (Yom-Tov and Mendelssohn, 1988).

Coastal Sand Dunes

Coastal sand dunes form a narrow strip along the Mediterranean coast of the Levant, and until the beginning of the twentieth century were almost uninhabited by man. Today, however, the majority of the population of Israel lives on the coastal plain. The many urban settlements built on the coastal sands have caused loss of sand dune areas to such considerable extent that of 110 km^2 of dunes which formerly occurred north of Nahal Soreq, only 3 km^2 have not been developed.

The land snail *Trochoidea picardi* was an endemic species whose entire small distribution area once ranged over a few kilometers of calcareous sand dune ("kurkar") along the coast of Israel. Today, the entire distribution range of *T. picardi* is buried beneath the motorways, skyscrapers, and residential areas of the Tel Aviv metropolitan area, and this species is almost surely extinct (**Figure 15**). Concerning nonendemics, the coastal sand dunes are the most eastern extension of Saharan sand-dwelling animals into the Mediterranean, temperate zone of the Levant: the Egyptian tortoise *Testudo kleinmanni*, snakes (*Cerastes vipera*, *Macroprotodon cucullatus*, and *Lytorhynchus diadema*), skinks (*S. scincus* and *Sphenops sepsoides*), geckos (*Stenodactylus sthenodacylus* and *S. petrii*), lizards (*Acanthodactylus scutellatus* and *Agama savignii*), a monitor (*Varanus griseus*), a chameleon (*Chamaeleo chamaeleon musae*), rodents (*Gerbillus gerbillus*, *G. pyramidum*, and *G. andersoni allenbyi*, the endemic jerboa *Jaculus jaculus schlüeri*, and the endemic jird *Meriones sacramenti*), and a hedgehog, *Hemiechinus auritus*. The destruction of the coastal sand dunes (accompanied by predation by feral domestic cats and dogs) has reduced the distribution and abundance of these species.

Wetland Habitats

Wetland habitats that existed in the Levant at the beginning of the twentieth century included swamps, lakes, rivers, springs, and temporary (mainly winter rain) pools; man-made fish ponds were then barely existent. Today, however, water in the Levant is a scarce and much needed resource. All these habitats have been affected.

Swamps

At the beginning of the twentieth century, the Hula swamp ranged more than 40 km^2 north of the (13 km^2) Hula Lake

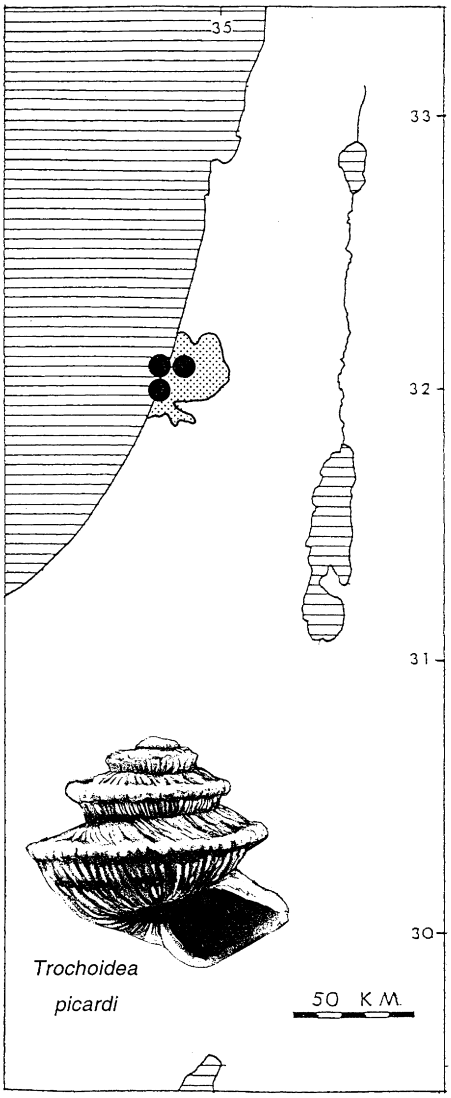


Figure 15 Extinction due to habitat destruction: total global distribution of the land snail *Trochoidea picardi* (museum records; each dot represents all samples collected within a 5 × 5 km area). Stippled area, Tel Aviv metropolis, 1999 (adapted from Heller, 1993). Reprinted with permission from Hebrew University of Jerusalem.

(**Figure 16**) and contained a rich diversity of (mainly widespread) animals. The swamp and lake were drained in the 1950s, but a small nature reserve (3 km^2) was created that represents an impoverished version of the original swamp. Even so, 22 species endemic to the Levant have not been recorded from the Hula since its drainage (**Table 6**). Although the Hula nature reserve is today rich in bird life, eight bird species that once bred in the Hula Valley do not breed there anymore.

Riverine Habitats

The rivers, streams, and wadis in the Levant drain, in general, either to the Mediterranean Sea or to the Rift Valley. Today, most Mediterranean-flowing rivers of the southern Levant are polluted, and consequently most fish have disappeared from them. Gray mullet (*Mugil cephalus* and *M. ramada*) and eel (*Anguilla anguilla*) fingerlings once used to spend their first years in these rivers (they reproduce in the sea) but now do not occur in most of them. Eight other, entirely freshwater, species are today very rare and the endemic *Acanthobrama telavivensis* now faces extinction. The soft-shelled turtle *Trionyx triunguis*, which once lived in many of the Mediterranean-flowing rivers and nested on their shores, has disappeared from most of them. Only in Alexander River does a breeding population persist, but many of the eggs fail to hatch and the population consists only of very large specimens. The river otter (*Lutra lutra*), a former inhabitant of all these rivers, survives only in the Bezet, Keziv, and Na'aman, in which it is only occasionally seen.

In the Rift Valley-flowing streams the major threat to biodiversity is the extreme reduction in the amount of water that flows in them since most springs are piped by man and used for irrigation. This has reduced and even eliminated the populations of several animal species. The river otter, once common in each major tributary in the Jordan River system, is now rare. The blue-cheeked bee-eater *Merops supeciliosus*, which bred along the Jordan River in the early 1930s, does not breed there any more (Yom-Tov and Mendelssohn, 1988).

Temporary Winter Rain Pools

A multitude of temporary winter rain pools once existed in temperate regions of the Levant. Many existed near Arab villages, where their waters once served for watering livestock and for irrigation. Many of these village pools were man-made, built as reservoirs by erecting a dam across a wadi or hewn into rock at the foot of a water-collecting slope, and were probably ancient. These pools supported a rich and varied invertebrate fauna and several species of amphibians.

The number of winter rain pools decreased sharply during the 1960s and 1970s. Many of the formerly undrained valleys were turned into cultivated fields. Other pools were eliminated due to road construction and building. Continuous spraying of herbicides on roadsides and of insecticides on pools near human settlements destroyed this ecosystem or turned these pools into breeding sites for pesticide-resistant mosquito strains. One result of these operations is that the once abundant amphibians have become a threatened group. The spadefoot toad *Pelobates syriacus* (**Figure 17**) breeds exclusively in rainwater pools and its tadpoles require a long (3 months) development period. Only large seasonal ponds of sufficient duration can support a viable population of spadefoot toads, and only 27 such ponds remain in Israel (including the Golan). Lack of suitable breeding localities has also caused the newt *Triturus vittatus* to become a threatened species, and even the formerly very common tree frog *Hyla arborea* and green toad *Bufo viridis* are becoming rare (Yom-Tov and Mendelssohn, 1988).

Poisoning

More than 700 compounds are registered in Israel's Ministry of Agriculture for agricultural use, including insecticides, acaricides, nematocides, fungicides, herbicides, bactericides, molluskicides, rodenticides, insect attractants, bird and mammal repellents, fumigants, plant growth regulators, and defoliants.



Figure 16 Extinction due to habitat destruction: Lake Hula, now drained (photograph by Kluger). Reproduced with permission from Central Zionist Archives.

Effects of pesticide residues on wildlife started shortly after DDT began to be used at the end of World War II. The most dramatic effect, however, was during the early 1950s, when thallium sulfate-coated grain was widely used to control rodents. These poison campaigns were not necessary. The main damage of these rodents was caused before 1950, when fields were plowed with the traditional shallow-plowing plows that did not disturb rodent burrows. When, after 1950, deep plowing was introduced, rodent burrows were destroyed, the rodents were exposed to predation, and the damage they caused decreased (pest control officers, however, attributed the decrease in rodent damage to their intense poison campaigns). In unplowed areas, as in alfalfa fields, damage caused by voles is still considerable and no rodenticides can prevent it. Since the 1960s, fluoracetamid has replaced thallium sulfate. The residues of both substances accumulate in bodies of secondary consumers such as raptors, bats, and carnivores.

Of 39 species of birds of prey that occurred in Israel before the use of pesticides, all but 2 were seriously affected. Species that once were common breeders became very rare breeders (the black kite *Milvus migrans*, griffon vulture *Gyps fulvus*, long-legged buzzard *Buteo ferox*, Bonnelli's eagle *Hieraeetus fasciatus*, Egyptian vulture *Neophron percnopterus*, kestrel *Falco tinnunculus*, lesser kestrel *Falco naumanni*, and lanner falcon *Falco biarmicus*). Rare breeders went extinct or their populations decreased drastically (the lapper-faced vulture *Torgos tracheliotus*, spotted eagle *Aquila clanga*, peregrine *Falco peregrinus brookei*, marsh harrier *Circus aeruginosus*, black eagle *Aquila verreauxi*, white-tailed eagle *Haliaetus albicilla*, bearded vulture lammergeier, and *Gypaetus barbatus*). Thallium sulfate also affected the wintering raptors, whose populations decreased considerably; some species disappeared completely for many years (the sparrowhawk *Accipiter nisus* and merlin *Falco aesalon*). Insect-eating birds, such as lesser kestrel *Falco naumanni* and scops owl *Otus scops*, were perhaps more affected by DDT and other persistent insecticides. The short-toed eagle *Circaetus gallicus* remained unaffected, probably due to its specialized reptilian diet and to its absence in winter when most of the poison grain was used. The latter factor also saved the hobby *Falco subbuteo*, the breeding population of which returns to the Levant only in May.

Today, after prohibition of the use of DDT, other chlorinated hydrocarbons, and thallium sulfate, some species have made a comeback but the existing populations of breeding raptors are only a fraction of their former populations. Hundreds of pairs of griffin vulture once bred in Galilee and on Mt. Carmel, but only 20 breed today; only one pair of lapper-faced vultures bred in 1986 in the Negev, where approximately 25 pairs had bred before the widespread use of agricultural poisons began. Continuous breeding attempts by long-legged buzzards, Bonnelli's eagle, and golden eagle *Aquila chrysaetus* often fail; only the kestrel seems to have almost fully recovered.

The removal of many raptors had secondary effects on the fauna. The population increase of the blackbird *Turdus merula*, the bulbul *Pycnonotus barbatus*, the palm dove *Streptopelia (Stigmatopelia) senegalensis*, the Syrian woodpecker *Dryobates syriacus*, and the jay *Garrulus glandarius* may be partly attributed to the decrease in the numbers of their predators (mainly the sparrowhawk *A. nisus*).

Secondary poisoning by insecticides also affected some insectivorous birds, particularly species that lived near fields and human settlements. Populations of the swallow *Hirundo rustica*, red-rumped swallow *Hirundo daurica*, white-throat *Sylvia communis*, rufous bushchat *Cercotrichas galactotes*, nubian shrike *Lanius nubicus*, spotted flycatcher *Muscicapa striata*, roller *Coracias garrulus*, bee-eater *Merops apiaster*, and Egyptian nightjar *Caprimulgus aegyptius* all decreased considerably.

Also, insectivorous bats (Microchiroptera) have been affected by chemicals. Twenty-eight species are known in the southern Levant, and many caves used to be inhabited by thousands of roosting bats. The numbers of both

Table 6 Aquatic species endemic to the Levant and not recorded from the Hula since Its drainage

Sponges (Porifera)
Ephydatia syriaca
Turbellarians (Turbellaria)
Dendrocoelum sp.
Dugesia biblica
Phagocata? armeniaca
Crabs (Crustacea)
Copepods Copepoda)
Attheyella trispinosa affinis
Insects (Insecta)
Dragonflies (Odonata)
Rhyothermis semihyalina syriaca
Urothermis edwardsi hulae
Calopteryx hyaline
Caddisflies (Trichoptera)
Orthotrichia? melitta
Beetles (Coleoptera)
Herophydrus cleopatre
Herophydrus galileae
Hydrovatus n.sp.
Laccobius levantinus
Canthydrus ornatus
Mosquitoes (Diptera)
Anopheles pharoensis
Bivalves (Bivalvia)
Potomida littoralis semirugata
Unio terminalis terminalis
Fish (Pisces)
Tristramella simonis intermedia
Nun qalilaeus
Mirogrex terraesanctae hulensis
Amphibians (Amphibia)
Discoglossus nigriventer
Mammals (Mammalia)
Arvicola terrestris hintoni
Birds (Aves) breeding
Aquila clanga
Aquila pomarina
Anhinga melanogaster chantrei
Fulica atra
Porzana pusilla
Chlidonias hybrida
Chlidonias niger
Sterna albifrons

Source: Adapted from Dimentman Ch, Bromley HJ, and Por FD (1992) *Lake Hula.* Jerusalem: Israel Academy of Sciences and Humanities.

bat-inhabited caves and individuals roosting in them have decreased drastically since the end of the 1950s and most species have become rare. This decline is attributed to two main factors. The first is fumigation of caves with ethylene dibromid and later with lindane (Gammexan). Fumigations were intended to control the fruit-eating bats *Rousettus aegyptiacus*, which were considered fruit pests and which often roost together with insectivorous bats. It was later found that the damage allegedly caused by fruits bats was marginal at best and had been grossly exaggerated by pest control officers. The population of one insectivorous bat, *Pipistellus kuhli*, remained less affected than the others, probably because it roosts not in caves but in hollow trees and in wooden houses. The second factor affecting insectivorous bats is secondary poisoning: Most bat species occur in the cultivated regions and are highly vulnerable to secondary poisoning because noctuid moths are an important part of their diet. Larvae of noctuids such as *Spodoptera littoralis* are serious agricultural pests, and fields are regularly sprayed with insecticides to eliminate them. Hence, secondary poisoning may be another factor contributing to the decrease of bat populations in the Levant.

The remaining bat populations suffer now from another threat. Increasing numbers of hikers visit various caves and disturb roosting and hibernating bats, thus causing wastage of fat reserves and desiccation of the hibernators.

Carnivores are another group which suffered from poisoning. In 1964, the Plant Protection Department of Israel decided that jackals (*Canis aureus*), one of the few mammals not protected at that time by the "Wild Animals Protection Law," were a nuisance. They were being blamed for damaging plastic sheets used to cover certain crops. A large antijackal campaign was started and tens of thousands of chicks injected with 1081 (fluoracetamid) were spread over the Mediterranean area of Israel in an effort to eradicate the jackals. Several mammal predators were affected, including the jackal, the wolf *Canis lupus*, the red fox *Vulpes vulpes*, the Egyptian mongoose *Herpestes ichneumon*, the jungle cat *Felis chaus*, and the African wild cat *F. sylverstris*. Most species recovered within a few years and their numbers today are similar to those before the poisoning campaign. The jackal, however, is much slower to recover and its numbers in Israel are still low. The wolf, whose populations in the Negev have increased in recent years due to greater food availability at garbage dumps, is endangered in central and northern Israel, where individuals are larger than those in the Negev and occasionally prey on livestock. Although the wolf is legally protected, stock owners often retaliate by poisoning.

The caracal *Caracal caracal*, which was formerly known only in the Negev, has increased its distribution area northwards into the Galilee. This range extension coincided with the population decrease in many Mediterranean carnivores following the jackal poisoning operation, suggesting that range extension of the caracal was possible due to the absence of competitors. Apparently, the main competitor of the caracal is the jackal, which preys on hares, the main food item of the caracal. After the jackal poisoning campaign, hares increased considerably. Today, after most predator species have recovered from the effects of poisoning, caracals still occur in the Mediterranean region of Israel but seem to be rare.

Changes in Agricultural Practices

Current agricultural practices in the southern Levant are different from those of the past mainly in that today approximately half of the cultivated area is irrigated. As a result, areas which formerly were left fallow in summer are now cultivated in the dry season.

Some species have responded to the change in agricultural practices with an increase in both abundance and distribution. Population size of the mountain gazelle *Gazella g. gazella* in central and northern Israel increased between 1948 and 1985 from less than 500 to 10,000. Availability of succulent,



Figure 17 Extinction due to habitat destruction: The amphibian *Pelobates syriacus* is now a threatened species because wetland habitats are disappearing (photograph by Y. L. Werner).

nutritious food and of water throughout the year in irrigated agriculture enables female gazelles to deliver their first fawns at the age of 1 year (2 years in nonagricultural areas) and average 1.8 young annually (only 1 in nonagricultural areas). Gazelles have become a pest in several areas because of their high numbers: The males damage fruit trees by rubbing their horns against the bark, and both sexes eat cotton, wheat, corn, and other crops. Increasing food and water availability in agriculture has also affected the dorcas gazelle *G. dorcas* in the southern Negev. Whereas populations near agricultural settlements increased between 1964 and 1985 by 9% annually, the rate of increase in other populations was only approximately 3%. Normally, dorcas gazelles in the southern Levant have only one fawn per year, in spring. Even if this fawn is lost, the dam does not become estrous until the normal breeding season in autumn. In recent years, increasingly more fawns are born in autumn so that either some females breed twice per year, as do mountain gazelles, or females become estrous again after losing their fawn. Autumn fawns are seen mainly near agricultural areas.

Availability of green food in summer and leftover grain in wheat and barley fields has enabled several seed-eating birds to increase their populations to such an extent that they have become agricultural pests: the collared dove *Streptopelia decaocto*, palm dove *Stigmatopelia senegalensis*, house sparrow *Passer domesticus*, and feral domestic pigeons *Columba livia domestica*.

The planting of exotic trees and ornamental plants in human settlements has made new areas available to several bird species. Typical woodland species, such as the syrian woodpecker *D. syriacus*, blackbird *Turdus merula*, great tit *Parus major*, and jay *G. glandarius*, once restricted mainly to woodlands of the Galilee, Mt. Carmel, and Judean hills, are now widespread and common in the gardens and parks that accompany many human settlements, sometimes even in the Negev and Jordan Valley. The Syrian woodpecker has become a pest in certain areas because it drills holes in irrigation pipes and in telephone cables.

Approximately 1500 plant species, mostly ornamental plants, were imported to Israel during the twentieth century. The orange-tufted sunbird *Nectarinia osea* is favorably affected by this enriched flowering vegetation of ornamental exotic plants in human settlements. When first discovered (in the nineteenth century) it occurred only in the lower Jordan Valley and near the Dead Sea, its distribution coinciding with that of the mistletoe *Loranthus acaciae*. Today, however, owing to the presence of ornamental plants with nectar-bearing flowers, the sunbird is widespread and common in settlements throughout much of the southern Levant.

Another species benefiting from the spread of plantations and gardens is the Arabian bulbul *Pycnonotus barbatus capensis*. It feeds mainly on fruits, flowers, and leaves of introduced plants such *as Melia azedarach*, *Lantana sp.*, *Erythrina sp.*, and other fruit trees and flowers. Due to the ample available food it has increased and become a pest in vineyards and orchards.

Man-constructed fish ponds and water reservoirs present an alternative habitat for some wetland species. The night heron *Nycticorax nycticorax*, little egret *Egretta garzetta*, and to a smaller extent squacco heron *Ardeola ralloides* are the main beneficiaries, but glossy ibis *Plegadis falcinellus* also breed near some reservoirs. All four species were rare breeders or nonbreeders in the Levant until the late 1950s, when they started breeding near fish ponds, which provide food and nesting sites. Fish ponds also enabled the pied kingfisher *Ceryle rudis* to extend its breeding distribution from the Jordan Valley to the coastal plain. Dirt roads surrounding fish ponds and reservoirs are a favorite nesting site of the spur-winged plover *Hoplopterus spinosus*, which has already extended its breeding distribution into the northern Negev. Stilts

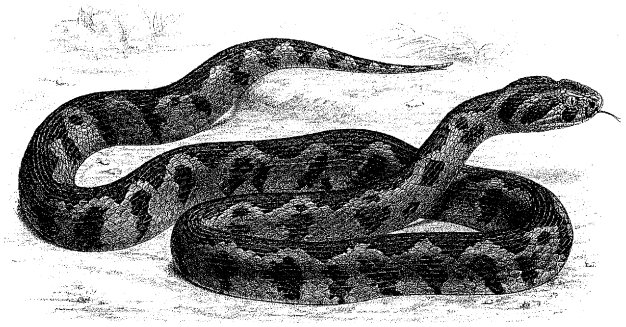


Figure 18 The viper *Vipera palaestinae* (reproduced from Tristram (1884) *Fauna and Flora of Palestine*. London: Committee of the Palestine Exploration Fund).

H. himantopus, which once bred only in the Hula swamp area, are now common breeders along the coastal plain on banks of fish ponds and reservoirs. Also, the little-ringed plover *Charadrius dubius* and the Kentish plover *C. alexandrinus*, which formerly bred mainly on the Mediterranean seashore, now breed on reservoir banks. Another wader that breeds on banks, especially of salt ponds in the coastal plain and in fish ponds with brackish water in the Bet Shean valley, is the avocet *Recurvirostra avocetta*.

Among mammals, the coypu *Myocastor coypu* was introduced into the southern Levant for the fur industry during the 1950s. Some individuals escaped from breeding farms, and others were released when their farming was discontinued. Currently, coypu are common near any water body, from northern Israel to the Gaza Strip. The jungle cat *Felis chaus*, which suffered heavily from the draining of wetlands and from poisoning (aimed at jackals and voles) during the 1960s, has fully recovered and occurs mainly near fish ponds. Also, the Egyptian mongoose *Herpestes ichneumon* has expanded, in numbers and distribution, due to the existence of fish ponds. This increase has affected one of its prey species, the water snake *Natrix tessellata*, so that from a very abundant species (traps in fish ponds yielded 30–40 snakes per pond in several days) it has become rather rare. However, the poisoning of jackals also reduced the number of mongooses. This was followed by an increase in the number of snake bites in Israel by the Palestine viper *Vipera palaestinae* (**Figure 18**) from approximately 200 per year during 1960–1964 to 430 in 1967. Recovery of the mongoose populations was followed by a decrease of snake bite cases to the former level.

In the future, animal diversity in the Levant will be far from the great ecological and zoogeographical variety which makes it so unique.

See also: Desert Ecosystems. Mediterranean-Climate Ecosystems. Near East Ecosystems, Plant Diversity

References

Arad Z (1990) Resistance to desiccation and distribution patterns in bush-dwelling land snails. *J. Zool.* 221: 113.

Ayal Y, Broza M, and Pener M (1999) Geographical distribution and habitat segregation of bushcrickets (Orthoptera: Tettigoniidae) in Israel. *Israel J. Zool.* 45: 49.

Benyamini D (1988) The zoogeography of the butterflies (Lepidoptera, Rhopalocera) of Israel and nearby areas. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Botosaneanu L (1992) *Trichoptera of the Levant*. Jerusalem: Israel Academy of Science and Humanities.

Dimentman Ch, Bromley HJ, and Por FD (1992) *Lake Hula*. Jerusalem: Israel Academy of Sciences and Humanities.

Dumont HJ (1991) *Fauna Palaestina. Insecta 5—Odonata of the Levant*. Jerusalem: Israel Academy of Sciences and Humanities.

Fishelson L (1985) *Orthoptera: Acridoidea*. Jerusalem: Israel Academy of Sciences and Humanities.

Freidberg A (1988) The zoogeography of the Diptera of Israel. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Freidberg A and Kugler J (1989) *Fauna Palaestina. Insecta 4—Diptera: Tephritidae*. Jerusalem: Israel Academy of Sciences and Humanities.

Heller J (1988) The biogeography of the land snails of Israel. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Kadmon R and Heller J (1998) Modelling faunal responses to climatic gradients with GIS: Land snails as a case study. *J. Biogeogr.* 25: 527.

Krupp F (1987) Freshwater ichthyogeography of the Levant. In: Krupp F, Schneider W, and Kinzelbach R (eds.) *Proceedings of the Symposium on the Fauna and Zoogeography of the Middle East, Mainz 1985*. Weisbaden: Ludwig Reichert Verlag.

Kugler J (1988) The zoogeography of social insects in Israel. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Levy G (1985) *Fauna Palaestina. Arachnida 3. Araneae: Thomisidae*. Jerusalem: Israel Academy of Sciences and Humanities.

Levy G (1991) On some new and uncommon spiders from Israel (Araneae). *Bull. Br. Arachnol. Soc.* 8: 227.

Levy G (1998) *Fauna Palaestina. Arachnida 3. Araneae: Theridiidae*. Jerusalem: Israel Academy of Sciences and Humanities.

Levy G and Amitai P (1980) *Scorpiones*. Jerusalem: Israel Academy of Sciences and Humanities.

Levy G and Shulov A (1964) The Solifuga of Israel. *Israel J. Zool.* 13: 102.

Martens K and Ortal R (1999) Diversity and zoogeography of inland-water Ostracoda (Crustacea) in Israel (Levant). *Israel J. Zool.* 45: 159.

Por FD (1975) An outline of the zoogeography of the Levant. *Zool. Scripta* 4: 5.

Shkolnick A (1988) Physiological adaptations to the environment: The Israeli experience. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Tchernov E (1988) The palaeobiogeographical history of the southern Levant. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Theodor O (1980) *Fauna Palaestina. Insecta 2—Diptera: Asilidae*. Jerusalem: Israel Academy of Sciences and Humanities.

Vet V, Polis GA, and Sissom D (1998) Life in sandy deserts: The scorpion model. *J. Arid Environ.* 39: 609.

Werner YL (1987) Ecological zoogeography of the Saharo-Arabian, Saharan and Arabian reptiles in the sand deserts of southern Israel. In: Krupp F, Schneider W, and Kinzelbach R (eds.) *Proceedings of the Symposium on the Fauna and Zoogeography of the Middle East, Mainz 1985*. Weisbaden: Ludwig Reichert Verlag.

Werner YL (1988) Herpetofaunal survey of Israel (1950–1985), with comments on Sinai and Jordan and on zoogeographical herpetogenety. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Yom-Tov Y (1988) The zoogeography of the birds and mammals of Israel. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Yom-Tov Y and Mendelssohn H (1988) Changes in the distribution and abundance of vertebrates in Israel during the 20th century. In: Yom-Tov Y and Tchernov E (eds.) *The Zoogeography of Israel*. Dordrecht: Junk.

Near East Ecosystems, Plant Diversity

Avinoam Danin, The Hebrew University of Jerusalem, Jerusalem, Israel

© 2013 Elsevier Inc. All rights reserved.

Glossary

Batha A biblical term for semishrub communities, mainly of seral vegetation of old field succession in the Mediterranean part of the Near East.

Contracted vegetation Vegetation restricted to wadis that receive additional water supply.

Diffused vegetation Vegetation occupying all slopes and most habitats.

Semisteppe batha Semishrub communities developing on most soil types at the boundary of the Mediterranean zone of the Near East.

Wadi Dry water course.

Environmental Conditions

Currently, the most important factors affecting the distribution of species are environmental conditions. The evolutionary history of the area reviewed influenced the composition of the flora, but the environment influences the composition and their ability to coexist and prosper. Before discussing the vegetation and its gradients of species diversity, the physical setup of the study area is discussed.

Topography

The topography of Israel and Jordan can best be described as north–south topographic belts, which are influenced by the geomorphological features of the area. The Mediterranean coastal plain is narrow in the north of Israel (**Figure 1**, 1) and becomes wide in the south (**Figure 1**, 5). Low hills with gentle topography and wide valleys with deep soil constitute the foothills. The mountainous area reaches elevations of approximately 1000 m at the Judean Mountains (**Figure 1**, 12) and the Negev Highlands (**Figure 1**, 15) and 1200 m at the Galilee (**Figure 1**, 7). The Jordan–Dead Sea–Arava rift valley is part of the Afro-Syrian Rift Valley and constitutes the lowest topographic element in the area. This valley is 400 m below sea level in the deepest part of the Dead Sea (**Figure 1**, 24). The terrain ascends to 200 m above sea level (m a.s.l.) at the foot of Mount Hermon (**Figure 1**, 26) in the north and at the water divide of the Arava Valley (**Figure 1**, 25) to the south. A few transversal valleys, such as the Esderaelon Plain (**Figure 1**, 9) running from east to west, dissect the mountainous area into large blocks. Steep escarpments, dissected by canyons, typify the areas east and west of the rift valley. There are a few internal lakes in this valley. Lake Hula between the Golan and Galilee was drained in the 1950s. The Kinneret, also known as the Sea of Galilee, is the largest freshwater natural reservoir of Israel. The Dead Sea is a highly salty water body. The large salt marshes at the Arava Valley are in fact a large underground lake. At the vicinity of these water bodies there are small or large springs. Mount Hermon, at the common border of Israel, Lebanon, and Syria, reaches an elevation of 2800 m and terminates sharply at the basalt-covered plateau of the Golan to the south. This plateau descends gently from an elevation of 1200 m toward the Yarmouch River and toward the Kinneret at 200 m below sea level. The Jordanian plateau starts south of the Yarmouch River (**Figure 1**, 28). Its escarpments delimiting the rift valley are as steep as, and in many places steeper than, the escarpments in Israel. The Jordanian plateau has several peaks 1200–1600 m high. A few deep rivers dissect the plateau from east to west toward the rift valley. East of the water divide, which is near the western edge of the plateau, the landscape becomes flat and gradually descends toward Mesopotamia.

Sinai has a wide coastal area along the Mediterranean Sea (**Figure 1**, 32), ascending gradually southward. The northern Sinai sand belt reaches the anticlinal ridges of Gebel Maghara, Gebel Halal, and Gebel Yiallaq (**Figure 1**, 33). The central Sinai gravel plains (**Figure 1**, 35) are surrounded by crescent-like ranges of mountains or plateaus (**Figure 1**, 36), dissected by a few valley passages such as those of Wadi Giddi, Wadi Mitlah, and Wadi Sidr. South of the gravelly plains there are two large plateaus ascending gradually in a north–south direction. The Gebel el Igma plateau (**Figure 1**, 37) peaks at 1600 m and the Gebel et Tih plateau peaks at 1400 m. The erosion escarpments of the two plateaus are steep and that of the latter descends to the sandstone belt. The morphology of the sandstone belt (**Figure 1**, 41) is highly diverse in different parts of Sinai. The Southern Sinai Massif (**Figure 1**, 42 and 43), built of magmatic and metamorphic rocks, is highly dissected topographically. The mountain peaks reach 2500–2600 m. The escarpments toward the Gulf of Elat-Aqaba are not as steep as those of the rift valley in Israel and Jordan. The coastal plain along this gulf is wide in the southern part of Sinai, whereas in many places the mountain slopes descend directly into the sea without any coastal plain. There is a wide coastal plain along the Gulf of Suez.

Rock Types, Geomorphology, and Edaphic Conditions

The most common rock types of northern Israel and Jordan are sedimentary limestones, dolomites, chalks, and marls of the Cretaceous and the Tertiary. Terra Rossa soils occur in the mountainous areas on hard rocks, and Rendzinas occur on the soft ones. Basalt rocks typify much of the Golan plateau and northeast Galilee: Both are covered with basaltic brown or red Mediterranean soils. Tertiary and Pleistocene rocks and

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00195-7

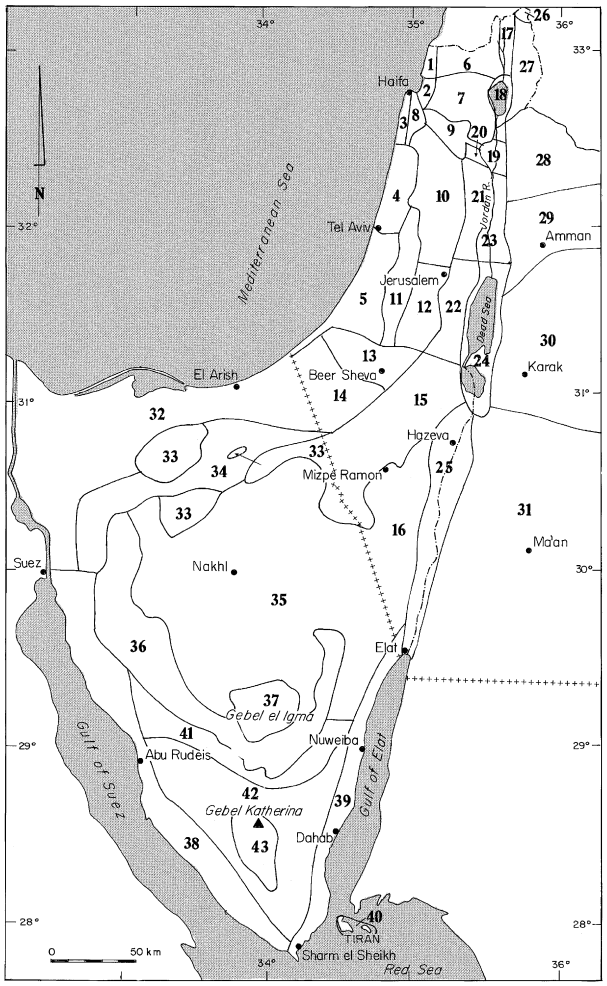


Figure 1 Geographical subdivision of the study area. Israel: 1, coastal Galilee; 2, Acco Plain; 3, coast of Carmel; 4, Sharon Plain; 5, Philistean Plain; 6, Upper Galilee; 7, Lower Galilee; 8, Mt. Carmel; 9, Esdraelon Plain; 10, Samaria; 11, Shefela; 12, Judean Mountains; 13, northern Negev; 14, western Negev; 15, Negev Highlands; 16, southern Negev; 17, Hula Plain; 18, Kinnroth Valley; 19, Beit Shean Valley; 20, Mt. Gilboa; 21, Samarian Desert; 22, Judean Desert; 23, Lower Jordan Valley; 24, Dead Sea Valley; 25, Arava Valley; 26, Mt. Hermon; 27, Golan. Jordan: 28, Gilead; 29, Ammon; 30, Moav; 31, Edom. Sinai: 32, Mediterranean sands and salt marshes; 33, anticlines of northern Sinai; 34, gravelly plains covered with shifting sands; 35, gravelly plains of central Sinai; 36, Table mountains of central and western Sinai; 37, Gebel el 'Igma; 38, coastal plain of the Gulf of Suez; 39, coastal plain and foothills of the Gulf of Elat; 40, Tiran and Sinafir Islands; 41, sandstone belt; 42, Lower Sinai Massif; 43, Upper Sinai Massif.

derived soils fill up the small and large valleys. Calcareous sandstone, locally known as kurkar, and sandy–loamy soils, known as hamra, typify the coastal plain near the coast and deep clay soils, known as grumusols or vertisoils, far from the Mediterranean coast. Cretaceous and older sandstone typify the vicinity of the rift valley of the Jordan River, the Dead Sea, and Arava in Jordan. The edaphic conditions in northern Jordan and south to Amman are similar to those of the area west of the Jordan River.

The steppe and desert areas of Israel, Jordan, and Sinai develop on a much more diverse assemblage of rocks. The main contributors are limestone, chalk, marl, chert, sandstone, magmatic, metamorphic rocks, and gravel of alluvial origin or rocks weathered *in situ*. Loess is important eolian sediment, composed mainly of silt and clay particles, that much influences the soils of the transition zone of the Mediterranean territory and the steppelands. Large flatlands south of this transition zone occur in the northern Negev of Israel and east of Irbid, Jarash, and Amman in Jordan. The poor moisture regime of loessial soils with low quantities of rainfall makes the desert boundary prominent.

Dan *et al*. (1975) provide soil maps of Israel, and Al-Eisawi (1996) provides a map of Jordan; a detailed soil map of Sinai is not available. The specific influence of each substratum type on the vegetation is discussed in detail elsewhere (Danin, 1983). The relationships between rock type, soil, and vegetation in the desert areas of Israel and Sinai are similar to those in many areas of Jordan.

Geomorphological structures in Israel are small, and they are larger in Sinai and Jordan; however, these are smaller than those in countries such as Saudi Arabia in which huge areas of similar structures occur. This is the expression of relatively high geomorphological diversity. One might expect to find a high diversity of rock and soil types under a certain climatic regime. Hence, habitat diversity in such areas is expected to be high.

Climate

Rainfall

The climate of Israel, Jordan, and Sinai is Mediterranean, characterized by a cold and rainy winter and a rainless and warm summer. Rainfall quantities vary in two directions: rainfall decreases gradually from north to south – sharply from the water divide eastward in Israel and gradually in Jordan from the water divide eastward (**Figure 2**). The north–south gradient is influenced mainly by the intensity and frequency of rain-contributing systems, whereas the west–east gradients are influenced by the topography. Thus, the ascent of air bodies from the Mediterranean Sea toward the mountains of Judea, Samaria, and Galilee results in cooling of the air and an increase in rainfall with increasing elevation. Air bodies descending toward the rift valley become warmer and drier; hence, the Samarian and Judean Deserts and the eastern Galilee occur at the rain shadow of the Judean, Samarian, and Galilee mountains, respectively. The ascent to the Jordanian plateau mirrors the rainfall map with maximum at the mountaintops. There is a gradual decrease in mean annual rainfall eastward with increasing continentality far from the Mediterranean toward the desert. The increase in mean annual rainfall in Moav and Edom, located south of the west–east section of the 300-mm isohyet in Israel, is related to the orographic influence of the high elevation of this area. Snow may fall in winters of cold and wet years on the highest summits of mountains in Israel, Jordan, and Sinai and remain for a few days.

Dew and Fog

Dew and fog are important sources of humidity for the poikilohydric organisms growing on rock outcrops. The measurement of dew by Duvdevani's dew gage indicated a high

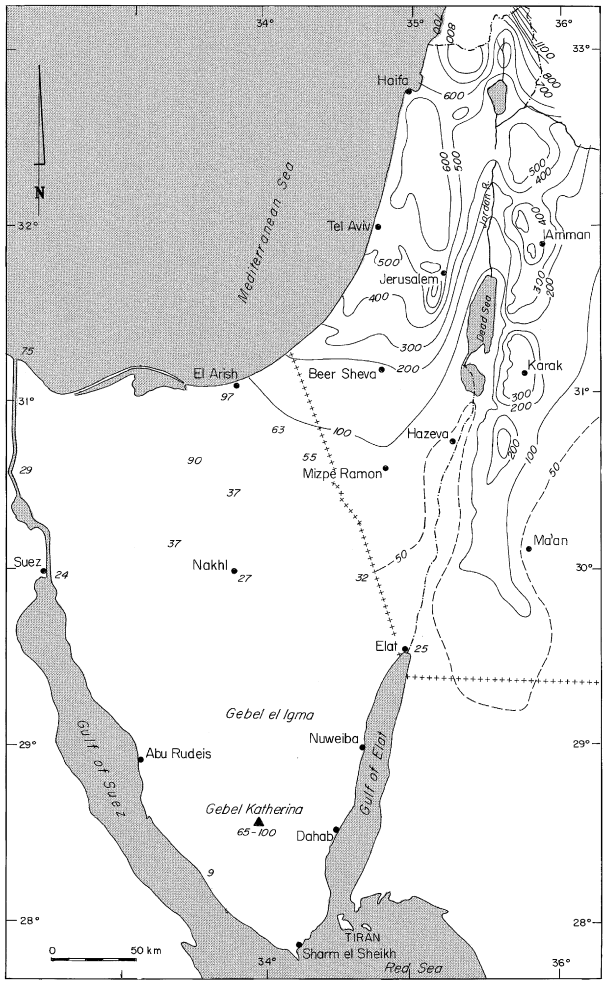


Figure 2 Mean annual rainfall map.

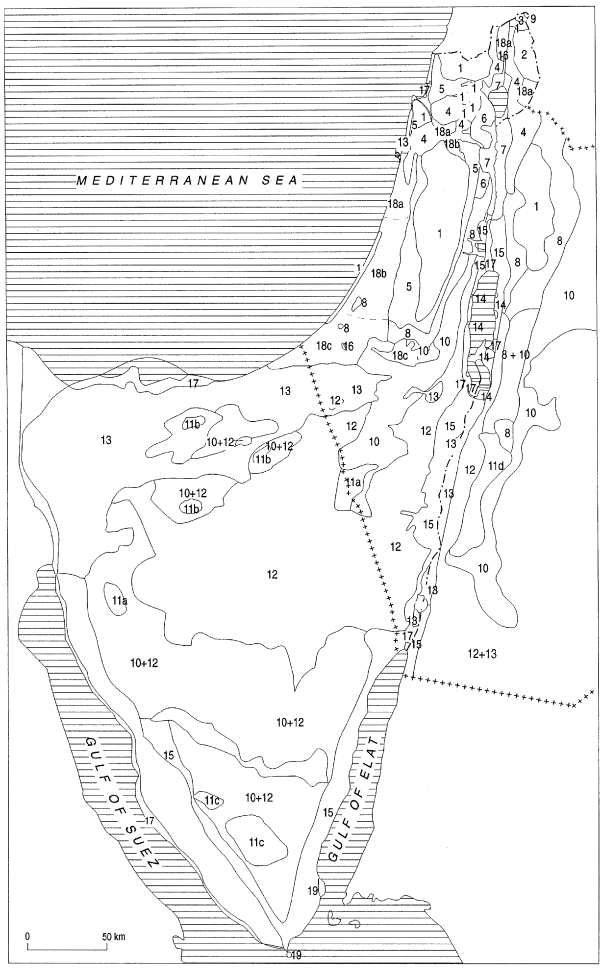


Figure 3 Vegetation map of Israel, Sinai, and western Jordan: 1, maquis and forests; 2, *Quercus calliprinos* woodlands on basalt; 3, montane forest of Mt. Hermon; 4, open forests of *Quercus ithaburensis*; 5, open forests of Ceratonia siliqua and Pistacia lentiscus; 6, Ziziphus lotus with herbaceous vegetation; 7, Mediterranean savannoid vegetation dominated by Ziziphus spina-christi; 8, Semisteppe batha; 9, tragacanth vegetation; 10, shrub–steppes; 11, shrub–steppes with trees of (a) Pistacia atlantica, (b) Juniperus phoenicea, (c) Pistacia khinjuk (Sinai), (d) *Q. calliprinos* and trees of 11a, 11b, and 11c (Jordan); 12, desert vegetation; 13, sand vegetation; 14, oases with Sudanian trees; 15, desert savannoid vegetation; 16, swamps and reed thickets; 17, wet salinas; 18, synanthropic vegetation; 19, mangroves.

similarity to events of efficient dew when lichens imbibed on stones and rocks at the Negev Highlands. The mean annual number of nights with more than 0.02 mm per night of dew has been 191 ± 22 for the past 15 years at Avdat; of these, there are only 124 ± 28 nights annually with dew amounts of 0.11–0.5 mm. Dew measurements in other parts of Israel, Jordan, and Sinai are not available. However, lithobiont communities near Avdat were used to extrapolate the regional distribution of dew. There are many similarities in weathering features found at the top of the northern Sinai anticlines (**Figure 1**, 33) and on limestone outcrops at the area marked 11d in Jordan (**Figure 3**). The author assumes that the dew and fog regime is similar in the latter two areas to that of the Avdat area where real measurements of dew and rainfall were carried out.

Temperature

Temperature regime in the study area is influenced by the altitudinal and latitudinal position of the site under review. Mean annual temperature in the desert areas varies from 9 to 25 °C. The lowest temperatures prevail in the peaks of southern Sinai and Jordan near Shoubak, whereas the highest are those of the Dead Sea and Arava Valleys, where elevation reaches 400 m below sea level. Mean annual temperature maps of Israel are presented in the atlas of Israel and of Jordan in Al-Eisawi (1996).

Flora

There are 2682 plant species in Israel, including approximately 200 species which occur in the Mt. Hermon area and are absent in the other districts of Israel. The number of species in Jordan is 2078, according to Al-Eisawi (1996). There are 2885 species in the combined lists of plants of Israel and Jordan

(A. Danin, unpublished data). The vast species richness of Israel, expressed as the parameter of species to area (**Table 1**), is related mainly to the following factors:

1. Its position in a meeting zone between plant geographical regions, each with its own typical flora.
2. The existence of many habitats needed to support these species: The wealth of habitats derives from the climatic transition between the relatively moist area in the northern part of the two countries and the extreme desert areas in their southern part and also in Jordan's eastern part. Topography is also a factor in creating the warm climates of the rift valleys and the relatively cold climate of the mountainous areas. Similarly, other highlands and lowlands have local climatic influence, which increases the habitat diversity of Israel and Jordan and increases the number of habitats which support much plant species diversity. The large number of rock types influences the development of many soil types in a small area, increasing the diversity of habitats available for plants.
3. A long history of human pressure of cultivation and grazing by domestic animals led to a strong stress on the existing flora and enabled the introduction of many alien species, many of which occupy habitats created by human activity.

According to the various studies of the flora of Israel, it may be divided into the following groups:

1. Mediterranean species, which are distributed around the Mediterranean Sea.
2. Irano-Turanian species, which also inhabit Asian steppes of the Syrian desert, Iran, Anatolia in Turkey, and the Gobi desert.
3. Saharo-Arabian species, which also grow in the Sahara, Sinai, and Arabian deserts.
4. Sudano-Zambesian species, typical of the subtropical savannas of Africa.
5. Euro-Siberian species, also known in countries with a wetter and cooler climate than that of Israel. These species grow mainly in wet habitats and along the Mediterranean coasts.
6. Biregional, triregional, and multiregional species that grow in more than one of the regions mentioned previously.
7. Alien species derived from remote countries; these plants propagate without human assistance.

Table 1 Species richness of the flora of several countries

Country/State	*Species*	*Area (km^2)*	*Species per 100 km^2*
Israel	2682	29,600	9.06
Sinai	889	61,100	1.45
Egypt (excluding Sinai)	2100	1,000,000	0.21
Saudi Arabia	2200	2,200,000	0.1
Syria	3060	185,000	1.65
Spain	5500	504,000	1.09
Tunisia	2250	164,000	1.37
Greece	4200	132,562	3.17
Italy	5600	301,100	1.86
Britain	1666	229,850	0.72
California	5057	411,000	1.23

Source: Reproduced from Danin A (1996) Vegetation of Israel and Sinai. *Botanicheskii Zhurnal* 81(11): 14–31, and Le Houérou HN (1995) Bioclimatologie et biogéographie des steppes arides du Nord del'Afrique; Diversité biologique, dévelopement durable et désertisation, OPTIONS méditerranéennes, serie b, CIHEAM (Montpellier) and ACCT (Paris), France (tudes et recherches No. 10).

Vegetation

Danin (1999) discusses the vegetation of the entire area in detail. The most prominent factors influencing the vast species diversity of the study area are discussed in the following sections. The reader may combine the general information on the vegetation of Israel (Danin, 1999) with that of Sinai (Danin, 1983) and Jordan (Al-Eisawi, 1996).

Maquis and Forests

The principal spontaneous woodland areas of Israel are found in the mountains of Judea, Carmel, and Galilee and at the foot of Mt. Hermon; those of Jordan occur north of Amman (**Figure 3**, 1). These forests or maquis are dominated by the sclerophyllous evergreen *Quercus calliprinos* and the deciduous *Pistacia palaestina* and grow on hard limestone with terra rossa soil. The companions of *Q. calliprinos* vary according to edaphic and climatic conditions. In the mesophytic aspect of the oak woodlands, found mainly in Upper Galilee, there are several trees, shrubs, vines, and geophytes that are not found in the xerophytic aspect of the southern Judean Mountains. In the driest maquis stands, *Rhamnus lycioides* subsp. *graecus* is the only arboreal companion of *Q. calliprinos*. Typical vines in these maquis are *Rubia tenuifolia*, *Lonicera etrusca*, *Asparagus aphyllus*, and *Ephedra foeminea*.

Marly chalk is a common rock type that has high moisture-holding capacity and is covered with a shallow light rendzina soil. The aeration of the rhizosphere, of trees and shrubs that penetrate into the soft rock, is poor and leads to nutritional stresses, which only specially adapted plants can withstand. Much of the nitrogen in these soils is in the form of ammonium ions, whereas in the terra rossa it is in the nitrate ion form. The vegetation cover of the light rendzinas on marly chalk is found poor when compared with that on terra rossa and includes very few annual species. The principal trees of these soils are *Arbutus andrachne* and *Pinus halepensis*.

In most of the area, cultivated plants have replaced the spontaneous trees. A few thousand years ago, people in the eastern Mediterranean countries started to clear the natural vegetation to create agricultural land. Trees such as olives (*Olea europaea*) and almonds (*Amygdalus communis*), which grow spontaneously in the area, have been domesticated. Timber derived from the forests and maquis was used for the construction of houses, for agricultural tools, and for fire fuel. For the past few millennia shepherds have burned large woodland areas to open paths for domestic animals, and the pasture quality has been improved through the replacement of trees and shrubs with palatable herbaceous plants.

As a result of thousands of years of deforestation and agricultural and urban development, large areas of Israel and Jordan look like mosaics of seral communities. Abandoned

cultivated ground becomes populated for dozens of years by herbaceous and low lignified plants. Habitat diversity, and thus species diversity, is extremely high due to the removal of the shading trees and the opening of poor and temporary fertile soils to annual plants which could not grow in the closed woodlands. Frequent human-induced fires reopen areas, which become populated with shading trees and shrubs and ephemeral plants that increase species diversity. Some terra rossa soils, which are leached and poor, become fertilized during the years of crop cultivation. Once abandoned, a series of replacement annual plant communities occurs, from those using fertile ground, including tall plants, to those using poor soil and being small in size. By sampling approximately five such stages, each a square of 1×1 m^2, in the vicinity of Jerusalem, the author found approximately 100 different species in 8 m^2 . Local enrichment of the poor soils by gazelle droppings and at the vicinity of nests of harvesting ants creates an ecotone of soil fertility, which enables the coexistence of many species.

Light rendzinas were often cultivated due to the high water-holding capacity of the substratum of soil plus rock. When abandoned, this soil may support the growth of a few annual species. After time passes and the aerated plowed soil becomes flattened, microbiotic crust, including cyanobacteria, lichens, and mosses, develops on the surface, thus influencing soil aeration and changing edaphic conditions. Such minute edaphic changes along the precipitation gradients of the mountainous areas (**Figure 2**) contribute to the vast species diversity of the country.

Q. calliprinos Woodlands on Basalt

The basalt flows of the northern Golan are younger and differ in rock types from those of the rest of the Golan. The vegetation constituting these woodlands (**Figure 3**, 2) is rich in herbaceous companions. In contrary to the vegetation of the calcareous rocks, they are almost devoid of semishrub communities at the early succession stages because of destruction and abandonment. The main environmental factor influencing this is the high phosphorous content of the rock and soil, which favors the development of ephemeral plants. Thus, a dense maquis of *Q. calliprinos* covers the gentle north-facing slope of the ancient volcanic cone of Har Odem, the Golan, at an elevation of 900–1000 m. Regarding the trees, the rich ephemeral vegetation includes approximately 20 species of Trifolium, many other Papilionaceae, and rich annual flora of other families. The phytomass is low, thus indicating some nutrient deficiency when compared to that of the vegetation of other basalt flows. Under such conditions, without one dominant species there is high species diversity.

Other basalt rocks derived from older basalt flows at lower elevation, down to 500 m a.s.l., support denser and taller ephemeral vegetation with different composition due to differences in soil fertility and climatic conditions, which follow elevation. The oak trees and their arboreal companions are found only occasionally and they have no impact as shadow creators in most of the Golan and eastern Galilee.

Ancient volcanoes provided much volcanic ash to large areas surrounding them. Because of a different pedogenetic nature, this substratum supports a rich flora, many components of which are not common in other kinds of soil in this area. Many of the volcanic ash layers have a porous and well-aerated texture. In contrast to the grumusols and protogrumusols of much of the area surrounding it, this substratum supports many narrow-range species.

Montane Forest of Mt. Hermon

The montane forest (**Figure 3**, 3) stretches at Mt. Hermon from 1300 to 1700 m a.s.l., where there are only scattered remnants of the woodlands which are believed to have been there before their destruction by human activity. The dominants are deciduous trees such as *Quercus boissieri* and *Quercus libani* and several species of *Crataegus, Amygdalus, Acer*, and *Prunus*. Their companions are mainly perennial grasses and many herbaceous species that differ from those of the rest of the Mediterranean territories of the Near East. These occur at lower elevation and presumably thrive on a different temperature and rainfall regime than that of the montane forest. There are no representative areas of this category in Jordan because there are no mountains of this elevation in the northern part of the country. However, in southern Jordan, there are such high mountains, but due to their drier climate woodlands of this kind are absent. Nonetheless, a few representative shrubs of the genera *Astragalus* and *Cotoneaster*, common at Mt. Hermon, occur at high elevations in crevices of smooth-faced rocks in southern Sinai and in southwestern Jordan.

Open Forests of *Quercus ithaburensis*

The tabor oak is the most frequent tree in the open forests prevailing in large areas of Israel and Jordan (**Figure 3**, 4). *Styrax officinalis* also appears with the tabor oak when developed on chalky ground of the Lower Galilee, Israel, and the Gilead in Jordan. *Pistacia atlantica* is the companion on basalt rocks of the Golan. Plenty of herbaceous plants cover the open space among the trees in all areas of this category. The few semishrubs found in this community when developing on chalky ground are mainly of *Majorana syriaca*. This is contrary to the dominance of *Sarcopoterium spinosum*, the typical component of seral communities on terra rossa. This type of forest, which once dominated the Sharon Plain, is restricted in the Sharon (**Figure 3**, 4) to solitary, sporadic tabor oak trees that survived urbanization and agricultural development. There are large woodlands of the tabor oak in the Lower Galilee and in the Golan, lower than 500 m. Some of the largest indigenous trees in Israel are those of tabor oak at the Hula and Dan Valleys in the northern section of the rift valley in Israel. The tabor oak woodlands of the northwest of Jordan are well preserved. They cover large areas of rocky terrain west of Irbid, at the western escarpments of the Jordanian plateau, from sea level to 500 m a.s.l. The drier and more continental climate of the Jordanian woodlands allows for a different assemblage of companions to the oak.

Open Forests of *C. siliqua* and *Pistacia lentiscus*

Open woodlands dominated by carob trees and shrubs of *P. lentiscus* cover the lower altitudinal belt of the main

mountain ranges, 0–300 m on both sides of the central mountain range on terra rossa soils. This community inhabits rendzina soils at the foothills of the Judean mountains and sandy soils of the Sharon Plain in the littoral aspect of the community (**Figure 3**, 5). Generally, the community is more drought and heat resistant than the communities dominated by *Q. calliprinos* and has a similar position in the aridity sequence of communities as those dominated by the tabor oak. Wild olive (*O. europaea* var. *sylvestris*), which resembles the cultivated olive but has much smaller fruits, is an important companion of the carob in rocky sites at Mt. Carmel and Galilee. In Jordan, this category is almost missing; however, scattered carob trees occur in the open woodlands of *Q. ithaburensis* of the Gilead, mainly in the transition zone from the belt dominated by the tabor oak to that of *Q. calliprinos* at elevations of 500–600 m. The main companion of the carob at the southern Judean foothills is *R. lycioides* subsp. *graecus*. The latter totally replaces all the other arboreal components in dry habitats.

The flora of this community is rich due to high habitat and microhabitat diversity. The hard limestone rocks have many crevices and soil pockets, which enable diverse microhabitats supporting different species to coexist that are isolated from each other by the rock body. Soft chalk rocks, covered by hard nari crust, also known as caliche, constitute additional diverse habitats. The nari layers have an upper hard crust layer of 5–10 mm, which covers softer rock but is still consolidated chalk. Trees and shrubs, the roots of which penetrate into the rock, may establish themselves, thus avoiding the competition of the rich herbaceous flora which accompany them in the deep soil pockets. Thus, in many places the local boundary of the land dominated by shrubs and trees coincides with that of the nari rocks. Various kinds of shade are provided by evergreen trees and shrubs, each having a different architecture that influences the companion life. Opened for grazing by domestic and wild animals, the potential strong competition ordained by a few herbaceous plants is decreased, thus enabling many species to coexist.

Ziziphus lotus with Herbaceous Vegetation

The phosphorus-rich soils developing on basalt rocks of southeastern Galilee and the slopes of the Golan near the Sea of Galilee and limestone and other sedimentary rocks down to the Samarian desert support grasslands dominated by Gramineae with large seeds. The most common grasses are wild wheat (*Triticum dicoccoides*), wild barley (*Hordeum spontaneum*), and wild oats (*Avena sterilis*). This area (**Figure 3**, 6) contains the highest genetic diversity of these species, which are believed to have been domesticated in the Near East thousands of years ago (Zohary and Hopf, 1993). A similar formation covers the west-facing slopes of the Gilead, Jordan, below the tabor oak belt. In drier and warmer sites or sites at which the soil is shallow, *Stipa capensis* is the dominant plant. The lignified and thorny shrub *Z. lotus*, spread as green patches over the area, typifies most areas of these grasslands. Sites of relatively high soil fertility, common in this plant community, are the nests of the harvesting ant *Messor messor* and piles of dung of the male gazelles (*Gazella gazella*). This category is presented as a mosaic with that of the savannoid Mediterranean vegetation (**Figure 3**, 7) in the steep topography at the vicinity of the rift valley north of Jericho to north of the Sea of Galilee.

Mediterranean Savannoid Vegetation

Warm, stony/rocky slopes of the Galilee, Golan, Gilead, and Samaria descending to the rift valley (**Figure 3**, 7) at sea level and below are covered with grasslands of Mediterranean annuals with large seeds, such as wild wheat, barley, and oats, and scattered *Ziziphus spina-christi* trees. The latter is a low, thorny tree with edible fruits that also grows in true savannas in Africa, where its companions are Sudanian perennial grasses. Therefore, the vegetation here is named "savannoid." Stands of this tree are also established along the transversal valleys on deep clay soils and along the Sharon Plain, where it grows on sandy–loamy soils together with the grass *Desmostachya bipinnata*. Local diversity of nutrients and derivative high species diversity are caused by harvesting ants and gazelles as discussed previously (*see* Maquis and Forests).

At the Bet-Shean Valley south of the Sea of Galilee there are considerable areas of slightly saline springs, the salts of which remain in the soil while water evaporation occurs. This results in a wide range of microhabitats, which vary in soil salinity and the availability of spring water. The northernmost site of *Balanites aegyptiaca*, an African savannah tree, is in the vicinity of one of these springs. The Sudanian mistletoe *Plicosepalus acaciae* (=*Loranthus acaciae*) has its northernmost site on these Balanites trees.

Semisteppe Batha

Semishrub communities of the *Ballotetea undulatae* at the boundary of the Mediterranean zone (**Figure 3**, 8), where mean annual rainfall is 300–400 mm, are regarded as semisteppe bathas. There are several communities dominated by Mediterranean plants such as *S. spinosum*, which dominates bathas in more mesic parts of the country. Others, such as *Artemisia sieberi* and *Noaea mucronata*, dominate steppe areas in the Negev, Sinai, Jordan, and eastward to Afghanistan. Many plants that play an important role in the seral communities in fallow fields at the center of the Mediterranean region grow here in what is regarded as their primary habitats. No anthropogenic disturbance assists or enables their growth here, and they occupy their typical habitats that are related to certain edaphic and climatic conditions. *S. spinosum*, which becomes ethiolant and dies under the shade of trees and shrubs in the course of plant succession in maquis and forests, has no such competitors in the semisteppe bathas. The semisteppe bathas extend further south and east of the regional boundary of the maquis in Jordan and in Israel. Because semishrubs are the dominant tallest growth form, the semisteppe bathas are rich in their flora. Many plants of mesic and xeric origin coexist here, contributing to the vast species diversity. Phytogeographical analysis of each association in this vegetation indicates a higher diversity in origin when compared to the associations of drier and moister conditions.

Tragacanth Vegetation of Mt. Hermon

Semishrubs that look like thorny cushions constitute the most prominent formation of vegetation developed on the windward slopes of the peaks of Mt. Hermon above 1900 m (**Figure 3**, 9). Many species of section *Tragacantha* in the genus *Astragalus* and of the genus *Acantholimon* are components of the cushion plants formation throughout the Middle East. The thorny cushion seems to have some biological advantages, which make it adaptable to some of the components of the harsh conditions of this habitat: cold winter with high-velocity winds, precipitation mainly as snow, and dry summer. The short growth season and the harsh environment are the main factors enabling the success of very few annual species as companions of the cushion plants.

Snow accumulates on the slopes in the wind shadow of small local ridges or crests because of the low wind velocity. Consequently, snow may reach a depth of 10 m for a few months. Water drainage from the slopes of this area is through karst systems and not through wadis as in many areas of the country. A common feature of the karst topography is the occurrence of small valleys, which are filled with fine-grained soil known as dolinas. The soil in these valleys is waterlogged for a long time, thus increasing habitat diversity and the growth of many additional species compared to those of the rest of Israel.

Several species of tragacanth *Astragalus* occur in areas which are far from Mt. Hermon and the Anti Lebanon mountains. *Astragalus bethlehemiticus* is a typical companion of steppes and rock vegetation in the shrub–steppes and semi-steppe bathas of the Negev Highlands and of southwestern Jordan; *Astragalus echinus* is a representative of this group at the high elevations of southern Sinai, where it is confined to crevices of smooth-faced granite.

Vegetation Patterns in the Dry Areas of Israel, Jordan, and Sinai

When dealing with the vegetation of drylands, which receive less than 200 mm mean annual rainfall, one has to consider general patterns of distribution of vegetation. The two principal patterns of distribution of vegetation in the dry areas of the Near East are related to dry watercourses, which are known as wadis from the Arabic language or arroyos in the southwestern US. Wherever plants grow on slopes and depressions, the pattern is regarded as "diffused." In extremely dry desert areas, where vegetation is restricted to wadis that receive additional water supply, the pattern is "contracted." In an area in which the climatic conditions enable the development of diffused vegetation on most soil types, silty or clayish soils, which are relatively dry soils in desert will locally support contracted vegetation. Plants demanding relatively high quantities of moisture may inhabit special habitats of wadis in zones of contracted vegetation. Such special habitats are derived from local contribution of runoff water from rocks to their crevices and soil pockets. In hard and fissured limestone, dolomite, granite, and metamorphic rocks, much of the rainfall is available to the semishrubs growing there because water infiltration is good and the soil in the rock fissures is leached. Accumulating deep in the soil and the weathered rock, this water is protected from direct evaporation by rocks and stones. Trees that grow on slopes of the Mediterranean zone that receives 500–700 mm of annual rainfall occur in desert areas with 100 mm or even less in proximity to outcrops of smooth-faced hard rocks. These rocks do not absorb much water and their crevices receive high amounts of water through runoff (Danin, 1999).

One of the ways in which plants are adapted to desert conditions is that they avoid the extremely dry season, which may last from 6 months to several years. Thus, annual plants may be seen in some habitats every few years. This depends on edaphic and climatic conditions; for example, annual halophytes may germinate and grow to produce seeds only on salty soils and only in rainy years when the soil solution is diluted to the appropriate level. However, annuals that develop in crevices of smooth-faced rocks may have sufficient resources to be sustained almost every year. Thus, species diversity in the desert may be high in areas in which habitat diversity is high, especially if there are sufficiently large outcrops of smooth-faced hard rocks, which contain many different microhabitats.

Shrub–Steppes

The area of the Negev Highlands, the Judean Desert, Sinai, and southwestern Jordan, which receives 80–250 mm of rainfall annually, is covered with semishrubs at a diffused pattern, thus forming shrub–steppes (**Figure 3**, 10). The most common dominants in these steppes are *A. sieberi*, *Noaea mucronata*, and *Gymnocarpos decander*. The phytomass produced by annuals in the plant communities developing on stony or rocky shallow soils is always small when compared with that of fine-grained and deep soils. The latter types hold much of their water close to the soil surface, thus losing much of it through direct evaporation. The minute quantities of 8 ppm of salts carried by clouds and later by rain from the Mediterranean Sea by climate systems remain in the soil and accumulate there. The soil may be too dry or too saline for the growth of annuals in regular or dry years. Nearly monospecific communities of semishrubs exist, each best adapted to the specific local saline conditions. The most common dominants under these conditions are *Reaumuria hirtella*, *Reaumuria negevensis*, *Salsola vermiculata*, *Bassia* (*Chenolea*) *arabica*, and *Atriplex glauca* on chalk- and marl-derived soils; *Anabasis syriaca* and *Haloxylon scoparium* are the shrubby dominants on loess-derived soils. However, in moist years there is much development of annuals, although in patches with high salinity there are monospecific patches of salt-resistant annuals. Showy geophytes such as species of *Tulipa*, *Iris*, *Ixiolirion*, *Ranunculus*, and *Anemone* may bloom in large quantities in the shrub–steppes in moist years.

Outcrops of smooth-faced hard limestone support plants that differ much from those on the other soil types. This vegetation is typically characterized by *Chiliadenus iphionoides*, *C. montanus*, *Globularia arabica*, *Stachys aegyptiaca*, *Polygala negevensis*, *Tanacetum sinaicum*, and *Capparis aegyptia*. Isolated populations of dozens of Mediterranean relicts and many rare desert plants are found in this habitat in the Negev, the Judean Desert, Sinai, and Jordan. The semishrub *S. spinosum* and the geophytes *Narcissus tazetta* and *Sternbergia clusiana* are examples of this phenomenon in the Negev (Danin, 1983).

Several species that were discovered as endemic and new to science and many species that were not previously collected in the steppe areas of the Near East are confined to this habitat (Danin, 1999, 2008; Danin and Künne, 1996).

Shrubs of *Retama raetam* and *Achillea fragrantissima* are the dominants in wadis of the terrain of hard limestones. *Atriplex halimus* prevails in this habitat, in which the catchment area is built up from the salty soils on chalk, clay, or marl. At lower elevations, *Acacia raddiana*, *Acacia pachyceras*, *Tamarix nilotica*, and *Tamarix aphylla* also occur. Many small springs exist in the limestone hills of western Sinai and the sandstone hills of southwestern Jordan. Most of these springs can be detected from afar by the date palm (*Phoenix dactylifera*), which is confined to sites with a high water table of freshwater. It is accompanied by *Nitraria retusa*, *Juncus arabicus*, *Phragmites australis*, and *Cressa cretica*. Canyons occur in many wadis and may have long-lasting water pools supplied by floods and support rich flora of hydrophytes such as *Zannichellia palustris* and *Potamogeton* spp. and green algae such as *Chara* spp.

Shrub–Steppes with Trees

Shrub–steppes similar to those discussed previously (*see* Shrub–Steppes) cover most of the area of this category (**Figure 3**, 11); however, prominent trees occur in special habitats sporadically. Approximately 1400 adult trees, most of which are *P. atlantica*, accompanied by a few *Amygdalus ramonensis* and *Rhamnus disperma* trees and shrubs (Danin, 1983), typify subunit 11a in the Negev Highlands and eastern Sinai. The trees are found in affinity to outcrops of smooth-faced rocks. The high yields of runoff water from the rocky outcrops to their crevices enable the successful germination and establishment of seedlings even outside of wadis. Sufficiently large outcrops of limestone support trees, some of which are several hundred years old. Wadis with such outcrops in their catchment area support large trees. Rare relicts, such as the vines of the Mediterranean maquis *Prasium majus* and *E. foeminea*, and endemic plants such as *Origanum ramonense*, *Bufonia ramonensis*, and *Ferula negevensis* also occur in these rocks.

Juniperus phoenicea is the tree of subunit 11b. Of the three anticlinal ridges of northern Sinai, Gebel Halal is the richest in trees and accompanying rare plants. The junipers occur in crevices of smooth-faced limestone outcrops and in wadis. Rare Mediterranean relicts which occur in the rock crevices are *E. foeminea*, *R. tenuifolia*, and *Astomaea seselifolium*. A rare component of the rock vegetation is *Origanum isthmicum*, which is endemic to a small area in Gebel Halal ridge. Its closest relative, *Origanum jordanicum*, was discovered in Jordan near Petra (Danin and Künne, 1996). The juniper populations of Gebel Maghara and Gebel Yiallaq grow mainly in wadis.

Subunit 11c is richer both in nondesert trees and shrubs and in companions. The large outcrops of smooth granite and the high elevation of southern Sinai Massif result in a high number of habitats available for the survival of rare species. The typical trees of the rocky environment in 11c are *Pistacia khinjuk*, *Crataegus sinaicus*, and *Ficus (pseudosycomorus) palmata*. The west-facing escarpments of Gebel Serbal are rich in *Moringa peregrina*, which also grows in the vicinity of Wadi Feiran oasis. The typical shrubs of the rocky habitat of 11c are *R. disperma*, *Rhus tripartita*, *Cotoneaster orbicularis*, *Periploca aphylla*, and *Sageretia thea*. Most of the endemic and rare species of Sinai occur in the crevices of rocks that also support trees.

The *Q. calliprinos* and *J. phoenicea* woodlands in Edom, Jordan, are marked 11d on the vegetation map. The climatically controlled belt of the arboreal Mediterranean vegetation terminates in the vicinity of Amman, approximately 120 km north of the Dana-Tafila area. Large areas of shrub-steppe, semisteppe batha, and steppe–forest typify the western ridge of the Jordanian plateau between At-Tafila and Petra. These are dominated by *A. sieberi*, *N. mucronata*, *Argyrolobium crotalarioides*, and thorny species of *Astragalus*. The occasional arboreal components include *P. atlantica*, *Crataegus aronia*, *J. phoenicea*, and *Q. calliprinos*. These formations develop on fissured limestone, basalt, and chalk rocks. The smooth-faced hard sandstone outcrops of the Dana-Petra area support the richest relict Mediterranean flora in the Near East. The trees and shrubs growing here are *Q. calliprinos*, *P. atlantica*, *P. palaestina*, *P. khinjuk*, *C. aronia*, *J. phoenicea*, *Amygdalus korschinskii*, *Ceratonia siliqua*, *O. europaea*, *A. andrachne*, *Rhamnus punctata*, *R. lycioides*, *R. disperma*, *Ficus carica*, *F. palmata*, and *Sageretia thea*. Typical root parasites of the Mediterranean maquis, such as *Osyris alba* and *Thesium bergeri*, occur in the rock crevices, as do typical vines of the maquis – *R. tenuifolia*, *E. foeminea*, *Hedera helix*, *Bryonia cretica*, and *L. etrusca*, which are represented by a higher number of individuals than other Mediterranean vines in any other refugia in the Near East. The rich Mediterranean flora with many endemics may be regarded as existing in a successful refugium, which has functioned for a long time in the evolutionary history of the region. Of the many endemic species of unit 11d species discovered recently are *Origanum petraeum*, *O. punonense*, *O. jordanicum*, *Pycnocycla saxatilis*, *Rubia danaensis*, *Micromeria danaensis*, and *Silene danaensis*.

Desert Vegetation

Sparsely vegetated shrub–steppes dominated by *Anabasis articulata* and *Zygophyllum dumosum* occur at the broad boundary area of the desert with the steppes zone (**Figure 3**, 12 and 10). These typical semishrubs of the desert grow in this transition zone in a diffused pattern. Chalk and marl outcrops are populated with the xerohalophyte communities of *Suaeda asphaltica*, *Salsola tetrandra*, *B.* (*Chenolea*) *arabica*, and *Haloxylon negevensis*. The nearly monospecific communities of semishrub xerohalophytes are accompanied by a diverse assemblage of herbaceous plants that grow on the leached soil only in rainy years.

The extreme desert areas may have no higher plants growing out of wadis, but the entire area may be covered by a highly diverse microbiotic crust. Dor and Danin (1996) studied the microbiotic succession and changes in floristic composition of the crust in the Dead Sea area.

The sequence of plant communities along the wadis of a more arid zone of this category is typified by a stretch of annuals which occur in rainy years. The nature of the annuals community is heavily influenced by the edaphic and climatic conditions of the site. A lower section of the wadi receives higher quantities of water and supports small and short-living semishrubs such as *Pulicaria incisa*, which may function locally as an annual. Further down, larger and long-living semishrubs

such as *A. articulata* or *Z. dumosum* grow. A section dominated by shrubs such as *R. raetam*, *Ochradenus baccatus*, or *Lycium shawii* prevails further down the wadi system. In the lowest section of the wadi system, trees, mostly *Acacia* species or *Tamarix* species, may be found. The nature of plant communities and the sequence of their occurrence along the wadis are in close affinity to rock and soil types, which greatly influence the moisture, salinity, and nutrient regime in the wadi systems.

In similar gravel plains in the large desert area south of Ma'an to the Saudi Arabian border, *A. articulata* grows with trees of *A. pachyceras* at the fifth-order section of the wadi system. The rocky terrain at the escarpments of southwestern Jordan to the Arava rift valley supports diverse communities in a diffused pattern as occurs in Sinai and the Negev.

Sand Vegetation

The main areas in which sand dunes or sand sheets occur in Israel (**Figure 3**, 13) are the Mediterranean coastal plain, the western Negev, a few valleys of the northeastern Negev, and the Arava Valley. Different climatic regimes prevail in each of these areas, sand texture differs, and hence the vegetation and processes of its development differ accordingly. Much of this vegetation, patterns of species diversity, and the nature of the microbiotic crust of the sand in desert areas are discussed in other publications. Thus, some background concerning the sand vegetation of southern Jordan is presented here.

Weathering of Nubian sandstones in the Jordanian plateau contributes mobile sand to the Arava Valley. The relatively high water table at the Arava enables the successful development of the tall shrub *Haloxylon persicum* at sites in which sand is sufficiently deep. However, most stands of this plant in Israel became intensively irrigated agricultural areas. Nonetheless, large areas of *Haloxylonetum persici* cover considerable areas of the Arava, southeastern Sinai, and the southeastern desert of Jordan (Al-Eisawi, 1996). Annuals accompany the shrubs of the sand sheets in rainy years. The sandstone area of the Jordanian plateau is built of sandstone inslebergs, which project from large flat valleys and are filled with stable sand. Huge areas of sand sheets in the Wadi Rum area are dominated by *H. salicornicum*, *A. articulata*, and occasional patches of *H. persicum*. The sandstone hills support many relict Mediterranean species and there is particularly interesting vegetation at the vicinity of the contact zone between the sandstone and the Precambrian crystalline rocks.

Oases with Sudanian Trees

The climate of the rift valley is much warmer than that of the hilly terrain surrounding it. Springs in which a high quantity of freshwater is available throughout the year (**Figure 3**, 14) support thermophilous Sudanian trees such as date palms (*P. dactylifera*), *A. raddiana*, *Acacia tortilis*, *Calotropis procera*, *Moringa peregrina*, *B. aegyptiaca*, *Cordia sinensis*, *Maerua crassifolia*, *Dalbergia sisoo*, *Capparis decidua* (in Jordan), and *Z. spina-christi*. Rich annual flora may accompany sites that are not under the influence of the spring water but successfully grow in rock crevices, soil patches, and abandoned cultivated land. In many freshwater springs/oases near the Dead Sea in Israel and Jordan and in southern Sinai, *Adiantum capillus-veneris* grows on dripping water at shady places with the orchid *Epipactis veratrifolia*. A prominent species of the freshwater springs in the Jordanian desert, *Nerium oleander*, is absent from the desert springs of Israel and Sinai.

Desert Savannoid Vegetation

A considerable part of the Arava and Dead Sea Valleys is covered by savannoid vegetation in which *Acacia* trees are accompanied by desert semishrubs (**Figure 3**, 15). In wadis, the upper soil layers support the typical desert vegetation. *Acacia pachyceras* is the dominant tree of areas that are at a relatively high elevation, such as the upper tributaries of Nahal Paran in Israel, Wadi Jirafi in Sinai, and large areas of gravel plains in Jordan south of Ma'an. *Acacia raddiana*, which is less resistant to low temperatures, is the dominant tree at lower elevations, and *A. tortilis*, which has the highest demands for high temperatures, grows in the southern Arava or below sea level in the northern Arava and Dead Sea Valley. In a few places between Wadi Watir and Sharm el Sheikh in eastern Sinai, rare trees of *C. decidua* grow, which is an important component of the savannah vegetation of southern Egypt and Sudan.

Swamps and Reed Thickets

Only small nature reserves, such as that of the Hula Lake, remain (**Figure 3**, 16) from the swamps that covered large areas at the beginning of the twentieth century. Freshwater springs still flow and their vegetation is composed of a small number of species that produce high quantities of phytomass. The most common hydrophytes are *P. australis*, *Phragmites frutescens*, *Arundo donax*, *Arundo mediterranea*, and *Typha domingensis*. Large areas of swamps in the Hula Valley supported *Cyperus papyrus*, which reached its northernmost station at this valley. Several rivers of the coastal plain that supported riparian vegetation in the past have become sewage canals and their polluted water has destroyed most of the past rivers' vegetation.

Wet Salinas

Wet, salty soils occur in places in which springs of salty water occur or at sites in which the water table is near the surface and the evaporating water leaves salt in the upper soil layers (**Figure 3**, 17). A salt crust may occur at the soil surface in arid regions and plants are restricted to small wadis in which leaching occurs. Most plants growing in the desert salt marshes are perennials that establish themselves in the rare event of leaching. The typical plants of desert salty soils are *Suaeda monoica*, *S. fruticosa*, *S. vermiculata*, *N. retusa*, *Seidlitzia rosmarinus*, and a few species of *Tamarix*.

Synanthropic Vegetation

The category of synanthropic vegetation (**Figure 3**, 18) is further divided in Israel into three subcategories according to

the remnants of trees found in the intensively cultivated areas: In 18a, it is trees of *Q. ithaburensis*, in 18b the tree is *Z. spina-christi*, and in 18c it is *A. raddiana* and *Z. spina-christi*. The synanthropic vegetation of Jordan and the current status of the synanthropic vegetation in Sinai need further investigation.

Mangroves

The mangrove of Nabq, eastern Sinai (**Figure 3**, 19), typically growing in muddy soils of the tidewater, constitutes the most northerly population of *Avicennia marina* on Earth at 28°10′ N. This species was recorded even farther from the equator in South Australia at 37° S. The only population found in the Gulf of Suez shares its water, and possibly warmth, with the Gulf of Elat at Ras Mohammed.

See also: Desert Ecosystems. Mediterranean-Climate Ecosystems. Near East Ecosystems, Animal Diversity

References

Al-Eisawi DM (1996) *Vegetation of Jordan*. Cairo: UNESCO.

Danin A (1983) *Desert Vegetation of Israel and Sinai*. Jerusalem: Cana.

Danin A (1996) Vegetation of Israel and Sinai. *Botanicheskii Zhurnal* 81(11): 14–31.

Danin A (1999) Desert rocks as plant refugia in the Near East. *Botanical Reviews* 65(2): 93–170.

Danin A (2008) Desert rocks – A habitat which supports many species that were new to science in the last 40 years. *Turkish Journal of Botany* 32(2008): 459–464.

Danin A and Künne I (1996) A new species of Origanum (Labiatae) from Jordan: *O. jordanicum* Danin et Künne sp.n., and notes on the species of section Campanulaticalyx. *Willdenowia* 25(2): 601–611.

Dor I and Danin A (1996) Cyanobacterial desert crusts in the Dead Sea Valley, Israel. *Algological Studies* 83: 197–206.

Le Houérou HN (1995) Bioclimatologie et biogéographie des steppes arides du Nord del'Afrique; Diversité biologique, dévelopement durable et désertisation, OPTIONS méditerranéennes, serie b, CIHEAM (Montpellier) and ACCT (Paris), France (tudes et recherches No. 10).

Neotropical Seasonally Dry Forests

Julio Calvo-Alvarado, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica
Arturo Sánchez-Azofeifa, University of Alberta, Edmonton, AB, Canada
Carlos Portillo-Quintero, Instituto Venezolano de Investigaciones Científicas (IVIC), Maracaibo, Venezuela

© 2013 Elsevier Inc. All rights reserved.

Glossary

Climate Change Refers to a statistically significant variation in either the mean state of the climate or in its variability, persisting for an extended period (typically decades or longer).
Deforestation The removal of a forest stand where the land is put to a nonforest use.
Ecological succession The process by which an ecological community undergoes more or less orderly and predictable changes following disturbance or initial colonization of new habitat.
Forest fragmentation When forests are cut down in a manner that leaves isolated patches of forest known as forest fragments or forest remnants.
Sustainable forest management The stewardship and use of forests and forest lands in a way, and at a rate, that maintains their biodiversity, productivity, regeneration capacity, vitality and their potential to fulfill, now and in the future, relevant ecological, economic and social functions, at local, national, and global levels, and that does not cause damage to other ecosystems.
Tropical seasonally dry forests Forests in the tropical region characterized by pronounced seasonality in rainfall distribution with several months of drought.

Introduction

Tropical seasonally dry forests (TSDFs) are considered one of the most threatened forest ecosystems in the world as a consequence of intensive anthropogenic disturbance (Janzen, 1986; Murphy and Lugo, 1986, 1995; Hoekstra *et al.*, 2005; Miles *et al.*, 2006; Portillo-Quintero and Sánchez-Azofeifa, 2010). Their fertile soils and mild climates make them highly suitable for agriculture and livestock (Murphy and Lugo, 1995; Ewel, 1999; Piperno and Pearsall, 2000). In TSDF, annual rainfall does not deviate greatly from potential evapotranspiration; therefore water for irrigation is needed in modest amounts only during the dry season (Ewel, 1999). Given its attractive environmental characteristics for human settlement and development, these ecosystems have historically supported extensive deforestation and high human population densities (Tosi and Voertman, 1964; Janzen, 1986; Quesada and Stoner, 2004; Sánchez-Azofeifa *et al.*, 2005b).

TSDFs have several definitions. According to the Holdridge life zone system, TSDFs have a biotemperature greater than 17 °C, an annual precipitation of 500–2000 mm, a potential evapotranspiration ratio of between 1 and 2 (ratio of mean potential evapotranspiration to mean annual precipitation; a measure of humidity), and a dry season that lasts 4–6 months (Holdridge, 1967, Janzen, 1983, Lüttge, 1997). According to Mooney *et al.* (1995) TSDFs are forests in the tropical region characterized by pronounced seasonality in rainfall distribution with several months of drought. Moreover, Sánchez-Azofeifa *et al.* (2005b) defines TSDF as a vegetation type dominated by deciduous trees where at least 50% of trees present are drought deciduous, the mean annual biotemperature is >17 °C, total annual precipitation ranges between 700 and 2000 mm, and there is the presence of three or more months every year with less than 100 mm of rainfall. Sánchez-Azofeifa *et al.* (2005b) also considers savannas, gallery forests, coastlines and mangroves among the associated vegetation types that can occur within the matrix of NSTFs. According to Pennington *et al.* (2006), TSDF is a tropical vegetation type that experiences a minimum dry season period of 5–6 months with concomitant strongly seasonal ecological processes and functions. The latter definition includes diverse vegetation formations such as forests within grasslands, shrublands, and savanna ecosystems as in Sánchez-Azofeifa *et al.* (2005b); from tall forests on the more humid dry forest sites to cactus scrub on the driest.

Despite TSDF remarkable richness of species, including endemic species, alarming deforestation rates, and the importance of its ecosystem services to sustain economic development for many regions with high population densities, it is surprising to realize how far knowledge of the ecology, the human and biophysical dimensions about this important ecosystem lags behind that of other tropical forest biomes (Murphy and Lugo, 1986; Quesada and Stoner, 2004; Sánchez-Azofeifa *et al.*, 2005b; Miles *et al.*, 2006, Quesada *et al.*, 2009; Portillo-Quintero and Sánchez-Azofeifa, 2010; Balvanera *et al.*, 2011; Chazdon *et al.*, 2011). As of 2005, only 14% of all scientific publications on world tropical forest research focused on TSDF, whereas 86% referred to wet forests (Sánchez-Azofeifa *et al.*, 2005b). In addition, no publications on the effect of climate change on TSDFs or the quantification of ecosystem services were part of the body of knowledge until that specific moment.

According to Miles *et al.* (2006), more than 54% of world's TSDFs are located in the Neotropics, where large continuous fragments still remain and most scientific studies and conservation efforts have taken place. This article refers to seasonally dry forests in the Americas or Neotropical seasonally dry forests (NSDFs). In the next sections, the general physical and biotic characteristics of NSDF as well as the characteristics of its successional stages are described. Patterns and processes of land-use/cover change such as extent, fragmentation

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00354-3

patterns, degree of protection, drivers of change, and issues aspects of silvicultural management are explored.

In addition, this article includes information on the extent and distribution of forest vegetation outside the strict NSDF bioclimatic limits following Portillo-Quintero and Sanchez-Azofeifa (2010). In their study, tropical and subtropical forests within important associated savannas, grasslands, and shrublands ecosystems such as the Chaco dry forests, Caatinga, Cerrado, the Llanos, and other savannas in South America were included. These ecosystems are considered areas where NSDF species could potentially occur (Miles *et al.*, 2006).

NSDF Structural and Compositional Characteristics

When compared with wet forests, NSDF present some unique characteristics:

a. NSDF defy the well-established trend of increasing species richness as one approaches the equator (Fischer, 1960; Gentry, 1988 and 1995; Chazdon and Denslow, 2002). The most diverse NSDF are located farther from the equator (Gentry, 1988 and 1995), been on the Pacific Coast of Mexico a hotspot of NSDF biodiversity worldwide.
b. In NSDF, species richness is higher among the driest sites. For example, a greater diversity has been found in Chamela Mexico and Quiapaca Bolivia than from sites that receive more precipitation such as Guanacaste, Costa Rica (Gentry, 1995; Lobo *et al.*, 2003).
c. NSDFs have a high degree of floristic endemism. For example, the estimates of endemism had been as high as 43% in Mexico (Rzedowski, 1991) or 70% in South America (Gentry, 1982; Joly, 1970). More recently endemism patterns for NSDFs have been confirmed by the study of Linares-Palomino *et al.* (2011). The main conclusions of Linares-Palomino *et al.* (2011) is that 12 of 23 neotropical geographical nuclei have more than 20% of unique species (a proxy for level of endemism), the highest nuclei being Insular Caribbean (77.5%) and Mexican Pacific (65.5%) in the north, and Equatorial Pacific (47.1%) and Peruvian Inter Andean Valley (46.4%) in the South.
d. NSDF plant species show specialized structural and physiological strategies to cope with the seasonal reduction of soil moisture and increase in evapotranspiration demands. These include higher wood density, greater root biomass, and deeper root systems than plant species in humid forests, stem water storage, and hydraulic architecture, as well as complex dynamics of leaf abscission, timing of leaf fall, flowering, and shoot expansion (Holbrook *et al.*, 1995).
e. The *Fabaceae* is consistently the dominant tree family (except for the Caribbean islands where the *Myrtaceae* family predominates). The *Bignoniaceae* is the dominant liana family, but their relative dominance is greater in NSDF as compared to humid forests due to its lower species diversity (Gentry, 1995; Lugo *et al.*, 2006).
f. NSDFs have extremely high β-diversity or spatial species turnover: of 21 NSDF geographic nuclei in the Neotropical region, 80% of all possible pairs (253) have dissimilarity values greater than 70% (Linares-Palomino *et al.*, 2011).

Table 1 Structural and functional characteristics of tropical and subtropical dry and humid (wet/rain) forest

Characteristics	*Forest type*	
	Dry[a]	*Wet*[b]
Structural characteristics		
Number of species[c]	35–90	50–200
Canopy height (m)	10–40	20–84
No. of canopy strata	1–3	3
Basal area ($m^2\ ha^{-1}$)	17–40	20–75
Leaf area index ($m^2\ m^{-2}$	3–7	5–8
Plant biomass ($t\ ha^{-1}$)		
Stems and branches	28–266	209–1163
Leaves	2–7	7–10
Roots	10–45	11–135
Total	78–320	269–1186
Root biomass as % of total	8–50	5–33
Functional characteristics		
Net total primary productivity ($t\ ha^{-1}\ yr^{-1}$)	8–21	13–28
Fine litter production ($t\ ha^{-1}\ yr^{-1}$)	3–10	5–14
Tree diameter growth ($mm\ yr^{-1}$)	1–3	2–6

[a]Annual rainfall 500–2000 mm; strongly seasonal; annual PET/P normally >1.
[b]Annual rainfall >2000 mm; little or moderate seasonality; annual PET/P normally ≤ 1.
[c]Based on surveys of 1–3 ha; includes trees at least as small as 10 cm dbh.
Source: Adapted from Murphy PG and Lugo AE (1986) Ecology of tropical dry forest. *Annual Review of Ecological Systematics* 17: 67–88.

g. Structure of NSDF is characterized by smaller trees, lower basal area, diameter growth, biomass, and productivity than wet forests (**Table 1**), but they tend to be more similar with respect to tree density (i.e., number of trees per hectare) (Gentry, 1995). For example; continental NSDFs have the same average number of plants 2.5 cm in dbh in 0.1 ha plots as do humid forests (370 vs. 373), the same number of trees 10 cm dbh (65 vs. 64), the same number of lianas 2.5 cm diameter (72 vs. 68), and only slightly lower basal area in $m^2\ ha^{-1}$ (34.7 vs. 40). However, average root biomass as a portion of total plant biomass is greater in TSDF (33.5% with a range of 8–50%) than in wet forests (16% with a range of 5–33%) (Holbrook *et al.*, 1995).

On average, NSDF species richness is close to 64 (range 21–121); subdivided in: tree richness of 50 sp. (range 21–81) and lianas richness 15 sp. (range 0–37) (Gentry, 1995). The most important families in order of dominance are: *Fabaceae*, *Bignoniaceae*, *Rubiaceae*, *Sapindaceae*, *Euphorbiaceae*, *Flacourtiaceae*, *Capparidaceae*, *Apocynaceae*, *Nyctaginaceae*, *Polygonaceae*, *Malpighiaceae*, and *Cactaceae*. The most representative genera on Gentry's plots were: *Tabebuia*, *Cordia*, *Casearia*, *Bauhinia*, *Trichilia*, *Erythorxylum*, *Randia*, *Hippocratea*, *Serjania*, *Croton*, *Zanthoxylum*, *Arrabidaea*, *Celtis*, *Capparis*, *Combretum*, *Pithecellobium*, *Pterocarpus*, *Paullinia*, *Bursera*, *Coccoloba*, *Macfadyena*, *Eugenia*, *Acacia*, *and Ficus*; *Tabebuia* being the most widespread genus.

NSDFs have larger diversity of life forms than humid forest ecosystems, likely a consequence of habitat heterogeneity in water availability, with woody deciduous components being the dominant type. Tropical moist and wet forest include six major life forms; evergreen woody plants, deciduous woody plants, lianas, C3 epiphytes, crassulacean acid metabolism (CAM)

epiphytes, and hemiparasites; whereas NSDF includes nine life forms: evergreen woody plants, deciduous woody plants (40% of the woody species), sclerophyllous woody plants, succulent woody plants (including cacti), herbs, rosettes, lianas, CAM epiphytes, and hemiparasites (Medina, 1995).

Other characteristics of NSDF are the high proportion of wind-dispersed species (i.e., anemochorous species). According to Gentry (1995), 80% of the lianas and a third to a quarter of the trees are wind dispersed, contrasting with the figures for the humid forest where 90% of the trees and 75% of the lianas are zoochorous (i.e., spores or seeds are dispersed by animals). Also, from about two-thirds to three quarters of NSDF woody taxa have conspicuous flowers, which are pollinated by hawkmoths, bats, and large and medium size bees (Bullock, 1995).

Most of the ecological processes in NSDF are seasonally driven, and depend on the length and intensity of their dry season. Leaf litter biomass builds up during the dry season, flowering and fruiting phenologies occur in specific moments of the year and many flower species blossom at the same time during the transition between dry and wet season (Pennington *et al.*, 2006; Bullock, 1995; Mooney, 1995). In the driest sites, there is an increase in the proportion of evergreen and succulent species.

Animal Diversity

According to available comparable data from 10 selected study sites (Ceballos, 1995), NSDF have on average less vertebrate species diversity than moist and wet forests: total species (443 vs. 630), mammal sp. (88 vs. 113), birds sp. (287 vs. 404) and herptiles sp. (67 vs. 112). All animals in NSDF exhibit diverse means to adapt to the reduction of available food and water resources during the dry season (Ceballos, 1995; Stoner and Timm, 2011). These adaptations might include one or a combination of several strategies: latitudinal or elevation movements, shifts in diet to utilize seasonally available food resources, seasonal accumulation of fat or food resources, reproductive delay, behavioral and physiological adaptations. Wildlife trade, hunting, forest fires, pesticide contamination, destruction and fragmentation of NSDF habitats are the major problems affecting conservation of biodiversity (Ceballos, 1995; Stoner and Timm, 2011; and Chazdon *et al.*, 2011).

A large proportion of all insects in tropical forests are phytophagous or parasitoids of other phytophagous insects. According to Hanson (2011), the majority of insect families are probably more species rich in humid forests than in NSDF because plant species richness is greater in humid forests. Some NSDF characteristics can affect insect diversity; for example, the quality of dead wood in the forest floor (Dynastinae family diversity), soil physical and chemical characteristics (insect taxa with soil-inhabiting larvae), and reduced number of epiphytes (ant diversity).

Secondary Succession in NSDF

Many NSDF can no longer be considered pristine old-growth forests, but rather as a mosaic of successional stages due to extensive anthropogenic disturbances, particularly Mesoamerica and the Caribbean (Janzen, 1986, 1988; Arroyo-Mora *et al.*, 2005a,b; Quesada *et al.*, 2009; Calvo-Alvarado *et al.*, 2009). Recent studies have also shown that NSDFs are recovering in some areas through natural restoration (Abizaid and Coomes, 2004; Arroyo-Mora *et al.*, 2005a,b; Tucker *et al.*, 2005), prompting the need to study the ecology of different NSDF successional stages (Quesada *et al.*, 2009). Murphy and Lugo (1986) pointed out that NSDF succession takes place at a slower rate in terms of plant growth but because the canopy of STDF has a relative low tree height and simple vertical structure, they have the potential to recover to a mature state more quickly than do humid forests, and therefore, they are considered more resilient (but see Quesada *et al.*, 2009).

One of the most important NSDF restoration efforts in the neotropics took place in the Santa Rosa National Park, Costa Rica, in 1984, under the initiative of Dr. Daniel Janzen. Restoration of the NSDF of Santa Rosa has been achieved mostly by fire control (Janzen, 1987). As a consequence, the wind-dispersed seeds from neighboring natural forests germinated and seedlings grew to provide shade and habitat for animals, which in turns brought in the seeds of their favorite fruits in their excreta and spittle. The defense of solitary trees in pastures (nuclear trees) from fires was also favorable for natural regeneration, since when the tree got large enough to provide shade for domestic grazing animals and habitat for birds, it attracted seed dispersing animals that brought in the second wave of seeds (Allen, 1988; Janzen, 1987, Castillo *et al.*, 2011).

To highlight the compositional and structural differences between NSDF successional stages, we show results from an ecological study of three successional NSDF stages in **Table 2**, with information obtained from 26 plots in Santa Rosa, Costa Rica. Measurements were obtained from 26 plots of 0.1 ha each and measuring all trees 10 cm diameter (Kalacska *et al.*, 2004a, b). In general, the results in Table 2 point to a progressive increase in forest structure, tree species richness, and diversity from early to late stages. It is important to notice from **Table 2** how pasture land can recover after 15 years of fire protection and become an early secondary forest that provides better habitat and environmental services than degraded pasture land.

Extent and Conservation of NSDF

Approximately 40% of the total world tropical and subtropical regions were once potentially covered by open or closed forest (Murphy and Lugo, 1986, in accordance to the Holdridge Life Zone System; Holdridge, 1967). Of this total potential forest cover, 47% belongs to the dry forest life zone and 53% belongs to the humid forests life zones (i.e., wet, moist, and rain forests).

A recently published study by Portillo-Quintero and Sánchez-Azofeifa (2010) provided new insights into the extent and fragmentation of NSDF in the Americas. This study focused exclusively on NSDF, including satellite imagery selection and classification algorithms that accounted for latitudinal differences in leaf phenology patterns (differences in the date of the range and peak of the dry season and wet season). Their study also focused exclusively on NSDF

Table 2 Structural and functional characteristics of three NSDF successional stages in Santa Rosa National Park, Costa Rica

Trait	*NSDF successional forest stage*		
	Early	*Intermediate*	*Late*
Structural characteristics[a]			
Species density, No./0.1 ha	15±7	29±5	29±7
Canopy height (m)	7.5±2.2	10.3±3.4	15.0±2.2
No. of canopy strata	1	1	2
Basal area (m^2 ha)	11.7±5.4	21.4±6.8	30.1±6.5
Stem density, No./0.1 ha	112±64	130±35	107±42
Shannon diversity	1.77±0.56	2.88±0.36	2.75±0.56
Above ground plant biomass, t/0.1 ha	5.6±3.3	10.9±4.6	16.0±4.3
Carbon (t ha^{-1})	31.8±1.8	60.9±2.5	88.9±2.0
Leaf area index (m^2 m^{-2}) (dry season)	0.3±0.7	0.3±0.6	1.5±0.9
Leaf area index (m^2m^{-2}) (wet season)	3.8±2.6	3.7±1.2	11.3±4.9
Functional characteristics[b]			
Litter fall (kg h yr^{-1})	152,08	483,93	454,80
Tree diameter growth (mm yr^{-1})	1.33±1.54	1.83±3.15	1±1.61

[a]Information adapted from Kalacska M, Sanchez-Azofeifa GA, Rivard B, Calvo-Alvarado JC, and Quesada M (2008) Baseline assessment of environmental services payments from satellite imagery: A case study from Costa Rica and Mexico. *Journal of Environmental Management* 88: 348–359; Kalacska M, Sánchez-Azofeifa A, Rivard B, Caelli T, White P, and Calvo-Alvarado J (2007) Ecological fingerprinting of ecosystem succession: Estimating secondary tropical dry forest structure and diversity using imaging spectroscopy. *Remote Sensing of Environment* 108: 82–96; Kalacska M, Sánchez-Azofeifa A, Calvo-Alvarado J, Quesada M, and Janzen D (2004a) Species composition, similarity and diversity in three successional stages of tropical dry forest. *Forest Ecology and Management* 200: 227–247, and Kalacska M, Sánchez-Azofeifa A, Rivard B, *et al.* (2004b) Leaf area index measurements in a tropical moist forest: A case study from Costa Rica. *Remote Sensing of Environment* 91: 134–152. Mean values with 1 standard deviation.
[b]Unpublished information for Santa Rosa provided by J. C. Calvo-Alvarado.

that fall within the floristic, zoogeographic, and bioclimatic parameters suggested by Olson *et al.* (2001) and by Sánchez-Azofeifa *et al.* (2005b) for North America, Central America, the Caribbean islands, and South America.

From the study of Portillo-Quintero and Sánchez-Azofeifa (2010) several important aspect of NSDF should be discussed.

NSDF Extent

The total extent of NSDF in the Americas as of 2010 was 519,597 km^2. North and Central American dry forests occupy 203,884 km^2 of land, this represents 39% of the total extent of all STDF, while South America contains 268,875 km^2 of NSDF forest representing 51% of the total. The Caribbean islands comprise 46,839 km^2, which represents 9% of the tropical dry forest remaining. From the analysis of the map in **Figure 1** the following information and conclusions can be derived: (a) NSDFs are far from being uniformly distributed through the continent; (b) Mexico has the largest amount of tropical dry forest within its boundaries, comprising 38% of all tropical dry forests; (c) Bolivia and Brazil harbors large portions of NSDF (25% and 17%, respectively); (d) Colombia and Venezuela includes lower percentages (6.5% and 6.2% respectively) of NSDF; (e) together, Mexico, Bolivia, Brazil, Colombia, and Venezuela contain 93% of the tropical dry forests in continental areas; (f) in the Caribbean islands, the majority of dry forests (92%) occur within Cuba and Dominican Republic. Cuba contains 79% (36,996 km2) of all Caribbean islands NSDF representing 7% of the NSDF in the Americas.

NSDF Loss

When comparing the current extent of NSDF in the Americas to its potential extent according to the delineation of the dry forest biome it appears that 66% of the area has already been converted to other land uses. This value is higher than the world estimate of 48% of NSDF loss made by Hoekstra *et al.* (2005). In North and Central America, 72% of dry forest has disappeared, while South America and the Caribbean islands have lost 60% and 66% of their potential cover, respectively. In continental countries, the magnitude of NSDF conversion ranges from 45% in Bolivia to 95% in Peru, with the next highest conversion in Guatemala with 86%. With the exception of Bolivia, all other continental countries have converted at least 52% of their NSDF cover, and the overall conversion for these countries is an impressive 66%. The five insular countries for which data are available show a magnitude of conversion that ranges from 54% in Jamaica to 78% in Haiti, with an overall conversion of 66%. In sum, this analysis suggests a dramatic rate of conversion, with only 44% of NSDF remaining in the region (**Table 3**).

It is surprising that given its importance for conservation, no accurate estimates of tropical deforestation exist for NSDF and no distinction or attention is given to it in recent global deforestation assessments where forest extent and deforestation are analyzed as a whole (FAO, 2011). Proper systematic evaluation of NSDF deforestation rates is needed in order to understand regional dynamics of NSDF loss. Costa Rica is probably the only country with accurate estimates of deforestation, with deforestation rates for the period 1986–98 estimated at around 0.2%. Then, for the 1998–2005 time period, deforestation drops close to 0.1%. Costa Rica is most likely the exception to the regional patterns, since much higher rates have been reported for Bolivia and Brazil (Killeen *et al.*, 2007; Steininger *et al.*, 2001).

The causes of deforestation for NSDF are different from those of tropical rain forests. In general, NSDF has been not the last, but the first frontier for economic development policies,

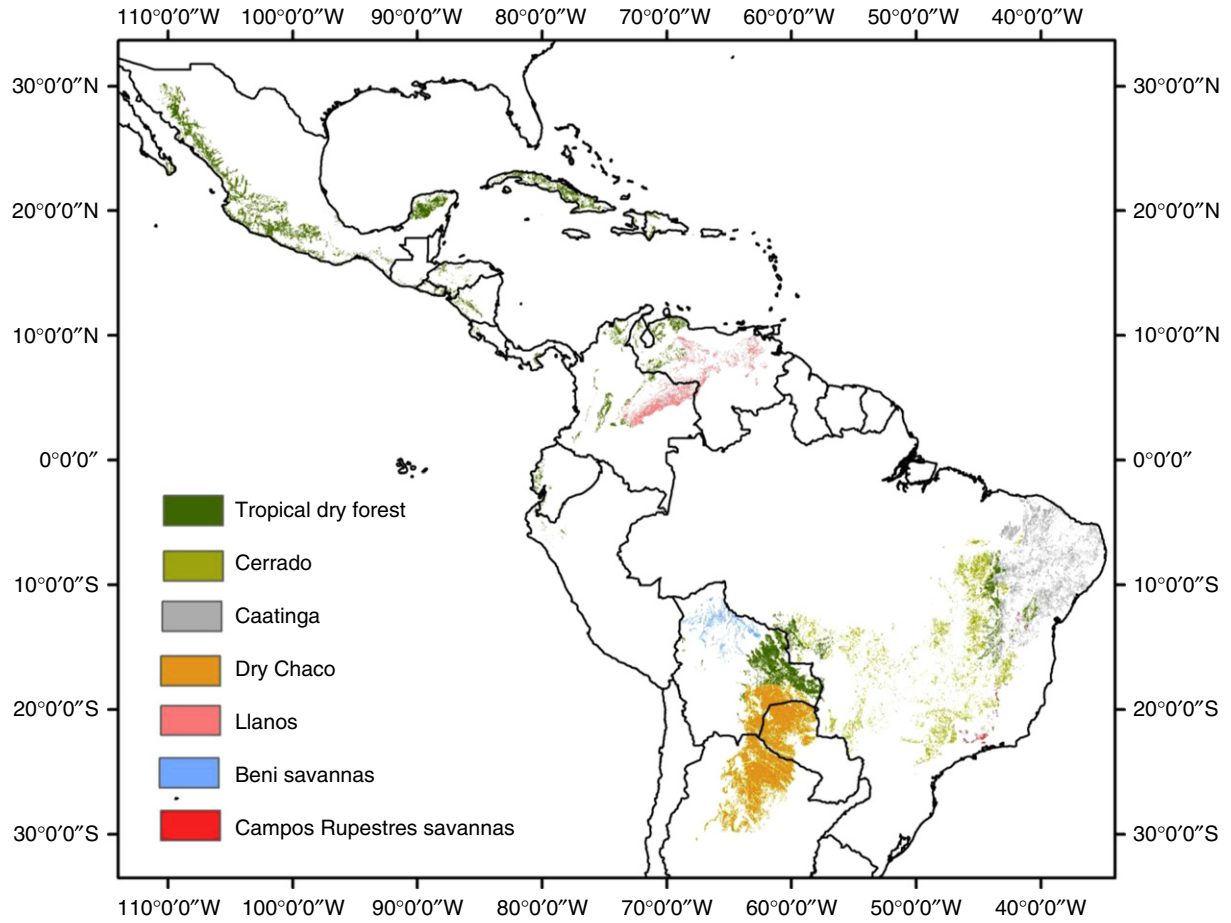


Figure 1 Land-cover map showing the extent and geographical distribution of NSDF derived from MODIS 500-m. Reproduced from Portillo-Quintero C and Sánchez-Azofeifa A (2010) Extent and conservation of tropical dry forests in the Americas. *Biological Conservation* 143(1): 144–155. The map also shows the extent and geographical distribution of seasonally dry forests in tropical and subtropical grassland, shrubland and savanna ecosystems.

and the degradation and deforestation rates in this ecosystem have been greater than for other types of tropical forests (Sánchez-Azofeifa *et al.*, 2005a,b; Portillo-Quintero and Sánchez-Azofeifa, 2010). It is important to separate the effects of continental and insular drivers of deforestation, since they are different. The main driver of change in NSDF is the expansion of the agricultural frontier in Continental America, since these forests are located in areas with excellent soils for agricultural industries and cattle expansion (Sánchez-Azofeifa *et al.*, 2005b; Kalacska *et al.*, 2004a, b; Calvo-Alvarado *et al.*, 2009). Arroyo-Mora *et al.* (2005a,b) and Calvo-Alvarado *et al.* (2009), examining deforestation processes in Guanacaste, Costa Rica, concluded that the expansion of the cattle frontier was the main force driving tropical deforestation in the late 1950s and 1960s, the former contributing to the creation of a highly fragmented landscape. Trejo and Dirzo (2000) studying a NSDF in central Mexico, and later Sanchez-Azofeifa *et al.* (2009) studying deforestation processes in the coastal region of Jalisco and the areas surrounding the Chamela–Cuixmala Biosphere Reserve, arrived at similar conclusions.

Although the forces driving land-use/cover change in NSDF environments appear superficially similar to those for other sites in the tropics, an analysis of drivers of change for 35 Neotropical NSDF ecoregions (Olson *et al.*, 2001) indicates that expansion of the agricultural frontier, cattle ranching, and selective logging emerge as the main dominant drivers in more than 50% of all NSDF sites studied, including both continental and insular sites. Nonetheless, when island and continental sites are analyzed separately, for insular regions the main driving forces are invasion of exotic species, urban sprawl, selective logging, agriculture, tourism development, and road construction. For continental sites, however, the main driving forces are agricultural expansion, cattle ranching, selective logging, urban sprawl, and hunting (**Figure 2**). When continental and insular drivers are compared, only two drivers of change emerge: selective logging and urban sprawl are common to each region, showing a sharp difference not found previously in the literature. The former has driven us to conclude that the expansion of the agricultural frontier and cattle ranching/grazing were the main forces of land-cover change in this ecosystem.

NSDF Protection

Only 4.5% (23,417 km^2) of the total NSDF current extension in the Americas is protected by nature reserves. In South America, 17,816 km^2 of dry forests are protected. Two

Table 3 Current NSDF forest extent (km^2) derived from MODIS 500-m data and area under protected areas at three levels: (a) North, Central, and South American countries; (b) Countries of the Caribbean islands; and (c) Summary of results per subregion

Country	*NSDF potential extent*[a]	*NSDF current extent*[b]	*% NSDF converted*	*NSDF protected (km^2)*	*% Under protection*
(a) North, Central, and South American countries					
Mexico	625,038	181,461	71	336	0.2
Bolivia	216,031	118,940	45	10,609	8.9
Brasil	168,164	81,046	52	5,015	6.2
Venezuela	113,143	29,396	74	302	1
Colombia	92,664	30,713	67	1555	5.1
Peru	48,914	2337	95	188	8.1
Nicaragua	32,277	7414	77		
Honduras	26,582	6280	76		
Ecuador	25,275	6443	75	147	2.3
El Salvador	11,291	3344	70	9	0.3
Guatemala	10,431	1463	86		
Costa Rica	7559	1795	76	279	15.6
Panama	6160	2128	65		
Total	1,383,529	472,759	66	18,620	3.9
(b) Countries of the Caribbean islands					
Cuba	109,879	36,996	66	4023	10.9
Dominican Republic	14,669	6194	58	368	6
Haiti	8971	2002	78	0	0
Jamaica	3438	1585	54	400	25
Cayman islands	173	63	64	3.5	5.6
Total	137,130	46,839	66	4797	10.2

(c) Summary of results per subregion

Subregion	*NSDF potential extent*[a]	*NSDF current extent*[b]	*% NSDF converted*	*NSDF protected (km^2)*	*% Under protection*
N and C America	719,338	203,884	72	624	0.3
South America	664,191	268,875	60	17816	6.6
C. islands	137,130	46,839	66	4797	10.2
Total	1,520,659	519,597	66	23417	4.5

[a]Information based on Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on earth. *BioScience* 51: 933–938.

[b]Information based on Portillo-Quintero C and Sánchez-Azofeifa A (2010) Extent and conservation of tropical dry forests in the Americas. *Biological Conservation* 143(1): 144–155, and Sánchez-Azofeifa A and Portillo-Quintero C (2011) Extent and drivers of change of Neotropical seasonally dry forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 259–277, 408 pp. London, UK: Island Press.

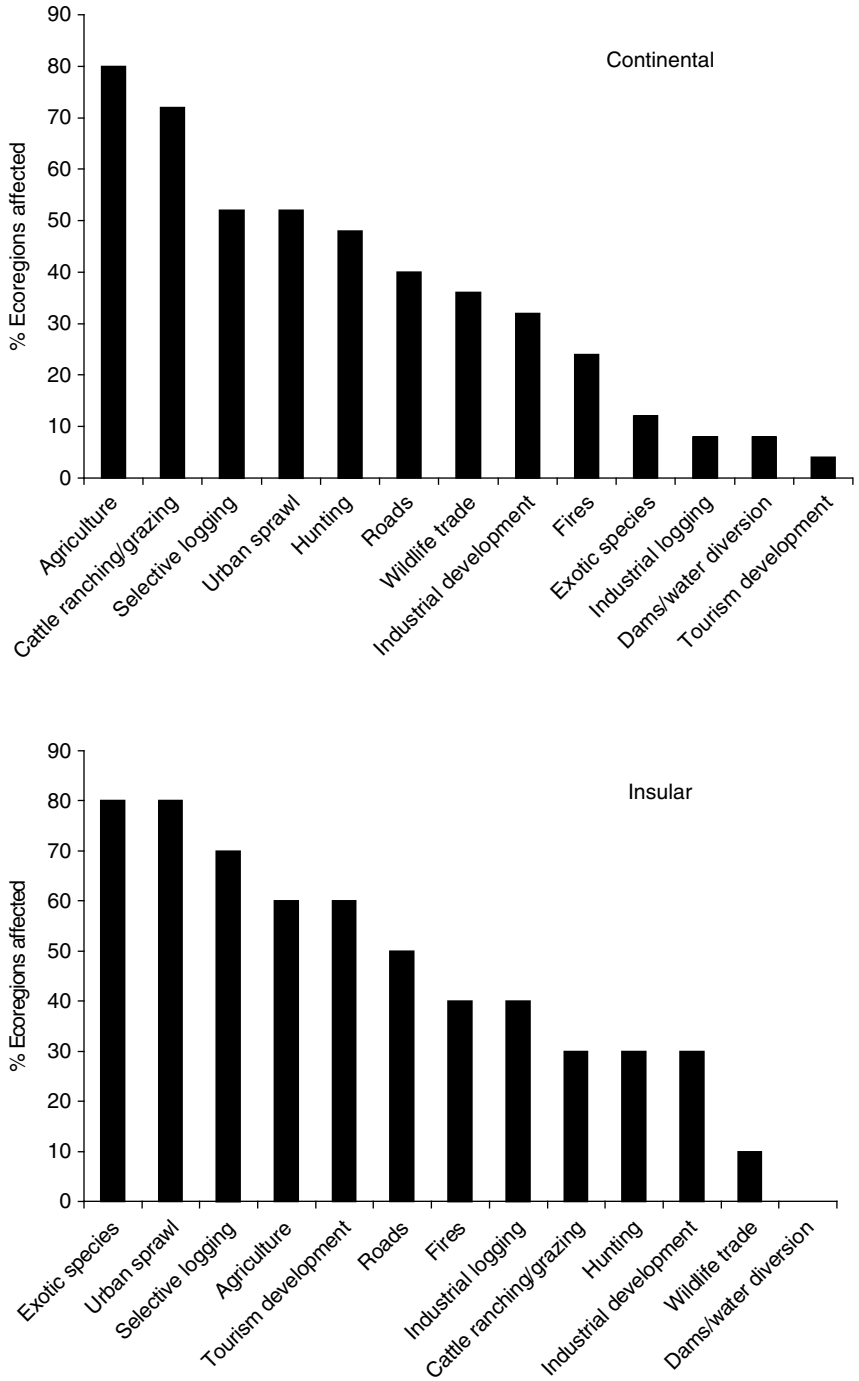


Figure 2 Drivers of forest cover change in NSDF. Reproduced from Sánchez-Azofeifa A and Portillo-Quintero C (2011) Extent and drivers of change of Neotropical seasonally dry forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests*: *Ecology and Conservation*, pp. 259–277, 408 pp. London, UK: Island Press.

countries, Bolivia and Brazil, jointly contain roughly 64% of all protected NSDF (15,000 km^2). Bolivia protects 10,609 km^2 of NSDF, with approximately 7600 km^2 being located within the Chiquitano dry forest ecoregion and protected by a single park: the Noel Kempff Mercado National Park. In North and Central America, where 72% of the dry forests have already been lost, only 804 km^2 (0.4%) are protected. Most of the protection occurs in Mexico and Costa Rica, where a similar extent is under protection in both countries (about 300 km^2). However, the proportion protected relative to their current dry forests extents indicates differences in the relative importance that countries give to the protection of NSDF (See **Figure 3**). These two countries have large differences in potential and current extent of tropical dry forests, yet Costa Rica protects 15% of its current extent, while Mexico protects only 0.2%. Costa Rica has the highest degree of protection of NSDF in

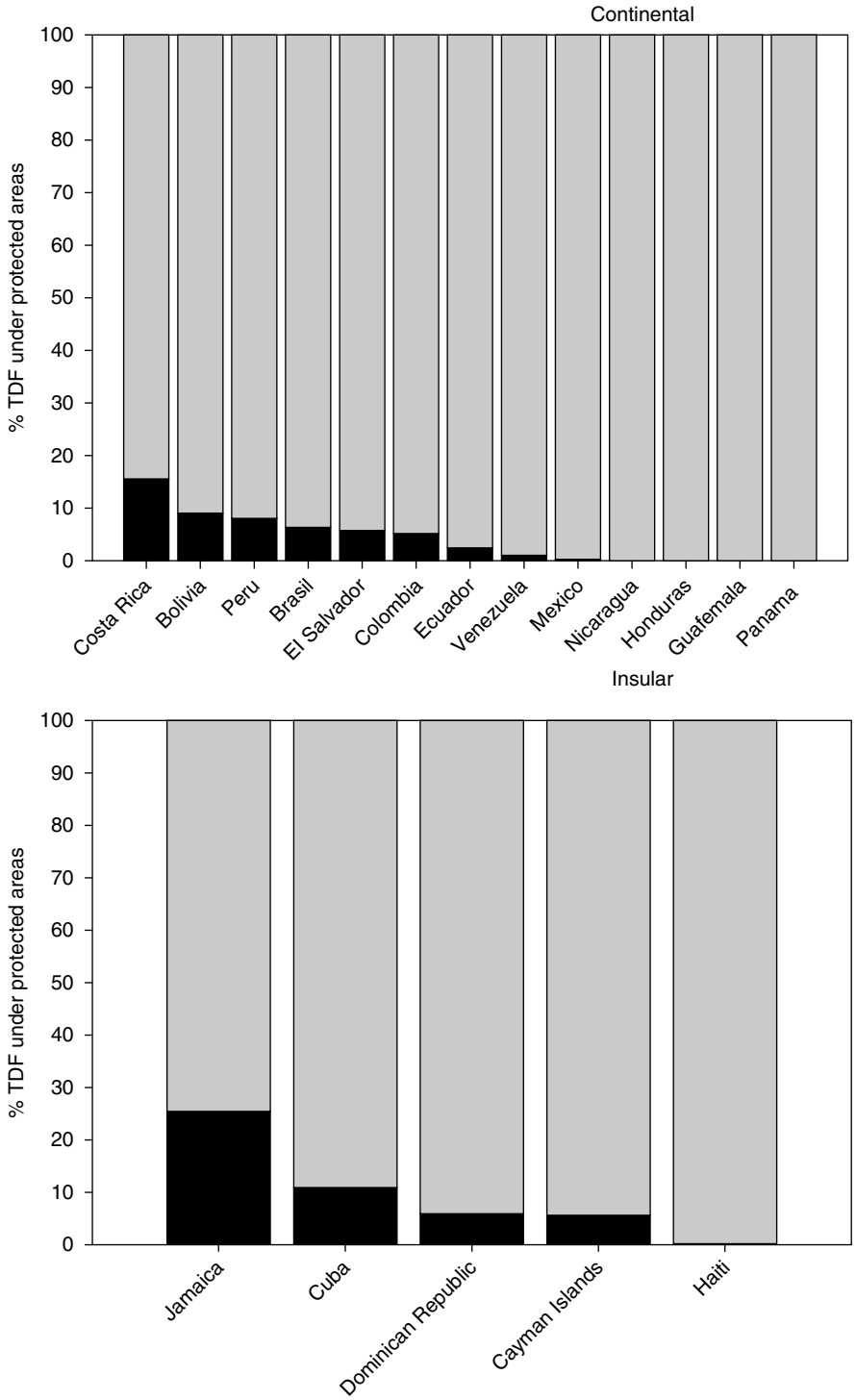


Figure 3 Percentage of NSDF under Protected Areas. Reproduced from Portillo-Quintero C and Sánchez-Azofeifa A (2010) Extent and conservation of tropical dry forests in the Americas. *Biological Conservation* 143(1): 144–155.

North and Central America, and this might be related to the national priorities placed on conservation. Cuba shows protected areas representing 10.9% and Jamaica about 25% of the NSDF extent. In general, countries in the Caribbean islands show the highest coverage of NSDF under protected status (10.2% of current extent is protected) and countries of North, Central, and South America protect 3.9% of the forests' extent, this is probably due to the fact that most of their forest cover is in fact dominated by NSDF. However, protection to forests differs qualitatively given that in North, Central, and South America, about 90% of dry forests are protected under the "National Park" status (IUCN Protected Area Category II),

while in the Caribbean islands, 66% are protected under "Sustainable Management areas" (IUCN Protected Area Category VI) and 20% under the "National Park" category.

NSDF Fragmentation

As is apparent in **Figure 1**, NSDF occur in disconnected patches scattered in the landscape, with sizes varying from small fragments (2.5 km^2), intermediate fragments (2.5 km^2 and 10 km^2), to large fragments (10 km^2). According to Laurance *et al.* (2002) and Rodríguez *et al.* (2007a,b) larger fragments protect better biodiversity and ecosystems functions, while intermediate and small fragments experience higher species extinction and deforestation rates. Given the story of land-cover transformation of NSDF landscapes Laurance *et al.* (2002) observations must be validated for NSDFs regions.

Bolivia has the highest proportion of NSDF in large fragments, followed by Brazil and Mexico (**Figure 4**). The proportions of larger fragments in the majority of countries is 60%, probably because of the way land-use change advances in agricultural lands, leaving patches of NSDF in areas of more difficult access. The highest proportion of small and intermediate fragments is found in countries like Nicaragua, Guatemala, Ecuador, Costa Rica, and Peru. Most Caribbean island countries have a high proportion of NSDF in large fragments (80%) except for Jamaica.

NSDF fragmentation is one of the biggest threats to the protection of biodiversity and ecosystem functions, but it can also be an opportunity for the design and establishment of nature reserves and biological corridors. According to Balvanera *et al.* (2011) agricultural and pastoral goods obtained by removing NSDF are their most commonly recognized ecosystem service, but NSDF also contribute to human well-being through a wide variety of services; including the regulation of a number of processes such as soil fertility, erosion, water cycling, climate, pollination, pests, disease control, invasion by exotic species, and buffering of impacts of extreme meteorological events, as well as many cultural services. Until now, in NSDF landscapes, agricultural production has been favored at the cost of losing most of the other services they provide, probably irreversibly.

Good management strategies of fragmented NSDF landscapes must recognize and recover traditional ethnobotanical information and use of local plant species. Peters (2011) indicated that local and indigenous traditional communities in Brazil have learned the properties, uses, and ecological requirements of hundreds of native species from NSDFs. For example, 80% of the species in Pernambuco, Brazil, have been shown to have local uses (Albuquerque *et al.*, 2005), more than 400 species have medicinal use in northeastern Brazil (Victor, 1990), and more than 176 species are used for food alone in Tarahumara, Chihuahua, Mexico (Bye, 1989). Chazdon *et al.* (2011) showed that agricultural landscapes in NSDF regions support a considerable proportion of the original biodiversity but fewer species than found in intact forest. Further research is needed in order to assess the diversity of NSDF species that survive and thrive using the resources available in agricultural areas.

NSDF in Tropical and Subtropical Grasslands, Savannas and Shrublands

Regarding deciduous forest vegetation outside the strict NSDF bioclimatic range, **Figure 1** (from Portillo-Quintero and Sánchez-Azofeifa, 2010) shows the extent and distribution of tropical and subtropical forests within important associated savanna, grassland, and shrubland ecosystems. These ecosystems are: the Chaco dry forests, Caatinga, Cerrado, the Llanos, the Beni Savannas, and the Campos Ruprestres Savannas. These ecosystems are considered areas where NSDF species could potentially occur (Miles *et al.*, 2006).

Results show that the total extent of forests in grassland, shrubland, and savanna ecosystems is 1,069,692 km^2, which doubles the extent of NSDF. The Dry Chaco comprises 41% of this total; the Caatinga and Cerrado both have very similar extent of forest vegetation, each of them representing about 22% of the total forests; the Llanos and Beni savannas are mostly represented by gallery forests and submontane forests encompassing 9.67% and 4%, respectively; and the Campos Rupestres savannas represents 1% of the total amount of forest in these ecosystems.

In the shrublands and savanna ecosystems, 35,820 km^2 of forests are under legal protection representing 3.3% of the total extent. The Dry Chaco and the Cerrado woodlands show the largest extent under protection with 13,310 km^2 and 11,635 km^2 protected, respectively, followed by the 5446 km^2 of forests protected in the Llanos ecosystems. The degree of protection is rather similar across ecosystems, although the Llanos and the Campos Rupestres savanna ecoregions have a slightly higher degree of protection. Approximately 90% of these protected forests exist under the "National Park" category status.

The Dry Chaco region contains most of its forest extent in large fragments (**Figure 4**). Caatinga ecosystems are more scattered and fragmented, although a significant proportion still remains as large fragments. A similar pattern is detected along the Cerrado, Beni Savanna, the Llanos, and the Campos Rupestres savanna ecosystems.

Fires and NSDF Land-Cover Change

Although the presence of fires is not considered a major driver in the evolutionary history of part of NSDF, it is considered one of the main factors driving land-cover change in this ecosystem (Otterstrom, 2006). Fire in tropical regions is generally considered a landscape-level disturbance that impoverishes ecosystems while it serves as an important element in the human domination of the landscape (Ehrlich *et al.*, 1997). Di Bella *et al.*, (2006) found that the Neotropical ecosystem that is most affected by the occurrence of fires is NSDF, a fact that was later confirmed by Wright *et al.* (2007). Given that the majority of the population in the Neotropics lives within the dry forest ecosystem in a matrix of high-intensity agriculture (Fajardo *et al.*, 2005), it is expected that fire occurrence will be higher here than in other tropical ecosystems.

In general, fires in tropical dry forests are initiated by farmers, either aiming to control pasture lands at the end of the dry season or as a slash-and-burn technique after clear-cutting, although the latter practice rather than the former is more used

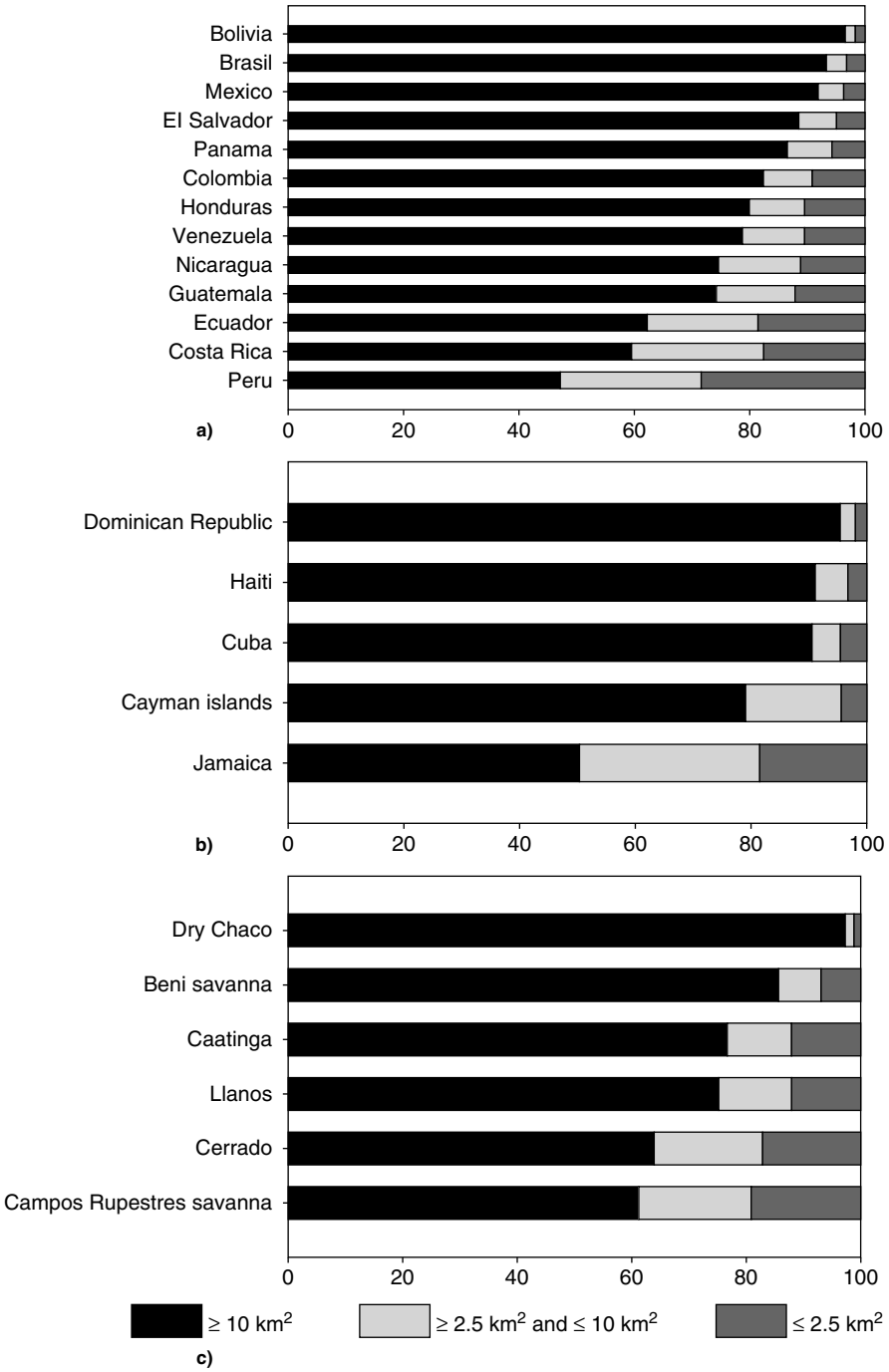


Figure 4 Percentage of NSDF under three levels of fragmentation for countries in North, Central, and South America, the Caribbean and within shrub land and savannas ecosystems. Reproduced from Portillo-Quintero C and Sánchez-Azofeifa A (2010) Extent and conservation of tropical dry forests in the Americas. *Biological Conservation* 143(1): 144–155.

in tropical dry forest environments. Fire control, in fact, is considered to be the most important achievement in the restoration of tropical dry forests at the Santa Rosa National Park after other significant initiatives for forest restoration failed. Unfortunately, our current knowledge of fire's impact on NSDF and their long-term resilience is limited (Otterstrom *et al.*, 2006), and additional work is necessary to better understand the long-term effects of this land-cover change driver.

Climate Change and NSDF

According to the thorough analysis recently carried out by Meir and Pennington (2011) of the possible consequences of climate change on NSDF using a Dynamic Global Vegetation Model (DGVM), the estimated warming rates for NSDF regions are consistent with, or slightly above, the global predicted means of approximately 2–4 °C warming by 2100.

However, the scenarios for regional alterations of rainfall are uncertain and there is no clear global pattern emerging so far from the IPCC (2007). The most recent Intergovernmental Panel on Climate Change (IPCC) examination of the twenty-first century hydrological cycle (Bates *et al.*, 2008) indicates that a 5–15% reduction in soil moisture availability is expected by the late twenty-first century across much of tropical Latin America. These results have also been confirmed by other studies in Amazonia (Timmermann, 1999; Cox *et al.*, 2008; Malhi *et al.*, 2008), but not for tropical dry forests. Other alterations to key climate variables, such as changes in radiation flux caused by differences in cloud cover, are thought to impact primary production (Mercado *et al.*, 2009) and may also influence tree reproduction (Wright *et al.*, 1999). These two effects could be important in seasonal tropical dry forest regions with very short growing seasons. Land-use change impacts are likely to exacerbate the effects of climatic warming, and this may be especially true for NSDF, given that they are mostly already highly disturbed and fragmented.

Although the findings above suggest an increased frequency of extreme events in NSDF at seasonal and longer timescales, predictions about their future status are uncertain for a number of reasons. First, climate models differ considerably in their predictions about rainfall change, the main controller of this system. Furthermore, major grid modeling efforts, and IPCC predictions fail to cover areas associated with NSDFs. Second, the present capacity to model vegetation response cannot fully capture even the information that we already have about the physiological responses of the many constituent species of these systems (Mooney, 2011). Much emphasis had been placed on the impacts of climate change on tropical rain forest, whereas less effort has been made on NSDF (Meir and Pennington, 2011).

Management Aspects of NSDF

An issue that deserves better understanding is the transformation of tropical forests through forest management practices that pursue the sustainable harvesting of timber from natural forests (Sheil and Van Heist, 2000; Quesada *et al.*, 2009). The main argument is that forest logging and silvicultural practices are extrinsic forest disturbances similar in size and intensity to intrinsic natural disturbances such as tree falls (e.g., Skorupa and Kasenenek, 1984). Such management practices propose that natural regeneration and gap forest dynamics can be applied to the sustainable management of tropical forests (e.g., Brokaw, 1987; Denslow, 1987; Brandani *et al.*, 1988). Some of the main ecological issues that question these forest management practices are: (1) natural tree falls occur because of dead, hollow, or vulnerable trees and hence forest management implies extra forest disturbance (Uhl and Guimarães-Vieira, 1989; Barrantes *et al.*, 1999; Putz *et al.*, 2008); (2) harvested forest gaps are generally larger than natural gaps favoring the regeneration of pioneer and sometimes exotic invasive species, which in turn simplify and homogenize the forest composition and structure (Barrantes *et al.*, 1999; Costa and Magnusson, 2003; Putz *et al.*, 2008); (3) removal of larger trees will negatively affect subsequent regeneration due to loss of fruit and seed sources (Barrantes *et al.*, 1999; Costa *et al.*, 2002; Lobo *et al.*, 2007); (4) forest management will negatively affect the role of many animal species in pollination and seed dispersal that is often species specific and difficult to replace; and (5) knowledge of the reproductive attributes of timber tree species is surprisingly limited (Barrantes *et al.*, 1999; Quesada and Stoner, 2004). Hence TSDF ecological research directed to generate additional information along these lines will be important to help resolve the many technical uncertainties in surrounding sustainable forest management and habitat maintenance.

Appendix

List of Courses

1. Forest Ecology
2. Tropical Forest Silviculture
3. Tropical Forest Conservation
4. Conservation Biology
5. Tropical Forestry
6. Remote Sensing of Environment
7. Environmental Monitoring
8. Introduction to Global Change

Acknowledgments

This work was conducted with the support of the Inter-American Institute for Global Change Research (IAI) under their Collaborative Research Network Program (CRN2) TropiDry: Human and Biophysical Dimensions of Tropical Dry Forests (CRN2–021), funded by the US National Science Foundation (GEO-0452325). Logistical support and funding from the University of Alberta and Instituto Tecnológico de Costa Rica is recognized.

See also: Biodiversity and Ecosystem Services. Central America, Ecosystems of. Climate, Effects of. Conservation Biology, Discipline of. Deforestation and Land Clearing. Diversity and Conservation of Neotropical Mammals. Endangered Ecosystems. Fires, Ecological Effects of. Forest Ecology. Global Species Richness. Land-Use Issues. Mammals, Conservation Efforts for. Restoration of Biodiversity, Overview. South American Natural Ecosystems, Status of. Tropical Forest Ecosystems. Tropical Forest Regeneration

References

Abizaid C and Coomes OT (2004) Land use and forest fallowing dynamics in seasonally dry tropical forest in the southern Yucatán Peninsula, Mexico. *Land Use Policy* 21(1): 71–84.

Albuquerque UP, Andrade LHC, and Silva ACO (2005) Use of plant resources in a seasonal dry forest (northeastern Brazil). *Acta Botanica Brasilica* 19: 27–38.

Allen WH (1988) Biocultural restoration of a tropical dry forest. *BioScience* 38: 156–161. *American Forest Ecology Management* 245: 88–95.

Arroyo-Mora J, Sánchez-Azofeifa A, Kalacska M, Rivard B, Calvo-Alvarado J, and Janzen D (2005b) Secondary forest detection in a Neotropical dry forest

landscape using Landsat 7 ETM + and IKONOS imagery1. *Biotropica* 37(4): 497–507.
Arroyo-Mora J, Sanchez-Azofeifa A, Rivard B, Calvo-Alvarado J, and Janzen D (2005a) Dynamics in landscape structure and composition for the Chorotega region, Costa Rica from 1960 to 2000. *Agriculture, Ecosystems and Environment* 106: 27–39.
Balvanera P, Castillo A, and Martínez-Harms J (2011) Ecosystem services in seasonally dry tropical forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 259–278. London, UK: Island Press.
Barrantes G, Jiménez Q, Lobo JA, Maldonado T, Quesada M, and Quesada R (1999) Evaluación de los planes de manejo forestal autorizados en el periodo 1997–1999 en la Península de Osa. Cumplimiento de normas té cnicas, ambientales e impacto sobre el bosque natural. *Informe para Fundación Cecropia*
Bates BC, Kundzewicz ZW, Wu S, and Palutikof JP (eds.) (2008) Climate change and water. *Technical Paper of the Intergovernmental Panel on Climate Change*. Geneva: IPCC Secretariat.
Brandani I, Hartshorgn S, and Orians H (1988) Internal heterogeneity of gaps and species richness in Costa Rican tropical wet forest. *Journal of Tropical Ecology* 4: 99–119.
Brokaw N (1987) Gap phase regeneration of three pioneer tree species in a tropical forest. *Journal of Ecology* 75: 9–20.
Bullock SH (1995) Plant reproduction in Neotropical dry forests. In: Bullock SH, Mooney HA, and Medina E (eds.) *Seasonally Dry Tropical Forests*, pp. 277–303. Cambridge: Cambridge University Press.
Bye R (1989) Plantas utiles del bosque tropical caducifolio de Chihuahua, Mexico. In: Cervantes L and Bye R (eds.) *Programa y Resúmenes, Reunión Etnobotánica Ecológica Regional de Selvas Bajas Caducifolias (Bosque tropical caducifolia) y vegetación asociada en México*, pp. 19–20. Mexico City: Instituto de Biologia, Universidad Nacional Autonoma de Mexico.
Calvo-Alvarado J, McLennan BJ, Garvin T, and Sanchez-Azofeifa AG (2009) Putting conservation policies in context: The social dynamics of forest regrowth in Guanacaste, Costa Rica. *Forest Ecology and Management* 258: 931–940.
Castillo-Nunez M, Sanchez-Azofeifa GA, Croitoru A, Rivard B, Calvo-Alvarado J, and Dubayah O (2011) Delineation of secondary succession mechanisms for tropical dry forests using LiDAR. *Remote Sensing of Environment* 115: 2217–2231.
Ceballos G (1995) Vertebrate diversity, ecology and conservation in neotropical dry forest. In: Mooney HA, Bullock SH, and Medina E (eds.) *Seasonally Tropical Forests*, 450 pp. Cambridge, UK: University of Cambridge Press.
Chazdon R, Harvey C, Martínez-Ramos M, *et al.* (2011) Seasonally dry tropical forest biodiversity and conservation value in agricultural landscapes of Mesoamerica. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 210–235. London, UK: Island Press.
Chazdon RL and Denslow J (2002) Floristic composition and species richness. In: Chazdon RL and Whitmore TC (eds.) *Foundations of Tropical Forest Biology. Classic Papers with Commentaries*, pp. 513–522. Chicago, IL: University of Chicago Press.
Costa FRC and Magnusson WE (2003) Effects of selective logging on the diversity and abundance of flowering and fruiting understory plants in Central Amazonian forest. *Biotropica* 35: 103–114.
Costa FRC, Senna C, and Nakkazono EM (2002) Effects of selective logging on populations of two tropical understory herbs in an Amazonian forest. *Biotropica* 34: 289–296.
Cox PM, Harris PP, Huntingford C, *et al.* (2008) Increasing risk of Amazonian drought due to decreasing aerosol pollution. *Nature* 453: 212–215.
Denslow JS (1987) Tropical rainforest gaps and tree species diversity. *Annual Review of Ecology, Evolution, and Systematics* 18: 431–451.
Di Bella CM, Jobbaggy EG, Paruelo JM, and Pinnock S (2006) Continental fire density patterns in South America. *Global Ecology and Biogeography* 15: 192–199.
Ehrlich D, Lambin EF, and Malingreau JP (1997) Biomass burning and broad scale land-cover changes in western Africa. *Remote Sensing of the Environment* 61: 201–209.
Ewel J (1999) Natural systems as models for the design of sustainable systems of land use. *Agroforestry Systems* 45: 1–21.
FAO (2011) *Global Forest resources Assessment 2011*. Rome: Food and Agriculture Organization of the United Nations.
Fajardo L, González V, Nassar J, *et al.* (2005) Tropical dry forests of Venezuela: Characterization and current conservation status. *Biotropica* 37(4): 531–546.
Fischer AG (1960) Latitudinal variations in organic diversity. *Evolution* 14: 64–81.
Gentry AH (1982) Neotropical floristic diversity: Phytographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of Missouri Botanical Garden* 69: 557–593.
Gentry AH (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of Missouri Botanical Garden* 75: 1–34.
Gentry AH (1995) Diversity and floristic composition of neotropical dry forests. In: Mooney HA, Bullock SH, and Medina E (eds.) *Seasonally Tropical Forests*, 450 pp. Cambridge, UK: University of Cambridge Press.
Hanson P (2011) Insect diversity in seasonally dry tropical forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 86–99. London, UK: Island Press.
Hoekstra J, Boucher T, Ricketts T, and Roberts C (2005) Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters* 8: 23–29.
Holbrook M, Whitbeck J, and Mooney H (1995) Drought responses of neotropical dry forest trees. In: Mooney HA, Bullock SH, and Medina E (eds.) *Seasonally Tropical Forests*, 450 pp. Cambridge, UK: University of Cambridge Press.
Holdridge L (1967) *Life Zone Ecology*, 206 pp. San Jose, Costa Rica: Tropical Science Center.
IPCC (2007) Climate Change 2007: The Physical Science Basis. In: Solomon S, Qin D, Manning M, *et al.* (eds.) *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
Janzen D (1987) How to grow a tropical national park. *Experentia* 43: 1037–1038.
Janzen DH (ed.) (1983) *Costa Rican Natural History*. Chicago: University of Chicago Press.
Janzen DH (1986) Tropical dry forests: The most endangered major tropical ecosystem. In: Wilson EO (ed.) *Biodiversity*, pp. 130–137. Washington: National Academy Press.
Janzen DH (1988) Management of habitat fragments in a tropical dry forest: Growth. *Annals of Missouri Botanical Garden* 75: 105–116.
Joly AB (1970) *Conhea a vegetação brasileira, EDUSP e Polígono*. Brasil: São Paulo.
Kalacska M, Sánchez-Azofeifa A, Calvo-Alvarado J, Quesada M, and Janzen D (2004a) Species composition, similarity and diversity in three successional stages of tropical dry forest. *Forest Ecology and Management* 200: 227–247.
Kalacska M, Sánchez-Azofeifa A, Rivard B, *et al.* (2004b) Leaf area index measurements in a tropical moist forest: A case study from Costa Rica. *Remote Sensing of Environment* 91: 134–152.
Kalacska M, Sánchez-Azofeifa A, Rivard B, Caelli T, White P, and Calvo-Alvarado J (2007) Ecological fingerprinting of ecosystem succession: Estimating secondary tropical dry forest structure and diversity using imaging spectroscopy. *Remote Sensing of Environment* 108: 82–96.
Kalacska M, Sanchez-Azofeifa GA, Rivard B, Calvo-Alvarado JC, and Quesada M (2008) Baseline assessment for environmental services payments from satellite imagery: A case study from Costa Rica and Mexico. *Journal of Environmental Management* 88: 348–359.
Killeen TJ, Calderon V, Soria L, *et al.* (2007) Thirty years of land-cover change in Bolivia. *Ambio* 36: 600–606.
Laurance W, *et al.* (2002) Ecosystem decay of Amazonian forest fragments: A 22-year investigation. *Conservation Biology* 16: 605–618.
Linares-Palomino R, Oliveira-Fiho A, and Pennington T (2011) Neotropical seasonally dry forests: diversity, endemism, and biogeography of woody plants, 18–36. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 210–235. London, UK: Island Press.
Lobo JA, Quesada M, Stoner KE, *et al.* (2003) Factors affecting phenological patterns of Bombacaceous trees in seasonal forests in Costa Rica and Mexico. *American Journal of Botany* 90: 1054–1063.
Lobo JG, Barrantes G, Castillo M, et al. (2007) Effects of selective logging on the abundance, regeneration and long-term survival of *Caryocar costaricense* (Caryocaceae) and *Peltogyne purpurea* (Caesalpinaceae), two endemic timber species of southern Central.
Lugo A, Medina E, and Trejos-Torres C (2006) Botanical and ecological basis for the resilience of Antillean dry forests. In: Pennington T, Lewis G, and Ratter (eds.) *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity*, 484 pp. FL, USA: Biogeography and Conservation CRC Press.
Lüttge U (1997) *Physiological Ecology of Tropical Plants*, 384 pp. Heidelberg: Springer-Verlag.
Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, and Nobre CA (2008) Climate change, deforestation, and the fate of the Amazon. *Science* 319: 169–172.

Medina E (1995) Diversity of life forms plants in neotropical dry forest. In: Mooney HA, Bullock SH, and Medina E (eds.) *Seasonally Tropical Forests*, 450 pp. Cambridge, UK: University of Cambridge Press.

Meir P and Pennington RT (2011) Climate change and seasonal dry tropical forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 294–315. London, UK: Island Press.

Mercado LM, Belluin N, Sitch S, *et al.* (2009) Impact of changes in diffuse radiation on the global land carbon sink. *Nature* 458: 1014–1017.

Miles L, Newton A, DeFries R, *et al.* (2006) A global overview of the conservation status of tropical dry forests. *Journal of Biogeography.* 33: 491–505.

Mooney H (2011) Synthesis and promising lines of research on seasonally dry tropical forest. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 259–277. London, UK: Island Press.

Mooney HA, Bullock SH, and Medina E (1995) Introduction. In: Mooney HA, Bullock SH, and Medina E (eds.) *Seasonally Dry Tropical Forests*, 450 pp. Cambridge, UK: Cambridge University Press.

Murphy PG and Lugo AE (1986) Ecology of tropical dry forest. *Annual Review of Ecological Systematics* 17: 67–88.

Murphy PG and Lugo AE (1995) Dry forest of Central America and the Caribbean. In: Bullock SH, Mooney HA, and Medina E (eds.) *Seasonally Dry Tropical Forest*, 450 pp. Cambridge, UK: Cambridge University Press.

Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on earth. *BioScience* 51: 933–938.

Otterstrom SM, Schwartz MW, and Velazquez-Rocha I (2006) Responses to fire in selected tropical dry forest trees. *Biotropica* 38: 592–598.

Pennington T, Lewis G, and Ratter J (2006) An overview of the plant diversity, biogeography and conservation of Neotropical Savannas and seasonally dry forests: Plant diversity, biogeography and conservation. In: Pennington T, Lewis G, and Ratter (eds.) *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity*, 484 pp. FL, USA: Biogeography and Conservation CRC Press.

Peters C (2011) Economic Botany and Management Potential of Neotropical Seasonally Dry Forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 239–257. London, UK: Island Press.

Piperno DR and Pearsall DM (2000) *Origins and Agriculture in the Neotropics*, 400 pp. San Diego, CA: Academic Press.

Portillo-Quintero C and Sánchez-Azofeifa A (2010) Extent and conservation of tropical dry forests in the Americas. *Biological Conservation* 143(1): 144–155.

Putz FE, Sist P, Fredericksen T, and Dykstra D (2008) Reduced-impact logging: Challenges and opportunities. *Forest Ecology Management* 256: 1427–1433.

Quesada M, Sánchez-Azofeifa GA, Alvarez-Anorve M, *et al.* (2009) Succession and management of tropical dry forests in the Americas: Review and New perspectives. *Forest Ecology and Management* 258: 1014–1024.

Quesada M and Stoner KE (2004) Threats to the conservation of tropical dry forest in Costa Rica. In: Frankie GW, Mata A, and Vinson SB (eds.) *Biodiversity Conservation in Costa Rica Learning the Lessons in a Seasonal Dry Forest*, pp. 266–280. Berkeley: University of California Press.

Rodríguez JP, Balch JK, and Rodríguez-Clark KM (2007a) Assessing extinction risk in the absence of species-level data: Quantitative criteria for terrestrial ecosystems. *Biodiversity and Conservation* 16: 183–209.

Rodríguez JP, *et al.* (2007b) Globalization of conservation: A view from the south. *Science* 317(5839): 755–756.

Rzedowski J (1991) El endemismo en la flora fanerogámica Mexicana: Una apreciación analítica preliminar. *Acta Botánica Mexicana* 15: 47–64.

Sánchez-Azofeifa A, Kalacska M, Quesada M, Calvo-Alvarado J, Nassa J, and Rodríguez J (2005a) Need for integrated research for a sustainable future in tropical dry forests. *Conservation Biology* 19(2): 1–2.

Sánchez-Azofeifa A and Portillo-Quintero C (2011) Extent and drivers of change of Neotropical seasonally dry tropical forests. In: Dirzo R, Young H, Mooney H, and Ceballos G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 259–277. London, UK: Island Press.

Sanchez-Azofeifa GA, Quesada M, Cuevas-Reyes P, Castillo A, and Sanchez-Montoya G (2009) Land cover and conservation in the area of influence of the Chamela–Cuixmala biosphere reserve, Mexico. *Forest Ecology and Management* 258: 907–912.

Sánchez-Azofeifa GA, Quesada M, Rodriguez JP, *et al.* (2005b) Research priorities for neotropical dry forests. *Biotropica* 37(4): 477–485.

Sheil D and Van Heist M (2000) Ecology for tropical forest management. *International Forestry Review* 2: 261–270.

Skorupa P and Kasenenek M (1984) Tropical forest management: Can rates of natural tree falls help guide us? *Oryx* 18: 96–101.

Steininger MK, Tucker CJ, Ersts P, Killeen TJ, Villegas Z, and Hecht SB (2001) Clearance and fragmentation of tropical deciduous forest in the Tierras Bajas, Santa Cruz, Bolivia. *Conservation Biology* 15: 856–866.

Stoner K and Timm R (2011) Seasonally dry tropical forest mammals: Adaptations and seasonal patterns. In: Dirzo R, Young H, Mooney H, and Ceballos. G (eds.) *Seasonally Dry Tropical Forests: Ecology and Conservation*, pp. 100–121. London, UK: Island Press.

Timmermann A (1999) Detecting the nonstationary response of ENSO to greenhouse warming. *Journal of the Atmospheric Sciences* 56: 2313–2325.

Tosi Jr. JA and Voertman RF (1964) Some environmental factors in the economic development of the tropics. *Economic Geography* 40: 189–205.

Trejo I and Dirzo R (2000) Deforestation of seasonally dry tropical forest: A national and local analysis in Mexico. *Biological Conservation* 94: 133–142.

Tucker C, Pinzon J, Brown M, *et al.* (2005) An extended AVHRR 8–km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing* 26(20): 4485–4498.

Uhl C and Guimarães-Vieira IC (1989) Ecological impacts of selective logging in the Brazilian Amazon: A case study from the Paragominas region of the state of Pará. *Biotropica* 21: 98–106.

Victor P (1990) *Plantas medicinais: Comparacao da flora de quarto municipios de Pernambco.* PhD Dissertation, Universidade Federal de Pernambuco.

Wright SJ, Carrasco C, Calderon O, and Paton S (1999) The El Nino Southern Oscillation, variable fruit production, and famine in a tropical forest. *Ecology* 80: 1632–1647.

Wright JC, Sanchez-Azofeifa GA, Portillo-Quintero C, and Davies D (2007) Poverty and corruption compromises tropical forest reserves. *Ecological Applications* 17: 1259–1266.

Nest Parasitism

Scott K Robinson, University of Florida, Gainesville, FL, USA
Stephen I Rothstein, University of California, Santa Barbara, CA, USA
Brian D Peer, Western Illinois University, Macomb, IL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Coevolution Cycles of adaptation and counteradaptation that occur between interacting lineages.
Coevolutionary arms race The continuing bouts of coevolving defenses and counterdefenses that occur between hosts and parasites.
Conspecific brood parasites Brood parasites that lay their eggs in the nests of other individuals of the same species.
Egg rejection Host responses to parasitism that include ejection of parasitic eggs, abandonment of parasitized nests (usually with renesting), or a new cup built over a parasitized clutch.
Evolutionary equilibrium hypothesis The hypothesis that frequencies of acceptance of brood parasitism reflect an equilibrium between costs and benefits of host defenses against parasitism.
Evolutionary lag hypothesis The hypothesis that hosts lack defenses against brood parasites because the defenses have not yet had time to evolve.
Generalist brood parasites Species that parasitize many (up to more than 200) host species.
Gentes (singular gens) Lineages of cuckoos in which individual females specialize on a single host and lay mimetic eggs.
Interspecific brood parasites Brood parasites that lay their eggs in the nests of other species.
Specialist brood parasites Species that parasitize only one or a few host species.

Natural History

Brood parasites lay their eggs in the nests of other individuals, which then raise the parasitic young at the expense of part or all of their own brood. Brood parasitism was noted by Aristotle and even earlier (approximately 2000 BC) in India. Brood parasitism is best known in birds, approximately 1% of which are obligate interspecific brood parasites that only lay their eggs in the nests of other species. Interspecific brood parasitism, however, has also been documented in insects and fish. Conspecific brood parasitism (CBP), in which individuals facultatively lay eggs in nests of conspecific individuals, is more widespread.

Brood parasites have generated intense interest in the public and scientific communities. Brood parasites tend to be vilified in the media because of a human tendency to moralize about such things as killing baby birds (which parasitic birds often do) and because at least some may pose a conservation threat to some of their hosts. Among scientists, brood parasites have generated intense debate about the coevolutionary processes that may be responsible for the seemingly maladaptive acceptance of parasitic eggs by hosts.

Adaptations of Obligate Avian Brood Parasites

Brood parasites search for host nests, synchronize their laying with that of the host, and often remove a host egg from nests they parasitize. Brood parasitic eggs often have unusually thick eggshells, are usually small relative to the size of the parasite (but large relative to the size of the host), often mimic the coloration of the hosts' eggs, and have rapidly developing embryos. All these traits increase fitness of the parasite by reducing competition with the host nestmates and making parasitic eggs more difficult to detect and remove. Egg mimicry is most pronounced in specialist brood parasites. In the well-studied common cuckoo, the species as a whole is a generalist, but each individual female is a specialist member of a "gens." Any one region has only one to several coexisting gentes, but different arrays of gentes occur in different regions.

Brood parasitic nestlings also have a variety of mechanisms that enhance their ability to compete with host nestlings. Cuckoo nestlings have concave backs that they use to push out host eggs and nestlings, whereas several other brood parasites have specialized bill hooks that they use to kill nestmates. Some brood parasites apparently increase the amount of food hosts bring to them by mimicking the juvenile plumages or the complex mouth markings (the "gapes") of host nestlings or by having large mouths and intense begging behavior.

Because brood parasites do not have to engage in costly breeding activities such as incubation and nestling feeding, they often have more time and energy to devote to egg production. Some generalist brood parasites lay 40 or more eggs in a season; estimates for one tropical brood parasite suggest that more than 100 eggs per year are typical. Other brood parasites, however, may be only slightly more fecund than their hosts.

A Survey of Avian Brood Parasites

There are approximately 102 species of obligate avian brood parasites in five unrelated families (**Table 1**).

Cuculidae

The cuckoos are a diverse family that contains both parasitic and nonparasitic species. The Cuculidae has traditionally been

Table 1 Major groups of avian brood parasites

Taxon (family/subfamily)	*Number of species*
Old World cuckoos (Cuculidae)[a]	56
New World cuckoos (Neomorphinae)	3
Honeyguides (Africa and Asia) (Indicatoridae)	17
Viduine finches (Africa) (Ploceidae)	19
Cuckoo-finch (Africa) (Ploceidae)	1
Cowbirds (New World) (Icteridae)	5
Black-headed duck (South America) (Anatidae)	1

[a]Recent work indicates that at least one genus in this subfamily belongs in another cuckoo subfamily (see text).

recognized as an Old World subfamily, all of whose approximately 56 species are obligate brood parasites. Most species are host specialists that are relatively uncommon and only a few have been well studied (the common cuckoo and the great spotted cuckoo). Another traditionally recognized subfamily, the Neomorphinae or New World ground cuckoos, has 11 species, three of which are relatively rare obligate parasites. The latter tend to parasitize hosts that build domed nests and are poorly known overall. Nestlings of one species, *Tapera naevia*, have pincer-like bills that they use to kill host nestlings, a case of convergence on honeyguides (Indicatorinae). Recently published DNA sequence data indicate that at least one genus (*Clamator*) within the Cuculidae is more closely related to the Phaenicophaeinae, a subfamily of New World cuckoos. None of the currently recognized members of the Phaenicophaeinae are obligate parasites but some parasitize conspecifics and occasionally other species. These new DNA data place obligate parasitism within three groups of cuckoos, two of which have both parasitic and nonparasitic species, and indicate that parasitism evolved separately up to three times in cuckoos or that some parasitic cuckoo lineages reverted to parental behavior.

Indicatoridae

The honeyguides, a family named for the habit of one species that guides people and the African honey badger or ratel to beehives, has 17 obligately parasitic species. The species of *Indicator* and three related genera parasitize cavity nesters, whereas three species of *Protodiscus* parasitize open-cup nesters. Most species are poorly known and the eggs of many species have never been described. Some honeyguides have raptorial hooks on their bills, which they use to kill host nestlings. Recent studies suggest there are two host-specific lineages of greater honeyguides, those that utilize burrow nesting hosts (bee-eaters and a kingfisher species) and those that use cavity-nesting hosts (woodhoopoes and another kingfisher species). These lineages have eggs that match host eggs based on size and shape rather than color because hosts nests in dark cavities in which color is not readily detectable.

Viduine Finches

The approximately 19 species of obligate brood parasitic viduines occur in Africa and provide some of the best examples of nestling mimicry, particularly of the intricate gape patterns of their hosts. However, the exact function of this mimicry is unclear because when the color of the parasitic nestling gapes is experimentally altered it does not decrease likelihood of them fledging from host nests. Compared with other brood parasites, viduines exert relatively low costs on their hosts, which can usually raise mixed broods. Most *Vidua* species specialize on single species of grassfinches (Estrildidae), to which they may be related. Unlike most brood parasites, which are insectivorous, hosts of viduines are granivorous. Some species of viduines incorporate the songs of their hosts into their own songs, which are then used by females to choose males that have been raised by the same species.

Cuckoo-Finch (Anomalospiza imberbis)

This African species is closely related to viduine finches and parasitizes approximately 15 species of the Cisticolidae in grasslands. Similar to the common cuckoo there are apparently host-specific gentes that have evolved to circumvent host egg recognition. Recently reported DNA evidence suggests that parasitism evolved once in a single lineage that then gave rise to both *Anomalospiza* and the viduines.

Icteridae

The five species of New World brood parasitic cowbirds include the most generalized of all brood parasites, the brown-headed and shiny cowbirds, both of which parasitize more than 200 host species. Another species, the tropical giant cowbird, parasitizes almost exclusively colonial American blackbirds. Of the two remaining species, the bronzed cowbird is a host generalist (>100 host species), whereas the screaming cowbird is one of the most specialized of all brood parasites. It mainly parasitizes a single species of blackbird, which was formerly thought to be a species of cowbird. Unlike most brood parasites, cowbirds are often more abundant than many of their hosts and can pose a significant threat to populations of rare, localized host species. Most cowbirds benefit greatly from changes in the landscape caused by humans, which has enabled several species to expand their geographic ranges and increase their population sizes. As a result, they are coming into contact with new hosts that have not had recent contacts with brood parasitism. Some of these new hosts suffer extremely high levels of parasitism. Cowbirds appear to be extremely fecund; the shiny cowbird may lay more than 100 eggs a year. The invasion of North America by this cowbird may pose an additional threat to species that are not currently being parasitized by the brown-headed cowbird. The brown-headed cowbird may be the most intensively studied North American bird

Black-Headed Duck (Heteronetta atricapilla)

This species is the only obligately parasitic species with precocial young that feed themselves when they hatch. For this reason, this brood parasite has little effect on the nesting success of its hosts, which only have to incubate an extra egg or two. *Heteronetta* has been recorded parasitizing 14 different hosts species, such as gulls and ibises, as well as other ducks.

CBP

Recent studies have shown widespread CBP, especially in waterfowl, gallinaceous birds, and in a small number of

songbirds (generally species with rare nest sites or that breed colonially).

Other Brood Parasite Systems

Fish

An African catfish is the only fish species known to be an obligate brood parasite. It parasitizes mouth-breeding cichlids and apparently consumes all of its host's young while in the mouth of its foster parent. CBP and occasional parasitism of other species occur in many nest-building fish. Some host defenses such as increased nest guarding may have developed in response to this parasitism but egg recognition has not been reported.

Insects

A range of insects practice brood parasitism, especially members of the order Hymenoptera (bees, wasps, and ants). Some of these species are obligate brood parasites, such as certain ants that kill the queens of nonparasitic species and use the workforce of the entire colony to rear their own young. Other species are facultative parasites that victimize conspecifics or raise their own young. To our perception, some of these parasitic insects look remarkably different from their hosts so their acceptance by hosts seems inexplicable. However, olfaction is critical to these insects and it is likely that parasitic species have evolved chemical cues that mimic those of their hosts. Because both humans and birds perceive the world largely via vision and audition, and have relatively poor olfactory abilities, it is not surprising that parasitic birds have been subjects of much more research than have parasitic insects, even though the latter undoubtedly have many equally interesting examples of coevolved adaptations.

Coevolution between Brood Parasites and Their Hosts

Cuckoo–Host Systems

Cuckoo–host systems are much more highly coevolved than the other well-studied systems, those of cowbirds, and DNA divergence data show that they are also much older. Most cuckoo species use only one or a few host species. Some species; such as the Eurasian common cuckoo, might be called generalists when viewed over their entire ranges because they parasitize 50–100 or more host species. However, even these cuckoos are more properly viewed as specialists because they use only one to several host species in each region and individual females, or members of a gens, specialize on a single host species or several similar host species. In regions where several cuckoo species coexist, they tend to show little or no overlap in host use, which reduces competition for hosts. Most cuckoos are much less abundant than their hosts and probably have little effect on host population dynamics because they affect only a small proportion of their host populations.

Some cuckoo–host systems may have reached an evolutionary equilibrium, whereas others show strong evidence of ongoing coevolution and provide some of the strongest examples in vertebrates of microevolution that has occurred during ongoing research.

Adaptations of Cuckoos

Behavior

Cuckoos typically approach nests stealthily and may quickly drop their thick-shelled eggs onto host eggs to increase chances of breaking host eggs. Some parasitic cuckoos resemble hawks, whereas nonparasitic species do not. This resemblance or simply host recognition of cuckoos incites host aggression or may frighten hosts, and males in some species use this aggression to draw hosts away from nests. This male distraction is coordinated with and facilitates nest entry by the stealthy and drab-colored females in several cuckoo species. Nestlings of some cuckoos mimic the begging calls or plumage of host young. There is evidence that great spotted cuckoos revisit nests they have parasitized and destroy host eggs or nestlings if the parasitic egg has been removed by the host. Such a protection racket or "mafia effect" could select for acceptance of host eggs.

Cuckoos find nests by watching hosts from hidden perches and laying is synchronized with that of the hosts. Most cuckoos lay one egg per nest, except in a few species in which cuckoo nestlings do not kill those of the host. Cuckoos defend territories against other cuckoos and usually remove at least one host egg before laying one of their own.

Fecundity

Cuckoos lay 8–25 eggs per season, which is generally more than their hosts lay. Eggs are laid every other day.

Egg Adaptations

Cuckoo eggs usually mimic the coloration of their hosts' eggs. Cuckoos that parasitize many host species have individually specialized females, with each group of females or gens (plural gentes) mimicking the eggs of a single host species or a group of hosts with similar eggs. Egg mimicry is usually assumed to have arisen in response to host egg recognition and recently the first evidence of disruptive selection by the common cuckoo and one of its hosts has resulted in egg polymorphism in both species. An alternative hypothesis argues that egg mimicry evolved to keep a second cuckoo from identifying and removing a cuckoo egg from an already parasitized nest and then laying its own egg in the nest. However, it is not clear that more than one cuckoo attempts to parasitize the same host nest often enough for this to be a significant selective pressure. Most cuckoos parasitize hosts that are much smaller than themselves and have correspondingly small eggs for their body sizes. However, cuckoo eggs are usually slightly larger than those of the hosts. The few species that parasitize hosts larger than themselves do not have smaller eggs. The thick shells of cuckoo eggs may be adaptive because they reduce chances of breakage of cuckoo eggs during rapid laying, facilitate breakage of host eggs (thereby reducing competition with hosts), or make cuckoo eggs resistant to pecking by hosts. A counteradaptation to egg mimicry that some hosts have developed is variable egg types. In response, some cuckoo species have evolved an equivalent range of egg types; but hosts still seem to benefit because individual female cuckoos lay only one type and may not always parasitize a host female who lays the same type. The highly variable host eggs may also function in the context of conspecific parasitism

Nestling and Fledgling Adaptations

Nestlings in most cuckoo species have concave backs that they use to push host nestlings or eggs out of nests. Cuckoos that parasitize larger hosts do not show this eviction behavior but can compete successfully with host nestlings because they beg more loudly and incessantly, hatch earlier, and develop more rapidly. Mimicry of host begging calls occurs mainly, but not exclusively, in species that do not evict host nestlings.

Host Defenses Against Cuckoos

Defenses Against Adults

Hosts often respond aggressively to cuckoos, but this may actually provide a cue to cuckoos about their stage of the nest cycle and the proximity of nests. Mobbing cuckoos has both genetic and learned components. One possible advantage of mobbing cuckoos is that it may trigger ejection of parasitic eggs.

Egg Rejection

Hosts reject cuckoo eggs by abandoning parasitized nests or ejecting eggs directly. Small hosts are more likely to abandon (and renest), perhaps because ejection results in the incidental breakage of some of their own eggs and is therefore costly or because they are simply too small to eject cuckoo eggs. Egg rejection is most prevalent in hosts that have a long coevolutionary history with cuckoos. These hosts, which often show fine abilities to discriminate among egg types, are species that have intrinsic characteristics that make them suitable for parasitism, such as animal food appropriate for the nestling parasites and nests accessible to adult cuckoos. Species unsuitable as hosts, such as ones that have specialized nestling diets (e.g., seeds in some finch species), have presumably never been parasitized intensely in the past and generally lack egg recognition altogether. Among suitable hosts, egg recognition is more prevalent in species that are currently rarely parasitized than in species that are currently common hosts. This trend suggests that cuckoos shift away from using suitable hosts with well-developed egg recognition. It is possible that cuckoos have a dynamic system of host usage in which they repeatedly switch from hosts with well-developed defenses only to switch back to these hosts after their defenses have declined and defenses in other hosts have increased. Such a system could result in never-ending cycles of host switches and coevolution, but it is also possible that hosts retain high levels of egg recognition for long periods in the absence of parasitism, which would force cuckoos to become increasingly more specialized. Long-term retention is indicated by high levels of egg recognition in New World magpies and shrikes, the latter of which have retained egg recognition for more than 1 million years (My). Neither of these hosts are currently parasitized by any brood parasites but are descended from Old World ancestors that are cuckoo hosts. Populations of some suitable hosts that have only recently come into contact with cuckoos have apparently undergone rapid increases in egg ejection and discrimination, apparently because they possessed some recognition even in the absence of parasitism. This possession could be due to retention of recognition from past bouts of parasitism or to gene flow from parasitized populations of the same species.

Nestling Discrimination

Despite there being many species that have some degree of mimicry of host nestling appearance or begging calls, there are relatively few cases in which cuckoo hosts show outright rejection (i.e., removal) of nonmimetic nestlings. Recent studies indicate two hosts of little bronze-cuckoos remove the parasitic nestlings from their nests, and another host deserts nests with lone parasitic Horsefield's bronze or shining bronze cuckoo nestlings using their begging calls as cues to elicit abandonment. It is possible that nestling rejection is more widespread given the difficulty in detecting it. However, it is also possible that the selective value of nestling mimicry may also relate to ensuring that parasitic nestlings receive high-quality care from their hosts rather than avoiding outright rejection.

Cowbird–Host Systems

The five brood parasitic cowbirds include one of the most specialized of all brood parasites and the two most generalized of all brood parasites. The authors have chosen to highlight cowbird–host interactions in this article because there is a vast literature on cowbird–host systems and because these systems are quite different from cuckoo–host systems.

Cowbird Adaptations

Egg Adaptations

As described previously (*see* Icteridae), some cowbirds are extremely fecund, an apparent reflection of a trade-off between egg production and parental care. In captivity, brown-headed cowbirds can lay up to 77 eggs in the short temperate breeding season and field estimates of shiny cowbird productivity exceed 100 eggs during the much longer tropical breeding season. As in cuckoos, the eggs of the smaller cowbirds are small relative to the size of the cowbird but are larger, thicker shelled, rounder, and faster to develop than those of most of their hosts. Only the giant cowbird, which parasitizes hosts as large as itself, has normally proportioned eggs. The other smaller species usually parasitize smaller hosts. Egg mimicry has not been firmly established in any cowbirds, but cowbird egg shape and patterns have evolved, perhaps in response to diffuse selection by dominant hosts or groups of hosts. The two most generalized cowbirds, the shiny and brown-headed, have generalized spotted egg colorations that are similar to those of many passerines. Although cowbird egg coloration is not finely tuned to host egg coloration to the same degree as in cuckoos, it has undergone some shifts, presumably in response to host use, during the evolution of the cowbird lineage. Different cowbird species have generalized spotted eggs or immaculate unspotted ones; however, two species, the giant and shiny cowbirds, have both kinds of eggs. There appears to be some host specialization by female shiny cowbirds and their eggs resemble host eggs in shape and background color in at least two host species.

Behavioral Adaptations of Adults

Similar to cuckoos, cowbirds search for nests primarily by observing the behavior of hosts. Cowbirds approach nests stealthily and lay eggs very early in the morning before most hosts appear at their nests during the egg-laying period. In contrast, giant cowbirds appear to circumvent the defenses of colonial hosts by visiting colonies as a group in which males appear to distract hosts while females stealthily enter nests. Screaming cowbirds also visit their communally breeding hosts in groups, but males do not appear to help distract the hosts. Cowbird eggs are usually synchronized properly with host laying but improperly timed eggs that appear before the host eggs or after hosts have finished laying occur more commonly than in cuckoos. Although generalist cowbird species avoid parasitizing some species that are clearly unsuitable as hosts, such as doves, there is little evidence that they select the best host species in a community. Many cowbird eggs are seemingly wasted in nests of hosts that feed their nestlings inappropriate diets of seeds or fruit or that eject cowbird eggs. However, parasitism of the latter species may sometimes be adaptive for two reasons. First, some species that exercise egg recognition and ejection have low nest predation rates. Second, these species learn the appearance of their own eggs by imprinting on the first egg or eggs that they lay during their lives. Accordingly, naive hosts parasitized about the time they begin laying may come to learn cowbird eggs as their own eggs and may provide cowbirds with nests that have a relatively low likelihood of failing due to nest predation. Multiple parasitism (two or more cowbird eggs in a nest) occurs in approximately one-third of all nests parasitized by cowbirds, especially with some large hosts. Several female cowbirds sometimes lay in the same nest even though multiple parasitism usually reduces the success rate of cowbird eggs and many host nests receive no cowbird eggs. Female cowbirds are usually highly aggressive toward one another in habitats in which hosts occur (but not in which cowbirds feed), but there is still substantial overlap among home ranges. Cowbirds sometimes, but not always, remove host eggs by puncturing them. Egg removal is rarely done during the same nest visits on which females lay but may occur the afternoon before or later in the morning on which cowbird eggs are laid. Egg removal may enhance incubation of the parasitic egg and females sometimes consume the host egg. There are many observations of cowbirds depredating unparasitized host nests and there is evidence that cowbirds do this regularly to stimulate renesting, which will increase the future availability of host nests to be parasitized. However, some studies have found no evidence for this depredation–renesting hypothesis and the behavior may occur only in special circumstances. Cowbirds have also been experimentally shown to engage in "mafia" behavior in which they depredate clutches from which a previously laid cowbird egg has been removed. This behavior would greatly increase the costs of egg rejection by hosts and might slow the evolution of counterdefenses in this species.

Nestling and Fledgling Adaptations

In the specialized screaming cowbird, there is almost perfect mimicry of host plumage and vocalizations of its usual host species, the bay-winged cowbird (which is not closely related to the parasitic cowbirds despite its name). This mimicry does not seem to be essential for nestling care but instead seems to be an absolute requirement for receiving host care after the parasites fledge from host nests. Such perfect mimicry is lacking in the other, more generalized cowbirds. However, giant cowbird nestlings have white bills, like those of their oropendola hosts. Bills turn the usual cowbird-black after the fledglings become independent. Nestling cowbirds grow more rapidly and beg more loudly than the nestlings of many, but not all, of their hosts. They have relatively large mouths and show intraspecific variation in their gape colors, unlike the majority of nonparasitic nestlings. This gape variation may relate to host use in some unknown way. There is a recent report of a cowbird ejecting a host nestling, although the ejection could have been inadvertent.

Host Defenses against Cowbirds

Preventing Access to Nests

Many hosts react aggressively to cowbirds and colonial nesting may provide protection for some hosts. Aggressive responses to cowbirds, however, may reveal nest locations to cowbirds and be counterproductive for small hosts, which may not be able to drive cowbirds away from nests. Some hosts sit on their nests when approached by a cowbird, but there are records of cowbirds pulling hosts off their nests. Some hosts reduce parasitism by nesting in cavities or in dense vegetation, but even cavity nests are often parasitized in many species.

Egg Rejection

Cowbird host species can be divided into accepters, which accept cowbird eggs, and rejecters, which abandon parasitized nests, build a new nest on top of the old nest, or, most commonly, eject cowbird eggs. Within accepter species, nearly 100% of the individuals accept cowbird eggs. Rejecter species show nearly 100% rejection, although an increasing number of "intermediate" rejecters have been described, and there is little geographic variation in a species' responses to cowbird eggs. Birds that eject cowbird eggs either puncture them or grasp them and usually drop them at least several meters from the nest. Smaller hosts tend to puncture–eject, whereas larger hosts grasp–eject. Nest abandonment occurs most frequently when cowbird eggs are laid too early in a host's laying cycle and when hosts are too small to eject eggs (or to eject without breaking too many of their own eggs). Nest desertion appears to be triggered by encounters at the nest between hosts and cowbirds. In nearly all host species, detection of cowbird eggs in nests does not appear to play a role in eliciting desertion, even in hosts whose eggs are highly divergent from cowbird eggs. Some hosts may begin incubating early in the morning to prevent undetected cowbird parasitism. Most species that reject cowbird eggs occur in open habitats where they may have a long coevolutionary history with cowbirds, all of which forage in open areas in association with grazing ungulates. Hosts that eject cowbird eggs recognize their own eggs and remove cowbird eggs even if these outnumber their own eggs, unless they have misimprinted on cowbird eggs laid in their very first nest. Similar to Old World cuckoo hosts, cowbird hosts retain egg rejection behavior in the absence of brood parasitism and for long periods. Island scrub-jays, Bohemian waxwings, and great-tailed and boat-tailed grackles all reject parasitism in areas where they are not parasitized by brood parasites and have

possibly retained this behavior between 150,000 and 3 My. This supports a single trajectory model whereby hosts of brood parasites are becoming increasingly intolerant of brood parasitism. As more hosts evolve rejection, cowbirds will not be able to return to parasitizing former hosts because they retain defenses, and instead will be forced to evolve specialized adaptations, such as egg mimicry, for just a few species.

Modeling Host–Parasite Coevolution

There is little doubt that the adaptations and counteradaptations described in the section Coevolution between Brood Parasites and their Hosts result from coevolution between parasites and their hosts. Parasitic counterdefenses to host defenses must correspond to the characteristics of hosts, as in mimicry of host eggs. Thus, parasites cannot have effective defenses against more than a few hosts at any one time because hosts vary in key characteristics such as egg coloration. The traditional view therefore hypothesizes that parasites specialize on increasingly fewer hosts as a parasite–host system becomes older and more hosts evolve defenses. This view is supported by studies showing that (1) recently derived brood parasites tend to be more generalized than older species (e.g., cowbirds vs. cuckoos); (2) the costs of parasitism are often much higher than the costs of potential host defenses, which suggests that there is a continuing evolutionary arms race; (3) some systems in which hosts and parasites have only recently come in contact show either an absence or low levels of host defenses; (4) host defenses have evolved since recent contact with a brood parasite; (5) parasites have evolved egg mimicry in response to host defenses; and (6) cuckoo–host systems are dynamic as evidenced by host shifts in parasitic cuckoos. In accord with these generalizations, most Old World passerines in Africa and Eurasia, which are exposed to numerous species of an ancient group (the cuckoos), possess egg recognition. However, most New World passerines, which are exposed to a smaller and relatively younger group of parasites (the cowbirds) lack egg recognition.

The question then arises, why do so many hosts accept easily distinguishable parasitic eggs? Two hypotheses have been proposed: "evolutionary lag" and "evolutionary equilibrium." There is a burgeoning literature by proponents of each hypothesis. The evolutionary lag hypothesis is that these accepters have not yet had time to evolve antiparasite adaptations. Acceptance of parasitic eggs by some hosts in heavily parasitized, unproductive populations may also result from gene flow from populations in areas where cowbirds are absent and reproduction is high. The lag hypothesis is clearly applicable to situations in which parasites have begun to use host species with no recent history of exposure to parasitism but may even apply to some host species with long histories of parasitism because egg recognition/rejection is a totally new feature that may be difficult to evolve. In the evolutionary equilibrium hypothesis, hosts accept parasitic eggs because of the high costs of rejection. In this scenario, parasitism is costly, but trying to reject parasitism would be even more costly, i.e., accepters are making the best of a bad situation. If the "mafia" behavior described earlier (*see* Behavioral Adaptations of Adults) by cowbirds is widespread, it would increase the costs of rejection and might lead to an evolutionary equilibrium.

Distinguishing between lag and equilibrium has proven to be difficult. Testing either hypothesis requires measurements of the costs of rejection in species that accept, which has not yet been done effectively. Sometimes, likely costs for accepters can be estimated from similar species that eject. For example, the number of their own eggs that small ejecting hosts break when they eject cowbird or cuckoo eggs can be extrapolated to accepter hosts of similar size. These ejection costs, generally less than 0.5 host eggs per ejection, do not seem to be sufficiently high to make acceptance a more adaptive option because acceptance usually results in the loss of all the host's young for cuckoo hosts and for small cowbird hosts. There is evidence that rejecter hosts mistakenly eject their own eggs on occasion, but it is not clear if this cost is sufficient to outweigh the demonstrable costs of egg acceptance. Equilibrium may be most likely in species for which parasitism is relatively infrequent and in which parasitic eggs are not easily ejected because of the host's small bill size and may result in damage to the host's own eggs, or when they resemble those of the host, which greatly increases the probability of mistakes. Nevertheless, distinguishing conclusively between lag and equilibrium will require new experimental and phylogenetic studies.

Impacts on Host Population Dynamics

Brood parasitism is undoubtedly almost always costly for individual hosts, but the extent to which brood parasites threaten host populations is much less clear. Most brood parasites have little impact on host populations, presumably because coevolutionary processes reduce the frequency of parasitism and because most parasites are much less abundant than their hosts. In Europe, for example, less than 5% of most host nests are parasitized by the common cuckoo. One brood parasite, the giant cowbird, has even been hypothesized to benefit its hosts because its nestlings remove parasitic botfly eggs from its host nestmates.

Nevertheless, brood parasites can pose threats to host populations, especially generalist brood parasites and those that have undergone recent range expansions and population increases. In a few cases, cuckoos have expanded their ranges as a result of human-caused increases of early successional habitat. In areas recently colonized by cuckoos in Japan, levels of parasitism often exceed 40%, although in most of these areas the incidence of host defenses appears to be increasing rapidly, probably because these host populations already had some level of defense before being parasitized. Such rapid development of host defenses, however, appears to be absent from hosts of the brown-headed and shiny cowbirds. Furthermore, range expansions are much more pervasive in these two generalist cowbirds than in any cuckoo species. Cowbirds have undergone enormous population increases as the result of human activities, especially those associated with cattle. As a result, many host populations with no recent history of parasitism are being exposed to massive levels of brood parasitism. For some species, especially forest-nesting ones, widespread parasitism may be occurring for the first time in

the history of the species' lineage. Parasitism levels of many forest species nesting in the midwestern US, for example, exceed 80% with most nests receiving multiple cowbird eggs. Most newly exposed host lineages lack effective defenses. Several recent studies have linked large-scale population declines of songbirds with increasing levels of cowbird parasitism. Concerns about cowbird parasitism have stimulated an enormous amount of research since the late 1980s and three major symposia have been published recently on this subject.

This increased focus on cowbirds has resulted in pressure to trap and kill cowbirds to reduce their impacts on hosts. However, host species that have declined have experienced massive loss of habitat and increased rates of nest predation in addition to increased rates of cowbird parasitism. For example, most passerine species that breed in riparian habitat in the southwestern US have declined in the past century but so too has their habitat. Dams, water diversions, overgrazing, urbanization, and exotic plants have resulted in the degradation or loss of more than 90% of the riparian habitat present a century ago in the Southwest. Even if cowbirds are not the primary cause of some or all declines, they may now be exposing some reduced host populations to additional stresses that threaten the populations with extinction. Importantly, a cowbird population can be stable or grow even as it pushes a rare host to extinction because individual female cowbirds do not specialize on single host species.

Nevertheless, it is not clear whether cowbird parasitism threatens more than a few host species. Species that are most at risk are those with small geographic ranges that are wholly included within areas that contain abundant cowbird foraging habitat (pastures, feedlots, mowed grass, and bare soil) and that cannot be rescued by emigrants from more productive populations. This list includes several endangered species that are brown-headed cowbird hosts (Kirtland's warbler, black-capped vireo, least Bell's vireo, and southwestern willow flycatcher) and many species restricted to islands in the Caribbean that have recently been invaded by shiny cowbirds (yellow-shouldered blackbird and Puerto Rican vireo).

Most species in North America have very large geographic breeding ranges that include regions in which cowbirds are rare and restricted to areas near human habitations. These refugia from cowbird parasitism tend to occur in large, unfragmented habitats that may act as "sources" of surplus host young that can recolonize populations in more fragmented habitats in which levels of parasitization (and nest predation) are often very high. In these population "sinks" in fragmented habitats, levels of reproduction may be too low to compensate for adult mortality; that is, such populations can only be sustained by immigration from source habitats. Evidence for this source–sink scenario derives from well-documented sink populations in fragmented midwestern US forests. These populations are nevertheless relatively stable, probably due to immigration from populations in large, unfragmented forests in the region in which levels of both parasitism and nest predation are very low and reproductive success is sufficiently high for populations to act as sources. Such large-scale source–sink population dynamics can slow the evolution of host defenses because most young are being produced in areas in which cowbird parasitism is rare.

Cowbird Management

Many cowbird control programs have been initiated as a result of concern regarding several endangered songbird species. Controlling cowbirds, which is easily done because their social nature attracts them into traps that contain cowbirds that function as decoys or "bait," has become a multimillion dollar a year business in the American Southwest and there are also active programs elsewhere in North America and in the Caribbean. Local control efforts, in conjunction with habitat restoration, may have prevented the extinction of several endangered host species and races, and two endangered hosts have increased in population since cowbird control began. However, control programs for two other endangered hosts resulted in no increases in the sizes of host breeding populations, even though all control programs have resulted in increases in host reproductive output. Unfortunately, even when host populations increase, cowbird control must be done every year because cowbird removal has little or no year-to-year effect on the numbers of cowbirds that occur in an area due to high dispersal by cowbirds. Therefore, most workers view cowbird control as a temporary, stop-gap measure, although a US Fish and Wildlife recovery plan for the least Bell's vireo in California advocates cowbird control in "perpetuity." Larger scale control programs, such as killing cowbirds by the millions at winter roosts, have been suggested by some workers but do not seem justified given that regional breeding season control programs are effective in eliminating nearly all cowbirds from the ranges of endangered hosts. Such large-scale killing programs may also raise important ethical issues and some workers have argued that an undue emphasis on cowbirds may detract from more productive and more long-lasting management actions such as habitat restoration.

Ecology and Social Behavior of Brood Parasites

The lack of parental care in brood parasites potentially sets them apart from most other birds in their foraging ecology, mating systems, spacing behavior, and vocal development.

Foraging Ecology

Most brood parasites have unusual diets and foraging behavior. Honeyguides eat wax, many cuckoos eat hairy and toxic caterpillars, and cowbirds prefer to forage in short grass close to ungulates such as cattle, horses, or bison. In each of these cases, either the diet of the brood parasite would be difficult for nestlings to digest (wax and hairy caterpillars) or the foraging habitat may be so ephemeral (the proximity of ungulates) that it may not be available for an entire nesting cycle. It is not known, however, if these unusual foraging ecologies were precursors of brood parasitism or if they were made possible by the evolution of brood parasitism.

Mating Systems

Mating system theory predicts that birds freed from the needs of parental care should be promiscuous. Nevertheless, avian brood parasites show a remarkable array of mating systems,

including monogamy and resource-based polygyny. This variation in mating behavior has been attributed to the diverse ways in which parasites gain access to nests (some of which require cooperation between several individuals), to the great variety of foraging ecologies of brood parasites, and to such unusual features of individual host–parasite systems as strongly male-biased sex ratios in cowbirds. In general, there have been few detailed studies of the mating systems of brood parasites. The mating system of the brown-headed cowbird has received the most study among parasitic birds and has been described as promiscuous, polygynous, or monogamous. Promiscuity has been proposed because it is a common sight to see two or more males associating with a single female. However, studies of color-marked cowbirds that enabled researchers to determine which males and females mated together and associated with one another the most have indicated monogamy. A recent DNA fingerprinting study confirmed that monogamy prevails and even found that cowbirds have fewer matings outside of their pair bonds than do most nonparasitic songbirds. Field observations show that the mate faithfulness in cowbirds is due both to males guarding females from the advances of other males and to females being reluctant to mate with males other than their usual consort. Because males outnumber females, many males do not acquire a mate. A single study demonstrated promiscuity in an area in which extremely high cowbird abundances may have made it difficult for the birds to maintain pair bonds.

In one Asian honeyguide, males defend beehives and only allow access to wax if the females mate with them. In some viduine weavers and cuckoos, males display from prominent perches and are chosen by females presumably on the basis of mate quality, which may be indicated by song elements or plumage.

Spacing Behavior

Many brood parasites defend breeding areas that are rich in hosts. At least some cowbirds and many cuckoos defend home ranges against conspecifics. However, the home ranges are technically not true territories because several males and females often occupy the same area. In areas of home range overlap, some host nests are often parasitized by several female cowbirds. Cuckoos may occupy areas that are more mutually exclusive and thus this is closer to classic territoriality. In areas of low abundance, cowbirds may occupy mutually exclusive areas and thus may appear to be territorial.

Cowbirds often breed and feed on a daily basis in different areas that can be separated by as much as 15 km. After searching for nests in the morning, cowbirds commute to feeding sites in pastures, plowed fields, and in other habitats with short mowed grass. The uncoupling of breeding and feeding areas made possible by brood parasitism allows cowbirds to select a wide variety of breeding habitats, even if there are only a few foraging sites in a region. Some cuckoos also appear to have very large home ranges. Home range size is related to a bird's body size and trophic level, with predators needing especially large areas. Although cowbirds and cuckoos are not predators of other birds, adult breeding birds are a key resource for them, and they have home ranges similar in size to those of raptors that feed on adult birds.

Vocal Behavior

Brood parasites have provided some of the clearest examples of genetically hard-wired vocal behavior, but recent studies have also shown a key role for subsequent learning in the modification of vocalizations. Because cowbirds are raised by different species, it is not surprising that at least one of their songs, the perched song, is genetically programmed. Cowbirds, however, learn to modify this song in response to female preferences and interactions with other males and develop individual repertoires of approximately five perched song types. A second song type, the flight whistle, is almost totally learned. It occurs as discrete spatial dialects generally tens of kilometers in diameter, which are so variable that trained observers may fail to recognize as cowbird vocalizations whistles from dialects they have not yet experienced. These dialects are true examples of culture and show that considerable amounts of biodiversity within a species can be related to learned/cultural differences rather than to genetic differences. Unlike the majority of songbirds that complete their vocal development by the time they are 1 year old, male cowbirds in some regions do not master local versions of whistles and perched songs until they are 2 years old.

Vocal learning occurs among one other parasitic group. Male viduine weavers incorporate elements of their host's song into their repertoire, which presumably enables females to choose a male raised by the same host species. As such, viduine weavers show a high potential for sympatric speciation by cultural learning. Such assortative mating maintains coevolutionary adaptations such as the gape mimicry that is specific to just one host. Assortative mating is essential because gape appearance is due to paternal and maternal genes. In contrast, common cuckoo and cowbird males do not incorporate host songs into their repertoires, perhaps because egg appearance, which is essential for acceptance by their hosts, is determined solely by the maternal genotype.

CBP

Female birds sometimes lay their eggs in the nests of conspecifics. This behavior has been documented in more than 200 bird species and has also been documented in some insects and fish. Among birds, CBP is most prevalent among precocial species, such as ducks and gallinaceous birds, and among altricial species that breed colonially and/or use specialized nest sites, such as cavities. CBP is rare or absent in most songbirds, even some that are colonial, and it has been well studied in swallows.

In many cases, CBP occurs when individuals lose their nests during the laying period or are unable to obtain a territory or nest site. In these situations, individuals may be making the best of a bad situation. In some species, CBP may be an alternative reproductive strategy that increases reproductive success by spreading the risk among several nests, reducing competition within a female's own nest, and exploiting the parental care of others. Recent theoretical work suggests that CBP is a necessary

precursor to the evolution of interspecific brood parasitism. CBP is especially prominent in New World cuckoos of the genus *Coccyzus*, which tend their own nests and also parasitize conspecifics and congeners. In at least some species, such as the ostrich and some ducks, individuals may benefit from CBP because extra eggs or young dilute losses to predators.

Nest defenses against CBP parallel those used to combat interspecific brood parasitism. Hosts guard their nests, especially at high population densities in barn swallows, which may exert a significant cost of nesting colonially. Parasitic females of another swallow species, the cliff swallow, actually carry their eggs between nests. The rejection of parasitic eggs laid before the host has begun laying is a common adaptation against CBP but does not necessarily involve egg recognition. True egg recognition, in which a bird discriminates among egg types, is a rare response to CBP because it is difficult to evolve because parasitic and host eggs are similar in appearance, given that they are from the same species. Also, CBP is less deleterious to hosts than is interspecific parasitism because it does not involve adaptations for killing host nestlings or asymmetries in size and incubation period that accomplish the same end. However, egg recognition in response to CBP does occur in some birds such as ploceid weaver finches of Africa and Asia, which have extensive variation among the eggs of conspecific females. Indeed, it has been hypothesized that the extreme intraspecific egg variation of these weavers is an evolutionary response to CBP. Some species with typical amounts of variation in egg appearance may lay smaller clutches than they could potentially feed to leave room for parasitic eggs. If hosts cannot recognize parasitic eggs, then laying a smaller clutch might avoid starvation of most or all nestlings when parasitism does occur. American coots also display remarkably refined abilities to recognize conspecific parasitism, possibly even by counting eggs. Coots are indeterminate layers and rejecter females not only reject the parasitic eggs by burying them or placing them at the periphery of the nest, but they also lay the normal number of eggs, whereas accepter females lay fewer eggs in response to the addition of parasitic eggs. CBP may actually affect population sizes. In species with frequent CBP, fitness may be reduced at high population densities, which may cause populations to be cyclic. Nests of wood ducks that are placed too close together are subject to extreme levels of CBP, which can reduce population fitness.

Conclusions and Research Needs

Brood parasitism raises fascinating questions about coevolution and conservation. Studies of brood parasitism provide some of the strongest evidence of microevolution yet documented in vertebrates. The extreme mobility of cowbirds and their potential threats to host populations illustrate the importance of landscape-level processes in conservation biology. Cowbird song development has become a model system showing that both genetic and environmental factors are important to the development of behavior in general. Despite these and other lessons learned from studies of brood parasitism, there are many unanswered questions that are vital to our understanding of this subject.

We still have much to learn about topics such as (1) why so many hosts of the cowbird in North America accept parasitism and whether an evolutionary equilibrium can account for this seemingly maladaptive behavior (2) the frequency of recognition errors in hosts with rejection behavior, (3) reasons for the lack of widespread parasite nestling recognition by hosts, (4) the frequency of "mafia"-like behaviors and nest predation by brood parasites, (5) how brood parasites choose from available hosts, (6) whether or not cowbirds pose a significant threat to populations of widespread host species, and (7) the genetic mating systems of brood parasitism. The recent invasion of the USA by a new generalist brood parasite, the shiny cowbird, also offers an excellent opportunity to study its interactions with other cowbirds and its new hosts. There is only indirect evidence that shiny cowbirds have begun parasitizing hosts in the USA and their invasion has largely stalled. Finally, most host–parasite systems remain very poorly known, especially those of the New World cuckoos, many Old World cuckoos, and some old honeyguides. Some members of the two latter groups are so poorly known that there is no direct evidence that they are parasitic. However, they are assumed to be so because closely related species are known to be parasitic.

Appendix

List of Courses

1. Ornithology
2. Evolution
3. Ecology
4. Conservation Biology

See also: Birds, Biodiversity of. Coevolution. Parasitism. Population Dynamics. Species Coexistence

References

Davies NB, Bourne AFG, and Brooke M de L (1989) Cuckoos and parasitic ants: Interspecific brood parasitism as an evolutionary arms race. *Trends in Ecology and Evolution* 8: 2–4.

Johnsgard PA (1997) *The Avian Brood Parasites: Deception at the Nest*. New York: Oxford University Press.

Ortega CP (1998) *Cowbirds and Other Brood Parasites*. Tucson: University of Arizona Press.

Payne RB (1997) Avian brood parasitism. In: Clayton DH and Moore J (eds.) *Host–Parasite Coevolution: General Principles Avian Models*, pp. 338–369. New York: Oxford University Press.

Payne RB (2005) *The Cuckoos*. Oxford: Oxford University Press.

Rothstein SI and Robinson SK (1998) *Parasitic Birds and their Hosts: Studies in Coevolution Oxford Ornithology Series*. Oxford: Oxford University Press.

Neutral Theory and Beyond

James O'Dwyer, Santa Fe Institute, Santa Fe, NM, USA
Ryan Chisholm, Smithsonian Tropical Research Institute, Panama, Republic of Panama

© 2013 Elsevier Inc. All rights reserved.

Glossary

Dispersal limitation The degree to which the location of an organism is constrained by the location of its parent through dispersal.
Equalizing mechanisms Biotic or abiotic forces giving rise to equal net per-capita population growth rates of individuals of different species in a community.
Local community A set of individual organisms in a region of space that is typically small, in which local speciation is assumed to have a negligible impact on community dynamics, and which may be a part of or connected to a much larger metacommunity.
Metacommunity A set of individual organisms in a region of space that is typically large and may comprise many local communities.
Spatially explicit model A model with an explicit spatial structure, so that individuals are characterized by their spatial location.
Spatially implicit model A model with limited spatial structure, and in which individuals are not characterized by an explicit spatial location.
Speciation The evolutionary process whereby new species arise in a biological system.
Species abundance distribution (SAD) The number of species in a community as a function of abundance.
Species-area relationship (SAR) The number of distinct species observed as a function of sample area.
Stabilizing mechanisms Biotic or abiotic forces that introduce a tendency for species to increase when rare, giving rise to niches and possible stable coexistence of multiple species in a community.

The neutral theory of biodiversity was introduced in its modern form by Hubbell in 2001, and has quickly become one of the most vigorously discussed and highly cited frameworks in theoretical ecology (Rosindell *et al.*, 2011). The development of neutral theory was motivated by the goals of explaining the apparent coexistence of large numbers of species in high-diversity communities, and the near-universality of certain large-scale patterns of biodiversity. Its central results characterize how the ecological processes of drift, dispersal, and speciation impact the structure of communities. This has allowed for the development of mathematically rigorous predictions for patterns like the species abundance distribution and the species-area relationship that are quantitatively consistent with patterns observed empirically in many species-rich communities. The rapid development of neutral models in the last decade has also seen the formulation of nonneutral models that retain some of the original neutral theory framework, but either refine or extend the scope of its predictions.

Introduction

The neutral theory of biodiversity is an ecological theory asserting that many observed biogeographical patterns in nature can be explained by simple models that ignore differences between species. Despite the apparently radical assumption of species equivalence, the predictions of neutral models often match empirical data. Stephen Hubbell's seminal monograph (Hubbell, 2001) on neutral theory generated a wave of creativity in community ecology, and there is now a rich assortment of theoretical tools that can be used to generate neutral predictions and compare these predictions to empirical data. This ongoing research program informs scientific understanding of when and where neutral processes can be the key determinants of biodiversity patterns in nature.

The remainder of this introduction outlines the motivations behind neutral theory, defines neutrality and describes key characteristics of neutral models. The section History gives an overview of the history of neutral theory. The section Neutral Models describes some of the models that have been used to develop neutral theory and reviews the performance of these models' predictions against biodiversity data. The section Nonneutral Extensions looks at some extensions of these models, which have in various ways included the effects of species differences while retaining some of the predictive power of neutral models. The section The Future of Neutral Theory discusses the future of neutral theory, with a particular focus on conducting more rigorous and extensive tests of neutral predictions. The article concludes with a summary and an examination of the consequences of neutral theory for scientific understanding of ecological communities.

Why Neutrality?

The ultimate goals of neutral theory, and of community ecology more generally, are to increase scientific understanding of species assemblages and to predict or manage their future behavior. How can hundreds of different tree species apparently coexist in a single 50-ha forest plot? What processes structure species abundance distributions (SADs) and species–area relationships (SARs) at different spatial scales? What governs the temporal dynamics of ensembles of species? There are both pure scientific benefits to developing this understanding, and practical conservation and management benefits as well. At what rate will species be lost with habitat destruction?

How should conservation reserves be designed to maximize biodiversity? The development of quantitative models like neutral theory can address these questions too.

The classical ecological perspective is that species coexist through their differences: each species has its own niche, and any two species too similar in form and function compete, with only the fitter surviving. The formalization of this competitive exclusion principle and the subsequent development of niche theory have given ecologists an intuitive view of how communities are structured, often borne out by empirical evidence.

There are, however, two clear challenges to this niche-structured picture. The first is empirical: in highly diverse communities it is difficult to reconcile the number of species with the number of niches. The second difficulty is theoretical: deriving predictions for classical ecological patterns like the SAR and SAD from niche theories is not only difficult, but also seems contingent on the complexities of a given community. The remarkable similarity of these and other patterns across multiple communities has led to many statistical models, but deriving these patterns from mechanistic processes in niche-based models has proved extremely difficult. Two questions therefore motivate neutral theory:

1. Is there an alternative general explanation for species coexistence that does not rely on species occupying distinct niches?
2. Is it possible to generate quantitative predictions of biodiversity patterns (such as SADs and SARs) from a mechanistic theory?

The neutral theory of biodiversity provides possible answers to these questions. To address the first question, neutral theory steps away from the idea of potentially indefinite coexistence and toward transient coexistence, also known as co-occurrence. If observed species richness arises from a simple balance between speciation and extinction, then coexistence no longer depends on each species having a distinct role in the community. Neutral theory answers the second question directly by accurately predicting certain biodiversity patterns from simple mechanistic models. This provides evidence that similar mechanisms may be responsible for generating biodiversity patterns across different communities, despite large inter-community variation in species identities, interactions, and environmental conditions. It is important to note that neutral theory does not deny the existence of nonneutral processes; it postulates that they are not important for predicting community structure at certain spatial and temporal scales.

What is Neutrality?

A clear definition of neutrality is an essential precursor to a broader discussion of the theory. An ecological community is neutral if its dynamics are independent of species' labels. This definition can be broken down into two conditions: first, the community dynamics are symmetric at the species level, so that demographic rates are equal for all species. This condition is fundamental to the idea of neutrality and formed Hubbell's original definition. However, a definition based only on this condition includes models in which diversity is maintained by strong intraspecific negative density dependence – models that would be considered niche models by most ecologists (e.g., Volkov *et al.*, 2005). Thus, a second condition for neutrality is that intraspecific density dependence is equal to interspecific density dependence. This means that there is no intrinsic 'rare species advantage' – a stabilizing mechanism that would differentiate the dynamics of individuals belonging to rare species from those belonging to abundant species. In effect, most neutral theories are symmetric not only at the level of species, but also at the level of individuals.

Key Characteristics of Neutral Models

Although the definition of neutrality provides an overarching guiding principle, specific neutral models are needed to generate predictions and test the theory, and the definition allows scope for substantial variation among these models. The three fundamental ecological processes that most neutral models include are stochastic drift, speciation, and dispersal. The authors give a schematic picture of a neutral model in **Figure 1**.

Stochastic Drift

In typical neutral models, stochastic birth and death processes approximate the complex, multifaceted and typically unpredictable fine-scale processes that occur in real ecosystems. The demographic stochasticity that governs the dynamics of such models is also referred to as drift. Drift is expected to drive the dynamics of species abundances when stabilizing forces, characteristic of deterministic niche theories, are weak. Some neutral models constrain the drift process by enforcing a zero-sum rule at the community level, whereas others do not.

Speciation

The definition of neutrality does not make any reference to the fundamental ecological process of speciation, but speciation may be introduced to balance extinction arising from stochastic drift, facilitating the maintenance of diversity. Point speciation, the mode of speciation in most published neutral models, allows a small but finite probability that an individual birth event will give rise to a new species. This is a schematic representation of true speciation processes, and more recent developments have added a variety of ways to model speciation in neutral theory: 'random fission' is an analog of allopatric speciation, whereas 'protracted' speciation distinguishes incipient from well-established species.

Dispersal

Neutral models treat dispersal in different ways. Hubbell's original approach has a two-scale spatial structure (*see* Neutral Models), in which dispersal is treated as immigration from a metacommunity into a local community. In other neutral models, dispersal is modeled as a spatially explicit process, in which case it is possible to give a more realistic treatment of neutral dynamics on a contiguous landscape.

History

The intellectual heritage of neutral ecology has two distinct strands: MacArthur and Wilson's theory of island

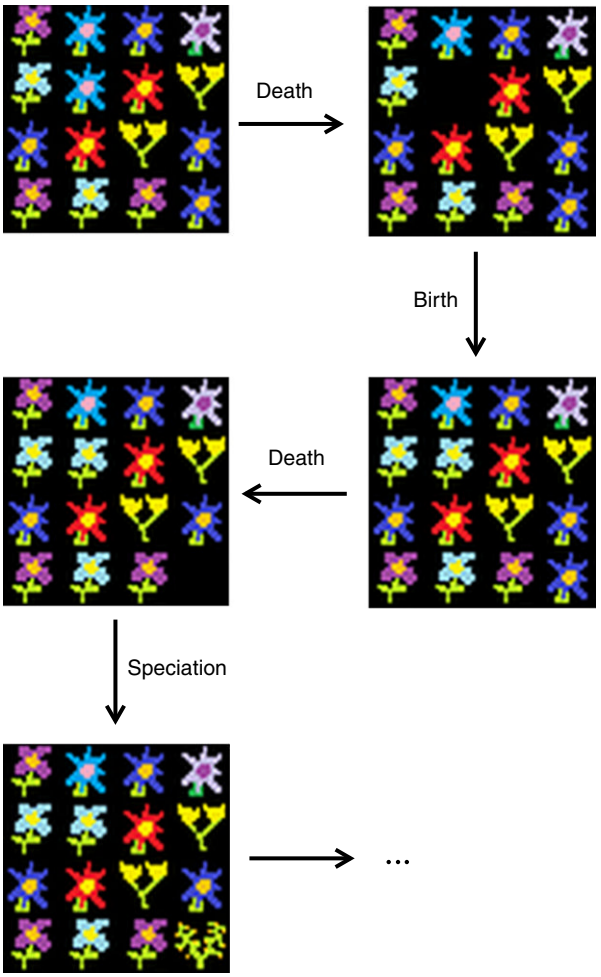


Figure 1 Schematic diagram of the processes of birth, death, speciation and dispersal, shown on a spatially explicit lattice. The death of an individual opens a space on the lattice, which is followed by a birth event: the space is colonized by the offspring of an existing organism. In spatially explicit models, dispersal limitation constrains the likelihood that a given organism's offspring will successfully colonize the space, as a function of separation. The lattice structure here imposes a zero-sum constraint, so that the total number of individuals is fixed. Subsequent models have relaxed this zero-sum constraint, so that the total number of individuals fluctuates around a fixed average.

biogeography (MacArthur and Wilson, 1967) and the neutral theory of molecular evolution (Kimura, 1968). Conceptually, the dispersal/immigration aspects of neutral ecology come from the former, whereas the speciation and drift aspects come from the latter.

Island biogeography is a seminal conceptual framework in theoretical ecology, which aims to explain variation in species richness on islands. Island biogeography predicts that steady-state species richness arises from a balance between stochastic extinction of species on the island, and immigration of new species from a mainland reservoir of biodiversity. The theory results in an explanation of the relatively low observed species richness on islands of a given size relative to equal-sized portions of a contiguous habitat, and has inspired the development of more rigorous theory in conservation biology and the framework for metapopulation theory. It also represents a departure from niche-based approaches to understanding community assembly. The obvious similarities between island biogeography and neutral theory are that species are treated neutrally and that random dispersal is the dominant driving force in determining local species richness. However, the dynamics of individuals are not explicitly considered in the island biogeography framework, making it difficult to go beyond species richness to make predictions for patterns that depend on species abundances.

The neutral theory of molecular evolution is the second strand of ecological neutral theory's intellectual heritage. It is also the origin of much of its mathematical framework, with species exchanged for alleles, speciation events exchanged for mutations, and stochastic drift in abundances analogous to genetic drift. The neutral theory of molecular evolution posits that a majority of evolutionary changes are due to stochastic drift of selectively neutral mutations; the neutralist– selectionist debate in molecular evolution parallels the current debate among advocates of niche and neutral approaches to community ecology.

The genesis of neutral ecology came with early attempts to synthesize these two disparate branches of biological theory (Caswell, 1976; Hubbell, 1979; Bell, 2001). More than two decades later, neutral ecology gained prominence with the publication of 'The Unified Neutral Theory of Biodiversity and Biogeography' (Hubbell, 2001), which presented mathematical and numerical analyses of spatially implicit and spatially explicit neutral ecological models and made quantitative predictions for SADs, SARs, and other biogeographical patterns. The primary focus was on a specific model that unifies local community and metacommunity scales: diversity in the metacommunity is maintained through a balance of extinction and speciation, whereas diversity in the (semi-isolated) local community is maintained through a balance of local extinction and immigration from the metacommunity (*see* Neutral Models).

The recent history of neutral theory has seen developments on a number of fronts. Hubbell's original model has been extensively tested, and maximum likelihood techniques have added a rigorous backbone to the estimation of neutral model parameters. At the same time, new mathematical tools have been developed to generalize the way dispersal and speciation are implemented in neutral models. Nonequilibrium field theory has provided a framework for spatially explicit neutral theory in the mathematical language of many-body physics. In addition, more biologically realistic speciation modes have generalized Hubbell's original point speciation model, with, for example, random fission speciation borrowing from fragmentation theory.

Neutral Models

Hubbell's Spatially Implicit Neutral Model

Hubbell's original formulation of neutral theory is often referred to as 'spatially implicit', because although the model is not explicitly spatial, there is a degree of spatial structure introduced through the separation of the local community from the metacommunity. The dynamics of the metacommunity operate as follows: in each timestep, a randomly

chosen individual dies and is replaced (with probability $1-v$) with the offspring of another randomly chosen individual or (with probability v) by a completely new species arising from a speciation event. The metacommunity is assumed to be unchanging on timescales relevant to the local community; the local community can then be thought of as a dispersal-limited sample from this larger, static metacommunity. In the local community, in each timestep, a randomly chosen individual dies and is replaced (with probability $1-m$) with the offspring of another randomly chosen individual in the local community or (with probability m) by an immigrant from the metacommunity.

SADs

The most extensively analyzed aspect of the spatially implicit neutral model is the SAD (Hubbell, 2001; Volkov *et al.*, 2003; Alonso and McKane, 2004). Let the probability that a species has abundance n_t at time t be $P(n_t)$. In the discrete-time formulation of the model, the metacommunity is governed by the following master equations (there is an analogous continuous-time formulation that is not presented here):

$$P(n_{t+1}=n_t-1)=\frac{n_t}{J_m}\left\{v+(1-v)\frac{J_M-n_t}{J_M-1}\right\}$$

$$P(n_{t+1}=n_t+1)=\frac{J_M-n_t}{J_M}(1-v)\frac{n_t-1}{J_M-1}$$

$$P(n_{t+1}=n_t)=1-P(n_{t+1}=n_t-1)-P(n_{t+1}=n_t+1)$$

where J_M is the metacommunity size and v is the speciation rate. The first equation is the probability that a species with abundance n_t will decrease by one in one timestep: the first factor on the right-hand side is the probability that an individual of the focal species dies; the second factor is the probability that the newly born individual is either a new species or the offspring of an individual not belonging to the focal species. The second equation is the probability that a species with abundance n_t will increase by one in one timestep: the first factor on the right-hand side is the probability that an individual not of the focal species dies; the second factor is the probability that a speciation event does not occur; and the third factor is the probability that the newly born individual is of the focal species. The third equation is the probability that the species' abundance does not change in one timestep, and is obtained from the normalization condition. This master equation can be solved for the steady-state SAD and the expected number of species with abundance n is given by

$$\langle\phi_n\rangle_M=\frac{\theta}{n}\frac{\Gamma(J_M+1)}{\Gamma(J_M+1-n)}\frac{\Gamma(J_M+\theta-n)}{\Gamma(J_M+\theta)}$$

The composite parameter $\theta=J_Mv/(1-v)$ is referred to as the fundamental biodiversity number, and $\Gamma(x)$ is the gamma function. In high-diversity communities, a very good asymptotic approximation to the SAD is given by

$$\langle\phi_n\rangle_M\sim\frac{\theta}{n}(1-v)^n$$

This is the log-series distribution, first proposed by Fisher as a mathematical representation of SADs. The appearance of this classic distribution in the neutral metacommunity is consistent with observations of log-series distributions in nature at large scales.

The equations governing the dynamics of the local community have a similar structure to those for the metacommunity and, while the mathematical analysis is more complicated, lead also to an analytical solution for the SAD, referred to as the zero-sum multinomial or the dispersal-limited multinomial. The key result is that the dispersal-limited multinomial resembles a log-series distribution for high levels of immigration and a lognormal distribution for intermediate levels of immigration. Again, this is qualitatively consistent with data, because lognormal-shaped SADs are ubiquitous at small scales in nature. Moreover, the neutral SADs have provided good quantitative fits to empirical SADs from tropical forest tree communities and other ecosystems. The authors show one of these fits in **Figure 2**, where the dispersal-limited multinomial is fitted to data from six different censuses of trees on Barro Colorado Island, Panama.

This good correspondence between neutral SADs and those from empirical data has been celebrated as perhaps neutral theory's greatest achievement. An important caveat to these successes is that they often rely on curve-fitting procedures and qualitative visual comparisons. The fitted parameters are the fundamental biodiversity number (θ) and the immigration rate (m) from the metacommunity to the local community. Multivariate techniques can quantify the probability that a given set of parameters will give rise to an observed SAD, and best estimates of parameters can then be determined from maximum likelihood (Etienne, 2007). In principle, these techniques facilitate more robust fitting exercises, but in applications to empirical data different combinations of

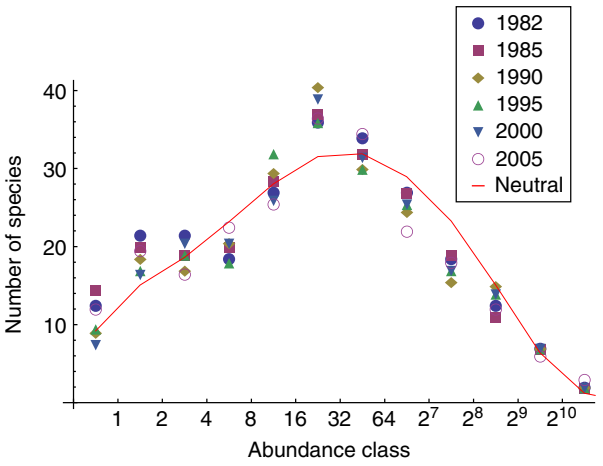


Figure 2 Comparison of the observed and theoretical neutral species abundance distribution of forest trees ($\geq$10 cm diameter at breast height) in the 50-ha permanent plot on Barro Colorado Island. The theoretical distribution from the spatially implicit neutral model is constrained only by the observed population size, species richness and mean dispersal distance; the line was not curve-fitted to the data. Each abundance class contains species with abundances falling between the values on the tick marks on either side. Species with exactly 2^x individuals are split between adjacent classes.

dispersal and speciation rate give rise to similarly plausible SADs (i.e., the maximum likelihood surfaces have broad plateaus or multiple peaks). This suggests that even the two free parameters of spatially implicit neutral theory may overfit the SAD.

One way to address the overfitting is to estimate the parameters of spatially implicit neutral theory independently from empirical data, and methods for estimating the immigration parameter from dispersal data are now available (Chisholm and Lichstein, 2009). This approach leaves the spatially implicit model with only one free parameter, the fundamental diversity number, which can be constrained on the observed species richness. Using these methods, the spatially implicit model continues to yield a good fit to the empirical SAD from Barro Colorado Island. Another way to address the overfitting problem is to incorporate other empirical data, such as phylogenetic trees, into the fitting process (Jabot and Chave, 2009).

Beyond the SAD

It is possible to make a range of other predictions from Hubbell's original model, most of which have been less explored than the SAD, and in some cases are less consistent with empirical data.

The Species-Area Relationship (SAR) quantifies the increase in the expected number of observed species with increasing sample area, and is another classic and widely investigated pattern in ecology. Predictions of SARs from the spatially implicit model do not generally match empirical SARs at large or intermediate scales, which have most often been modeled as a power-law dependence of species richness on area.

Most previous tests of neutral theory have been based on its static predictions, but its dynamic predictions can facilitate a potentially much broader set of tests. On short timescales, species-turnover does not appear to be consistent with the predictions of neutral models (Leigh, 2007). On geological timescales dynamic predictions of neutral models also appear to be unsuccessful. Analyses based on the best fit parameters of SADs for tropical forests apparently overestimate species lifetimes, and community phylogenies are more balanced than neutral predictions (Jabot and Chave, 2009). Reconciling neutral models with these data may be possible by fine-tuning parameters, such as the immigration rate (Jabot and Chave, 2009), but it seems likely that a predictive dynamic theory will need to incorporate nonneutral mechanisms.

Alternative Neutral Models

The spatially implicit neutral model, while the focus of much research, is one member of a vast set of potential neutral models. As outlined in the Introduction, other neutral models differ from the spatially implicit model and from one another primarily in their characterization of the dispersal and speciation processes.

Spatially Explicit Neutral Models

Spatially explicit models in ecology describe the abundance and distribution of species across a spatial landscape, in contrast to nonspatial models, which describe only total abundances. With regards to neutral theory, the advantages of a spatially explicit description are threefold. Firstly, spatially explicit models facilitate more rigorous confrontations with empirical data because the process of dispersal is treated mechanistically. This permits a more explicit parameterization of dispersal: parameters can be estimated directly from field data, whereas the more abstract immigration of spatially implicit neutral theory has most often been treated as free parameter to fit neutral predictions for the SAD. Secondly, spatially explicit models, due to their increased biological realism, can potentially provide more accurate predictions of diversity patterns. This is especially true for patterns with a strong spatial component, such as SARs and spatial scaling of SADs. Third, spatially explicit models can predict a wider range of patterns, such as similarity-distance decay relationships, which usually cannot be predicted at all from nonspatial or spatially implicit models.

The potential for more, and more accurate, predictions from spatially explicit models is promising, but in practice the increased complexity of spatially explicit models makes it difficult to derive analytical results. Some models extend spatially implicit neutral theory by including a network of interconnected local communities, and analytical results are readily obtained in this framework. But given that the locations of individual organisms are not tracked in these models, they are still not truly spatially explicit. Mathematical techniques from spatial population genetics have been successfully adapted to derive analytical results for similarity-distance decay relationships from spatially explicit neutral models (Condit *et al.*, 2002). These predicted relationships closely match empirical data from tropical forest tree communities at scales beyond the typical dispersal distance. More recently, the formalism of nonequilibrium field theory (O'Dwyer and Green, 2010) has pointed toward a more general mathematical framework for spatially explicit neutral models, and approximate analytical results successfully reproduce the classic three-phase SAR observed in nature. This three-phase relationship encompasses a linear sampling phase at small scales, an approximate power-law relationship at intermediate scales, and a linear relationship at scales beyond the typical species range. It should be noted that these techniques are still in the early stages of their development, and future improvements will need to go beyond, or justify, the approximations currently necessary to obtain analytical results.

The difficulty in obtaining analytical results for spatially explicit models has also spurred the development of simulated, lattice-based spatially explicit neutral models. Drawing on the techniques of coalescent theory in population genetics, SARs and the spatial dependence of the SAD can be simulated for a wide range of speciation rates and dispersal parameters. The resulting relationships are qualitatively (and in some cases quantitatively) in agreement with empirical data, with a three-phase SAR again emerging (Rosindell and Cornell, 2007).

Currently a major limitation with testing spatially explicit neutral models is simply a lack of appropriate data at sufficiently large scales. For instance, the tropical forest data set from Barro Colorado Island can be used to generate small-scale SARs, but intermediate- and large-scale SARs are less well determined empirically. More definitive tests of

spatially explicit neutral models are thus dependent on future advances in data collection technology (*see* The Future of Neutral Theory).

Generalizing Speciation Mode

Speciation is a key process in neutral theory. The majority of neutral models, both spatially implicit and spatially explicit, have continued to use the original point speciation mode introduced by Hubbell. This approach was inspired by the process of point mutation in population genetics, but some alternative and more ecologically realistic speciation modes have since been introduced. The primary examples are random fission speciation and protracted speciation.

Random fission is intended to describe the biological process of allopatric speciation, under which geographic changes cause a population to become separated into two parts that then diverge genetically, ultimately becoming distinct species. In terms of mathematical implementation, random fission models allow for an extant species to split into two new species, with total abundance conserved through the process. Given that random fission speciation is implemented as an instantaneous process in neutral models, it still must be considered as an effective speciation mode, rather than an accurate description of allopatric speciation. Neutral models incorporating random fission speciation (Etienne and Haegeman, 2010) exhibit SADs that are qualitatively different from those of standard neutral models, potentially providing a way to distinguish between these different speciation modes.

Protracted speciation (Rosindell *et al.*, 2010) distinguishes incipient species arising from point speciation from full species, defined to be those exceeding a fixed threshold abundance. The average lifetime of an incipient species in a neutral model is shorter than that of a fully mature species, and this leads to modified predictions of SADs and the distribution of species lifetimes. These results potentially address the discrepancies between neutral theory and phylogenetic data discussed by Rosindell, though a more complete analysis has yet to be carried out.

So far, alternative speciation modes have been incorporated only in spatially implicit models, and the parallel development in spatially explicit models is likely to present further challenges, particularly for the explicit spatial process of allopatric speciation. A deep issue is the general lack of a rigorous biological definition of speciation, which stymies attempts to estimate species lifetimes from the fossil record in a consistent and meaningful manner or to devise a speciation model that can be considered objectively realistic.

Nonneutral Extensions

As discussed, neutral theory apparently provides a good description of biogeographical processes at some spatial and temporal scales, but its predictions perform less well at others. Furthermore, there are some questions that neutral theory is not designed to directly address; for example, those relating to the impact of climate change or increased nitrogen deposition on ecological communities. Accordingly, the development of nonneutral models building on the foundation of neutral theory has been motivated by a desire to expand both the scope and accuracy of these models.

As discussed in the Introduction, neutral models include some combination of drift, dispersal and speciation, but ignore a fourth fundamental ecological process, selection, which is essentially synonymous with niche processes in an ecological community (Vellend, 2010; Tilman, 2004). When neutral theory fails, then, a possible next step is to include some kind of niche structure, where species are selected for in different ways. There are many approaches to niche models in the literature, but, in contrast to neutral models, these are often complex, highly parameterized, difficult to solve analytically and difficult to test empirically: the general picture is a set of interacting species with differing resource requirements living on a heterogeneous landscape. A key challenge is therefore to develop models that combine the greater biological realism of niche models with the parsimony, analytical tractability and quantitative testability of neutral models: ecologists seeking to bridge niche and neutral theory must decide which niche processes to include, how to model these processes and how to test the models' predictions. This section reviews four distinct approaches to building niche structure into a neutral framework: directly incorporating negative density dependence into species demographic rates; adding species interactions in the form of nonlinearities in the basic stochastic equations describing community dynamics; modeling explicit habitat differences and environmental heterogeneity; and including species or individual-level variation driven by traits such as body size. The common feature in each of these models is that they all partially break the neutral symmetry introduced in the Introduction, but retain some symmetry in order to make the models more tractable.

The first of these four types of extension was developed by Volkov *et al.* (2005). In this approach intraspecific density dependence is introduced by allowing per-capita birth and mortality rates to be general functions of abundance. The model retains a species-level symmetry because these general functions are the same for all species; it retains a strong degree of symmetry because each species' dynamics is governed by the same rules. But it breaks neutrality according to the definition given in the Introduction, because intraspecific density dependence is stronger than interspecific density dependence, giving rare species an advantage. The central result is that this drift-selection model and the spatially implicit neutral model can both fit tropical forest SADs with similar accuracy. This again provides evidence that a single snapshot of an SAD does not provide sufficient information to distinguish different underlying processes.

More recently, Haegeman and Loreau (2010) have modeled a stochastic community with general quadratic interactions; in effect, a stochastic, multispecies Lotka–Volterra system. To restrict the set of possible interactions, symmetry is again retained at the species level: the interactions are structured so that swapping species labels makes no difference to the dynamics. When intra-specific and interspecific interactions are equal, the results of this model are very similar to those of the spatially implicit neutral model: density dependence for each species reduces to a dependence on overall community size, which in effect acts as a constant per-capita mortality rate when fluctuations in total community size are small. For the

nonneutral case in which intraspecific interactions are stronger than interspecific interactions, the model makes predictions that deviate from those of neutral theory. Although the predictions of this particular model have yet to be tested against empirical data, the theoretical results do demonstrate the general point that interspecific interactions can be important even in highly symmetric models.

A third way to build niche processes on top of a neutral framework is to differentiate groups of species according to their niche or function (Chisholm and Pacala, 2010). In the limit of strong selection, each species is assigned to one of several niches, and the species can survive only in this single niche. The neutral feature of this model is that within a given niche, species are assumed to be neutral relative to one another, so that the within-niche dynamics are then dominated by random drift. In the limit of high diversity, community patterns in this model are identical to those of neutral theory. For finite diversity, signatures of niche processes arise from the fact that maximum species abundances are constrained by niche size rather than overall community size. This yet again reinforces the notion that single snapshots of SADs are not sensitive tests of underlying processes. Perhaps more importantly, it provides an insight as to why neutral theory is a good approximation to community dynamics in high-diversity ecosystems: species can drift relative to one another even in strongly niche-structured systems, as long as the number of species per niche is sufficiently high.

Finally, O'Dwyer *et al*. (2009) have introduced fitness variation into a neutral framework by allowing individual birth, death, and ontogenetic growth rates to vary by body size. Symmetry is thus broken at the individual level but retained at the species level, because each species has the same scaling of demographic rates with body size. Empirical data from tropical forests do indeed indicate that demographic rates can vary by body size, so in that sense this model is more realistic than a neutral model. The model leads to some predictions that differ from or extend those of neutral models, but these predictions have not yet been tested against empirical data.

In the four examples above, a common theme is the breaking of neutrality but the retention of strong symmetry. The major advantage of symmetric models is that they are often analytically tractable because they have fewer parameters. A disadvantage of symmetric models is that, heuristically, they are similar to neutral models in many respects and hence lead to predictions that are often similar to neutral predictions. The choice of which symmetries to break and which to retain has so far been made in a somewhat ad hoc fashion, meaning that approaches to breaking neutral symmetry have lacked a deeper, guiding principle.

The Future of Neutral Theory

At this stage in the development of neutral theory, Hubbell's classic spatially implicit neutral model has been extensively investigated and tested, and many other theoretical developments are beginning to come to fruition. As stated in the Introduction, the two goals of neutral theory are to improve scientific understanding of species assemblages and to predict their behavior. Looking to the future then, it will be necessary to test the predictions of neutral models more extensively in order to identify the conditions under which they do and do not provide adequate descriptions of real communities. This, in turn, will continue to motivate the exploration of more complex models. This section draws together some of the issues identified in the previous sections, and outlinines some of the most promising prospects in current neutral theory research.

Testing the Theory

As mentioned earlier, neutral theory has been tested against a relatively limited range of empirical data. Although the good fits to SADs are in many cases impressive, their limitations as tests of neutral theory have been identified by a number of authors. Three directions of research promise to extend these empirical tests to provide more rigorous tests of neutral theory.

Firstly, future tests of neutral theory should encompass a broader range of spatial and temporal biodiversity patterns. Although some studies have looked at SARs, spatial correlation functions and fluctuations in species abundances over time, these studies are in the minority. One current limitation is that relatively few analytical predictions of these spatial and temporal patterns are available. Another is that the spatial and temporal scales of empirical data are typically limited for practical reasons. Future advances in this direction depend on the continued exploration of spatially explicit and dynamical neutral models, and the development of large-scale data collection techniques.

Secondly, future tests of neutral theory will be more powerful if they are based on multiple data sets and multiple types of data rather than single snapshots of particular distributions. Comprehensive tests of any stochastic model depend on an accurate empirical characterization of the full distribution of test statistics. When it is said that SADs provide insensitive tests of neutral theory, perhaps what is really meant is that single snapshots of SADs provide insensitive tests. Mathematical machinery for performing tests on multiple SADs is likely to be more widely employed as data availability increases: the development of remote sensing and other technologies will potentially provide the extensive data necessary to test and distinguish the predictions of neutral and nonneutral theories in a more satisfactory way.

Lastly, future tests of neutral theory should include predictions of data that have yet to be observed. Such tests provide much more convincing model validation than the mostly post-hoc fitting exercises that have been performed to date. One example of an analysis that extends neutral theory in the directions outlined above is the spatial scaling of SADs. If it is possible to predict SADs at multiple spatial scales from a single neutral model, it will become possible to use an empirical SAD at one spatial scale to make a neutral prediction for the SAD at larger spatial scales.

Developing the Theory

Future theoretical developments will generate new predictions that can be used to test neutral theory more rigorously.

Especially fruitful may be the nonneutral models (*see* Nonneutral Extensions) that build on neutral theory's mathematical framework. In creating such models, a particular challenge is to decide which biological processes should be included and how they should be represented mathematically. In the spirit of neutral theory, it would seem prudent to follow the general principles of parsimony and analytical tractability. The goal is not to include all available biological information but rather to distil the essence of a complex system into a few key processes and parameters that, when chosen correctly, characterize the model's behavior over broad spatial and temporal domains.

There is also a need to attain a better understanding of why neutral theory often works so well, despite its simplicity and lack of biological detail. Analogies have been drawn between neutral theory in ecology and ideal gas theory in physics: both make assumptions that are demonstrably false but lead nevertheless to large-scale effective theories that have remarkable explanatory power. But, such analogies highlight the lack of knowledge about how, and under what circumstances, neutral theory actually could emerge from more complex and realistic niche-based theories. Indeed, it is unclear what such niche-based theories should even look like. It is possible that a deep understanding of neutral theory can be derived from top-down observations, such as the ability of many species to persist on relatively few limiting resources or the inevitable low frequency of interaction of any two species in a high-diversity system.

Applying the Theory

To date, neutral theory has rarely been used to solve problems in conservation or other applied contexts. However, the field is relatively young. In terms of direct applications, neutral theory is most applicable as a first approximation to problems that involve large numbers of species with similar ecological function and in which the driver of change is habitat destruction or fragmentation. Neutral approaches will inevitably struggle to accurately predict the dynamics of an individual species, or to predict ecological communities' responses to climate change, nitrogen deposition and other processes that almost certainly require a niche perspective. But viewed as basic science, neutral theory can inform ecologists' understanding of fundamental ecological processes, thereby contributing indirectly to the solution of applied problems.

Summary

The development of neutral theory was driven by a desire to understand the maintenance of diversity in species-rich communities, and also by the need to make quantitative predictions for patterns of biodiversity. It has been extraordinarily successful in meeting these two objectives: the emphasis on transient coexistence has highlighted an alternative perspective on species diversity, previously overshadowed by the classic niche picture of indefinite coexistence. At the same time, the equalizing mechanisms that are necessary to prolong this transient coexistence lead to a set of highly symmetric models. This symmetry has been fundamental to the success of neutral theory, as it has helped to furnish a wide range of analytical predictions. In the case of the SAD and SAR, these predictions are surprisingly consistent with empirical patterns.

Despite these successes, there remains an apparent inconsistency between neutral theory and the vast array of species differences observed by ecologists in the field. To reconcile this, it must be emphasized that neutral theory merely hypothesizes that species differences do not impact large-scale patterns – it does not state that the differences do not exist. Nevertheless, this issue highlights the need to understand precisely which features of a real, nonneutral community lead to the emergence of neutral patterns. The universality of certain large-scale patterns across different communities provides some evidence that not all the details of a nonneutral community are equally important, but the complete picture is lacking.

The limitations of neutral theory must also be acknowledged. Neutral theory is designed to answer questions that represent only a subset of the questions of interest to ecologists. Statistical patterns of species richness and abundance (e.g., SADs and SARs) have been the subject of extensive empirical and theoretical investigation for decades, but ecologists are also interested in questions that depend intrinsically on species differences and therefore cannot be answered by neutral theory. For example, neutral models cannot predict when a certain function will be lost from an ecosystem, or under what circumstances a food web will collapse. Future progress in this direction may be achieved by wedding neutral theory's quantitative, predictive approach with insights from other branches of ecological theory.

Neutral ecology remains a work in progress, and the full range of predictions of neutral models has yet to be explored. New mathematical and computational techniques are likely to accelerate progress in the understanding of both neutral models and their nonneutral extensions in the coming years. Above all, the success of this research program will rest on its ability not only to predict patterns consistent with those already observed in nature, but also to make predictions that are unexpected or unintuitive, and yet turn out to be correct.

Acknowledgments

The BCI forest dynamics research project was made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Center for Tropical Forest Science, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Celera Foundation, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past two decades. The plot project is part of the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

Appendix

List of Courses

- Community Ecology (Undergraduate or Graduate)
- Theoretical Ecology (Undergraduate or Graduate)

See also: Biodiversity, Definition of. Biogeographical Models. Biogeography, Overview. Island Biogeography. Species–Area Relationships

References

Alonso D and McKane AJ (2004) Sampling Hubbell's neutral theory of biodiversity. *Ecology Letters* 7: 901–910.

Bell G (2001) Ecology – Neutral macroecology. *Science* 293: 2413–2418.

Caswell H (1976) Community structure – Neutral model analysis. *Ecological Monographs* 46: 327–354.

Chisholm RA and Lichstein JW (2009) Linking dispersal, immigration and scale in the neutral theory of biodiversity. *Ecology Letters* 12: 1385–1393.

Chisholm RA and Pacala SW (2010) Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proceedings of the National Academy of Sciences of the United States of America* 107: 15821–15825.

Condit R, Pitman N, Leigh EG, *et al.* (2002) Beta-diversity in tropical forest trees. *Science* 295: 666–669.

Etienne RS (2007) A neutral sampling formula for multiple samples and an 'exact' test of neutrality. *Ecology Letters* 10: 608–618.

Etienne RS and Haegeman B (2010) The neutral theory of biodiversity with random fission speciation. *Theoretical Ecology* 4: 87–109.

Haegeman B and Loreau M (2010) A mathematical synthesis of niche and neutral theories in community ecology. *Journal of Theoretical Biology* 260: 150–165.

Hubbell SP (1979) Tree dispersion, abundance, and diversity in a tropical dry forest. *Science* 203: 1299–1309.

Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography.* Princeton, NJ: Princeton University Press.

Jabot F and Chave J (2009) Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. *Ecology Letters* 12: 239–248.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature* 217: 624–626.

Leigh EG (2007) Neutral Theory: a historical perspective. *Journal of Evolutionary Biology* 20: 2075–2091.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Princeton, NJ: Princeton University Press.

O'Dwyer JP and Green JL (2010) Field theory for biogeography: A spatially explicit model for predicting patterns of biodiversity. *Ecology Letters* 13: 87–95.

O'Dwyer JP, Lake JK, Ostling A, Savage VM, and Green JL (2009) An integrative framework for stochastic, size-structured community assembly. *Proceedings of the National Academy of Sciences United States of America* 106: 6170–6175.

Rosindell J and Cornell SJ (2007) Species–area relationships from a spatially explicit neutral model in an infinite landscape. *Ecology Letters* 10: 586–595.

Rosindell J, Cornell SJ, Hubbell SP, and Etienne RS (2010) Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology Letters* 13: 716–727.

Rosindell J, Leigh EG, and Etienne RS (2011) The unified neutral theory of biodiversity and biogeography at age ten. *Trends in Ecology and Evolution* 26: 340–348.

Tilman D (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences United States of America* 101: 10854–10861.

Vellend M (2010) Conceptual synthesis in community ecology. *Quarterly Review of Biology* 85: 183–206.

Volkov I, Banavar JR, He F, Hubbell SP, and Maritan A (2005) Density dependence explains tree species abundance and diversity in tropical forests. *Nature* 438: 658–661.

Volkov I, Banavar JR, Hubbell SP, and Maritan A (2003) Neutral theory and relative species abundance in ecology. *Nature* 424: 1035–1037.

Nitrogen Deposition and Terrestrial Biodiversity

Christopher M Clark, US Environmental Protection Agency, Washington, DC, USA
Yongfei Bai, Chinese Academy of Sciences, Beijing, China
William D Bowman, INSTAAR University of Colorado, Boulder, CO, USA
Jane M Cowles, University of Minnesota, Saint Paul, MN, USA
Mark E Fenn, Pacific Southwest Research Station, Riverside, CA, USA
Frank S Gilliam, Marshall University, Huntington, WV, USA
Gareth K Phoenix, University of Sheffield, Sheffield, UK
Ilyas Siddique, Universidade Federal de Santa Catarina, Brazil
Carly J Stevens, The Open University, Milton Keynes, UK, and Lancaster University, Lancaster, UK
Harald U Sverdrup, Lund University, Lund, Sweden
Heather L Throop, New Mexico State University, Las Cruces, NM, USA

Published by Elsevier Inc.

The views expressed in this article are those of the authors and they do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency or other federal agencies.

Glossary

Acidification The process by which soil pH is reduced, potentially causing release of toxic minerals into the soil, base cation depletion, losses of plant biodiversity, and dominance by acid-tolerant species.
Calcareous A soil property of high calcium carbonate ($CaCO_3$) which buffers the soil against changes in pH.
Critical Load A quantitative estimate of an exposure to a pollutant below which significant harmful effects on specified elements of the environment do not occur according to present knowledge.
Denitrification A biogeochemical process mediated by soil microbes in which nitrate (NO_3^-) is converted to dinitrogen gas (N_2) along with other intermediary molecules.
Eutrophication The process by which enrichment with nutrients such as nitrogen stimulates plant growth, often leading to losses of plant biodiversity and dominance by weedy species.
Limiting resource The resource that most limits primary production in ecosystems (often nitrogen).
Nitrification A biogeochemical process mediated by soil microbes in which ammonium (NH_4^+) is converted to nitrate (NO_3^-) along with the release of protons into the soil.
Nitrogen deposition The process by which reactive forms of nitrogen are deposited to the earth's surface through either wet or dry deposition.
Reactive nitrogen All forms of nitrogen except atmospheric dinitrogen gas, includes all radiatively, photochemically, and biologically active inorganic forms (e.g., NH_3, NH_4^+, NO_x, HNO_3, N_2O) and organic molecules (e.g., proteins, urea, etc.).

Overview of the Issue

Nitrogen deposition, along with habitat losses and climate change, has been identified as a primary threat to biodiversity worldwide (Butchart *et al.*, 2010; MEA, 2005; Sala *et al.*, 2000). The source of this stressor to natural systems is generally twofold: burning of fossil fuels and the use of fertilizers in modern intensive agriculture. Each of these human enterprises leads to the release of large amounts of biologically reactive nitrogen (henceforth contracted to "nitrogen") to the atmosphere, which is later deposited to ecosystems. Because nitrogen is a critical element to all living things (as a primary building block of proteins among other biological molecules), nitrogen availability often limits primary production and is tightly recycled in many natural ecosystems. This is especially true in temperate ecosystems, though it may also be true for some areas in the tropics that are not phosphorus-limited (Adams *et al.*, 2004; Matson *et al.*, 1999). Thus, the large increase in availability of this critical nutrient as a result of human activity has profound impacts on ecosystems and on biodiversity.

Once nitrogen is deposited on terrestrial ecosystems, a cascade of effects can occur that often leads to overall declines in biodiversity (Bobbink *et al.*, 2010; Galloway *et al.*, 2003). For plants, nitrogen deposition can impact biodiversity generally through four processes: (1) stimulation of growth often of weedy species that outcompete local neighbors (termed "eutrophication"), (2) acidification of the soil and consequent imbalances in other key nutrients that favors acid tolerant species (termed "acidification"), (3) enhancement of secondary stressors such as from fire, drought, frost, or pests triggered by increased nitrogen availability (termed "secondary stressors"), and (4) direct damage to leaves (termed "direct toxicity") (Bobbink, 1998; Bobbink *et al.*, 2010). For animals, much less is known, but reductions in plant biodiversity can lead to reductions in diversity of invertebrate and other animal species, loss of habitat heterogeneity and specialist habitats, increased pest populations and activity, and changes in soil microbial communities (McKinney and Lockwood, 1999; Throop and Lerdau, 2004; Treseder, 2004).

In total, it is estimated that nearly 16.3 million km^2 or 11% of the terrestrial land surface is currently exposed to high levels of nitrogen deposition that could threaten biodiversity (Dentener *et al.*, 2006). Regions vary globally in the amount of area exposed, ranging from eastern Europe (80%), to South Asia (60%), east Asia (40%), southeast Asia (30%), western Europe (30%), and the US (20%), with the remaining global regions being generally low (<10%). Thus, nitrogen deposition potentially impacts biodiversity over much of the globe, from tropics to tundra. Although the biodiversity of most systems responds negatively to nitrogen deposition, the magnitude and exact nature of the effect can vary widely depending on interactions between nutrient availability and other factors such as climate, disturbance, and plant community composition (Bobbink *et al.*, 2010; Dise, 2011).

It is unclear to what degree recovery of biodiversity is possible from long-term nitrogen deposition. Recovery is anticipated to be enhanced through reduction of soil nitrogen availability, restoration of soil pH and other nutrient conditions, and addition of species formerly lost (Bakker and Berendse, 1999). In practice, it is unresolved to what degree this process occurs naturally if nitrogen deposition were reduced through regional, national, or international regulation. However, over time periods of interest to decision makers (years to decades) active management may be necessary to restore biodiversity within affected areas (Dise, 2011).

Many of the world's "hot spots" of biodiversity are either currently exposed, or are expected to be exposed in the near future, to potentially high levels of nitrogen deposition as industrialization continues to occur in developing nations of the Tropics and Asia (Phoenix *et al.*, 2006). Most of our knowledge on the impacts from nitrogen deposition comes from Europe and the US – areas that have already experienced modern industrialization and widespread losses of biodiversity – and, using short-term experiments in which high levels of N in excess of deposition were added. Thus, there is an urgent need for a greater understanding of the long-term impacts from low levels of nitrogen deposition in all systems, with particular emphasis on understudied biomes and geographic areas such as the Tropics, Asia, and Africa (Bobbink *et al.*, 2010).

Background on Nitrogen as a Nutrient and Pollutant in Ecosystems

Nitrogen as a Nutrient and Resource Limitation

Nearly 99% of the nitrogen (N) on the planet is in the atmosphere as highly stable dinitrogen gas (N_2), where two atoms of N are triple-bonded together. N is the most abundant element in our atmosphere, making up approximately 78% by volume, followed by oxygen (~20%) and Argon (~1%) (Galloway *et al.*, 2003). N is also one of the most critical elements for life, constituting the elemental basis for peptide bonds between amino acids, which combine to form proteins – the basic building blocks for the biochemical reactions underpinning all life.

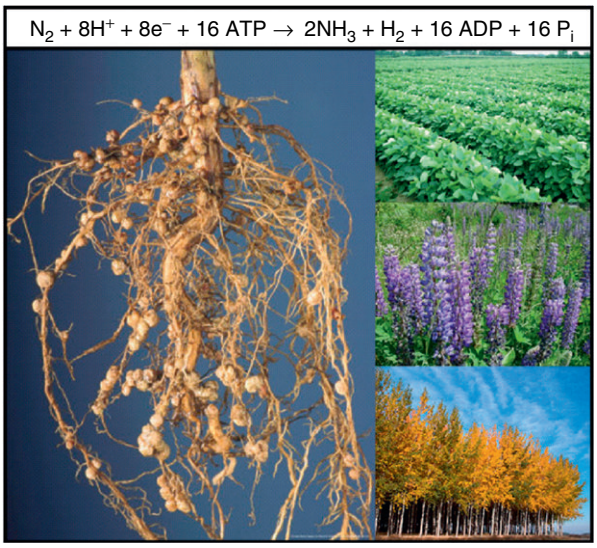


Figure 1 Illustration of biological nitrogen fixation (BNF). Clockwise from the top: Simplified chemical equation of BNF; some common N fixing plants, soybean, lupine, alder; and a closeup of plant roots and the root nodules where BNF takes place for leguminous species.

Ironically, however, atmospheric nitrogen is unavailable to approximately 99% of living organisms (Galloway *et al.*, 2003). The ability to break this triple bond, thereby converting N_2 into more reactive forms, is relatively rare in nature. It requires significant energy, low oxygen levels, specialized enzymes, and is almost entirely the purview of bacteria and archea (the group capable termed diazotrophs). These organisms generally live either freely in the soil or in water, or in close association with plant roots belowground, and are responsible for this biological nitrogen fixation (BNF, **Figure 1**). Lightning can also break the triple bond of N_2, though the importance of lightning to global N supplies is small by comparison to BNF. Thus, N_2 is commonly termed "nonreactive N" because it is inert to all but a small fraction of organisms. The remaining 1% of planetary N includes all biologically, photochemically, and radiatively active compounds in the atmosphere, hydrosphere, geosphere, and biosphere (termed "reactive N" or "Nr"). Nr includes dozens of different molecular forms, including inorganic oxidized N (e.g., nitric acid (HNO_3), nitrogen oxides (NO_x), nitrous oxide (N_2O), and nitrate (NO_3^-)), inorganic reduced N (e.g., ammonium (NH_4^+) and ammonia (NH_3)), and organic N (e.g., proteins, urea, or amines) (Galloway *et al.*, 2003).

The importance of nitrogen to all biological functioning, and its relatively restricted supplies, means that its availability often limits primary production in natural ecosystems (Vitousek *et al.*, 2002). Other nutrients can also limit production, especially phosphorus, though nitrogen limitation or colimitation is widespread in terrestrial ecosystems (Elser *et al.*, 2007; LeBauer and Treseder, 2008). Because the ultimate source of phosphorus is chemical weathering of mineral rocks, older systems that have not been glaciated for millennia can develop P limitation, a condition often observed in tropical regions (Matson *et al.*, 1999). Nonetheless, nitrogen limitation is also observed in many tropical areas.

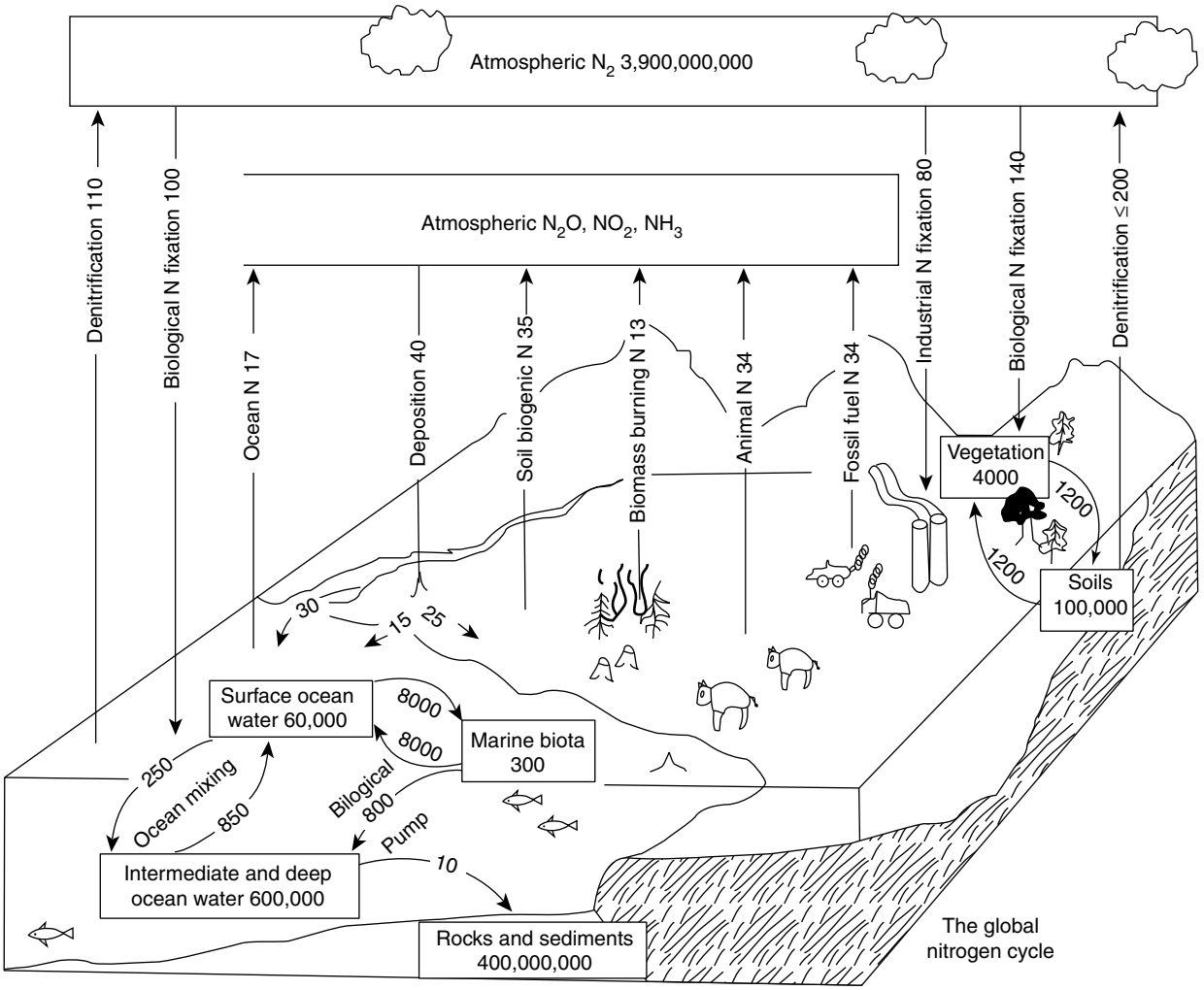


Figure 2 The global nitrogen cycle showing approximate magnitudes of major pools (boxes) and fluxes (arrows) in teragrams per year ($1Tg = 10^{12}$ g). The atmosphere contains the vast majority of Earth's nitrogen, followed by oceanic rocks and sediments and the soil. The amount of N that cycles in terrestrial and marine systems is much greater than N inputs from BNF (nine-fold and 80-fold respectively). Reproduced with permission from Chapin FS, Matson PA, and Mooney HA (2002) *Principles of Terrestrial Ecosystem Ecology*. New York: Springer-Verlag.

Nitrogen Cycling: Preindustrial and Postindustrial

Prior to the industrial revolution, creation of Nr came from three sources: BNF, lightning, and preindustrial agriculture. Together, these processes added Nr to terrestrial ecosystems at roughly 141 TgN/year ($Tg = 10^{12}$ g), which was predominantly natural BNF (BNF: 92%; preindustrial agriculture: 6%; lightning: 2%; Galloway *et al.*, 2004). There was some anthropogenic-driven BNF through the cultivation of rice paddies, but this was small by comparison. Most of these nitrogen inputs were transferred to the ocean, denitrified to the atmosphere, or accumulated slowly as organic material in soils. All these processes are part of the global nitrogen cycle (**Figure 2**).

Human activity began to substantially impact the global nitrogen cycle following the industrial revolution from the burning of fossil fuels for energy and the advent of modern agriculture. When fossil fuels are burned at high temperatures in the presence of atmospheric oxygen, many oxidized forms of Nr are produced (e.g., nitrogen oxides such as NO_x). These molecular forms of Nr disperse and react with the atmospheric constituents, eventually being deposited through either wet or dry nitrogen deposition.

The advent of modern agriculture also had a major impact on the global nitrogen cycle (Smil, 2001). In the late nineteenth and early twentieth centuries, nitrogen for fertilizer and explosives came mainly from nitrate salts mined in the Atacama desert of Chile and from animal droppings on islands off the coasts of Chile and Peru. There was a global shortage developing for these resources, provoking a search for substitutes. Soon thereafter, German scientists Fritz Haber and Carl Bosch invented a process for synthesizing NH_3 from N_2 using high temperature, pressure, catalysts, and an abundant hydrogen source. This abundant and affordable Nr source was a major driver for increases in global food supply as a primary constituent of fertilizer. Some fertilizer N volatilizes into the atmosphere, leading to wet or dry deposition of reduced N to terrestrial ecosystems, though much of it leaches to aquatic ecosystems. Vitousek (1997) estimated that sometime in

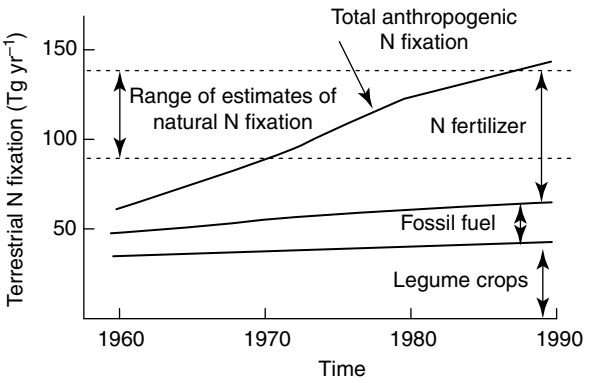


Figure 3 Increases through time in anthropogenic N fixation from planting of leguminous crops, fossil fuel combustion, and fertilizer, as compared with all natural processes combined. Reproduced with permission from Galloway JN, Schlesinger WH, Levy H III, Michaels A, and Schnoor JL (1995) Nitrogen fixation: anthropogenic enhancement-environmental response. *Global Biogeochemical Cycles* 9: 235–252.

the past few decades, human additions to the nitrogen cycle exceeded all natural processes combined (**Figure 3**). A budget of the nitrogen cycle in 1890 (**Figure 4(a)**) and 1990 (**Figure 4(b)**) highlights the magnitude of these effects.

Nitrogen deposition is expected to continue to increase as developing nations industrialize (**Figure 5**). Because most terrestrial ecosystems developed under conditions of nitrogen scarcity, enriching global ecosystems with Nr can have dramatic effects.

Biodiversity and Nitrogen Deposition

Here the authors define biodiversity or biological diversity simply as the diversity of life within a particular system, including genes, species, communities, and ecosystems. Most research on nitrogen deposition has focused on the number of species within a particular area, termed species richness. The authors focus largely on impacts to plants, also because this is where most research focuses, and note that impacts on other trophic levels often stem from the impacts on the plant community. Some impacts that occur on other trophic levels are elaborated below.

How Nitrogen Deposition Impacts Terrestrial Biodiversity

Nitrogen deposition, after habitat losses and climate change, is considered the major threat to biodiversity worldwide, with increasing stresses on some of Earth's most diverse areas (**Figure 6**, **Table 1**). For plants, N deposition affects terrestrial biodiversity through four primary mechanisms: (1) eutrophication, (2) acidification, (3) exacerbation of secondary stress, and (4) direct toxicity (**Figure 7**). These mechanisms will not operate – or have equal importance – in all ecosystems. The strength of each of these four mechanisms are influenced by other modifying factors discussed in Characteristics Describing Sensitivity to Nitrogen Deposition. Impacts on animals are less well studied, and are presented in the taxa-specific subsections of the Section Conditions that Alter the Magnitude of Impacts on Biodiversity (e.g., for soil biota, insects, mammals, etc.). Here the authors describe these mechanisms affecting plants and their properties generally.

Eutrophication

Eutrophication describes the process of N loading increasing availability of N in the soil to plants, which often leads to a cascade of effects. Because plant growth in many ecosystems is limited by the availability of nitrogen (Elser *et al.*, 2007), this is a direct process where concentrations increase in the soil and plant growth increases. Over time, a positive feedback can emerge, where increases in plant tissue N stimulates further increases in decomposition and liberation of more N. Overall, this increase in N availability often stimulates the growth of fast-growing species (termed "nitrophilous" species), resulting in competitive exclusion of less responsive species. Species that are rare, slow growing, and native are often lost more than other species, though this is not always the case (Suding *et al.*, 2005). Because most species are adapted to low nutrient availability (the prevailing condition), they are less responsive than weedy species, and are outcompeted through competition either for light aboveground and/or nutrients belowground. Eutrophication may result in expansion of aggressive species already in the plant community, or facilitate invasion by species not originally present. Eventually, ecosystems become saturated with N, and their productivity becomes limiting by other factors such as water or P. Even so, tissue concentrations of N may further increase, leading to potential nutrient imbalances, physiological stresses, and/or increased losses to herbivory (Dise, 2011).

Acidification

Acidification describes the process by which addition of N decreases soil pH, which can have a variety of direct and indirect effects on plant growth. Generally during acidification, changes in soil pH are mitigated by the release of carbonates and base cations from the soil (Bowman *et al.*, 2008). Once these are exhausted, clay minerals in the soil can breakdown leading to the release of toxic minerals into the soil (especially aluminum, Al^{3+}). N deposition can result in acidification through a number of mechanisms, including (1) stimulation of nitrification which yields protons (H^+), (2) roots uptake of NH_4^+ releasing H^+ as a counter ion, and (3) binding of NO_3^- with base cations and subsequent loss via leaching (reduces soil buffering capacity) (Dise, 2011; Ulrich, 1983). Over long time periods, acidification can suppress nitrification and plant uptake of nitrogen, leading to further accumulation of acidifying compounds such as NH_4^+ and a buildup of undecomposed material (Roelofs *et al.*, 1985). Acidification generally reduces biodiversity because there are fewer plant species adapted to more acidic soils, through suppression of germination, and through changes in the concentrations of either toxic minerals (e.g., Al^{3+}) or nutrients (e.g., N, P, base cations) in the soil (Horswill *et al.*, 2008; Stevens *et al.*, 2010a).

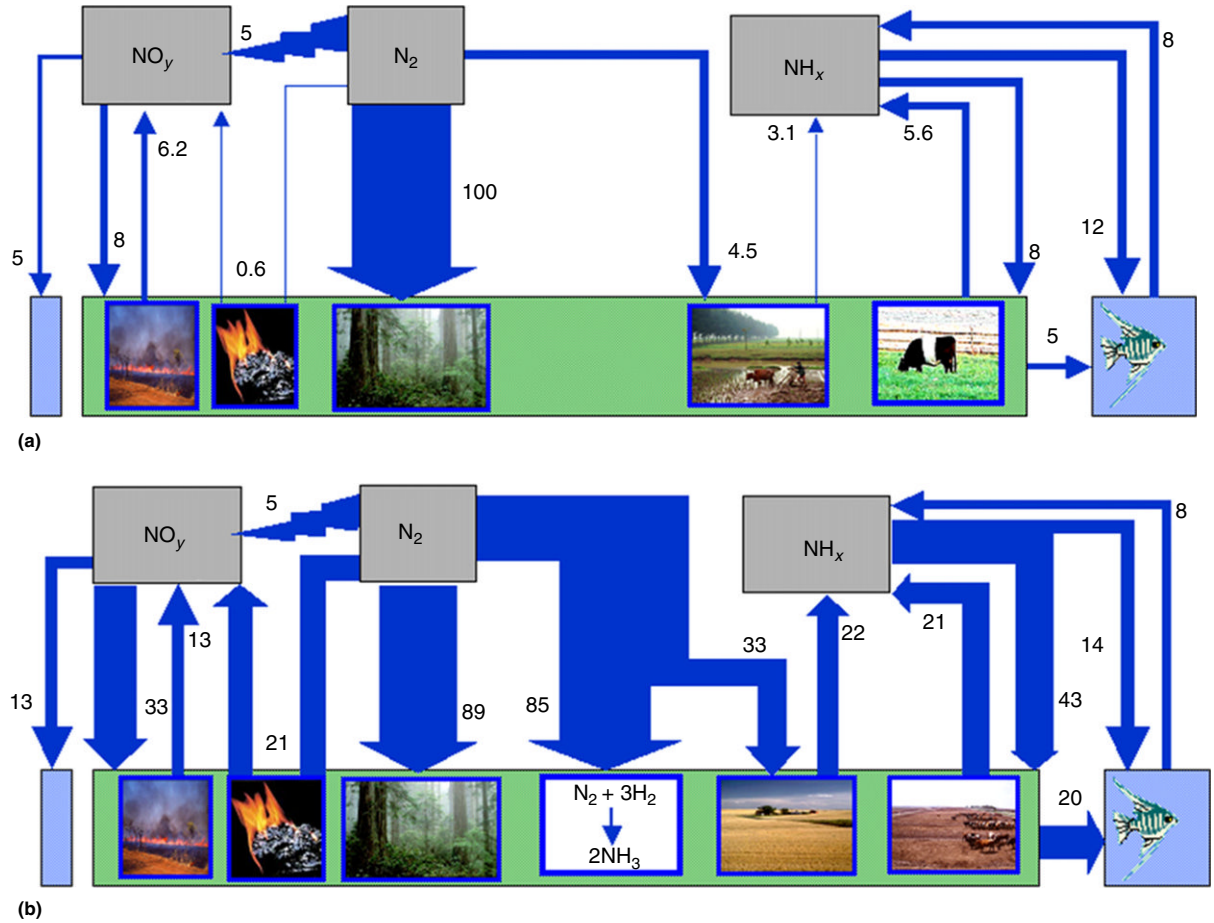


Figure 4 Global nitrogen budget for (a) 1890, and (b) 1990 Tg N per year. Down arrows represent inputs to terrestrial and aquatic systems, up arrows represent inputs to the atmosphere, lateral arrows represent transfers (two shown: Via rivers from terrestrial to coastal marine systems, or via lightning). Systems and subsystems are (from left to right): open ocean, terrestrial (land burning, fossil fuel combustion, natural ecosystems, fertilizer production (absent in 1890), crop agriculture, and animal husbandry), and coastal marine systems. Arrow sizes are relative to flux rates. The emissions to the 'NO_y box' from the 'coal' reflect fossil–fuel combustion. Those from the "vegetation" include agricultural and natural soil emissions, and combustion of biofuel, biomass (savannah and forests) and agricultural waste. The emissions to the "NH_x box" from the "agricultural field" include emissions from agricultural land and combustion of biofuel, biomass (savannah and forests) and agricultural waste. The NH_x emissions from the "cow" and "feedlot" reflect emissions from animal waste. Reproduced from Galloway JN and Cowling EB (2002) Reactive nitrogen and the world: 200 years of change. *Ambio* 31: 64–71, with permission from Royal Swedish Academy of Sciences.

Secondary Abiotic and Biotic Stressors

Secondary stresses may also be exacerbated by N deposition. For instance, increased winter injury and summer drought damage has been observed on *Calluna vulgaris* in heathland and bog ecosystems (Britton and Fisher, 2007; Sheppard *et al.*, 2008). The same species has shown increase infection from *Botrytis* and *Phytophthora* pathogens under enhanced N deposition (Sheppard *et al.*, 2008). The mechanisms causing winter, drought, and pathogen damage remain unclear, though greater stress sensitivity of more luxuriant growth, reduced biomass allocation to roots, lower mycorrhizal infection, shifts toward more parasitic associations belowground, and loss of essential nutrient ions such as Ca^{2+} have all been implicated (Bobbink *et al.*, 2010). N deposition has also been shown to lead to greater damage from invertebrate herbivores, which appears to be driven by greater foliar nutrient quality, reduced secondary defense compounds, and in some cases, greater invertebrate herbivore growth rates when feeding on nitrogen enriched foliage (Power *et al.*, 1998; Throop and Lerdau, 2004).

Direct Foliar Damage

Although direct foliar toxicity is not generally assumed to be a prominent driver of biodiversity changes, impacts can occur when atmospheric N compounds are found at high concentrations, usually close to emissions sources, and for especially sensitive taxa that lack protective tissues and structures such as moss and lichens (Bobbink *et al.*, 2010). For higher plants, outer tissues are relatively impervious (e.g., cuticle layers of leaves) to Nr (e.g., NH_3), with impacts occurring following direct entry through the stomata (Krupa, 2003). Following entry, NH_y can have a variety of effects on all plant types including inducing stomatal opening, nutrient imbalances, and disruption of cell membrane integrity, in addition to the

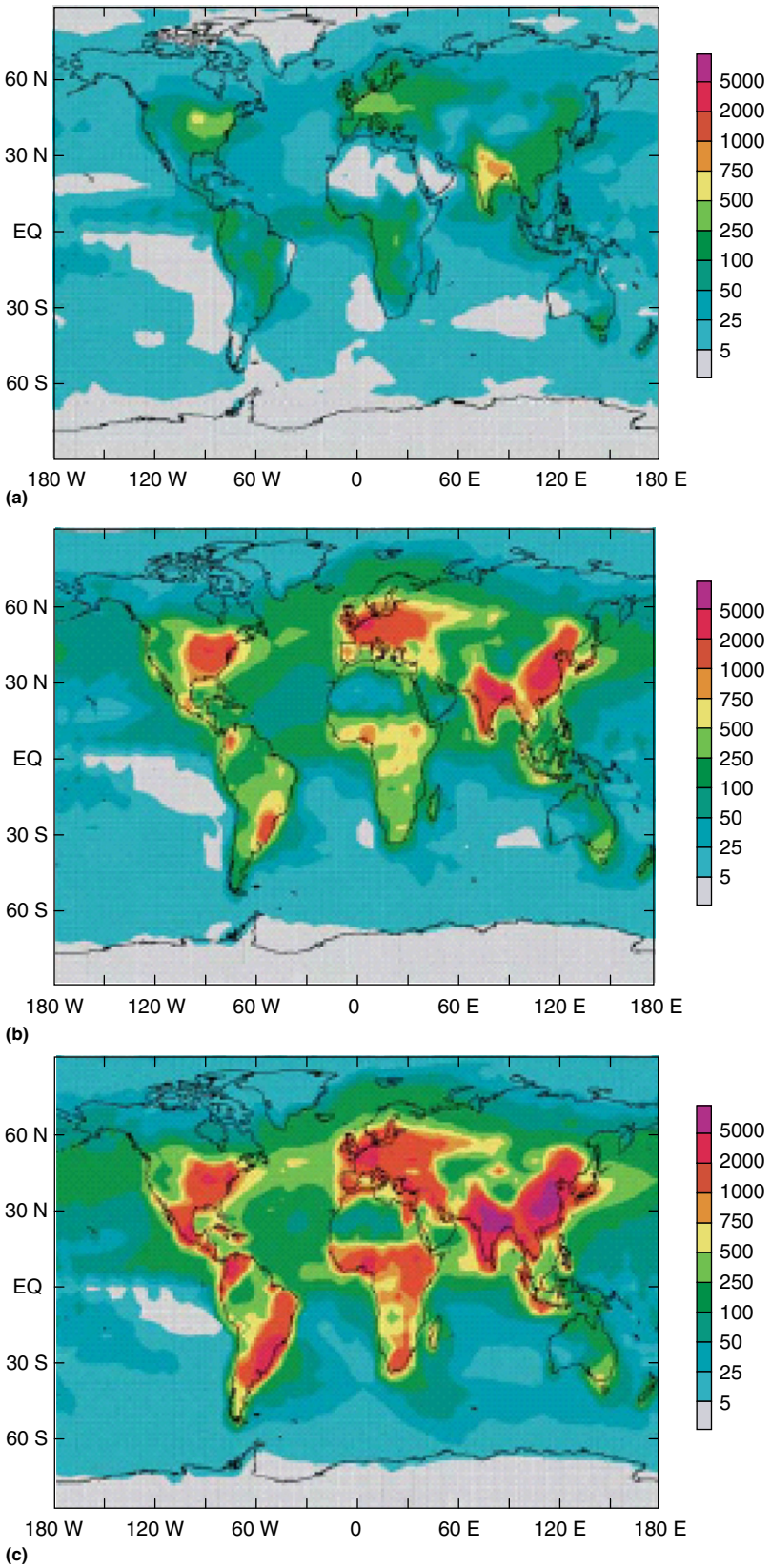


Figure 5 Distribution of total inorganic N deposition estimated in (a) 1860, (b) early 1990s, and (c) 2050 in mg N m^{-2} per $year^2$. Reproduced from Galloway JN, *et al.* (2004) Nitrogen cycles: Past, present, and future. *Biogeochemistry* 70: 153–226, with permission from Springer.

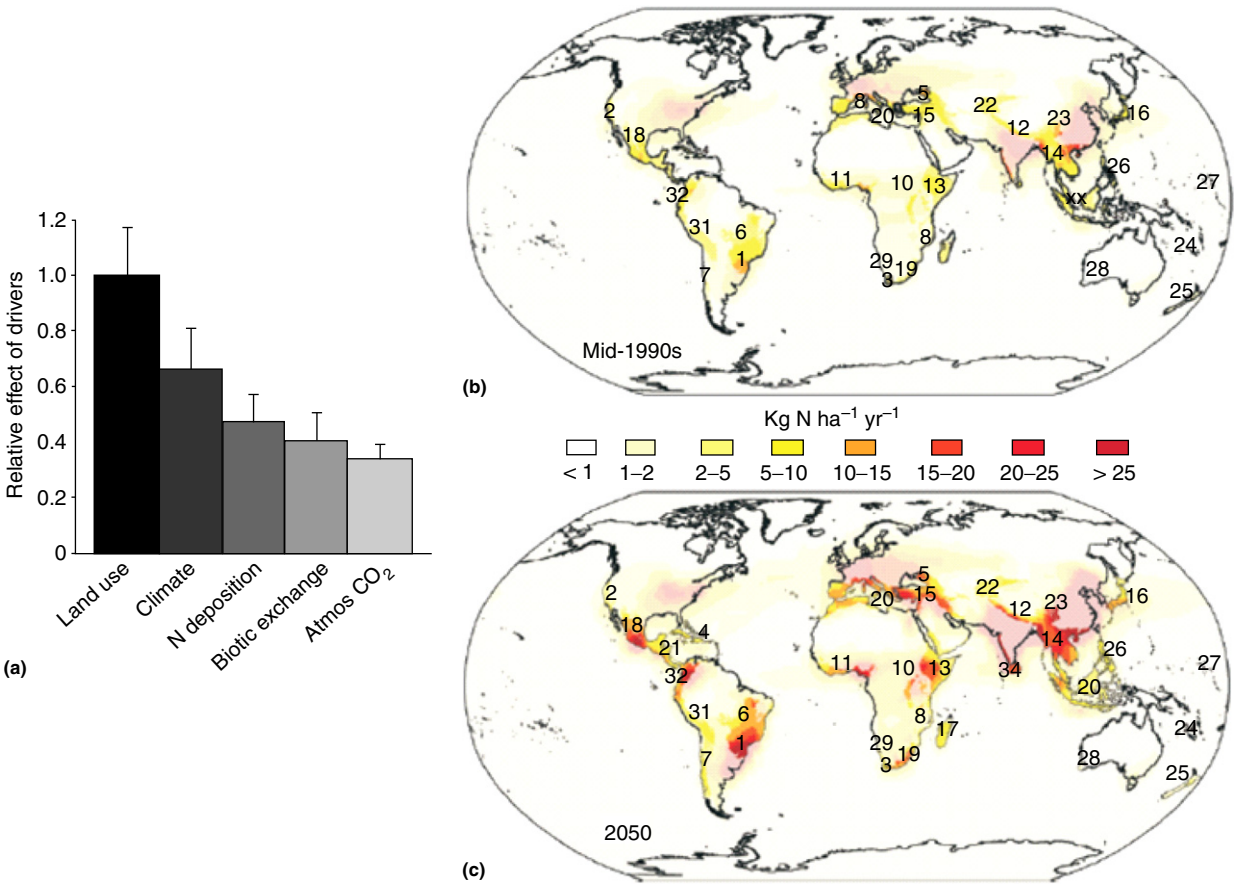


Figure 6 Relative effects estimated from major drivers of change on biodiversity in 2100. (a) Reproduced from Sala OE, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774, with permission from AAAS. Effects are scaled relative to land use. Also shown are estimated deposition levels in 1990 (b) and 2050 (c) for 34 biodiversity hotspots as defined by Conservation International (See **Table 2** for names, descriptions, and additional information. To aid in identification of deposition in biodiversity hotspots, coloring is masked (paler) for deposition outside hotspot boundaries. (b, c) Reproduced from Phoenix GK, *et al.* (2006) Atmospheric nitrogen deposition in world biodiversity hotspots: The need for a greater global perspective in assessing N deposition impacts. *Global Change Biology* 12: 470–476.

secondary stresses highlighted above following N assimilation into plant tissue (Krupa, 2003).

Research Approaches: How Do We Know What We Know?

Research efforts investigating ecological impacts associated with chronically elevated N deposition began decades earlier in Europe than in the rest of the world, in part because these problems arose earlier and were relatively more widespread. Most of the research on effects of excess N on terrestrial ecosystems has focused on biogeochemical responses, with far less research on effects on biodiversity. Studies examining biodiversity are heavily skewed toward plants. Even so, experimental studies examining the effects of N on species diversity of plant communities have a long history, with the first being the Park Grass Experiment in Rothamsted, England, established in 1858 (Silvertown *et al.*, 2006). More recent studies around the world find similar results of this earlier work – experimentally-added N profoundly alters species composition often decreasing species diversity of plant communities. In terms of which ecosystems and regions are studied, most research has been carried out on herb-dominated communities in temperate areas, and far fewer have been done in forested ecosystems and in the tropics (Bobbink *et al.*, 2010; Gilliam, 2007). These are significant discrepancies, considering that temperate forests often occupy areas receiving high rates of atmospheric deposition of N, and many tropical areas are expected to experience increased N deposition as development intensifies in the coming decades (Phoenix *et al.*, 2006).

Research approaches toward determining the response of plant biodiversity to increasing N deposition can be divided into three broad categories: observational studies, manipulative studies, and modeling. Observational studies are divided into two types, often termed "gradient" and "resampling" studies. Gradient studies examine biodiversity patterns across N deposition gradient(s) from areas of low to high N deposition. Resampling studies measure biodiversity at a particular location, comparing patterns when N deposition was low (usually in the past) with patterns when N deposition is/was high (usually current). Manipulative, or experimental, approaches involve controlled addition of N in various forms and amounts to plots or watersheds (much more rare) and measuring the

Table 1 Percent area of hotspots exposed to N deposition in excess of a commonly cited harmful level (10 kg N ha^{-1} per year, Bobbink *et al.* (2010) and other hotspot properties

	Mid 1990s	*2050*	*# Endemics*	*# Species*	*Predominant biome receiving excess N deposition in 2050*
Western Ghats and Sri Lanka (34)	66.6	100	3049	5916	Tropical and subtropical moist broadleaf forest
Indo-Burma (14) 3	34.8	97	7000	13,500	Tropical and subtropical moist broadleaf forest
Atlantic forest (1)	31.7	94.8	8000	20,000	Tropical and subtropical moist broadleaf forest
Mountains of Southwest China (23)	47.9	90.2	3500	12,000	Temperate coniferous forest
Tumbes-Choco'-Magdalena (32)	–	82.7	2750	11,000	Tropical and subtropical moist broadleaf forest
Maputaland-Pondoland-Albany (19)	–	78.4	1900	8100	Montane grasslands and shrublands
Irano-Anatolian (15)	–	78.3	2500	6000	Temperate broadleaf and mixed forest
Madrean Pine-Oak Woodlands (18)	3.8	78.3	3975	5300	Temperate coniferous forest
Guinean forests of West Africa (11)	19.8	75.5	1800	9000	Tropical and subtropical moist broadleaf forest
Mediterranean basin (20)	12.6	68.9	11,700	22,500	Mediterranean Forests, woodlands, and scrub
Cerrado (6)	0.3	68.7	8000	20,000	T & s-t grasslands, savannas, and shrublands
Eastern Afromontane (10)	–	68.1	2356	7598	Montane grasslands and shrublands
Himalaya (12)	5.9	59.4	3160	10,000	Temperate broadleaf and mixed forest
Caucasus (5)	18.9	49.7	1600	6400	Temperate broadleaf and mixed forest
Japan (16)	3.2	46.6	1950	5600	Temperate broadleaf and mixed forest
Mesoamerica (21)	1.5	46.1	2941	17,000	Tropical and subtropical dry broadleaf forest
Cape floristic region (3)	–	41.7	6210	9000	Mediterranean Forests, woodlands, and scrub
Horn of Africa (13)	–	37.1	2750	5000	T & s–t grasslands, savannas, and shrublands
Tropical Andes (31)	2.1	30.4	15,000	30,000	Tropical and subtropical moist broadleaf forest
Succulent Karoo (29)	–	18.5	2439	6356	Deserts and xeric shrublands
Sundaland	–	15.3	15,000	25,000	Tropical and subtropical moist broadleaf forest
Coastal forests of Eastern Africa (8)	–	11	1750	4000	Tropical and subtropical moist broadleaf forest
Caribbean islands (4)	–	1	6550	13,000	Tropical and subtropical dry broadleaf forest
Philippines (26)	–	0.1	6091	9253	Tropical and subtropical moist broadleaf forest
Chilean winter rainfall-Valdivian forests (7)	–	<0.1	1957	3892	Temperate broadleaf and mixed forest

Numbering in parentheses refers to the map in Figure 6b and c.

Source: Modified from Table 1 from Phoenix GK, *et al.* (2006) Atmospheric nitrogen deposition in world biodiversity hotspots: The need for a greater global perspective in assessing N deposition impacts. *Global Change Biology* 12: 470–476.

biodiversity response. Integrated soil-vegetation models show much promise in integrating our combined understanding of the process of N induced changes in biodiversity, though significant challenges exist. Most modeling approaches combine two modeling phases: (1) examination of the impacts of N deposition on soil solution N, water, and soil pH, and (2) the impacts of these changes on plant community structure (De Vries *et al.*, 2010). Models differ in many substantial ways, including the use of statistical relationships to derive results, the degree of resolution for ecological processes, the input parameters and variables computed (De Vries *et al.*, 2010).

Each of these approaches has strengths and weaknesses with respect to each other, which are often exclusive (**Table 2**). As reviewed by Gilliam (2006), modeling studies are much less prevalent than empirical work, with observational studies tending to be more common in Europe, and manipulative studies are generally more common in North America, South America, and Asia.

Conditions that Alter the Magnitude of Impacts on Biodiversity

Ecosystem-specific Effects of Nitrogen Deposition on Biodiversity

The effects of N deposition on ecosystems worldwide can depend on a variety of abiotic and biotic factors, including climatic factors, soil properties, preexisting nutrient limitations, productivity, and history of N deposition (see Bobbink *et al.*, 2010 for a comprehensive review). In many ecosystems including deserts, temperate grasslands, savannahs, shrublands, and Mediterranean systems, N enrichment can lead to increases in nonnative grasses, at the expense of native forb abundance and diversity. Many times this leads to an overall decrease in biodiversity (Bai *et al.*, 2010; Clark and Tilman, 2008, Allen *et al.*, 2009, Zavaleta *et al.*, 2003). Generally speaking, temperate ecosystems appear more sensitive to N-induced species declines than tropical systems because the former are more N-poor (as opposed to the more P-poor tropics) (Matson *et al.*, 1999). Arctic systems may be particularly sensitive for similar reasons, though low deposition rates in these areas and short growing seasons may limit responses in these ecosystems. Reponses in montane systems such as high alpine meadows may be similarly limited, although orographic lifting of air masses leads to disproportionately high levels of N deposition compared with lowlands below (Weathers *et al.*, 2006). The effects of N deposition on acid soils such as those found in Europe is predominantly due to soil acidification, whereas in well-buffered calcareous soils the major mechanism is eutrophication. Finally, wetter systems appear more sensitive than drier systems within a biogeographic region because the latter are more likely to be colimited by water and therefore less responsive to added N (Clark *et al.*, 2007).

These generalities above, however, belie complex responses that can occur for any of these ecosystems and regions.

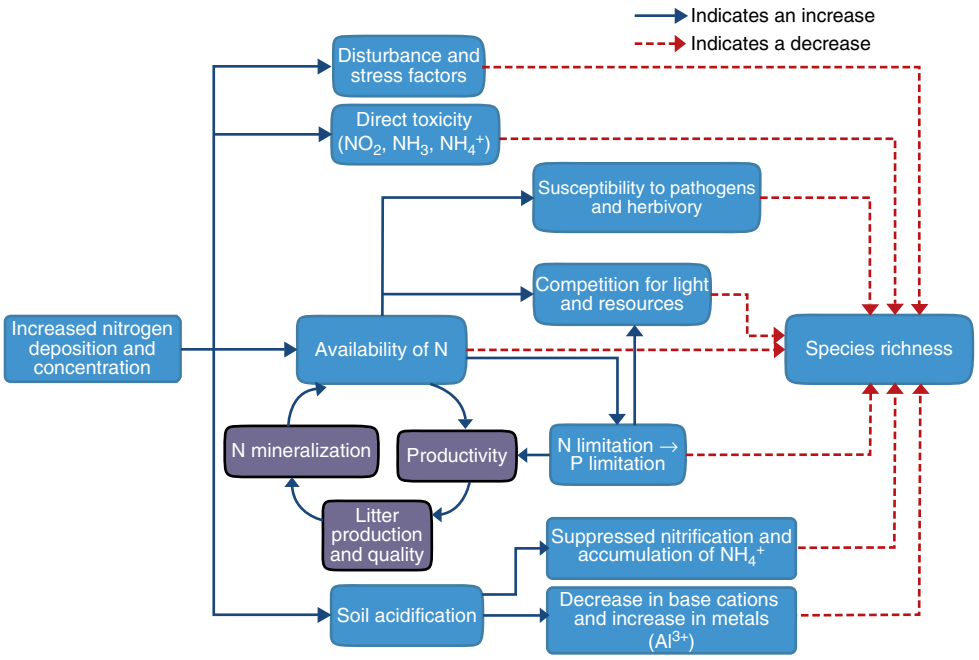




Figure 7 Schematic of the primary drivers of biodiversity decline from nitrogen deposition (top, Reproduced with permission from Dise N, Mike A, Salim B, *et al.* (2011) Nitrogen as a threat to European terrestrial biodiversity. In: Sutton MA (ed.) *The European Nitrogen Assessment.* Cambridge: Cambridge University Press. (Also shown are two examples of responses: From species rich nutrient-poor grasslands in Minnesota and from a high elevation spruce fir forest in Vermont. Photos on the left are control plots and photos on the right are plots receiving nitrogen fertilizer. Reproduced from Pardo LH, Robin-Abbott MJ, and Driscoll CT (2011) Assessment of nitrogen deposition effects and empirical critical loads of nitrogen for ecoregions of the United States. General Technical Report NRS-80, U.S. Department of Agriculture, Forest Service, Northern Research Station, with permission from CUP.

Table 2 A survey of the major approaches to studying the impacts of nitrogen deposition on biodiversity

Type of study	*Brief description*	*Strengths*	*Weaknesses*	*Examples (**Figure 8**)*
Observational-gradient	Measure biodiversity across a transect from high to low nitrogen deposition at one point in time	• Realistic nitrogen deposition profile (amount, form, timing, etc.) • Large scale represents dispersal limitations	• Other factors change along the gradient that may explain the biodiversity pattern (e.g., soil, land-use, climate, plant community) • More difficult to detect pattern because of low signal to noise ratio	Maskell *et al.* (2010), Stevens *et al.* (2004, 2010b)
Observational-resampling	Measure biodiversity at one location comparing when deposition was low (e.g., the past) with when deposition is high (e.g., current)	• Realistic nitrogen deposition profile (amount, form, timing).	• Other factors that change through time may explain the biodiversity pattern (e.g., land-use, climate) • More difficult to detect pattern because of low signal to noise ratio	Bennie *et al.* (2006), Dupre *et al.* (2010), Smart *et al.* (2005)
Manipulative	Add controlled amounts of N to plots or watersheds and measure biodiversity response	• Greater isolation of the effect of N, fewer confounding factors • Replication allows for greater statistical strength and higher signal:noise • If watersheds are the experimental unit, large scale realistically represents deposition	• Treatments often do not accurately represent deposition (one time addition of often large amounts of N in solid granular form) • Usually replicate plots are small (e.g., from 10 m $\times$ 10 m to 1 m $\times$ 1 m); or, large watersheds are unreplicated	Bowman *et al.* (2006), Clark and Tilman (2008) Gilliam (2006), Morecroft *et al.* (1994), Suding *et al.* (2005)
Modeling	Process and/or statistical models relating deposition to biodiversity	• Captures the full dynamics of how nitrogen impacts biodiversity through eutrophication and acidification pathways	• Based on current, often incomplete knowledge • Large data input requirements that are often lacking • Secondary factors and direct toxicity not currently modeled	Belyazid *et al.* (2011), De Vries *et al.* (2010), Sverdrup *et al.* (2007)

Source: Modified from Table 1 from Phoenix GK, *et al.* (2006) Atmospheric nitrogen deposition in world biodiversity hotspots: The need for a greater global perspective in assessing N deposition impacts. *Global Change Biology* 12: 470–476.

For example, although deserts as a whole may be less responsive than nondeserts, deserts in southern California have been invaded by fast-growing grasses are very responsive. This leads to a strong reduction in diversity and increases risk of fire that may additionally impact diversity (Rao *et al.*, 2010). In addition, some European heathlands may show little changes in diversity initially. But, following increases in plant N over years of N enrichment, pest outbreaks can lead to increased light levels at the soil and rapid changes in biodiversity including increased grass dominance (Strengbom *et al.*, 2002; Bobbink *et al.*, 2010). In summary, although there is a wide range of potential responses to N deposition, terrestrial biodiversity in most systems is negatively affected through one or more mechanisms.

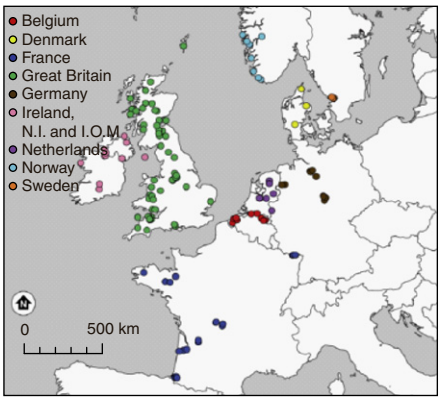
Taxa Specific Responses to Nitrogen Deposition

Nonvascular Plants

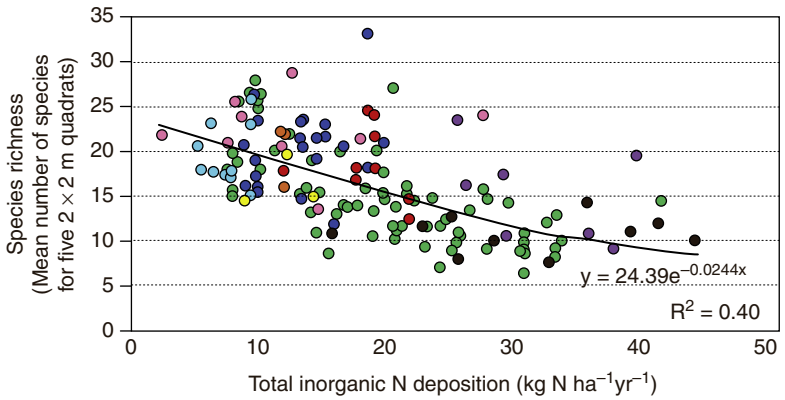
In general, nonvascular plants are the most sensitive to nitrogen enrichment, followed by herbaceous plants and shrubs, with trees being the least sensitive. The unique anatomy of nonvascular plants (e.g., lichens, bryophytes, liverworts, and mosses) makes them highly sensitive to fluctuations in atmospheric sources of N. Nonvascular plants lack root structures to access soil nutrient pools, and instead rely on nutrients directly absorbed from deposition, throughfall, and leachates from overstory vegetation. Their lack of a cuticle and vascular structures allows the passive, rapid absorption of water over their entire surface. Thus, they are particularly sensitive to deposited N. These inherent sensitivities causes nonvascular plants to respond to extremely low deposition levels, for instance, <3 kg N ha^{-1} per year for epiphytic lichens in the Sierra Nevada, CA (Fenn *et al.*, 2008). Degradation of these nonvascular species communities has far-ranging consequences that are often overlooked because of their supposed diminutive role in ecosystem function. For example, *Sphagnum* bogs are major carbon sinks in temperate areas, reindeer lichen are critical forage for these ungulates during the winter months, and many lichen species are used by birds for nesting material.

Herbaceous Plants

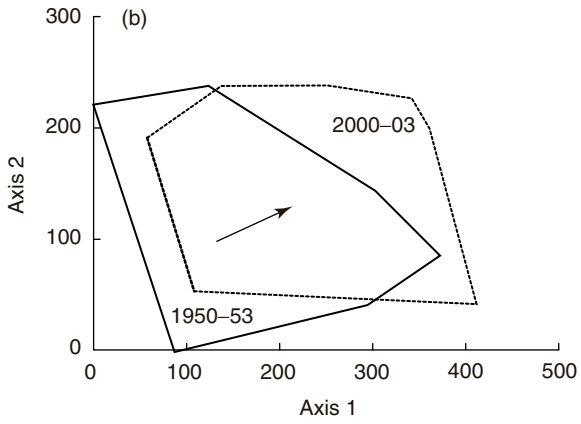
Because herbaceous plants and shrubs have vascular systems and protective epidermal layers, they access most of their nitrogen through the soil and are not as sensitive as nonvascular species to high concentrations of nitrogenous compounds in the air. Once deposited, however, large impacts can occur because of their shallow root systems, short life spans, and rapid growth rates compared with forest trees. Some plants



(a)



(b)



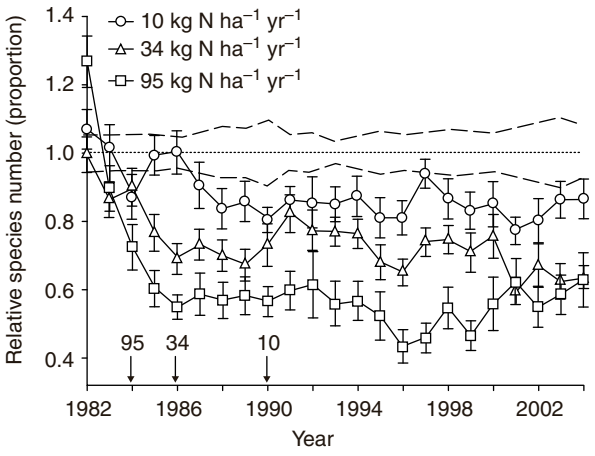
(c)



(d)



(e)



(f)

Figure 8 Example studies of N deposition effects on plant biodiversity from observational spatial gradients in Europe (a, b: Reproduced from Stevens CJ, *et al.* (2010b) Nitrogen deposition threatens species richness of grasslands across Europe. *Environmental Pollution* 158: 2940–2945.), temporal resampling of British chalk grasslands in the 1950s compared with 2000s (c, d: Reproduced from Bennie J, Hill MO, Baxter R, and Huntley B (2006) Influence of slope and aspect on long-term vegetation change in British chalk grasslands. *Journal of Ecology* 94: 355–368, with permission from Wiley.), and experimental manipulations (e, f: Reproduced with permission from Clark CM and Tilman D (2008) Loss of plant species after chronic low-level nitrogen deposition to prairie grasslands. *Nature* 451: 712–715.). Total species richness (b) declined with total inorganic N deposition. In (b) there was an exponential decline in species richness with N deposition. From the 1950s to 2000s (c), there was a shift in the community composition toward nitrophilic species as estimated using statistically distinct axes of composition. In (f) total species richness declined through time at different rates from annual additions of fertilizer N to experimental plots.

respond negatively to N deposition, declining in occurrence and/or abundance when N inputs are high, others show positive responses benefiting from the additional N through direct or indirect mechanisms. In many temperate grasslands, savannahs, and shrublands, grasses become more dominant whereas the cover and biodiversity of forbs and other species declines. Some species are particularly sensitive, such as slow growing long-lived species that historically dominated much of the US great plains, and/or acid-sensitive species from heathlands and acid grasslands of Europe. In forests, herbaceous species (e.g., the understory) have a disproportionate significance to the biome compared with their abundance (Gilliam, 2007). Indeed, although they only represent ~0.2% of aboveground biomass, herbaceous understory species make up 90% of plant biodiversity and produce >15% of the litter biomass. Thus, their losses may have large effects on many species that depend on them for forage and habitat.

Trees

Trees have long life spans, deep roots, and lower growth rates, characteristics that make them less susceptible to rapid changes in composition from N deposition. Nonetheless, shifts in relative growth rates among species within the community (stemming from rapid changes in N availability with N deposition) may portend to large biodiversity effects in the future (Bobbink *et al.*, 2010; Pardo *et al.*, in press; Thomas *et al.*, 2010). A seminal study by Thomas (2010) documented changes in forest tree growth rates with N deposition over much of the northeastern US, reporting that growth of many coniferous species declined whereas growth of tree species with arbuscular mycorrhizae increased, suggesting a strong potential for large-scale changes in forest composition. These patterns likely occurred because coniferous species are generally slower growing than broadleaf or deciduous species, and because trees with arbuscular mycorrhizae are better able to capitalize on deposited N (because these fungi can produce enzymes to break down N sources), whereas trees with other fungal associations such as ectomycorrhizal fungi cannot. Negative impacts from N enrichment in a tropical forest were found on a nutrient-poor acidic soil that was dominated by fast-growing pioneer species, all conditions leading to higher sensitivity (Siddique *et al.*, 2010).

Soil Microbes

Soil biota are capable of rapid changes in growth due to their small sizes and intimate association with soil particles and plant roots. They perform or aid in many critical ecosystem functions such as decomposition and nutrient uptake of N and P. Across various ecosystems, research has shown mixed effects on soil biotic communities, with some taxa increasing in abundance and diversity and others decreasing with N enrichment. However, for mycorrhizal fungi, many studies demonstrate a decline in reproductive output of species adapted to N-poor environments and their subsequent loss (Pardo *et al.*, in press). This is often associated with an increase in parasitic fungal species, declines in microbial diversity, and/or shifts in the soil community toward more bacterial dominance (Pardo *et al.*, in press). These patterns may be especially prevalent in systems with P-rich soils (Johnson *et al.*, 2003). Unfortunately, given the need for additional study, few generalizations can be made about the response of soil biota.

Other Organisms

Higher trophic levels (hereafter consumers, including herbivores, carnivores, pests, etc.) are primarily affected indirectly by nitrogen deposition via nitrogen-induced changes in food quality or quantity (Throop and Lerdau, 2004). N concentration of plant tissue often increases with nitrogen enrichment, which can strongly, and typically positively, influence the individual performance, feeding behavior, and population dynamics of herbivores. Individual-level responses of insect herbivores can drive population-level increases and increased herbivory may in turn suppress positive impacts of N on plant biomass (Bertness *et al.*, 2008). These changes can subsequently alter ecosystem-level patterns of carbon and nitrogen cycling (Throop *et al.*, 2004). However, reductions in plant diversity and habitat homogenization has been found to reduce diversity of insects (Haddad *et al.*, 2000), which may extend to other trophic levels. Many responses of insect herbivores to N deposition are species-specific, and differential responses to N deposition can affect community composition of herbivorous insects. Considerably, less is known about the responses to N deposition on noninsect consumers. However, as highlighted in the European heathlands, elevated N has been associated with increased fungal pests (Nordin *et al.*, 2005; Pardo *et al.*, in press). Little work to date has documented whether N deposition effects on insect herbivores extend to higher-level consumers in field situations, although changes in prey quality from N additions may affect predator feeding behavior (Throop and Lerdau, 2004).

Overview on Vulnerability

The vulnerability of biodiversity to nitrogen deposition has two components: exposure and sensitivity. Exposure describes the amount, duration, form, and mechanism of nitrogen deposition. The sensitivity of the system describes the intrinsic properties of the ecosystem that may preclude a larger or smaller impact for a given amount of the stressor. Generally, this is described by properties related to abiotic and biotic characteristics of the community.

Characteristics Describing Exposure to Nitrogen Deposition

The exposure characteristics of nitrogen deposition can generally be described by the amount (rate), duration, timing, chemical form, and mechanism of deposition. These characteristics, in turn, are affected by regional land use practices (e.g., agricultural vs urban), industrial activities, climate, and orographic effects among others. A large number of experimental N additions and surveys have found a "dose-dependent" response to N deposition (e.g., Stevens *et al.* 2004), with larger effects at higher rates of nitrogen addition. The more nitrogen is added, the greater the total effect. Timing of inputs also matters, where a greater effect might be expected if nitrogen is deposited during periods of active plant growth such as the spring and summer.

The chemical form of N deposition can also be an important determinant of impact. Differences have been observed in the impact of reduced and oxidized deposition (NH_x and NO_y) (summarized in Stevens *et al.* (2011)). Some plants have clear preferences for different N forms, and the form of N taken up by a plant may affect its health and performance. Also, the mechanism of deposition, whether deposited as wet deposition in rain, snow, or fog, on the leaf or soil surface, or as dry deposition onto leaf surfaces or the soil, may influence the impact of nitrogen deposition (Dise, 2011).

Characteristics Describing Sensitivity to Nitrogen Deposition

Abiotic Factors Affecting Sensitivity

Many abiotic factors influence the effect of a given amount of N on terrestrial biodiversity. The relative importance of each of these factors depends on which of the four dominant mechanisms (**Figure 7**) is driving changes in biodiversity. For systems in which eutrophication and competitive exclusion is the dominant mechanism, abiotic factors include the presence and strength of nitrogen limitation, the availability of open spaces for invasion by new species or expansion by existing species, and the availability and timing of other potentially limiting resources. For example, drier systems in the western US prairie tend to respond more weakly than wetter systems in the eastern plains (Clark *et al.*, 2007). Presumably, this occurs because the western plains are strongly colimited by nitrogen and water, whereas the eastern plains are primarily limited by nitrogen and therefore are better able to respond after N increases. Deserts have also been found to be less sensitive to N induced species declines in some but not all cases.

For systems in which acidification dominates, abiotic factors include soil pH, soil buffering capacity, weathering rates, as well as the availability and mobility of nutrient cations and toxic minerals in the soil. For example, systems with an already low pH and a low soil buffering capacity, might be more vulnerable to a given amount of nitrogen deposition than a more buffered soil all else being equal. This has been observed in grassland studies in Europe, where grasslands on poorly buffered acidic soils are more sensitive than grasslands on well-buffered calcareous soils (Maskell *et al.*, 2010). Abiotic factors may also affect the impact of direct toxicity, such as climate and base cation availability.

In systems where secondary stress dominates the ecosystem response to nitrogen deposition (e.g., through drought, frost, pathogens, herbivores, etc.), many of the same abiotic factors already mentioned above (e.g., climate and soil influencing the degree of N limitation affecting leaf palatability to herbivores and pathogens) operate to influence ecosystem sensitivity.

Biotic Factors Affecting Sensitivity

In addition to abiotic factors described above, several biotic factors affect sensitivity to N deposition (Dise, 2011; Gilliam, 2006). These biotic controls over the diversity response to N are in turn associated with the underlying mechanisms that alter changes in diversity.

Because growth and reproductive rates are linked with competitive ability, and not identical among competing species, the differential ability to increase growth with N deposition affects local extinctions from competitive exclusion (Suding *et al.*, 2005). This variation in responsiveness to N is associated with adaptation to soil nutrient conditions (Aerts and Chapin, 2000). Species characteristic of infertile soils generally have low growth rates and low tissue nutrient concentrations, which lowers their demand for nutrients, and lowers the rate of supply during decomposition of senesced material, thereby promoting their own persistence. Species characteristic of fertile soils follow opposite patterns, with high growth rates and nutrient tissue creating conditions favorable for their persistence. Patterns of biomass allocation and the ability to form new meristems in response to variation in N supply is another important determinant of a species growth response (Bowman and Bilbrough, 2001). Woody species and forbs generally have a lower capacity to form new meristems relative to grasses.

These physiological patterns are likely responsible for reported shifts among functional response types and traits under N enrichment. For instance, N enrichment tends to favor grasses, especially annual and tall or shade-tolerant grasses, nonlegumes (legumes fix atmospheric N), sedges , and broad-leaved trees (Fynn and O'Connor, 2005; Xia and Wan, 2008). Conversely, forbs, legumes, and perennials may be competitively suppressed by N enrichment (Xia and Wan, 2008). Thus, a first order approximation of the potential losses of biodiversity is whether there are few or many sensitive species present. However, a species' risk of loss is not solely determined by its individual traits, but also on how those traits compare with the species around it. For example, annuals are expected generally to respond strongly to nitrogen deposition. However, in communities dominated by many annual species (e.g., California Mediterranean grassland) there is a relatively weak response because species are generally similar to one another. Contrast this to the strong response in the perennial-dominated (*Leymus chinensis*) mature steppe of China, which experienced a strong increase of previously rare annual species and reductions in diversity (Bai *et al.*, 2010).

The type of microbial association also appears to influence a species' growth response. Trees in the eastern US with arbuscular mycorrhizal interactions have a greater capacity to increase growth in response to N deposition than ectomycorrhizal species (Thomas *et al.*, 2010). Although N enrichment often suppresses arbuscular mycorrhizae more strongly than ectomycorrhizae because of reduced C from plant hosts (Treseder, 2004), this may not have been the case in Thomas *et al.* (2010) possibly because of low soil P. The growth of species with symbiotic N-fixing bacteria are usually limited by P or micronutrients such as molybdenum, and are generally more likely to experience local extinction with increases in N availability than species that are N limited (Suding *et al.*, 2005).

Species also vary in the degree to which N deposition affects their susceptibility to environmental stress and soil acidification. Generally evergreen species exhibit greater deposition-induced susceptibility to stress than deciduous species (Evans *et al.*, 2001; Sheppard *et al.*, 2008). Species buffer the soils in the vicinity of their roots through ion exchange, and regulate tissue nutrient balances through similar ion

exchanges, each of which may differ among species (Stevens *et al.*, 2011).

Although globally herbaceous species show a greater biomass response to N enrichment than woody species (Xia and Wan, 2008), within an ecosystem N enrichment may shift dominance among plant life forms in the opposite direction. For instance, in some Arctic tundra communities enhanced coverage of woody species occurs in response to greater N availability, leading to declines in herbaceous species (Bret-Harte *et al.*, 2008). In tropical secondary forest, combined enrichment with nitrogen and phosphorus (N + P) neutralized the negative effects of N-only enrichment, associated with dramatic grass biomass responses to N + P (Siddique *et al.*, 2010). Thus, although there are some patterns that appear regularly in the literature, site-specific responses are dependent on many interacting processes that can yield a variety of patterns.

Finally, N deposition may influence diversity through interactions between plants and consumers (Throop and Lerdau, 2004). Increases in deposition have the potential to mitigate losses associated with insect herbivory through increased plant production (Throop, 2005), but may also amplify losses through increased feeding rates and pest populations associated with increased amount and nutrient content of foliage (Throop and Lerdau, 2004; Xia and Wan, 2008). Differential responses in phenology may amplify competitive interactions in some systems. In a Mediterranean California grassland, N addition delayed the early activity and flowering of grasses and brought on earlier flowering for forbs (Cleland *et al.*, 2006), enhancing competition for these two functional groups. Thus, there can be a greater potential for interaction through pests, pollinators, and herbivores, as well as for soil nutrients.

Interactions with Other Factors

Site History

Disturbance and management history may modify a site's susceptibility to N deposition by shifting relative resource limitation in relation to N supply or demand, or changing soil pH (Bobbink *et al.*, 2010; Dise, 2011). Management factors altering the potential impact of N deposition include the history of N fertilization, burning, grazing, mowing, and modification of vegetation and soil properties. In systems that are strictly N limited, practices that further reduce N availability (e.g., fire, mowing), might be expected to enhance sensitivity to N deposition, whereas practices that increase N availability (e.g., historical N fertilization) might be expected to reduce sensitivity to additional deposition (Bobbink, 1998). However, availability of other resources such as light and P are also affected, thus responses may be far more complex. Grazing is an especially dynamic process, and can increase N availability (through urine and feces), as well as decrease N availability and increase light (through biomass removal). Although the former tends to reduce biodiversity, the latter tends to increase it (Collins *et al.*, 1998). Historical addition of lime ($CaCO_3$) would likely reduce sensitivity to acidification and subsequent cation depletion. In total, there are numerous factors related to site history that can modify the impact of nitrogen deposition on the biodiversity of a particular area.

Can Systems Naturally Recover from Nitrogen-deposition Induced Changes in Biodiversity?

The potential for terrestrial biodiversity to recover following reductions in N deposition is an active and relatively new area of research. As described above, few studies have examined the impacts of added N on biodiversity at levels of N input comparable to N deposition; and, even fewer have examined recovery patterns. Nonetheless, a handful of studies globally are beginning to yield critical information.

For plants, three factors may slow or prevent biodiversity recovery (Bakker and Berendse, 1999; Clark and Tilman, 2010). First, long term N addition may increase N cycling via increases in plant and soil N content, and changes in plant community composition toward more N-rich species (**Figure 7**). Thus, merely stopping N deposition may not effectively reduce N cycling. Second, the availability seeds or propagules of the original species may be limiting, slowing their reestablishment. Third, acidification, toxic mineral buildup, and depletion of base cations could make a region unsuitable to the original species. In grasslands, accumulated litter can also inhibit germination through reducing light levels at the soil surface (Facelli and Pickett, 1991).

A large scale experiment in a Dutch pine forest that reduced N deposition via shelters found after 6 years that nitrogen leaching losses decreased, fungal populations increased, and pine growth and cation balance increased (Boxman *et al.*, 1998). Some other studies have reported similar trends, although others have not (Dise, 2011). For example, in two Swedish forest sites, fungal populations and understory vegetation were still degraded even though nitrogen fertilizer treatment had ceased for 9 and 47 years (Strengbom *et al.*, 2001). Other studies from the US and Europe have found that some important N cycling processes remained elevated 8, 10, 14, and 25 years after treatments ceased to heathland, prairie, shortgrass steppe, and a northeastern forest (reviewed in Clark *et al.*, 2009). An analysis of recovery in a UK heathland found that whereas shoot length, soil pH, and lichens had generally recovered, plant phenology, soil N, and soil microbial activity had not (Power *et al.*, 2006). Analyses of soil seed banks suggest that seeds of many target species (e.g., perennial forbs) do not survive more than a few years to a decade in the soil, and their germination can be suppressed after long-term nitrogen deposition (Thompson *et al.*, 1998). Thus, unless there are refugia nearby of target populations, once lost from the landscape, species recovery may be particularly slow.

Much less is known about how species other than plants and soil biota may or may not recover following reductions in N deposition. For insects and other animals, it is generally assumed that recovery of the plant community would promote recovery in the insect and animal communities. However, especially for large and/or nonflying animals, recovery may be additionally impaired by their ability to move to the recovered habitat. With extensive modification of the landscape from agriculture and urbanization, such movement may be difficult.

It is clear that the recovery potential varies widely among systems and for different processes within systems. Generally, recovery strongly depends on the degree of degradation that has occurred, and the strength of the aforementioned

processes in maintaining the degraded community. It appears that fast-cycling processes such as nitrate leaching and plant nutrient concentrations may recover fairly quickly, whereas slower cycling processes such as decomposition and plant populations might recover much more slowly if at all. Thus, recovery of terrestrial biodiversity over time scales of interest to land managers (years to a few decades) may require management intervention (Dise, 2011).

Management Options to Prevent Degradation and Restore Biodiversity

Monitoring and Modeling

Monitoring networks that measure N deposition rates have been established within the US through the Environmental Protection Agency's (EPA) National Atmospheric Deposition Program (NADP) and the Clean Air Status and Trends Network (CASTNET), and in Europe through the European Monitoring and Evaluation Program (EMEP). These networks provide national scale data on rates of nitrogen deposition. Similar networks are rare in the rest of the world, with only scattered monitoring stations available in most other regions. These networks are critical toward advancing our understanding of nitrogen deposition. Nonetheless, they do have limitations, including: (1) not all nitrogenous species are measured (e.g., NH_3, organic N), (2) not all mechanisms of deposition are accurately and regularly assessed (esp. dry and fog deposition), (3) monitoring stations are generally lacking in remote areas or areas with complex terrain (Pardo *et al.*, in press; Weathers *et al.*, 2006). Several modeling efforts have been developed to try and address some of these issues (esp. the sparse coverage), including the EPA's Community Multiscale Air Quality (CMAQ). These models are complex three-dimensional atmospheric transport and chemistry models that simulate deposition from emission sources to deposition sites. Though these modeling efforts are major contributions, they are also limited by our own understanding of process and by a lack of data to calibrate modeling runs.

Critical Loads

In Europe work has been ongoing over the past few decades to establish critical loads for atmospheric pollution under the framework of the Convention on Long-range Transboundary Air pollution (Dise, 2011). Critical loads are defined as "a quantitative estimate of an exposure to one or more pollutants below which significant harmful effects on specified sensitive elements of the environment do not occur according to present knowledge (Nilsson and Grennfelt, 1988). They are used as a guide to determine when and where ecosystems are vulnerable to degradation. Potential end points include enhanced leaching of soil nitrate, soil acidification, losses of biodiversity, and changes in species composition. Generally, critical loads for biodiversity are estimated based on empirical estimate, using experiments and observations (Bobbink *et al.*, 2010; Pardo *et al.*, in press). Models have also been employed that include biogeochemical cycling models and/or vegetation models (De Vries *et al.*, 2010). Similar efforts in the US have been much less developed until only recently (Pardo *et al.*, 2011, in press). A summary of these critical loads for the US and Europe is shown in **Figure 9**. For many systems, the critical loads in the US are estimated to be lower than in Europe, perhaps because the EU has experienced high N deposition levels for longer periods, with observed systems already impacted and changes only being detectable at higher input rates (Dise, 2011; Pardo *et al.*, in press). In either case, much of Europe and the Eastern US experience N deposition at or above suggested critical loads.

Critical loads can be an effective tool for protection of biodiversity when are used to guide air pollution regulatory policy. Application of critical loads in Europe has connected science to policy by providing scientific methodologies to define pollution limits and to assist in setting reduction targets within a broad multination policy framework. In the US, nitrogen critical loads are only beginning to be developed nationally, and are not currently used as a basis for regulatory policy. Efforts are underway to set a secondary standard for SO_2 and NO_2 concentrations based on ecological effects under the National Ambient Air Quality Standards in the US. However, terrestrial biodiversity impacts may not be included in setting secondary standards, largely because of insufficient data for setting N deposition thresholds for terrestrial biodiversity effects in the US.

Intervention and Policy

Intervention approaches generally aim to either reduce N deposition or enhance recovery. Reducing N deposition can occur through many policy approaches, including the establishment of the aforementioned critical loads, and through allowing tradable permits for pollution which are slowly removed from the market (thereby reducing pollution). An example is the Clean Air Markets Division of the US EPA. In many countries emissions and deposition of NO_x (and especially SO_x) have decreased in the past 20 years as a result of regulatory policies. However, similar controls for emissions of ammonia are less prevalent and the proportion of N deposition occurring in reduced forms (NH_x) is increasing in many areas above levels known to have ecological effects on sensitive taxa (Clarisse *et al.*, 2009; Fenn *et al.*, 2010). This highlights that special consideration as to which receptor to use is required prior to implementing critical loads. Plant and lichen biodiversity are impacted at lower air pollution levels than for human health impacts, and nonvascular biodiversity impacts are generally lower than vascular biodiversity impacts. Thus, when air pollution standards are determined primarily or solely by human health impacts, in many cases sensitive ecosystems and biodiversity will not be effectively protected.

Reduction of N deposition, however, may not be sufficient after decades of exposure rendering management efforts as necessary to promote recovery. Recovery can be promoted generally through two processes: (1) restoring the nitrogen cycle and other resource conditions to their predeposition state, and (2) enhancing the growth and productivity of target species of value. Restoring the nitrogen cycle can be fairly difficult, because many ecosystems are very efficient at retaining this critical nutrient (Chapin *et al.*, 2002; Vitousek

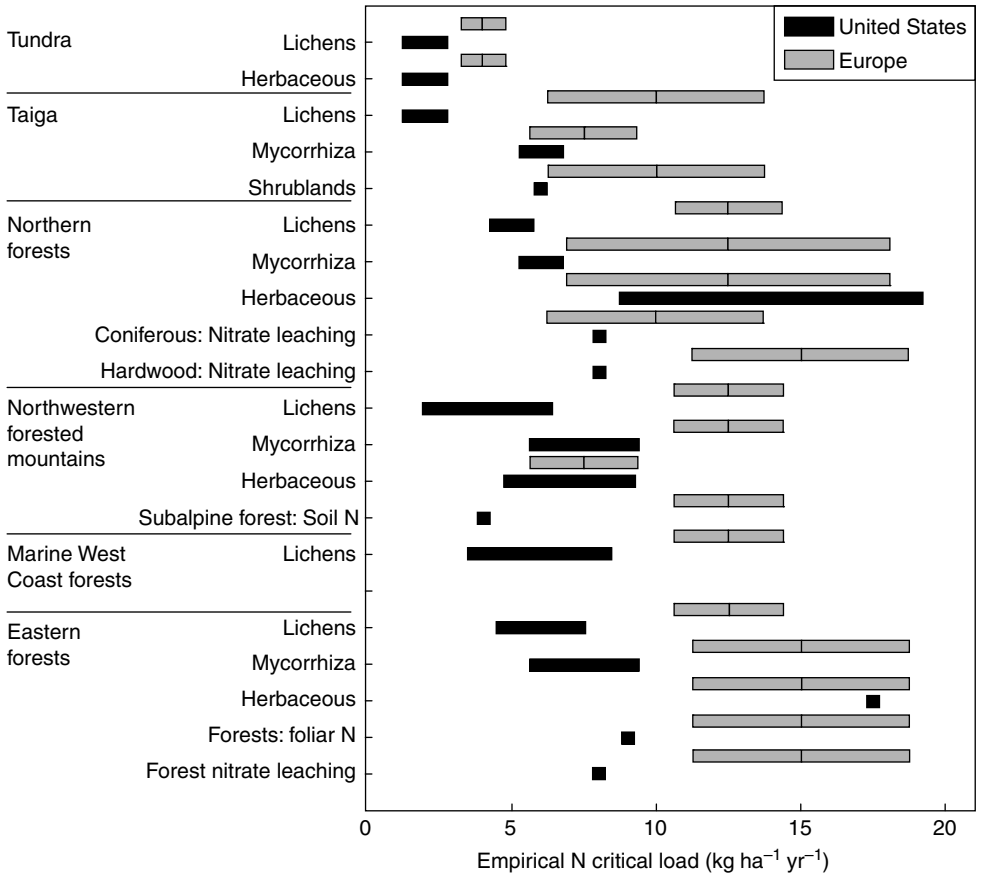


Figure 9 Comparison of critical loads from the US and Europe (Pardo *et al.*, in press, 2011). Included in the bars are many target receptors for nutrient nitrogen (e.g., nitrate leaching, plant biodiversity decline, etc.). Reproduced from Pardo LH, Fenn M, Goodale CL, *et al.* (2011) Effects of nitrogen deposition and empirical nitrogen critical loads for ecoregions of the United States. *Ecological Applications* 21: 3049–3082.

et al., 2002). However, several approaches have been explored to decrease nitrogen availability and restore predeposition conditions including increasing N export through harvesting or fire, increasing N leaching to the groundwater through flushing with aqueous solutions, and decreasing N availability through the addition of carbon (Bakker and Berendse, 1999; Blumenthal *et al.*, 2003). The addition of a carbon source to the soil often triggers soil microbes to take up soil N, thereby decreasing N availability to plants. Other soil and ecological conditions may have to be restored to promote recovery, including increasing pH through the addition of lime, or decreasing pest pressures through application of pesticides. However, restoration of nitrogen, soil, and other ecological conditions has no guarantee that the original species will return. This is more of a concern for grasslands than for forests, because forests respond much more slowly and have not generally experienced large changes in composition (Pardo *et al.*, in press). Because adults of grassland species may no longer be present in the regional landscape, and seeds in the seedbank may no longer be viable, seed addition may be required. For example, in experimental plots in Minnesota and in Kansas, seed addition was required to increase the biodiversity of target species even though individuals in undisturbed areas were less than a few hundred meters away (Clark and Tilman, 2010; Foster *et al.*, 2007). An experiment in Minnesota that isolated the effects from several aforementioned mechanisms found that seed addition and successful germination led to the greatest recovery of biodiversity; and, studies from Australia and the Netherlands found that restoration of soil conditions was not sufficient to induce community recovery.

In total, research suggests that reduction in N deposition is a necessary, but possibly not sufficient, condition for recovery of terrestrial biodiversity. Additional intervention to restore soil conditions and/or plant populations may be necessary (esp. for nontree species), depending on many plant, soil, and ecosystem characteristics.

Conclusions and Next Steps

Nitrogen deposition, along with habitat losses and climate change, is known to be a primary threat to terrestrial biodiversity worldwide. Plant and animal biodiversity usually decline with elevated N in most biomes around the world. However, there is substantial variation in the magnitude of response from system to system and taxa to taxa, depending on many characteristics that influence ecosystem exposure and sensitivity to this critical nutrient. It is unknown to what degree recovery of biodiversity is likely if policies are put in place (or strengthened) to reduce deposition. Greater international coordination of research efforts

and policy design, especially in the area of impacts assessment and critical loads estimation, could enhance management of the impacts of nitrogen on terrestrial biodiversity.

Appendix

List of Courses

1. Ecology
2. Global Change Ecology
3. Atmospheric Pollution

See also: Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems. Air Pollution. Biogeochemical Cycles. Eutrophication and Oligotrophication. Government Legislation and Regulations in the United States. Human Impact on Biodiversity, Overview. Nitrogen, Nitrogen Cycle. Terrestrial Ecosystems

References

Adams M, Ineson P, Binkley D, *et al.* (2004) Soil functional responses to excess nitrogen inputs at global scale. *Ambio* 33: 530–536.

Aerts R and Chapin FS (2000) The mineral nutrition of wild plants revisited: A re-evaluation of processes and patterns. *Advances in Ecological Research* 30.

Allen E, Rao LE, Steers RJ, Bytnerowicz A, and Fenn M (2009) Impacts of atmospheric nitrogen deposition on vegetation and soils at Joshua Tree National Park. In: Webb RH, Fenstermaker LF, Heaton JS, Hughson DL, McDonald EV, and Miller DM (eds.) *The Mojave Desert: Ecosystem Processes and Sustainability*. Las Vegas, NV: University of Nevada Press.

Bai YF, Wu JG, Clark CM, *et al.* (2010) Tradeoffs and thresholds in the effects of nitrogen addition on biodiversity and ecosystem functioning: Evidence from inner Mongolia Grasslands. *Global Change Biology* 16: 358–372.

Bakker JP and Berendse F (1999) Constraints in the restoration of ecological diversity in grassland and heathland communities. *Trends In Ecology & Evolution* 14: 63–68.

Belyazid S, Kurz D, Braun S, Sverdrup H, Rihm B, and Hettelingh JP (2011) A dynamic modelling approach for estimating critical loads of nitrogen based on plant community changes under a changing climate. *Environmental Pollution* 159: 789–801.

Bennie J, Hill MO, Baxter R, and Huntley B (2006) Influence of slope and aspect on long-term vegetation change in British chalk grasslands. *Journal of Ecology* 94: 355–368.

Bertness MD, Crain C, Holdredge C, and Sala N (2008) Eutrophication and consumer control of New England salt marsh primary productivity. *Conservation Biology* 22: 131–139.

Blumenthal DM, Jordan NR, and Russelle MP (2003) Soil carbon addition controls weeds and facilitates prairie restoration. *Ecological Applications* 13: 605–615.

Bobbink R (1998) Impacts of tropospheric ozone and airborne nitrogenous pollutants on natural and seminatural ecosystems: A commentary. *New Phytologist* 139: 161–168.

Bobbink R, *et al.* (2010) Global assessment of nitrogen deposition effects on terrestrial plant diversity: A synthesis. *Ecological Applications* 20: 30–59.

Bowman WD and Bilbrough CJ (2001) Influence of a pulsed nitrogen supply on growth and nitrogen uptake in alpine graminoids. *Plant and Soil* 233: 283–290.

Bowman WD, Cleveland CC, Halada L, Hresko J, and Baron JS (2008) Negative impact of nitrogen deposition on soil buffering capacity. *Nature Geoscience* 1: 767–770.

Bowman WD, Gartner JR, Holland K, and Wiedermann M (2006) Nitrogen critical loads for alpine vegetation and terrestrial ecosystem response: Are we there yet. *Ecological Applications* 16: 1183–1193.

Boxman AW, van der Ven PJM, and Roelefs JGM (1998) Ecosystem recovery after a decrease in nitrogen input to a Scots pine stand at Ysselsteyn, the Netherlands. *Forest Ecology and Management* 101: 155–163.

Bret-Harte MS, Mack MC, Goldsmith GR, *et al.* (2008) Plant functional types do not predict biomass responses to removal and fertilization in Alaskan tussock tundra. *Journal of Ecology* 96: 713–726.

Britton AJ and Fisher JM (2007) Interactive effects of nitrogen deposition, fire and grazing on diversity and composition of low-alpine prostrate *Calluna vulgaris* heathland. *Journal of Applied Ecology* 44: 125–135.

Butchart SHM, *et al.* (2010) Global Biodiversity: Indicators of recent declines. *Science* 328: 1164–1168.

Chapin FS, Matson PA, and Mooney HA (2002) *Principles of Terrestrial Ecosystem Ecology*. New York: Springer-Verlag.

Clarisse L, Clerbaux C, Dentener F, Hurtmans D, and Coheur PF (2009) Global ammonia distribution derived from infrared satellite observations. *Nature Geoscience* 2: 479–483.

Clark CM, Cleland EE, Collins SL, *et al.* (2007) Environmental and plant community determinants of species loss following nitrogen enrichment. *Ecology Letters* 10: 596–607.

Clark CM and Tilman D (2008) Loss of plant species after chronic low-level nitrogen deposition to prairie grasslands. *Nature* 451: 712–715.

Clark CM and Tilman D (2010) Recovery of plant diversity following N cessation: Effects of recruitment, litter, and elevated N cycling. *Ecology* 91: 3620–3630.

Cleland EE, Chiariello NR, Loarie SR, Mooney HA, and Field CB (2006) Diverse responses of phenology to global changes in a grassland ecosystem. *Proceedings of The National Academy of Sciences of The United States of America* 103: 13740–13744.

Collins SL, Knapp AK, Briggs JM, Blair JM, and Steinauer EM (1998) Modulation of diversity by grazing and mowing in native tallgrass prairie. *Science* 280: 745–747.

De Vries W, *et al.* (2010) Use of dynamic soil-vegetation models to assess impacts of nitrogen deposition on plant species composition: An overview. *Ecological Applications* 20: 60–79.

Dentener F, *et al.* (2006) Nitrogen and sulfur deposition on regional and global scales: A multimodel evaluation. *Global Biogeochemical Cycles* 20.

Dise N, Mike A, Salim B, *et al.* (2011) Nitrogen as a threat to European terrestrial biodiversity. In: Sutton MA (ed.) *The European Nitrogen Assessment*. Cambridge: Cambridge University Press.

Dupre C, Stevens CJ, Ranke T, *et al.* (2010) Changes in species richness and composition in European acidic grasslands over the past 70 years: The contribution of cumulative atmospheric nitrogen deposition. *Global Change Biology* 16: 344–357.

Elser JJ, Bracken MES, Cleland EE, *et al.* (2007) Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters* 10: 1135–1142.

Evans CA, Miller EK, and Friedland AJ (2001) Effect of nitrogen and light on nutrient concentrations and associated physiological responses in birch and fir seedlings. *Plant and Soil* 236: 197–207.

Facelli JM and Pickett STA (1991) Plant litter – its dynamics and effects on plant community structure. *Botanical Review* 57: 1–32.

Fenn ME, Jovan S, Yuan F, Geiser L, Meixner T, and Gimeno BS (2008) Empirical and simulated critical loads for nitrogen deposition in California mixed conifer forests. *Environmental Pollution* 155: 492–511.

Fenn ME, *et al.* (2010) Nitrogen critical loads and management alternatives for N-impacted ecosystems in California. *Journal of Environmental Management* 91: 2404–2423.

Foster BL, Murphy CA, Keller KR, Aschenbach TA, Questad EJ, and Kindscher K (2007) Restoration of prairie community structure and ecosystem function in an abandoned hayfield: A sowing experiment. *Restoration Ecology* 15: 652–661.

Fynn RWS and O'Connor TG (2005) Determinants of community organization of a South African mesic grassland. *Journal of Vegetation Science* 16: 93–102.

Galloway JN, Aber JD, Erisman JW, *et al.* (2003) The nitrogen cascade. *Bioscience* 53: 341–356.

Galloway JN and Cowling EB (2002) Reactive nitrogen and the world: 200 years of change. *Ambio* 31: 64–71.

Galloway JN, Schlesinger WH, Levy H III, Michaels A, and Schnoor JL (1995) Nitrogen fixation: anthropogenic enhancement-environmental response. *Global Biogeochemical Cycles* 9: 235–252.

Galloway JN, *et al.* (2004) Nitrogen cycles: Past, present, and future. *Biogeochemistry* 70: 153–226.

Gilliam FS (2006) Response of the herbaceous layer of forest ecosystems to excess nitrogen deposition. *Journal of Ecology* 94: 1176–1191.

Gilliam FS (2007) The ecological significance of the herbaceous layer in temperate forest ecosystems. *BioScience* 57: 845–858.

Haddad NM, Haarstad J, and Tilman D (2000) The effects of long-term nitrogen loading on grassland insect communities. *Oecologia* 124: 73–84.

Horswill P, O'Sullivan O, Phoenix GK, Lee JA, and Leake JR (2008) Base cation depletion, eutrophication and acidification of species-rich grasslands in response to long-term simulated nitrogen deposition. *Environmental Pollution* 155: 336–349.

Johnson NC, Rowland DL, Corkidi L, Egerton-Warburton LM, and Allen EB (2003) Nitrogen enrichment alters mycorrhizal allocation at five mesic to semiarid grasslands. *Ecology* 84: 1895–1908.

Krupa SV (2003) Effects of atmospheric ammonia (NH_3) on terrestrial vegetation: A review. *Environmental Pollution* 124: 179–221.

LeBauer DS and Treseder KK (2008) Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology* 89: 371–379.

Maskell LC, Smart SM, Bullock JM, Thompson K, and Stevens CJ (2010) Nitrogen deposition causes widespread loss of species richness in British habitats. *Global Change Biology* 16: 671–679.

Matson PA, McDowell WH, Townsend AR, and Vitousek PM (1999) The globalization of N deposition: Ecosystem consequences in tropical environments. *Biogeochemistry* 46: 67–83.

McKinney ML and Lockwood JL (1999) Biotic homogenization: A few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution* 14: 450–453.

MEA (2005) *Millenium Ecosystem Assessment*. Washington DC: Island Press.

Morecroft MD, Sellers EK, and Lee JA (1994) An experimental investigation into the effects of atmospheric nitrogen deposition on two semi-natural grasslands. *Journal of Ecology* 82: 475–483.

Nordin A, Strengbom J, Witzell J, Nasholm T, and Ericson L (2005) Nitrogen deposition and the biodiversity of boreal forests: Implications for the nitrogen critical load. *Ambio* 34: 20–24.

Pardo LH, Fenn M, Goodale CL, *et al.* (2011) Effects of nitrogen deposition and empirical nitrogen critical loads for ecoregions of the United States. *Ecological Applications* 21: 3049–3082.

Pardo LH, Robin-Abbott MJ, Driscoll CT (2011) Assessment of nitrogen deposition effects and empirical critical loads of nitrogen for ecoregions of the United States. U.S. Department of Agriculture, Forest Service, Northern Research Station. Report no. NRS-80.

Phoenix GK, *et al.* (2006) Atmospheric nitrogen deposition in world biodiversity hotspots: The need for a greater global perspective in assessing N deposition impacts. *Global Change Biology* 12: 470–476.

Power SA, Ashmore MR, Cousins DA, and Sheppard LJ (1998) Effects of nitrogen addition on the stress sensitivity of *Calluna vulgaris*. *New Phytologist* 138: 663–673.

Power SA, Green ER, Barker CG, Bell JNB, and Ashmore MR (2006) Ecosystem recovery: Heathland response to a reduction in nitrogen deposition. *Global Change Biology* 12: 1241–1252.

Rao LE, Allen EB, and Meixner T (2010) Risk-based determination of critical nitrogen deposition loads for fire spread in southern California deserts. *Ecological Applications* 20: 1320–1335.

Roelofs JGM, Kempers AJ, Houdijk A, and Jansen J (1985) The effect of air-borne ammonium-sulfate on pinus-nigra-var-maritima in the Netherlands. *Plant and Soil* 84: 45–56.

Sala OE, *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.

Sheppard LJ, Leith ID, Crossley A, *et al.* (2008) Stress responses of *Calluna vulgaris* to reduced and oxidised N applied under 'real world conditions'. *Environmental Pollution* 154: 404–413.

Siddique I, Vieira ICG, Schmidt S, *et al.* (2010) Nitrogen and phosphorus additions negatively affect tree species diversity in tropical forest regrowth trajectories. *Ecology* 91: 2121–2131.

Silvertown J, Poulton P, Johnston E, Edwards G, Heard M, and Biss PM (2006) The Park Grass Experiment 1856–2006: Its contribution to ecology. *Journal of Ecology* 94: 801–814.

Smart SM, Bunce RGH, Marrs R, *et al.* (2005) Large-scale changes in the abundance of common higher plant species across Britain between 1978, 1990 and 1998 as a consequence of human activity: Tests of hypothesised changes in trait representation. *Biological Conservation* 124: 355–371.

Smil V (2001) *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of Food Production*. Cambridge: MIT Press.

Stevens CJ, Dise NB, Mountford JO, and Gowing DJ (2004) Impact of nitrogen deposition on the species richness of grasslands. *Science* 303: 1876–1879.

Stevens CJ, Thompson K, Grime JP, Long CJ, and Gowing DJG (2010a) Contribution of acidification and eutrophication to declines in species richness of calcifuge grasslands along a gradient of atmospheric nitrogen deposition. *Functional Ecology* 24: 478–484.

Stevens CJ, *et al.* (2010b) Nitrogen deposition threatens species richness of grasslands across Europe. *Environmental Pollution* 158: 2940–2945.

Stevens CJ, *et al.* (2011) Ecosystem responses to reduced and oxidised nitrogen inputs in European terrestrial habitats. *Environmental Pollution* 159: 665–676.

Strengbom J, Nordin A, Nasholm T, and Ericson L (2001) Slow recovery of boreal forest ecosystem following decreased nitrogen input. *Functional Ecology* 15: 451–457.

Strengbom J, Nordin A, Nasholm T, and Ericson L (2002) Parasitic fungus mediates change in nitrogen-exposed boreal forest vegetation. *Journal of Ecology* 90: 61–67.

Suding KN, Collins SL, Gough L, *et al.* (2005) Functional- and abundance-based mechanisms explain diversity loss due to N fertilization. *Proceedings of the National Academy of Sciences of the United States of America* 102: 4387–4392.

Sverdrup H, Belyazid S, Nihlgard B, and Ericson L (2007) Modelling change in ground vegetation response to acid and nitrogen pollution, climate change and forest management at in Sweden 1500–2100 A.D. *Water, Air, and Soil Pollution Focus* 7: 163–179.

Thomas RQ, Canham CD, Weathers KC, and Goodale CL (2010) Increased tree carbon storage in response to nitrogen deposition in the US. *Nature Geoscience* 3: 13–17.

Thompson K, Bakker JP, Bekker RM, and Hodgson JG (1998) Ecological correlates of seed persistence in soil in the north-west European flora. *Journal of Ecology* 86: 163–169.

Throop HL (2005) Nitrogen deposition and herbivory affect biomass production and allocation in an annual plant. *Oikos* 111: 91–100.

Throop HL and Lerdau MT (2004) Effects of nitrogen deposition on insect herbivory: Implications for community and ecosystem processes. *Ecosystems* 7: 109–133.

Throop HL, Holland EA, Parton WJ, Ojima DS, and Keough CA (2004) Effects of nitrogen deposition and insect herbivory on patterns of ecosystem-level carbon and nitrogen dynamics: Results from the CENTURY model. *Global Change Biology* 10: 1092–1105.

Treseder KK (2004) A meta-analysis of mycorrhizal responses to nitrogen, phosphorus, and atmospheric CO_2 in field studies. *New Phytologist* 164: 347–355.

Ulrich B (1983) Soil acidity and its relation to acid deposition. In: Ulrich B and Pankrath J (eds.) *Effects of Accumulation of Air Pollutants in Ecosystems*, pp. 127–146. Boston: Reidel Publishing.

Vitousek PM, Aber JD, Howarth RW, *et al.* (1997) Human alteration of the global nitrogen cycle: Sources and consequences. *Ecological Applications* 7: 737–750.

Vitousek PM, *et al.* (2002) Towards an ecological understanding of biological nitrogen fixation. *Biogeochemistry* 57: 1–45.

Weathers KC, Simkin SM, Lovett GM, and Lindberg SE (2006) Empirical modeling of atmospheric deposition in mountainous landscapes. *Ecological Applications* 16: 1590–1607.

Xia JY and Wan SQ (2008) Global response patterns of terrestrial plant species to nitrogen addition. *New Phytologist* 179: 428–439.

Zavaleta ES, Shaw MR, Chiariello NR, Mooney HA, and Field CB (2003) Additive effects of simulated climate changes, elevated CO_2, and nitrogen deposition on grassland diversity. *Proceedings of the National Academy of Sciences USA* 100: 7650–7654.

Nitrogen, Nitrogen Cycle

Sandy L Tartowski, US Department of Agriculture, USA
Robert W Howarth, Cornell University, Ithaca, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biogeochemistry The discipline which studies biotic controls on the chemistry of the environment and geochemical controls on the structure and function of ecosystems.
Deposition The delivery of material inputs to the earth's surface from the atmosphere.
Eutrophic High in production or species typical of high-production environments.
Mineralization The conversion of an element from an organically bound form to an inorganic form.
Nitrification The oxidation of ammonia to nitrite and nitrate by bacteria.
Nitrogen fixation The reduction of atmospheric N_2 to ammonia or other organic or inorganic compounds by bacteria, lightning, or photochemistry.
Nitrogen saturation Nitrogen supply in excess of the capacity of the ecosystem to retain nitrogen.
Nutrient limitation The addition of a nutrient or nutrients to an ecosystem causes an increase in net primary production.
Oligotrophic Low in production or species typical of low-production environments.
Opportunistic species Species typical of transient, unstable, unpredictable, frequently disturbed, or periodically extreme environments, usually having strong dispersal abilities and faster growth, smaller size, and shorter life spans than potential competitors.

Introduction

Nitrogen is an essential element for life and it frequently controls productivity and community structure in both terrestrial and aquatic ecosystems. During the past several decades, human activity has changed the nitrogen cycle more than any other element cycle, resulting in many deleterious environmental changes, including decreased biotic diversity.

An Overview of the Nitrogen Cycle and its Ecological Importance

Nitrogen (N) is an essential element for all life on Earth, a required component of deoxyribonucleic acid (DNA), all proteins (including enzymes), chlorophyll, and many other critical constituents of biological structure and function. More than 99% of nitrogen on Earth is present as molecular N_2, with most of this in the atmosphere and some dissolved as a gas in the oceans. However, most organisms cannot use the relatively inert N_2 molecule. Paradoxically, molecular N_2 is not a thermodynamically stable compound in the presence of oxygen, but these two gases coexist and comprise virtually all the atmosphere. The nitrogen in N_2 becomes biologically available only when the N–N bond is broken and nitrogen is combined with other elements to form compounds such as ammonium (NH_4^+) and nitrate (NO_3^-) in a process called nitrogen fixation. Prior to the massive alteration in the global nitrogen cycle by human activity, small amounts of nitrogen were fixed by lightning and by volcanic activity, but the vast majority of nitrogen fixed on Earth each year (more than 95%) was fixed by bacteria. Not all bacteria are capable of fixing nitrogen, although the ability is fairly widespread among diverse groups of bacteria including cyanobacteria, nonoxygenic photosynthetic bacteria (such as purple sulfur bacteria), and many classes of heterotrophic bacteria. No other organisms are capable of nitrogen fixation, although many plants, lichens, mosses, and some animals have bacterial symbionts which fix nitrogen. These plants with symbiotic nitrogen-fixing bacteria include many species of trees (such as alders in the temperate zone and many species in moist tropical forests) and many agricultural crops, including alfalfa, peas, peanuts, soybeans, and rice. Plants with symbiotic nitrogen-fixing bacteria are often loosely called 'nitrogen fixers,' although it is always the bacteria that are actually fixing nitrogen.

For most of the history of life on Earth, the demand by plants for biologically available forms of nitrogen such as nitrate and ammonium has been greater than the rate of supply by nitrogen fixation. Consequently, nitrogen limits the primary productivity in many ecosystems, and when more biologically available nitrogen is added to these systems plant production increases. This relative lack of availability of nitrogen globally has also been a major factor shaping the evolution of life and structuring the communities in many different types of ecosystems, both terrestrial and aquatic.

The general patterns which relate productivity and diversity in ecosystems are well established. In aquatic ecosystems, species diversity declines with increasing productivity, and diversity is greatest in systems of very low productivity, such as oligotrophic lakes and areas of low productivity in the oceans in which colimitation by multiple nutrients (including nitrogen) may be common. In terrestrial ecosystems, the greatest diversity of plant species usually occurs in sites of intermediate fertility and productivity. In sites of very low fertility, the resources can support only a few viable populations able to tolerate the severe conditions. In very high fertility sites, competition intensifies and may shift from competition for nutrients to the more asymmetrical competition for light or space, resulting in dominance by a few species. For both

aquatic and terrestrial systems, when nitrogen is limiting production, it is also influencing diversity.

The importance of nitrogen as a factor limiting productivity and regulating community structure varies among the earth's ecosystems, with nitrogen being of paramount importance in some systems and not limiting at all in others. Although there are exceptions, among terrestrial ecosystems nitrogen availability generally controls production in temperate forests, boreal forests, arctic and alpine tundra systems, temperate and tropical grasslands, and agroecosystems. However, phosphorus or other elements – and not nitrogen – tend to limit production in most tropical forests if any element is limiting, except in forests on very young soils (less than a few tens of thousands year old), where nitrogen is commonly limiting. Among aquatic ecosystems, nitrogen limits production in most estuaries and coastal seas of the temperate zone and perhaps in some tropical seas, although these are sometimes phosphorus limited. Nitrogen with phosphorus as a colimiting element may regulate production in much of the oceans away from shore as well. However, phosphorus is generally more important in regulating production in freshwater lakes, at least in moderately or highly productive lakes in the temperate zone.

These patterns of nutrient limitation are the result of many processes which regulate the availability of nitrogen in comparison to other elements essential for plant growth. For terrestrial ecosystems, one important aspect is the change in phosphorus biogeochemistry as soils develop over geological time. In a geologically young soil, the initial amount of phosphorus is often high, but as the soil is weathered this phosphorus is exported from the system or stored in mineral forms which make the phosphorus biologically unavailable. In the temperate zone, soils are often young due to glaciation within approximately the past 20,000 years. However, many tropical soils are extremely old and highly weathered and therefore have very low availabilities of phosphorus. This is one factor which tends to make tropical forests phosphorus limited and temperate forests nitrogen limited.

Differences in nitrogen processes among systems are also important in determining which element is most limiting to primary productivity. Nitrogen fixation is one critically important process, and it is much more prevalent and occurs at higher rates in tropical forests than in temperate forests. This contributes to the greater tendency for nitrogen limitation in temperate forests. An interesting paradox is that nitrogen fixation does not alleviate nitrogen shortages in temperate forests and thereby alleviate nitrogen limitation. Why should any ecosystem be nitrogen limited if nitrogen fixation could provide sufficient inputs? Why is nitrogen fixation in temperate forests so much less than that in tropical forests? The reasons for the relatively low rates of nitrogen input to nitrogen-limited temperate forests remain poorly known and are an active area of research. A similar paradox occurs in aquatic ecosystems. One of the reasons that freshwater lakes tend to be limited by phosphorus is that planktonic cyanobacteria in these systems can often fix sufficient nitrogen to alleviate any potential nitrogen limitation. In contrast, nitrogen fixation by planktonic cyanobacteria rarely occurs in even strongly nitrogen-limited estuaries. The relative lack of nitrogen fixation in estuaries is probably due to an interaction of two factors: trace metals that are required for the process of nitrogen fixation (molybdenum and iron) are less available in estuaries than in lakes, slowing the potential growth rate of nitrogen-fixing organisms, and this slow growth rate makes cyanobacteria highly susceptible to grazing mortality by zooplankton and benthic animals.

Nitrogen fixation is only one process affecting nitrogen availability in an ecosystem. Also important are the processes which regulate the loss of nitrogen from the system and the input of biologically available forms of nitrogen from neighboring ecosystems and from the atmosphere. A brief summary of the nitrogen cycle illustrates the complexity. The nitrogen that is fixed accumulates over time in both the living and nonliving components of ecosystems and is transported and transformed by numerous processes (**Figure 1**). Ammonium and nitrate are taken up by plants as well as by microorganisms and converted to organic nitrogen compounds. This organic nitrogen is cycled through food webs but also accumulates in soils, water, and sediments as organisms die, excrete matter, or drop tissues such as the fall of leaves from trees. Much of the nitrogen in this nonliving organic matter is not readily available for use by organisms, but microbial decomposition of the organic matter releases ammonium, which can then be oxidized to nitrate by bacteria. Both nitrate and ammonium are highly soluble in water, but the negatively charged nitrate ion is more easily leached from soil and transported to aquatic ecosystems. In ecosystems where oxygen concentrations are low (such as in the sediments of wetlands, lakes, and estuaries), the nitrogen in nitrate can be converted back to molecular N_2 (or the slightly oxidized gas N_2O, which is a very powerful greenhouse gas) by bacteria in a process called denitrification. Also, in soils and waters where the pH is sufficiently high to favor the dissociation of ammonium to ammonia (NH_3), nitrogen can be lost to the atmosphere as ammonia gas. Fires can also volatilize both ammonia and oxidized gases of nitrogen (such as NO and N_2O) to the atmosphere. Most of these forms of nitrogen are redeposited from the atmosphere onto the earth's surface in precipitation or as dry deposition.

Human Impacts on the Nitrogen Cycle

Human activity has altered the nitrogen cycle globally more than that of any other element cycle, and much of the change has occurred during the past 50 years. The increase in carbon dioxide and other greenhouse gases in the atmosphere is, appropriately, a cause of widespread alarm, but the rate at which humans have altered nitrogen availability on land is even greater. As recently as 40–50 years ago, the major source of fixed nitrogen was still natural biological nitrogen fixation, but today more nitrogen is fixed through human activities than by all nitrogen fixation in natural terrestrial ecosystems on Earth (**Figure 2**). The greatest change has been the widespread use of synthetic nitrogen fertilizer, and the production of such fertilizer today accounts for more than half of the total amount of nitrogen fixed by all human activities. The process for making nitrogen fertilizer from atmospheric N_2 was invented in the early twentieth century, but such fertilizer was not widely used until the 1950s. The rate of use has increased

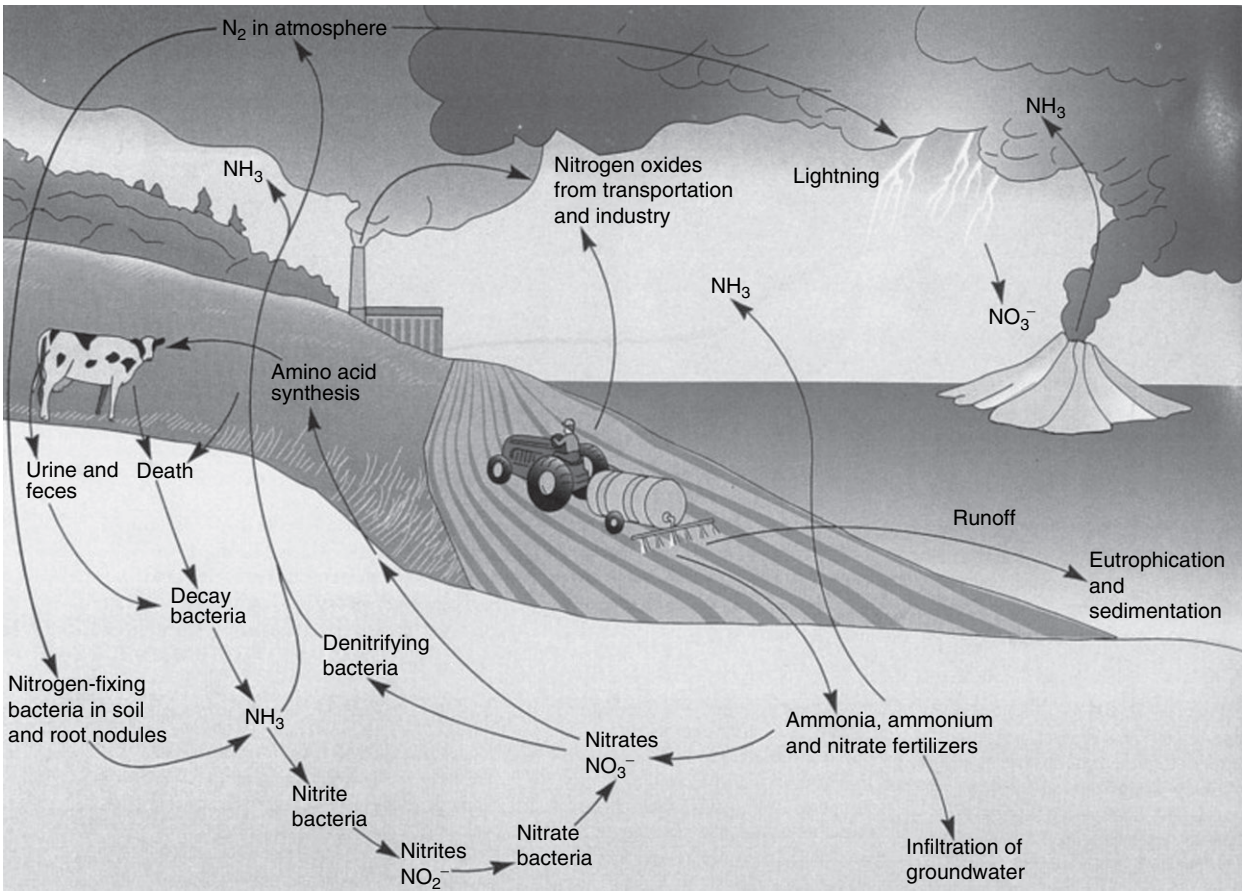


Figure 1 Simplified diagram of the nitrogen cycle. Reproduced with permission from Cunningham and Saigo (1995).

steadily since then, with only a modest interruption caused by the collapse of the former Soviet Union and the disruption of agricultural in Russia and eastern Europe in the early 1990s. Most of the synthetic nitrogen fertilizer that has ever been used on the planet has been used during the past 25 years.

The synthesis of inorganic fertilizer accounts for 60% or more of the total amount of nitrogen fixed globally through human activities, with the rest coming primarily from the production of nitrogen-fixing crops in agriculture and from the combustion of fossil fuels. Fossil fuel combustion not only releases nitrogen from geological storage in the fuel but more importantly catalyzes the reaction of O_2 and N_2 as fossil fuels are burned. This nitrogen is released as oxidized nitrogen gases, generically referred to as NO_x, which are subsequently deposited onto the earth's surface in precipitation or dry deposition. Overall, human fixation of nitrogen (including the production of fertilizer, combustion of fossil fuel, and production of nitrogen-fixing agricultural crops) increased globally approximately three- or fourfold between 1960 and 2010 and continues to increase. By the year 2000, human activities made fixed nitrogen available at a rate of approximately 170 Tg year^{-1}, or more than 1 g N m^{-2} year^{-1} when averaged over the entire land surface of the earth.

Nitrogen from human sources can travel for fairly long distances in the environment. For instance, much of the nitrogen deposited from the atmosphere in the northeastern USA comes from the combustion of fossil fuels in the Midwest, and the majority of nitrogen flowing down the Mississippi River and causing eutrophication in the Gulf of Mexico originates from agricultural sources in Illinois, Iowa, Indiana, Minnesota, and Wisconsin. Of the nitrogen fertilizer applied to agricultural fields, some leaches into ground water and surface water and affects downstream ecosystems; on average in the USA, approximately 20% of nitrogen fertilizer is thus exported from agroecosystems, although this percentage ranges from only 3% to 80% depending on the soil type, climate, and agricultural practices. Nitrogen from agricultural sources also gets volatilized into the atmosphere and thereby redistributed onto nonagricultural lands, including forests. Some of this nitrogen is volatilized directly from the agricultural fields, as both ammonia and NO_x, with the flux of NO_x being particularly important in tropical areas. Also of importance is the volatilization of ammonia to the atmosphere from animal wastes. In developed countries, approximately half of the nitrogen used as fertilizer is exported from the field in crop harvest, on average, with most of these crops being fed to livestock and poultry. Much of the nitrogen in these feedstocks ends up in the animal wastes, with a high percentage being volatilized. The total flux of such volatilization in the USA is almost as large as the leaching of nitrogen directly from fertilized fields. Increasingly over the past several decades, crops may be transported hundreds of

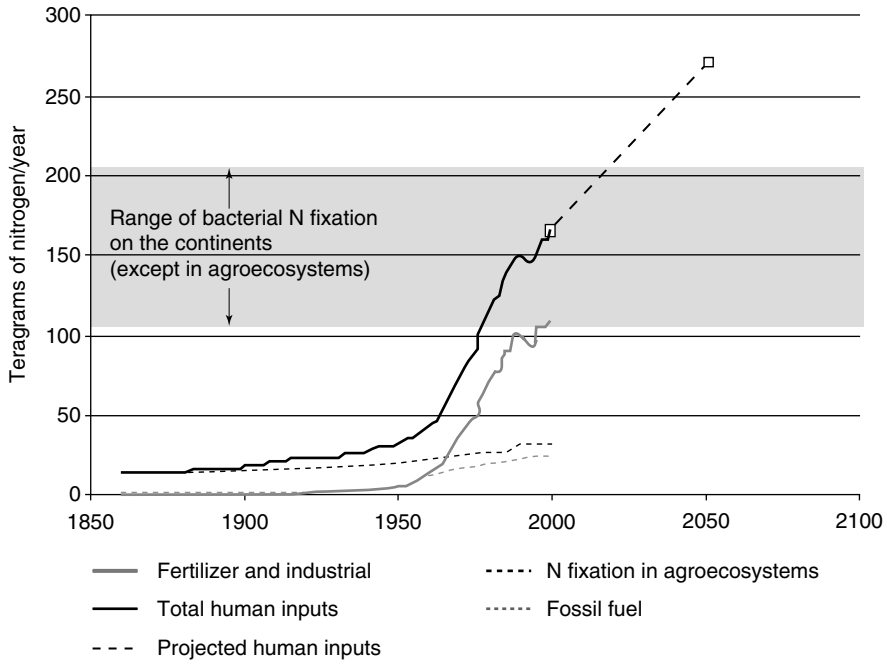


Figure 2 Anthropogenic fixation of nitrogen compared to the range of estimates of natural biological fixation on land. Reprinted with permission from Howarth RW, Ramakrishna K, Choi E, *et al.* (2005) Nutrient management, responses assessment. In: *Ecosystems and Human Well-being*, ch. 9, vol. 3, pp. 295–311. Policy Responses, the Millennium Ecosystem Assessment. Washington, DC: Island Press.

kilometers or more to animal feedlot operations, making it unlikely that the animal wastes are ever transported back to the fields, thereby compounding nitrogen pollution problems. The steady input of new synthetic nitrogen fertilizers makes such practices possible.

Reactive nitrogen compounds are usually transported for only a few hundred kilometers to at most approximately 1000 km through the atmosphere, and rivers and ocean currents transport biologically available forms of nitrogen on the same sort of spatial scales. As a result, the alteration of the nitrogen cycle is not uniform over the earth, and the greatest changes are concentrated in the areas of greatest population density and greatest agricultural production. Generally, the largest changes have occurred in the northern temperate zone. The tropics have seen less change through the twentieth century, but that has been changing over the past decade. Most terrestrial ecosystems and most coastal marine ecosystems in the temperate zone are nitrogen limited, and therefore, the acceleration of nitrogen cycling in the northern temperate zone has had a great impact. Current deposition of nitrogen from the atmosphere onto the landmasses of the Northern Hemisphere temperate zone is almost seven times more than the preindustrial deposition, and in some regions, such as the northeastern USA, Western Europe, and east Asia, the increase has been far greater. The natural rate of biological nitrogen fixation in European watersheds of the North Sea and in the northeastern USA is approximately 0.5 g N m^{-2} $year^{-1}$ or less and human activity has increased nitrogen input to the watersheds of the North Sea to 7.5 g N m^{-2} $year^{-1}$ and in the northeastern USA to 4.1 g N m^{-2} $year^{-1}$. The export of nitrogen in major rivers to coastal oceans in both Europe and North America is closely related to the increased inputs of nitrogen to the basins from human activities (**Figure 3**). Downstream transport of nitrogen to estuaries and coastal oceans has increased up to 20-fold in some areas such as the North Sea. **Figure 4** shows the spatial pattern of these nitrogen inputs from human activity across the lower 48 states of the USA, which drive the increased nitrogen export downstream. Note in particular the high levels of net anthropogenic nitrogen inputs in the corn belt of the upper mid-west, and hot spots along the east and west coasts of the USA associated with concentrated animal agricultural operations and high density of human populations.

The rate at which human activity has accelerated nitrogen cycling varies widely among regions of the world. Globally, the use of synthetic nitrogen fertilizer continues to increase, but in the USA and many other developed countries fertilizer use has changed little since 1980, and fertilizer use has actually fallen in western Europe and even more so in the former Soviet Union. In China, however, the use of synthetic nitrogen fertilizer has increased 2.5-fold from 1980 to 2000 and continues to increase at a rate of over 4% per year. For the past decade or two, China has been using one-third of all inorganic nitrogen fertilizer produced globally. The acceleration of nitrogen use in developing countries has allowed major increases in food production and has greatly reduced starvation, but the environmental consequences are becoming increasingly apparent.

Effects of Nitrogen on Biodiversity

Human acceleration of the nitrogen cycle has tremendous implications for biodiversity at all scales, from the genome to

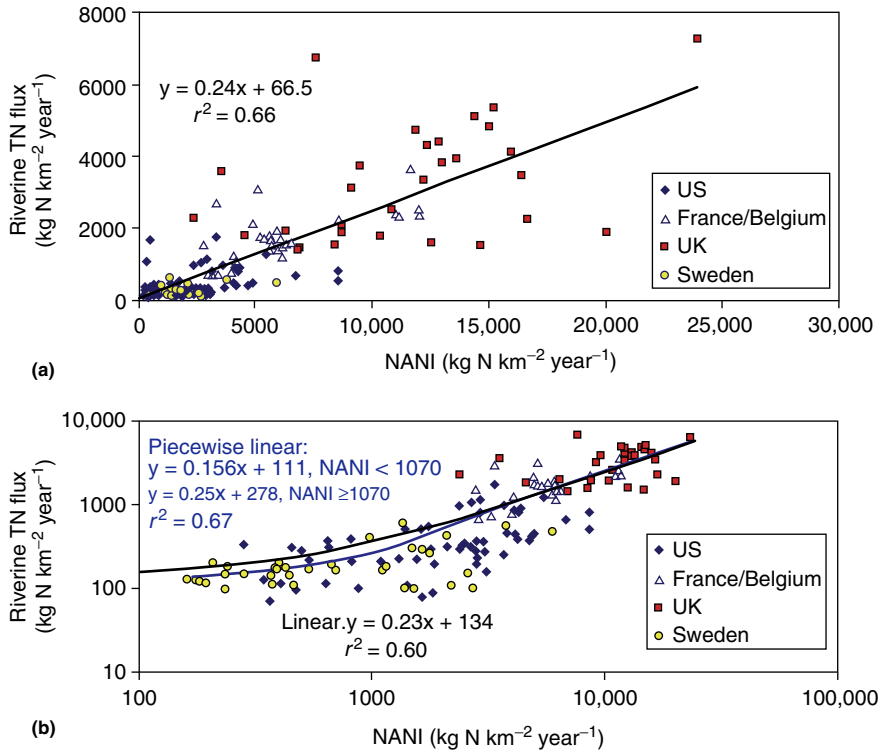


Figure 3 Export of nitrogen from rivers in both North America and Europe is strongly correlated with the net nitrogen inputs to the watersheds (NANI) caused by human activities, especially agriculture and combustion of fossil fuels, on both a linear and log–log scales. Reprinted from Howarth RW, Swaney D, Billen G, *et al.* (2012) Nitrogen fluxes from large watershed to coastal ecosystems controlled by net anthropogenic nitrogen inputs and climate. *Frontiers in Ecology & Environment* doi:10.1890/10017.

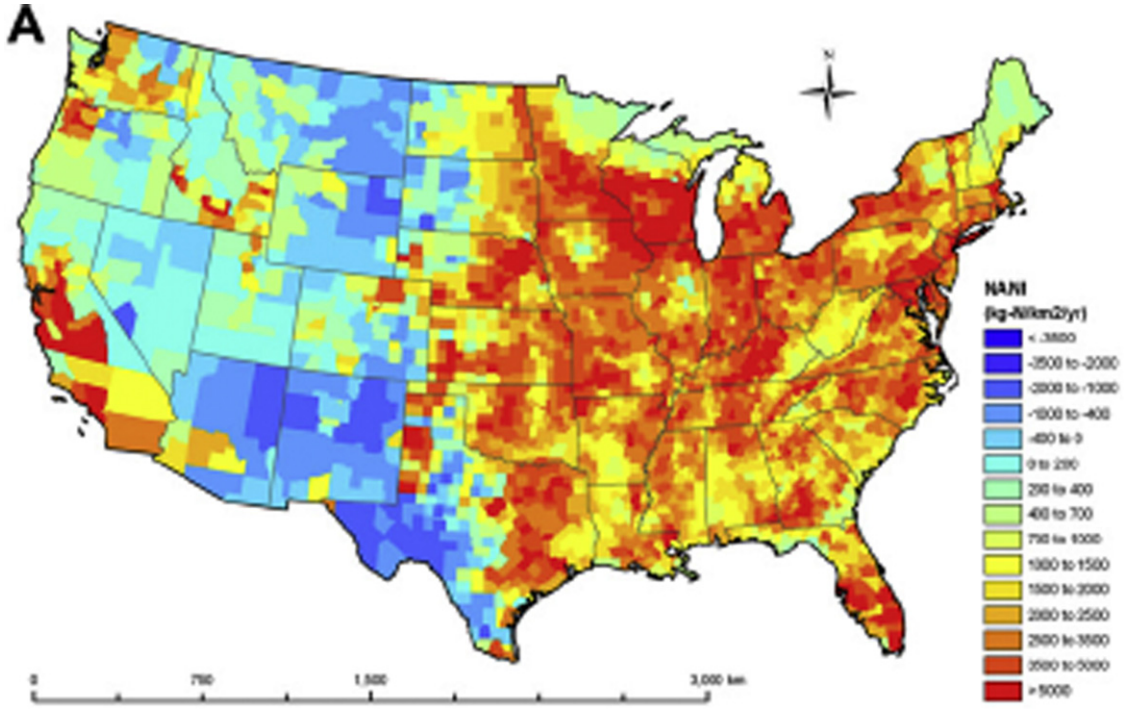


Figure 4 The spatial pattern of the net anthropogenic nitrogen inputs to the landscape across the lower 48 states of the US. Reprinted from Hong B, Swaney D, and Howarth RW (2011) A toolbox for calculating net anthropogenic nitrogen inputs (NANI). *Environmental Modeling and Software* 26: 623–633.

the globe. Reduced genetic diversity within species exposed to high levels of nitrogen deposition has been documented in tree species, soil mycorrhizae, and *Rhizobium spp*. in beans. Changes in species composition and reductions in species diversity are obvious and widespread. Declining ecosystem and landscape diversity are becoming evident. In return, changes in biodiversity influence nitrogen dynamics and processes, creating complex nonlinear interactions and feedbacks between nitrogen cycling and biodiversity. Nitrogen enrichment can reduce biodiversity by several mechanisms including acidifying soil and water as well as increasing the growth and dominance of a few particularly responsive species in nitrogen-limited ecosystems.

Effects of Eutrophication on Biodiversity

The substantial increase in nitrogen inputs to nitrogen-limited ecosystems as a result of human activities alters species composition (which species are present) and decreases species diversity (the number of species and the evenness of their relative abundance). A few fast-growing opportunistic species, with greater ability to take advantage of the increased availability of nitrogen to produce additional biomass, usually become dominant in fertilized ecosystems, while the species characteristic of less fertile environments (oligotrophic species) decline and disappear. Often, a relatively diverse flora, including rare endemic plants, characteristic of an oligotrophic ecosystem is replaced by a more productive, but less diverse flora dominated by a few common weedy species, normally found in mesotrophic or eutrophic sites. Usually, invasive pest plants respond strongly to increased availability of nitrogen, and generally eutrophication facilitates the spread of introduced and native pest plants. As plant species are lost from an ecosystem, additional animal species associated with those plants may be lost as well, further decreasing species diversity. The conversion of whole ecosystems from oligotrophic to eutrophic and the disappearance of entire types of ecosystems encompass the extinction of large groups of species, not just a few particularly vulnerable endangered species, and reduce biodiversity on the local, landscape, and regional scale.

The decrease in species diversity caused by the eutrophication of terrestrial and aquatic ecosystems has been observed consistently across a wide variety of nitrogen-limited ecosystems. However, the speed and magnitude of the effects of increased inputs of nitrogen on biodiversity depend on the characteristics and history of the affected ecosystem. The mobilization and redistribution of nitrogen by human activity can cause local areas of nitrogen depletion, even though the amount of fixed nitrogen has increased overall. Semiarid and arid lands and unfertilized agricultural and overgrazed systems are particularly susceptible to the loss of nitrogen-containing topsoil. In some cases, nitrogen depletion can threaten biodiversity much as nitrogen addition reduces biodiversity. In very unproductive ecosystems, the addition of nitrogen may cause an initial, usually transient, increase in species diversity by allowing the invasion of native species from more fertile sites in the same region or exotic species introduced from other regions. Even when species diversity remains elevated, often the oligotrophic species are lost from the ecosystem, contributing to a regional decline in species diversity. The effects of nitrogen enrichment may be delayed in systems which are severely nitrogen limited due to previous depletion of nitrogen by fire or human activities, such as forest clearing or timber harvesting.

Enrichment of grasslands with nitrogen causes productivity increases in a few dominant grasses and severe declines in species diversity, as shown by experiments in North America, Australia, and Europe (**Figure 5**). In one long-term study of fertilizer addition to a grassland in England, species diversity was reduced by more than fivefold. Oligotrophic, calcareous grasslands in Europe, with high biodiversity and many endemic species, are particularly sensitive to species loss due to eutrophication and most sites have already experienced changes in species composition. Fertilization and increased productivity decrease species diversity by allowing fast growing, tall plants to shade competitors and sometimes by increasing production of litter which prevents the establishment of seedlings. Increased availability of nitrogen suppresses nitrogen-fixing plant species and consequently many of the forbs lost from eutrophied grasslands and heaths are nitrogen-fixing species. The mechanisms by which nitrogen enrichment reduces biodiversity may be quite specific to the circumstances of each ecosystem. For example, when nitrogen was added to fertile wet and infertile dry sedge meadows in Colorado, diversity declined in the wet meadow but not in the dry meadow. In the infertile dry meadow, nitrogen additions released rare species from nitrogen limitation, increasing the evenness of species abundances, but in the more fertile wet meadow the already dominant species increased in abundance, suppressing subordinate species. Nitrogen inputs affect not only plant diversity but also the diversity of the animal community. It has recently been suggested that nitrogen deposition near San Francisco may be a contributing factor leading to the decline of a threatened, endemic butterfly.

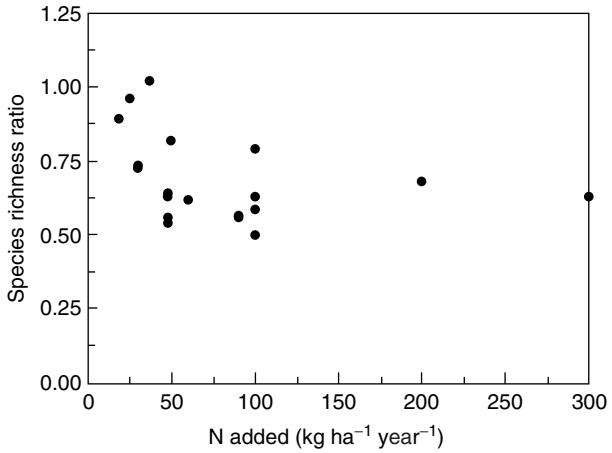


Figure 5 Species richness of plants in grasslands across Europe as a function of nitrogen added in a variety of experiments. Reprinted from Figure 2 in Bobbink R, Hicks K, Galloway J, *et al.* (2010) Global assessment of nitrogen deposition effects on terrestrial plant diversity: A synthesis. *Ecological Applications* 20: 30–59, with permission from ESA.

In many ecosystems, moderate levels of grazing – within the range typical of the evolutionary history of the system and plants – tend to increase plant species diversity. Grazing can interact with nitrogen availability to affect diversity, and ungrazed or lightly grazed grasslands are more susceptible to decreases in biodiversity caused by nitrogen enrichment because heavy grazing usually suppresses the dominant grass species stimulated by nitrogen. Nitrogen enrichment usually reduces the spatial heterogeneity of nitrogen and therefore reduces the potential for different species to coexist using microsites with different availabilities of nitrogen. Grazing redistributes nitrogen and usually increases the spatial heterogeneity and temporal variability of nitrogen availability, contributing to the maintenance of biodiversity. However, when fertilization increases productivity sufficiently to support greatly increased grazing, grazing-intolerant plant species may be lost from the ecosystem.

Diversity in forests is also affected by nitrogen additions, as has been particularly well studied in parts of Europe in which nitrogen deposition from the atmosphere is high. In both deciduous (the Netherlands and France) and coniferous forest (Sweden and Finland), nitrogen enrichment has caused shifts in the species composition of the understory, favoring opportunistic species with high nitrogen requirements and reducing biodiversity of lichens, mosses, and vascular plants. Increased nitrogen deposition has caused weakened tree trunks, decreased frost hardiness, and increased pest outbreaks in some forests. After many years of high nitrogen deposition, some forests of Western Europe have become 'nitrogen saturated,' unable to store additional nitrogen in soil or vegetation. In nitrogen-saturated ecosystems, continuing nitrogen inputs are balanced by comparable nitrogen outputs, accelerating both the eutrophication of downstream waters and the loss of other essential elements from the forest.

In the Netherlands, the high population density, with accompanying industry and vehicles, is combined with intensive livestock production to generate perhaps the highest rates of nitrogen deposition in the world. More than 35% of oligotrophic, diverse heathlands in the Netherlands have suffered nitrogen accumulation, invasion by grasses, and reduced diversity. In dry heath-lands, this nitrogen fertilization has stimulated outbreaks of the heather beetle, accelerating the conversion to grassland. Some heaths have been completely converted to much less diverse grasslands and forest, reducing the landscape and regional diversity as well as the diversity within the heathlands. The beginnings of similar eutrophication of heathlands have been observed in Norway and in the UK.

Estuaries and other coastal waters receive a variety of insults from human activities. Perhaps the greatest pollution problem, however, is the eutrophication that results from increased nitrogen inputs. Eutrophication of coastal marine ecosystems frequently results in hypoxia (the depletion of dissolved oxygen) or anoxia (the absence of dissolved oxygen), both of which result not only in fish kills and the loss of other biotic resources but also in sharp decreases in biotic diversity. Hypoxia has severe effects on the diversity and abundance of benthic species, shifting dominance from large, long-lived, less mobile species to smaller, mobile, opportunistic, short-lived species. Frequent hypoxia may prevent successional development beyond the early colonizing community. The loss of burrowing benthic organisms that irrigate and oxygenate the sediments may strongly influence biogeochemical processes such as phosphorus adsorption and nitrification and denitrification, which can feed back to influence diversity. Increased nitrogen inputs have caused increased hypoxia and anoxia in Chesapeake Bay, the Baltic Sea, the Black Sea, Long Island Sound, Florida Bay and the Florida Keys, the northern Adriatic Sea, and many other areas globally. The hypoxic 'dead zone' in the northern Gulf of Mexico increased from 9500 km^2 when first reported in 1991 to well over 20,000 km^2 in most years since 1999 and is the direct result of nitrogen inputs coming down the Mississippi River.

Not all the effects of eutrophication on biodiversity are due to hypoxia and anoxia. Nutrient enrichment also leads directly to changes in the composition of the phytoplankton community, and these can cascade up the food web generally leading to lowered diversity at each trophic level. Eutrophication of coastal waters is often accompanied by a decrease in silica availability, due both to increased sedimentation of silica within estuaries and to trapping of silica in upstream, eutrophic fresh waters. Since silica is required by diatoms, but not by other types of phytoplankton, eutrophication often results in a relative loss of diatoms from the community. A fourfold increase in the relative inputs of nitrogen compared to silica along the German coast during the past few decades was accompanied by a 10-fold decrease in diatoms and a comparable increase in flagellates, which are less conducive to supporting food webs leading to commercially valuable fisheries. Decreased silica availability and the concomitant loss of diatoms may also be responsible, at least in part, for the increased frequency of occurrence and duration of blooms of harmful algae that seem to accompany coastal eutrophication. Harmful algal blooms ('red tides,' 'brown tides,' and 'green tides') can kill animals, such as sea lions, far up the food chain. The resulting loss of important top predators such as sea otters can cascade back down the food chain, altering species composition and reducing biodiversity.

Coastal systems that are both high in biodiversity and severely affected by eutrophication are seagrass beds and coral reefs. Temperate seagrass beds are usually nitrogen limited, whereas tropical seagrass beds and coral reefs are often phosphorus limited. However, even these tropical systems probably become nitrogen limited once eutrophication begins, and therefore additional inputs of nitrogen can cause immense harm. For the seagrasses, eutrophication leads to an immediate loss in diversity and can lead to a complete loss of the grasses due to shading by the increased biomass of phytoplankton in the water column, shading by mats of opportunistic species of macroalgae, or an accumulation of toxic decomposition products in the sediments. Eutrophic shallow estuaries and lagoons with large accumulations of macroalgae may experience frequent episodic oxygen depletion, further reducing biodiversity. Seagrasses provide food and shelter for a rich and diverse fauna, and reduced seagrass depth distribution or replacement by macroalgal blooms can result in marked declines in the abundance and biodiversity of the associated fauna.

Coral reefs occur in the oligotrophic shallow waters of the tropics, support extraordinary biodiversity, and are extremely

sensitive to damage from eutrophication. The biodiversity of coral reefs declines dramatically with eutrophication from nitrogen inputs, especially when species-poor turf algae or macroalgae communities overgrow and replace corals and coralline algae. Nutrient enrichment disrupts the coral–zooxanthellae symbiosis, inhibiting calcification and contributing to increased coral 'bleaching' (loss of zooxanthellae). Increased phytoplankton biomass and production increase turbidity and sedimentation. The decreased quality and quantity of light penetrating to the corals slow growth and reduce the maximum depth at which the corals can survive, reducing the available habitat. Algal blooms can cause hypoxia, decreasing the biodiversity of reef organisms through mortality and reduced habitat quality. Also, eutrophication can damage reefs indirectly by increasing predation on corals, facilitating the expansion of opportunistic filter feeders, and reducing the recruitment of corals.

Effects of Acidification by Nitrogen on Biodiversity

A variety of oxidized nitrogen compounds (NO_x) are released into the atmosphere in the exhaust of vehicles, electric power plants, and other fossil fuel combustion. Oxides of nitrogen are also released from soils and burning vegetation. In the presence of sunlight, NO_x catalyzes the formation of photochemical (or brown) smog and reacts with oxygen and hydrocarbons from automobile exhausts to form ozone, an air pollutant damaging to many plant species. In the atmosphere, NO_x reacts to form nitric acid, one of two major components of acid rain (the other is sulfuric acid).

At least 70% of global ammonia (NH_3) emissions are caused by humans, mostly by volatilization from fertilized fields, animal wastes, and forest burning. In the atmosphere, NH_3 neutralizes some of the acidity in aerosols, cloud water, and precipitation. However, once deposited on the earth's surface, ammonium (NH_4^+) is taken up by plants or converted to nitrite (NO_2^-) and then nitrate (NO_3^-) by bacteria. These processes of biological uptake and nitrification release hydrogen ions, acidifying the soil and downstream aquatic systems. Other processes, such as increased biological nitrogen fixation or increased rates of decomposition, which increase the amount of ammonium in the soil, also contribute to acidification.

As nitrogen inputs to ecosystems increase and the capacity of the ecosystems to retain nitrogen is lessened, ecosystems begin to release increasing amounts of NO_3^-. When the negatively charged ion leaves the ecosystem, it carries with it positively charged base cations, such as Ca^{2+}, Mg^{2+}, and K^+. These elements are essential plant nutrients and their removal can deplete soil fertility and cause serious nutrient imbalances which are detrimental to plants. As the other soil cations are depleted, toxic Al^{3+} is mobilized. Soils with low capacity to neutralize or buffer the increased acidity from nitrogen and acid deposition are especially sensitive to acidification. Increased acidification may be the major impact of increased nitrogen deposition in ecosystems which are not nitrogen limited, such as many tropical forests and most temperate lakes, or are already nitrogen saturated, such as some temperate forests.

Until recently, most of the impacts of humans on nitrogen dynamics were focused in the temperate zone, but industrial and agricultural practices are changing rapidly in the tropics, where most of the earth's biodiversity resides and for which our ecological knowledge is rudimentary. By 2020, approximately two-thirds of the global application of industrial nitrogen fertilizer and energy-related nitrogen inputs will occur in the tropics and subtropics. The majority of tropical forests are probably not nitrogen limited and nitrogen enrichment will likely have little direct effect on plant production but instead will substantially increase the output of dissolved and gaseous nitrogen from these ecosystems. The acid soils of the tropics may be particularly susceptible to further acidification, depletion of base cations, decreased availability of plant nutrients, and release of toxic Al^{3+}, all possibly reducing productivity and biodiversity.

The effects of acidification on sensitive terrestrial and aquatic ecosystems are clear and severe. Increasing acidity of thousands of lakes and streams and the loss of fish and amphibian populations have been documented in Scandinavia and eastern North America. Experimental acidification of an entire Canadian lake eliminated shrimp, minnow, and crayfish species, all important food sources for the lake trout, which subsequently ceased reproduction. Increased nitrogen deposition has been implicated in the dieback of poorly buffered high-elevation forests in Europe and damage to high-elevation red spruce and fir forests in the northeastern USA. Acid deposition causes direct damage to trees by leaching base cations from leaves. Exposure to acid fog and cloud water, usually much more acidic than rain, increases sensitivity to frost damage. Acidification of forest soils causes nutrient imbalances, cation loss, aluminum toxicity, reduced growth rates, and increased mortality in trees. The diversity of lichens is greatly reduced by acidic deposition and the species composition of vascular plants, soil animals, and microbes shifts toward a less diverse community of acid-tolerant species. Acidification of grasslands and heathlands is widespread in Europe and results in decreased plant diversity when acid-tolerant species replace the species typical of less acid soils.

Terrestrial and aquatic ecosystems subjected to increased inputs of nitrogen often experience both eutrophication and acidification. The combined effects on biodiversity may be independent, additive, or synergistic depending on the characteristics of the ecosystem, but interactions between eutrophication and acidification are common. For example, nitrogen saturation increases the loss of base cations from the soil, accelerating acidification, while acidification strips base cations from plants, increasing the severity of the nutrient imbalances caused by eutrophication.

Effects of Biodiversity on Nitrogen Dynamics

The global biodiversity crisis has prompted increased concern for the possible effects of extinction of native species and introduction of invasive species on ecosystem function and the provision of ecosystem services essential to humans. The effects of biodiversity on nitrogen dynamics have not been studied as thoroughly as the converse effects of nitrogen on biodiversity, but there are numerous examples of significant

effects of particular species or functional groups of species on ecosystem processes, including nitrogen dynamics. There are far fewer examples of the effects of species diversity (the number and relative abundance of species), independent of the species composition (specific identity of the species present), on ecosystem processes such as primary production and biogeochemistry.

At larger scales, the effects of ecosystem and landscape diversity on nitrogen dynamics vary with the specific ecosystems and their spatial relationships. Ecosystems which capture or accumulate nitrogen, such as riparian zones bordering farm fields or wetlands upstream of estuaries, can be very important in controlling the transport and downstream impact of nitrogen. Declines in biodiversity associated with the fragmentation and homogenization of landscapes, land clearing, and development may increase the magnitude and variability of exports of nitrogen from the disturbed or converted ecosystems.

Particular species sometimes play unique or dominant roles in ecosystems and the introduction or elimination of these key species can dramatically alter the structure and function of particular ecosystems. For example, the introduction of a nitrogen-fixing tree to a Hawaiian ecosystem which contained no other nitrogen-fixing trees rapidly increased the availability of nitrogen and led to other changes in species composition of the vegetation. Burrowing animals, such as pocket gophers or prairie dogs, increase the rate of mineralization of nitrogen and increase the heterogeneity and availability of nitrogen. In many grasslands, grazers accelerate the rate of nitrogen cycling and influence the distribution and availability of nitrogen as well as the species composition of the plant community. In the boreal forest, in contrast, selective browsing by moose on deciduous trees containing more readily mineralized nitrogen slows nitrogen cycling and speeds succession toward coniferous forest. Addition or removal of whole groups of species that are performing the same functional role ('functional group') has effects similar to removing or adding a single species which is performing a unique function (a functional group with one member). For example, functional groups including nitrogen fixers, leaf-chewing insects in streams, and early spring ephemeral plants of the forest understory all have significant effects on nitrogen dynamics. To date, most of the investigations of the effects of diversity on ecosystem processes have focused on the diversity within groups of primary producers, usually herbaceous plants, but the effects of diversity at other trophic or organizational levels (consumers, predators, microbes, soil invertebrates, or whole ecosystems) may differ from the effects on primary producers. The influence of species or functional group diversity on ecosystem processes may differ for each functional group.

Undoubtedly, the idiosyncratic traits of individual species can influence nitrogen dynamics, but species also share many similar characteristics, compete for some of the same limited resources, and carry out some of the same ecosystem functions. If species are very similar, then redundancy is high and the removal of a single species would have little impact on ecosystem function. Alternatively, if species are very different (idiosyncratic) and there is little niche overlap (rather, niche differentiation), then the removal of a particular species would have a greater impact on ecosystem function. As the number of species increases, the probability of including an extremely effective dominant species in the more diverse community increases. This simple 'sampling effect' is a result of changing species composition which occurs simultaneously with increasing species diversity.

Spatial and temporal variability in resource availability and environmental conditions are characteristic of most ecosystems; therefore, there are many ways in which species might differ in their use of resources and the environment. One species may be capable of symbiotic nitrogen fixation, whereas another can use organic nitrogen and another prefers NO_3^- or NH_4^+. One species may be better adapted to drier environments or another may flower earlier in the growing season. Theoretically, a greater number of species, with different traits, will make more complete use of the available resources, increasing productivity beyond the maximum of any less diverse mixture of species. Complementarity of species in the use of nitrogen as a limiting resource and facilitation of one species by another as it occurs when a nitrogen fixer increases the availability of nitrogen to other species can cause productivity of diverse mixtures of species to be greater than the productivity of any of the single species.

Species diversity is often correlated with other factors, such as productivity, biomass, predation, or resource availability, that complicate the assessment of the effects of species diversity separate from other factors. Only a few studies have examined the effect of species diversity, independent of species composition, on nitrogen cycling. These initial results suggest that greater species diversity among the primary producers more completely uses the available nitrogen and increases plant uptake of nitrogen. In a Minnesota grassland, the NO_3^- concentration in the soil of the rooting zone (**Figure 6**) and below the rooting zone decreased with increasing diversity, indicating that less NO_3^- was lost via leaching and more nitrogen was retained within the ecosystem. In these same Minnesota grasslands, the nitrogen concentration in plants and the total nitrogen content of the plant biomass increased with the number of functional groups as well, and the

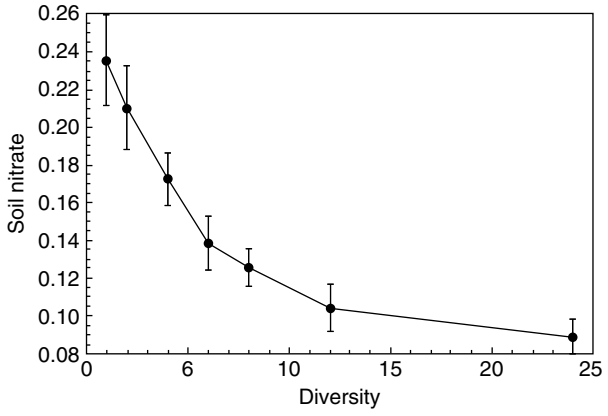


Figure 6 Effect of plant species diversity (number of species) in grassland plots on the concentration of nitrate in the soil (mg N/kg) containing most of the plant roots (0–20 cm). Reproduced from Tilman DG (1999) The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80: 1455–1474.

composition and diversity of the functional groups had approximately equal influences on ecosystem processes. In California serpentine grasslands, temporal partitioning of resources and facilitation by nitrogen fixers allowed more complete use of the greater available nitrogen, but the identity of the functional groups was more important than the number of functional groups. Hypothetically, increased diversity and resource exploitation increase nitrogen retention, thereby creating a positive feedback loop of increased fertility and productivity. With more complete capture or immobilization of a limiting resource such as nitrogen, more diverse communities may be less susceptible to invasion.

Variability of NO_3^- in the rooting zone decreases as the number of plant species increases, but this is partly due to a simple statistical artifact – a 'portfolio effect' in which the aggregate variance of several species (like the diversified portfolio) is less than the variance of individual species. More diverse communities may show less variability in ecosystem function under the varying conditions typical in the natural world. Thus, species diversity may provide 'insurance' that regardless of the varying environmental conditions, ecosystem functions will be carried out by some species suited to the prevailing conditions.

In general, it appears that the number of species or functional groups may influence nitrogen dynamics, but these are probably less influential than the particular identities of the species or functional groups, especially within the context of the many other factors which strongly influence nitrogen dynamics, including climate, geology, hydrology, type and history of disturbance, etc. Species and functional group diversity may be most important at the very low diversity characteristic of intensively managed and exploited ecosystems typical of modern agriculture or very degraded ecosystems. It appears that the upper limit of the number of plant species for species diversity, independent of species composition, to have an impact on nitrogen cycling may be quite low (1–12) in most ecosystems. Of course, there may be some ecosystems in which nitrogen cycling cannot be maintained by a few species – perhaps extremely variable and heterogeneous ecosystems such as semiarid grasslands.

Considerations for the Future

Unfortunately, there is no obvious reason to expect human impact on nitrogen cycling or other human impacts on biodiversity to decrease. Rather, we can expect modification of the nitrogen cycle and declining biodiversity to continue, within the context of all the other modifications which we humans are making to our environment, including climate change, air and water pollution, and overexploitation of natural resources. Many indirect effects of alterations of the nitrogen cycle on biodiversity are possible. For example, human activities have increased the release of N_2O, a greenhouse gas which contributes to climate change. Climate change has potentially major effects on biodiversity. In addition, N_2O which reaches the stratosphere participates in the destruction of the stratospheric ozone shield, allowing more ultraviolet light to penetrate to the earth's surface, potentially decreasing biodiversity. The variety of important ecological interactions which involve nitrogen, the complexity of controls on nitrogen dynamics, and the huge scale of human alterations to the biogeochemistry of nitrogen imply that many more interactions and feedbacks between nitrogen and biodiversity remain to be discovered as the domination of the environment by humans intensifies. The interactions between nitrogen and biodiversity are multiple and complex, but it is clear that humans cause both the changes in the biogeochemistry of nitrogen that usually reduce biodiversity and the changes in biodiversity that alter the biogeochemistry of nitrogen.

Appendix

List of Courses

1. Ecology
2. Biogeochemistry
3. Environmental Sciences
4. Diversity
5. Global Changes

See also: Atmospheric Gases. Bacterial Biodiversity. Biogeochemical Cycles. Carbon Cycle. Estuarine Ecosystems. Eutrophication and Oligotrophication. Soil Conservation

References

Dise NB, Ashmore M, Belyazid S, *et al.* (2011) In: Sutton MA, *et al.* (eds.) *The European Nitrogen Assessment: Sources, Effects, and Policy Perspectives*, pp. 463–494. Cambridge University Press.

Galloway JN, Dentener FJ, Capone DG, *et al.* (2004) Nitrogen cycles: Past, present, and future. *Biogeochemistry* 70: 153–226.

Hooper DU and Vitousek PM (1998) Effects of plant composition and diversity on nutrient cycling. *Ecological Monographs* 68: 121.

Huston MA (1997) Hidden treatments in ecological experiments: Re-evaluating the ecosystem function of biodiversity. *Oecologia* 110: 449–460.

Naeem S, Thompson LJ, Lawler SP, Lawton JH, and Woodfin RM (1994) Declining biodiversity can alter the performance of ecosystems. *Nature London* 368: 734.

National Research Council (1996) *Understanding Marine Biodiversity*. Washington, DC: National Academy Press.

National Research Council (2000) *Nutrient Over-Enrichment in Coastal Waters*. Washington, DC: National Academy Press.

Stevens CJ, Duprè C, Dorland E, *et al.* (2011) The impact of nitrogen deposition on acid grasslands in the Atlantic region of Europe. *Environmental Pollution* 159: 2243–2250.

Townsend A and Howarth RW (2010) Human acceleration of the global nitrogen cycle. *Scientific American* 302: 32–39.

Townsend AR, Martinelli LA, and Howarth RW (2009) The global nitrogen cycle, biodiversity, and human health. In: Sala OE, Meyerson LA, and Parmeson C (eds.) *Biodiversity Change and Human Health. SCOPE #69*, pp. 159–178. Washington, DC: Island Press.

Vitousek PM, Aber JD, Howarth RH, *et al.* (1997) Human alteration of the global nitrogen cycle: Source and consequences. *Ecological Applications* 7: 737–750.

Nomenclature, Systems of

David L Hawksworth, IUBS/IUMS International Committee on Bionomenclature, Ciudad Universitaria, Spain, and Natural History Museum, London, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Available Of a scientific name of an organism, one that must be taken into account when deciding the correct (q.v.) name to be used for it.
Binomen A binominal or binary name of a species, consisting of the name of the genus in which it is placed followed by a word peculiar to that species.
Code In nomenclature, one of the sets of internationally agreed rules governing the scientific names of organisms.
Conserved Of a scientific name, one that the appropriate international body has decided should continue to be used for an organism in cases where a strict application of the Code (q.v.) would mean it had to be replaced.
Correct Of a scientific name of an organism, one that conforms to the appropriate international Code (q.v.) for the position in which it is placed.
Priority Of a scientific name, its date of valid publication; the first published name is generally the one to be used, subject to the provisions of the pertinent Code (q.v.).
Synonym One of two or more scientific names applied to the same organism.
Taxon A taxonomic entity or group of any rank, including all subordinately ranked taxa placed within it.
Type Of an organism, the specimen, culture, or other element on which a scientific name of an organism is based and that fixes its application; the name-bearing type.
Valid Of a scientific name, one that conforms to the conditions of valid publication proscribed in the relevant Code (q.v.).

Introduction

Systems of nomenclature are the procedures developed to provide the scientific names of organisms. This article concentrates on the consensual systems that operate in different groups of organisms through a series of published, internationally adopted Codes. It also discusses why scientific names change, the procedures involved in the naming of newly discovered species, and provides hints for ascertaining the current name of an organism. The article is concerned only with the nomenclature of organisms, that is, bionomenclature, and not with systems developed for naming parts of the genome, parts of the organisms, or communities of plants or animals.

Purpose

The purpose of systems of nomenclature is to provide an unambiguous mechanism that enables biologists and all others who work with organisms to communicate by a scientific name so as to avoid misunderstandings and confusion. Scientific names are Latinized (though not necessarily based on Latin words) and given to organisms of all kinds, whether living or fossil, and are in effect a universal currency – the means of connecting together and accessing the accumulated knowledge of any particular organism. The same species may have different common or colloquial names in a variety of current languages, but can bear only a single correct scientific name.

Nomenclature is not the same as taxonomy, although the two are often confused. Taxonomy, a component of systematics (or biosystematics), is concerned with drawing up classifications, with deciding, for example, to what genus a species belongs, whether two species are really the same, or whether one species should be divided into two. Nomenclature is the process of determining what scientific names, in effect labels, should be applied to the units that taxonomic research considers to merit independent names. Nomenclature is subservient to taxonomy; it is not science itself, but the method by which scientific results are made available for general use.

The application of naming systems is complicated because new research may show that a particular taxonomy does not reflect the evolutionary (i.e., phylogenetic) relationships of the organisms concerned. Further, the same species may have been described more than once by different scientists, and placed in a number of genera according to different taxonomic opinions. The net result is that there are many more scientific names than known species in the scientific literature. The full extent of the problem is unclear, but in the case of non-fossil groups traditionally viewed as "botanical" (i.e., algae, fungi, and plants) there are approximately 1.7 million species names available for 406 000 accepted species. It is the challenge of nomenclatural systems to provide clear procedures of how to handle a morass of names that may have been applied over the centuries to the same organism.

Origins

From the earliest times, humans needed to communicate taxonomic information: to tell each other which animals or plants are a threat, which can be eaten, which make tools or dyes, or which can help cure particular ailments. Indigenous peoples today confront the same problems, and often have

sophisticated naming systems for organisms of importance to them; interestingly, the species concepts they employ are often not too divergent from those a taxonomist would adopt. The basic need to name organisms was recognized even in The Bible, where one of Adam's first tasks was to name the animals, and there are now indications that one area of the human brain is devoted to discriminating living organisms.

Problems arose when different peoples needed to communicate with each other, and even the use of Latin as a language of scholars failed to solve the problem. Descriptive phrases written in Latin came to be used, but these polynomials were cumbersome and difficult to remember. For instance, the tea plant was referred to as *Evonymus affinis arbor orientalis nucifera, flore roseo* by Leonard Plukenet in 1696. By the early eighteenth century the situation was becoming impossible, and it was Carl Linnaeus (1707–1778) in Sweden who developed a pragmatic solution that remains the basis of our current nomenclatural systems. Linnaeus at first followed several earlier authors in grouping his species into categories that had a single name, in effect a genus. However, in 1737 he started to use a single second name to denote the species in the indexes in his books. This shorthand practice entered the main texts of some of the works he was associated with in 1751, and he then went on to produce accounts of all organisms known throughout the world using this approach; the species names were printed in the margins offset from the main text. Instead of Plukenet's polynomial for the tea plant, Linnaeus simply used the binominal, or binary name, *Thea sinensis*. Such a radical change was not well received by all of his contemporaries, who regarded him as autocratic, but his *Species Plantarum* (1753) and *Systerna Naturae* (10th edition, 1758) were destined to become designated as the starting points of the modern nomenclature of most forms of life.

Hierarchical Systems

Nomenclatural systems are based on a series of ranks. Each species is referred to a genus, a genus to a family, a family to an order, and so on. The principal ranks available are, in descending order: domain, kingdom, phylum (or division), order, class, family, tribe, genus, section, species, and subspecies. Additional ranks can be intercalated using the prefixes super- and sub- (e.g., superfamily above family, subfamily below family). In practice, all the possible ranks are rarely used in any particular case.

Though the number of available ranks may seem large, the current system cannot adequately satisfy those who wish to recognize each branching point in a cladogram, and proposals for a modified rankless system has been made in the PhyloCode (*see* The PhyloCode), which is independent of the organismal Codes and does not at present deal with species names.

Some of the ranks above genus have standardized suffixes that enable them to be recognized (**Table 1**), but there are exceptions for some well-known names approved by the relevant international nomenclatural bodies; for example, the grass family *Gramineae* (not *Graminaceae*). There are no rank-indicating endings used above phylum in botany, none over order in bacteriology, nor any over superfamily in zoology – although traditions have in some cases developed.

Table 1 Hierarchical ranks and suffixes indicating them

Rank	*Bacteriology*	*Algae, fungi, and plants*	*Virology*	*Zoology*
Domain				
Kingdom				
Subkingdom		-mycetia (fungi)		
Phylum		-mycota (fungi) -phyta (algae)		
Subphylum		-mycotina (fungi) -phytina (algae)		
Class		-mycetes (fungi) -opsida (ferns) -phyceae (algae)		
Subclass		-idae (ferns) -mycetidae (fungi) -phycidae (algae)		
Order	-ales	-ales	-virales	
Suborder	-ineae	-ineae		
Superfamily				-oidea
Family	-aceae	-aceae	-viridae	-idea
Subfamily	-oideae	-oideae		-inae
Tribe	-eae	-eae		-ini
Subtribe	-inae	-inae		
Genus			-virus	
Subgenus				
Section				
Species				
Subspecies				
Variety				
Form				

The names of families are based on the name of an included genus, such as *Nostocaceae* on the genus *Nostoc*. Order names are also based on generic names in algae, fungi, and plants, but not necessarily so in animals, where the characters of the order tend to be employed. At ranks above order, there are no strict rules, but the names usually reflect some key distinguishing feature.

The only rank below species that is formally recognized in bacteriology and zoology is that of subspecies. In zoology the subspecies name follows that of the species without any indication of rank to form a trinomial, for example, *Mus musculus domesticus*, whereas under the other main Codes the rank is made clear by the insertion of subsp. or ssp., as in *Silene vulgaris* ssp. *maritima*. The insertion of a rank term in those cases is essential to make clear which is referred to: "var." denotes variety and "f." denotes a form. In botany, additional terms are used in particular groups, most notably "cultivar" ("cv.") for stable propagated cultivated plants and "special form" ("f. sp.") for pathogenic fungi attacking different host plants that cannot be separated morphologically. However, with the advent of molecular techniques the use of special form is becoming obsolete as species concepts become better understood.

Scientific names in the rank of genus and above always start with a capital letter. Generic, specific, and formal ranks below species (see **Table 1**) are always printed in italic script. The practice at other ranks varies, and is a matter of editorial

style as this is not ruled on in any of the existing Codes. Although in zoology it is not usual to italicize any other ranks, the scientific names of algae, fungi, and plants at all ranks are italicized in the most recent Code (*see* Codes of Nomenclature), and this practice is increasingly being adopted in botanical and mycological publications as well as being advocated in the interdisciplinary BioCode (*see* The Draft BioCode).

Principles

All Codes share a common objective: the scientific name of an organism must be unique and stable in a particular classification (i.e., a taxonomy). Nomenclature can be thought of as a tool used when it is necessary to decide on the name to apply to a particular organism or other hierarchical rank; it must not constrain the science of taxonomy. To achieve this objective, each name needs to be formally introduced by agreed upon procedures. There must also be a means of ensuring that the application of every name is fixed so that it can always be applied unambiguously. In cases where more than one name has been proposed for the same organism, the basis on which to make a choice must also be unambiguous so that different researchers will be led to apply the same name where they have identical taxonomies.

Names of genera and species can be formed from any source, but are treated as if they were Latin. Ideally they should be informative, giving some clue to their features, but they are often based on the names of their discoverers or the places where they were found. However, this is not always the case; they can even be derived from acronyms or be anagrams. Today, the etymology of newly coined names is generally expounded when they are first introduced.

A consequence of treating species names as if Latin is that their endings must agree with the gender of the genus in which they are placed. The ending of the name of the same species may thus change if it is transferred from a genus that has a female gender to one that is male.

The name of the person (or persons) who first introduced a scientific name is often indicated after the name itself, sometimes with the date of the publication. This device, referred to as an author citation, is essentially a much abbreviated reference to a particular publication, and although often interpreted as recognizing the contribution of that person, it can also be viewed as attributing blame to the person who burdened the scientific literature with it – especially as so many names prove to be unnecessarily coined for species that are later found to have previously been described. It is much easier and takes less time to describe an organism as new to science than to determine if it has already been named. If an author's name appears without brackets, it was introduced just like that. If the name is in brackets, the name was originally in another genus or rank.

The citation *Simulium albellum* Rubtsov tells us that Rubstov published that species name in the genus *Simulium*; *Obuchovia albella* (Rubstov) makes clear that Rubstov introduced the species name of this fly but did not refer it to the genus *Obuchovia*. In bacteriology and botany, the name of the person making the change in generic placement or rank follows any name placed in brackets. For example, the citation of the Fool's Watercress, *Apium nodiflorum* (L.) Lagasca, indicates that L. (Linnaeus) coined the name (actually as *Simulium nodiflorum* L.) but that Lagasca first placed it in the genus *Apium*. Names such as *Simulium albellum* and *Sium. nodiflorum*, which give the name being used in a different rank or combined in other genera, are termed "basionyms;" names using such basionyms in a different rank or position are "combinations."

The abbreviation of the names of authors, such as L. for Linnaeus in this case, is commonplace, especially in algae, fungi, and plants where an internationally recognized list of abbreviations to be used for particular authors was issued in 1992; initials are inserted where there is more than one author with the same surname. An ex will sometimes appear inserted between the names of two authors; this means that the one before it first used the name but did not meet some requirements of validity or availability (see the following), whereas the second author did. However, there are instances in zoology where the role of the authors as given in the citation is reversed, with the first being the one first correctly publishing the name.

The way in which names are published is also specified in the different Codes. These relate to the publication itself (effective publication), which must be available to the scientific community, and the matter that must be associated with a name when it is introduced (valid publication), for example, a description and details of the material on which the name was based (the name-bearing type, discussed next). Only if the specified criteria are met does a name formally enter the body of scientific literature and become available for use. From 1st January 2012, electronic-alone publication is considered effective for algae, fungi, and plants provided the journal has an ISSN number, and proposals to change this are under active consideration for inclusion in the next editions of the zoological Code. The reluctance to immediately accept electronic-only publication is due to taxonomists being concerned about the accessibility of electronic-only resources for centuries to come as technologies are evolving so rapidly.

The name-bearing type irrevocably now fixes the application of a name. This is usually a single specimen permanently preserved in a museum or other public institution, but in some cases it can also be a microscopic preparation, drawing, photograph, or microbial culture. The original material used and specified by the author is a holotype; a lectotype is material from the original author's collection selected to act as the name-bearing type by a later researcher; a neotype is a later collection chosen to serve as the name-bearing type where no original material remains. Types are not necessarily representative of the range of a species or subspecies, but are simply a nomenclatural device to unambiguously fix the application of names. They are consequently a very important reference point in biological communication and merit careful preservation and handling as they will remain of relevance for centuries.

A key feature in all Codes is the priority by date accorded to the names to be used for an organism. The name to be adopted is generally the first published in the same rank or group of ranks, depending on the Code.

Surprisingly, with the exception of bacteriology, mycology, and some groups of cultivated plants, there is no obligation

on a person introducing a new scientific name to register it with some international authority. Proposals to introduce such a system for botanical groups were accepted in principle in 1993 and were to be introduced for the year 2000, but the 1999 International Botanical Congress reversed that decision. Proposals to make this mandatory for fungi were, however, adopted at the Melbourne Congress in 2011 and are required from 1st January 2013. The names of fungi after that date must be registered to be considered validly published; registration in MycoBank has been tried and tested was already being enforced by the leading mycological journals and information on approximately 80% of fungal names proposed each year was already being deposited voluntarily.

Conscious that the application of a rigid system of rules could lead to unwelcome and confusing name changes, the Codes have developed a variety of mechanisms for overcoming this. These involve cases being published for open comment and consideration by international committees who debate them and express an opinion or recommend a particular course of action. Names can be conserved against ones that otherwise would have to be used, or rejected, in which event they cannot be taken up. In some instances, whole publications can be suppressed from usage.

An Approved List of names is published in bacteriology. Virus Taxonomy approaches this, and there is an Official Lists and Indices of Names and Works in Zoology. There is no exact equivalent for algae, fungi, and plants, although some of the appendices of protected and rejected names to that Code approach them, at least in part and approval for Lists to be developed for fungi was granted at the Melbourne Congress in 2011. Proposals to develop protected lists of botanical Names in Current Use were debated, but finally rejected by the International Botanical Congress in 1993. However, Adopted Lists of Names are a feature of the Draft BioCode (*see* The Draft BioCode), currently under discussion.

Names other than the one to be used that have been proposed for the same biological unit are termed "synonyms." These can be of two types: "homotypic" ("objective" or "nomenclatural" synonyms) and "hetero-typic" ("subjective" or "taxonomic"). Homotypic synonyms are based on the same name-bearing type, whereas heterotypic ones are based on different name-bearing types. Thus, homotypic synonyms are irrevocably linked, whereas heterotypic synonyms are dependent on the opinions of taxonomists on the placement of the different name-bearing types.

Codes of Nomenclature

The five current organism-related Codes differ in their detailed provisions. All are complex documents approved by particular internationally mandated bodies. Each has its own history of development and special characteristics. Some of their distinctive features, differences, and procedures are summarized here.

Algae, Fungi, and Plants

A single Code (McNeill *et al.*, 2012) covers algae, fungi, and plants; the organisms traditionally studied by botanists even though many are no longer classified in the plant kingdom (*Plantae*). It consequently embraces cyanobacteria (*Bacteria*), fungi (*Fungi*), slime-molds (*Protozoa*), and various algal groups (*Chromista, Heterokonta,* or *Straminipila*). The name of this Code was changed from the International Code of Botanical Nomenclature in July 2011 to reflect the disparate organisms it covers.

The nomenclature of these groups is considered as starting from the publication of Linnaeus' *Species Plantarum* in 1753, but the first internationally agreed Code dates from the *Lois de La Nomendatue Botanique* prepared by Alphonse de Candolle in 1867 and adopted by the International Botanical Congress in Paris that year. Later starting dates, all linked to particular major publications by authors other than Linnaeus are used for some groups, notably mosses (1801), certain groups of algae (1848–1900), and fossils (1820). Fungal names were formerly dated from 1801 or 1821, depending on the group, but that practice was discontinued in 1981. Names of fungi accepted in the previous starting point works by Christian Persoon (1801) and Elias Fries (1821–1832) remain sanctioned for continued use even if earlier competing names exist. Traditionally, the distribution of hard-copy publications, i.e. books or journals, was required for effective publication under the Code. However, this situation changed. From 1st January 2012, electronic publications are accepted as a means of effective publication, provided that certain criteria are met. The work must have an ISBN or ISSN number, as appropriate, and be in its final form and archived. Articles in journals that are made available online and in an unalterable form prior to distribution of printed copies are now considered effectively published and date from the date on which the online version was published.

In addition to the insertion of terms to denote ranks below species and differences in the practice of how describing authors are cited, several features of this Code are unique to it. New scientific names published after 1935 and before 1st January 2012 must, with a few special exceptions, have a description or diagnosis (i.e., a statement of how the organism differs from others) in Latin. However, on or after 1st January 2012, the validating diagnosis or description must be in either English or Latin.

This Code recognizes the priority of names only within the particular rank under consideration. This means, for example, that even if a plant was recognized as a subspecies or variety long before a species name was coined, the species name is nevertheless the one to be used. In addition, this Code rules as illegitimate names that have been introduced unnecessarily when another should have been adopted by the author, for instance, by placing an earlier validly published name as a synonym when it was introduced. Also illegitimate are names spelled in either exactly the same way or are so similar in spelling that they are likely to be confused; these names are termed "homonyms", and only the oldest is generally available for use. For example, *Erica hibernica* (Hook and Arnott) Syme 1866 is illegitimate and to be rejected because of the existence of *Eri. hibernica* Utinet 1839, which represents a different species and is based on a separate name-bearing type. That Syme's name was based on *Eri. mediterranea* var. *hibernica* Hook and Arnott 1835 does not affect the situation, as that name has priority from 1835 only in the rank of variety and not of species.

Name-bearing types have had to be designated when describing new species and infraspecific taxa in this Code since 1958, and from 1990 the institution where they are preserved must also be cited. Living type material is not permitted, but dating from 1993 freeze-dried (lyophilized) material or specimens of algae or fungi preserved in liquid nitrogen are acceptable provided they are maintained in a metabolically inactive state; resuscitated cultures from such material are then referred to as ex-holotype, ex-isotype, etc.

This Code has special provisions for hybrids, fungi and fossils, which are considered separately below (*see* Special Cases). Appendices listing conserved and rejected names, and suppressed publications form a part of the Code, but at the International Botanical Congress in Melbourne in 2011, it was decided that the appendices need not be printed in the Code itself but could be published separately. The provisions of this Code are debated at each 6-yearly International Botanical Congress, after which a new edition of the Code is issued. Any changes proposed have to be published, something generally done in the journal *Taxon*, are balloted on first by mail (by individual members of the International Association for Plant Taxonomy), and then at the Nomenclature Section meeting of the Congress itself, where a 60% majority is normally required to effect any change. The Congress establishes a series of permanent Committees that are charged with considering and making recommendations on proposals to reject or conserve names in different groups; those are also generally published in *Taxon* and are ratified by the subsequent Congress.

New proposals to be placed before the next Congress in 2011 include ones to permit electronic-only publication under particular circumstances, a move from mandatory Latin diagnoses (see above), the transfer of decision making in relation to matters concerning fungi to International Mycological Congresses (now organized every 4 years under the auspices of the International Mycological Association), and to require the prior deposit of key nomenclatural information in a recognized online repository (currently MycoBank) for the valid publication of fungal names.

Cultivated Plants

The cultivated plant Code (Brickell *et al*., 2009) split from the International Code of Botanical Nomenclature in 1952. It is primarily concerned with the naming of cultivated varieties or cultivars, which are propagated horticulturally by any method that retains their characteristics in a stable way. It is now complementary to the International Code of Nomenclature for algae, fungi, and plants, and does not compete with it, as species names and those in other formal recognized ranks of cultivated plants remain subject to the same rules as wild plants.

Cultivar names, with a few established exceptions, are not Latinized and since 1958 have been required to be in a modern language; they are placed within single quotation marks and never printed in italic type, for example, *Rubus nitidoides* ''Merton Early,'' in which ''Merton Early'' is the cultivar name and *R. nitidoides* is the botanical species name. The abbreviation cv. can be inserted before the cultivar name but is generally omitted. It is also not uncommon for the species epithet to be dropped as well in garden catalogs (e.g., Rubus ''Merton Early''). Similar cultivars can be grouped into cultivar-groups; this is achieved by placing a group name in brackets before the cultivar name and without quotation marks, for instance, Rosa (Hybrid Tea Group) ''Richmond''; if the particular cultivar is not being cited, the brackets are omitted (e.g., Rosa Hybrid Tea Group).

In the case of orchids, groupings of cultivars based on parentage are called grexes (or more correctly greges); their application follows separate provisions in The Handbook of Orchid Nomenclature and Registration (1993).

Graft-chimeras, in which living tissues of two species are made to grow together by grafting, can also be named under this Code. These are indicated by the use of a + sign in a manner equivalent to the use of × in the names of hybrids (*see* Hybrids). The form *Crataegus monogyna* + *Mespilus germanica*, for example, indicates a graft with tissues of both of those plants. A Latinized name can be used instead provided that the + sign is retained and the resultant name is not identical to a name of a hybrid; in this example, the name + *Crataegomespilus dardarii* has been coined. Cultivar names can be added after the name of the graft-chimera.

Names of cultivars have to be established through publication and not by garden labels or confidential trade lists. There must also be a reference point for their application, a standard dried specimen preserved in a herbarium, or in some cases an illustration. There is then a system of 72 International Cultivar Registration Authorities (ICRAs), each dealing with particular groups of cultivated plants, that register newly proposed names and make lists of these available. In addition, there are also national Statutory Plant Registration Authorities recognized by the International Union for the Protection of New Varieties of Plants (IPOV).

The cultivated plant Code operates under the auspices of the International Commission for the Nomenclature of Cultivated Plants.

Prokaryotes

The Bacteriological Code (Lapage *et al*., 1992) arose as an offshoot of the International Code of Botanical Nomenclature in 1939, in large part due to the need to accept living cultures as name-bearing types. In consequence, there are many similarities between the two Codes, although they have diverged in some respects. Although the printed version is labeled Bacteriology, it covers both *Archaea* and *Bacteria*, and so is being increasingly referred to as the Prokaryote Code under which name the next hard-copy version will appear. The most striking feature of this Code, and a response to the problem of the large numbers of bacterial names of uncertain application, a start date of 1980 was established and linked to an Approved List of Bacterial Names; names published prior to 1980 and not on the list do not exist for nomenclatural purposes, that is, they are not treated as validly published and so are no longer available for use. At a stroke, the number of species names that had to be considered by bacteriologists was reduced from approximately 32 000 to 3000.

New names of bacteria now have to be published in or be registered by the *International Journal of Systematic and*

Evolutionary Microbiology, and will then be added to the next edition of the Approved List; a continuously updated online system is available through the Internet.

The realization that many bacteria, and especially archaea, can be identified on the basis of molecular sequence data but not grown in culture is challenging the requirement for living cultures to be designated as name-bearing types. In addition, many cyanobacteria described under the botanical Code have name-bearing types that are dried specimens or microscopic preparations. Prokaryotic organisms that cannot yet be fully described and have designated type cultures can, however, be given a name but in a status termed "candidatus" – candidates for species.

Although the only rank below species to be formally recognized in the prokaryote Code is that of subspecies, "pathovar" ("pv.") is employed for plant pathogens essentially distinguished only by the ability to cause diseases in particular plants; for instance, *Xanthomonas campestris* pv. *durantae* (on *Duranta repens*) and pv. *erythrinae* (on *Erythrina indica*). Regulations regarding pathovar names and lists of those proposed are prepared by a committee of the International Society for Plant Pathology. Strain or serological type names or numbers are commonly used in bacteria of medical importance, but these have no standing under the Code.

The prokaryote Code operates under the aegis of the International Committee on Prokaryote Systematics (ICPS), which is a part of the International Union of Microbiological Societies (IUMS). The Committee has a Judicial Commission, which reports back to the International Committee, and has wide powers to propose changes to the Code, which do, however, have to be subsequently ratified by IUMS congresses.

Viruses

The naming of viruses was controversial for many years. Two different systems were operating in parallel: a traditional system recognizing families, genera, and species with Latinized names, and a divergent one employing the categories of group, subgroup, and type (or virus) and in a modern language. The former held sway among those working with viruses in animals, and the latter among plant pathologists. In the case of plant viruses, a scientific name indicative of the structure of the virus was generally preceded by an indication of the host and (or) the symptoms caused; for instance, the Desmodium mosaic potyvirus and the Desmodium yellow mottle potyvirus.

The situation started to change rapidly in 1991 when the International Committee on the Taxonomy of Viruses, which operates within the International Union of Microbiological Societies, decided to use the traditional approach with Latinized family and genus names, but to retain English names for species. New guidelines were drawn up and adopted by the International Congress of Virology in 1993, and the latest revision of these was published in 2011 (King *et al.*, 2011). Proposals to modify the rules relating to virus names may be aired in *Archives of Virology*, which also includes reports of meetings of the Committee and periodically a survey of all accepted virus taxa.

Virus names can be recognized by their endings even down to the rank of genus; generic names have the suffix *-virus* (see **Table 1**), whereas the species names remain in English. For example, the genus *Ipomovirus* has as its type species Sweet potato mild mottle virus, whereas that of the genus *Roseolovirus* is Human herpesvirus 6.

Zoological Organisms

Just as the International Code of Nomenclature for algae, fungi, and plants covers all groups traditionally studied by botanists, so the International Code of Zoological Nomenclature (International Commission on Zoological Nomenclature, 1999) embraces all that have been considered by zoologists. Although the two Codes both arose in the mid-nineteenth century, their provisions have diverged in some significant ways through subsequent modifications. The first zoological Code appeared in 1843, and revisions were considered at each International Congress of Zoology until these were discontinued after 1972. New editions are now prepared more irregularly. The starting point for zoological nomenclature is the 10th edition of Linnaeus' *Systerna Naturae* published in 1758; there are no exceptions for particular groups.

The zoological and algal, fungal, and plant Codes rule that they are independent. This means, for example, that the same generic name can be used for disparate organisms provided they belong to groups treated under different Codes; a classic case is *Drosophila*, a genus of mushrooms and also one of flies. The occurrence of such cross-Code homonyms can cause confusions in database searching, and the algal, fungal, and plant Code now recommends that taxonomists publishing new generic names should not use ones already existing in zoology.

The zoological Code is distinctive in operating the principle of coordination of ranks. Three series of names are recognized as coordinate: family-, genus-, and species-groups. Each group includes names of several ranks, for example, species includes superspecies, species, and subspecies; these are treated as equivalent for nomenclatural purposes, such as assigning priority of publication and author citations. No rules relating to names above the family-group level are included.

The description or diagnosis accompanying the publication of new scientific names in zoology can be in any language that uses words, although English, French, German, Italian, or Latin are recommended. Surprisingly, the designation of name-bearing types, although recommended, was not made mandatory until the 1999 edition of this Code. In contrast to the other Codes, a name-bearing type can represent the "work of an animal" rather than the organism itself, for example a bite mark left by an animal in a leaf. Further species names in which the generic name and specific epithet are identical, for example, *Troglodytes troglodytes*, termed tautonyms, are permitted in zoology but not in the other Codes.

There are also differences when a species is moved from one genus to another. In zoology, the earliest name takes precedence and is to be taken up in the new genus even if an identical name based on a different name-bearing type already exists in the genus; in this way, secondary homonyms can be created.

In order to minimize the disruption of names, this Code rules that names which have not been adopted after 1st

Table 2 Equivalent terms used in the current Codes of nomenclature (excluding the PhyloCode)

Draft BioCode	*Bacteriology*	*Algology, Botany, and Mycology*	*Cultivated plants*	*Zoology*
Publication and dates of names				
Published	Effectively published	Effectively Published	Published	Published
Registerable	Effectively published			
Date	Date	Date (or priority)	Date	Priority
Priority	Priority	Priority		Priority
Precedence	Priority	Priority	Precedence	Precedence
Earlier	Senior	Earlier	Earlier	Earlier
Later	Junior	Later	Later	Junior
Nomenclatural status				
Established	Validly published	Validly published	Established	Available
Registration	Validation		Registration	
Acceptable	Legitimate	Legitimate	Acceptable	Potentially valid
Taxonomic status				
Accepted	Correct	Correct	Accepted	Valid
Types of names				
Name-bearing type	Nomenclatural type	Nomenclatural type	Name-bearing type	Name-bearing type
Nominal taxon	Name and type	Name and type		Nominal taxon
Synonymy				
Homotypic	Objective	Nomenclatural	Homotypic	Objective
Heterotypic	Subjective	Taxonomic	Heterotypic	Subjective
Replacement name		Avowed substitute		Explicit replacement
Setting aside the rules				
Conserved	Conserved	Conserved	Conserved	Conserved
Rejected	Rejected	Rejected	Rejected	Conditionally suppressed
Suppressed	Rejected	Explicitly rejected		Suppressed

January 2000 cannot take precedence over a different name in current use; there are stringent guidelines as to how names qualify as ''forgotten''. Also in the interests of stability, a series of Official Lists are maintained, dealing with names in particular ranks and also particular works. Terminological differences between the different Codes are a particular source of confusion to biologists; the same term can have different meanings and the same concept different terms (**Table 2**). Proposals to harmonize terminology have been made but have not yet been adopted across the Codes.

The International Commission on Zoological Nomenclature has wide powers to set aside provisions of the Code in particular cases where to follow the rules would lead to changes in well-established names. Applications have to be published in the *Bulletin of Zoological Nomenclature*, and after a period to permit zoologists to publish Comments in the Bulletin, the Commission expresses and publishes an Opinion that is binding. After the long series of International Congresses of Zoology ended in 1973, the Commission has worked to prepare new editions of the Code largely by correspondence and discussions of proposals in the *Bulletin*. New editions are now subject to ratification at a General Assembly of the International Union of Biological Sciences, under whose aegis it operates. Proposals relating to electronic publication, and the requirement that new scientific names must be registered in ZooBank (an online system similar to MycoBank for fungi) in order to be available, are currently under active debate and likely to be implemented in the next edition of this Code.

Interdisciplinary Codes

The Draft BioCode

The need for a more unified approach to biological nomenclature has long been recognized, and this has been of particular concern to the International Union of Biological Science (IUBS). Concerted action to address the problem emerged from a symposium held during the International Congress of Systematic and Evolutionary Biology in 1985. Pressure for harmonization between the various Codes arises from the unification of biology as a discipline and the consequent difficulties in teaching nomenclature, the mismatch between the Codes and the kingdoms of life now recognized, the needs of users for increased stability in names, the needs of systematists to spend less time on nomenclature, and the difficulties faced by developing nations. In addition, the existing Codes were finding that they had to confront similar issues, for example, in relation to electronic publication, living name-bearing types, organisms that could be treated under different Codes depending on whether they were interpreted as animals or plants, protected lists of names, and the registration of newly proposed names.

The move toward harmonization is operating along three fronts: harmonization of terminology (see **Table 2**), bringing similar provisions into all Codes where new problems are being confronted, and developing a unified Code that would apply to all kinds of organisms. In 1995 the IUBS and the International Union of Microbiological Societies, the two bodies that

between them oversee all five Codes, jointly established an International Committee on Bionomenclature to promote these initiatives. This Committee, which consists of representatives of the bodies responsible for the five organismal Codes, issued a Draft BioCode: The Prospective International Rules for the Scientific Names of Organisms for discussion in 1996, and a revised version was then prepared in 1997. There was little progress toward harmonization made at that time, as the mechanisms for the production of lists of available names and registration were not then available. However, that position has changed as it became increasingly evident that the different Codes had similar issues to contend with, and that there were increasing numbers of cases where there were problems of deciding if particular groups of organisms were a part of botany or of zoology – when they actually belonged to neither the plant nor the animal kingdom. Consequently, in 2009 the International Union of Biological Sciences initiated a program to revise the Draft BioCode (1997).

A revised version of this Code, the Draft BioCode (2011), prepared by a working group of the International Committee on Bionomenclature, and taking note of new developments in technology, was released in December 2010 (Greuter *et al.*, 2011), and was discussed at the International Congress of Systematic and Evolutionary Biology in Berlin in February 2011. That Congress commended the Code for further revision following feedback from individuals, and a new version is anticipated in 2012 for presentation to the upcoming IUBS General Assembly. The BioCode does not aim to replace the current series of Codes, but to provide them with an overarching framework. Some changes in the special Codes will be needed before it can take effect – especially the requirement to register new names and nomenclatural acts, the restriction of diagnoses to English or Latin, and the ability to protect names and their usages by Adopted Lists. These changes to promote common practices and rules are currently in discussion or have been implemented by the different responsible bodies. This Code, in addition to other matters, also seeks to prohibit the same names being available for use for different organisms under the separate Codes, and adopts the zoological Code principle of coordinated status of names within rank groups. It is anticipated that, during the next few years, a finalized version of the *BioCode* will, subject to approval of the appropriately mandated bodies, be proposed for adoption by the International Congress of Systematic and Evolutionary Biology, the International Union of Biological Sciences, and the International Union of Microbiological Societies.

A by-product of the endeavors of the International Committee on Bionomenclature to harmonize terminologies between the different Codes has been the compilation of a glossary of terms. The first draft of 1994 included 1175 terms, which swelled to over 2100 in an extended version, "Terms used in Bionomenclature: The naming of organisms (and plant communities)" (Hawksworth, 2010; also available as a free PDF to download from or interrogate on the website of the Global Biodiversity Information Facility).

The PhyloCode

The advent of molecular systematics has accelerated the elucidation of phylogenetic relationships of organisms, and some systematists have felt that there is a need for a separate system for the naming of evolutionary lines (i.e., clades) that would allow the use of unlimited numbers of hierarchical categories – and enable the same names to be retained for a particular clade when its rank was changed. If formal scientific hierarchical names were to be given to equivalent branching points in phylogenetic trees, it has been argued that considerable numbers of new ranks would be needed and there would be an exponential expansion in the numbers of such names. This problem is addressed by the PhyloCode (Cantino and De Queiroz, 2010), the International Code of Phylogenetic Nomenclature, the first draft of which was prepared in 2000. This is seen as a complement to the existing Codes, differing fundamentally in being based on the recognition and naming of groups of organisms circumscribed by descent, rather than the linking of names to name-bearing types. At present, the PhyloCode does not make provisions for the naming of species, although different ways in which that might be handled have been proposed.

The content of a taxon is therefore delimited by a definition based on ancestry and uses "specifiers" (e.g., species, specimens, and shared derived characters) to indicate actual organisms. The phylogenetically defined taxon is an ancestor and all of its descendants. The writers of the PhyloCode used the Draft BioCode (1997) in its preparation, and consequently its organization, some of its terminology, and the wording of certain rules are derived from the BioCode. It also envisages a system for the registration of new clade names.

The PhyloCode has occasioned much debate and acrimony among taxonomists, most of whom feel that the hierarchical systems of the other Codes are sufficiently flexible to accommodate the results of phylogenetic studies. The PhyloCode has not been adopted by most taxonomists, including those concerned with elucidating phylogenies, but it may have applications in phylogenetic discussions in cases where the clades are particularly robust. In many groups of organisms, however, the current sampling is so incomplete that these are not well supported and are subject to change as new data sets are included.

Special Cases

Particular provisions or problems in the Codes relate to the special cases of ambiregnal organisms, fossils, hybrids, lichens, and pleomorphic fungi.

Ambiregnal Organisms

Ambiregnal organisms are ones that can be treated under different Codes depending on whether they are considered to be plants or animals, algae or bacteria, etc. The problem is particularly acute in the case of unicellular and often flagellate organisms now placed in the kingdom Protista (or Protoctista or Protozoa), and especially where related species may or may not have chloroplasts and so have traditionally been treated as plants or animals, respectively. Because different rules apply when a plant is found to be an animal, name changes in both the generic and species names may result. For example, the

euglenoid name *Entosiphon* B. Stein 1878 is acceptable under the zoological Code, but when the organism is treated as a plant it has to be replaced by *Entosiphonomonas* Larsen & Patterson 1991, as there is an earlier flowering plant genus *Entosiphon* Beddome 1864.

The resultant frustrations and confusions have led to demands for a separate Code for protoctistan groups, but the preferred route forward is to have parallel provisions in the existing Codes. The problem would not arise under the Draft BioCode as no decisions as to whether an organism is a plant or an animal would be needed as Adopted Lists of Names could be proposed for protection for such organisms.

In the case of slime molds (*Myxomycota*), because these organisms have been traditionally studied by mycologists, they are treated under the International Code of Nomenclature for algae, fungi, and plants; that practice is continued to prevent the disruption in names that a switch to the zoological Code would entail. The *Cyanobacteria*, traditionally treated under what is now the International Code of Nomenclature for algae, fungi, and plants as blue-green algae, are pragmatically best left there rather than transferred to the prokaryote Code in order to avoid unfortunate and unnecessary name changes – although that issue is still contentious and to be addressed by an inter-Code committee.

Fossils

Special provisions for fossils are made only under the International Code of Nomenclature for algae, fungi, and plants. Pieces of trunks, stems, roots, and leaves, and seeds, pollen, and spores can all be given Latinized binomials. These have been considered to be form- or organ- genera and species in the absence of intact plants. Even when such pieces are later found connected together, as in the case of the *Archaeopteris* leaves and *Callixylon* wood, in the past the two systems were often maintained because of the time spans involved. For instance, it is not known whether a *Callixylon* leaf always had an *Archaeopteris* leaf system. However, this provision was deleted from the latest version of the Code (McNeill *et al.*, 2012) so that names based on fossil plant parts are subject to the same rules as other names.

Where names of living and fossil algae, fungi, or plants compete, those of the living one always have precedence. The names of fossils thought to be close but not necessarily identical with a modern genus are often given the suffix *-ites* as in *Ginkgoites* for fossil plants and *Ginkgo* for the single species now living.

Hybrids

Only the International Code of Nomenclature for algae, fungi, and plants has special arrangements for the naming of hybrids. The zoological Code excludes them from consideration.

Under that Code, hybrids can be referred to by indicating their parentage and the use of the multiplication symbol × to make a hybrid formula. Alternatively, and particularly if the hybrid is regularly encountered in nature, it can be given a separate scientific name prefixed by the same symbol. For example, the hybrid between *Potentilla anglica* and *Potentilla erecta* can be indicated as *Po.. anglica* × *Po. erecta* or alternatively as *Potentilla* ×*suberecta*. The same principle applies to generic names, for instance, ×*Agroelymus* is a hybrid between species belonging to the genera *Agropyron* and *Elymus*; the name ×*Agroelymus* is called a condensed formula.

In order to be established, the separate names of hybrids have to conform to other parts of the Code that apply to their rank. In the case of condensed formulae for generic names, the parent genera have to be indicated.

Lichens

Although generally appearing as if they were discrete organisms, lichens are mutualistic associations composed of two, or sometimes more, different kinds of organisms. A fungal partner combines with a green alga or cyanobacterium, or both, to form characteristic shapes for the particular association. The different partners, when isolated into pure culture, almost invariably fail to produce the same morphological structures. The classification of both the fungi and algae or cyanobacteria involved is based on their own characteristics and not those of the combined association; both have independent scientific names.

Before the dual nature of lichens was recognized in 1867, and for many years afterward, names were given to the composite lichen structure. The International Code of Nomenclature for algae, fungi, and plants rules that, for nomenclatural purposes, the names given to lichens are to be considered as belonging to the fungal partner alone – and the photosynthetic partner(s) can have separate scientific names. Pedantically, this means that the lichen associations do not have names, only their separate components – so it is impossible to name a lichen.

Fungi

Fungi that have a life cycle with separate sporulating stages, the pleomorphic fungi, have presented a particular problem in naming. In many cases the different stages were first named separately in different genera and were never thought to be connected until fresh research was carried out, for example, by germinating single sexual spores. In many cases the names given to the sexual (teleomorph) and asexual (anamorph) stage have both become well established, and mycologists have been reluctant to simply apply the earliest name regardless of the stage it represents. This resistance is understandable, as it is often a particular sporulating form that is associated with a plant disease or the production of toxins. However, at the International Botanical Congress in Melbourne in 2011, the system of allowing separate names for different morphs of the same fungus species was ended.

Prior to the Melbourne Congress, the Code had ruled that in basidiomycete and ascomycete fungi, with the exception of lichen-forming species, the Code has ruled that the correct name for a fungus with a pleomorphic life cycle that covers all its stages (the holomorph) was that which produces sexual spores, that is, the name of the teleomorph. The anamorph (or anamorphs) were permitted to have independent scientific names that could be used when only those stages are being referred to; the name-bearing types of the names of anamorphs had to represent only the asexual (mitosporic) sporing stage, and the sexual stage must not be mentioned in its establishing

description. For example, when the name *Penicillium brefeldianum* B. O. Dodge 1933 was introduced, the description and type covered both the teleomorph and anamorph; that name is therefore applicable to the teleomorph, even though the generic name *Penicillium* is based on an asexual species. In this case, the teleomorph belongs to the sexual genus *Eupenicillium* and is correctly named *Eupenicillium brefeldianum* (B. O. Dodge) Stolk & Scott 1967. The anamorphic name *Penicillium dodgei* Pitt 1980, with only asexual structures mentioned in its description and represented on the type, was introduced for those wishing to refer just to the asexual *Penicillium* stage. The main reason this system was ended was that molecular techniques now enable the phylogenetic position of asexual fungi to be determined with confidence in overall fungal classifications, in the absence of a sexual stage. The removal of this provision means that the names given to different morphs of the same species now compete on an equal footing. In the example given above, this means that the name *Penicillium dodgei* now has to be treated as a younger synonym of *Eupenicillium brefeldianum*. In order to minimize the potential disruption of a system that had been operating since 1912, the Code instructs mycologists not to take up asexually typified names for species where the sexual names are widely used without reference to the General Committee on Nomenclature (via the Nomenclature Committee for Fungi). Further, the option of developing Lists of approved or rejected names which would receive special protected status and be treated as appendices to the Code was provided. This provision is not restricted to names of fungi with pleomorphic life-cycles, but does exclude lichen-forming fungi.

A further major breakthrough was made at the Melbourne congress in that after 1st January 2013, all newly proposed scientific names of fungi at all ranks have to be deposited in a recognized online repository, along with information required for valid publication (including a description or diagnosis and details of the name-bearing type). The accession number in the database must be cited in order for the name to be validly published after that date. It is the responsibility of the Nomenclature Committee to recognize one or more repositories, and an announcement on this matter is expected during 2012. The MycoBank repository of the International Mycological Association will almost certainly be authorized to carry out this role, but perhaps in conjunction with other indexing centres.

A synopsis of these and some other changes relating to the regulation of fungal names is provided elsewhere (Hawksworth, 2011).

Other Systems

This article focuses on the nomenclatural systems used in and operating under the mandate of internationally mandated committees established by the scientific community. However, in addition to infraspecifc systems such as those for special forms of fungi and pathovars of bacteria that are specifically excluded from these Codes, some alternative nomenclatural systems have been proposed, although none has found general acceptance.

A system proposed by Kinman KE in 1994 aimed to take the best from cladistic analyses and include them in a practical system, with markers used to indicate where orders or higher ranks included organisms belonging to different evolutionary lines (i.e., are paraphyletic) while retaining established names.

More radical is the New Biological Nomenclature, which arose in 1971 out of the frustration felt by Belgian zoologists with the current system of Codes. Names being introduced have to be agreed on by a group of specialists and are presented in Esperanto and not Latinized. Generic names under this system relate more to common perceptions than those of scientists and can embrace several traditionally accepted genera. For example, the New Biological Nomenclature Esperanto generic name *Delfeno* embraces 17 genera of cetaceans; *Delfeno diverskolora* is used for the common dolphin (i.e., *Delphinus delphis*) and *Delfeno ordotipa* for the bottle-nosed dolphin (i.e., *Tursiops truncatus*). The system has not gained general acceptance, but does have an Association devoted to its promotion; detailed rules have been issued, the latest edition in 1991, and 522 names had been officialized by the end of 1990.

A novel approach to the nomenclature of fossils of all groups, a Paleontological Data Handling Code, was proposed by NF Hughes in 1989. Instead of first matching a newly discovered fossil to what has already been described, each specimen is recorded separately with particular reference to its stratigraphic position; only then are comparisons made with what has already been described in that particular position in the fossil record.

Various systems of coding using letters or numbers rather than Latinized names have been proposed, the first by P Hartig in 1871. A particularly elaborate system was that published by G Tornier in 1898, which took the first letters of words of the higher ranks to which a species belongs, followed by a number indicating the particular species; where more than one name started with the same first letter, subscripts were employed. Numerical codes regained popularity in the 1960s in response to the need to compress data in the computing systems in use at the time; one such code extended to 36 digits. With the advances in current technology, the interest in alphanumeric systems has now waned. Humans find it easier to remember words than numbers.

Although concerned with attributing names to assemblages of organisms rather than individual species, attention should also be drawn to the International Code of Phytosociological Nomenclature (Weber *et al.*, 2000), produced under the auspices of the International Association for Vegetation Science. This Code, the first version of which appeared in 1976, governs the allocation of names to plant communities. The names used in this system could be confused with scientific names of organisms, and the rules that control them are based on similar principles to the International Code of Nomenclature for algae, fungi, and plants. The system is hierarchical, and that the names are phytosociological and not organismal can be recognized by their suffixes, which in descending order are *-etum*, *-ion*, *-etalia*, *-etea*, *-etosum*, *-enion*, *-enalia*, and *-enea*.

Why Names Change

Changing scientific names are a constant cause of irritation and frustration to all who use them. That this occurs at all is sometimes interpreted as a criticism of the discipline of

taxonomy, and while that may be justified in some instances, in most cases it is in the long-term interests of users.

New research can show that a certain species is actually a member of a different genus than the one in which it was placed, and it is therefore transferred. Because classifications are predictive of the properties of an organism, accurate placement will benefit users searching for species with particular attributes or wishing to determine the risks they may pose. Detailed studies may also show that one species is a complex of several that merit independent recognition, or that two species currently treated as distinct are really the same; again, these are categories of information that are relevant to users.

Name changes that arise from new taxonomic research, or taxonomic changes, are therefore to be welcomed, but that is not so for nomenclatural changes. Nomenclatural changes, in contrast, are ones that arise from the application of the rules in the Codes, and not from any new scientific data. Two categories of nomenclatural changes can be distinguished. First, some changes are due to a failure to follow the rules of the Codes, for example, the discovery of an earlier name for the same species, that a name did not meet all the requirements for establishment, or, in the case of algae, fungi, and plants, that a name was superfluous when published. Second, some changes arise from rules of the Code itself that are retroactive, for example, in relation to starting point dates or new provisions relating to issues ranging from recommendations on spelling to complete revisions of particular sections.

Problems leading to name changes also arise from the application of the system of name-bearing types. Species names are sometimes found to have come to be applied to a species different from that represented by its type. This category of changes is especially unfortunate as it can mean that a name that has been used in one sense has to be applied in a different one.

The responsibility for deciding when enough data are at hand to justify a change in taxonomy will always lie with the taxonomist. Provided that the taxonomist is self-critical in approach and conscious of the need to be conservative in the interests of the image of the subject, most taxonomic changes will prove to be in the users' interests. That is not the case with nomenclatural changes where all Codes, and particularly the Draft BioCode (2011), are increasingly addressing this fundamental issue. The powers now given to the nomenclature committees and commissions mean that, provided a sound case is made, rules may be set aside to protect familiar names; this can even extend to the designation of different name-bearing types in order to maintain the current usage of names. Unfortunately, not all taxonomists take the route of making formal proposals to Committees to safeguard names, some apparently even taking delight in finding a reason to change a name. This practice is not, however, one condoned by the systematics community as a whole.

Naming Newly Discovered Species

As only about 1.8 million of the estimated 13.7 million species on Earth have yet been given scientific names, the task of naming the balance is one of the major challenges facing biology in the twenty-first century. However, it is always easier to describe an organism as new to science than to establish if it really has never been described before, albeit in a different genus or even in another family. The repeated unnecessary description of species that are not new is a major cause of the inflation of taxonomic literature. In the case of the fungi, one analysis showed that each accepted species had been on an average described 2.5 times. Once introduced into the body of scientific knowledge, names can never be expunged and will have to be taken note of by all future workers on the group concerned. Before deciding to introduce a new name, careful checking is a prerequisite.

This checking proceeds in a series of steps. First, see if the prospective species matches any in the same (or allied) genera in any checklists for the geographical area from which it comes. Then proceed to any regional or world monographs of the group if they exist. If these steps yield no matches, check names in the world indices of names (*see* Determining Current Names), investigating possible names in allied genera and under names no longer in use. This will entail making lists of possible names and their places of publication, looking up original descriptions, and locating and making comparisons with name-bearing types for likely candidates. If a specialist in the group concerned can be located, that can be a most valuable shortcut as the checking process can be exorbitantly time-consuming if dealing with, for example, a genus of over 1000 described species.

When novelty has been confirmed in the checking phase, the procedures detailed in the appropriate Code for introducing (and in some cases registering) new names must be followed. These include clear statements of why the species is different, a full description (which may have to be in Latin or English), illustrations (photographs as well as drawings are recommended), and a name-bearing type deposited in a secure public institution (ideally with duplicates deposited elsewhere for security). In some groups, molecular sequence data are now being routinely sought to justify the recognition of new species. Publication should ideally be in a journal that will be readily accessible to other specialists in that group and geographical region, although this is becoming less of a problem with electronic publications being made available on the web. Much can be learned from scrutinizing recent publications of new species in the same group by other authors when preparing accounts for publication.

In some instances, notably certain bacteria, internationally developed minimum standards for descriptions are available and should be followed. In some cases codes of practice have been developed, for example, there is one for the description of new fungi developed by the International Commission on the Taxonomy of Fungi.

Determining Current Names

When taxonomic work is undertaken, a number of different names are often found to have been given to the same species. The steps to be followed in determining which is to be used can be considered as a series of sieves, a nomenclatural filter. Although there are differences in the detailed requirements at

each step, and some will depend on the rank, the sequence in which these need to be followed is the same.

Step 1: *Publication* If a name has not met the requirements for effective publication, either from being printed in a book or journal available to other researchers, or published electronically after 1st January 2012 for algae, fungi, and plants, it need not be considered further.

Step 2: *Establishment* All of the following requirements must be met: species description (and in some cases illustrations), clarity concerning the rank (or coordinate group of ranks in zoology) and the intention to introduce a new scientific name, registration where appropriate (i.e., for prokaryotes, fungi, and some cultivated plants), citation of the name-bearing type, and in some cases its place of preservation. If any requirement is not fulfilled, those names also need not be considered further.

Step 3: *Typification* The most critical step and the one that requires scientific judgment is checking the status of the type and examining the specimen, microscopic preparation, illustration, and living culture as appropriate. In some cases, no name-bearing type may have been designated and it may be necessary to select a lectotype or neotype to fix the application of the name. Names whose types are different from the taxon to be named can be excluded at this step.

Step 4: *Acceptability* Does the name contravene any special provisions, such as being superfluous under the International Code of Nomenclature for algae, fungi, and plants? Is the spelling identical to another name of the same rank (or coordinate group of ranks in zoology)?

Step 5: *Precedence* What is the date of establishment of each name and which is the earliest in the same rank (in algology, botany, and mycology) or family-, genus-, or species-group (zoology)?

Step 6: *Accepted Name* The name to be used under the Code will be the one with precedence, that is, the earliest acceptable name to be established in the same rank or group of ranks. This name may have to be used in a new combination. Under the International Code of Nomenclature for algae, fungi, and plants, a new combination cannot be made if the resultant binomial would be spelled exactly like an already existing name (i.e., a homonym). The situation in zoology differs in that a new homonym can be made if the basionym takes precedence by date; the already existing homonym becomes a secondary homonym and can no longer be used. Names applied to the same suprageneric rank or organisms other than the accepted one are synonyms. Synonyms are of two main types: homotypic when based on the same basionym or name-bearing type, and heterotypic when based on different basionyms or name-bearing types. Homotypic synonyms will always appear together, as their applications are irrevocably linked by a common type, whereas heterotypic synonyms may be treated differently depending on the disparate views of taxonomists on the particular types. If the steps potentially result in the disruption of a well-known name, the procedures for the conservation and rejection of names are available as a safeguard (*see* Principles).

This daunting series of steps, once performed, leads to an accepted name. The advantage of the internationally agreed on Codes is that whoever follows the steps faithfully will come to the same answer regarding the name to be used in a particular taxonomy. That name will remain stable unless new nomenclatural or taxonomic information comes to light (*see* Why Names Change) and the taxonomy is changed.

Faced with such complexity, users of names wishing to check the currently accepted name for an organism need shortcuts. Recent world monographs and checklists are the first port of call, and then any regional or national treatments. If these are not available, and there is no taxonomist specializing in the group that can be consulted, it will be necessary to check the main indices of names and supporting bibliographic databases or abstract publications (e.g., Biological Abstracts, Kew Record, or Bibliography of Systematic Mycology). The possibilities for the online checking of names is improving year on year (*see* Indices of Names). However, that will not obviate the need for checking, which should be viewed as a necessary evil in all biodiversity inventory work and in biological papers in general. If research is published using an unfamiliar obsolete name, it may never be retrieved from the hundreds of thousands of scientific articles that appear each year, or linked to other pertinent information.

Indices of Names

There is a series of major online and printed reference works of the scientific names of organisms that focus on different major groups and that vary with respect to the information captured. There are too many to list here, but especially important are: AlgaeBase, FishBase, Index Kewensis, Index Fungorum, Integrated Taxonomic Information System, International Plant Names Index, Zoological Record, and the Approved Lists of Bacterial Names.

Major progress has been made toward a one-stop shop for information on names of all kinds of organisms in the *Catalogue of Life*, a collaborative project that brings together data on names from 77 global species databases: the 2010 annual checklist includes data on the names of a staggering 1.25 million currently accepted species.

See also: Species, Concepts of. Subspecies, Semispecies, Superspecies. Systematics, Overview. Taxonomy, Methods of

References

Brickell CD, Alexander C, David JC, *et al.* (eds.) (2009) *International Code of Nomenclature for Cultivated Plants*, 8th edn. Scripta Horticulturae No. 10 and Regnum Vegetabile No. 151. Leuven: International Society for Horticultural Science. http://www.ishs.org/sci/icracpco.htm

Cantino PD and De Queiroz K (2010) *PhyloCode: International Code of Phylogenetic Nomenclature, Version 4c.* http://www.ohio.edu/phylocode/index.html or http://phylocode.org

Greuter W, Garrity G, Hawksworth DL, *et al.* (2011) Draft BioCode (2011): Principles and rules regulating the naming of organisms. *Bionomina* 1 (2011), 26–44. *Taxon* 60 (2011), 201–212. *Bulletin of Zoological Nomenclature* 68: 10–28. http://www.bionomenclature.net/biocode2011.html

Hawksworth DL (ed.) (1991) *Improving the Stability of Names: Needs and Options*, Regnum Vegetabile No. 123. Königstein: Koeltz Scientific Books.

Hawksworth DL (ed.) (1997) *The New Bionomenclature: The BioCode Debate*, Biology International Special Issue No. 34. Paris: International Union of Biological Sciences.

Hawksworth DL (2010) *Terms used in Bionomenclature: The Naming of Organisms (and Plant Communities)*. Copenhagen: Global Biodiversity Information Facility. http://www.gbif.org/communications/resources/print-and-online-resources/bionomenclature/

Hawksworth DL (2011) A new dawn for the naming of fungi: Impacts of decisions made in Melbourne in July 2011 on the future publication and regulation of fungal names. *MycoKeys* 1: 7–20; *IMA Fungus* 2: 155–162.

International Commission on Zoological Nomenclature (1999) *International Code of Zoological Nomenclature*, 4th edn. London: International Trust for Zoological Nomenclature. http://iczn.org

Jeffrey C (1989) *Biological Nomenclature*, 3rd edn. London: Edward Arnold.

King AMQ, Lefkowitz EJ, Adams MJ, and Carstens EB (eds.) (2011) *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. Amsterdam: Elsevier Academic Press. http://www.ictvonline.org/codeOfVirusClassification_2002.asp?bhcp = 1

Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, and Clark WA (eds.) (1992) *International Code of Nomenclature of Bacteria (1990 Revision)*. Washington DC: American Society for Microbiology. http://www.ncbi.nlm.nih.gov/books/NBK8817/

McNeill J, Barrie FR, Buck WR, *et al.* (eds.) (2011) *International Code of Nomenclature for algae, fungi, and plants (Melbourne Code) Adopted by the Eighteenth International Botanical Congress, Melbourne, Australia, July 2011*. Regnum Vegetabile (in press). Ruggell: ARG. Gantner Verlag. http://ibot.sav.sk/icbn/main.htm

Melville RV (1995) *Towards Stability in the Names of Animals*. London: International Trust for Zoological Nomenclature.

Minelli A (1993) *Biological Systematics: The State of the Art*. London: Chapman & Hall.

Nicolson DH (1991) A history of botanical nomenclature. *Annals of Missouri Botanical Garden* 78: 33–56.

Polaszek A (ed.) (2010) *Systema Naturae 250: The Linnean Ark*. Boca Raton: CRC Press.

Ride WDL and Younés T (eds.) (1986) *Biological Nomenclature Today*, IUBS Monograph No. 2. Oxford, UK: IRL Press.

Stafleu FA (1971) *Linnaeus and the Linneans A*. Utrecht, The Netherlands: Oosthoek.

Weber HE, Moravec J, and Theurillat J-P (2000) International Code of Phytosociological Nomenclature, 3rd edition. *Journal of Vegetation Science* 11: 739–768. http://www.iavs.org/pdf/Code.pdf

Winston JE (1999) *Describing Species: Practical Taxonomic Procedure for Biologists*. New York: Columbia University Press.

Yoon CK (2009) *Naming Nature: The Clash between Instinct and Science*. New York: W Norton.

North America, Patterns of Biodiversity in

Martin J Lechowicz, McGill University, Montreal, QC, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Beringia A region at the northwestern corner of North America separated from Asia only by the shallow waters of the Bering Strait. When ocean levels decline during glaciation, Asia and North America are connected by land in this region.

Bioregion A landscape subunit at the scale of the continent that is set apart by its coherent geological and biotic history.

Continental shelf The continental margin, which is submerged to depths of approximately 180 m at current sea levels and is bounded by the abrupt drop to the abyssal depths of the oceans.

Ecoregion A landscape unit within a bioregion in which a distinct assemblage of organisms interacts ecologically, usually within a spatial context defined by drainage systems, mountain ranges, or similar natural boundaries.

Holarctic A biotic realm in the temperate, boreal, and arctic regions of the Northern Hemisphere with strong affinities among the regional floras and faunas.

Neotropics A biotic realm in tropical South and Central America with strong affinities among the regional floras and faunas but different from the tropical biota of Africa and Southeast Asia (the Paleotropics).

Physiography The landforms that lend shape and character to the continental landscape, for example, the configuration of mountains and drainage basins.

Continental Diversity

Mainland North America, with an area of 24,258,000 km^2, accounts for 16% of the land surface on Earth and stretches over 65° of latitude. The ancient core of North America is well over 3 billion years old, but parts of the continent are much younger. The continent is bounded on the west by a series of mountain ranges along the coast of the Pacific Ocean. A more extensive coastal plain and less rugged mountains characterize the Atlantic coast of North America. The Arctic Ocean lies at the northern limit of the continent, which includes the extensive Arctic Archipelago. Rugged mountains mark the commonly accepted limit of the continent in southern Mexico, but in fact distinguishing the boundary in this mountainous and geologically complex region where North America grades into Central America is arbitrary. The continental interior is an extensive plain with relatively little topographic relief, although highlands, escarpments, and erosional features occur regionally. Major drainage systems flow both north and south from the interior and also through the Great Lakes basin to the east. The Mississippi River system flowing south into the Gulf of Mexico is among the largest drainage basins in the world. The Caribbean Sea lies southeast of the continent. North America has a land connection to South America through the Isthmus of Panama, a link that formed only approximately 3 million years ago (Ma). In the past few million years, as ocean levels have fluctuated during glacial cycles, there has also been an intermittent land connection to Asia through Beringia. People arrived in North America by this route, and perhaps others, only very recently – sometime in the past 15,000–35,000 years. With the exception of Antarctica, the Americas were the last of the continents to be colonized by people. Although recently much influenced by people, the biota of North America evolved for hundreds of millions of years without a human presence. The current diversity of the continental biota is founded on this long period of geological and evolutionary history.

Two biological processes interact over millions of years in the evolutionary history of the continental biota: speciation and extinction. Populations, isolated from one another by changing landforms or climates, diverge to form new species. Species become extinct in the face of changing environments or by chance events in small populations. The number of native species in a continental biota at any point in time reflects in part an ongoing balance between speciation and extinction on the continent as a whole. The evolutionary history of a continent is intimately interwoven with its geological history, especially through the influence of continental drift. In the 400 My since life appeared on land, interchange among different continental biota has been possible whenever continents have drifted into contact with one another. Species that have evolved separately on other continents have had opportunities to disperse and colonize North America. This occasional mixing of continental biota together with *in situ* evolution in North America has created the total pool of native species on the continent today. Although this pool of species defines the current magnitude of diversity in North America, the regional patterns of diversity are set by the different distributions of individual species within the continent. Some regions have more species and others fewer. The range of a species within a continent reflects both environmental requirements and the ability to disperse to suitable environments. Species differ in their rates of dispersal and establishment in the face of changing environments and in their ability to cross barriers such as mountains, bodies of water, or inhospitable habitat. The physiography of the continent is thus an important influence on the current patterns of species diversity within North America. Species distributions also depend on climatic patterns within the continent, which arise from both the local influence of physiographic features and the global patterns of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00100-3

atmospheric circulation from season to season. In the end, the complement of species found in a region depends on the ability of species to disperse to the region, establish there, and survive and reproduce there. Ecological processes acting on relatively short timescales (years to millennia) are the proximate controls on species distribution, but the availability of the species and the distribution of suitable habitats on the continent ultimately have arisen in geological and evolutionary processes acting over much longer timescales (millions of years). It is this interplay of processes on ecological and evolutionary timescales that affects the natural patterns of diversity in North America.

Terrestrial Bioregions of North America

Mainland North America and its offshore islands can be divided into six bioregions: (1) the Canadian Shield in the north, (2) eastern North America, (3) western North America, (4) southwestern North America, (5) a northern extension of the Central American biota in southern Mexico, and (6) outliers of the Caribbean biota on the southern tip of the Florida peninsula. **Figure 1** illustrates and depicts in a general way the boundaries of these six terrestrial bioregions in North America. There is, of course, substantial physiographic and biotic diversity within each bioregion, and the boundaries between adjacent bioregions are not precisely identifiable. Fairly distinct ecoregions can be recognized at smaller spatial scales within each bioregion, and these are important for understanding and conserving biodiversity (Abell *et al.*, 1999; Ricketts *et al.*, 1999). Larger drainage systems, such as that of the Mississippi River and its tributaries or the Great Lakes basin, subsume many terrestrial ecoregions. The geography of the terrestrial versus freshwater biota of North America differs at the scale of ecoregions and to some degree at the scale of the continental bioregions. Similarly, the recognition of bioregion boundaries at the current continental margins is arbitrary. Continental watersheds are directly linked ecologically to the marine diversity of the continental shelf, and this shelf typically shares a geological history with the adjacent continent. Despite these uncertainties in the boundaries and internal structure of the six bioregions, each does have a reasonable coherence in its geological history, current physiography, and current biota. These bioregions therefore provide a useful working framework in which to summarize the general patterns of terrestrial and freshwater species diversity within North America.

The first three of these bioregions are part of the Holarctic, a zonal assemblage of flora and fauna that has coherence across the whole of the middle and higher latitudes of the Northern Hemisphere. Species may differ among North America, Europe, and Asia, but many families and genera of plants and animals are represented throughout the Holarctic. Higher latitudes within the Holarctic realm, including the Canadian Shield, have especially strong biotic affinities. High-latitude sites in North America, Europe, and Asia are in close proximity around the polar region and share common climatic regimes; biota can disperse and establish among these high-latitude regions more easily than they can move between the more distant and climatically disparate latitudinal extremes within North America. The biota at midlatitudes in the Holarctic, however, are more strongly influenced by

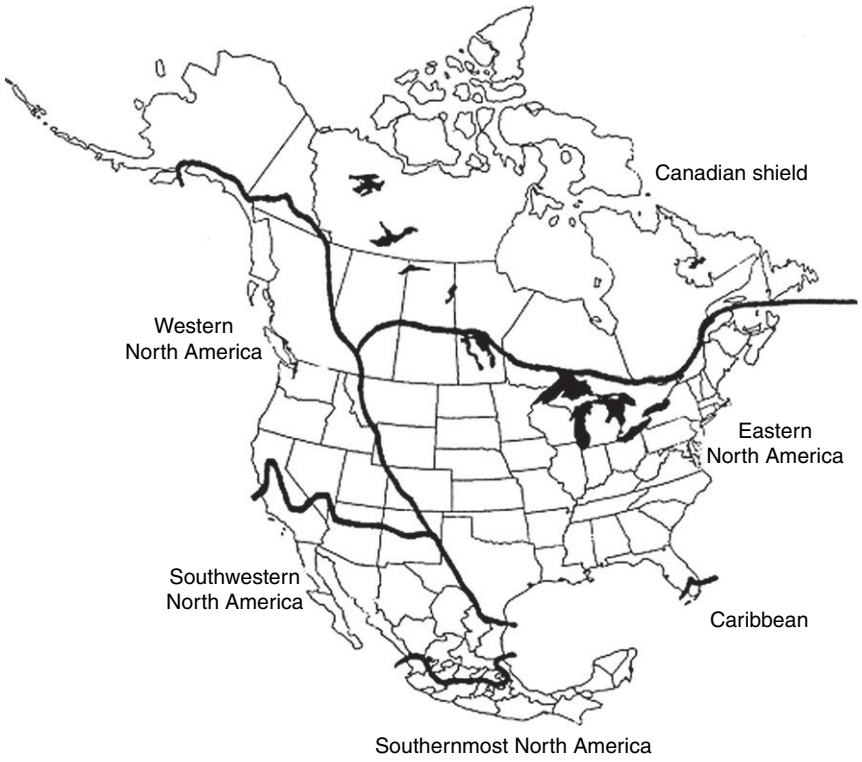


Figure 1 Schematic diagram of the six terrestrial bioregions of North America.

longitudinal climatic gradients and longitudinal contrasts in geologic history and physiography. Hence, we distinguish eastern and western bioregions in the midlatitudes of North America, whereas the Canadian Shield bioregion stretches across the top of the entire continent. Some biogeographers, recognizing similar longitudinal differences at midlatitudes on a global scale, divide the Holarctic into the Nearctic and Paleoarctic regions. North America comprises the Nearctic.

The last three North American bioregions have strong affinities with the Neotropics, the biotic realm to which South America belongs. The biota of equatorial regions does not have the global coherence that characterizes the Holarctic. Tropical regions of Africa and Southeast Asia, which have a flora and fauna that is distinct from that in the Neotropics, are distinguished as the Paleotropics. The Caribbean and southernmost North American bioregions are unambiguously Neotropical, basically representing the northern extent of floras and faunas centered on the continent of South America. The southwestern North American bioregion, however, is a mix of Neotropic and Holarctic biota and is one of the most species-rich regions in the world.

Canadian Shield

This bioregion takes its name from the Canadian Shield, which is the ancient core, or craton, of the North American continent. The bioregion and the craton are not geographically identical. The craton extends as basement rock well under the central and eastern parts of the continent, but this does not affect the regional development of the modern biota. The southern boundary of this bioregion is marked by the shift in surface geology from the ancient, crystalline bedrocks of the craton to younger, sedimentary bedrocks. The bioregion actually extends beyond the craton in northern Alaska and the Yukon, even though the bedrock in these western regions is not nearly as ancient. Two major biomes occur with the Canadian Shield bioregion: the boreal forest and the arctic tundra.

Tundra occupies the coasts and islands of the Arctic Ocean and higher elevations in interior Alaska and the Yukon. These barren lands, with a short, cold growing season, are relatively poor in plant and animal species. From one ecoregion to another within the Canadian Shield bioregion, there are only 100–750 plant species in tundra. Some tundra ecoregions have as many as 50 mammal and 160 bird species, and from 0 to 75 butterfly species, but very few have any reptile or amphibian species. There are usually only approximately 5–25 fish species in tundra. Tundra diversity decreases seasonally as birds migrate to more southern bioregions for the winter. The most diverse tundra biota occurs at high elevations in the Wrangell–St. Elias Range in southern Alaska and the adjacent Yukon. The least diverse tundra occurs along the coasts of the heavily glaciated Axel Heiberg and Ellesmere Islands, both of which extend north of 80° N latitude. There are only approximately 110 species of vascular plants in this northernmost part of North America.

The boreal forest, which occupies the more southern parts of the Canadian Shield bioregion, generally has greater biodiversity than the tundra. Different boreal ecoregions have between 250 and 1250 vascular plant species that include 5–30 tree species. There are never more than 10 coniferous tree species in any ecoregion, even though conifers dominate much of the boreal landscape. Ecoregions in the boreal forest usually have approximately 20–50 mammal species, 105–185 bird species, 35–95 butterfly species, and 2060 fish species. Amphibians are less diverse, usually 1–5 species in an ecoregion and occasionally more than 10; there may be 1–3 reptile species, but often there are none in a given ecoregion. The boreal forests of central and eastern Canada are the most diverse overall, and those of Newfoundland are the least diverse.

Eastern North America

The Appalachian Orogeny essentially created the character of this bioregion. These mountains on the eastern margin of North America arose approximately 300 Ma in the collision of all the ancient continents to form the supercontinent Pangea. The Appalachians are in fact a part of the African continent left behind as the mid-Atlantic Rift developed beginning approximately 240 Ma, gradually breaking up Pangea, creating the Atlantic Ocean and moving Europe, Africa, and the Americas to their current geographic positions. As the young and rugged Appalachians eroded, huge deposits of sediment accumulated in the continental interior, forming the bedrock of much of the interior plains that comprise the central part of this bioregion. Their sedimentary deposits also account for an extensive continental shelf, rich in marine biota today, and also an important corridor for plant and animal migrations during and after glacial peaks when ocean levels are low. From approximately 170 to 70 Ma, much of the more western interior of this bioregion lay beneath shallow continental seas in which marine biota built up extensive reef systems and sediments. The Canadian Shield forms the basement rock of essentially the entire bioregion, but it is these much younger mountains and sedimentary features that influence the current patterns of biodiversity.

The eastern parts of this bioregion are generally forested with some combination of conifer and broadleaf trees. The western parts are more often grasslands, except for gallery forests along watercourses. The transition from forest to grasslands, which is more a mosaic dependent on local topography and drainage than a smooth gradient, generally occurs to one side or the other of the Mississippi River. Savannas, grasslands with a scattering of trees, are common in the transitional ecoregions. Many of these transitional ecoregions, such as the Edwards Plateau in Texas, have exceptionally high biodiversity. There are approximately 650–3350 species of plants in the different ecoregions of eastern North America, 30–85 mammals, 160–270 birds, 5–65 amphibians, 5–85 reptiles, 75–230 butterflies, 40–230 fish species, 2–260 land snails, 5–125 mussel species, and 1–65 crayfish. The mixed forests of the southeastern US comprise the ecoregion most rich in plant species within eastern North America, and the Pine Barrens of the Atlantic Coastal Plain are the least rich. The western short grasslands are richest in mammal species, and the Florida Pine Scrub is the least rich. These western short grasslands also have the most butterfly species in this bioregion, whereas the lowland

forests of the Gulf of St. Lawrence have the least. The grasslands of the western coast of the Gulf of Mexico are the most rich in bird species, and the Nebraska Sand Hills are least rich. Very high numbers of both amphibian and land snail species occur in the forested Blue Ridge and southern Appalachian Mountains. The Tennessee and Cumberland River systems, which drain the unglaciated Appalachian–Blue Ridge region, are the most rich in fish, mussel, and crayfish species of any region in the bioregion; the glaciated, northern parts of the bioregion are most poor in freshwater biota.

Western North America

Western North America is younger and more rugged than the northern and eastern parts of the continent, with some coastal mountains having been created only in the past 1 My. All of western North America was created as the Farallon Plate, which has been squeezed between the Pacific and North American Plates, gradually subducted under the North American Plate. The subduction of the Farallon Plate caused a series of major episodes of mountain building in western North America: the Sonoran Orogeny approximately 245 Ma, the Nevadan Orogeny approximately 150 Ma, the Sevier Orogeny approximately 120–90 Ma, the Laramide Orogeny approximately 70–60 Ma, and the Basin and Range Orogeny approximately 15 Ma. The continued subduction of the Juan de Fuca Plate, a remnant of the now broken-up Farallon Plate, accounts for the current geologic activity in the Pacific Northwest. All this mountain building associated with the expansion of the North American continent has created and defined the essential character of this bioregion. The landscape is physiographically diverse, with a high degree of regional isolation leading to many distinct ecoregions and a biota rich in local endemics.

The mountains that dominate the landscape in the bioregion are predominantly forested, but various grasslands, shrublands, and deserts occur in basins and on plateaus amid the mountains. Between ecoregions, there is a considerable variation in species diversity. There are approximately 450–2550 species of plants in the different ecoregions of western North America, 30–110 mammals, 140–240 birds, 1–20 amphibians, 0–60 reptiles, 5–225 butterflies, 0–50 land snails, 5–60 fishes, 1–5 mussels, and 0–5 crayfish. Plant species diversity is the lowest on the isolated Queen Charlotte Islands and high in the Sierra Nevada, in the Great Basin, and on the Colorado Plateau. The Colorado Plateau and the Sierra Nevada are notably rich in bird species. The Colorado Plateau and the Colorado Rocky Mountains are rich in butterfly species, whereas the Queen Charlotte Islands are very poor in butterfly species. The forested coastal mountains have the most amphibians in the bioregion, and the Colorado Plateau and the Great Basin have the most reptile species. The Siskiyou Mountains, the Sierra Nevada, and the north Central Rocky Mountains are all notably rich in species of land snails. Although low in terrestrial biodiversity within the bioregion, the watersheds of the mountains along the northern Pacific Coast are the most rich in aquatic species; the isolated and arid watershed of south-central Oregon is the least rich. The Colorado Plateau shrublands are the most species-rich ecoregion overall, whereas the Queen Charlotte Islands and the adjacent northern Pacific Coast are the most species poor.

Southwestern North America

Northern and central Mexico and immediately adjacent parts of the American Southwest form a distinct bioregion within North America, which, like western North America, has rugged and relatively young terrain. The Sierra Madre Occidental and Sierra Madre Oriental running down the western and eastern coasts converge on a belt of high mountains of volcanic origin in southern Mexico, which form a natural boundary at the southern limits of this bioregion. These mountainous regions surround an interior plateau that grades north into increasingly arid deserts. The mountain ranges stand in contrast to the low relief of the coastal plains along the Pacific Ocean and the Gulf of Mexico and to the Baja California Peninsula. Ecoregions in southwestern North America are especially rich in bird, mammal, reptile, and butterfly species. Among the ecoregions, bird species number 90–280, mammals 50–110, reptiles 10–105, fishes 10–50, and there are at least 130–260 butterflies. Numbers of plant species are also high – 1500–4000 species among the ecoregions – but there are as many as 8000 species along the Sierra Madre Oriental. Not surprisingly, given the aridity of this bioregion in general, the diversity of amphibians is low – only 5–20 species in most ecoregions. There are insufficient data to designate an ecoregion that is the most rich in biodiversity overall, but the more diverse ecoregions are almost certainly in Mexico as opposed to the southwestern US.

Southernmost North America

The tropical highlands of the Sierra Madre del Sur and the Sierra Madre de Chiapas and the tropical lowlands on the Pacific and Gulf Coasts of Mexico comprise this bioregion, which is located at the southern limits of North America. This part of North America has a complex and poorly understood geological history involving the convergence of five plates and various crustal fragments of uncertain continental affinity. It appears certain, however, whether the rugged mountains of Central America and the land connection between North and South America have developed only in the past 2–5 Ma. For the preceding 70 Ma or more, North and South America had been separated by an ocean channel deep and wide enough to severely limit biotic exchange between the two continents. Once Central America arose as a bridge between the two continents, its biota was enriched by migrations in both directions, but especially by Neotropical flora and fauna from South America that were well suited to its tropical climate. Most ecoregions in this part of southern Mexico have in the order of 8000 plant species, although this decreases to only approximately 2000 in the Yucatan Peninsula, where water quickly drains deep underground in the karst landscape. Some ecoregions have as many as 125 amphibian species, but most have only 10–30. Many ecoregions have as many as 120–160 reptile species, 100–300 bird species, and 80–160 mammal species. The overall diversity of this bioregion is high and increases toward the south as the influence of the South

American biota strengthens. Defining the southern limit of North America at the Isthmus of Panama would considerably increase the biodiversity of the continent as a whole.

Caribbean

The Caribbean Plate has a poorly understood geological history. Most of the plate is submerged in a shallow sea dotted with island arcs that are closest to North America at the Florida Peninsula. The southern tip of Florida, the only part of North America that is located in the Caribbean bioregion, has a biota that is a mix of continental and Caribbean elements. There are approximately 1000 species of plants, 25 species of mammals, 175–200 species of birds, 15–20 species of amphibians, 45–50 species of reptiles, 130–135 species of butterflies, and 60–65 species of land snails. Despite the addition of Caribbean species, the biota is not exceptionally rich compared with adjacent parts of the continent.

Diversity in Major Groups of Organisms

An alternative perspective on the general levels and patterns of biodiversity in North America is to focus on the distribution of species in particular taxonomic or functional groups of organisms. Although our knowledge of the biodiversity of North America is better than that for many other parts of the world, it is far from complete and we can only adopt this perspective for the better studied organisms in North America. A few groups of vertebrate animals have been fairly well studied, but the diversity of even these is poorly known in parts of Mexico. Some groups of insects and vascular plants have been fairly well studied throughout the continent, but the sheer numbers of species in these groups undoubtedly indicate that many have yet to be discovered and properly mapped. Well-founded summaries of diversity in microbial groups, many invertebrate animals, fungi, and nonvascular plants currently do not exist for North America as a whole. Despite our incomplete knowledge of North American diversity in individual groups of organisms, the numbers of species and distributional patterns in the better known groups yield insights and provide guidance for conservation efforts.

Plants

There are approximately 1320 species of mosses in North America north of Mexico and approximately 1200 species known from Mexico. The liverworts and hornworts in North America north of Mexico number approximately 650 species, with about 800 species known from Mexico. All these numbers are expected to increase as the continental bryoflora is better explored. There is no comprehensive North American flora for these plant groups.

The pteridophytes (ferns, horsetails, club mosses, spike mosses, and quillworts) are among the most primitive vascular plants represented in the North American flora. There are approximately 440 species of pteridophytes in North America north of Mexico, more than 75% of which are ferns. Mexico has in the order of 1000 pteridophytes species, most of which are ferns of moist habitats and Neotropical affinity. For example, there are only approximately 125 pteridophyte species in the Chihuahuan Desert of northern Mexico compared with almost 700 species in the southern state of Oaxaca. The moist, tropical forests of southern Mexico harbor the greatest diversity of pteridophytes in North America.

The gymnosperms, an ancient group of plants that includes the pines, spruces, and other conifers, are commercially important and especially well known in North America. There are approximately 90 conifer species in North America north of Mexico; the numbers of conifers in Mexico are higher but uncertain. The well-studied pines of Mexico offer one specific point of comparison with the rest of North America. There are approximately 35 pine species in North America north of Mexico, and only about one-third of these have ranges extending into Mexico. There are another 60 species of pine in Mexico! However, there are 11 spruce species in North America, but only two occur in Mexico. The frequently isolated, arid mountain habitats of Mexico clearly have provided a center for the evolutionary diversification of pines in North America. North of Mexico, the greatest concentration of conifer species is found in the Klamath–Siskiyou ecoregion on the mountainous coast at the California–Oregon border; approximately 30 species of conifer are found there. Conifer diversity is generally greater in the western mountains than in eastern or northern North America.

Trees in general, angiosperm as well as gymnosperm species, have been well studied in North America. There are approximately 650 native species of trees in North America north of Mexico, approximately 85% of which are angiosperms; the arboreal flora of Mexico is not known with certainty but there are at least as many, and probably substantially more, tree species as there are in the rest of the continent. The oaks, which have been well studied in North America including Mexico, provide a useful point of comparison. As many as 150 species of oaks are known from Mexico, approximately 30 of which are also found farther north. There are only about 60 additional oak species in the rest of North America. The Mexican oaks predominate in arid habitats, with the mountains of central and eastern Mexico being a center of diversity. The center of tree diversity north of Mexico is in the southeastern US, where as many as 190 tree species can be found in a given ecoregion. With the exception of Mexico, the diversity of hardwood species declines steadily to the west and north of this concentration of species in the southeastern US.

The angiosperms, both herbaceous and woody, are the most recently evolved and most species rich of the vascular plant groups in North America and indeed throughout the world. The angiosperms comprise the largest part by far of any regional flora. They are also the least completely identified and mapped in terms of their North American patterns of biodiversity. There is no comprehensive floristic inventory for North America, only a series of regional descriptions. There may be as many as 17,000–20,000 native plant species in North America north of Mexico. At midlatitudes, the highest diversity of angiosperm species in an ecoregion is in the forests of the southeastern US, where there are approximately 3000 native angiosperms; species diversity decreases to the west and north on the continent but increases dramatically in Mexico. The state of California, which includes many diverse

ecoregions, has more than 5000 native plant species. The whole of Canada has about only 3125 native angiosperm species, and the Canadian Arctic Archipelago has only 330. Mexico, however, has approximately 19,000 angiosperm species; the state of Oaxaca alone has about 8000. Mexico has approximately 2600 of the 5500 species in the Asteraceae (sunflower family); there are only 400 native species of Asteraceae in Canada. The Fabaceae (bean family) has approximately 1700 species in Mexico, with more than half occurring only in Mexico, compared with only about 130 in Canada. The 900 species of Cactaceae (cacti) in Mexico represent more than half the species in the world. The greatest angiosperm diversity in North America can be clearly found at the southern end of the continent.

Butterflies

There is no reliable estimate of the total number of butterfly species in North America. Comprehensive surveys of western and southwestern North America report only approximately 530 species of butterflies, although this is believed to be from the part of the continent north of Mexico where butterfly diversity is greatest. Surveys of remnant prairie in the midcontinent yielded only approximately 90 species of butterflies, but the prairie ecosystems have been much disrupted in the past 200 years. There are estimated to be approximately 2200 butterfly species in Mexico, and it is certain that butterfly diversity within North America is greatest in Mexico. The swallowtails (Papilionidae), which are among the more conspicuous and better known butterflies, provide evidence of the relative richness of the Mexican butterfly fauna. There are almost 60 swallowtail species in Mexico, about twice as many as in the more northern parts of the continent.

Land Mammals

There are approximately 4500 mammal species in the world. Canada has approximately 140 species of mammals, whereas the US has about 350 and Mexico about 440–450. There is no good composite estimate for the continent as a whole. North of Mexico, the Colorado Plateau harbors the greatest diversity of mammal species but the Chihuahuan Desert, which extends into Mexico, is equally rich in mammal species. The species richness of mammals in North America decreases steadily to the north, and at midlatitudes, it is relatively low in the east compared with the west.

Birds

There are approximately 9000 bird species in the world. Canada has approximately 425 species of birds, the US approximately 650 species, and Mexico approximately 960–1050 species. There is no good composite estimate for the number of bird species on the continent as a whole. The geographic pattern of species richness for birds in North America is generally similar to that of mammals: richest in Mexico and the adjacent southwestern US and steadily decreasing to the north.

Amphibians

The amphibians in North America primarily include frogs and salamanders. There are approximately 4000 amphibian species in the world. Canada has approximately 40 species of amphibians, none of which also occur farther south on the continent. Mexico, in contrast, has approximately 285 amphibian species, approximately 170 of which do not occur elsewhere in North America. The US has approximately 230 amphibian species. There is no good estimate of the number of amphibian species for the continent as a whole. With 55–70 amphibian species, the Appalachian Mountains comprise the region north of Mexico that is most rich in amphibian species, but this is only one-fourth of the number of species in the tropical highlands of central and southern Mexico.

Reptiles

The reptiles in North America include turtles, lizards, snakes, and a negligible number of crocodilians. There are about 6000 reptile species in the world. Canada has only about 40 species of reptiles, none of which also occur farther south on the continent. The US has approximately 280 species of reptiles. Mexico has about 690–720 reptile species, approximately 370 of which do not occur elsewhere in North America. The Chihuahuan Desert is the part of the continent most rich in reptiles.

Freshwater Fishes

North America has approximately 1200 species of freshwater fish. There is, however, a substantial variation in the diversity of fishes in glaciated and unglaciated parts of the continent. The glaciated regions are notably poor in fish species. Canada, despite its very large land area rich in lakes and river systems, has only approximately 180 native fish species. The extensive Hudson Bay Basin, which was entirely glaciated 18,000 years ago, has only approximately 100 fish species and most of these occur in the southern headwaters of the drainage system. The Canadian Arctic Archipelago has fewer than 10 fish species. In contrast, the largely unglaciated Mississippi River Basin has approximately 375 fish species. The southeastern US has 485 freshwater fish species, with the greatest diversity in the Appalachian and adjacent Interior Plateau. Despite the aridity of much of Mexico, about 380–500 freshwater fish species occur there, with notably high numbers of species in the Panuco and Papaloapan River systems along the southwestern coast of the Gulf of Mexico. The arid regions of southwestern North America had wetter climates and more extensive wetland systems as recently as 7500 years ago, but the fish fauna is now restricted to remnant bodies of permanent water. Although the habitats available for fishes in southwestern North America are currently limited, there are many remnant species with a high degree of endemism.

Changes in North American Biodiversity

There is a long history of patterns of biodiversity in North America. The native species that occupy North America today

belong to groups of organisms that have developed over the long history of life on this continent. Some of the earliest life on land is represented by fossil plants more than 400 Million years old that are found along the Acadian coast of North America. The continent has shared in the evolutionary diversification of life on land. Ancestors of conifers in the family Pinaceae grew in what is now North America 135 Ma. Angiosperm fossils from North America date to approximately 120 Ma, near the origins of this now dominant group of plants. Dinosaurs that dominated the animal life of North America beginning approximately 200 Ma are ancestors to the modern birds; the mammals rose to dominance beginning 65 Ma with the mass extinction of dinosaurs. Primitive insects are among the earliest fossils of terrestrial life in North America, but the rapid increases in species numbers of groups such as the butterflies and beetles that have high diversity in North America today also began only 65 Ma. Patterns of continental drift and climate change during the 400 My that life has been on land have facilitated the exchange of biota from different continents. The strong affinity in the flora and fauna of the Holarctic regions of North America, Europe, and Asia has its origins in the period 170–80 Ma, when these continents were still in close proximity to one another. Similarly, the richness of the Central American flora and fauna in southernmost North America stems in part from biotic exchanges initiated in the past 2 or 3 Ma when the Americas were reconnected by land after 60–70 My of relative isolation from one another. Some contemporary patterns of diversity in North America that we might seek to attribute to current environmental conditions in fact originated in the ancient history of life on the continent.

The influence of past events on contemporary patterns of biodiversity has been most marked during the past 1 My, during which climatic changes associated with glacial cycles have repeatedly disrupted the North American biota. Species ranges in North America have shifted dramatically from glacial to interglacial, and current distributions of species reflect patterns of dispersion from the most recent glacial refugia. The relative poverty of species in the more northern parts of North America in part stems from the failure of some species to disperse back into these glaciated regions in the 5000–14,000 years since the ice melted. Major climatic changes from glacial to interglacial periods have also influenced the biodiversity of unglaciated regions of the continent. For example, during the most recent glacial maximum, parts of western and southwestern North America had a wetter climate. Large pluvial lakes and extensive river systems existed that are now mostly reduced to playas and dry channels. Today, desert pupfish occur only in isolated springs and pools, whereas once they were widespread in extensive lake and river systems throughout the Southwest. Once widespread palms now exist only as remnant populations in isolated canyons. In another cycle of changing climate, such remnant populations may again become common, but now we value them as rare species that enhance the biodiversity of North America. The rich and sometimes peculiar patterns of biodiversity in North America are built on a dynamic history of changing environments and changing species distributions.

In terms of the most recent history of the North American biota, the arrival of humans coincident with the end of the most recent glacial cycle is especially noteworthy. As the interglacial was beginning, hundreds of the larger animals in North America became extinct. It is difficult to account for these extinctions by changing climate alone, and it appears that human hunting pressures were a significant factor. In a brief period, the mammoths, mastadons, camels, horses, saber-tooth tigers, giant ground sloths, and many other large mammals that undoubtedly were major elements in the functioning of North American ecosystems simply disappeared from the continent. The patterns of biodiversity at the time of European settlement not many thousands of years later must reflect this fairly recent mass extinction as well as the continued influence of the few million aboriginal people who already lived in North America at that time. Whatever patterns of continental diversity we identify, we should not assume that they are independent of ancient human influence.

Human influence on patterns of diversity in North America has increased considerably in the past 500 years. The European colonists consciously or inadvertently brought many new plant and animal species to this continent. Indeed, people continue to introduce species from outside North America, some of which establish and spread to natural ecosystems well beyond the confines of urban or agricultural lands. Approximately 25% of the current flora of Canada consists of alien species; the US has approximately 3700 alien plant species, of the order of 20% of the flora. Approximately 6% of the fishes of Canada are not native species; the US has about 75 fish species native outside of North America and approximately 200 that have been introduced outside their native North American ranges. The fishes currently dominating the ecosystems of the Great Lakes are entirely the outcome of the introduction and management of nonnative species by people. Inadvertently introduced zebra mussels have an altered ecosystem function in the Great Lakes Basin and ousted many native mussel species. Introduced starlings and sparrows dominate the avifauna in many settled landscapes. The landscapes and habitats of large parts of the continent have been drastically altered from preColumbian times. Almost all rivers in the US have had their flow altered by dams or their drainage basin hydrology has been manipulated. Terrestrial habitats reasonably unaltered from preColumbian times now occupy only 48% of North America north of Mexico, and these are largely concentrated in remote mountainous or northern regions. Remnants of natural prairie and forest habitats represent only a few percent of many ecoregions in eastern North America. In western North America, natural habitats are essentially obliterated in the grasslands of central California and in the interior grasslands of the Pacific Northwest. We are faced with the challenge of identifying natural patterns in the biodiversity of our native flora and fauna while the continental biota is increasingly disrupted by the introduction of alien species and the alteration of natural landscapes.

Conclusion

From the preceding discussion, assembled from a variety of sources, one point is most clear: we have really only begun to know the patterns of biodiversity in North America. The biodiversity of no group of organisms in North America is

completely characterized. Some of the groups of organisms discussed in this review have been reasonably well studied, but all are incompletely known. Remote parts of arctic North America are little explored and may yield new species despite the general tendency of species diversity to decrease at higher latitudes. There are unusual habitats in the high arctic that are unlike anywhere else on Earth, such as the mineral-rich hot springs on Axel Heiberg Island that emerge through hundreds of meters of permafrost. Such habitats, which are being studied as analogs for Martian polar environments, may well harbor previously unknown microbial life. The more heavily settled and biologically explored midcontinent seems well known, but every year new localities and range extensions, and even new species, are recorded. The remote and rugged parts of southern North America undoubtedly harbor unknown species even in groups that have been well studied on the continent as a whole. Mexico is recognized as one of the top five countries in the world in terms of the overall levels of biodiversity. By far the most species-rich part of North America, Mexico is ironically the least well studied.

In addition to our incomplete knowledge of species diversity in well-studied groups of organisms in North America, we must recognize that there are many other groups of organisms on the continent whose distribution and abundance are scarcely known. There are almost 91,000 species of insects described from North America north of Mexico, but this number is certainly only a small fraction of the total insect species on the continent. Almost all the insect groups have been too little studied to assemble any sort of meaningful summary of their patterns of continental biodiversity. Similarly, there may be as many as 120,000 species of microfungi in the US alone, only one-fourth of which have been described, and far fewer have been studied sufficiently to map and inventory them for the continent. The 5000–10,000 species of macrofungi (mushrooms) estimated to occur in the United States are only marginally better known. It is too early to write the definitive treatise on patterns of biodiversity in North America. There is much exciting and worthwhile work yet to do.

See also: Central America, Ecosystems of. Ecosystems of South America

References

Abell R, Olson DM, Dinerstein E, *et al.* (1999) *Freshwater Ecoregions of North America. A Conservation Assessment.* Washington, DC: Island Press.

Boyce SG and Martin WH (eds.) (1993) *Biodiversity of the Southeastern United States: Lowland Terrestrial Communities.* New York: Wiley.

Brown DE (ed.) (1994) *Biotic Communities: Southwestern United States and Northwestern Mexico.* Salt Lake City: University of Utah Press.

Brown DE, Reichenbacher F, and Franson SE (1998) *A Classification System of North American Biotic Communities.* Salt Lake City: University of Utah Press.

Cartron J-LE, Ceballos G, and Felger RS (eds.) (2005) *Biodiversity, Ecosystems, and Conservation in Northern Mexico.* New York: Oxford University Press.

Hackney CT, Adams SM, and Martin WH (eds.) (1993) *Biodiversity of the Southeastern United States: Aquatic Communities.* New York: Wiley.

Harper KT, St. Clair LL, Thorne KH, and Hess WW (eds.) (1994) *Natural History of the Colorado Plateau and Great Basin.* Niwot: University Press of Colorado.

La Roe ET, Farris GS, Puckett CE, Doran PD, and Mac MJ (eds.) (1995) *Our Living Resources: A Report to the Nation on the Distribution, Abundance, and Health of U.S. Plants, Animals, and Ecosystems.* Washington, DC: U.S. Department of the Interior, National Biological Service.

Lawford R, Alaback P, and Fuentes ER (eds.) (1993) *High Latitude Rain Forests and Associated Ecosystems of the West Coast of the Americas: Climate, Hydrology, Ecology, and Conservation.* New York: Springer-Verlag.

Martin WH, Boyce SG, and Echternacht AC (eds.) (1993) *Biodiversity of the Southeastern United States: Upland Terrestrial Communities.* New York: Wiley.

Ramamoorthy TP, Bye R, Lot A, and Fa J (1993) *Biological Diversity of Mexico: Origins and Distribution.* New York: Oxford University Press.

Ricketts TH, Dinerstein E, Olson DM, *et al.* (1999) *Terrestrial Ecoregions of North America. A Conservation Assessment.* Washington, DC: Island Press.

Rzedowski J (1994) *Vegetación de México.* México: Editorial Limusa.

Schoenherr AA (1992) *A Natural History of California.* Berkeley: University of California Press.

Sisk TD (ed.) (1998) Perspectives on the land use history of North America: A context for understanding our changing environment. *U.S. Geological Survey, Biological Resources Division, Biological Science Report USGS/BRD/BSR.* 1998-0003.

Stehli FG and Webb SD (eds.) (1985) *The Great American Biotic Interchange.* New York: Plenum.

West FH (ed.) (1996) *American Beginnings: The Prehistory and Palaeoecology of Beringia.* Chicago: University of Chicago Press.

Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms

Laura F Landweber and Tamara L Horton, Princeton University, Princeton, NJ, USA
Jonatha M Gott, Case Western Reserve University, Cleveland, OH, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Tamara L. Horton and Laura F. Landweber, volume 4, pp 415–426, © 2001, Elsevier Inc.

Glossary

DNA Deoxyribonucleic acid; a linear polymer of nucleotides that encode information for a cell.
Eukaryote An organism whose cell or cells have a membrane-bound nucleus.
Mitochondrion A eukaryotic cellular organelle that is used for cellular respiration and energy production.
Nucleus Compartment of a cell in which DNA is stored on chromosomes.
Protein A three-dimensional macromolecule constructed of amino acids that is formed based on an RNA sequence.
Protist A member of a diverse collection of eukaryotes that are defined only by their exclusion from other groups – plants, animals, and fungi.
RNA Ribonucleic acid; a linear polymer of nucleotides that is transcribed from DNA.

Introduction

DNA is often described as a blueprint for life, implying that knowledge of the primary sequence of nucleic acids in a genome can give biologists a complete picture of the organism built from these plans. However, neither life nor DNA is that simple. In reality, the sequence of nucleotides in a genome is often just a starting point for the construction of DNA genes and RNA transcripts, which undergo many alterations before organisms actually use the information to fashion their building materials.

Two fascinating modes of nucleic acid sequence modification are gene scrambling and RNA editing. Gene scrambling is the rearrangement of DNA segments between a transcriptionally active copy and an archived germline copy. Some genes are broken into more than 50 unordered fragments along a germline chromosome and then unshuffled during formation of an active somatic chromosome. Although gene unscrambling assembles existing DNA information into a new order, RNA editing of transcripts creates completely novel sequences that are not even found in the genome. Processing of edited RNAs alters the transcript length and nucleotide identity, and the transformations can render the expressed RNA form unrecognizably different from the DNA molecule of its origin. Although gene scrambling and RNA editing complicate our analysis of genomes, understanding the methods by which organisms achieve these genetic revisions gives us an appreciation for the diversity of ways by which nucleic acids can store and recombine information.

Gene Scrambling

Gene scrambling occurs when coding segments of DNA are mixed in a randomly or a nonrandomly shuffled order along a chromosome. Before the stored information can be expressed by the organism, the pieces of the gene must be cut apart and reassembled in the proper sequential arrangement. These stunning acrobatics are performed in the genomes of spirotrichous ciliates, protists that have many unique and mysterious features. However, the fundamental lessons learned from spirotrichs about the flexibility of nucleic acid storage mechanisms are universally important.

Ciliates and Nuclear Dualism

Ciliates are unicellular protists that on phylogenetic trees diverge together with apicomplexan parasites and dinoflagellates, all members of the alveolates. The ciliates are a diverse monophyletic group, with certain species estimated to be as evolutionarily distant from one another as corn from rats. All ciliates share two features: a coating of cilia on their cell surfaces and two types of nuclei within single cells.

The two nuclei types in each ciliate cytoplasm are different sizes; they are called the *micronucleus* and the *macronucleus*. The tiny germline micronucleus is usually transcriptionally inert and functions solely in sexual exchange. In contrast, the large somatic macronucleus is responsible for gene expression, but its contents are only transmitted to asexual offspring. Ciliates reproduce asexually but are capable of exchanging genetic information in a sexual manner independent of reproduction. Conjugation between ciliates leads to an exchange of haploid micronuclei that fuse to form a zygotic nucleus (**Figure 1**). The biparentally created zygotic nuclei in each mating partner form new micronuclei and macronuclei, and the old macronuclei are destroyed.

DNA Processing during Macronuclear Formation

The micronuclear and macronuclear genomes do not have the same chromosomal structure. As the new macronucleus is

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00313-0

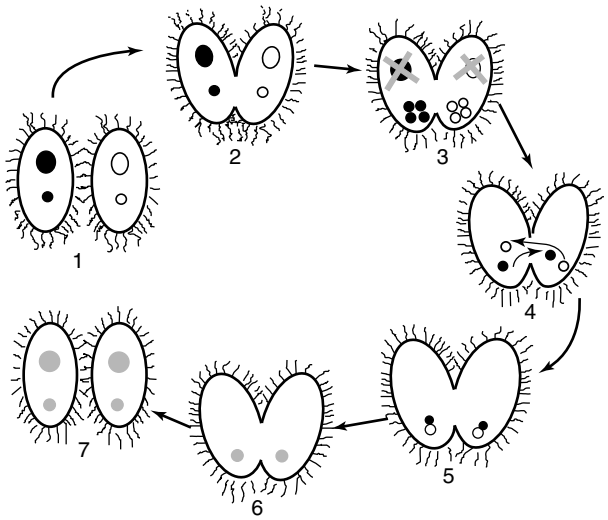


Figure 1 Ciliate sex. The micronuclei of conjugating ciliates undergo meiosis, exchange, and fusion to form new genetic combinations. Between steps 1 and 2, the ciliates conjugate. In the transition from step 2 to step 3, the micronuclei have undergone meiosis to form haploid micronuclei while the old macronuclei have been destroyed. In step 4, the haploid micronuclei are exchanged, and in step 5 they fuse. By step 6, two unique diploid micronuclei are formed with genetic material from both parents. At step 7, a new macronucleus is formed from each new micronucleus.

formed, diploid micronuclear chromosomes become polytene and reproducibly break at certain sites, portions of DNA undergo deletion, chromosomes are rejoined, and new chromosome ends are repaired by telomerase. DNA in the new, shortened chromosomes is differentially amplified so that copy numbers of individual chromosomes in *Tetrahymena thermophila* range from 45 to 9000. Spirotrichous ciliates tend to have higher copy numbers of chromosomes than the oligohymenophorans such as *Tetrahymena*, and these copy numbers are regulated by noncoding RNA (Nowacki *et al.*, 2010). Disparities in DNA copy number can relate to levels of gene expression, because the most highly amplified DNAs encode the highly expressed ribosomal RNAs, and the quantities of mRNA synthesis and cognate protein secretion from *Euplotes raikovi* pheromone genes directly correlate with the genes' individual macronuclear copy numbers.

Coding sequences in micronuclear genes are called MDSs for *macronuclear destined sequences* (or *segments*), whereas IESs are *internal eliminated* (or *excised*) *sequences*. MDSs and IESs are superficially analogous to exons and introns, respectively, although the latter two terms only refer to a particular type of RNA processing and do not apply to this unrelated restructuring of DNA. The amount of DNA eliminated during macronuclear formation varies extensively across diverse ciliates: the oligohymenophoran *T. thermophila* eliminates approximately 15% of its micronuclear genome, whereas spirotrichous ciliates such as *Stylonychia* and *Oxytricha* eliminate 95–98% of their micronuclear genomes to create macronuclear chromosomes that typically bear single genes and very little noncoding sequence (Prescott, 1994).

Scrambled Genes and the Unscrambling Process

The drastic genome rearrangements of spirotrichous ciliates are not confined to the extreme quantity of DNA deletions. The protein-coding MDSs in *Oxytricha* and *Stylonychia* species are sometimes disordered relative to their final position in the macronuclear copy. For example, Prescott and colleagues found that in *Oxytricha nova*, the micronuclear copy of three genes (actin I, α telomere-binding protein, and DNA polymerase α) must be reordered while intervening DNA sequences are removed in order to construct functional macronuclear genes. In *O. nova*'s micronuclear genome, the MDSs destined to construct the gene for α telomere end-binding protein (TEBPα) are arranged in the cryptic order 1–3–5–7–9–11–2–4–6–8–10–12–13–14 relative to their conventional order in the macronucleus of 1–2–3–4–5–6–7–8–9–10–11–12–13–14. Most impressively, the gene encoding DNA polymerase α (DNA pol α) in *Oxytricha trifallax* and *Stylonychia lemnae* is scrambled in 48 or more pieces in their germline nuclei (**Figure 2**).

Genome rearrangements in spirotrichous ciliates are guided by maternally inherited noncoding RNAs (Nowacki *et al.*, 2008) – the same noncoding RNAs that regulate chromosome copy number (Nowacki *et al.*, 2010). Homologous recombination at MDS boundaries is also part of the unscrambling process. A segment of DNA sequence at the junction between a particular MDS and its downstream IES usually matches the sequence at the junction of the next MDS and its upstream IES, leading to the correct ligation of the two MDSs over a distance. However, the presence of shared repeat regions as short as an average of 4 base pairs (bp) for nonscrambled MDSs and 9 bp for scrambled MDSs suggests that although these recombination guides may be necessary, they are certainly not sufficient to guide accurate assembly of the genes. Hence, it is more likely that the repeats satisfy a structural requirement for MDS joining rather than perform any role in substrate recognition, which the noncoding template RNAs provide instead. Otherwise, incorrectly spliced products of promiscuous recombination would dominate genomes since 2–4 bp repeats occur many thousands of times throughout the micronucleus. This incorrect hybridization could be a driving force in the production of newly scrambled patterns in evolution. Nonetheless, if this sort of ambiguous unscrambling actually occurs during macronuclear development, then only unscrambled molecules that contain both 5 and 3′ telomere addition sequences are selectively retained in the macronucleus, ensuring that most haphazardly ordered genes would be lost.

The pattern of MDSs in the micronuclear genome provides a strong clue to both the ciliates' orchestration of the unscrambling process and the mutational events that led to the scrambling in each gene sequence. For example, in the previously mentioned gene encoding *O. nova*'s TEBPα protein, the micronuclear order of MDSs (1–3–5–7–9–1–2–4–6–8–10–12–13–14) predicts a spiral mechanism in the unscrambling path to link odd and even segments in order (**Figure 3**).

In contrast, the DNA polymerase α gene has at least 44 MDSs in *O. nova* and 50 in *O. trifallax* scrambled in a nonrandom order with an inversion in the middle; some MDSs

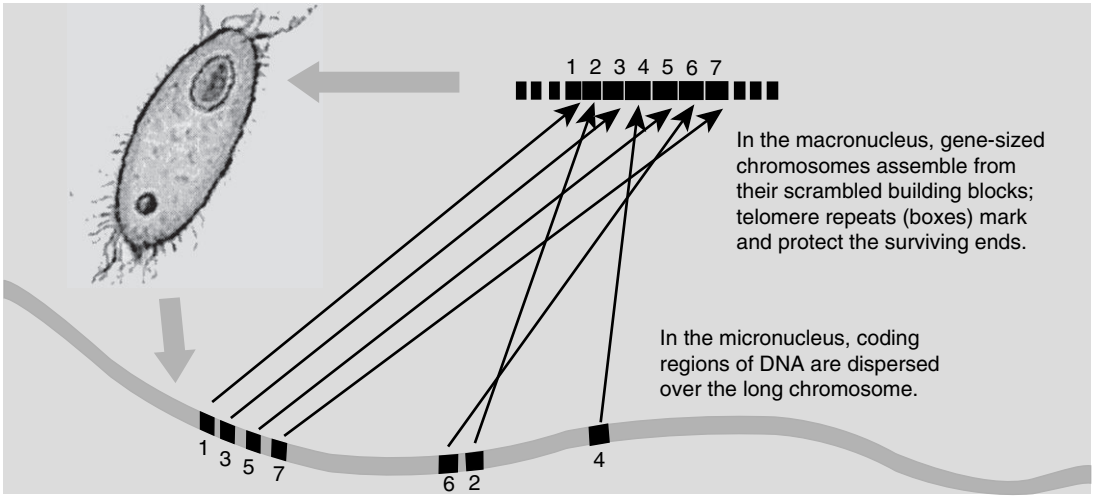


Figure 2 Overview of gene unscrambling. Dispersed coding MDSs 1–7 reassemble during macronuclear development to form the functional gene copy (top), complete with telomere addition to mark and protect both ends of the gene.

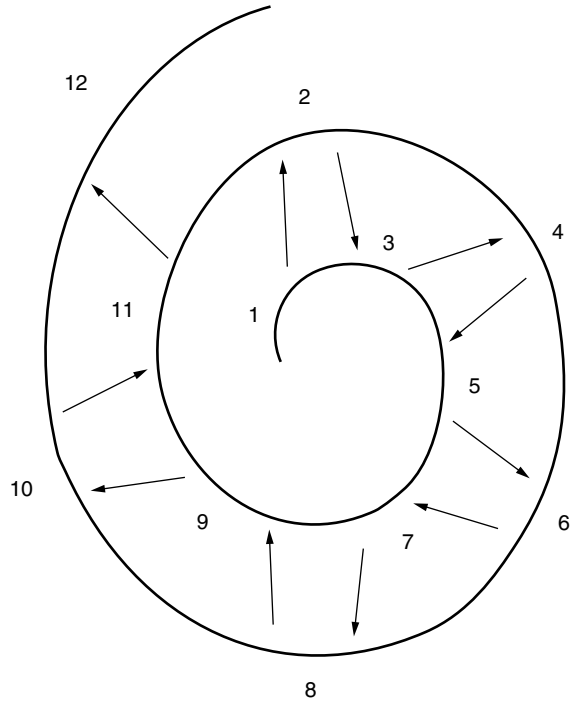


Figure 3 Spiral model for unscrambling in TEBPα.

are located at least several kilobases away from the main gene in an unmapped fragment. A hairpin structure could help resolve MDSs during unscrambling (**Figure 4**). Comparison of the *O. nova* sequence with that of *O. trifallax* allows precise predictions for the origin of new scrambled segments (Figures 4 and 5). Micronuclear junction sequences promote pairing between each MDS and the noncoding IES on the opposite side of a hairpin (**Figure 4**). Any mutation that leads to an increased stabilization of this hairpin during macronuclear chromosome formation would also increase the chance of homologous recombination between correct segments within the micronuclear genome. Consequently, selection would favor the appearance of more scrambled MDS segments in such a nonrandomly scrambled gene since each additional MDS adds more paired junction sequences to stabilize the hairpin necessary for unscrambling the already existing MDSs. This explanation is consistent with the additional MDSs in the DNA pol α gene in *O. trifallax*. The arrangement of MDSs 2, 6, and 10 in *O. nova* could have given rise to the arrangement of eight new MDSs in *O. trifallax* (**Figure 4**) by multiple crossovers in the germline micronucleus. Thus, the appearance of an inversion leads to the introduction of new MDSs in a nonrandomly scrambled pattern.

Gene scrambling in ciliates may have evolved as a product of an increased capacity for homologous recombination. The reason why it has been detected in only a restricted group of ciliates might reflect their increased levels of recombination, which generate both the scrambled arrangements and the subsequent process of unscrambling them. Although it is difficult to propose an adaptive argument for the presence and maintenance of such a complicated gene-decoding procedure, the forces that have led to ciliate nuclear dualism might be at the root of its origin. For example, Yao and Doerder independently proposed that the different roles required of the micronucleus and macronucleus, as well as intragenomic competition, may have led to disruptive selection acting on their genetic organization and chromatin structure. The accommodation of both types of nuclei within a single cytoplasm may promote or at least permit profound differences to arise that distinguish active genes from transmitted genes. Likewise, the gene scrambling present in certain spirotrichous ciliates may be a profound exaggeration of this solution to problems presented by ciliate life.

RNA Editing

RNA editing is the alteration of RNA sequences by base modifications, substitutions, insertions, and deletions (see Knoop, 2011

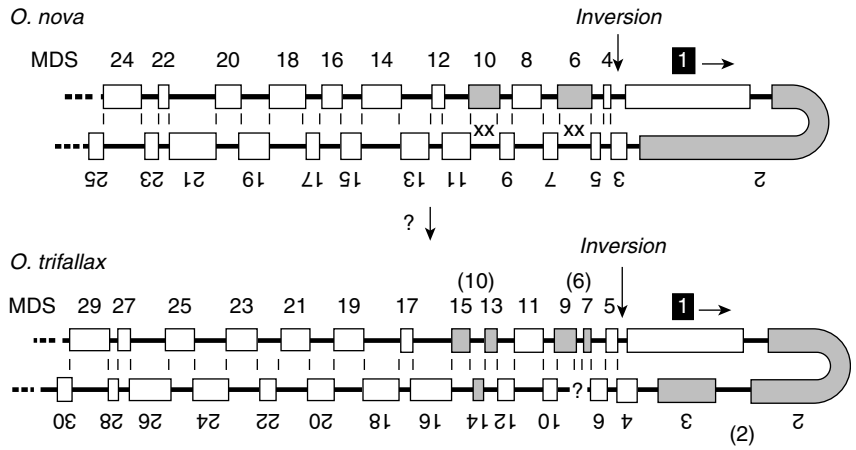


Figure 4 Model for scrambling of DNA pol α. Vertical lines indicate recombination junctions between scrambled MDSs guided by direct repeats. MDS 1 contains the start of the gene. MDS 10 in *O. nova* can also give rise to three new MDSs (13–15) in *O. trifallax*, one scrambled on the inverted strand, by two spontaneous intramolecular recombination events (x's) in the folded orientation shown. *Oxytricha nova* MDS 6 can give rise to *O. trifallax* MDSs 7–9 (MDS 8, shaded, is only 6 bp and was not identified in Hoffman and Prescott, 1997). *Oxytricha trifallax* nonscrambled MDSs 2 and 3 could be generated by the insertion of an IES in *O. nova* MDS 2 (similar to a model suggested by M. DuBois as cited in Hoffman and Prescott, 1997).

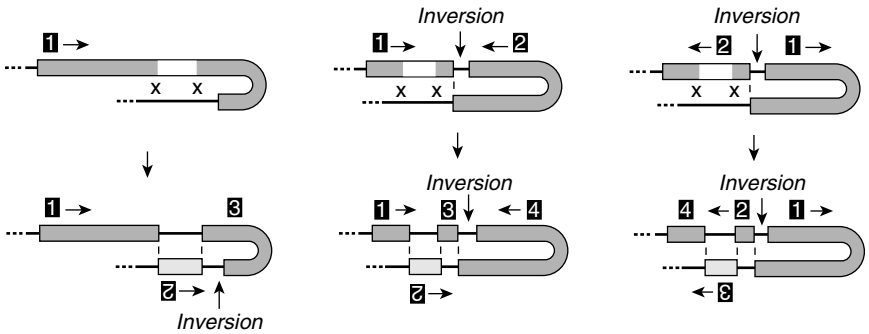


Figure 5 Proposed model for the origin of a scrambled gene. (Left) Birth of a scrambled gene from a nonscrambled gene by a double recombination with an IES or any noncoding DNA (new MDS order 1–3–2 with an inversion between MDSs 3 and 2). (Middle) Generation of a scrambled gene with a nonrandom MDS order from a nonscrambled gene with an inversion between two MDSs. (Right) Creation of new scrambled MDSs in a scrambled gene containing an inversion. Inversions may dramatically increase the production of scrambled MDSs by stabilizing the folded conformation that allows reciprocal recombination across the inversion.

for a comprehensive review). The process of rewriting RNA transcripts by editing produces major effects, even adding more than half of the nucleotides in some mitochondrial transcripts of kinetoplastid protozoa. In contrast, the impact of editing can still be large even in cases in which the physical extent of editing is small. For instance, in human apolipoprotein B transcripts, the replacement of a single cytidine (C) by uridine (U) results in the conversion of a glutamine codon to a stop codon; the early termination of apo B translation shortens the resultant polypeptide by half and removes a functional domain. Since the initial discovery of RNA editing as extensive U insertions and deletions in trypanosome mitochondrial mRNAs, many additional and apparently unrelated examples of editing have been found in organisms ranging from Ebola virus to humans. Substitution or modification editing exists in many nuclear and organellar RNAs among a diverse set of eukaryotes, whereas insertion or deletion editing is found in a much more limited set of organisms (Table 1).

On first inspection, the use of RNA editing in gene expression seems inefficient. Why not encode genes in their final edited form rather than require an additional revision step? However, organisms exploit the ability to edit RNA in amazingly clever ways. RNA editing allows the persistence of *deleterious* mutations in DNA genomes: new point mutations and frameshifts may persist if they are repaired at the RNA level, and DNA copies of genes and control sequences in crowded genomes may overlap but generate two or more discrete RNA sequences through editing. RNA editing also offers an array of posttranscriptional modes of genetic regulation through the formation of start and stop codons, intron splice sites, and open reading frames and the potential alteration of microRNA (miRNA)–mRNA interactions, mRNA localization, and stability. Editing even permits

Table 1 Types and distribution of RNA editing

Organism	*A>I*	*C>U*	*U>C*	*+A*	*+C*	*+G*	*+U*	*−A*	*−U*	*N to N*
Metazoa	mRNA[1] tRNA[2] ncRNA[3]	mRNA[4] mt tRNA[5]	mt mRNA[6]							mRNA[7] mt tRNA[8]
Viruses	HDV[9] HIV[10]			Ebola[11]		Paramxyoviruses[12]				
Plants (Plantae)	cp tRNA[2]	mt mRNA[13] mt tRNA[14] cp mRNA[15] cp tRNA[16]	mt mRNA[17] cp mRNA[18]							
Kinetoplastids (Excavata)	tRNA[19]	tRNA[19] mt tRNA[20]					mt mRNA[21]		mt mRNA[22]	
Naegleria (Excavata)		mt mRNA[23]								
Myxomycetes (Amboebozoa)		mt mRNA[24]		mt mRNA[24,25] mt rRNA[26]	mt mRNA[27] mt rRNA[28] mt rRNA[26]	mt mRNA[24,25] mt rRNA[25]	mt mRNA[24] mt tRNA[28] mt rRNA[26]	mt mRNA[29]		mt tRNA[30]
Acanthamoeba (Amboebozoa)										mt tRNA[31]
Spizellomyces (Fungi)										mt tRNA[32]
Dinoflagellates (Protista)										mt mRNA[33] mt rRNA[34] mt tRNA[34]

[1]Sommer *et al.* (1991).
[2]See Jühling *et al.* (2009).
[3]Luciano *et al.* (2004).
[4]Powell *et al.* (1987); Chen *et al.* (1987).
[5]Janke and Pääbo (1993).
[6]Burger *et al.* (2009).
[7]Li *et al.* (2011).
[8]Yokoboro and Pääbo (1995a, b).
[9]Luo, et al. (1990).
[10]Doria *et al.* (2009).
[11]Volchkov *et al.* (1995).
[12]Thomas *et al.* (1988).
[13]Covello and Gray (1989); Gualberto *et al.* (1989); Hiesel (1989).
[14]Marechal-Drouard *et al.* (1993); Binder *et al.* (1994).
[15]Hoch *et al.* (1991).
[16]Kugita *et al.* (2003).
[17]Gualberto *et al.* (1990); Schuster *et al.* (1990).
[18]Yoshinaga *et al.* (1996).
[19]Rubio *et al.* (2006).
[20]Alfonzo *et al.* (1999).
[21]Benne *et al.* (1986).
[22]Shaw *et al.* (1988).
[23]Rüdinger *et al.* (2011).
[24]Gott *et al.* (1993).
[25]Bundschuh *et al.* (2011).
[26]Mahendran *et al.* (1994).
[27]Mahendran *et al.* (1991).
[28]Antes *et al.* (1998).
[29]Gott *et al.* (2005).
[30]Gott *et al.* (2010).
[31]Lonergan and Gray (1993).
[32]LaForest *et al.* (1997).
[33]Lin *et al.* (2002).
[34]Zauner *et al.* (2004).
Abbreviations: mRNA, messenger RNA; tRNA, transfer RNA; rRNA, ribosomal RNA; ncRNA, non-coding RNA; mt, mitochondrial; cp, chloroplast.

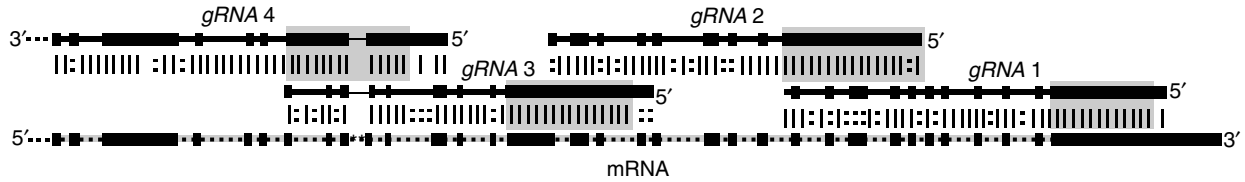


Figure 6 Editing of a gene region by four overlapping gRNAs. Thick lines in the mRNA are encoded in the mitochondrial DNA. Thin shaded lines are inserted U's; the two asterisks are deleted U's (Reproduced from Maslov DA and Simpson L (1992) The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* 70: 459–467.). Thin lines in the gRNAs are guide nucleotides (A or G) that pair with inserted U's. Vertical lines indicate Watson–Crick base pairs; colons indicate G:U wobble base pairs, illustrating formation of well-paired "anchors" between the 5′ ends of gRNAs and the corresponding region of the mRNA.

single genes to produce multiple peptides, allowing combinatorial protein diversity.

The following sections review several models that explain our current understanding of the molecular and evolutionary basis of numerous forms of RNA editing.

Kinetoplastids: Cryptogenes to Proteins by U Insertion or Deletion

Kinetoplastids are a group of unicellular protists that include the pernicious trypanosome and leishmania parasites responsible for deadly human diseases such as African sleeping sickness and Chagas disease. Translation of mitochondrial mRNAs of kinetoplastids is impossible without massive RNA editing of the transcripts by insertion and deletion of uridine (reviewed in Aphasizhev and Aphasizheva, 2011). This editing drastically rewrites the coding regions of the transcripts by introducing and fixing frameshifts as well as creating stop and start codons. In some kinetoplastid mRNAs, editing creates more than 90% of the amino acid codons. Since the DNA copies of many genes are barely recognizable as sources of their cognate mRNAs, they have been named *cryptogenes*.

Although all kinetoplastids display specific, reproducible editing patterns in certain mitochondrial mRNAs, comparison of the editing patterns in a particular gene transcript across a variety of species reveals a trend in the extent of RNA editing within each species. The earlier diverging kinetoplastids exhibit more editing within the cytochrome c oxidase III transcript than their later diverging relatives. This implies that this type of RNA editing is ancient within this lineage and that its use has decreased over evolutionary time. However, the lack of U insertion or deletion editing in any other type of organism suggests that it arose specifically within the kinetoplastid lineage. Cavalier-Smith proposed that glycosomes, special organelles found only in kinetoplastids, may provide a clue to the origin of editing in these protists. Glycosomes permit an efficient anaerobic lifestyle, during which RNA editing may have become fixed by drift in the seldom-used mitochondria of the early kinetoplastids.

Guide RNAs (gRNAs) provide the specificity of the kinetoplastid editing mechanism (Blum *et al.*, 1990). They are small RNA transcripts that guide the editing machinery through base pairing with the mRNAs in edited regions. The gRNAs are complementary to small segments of the fully edited mRNAs and thus serve as minitemplates to guide the addition and deletion of uridines from the preedited mRNA as they form a locally double-stranded RNA helix. An unpaired A or G in a gRNA signals for an editing event to insert a U into the mRNA. Complete editing of a mRNA often requires a set of many overlapping gRNAs. The affiliated set of overlapping gRNAs sequentially edit the mRNA from its 3′ end to its 5′ end. During editing, a complex of proteins called the *editosome* catalyzes the insertion and deletion of uridines until the length of each paired gRNA–mRNA region is maximized. The extent of G–U wobble pairs and conventional A–U and G–C pairs of each gRNA–mRNA pair directs the cascade of editing. Each subsequent gRNA binds the mRNA more stably than the last by possessing more Watson–Crick base pairs (A–U and G–C) in the anchor portion of the gRNA, leading to an overall 3′–5′ directionality of editing (**Figure 6**).

Guide RNAs are encoded in the kinetoplastid mitochondrial genome. Kinetoplastids are named after the unusual structure of their mitochondrial genomes, which consist of a *kinetoplast*, a network of DNA inside a single, large mitochondrion located at the base of the flagellum. In one kinetoplastid subdivision, the trypanosomatids, the kinetoplast consists of a few identical maxicircles of 20–40 kb intertwined with thousands of heterogeneous minicircles of 0.5–3 kb. The maxicircles encode the mitochondrial genes and cryptogenes for respiratory proteins, ribosomal proteins, and rRNA. The only known role of the minicircles is to encode gRNAs. Bodonids, the other group of kinetoplastids, share the presence of a DNA network with the trypanosomatids but have a different structure for their minicircle homologs. In Bodo caudatus, the *minicircles* are noncatenated 1.4-kb circles. *Trypanoplasma borreli* has minicircle-like structures of 170 and 200 kb, with tandem 1-kb repeats encoding most gRNAs.

How did the gRNAs and minicircles come to exist as they do today? A product of mitochondrial DNA recombination provides a plausible explanation. Lunt and Hyman (1997) identified a minicircle of DNA excised from the major mitochondrial genome circle in a nematode. Similar intramolecular recombination within a kinetoplastid maxicircle might have led to the formation of minicircles encoding tiny portions of mRNAs whereas other unrecombined maxicircles still contained complete functional copies of the mRNA. If the minicircle copy of a gene fragment acquired a promoter that allowed transcription of an antisense RNA, then such a gene-encoding complementary RNA could have

given rise to a proto-gRNA. If the protein equipment responsible for catalyzing RNA insertions and deletions simultaneously arose or was recruited from another function to serve in RNA editing, then mutations in the mitochondrial mRNA could have been repaired by editing. Guide RNA-containing minicircles would be selectively retained in the genome, ensuring their survival.

Myxomycetes: Multiple Editing Types in Coding and Noncoding RNAs

Physarum polycephalum is a myxomycete, or plasmodial slime mold. It takes on many shapes and sizes throughout its life, morphing from microscopic amoeba to a multinucleate syncytium that can be several feet across and then forming millimeter-scale delicate, mushroom-like fruiting bodies. Mitochondrial transcripts for almost all messenger and structural RNAs in *Physarum* and related myxomycetes require several types of RNA editing to create functional products (reviewed in Gott and Rhee, 2007). Editing is extensive in *Physarum*; deep sequencing of total plasmodial mitochondrial RNA revealed 1333 sites of editing (Bundschuh *et al.*, 2011). Most editing events involve the insertion of nonencoded nucleotides, with a total of 1347 nt added at 1324 sites. The vast majority are insertions of cytidine, but there are also a small number of specific and reproducible insertions of uridines, guanosines, an adenosine, and dinucleotides. The remaining editing sites involve three instances of A deletion, the replacement of the 5′ nucleotide of two of the mitochondrially encoded tRNAs (*see* Mitochondrial tRNA Editing), and four sites of C to U changes. A unique feature of myxomycete editing is that they are the only organisms known to combine multiple types of editing within the same transcript. For example, *P. polycephalum's* 1.5-kb transcript encoding cytochrome c oxidase subunit I (coI) undergoes 59 C insertions, a single U insertion, three dinucleotide insertions, and four sites of C→U base conversion (**Figure 7**). Insertion and base conversion editing have distinct evolutionary histories. The col gene of *Stemonitis flavogenita*, another myxomycete, shares all three types of insertional editing with *P. polycephalum* but lacks the C→U conversion editing in this transcript. Even the three types of insertional editing did not all evolve simultaneously because U insertional editing is found in all myxomycetes to date, whereas C insertion and dinucleotide insertion are found only in some slime molds.

Studies in isolated mitochondria have demonstrated that multiple editing mechanisms are at work in *P. polycephalum* mitochondria. Nucleotide insertions occur cotranscriptionally, whereas C→U changes and 5′ tRNA editing are posttranscriptional processes. Sites of nucleotide insertion are sprinkled throughout structural RNAs and in the coding regions of edited mRNAs but are never closer than nine nucleotides. Template regions required for editing fall within ~9 bp upstream of and downstream from nucleotide insertion sites, but no consensus sequences have been identified. Nucleotide insertion is remarkably accurate and efficient, with virtually all steady-state RNAs completely edited. Myxomycete insertional editing is clearly different from the posttranscriptional uridine insertion and deletion editing found in the kinetoplastids and, not surprisingly, no evidence for guide RNAs has been found by either deep sequencing or labeling studies. It is also mechanistically distinct from the cotranscriptional polymerase stuttering observed in paramyxoviruses (*see* Paramyxovirus and Ebola Virus: Polymerase Stutter Unites Reading Frames).

The sites of C→U editing in *P. polycephalum* are all in first or second codon position, similar to the case in plant organelle editing. Also similar is that the process of editing restores the conserved peptide sequence in this region so that the unedited *S. flavogenita* transcript codes for the same amino acids at these positions as does the edited *Physarum* transcript. In contrast, insertion sites are biased toward the third codon position, with roughly half of the added nucleotides found at this position.

Plant Organelles: C→U and U→C Editing Restores Conserved Amino Acids

Plant mitochondria employ rampant RNA editing, often involving hundreds of cytidine to uridine conversions (reviewed in Takenaka *et al.*, 2008; Chateigner-Boutin and Small, 2010). Plant chloroplast RNAs are also edited by C to U changes, but in higher plants these changes are much less frequent (see

CUAUUUUUAGU**C**UGCAUCUAGCUGGUGUUUCUUCUAUGUUAGGUGCUAU**C**A
AUUUCAUUUGUACCAUUAAAAAUAUGCGU**C**UUAAAGGAUUAACAGGAGAAC
GUUUAUCUUUAU**U**UGUUUGGGCUGUAUUAGUAACUGUGAUUUUAUUAUUAC
UUUCACUGCCUGU**C**UUAGCAGGUGCUAU**C**ACUAUGUUAUUAACUGAU**C**GUA
AUUUUAAUACAUCUU**U**U**U**UGAUGCAAC**C**GGUGGUGGAGAUCCUAUUUUAU
AUCAACAUUUGU**U**UUGGUUUUUUGGCCAUCCAGAAGUU**UA**CAUUUUAAUUU
UACCUGGUUUUGGUAU**C**GUUUCUAUUAUUAUUCAAGC**C**UAUGCUAAUAAAG
CUAUUUUUGGUUAUUUAGGUAUGGUGUAUGCUAUGUUGUCUAUUGGUAU**C**U
UGGGUUUUAUAGUGUGGGCUCAU**C**AUAUGUAUACUGUAGGAUUGGAUGUGG
AUA**CU**CGCGCUUAUUUCAC**C**GCUGCUACUAUGAU**C**AUUGCUGUGCCAAC**C**G

Figure 7 *RNA editing in P. polycephalum.* Partial *P. polycephalum* mRNA sequence for gene encoding cytochrome c oxidase subunit I, with underlined text indicating inserted nucleotides and boxed text indicating C→U substitution events. Reproduced from Gott JM, Visomirski LM, and Hunter JL (1993) Substitutional and insertional RNA editing of the cytochrome c oxidase subunit 1 mRNA of Physarum polycephalum. *Journal of Biological Chemistry* 268: 25483–25486.

Stern *et al.*, 2010). A limited number of plant lineages also contain uridine to cytidine changes in chloroplast or mitochondrial RNAs. Although the mechanism for plant organellar editing is not fully understood, when cytidine triphosphate (CTP) residues of *in vitro* transcripts were labeled on their α-phosphate and on their cytosine base, the labels were retained after editing. This indicates that the C→U changes occur through a deamination of cytidine rather than a base or nucleotide substitution. *In vitro* and *in organello* experiments have shown that editing signals in plant mitochondria are largely local, with essential sequences falling within ~16–20 nt upstream and 0–6 nt downstream of the edited C. The involvement of pentatricopeptide repeat proteins (PPR) in plant organellular editing was first demonstrated by Kotera *et al.* (2005), who showed that editing of a specific site within the *Arabidopsis thaliana* chloroplast ndhD gene was dependent on the PPR protein CRR4. Since that time, many other members of the extensive PPR protein family have been linked to editing of specific editing sites in mitochondria or chloroplasts (reviewed in Fujii and Small, 2011). As of yet, however, no enzymatic activity has been demonstrated for these PPR proteins, leaving open the identity of the factor(s) involved in the deamination (or potentially transamination) and amination reactions.

RNA editing is present in both the mitochondrial and the chloroplast genomes of land plants, but it is absent in green algae and the liverwort *Marchantia polymorpha*. This distribution suggests that the plant editing systems share common components and may have arisen simultaneously in both organelles or have been transferred from one to the other. Plants that possess the ability to perform base conversion may exploit the conversions as a mechanism for repairing sequences disturbed by mutational drift. The observation that most of the edited sites in plant genes are in first or second codon positions supports this hypothesis because editing produces nonsynonymous changes in amino acids. Furthermore, Malek *et al.* (1996) found that the level of mitochondrial editing in plants correlates with G/C content. Indeed, Lu and associates (1998) determined that editing "reverses" all nonsynonymous U→C substitutions at the RNA level in col genes from eight gymnosperm species, eliminating almost all variation in the predicted protein sequences. Editing can regulate translation or enzyme activity; it creates stop and start codons and proteins of varying function. However, the pressure toward loss of editing appears to be greater than any selective advantage it confers since Shield and Wolfe's (1997) comparison of edited sites over many plant species reveals a high rate of mutation of the genes toward the edited (uridine) nucleotide.

Mitochondrial tRNA Editing: Acceptor Stems and Anticodons

Mitochondrial genomes of some organisms seem pressed to economize space. "Junk" DNA is held to a minimum, and genes immediately abut their neighbors. Some tRNA genes overlap their 5′ and 3′ extremities. Base pairing of a tRNA's 5′ and 3′ ends is essential to form an acceptor stem for aminoacylation. In mitochondria, incompletely paired tRNA acceptor stems are repaired by exchanging mismatched bases on one side of the acceptor stem for ones that complement those bases present on the opposing stem (see Knoop, 2011). Thus, editing and the presence of an internal acceptor stem template allow the genome to repair mutations that may occur in response to selection for either nucleotide compositional bias or sequence compression.

In land snails, squids, and chickens, the overlapping ends of certain tRNAs disturb the crucial base pairing of the acceptor stems. RNA editing activity restores complementarity to these tRNAs by replacing mismatched 3′ guanosine, uridine, and cytidine nucleotides with adenosines, possibly by some form of polyA polymerase. In a platypus tRNA, there are three exchanged nucleotides (U→A, U→C, and A→C) in the 3′ half of the stem to complement the 5′ half of the acceptor stem, leading to the suggestion that this reaction might be carried out by a CCA-adding enzyme. The introduction of all four nucleotides in the 3′ portion of the acceptor stem is required for the maturation of mitochondrial tRNAs in the centripede *Lithobius forficatus* and the jakobid protist *Saculamonas ecuadoriensis*, suggesting a requirement for an RNA-dependent RNA polymerase in these cases. A different mechanism is required when nucleotides within the 5′ portion of the acceptor stem are replaced using the 3′ part of the acceptor stem as template. Initially described in the ameboid protist *Acanthamoeba* and the fungus *Spizellomyces*, this form of editing has also been found in the myxomycete *P. polycephalum* and is predicted to occur in many other species (**Table 1**). Editing of the *Acanthamoeba* and *Spizellomyces* tRNAs involves the removal of uridines, adenosines, and cytidines within the first three 5′ positions and replacement with purines complementary to a corresponding pyrimidine on the 3′ template side of the stem. This unusual 3′–5′ polymerization reaction is likely to be catalyzed by proteins related to Thg1, the enzyme responsible for the addition of a G_{-1} residue to the 5′ end of $tRNA^{His}$. Thg1-like proteins (or TLPs) from *Dictyostelium discoideum* have recently been shown to accurately repair truncated tRNA editing substrates (Abad *et al.*, 2011).

Mitochondrial tRNAs are also subject to A-to-I and C-to-U editing events (reviewed in Paris *et al.*, 2012). An extreme example of tRNA C to U editing is provided by the $tRNA^{Pro}$ from the lycophyte *Isoetes engelmannii*, which differs from its mitochondrial template at 18 different sites (Grewe *et al.*, 2009). More commonly, editing changes the decoding properties of a tRNA by affecting the anticodon loop.

In opossums and other marsupials, the DNA encoding mitochondrial $tRNA^{Asp}$ has the wrong anticodon. Although the rest of the tRNA has a canonical $tRNA^{Asp}$ sequence, the anticodon, GCC, will pair with glycine codons rather than aspartic acid codons. About half of the $tRNA^{Asp}$ transcripts are not edited and consequently are aminoacylated with glycine. The remainder are edited by a C→U change that restores the GUC aspartic acid anticodon. These are charged correctly with aspartic acid since this anticodon is used as a determinant of tRNA charging. Surprisingly, the standard $tRNA^{Gly}$ (with a UCC anticodon) is present in the genome as well, and it could theoretically recognize all four GGN glycine codons by wobble. However, a mutation just outside of the anticodon region of this $tRNA^{Gly}$ renders it incapable of recognizing the two glycine codons which the mutated $tRNA^{Asp}$ can bind. A second example of anticodon editing occurs in trypanosomatid $tRNA^{Trp}$. In trypansomes, all tRNAs are imported from the

cytoplasm. To decode the UGA codon as tryptophan in the mitochondria (but not the cytoplasm), a portion of the imported $tRNA^{Trp}$ is edited within the anticodon, which results in a change from CCA to UCA.

Böerner and Pääbo (1996) suggest that marsupial mitochondrial editing became fixed through two mutational steps. First, a mutation occurred in the anticodon for $tRNA^{Asp}$, transforming it into a functional $tRNA^{Gly}$. Genomes with this mutation were still viable since the altered tRNA could regain its necessary function as $tRNA^{Asp}$ through editing. A subsequent mutation in the original $tRNA^{Gly}$ was not lost because of redundancy in $tRNA^{Gly}$ activity: a mutation outside the anticodon of the original $tRNA^{Gly}$ eventually left it unable to recognize the two glycine codons that the newly mutated $tRNA^{Asp/Gly}$ could translate. After this second mutation in the original $tRNA^{Gly}$, back mutation of the $tRNA^{Gly/Asp}$ anticodon (to form simply a $tRNA^{Asp}$ anticodon) would be deleterious since the mutant $tRNA^{Gly/Asp}$ serves double duty by pairing with both aspartic acid and glycine codons. Thus, RNA editing permitted a deleterious change in a genome and then became fixed when a second mutation made editing a requirement for expression of mitochondrial proteins.

Human Apo B and Beyond: Shorter Peptides and Regulatory Possibilities

Apolipoprotein B is present in two different forms in both humans and rodents. The long form, apo B100, is part of very-low-density lipoprotein particles that have a role in cholesterol metabolism, whereas the short form, apo B48, contributes to chylomicrons that transport dietary lipids. The two forms of the apo B protein are actually encoded by a single gene with mRNA that undergoes tissue-specific C→U RNA editing at nucleotide position 6666. The editing converts an encoded glutamine codon into a stop codon to create the short form of the peptide (reviewed in Blanc and Davidson, 2010).

Editing of the apoB mRNA requires both a cytidine deaminase and an auxiliary factor. The catalytic peptide is apo B RNA editing cytidine deaminase subunit 1 (APOBEC-1), whereas site recognition is conferred by an RNA-binding protein, ACF (APOBEC-1 complementing factor). Specificity of the cytidine deamination is determined by the primary sequence of the apo B RNA transcript, particularly an 11-base *mooring sequence*, located just five nucleotides downstream of the edited site. In addition to this required mooring sequence, three other *efficiency* elements are found in the 140-base region surrounding the edited cytidine. These upstream and downstream elements increase the effectiveness of the editing reaction.

Although the apoB transcript is the only known APOBEC-1 target within coding regions in normal cells, a recent transcriptome-wide comparison of intestinal mRNAs from wild type and apobec-1 deficient mice identified another 32 APOBEC-1 targets within 3′ untranslated regions (3′UTRs). The majority of these new sites occur near sequence elements with significant homology to the apoB mooring sequence, making it likely that these C to U changes are catalyzed by the same enzymatic machinery (Rosenberg *et al.*, 2011). Although the functionality of these editing events has not been demonstrated as yet, most fall within conserved noncoding regions, suggesting possible roles in message stabilization, translational efficiency, and localization.

Cytidine deamination may also play a role in human tumorogenesis. For example, an imperfect APOBEC-1 mooring sequence and efficiency elements are present within the coding region of the neurofibromatosis type I (NF1) tumor suppressor gene. Despite slight differences in position and sequence context, normal individuals exhibit very low levels of editing. In comparison, tumor tissues from NF1-affected individuals showed more than eight times the normal quantity of edited transcript. The NF1 gene encodes the neurofibromin protein, which is a putative homolog of yeast proteins in the ras signal transduction pathways. The proposed GTPase activating domain of the neurofibromin lies just downstream of the NF1 editing site, and indeed C→U RNA editing at the site transforms an arginine codon into a premature translation stop upstream of the domain. Thus, the editing of the NF1 transcript most likely cripples the tumor-suppressing activity of neurofibromin.

APOBEC-1 is part of a family of cytidine deaminases. The related activation-induced cytidine deaminase (AID) also works at the polynucleotide level, but its substrate is DNA rather than RNA. AID is involved in antibody diversification, playing roles in both somatic hypermutation (SHM) and class switch recombination (CSR) via the deamination of genomic cytidines to uridines. A third set of related proteins, the APOBEC3 subfamily, is involved in the suppression of retroelements and restriction of retroviral replication.

A→I Deamination: Multiple Peptides, Multiple Regulatory Possibilities

Deamination of another type alters additional RNA sequences in a range of metazoans. Removal of adenosine's amino group yields inosine, a nucleotide that the translation machinery reads as guanosine (reviewed in Nishikura, 2010). A number of the initially identified A→I transitions cause predicted amino acid changes, most frequently in neurotransmitter receptors and ion channels (reviewed in Pullirsch and Jantsch, 2010). Editing at most sites is not complete, resulting in the generation of multiple protein isoforms within the cell. Editing events within glutamate receptors, for example, adjust the calcium channel permeability and the speed of the desensitization response. Likewise, in the serotonin receptors, protein loops encoded by the edited region of the RNA interact with G proteins to turn on signaling pathways. The editing, along with alternative splicing, allows the exquisite fine-tuning of neural responses by increasing combinatorial protein diversity available from a single transcript. Posttranscriptional regulation demands the expenditure of less time and energy than transcription of multiple gene sequences.

More recently, transcriptome-wide and bioinformatics studies have revealed many thousands of additional sites of A to I editing in humans, nearly all within introns and 3′ UTRs (e.g., Li *et al.*, 2009). The majority of these changes fall within repetitive elements (especially Alu elements) capable of generating long, partially double-stranded regions of RNA, a key feature of editing targets (see next paragraph). Possible functions of editing within these noncoding regions include alterations in miRNA–mRNA interactions (potentially altering

the level of translation), creation or elimination of splice sites, and changes in nuclear retention or stability.

A family of enzymes called ADARs (adenosine deaminase acting on RNA) is responsible for A→I editing activity. The largely double-stranded RNA regions required for ADAR action within coding regions are elegantly provided by pairing of the exonic sequences with intronic sequences in the pre-mRNAs. Neither mature mRNAs whose introns have been removed by splicing nor pre-mRNAs with mutations disturbing intron–exon complementarity act as substrates for editing in transfected cell lines. Mice genetically disabled for A→I editing at a single site demonstrate the importance of RNA editing in glutamate receptors. Brusa *et al.* (1995) removed an intron region crucial for creating a double-stranded editing substrate from the mouse GluR-B gene. Heterozygotes for this editing-disabled allele displayed severe epileptic seizures and premature death within 3 weeks of birth. The decrease in edited mRNA levels led to a fivefold increase in the calcium flow into their neurons and serious damage to their nervous systems.

At least three different ADAR enzymes are expressed differentially in human tissues, and they display discrete editing abilities at various mRNA sites *in vitro*. Mice engineered to lack ADAR2 demonstrate a phenotype similar to the GluR-B intron deletion described previously, but these mice can be "rescued" by the introduction of a GluR-B gene that is edited at the Q/R site, indicating that this target is acted on by ADAR2. ADAR3 knockout mice appear normal, but ADAR1 deficient mice are not viable, displaying widespread apoptosis and defects in hematopoiesis.

Scott (1995) suggested that ADARs might have evolved for an antiviral function, destabilizing RNA genomes by modification and subsequent unwinding. Interestingly, a devious viral genome coopted this function and uses host adenosine deaminases to its own benefit. Hepatitis delta virus (HDV), a single-stranded RNA subvirus of hepatitis B, has only one known gene product, the delta antigen. The short form of delta antigen stimulates replication of the genome, whereas the long form of the antigen suppresses replication and is necessary for packaging HDV. The short and long forms of the delta antigen differ only by a single A→I base change of the negative strand, or antigenome. The base conversion replaces a stop codon at the end of the shorter peptide's mRNA with a tryptophan codon to allow read through of the longer peptide. *In vitro*, ADAR1 is capable of inducing this A→I change on the antigenome of the HDV virus, suggesting that either this or a related enzyme edits HDV *in vivo*. In exquisite control of the reaction, the long-form delta antigen acts as a repressor of editing in a negative feedback loop. The repression prevents accumulation of unduly high levels of long-edited antigen, which would lower the level of viral replication.

Paramyxoviruses and Ebola Virus: Polymerase Stutter Unites Reading Frames

Two other types of viruses also capitalize on RNA editing processes as a method of increasing information storage while under pressure for genome size constraint. The RNA genomes of the paramyxoviruses have several overlapping genes. Some of these genes are activated for expression by ribosomal choice, whereas others become available by cotranscriptional RNA editing. The polymerase may slip as it travels through a purine run, adding from one to six extra guanosines, depending on the sequence present in the particular virus species. The stuttering polymerase thereby fuses coding regions together to make extended versions of genes.

Many viruses use an apparently similar mechanism to add poly A tails to mRNA transcripts. When the RNA polymerase reaches the U-rich regions, it adds additional nontemplated adenines to create the tails. In paramyxoviruses, RNA editing activity may be related to a process that corrects genome length, maintaining length in multiples of six nucleotides. The substrate for the RNA polymerase is the RNA genome complexed with the capsid in hexamer-length segments. Hausmann and colleagues (1996) note that if the genome length is not a multiple of six, then the polymerase inserts or deletes guanosines or adenines from the same region in which the RNA editing occurs.

Another RNA virus, the infamous Ebola virus, uses a comparable mechanism to insert adenines into a sequence to unite two reading frames. In this case, the sequence of the site where adenine addition occurs appears similar to viral poly A addition sites. In contrast to the paramyxovirus editing, the additions occur even when the RNA is produced by a non-native polymerase and thus must be intrinsic to the template sequence or structure. T7 transcription of the Ebola mRNA *in vitro* results in edited product, although in a lower quantity than when transcribed by Ebola's RNA polymerase.

Thus, once considered a molecular anomaly unique to protists, RNA editing is now recognized as a vital part of gene expression in a wide distribution of eukaryotes and their viruses. Organisms can alter RNA sequences through many different mechanisms by recruiting enzymes such as deaminases, nucleases, ligases, and special polymerases to specific RNA editing tasks. Editing clearly arose multiple times in evolutionary history in the various forms of both insertional and substitutional editing, all of which profoundly affect the expression of RNAs, the products they form, and the biological processes in which they participate. RNA editing is significant in cancers, cholesterol regulation, and neural function in vertebrates as well as in the *de novo* creation of coding sequences from obscured mitochondrial genes.

Furthermore, the impact of editing on the genomes that employ it is astounding. Genomes that use editing may become increasingly lenient toward persistence of deleterious mutations. Large amounts of genetic information and control devices may be confined to small spaces. Lastly, RNA editing, like gene scrambling, establishes a device for generating combinatorial sequence diversity.

Appendix

List of Courses

1. Genetics
2. Microbial Diversity

See also: Diversity, Molecular Level. Genes, Description of. Genetic Diversity

References

Abad MG, Long Y, Wilcox A, Gott JM, Gray MW, and Jackman JE (2011) A role for tRNA(his) guanylyltransferase (Thg1)-like proteins from Dictostelium discoideum in mitochondrial 5′-tRNA editing. *RNA* 17: 613–623.

Alfonzo JD, Blanc V, Esfevez AM, Rubio MA, and Simpson L (1999) C to U editing of the anticodon of imported mitochondrial tRNA(Trp) allows decoding of the UGA stop codon in *Leishmania tarentolae*. *EMBO Journal* 18: 7056–7062.

Antes T, Costandy H, Mahendran R, Spottswood M, and Miller D (1998) Insertional editing of mitochondrial tRNAs of Physarum polycephalum and Didymium nigripes. *Molecular and Cellular Biology* 18: 7521–7527.

Aphasizhev R and Aphasizheva I (2011) Mitochondrial RNA processing in trypanosomes. *Research in Microbiology* 162: 655–663.

Benne R, Van Den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, and Tromp MC (1986) Major transcript of the frame shifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* 46: 819–826.

Binder S, Marchfelder A, and Brennicke A (1994) RNA editing of tRNA(Phe) and tRNA(Cys) in mitochondria of Oenothera berteriana is initiated in precursor molecules. *Molecular & General Genetics* 244: 67–74.

Blanc V and Davidson NO (2010) Apobec-1 mediated RNA editing. *Wiley Interdisciplinary Reviews Systems Biology and Medicine* 2: 594–602.

Blum B, Bakalara N, and Simpson L (1990) A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* 60: 189–198.

Börner GV and Pääbo S (1996) Evolutionary fixation of RNA editing. *Nature* 383: 225. http://www.ncbi.nlm.nih.gov/pubmed/8805696

Brusa R, Zimmermann F, Koh DS, *et al.* (1995) Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science* 270: 1677–1680.

Bundschuh R, Altmuller J, Becker C, Nurnberg P, and Gott JM (2011) Complete characterization of the edited transcriptome of the mitochondrion of *Physarum polycephalum* using deep sequencing of RNA. *Nucleic Acids Research* 39: 6044–6055.

Burger G, Yan Y, Javadi P, and Lang B (2009) Group I-intron trans-splicing and mRNA editing in the mitochondria of placozoan animals. *Trends in Genetics* 25: 381–386.

Chateigner-Boutin A-L and Small I (2010) Plant RNA editing. *RNA Biology* 7: 213–219.

Chen S, Habib G, Yang C, *et al.* (1987) Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science* 238: 363–366.

Covello PS and Gray MW (1989) RNA editing in plant mitochondria. *Nature* 341: 662–666.

Doria M, Neri F, Gallo A, Farace MG, and Michienzi A (2009) Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection. *Nucleic Acids Research* 37: 5848–5858.

Fujii S and Small I (2011) The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytologist* 191: 37–47.

Gott JM, Parimi N, and Bundschuh R (2005) Discovery of new genes and deletion editing in Physarum mitochondria enabled by a novel algorithm for finding edited mRNAs. *Nucleic Acids Research* 33: 5063–5072.

Gott JM and Rhee AC (2007) Insertion/deletion editing in *Physarum polycephalum*. In: Göringer HU (ed.) *RNA Editing*. Berlin: Springer.

Gott JM, Somerlot BH, and Gray MW (2010) Two forms of RNA editing are required for tRNA maturation in Physarum mitochondria. *RNA* 16: 482–488.

Gott JM, Visomirski LM, and Hunter JL (1993) Substitutional and insertional RNA editing of the cytochrome c oxidase subunit 1 mRNA of Physarum polycephalum. *Journal of Biological Chemistry* 268: 25483–25486.

Grewe F, Viehoever P, Weisshaar B, and Knoop V (2009) A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte *Isoetes engelmannii*. *Nucleic Acids Research* 37: 5093–5104.

Gualberto JM, Lamattina L, Bonnard G, Weil J-H, and Grienenberger J-M (1989) RNA editing in wheat mitochodria results in the conservation of protein sequences. *Nature* 341: 660–662.

Gualberto JM, Weil J-H, and Grienenberger J-M (1990) Editing of the wheat coxIII transcript: evidence for twelve C to U and one U to C conversions and for sequence similarities around editing sites. *Nucleic Acids Res* 18: 3771–3776.

Hausmann S, Jacques JP, and Kolakofsky D (1996) Paramyxovirus RNA editing and the requirement for hexamer genome length. *RNA* 10: 1033–1045.

Hiesel R, Wissinger B, Schuster W, and Brennicke A (1989) RNA editing in plant mitochondria. *Science* 246: 1632–1634.

Hoch B, Maier R, Appel K, Igloi G, and Kössel H (1991) Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* 353: 1780180.

Hoffman DC and Prescott DM (1997) Evolution of internal eliminated segments and scrambling in the micronuclear gene encoding DNA polymerase alpha in two Oxytricha species. *Nucleic Acids Research* 25: 1883–1889.

Janke A and Pääbo S (1993) Editing of a tRNA anticodon in marsupial mitochondria changes its codon recognition. *Nucleic Acids Research* 21: 1523–1525.

Jühling F, Mörl M, Hartmann R, Sprinzl M, Stadler P, and Pütz J (2009) tRNAdb 2009: Compilation of tRNA sequences and tRNA genes. *Nucleic Acids Research* 37: d159–d162.

Knoop V (2011) When you can't trust the DNA: RNA editing changes transcript sequences. *Cellular and Molecular Life Sciences* 68: 567–586.

Kotera E, Tasaka M, and Shikanai T (2005) A pentatricopeptie repeat protein is essential for RNA editing in chloroplasts. *Nature* 433: 326–330.

Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, and Yoshinaga K (2003) RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Research* 31: 2417–2423.

Laforest M, Roewer I, and Lang B (1997) Mitochondrial tRNAs in the lower fungus spizellomyces punctatus: tRNA editing and UAG 'stop' codons recognized as leucine. *Nucleic Acids Research* 25: 626–632.

Li JB, Levanon EY, Yoon J-K, *et al.* (2009) Genome-wide identiification of human RNA editing sites by parallel DNA capturing and sequencing. *Science* 324: 1210–1213.

Li M, Wang I, Li Y, *et al.* (2011) Widespread RNA and DNA sequence differences in the human transcriptome. *Science* 333: 53–58.

Lin S, Zhang H, Spencer D, Norman J, and Gray MW (2002) Widespread and extensive editing of mitochondrial mRNAs in dinoflagellates. *Journal of Molecular Biology* 320: 727–739.

Lonergan KM and Gray MW (1993) Editing of transfer RNAs in acanthamoeba castellanii mitochondria. *Science* 259: 812–816.

Luciano DJ, Mirsky H, Vendetti NJ, and Maas S (2004) RNA editing of a miRNA precursor. *RNA* 10: 1174–1177.

Luo G, Chao M, Hsieh S-Y, Sureau C, Nishikura K, and Taylor J (1990) A specific base transition occurs on replicting hepatitis delta virus RNA. *Journal of Virology* 64: 1021–1027.

Mahendran R, Spottswood MS, Ghate A, Ling ML, Jeng K, and Miller DL (1994) Editing of the mitochondrial small subunit rRNA in physarum polycephalum. *The EMBO Journal* 13: 232–240.

Mahendran R, Spottswood MR, and Miller DL (1991) RNA editing by cytidine insertion in mitochondria of physarum polycephalum. *Nature* 349: 434–438.

Malek O, Lättig K, Hiesel R, Brennicke A, and Knoop V (1996) RNA editing in bryophytes and a molecular phylogeny of land plants. *The EMBO journal* 15: 1403–1411.

Maréchal-Drouard L, Ramamonjisoa D, Cosset A, Weil J, and Dietrich A (1993) Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. *Nucleic Acids Research* 21: 4909–4914.

Maslov DA and Simpson L (1992) The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* 70: 459–467.

Nishikura K (2010) Functions and regulation of RNA editing by ADAR deaminases. *Annual Review of Biochemistry* 79: 321–349.

Nowacki M, Haye JE, Fang W, Vijayan V, and Landweber LF (2010) RNA-mediated epigenetic regulation of DNA copy number. *Proceedings of the National Academy of Sciences* 107: 22140–22144.

Nowacki M, Vijayan V, Zhou Y, Schotanus K, Doak TG, and Landweber LF (2008) RNA-mediated epigenetic programming of a genome-rearrangement pathway. *Nature* 451: 153–158.

Paris Z, Fleming IMC, and Alfonzo JD (2012) Determinants of tRNA editing and modification: Avoiding conundrums, affecting function. *Seminars in Cell and Developmental Biology* 23: 269–274.

Powell L, Wallis S, Pease R, Edwards Y, Knott T, and Scott J (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell* 50: 831–840.

Prescott DM (1994) The DNA of ciliated protozoa. *Microbiology and Molecular Biology Reviews* 58: 233–267.

Pullirsch D and Jantsch MF (2010) Proteome diversification by adenosine to inosine RNA-editing. *RNA Biology* 7: 205–212.

Rosenberg BR, Hamilton CE, Mwangi MM, Dewell S, and Papavasiliou FN (2011) Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3′ UTRs. *Nature Structural & Molecular Biology* 18: 230–238.

Rubio MAT, Ragone FL, Gaston KW, Ibba M, and Alfonzo JD (2006) C to U editing stimulates A to I editing in the anticodon loop of a cytoplasmic threonly tRNA in *Trypanosoma brucei*. *Journal of Biological Chemistry* 281: 115–120.

Rüdinger M, Fritz-Laylin L, Polsakiewicz M, and Knoop V (2011) Plant-type mitochondrial RNA editing in the protist *Naegleria gruberi*. *RNA* 17: 2058–2062.

Schuster W, Hiesel R, Wissinger B, and Brennicke A (1990) RNA editing in the cytochrome b locus of the higher plant Oenothera berteriana includes a U-to-C transition. *Molecular and Cellular Biology* 10: 2428–2431.

Scott J (1995) A place in the world for RNA editing. *Cell* 81: 833–836.

Shaw JM, Feagin JE, Stuart K, and Simpson L (1988) Editing of the kinetoplastid mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG initiation codons. *Cell* 53: 401–411.

Sommer B, Kohler M, Sprengel R, and Seeburg PH (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* 67: 11–19.

Stern DB, Goldschmidt-Clermont M, and Hanson MR (2010) Chloroplast RNA metabolism. *Annu. Rev. Plant Biol.* 61: 125–155.

Takenaka M, Verbitskiy D, van der Merwe JA, Zehrmann A, and Brennicke A (2008) The process of RNA editing in plant mitochondria. *Mitochondrion* 8: 35–46.

Thomas SM, Lamb RA, and Paterson RG (1988) Two mRNAs that differ by two nontemplated nucleotides encode the amino co-terminal proteins P and V of the paramyxovirus SV5. *Cell* 54: 891–902.

Volchkov VE, Becker S, Volchkova VA, *et al.* (1995) GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. *Virology* 214: 421–430.

Yokobori S-I and Pääbo S (1995a) Transfer RNA editing in land snail mitochodria. *Proceedings of the National Academy of Sciences of the USA* 92: 10432–10435.

Yokobori S-I and Pääbo S (1995b) tRNA editing in metazoans. *Nature* 377: 490.

Yoshinaga K, Iinuma H, Masuzawa T, and Uedal K (1996) Extensive RNA editing of U to C in addition to C to U substitution in the rbcL transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. *Nucleic Acids Research* 24: 1008–1014.

Zauner S, Grellinger D, Laatsch T, Kowallik KV, and Maier UG (2004) Substitutional editing of transcripts from genes of cyanobacterial origin in the dinoflagellate Ceratium horridum. *FEBS Letters* 577: 535–538.

O

Ocean Ecosystems

Richard T Barber, Duke University, NC, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 427–437, © 2001, Elsevier Inc.

Glossary

Convective mixing Vertical mixing produced by the increasing density of a fluid in the upper layer, especially during winter in temperate and polar regions.
Euphotic zone The surface layer of the ocean in which there is adequate light for net positive photosynthesis.
Nutrients Dissolved mineral salts necessary for primary productivity and phytoplankton growth: Macronutrients are phosphate, nitrate, and silicate; micronutrients are iron, zinc, manganese, and other trace metals.
Phytoplankton Photosynthetic, usually single-celled, plants that drift with ocean currents.
Pycnocline The layer in which density changes most rapidly with depth and separates the surface mixed layer from the deep ocean waters.
Southern Ocean The circumpolar ocean in the Southern Hemisphere between the Subtropical Front and the continent of Antarctica.
Stratification The formation of distinct layers with different densities; stratification inhibits mixing.
Subpolar Pertaining to the regions between the polar and temperate zones, but for the oceans the boundaries are the Subtropical Front and the Polar Front.
Subtropical Pertaining to the regions which, under the influence of the trade winds, are permanently stratified.
Upwelling Upward vertical movement of water through the bottom of the surface mixed layer produced by a divergence at the surface.
Zooplankton Animals that float or drift with ocean currents: Microzooplankton are protozoan plankton that graze on small phytoplankton; mesozooplankton are crustaceans that graze on larger phytoplankton such as diatoms.

History of the Ecosystem Concept

Natural scientists have long recognized that beyond populations or communities there is a higher level of organization, but it has been difficult for physical and biological scientists to reach consensus about its nature. The origins of the ecosystem concept are indistinct and, indeed, involved polemics between some of the principals. Ernst Haeckel (1838–1919) coined the word "ecology" and he envisioned ecology as a science that studies the environment as a stage within which selection pressures and adaptations affect the evolution of species. Haeckel brought the physical environment to the forefront, but he was interested primarily in evolution, which acts on species or genetically interacting populations. Victor Hensen (1835–1924) coined the word "plankton" to describe the organisms that drift in the ocean following the paths dictated by currents and mixing. Hensen was interested in the integrated working of natural systems, the subject today called "ecosystem science." He was particularly interested in the functional behavior of ocean systems from the smallest phytoplankton up through the food web and leading finally to fish, birds, and marine mammals. Hensen and Haeckel had a long professional dispute which seems ironic today because the eventual synthesis of their ideas by others led indirectly to the ecosystem concept.

The linguistic route to the modern word "ecosystem" was tortuous. From 1877 to 1939 the words "biocoenose," "microcosm," "naturkomplex," "holocen," and "biosystem" were all proposed as names for this idea. In 1935 an English scientist, Tansley, introduced the neologism "ecosystem" and this term rapidly eclipsed competing names, but ecosystem science was not a well-organized endeavor until 1959, when Eugene Odum and Howard Odum published a groundbreaking textbook, *Fundamentals of Ecology*. There were many treatises on ecology before the Odums' book, but their book stressed the principle that the ecosystem (not the population or community) is the fundamental unit of ecology and indeed of biology. The 1959 text, which develops this concept in each of its chapters, has remained in print for many years. Although Eugene Odum's definition of ecosystem (given at the beginning of this article) is widely accepted, the principle of the ecosystem as the

fundamental unit of ecology is controversial. Populations, communities, adaptation, and evolution have all been explosively successful research areas. Ecosystem research, because of its inherent complexity and requirement for interdisciplinary work, has not enjoyed the success of the more reductionist areas of ecological research. The study of ecosystems is usually supported by encouraging words, but our institutions and agencies, which are organized along disciplinary lines, have trouble coping with a science composed of equal parts of physics, chemistry, and biology. Ecosystem studies are placed with the life sciences, in which there is little institutional or agency enthusiasm for the complex and expensive atmospheric and physical oceanographic work required by the study of ocean ecosystems.

Utility of the Concept

The ecosystem concept was originally developed for terrestrial, intertidal, and benthic habitats, in which ecological succession on timescales of 10–100 years is an obvious and salient characteristic of the ecosystem. Succession on this timescale is difficult or impossible to detect in the fluid medium of pelagic ocean ecosystems. In the ocean, winds, mixing, and currents appear to reset the successional time clock to time zero each annual cycle or, perhaps, with each storm or passage of a front or large eddy. It has been shown that there are low-frequency (over the timescale of 10–100 years) changes in species abundance and community structure. These biological changes are sometimes so far reaching that they are called "regime shifts," the term used by John Steele in 1998, but they are not functionally analogous to succession, which is (i) orderly and directional, (ii) involving biological modification of the physical environment, and (iii) resulting in a more stable climax community (Odum, 1969). Succession in ocean ecosystems is dramatically evident on the day to month timescale after spring stratification in temperate and subpolar waters or after an episode of upwelling. This short timescale succession, which appears to be cyclic, is not unidirectional and definitely not orderly.

If long-term directional succession does not appear to occur in pelagic ocean ecosystems and succession is a central tenet of ecosystem theory, isn't the very existence of ocean ecosystems in doubt? The answer is that ocean ecosystems, as Steele (1985) noted, are clearly different from terrestrial, aquatic, intertidal, or benthic ecosystems, but ocean ecosystems meet most of the requirements of the definition set forth by Eugene Odum. The ocean ecosystems described here all have a characteristic and distinct trophic structure, characteristic and distinct material cycles, some degree of internal homogeneity and commonality, and definable hydrographic boundaries.

There is considerable heuristic power in the ecosystem concept because understanding gained in one ocean ecosystem can be used to predict the response of another ecosystem of the same kind that is geographically distinct from it. This predictive power is perhaps the greatest benefit of ecosystem theory and provides evidence that each distinct kind of ocean ecosystem has characteristics that can be generalized and used in prediction.

Pomeroy and Alberts (1988) emphasized that the concept involves emergence of new properties. One consequence of the hierarchic organization of ecosystems is that as components (both biotic and abiotic) are integrated into a larger functional unit, new properties emerge that cannot be detected by study of the component populations or processes, no matter how thorough the reductionist study. This aspect of ecosystem theory makes the concept useful, even necessary, for predictive understanding of ocean ecosystems.

Determining specific emergent properties of a system as large as ocean ecosystems is difficult. Ocean ecosystem spatial domains of thousands of kilometers are difficult to sample adequately with ships; however, with satellites that measure wind, ocean temperature, ocean currents (from sea surface topography), and phytoplankton biomass, a new era has begun in understanding ocean ecosystems. Are there appropriate benefits to justify the large societal investment in remote sensing and data handling required to achieve this new level of understanding? One benefit deals with fisheries management. The historic approach to management of these fisheries has involved analysis of local populations and environment. It has become clear that ecosystem properties, especially physical conditions, operating on scales much larger than the range of the exploited population, may be responsible for changes in reproductive success and adult abundance. In this context, a valuable societal payoff of understanding ocean ecosystems is improved ability to predict local variations in living resources as shown by Sherman and others in 1986. Odum repeatedly emphasized that ecosystem science and economics are parallel disciplines and expressed regret that they are not perceived as such, particularly by the economic community. Investors and political leaders making policy decisions on resource development need an understanding of the probability, frequency, and intensity of natural variability for realistic economic decision making. In this context, economics and the ecosystem concept are closely related: By understanding the variability of ocean ecosystems, decision makers will also understand that variation in marine resources is a normal and inevitable characteristic that must be accommodated in economic plans. This kind of resource-related benefit, although important, is not the only societal benefit that will accrue.

Odum (1977) said, "It is the properties of the large-scale integrated systems that hold solutions to most of the long-range problems of society." Although few in the scientific community recognized the wisdom of this comment two decades ago, the validity of Odum's prediction has now been well demonstrated. For example, carbon dioxide modification of the planet's heat budget, and therefore climate, is a phenomenon that can be understood only if the emergent properties of large-scale biogeochemical systems are understood.

Partitioning the Ocean into Natural Functioning Units

Central Problem

The central problem for the lower trophic level of ocean ecosystems is obtaining light (energy) and inorganic nutrients (mass). Odum's definition requires that for a functioning system to be a distinct ecosystem it must possess characteristic trophic structure and material cycles. That is, how one kind of ocean ecosystem captures light and passes that energy on in

the form of primary productivity, secondary productivity, and so forth is different from how another kind of ocean ecosystem processes and transfers its energy. Likewise, how mass (C, N, P, and Si), initially in the form of inorganic compounds, is taken up and transferred through the food web and eventually released back to the environment is different and very poorly understood.

What controls the supply of light and nutrients to an ocean ecosystem? Sverdrup in 1955 was the first oceanographer to note that the spatial and temporal patterns of physical processes, particularly the seasonal patterns of mixing, stratification, and upwelling as well as the seasonal pattern of irradiance, control the patterns of biological organization. The division of ocean ecosystems into six distinct types is based fundamentally on the pioneering work of Sverdrup and that of others in the intervening years.

Biome Concept

Longhurst (1998) presented a scholarly discussion of the attempt to partition the oceans into natural functional provinces. He described with considerable elan a regional ecology of the oceans which is clearly distinct from partitioning the ocean into ecosystems. The difference can be illustrated using one of the best described ocean ecosystems, the coastal upwelling ecosystem. It is well accepted that there is one coastal upwelling ecosystem and that it is replicated at five coastal regions in the world ocean: the California Current off California and Mexico, the Peru Current off Peru and Chile, the Benguela Current, the Canary Current, and the Arabian Sea off Oman and Yemen. Longhurst recognized the same regions but counted them as five distinct ocean "provinces." The ecosystem concept emphasizes that there are geographic ocean regions that have much in common; ecological geography emphasizes the discreteness or autonomy of various ocean provinces. Both concepts are useful.

Longhurst (1998) proposed that the following information is required to define the functional nature of an ocean region:

Latitude, which determines if wind stress induces mixing or wind-driven advection

Depth of water because stratification may be broken down in shallow seas by tidal mixing

Proximity to coastline, which determines the effects of terrestrial runoff, river discharge, and release of nutrients (especially Fe) from sediments

Seasonal irradiance, which forces photosynthesis, stratification, and freezing or melting of ice

Winds, which force mixing or upwelling of subsurface waters and their nutrients up to the euphotic zone

Precipitation, which may induce strong stratification by making the surface layer less salty

Nutricline depth, which modifies the vertical flux of nutrients by wind mixing and upwelling

Strength of the vertical nutrient gradient, which determines the magnitude of the upward nutrient flux

External source of iron, because insufficient iron may limit uptake of macronutrients, phytoplankton growth, or primary productivity

Using these criteria, Longhurst partitioned the ocean into four major biomes: the westerlies biome, in which convective mixing and stratification are forced largely by the strong seasonal progression of winds, irradiance, and heat flux; the trades biome, in which upwelling and mixing are forced on the ocean-basin scale by both local and remote wind forcing; the polar biome, in which there is no significant thermal stratification and mixed layer depth is constrained by a surface low-salinity layer which forms each summer as the marginal ice melts; and the coastal biome, in which coastal processes such as tides or currents break down stratification and force mixing.

Longhurst's (1998) partitioning of the ocean into four biomes and 55 provinces is a significant accomplishment. Eventually, his ideas will be merged with the ecosystem concept to produce an internally consistent hierarchy of biomes, ecosystems, and provinces.

Ecosystem Concept

Using the formal Odum definition of ecosystem, six more or less well-defined ocean ecosystems can be delineated (**Table 1**), but note that the boundaries of these ecosystems are usually oceanographic, not geographic, features. (See Tomczak and Godfrey (1994) for a description of these oceanographic features). Although it is possible to describe approximately where these ecosystems exist, the actual domain is determined by dynamic processes such as fronts where two kinds of ocean water converge. The approach is to define the characteristic trophic structure and material cycles by applying the same analysis to each system. Each of the six systems has a different combination of stratification, nutrient supply (**Table 1**), primary productivity (**Table 2**), and biotic characteristics (**Table 3**). This analysis indicates that there are six distinct ocean systems that meet Odum's criteria:

1. A low-latitude gyre ecosystem is present in each of the five great low-latitude gyres—North Atlantic, South Atlantic, North Pacific, South Pacific, and South Indian Oceans—as well as in the warm pool of the western Pacific Ocean, the equatorial Indian Ocean, and in large marginal seas such as the Mediterranean Sea and the Gulf of Mexico. This ecosystem is also present in the Western Boundary Current regions between the western edge of each of the five great gyres and the coastal waters of the adjacent continent.
2. The Southern Ocean ecosystem occupies the circumpolar area between the continent of Antarctica and the Subtropical Front at approximately 40° S latitude; the Southern Ocean ecosystem has a subantarctic region from the Subtropical Front south to the Polar Front and an Antarctic region from the Polar Front to the Antarctic continent.
3. The equatorial upwelling ecosystem occupies an equatorial band from 5° N to 5° S and from South America westward to 180° in the eastern and central Pacific Ocean. This region is often called the "cold tongue" because upwelling keeps these tropical waters surprisingly cool. In the Atlantic the 5° N to 5° S band of equatorial upwelling reaches from Africa across to South America. In the Indian

Table 1 Summary of heat flux, stratification, nutrient, and light characteristics of ocean ecosystems

Oceanecosystem[a]	*Heat flux*	*Stratification*		*Nutrient*		*Light*	
		Strength	*Duration*	*Level*	*Source*	*Level*	*Pattern*
Low-latitude gyre	Neutral; negative in Western Boundary Current	Strong	Permanent	Low ($\ll K_s$)[b]	Eddy diffusion, very weak convection	High ($\gg E_k$)[c]	Continuous
Southern Ocean	Negative (ocean loses heat)	Very weak, except strong when ice melts in summer	Seasonal	High ($> K_S$), except $Si(OH)_4$	Mixing and upwelling	Moderate in summer; low rest of year	Strongly seasonal
Equatorial upwelling	Positive (ocean gains heat)	Strong stratification following vertical transport	Permanent	High ($> K_S$), except $Si(OH)_4$	Upwelling and mixing	High ($\gg E_k$)	Continuous
Subarctic gyre	Seasonally positive and negative	Moderate stratification following winter mixing	Seasonal convective mixing	High ($> K_S$ in winter; $\approx K_S$ rest of year)	Convective mixing and eddies	Low in winter ($\ll E_k$); moderate rest of year ($\approx E_k$)	Seasonal
Eastern Boundary Current	Positive	Medium	Permanent	Medium ($> K_S$)	Upwelling and lateral advection	Moderate ($> E_k$), except in winter	Seasonal
Coastal upwelling	Positive	Strong stratification following vertical transport	Continuous	High ($\gg K_S$)	Upwelling	High ($> E_k$)	Weakly seasonal

[a]See text for the location and boundaries of these ecosystems.
[b]K_s is the nutrient concentration at which nutrient uptake occurs at one-half the maximal rate; at concentrations $< K_s$, uptake is limited.
[c]E_k is the light level at which the photosynthetic rate is light saturated; at light levels $< E_k$, photosynthesis is light limited.

Table 2 Summary of size, primary productivity, export, and limiting factors of ocean ecosystems

Ocean ecosystem	*Size*[a]		*Primary productivity amount and pattern (mmol C m^{-2} day^{-1})*[b]	*Export ratio*[c]	*Factors limiting primary productivity*	*Factors limiting fish yield*
	Area ($m^2 \times 10^{12}$)	*%*				
Low-latitude gyre	164	52	35 continuously	Low; weak seasonality	Macronutrients	Low primary productivity, small size of organisms and long food web
Southern Ocean	77	25	120 summer mean; ≈35 in spring and fall; ≈0 in winter	High in summer; very low in winter	Iron and light	Short duration of summer bloom
Equatorial upwelling	22	7	90 continuously	Low continuously	Iron and silicate	Low export productivity
Subarctic gyre	22	7	150 spring bloom mean; ≈50 rest of year	Episodically high; moderate rest of year	Depth of winter mixing	Short duration of spring bloom
Eastern Boundary Current	21	7	150 summer mean; 75 rest of year; grades into gyre ≈35	High in summer; moderate rest of year	Unclear, but iron is a factor, and light is also in winter	Unclear
Coastal upwelling	6	2	300 continuously close to coast; grades into Eastern Boundary Current ≈150	High, but spatially variable	Iron	Small size of ecosystem

[a]Size was calculated for the world ocean exclusive of noncoastal upwelling, continental shelf regions; total area was 312×10^{12} m^2.

[b]Based mainly on recent measurements made by the author in the Joint Global Ocean Flux Study.

[c]Export ratio is the relative proportion of primary productivity that is exported vertically, horizontally, or to higher trophic levels. The maximum observed export ratios relative to total primary productivity are approximately 0.50.

Table 3 Summary of biotic characteristics, the zooplankton component of ocean ecosystems

Ecosystem	*Endemic spp.*	*Species richness*	*Species evenness*	*Variability of biomass*	*Size of organisms*	*Importance of protozoan micrograzers*
Low-latitude gyre	Low	High	High	Low	Small	Dominant
Southern Ocean	High	Low	Very low	Very high ($>5\times$)	Large	Moderate to low
Equatorial upwelling	Moderate	High	High	Low	Small	Dominant
Subarctic gyre	High	Low	Low	High ($>2\times$)	Moderate	Moderate to low
Eastern Boundary Current	Moderate	Moderate	Moderate	Low ($<2\times$)	Moderate	Dominant to low
Coastal upwelling	Low	High	Moderate	Moderate ($\approx 2\times$)	Large	Low

Ocean equatorial region there is no manifestation of this ecosystem because the upwelling, if present, does not extend upwards into the euphotic zone.

4. A subarctic gyre ecosystem is present in both the North Atlantic and the North Pacific in the region north of the Subtropical Front at approximately 40° N. In the Atlantic, the subarctic gyre ecosystem is bounded on the south by the North Atlantic Current and Subtropical Front, on the west by the Labrador Current, on the north by the East Greenland Current, and on the east by Europe. The North Atlantic subarctic gyre is a complex, but very well-delineated, ocean feature. The North Pacific subpolar gyre is bounded on the south at approximately 40° N by the North Pacific Current and the Subtropical Front, on the west by Siberia, on the north by the Aleutian Islands and Alaska, and on the east by Canada.
5. An Eastern Boundary Current (EBC) ecosystem is represented in each of the four great EBCs: the California Current off the west coast of the United States and Mexico, the Peru Current off Peru and Chile, the Canary Current off Northwest Africa, and the Benguela Current off Namibia and South Africa. These locations are characterized by a broad and weak equatorward-flowing current that usually extends offshore 400–600 km. Each of these ecosystems receives upwelled water from a coastal upwelling ecosystem and, in turn, exports water to the adjacent low-latitude ecosystem.
6. There are four great coastal upwelling ecosystems: Peru Current, Benguela Current, Canary Current, and California Current. This ecosystem is narrow but long; it extends for great distances along the west coasts of South America, North America, northwest Africa, and South Africa. This ecosystem is also present in the northwestern Arabian Sea, along the coasts of Oman and Yemen, where evidence of upwelled water is present up to 600 km off the Omani coast.

Characteristics of Ocean Ecosystems

Shared Characteristics

The following characteristics of ocean ecosystems are based on a 1974 synthesis by McGowan for the oceanic Pacific Ocean:

They are few in number (only six). This small number of ocean ecosystems contrasts dramatically with the much larger number of well-defined terrestrial, intertidal, and coastal habitats that occur in much smaller areas. The reason, of course, is that the fluid medium of the open ocean is fundamentally partitioned by the dominant physical processes that occur in an oceanic area, and there are only a few (six) of these overriding physical patterns. In contrast, terrestrial, intertidal, or coastal ecosystems are fundamentally partitioned by spatial or geographic features and there is a much larger number of unique geographic features associated with the solid surface of the earth's crust. The same fundamental concept pervades all aspects of the comparison of oceanic biodiversity with that of terrestrial, intertidal, or coastal biodiversity. The fluid medium of the ocean is relatively homogeneous, and the boundaries that do exist are dynamic processes such as the presence or absence of winter convective mixing. It is obvious that dynamic processes of this nature form physical boundaries that permit much more biological exchange than the physical barrier associated with the solid earth.

They are large relative to terrestrial, intertidal, or coastal ecosystems. The largest ocean ecosystem, the low-latitude gyre ecosystem, occupies 52% of the open ocean area of the world ocean. This ecosystem's size is a simple reflection of the phenomenon that the earth has vast surface area occupied by the great circulation features called gyres which are slow-turning circular ocean current patterns set in motion ultimately by wind and the rotation of the earth. These gyres, together with the tropical warm pool regions and certain of the marginal seas, occupy a large portion of the earth's surface. Because of the size of the low-latitude gyre ecosystem, it plays an important role in the global heat budget and in exchanges with the atmosphere, and it is the single most important ecosystem for understanding future global change. Ironically and unfortunately, it is one of the least understood ecosystems on Earth. Its size and remoteness from societal activity, especially direct economic activity, have led both scientific and political policymakers to assign the large low-latitude gyre ecosystem a low priority for scientific effort.

On the other hand, the smallest ocean ecosystem, the coastal upwelling ecosystem, has received thorough, comprehensive, and multinational study during the past three decades. This ecosystem is small relative to other ocean systems; it has very strong physical, chemical, and biological processes; it has huge economic importance and considerable geological importance because much of the earth's oil is assumed to have been formed in past upwelling ecosystems. Coastal upwelling ecosystems do have great societal importance, and the global effort to study and understand these

important systems reflects well on the wisdom of scientific and political decision makers.

They are geologically old (with the exception of coastal upwelling). Oceanic cores suggest that the defining oceanographic features have been in place for the past 200,000 years and perhaps for the past 1 million years. The coastal upwelling ecosystem, however, may have disappeared entirely during periods of lower sea level such as during the last Glacial Maximum approximately 18,000 years ago.

They respond to climate but not to weather. Subtle and pervasive changes in mixed layer depth, the depth of convective mixing, or the strength of stratification cause profound changes in ocean ecosystems. On the other hand, strong storms or the passage of a violent hurricane cause no change that is detectable in the ocean ecosystem 2 weeks after the event. El Niño events cause profound changes in equatorial upwelling, EBC, and coastal upwelling ecosystems, but when El Niño ends the lower trophic levels of these ecosystems (phytoplankton, bacteria, microzooplankton, and mesozooplankton) return to their pre-El Niño condition within 1 month. Higher trophic levels, of course, require several years to recover.

They have considerable internal homogeneity, i.e., they tend to be monotonous. The key word is "internal." There are clear changes in almost all biological properties when physical or oceanographic boundaries are crossed; however, if oceanographic boundaries are not crossed, biological properties will remain surprisingly similar over distances of thousands of kilometers.

They are relatively undisturbed by anthropogenic processes compared to other ecosystems. The largest human disturbance thus far has been ruthless overfishing that removes top predators. The removal of these long-lived and slow-growing carnivores appears to set off a trophic cascade that changes the ecosystem even down to the level of nutrient regeneration by bacteria and protozoa. A second and more threatening change is that as the earth's heat budget changes the processes of precipitation, winter mixing, stratification, and depth of the nutricline all change in such a way that there is less transport of nutrients into the euphotic zone. This change, already apparent in the North Pacific Ocean, will reduce productivity and, hence, the yield of food resources.

The basic organization (**Figure 1**) *is like that of other ecosystems*. It appears that the basic assembly rules for all ecosystems are very similar. There has to be a large number of primary producers, somewhat fewer herbivores, still fewer carnivores, and many fewer top predators. There have to be efficient recyclers to return a portion of the nutrients back to the euphotic zone. In addition, the ecological "rules" regarding diversity, species composition, and adaptation are expressed similarly in terrestrial and ocean ecosystems.

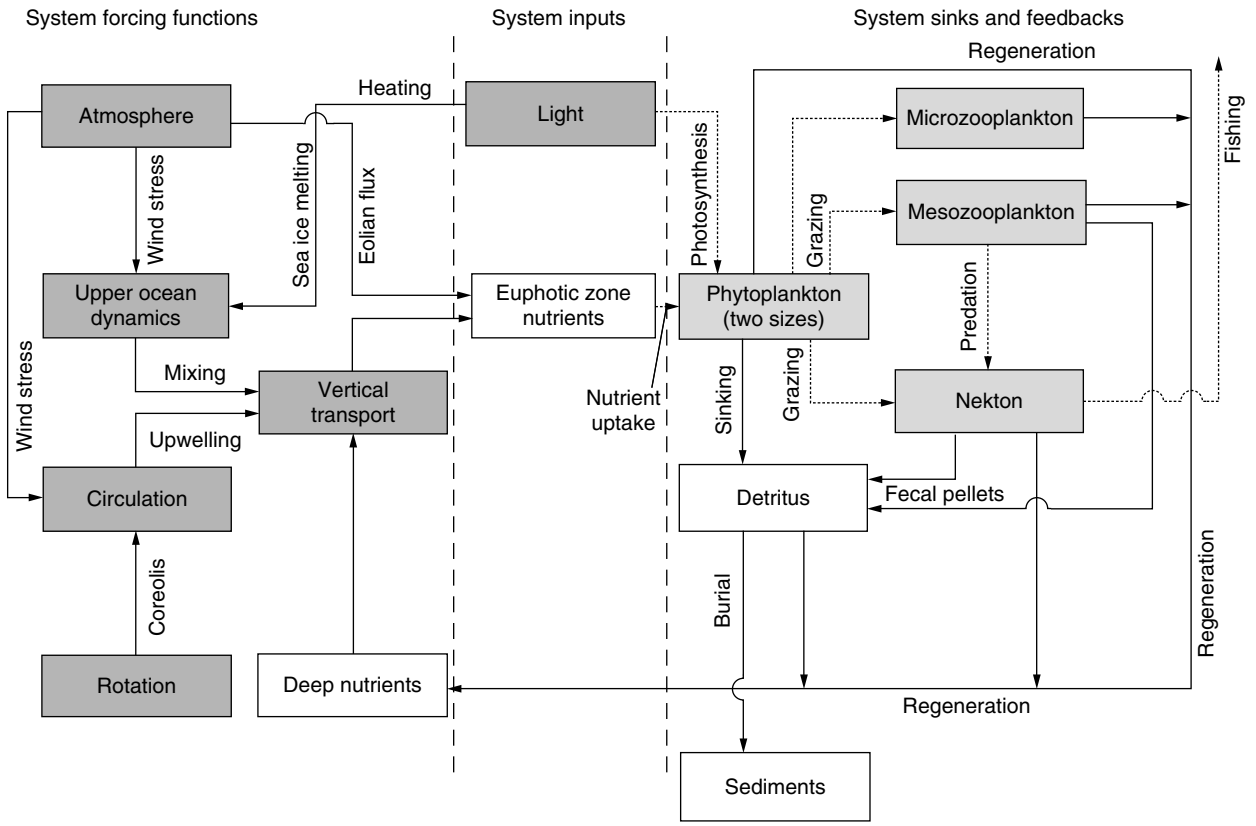


Figure 1 Generalized organization of ocean ecosystems. Physical forcing and input functions are shown in dark gray, chemical and geologic input functions are white, and biological components are in light gray. Physical transfers of energy, momentum, or mass are shown by arrows with solid lines; biological transfers of mass and energy between living components are shown by arrows with dotted lines. The export process at the extreme right, labeled fishing, is a proxy for all processes that remove biomass from the euphotic zone of the represented ecosystem.

Distinguishing Characteristics

There are qualitative differences in the basic biological processes. The most dramatic illustration of this is the observation that processes limiting productivity, biomass, export, and yield are different in almost every one of these six ecosystems. Although our understanding of limiting factors may change, currently it appears that the low-latitude gyres are limited by fixed nitrogen, especially nitrate and ammonia; the Southern Ocean appears to be limited by iron and light; the equatorial upwelling ecosystem appears to be limited by iron and silicon; coastal upwelling ecosystem rates appear not to be limited by any macronutrient, but the space where optimal coastal upwelling occurs is highly constrained and iron supply is involved; the EBC ecosystem appears to be limited by light and iron in winter and by iron alone in the summer; and the subarctic gyre ecosystem appears to be limited by the depth of the winter mixed layer or convection.

There are quantitative differences in basic processes such as primary productivity among different kinds of ocean ecosystems. One of the important milestones in the study and understanding of these differences was the work of John Ryther, who in 1969 published a work that provided a quantitative explanation of why fish yields vary by approximately 200-fold from the richest ocean ecosystems to the poorest. Variations in productivity, of course, are well-known from terrestrial ecosystems, but on land either aridity in deserts or freezing in polar regions is responsible for the low productivity of the poorest regions. Understanding why the benign low-latitude gyre ecosystem was so poor in fish production was much more difficult than understanding why productivity was low in deserts and polar regions. Part of the explanation was proposed in 1955 by Sverdrup, who said simply that the physical supply of nutrients to the euphotic zone is the reason for low fish yields in stratified ocean ecosystems. Ryther's contribution amplified the physical explanation by considering the nature of the ecological processes that lead to fish production.

First, Ryther estimated that approximately half the fish caught in the world are caught in coastal upwelling ecosystems, the smallest of the ocean ecosystems. Why? To begin, Sverdrup was correct: The physical processes of upwelling bring abundant nutrients to the surface layer, so primary productivity is very high in upwelling ecosystems. However, much more is involved. The phytoplankton that thrive in the rich coastal upwelling ecosystems are very large—so large that some are eaten directly by fish. This means that in coastal upwelling ecosystems the food chain leading to fish is very short. Ryther estimated that half the fish diet was phytoplankton and half was zooplankton. On average, then, the length of the food chain leading to fish had 1.5 transfers: large phytoplankton to fish or large phytoplankton to zooplankton to fish. At each ecological transfer, a large portion of the energy of the food is used to support the organism and this portion cannot be passed up the food chain. The shortness of the food chain is a factor that multiplies the high nutrient effect. Ryther also noted that in the small and food-rich coastal upwelling ecosystem the fish and zooplankton did not have to work as hard to get food, so the efficiency of transfer was increased relative to that of a poor environment such as the low-latitude gyre. Ryther proposed that fish yields were high in the coastal upwelling ecosystem because of high nutrients, high primary productivity, large size of the primary producers, short food chains with few transfers, and increased efficiency of transfer. These effects multiply each other, leading to very high yields of fish. For the same reason, the abundance of seabirds and marine mammals is also very high in the coastal upwelling ecosystem.

The same arguments in reverse explain the low fish yields of the low-latitude gyre ecosystem. The other four ocean ecosystems range between these two extremes. The order of fish yield per unit area can be approximately estimated, from highest to lowest, as (1) coastal upwelling ecosystem, (2) subarctic gyres ecosystem, (3) EBC ecosystem, (4) Southern Ocean ecosystem, (5) equatorial upwelling ecosystem, and (6) low-latitude gyre ecosystem.

Ocean Ecosystems and Global Change

Human intervention in the material cycles and trophic structure of ocean ecosystems may already have caused some changes. The interventions are so varied and play out at so many different scales that it is difficult to know how to describe their impact on future climate change. One approach is to focus on model studies of how anthropogenic ocean warming will affect ocean ecosystems. Considerable effort has gone into investigating how increased atmospheric carbon dioxide will affect the physical ocean–atmosphere system. Other effort has gone into determining how the ocean's carbon dioxide system will behave in an anthropogenically warmed ocean. Today's atmosphere–ocean climate models are fully coupled to an ocean carbon model, including a full carbon system and the biological transfers of the pelagic food web. Results of these model runs are dramatic and, as expected, they predict (i) a warming of global mean sea surface temperature by 2.5 °C, (ii) a slight decrease in light because of increased cloud cover, (iii) a large increase in vertical stratification because of increased precipitation which lowers salinity, and (iv) increased heating of the surface layer. The increase in stratification and decrease in convective mixing are predicted to cause a 37% reduction in the supply of nutrients to the euphotic layer. All physical processes that transport nutrients from the deep ocean reservoir to the surface layer are affected: upwelling, wind and tidal mixing, and convective mixing.

Of the three classes of impacts—warming, light, and increased stratification—increased stratification will have the major impact on ocean ecosystems. The changes predicted vary in intensity from one ocean ecosystem to another, but they are most severe in productive ecosystems such as the subantarctic gyre, equatorial upwelling, and coastal upwelling ecosystems. Very little change in nutrient supply is predicted for the world ocean's five great low-latitude gyres. The predicted change is an expansion of the size of the world ocean's most oligotrophic ecosystems: Approximately half of the ocean area will have decreased nutrient supply; the low-latitude gyres, the poorest half, will have no change; and a very small area (less than 5%) in the high Arctic and Antarctic will have an increase in nutrient flux to the surface.

The El Niño phenomenon has provided evidence of the biological consequences of reducing an ocean ecosystem nutrient supply. New populations that became established under the low nutrient conditions are healthy, highly diverse communities, but they are dramatically different. Another consequence of reduced nutrients is that the biological pump sequesters less carbon dioxide. This change will feed back into the global carbon system to accelerate the increase in concentration of atmospheric carbon dioxide.

A reduction of new nutrients by 37% will have a strong impact on both the quantity and the kind of fish present. Because the increase will occur in what are now rich fishing regions, the coastal upwelling and subarctic gyre, the societal impact will be larger than the biological impact. A change of 37% in new nutrients is significant, but it is not enough to destroy ocean ecosystems. The current gradient in nutrient flux from the low-latitude gyre ecosystem to a coastal upwelling ecosystem region is more than a factor of 10. The predicted changes will reduce the yield from the world's rich fishing banks, but the expanded low-latitude gyre ecosystems should maintain their ecological integrity.

See also: Climate Change and Ecology, Synergism of. Ecosystem, Concept of. Marine Ecosystems. Pelagic Ecosystems. Plankton, Status and Role of

References

Barber RT (1988) Ocean basin ecosystems. In: Pomeroy LR and Alberts JJ (eds.) *Concepts of Ecosystem Ecology*, pp. 171–193. New York: Springer-Verlag.

Barber RT and Smith RL (1981) Coastal upwelling ecosystems. In: Longhurst A (ed.) *Analysis of Marine Ecosystems*, pp. 31–68. San Diego: Academic Press.

Karl DM (1999) A sea of change: Biogeochemical variability in the North Pacific Subtropical Gyre. *Ecosystems* 2: 181–214.

Longhurst A (1998) *Ecological Geography of the Sea*. San Diego: Academic Press.

McGowan JA (1974) The nature of oceanic ecosystems. In: Miller C (ed.) *The Biology of the Oceanic Pacific. Proceedings of the 33rd Annual Biology Colloquium*, pp. 9–28. Corvallis: Oregon State Univ. Press.

Odum EP (1969) The strategy of ecosystem development. *Science* 164: 262–270.

Odum EP (1977) The emergence of ecology as a new integrative discipline. *Science* 195: 1289–1293.

Odum EP and Odum HT (1959) *Fundamentals of Ecology*. Philadelphia: Saunders.

Pomeroy LR and Alberts JJ (eds.) (1988) *Ecological Studies 67, Concepts of Ecosystem Ecology, A Comparative View*. New York: Springer-Verlag.

Ryther JH (1969) Photosynthesis and fish production in the sea. *Science* 166: 72–76.

Sherman K and Alexander LM (eds.) (1986) *Variability and Management of Large Marine Ecosystems*, AAAs Selected Symposium No. 99. Boulder, CO: Westview.

Steele JH (1985) A comparison of terrestrial and marine ecological systems. *Nature* 313: 355–358.

Steele JH (1998) From carbon flux to regime shift. *Fisheries Oceanogr.* 7: 176–181.

Sverdrup HU (1955) The place of physical oceanography in oceanographic research. *J. Mar. Res.* 14: 287.

Tomczak M and Godfrey JS (1994) *Regional Oceanography: An Introduction*. Oxford: Pergamon.

Oceanic Islands: Models of Diversity

Rosemary G Gillespie, University of California, Berkeley, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–13,
© 2001, Elsevier Inc.

Glossary

Adaptive radiation Evolution of ecological and phenotypic diversity within a rapidly multiplying lineage.
Allopatric Occurring in separate, nonoverlapping geographic areas.
Dioecious In plants, having the male and female reproductive organs borne on separate individuals of the same species.
Disharmony The nonrandom representation of biota among natural colonists of oceanic islands as compared with mainland systems. Species that are more dispersive will have a higher representation in the biota, the effect being accentuated by the tendency of these few successful colonists to diversify into multiple species through adaptive radiation.
Ecomorph A local population or group associated with a particular ecological niche; the term has no systematic or taxonomic significance.
Endemic/endemism Unique to a given locale; occurs in the site at which it is endemic, and nowhere else.
Impoverishment On an island, paucity of species relative to adjacent mainland.
Invasional meltdown The compounded effect on an ecological community of successful and successive invasions of alien species benefiting from the previous introduction of the other.
Microcosm A small, representative system that has analogies to a larger system. Oceanic islands are often considered microcosms of continental areas because the composition of species on islands, and the processes of formation and decline, are analogous but occur on a very small scale.
Parapatric Occurring in geographical areas that abut but do not overlap.
Phylogenetics The study of evolutionary relatedness among various groups of organisms (e.g., species and populations). Phylogenetics, also known as phylogenetic systematics, treats a species as a group of lineage-connected individuals over time.
Relictualization Survival of species on islands after those on adjacent mainland areas have suffered extinction.
Sympatric Occurring in the same geographic areas without interbreeding.
Taxon cycle A hypothesis proposed by E. O. Wilson to describe the common niche expansion among colonizers, and subsequent long term toward differentiation, with isolated species narrowing their niches and restricting their ranges resulting in eventual extinction.

Introduction

There is a huge diversity of types of islands, but the common denominator is that they are isolated, well-defined geographically, and have distinct boundaries. These characteristics can result in properties such as of a microcosmal nature and a uniquely evolved biota. Indeed, the flora and fauna of oceanic islands inspired scientific interest from the moment explorers first began sailing the oceans. The importance of evolutionary processes on islands was first recognized in Charles Darwin's work on the Galapagos, which gave rise to the theory of evolution by means of natural selection. Likewise, Alfred Russell Wallace contemporaneously developed similar theories based on studies in the Spice Islands of Indonesia. Over the last century, research has focused on multiple facets of islands—from the mechanism of formation of the isolated habitat and how its geological history has been shaped, to the colonization of species and the development of uniquely evolved biotas in similarly unique communities and ecosystems (Keast and Miller, 1996). Island ecosystems tend to be simple when compared with continental ecosystems; thus, the way ecosystems function can be more easily studied. In essence, islands can be considered as Nature's test tubes. Each island represents a trial in an experiment and each new island is the repeat of one of these experiments.

Islands can be broadly classified as one of the two types, continental or oceanic. Continental islands are fragments of a larger landmass, often created by a rise in sea level or when land breaks off and drifts from the mainland. Oceanic islands (e.g., **Figure 1**) can result from volcanoes rising above the water on or near a mid-ocean ridge, or where lithospheric plates converge (Nunn, 1994). The key biological difference between a continental and an oceanic island is that the former, when created, has a full complement of species (or at least a sample of what occurred on the continent before the fragment became isolated), and the number of species, at least initially, will tend to decline as a result of stochastic events. In contrast, oceanic islands are formed without life, so the number of species can only increase, and will do so at a rate that will depend on the interplay of isolation and time (Gillespie and Roderick, 2002). If the islands are not extremely isolated, equilibrium will be reached between immigration of colonists and extinction. With greater isolation (and lower immigration rates), and given sufficient topographical diversity, local endemism will increase over time. Extreme isolation, with the frequency of colonization being insufficient to allow the occupation of available ecological niche space

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00231-8



(a)



(b)

Figure 1 Comparison of oceanic islands. (a) View from the top of Mt. Rotui, Moorea (Society Island archipelago, French Polynesia) showing topographical diversity characteristic of younger volcanic islands; higher elevation forested habitat is shown in the foreground. (b) Atoll in Micronesia. Such oceanic islands, which show little topographic diversity, harbor relatively few terrestrial species, with little endemism. Photos by G.K. Roderick and R.G. Gillespie.

through migration, can provide the opportunity for adaptive radiation, and has been called the "radiation zone" (Whittaker, 1998). Here, species diversity arises through the formation of multiple new species adapted to exploit the available ecological space.

Over the last few years, geological understanding of islands has improved immensely. For many oceanic archipelagoes, the age of each island is often known with some level of precision. This, coupled with the fact that diversification on these islands has a clearly recognizable beginning (colonization from outside), has given biologists a temporal framework within which to examine ecological and evolutionary processes, and the formation of communities over time.

Characteristics of Biodiversity

Community Characteristics

Disharmony

On oceanic islands, the biota is often considered to be "disharmonic." This phenomenon arises partly because taxa differ considerably in their ability to colonize remote islands, leading to attenuation in the number of groups represented on more isolated islands. Moreover, as a consequence of the isolation, islands generally show a degree of impoverishment when compared with continental communities. Adaptive radiation accentuates this effect, with multiple species arising from the few that successfully colonize, the result being very

few families and genera compared with overall diversity. This pattern has been noted in the Canary Islands, (Juan *et al.*, 2000) but the effect is particularly well illustrated as one crosses the Pacific Ocean from west to east in the direction of increasing isolation of an oceanic island (**Figure 2**; it should be noted that this figure excludes continental islands such as New Caledonia and New Zealand, in which ancient connections to continental landmasses also appear to have allowed high levels of endemism).

Endemism

The pattern of species accumulation on oceanic islands depends on the rate of evolution relative to the frequency of island colonization, the extreme cases leading to extensive adaptive radiation as a result of *in situ* evolution with associated adaptation to occupy the available ecological space. On remote islands, the frequency of colonization becomes vanishingly rare. As a result, the biodiversity of remote islands has largely arisen through evolution and adaptation of the few initial colonists, and levels of endemism are consequently very high (see **Table 1** for a comparison of endemism among different groups of species in the Canary, Galapagos, and Hawaiian islands). The effect of isolation on endemicity is shown in **Figure 2**, which illustrates how endemism increases with isolation in the oceanic islands of the Pacific.

Relictualism

The mainland taxa that give rise to the island colonists exist in an environment that is often very different from that of species on islands. Accordingly, species that colonize islands may suffer extinction, though this cannot generally be distinguished from the effects of disharmony. Likewise, island progenitors may become extinct on the mainland yet remain extant on islands, because competition may not be as severe, the climate is less extreme; in this case the species on the islands can be considered "relicts." Clearly, this effect will be much more pronounced on older islands or archipelagoes.

Species Characteristics

Reduced Dispersal Capabilities

Loss of dispersal ability is a common feature of taxa on oceanic islands (Williamson, 1981). The phenomenon was first documented by Darwin, who suggested that "powers of flight would be injurious to insects inhabiting a confined locality, and expose them to be blown to the sea." For example, island plants are renowned for their large seed size and associated loss of dispersal ability. Among endemic/indigenous

Table 1 Approximate levels of endemism (%) for different groups of native organisms on three of the better-known oceanic archipelagoes

	Birds (all species)	*Flowering plants*	*Terrestrial invertebrates*
Canary Islands	45	40	58
Galapagos Islands	48	36	50
Hawaiian Islands	81	91	98

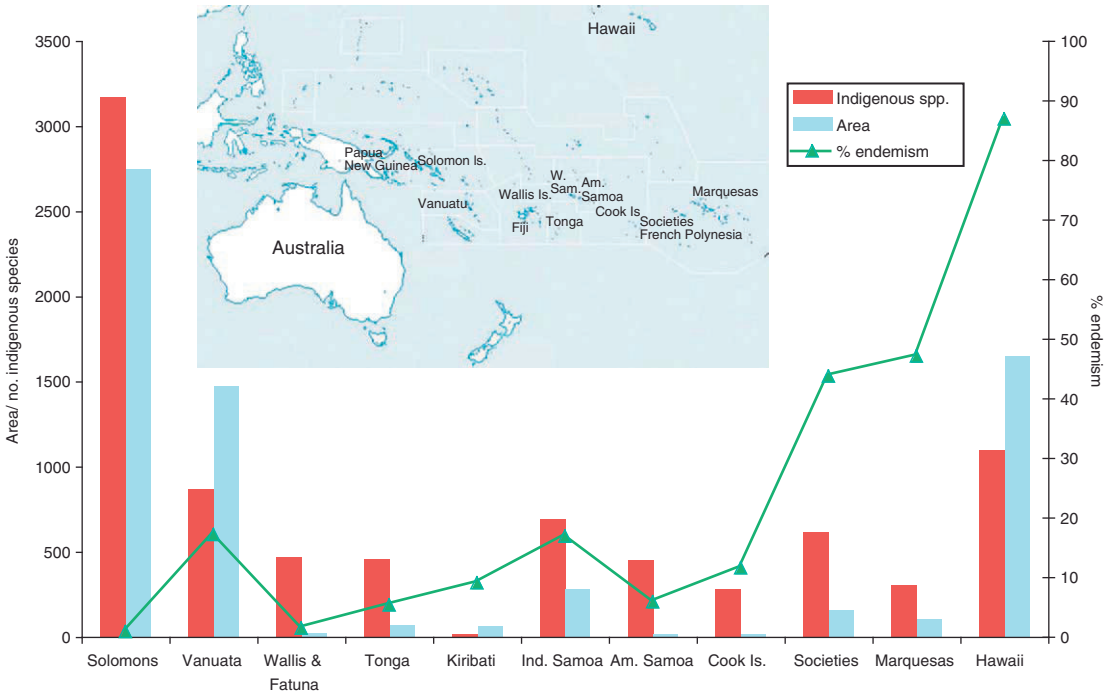


Figure 2 Plant transect across selected Pacific islands: indigenous species, area, and endemism. Larger islands have disproportionately larger numbers of species, as documented in MacArthur and Wilson (1967)'s "*Equilibrium Theory of Island Biogeography*." The initial idea of equilibrium numbers of species on islands was based on a balance between immigration (decreases with distance from a source of colonists) and extinction (decreases with island size). However, the high levels of endemism on remote islands also reflects the role of evolution within these islands, which may compensate for low levels of immigration with numerous species arising through adaptive radiation.

insects, 90% of those on Tristan de Cunha and 40% on Campbell Island show some degree of wing reduction. Among birds, flightlessness has evolved repeatedly on islands where predators are absent. Such reduction in dispersal ability may accelerate species diversification. However, many of the flightless birds on islands have been driven to extinction as human occupation spread, for example, in Hawaii the flightless rails and geese and the large flightless waterfowl the moa nalos, in Mauritius the Dodo, and in Madagascar the elephant bird.

Innovations

The known disharmony of islands appears to stimulate the development of some unusual traits in island groups. For example, Hawaiian birds have developed a remarkable diversity of elongate and decurved bill shapes, presumably because of the availability of nectar resources. The Hawaiian silversword plant alliance includes trees, shrubs, subshrubs, rosette plants, cushion plants, and a vine, that occur in some of the wettest sites recorded on Earth to extreme desert-like environments. Among insects, the absence of native social insects in some of the more remote islands of the Pacific may have been largely responsible for the adaptive shift to predation in some terrestrial insect groups and the diversity of other predatory arthropods on these islands. One of the most striking innovations is the development of predatory behavior in a lineage of Hawaiian caterpillar moths in the genus *Eupithecia* (Howarth and Mull, 1992). Another innovation in Hawaii has been the development of odonate damselflies with terrestrial larvae, associated with the paucity of lakes in the upland yet moist forests of the islands. Similarly, the unusual habitat of cryptogram herbivory in beetles in the subantarctic islands is considered to be a result of the disharmony of the flora.

Size Changes

Species on islands show a tendency toward size extremes, both gigantism and dwarfism. Overall, among mammal species that colonize islands, large species tend to get smaller while small ones tend to get larger. Included among the islands' giants are the giant flightless birds of New Zealand (moa) and Hawaii, giant sheet-web spiders of the remote Pacific islands, and hissing cockroaches of Madagascar. Dwarf species include elephants, foxes, rabbits, and snakes as well as the pygmy mammoth, which once inhabited California's Channel Islands. It appears that size changes are associated with the very different conditions experienced by island species, often with fewer constraints on size extremes.

Reproductive Changes

A common pattern seen in plants that colonize islands is the development of dioecy, asexual reproduction, and low rates of reproduction, each of which may arise as a result of selection for loss of dispersal ability, together with high selfing rates, after lineages have become established on remote islands. The Hawaiian Islands, for example, are particularly rich in dioecious species with about 15% dioecy in native species of flowering plants (compared with 6% in mainland species).

Island Dynamics

Colonization

Given a suitable abiotic environment, the success of any given colonization event depends largely on the availability of ecological space, although chance clearly plays a role. The importance of available ecological space is clearly illustrated on isolated archipelagoes that are formed serially, some in a linear fashion as they emanate from an oceanic hotspot. There is generally a heavy bias in the direction of more recent colonization toward younger islands (Wagner and Funk, 1995), which suggests that, once ecological space is "filled" on these islands, there is little further colonization.

Species Change Subsequent to Colonization

What happens to a population subsequent to successful colonization? Initially, the population may change ecologically, but over time evolutionary changes occur.

Ecological Change

Subsequent to the initial establishment of species in new habitats, a common outcome is ecological (or competitive) release, with colonizers expanding their range to fill the available ecological space. Regular cycles of distributional change following colonization of islands have been proposed several times in the literature, starting with the idea suggested by E. O. Wilson for a "taxon cycle" in Melanesian ants: widespread, dispersive populations give rise to many more restricted and specialized populations or species (Wilson, 1961). Although the generality of the taxon cycle has been questioned, the tendency for species to expand their range to fill the available ecological space upon initial colonization of an area is quite predictable.

Evolutionary Change

As described above, the remoteness of many oceanic islands allows evolution to play a prominent role in the addition of species to such islands, and adaptive radiation of species from single colonizers is a frequent characteristic of the more remote archipelagoes. Mechanisms underlying adaptive radiation on oceanic islands have been the subject of considerable interest ever since Darwin described the radiation of finches in the Galapagos (Grant, 1986). However, the extreme diversity of form and modifications associated with island colonization, coupled with the superficially similar appearances of many species on different islands within an archipelago, made it very difficult to understand how, and from where and what the diversity had arisen. The advent of molecular genetic methods has led to a huge growth in research on oceanic islands and has allowed the evolutionary history of organisms to be assessed (Givnish and Sytsma, 1997). Given the microcosmal nature of islands, with a finite pool of colonizers, and the diversification of these into multiple closely related species confined to small areas, knowledge of the evolutionary history of the biotas has allowed some profound insights into the mechanics of population divergence, and the interplay between isolation, sexual selection, and ecological shifts in allowing species formation.

Adaptive radiation, by definition, involves ecological diversification (Schluter, 2000). On oceanic islands, ecological isolation is clearly demonstrated by numerous sympatric members of an adaptive radiation. Within a radiation of long-jawed spiny-legged spiders in the Hawaiian Islands, 5–8 species co-occur at a given site, each species ecologically distinct from the others with which it co-occurs. Likewise, among the Hawaiian honeycreepers, seven ecologically distinct species co-occur at a single location in high-elevation forests on the island of Hawaii. The ecological differences among species are often profound, particularly on more remote islands where a given lineage can sometimes monopolize the ecological arena that would be used by multiple genera, families, or even orders, in continental areas.

Adaptive radiation is frequently associated with an acceleration in rate of diversification. However, the rate of diversification within an adaptive radiation may be further enhanced by sexual selection. Recent research has shown that crickets in the Hawaiian Islands have the highest documented rate of speciation known, which has been attributed to the role of courtship or sexual behavior in accelerating species divergence (Mendelson and Shaw, 2005). Likewise, sexual selection is thought to have played a prominent role in the diversification of Hawaiian drosophilid fruit flies (Kaneshiro and Boake, 1987), leading to the large number of native species (~800 in two genera, *Drosophila* and *Scaptomyza*) (Carson and Kaneshiro, 1976).

The initial processes of divergence leading to species formation on oceanic islands is still not entirely understood, particularly with respect to the relative importance of allopatric, parapatric, and sympatric speciation. Adaptive shifts clearly underlie species diversification in many radiations of species in isolated islands, although separation of gene pools is generally considered to be a requirement in initiating the process and has been demonstrated in the initial diversification of finches in the Galapagos, beetles in the Canary Islands, and spiders in the Hawaiian Islands, among other groups. In each case, subsequent sympatry appears to have led to ecological differentiation. In the Hawaiian Islands, differentiation among marginal isolates or founder populations appears to have been involved in species formation within radiations of *Drosophila* and carabid beetles. Parapatric divergence through adaptive shifts may have played a role in the initial divergence of cave-adapted species and has been implicated in the formation of cave endemics from surface relatives in the different Hawaiian Islands (lycosid spiders and isopods) and the Canary Islands (dysderid spiders) (Roderick and Gillespie, 1998). The idea that species on islands may be able to diverge in sympatry held considerable popularity for many years, but increasing evidence from phylogenetic studies has provided little support to the idea. Cases where multiple sister species co-occur within a single island are most commonly found in arthropods, and may be attributed to parapatric or "micro-allopatric" speciation, although some authors have suggested that sympatric speciation may have played a role in the diversification of palms on Lord Howe Island (Savolainen *et al.*, 2006) and lineage of tiny weevils (genus *Miocalles*) on the single small island of Rapa in southern French Polynesia. Together, these studies show that whatever the mechanism allowing initial divergence of populations, ecological segregation clearly plays the primary role in allowing coexistence of close relatives of an adaptive radiation.

Species Diversification

Species and the communities in which they live are dynamic entities, their existence, distribution, and abundance, dictated by evolutionary and ecological responses to both natural and anthropogenic environmental change. For hotspot archipelagoes—those in which islands emanate from a single volcanic hotspot from which they progressively move away—afford a unique opportunity: they make it possible to visualize snapshots of evolutionary history. For example, the geological history of the Hawaiian archipelago is well understood, with individual islands arranged linearly by age (**Figure 3**), allowing early stages of diversification and community formation to be studied on the island of Hawaii and compared with progressively later stages on the older islands of Maui, Lanai, Molokai, Oahu, and Kauai (Roderick and Gillespie, 1998). A similar chronological arrangement is found in the archipelagoes of both the Marquesas and the Societies in French Polynesia. Because of this geological history, coupled with their extreme isolation, the Pacific hotspot archipelagoes afford the opportunity for phylogenetic study of community formation within a temporal context. Other hotspot archipelagoes, such as the Galapagos and the Canary Islands, show temporal variation although islands in these groups are arranged as a cluster rather than a linear series, yet on these islands also the evolutionary progression of species tends to be from older to younger islands, and it may be possible to use the history of these islands in the same way as that in hotspot islands, to examine stages in adaptive radiation and community formation.

Extinction

Species on islands tend to have a limited geographic distribution and the naturally small population sizes means that most of these populations will not suffer additional inbreeding depression. However, they are vulnerable to extinction as a result of demographic stochasticity. Accordingly, natural extinction rates on islands are expected to be high. In the tropical islands of the Pacific, extinction has been much higher than on the islands of the north Atlantic, where species ranges are much broader. The naturally small population sizes and ranges make island species more vulnerable to extinction as a result of natural competition or habitat degradation due to geological and climatic changes, and such effects have played a pivotal role in shaping island biotas. At the same time, these same attributes of island species makes them more vulnerable to human-mediated impacts (see below).

Processes of Community Formation

Over the last few years, studies have started to incorporate phylogenetic history into the investigation of community assembly (Losos, 1992). While most of these studies argue for

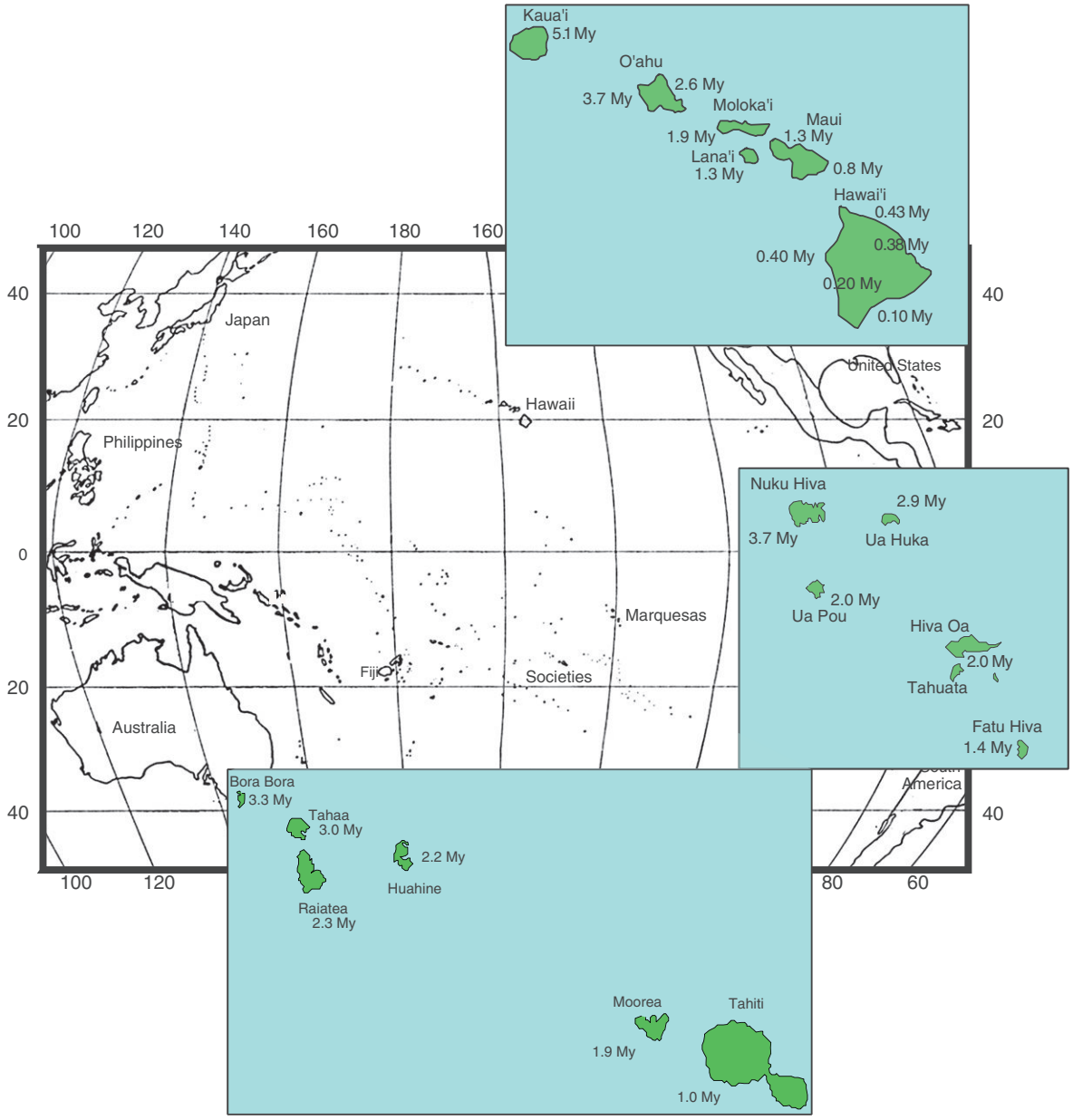


Figure 3 Major hotspot archipelagoes in the Pacific. The archipelagoes of the Hawaiian Islands, Marquesas, and Society Islands are shown on the map of the Pacific, with approximate geological ages of each island indicated. Within each archipelago, the oldest island is to the northwest, the youngest to the southeast.

the nonrandom nature of community composition, they have demonstrated opposing forces in shaping communities. For example, habitat filtering in which only taxa with certain traits colonize a community, can lead to phylogenetic clustering within that community. However, competitive exclusion, in which taxa that exploit similar resources will compete and exclude each other, can lead to phylogenetic overdispersion (Figure 4). Oceanic islands have played a key role in the integration of phylogenetics and evolutionary biology into studies of community composition. Using the fact that a given island system often includes multiple co-occurring representatives of the same lineage it is possible to infer how characters have changed to allow multiple species to co-occur in a community. In general, such studies have highlighted the role of convergent evolution, and suggest that community formation is somewhat deterministic. Accordingly, in a radiation of Hawaiian *Tetragnatha* spiders, ecologically and morphologically similar sets of species (ecomorphs) occur in most habitats on all islands: one "green spiny," one "large brown spiny," one "maroon spiny," and one "small brown spiny." These sets of species appear to have arrived in a habitat, either by direct colonization from an older island, or by evolution of one ecomorph from another on the same island, suggesting convergence of the same set of ecological forms on each island (Gillespie, 2004).

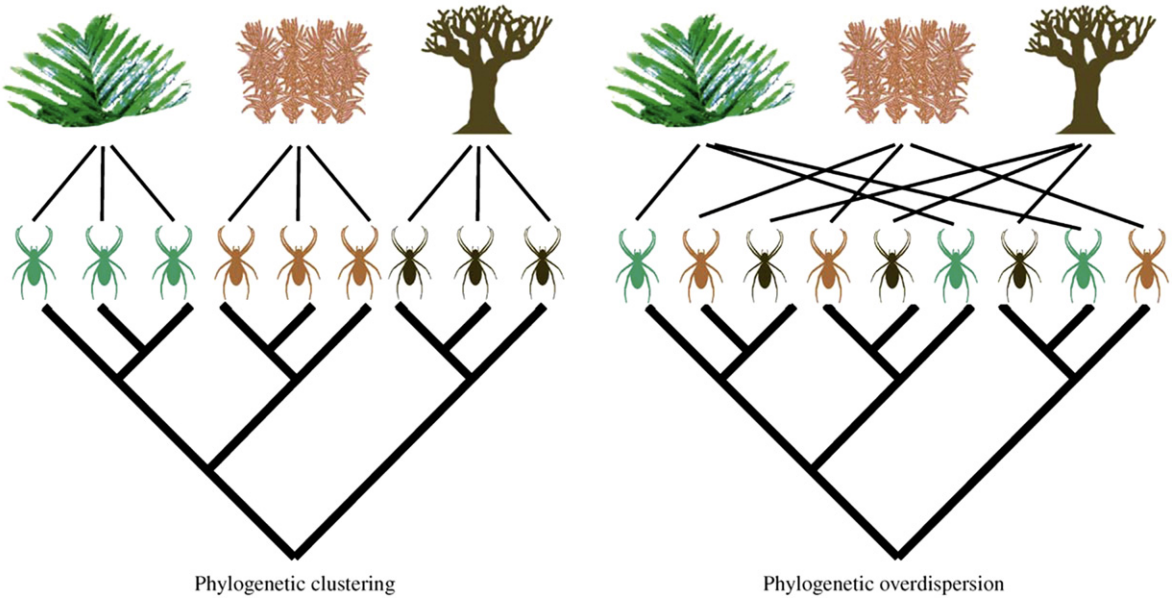


Figure 4 Hypothetical scenario of phylogenetic relationships among spiders showing how different ecological forces give rise to opposite expectations about the similarity of species traits within communities. Phylogenetic clustering may occur if closely related species share similar physiological limitations; similar adaptations allow species with similar traits to co-occur in the same habitat. Phylogenetic overdispersion may occur as a result of competitive exclusion among related species preventing coexistence of close relatives that share a resource, leading to ecological divergence among close relatives.

Species Loss

While the isolation of oceanic islands has made them living laboratories for understanding species adaptation and evolution, it has also made them extremely vulnerable to invasive species and other stresses. The vulnerability of islands to extinction is now widely recognized. Islands have suffered much higher rates of extinction than continents (Steadman, 1995). Of approximately 100 known land bird species in the Hawaiian Islands, almost 75% of the species present at first human contact have disappeared completely and many of those that survive are in danger of extinction (James, 1995). The situation is similar for native plants, in which about 120 of 1000 flowering species have fewer than 20 individuals left in the wild, and Hawaii has lost more documented arthropod species than the entire continental United States. Indeed, the Hawaiian Islands have been dubbed the "extinction capital of the world." Given the information on species diversification that scientists have been able to determine from oceanic islands, can we also use the islands as tools to understand human-mediated extinction? Extinction is clearly more gradual in continents than on many islands, but the trajectory is similar. Accordingly, understanding of the processes involved in species extinction in the islands, and replacement of largely native environments by species from elsewhere, may allow the development of effective conservation strategies in mainland habitats.

Alien Species Invasion

One of the most severe impacts affecting many island systems comes from nonnative species (Sax *et al.*, 2005). For insects, extinctions attributed to alien species invasion have been noted on Guam and the Galapagos. However, perhaps because it is the most remote archipelago, alien species in Hawaii appear to have taken a much greater toll on the native biota. The numbers now estimated for organisms purposely or accidentally introduced into Hawaii and established are 3046 arthropods, 20 reptiles, 46 land birds, 19 mammals, and 927 plants.

Community dynamics, and the ability of invasive species to penetrate certain native communities, and of certain native communities to withstand invasion, has been the subject of numerous articles (Stone and Scott, 1985). However, extensive research over the past few decades has shown that multiple parameters may act (or interact) to dictate invasion success, the most common being: (i) species diversity itself has been found, under some conditions, to serve as an ecological barrier to invasion; this effect has been used to explain the higher frequency and impact of invasions on islands which are considered to be relatively species poor; (ii) disturbance and the opening of ecological space can facilitate invasion; (iii) propagule pressure, in particular the relative abundance of native versus nonnative propagules, can permit infiltration of alien species into native habitats; (iv) characteristics of propagules, such as more generalist habitat requirements, can affect invasion success; and (v) novelty of perturbations and "naïveté" of native biota can accentuate the impact of an invasive species.

It is widely cited that the worst kind of introduction to an isolated archipelago is that of a generalist predator or competitor capable of exploiting a broad array of habitats and causing secondary impacts. This can sometimes result in an effect known as "invasional meltdown." For example, the crazy

ant *Anoplolepis gracilipes* has invaded Christmas Island, resulting in complete modification of the ecosystem in just 2 years. Supercolonies of the ants extirpate the dominant native omnivore, a large land crab, which in turn leads to increased seedling recruitment but slows litter decomposition. At the same time, the association of the ants with plant-feeding insects increases the abundance of Homoptera, with associated increased impacts on plants including honeydew productions, which in turn enhances sooty moulds.

Devastating introductions by generalist predators/competitors such as pigs are largely (though by no means completely) historical, occurring before protocols were established for the introduction of nonnative species. However, abundant, small human commensals, such as ants (**Figure 5**), mosquitoes, and agricultural and horticultural pests, are continually expanding their range on to oceanic islands. Moreover, although the introduction of highly specialized species for biocontrol purposes has been believed to be safe in that they are unlikely to go beyond the confines of their specialized host or habitat, evidence is accumulating in Hawaii to suggest that specialized species can expand their range when faced with novel conditions, and may do so quite dramatically.

Demographic Effects—Small Population Sizes

The limited geographic range and small population sizes that characterize islands make them vulnerable to demographic stochasticity (see above). Species with naturally small population sizes are similarly more vulnerable to habitat modification, simply because loss of even a small amount of habitat for a geographically restricted species could easily reduce numbers below sustainable levels.

Abiotic Effects—Climate Change

Islands can also be highly climate-sensitive. In the past two decades, for example, short-term extreme high temperatures contributed to a decline of coral reefs throughout the tropics. Corals, stressed by high temperatures, may eject their symbiotic algae resulting in coral bleaching, which renders the corals vulnerable to any further physiological stress and many of the colonies die. In 1994, elevated sea temperatures killed over 90% of the living corals of American Samoa from the intertidal zone to a depth of 10 m and fishing catches declined drastically following the coral death. Likewise, during the El Niño of 1997–98, coral bleaching in Palau, known for its spectacular coral reefs, was extensive.

Other concerns of climate change relate to the fact that most of the island species have restricted distributions such that local vagaries of climate can eliminate species so restricted as compared with more widespread taxa. In particular, the high-elevation cloud forests that characterize many oceanic islands have a narrow geographical and climatological niche that may be eliminated with even a slight climatological change. Climate change may also precipitate declines in forests due to floods, droughts, or increased incidence of pests, pathogens, or fire. In addition, increases in the frequency or intensity of hurricanes may cause disturbance that favors invasive species.

Sea level rise is also a concern for oceanic islands, resulting in coastal erosion and salt water intrusion into freshwater lenses. However, it is the low-lying islands and atolls that are the most vulnerable from these effects.

The Future of Biodiversity on Oceanic Islands

Oceanic islands have served as a source of intrigue for biologists for centuries. Not only are they discrete units within which the biota can be quantified and compared, but the interplay of isolation and time has allowed the development of uniquely derived faunas on some islands. Understanding of the interplay between time and isolation in allowing these unique sets of biota to develop is still limited. Even more limited is our knowledge of the nebulous concept of "vacant niche space," which appears to play a critical role in fostering colonization and diversification. There is much to be learned from the natural patterns of diversification on islands. At the same time, the current rate of anthropogenically induced habitat degradation and homogenization of the world's fauna threatens the biota of the world, particularly that of islands. So what are the priorities for addressing the ongoing extinction crises on these islands?

Taxonomic Knowledge

On many islands, the taxonomic impediment is huge. Accordingly, determining the status of a species, whether it is native or not, remains a major challenge for a conservation biologist on oceanic islands, particularly for those working on arthropods or other megadiverse groups. This problem is especially acute in tropical areas where knowledge of both potentially native and the pool of nonnative species is frequently sparse and unreliable. Research is needed to determine the identity and distribution of native and alien species on different islands, and subsequently to determine their evolutionary history and assess their future trajectory.

Development of Realistic Evolutionary Trajectories

In a world subject to accelerating global and local environmental change, policy makers and planners are struggling to maintain the status quo of diversity in natural systems. At the same time, there is growing appreciation that biological diversity is inherently dynamic. Conservation strategies that do not incorporate these evolutionary and ecological dynamics and that ignore the spatial flux in environmental quality will inevitably fail. Accordingly, an explicitly dynamic approach is required to allow the assessment of the conditions under which areas of endemism are either sensitive or resilient to environmental change. Oceanic islands provide circumscribed areas of discrete, highly endemic, assemblages of organisms within which understanding of evolutionary and ecological dynamics shaping the biota is an achievable goal.

It is not clear what such an assessment will tell us. For example, considering the Hawaiian Islands, it is estimated that the existing angiosperm flora developed from 265 colonization events (not including introductions or doubtfully native taxa),

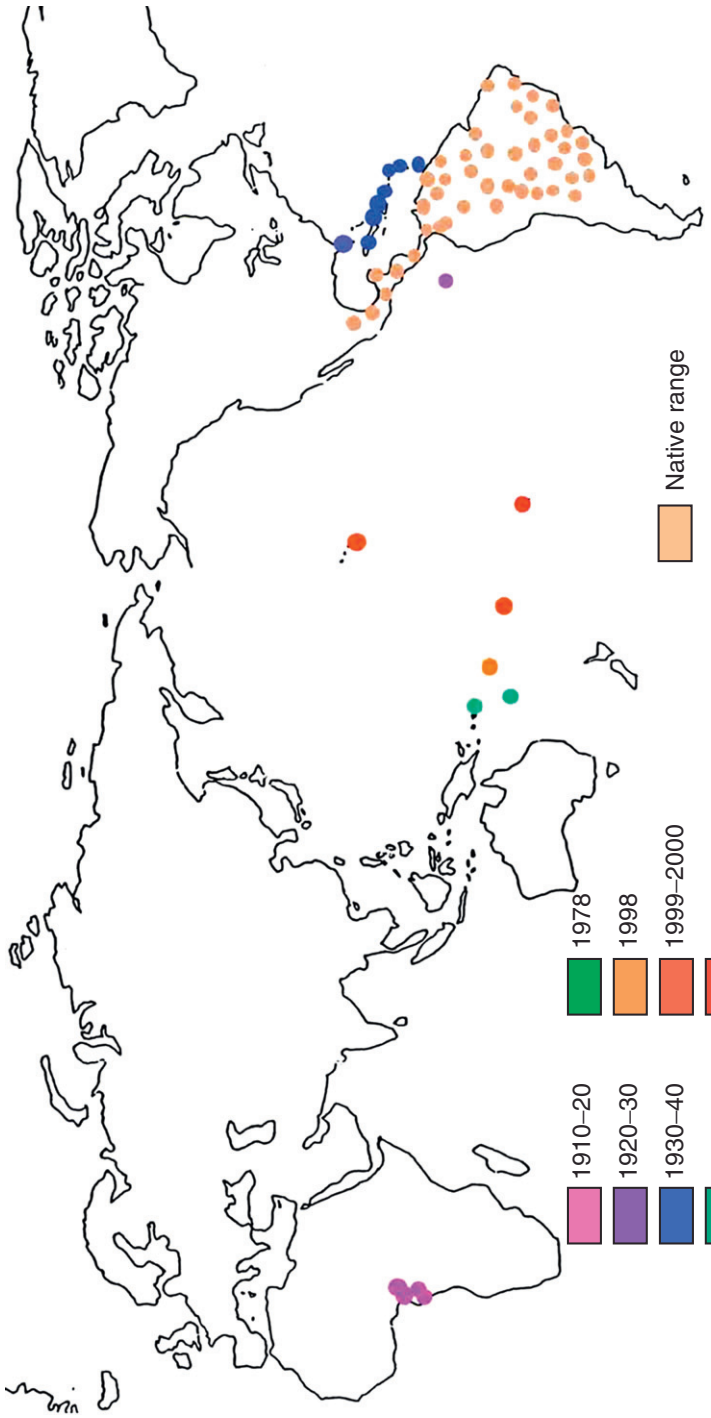


Figure 5 Invasion of the little fire ant, *Wasmannia auropunctata*, in the Pacific. *W. auropunctata* is native to South and Central America. Over the last century, the species has invaded multiple locations worldwide, and in each location its invasion has been associated with substantial impacts on the native biota. (Note that lower green represents New Caledonia and higher green the Solomons; similarly 1999–2000 represents Fiji and Hawaii, and 2005, Tahiti.)

with the number of species arising per colonist varying from 1 to 91. If colonization commenced on the oldest current high-elevation island of Kauai (which appears to be true for most groups), 53 successful colonizations per million years would be required to give rise to the current biota. In 2005, there were 1,162,000 aircraft operations (takeoffs and landings) in the Hawaiian Islands, carrying approximately 33,300,000 passengers, 500,000 t of cargo, and 130,000 t of mail. If only one out of every 60 passengers plus one out of 200 packages (assuming the average package weighs 2 lb) each carried a stray seed, and just half of these seeds became established, this would still represent an increase in the rate of colonization to the islands of 1 billion %! Clearly the Hawaiian Islands are no longer very isolated, and, even with the utmost care in preventing the introduction of new species, the ancient processes of evolution in isolation can never be recovered.

This does not mean, however, that the native environment on oceanic islands has "lost the game," nor that we should abandon the idea of long-term preservation of the biota. It simply tells us that we need to act quickly, and that oceanic islands can provide much-needed insights into factors underlying community resistance to change, and also its ability to recover completely under intensive restoration.

Restoration

Again because of their small size and unique biota, oceanic islands provide an opportunity to examine the reality and practicality of restoration efforts. For example, d ryland forests are among the most threatened of Hawaiian ecosystems, with one of the only remaining patches of the original dryland forest in Auwahi, an area on the south slope of Haleakala on Maui. Although harboring a rich diversity of native tree species, the area is largely pasture and dominated by kikuya grass. However, excellent seed sources for most species still exist. So, over the last two decades, scientists have fenced 30 acres for restoration, killed the grass, and replanted native saplings. The results have been quite remarkable and suggest that the entire area, complete with its arthropod fauna, can be recovered. The unique microcosmal nature of oceanic islands allows restoration efforts such as these, not only simply to be conducted, but also to be analyzed and interpreted to allow extrapolation of the parameters to larger and more complex continental settings.

Recognition of the acute conservation concerns impacting islands, and the lessons that they offer, is becoming increasingly recognized. A recent Conference of the Parties (COP8) to the Convention on Biological Diversity (CBD) placed island biodiversity high on the international conservation agenda. Immediate action is clearly required, not only to document and understand the unique biota that characterizes oceanic islands, but also to use the lessons learned from these relatively simple systems potentially to address more complex global issues of conservation.

See also: Adaptation. Adaptive Radiation. Biodiversity, Definition of. Biodiversity, Evolution and. Biodiversity, Origin of. Biogeography, Overview. Darwin, Charles (Darwinism). Dispersal Biogeography. Diversity, Community/Regional Level. Endemism. Extinction, Causes of. Hotspots. Introduced Plants, Negative Effects of. Island Biogeography. Loss of Biodiversity, Overview. Market Economy and Biodiversity. Speciation, Process of. Species–Area Relationships. Species Diversity, Overview

References

Carson HL and Kaneshiro KY (1976) *Drosophila* of Hawaii: Systematics and ecological genetics. *Annu. Rev. Ecol. Syst.* 7: 311–345.
Gillespie RG (2004) Community assembly through adaptive radiation in Hawaiian spiders. *Science* 303(5656): 356–359.
Gillespie RG and Roderick GK (2002) Arthropods on islands: Evolution and conservation. *Annu. Rev. Entomol.* 47: 595–632.
Givnish TJ and Sytsma KJ (eds.) (1997) *Molecular Evolution and Adaptive Radiation*. Cambridge, UK: Cambridge University Press.
Grant PR (ed.) (1986) *Ecology and Evolution of Darwin's Finches*. Princeton: Princeton University Press.
Howarth FG and Mull WP (1992) *Hawaiian Insects and Their Kin*, 160pp. Honolulu: University of Hawaii Press.
James HF (1995) Prehistoric extinctions and ecological changes on oceanic islands. In: Vitousek P (ed.) *Ecological Studies*, pp. 88–102. Berlin: Springer.
Juan C, Emerson B, Oromí P, and Hewitt GM (2000) Colonization and diversification: Towards a phylogeographic synthesis for the Canary Islands. *Trends Ecol. Evol.* 15: 104–109.
Kaneshiro KY and Boake CRB (1987) Sexual selection and speciation: Issues raised by Hawaiian *Drosophila*. *Trends Ecol. Evol.* 2: 207–212.
Keast A and Miller SE (eds.) (1996) *The Origin and Evolution of Pacific Island Biotas, New Guinea to Eastern Polynesia: Patterns and Processes*, pp. 67–90. Amsterdam: SPB Academic Publishing.
Losos JB (1992) The evolution of convergent structure in Caribbean *Anolis* communities. *Syst. Biol.* 41: 403–420.
MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton: Princeton University Press.
Mendelson TC and Shaw KL (2005) Sexual behavior: Rapid speciation in an arthropod. *Nature* 433: 375–376.
Nunn PD (1994) *Oceanic Islands*, 418pp. Oxford: Blackwell.
Roderick GK and Gillespie RG (1998) Speciation and phylogeography of Hawaiian terrestrial arthropods. *Mol. Ecol.* 7: 519–531.
Savolainen V, Anstett MC, Lexer C, *et al.* (2006) Sympatric speciation in palms on an oceanic island. *Nature* 441(7090): 210–213.
Sax DF, Stachowicz JJ, and Gaines SD (eds.) (2005) *Species Invasions: Insights into Ecology, Evolution, and Biogeography*. Sunderland, MA: Sinaur Associates.
Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford: Oxford University Press.
Steadman DW (1995) Prehistoric extinctions of Pacific Island birds: Biodiversity meets zooarcheology. *Science* 267: 1123–1131.
Stone CP and Scott JM (1985) *Hawaii's Terrestrial Ecosystems: Preservation and Management*, pp. 149–179. Cooperative National Park Resources Studies Unit, University of Hawaii, Honolulu.
Wagner WL and Funk VA (eds.) (1995) *Hawaiian Biogeography: Evolution of a Hot Spot Archipelago*. Washington, DC: Smithsonian Institution Press.
Whittaker RJ (1998) *Island Biogeography: Ecology, Evolution, and Conservation*. Oxford: Oxford University Press.
Williamson MH (1981) *Island Populations*, 286pp. Oxford: Oxford University Press.
Wilson EO (1961) The nature of the taxon cycle in the Melanesian ant fauna. *Am. Nat.* 95: 169–193.

Oil-Palm Plantations in the Context of Biodiversity Conservation

Erik Meijaard, People and Nature Consulting International, Jakarta, Indonesia; University of Queensland, Brisbane, QLD, Australia, and Center for International Forestry Research, Bogor, Indonesia
Douglas Sheil, Mbarara University of Science and Technology, Kabale, Uganda; Center for International Forestry Research, Bogor, Indonesia, and Southern Cross University, Lismore, NSW, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biofuel Wide range of fuels that are in some way derived from biomass.
Endosperm Nutritive storage tissue in the seeds of most angiosperms.
Epiphyte Plant that grows on another plant nonparasitically or sometimes on some other object.
Hectare (ha) Area equal to 2.47 acres.
Mesocarp Botanical term for the middle layer of the pericarp – for example, comprising the flesh of fruits such as plums and cherries.
Monoecious In the current context, having male and female flowers on the same plant.

Oil Palm: Green Gold or Great Evil?

An Introduction to the Green Gold

Few topics provide as much controversy in tropical forest and wildlife conservation as the rapid expansion of oil palm (*Elaeis guineensis*) plantations. On the one hand, oil palm has been linked to deforestation, peat degradation, biodiversity loss, forest fires, and a range of social issues (Danielsen *et al.*, 2009; Koh and Wilcove, 2008, 2009; Sheil *et al.*, 2009; Sodhi *et al.*, 2010). On the other hand, oil-palm expansion is considered a powerful driver of economic development in tropical countries with low levels of welfare (Casson, 2000; McCarthy and Zen, 2010; Sheil *et al.*, 2009; World Growth, 2011), and it has been referred to as "green gold" (Friends of the Earth, 2008). Economic development can lead to reduced levels of forest loss, and biofuels from oil palm can reduce global carbon emissions, but the unanswered question is whether, at a global scale, do the benefits of oil palm outweigh the environmental costs? With much of Earth's species diversity residing in tropical areas where oil palm thrives, there seems ample reason to closely assess the role that oil palm has played in tropical deforestation and loss of wildlife. Here the authors review the role of oil palm in biodiversity loss and conservation by assessing its impacts over a range of different spatial scales and in different socioecological contexts.

Basics

The origin of oil palm lies in the tropical rain forest region of West Africa in a region about 200–300 km wide along the coastal belt from Liberia to Angola (Duke, 1983). It has been described as "probably the most useful tree in West Africa" (Irvine, 1961). In prehistory, the palm was likely spread by people to a much larger area in Africa, ranging from 16° N latitude in Senegal to 15° S in Angola and eastward to the Indian Ocean, Zanzibar, and Madagascar. It has also been introduced and cultivated outside Africa and now occurs throughout the tropics between 16° N and 16° S latitudes. A distinct, closely related species of the oil palm, *Elaeis oleifera* (also known as *Elaeis melanococca*), is indigenous to Latin America. We will limit our discussion to the African species and refer to it as "oil palm."

Oil palm is a pioneer species that historically appears to replace evergreen rain forest under drier climatic conditions. For example, during the mid-Holocene in western Africa, changes in African monsoon conditions, decreased humidity, and increased fire led to the contraction of wet, evergreen rain forest and the expansion of woodland savannas. On these more-open savanna-type lands, oil palms were the dominating species (Maley, 2002; Ngomanda *et al.*, 2009; Salzmann and Hoelzmann, 2005). These vegetation shifts occurred alongside relatively "warm" regional and global conditions and could be an "analog" to events that might occur under global warming (Maley, 2002). Land clearance and burning act to increase the conditions under which oil palm thrives (Sowumni, 1999).

When fully grown, oil palms are tall, erect, single-stemmed trees that vary in heights from 8 to 20 m, with a stem diameter of as much as 50 cm. The tree is monoecious, with male and female flowers in separate clusters but on the same tree. Ecologically, this is a species of riverine forests and freshwater swamps that can tolerate temporary flooding and a fluctuating water table. The species does not do well in closed forest conditions and requires adequate light and generally open canopy conditions. It grows best in lowland areas with 1780 to 2280 mm rainfall per year, with a 2–4 month dry period, and a mean minimum temperature of 21–24 °C, but the species is adaptable and with proper care can be grown in climatic conditions outside these ranges (Duke, 1983). Its ecological adaptability is also clear from the wide range of tropical soils on which the species grows and thrives, with only waterlogged, highly lateritic, extremely sandy, stony, or peaty soils providing suboptimal growth conditions. Considering the rapid expansion of oil palm into Indonesian and Malaysian peat swamp areas (Koh *et al.*, 2011), it seems clear that even these acidic and often waterlogged conditions under appropriate silvicultural care provide suitable growing conditions for oil palm (Sheil *et al.*, 2009).

The use of oil palm by early humans is well known from the archeological record. Such uses date back to at least 4000

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00340-3

years BP (Logan and D'Andrea, 2012), and it appears that people in West Africa were actively cultivating oil palm as early as 3600–3200 BP (D'Andrea *et al.*, 2007). These early people were likely encouraging the growth of oil palms and achieving higher yields by clearing land (Logan and D'Andrea, 2012). Oil palm was a "camp follower" because of its ability to regenerate from discarded seeds without any particular horticultural treatment (Zeven, 1972). It was also traded widely, as indicated by finds of palm-oil residues in 5000-year-old Egyptian tombs (Friedel, 1897) far from where the oil was likely produced.

In Africa, palm oil has many traditional uses (Maley and Chepstow-Lusty, 2001). A few written records of the local food use of a palm oil (presumably from *Ela. guineensis*) are available in accounts of European travelers to West Africa from the middle of the fifteenth century (Hartley, 1988). One source describes how oil is produced from seeds by boiling, as well as how the oil-palm kernels are roasted and either eaten directly or made into flour (De Hondt, 1749). Palm oil later became an important item in the provisioning trade supplying the caravans and ships of the Atlantic slave trade, and it apparently remains a popular foodstuff among people of African descent in the Bahia region of Brazil (Northrup, 1978). Palm oil also found its way to Europe. James Welsh first brought 32 barrels of palm oil to England in 1590, and use grew rapidly after that. By the early nineteenth century, palm oil was being used to make soap and candles; later it was used for heating and cooking and in many other products from dynamite to margarine (Henderson and Osborne, 2000).

The increasing commercial use of palm oil is shown in early trade data. In the 1840s, the West African regions of Dahomey and the Niger delta exported approximately 1000 and 13,000 tons per year, respectively; by the 1880s these totals had risen to 5000 and 20,000 (Kiple and Ornelas, 2011). After 1900, European-run plantations were established in Central Africa and Southeast Asia, and the world trade in palm oil continued to grow slowly, reaching a level of 250,000 tons per year by 1930 (Hartley, 1988), still only about 0.5% of what was produced in the early twenty-first century (*see* The Modern Expansion).

The Modern Expansion

Plantations throughout Southeast Asia originate from the seeds of only four trees planted in Java, in present-day Indonesia, in 1848 (Henderson and Osborne, 2000). In 1905, a Belgian agricultural engineer, Adrien Hallet, arrived in Sumatra, another Indonesian island, and noticed that local palms that had originated from the small Javan gene pool grew more quickly and bore a richer fruit than counterparts in the Congo, where he had previously worked (Leplae, 1939). It was obvious that under Asian equatorial conditions, the locally cultured palms held a distinct advantage over the ordinary palms of Africa (Kiple and Ornelas, 2011). Reduced seasonality in island Southeast Asia compared to west Africa has a big impact on yield, with any drought (or even loss of humidity) reducing fruit set. Also, the fact that all the Asian palms were descended from so few parents meant that the early planters could expect fairly uniform results (Kiple and Ornelas, 2011), ensuing easier management. This lowered the risks associated with plantation cultivation, an effect reinforced by the absence of the palm's usual pests and diseases in its new geographic setting. The success of oil palm was quickly noted in neighboring Malaysia, and the first plantations were established in peninsular Malaysia in 1917. By 1919, more than 6000 ha had been planted in Sumatra, rising to 32,000 in 1925, by which time 3400 ha had come under cultivation in Malaysia. Over the next 5 years, a further 17,000 ha were planted in Malaysia, whereas the Sumatran area doubled (Kiple and Ornelas, 2011). By 1998, palm oil contributed more than 5% to Malaysia's gross domestic product (Yusoff, 2006).

Oil-palm seeds were introduced to Central America by the United Fruit Company, which brought seeds from Sierra Leone to Guatemala in 1920, and from Malaysia to Panama in 1926 and Honduras in 1927 (Kiple and Ornelas, 2011). Other introductions from Java and the Belgian Congo followed, but the first commercial planting of 250 ha only took place in Guatemala in 1940. In its tropical American setting, the oil palm, however, proved vulnerable to disease – possibly due to the native American species being almost the same – and difficulties were encountered in identifying suitable growing conditions (Hartley, 1988). By 1992, the total area of oil palm planted in Latin America had grown to 390,000 ha. This is a small fraction of the area in Africa and Southeast Asia (Kiple and Ornelas, 2011), but oil-palm production in the neotropics is viewed by many as a major new force for land-use change and forest conversion in that region.

Production of palm oil in Indonesia rose from 168,000 tons grown on 105,808 ha in 1967, to roughly 16.4 million tons grown on 6.2 million ha in 2006 (Sheil *et al.*, 2009). By 2011, an annual production of 25.4 million tons was estimated for Indonesia, 18.4 million tons for Malaysia, 1450 tons for Thailand, 880 tons for Colombia, and 850 tons for Nigeria, with an additional 3281 tons from a range of countries, adding up to a global production of 50.3 million tons (USDA, 2011). Palm oil takes up about 10% of the global production of vegetable oils, which remains dominated by soybean oil (USDA, 2011). These figures suggest that Indonesia alone underwent a 150-fold increase in palm oil production in 34 years.

Currently, Indonesia is the world's largest and most rapidly growing producer. Indonesia's wet tropical climate provides ideal growing conditions for oil palm. Land is abundant, and labor is cheap (Sheil *et al.*, 2009). About 10% of Indonesia's palm oil production comes from government plantations, 40% from small holders, and 50% from private plantations (IPOC, 2006). Malaysia is the world's second-largest individual palm oil–producing nation. Together, Indonesia and Malaysia account for about 90% of crude palm oil produced globally per annum (Sheil *et al.*, 2009). In the Southeast Asian region, a total of 8.3 million ha of closed canopy oil-palm plantations occur in peninsular Malaysia (2 million ha), Borneo (2.4 million ha), and Sumatra (3.9 million ha) (Koh *et al.*, 2011), suggesting that oil palm takes up about 6.2% of the total landmass of these three regions. We note that the study by Koh and colleagues was unable to detect newly planted oil palm, so that total area of oil palm may be larger.

As a region, Africa is the second-largest producer of oil palm in the world. Data from the Food and Agricultural Organization's FAOSTAT database indicate that about 4.5 million ha of productive oil-palm plantation existed on the continent

in 2009. Some 71% of African oil palm is produced in Nigeria, with Ghana, Guinea, Côte d'Ivoire, and the Democratic Republic of the Congo being other important producers.

In addition to Africa and Asia, oil-palm production is also rapidly expanding in the neotropics, with some 700,000 ha of productive plantation in 2009 (FAOSTAT data). Nearly half of the Amazon basin, around 2.3 million km^2, appears suited in terms of climate and soils for oil-palm cultivation (Stickler *et al.*, 2008). Even though the total oil-palm area remains small compared to Asia, the mean annual rate of expansion was an astonishing 7.9% between 1991 and 2001 (Bolivar and Cuellar-Mejia, 2003). Large-scale plantations are already established in Colombia, Ecuador, and Brazil, although the latter was only the world's 14th biggest producer of palm oil in 2009. If the full potential of the Amazon basin was utilized, however, Brazil alone could dwarf the current production of Asia (Butler and Laurance, 2009). Oil-palm planting has been promoted in Colombia, where it is seen as a relatively profitable alternative to cocaine (Gómez *et al.*, 2005). To differentiate it from the less-productive but similar native species *Ela. oleifera* (*see* Uses) *Ela. guineensis* is commonly referred to as the "African palm" or "dendezeiro" (Lopes and Steidle Neto, 2011). As in Asia, oil palm is viewed as a crop that can be profitable under many different levels of management intensity (including small-holders) in a wide range of contexts (Wolff, 1999). However, there are concerns over disease – though it is likely that breeders will be able to develop healthier and more-resistant varieties and hybrids (de Franqueville, 2003).

Uses

In well-managed plantations, oil palm produces 3–8 times more oil from a given area than any other tropical or temperate oil crop (Sheil *et al.*, 2009; Yusoff, 2006). Oil can be extracted from fruit and seed, palm-fruit oil from the outer mesocarp, and palm-kernel oil from the endosperm. Most palm-fruit oil is used in foods. In contrast, most palm-kernel oil is used in various nonedible products such as detergents, cosmetics, plastics, surfactants, and herbicides, as well as in industrial and agricultural chemicals (Wahid *et al.*, 2005). The use of palm oil as a biofuel is also increasing (Persson and Azar, 2010), giving oil palm an aura of environmental sustainability. In fact, if biodiversity losses from land-use changes are disregarded, oil palm is one of the most environmentally sustainable among a range of global biodiesel and ethanol crops (de Vries *et al.*, 2010). Together with sugarcane grown in Brazil and sweet sorghum grown in China, oil palms makes the most efficient use of land, water, nitrogen, and energy resources, whereas pesticide applications are relatively low in relation to the net energy per hectare produced (de Vries *et al.*, 2010).

The traditional red palm oil produced by West African village methods has a wide range of applications. It is mostly used for food (Kiple and Ornelas, 2011). This type of oil, however, has not proved suitable for food use in the importing countries of the West, where consumers require a bland, nearly white cooking fat. Today's plantation-produced palm oil can be treated to meet Western requirements, but this was not possible before the early twentieth century (Vanneck and Loncin, 1951). Once technology had advanced enough, European food manufacturers could exploit palm oil, replacing more-expensive fats such as butter, beef tallow, and lard in central and northern Europe and olive oil in southern Europe (Kiple and Ornelas, 2011). Palm oil was suitable as both liquid oil and solid fat.

Since the late 1960s, plant breeders have taken an interest in the American oil palm *Ela. oleifera* because its oil has a high iodine value and unsaturated fatty acid content, making it especially suitable for food use (Kiple and Ornelas, 2011). However, the fruit is often small, with a thin, oil-yielding mesocarp surrounding a large, thick-shelled kernel. Harvested bunches often contain a low proportion of fruit of quite variable quality. Hybrids between *Ela. guineensis* and *Ela. oleifera* have been trialed and show some advantages over *Ela. guineensis*, despite higher production costs (Amblard *et al.*, 1995).

Concern over greenhouse gases and high prices for fossil fuel have spurred interest in biofuels and alternative sources of energy. Biodiesel from palm oil (palm oil methylester) is currently leading the pack, and major investments are already planned to convert millions of hectares of tropical forests and other land types to oil-palm plantations (Sheil *et al.*, 2009).

Biofuels may have major positive or negative effects on natural forests, forest dwellers, and owners. On the one hand, biofuel from oil-palm plantations could help to promote economic prosperity and alleviate poverty (World Growth, 2011). On the other hand, demand for biofuels could increase competition for land, threaten food production, and exacerbate inequities between rich and poor (Astyk, 2006). Whether or not the use of palm oil as biodiesel yields a net reduction of greenhouse gas emissions remains debated (de Souza *et al.*, 2010; Gibbs *et al.*, 2008) but depends a lot on the type of vegetation that existed prior to oil-palm development.

High global demand feeds the current oil-palm boom. Despite many anti–oil palm campaigns targeting palm-oil consumers and importing countries, it is likely that the sector will expand further, either in Southeast Asia or, if the land bank becomes limited there, in the African and American tropics. At current prices, it has recently been estimated that the opportunity costs of conserving forests in Southeast Asia are US \$9860–12,750 ha^{-1} from logging and a further US \$11,240 ha^{-1} from subsequent conversion into oil palm plantations (Fisher *et al.*, 2011). Others have argued that these figures are overly pessimistic (from a forest conservation point of view) and that payments for carbon sequestration and other environmental services such as clean water supply from forests could realistically offset the opportunity costs of forest development (Ruslandi *et al.*, 2011; Venter *et al.*, 2009).

Production in Small-Holder and Large-Scale Contexts

Oil-palm seedlings are typically raised in a nursery for 1 year before planting out. Planting densities range from 110 to 150 stems per ha (Basiron, 2007). In small-holder settings in Africa, planting densities can be considerably higher; densities of 200 palms per ha were common in the late 1940s, and densities of more than 300 palms per ha were not unknown (Hartley, 1988). Most commercially used oil palms mature rapidly, and fruit can be harvested only 2–3 years after planting (Basiron, 2007) although 9–15 year old trees are most productive (BisInFocus, 2006). After 25–30 years, trees become too tall to harvest and are replaced. Some long-established

plantations in Malaysia have already been replanted for the third time (Basiron, 2007).

Labor input over the life of an oil-palm project is about 1397 person day per ha in Southeast Asia; divided by 25 years, this suggests that on average each hectare of oil palm has someone working on it – and thus earning income – 56 days of the year (Ginoga *et al.*, 1999). Unlike most other crops, oil-palm production is not very seasonal, allowing more-efficient, year-round use of labor.

Once harvested, fruit deteriorates rapidly and must be processed within 24 h (Vermeulen and Goad, 2006), so access to a mill is a major factor in determining where commercial plantations can be established. Palm-oil production is therefore most efficient when the crop is grown in a large monoculture around a central processing mill rather than in small holdings interspersed with other vegetation (Maddox *et al.*, 2007). The development of small-scale floating mills may allow companies to plant and process oil-palm fruits in remote areas at smaller scales, but such initiatives have not been taken up yet and are presumably less cost-effective than large-scale plantations.

Public Perceptions

Oil palm is hotly debated. Any internet search on keywords "orangutan" and "oil palm" reveals a plethora of mostly negative attitudes toward this palm and the people behind its boom. Internet titles such as "Palm oil costs the lives of about 50 orangutans every week and its cultivation is a major cause of global warming" and "Orangutans struggle to survive as palm oil booms" further suggests that conservation and oil palm are not happy bed fellows (EIA, 1998; Robertson and van Schaik, 2001; World Growth, 2009). Oil palm has its proponents too, however. These proponents not only include obvious ones such as palm oil producers and their support organizations but also the national governments of Indonesia and Malaysia, which earn significant revenues from palm oil production. It is becoming increasingly clear that small-scale farmers in these countries prefer oil palm to other crops because of high relative returns (Feintrenie *et al.*, 2010; Rist *et al.*, 2010). The strongly divergent viewpoints about the environmental and social costs of oil palm versus its benefits (Koh *et al.*, 2009; Meijaard and Sheil, 2011) has resulted in a situation in which middle-ground solutions of minimizing oil palm's impact have become increasingly difficult (Meijaard, 2010). The situation is not helped by the significant disinformation created on both sides of the debate (Koh and Wilcove, 2009; Sheil *et al.*, 2009). Better science-based information about the positive and negative impacts of oil palm over different temporal and spatial scales is urgently needed for more-informed discussion on the impact of this palm on global biodiversity (Sheil *et al.*, 2009).

Oil Palm and Biodiversity

Value as Wildlife Habitat

Not all aspects of biodiversity are negatively impacted by oil palm, and oil-palm plantations have some conservation benefits. Similar to fig and nectar, palm nuts are considered to be keystone ecological resources, providing crucial links between plant and animal communities (Terborgh, 1986). For example, in its native West Africa, oil palm provides important resources to the chimpanzee (*Pan troglodytes verus*) (Humle and Matsuzawa, 2004; Leciak *et al.*, 2005; Sousa *et al.*, 2011). This mostly occurs in patchy oil-palm groves in a matrix of agricultural land and forests rather than the extensive areas of single-species oil-palm plantation generally found in Asia. Chimpanzees seem to prefer oil palms for building their sleeping platforms or "nests," even when they have access to natural forests; they also use palm fruits as fallback resources (Sousa *et al.*, 2011).

Other African species that feed on oil palm include Thomas's rope squirrels (*Funisciurus anerythrus*) (Pettet, 1969); white-throated bee-eaters (*Merops albicollis*), which catch and eat the epicarp of the fruit dropped by the squirrels (Fry, 1964); southern yellow-billed hornbills (*Tockus leucomelas*); and the aptly named oil-palm vulture (*Gypohierax angolensis*) (Landsborough and Moreau, 1957). Black vultures (*Coragyps stratus*) in northwestern Colombia feed heavily on oil-palm fruit and appear to prefer it to carrion (Elias and Dubost, 1982), whereas several raptor species that feed on rats thrive in oil palm in Honduras (Padilla *et al.*, 1995). In Central America, the white-faced capuchins (*Cebus capucinus*) are well-known users of oil-palm areas (McKinney, 2010; Williams and Vaughan, 2001).

Southeast Asian oil-palm areas also provide resources to certain species, and many species use the oil-palm matrix to move between forest patches, something they might not do in plantings of annual crops or grasslands. A study in Sumatra showed a wide range of species inhabiting the area of an oil-palm plantation, with 40 mammals listed in total (38, not including domestic species) (Maddox *et al.*, 2007). Of these, 63% have an important conservation value or are protected under national law, and 25% are listed as vulnerable or higher on IUCN red lists. The tiger was the most endangered species recorded on site, rated as critically endangered. Asian elephants (*Elephas maximus*) and dhole or wild dog (*Cuon alpinus*) are the next most endangered. Tigers (*Panthera tigris*) and leopards (*Panthera pardus*) in peninsular Malaysia frequently move into oil-palm estates from surrounding forest areas to prey on wild ungulates such as pigs and deer or on domestic cattle (Azlan and Sharma, 2006). In fact, a study in peninsular Malaysia suggested that a hyperabundance of the banded pig (*Sus scrofa vittatus*) in a forest reserve surrounded by oil palm was caused by abundant year-round food supply of oil-palm fruits from the extensive plantations bordering the reserve (Ickes, 2001). The presence of prey species in oil palm is both a benefit and threat to large predators (*see* Charismatic Species).

Considering that oil palm produces highly nutritious nuts, it is surprising that few records exist of Southeast Asian species feeding on oil palm. There are indications that orangutans occasionally eat oil-palm nuts (M. Ancrenaz, pers. comm.), but such use is not extensively documented. Considering that chimpanzees use these fruits extensively, it may just be a matter of time until orangutans similarly learn to do so. Observations in Sumatra suggest that both long-tailed (*Macaca fascicularis*) and pig-tailed (*Macaca. nemestrina*) macaques feed

extensively on fallen palm fruit, as do pig species (both *S. scrofa* and *Sus barbatus*) (EM, pers. obs.). A lot more records of species feeding on oil palm are available in the literature on pest species. A book on pest species that affect oil palm globally (Hill, 2008) lists squirrels (*Callosciurus* spp.), rats (*Rattus* spp.), various parrots and parakeets, porcupines (*Hystrix* sp.), and a host of invertebrate species, such as the coconut case caterpillar (*Mahasena corbetti*), the African rhinoceros beetle (*Oryctes boas*), the coconut palm borer (*Melittomma insulare*), the palm leafminer (*Promecotheca cumingii*), the South American palm weevil (*Rhynchophorus palmarum*), and the African palm weevil (*Rhynchophorus phoenicis*). As documented in Southeast Asia, these invertebrates in turn attract bird species such as *Pycnonotus goiavier*, *Prinia* spp., *Parus major*, *Copsychus saularis*, and *Halcyon smyrnensis*, all species feeding primarily on insects and normally common outside forests (Desmier de Chenon and Susanto, 2006).

The obvious issue with some species in an oil-palm plantation context is that they cause damage to plants and palm nuts. For example, population densities of *Rattus tiomanicus* are between 100 and 600 animals per ha in Southeast Asian plantings of a range of ages and localities (Wood and Fee, 2003), and losses in Malaysian palm oil caused by these rodents were valued at US $32 million annually in the 1980s (Basri and Halim, 1985). This benefits threatened species such as blood pythons (*Python brongersmai*) and short-tailed pythons (*Python curtus*) that feed on these rats, and which in Sumatra have increased in abundance because of the establishment of oil-palm plantations (Shine *et al.*, 1999).

Impact on Species Diversity

Most of the world's species diversity is concentrated in humid tropical forest (Hoffmann *et al.*, 2010; Leadley *et al.*, 2010), the ideal habitat for oil-palm fruit production. The expansion of oil palm is therefore most likely to directly impact tropical biodiversity. The same tropical region is also an area where the majority of people are primarily concerned with meeting their basic needs (Kaimowitz and Sheil, 2007; Millennium Ecosystem Assessment, 2005). Economic development in many countries in this region is driven by natural resource exploitation, adding to the pressure on remaining forest areas. With regard to oil palm, this is especially evident in Southeast Asia, where the largest areas have so far been developed. Indonesia and Malaysia's lowland forests are among Earth's most species-rich terrestrial habitats (Sodhi *et al.*, 2004; Whitten *et al.*, 2004). The loss of Southeast Asia's lowland forests threatens the region's exceptional conservation value (Curran *et al.*, 2004; Tinker, 1997) and has long been the principal conservation concern in the region (Jepson *et al.*, 2001).

Surprisingly, despite the apparent impact of oil palm on biodiversity, conservation science is a relative newcomer to this topic. A 2008 review of 678 publications on oil palm published over 35 years found that only six of the publications specifically addressed the biodiversity and species conservation aspects of oil palm (Turner *et al.*, 2008). Since that time, there have been many more scientific study of species diversity and abundance in oil palm.

Because of oil palm's light requirements, plantation development generally requires that all other vegetation is removed. Oil-palm plantations are thus dominated by only one plant species (Danielsen *et al.*, 2009; Fitzherbert *et al.*, 2008; Gillison and Liswanti, 1999). Oil-palm plantations are also structurally less complex than natural forests, with a uniform tree age structure, lower canopy, sparse undergrowth, less-stable microclimate, and greater human disturbance (Danielsen and Heegaard, 1994; Fitzherbert *et al.*, 2008; Peh *et al.*, 2006), and they are cleared and replanted on a 25–30 year rotation (Sheil *et al.*, 2009). It is therefore not surprising that the floral and faunal diversity of these plantations is very low when compared to tropical lowland rain forests.

To give examples, researchers in the province of Jambi recorded 75% less plant diversity in oil-palm plantations than in natural forest (Gillison and Liswanti, 1999). Mammals are also affected, and a 4-year study of terrestrial mammals living in and around an oil-palm plantation concession in Jambi concluded that oil-palm monocultures are very poor habitats for most terrestrial mammal species (Maddox, 2007). Only four mammal species (10% of the number detected within the approximately 80,000 ha landscape) were regularly detected in the oil palm itself, and none of these species had a high conservation value. Some species, including deer (*Cervus unicolor*), macaques (*Macaca* spp.), and pangolin (*Manis javanica*) showed limited tolerance, but, with the exception of pigs (*Sus* spp.), all species showed a general preference for non–oil palm habitats – even heavily degraded forests (Maddox, 2007). In fact, the study highlighted the conservation importance of marginal or degraded habitats often found within palm-oil concessions and highlighted that these areas can retain high conservation values (Maddox, 2007).

Most studies of oil-palm biodiversity show large differences in faunal species composition between oil palm and forests (Fitzherbert *et al.*, 2008). The animal species lost tended to include species with specialized diets and reliance on habitat features not found in plantations (such as large trees for cavity-dwelling species) and also species with the smallest range sizes and those of highest conservation concern (Fitzherbert *et al.*, 2008). Plantation assemblages were typically dominated by a few abundant generalists, nonforest species (including alien invasives), and pests.

These findings of reduced species diversity in oil palm correspond with studies elsewhere. In Malaysia, researchers found that fewer than 20 of 75 mammal species encountered in primary forest also used oil palm (PORIM, 1994). Birds are also negatively affected, with one study in a 5000-ha study site of forests, oil palm, and agricultural lands reporting that conversion of forest to plantations resulted in reduced species richness of at least 60%, which especially affected threatened forest-dependent birds (Aratrakorn *et al.*, 2006). A review study of bird faunas in oil palms and forests found that although bird species richness is lower in oil palm than in forests, bird abundance does not appear to be. Species found in plantations are generally of lower conservation concern than those from forests (Najera and Simonetti, 2010).

Invertebrate communities in oil-palm plantations seem to be similarly influenced. Beetle assemblages in habitat types in Sabah, Malaysia, ranging from primary forest, logged forest, and acacia plantation to oil-palm plantation, had the lowest species diversity in oil palm, with a few species becoming numerically dominant (Chung *et al.*, 2000). Ant species

richness in Malaysian Borneo decreased from 309 to 110 species (−64%) between a 43,800-ha primary forest areas and a 2576-ha oil-palm plantation (Fayle *et al.*, 2010), and oil palm can sustain only about 5% of the ground-dwelling ant species of the forest interior (Bruhl and Eltz, 2010). However, the impact of oil palm on species diversity was not the same across all microhabitats that were investigated, with bird's nest ferns occurring in both forests and oil palm maintaining almost the same number of ant species in both vegetation types. Species losses were much more pronounced in canopy and leaf-litter faunas (Fayle *et al.*, 2010). Also, ant communities in oil palm are dominated by nonforest species, with nine of the 23 ant species baited in the plantations never having been recorded inside the forest (Bruhl and Eltz, 2010).

Species diversity *per se* may not always be a relevant measure for ecosystem health. A study of bee diversity in a range of vegetation types, including oil palm, in peninsular Malaysia found that the diversity in oil palm, as measured by a wide range of diversity and evenness indices, was considerably higher than in primary forest, although the absolute abundance of bees was much lower (Liow *et al.*, 2001). The 2500-ha monocultural oil-palm plantation had 17 species of bee, whereas the two natural forest sites (each >2000 ha) had nine and seven, respectively. The absolute number collected, however, was 64 for the oil-palm site and 419 and 444 for the natural forest sites. The authors suggest that absolute numbers of bees rather than species diversity may be more important for maintaining the ecosystem and ecological processes than the absolute number of species, because of their role in pollination.

A recent review of 13 studies summarized how species diversity in oil palm compared to that in other plantation crops (Fitzherbert *et al.*, 2008). Because of the small sample size, control for locations and context was not possible, and the review findings need to be interpreted with caution. Rubber (*Hevea brasiliensis*) supported as many or more species as oil palm and more forest species. Cocoa (*Theobroma cacao*) had similar or higher species richness but not always more forest species. Coffee (*Coffea canephora*) supported higher ant species richness and more forest species than oil palm. Rubber, cocoa, and coffee are often grown in small-holder settings or agroforestry landscapes. Compared to oil palm, their scale of development is generally smaller, and these crops often occur in a matrix of secondary forest regrowth. This might at least partly explain why their species richness is higher than in oil palm. *Acacia mangium* plantations, which are planted for pulp and paper, are generally developed as large (>10,000 ha) industrial plantations. In Indonesia, oil palm is established in monoculture plantations ranging in size from 4000 to more than 20,000 ha (Sheil *et al.*, 2009), which is on a scale similar to industrial tree plantations. Still, acacia plantations have higher beetle species richness than oil palm, and species composition is closer to that in forest (Chung *et al.*, 2000). Similar results were found for studies of birds, which in acacia and albizia (*Paraserianthes falcataria*) plantations resembled the avifauna of secondary forest regrowth, whereas oil palm attracted few bird species (Sheldon *et al.*, 2010).

Fitzherbert *et al.*'s (2008) review suggested that only pasture and urban mown grassland had lower species diversity than oil palm, whereas gardens of mixed crops had similar or higher species richness, and abandoned pasture had more species than oil palm. *Imperata cylindrica* grasslands, a fire-induced vegetation type that commonly replaces deforested land, had more species of ants but fewer forest ant species than oil palm (Fitzherbert *et al.*, 2008). Compared to other monocultural plantation species that harbor significant native species diversity (Hobbs *et al.*, 2006; Lugo, 1992), oil-palm plantations appear to resemble the other extreme of exotic plantation species that have limited value to native biodiversity conservation (Mascaro *et al.*, 2008). The overall conclusion about biodiversity in oil-palm plantations is that, at a local scale, it is as low as the most degraded and human-altered tropical vegetation types and therefore has limited local conservation importance.

Charismatic Species

A number of species, including orangutans (*Pongo* spp.) and the Sumatran tiger (*P. tigris sumatrae*) are the focus of international concern. The conservation of these species is often mentioned in relation to the expansion of oil palm (Linkie *et al.*, 2003; Nantha and Tisdell, 2009; WWF, 2011), and these species have played an important role in shaping the public attitude toward oil palm. Although industrial oil-palm development has been ongoing for decades, it was not until the 1990s when environmental campaigns started to focus on the role that oil palm plays in the demise of iconic conservation species and their forest habitats that the public mood began to change. These campaigns initially focused on the impacts on orangutans (Buckland, 2006; EIA, 1998) but also addressed other species, primarily tigers and elephants (Friends of the Earth, 2005). The authors discuss the impacts of oil palm on these species relative to other threats.

Orangutans

The main impact of oil palm on orangutans is habitat loss, with human–orangutan conflicts associated with oil-palm development a secondary threat (Meijaard *et al.*, 2011, 2012). Orangutans are primarily arboreal creatures, using relatively large territories and mostly feeding on fruits, leaves, and barks originating from hundreds of plant species (Rijksen and Meijaard, 1999). In 2008 in Kalimantan, oil palm threatened 750,000 ha of orangutan forest, representing 5.5% of the Bornean orangutan distribution (Venter *et al.*, 2009).

Recent studies have shown unexpected ecological resilience in orangutans in selectively harvested timber concessions and plantations of *Acacia mangium* (Ancrenaz *et al.*, 2010; Meijaard *et al.*, 2010). Surprisingly, very few studies exist of orangutan use of oil-palm habitats. One report focuses on management and the avoidance of human–orangutan conflict in oil-palm areas (Yuwono *et al.*, 2007), but it does not clarify how orangutans are affected. A recent study in Sumatra investigated crop-raiding by a population of Sumatran orangutans (*Pongo abelii*) that had become isolated from natural forest in an agricultural landscape, including oil-palm plantations (Campbell-Smith *et al.*, 2011). This study showed that the oil-palm patches in this landscape offered few, if any, benefits to orangutans.

Aerial surveys in eastern Sabah, Malaysian Borneo (M. Ancrenaz, unpublished data), identified large numbers of orangutan nests in oil-palm plantations, especially in small forest patches within the oil-palm matrix. The size of these patches fluctuated from a single tree to a few hectares and the forest was highly degraded and lacked the typical forest structure. It was estimated that at least a couple of hundred individuals were using the oil-palm landscape at the time of the surveys.

As long as oil palm does not offer a food resource to orangutans and forest fragments within the oil palm are small, degraded, and few, it is doubtful that an oil-palm landscape can sustain a viable resident orangutan population in the long-term. The nests seen during aerial surveys were most probably built by "transient" orangutans that are roaming through the oil-palm estates in search of forest during their dispersal phase. Indeed, young males leave their native community when they become mature and establish their own territory in a new forest area (Goossens *et al.*, 2006). These orangutans are "connectors" in fragmented metapopulations, and oil palm could therefore have some benefits in maintaining overall connectivity.

Tigers

Tigers are threatened worldwide by habitat loss, reduction in prey, and hunting (Chundawat *et al.*, 2010) and in Malaysia and Indonesia also by expansion of oil palm (Linkie *et al.*, 2003). Like orangutans, tigers do reasonably well in selectively logged or otherwise degraded forests, but they favor areas with little human use (Linkie *et al.*, 2008). Compared to natural forests, oil-palm estates have relatively high human use. It is therefore not surprising that tigers have much higher densities in forest than in oil palm (Maddox *et al.*, 2007).

Tigers do use oil-palm areas, however, especially when these are adjacent to good-quality forest. The attraction is in the food resources such as deer, pigs, and also domestic animals. For example, tigers killed at least 60 cattle in a 27-month period in an oil-palm estate in peninsular Malaysia (Azlan and Sharma, 2006). Where large predators and oil palm coincide, this often leads to conflict and there are regular reports in Malaysian and Indonesian newspapers of oil-palm workers having been killed. Generally, tigers are unwelcome in oil palm and are often killed if they threaten workers (Brown and Jacobson, 2005). Also, crop predation by wild ungulates such as pigs and deer leads to crop protection measures, which often include nonselective techniques such as snaring, poisoning, and drive netting. These, in turn, harm or kill tigers and reduce their prey (Wibisono, 2005). In fact, it is thought that one of the main threats to the conservation of Sumatran tigers is the response to crop depredation by large ungulates in agricultural lands, including oil-palm plantations, near protected areas (Wibisono and Pusparini, 2010).

A recent study modeled extinction risk of Sumatran tiger in a landscape containing a protected area, logging concessions, pulp-wood plantations, agroforestry, oil palm, and settlements (Imron *et al.*, 2011). The study used information on tiger hunting and breeding behavior and found that the longest survival times occurred in mixed landscapes of protected areas, logging concessions, and pulp-wood plantations rather than models based on a single land use. Selectively logged forests contributed most to the survival chances of tigers in the protected area, concurring what was found by Meijaard and Sheil (2008) elsewhere. The settlement and oil-palm plantation scenarios clearly showed the detrimental effect of these land-uses on tiger persistence. Both single land use and combined scenarios resulted in extinction within a relatively short period of time, confirming that oil-palm plantations do not provide good habitat for tiger prey, provide poor tiger habitat, and experience high human pressure, which lead to the absence of tigers (Imron *et al.*, 2011).

Asian Elephants

Asian elephants (*Ele. maximus*) are primarily a forest-edge species (Rood *et al.*, 2011), suggesting they prefer to feed on the type of vegetation found in disturbed areas. Potentially, this could include oil-palm areas, but the evidence for this is unclear. On the one hand, they are reported to avoid oil palm. In a study in Sumatra, elephants were only ever recorded once on the fringes of the oil palm (Maddox *et al.*, 2007). On the other hand, another study reported that elephants are considered to pose a risk to oil-palm plantations because they often destroy palms and feed on the oil-rich palm nuts (Susanto and Ardiansyah, 2003). In fact, it has been suggested that such agricultural conflicts may pose as big a threat to Asian elephants as habitat loss (Hedges *et al.*, 2005; Linkie *et al.*, 2007). An internet search reveals many stories of elephants causing damage to oil palm and dead elephants being found in or close to oil-palm plantations, several reportedly killed by poisoning. Often the conservation authorities assist local farmers and oil-palm companies by capturing elephants and either moving them to other areas or keeping them in captivity. Trials in Malaysian Borneo, where the tamer Bornean subspecies of *Ele. maximus* uses oil-palm areas to move between forest patches, show that the use of electrical fencing to protect small-holder crops combined with the replanting of forest corridors provides an effective means to reduce elephant conflict (Ancrenaz and Lackman, 2011). This is expensive, however, and may only work in small plots. Chili grease–covered fences may be a cheaper alternative (Hedges and Giunaryadi, 2010).

Beneficial Wildlife

Oil-palm estate managers actively promote the presence of some species because they increase the production of oil palm or at least make it cheaper. Owls and snakes are the most important among these beneficial species. Barn owls (*Tyto alba javanica*) have been widely encouraged in Malaysian oil-palm plantations to control rodent pests. They were formerly considered vagrants in peninsular Malaysia, but they became established following the increase in rats with the advent of oil-palm plantations (Lenton, 1984). It is estimated that a pair of barn owls together with their chicks consume around 1300 rats per year (Duckett and Karuppuah, 1989), but it apparently remains doubtful whether these owls truly regulate rodent populations or whether rodent populations are more strongly affected by other factors such as food supply (Puan *et al.*, 2011).

Certain species of snakes are also attracted to the many rodents and other species feeding in oil-palm areas (Akani *et al.*, 2008; Shine *et al.*, 1999), and some plantations actively use snakes to control rodents, although not as commonly as owls or baiting (Hafidzi and Saayon, 2001). How effective such pest control is remains unclear.

Exclosure studies in Sabah, Malaysia, show that insectivorous birds deliver a natural pest-control service for oil-palm agriculture (Koh, 2008a). Where birds were excluded from oil-palm seedlings, herbivory rates from insects increased between 1.2- and 17.2-fold – significantly higher than that in control treatments.

Koh (2008a) reports that many companies adopt an integrated pest-management approach that favors the use of nonchemical pest control methods such as the establishment of "beneficial plants" (e.g., *Euphorbia heterophylla*) to attract the insect predators and parasitoids of oil-palm pests such as the wasp *Dolichogenidea metesae* (Basri *et al.*, 1995).

Finally, the native pollinator of oil palm (the weevil *Elaeidobius kamerunicus*) did not originally occur in Asia. When it was introduced from Africa, production increased and the cost of artificial pollination was saved (Dhileepan, 1994; Southworth, 1985).

Notes of Caution

One of the constraints on interpreting research on species diversity in oil palm is that there are few scientific case studies (Fitzherbert *et al.*, 2008). For example, there are no scientific studies that address plant diversity in oil palm. This introduces confounding factors that often cannot be controlled for. Research is required that addresses questions such as, what is the effect of area on species diversity when one compares species in a 50,000-ha natural forest with those in a 2500-ha oil-palm plantation? Would the species diversity of a 1000-ha oil-palm plantation be the same as a 10,000-ha one? What is the effect of fragmentation when areas of natural forests to which oil-palm diversity is compared are fragments themselves in a matrix of nonforests (Liow *et al.*, 2001)? How does species diversity vary in different oil-palm contexts, from the mixed-forest gardens settings often found in Africa to the large (>20,000-ha) monocultural plantings sometimes found in Indonesia?

Broader Environmental Impacts of Oil Palm

Oil Palm and Deforestation

Information on how much forest has been displaced by oil palm is hard to come by. Considering that oil palm is a crop of the humid tropics, one could argue that all planted oil palm has ultimately replaced tropical forest. Some forests, however, were cut down centuries ago and only recently planted with oil palm. Oil palm is developed under a wide range of field conditions, varying from old degraded grasslands, secondary scrubland, forest regrowth, degraded and overlogged forest, and relatively intact forests. In our experience, rarely has oil palm been established in areas that were primary forest (i.e., visually untouched by human activities) directly prior to oil-palm development. Therefore, the more-pertinent question regarding oil palm and forest wildlife is how much forest has recently been cut down and directly been replaced by oil palm? A recent analysis of agricultural and deforestation statistics for the period 1990–2005 suggested that more than half the area of oil-palm developed in Malaysia and Indonesia had resulted in deforestation (Koh and Wilcove, 2008). Others, however, argue that the data are too poor to draw such conclusions and that these estimates do not account for other causes that triggered deforestation before oil-palm plantations were established (Wicke *et al.*, 2011).

To estimate future impacts, we need to know how much of the oil-palm expansion will be in forested areas. Future demand for edible oil is estimated at around 240 Mt in 2050, requiring an additional 12 million ha of palms, if average yields continue to rise as in the past (Corley, 2009). This demand could at least partly be met on existing nonforest lands (Wicke *et al.*, 2011). However, Corley (2009) also points out that biofuel demand might greatly exceed that for edible use, and the interchangeability of the major oils for edible and biofuel uses means that this demand will drive oil-palm expansion, whether or not palm oil is actually used for biodiesel.

Without a clear definition of oil palm–induced deforestation, better data on forest cover and the distribution of oil-palm plantations, and future expansion potential of oil palm, it remains impossible to accurately quantify the impact of oil-palm development on forest wildlife.

Some have argued that oil-palm plantations are forests. Malaysia, for example, has considered (but ultimately rejected) including oil-palm plantations in the country's national statistics on forest cover (Simamora, 2010). Many conservation bodies highlighted this as unacceptable (Biofuels Watch, 2010; World Rainforest Movement, 2010), and the Food and Agricultural Organization excludes oil palm from global forest estimates because it considers it an agricultural crop, not a planted forest (FAO, 2010). Meijaard and Sheil (2011) pointed out that in much of the temperate world pulpwood plantations are included as forests, and there is an obvious need to develop and agree on such definitions (Sasaki and Putz, 2009).

Broader Environmental Impacts of Oil-Palm Plantations

Palm-oil production has environmental impacts that could potentially affect wildlife beyond the actual plantation. Extraction of palm oil results in large amounts of effluent that is often returned to natural water courses without treatment (Sheil *et al.*, 2009). Palm-oil mill effluent is a colloidal suspension of water, oil, grease, and solids: it is fairly acidic (pH 4–5) and is typically discharged hot (80–90 °C) (Ahmad *et al.*, 2005). Although most mills have treatment areas, leaks of effluent can have significant negative impacts on water quality. How this affects the ecological functioning of waterways remains largely unstudied (Sheil *et al.*, 2009).

The oil-palm industry is one of the largest consumers of mineral fertilizers in Southeast Asia (Hardter and Fairhurst, 2003). A typical oil-palm plantation planted on both mineral and peat soils requires around 354 kg ha^{-1} of nitrogen over the first 5 years to increase and maintain yields (Guyon and

Simorangkir, 2002). Pesticides and herbicides also increase pollution, especially with repeated use (Hartemink, 2005). Most of the reports on impacts are generated by companies and may not be objective because they wish to be seen as minimizing damage to the environment (Sheil *et al.*, 2009).

The environmental impact of oil-palm plantations could be less than most alternative crops if considered in terms of production – more can be produced on less land. Given the necessary trade-offs between conservation and economic growth, this is important. Better management, higher yields from improved varieties, and planting on land that is already degraded could improve yields significantly without further deforestation (Hardter *et al.*, 1997). Concentrating oil-producing crops on those lands with the highest yields could reduce the need for land elsewhere, offering potential conservation benefits

Could Oil-Palm Development Reduce Biodiversity Impacts Elsewhere?

Large-scale oil-palm production has documented benefits. The plantation sector in Malaysia is one of the largest employers, providing income and employment for many rural people. Basiron (2007) comments that "involvement in cultivation or downstream activities has uplifted the quality of life of people." Decreasing rural poverty may reduce deforestation, although this is highly context-specific (Sunderlin *et al.*, 2007; Wunder, 2001). Also, assuming a certain global demand for vegetable oil – for food and biofuel – producing it in areas with plant species that maximize yields could potentially reduce pressure on land elsewhere. The interactions between the various economic, trade, environment, and political factors remain too complex to reliably determine overall global impacts of oil palm on biodiversity compared to the alternative of producing oils with different crops. This is an important area of research to guide the different oil industries.

Enhancing the Biodiversity Values of Oil Palm

An important question regarding the biodiversity of oil-palm plantations is whether this can be boosted by retaining patches of natural forest within the oil-palm matrix, the so-called wildlife-friendly strategy (Edwards *et al.*, 2010; Fitzherbert *et al.*, 2008; Koh, 2008b). Oil palm developed in large estates can create monocultural stands of 50,000 ha or more. Such areas have very limited ecological variation and create large areas mostly devoid of wildlife. In a small-holder setting, oil palm is planted on much finer scales, often in plantations of 1 or 2 ha. If such plantations are part of a broader multifunctional landscape with remaining forest stands and secondary regrowth, the overall species diversity is likely to be higher. If a certain total area of oil-palm plantation is targeted to fulfill global demands, an important ecological question is whether for wildlife conservation purposes it is better to concentrate all oil palms into large monocultural stands (potentially leaving more space for natural forests) or to spread oil-palm plantings over much larger multifunctional landscapes.

A study comparing bird diversity in oil palm, forest fragments within oil palm, and contiguous natural forest indicated that abundances of imperiled bird species in oil palm were 60 times lower in fragments and 200 times lower in oil palm than in contiguous forest. Forest fragments did not increase bird abundances in adjacent oil palm, and they had lower species richness than contiguous forest and an avifaunal composition that was more similar to oil palm than to contiguous forest. The study concluded that, from a perspective of maximizing biodiversity conservation, any investment in the retention of fragments would be better directed toward the protection of contiguous forest (Edwards *et al.*, 2010) – that is, the land-sparing strategy.

Increasing the productivity of existing oil-palm plantations – for example, by better management of harvesting to improve oil yield – could potentially reduce the need for more land to be cleared. However, this will only generate a conservation gain if it is linked to the protection of natural habitats – for example, through strategic land-use planning and implementation (Fitzherbert *et al.*, 2008). Fitzherbert et al (2008) argue that with higher yields per unit area for both large and small scale enterprises, oil palm might provide a substitute for traditional subsistence agriculture and could reduce the area of land needed to support each household. They also point out that successful land sparing is contingent on inelasticity of demand for agricultural products (Green *et al.*, 2005). The substitutability of vegetable oils ensures that demand for any one oil is elastic and, although future global requirements for edible oils – depending much on demand from China and India – may be reasonably predictable, demand will become effectively limitless if driven by new biofuel markets. Proposals for nongovernmental organizations to use oil-palm agriculture to acquire private reserves (Koh and Wilcove, 2007) are unlikely to be the most cost-effective approach (Venter *et al.*, 2008).

Meanwhile, several new international and national initiatives are under way to improve practices in establishing oil-palm plantations and using forests. One national initiative is Sawit Watch (*sawit* meaning oil palm), which campaigns for the rights of indigenous people in land disputes and highlights the social ramifications of oil-palm developments in Indonesia (Sheil *et al.*, 2009). International initiatives include the Roundtable for Sustainable Palm Oil (RSPO), which was established in 2004 by Malaysian and Indonesian companies to ensure palm oil "contributes to a better world." The RSPO has developed a verifiable standard for sustainable palm oil and encourages oil-palm companies to adopt more-responsible practices. This standard consists of the RSPO Principles and Criteria (P&C) for Sustainable Palm Oil Production, which set out the requirements that must be met and against which certification assessments are made. To define sustainability in the oil-palm sector, the RSPO has developed 39 sustainability criteria organized under eight general principles that are designed to limit environmental impacts of growing and processing palm oil (Laurance *et al.*, 2010). Of these, principle four is of direct relevance to biodiversity. Among others, it requires that growers maintain soil fertility, minimize and control erosion and degradation of soils, maintain the quality and availability of surface and groundwater, regulate the use of agrochemicals, and effectively

manage pests, diseases, weeds, and invasive introduced species (RSPO, 2007). Most importantly for biodiversity in oil palm, however, is principle five, which concerns the environmental responsibility and conservation of natural resources and biodiversity. This focuses primarily on the design of plantations, most relevantly the clearing of natural vegetation and how this affects the status of rare, threatened, or endangered species and habitats of high conservation value. Specifically, if such species or habitats are present, the standard requires that any legal requirements relating to the protection of the species or habitat are met, damage to and deterioration of applicable habitats is avoided, and any illegal or inappropriate hunting, fishing, or collecting activities is controlled, including the development of responsible measures to resolve human–wildlife conflicts. Such conflicts are frequent – as has, for example, been indicated by the many reported cases of orangutan killing in association with oil-palm development (Meijaard *et al.*, 2011).

Despite its ambitious environmental goals, the RSPO has been criticized for failing to stop clearing of natural forests and, more generally, for noncompliance by its members (Laurance *et al.*, 2010). Also, many companies have experimented with the RSPO standard since it was ratified in November 2005 but have found it to be complicated, costly, and hard to implement (Nikoloyuk *et al.*, 2010; Paoli *et al.*, 2010). Recently, RSPO has channeled activities toward developing a standard for smallholders because they cannot afford the additional oversight required for mainstream RSPO certification. Smallholders also struggle to adopt best practices, such as zero burning, because such practices require upfront capital and are more expensive at the onset. It remains to be seen whether the lofty goal of "sustainable" palm-oil management can be attained through the RSPO process. Countries such as Indonesia that investigated ways of integrating RSPO principles into current policies (McCarthy and Zen, 2010) have apparently concluded that this was not possible and subsequently developed their own standards: the Indonesian Sustainable Palm Oil Foundation (ISPO).

Conclusions

The scientific evidence suggests that oil-palm plantations in equatorial Asia have low biodiversity value compared to most other tropical land uses. A few species do well in oil palm, but these generally have little conservation value. Species that lose out in oil palm are forest-dependent species with specific habitat requirements and low abundance, and many are of conservation significance (Persey, 2011).

It is possible to make the oil-palm industry more biodiversity-friendly. It is most important, however, that oil palm should be developed on already deforested or degraded lands rather than in areas of tropical forest. Oil palm itself can also be made more hospitable for biodiversity – for example, by increasing structural and faunistic diversity (e.g., allowing the growth of epiphytic ferns and maintaining weed cover) and retaining as much natural forest in and around the planted areas as possible.

Ultimately, the global impact of oil palm on biodiversity can only be judged in relation to the alternatives. There is an increasing demand for vegetable oils for food and other uses, and demand for biofuels is growing. Oils and biofuels can be generated with different crops, and oil palm has the highest yield per unit land area and per unit of financial investments. If oil palm is not expanded further, then either the demand for oil will not be met or it will be fulfilled with other crops that require more land than would the oil palm.

See also: Agriculture, Sustainable. Agrobiodiversity. Biodiversity-Rich Countries. Deforestation and Land Clearing. Hotspots. Land-Use Issues. Mammals, Conservation Efforts for. Market Economy and Biodiversity. Poverty and Biodiversity. Primate Populations, Conservation of. Rainforest Ecosystems, Animal Diversity. Rainforest Ecosystems, Plant Diversity. Rainforest Loss and Change. Sustainability and Biodiversity

References

Ahmad AL, Sumathi S, and Hameed BH (2005) Adsorption of residue oil from palm oil mill effluent using powder and flake chitosan: Equilibrium and kinetic studies. *Water Research* 39: 2483–2494.

Akani GC, Ebere N, Luiselli L, and Eniang EA (2008) Community structure and ecology of snakes in fields of oil palm trees (*Elaeis guineensis*) in the Niger Delta, southern Nigeria. *African Journal of Ecology* 46: 500–506.

Amblard P, Noiret JM, Kouame B, Potier F, and Adon B (1995) Comparative performances of interspecific hybrids and the commercial material. *Oléagineux Corps Gras Lipides* 2: 335–340.

Ancrenaz M, Ambu L, Sunjoto I, *et al.* (2010) Recent surveys in the forests of Ulu Segama Malua, Sabah, Malaysia, show that orang-utans (*P. p. morio*) can be maintained in slightly logged forests. *PloSOne* 5: e11510. doi:11510.11371/journal.pone.0011510.

Ancrenaz M and Lackman I (2011) The HUTAN "Elephant Conservation Unit". *2010 Yearly Activity Report*. Kota Kinabalu: HUTAN.

Aratrakorn S, Thunhikorn S, and Donald PF (2006) Changes in bird communities following conversion of lowland forest to oil palm and rubber plantations in southern Thailand. *Bird Conservation International* 16: 71–82.

Astyk S (2006) Ethics of biofuels. *Energy Bulletin*. http://www.energybulletin.net/node/24169 (accessed on 25 May 2011).

Azlan JM and Sharma DSK (2006) The diversity and activity patterns of wild felids in a secondary forest in Peninsular Malaysia. *Oryx* 40: 36–41.

Basiron Y (2007) Palm oil production through sustainable plantations. *European Journal of Lipid Science and Technology* 109: 289–295.

Basri MW and Halim AH (1985) The effects of *Elaeidobius kamerunicus* Faust on rat control programmes of oil palm estates in Malaysia. *PORIM Occasional Paper* 14, pp. 1–50.

Basri MW, Norman K, and Hamdan AB (1995) Natural enemies of the bagworm, *Metisa plana* Walker (Lepidoptera: Psychidae) and their impact on host population regulation. *Crop Protection* 14: 637–645.

Biofuels Watch (2010) Oil palm plantations defined as "forets" [sic]. http://bio-fuel-watch.blogspot.com/2010/02/oil-palm-plantations-defined-as-forets.html

BisInFocus (2006) *Prospek Perkebunan & Industri Minyak Sawit Di Indonesia 2006–2020*. Jakarta, Indonesia: PT BisInFocus Data Pratama.

Bolivar EC and Cuellar-Mejia M (2003) Evolution of the Latin American oil palm sector during the last decade (1991–2001). *Oil Palm Industry Economic Journal* 3: 1–14.

Brown E and Jacobson MF (2005) *Cruel Oil. How Palm Oil Harms Health, Rainforest & Wildlife*. Washington, DC: Center for Science in the Public Interest.

Bruhl C and Eltz T (2010) Fuelling the biodiversity crisis: Species loss of ground-dwelling forest ants in oil palm plantations in Sabah, Malaysia (Borneo). *Biodiversity and Conservation* 19: 519–529.

Buckland H (2006) *The Oil for Ape Scandal. How Palm Oil is Threatening the Orang-utan*. London, UK: Friends of the Earth Trust.

Butler RA and Laurance WF (2009) Is oil palm the next emerging threat to the Amazon? *Tropical Conservation Science* 2: 1–10.

Campbell-Smith G, Campbell-Smith M, Singleton I, and Linkie M (2011) Apes in space: Saving an imperilled orangutan population in Sumatra. *PloS ONE* 6: e17210.

Casson A (2000) *The Hesitant Boom: Indonesia's Oil Palm Sub-Sector in an Era of Economic Crisis and Political Change*, CIFOR Occasional Paper No. 29. Bogor, Indonesia: Center for International Forestry Research.

Chundawat RS, Habib B, Karanth U, *et al.* (2010) *Panthera tigris* in IUCN, editor. IUCN Red List of Threatened Species. Version 2010.4. Downloaded on 3 June 2011. http://www.iucnredlist.org/apps/redlist/details/15955/0

Chung AYC, Eggleton P, Speight MR, Hammond PM, and Chey VK (2000) The diversity of beetle assemblages in different habitat types in Sabah, Malaysia. *Bulletin of Entomological Research* 90: 475–496.

Corley RHV (2009) How much palm oil do we need? *Environmental Science & Policy* 12: 134–139.

Curran LM, Trigg SN, McDonald AK, *et al.* (2004) Lowland forest loss in protected areas of Indonesian Borneo. *Science* 303: 1000–1003.

D'Andrea AC, Kahlleber S, Logan AL, and Watson DJ (2007) Early domesticated cowpea (*Vigna unguiculata*) from central Ghana. *Antiquity* 81: 686–698.

Danielsen F and Heegaard M (1994) The impact of logging and forest conversion of lowland forest birds and other wildlife in Seberida, Riau Province, Sumatra. In: Sandbukt OW (ed.) *Rainforest and Resource Management. Proceedings of NORINDRA Seminar*, pp. 59–60. Jakarta: Indonesian Institute of Sciences.

Danielsen F, Beukema H, Burgess ND, *et al.* (2009) Biofuel Plantations on Forested Lands: Double Jeopardy for Biodiversity and Climate. *Conservation Biology* 23: 348–358.

De Franqueville H (2003) Oil palm bud rot in Latin America. *Experimental Agriculture* 39: 225–240.

De Hondt P (1749) Historische Beschrijving der Reizen of Nieuwe en Volkoome Verzameling van de Aller-Waardigste en Zeldsaamste Zee- en Landtogten. vol 7. Beschrijving der Koningryken Loango, Kongo, Angola, Benguela, en der Nabuurige Landen, 's Gravenhage.

de Souza SP, Pacca S, de Avila MT, and Borges JLB (2010) Greenhouse gas emissions and energy balance of palm oil biofuel. *Renewable Energy* 35: 2552–2561.

de Vries SC, van de Ven GWJ, van Ittersum MK, and Giller KE (2010) Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. *Biomass and Bioenergy* 34: 588–601.

Desmier de Chenon R and Susanto A (2006) Ecological observations on diurnal bird in Indonesian oil palm plantations. *Journal of oil palm Research* 18: 122–143.

Dhileepan K (1994) Variation in populations of the introduced pollinating weevil (*Elaeidobius kamerunicus*) (Coleoptera, Curculionidae) and its impact on fruitset of oil palm (*Elaeis guineensis*) in India. *Bulletin of Entomological Research* 84: 477–485.

Duckett JE and Karuppuah S (1989) A guide to the planter in utilizing barn owl (*Tyto alba*) as an effective biological control of rats in mature oil palm plantations. *Proceedings of the PORIM International Oil Palm Development Conference*, Malaysia, pp. 357–372. Kuala Lumpur, Malaysia: Palm Oil Research Institute of Malaysia.

Duke JA (1983) *Handbook of Energy Crops*. Published online at http://www.hort.purdue.edu/newcrop/duke_energy/elaeis_guineensis.html

Edwards DP, Hodgson JA, Hamer KC, *et al.* (2010) Wildlife-friendly oil palm plantations fail to protect biodiversity effectively. *Conservation Letters* 3: 236–242.

EIA (1998) *The Politics of Extinction. The Orangutan Crisis. The Destruction of Indonesia's Forests*, p. 53. London, UK: Environmental Investigation Agency.

Elias DJ and Dubost DVG (1982) Unusual feeding behavior by a population of Black Vultures. *Wilson Bulletin* 94: 214.

FAO (2010) *Global Forest Resources Assessment 2010. Progress Towards Sustainable Forest Management*. Rome, Italy: Food and Agricultural Organization of the United Nations. FAO Forest Paper 163.

Fayle TM, Turner EC, Snaddon JL, *et al.* (2010) Oil palm expansion into rain forest greatly reduces ant biodiversity in canopy, epiphytes and leaf-litter. *Basic and Applied Ecology* 11: 337–345.

Feintrenie L, Chong WK, and Levang P (2010) Why do farmers prefer oil palm? Lessons learnt from Bungo district, Indonesia. *Small Scale Forestry* 9: 379–396.

Fisher B, Edwards DP, Giam X, and Wilcove DS (2011) The high costs of conserving Southeast Asia's lowland rainforests. *Frontiers in Ecology & The Environment* 9: 329–334.http://dx.doi.org/10.1890/100079.

Fitzherbert EB, Struebig MJ, Morel A, *et al.* (2008) How will oil palm expansion affect biodiversity? *Trends in Ecology & Evolution* 23: 538–545.

Friedel MC (1897) Sur des matières grasses trouvées dans des tombes égyptiennes d'Abydos. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* 24: 648–653.

Friends of the Earth (2005) *Greasy Palms. The Social and Ecological Impacts of Large-Scale Oil Palm Plantation Development in Southeast Asia*. London, UK: Friends of the Earth.

Friends of the Earth (2008) *Malaysian Palm oil – Green Gold or Green Wash? A Commentary on the Sustainability Claims of Malaysia's Palm Oil Lobby, With a Special Focus on the State of Sarawak*. Amsterdam, The Netherlands: Friends of the Earth. www.foeeurope.org/publications/2008/malaysian-palm-oil-report.pdf

Fry CH (1964) White-throated bee-eater eating oil-palm nut fibers. *Bulletin of the Nigerian Ornithologists' Society* 1: 16.

Gibbs HK, Johnston M, Foley JA, *et al.* (2008) Carbon payback times for crop-based biofuel expansion in the tropics: The effects of changing yield and technology. *Environmental Research Letters* 3: 34001.

Gillison A and Liswanti N (1999) *Impact of oil palm plantations on biodiversity in Jambi, Central Sumatra, Indonesia*. Bogor, Indonesia: Center for International Forestry Research.

Ginoga K, Cacho O, Erwidodo, Lugina M, and Djaenudin D (1999) Economic performance of common agroforestry systems in Southern Sumatra: Implications for carbon sequestration services. *Working Paper CC03*. ACIAR project ASEM 1999/093. Canberra, Australia: Australian Centre for International Agricultural Research. http://www.une.edu.au/febl/Econ/carbon/

Gómez PL, Mosquera M, and Castilla C (2005) Oil palm: A sustainable agro-industry in Colombia. *Oléagineux, Corps Gras, Lipides* 12: 121–124.

Goossens B, Chikhi L, Ancrenaz M, Lackman-Ancrenaz I, Andau P, and Bruford MW (2006) Genetic signature of anthropogenic population collapse in orang-utans. *Plos Biology* 4: 285–291.

Green RE, Cornell SJ, Scharlemann JPW, and Balmford A (2005) Farming and the fate of wild nature. *Science* 307: 550–555.

Guyon A and Simorangkir D (2002) *The Economics of Fire Use in Agriculture and Forestry: A Preliminary Review for Indonesia*. Jakarta, Indonesia: Project FireFight South East Asia.

Hafidzi MN and Saayon MK (2001) Status of rat infestation and recent control strategies in oil palm plantations in Peninsular Malaysia. *Pertanika Journal of Tropical Agricultural Science* 24: 109–114.

Hardter R, Chow WY, and Hock OS (1997) Intensive plantation cropping, a source of sustainable food and energy production in the tropical rain forest areas in Southeast Asia. *Forest Ecology and Management* 91: 93–102.

Hardter R and Fairhurst T (2003) Introduction: Oil palm: Management for large and sustainable yields. In: Fairhurst T and Härdter R (eds.) *Oil Palm: Management for Large and Sustainable Yields*, pp. 1–12. Singapore: PPI/PPIC and IPI.

Hartemink AE (2005) Plantation agriculture in the tropics – environmental issues. *Outlook on Agriculture* 34: 11–21.

Hartley CWS (1988) *The Oil Palm (*Elaeis guineensis *Jacq.)*. New York, USA: Longman Scientific & Technical.

Hedges S and Giunaryadi D (2010) Reducing human–elephant conflict: Do chillies help deter elephants from entering crop fields? *Oryx* 44: 139–146.

Hedges S, Tyson MJ, Sitompul AF, Kinnaird MF, Gunaryadi D, and Aslan (2005) Distribution, status, and conservation needs of Asian elephants (*Elephas maximus*) in Lampung Province, Sumatra, Indonesia. *Biological Conservation* 124: 35–48.

Henderson J and Osborne DJ (2000) The oil palm in all our lives: How this came about. *Endeavour* 24: 63–68.

Hill DS (2008) *Pests of Crops in Warmer Climates and their Control*. New York: Springer.

Hobbs RJ, Arico S, Aronson J, *et al.* (2006) Novel ecosystems: Theoretical and management aspects of the new ecological world order. *Global Ecology & Biogeography* 15: 1–7.

Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science* 330: 1503–1509.

Humle T and Matsuzawa T (2004) Oil palm use by adjacent communities of chimpanzees at Bossou and Nimba Mountains, West Africa. *International Journal of Primatology* 25: 551–581.

Ickes K (2001) Hyper-abundance of native wild pigs (*Sus scrofa*) in a lowland dipterocarp rain forest of Peninsular Malaysia. *Biotropica* 33: 682–690.

Imron MA, Herzog S, and Berger U (2011) The influence of agroforestry and other land-use types on the persistence of a Sumatran Tiger (*Panthera tigris sumatrae*) population: An individual-based model approach. *Environmental Management* 48: 276–288.

IPOC (Indonesian Palm Oil Commission) (2006) *Statistik Kelapa Sawit Indonesia 2005*. Jakarta, Indonesia: Department of Agriculture.

Irvine FR (1961) *Woody Plants of Ghana: With Special Reference to their Uses*. London, UK: Oxford University Press.
Jepson P, Jarvie JK, MacKinnon K, and Monk KA (2001) The end for Indonesia's lowland forests? *Science* 292: 859–861.
Kaimowitz D and Sheil D (2007) Conserving what and for whom? Why conservation should help meet basic human needs in the tropics. *Biotropica* 39: 567–574.
Kiple KF and Ornelas KC (2011) *The Cambridge World History of Food. II.E.3. – Palm Oil*. Cambridge, UK: Cambridge University Press. http://www.cambridge.org/us/books/kiple/palmoil.htm (accessed 25 May 2011).
Koh LP (2008a) Birds defend oil palms from herbivorous insects. *Ecological Applications* 18: 821–825.
Koh LP (2008b) Can oil palm plantations be made more hospitable for forest butterflies and birds? *Journal of Applied Ecology* 45: 1002–1009.
Koh LP and Wilcove DS (2007) Cashing in palm oil for conservation. *Nature* 448: 993–994.
Koh LP and Wilcove DS (2008) Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letters* 1: 60–64.
Koh LP and Wilcove DS (2009) Oil palm: Disinformation enables deforestation. *Trends in Ecology & Evolution* 24: 67–68.
Koh LP, Ghazoul J, Butler RA, *et al*. (2009) Wash and spin cycle threats to tropical biodiversity. *Biotropica* 42: 67–71.
Koh LP, Miettinen J, Liew SC, and Ghazoul J (2011) Remotely sensed evidence of tropical peatland conversion to oil palm. *Proceedings of the National Academy of Sciences of the United States of America* 108: 5127–5132.
Landsborough TA and Moreau RE (1957) Feeding habits of the Palm-Nut Vulture *Gypoheerax*. *Ibis* 99: 608–613.
Laurance WF, Koh LP, Butler R, *et al*. (2010) Improving the performance of the Roundtable on Sustainable Palm Oil for nature conservation. *Conservation Biology* 24: 377–381.
Leadley P, Pereira HM, Alkemade R, *et al*. (2010) *Biodiversity Scenarios: Projections of 21st Century Change in Biodiversity and Associated Ecosystem Services*. CBD Technical Series No. 50, p. 132. Montreal, Canada: Secretariat of the Convention on Biological Diversity.
Leciak E, Hladik A, and Hladik CM (2005) The oil palm (*Elaeis guineensis*) and the cores of high biodiversity in gallery forests of Guinea in relation to human and chimpanzees commensalism. *Revue d'Ecologie-La Terre et la Vie* 60: 179–184.
Lenton GM (1984) The feeding and breeding ecology of Barn Owls *Tyto alba* in peninsular Malaysia. *Ibis* 126: 551–575.
Leplae E (1939) *Le palmier à huile en Afrique, son exploitation au Congo belge et en Extreme-Orient*. Brussel, Belgium: Librairie Falk fils.
Linkie M, Dinata Y, Nofrianto A, and Leader-Williams N (2007) Patterns and perceptions of wildlife crop raiding in and around Kerinci Seblat National Park, Sumatra. *Animal Conservation* 10: 127–135.
Linkie M, Haidir WA, Nugroho A, and Dinata Y (2008) Conserving tigers *Panthera tigris* in selectively logged Sumatran forests. *Biological Conservation* 141: 2410–2415.
Linkie M, Martyr DJ, Holden J, *et al*. (2003) Habitat destruction and poaching threaten the Sumatran tiger in Kerinci Seblat National Park, Sumatra. *Oryx* 37: 41–48.
Liow LH, Sodhi NS, and Elmqvist T (2001) Bee diversity along a disturbance gradient in tropical lowland forests of Southeast Asia. *Journal of Applied Ecology* 38: 180–192.
Logan AL and D'Andrea AC (2012) Oil palm, arboriculture, and changing subsistence practices during Kintampo times (3600–3200 BP, Ghana). *Quaternary International* 249: 63–71. doi:10.1016/j.quaint.2010.12.004.
Lopes DC and Steidle Neto AJ (2011) Potential Crops for Biodiesel Production in Brazil: A Review. *World Journal of Agricultural Sciences* 7: 206–217.
Lugo AE (1992) Comparison of tropical tree plantations with secondary forests of similar age. *Ecological Monographs* 62: 1–41.
Maddox T, Priatna D, Gemita E, and Salampessy A (2007) *The Conservation of Tigers and Other Wildlife in Oil Palm Plantations Jambi Province, Sumatra, Indonesia*. London, UK: The Zoological Society of London. ZSL Conservation Report No. 7.
Maddox TM (2007) Oil palm and mammal conservation. *Paper to the International Conference on Oil Palm and Environment*. Bali, Indonesia.
Maley J (2002) A catastrophic destruction of African forests about 2,500 years ago still exerts a major influence on present vegetation formations. *IDS Bulletin-Institute of Development Studies* 33: 13.
Maley J and Chepstow-Lusty A (2001) *Elaeis guineensis* Jacq. (oil palm) fluctuations in central Africa during the late Holocene: Climate or human driving forces for this pioneering species? *Vegetation History and Archaeobotany* 10: 117–120.
Mascaro J, Becklund KK, Hughes RF, and Schnitzer SA (2008) Limited native plant regeneration in novel, exotic-dominated forests on Hawai'i. *Forest Ecology & Management* 256: 593–606.
McCarthy J and Zen Z (2010) Regulating the oil palm boom: Assessing the effectiveness of environmental governance approaches to agro-industrial pollution in Indonesia. *Law & Policy* 32: 153–179.
McKinney T (2010) The effects of provisioning and crop-raiding on the diet and foraging activities of human-commensal white-faced capuchins (*Cebus capucinus*). *American Journal of Primatology* 73: 439–448.
Meijaard E (2010) The Thinker: The Middle Way http://www.thejakartaglobe.com/columns/the-thinker-a-middle-way/403730. *The Jakarta Globe*, 28 October 2010.
Meijaard E and Sheil D (2008) The persistence and conservation of Borneo's mammals in lowland rain forests managed for timber: Observations, overviews and opportunities. *Ecological Research* 23: 21–34.
Meijaard E and Sheil D (2011) Modest proposal for wealthy countries to reforest their land for the common good. *Biotropica* 43: 544–548.
Meijaard E, Albar G, Rayadin Y, Nardiyono Ancrenaz, M, and Spehar S (2010) Unexpected ecological resilience in Bornean Orangutans and implications for pulp and paper plantation management. *PloSONE* 5: e12813.
Meijaard E, Buchori D, Hadiprakoso Y, Utami-Atmoko SS, Tjiu A, and Prasetyo D (2011) Quantifying killing of orangutans and human–orangutan conflict in Kalimantan, Indonesia. *PLoS ONE* 6: e27491.
Meijaard E, Wich S, Ancrenaz M, and Marshall AJ (2012) Not by science alone: Why orangutan conservationists must think outside the box. *Annals of the New York Academy of Sciences* http://dx.doi.org/10.1111/j.1749-6632.2011.06288.x.
Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Biodiversity Synthesis*. Washington, DC: World Resources Institute.
Najera A and Simonetti JA (2010) Can oil palm plantations become bird friendly? *Agroforestry Systems* 80: 203–209.
Nantha HS and Tisdell C (2009) The orangutan-oil palm conflict: Economic constraints and opportunities for conservation. *Biodiversity & Conservation* 18: 487–502.
Ngomanda A, Chepstow-Lusty A, Makaya M, *et al*. (2009) Western equatorial African forest-savanna mosaics: A legacy of late Holocene climatic change? *Climate of the Past* 5: 647–659.
Nikoloyuk J, Burns TR, and de Man R (2010) The promise and limitations of partnered governance: The case of sustainable palm oil. *Corporate Governance* 10: 59–72.
Northrup D (1978) *Trade Without Rulers: Precolonial Economic Development in Southeastern Nigeria (Oxford Studies in African Affairs)*. Oxford: Oxford University Press.
Padilla M, Chinchilla C, Arias E, and Flores I (1995) Diurnal predatory birds and damage caused by rats in oil palm (*Elaeis guineensis* Jacq.) in Honduras. *ASD Oil Palm Papers* 10: 1–12.
Paoli GD, Yaap B, Wells PL, and Sileuw A (2010) CSR, oil palm and the RSPO: Translating boardroom philosophy into conservation action on the ground. *Tropical Conservation Science* 34: 438–446.
Peh KSH, Sodhi NS, de Jong J, Sekercioglu CH, Yap CAM, and Lim SLH (2006) Conservation value of degraded habitats for forest birds in southern Peninsular Malaysia. *Diversity & Distributions* 12: 572–581.
Persey S (2011) Can oil palm & biodiversity co-exist? *Symposium on Sustainable Palm Oil: Challenges, A Common Vision and the Way Forward*. 5–6 May 2011. London, UK: Zoological Society of London.
Persson UM and Azar C (2010) Preserving the world's tropical forests: A price on carbon may not do. *Environmental Science and Technology* 44: 210–215.
Pettet A (1969) Feeding association of *Aerops albicollis* and *Cinnamopteryx castaneofuscus* with the squirrel *Punisciurus anerythrus*. *Ibis* 111: 98–101.
PORIM (1994) *Annual Research Report for 1994*. Bangi, Malaysia: Biology Division, PORIM. Palm Oil Research Institute of Malaysia.
Puan CL, Goldizen AW, Zakaria M, Hafidzi MN, and Baxter GS (2011) Absence of differential predation on rats by Malaysian barn owls in oil palm plantations. *Journal of Raptor Research* 45: 71–78.
Rijksen HD and Meijaard E (1999) *Our Vanishing Relative. The Status of Wild Orang-Utans at the Close of the Twentieth Century*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
Rist L, Feintrenie L, and Levang P (2010) The livelihood impacts of oil palm: Smallholders in Indonesia. *Biodiversity & Conservation* 19: 1009–1024.
Robertson JMY and van Schaik CP (2001) Causal factors underlying the dramatic decline of the Sumatran orang-utan. *Oryx* 35: 26–38.
Rood E, Ganie AA, and Nijman V (2011) Using presence-only modelling to predict Asian elephant habitat use in a tropical forest landscape: Implications for conservation. *Diversity & Distributions* 16: 975–984.

RSPO (2007) *RSPO Principles and Criteria for Sustainable Palm Oil Production Including Indicators and Guidance*. Kuala Lumpur, Malaysia: Round Table for Sustainable Palm Oil.

Ruslandi, Venter O, and Putz FE (2011) Over-estimating the costs of conservation in Southeast Asia. *Frontiers in Ecology and the Environment* 9: 542–544.

Salzmann U and Hoelzmann P (2005) The Dahomey Gap: An abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *Holocene* 15: 190–199.

Sasaki N and Putz FE (2009) Critical need for new definitions of "forest" and "forest degradation" in global climate change agreements. *Conservation Letters* 2: 226–232.

Sheil D, Casson A, Meijaard E, *et al.* (2009) The impacts and opportunities of oil palm in Southeast Asia. What do we know and what do we need to know? *CIFOR Occasional Paper No.* 51. Bogor, Indonesia: Center for International Forestry Research.

Sheldon FH, Styring A, and Hosner PA (2010) Bird species richness in a Bornean exotic tree plantation: A long-term perspective. *Biological Conservation* 143: 399–407.

Shine R, Ambariyanto, Harlow PS, and Mumpuni (1999) Ecological attributes of two commercially-harvested python species in northern Sumatra. *Journal of Herpetology* 33: 249–257.

Simamora AP (2010) Palm estate is forest, says ministry. *The Jakarta Post*. Tuesday 16 February 2010. http://www.thejakartapost.com/news/2010/02/16/palm-estate-forest-says-ministry.html

Sodhi NS, Koh LP, Brook BW, and Ng PKL (2004) Southeast Asian biodiversity: An impending disaster. *Trends in Ecology & Evolution* 19: 654–660.

Sodhi NS, Koh LP, Clements R, *et al.* (2010) Conserving Southeast Asian forest biodiversity in human-modified landscapes. *Biological Conservation* 143: 2375–2384.

Sousa J, Barata AV, Sousa C, Casanova CCN, and Vicente L (2011) Chimpanzee Oil-Palm Use in Southern Cantanhez National Park, Guinea-Bissau. *American Journal of Primatology* 73: 485–497.

Southworth A (1985) Palm oil and palm kernels. *Journal of the American Oil Chemists' Society & Natural Resources* 62: 250–254.

Sowumni MA (1999) The significance of the oil palm (*Elaeis guineenses* Jacq.) in the late Holocene environments of west and central Africa: A further consideration. *Vegetation History and Archaeobotany* 8: 199–210.

Stickler C, Coe M, Nepstad D, Fiske G, and Lefebvre P (2008) *Ready for REDD? A Preliminary Assessment of Global Forested Land Suitability for Agriculture*. Massachusetts: Woods Hole Research Center. http://whrc.org/BaliReports/assets/Bali_crop_suitability.pdf.

Sunderlin WD, Dewi S, and Puntodewo A (2007) *Poverty and Forests. Multi-Country Analysis of Spatial Association and Proposed Policy Solutions*, p. 44. Bogor, Indonesia: Center for International Forestry Research (CIFOR).

Susanto P and Ardiansyah F (2003) *The Palm Oil Industry in Riau: Dari Hutan Sampai Ke Wajan, Laporan Yayasan*. Indonesia, Jakarta: WWF.

Terborgh J (1986) Keystone Plant Resources in the Tropical Forest. In: Soulé ME (ed.) *Conservation Biology: The Science of Scarcity and Diversity*, pp. 330–344. Sunderland, MA: Sinauer Associates.

Tinker PB (1997) The environmental implications of intensified land use in developing countries. In: Greenland DJ, Gregory PJ, and Nye PH (eds.) *On the Edge of the Malthusian Precipice?*, pp. 1023–1033. London, UK: The Royal Society.

Turner EC, Snaddon JL, Fayle TM, and Foster WA (2008) Oil palm research in context: Identifying the need for biodiversity assessment. *PlosONE* 3: e1572.

USDA (2011) *Oilseeds: World Markets and Trade*. http://www.fas.usda.gov/oilseeds_arc.asp (downloaded 25 May 2011). Washington, DC: US Department of Agriculture.

Vanneck C and Loncin M (1951) Considérations sur l'alteration de l'huile de palme. *Bulletin Agricole du Congo Belge* 42: 57–64.

Venter O, Meijaard E, Possingham HP, *et al.* (2009) Carbon payments as a safeguard for threatened tropical mammals. *Conservation Letters* 2: 123–129.

Venter O, Meijaard E, and Wilson K (2008) Strategies and alliances needed to protect forest from palm-oil industry. *Nature* 451: 16.

Vermeulen S and Goad N (2006) *Towards Better Practice in Smallholder Palm Oil Production*. London, UK: International Institute for Environment and Development (IIED).

Wahid MB, Abdullah SNA, and Henson IE (2005) Oil palm – Achievements and potential. *Plant Production Science* 8: 288–297.

Whitten T, van Dijk PP, Curran L, Meijaard E, Supriatna J, and Ellis S (2004) Sundaland. In: Mittermeier RA, Gil PR, and Hoffmann M (eds.) *Hotspots Revisited: Another Look at Earth's Richest and Most Endangered Terrestrial Ecoregions*. Mexico: Cemex.

Wibisono HT (2005) *Population Ecology of Sumatran tigers* (Panthera tigris sumatrae) *and their Prey in Bukit Barisan Selatan National Park, Sumatra, Indonesia (Masters Thesis)*. Amherst, MA, USA: University of Massachusetts.

Wibisono HT and Pusparini W (2010) Sumatran tiger (*Panthera tigris sumatrae*): A review of conservation status. *Integrative Zoology* 5: 313–323.

Wicke B, Sikkema R, Dornburg V, and Faaij A (2011) Exploring land use changes and the role of palm oil production in Indonesia and Malaysia. *Land Use Policy* 28: 193–206.

Williams HE and Vaughan C (2001) White-faced monkey (*Cebus capucinus*) ecology and management in neotropical agricultural landscapes during the dry season. *Revista de Biologia Tropical* 49: 1199–1206.

Wolff C (1999) *The Economics of Oil Palm Production in Chiapas, Mexico*, Working Paper. Ithaca, NY, USA: Department of Agricultural, Resource, and Managerial Economics Cornell University.

Wood BJ and Fee CG (2003) A critical review of the development of rat control in Malaysian agriculture since the 1960s. *Crop Protection* 22: 445–461.

World Growth (2009) *Collateral Damage: How the Bogus Campaign Against Palm Oil Harms the Poor*. Arlington, VA: World Growth.

World Growth (2011) *World Bank's New Anti Poor Palm Oil Policy. Green Papers Issue VIII*. World Growth.

World Rainforest Movement (2010) Hiding monoculture oil palm plantations under a business-friendly "forest" definition. *WRM Bulletin No. 151*, February 2010. Montevideo, Uruguay: World Rainforest Movement. http://www.wrm.org.uy/bulletin/151/oil_palm_plantations.html

Wunder S (2001) Poverty alleviation and tropical forests – What scope for synergies? *World Development* 29: 1817–1833.

WWF (2011) *Orangutans and Oil Palm Plantations*. http://wwf.panda.org/about_our_earth/about_forests/deforestation/forest_conversion_agriculture/orang_utans_palm_oil/ (accessed 15 February 2011). Gland, Switzerland: WWF International.

Yusoff S (2006) Renewable energy from palm oil – An innovation on effective utilization of waste. *Journal of Cleaner Production* 14: 87–93.

Yuwono EH, Susanto P, Saleh C, Andayani N, Prasetyo D, and Atmoko SSU (2007) *Guidelines for the Better Management Practices on Avoidance, Mitigation and Management of Human–Orangutan Conflict in and around Oil Palm Plantations*. Jakarta, Indonesia: WWF.

Zeven AC (1972) The oil palm in West Africa. *Economic Botany* 26: 274–279.

Origin and Diversification of Angiosperms

William L Crepet, Cornell University, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bennettitales Synonymous with Cycadeoidales, an order of extinct gymnosperms composed of two families having both unisexual and bisexual reproductive structures.
Gnetales A small relict order of chiefly tropical or xerophytic gymnosperms that includes the genera *Gnetum*, *Welwitschia* and *Ephredra*.
Homology Similarity due to common ancestry.
Phylogeny The evolutionary history of species and higher taxa.
Synapomorphies Shared characteristics that define a cladistic grouping or clade.

Introduction

Only a few years ago, a meaningful understanding of angiosperm relationships and their diversification might have seemed out of reach and questions swirled around the origin with little focus. Stunning advances in molecular systematics (e.g., Palmer and Zamir, 1982), including those techniques that now allow rapid analysis of large molecular datasets (the "ratchet," Nixon, 2003), have resulted in remarkably well-resolved assessments of within-angiosperm relationships. These advances accompanied by a much improved fossil record of angiosperms now also allow us to estimate timing in their diversification. The origin of angiosperms, however, is another matter. Just a few years ago, the angiosperms were comfortably nested in a clade called "anthophytes" with living Gnetales and extinct Cycadeiodales (Bennettitales) (Friis *et al.*, 2009). The same advances in molecular systematics that have provided valuable insights into within-angiosperm relationships and timing in angiosperm history have challenged the integrity of the anthophytes removing Gnetales to the rest of the gymnosperms. Cycadeoidales, on a closer look, may also be more appropriately grouped with other gymnosperms (Crepet and Stevenson, 2010). Thus, new methods have helped to focus our questions on angiosperm ancestry but have not answered them. There remains a gulf between angiosperms and their possible ancestors that is populated by fossil taxa with arguable transitional affinities based on logical but inconclusive assertions of connecting homologies. Nonetheless, new breakthroughs in the molecular genetics of development and exciting, if controversial, new fossils are improving the prospects for a near-term solution to what is debatably one of evolutionary biology's greatest remaining mysteries.

In summarizing our state of understanding of angiosperm origins and diversification, it is important to note at the onset that these areas of inquiry remain open to controversy even as exciting new modes of investigation and analysis have become available. Perhaps because of the importance and scope of the questions that remain unresolved concerning angiosperm relationships and radiation, these questions often inspire passionate advocates of differing points of view. Although a consensus is being approached on certain aspects of angiosperm history, and relationships among the angiosperms are better understood, others are open to controversy and change in the face of an increasing pace of investigation. Thus, this may be regarded as a status quo summary and change in our understanding of these important biological questions is highly probable in the near future.

Flowering Plants and the Abominable Mystery

The flowering plants are preeminent among vascular plants in numbers of species (250,000–300,000), they define terrestrial ecosystems at most latitudes, and exhibit more morphological diversity than any other group of plants. Economically, they are of primary importance as food, fodder, fiber, and building materials, in addition to serving as drug sources. It is thus interesting that, of all plant groups, they have the most elusive evolutionary history and, until recently, perhaps have been the most poorly understood with respect to their interrelationships.

The uncertainty surrounding angiosperm origin extends to the nineteenth century when early paleobotanical studies relied on gross morphological similarities in assigning affinities to angiosperm fossils. This approach resulted in skewed identifications, giving the appearance that apparently modern taxa were present relatively early in the fossil record, and suggesting a rather abrupt appearance of modern angiosperms. This apparent burst of modern taxa without any identifiable preceding ancestral lineage led Darwin to observe, in a now famously hackneyed quote from a letter to Heer, that the origin of angiosperms was both an "abominable mystery" and "a perplexing phenomenon" (Darwin, 1903; Friedman, 2009).

Since the nineteenth century, approaches to understanding the origin and diversification of the angiosperms have depended on paleobotany, systematics methodologies, and molecular genetics successively and, as they became available, in combination. New developments in each of these areas have increased our understanding of angiosperm diversification and relationships (Crepet and Niklas, 2009), but interestingly, have deepened the uncertainty surrounding angiosperm origin. Even with valuable new techniques and methodologies, the angiosperm fossil record remains central to understanding their origin. Beyond certain generalizations, ancestral types cannot be reconstructed from phylogenetic

trees either by character optimization on morphology/structure-based trees or by inference from trees based on molecular data alone (gene sequences). Yet, although the fossil record holds the best potential for identifying angiosperm ancestors, it has failed to resolve questions surrounding angiosperm origin. In addition, at its face value – that is based on the totality of literature on the angiosperm fossil record – it has been proven unreliable for documenting angiosperm diversification because, historically, the purported affinities of many angiosperm fossils have not been based on sufficiently rigorous analyses and are unreliable in estimating the timing (Crepet *et al.*, 2004).

Studies of Diversification

One must be selective in attempting to derive a reasonably accurate picture of angiosperm diversification from published fossil evidence. Criteria used to identify the affinities of particular fossil taxa should conform with those used in establishing the relationships among the living ones today, that is, fossils must have a sufficient number of key characters (synapomorphies) to warrant identification via phylogenetic analysis. Developments in both paleobotanical and systematic studies have made it practical to employ these criteria in confirming fossil affinities. Increasingly careful comparative analyses of angiosperm fossils began in earnest in the 1960s with emphases on leaf venation, epidermal features, and pollen characters (Dilcher, 1974; Hickey, 1973; Walker and Doyle, 1975; Hickey and Wolfe, 1975). In the mid-1970s, fossil flowers, important because of their taxonomic and reproductive implications, became serious and sustained objects of study, resulting in a more informative fossil record of angiosperms and a better understanding of their diversification pattern (Crepet *et al.*, 2004; Crepet, 2008, Crepet and Niklas, 2009).

Studies of Angiosperm Origin

In recent years, new analytical methods, new data sources, and newly discovered fossils have brought us much closer to an understanding of the origin of angiosperms. Nevertheless, the origin of angiosperms remains unresolved. For example, even though there has been a breakthrough in identifying a number of genes controlling floral development in angiosperms and some of these genes have been identified in nonangiospermous taxa, such discoveries have not yet provided definitive insights into angiosperm origin. So far, there have been no successful attempts to reconstruct transitional morphologies linking angiosperms to gymnospermous ancestors based on our understanding of genes controlling floral development. Although phylogenetic studies using molecular data have refined our understanding of the relationships among and within all major clades of plants, including angiosperms, they have also separated gymnosperm and angiosperm lineages at the base of the phylogenetic tree of extant seed plants (e.g., Chaw *et al.*, 2000). In doing so, these studies have distanced modern Gnetales (composed of the genera *Gnetum*, *Welwitschia*, and *Ephedra*) from the angiosperms. The Gnetales are now considered to be more closely allied to other gymnosperms despite the shared suite of morphological/structural characters that was once interpreted as evidence of a close relationship with angiosperms (e.g., Crane, 1985; Doyle and Donoghue, 1986; Nixon *et al.*, 1994; Rothwell and Serbet, 1994). This complexity has further deepened the mystery of angiosperm relationships with other seed plants (angiosperm origins).

This separation between angiosperms and Gnetales is both interesting and perplexing because of the apparent conflict between the previous morphological analyses that consistently grouped Gnetales with angiosperms in the anthophytes (Crane, 1985; Doyle and Donoghue, 1986; Nixon *et al.*, 1994; Rothwell and Serbet, 1994; Rothwell *et al.*, 2009) and those more recent analyses based on gene sequence data that separate all gymnosperms from angiosperms early in the diversification of the seed plants (e.g., Chaw *et al.*, 2000; Mathews, 2009). Given the implications of nucleic acid sequence based phylogenies, the characters in Gnetales that putatively linked them with angiosperms (e.g., including net-veined leaves, a form of double fertilization, bisexual reproductive structures (**Figure 1**), the presence of vessels in the wood, and sterile structures surrounding ovules) are now



Figure 1 *Gnetum gnemon* pollen-bearing organs. Note that whorls of synangiate pollen-bearing organs subtend whorls of (sterile) ovules illustrating the juxtaposition of structurally male and female reproductive organs in the Gnetales. Photo courtesy of Nixon, KC (2006). Image of *Gnetum Gnemon*, http://www.PlantSystematics.org, image ref.: DOL10045 (accessed November 2006).

widely regarded as parallel developments. Yet, in the face of apparently overwhelming evidence from numerous molecular analyses, there are stalwarts who still advocate the relationships based on morphology trees, and new fossil evidence has been cited in support of the anthophyte concept (Friis *et al.*, 2007). However, more recent analysis of the data presented by these authors does not add support to the anthophyte clade concept (Rothwell *et al.*, 2009). At the same time, comparative studies of within-modern angiosperm relationships have also failed to illuminate angiosperm ancestry because the modern representatives of the earliest branching lineages of angiosperms, *Amborella* or the Nymphaeales, depending on the analysis, do not have characters that provide insights into the relatives of angiosperms (i.e., they do not share significant characters with any nonangiospermous taxa). Given that no living taxon can be identified as linking angiosperms and other seed plants, and that the reconstruction of possible links based on our existing understanding of character distribution and genes controlling floral development have not been possible, hope for a solution remains with the fossil record. Fossil evidence, however, is problematic and as with comparative studies of extant taxa, there are no obvious fossil intermediates between the gymnosperms and the angiosperms, leaving the field open to speculation and a variety of competing hypotheses.

Why has the Fossil Record Failed So Far?

Why has the fossil record failed us with this most important group of plants despite the fact that it is replete with transitional mosaics (missing links) between other groups of vascular plants? There are no generally accepted fossil intermediates between angiosperms and other seed plants, and attempts to identify fossils that represent extinct taxa transitional to angiosperms are complicated by a number of issues. A traditional difficulty has been attempting to identify an angiosperm sister group on the basis of the preserved morphological/structural characters. A weakness to this approach lies in the possibility that underlying genetic changes could have been involved in such transformations without leaving a sequence of morphologically expressed connecting intermediates.

Thus, from the onset one must acknowledge the possibility that there were no morphoclines leading from gymnosperms to angiosperms, that is, that actual intermediates may not be recognizable on morphological grounds even if discovered in the fossil record. If so, clarification of angiosperm origin must await advances in understanding more completely the genetics underlying floral development, factors controlling gene expression, and the distribution of these genes in existing seed plants. But assuming that intermediacy was expressed in a morphological transition to angiospermy, and that fossils representing transitional taxa have either been undiscovered or gone unrecognized, there are a number of relevant questions, including: What subset of characters is required to reasonably identify a potential angiosperm sister group and for that matter what subset might be necessary even in identifying a very early angiosperm? What combinations definitely indicate a relationship to the angiosperms? And, given that preservation is most often incomplete limiting the number of available characters, which characters are definitive indicators of angiosperms? Given these questions, it has been traditionally difficult to avoid subjective judgments in evaluating the significance of fossils that are possible angiosperm ancestors. However, questions regarding relationships of fossils can now be addressed more objectively by phylogenetic analysis, a methodology that precisely determines relationships among taxa and purportedly minimizes the influences of subjective interpretation. However, even with the advent of phylogenetic analysis, there has been no consensus on the closest fossil relatives of angiosperms.

There are a number of plausible reasons for this disappointing situation. Firstly, in addition to the issues cited above (*see* Introduction), there are general problems inherent to fossil evidence: the record is incomplete, reflecting variation in optimum conditions for fossilization through time, and is undersampled. In addition, there have been relatively few paleontologists and available fossil sites for investigation. Furthermore, the fossil record has been poorly sampled in the time periods likely to be important in discovering extinct angiosperm relatives (Jurassic–Early Cretaceous). In addition, given the morphological gap between angiosperm reproductive structures and all the known gymnospermous ones (fossil and modern), and in the absence of a generally accepted transformation series (or morphocline) linking key nonangiospermous reproductive structures with angiospermous ones, it is difficult to identify characters in nonangiosperms that can be regarded as homologous with angiosperm defining characters. For example, based on morphological transformation models, multistepped and tortuous transformations are necessary to support hypotheses linking various (and virtually all possibly ancestral) fossil taxa to the angiosperms (e.g., *Cycadeoidea* (Bennettitales), Corystospermaceae, Glossopteridales, and Caytoniaceae (a good example is in the sometimes hypothesized transition from the caytonialean cupule to anatropous ovule + carpel of angiosperms)). Thus, although phylogenetic analysis has provided a powerfully objective tool for determining relationships, the assignment of character equivalencies needed for phylogenetic analysis (the determination of homologies) remains highly subjective in the absence of unequivocal links and prone to the influence of the currents of tradition, and so there is no consensus on extinct angiosperm relatives (Taylor and Taylor, 2009).

Fossil Evidence – Status Quo

What is the status quo in our understanding of fossil evidence? In approaching the fossil record, as with evaluations of extant taxa that might be germane to angiosperm origin, we might consider what to look for in addressing the question of angiosperm origin and then consider appropriate fossils and the implications of these fossils. In general, there are two possible subgroups that would be relevant to angiosperm ancestry: fossil taxa that are clearly not angiosperms, but that have synapomorphies (shared derived characters) that may link them to angiosperms; and early angiosperm fossils that may possess characters linking them to nonangiospermous ancestors.

Potential Angiosperm Sister Groups (Ancestors)

In the first category above (nonangiosperms), known fossils include principally:

The Mesozoic seed fern *Caytonia* (**Figure 2**). *Caytonia* was originally described as an angiosperm because the ovule-enclosing cupules were confused with enclosed carpels (Thomas, 1925). Even after the discovery that the cupules were not closed led to the conclusion that *Caytonia* was a seed fern (Harris, 1940, 1964), the taxon continued to be linked with angiosperms. Some authors felt this because of a number of similarities such as net-veined leaves (although *Caytonia* veins are not hierarchical by size like those of the angiosperms), stomatal subsidiary cell configuration (paracytic – a generalized character occurring in a number of taxa) and, most significantly, by a superficial similarity between the adaxially incurved cupules containing numerous seeds and the anatropous ovule characteristic of many, but not all angiosperms (interestingly not found in *Amborella*, the taxon at the base of most recent angiosperm phylogenies). The proposal equating the cupule of *Caytonia* with an anatropous angiosperm ovule posits that the cupule itself represents the second integument of the angiosperm ovule (even though not all angiosperms are bitegmic), and requires several additional morphological changes, including the reduction of ovules within the cupule to one, and the *de novo* origin of the carpel. Thus, beyond what has been observed in *Caytonia*, there are several steps needed for the transformation, and yet there is no fossil evidence of any intermediates to support the hypothesized transformation. Furthermore, the combination of ovules with adaxially disposed micropyles and subtending sterile structures evocative of anatropy is not unique to the Caytoniaceae in gymnosperms or even among Mesozoic seed ferns (Taylor and Taylor, 2009) Thus, although plausible and possible depending on underlying genetics, this proposed homology is not compelling enough to be regarded as solving the issue of angiosperm origins, especially given the distinctly pinaceous saccate pollen found in *Caytonia* (adding additional steps to any transformation).

The Cycadeoidales (or Bennettitales, **Figure 3**). This is the other major fossil group often associated with angiosperms. One of the principal genera of the order, *Cycadeoidea* (**Figure 3**), was thought to be a link between angiosperms and gymnosperms when it was first found to have bisexual reproductive structures (**Figure 3**). Subsequently, its relationship to angiosperms has been debated. Although the Cycadeoidales are part of the anthophyte clade recognized in morphology-based phylogenetic analyses, the integrity of the group is in question based on molecular systematics analyses and recent reanalyses of morphology-based trees challenge the position of Cycadeoidales in an anthophyte clade (**Figure 4**; Crepet *et al.*, 2011; Crepet and Stevenson, 2010). Nonetheless, as with *Caytonia*, there are some appealing similarities between angiosperms and Cycadeoidales. The cycadeoidalean genera *Cycadeoidea*, *Wielandiella*, and *Williamsoniella* each have bisporangiate cones with an ovulate receptacle subtended by variously pinnate microsporophylls bearing synangiate pollen organs that are in turn subtended by sterile bracts. The congruence between this arrangement of organs and the succession found in the typical angiosperm flower (carpels surrounded by stamens surrounded by sterile petals/sepals) raises the question of affinities: are the similarities based on homology or parallel evolution?

These similarities are even more compelling in light of other shared characteristics including some pollen characters, stomatal subsidiary cell configuration, and dicotyledonous

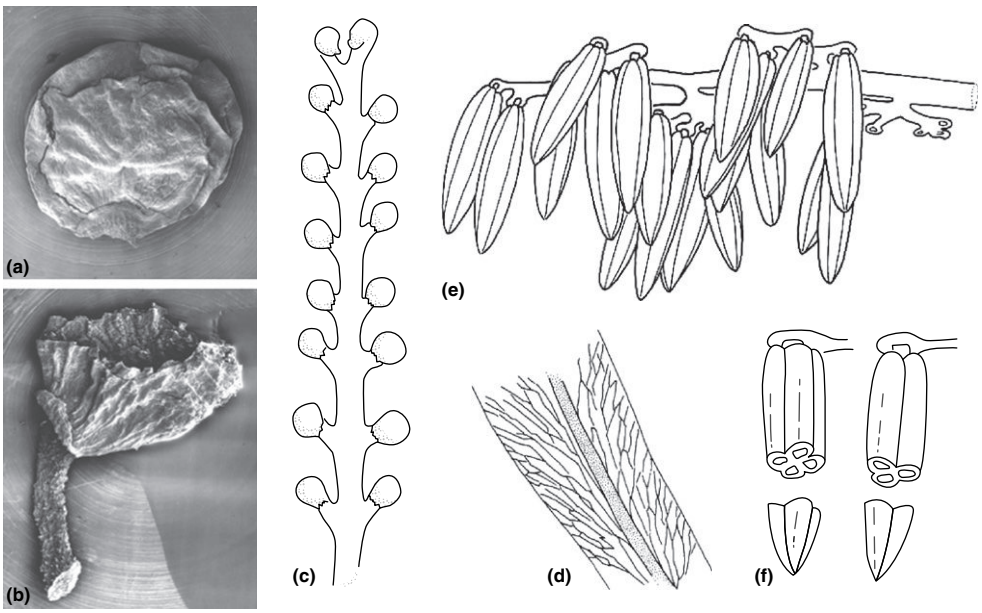


Figure 2 (a)–(c) *Caytonia*. (a), (b) SEM photographs of the seed-bearing cupules. (c) Line drawing of cupule-bearing axis. (d) *Caytonia* leaf (*Sagenopteris*) venation pattern. (e), (f) *Caytonia* pollen-bearing organs (*Caytonanthus*). (e) A synangia-bearing frond. (f) Several synangia showing that they are composed of four fused and radially arranged pollen organs. (c), (d) Redrawn from Thomas HH (1925) The caytoniales, a new group of angiospermous plants from the Jurassic rocks of Yorkshire. *Philosophical Transactions of the Royal Society of London* 213B: 299–363, and (f) redrawn from Harris TM (1937) The fossil flora of Scores by Sound, east Greenland. Part 5. Stratigraphic relations of the plant beds. *Meddeleser om Grønland*, *Kjobenhavn* 112(2): 1–114.

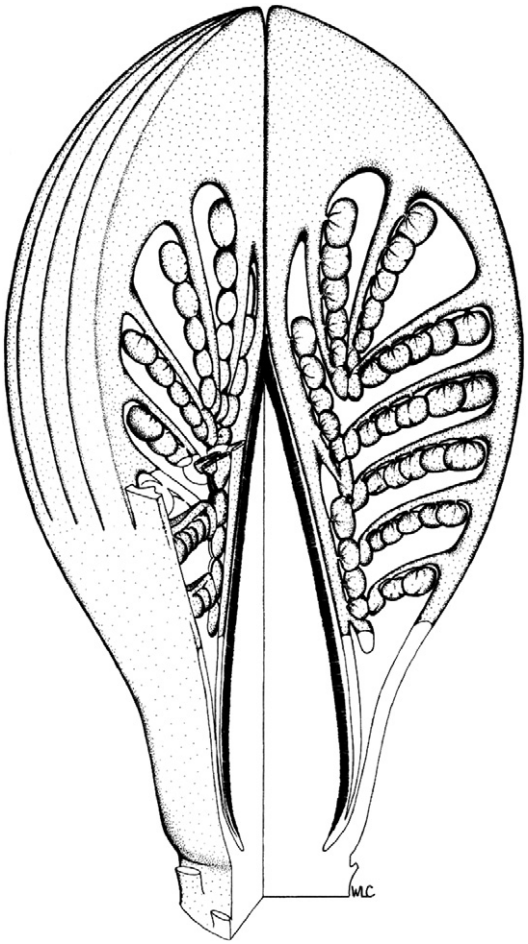


Figure 3 A reconstruction of the bisporangiate cone of the genus *Cycadeoidea* showing pollen-bearing organs subtending a conical ovule-bearing receptacle. This arrangement of pollen-bearing organs and ovules is unique to cycadeoids, angiosperms, and is approached in Gnetales.

embryos, although such embryos are not unique to angiosperms and cycadeoids and are not even ubiquitous within angiosperms, and, transformation of a cycadeoid flower would be complex and once more require the *de novo* evolution of a carpel, dramatic reduction of the microsporophylls and a change in symmetry in both pollen organs and ovule-bearing structures. Moreover, careful comparisons between Cycadales and Cycadeoidales reveal more similarities and fewer differences than previously recognized in analyses leading to the identification of an anthophyte clade. Recent analyses that incorporate these observations suggest closer ties between Cycadales and Cycadeoidales in a clade that also includes Mezozoic pteridosperms (Crepet and Stevenson, 2010) (**Figure 4**). Yet, even if Cycadeoidales turned out to be more closely related to Cycadales than previously thought as indicated by this analysis, a link between these lower seed plants and angiosperms might not be out of the question in the context of the early split in seed plants between the angiosperm and gymnosperm lineages, that is, any group at the base of gymnosperms could plausibly be the one most closely related to the angiosperms.

Corystospermaceae: These Mesozoic seed ferns have been the focus of an attempt to synthesize knowledge of MADS Box floral development genes, and fossil evidence in hypotheses on angiosperm origin that suggest ways that corystosperm pollen-bearing organs might be transformed into bisexual angiosperm flowers. This hypothesis draws attention to the importance of underlying genetics and to the possibility that transitional morphologies have not been discovered in the fossil record because they never existed. This general approach will become more important as details of genes controlling floral development are increasingly revealed and when there is a better understanding of the distribution and nature of these genes among extant taxa (Frohlich and Parker, 2000).

Glossopteridales: The seed-bearing structures of glossopterids were known for a long time from only relatively poorly preserved impression fossils. These fossils were alluringly carpel-like in appearance and quite variable in morphology. They are now more completely understood based on careful studies of well-preserved petrified fossils from Australia and Antarctica (e.g., Gould and Delevoryas, 1977). Although early suggestions that glossopterids might have been related to the angiosperms were based on putative homologies between the seed-bearing organs of *Glossopteris* and the carpels of angiosperms and on their superficially common net-veined leaves (Retallack and Dilcher, 1981), glossopterids have not been placed near angiosperms in comprehensive phylogenetic analyses (e.g., Crane, 1985; Doyle and Donoghue, 1986; Nixon *et al.*, 1994) (**Figure 4**).

In summary, there is no clear consensus on the closest fossil relatives of angiosperms, with unavoidably subjective assessments of homology clouding the results of the various, sometimes conflicting phylogenetic analyses. As it stands, there are a number of possibilities that cannot be definitively ruled out because the genetics underlying the development of key reproductive structures is poorly understood and available fossil (and other) evidence does not favor one fossil taxon over another (e.g., Taylor and Taylor, 2009).

Early Angiosperms – Imagined and Real

The second category of potentially informative fossils relevant to angiosperm origin, that is, early putative angiosperm fossils, includes a number of taxa. Some of the better known ones are interesting, but they are not understood well enough to be unequivocally regarded as angiospermous because the fossils have too few characters to be convincingly placed with angiosperms in phylogenetic analysis. Notable among these fossils are several of the Triassic age, and thus they are older than the earliest generally accepted angiosperm fossils (these are Lower Cretaceous). Such fossils include *Furcula*, an organ taxon for dichotomous leaves with net venation (Harris, 1932), isolated pollen grains with angiospermous pollen characteristics (Cornet, 1989a), and a taxon first described from leaves (*Sanmiguelia*) but later descriptions included purportedly related reproductive structures (*Axelrodia*) as well as almost entire plants (Brown, 1956; Cornet, 1986, 1989b; Tidwell *et al.*, 1977). *Sanmiguelia* is intriguing because the plicate fan-shaped leaves are similar to those of some extant monocots (e.g., Liliaceae). The fossils, however, are poorly preserved and the

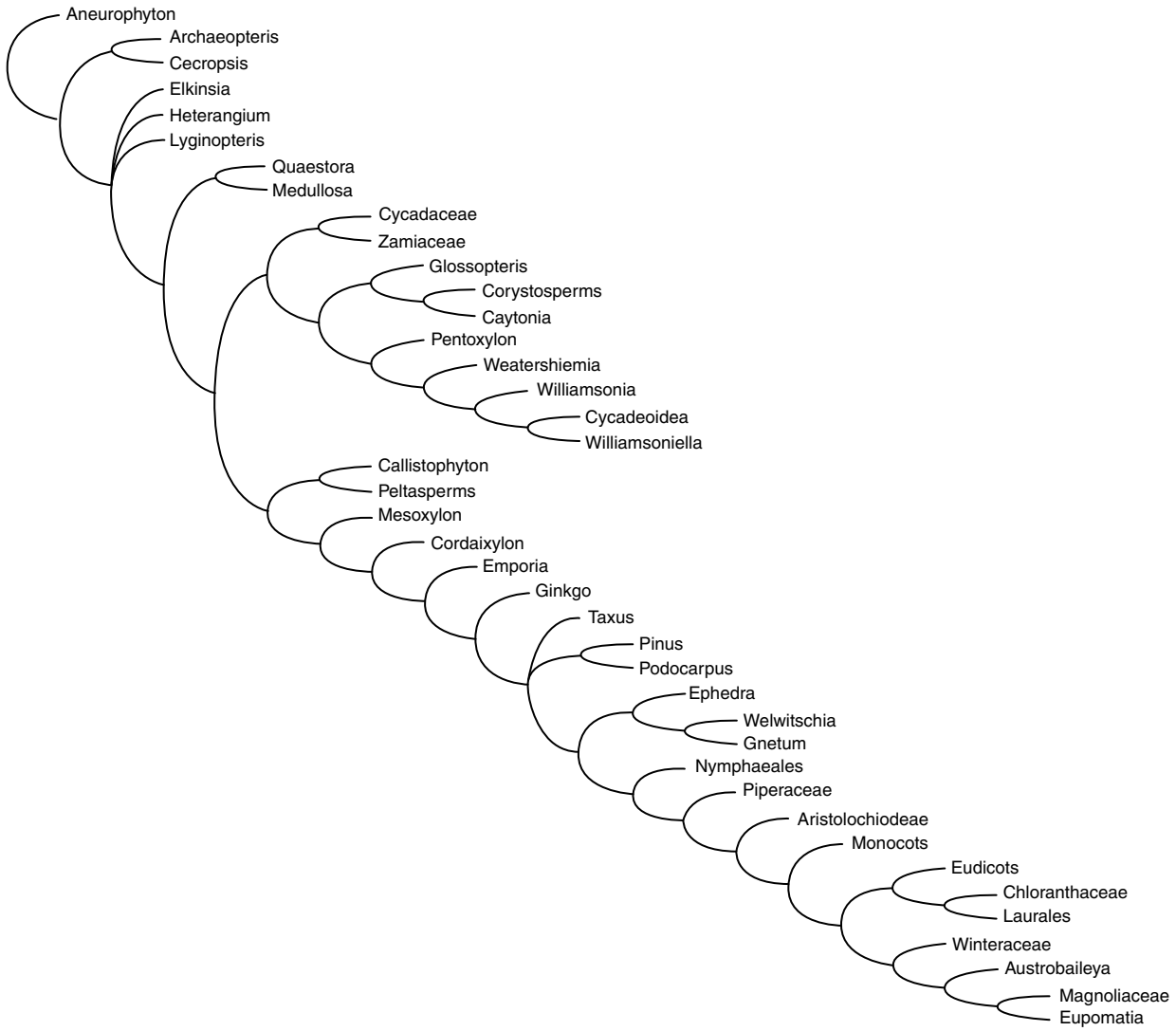


Figure 4 Phylogeny of seed plants based on a revised and expanded morphology matrix illustration the position of Caytoniales, Glossopteridales and the instability of the Cycadeoidales (Bennettitales).

reproductive structures have been unconvincingly reconstructed. In the absence of sufficient data and analyses, there is no consensus on their significance at this time.

A relatively recently discovered fossil, *Archaefructus*, is one of the most potentially important of the known fossil angiosperms. Although there is no doubt of its angiospermous nature, there is a debate as to its phylogenetic position. The question at issue: is it the sister group to all remaining angiosperms or is it nested within an extant taxon? (Friis *et al.*, 2003; Endress and Doyle, 2009). *Archaefructus* is significant because it is completely preserved with reproductive organs attached to the vegetative part of the plant and because it comprises an unusual set of characteristics. Initially reported as Jurassic in age, it is now regarded as Lower Cretaceous and although early for an angiosperm, it is not uniquely the oldest of generally accepted angiosperm fossils. From specimens preserved in their entirety, it is obvious that *Archaefructus* was an aquatic plant (the leaves have air bladders), and that it lacked the typical organization of angiosperm flowers because naked carpels are only subtended by stamens along elongate axes. There is no evidence of any kind of floral envelope. *Archaefructus* has been interpreted both as an angiosperm that is a sister group to the rest of the angiosperms (Sun *et al.*, 1998) and as an aberrant member of the Nymphaeales, morphologically distinct from any modern member of that group with flowers that are so highly reduced that there is no evidence of usual angiosperm floral organization (Friis *et al.*, 2003). The latter interpretation is based on a phylogenetic reanalysis of the Sun *et al.* (2002) character matrix (Friis *et al.*, 2003). However, this reanalysis included a mistake in coding of a significant character (leaves) in a taxon that was added to the original analysis (*Cabomba*). When the matrix is corrected to reflect actual variation in *Cabomba* leaf morphology, subsequent analyses uniformly indicate that *Archaefructus* is the earliest branch in the angiosperm lineage. Once again it becomes a sister group to the angiosperms in 100% of the shortest trees. It is never placed with *Cabomba* or in the Nymphaeales, thus conforming with the results of Sun *et al.*

(2002) and supporting original interpretations of its reproductive structures as innately simple and not simple owing to the process of evolutionary reduction that might have been associated with adaptation to an aquatic environment. In another recent analysis (Endress and Doyle, 2009), *Archaefructus* also comes out with Nymphaeales, but the analysis is fragile and there are coding issues that call the results into question that, when corrected, again result in a phylogeny placing *Archaefructus* at the base of the angiosperms (Crepet *et al.*, 2011). Even so, it is likely that more characters (details of pollen structure, etc.) and subsequent supporting analyses will be needed to achieve a consensus view on the phylogenetic significance of *Archaefructus*. If confirmed as representing the earliest branch in the angiosperm lineage, *Archaefructus* would provide some unique insights into the nature of early angiosperms, but based on our present understanding of its morphology, would not unambiguously link angiosperms to any group of nonangiospermous seed plants. In general aspect, *Archaefructus* is evocative of seed ferns, but the gap between known seed fern reproductive structures and those of *Archaefructus* is significant, and evidence for transformations such as those invoked in the putative transition from *Caytonia* to an angiosperm is missing from the fossil record.

Should additional characters and analyses continue to support the phylogenetic placement of *Archaefructus* on the stem lineage below all other angiosperms, the implications of the aquatic habit in *Archaefructus* and questions as to whether an aquatic bottleneck was involved in the consolidation of angiospermous reproductive features will undoubtedly be pursued (e.g., Did carpels evolve in response to selection for protecting developing ovules while still submerged? Is insect pollination more efficient at the water level? What are the implications for identifying possible angiosperm ancestors? Were there earlier stem lineage angiosperms that were not aquatics?).

Conclusions – Angiosperm Origin

There is no convincing evidence as to the closest group of nonangiospermous seed plants living or fossil. Recent developments have raised questions and revealed aspects of seed plant relationships that narrow the search and pose questions that can be foci of more intense investigations. Fossil studies have the potential for identifying the ancestral angiosperm lineage, especially in the context of molecular genetics, but, in spite of exciting new developments, the issue of angiosperm origin remains unresolved.

Angiosperm Diversification

Analytical Techniques Relevant to Understanding Diversification

Ideally, increasingly complete and reliably identified fossil evidence and advances in molecular systematics can be combined with the aim of improving our understanding of angiosperm diversification. Although timing was once taken at face value from the fossil record, today we recognize that such an approach is compromised by a literature that is replete with misidentifications based largely on superficial comparisons between fossil angiosperms and the corresponding organs of modern angiosperm taxa. A selective approach to fossil evidence focusing on accurately identified fossils would improve our understanding of angiosperm history and would provide a better understanding of the diversification of the angiosperms than an uncritical review of the literature. However, the fossil record is incomplete and such a strict interpretation of the literature would leave many gaps in our understanding of the timing of events in angiosperm radiation. The advent of molecular systematics provides better ways to estimate both diversification pattern and timing, especially now that it has engendered a growing consensus on the relationships among modern angiosperm taxa. There are a number of ways to combine phylogenies based on gene sequence data and the fossil record of angiosperms to reveal timing in angiosperm diversification. Different methods have produced different estimates and a lively debate is underway as to which promises to be the most accurate. A "consenus phylogeny" may be used either with selected fossil evidence in molecular clock-derived models (e.g., Sanderson, 1997; Sanderson *et al.*, 2004) or alternatively, in conjunction with the carefully screened entire fossil record of angiosperms in the "minimum-age node mapping" methodology (Crepet *et al.*, 2004). Depending on the details of the applications, the two methods can produce dramatically contrasting estimates of timing in angiosperm history or within particular clades. Depending on the model and the analysis, there is considerable variation in the estimation of timing in angiosperm history, and some of this variation is related to the specific methods used to calculate timing and some to the choice of fossils used in calibration. For these reasons, estimates based on molecular clock-derived models have been subjected to intense scrutiny and criticism (Graur and Martin, 2004), but these clock-based methods are still evolving with different approaches showing promise (Sanderson *et al.*, 2004; Lavin *et al.*, 2005). The results of certain analyses are roughly congruent with timing suggested by fossil evidence alone, but timing estimates based on other analyses project much greater ages for various lineages or for the entire angiosperms than those revealed by the fossil record (Savolainen *et al.*, 2000). These discrepancies in timing angiosperm diversification have significant implications because differences in perceived timing affect our understanding of important events in angiosperm evolution and therefore of factors that may have been relevant to angiosperm success (e.g., possible synergy with insect pollination, animal fruit dispersal, and climatic conditions). Such differences also affect our understanding of angiosperm biogeography (e.g., timing is important in evaluating the potential significance of various land bridges to angiosperm biogeography or in estimating the effects of vicariance; Tiffney, 2008).

One of the major criticisms of analyses that have predicted greater ages for taxa than those verifiable by fossil evidence is that these discrepancies have not been satisfactorily explained except by the untested premise that the record itself lags the pattern of evolution, often dramatically. This notion is based on the assumption that there is a significant delay between the origin of a new taxon and its first appearance in the fossil record owing to the time it takes a new taxon to reach a population size and distribution that make deposition,

preservation, and discovery likely. This explanation of lagging fossil evidence is unconvincing because the time needed for population size to grow to the point where preservation (and, ultimately, fossil discovery) is likely is going to be miniscule relative to geologic time. Thus, discrepancies in timing on the scales projected by some clock-based models (tens of millions of years in some cases) seem unlikely given that a new taxon may have increased in population size and range to the point where preservation in the fossil record was likely in an interval that would be essentially instantaneous in the frame of geologic time (Burnham, 2008). There are cases where preservation seems unlikely in widely dispersed herbaceous taxa that are insect-pollinated, for example, and instances such as these are partially responsible for gaps in the fossil record that may be addressed by the methods discussed above (*see* Analytical Techniques Relevant to Understanding Diversification).

As molecular clock-based models continue to be refined and are becoming more accurate, an alternative and complementary way to combine fossil evidence with molecular systematics-generated phylogenies of living taxa already exists in "minimum-age node mapping" (Crepet *et al.*, 2004). Minimum-age node mapping methodology is based on the straightforward premise that, in phylogenetic context (and the success of the method depends on the framework of an accurate phylogeny for angiosperms made possible through molecular systematics such as that of Soltis *et al.*, 2000), a taxon's immediate ancestor (i.e., the taxon represented by the node proximally subtending the taxon) has to be as old or older than that taxon. By plotting reliably identified fossils on a consensus molecular systematics phylogeny, this methodology provides minimum times of appearance even for taxa that do not have fossil records (of course, such predictive power is also a goal of variously modified molecular clock models).

Although some clock-based estimates of timing in angiosperm evolution seem outlandish when compared with the fossil record, others, using reliably identified fossils and refined methodologies (e.g., Lavin *et al.*, 2005; Magallon-Puebla *et al.*, 1999; Sanderson *et al.*, 2004), have generated results that are more consistent with those of minimum-age node mapping. As increasing numbers of reliably identified fossil taxa are used in calibrating such models, and as the models themselves become further refined, results might be expected to converge with the fossil record of angiosperms and might also be effectively used for filling in gaps in understanding timing in particular clades that have relatively good fossil records (as in Lavin *et al.*, 2005). Currently, the results of node mapping methodology provide the most conservative assessment of timing in angiosperm radiation. Given the interest in improving molecular clock-based timing models, however, and their popularity (easy to use, appealing (especially rate smoothing) premise, they fill in the gaps in the record), assessments of timing are likely to be in flux and somewhat controversial in the near future.

Major Features of Angiosperm Diversification

As consensus is approached in estimates of the timing and pattern of angiosperm diversification, major events in the evolution and radiation of the flowering plants can be conservatively, but reliably, placed in a temporal perspective allowing for the evaluation of various ecological/evolutionary hypotheses that have been proposed to explain angiosperm success (Crepet and Niklas, 2009).

Early Angiosperm Diversification

Timing and diversification in angiosperm history is summarized in the phylogenies presented in **Figures 4–6** based on a subset of carefully identified angiosperm fossils plotted according to minimum-age node mapping methodology using the phylogeny of Soltis *et al.* (2000) and Crepet *et al.* (2004). The early radiation of bona fide angiosperms illustrated in **Figures 4** and **5**, began in, and then characterized the Lower Cretaceous. In **Figure 4**, the emphasis is more on basal angiosperms followed closely and overlapping temporally with the subsequent radiation of the remaining tricolpate pollen-bearing clades (the "eudicots") illustrated in **Figure 5**. The numbers in the figures represent ages of the clades in millions of years. Numbers on a terminal branch represent the age of the terminal taxon (e.g., Chloranthaceae 98). Terminals on unnumbered branches or clades without numbers (e.g., *Hedycarya* + *Peumus*) have the same minimum ages as the node below. *Archaefructus* is not included in the cladogram, but as a sister group to the rest of the angiosperms, a position now favored by the preponderance of available evidence, it would push back the timing of angiosperm radiation (still within the Cretaceous) based on bona fide fossil evidence but would not affect the timing in the rest of the phylogeny.

The estimates illustrated in **Figures 4** and **5** suggest extensive diversification in several important angiosperm clades during the Lower Cretaceous to Turonian interval that included magnoliids, monocots, and several tricolpate clades. The possibility of an Early Cretaceous monocot diversification is supported by the presence of Triuridaceae floral evidence in the Turonian (because the triurids represent a derived taxon), but Early Cretaceous monocot diversification is not reflected by the mega or mesofossil records. Although many reports of pre-Turonian monocot pollen are based on characters that are not exclusive to monocots (Gandolfo *et al.*, 2000), a recent report of dispersed but distinctive Araceae-like pollen (Friis *et al.*, 2004) suggests that at least some Early Cretaceous palynomorphs represent monocots, a finding that would be consistent with node mapping projections of monocot minimum ages (**Figure 5**).

The occurrence of numerous platanoid fossil leaves, flowers, fruits, and pollen among the oldest representatives of the "tricolpate" clade (or "eudicots" of some authors) at about 100 Ma is consistent with, and even slightly younger, than the age of the sister clade of the tricolpates + Ceratophyllum based on fossil evidence of winteraceous pollen (Walker *et al.*, 1983) from the Early Cretaceous of Israel (4105 MyBP).

Minimum-node mapping methodology, like molecular clock-based models, can calculate minimum ages for the stem groups of some families that lack a fossil record revealing that the minimum calculated ages for some families are greater than the ages of the oldest known fossils for the group. Sabiaceae, for example, has a calculated minimum age of 98 MyBP for its divergence from other extant angiosperms,

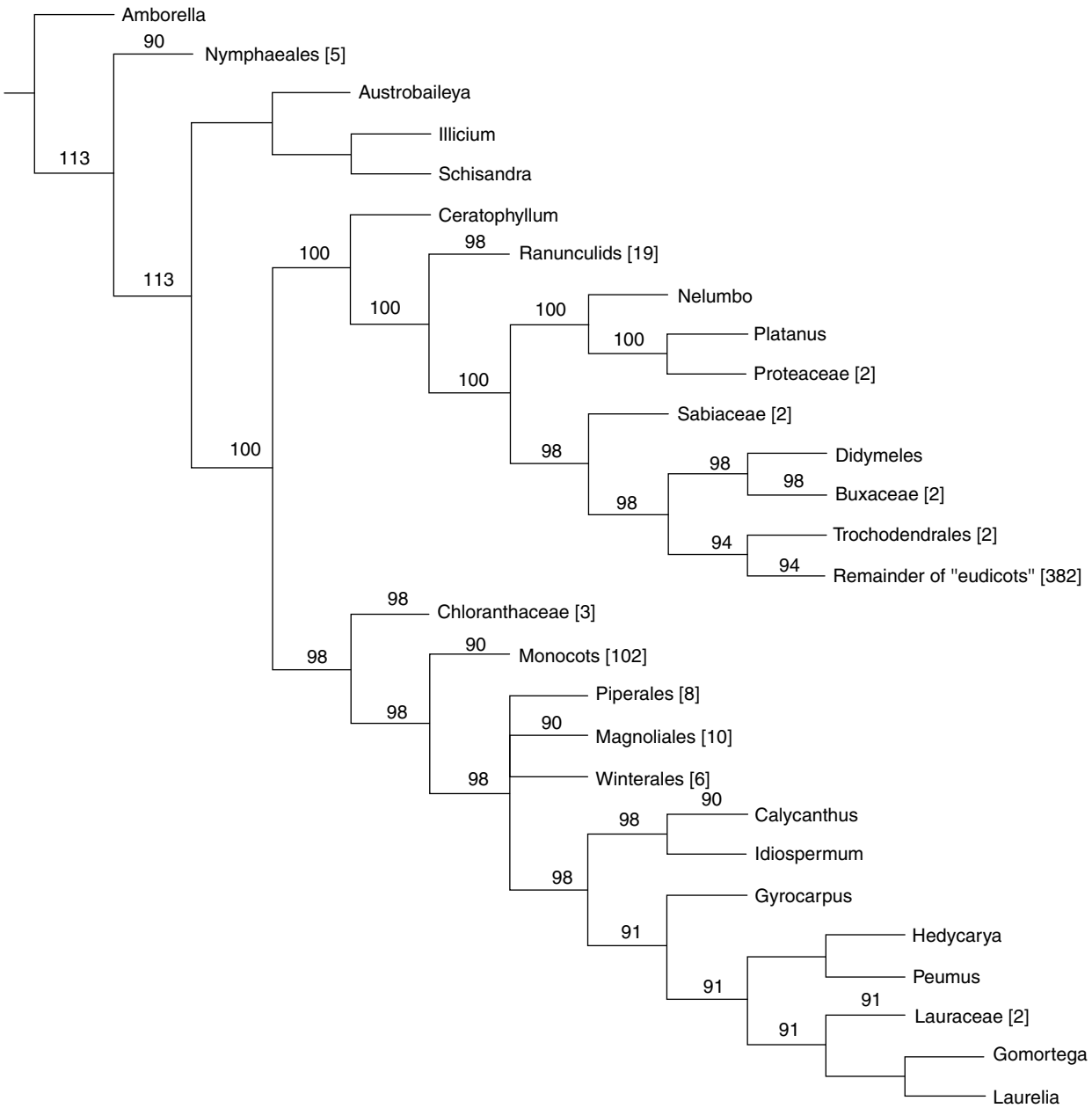


Figure 5 Minimum ages for major ''basal'' lineages of angiosperms (113–90 MyBP) based on the best available fossil evidence mapped onto the three-gene 560 angiosperm tree (Soltis *et al.*, 2000). The tree has been collapsed to show major clades of interest. Numbers in brackets following a name indicate the number of distinct sequences that were used in the original molecular analysis.

based on fossils assignable to other families (**Figure 5**), even though the oldest putative fossils known for Sabiaceae seeds from the Late Cretaceous, are at least 27 My younger (Knobloch and Mai, 1986).

As with all minimum ages for clades predicted by minimum-age node mapping or other (molecular clock-based) models, this does not imply that the ''crown group,'' in this instance modern Sabiaceae, had diverged by this time, but that the divergence of Sabiaceae from other extant angiosperms occurred at least 98 Ma. In fact, and again a good example for similar instances with other clades, the modern family may have diverged much more recently, and the perceived hiatus may be explained by numerous factors, including lack of identifiable features of modern Sabiaceae in the early members of the clade (e.g., as in legumes), preservational bias owing to morphology, mode of pollination, ecological distribution, or random events influencing the preservation of fossils (Crepet *et al.*, 2004).

Rosid 1 Clade Diversification

One of the most impressive aspects of angiosperm diversification is the rapid radiation of the Rosid 1 Clade – the clade

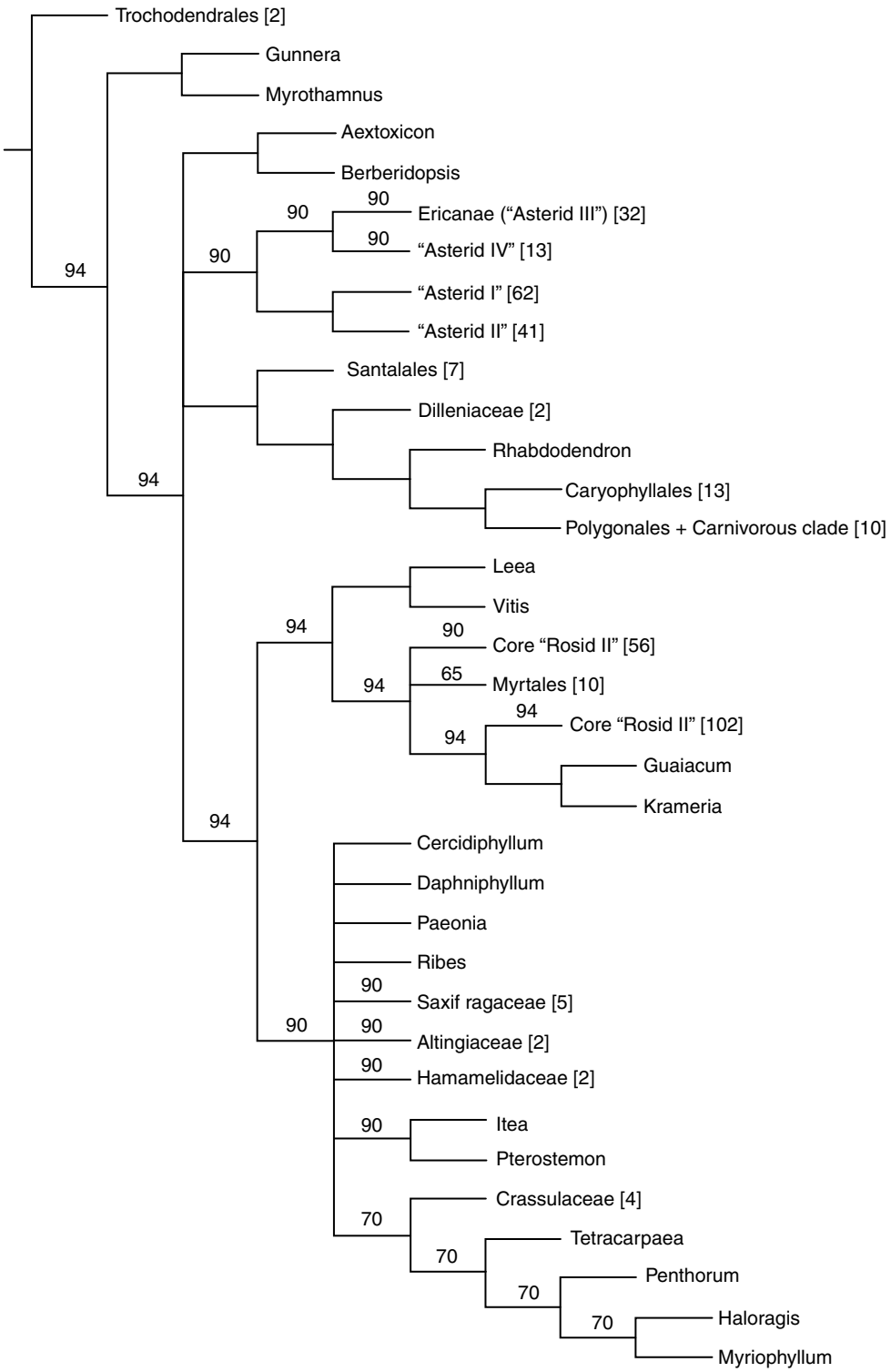


Figure 6 Minimum ages in million years before present for major lineages within the tricolpate or ‘‘eudicot’’ clade (some collateral groups not shown, such as Platanaceae, Buxaceae; see **Figure 5**).

that includes Rosaceae, Malpighiaceae, and a host of other families traditionally placed in either the Rosidae or the ‘‘Higher Hamamelididae’’ (**Figure 7**). Beginning at least by the mid-Cretaceous, Rosid 1 diversity was already well established by the Turonian with the existence of two firmly identified fossils in the lower part of this subtree (Cunoniaceae and Clusiaceae) allowing estimates of the minimum divergence times for the stem lineages of several families owing to the

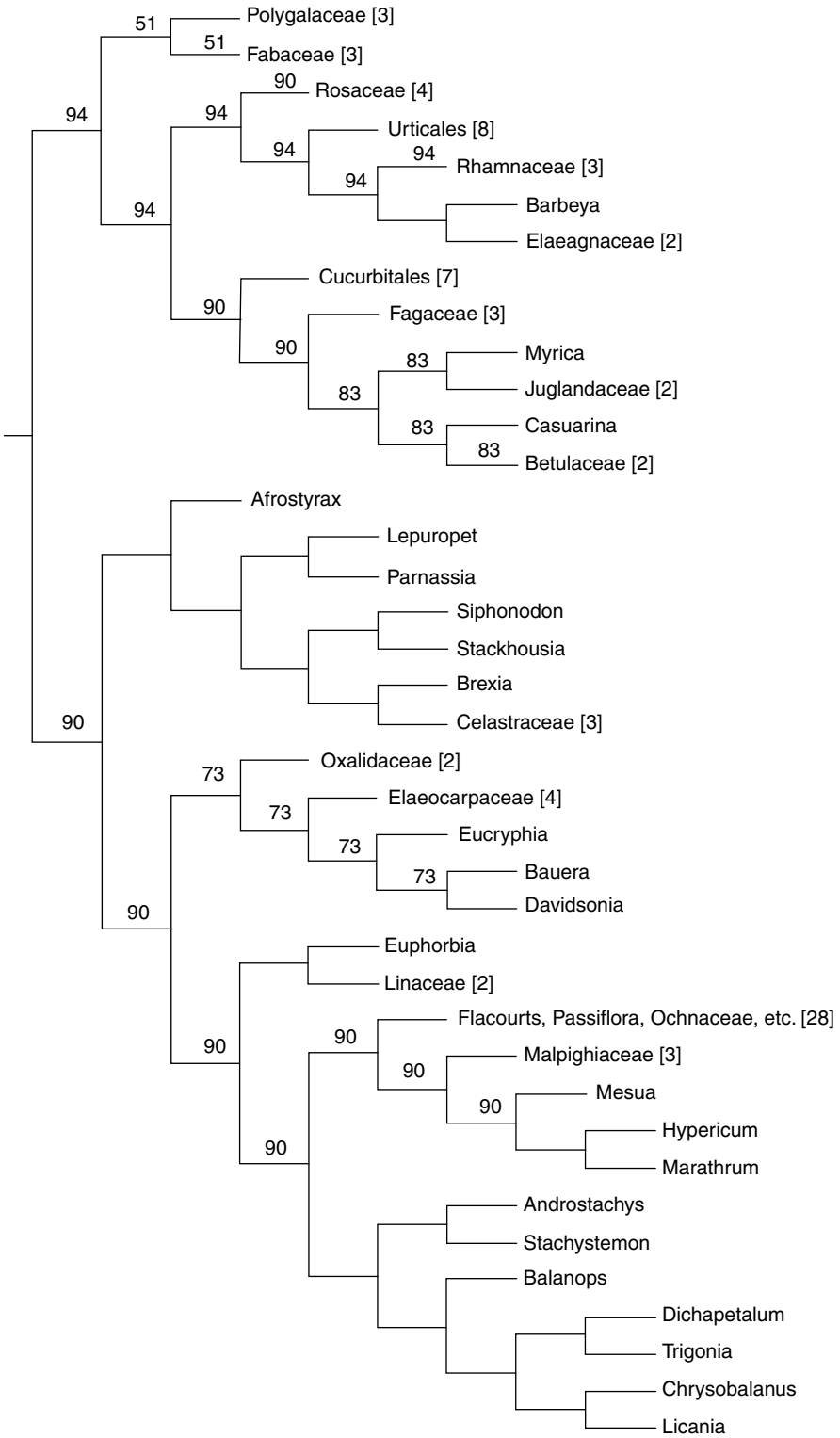


Figure 7 Minimum ages in million years before present of taxa within the "Rosid I" (sensu APG II, 2003) clade.

pectinate topology. The divergence of the stem groups for Oxalidaceae, Elaeocarpaceae s.s., Eucryphiaceae, and Cunoniaceae can be estimated at 73 MyBP, and thus, any modern species in those families has a minimum divergence of 73 MyBP from any other angiosperm species not in the same family. Likewise, a minimum age of 90 MyBP for the divergence of the stem group of Malpighiaceae can be estimated based on Clusiaceae fossils from the Turonian of New Jersey



Figure 8 Reconstruction of the fossil flower *Paleoclusia chevalieri* (Crepet and Nixon, 1998). Illustration courtesy of M. Rothman.

(**Figure 8**; Crepet and Nixon, 1998; also note **Figure 9** for an example of the kind of preservation that has improved our understanding of fossil floral history).

Another important clade within the "Rosid I" group includes Rosaceae, Fabaceae, Urticales, Cucurbitales, and the families formerly included in the "Higher Hamamelididae" (e.g., Fagaceae, Nothofagaceae, Juglandaceae, Myricaceae, Betulaceae, and Casuarinaceae). This clade includes a diverse assemblage of economically important taxa with an exceptional fossil record. The minimum age of the stem group of the clade that includes Rosaceae and Fagales but excludes Fabaceae and Polygalaceae is estimated at 94 MyBP, whereas the first identifiable fossil of Fabaceae is from the Tertiary, *ca.* 51 MyBP (Herendeen and Crane, 1992), or at least 40 My after the minimum estimated divergence for the clade. It should be noted in this context that strongly zygomorphic flowers are not present in the best known Late Cretaceous deposits, and the radiation of both Polygalaceae and zygomorphic Fabaceae (i.e., papilionoid and caesalpinoid flowers) may have been a Tertiary event, possibly in response to selective pressures associated with the increased availability of specialized hymenopteran pollinators (apparently available at some level since at least Turonian times, Crepet, 2000). It is likely that members of the "stem group" subtending Polygalaceae–Fabaceae were actinomorphic and closely resembled modern Rosaceae, and thus identification of such taxa would be problematic without extensive characters for evaluation as might not be available in fossils (Lavin *et al.*, 2005).

The extensive record of the *Normapolles* palynomorphs in the Late Cretaceous (e.g., Christopher, 1979; Schönenberger *et al.*, 2001) seems to reflect diversification of the Juglandales lineage, which is well nested within the "Fagales" (sensu the APG II, 2003). Although *Normapolles* palynomorphs are diverse and difficult to place taxonomically from the characters of the pollen alone and there are only a few cases where pollen has been found in inflorescences, there is little doubt that at least some of the *Normapolles* palynomorphs are correctly placed in the broad clade that includes Juglandaceae, Betulaceae, and Myricaceae (Crepet *et al.*, 2004), implying a minimum age for the clade of 83 MyBP. Megafossils identifiable to those particular families do not occur until later, with the exception of cupulate Fagaceae from the Turonian of New Jersey (Crepet *et al.*, 2001; Nixon *et al.*, 2001).



Figure 9 An example of the exquisite preservation by charcoalification that has added so much to our understanding of the angiosperm fossil record. This 90 My old flower, *Pentapetalum trifasciculandricus* (Martinez-Millán *et al.*, 2009) related to modern Theales, is three dimensionally preserved and shows the preservation of delicate petals and stamens. Courtesy of Crepet and Martinez Milan, Photo by S. Sritko.

Notably absent are fossils of the Cucurbitales clade, which is mostly herbaceous, and for which the stem group should have been present at least by 90 MyBP. The pattern within these clades nicely illustrates the probable effect of depositional, preservational, and ecological bias in the fossil record. The largely herbaceous and insect-pollinated Cucurbitales is essentially without a fossil record (with the exception of Tertiary fossils of the woody genus *Tetrameles* (Datiscaceae), Lakhanpal and Verma, 1966), whereas the concomitant "Fagales" group, which is entirely woody and mostly wind pollinated has one of the most extensive fossil records for angiosperms, in terms of leaf, pollen, and reproductive structures, beginning in the Late Cretaceous and extending throughout the Tertiary (Crepet and Nixon, 1989a, b). Another factor that might contribute to the abundance of fossil Fagaceae and other members of the woody "Fagales" clade is their tendency to form extensive dominant forests (or riparian forests), whereas the Cucurbitales (e.g., Cucurbitaceae and

Begoniaceae in particular) are generally scattered individuals with far less cumulative biomass.

Ericanae (ASTERID III) and ASTERID IV Clades Diversification

The Ericales are well represented in the Turonian deposits of New Jersey (Crepet, 2000; Nixon and Crepet, 1993), providing an estimate of 90 MyBP for the minimum divergence of that group as well as for the broader group known as ASTERID III in recent works. The term "Ericanae" is used in **Figure 5** in place of Asterid III because it is more descriptive of the families that were not initially placed in Asteridae that, for the most part, lack fused corollas. The Asterid IV group includes Hydrangeales known from well-established fossils from Turonian deposits in New Jersey providing an estimate of *ca.* 90 MyBP for the divergence of this group.

ASTERID I and ASTERID II Clades Diversification

These groups include families that were traditionally grouped under Asteridae (Asteraceae, Solanaceae, Lamiaceae, Rubiaceae, and a host of other families that together include enormous species diversity, particularly in the tropics, Cronquist, 1981), and that typically have gamopetalous and often tubular and/or zygomorphic flowers. They are not reliably represented in Cretaceous deposits. The diversification of these asterids, with the associated floral features and various correlated often highly specialized pollination syndromes, very likely took place in the Tertiary unless early asterids were herbaceous, characterized by low population densities and were insect-pollinated and thus have gone undetected in the fossil record. Nonetheless, the apparently parallel radiation of legumes, a group with generally similarly specialized pollinators, is consistent with a radiation of asterids in the Tertiary (e.g., Herendeen and Crane, 1992; Lavin *et al.*, 2005), possibly for similar reasons.

Various complex pollination syndromes were likely to have been present by 90 MyBP given the diversity of Turonian floral morphologies (Crepet, 2000; Crepet and Nixon, 1998; Gandolfo *et al.*, 1998a, b; Nixon and Crepet, 1993). These morphologies have characteristics that imply high pollinator specificity, such as viscin threads (Nixon and Crepet, 1993), enclosed floral chambers that imply entrapment pollination (Gandolfo *et al.*, 2004), nonnectar floral rewards (Crepet and Nixon, 1998), and at least some sympetaly (Nixon and Crepet, 1993). The absence of strongly zygomorphic or bilabiate flowers at this time suggests that floral zygomorphy was, and remains, an ultimate refinement in pollination syndrome that followed the development of other features. It is worth noting that the groups that later developed zygomorphy (again, probably almost entirely in the Tertiary based on current evidence) have become some of the most diverse and successful clades of modern angiosperms, including a large portion of the asterids (I and II) as well as the papilionoid and caesalpinoid legumes. Just as striking, however, is the success and persistence of numerous clades in which zygomorphic floral presentation is absent or rare, and which generally have less-stringent pollinator specificity (or passive pollination syndromes), such as the majority of rosid and ranunculid lineages, including the mimosoid legumes.

Conclusions – Diversification Pattern

Conservative estimates of timing in angiosperm history suggest a rapid diversification of angiosperm groups between 113 and 80 MyBP, or alternatively, indicate an earlier diversification that left an extremely poor, even invisible fossil record.

The observed timing in angiosperm diversification has potential implications for understanding angiosperm success and distribution. For example, timing is critical in evaluating possible reasons for the rapid mid-Upper Cretaceous radiation of the tricolpate pollen-bearing angiosperms (the "eudicots"). Among other factors (Crepet and Niklas, 2009), this radiation appears to have been related to the availability of efficient pollinators; a relationship suggested by the timely appearance in the fossil record of characteristic floral morphologies, pollen characters, and taxa that are today specifically associated with such pollinators (Crepet, 2000; Crepet and Nixon, 1998; Gandolfo *et al.*, 1998a, b; Nixon and Crepet, 1993). This coincidence has been addressed in numerous papers and is consistent with the possibility that insect pollinators, especially those likely to demonstrate pollinator fidelity, have been important in at least one major diversification of flowering plants and ultimately have contributed to establishing their dominance in numbers of species (Crepet, 2000; Grimaldi and Engel, 2005; Crepet and Niklas, 2009).

It appears that angiosperms that developed zygomorphy, and extensively fused corollas, went through a second major radiation with highly specific insect pollinators almost entirely during the Tertiary. These groups have become some of the most diverse and successful clades of modern angiosperms and they include a large portion of the asterids (I and II) as well as the papilionoid and caesalpinoid legumes.

However, numerous clades which generally have less specific modes of pollination including the majority of rosid and ranunculiid lineages have been successful in attaining great diversity and temporal duration. In the same vein, it is interesting that other major groups radiated rapidly during the Tertiary and for apparently different reasons (Crepet and Niklas, 2009). Monocots with a spotty Cretaceous record nonetheless demonstrate considerable diversity by the end of the Cretaceous and an ensuing radiation during the Tertiary that includes palms and grasses. Although Juglandales/Fagales were apparently insect-pollinated early in their history (Friis, 1983), they diversified in the Tertiary in association with climatic change, animal seed dispersal, and significant adaptation to wind pollination.

Certainly, the fossil record indicates that angiosperms had the capacity to facilitate the evolution of manifold abiotic and biotic pollination syndromes that were critical to the success of the earliest flowering plants in light of a reciprocal driving mechanism for angiosperm and animal diversifications. However, if this was the single evolutionary innovation of the flowering plants, we would expect to see a monotonically decreasing pattern in the rate of angiosperm diversification throughout the Cretaceous–Tertiary, but the available (albeit limited) paleobotanical data indicate a pattern of saltational evolution and species diversification more in keeping with repeated bursts of phenotypic and correlated behavioral innovations resulting from the evolutionary introduction of novel functional traits throughout the history of the flowering

plants (Crepet and Niklas, 2009). These include features such as polyploidy and its possible significance to speciation, the attainment of the herbaceous habit-unique to angiosperms in living seed plants, the annual habit, also exclusive to angiosperms, the allopatric speciation-promoting influence of animal-mediated fruit dispersal and numerous other adaptations (Crepet and Niklas, 2009). The obvious and enduring question remains: Why did angiosperms have the capacity to accumulate these desirable and apparently speciation-promoting characteristics? We have not yet ruled out the possibility that a relatively superior reproductive biology characterized animal mediation somehow catalyzed the acquisition of these characteristics within the angiosperm lineage.

See also: Phylogeny. Pollinators, Role of. Systematics, Overview

References

APG II (2003) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436.

Brown RW (1956) Palm-like plants from the Dolores Formation (Triassic), southwestern Colorado. *US Geological Survey Professional Paper* 274H: 205–209.

Burnham R (2008) Hide and go seek: What does presence mean in the fossil record? *Annals of the Missouri Botanical Garden* 95: 51–71.

Chaw SM, Parkinson CL, Cheng YC, Vincent TM, and Palmer JD (2000) Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences of the United States of America* 97: 4086–4091.

Christopher RA (1979) *Normapolles* and triporate pollen assemblages from the Raritan and Magothy Formations (Upper Cretaceous) of New Jersey. *Palynology* 3: 73–121.

Cornet B (1986) The reproductive structures and leaf venation of a Late Triassic angiosperm *Sanmiguelia lewisii*. *Evolutionary Theory* 7: 231–309.

Cornet B (1989a) Late Triassic angiosperm-like pollen from the Richmond Rift Basin of Virginia, USA. *Paläontographica* B213: 37–87.

Cornet B (1989b) The reproductive morphology and biology of *Sanmiguelia lewisii*, and its bearing on angiosperm evolution in the Late Triassic. *Evolutionary Trends in Plants* 3: 25–51.

Crane PR (1985) Seed plant phylogenetics. *Annals of the Missouri Botanical Garden* 72: 716–793.

Crepet WL (2000) Progress in understanding angiosperm history, success, and relationships: Darwin's abominably "perplexing phenomenon". *Proceedings of the National Academy of Sciences of the United States of America* 97: 12939–12941.

Crepet WL (2008) The fossil record of angiosperms: Requiem or renaissance? *Annals of the Missouri Botanical Garden* 95: 3–33.

Crepet WL and Niklas KJ (2009) Darwin's second "abominable mystery": Why are there so many angiosperm species? *American Journal of Botany: Darwin Bicentenial* 96: 366–381.

Crepet WL and Nixon KC (1989a) Earliest megafossil evidence of Fagaceae: Phylogenetic and biogeographic implications. *American Journal of Botany* 76: 842–855.

Crepet WL and Nixon KC (1989b) Extinct transitional Fagaceae from the Oligocene and their phylogenetic implications. *American Journal of Botany* 76: 1493–1505.

Crepet WL and Nixon KC (1998) Fossil Clusiaceae from the Late Cretaceous (Turonian) of New Jersey and implications regarding the history of bee pollination. *American Journal of Botany* 85: 1122–1133.

Crepet WL and Stevenson DW (2010) The Bennettitales (Cycadeoidales): A preliminary perspective on this arguably enigmatic group. In: Gee CT (ed.) *Plants in Mesozoic Time: Morphological Innovations, Phylogeny, Ecosystems*, pp, 215–246. Bloomington: Indiana University Press.

Crepet WL, Nixon KC, and Gandolfo MA (2001) Turonian flora of New Jersey, USA. In: *VII International Symposium on Mesozoic Terrestrial Ecosystems*, vol. 7, pp. 61–69. Asociación Paleontológica Argentina, Publicación Especial.

Crepet WL, Nixon KC, and Gandolfo MA (2004) Fossil evidence and phylogeny: The age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *American Journal of Botany* 91: 1666–1682.

Crepet WL, Nixon KC, and Gandolfo MA (2011) The mismeasure of a mystery: Angiosperm origins and phylogentics. *Botany* St. Louis, July 9–13.

Cronquist A (1981) An Integrated System of Classification of Flowering Plants. New York: Columbia University Press.

Darwin C (1903) In: Darwin F and Seward AC (eds.) *In More Letters from Charles Darwin*, and 239–240, pp. 20–22. New York: Appleton.

Dilcher DL (1974) Approaches to the identification of angiosperm leaf remains. *Botanical Review: Interpreting Botanical Program* 40: 1–157.

Doyle JA and Donoghue MJ (1986) Seed plant phylogeny and the origin of the angiosperms: An experimental cladistic approach. *Botanical Review* 52: 321–431.

Endress PR and Doyle JA (2009) Reconstructing the ancestral angiosperm flower and its initial specializations. *American Journal of Botany* 96: 5–21.

Friedman WE (2009) The meaning of Darwin's "abominable mystery". *American Journal of Botany* 96: 5–21.

Friis EM (1983) Upper Cretaceous (Senonian) floral structures of juglandalean affinity containing Normapolles pollen. *Review of Palaeobotany and Palynology* 39: 161–188.

Friis EM, Crane PR, Pedersen KR, *et al.* (2007) Phase contrast X-ray microtomography links Cretaceous seeds with Gnetales and Bennettitales. *Nature* 450: 549–552.

Friis EM, Doyle JA, Endress PK, and Leng Q (2003) *Archaefructus* – angiosperm precursor or specialized early angiosperm? *Trends in Plant Science* 8: 369–373.

Friis EM, Pedersen KR, and Crane PR (2004) Araceae from the Early Cretaceous of Portugal: Evidence on the emergence of monocotyledons. *Proceedings of the National Academy of Sciences of the United States of America* 101: 16565–16570.

Friis EM, Pedersen KR, and Crane PR (2009) Early Cretaceous mesofossils from Portugal and eastern North America related to the Bennettitales–Erdtmanithecales–Gnetales group. *American Journal of Botany* 96: 252–283.

Frohlich MW and Parker S (2000) The mostly male theory of flower evolutionary origins: From genes to fossils. *Systematic Botany* 25: 155–170.

Gandolfo MA, Nixon KC, and Crepet WL (1998a) *Tylerianthus crossmanensis* gen. et sp. nov. (aff. Hydrangeaceae) from the Upper Cretaceous of New Jersey. *American Journal of Botany* 85: 376–386.

Gandolfo MA, Nixon KC, and Crepet WL (1998b) A new fossil flower from the Turonian of New Jersey: *Dressanthia bicarpellata* gen. et sp. nov. (Capparales). *American Journal of Botany* 85: 964–974.

Gandolfo MA, Nixon KC, and Crepet WL (2000) Monocotyledons: A review of their Early Cretaceous record. In: Wilson K and Morrison D (eds.) *Proceedings of the Second International Conference on the Comparative Biology of the Monocotyledons*. Sydney, Australia: CSIRO.

Gandolfo MA, Nixon KC, and Crepet WL (2004) The oldest complete fossil flowers of Nymphaeaceae and implications for the complex insect entrapment pollination mechanisms in Early Angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* 101: 8056–8060.

Gould RE and Delevoryas T (1977) The biology of *Glossopteris*: Evidence from petrified seed-bearing and pollen-bearing organs. *Alcheringa* 1: 387–399.

Graur D and Martin W (2004) Reading the entrails of chickens: Molecular timescales of evolution and the illusion of precision. *Trends in Genetics* 20: 80–86.

Grimaldi D and Engel MS (2005) *Evolution of the Insects*. New York: Cambridge University Press.

Harris TM (1932) The fossil flora of Scoresby Sound, east Greenland. Part 2: Description of seed plants incertae sedis together with a discussion of certain cycadophyte cuticles. *Meddelelser om Grønland* 85: 1–112.

Harris TM (1937) The fossil flora of Scoresby Sound, east Greenland. Part 5. Stratigraphic relations of the plant beds. *Meddeleser om Grønland* 112(2): 1–114.

Harris TM (1940) Caytonia. *Annals of Botany.* 4: 713–734.

Harris TM (1964) The Yorkshire Jurassic Flora. 11. Caytoniales, Cycadales and Pteridosperms. London: British Museum (Natural History).

Herendeen PS Crane PR (1992) Early Caesalpiniod fruits from the Palaeogene of Southern England. *Advances in Legume Systematics 4: The Fossil Record*. In: Herendeen PS and Dilcher DL (eds.) Kew Royal Botanic Gardens, pp. 57–68.

Hickey LJ (1973) Classification of the architecture of dicotyledonous leaves. *American Journal of Botany* 60: 17–33.

Hickey LJ and Wolfe JA (1975) The bases of angiosperm phylogeny: Vegetative morphology. *Annals of the Missouri Botanical Garden* 62: 538–589.
Knobloch E and Mai DH (1986) Monographie der früchte und samen in der Kreide von Mitteleuropa. *Edice Rozpravy ústředního ústavu Geologického* 47: 1–219.
Lakhanpal RN and Verma JK (1966) Fossil wood of Tetrameles from the Deccan Intertrappean beds of Mohgaonkalan, Madhya Pradesh. *Paleobotanist* 14: 209–213.
Lavin M, Herendeen PS, and Wojciechowski MF (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Botany* 54: 575–594.
Magallon-Puebla S, Crane PR, and Herendeen PS (1999) Phylogenetic pattern, diversity, and diversification of Eudicots. *Annals of the Missouri Botanical Garden* 86: 297–372.
Martínez-Millán M, Crepet WL, and Nixon KC (2009) *Pentapetalum trifasciculandricus* gen. et. sp. nov., a thealean fossil flower from the Raritan formation, New Jersey, USA (Turonian, Late Cretaceous). *American Journal of Botany* 96: 933–949.
Mathews S (2009) Phylogenetic relationships among seed plants: Persistent questions and the limits of DNA sequence data. *American Journal of Botany* 96: 228–236.
Nixon, KC (2003) *Winclada. Computer program for phylogenetic analysis.* www.cladistics.com (accessed November 2006).
Nixon KC and Crepet WL (1993) Late Cretaceous fossil flowers of ericalean affinity. *American Journal of Botany* 80: 616–623.
Nixon KC, Crepet WL, Stevenson DM, and Friis EM (1994) A reevaluation of seed plant phylogeny. *Annals of the Missouri Botanical Garden* 81: 484–533.
Nixon KC, Gandolfo MA, and Crepet WL (2001) Origins of Fagaceae: A review of relevant Turonian fossil material from New Jersey. *American Journal of Botany* 88: 68 (abstract).
Palmer JD and Zamir D (1982) Chloroplast DNA evolution and phylogenetic relationships in Lycopersicon. *Proceedings of the National Academy of Sciences of the United States of America* 79: 5006–5010.
Retallack GJ and Dilcher DL (1981) A coastal hypothesis for the dispersal and rise to dominance of flowering plants. In: Niklas KJ (ed.) *Paleobotany, Paleoecology and Evolution.* Connecticut: Praeger Publishers.
Rothwell GW, Crepet WL, and Stockey RA (2009) Is the anthophyte hypothesis alive and well? New evidence from the reproductive structures of Bennettitales. *American Journal of Botany: Darwin Bicentenial* 96: 296–322.
Rothwell GW and Serbet R (1994) Lignophyte phylogeny and the evolution of spermatophytes: A numerical cladistic analysis. *Systematic Botany* 19: 443–482.
Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14: 1218–1232.
Sanderson MJ, Thorne JL, Wilk N, and Bremer KG (2004) Molecular evidence on plant divergence times. *American Journal of Botany* 91: 1656–1665.
Savolainen V, Fay FM, Albach DC, *et al.* (2000) Phylogeny of the eudicots: A nearly complete familial analysis based on rbcL gene sequences. *Kew Bullettin* 55: 257–309.
Schönenberger J, Pedersen KR, and Friis EM (2001) Normapolles flowers of fagalean affinity from the Late Cretaceous of Portugal. *Plant Systematics and Evolution* 226: 205–230.
Soltis DE, Soltis PS, Chase MW, *et al.* (2000) Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. *Botanical Journal of the Linnean Society* 133: 381–461.
Sun G, Dilcher DL, Zheng S, and Zhou Z (1998) In search of the first flower: A Jurassic angiosperm, Archaefructus, from northeast China. *Science* 282: 1692–1695.
Sun G, Ji Q, Dilcher DL, Zheng SL, Nixon KC, and Wang XF (2002) Archaefructaceae, a new basal angiosperm family. *Science* 296: 899–904.
Taylor EL and Taylor TN (2009) Seed ferns from the late Paleozoic and Mesozoic: Any angiosperm ancestors lurking there? *American Journal of Botany* 96: 237–251.
Thomas HH (1925) The Caytoniales, a new group of angiospermous plants from the Jurassic rocks of Yorkshire. *Philosophical Transactions of the Royal Society* 213B: 299–363.
Tidwell WD, Simper AD, and Thayn GF (1977) Additional information concerning the controversial Triassic plant: Sanmiguelia. *Palaeontographica* 163B: 143–151.
Tiffney BH (2008) Phylogeography, fossils, and Northern Hemisphere biogeography: The role of physiological uniformitarianism. *Annals of the Missouri Botanical Garden* 95: 135–143.
Walker JW and Doyle JA (1975) The bases of angiosperm phylogeny: Palynology. *Annals of the Missouri Botanical Garden* 62: 664–723.
Walker JW, Brenner GJ, and Walker AG (1983) Winteraceous pollen in the Lower Cretaceous of Israel: Early evidence of a magnolacean angiosperm family. *Science* 220: 1273–1275.

Origin of Life, Theories of

Susanne Brakmann, Evotec Biosystems AG, Germany

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 439–450, © 2001, Elsevier Inc.

Glossary

Autocatalysis Self-acceleration of certain chemical reactions that yield catalytically active products. The reaction rate of autocatalytic processes increases exponentially.

Hypercycle Cyclic sequence of self-supporting reactions between primitive prebiotic biomolecules, nucleic acids, and proteins. Hypercyclic organization of reacting systems was postulated to explain the spontaneous emergence of replicating systems. As a result of the hypercyclic reaction principle, functional characteristics of replicating molecules (i.e., phenotypes) are connected to their hereditary characteristics or their genotypes. Consequently, both phenotypes and appertaining genotypes are evaluated and selected in feedback loops.

Natural selection Preferential reproduction and survival of certain species under given environmental conditions. Growth advantages of selected individuals reflect the efficiency of their reproduction pathways.

Quasispecies Hierarchically ordered population of mutants that results from erroneous copying of a genotypic ancestor. Whole mutant distributions behave like single species. They are subjected to natural selection. The average sequence is defined as a consensus sequence that represents the "center of gravity" of the distribution. The consensus sequence is synonymous to a wild type.

Self-organization Spontaneous formation of complex structural aggregates from biological macromolecules such as proteins, nucleic acids, or lipids. The assembly pathways are usually determined by physicochemical properties of the constituent molecules.

Historical Overview

What is Life?

Currently, life appears on Earth with an immense variety of structures and functions. Despite this diversity, some invariant features exist that are common to living systems and distinguish them from non-living matter.

All life takes shape with defined individuals. Cells are the smallest individual entities capable of performing all processes for living and reproducing. They represent the elementary structural units of life. Cells do not spontaneously originate from suitable mixtures of nonliving matter. Louis Pasteur (1822–1895) was the first to disprove the idea of abiogenesis by showing that each cell results from cell division. Rudolf Virchow (1821–1902) confirmed this, noting that fusion of cellular precursors (germ cells) is the only alternative for generating new cells. Two general types of cellular organization are known: the prokaryotic type, which lacks a nucleus, and the more highly developed eukaryotic type, which harbors a nucleus. Basically, both cell types exhibit the same structure and functional organization suggesting a common origin. Simple organisms such as bacteria consist of single cells which are either prokaryotic or eukaryotic, whereas complex organisms, such as plants and animals, are highly organized multicellular assemblies of the eukaryotic type.

The different forms of life do not originate from nothing as Charles Darwin (1809–1882) discerned in his epoch-making work. He, and independently Alfred Russel Wallace (1823–1913), postulated that, instead, a continuous process of successive adaptation to environmental conditions led to the origin and development of species. Darwin also realized that evolutionary progress results from mutageneous reproduction and natural selection. Today, this theory can be explained with twentieth-century insights into structural and chemical characteristics of living cells.

Each single cell contains an extraordinarily efficient machinery for respiration and production of energy; for absorption and assimilation of nourishment; for synthesis, storage, utilization, and transport of new products; and for production of more or less identical copies of itself. The cellular machinery is instructed by hereditary information that is carried by large molecules which are composed of nucleic acid and known as genes. They comprise the architect's plan of all cellular components and also include structural elements regulating the formation of the different characteristics of an organism. During the process of reproduction, genes also reproduce, thereby passing all instructions on to the next generation. Proteins represent the executive that is coded by genetic information. They cooperate closely with nucleic acids in order to fulfill the diverse functional and structural tasks that maintain the cell machinery. Many proteins are enzymes that act as catalysts for the reproduction of nucleic acids and for the synthesis of essential organic compounds, such as carbohydrates, amino acids, lipids, and hormones. Other proteins care for the cellular metabolism and the production of energy. The teamwork of these two molecular species, polynucleotides and polypeptides, predominantly characterizes the chemistry of living matter as we know it.

Generally, high-energy macromolecular compounds play a major role in living organisms. In addition to nucleic acids and proteins, biomolecules such as lipids, polysaccharides, and phosphates are found. They serve for demarcation of individual cells, for their division into different subcompartments, for their structural stability, and for energy transfer

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00101-5

processes. However, the splendid variety of life on Earth is mainly due to the heteropolymeric nature of proteins. They consist of long chains that are built from 20 constituents, chemically defined as amino acids. Similarly, nucleic acids encoding the natural proteins are polymeric chains based on 4 different monomers. The building principle realized with both molecular species allows a variability that is far beyond the imagination—even with chain lengths of 20 monomers, more than 10^{12} different nucleic acids (with 4 different building blocks) or 10^{26} alternative proteins (with 20 different building blocks) can exist.

Structural variability is not the only feature that characterizes the molecules of life. Nucleic acids are also capable of serving as templates for their own replication, thereby guaranteeing the reproduction of living organisms. Self-reproduction is an inherent property of polynucleotides that is based on physical interactions between their complementary constituents. This phenomenon is known as base pairing and allows the synthesis of a unique complementary copy of each molecule. Replication of this copy then yields a copy of the original chain. The complete procedure very much resembles the positive–negative printing used in photography. Occasionally, the copying process is imperfect, leading to mutations in gene replication. A mutation alters the genetic information and therefore the instructions for one or more characteristics. Continuous replication and mutation can then produce genetic variability. Following expression (i.e., translation of nucleic acids into proteins), some mutations will produce altered characteristics that are favorable for the organism. This organism will reproduce preferentially over those with the unchanged gene. By far most mutations are deleterious, leading to reduced replication or even cell death. Only beneficial mutations which occur accidentally result in organisms with better adaptation to their environment. In this way, organisms evolve toward higher fitness and often toward greater complexity. In summary, complex organisms evolved through time according to Darwin's theory because of replication, mutageneous replication, natural selection, and replication of resulting mutants. During the process of Darwinian evolution, information is created continually.

Another view of "life" is presented by the discipline of thermodynamics that distinguishes between open and closed systems. Corresponding to the second law of thermodynamics, no processes can occur in closed systems that increase the net order of the system or that decrease their entropy. In this perspective, any living organism is an open system that exchanges energy and matter with its surrounding, thereby establishing and maintaining a highly ordered structural assembly at the expense of a larger decrease of order of the universe outside. Living organisms participate in ecosystems and depend on resources found in their environment. In particular, most organisms are dependent on the flow of sunlight, which is absorbed by plants and utilized by them to synthesize high-energy molecules from simpler ones. These products serve as a source of food and energy for other links of the food chain (e.g., animals). By excretion and degradation of dead organic compounds, all material will be recycled without altering the organisms' state of life, which is characterized as being far from thermodynamic equilibrium. A state of thermodynamic equilibrium would be identical to a state of maximum entropy and thus correspond to cell death.

All organisms on Earth are closely related and the fundamental pattern of their life is essentially identical for all of them. The similarity among all living organisms probably implies that life is descended from a single instance of origin. This incident is central to many scientific and philosophical problems and the object of various different and contradictory hypotheses which will be described in the following section.

Hypotheses and Theories on the Origin of Life

The traditional position of theology and some philosophy views the origin of life as the result of a supernatural event which is permanently beyond the descriptive powers of chemistry and physics. In its most general form, this view is not necessarily contradictory to contemporary scientific knowledge about prebiotic evolution, although the biblical descriptions of creation given in the first two chapters of Genesis, taken literally and not metaphorically, are inconsistent with modern knowledge.

Until the mid-seventeenth century, the prevailing opinion was that God created man together with higher animals and plants, but that simple forms of life such as worms and insects arise steadily from mud, waste, and putrefied matter during short periods of time. The physiologist William Harvey (1578–1657), who studied reproduction and development of deer, was the first to challenge this view by postulating that every animal comes from an egg ("*omnia viva ex ovo*") a long time before Karl-Ernst von Baer (1792–1876) discovered the existence of human egg cells by microscopy. An Italian scientist, Francesco Redi (1626–1698), found Harvey's idea to be true, at least for insects; he found that maggots in meat arise from fly eggs. Later, Lazzaro Spallanzani (1729–1799) discovered that spermatozoa were necessary for the reproduction of mammals. Before Pasteur, Spallanzani also showed that living matter ("infusories") does not originate from boiled fluids kept in closed containers. Although Redi's and Spallanzani's findings definitely proved that insects and larger animals develop from eggs, it remained obvious to a large majority that at least microorganisms, because of their ubiquity, are generated continually from inorganic material. The debate of whether life is spontaneously generated from non-living matter or not culminated in the famous controversy between Louis Pasteur and Félix-Archimède Pouchet (1800–1872) which Pasteur won triumphantly. He showed that even microorganisms in fluids come from germs floating in the air, and he also demonstrated that nutrient solutions could be guarded against these creatures by suitable sterilization such as filtration or boiling. However, contemporary scientists were not satisfied by Pasteur's experiments because a delicate question remained: If living organisms do not arise from non-living matter, how had life come about in the first place?

In the late nineteenth century, another hypothesis was initiated by the Swedish chemist Svante Arrhenius (1859–1927). He strongly believed that the whole universe is replenished with living germs, a phenomenon that he called "panspermia." He suggested that microorganisms and spores

of cosmic origin spread from solar system to solar system, and thus they arrived on Earth. Although Arrhenius' view avoids rather than solves the problem of the origin of life, and despite the extreme unlikeliness of microorganisms surviving the interstellar effects of cold, vacuum, and radiation, a few twentieth-century members of the scientific community returned to the idea of panspermia. Among these scientists are astronomer Fred Hoyle (1915–) and molecular biologist Francis Crick (1916–), who are convinced that the time span between the origin of Earth and the appearance of first cellular organisms on this planet was too short for life to have occurred spontaneously.

Darwin's theory of "natural selection as motive power for evolution" resulted in a new view on the phenomenon of life that is still valid. Although Darwin did not commit himself on the origin of life, contemporary scientists such as Thomas Huxley (1825–1895) extended his idea, asserting that life could be generated from inorganic chemicals. Pursuing this opinion, Alexander Oparin (1894–1980) was the most influential advocate of the successive origin of cellular organisms from non-living matter. He suspected this transition was proceeded by a series of regular and progressive chemical reactions under the physical and chemical conditions on early Earth. Together with John Scott Haldane (1860–1936), Oparin recognized that the abiological production of organic molecules in the current oxidizing atmosphere of Earth is highly unlikely. Instead, both suggested that the beginning of life occurred in primordial hot waters under more reducing (i.e., hydrogen-rich) conditions. Furthermore, Oparin postulated the existence of pre-cellular coacervates—globular units with membrane-like surface structures—that may have high concentrations of certain chemical compounds. Coacervates indeed form spontaneously from colloidal aqueous solutions of two or more macro-molecular compounds.

However, many fundamental problems on the transition from non-living to living matter remained unsolved. The central question concerned the role of the second law of thermodynamics, which defines the equilibrium in an isolated system as a state of maximum entropy that appears to contradict the origin and existence of highly ordered living organisms. Erwin Schrödinger (1887–1961) gave a decisive answer to this question, stating that "living matter evades the decay to equilibrium" or death by steadily compensating for the production of entropy. In any organism, this is achieved by feeding it free energy or energy-rich matter which is used by the cellular machinery to drive essential chemical reactions. Schrödinger and others also realized that living organisms can thermodynamically be described as open systems, but they could not explain the general physical conditions for self-ordering processes. These were perceived by Ilja Prigogine (1917–) and Paul Glansdorff (1904–1999), who worked on a thermodynamic theory of irreversible processes. According to Prigogine, selection and evolution cannot occur in equilibrated or nearly equilibrated reaction systems, even if the right types of substances are present. Instead, certain combinations of autocatalytic reactions with transport processes may lead to peculiar spatial distributions of reaction partners, called "dissipative structures." These ordered structures are of importance for the formation of functional order in the evolution of life, especially for early morphogenesis. However, the first steps of self-organization probably involved little organization in physical space but extensive functional ordering of a tremendously complex variety of chemical compounds. Manfred Eigen (1927–) explained the process of ordering among molecules by augmenting the Prigogine–Glansdorff principle with phenomenological considerations on the behavior of self-replicating molecules: A certain quantity approaches a maximal value in any open system that is replicating autocatalytically with sufficient fidelity, and thereby continually consuming energy and matter. This quantity is called "information" and is closely related to the "negative entropy" postulated by Schrödinger. In addition to setting the stage for a molecular interpretation of biological information, Eigen developed the mathematical models for describing "selection." According to Eigen's theory, selection is the fundamental natural principle that brings order into any random arrangement of autocatalytically replicating species. With selection, information is generated successively, leading to a steady optimization of species, which can either be organisms or molecules.

The mathematical models developed by Eigen support a detailed hypothesis of the origin of life which comprises multiple, successive steps for the transition from inorganic to living matter. However, it should be mentioned that some scientists have theories on the emergence of life that differ from Eigen's theory. Among these is Stuart Kauffman (1939–), who believes that natural selection is important but not the sole ordering principle of the biological world. Instead, he considers spontaneous self-organization to represent the predominant source of natural order. Kauffman demonstrated that sets of inter-related autocatalytic reactions can undergo a transition to a newly ordered (i.e., self-organized) state as soon as their connectivity reaches a certain threshold value. Furthermore, Kauffman emphasizes that the phenomenon of autocatalysis, which plays the central role in his theory, is not limited to nucleic acids. Therefore, he concludes that even genes were not necessary for the origin of life. In contrast to Kauffman, Eigen distinguishes "random" autocatalytic or self-replicating activity which is observed for a variety of molecular species from the "inherently" self-replicating nucleic acids. Inherent capability for self-replication, in turn, represents the molecular basis for natural selection according to Eigen's theory.

Well-defined experiments were invented in order to simulate the principles that were postulated for molecular evolution. With certain experimental set-ups, replication and selection can be performed in a test tube. Similarly, the chemical conditions on primordial Earth can be mimicked in the laboratory. Several scientists attempted to verify experimentally the twentieth-century ideas on biogenesis. Their experiments are discussed in the following section.

Experiments

Production of Building Blocks and Polymers

The ideas of Oparin and Haldane inspired the first simulations of prebiotic conditions on early Earth. Under the auspices of Harold Clayton Urey (1893–1981), graduate student

Stanley Miller (1930–) attempted to simulate the primordial chemical processes in a famous experiment. He continually exposed a mixture of methane, ammonia, hydrogen, and water vapor above an "ocean" of boiling water to sparks of a corona discharge in a closed apparatus (**Figure 1**). The gaseous reaction products were condensed in a cooler, dissolved in boiling water, and recirculated to the gaseous atmosphere. After several days of exposure to sparking, the aqueous solution changed color. Subsequent analysis indicated that several amino and hydroxy acids had been produced by this simple procedure. Because on early Earth probably much more energy was available in ultraviolet light than in lightning discharges, astrophysicist Carl Sagan (1934–1996) and colleagues altered Miller's experimental conditions. They synthesized amino acids by applying long-wavelength ultraviolet irradiation to a gaseous mixture of methane, ammonia, water, and hydrogen sulfide. Similar to Miller's experiment, these attempts also revealed that amino acids, particularly the biologically relevant ones, form readily under simulated primitive conditions. Also, it became obvious that reducing conditions were necessary for prebiological formation of amino acids since no yield was observed with oxidizing atmosphere. Further modification of these types of experiments showed that alkaline conditions led to the spontaneous synthesis of a variety of sugar molecules, including the five-carbon sugars fundamental for the formation of nucleic acids and six-carbon sugars such as fructose and glucose, which represent common metabolites and structural building blocks in contemporary organisms. As was shown by Juan Oró (1923–) and several other investigators, nucleotide bases and porphyrins can also be synthesized from simple mixtures of ammonia, water, and hydrogen cyanide. Therefore, it seems probable that most—if not all—essential building blocks of proteins, carbohydrates, and nucleic acids can be produced under quite general primitive reducing conditions.

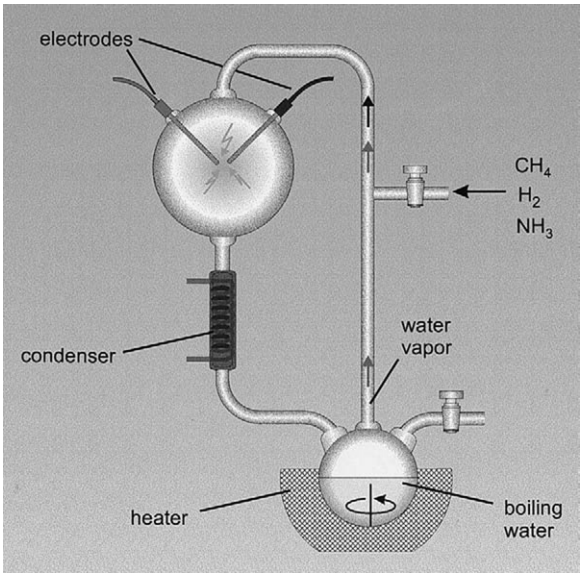


Figure 1 Simulation of prebiotic chemistry by Miller and Urey. Methane, ammonia, and hydrogen are led into the apparatus via stopcock 1 (top), driven into the reaction vessel by water vapor, and exposed to sparks of a corona discharge. The organic gaseous reaction products are condensed and dissolved in boiling water. Samples for analysis can be taken through stopcock 2 (bottom).

The formation of polymers, long-chain molecules made of repeating units of the building blocks, would have to be the next major step during chemical evolution. However, polymerization of monomers was not achieved in the simple experiments described previously. These reactions involve the loss of one water molecule during the formation of each two-unit product and therefore are facilitated under dehydrating conditions. By dry heating of amino acid mixtures, Sidney Fox (1912–) accomplished the condensation reaction in the laboratory to yield polyamino acids. He observed that these molecules were not random polymers and they also exhibited some distinct enzyme-like catalytic activities; therefore, he called them proteinoids. In addition to this work, various experimental approaches concerned the catalysis of polymerization reactions by mineral surfaces, which may also protect growing chains from degradation by water molecules, as well as combinations of heating under dry and wet conditions mimicking the environment of submarine hydro thermal vents.

Some scientists cast considerable doubt on any contribution of prebiotic syntheses of organic compounds to the origin of life. Predominant among these is Günther Wächtershäuser (1938–), who imagines an origin without reducing atmosphere and without the primordial soup. According to Wächtershäuser, prebiotic chemistry began on the surface of minerals in the lithosphere. If these minerals were positively charged, they would form ionic bonds with negatively charged chemical groups of phosphate, carbonyl, or sulfide. As a result, a layer of negatively charged molecules would form, coating the minerals in two dimensions. Once a diversity of molecules accumulated, inorganic reactions involving the minerals could supply reducing power for the synthesis of more complex organic molecules from various possible interactions within the two-dimensional layer.

The most important implication of Wächtershäuser's theory is that the first forms of life could have been autotrophic instead of heterotrophic. Autotroph organisms get their carbon from carbon dioxide, whereas heterotrophs get it from organic molecules such as glucose. The fact that the oldest microfossils appear to be photoautotroph blue-green bacteria supports the autotroph hypothesis. Nevertheless, these ideas have not been adequately tested.

Evolution Experiments

The onset of self-replication—following the formation of polymeric nucleic acids—represents a major difficulty for the explanation of the origin of life. Experimental attempts by Leslie Orgel (1927–) and coworkers revealed that, under prebiotic conditions, it is possible to add monomeric units, called nucleotides, to complement a small template nucleic acid molecule and to combine the adhering nucleotides to the correct negative copy of the positive original. This is merely the first step of a self-replicating process that must be continued by a separation of both strands, positive and negative, and by repeated addition and concatenation of monomers. Without proteins supporting the polymerization of the building blocks, identical copies of the original nucleic acid

molecule, the positive strand, could not be synthesized. However, because the simultaneous emergence of self-replication of nucleic acids and specific catalysis of their duplication by proteins seems very unlikely, Orgel, together with Francis Crick and Carl Woese (1928–), supposed that the first and primitive self-replicating systems involved nucleic acids only. The discovery of catalytically active nucleic acids, or ribozymes, by Thomas Cech (1947–) and Sidney Altman (1939–) decisively supports this hypothesis. Much contemporary work therefore focuses on searching for ribozymes catalyzing the replication of nucleic acids.

Experimental studies on selection among replicating polynucleotides were first performed by Solomon Spiegelman (1914–). He purified nucleic acid and the replicating enzyme of bacteriophage Qβ and implemented their replication reaction in a test tube which contained all essential reagents. Then he replicated the viral nucleic acid in serial transfer experiments by applying constraints that select exclusively for higher replication rates (**Figure 2**). Beginning with infectious nucleic acid molecules, sequences were obtained after several rounds that replicated several times faster and had much lower molecular weights than the original molecules but were unable to infect bacteria. On the basis of more detailed knowledge of this replication reaction, Christof Biebricher (1941–) and coworkers, together with Manfred Eigen, repeated Spiegelman's experiments. They applied conditions which were chosen to be optimal with respect to rate of evolution, and they analyzed the stepwise progress of the optimization procedure. Their attempts at facilitating artificial evolution established an "irrational" technique for the design of new drugs known as directed molecular evolution. Eigen and coworkers also performed evolution experiments with various different replicators, either molecules or viruses, and they observed functional organization between replicators and their catalysts as well as some of the molecular mechanisms of evolutionary optimization.

Theoretical Concepts

Quantitative Treatment of Darwinian Evolution

The process of complementary reproduction of self-replicating molecular assemblies can be described in detail by applying the mathematical formalism of chemical kinetics. For any population of molecules within a certain environment, consideration from the viewpoint of kinetics accounts for three necessary conditions: metabolism, self-reproduction, and mutability.

Metabolism describes the rates of formation and decomposition of each molecular species taking part in the competition. The term metabolism expresses the continuous influx of energy-rich matter, which maintains the chemical reaction, and the corresponding efflux of decomposition products. In the case of nucleic acids, the influx is due to their energy-rich building blocks, nucleoside triphosphates. The respective mono-phosphates result from decomposition by hydrolysis, or by enzymatic cleavage of polymeric strands, and represent the efflux. Metabolism by consuming and processing resources thus effects a dependence of the replicating individuals on their environment.

Self-reproduction is due to the self-accelerating or autocatalytic replication of nucleic acids. The amount of descendant molecules in a defined environment depends on the rate of the replication reaction, which is dependent on the sequence of each replicating molecule. Fluctuations of the reaction rate may also be due to variations of substrate concentrations resulting from influx and outflux within an open system.

Mutability takes into account the quality of replication. The copying fidelity is physically limited because irregularities occasionally occur during the recognition and interaction of complementary strands. Therefore, errors are introduced into the replicating molecules which lead to an increase in the variety of the offspring.

The combined considerations can be expressed in differential equations, which are used to calculate the parallel

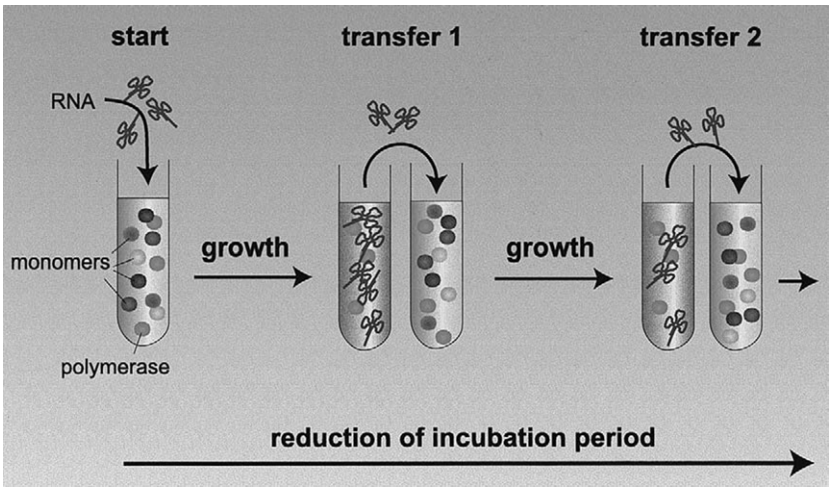


Figure 2 Scheme of the serial transfer principle performed by Spiegelman. A series of test tubes contain Qβ replicase, RNA building blocks (nucleoside triphosphates of A, G, C, and U), and necessary growth factors. The mixture is supplied with RNA template, and the reaction is started by increasing the temperature to 30 °C. When a certain product concentration is reached, a small fraction of the mixture is transferred to the next tube and incubated in the same way. The procedure is repeated many times, continually reducing the incubation period. Eventually, a single optimal RNA product is selected.

growth of molecular individuals within an expanding population. Similar to the constancy of generalized forces or fluxes in the thermodynamics of irreversible processes, the reaction equations describing autocatalytic growth of nucleic acids can be simplified on the basis of certain constant boundary conditions. In the case of constant fluxes or constant organization, the equations can be solved exactly and result in explicit expressions. The physical interpretation of the results reveals that selection in the Darwinian sense is an inherent property of autocatalytically replicating populations of molecules. Selection can be characterized by an extremal principle which mathematically describes the preferential treatment of those replicators which maximize their fitness with respect to the environmental constraints. These optimal mutants, or more precisely mutant distributions, succeed at the total expense of the others. Analogous to the concept of species that Darwin believed to be the target of "survival of the fittest" within the process of natural evolution, the selected mutant distribution with the most efficient replication pathways behaves like an individual species; therefore, it was called "quasispecies" by Eigen. The substitution of less advantageous by more advantageous mutant distributions results in reaction rate changes in opposite directions—similar to those calculated by Prigogine and Glansdorff for simple autocatalytic processes far from thermodynamic equilibrium. Hence, the instability associated with the selection process is in accordance with the expectations drawn from the Prigogine–Glansdorff principle: It results in the breakdown of the formerly (meta-) stable quasispecies and its substitution by a new quasispecies.

The quasispecies behaves physically and mathematically like a single species, or more precisely like the selected wild type of a species, but does not define a single species. It includes all mutants of the dominant form and represents a widely dispersed, not necessarily symmetrical, distribution of similar and related, but by no means identical, sequences. A quasispecies is usually characterized by a defined average sequence, called a consensus sequence, and can be dominated by one or more so-called master sequences. Evolutionary progress of the quasispecies results in new mutants that appear on the periphery of the distribution. Selection, or the emergence of mutants which replicate more efficiently than their ancestors, is a manifestation that new genetic information has been generated.

The Informational Aspect

What Is Information and How does it Originate?

In the most common sense, information is associated with the content and meaning of a "news" message. From the viewpoint of a recipient, the news can be understood only if the recipient and the transmitter agree on the meaning of the symbols used to code the message. In other words, new information is gained by a recipient when he or she can interpret the received symbols or the elementary units of information. Genetic information is stored in the linear copolymers of RNA and DNA. The number of monomers, representing the symbols, and their sequence within the polymeric chains designate whether or not a molecule encodes a message, and the arrangement of symbols determines the molecule's information content. The meaning of the symbols originates with the unique ability of RNA and DNA to recognize and "read" themselves. Their reading is based on particular chemical interactions mediated by specific hydrogen bonding that links the four monomers, A, T (U in the case of RNA), G, and C, into the complementary pairs AT (AU) and GC. The encoded messages transmitted by RNA and DNA include instructions that initiate and control all biochemical reaction cascades taking place within each organism.

During the process of selection, mainly the individual abilities to replicate are assessed with respect to the requirements of a given environment. The most efficient replication pathways, in turn, are "instructed" by genetic information. Therefore, new information is generated by erroneous replication and selection.

Maximal Information Content of the "Species"

Evolutionary flexibility represents the pivot of the selection process. It is guaranteed by replication error rates that produce a sufficiently large variety of mutants. However, as the mutation rates increase, the accumulation of errors outweighs the selective narrowing down. As a consequence, the genetic information "melts" away into random combinations of symbols that correspond to nonsense messages without any ability to generate and maintain necessary reaction pathways. This is a catastrophe for every living organism because it leads to the evolutionary decay of information. Therefore, natural evolution processes are bounded by an error threshold. Far below this boundary, selection can be reduced to all-or-none decisions. Evolutionary progress is achieved at reasonable velocities only when the error rate is near the threshold value. Indeed, it has been established experimentally by Esteban Domingo (1942–), Charles Weissmann (1931–), Lawrence Loeb (1936–), and several other research groups that viruses (which, due to their short replication periods, are ideally suited for evolutionary studies) behave like quasispecies and replicate close to their error thresholds. In accordance with theory, the reciprocals of the error rates that have been experimentally determined for many viruses define the limiting information capacity of their genomes. Loeb was also the first to prove the hypothesis that an additional increase in the mutation rate would abolish viral replication. By applying deoxynucleoside analogs to HIV replicating *in vitro*, he observed a complete loss of viral replicative potential and thus a lethal mutagenesis.

Generally, the theoretical and experimental results on replication error thresholds are not just true for viruses. They also remain valid for all replicating individuals that can be described by the formalism developed by Eigen (**Table 1**).

The complexity that life has developed by successive evolutionary optimization does not just result from vertically transmitted point mutations. ("Vertically" expresses the hereditary transmission of errors to the direct offspring.) Additional principles for the modification and expansion of genomes exist that enable the horizontal transfer of genetic information, i.e., the transfer between individuals or species. These additional principles are homologous genetic recombination, the basic mechanism of sexual heredity, and the transposition and integration of genes via mobile information-carrying elements, such as plasmids and viruses.

Table 1 Stepwise Expansion of information capacity during the evolution of life

Replication process	*Subject/organism*	*Error rate*	*Maximal genome length*	*Example*
Enzyme-free replication (base pairing)	AU polymer	10^{-1}	10	80 (tRNA ancestors)
	GC polymer	10^{-2}	100	
Enzyme-catalyzed RNA replication	RNA viruses	10^{-4}	10^{4}	4500 (bacteriophage Qβ)
Enzyme-catalyzed DNA replication	Prokaryotes	10^{-7}	10^{7}	4×10^{6}(*Escherichia coli*)
Recombinative, enzyme-catalyzed DNA replication	Eukaryotes	$\leq 10^{-10}$	10^{10}	3×10^{9}(man)

Hypotheses on the Hypercycle as an Ordering Principle

Erroneous autocatalytic copying is the reason for selection and for the emergence of cooperation between nucleic acids and proteins. In the beginning of the prebiotic era, nucleic acids probably were the first self-replicating entities. The fidelity of their replication was mainly dependent on the stability of the complementary base pairs AU and GC that formed during the copying procedure. Based on their different hydrogen-bonding energies, very high error rates resulted that reached 1:10 for AU-rich polymers and 1:100 for GC-rich polymers. According to theory, these error frequencies led to a maximal information content comprising 10–100 symbols for nucleic acid replicators. However, even small enzyme molecules are coded by more than several hundred nucleotides, i.e., far beyond the critical value that is defined by enzyme-free replication. Thus, an information crisis resulted from the error threshold relation: Any increase in the coding capacity required a decrease in the replication error rate, and this in turn could be achieved only by application of error-correcting mechanisms. Contemporary replication machineries achieve "proof-reading" with protein enzymes replicating the genetic information which is stored in double-stranded DNA. In contrast to RNA, a newly synthesized daughter strand of DNA remains associated to its parental strand and thereby enables the comparison of both sequences. False incorporations can be identified by enzymes that test the shape of the double strand for true base pairing and that correct unpaired or mispaired regions by substitution. However, how did the first replication machinery with error-correcting facilities evolve if the coding capacity needed for it could not faithfully be replicated without the very machinery? RNA might represent a major connecting link as was demonstrated recently by David Bartel (1963–) and colleagues. They generated RNA that synthesizes RNA using the same reaction as that employed by protein enzymes that catalyze RNA polymerization. In the presence of appropriate template RNA and nucleoside triphosphates, the ribozyme extends an RNA primer by successive addition of up to six mononucleotides, thereby showing significant template fidelity.

Eventually, the limiting chain length of approximately 100 nucleotides was overcome by the capability for translation of nucleic acid into protein. The appearance of RNA molecules which acted as adaptors or interpreters between nucleic acid and protein enabled the development of feedback loops and therefore the basis for cooperation. Together with these adaptor molecules, known as transfer-RNA (tRNA), the genetic code originated. tRNAs probably arose early and played an essential role in ancient replicating systems. Even today, tRNA and tRNA-like structures are involved in a variety of replicative processes, including replication of RNA viruses of bacteria, plants, and possibly mammals. Alan Weiner (1951–) therefore proposed that tRNA-like motifs emerged as 3′terminal structures that tagged RNA genomes for replication before the advent of protein synthesis. These tags could have served two main roles—providing an initiation site for replication and functioning as a simple telomere, i.e., ensuring that critical terminal regions were not lost during replication.

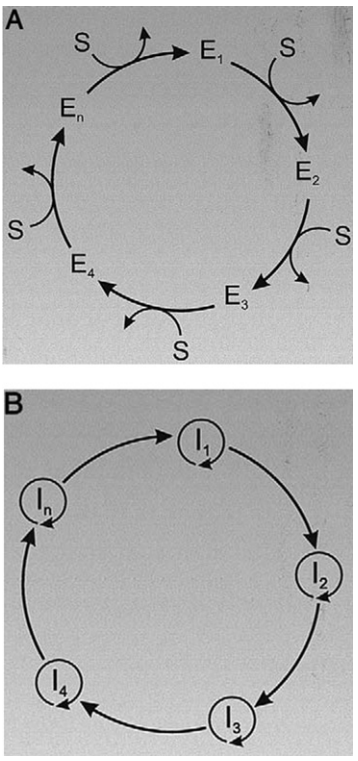


Figure 3 (A) Schematic representation of a catalytic cycle which represents a higher level of organization in a hierarchy of catalytic schemes. The constituents of the cycle ($E_1 - E_n$) are catalysts which are formed from energy-rich substrates (S). Each intermediate E_i acts as a catalyst for the formation $E_i + 1$. The catalytic cycle as a whole is equivalent to an autocatalyst. Replication of single-stranded RNA that involves mutual instruction (self-instruction) by complementary interaction represents an example of biological importance. (B) Schematic representation of a catalytic hypercycle that consists of self-instructive units with two-fold catalytic function. Each of the intermediates I_i which represent autocatalysts or catalytic cycles of the type shown in Figure 3A is able to instruct its own replication and to provide catalytic support for the production of the subsequent intermediate. For example, a simple hypercyclic process is found in nature with the infection cycle of bacteriophage Qβ.

The first proteins that arose probably required more genetic information than a primitive molecular replicator was able to provide. Because more information needed to be stabilized, cooperation of differentiated nucleic acid molecules emerged that was mediated and regulated by proteins, their translation products. The enforced cooperation of otherwise competing genes allowed their mutual survival and regulated their growth. It also enabled a more refined competition–selection behavior than that among the individuals of a quasispecies. Instead of single RNA mutants, cooperative ensembles of nucleic acid and protein were evaluated according to their self-replication qualities and according to their stabilities. First, those sequences were selected as fittest that were best able to get themselves replicated as quickly and as accurately as possible by the enzyme responsible for their replication. Second, with the continual introduction of new nucleic acid mutants, new catalytic couplings were constantly being tested.

The cooperation results in a double-feedback loop in which both the enzyme encoded by the nucleic acid template and the sequence information contribute to the replication of the template, a process called second-order autocatalysis. Phenomena of this particular kind were named "hypercycles" by Eigen and coworkers.

Hypercycles exhibit three major characteristics. First, they consist of coexisting and cooperating quasispecies, each of which is replicating autocatalytically and thereby maintaining competitive growth, selection, and its specific genetic information. Second, some of the coexisting quasispecies replicate and evolve independently. Finally, others form cooperative units within the hypercyclic organization and evolve together with mutual advantage. In an extreme case, some species as

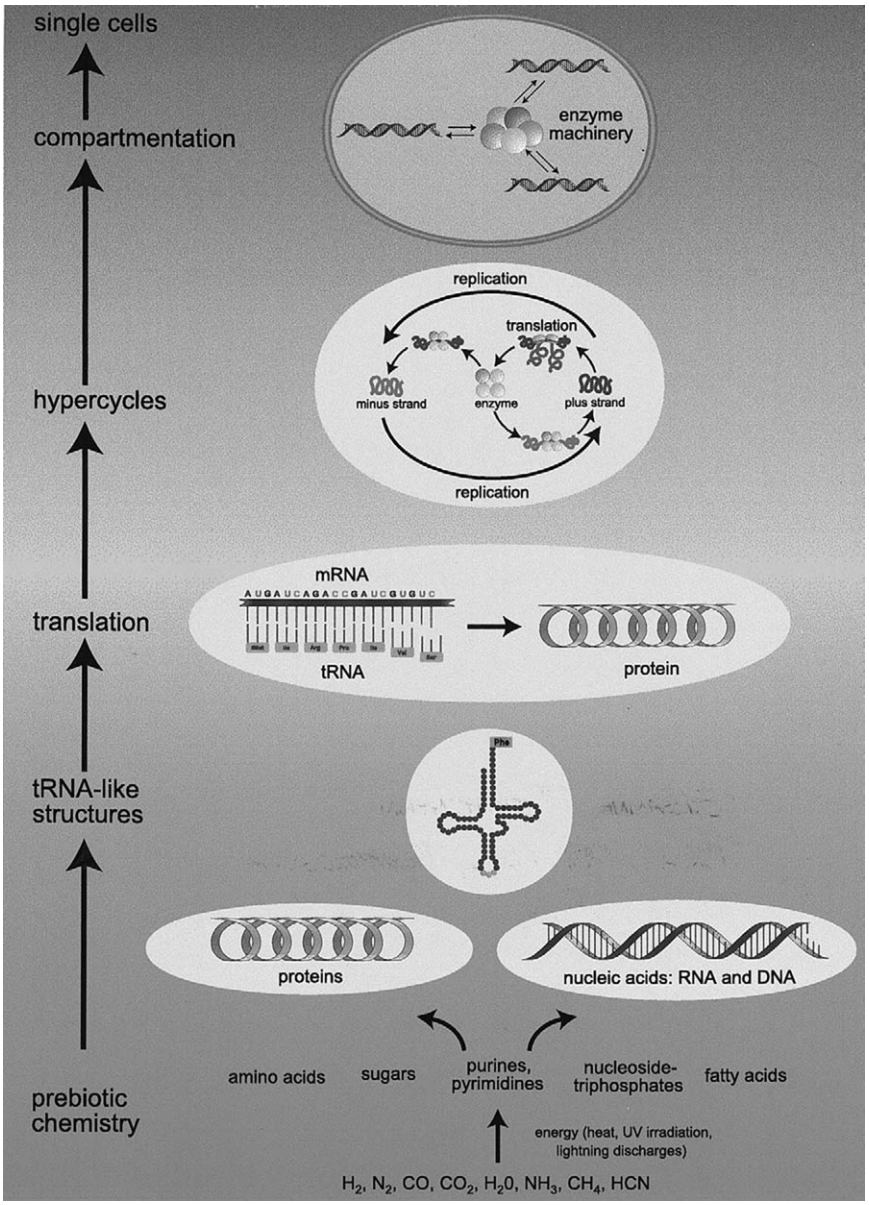


Figure 4 Stages of emerging life.

well as quasispecies may exist only by participating in hypercyclic organization. As a whole, the cooperating units can compete with other hypercyclically replicating systems.

The overcoming of the information crisis of self-replicating nucleic acids probably proceeded via hypercyclic coupling as shown in **Figure 3**, and it supported the emergence of "translation" from nucleic acid into protein. It is assumed that hypercycles represent one major step in the organization of complex reaction networks which arose as naturally and continuously as did quasispecies. Experimentally, hypercyclic organization and its efficiency were demonstrated by Eigen and coworkers to mainly determine the infection and replication cycle of a simple bacteriophage such as Qβ.

From Hypercycles to Cells

All life on Earth is cellular and reproduces only when this complex replicative unit, the cell, divides. Among the obvious advantages of cellular organization are protection of the cell's content from fluctuations of resources in the external environment and the maintenance of internal concentration gradients necessary to drive chemical reactions. Nevertheless, these facts do not explain the necessity for cellular organization: Instead, the origin of spatially limited reaction systems from a homogeneous prebiotic soup appears to be a result of problems in information processing.

Although hypercycles are the essential organizational form for the transition from self-replicating molecules to reproductive, multi-molecular machineries, they do not represent the ultimate optimum of organization. Hypercycles and quasispecies share a substantial evolutionary disadvantage: Both quasispecies competition and hypercyclic cooperation evaluate only the phenotypic properties of replicating nucleic acids, i.e., their replication rates and their stabilities. Hypercyclic coupling enables the discovery of molecular replicators which are beneficial to one or more others, and it supports the selection of those coexisting replicator combinations which gain mutual advantage by amplifying themselves. However, hypercycles cannot evaluate genotypic properties of nucleic acid sequences or their genetic messages. In other words, they do not selectively amplify mutant sequences that demonstrate their advantages only before translation.

Primitive translation mechanisms probably gave rise to proteins that were more helpful to self-replication than miscellaneous proteins randomly occurring in a homogeneous distribution. With time, preferences between translation products and certain sequences became pronounced, and these more distinct interactions effected advantages due to more specific catalysis. Finally, the differences among the various template–catalyst interactions became so influential that each enzyme had a particular catalytic role. At this stage, selection acted not only on the kinetic characteristics (i.e., the phenotype) of the replicating sequences but also on the information content (i.e., the genotype) of the sequences. Within the enlarged information system, the amount and quality of coding genes were strongly determined by efficiency, fidelity, and rate of the catalyzing enzymes, and these in turn were the translation products of their templates. By spatial separation of individual hypercycles, selective genotypical advantages in competition with other compartments were exploited. Favorably mutated compartments could develop more strongly than others and thus evolve more quickly. The compartmentalization was accomplished by individualization of all replicative units represented by a hypercycle. This event can be viewed as the birth of the cell, the smallest living entity known to us today. **Figure 4** summarizes the complete view that results from these theoretical descriptions.

See also: Archaea, Origin of. Biodiversity, Origin of. Darwin, Charles (Darwinism). Eukaryotes, Origin of. Evolution, Theory of. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms

References

De Duve C (1991) *Blueprint for a Cell: The Nature and Origin of Life*. Burlington, NC: Carolina Biological Supply.

Eigen M and Schuster P (1979) *The Hypercycle. A Principle of Natural Self-Organization*. New York: Springer.

Eigen M and Winkler-Oswatitsch R (1996) *Steps towards Life: A Perspective on Evolution*. Oxford: Oxford Univ. Press.

Origins, Evolution, and Breakdown of Bacterial Symbiosis

Joel L Sachs, University of California, Riverside, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bacterial mutualist A bacterium that interacts closely with a host and significantly enhances the host's fitness.

Bacterial parasite A microbe that interacts closely with a host and significantly reduces the host's fitness.

Endosymbiotic bacteria Bacterial species that live within the cells of hosts, are transmitted by hosts to their offspring, and often cannot survive outside of the host.

Infection Colonization of a host organism by bacteria that reproduce within the host and are ultimately transmitted to new hosts.

Introduction

The primitive earth was dominated by single-celled bacteria and archaea; the two major lineages of unicellular organisms that lack nuclei and the membrane-bound organelles of the eukaryotes. Beginning with the first fossil evidence of life, approximately 3.5 billion years ago and continuing for 2 billion years thereafter, most evidence of life appeared bacterial and archaeal. Not until a period that began about 1.5 billion years ago did eukaryotic lineages emerge and only in the last 1 billion years did multicellular plants and animals begin to diversify and ultimately to dominate. Thus, we can imagine a period of at least 2 billion years in which bacteria and archaea persisted mostly as free-living cells in the terrestrial and aquatic habitats. The last 500 million years of earth's history has witnessed an explosion in eukaryotic diversity including major radiation events of animal and plant lineages. This diversification of eukaryotic lineages has also driven a massive and parallel radiation of symbiotic bacteria (Sachs *et al.*, 2011b). Although archaea can also form symbioses with eukaryotic hosts, these interactions do not appear as common or diverse.

In the last billion years bacterial lineages have evolved a diverse array of mechanisms to gain entry into the tissues and cells of eukaryotes and to proliferate within these hosts (Medina and Sachs, 2010; Toft and Andersson, 2010). Here, the author refers to both harmful and beneficial bacterial infections as symbioses, a term used to describe any persistent and intimate association between microbes (the symbionts) and host species (animals, plants, or other eukaryotes; Medina and Sachs, 2010). Much of the research uncovering the mechanisms of symbiotic infection has focused on microbial parasitism (Sachs *et al.*, 2011a). In parasitic interactions, the microbes infect and actively exploit hosts resulting in reduced host survival and reproduction (evolutionary fitness). Bacterial parasites are a major challenge to human health and also represent a huge cost in terms of lost crops and animal stocks. However beneficial bacterial infections are known to be just as ubiquitous as parasites and can be found in virtually any eukaryotic species and in any ecosystem (Sachs *et al.*, 2011b). Such mutualistic infections occur when microbes infect and offer resources and services to hosts that enhance host fitness. Here the author's goal is to trace the deep evolutionary history of bacterial symbiosis and particularly to examine events leading to the origins and macroevolution of the bacterial mutualism.

Bacterial mutualists exhibit a great deal of lifestyle variation in their interactions with the eukaryotic hosts. Three key sources of variation are thought to bear particular importance to the evolution of mutualist bacterial lineages (Sachs *et al.*, 2011b). One source of variation is the degree to which the bacteria reproduce inside their host versus free in the environment. Many (and perhaps most) mutualist bacteria infect hosts but also spend a large portion of their lifecycle free in soils or aquatic environments (Sachs *et al.*, 2011a). The animal and plant hosts of these bacteria are born symbiont-free and thus must acquire their infections from the environmental pools (Szathmary and Maynard-Smith, 1995; Nyholm and Mcfall-Ngai, 2004; Sachs *et al.*, 2011a). At the opposite end of the spectrum are endosymbiotic bacteria and bacterially-derived organelles (Sagan, 1967; Keeling, 2010), which are the bacteria that cannot live independently without the hosts, only divide within host tissue and can only form new infections via transmission from host parent to offspring. A second source of variation is the array of animal and plant host tissues that bacterial mutualists can inhabit, including cell surfaces ranging from skin, mucosa, leaves and roots, and spaces between and within cells. In some cases the hosts have evolved specialized structures to house beneficial symbionts (e.g., Douglas, 1989; Bright and Sorgo, 2003; Nyholm and Mcfall-Ngai, 2004; Currie *et al.*, 2006; Goettler *et al.*, 2007; Vaishnava *et al.*, 2008) whereas in other cases beneficial bacterial symbionts can range widely within unstructured host tissues (Hirose *et al.*, 2009; Kaltenpoth *et al.*, 2009; Sachs *et al.*, 2011b). A final source of variation is represented by the array of mechanisms by which bacterial mutualists can provide fitness benefits to hosts, including the transfer of key nutrients, protection via toxicity as well as bioluminescence (Sachs *et al.*, 2011a, b). As described below each of these variables can have implications for the long-term evolution of symbiotic bacterial lineages.

The author describes how the study can be done on the diversification of such mutualist bacterial traits over the millions of years that they have developed and moreover how the study can be done on the origins of bacterial mutualist lineages and their persistence over evolutionary time. One approach to characterize such evolutionary history is to distill the key events and investigate the forces that shaped them (Szathmary and Smith, 1995). Bacteria can be understood to have undergone several major transitions in their evolutionary history and route to become mutualists with the eukaryotic

hosts. Some of these transitions represent key evolutionary milestones that have occurred many times in the history of bacteria. Five of these evolutionary transitions will be explored in depth over the course of this article (Sachs *et al.*, 2011b).

Evolutionary Origins of Bacterial Associations with Hosts

In the past history of bacteria, evolutionary origins of host association have occurred when environmentally existing bacteria initially evolved to form persistent and intimate interactions with the eukaryotic hosts (Sachs *et al.*, 2011b). Since the planet was first dominated only by the bacterial and archaeal cells, none of these organisms had large bodied hosts to live in or infect. Once eukaryotic lineages evolved, in particular the diversification of animals and plants, there were many evolutionary origins in which bacteria that once lived independently in the environment evolved to form intimate and persistent associations with these hosts (Sachs *et al.*, 2011b). For bacteria to undergo the transition to host association they must have had to overcome major hurdles, including evasion of host defenses, competition with other microbes to inhabit host surfaces, uptake of novel compounds in hosts and the ability to gain transmission from one host to the next (Sachs *et al.*, 2011b). To examine the difficulty of these potential hurdles, evolutionary methods can be used to assess the degree to which origins of host association are rare in bacterial lineages (Sachs *et al.*, 2011b).

One method to quantitatively investigate past evolutionary events is to build an evolutionary tree of genetically related species (a phylogeny) and then use the tree as a tool to test evolutionary hypotheses. The first step in this case is to reconstruct a species phylogeny of bacteria. This is becoming increasingly easier with the great availability of whole genome sequences (e.g., Wu *et al.*, 2009). The recent analysis by Wu and colleagues (2009) used genetic data from fully sequenced bacterial genomes to reconstruct an evolutionary tree of over 350 bacterial species. Because of the large amount of data that is available from the whole genome sequence, a very robust phylogeny can result (Wu *et al.*, 2009; Sachs *et al.*, 2011b; see **Figure 1**). The next step is to map species traits that can be added to the tree to study past events. In this case the traits of importance characterize each bacterial species in terms of their interactions with eukaryotic hosts (Boussau *et al.*, 2004; Merhej *et al.*, 2009; Toft and Andersson, 2010; Philippot *et al.*, 2010). Many bacterial species in this analysis exhibit no symbiotic interactions and are thus characterized as environmental whereas other species interact intimately with eukaryotic hosts and are categorized as parasites and mutualists depending on their fitness effects on hosts during infection (Sachs *et al.*, 2011b). Once a phylogenetic tree has been reconstructed for a lineage and phenotypes (traits) became known for each of the species, algorithms can be used to infer ancestral character states onto the tree. Such inference methods (known as ancestral state reconstruction) can be used to retrace the order of evolutionary events as well as the number of times the particular traits have independently evolved within a lineage.

Analyzing the evolution of host association origins across bacteria (Wu *et al.*, 2009; Toft and Andersson, 2010) a couple of conclusions can be made. First of all, simply from viewing the tree (**Figure 1**) it becomes clear that many independent origins of bacterial symbiosis have evolved. These origins can be found in almost every well-known lineage. For instance, on the 350 species tree a minimum of 42 origins of host-association are inferred from the environmental (free living) ancestral condition (Sachs *et al.*, 2011b). This pattern is interesting because it suggests that the hurdle in evolving from an environmental bacterium into a symbiont might not be as difficult as first imagined. Secondly, it appears that these origins of host association are more-or-less randomly distributed among bacteria (Sachs *et al.*, 2011b). Whereas some well-studied lineages appear to have many origins (Proteobacteria, Actinobacteria, and Firmicutes; **Figure 1**) and a few lineages (such as Chlorobi, Chloroflexi, and Planctomycetes) have never evolved the host association (Madigan *et al.*, 2009; Wu *et al.*, 2009; Toft and Andersson, 2010) there does not seem to be any significant bias where the origins occur on the tree (Sachs *et al.*, 2011).

To explore the mechanisms of symbiosis origins, researchers have begun to investigate the changes within the bacterial genomes that are correlated with these evolutionary transitions. One newly available approach is to study bacteria for which whole genome datasets are published and to compare the genomes of bacteria with different lifestyles. The goal here is to specifically retrace the particular sets of genes or genetic loci that are gained or lost concurrently with the origins of host association. Two interesting results have come from these studies so far. One pattern is that there are often few consistent genetic differences between host associated and environmental bacterial species (Carvalho *et al.*, 2010). Similarly, few genetic differences are often found between related bacteria that are beneficial versus those that are harmful to the hosts during infection (Sachs *et al.*, 2011a). Among the small set of genes that do differ in host associated and environmental species, many of them exhibit a strong evidence of horizontal gene transfer (movement of genes from one organism to an unrelated organism; Dale *et al.*, 2001; Horn *et al.*, 2004; Frank *et al.*, 2005; Ruby *et al.*, 2005; Ma *et al.*, 2006). Whereas only a few studies have rigorously compared the genetics of environmental versus beneficial or harmful host-associated bacteria (Dale *et al.*, 2001; Horn *et al.*, 2004; Frank *et al.*, 2005; Ruby *et al.*, 2005; Ma *et al.*, 2006; Carvalho *et al.*, 2010), one initial conclusion that can be made from these studies is that horizontal gene transfer is likely an important mechanism that promotes shifts in the lifestyles of bacteria. This pattern suggests that transitions to host association might be caused by a relatively smaller set of genes that are mobilizable among the bacterial lineages. At a broader level, these data suggest that very simple ecological factors might consistently promote or constrain a bacterium to gain the ability to infect eukaryotic hosts. These ecological factors might be as simple as access to habitats with compatible hosts and access to compatible genes (gained from other bacterial lineages) that allow an environmental bacterium to infect the hosts (Sachs *et al.*, 2011b).

The Origins of Bacterial Mutualism

Among the host-associated lineages, bacterial mutualists are incredibly diverse. Beneficial bacteria are diverse in the sense

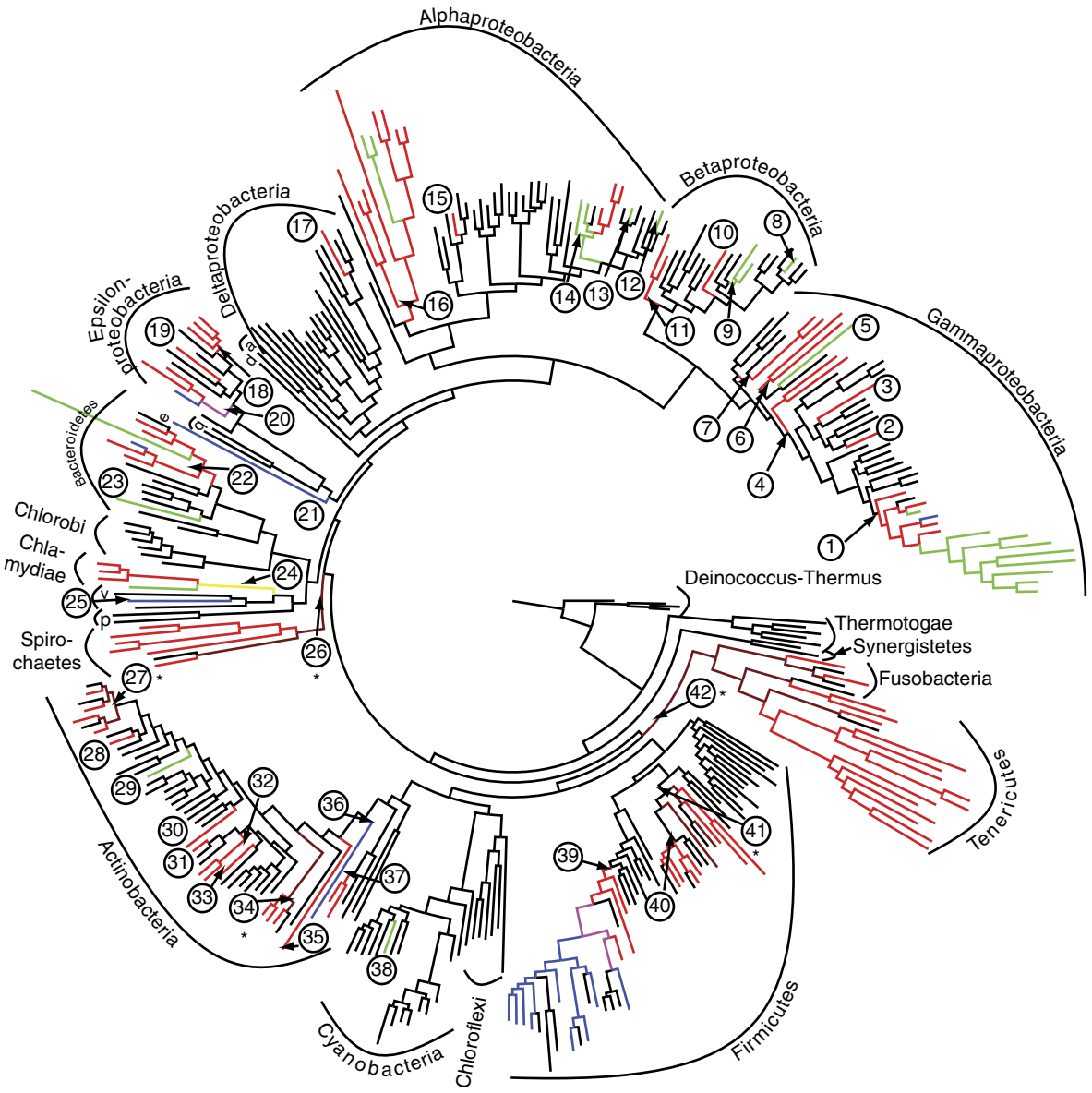


Figure 1 Inferred evolutionary history of bacterial host-association. Ancestral states of symbiosis are inferred on a bacterial evolutionary tree (modified from a previously published figure; Sachs JL, Skophammer RG, and Regus JU (2011b) Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences USA* 108:10800–10807). The tree is reconstructed from 31 single copy genes taken from 350 bacterial species for which whole genome data are available (Wu *et al.*, 2009). Phyla and proteobacterial classes are labeled with their full names (e.g. Gammaproteobacteria; Firmicutes) or single-letter abbreviations (a = Acidobacteria; d = Defferribacteres; q = Aquificae; e = Elusimicrobia; v = Verrucomicrobia; p = Planctomycetes). Branch colors represent host-associated traits on the tips of the tree and inferred states on ancestral nodes (black = environmental; blue = commensal; green = mutualist; red = parasite). Host association traits were obtained from a prior review (Toft and Andersson, 2010). Sachs and colleagues (2011b) inferred a minimum of 42 origins of host-association (labeled 1–42).

that this strategy has not only evolved independently in many anciently diverged lineages but also because so many types of hosts are infected in many different ways (Sachs *et al.*, 2011a). Other questions risen by the author include: what evolutionary process brought about these interactions in the first place? In other words, from the perspective of the bacterial evolutionary tree, what are the ancestral conditions that predate the origins of bacterial mutualists? Furthermore can evolutionary predictions be made about scenarios that favor the evolution of bacterial mutualists from other lifestyles?

Most of these questions remain unresolved. However, just as the author has done above for the origins of host association, information from bacterial phylogenies can be used to begin retracing the origins of bacterial mutualism. Perhaps the most perplexing question about bacterial mutualists is where they came from (Ewald, 1987; Szathmary and Smith, 1995; Corsaro *et al.*, 1999; Moran and Wernegreen, 2000; Medina and Sachs,

2010). Two competing hypotheses describe evolutionary scenarios for the origins of bacterial mutualism. One hypothesis proposes that bacterial mutualists have come about from parasitic ancestors that have evolved to be less harmful over time to the point that beneficial interactions could evolve with the hosts (Ewald, 1987). This attenuation of virulence is thought to be linked with the type of transmission that the bacteria exhibit (Ewald, 1987). Many bacteria are horizontally transmitted, meaning that individual bacteria travel from one host to a new and unrelated host to initiate a new infection. Horizontal transmission can actually promote parasitism of hosts because the ability to escape and infect a new host allows the bacterium to exploit the current host without harming itself (Fine, 1975). In contrast, vertical transmission occurs when bacteria are transmitted from a host to its offspring and in this case the bacteria can benefit themselves by being mutualistic (Fine, 1975). In a seminal paper Paul Ewald hypothesized that microbial mutualism can evolve from parasitism when there are evolutionary shifts from horizontal to vertical transmission among the hosts (Ewald, 1987). The author reviews data for and against this hypothesis below. The competing hypothesis states that bacterial mutualists must evolve directly from ancestors that lived free in the environment (Ewald, 1987). This hypothesis is based on the idea that it is difficult or impossible for bacteria switching from parasitism to mutualism because of constraints of genetic architecture (Moran and Wernegreen, 2000).

To examine empirical support for these scenarios we can again use the evolutionary reconstruction of bacterial species to infer past evolutionary transitions. However, this kind of analysis (that investigates transitions between different bacterial lifestyles) must be used with caution. The goal of hypothesis testing here is to examine general patterns of evolutionary transitions, however the dataset that is being used is relatively small because there are very few bacterial species on the tree (350 compared to the likely millions of actual bacterial species on earth), and because only a fraction of them ever associate with hosts. This sparse sampling means that many transitions are missing and that some inferred transitions might be spurious because intermediate steps have not been sampled. With these caveats in mind it can seen in the dataset that both the above hypotheses are supported. Three independent transitions from parasitism to mutualism can be observed on the tree and also mutualist lineages are inferred to have originated directly from the environmental ancestors (**Figure 1**). Unfortunately, there is not enough data on the tree to sufficiently test Ewald's hypothesis (1987) that the evolution of vertical transmission is responsible for the transitions from parasitism to mutualism. Phylogenies that examine specific bacterial lineages that include different host-associated lifestyles have also been used to examine and test these hypotheses and are consistent with the conclusion that mutualist bacteria can evolve from both environmental and parasitic ancestors (Sachs *et al.*, 2011b).

The Evolutionary Maintenance of Bacterial Mutualism

Mutualism is inherently unstable over evolutionary time (Sachs and Simms, 2006). For any cooperative interaction among species to be maintained, a mechanism must be in place to prevent the invasion of individuals that gain benefits from the interaction but do not pay any costs (cheaters). Since cheaters only take but do not give in an interaction they can be strongly favored by natural selection. The challenge then, is to explain how natural selection can promote cooperative individuals in a mutualism over cheaters. Three types of models for the maintenance of mutualism have been proposed (Axelrod and Hamilton, 1981; Bull and Rice, 1991; Sachs *et al.*, 2004; Foster and Wenseleers, 2006). The simplest of these models is called byproduct cooperation, which occurs when the bacterial partner cooperates with hosts as an automatic consequence of an unrelated and selfish trait (Sachs *et al.*, 2004). For instance, many bacteria produce toxins and antibiotics to protect themselves in the environment and these bacterial traits can automatically offer protection to hosts. One interesting outcome of such 'selfish cooperation' is that there is no opportunity for the bacterium to cheat since the cooperative trait usually bears no net cost of the bacterium. The next type of model, called partner fidelity feedback, works under the idea that there is a feedback in benefits between the bacterial and host partners. Feedbacks can work whenever the host and symbiont interact repeatedly or over a long enough period that their fitness interests become linked. The best example of this linkage has already been described in the example of vertical transmission. When hosts transmit bacteria from parent to offspring, both the bacterium and host can share a common interest in maximizing each other's survival and reproduction. The last type of model, called partner choice, occurs when the host can exert some level of control over the bacteria and has the ability to act as a selective force in the bacterial population. Specifically, with partner's choice the host can selectively favor cooperative over uncooperative infections and thus select for mutualists (and select against cheater mutants). To some degree, each of these models was developed without any particular biological interaction in mind and none of these models was initially created to study microbial symbioses. Hence, an interesting question to address is the degree to which different bacterial–host mutualisms are stabilized by byproducts, partner fidelity, or partner choice (Sachs *et al.*, 2011b).

Byproducts cooperation has not had a great deal of attention from empirical biologists. In part, this might be because it can be difficult to resolve byproducts cooperation into clear mechanisms. Yet, a feature that differentiates byproducts from partner fidelity feedback and partner choice is that the cooperative act bears no net cost for the bacterium; the trait is beneficial to the bacterium for some purpose unrelated to the host and exists irrespective of the host interaction. As described above, bacteria that exhibit toxicity or antibiotics might fulfill this description well. For instance, bacteria in the lineage Actinomycetes produce antibiotics as a benefit for their ant hosts. In this case the ants are fungus farmers; they grow fungi from the leaf litter collected in the forest and then eat the fungus for sustenance. The farming ants use the Actinomycete bacteria that grow on their exoskeleton to keep their fungal gardens pathogen-free (Currie *et al.*, 1999). As described above the trait of antibiotic production is a function that also benefits bacteria directly when it is in the environment, and this trait is expressed whether the bacteria are in the host or

not (Currie *et al.*, 1999; Sachs *et al.*, 2011b). Another likely example comes from bacteria that inhabit the lower gastrointestinal tracts of humans, *Bacteroides thetaiotaomicron*. One of the benefits that these bacterial mutualists are thought to provide for humans is to breakdown macromolecules that we cannot otherwise digest because we lack the enzymes to do it (Sonnenburg *et al.*, 2005). This is likely true for many functions that occur within the human body as the bacteria that inhabit us and carry our beneficial functions within the human body contain about 100 times as many genes as are found in our genome. Specifically in the case of *B. thetaiotaomicron*, the action of breaking down complex energy-filled food molecules into simpler compounds also must benefit the bacteria directly (Sachs *et al.*, 2011b). As there has been so little research on this topic there are likely many more examples such as these, but more work needs to be done to understand the breadth and importance of the byproducts cooperation in bacterial mutualists.

Partner-fidelity between bacteria and their hosts occurs when the bacteria are vertically transmitted from hosts to their offspring (Fine, 1975) and more generally in any scenario in which the symbiont interacts intimately with the host for a prolonged period (Sachs *et al.*, 2004; Sachs and Wilcox, 2006). Whereas vertical transmission is only one mechanism to maintain partner fidelity between bacteria and hosts, it remains the only driver that is well characterized or understood. In most cases vertically transmitted bacteria are very cooperative with hosts, but some exceptions exist. Some vertically transmitted bacteria harmfully manipulate the host's reproduction to maximize their own spread in the host population. For instance, many vertically transmitted symbionts are only passed on through females to offspring (via the ovum) and thus male hosts represent an evolutionary dead-end to these symbionts. In the case of *Wolbachia*, a common bacterial symbiont of most insects, have sometimes evolved mechanisms to bias host sex ratio toward females in order to maximize their own transmission (Stouthamer *et al.*, 1999).

Partner choice occurs when hosts exhibit mechanisms to selectively benefit beneficial bacterial partners and or punish harmful strains (Sachs *et al.*, 2004). Yet, how can a large eukaryotic host such as a plant or animal individually recognize individual bacterial genotypes and species and react toward them differentially in an adaptive manner? Despite a great deal of research, this question remains one of the major unsolved mysteries of symbiosis. One recent idea of how hosts can manage a complex group of genetically diverse symbionts is that hosts spatially separate their symbionts so that each genotype can be dealt with on an individualized basis (Denison, 2000; Sachs *et al.*, 2004, 2011b; West *et al.*, 2002a, b). Consistent with this hypothesis, many hosts of beneficial bacteria have evolved specialized structures that only exist in the context of a bacterial interaction and that have the potential to order symbionts into a fine level of spatial structure (e.g., Douglas, 1989; Bright and Sorgo, 2003; Nyholm and Mcfall-Ngai, 2004; Currie *et al.*, 2006; Goettler *et al.*, 2007; Vaishnava *et al.*, 2008). In many of these hosts, little is known about how or whether the structure promotes partner choice. Perhaps the best understood system is in legumes that are infected with root-nodule forming bacteria (rhizobia) that fix nitrogen for the host in exchange for plant sugars (Sprent *et al.*, 1987). In this case each nodule is most often infected by a single rhizobial genotype and thus nodules serve as a structural means for the plant to reward or punish rhizobial strains depending on how cooperative they are (Kiers *et al.*, 2003; Simms *et al.*, 2006; Sachs *et al.*, 2011a). The question is whether other hosts can wield such power over their bacterial symbionts. Some animals that host gastrointestinal bacteria have evolved complex crypt structures within their guts that appear specially to compartmentalize bacterial symbionts. But whether these crypt structures actively enforce partner choice is poorly understood. Humans exhibit crypt structures in their gastrointestinal tracts that bear importance to interaction with gut flora (Vaishnava *et al.*, 2008) but it is not clear whether these crypts are necessary for the maintenance of the mutualism or not. In some lineages of stinkbugs, the hosts have bacterial symbionts in the genus *Burkholderia* (Kikuchi *et al.*, 2011) that proliferate in the specialized gut crypts. Interestingly, only the stinkbug species that exhibit these crypts form beneficial symbioses with *Burkholderia*. This pattern suggests that the structural separation of the symbionts is a prerequisite to the evolutionary maintenance of the interaction (Sachs *et al.*, 2011b), but more data are needed.

The Evolutionary Capture of Bacterial Symbionts within Host Lineages

Most beneficial bacteria exhibit a rich existence both inside and outside their hosts. Since symbiotic bacteria are often environmentally acquired by hosts, the bacteria experience long phases in which they persist in soils or water separated by infection and growth within the hosts. To maintain this dual lifestyle the host-associated bacteria must be able to replicate in diverse environmental settings and must also be able to successfully localize and infect new hosts (Bright and Bulgheresi, 2010). As discussed above in the paragraph on partner–fidelity feedback, a subset of bacteria have evolved the ability to be vertically transmitted within host lineages (parent to offspring) and thus have lost their independence. The evolution of strict vertical transmission allows the symbionts to completely lose any environmental phase and thus persist only within the tissues or cells of hosts. How does this process of symbiont capture evolve and is it driven by the hosts or bacteria themselves (Sachs *et al.*, 2011b)?

A conflict of interests over symbiont capture appears to exist. Whereas vertically transmitted symbionts appear to suffer reduced fitness over time, their hosts experience clear benefits (Frank, 1996; Sachs *et al.*, 2011b). A great deal of data has accumulated regarding the genomic consequences of evolving vertical transmission for bacteria. When bacteria make the transition from free-living to strict vertical transmission their population size drops as a consequence of losing environmental phases. This simple change is thought to have enormous consequences. One of the fundamental rules of population biology is that natural selection only works efficiently on large populations: once populations drop to a small enough size they are only effected by random genetic change and tend to evolve reduced fitness over time (Nilsson *et al.*, 2005). Consistent with this theory, vertically transmitted

bacteria are known to accumulate harmful mutations over time, lose genetic pathways (Moran, 2003; Toh *et al.*, 2006), and ultimately lose so much function that they are obliged to rely on the host for basic nutrient synthesis (Shigenobu *et al.*, 2000). Because of this evolutionary pattern of genome degradation, once vertical transmission evolves it can lead to an irreversible evolutionary endpoint for bacteria that ends in reliance on hosts (Sachs *et al.*, 2011b). Whereas all of these changes appear to be harmful for vertically transmitted symbionts, the hosts mostly experience benefits. Hosts that transmit their beneficial bacteria vertically guarantee that their offspring will also receive a beneficial infection and by virtue of partner fidelity promote selection for beneficial traits in their bacterial symbiont population (Sachs *et al.*, 2004; Sachs and Bull, 2005; Sachs, 2006).

Hosts appear to have evolved many types of structures and mechanisms that promote symbiont capture (Bright and Bulgheresi, 2010). In many cases hosts have specific structures that appear to have no purpose at all except to efficiently transfer bacteria to offspring (Sachs *et al.*, 2011b). It is fascinating to observe the diversity of specialized structures that animal hosts have evolved to engage in vertical transmission of bacteria (Bright and Bulgheresi, 2010). Moreover, animals have also evolved complex behaviors to pass on symbionts to offspring, for instance by smearing symbionts onto eggs, onto egg cases, into cocoons, or directly on growing offspring (Douglas, 1989; Kaltenpoth *et al.*, 2005; Hirose *et al.*, 2009; Kikuchi *et al.*, 2009; Kojima and Hirose, 2010; Kaltenpoth *et al.*, 2010). In contrast to these well-studied mechanisms, little is known about the ability of bacteria to promote or counteract symbiont capture. In some interesting cases bacteria can be seen to actively migrate within their host to get to the ovaries (Douglas, 1989) or even move outside the parent to infect offspring (Ran *et al.*, 2010). Nonetheless, these examples are not necessarily evident against hosts being responsible for symbiont capture. This is because once capture has evolved the bacteria have no other option to infect new hosts other than vertical transmission (Sachs *et al.*, 2011a) Whereas hosts clearly benefit from the structures and behaviors that promote symbiont capture, little is known about the evolution of capture from the symbiont's perspective.

Evolutionary Breakdown of Bacterial-eukaryotic Symbioses

A great deal of theoretical work has modeled the ecology and evolution of mutualistic interactions. Although bacterial mutualists are extremely common, much of this theory predicts that evolutionary instability which leads to the breakdown of mutualism over time (Sachs and Simms, 2006). One major prediction that has come from this work is that mutualisms are prone to shift into parasitism as cheater mutants invade mutualist populations (Axelrod and Hamilton, 1981; Bull and Rice, 1991). Another prediction is that mutualists can evolve to abandon their partners if costs of the interaction come to outweigh potential benefits (Keeler, 1985; Holland *et al.*, 2004). To what degree does evidence support these hypotheses of 'mutualisms-breakdown' (e.g., Sachs and Simms, 2006)? As described above, phylogenetic analyses can be used to retrace the steps of evolution and infer how often mutualist bacterial lineages shift into other lifestyles. Interestingly, on the phylogenetic tree described above that samples lineages across bacteria (Wu *et al.*, 2009; Toft and Andersson, 2010; **Figure 1**) there are only two observed switches in which bacterial mutualists shift into other lifestyles. One of these shifts is a transition from a mutualist to a parasite and the other is a reversion back to an environmental (nonhost-associated) lifestyle (Sachs *et al.*, 2011b). Compared to the total number of lifestyle switches on the tree (72; Sachs *et al.*, 2011b) this dataset makes mutualism-breakdown seem surprisingly rare among bacterial lineages. Yet, it is difficult to tell whether this pattern represents an accurate reflection of bacterial evolution, or if it is an artifact of the dataset. As the author described above, the one big challenge to this kind of analysis is that so much of the dataset is missing: the tree represents a very sparse sampling of the actual bacterial diversity that exists.

For filling the potential gaps on the bacterial tree (Wu *et al.*, 2009; Toft and Andersson, 2010; **Figure 1**) one option is to 'zoom in' and investigate the potential for mutualism among much more closely related species or even strains of bacteria. Only in the last several years has there been a pulse of phylogenetic studies that have sampled bacterial lineages deeply enough to investigate the losses or gains of mutualism with hosts. Although several recent studies show repeated losses (abandonment) of bacterial mutualism (Nishiguchi and Nair, 2003; O'Brien *et al.*, 2005; Mueller *et al.*, 2010; Sachs *et al.*, 2010; Kikuchi *et al.*, 2011) no phylogenetic study that the author is aware of has uncovered a transition from mutualism to parasitism within a well-studied lineage. Hence, mutualism-breakdown appears to occur when bacterial mutualists evolve to abandon hosts (and revert to an environmental lifestyle) but there is extremely little evidence to support the idea that cheaters evolve within bacterial mutualist populations and drive transitions from mutualism to parasitism (Sachs *et al.*, 2011b).

Discussion

The history of bacterial interactions with animal, fungal, and plant hosts likely stretches back to the very origins of these lineages. It is virtually impossible to imagine organisms in any of these lineages existing without many intimate interactions with bacterial species (both harmful and beneficial). In the article the author has reviewed what is known about the evolution of bacterial interactions with eukaryotic hosts with a special focus on bacteria that cause beneficial infections (mutualists). Particularly, the author has explored the origins, maintenance, and ultimate breakdown of the bacterial interactions with hosts.

From the available data it can concluded that the evolutionary milestone of bacteria first achieving symbioses with hosts has evolved many times in almost every bacterial lineage that is known to science. The repeated and convergent evolution of persistent and intimate interactions between bacteria and hosts suggests that these interactions are highly advantageous to bacteria, relatively easier to evolve and mostly independent of the specific habitat requirements (Sachs *et al.*, 2011b). Among the many intimate interactions that bacteria

exhibit with eukaryotes a large subset are beneficial to hosts. These bacterial mutualists have evolved from the diverse ancestors that include parasitic species (that evolved over time to offer benefits to hosts) as well as environmental bacteria (that evolved from unassociation with hosts directly to a beneficial interaction; Ewald, 1987; Sachs *et al.*, 2011b). Among these types of origins it appears that bacterial mutualism most often evolves directly from the ancestors that live free in the environment.

Bacteria that have evolved mutualist interactions with hosts are incredibly diverse and provide almost any kind of nutrient or biological service to hosts that can be imagined (Sachs *et al.*, 2011a). In some cases hosts have evolved mechanisms to actively select bacterial cooperation, for instance by selectively rewarding beneficial strains or by punishing cheaters (Sachs *et al.*, 2004). In other cases bacterial cooperation is maintained in a more passive manner because host and symbiont exhibit a feedback in terms of fitness benefits (Fine, 1975). Most bacteria that provide beneficial infections to hosts are acquired from environmental sources and exhibit horizontal transmission among hosts. In some cases, bacterial symbionts have switched to vertical transmission and thus can lose any environmental phase and evolve to live only within the hosts. Although transitions from horizontal to strict vertical transmission (symbiont capture) involve both host and symbiont evolution, both theory and evidence suggest that hosts are likely to dominate control of these transitions. Finally, the evolutionary breakdown of bacterial mutualism appears to be relatively rare compared to other transitions in bacterial symbiosis. This lack of transitions from mutualism to parasitism or other bacterial lifestyles suggests either that bacterial mutualisms are evolutionarily robust or that transitions from mutualism to other lifestyles are themselves unstable (and lead to extinctions or other stable states; Sachs and Simms, 2006).

See also: Adaptive Radiation. Geologic Time, History of Biodiversity in. Phylogeny

References

Axelrod R and Hamilton WD (1981) The evolution of cooperation. *Science* 211: 1390–1396.

Boussau B, Karlberg EO, Frank AC, Legault BA, and Andersson SGE (2004) Computational inference of scenarios for alpha-proteobacterial genome evolution. *Proceedings of the National Academy of Sciences of the United States of America* 101: 9722–9727.

Bright M and Bulgheresi S (2010) A complex journey: Transmission of microbial symbionts. *Nature Reviews Microbiology*. 8: 218–230.

Bright M and Sorgo A (2003) Ultrastructural reinvestigation of the trophosome in adults of *Riftia pachyptila* (Annelida, Siboglinidae). *Invertebrate Biology* 122: 347–368.

Bull JJ and Rice WR (1991) Distinguishing mechanisms for the evolution of cooperation. *Journal of Theoretical Biology* 149: 63–74.

Carvalho SFM, Souza RC, Barcellos FG, Hungria M, and Vasconcelos ATR (2010) Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order Rhizobiales. *Bmc Microbiology* 10: 37.

Corsaro D, Venditti D, Padula M, and Valassina M (1999) Intracellular life. *Critical Reviews in Microbiology* 25: 39–79.

Currie CR, Poulsen M, Mendenhall J, Boomsma JJ, and Billen J (2006) Coevolved crypts and exocrine glands support mutualistic bacteria in fungus-growing ants. *Science* 311: 81–83.

Currie CR, Scott JA, Summerbell RC, and Malloch D (1999) Fungus-growing ants use antibiotic-producing bacteria to control garden Parasites. *Nature* 398: 701–704.

Dale C, Young SA, Haydon DT, and Welburn SC (2001) The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. *Proceedings of the National Academy of Sciences of the United States of America* 98: 1883–1888.

Denison RF (2000) Legume sanctions and the evolution of symbiotic cooperation by rhizobia. *American Naturalist* 6: 567–576.

Douglas AE (1989) Mycetocyte symbiosis in insects. *Biological Reviews of the Cambridge Philosophical Society* 64: 409–434.

Ewald PW (1987) Transmission modes and evolution of the parasitism–mutualism continuum. *Annals of the New York Academy of Sciences* 503: 295–306.

Fine PEM (1975) Vectors and vertical transmission: An epidemiological perspective. *Annals of the New York Academy of Sciences* 266: 173–194.

Foster KR and Wenseleers T (2006) A general model for the evolution of mutualisms. *Journal of Evolutionary Biology* 19: 1283–1293.

Frank AC, Alsmark CM, Thollesson M, and Andersson SGE (2005) Functional divergence and horizontal transfer of type IV secretion systems. *Molecular Biology and Evolution*. 22: 1325–1336.

Frank SA (1996) Host control of symbiont transmission: The separation of symbionts into germ and soma. *American Naturalist* 148: 1113–1124.

Goettler W, Kaltenpoth M, Herzner G, and Strohm E (2007) Morphology and ultrastructure of a bacteria cultivation organ: The antennal glands of female European beewolves, *Philanthus triangulum* (Hymenoptera, Crabronidae). *Arthropod Structure and Development* 36: 1–9.

Hirose E, Neilan BA, Schmidt EW, and Murakami A (2009) Enigmatic life and evolution of Prochloron and related Cyanobacteria inhabiting colonial Ascidians. In: Gault PM and Marler HJ Hauppauge (eds.) *Handbook on Cyanobacteria: Biochemistry, Biotechnology and Applications*. New York: Nova Sciences Publishers Inc. pp. 161–189.

Holland JN, DeAngelis DL, and Schultz ST (2004) Evolutionary stability of mutualism: Interspecific population regulation as an evolutionarily stable strategy. *Proceedings of the Royal Society of London Series B – Biological Sciences* 271: 1807–1814.

Horn M, Collingro A, and Schmitz-Esser S (2004) Illuminating the evolutionary history of chlamydiae. *Science* 304: 728–730.

Kaltenpoth M, Gottler W, Herzner G, and Strohm E (2005) Symbiotic bacteria protect wasp larvae from fungal infestation. *Current Biology* 15: 475–479.

Kaltenpoth M, Goettler W, Koehler S, and Strohm E (2010) Life cycle and population dynamics of a protective insect symbiont reveal severe bottlenecks during vertical transmission. *Evolutionary Ecology* 24: 463–477.

Kaltenpoth M, Winter SA, and Kleinhammer A (2009) Localization and transmission route of *Coriobacterium glomerans*, the endosymbiont of pyrrhocorid bugs. *Fems Microbiology Ecology* 69: 373–383.

Keeler KH (1985) Cost: Benefitmodels of mutualism. In: Boucher DH (ed.) *The Biology of Mutualism, Ecology and Evolution*, pp. 100–127. Oxford: Oxford University Press.

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B – Biological Sciences* 365: 729–748.

Kiers ET, Rousseau RA, West SA, and Denison RF (2003) Host sanctions and the legume–rhizobium mutualism. *Nature* 425: 78–81.

Kikuchi Y, Hosokawa T, and Fukatsu T (2011) An ancient but promiscuous host–symbiont association between Burkholderia gut symbionts and their heteropteran hosts. *ISME Journal* 5: 446–460.

Kikuchi Y, Hosokawa T, Nikoh N, Meng XY, Kamagata Y, and Fukatsu T (2009) Host–symbiont co-speciation and reductive genome evolution in gut symbiotic bacteria of acanthosomatid stinkbugs. *Bmc Biology* 7.

Kojima A and Hirose E (2010) Transfer of prokaryotic algal symbionts from a tropical Ascidian (*Lissoclinum punctatum*) colony to its larvae. *Zoological Science* 27: 124–127.

Ma WB, Dong FFT, Stavrinides J, and Guttman DS (2006) Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *Plos Genetics* 2: 2131–2142.

Madigan MT, Martinko JM, Dunlap PV, and Clark DP (2009) *Brock Biology of Microorganisms*. San Francisco: Pearson Benjamin-Cummings.

Medina M and Sachs JL (2010) Symbiont genomics, our new tangled bank. *Genomics* 95: 129–137.

Merhej V, Royer-Carenzi M, Pontarotti P, and Raoult D (2009) Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. *Biology Direct* 4.

Moran NA (2003) Tracing the evolution of gene loss in obligate bacterial symbionts. *Current Opinion in Microbiology* 6: 512–518.

Moran NA and Wernegreen JJ (2000) Lifestyle evolution in symbiotic bacteria: Insights from genomics. *Trends in Ecology and Evolution* 15: 321–326.

Mueller UG, Ishak H, Lee JC, Sen R, and Gutell RR (2010) Placement of attine ant-associated Pseudonocardia in a global Pseudonocardia phylogeny (Pseudonocardiaceae, Actinomycetales): A test of two symbiont-association models. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* 98: 195–212.

Nilsson A, Koskiniemi S, Eriksson S, Kugelberg E, Hinton JCD, and Andersson DI (2005) Bacterial genome size reduction by experimental evolution. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12112–12116.

Nishiguchi MK and Nair VS (2003) Evolution of symbiosis in the Vibroonaceae: A combined approach using molecules and physiology. *International Journal of Systematic and Evolutionary Microbiology* 53: 2019–2026.

Nyholm SV and McFall-Ngai MJ (2004) The winnowing: Establishing the squid-Vibrio symbiosis. *Nature Reviews Microbiology* 2: 632–642.

O'Brien H, Miadlikowska J, and Lutzoni F (2005) Assessing host specialization in symbiotic cyanobacteria associated with four closely related species of the lichen Peltigera. *European Journal of Phycology* 40: 363–378.

Philippot L, Andersson SGE, and Battin TJ (2010) The ecological coherence of high bacterial taxonomic ranks. *Nature Reviews Microbiology* 8: 523–529.

Ran LA, Larsson J, and Vigil-Stenman T (2010) Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. *Plos One* 57: e11486.

Ruby EG, Urbanowski M, and Campbell J (2005) Complete genome sequence of *Vibrio fischeri*: A symbiotic bacterium with pathogenic congeners. *Proceedings of the National Academy of Sciences of the United States of America* 102: 3004–3009.

Sachs JL (2006) Cooperation within and among species. *Journal of Evolutionary Biology* 19: 1415–1418.

Sachs JL and Bull JJ (2005) Experimental evolution of conflict mediation between genomes. *Proceedings of the National Academy of Sciences of the United States of America* 102: 390–395.

Sachs JL, Ehinger MO, and Simms EL (2010) Origins of cheating and loss of symbiosis in wild Bradyrhizobium. *Journal of Evolutionary Biology* 23: 1075–1089.

Sachs JL, Essenberg CE, and Turcotte MM (2011a) New Paradigms for the Evolution of Beneficial Infections. *Trends in Ecology Evolution* 26: 202–209.

Sachs JL, Mueller UG, Wilcox TP, and Bull JJ (2004) The evolution of cooperation. *Quarterly Review of Biology* 79: 135–160.

Sachs JL, Russell JE, Lii YE, Black KC, Lopez G, and Patil SA (2010) Host control over infection and proliferation of a cheater symbiont. *Journal of Evolutionary Biology* 23: 1919–1927.

Sachs JL and Simms EL (2006) Pathways to mutualism breakdown. *Trends in Ecology and Evolution* 21: 585–592.

Sachs JL, Skophammer RG, and Regus JU (2011b) Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences USA* 108: 10800–10807.

Sachs JL and Wilcox TP (2006) A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proceedings of the Royal Society B – Biological Sciences* 273: 425–429.

Sagan L (1967) On Origin of Mitosing Cells. *Journal of Theoretical Biology* 14: 225–274.

Shigenobu S, Watanabe H, Hattori M, Sakaki Y, and Ishikawa H (2000) Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp APS. *Nature* 407: 81–86.

Simms EL, Taylor DL, and Povich J (2006) An empirical test of partner choice mechanisms in a wild legume–rhizobium interaction. *Proceedings of the Royal Society B – Biological Sciences* 273: 77–81.

Sonnenburg JL, Xu J, and Leip DD (2005) Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. *Science* 307: 1955–1959.

Sprent JI, Sutherland JM, and Faria SM (1987) Some aspects of the biology of nitrogen-fixing organisms. *Philosophical Transactions of the Royal Society of London, B. Biological Sciences* 317: 111–129.

Stouthamer R, Breeuwer JAJ, and Hurst GDD (1999) *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. *Annual Review of Microbiology* 53: 71–102.

Szathmary E and Smith JM (1995) The major evolutionary transitions. *Nature* 374: 227–232.

Toft C and Andersson SGE (2010) Evolutionary microbial genomics: Insights into bacterial host adaptation. *Nature Reviews Genetics* 11: 465–475.

Toh H, Weiss BL, and Perkin SAH (2006) Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Research* 16: 149–156.

Vaishnava S, Behrendt CL, Ismail AS, Eckmann L, and Hooper LV (2008) Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host–microbial interface. *Proceedings of the National Academy of Sciences of the United States of America* 105: 20858–20863.

West SA, Kiers ET, Pen I, and Denison RF (2002a) Sanctions and mutualism stability: When should less beneficial mutualists be tolerated. *Journal of Evolutionary Biology* 15: 830–837.

West SA, Kiers ET, Simms EL, and Denison RF (2002b) Sanctions and mutualism stability: Why do rhizobia fix nitrogen. *Proceedings of the Royal Society of London Series B – Biological Sciences* 269: 685–694.

Wu DY, Hugenholtz P, Mavromatis K, *et al.* (2009) A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462: 1056–1060.

P

Paleoecology

Thompson Webb, Brown University, RI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

AMS dates Radiocarbon dates obtained by directly measuring the amount of carbon 14 in a sample by using an accelerator mass spectrometer.
Bryophytes Mosses and liverworts; nonvascular plants.
Cyperaceae Sedge family.
Fagus grandifolia American beech trees.
Late Quaternary The past 21,000 years since the last time of maximum glaciation.
Palynology The study of pollen and spores.
Pliocene The period of geological time from 3.5 to 1.8 Ma.
Quaternary The past 1.8 My of geological time.
Radiocarbon dates Dates of organic matter in geological samples using the radioactive decay of carbon 14, which has a half-life of 5750 years.
Tracheophytes Vascular plants.

Introduction

Past ecosystems are a continuous creation of those who study them. What is mapped through time or reduced to time series depends on the data available and the training of those who interpret and display the data. Studies in paleoecology, therefore, are an exercise in perception and interpretation that depend, to some extent, on the sensitivity of the data in time, space, and taxonomy. The ultimate goal is to understand what occurred ecologically in the past, but this understanding can only be obtained if reliable records exist of past taxon distributions (or other components of ecosystems) that can be displayed and interpreted in informative ways. Simple time series are not enough because of the many spatial variations and processes that affect the temporal changes at a site. Geographic networks of temporally varying data are required to show how past and present taxa have varied in abundance, location, and association within and among their associated ecosystems.

The focus here is on the rich data sets of Quaternary pollen and plant macrofossils that provide records of the major changes in vegetation at a variety of space and timescales. These data (*see* Neotoma Paleoecology Database – http://www.neotomadb.org/) provide a valuable bridge between modern ecological data and the fossil evidence in the Paleobiology Database, which paleontologists are compiling. The Quaternary data sets yield a remotely sensed view of past vegetation, landscapes, taxon migrations, and invasions and illustrate various temporal changes in ecological assemblages and communities (Huntley and Birks, 1983; Williams *et al.*, 2004; Binney *et al.*, 2009) – all of which show changes and patterns in biodiversity. Direct measurement of the changing numbers of species through time, however, is not always possible with pollen data, even when supplemented by plant macrofossil data, although the extinction and arrival of some species are evident in records dating back to the Pliocene (Hoogheimstra and Ran, 1994). The focus with these paleoecological data, therefore, is on vegetation and its diverse changing patterns of composition and distribution through time and space and how environmental factors, such as climate, disease, fire, and human activities, have influenced these changes. Other types of data such as ground-penetrating radar records, sediment characteristics, diatoms, ostracodes, and various geochemical measures allow independent interpretations of past environmental changes and thus permit the pollen and plant macrofossil data to show the taxon-level and vegetational response to these past environmental changes (Shuman *et al.*, 2002, 2004). The combinations of data sets illustrate the ecological and environmental setting within which modern patterns of biodiversity have arisen. They also demonstrate, as long claimed by Richard West and Margaret Davis, that the "present plant communities have no long history."

Data and Sensing System

Just as satellite-derived vegetation data today can be displayed in time series of maps, the pollen and plant macrofossil data can be viewed as if they came from a video recorder from high in space with resolution in places of up to 10 m. The images that are retrieved can be of high or low resolution temporally, spatially, taxonomically, and numerically, and they can illustrate local to global changes in plant populations, vegetation,

biodiversity, human activity, fire frequency, and plant diseases over decades to millions of years. Because each of these entities or phenomena varies spatially and temporally, records of data covering a breadth of scales in space and time are needed. To obtain the highest quality images of a specific phenomenon requires an understanding of the sensing system that accumulated the data. How does the paleoecological video-recording system work and what are the scaling characteristics of the images that it registers? These characteristics include breadth of coverage, sampling resolution, and sampling density in time, space, and taxonomy. Studies of pollen and plant macrofossil data covering a variety of temporal and spatial scales within the Quaternary have helped provide this understanding.

Quaternary paleovegetation data pose several problems for interpretation. The main problem arises because pollen samples in sediments represent death assemblages of immature microgametophytes (i.e., pollen grains) that differ manifestly from the assemblages of sporophytes (i.e., plants) that produced the pollen. Studies of modern data from surface sediments have therefore focused on identifying the features of plant assemblages that appear in pollen records. This work parallels ground-truthing studies by scientists deciphering the remotely sensed data from satellites because pollen data are a form of remotely sensed data from plant populations and vegetation. Just as the current vegetation emits or reflects radiation that remote sensors on satellites intercept, so too does (and has) the current (and past) vegetation shed pollen that accumulates remotely (i.e., well away from the source) in lakes and bogs. Both types of remote sensors (satellite instruments and lake sediments) record data with certain sampling characteristics (e.g., spatial and temporal resolution), and their data need calibration and ground-truthing in terms of vegetation attributes such as composition (taxon abundances), structure (height and mixture of growth forms), and pattern (geographic gradients and mosaics and aspect of beta and gamma diversity). Studies of spatial arrays of modern data by Bradshaw, Prentice, Hicks, Calcote, and Jackson have provided this calibration and ground-truthing.

Down-core studies of pollen bring time into the picture, and temporal resolution becomes one of the factors controlling what is recorded in the data. Pollen in annually laminated sediments can provide seasonally distinct samples as neatly shown by Peglar *et al.* (1984), but most samples integrate 10 or more years and can be independently dated back 40,000 years with an average precision of ± 200 years for data from within the past 12,000 years. However, if vegetational and ecological phenomena are the target for study, then spatially distributed data are needed because vegetation and biodiversity are inherently spatial entities that vary on virtually all time and space scales. Recording their full dynamics requires time series of geographically distributed data that yield a zoom lens space–time perspective. The author focuses on how the different scaling characteristics of the data sets control what we see and on how data can be organized to yield such a zoom lens perspective. The author discusses each of the sampling characteristics in taxonomy, space, and time. The author then describes different ways to display and visualize the data and some of the vegetational dynamics that they have revealed.

Scaling Factors

Most Quaternary palynologists obtain their data from sediments from lakes, mires, or estuaries where (1) sediments and pollen grains accumulate relatively continuously, (2) the sediments are organic enough for radiometric dating, and (3) the pollen data blow or wash in from the surrounding vegetation. In these sediments, pollen grains are morphologically distinct and numerous, and plant macrofossils are often found. (Pollen and plant macrofossil data are also available from deposits that are discontinuous in time but can be organized into time series; Baker *et al.*, 2002; Betancourt *et al.*, 1990). Once analyzed, the data exist as point estimates of the abundance for each taxon in time series of samples at a site (or at nearby sites and middens for some discontinuous records). These time series can be expanded to transects of time series in latitude–, longitude–, or elevation–time diagrams, to networks of samples from an area for mapping, or to networks of time series, which are also time series of maps, to form a space–time box for displaying the data (**Figure 1**). Each of these displays illustrates different views in different dimensions of the space–time variations in the data. A zoom lens perspective can then be achieved in space, time, or both simultaneously by moving from data at a single site for a short time interval to progressively longer time series for either that site or for data on maps from increasingly larger areas. In this way, records of local succession can be seen in the broader context of long-term migrations and global climate changes.

Palynologists control the ability of their data to display selected patterns in vegetation and biodiversity (both in time and in space) by making choices about the numerical aspects of their data and by choosing the temporal, spatial, and taxonomic sampling characteristics of their data. These sampling characteristics have three elements – breadth of coverage, resolution of individual samples, and sampling density – which for a photograph are comparable to its frame of reference, grain size, and number of grains (or pixels) that are exposed, respectively. The temporal and spatial breadth of the data is the total time or area covered by a data set, and the taxonomic breadth is the total set of taxonomic groups (e.g., seed plants) included. The temporal and spatial resolution of each sample is defined in terms of how much time or area is represented in each individual sample within the data set, and taxonomic resolution depends on whether the data are lumped into groups or listed at their finest level of morphological distinction. The uncertainties of radiocarbon dates and age estimates also affect temporal resolution in correlations among sites. The number of samples in time and space defines the temporal and spatial sampling density. The total number of taxa listed in a data set defines the taxonomic sampling density. Choices about these characteristics can influence what types of changes in vegetation and biodiversity appear in paleoecological times series and map sequences (**Figures 1–3**).

Taxonomic and Numerical Characteristics

In most studies of lakes and bog sediments, the potential taxonomic breadth is Tracheophytes and Bryophytes along

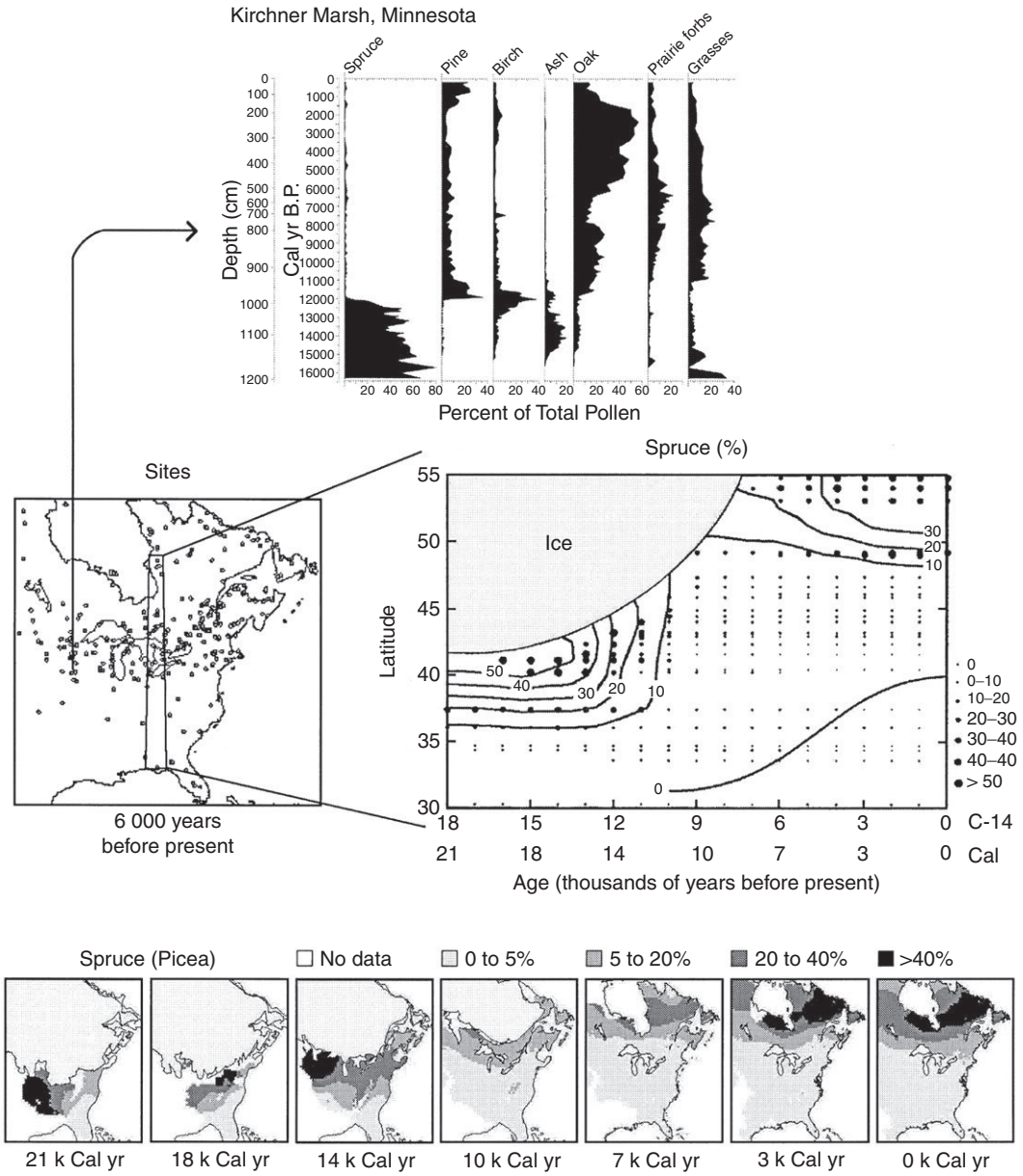


Figure 1 Pollen data in time and space. Different views from time series for multiple taxa at a single site to a single taxon (spruce) displayed either for a transect of sites in a latitude–time diagram or for a network of sites in the time series of maps from 21 k Cal yr (thousands of calendar years ago) to present. The shading on the maps indicates pollen abundances of 1–5% (light gray) and >5% (black). The white area at the top of the maps for 21, 18, 14, and 10 ka is the Laurentide ice sheet. When stacked vertically, these maps form a space–time box in which the three-dimensional contours can show the four-dimensional distribution of spruce pollen percentages in space–time.

with certain algal remains (*Pediastrum*) and fungal spores. When nonwetland vegetation is a primary focus, spores, algal remains, and pollen and plant macrofossils from aquatic plants are excluded from detailed study. Careful study and morphological characteristics of pollen grains determine the taxonomic resolution in samples and permit identification of genera and some species, although some grains can only be determined to the family level (e.g., Poaceae and Cyperaceae). Regional restriction of a single species, otherwise morphologically identifiable pollen only at the genus level (e.g., *Fagus grandifolia* within the northeastern US), can allow designation of plant species from pollen, but plant macrofossils, if present, allow species to be identified directly.

Palynology is a direct beneficiary of the ubiquity and inefficiency of wind pollination. Millions of grains are produced for every successful pollination, and some of the abundant residual reaches the sediments of lakes and bogs where the pollen grains are well preserved. Entomophilous grains are abundant in honey and on the legs of bees, but too few insects perish in lakes and bogs to leave a record that can be discerned among the overwhelming numbers of anemophilous grains. From the point of view of the remote-sensing metaphor,

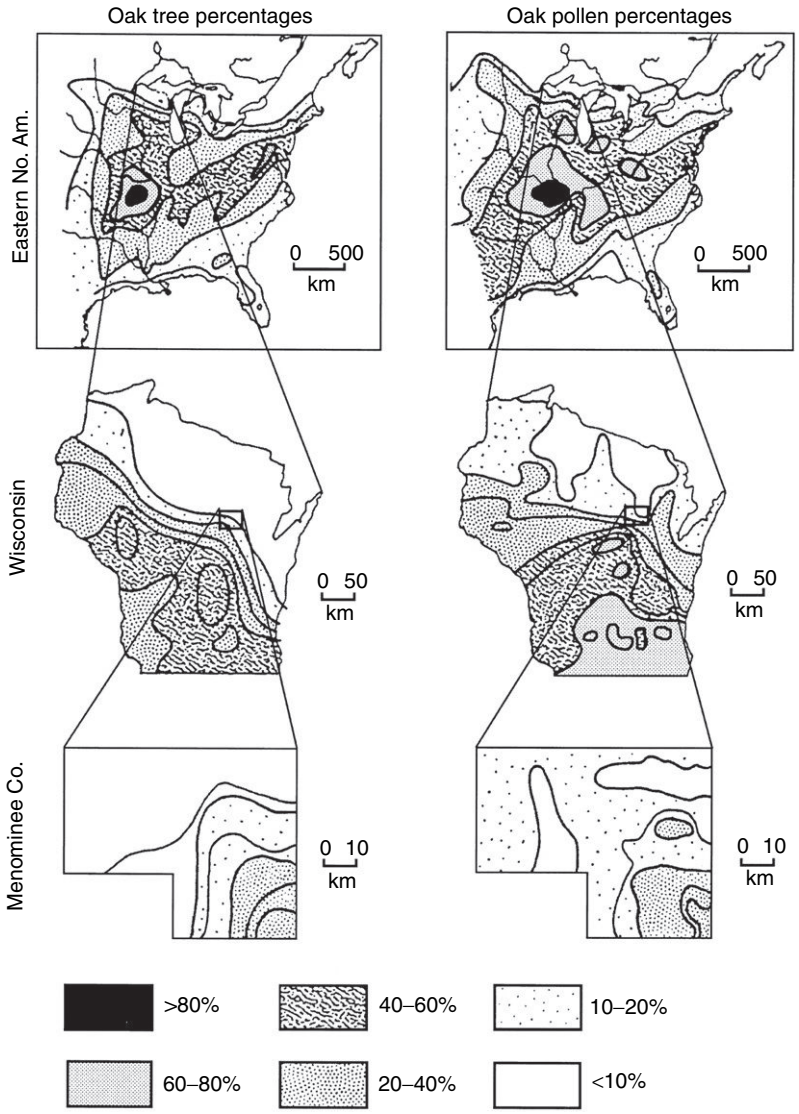


Figure 2 Zoom lens view of the mapped distributions of oak tree and pollen percentages at the scales of a subcontinent (107 km^2), state (105 km^2), and county (103 km^2). These maps show how well pollen percentages can reflect the distribution of tree percentages. Data used in contouring at the subcontinental scale were also used in contouring at the state scale, but a different data set was used at the county scale. Modified from Solomon and Webb (1985).

wind-pollinated plants are bright lights on the landscape and little or no signal comes from the other plants. As a result of this bias, the focus is on using pollen data to record the vegetation rather than species lists. For temperate to boreal forests in which the diversity of wind-pollinated trees is greatest globally, pollen records provide a fairly direct representation of the vegetation; however, in tropical regions and deserts, the pollen records yield a much more indirect representation of the vegetation because so many species are insect-pollinated.

In arid lands of southwestern North America, where lakes are rare, packrats create middens with embedded plant macrofossils that can be radiocarbon dated and identified to the species level for those plants collected (Betancourt, 2004). These records provide valuable information about arid vegetation and environments including studies of Holocene plant migration and biogeography (Lyford *et al.*, 2004; Gray *et al.*, 2006).

In most studies of Quaternary pollen, taxon percentages are used. Three hundred to five hundred grains are typically counted in a sample, and the counts for each taxon are divided by the total count. Because counts for individual taxa can be thought of as being binomially distributed (i.e., each grain is either taxon x or not taxon x), the percentages for multiple taxa are multinomially distributed, thus permitting direct calculation of confidence intervals. Pollen concentrations (grains cm^{-3} or grains g^{-1}) are seldom used because of their dependence on sedimentation rates. When radiocarbon or stratigraphic dates allow estimation of sedimentation rates, pollen accumulation rates (grains cm^{-2} $year^{-1}$) can be calculated, and procedures exist for calculating and potentially minimizing the confidence intervals for both pollen concentrations and accumulation rates.

Pollen accumulation rates, first introduced by Margaret Davis, have proven valuable for checking the ambiguity of

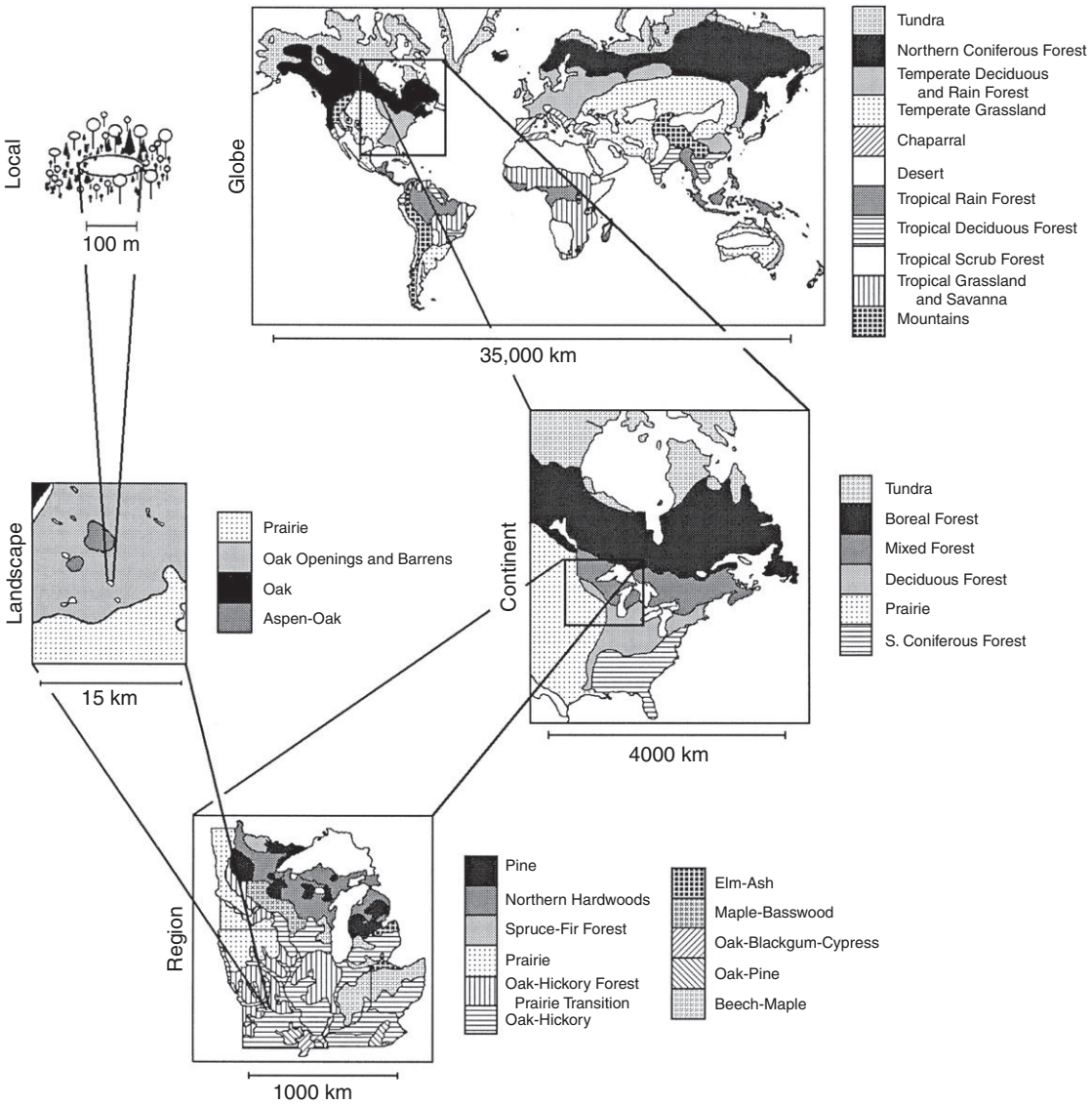


Figure 3 Zoom lens view of the vegetation from the scale of single site (where pollen accumulates) up to the global scale. No one view or interpretation of the vegetation is possible because the composition and gradients vary with scale and reflect climate gradients at the global and continental scales, a mixture of soil and climate gradients at the regional scale, and difference in soils and disturbance history at the landscape and local scales. Modified from Kutzbach and Webb (1991).

certain changes in pollen percentages. However, pollen accumulation rates are unsuitable when large sets of comparably sampled pollen data are required for mapping. These rates are highly sensitive to sedimentation differences within and between lakes as shown by Pennington and Davis and thus vary dramatically for nonvegetational reasons. Percentages of pollen taxa, which serve to cancel out these variations, are therefore best used in mapping studies (**Figures 1–3**). Many studies show that the percentages for pollen taxa represent well the relative abundance of plants on the landscape. The relationships vary with spatial scale and pollen type, and the uncertainties of these relationships add to the number of uncertainties for the pollen data when estimates of the vegetation are attempted.

Finding reliable quantitative measures for plant macrofossils is much more difficult than for pollen data. Most data are recorded in presence/absence terms or categorically, but sometimes percentages or concentrations are calculated. Local biases can make the data difficult to interpret because, for example, 1000 needles may all come from the same nearby tree.

Spatial Characteristics

The breadth of coverage for sets of pollen data can be broad or fine and thus match the different scales for mapping the vegetation (**Figure 3**). Continental and global data sets exist (Prentice *et al.*, 2000) along with those at regional (1000–300,000 km^2) and local (10^{-3} km^2) scales, and an embedded series of these data sets can provide a zoom lens view of past vegetation change if the sample resolution and sampling density are appropriate.

The potential spatial resolution of individual pollen samples varies among pollen taxa depending on the dispersibility of their pollen. Pollen from taxa with large grains is dispersed less far than that with small grains. Individual pollen samples are therefore variable-area samples, with the area of vegetation contributing pollen depending on which pollen type is studied; the percentages for pine pollen, for example, represent a larger area than those for beech or maple. Such differences in sampling area among pollen types can be important for studies of local or even regional vegetation but are less important for networks of data in which the distance between samples is much larger than the average dispersal distances for the different taxa.

The potential spatial resolution also depends on the basin characteristics. Key factors include the presence or absence of a canopy over the basin, the size of the water or wetland surface accumulating airborne pollen, the presence or absence of horizontal mixing of sediments within the basin, and the watershed area that supplies additional waterborne pollen to lakes. Samples from within a forest canopy accumulate most of their pollen from plants growing with 10–100 m of the coring site, and data from samples that preserve pollen are comparable spatially to vegetation data from large permanent plots in forests. In contrast, samples from lakes accumulate their pollen from distances of 100 m to 30 km. The within-canopy data have enough spatial resolution to illustrate succession within forests (**Figure 4**), whereas pollen data from lakes integrate across the mosaic of vegetation on landscapes and are sensitive to succession only if it is occurring across much of the pollen-source area. Bogs, mires, and shallow wetlands accumulate pollen at both regional (1 to >10 km) and local (10 m) scales because many plants grow on these wetlands. In samples within mires, pollen from local taxa can vary systematically and abruptly, whereas the pollen from regional taxa remains reasonably uniform. (This behavior of pollen percentages for regional taxa led von Post to develop pollen analysis as a method in 1916.) The spatial resolution of pollen samples, therefore, varies with pollen type within certain basin types, but selection of samples by pollen type and basin type can help keep the spatial resolution reasonably uniform. Such choices are key to seeing well with the observations afforded by sets of pollen samples.

For plant macrofossils, Dunwiddie and Jackson have shown that needles and seeds in lake sediments have a spatial sampling resolution of 10–20 m, which is much finer than the distance sampled by most pollen in lakes or open wetlands. Plant macrofossils can therefore help resolve local changes at or near a site and fine-scale elevation differences (Jackson and Whitehead, 1991). Sampling schemes can use a mixture of open and canopy-covered basins, mire and lake sediments, local and regional pollen types, and plant macrofossils in order to represent different scales of spatial pattern in the vegetation. Such mixtures of data sets can show vegetation differences that reflect specific soil and elevation differences within a broader geographic data set. Mapping of plant macrofossils can also show the range of species and of haplotypes and add precision to what maps of pollen data show (Jackson *et al.*, 1997; Binney *et al.*, 2009; Magri *et al.*, 2006).

The geographical sampling density for pollen data varies. To illustrate the variations in recent pollen within a basin or a forest, some studies include samples at 10 m or finer intervals along 500 m or longer transects through a basin or forest. For selected taxa, these studies show high sensitivity in the pollen data to local variations in vegetation cover and biodiversity. The variations are also evident in sets of time series from closely spaced cores in a basin (Simmons, 1993). For studies whose breadth of coverage is regions to continents (**Figure 3**), sets of modern data exist with average densities of 1/14 km^2 in 1000 km^2, 1/500 km^2 in 100,000 km^2, and 1/5000 km^2 in 107 km^2. These densities are sufficient for the contour patterns of oak pollen percentages to match the corresponding patterns of oak tree percentages at each spatial scale (**Figure 2**). Sets of fossil data provide less dense geographic coverage and vary in density from 1/50 km^2 for the Adirondacks to 1/6000 km^2 for the northern Midwest and 1/40,000 km^2 for eastern North America and Europe (Gaudreau *et al.*, 1989; Williams *et al.*, 2004).

These data sets have been used to illustrate how well pollen data represent or remotely sense spatial vegetational features such as range boundaries, ecotones, and abundance gradients (**Figures 2** and **3**) (Williams *et al.*, 2004; http://www.ncdc.noaa.gov/paleo/pollen/viewer/webviewer.html). Detecting the range boundary with pollen can be difficult because of pollen transport and relatively high counting uncertainties at low pollen percentages, but Davis *et al.* (1991) used a data set from the Midwest with a sampling density of one sample per 1000 km^2 to show that the past species limit for beech and hemlock can be identified within an area 20 km wide. For widely dispersed taxa such as oak and pine, such fine-scale resolution may not be possible, but the abundance gradients for these taxa can be used to locate ecotones at scales of 10–100 km (**Figure 2**).

Temporal Characteristics

The time range for records from individual cores can vary from 50 years to millions of years ago (**Figure 4**), with the bulk of the late Quaternary cores covering 14,000 years and fewer extending back 21,000 years, the last glacial maximum. Temporal density generally correlates with time range because short cores are often taken to study short-term changes and may contain 10–20 samples for 500 years, which yields one sample per 25–50 years, whereas 200 samples in a 2-million-year (My) record can only yield one sample per 20,000 years. Twenty to 40 samples for 10,000 years yield one sample per 250–500 years, which is typical for most cores in the data sets from eastern North American and European (Huntley and Webb, 1988). At a high level of temporal density are records from millimeter-scale sampling of peats that have yielded 1- to 5-year sampling for 50–200 years for selected times in the past, as reviewed by Turner and Peglar (1988) and Simmons (1993).

One of the longest continuous records for the Quaternary is from Sabana de Bogota in Colombia and covers 1.5 My (**Figure 4**). This study is remarkable in having relatively high sampling density of one sample per 1000 years and shows the appearance from North America of Alnus and Quercus at 1.3

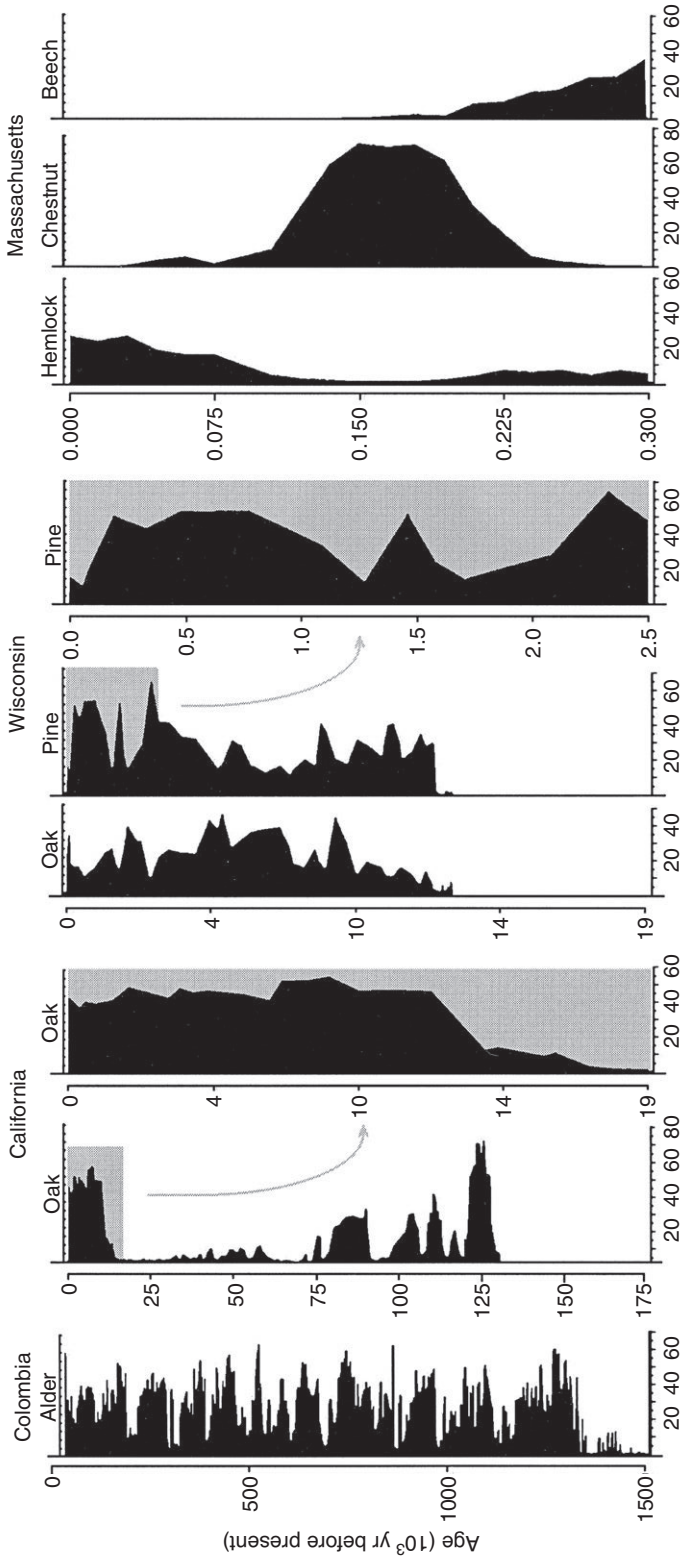


Figure 4 Zoom lens view of pollen data in time from a 1.5-My record down to a 300-year record. Different ecological processes are evident on different timescales. Temporal coverage, resolution, and density all vary appropriately for the Colombian lake record of alder pollen to show (a) alder's arrival at the site long after North and South America joined and (b) its response to orbitally driven climate changes (Hooghiemstra and Ran, 1994). The California lake record of oak pollen shows its response to orbitally driven climate variations over 130,000 years and the transition from glacial to interglacial climate over 16,000 years. In Wisconsin, pollen data from a lake that is free from bottom-organism mixing are sufficiently high in resolution to record how oak and pine populations responded to climate and regional disturbance events (e.g., fire) over 11,000 and 2500 years, respectively. Local succession is evident in the pollen data from a 300-year record from a small hollow within the forest canopy in Massachusetts (Foster and Zebryk, 1993). Gray indicates a portion of a time series that is expanded in the graph on the right.

million and 250,000 years ago, respectively. Other long cores with records of 50,000 years or longer exist on all continents but Antarctica; however, only in Europe are these long cores numerous enough to allow some mapping of vegetation from previous glacial and interglacial periods.

At the other extreme in terms of temporal coverage are (1) short cores of the upper sediments of lakes or bogs and (2) cores from forest soils that preserve pollen. The former provide records of human impact on the vegetation and of recent changes, whereas the latter show combined fine-scale sampling in both time and space and provide the view most closely approximating scale within which succession into forest gaps occurs. One of the ironies of paleoecological research is how seldom the data reveal succession as the dominant dynamic in the vegetation even though much ecological research has focused on succession.

These short-term records are often analyzed in high temporal density, for example, one sample per 50 years or even one per decade. The short cores (50 cm to 2 m) from lakes give coverage in high detail for 100–500 years or more in transects or in networks of relatively high spatial density (one sample per 1000 km^2) from states or regions (Russell *et al.*, 1993; Foster and Aber, 2006). Dating can be fairly accurate and precise because historical dates can be assigned to events in cores, and the date for the core top is usually known. High-resolution mapping in 50-year intervals is possible.

Temporal Resolution

A key metaphor for thinking about temporal resolution within individual samples and cores is to consider stratigraphic sections of lake sediments as strip charts along which pollen data have accumulated to form palygraphs. The sediment accumulation rate (which varies but averages 0.7 mm per year in most Holocene lakes) measures the drum speed (as on a thermograph) that drives the strip chart past the recorder, and the mixing of sediments between one and a few cm of the sediment surface after burial measures the degree to which the pen recording the data wiggles up and down vertically in time while registering horizontally the changing abundances of pollen. Within this sediment mixing zone, older sediments are moved both up from below (2–6 cm) and in from other parts of the basin by reworking. Such processes can make for tricky pen work on the palygraph. If the paper in the imagined palygraph were to absorb the ink slowly such that the scatter of tracings in the upper mixed zone is reduced to a single tracing of the abundances below this zone, then our metaphorical instrument would yield what is finally observed downcore. This image of the palygraph helps by indicating the separate components controlling temporal resolution in a core. The mixing depth and sedimentation rate largely control postdepositional time averaging, which regulates the time interval (i.e., temporal resolution) represented by pollen in a sample, whereas the temporal density of samples recorded and the accuracy and precision of assigning dates to each of many cores (strip charts) control analytical time averaging, which determines the temporal resolution among a set of cores used for mapping.

Dating Uncertainties

The dating uncertainty of an arbitrary pollen sample at some depth in a core depends on the uncertainties of the dating method, what is dated, the uncertainties of the age model, and variations in the sediment accumulation rate. The most accurate and precise dates are from lakes with annually laminated sediments, whose annual layers can be counted. These dates are in calendar years and accurate, with dating errors of only 1% or 2% at most. Unfortunately, lakes with such precision are too sparsely distributed for detailed mapping work. For lakes without varves, the dating near the tops of cores is more accurate and precise than that lower in the core because historical dates can be assigned to events in the core. These dates have uncertainties of 50 years or less and can be pinpointed in the core to 1 or 2 cm.

In studies of Late Quaternary data, most cores have been dated by conventional radiocarbon dates of bulk sediments. The bulk dates give the dates for the sediment matrix (i.e., strip chart material) and not for the pollen embedded within the sediments (i.e., what the pen records). The bulk dates have counting uncertainties of 20–400 years, with most approximately 100 years. These average errors give $\pm$200 years as the 95% confidence interval for each date. The dates can also be derived from 2 to 10 cm of sediment and therefore may average sediments covering 20–1000 years in age depending on the sediment accumulation rate. When dates for synchronous pollen events are averaged, such as the elm decline in Europe and the hemlock decline in eastern North America, standard deviations of *ca.* 300 years result. This uncertainty incorporates both the depth and the counting uncertainties. As Olsson (1986) observed, "There is … no point in determining the radiocarbon age with an uncertainty of $\pm$50 years if the relevant pollen-analytical level has an uncertainty of $\pm$200 years."

Olsson (1986) and Grimm *et al.* (2009) also note the problems with contamination, old carbon, and sampling depths. Old carbon can lead to systematic errors of 2000 or more years and requires correction. Dates on wood and accelerator mass spectrometer (AMS) dates on macrofossils in cores are helping to reduce these errors, and AMS radiocarbon dates are allowing depth intervals of 1 cm or less to be dated. Pilcher (1993) concludes that standard errors for radiocarbon dates from sediments are potentially $\pm$200 years but not more than $\pm$225 years. Recent work using many AMS radiocarbon dates on selected macrofossils and wiggle matching can lower this uncertainty however, and holds much promise for studying century-scale phenomena (see Hormes *et al.*, 2009).

Age Models and Mapping Intervals

Most pollen samples (30–50 per core) are not dated directly by a radiocarbon date (3–6 per core). Rather, their age is estimated by an age model based on linear interpretation, regression, or some other curve-fitting method. For cores with continuous sedimentation, the age model estimates the drum speed for the strip chart on the palygraph.

The uncertainties for all data plotted on the map for 12,000 years ago give an estimate for the amount of time before and after 12,000 years ago that the pollen data may cover on the

map. The uncertainties are likely to be between 300 and 500 years. This dating uncertainty for the data mapped matches the average temporal sampling density of one sample per 300–500 years in Late Quaternary pollen studies. Given such large uncertainties, maps to-date are best restricted to 1000-year intervals. Attempts to map at finer intervals will only produce maps whose data are not independently observed from the data in the maps for the next earlier and later dates.

As more studies get done at higher temporal resolution and with greater sampling interval, mapping can be attempted at finer time intervals. For sets of short cores in which historical dates yield a precision of 50 years or less, age models back to 500 years can yield dating sufficiently precise for maps in 100- to 50-year intervals. For regions over which stratigraphic markers such as volcanic ash represent synchronous layers in cores, some fine-scale mapping intervals are possible for the distant past, but even the dating of tephras has uncertainties of a century or more. Pollen-assemblage zones are sometimes used chronostratigraphically and can yield temporal resolution to 50 years or less among nearby cores, but synchroneity must be assumed for processes (e.g., abundances changes for individual taxa) that are ultimately time transgressive, especially at scales of 200 km or more. The geographic area for maps at a fine time resolution is therefore relatively small, but may be enlarged as more studies like that by Hormes *et al.* (2009) are completed.

Zoom Lens Perspective

Time series and mapped networks of pollen and plant macrofossils record Late Quaternary vegetation patterns at the scale of local vegetation, landscapes, regions, and continents (**Figures 1–4**). A hierarchy of nested data sets can reveal vegetation patterns at a variety of scales and provide information from general to detailed about the spatial vegetation patterns (**Figures 2** and **3**). Such hierarchies can link local changes to global events by recording each simultaneously (Kutzbach and Webb, 1991). A paleoecological zoom lens can then be used to examine the vegetation patterns that dominate different spatial scales for areas of $100–10^7$ km^2 with a spatial resolution of $10^{-2}–10^4$ km over timescales covering $10–10^4$ years (with time series to 10^6 years) with resolution from 1 to 500 years (**Figure 4**). From a global perspective over thousands of years with a relatively coarse grid of temporally and spatially averaged samples, investigators can zoom in ultimately to the locally small but statistically significant short-term changes in species abundances within one wetland or forest stand. A zoom lens perspective is possible for eastern North America and Europe and is becoming possible in western North America and other regions.

Two examples using current data illustrate this zoom lens view. The first holds time relatively constant and focuses up spatially from local to continental (**Figures 1–3**), and the second zooms in from a 1-My timescale down to successional changes during the past 300 years (**Figure 4**). The first example covers the past 21,000 years since the last glacial maximum in the northern Midwest and focuses up from a local to a continental perspective. It shows how the local changes in a wetland during the past 12,000 years in central Minnesota are part of regional climate changes that ultimately reflect the global climate changes during deglaciation and the current interglacial period.

At the local scale of the sampling basin, plant macrofossils, aquatic pollen, and diatoms at Kirchner Marsh in east-central Minnesota illustrate that 3000 years after the Laurentide ice sheet retreated from the sites *ca.* 19,000 years ago, a buried ice block melted to form the kettle basin and its associated lake (Kutzbach and Webb, 1991). Water levels were high until approximately 11,000 years ago when the seeds of damp ground and weedy annuals began appearing in the core (Watts and Winter, 1966). These appeared when the mixture of oak and herb pollen data show the first evidence of oak savanna conditions within the landscape and region near the site (**Figure 1**). When the herb pollen values increased sufficiently to indicate prairie vegetation near the site, large fluctuations in aquatic seeds indicated intermittent droughts that continued until a marsh developed at the site 2000 years ago. This history of local wetland development is closely linked to the vegetation and climate changes indicated at the landscape, regional, and continental levels (**Figures 1–3**).

Here, there is a shift from the time series view provided by local pollen and macrofossils to the continental view that the maps of pollen data provide. These show how 15,000 years ago spruce trees grew initially in a parkland that became populated with deciduous trees, first ash and then birch and oak, from 14,500 to 12,000 years ago as forests began to develop near the site for the first time (**Figure 1**). This early vegetation was unlike any growing today and illustrates the major compositional changes in vegetation that accompanied the major changes in climate (Williams and Jackson, 2007). The sudden decrease in spruce abundance near Kirchner Marsh 12,500 years ago reflects the general northward movement and decrease in regional and continental spruce populations as climate warmed and the ice sheet retreated. Pine populations were then invading from the east to replace the birch trees that briefly grew abundantly near the site and in the northern Midwest generally. Oak and elm populations then increased to replace the pines as the climate warmed. The regional addition of herb pollen after 11,000 years ago signals the development of savanna conditions as the climate dried and the ice sheet retreated further north. The savanna grew until prairie vegetation developed approximately 8000 years ago as the grasslands spread from the west across Minnesota and the northern Midwest. By then, the ice sheet had almost disappeared and spruce populations were beginning to grow abundantly in eastern and central Canada where the boreal forest has developed today. With the retreat westward of the prairie–forest ecotone after 6000 years ago, open oak forests began growing near the lake that with further growth in aquatic vegetation became a marsh after 2000 years ago. Oak populations then decreased regionally as part of a general retreat of oak in the north as conifer populations increased and spruce populations increased south from Canada into the northern Midwest and New England. The broad-scale regional and continental maps show how the local and landscape changes in vegetation fit within the broader context of changes.

The second example is a zoom lens perspective in time (**Figure 4**) and shows time series for taxa reflecting the long-

term temporal beat in climate at long time scales (1.5 My to 16,000 years) but then shifting to reflect fires, disease, and other disturbances at shorter timescales, especially when the pollen data are derived from within a forested hollow. Within the latter, the pollen reflects local changes in trees next to the site.

This arrangement and description of the data allow paleoecologists to zoom in or out in space–time to observe different aspects of vegetation dynamics and patterns of biodiversity from long-term changes on several spatial scales (from continental to inside the forest canopy) to competitive interactions after disturbance within communities. With such an ordering of data and images, the interconnected roles and impacts of climate change, disturbance, disease, succession, soil development, competition, and evolution can all be observed and potentially distinguished from one another in studies of past biodiversity. To this end, recent studies of sediment records are including data on charcoal, sediments, biomarkers, isotopes, biogeochemistry, and plankton that provide evidence for environmental changes from fires to climate that can be used to help interpret the changes in the vegetation (Whitlock *et al.*, 2003; Webb *et al.*, 2003).

Mapping and studying pollen and plant macrofossil data in space–time have been convincing of the potential for resolving many different views of the past. No single fixed picture is possible because there is no preferred viewing scale or perspective, nor is there any one way to describe the vegetation and biodiversity. Many conventional displays and perspectives are used, but some, such as traditional pollen diagrams, may not be optimal for interpreting key variations in the vegetation. A zoom lens perspective in space and time is required to help explore all the possibilities and to allow for diverse narratives and explanations. Succession and patch dynamics are evident with data scaled to meters and decades, whereas climatically induced migrations and abundance changes appear best over thousands of kilometers and years. To zoom out from a local to a global view, the paleoecological data need to be organized so that highly resolved images of species near a site give way to more generalized averages of data values whose patterns will stand out sharply when viewed across continents or the world. Appropriate study of assembled databases should allow ecologists to resolve and perceive a variety of phenomena, such as ecotones, disturbance horizons, species distributions, and migration patterns, that are important for biodiversity.

See also: Biodiversity, Evolution and. Fossil Record. Remote Sensing and Image Processing

References

Baker RG, Bettis III EA, Denniston RF, Gonzalez LA, Strickland LE, and Krieg JR (2002) Holocene paleoenvironments in southeastern Minnesota – Chasing the prairie–forest ecotone. *Palaeogeography Palaeoclimatology Palaeoecology* 177: 103–122.

Betancourt JL (2004) Arid lands paleobiogeography: The fossil rodent midden record in the Americas. In: Lomolino MV and Heaney LR (eds.) *Frontiers in Biogeography: New Directions in the Geography of Nature*, pp. 27–46. Sunderland, MA: Sinauer Associates Inc.

Betancourt JL, Van Devender TR, and Martin PS (eds.) (1990) *Packrat Middens: The Last 40,000 Years of Biotic Change*. Tucson: University of Arizona Press.

Binney HA, Willis KJ, Edwards ME, *et al.* (2009) The distribution of late Quaternary woody taxa in northern Eurasia: Evidence from a new macrofossil database. *Quaternary Science Reviews* 28: 2445–2464.

Davis MB, Schwartz MW, and Woods K (1991) Detecting a species limit from pollen in sediments. *Journal of Biogeography* 18: 653–668.

Foster DR and Aber JD (eds.) (2006) *Forests in Time: The Environmental Consequences of 1,000 Years of Change in New England*. New Haven, CT: Yale University Press.

Foster DR and Zebryk TM (1993) Long-term vegetation dynamic and disturbance history of a Tsuga-dominated forest in New England. *Ecology* 74: 982–998.

Gaudreau DC, Jackson ST, and Webb III T (1989) The use of pollen data to record vegetational patterns in regions of moderate to high relief. *Acta Botanica Nederlandica* 38: 369–390.

Gray ST, Betancourt JL, Jackson ST, and Eddy RG (2006) Role of multidecadal climatic variability in a range extension of pinyon pine. *Ecology* 87: 1124–1130.

Grimm EC, Maher Jr. LJ, and Nelson DM (2009) The magnitude of error in bulk-sediment radiocarbon dates from central North America. *Quaternary Research* 72: 301–308.

Hoogheimstra H and Ran ETH (1994) Late Pliocene–Pleistocene high resolution pollen sequence of Colombia: An overview of climatic change. *Quaternary International* 21: 63–80.

Hormes A, Blaauw M, Dahl SO, Nesje A, and Possnert G (2009) Radiocarbon wiggle-match dating of proglacial lake sediments – implications for the 8.2 ka event. *Quaternary Geochronology* 4: 267–277.

Huntley B and Birks HJB (1983) *An Atlas of Past and Present Pollen Maps for Europe: 0–13000 Years Ago*. Cambridge: Cambridge University Press.

Huntley B and Webb III T (eds.) (1988) *Vegetation History*. Dordrecht: Kluwer.

Jackson ST and Whitehead DR (1991) Holocene vegetation patterns in the Adirondack Mountains. *Ecology* 72: 641–653.

Jackson ST, Overpeck JT, Webb III T, Keattch S, and Anderson KH (1997) Mapped plant macrofossil and pollen records of late-Quaternary vegetation change in eastern North America. *Quaternary Science Review* 15: 1–71.

Kutzbach JE and Webb III T (1991) Late Quaternary climatic and vegetational change in eastern North America: Concepts, models, and dates. In: Shane LCK and Cushing EJ (eds.) *Quaternary Landscapes*, pp. 175–217. Minneapolis: University of Minnesota Press.

Lyford ME, Jackson ST, Gray ST, and Eddy RJ (2004) Validating the use of woodrat (Neotoma) middens for documenting natural invasions. *Journal of Biogeography* 31: 333–342.

Magri D, Vendramin GG, Comps B, *et al.* (2006) A new scenario for the Quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. *Journal of Biogeography* 171: 199–221.

Olsson IU (1986) Radiometric dating. In: Berglund BE (ed.) *Handbook of Holocene Palaeoecology and Palaeohydrology*, pp. 110–120. Chichester, UK: Wiley.

Paleobiology Database. http://paleodb.org/cgi-bin/bridge.pl

Peglar SM, Fritz SC, Alapieti A, Saarnisto M, and Birks HJB (1984) Composition and formation of laminated sediments in Diss Mere, Norfolk, England. *Boreas* 13: 13–28.

Pilcher JR (1993) Radiocarbon dating and the palynologist: A realistic approach to precision and accuracy. In: Chambers FM (ed.) *Climate Change and Human Impact on the Landscape*, pp. 23–32. London: Chapman & Hall.

Prentice IC, Jolly D and BIOME 6000 participants (2000) Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *Journal of Biography 27*: 507–519. http://www.neotomads.org

Russell E, Davis RB, Anderson RS, Rhodes TE, and Anderson DS (1993) Recent centuries of vegetational change in the glaciated north-eastern United States. *Journal of Ecology* 81: 647–664.

Shuman BN, Bartlein PJ, Logar N, Newby P, and Webb III T (2002) Parallel vegetation and climate responses to the Early-Holocene collapse of the Laurentide Ice Sheet. *Quaternary Science Reviews* 21: 1793–1805.

Shuman BN, Newby P, Huang Y, and Webb III T (2004) Evidence for the close climatic control of New England vegetation history. *Ecology* 85: 1297–1310.

Simmons IG (1993) Vegetation change during the Mesolithic in the British Isles: Some amplifications. In: Chambers FM (ed.) *Climate Change and Human Impact on the Landscape*, pp. 109–118. London: Chapman & Hall.

Turner J and Peglar SM (1988) Temporally precise studies of vegetation history. In: Huntley B and Webb III T (eds.) *Vegetation History*, pp. 753–777. Dordrecht: Kluwer.

Watts WA and Winter TC (1966) Plant macrofossils from Kirchner Marsh, Minnesota – a paleoecological study. *Bulletin Geological Society of America* 77: 1339–1360.

Webb III T, Shuman BN, and Williams JW (2003) Climatically-forced vegetation dynamics in North America during the late-Quaternary. In: Gillespie A, Porter SC, and Atwater BF (eds.) *The Quaternary Period in the United States*, pp. 459–478. New York: Elsevier.

Williams JW and Jackson ST (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* 5: 475–482.

Williams JW, Shuman BN, Webb III T, Bartlein PJ, and Leduc PL (2004) Late Quaternary vegetation dynamics in North America: Scaling from taxa to biomes. *Ecological Monographs* 74: 309–334.

Whitlock C, Shafer SL, and Marlon J (2003) The role of climate and vegetation change in shaping past and future fire regimes in the northwestern US and the implications for ecosystem management. *Forest Ecology and Management* 178: 5–21.

Parasitism

Klaus Rohde, University of New England, NSW, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adult parasite A parasite associated with a host during part or the whole of its mature phase.
Commensalism An association of animals in which one uses food supplied in the internal or the external environment of a host without affecting the host in any way.
Ectoparasite A parasite living on the surface of a host.
Endoparasite A parasite living inside a host.
Facultative parasite A parasite that can also live without a host.
Final (= definitive) host A host that harbors sexually mature stages of a parasite.
Hyperparasite (of first, second, etc. degree) A parasite living on or in another parasite.
Intermediate host A host that harbors sexually immature, developing stages of a parasite.
Intraspecific parasitism A parasitic association of members of the same species.
Latent parasitism Parasitism without obvious symptoms.
Mutualism An association of organisms in which both partners benefit from the association.
Obligatory parasite A parasite that cannot survive without a host.
Periodic parasite A parasite visiting a host at intervals.
Permanent parasite A parasite associated with a host for long periods.
Phoresis An association in which one organism uses another as a means of transport and protection.
Symbiosis (*sensu lato*) Any association between organisms (parasitism, commensalism, mutualism, phoresis).
Symbiosis (*sensu strictu*) An association of organisms in which both partners benefit from the association and cannot live without each other.
Transport host A host that harbors sexually immature stages of a parasite that do not develop.

Parasitism is defined in different ways by different authors, usually reflecting their research interest and bias. Parasitism, as used here, is defined as a close association between two organisms in which one, the parasite, depends on the other, the host, deriving some benefit (usually food) from it without necessarily damaging it. Traditionally, fungi, bacteria, and viruses, many of which are parasitic, are studied by microbiologists, whereas parasitologists study protozoan and metazoan parasites. In this contribution, only protozoan and metazoan parasites as well as higher plants (angiosperms) are included. Several types of associations resemble parasitism in various ways and cannot always be clearly distinguished from it, either because of insufficient knowledge or because genuine intermediate forms exist. Such associations are discussed in the following section.

Related Phenomena

Commensalism

A commensal is an organism that uses food supplied in the internal or the external environment of the host, without establishing a close association with the host, for instance by feeding on its tissues. Examples are the ameba *Entamoeba coli*, an endocommensal of humans feeding on bacteria in the lumen of the intestine, and the ciliate protozoan *Ephelota gemmipara*, an ectocommensal on various marine invertebrates.

Phoresis (Phoresy)

In a phoretic association, one organism uses another as a means of transport and protection. An example is barnacles living on whales.

Mutualism

A mutualistic association is one in which both organisms derive a benefit, but the association is not compulsory. The cleaner fish *Labroides dimidiatus* feeds on parasites and diseased tissues of various marine fishes, and both partners derive a benefit, the cleaner obtaining food and the host fish getting rid of their parasites and diseased tissues.

Symbiosis

Symbionts live in a compulsory association in which both partners derive a benefit. An example is the symbiosis of fungi and algae in lichens. It should be noted, however, that the term symbiosis is sometimes used in a wider sense, including all types of associations between organisms (parasitism, commensalism, phoresis, and mutualism).

Predation

A predator is an organism that attacks another, the prey, and usually kills and eats it. Most predators are larger than their prey.

An organism may be a parasite under certain conditions, but a commensal, a mutualist, or a predator when conditions change. For example, *Entamoeba histolytica* is often a harmless commensal feeding on bacteria in the intestine of humans, but may become a dangerous parasite feeding on red blood cells, apparently induced by some changes in the host that are not fully understood. A number of normally pathogenic parasites even improve the health and fitness of their hosts at low infection intensities. Thus, Lincicome (1971, cited in Rohde, 1993) conducted some controlled experiments using rats and mice infected with two species of trypanosomes and the

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00103-9

nematode *Trichinella spiralis* that showed that parasitized animals grew faster, ate more, could compensate better for deficiencies in the diet, had livers richer in certain vitamins, and were more active and more responsive to the human presence. For these reasons, it is best to consider parasitism as a type of association that is not clearly delimited from the others.

Types of Parasites

There are many types of parasites, distinguished by the site of infection, kinds of hosts, state of development, etc. Ectoparasites are parasites that live on the external surface of hosts, for example, fleas and lice of various terrestrial vertebrates, and Monogenea and Copepoda of freshwater and marine fishes. Endoparasites are parasites that live in the tissues and organs of their hosts, such as tapeworms, flukes, and protozoans of vertebrates. An obligatory parasite is a parasite that cannot survive without a host, such as the malaria parasite, and a facultative parasite is a parasite that can also live without a host, such as maggots, which are normally saprophagous but can infect living hosts as well. A permanent parasite is associated with a host for long periods, whereas a temporary parasite is found in or on a host only for short periods: examples of the former are human helminths and blood protozoans and examples of the latter are mosquitoes and leeches that visit hosts for blood-sucking only for short periods. Larval parasites, like the praniza larva of isopods, are parasitic only at a larval stage, and adult parasites, to which most metazoan parasites belong, are associated with a host during part or the whole of their mature phase. Periodic parasites (leeches, mosquitoes) visit a host at intervals. Intraspecific parasites parasitize individuals of the same species; for example, males of certain deep-sea fish are permanently attached to females of the species and derive food from them. Hyperparasites (of the first, second, etc. degrees) are parasites living on or in other parasites. An example of a hyperparasite of the first degree is Udonella, a monogenean parasitic on copepods, which themselves parasitize marine fishes. Microparasites, i.e., protozoans, bacteria, viruses, and some helminths, are small, with short generation times; they reproduce on or in a host usually at high rates, the duration of infection is usually much shorter than the life span of the host, and they induce immune responses in vertebrates. Macroparasites, i.e., most helminths and arthropods, do not reproduce on or in the host, they have longer generation times than microparasites, immune responses are lacking or weak and depend on infection intensities, and infections are often chronic and lead to morbidity rather than mortality. Parasitoids, many species of Hymenoptera, lay their eggs into insect hosts that may survive for some time but are invariably killed by the growing larvae of the parasitoids. Hence, parasitoids are predators rather than genuine parasites.

Adaptations to Parasitism

Size of Parasites

Most parasite species are much smaller than their hosts. Malaria parasites, for example, are microscopic protozoans, and human pinworms are less than 1 cm long. Nevertheless, some species reach a remarkable size. For example, a didymozoid trematode infecting the sunfish, *Mola mola*, reaches a length of 12 m, although its diameter is very small and the volume of the parasite is still much smaller than that of the fish, which reaches a weight of 1 ton. Likewise, the broad fish tapeworm, *Diphyllobothrium latum*, which lives in the intestine of various fish-eating mammals, including humans, reaches a length of greater than 10 m, but its volume is much smaller than that of its hosts. At first glance, perhaps surprisingly, parasites are often considerably larger than their free-living relatives. This phenomenon is clearly shown in flatworms, phylum Platyhelminthes. Most free-living flatworms are very small, from less than 1 mm to a few millimeters long, whereas parasitic flatworms such as flukes and tapeworms, as a rule, are much larger, up to several centimeters in length for the flukes and many meters long for the tapeworms. There may be two reasons for this. First, parasitic flatworms in the organs and tissues of their hosts have a much richer and consistent food supply than do free-living species, and food supply therefore does not present a limit to size. Second, selection may have favored multiple and larger gonads and, therefore, a large body size because parasites have to produce many offspring in order to overcome the hazards involved in infecting other hosts. But selection may also favor a smaller body size, because parasites depend on living hosts and it is important that hosts survive at least until the parasite has produced offspring that can infect other hosts. A smaller size of parasites relative to that of their hosts is therefore an advantage.

Reduction and Increase in Complexity

A general misconception is that all parasites have a less complex structure than free-living forms, a phenomenon that has been named sacculinization after the parasitic barnacle Sacculina, which infects marine crabs and which indeed shows a remarkable reduction in complexity: the only parts of the parasite visible on the outside of the host are the so-called externa, a sac-like structure containing the gonads. Most of the parasite consists of an extensive system of cytoplasmic processes reaching into the various host tissues (**Figure 1**). All the crustacean characteristics have been lost, and only the free-living larval stage indicates that the parasite is indeed a crustacean, related to barnacles. However, in most parasites, such a reduction in complexity has not occurred, and in many species there is, in fact, an increase in complexity. A well-studied example of the latter case is the trematode *Lobatostoma manteri*, a parasite of marine fishes, about 3–5 mm long, which not only has a remarkably complex nervous system with more longitudinal nerve cords than free-living flat-worms but also about 20,000–40,000 sensory receptors, belonging to about a dozen different types distinguishable under the electron microscope (**Figure 2**). These numbers are much larger than those of most related free-living flatworms, and this in spite of the fact that there is not a single free-living stage in the life cycle of this parasite. The adult worm lives in the small intestine of marine fish, it produces eggs containing an infective larva, the egg is eaten by a snail in which the larva hatches and develops to a stage infective to fish, and fish become infected by eating snails.

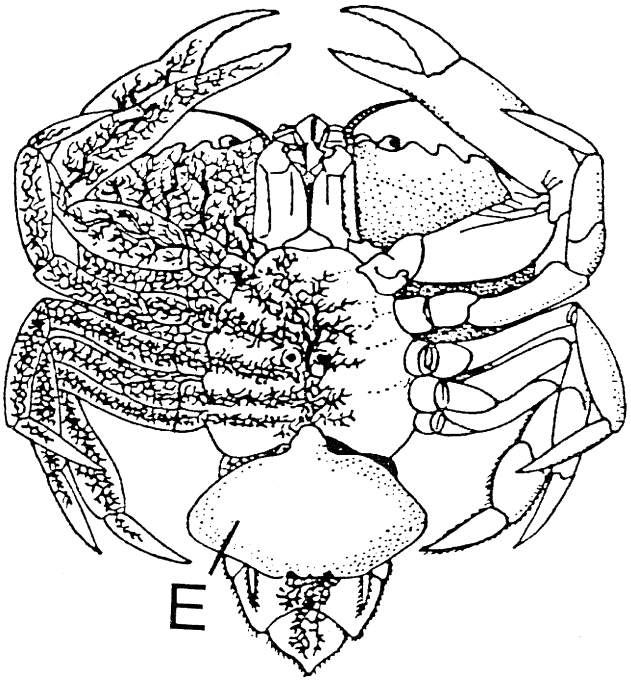


Figure 1 The rhizocephalan *Sacculina carcini* parasitic on and in the crab *Carcinus maenas*. E, externa of the parasite containing the gonads. Processes of parasite in crab drawn only on one side of the body. After Boas, reproduced from Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

Increase in Reproductive Capacity

Almost all parasites that have been studied produce a remarkable number of offspring, greater than that of related free-living forms. An example is given in **Table 1**; free-living turbellarians produce the fewest offspring of any flatworms, ectoparasitic Monogenea produce more, and the highest number of offspring is produced by endoparasitic trematodes and cestodes. Most trematodes and some tapeworms not only produce large numbers of eggs, but also there is a secondary increase in the numbers of offspring in the intermediate hosts by asexual or parthenogenetic reproduction. Thus, a single egg of a trematode can produce thousands of cercariae, i.e., larval stages infective to the final, vertebrate host. A prerequisite for this increased reproductive capacity is the safe and rich food supply available to parasites as well as the relatively large body size of many parasites compared with their free-living relatives (*see* Size of Parasites). However, it is likely that selection has favored an increased reproductive output, because the hazards encountered by parasites in their often complex life cycles are enormous. Very few larvae will survive and establish an infection.

Mechanisms of Dispersal

Dispersal is important for any species, whether free-living or parasitic, because a population restricted to one small area risks becoming extinct if conditions become unfavorable and because dispersal reduces inbreeding and the loss of evolutionary adaptability. For parasites, a third point is important: dispersal may reduce the chances of hosts becoming overinfected. Three aspects of dispersal are important: dispersal over short distances away from an individual host, dispersal in space and range extension over larger distances, and dispersal in time. Trematode larvae illustrate that all three aspects of dispersal can be brought about by the same stage. Larvae (cercariae) are often forcibly ejected into the respiratory currents of the snails in which they have developed, bringing about dispersal away from the host. They actively swim and keep afloat by means of their tails and can thus be dispersed over long distances by water currents. In many species, special flotation devices of the tail prolong the duration of floating (**Figure 3**). Adult flukes produce eggs, and larvae in the snail hosts are produced over long periods, months, or even many years, leading to dispersal in time.

Mechanisms of Infection

It is essential for a parasite to ensure entry into a host, and this is achieved by an amazing variety of mechanisms. **Table 2** lists examples of the infection mechanisms used by human parasites. Many parasite species have evolved remarkable behavioral adaptations that facilitate transmission to a host. For example, microfilariae, i.e., larvae of various species of filariae (nematodes), circulate in the peripheral blood of vertebrates, where they are ingested by mosquitoes for further development. Depending on the activities of the mosquito species involved, the microfilariae appear in the peripheral blood either during the day or the night, and strains of the same species may be nocturnal in one area and diurnal in another. Larvae of some monogeneans have endogenous hatching rhythms (rhythms not induced by external factors) adapted to the hosts' behavior. Some species (and strains) hatch in the

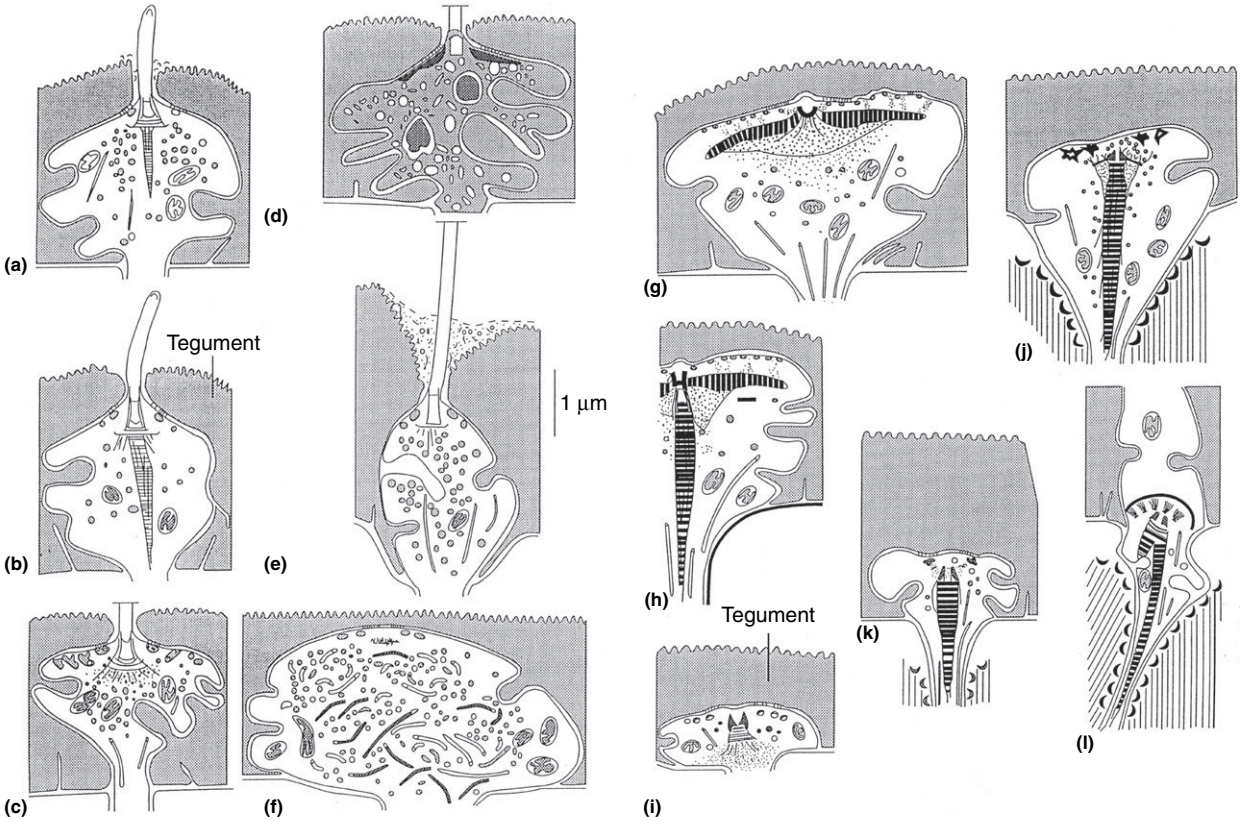


Figure 2 Sensory receptors of adult *Lobatostoma manteri* (Trematoda, Aspidogastrea). After Boas, redrawn from Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

Table 1 Approximate estimates of fecundity of free-living and parasitic flatworms

	Number of eggs	*Multiplication of larvae*
Free-living Turbellaria	10	None
Ectoparasitic Monogenea	1000	None
Endoparasitic trematodes	10 million	× 1000 at least
Endoparasitic tapeworms	10 million	× 1–1000

evening and others in the morning, whenever chances to infect a host fish are greatest. Perhaps most remarkable are the adaptations of some trematodes. For example, the liver fluke of sheep, *Dicrocoelium lanceolatum*, uses terrestrial snails as the first and ants as the second intermediate hosts. Cercariae produced in the snails aggregate in slime balls inside the snails, are expelled by the snails, and are eaten by ants. The first cercariae that enter an ant migrate into the subesophageal ganglion, inducing spastic behavior of the ant, which makes it cling to a plant, where their chances of being eaten by sheep are enhanced. The trematode *Leucochloridium macrostomum* has a larval stage, the sporocyst, which forms colorful outgrowths that extend into the snail's tentacles. They pulsate rhythmically, mimicking worms, the natural food of small birds, which are the final hosts. Birds bite off the tentacles and become infected. Infection also induces snails to move to more exposed sites.

Aggregation

The distribution of organisms in space may be even (more or less equal distances between individuals), random (random distances between individuals), or aggregated (or clustered, or overdispersed; individuals occur in clusters). Applied to populations of parasites, as a rule, parasites show an aggregated distribution in host populations: some individuals are heavily infected, but most are very lightly infected or not at all. Very often, such distributions can be described best by a negative binomial distribution. The reasons for aggregated distributions are manifold: a series of exposures, each with different chances of infection; a nonrandom distribution of infective stages; an increase or decrease in chances for further infections by a first infection; variations in the susceptibility of host individuals; and changes of infection of individual hosts over time. Selection may even have favored aggregation in order to restrict damage to a few heavily infected individuals or to facilitate mating, but experimental evidence does not exist.

Hermaphroditism, Parthenogenesis, and Asexual Reproduction

Contact between parasites and hosts is usually sporadic and, in most cases, only a single or a few parasites will manage to infect a host. It is important that populations can be built up from these individuals. A single individual can produce

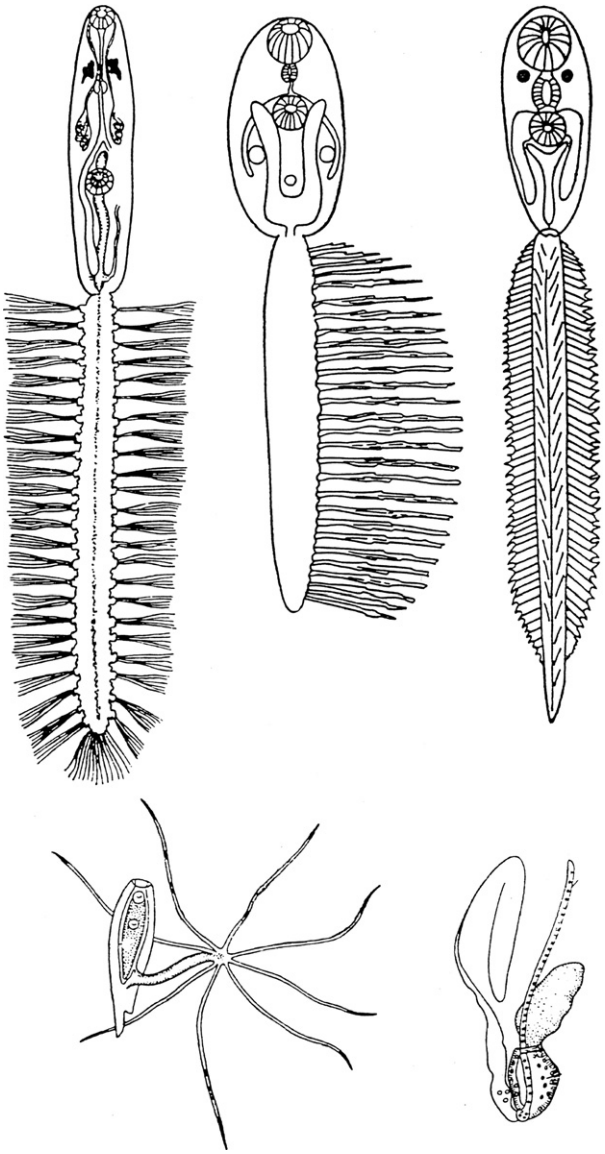


Figure 3 Marine cercariae with various flotation mechanisms on the tail. After various authors, reproduced from Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

large populations by parthenogenetic or asexual reproduction. In the first case, gametes develop without fertilization; in the second, somatic cells develop. An example of the second case is malaria parasites developing by schizogony in human red blood cells. An example for the first case is (probably) trematode larvae developing in snails. Almost all trematodes, as well as the cestodes and monogeneans, are hermaphroditic, thereby doubling the chance of meeting a mating partner. Furthermore, some species can at least sometimes self-fertilize.

Host Specificity

All parasites are restricted to certain host species, i.e., they cannot infect all species available, although some parasite species are more restricted than others. For example, the human pinworm, *Enterobius vermicularis*, is found only in humans, whereas the protozoan *Toxoplasma gondii* has been found in a wide range of mammals and birds.

Some parasites, although found in many hosts, nevertheless infect a single or a few host species much more strongly than others. It is therefore useful to distinguish host range (the number of host species infected irrespective of how frequently and how strongly they are infected) and host specificity (taking frequency, or prevalence, and intensity of infection into account). A parasite that infects many host species, but one of them much more strongly than the others, may have a greater host specificity than a parasite species that infects fewer host species, but all of them strongly. Indices to measure host specificity are available (Rohde, 1993).

Table 2 Infection mechanisms of some human parasites

Means of infection	*Parasite*
Inoculation by arthropod hosts	Malaria parasite (Plasmodium); sleeping sickness
Fecal contamination of wounds	Chagas disease (*Trypanosoma cruzi*)
Retrofection (through the anus into the large intestine)	Pinworm (*Enterobius vermicularis*)
Ingestion of cysts	Amebic dysentery (*Entamoeba histolytica*)
Ingestion of eggs	Roundworm (*Ascaris lumbricoides*)
Ingestion of spores	Microsporans
Ingestion of transport hosts	Anisakis
Ingestion of intermediate hosts	Broad fish tapeworm (*Diphyllobothrium latum*)
Inhalation	*Pneumocystis carinii* (likely, but no experimental evidence)
Contact transfer	Mange mite (*Sarcoptes scabiei*), lice
Kissing or joint use of eating or drinking utensils	*Trichomonas tenax*
Sexual intercourse	*Trichomonas vaginalis*
Penetration through the skin	Bilharzia (Schistosoma spp.); hookworm (Ancylostoma, Necator)
Penetration into the nasal passages	Primary amebic meningoencephalitis (*Naegleria fowleri*)
Intrauterine infection	Toxoplasma, malaria

Site Specificity

There is no universal parasite that infects all organs and tissues of a host equally, but the degree of site specificity within or on a host varies considerably. *Entamoeba histolytica* infects not only the intestine of humans but also the liver, lungs, brain, and other organs, whereas adult schistosomes are restricted to the blood vessels. Site specificity is extreme in many Monogenea on the gills of fishes. Different species occupy different parts of the gill on the same host (**Figure 4**).

Simple and Complex Life Cycles

Many parasites use a single host; i.e., they have a direct life cycle. Others use a final (definitive) host as well as one or several intermediate hosts; i.e., they have indirect life cycles. Parasites with direct life cycles include fleas and lice on various vertebrates as well as many intestinal nematodes and Monogenea on the skin and gills of fish. Usually, the adult stage is parasitic. It produces eggs or larvae that infect the same or other host individuals. In the case of some isopods, the only parasitic stage is the so-called praniza larva: adults live on the sea floor, larvae attach themselves to the gills of fish, suck blood, and drop off to mature in the benthic environment. Parasites with indirect life cycles include the trematodes, some with a single intermediate host and others with several. An example of the former is the aspidogastrean trematode *Lo. manteri*: the adult infects the intestine of marine fish; eggs are produced and shed in the feces; they are eaten by marine snails, in which the larvae hatch and grow to (almost) adult body size; and fish become infected by eating snails **Figure 5**). An example of a trematode with two intermediate hosts is the Chinese liver fluke, *Opisthorchis* (*Clonorchis*) *sinensis*. Adults infect the liver of fish-eating mammals, including humans; snails are the first and fish the second intermediate hosts. An example of a trematode with three intermediate hosts is the bird fluke *Strigea falconispalumbi*. Adults live in predatory birds, the first intermediate hosts are snails, the second intermediate hosts are tadpoles/frogs, and the third intermediate hosts are amphibians, snakes, mammals, and birds, which are eaten by the final hosts (**Figure 6**). The life cycles of some trematodes with four hosts can be extended by incorporating various transport hosts in which larvae accumulate but do not develop further. Nematodes with indirect life cycles include filariae, which are transmitted by blood-sucking dipterans. Some nematodes with direct life cycles include species whose larvae undergo a peculiar migration through the body of the host. The human roundworm *Ascaris lumbricoides*, for instance, produces eggs that are swallowed by the host. They hatch in the digestive tract but do not grow there to the adult stage; instead, they penetrate the wall of the intestine, invade the blood vessels, are carried into the lungs, penetrate through the alveoli, migrate up the trachea, are swallowed a second time, and now develop to the adult stage in the intestine. The reasons for this curious phenomenon are not known, but it may reflect a different kind of life cycle in the evolutionary past. Hookworms, which infect their hosts by penetrating through the skin, undergo a similar body migration involving passage through the lung alveoli and final maturation in the intestine.

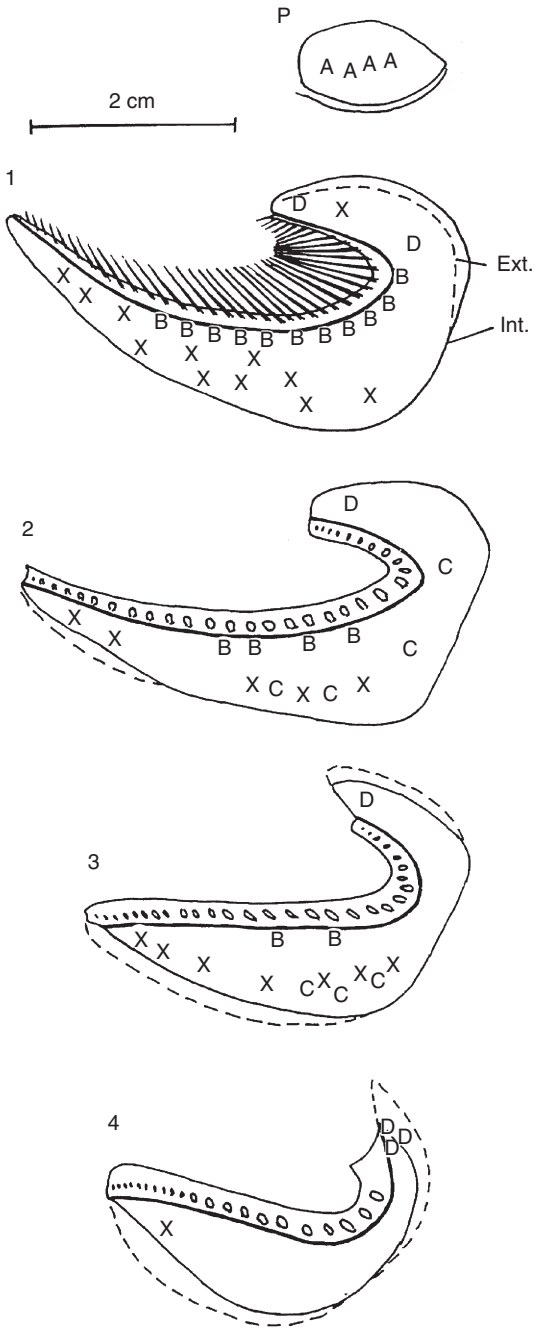


Figure 4 Monogenean gill parasites on the gills of mackerel, *Scomber australasicus*, off southeastern Australia. P, pseudobranch; 1–4, gills 1–4; ext, external gill filaments; int, internal gill filaments. A, *Kuhnia sprostonr*, B, *Kuhnia scombri*; C, *Kuhnia scombercolias*; D, *Grubea australis*; x, *Pseudokuhnia minor*. Note that species A–D have identical copulatory organs, and species x has different copulatory organs; A–D are spatially segregated from each other, and species x overlaps with B–D.

Some Physiological Adaptations

Parasites belong to a wide range of taxa and infect a wide range of hosts. Hence, their physiological adaptations are also diverse and cannot be discussed in any detail here. Jennings (in Rohde, 1997) has carried out detailed comparative studies of free-living, ecto-, and endoparasitic flatworms. These studies are particularly informative because they show the transition from free-living to commensal to ecto- and endoparasitic forms. Some free-living turbellarians living on the body surface and in the gill chambers of their hosts feed on the same kind of food as their free-living relatives, but in addition, they also feed opportunistically on food scraps of the host. Food reserves and digestive physiology do not differ from those of free-living species. Turbellarians living in the interior of mollusks, arthropods, and echinoderms have become increasingly dependent on their hosts. Some feed on protozoans also

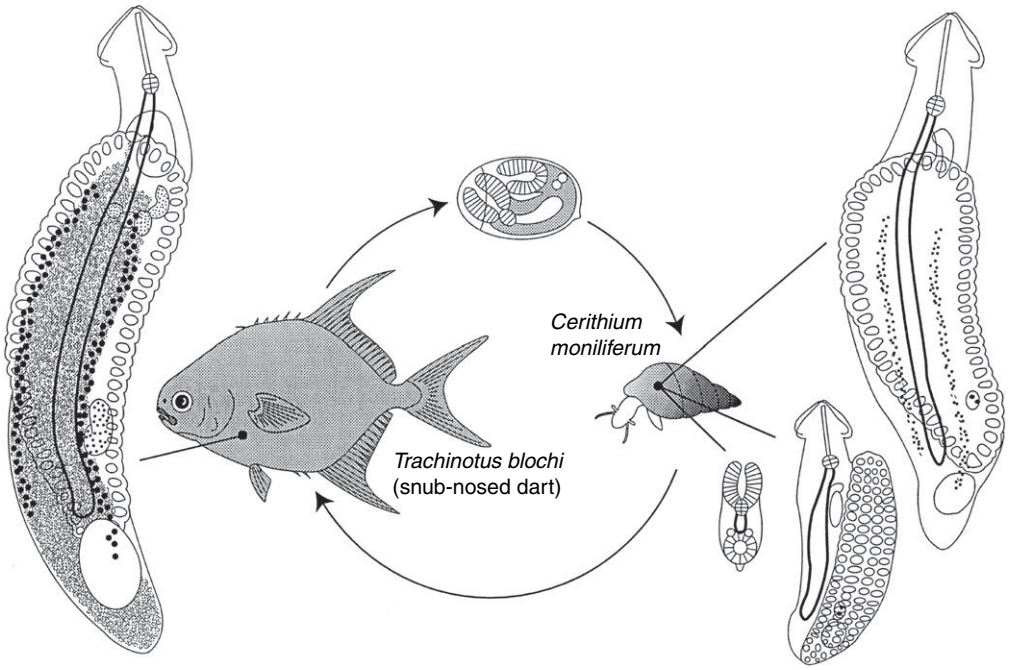


Figure 5 Life cycle of *Lobatostoma manteri* (Trematoda, Aspidogastrea). Note one intermediate (snail) and a final host (fish). No multiplication of larvae occurs in the intermediate host.

found in the hosts as well as on food ingested by the host, intestinal cells, and coelomocytes. Others feed mainly on intestinal cells, using some digestive enzymes from the host, and others entirely lack digestive enzymes and depend on enzymes ingested with host tissues. Most of the endoparasitic turbellarians do not store lipids (as free-living and ectocommensal forms do), but they store glycogen instead. Glycogen storage is also characteristic of the Neodermata (the major groups of parasitic flatworms, including the Monogenea, Trematoda, and Cestoda). Some endoparasitic turbellarians possess physiologically active hemoglobin, which allow preferential abstraction of oxygen from host tissues.

Adaptations of Flowering Plants

Parasitic plants have haustoria, which form a close connection with the vascular system of host plants, either in the roots or in the shoots. They depend entirely or partly on the host for water and inorganic and organic solutes. Hemiparasites, i.e., plants only partly dependent on the host, have chlorophyll, whereas holoparasites, i.e., plants entirely dependent on the host, lack chlorophyll.

Origins of Parasitism and Complex Life Cycles, The Evolution of Virulence, and Coevolution of Hosts and Parasites

Origins of Parasitism and Complex Life Cycles

Few fossil parasites are known. They include schistosome eggs from ancient Egyptian mummies a few thousand years ago and galls on the arms of feather stars, probably produced by Myzostomida (parasitic annelids) from the Silurian and Devonian periods, 350–430 million years ago (Ma). Conclusions on the origins of parasitism and parasite life cycles must therefore be based on inferences from comparative studies of extant species. The Platyhelminthes have been studied most thoroughly, using DNA studies of several genes, in particular 18-S rDNA and 28-S rDNA, and phylogenetic systematics (cladistics). Phylogenetic systematics seeks to establish branching patterns in phylogeny on the basis of shared acquired characteristics (synapomorphies). Ultra-structural characteristics are of particular use because of their complexity. Cladistic and DNA studies by many authors have consistently shown that the Neodermata, the major groups of parasitic flatworms (Trematoda, Monogenea, and Cestoda), all share a common ancestor, i.e., are monophyletic. The Neodermata, or the Neodermata plus some of the parasitic turbellarians, are the sister group of a very large taxon, including most Turbellaria, with which they share a common ancestor. This means that the parasitic groups evolved very early in evolutionary history. Among the Neodermata, the trematodes are the sister group of the other Neodermata and, among the trematodes, the Aspidogastrea are the sister group of the other trematodes or Digenea. Three of the four families of Aspidogastrea occur in elasmobranchs, whereas almost all digeneans parasitize teleost fishes, amphibians, reptiles, birds, and mammals; very few species of Digenea have been recorded from elasmobranchs, to which they have secondarily adapted. Fossil records indicate that elasmobranchs are 450 My old and teleosts are 210 My old. This suggests that the Aspidogastrea are the oldest extant trematodes and possibly neodermatans. It also suggests that the simple life cycle of Aspidogastrea (**Figure 5** is the original life cycle of trematodes, including a final and an intermediate host (which, in some species, is not obligatory),

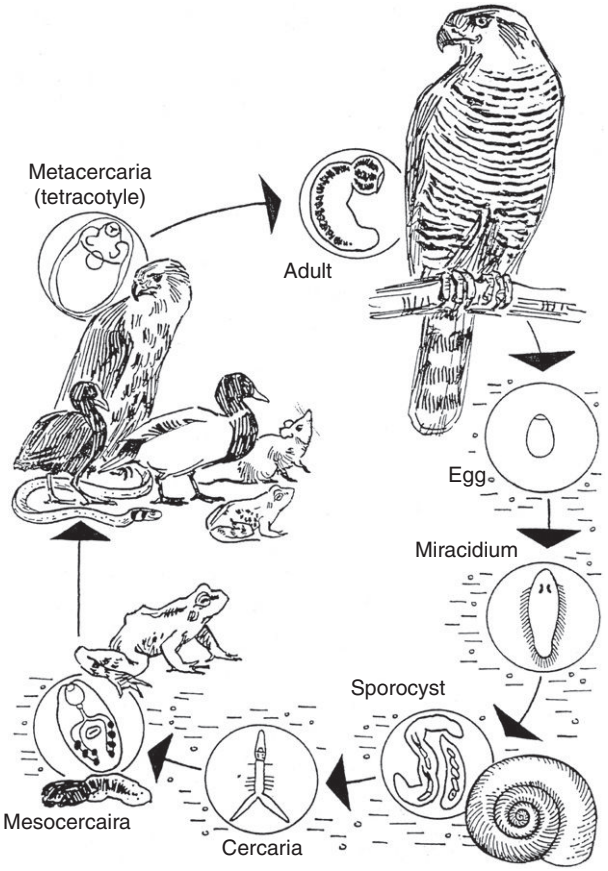


Figure 6 Life cycle of the trematode *Strigea falconispalumbi*. *Note*: Final hosts (predatory birds) containing adult worms that produce eggs in which miracidia develop, first intermediate host (snails) containing the sporocysts that produce cercariae, second intermediate host (tadpoles/frogs) containing mesocercariae, and third intermediate hosts (amphibians, snakes, birds, mammals) containing metacercariae. Modified from Odening K (1969) *Entwicklungswege der Schmarotzerwürmer.* Leipzig: Akademische Verlagsgesellschaft Geest & Portig K. G. © Spektrum Akademischen Verlag, Heidelberg, Berlin, with permission from Wiley.

without multiplication of larval stages in the intermediate host. The complex life cycles of digenean trematodes, including a final and at least one, and up to three, intermediate hosts and several transport hosts, with multiplication of larval stages in the first intermediate host (**Figure 6**), may have evolved from this primitive kind of life cycle to make transmission to the final host more effective.

Among the crustaceans, most barnacles are free-living, attached to rocks or other hard substrata. Some barnacles live in a phoretic association, for instance with whales, attached to their skin and feeding on plankton in the environment. Other, closely related species have become parasites. Thus, Anelasma parasitizes the skin of sharks, processes of its stalk branching in the host's muscles and extracting food from it. All the approximately 120 species of rhizocephalans, also related to the barnacles, are parasites and strongly modified in adaptation to their way of life (see, e.g., Sacculina, **Figure 1**). This suggests that parasitism in the barnacles may have evolved from free-living to phoretic to parasitic.

Evolution of Virulence

Intuitively, one might suspect that it is not in the parasite's interest to severely damage or even kill its host, because this would also affect the fitness of the parasite. However, parasite transmission to another host may well be facilitated by such damage to an intermediate or a transport host. In other words, evolutionary pressure may have led to an increase in virulence in some cases and to a decrease in others. Anderson and May have developed an epidemiological model that considers virulence, as follows: $R_0 = \beta(N)/(\mu + a + v)$. R_0 is the fitness of the parasite, its lifetime reproductive success; β is the rate of transmission of the parasite by its host to other hosts, which is dependent on host density N; μ is the mortality rate of uninfected hosts; a is the virulence or the mortality rate induced by the parasite; and v is the recovery rate of the host. This model (applied and further developed by various authors) allows some predictions on the evolution of virulence under different conditions. For example, a parasite that can achieve a large increase in the transmission rate by a small increase in virulence should have lower optimal virulence than one that achieves less by such an increase. The transmission rate depends, among other factors, on the likelihood of transmission for instance facilitated by the behavior of the host. Low virulence will evolve when contacts between hosts and thus opportunities for transmission are frequent. High virulence may evolve in vector-transmitted infections, and optimal virulence may differ in parasites that are vertically transmitted (from parents to offspring) from those that are horizontally transmitted. Although the parameters in the model are difficult to measure, some of the predictions have been verified empirically. Thus, parasites transmitted by contact (e.g., lice) are usually less virulent than those transmitted by vectors (e.g., malaria), and sexually transmitted parasites (e.g., *Trichomonas vaginalis*) are usually less virulent and longer lasting than parasites transmitted in other ways. In vertically transmitted nematodes of fig wasps and in vertically transmitted microsporans of mosquitoes, virulence is less than that in horizontally transmitted related species.

Coevolution of Hosts and Parasites

The comparison of phylogenetic trees of hosts and parasites has yielded contradictory results concerning possible coevolution. The reasons may be biological differences between the groups studied or they may be due to different interpretations of the results by different authors. For example, according to several authors, chewing lice infecting rodents of the family Geomyidae, as well as lice of seabirds, seem to have coevolved closely with their hosts. In contrast, lice of rock wallabies in Australia, apparently, have switched hosts frequently and there is little evidence for coevolution. Altogether, there is insufficient knowledge indicating how common coevolution is.

Host–Parasite Interactions

Cleaning Symbiosis

Hosts use a variety of behavioral methods to rid themselves of parasites. Preening of birds, bathing of birds in dust and water,

and passive and active anting (letting ants passively crawl over the body or actively squeezing ants over the plumage, respectively) are thought to help in reducing parasite loads, although evidence is scarce. Some other behavioral patterns (dolphins rubbing against rocks, fish jumping out of the water, etc.) may also play a role. Most widely distributed and well studied, particularly in the marine environment, is cleaning symbiosis, in which one animal (the cleaner) cleans another (the host), removing its parasites and diseased tissues. Birds remove ticks and other ectoparasites from cattle, hippopotamus, and even large marine fish floating on the surface, and several species of shrimps, as well as greater than 100 species of fish, are cleaners in freshwater but mainly in the sea. Besides fishes, the Galapagos marine iguana, whales and dolphins, invertebrates, etc. are hosts to cleaners. Cleaner fish often have special morphological adaptations, such as a terminal mouth and fused anterior teeth and conspicuous color patterns, the so-called guild signs. In addition, the Indo-Pacific cleaner wrasse, *L. dimidiatus*, performs a cleaning dance that attracts host fish. Hosts, on their part, show invitation postures signaling to the cleaner that they are ready to be cleaned, and fish of many species, usually hostile to each other, queue peacefully up at cleaning stations (territories where cleaning occurs). They even allow cleaners to enter their mouth cavity. In well-established cleaning symbioses, hosts rarely or never eat cleaners. Some fish were observed to spend as much time at cleaning stations as they spend on feeding, and some cleaner species feed exclusively on parasites and diseased host tissue. Certain fish mimic the color pattern and behavior of cleaners to approach hosts in order to attack them and bite off pieces of the fin or skin.

Immune and Tissue Reactions, Resistance

Defense reactions to parasites at the humoral and tissue levels are based on the ability of the host to distinguish self (its own cells) from nonself (foreign cells and material). Vertebrates have three types of such reactions: phagocytosis, inflammation, and adaptive immunity. The first two are nonspecific tissue reactions and the third is specific to a certain type of nonself material. They are found in all vertebrates, but (particularly immunity reactions) are best developed in birds and mammals. All three defense mechanisms usually interact and occur in most tissues and organs. Typically, the reactions occur in a certain sequence: degeneration or necrosis of cells due to the infection leads to an inflammatory response with edema (swelling of tissue). Phagocytic cells engulf small parasites, but if parasites are not eliminated, a chronic inflammation develops, leading to a connective tissue capsule around the parasite; macrophages in the capsule engulf damaged cells and often the parasite. Special defense mechanisms are active on the surface. Thus, fish continually shed mucoid material from the skin even if uninfected, but in infected fish, the slough increases and leads to the removal of monogenean and other ectoparasites.

Immune reactions are induced by antigens of the parasite that lead to the formation of specific antibodies in the host. In microparasites (protozoans and some helminths, e.g., *Strongyloides stercoralis*, but particularly in bacteria and viruses, which are not discussed here), immune responses are more effective than in macroparasites (most helminths and arthropods), where they are either lacking or only short-lasting. Immune responses have been particularly well studied in trypanosomes. The antigen is a surface glycoprotein that completely covers the parasite. The host's antibodies eliminate most parasites, but a few trypanosomes of different antigenic types survive and build up a new population. This is repeated over and over again, leading to marked fluctuations in infection intensities. The host finally dies, because the immune reactions are not sufficient to destroy the entire parasite population. The number of antigenic types is very large and apparently limited only by the host's life span: more than 100 types were shown to develop in a single clone strain. In parasites with complex life cycles, such as the malaria parasite, each stage in the life cycle has different antigenic properties, and there are many variants because of the large numbers of daughter cells produced by schizogony. These are the reasons why attempts to develop effective vaccines have failed so far.

Invertebrates, apparently, cannot acquire specific immune responses; they rely on phagocytosis and capsule formation. Pearl formation in bivalves, for example, is the result of encapsulation of foreign bodies.

Hosts show different degrees of resistance to infections that are not due to acquired immunity. For example, some sheep may be less susceptible to nematodes than others because of their genetic makeup, and individuals of the same species of different ages may differ in susceptibility. Older individuals are often less infected than young ones, a phenomenon referred to as age resistance. For example, the cestode *Austramphilina elongata* uses turtles as final and crayfish as intermediate hosts, but only young crayfish can be infected experimentally: the cuticle of older individuals prevents successful penetration of the larva.

Effects on Host Individuals and Populations

There is a huge variety of effects on host individuals, depending on the parasite and host species, site of infection, virulence of parasites, and susceptibility of hosts. A few examples of human parasites may illustrate this. The mite *Demodex folliculorum*, intestinal nematodes at low infection intensities, Toxoplasma, etc. often do not cause symptoms. However, Toxoplasma and the nematode *Onchocerca volvulus* can lead to blindness and filariae can cause elephantiasis, a sometimes enormous swelling of the scrotum, legs, arms, etc. Malaria causes a range of symptoms from influenza-like to death. Hookworms, depending on the infection intensity, can cause severe anemia and death. Cysticerci of *Taenia solium*, infecting the brain, may lead to epilepsy-like symptoms and death. Nematodes of sheep often cause death, and monogeneans may severely damage the skin, fins, and gills of fish, particularly in aquaculture. Larval trematodes often cause partial or complete castration of their snail hosts, and they sometimes induce gigantism in the snails, i.e., a markedly larger body size of infected than uninfected snails.

Models predict that microparasites should be highly effective in controlling host populations, and this is indeed often the case. For example, trypanosomes and malaria

severely affect human populations and the former also affect livestock, and oyster beds were decimated by various protozoan parasites in several countries. However, almost all such reports deal with populations in abnormally high densities (humans, oysters) or with populations affected after the introduction of a parasite into an area where it was not present originally or after the introduction of a host species into a new area (livestock in Africa). The same applies to macroparasites. Host populations, under natural conditions, may sometimes be severely affected by macroparasites, but evidence is usually circumstantial and mass mortalities may not be due to parasites alone, but due to synergistic effects involving, for instance, environmental degradation, as well as parasites. Best documented are cases of mass mortalities caused by parasites in livestock and aquaculture, as well as after the introduction of parasites or hosts into a new habitat. Of historical interest is an epizootic caused by the liver fluke *Fasciola hepatica*, which caused the death of 3 million sheep in Great Britain in 1879/1880 and led to the clarification of the life cycle of this parasite. A well-documented case of an epizootic due to the introduction of a parasite is that of the monogenean *Nitzschia sturionis*, which was introduced into the Aral Sea with sturgeon from the Caspian Sea in the early 1930s. It devastated the local sturgeon population and led to the collapse of the sturgeon and caviar industry in the Aral Sea in 1937, which did not recover for 20 years. Recently, the first experimental proof was provided showing that a gastrointestinal roundworm can indeed regulate wildlife populations of red grouse in England.

The Ecological Niches of Parasites

The Niche Concept: Niche Dimensions of Parasites

A niche is defined as the total of an organism's relations to its biotic and abiotic environment. These relations define the organism's place in nature or, in ecological jargon, its place in a multidimensional niche space. Important niche dimensions of parasites are hosts, microhabitats, geographical range, sex of host, age, season, and food. Different parasite species use different host species and different microhabitats within or on the host. Thus, particular nematodes of the cat use different microhabitats (**Table 3**), and even within each microhabitat, there is further subdivision, some species using certain parts of the small intestine, the stomach, etc. Male lambs in the USA have more nematodes of certain species than females, and many parasites prefer hosts of a certain age, occur only at certain seasons, or are restricted to host populations that use certain food.

Saturation of Niches with Parasites

An important question in ecology is whether habitats are saturated with species or whether empty niches exist. The most extensive and intensive studies were carried out on parasite communities on the heads and gills of marine fish. Results on 112 host species showed that the maximum component species richness was 27 parasite species but that most species had less than six (**Figure 7**). Similar differences were found for abundances (intensities of infection with all parasite species), which ranged from less than five for most host species to greater than 3000, with no apparent signs of damage in the most heavily infected fish. This strongly suggests that most fish species, at least, could accommodate more parasite species and individuals, and even in fish with the greatest parasite species richnesses and abundances, some parts of the gills that are occupied in other species were not infected.

Proximate and Ultimate Causes of Niche Restriction

Proximate causes are physical and chemical factors determining niche selection of a parasite, whereas ultimate causes refer to the biological function of niche selection: why has selection favored one niche over another? There are many proximate causes responsible for niche selection of different parasites, although they have been poorly studied. Parasite larvae, for example, may use chemical stimuli to locate a particular host and, even if they can infect many hosts, they may survive only on the correct ones because a factor or factors produced by the wrong hosts may kill

Table 3 Nematodes of the cat: sites of infection

Site	*Species of nematode*
Kidney	Dioctophyma renale
Colon	*Strongyloides mystax*, *S. tumefaciens*, *Trichuris campanula*
Stomach and small intestine	Abbreviata gemina, *Ancylostoma braziliense*, Gnathosoma spinigerum, *Mastophorus muris*, Ollulanus tricuspis, *Physaloptera brevispiculum*, *P. felidis*, *P. pacitae*, *P. canis*, *Rictularia cahirensis*, *Soboliphyme baturini*, *Spirura rytipleurites*, *Toxascaris leonina*, *Toxocara canis*, *T. mystax*, *Trichinella spiralis* (adults), *Uncinaria stenocephala*, *Ancylostoma tubaeforme*, *Physaloptera praeputiale*
Skin	*Dirofilaria repens*
Lungs	*Aelurostrongylus abstrusus*, *Anafilaroides rostratus*, *Capillaria aerophila*, *Troglostrongylus subcrenatus*, *Vogeloides massinoi*, *V. ramanujacharii*
Heart	*Dirofilaria immitis*
Middle ear	*Mammomongamus auris*
Spinal cord	*Guretia paralysans*
Lymph vessels	*Brugia malayi*
Diaphragm and other striated muscles	*T. spiralis* (juveniles)

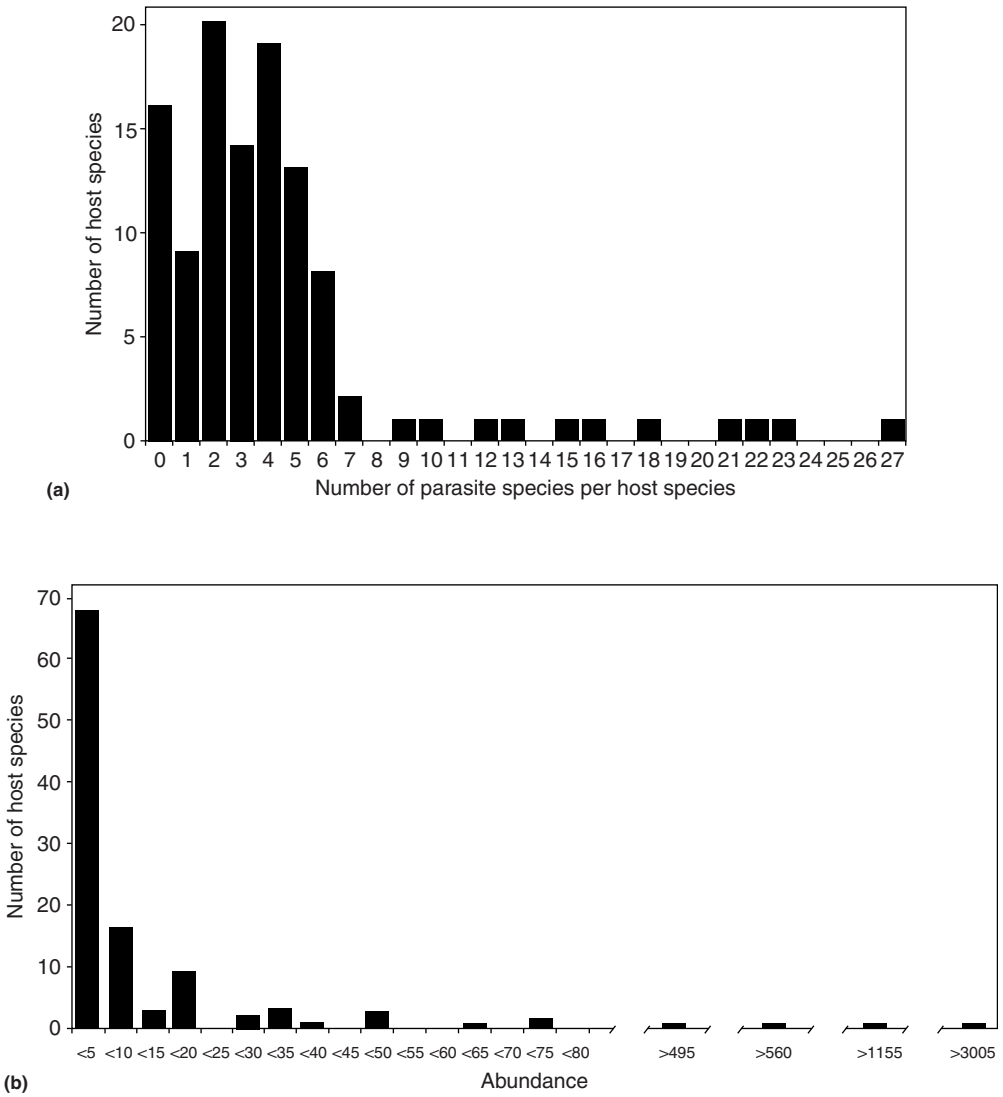


Figure 7 Number of ectoparasite species per host species (a) and abundances (b) of infection on 112 teleost species (5666 fish). Reproduced from Figure 1 in Rohde K (1998) Latitudinal gradients in species diversity. Area matters, but how much? *Oikos* 82: 184–190, with permission from Nordic Society Oikos.

them. Exsheathment and development of parasite larvae of different species may depend on different stimuli, redox potential, and pH in different parts of the intestine, thus determining microhabitat differences. Males and females of a host species may simply acquire different parasites because of different feeding habits, etc. Ultimate causes may be extrinsic, due to interspecific effects, or intrinsic, due to intraspecific effects. Among the interspecific effects, interspecific competition is generally considered to be the most important one, as evidenced by the occasional exclusion of one species by another, microhabitat shifts in the presence of other species, etc. However, altogether, evidence is scarce. Many cases of microhabitat differences that have been explained by interspecific competition are likely to be the result of reinforcement of reproductive barriers. Thus, only those monogeneans that have identical copulatory organs are spatially segregated in different microhabitats, whereas species with dissimilar copulatory organs coexist in the same microhabitat (**Figure 8**, also **Figure 4**), suggesting that microhabitat segregation does not have the function to avoid competition but hybridization between closely related species. Only one intrinsic factor has been suggested: facilitation of mating. Many parasites live at low infection intensities and prevalences (**Figure 7**). Restriction to certain hosts and microhabitats may therefore vastly increase the chances of meeting a mating partner. Experimental evidence, however, is scarce.

The Structure of Parasite Communities

Concepts of Community Ecology

Community ecology deals with the relations between organisms in a certain habitat, in the case of parasites with relations between parasites infecting a certain host. Such relations can be studied at different levels, the level of infracommunity, component community, and compound community.

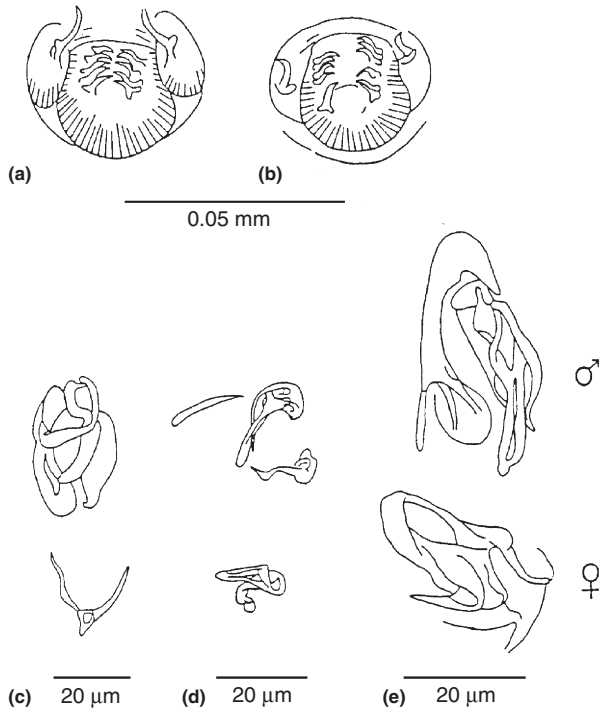


Figure 8 (a, b) Copulatory organs of two species of the monogenean Kuhnia inhabiting different parts of the gills of mackerel, Scomber spp. (c–e) Male and female copulatory organs of three species of the monogenean Lamellodiscus inhabiting the same parts of the gills of bream, *Acanthopagrus australis*. Reproduced with permission from Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

Infracommunities of parasites consist of all the infrapopulations within a host individual. Component communities consist of all infrapopulations within a host population, and compound communities consist of all parasite communities within an ecosystem. An infrapopulation is the total of all the individuals of a parasite species within a host individual. Two of the major questions of community ecology are (1) whether communities show predictable patterns of species composition, relative abundances, and resource use and (2) what processes are responsible for the patterns. Studies of endo- and ectoparasite communities of various vertebrate hosts have led authors to distinguish interactive communities, characterized by species regularly occurring in huge densities with much interaction (interspecific competition) between them, and isolationist communities, characterized by species occurring at low densities with little interaction between them. Both types, supposedly, are extremes at the ends of a spectrum, with many intermediate kinds of communities between them. Authors also distinguished between core and satellite parasite species, the former dominant species with high prevalences and intensities of infection and the latter with low prevalences and intensities of infection.

Empirical Evidence

Even in host species with extremely species-rich parasite communities, core and satellite species usually cannot be distinguished, and evidence indicates that many more parasite species than actually present could be accommodated (i.e., that vacant niches exist). This, and the frequency of positive and the scarcity of negative associations between species, scarcity of nestedness, greater intra- than interspecific aggregation, and no or only minor effects of the number of parasite species on microhabitat size, strongly suggest that interspecific competition is not of great importance. Concerning the structure of communities as indicated by predictable co-occurrence and the relative abundance of species, studies of ectoparasite communities of marine fish showed that the most dominant species usually represented between 60% and 80% of all parasites in an infracommunity but that different species were dominant in different infracommunities. For communities of larval trematodes in snails, Sousa (see Rohde, 1993) has shown that there is a significant difference between communities studied at different scales. Interactions are important at the infracommunity level, but at the level of component community, interactions are not important and communities are largely structured by external processes.

Parasite Communities as General Ecological Models

The vast majority of animal species are probably arthropods parasitizing plants, and they share many characteristics with parasites of animals: their habitats are resource-rich and seldom exhausted by the parasites. The major problem for such parasites is not to avoid competition with other species but to find the appropriate hosts and sites for feeding and mating. The conclusions based on the study of parasites of animals can therefore be applied to them.

Parasite Population Dynamics

Concepts of Population Growth

Population growth is described by the logistic growth equation $dN/dt = rN[(K-N)/K]$. dN/dt is the rate of population growth, N is the number of individuals at the time t, r is the *per capita* rate of natural population increase, and K is the carrying capacity of the habitat (the maximum number of individuals a habitat can support). The equation shows that population growth is exponential when population density is small and that it decreases with increasing N and approaches 0 when N approaches K. In a graphic representation, the curve becomes asymptotic, i.e., it flattens out when population density reaches the carrying capacity. In habitats in which, for example, populations repeatedly become impoverished as a result of external disturbances, species will be favored that can reproduce and develop rapidly, without huge investment in the individual offspring. In stable habitats where population size is usually close to the carrying capacity, it pays off to have few but well-adapted offspring. Selection for large numbers of fast-developing offspring is referred to as r-selection, whereas selection for few well-adapted offspring is referred to as K-selection. However, the general applicability of this distinction has been questioned because population density is only one of the factors that determine selection pressure.

Nevertheless, the distinction is still useful as an approximation to the real scenario.

Ecological Strategies of Parasites

Most parasites rely on the production of many offspring and therefore tend to be r-strategists. In some cases, numbers are controlled by abiotic factors and not by population size; i.e., the factors are density independent. This has been demonstrated for some helminth species, for example, the cestode *Bothriocephalus acheilognathi* in the freshwater fish *Gambusia affinis*, whose prevalence and intensity of infection varied strongly with water temperature over several years in a lake in North Carolina. However, few such studies have been conducted and generalizations are therefore premature. Density-dependent factors, i.e., factors dependent on the population size of the parasite, also play a role in some cases. For example, predation by trematode larvae in snails on other trematode larvae, immune responses of hosts that depend on infection intensities (as in many helminth infections), and competition between parasites of the same and of different species are such density-dependent factors. There is experimental evidence for all of these, but studies are too few to allow generalizations on how important such effects are. For example, high infection intensities often lead to stunted growth and a reduction in the number of offspring, an example of intraspecific competitive effects. The microhabitat of a parasite species may shift or become smaller if other species are present, or one parasite species may reduce the numbers of another or even completely eliminate it, examples of interspecific competition. However, the importance of interspecific competition, in particular, generally seems to have been overestimated.

The Diversity of Parasites (Distribution of Parasites in the Animal and Plant Kingdoms)

The first estimate of the total number of parasite species in a fauna was made by Arndt in 1940. He counted 10,000 parasitic species among the total of 40,000 species in Germany but did not include insect herbivores closely associated with plants and therefore, according to our definition, parasitic. Price, in 1977, including such species but excluding temporary parasites such as mosquitoes and leeches, estimated that more than half of all British species are parasitic. Parasites, because of their small size, are less well known than free-living forms, and this applies especially to parasites of invertebrates, which have been little studied. Estimates of parasitic species may therefore well be too low. However, recent studies of the fauna of tropical rain forest canopies have revealed an enormous number of insect species and we do not know how many of these are parasitic. There are no studies of parasites of these insects. Hence, even approximate estimates of the proportion of parasitic species are premature, although we can state with confidence that parasitism is by far the most successful way of life on earth.

Data in **Table 4** show that 13 large taxa (phyla, subphyla, or classes), some of them are very large, consist entirely of parasites, and many other groups include a high proportion of parasitic species. Even some vertebrates are parasitic. Males of a number of deep-sea fishes parasitize females of the same species (**Figure 9**). The needle fish, *Carapus acus*, feeds on the viscera of sea cucumbers, which represent an almost inexhaustible source of food because they regenerate, cleaner mimics (fish that mimic genuine cleaners) approach fish and bite off pieces of the fin or skin, and gulls were observed feeding on the flesh of whales. Cowbirds and about 50 species of cuckoos are brood parasites; i.e., they lay their eggs into the nests of other birds, which then incubate them. Skuas and frigate birds chase other birds in flight and feed on their regurgitated food (kleptoparasitism). Kleptoparasitism in its various varieties is widespread among humans.

Concerning plants, estimates are that 1% of flowering plants, about 3000–4000 species, are parasitic. Parasitism has evolved at least eight times in the angiosperms, and 16 plant families have parasitic species.

Zoogeography of Parasites

Latitudinal Gradients in Species Richness and Abundance

The most dangerous parasites of humans are found in warm countries: malaria (originally also found in temperate countries such as those of northern Europe), filariasis, guinea worm, sleeping sickness, Chagas disease, and several other diseases are restricted to tropical–subtropical countries. However, latitudinal gradients in species richness have been quantified best for parasites of marine fishes. Species number of marine fishes is greatest in the tropics. Species richness of metazoan ecto- and endoparasites also increases toward the tropics. However, interestingly, both the number of ectoparasitic species per host fish (their relative species richness) and their abundance increase at lower latitudes, whereas this trend does not exist for the endoparasites (**Figure 10**). The reasons for these differences are not understood.

Latitudinal Gradients in Reproductive Strategies

Such a gradient has been studied only in marine Monogenea. Most monogeneans produce eggs in which ciliated larvae (oncomiracidia) develop; one family, the Gyrodactylidae, consists exclusively of small viviparous species. Whereas gyrodactylids are very rare in warm waters, they represent the vast majority of all species in cold waters (75–89% in the Bering, White, and Barents Seas). This corresponds to a trend in marine benthic invertebrates, which produce large numbers of eggs from which pelagic larvae develop in warm waters and small numbers of larvae by viviparity, etc. in cold waters (Thorson's rule).

Latitudinal Gradients in Host Ranges and Host Specificity

Such trends have not been studied and are not obvious in terrestrial and freshwater parasites. Studies of trematodes and monogeneans of marine fish have shown that the former have greater host ranges, i.e., they use more host species, in cold than in warm waters, whereas monogeneans do not show such a trend, using very few hosts at all latitudes. Host specificity

Table 4 Estimated number of parasitic species among various groups in the animal kingdom[a]

Phylum	*Subphylum (Protoza) of class*	*Total number of species*	*Number of parasite species*
Subkingdom Protozoa		70,000	
Sarcomastigophora	Mastigophora	?	2000
	Opalinata	150	**150**
	Sarcodina	?	**250**
Labyrinthomorpha		**Some**	**Some**
Apicomplexa			
Microspora (Microsporidia)		>7000	>7000
Ascetospora (Haplosporidia)			
Ciliophora		?	2500
Subkingdom Metozoa		**55**	**55**
Mesozoa	**Dicyemida**	**40**	**40**
	Orthonecctida	**15**	**15**
		5000	1 (few?)
Porifera		8900	20
Crudaria	Hydrozoa	2700	few
	Scyphozoa	200	0 (?)
	Anthozoa	6000	few
Ctenophora		80	1 (few?)
Myxozoa (**Myxosporidia**)		>1500	>**1500**
Platyhelminthes		>*ca.* 80,000	>*ca.* 65,000
	Turbellaria	>15,000	>200
	Trematoda	>25,000	>25,000
	Monogenea	>35,000	>35,000
	Cestoda	>2500	>2500
Gnathostomulida		80	0
Priapulida		3	0
Entoprocta		60	0
Nemertina		750	10
Nemathelminthes		>12,000	*ca.* 5600
	Rotatoria	1500	20
	Gastrotricha	150	0
	Nematoda	>50,000	>20,000
	Nematomorpha	<**100**	<**100**
	Kinorhyncha	100	0
	Acanthocephala	***ca*. 500**	***ca*. 500**
Annelids		>7000	*ca.* 420
	Polychaeta	>4000	20
	Myzostomida	**111**	**111**
	Oligochaeta	2400	40
	Hirudinca	300	250
Onychophora		70	0
Arthropoda		Many millions	Many millions
	Xiphosura	5	0
	Arachnida	30,000	4300
	Pycnogonida	350	<350
	Crustacea	20,000	<2500
	Myriapoda	10,500	0
	Insecta	Many millions	Many millions[b] (>50%)
Tardigrada		180	1(?)
Pentastomida		75	75
Tentaculata	Phoronidea	18	0
	Bryzoa (Ectoprocta)	>4000	0
	Brachiopoda	280	0
Mollusca		112,000	>100
	Solenogastres	140	Few(?)
	Placophora	1000	0
	Gastropoda	85,000	100
	Scaphopoda	300	0
	Bivalvia	25,000	>10
	Cephalopoda	>600	0
Echiurida		70	Few (intraspecific parasite)

(*Continued*)

Table 4 Continued

Phylum	*Subphylum (Protoza) of class*	*Total number of species*	*Number of parasite species*
Sipuncula		250	0
Hemichordata		80	0
	Enteropneusta	60	0
	Pterobranchia	20	0
Echinodermata		*ca.* 6000	≥ 2
	Crinoidea	620	0
	Holothuroidea	1100	1(?)
	Echinoidea	860	0
	Asteroidea	1500	0
	Ophiuroidea	1900	2
Pogonophora		47	0
Chaetognacha		50	0
Chordata		62,000	Some
Total		Many millions	Millions

[a]Groups that consist entirely of parasitic species are printed in bold. After various authors. Only extant species are included.
[b]Many authors refer to many of the plant-parasitic insects as herbivorous insects.

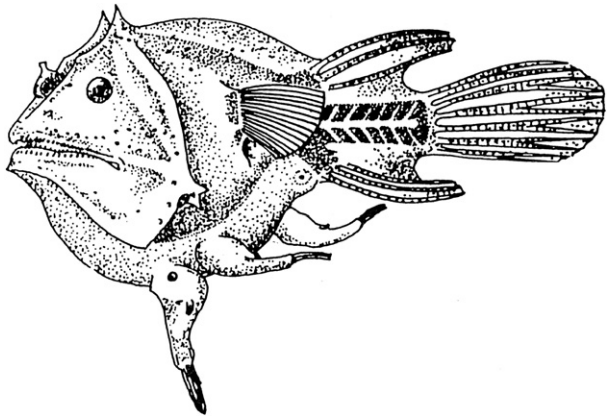


Figure 9 Intraspecific parasitism of vertebrates. Three males of the deep-sea fish *Edriolychnus schmidti* fused to a female. Reproduced from Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

(i.e., the number of host species infected taking the prevalences and intensities of infection in different host species into account) is the same at all latitudes for both groups.

Parasites as Biological Markers

Parasites are frequently used as biological markers to study host populations and migrations, particularly in the sea. A prerequisite for such a study is that parasites have a long life and are acquired only in the population or area studied. Larval helminths fulfill these conditions and are therefore usually used. Examples for the successful use of parasites as biological markers are studies that demonstrated different populations of herring with different Anisakis infections in the Baltic Sea, different populations of Atlantic salmon from different tributaries of the Miramichi River system in Canada, and different populations of sockeye salmon in the North Pacific, one harboring the nematode *Dacnitis truttae* acquired in Kamchatka and one harboring the larval cestode *Triaenophorus crassus* acquired in western Alaska.

The von Ihering Method

About 100 years ago, von Ihering first used parasites and ectocommensals to clarify places of origin and dispersal of hosts and ancient land connections between land masses. One assumption of the method is that animals have acquired the greatest diversity in the area where they have been the longest. Mamaev and Margolis applied the method to Pacific salmon. On the basis of their finding that salmonids harbor many freshwater parasites, some of them with complex life cycles, but no marine endoparasites, they concluded that salmonids are of freshwater origin. Many similar studies have been carried out.

Economic and Hygienic Importance of Parasites

Human Parasites are Among the Most Important Disease Agents of Man

The web pages of The World Health Organization, Division of Tropical Diseases (CTD), and of the Centers for Disease Control (CDC) contain information about the current status of the important parasitic diseases, which is continually updated. Information about the most important parasites of humans is given in **Table 5** and **6**. Note that some of the most widespread species, such as *D. folliculorum*, a mite infecting the skin, and *To. gondii*, are not included because of lack of data and usually symptomless infections.

A number of parasite species, most of them protozoans, including the nematode *S. stercoralis*, have recently become more important because of AIDS. All these species multiply within the human host and may become fatal in immunodepressed patients. The protozoans involved are, among others, Cryptosporidium and Giardia, sometimes acquired in polluted drinking water, *Pneumocystis carinii*, Microspora, and *Entamoeba histolytica*.

Resistance has developed to drugs used against various parasite species and to insecticides used against various vectors.

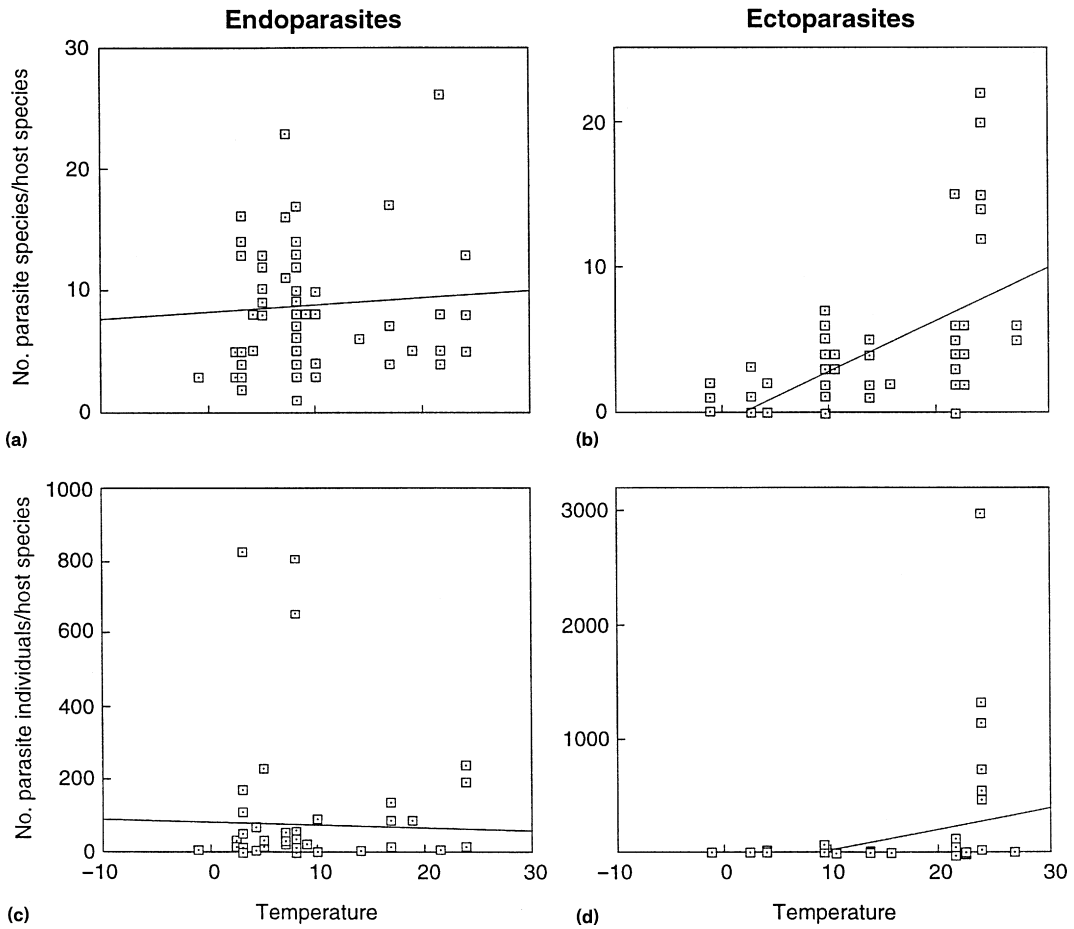


Figure 10 Metazoan ecto- and endoparasites on the heads and gills of marine teleost fish at various latitudes. Reproduced from Figure 1 in Rohde K and Heap M (1998) Latitudinal differences in species and community richness. *International Journal for Parasitology* 28: 461–474.

Table 5 Human infections with the most important parasites worldwide

Parasitic infection	*Number infected (million)*
Intestinal roundworms	1400
Malaria	300
Bilharzia	200
Lymphatic filariasis	120
Amebiasis	40
Food-borne trematodes	40
Chagas disease	16–18
Leishmaniasis	12
Sleeping sickness	0.3
Guinea worm	0.1

Source: Reproduced from McGill: The World of Parasites (http://martin.parasitology.mcgill.ca).

Hence, the epidemiological situation is becoming worse. Malaria is becoming more prevalent in many countries, not only due to insufficient financial and human resources for control but also because of human-induced climatic and environmental changes, migration, and war. Likewise, the prevalence of intestinal parasites worldwide is increasing, partly due to increasing

Table 6 Infections of the most common parasites in some countries/regions

Region	*Infection*	*Number infected*
United States	Pinworms (*Enterobius vermicularis*)	50 million
	Strongyloides stercoralis	<1 million
South America	Bilharzia	45 million
	Chagas disease	16–18 million
	Strongyloides stercoralis	ca. 7.4 million
Africa	*Ascaris lumbricoides*	>200 millon
	Bilharzia	100 million (mortality, 0.01 million)
	Malaria	23 million (mortality, 2.6 million)
China	*Ascaris lumbricoides*	ca. 100 million
	Opisthorchis sinensis	5 million
Middle East	*Ascaris lumbricoides*	100 million
	Hookworms	60 million
	Bilharzia	50 million

Note: *Demodex folliculorum* and Toxoplasma not included.
Source: Reproduced from McGill: The World of Parasites (http://martin.parasitology.mcgill.ca).

urbanization. In contrast, the global prevalence of guinea worm (*Dracunculus medinensis*) has decreased from about 10 million 10 years ago to about 150,000 (most of them in the Sudan), as a result of a campaign to eradicate the infection. Almost certainly, malaria and several other important parasite diseases would spread into areas presently not affected if global warming as a result of the greenhouse effect should occur.

Parasites of Livestock also Have Very Great Economic Importance

For example, nagana, a disease of many domestic animals caused by *Trypanosoma brucei brucei* and transmitted by tsetse flies in Africa, has made large areas of sub-Saharan Africa unsuitable for livestock production, and nematodes of sheep and cattle cause large economic losses in many countries. *Ostertagia ostertagi*, one of the nematodes infecting cattle, is estimated to cause an annual loss of $600 million to the cattle industry in the USA alone. The situation is not improving because of drench resistance of nematodes. Thus, in Australia, about 90% of sheep farms have worms resistant to one or more drench classes.

Parasites Have Repeatedly Caused Large Economic Losses in the Fisheries Industries

In particular, aquaculture is affected. On the East Coast of North America, for example, several protozoan parasites have repeatedly decimated oyster culture in several areas. The haplosporidian *Haplosporidium nelsoni* caused a decline in oyster production in the New Jersey waters of the Delaware Bay from about 5–8 million pounds between 1950 and 1955 to 167,000 pounds in 1960, which has never fully recovered (**Figure 11**).

Many Parasites, Particularly Nematodes, Are Among the Most Important Pests of Plants

For example, wheat is attacked by the nematode *Anguina tritici*, potatoes are attacked by the potato cyst nematode

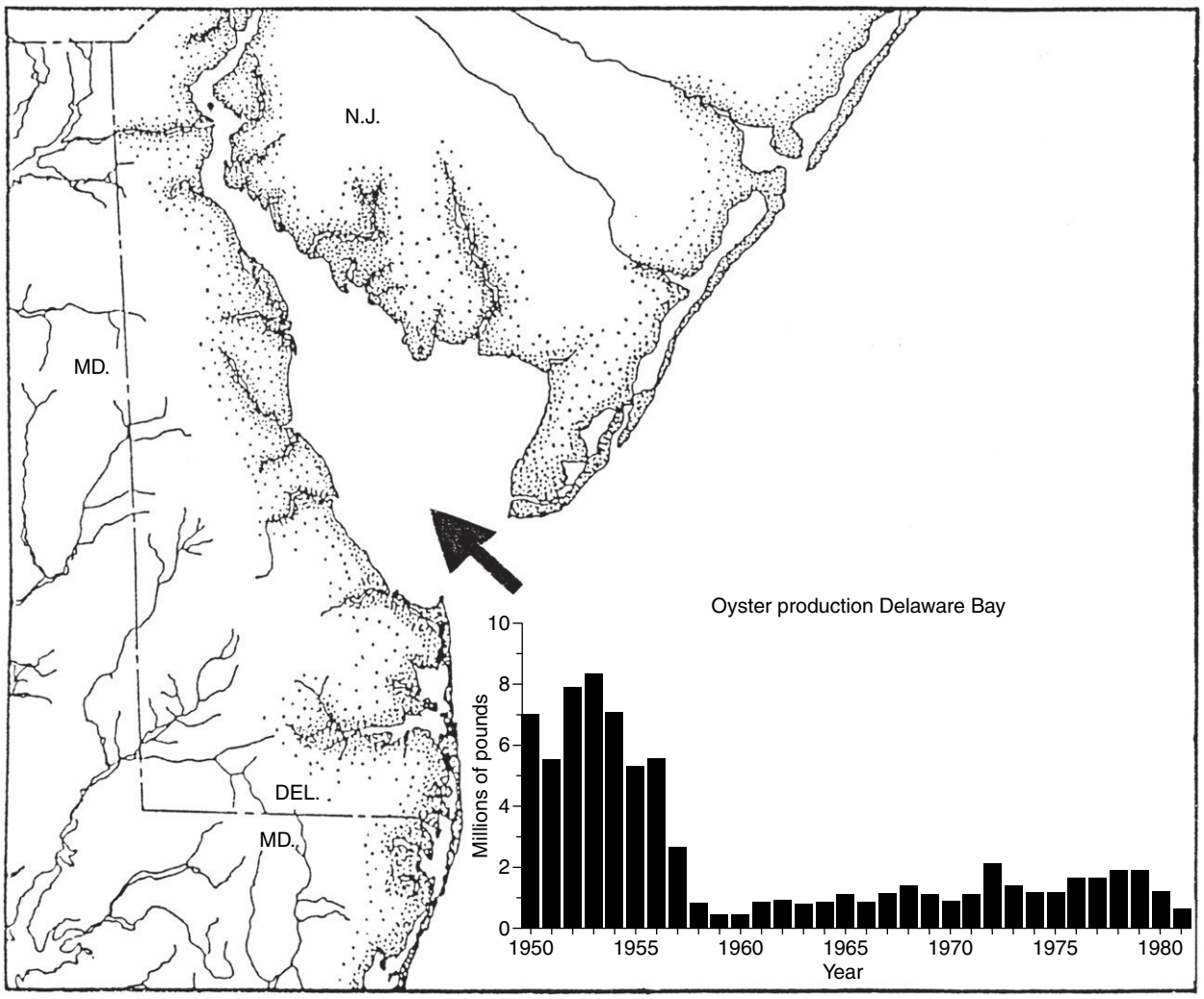


Figure 11 Oyster production in Delaware Bay, East Coast of the USA. *Note*: Collapse of oyster production in the late 1950s due to the haplosporidian *Haplosporidium nelsoni*. Modified after Sindermann, with permission from Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

Globodera rostochiensis, and rape, rice, etc., all have their specific nematodes that lead to large economic losses. The estimated annual loss to crop production due to nematodes in the USA is $8 billion (12%), and it is $78 billion globally (http://ianrwww.unl.edu/son/nn_nema.htm).

Parasitic Angiosperms Attack Many Crops

At least 11 species of witchweed, Striga spp., attack crops, including all the important tropical cereals; control is difficult because of the high production of seeds and longevity of seeds up to 20 years in the soil. Dwarf mistletoes almost always cause deformities of trees, leading to a substantial reduction in the yield and quality of timber. The mistletoe *Viscum album* reduced the yield of a particular apple variety in the UK by 7–54% (Press and Graves, 1995).

Parasites, Including Nematodes, Are Increasingly being Used to Control Insect Pests

Neoaplectana carpocapsae is bred in the laboratory and sprayed on rape to combat insects attacking it, and Howardula sp. effectively kills the dried-fruit beetle *Carpophilus mutulatis*. *Heterorhabditis bacteriophora kills* insects by carrying bacteria into them.

See also: Adaptation. Biogeography, Overview. Coevolution. Competition, Interspecific. Diversity, Community/Regional Level. Diversity, Organism Level. Global Species Richness. Habitat and Niche, Concept of. Hemiparasitism. Marine Ecosystems. Nest Parasitism. Parasitoids. Population Dynamics. Worms, Nematoda. Worms, Platyhelminthes

References

Bogitsh BJ and Cheng TC (1990) *Human Parasitology*. New York: Saunders College Publishing.

Esch G, Bush A, and Aho J (eds.) (1990) *Parasite Communities: Patterns and Processes*. London/New York: Chapman & Hall.

Esch GW and Fernandez JC (1993) *A Functional Biology of Parasitism: Ecological and Evolutionary Implications*. London/New York: Chapman & Hall.

Kinne O (1980) *Diseases of Marine Animals*, vols. 1–4. New York and Hamburg: Wiley and Biologische Anstalt Helgoland.

Miyazaki I (1991) *Helminthic Zoonoses*. Tokyo: International Medical Foundation of Japan.

Poulin R (2006) *Evolutionary Ecology of Parasites*, 2nd edn. Princeton, NJ: Princeton University Press.

Press MC and Graves JD (eds.) (1995) *Parasitic Plants*. London/New York: Chapman & Hall.

Rohde K (1993) *Ecology of Marine Parasites*, 2nd edn. Oxford: CAB International.

Rohde K (1994) Niche restriction in parasites: Proximate and ultimate causes. *Parasitology* 109: 69–84.

Rohde K (ed.) (1997) The origins of parasitism in the platyhelminthes, Proceedings of a Special Session of the VIIIth International Symposium on the Biology of the Turbellaria. *International Journal of Parasitology* 27(6): 677–746.

Rohde K (1998) Latitudinal gradients in species diversity. Area matters, but how much? *Oikos* 82: 184–190.

Rohde K (ed.) (2005) *Marine Parasitology*. Melbourne and Wallingford, Oxon: CSIRO Publishing and CABI.

Rohde K and Heap M (1998) Latitudinal differences in species and community richness. *International Journal for Parasitology* 28: 461–474.

Schmidt GD, Roberts LS, and Janovy J (1995) *Foundations of Parasitology*, 5th edn. New York: McGraw-Hill.

Urquhart GM and Jennings FW (1996) *Veterinary Parasitology*, 2nd edn. Ames, IA: Iowa State University Press.

Parasitoids

HCJ Godfray, University of Oxford, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 1–13, © 2001, Elsevier Inc.

Introduction

A parasitoid is a type of animal that has a life history intermediate between that of a predator and prey. In general, parasitoids are invariably insects and their larvae feed at the expense of other insects, with the exception of a few species that attack different types of arthropods or mollusks. The adult female parasitoid lays her eggs on, in, or occasionally near the body of the host; the eggs hatch and the developing parasitoid larvae consume the host, eventually killing it. Like a parasite, only a single host is required for full development, but like a predator the host is invariably destroyed. The term parasitoid, originally coined by Reuter at the beginning of the twentieth century, is now used nearly universally to describe species with this life cycle, although in the older literature they are sometimes just called insect or protean parasites. Some solitary wasps have a similar life cycle except that the female parent carries paralyzed hosts to prepared nests or caches. Normally the requirement that hosts are attacked and oviposited on *in situ* is included in the definition of a parasitoid.

More than 50% of parasitoids belong to the insect order Hymenoptera: the sawflies, ants, bees, and wasps. The least-derived suborder is that of the sawflies (Symphyta), nearly all of which are phytophagous, although one family, the Orussidae, contains species whose larvae are parasitoids. From a lineage that was probably related to the Orussidae, the remainder of the Hymenoptera (the Aprocrita) evolved, and these are in turn divided into a huge group the Parasitica or parasitoid wasps, which are almost exclusively parasitoids, and the smaller Aculeata in which the ovipositor has become a sting. The Aculeata, derived from the Parasitica, are primitively parasitoids that have radiated into a wide range of feeding habitats, including predators, scavengers, and pollen feeders. The Aculeata contain nearly all the Hymenoptera familiar to the general public, including all the species of ants, bees, and wasps that show advanced sociality (eusociality). The Parasitica are classified into a large number of families and superfamilies, the most important of which are the Ichneumonidae, Braconidae, and Chalcidoidea.

The second major group of parasitoids belongs to the two-winged flies or Diptera. The parasitoid habit has evolved on many occasions in this order, with the most important taxon being the family Tachinidae: these are almost exclusively parasitoids, with the adults frequently resembling houseflies (Muscidae, like the Tachinidae in the Calypterata). The parasitoid habit has also evolved on many occasions in the beetles (Coleoptera), although a relative small number of species feed in this way. The most frequently encountered beetle parasitoids are probably those in the rove beetle family (Staphylinidae). Outside the three insect orders just mentioned, parasitoids are extremely scarce and taxonomically scattered.

As described in more detail later, parasitoids are important members of nearly all terrestrial ecosystems, but there is little agreement about how many species exist. Currently, about 10% of the approximately 1 million described insects are parasitoids, but parasitoids are almost certainly less known compared with other groups—parasitoid taxonomy is legendarily difficult. Most workers today would argue that there are between 4 and 8 million species of insects on Earth, of which 1.5–2 million are parasitoids. To put this into perspective, recall that there are less than 10,000 species of birds and approximately 5000 species of mammals.

Life History Variation

A simple way to classify parasitoids is by the host stage that they attack. Holometabolous insects have three juvenile stages (egg, larva, and pupa), whereas hemimetabolous insects have nymphal stages with no pupal metamorphosis. The most common category of parasitoids is composed of those that attack larval or nymphal stages, although egg and pupal parasitoids are also common. The least common type is adult parasitoids, probably because the adult is best able to defend itself against parasitoid attack. Egg parasitoids deposit their own eggs within that of their host and are obviously very small insects. This life history has evolved several times but is most common in two families of chalcidoid wasp: the Trichogrammatidae and the Mymaridae. Members of the latter family, which have very narrow wings with pronounced cilia (fine hairs) and are known as fairy flies, include the smallest of all known insects. Some parasitoids oviposit into the eggs of their host, but their larvae only kill the host later when it is a larva or pupa. These are called egg–larval or egg–pupal parasitoids, whereas another common life history is found in larval–pupal parasitoids.

Most parasitoid eggs are deposited in or on the host, but a substantial minority (especially among the flies and beetles) is deposited in the vicinity of the hosts, and the final stages of host location are carried out by a specialized first-instar larva (a triangulum or planidial larva). There are two main modes of parasitoid larval development: as endoparasitoids or as ectoparasitoids. As their names suggest, the former feeds internally within the host and the latter externally. Occasionally, a parasitoid species may begin life as an ectoparasitoid and switch midway through development to an endoparasitoid or vice versa. Ectoparasitism is found most frequently in species that attack concealed hosts in plant tissue or other refuges, where the external feeding parasitoid has some protection from the environment and predators. Parasitoids can also be classified as idiobionts or koinobionts. Idiobionts kill or permanently paralyze their host at oviposition, whereas

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00152-0

koinobionts either do not or only temporarily paralyze the host so that it recovers and continues feeding. The parasitoid suspends its own development, normally as a first instar, and waits for the host to reach a size at which it can support its development, whereupon it resumes growth that results in host death. The koinobiont life history, normally found only in endoparasitoids, allows the adult parasitoid to attack hosts when they are still too small to support the full development of the progeny. However, while in a state of suspended development, the parasitoid larva is at risk from the host insect's immune system.

Part of the definition of a parasitoid is that its larva requires just a single host for development; however, it is possible for several parasitoids to share the same host. Species in which only a single parasitoid develops per host are termed solitary as opposed to gregarious parasitoids. The larvae of solitary parasitoids tend to have large mandibles which they use to attack other parasitoids in the host, whether conspecific or of a different species, and thus the difference between the two types of parasitoids is more than just one of degree. Occasionally, very large numbers of parasitoids emerge from a single host, for example, several hundred braconid wasps from a single hornworm (hawkmoth, Sphingidae) larva, although the largest broods are known to be associated with certain specialized wasps with polyembryonic development. In these species (particularly in certain genera of chalcidoid Encyrtidae), one or two eggs are laid into the host, but these divide asexually to produce many embryos—occasionally, hundreds and sometimes more than a thousand.

Unlike predation, in which feeding removes a prey item, a parasitized host may be discovered by a second parasitoid individual prior to its destruction. Sometimes, parasitoids avoid previously parasitized hosts, a practice called host discrimination, and this may be aided by chemical marks deposited by the first female. However, in many circumstances the second female will be selected to add its own eggs to the host. If the second individual is of the same species as that of the first, this is termed superparasitism, whereas if it is of a different species the term multiparasitism is used. Occasionally, a female may parasitize a host that she had previously attacked (self-superparasitism). Some, but not all, parasitoid species can distinguish between hosts previously parasitized by self and by other conspecifics. In superparasitism and multiparasitism, the parasitoid larvae compete with each other for host resources; when hyperparasitism occurs, one parasitoid larva actually feeds on the other. There are two main categories of hyperparasitism: obligate and facultative. As the name suggests, an obligate species can only develop as a parasitoid of a parasitoid, whereas a facultative species develops as a normal parasitoid when it encounters an unparasitized host and as a hyperparasitoid when it discovers a previously attacked host. Hyperparasitism is sometimes called secondary parasitism (as opposed to primary parasitism) and cases of tertiary and even facultative quaternary parasitism have been recorded. An unusual life history is found in some chalcidoid wasps of the family Aphelinidae: females develop as primary parasitoids, but males develop as secondary parasitoids of other aphelinids (including females of the same species).

Host Location

Parasitoids are small to tiny animals that search for hosts that are themselves small or tiny animals in a very complex environment. They have solved the formidable problems of host location by making use of a battery of clues and stimuli that reveal information about the presence or potential presence of hosts. The majority of these cues, especially those acting over long ranges, are chemical, but visual, tactile, thermal, and auditory information is also used. Moreover, the parasitoid is not a hard-wired automaton but constantly updates its estimates of the values of different stimuli to modulate its strategies for host location.

One can envisage a parasitoid as utilizing a hierarchy of stimuli with increasing information content. Perhaps the lowest-grade stimuli point to habitats or microhabitats that might contain hosts. Many parasitoids attacking plant-feeding hosts orientate toward green objects, whereas the parasitoids of *Drosophila* that feed on different types of fermenting substrates are attracted to volatile chemicals characteristic of their hosts' microhabitat. Plants damaged by herbivores release a cocktail of volatiles, which are often highly attractive to parasitoids. In some cases, parasitoids respond to artificially damaged hosts, whereas in other cases they are only attracted to the "host–host plant" complex. It has been shown that the plant's response to particular herbivores and their salivary and other secretions can be very specific, prompting the intriguing suggestion that the plant may have been selected to recruit "bodyguards" to protect itself against its herbivores.

Host waste products reveal valuable information about their originator's whereabouts. Many parasitoids are highly attracted to host frass (feces), while volatiles associated with salivary or mandibular gland secretions, or with silk, are also used in host location. Some hosts have symbiotic relationships with other organisms that may provide clues to their location: for example, wood-boring sawflies are associated with symbiotic fungi that digest the wood, and their parasitoids are strongly attracted to chemicals released by the fungi.

Clearly, hosts will be under strong pressure to reveal as little about their location as possible to searching parasitoids. Many parasitoids use vision in host location and camouflage and immobility may help protect hosts against parasitoids and more traditional predators. Vibration is particularly important for species attacking hosts feeding on concealed material, and some hosts "freeze" if they detect the presence of a searching parasitoid. There is evidence that parasitoids may also use thermal gradients to detect the presence of concealed hosts.

Hosts should maintain as far as possible the chemical equivalent of "radio silence," although sometimes hosts have to advertise themselves for different reasons and in doing so may provide clues to searching parasitoids. Some bark beetles (Scolytidae) can only overcome tree defenses by mass attacks of many hundreds or thousands of individuals, and these are coordinated by the beetles releasing an aggregation pheromone. However, not only other beetles but also their larval parasitoids are attracted to the pheromone. Similarly parasitoids may be sensitive to their host's sex pheromone. This is useful both for species that attack the adult host and those that use the presence of the adult as a cue for finding other life history stages. A particularly devilish instance of parasitoids

subverting their hosts' sexual signaling is that of the dipteran parasitoids of grasshoppers that home in on stridulating males.

Parasitoids are born with an innate hierarchy of responses to different stimuli, but they will change the position of cues in this hierarchy and incorporate new stimuli as they gain experience in host location. For example, if a novel volatile chemical such as vanilla is presented to the parasitoid at oviposition, the parasitoid learns to associate this chemical with hosts and subsequently flies to baits releasing this chemical. In a more natural setting, *Drosophila* parasitoids preferentially fly toward certain potential host microhabitats if they have previously been successful in finding hosts in that microhabitat. Many parasitoids, as they emerge from their pupa or cocoon, carefully examine with their antennae any remains of the host at the pupation site. They appear to be learning potential host cues because they subsequently fly toward sites containing the same volatile chemicals. An interesting corollary to this behavior is that it will tend to promote the formation of host races or biotypes.

Hosts are often found in patches that parasitoids exploit, but with decreasing returns as the fraction of unparasitized hosts declines. There will thus come a time when the parasitoid will be selected to abandon the current patch and search for a new site. Studying parasitoid patch use behavior has many parallels with predators searching for prey that are distributed in patches. This is one of the classical problems of optimal foraging theory, a field of behavioral ecology. In the simplest case, predators should leave the patch when their instantaneous rate of gain of food equals the maximum long-term rate that can be achieved in that environment. Longer patch residence is thus predicted in poorer environment, for example, when interpatch travel times are long.

The qualitative predictions of patch use theory have been supported by studies of parasitoids, but specific aspects of their searching behavior complicate the theory's application. For example, previously parasitized hosts do not disappear (as do eaten prey) but remain as potential sites for superparasitism. The parasitoid may also have to waste time in identifying them as parasitized, although they may provide useful information about host distributions. Parasitoid behavior may also change critically in the presence of conspecifics: when a parasitoid is searching a patch alone, any superparasitism is normally damaging self-superparasitism, which is not the case if other females are active at the site. Current research in this area using statistical models based on survival analysis seeks to identify the precise factors that determine when a parasitoid abandons the patch.

Host Acceptance and Oviposition Strategy

After a parasitoid locates a host, it then has to make a series of decisions that are very important for its future reproductive success. First, it must decide whether the host is of sufficient quality to be used as an oviposition site, or whether to search for a better host, perhaps using the current host as food, so as not to waste eggs. Given that the host is of sufficiently good quality, gregarious parasitoids have to decide how many eggs to lay, while hymenopterous parasitoids have to decide whether to lay male or female eggs. Only Hymenoptera have to make this last decision because they have a genetic system called haplodiploidy in which females develop from fertilized eggs and males from unfertilized eggs. By choosing whether or not to fertilize the egg, the parent can control the sex of its offspring. Unlike groups in which sex is determined by the random segregation of sex chromosomes, natural selection can operate on the behavior of haplodiploid females to produce sex ratios adapted to local conditions.

Before laying an egg, most parasitoids determine whether a potential host is suitable by examining it externally, and in the case of endoparasitoids often internally as well by using sensillae on their ovipositors. In either case, a host may be rejected if unsuitable, perhaps because it is of the wrong species. As part of biological control campaigns, parasitoids sometimes have to be mass reared, and much effort has gone into investigating whether they can be reared on artificial hosts. Critical to a successful strategy of this type is an understanding of exactly what host stimuli trigger oviposition.

Even if a host can potentially support a parasitoid larva, it may still not be worth laying an egg if the resulting offspring is of poor quality, perhaps very small in size. Understanding how natural selection molds host acceptance behaviors has many parallels with understanding how natural selection influences diet breadth in predators and other consumers, a second major strand of optimal foraging theory. Theory assumes that predators should maximize their rate of gain of energy (or other resources) and hence should not take "time out" from searching for good-quality prey items to attack poorer-quality types as the energy returns from the latter do not justify the time it takes to subdue and process them. Applying these ideas to parasitoids, searching females should not waste time ovipositing on poor-quality hosts when that time could be better spent searching for good hosts. Exactly what constitutes a poor host thus depends on the quality of the environment: attacking a particular host of medium quality might be economically advantageous in a habitat in which very good hosts are rare or absent, but it may not make sense in an environment in which the best hosts are abundant. There is experimental support for this prediction.

In many cases, the simple assumption that parasitoids should seek to maximize their rate of gain of reproductive success with time is probably overly simplistic. All parasitoids have a limited number of eggs, and maximizing the rate of gain of fitness may lead to the exhaustion of egg reserves prior to death. More sophisticated models have been developed that maximize reproductive success over the lifetime of the insect and predict that the parasitoid may become more selective as its egg supply diminishes. Again, there is experimental data supporting this prediction, although whether egg limitation is a major feature of parasitoid biology is a highly contentious topic. Another factor that may influence oviposition decisions is host feeding. Many adult parasitoids can feed from hosts, although this normally results in the host's destruction. Host feeding provides the parasitoid with energy and nutrients and hence allows it to produce more eggs. Thus, there is a trade-off: should the parasitoid use a host for reproduction and possibly risk later starvation or egg exhaustion, or should it forgo an immediate increment in fitness in the hope of deferred payback? In these circumstances, it clearly makes sense to feed from suboptimal hosts.

The question of optimal clutch size in parasitoids can be explored using similar economic arguments. In this case, the ideas originally developed to understand how natural selection operates on bird clutch size. An obvious prediction is that parasitoids should adjust their clutch size to the amount of resources available, and there is ample evidence that parasitoids do this, with the numbers of eggs laid typically being closely correlated to the size of the host. Careful experiments have shown that parasitoids can often estimate host size from very subtle cues; for example, some egg parasitoids can assess host size by very limited information about the local curvature of the host egg. As eggs are added to the host, the fitness returns from each egg typically decline because the members of the brood compete with each other for a fixed pot of resources. The optimal clutch size is defined as the number of eggs for which the marginal returns from adding additional eggs exactly equal the marginal opportunity costs in terms of wasted time or wasted eggs on the current host. This argument suggests that in good-quality environments in which fresh hosts are easy to discover, smaller clutches should be laid. Similarly, small clutches should be found in circumstances in which egg reserves are low. Experiments on gregarious parasitoids have supported both predictions.

Are solitary and gregarious parasitoids just part of a single continuum in clutch-size behavior? There are arguments that suggest this is not so: solitary parasitoids typically have large mandibles which they use to attack any other parasitoids in the host, including siblings, whereas gregarious parasitoids do not have this aggressive behavior. For evolution to convert a solitary parasitoid to a gregarious parasitoid, natural selection must operate not only on parental clutch size behavior but also on larval morphology and fighting. Although adult parasitoids and their offspring have many evolutionary interests in common, these interests are not identical and there is the potential for evolutionary parent–offspring conflict. Models incorporating this genetic conflict predict that the solitary life history is qualitatively distinct from gregariousness, and that there is hysteresis in the transition between the two states with gregariousness being far more difficult to evolve from the solitary state than vice versa.

The study of the sex ratio is one of the most successful areas of evolutionary ecology. The problem is to predict why some species of plants and animals produce sex ratios that differ from the 50:50 ratio. For a deviation to occur, there must not only be the selection of a biased sex ratio but also the flexibility to control son and daughter ratios. Haplodiploid parasitoid wasps have been critical in developing this field because they have both the incentive and the means to produce unusual sex ratios.

As Fisher demonstrated in the 1930s, a 50:50 sex ratio is expected to evolve because were the population to be biased toward females, sons would be a more efficient way of getting genes into future generations (because they would, on average, mate with more than one female), whereas if the population were to be biased toward males, daughters would be a more efficient way of getting genes into future generations (because sons would mate, on average, with less than one female). However, this argument assumes that the population is completely mixed and that all males compete for all females. This assumption breaks down if sons compete for mates, including their sisters. In such situations, the ovipositing female is selected to produce fewer sons to avoid wasteful within-family conflict and also to provide more potential mates for the male offspring. Hamilton, who in the late 1960s developed these ideas, called this process as local mate competition and marshaled a long list of examples in which mating among siblings was associated with female-biased sex ratios: 18 of the 26 examples he quoted were parasitoids.

Recent studies have shown that some parasitoids are able to assess the degree of local mate competition that their progeny is likely to experience and to adjust their sex ratios accordingly. For example, *Nasonia vitripennis* is a gregarious parasitoid of fly pupae (e.g., of blowflies that live in birds' nests). Wasps emerge in the bird's nest and mate prior to dispersal (the male has limited powers of flight). If a single female colonizes a bird's nest, all her offspring will mate among themselves and the wasp should produce just enough sons to mate all her daughters; if two females colonize a nest then local mate competition still occurs, although it is less intense, and a sex ratio of approximately 25% males is predicted. With greater numbers of females, the sex ratio quickly asymptotes at 50%. In laboratory experiments in which the number of insects searching together is manipulated, females respond to the presence of conspecifics by producing less female-biased sex ratios.

It had long been noted that parasitoid wasps often lay male eggs in small hosts and female eggs in large hosts, but why this evolved was only understood in the late 1970s by Charnov as part of a wider theory of environmentally correlated sex ratios. In parasitoids, there is normally a strong correlation between host size and the size of the adult wasp that eventually emerges from it. In general, wasps suffer from being small; they tend to have reduced longevity and fecundity and are less able to withstand the insults thrown at them by the environment. However, the negative consequences of being small are probably not experienced equally by the two sexes. In particular, female reproductive success may depend more critically on size than does male reproductive success for the simple reason that eggs are expensive and sperm cheap. This assumption is difficult to test because it requires size-dependent fitness to be measured in the field, which is difficult for such small animals, but the limited amount of available evidence supports this notion. If female fitness increases more rapidly with size than does male fitness, then Charnov's theory predicts that males should be laid in relatively small hosts and females in relatively large hosts. Moreover, there should be a sharp threshold in host size below which only males, and above which only females, are produced.

Laboratory experiments tend to show a threshold in host size as predicted by theory, although typically it is not as sharp as expected, perhaps because the wasp is using subtly different cues to assess host size from those used by the experimenter. To test whether relative or absolute size is important, many workers have presented medium-sized hosts to wasps in conjunction with either larger or smaller hosts. Theory predicts that the wasp should produce male eggs in medium-sized hosts when they are relatively small and female eggs when relatively large. Some wasps do indeed do this, although others have a much more fixed behavior, responding only to absolute host size. Possibly, these wasps always encounter the

same host distribution in the field and hence have never evolved the flexibility to deal with variable patterns of host size.

Resistance and Virulence

Endoparasitoids that delay their development while the host continues to feed and grow in size have to protect themselves from the host's defenses. The main host defense against parasitism is a cellular immune response called encapsulation. Cells (hemocytes) circulating in the blood cavity (hemocoel) recognize a parasitoid egg or larva as foreign and cause other cells to aggregate against the intruder, forming a capsule. The cells in the capsule fuse, and the whole structure becomes hardened and melanized, thus leading to the death of the parasitoid either through suffocation or through the release of necrotizing substances by the parasitoid. Encapsulation is not a specific antiparasitoid measure but rather operates against any foreign body.

There are two broad parasitoid strategies to counter host resistance: passive and active defense. In the first case, parasitoids either camouflage themselves so that they are not recognized as foreign or insert their eggs in specific organs away from circulating hemocytes. How this camouflage works can be seen in some species in which the egg is coated with a layer of protein prior to injection into the host. If the egg is dissected out of the wasp and the protein layer removed, the egg becomes susceptible to encapsulation if artificially "oviposited" into the host. Many parasitoids carefully position their eggs in the host brain or ganglia where they avoid hemocytes, while even eggs placed in the hemocoel may obtain some protection by sticking to fat body and other host organs.

The most common form of active defense is the injection of toxins at oviposition or their later secretion by the developing larva. It is common for the ovipositing female to temporarily paralyze the host so that it is easier to handle, but the same or other substances can also damage the host's immune system. Many parasitoids, though so far only Hymenoptera (particularly the Ichneumonoidea), inject viruses into the host at oviposition. The best-studied type is polydnavirus, so named (poly-DNA-virus) because the genetic material is composed of many separate DNA molecules within a single protein coat. The viral DNA is stably integrated within the wasp's genome, but in certain cells of the female reproductive tract it is copied, greatly amplified, and encapsidated. In the host, the virus does not replicate, but it does invade the hemocytes responsible for encapsulation which it can destroy by triggering apoptosis. How polydnaviruses evolved is not clear: in particular, are they part of the wasp's genome that has acquired virus-like properties or are they originally autonomous viruses that have been captured and tamed by the wasp? Instances are known of wasps injecting viruses that appear unrelated to polydnaviruses and also of wasps injecting virus-like particles that resemble the protein coats of viruses but contain no DNA.

Parasitoid eggs are surrounded by embryonic membranes which normally disintegrate after hatching. However, in some species, the cells of the membranes dissociate and float free in the hemocoel where they grow enormously in size. These cells are called teratocytes, and their primary role is most likely secretory (they are packed full of endoplasmic reticulum), producing substances that limit the host's immune defense. When the parasitoid is fully grown it consumes the teratocytes, which may thus also have a nutritive function.

Hosts and parasitoids almost certainly exert very strong selection pressures on each other, and observed levels of resistance and virulence will be determined by this coevolutionary interaction. Two major factors determining the outcome are the asymmetry in the interaction and the nature of the costs of resistance and virulence. The interaction is asymmetric in that every parasitoid has to overcome its host's defenses to survive, whereas not every host is parasitized. Models of the dynamics of host–parasitoid coevolution predict that in certain circumstances hosts may be selected not to defend themselves but to "gamble" on not being attacked. A major determinant of whether this strategy is favored is the nature of the costs to the host of investing in resistance mechanisms and to the parasitoid in investing in countermeasures. Costs are difficult to measure, but artificial selection experiments in the fruit fly, *Drosophila*, have shown that increased resistance to its parasitoids is accompanied by a decrease in competitive ability under conditions of resource stress. It appears that the fly switches limiting resources from optimizing feeding efficiency to immune function.

Associations with Microorganisms

The last 10 years have seen an enormous growth in our knowledge about associations between insects and microorganisms, particularly bacteria. This research programme has revealed a variety of interactions between parasitoids and microorganisms that are affecting our understanding of numerous aspects of their biology.

We now know that many types of bacteria are associated with insects and are transmitted vertically to their offspring, typically with some horizontal transmission occurring, though much more rarely. In order for a bacterium that is always, or nearly always, transmitted vertically to spread and maintain itself it must either increase the fitness of its host, or in some way manipulate its host's reproduction such that more daughters are produced by infected compared to uninfected females. It is this latter route that many symbiotic bacteria of parasitoids have taken.

The most common bacterial associate of parasitoids is *Wolbachia*, which is found in numerous species as well as in many nonparasitoids. The dominant means by which *Wolbachia* spreads is through cytoplasmic incompatibility. In a way not yet fully understood *Wolbachia* in males modifies the sperms such that they can only be used by infected females. Because uninfected females can only mate successfully with uninfected males while infected females can mate successfully with everyone, the latter have an advantage (at least when the infection is above a threshold density). Work on *Wolbachia* in parasitoids established much of the biology of the bacterium, as well as some of its possible consequences for population differentiation and speciation. Typically, researchers use antibiotics to "cure" species of *Wolbachia*, and then observe how

this affects mate use. Curiously, in *Asobara*, a species that attacks *Drosophila*, removing the *Wolbachia* leaves the wasp unable to reproduce. Recently, a second bacterium that also spreads through cytoplasmic incompatibility, *Cardinium*, has also been discovered in a parasitoid wasp.

Because parasitoid wasps have a haplodiploid genetic system in which males develop from unfertilized eggs, it is probably easier, compared with other organisms, for bacteria to spread by causing its host to reproduce parthenogenetically. This strategy has been adopted by a number of *Wolbachia* strains infecting different parasitoid species, especially the egg parasitoid *Trichogramma*. These findings validate several observations in the old biological control literature that heat shock could cause parthenogenetic species to produce the occasional male: the heat shock temporarily incapacitates the bacteria. Typically, curing the wasp of *Wolbachia* restores male production though in some, presumably very old associations, the ability to produce males seems to have been lost.

Finally, some bacteria spread by "male-killing"; destroying sons so that daughters benefit by having more resources. In parasitoids this was first observed in the gregarious wasp *Nasonia*, which oviposits a mixture of male and female eggs into a fly pupa. A bacterium called *Asenophonus nasoniae* causes the death of males leading to bigger and fitter females because they do not have to compete with their brothers. Spread through male-killing requires interactions among broods, and this is surprisingly widespread in insects with gregarious larvae.

As noted above, not only parasitoids but many other insects have symbiotic bacteria, and phylogenetic studies suggest that horizontal transfer must occur, if only occasionally. Frequently, the mode of horizontal transfer is a mystery, particularly as most bacteria are obligate intracellular parasites. It has often been suggested that parasitoids may vector bacteria between species, or may contract bacteria from their host, and there is some phylogenetic evidence for this in species infected by *Wolbachia*.

Lastly, symbiotic bacteria may also spread by providing benefits for their host. Virtually all aphids possess a primary symbiont that provides essential nutrients missing from its phloem diet. But many clones possess secondary, nonobligate symbionts. Recently, it has been shown that two facultative symbiont bacteria increase aphid resistance to parasitoids, although how they do so is not yet known.

Population Dynamics

The study of host–parasitoid population dynamics has received considerable attention for two main reasons. First, parasitoids are important biological control agents and understanding how they may regulate their host has clear applied relevance. Second, host–parasitoid interactions are a useful model system for answering broader questions in population biology applicable to all resource–consumer interactions. Host–parasitoid interactions are attractive as model systems because of the simplicity of the trophic relationship: in most cases, a parasitoid attack results in one (solitary species) or a certain number (gregarious species) of recruits to the next generation of parasitoids. This is in contrast with predation, in which it is far more difficult to predict how a single instance of prey capture influences the future number of predators.

Modern consideration of host–parasitoid population dynamics began in the 1930s, particularly through the work of the Australian ecologist Nicholson and his collaboration with the mathematician Bailey. They considered the case of a host–parasitoid interaction with discrete, synchronized generations. Hosts are exposed to parasitoids during a fixed developmental stage, and those that survive parasitism and other sources of mortality produce the next generation of hosts the following years. Parasitized hosts are the source of the subsequent parasitoid generations. In their basic formulation, known as the Nicholson–Bailey model, they assumed that parasitoids searched randomly in the environment. The average number of times a host is encountered is then simply a product of parasitoid density (P) and a constant (a) known as the attack coefficient. Provided that the parasitoid search is random, the probability of escaping parasitism is simply the zero term of a Poisson distribution with this mean (i.e., e^{-aP}).

The Nicholson–Bailey model is important because it encapsulates the minimum bare bones of a host–parasitoid interaction. It tells us what to expect in the absence of any biological complications, and what it tells us is perhaps surprising. If the Nicholson–Bailey model is iterated over many generations, the densities of hosts and parasitoids oscillate with ever-increasing amplitude until one or both species go extinct. In essence, the parasitoid overexploits the host, which crashes to low densities, and this is followed by a drastic decline in the parasitoid number. The host then recovers and in the temporary absence of the parasitoid increases to higher densities than before; this is followed, after a lag, by a huge growth in parasitoid numbers and even more cataclysmic overexploitation. The combination of random search and the time lags inherent in this discrete-generation framework does not allow a host–parasitoid interaction to persist.

The intrinsic instability of the Nicholson–Bailey framework has set the agenda for much of host–parasitoid population dynamics in the past 60 years: what aspects of real host–parasitoid interactions allow the persistence that is so palpably clear in the field? Introducing a realistic functional response (a limit on the number of hosts a parasitoid can attack due to egg or time limitation) does not help, and if anything, it makes nonpersistence more likely. In the 1960s and 1970s, many workers thought that interference between parasitoids was the key. Interference is said to occur if female parasitoids disrupt each other's searching, perhaps because when two parasitoids meet they fight or in other ways interact. Such interference might reduce the efficiency of high-density parasitoid populations and break the cycle of overexploitation and recovery. However, although interference occurs and can be measured in the laboratory, a consensus soon formed that in the field it was unlikely to be strong enough to allow persistence. Another idea was to assume that in addition to parasitism, hosts were also subject to direct density dependence due to competition for food. This additional density dependence can stabilize the Nicholson–Bailey interaction and may very well be important in some host–parasitoid interactions in the field. However, it is unlikely to be the complete answer because at the stable population densities that are predicted by this model, there is still substantial

competition for food. This theory can neither explain why so many hosts are regulated at levels well below their food-carrying capacity nor why parasitoids can cause such dramatic reductions in host density after their release in biological control programs.

Most workers today believe that host–parasitoid persistence involves a general phenomenon called heterogeneity of risk (this idea actually dates back to the last papers of Nicholson and Bailey in the 1960s). Overexploitation of the host population occurs because all hosts are equally susceptible to parasitoid attack, and therefore nearly everyone succumbs when parasitoid densities are high. However, suppose that some hosts are in a refuge and protected from parasitoids. These individuals can survive periods of high parasitoid densities and prevent the crash in host density that initiates the diverging oscillations (of course, if too many hosts are in a refuge then the parasitoid will not be able to prevent the host population from increasing without restraint).

What might these refuges be? There are many possibilities. First, there may be physical refuges in which hosts are protected from parasitism. For example, the host might be a stem borer and a certain fraction of hosts may burrow so deeply in the plant tissues that they are inaccessible to parasitoids. Second, a certain fraction of the hosts may be physiologically immune to parasitoid attack. Third, there may be a phenological mismatch between host and parasitoid: some hosts may avoid parasitoids by emerging either sufficiently early or late in the season so that they miss all or most searching parasitoids.

Finally, and probably the most important, the refuge may not be physical, physiological, or phenological but statistical: a certain number of hosts, more than would occur if the parasitoid search was random, may escape parasitism by chance each year. For example, let us suppose that hosts are distributed across the environment in patches, and that certain patches are more attractive than others to parasitoids (for reasons not associated with host densities, some patches may be in areas that are more likely to be encountered by parasitoids). Provided enough hosts occur in rarely encountered patches—statistical refuges—the interaction will be stable. (A popular way to model this is to replace the Poisson distribution in the Nicholson–Bailey model with a different, more aggregated statistical distribution such as the negative binomial. The resulting interaction is stable provided the variance of the distribution is sufficiently large.) All these ways of stabilizing the Nicholson–Bailey model rely on there being differences among individuals in the risk of parasitism; hence their general labeling as models of heterogeneity of risk.

Recent developments in host–parasitoid population dynamics have largely involved the study of interactions with overlapping rather than discrete, synchronized generations and interactions in which the spatial distribution of the population needs to be taken into account. Models for species with overlapping generations typically abandon the discrete time difference equations of the Nicholson–Bailey model and use instead differential equations, although they often incorporate time delays to represent developmental lags. Two new insights have come from this approach. First, an interaction that is otherwise identical to the Nicholson–Bailey model may be persistent if there is a long-lived, invulnerable host stage. The reason for this relates to the arguments regarding the heterogeneity of risk discussed previously: the long-lived invulnerable stage provides a refuge that allows the host population to ride out periods of very high parasitoid densities.

The second insight to emerge concerns when hosts and parasitoids differ in generation time, in particular, when the parasitoid generation time is approximately half that of the host. The population can then display persistent cycles with a period approximately equal to the host generation time. This is of more than technical interest because such cycles are common among pests of tropical plantation trees such as coconut, oil palm, and cocoa, and these pests tend to be attacked by parasitoids with shorter generation times. The cycles arise because of dynamic interference between the two time lags in the system. Suppose there is a temporary increase in the densities of host: the progeny of this "blip" will form a secondary blip one host generation later. However, the parasitoids will also be able to take advantage of the increase in host population density, and this will result in a secondary increase in their densities one parasitoid generation later. If the two generation times are equal, the two secondary blips coincide (the Nicholson–Bailey situation), but if the parasitoid generation time is approximately half that of the host the two blips do not coincide, and the parasitoid causes a trough in host density, an "antiblip," half a host generation later.

There are two main ways to incorporate a spatial component into host–parasitoid models. The first is to assume that the system consists of discrete subsystems linked by migration, whereas the second is explicitly to incorporate space and write equations that describe how host and parasitoid densities change at all spatial coordinates. Species that have isolated populations linked by migration are normally described as having a metapopulation structure. Metapopulations are an area of very active research in population biology, although studies of interacting metapopulations, such as that of a parasitoid and its host, are still rare. The most interesting aspect about host–parasitoid metapopulations is that the interaction they describe can persist, even if every component population had been nonpersistent in isolation. For this to happen, component populations must fluctuate out of synchrony so that one population doomed to extinction can be rescued by immigration from a different population in a different phase of the host–parasitoid cycle. Random, local effects that promote asynchrony favor the persistence of a metapopulation, whereas regional fluctuations in characteristics such as climate will tend to synchronize populations and hence act against persistence.

In fully spatial models, the rescue effect can still operate, but now more elaborate spatial processes can also be observed. For example, a wave of hosts can spread through the environment, chased by a following wave of parasitoids. The hosts escape parasitoids by moving into uncolonized areas ahead of the advancing waves, whereas the parasitoid destroys the host population behind the wave front and would itself be destroyed were it not able to chase the host population. These waves tend to form spiral patterns, a feature common to many spatial interactions in population dynamics, physiology, and even inorganic chemistry. For certain parameter combinations, the organized spiral waves break down into a seemingly random jumble of interactions and short-lived wave fronts called spatial chaos. Experimental studies of spatially extended parasitoid populations are still somewhat behind theory and

one of the greatest challenges of contemporary host–parasitoid population dynamics is to devise ways of studying spatial dynamics in the field.

Community Ecology

Few, if any, host–parasitoid interactions exist in isolation: most parasitoids attack more than one species of host, and the majority of hosts is attacked by several species of parasitoid. How does one begin to understand the workings of a large community of interacting hosts and parasitoids? One approach is to build up from the simple one-host, one-parasitoid interactions discussed previously. Alternatively, one can survey the properties of real communities and try to deduce patterns that may reveal how they are structured.

The simplest possible community that is more complex than those discussed previously consists of one host and two parasitoids, or two hosts attacked by a common parasitoid. Both tell us interesting things. In the first case, the simplest models predict that one of the two parasitoids will invariably go extinct. This is a corollary of a basic finding in population ecology that identical species cannot coexist on the same resource—the competitive exclusion principle. For coexistence to occur, other factors must be brought into play. In particular, coexistence can occur if the ecological niche (i.e., the host) is divided in such a way that the parasitoids are no longer identical. Thus, one parasitoid may use the host at one time of year or in one microhabitat, or it may specialize on a different developmental stage of the host from the other. Coexistence can also occur by a more statistical route if hosts differ in their susceptibility to attack by parasitoids. If some hosts happen to be protected from one species of parasitoid by being in the "statistical refuge" discussed above, and if others are independently protected from a second parasitoid species, then this can promote coexistence. Spatial processes can also promote coexistence. Consider two species of parasitoid, one of which is always the winner in a straight competition at a single locality, while the other has a superior dispersal capability. Coexistence is now possible in an environment in which new host populations are constantly being generated because the poorer competitor is able to find unexploited hosts: this is known as a competition–colonization trade-off.

I now discuss the two-host, one-parasitoid case. Again, the simple models predict extinction, this time of one of the host populations. This occurs because the equilibrium population of parasitoids maintained on the surviving host species is sufficient to prevent the second host species from replacing itself. A recent laboratory experiment cleverly illustrated this using two flour moths attacked by the same species of wasp. This is a similar situation to competitive exclusion, but the trophic structure is reversed: instead of two species competing for the same resource, two species are subject to the same natural enemy. In fact, there are deep biological and mathematical symmetries between these two cases, and this has led to the phenomenon of two species interacting through a common natural enemy being referred to as apparent competition to stress the parallels with direct competition, in which two species interact through a common resource. Again, for a two-host, one-parasitoid system to persist, something extra must be added to the simplest models: the two hosts might be present at different times of the year or might be spatially segregated, or the parasitoids might preferentially exploit the most abundant host (behavioral switching).

Much of traditional community ecology has stressed how biological communities may be structured by resource competition, and thus it does not apply to many insect communities in which the majority of species feed on different host plants and thus never come into contact. Apparent competition is significant because it can at least potentially structure a community in exactly the same way as direct competition does. The degree to which this actually occurs is a major theme in current insect ecology.

Progressing from models of three species to those with larger numbers of components becomes increasingly difficult. There are two problems: first, more assumptions have to be made about how different species interact and about the values of large numbers of parameters. Seldom are there field data to reduce this burden of supposition. Second, the dynamic behavior of larger communities becomes increasingly more complicated. For example, a model of a five-species community consisting of two hosts, two specialist parasitoids, and a generalist parasitoid showed the same range of population dynamic behaviors as those exhibited by simpler communities. However, it also showed more complex behaviors in which the full five-species community was unstable with one or more species going extinct, but with the resultant smaller communities, after they had reached equilibria, then being susceptible to invasion by the species that had recently gone extinct. Other multispecies models have shown complex chaotic dynamics. Currently, it is unclear regarding the extent to which the bottom-up approach, modeling explicitly the dynamics of each member of a large community, is a feasible way to approach parasitoid community ecology.

The top-down approach to community ecology consists of searching for patterns in multispecies assemblages that provide evidence of structuring forces. For example, workers have searched for patterns in the number of species of parasitoids attacking different species of hosts. The attraction of this approach is that there are numerous studies in the literature providing information on the parasitoid complexes of different insects. The major result to emerge from these studies is that the host-feeding niche influences parasitoid species numbers. Leaf-mining insects are attacked by the largest number of species, with successively smaller numbers attacking more concealed hosts (gall formers, shoot borers, and root borers) and less-concealed hosts (species living in leaf rolls and ties and those living externally like typical caterpillars). There are two explanations for this pattern. One suggests that the number of parasitoid species that can coexist on a single host is influenced by the fraction of that host's population that inhabits a refuge from (all) parasitoid attack. Proponents of this view argue for a correlation between feeding niche and the size of the refuge. Alternatively, many species are likely to occur on hosts that are taxonomically and ecologically similar to other hosts because there is less of an evolutionary hurdle to overcome in incorporating a new species in a parasitoid's host range. Leaf-mining insects have a far more uniform ecology than insects feeding in other host niches. They are also relatively

taxonomically homogeneous, and this may have facilitated host transfer and broad parasitoid host ranges.

There are less data in the literature on the host range of parasitoids because this requires all potential hosts in an area to be surveyed. However, the data that are available support one major conclusion: idiobionts have broader host ranges than koinobionts. Recall that idiobionts kill or permanently paralyze their host at oviposition, whereas koinobionts delay development until the host is fully grown. During this suspension of growth, the parasitoid has to cope with the attack by the host's immune system, and the need to evolve to be finely attuned to the host to counteract its defenses probably limits the koinobiont host range.

A different top-down approach to parasitoid community ecology has recently been taken by a few groups, although it is still too soon to assess its value. There is a long tradition of constructing food webs in ecology, and one of the aims of this research program is to search for patterns that are common across different webs. A major limitation of this research is the heterogeneity in published food webs, which are typically collected in very different ways and which differ greatly in their taxonomic resolution. Host–parasitoid interactions have many advantages for food web studies, the most prominent of which is the relative ease with which trophic links can be identified and quantified. This has led to the construction of several quantitative food webs in which all hosts, parasitoids, and links are expressed in the same units. The webs published to date illustrate the extent to which different hosts are linked by shared natural enemies and also the extent to which indirect effects such as apparent competition may act as forces structuring the community. Of course, a limitation of this approach is that only one guild of natural enemies is included in the web, but it will be interesting to determine whether common patterns emerge as more communities are studied.

The Importance of Parasitoids

Why are parasitoids important to students of biodiversity? First, they are abundant members of nearly all terrestrial communities. The populations of few herbivorous insects are unaffected by parasitoid attack, and understanding how they influence and possibly regulate the densities of their host is immensely important in understanding the mechanisms underlying biodiversity. Much of the modern work in parasitoid ecology has been motivated by these considerations.

The second reason why parasitoids are important and why there is such an enormous and valuable literature on parasitoids dating back nearly a century is that parasitoids are our allies in the war against agricultural and forestry pests. There are two main strategies of parasitoid biological control: classical and inundative. In classical biological control, a parasitoid is released into an area to control a pest. Typically, the pest is a nonnative species that has been accidentally introduced into a new region, and normally the parasitoid is a species that attacks the pest in its region of origin. The parasitoid is released once or on a few occasions until it is established. Although only 10–20% of biological attempts are successful, they can have dramatic effects in controlling pests, and the savings in crop yield are normally many multiples of the investment in the biological control program. There are interesting patterns in the probability of success against different hosts, for example, biological control has an excellent record against scale insects and mealy bugs but is far less successful against aphids.

The second strategy is inundative release in which large numbers of parasitoids are released onto a crop to destroy a pest. Applications are made as often as required, and long-term establishment of the parasitoid (which is normally already there at low density) is not a goal. In effect, the parasitoid is acting here as a biological insecticide. There are two main situations in which this strategy is used today. The first is when labor is relatively cheap so that the costs of mass rearing the parasitoids are not too great. In many tropical countries, huge numbers of egg parasitoids are mass reared and released to attack cutworms, armyworms, and bollworms. Advances in the mechanization of mass rearing are making inundative release more attractive in developed, high-wage countries. The second situation concerns high-value crops in enclosed growing conditions such as greenhouses. Control of tomato whitefly is typically by parasitoids in northern Europe. Inundative control has been favored in these situations because of the decreased tolerance of insecticide residues by regulatory authorities and because of the general need to reduce chemical input so as not to disrupt insect pollination and the existing successful biological control of other pests.

In the post-"Silent Spring" era, biological control has often been viewed as an unmitigated good (particularly by biocontrol workers) because it does not involve pesticides. However, in recent years, increasing attention has been paid to the dangers of biological invasions, particularly by plants, and this has led to a reassessment of biological control and to worries regarding the effect of introduced parasitoids on native insects. These are important issues and need to be carefully scrutinized, although there is a real danger of an overreaction and a return to the automatic reliance on environmentally harmful chemical insecticides. Currently, the major problem is a lack of good data on the consequences of biological control introductions. The challenge to applied ecological entomologists is to develop a sufficient understanding of parasitoid population and community ecology building on the ideas described previously, so that they can provide rules and guidelines that will ensure that biological control has a net positive rather than negative effect on biodiversity.

See also: Hymenoptera. Parasitism. Species Interactions

References

Godfray HCJ (1994) *Parasitoids, Behavioral and Evolutionary Ecology.* Princeton, NJ: Princeton University Press.

Hassell MP (2000) *Insect Parasitoid Population Dynamics.* Oxford: Oxford University Press.

Hawkins BA (1994) *Pattern & Process in Host–Parasitoid Interactions.* Cambridge, UK: Cambridge University Press.

Quicke D (1997) *Parasitic Wasps.* London: Wiley.

Waage JK and Greathead D (eds.) (1986) *Insect Parasitoids.* London: Academic Press.

Pelagic Ecosystems

Andrea Belgrano, Institute of Marine Research, Lysekil, Sweden
Sonia D Batten and Philip C Reid, Sir Alister Hardy Foundation for Ocean Science, Plymouth, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Calanoid copepods These belong to the Crustacea and are very abundant in the zooplankton. They play a major ecological role in the food web of the pelagic ecosystem as grazers of phytoplankton and as a food source for, for example, larval and adult fish.

Clupeid fish Pelagic fish, including the anchovy, sardine, and herring, which feed on plankton.

El Niño Southern Oscillation (ENSO) Regarded as quasiperiodic fluctuations occurring in the equatorial region of the Pacific Ocean. The ENSO is associated with changes in the sea surface temperature, sea surface levels, and rainfall patterns in the tropics and with a possible influence on the weather at higher latitudes. ENSO events have also been associated with changes in fisheries and ecosystem processes.

Gadoid fish Includes the cod Gadus morhua L.; the life cycle of cod can be summarized as four stages: spawning, larvae, juveniles, and adults. During the larval and early juvenile phases they feed largely on copepods on the nursery grounds; after this phase they become largely stationary and feed on benthic or epibenthic animals.

North Atlantic Oscillation (NAO) The difference in atmospheric pressure between Ponta Delgada, Azores, and Stykkisholmur, Iceland. The variations in the NAO are usually associated with a positive or negative phase related to changes in the direction and strength of the westerly wind as well as in sea surface temperature. During a positive phase, winters over Scandinavia are warmer and vice versa. Recently, the fluctuations observed in the NAO have been related to changes in ecosystem processes.

Phytoplankton Unicellular photoautotrophic plant cells ranging in size from 1 μm to 1 mm. They are divided into 12 taxonomically defined divisions, including diatoms, dinoflagellates, and coccolithophorids.

Primary production An estimate of the rate of carbon fixation by phytoplankton; this can account for up to 75% of the photosynthetic production on Earth.

Russell cycle Named after F. Russell, the Russell cycle describes a major change in the biota of the English Channel from 1925 to 1972.

Zooplankton From the Greek zoon meaning animal, zooplankton comprises those animals that are found passively drifting or weakly swimming in the water column. Zooplankton can be divided into two major categories: holoplankton, which are organisms that spend their entire lives as plankton, and meroplankton, which are organisms that spend part of their life cycle as plankton and part on the seafloor as benthic invertebrate larvae or as nekton (e.g., fish larvae).

Phytoplankton and Primary Production

The etymology of the word "plankton" derives from the ancient Greek meaning wandering. Plankton have very limited motility and are dependent on currents and the physical environment for their location. Phytoplankton comprises a very diverse group of single-celled photoautotrophic planktonic plants that are divided into 12 taxonomic divisions (3500–4500 species of oceanic plankton) and live in the surface water of the oceans. The role played by phytoplankton is at the base of the biogeochemical cycles in the ocean. Species range in size from small prokaryotic and eukaryotic cells comparable to bacteria to larger organisms such as diatoms, dinoflagellates, and coccolithophorids (**Figures 1** and **2**). A typical phytoplankton community contains a mixture of these, although the species which are dominant vary through time. The majority of the species present in the ocean gain their energy *via* photosynthesis and, therefore, environmental factors such as light availability, temperature, and the supply of major nutrients in the form of ions are important factors that influence their distribution on both temporal and spatial scales. Several hypotheses have been proposed which explain the coexistence of several species; one, "temporal succession," supposes that the nutrient regime of the environment changes rapidly so that all species are in different states of approaching to or declining from their maximum abundance. In low-turbulence environments microhabitats may develop which favor the growth of a distinct patch of plankton. Alternatively, different species may be limited in growth by the availability of different nutrients so that coexistence is possible. In neritic temperate waters of the North Sea and in the open temperate North Atlantic, the mean seasonal pattern of phytoplankton abundance shows a distinct peak in the spring, followed by a summer decline and a second, lesser, peak in the autumn. Work by Colebrook (1986) showed a clear differentiation between those species associated with the spring bloom, which tended to be diatoms, and those associated with the autumn bloom, which tended to be dinoflagellates. There were changes in the North Sea in terms of phytoplankton species composition after 1987, for example, *Ceratium* spp. (Dickson *et al.*, 1992) and Thalassiothrix longissima (Reid *et al.*, 1992), which signaled a change in the ecosystem. These changes have been related to an increase in the sea surface temperature (SST) and to changes in the extension of the Baltic/Norwegian waters in the North Sea and an increased input of more oceanic water. Variations in the seasonal cycle are caused by different preferences of the groups for light intensity, water stability, temperature, and nutrient availability.

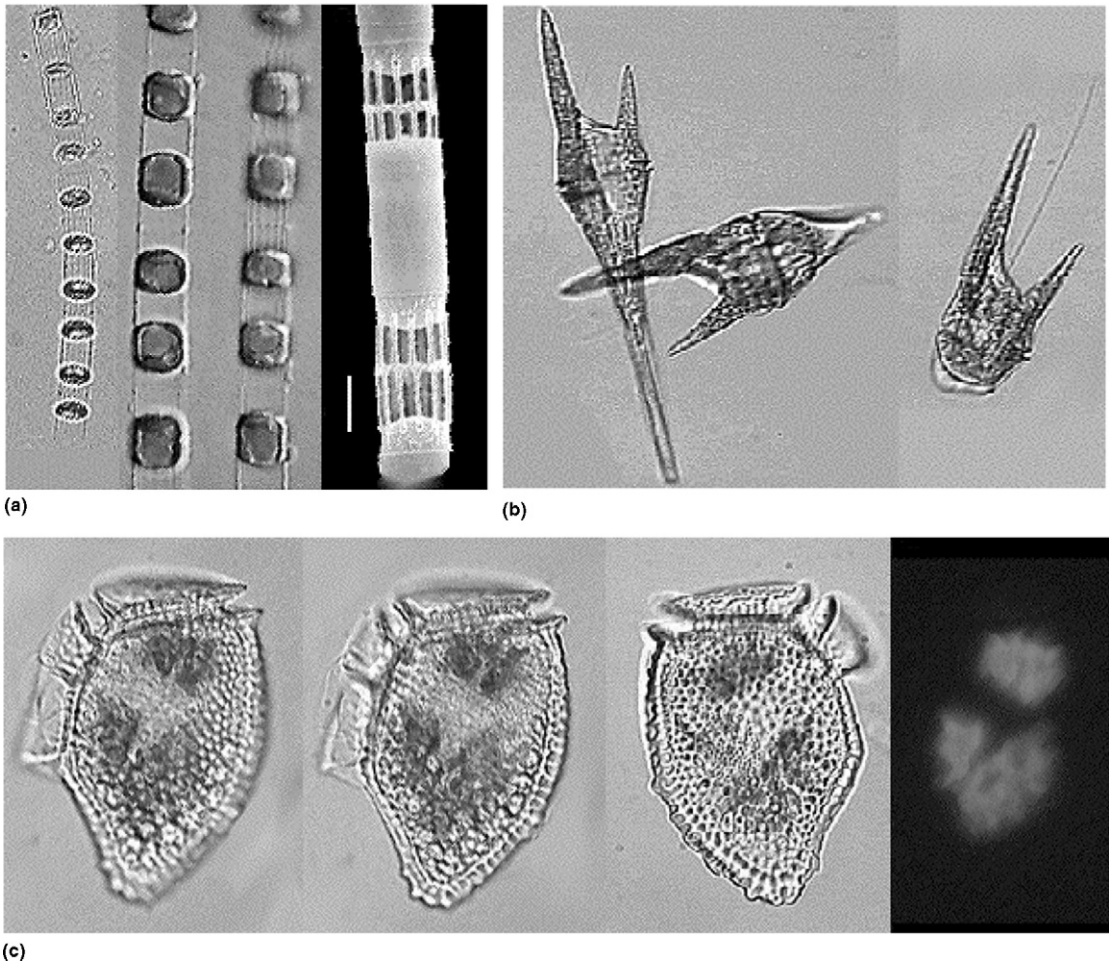


Figure 1 Phytoplankton species illustrating the taxonomic diversity as well as the shape and size of these organisms, which are an important component of the pelagic food web. (a) Diatomophyceae, *Skeletonema costatum*, a worldwide diatom species with distinctive features such as long chains of small cells with long external tubes. Diameter varies between 2 and 21 μm. (b) Dinophyceae, *Ceratium furca*, a worldwide dinoflagellate with a solitary or paired life-form, with a straight body, and the epitheca gradually tapers into an anterior horn. The size ranges in length between 70 and 200 μm and in width between 30 and 50 μm. (c) Dinophyceae, *Dinophysis norvegica*, a solitary dinoflagellate, normally found in cold waters, varies in size with length between 48 and 67 μm and width between 39 and 53 μm. Dinophysis norvegica can be regarded as a toxic species since it produces a toxin that causes diarrohetic shellfish poisoning and therefore it is a potentially dangerous species when found in coastal waters in large numbers. (Photos courtesy of Mats Kuylenstierna and Bengt Karlson).

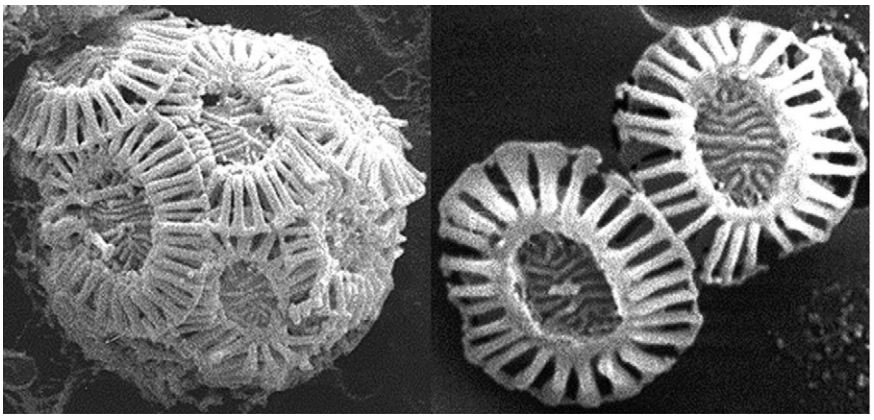


Figure 2 A (SEM) photo of a Haptophyta coccolithophorid, *Emiliana huxleyi*, found worldwide; the diameter varies between 5 and 10 μm, and the coccoliths are 3.5×3 μm. (Photos courtesy of Mats Kuylenstierna and Bengt Karlson).

A bimodal pattern of the seasonal cycle is evident across the temperate North Atlantic but other patterns exist in different geographical regions. Polar regions show a single summer peak because light intensity is only sufficient for a short period of phytoplankton growth. The north Pacific Ocean shows no spring peak, and in fact there is little change in the phytoplankton standing stock throughout the year. The principal reason is that the dominant zooplankton (the copepod genus Neocalanus) overwinter as adults which produce young in the spring without feeding. The young are then able to immediately take advantage of any increase in the phytoplankton and effectively prevent a spring bloom. The tropics show little seasonality and both phytoplankton and zooplankton fluctuate slightly throughout the year, dependent on local processes. Large blooms of phytoplankton can be monitored *via* satellite remote sensing (e.g., the coccolithophorid *Emiliana huxleyi*). They have the potential to modify, to some extent, the carbon exchange taking place between the ocean and the atmosphere. **Figure 3** shows an example of a cyanobacterium, *Synechococcus* sp., which is an extremely small photoautotrophic cell 1–3 μm in length. Picocyanobacteria and nanoflagellates are responsible for assimilating excretory products of zooplankton consumers in the euphotic zone in nutrient-poor waters characteristic of the central ocean gyres. As noted by Falkowski *et al.* (1998), the fixation of carbon by phytoplankton results in approximately 45 gigatons of organic carbon per year, of which approximately 16 gigatons is exported out of the surface waters. Changes in the total and export production can affect the marine food web structure, ultimately leading to changes in the fish stocks. The photosynthetic pigments such as chlorophyll *a* can be used as a proxy measurement of phytoplankton biomass, and with the aid of satellite-derived measurements a color index can be derived to illustrate changes in primary production. Behrenfeld and Falkowski (1997) generated several maps of primary production estimates for the world's ocean (**Figure 4** and **Figure 5**). Measurements of phytoplankton color on samples taken by the Continuous plankton recorder (CPR) survey from the central northeast Atlantic and North Sea (**Figure 6**) shows a similar increasing trend and longer growing season as that of the satellite-derived vegetation index and this trend dates back to 1948. The phytoplankton increase is not ubiquitous because an inverse trend is evident in eastern Atlantic waters between latitudes 59° N and 63° N. This index of phytoplankton change may reflect a possible long-term



Figure 3 An example of a cynobacteria species, Synechococcus, found in the central Skagerrak; the size range is between 1 and 3 μm in length. (Photos courtesy of Mats Kuylenstierna and Bengt Karlson).

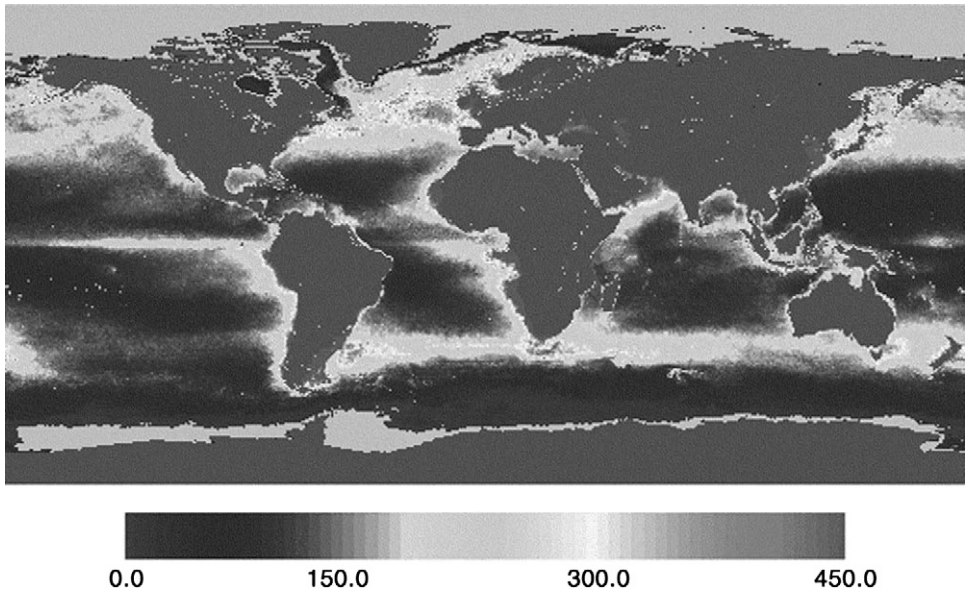


Figure 4 A map of annual primary production (PP) generated for the standard parmeterization of the VGPM, where the PoptB is modeled according to Behrenfeld and Falkowski (1997) and the surface irradiance is corrected for cloudiness (map courtesy of Paul G. Falkowski).

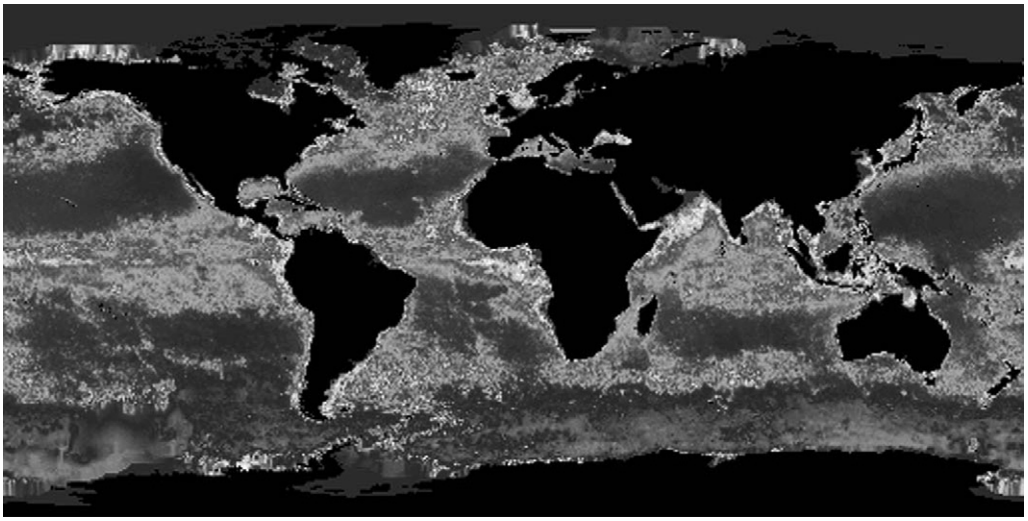


Figure 5 A map of primary production generated when PoptB is estimated using the relationship between temperature and the maximum phytoplankton specific growth rate and modified for photosynthesis (map courtesy of Paul G. Falkowski).

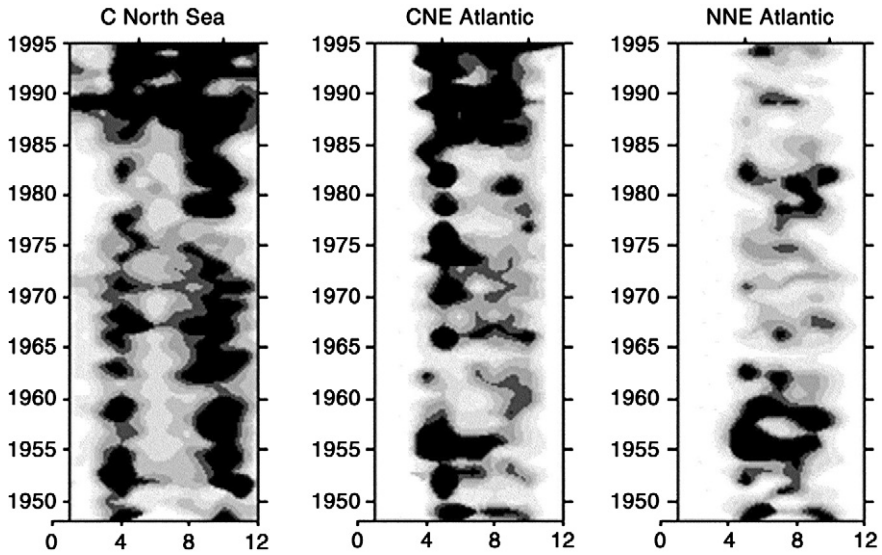


Figure 6 Phytoplankton color distribution from the CPR survey from 1948 to 1995 showing an increase in the late 1980s in the central part of the North Sea and in the central northeast Alantic coinciding with a change in the NAO index from a negative to a positive phase. A reverse trend was observed for the north-northeast Atlantic (redrawn from Reid *et al.*, 1998, courtesy of SAHFOS, with permission from Nature 391, 546 (1998)).

biological response in the North Atlantic to increasing global temperatures which are causing melting of ice, sea ice, and permafrost in the Arctic (Reid *et al.*, 1998).

In the central North Pacific environment, Venrick (1990) observed 245 species of phytoplankton in the shallow waters and 231 in the deep zone during the period between 1972 and 1985. In both areas, 21 species accounted for 90% of the individuals and most of the variance in abundance occurred at small spatiotemporal scales and samples within a zone were similar. This study showed that epiplanktonic populations of the North Pacific central environment were more stable when compared with planktonic populations from other ecosystems. In relation to the pelagic food web structure, phytoplankton is an important food source for the larval stage of pelagic fish and thus plays an important role in the trophic structure of pelagic communities.

Zooplankton

Zooplankton are the secondary producers in pelagic ecosystems and comprise an extraordinarily wide range of organisms. The zooplankton community of continental shelf waters, for example, may contain larval stages of littoral and benthic invertebrates (meroplankton) in addition to the species that spend all their lives in the plankton (holoplankton). Phytoplankton production in these nutrient-rich waters is generally regulated by the grazing activity of zooplankton

(Figure 7). Most zooplankton belong to the crustacean, with the principal organism group being copepods. Two copepod species, *Calanus finmarchicus* and *C. helgolandicus*, constitute the major components of the northeast Atlantic and North Sea zooplankton in terms of biomass, abundance, and trophic role (Marshall and Orr, 1972). These populations follow distinct spatial and seasonal dynamics such that *C. finmarchicus* is a northern spring species, whereas *C. helgolandicus* is located in warm-temperate waters and reaches its maximum abundance in autumn. The distribution of these species of Calanus has been studied during the period 1958–1995 in the North Atlantic (Figures 8–10) in relation to the year-to-year variability in the North Atlantic Oscillation (NAO). The NAO is a climatic oscillation affecting the hydrology and climate in the North Atlantic Ocean and adjacent region, and it can be compared to El Niño in the Pacific Ocean. After 1987, the zooplankton community in the North Sea changed with, for example, an increase in the abundance of the copepod *Corycaeus* spp. and it may have also affected the distribution and abundance of the two Calanus species. The increase in temperature, wind, and alteration of the winter circulation pattern observed during years of a high positive NAO index have resulted in unfavorable conditions for the *C. finmarchicus* population leading to a significant decrease in the abundance of the species. Conversely, these hydro-climatic shifts have proved beneficial to *C. helgolandicus*, the abundance of which has increased during these years (Planque and Taylor, 1998; Stephens *et al.*, 1998).



Figure 7 An example of copepod species found in the zooplankton; two Acartia tonsa copepods with ingested food item *Thalassiosira weissflogii* (photos courtesy of Kajsa Tönnesson).

In the northeast Pacific the plankton community structure (Figure 11) responded to the warming in 1958 to 1960 in the California Current in relation to an El Niño event that caused a sudden change in the SST distribution (McGowan *et al.*, 1998). The zooplankton abundance has generally decreased in the California Current system while, in a synchronous manner, the zooplankton abundance in the Gulf of Alaska gyre increased. These interdecadal regime shifts in these two systems driven by

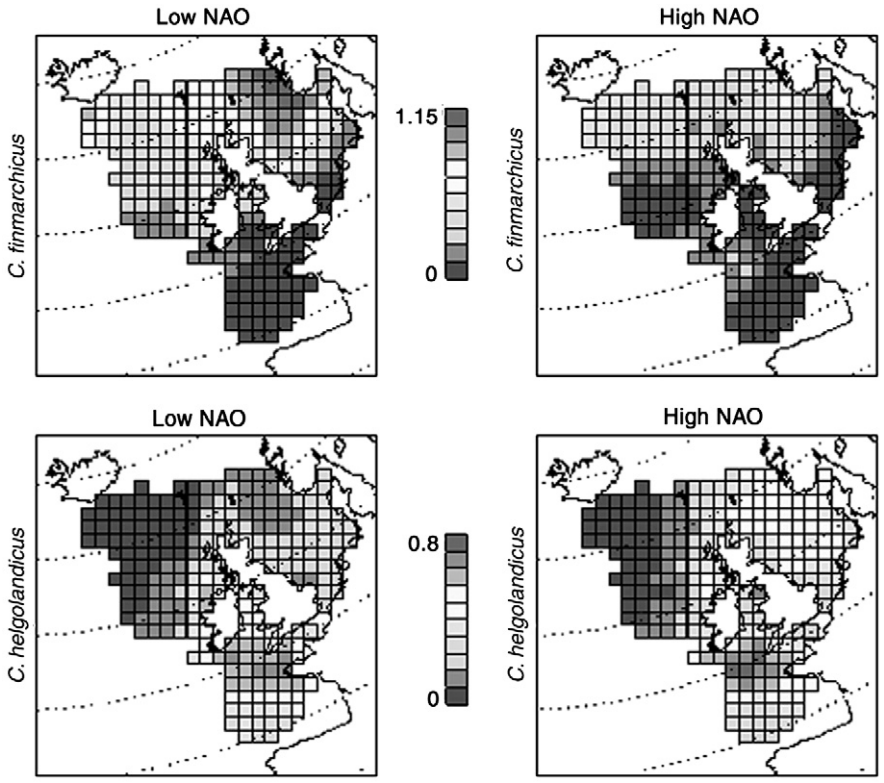


Figure 8 The mean spatial distribution of the abundance of two species of calanoid copepod Calanus finmarchicus and Calanus helgolandicus in the northeast Atlantic and the North Sea, in relation to years of both low and high NAO. The color scale is proportional to the log abundances of the species. High NAO years, 1983 and 1989–1992; low NAO years, 1963, 1964, 1966, 1977, and 1979 (maps from Planque, 1997, courtesy of SAHFOS).

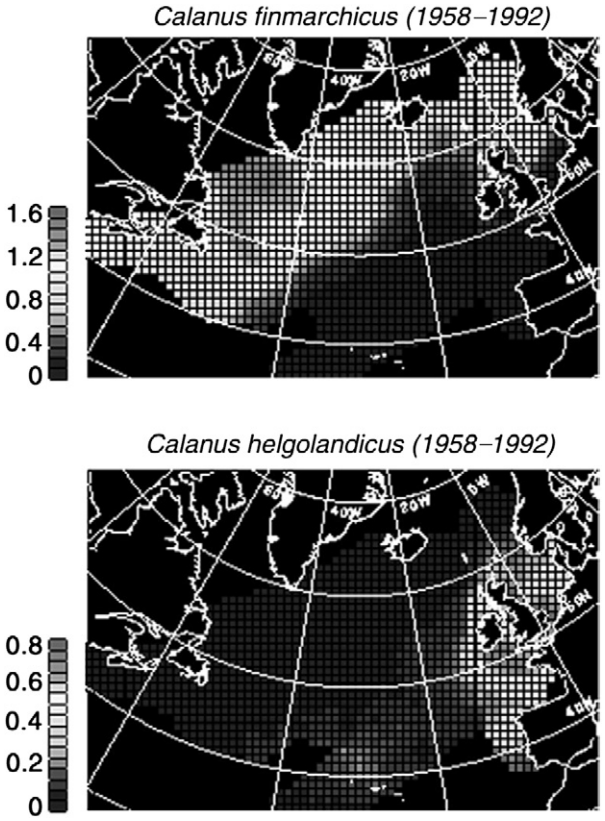


Figure 9 Regional distribution of Calanus finmarchicus and Calanus helgolandicus for the years 1958–1992 (maps from Planque, 1997, courtesy of SAHFOS).

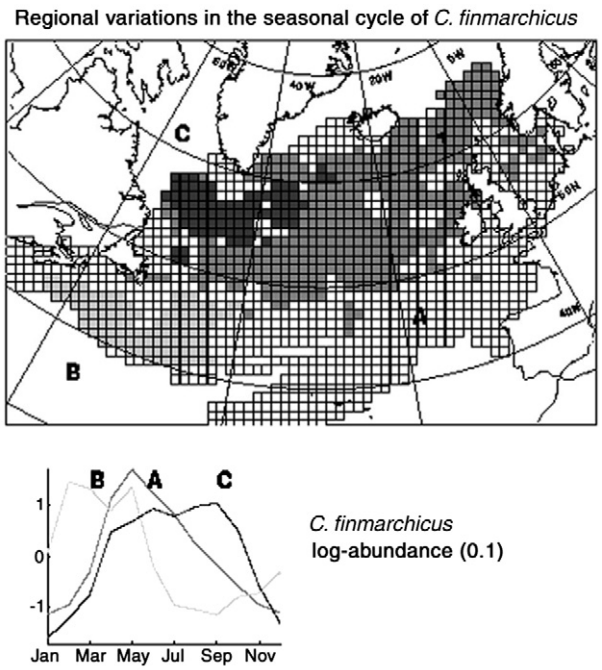


Figure 10 Regional variation in the seasonal cycle of Calanus finmarchicus showing how the distribution varies in the North Atlantic and how the peak of the log abundance differs between different locations (maps from Planque, 1997, courtesy of SAHFOS).

climatic variations in the atmosphere are reflected first in changes in the physical structure of the ocean resulting in large-scale biological responses in the ocean. These regime shifts have a severe effect on the temporal and spatial distribution of planktonic species, leading to changes in the secondary production as well as in community structure (McGowan *et al.*, 1998). In terms of species richness and diversity in the open area of the Atlantic Ocean, the maximum species richness is found at depths between approximately 1000 and 1500 m. Angel (1991) showed, for example, that the species richness of planktonic ostracods in the northeastern Atlantic at 20° W occurred between 500 and 1000 m. Moving up the water column, and at higher latitudes, there tend to be fewer species but these occur at higher abundances. The zooplankton, and especially the Calanus species, are an important part of the pelagic food web since they are the major source of food for fish species including cod. The nauplii of Calanus in particular are a key component of the diet of fish larvae and a recent model of Calanus off the Georges Bank (Miller *et al.*, 1998) showed clearly the importance of the availability of *C. finmarchicus* for the cod and haddock populations along the bank (Lough, 1984).

Recent evidence shows that SST explains almost 90% of the geographic variation in planktonic foraminiferal diversity for the whole Atlantic Ocean (Rutherford *et al.*, 1999). The peak in foraminiferal diversity was found at middle latitudes in the Atlantic, and this pattern may be extended to other taxa such as euphausiids (krill species), pteropods, and chaetognaths in the North Pacific. This is additional evidence that changes in pelagic diversity in the world's ocean result from changes in the upper-ocean physical structure and especially that water column and ecosystem structure may depend on vertical niche separation.

In the North Sea, particularly the northern areas, there has been an increase in zooplanktonic species richness, as recorded by the CPR survey, since the 1950s. The increase, particularly of calanoid copepods, in the most northerly areas can be attributed to increased inflow from the Atlantic Ocean, but currently the species which are increasing or adding to the numbers are not permanent components of the North Sea plankton. Resident and colder water holoplanktonic species have declined in abundance, and meroplankton (adult and larval) and expatriates from warmer oceanic and mixed waters have increased. This constitutes a significant challenge to analysis of pelagic diversity because the species richness of the meropanktonic groups may exceed that of the holoplanktonic groups. For example, in a region of the northwest North Sea, 17 calanoid copepod species were recorded in 10 years of CPR samples, but 24 species of decapod larvae were recorded in 2 years and 34 species of bivalve larvae in just 1 year. Meroplankton species, larval and adult, are linked to the more permanent habitats of the benthic phases of their life cycle, whereas holoplankton live entirely within a dynamic water mass, complicating the description of plankton diversity at a specific geographical location.

Pelagic Fish

In the marine environment, approximately 80% of the known fish species are found in coastal areas, whereas approximately

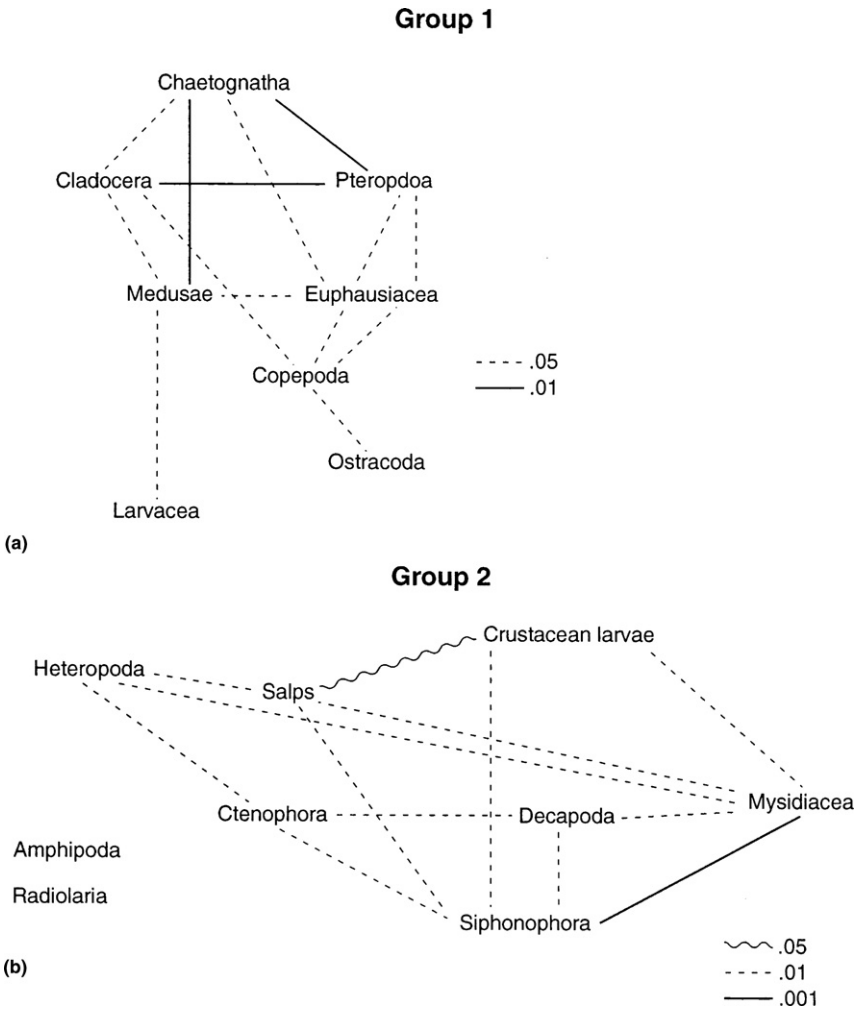


Figure 11 An example of changes in a pelagic food web in relation to a large El Niño event in California, (a) Correlations of abundance for zooplankton taxa between the years 1955 and 1959. (b) Similarities found using rank-order abundance of plankton. The years 1955–1957 showed no correlation with years 1958 and 1959. Note that 0.05, 0.01, 0.001 and the respective thickness of the line indicate the significance level of the correlations (redrawn from McGowan *et al.*, 1998).

2% are epipelagic and the remaining are deep-sea species. For example, at a 1000-m depth in the North Atlantic 44° N 13° W, the fish were the most taxonomically diverse group found in the micronekton (Angel, 1993). In the pelagic food web, pelagic planktivorous fish are regarded as predator species and changes in their prey items in relation to physical changes in the ocean circulation pattern are reflected in fish stock fluctuations (**Figure 12**) and changes in the catches of one species of sardine, *Sardinops sagax* (Kawasaki, 1991; Sharp and McLain, 1993; Strömberg, 1997). The increase in the sardine populations may reflect changes in phytoplankton biomass in the Pacific Ocean during the 1970s and 1980s (Venrick *et al.*, 1987) as well as changes in zooplankton abundances in the northeast Pacific. Zooplankton populations off the Peruvian coast declined from 1972 coincident with the severe decline in the anchoveta fishery. Shifts in the Kuroshio current southeast of Japan, and shifts in the wind stress, may be related to the fluctuations in sardine populations. The effects of El Niño Southern Oscillation events on the fluctuations observed in the fish stocks are unclear as, for example, in the case of the Chilean sardine and the Peruvian anchovetta that do not indicate a clear connection. During the El Niño events in the late 1950s, 1960s, and early 1970s, the Peruvian anchovetta stock was high, whereas the Chilean sardine stock was low. In 1971 and 1972, the El Niño event decreased dramatically and as a result the Chilean sardine stock increased while the Peruvian anchovetta stock crashed. At the same time, both the Japanese sardine and the Californian sardine increased. This trend was confirmed during the El Niño events that took place in 1982 and 1983. In the Black Sea the introduction of a new species, the comb jellyfish Mnemiopsis leidy, caused a strong decline in the Black Sea anchovy stock. As the abundance of the jellyfish increased, it had a great impact on the recruitment of the anchovy due to changes in the food web and competition for food. These population collapses are very important since pelagic fish such as sardine and anchovy species are among the top 10 species that dominate the world catches.

If we examine the long-term studies of plankton in the North Atlantic observed by the CPR, we find that in the North

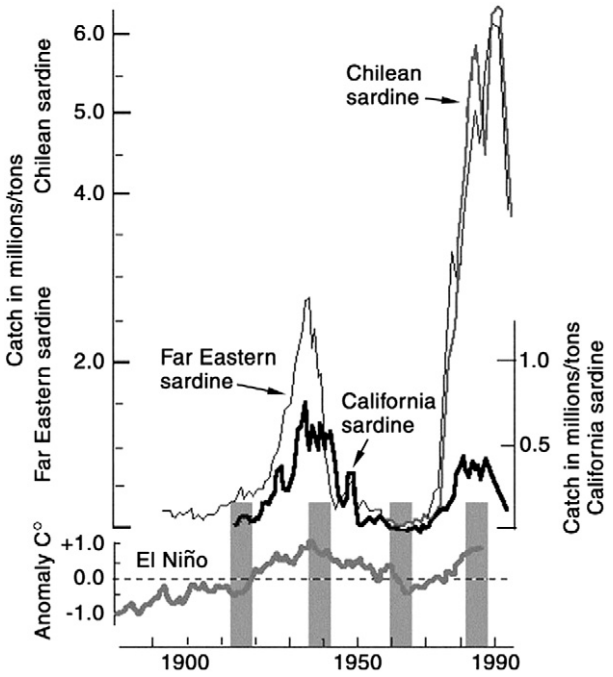


Figure 12 The covariation of three different populations of Pacific sardine compared with the temperature anomaly from the North Pacific in relation to ENSO events. The data for the California sardine are part of the California Cooperative Fisheries Investigations (CALCOFI) (data from Kawasaki T (1991) Long-term variability in the pelagic fish populations. In: Kawasaki T, Tanaka S, Toba Y, and Taniguchi A (eds.) *Long-Term Variability of Pelagic Fish Population and Their Environment*. New York: Pergamon, and Sharp GD and McLain DR (1993) Fisheries, El ninō-Southern Oscillation and upper-ocean temperature records: An eastern Pacific example. *Oceanography* 6:13–22; redrawn from Strömberg J-O (1997) Human influence or natural perturbation in oceanic and coastal waters—Can we distinguish between them? *Hydrobiologia* 352: 181–193).

Sea after 1987 an increase in phytoplankton biomass coincided with taxonomic changes in the phytoplankton and zooplankton species and a large increase in the catches of the western stock of the horse mackerel *Trachurus trachurus* L. in the northern North Sea. This reflected a northerly expansion of the stocks along the shelf edge from the Bay of Biscay to the North Sea after 1987. The increase in the western horse mackerel may be related to the fact that this species tracked down an increase in food resources available in the North Sea during high NAO years. The presence after 1988 of an unusual plankton species (for the North Sea), *Eucheta hebes*, which normally has a more southerly/oceanic distribution, suggests an increase in the shelf edge current that in turn favored the advection of western horse mackerel from the western margin of Europe, allowing them to reach the northern North Sea. As noted by Strömberg (1997), the causes and events related to the fluctuations in the standing stocks of pelagic fish in relation to changes in the plankton and to production are very complex, and these changes may be related to wind-driven advection processes and to a change in the behavior of predator and prey in response to changes in the small turbulent flow. In the coastal areas, and in particular in the shelf seas, very abrupt changes in the community structure can occur and persist on a decadal scale (Steele, 1998). The observed changes in the dynamics of pelagic fish stocks may be related to low- and high-frequency climatic variability that in turn affects the interannual variation in recruitment, growth, and abundance. For example, in a time series analysis of variation in O-group cod Gadus morhua along the Norwegian Skagerrak coast, variation was found to not be directly associated with the NAO and with the abundances of *C. finmarchicus*, but the fluctuations in the cod stock may have been related to changes in the bottom habitat and especially in the seagrass coverage by *Zostera marina* (Fromentin *et al.*, 1998). Further analysis of this data set demonstrated the importance of age-structured interactions such as asymmetric competition and cannibalism between cohorts, causing fluctuations in the abundance of the O-group and 1-group juvenile cohorts of cod (Bjørnstad *et al.*, 1999).

The northern stock of cod on the Labrador Shelf and Grand Banks of the northwest Atlantic declined and, although overfishing occurred, other factors also affected stock recruitment. The match/mismatch hypothesis was revisited by Conover *et al.* (1995). This hypothesis basically states that the spawning of cod occurs at the same time each year and if the spawning of copepods (the major food source) occurs at the same time there is a "match" and a successful recruitment. If the spawning of copepods is delayed through dependence on the occurrence of the phytoplankton spring bloom there is a "mismatch" and the recruitment of cod larvae will be low. Changes in the physical structure in the Labrador Sea and Grand Banks occurred with unusually low temperatures in the Labrador current and this was followed by the appearance in the zooplankton of the copepod *Calanus glacialis* rather than *C. finmarchicus*. *Calanus glacialis* spawns later than *C. finmarchicus* and therefore does not provide a good food source for the cod larvae. Changes in the Labrador current may be the result of climate change and in turn this affects the structure of the pelagic food web by creating a shift in the species composition as just described, thus resulting in a change in cod recruitment. Ultimately, the increase in fishing activities may have a severe effect on fish populations (Strömberg, 1997) since indirect effects considerably alter the biogenic habitat for fish recruitment, as does the removal of top predators. These indirect interactions by the fisheries can cause a trophic change in the food web as a result of top-down control of community structure, whereas a change in the resource availability induced, for example, by the El Niño event represents a bottom-up process that influences the dynamics of pelagic ecosystems. The interplay between these two types of perturbation may be the cause of shifts in species composition which result in changes in the food web structure. In light of these processes, we should also consider the importance of the microbial loop (Azam, 1998) and be aware of bacterial control of the fluxes of organic matter and its role in ecosystem dynamics. For example, fish production in the eastern Mediterranean decreased through a dominant microbial loop, and the uncoupling of bacteria from primary productivity was found in a fishery off the coastal region of Newfoundland that was causally related to the intensity of the fishery.

Outlook

The structure of pelagic ecosystems varies at both temporal and spatial scales. At an ocean basin spatial scale, phenomena such as El Niño can be linked at a temporal scale of approximately 1 year to the stock recruitment of pelagic fish, and species and single species models can be used as predictive tools. At a larger spatial scale and on a decadal timescale, climatic oscillations such as the NAO can be linked to regime shifts in the species composition and community structure of ecosystems, and models should be at a community or ecosystem level. A retrospective analysis of historical data could reveal important biodiversity patterns.

The Russell cycle (Russell *et al.*, 1973) explained the changes in the biota of the English Channel very well. In 1925, a decline in the herring stock was observed and the Plymouth herring stock (on the southwest coast of the United Kingdom) collapsed in 1936. In 1931, the decline in the macroplankton resulted in a change in the presence of the arrowworm from the species *Sagitta elegans* to *S. setosa*. In the following decade, there was a drastic decrease in the standing stocks of gadoid fish such as cod (*G. morhua*). The English Channel ecosystem essentially shifted from one dominated by herring and that had a good supply of dissolved nutrients to one with less nutrients and that was dominated by pilchards and smaller planktonic organisms. It is interesting that in 1965 there was an increase in the fish larvae concomitant with an increase in nutrient supply. These events were followed in the 1970s by a reverse switch in the Sagitta species and an increase in copepod abundance, whereas the number of pilchard eggs decreased. The Russell cycle is a well-documented example that shows how ecosystems can shift or oscillate from one state to another after reaching a bifurcation point (May and Oster, 1976).

Pelagic ecosystems seem to show decadal or longer oscillations and this is reflected in the fossil records and paleoceanographic studies. Pelagic ecosystem diversity and the maintenance of biodiversity appear to be influenced by climatic oscillations and through pressure exerted by overfishing and removal of top predators from the food web (Casini *et al.*, 2009). These can be regarded as macroscopic system properties that in turn are related to different flux that may have an effect on the structure and ecosystem functioning (Levin, 1998). An understanding of changes in biodiversity patterns in relation to the variability in ecosystem structure and function needs to be incorporated in a sustainable global ecosystem policy for the next millennium.

See also: Fish Stocks. Fishes, Biodiversity of. Food Webs. Ocean Ecosystems

References

Angel MV (1991) Variations in time and space: Is biogeography relevant to studies of long-time scale change? *Journal of the Marine Biological Association of the United Kingdom* 71: 191–206.

Angel MV (1993) Biodiversity of the pelagic ocean. *Conservation Biology* 7: 760–772.

Azam F (1998) Microbial control of oceanic carbon flux: The plot thickens. *Science* 280: 694–696.

Behrenfeld MJ and Falkowski PG (1997) Photosynthetic rates derived from satellite-based chlorophyll concentrations. *Limnology and Oceanography* 42: 1–20.

Bjørnstad ON, Fromentin J-M, Stenseth NC, and Gjøsæter J (1999) Cycles and trends in cod populations. *Proceedings of the National Academy of Sciences USA* 96: 5066–5071.

Casini M, Hjelm J, Molinero J-C, *et al.* (2009) Trophic cascades promote threshold-like shifts in pelagic marine ecosystems. *Proceedings of the National Academy of Sciences USA* 106: 197–202.

Colebrook JM (1986) Environmental influences on long-term variability in marine plankton. *Hydrobiologia* 142: 309–325.

Conover RJ, Wilson S, Harding CH, and Vass WP (1995) Climate, copepods and cod: Some thoughts on the long-range prospects for a sustainable northern cod fishery. *Climate Research* 5: 69–82.

Dickson RR, Colebrook JM, and Svendsen E (1992) Recent changes in the summer plankton of the North Sea. *ICES Marine Science Symposia* 195: 232–242.

Falkowski PG, Barber RT, and Smetacek V (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* 281: 200–206.

Fromentin J-M, Stenseth NC, Gjøsæter J, Johannessen T, and Planque B (1998) Long-term fluctuations in cod and pollock along the Norwegian Skegerrak coast. *Marine Ecology Progress Series* 162: 265–278.

Kawasaki T (1991) Long-term variability in the pelagic fish populations. In: Kawasaki T, Tanaka S, Toba Y, and Taniguchi A (eds.) *Long-Term Variability of Pelagic Fish Population and Their Environment*. New York: Pergamon.

Levin S (1998) Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* 1: 431–436.

Lough RG (1984) Larval fish trophodynamic studies on Georges Bank: Sampling strategy and initial results. In: Dahl E, Danielssen S, Moksness E, and Solemdal P (eds.) *The propagation of Cod, Gadus morhua*. Flodevigen, Norway: Flodevigen Rapp.

Marshall SM and Orr AP (1972) *The Biology of a Marine Copepod*. New York: Springer-Verlag.

May RM and Oster GF (1976) Bifurcations and dynamic complexity in simple ecological models. *The American Naturalist* 110: 573–599.

McGowan JA, Cayan DR, and Dorman L (1998) Climate–ocean variability and ecosystem response in the northeast Pacific. *Science* 281: 210–217.

Miller CB, Lynch DR, Carlotti F, Gentleman W, and Lewis CVW (1998) Coupling of an individual-based population dynamic model of Calanus finmarchicus to a circulation model for the Georges Bank region. *Fisheries Oceanography* 7(3/4): 219–234.

Planque B (1997) Spatial and Temporal Fluctuations in Calanus Populations Sampled by the Continuous Plankton Recorder. PhD Thesis. France: University of Pierre et Marie Curie Paris VI.

Planque B and Taylor AH (1998) Long-term changes in plankton and the climate of the North Atlantic. *ICES Journal of Marine Science* 78: 1015–1018.

Reid PC, Surey-Gent SC, Hunt HG, and Durrant AE (1992) Thalassiothrix longissima, a possible oceanic indicator species in the North Sea. *ICES Journal of Marine Science* 195: 268–277.

Reid PC, Edwards M, Hunt HG, and Warner AJ (1998) Phytoplankton change in the North Atlantic. *Nature* 391: 546.

Rutherford S, D'Hondt S, and Prell W (1999) Environmental controls on the geographic distribution of zooplankton diversity. *Nature* 400: 749–753.

Sharp GD and McLain DR (1993) Fisheries, El niño-Southern Oscillation and upper-ocean temperature records: An eastern Pacific example. *Oceanography* 6: 13–22.

Steele JH (1998) From carbon flux to regime shift. *Fisheries and Oceanography* 7(3/4): 176–181.

Stephens JA, Jordan M, Taylor A, and Proctor R (1998) The effects of fluctuations in North Sea flows on zooplankton biomass. *Journal of Plankton Research* 20: 943–956.

Strömberg J-O (1997) Human influence or natural perturbation in oceanic and coastal waters—Can we distinguish between them? *Hydrobiologia* 352: 181–193.

Venrick EL (1990) Phytoplankton in an oligothrophic ocean: Species structure and interannual variability. *Ecology* 71(4): 1547–1563.

Venrick EL, McGowan JA, Cayan DR, and Hayward TL (1987) Climate and chlorophyll-a: Long-term trends in the central North Pacific Ocean. *Science* 238: 70–72.

Pesticides, Uses and Effects of

Paul C Jepson, Oregon State University, OR, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 509-522, © 2001, Elsevier Inc.

Pesticides have been in recorded use since 1000 B.C. Arsenic was in regular use as a garden insecticide in China by A.D. 900 and chemicals of one form or another have protected humans and their crops and livestock throughout the development of modern civilization. In comparison with this long time scale, we are still in the earliest phases of the use of synthetic organic pesticides, which were first used over large areas in the 1940s. There has, nonetheless, been sufficient time for several generations of pesticide chemistry to evolve, and pesticides have influenced all habitats and have affected the lives of all their inhabitants over this period.

Classification and Uses

Nowhere have the costs and benefits of modern technology been more difficult to reconcile than with pesticides. Benefits that include improved yield, crop quality, and food safety and reductions in vector-borne disease incidence have driven, and will continue to drive, their use. The earlier synthetic organic compounds were, however, flawed in their environmental behavior. They were persistent, they had a very broad spectrum of toxicological activity, and they displayed a tendency to be magnified in concentration through food chains, such that damage was inflicted to animal populations that lived beyond the treated area in habitats that were not intentionally contaminated. The discovery of some of these limiting impacts was only made possible through technical advances in, for example, analytical chemistry and conceptual advances where, for example, the ability to predict environmental behavior from chemical properties only developed after pesticides had been in use for decades. New pesticide discoveries are no longer accompanied by the marvel and optimism that characterized the first synthetic insecticides. Our ability to exploit these materials has, however, advanced considerably in recent years, and the chemicals themselves are increasingly specific in their impacts and appear to pose reduced risks. Scientists still question the value of reliance upon chemical pesticides, however, and modern pest control is characterized in general by a cautious approach to their management and use.

Set in a volume that will be consulted largely by biologists and ecologists with an interest in biodiversity, this article summarizes the chemicals that are in most widespread use and outlines the processes that contribute most to efficient delivery to the biological target. The nature and importance of toxicological and ecological effects are then reviewed, followed by a summary of procedures through which pesticides are regulated and managed. This article does not provide detailed reviews of biochemical mode of action, application, formulation, or environmental fate and behavior in any detail, and readers are recommended to pursue some of the literature in the References section to gain insight into these important areas of pesticide science.

Uses of Pesticides

Pesticides may be classified by the types of pest they control (terms often bearing the suffix -cide; **Table 1**) or by the effects that they have upon the pest organism (terms that do not bear the suffix -cide; **Table 2**).

Table 1 Classification of pesticides by the type of pest controlled

Acaricide	Mites, ticks, and spiders
Adulticide	Adult insects
Algicide	Algae
Arboricide	Trees, brush, and scrub
Avicide	Birds
Bactericide	Bacteria
Fungicide	Fungi
Herbicide	Plants
Insecticide	Insect and sometimes related arthropod pests, including mites
Ixodicide	Ticks
Larvicide	Insect larvae
Miticide	Mites, ticks, and spiders
Molluscicide	Mollusks, such as slugs and snails
Nematicide	Nematodes
Ovicide	Invertebrate eggs
Piscicide	Fish
Predacide	Vertebrate Predators
Rodenticide	Rodents, including rats and mice
Silvicide	Trees, brush, and scrub
Termiticide	Ants and termites

Table 2 Classification of pesticides by effects on pests

Antifeedant	Inhibits feeding while insects remain on the treated plant
Antitranspirant	Reduces transpiration
Attractant	Lures pest to a specific location
Chemosterilant	Prevents reproduction
Defoliant	Removes foliage without immediately killing plant
Desiccant	Causes plant parts to dry
Disinfectant	Destroys or inactivates harmful organisms
Feeding stimulant	Causes vigorous feeding
Growth regulator	Stops, speeds up, or retards growth in insects or plants
Repellent	Drives away pests without killing them
Semiochemical	Pheromones and other substances emitted by plants or animals that alter animal behavior
Synergist	Substance that enhances the effects of a pesticide

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00105-2

Classification of Pesticides

Formulations

Pesticides are marketed in complex mixtures, or formulations, containing the pesticide chemical itself, the active ingredient (AI), and additives that enhance mixing and dilution in water or oil and release or deliver the toxic material once it has been applied. Common ingredients in a pesticide formulation include the AI, solvents, carriers, surface-active agents, and specialized additives. Solvent choice is determined by the solubility of the AI, potential toxicity to target plants (phytotoxicity), toxicology, flammability, volatility, and cost. Some solvents are not miscible in water and cause emulsions to be formed (e.g., xylene), whereas others are selected because of their ability to dissolve the AI and their ability to dissolve in water (e.g., isopropyl alcohol). Carriers can include inert clays that disperse the active ingredient through a powder or granular formulation. Surface-active ingredients are used to assist in the process of emulsion formation (the dispersion of the pesticide liquid within the liquid diluent), but they also include wetting agents (materials that reduce surface tension and enhance wetting and coverage of surfaces), dispersing agents (materials that maintain the emulsion as microscopic droplets within the diluted formulation), and spreading agents (materials that enhance coverage of waxy plant foliage and insect cuticles).

The pesticide formulation may be in solid (i.e., powder or granule) form, or it may be a liquid or gas concentrate. Modern pesticides may achieve their intended effects at rates of less than 10 g per hectare (10 000 m^2=1 ha), and the formulation blend enables these tiny quantities to be distributed evenly over the intended target surface.

A two-letter code denotes the formulation type on all pesticide labels, and this is a fundamentally important aspect of the selection of a pesticide for a particular use. There are four groups of formulation types:

- Group 1: Concentrates for dilution in water, including DC (dispersable concentrate), EC (emulsifiable concentrate), SC (suspension concentrate), and WP (wettable powder)
- Group 2: Concentrates for dilution with organic solvents, including OL (oil-miscible liquid) and OP (oil-dispersable powder)
- Group 3: Formulations to be applied undiluted, including GR (granules) and UL (ultra-low-volume (ULV) liquids)
- Group 4: Miscellaneous formulations, including RB (bait) and AE (aerosol dispenser)

Pesticide Types

Pesticides are characterized by their chemical diversity also. The summary below includes many of the most important groups of chemicals used in crop protection, but it is not exhaustive. Some pesticide properties are sufficiently uniform within the major classes of chemicals for this broad classification to be used in selection of chemicals for a particular use. Biochemical mode of action, for example, tends to be similar within major classes, and many crop protection programs use materials from several classes to avoid excessive selection pressure for resistance to pesticides. Other properties, are, however highly variable within pesticide classes as well as between them. These include vapor pressure (volatility) and the various partitioning coefficients that determine the distribution and fate of the active ingredient in the environment. These properties determine the effectiveness of the pesticide against a specific target or in a specific climate type or habitat, and detailed knowledge of these properties is required for pesticide selection to be effective. Some aspects of properties will be dealt with later (*see* Efficiency).

Major Classes of Insecticides

Inorganic

These are mainly nonvolatile and water-soluble compounds that often have high mammalian toxicity and may also be cumulative poisons. Many are now discontinued or banned. They include materials such as boric acid, copper sulfate, mercurous chloride, and sodium arsenite.

Organic

Botanical These pesticides are derived from plant materials, some of which have low mammalian toxicity. Some have adverse effects on wildlife, including toxicity to fish. Examples include neem tree (*Azadiracta indica*) oil, which is used to protect stored products from insects; pyrethrum, from *Chrysanthemum cinerariaefolium*, a powerful but rapidly degraded insecticide that affects the peripheral nervous system, causing paralysis, known as "knockdown"; nicotine (from *Nocotiana tabacum*), an alkaloid insecticide that is also a neuromuscular poison in vertebrates; and rotenone, from the roots of *Derris* and *Lonchocarpus* spp., also used as a piscicide. Some of these materials are widely used in locally made or commercial formulations, and although they may be highly toxic, many are not persistent in the environment and degrade rapidly in sunlight or when exposed to microbial activity in soil or water. Each material has unique properties, and it is not possible to make generalizations about toxicology or environmental impact.

Organochlorines Organochlorines are synthetic pesticides, among the first to be developed. Early materials, including DDT, had dramatic early successes in vector-borne disease control, but most have adverse environmental impacts, resulting from persistence in the environment, accumulation through food chains, global redistribution of residues, and ecotoxicological impacts. Many organochlorines are now banned internationally. Examples include diphenyl aliphatics (DDT, dicofol, methoxychlor), benzene derivatives (HCH, pentachlorophenol), cyclodienes (chlordane, endosulfan, endrin), and polychloroterpenes (camphechlor).

Organophosphates Organophosphates, or OPs, are esters of phosphoric acid. Many have high mammalian toxicity and may require frequent application because they are generally not persistent. Organophosphates are nerve poisons, acting through inhibition of cholinesterase. They fall into three groups: (1) Aliphatic organophosphates, which are the oldest group, some with low mammalian toxicity (e.g., malathion, which has been in use for 40 years) but others having high mammalian toxicity, but short persistence. Short persistence makes many OPs of use in short-season crops, where the plant or the consumer has a low tolerance for pests

(e.g., vegetable crops). Examples include, acephate, dichlorvos, dimethoate, malathion, and phorate. (2) Phenyl organophosphates, which are more stable and persistent but which include materials with high mammalian toxicity. Examples include fenitrothion, methyl parathion, and temephos. (3) Heterocyclic organophosphates, which may be more persistent and active in soil. Examples include azinphos-methyl, chlorpyriphos, phosmet, and pyrazophos.

Carbamates Similar to organophosphates, carbamates are esters of carbamic acid, with short residual life and a wide spectrum of activity. This group also includes some herbicides and fungicides. Some of the pesticides with highest mammalian toxicity are carbamates. Again, they divide into three groups: (1) Methyl carbamates with a phenyl ring structure, including carbaryl and methiocarb; (2) methyl and dimethyl carbamates with heterocyclic structures, including aldicarb and methomyl; and (3) methyl carbamates of oximes, having a chain structure, including bendiocarb.

Formamidines Formamidines are used as insecticides and acaricides, have the characteristic nitrogen structure –N=CHN–, and are effective against the eggs and larvae of Lepidoptera (butterflies and moths). They are useful alternatives to organophosphates, where resistance has developed. Examples include amitraz.

Dinitrophenols Dinitrophenols are broad-spectrum insecticides that have two nitro groups (NO_2) attached. Examples include binapicryl and dinocap.

Organotins Organotins are tin-based organic compounds that act as acaricides and fungicides, some with long residual activity. Examples include cyhexatin.

Pyrethroids Pyrethroids are synthetic compounds based upon pyrethrins, found in chrysanthemum flowers. They have similar toxicological properties but tend to be much more photostable and persistent. Natural pyrethrins occur in mixtures of six esters, and resistance is rare in the insects that they are used against. The synthetic pyrethoids are generally marketed as a single ester, and resistance may develop rapidly. Pyrethoids are generally of low mammalian toxicity but are highly toxic to fish and bees and beneficial predatory or parasitic invertebrates. Pest resurgence, or population outbreaks that result from the destruction of natural enemies, may occur in certain crops. Examples include allethrin, bioresmethrin, cyfluthrin, deltamethrin, and permethrin.

Fumigants Fumigants are insecticides that kill by vapor or gas action, have low molecular weight, and often contain halogen radicals. They can be highly toxic to vertebrates, particularly in enclosed spaces. Examples include chloropicrin, ethylene dibromide, and methyl bromide.

Petroleum Oils Petroleum oils, also known as mineral oils, have insecticidal properties. Light applications of refined paraffin oils can be made to trees in leaf. Lower viscosity, semirefined oils can be applied to dormant trees in winter to kill invertebrates and their eggs.

Antibiotics One antibiotic, abamectin, derived from *Streptomyces*, is an effective insecticide and antiparasitic agent also in veterinary use. There is evidence that when used in cattle, it can harm dung-feeding invertebrates. It has some systemic properties (i.e., it can penetrate leaves and be carried within the plant vascular system).

Semiochemicals Semiochemicals are behavior-modifying compounds, including pheromones, particularly the sex attractants of female moths. These may attract male moths for monitoring or be dispersed by spraying or formulation into controlled-release devices to establish false trails and disrupt mating of pest species. These compounds have very low environmental hazards, and the development of resistance is rare.

Insect Growth Regulators Insect growth regulators are compounds derived from, or inhibitory to, insect hormones and include methoprene, a juvenoid or juvenile hormone mimic, effective against mosquitoes, and diflubenzuron, which disrupts the synthesis of chitin and kills insects when they molt.

Microbials Included in this group are bacterial agents, particularly *Bacillus thuringiensis*, which is used against Lepidoptera, Diptera, and Coleoptera. A proteinaceous inclusion or endotoxin, which constitutes 40% of cell weight at sporulation, breaks down in insects with alkaline midguts to an active protoxin. This category also includes fungal diseases of insects, including *Beauvaria bassiana*, *Metaryzium anisopliae*, and *Verticillium lecanii*, some of which can be formulated and sprayed like conventional insecticides. Some viral agents are also used on a large scale, including nuclear polyhedrosis viruses (NPV) of Lepidoptera. Microbial insecticides tend to have very limited effects on organisms that they do not directly parasitize, and vertebrate effects are almost unknown.

New Insecticides A series of new insecticides belonging to several previously unexploited classes of chemistry have been introduced in recent years. Some of these pose greatly reduced environmental risks, although this does not apply to all of the newer materials. These include imidacloprid (in the class chloronicotinyls), a highly systemic pesticide that is now used on a very wide scale. Research on these newer insecticides is limited at present, but there is optimism that ecological impacts and risks to wildlife will be greatly reduced as they become more widely used.

Major Classes of Fungicides

Inorganic

Inorganic fungicides are derived from sulfur or simple metal salts. They are generally stable, persistent, and insoluble in water. They include sulfur, which was originally applied as flowers of sulfur, in dust form, and which is still used, but in a more highly ground colloidal suspension. It has both direct contact and fumigant activity at temperatures above 20 °C, but above 32 °C the vapor may cause phytotoxicity (toxic harm to the target plant). Environmental damage and toxicological impacts to nontarget organisms are otherwise limited. Copper-based fungicides include Bordeaux mixture, an early fungicide that consists of a solution of copper sulfate and

hydrated lime. With 12% copper, this fungicide has low mammalian toxicity. A more stable form of copper (e.g., copper oxychloride) is used in modern formulations, enabling the slow release of copper into leaf surface water film and toxic buildup in fungal tissue. Other inorganic fungicides have included heavy-metal-containing materials that incorporate mercury, zinc, nickel, or chromium. These are normally highly toxic and persistent, and they have been banned internationally.

Organic

Organic fungicides have the reputation of being safer and less persistent than some of their inorganic counterparts, and they are often used at very low doses.

Dithiocarbamates Dithiocarbamates are derivatives of sulfur-containing dithiocarbamic acid, in combination with zinc salts, ferric salts, and manganese salts. These fungicides have greater efficacy, better stability, and less phytotoxicity than elemental sulfur. Their toxicity derives from the formation of the isothiocyanate radical (–N=C=S–) in breakdown. Compounds in this group include thiram, maneb, metham-sodium, and zineb.

Organometallics This group of fungicides includes the following: (1) Mercury compounds, formerly popular for disinfective and protective action as well as volatility. They have high mammalian toxicity and are no longer available for any purpose. (2) Organocopper compounds, including copper acetate, which was first synthesized in 1899. These fungicides are not easily washed off leaves, being insoluble in water, and give persistent protection. They act by the nonspecific denaturing of proteins. Examples include cuprobam. (3) Organotins, triphenyltin salts, including fentin, which can be both toxic and phytotoxic.

Substituted Aromatics These compounds include derivatives of benzene and phenol, with hydrogen atoms replaced by chlorine, nitrogen, and oxygen, and are suited for seed treatment and control of soil-borne fungi. Examples include chlorthalonil and pentachlorophenol (PCP).

Dicarboximides Also called sulfenamides, these compounds are considered to be among the safest pesticides in seed treatment and protectant sprays. Examples include iprodione and vinchlozin.

Phthalamides Phthalamides are used as nonsystemic, broad-spectrum, foliar fungicides on fruit, vegetables, and ornamentals. Examples include captafol and captan, which is used widely in the tropics, but increasingly subject to restrictions.

Dinitrophenols These nonsystemic fungicides are used against powdery mildew and include the compounds binapacryl and dinocap.

Triazines This group consists mainly of herbicides, with anilazine as the only fungicide, used as a protectant treatment in vegetables.

Systemic Compounds

These fungicides are absorbed and translocated through plant tissues and provide longer term, protective control, with some curative and therapeutic effects for plants that are already infected. There are many groups, some of which are summarized below:

Oxathiins Oxathiins control basidiomycete fungi and include the compounds carboxin and methfuroxam.

Benzimidazoles and Thiophantes Benzimidazoles and thiophantes are broad-spectrum fungicides and are widely used in the tropics; intensive use has often led to resistance. Some of these compounds replaced organomercury fungicides as seed dressings. They include benomyl, carbendazim, and thiabendazole.

Pyrimidines Included in this class of compounds are bipirimate and ethirimol.

Acylalanines This class of compounds includes metalaxyl.

Ergosterol Biosynthesis Inhibitors (EBIs) Ergosterol biosynthesis inhibitors are a heterogeneous group of compounds with a common mode of action. They can have systemic, protective, and curative properties. They include (1) imidazoles (e.g., imazalil and prochloraz), (2) piperazine, pyridine, and pyrimidine compounds (e.g., pyrifenox and triforine), (3) morpholines (e.g., dodemorph and tridemorph), and (4) triazoles (e.g., flutriafol, myclobutanil, and propoconazol).

Organophosphates Organophosphates are a group consisting mainly of neurotoxic insecticides but also include fungicides such as pyrazophos.

Phenylamide and other Fungicides against Oomycetes This is another heterogeneous group, sharing the property of toxicity to oomycete fungi. High specificity and systemic activity help to confer resistance when used intensively. Included in this group are (1) phenylamides (e.g., benalaxyl and metalaxyl) and (2) the compounds cymoxanil and prothiocarb.

2-Aminopyrimidines This small group of systemic fungicides has activity against powdery mildew and includes bupirimate.

Quinones This group includes benodanil and futonil.

Other Organic Fungicides A number of important fungicides do not belong to the groups above, but comprise a random selection of compounds with unrelated chemical structures. They include chlorothalonil, dodine, guazatine, and thiocyclam. Most are nonsystemic, protective fungicides.

Major Classes of Herbicides

Herbicides kill or interrupt the growth of plants. Some are selective, but others kill all plants that come into contact with them and are used in industrial settings and in rights-of-way maintenance.

Inorganic Herbicides

This class includes mostly salts that have been in use for a considerable time. Sodium arsenite solutions and arsenic trioxide were popular in the 1960s but are now banned. This class also includes iron and copper sulfates, which are used for foliar application. Sulfuric acid has been widely used in horticulture and cereal crops as a selective herbicide. These materials are water soluble and readily leach from soil. Further examples include sodium salts of boric acid (borates) and sodium chlorate.

Organic Herbicides

Organic Arsenicals These materials are much less toxic to mammals than inorganic arsenic salts. They inhibit metabolism, competing with phosphate in essential reactions. These compounds include disodium methanearsenate (DSMA).

Phenoxyaliphatic Acids Herbicides in this group are a series of compounds in which the phenoxy nucleus is linked with acetic, propionic, and butyric acids. Solubility in water is high relative to that of many herbicides. They have selective, hormone-type effects on broad-leaved weeds, but grasses are tolerant. 2,4-D, 2,4,5-T, and MCPA were immensely popular and once regarded as safe. In the 1970s, dioxin contamination was suspected in 2,4,5-T and its registration was canceled.

Substituted Amides This is a large group of organic nitrogenous herbicides, including amides and anilides. Amides primarily act in soil against annual grass weeds, causing stunting. They have low mammalian toxicity and include chlorthiamid and propyzamide. Anilides, which have numerous subgroups, are used in postemergence grass and broad-leaved weed control as well as in preemergence compounds to control germinating weeds and grasses. All have low mammalian toxicity (e.g., alachlor and propanil).

Diphenyl Ethers These compounds possess two benzene rings, joined through oxygen or a more complex chain of molecules. They include preemergence and selective postemergence compounds of low mammalian toxicity. They are fairly insoluble in water, do not leach, and may be persistent for several months. Examples include diclofop methyl and oxyfluorofen.

Dinitroanilines Dinitroanilines are the most widely used group of herbicides in agriculture. They have a dinitroaniline nucleus in common and are effective against annual grasses and broad-leaved weeds when applied preemergence. Degradation occurs through volatilization and photodecomposition. Mammalian toxicity is low, and in soil persistence can be quite long. This class of compounds includes benfluralin and trifluralin.

Substituted Ureas Based on the simple nitrogen-containing molecule urea, the first example, monuron, was discovered in the 1950s as a total herbicide. More modern compounds are used as selective, preemergence herbicides that are strongly adsorbed to soil. They inhibit photosynthesis, causing chlorosis. These compounds have low mammalian toxicity and their efficacy is influenced by temperature, rainfall, and soil type. Examples include chloroxuron, diuron, isoproturon, and linuron.

Carbamates In addition to many insecticides and fungicides, a number of herbicides have been developed from this chemical group. They include pre- and postemergence materials that inhibit germination and cell division. They have low mammalian toxicity and short persistence. Examples include asulam and phenmedipham.

Thiocarbamates The thiocarbamates are a group of carbamates containing sulfur, including selective pre- and postemergence compounds. The -allate compounds may persist for many months in soil. Metham sodium is a soil fumigant, converted in the soil to methyl isothiocyanate. Examples include diallate, metham sodium, and thiobencarb.

Heterocyclic Nitrogens These compounds have a ring structure where the carbon atoms are replaced by nitrogen or sulfur. They include triazines, triazinones, triazoles, pyridines, and uracils. *Triazines* are a very selective group of compounds; selectivity depends upon the plant's ability to metabolize the AI. They are soil applied and absorbed through roots, inhibiting photosynthesis once the plant has emerged. Mammalian toxicity is generally low, and intense application in some areas has led to groundwater contamination and restrictions upon use. Examples include atrazine and simazine. *Triazinones* have six-membered rings and include hexazinon. *Triazoles*, with a five-membered-ring structure, are active against broad-leaved weeds (e.g., amitrole). Some *pyridines* are effective against deep-rooted herbaceous weeds, and others are selective brush killers (e.g., picloram and trichlopyr). *Uracils*, another substituted, six-membered-ring family of chemicals with two nitrogen atoms and a double bond, are primarily for preemergence application and uptake by roots. They are effective at controlling annual grasses and broad-leaved weeds over an extended period by inhibition of photosynthesis. Uracils may be very persistent. Examples include bromacil and lenacil.

Bipyridiliums, also Termed Pyridines Containing two pyridyl rings, this group includes some extremely well known and widely used compounds (e.g., diquat and paraquat). They have contact action on the above-ground parts of plants and in the plant they are reduced to free radicals that destroy tissue under light. They are widely used as desiccants. These compounds are deactivated by sorption to the soil and slow to degrade, leading to increasing restrictions. Both AIs are hazardous, being the main cause of pesticide-related death in many countries. There is no known antidote.

Aliphatic Acids This group includes chlorinated derivatives of acetic acid, trichloroacetic acid (TCA) and dalapon. These compounds are used on non-cropland and in forestry, killing the plant by causing precipitation of proteins within cells.

Phenol Derivatives Herbicides in this group are highly toxic, selective, foliar herbicides and include dinitrophenols

and one chlorinated phenol, pentachlorophenol. Dinitrophenols, primarily contact herbicides, inhibit respiration and photosynthesis (e.g., dinoseb (DNBP), DNOC) and are now banned in the United States. Pentachlorophenol is now being banned in Europe and the United States.

Benzonitriles or Substituted Nitrites These compounds consist of a benzene ring with a cyanide (C≡N)- radical and are broad spectrum, acting on various processes of growth and tissue disruption (e.g., dichlobenil and ioxynil).

Miscellaneous Herbicides Herbicides with unrelated chemical structures include some widely used materials such as endothal sodium, an aquatic weed killer, glyphosate, a nonselective, residual, postemergence herbicide, and alloxydim sodium, fluazifop-butyl, and other selective, postemergence systemic herbicides, used against perennial and annual grasses.

Efficiency

Ideally, the receiving organism receives a toxic dose of pesticide and nontarget organisms escape injury, achieving high efficiency of use. Efficiency is, however, very rarely measured. When it has been, efficiency is far from ideal when expressed as the proportion of the applied dose taken up by the target organism or the amount needed to directly combat an infestation by treating it directly (Table 3).

What should our expectation for efficiency be? High efficiency cannot be expected in all circumstances. For example, many applications are made before serious infestations ensue, and pests are few in number or widely dispersed. It is generally agreed, however, that there is considerable scope for improvement in the efficiency of pesticide delivery, through attention to application, formulation, and the optimum delivery of the AI through attention to the physicochemical properties of the compound. The interaction between the pesticide and organism that leads to chemical exposure and uptake is termed bioavailability. The term availance describes the profile of chemical concentration over time, a function of physicochemical properties, transfer, and sorption within the treated environment.

Delivery and Bioavailability

The toxic effects of a pesticide against any intended target or unintended nontarget organism are a function of intrinsic toxicity (the activity when it is directly applied to the organism) and the amount that reaches the organism under the conditions of application. The amount that the organism is exposed to depends upon placement (which is a function of application method), pattern of release (a function of the formulation), redistribution and transfer processes (a function of the physicochemical properties of the pesticide, environmental conditions, and the nature of the substrate), and the location and receiving characteristics of the organism. The latter is a function of the degree of contact or avoidance of the chemical, the activity of the organism in question, and, on a large scale, the spatial dynamics of the organism as it relates to the pattern of treatment and the persistence of the chemical.

Application, Formulation, and Delivery of the Toxic Dose

Through application and formulation, it is possible to manipulate the placement, size, and composition of the discrete units or drops of pesticide in which it is dispensed and the rate of release from those units or particles.

Soil Treatment

Some of the most difficult challenges relate to the problem of delivering sufficient pesticide to small targets distributed in a bulk medium in which mobility is restricted and the chemical subject to rapid degradation. It is therefore important to localize the pesticide to the vicinity of the target and plant part that needs protection. For example, granular formulations and seed treatments can both be used to increase efficiency of transfer of pesticide to organisms in the soil. Once applied, the chemical must move into the soil matrix and the rate of release determines the concentration in the soil. Diffusion and redistribution can be achieved by bulk flow of water moving down through the soil profile after rain or irrigation. Contact

Table 3 Utilization efficiency of pesticides

Pesticide	*Method of application*	*Target organism*	*Utilization efficiency (%)*	*Explanation*
Demeton-S-methyl	Foliar spray	Aphids on sugar beet	0.000008	Insects located in heart leaves, an effective refuge from direct spraying
Dieldrin	Seed treatment	Wheat bulb fly larvae	0.0015	Pests less likely to encounter seed as plant develops
Dimethoate	Foliar spray	Aphids on field beans	0.03	Relatively low efficiency of spray retention by plant
Lindane	Foliar spray	Capsids on cocoa	0.02	Pesticide losses due to drift when treating large trees
Lindane/dieldrin	Aerial spraying of swarms	Locusts	6.0	Pesticide applied by aircraft within swarm, maximizing chance of contact

Source: Data from Graham-Bryce IJ (1977) *Philos. Trans. R. Soc. London, B* 281: 163–179.

between the water network in soil pores and the coated seed is very important.

Efficiency, or the toxicity of soil-applied AI per unit of material released, is determined by the way in which the chemical partitions between the solid, liquid, and air phases of the soil. The partition coefficients between these different phases can be used to accurately predict the rate of movement of pesticides within a given soil type.

Spray Application

Spray application is the commonest method of pesticide delivery. It is convenient, flexible, and simple, but it lacks selectivity and there is a risk of contaminating nontarget areas. Hydraulic sprayers give a range of pesticide drop sizes, at the smallest extreme of which there is a tendency to drift away from the target area. Controlled droplet application (CDA) sprayers are much rarer, but they control drop size and as a consequence may reduce drift and increase efficiency.

The main function of spray machinery is to transfer energy to the liquid pesticide formulation to atomize it into droplets by means of a nozzle. The drop size distribution is typical for the nozzle in question and it is characterized by the volume median diameter and the number median diameter of the drops.

- Volume median diameter (VMD): If a sample of spray is divided into two equal volumes, then half of the volume contains drops with a diameter less than the VMD, and the other half contains drops with a diameter greater than the VMD.
- Number median diameter (NMD): If the drops are divided in half by number, independent of volume, then the diameter of half of the drops is larger than the NMD, and the diameter of the other half is smaller.

The NMD : VMD ratio indicates the spread of drop sizes: the larger the ratio, the broader the spectrum. If the ratio approaches 1, the drops are all of a similar size and behave in a similar way.

Conventional spray application through hydraulic nozzles presents a major challenge for optimization of pesticide use efficiency, and their widespread use accounts for a great deal of pesticide drift and off-target exposure and contamination. **Table 4** provides details of the drop spectrum from a conventional fan jet hydraulic spray nozzle. Drops of between 1 and 50 μm in diameter account for only 4.2% of the volume, but 90.8% of the total drops in the sample. In addition, 61.9% of the volume of the spray liquid is in drops of greater than 160-μm diameter. In 1 liter of spray liquid, this would yield nearly 90 million drops with a diameter of less than 50 μm, and therefore would be susceptible to drift.

Drop behavior has been the subject of detailed research. Drops are subjected to various forces as they leave the nozzle and before they impact on a surface.

- Gravitational forces: Small drops have a low terminal velocity. For example, a 5-μm-diameter drop falls at 0.075 cm/sec and often drifts from the target area, whereas a 500-μm-diameter drop falls to the ground at 213.9 cm/sec.
- Movement induced by air movement: Wind carries drops away from the target area. For example, in a wind of 1 m/sec, a 5-μm-diameter drop released from 1 m experiences 350 m of sideways transport, whereas a 500-μm drop moves 0.48 m. To deposit within 5 m of the target area, a drop must have a diameter of greater than 67 μm in a wind speed of 0.7 m/sec and it must have a diameter of greater than 168 μm in a wind traveling at 3 m/sec. The friction of the air rapidly decelerates drops from the velocity at which they left the nozzle. For example, a 5-μm drop is decelerated to its deposition velocity in 0.33 cm and a 200-μm drop is brought to the equivalent velocity in 630 cm.

Drift is worse where relative humidity is low. Drops become smaller as spray liquid evaporates and sedimentation takes longer. If 1 liter of liquid is atomized to 10-μm drops, then the cumulative surface area is 600 m^2. If it is atomized to 100-μm drops, then the surface area falls to 60 m^2. Evaporation is therefore exacerbated by the shrinkage of drops and the increase of relative surface area.

Large drops pose an equivalent problem: not deflected by air currents, they gain momentum and either miss the plant target and hit the soil or strike the plant, shatter, and fall to the ground. The cubed relationship between diameter and volume means that the volume of a 100-μm-diameter drop is 1 million times that of a 1-μm-diameter drop. The variation in pesticide dose between a 4- and 400-μm-diameter drop (common in a hydraulic nozzle drop spectrum) is 1 million fold. Many drops of a diameter greater than 300 μm fall to the ground, and a large proportion of the pesticide applied by a hydraulic nozzle

Table 4 Drop sizes from a hydraulic spray nozzle measured with a laser particle size analyzer

Drop range (μm)	*Volume (%)*	*Cumulative volume (%)*	*Number (%)*	*Cumulative number (%)*
563–262	26.17	61.94	0.07	0.80
262–160	35.77		0.75	
160–113	17.52		1.36	
113–84	8.78	33.84	1.82	8.39
84–65	4.81		2.31	
65–50	2.73		2.86	
50–30	2.66		8.45	
30–15	1.12	4.22	21.90	90.81
<15	0.27		60.46	

Source: Data from Micron Sprayers, UK. Nozzle, Fan Jet SS8002 operating at a pressure of 3 bar and a flow rate of 700 ml/min. VMD = 192 μm NMD = 12 μm, VMD:NMD ratio-16.

never comes into contact with the pest, disease, or weed it is intended to control.

Deposition on Obstacles

Objects are surrounded by a cushioning layer of air, and spray drops require considerable momentum to hit a target. Larger drops strike targets, and smaller drops move around them. The probability of striking depends upon the shape of the object; hairs, for example, pick up small drops. Flying insects such as mosquitoes are best hit by drops of 10–20 μm in diameter, resting flies require drops of 30–40 μm, diameters of 90–130 μm are needed where crop penetration is needed, and a diameter of >250 μm is optimal for horizontal surfaces such as weeds.

Continuing reliance upon hydraulic spray technology limits the efficiency, effectiveness, and safety of the spray application process.

Ecotoxicology and Management

Pesticides pose challenges through being toxic to organisms other than the specific pests, diseases, or weeds they are targeted against. Relative toxicity is expressed in a number of ways, but the most prevalent relates to oral and dermal exposure of mammals, which equates to both human and vertebrate wildlife toxicological risks for organisms that are directly exposed (**Table 5**).

Table 6 lists some representative active ingredients by toxicity and demonstrates that in all major pesticide types and range of classes, there is a considerable variety in intrinsic toxicities to mammals.

It is extremely difficult to translate these data, or toxicological data for any other species, into predictions of ecological risk. Very toxic materials may be relatively nonhazardous in the environment, because they are rapidly sorbed or dissipated, whereas less toxic materials may prove hazardous because they readily leach or runoff into water or because they are persistent.

Recent attempts have been made to bridge the gap that exists between basic toxicology and environmental chemistry and the potential ecological impact of pesticides in ecosystems. Van Straalen and van Rijn (1998) computed species sensitivity distributions for soil fauna (soil invertebrates that contribute to biogeochemical cycling), based upon laboratory

Table 5 Pesticide toxicity classes

Toxicity rating	*LD_{50}, single oral dose for rats (mg/kg)*	*LD_{50}, single dermal dose for rabbits (mg/kg)*	*Probable lethal oral dose for humans*
6, supertoxic	<5	<20	A taste, a grain
5, extremely toxic	5–50	20–200	A pinch, 1 teaspoon (5 ml)
4, very toxic	50–500	200–1000	1 teaspoon (5 ml) to 2 tablespoons
3, moderately toxic	500–5000	1000–2000	1 liquid ounce (30 ml) to 1 pint (470 ml)
2, slightly toxic	5000–15,000	2000–20,000	1 pint (470 ml) to 1 quart (950 ml)
1, practically nontoxic	>15,000	>20,000	>1 quart (950 ml)

Table 6 Examples of pesticide toxicities to vertebrates from laboratory test data used in registration

Pesticide type and class	*Pesticide technical name*	*Oral LD_{50} (rat, mg/kg)*	*Dermal LD_{50} (rat (a), rabbit (b))*
Insecticides			
Organochlorine	Dieldrin	46	10 (a)
	Endosulfan	80	359 (b)
Organophosphate	Azinphos-methyl	16	222 (a)
	Chlorpyrifos	135	2000 (b)
	Dichlorvos	56	75 (a)
	Dimethoate	320	<130 (b)
	Malathion	2800	4100 (b)
Carbamate	Aldicarb	0.7	5 (b)
	Carbaryl	850	>2000 (b)
Pyrethroid	Deltamethrin	135	>2000 (b)
	Allethrin	930	–
	Bioresmethrin	7070	–
Fungicides			
Dithiocarbamates	Thiram	780	>5000 (a)
	Maneb	7990	>5000 (a)
Dinitrophenol	Dinocap	980	–
Herbicides			
Inorganic	Sodium arsenite	39	–
Bipyridiliums	Paraquat	157	–
Triazines	Simazine	5000	>3100 (b)

Table 7 Rankings of 10 selected pesticides according to three criteria[a]

Pesticide	*Rank based on toxicity*	*Rank based on persistence*	*Joint assessment (recovery time)*
Lindane (organochlorine insecticide)	6	2	2
Dimethoate (organophosphate insecticide)	8	8	7
Parathion (organophosphate insecticide)	5	6	6
Chlorpyrifos (organophosphate insecticide)	1	7	4
Carbofuran (carbamate insecticide)	4	3	3
Carbaryl (carbamate insecticide)	2	10	5
Methomyl (carbamate fungicide)	7	9	8
Benomyl (benzimidazole fungicide)	3	1	1
Atrazine (triazine herbicide)	10	4	10
Fentin (organotin fungicide)	9	5	9

[a]The score 1 is given for the highest toxicity (the 95% protection level referred to in the text), the greatest persistence (soil half-life), and the longest predicted ecotoxicological recovery time (i.e., the time it takes the pesticide concentration to decline to the 95% protection level).

Table 8 Summary of data from farming system experiment in the United Kingdom (Boxworth, SCARAB, TALISMAN, RISC, LIFE), Germany (INTEX), The Netherlands (Nagele), and Switzerland (Third Way): Results of integrated farming, with reduced pesticide inputs, compared with the conventional levels of use[a]

Project	*Beneficial arthropods*	*Birds and mammals*	*Earthworms*	*Soil microorganisms*	*Soil minerals*
Boxworth	+	o			
SCARAB	+		o	o	
TALISMAN	=				
RISC	o				
LIFE	+		+	=	=
Lautenbach	+		+		=
INTEX	+				=
Netherlands	+	+	+		
Third Way	+		+	+	

[a]Key: +, increase; =, no change; o, variable result over the period of the study.

test data for toxic effects in soil, and calculated the concentration that would be below the no-effect concentration (NOEC: the highest dose not to cause a specific effect) for reproductive inhibition for 95% of the species that are theoretically present within the soil arthropod community. Using a simple model for pesticide fate, based upon exponential decay, they then calculated for a series of compounds the time that each product would take to reach this "95% protection level." This time, referred to as the "ecotoxicological recovery time," can be used as a basis for calculating the minimum time for ecological recovery to take place after pesticide application. There is considerable scope for further development of this technique, built upon readily collected toxicological data sets. Such data are, however, surprisingly difficult to find in the published literature, and they are not routinely collected for suites of compounds in such a way that they may be exploited within this new approach as a decision-making aid.

Table 7, adapted from van Straalen and van Rijn (1998), summarizes the differences in ranking of toxicity that result if pesticide toxicity to soil organisms and chemical persistence (half-life) in soil are integrated to predict the time when recovery could take place by organisms. A compound with high intrinsic toxicity, such as chlorpyrifos, has a lower ranking once its limited persistence is taken into account, whereas benomyl climbs higher up the rankings because of its persistence. Analytical procedures of this form are beginning to forge a link between the large amounts of laboratory-obtained toxicological data that exist and the field, where functioning communities are exposed to the toxin, rather than individual species.

Farm-Scale Observations of Ecotoxicological Impact

A number of research projects investigating the ecology of farming systems were initiated throughout the 1980s and 1990s in Europe (Holland *et al.*, 1994). Pesticide impacts were investigated in all of these studies, and comparisons between conventional and reduced pesticide inputs were made in at least 13 sites, in Germany, The Netherlands, Switzerland, and the United Kingdom (**Table 8**). In all cases, beneficial nontarget invertebrate densities were higher in areas where the total regime of pesticide input had been reduced. In the larger scale studies (e.g., the Boxworth study, UK), with 5.6- to 15.7-ha treatment area sizes, some beneficial species were rendered locally extinct for the full 5-year treatment phase of the project, and recovery was subsequently slow. The level of impact on individual species was shown to be a function of life history attributes that affected pesticide exposure and capacity of that organism to reinvade the treated area. Predatory capacity was inhibited in the highest pesticide regimes, and there was evidence that this contributed to higher pest densities in some years. Although similar data have been obtained for the invertebrate community of rice systems in the

tropics, data sets of this level of complexity are rare in the investigation of pesticide impacts in agroecosystems.

These investigations have revealed the subtler impact of second- and third-generation active ingredients that followed the organochlorines. They reinforce the need to measure ecological impacts on scales that reflect agricultural practice and pesticide use in the real world and that are tuned to the scale of dispersive movement of nontarget taxa. They have also triggered considerable interest in the mechanisms that underlie long-term depletions, even local extinction of certain species. Both scale of treatment and the mode and rate of dispersal of arthropods influence rates of population recovery following pesticide use, and the proximity of local refugia from which recovery can occur, and landscape features conducive to movement and colonization are important factors underlying extinction risk. Modeling may provide an appropriate tool for testing our understanding of invertebrate population processes at the agroecosystem level, but it does not substitute for the need to undertake a far greater number of manipulative experiments and monitoring programs that examine the spatial dynamics of nontarget organisms in sprayed farming systems.

Avian Impacts

The requirement for large-scale, field-based monitoring and experimentation is not restricted to terrestrial nontarget invertebrates. Avian toxicologists have long recognized that field-based studies provide unique and indispensable data for interpretation of pesticide effects on birds (Taub, 1997). Many of the most important effects of pesticides on birds, including eggshell thinning as a result of the bioconcentration of organochlorines through food chains, were originally detected and documented as a result of field-based observations. These effects also include endocrine disruption, food supply reduction or alteration, variation in sensitivity between species, and the synergistic effects of exposure to multiple compounds.

Aquatic Systems

Similar arguments may be applied to the investigation of pesticide effects on aquatic organisms. The recent decision in the United States to reduce the requirements for ecological data from multispecies test systems in the regulatory process has been criticized because it will reduce the probability of detecting biologically significant effects (Taub, 1997). These effects include indirect trophic-level impacts, compensatory shifts within a trophic level, responses that are associated with seasonal trends in populations, chemical transformations effected by organisms in the exposed system, and impacts that result from long persistence of either the parent product or toxic breakdown products.

In conclusion, there is now abundant evidence that pesticide impacts can evolve at the agroecosystem scale and that this requires the development of appropriately scaled monitoring or experimental systems. Toxicological data are of fundamental value in the initial evaluation of pesticides, but in their real-world applications, pesticides may also elicit ecological effects that reverberate through the system, long after the chemical residues have become undetectable. Laboratory-based, single-species tests are limited in their predictive power for impacts in the field, and further development of the theory and methodology associated with ecological effects is required.

Tools for Pesticide Management

It is questionable whether we apply our knowledge about pesticides effectively at any stage following the regulatory permission to use a product in a specified way. Given the need to combine knowledge of chemical properties, fate and behavior, environmental attributes such as soil type, toxicology, and ecology, the challenge is considerable. In each of these disciplinary areas, however, there are considerable databases of knowledge and predictive models that can be used to interpret how a pesticide will behave in a given set of field conditions.

Attempts have been made to summarize clusters of pesticide properties in databases that may be used to compare the environmental and ecological risks posed by lists of candidate compounds for specific uses. These databases are not specifically tuned to local conditions; they do, however, combine many factors, including risks to wildlife, in the rankings that they generate, and enhance the capacity of the end user to make more informed decisions about particular uses. The most significant development in this field is the proposal to derive Environmental Impact Quotients (EIQ) for compounds from established databases of pesticide properties (Kovach *et al.*, 1992). The quotient combines farmworker effects (built from acute and chronic toxicity and plant surface half-life) with consumer effects (built from systemicity, soil and leaf surface persistence, and potential for groundwater contamination) and ecological effects (built from aquatic and terrestrial ecotoxicology). There is considerable scope for toxicologists to develop approaches such as this from first principles in order to make data available in a form that can be used to assist decisions in the field. Guided by EIQs, growers and their advisors could then plan pest control tactics that attempt to minimize nontarget impacts and track their progress with a rigorous and quantitative methodology that could also be explained to consumers.

See also: Agriculture, Industrialized. Agricultural Invasions. Ecotoxicology. Herbicides. Insecticide Resistance

References

Bellows TS and Fisher TW (1999) *Handbook of Biological Control*. San Diego: Academic Press.

Benbrook CM (1996) *Pest Management at the Crossroads*. New York: Consumers Union.

Calow P (1998) *Handbook of Ecotoxicology*, vols. 1 and 2. Oxford: Blackwell Sci.

Conway GR and Pretty J (1991) *Unwelcome Harvest: Agriculture and Pollution*. London: Earthscan Publications.

Croft BA (1990) *Arthropod Biological Control Agents and Pesticides*. New York: Wiley.

Farm Chemicals Handbook (2000) Willoughby, Ohio: Meister Publishing Co.

Greig-Smith PW, Frampton GK, and Hardy AR (eds.) (1992) *Pesticides, Cereal Farming and the Environment*. London: H. M. Stationery Office.

Hartley GS and Graham Bryce IJ (1980) *Physical Principles of Pesticide Behavior*. San Diego: Academic Press.

Holland JM, Frampton GK, Cilgi T, and Wratten SD (1994) Arable acronyms analysed: A review of integrated arable farming systems research in W. Europe. *Ann. Appl. Biol.* 125: 399–438.

Kogan M (ed.) (1986) *Ecological Theory and Integrated Pest Management Practice*. New York: Wiley.

Kovach J, Petzoldt C, Degni J, and Tette J (1992) A method to measure the environmental impact of pesticides. *New York's Food Life Sci. Bull.* No. 139.

Marco GJ, Hollingworth RM, and Durham W (1987) *Silent Spring Revisited* Washington, DC: Am. Chem. Soc.

Matthews GA (1992) *Pesticide Application Methods*, 2nd edn. New York: Longmann.

Taub FB (ed.) (1997) Invited feature: Are ecotoxicological studies relevant to pesticide registration decisions? *Ecol. Appl. 7* 1083–1132.

Van Straalen NM and van Rijn JP (1998) Ecotoxicological risk assessment of soil fauna recovery from pesticide application. *Rev. Environ. Contam. Toxicol.* 154: 83–141.

Ware GW (1999) *The Pesticide Book*, 5th edn. Washington, DC: Thomson Publications.

Pharmacology, Biodiversity and

Paul Alan Cox, Institute for Ethnomedicine, Jackson, WY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bioassay A test for pharmaceutical activity in substances conducted either in living organisms *(in vivo* assays) or in the test tube *(in vitro* assays).

Coevolution A heritable response made by a species through evolutionary time to a different species that is then reciprocated.

Ethnobotany The study of the uses of plants by different groups of peoples, often indigenous peoples.

Extremophile An organism that lives and thrives in a habitat characterized by extremes of temperature, acidity, alkalinity, salinity, light, or pressure that would prove lethal to most other organisms.

Fractionation A method of obtaining a pure compound by extracting components of a mixture with solvents of different polarity.

Herbal A compilation of medicinal plants and their properties often used in earlier times as a reference for physicians prescribing medical treatments based on plant therapies.

Indigenous intellectual property rights Rights to intellectual properties belonging to indigenous peoples, including but not limited to iconographical representations, terminologies, names, phrases, legends, methods, and techniques of traditional cultivation, healing, identification, preparation, and other use of biodiversity.

Indigenous peoples Peoples who have resided in the same geographical area for many generations and who possess legends, proverbs, genealogies, languages, and other unique cultural features linking them to the land.

Shotgun DNA A technique that isolates deoxyribonucleic acid (DNA) fragments from soil or water samples and compares them to known DNA libraries to determine the presence of novel species.

Voucher A representative specimen of a plant or an animal that is properly collected, prepared, and preserved in a herbarium or museum to facilitate expert identification of the species.

Introduction

Pharmacology, the science of drugs, studies the ways in which natural products can be employed as medicines. Since ancient times, humans have recognized the pharmaceutical properties of certain compounds derived from plants and animals. Current research now also recognizes that biodiversity must be maintained in order for the environment to continue to be a source of such medicinal substances.

Evolutionary Perspectives

In the 4.5 billion years since the Earth was formed, living organisms have evolved a plethora of complex molecules to help mediate interactions with their physical and biotic environments. Approximately 3.6 billion years ago, organisms originated that use deoxyribonucleic acid (DNA) to transmit genetic information. However, it is believed that before the rise of DNA-based organisms, perhaps as early as 3.8 billion years ago, Earth was inhabited by organisms that solely used ribonucleic acid (RNA) as their primary genetic material (Joyce, 2002). Echoes of this previous RNA world can be found in ribosomal RNA (rRNA), which facilitates protein assembly. Even earlier, perhaps 4 billion years ago, before the rise of the RNA world, there was an era of self-replicating peptide nucleic acids (PNA), such as *N*-(2-aminoethyl)glycine units with the bases attached, bridging the abiotic and biotic eras. The interplay between increasing molecular diversity and the concomitant increase in biodiversity, which characterizes the history of life on Earth, has resulted in an astonishing array of complex molecules, many of which cannot be synthesized in modern laboratories. This molecular cornucopia can be used to develop new drugs and new approaches to the treatment of disease, drawing from molecules that are themselves of ancient date. For example, the use of antisense RNA, RNA interference, or even antisense PNA, for novel antibiotics and gene silencing, has been informed by studies of "primitive" prokaryotes. And the relatively recent discovery of an entire new domain of life – the Archaea – which were initially believed to be limited to extreme habitats, but are now known to be relatively ubiquitous throughout Earth, has opened new approaches for drug discovery.

The evolutionary divergence approximately 3.5 billion years ago of the three major domains of life, the Eukarya, the Bacteria, and the Archaea, led to dramatic differences in species and molecular diversity among these domains. The International Union for Conservation of Nature (IUCN) estimates that there are 1.7 million known species of Eukarya, which include plants, animals, fungi, and algae. With an estimated 5.5 million eukaryotes remaining to be described, there are perhaps 5–15 million species of eukaryotes on Earth. This diversity, however, pales when compared to that of the prokaryotes. Based on shotgun DNA techniques, estimates of the number of species of bacteria have increased in the past three decades from several thousand to as many as a billion. Understandings of species diversity of Archaea, while nascent, are also increasing. Recently, the smallest known prokaryotes, with diameters as small as 500 nm, were discovered in acidic mine tailings in northern California by UC Berkeley researcher

Brett Baker (Baker *et al.*, 2010). An analysis of the Archaeum *Pyrococcus furiosus* by the biotech company Strategene led to the discovery of the enzyme Pfu DNA polymerase, which is useful in DNA fingerprinting and genetic engineering.

The inescapable conclusion of these considerations is that our planet is biochemically diverse and, as far as it is known, the most molecularly complex assemblage in the universe. Evolution has driven organisms to increasingly sophisticated levels of chemical interaction, with some of the resultant molecules functioning as materials for membranes and cell walls, quorum signals, photosynthetic pigments, energy storage compounds, neurotransmitters (in animals) and hormonal signals (in plants as well as in animals), defensive compounds, a variety of protective pigments, and structural materials. The resultant juxtaposition of similarity and diversity of the chemical compounds created by different life forms is striking. The ancient evolutionary divergences resulted in differential vulnerability to toxic compounds, with ranges in LD50s of specific compounds varying several orders of magnitude between different phylogenetic groups. For example, the concentration of penicillin that completely stops cell wall formation and growth in *Streptococcus* bacteria shows little effect when administered to a centipede or to a geranium, whereas the dose of 2,4-D sufficient to kill a geranium would have negligible impact on either centipedes or bacteria.

Throughout the history of life on Earth, such differences in biochemical vulnerability have generated abundant possibilities for chemical coevolution between different species. Even though life on Earth appears to share a common phylogeny, biochemical differences between species leave open the possibility that a mutation in the biosynthetic pathway of one organism may yield a compound toxic to another. Selection for such protective molecules from predation is perhaps the greatest in organisms that are unable to flee or lack alternative forms of defense. Among eukaryotes, sessile organisms such as corals, sponges, tunicates, and other marine invertebrates and most plants have been particularly rich sources of pharmaceuticals. Although poisonous birds have been found in New Guinea, toxic frogs are prominent in the Neotropics, and brightly colored toxic insects are known to most amateur naturalists, biochemical protection against predation appears to be the exception rather than the rule for motile organisms.

A plant species under pressure from a predator or parasite may perish unless individuals in the plant population possess, by mutation or some other evolutionary accident, a compound toxic or unpalatable to its enemies. If the mutation is heritable, progeny of such plants will increase in time throughout the population, replacing those individuals previously lost to predation. If the species of predator or parasite are highly selective in their choice of prey, they may in turn become imperiled by such chemical innovations, unless individuals in the predator or parasite population possess a means of detoxifying the defensive chemistry generated by the plants. The resultant coevolutionary race between two species, sometimes termed a "red queen race" after the never-ending contest described by Lewis Carroll in *Alice in Wonderland*, has been documented in the interaction between *Passiflora* vines and *Heliconius* butterflies in the Neotropics by Larry Gilbert and coworkers at the University of Texas. Other such tight coevolutionary links are the topic of much scientific interest.

More common than species–species coevolution is what might be termed diffuse chemical coevolution. Instead of evolving defensive chemistry against a single species of predator or parasite, a plant species generates compounds effective against a wide variety of plant enemies. The presence of toxic compounds in a plant cannot, however, be considered *prima facie* evidence of an evolutionary response to predation. Bioactive molecules occur in plants as secondary metabolites as well as evolved chemical defenses against predation, fungal attack, microbial invasion, and viral infection. It is also unlikely that plant compounds known to be toxic to humans represent evolutionary responses to anthropogenic harvesting; given the omnivorous, generalist foraging patterns of higher primates, it is unlikely (with the possible exception of agricultural weeds) that, absent artificial selection, any plants have evolved a specific biochemical response to consumption by humans. However, hundreds of thousands – perhaps even millions – of bioactive molecules in nature result from plant chemical warfare with viruses, bacteria, fungi, and animals. Many of these compounds, although toxic at certain doses, can be beneficial to humans at lower doses.

The emerging picture of chemical interactions between sessile prey and motile predators is compelling. What might otherwise appear to be a quiet forest glen or a tranquil coral reef may be the evolutionary equivalent of a battlefield rife with chemical warfare. Sessile prey appear to sequester toxic agents, and motile predators develop detoxification strategies at a prodigious rate through evolutionary time. Since many of these bioactive molecules have relevancy to human disease, a coral reef or forest glen could also be said to resemble a large pharmaceutical storehouse, perhaps one that is in great disarray. Tens of thousands of vials of pharmacologically active compounds litter the factory floor, but the labels identifying the contents and their therapeutic utility have been lost. How can one determine which molecules are useful and for which diseases?

Eloy Rodriguez and coworkers at Cornell University have found evidence that some vertebrates self-dose with pharmacologically active compounds produced by plants. Many mammals, including primates, engage in self-medication when ill. Ongoing pharmacological self-experimentation by animals may have been observed and mimicked by early humans. Polynesian legends indicate that kava, a beverage made from the roots and rhizomes of *Piper methysticum* rich in tranquilizing sesquiterpenes and kava lactones, was discovered by watching the calming effect the roots had on rats. Legends from India claim that in ancient times, mongooses were observed to feed on *Rauvolfia serpentina* before engaging in combat with cobras. Copying the reputed feeding behavior of the mongoose, local people found that the shrub could serve as a potent antidote to snakebite. The veracity of such legends is unclear, but it seems useful to trace the use of plant-derived bioactive molecules in medicine whose use in prehistory by indigenous peoples must first be examined.

As powerful as the paradigm of competition was for pioneers in the theory of evolution, more recent focus has been directed toward evolutionary cooperation. Mutualisms and symbioses may prove to be much more common than had previously been suspected, particularly when beneficial evolutionary relationships between prokaryotes and eukaryotes

are considered. For example, both termites and cattle depend on the presence of methanogenic archaea in their guts to digest cellulose. Endophytic cyanobacteria in the roots of cycad trees produce a potent motor neuron toxin, *β-N*-methylamino-L-alanine (BMAA), which once likely served to protect the ancestors of modern cycads from dinosaur herbivory. And the emerging study of biofilms has profound implications for human health ranging from new antibiotics to strategies for implantation of artificial organs. The molecular signaling involved in the establishment and maintenance of symbioses and biofilms represents a new and largely untapped frontier for pharmacology. Already disruption of quorum sensing molecules has been used to target biofilms characteristic of chronic human infections (Hentzer *et al*, 2003).

History

Early Understandings

All known indigenous peoples have ethnomedical traditions based, at least in part, on pharmacologically active compounds found in plants and animals. It is therefore safe to assume that early humans also developed significant ethnopharmacopoeias; this assumption is corroborated by archeological findings of what appear to be medicine pouches in ancient burial sites. Results of successful experiments with pharmacologically active plants and animals were likely passed from generation to generation. This accumulation of oral information resembles in part the results of a vast, uncontrolled, and unwritten human bioassay experiment. Later knowledge about medicinal plants was transmitted in written form. By the fifth century BC, Hippocrates codified and disseminated knowledge from earlier times concerning plant medicine in the Mediterranean area. During this period, commerce in medicinal plants was vigorous, especially with traders called *rhizotomii* dealing principally in plant roots. The folk knowledge accumulated by the *rhizotomii* was subsequently studied by scholars, including Aristotle (384–322 BC), who laid a foundation for the philosophical consideration of plants. Aristotle bequeathed his library to his student, Theophrastus, who wrote the most important book on plants up to that time, *De Causis Plantarum*, and who is now regarded to have been the first academic botanist. In his *Inquiry into Plants*, which in the ninth book catalogs information on medicinal plants, Theophrastus incorporated information from the *rhizotomii*. He was, however, skeptical of folk superstitions concerning the gathering of medicinal plants.

In the first century, Pedanius Dioscorides, a Greek physician, wrote a seminal work on medicinal plants titled *De Materia Medica*, which describes more than 500 medicinal plants and includes many drawings. Until the beginning of the Renaissance, Dioscorides's work was the final word in medicinal plants for more than 1000 years.

Such early compilations of folk wisdom concerning medicinal plants were not confined to the West. In 2000 BC, the Chinese emperor Chi'en Nung compiled the *Pen Tsao*, which is perhaps the earliest known herbal. In India, Sanskrit texts on medicinal plants date to 3000 years ago. In the Americas, the Mayans also wrote extensive manuscripts on medicinal plants. Unfortunately, many of these codices were burned by the early Spanish conquistadors who took such examples of scholarly achievement by their unchurched and "uncivilized" indigenous opponents as evidence of demonic influence.

Herbals and Medicinal Plants

In Europe, an increase in scholarly interest in herbals and medicinal plants began during the Renaissance, when the first physic gardens (gardens of medicinal plants for therapeutic use and study) were planted in Italy and Germany. Most early Renaissance herbals were largely based on the work of Dioscorides's book, *De Materia Medica*, with additions made from the author's own knowledge. As can be seen in *De Materia Medica*, two pieces of information were disseminated concerning each plant: its purported healing properties and its identification.

Arber (1953) pointed out that the necessity for precise identification led to an unusual feature of early botanical iconography: For nearly 2000 years, plants in herbals were pictured with the part most likely to be in commerce: their roots. Although these earlier herbals were lovingly copied throughout the ages, the quality of illustrations declined until, in the fourteenth century, it had become increasingly difficult to identify the plants. Innovators such as Fuchs and Mattioli began to inscribe illustrations drawn from living plants on wood blocks for printing. This was a major boon for botanists since images reproduced by wood block printing did not suffer the same image deterioration in successive copies as occurred in hand-inked illustrations. The new herbals also contained current information gleaned from European folk medicine.

In the English-speaking world, the first herbal is an eleventh-century Anglo-Saxon codex known as the *Herbarium of Apuleius Platonicus*. The earliest printed English herbal is an anonymous quarto from 1525 imprinted by Richard Banckes (as quoted in Arber, 1953, p. 41): "Here beynneth a newe mater, the whiche sheweth and treateth of y vertues and proprytes of herbes, the whiche is called an Herball."

A year later, a translation of a French herbal was published by Peter Treversi, and in 1538 William Turner published *Lebellus de re Herbaria Nova*. In 1551, Henry F. Lyte published a translation of Dodoen's famous herbal, and in 1555 Anthony Askham published a herbal derived from the 1525 English work of Banckes. However, the most popular of all sixteenth-century herbals was that of John Gerard, published in 1597.

Born in 1545, at the age of 32 Gerard was appointed superintendent of the gardens of Lord Burleigh at the Strand in London. Later, he was appointed as the curator of the Physic Garden of the College of Physicians of London. A catalog of plants in his own garden appeared in 1596, and in 1597 he published *The Herball, or General Natural History of Plants*, one of the most quoted botanical works ever published. The only published text to have remained constantly in print for such a long period is the Bible. Gerard's *Herball*, with 1392 pages and 2200 woodcut images of medicinal plants, was greeted with tremendous enthusiasm by medical practitioners and the general public.

The importance of Gerard's *Herball* in the history of discovery of novel bioactive compounds from biodiversity cannot be overstated. His description of the medicinal properties of *Filipendula* – now called *Spirea* – a genus in the rose family, led to the isolation of salicin and the eventual synthesis of "*A Spirea* n" or aspirin. Another example is Gerard's entry on page 646 "of Foxe gloves":

> Foxe gloue boiled in water or wine, and drunken, doth cut and consume the thick toughness of grosse and slimie flegme and naughtie humors; it openeth also the stopping of the liver, spleene and milt of other inward parts.

Gerard's recording of foxglove opening the stopping of "other inward parts" was not systematically examined for 200 years. In 1785, William Withering published *An Account of the Foxglove and Some of Its Medical Uses: With Practical Remarks on Dropsy and Other Diseases*, in which he quoted Gerard's account of the "vertues" of foxglove and proposed that foxglove could be an important medicine for dropsy, an ailment caused by the retention of fluid due to inadequate pumping action by the heart. The connection between dropsy and heart disease was not properly understood in Withering's day, but Withering observed the action of foxglove on the heart: "It has a power over the motion of the heart, to a degree yet unobserved in any other medicine, and … this power may be converted to salutary ends." By any standard, foxglove as administered by Withering was an astonishingly successful treatment for dropsy. J. K. Aronson at Oxford recently reanalyzed data from Withering's cases and found a cure rate of between 65% and 80%.

The Linnaean name given for foxglove, *Digitalis*, was affixed to the crude drug as well as to cardiac glycosides (steroidal compounds with sugars attached at the 3-position) isolated from foxglove in the early twentieth century. More than 30 cardiac glycosides have been isolated from dried foxglove leaves, including digitoxin and digoxin. These two drugs have never been synthesized, and they are still extracted from dried foxglove leaves. Each year, more than 1500 kg of pure digoxin and 200 kg of digitoxin are prescribed to thousands of heart patients.

Digitalis, however, is but one drug inspired by Gerard's *Herball*. Since the publication of the *Herball*, 18 different pharmaceutical compounds have been isolated from plants that Gerard described, including 16 drugs prescribed today. It is likely that the English translation of *Herbarium Amboinense* by Rumphius will also lead to the discovery of new drugs from an ancient text as will analyses of other historical herbals (Buenz *et al.*, 2004).

European herbals and folk medicine were not the only source of pharmaceuticals; drug discovery has been greatly enhanced by non-Western traditions such as Ayurvedic medicine in India. For example, *Rauvolfia serpentina* has long been used in India to treat snakebite and insomnia and as a sedative for hyperactive children. In 1931, Indian chemists isolated a variety of molecules from the plant and later found that *Rauvolfia* powder lowered blood pressure. In 1949, R. J. Vakil published a clinical study of *Rauvolfia* in the *British Heart Journal*. Following up on Vakil's research, Dr. Emil Schlittler and Hans Schwarz at CIBA extracted from *Rauvolfia* roots an alkaloid, reserpine, which at low doses shows strong activity in lowering blood pressure. CIBA soon introduced reserpine to commerce. Reserpine had a direct effect on the hypothalamus and recently has been prescribed in combination with other antihypertensive drugs such as hydralazine hydrochloride.

Modern Chemical Approaches to Drug Discovery

Digitalis was discovered through the same sequence of steps used in modern ethnobotanical drug discovery programs: (1) folk knowledge accumulates concerning possible pharmacological activity of a plant, (2) the plant is used therapeutically by a healer, (3) the healer communicates this knowledge to a scientist, (4) the scientist collects and identifies the plant, (5) plant extracts are tested with a bioassay (a preliminary screen for desired pharmacological activity), (6) a pure compound is isolated using the bioassay to trace the source of the activity in the plant extract, and (7) the structure of the pure substance is determined.

This method of discovery and isolation of new drugs from plants persisted until the 1950s because most pharmaceutical research relied heavily on compounds derived from vascular plants, particularly plants used in folk medicine. Indeed, flowering plants and ferns (as opposed to microscopic organisms and fungi) have given rise to approximately 120 commercially sold drugs. Of the top 150 prescription drugs in the US, 57% contain at least one compound derived from biodiversity, including plants, animals, fungi, and bacteria (Grifo and Rosenthal, 1997). One-fourth of all US prescription drugs contain molecules derived from, or modeled after, naturally occurring molecules in higher plants (Grifo and Rosenthal, 1997), including reserpine, digitalis, vincristine, and many other compounds derived from plants used in traditional medicine.

The importance of biodiversity-derived drugs is particularly evident in cancer therapy. From 1981 to 2002, 62% of novel molecular entities showing anticancer activity were nonsynthetic. And of 74 synthetic drug candidates developed during the same time frame, 48 mimic natural products (Newman *et al.*, 2003; Newman and Cragg, 2007).

In the late 1950s and early 1960s, the search for bioactive molecules expanded to the marine environment. Increased access to underwater breathing apparatus and small submersible craft made studies of novel bioactive chemicals produced by marine organisms far easier. In the underwater realm, most discoveries of novel bioactive molecules had no basis in folk knowledge, even though evolutionary processes similar to those that produced large chemical diversity in higher plants are also believed to be operant in the marine environment, particularly among sessile organisms such as corals, sponges, and tunicates. It is somewhat surprising that few indigenous cultures rely on marine organisms as part of their pharmacopoeias, but perhaps difficulties in harvesting marine resources in the absence of SCUBA technology and the unpredictability of capturing mobile marine organisms resulted in greater indigenous reliance on terrestrial plants.

Natural products derived from biodiversity may contribute to the search for new drugs in three ways: (1) by producing new drugs used in an unmodified state (e.g., vincristine from

Catharanthus roseus (Apocynaceae)), (2) by providing chemical "building blocks" used to synthesize more complex compounds (e.g., the synthesis of oral contraceptives from diosgenin derived from *Dioscorea composita* and *Dioscorea mexicana* (Dioscoreaceae)), and (3) by indicating new modes of pharmacological action that allow complete synthesis of novel analogs (e.g., synthetic analogs of reserpine from *Raulvolfia serpentina* (Apocynaceae)).

Regardless of their origin, recent discoveries of potential new pharmaceuticals have been facilitated by increasingly sophisticated techniques of vouchering and species identification, bioassay, fractionation, and structural elucidation of molecular entities.

Vouchering

Proper vouchering in pharmaceutical research has sometimes been overlooked with disastrous effects because accurate identification of a species of plant or marine organism has been rendered impossible due to a missing or inadequate voucher specimen. The importance of adequate herbarium voucher specimens cannot be overstated: Any subsequent question or dispute concerning the identity of the species involved can be unequivocally settled by expert examination of a properly collected voucher specimen. Vouchers for fish and marine invertebrates are usually stored as dried or fluid preserved (FAA or aqueous ethanol) specimens deposited in zoological or natural history museums. Specimens of algae, lichen, and vascular plants are preserved as dried specimens in herbaria. Because of their scientific value for the future (sophisticated bioassays of the future may require only micrograms of material), voucher specimens should be deposited and preserved in well-curated herbaria or museums with duplicate specimens deposited in geographically distant institutions, including institutions in the countries in which the material was collected. Addresses and contact information for herbaria throughout the world are listed in *Index Herbariorum*, published by the International Society of Plant Taxonomists and the New York Botanical Garden.

The collection number of the specimen should be used to label all subsequent pharmacological fractions and residues so that any new discovery or question can be immediately referred to the original specimen. Herbarium specimens should consist of the leaf, fruit, flower, and other plant parts necessary for proper identification by a botanist. Plants are pressed in newspaper and dried in the field using plant presses and a heater. When it arrives at the herbarium, the dried plant material is glued to high-quality rag bond paper and stored in files in herbarium cases. Properly stored under the correct conditions, herbarium specimens can retain their scientific value for centuries. In the field, it is also important to take careful notes to accompany the voucher specimens, including precise information on the collection location and other information on the plant, such as floral odor or color, that might be lost during drying. The collector's name and date of collection should always be included. For plants or animals used in folk medicine, ethnobiological data on the diseases treated, mode of formulation, and methods of administration should be sufficiently detailed to guide subsequent investigations.

In preparing samples for pharmaceutical testing, drying is used to stop enzymatic processes of degradation: Most cease when water content is reduced to 10% by volume. R apid drying is preferred, although high temperatures can alter pharmaceutical constituents. Air drying at 20–40 °C is preferred, but sometimes in the moist tropics, temperatures must be raised far higher to promote drying of bark or roots. With access to the laboratory, freeze-drying of a sample is an attractive alternative, whereas under field conditions in the humid tropics, fluid preservation in aqueous ethanol or formalin is often necessary to prevent spoilage of the sample.

The vouchering of prokaryotic samples is more difficult, particularly given the rise of shotgun DNA techniques that can identify new species of archaea and bacteria in the absence of the entire organism. In shotgun DNA techniques, pioneered by Craig Venter and his colleagues, DNA from soil or water is split into small random fragments. These fragments are attached to viral DNA and packaged to create DNA libraries. The resultant DNA libraries are sequenced, with overlapping DNA sequences assembled by computer. Each full DNA sequence is then compared to known DNA sequences, and novel DNA sequences are placed in the database as evidence of a new species.

The increase in species discovery using shotgun DNA techniques is phenomenal, with individual scientists being able to collect up to 100 new species per day. While the establishment of an axenic culture deposited with either the American Type Culture Collection (ATCC) or the Pasteur Culture Collection (PCC) remains the gold standard of prokaryotic vouchering, difficulties in culturing, let alone describing new bacteria or archaea identified through shotgun DNA techniques, suggest that vouchering in the future may revolve around publicly accessible databases for DNA sequences.

Bioassays

Bioassays vary with regard to type and precision. *In vivo* bioassays involve administering the test substance to a living organism to determine the substance's pharmacological activity. One of the more common *in vivo* bioassays that has been used to great advantage by some investigators is a brine shrimp bioassay, in which the test substance is administered to a culture of brine shrimp, available from any aquarium supplier. The percentage death of the brine shrimp within a short period is recorded, with brine shrimp mortality yielding a crude measure of cytotoxicity. Other *in vivo* bioassays include the Hippocratic screen, a test protocol for oral or peritoneal administration of the test substance in rats or mice. Various behavioral and physiological parameters are measured before and after administration, allowing the type and mode of pharmacological activity to be adduced. The most important *in vivo* bioassays for pharmaceutical research are human clinical trials, which are almost always administered after reasonable expectations of safety and efficacy have been determined from animal trials. The power of *in vivo* assays is their ability to reveal unexpected or novel modes of pharmacological activity since the impact of the test substance on the whole organism can be studied.

More common in modern pharmaceutical use, though, are *in vitro* assays – bioassays that do not depend on a living organism. Although some broad-based *in vitro* assays depend on living tissues or cells, such as the guinea pig ileum, frog sciatic nerve, or human cancer cell lines, far more common is the use of target-specific molecular assays. In such assays, a single point of action in a given biochemical pathway, such as the role of the enzyme phospholipase A2 in mediating cellular inflammation, has already been identified. Various substances are then screened to determine if they impact the enzyme, receptor, site, or rate-limiting step of interest. Modern high-throughput *in vitro* assays are typically costly (given the immense research effort required to identify the best target along a given biochemical pathway), proprietary, and designed for high-throughput rapid testing of numerous samples for the desired pharmacological activity.

In a typical high-throughput bioassay, minuscule portions of a given sample are tested for an enzymatic color change when combined with the test system. These bioassays test by screening, using a computer-controlled camera, hundreds and even thousands of different samples deposited in the microwells of a glass dish that are manipulated by robotic arms. The advantages of modern high-throughput bioassays are their speed, efficiency, repeatability, and specificity. The number of false positives (initial indications of pharmacological activity that do not pan out in subsequent testing) in modern receptor-binding assays is low; however, such high specificity entails a research liability since, unlike tests on living animals, a molecular bioassay is unlikely to reveal any mode of pharmacological action other than the specific activity it is designed to detect.

The use of precise high-throughput bioassays designed by molecular biologists has resulted in some spectacular successes, such as the discovery of the cholesterol-lowering drug Zocor or the generation of antiinflammatory drugs based on COX-2 inhibition such as Celebrex. However, such screens almost certainly miss many important modes of action. It is unlikely, for example, that the psychotropic action of LSD or the utility of phosphodiesterase inhibitors such as Viagra to treat male erectile dysfunction would have ever been discovered had their effects not been noted through inadvertent administration to living human beings.

Isolation

Once a "hit" or pharmacological activity is detected in a test substance during a bioassay, the substance, particularly if it is a natural product of unknown chemical composition, must undergo fractionation. During fractionation, solvents of different polarity are used to produce increasingly refined chemical samples of the test substance. Each fraction is then tested against the original bioassay to determine its bioactivity. Fractions exhibiting significant bioactivity are further fractionated until eventually a purified compound is obtained. High-pressure liquid chromatography, gel electrophoresis, and other chemical techniques are also employed in addition to solvent fractionation to purify samples. This approach to isolating pure bioactive molecules by tracing bioactivity through the purification process is called bioassay-guided fractionation. Typical yields of pure bioactive molecules under research settings are low: Depending on the nature of the active molecule, 1 kg of dried plant material can yield as little as 1–10 μg of pure compound. Thus, on detection of a hit in a bioassay, recollection of a bulk sample (10–40 kg) is often required to produce sufficient pure substance for structural determination.

The advent of tandem mass spectroscopy with time-of-flight analysis coupled with computer-intensive pattern analysis has made the detection of small molecule metabolites much more accessible to investigators. By comparing genotypes with and without a known pharmacological activity, subtraction of all possible metabolites can result in identification of active molecules. However, as common to most emerging fields, many difficulties in both technique and analysis remain. Similarly, an analysis of the transcriptome, a set of all RNA molecules produced, can facilitate the analysis of genes and gene products involved in a disease. This in turn can be used as both a bioassay and an isolation technique for natural products that alter the transcriptome by modulating gene expression. Richard Weindruch and his colleagues in Madison, Wisconsin, are exploring how to achieve antiaging effects analogous to severe caloric restriction via changes to the transcriptome induced by natural products. They report that consumption of resveratrol, a molecule derived from red wine, mimics the gene expression of mice subjected to a severely calorie restricted diet (Barger *et al.*, 2008).

Structural Determination

Once a pure bioactive molecule is obtained through fractionation or chromatographic techniques, the next step is to determine its molecular structure. A definitive technique for determining the structure (if the purified molecule can be crystallized) is X-ray crystallography, but more common techniques involve the use of high-pressure liquid chromatography, gas chromatography (GC), mass spectrometry (MS), or, for novel structures, nuclear magnetic resonance (NMR) spectroscopy. Triple quadrupole mass spectrometry/mass spectrometry has been the most useful in detecting small molecules in complex matrices if purification is not possible (Domon and Aebersold, 2006). Most modern MS instruments have a computer-based library documenting molecular weights of known molecules; observed peaks are quickly and easily compared to those in the computer library. This technique is particularly useful for forensic toxicologists who do not typically expect to find novel molecules in their work. Structural determination in most biodiversity studies, however, relies on NMR since novel rather than known structures are often discovered during the course of pharmaceutical research. The frequencies and resolution capabilities of NMR machines differ, as do the types of experiments possible on NMR equipment, but of particular interest to natural product investigators is the INADEQUATE experiment, which, through supercomputer analysis, can resolve structures for molecules in substances that are parts of mixtures rather than pure samples. X-ray crystallography can resolve structures that are difficult to resolve through NMR.

Rational Screens of Biodiversity

Sometimes, such as in the case of penicillin, discovery of new pharmaceuticals from biodiversity is the result of serendipity. More often, though, success in natural product research is the result of carefully designed screening processes. Such planned search strategies to select appropriate organisms for study are termed "rational screens," of which there are several varieties: random, phylogenetic, ecological, and ethnobotanical.

Random Screens

Random screens, despite their name, involve a great deal of planning and design. To be successful, they must have a clearly defined disease target and a clear and unambiguous bioassay. One of the best examples of a successful random screen in recent times is the effort of the Natural Product Branch of the U.S. National Cancer Institute (NCI) to evaluate plants and animals for anticancer activity. Using a variety of *in vitro* cancer cell cultures as bioassays, the NCI screened more than 35,000 plant accessions and 6000 marine organisms for activity against human cancers. Although the screen was not truly random in that it was designed to sample a wide range of geographical and taxonomic diversity, it allowed the NCI (with a variety of contracting collecting institutions) to assemble a large archive of materials that could be retested as bioassays continued to evolve. The geographical reach of the NCI was truly broad and led to pioneering efforts to negotiate international protocols for biodiversity prospecting. The NCI "letter of intent," which collaborators used to obtain collecting permits in foreign countries, pledges equitable sharing of data and economic benefits with host countries. This letter has emerged as a model for recent international agreements.

Although the low hit rate of the NCI random screens led to criticism of the approach, few can argue with the stellar success of the NCI program in producing what may be one of the most important chemotherapeutic agents discovered in the latter part of the twentieth century, taxol, which was initially isolated from the Pacific yew tree (*Taxus brevifolia* (Taxaceae)) and which has been licensed by the NCI to Bristol Meyers and approved by the Food and Drug Administration (FDA) for the treatment of ovarian and breast cancers. However, because of the low success rates and high costs of collection, random screens are currently suitable only for large institutions, even given the advent of high-throughput bioassays.

Phylogenetic Screens

Phylogenetic screens involve the pharmacological testing of related groups of organisms. Increased precision in elucidating phylogenies, largely due to rapid computer programs for cladistic analysis and the advent of molecular techniques for phylogenetic determination, facilitates the identification of relatives of any species showing pharmacological value. Phylogenetic screens, albeit in a crude sense, have long been utilized. For example, plants in the Apocynaceae, or milkweed family, have always merited special attention due to the family's abundance of alkaloid-producing species such as *Catharanthus*, which produces the antileukemia drug vincristine. Only recently have modern techniques encouraged investigators to study close relatives of a species for either (1) increased abundance of an important bioactive compound (such as species of *Taxus* for the presence of taxol) or (2) natural homologs of known pharmaceuticals (such as species of *Catharanthus* that may produce variant forms of vincristine or vinblastine).

Ecological Screens

An ecological screen depends on the environmental setting of the organism for clues as to its possible pharmaceutical value. Recently, such screens have focused on extremophiles – organisms that survive under extreme conditions of high temperatures, acidity, alkalinity, or salinity. For example, the polymerase chain reaction, invented by Kary Mullis at Cetus Corporation and widely used for genetic studies, medical diagnosis, or gene therapy, depends on an enzyme, Taq polymerase, a heat-stable DNA polymerase derived from a bacterium discovered in the thermal springs of Yellowstone National Park by Thomas Brock of the University of Wisconsin. The bacterium that produces the enzyme, *Thermus aquaticus*, can withstand temperatures of up to 95 °C, and its enzymes are similarly heat resistant, an important feature for the rapid thermal cycles used in the polymerase chain reaction.

Ecological screens have also been conducted underneath the sea near hydrothermal vents where temperatures can reach more than 350 °C. The enzyme Pfu DNA polymerase, derived from the Archaeum *Pyrrhus furiosus* and discovered by Karl Stetter of the University of Regensburg during exploration of undersea thermal vents, is of increasing interest because of its precision in polymerase chain reactions due to a built-in DNA repair mechanism. *P. furiosus* can survive temperatures more than 100 °C and is resistant to high pH and radioactivity. Enzymes derived from other organisms that survive conditions of extreme cold, alkalinity, acidity, or salinity are of increasing interest to the biotechnology industry (Madigan and Marrs, 1997).

Other ecological screens focus on searches for antifungal compounds from plants that grow in moist tropical environments (successes in this area have been achieved by Alice Clark at the University of Mississippi) as well as searches for inexpensive molluscicidal compounds useful in the fight against schistosomiasis. Kurt Hostettmann and his colleagues at the University of Lausanne discovered saponin-like compounds in several African tree species that show intense toxicity to schistosomiasis-carrying snails (Teryvaud *et al.*, 2000). They are investigating the possibility of supplying villagers with seeds and the simple techniques to prepare the crude solutions of these compounds, which could be used to dose village bathing and washing areas.

Ecological screens can also identify those few animals that produce useful pharmaceuticals. The Gila monster, a venomous lizard of the southwestern USA, produces venom in modified salivary glands resulting in severe pain and a drop in blood pressure in their prey. In 2005, a synthetic version of the protein exenatide derived from saliva was approved for treatment of type 2 diabetes. Marketed as Byetta, the drug regulates

glucose and causes decrease in appetite, leading to weight loss (DeFronzo *et al.*, 2005). Similarly, studies of toxins used by marine cone shells have led to the discovery of neuropeptides that show great potential in treating chronic pain (Olivera, 2006). The drug Prialt (generically called ziconitide) from the venom of *Conus majus* was approved in 2004 by the FDA for treatment of severe, chronic pain.

Ethnobotanical Screens

An ethnobotanical screen is based on the finding that indigenous peoples through generations have accumulated useful information about pharmaceutically active plants and animals (Cox, 2000a). Since ethnobotanical leads have resulted in the discovery of the active compounds for 25% of all prescription drugs, there is little question as to the efficacy of ethnobotanical approaches to drug discovery (Cox and Balick, 1994). However, although there are strengths to ethnobotanical screens, significant limitations also exist.

In general, cultures with three characteristics seem most likely to have discovered pharmacologically active plants: (1) residency within an area of high biodiversity, (2) an extended history of residence within the area, and (3) a cultural mechanism for accurate transmission of ethnomedical information from generation to generation. Under the first criterion, the Kayapo Indians of Brazil merit more attention than the Aleuts in Alaska because of the greater biodiversity of tropical Brazil. Under the second criterion, the aboriginal peoples of North Arnhem Land who have been resident in Australia for thousands of years would be of more interest than the Pitcairn islanders who arrived in their island home only a few centuries ago. Under the third criterion, the Samoan islanders, with a precise matrilineal system of transmitting healing knowledge from generation to generation, would be of more interest than faith healers in the Philippines, who depend on dreams for their knowledge of healing plants.

In an ethnobotanical screen, healing information is recorded by a linguistically adept ethnobotanist trained both in the techniques of ethnographic interviews and in field botany. Often, there is little overlap between indigenous and Western disease concepts. For example, there is no Western equivalent of "susto" (a type of soul loss with diverse symptoms) that afflicts indigenous peoples in Central America, so the ethnobotanist, often in working with a physician, must be adept at mapping signs and symptoms of indigenous disease states onto concepts of illness understood by Western peoples. Those plants identified by the indigenous peoples as containing healing properties are collected and identified, with samples prepared for bioassay-guided fractionation. In addition, careful notes about possible adverse reactions, dosages, preparation, and application techniques are carefully prepared to assist other members of the drug discovery team. Although the ethnobotanical approach seems to work well for diseases (such as skin fungus or acute viral syndrome) that are easily recognizable by indigenous peoples, other diseases, such as brain tumors or lymphoma, which require advanced techniques for diagnosis, show less hope of being successively treated by drugs discovered during ethnobotanical research.

There are different approaches to ethnobotanical screens. In the consensus approach, developed by Brent Berlin at the University of Georgia, numerous villagers are interviewed with regard to their knowledge of healing plants. Those medicinal plants that are repeatedly mentioned in interviews are prioritized for testing due to their high saliency in the culture. A variant of this approach is to prioritize for study those plants used in common by different villages or even cultures; an ethnobotanical screen based on the latter approach was successfully used by Napo Pharmaceuticals in the discovery of Crofelemer, an antidiarrheal drug derived from *Croton lechleri* (Euphorbiaceae), which successfully completed double-blinded phase III clinical trials.

A different approach to consensus techniques, developed by Michael Balick of the New York Botanical Gardens and Paul Cox of the Institute for Ethnomedicine in Jackson Hole, relies on specialist knowledge known only to healers rather than on knowledge held in common by villagers. In this approach, the healers who are most highly regarded by their societies, who have the greatest knowledge of diseases, and who use the largest repertoire of plant species in their therapeutic practices are asked to rank the efficacy of their healing plants. In a careful trial in Belize conducted in conjunction with the NCI, Balick found that such "powerful plants" identified by healers had a far greater likelihood of generating hits in bioassays.

Nature of Biodiversity-Derived Pharmaceuticals

Microbial Products

In 1929, Alexander Fleming discovered that a petri dish of *Staphylococcus* had become contaminated with a *Penicillium* fungus: Around the fungus in the culture medium was a zone resistant to *Staphylococcus* growth. At Oxford University, a policeman suffering sepsis from a scratch with a rose thorn was successfully treated with a purified extract of *Penicillium*. During World War II, the antibiotic penicillin was isolated. This led to dramatic treatment of infections from wounds and other trauma. Since that time, more than 6000 different antibiotics have been isolated, with more than 1000 having been commercially produced (Samuelsson, 1992). Only a single "natural" penicillin has ever been prescribed as a drug, but more than 100 semisynthetic penicillins have been produced by adding other carboxylic acids to the basic penicillin structure (Sameulsson, 1992). The structure of penicillin, a peptide with a thazolidine ring, is similar to the structure of the cephalosporins, in which the thazolidine ring is substituted by a dihydrothiazine ring. *Bacillus brevis* and *Bacillus licheniformis* have been used to produce gramicidin and bacitracin, respectively.

Many new antibiotics and other pharmaceutical products have been obtained from soil samples that are carefully cultivated in sterile petri dishes and fermentation vats. One of the more successful products to be produced from a soil culture is cyclosporin, a cyclic peptide produced from the fungus *Cylindrocarpon lucidum*, originally derived from a Norwegian soil sample. Cyclosporin, a powerful immunosuppresant, was initially seen to have little therapeutic value and remained archived as a refrigerated fungal culture at

Sandoz Pharmaceuticals (now Novartis) in Basel, Switzerland, until mycologist Michael Dreyfuss convinced the management that the product had utility. Today, nearly every heart or kidney transplant patient in the world takes a daily dose of cyclosporin or a similar compound to prevent rejection of the transplanted organ.

Soil samples continue to yield new and interesting bacteria and fungi, whereas exploration of extreme habitats promises to yield important new useful compounds, including those produced by Archea, the recently discovered third domain of life, which produced the enzyme Pfu DNA polymerase. Using recombinant DNA technology, microorganisms can be employed to produce pharmaceutical compounds that they do not ordinarily produce in nature.

Yet the need for new antibiotics looms as a major human health crisis. The routine feeding of antibiotics to livestock to increase their weight, and indiscriminate the use of antibiotics by people around the world, have led to the evolution of antibiotic resistance among many species of bacteria. Most serious is community-acquired methicillin-resistant *Staphylococcus aureus* (CA-MRSA), which is approaching epidemic levels in US hospitals, prisons, athletic teams, and newborn nurseries. CA-MRSA causes rapidly progressive pneumonia and tissue necrosis. Entirely new classes of antibiotics are required to treat CA-MRSA infections that can be quickly fatal. Cyanobacteria, which produce a range of toxic compounds, such as BMAA – which has been linked to motor neuron disease (Cox *et al.*, 2005) – may be one possible source for new antibiotic and antiviral compounds. Recent studies by William Gerwick and his colleagues at the Scripps Institute of Oceanography have yielded a number of anticancer drug candidates from cyanobacteria (Simmons *et al.*, 2005).

Marine Organisms

Marine organisms, particularly invertebrates, are known to produce extraordinarily toxic compounds. Currently, three compounds isolated from marine resources are in preclinical development as possible anticancer agents by the NCI. The marine algal species *Portieria hornemannii* from the Philippines has yielded a drug called halomon, a species of the sponge genus *Lissodendoryx* from New Zealand has produced halichondrin B, and a Caribbean tunicate *Ecteinascidia turbinata* has yielded Trabectedin (ecteinascidin 743), which has been approved in Europe, where it is sold by Zeltia and Johnson & Johnson under the brand name Yondelis for the treatment of soft tissue sarcoma. It is currently being evaluated for treatment of breast, prostate, and pediatric cancers. Another promising marine pharmaceutical is bryostatin from the Californian bryozoan *Bugula neritina.* Bryostatin inhibits melanoma, reticulum cell sarcoma, and lung carcinoma. The discovery of bryostatin will likely lead to new therapeutic concepts in chemotherapy. As previously mentioned (*see* Ecological Screens), Prialt, a neuropeptide derived from cone snail venom, is used to treat intractable pain. Originally marketed by Elan Pharmaceuticals, Prialt was recently acquired by Azur Pharma. It currently is the only nonopiate intrathecal therapy approved for severe chronic pain. Currently, 6000 marine samples are archived at the NCI for testing, and there is continuing industrial interest in natural products derived from marine organisms (Molinski *et al.*, 2009).

Terrestrial Animals

Terrestrial animals have historically played an important role in the production of certain therapeutic hormones and steroids. Progesterone was first isolated from pig ovaries in 1934. Animals were eclipsed by plants, however, as a source of steroidal precursors when in 1940 Russell Marker discovered that Mexican yams of the genus *Dioscorea* could be utilized to produce diosgenin. There has also been a long-term research interest in venoms produced by snakes as sources of pharmaceuticals. Ancrod, a drug used to treat circulatory diseases, is isolated from the venom of the Malayan pit viper *Agkistrodon rhodostoma.* Recent research, such as that conducted by Matt Ritter and Francis Markland at the University of Southern California, indicates that proteins such as contortrostatin found in certain snake venoms can treat cancer tumors (Ritter *et al.*, 2000). And as previously mentioned (*see* Ecological Screens), Byetta, a synthetic form of the peptide exenatide derived from Gila monster saliva, is transforming the treatment of type 2 diabetes. The poisonous secretions of some toads and frogs contain bufotenin. Ethnozoologist Wade Davis has found that secretions of species of *Bufo* were used as psychotropic substances by indigenous peoples in Central America. Epibatidine, from the poison dart frog, is being developed as a possible analgesic. It is likely that other venoms and stings will produce new pharmacological insights.

Terrestrial Plants

Plant-derived drugs are used for a broad spectrum of diseases and include quinine (for malarial suppression), digitalis (for treatment of rapid atrial fibrillation), vincristine (for treatment of pediatric leukemia), tubocurarine (a muscle relaxant used in anesthesia), pilocarpine (for the treatment of glaucoma), and γ-strophanthin (for congestive heart failure) (**Table 1**). However, in the latter quarter of the twentieth century, very few plant-derived drugs were released to market, even though there was an increase in consumer interest in North America, Europe, and Japan in "natural" plant medicines. Given that less than 1% of the world's plants have ever been carefully studied for pharmacological activity, it is puzzling that pharmaceutical firms are not rigorously investigating plants as the sources of new pharmaceuticals.

Flowering plants and ferns have traditionally provided the bulk of biodiversity-derived pharmaceuticals, but recent advances in molecular biology and combinatorial chemistry have reduced the pharmaceutical industry's ardor in studying plants. The popular mythology of industry-supported botanists combing the rainforests for new plant-based leads is simply not true. Currently, there is not a single major pharmaceutical firm that lists among its employees a Ph.D. botanist or an ethnobotanist who works in a drug-discovery program based on plants. Despite the lack of interest by the pharmaceutical industry, both the promise and the impact of plant-derived pharmaceuticals are profound. The World

Table 1 Fifty drugs discovered from ethnobotanical leads

Drug	*Medical use*	*Plant source*
Ajmaline	Heart arrhythmia	*Rauvolfia* spp.
Aspirin	Analgesic, inflammation	*Filipendula ulmaria*
Atropine	Ophthalmology	*Atropa belladonna*
Benzoin	Oral disinfectant	*Styrax tonkinensis*
Caffeine	Stimulant	*Camellia sinensis*
Camphor	Rheumatic pain	*Cinnamomum camphora*
Cascara	Purgative	*Rhamnus purshiana*
Cocaine	Ophthalmic anesthetic	*Erythroxylum coca*
Codeine	Analgesic, antitussive	*Papaver somniferum*
Colchicine	Gout	*Colchicum autumnale*
Demecolcine	Leukemia, lymphomata	*Colchicum autumnale*
Deserpidine	Hypertension	*Rauvolfia canescens*
Dicoumarol	Thrombosis	*Melilotus officinalis*
Digoxin	Atrial fibrillation	*Digitalis purpurea*
Digitoxin	Atrial fibrillation	*Digitalis purpurea*
Emetine	Amebic dysentery	*Cephaelis ipecacuanha*
Ephedrine	Bronchodilator	*Ephedra sinica*
Eugenol	Toothache	*Syzygium aromaticum*
Gallotannins	Hemorrhoid suppository	*Hamamelis virginiana*
Hyoscyamine	Anticholinergic	*Hyoscyamus niger*
Ipecac	Emetic	*Cephaelis ipecacuanha*
Ipratropium	Bronchodilator	*Hyoscyamus niger*
Morphine	Analgesic	*Papaver somniferum*
Noscapine	Antitussive	*Papaver somniferum*
Papain	Attenuate mucous	*Carica papaya*
Papaverine	Antispasmodic	*Papaver somniferum*
Physostigmine	Glaucoma	*Physostigma venenosum*
Picrotoxin	Barbiturate antidote	*Anamirta cocculus*
Pilocarpine	Glaucoma	*Pilocarpus jaborandi*
Podophyllotoxin	Condylomata acuminata	*Podophyllum peltatum*
Proscillaridin	Cardiac malfunction	*Drimia maritima*
Protoveratrine	Hypertension	*Veratrum album*
Pseudoephedrine	Rhinitis	*Ephedra sinica*
Psoralen	Vitiligo	*Psoralea corylifolia*
Quinidine	Cardiac arrhythmia	*Cinchona pubescens*
Quinine	Malaria prophylaxis	*Cinchona pubescens*
Rescinnamine	Hypertension	*Rauvolfia serpentina*
Reserpine	Hypertension	*Rauvolfia serpentina*
Sennoside A, B	Laxative	*Cassia angustifolia*
Scopolamine	Motion sickness	*Datura stramonium*
Stigmasterol	Steroidal precursor	*Physostigma venenosum*
Strophanthin	Congestive heart failure	*Strophanthus gratus*
Teniposide	Bladder neoplasms	*Podophyllum peltatum*
THC	Antiemetic	*Cannabis sativa*
Theophylline	Diuretic, asthma	*Camellia sinensis*
Toxiferine	Surgery, relaxant	*Strychnos guianensis*
Tubocurarine	Muscle relaxant	*Chondrodendron tomentosum*
Vinblastine	Hodgkin's disease	*C. roseus*
Vincristine	Pediatric leukemia	*C. roseus*
Xanthotoxin	Vitiligo	*Ammi majus*

Source: Reproduced from Balick M and Cox P (1997) *Plants, People, and Culture: The Science of Ethnobotany*, pp. 34–35. New York: Scientific American Library.

Health Organization estimates that 85% of the world's population depends directly on plants for medicine, and more than 25% of current prescription drugs have at least one active component derived from a flowering plant. The former statistic, derived principally from developing countries, has provided significant opportunities to contribute to world health by assisting different nations to evaluate the safety and efficacy of their own pharmacopoeias. Currently, China, Mexico, Thailand, and Nigeria have incorporated traditional plant-based medicines directly into primary health care. An example of the possibilities inherent in this approach has been demonstrated in Thailand, where a potent antiinflammatory compound from the beach plant *Ipomoea pes-caprae* (Convolvulaceae) long used by fishermen to treat stings from

the Portuguese man-of-war was isolated and purified using bioassay-guided fractionation with its structure being determined using NMR spectroscopy. However, rather than testing a pure compound, the Thai researchers conducted a careful, double-blind controlled study of a crude tincture of the plant. It was found to be both safe and efficacious. The resultant tincture can now be bought in any Thai drugstore for a price far less than that of a synthetic drug. The widespread European acceptance of the German Commission E herbal medicines in primary health care is a likely portent of broader acceptability of phytopharmaceuticals in the industrialized world (Blumenthal, 1998).

However, several plant-based pharmaceuticals have been recently released or are under current preclinical or clinical evaluation. Taxol, discovered during the NCI random screen, has received FDA approval for treatment of breast and ovarian cancers. For AIDS, currently five plant-derived compounds are in preclinical development. Derived from the resin of *Calophyllum lanigerum* (Guttiferae) in Malaysia, calanolide A has shown intense activity against HIV-1, as has costatolide derived from a related tree *Calophyllum teysmannii* from the same region. Michellamine B from the leaves of the Cameroon vine *Ancistrocladus korupensis* (Ancistrocladaceae) has also shown antiHIV activity, as has conocurvone derived from *Conospermum* (Proteaceae) shrubs found in Western Australia. Because of its use in indigenous medicine, the fifth member of this anti-AIDS quintet, prostratin, perhaps deserves in-depth discussion as a case study of recent developments in this field.

For many years, Samoan healers have used the stem wood of a small understory rainforest tree, *Homalanthus nutans* (Euphorbiaceae), to treat a disease known as "fiva samasama," characterized by yellowing of the eyes, dark urine, jaundice, and fevers; this disease was subsequently determined to be hepatitis. To prepare the remedy, the wood of *Homalanthus* is macerated into a clean cloth, which, like a tea bag, is immersed into boiling water. After steeping, the contents of the tea bag are discarded, and the resultant tea is drunk by the patient. Samples of the healer preparations and the plants were tested against an *in vitro* HIV-1 bioassay in a collaboration between Paul Cox and NCI researchers Kirk Gustafson, John Cardellina, II, John Beutler, Peter Blumberg, Gordon Cragg, Michael Boyd, and others at the Natural Products Branch of the NCI (Gustafson *et al.*, 1992). The healer mixture and the crude plant extracts were found to protect cells from death by HIV-1, even though there was no evidence of reverse transcriptase inhibition or other marked inhibition of the mixtures to the virus. Bioassay-guided fractionation yielded a pure compound, prostratin (12-deoxyphorbol-13-acetate). Although phorboIs, similar in structure to prostratin, are potent tumor promoters, *in vivo* studies by Peter Blumberg at the National Institutes of Health showed prostratin to be an antipromoter, even though it activates protein kinase C. The NCI advertised prostratin as a potential new component of a combination therapy for AIDS. The NCI required any pharmaceutical developer of prostratin to negotiate directly with the Samoan government for a fair and equitable return of a portion of any royalties. Such negotiations between the Samoan government and the AIDS Research Alliance (ARA) led to an agreement to return 20% of ARA's profits from prostratin to Samoa to be divided between the government, the village, and the healers. Similar negotiations led the University of California, Berkeley, to sign an agreement with the Samoan government to return 50% of the commercial proceeds of the prostratin genes to Samoa (Cox, 2001). In addition, preroyalty benefits provided to the village through the conservation organization Seacology have now exceeded $500,000. Although prostratin is currently being developed as an antiHIV drug, its utility in the treatment of hepatitis has yet to be determined.

Biodiversity and Ethics

The Convention on Biodiversity

The Convention on Biological Diversity (CBD), often referred to as the Rio Treaty, provided a broad basis in international law for the discovery and development of pharmaceutical compounds from biodiversity. To promote the conservation and study of biodiversity, the parties to the CBD agreed that each nation has sovereignty over its own biological resources. Collection and development of these resources can thus proceed only with express permission of the national government. What initially was seen as a mechanism to accelerate investigation and conservation of biological resources by facilitating international cooperation has been criticized by some as an impediment to research. Many countries, fearing uncompensated exploitation of their national biological resources, have either closed their doors to international scientific exploration or caused the process of obtaining research permits to be extraordinarily slow and difficult. Given the history of colonial exploitation of biological resources, with little thought given to equitable sharing of benefits of discoveries based on biodiversity, such reaction by developing countries is understandable. If the processes of biodiversity were static and extinction quiescent, closing national boundaries to scientific investigation would have little long-term consequence; however, given the rapid rate of extinction (the IUCN estimates that one-eighth of the world's plants are threatened or endangered), nations that close their doors to scientists may unwittingly ensure that their biodiversity will vanish without any significant appraisal of its economic value. An interesting exercise would be to compare the relative difficulties of obtaining permits to collect and pharmacologically analyze 50 mg plant samples from various countries versus the difficulties in obtaining permits to clear cut and export entire rainforests: Too often, those who destroy biodiversity are facilitated while those who seek to study it are hindered.

Given the relative lack of interest of the pharmaceutical industry in investigating biodiversity as a source of new pharmaceuticals, and the occasionally unreasonable expectations placed on pharmaceutical firms for large initial payments and high percentages of royalties, many pharmaceutical firms that once had nascent interest in natural product chemistry have retreated to computer modeling and combinatorial chemistry as the sole engines of pharmaceutical discovery. Although this situation may change if new, important drugs are discovered from biodiversity, it is crucial that

investigators adhere to high ethical standards. Training local scientists, obtaining informed consent from village chiefs or elders before initiating research, depositing duplicates of all collected specimens with local or regional herbaria, and vigorously protecting indigenous intellectual property all help to promote understanding and collaboration. Investigators who fail to heed such standards or, worse, ignore local laws and international treaties, imperil not only their own research programs but also conservation of the very biodiversity they purport to study.

An interesting, and perhaps unintended, consequence of the CBD has been its impact on the rights of indigenous peoples. By granting sovereignty over all biological resources to national governments, the convention in a sense can disenfranchise indigenous peoples, particularly in countries with a history of mistreating them. Is it reasonable that a plant that has been developed and studied through generations of indigenous healers should be considered the sole property of a distant (and perhaps hostile) national government? However, by stating in article 8(j) that indigenous peoples bear some consideration, the architects of the Convention on Biodiversity provided indigenous peoples with one of the first broadly accepted international treaties that explicitly recognizes their existence and aspirations. Article 8(j) requires signatory nations to (1) respect, preserve, and maintain traditional knowledge; (2) promote wide application of traditional knowledge; and (3) encourage equitable sharing of benefits from traditional knowledge. Many indigenous groups hope that their voices will increasingly be heard and considered at meetings of the parties to the convention.

In October 2010, the COP 10 Biodiversity meetings in Aichi-Nagoya, Japan, resulted in the adoption of the Protocol on Access and Benefit Sharing by the parties to the convention. It is hoped that this international agreement will continue to encourage analysis of biodiversity for new pharmaceuticals and the equitable sharing of benefits with biodiverse-rich countries and indigenous peoples. Benefits in the COP 10 agreement include technology transfer, research results, and royalty sharing from patented products.

Bilateral Agreements on Pharmaceutical Research

There have been attempts to directly negotiate with national governments for rights to perform pharmaceutical research on biodiversity. One of the best known of these arrangements was a remarkable agreement entered into by the government of Costa Rica and Merck Pharmaceuticals. Under the INBio Agreement, Merck received rights to survey the flora and fauna of Costa Rica in return for a share of the proceeds of any new discovery and for a payment of $1 million used to support the training of local investigators and development of a national inventory of Costa Rican biodiversity. Costa Rica has been very innovative in training local people in biological survey techniques. Termed "parataxonomists," these local people, often with profound folk knowledge of organisms in their environment, are trained in modern techniques of collection and vouchering. By devoting the bulk of the Merck funds to biodiversity research, the government of Costa Rica has guaranteed that the people of Costa Rica and indeed the world will benefit for many years.

Indigenous Intellectual Property Rights

In ethnobotanical approaches to drug discovery, it is important that the contributions of indigenous people be appropriately recognized and compensated. Often, indigenous values, particularly views of fairness and equability, differ dramatically from those of Western societies. Many Western societies highly value individualism, rationalism, and candor in expression, whereas in many indigenous societies, values such as respect for village elders, proper regard for ancestors and deities, and recognition of the sacred nature of living organisms play a crucial role. Frequently, indigenous peoples also have very distinct views about appropriate behavior and dress within a village.

It is important that Western investigators dealing with indigenous peoples show appropriate respect for indigenous values. Indigenous mores and folkways should not be violated by investigators, and the position of traditional leaders, chiefs, and village councils should be respected rather than eroded. This may mean, for example, that Western scientists do not publish biodiversity information considered secret by indigenous people or do not venture into areas considered to be taboo. Respect can be communicated by completely informing such leaders and village councils about the scope, nature, and possible impact of biodiversity research. Informed consent of traditional leaders is crucial when discussions are held concerning fair and equitable returns for indigenous intellectual property: Traditional leaders cannot make informed decisions concerning biodiversity research unless they possess accurate information concerning the nature and value of the research. Rather than imposing Western values of fairness and equability on indigenous peoples, investigators should regard indigenous peoples as coequals in any negotiations and seek solutions that meet indigenous rather than Western concepts of fairness and equability. Often, traditional leaders are not familiar with Western culture or conversant in Western languages, but the temptation to use Western-educated elites as surrogates for traditional leaders in negotiations should be avoided (Cox, 2000b).

For many indigenous peoples, conservation is a keenly felt need, a means of both perpetuating their culture and demonstrating respect to Earth and to their ancestors. Innovative approaches to assisting indigenous peoples in conservation can be successful if indigenous peoples truly control the conservation initiatives. The Terra Nova biomedical reserve in Belize and the Falealupo Rain Forest Reserve in Samoa are examples of indigenous conservation begun by Western scientists that are now completely controlled by indigenous peoples (Cox, 2001). Culturally appropriate acknowledgment of indigenous collaborators such as including publication of scientific abstracts and reviews in indigenous languages and coauthorship of papers and books with indigenous collaborators are examples of the best practices currently employed by ethnobotanists. Protection of indigenous financial interests often requires vigorous efforts on the part of Western scientists, who need to assist indigenous leaders in dealing with

lawyers, commercial firms, or governments that may not completely appreciate indigenous property rights. Entering into any collaborative project, including ethnobotanical research, with indigenous peoples imposes a responsibility on the part of scientists to protect indigenous rights. It should not be surprising, then, that some of the most persuasive Western advocates of indigenous rights have been scientists who have lived and worked closely with indigenous peoples.

Rights of the Sick and Afflicted

In an era in which life itself increasingly appears to be a commercial commodity, it might appear quaint to raise the possibility that stakeholders in biodiversity research include others than sovereign nations, pharmaceutical firms, and indigenous peoples. However, moral rights are not always harmonious with property rights: Most societies, for example, believe that small children have an inalienable claim on their parents for support, even though children do not possess legal title to any of their parents' possessions. Similarly, the constitutions of many nations now recognize rights of mothers, children, and other citizens to adequate nutrition and minimal health care. In former eras, the sick and afflicted were believed to have certain claims on healing plants and substances. Although it is unlikely in the modern market economy that the sick and afflicted will ever be granted a significant voice in biodiversity research, ill people remain often unseen but tangible stockholders in biodiversity-based pharmaceutical research. The sick and afflicted, and their families and societies that support them, stand to lose if the world's biodiversity is thoughtlessly destroyed before it can be evaluated for pharmaceutical potential.

See also: Biodiversity, Human Well-Being, and Markets. Biodiversity-Rich Countries. Bioprospecting. Diseases, Conservation and. Diversity, Molecular Level. Economic Value of Biodiversity, Measurements of. Marine Ecosystems. Market Economy and Biodiversity. Property Rights and Biodiversity. The Value of Biodiversity. Thermophiles, Origin of

References

Arber A (1953) *Herbals, Their Origin and Evolution*. Cambridge, UK: Cambridge University Press.

Baker BJ, Comolli LR, Dick GJ, *et al.* (2010) Enigmatic, ultrasmall, uncultivated Archaea. *Proceedings of the National Academy of Sciences* 107: 8806–8811.

Barger JL, Kayo T, Vann JM, *et al.* (2008) A low dose of dietary resveratrol partially mimics caloric restriction and retards aging parameters in mice. *PloS ONE* 3(6): e2264.

Blumenthal M (ed.) (1998) *The Complete German Commission E Monographs: Therapeutic Guide to Herbal Medicines*. Austin, TX: American Botanical Council.

Buenz EJ, Schnepple DJ, Bauer BA, *et al.* (2004) Techniques: Bioprospecting historical herbal texts by hunting for new leads in old tomes. *Trends in Pharmacological Sciences* 25(9): 494–498.

Cox PA (2000a) Will tribal knowledge survive the millennium? *Science* 287(5450): 44.

Cox PA (2000b) A tale of two villages: Culture, conservation, and ecocolonialism in Samoa. In: Zerner C (ed.) *People, Plants, and Justice: The Politics of Nature Conservation*, pp. 330–344. New York: Columbia University Press.

Cox PA (2001) Ensuring equitable benefits: The Falealupo covenant and the isolation of anti-viral drug Prostratin from a Samoan medicinal plant. *Pharmaceutical Biology* 39(Supplement): 33–40.

Cox PA and Balick MJ (1994) The ethnobotanical approach to drug discovery. *Scientific American* 270(6): 82–87.

Cox PA, Banack SA, Murch SJ, *et al.* (2005) Diverse taxa of cyanobacteria produce *β-N*-methylamino-L-alanine, a neurotoxic amino acid. *Proceedings of the National Academy of Sciences* 102(14): 5074–5078.

DeFronzo RA, Ratner RE, Han J, *et al.* (2005) Effects of exenatide (Exendin-4) on glycemic control and weight over 30 weeks in metformin-treated patients with type 2 diabetes. *Diabetes Care* 28(5): 1092–1100.

Domon B and Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312: 212–217.

Grifo F and Rosenthal J (eds.) (1997) *Biodiversity and Human Health*. Washington, DC: Island Press.

Gustafson KR, Cardellina JH, McMahon JB, *et al.* (1992) A non-promoting phorbol from the Samoan medicinal plant Homalanthus nutans inhibits cell killing by HIV-I. *Medicinal Chemistry* 35: 1978–1986.

Hentzer M, Eberl L, Nielsen J, and Givskov M (2003) Quorum sensing: A novel target for the treatment of biofilm infections. *BioDrugs* 17(4): 241–250.

Joyce G (2002) The antiquity of RNA-based evolution. *Nature* 418: 214–221.

Madigan MT and Marrs BL (1997) Extremophiles. *Scientific American* 276(4): 82–87.

Molinski TF, Dalisay DS, Lievens SL, and Saludes JP (2009) Drug development from marine natural products. *Nature Reviews Drug Discovery* 8: 69–85.

Newman DJ and Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *Journal of Natural Products* 70(3): 461–477.

Newman DJ, Cragg GM, and Snader KM (2003) Natural products as sources of new drugs over the period 1981–2002. *Journal of Natural Products* 66: 1022–1037.

Olivera BM (2006) Conus peptides: Biodiversity-based discovery and exogenomics. *The Journal of Biological Chemistry* 281: 31173–31177.

Ritter MR, Zhou Q, and Markland FS (2000) Contortrostatin, a snake venom disintegrin, induces αvβ3-mediated tyrosine phosphorylation of CAS and FAK in tumor cells. *Journal of Cellular Biochemistry* 79(1): 28–37.

Samuelsson G (1992) *Drugs of Natural Origin*. Stockholm: Swedish Pharmaceutical Press.

Simmons TL, Andrianasolo E, McPhail K, Flatt P, and Gerwick WH (2005) Marine natural products as anticancer drugs. *Molecular Cancer Therapeutics* 4: 333–342.

Treyvaud V, Marston A, Dyatmiko W, and Hostettmann K (2000) Molluscicidal saponins from Phytolacca icosandra. *Phytochemistry* 55(6): 603–609.

Phenological Shifts in Animals Under Contemporary Climate Change

Marcel E Visser, Netherlands Institute of Ecology, The Netherlands

© 2013 Elsevier Inc. All rights reserved.

Glossary

Circannual rhythms Recurrence of a phenomenon in cycles of approximately 1 year.
Phenological shifts Change over years in the mean of the phenological value of a trait within a population.
Phenology The study of the seasonal timing of natural events, such as the return date of migrant birds or the initiation of reproduction.
Phenotypic plasticity The plasticity of a genotype to express a different phenotype under different environmental conditions.
Reaction norm The relationship between the phenotypic expression of a single genotype and the environmental conditions under which the phenotype is shaped.

Introduction

Over the past three decades, global climate change has had a clear impact on the biology of wild species. One of the most pronounced effects has been on the phenology, or seasonal timing, of species. There are numerous examples of animal and plant species that have shifted their phenology to an earlier date in spring, such as the return date of avian migrants, the termination of hibernation periods, or the hatching date of overwintering insects. Before summarizing these observed shifts in animal species, the article first discusses a general framework for phenology, which can be used to understand why phenology is likely to be affected by climate change and the mechanisms by which this may occur.

Animals go through distinct annual life-history cycles. There are times of the year when they reproduce, times when they molt their feathers or change their coat, and times when they hibernate or migrate to overwintering areas. These life history stages often have a strict order and happen at relatively fixed periods within a year: In harsh winter, marmots hibernate, and in spring, when food is plentiful, they have their offspring. This annual life history cycle is tuned to the variation in environmental conditions: The most demanding life history stages have to match the times within a year when the environment is most favorable. As the timing of these favorable periods often varies from year to year, so too does the phenology of the life cycle: In some years, migrant birds return earlier to their breeding grounds than in other years, and ultimately this will be because favorable conditions occur at different times in different years.

Phenological phenomena are studied by different biological disciplines, which differ in their approaches and in questions investigated (Visser *et al.*, 2010). Phenology research as a discipline observes the periodic appearance of life cycle events and relates their annual variation to the variation in climate. Many of the reports on shifts in phenology simply regress the mean annual value of the phenology of a trait against mean temperature over a specific period and demonstrate that this mean date is, for example, earlier at higher spring temperatures. But phenology is also studied by evolutionary ecologists, who term it "seasonal timing" and emphasize the variation in timing among individuals and its fitness consequences. Phenology as it is studied by chronobiologists is viewed as the periodic nature of life cycle stages with an emphasis on their underlying endogenous timing programs (the so-called circannual rhythms). Finally, on a more mechanistic level, the (neuro-)endocrine processes underlying these life cycle events are studied by physiologists, and these processes are linked to genes by molecular geneticists (see Visser *et al.*, 2010 for a full discussion on an integrated framework of these disciplines). This article will apply the framework used by evolutionary ecologists as this will help explain, rather than just describe, the patterns in the observed shifts in phenology due to climate change.

Conceptual Framework for Shifts in Phenology

Animals have evolved seasonal patterns of activity that allow them to match their energetic needs, or that of their offspring, with the temporal pattern in environmental conditions. To optimize the annual scheduling of reproduction, migration, hibernation, etc., animals use photoperiod (i.e., day length, a reliable indicator of time of year) to anticipate seasonal changes in conditions. Photoperiod does not vary from year to year, however, so it cannot account for year-to-year variation in phenology (Bradshaw and Holzapfel, 2006). In the temperate zones, temporal changes in the environment are closely related to changes in temperature. In a warm spring, for instance, trees develop their leaves early and thus insects that feed from these newly developed leaves also need to appear early. For this reason, many temperate zone species use temperature as a cue: Their phenology is causally affected by temperature or by an environmental variable closely associated with temperature. Note, however, that in some species, for example, those inhabiting desert environments, their relevant environmental variation might be affected by variables other than temperature. For instance, variation in food abundance for zebra finches is more strongly affected by rainfall than by temperature, and thus environmental variables associated with rainfall are used instead (Zann *et al.*, 1995).

The relationship between phenology and temperature is mainly caused by individuals that change their phenology via

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00376-2

a mechanism termed "phenotypic plasticity" (Pigliucci, 2001): the same genotype (individual) expresses a different phenotype (phenological trait) under different environmental conditions (spring temperatures). The relationship between the phenological trait and the environmental variable affecting its value is termed a "reaction norm." Thus, the often observed negative relationship between a phenological trait and the temperature for a population is the result of individual reaction norms; for example, individuals respond by breeding earlier in warmer springs. Note, however, that there may be significant individual variation in the steepness (Nussey *et al.*, 2005) or shape (Reed *et al.*, 2009) of these reaction norms within a population although this is not always the case (Charmantier *et al.*, 2008; Reed *et al.*, 2006), with the response at the population level being the combined effect of all individual responses.

Climate change has led to a clear increase in mean temperatures worldwide; however, with striking variation both spatially (in some parts of the world, temperatures have increased more than in other parts) and temporally (within a single location, temperatures have increased more acutely in some parts of the year than in other parts). This increase in ambient temperature will, for the majority of species, shift their phenology to an earlier date in the year, because most species use temperature or a temperature-related environmental variable as a cue. Whereas among vertebrates, temperature may affect phenology mainly because it is used as a cue, for insects, temperature may (also) play a more direct role as the rate of many developmental processes in ectothermic species is temperature dependent. Global warming will thus directly affect the phenology of insects' life cycles, such as egg hatching date, when temperatures in the period after breaking diapause increase.

Methodological Issues in Studying Shifts in Phenology

Many studies have estimated shifts in phenology over recent decades. Before addressing these, however, it is useful to discuss a few methodological issues related to these estimates. One of the main methodological problems is that many datasets used to estimate phenological shifts are based on the first observations of an event: The first birds of a certain species returning from their wintering areas, the first butterfly seen in spring, etc. In a methodological paper, Moussus *et al.* (2010) used simulated data to compare 10 phenological estimators and found that first event dates are an inaccurate and biased estimator for the phenology of entire populations. This result was confirmed empirically in a few studies for which the entire distribution of phenological events was available. These show that first events shift more rapidly than middle or peak events (Miller-Rushing *et al.*, 2008; van Buskirk *et al.*, 2009; Thackeray *et al.*, 2010).

A second potential methodological problem is the strong increase in amateur observers, who have collected most of the long-term data on plants, birds, and butterflies, over the past decades. The date of first events, in particular, is strongly affected by the number of observers, and as a result, shifts in phenology might reflect the increased probability of observing a phenological event rather than a real change in the phenology of a species. This problem is also best solved by directing efforts to measure the entire distribution of phenological events rather than observing first events only (Moussus *et al.*, 2010).

A third methodological problem is also related to observing first events. With an increasing population size, it is likely that the date of first event shifts forward. This was shown by Tryjanowski and Sparks (2001), who analyzed migration phenology of red-backed shrike (*Lanius collurio*) in Poland from 1983 to 2000. The red-backed shrike both returned earlier and increased in population size, and a further analysis showed that there was a negative correlation between arrival phenology (Julian date of arrival) and population size. Another example comes from a study of 32 migratory North American bird species (1970–2002), which showed that changes in first arrival date differed strongly from changes in mean arrival date (0.2 days per decade later vs. 0.8 days per decade earlier). The difference in these two trends was due to declining cohort sizes over time (Miller-Rushing *et al.*, 2008). Thus, reports on shifts in phenology in first events need to take population increases or decreases into account, although this is seldom done.

A final pitfall is publication bias. Meta-analyses including single-species studies, in particular, may be a potential problem as positive results are more likely to be published than no change (Parmesan, 2007). Such a positive publication bias will lead to estimates of shifts based on datasets that underrepresent taxa showing little or no phenological change (Thackeray *et al.*, 2010). One way to prevent a publication bias is to use in meta-analyses only multispecies studies that report data on both shifting and nonshifting species (Parmesan and Yohe, 2003).

Observed Phenological Shifts Under Climate Change

There is an overwhelming body of literature on shifts in phenology in response to climate change. This article first presents an overview of the general review papers and then discusses some trends that emerge from these reviews. Next is a discussion on the variation in phenological shift among species within taxonomic groups and variation among populations within species.

General Reviews

The first two general reviews on the shift in phenology in response to climate change were published in 2003: the paper by Parmesan and Yohe (2003) and by Root *et al.* (2003) both reviewed the peer-reviewed literature, but their selection criteria for inclusion of studies differed. Although in both cases the overall conclusion was that there were clear shifts in phenology, the estimates of the magnitude of this shift differed: Parmesan and Yohe (2003) estimated the shift to be 2.3 days per decade across all species, whereas Root *et al.* (2003) calculated a shift of 5.1 days per decade for the subset of species showing a shift of at least 1 day per decade. The two papers were explicitly compared by Parmesan (2007), who attributed the difference in these shifts to the criteria used:

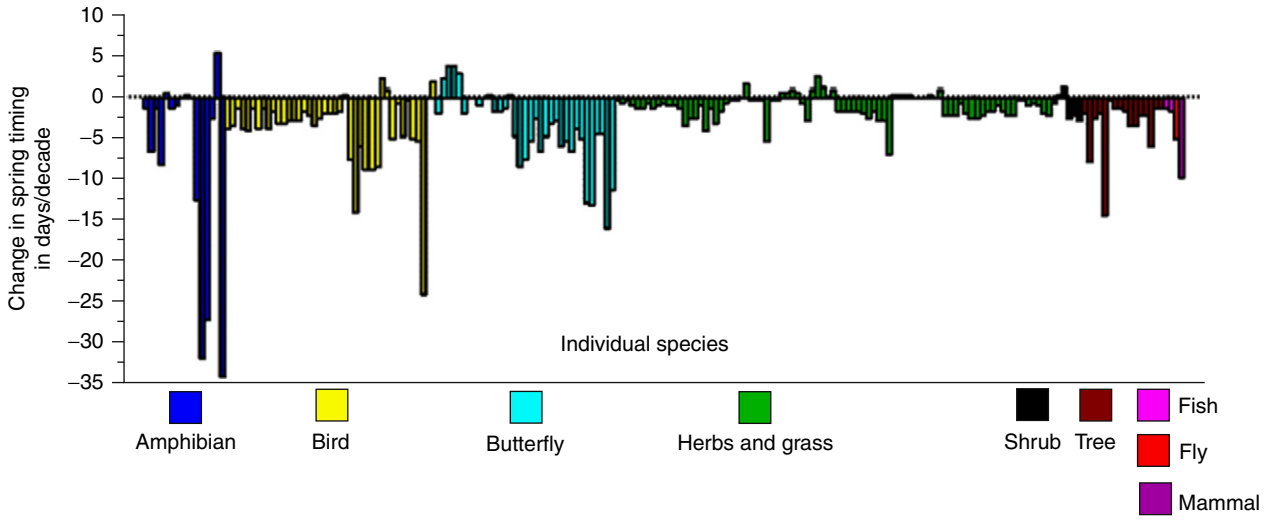


Figure 1 Shifts in spring phenology (days per decade) for individual species grouped by taxonomy or functional type. Each bar represents a separate, independent species. Negative values indicate a shift toward earlier phenology over time, positive values a shift toward later phenology (reproduced from Parmesan C (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology* 13: 1860–1872, with permission from Wiley).

Root *et al.* (2003) used both single species as well as multispecies studies and included only studies in which the shift was larger than 1 day per decade, whereas Parmesan and Yohe (2003) used only multispecies studies, independent of whether species shifted phenology or not. To obtain a better estimate for the shift in phenology and the variation in this shift among taxonomic groups, Parmesan (2007) used a combined data set of 203 Northern Hemisphere species in which only a single data point for each species was included. Using this combined database, the best estimate for the shift in spring phenology was 2.8 days per decade. However, there is substantial taxonomic variation in this rate: The shifts in phenology in amphibians were more than twice that in trees, birds, and butterflies, and almost eight times as fast as in herbs, grasses, and shrubs (**Figure 1**).

Thackeray *et al.* (2010) analyzed the phenological shifts in 726 terrestrial, freshwater, and marine species in the UK. They compiled more than 25,000 shifts calculated from the raw observations from 18 UK monitoring programs for a period of 30 years. Their estimate across taxonomic groups was a shift of 3.9 days per decade and with 84% of the observed shifts toward an earlier seasonal timing. This shift is somewhat higher than the shift of 2.8 days per decade estimated by Parmesan (2007). This may be either due to the time period used (1976–2005) by Thackeray *et al.* (2010), a period in which climate change was more pronounced, or due to the difference in the geographical region: In the UK data set used by Thackeray *et al.* (2010), shifts in phenology may be stronger than in the rest of the Northern Hemisphere data set used by Parmesan (2007). Like Parmesan (2007), Thackeray *et al.* (2010) also found large differences in the phenological shift among species from different taxonomic groups (**Figure 2**). Terrestrial plants shifted the most, whereas freshwater plant species shifted the least. Terrestrial invertebrates shifted more (4.1 days per decade) than terrestrial vertebrates (2.6 days per decade).

A few more general reviews have been published, albeit more limited in taxonomic diversity or geographical range (Walther *et al.*, 2002; Menzel *et al.*, 2006; Penuelas *et al.*, 2002). Menzel *et al.* (2006) presented phenological shifts from 21 European countries for the period 1971–2000, using national schemes with standardized datasets, but although data on 19 animal species were available, they report only on the data for the 542 plant species. Penuelas *et al.* (2002) studied phenological shifts in plants and arrival dates of birds in the Mediterranean region with the surprising result that the plants advanced their bud burst, flowering, and fruiting, but the sub-Saharan migratory birds now arrive 15 days later than in 1950.

From these reviews, a number of general conclusions can be drawn. First, there is a very clear shift in phenology toward an earlier date in response to climate change of between 3 and 4 days per decade, and second, that species from different taxonomic groups shift at different rates. However, these conclusions are to a very large extent based on data from the Northern Hemisphere. What little data there are from the Southern Hemisphere seem to point in a different direction. For instance, Barbraud and Weimerskirch (2006) show that for nine seabird species breeding in the Antarctic, both the breeding and the migratory phenology have shifted to later dates over the period 1950–2004. The overall shifts indicated that birds arrive 1.7 days per decade later and lay eggs 0.4 days per decade later. Another limitation of the data is that the majority of the documented shifts concern spring phenology. There are a few documented autumn shifts that show a shift in phenology to a later date such as leaf fall dates (Penuelas *et al.*, 2002) and the date of autumn migration in birds (Jenni and Kéry, 2003; Lehikoinen *et al.*, 2004).

Between-Species Variation Within Taxonomic Groups

There are a number of review papers on specific taxonomic groups, most of them on laying date in birds (Dunn, 2004; Crick *et al.*, 1997; Crick and Sparks, 1999). The first of these

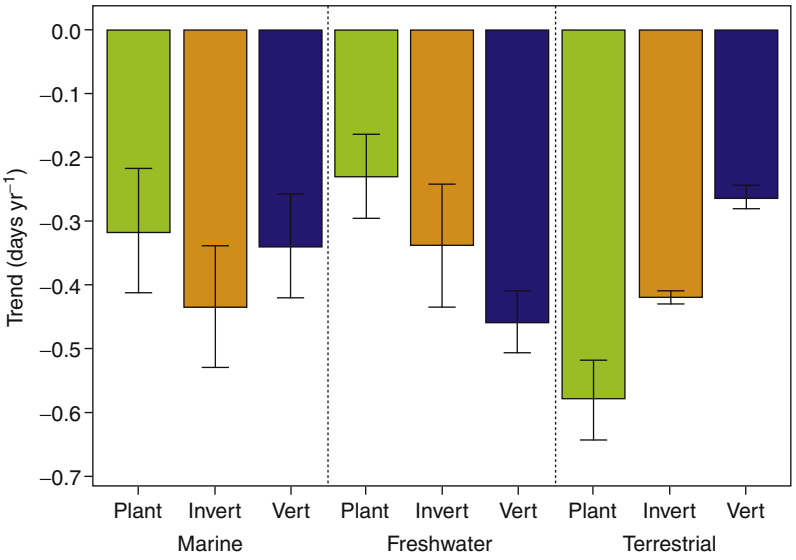


Figure 2 Phenological shifts from 1976 to 2005 for UK plants/phytoplankton (green), invertebrates (yellow), and vertebrates (blue) in marine, freshwater, and terrestrial environments. Negative values indicate a shift toward earlier phenology over time (reproduced from Thackeray SJ, Sparks TH, Frederiksen M, *et al.* (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology* 16: 3304–3313, with permission from Wiley).

was by Crick *et al.* (1997), who analyzed laying date from more than 74,000 nest records from 65 bird species from the UK between 1971 and 1995. In 51 of these species, there was a trend toward earlier nesting, with an average advancement of 3.5 days per decade for the 20 species for which the shift in phenology was significant, whereas just a single species, the stock dove (*Columba oenas*), laid significantly later. In a follow-up paper on an extended data set for the period 1939–1995, Crick and Sparks (1999) found that 19 out of 36 species showed an advance of laying date over the period 1970–1995. Dunn (2004) brought together data on shifts in the phenology of laying dates or hatch dates for an additional eight species from various locations, and in all of these, there is evidence for a shift to an earlier date.

These overviews of avian reproductive phenology thus clearly show that phenology has shifted, but perhaps more strikingly, they demonstrate a large variation in the rate of the shifts between species within a taxonomic group, even within a single geographic location (as in the Crick *et al.*, 1997, study). There are two potential explanations for this variation: Either the magnitude at which temperature is increasing varies among species' distribution ranges or species vary in their sensitivity to temperature. As there is certainly variation in the phenological shift among species at the same location, the latter explanation is more likely. The conceptual framework for the shifts in phenology (*see* Conceptual Framework for Shifts in Phenology) also predicts variation in temperature sensitivity among species. The sensitivity of a species to temperature, and thereby the rate at which its reproductive phenology will shift, will ultimately depend on the extent to which the phenology of its food depends on temperature (Visser and Both, 2005). The food phenology of some species is highly temperature sensitive, such as insectivorous forest passerines, whereas that of other species may not depend on temperature at all. This, for instance, explains why the phenology of the stock dove (*C. oenas*) did not advance (Crick *et al.*, 1997) as this is a species that feeds its nestlings with milk from the crop-sac (Dawson *et al.*, 2001) and hence can feed on a wide range of food for which the phenology will not have changed. Other opportunistic breeders, like crossbills, are also not likely to advance their phenology if climate change leads to increased temperatures.

A number of overviews on the shifts in avian migration phenology make use of the banding stations along migratory flyways (Jenni and Kéry, 2003; Hüppop and Hüppop, 2003; Jonzen *et al.*, 2006; Sokolov, 2006; Lehikoinen *et al.*, 2004). Hüppop and Hüppop (2003) showed for 24 migratory species that they now pass through Helgoland (Germany) in spring 0.5–2.8 days per decade earlier than in 1960, with no difference between short/medium-distance and long-distance migrant species. In a study in Pennsylvania (USA), van Buskirk *et al.* (2009) found that spring migration advanced in a set of 58 songbird species over a 46-year period with on average 0.7 days per decade. In a meta-analysis of 21 studies recording first arrival dates of 983 species/study combinations, Lehikoinen *et al.* (2004) found an advancement of 3.7 days per decade, whereas a meta-analysis of five studies (with 222 species/study combinations) on median/mean migration time showed an advancement of 1.0 day per decade. Shifts in autumn migration have been documented less frequently and are often more variable (van Buskirk *et al.*, 2009). Jenni and Kéry (2003) show for 65 bird species that, over a 42-year period, birds passing through Switzerland altered their passage time. Although long-distance migrants pass through earlier in recent years, medium-distance migrants pass through later, especially the species that have multiple broods per year. For these species the breeding season has become prolonged as climate change has led to a longer period of favorable conditions. This study again shows that there may be large differences among species in their shifts in phenology, related to their ecology.

Another source of data on shifts in avian migratory phenology comes from studies on arrival dates (Cotton, 2003; Penuelas *et al.*, 2002; Gordo and Sanz, 2005). In the UK, a study on 20 migrant species showed an average advancement of 2.7 days per decade (Cotton, 2003). Arrival dates of 23 species of passerines at the Courland Spit (Russia) advanced in 22 species (15 significantly), an average of 3.8 days per decade, with no difference between median- and long-distance migrants (Sokolov, 2006). However, not all studies show an advancement of arrival date. In Spain, five out of six long-distance migrant species now arrive later rather than earlier over a 50-year period (Penuelas *et al.*, 2002). This again emphasizes the substantial variation across geographical scales, which for migratory birds is even more pronounced as it is the variation in climatic change not only in the breeding location but also in the overwintering location that contributes to the variation in the shift in phenology.

Butterflies are the only other taxonomic group of animals for which an overview of phenological shifts for a reasonable number of species is available (Roy and Sparks, 2000; Stefanescu *et al.*, 2003; Forister and Shapiro, 2003). In all these studies, the phenology of first appearance has advanced. In the UK, this was true for 26 out of 35 species, of which 13 were statistically significant (mean advancement over all species 3.7 days per decade; Roy and Sparks, 2000); in Spain, for all of the 17 species included in the study, of which five were statistically significant (Stefanescu *et al.*, 2003); and in California, 16 out of 23 species advanced their phenology, of which four were statistically significant (mean advancement over all species 1.7 days per decade; Forister and Shapiro, 2003). Despite the clear overall advancement in phenology over the past decades, butterflies also show clear among-species variation in the rate (and to a lesser extent the direction) of these shifts. Forister and Shapiro (2003) tested whether species that appear early in the season shifted more, which was not the case. However, species that overwinter as pupae did advance their phenology more than those overwintering as larvae.

Between-Population Variation Within Species

Comparisons of the shifts in breeding phenology among populations within species are quite rare as the comparisons depend on long-term studies of a single species at several locations. Data have been analyzed for four bird species, all hole-breeding passerines: great tit (*Parus major*), blue tit (*Cyanistes caeruleus*), pied flycatcher (*Ficedula hypoleuca*), and tree swallow (*Tachycineta bicolor*) (Sanz, 2002; Both *et al.*, 2004; Visser *et al.*, 2003; Dunn and Winkler, 1999; Sanz, 2003). Tree swallows advanced their laying date 2.7 days per decade over the period 1959–1991, based on an analysis over their entire North American breeding range (Dunn and Winkler, 1999). In a Europe-wide analysis of pied flycatchers, Both *et al.* (2004) focused on geographical variation in the shift in breeding phenology. They found that 9 out of the 23 study population significantly advanced their phenology in the period 1980–2002, with an average shift over all populations of 1.6 days per decade. More importantly, they showed that the rate of advancement was correlated with the rate at which temperature had increased at that location, again emphasizing that the variation in the change in phenology is related to geographical variation in the change in climate. Visser *et al.* (2003) analyzed laying dates of 14 European populations of great and blue tits for 1979–1998 and also found that despite a general advancement in phenology (over all populations, advance for great tits was 3.1 days per decade and for blue tits 3.5 days per decade), there was ample variation among populations in the rate of change in breeding phenology. This variation was correlated with the degree to which populations were formerly double brooded: Populations that historically were single brooded advanced their laying dates more than those that used to be double brooded and have now become single brooded (Husby *et al.*, 2009). Thus, not only these within-species studies echo the results of among-species studies, showing that variation in the shift in phenology is correlated with geographical variation in change in climate, but the within-species variation in general life-histories also plays an important role.

Linking Observed Shifts in Phenology to Climate Change

This section discusses how shifts in phenology of three major phenological life history events – timing of reproduction, timing of migration, and timing of winter rest – are linked to climate change. There are two general approaches to linking phenology to climate change. The more common one is to correlate shifts in phenology to some measure of climate, often with local temperatures. In most studies, phenology is correlated with temperature by averaging temperature over some period. Although there are ample problems with this, (*see* Outlook), the author includes these studies as they form the bulk of the literature. Some studies correlate phenology with large-scale atmospheric phenomena, such as the North Atlantic Oscillation (NAO; Hurrell *et al.*, 2001) or the El Niño-Southern Oscillation (ENSO; Trenberth and Caron, 2000). Winter NAO, defined as the difference in the sea-level air pressure between Portugal and Iceland from December to March (Hurrell *et al.*, 2001), captures variation in climatic conditions on a large scale. In years with positive NAO values, weather in Europe is more affected by Atlantic influences, leading to wetter winters with higher temperatures, whereas in years with negative values the weather is more affected by continental influences, characterized by cold, dry winters in the north of Europe. The ENSO is characterized by variations in the surface temperature of the tropical eastern Pacific Ocean (of which the warming and cooling phases are known as El Niño and La Niña, respectively) and by air surface pressure in the tropical western Pacific (the Southern Oscillation). In years with positive anomalies, there are drier conditions in the Gulf of Guinea and Sahel rainfall is reduced. These large-scale atmospheric phenomena are used to correlate phenology with a set of climatic variables on a larger geographical scale, rather than with specific local temperatures. The atmospheric approach may be more appropriate when analyzing phenological data on larger scales or when trying to capture the effect of more climatic variables than temperature (Hallett *et al.*, 2004).

A second approach is to study, under controlled conditions, the environmental variables involved in the phenotypic plasticity of individuals, which underlies the changes in phenology. Although this approach allows for causal rather than correlational work, the number of species studied in sufficient detail to make the link between plasticity and phenological shifts is very limited, also because it turns out that causal links of phenology to temperatures are difficult to demonstrate in vertebrates (Visser *et al.*, 2009). This section therefore concentrates on the approach in which phenology is correlated with environmental variables, either as temperature or as NAO and ENSO, but returns to the latter approach in the section Outlook.

Timing of Reproduction

Shifts in the reproductive phenology (timing of reproduction) have been linked to temperatures for many taxonomic groups. For birds, Crick and Sparks (1999) showed that for 31 out of 36 species, the breeding phenology was correlated with temperature and precipitation. In the overview of Dunn (2004), in 50 out of 63 studies, temperature was correlated with laying date. An even more direct indication that shifts in laying dates are due to climate change comes from the correlation between the change in laying date and the change in local temperature in a Europe-wide comparative study of pied flycatchers (Both *et al.*, 2004). In four species of amphibians, there was also a clear correlation between spawning date and temperature (Beebee, 1995), which could account for the advancement of their phenology. In a mammalian species, the caribou (*Rangifer tarandus*), the timing of onset of calving was correlated with temperatures in the period from March to May (Post and Forchhammer, 2008). There is thus ample evidence that reproductive phenology is correlated with temperatures, albeit that the period over which these temperatures are averaged varies from species to species (*see* Outlook).

Using NAO rather than local temperatures, Sanz (2002) found that the breeding phenology of great and blue tits was correlated with winter NAO (with earlier laying at positive NAO values). A similar result was found by Forchhammer *et al.* (1998) for the corn bunting (*Miliaria calandra*) and the chiffchaff (*Phylloscopus collybita*), as well as for the spawning date of three species of amphibians.

Timing of Migration

One of the key problems with linking migration phenology to climatic conditions is that potential climatic conditions at the wintering grounds, the breeding grounds, and along the migration route all play a role (Gordo, 2007). There are ample studies showing a correlation between arrival or passage dates and local temperatures. In a meta-analysis on migratory birds, Lehikoinen *et al.* (2004) showed that both first arrival dates and median/mean migration phenology were negatively correlated with local temperatures. As these local temperatures cannot causally affect the birds at departure or *en route*, it is likely that, due to large-scale spatial autocorrelations in climate, the temperatures experienced by the birds at departure or *en route* were correlated with these local temperatures.

Given that migratory species are affected in their phenology by conditions on a larger spatial scale, many studies have correlated migratory phenology with large-scale atmospheric phenomena, mainly with the NAO. In their meta-analysis, Lehikoinen *et al.* (2004) found a clear negative correlation between phenology and NAO, both for first date of arrival and for median/mean migration phenology. One of the studies in their meta-analysis, in which 40 years of observations on Helgoland (Germany) were analyzed, found that mean spring passage time for no less than 23 out of 24 birds species had a negative relationship with NAO (Hüppop and Hüppop, 2003). In another meta-analysis on avian migration phenology Gienapp *et al.* (2007) analyzed data from 18 studies on 249 species and also found a negative relationship between migration phenology and NAO, which did not differ for European long- and short-distance migrants.

Considerably fewer studies have correlated migration phenology with another large-scale atmospheric phenomenon, the ENSO. For instance, Barbraud and Weimerskirch (2006) related the breeding and the migration phenology of six bird species breeding in the Antarctic to ENSO, and in only one species, the Cape petrel (*Daption capense*), there was a significant correlation. MacMynowski and Root (2007) found a correlation between migratory phenology and ENSO for only 2 out of 22 North American species. Another study on 23 North American migrants showed that variation in arrival date of short-distance migrants correlated with variation in temperature, whereas for medium-distance migrants, a correlation was found with the Southern Oscillation Index (Miller-Rushing *et al.*, 2008).

Clearly, European and North American avian migratory phenology is correlated with both local temperatures and NAO. There is not much evidence for correlations with ENSO, but given that the overwhelming majority of studies are on Northern Hemisphere populations, clearly more data are needed. There is some evidence that the phenology of short-distance migrants correlates best with local temperature, whereas that of medium- or long-distance migrants is correlated better with the NAO (MacMynowski and Root, 2007; Miller-Rushing *et al.*, 2008). This would make sense if short-distance migrants overwinter so close enough to their breeding grounds that there is spatial autocorrelation in climatic conditions between their wintering and breeding areas. In long-distance migrants, autocorrelation is less likely, so that a large-scale atmospheric phenomenon, such as the NAO, captures the relevant climatic variables better.

For taxonomic groups other than birds, there are very few studies linking migration phenology to either temperatures or NAO, but there is some evidence that the migration phenology in veined squid (*Loligo forbesi*) is correlated with sea water temperature and with NAO (Sims *et al.*, 2001), and similar correlations were found for a fish, the flounder (*Platichthys flesus*) (Sims *et al.*, 2004). In pink salmon (*Oncorhynchus gorbuscha*) in Alaska, fry migration phenology advanced in response to increasing water temperatures 5.0 days per decade in the period 1972–2005 (Taylor, 2008).

Timing of Onset and Termination of Winter Rest

Many species have some sort of winter rest. Insects may go into diapause as an egg, a larva, a pupa, or an adult, whereas mammals may hibernate throughout the winter. Such a life

history requires two timing decisions: When to go into winter rest and when to become active again. In many species the induction and the breaking of diapause or the start and end of hibernation are triggered by photoperiod or by some internal clock. However, the developmental rate in, for instance, insect eggs is strongly affected by temperature, and hence the length of time until these eggs hatch will be directly related to temperature.

Adult winter moths (*Operophtera brumata*) enclose from their pupae in November/December, after which they mate and the wingless females lay their eggs on the branches of trees. These eggs then develop slowly until they hatch in April, at the time the buds of the trees open. The phenology of egg hatching is clearly affected by temperature, as shown both in the wild and under controlled conditions (Visser and Holleman, 2001), and thus egg hatching date has advanced due to climate change. Butterfly larvae come out of winter egg rest in spring, and their date of first appearance is correlated with spring temperatures. Many species thus have advanced their flight period as a consequence of climate change (Roy and Sparks, 2000; Stefanescu *et al.*, 2003; Forister and Shapiro, 2003).

Termination of hibernation of the yellow-bellied marmot (*Marmota flaviventris*) in the Colorado Rocky Mountains is strongly affected by April temperatures (Inouye *et al.*, 2000). As these temperatures increased, the marmots emerged 38 days earlier over the period 1975–1999 (15.2 days per decade).

Overall Conclusion

In all taxonomic groups there is clear evidence that the phenology of a majority of species is correlated with one or another temperature or, especially for migratory species, with large-scale climatic patterns. However, it is not always clear over which period predictive temperature needs to be averaged as this varies from species to species depending on their ecology. Additionally, correlating a phenological event with the mean temperature over a fixed period has clear methodological problems (*see* Outlook). Despite this general pattern there are also species whose phenology is not affected by temperature. These are likely species for which temperature does not predict the period of favorable conditions (*see* Conceptual Framework for Shifts in Phenology). Furthermore, for the few species for which this has been studied, temperature sensitivity of phenology varies on a geographical scale with populations at more northern locations being less temperature sensitive (Silverin *et al.*, 2008). Because climate change scenarios predict the strongest temperature changes in the north, this lack of sensitivity may lead to insufficient shifts in phenology in northern locations (*see* Phenological Mismatches Due to Climate Change).

Phenological Mismatches Due to Climate Change

Mismatches among Trophic Levels

Although many species shift their phenology in response to contemporary climate change, this does not mean that these shifts are necessarily adaptive, that is, that there are no negative effects on the reproductive success or survival of these shifts. To determine whether the observed shifts in phenology are sufficiently compensatory, or perhaps even too strong, some kind of yardstick is needed. Visser and Both (2005) proposed to use the shift in phenology of the food of animals as such a yardstick. For instance, to assess whether small insectivorous birds shift their breeding phenology sufficiently, their rate of advancement is compared to the rate at which the phenology of their food source, caterpillars, is advancing. Of the 10 pairs of species interaction for which data on phenological shifts were available, in five the shift of the predator was not strong enough, in two cases it was too strong, and only in three cases the shifts of prey and predator were of similar strength (Visser and Both, 2005). Thus, in the majority of the cases, climate change has led to a so-called phenological mismatch (Visser *et al.*, 1998; Stenseth and Mysterud, 2002; van Asch and Visser, 2007; Harrington *et al.*, 1999).

The generality of the phenological mismatch problem is confirmed by the review of phenological shifts reported by Thackeray *et al.* (2010). In their analysis, the phenological shifts varied significantly among the different trophic levels: secondary consumers advanced less strongly than primary producers and primary consumers (**Figure 3**). This difference was found in marine, freshwater, and terrestrial ecosystems, indicating that phenological mismatches are likely in all ecosystems.

It is somewhat surprising that not all species within a food chain shift their phenology at the same rate, as there has been extensive selection on predators to keep their phenology matched to that of their prey under the natural range of annual temperature variation. This variation is normally larger than the increase in temperature due to climate change.

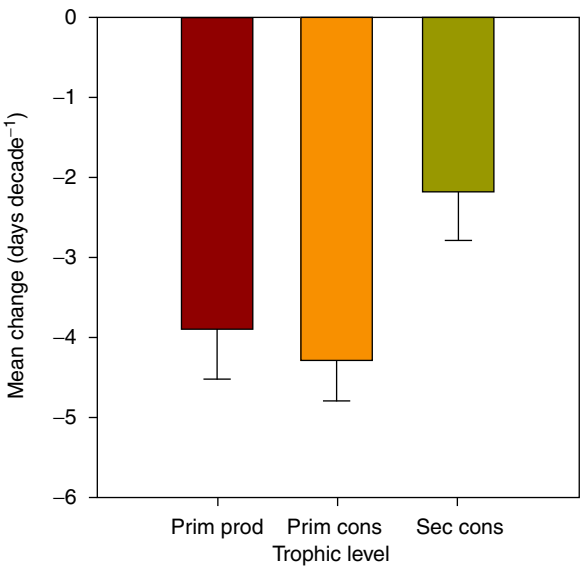


Figure 3 Phenological shifts from 1976 to 2005 for UK primary producers (red), primary consumers (orange), and secondary consumers (green). Negative values indicate a shift toward earlier phenology over time (reproduced from Thackeray SJ, Sparks TH, Frederiksen M, *et al.* (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology* 16: 3304–3313, with permission from Wiley).

Although there may not be a single answer for all species, two elements are of importance. First, "decisions" on the timing of life-history stages often have to be made weeks before any chance of a phenological match with the food. For instance, small passerine birds have to lay their eggs approximately a month before the time that they need ample food for provisioning their chicks, the time when they need to be matched to the phenology of their food. Similarly, migratory species often have to set their departure time from the wintering areas weeks before they have to match their reproductive phenology with favorable conditions. Animals thus have developed mechanisms that use predictive cues to time their reproduction or migration. The second element is that climate change has not led to a uniform increase in temperatures over the entire year. This lack of uniformity may thus lead to situations in which temperature increases less at the time of decision making, thus only marginally affecting the cues that are used, but in which temperature increases much more at the time of selection, thus affecting the phenology of the food (Visser *et al.*, 2004). Under these conditions, the cues are no longer accurately predicting the favorable period. One potential explanation why secondary consumers, in particular, are advancing their phenology too little is that these are often vertebrates that use photoperiod as one of the main cues in their timing (Bradshaw and Holzapfel, 2008), and photoperiod is obviously not changing due to climate change.

Studies on potential phenological mismatch have been carried out mostly on birds. Visser *et al.* (2011) evaluated 16 species interactions for which data on the shift in phenology of both the bird species and its prey species are known and found that in seven of these, climate change leads to a phenological mismatch. This review was, however, strongly biased to small forest insectivores (13 out of 16 studies), such as the great tit. Phenological mismatches do not always develop between the phenologies of breeding and food abundance. The American robin (*Turdus migratorius*), an altitudinal migrant, has advanced its arrival date at its breeding grounds 14 days over a 26-year period (1974–1999), although the date of snow melt has not changed, and as a consequence the delay between arrival and the date of bare ground is now 18 days (Inouye *et al.*, 2000). A second example comes from black-tailed godwits (*Limosa limosa*), which have not advanced their breeding phenology, but at the time the chicks hatch, the grass is taller and denser than 25 years ago, making it difficult for the nidifugous chicks to forage (Kleijn *et al.*, 2010).

There are just a few mammal studies on trophic mismatches. Yellow-bellied marmots (*M. flaviventris*) now come out of hibernation 3 weeks earlier, whereas the flowering time of their food plants has not changed, thus altering the relative phenology of the marmots and their food plants (Inouye *et al.*, 2000). Caribou (*R. tarandus*) in Greenland need to time the birth of their calves with the plant-growing season in order for their calves to have sufficient suitable food. Calving date advanced just 3.8 days between 1993 and 2006 (2.7 days per decade), whereas the phenology of onset of the plant-growing season advanced by 4.6 days (3.3 days per decade). Thus, caribou are getting increasingly mismatched (Post and Forchhammer, 2008).

Among amphibians, breeding phenology has shifted in some species, but not in all (Beebee, 1995), and this has affected the trophic interactions between prey and predators. In the UK, newts (*Triturus* spp.) now enter ponds earlier, whereas frogs (*Rana temporaria*) did not advance their breeding phenology. As a consequence, the larvae of the frogs are exposed to higher levels of newt predation.

For insects, the phenological match between the leaf-feeding moths or aphids and the bud opening of their host plants, or of flower- and seed-eating weevils and the flowering or seed setting of their host plant, has major fitness consequences (van Asch and Visser, 2007). As the phenology of insects and plants are affected by temperatures of different periods of spring, and because climate change leads to differential warming during these different periods, there is ample opportunity for mismatching to occur. This has been shown for the winter moth (*O. brumata*), for which the egg hatching date has advanced more in the past 25 years than the bud burst date of the oak (*Quercus robur*) (Visser and Holleman, 2001).

Mismatches among Life-History Stages

Phenological events need to be matched not only across species in a food chain but also with abiotic conditions, within trophic levels or even within-individual life histories (Miller-Rushing *et al.*, 2010). Examples of such within-individual mismatches are migrant species that first need to arrive at the breeding ground before they can start breeding. Similarly, the phenology of different life-history stages such as breeding and molt (fur in mammals or feathers in birds) need to be phenologically matched. Not many studies have addressed this explicitly, but as some life-history stages are affected by temperature whereas others are not, climate change could potentially disrupt the phenological matches within individuals' life cycles (Crozier *et al.*, 2008; Miller-Rushing *et al.*, 2010).

The best known examples of differential shifts in the phenologies of coupled life-history traits are arrival date and breeding date in migratory birds, where a lack of the shift in arrival date could potentially constrain shifts in breeding phenology (Both and Visser, 2001). Another example comes from a study on red deer (*Cervus elaphus*) (Moyes *et al.*, 2011), in which the phenologies of six traits were measured over a 28-year period. In female deer, parturition dates advanced by 11.8 days (4.2 days per decade), whereas the date of first oestrus advanced by only 7.3 days (2.6 days per decade). In male deer, the phenologies of antler casting and cleaning advanced by 5.6 (2.0 days per decade) and 7.3 days (2.6 days per decade), respectively. Interestingly, the end of the rut shifted twice as fast as its start dates: 12.0 versus 5.9 days (4.3 vs. 2.1 days per decade), respectively. It is likely that the phenologies of these traits are differentially affected by temperature or by different temperature periods, leading to a potential mismatch in the phenology of the deer's life-cycle stages.

Consequences of Shifts in Phenology

Consequences for Population Viability

Climate change leads to major shifts in phenology, and in many cases these shifts do not seem to be of the appropriate

magnitude, leading to a phenological mismatch (Visser and Both, 2005). As there is ample evidence that mismatched reproduction has consequences for reproductive success and survival (Thomas *et al.*, 2001), it is likely that climate change will affect population viabilities by inducing these nonadaptive shifts in phenology (Miller-Rushing *et al.*, 2010). There are, however, only a few studies that have addressed the consequences of phenological shifts on population processes.

In a study on 100 European migratory bird species, Moller *et al.* (2008) found that the species that declined the most in breeding numbers over the period 1990–2000 were those that advanced their migration phenology the least (**Figure 4**). Although it is not clear how the lack of a shift in phenology affects population numbers, this multispecies analysis does indicate that there is a link between shifts in phenology and population viability.

A number of single-species studies confirm a link between shifts in phenology and population numbers. In a study on a number of pied flycatcher (*F. hypoleuca*) populations in the Netherlands, Both *et al.* (2006) found a significant relationship between the phenology of the food peak and the population trend in the flycatchers: populations declined the most in the areas with the largest phenological mismatch. In another analysis, population trends of six long-distance migrant species between 1984 and 2004 were compared in forests, where there is a sharp invertebrate food peak, and marshes, where the food peak is much less pronounced. As predicted, the populations in forests declined (by 38%), whereas those of the same species in marshes did not systematically decline (Both *et al.*, 2010).

The reproductive phenology of caribou (*R. tarandus*) in Greenland is under control of photoperiod, whereas the phenology of plant growth is affected by spring temperatures. As temperatures have increased, there is now a phenological mismatch between the timing of calving by caribou and the timing of plant growth (Post and Forchhammer, 2008). This has had consequences for reproduction: offspring mortality increased and offspring production has dropped fourfold. Another study on mammals, however, did not show clear population effects of phenological mismatches. In Soay sheep (*Ovis aries*), there was no effect of the phenological mismatch on offspring survival, although this was not unexpected as their environment is only weakly seasonal with vegetation being available all year round (Durant *et al.*, 2005). Another study, on red deer (*C. elaphus*), which does show shifts in phenology of various life-history traits but does not report whether or not animals get mismatched, found no changes over time in offspring birth weight or offspring winter survival (Moyes *et al.*, 2011).

It is clear that more studies are needed that link shifts in phenology to population viability. Given the frequent occurrence of phenological mismatches, it is likely that there will be consequences for population viability, with potential knock-on effects for biodiversity in general, but at present it is not possible to quantify these effects.

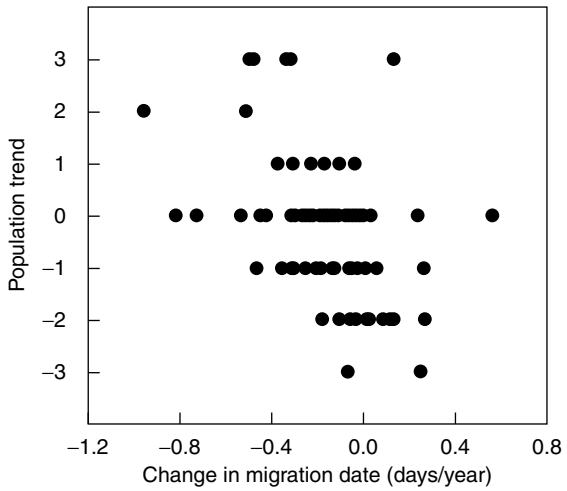


Figure 4 Population trends of European migratory bird species (1990–2000) in relation to shifts in spring migration phenology (days per year; 1960–2006). Negative values in the population trend indicate a decline in numbers, positive values an increase. Negative values in the shift in phenology indicate a shift toward earlier phenology over time (reproduced with permission from Moller AP, Rubolini D, and Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are declining. *Proceedings of the National Academy of Sciences of the United States of America* 105: 16195–16200).

Consequences for Microevolution

Shifts in phenology due to climate change can be caused by phenotypic plasticity, in which individuals change their timing in response to temperature (*see* Conceptual Framework for Shifts in Phenology), or it can be due to genetic changes in the population in response to selection (microevolution). By far, phenotypic plasticity is the most common mechanism underlying phenological shifts. Gienapp *et al.* (2008) reviewed the evidence for genetic responses in long-term studies of vertebrates and found that there was very little evidence for genetic responses. One of the best known examples of genetic adaptation comes from an insect, the pitcher plant mosquito (*Wyeomyia smithii*) in which the phenology of entering diapause has changed genetically in response to a longer growing season. The mosquitoes now enter winter diapause at a shorter day length than 24 years ago (Bradshaw and Holzapfel, 2001).

Although the majority of phenological shifts are due to phenotypic plasticity, this does not mean that this plasticity is sufficient to keep species matched to their abiotic and biotic environments. Often the phenology of species responds either too strongly or too weakly to climate change, resulting in a phenological mismatch, indicating that plasticity is no longer adaptive. An example comes from great tits (*P. major*) in the Netherlands, where the timing of reproduction is phenotypically plastic but still the shift in phenology was insufficient to keep the breeding matched with the phenology of their food. As a result, there is now selection on the plasticity: The birds should respond more strongly to temperature, as is clear from the directional selection for the degree of plasticity (Nussey *et al.*, 2005). Interestingly, in a great tit population in the UK, plasticity is still adaptive and the breeding phenology changes at a similar rate as the phenology of the food (Charmantier *et al.*, 2008).

Although in the great tits a response to this selection on the degree of plasticity is estimated to be slow, in another example, on winter moth egg hatching date, plasticity is estimated to genetically shift rather rapidly (van Asch *et al.*, 2007). In this system, egg hatching date is too early, as the winter moths respond too strongly to temperature in their phenology compared to the phenology of bud burst of the oaks that the caterpillars feed on. However, based on the heritability of and the selection on egg hatching date, a quantitative genetic model predicts rapid adaptation in the degree of plasticity, which will restore the phenological match.

The key issue is thus at what rate species can change genetically. It is this rate of microevolution that will determine whether or not species will be able to adapt fast enough to prevent a decline in population viability and thereby a loss of biodiversity (Visser, 2008) as shifts in phenology due to phenotypic plasticity alone will not be sufficient because this plasticity will no longer be adaptive (*see* Consequences of Microevolution). Adaptation via immigration of individuals with genotypes adapted to warmer environments will also require microevolution as these genotypes will be better adapted to current temperature, but not, for instance, to the photoperiod of their new environment.

Outlook

There is ample evidence for phenological shifts due to contemporary climate change (*see* Observed Phenological Shifts Under Climate Change), but the data available are strongly skewed to plants and birds. Mammal, fish, amphibian, reptile, and invertebrate studies are quite rare, with the exception of studies on butterflies. One clear reason for this pattern is that birds and butterflies receive ample interest from the general public and can be easily observed, and as a consequence, long-term datasets based on monitoring programs in which the data are collected by amateurs are available for analysis. Also, data from aquatic and marine systems are less common than data from terrestrial systems. However, from the data that are available, it seems that the patterns found in these well-known groups are not fundamentally different from the less-studied groups or habitats (**Figure 2**). A possible exception is the degree to which climate change induces genetic changes, which seems to be higher in vertebrates.

There is also ample evidence that the observed shifts in phenology over the past decades are indeed due to climate change (*see* Linking Observed Shifts in Phenology to Climate Change). There are, however, important methodological problems with linking phenology to temperatures (Gienapp *et al.*, 2005). Often phenological events are simply correlated with the mean temperature or the sum of temperatures over a fixed period. This period is often a single month or a combination of months. One problem is that the correlation found depends on the choice of this period. The period yielding the best correlation can be established using a sliding-window approach to test all possible periods between two dates. But even within species, this best correlating period may differ among populations (Husby *et al.*, 2010). A second problem is that often the earliest individuals in a population already expressed their phenological trait before the end of the period over which temperature is being averaged, and thus their phenology correlates with temperatures occurring after they expressed their trait. A final problem is that this fixed-period approach does not allow tests for an interaction with, for instance, photoperiod, which is an important cue in the phenology of many species.

These problems with a sliding-window approach can be avoided by analyzing phenology using proportional hazard models (Gienapp *et al.*, 2005). These models calculate the probability that the phenological event will occur every day under the assumption that it has not yet occurred. The model thus calculates how, for instance, temperature affects this probability, not only using the temperature of that day, but also integrating the temperatures of previous days. The weight of these previous days decreases with each passing day Gienapp *et al.* (2005) assumed that this decrease follows a specific rule, whereas this assumption was relaxed in a model that combines a memory function of past temperatures with a sliding-window approach (van de Pol and Cockburn, 2011). Although these two methods have clear advantages over the classical sliding-window approach, they are not widely used, probably because of the ease with which the sliding-window model can be calculated and interpreted.

An approach fundamentally different from these statistical correlations that can be used to assess the effects of temperature on phenology comes from (neuro)physiological studies. Such studies will ultimately determine which environmental variables affect the phenology of a trait via the mechanisms underlying plasticity (Visser *et al.*, 2010). Obviously, this approach will not be possible for a large set of species, but by choosing the few species where research is feasible based on their differences in life-history traits (long- vs. short-lived, etc.), it may be possible to obtain general insights into whether temperature plays a direct role in the timing of life cycle events or, instead, whether this effect comes about indirectly, for instance, via the phenology of the vegetation (Schaper *et al.*, 2011).

The observed shifts in phenology are mainly due to phenotypic plasticity, as there is very little evidence for genetic changes (*see* Consequences of Shifts in Phenology). As discussed earlier, this plasticity is often no longer optimal, leading to phenological mismatches. If temperatures keep on increasing due to climate change, such suboptimal or maladaptive shifts in phenology will become more common and will have more severe consequences, both for individuals (Visser *et al.*, 2004) and for populations (Reed *et al.*, 2010). This is simply because organisms have evolved under a certain set of correlated abiotic variables and it would be very unlikely that any response to these abiotic variables would still be optimal when climate change leads to a set of abiotic variables that are correlated differently (Visser *et al.*, 2004). Moreover, there may be nonlinear effects or constraints in plasticity that will prevent any further advancement of phenology at higher temperatures. At what temperatures this may become important can be predicted only by in-depth causal studies on the relationship between phenology and temperature (*see* Outlook).

There is already some evidence that shifts in phenology have consequences for population viability and from there for biodiversity (*see* Consequences of Shifts in Phenology).

Clearly, more studies are needed to assess how large this impact may be. But the next challenge will be to forecast the impact of future climate change on biodiversity. This is particularly important because future climate change is not determined but will depend on the greenhouse gas emission targets that are set to reduce global climate change. It is of crucial importance that ecological consequences of socio-economic development are included while setting these climate targets (Visser, 2008). Phenology will play an important role in the models that assess climate change impact because of the large fitness consequences of phenology. Ultimately, models that combine evolution and ecological processes (e.g., Reed *et al.*, 2011) have to be developed (Jenouvrier and Visser, 2011), and from these the consequences of climate change, via shifts in phenology, of the different climate scenarios (IPCC, 2007) on biodiversity have to be estimated.

See also: Climate Change and Wild Species. Evolution in Response to Climate Change. Plant Phenology Changes and Climate Change

References

Barbraud C and Weimerskirch H (2006) Antarctic birds breed later in response to climate change. *Proceedings of the National Academy of Sciences of the United States of America* 103: 6248–6251.

Beebee TJC (1995) Amphibian breeding and climate. *Nature* 374: 219–220.

Both C, Artemyev AV, Blaauw B, *et al.* (2004) Large-scale geographical variation confirms that climate change causes birds to lay earlier. *Proceedings of the Royal Society of London Series B – Biological Sciences* 271: 1657–1662.

Both C, Bouwhuis S, Lessells CM, and Visser ME (2006) Climate change and population declines in a long-distance migratory bird. *Nature* 441: 81–83.

Both C, van Turnhout CAM, Bijlsma RG, Siepel H, Van Strien AJ, and Foppen RPB (2010) Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. *Proceedings of the Royal Society B – Biological Sciences* 277: 1259–1266.

Both C and Visser ME (2001) Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature* 411: 296–298.

Bradshaw WE and Holzapfel CM (2001) Genetic shift in photoperiodic response correlated with global warming. *Proceedings of the National Academy of Sciences of the United States of America* 98: 14509–14511.

Bradshaw WE and Holzapfel CM (2006) Climate change – Evolutionary response to rapid climate change. *Science* 312: 1477–1478.

Bradshaw WE and Holzapfel CM (2008) Evolution of animal photoperiodism. *Annual Reviews of Ecology, Evolution and Systematics* 38: 1–35.

Charmantier A, Mccleery RH, Cole LR, Perrins C, Kruuk LEB, and Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* 320: 800–803.

Cotton PA (2003) Avian migration phenology and global climate change. *Proceedings of the National Academy of Sciences of the United States of America* 100: 12219–12222.

Crick HQP, Dudley C, Glue DE, and Thomson DL (1997) UK birds are laying eggs earlier. *Nature* 388: 526.

Crick HQP and Sparks TH (1999) Climate change related to egg-laying trends. *Nature* 399: 423–424.

Crozier LG, Hendry AP, Lawson PW, *et al.* (2008) Potential responses to climate change in organisms with complex life histories: Evolution and plasticity in Pacific salmon. *Evolutionary Applications* 1: 252–270.

Dawson A, King VM, Bentley GE, and Ball GF (2001) Photoperiodic control of seasonality in birds. *Journal of Biological Rhythms* 16: 366–381.

Dunn P (2004) Breeding dates and reproductive performance. *Advances in Ecological Research* 35: 69–87.

Dunn PO and Winkler DW (1999) Climate change has affected the breeding date of tree swallows throughout North America. *Proceedings of the Royal Society of London Series B – Biological Sciences* 266: 2487–2490.

Durant JM, Hjermann DO, Anker-Nilssen T, *et al.* (2005) Timing and abundance as key mechanisms affecting trophic interactions in variable environments. *Ecology Letters* 8: 952–958.

Forchhammer MC, Post E, and Stenseth NC (1998) Breeding phenology and climate. *Nature* 391: 29–30.

Forister ML and Shapiro AM (2003) Climatic trends and advancing spring flight of butterflies in lowland California. *Global Change Biology* 9: 1130–1135.

Gienapp P, Hemerik L, and Visser ME (2005) A new statistical tool to predict phenology under climate change scenarios. *Global Change Biology* 11: 600–606.

Gienapp P, Leimu R, and Merila J (2007) Responses to climate change in avian migration time – Microevolution versus phenotypic plasticity. *Climate Research* 35: 25–35.

Gienapp P, Teplitsky C, Alho JS, Mills JA, and Merila J (2008) Climate change and evolution: Disentangling environmental and genetic responses. *Molecular Ecology* 17: 167–178.

Gordo O (2007) Why are bird migration dates shifting? A review of weather and climate effects on avian migratory phenology. *Climate Research* 35: 37–58.

Gordo O and Sanz JJ (2005) Phenology and climate change: A long-term study in a Mediterranean locality. *Oecologia* 146: 484–495.

Hallett TB, Coulson T, Pilkington JG, Clutton-Brock TH, Pemberton JM, and Grenfell BT (2004) Why large-scale climate indices seem to predict ecological processes better than local weather. *Nature* 430: 71–75.

Harrington R, Woiwod I, and Sparks T (1999) Climate change and trophic interactions. *Trends in Ecology and Evolution* 14: 146–150.

Hüppop O and Hüppop K (2003) North Atlantic Oscillation and timing of spring migration in birds. *Proceedings of the Royal Society of London Series B – Biological Sciences* 270: 233–240.

Hurrell JW, Kushnir Y, and Visbeck M (2001) The North Atlantic Oscillation. *Science* 291: 603–605.

Husby A, Kruuk LEB, and Visser ME (2009) Decline in the frequency and benefits of multiple brooding in great tits as a consequence of a changing environment. *Proceedings of the Royal Society B – Biological Sciences* 276: 1845–1854.

Husby A, Nussey DH, Visser ME, Wilson AJ, Sheldon BC, and Kruuk LEB (2010) Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (*Parus major)* populations. *Evolution* 64: 2221–2237.

Inouye DW, Barr B, Armitage KB, and Inouye BD (2000) Climate change is affecting altitudinal migrants and hibernating species. *Proceedings of the National Academy of Sciences of the United States of America* 97: 1630–1633.

IPCC (2007) Summary for policymakers. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL (eds.) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press.

Jenni L and Kéry M (2003) Timing of autumn bird migration under climate change: Advances in long-distance migrants, delays in short-distance migrants. *Proceedings of the Royal Society of London Series B-Biological Sciences* 270: 1467–1471.

Jenouvrier S and Visser ME (2011) Integrating ecological and evolutionary processes to predict population responses to climate change. *International Journal of Biometeorology* 55: 905–919.

Jonzen N, Linden A, Ergon T, *et al.* (2006) Rapid advance of spring arrival dates in long-distance migratory birds. *Science* 312: 1959–1961.

Kleijn D, Schekkerman H, Dimmers WJ, Van Kats RJM, Melman D, and Teunissen WA (2010) Adverse effects of agricultural intensification and climate change on breeding habitat quality of Black-tailed Godwits Limosa l. limosa in the Netherlands. *Ibis* 152: 475–486.

Lehikoinen E, Sparks TH, and Zalakevicius M (2004) Arrival and departure dates. *Birds and Climate Change* 35: 1–31.

Macmynowski DP and Root TL (2007) Climate and the complexity of migratory phenology: Sexes, migratory distance, and arrival distributions. *International Journal of Biometeorology* 51(5): 361–373.

Menzel A, Sparks TH, Estrella N, *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology* 12: 1969–1976.

Miller-Rushing AJ, Hoye TT, Inouye DW, and Post E (2010) The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365: 3177–3186.

Miller-Rushing AJ, Lloyd-Evanss TL, Primack RB, and Satzinger P (2008) Bird migration times, climate change, and changing population sizes. *Global Change Biology* 14: 1–14.

Moller AP, Rubolini D, and Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are

declining. *Proceedings of the National Academy of Sciences of the United States of America* 105: 16195–16200.
Moussus JP, Julliard R, and Jiguet F (2010) Featuring 10 phenological estimators using simulated data. *Methods in Ecology and Evolution* 1: 140–150.
Moyes K, Nussey DH, Clements MN, *et al.* (2011) Advancing breeding phenology in response to environmental change in a wild red deer population. *Global Change Biology* 17: 2455–2469.
Nussey DH, Postma E, Gienapp P, and Visser ME (2005) Selection on heritable phenotypic plasticity in a wild bird population. *Science* 310: 304–306.
Parmesan C (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology* 13: 1860–1872.
Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.
Penuelas J, Filella I, and Comas P (2002) Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. *Global Change Biology* 8: 531–544.
Pigliucci M (2001) Phenotypic Plasticity; Beyond Nature and Nurture. Baltimore: John Hopkins University Press.
Post E and Forchhammer MC (2008) Climate change reduces reproductive success of an Arctic herbivore through trophic mismatch. *Philosophical Transactions of the Royal Society B – Biological Sciences* 363: 2369–2375.
Reed T, Waples R, Schindler D, Hard J, and Kinnison M (2010) Phenotypic plasticity and population viability: The importance of environmental predictability. *Proceedings of the Royal Society B: Biological Sciences* 277: 3391–3400.
Reed TE, Schindler DE, Hague MJ, *et al.* (2011) Time to evolve? Potential evolutionary responses of Fraser River Sockeye Salmon to climate change and effects on persistence. *PLoS ONE* 6: e20380.
Reed TE, Wanless S, Harris MP, Frederiksen M, Kruuk LEB, and Cunningham EJA (2006) Responding to environmental change: Plastic responses vary little in a synchronous breeder. *Proceedings of the Royal Society B-Biological Sciences* 273: 2713–2719.
Reed TE, Warzybok P, Wilson AJ, Bradley RW, Wanless S, and Sydeman WJ (2009) Timing is everything: Flexible phenology and shifting selection in a colonial seabird. *Journal of Animal Ecology* 78: 376–387.
Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, and Pounds JA (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421: 57–60.
Roy DB and Sparks TH (2000) Phenology of British butterflies and climate change. *Global Change Biology* 6: 407–416.
Sanz JJ (2002) Climate change and breeding parameters of great and blue tits throughout the western Palaearctic. *Global Change Biology* 8: 409–422.
Sanz JJ (2003) Large-scale effect of climate change on breeding parameters of pied flycatchers in Western Europe. *Ecography* 26: 45–50.
Schaper SV, Rueda C, Sharp P, Dawson AS, and Visser ME (2011) Spring phenology does not affect timing of reproduction in the great tit *(Parus major). Journal of Experimental Biology* 214: 3664–3671.
Silverin B, Wingfield JC, Stokkan KA, *et al.* (2008) Ambient temperature effects on photo induced gonadal cycles and hormonal secretion patterns in Great Tits from three different breeding latitudes. *Hormones and Behavior* 54: 60–68.
Sims DW, Genner MJ, Southward AJ, and Hawkins SJ (2001) Timing of squid migration reflects North Atlantic climate variability. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268: 2607–2611.
Sims DW, Wearmouth VJ, Genner MJ, Southward AJ, and Hawkins SJ (2004) Low-temperature-driven early spawning migration of a temperate marine fish. *Journal of Animal Ecology* 73: 333–341.
Sokolov LV (2006) Effect of global warming on the timing of migration and breeding of Passerine birds in the 20th century. *Entomological Review* 86: S59–S81.
Stefanescu C, Penuelas J, and Filella I (2003) Effects of climatic change on the phenology of butterflies in the northwest Mediterranean Basin. *Global Change Biology* 9: 1494–1506.
Stenseth NC and Mysterud A (2002) Climate, changing phenology, and other life history and traits: Nonlinearity and match–mismatch to the environment. *Proceedings of the National Academy of Sciences of the United States of America* 99: 13379–13381.
Taylor SG (2008) Climate warming causes phenological shift in Pink Salmon, *Oncorhynchus gorbuscha*, behavior at Auke Creek, Alaska. *Global Change Biology* 14: 229–235.
Thackeray SJ, Sparks TH, Frederiksen M, *et al.* (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology* 16: 3304–3313.
Thomas DW, Blondel J, Perret P, Lambrechts MM, and Speakman JR (2001) Energetic and fitness costs of mismatching resource supply and demand in seasonally breeding birds. *Science* 291: 2598–2600.
Trenberth KE and Caron JM (2000) The Southern Oscillation revisited: Sea level pressures, surface temperatures, and precipitation. *Journal of Climate* 13: 4358–4365.
Tryjanowski P and Sparks TH (2001) Is the detection of the first arrival date of migrating birds influenced by population size? A case study of the red-backed shrike Lanius collurio. *International Journal of Biometeorology* 45: 217–219.
Van Asch M, Van Tienderen PH, Holleman LJM, and Visser ME (2007) Predicting shifts in phenology in response to climate change, an insect herbivore example. *Global Change Biology* 13: 1596–1604.
Van Asch M and Visser ME (2007) Phenology of forest caterpillars and their host trees: The importance of synchrony. *Annual Review of Entomology* 52: 37–55.
Van Buskirk J, Mulvihill RS, and Leberman RC (2009) Variable shifts in spring and autumn migration phenology in North American songbirds associated with climate change. *Global Change Biology* 15: 760–771.
Van de Pol M and Cockburn A (2011) Identifying the critical climatic time window that affects trait expression. *American Naturalist* 177: 698–707.
Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society of London Series B-Biological Sciences* 275: 649–659.
Visser ME, Adriaensen F, Van Balen JH, *et al.* (2003) Variable responses to large-scale climate change in European Parus populations. *Proceedings of the Royal Society of London Series B-Biological Sciences* 270: 367–372.
Visser ME and Both C (2005) Shifts in phenology due to global climate change: The need for a yardstick. *Proceedings of the Royal Society of London Series B – Biological Sciences* 272: 2561–2569.
Visser ME, Both C, and Lambrechts MM (2004) Global climate change leads to mistimed avian reproduction. *Advances in Ecological Research* 35: 89–110.
Visser ME, Caro SP, Van Oers K, Schaper SV, and Helm B (2010) Phenology, seasonal timing and circannual rhythms: Towards a unified framework. *Philosophical Transactions B* 365: 3113–3127.
Visser ME and Holleman LJM (2001) Warmer springs disrupt the synchrony of oak and winter moth phenology. *Proceedings of the Royal Society of London Series B – Biological Sciences* 268: 289–294.
Visser ME, Holleman LJM, and Caro SP (2009) Temperature has a causal effect on avian timing of reproduction. *Proceedings of the Royal Society B – Biological Sciences* 276: 2323–2331.
Visser ME, Te Marvelde L, and Lof M (2011). Adaptive phenological mismatches of birds and their food in a warming world. Journal of Ornithology, Subm.
Visser ME, Van Noordwijk AJ, Tinbergen JM, and Lessells CM (1998) Warmer springs lead to mistimed reproduction in great tits (*Parus major*). *Proceedings of the Royal Society of London Series B – Biological Sciences* 265: 1867–1870.
Walther GR, Post E, Convey P, *et al.* (2002) Ecological responses to recent climate change. *Nature* 416: 389–395.
Zann RA, Morton SR, Jones KR, and Burley NT (1995) The timing of breeding by Zebra finches in relation to rainfall in Central Australia. *Emu* 95: 208–222.

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF
Simon A Levin
Princeton University, Princeton, NJ, USA

VOLUME 6



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xv

Contributors to Volume 6 xxi

Foreword to the First Edition xxv

Preface to the First Edition xxvii

Preface to the Second Edition xxix

How to use of the Encyclopedia xxxi

VOLUME 6

Phenotype, a Historical Perspective 1
Robert J Berry

Photosynthesis, Mechanisms of 9
John A Raven

Phylogeny 16
Kevin C Nixon

Plankton, Status and Role of 24
Colin S Reynolds and Judit Padisák

Plant–Animal Interactions 39
Ellen L Simms

Plant Biodiversity, Overview 56
Jeannette Whitton

Plant Communities, Evolution of 65
Brian J Enquist, Bruce H Tiffney, John Haskell, and Karl J Niklas

Plant Conservation 76
Mike Maunder

Plant Invasions 90
David M Richardson and Petr Pyšek

Plant Phenology Changes and Climate Change 103
Tim Sparks and Annette Menzel

Plant–Soil Interactions 109
Joan G Ehrenfeld

Plant Sources of Drugs and Chemicals 129
William H Gerwick

Pollinators, Role of 140
David W Inouye

Pollution, Overview 147
William H Smith

Population Density 157
Brian H McArdle

Population Diversity, Overview 168
Jennifer BH Martiny

Population Dynamics 175
Alan Hastings

Population Genetics 182
Brian Charlesworth

Population Stabilization, Human 199
Samir KC

Population Viability Analysis 210
Hugh P Possingham, Michael A McCarthy, and David B Lindenmayer

Poverty and Biodiversity 220
Madhav Gadgil

Predators, Ecological Role of 229
James Estes, Kevin Crooks, and Robert D Holt

Primate Populations, Conservation of 250
Russell A Mittermeier and Matthew C Richardson

Priority Setting for Biodiversity and Ecosystem Services 261
Derric Pennington, Taylor Ricketts, and Robin Naidoo

Property Rights and Biodiversity 273
Susan S Hanna

Proteorhodopsins: Widespread Microbial Light-Driven Proton Pumps 280
Oded Béjà, Jarone Pinhassi, and John L Spudich

Protozoa 286
Bland J Finlay and Genoveva F Esteban

Psychrophiles 298
Richard Y Morita and Craig L Moyer

R

Rainforest Ecosystems, Animal Diversity 304
Gregory H Adler

Rainforest Ecosystems, Plant Diversity 313
Ian M Turner

Rainforest Loss and Change 322
Karndeo D Singh

Range Ecology, Global Livestock Influences 330
J Boone Kauffman and David A Pyke

Rare Biosphere 345
Carlos Pedrós-Alió

Recombination 353
Abraham B Korol

Reforestation 370
David Lamb

Religious Traditions and Biodiversity 380
Fikret Berkes

Remote Sensing and Image Processing 389
Ruth DeFries

Reptiles, Biodiversity of 400
F Harvey Pough

Residents and Transients in the Fossil Record 414
Albert J van der Meulen and Pablo Peláez-Campomanes

Resource Exploitation, Fisheries 420
John Beddington and Ray Hilborn

Resource Partitioning 429
FA Bazzaz and S Catovsky

Restoration of Animal, Plant, and Microbial Diversity 438
Edith B Allen, Joel S Brown, and Michael F Allen

Restoration of Biodiversity, Overview 453
Joy B Zedler and Roberto Lindig-Cisneros

Riparian Landscapes 461
Robert J Naiman, Henri Décamps, and Michael E McClain

River Ecosystems 469
KE Limburg, DP Swaney, and DL Strayer

Role and Trends of Protected Areas in Conservation 485
Timothy M Boucher, Mark Spalding, and Carmen Revenga

Role of Botanic Gardens 504
Peter Wyse Jackson and Lucy A Sutherland

S

Salmon 522
Michael H Schiewe

Scale, Concept and Effects of 532
David Clayton Schneider

Seagrasses 540
Carlos M Duarte

Slash-and-Burn Agriculture, Effects of 551
Stefan Hauser and Lindsey Norgrove

Social and Cultural Factors 563
Jeffrey A McNeely

Social Behavior 571
Daniel I Rubenstein and Dustin R Rubenstein

Soil Biota, Soil Systems, and Processes 580
David C Coleman

Soil Conservation 590
Dorota L Porazinska and Diana H Wall

South American Natural Ecosystems, Status of 599
Philip M Fearnside

Southern (Austral) Ecosystems 612
Peter H Weston and Robert S Hill

Speciation, Process of 620
Guy L Bush

Speciation, Theories of 629
Hope Hollocher

Species–Area Relationships 640
EF Conor and ED McCoy

Species Assemblages, Macroecology, and Global Change 651
Brian J McGill

Species Coexistence 667
Robert D Holt

Species, Concepts of 679
James Mallet

Species Distribution Modeling 692
Jane Elith and Janet Franklin

Species Diversity, Overview 706
A Ross Kiester

Species Interactions 715
Jessica J Hellmann

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 6

Gregory H Adler
University of Wisconsin at Oshkosh, Oshkosh, WI, USA

Edith B Allen
University of California, CA, USA

Michael F Allen
Center for Conservation Biology, University of California, CA, USA

FA Bazzaz
Harvard University, MA, USA

John Beddington
Imperial College of Science, Technology and Medicine, London, UK

Oded Béjà
Technion – Israel Institute of Technology, Haifa, Israel

Fikret Berkes
University of Manitoba, Winnipeg, Canada

Robert J Berry
University College London, London, UK

J Boone Kauffman
Oregon State University, OR, USA

Timothy M Boucher
The Nature Conservancy, VA, USA

Joel S Brown
University of Illinois, IL, USA

Guy L Bush
Michigan State University, MI, USA

S Catovsky
Harvard University, MA, USA

Brian Charlesworth
University of Edinburgh, Edinburgh, UK

David C Coleman
University of Georgia, GA, USA

EF Conor
San Francisco State University, CA, USA

Kevin Crooks
Colorado State University, CO, USA

Henri Décamps
Université Paul Sabatier, Toulouse, France

Ruth DeFries
Columbia University, NY, USA

Carlos M Duarte
IMEDEA, CSIC-Universitat de les Illes Balears, Spain

Joan G Ehrenfeld[†]
Rutgers University, NJ, USA

Jane Elith
The University of Melbourne, Parkville, VIC, Australia

Brian J Enquist
University of California at Santa Barbara, CA, USA

Genoveva F Esteban
Bournemouth University, Poole, UK

James Estes
University of California, CA, USA

Philip M Fearnside
National Institute for Research in the Amazon (INPA), Brazil

Bland J Finlay
Queen Mary University of London, Wareham, UK

Janet Franklin
Arizona State University, Tempe, AZ, USA

Madhav Gadgil
Indian Institute of Science, Bangalore, India

William H Gerwick
University of California, San Diego, CA, USA

Susan S Hanna
State University Corvallis, OR, USA

[†]Deceased.

F Harvey Pough
School of Life Sciences, Rochester Institute of Technology, NY, USA

John Haskell
University of New Mexico, NM, USA

Alan Hastings
University of California, Davis, CA, USA

Stefan Hauser
IITA Democratic Republic of Congo, Kinshasa, DR Congo

Jessica J Hellmann
University of Notre Dame, Notre Dame, IN, USA

Ray Hilborn
University of Washington, WA, USA

Robert S Hill
University of Adelaide, SA, Australia

Hope Hollocher
University of Notre Dame, Notre Dame, IN, USA

Robert D Holt
University of Florida, Gainesville, FL, USA

David W Inouye
University of Maryland, MA, USA

Peter Wyse Jackson
Missouri Botanical Garden, St Louis, Missouri, USA

Samir KC
International Institute for Applied Systems Analysis, Laxenburg, Austria

Abraham B Korol
University of Haifa, Haifa, Israel

David Lamb
University of Queensland, Brisbane, Australia

KE Limburg
State University of New York, NY, USA

David B Lindenmayer
The Australian National University, Canberra, Australia

Roberto Lindig-Cisneros
Universidad National Autónomo de México, México

James Mallet
University College London, London, UK

Jennifer BH Martiny
University of California, Irvine, CA, USA

Mike Maunder
Florida International University, FL, USA

Brian H McArdle
University of Auckland, Auckland, New Zealand

Michael A McCarthy
University of Melbourne, Melbourne, Australia

Michael E McClain
UNESCO-Institute for Water Education (IHE), The Netherlands

ED McCoy
University of South Florida, FL, USA

Brian J McGill
University of Maine, ME, USA

Jeffrey A McNeely
The International Union for Conservation of Nature (IUCN), Gland, Switzerland

Annette Menzel
Technische Universität München, Germany

Albert J van der Meulen
University of Utrecht, Netherlands

Russell A Mittermeier
Conservation International, VA, USA

Richard Y Morita
Oregon State University, OR, USA

Craig L Moyer
Western Washington University, WA, USA

Robin Naidoo
World Wildlife Fund – US, Washington, DC, USA

Robert J Naiman
University of Washington, WA, USA

Karl J Niklas
Cornell University, NY, USA

Kevin C Nixon
Cornell University, NY, USA

Lindsey Norgrove
CABI, Delémont, Switzerland

Judit Padisák
University of Pannonia, Hungary

Carlos Pedrós-Alió
Institut de Ciències del Mar, CSIC, Passeig Marítim de la Barceloneta, Barcelona, Spain

Pablo Peláez-Campomanes
National Museum of Natural History, CSIC, Spain

Derric Pennington
University of Minnesota, St Paul, MN, USA, and World Wildlife Fund – US, Washington, DC, USA

Jarone Pinhassi
Linnaeus University, Kalmar, Sweden

Dorota L Porazinska
University of Florida, FL, USA

Hugh P Possingham
The University of Queensland, Queensland, Australia

Petr Pyšek
Academy of Sciences and Charles University, Prague, Czech Republic

David A Pyke
USGS Forest and Rangeland Ecosystems Science Center, OR, USA

John A Raven
University of Dundee, Dundee, UK

Carmen Revenga
The Nature Conservancy, VA, USA

Colin S Reynolds
Freshwater Biological Association and NERC Institute of Freshwater Ecology, Cumbria, UK

David M Richardson
Stellenbosch University, Matieland, South Africa

Matthew C Richardson
Conservation International, VA, USA

Taylor Ricketts
University of Vermont, Burlington, VT, USA

A Ross Kiester
Turtle Conservancy, NY, USA

Daniel I Rubenstein
Princeton University, Princeton, NJ, USA

Dustin R Rubenstein
Columbia University, New York, NY, USA

Michael H Schiewe
Anchor QEA, WA, USA

David Clayton Schneider
Memorial University of Newfoundland, Newfoundland and Labrador, Canada

Ellen L Simms
University of California, Berkeley, CA, USA

Karndeo D Singh
Academy of Forests and Environmental Sciences, New Delhi, India

William H Smith
Yale University, CT, USA

Mark Spalding
The Nature Conservancy, Newmarket, UK

Tim Sparks
Technische Universität München, Germany; Poznań University of Life Sciences, Poland, and University of Cambridge, UK

John L Spudich
The University of Texas Medical School, Houston, TX, USA

DL Strayer
Cary Institute of Ecosystem Studies, Millbrook, NY, USA

Lucy A Sutherland
Australian National Botanic Gardens, Canberra, ACT, Australia

DP Swaney
Cornell University, NY, USA

Bruce H Tiffney
University of California at Santa Barbara, CA, USA

Ian M Turner
Singapore Botanic Gardens, Singapore

Diana H Wall
Colorado State University, Fort Collins, CO, USA

Peter H Weston
University of Adelaide and National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, NSW, Australia

Jeannette Whitton
University of British Columbia, Vancouver, Canada

Joy B Zedler
University of Wisconsin, WI, USA

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

P

Phenotype, a Historical Perspective

Robert J Berry, University College London, London, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 1, pp 537–547, © 2001, Elsevier Inc.

Glossary

Allele (originally **allelomorph)** A form of variant (sometimes called a mutation) of a gene.
Diploid Individual who has two sets of chromosomes, usually received from different parents, in contrast to a **haploid** organism such as most microorganisms, the gametophyte stage of higher plants and some parthenogenetic forms (e.g., male ants and honeybees). It may have the same form or allele of a gene on both members of a chromosome pair (in which case it is homozygous) in different forms (alleles) (in which case it is heterozygous).
Gene Historically, the inherited factor which determines a trait. Tends to be used somewhat loosely; more strictly represents a place or locus on the chromosomes which codes for a particular function.
Gene (strictly **allele**) **frequency** The frequency of an allele in a population.
Genome The total genetic composition of an individual.
Genotype The allelic composition of an individual at a particular locus.
Mutation Change in an allele, producing a different allele; rate of occurrence affected both physically (especially by ionizing radiation) and by many chemicals. It may also refer to changes in a chromosome (involving duplication, deletion, or inversion of a segment).
Phenetics The study of phenotypes, usually describing the grouping of organisms into taxa on the basis of estimates of similarity.
Phenotype The appearance (function or behavior) of an organism.

Background

Phenotypes are the appearance and properties of real animals and plants—living and dying, reproducing and bearing fruit, and succeeding or failing—but in themselves they are the products of interactions between the inherited (genetic) material and the environment, both before and after birth. The distinction between genotype and phenotype is one of the more important ones in biology: A phenotype may be produced by several different genotypes, whereas a genotype may manifest as several different phenotypes, particularly when reacting with different environments.

The distinction between inherited constitution and external appearance (i.e., between genotype and phenotype) was one of the most important demonstrations of Gregor Mendel (1822–1884) in his series of breeding experiments with peas which laid the foundation for genetics, and in which he recognized that a character may be inherited in a dominant or a recessive manner. An organism manifesting a dominantly inherited trait (in Mendel's case, round as opposed to wrinkled pods or colored as opposed to white flowers; more familiar examples are brown versus blue eyes in humans or brown coat color versus albino in rabbits or mice) may carry both alleles for the dominant trait or one for the dominant and one for the recessive trait. That is, it will have the same phenotype but could be genetically homozygous or heterozygous. A distinction between genes and their manifestation was also implicit in August Weismann's (1834–1914) embryological division between germplasm (which gives rise to reproductive cells) and soma (or body). However, the formal nomenclature phenotype and genotype was devised by the Danish botanist Wilhelm Johannsen (1857–1927), who introduced the word gene for the material basis of an inherited character, and thence the terms genotype and phenotype.

Johannsen set out to investigate the relationship between (phenotypic) variation and selection. The rediscovery of Mendel's work in 1900 led to a rift between the biometricians [notably Karl Pearson (1857–1936) and W. F. R. Weldon (1860–1906)], who followed Darwin (1809–1882) in viewing small continuous variation as the raw material of evolution, and the "mendelists" (or geneticists) [led by William Bateson, (1861–1926)], who believed that large discontinous saltations (or mutations) were the main cause of variation. The problem was the maintenance of continuous variation. Darwin had postulated a continuous replenishment of variation in order for selection to act, but he did not know its source. His

proposal of "pangenes" was an effort to solve the problem. Weismann's demonstration of the early separation in development of the reproductive tissue from the rest of the body supported the general assumption at the time that continuous variation was produced by environmental effects. Intraspecific variation (i.e., subspecies or local races) was therefore regarded as environmentally caused; the species was viewed as monotypic in a way that had much in common with the pre-Darwinian ideal or Linnean type.

Johannsen experimented with a self-fertilized cultivar of the bean *Phaseolus vulgaris*. The implication of this was that all the descendants of a single individual would have the same genes and constitute what Johannsen called a "pure line." Although individual beans might be different (due, for example, to their place in the pod), the mean and variance of all the characters of plants in a pure line were the same and were not affected by attempts at selection, i.e., plants grown from both small beans and large beans produced beans of the same average weight as that of all the plants in the pure line. In contrast, plants grown from crosses between pure lines had different (usually intermediate) characteristics from the parents, although these characteristics remained constant in pure lines derived from each cross. Johannsen argued that selection on continuous variation was inevitably ineffective, and the only variation on which selection could work depended on new mutation. This strengthened the contemporary assumption that evolution was bound up with mutations and their rate of occurrence, and that individuals were largely genetically uniform (i.e., homozygous at most loci for "wild-type" alleles).

This assumption became built into conventional population genetic theory so that when recessive traits manifested in inbred populations (usually in the laboratory or garden, but occasionally appearing under wild conditions) they were assumed to be recent mutations in the process of elimination by natural selection. Most mutations seemed to be deleterious to their carriers (which would be expected if they were random changes in a functioning organism), which meant that high mutation rates would inevitably impose a burden on a population. In 1950, at a time of acute concern about the genetic effects of atomic warfare, H. J. Müller (1890–1967) proposed the concept of "genetic load" and showed that a doubling of the mutation rate in a slow-reproducing species such as humankind could lead to extinction through genetically caused death.

The work of Bateson, Johannsen, and other early geneticists (one of the most influential was the Dutchman Hugo de Vries, 1848–1935) led to the isolation of genetics from evolutionary studies, particularly as represented by paleontology. This was resolved by the theoretical work of Sewall Wright (1889–1988), J. B. S. Haldane (1892–1964), and especially R. A. Fisher (1890–1962). Fisher published significant papers in 1918 and 1922 describing the expected biometrical properties of a Mendelian (i.e., breeding) population and the effects of an allele substitution on a quantitative (i.e., continually varying) character. (The papers were published by the Royal Society of Edinburgh; the first one was rejected by the Royal Society of London on the advice of Karl Pearson and the geneticist R. C. Punnett.) He went on to argue that dominance (and recessivity) are traits that have evolved and this explains why the majority of new mutations are recessive, detrimental, and have a major effect (i.e., on the phenotype); most of the data available at the time were from laboratory breeding of *Drosophila melanogaster*. Fisher reasoned that there was no intrinsic reason for a mutation occurring for the first time to be either dominant or recessive; the greatest probability is that it will be intermediate, with an effect somewhere between its expression in double dose (i.e., when homozygous) and the unmodified condition. However,

1. Mutations occur repeatedly at virtually every locus. The rare (approximately 1 in 10^{-5}) mutational events we observe are recurrences of something that has happened thousands of times in the past.
2. When a mutation occurs, it will almost always be present in the heterozygous condition: If an allele has become relatively common in a population so that a fresh mutation to it has a reasonable chance of occurring in an existing heterozygote, then mutation cannot be the only force influencing its frequency.
3. If a newly arisen allele has a beneficial effect on its carrier, combinations of it with other alleles that increase its effect will have a higher fitness than any which decrease it. This will repeatedly occur so that the architecture of the species will become modified to the extent that any new occurrences of that advantageous allele will always produce the maximum effect in its carrier; this will almost always be in the heterozygous condition. In other words, beneficial characters will be selected for dominance and will also spread to replace the previous expression of the trait. Conversely, alleles which are deleterious in the heterozygote are only likely to be transmitted in combinations in which their effect is least, i.e., there will be selection for a small heterozygous effect (in the direction of recessivity).

Fisher put forward these ideas from first principles. They were received skeptically because of the difficulty in believing that selection pressures would be strong enough to allow genes which modify dominance to spread. In pre-1950 days primary selection coefficients were thought to be approximately 0.1 to 1.0% and second-order effects were believed to be much less. Haldane suggested that dominance was more likely to be the effect of alleles with a biochemical or developmental margin of safety becoming the normal allele. Since they could exercise undiminished action when heterozygous, mutant alleles would be recessive and deleterious. However, Fisher's theory has been proven to be correct on many occasions, and although it may not apply for every allele at every locus it has considerable historical significance in bringing together paleontologists and geneticists, and it has relevance in highlighting an important factor influencing genetical architecture.

Fisher was the first to demonstrate the experimental modification of dominance by crossing domestic poultry with wild jungle fowl for five generations. This changed the inheritance of certain characters so that a degree of heterozygous manifestation occurred where complete dominance had previously prevailed. Fisher suggested that the dominance of the traits he studied had been attained during domestication as a result of selection for the more striking heterozygotes.

A more complete demonstration of the influences of modifying genes on dominance was performed by E. B. Ford (1901–1985) by breeding from the greatest and least expressions of a variable yellow variant (*lutea*) of the Currant Moth (*Abraxas grossulariata*). Although the difference between *lutea* and non-*lutea* can be regarded as caused by a single allelic difference, after only three generations of selection Ford produced heterozygotes virtually indistinguishable from the *lutea* homozygote in selection for the yellowest individuals and ones most like the typical homozygote in the white selection line. In other words, he had changed the heterozygote from a position of no (or intermediate) dominance to complete dominance or recessivity respectively. Ford then crossed his modified heterozygotes with unselected stock, and by the second generation (when the selected modifiers would have a chance of segregating independently) the original variable heterozygotes reappeared: He thus showed that it was the response of the organism rather than the gene that had changed.

Laboratory experiments of this nature have been carried out on a variety of organisms and a range of characters. An experiment particularly informative with regard to genetical architecture in the wild was carried out by H. B. D. Kettlewell (1907–1979) using British and Canadian peppered moths (*Biston betularia* and *Amphydasis cognataria*, respectively; these are fully interfertile). Although there is a melanic form (*swettaria*) in the North American species, it is comparatively restricted in its distribution, and only pale (or typical) moths occur over vast tracts of Canada.

When a melanic heterozygote and a typical homozygote of British origin are mated, the offspring are clearly dark or light: The melanic character is a straightforward dominant. Even when the typical moths come from Cornwall in extreme southwest England where melanics have never been reported, only a slight loss of complete dominance occurs and that only after several generations of crossing melanics back to Cornish stock (i.e., back-crossing a melanic from several generations of melanic × Cornish cross with a "pure" Cornish parent). This modification consists of some white dots on the normally jet-black wings of the heterozygote. Perhaps significantly, this slight heterozygous expression of the gene gives specimens similar to those caught in the early days of the spread of peppered moth melanism in the mid-nineteenth century and which are now prized collectors' specimens. This white speckling has long since disappeared in wild-caught British specimens, and modern heterozygotes are indistinguishable from melanic homozygotes. Melanism has become fully dominant during the ensuing decades.

However, there has been no opportunity for dominance to evolve in Canadian peppered moths. When British melanics are crossed to Canadian stock (from areas where *swettaria* does not occur) the first-generation progeny segregate as dark or light in the same way as in a cross between British moths. In the first generation, the dominance modifiers in the British parent will be carried on the chromosomes in the same order as in British stock and produce dominance in the same way. In the next generation, the gametes contain chromosomes which have crossed over between the British and Canadian grandparents. Consequently, in the second generation, the "switch" between pale and black forms does not operate as efficiently. Kettlewell crossed heterozygotes from a British × Canadian mating with Canadian moths and repeated this back-crossing for four consecutive generations, after which the heterozygotes ranged from black to pale—there was no sign of dominance. He then reversed the procedure and mated his "broken-down" melanics to British typicals. The dominance of the condition was immediately reestablished: The architecture of the British chromosomes shaped a clear segregation between a dark heterozygote and a pale homozygote.

Fisher's theory has proved correct in many similar experiments in both plants and animals, and although it may not be universally operative it provided the genetical basis for the understanding between disciplines that was needed before the neo-Darwinian synthesis could occur.

As late as 1932, T. H. Morgan was asserting that "natural selection does not play the role of a creative principle in evolution," but 10 years later all but a very few biologists were agreed on an evolutionary theory based firmly on Darwin's ideas knitted with subsequent developments in genetics. This coming together was described by Julian Huxley as the "modern synthesis" in a book with the same name published in 1942. The synthesis can first be seen in three English books: R. A. Fisher's *Genetical Theory of Natural Selection* (1930), E. B. Ford's *Mendelism and Evolution* (1931), and J. B. S. Haldane's *Causes of Evolution* (1932). It was consolidated in three works from America: Theodosius Dobzhanky's *Genetics and the Origin of Species* (1937), Ernst Mayr's *Systematics and the Origin of Species* (1942), and George Gaylord Simpson's *Tempo and Mode in Evolution* (1944). As Mayr noted, it did not occur as a result of one side being proved right and the others wrong but rather from "an exchange of the most viable components of the previously competing research traditions."

The Making of a Phenotype

In the simplest microbial systems, there is a one to one relationship between gene action and phenotype: Change in a gene is likely to produce a change in its product, which will manifest directly in the organism as an altered character, perhaps the loss or modification of an enzyme. Such simple relationships exist in all organisms, however complex. There are more than 200 "inborn errors of metabolism" in humans, each resulting from changes in a particular gene, producing a change of phenotype in the whole organism. For example, albinism is due to the inability to synthesize melanin, which is made from tyrosine under the influence of tyrosinase, which is under the control of a gene on chromosome 11 in humans[1]; "classical" hemophilia results from the absence of a protein (factor VIII) coded by a gene on the X chromosome. However, most traits are affected by many genes. For example, blood clotting in mammals is effected by a cascade of physiological reactions, each under the control of a different gene(s). Efficient clotting requires all the different stages to be operating, and defects in any stage (especially the genes directing the relevant proteins and enzymes)

[1]In fact, human albinism can be due to genes on at least two different chromosomes. There are also mutations at another locus on chromosome 11 which cause albinism, but in this case the tyrosinase enzyme is apparently normal. Both genes produce an albino phenotype, but the phenotypes can be distinguished clinically.

will lead to a "bleeding" phenotype. In blood clotting it is possible to detect where the error lies because we know the normal determinants of clotting in detail. For most traits we do not have information about the steps in their formation. One of the benefits of genome mapping is that the genetical determinants of complex characters will be analyzable in an orderly way. Currently, most phenotypic analysis is biometric rather than genetic.

For example, the tails of house mice may be shortened by mutations at approximately 40 loci. We know the action of many of these genes: Some affect the notochord, some cell division rates, and some inductive relations between endoderm and mesectoderm. We also know that the tail length of mice varies between different inbred strains of mice (which are similar in the genetic sense to Johannsen's pure lines) and that we can increase or decrease tail length by selecting wild-caught animals or the products of crosses between inbred strains. We do not know which genes are variable (or segregating) in any one population or which genes are being affected by selection. However, it is clear from many selection experiments that the genes affecting any complex trait are distributed throughout the genome (i.e., over many chromosomes) and that selection for any one trait may "unbalance" the genome.

The concept of genomic balance (often called genomic or genetic architecture) is important. When a normal outbreeding species, such as maize, sugar beet, poultry, or *Drosophila*, is made to inbreed by manipulating its breeding system, the individuals characteristically decline in vigor and fertility until they stabilize at a stage before complete homozygosis is approached. The amount of this inbreeding depression varies from one line to another within any species. When two inbred lines are intercrossed, the F_1 (i.e., the first-generation progeny) show a considerable increase in vigor and fertility, known as hybrid vigor or heterosis. In general, the F_1 phenotype is similar to that in the population from which the inbred strains were derived. If F_2's are raised and inbreeding is resumed, inbreeding depression will again occur, although not necessarily to the same extent as in the original lines. Inbreeding depression is often reflected in increased variability not only between individuals but also among repetitive parts such as bilateral characters in animals and floral morphology in plants; fluctuating asymmetry, in which there are differences between the right and left sides of individuals, is commonly used as a rough indicator of inbreeding. F_1's tend to show decreased variability. Michael Lerner argued that this indicates an innate superiority of heterozygotes over homozygotes, perhaps because of a greater biochemical flexibility in the former. It is a phenomenon which has produced considerable research interest, particularly among marine biologists. Lerner argued that if artificial selection is suspended before much of the variation has been lost (as homozygosity increases) natural selection will tend to restore the character (and hence its genetic determinants) to an equilibrium value, with the mean of the character which was artificially selected tending to revert toward its original value; he called this genetic homeostasis.

Artificial selection tends to accumulate alleles which act on a character in a particular way, e.g., to increase its size of or the number of elements making up a repeated trait. If we make the reasonable assumption that any population needs to combine phenotypic uniformity in a stable environment with long-term flexibility should the environment change, the easiest way to do this is for the character in question to be controlled by alleles at different loci, with some alleles acting to increase the expression of the character and others acting to decrease it. In such a case, Fisher showed that selection will favor linkage between the loci responsible, with the evolution of "balanced" chromosomes containing "positive" and "negative" alleles. The simplest situation will involve two segregating loci, A,a and B,b with A,B acting in one direction and a,b in the other (i.e., additive genes). The intermediate type can be either the attraction or the repulsion heterozygote, AB/ab or Ab/aB. However, the latter will be favored since it will be less likely to produce zygotes giving the extreme phenotypes (**Figure 1**). For the same reason, any mechanisms bringing about closer linkage between the loci concerned (such as a chromosomal inversion) will be favored.

This theoretical arrangement has been subjected to experimental analysis, mainly through selection experiments in *Drosophila*, and has been generally confirmed; it receives support from the distribution of "quantitative trait loci" identified by molecular techniques. It also accounts for other properties:

1. Selection (natural or artificial) for any character almost invariably produces correlated responses in additional developmentally independent traits of the phenotype.
2. Gene loci affecting viability traits are interspersed along the chromosomes with loci affecting other characters.
3. The highest rate of artificially induced new variation by mutation is many times less than that occurring spontaneously through recombination.
4. There is widespread occurrence of chromosomal inversions in nature.
5. Deleterious gene combinations (even to the point of lethality) may occur solely as a result of recombination.

There are a few cases in which the different genes contributing to complex variation have been identified. This has been done for the determinants of the pin thrum polymorphism in *Primula vulgaris*, in which different members of a linked group of loci control anther height, style length, pollen size, rate of pollen tube growth, and length of the papillae on the stigma; for mimetic patterns in the African swallowtail butterfly *Papilio dardanus*; and for color and banding patterns in the land snail *Cepaea nemoralis*. In house mice, approximately 16% of genes with an identified function are concerned with behavior and have been located on 19 of the 20 chromosomes of the species. However, in most cases it is only possible to conclude that different components of a character complex are inherited as a unit, i.e., that the gene complex is coadapted.

Coadaptation is presumably the reason why the genomes of interfertile species do not always merge when they meet. For example, races of dark- and light-bellied house mice (referred to as *Mus musculus domesticus* and *M. m. musculus*, respectively, although they should perhaps be given full species status) meet in a narrow zone of intergradation across Jutland (Denmark) and south through Germany. Although the two forms readily interbreed in the laboratory, the hybrid zone has apparently been constant for at least 50 years. Gene frequencies on the two sides of the hybrid zone are very different. Remarkably, the frequencies in the light-bellied form in California, which are descended from the same stock as the light-bellied Danish

Parents		Offspring					
		From non-recombinant gametes		With recombination in one parent		With recombination in both parents	
Chromosomes	Phenotypic score	Chromosomes	Phenotypic score	Chromosomes	Phenotypic score	Chromosomes	Phenotypic score
Repulsion							
A b / a B	2	Ab / aB	2	AB / aB	3	AB / ab	2
		Ab / Ab	2	AB / Ab	3	AB / AB	4
A b / a B	2	aB / aB	2	ab / aB	1	ab / ab	0
		aB / Ab	2	ab / Ab	1	AB / ab	2
Attraction							
A B / a b		AB / ab	2	AB / Ab	3	Ab / Ab	2
		AB / AB	4	AB / aB	3	Ab / aB	2
A B / a b		ab / ab	0	ab / Ab	1	aB / Ab	2
		ab / AB	2	ab / aB	1	aB / aB	2

Figure 1 Effect of linkage on the incidence of the extremes of manifestation of an additively inherited character. If A,B produce a phenotype effect of 1 and a,b an effect of 0, the matings between two repulsion heterozygotes are least likely to produce extreme phenotypes (they will only appear if recombination between the two loci occurs in both parents).

animals, are more like those in the light-bellied Danish population than are the light-bellied Danish ones from their dark-bellied neighbors. A similar situation exists in deer mice (*Peromyscus polionotus*) in Florida and also in carrion and hooded crows in the same area in which the dark- and light-bellied mice occur in Europe. However, some species lose their identity wholly when brought into breeding contact. This has happened commonly in New Zealand, where much of the terrestrial biota has been introduced. Hybrid zones are common. Presumably these occur when coadaptation has not evolved.

Formal Analysis

In the past, it was suggested that the genes which control major or qualitative traits are different from those which affect quantitative ones (oligogenes and polygenes, respectively). This distinction is now rarely made. In other words, it is assumed that genes affecting quantitative traits (such as weight or size or physiological properties such as metabolic rate) follow Mendelian patterns of inheritance, may have multiple alleles, can mutate, change in gene frequency, show dominance, etc. Quantitative inheritance is merely a general case of the interaction of genes in which the interacting components are little or wholly known.

The number of genes which affect a trait can be estimated by the amount and speed of response in a selection experiment or by the mean and variance of the character as measured in a population. Using these techniques it has been calculated that human skin color may be determined by only 5 or 6 loci, whereas the number of genes affecting oil and protein production in maize may be as high as 54 and 122, respectively. However, such estimates are very dependent on the nature of interactions between the loci concerned and should be regarded as no more than suggestive of a large or small number.

The simplest assumption in multigenic trait determination is that all the loci affect the trait equally and therefore additively. However, detailed analysis has shown that in many (perhaps most) cases, a few genes have a major effect and many genes have a minor effect. For example, in the well-studied case of variation in sternopleural bristle number in *D. melanogaster*, approximately 10 loci account for 75% of the genetic variation in number. The situation is further complicated by pleiotropy: A gene may have a major effect on one character but minor effects on others. For example, phenylketonuria is an inborn error of metabolism producing severe mental retardation in humans and is produced by the nonfunctioning of phenylalanine hydroxylase, which is controlled as a recessive trait by a single gene on chromosome 12. The same enzyme is involved in melanin synthesis, and phenylketonurics have slightly paler hair and complexion than their normal sibs. The gene can therefore be regarded as having a major effect on intelligence but a minor effect on pigmentation.

However, a quantitatively inherited trait will be more likely than a qualitative one to be affected by environment. A group of individuals having identical genes for growth (i.e., a pure line or a clone) may show considerable variation in size due to differences in available nutrients. Although the same potential for size is present in the initial gene products, the manifestation of the phenotype will be limited by such factors as food availability.

Conversely, a population of genetically heterogeneous individuals may grow to the same size if no gene–environment interaction is limiting (or if different interactions compensate for each other). Generalizing, we can express the phenotypic value P for individual i in environment j as

$$P_{ij} = G_i + E_j$$

where G_i is the genetic contribution of the jth genotype and E_j is the environmental deviation resulting from the jth environment.

A particular genotype may do well in a particular environment, implying a specific interaction between the two. In this case,

$$P_{ij} = G_i + E_j + GE_{ij}$$

In practice, there will be variance of these components so that

$$V_\mathrm{p} = V_\mathrm{G} + V_E + 2\mathrm{Cov}_{GE}$$

where V_p, V_G, V_E, and $2\mathrm{Cov}_{GE}$, are the phenotypic, genetic, and environmental variance and the genotype–environment covariance, respectively. The genotype–environment covariance is positive when genotypes with higher values are in better environments and poorer genotypes have poorer environments. This may occur in animals when one member of a litter is large because of its genes and gets more food from its parents or when an animal is socially dominant for genetic reasons and therefore has more resources in food and spaces. A plant genotype which grows faster may have a better environment because it is less likely to be shaded.

In controlled plant and animal breeding, efforts are made to randomize genotypes and environments, so that Cov_{GE} is minimized. In this situation, Cov_{GE} can be neglected, and

$$V_\mathrm{p} = V_\mathrm{G} + V_E$$

Conventionally,

$$V_\mathrm{G}/V_P + V_E/V_P = h^2 + e^2$$

where h^2 and e^2 are the proportion of phenotypic variation due to genetic and environmental factors, respectively. The term h^2 is known as heritability in the broad sense,

$$h^2_\mathrm{B} = V_\mathrm{G}/V_P$$

In practice the genetic variance is composed of a range of different interactions between loci, which may be additive, dominant, or epistatic. Variability in the narrow sense is defined as

$$h^2_\mathrm{N} = V_A/V_P$$

where V_A is the variance due to additive genetic factors. It is an important statistic in determining the rate and amount of response to directional selection in breeding programs.

Heritability is a population-specific measurement. It does not measure an invariant property of a particular trait but only the relative contributions of genetic and environmental differences to phenotypic variation in a specific situation. If either genetic or environmental variation changes, heritability estimates will also change; heritability measures the proportion of phenotypic variation in a particular population due to genetic variation.

Genetic and Phenotypic Variability

The assumption of the early geneticists that most populations carry little variation and that most individuals are homozygous at all but a few loci at which alleles are either recent mutants in the process of being eliminated by selection or maintained by opposing selection pressures or heterosis is clearly wrong. Evolutionary theory (requiring modifying genes to change phenotype expression) and artificial selection practice (revealing considerable potential for inherited response to selection) imply the existence of considerable genetic variation in and between populations. In the 1950s and 1960s assumptions about genetic load suggested there was a maximum amount of variation which could be tolerated in any population, but the application of protein and later DNA electrophoresis showed that load theory was too narrow and deterministic and led to reinterpretation so as to incorporate ecological factors (including heterogeneity in time and space and variable stress from biotic and abiotic agents). Empirical data have shown that even inbred organisms (such as obligate self-fertilizers) or ones living in an apparently constant environment (such as the deep sea) may still be heterozygous at a significant proportion of their loci (up to one-fourth for enzyme loci in plants and invertebrates and less for vertebrates).

The implications of load theory led to debate among population geneticists and evolutionists about the significance of the observed variation. Theoreticians impressed by the apparent rigor of load theory and biochemists not used to meaningful variation in their study of chemical pathways tended to dismiss the bulk of genetic variation as neutral and irrelevant to the organism, whereas evolutionary ecologists and practical breeders regarded it as potentially adaptive and as expected by Darwinian understanding. The latter emphasis has been shown to be correct: Far from being constrained by invariant genotypes, phenotypes can (and should) be treated as capable of rapid response to the environment and hence permissive rather than determinative of survival. A phenotype can be interpreted as the consequence of developmental reaction norms, facilitated in most cases by the width of its underlying genetic base.

This concept is supported by a series of experiments carried out by Bruce Wallace to test if radiation-induced mutations in *D. melanogaster* were inevitably detrimental. He found that mutations induced in flies made homozygous by artificial breeding were often heterotic, although mutations occurring in flies with a heterozygous background were usually deleterious to their carriers, presumably since their gene expression had been adjusted by normal selection. This led Wallace to emphasize the importance of balancing (or stabilizing) selection rather than the directed or cleansing selection which removes unwanted variants.

The concept of reaction norm incorporates the idea of phenotypic plasticity championed by Bradshaw and Levins in the 1960s and provides a much needed synthesis of allometry and ontogeny with ecological realism and fitness components. A particularly good example is the seasonal polyphenism of the African satyrine butterfly, *Bicyclus anyana*, studied by Paul Brakefield and colleagues. The wet and dry season forms of this species are phenotypically distinct in the size of the

eyespots and the banding patterns on the wings. Seasonal shifts in rainfall are associated with changes in temperature and with butterfly behavior. In the cooler dry season, the insects rest on dry grass or leaf litter and rarely fly; they do not breed at this time. In contrast, the adults fly actively in the warmer wet season, searching for mates and oviposition sites in the lush vegetation. Survival differences between the two forms suggest that cryptic matching is more important in the dry season and deception mediated by the false eyes in the wet season.

The wet season form arises through an acceleration of development which can result from an increased or fluctuating temperature, food quality, or hormonal influence. However, eyespot size and plasticity are affected by genetic variation as shown by variation between families in different temperature regimes and by directed selection. The switch between different wing spotting phenotypes in the butterfly *Maniola jurtina* studied over many years by E. B. Ford is probably also of this nature. Ford's extensive field studies showed the interaction between founder population differentiation, natural selection in particular localities, and environmental determinants, although he did not interpret them in this way. Unlike the *Bicyclus* case, the changes in phenotype in *Maniola* are small and unlikely to be important in survival or fitness (although it would be wrong to be dogmatic about this conclusion). However, they serve as a marker of a variable developmental system which may well have a range of functional responses. There are many such examples of trivial (or even wholly cryptic) differences between individuals which have different functional properties (e.g., the possession of B chromosomes or resistance to a newly introduced pesticide or pathogen).

The complicated response system in *Bicyclus* (and other butterflies) is similar in principle to the much simpler situation of eye pigmentation in the amphipod *Gammarus chevreuxi* studied classically by Julian Huxley and Ford. Here, the unpigmented, red-eye form is produced by a single mutant allele which has a slower rate of melanin synthesis than normal. However, pigmented eyes like the wild type arise if genetically red-eyed animals are raised at higher temperatures than normal or if they also carry a gene for small eyes which enables the small amount of pigment produced to cover the whole eye. Another example is the "himalayan" mutation of mammals frequently described in elementary genetics textbooks in which an unpigmented (white) coat becomes pigmented at the cold extremities (feet and ears), as in Siamese cats. Pigmentation may also develop in hair which grows in a shaved area which has a lower temperature than usual. The tyrosinase in this genotype is heat labile and has a maximum activity below normal body temperature. Rate genes of this nature have been very important in evolution and can be shown to produce major alterations in body form through various forms of allometry.

Phenotypes and Evolution

It is a truism that natural selection can only operate where phenotypic variation exists. Evolution results from changes in gene (strictly allele) frequencies, but these are the consequence rather than the cause of phenotypic differences. It follows that selection on local forms (or ecotypes) which are solely the environmental consequence of a particular environment (e.g., plants grown in sheltered or exposed conditions or subject to a particular dietary lack or a behavioral reaction in an animal) will not lead to adaptive adjustment unless the phenotype in question depends on different genotypes. However, a phenotypic response to environmental conditions may allow a genetically nonadapted population to survive long enough to accumulate variants (through mutation, recombination, or immigration) and then adapt genetically. This idea was put forward independently by Baldwin (1896), Osborn (1897), and Lloyd Morgan (1900): It is commonly known as the Principle of Organic Selection or the Baldwin effect. This has been claimed by some to bridge the directed response which is the basis of Lamarckism and the random source of inherited variation which underlies Darwinism. Certainly, it may mimic Lamarckism, although it is of course wholly Darwinian in its operation.

Most of the early studies of phenotypic plasticity were carried out by botanists: Gaston Bonnier in the Alps and Pyrenees, F. E. Clements in Colorado and California, Turesson in Sweden, and Clausen, Keck, and Hiesey in California. However, the evolutionary implications were not pursued until Gause and Schmalhaussen in Russia and Waddington in Britain began to examine the relationships between genotype, phenotype, and environment in animal experiments. They were able to show not only that natural selection could lead to a character originally induced by the environment becoming an inherited character (Waddington called this "genetic assimilation") but also that there is a causal connection between the environmentally induced change and subsequent genetic changes. They argued that since adaptability—the ability to acquire an adaptive variant during an animal's lifetime—has a genetic basis, the genes underlying flexible adaptive variations may ultimately be responsible for the evolution of fixed adaptations to a new environment, i.e., the environment is much more than a sieve selecting (or eliminating) chance mutations. These ideas have been criticized (notably by G. C. Williams) on two grounds: That most changes resulting from environmental challenges are not adaptive and that fixation of a genetic response results in a decrease in genetic potential because less (genetic) information is needed to specify a fixed, than a variable response. There is truth in both these objections, although the assumption that plasticity requires more inherited variation than a fixed state is debatable. However, there is no doubt that the link between stress-generated responses and subsequent adaptation requires more study. The problem historically is that ecologists have been primarily concerned with pattern (i.e., spatial relationships within communities), whereas evolutionists have concentrated on processes (i.e., temporal changes within communities), and both have used restrictive definitions of stress. This division between disciplines is not, of course, absolute, but it has led to a rift within population biology and a lack of understanding of coevolutionary possibilities and constraints.

The divergence between ecologists and evolutionists is seen in the models developed by each. In general, ecological models are self-contained because the dependent variables (numbers or density) and parameters (birth and death rates, dispersion differences, rates of predation, etc.) are all

measurable by ecological methods. Ecologists have been particularly concerned with identifying criteria for stability or fluctuation when species interact, for invadability, and for extinction when competition occurs. Such models are limited because the parameters are all manifestations of the phenotypic properties of individuals and hence the product of evolutionary processes. Although lip service is paid to the fact that phenotypes may change, in practice they are assumed to be effectively constant in ecological time. Evolutionary models originally concentrated on examining possible rates of evolution; then they became concerned more with the factors involved in the maintenance of genetic polymorphism. Their variables are the frequencies of alleles and genotypes; their parameters are relative fitnesses and rates of mutation, migration, and recombination. For a long time, fitnesses were regarded by the model builders as virtually constant, and it is only comparatively recently that the dependence of selection coefficients on density and frequency has been incorporated. This false assumption of constancy led to the problems associated with genetic load and to controversies over neutralism.

Genetical (evolutionary) models have become increasingly realistic and treat relative fitnesses as capable of varying in time and space as well as with gene frequency and population density. Such models are both genetical and ecological, and therefore they are intrinsically more informative than purely ecological ones. However, they still suffer from a major problem in that they tend to assume a degree of genetical equilibrium or stasis which is unjustified.

The way forward will be to increase ecological reality of evolutionary models, taking into account the characteristics of the niche for particular populations and communities, particularly any genetic constraints. Such models obviously need to include behavioral (sociobiological) input and the notion of evolutionary "strategies."

Taxonomy

Classification has become increasingly sophisticated with the increase in traits which can be used to characterize a group (or taxon). The problem is that conventional taxonomic (phenotypic) diversity may bear little relationship to genetic diversity. For example, the seaside sparrow (*Ammodramus nigrescens*) is common on the eastern and southern coasts of the United States. Nine subspecies have been described, including a rare and recently extinct (1987) dusky form, originally regarded as a separate species. Molecular analysis (based on mitochondrial DNA sequences) showed that the dusky sparrow was indistinguishable from other Atlantic forms, but that there was a major and previously unsuspected distinction between Atlantic and Gulf Coast groups.

In contrast, New Zealand tuatara lizards (*Sphenodon punctatus*) are conventionally treated as a single taxon, but molecular (and morphological) criteria indicate they comprise at least three species. For conservation management, it is obviously important that each taxon be considered separately.

There is no clear correlation between genetic and taxonomic diversity. The problems of classification that have always challenged museum workers have been compounded by the possibility of using genetic factors as additional or substitute taxonomic criteria: Closely related species in the genetic sense may have very different niches than those of their near relatives, whereas genetically more distant forms may look alike and interact strongly. There are no clear rules to link phenotype with genotype or phenotypic variety with genetic variety. There is plenty of scope for better multidisciplinary understanding of phenotypes, this will have to involve genetics, development, behavior, biotic and abiotic environments, life history, and phylogeny. Notwithstanding case by case examinations of the determinants and plasticity of the phenotypes in particular species and species groups are still needed.

See also: Evolution, Theory of. Genes, Description of. Genetic Diversity

References

Avise JC (1994) *Molecular Markers, Natural History and Evolution.* New York: Chapman & Hall.
Berry RJ, Crawford TJ, and Hewitt GM (eds.) (1992) *Genes in Ecology.* Oxford: Blackwell.
Briggs D and Walters SM (1984) *Plant Variation and Evolution*, 2nd ed. Cambridge, UK: Cambridge Univ. Press.
Falconer DS (1989) *Introduction to Quantitative Genetics*, 3rd ed. New York: Longman.
Feder ME, Bennett AF, Burggren WW, and Huey RB (eds.) (1987) *New Directions in Ecological Physiology.* Cambridge, UK: Cambridge Univ. Press.
Gould SJ (1977) *Ontogeny and Phylogeny.* Cambridge, MA: Harvard Univ. Press.
Hamilton WD (1996) *Narrow Roads of Gene Land.* New York: Freeman.
Hoffman AA and Parsons PA (1997) *Extreme Environmental Change and Evolution.* Cambridge, UK: Cambridge Univ. Press.
Jablonka E and Lamb MJ (1995) *Epigenetic Inheritance and Evolution.* Oxford: Oxford Univ. Press.
Mather K (1973) *Genetical Structure of Populations.* London: Chapman & Hall.
Mayr E (1982) *The Growth of Biological Thought.* Cambridge, MA: Harvard Univ. Press.
Raff RA (1996) *The Shape of Life.* Chicago: Univ. of Chicago Press.
Schlichting CD and Pigliucci M (1998) *Phenotypic Evolution.* Sunderland, MA: Sinauer.
Waddington CH (1957) *The Strategy of the Genes.* London: Allen & Unwin.
Yablokov AV (1986) *Phenetics.* New York: Columbia Univ. Press.

Photosynthesis, Mechanisms of

John A Raven, University of Dundee, Dundee, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 549–558, © 2001, Elsevier Inc.

Glossary

ATP synthetase Membrane-associated protein complex that couples exergonic fluxes of H^+ across the membrane to ADP phosphorylation, or vice versa.

Carboxylase Enzyme that catalyzes the formation of a C—C bond between inorganic C (CO_2 or HCO_3^-) and some organic molecule.

Chemiosmotic energy coupling Exergonic (photo)-chemical redox reaction, or an exergonic hydrolysis of a phosphate anhydride or thioester, coupled to endergonic ion (normally H^+) transfer across a membrance, or vice versa.

Chromophore Organic molecule that absorbs photosynthetically active radiation and either carries out photochemistry or transfers excitation energy to another chromophore that carries out photochemistry. Always bound to a polypeptide in the photosynthetic apparatus.

Reaction center (Bacterio)chlorophyll–protein complex that performs the primary photochemical reactions of photosynthesis, the earliest products being the oxidized reaction center (bacterio) chlorophyll at the outside/thylakoid lumen side and a reduced (bacterio) chlorophyll (RC1) or reduced (bacterio) phaeophytin (RC2) on the cytosol/stroma side of the photosynthetic membrane.

Photosynthesis, or building with light, involves the use of electromagnetic energy to bring about endergonic reactions. The concept is generally extended to the use of the more immediate photochemical products to reduce CO_2 and (in most photosynthetic organisms) oxidize H_2O and evolve O_2, this extended definition is used here. This article deals with the partial processes of photosynthesis in the order in which they occur subsequent to initial photon absorption. Starting with the characteristics of the photon-harvesting apparatus, we then consider the photochemical reaction centers and their immediate products, which are an oxidant and a reductant oxidation–reduction and/or an ion (H^+ or, rarely, Cl^-) gradient. These light-harvesting and photochemical reactions are invariably associated with membranes. Next in temporal sequences are the reactions that lead to the phosphorylation of ADP and generation of strong (NAD(P)H) reductant; these reactions are also membrane related. Finally, we deal with the use of ATP and NADPH in CO_2 fixation and other endergonic reactions; these reactions are either in aqueous solution or associated with membranes other than those in which primary photochemistry occurs. In O_2-evolving organisms, the basic photosynthetic equation is (1)

$$C^{\dagger}O_2 + 2H_2^{*}O + 8\,\text{photons} \rightarrow (CH_2{}^{\dagger}O) + {}^{*}O_2 + H_2{}^{\dagger}O \qquad (1)$$

Light Harvesting and Excitation Energy Transfer to Reaction Centers

For photochemical reactions to occur, photons must be absorbed by a pigment. In the majority of photosynthetic organisms, only one in 100 to one in 1000 pigment molecules can actually bring about useful (to the organisms) photochemical reactions. The photochemically inactive antenna pigments serve to absorb radiation and transfer the energy as excitation energy (i.e., not in stable chemical forms) to photochemically active pigments in reaction centers. This division of labor presumably relates to the generally rather "dilute" nature of solar radiation, such that each pigment molecule absorbs a photon on the order of a few times each second, yet the photochemical machinery and the immediate downstream reactions can react hundreds or thousands of times per second. Accordingly, the resource-expensive (energy, nitrogen, iron) reaction centers and the related catalytic machinery that is energetically and temporally downstream of the reaction centers are present as a small number of copies of the catalysts relative to the number of molecules of light-harvesting pigment; this arrangement is presumably related to optimal allocation of resources in natural selection.

A very small number of species, all of them members of the Archaea, do not accord with this generalization about antenna pigments and reaction centers. These Halobacteria use carotene pigments associated with membrane-spanning proteins to catalyze their primary photochemical reaction, with each protein molecule associated with its pigment molecule acting as a light-energized H^+ (bacteriorhodopsin) or Cl^- (halorhodopsin) pump (*see* Reaction Centers and Primary Photochemistry).

By contrast, all other photosynthetic organisms are (eu)Bacteria and their symbiotic eukaryotic descendants that use the antenna plus reaction center system. In all of these organisms the pigment involved in the primary photochemical reactions is a Mg- (very rarely Zn-) binding cyclic tetrapyrol, chlorophyll (in O_2-evolvers) and bacteriochlorophyll (in non-O_2-evolvers). Chlorophylls always occur in a functional photosynthetic apparatus in association with proteins in, or more rarely on, bilayer membranes. In all chlorophyll-containing organisms there are some chlorophyll molecules that act as antenna pigments. In many such organisms (e.g., all green algae and higher plants) the great

Table 1 Chemistry, taxonomy, and location of light-harvesting pigment–protein complexes

Chromophores	*Taxonomy*	*Location*
O_2-Evolvers		
Chlorophylls *a* and *b*	Chlorophyta, Euglenophyta, Chlorarachniophyta, Embryophyta; "chloroxybacterial" cyanobacteria	Integral membrane protein of thylakoid
Chlorophylls *a* + (usually) *c*	Cryptophyta, Heterokontophyta (Bacillariophyceae, Chrysophyceae, Phaeophyceae, Tribophyceae), Haptophyta, Dinophyta	Integral membrane protein of thylakoid
Chlorophyll *a*, peridinin	Dinophyta	Peripheral protein on stromal surface of thylakoid
Phycobilinogen	Cyanobacteria *sensu stricto*, Rhodophyta	Phycobilisomes on stromal/cytosol surface of thylakoid
	Cryptophyta	Thylakoid lumen
Non-O_2-Evolvers		
Bacteriochlorophylls *a* and *b*	Proteobacteria: purple sulfur bacteria (Chromatiaceae), purple nonsulfur bacteria (Rhodospirillaceae)	Integral membrane protein of "chromatophores"
Bacteriochlorophyll *g*	Heliobacteriaceae	Integral membrane protein of plasmalemma
Bacteriochlorophylls *c*, *d*, and *e*	Chlorobiaceae, Chloroflexaceae	Chlorosomes on the cytosol side of plasmalemma

majority of light harvesting is carried out by chlorophyll–protein complexes, in this case with chlorophylls *a* and *b* as the major light-absorbers with minor contributions from carotenoids. In the case of algae with chlorophyll *a* and, usually, chlorophyll *c*, carotenoids generally have a more significant role in light harvesting by pigment–protein complexes. In a final group of organisms (mainly cyanobacteria and red algae), the light-harvesting pigments in addition to chlorophylls are open-chain tetrapyrol pigments covalently bound to proteins, attached to the outer (stromal in red algae, cytosolic in cyanobacteria) side of the photosynthetic membranes, or in the thylakoid lumen of cryptophyte algae. The variations in the chemistry and location of light-harvesting pigment–protein complexes are indicated in **Table 1**.

A number of energetic and structural constraints determine the spatial relationship of the different pigments in light-harvesting complexes. The energetic constraint is that energetically efficient excitation energy transfer among antenna pigments, and from antenna pigments to reaction centers, involves an unavoidable energy loss during the transfer. Thus, the peripheral pigments in the antenna have a higher energy content of their excited states than do the inner antenna pigments, and those in turn have a more energetic excited state than the reaction center pigment. The situation is complicated by the occurrence of two readily accessible excited states in some photosynthetic pigments, and especially in the chlorophylls. Here the second excited state, produced by absorption of blue radiation, has a higher energy content than the first excited state, corresponding to the *in vivo* absorption in the red (chlorophylls) or infrared (bacteriochlorophylls) region of the electromagnetic spectrum. It is the first excited state that performs the primary photochemistry in (bacterio) chlorophyll-based photosynthesis. Thus, photons that excite (bacterio) chlorophylls to the second excited state are not directly used in photochemistry via this state, but only after they become converted to the first excited state; this also applies to excitation energy transfer among pigments.

These energetic requirements are reflected in the structure of the light-harvesting apparatus. Thus, the pigments are arranged spatially such that those absorbing shorter-wavelength (thus higher energy) photons are generally at a greater distance from the reaction center, whereas those absorbing at longer wavelengths (more similar to that of the reaction center pigments) are close to the reaction center. Furthermore, the smaller-scale arrangement of individual chromophores in the pigment–protein complexes is also a function of the distance between chromophores over which effective excitation energy transfer can occur; these arrangements have been shown for a number of light-harvesting pigment–protein complexes using X-ray crystallographic techniques.

A related consideration is that of how many chromophore molecules can be involved in light harvesting for a given reaction center. A *minimum* estimate of the upper limit of this number comes from determining the ratio of pigment molecules serving a given kind of reaction center and dividing it by the number of the type of reaction center that they serve. This is clearly most readily determined when there is only one kind of reaction center in the organisms concerned, as is the case in organisms that do not evolve O_2, but it is also applicable to O_2-evolvers, which invariably have two kinds of reaction center (*see* Reaction Centers and Primary Photochemistry). Such computations suggest that up to 1000–2000 chromophores occur per reaction center of a given type. Even more pigment molecules can be involved in excitation energy transfer to a given reaction center if pigment molecules are not all dedicated to a given reaction center but can be involved in excitation energy transfer to several nearby centers.

Another size-related aspect of light harvesting occurs at all scales, but becomes most obvious at a larger scale: this concerns the optical thickness of the cell or tissue in which the photosynthetic apparatus occurs. A small optical thickness involves a smaller physical thickness for a given pigment concentration per unit volume, or a lower concentration of pigment per unit volume if there is a fixed physical thickness, or a combination of the two. A large optical thickness means more self-shading of pigment molecules, so that the average specific absorption coefficient of each pigment molecule is

lower. Furthermore, in a given radiation environment, it takes longer for a pigment molecule to absorb enough photons to cover the energy cost of its synthesis in a cell or organ with more self-shading. Finally, the cells or tissues with a small optical thickness, and hence smaller package effect, have a greater opportunity for the spectral diversity of pigments that they contain to be manifest in the specific wavelengths of radiation that are absorbed. The ultimate in large optical thickness is found in the thalli of some seaweeds that, with more than one millimole of pigment per square meter, look black to the human observer (because of low reflectance or transmittance, with minimal wavelength dependence of these two processes due to high absorptance at all visible wavelengths) regardless of whether their pigments are of the kind found in green, brown, or red algae.

A final aspect of light harvesting concerns the occurrence of pigment molecules in photosynthetic cells that (of necessity) absorb radiation but that show little or no transfer of excitation energy to reaction centers. Examples of such pigments are β-carotene, and at least some portion of the other carotenoids. The pigments serve as photoprotectants, quenching the triplet excitation states of chlorophyll and singlet oxygen in the case of β-carotene, and as nonphotochemical quenchers of excess (singlet) excitation energy of antenna pigments in the case of phototransformable carotenoids such as violoxanthin–antheroxanthin–zeaxanthin and diadinoxanthin–diatoxanthin in certain O_2-evolvers.

Reaction Centers and Primary Photochemistry

We have seen (*in* Light Harvesting and Excitation Energy Transfer to Reaction Centers) that the halobacteria, with their halorhodopsin or bacteriorhodopsin pigments, do not have separate antenna and reaction center pigments. These organisms have active transport of ions (H^+ out of the cell or Cl^- into it) as the sole product of photochemistry external to the pigment–protein complex. The other photosynthetic organisms, with (bacterio) chlorophyll as their photochemically active pigment, do have reaction centers; these organisms are predominant in terms of number of species, habitats occupied, and quantity of energy and materials transformed.

In all of the (bacterio)chlorophyll-based reaction centers the primary photochemical event is the production of a reduced compound on the side of the membrane abutting the cytosol (prokaryotes) or chloroplast stroma (eukaryotes) and an oxidized compound on the side of the membrane adjacent to the external medium (many prokaryotes) or thylakoid lumen (almost all O_2-evolvers). The photochemical redox process thus involves the transfer of an electron from near one side of the membrane to near the other side of the membrane. Furthermore, the photochemically active pigment is bacteriochlorophyll *a*, *b*, or *g*, or chlorophyll *a* (or apparently, in one case, chlorophyll *d*), bound to protein (**Table 2**).

While all (bacterio)chlorophyll-based reaction centers share this basic mechanism, such centers can be divided into two types termed RC1 and RC2 (**Table 2**). In the RC1 type (e.g., the reaction centers of the Chlorobiaceae and Heliobacteriaceae, and PSI of O_2-evolving organisms), the (bacterio)chlorophyll catalyst of primary photochemistry that is oxidized generates an oxidant (oxidized (bacterio)-chlorophyll) at a redox potential of $+0.25$–$+0.45$ V, while the primary reductant is another form of (bacterio)-chlorophyll with a redox potential of -0.6–-0.7 V. In the RC2 type (e.g., the reaction centers of the Chloroflexaceae and Proteobacteria, and PSII of O_2-evolvers), the oxidized (bacterio) chlorophyll product of primary photochemistry has a redox potential of $+0.45$ (bacteria that cannot evolve O_2) to $+0.1$V (O_2-evolving organisms), while the reductant (a (bacterio)pheophytin, i.e., a (bacterio)chlorophyll without Mg) has a redox potential of -0.8 V (bacteria that cannot evolve O_2) to -0.5V (O_2-evolvers).

The primary charge separation is relatively unstable, and reduction of a more stable compound with a higher redox potential, and which is less prone to back-react yielding chemiluminescence, occurs within $\sim$100 ps. For RC1 this secondary acceptor is a phylloquinone (vitamin E) followed, within $\sim$100 ns, by electron transfer to iron–sulfur centers. For RC2 the secondary acceptor is a ubi- (non-O_2-evolvers) or plasto- (O_2-evolvers) quinone bound to the reaction center; the resulting semiquinone free radical reduces a second, chemically identical quinone within $\sim$300 μs.

The reduction of the primary oxidized product of photosynthesis takes longer than does oxidation of the primary reduced product. For RC1, and for the RC2 of organisms that do not evolve O_2, the compound that reduces the primary oxidant is a c-type cytochrome or (in many O_2-evolvers) the cuproprotein plastocyanin; this reduction takes $\sim$200 μs. For RC2 in O_2-evolvers (i.e., PSII) the reductant of the primary oxidant is a tyrosine residue termed Y_Z in a polypeptide (Dl), which, with polypeptide D_2, binds the reaction center chlorophyll *a*. The action takes $\sim$200 ns, and the oxidized

Table 2 Reaction centers of various photosynthetic organisms that contain (bacterio)chlorophyll

Type of reaction center	*Chromophore involved in primary photochemistry*	*Taxonomy*
O_2-Evolvers		
RC1 (as PSI)	Chlorophyll *a* (P700)	All O_2-evolvers
RC2 (as PSII)	Chlorophyll *a* (P680)	All O_2-evolvers
Non-O_2-Evolvers		
RC1	Bacteriochlorophyll *g* (P798)	Heliobacteriaceae
RC1	Bacteriochlorophyll *a* (P840)	Chlorobiaceae
RC2	Bacteriochlorophyll *a* (P840)	Chloroflexaceae
RC2	Bacteriochlorophyll *a* or *b* (P870 *or* P890)	Chromatiaceae and Rhodospirillaceae

tyrosine is re-reduced by a protein-bound cluster of four Mn in ~200 μs. All of the RC1 reaction centers have, as known from molecular genetic data, a common ancestry, whereas RC2 reaction centers also have a common ancestry that differs from that of RC1. Further discussion of evolutionary points are beyond the scope of this article.

Membrane-Associated Reactions Leading from Primary Photochemistry to ATP and NAD(P)H

Dealing first with RC1, a role that these reaction centers can perform in most, if not all, of the organisms in which they occur is cyclic electron transport. This process involves ubi- (non-O_2-evolvers) or plasto- (O_2-evolvers) quinone and a cytochrome b–c_1 (non-O_2-evolvers) or cytochrome b_6–f (O_2-evolvers) complex. The link between the iron–sulfur reductant of RC1 and the quinone involves the soluble FeS protein ferredoxin (less commonly the Fe-free flavodoxin), and other redox catalysts that are incompletely characterized and that are probably variable among or even within taxa. One of these could be NAD(P)H dehydrogenase. The oxidizing end is much better understood, with the cytochrome b–c_1 (b_6–f) complex reducing a cytochrome c (or plastocyanin) and hence the primary oxidant of the photochemical reaction.

The bioenergetic role of cyclic electron transport is the pumping of H^+ from the cytosol (or stroma) to the external environment (or thylakoid lumen) compartments. Two H^+ per electron are pumped via the quinone and the cytochrome complex. If the NAD(P)H dehydrogenase is involved, then a further $2H^+$ per electron could be transferred. A major role for these pumped H^+ is to move back across the membrane through the ATP synthetase; $4H^+$ energize the phosphorylation of one ADP. If all the absorbed photons are used to energize RC1, each photon could produce 0.5–1.0 ATP, depending on the involvement of NAD(P)H dehydrogenase. An alternative use of the H^+ gradient is to energize secondary active transport coupled to exergonic H^+ fluxes.

Another role for RC1, and apparently the major one in almost every situation in which RC1 occurs, is to generate a reductant at a low enough redox potential to reduce CO_2 to carbohydrate. As will be seen (*in* Use of NAD(PH) and ATP in CO_2 Fixation and Other Reactions), in many cases the reductive step in CO_2 fixation involves an energetic input from ATP, and here the reductant is NAD(P)H. The path from the reduced protein-bound iron–sulfur involves ferredoxin and a membrane-bound ferredoxin–NAD(P)H reductase. The reduction of ferredoxin and NAD(P)H, and hence the reduction of CO_2 (or NO_2^-, or other electron acceptors), requires an input of (weaker) reductant at the oxidizing end of the photosystem. For members of the Chlorobiaceae and Heliobacteriaceae, the reductant is ultimately some exogenous organic or inorganic electron donor (other than H_2O) such as S^{2-}, which supplies electrons to the oxidizing end of the photosystem via cytochrome c, with or without the involvement of the cytochrome b–c_1 complex. For the O_2-evolvers, PSI (photoreaction I) is supplied with reductant by PSII (photoreaction II) via plastoquinone, the cytochrome b_6–f complex, and cytochrome c_6 or plastocyanin.

RC2 does not produce a stable reductant with a redox potential that is sufficiently negative to reduce ferredoxin or $NAD(P)H^+$. For the photosynthetic Proteobacteria and the Chloroflexaceae, a major function of the RC2 reaction center is in a cyclic electron transport pathway. This uses the ubiquinol produced at the reducing end of the reaction center to reduce the cytochrome b–c_1 complex, and then cytochrome c, and finally the oxidant produced by the reaction center is reduced. This mechanism pumps $2H^+$ per electron, and hence per photon, whose excitation energy is transferred to the reaction center. One use of the H^+ gradient is, as for cyclic electron flow and H^+ pumping by RC1, to generate ATP. Another is to use a weak reductant (e.g., S^{2-}) to generate a strong reductant (NADH). Here electrons are transported against an energy difference using the energy from the flux of H^+ from the medium back to the cytosol. This process is the reverse of the active H^+ transport caused by the respiratory, exergonic, electron transport from NADH to S^0 (generating S^{2-}). Similar uses of the H^+ gradient can be envisaged in the Halobacteriaceae.

The other, and quantitatively predominant, occurrence of RC2 in the biosphere is as PSII of O_2-evolvers. Here the usual immediate sink for the electrons from reduced plastoquinone (= plastoquinol) is the cytochrome b_6–f complex, with subsequent sequential reduction of cytochrome c_6 or plastocyanin, PSI, ferredoxin, and $NADP^+$, or, finally to a smaller extent, some other oxidant such as NO_2^- or O_2 (see the following). The corresponding electron donor to the oxidizing side of PSII, after the polypeptide-bound tyrosine Y_Z and polypeptide-bound Mn, is the very weak reductant H_2O. It appears that PSII does not engage in cyclic electron transport involving the cytochrome b_6–f complex in the manner of RC2 in the Proteobacteria and Chloroflexaceae, although it may perform cyclic electron transport via cytochrome b_{559}, a redox catalyst with no analog in the RC2 of non-O_2-evolvers. This cyclic electron transport probably occurs as a photochemical means of dissipating excess, and potentially damaging, excitation of PSII.

Putting PSII and PSI together, in terms of photon requirement, the transport of one electron from H_2O to $NADP^+$, and hence CO_2, and the associated H^+ transport, requires the energy of two photons, one used by PSII and the other by PSI. The stoichiometry of H^+ transport is that one H^+ is left in the thylakoid lumen, and one H^+ is taken up in the cytosol or stroma, per electron moved from H_2O to $NADP^+$ and then to CO_2. In addition, for each electron moving through plastoquinone and the cytochrome b_6–f complex, two H^+ are transferred from the cytosol or stroma to the thylakoid lumen. Thus, per electron moved from H_2O to $NADP^+$ (and thence CO_2), $3H^+$ are moved from the cytosol or stroma to the thylakoid lumen:

$$2NADP^+_{(s)} + 2H_2^*O_{(t)} + 12H^+_{(s)} + 8\,\text{photons} \rightarrow$$

$$2NADPH_{(s)} + 2H^+_{(s)} + 12H^+_{(t)} + {}^*O_2 \qquad (2)$$

where the subscript (s) means stroma (or cytosol of cyanobacteria) and the subscript (t) means thylakoid lumen.

With an H^+/ATP of 4 in ATP synthetase, the ADP phosphorylation corresponding to the recycling of $12H^+$ can be

described by

$$12H^{+}_{(s)} + 3ADP_{(s)} + 3Pi_{(s)} \rightarrow$$
$$12H^{+}_{(t)} + 3ATP_{(s)} + 3H_2O \quad (3)$$

Thus, each electron moved from H_2O to $NADP^+$ generates 0.75 ATP, so that 3ADP are phosphorylated per $2NADP^+$ (and hence $1CO_2$) reduced. This can be seen when Eqs. (2) and (3) are added together to yield

$$2NADP^{+}_{(s)} + 2H_2{}^{*}O_{(t)} + 3Pi_{(s)} + 8photons \rightarrow$$
$$2NADPH_{(s)} + 2H^{+}_{(s)} + 3ATP_{(s)} + 3H_2O_{(s)} + {}^{*}O_{2(t)} \quad (4)$$

This has important implications for the minimum photon cost of CO_2 fixation in O_2-evolvers, which is 8 photons absorbed per CO_2 reduced to carbohydrate.

In addition to the reduction of CO_2 and NO_2^- by PSI, there is also the reduction of O_2 to form, ultimately, H_2O. The Mehler Peroxidase reaction involves two electrons from the reducing end of PSI passing to two O_2 to form two superoxide anions, which are then dismuted by the enzyme superoxide dismutase to form one H_2O_2 and one O_2. The H_2O_2, plus ascorbate, are then acted on by the enzyme ascorbate peroxidase, which converts one H_2O_2 and one ascorbate to two H_2O and one dehydroascorbate, involving (via other enzymes) two more electrons from PSI to regenerate the ascorbate co-substrate from dehydroascorbate. Overall, the reaction involves the evolution of one O_2 at the oxidizing end of PSII and the uptake of one O_2 at the reducing end of PSI, so that there is no net evolution or uptake of O_2. The reaction can be quantified by the use of $^{18}O_2$ tracer. The Mehler Peroxidase reaction is a means of disposing of active, and potentially damaging, oxygen species produced in thylakoid redox reactions and, to the extent that superoxide production by PSI can be increased above the unavoidable basal rate, a means of photochemical energy dissipation and of ATP generation not paralleled by $NADP^+$ reduction as an alternative to cyclic photophosphorylation.

The photon cost of CO_2 fixation in terms of incident irradiance depends on the photon cost of CO_2 fixation in terms of absorbed photons (8 in the case mentioned here) and the fraction of the incident photons that are absorbed (*see* Light Harvesting and Excitation Energy Transfer to Reaction Centers). The catalytic capacity of the reactions generating NAD(P)H and ATP can, under many circumstances, constrain the rate of CO_2 fixation when photon supply is not limiting. However, in some cases it is the supply of CO_2, or the capacity of CO_2-assimilation reactions (*see* Use of NAD(P)H and ATP in CO_2 Fixation and Other Reactions), that constrains the CO_2 fixation rate at saturating light supply. The capacity for light harvesting relative to that for primary photochemical reactions and downstream thylakoid reactions and CO_2 fixation reactions (*see* Use of NAD(P)H and ATP in CO_2 Fixation and Other Reactions) depends on the irradiance that the plant has encountered in its evolutionary history (genetic adaptation) and during the life of an individual (phenotypic acclimation). In both adaptation and acclimation, the response to low light is an increased capacity for light harvesting relative to downstream reactions, and vice versa for the response to high light.

In O_2-evolvers the rate and (incident) photon cost of CO_2 fixation at limiting incident irradiances, and the rate of CO_2 fixation at saturating irradiances, can be reduced by photoinhibition. This phenomenon has its basis in damage to the PSII reaction center by one in 10^6 to 10^7 of the photons whose excitation energy reaches the reaction center. In the absence of any protective mechanisms, this photodamage would occur more rapidly at higher incident photon flux densities, but would result in a greater fractional inhibition of photosynthesis when measurements on organisms exposed to high incident photon flux densities are subsequently made at lower (rate-limiting) irradiances. The observed phenomenon of photoinhibition, that is, a reduced rate of photosynthetic CO_2 fixation during, and following, exposure to (especially high) irradiances, is not solely a function of photodamage occurring faster than it can be repaired. Thus, the decreased photosynthetic rate can be a result of diversion of the products of PSII photochemistry to processes other than CO_2 fixation at high incident irradiances, for example, to cyclic electron flow round PSII or to the Mehler Peroxidase reaction, decreasing the photochemical unusable "excess" excitation of PSII. More commonly, excitation energy is prevented from reaching the PSII reaction center by dissociation of some of the light-harvesting pigment–protein complexes (**Table 2**) from the reaction center, or by nonphotochemical quenching of excitation in the light-harvesting complexes by processes triggered by very low pH in the thylakoid lumen. This increased acidification of the lumen is a result of photochemically driven H^+ pumping in excess of the rate at which H^+ is recycled through the ATP synthetase due to restricted ATP consumption and hence limited ADP availability, and can dissipate excitation energy via a xanthophyll cycle mechanism in all but cyanobacteria and red algae and, in all O_2-evolvers, via mechanisms not related to a xanthophyll cycle.

Use of NAD(P)H and ATP in CO_2 Fixation and Other Reactions

Most of the ATP and NAD(P)H generated in the thylakoid reactions in O_2-evolvers is used in CO_2 reduction to (CH_2O). Other uses for the ATP resulting from photochemical reactions include photoassimilation of exogenous organic compounds, as well as other transport and biochemical reactions. Photochemically produced reductant can be used for reductive biosyntheses, such as reduction of NO_3^-, NO_2^- and SO_4^{2-}. These uses of ATP and reductant notwithstanding, the major global biogeochemical role of NAD(P)H and ATP generated in photochemical reactions is the conversion of CO_2 to carbohydrates; a net primary production of $\sim$100 Pg carbon per year worldwide is carried out by O_2-evolvers. However, some CO_2 is fixed by non-O_2-evolving autotrophs. In the Chlorobiaceae, the reverse tricarboxylic acid cycle is used, whereas the Chloroflexaceae use the hydroxypropionate cycle (**Table 3**). These two processes each involve more than one carboxylation reaction in converting CO_2 to carbohydrates. By contrast, the photosynthetic carbon reduction cycle used in the Proteobacteriaceae and in all O_2-evolvers involves but a single carboxylase, that is, ribulose-1,5-bisphosphate carboxylase–oxygenase (RUBISCO), in converting CO_2 to

Table 3 Taxonomic distribution and energy costs of CO_2 assimilation reactions in photosynthetic organisms

CO_2 assimilation reaction	*Taxonomy*	*NAD(P)H cost per net CO_2 fixed*	*ATP cost per net CO_2 fixed*
Reverse tricarboxylic acid cycle	Chlorobiaceae	2	?
Hydroxypropionate pathway	Chloroflexaceae	2	?
RUBISCO and photosynthetic carbon reduction pathway			
(i) In high CO_2–low O_2			
(ii) In low CO_2–high O_2 (present atmosphere or water in equilibrium with it)	Proteobacteriaceae (Chroma tiaceae Rhodospirillaceae)	2	3
(a) Diffusive CO_2 entry to RUBISCO, with photorespiratory carbon oxidation cycle	C_3 land plants; some algae	2.8	4.2
(b) Active transport of inorganic C to RUBISCO by a process that does not involve C—C bond formation	Cyanobacteria; many algae	2	4–5
(c) Active transport of inorganic C to RUBISCO by a process that does involve C—C bond Formation	C_4 higher plants	2	5
	CAM higher plants	2	6.5

carbohydrates, although some essential biosyntheses involve parallel, quantitatively minor, carboxylation reactions catalyzed by other carboxylases.

RUBISCO, as its name suggests, not only catalyzes a carboxylase reaction but also has an oxygenase activity. The oxygenase activity is competitive with the carboxylase activity. Thus, the oxygenase is not a significant reaction in proteobacterial photosynthesis in their typical high-CO_2, low-O_2 habitats where heterotrophic organisms dominate and organic carbon sedimentation into the habitat exceeds O_2 diffusion into it. Furthermore, the oxygenase activity would not have been significant in the high-CO_2, low-O_2 world prior to 2 billion years ago, that is, for the first 1.5 billion years or so of the occurrence of RUBISCO. When oxygenase activity is negligible, the conversion of $1CO_2$ to the carbohydrate redox level by the photosynthetic carbon reduction cycle needs 2NAD(P)H and 3ATP:

$$2\text{NADPH} + 2\text{H}^+ + 3\text{ATP} + 3\text{H}_2\text{O} + \text{C}^{\dagger}\text{O}_2 \rightarrow$$
$$(\text{CH}_2{}^{\dagger}\text{O}) + 2\text{NADPH}^+ + 3\text{ADP} + 3\text{Pi} + \text{H}_2{}^{\dagger}\text{O} \quad (5)$$

Equation (5), showing the consumption of NADPH and ATP in CO_2 fixation, can be combined with Eq. (4), describing the production of NADPH and ATP by light-driven processes in the thylakoid, to give the overall equation for photosynthesis (Eq. (1)).

While the reductant requirement per CO_2 assimilated is the same as for the reverse tricarboxylic acid cycle and the hydroxypropionate cycle, the ATP requirement is greater for the photosynthetic carbon reduction cycle (see **Table 3**).

The greater energy requirement for the photosynthetic carbon reduction cycle than for the other carbon fixation cycles is exacerbated in the present atmosphere, or in natural waters in equilibrium with it. If CO_2 enters by diffusion (C_3 physiology) then there is significant oxygenase activity of RUBISCO, although there is significant variability among RUBISCOs of different organisms in selectivity for CO_2 relative to O_2. The oxygenase activity produces phosphoglycolate, and thence glycolate. Recycling this glycolate to carbohydrate has a significant energy cost; the excretion of glycolate in aquatic organisms imposes an even greater energy cost for net CO_2 fixation. The energy cost of the net fixation of $1CO_2$ from the present atmosphere by a vascular land plant with C_3 physiology (CO_2 reaching RUBISCO by diffusion) is some 2.8 NADPH and 4.2 ATP. This contrasts with the 2NADPH and 3ATP required when CO_2 is saturating and, regardless of O_2 concentration, there is no oxygenase activity (see **Table 3**).

C_3 physiology occurs in the great majority of species of terrestrial embryophytes, a few lichens and asymbiotic terrestrial algae, and some aquatic algae and embryophytes. Some red algal RUBISCOs have the highest reported selectivity for CO_2 over O_2, so that glycolate production may be minimal in red algae with C_3 physiology. The algae, lichens, and bryophyte gametophytes with C_3 physiology have limitations on photosynthesis by CO_2 diffusion that vary with the structure of the photosynthetic apparatus, whether the medium supplying inorganic carbon ($CO_2 + HCO_3^-$) is air or water, the CO_2 concentration in the medium, and the thickness of the diffusion boundary around the photosynthetic structure. In sporophytes of embryophytes in water the situation is as for gametophytes, whereas for sporophytes in air there is the additional complication of stomata. These variable conductance apertures connecting intercellular gas spaces to the atmosphere are crucial in maintaining land plants in a hydrated state when soil water availability is low and/or the atmosphere is very dry, and permit CO_2 fixation in the light when the plants can remain hydrated with open stomata. In most cases the C_3 vascular plant sporophytes on land fix CO_2 under conditions in which the rate of fixation is about one-third constrained by CO_2 diffusion and two-thirds constrained by biochemical reactions. In other words, were CO_2 diffusion constraints to be reduced by 10%, the rate of photosynthesis would increase by 3%; if the constraints from biochemistry were reduced by 10%, the rate of CO_2 fixation would increase by 7%. The fractional limitation of photosynthesis by CO_2 diffusion in aquatic plants with C_3 physiology is more variable than in the terrestrial vascular plant sporophytes with this photosynthetic physiology.

In a number of species of O_2-evolvers the supply of CO_2 from the environment to RUBISCO involves some energized

process that gives a higher CO_2 concentration, and CO_2/O_2 ratio, near RUBISCO than that in the medium. Such mechanisms suppress oxygenase activity, in some cases almost completely, and frequently increase the affinity of *in vitro* photosynthesis for CO_2 relative to the situation with C_3 physiology (diffusive supply of CO_2 to RUBISCO).

One group of these CO_2-concentrating mechanisms does *not* involve covalent bond formation between CO_2 and an organic compound before RUBISCO activity. These mechanisms have considerable diversity, and can involve active transport of CO_2, HCO_3^-, or H^+ as means of concentrating CO_2. Such processes occur in many aquatic O_2-evolvers, including all of the cyanobacteria, many of the algae, and some of the submerged vascular plants found in aquatic environments, and in a few terrestrial O_2-evolvers, including all of the cyanobacteria, many of the terrestrial algae and lichens, and, among embryophytes, certain hornworts. There is an energetic cost of operating the CO_2-concentrating mechanism; this has not been well defined for any mechanistic variant. In the case of cyanobacteria and dinoflagellates, the kinetic properties of RUBISCO mean that C_3 physiology is not an option for growth in an air-equilibrium solution. In some of the other organisms, the selective significance of a CO_2-concentrating mechanism is quantitatively less clear (see **Table 3**).

A second group of CO_2-concentrating mechanisms *does* involve covalent bond formation between CO_2 and an organic compound in the delivery of CO_2 from the environment to RUBISCO. These mechanisms are essentially restricted to vascular plants, with C_4 metabolism occurring only in certain flowering plants, whereas CAM (Crassulacean Acid Metabolism) is found in a wider range of vascular plants. Here the initial carboxylation reaction preceding RUBISCO is almost invariably catalyzed by phosphoenolpyruvate carboxylase (PEPC) to yield oxaloacetate, which is subsequently converted to malate or aspartate. In the case of C_4 metabolism, the formation of malate or aspartate occurs in mesophyll cells, with transfer of the C_4 acids to bundle sheath cells that contain RUBISCO but have little capacity for CO_2 exchange with the atmosphere. In the bundle sheath cells the C_4 dicarboxylate releases CO_2, which is then fixed by RUBISCO, and the C_3 monocarboxylate residue is returned to the mesophyll, where it can be used to form more C_4 dicarboxylate. In C_4 plants the spatial separation of carboxylation and decarboxylation between two cell types involves only a temporal separation of seconds or tens of seconds. For CAM carboxylation to produce the C_4 dicarboxylate, malate and its decarboxylation to regenerate CO_2 occurs in a single cell type, but with an $\sim$12-h diel separation of carboxylation (in the dark) and decarboxylation (in the light). The energy cost of these biochemical CO_2-concentrating mechanisms in addition to the 3ATP and 2NADPH required for the photosynthetic carbon reduction cycle in the absence of oxygenase activity is $\sim$2ATP for C_4 and $\sim$3.5ATP for CAM, giving a total per CO_2 fixed of 5ATP and 2NADPH for C_4 and 6.5 ATP and 2NADPH for CAM (see **Table 3**). C_4 plants are almost all terrestrial; while there are very large numbers of terrestrial CAM plant species (more than 20,000), there is also a significant aquatic representation ($\sim$100 species). In the case of terrestrial plants, C_4 plants lose less water in transpiration per unit carbon fixed than do C_3 plants, but with the similar or greater productivities; CAM plants are even more economical in terms of water loss, but generally have lower productivities when similar life-forms are compared.

Conclusions

The core reactions of primary photochemistry (RC1, RC2 in Bacteria and eukaryotes; bacteriorhodopsin and halorhodopsin in Archaea) show relatively little variation in structure or overall function, although there are significant differences (e.g., RC2 in O_2-evolvers can dehydrogenate water). Similarly, the quinones, the cytochromes b–c_1 and b_6–f, cytochrome c, and ferredoxin also show little variation, as do plastocyanin (alternative to cytochrome c_6 in some O_2-evolvers) and flavodoxin (alternative to ferredoxin in some phototrophs). ATP synthetase also shows great similarity among bacterial and eukaryotic phototrophs. There is more variation among the reactions upstream of the photochemical reactions that are involved in photon harvesting and excitation energy transfer.

The core CO_2 fixation reaction using RUBISCO in O_2-evolvers and Proteobacteria is supplemented in some O_2-evolvers by a diversity of CO_2-concentrating mechanisms. In many aquatic plants there are CO_2 pumps not based on preliminary carboxylation and decarboxylation, whereas these reactions are found in C_4 and CAM. Bacteria other than Proteobacteria fix CO_2 via reactions not involving RUBISCO.

See also: Archaea, Origin of. Carbon Cycle. Eukaryotes, Origin of

References

Badger MR, Andrews TJ, Whitney SM, *et al.* (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2 concentrating mechanisms in algae. *Can. J. Bot.* 76: 1052.

Blankenship RE, Madigan MT, and Bauer CE (eds.) (1993) *Anoxygenic Photosynthetic Bacteria.* Dordrecht, Netherlands: Kluwer.

Falkowski PG and Raven JA (1997) *Aquatic Photosynthesis.* Malden, Massachusetts: Blackwell Science.

Hall DO and Rao KK (1994) *Photosynthesis*, 5th ed. Cambridge, United Kingdom: Cambridge University Press.

Lawlor DW (1993) *Photosynthesis: Molecular, Physiological and Environmental processes*, 2nd ed. London: Longman Scientific.

Nicholls DG and Ferguson SJ (1992) *Bioenergetics 2.* London: Academic Press.

Phylogeny

Kevin C Nixon, Cornell University, NY, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 559–568, © 2001, Elsevier Inc.

Glossary

Homology Any similarity that is explained by common ancestry.

Monophyletic group, monophyly A group of taxa that includes all of the descendants of the most recent common ancestor for the group.

Paraphyletic group, paraphyly A group of taxa that includes some, but not all, of the descendants of the most recent common ancestor for the group.

Synapomorphy A character state that is derived relative to the ancestral state (plesiomorphy). Synapomorphies are special cases of homology that are evidence of monophyly among the taxa that bear them relative to taxa that bear the plesiomorphic state(s).

Historical Overview

In the late eighteenth century, taxonomy became stabilized with the work of Linnaeus and others, and the modern hierarchic system of classification (often called the Linnean system) became standard. Hierarchical classification systems in the late eighteenth and early nineteenth centuries were often illustrated with treelike diagrams, but in general these were not intended to show evolutionary relationship and should not be interpreted as such. Following Darwin and the ultimate acceptance of evolution as the explanation for the diversity and variation of life, efforts began to produce phylogenies (diagrams of higher level evolutionary divergence among taxa), which also were often presented as treelike diagrams. The methods by which these early phylogenies were produced varied, most often based on a combination of raw similarities among taxa and either implicit (or explicit) ideas about the way in which characters had evolved. For example, Bessey in the early twentieth century published a list of dicta indicating trends in morphological characters of flowering plants (e.g., flower parts primitively with many parts, becoming fewer) and used these dicta as an implicit means of deriving phylogenies. Anecdotal theories about character evolution have been utilized by many systematists as an explicit or implicit basis for major phylogenies even to the present day (e.g., Cronquist, 1981).

Early phylogenies were usually considered independently from systems of classification, or classifications were only partially based on the results of phylogenetic analysis. One of the most prominent features of classifications from this period that persists in many current classifications is the recognition of paraphyletic groups (see below).

In a modern discussion of phylogeny, there are three main aspects that need to be addressed. First is the issue of how the phylogeny is derived: the source of evidence and how that evidence is analyzed. Second, there is the nature of the resulting phylogenetic diagram, or pattern, and what information such a diagram conveys. Third, there is the relationship between a phylogenetic system (a phylogenetic tree) and a classification that may or may not be derived from such a phylogenetic pattern. All of these issues remain controversial at some level, with different workers taking different combinations of views on each topic. While some may view these controversies as overwrought, one healthy consequence for science is the competitive improvement of methods and the need for explicit theoretical justifications by all participants in these arguments.

Modern Phylogenetic Analysis

Beginning with the popularization of the works of Hennig in the late 1960s and early 1970s, the idea that raw similarity should be the basis for phylogenetic reconstruction was quickly overturned. One might refer to this as the "cladistic revolution." By the mid-1980s Hennig's notion that only special similarities, or synapomorphies, should be used to reconstruct relationships had become the dominant force in phylogenetic analysis. Although one still finds phenetic analyses published, these are often focused on infraspecific patterns where cladistic methods have not yet been fully developed. The vast majority of phylogenetic analyses in the late 1980s to the present were undertaken with cladistic methods, in particular parsimony (see below). Other related methods such as neighbor-joining and maximum likelihood have been embraced by a subset of workers, but for various reasons are still not as widely used as parsimony. Thus, the discussion here will focus primarily on parsimony.

Analytical Methods

Phenetic Methods

Phenetics had its basis in the statistical literature that developed out of the "New Synthesis" and came to fruition with the advent of large and accessible computers in the late 1950s

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00108-8

and early 1960s. The various methods of phenetic analysis are too many to enumerate here but generally can be classified into methods that produce treelike diagrams (clustering) and methods that produce scatter diagrams on two or more axes (ordination). Phenetic methods do share the premise that all similarities, whether derived or primitive, should be used equally in an analysis. It is debatable as to whether the original proponents of phenetic methods actually intended the results of such analyses to be interpreted as phylogenies, but cluster methods produce treelike diagrams that were often directly interpreted as phylogenetic trees. Various justifications for phenetic methods have been proposed, and indeed, the methods as well as the theoretical underpinnings are eclectic. It became apparent in the 1970s and early 1980s that phenetic methods were not defensible as methods for recovering phylogenetic history, and parsimony analysis became the method of choice for most systematists interested in reconstructing patterns of divergence. Those interested in these debates can find entry into the literature by examining almost any issue of the journal *Systematic Zoology* from ca. 1970 to 1985. By the 1990s, phenetic methods had become rarely used except in particular subspecific and populational studies, and need not be considered further here.

Phylogenetic Methods

From the standpoint of methods of analysis, there is little to be gained by separating phylogenetics from cladistics although it should be noted that cladistics is the broader of the two. Cladistics is derived from the Greek term *clado* for "branch" and was coined by opponents of such methods to refer to those who were more interested in the pattern of divergence than in the process and degree of divergence. Originally the term cladistics was mostly applied to parsimony analysis (and some related methods that have fallen out of favor) but at the present time cladistics might be more broadly interpreted to include at least two other methods, neighbor-joining and maximum likelihood. However, as with all such terminology, there are differences of opinion, and some workers would rather restrict the term cladistics to parsimony-based methods only. For the purposes of this discussion, we will use the term in the broader sense. Any method that results in a (nonreticulate) tree diagram that is interpreted as representing the evolutionary history and relationships of taxa can be considered phylogenetic. Because of problems in interpreting most phenetic diagrams as phylogenies, those methods cannot be considered phylogenetic.

Parsimony

Parsimony analysis remains the method of choice for analysis of character data by most morphological systematists, and perhaps still even for molecular systematists. Modern parsimony analyses utilize matrices of either molecular or morphological data, or a combination of the two, to produce cladograms. Although Hennig did not propose parsimony analysis as we now implement it, his works provided the basis for the mathematical and theoretical underpinnings of parsimony analysis that were largely developed by Farris and others from 1969 to 1981 in a series of papers published in *Systematic Zoology* and elsewhere. Parsimony is derived from Occam's razor—the idea that the simplest explanation is always preferable. Parsimony analyses (when successful) result in cladograms (branching diagrams) that have the fewest number of steps (character transformations) when characters are optimally mapped onto the diagrams. Such trees are usually more or less directly interpreted as phylogenies, and character steps on these trees represent the minimum possible number of evolutionary changes to explain the distribution of the character. Such an interpretation does not require the assumption that evolution itself is parsimonious; indeed, most such trees based on real data have multiple origins for many, if not most, characters in the matrix. Parsimony simply eliminates the need for ad hoc explanations for multiple character origins by reducing them to the minimum possible number.

Parsimony Analyses

One major difficulty in parsimony analyses (and to an even greater degree in maximum likelihood) is the computational difficulty in finding most parsimonious solutions. There currently exist no methods for directly calculating most parsimonious trees. Because of this, the problem is directly related to the number of possible trees to be examined and is considered to be NP-complete. Because there is no direct way to calculate shortest trees, brute force methods that examine many possible trees must be used. These methods are generally implemented by "branch swapping" on a starting tree and examining the length of each new tree generated by moving branches to different positions on the tree. For very small data sets (<15 taxa), it is possible to examine all possible trees and be certain that one has collected all of the most parsimonious trees. For slightly larger data sets (possibly up to 30 taxa, depending on the quality of data), it is possible to use a "branch and bound" algorithm that reduces the number of trees examined but still guarantees that all shortest trees are found. However, for the vast majority of analyses, more than 30 taxa, and in recent analyses, as many as several hundred taxa (or "terminals"), are included. It is important to note here that the number of characters (e.g., the length of a DNA sequence) only affects computation in a roughly linear manner, while computation time is affected logarithmically as the number of taxa increases. With larger analyses, branch swapping is used under various "quick" strategies to find locally optimal trees and these trees are then used as starting points for additional searches. Recent advances in search strategies have resulted in programs that can easily analyze data sets of several hundred taxa with a very high confidence that shortest trees have been found (see Goloboff, 1999). Currently, for matrices larger than 1000 taxa, quick estimation methods such as the jackknife are used instead of directly searching for shortest trees.

Homology

The term homology, like the term monophyly, is used in various ways in the systematic literature. However, most recent workers accept the term to mean a shared similarity that is due to (explainable by) common descent. Thus, any similarity that is not explained by common ancestry is not homology. Characters that are initially thought to be homologous but are found by analysis to be nonhomologous are termed homoplasious. Homoplasy is best considered as error in homology

assessment, which can be explained by parallelism, convergence, or other more complex processes such as hybridization and retained polymorphism. Homologous characters by definition are evidence of relationship. One theoretical basis for cladistic analysis using parsimony is that parsimony maximizes the homology statements presented in the data matrix. Thus, when we score two terminals (taxa, species, etc.) as having the same state for a character, we are asserting a hypothesis of homology (the "primary" homology of some workers). Parsimony then resolves conflict among such statements by picking the tree (cladogram) that maximizes the number of correct homology assessments (or minimizes the number of incorrect assessments). In so doing, it minimizes the number of ad hoc explanations that are necessary to explain similarities that appear to be homologous at first, but are not based upon the most parsimonious tree. Farris and others have demonstrated that the shortest trees maximize explanatory power as well as information content of a classification and are therefore preferable even outside of an evolutionary context. It is important to note that parsimony does not assume that evolution is parsimonious, only that parallelism and convergence are not correlated in such a manner as to be misleading. Given randomly distributed parallel and convergent false homologies, with enough "true" homologies parsimony will resolve the correct tree (the tree that most accurately reflects phylogenetic history). This assertion has been challenged recently by proponents of maximum likelihood, but this controversy remains unresolved at this time and it appears that all methods (parsimony and maximum likelihood included) fail under certain circumstances.

Neighbor-Joining (NJ)

Neighbor-joining is a method that is arguably cladistic and is favored by many molecular biologists because it is rapid and generally (at least as commonly implemented) results in a single tree. The basic algorithm of NJ is similar to the Distance Wagner method of Farris and constructs a tree by calculating distances among terminals and adding terminals in such a way as to minimize overall distance. While NJ is an expedient method for quickly obtaining an estimate of the most parsimonious tree, there seems to be no justification for preferring an NJ tree over a shorter tree obtained by parsimony analysis.

Maximum Likelihood (ML)

Maximum likelihood as a method of analysis is popular among many molecular biologists, particularly those who are more interested in models of evolution than in the actual phylogenetic pattern of taxon relationships. ML requires an explicit model of character transformation, with associated probabilities for each possible transformation from one state to another. Trees are searched in a manner similar to methods used in parsimony (e.g., branch swapping) but each tree is evaluated not by overall length, but instead by a measure of compound probability. Those trees with the highest compound probabilities (maximum likelihood) of character distribution are selected as best. Within certain theoretical frameworks, parsimony can be viewed as a particular form of maximum likelihood with reduced assumptions and infinite parameters. Under such circumstances, parsimony analyses and maximum likelihood analyses will give the same results. Under most circumstances, maximum likelihood and parsimony analyses of the same data sets have provided very similar results. However, at the present time maximum likelihood is not feasible for larger data sets due to massive computation times (at least with today's hardware and software). The computational problems are unlikely to be resolved by the mere improvement of hardware and will require advances in software that are probably not possible with current engineering capabilities in this field. However, it is possible that recent advances in algorithms for parsimony searches can be incorporated into maximum likelihood programs with similar relative levels of improvement. Because of the additional complexities of ML, even with such improvements, the ability to perform maximum likelihood analyses on the ever-larger data sets being produced with molecular techniques will lag behind parsimony for the foreseeable future.

Measures of Support

The Bootstrap

Many workers wish to place some measure of statistical confidence on the phylogenetic trees that they construct. Because of theoretical issues, it is preferable to refer to such statistical measures as support values as opposed to confidence intervals. The most widely used support measure in recent literature is the bootstrap. The bootstrap is calculated by permuting the original data set numerous times by sampling characters with replacement and recalculating the most parsimonious trees for each permuted data set. Usually, from 100 to 1000 separate permuted data sets are analyzed, and support for a particular clade can be evaluated as the number of separate replicates in which that clade occurs in the most parsimonious solutions. Thus, if a group is found in parsimony analysis of 90 of 100 permuted data sets, its bootstrap value would be 90. Theoretical issues surrounding the bootstrap include the assumptions about character distribution and sampling and the exact interpretation of the bootstrap values that are obtained for each branch in the tree. Because the relationship between the bootstrap values and the underlying "true tree" is not apparent, many workers prefer to view bootstrap values as an assessment of the strength of signal within a particular data set, and not as a statistical measure of confidence except in a very restricted sense. This has resulted in development of an alternative terminology, under which bootstrap (and jackknife) values are referred to as "measures of support."

The Jackknife

The jackknife is closely related to the bootstrap and is performed also by permuting the data set and counting the occurrence of groups (clades) among replicates. The jackknife differs from the bootstrap in that instead of resampling characters, characters are removed at a set level of probability—the probability of removal usually set to be the same as the probability of not being sampled in a bootstrap. The jackknife has some theoretical advantages over the bootstrap and is generally faster to calculate because of very fast software

that is available. The jackknife has been used recently as the sole analytical method for some very large molecular data sets (e.g., Kallersjo's analysis of an rbcL data set for green plants with more than 2000 taxa). Indeed, some proponents of the jackknife have suggested its exclusive use in lieu of more intensive parsimony analyses, under the view that only well-supported clades need be recovered. This is not an abandonment of parsimony, but rather a rejection of poorly supported conclusions under a parsimony optimality criterion.

Bremer Support

Another measure of support commonly used in conjunction with parsimony analyses is Bremer support, sometimes incorrectly termed the "decay index." Bremer support is measured by collecting suboptimal trees (trees of longer length than the shortest) and determining which groups are present. If, for example, a clade is present in shortest trees and all trees up to 2 steps longer, but is not present in all trees 3 steps longer, the Bremer support value for that group is 3 steps. While Bremer support is appealing because it utilizes parsimony directly, it is difficult to compare Bremer support values among different data sets. In general, the bootstrap and jackknife have been preferred in recent years as measures of support, and especially for large data sets, the former two are much easier to calculate than Bremer support.

Data Sources

Phylogenetic analyses have evolved from being entirely or largely based on morphological data to being more and more based on molecular data, in particular DNA sequence data. There are both theoretical and practical reasons for this shift. First, many believe that molecular data are more objective and that character definition and scoring of morphological characters is both more subjective and subject to error. Because the expense of sequencing DNA has dropped dramatically, it is arguable now that it is actually more cost-effective to extract and sequence DNA than to spend large amounts of time tediously collecting morphological data. One problematic aspect of DNA sequence data is the lack of such data for fossil taxa. Fossil taxa, when available and when well preserved, can often be included in morphological cladistic analyses and may provide additional insight into patterns of relationship. Occasionally, such fossils even may have dramatic effects on the position of particular taxa. A major caveat of the inclusion of fossil taxa in cladistic analyses is the large amount of missing or incomplete data for fossils, especially in "soft" features that do not preserve well. Terminals with large amounts of missing data have the potential to "move" on the most parsimonious trees, resulting in many equally parsimonious solutions and in essence "deresolving" the tree. Such taxa that move because of ambiguity in data are often referred to as "wildcards" because of the propensity for them to match multiple terminals in the analysis. In general, it is best not to include such fossils when there is evidence that they are not contributing relevant information to the analysis but instead are reducing the precision of the results.

In the early 1990s a debate developed about the best way to treat disparate data sets, such as different gene sequences, or gene sequences and morphology. One side suggested that data sets should be analyzed separately unless it could be shown that they were congruent, while the other proposed that data sets should be combined as the best method to resolve such incongruence. Currently, it appears as though the latter position is more popular, since most multigene studies combine data sets in a simultaneous analysis regardless of whether tests indicate congruence.

The availability of vast quantities of molecular data across broad groups of taxa holds the promise of providing more or less complete large-scale phylogenies in the near future. Theoretically, larger analyses with more taxa should provide better estimates of relationship that are less sensitive to the addition of new data. Using parsimony jackknifing, Lipscomb and coworkers (1998) analyzed all of the eukaryotes in the Ribosomal Database Project (480 eukaryotes, 15 prokaryotes as outgroups). This analysis confirmed previous morphological and molecular studies that indicated that the protista and fungi were not monophyletic. However, this analysis showed that several of the startling and "new" findings of molecular systematics were probably the result of analysis of small data sets. In particular, there was no evidence for placing Giardia and its relatives as the most primitive eukaryotes nor any clear indication that the Dinoflagellates were sister to the Apicomplexa.

Phylogenetic Interpretation

One aspect of phylogeny and phylogenetic reconstruction that often presents difficulty is the interpretation of phylogenetic trees as evolutionary diagrams. Imprecision in both terminology and concepts has resulted in interpretations that are often inaccurate and unsupported. One of the most common mistakes in interpreting cladograms (or phylogenies) is the tendency to view the trees as providing a linear sequence of events, with a "first branch" that bears a more primitive group, followed by a "second branch," etc. Such terminology generally confounds group size with time of origin, usually placing the smaller extant group of two concordant groups at the "base" of the tree. Thus, recent molecular analyses of the angiosperms have often been characterized as placing Amborella at the base or as the first branch of the angiosperms (see **Figure 1**). However, the correct way to view such diagrams is in a *dichotomous* linear time scale, with each successive dichotomy younger than the preceding. In **Figure 1**, time can be viewed as proceeding from left to right. Amborella is not itself the first branch anymore than the angiosperms minus Amborella is the first branch—the first two branches of the angiosperm tree bear a single species, Amborella, on one side and a few hundred thousand species (the remainder of angiosperms) on the other side. In terms of primitive or derived features, Amborella might be more derived than numerous persisting species on the other, larger branch of angiosperms and should not be viewed simply as the "most primitive" angiosperm. For instance, any one of several taxa near the base of the diagram (e.g., Nymphea) might be more similar to the primitive angiosperm than is Amborella.

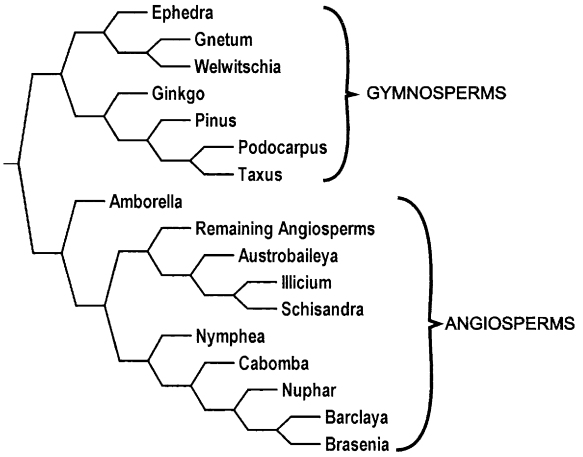


Figure 1 Reduced phylogenetic tree based on a recent large-scale molecular parsimony analysis of seed plant relationships. The results suggest that Amborella is the remaining extant sister taxon of the remainder of angiosperms; note that this tree does not indicate that Amborella is "the most primitive" flowering plant. See text for discussion.

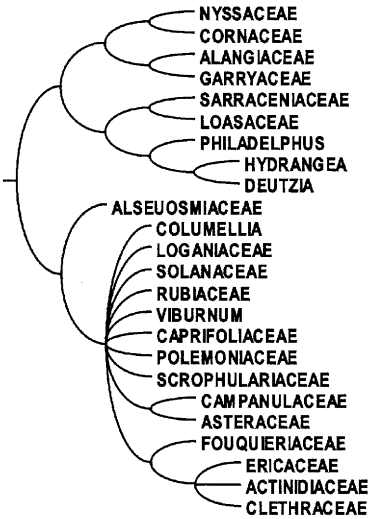


Figure 2 A portion of a phylogenetic tree based on a consensus tree from a parsimony analysis of a selected group of families of flowering plants. Note the large polytomy, or "star phylogeny" as such patterns are sometimes called. Such structures do not necessarily indicate rapid divergence from a common ancestor, but instead reflect lack of sufficient evidence, or conflicting evidence, regarding relationships. The graphic tree style was selected to more clearly indicate polytomies.

However, Amborella might be considered to be the "most unique" angiosperm since it is the sole representative of one of the two basal clades (based on extant species alone) of angiosperms, and in that sense is the equivalent of the remainder of angiosperms. This simple and direct view of branching patterns in a phylogeny is not only the correct way to interpret phylogenetic diagrams, but it also provides the basic concept of phylogenetic diversity, or the measure of phylogenetic uniqueness, that is outlined below.

Another common mistake in dealing with phylogenies is the interpretation of polytomies, or multifurcations in cladograms, as indicating rapid radiation from a single ancestral plexus (see **Figure 2**). Often, such patterns are referred to as "star phylogenies." Although evolution may not always produce dichotomous divergences, it is not possible to positively identify multifurcations as rapid divergences. Instead, such patterns may simply reflect a lack of sufficient data to resolve more completely bifurcating trees. In general, polytomies reflect a lack of conclusive evidence that is due to either character conflict or lack of data. Polytomies are best viewed simply as areas of the tree that need additional data to resolve, unless other independent data support a hypothesis of rapid divergence at the point of the polytomy.

Another frequently misunderstood aspect of phylogenetic analysis is the role of *a priori* character polarization. In the 1970s, many of the computer programs used to calculate phylogenetic trees required *a priori* character polarity (the assignment of "primitive" and derived states). However, strict parsimony analysis will produce the same topology regardless of the ultimate character polarity of the characters; in fact, parsimony computer programs do not consider character polarity during tree searches. Beginning in the late 1970s, it became apparent that the correct method for rooting trees was to select one or more outgroups that were included in the analysis, and the resulting trees rooted between the outgroup(s) and the ingroup. Although the best choices for outgroup taxa are those that are closely related to the ingroup, virtually any taxa not part of the ingroup may be used. However, the more distant the outgroups, the less accurate the rooting is likely to be, and particularly with molecular data, if the outgroups are too far removed, virtually random patterns of rooting may emerge. Character polarity becomes a matter of reading character distributions from the resulting trees when the outgroup method is employed, not an *a priori* exercise. Not only is this far easier than attempting to assign polarities to characters directly, but it reduces a major source of subjectivity in the analysis of data. Maximum likelihood models that allow equal probability for changes in any direction are also calculated independent of *a priori* rooting, but any models that have asymmetric models of character change are sensitive to a choice of a root, or starting point.

Phylogeny and Classification

The issue of how to classify taxa into formal named systems may seem trivial to some biologists, but it has a profound impact on the way that we think about living things, and therefore does have an impact on the design and implementation of many biological studies. Statements such as "reptiles are cold blooded" or "reptiles have scales" or "the dinosaurs went extinct in the late Cretaceous" all have different interpretations depending on the classification in use. Characterization of paraphyletic groups on the basis of widespread and notable characters (e.g., the scales of "reptiles" or the naked seeds of "gymnosperms") may provide useful means of communicating, but in formal classifications paraphyletic groups often provide misleading implications about relationship. Thus, if we include crocodiles in the group reptiles, we might

guess that crocodiles are more closely related to lizards than to birds; this in fact is not the case, based on several lines of evidence when analyzed cladistically that reveal that crocodiles are more closely related to birds than to other "reptiles"—in other words, crocodiles and birds form a monophyletic group that excludes other reptiles and mammals (see **Figure 3**). The advantage of a phylogenetic classification lies in the information content; if we know that a classification adheres strictly to the rule that all named groups are monophyletic, then all members of the group will be more closely related to each other than to any taxa *not* included in the group. Traditional groups such as "reptiles" (or formally, Reptilia) do not meet such criteria, and unfortunately it is impossible to know from such a nonphylogenetic classification that crocodiles (when placed within the "reptiles") are more closely related to birds than to other "reptiles." This observation at once supports both the need to abandon paraphyletic groups in formal classifications, and in doing so, to adopt a system of classification that is based consistently on phylogeny—with named groups always corresponding to groups based on available evidence that are believed to be monophyletic.

Much of the difficulty that some systematists and many outside of systematics have with the abandonment of paraphyletic groups comes from the intuitive nature of such groups and their often easy characterization (e.g., reptiles are scaly, air-breathing, cold-blooded vertebrates that lay eggs). A classification that recognizes paraphyletic groups must do so based on primitive (plesiomorphic) features, as in the example just provided. The difficulty, and the arbitrary factor in such classifications, comes not in what to include in a paraphyletic group but instead in what to exclude. Because there is no objective way to determine what groups are significant enough to be excluded, the final decision is left to the authority, and in the strictest sense of the term, such classifications are authoritarian—and each authority may recognize different groups. Thus, removal of the birds from the group "reptiles" is based on a particular viewpoint about the importance of the characters (e.g., feathers) that the groups possess. However, another authority might just as easily argue to remove snakes (they lost those encumbering limbs) or turtles (the ultimate innovation in defensive armor). Decisions to exclude (and thereby elevate) groups such as birds are ultimately based on a view on the importance of particular characters, and such classifications do not accurately convey information about relatedness. Unfortunately, because paraphyletic classifications are based on arbitrary assignments of significance to characters, they are not even useful in the prediction of innovation or levels of anagenetic change (there is no measure of the amount of change involved in the removal of concordant groups, only its significance). The realization that paraphyletic classifications are at their base arbitrary is the first step toward the acceptance of the need for classifications based solely on monophyly.

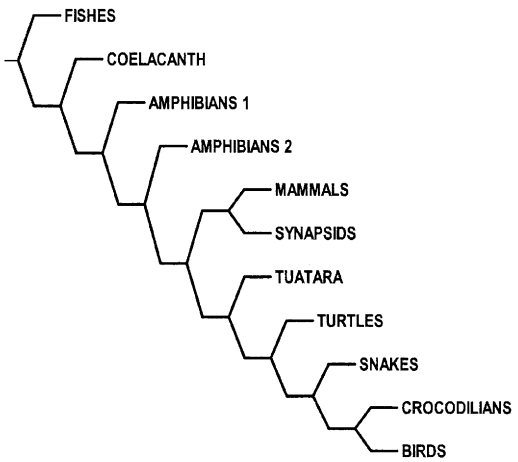


Figure 3 A phylogenetic representation of relationships among major groups of vertebrates, based on several published studies of morphological and molecular data. Note that crocodilians are more closely related to birds than to other "reptiles."

Ranks and Priority

The need for a useful classification to be based on phylogeny has been confounded in recent years with other aspects of classification, such as ranking and the priority of names. Recently, a movement has emerged that has been self-proclaimed as "Phylogenetic Systematics," which putatively embodies the ideals of Hennig and requires strict translation of phylogenies into named classifications. However, the need for maintaining only monophyletic named groups has been intertwined with other proposals, such as the abandonment of ranks, and proclamations that the existing Linnean system is incapable of accommodating a truly hierarchic system. Of course, the Linnean system is a strictly hierarchic system, and the claims that it must be abandoned are both premature and poorly reasoned. Unfortunately, many outside of this controversy have been misled to believe that the new proposals will provide greater stability in names, which is untrue. Stability in names in any phylogenetic system will come with stability in the underlying phylogenies on which the system of classification is based, and we are on the verge of achieving such stability with the vast quantities of molecular data and improved analytical algorithms that are becoming available.

Applications of Phylogeny

General Evolutionary Context

In recent years, phylogeny has become an important aspect of many evolutionary and ecological studies. Phylogenetic studies focus on the pattern of history, but provide the foundation for both generalizations and specific conclusions about processes of evolution and diversification. Generalizations about evolutionary process based on phylogenetic trees are more predictive because phylogenetic trees provide information about character distributions as well as relationship. Phylogenies also provide the ultimate test of particular ideas or theories about the course of evolution. By mapping character distributions in the most parsimonious possible manner on a well-supported phylogeny, it is possible to understand whether characters of interest have evolved once or more than once and whether particular characters are correlated in a

historical sense. Attempts in the past to accomplish these goals were often based on paraphyletic classifications that were misleading or based merely on ranked classifications with unknown or unstated bases. Without a clear understanding of phylogenetic pattern, these studies often miscounted the number of times evolutionary steps occurred within and among taxa and came to incorrect conclusions about character evolution. Precise phylogenies now provide a means to improve such studies.

Adaptation

Phylogenetic trees have become a standard tool in the study of adaptation, and such uses are often referred to as the "comparative method." First, it is necessary to establish that a particular "adaptation" is distributed as an apomorphy within the group in question and then, if there are multiple origins, to determine if these origins are correlated with other characters and/or environmental variables. While numerous statistical approaches have been suggested for such studies, they all assume that multiple independent origins of characters correlated with environmental or historical factors are evidence of adaptation. Indeed, some workers maintain that it is only possible to discuss adaptation in a historical context, i.e., based on explicit phylogenetic trees. Undoubtedly continued work in these areas will result in improved statistical tests for adaptation based on character distributions on phylogenetic trees.

Vicariance Biogeography

Phylogenetics has also become the main tool for biogeographic studies, and indeed the field of biogeography is now almost synonymous with vicariance biogeography. Vicariance biogeography is rich in theory and a thorough explication is beyond the scope of this article. The basic premise is that historical patterns of land connection and movement are recoverable through finding repeated patterns among the phylogenies of various plant and animal groups. Repeated patterns of clade distribution suggest a common history that is interpretable as supporting wider continuous distributions in the past. A well-known example of a broad vicariant pattern of distribution is found in the plant genus *Nothofagus* (Nothofagaceae) in the Southern Hemisphere, with species distributed in South America, New Guinea, Australia, New Caledonia, and New Zealand and fossils known from Antarctica. Similar patterns in other plants (e.g., Proteaceae) and animals (e.g., ratite birds) can be evaluated in a phylogenetic context to provide evidence of interchange among different "areas of endemism" and develop an "area cladogram" that presents a pattern of relationship among the land areas. These patterns can then be interpreted in the context of plate tectonics as indicators of historical fission of land masses and subsequent plate movements.

Phylogeny and Biodiversity: Phylogenetic Diversity

Phylogenies also provide a framework for alternative ways of looking at biodiversity. Most measures of biodiversity use species richness, either in a geographic (either broad or local) or ecologic sense. Other views of diversity may focus on the nature and breadth of adaptation. Such measures unfortunately require a subjective view of the importance of particular adaptations—e.g., the birds might be considered "diverse" because they include so many adaptations for use of the bill. However, phylogeny provides another perspective on biodiversity that allows an objective way to compare uniqueness and diversity of taxa. Although various specific measures of phylogenetic diversity have been proposed, most share a basic approach by which phylogenetic trees are used to evaluate species richness in concordant groups. It means little to say that "orchids are highly speciose" or "monotremes are species depauperate" unless we have some idea as to the relationships of the two groups being compared. Phylogenetic pattern provides the basis for such comparisons.

Currently, the use of phylogenetic diversity measures is largely limited to theoretical discussions and there have been few efforts to actually apply such measures to conservation. This is partly due to the relative paucity of high-quality phylogenies that are available across broad groups of taxa and partly because of a distinctly ecological bias in most studies of biodiversity. As molecular data sources provide better and more complete phylogenies for use by other workers, this is likely to change. It is probable that in the near future measures of phylogenetic diversity will become standard components, in combination with more traditional measures of ecologic uniqueness, species richness, and sensitivity, in the formulae that are used to evaluate conservation priorities for areas and endangered species.

Conclusions

An understanding of phylogeny is increasingly important as a basis for experimental design and for interpretation of results in a broad array of biological studies. New methods of analysis, improved computer hardware and software, and new and improved sources of data, most notably DNA sequences, have revolutionized our ability to construct phylogenetic trees. It is likely that our understanding of the origins and diversification of most major lineages of plants, animals, and microorganisms will improve dramatically over the next century, and we will arrive at a stable and robust reference system for all of life on earth.

See also: Biodiversity, Evolution and. Cladogenesis. Darwin, Charles (Darwinism). Evolution, Theory of. Taxonomy, Methods of. Vicariance Biogeography

References

Cronquist A (1981) *An Integrated System of Classification of Flowering Plants*. New York: Columbia Univ. Press.

Goloboff PA (1999) Analyzing large data sets in reasonable times: Solutions for composite optima. *Cladistics* 15(4): 415–428.
Hennig W (1966) *Phylogenetic Systematics*. Urbana, IL: Univ. of Illinois Press.
Lipscomb D, Farris JS, Kallersjo M, and Tehler A (1998) Support, ribosomal sequences, and the phylogeny of the eukaryotes. *Cladistics* 14: 303–338.
Nixon KC (1996) Paleobotany in cladistics and cladistics in paleobotany: Enlightenment and uncertainty. *Rev. Paleobot. Palynol* 90: 361–373.
Schuh RT (2000) *Biological Systematics: Principles and Applications*. Ithaca, NY: Cornell Univ. Press.

Plankton, Status and Role of

Colin S Reynolds, Freshwater Biological Association and NERC Institute of Freshwater Ecology, Cumbria, UK
Judit Padisák, University of Pannonia, Hungary

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Colin. S. Reynolds, volume 4, pp 569–599, © 2001, Elsevier Inc.

Glossary

Autotrophy The ability of organisms to grow and reproduce independent of external sources of organic carbon compounds.
Eukaryote An organizational state of cellular organisms in which the genome of the cell is stored in chromosomes enclosed in a membrane-bound nucleus; all protists (algae and protozoa), fungi, plants, and animals are eukaryotes.
Euphotic The top layer of a water body through which sufficient light penetrates to support net photosynthesis. Rarely more than 100 m in depth, the euphotic layer can be as little as 1 m in turbid waters.
Heterotrophy The ability of organisms to grow and reproduce on organic carbon sources, taken in dissolved or particle form.
Metazoan Literally, a multicelled animal.
Mixotrophy The ability of a normally autotrophic organism to switch, circumstantially, to phagotrophy, or to support an otherwise meager food supply by resorting to the ingestion and assimilation of bacteria or their products.
Phagotrophy A type of heterotrophy that involves the consumption of protists, plants, or animals as food.
Photoautotrophy A type of autotrophy in which organisms gather light energy in order to reduce carbon dioxide to organic carbon; characteristic of green plants, most algae, and some prokaryotes.
Picophytoplankton The smallest (<2 μm) size class of photoautotrophic plankton.
Prokaryote Organizational state of cells lacking a membrane-bound nucleus and certain other organelles. Bacteria, including the cyanobacteria, are typically prokaryotic.

Introduction

"Plankton" is a collective term for organisms adapted specifically for a life in suspension in the open waters (the pelagic zone) of the sea and of such inland waters as lakes, reservoirs, and rivers. Planktonic organisms include protists (allegedly simple, unicellular, or colony-forming algal primary producers and their protozoan consumers), microorganisms, and certain types of small metazoan animals, all sharing a common liability to passive entrainment in water currents, generated by tide, wind, convection, gravity, and the rotation of the earth. The inherent physical variability of open-water habitats typically favors short life histories; rapid changes in dominant species composition, in response to fluctuating environmental conditions, contribute to the maintenance of high biological diversity in individual habitats and to the survival of a high species richness among planktonic assemblages in general. Genuinely, planktonic organisms – which include the plantlike chlorophyll-containing primary producers of the phytoplankton, the heterotrophic decomposer microorganisms of the bacterioplankton, and the more animal-like consumers of the zooplankton – are too small (often <20 mm) for their own intrinsic movements to be able to overcome, often or at all, the dispersive effects of water movement. Thus, "embedded" within the tireless and unconstrained motion of open water, planktonic organisms broadly go wherever the flow takes them. In this way, the ecology of plankton is inextricably related to the physical properties of the medium, the extent and limits of its motion, and the environmental conditions set within these bounds (Reynolds, 1997). This situation contrasts with that of most fish and other larger (>20 mm) animals of open water – the "nekton" – whose swimming strength is usually able to overcome normal movement of the water.

The older literature promulgated a view that suspension in water was necessarily beneficial, supposing water to be something of an ideal habitat. Living in water does confer some notable positive advantages over terrestrial or aerial habitats. These include the mechanical ("Archimedean") support water provides, as a consequence of its much greater density in comparison with air; its slow temperature fluctuations, as a consequence of its much higher specific heat than air; and its solvent properties, which maintain nutrients and metabolic gases in readily assimilable state. However, the planktonic ways of life have evolved to accommodate the several problems and drawbacks associated with living in open water. These concern (1) remaining in suspension, (2) exploiting nutrient resources, and (3) energy.

Specific gravity of planktonic organisms is typically higher than that of the water. Therefore, in calm water, planktonic organisms will ultimately sink. Evolution of plankton was largely driven by overcoming sinking loss and includes a number of mechanisms (active movements by means of flagella, cilia, contracting cells/tissues, shape, gas vacuoles, chemical composition of storage materials, decreasing sinking loss by developing complex forms to increase form resistance; Padisák *et al.*, 2003a, b).

Concentrations of obligate material components of living cells, especially the aqueous concentrations of some of these (especially carbon, nitrogen, phosphorus, iron, and 15 or so micronutrient elements) are often so dilute that their availabilities place a severe constraint on the assembly of planktonic biomass. Moreover, the absorbance of solar energy by

pure water is such that, at depths of > 100 m it is always as dark as night. Biological productivity in lakes is often severely constrained by rarefied resources or by deficiencies in processing energy, or by both. Far from being an ideal environment, the pelagic is a rather unpromising medium for successful exploitation by living communities.

Yet, within this general constraint, there is a remarkable richness of individual species inhabiting the plankton of the world's lakes and seas. As a preface to the planktonic biota, it is necessary to emphasize that the familiar separation, first of plants and animals and then their subdivision among phyletic divisions, cannot be applied too rigidly. The convenience of distinguishing planktonic "plants" and "animals" by function does not necessarily correspond to any fundamental evolutionary or phyletic separation. The first attempts at describing planktonic diversity by tools of molecular genetics resulted in recognition of much higher number of heterotrophic protists than was revealed by microscopic methods (Luo *et al.*, 2011). The perplexity is yet greater when referring to the bacteria: Although the lack of a membrane-bound nucleus and of certain other intracellular organelles provides a robust separation of their prokaryotic organization from the cells of eukaryotes, it is still difficult to distinguish among most bacteria other than by their biochemical activities and their affinities at the molecular level. Relatively "safe ground" is reached only among the more distinctive metazoan phyla, though even there, affinities among the main groups are often still obscure.

The Structure of Planktonic Communities

Phytoplankton is extremely diverse. More than 4000 species of marine phytoplankton have been named and described (Sournia *et al.*, 1991). The total number recorded from inland waters is not certainly known, but it is estimated that there are about 4000 of these as well (Reynolds, 1996). Few genera and still fewer species are common to both fresh and salt waters. The purple (Chromatiaceae) and green sulfur bacteria (Chlorobiaceae) are represented in specialized, anoxic habitats. Of the planktonic genera of cyanoprokaryotes (formerly classed as Cyanophyceae, or "blue green algae," and now most commonly referred to as Cyanobacteria), most occur in lakes, though several are also common in the low-salinity (< 11 parts per thousand) areas of the Baltic. The cyanoprokaryotes are also well represented among the smallest marine and freshwater primary producers (the picophytoplankton: cells < 2 μm in diameter; Waterbury *et al.*, 1979; Callieri, 2008).

The more conspicuous components of the phytoplankton of the open sea belong to the Pyrrhophyta (including a wide variety of dinoflagellate species) and Heterokontophyta. The latter large division is taken to include the large number of diatom species and the diversity of elaborate, scale-bearing Coccolithophorids is also remarkable. Besides diatoms and Cyanoprokaryotes, the more conspicuous components of the freshwater phytoplankton may be contributed from among the many chlorophyte and chrysophyte orders. However, cryptomonads, peridinians, chloromonads, and euglenoids can all occur in large numbers at certain locations (Krienitz, 2009).

The number of species represented in the marine zooplankton of the sea is considerably enriched by the distinctive dispersal stages of many marine animals that spend their adult lives in the littoral or the benthos. To differing extents, these larvae (the amphiblastulae of sponges, the medusae and ephyrae of the cnidarian coelenterates, the pilidia of nemerteans, the trochospheres of polychaetes, the cypris larvae of cirripedes, the phyllosomae and zoeae of the eucarid malacostraca, the veligers of the lamellibranch mollusks, the various auriculariae, bipinnariae, and plutei of the echinoderms, and the appendicularian larvae of the ascidians) share the diminutive size ranges and feeble swimming movements characteristic of the species which are planktonic throughout their lives. The microzooplankton fraction (in the size range, 20–200 μm) includes the rhizopod foraminiferans and radiolarians, and a range of ciliate and suctorian ciliophorans. More conspicuous (0.2–20 mm) in the marine zooplankton are the ctenophoran "comb jellies" and sea-gooseberries, the chaetognath "arrow worms," some specialized turbellarians and polychaetes, certain opisthobranch gastropods, and the larvaceans and salps.

The most prominent of the animals of the marine plankton, however, are crustaceans. Most of the species are copepods, including calanoids and cyclopoids. There are also some representative cladoceran and ostracod genera. The malacostracan orders are distinctively represented by organisms that remain planktonic in their adult stages. Mysidaceans are locally abundant and a few cumaceans are remarkable for being nocturnally planktonic and diurnally benthic in coastal waters. Perhaps the most renowned component of the plankton of the high-latitude oceans, mainly for being one of the major food sources of the filter-feeding whales, is krill (*Euphausia* spp.). Some of the cnidarian jellyfish may be regarded as being the largest animals in the plankton (> 20 mm, perhaps to 2 m). Though some of these are rather larger than some of the swimming organisms ("nekton": fish, cephalopods) that are excluded from the understanding of "plankton" (discussed earlier), the large jellyfish qualify for their poor ability to control their own movements in the sea.

Finally, the young stages of several species of pelagic fish are of such diminutive size and swim so feebly and with weakness of motility that, for the first part of their lives, they are reasonably included among the plankton.

The phyletic representation in the freshwater zooplankton (Havel, 2009; Rudstam, 2009; Wallace and Smith, 2009; Williamson and Reid, 2009) is known to be relatively much poorer than in the sea, but this reflects mostly the poorer representation of the animal phyla among freshwaters. Protists usually figure prominently in the plankton of inland waters, supposedly as a function of the organic matter available, including detrital particles and their associated bacteria. The smaller heterotrophic flagellates generally consume free-living bacteria, but many of the common planktonic microciliates, like *Coleps* and *Tintinnidium*, feed also on small algae and flagellates. In doing so, they fulfill an important linkage in the so-called microbial loop (Azam *et al.*, 1983; Fenchel, 2008): In many systems, the excretion of excess organic carbon fixed in photosynthesis, its assimilation by bacteria and its successive transfer through closely coupled flagellate–ciliate consortial linkages to copepod consumers and, eventually, planktivorous

fish, demonstrably exceeds the carbon transfers along the conventional phytoplankton → zooplankton → fish trophic pathway (grazing chain). Planktonic fungi (especially chytrids) feeding on large sized, ungrazeable phytoplankton species are linked to the trophic system as a "mycoloop" sometimes retaining considerable amount of carbon in the pelagic systems (Kagami *et al.*, 2007).

Occasionally, the larger Amoebae and ciliates may dominate the zooplankton; this may owe to the prevalence of a particular food source, or to the fact that a crustacean plankton has not yet developed, or that some other factor (e.g., low oxygen concentration) restrains potential competitors for the same food resource.

At least two freshwater coelenterate genera have planktonic medusae. The Gastrotricha phylum of minute, wormlike but unsegmented animals is occasionally represented by planktonic specimens. Rotatoria is prominently represented in the freshwater plankton by some two dozen genera. The most ecologically important genera include *Asplanchna*, *Brachionus*, *Filinia*, *Keratella*, *Kellicottia*, and *Synchaeta*. Some are specialist feeders; most of those mentioned browse or filter-feed on bacteria, detritus, and planktonic algae, generally within species-specific size ranges (Pourriot, 1977).

As in the sea, the most prominent planktonic animals are crustaceans. The branchiopods are represented by the anostracan "brine shrimps," especially in temporary waters, and by the Cladocera, the familiar "water fleas." This grouping (its monophyletic origin is now doubted) includes the mainly herbivorous, filter-feeding species of *Daphnia*, *Ceriodaphnia*, *Moina*, *Simocephalus*, *Bosmina*, and *Holopedium*, and the predatory *Bythotrephes* and *Leptodora*. Planktonic ostracods are noted in lakes in Southeast Asia (*Cypria javensis*) and in Laguna de Petén, Guatemala (*Cypria petenensis*). The copepods are particularly well represented by calanoids (such as *Eudiaptomus*, *Eurytemora*, and *Boeckella*) and by cyclopoids (e.g., *Mesocyclops*); some parasitic cyclopoids (e.g., *Ergasilus*) are dispersed through the plankton. The branchiurans (e.g., *Argulus*), which are ectoparasitic on open-water fish, are certainly considered essentially planktonic. Though principally marine, the mysids are represented in the plankton of several high-latitude lake systems, where *Mysis* is regarded as a relict from the last glaciation. The Amphipod flag is carried by *Macrohectopus*, which is endemic to the plankton of Baykal Lake.

Among the arthropod insects, several genera of Diptera have larvae, which pass most of their time in the plankton. The most specialized of these are the juvenile chaoborines (e.g., *Chaoborus*) or phantom midges, whose transparent bodies reveal the internal provision of buoyancy-providing air sacs. The dispersal stages of aquatic larvae of other orders of insects sometimes show adaptations that are unmistakably planktonic: A striking instance is provided by the first instars of *Sialis* (Megaloptera; see Elliott, 1996).

Explaining Species Diversity

The foregoing passages confirm the richness of phyletic representation and the very large number of individual species that collectively contribute to the overall diversity of planktonic organisms. At one level, we might be surprised that this problem arises at all. It should not be unreasonable at all to anticipate that in a supposedly fluid and isotropic medium, fully accessible to suitably adapted species, Darwinian selection should move the structure of the assemblage toward just a small number of specialists, each being the best-fit survivor in its key community role. Every less-fit competitor might be supposed to suffer progressive exclusion by the stronger species; overall, diversity should be suppressed. Thus, to confront the remarkable richness of planktonic plant and animal species surveyed previously, is counterintuitive and paradoxical (Hutchinson, 1961).

Many theories have been advanced to explain Hutchinson's (1961) paradox of the plankton, but the conundrum for a long time remained unsolved. Either the number of occupiable niches must be far more numerous than had been supposed or the competition was somehow incomplete in its effects.

Hutchinson himself suspected that the assumption of a homogeneous environment with steady-state properties was flawed. Collectively, the environments that planktonic populations inhabit are subject to huge variability in their chemical makeup and in their physical characters. They may be physically separated from each other (as are freshwater catchments), enjoying quite different climates or, even if contiguous (like the seas), their systems may be close to mutual isolation by currents and circulations. On the continents, lake basins are created and destroyed at, relative to evolutionary rates, high frequency. The skeletal understanding of the history of the world's oceans is that they have undergone large oscillations in metabolism and productivity associated with changes in the biospheric carbon cycles (Thierstein, 1989). Superimposed on the longer term changes are relatively higher frequencies of periodic forcing. Within these smaller habitat units, the higher frequency environmental oscillations might lead to alternations in the favored specialisms and consequent interspecific transfers of competitive advantage. Physical limitation of range (endemism) or otherwise (cosmopolitanism) would then represent the two extremes of dispersal efficiency across physical barriers and habitat preferences.

Current understanding of plankton ecology conforms to this general view in that it is certainly possible for the experienced observer to determine the broad habitat provenance of given planktonic assemblages: Their species composition has an indicative value because similar assemblages characterize similar pieces of water. Most university textbooks accepted a prevailing view that planktonic species are generally cosmopolitan and disperse freely, so that, on balance, most are able to establish quickly wherever suitable conditions arise. Dispersal mechanisms among the protists and prokaryotes are certainly effective (Padisák, 2009), and it is true that many conventionally identified morphotypes enjoy worldwide distributions. As more molecular information becomes available, the cosmopolitan nature of plankton is increasingly called into question. Besides, the notion that "everything is everywhere" is surely a matter of degree: Ease and frequency of dispersal are not equal among planktonic species. Isolation and regional endemism occur (Tyler, 1996), especially where long physical distances or significant hostile

barriers separate sites of suitable habitat. It often takes anthropogenic intervention to bridge these gaps, as the recent "breakout" of the lamellibranch, *Dreissena polymorpha*, and hitherto endemic mysids from their original Caspian-basin locations into the waterways of Europe and, now, North America graphically reminds us. Equally remarkable is the arrival of the first freshwater Cladocera on Easter Island in 1780, which had failed to bridge the distance from the next nearest lake (more than 3000 km distant) prior to Captain James Cook's requirement to replenish his ship's supply of drinking water (Dumont, 1999). Energetic field naturalists carrying algae on their plankton nets from one lake to another can also be efficient dispersal agents (Talling, 1951).

Planktonic communities change frequently and fast within individual years. The differential responses of individual species of plankton to temperature, day length, or nutrient resources are generally recognized to underpin what is perhaps the most familiar feature of planktonic communities – the so-called seasonal succession. As with other aspects of plankton ecology, the identity of its driving variables has been actively pursued and described by conceptual word models; the most compelling of these has been the PEG-model of Sommer *et al.* (1986). It is founded on an implicit acceptance of the differences among planktonic species and the suitability of species-specific adaptations to particular habitat constraints.

The way to explain planktonic biodiversity is through a simultaneous recognition of the variables constraining habitat suitability and the adaptive specialisms and limitations of species for which the habitat constraints will select and those against which they will discriminate.

Habitat Constraints in the Plankton

The assumption of a uniform, hospitable steady state is erroneous so far as most open-water habitats are concerned. On the contrary, many are as hostile as their desert-like barrenness conveys. Relative infertility may be locally or regionally attributable to an inadequacy of light energy to sustain planktonic primary production and the import of organic carbon is too modest to support the heterotrophs. Light income may be scarce for reasons of latitude (e.g., short day length and low solar declination), or it may be severely "diluted" over a deep mixed layer, or its penetration may be abruptly curtailed by a high concentration of inert particles (turbidity). Commonly, the concentrations of assimilable sources of certain primary nutrients place a very low ceiling on the biomass that can be assembled (nitrogen, phosphorus, and iron are cited most frequently as capacity-limiting factors). For the planktonic animals, the problem is to be able to forage sufficient food; the concentrations and the size distributions of potential food particles influence profoundly the nature and abundance of the consumer populations.

Besides the spatial differentiation of planktonic habitats, there is usually a considerable temporal variability. Conspicuous are the seasonal changes consequent on latitude: Lengthening spring days and higher flux densities are generic to all temperate ecosystems but for small organisms with short life spans, for example, the changes are perceived by successive generations and not as some perennial amplitude of fluctuation registered by a forest tree. The onset of thermal stratification, as a consequence of a shifting balance between the buoyancy forces brought through surface warming and the kinetic dissipation of the work of wind energy, precipitates a train of environmental effects leading to enhancement of the segregation of the warm, insolated, aerated, and increasingly resource-depleted epilimnion from a colder, darker, and potentially less oxidizing hypolimnion. Also, such seasonality is not confined to high latitudes: Even quite small differences in wind prevalence, cloud cover, humidity, and hydraulic exchange result in seasonal variations in thermal stability and downmixing in the tropics (e.g., see Talling and Lemoalle, 1998), and convectional currents generated by day–night temperature differences (atelomixis, see Lewis, 1973) may lead to selection of species that are particularly adapted to this environment (functional group $\mathbf{N_A}$, see Souza *et al.*, 2008).

So it is that the different kinds of water bodies (lake or pond, river, estuary, coastal shelf, upwelling, or the open oceans) offer, to the most appropriately adapted species, or to those simply furnishing the largest inocula, varied and sometimes very transient opportunities to build their local populations. Furthermore, because the medium is fluid and the movement induced by wind or gravity is subject to fluctuations in strength, the persistence and vertical extent of an upper differentiated layer is a highly variable character of the environment, which alters not just from season to season but also from day to day and from hour to hour. Interest is also growing in the effect of year-to-year variations. These might be the result of mesoclimatic variability acting through salinity (Padisák, 1998) or extension of ice-cover (Padisák *et al.*, 2010) fluctuating over decadal scales.

It follows that selective advantage is likely to move among species. Environmental variability is a powerful influence on the assembly of planktonic communities and thus on their biological diversity. The causal linkages between observable patterns and processes in the maintenance of species diversity in the plankton may be usefully explored through a diagrammatic representation of the two principal variables characterizing particular open-water environments, namely, the fluctuating fluidity or vertical extent of the uppermost water layer and the availability of the resources to support the assembly of biomass. The layout of **Figure 1** is descended from Margalef's (1978) original scheme for marine phytoplankton, in which a "nutrient" axis is set against one of "turbulence;" the arrangement is amenable to tracking habitat variability and the changing species composition through time. The scheme has been developed for freshwaters to accommodate habitat conditions in which nutrient resources and the light availability in the surface-mixed layer are sufficient to saturate the fastest *in situ* algal growth rates (in the top left corner of **Figure 1**) and to distinguish them from conditions that become either increasingly resource-constrained (moving downward in the matrix) or increasingly energy-deprived (moving rightward), or fall deficient in both (bottom right corner). The selection of given species or groups of species was found broadly to be correlated with the matrix space thus represented; high-performance, fast-growing algae would be favored in the relative paradise of the upper left-hand corner; algae favored by the second and third contingencies show some specialist adaptations, respectively, for operating under

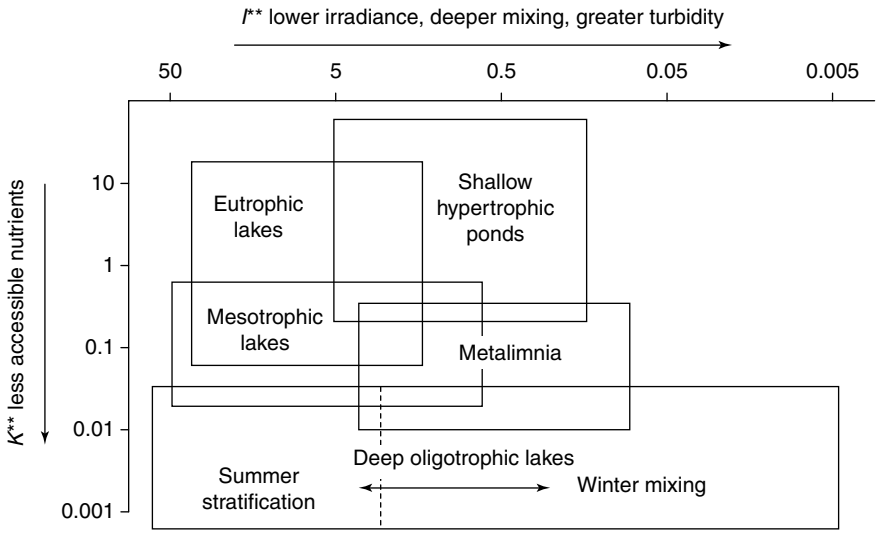


Figure 1 Two-dimensional representation of limnetic habitats, which, starting in the top left, distinguishes well-insolated, resource-replete water masses from (working down) long-stratified, resource-segregated water columns and from (working across) increasingly light-deficient environments.

conditions of low concentrations and remote sources of nutrients, or for harvesting light from low irradiances or from infrequent short periods in the light field. The fourth, low-nutrient, low-light contingency is unfilled and is something of a phycological desert, evidently untenable as a habitat for a photoautotroph (Reynolds, 1988b).

Effort has been invested to improve both the utility and the quantification of the axes. The horizontal axis in **Figure 1** describes the integral, I^{**}, which, as the product of the physical depth of mixing and the vertical attenuation of a finite input of light energy, is sensitive to the constraints of both deep mixing and high turbidity; quantities diminish rightward on a logarithmic scale. The vertical axis in **Figure 1** accommodates another composite scale, K^{**}, being the quotient of the concentration of the critical nutrient in the medium (usually the surface-mixed layer) and the gradient of concentration between the top and bottom of the entire trophogenic layer. This scheme differentiates sites that are chronically deficient in nutrients (points plotted well down the K^{**} axis) from, on the one hand, those in which steep gradients develop as a consequence of near-surface uptake (points track from high to low on the K^{**} axis) and, on the other, those in which the nutrient is scarcely exhausted (points located consistently toward the top of the K^{**} axis).

Against these axes, it is possible to characterize the "signatures" of seasonal changes in various kinds of water body, including of small, "eutrophic" lakes, in which nutrient availability is seasonally reduced, of shallow, fertile systems, in which the turbidity is the predominant variable, and of deeper, oligotrophic systems, where the depth of mixing is the strongest seasonal variable. An analogous (and in many ways, a more self-evident) approach has been developed for the sea (**Figure 2**). This distinguishes energy-limited, well-mixed, high-latitude oceans from the highly stratified, nutrient-segregated waters of the tropics and will represent processes and interactions in coastal and shelf waters, including frontal zones and major upwellings of deep circulation currents.

Form, Function, and Selection in the Phytoplankton

The essential environmental requirements of planktonic photoautotrophs, to be able to grow in numbers and increase in total biomass, do not differ fundamentally from those of plants in other ecosystems: Access to water, exposure to adequate levels of photosynthetically active wavelengths of light, a source of assimilable carbon and adequate supplies of each of a score of other elements. In the dilute world of the plankton, diffusion alone is rarely able to supply the resource requirements of plankton. Adaptive mechanisms for gathering chronically deficient resources include the enhancement of uptake affinity and of the maximization of storage. Mechanisms enhancing access to remote reserves or for exploiting alternative sources of nutrients are valuable against a depleting resource base.

The distinguishing feature of the photoautotroph is the ability to first generate the carbohydrate through the photosynthetic reduction of carbon dioxide. The energy to do this comes from sunlight. Because of the sharp underwater attenuation of downwelling sunlight, only the near-surface layer is readily supportive of net primary production. Even then, seasonal, diurnal, and weather-determined fluctuations in the intensity and duration of incident sunshine means that the spatial and temporal extent of this euphotic layer is highly variable. To operate habitually at more rarefied light levels clearly requires the provision of more, or more efficient, light-harvesting capacity. Moreover, because it is frequently the case that the depth of the wind-mixed layer extends beyond the euphotic layer, convective entrainment of phytoplankton often diminishes its aggregate exposure to the light field, at times leaving cells unable even to compensate their basic metabolic energy requirement. At the same time, the planktonic photoautotroph still needs a suite of biochemical and physiological defenses to deal with the risk of overexposure of enhanced light.

Assuming the theoretical output of 1 mol of reduced carbon for the investment of 8 mol photons of photosynthetically

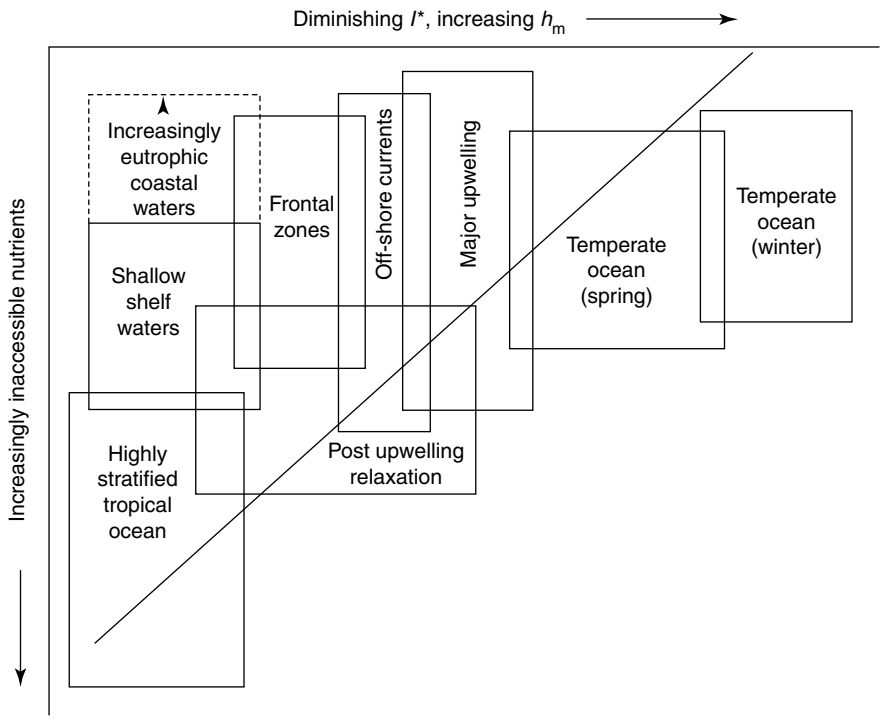


Figure 2 Two-dimensional representation of some marine habitats, which separates the mixed (light deficient) and well-stratified (nutrient-segregated) from shallow, nutrient-rich shelf waters, as well as upwellings and frontal zones. As in **Figure 1**, moving away from the top left corner, habitats become increasingly resource-stressed (working down) or processing-limited (working across). Toward the bottom right (beyond the diagonal), habitats are ultimately untenable to photosynthetic autotrophs.

active radiation harvested, taking an average 50 : 1 carbon-to-chlorophyll ratio, exponential cell-specific growth capacities of 0.04–0.26 h^{-1} may be deduced, which potentially might support biomass doublings from one every 17 h to one every 3 h.

For such rates of cell replication to be realized and sustained, not only must the light levels and the carbon dioxide supply be upheld, but also all the other chemical resources have to be available in concentrations that will also saturate the uptake demand. For instance, the widely accepted stoichiometry of Redfield (1958) leads us to suppose that for every 42 mg C incorporated, the new generation would, ideally, also require roughly 1 mg P and 7 mg N if the normal cell stoichiometry was to be preserved. Such theoretical calculations of the stoichiometric capacities are helpful to the identification of which of them is most likely to set the constraint on growth generally and, hence, that to variations in which will evoke the most sensitive changes in growth rate and supportable biomass. In many kinds of lakes and seas, the total annual area-specific loads (including recycled nutrient) of either nitrogen or phosphorus or both fall far short of the respective hypothesized saturation requirements.

It is possible to show consistent morphological and adaptive traits among phytoplankton species that are variously exploitative, disturbance tolerant, and stress tolerant. It is also possible to quantify the impacts of adaptation and to demonstrate the satisfying coherence among the form, function, and ecology of planktonic algae (Reynolds, 1988b, 1995). For example, the species that appear in "new water bodies" (from rain puddles to tidal pools), in hydraulically (flood plain lakes after the river drops back) or seasonally refreshed systems (lakes at the onset of thermal stratification) are typically characterized by an ability for rapid growth. Typical freshwater representatives include such algae as *Chlorella*, *Koliella*, *Chlamydomonas*, *Rhodomonas*, and *Chrysochromulina*,. They are unicellular or form small coenobia (generally $<10^3\ \mu m^3$), offering a surface/volume ratio ($>0.5\ \mu m^{-1}$) favorable to rapid solute exchange and nutrient assimilation. The area (e.g., of light field) projected by the algal cell mass, usually a mark of its potential efficiency as a light-gathering antenna, is equivalent to at least 6.5 m^2 (mol cell C)$^{-1}$. Reynolds (1995) proposed that these traits were indicative of an "invasive" or, in the terminology of Grime (1979), C-type life-history strategy.

The adaptations that help species to survive developing nutrient stress include the physiological flexibility to overcome deficiencies in the supply of carbon (such as the ingestion of bacteria by the facutatively phagotrophic dinoflagellates and chromulines), nitrogen ("fixing" the gas dissolved in the water, as occurs in the facultatively produced heterocytes of Nostocalean Cyanobacteria), and phosphorus (producing osmotically inactive polyphosphates). The ability to conserve assembled biomass through reduction in sedimentary and grazing losses is contributed by combining motility (usually swimming) with large size: Ease of disentrainment and self-regulated migratory ranges give access to resources in the remoter parts of the water column that may be denied to smaller species. Freshwater examples (which include larger species of

Peridinium, *Ceratium*, *Microcystis*, and other colonial, bloom-forming Cyanobacteria and, arguably, such colony-forming algae as *Uroglena*), are characteristically "large," having cells or coenobia that are $>10^4$ μm^3 in volume and often much more than 30 μm in diameter. Their consequent rather low surface/volume ratios (<0.2 μm^{-1}) leave the algae with slow rates of growth ($<8\times10^{-6}$ s^{-1} at 20 °C), sensitivity to low temperatures, and poor antennal projection (≤2.5 m^2(mol cell C)$^{-1}$). The occurrence of growing populations of these stress-tolerant – S-type, (or "acquisitive;" Reynolds, 1995) – strategists, tends to be restricted mainly to warm, well-insolated shallow lakes and epilimnia.

The phytoplankton species tolerant of, if not dependent on, near-continuous entrainment within deep or turbid mixed layers have well-developed capabilities for maintaining growth despite intermittent brief exposure to light. "Attuning" (Reynolds, 1995) strategists tolerate this scale of pelagic disturbance through the projection of a large mass-specific surface area (>8 m^2 (mol cell C)$^{-1}$ in the case of 1-mm threads of *Planktothrix agardhii* and ~30 m^2(mol cell C)$^{-1}$ in the case of an eight-celled colony of *Asterionella formosa*). They have high surface-to-volume ratios (generally >0.5 μm^{-1}) though these are not necessarily attained through small unit-size but by morphological complexity (filaments, attenuated or fenestrated coenobia, protuberances, etc.) and so continue to benefit from reasonably rapid rates of cell-specific nutrient uptake and growth. The potential for energy conversion may be enhanced by an increased cell-specific chlorophyll content and accessory pigmentation to widen the spectrum of harvestable wavelengths.

Not all phytoplankton species fit perfectly within one or other of the three categories; some show properties intermediate between them. What is interesting, however, is that the intermediacy in morphological and physiological adaptation matches well the intermediacy in their ecologies. The "space" between the invasive and the stress-tolerant acquisitive species is occupied by genera such *as Dinobryon*, *Dictyosphaerium*, *Sphaerocystis*, *Gemellicystis*, arguably *Volvox*, *Eudorina*, *Aphanizomenon*, and *Gloeotrichia*, which diminish in surface area/volume ratio and maximum growth rate but increase in their abilities to exploit and conserve nutrient resources. Freshwater algae between the invasive and the attuning poles include species of *Cyclotella*, *Scenedesmus*, and *Coelastrum*, which could be said to be increasingly large, more convoluted, and increasingly disturbance–turbidity-tolerant algae. The axis between stress- and disturbance-tolerance is occupied by relatively large, acquisitive, but self-regulating species, which can persist for months to years on stable density gradients, notably *Planktothrix rubescens* and other cyanoprokaryotes including the notorious *Cylindrospermopsis raciborskii* (Padisák *et al.*, 2003c).

Despite its celebrated species richness, the phytoplankton is usually dominated by very few genera at a time (it has been suggested that 95% of the extant standing biomass will be incorporated in no more than eight species at any one time; often it will be in rather fewer, as few as one or two; Reynolds, 1997). Yet it is well recognized that the dominance moves among different species through time and that, in a given system, the sequence will be similar from one calendar year to the next. Moreover, similar patterns may be observed in similar but often geographically remote lakes. Numerous such cycles have been described in the literature (Reynolds, 1984; Sommer *et al.*, 1986). No two years will be exactly the same, and the relative proportions of simultaneously dominant and codominant species will fluctuate.

Preliminary considerations of the types of phytoplankton characterizing the various subdivisions of the sea do reveal parallel trends between morphology and performance. For instance, as days lengthen over the open seas of the temperate latitudes, an initial typical spring bloom of such chain-forming diatoms as *Thalassiosira nordenskioldi*, *Skeletonema costatum*, and *Chaetoceros* (all manifestly having R-type, attuning characters) gives place in summer to relatively large, motile, S-like peridinians (e.g., *Scrippsiella trochoidea*) and ceratians (e.g., *Ceratium tripos*), which are often followed, in autumn, by chain-forming (*Chaetoceros* spp.) diatoms with high surface-area-to-volume ratios and antennal enhancement. In the high-nutrient, high-energy waters of coastal margins, potential dominants are generally small, usually unicellular, centric diatoms (e.g., *Cyclotella caspia* and *Thalassiosira weissflogii*), green flagellates (e.g., *Dunaliella* and *Nannochloris*), euglenoids (*Eutreptia*), and gymnodinioids (*Gymnodiniurri*); all qualify as fast-growing opportunist (C-) species. The estuarine prorocentroids are also conspicuous. Blooms of *Phaeocystis* may follow, before giving place to ceratians. In the highly stratified, tropical oceans, dominance passes quickly from diatoms to *Emiliana* and to the large *Ceratium* spp. and *Ornithocercus*. The big, buoyancy-regulating, stress-tolerant, S-species of diatom (*Ethmodiscus*) and dinoflagellate (*Pyrocystis*) are at the culmination of this sequence.

Spatial differences – among latitudes or between neritic shelf waters and the open oceanic systems – probably also attest to the functional differentiation of the phytoplankton. Areas of frontal activity and of surface upwellings of the deep oceanic circulation currents tend to combine intensive mixing with renewed resources and to select species in the C–R intermedium through much of the year: *Gyrodinium* spp., *Alexandrium tamarense*, and *Thalassiosira leptoporous* can be typical assemblage members. Away from the epicenters of the upwelling areas and where, indeed, the water is less mixed but still charged with nutrients, such taxa as *Dinophysis*, *Gonyaulax* and *Gymnodinium catenatum* provide prominent markers of the (R–S) transition between stressed and disturbed conditions.

The functional group concept in phytoplankton was initiated by Reynolds (1980), and after some modifications (Reynolds, 1984) it was published in full detail.

Two main ideas lie beneath the functional groups theory: (1) a functionally well-adapted species is likely to tolerate the constraining conditions of factor deficiency more successfully than individuals of a less well-adapted species; (2) a habitat shown typically to be constrained by light, P, C, N, or whatever, is more likely to be populated by species with the appropriate adaptations to be able to function there (this of course does not imply that those species will be there). As a consequence, the term "functional group" is sensitive to the sets of appropriate adaptive specialisms and the clusters of species that exhibit them.

The article "Toward a functional classification of the freshwater phytoplankton" published in 2002 by Reynolds

and coauthors has received a great deal of attention from phytoplankton ecologists.

Until now, 37 functional groups are described (Padisák *et al.*, 2009) and the application of the concept to monitoring purposes is also developed (Padisák *et al.*, 2006; Borics *et al.*, 2007). Though splitting species into different functional groups requires good taxonomic skills and a deep knowledge of the autoecology of single species or species groups, application of the functional group concept is being applied to an increasing extent.

Form, Function, and Selection in the Zooplankton

In terms of species success, the survival strategies, and selective biases, the ecological constraints governing the lives of zooplankton are wholly analogous to those acting on the phytoplankton. In the case of a heterotroph, additional energy is consumed in the foraging of foods – the essential requirement is that the investment of energy yields a greater return of potential energy in the organic carbon thus derived. Resource gathering carries a high energy demand: "Maintenance costs" of phagotrophs are proportionately high, not untypically accounting for some 90% of the resources consumed. Specific requirements differ substantially, as do the time-scales over which the organisms respond to environmental fluctuations.

The size range of planktonic animals embraces the flagellated protists, whose dimensions leave them substantially embedded within the viscous range of aquatic environments, through to crustaceans and young fish whose lives reach well into the turbulent range. Although the former rely on encounters of suitable food particles (organic detritus, not necessarily autochthonous, bacteria, small algae) within the same viscous neighborhood, animals living mainly within the boundaries of the turbulent world have a modicum of control over their own movements.

There is also considerable differentiation among the strategies for survival and resource exploitation. In addition, predation by planktivorous fish and invertebrates has a central role in structuring zooplankton assemblages: Although larger animals ostensibly have more foraging opportunities for a wider choice of potential foods, they are also more vulnerable to visual predators (Brooks and Dodson, 1965; Hall *et al.*, 1976). Romanovsky's (1985) analysis of life-history strategies of the planktonic Cladocera distinguished among those species that invest in high survivorship rather than in rapid juvenile growth (exemplified by *Diaphanosoma brachyurum*), those ultimately large-bodied species that sustain rapid rates of juvenile recruitment and development (typified by *Daphnia hyalina*), and those smaller-bodied, fast-reproducing inhabitants exploiting the fluctuating opportunities of temporary ponds and water bodies (including species of *Moina*). Taking particular account of the interspecific trait variability in adult cladoceran size and the negative correlations ("consistent trade-offs") between egg number and yolk investment and between larval development time and resistance to starvation, Romanovsky recognized three basic life-history strategies. He referred to these as being "patient," "violent," or "explerent." It needs little insight, however, to recognize that this imaginative perception is but a quite independent assessment of the expression of "resource-stress tolerance," "exploitative invasiveness," and "disturbance tolerance" among animals. In this way, it is possible to fit planktonic animals to a habitat template, analogous to that for the phytoplankton, on the basis of a wholly functional classification.

Zooplankton of the open seas and of larger oligotrophic lakes live in environments where the producer biomass is chronically resource limited: The potential food resources for zooplankton (planktonic algae, bacteria, and particulate organic detritus) are generally <0.1 g C m^{-3}. To be successful under persistent conditions of dilute resource, animals must blend energy-efficient foraging with controlled resource exploitation – they must adopt the "patient" strategy of resource-stress tolerance, especially with respect to the recruitment of juveniles. Investment in a small number of relatively large eggs with a relatively long development time is typical. Physical and behavioral defenses reduce the vulnerability to predation. Such traits distinguish the life histories not only of *Diaphanosoma* but also of many of the calanoid copeods. The sophisticated browsing mode of feeding that they adopt has been shown to be sufficient to sustain natural populations and support their reproduction at food concentrations in a range equivalent to 0.01–0.08 g C m^{-3} in both the sea (Huntley and Lopez, 1992) and lakes (Hart, 1996). Arguably, the ciliates (e.g., *Halteria* and *Strombidium*), which constitute a significant part of the diet of calanoids in oligotrophic lakes, must be, similarly, stress-tolerant patient strategists.

In the more productive waters of fertile lakes and rivers, estuaries, nutrient-enriched coastal-shelf waters, and oceanic upwellings, the capacity to support autotrophic biomass may, for at least some of the time, significantly exceed 0.1 C g m^{-3}. This is an approximate threshold, above which the energetic return from filter-feeding becomes steadily more favorable. Among lakes, for example, daphniids tend to become much more abundant, absolutely, and relative to calanoids. Most species of *Daphnia*, when adequately nourished, are able to grow and mature rapidly to increase egg production and to recruit the next filter-feeding generation over a few days (George and Reynolds, 1997). The 12-fold increase in the *Daphnia*-dominated community-filtration rate in 13 days, measured by Reynolds *et al.* (1982; note, aggregate filtration capacity doubles every 3–4 days), conforms to the understanding of an exploitative, or "violent," life-history strategy. Among the noncrustaceans, the ciliates of microaerophilous environments (e.g., *Metopus* and *Caenomorpha*) react in an analogous way to food abundance in media from which other sorts of consumer are mainly excluded.

Many noncrustacean members of the zooplankton take advantage of short generation times to track bursts of algal or bacterial abundance. Protists including *Difflugia*, *Coleps*, *Tintinnidium*, and *Nassula* and rotifers such as *Keratella*, *Kellicottia*, and *Brachionus* will often respond to an expansion in suitable algal foods, but this occurs before the daphniids can ascend to dominance. Their disturbance-accommodating lifestyles are well suited to the direct exploitation of small phytoplankton and organic detritus in the middle reaches of large rivers, provided travel times permit (Viroux, 1997).

Function in the Bacterioplankton

Microbial populations constitute the third integral component of the planktonic communities of lakes and seas, where they are often numerous (typically within the range 10^5–10^7 cells ml^{-1}). Microbial diversity, however, is not yet well understood, principally because the bacterial taxa are still insufficiently distinguishable to be separated routinely to species level. Given that, for most of them, the utilization of dissolved reduced-carbon substrates provides the main source of energy, even the functional approach appears to have its limitations for bacterioplankton. "Organic carbon compounds" are derived from a very wide range of the breakdown products of biogenic materials, many originating from the land, as well as a host of very unnatural anthropogenic organic substances. The ability to break down particular classes of compounds or kinds of bonds is unlikely to be shared by all bacteria. Every such process presumably requires the action of one or more discrete enzymes, and the ability to produce each enzyme requires one or more genes dedicated to its production. As the genetic complement of microbes is relatively modest, it follows that the number of separate, substrate-specific bacterial strains is likely to be high.

Modern molecular approaches are slowly sorting out "who does what" but, for the present, the diversity of the bacterioplankton is anticipated to be high.

In smaller lakes and ponds, the nature and diversity of organic carbon sources are enriched by direct transport of biogenic materials from the catchment area. However, current understanding is also having to accommodate the recognition that often the largest fraction of organic carbon present in stream and lake waters comprises plant-derived humic and fulvic acids washed in from catchment soils (Wetzel, 1995). They are particularly abundant in brown and "black" waters draining forests and peatlands. The relative size of this fraction indicates its high resistance to microbial breakdown. Refractory dissolved organic carbon from terrestrial sources break down slowly; some is rendered labile by exposure to ultraviolet radiation.

In the microaerophilous environments of the bottom sediments and in the deep water of stratified, productive lakes, anaerobic microbes are common and typically numerous; they include *Clostridium*, the sulfate-reducing *Desulfovibrio*, and the archean methanogens. Interacting gradients of light and of redox superimposed on relatively stable density gradients may also provide vertical sequences of microhabitats, each niche potentially differentiated according to the microorganisms it supports and whose physiological adaptations it most suits. To be able to carry out photosynthesis in a reducing environment favors *Rhodospirillum*; but to be able to use sulfide or sulfur, as photosynthetic electron donor, can bias the selection in favor of Chromatiaceae and Chlorobiaceae. The involvement of specialist chemolithotrophs in nitrification (e.g., *Nitrosomonas* and *Nitrobacter*) and in the oxidation of sulfur (e.g., *Thiobacillus* and *Beggiatoa*) and iron compounds (e.g., *Ochrobium*) further adds to the microbial diversity of freshwaters.

In the open waters of the sea, most bacteria are aerobic – relatively few anaerobes are ever found in the surface waters. Many of the same genera represented in freshwaters include marine species: *Pseudomonas* and *Vibrio* are often found to dominate and species of *Flavobacterium*, *Alcaligenes*, and *Cytophaga* are found in high numbers (Atlas and Bartha, 1993). Sediments receiving organic inputs are bacteria-rich; moreover, where these fall anaerobic, marine sulfate-reducers and methanogens are present in substantial numbers.

In both marine and freshwaters, the vital contribution of the bacterioplankton to ecosystem function arises not only from the mineralization of organic carbon products but also from the potential remobilization and renewed bioavailability of nitrogen, phosphorus, and the metals required in the assembly of new biomass. The largely unexplored world of deep-sea hydrothermal vents may represent a substantial addition to marine bacterial biodiversity (Martin *et al.*, 2008).

Temporal Patterns in the Organization and Diversity of Planktonic Communities

Diversity is a concept drawn from communication theory (Shannon, 1948), referring to the amount of discrete information in a particular location. In communities, it could refer to the different fragments of genetic information. A diverse community is one with many different kinds of genes (species) present simultaneously. This biotic diversity (H'') can be quantified with great precision, using the specialized Shannon–Weaver function:

$$H'' = -\Sigma b_i/B \log_2(b_i/B)$$

Diversity increases with the number of species (s) each contributing biomass (b_i) to the total, B, in the sample of lake or seawater. The more species present, the greater is the diversity, until $H''_{max} = \log_2 s$. In practice, it is necessary to set a bar on the efficiency of search for the rarer species (more diligent searchers would always score higher diversities). Nevertheless, the evenness of the interspecific distribution, $E = H''/H''_{max}$, is itself a useful measure.

Local species richness (the sample, the pond, and the lake) may run to a few tens to hundreds of phytoplankton and, if the phagotrophic protists are included, perhaps 100 or so species of zooplankton. Regular sampling may raise this total considerably, as seasonal changes are encountered and the probability of encountering rare species is increased (Padisák, 1992; Padisák *et al.*, 2010). "Rare" in this context may mean one individual per liter, but there could still be a billion such individuals in a small, 10 ha lake. As observed earlier, most ($\geq$95%) of the total biomass of phytoplankton and that of zooplankton will each be invested in eight or fewer species. Most of the consideration of evenness concerns how the total biomass is shared among those eight. Thus, whether local species composition is species rich or otherwise, the perception of its diversity is greatly affected by the relative evenness of the most common species.

The essential step in logic that separates the niche-diversification and the persistent-disturbance theories of species richness rests on an acceptance that the supportive capacity is not always filled; conversely, although the capacity of the resources is always finite, the underexploited resource base is, by definition, *not* "limiting." Phytoplankton

concentration is able to increase at all because, in reality, there are times when nutrient resources are available to support the additional biomass and unharvested light energy to fuel its assembly. The spring bloom, at least initially, is the response of the producer community to an expanding capacity of the resource and energy bases. Later in the year, the resource supply may well fail to meet the biological demand and severe constraints are then imposed on its further growth and maintenance. Alternatively, it is left to shortening days and convective mixing in the autumn months to erode the supportive capacity of the insolation flux.

Rapidity of growth and reproductive efficiency are crucial to the exploitative outcome: Having a large starting inoculum and, especially, a high yield/resource conversion potential is vital to the fitness of colonist and invasive species. If we judge it to be "outcompeting" the slow-growing species present, we are giving a second nuance to the word "competition." By exploiting the structure of the water column to the capacity of its resources and by continuing to grow for longer, often in the face of intensifying resource shortage, the S-type adaptive strategy allows the organisms to achieve a larger, climactic biomass than that achieved by any of the earlier invasive species. The dominance can become total, other species are effectively excluded by the ultimately superior competitor, and the Shannon diversity falls to a minimum.

Diversity within Habitats

For the geometric representation of environmental variability and of its impact on species selection, it is useful to bring the concept of fluctuation tracking back to the habitat template (of **Figure 3(a)**), because it is easy to relate to the environments it seeks to represent (**Figures 1** and **2**) and because sufficient preliminary knowledge exists about how the tracking thus represented actually selects for the preferred traits of species, at least of the freshwater phytoplankton (**Figure 3(b)**).

The organizational trends imposed by disturbance and stress press the selective bias toward the R or S apices of the triangular template (shown in **Figure 4(a)**). We have already seen that seasonal trends track through the template matrix in broad, predictable ways (summarized as S-ward and R-ward trends in **Figure 4(b)**), but substantial within-season variability will lead to almost chaotic short-term time tracks, such as the fragment included in **Figure 4(c)**.

Our deduction is again that the effect of environmental variability is to move any selective advantage among species at a faster rate than opportunities narrow or that competition can forge a low-diversity monoculture dominated by a single well-adapted species. We know that this theoretical outcome, correctly anticipated by Hutchinson (1961), certainly is achievable and good descriptions of local, competitively excluded community structures are described in the literature. These cover the sort of arrested "plagioclimaces" of overwhelming *Microcystis* dominance of tropical eutrophic Lake George (Ganf and Viner, 1973) year-round and *Planktothrix* dominance of exposed, continuously mixed hypertrophic polder lakes in the Netherlands (first described by Berger, 1975). At one extremity, diversity may drop to zero, as was

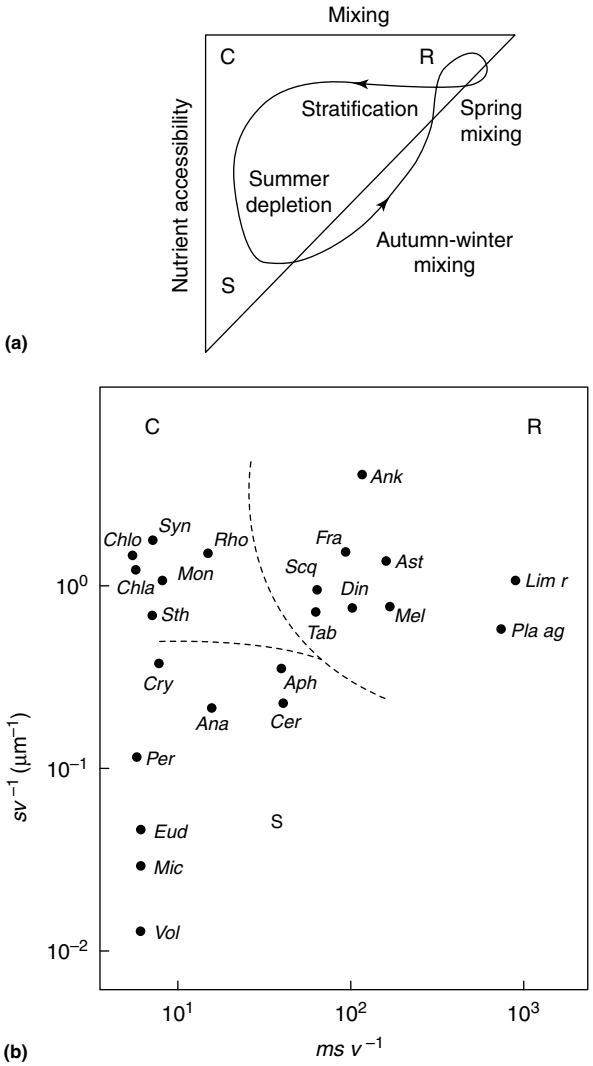


Figure 3 (a) How growth conditions in a monomictic deep lake (mixed in winter) track through a calendar year, selecting sequentially for a low-light tolerant spring bloom (of R species), then opportunist C species after the onset of stratification, leading to the biologically enhanced segregation of nutrient resources tolerated by S species, before autumnal downmixing again forces conditions back toward the top right: Nutrients may be further increased (by inflow) in winter, when primary production is still increasingly light limited. (b) Some of the algal species that might be selected are shown in an analogous matrix that is based purely on shape- and size-characters of the algae concerned (sv^{-1} vs. msv^{-1}, where s is the surface area of the alga in μm^2, v is its volume in μm^3, and m is the maximum length dimension, in μm). A strong correlation exists between the morphologies of these algae and their functional ecologies. The algae shown are represented as follows: *Ana* = *Anabaena flos-aquae*, *Ank* = *Ankistrodesmus falcatus*, *Aph* = *Aphanizomenon flos-aquae*, *Ast* = *Asterionella formosa*, *Cer* = *Ceratium hirundinella*, *Chla* = *Chlamdomonas*, *Chlo* = *Chlorella*, *Cry* = *Cryptomonas ovata*, *Din* = *Dinobryon divergens*, *Eud* = *Eudorina unicocca*, *Fra* = *Fragilaria crotonensis*, *Lim R* = *Limnothrix redekei*, *Mel* = *Aulacoeira subarctica*, *Mic* = *Microcystis aeruginosa*, *Mon* = *Monodus*, *Pla AG* = *Planktothrix agardhii*, *Per* = *Peridinium cinctum*, *Rho* = *Rhodomonas pusilla*, *Scq* = *Scenedesmus quadricauda*, *Sth* = *Stephanodiscus hantzschii*, *Syn* = *Synecococcus*, *Tab* = *Tabellaria flocculosa*, *Vol* = *Volvox aureus*.

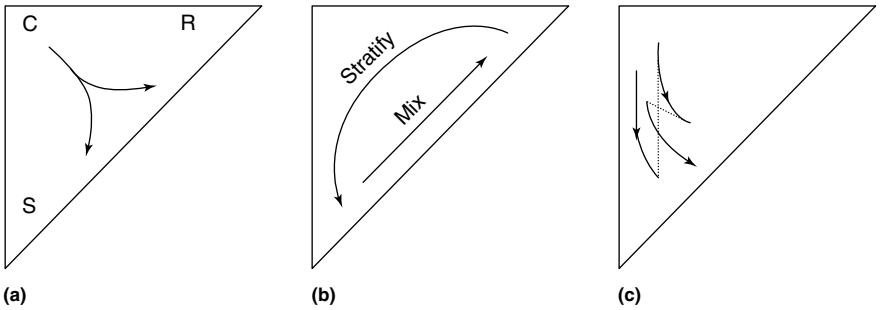


Figure 4 (a) From a starting condition, near the top left (c) corner of the triangle, environmental change moves to select increasingly for (S-) species tolerant of resource stress or for (R-) species tolerant of energy limitation (processing constraints); (b) major environmental restructuring alters the trajectory, changing the selection significantly; (c) under more frequent and perhaps less severe forcing, trajectories are redirected erratically, delaying progress to the competitive exclusion signified by the R and S apices.

observed in a small Hungarian fishpond when *Cylindrospermopsis* reached overwhelming dominance (Borics *et al.*, 2000). No less impressive is the striking commonality of organization of vertically segregated layers (or "plates") of algal and microbial producers on the stable physicochemical gradients in permanently ice-covered lakes in Antarctica (e.g., Vincent, 1981), in tropical forest lakes (Reynolds *et al.*, 1983), and in midlatitude karstic dolines (Vicente and Miracle, 1988). However, such steady states do not occur frequently. In contrast, the great majority of assemblages that have been sampled, identified, and described in the literature are, simply, weakly organized. There need not be any paradox in this once it is appreciated that in such accumulating, mid-successional assemblages most of the specific populations are not in any steady state but rather in a state of flux, either increasing or decreasing in response to events in the recent past. Thus, most so-called communities comprise populations which, at the given point in time, are in some stage between being dismantled and reassembled.

The crucial questions about local diversity in the plankton should be "How long has it been since a restructuring was initiated by external forcing of a magnitude sufficient to create a disturbance?" and "At what rate is the restructuring taking place?" The field-scale experiments carried out in the large limnetic enclosures in Blelham Tarn, England, pointed to a progress to steady-state monocultural dominance requiring from 12 to 16 generations of the eventual dominant (Reynolds, 1988a). The time required for this could be as few as 35–60 days, provided growth rates could be sustained by water temperatures $\geq 20\,^{\circ}C$ and a supply of carbon and nutrients of a capacity to support the last population doubling. Accordingly, limitations on growth rate imposed by low temperature, slow carbon renewal, or nutrient depletion extend the generation times reciprocally, delaying the onset of the climactic condition and slowing the exclusion of diversity. More likely, the intervention of external forcing (e.g., strong winds and heavy rain) would interrupt the successional maturation at an earlier date, and if of sufficient magnitude, select for alternative species with a new target outcome. It was found consistently that the community response, represented by two to four divisions of the newly selected species simultaneously with the dieback of the erstwhile-selected species, took some 5–15 days before it was clearly manifested. It was deduced that a disturbance frequency of this magnitude (two to four generations) is sufficient to maintain an optimal local species diversity (Reynolds, 1988a).

Analyses of temporal diversity fluctuations in small lakes in locations as far apart as Canada (Trimbee and Harris, 1983) and central Hungary (Padisák, 1993) confirm the validity of this deduction, showing diversities in phytoplankton composition peaking within 11 days of a recognized physical stimulus. These and numerous other case studies have been brought together to demonstrate this planktonic validation of Connell's (1978) intermediate disturbance hypothesis. The essence of Connell's hypothesis is that frequent disturbance excludes all but fast-maturing species; very infrequent disturbance allows competitive exclusion to reduce diversity; therefore, maximum diversity is maintained at intermediate frequencies. At both high and low frequencies of forcing, the numbers of intermediate species are reduced so that the inocula are no longer readily available to take advantage of conditions under which their growth might be favored. In contrast, local stocks of mid-successional species are enhanced by the sorts of opportunities that might be provided every two or three generations by the (intermediate) rejuvenation of appropriate growing conditions and with a marked alleviation of severe interspecific competition in the early postdisturbance sequel.

Species Richness Among Habitats

It can be seen that the species representation contributing the local diversity remains heavily biased toward those that have been well represented in the recent past, and which, of those whose growth should be favored by the onset of the appropriate environmental characteristics, remain potentially capable of seeding the largest inocula from standing stocks of vegetative cells and resting propagules. With the same kinds of environmental variability affecting Shannon-type species diversity in each tangible locality, there is still no clear explanation as to why the collective representation of planktonic organisms is so relatively species rich. We need a proposition for separate species representation in separated localities which recognizes that communication between localities is sufficient to maintain the high level of apparent cosmopolitanism among planktonic organisms. In other words, we need to be aware of the roles of perennation and dispersal in relation to the maintenance of biodiversity.

The survival of any species whose range of habitat suitability (as defined in the section Explaining Species Diversity) is discontinuous, in time as well as space, involves the separate development of separated populations, but with a measure of regenerative connection and gene renewal that resists their permanent divergence. The metapopulation ecology of plankton is not, formally, a well-rehearsed topic, though there have been numerous studies to contribute a general appreciation of the principles. Although the existence of liquid water on the planet has a very long geological history, individual bodies of freshwaters are extremely transient. With the exception of the basins formed by tectonic movements, a majority of these is less than 20,000 years in age. The idea of lakes as islands in a terrestrial sea is an easy one to assimilate. The actual oceans are physically and temporally contiguous, but the patterns of global circulation allow significant habitat differentiation to be maintained at the scales at which planktonic organisms live their lives. The principles of island biogeography propounded by MacArthur and Wilson (1967) provide a good model for plankton ecology.

The temporal discontinuities are most appropriately bridged by the production of resistant, resting propagules. Among the freshwater phytoplankton, almost all the major groups produce some kind of physiological resting stage, if not a discrete resting spore, cyst or akinete, some of which can remain dormant for many years and still be fully viable. Eggs may fulfill a similar role for zooplankton, maintaining banks of inocula and pending the restoration of conditions favorable to growth and recruitment of successive generations (Hairston, 1996). Produced in adequate numbers, such dormant life-history stages provide a significant survival "hedge" through periods of hostile conditions (low temperature, low light, poor resources or food availability, drought, and so on). The inocula potentially supplied by subsequent germination and egg hatching provide a powerful mechanism for maintaining diversity in seasonally fluctuating environments.

In many cases, the propagules facilitate spatial transfers as well. Dispersal in water droplets, or in dust, or in or on the bodies of animals have been shown to be effective pathways of planktonic organisms (review of Kristiansen, 1996) and to have been directly implicated in the establishment of populations in new or isolated bodies of water (Maguire, 1963, 1977). Planktonic species are not equally amenable to dispersal; for instance, effective perennation and dispersion are essential properties of species of temporary or temporally variable habitats. In general, invasiveness of propagules is dependent on a raft of species-specific features such as their sizes, their resistance to desiccation, and the numbers and frequencies with which they are produced. The probability (p) of a given species (A) being able to establish itself in another water body will be determined as a function, partly of its relevant species-specific traits (T_A) and partly of the problems posed to all would-be invaders of the distance (d) and size (a) of the new (or "target") site from the existing (or "source" MacArthur and Wilson, 1967) population:

$$p = f(1/d, a, T_A)$$

An important converse of this relationship also manifests itself in relation to the ongoing suitability or effectiveness of source and target sites as a habitat for the given species. Either because the habitat changes in consequence of autogenic properties or because species A is prevented from completing its perennation in that habitat, through the intervention of a facultative predator or pathogen, it is ultimately likely that the survival of the species in that location is at risk. Repeated at several sites, this process would threaten the survival of the species in its entirety and bring about a diminution in total species richness (i.e., the biodiversity). The resistance to that comes, quite literally, from the patch dynamics of habitat distribution and their temporal suitability to species A. Thus, the status of species A is a function of the number of possible habitat patches (N), the number that is occupied (O), the rate at which they become excluded there from (e), and the probability (p) of colonizing the (N–O) unoccupied habitats. The rate of expansion (or loss) of the metapopulation in time ($\mathrm{d}O/\mathrm{d}t$) may then be symbolized:

$$(\mathrm{d}O/\mathrm{d}t) = pO(N - O) - cO$$

For the more familiar and (supposedly) more cosmopolitan planktonic forms, relative ease of dispersal is, manifestly, the dominant factor upholding the widespread occupancy among suitable habitats. For the majority of known species that are relatively rare (that is, there is a relatively low occupancy among the total number of patches), effective dispersal may be resisted by specific traits (T_A) or simply by low numbers. In this case, survival may be considered more tenuous. However, from the limited investigative evidence available, many of these rarer species have been numerous in the past and occasionally continue to perform relatively well on the inertia of persistent, viable propagules. While the length of this "ecological memory" (Padisák, 1992; Padisák *et al.*, 2010) is not certainly quantified, we may be aware of the fact that it is generally the rarer species that contribute most to the total biodiversity of planktonic organisms. Although the metabolically significant planktonic biomass relies on the productivity of a relatively small number of species in a series of variable and renewable habitats, the majority of species is limited to habitats in which temporal variability is less extreme or exclusion is long protracted.

Conclusions and Implications

The feature distinguishing plankton-based ecosystems from terrestrial ones is that the relevant temporal scales are mostly much shorter in open waters than on the land. Not only are the mechanisms of planktonic biodiversity conveniently observed, but also the study of planktonic systems may be rewarded with insights into current issues about the importance and the protection of biodiversity.

Of the many propositions that have been advanced about the role of diversity in upholding ecosystem processes, current attention focuses mainly on four broad postulates (Lawton, 1997). In various ways, they relate to Darwin's (1859) proposition that a large number of species imparts a higher level of functional stability than does a small one. Thus, to lose species inevitably impairs the functional integrity of the ecosystem. The "redundant species hypothesis" counters that, because

species are not equally represented, some are contributing much more than others to ecosystem function. Indeed, some "keystone species" (Paine, 1969) drive the main functions and have a much stronger influence than "passenger" species (Walker, 1992). Then, logic suggests, a minimal diversity is essential to adequate ecosystem functioning and most of the species are really redundant in their roles (Lawton and Brown, 1993; Walker, 1992). Almost diametrically opposed to this view is the "rivet hypothesis" (Ehrlich and Ehrlich, 1981), which accords to each species an essential contributory role, and which, if lost from the whole, like rivets lost from the structure of an aircraft, quickly leads to serious functional impairment and failure. A third hypothesis takes a much looser view, suggesting that function is modified by changes in the richness of species composition but in unpredictable ways ("idiosyncratic"; Lawton, 1994) because the contributions of individual species to system function are unequal. The fourth possibility is the null hypothesis that ecosystems are insensitive to changes in species composition. This seems increasingly to be implausible and will not be discussed.

One of the important lessons from the plankton scale is that a distinction should be made between structural stability and functional stability. Taking the basic function of the plankton to be the material cycling in open water, driven by solar energy invested in carbon bonds, the photoautotrophic, heterotrophic, and phagotrophic roles are fulfilled, respectively, by any or all of a large number of species of phytoplankton, bacterioplankton, and zooplankton. Despite having vastly different and frequently changing species compositions, studies of the productivity of planktonic systems demonstrate a remarkable level of functional coherence (Schindler, 1990). In this way, pelagic fish may continue to feed planktivorously and to respire photosynthetically generated oxygen almost without reference to the planktonic structure, as long as similar functions are maintained.

At the level of planktonic communities, a decline in the population of a dominant species will carry fewer implications if a second species is poised to substitute quickly. Frost *et al.* (1995) made the case that species complementarity is made for functional compensation of the smaller scales of structural variability. The case is analogous to the arguments about the flexibility of communicating information through networks rather than along a single pathway (Pahl-Wostl, 1990), or what has been called the "World Wide Web" explanation of biodiversity (Reynolds, 1997). This recognizes the idea that some species may contribute less to overall ecosystem function than other, more functionally creative keystone species (following Paine, 1969) or "ecosystem engineers" (Lawton, 1997), and providing opportunities to others may also serve to promote and maintain the species richness. It also values the role of the reserve (richness) of complementary species able to fulfill the compensatory function.

What clues does planktonic diversity give us about the approaches to uphold species richness in other ecosystems? The advice, "Keep up the disturbance," is a little trite and somewhat counterintuitive to traditional conservation attitudes. Only at the global scale of aquatic systems is it possible to conceive the simultaneous existence of sufficient a range of habitats to accommodate the entirety of species. No single location, no matter how variable, offers habitat or refuge for more than a minority representation of the aquatic biota. Aquatic biodiversity has to be pursued at the scale of fluvial catchments and regional seas – every lake, pond, evolving flood plain, river estuary, coastal shelf, and ocean circulation is but a piece in the biospheric mosaic. The essential general aim should be to ensure that suitable habitat elements persist, each in a range of developmental states. Local extinctions are resisted by good between-locality communication of propagules. In this sense, managing biodiversity has to adopt the philosophies of patch dynamics, taking full account of the longevities and the invasiveness of species-specific populations in relation to the externally influenced availability and renewal of suitable patches.

See also: Estuarine Ecosystems. Lake and Pond Ecosystems. Ocean Ecosystems. River Ecosystems

References

Atlas RM and Bartha R (1993) *Microbial Ecology – Fundamentals and Applications*, 3rd edn. Redwood City: Benjamin Cummings.

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, and Thingstad F (1983) The ecological role of water-column microbes in the sea. *Marine Ecology, Progress Series* 10: 257–263.

Berger C (1975) Occurrence of oscillatoria agardhii Gomont in some shallow eutrophic lakes. *Verhandlungen der Internationale Vereinigung für Theoretische und Angewandte Limnolgie* 26: 97–113.

Borics G, Grigorszky I, Szabó S, and Padisák J (2000) Phytoplankton associations under changing pattern of bottom-up vs. top-down control in a small hypertrophic fishpond in East Hungary. *Hydrobiologia* 424: 79–90.

Borics G, Várbíró G, Grigorszky I, Krasznai E, Szabó S, and Kiss KT (2007) A new evaluation technique of potamoplankton for the assessemnt of the ecological status of rivers. Large Rivers, 17. *Archiv für Hydrobiologie* supplement 161: 465–486.

Brooks JL and Dodson SI (1965) Predation, body size and composition of the plankton. *Science* 150: 28–35.

Callieri C (2008) Picophytoplankton in freshwater ecosystems: The importance of small-sized phototrophs. *Freshwater Reviews* 1: 1–28.

Connell JH (1978) Diversity in tropical rain forests and coral reefs. *Science* 199: 1302–1310.

Darwin CR (1859) *The Origin of Species by Natural Selection.* London: Penguin Books.

Dumont H (1999) Effects of reservoirs on faunal richness and dispersal. In: Straškraba M and Tundisi JG (eds.) *Theoretical Reservoir Ecology*, pp. 477–491. São Carlos: IIE.

Ehrlich PR and Ehrlich AH (1981) Extinction. *The Causes and Consequences of the Disappearance of Species.* New York: Random House.

Elliott JM (1996) British freshwater Megaloptera and Neuroptera. *Scientific Publications of the Freshwater Biological Association* 54: 70.

Fenchel T (2008) The microbial loop – 25 years later. *Journal of Experimental Marine Biology and Ecology* 366: 99–103.

Frost TM, Carpenter SR, Ives AR, and Kratz TR (1995) Species compensation and complementarity in ecosystem function. In: Jones CG and Lawton JH (eds.) *Linking Species and Ecosystems*, pp. 224–239. New York: Chapman and Hall.

Ganf GG and Viner AB (1973) Ecological stability in a shallow equatorial lake (Lake George, Uganda). *Proceedings of the Royal Society of London B* 184: 321–346.

George DG and Reynolds CS (1997) Zooplankton–phytoplankton interactions: The case for refining methods, measurements and models. *Aquatic Ecology* 31: 59–71.

Grime JP (1979) *Plant Strategies and Vegetation Processes.* Chichester, UK: Wiley Interscience.

Hairston Jr NG (1996) Zooplankton egg banks as biotic reservoirs in changing environments. *Limnology and Oceanography* 41: 987–1092.

Hall DJ, Threlkeld ST, Burns CW, and Crowley PH (1976) The size-efficiency hypothesis and the structure of zooplankton communities. *Annual Review of Ecology and Systematics* 7: 177–208.
Hart RC (1996) Naupliar and copepodite growth and survival of two freshwater calanoids at various food levels: Demographic contrasts, similarities and food needs. *Limnology and Oceanography* 41: 648–658.
Havel JE (2009) Cladocera. In: Likens Gene E (ed.) *Encyclopedia of Inland Waters,* vol. 3. pp. 611–622. Oxford: Elsevier.
Huntley ME and Lopez MDG (1992) Temperature-dependent production of marine copepods: A global synthesis. *American Naturalist* 140: 201–242.
Hutchinson GE (1961) The paradox of the plankton. *American Naturalist* 95: 137–147.
Kagami M, de Bruin A, Ibelings BW, and van Donk E (2007) Parasitic chytrids: Their effects on phytoplankton communities and food-web dynamics. *Hydrobiologia* 578: 113–129.http://dx.doi.org/10.1007/s10750-006-0438-z.
Krienitz L (2009) Algae. In: Likens Gene E (ed.) *Encyclopedia of Inland Waters,* vol. 1. pp. 103–113. Oxford: Elsevier.
Kristiansen J (1996) Dispersal of freshwater algae: A review. *Hydrobiologia* 336: 151–157.
Lawton JH (1994) What do species do in ecosystems? *Oikos* 71: 367–374.
Lawton JH (1997) The role of species in ecosystems: Aspects of ecological complexity and biological diversity. In: Abe T, Levin SR, and Higashi M (eds.) *Ecological Perspectives of Biodiversity*, pp. 215–228. New York: Springer.
Lawton JH and Brown VK (1993) Redundancy in ecosystems. In: Schulze ED and Mooney HA (eds.) *Biodiversity and Ecosystem Function*, pp. 255–270. New York: Springer University Press.
Lewis Jr WM (1973) The thermal regime of lake Lanao (Philippines) and its theoretical implications for tropical lakes. *Limnology and Oceanography* 18: 200–217.
Luo W, Bock C, Li H, Padisák J, and Krienitz L (2011) Molecular and microscopic diversity of planktonic eukaryotes in the oligotrophic Lake Stechlin (Germany). *Hydrobiologia* 661: 133–143.http://dx.doi.org/10.1007/s10750-010-0510-6.
MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Princeton: Princeton University Press.
Maguire B (1963) The passive dispersal of small aquatic organisms and their colonization of isolated bodies of water. *Ecological Monographs* 33: 161–185.
Maguire B (1977) Community structure of protozoans and algae with particular emphasis on recently colonized bodies of water. In: Cairns J (ed.) *Aquatic Microbial Communities*, pp. 355–397. New York: Garland.
Margalef R (1978) Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica Acta* 1: 493–509.
Martin W, Baross J, Kelley D, and Russell MJ (2008) Hydrothermal vents and the origin of life. *Nature Reviews Microbiology* 6: 805–814.
Naselli-Flores L, Padisák J, Dokulil MT, and Chorus I (2003) Equilibrium/steady-state concept in phytoplankton ecology. *Hydrobiologia* 502: 395–403.
Padisák J (1992) Seasonal succession of phytoplankton in a large, shallow lake (Balaton, Hungary): A dynamic approach to ecological memory, its possible role and mechanisms. *Journal of Ecology* 80: 217–230.
Padisák J (1993) The influence of different disturbance frequencies on the species richness, diversity and equability of phytoplankton in shallow lakes. *Hydrobiologia* 249: 135–156.
Padisák J (1998) Sudden and gradual responses of phytoplankton to global climate change: Case studies from two large, shallow lakes (Balaton, Hungary and the Neusiedlersee Austria/Hungary). In: George DG, Jones JG, Puncochar P, Reynolds CS, and Sutcliffe DW (eds.) *Management of Lakes and Reservoirs During Global Change*, pp. 111–125. Dordrecht, Boston, London: Kluwer Academic Publisher.
Padisák J, Soróczki-Pintér É, and Rezner Zs (2003a) Sinking properties of some phytoplankton shapes and relation of form resistance to morphological diversity of plankton – an experimental study. *Hydrobiologia* 500: 243–257.
Padisák J, Soróczki-Pintér É, and Rezner Zs (2003b) Erratum. *Hydrobiologia* 501: 219.
Padisák J, Barbosa FAR, Koschel R, and Krienitz L (2003c) Deep layer cyanoprokaryota maxima are constitutional features of lakes: examples from temperate and tropical regions. *Archiv für Hydrobiologie, Special Issues Advances in Limnology* 58: 175–199.
Padisák J, Grigorszky I, Borics G, and Soróczki-Pintér É (2006) Use of phytoplankton assemblages for monitoring ecological status of lakes within the Water Framework Directive: The assemblage index. *Hydrobiologia* 553: 1–14.
Padisák J (2009) The Phycogeography of Freshwater Algae. In: Likens Gene E (ed.) *Encyclopedia of Inland Waters,* vol. 1. pp. 219–223. Oxford: Elsevier.
Padisák J, Crossetti LO, and Naselli-Flores L (2009) Use and misuse in the application of the phytoplankton functional classification: A critical review with updates. *Hydrobiologia* 621: 1–19.
Padisák J, Hajnal É, Krienitz L, Lakner J, and Üveges V (2010) Rarity, ecological memory, rate of floral change in phytoplankton – and the mystery of the Red Cock. *Hydrobiologia* 653: 45–67.http://dx.doi.org/10.1007/s10750-010-0344-2.
Pahl-Wostl C (1990) Temporal organisation: A new perspective on the ecological network. *Oikos* 58: 293–305.
Paine RT (1969) A note on trophic complexity and community stability. *American Naturalist* 103: 91–93.
Pourriot R (1977) Food and feeding habits of Rotifera. *Ergebnisse der Limnologie* 8: 243–260.
Redfield AC (1958) The biological control of chemical factors in the environment. *American Scientist* 46: 205–221.
Reynolds CS (1980) Phytoplankton assemblages and their periodicity in stratifying lake systems. *Holarctic Ecol.ogy* 3: 141–159.
Reynolds CS (1984) Phytoplankton periodicity: The interaction of form, function and environmental variability. *Freshwater Biology* 14: 111–142.
Reynolds CS (1988a) The concept of ecological succession applied to the seasonal periodicity of freshwater phytoplankton. *Verhandlungen der internationale Vereinigung für theoretische und angewadte Limnologie* 22: 683–691.
Reynolds CS (1988b) Functional morphology and the adaptive strategies of freshwater phytoplankton. In: Sandgren CD (ed.) *Growth and Reproductive Strategies of Freshwater Phytoplankton*, pp. 388–433. New York: Cambridge University Press.
Reynolds CS (1995) Successional change in the planktonic vegetation: Species, structures, scales. In: Joint I (ed.) *The Molecular Ecology of Aquatic Microbes*, pp. 115–132. Berlin: Springer-Verlag.
Reynolds CS (1996) The plant life of the pelagic. *Verhandlungen der internationale Vereinigung für theoretische und angewandte Limnolgie* 26: 97–113.
Reynolds CS (1997) *Vegetation Processes in the Pelagic.* Oldendorf: ECI.
Reynolds CS, Thompson JM, Ferguson AJD, and Wiseman SW (1982) Loss processes in the population dynamics of phytoplankton maintained in closed limnetic systems. *Journal of Plankton Research* 4: 561–600.
Reynolds CS, Tundisi JG, and Hino K (1983) Observations on a metalimnetic Lyngbya population in a stably stratified tropical lake (Lagoa Carioca, Eastern Brasil). *Archiv für Hydrobiologie* 97: 7–17.
Yu E (1985) Romanovsky, Food limitation and life-history strategies in cladoceran crustaceans. *Ergebnisse der Limnologie* 21: 363–372.
Rudstam LG (2009) Other Zooplankton. In: Likens Gene E (ed.) *Encyclopedia of Inland Waters,* vol. 3. pp. 667–677. Oxford: Elsevier.
Schindler DW (1990) Experimental perturbations of whole lakes as tests of hypotheses concerning ecosytem structure and function. *Oikos* 57: 25–41.
Shannon CE (1948) A mathematical theory of communication. *Bell Systems Technical Journal* 27: 623–656.
Sournia A, Chrétiennot-Dinet MJ, and Ricard M (1991) Marine plankton: How many species in the world oceans? *Journal of Plankton Research* 13: 1093–1099.
Souza MBG, Barros CFA, Barbosa FAR, Hajnal É, and Padisák J (2008) The role of atelomixis in phytoplankton assemblages' replacement in Dom Helvécio Lake, South-East Brazil. *Hydrobiologa* 607: 211–224.
Sommer U, Gliwicz ZM, Lampert W, and Duncan A (1986) The PEG model of seasonal succession of planktonic events in fresh waters. *Archiv für Hydrobiologie* 106: 433–471.
Talling JF (1951) The element of chance in pond populations. *The Naturalist* 1951: 157–170.
Talling J and Lemoalle J (1998) *Ecological Dynamics of Tropical Inland Waters.* Cambridge: Cambridge University Press.
Thierstein HR (1989) Inventory of palaeproductivity records. In: Berger WH, Smetacek VS, and Wefer G (eds.) *Productivity of the Ocean Present and Past*, pp. 355–375. Chichester, UK: John Wiley and Sons.
Trimbee AM and Harris GP (1983) Use of time-series analysis to demonstrate advection rates of different variables in a small lake. *Journal of Plankton Research* 5: 819–833.
Tyler PA (1996) Endemism in freshwater algae. *Hydrobiologia* 336: 127–135.
Vicente E and Miracle MR (1988) Physicochemical and microbial stratification in a meromictic karstic lake of Spain. *Verhandlungen der internationale Vereinigung für theoretische und angewandte Limnologie* 23: 522–529.
Vincent WF (1981) Production strategies in Antarctic inland waters: Phyto-plankton ecophysiology in a permanently ice-covered lake. *Ecology* 62: 1215–1224.
Viroux L (1997) Zooplankton development in two large lowland rivers, the Moselle (France) and the Meuse (Belgium) in 1993. *Journal of Plankton Research* 19: 1743–1762.

Walker BH (1992) Biodiversity and ecological redundancy. *Biological Conservation* 6: 18–23.

Wallace RL and Smith HA (2009) Rotifera. In: Likens Gene E (ed.) *Encyclopedia of Inland Waters,* vol. 3. pp. 689–703. Oxford: Elsevier.

Waterbury J, Watson S, Guillard RRL, and Brand L (1979) Widespread occurrence of a unicellular, marine, planktonic cyanobacterium. *Nature* 277: 293–294.

Wetzel RG (1995) Death, detritus and energy flow in aquatic ecosystems. *Freshwater Biology* 33: 83–89.

Williamson CE and Reid JW (2009) Copepoda. In: Likens Gene E (ed.) *Encyclopedia of Inland Waters,* vol. 3. pp. 633–642. Oxford: Elsevier.

Plant–Animal Interactions

Ellen L Simms, University of California, Berkeley, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allee effect For population size to be regulated, they must exhibit negative density dependence. That is, the population growth rate must decline as the population gets larger. However, under certain circumstances, some populations exhibit positive density dependence. This phenomenon, in which population growth rate increases as the population gets larger, is called an Allee effect. An Allee effect may generate a critical minimum population size, below which extinction will occur.

Angiosperm Flowering plants. A lineage characterized by flowers, seeds enclosed in carpels, specialized conducting elements in the phloem (sieve tube members) and xylem (vessels), presence of endosperm, double fertilization, and tectate pollen.

Diaspore A plant part distributed by dispersal, regardless of its developmental and morphological origins. A diaspore may be a naked seed, a seed enclosed in a fruit, or many seeds enclosed in a fruit. It may also mean bulbs or lengths of rhizomes. A good synonym is propagule.

Fitness Relative contribution of offspring to the next generation. An individual, genotype, or phenotype whose progeny constitutes a large proportion of the succeeding generation has high fitness.

Frugivores Animals that eat fleshy fruits.

Granivores Animals that eat seeds or achenes (grass fruits).

Herbivores Animals that eat plants. Usually excludes instances in which a single animal eats an entire plant, which is categorized as predation.

Phylogeny The evolutionary relationships among taxa (groups of related organisms), often portrayed with some kind of branching diagram, with branches representing speciation events. Typically, the true phylogeny of a group is hidden deep in the past, and evolutionary biologists must infer relationships statistically from DNA sequences and traits of modern organisms and, sometimes, traits of fossils.

Symbiosis Interaction in which two organisms live in close proximity. Symbiosis can be antagonistic or mutualistic. Often, the larger individual is called the *host*, and its inhabitant is called the *guest*.

Trophic level Position of a species in the food web (Figure 1). Plants – autotrophs which convert solar energy to chemical energy and utilize mineral nutrients – constitute the first trophic level. Primary consumers – animals that feed on living plants – constitute the second trophic level. The third trophic level is composed of secondary consumers – animals that feed on primary consumers. Predation, parasitism, grazing, and herbivory are intertrophic level interactions; competition occurs among species in the same trophic level.

Types of Interactions

Biotic interactions can be categorized by their effects on the interacting parties (Table 1). An interaction may not affect a species, or may be beneficial or detrimental. Antagonistic interactions negatively influence one or both species. Some interactions are clearly antagonistic: When a vole consumes an oak seedling, the rodent benefits whereas the plant dies. In other cases, it seems clear that both parties benefit, as when a hummingbird obtains nectar while transporting pollen from one plant to another. These interactions are termed mutualistic. The effects of other interactions may be more ambiguous. For example, Clark's nutcrackers not only consume pine-nuts but also disperse its seeds, and it is unclear whether the interaction is mutually beneficial to both species or whether the birds benefit at a net expense to the tree. The fitness costs and benefits to the parties involved in such an interaction may also be conditional on the current environment. For example, during years of heavy seed production, birds may provide trees with a net benefit, whereas during poor seed years, the net effect of birds on the tree may be negative (Figure 1).

Ecologists seek to categorize interactions because they have distinct ecological and evolutionary implications. Within biotic communities, certain kinds of interactions increase species diversity whereas others reduce it. Similarly, the nature of interactions determines their impact on genetic diversity within populations.

Table 1 Pairwise ecological interactions between plants and animals, categorized by their effects on the fitness of each party

		Plant effect on animal	*Animal effect on plant*
Antagonisms	Herbivory	+	−/0
	Carnivory	−	+
Mutualisms		+	+

Antagonistic Interactions

Plant Consumers – Herbivores and Granivores

Types of Plant Consumers

Animals consume plants in all kinds of habitats, including marine, terrestrial, and freshwater, and do so in wildly diverse ways. For some plants, there is at least one animal species devoted to consuming each type of organ. Some animals

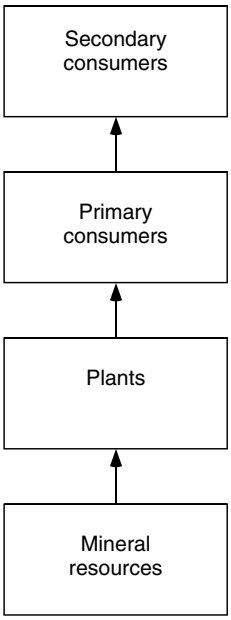


Figure 1 Trophic levels within a food web. Each trophic level is comprised of one or more species that consume individuals or resources at the next lower trophic level and are, in turn, consumed by species in the next higher trophic level. Arrows indicate the direction of flow of energy and resources in the food web.

remove tissues by chewing, others suck plant sap. Grazing individuals feed from many different individual plants. Some animals live within plants, literally surrounded by food. These include borers, gallers, and leaf miners.

Many animals chew on plants, just as we do. However, another important mode of consumption is to use straw-like mouthparts to pierce and suck fluids from vascular structures such as xylem and phloem saps, which transport water, minerals, and other compounds throughout the plant. Aphids are common sap feeders. Spider mites are another type of plant-sucking arthropod that may be familiar to unhappy owners of house plants. The origin of this feeding mode probably extends back to the Carboniferous with the Paleodictyopteroidea (Taylor and Scott, 1983; Labandeira, 1998), an assemblage of insects with sucking mouthparts.

Sap feeders share an interesting problem in common with many blood feeders: Neither blood nor plant sap provides a balanced or complete supply of vitamins and amino acids. One solution to this problem among aphids and their allies has been to host intracellular microbial symbionts with the enzymes necessary to convert common nonessential amino acids into essential rare ones, much as Midas converted base metals to gold (Douglas, 1994; Gündüz and Douglas, 2009). These intracellular guests inhabit special cells near the gut and oviducts and are transmitted by the mother to her eggs.

Perhaps the best-known mode of plant consumption is grazing. All of us are familiar with picturesque scenes of cows grazing or deer browsing in verdant pastures. Far from peaceful, these are scenes of graphic violence. Hundreds, perhaps thousands of plants are being eaten alive! Each grazer is eating photosynthetic organs from numerous living plant individuals. However, because the aboveground parts of most plants are constructed of repeating, renewable modules called shoots, a single episode of grazing or browsing rarely kills a plant.

Less well-known grazers include parrot fish, which maintain well-mown lawns of algae on coral reefs. When portions of the reef are protected from parrot fish, luxuriant algal growth can smother the coral. Sea urchins graze kelp forests and, when their natural predators are eliminated, may create vast barrens on the ocean floor. On land, many plant-eating insects, such as grasshoppers, katydids, and some beetles, are grazers. Some grazers feed on roots. These include tiny soil insects called springtails, relatives of silverfish, which move around in the air spaces between soil particles, nibbling on roots. Microscopic crustaceans in the water column of oceans and freshwater lakes and ponds graze on tiny photosynthetic single-celled organisms called phytoplankton. In this case, however, the animals function more like predators because they must kill their prey to eat it.

In contrast to grazers, which move from plant to plant and eat only a portion of each plant, some animals feed on only one plant during their lifetime. In some cases, many generations of the consumer will occupy the same tree, evolving greater and greater adaptation to that one individual (Edmunds and Alstad, 1978). The problem can be especially acute for an isolated tree, because adaptive traits in its local pest population might not be diluted by maladapted immigrants from nearby trees (Tack and Roslin, 2010). An important unanswered question in evolutionary biology is how long-lived plants such as trees avoid being destroyed by rapidly evolving pests.

Some consumers live inside their food. Perhaps the most intriguing are gall-forming arthropods: wasps, flies, and mites that induce plants to form elaborate structures within which the animal munches away on the host. Galls occur in many forms, including small red swellings on leaf blades, large globular swellings on branches, or bristly structures in odd places. In most cases, the female insect induces the gall when she injects her eggs into the plant tissues. Her young hatch and feed inside the gall, where they are relatively protected from enemies and the rigors of the physical environment. The relative frequency of gall-forming herbivores increases with increasing aridity of the environment, presumably because insects in galls are less vulnerable to desiccation. Plant-feeding nematodes, called vinegar eels, often live in galls on roots. The earliest know gall was produced by an insect feeding on tree-fern fronds in the Carboniferous (Labandeira, 2006).

The same species of galling insect can produce very different-looking galls on different plant species, suggesting that the plant determines gall form. But, when the same host is attacked by different insects, very different-looking galls can be formed. Detailed work on the goldenrod ball galler indicates that both insect and plant genes interact to determine gall size (Abrahamson and Weis, 1997). Although gallers effectively manipulate plant development, which is a goal of many biotechnologists, little is known about the molecular mechanisms of this fascinating interaction (Shorthouse *et al.*, 2005). The complex structures of some galls have even led to the tantalizing hypothesis that viruses carried by the animal induce and direct gall development (Cornell, 1983).

Another group of consumers that lives inside plants are leaf miners. As their name suggests, the larvae of these flies and

moths burrow within a leaf like a miner, eating the tissue they excavate and leaving behind their waste material called frass. Because they must live between the top and bottom surfaces of leaves, leaf miner larvae are very flattened from top to bottom. Some leaf miners pupate within the mine; others exit the mine and pupate in the soil beneath the host plant. In addition to mining leaves, some other insects mine plant stems, burrowing just beneath the stem surface and leaving similar-looking feeding galleries.

Some miners snake along, leaving a long serpentine mine in their wake. Others scrape away at the sides of their chamber, creating a large blotch in the leaf. In many cases, miners can be identified to species from their mine appearance and host plant, which has also made them easy to detect in the fossil record (Labandeira *et al.*, 1994).

Bark beetles are economically very important plant consumers (Paine *et al.*, 1997). Adults attempt to bore into tree trunks, but healthy trees generally can "pitch" them out, literally flooding the bore holes with sap. However, the sap attracts more adults of these gregarious beetles and an ailing tree's defenses can be quickly overcome. Female beetles then excavate extensive galleries just beneath the bark and lay their eggs in the termini of the tunnels. The larvae feed on both the host wood and on tree-feeding fungi that adult beetles transport to the tree in special structures called mycangia. Often, bark beetles do not kill their host but are indirectly responsible for its death because they introduce deadly fungal pathogens.

Evolutionary Responses of Plants

Plants have evolved numerous responses to their consumers. Primary among these is resistance. Plants resist consumers by defending themselves chemically or mechanically, or they may escape damage by being difficult for consumers to discover.

Chemical Defenses

Plants are highly proficient chemists. Early researchers were puzzled as to why plants produce such a startlingly diverse array of chemicals, which, because they had no known function in primary metabolism, became known as secondary compounds. Early theories proposed that these compounds were waste products from "pathological overproduction of carbon." But this explanation begged the question, "Why so many ways of throwing out the trash?"

By the 1950s, phytochemists had begun speculating that secondary chemicals were important in plant defense, and this theory has dominated the latter half of the twentieth century. Nevertheless, competing hypotheses have held their own – notably that secondary chemicals are important in protecting plants from physical dangers such as ultraviolet light. Some plant chemicals do perform such functions, but it is certainly not clear why plants should have so many kinds of sun screen. Perhaps the most defensible alternative hypothesis is that secondary compounds are important in defense against microbial enemies such as fungal, bacterial, and viral pathogens. It is quite likely, in fact, that many compounds defend plants against both these enemies and animal consumers.

Chemical defenses are classified into three categories: digestibility reducers, toxins, and repellents. Many plants produce digestibility reducers such as tannins, which are large molecular-weight carbon-rich compounds that bind proteins and make them difficult to digest. Thus, these compounds do not directly kill the consumer, but make the plant less nutritious to eat. This indirect mechanism of defense leads to important questions about the selection pressures that might have led to their evolution. In particular, it is not always clear that individual consumers know the nutritional value of their host. Moreover, in many cases the individual that chooses the plant is not the one destined to feed on it. For example, when choosing among a population of the same plant species, female butterflies do not necessarily lay their eggs on the individuals on which their larvae can develop best. Another question is, "If the chemical does not kill the consumer, why would it be evolutionarily advantageous to plants possessing it?" This question becomes even more vexing in light of considerable evidence that, when protected from other causes of mortality, consumers can compensate for poor food quality by consuming more (not less!) plant tissue. The leading hypothesis to explain this paradox is that consumers feeding on a plant with a digestibility-reducing compound will grow more slowly, making them more vulnerable to their enemies, such as predators, parasites, or diseases, and therefore die before consuming much plant tissue.

Toxins, as their name implies, are poisons that have direct negative effects on animals that consume them. Some plants are so toxic that apparently no animals will eat them. However, in many cases, certain animals have evolved mechanisms that detoxify extremely potent poisons. Often, these animals are specialists on that host plant. There may be strong evolutionary benefits to feeding on such a previously unexploited resource. The most obvious benefit is that no one else is using the plant, which reduces competition for food. However, because a toxic plant harbors so few consumers, it may also provide "enemy-free space" in which consumers are less likely to be discovered by predators or parasites (Bernays and Graham, 1988). Enemy-free space could also arise by a slightly different mechanism, which is nicely illustrated by small consumers on seaweeds. Small crustaceans called amphipods live on the algae they eat. Parrot fish also eat algae, but when they discover an amphipod, they snap it up like candy. Thus, amphipods that feed on very toxic species of algae are less likely to be consumed by herbivorous fish that are opportunistically omnivorous (Duffy and Hay, 1991).

Perhaps because of these kinds of benefits, many animals that have evolved the ability to consume a toxic plant have also evolved specialization to that host. For example, specialist herbivores may use the toxic compound as a cue to find their host. Additionally, the compounds may stimulate feeding or egg laying. This level of specialization sets up an important evolutionary trade-off for toxic plants. A compound that previously killed all consumers has become an attractant to at least one consumer. Individual plants that have higher levels of the compound may be better defended against most consumers but may be more attractive to the specialist. Conversely, plants with lower toxin concentrations will be less attractive to the specialist but may become vulnerable to other generalist consumers. This type of situation may create stabilizing selection and cause the plant population to evolve an intermediate concentration of the compound (Simms, 1992). Alternatively, the specialist could become a major cause of

death and seriously reduce abundance of the host. Indeed, when plants introduced into new habitats become noxious weeds, specialist consumers are sometimes imported to control them.

A final class of chemical defenses is repellents, which as their name suggests, repel animals from laying eggs on or eating plants that possess them. Some evolutionary biologists have argued that herbivores will not quickly evolve mechanisms that overcome repellents. This argument is based on the understanding that the rate of evolution depends partly on the strength of selection. Toxins impose strong selection for detoxification mechanisms because they kill animals. Repellents simply cause animals to look elsewhere for food, which might impose weaker selection. The strength of selection imposed by repellents on consumers will, however, depend critically on the fitness cost to the animal for finding an acceptable alternate host.

Animal uses of plant defensive compounds Whatever the evolutionary reasons for these compounds, plant secondary chemicals have enormous value to humans. For example, synthetic pyrethroids, which were originally extracted from a species of chrysanthemum, are valuable insecticides because they are effective but degrade quickly and so do not accumulate in the environment. Plant secondary chemicals are the source of virtually all of the herbs and spices that make food interesting to eat. In many cases, these compounds were first important in preserving foods and keeping them safe to eat. Finally, plant secondary compounds are an important source of pharmaceuticals (Farnsworth *et al.*, 1985; Harvey, 2008).

Other animals also use plant secondary compounds in interesting ways. Many consumers sequester plant compounds from their food and use them for their own defense; the best-known example being the monarch butterfly, which harbors cardiac glycosides from its milkweed hosts. Some animals even self-medicate with secondary compounds (Rodriquez and Wrangham, 1993). For example, healthy wooly bear caterpillars avoid poison hemlock, the plant used to execute Socrates. But wooly bears infested with a lethal parasitoid will preferentially consume poison hemlock (*Conium maculatum*), which can kill the parasite and allow the caterpillar to survive to adulthood (Karban and EnglishLoeb, 1997). European starlings protect their nestlings by lining their nests with fresh plant materials that inhibit mites from hatching (Clark and Mason, 1988). Animals may also use plant secondary compounds to preserve their food. Pikas, relatives of rabbits, live in burrows in talus slopes. To survive the winter in their alpine homes, they harvest and store enormous quantities of vegetation in haypiles. Pikas prefer to eat grass hay, but they often will add toxic herbs to haypiles, which reduces the chance of the pile becoming moldy (Dearing, 1997).

Mechanical Defenses

Mechanical resistance to consumers may be obvious, as in the case of spiny cacti and thorny shrubs, or more subtle, as in the case of silica bodies that wear down the teeth and mandibles of many consumers. Some plants combine both chemical and mechanical defenses. For example, Wright's datura (jimson weed) possesses both toxic alkaloids and leaf hairs called trichomes (van Dam and Hare, 1998). It is generally supposed that mechanical defenses are more difficult for consumers to overcome (Coley, 1983).

When animals consume other animals, they eat tissue that has a composition similar to their own. However, plant tissues are generally richer in carbon and poorer in nitrogen than animal tissues. Partly because of mechanical defenses, a high proportion of the carbon in terrestrial plants is devoted to structural molecules such as cellulose and lignins. Animals lack the enzymes necessary to digest these compounds. Microbes do have such enzymes, however, and herbivores often have elaborate modifications of their guts which house microbial symbionts that can digest the fibrous fare.

Escape

In addition to defending tissues, plants may also escape consumers in time or space. Ephemeral plants, especially those with annual lifecycles, may obtain temporal escape with a lifecycle that does not sustain consumers long enough for them to complete their lifecycles. This constraint prevents the build-up of dense pest populations which short-lived consumers can attain on long-lived plants. Plants might also escape by appearing only during seasons when consumers are rare. For example, many cool-season plants are relatively unaffected by insect consumers, which are far more abundant during warm weather.

Highly dispersed plants may also escape consumption if the distances between them are greater than the average distance traveled by their consumers. This concept is embodied in the "resource-concentration hypothesis" (Root, 1973), which states that dense concentrations of host plants will harbor the highest densities of consumers. If a plant species is both short lived and rare, it may be difficult to find that it can complete its lifecycle before being found by consumers. This mechanism may in part be responsible for the astonishing diversity of certain tropical forests (Janzen, 1970; Connell, 1971), in which only one individual of each tree species will be found in a large area.

Finally, plants might escape herbivory when the enemies of herbivores are attracted to plant volatiles emitted from damaged tissues. Debate continues as to whether natural enemies of herbivores have evolved the ability to detect compounds passively emitted by damaged tissues or such volatile compounds are a plant adaptation that attracts "mobile body guards" (Dicke *et al.*, 1990), but these compounds are very promising as tools to reduce pest damage to crops.

Tolerance and Compensatory Growth

Even if plants are damaged by consumers, they may evolve mechanisms that allow them to maintain fitness in the face of damage (Simms and Triplett, 1994). Plants may tolerate damage through various compensatory mechanisms, including reallocating resources from undamaged plant parts to replace damaged tissues (Stowe *et al.*, 2000). Tolerance of consumers may be especially important among fast-growing plants living in resource-rich environments (Coley *et al.*, 1985).

Coevolution of Plants and Herbivores

Ehrlich and Raven (1964) conceived a coevolutionary hypothesis to explain the magnificent diversity of plant chemistry. They postulated that herbivorous insects cause plant mortality and that any novel trait that protected a plant from its herbivorous insects would quickly spread in the plant population. Further, they argued that escaping its herbivores would create for the plant the opportunity for a period of rapid speciation called an adaptive radiation. For the herbivores, newly evolved plant species represent unused resources. Any insect trait that allowed an insect to exploit these plants would likewise result in an adaptive radiation of herbivores. As Janzen (1980) later argued, the crucial characteristic of this process that distinguishes it from ordinary evolution by natural selection is its reciprocal nature. The plant evolves resistance to the herbivore; the herbivore then evolves a mechanism that negates the resistance, after which the plant evolves resistance to the herbivore, and so on, *ad infinitum*. With every crank of the coevolutionary process, new species arise through adaptive radiation.

This hypothesis has excited considerable controversy. Some authors argue that reciprocal coevolution is rare or nonexistent because insect herbivores do not impose sufficiently strong selection pressures on plants (Jermy, 1976). Others argue that herbivores experience more selection from their natural enemies than from plants, and that moving to new plants is driven by the adaptive advantage of enemy-free space (Bernays and Graham, 1988). Further, Labandeira and Sepkoski (1993) pointed out fossil evidence that the great radiation of modern insects began 245 million years (My) ago and was not accelerated by the expansion of angiosperms during the Cretaceous period. However, neither insects nor plants have stopped evolving, and currently evolving systems provide the best tests of the coevolutionary hypothesis.

One way that scientists have tried to test the coevolutionary hypothesis is to compare the evolutionary lineages (phylogenies) of host plants and insect consumers that narrowly specialize on that group of hosts. Coevolution can lead to a pattern called cospeciation (Brooks, 1979), in which the two phylogenies match, much like fingers do when the palms are placed together. One hand represents the plant lineages, the other the insects. At the base of the palm is the ancestral species; the fingers represent various derivative lineages. An alternative phylogenetic pattern, in which lineages do not match, is produced by host switching. Host switching is the phenomenon whereby specialist consumers shift host species and then speciate on that new species, without any speciation by the host plant.

The alliance of Hawaiian silverswords and the plant hoppers that live and feed on them provide a particularly exciting pair of phylogenies to test the coevolutionary hypothesis. The Hawaiian silverswords are derived from a pair of ancestral species in the rather prosaic group of California plants called the tarweeds (Baldwin, 1997). Over the past 5 My, the Hawaiian descendents have rapidly radiated into a stunningly diverse array of species. Some of these Hawaiian descendents are magnificent rosette plants with life histories much like century plants – they live long lives terminating in the production of giant flowering stalks. Others are multibranched perennial shrubs. Most are attacked by members of the plant hopper genus, *Nesosydne*, in the family Delphacidae. The plant hoppers are highly host specific; each species feeds on one or a few closely related plant species (Roderick, 1997). Phylogenetic analysis of molecular data from the plant hoppers and their hosts reveals a pattern of cospeciation: Each plant hopper species is most likely to use the host species that is most closely related to the host of its closest relative. This pattern is exciting, because it is expected to arise from reciprocal coevolution.

However, other mechanisms could produce the same pattern. For example, the plants could be evolving in response to some other selective pressure, with the insects following along behind. Matching phylogenies can also arise when both an insect and its host plant speciate simultaneously in response to some external event as, for example, when they become geographically isolated from their main populations by a geologic event such as mountain building (Roderick, 1997). Thus, phylogenetic comparisons alone cannot deduce whether coevolution has occurred. Other types of data must be examined. Careful observation of host used across hybrid zones suggests that the most likely explanation of the match between silversword and plant hopper phylogenies is that plant hosts are speciating in response to some external pressure and plant hoppers are tagging along behind plant host speciation (Roderick, 1997).

Community and Ecosystem Effects of Plant–Consumer Interactions

Considerable evidence suggests that consumers can reduce the fitness of individual plants and thereby impose selection pressures that produce evolutionary changes in plant traits. However, it is less clear whether consumers influence plant abundance in the landscape. This issue is at the heart of a controversy that has raged among community ecologists for decades regarding the relative importance to community structure of top–down control by consumers of their resource populations (e.g., Hairston *et al.*, 1960) versus bottom–up control of consumer populations by resources at the base of the food web (e.g., Lindeman, 1942). The current view is that both types of forces interact in complex ways to structure biotic communities (e.g., Oksanen *et al.*, 1981). However, several competing hypotheses aim to explain how these forces interact. Three of the major models are summarized here.

Donor control models predict that while plants feed their consumers, the latter have little effect on plant abundance. Thus, biomass of organisms at one trophic level is a function of the productivity of their resource base at lower trophic levels. This means that adding resources to the base of a food web will trickle up the web, increasing biomass at all trophic levels. In contrast, consumer control models predict that each trophic level can be controlled by either its resources or its consumers, but not by both. They further predict that the direction in which the control moves depends on the trophic level being examined and the number of trophic levels in the ecosystem. In particular, plants, at the base of the food web, are expected to dominate in ecosystems with odd numbers of trophic levels, whereas herbivores will dominate in ecosystems with even numbers of trophic levels. Thus, increasing abundance in a particular trophic level will cascade up and down the food web, alternately expanding or shrinking trophic levels. Finally, keystone predation models predict that the

species composition at each trophic level modifies the relative effects of resources and consumers.

The last model is more complex than the previous two, incorporating aspects of each. It is of particular interest here because it incorporates information about the relative vulnerability to consumers (i.e., resistance) of different resource (e.g., plant) species. The keystone model predicts that species diversity in resource populations can be maintained if resource species exhibit trade-offs between their relative competitive abilities and their relative resistance to consumers (Paine, 1966). Further, the model predicts that when resources are scarce, consumer populations will be small, the plant community will be dominated by a few fast-growing, strong competitors that are highly vulnerable to consumers, and is therefore consumer controlled. Under nutrient-rich conditions, consumer populations will be large and the plant community will consist primarily of a few slow-growing, well-defended species that are resource controlled because of their heavy investment in resistance. At intermediate levels of nutrient availability, both types of plants can coexist because of the trade-off between competitive ability and resistance to consumption (Leibold *et al.*, 1997). Consequently, species diversity will be greatest at intermediate levels of productivity (Grover, 1994; Holt et al., 1994; Grover, 1995; Leibold, 1996). Other factors will determine whether these diverse communities are controlled by consumers (top–down) or by resources (bottom–up).

Leibold (1999) examined this prediction in planktonic communities of fishless ponds that varied in their level of mineral nutrient availability. These communities consist of photosynthetic planktonic algae (phytoplankton; single-celled green plants) that are grazed by herbivorous microarthopod zooplankton. Leibold found that algae in low-nutrient ponds consisted primarily of small, unprotected forms thought to be fast-growing but susceptible to grazing. Algae from more eutrophic (nutrient-rich) ponds were larger and often sheathed or gelatinous forms thought to be slow-growing but resistant to grazers.

In another analysis, Chase *et al.* (2000) reviewed studies of temperate terrestrial grasslands to determine whether the effects of consumers on plant biomass fit the keystone herbivore model, which predicts that consumer control should be strongest at high levels of resource availability and decline with declining productivity. They reviewed the results of experiments that manipulated the presence or absence of large grazers. Because most temperate grasslands are water-limited, they sought evidence for a correlation between consumer effect and precipitation. As predicted, the proportional effect of consumers on plant biomass declined significantly with increasing precipitation. Schmitz (1994, 1997) found a similar relationship between plant productivity and the effect of insect herbivores on plant biomass. Further, Chase and colleagues found a turnover in species composition among plants along the precipitation gradient, as predicted by the keystone model.

Plants as Consumers – Carnivorous Plants

The most ubiquitous interaction between plants and animals is the use of plants by animals as sources of material resources and energy. However, there are a few plants that turn the tables. In a world of plant-eating animals, carnivorous plants eat animals (Givnish, 1989). More than 500 species in nine plant families have evolved the carnivorous habit.

Mechanisms of Prey Capture

Carnivorous plants capture prey in several remarkable ways (**Table 2**). The evolutionarily independent origin of carnivory is demonstrated by the many ontogenetic origins of the traps. A pitfall trap is a tubular structure, often containing liquid, which prey can enter but have difficulty leaving. Five different plant families capture prey in some kind of pitfall trap, called tanks or pitchers, and these traps may be comprised of leaf

Table 2 Presence (+) or absence (−) of adaptations for active prey attraction, capture, and digestion in carnivorous plant genera and species

Genera (number of spp.)	*Attraction*	*Digestion*	*Type of trap*
Bromeliaceae			
Catopsis (1)	+	−	Pitfall
Brocchinia (2)	+	−	Pitfall
Eriocaulaceae			
Paepalanthus (1)	+	−	Pitfall
Sarraceniaceae			
Heliamphora (6)	Chemical and visual	−/+	Pitfall
Darlingtonia (1)	+	−	Pitfall
Sarracenia (9)	+	+/−	Pitfall
Nepenthaceae			
Nepenthes (82)	+	+	Pitfall
Cephalotaceae			
Cephalotus (1)	+	+	Pitfall
Droseraceae			
Drosera (90)	−	+	Active flypaper
Aldrovanda (1)	−	+	Steel trap
Dionaea (1)	+	+	Steel trap
Dioncophyllaceae			
Triphyophyllum (1)	−	+	Passive flypaper
Drosophyllum (1)	+	+	Passive flypaper
Roridulaceae			
Roridula (2)	+	+	Passive flypaper[a]
Lentibulariaceae			
Pinguicula (35)	−	+	Active flypaper
Utricularia (280)	+/−	+	Bladder trap
Genlisea (35)	−	+	Lobster-pot
Biovularia (1)	−	+	Bladder trap
Polypompholyx (2)	−	+	Bladder trap
Byblidaceae			
Byblis (2)	−	+	Passive flypaper

[a]Nutrient uptake apparently assisted by exudations of kleptoparasitic bug, *Pameridea roridulae*.

Source: Modified from Givnish TJ, Burkhardt EL, Happel RE, and Weintraub JD (1984) Carnivory in the Bromliad *Brocchinia reducta*, with a cost–benefit model for the general restriction of carnivorous plants to sunny, moist, nutrient-poor habitats. *American Naturalist* 124: 479–497, and Givnish TJ (1989) Ecology and evolution of carnivorous plants. In: Abrahamson WG (ed.) *Plant–Animal Interactions*, pp. 243–290. New York: McGraw-Hill.

rosettes (e.g., *Brocchinia*), modified leaves (e.g., *Sarracenia*), or modified leaf-tips (e.g., *Nepenthes*). Some plants have active mechanisms to trap prey. For example, the leaves of the Venus flytrap (*Dionea miscipula*) function like a miniature steel-jawed trap when tripped by the hapless prey. There are even aquatic plants (e.g., *Utricularia*) with sophisticated underwater traps that slurp up prey unlucky enough to trip them.

Costs of Carnivory

Carnivorous plants are usually restricted to sunny, moist, nutrient-poor habitats, such as bogs and fens. Their slow growth and restricted distribution suggests that there are fitness costs associated with carnivory. For example, it has been argued that leaves morphologically specialized for prey capture have compromised photosynthetic abilities, making carnivorous plants poor competitors. This hypothesis is supported by studies of the pitcher plant *Sarracenia alata*, which produces two kinds of leaves, a "regular" leaf and one that is modified into a pitcher. The plant responds to competition from neighboring vegetation by diverting resource allocation from pitchers to regular leaves (Brewer, 1999a, b). Small stature and slow growth also make many carnivorous plants (e.g., *Pinguicula* and *Utricularia*) vulnerable to being buried by litter fall (Brewer, 1998, 1999a, b). Because of their compromised competitive ability, carnivorous plants generally respond poorly to addition of nutrients to their habitat, being easily outcompeted by plants that thrive under richer conditions. This characteristic creates important conservation concerns in many parts of the world where atmospheric input of anthropogenic nitrogen is significantly increasing nitrogen availability in previously nitrogen-poor bogs and fens.

Another fitness cost of carnivory was identified by Zamora (1999), who found that *Pinguicula vallisneriifolia* tends to trap its own pollinators. He also found that reproduction in the plant is limited by pollen availability, indicating that feeding on its pollinators reduces plant fitness.

Benefits of Carnivory

As with so many plant novelties, plant carnivory was noticed by the inquiring mind of Darwin (1875), pp. 1–2, who demonstrated that the sticky traps of *Drosera rotundifolia* do indeed capture and digest animals. Darwin's son, Francis Darwin (1877b), first demonstrated experimentally that prey capture enhances the growth and reproduction of this species. Similar studies have subsequently found that in most circumstances, carnivorous plant growth benefits from prey capture.

In many cases, the majority of a carnivorous plant's nitrogen and phosphorous is obtained from prey, but only a small proportion of other necessary nutrients come from prey. Experiments suggest that plant growth is generally restricted by the rate of prey capture; plants could utilize many more prey than they were able to catch. Carnivorous plants tend to be frugal with their nutrients, practicing particularly efficient internal recycling of nitrogen and phosphorous.

Many elegant methods have been used to examine in more detail nutrient uptake from prey in carnivorous plants (Adamec, 1997). In particular, putatively carnivorous plants can be offered insects reared on media enriched in the stable nitrogen isotope N^{15}. If the plant tissues subsequently become N^{15} enriched, this indicates that their nitrogen supply has been supplemented by insect proteins. Using this method, Hanslin and Karlsson (1996) found that *D. rotundifolia* and several species of *Pinguicula* in a subarctic environment took up 29–41% of the nitrogen available in insect prey they were offered. Further, root uptake of nitrogen was stimulated by prey capture, an unanticipated additional benefit of carnivory. Experiments performed in glasshouse or laboratory environments generally reveal even greater uptake efficiencies.

Carnivorous plants show interesting developmental changes during maturation of their carnivorous organs. Using fluorescent dye tracers, Owen *et al.* (1999) found developmentally regulated bidirectional transport by leaf glands in the pitcher vine, *Nepenthes alata*. In mature leaves, the glands transport fluids directly from the pitcher fluid to the plant vasculature (internal plumbing system), apparently functioning in nutrient uptake. However, in immature closed leaves, the glands secrete fluid from the vascular tissues into the pitcher, building up a supply of fluid in which to eventually trap prey. Expression of hydrolase, an enzyme involved in prey digestion, commenced when the pitcher first opened on maturity, increased for several days, and would largely cease after 2 weeks without prey (Gallie and Chang, 1997). However, adding prey-derived resources such as amino acids to the pitcher fluid could induce hydrolase expression in pitchers that had ceased expression due to lack of prey.

Nonprey Guests – Iniquiline Communities

Prey-digesting guests (iniquilines) are very common in pitfall traps, and their bodily wastes provide for some carnivorous plants the sole means of benefiting from prey. The foodwebs of iniquilines in pitcher plant (*Sarracenia* spp.) traps have been the subject of numerous highly informative community and population ecology studies.

In one of the most arcane modes of nutrient uptake, *Roridula gorgonias* hosts a bug, *Pameridea roridulae*, that feeds on insects trapped on its sticky leaves. Although the plant has no known method of digesting its prey, stable isotope studies indicate that it acquires N^{15} label from prey. It apparently derives nutritional benefit via exudations from the bug guest which has dined on the labeled prey (Ellis and Midgley, 1996; Midgley and Stock, 1998). This example, however, highlights the vulnerability of insectivorous plants to kleptoparasites (animals that steal prey). Spiders, in particular, frequently compete with carnivorous plants for prey that have been attracted to the plant.

Mutualistic Interactions

In addition to antagonistic interactions in which one party feeds on the other, plants and animals may also interact in ways that can benefit both parties. Far from being pleasant

affairs, however, mutualistic associations can be highly vulnerable to cheating, which often makes them evolutionarily uneasy truces between parties. The delicate evolutionary and ecological balances that can be achieved by these organisms are truly fascinating and lead to some common evolutionary issues.

One important question regards the degree of specialization between mutualists. Specialization is useful because it allows the evolutionary development of elaborate lock-and-key mechanisms that exclude cheaters and maintain the mutualism. However, specialization is an evolutionarily vulnerable position, because extinction of one partner species can spell doom for the other. Further, in mutualistic interactions that must be reconstituted each generation, which is the case for all animal–plant interactions, specialization might make it more difficult for individuals to find the correct partner.

Plant-Protecting Ants and Plant-Feeding Ants

As described previously, plants have evolved a variety of mechanisms that defend them against herbivores. One of the strangest defenses, though, is provided when plants are guarded by ants. Once thought to be rare and unusual, myrmecophytes (ant plants) are actually widespread and ecologically important. In many cases, these relationships appear to be quite casual. Visiting ants may rob nectar from flowers, but perform some guarding services in return. However, the term myrmecophyte is generally reserved for the more specialized case in which an ant colony resides in special structures provided by the plant.

Benefits to Ants – Costs to Plants

Tropical myrmecophytes display the most sophisticated development of this type of interaction. In the neotropical regions alone, associations between ants and myrmecophytes have been described for about 250 plant species from 19 families, and up to 180 ant species from five subfamilies (Davidson and McKey, 1993). These ant plants may provide ant housing in specially modified stems, hollow thorns, or specialized leaf pouches called leaf-domatia. The trees or shrubs often feed ants with amino acid- or sugar-based solutions produced by extrafloral nectaries. The most developed myrmecophytes may also provide food in the form of specialized structures composed of lipids (Beccarian bodies), proteins (Beltian bodies), glycogen (Müllerian bodies), or some combination. In the most elaborate cases, plants may provide everything – room, board, and drink – to their ants.

Food provision for ants may be quite costly to myrmecophytes and studies show that plants regulate production of food structures. Removing Müllerian bodies from Central American *Cecropia* (Moraceae) trees stimulates their production. In contrast, when Müllerian bodies accumulate, which would happen if ants were not present, plants cease production (Folgarait *et al*., 1994).

Many ant–plant interactions have an important third partner through which the ants obtain benefit: sap-sucking homopterans tended by ants. Ants derive benefit from these homopterans in two ways. They may keep "milk herds" from which they obtain honeydew, or they may be in the "beef business" and eat the homopterans they tend. Rather than simply supporting ants, the plants in these situations must also support homopteran consumers. Homopterans increase plant risks too, as they are often important vectors of plant diseases.

The homopteran mode of ant benefit provides few options for control by the plant. This problem is illustrated with the African myrmecophtye *Leonardoxa africana*, on which the same ant species (*Aphomyrmex afer*) may tend one or both of two different homopteran species. Gaume *et al*. (1998) found that homopteran identity influenced the costs and benefits to the plant of ant patrol. One homopteran, the pseudococcid, could support larger colonies of ants, leading to better plant defense. This homopteran was also more efficient at producing ant biomass; ants tending pseudococcids did not use other plant resources. Moreover, the plant could directly control pseudococcid colony size by regulating domatia volume. Plants that provided a smaller total volume of the swollen stems used as domatia supported fewer pseudococcids and therefore fewer ants. However, when ants tended coccids, the other homopteran, they also used plant resources from the extrafloral nectaries. Thus, when ants tended coccids, the only control that plants had over homopteran feeding was indirectly through nectary production. Plants that produced fewer nectaries supported fewer ants, and fewer ants could tend fewer coccids.

Benefits to Plants – What Motivates Ants to Patrol?

Ants have frequently been observed killing and removing insect herbivores, and numerous experiments demonstrate the efficacy of this defense. Fonseca (1994) observed more than four times as many herbivores on *Tachigali myrmecophila* plants from which he removed *Pseduomyrmex concolor* ants as on plants with intact ant colonies. Further, the daily rate of herbivory was about 10 times lower when ants were present, resulting in experimental plants without ants exhibiting about twice as much cumulative herbivore damage during the 18-month experiment. Leaf longevity was also substantially higher on plants with ants. It is interesting to note that these ants do not eat the herbivores they kill. Instead, they feed exclusively on catenococcid insects they tend inside the domatium, which is the hollow rachis of the compound leaf.

Ants are also effective deterrents of mammalian herbivory. For example, the African myrmecophyte, *Acacia drepanolobium*, possesses two kinds of thorns. The swollen thorns are domatia in which *Crematogaster* ants live and rear their broods. The unswollen thorns slow plant damage by browsing mammals, but browsers may compensate by feeding longer. Ants provided far more effective defenses (Stapley, 1998). When a browsing mammal was encountered and stung by ants, it stopped feeding immediately and could not be induced to feed further on that tree.

A second benefit that patrolling ants may provide is in competition with neighboring plants. Ants will prune vines (lianas) and branches of neighboring trees, effectively preventing their host tree from being overgrown (Janzen, 1966). The result of this vigilance is that the host tree might occupy a dramatically open cylinder of space amid otherwise densely packed tree canopies. Although such pruning of neighbors clearly benefits the host tree, it also benefits the ant colony by

reducing the number of directions from which it may be attacked by competing or predatory ants (Davidson *et al.*, 1988).

However, ants sometimes harm their own host by pruning it rather than its neighbors. For example, *Crematogaster nigriceps* so severely prunes its host tree, *A. drepanolobium*, that the tree cannot flower and is sterilized (Stanton *et al.*, 1999). In the habitat studied, four species of ants compete strongly for hosts, and *C. nigriceps* fares poorly in the violent conflicts over nest space. Instead of pruning neighboring trees, *C. nigriceps* prunes its own tree, apparently because it cannot prune neighboring trees occupied by competitively dominant ants. Indeed, careful observation of a large number of trees occupied by *C. nigriceps* revealed that these trees were always pruned in such a way as to avoid canopy contact with adjacent trees occupied by competing ant colonies. Canopy pruning of its own tree appears to be a defensive response by a *C. nigriceps* colony to competition with dominant ants which prevent it from pruning their trees.

Ants Feeding Plants – Myrmecotrophy

Finally, a very different group of plants receives nutrients from the ants they house. These plants are also known as ant epiphytes. Epiphytes are plants that live on, but derive no nutritional benefit from, the branches of other plants. Myrmecotrophic epiphytes provide ant domatia in hollow or inflated roots, hollow rhizomes, or folded leaves. Ants then act as "mobile roots," gathering food items for the nest, processing them, and then depositing the resulting waste and fecal matter within the plant. The best studied of these systems is *Myrmycodia tuberosa* (Rubiaceae), an epiphytic shrub of southeastern Asia and northern Australia. This species also has elaiosome-bearing seeds, a typical feature of ant-dispersed seeds. The ants feed on these food bodies and then "plant" the seeds along the walkways they create in the canopies of the trees their plant "homes" inhabit.

Plant–Pollinator Interactions

An important and obvious characteristic of plants is that, with few exceptions, they are rooted to the ground and cannot move. This poses a crucial problem for sexual reproduction: Immobile mates must exchange gametes, meaning that pollen must be moved to the ovule. Pollen can move passively with fluid flow in the physical environment. Wind pollination has been successful among the gymnosperms and water pollination is found among many aquatic angiosperms. However, as many allergy sufferers know too well, most of the pollen produced by wind-pollinated plants never reaches its intended target. Many land plants avoid this inefficiency by using animals as pollen vectors. Most temperate angiosperms and almost all tropical angiosperms are animal pollinated.

Of course, animals do not move pollen around as a favor to the plant. Attracting pollen vectors has been an important evolutionary problem for plants and the need to attract pollinators has probably been a primary driver in the evolution of flower morphology. Animal-pollinated flowers often have large, brightly colored structures that function as "advertising" for the goodies available to the visitor. The rewards may be nectar, other more specialized chemicals, or even the pollen itself. As with any purveyor of treats, the flower also must contend with thieves. For example, nectar-robbing bees may drill through the flower wall and gain access to the nectar without transporting pollen (Inouye, 1980). Other visitors may obtain rewards yet be too small to trip the elaborate pollen-application mechanisms of some flowers (Bertin, 1989).

The potential for cheating selects for specificity in pollinator attraction. Specificity can also promote pollinator fidelity. Simply attracting a pollinator once is not sufficient. To ensure fatherhood, the plant must attract an animal that will visit other flowers of the same species. Moreover, to avoid inbreeding, the recipient flower must be on a different individual. These competing needs create strong evolutionary pressure for plant traits that promote pollinator specificity. As might be expected when sex is involved, response to this selection pressure has led many plants and animal pollinators to exhibit fascinating and baroque relationships (Darwin, 1877a, b, 1984; Grant and Grant, 1965). For example, male euglossine bees depend on flowers of plants in the euphorb and orchid families for fragrances which they use as pheromones to attract mates (Williams and Whitten, 1983; Armbruster, 1993). In another instance, male insects are tricked by orchids into thinking that they have found a mate, when in fact all they have discovered is a cleverly shaped and scented mimic (Peakall and Beattie, 1996).

Role of Plant Pollinators in Plant Diversification

An important aspect of these highly specialized relationships is that they can sometimes prevent mating between genetically compatible individuals, a phenomenon known as reproductive isolation. Evolution in a trait that promotes reproductive isolation can quickly lead to speciation. Traits that create the opportunity for rapid speciation by exploiting novel resources such as new pollinators are considered "key innovations."

Identifying such key innovations is a sticky problem in evolutionary biology. Circularity arises because the characteristic that best defines a group, and therefore allows it to be identified as speciose (having lots of species), is frequently also the character postulated to be the key innovation responsible for the radiation. One way out of this thicket is to identify a causal link between the putative key innovation and one of the processes that determines diversity.

Floral Nectar Spurs as a Key Innovation

One especially persuasive example of a key innovation is floral nectar spurs. These structures are critically involved in pollinator specialization because they hold nectar deep within the flower and make it available to only a narrow range of floral visitors. Animals must either be small enough to enter the spur or have sufficiently long and narrow mouthparts to sip nectar from the spur. Observing the 12-inch long nectaries of a Madagascar orchid, *Angraecum sesquipedale*, prompted Charles Darwin (1877a, 1984, pp. 162–163) to predict correctly the existence of a moth with a sufficiently long proboscis to pollinate this fantastic flower. Nectar spurs have evolved independently in several distantly related families and genera of flowering plants and, as would be expected were spurs driving diversification, spurred groups significantly have

more species than closely related groups without spurs (Hodges, 1997).

Experimental research has corroborated the putative link between nectar spur morphology and pollinator fidelity. Within two groups of orchids, experimental manipulation of spur length significantly changed rates of both pollinia (pollen bags) removal by pollinators and fruit set (Nilsson, 1988; Johnson and Steiner, 1997), demonstrating that spur morphology directly influences pollinator-mediated reproductive success. In a study of columbines, hybrids varied in floral characters, including spur length and orientation, and these morphologies differentially attracted either hummingbirds (primary pollinators of *Aquilegia formosa*) or hawk moths (primary pollinators of *Aquilegia pubescens*), and thereby promoted reproductive isolation (Hodges and Arnold, 1994). However, the presence of a hybrid zone indicates that floral morphology has not prevented pollination "mistakes." Nevertheless, comparative analysis suggests that North American columbines have, over the past 2 My, radiated into species that are primarily bee, hummingbird, or hawkmoth pollinated and have corresponding floral morphologies (Whittall and Hodges, 2007).

Insect Pollination in the Fossil Record and the Angiosperm Radiation

The role of plant–pollinator interactions in reproductive isolation has also led to the much grander hypothesis that insect pollination was a "key innovation" leading to the coradiation of flowering plants (Angiosperms) and anthophilous insects (the groups most involved in pollination), including certain bees and wasps (Hymenoptera), various families of flies (Diptera), and butterflies and moths (Lepidoptera).

One necessary prediction of the insect-pollination Angiosperm-radiation hypothesis is concurrent diversification of the angiosperms and anthophilous insects. Despite recent progress in our understanding of flower evolution (Specht and Bartlett, 2009; Endress, 2011), controversy still surrounds the dates of angiosperm diversification (Crepet, 2008). Nevertheless, recent analysis of the land plant record has produced a richer story that even more powerfully supports a tight evolutionary relationship between plants and pollinators (Labandeira and Sepkoski, 1993; Crepet, 2008). Indeed, insect pollination seems to have preceded the origin of angiosperms and insect diversity may have waxed, waned, and resurged with the fortunes of different plant groups.

Fossil plant and insect morphology suggests that many ancient gymnosperms, cycads, and early seed plants from the Middle Jurassic and Early Cretaceous were pollinated by insects with body parts specialized to obtain pollen, nectar, or other plant rewards (Labandeira, 2010). However, the global turnover from gymnosperm to angiosperm floras in the early Cretaceous apparently extirpated much of this diverse plant–pollinator biota and reduced insect diversity for 20 My until the subsequent diversification of angiosperm flowers and pollinators.

The earliest angiosperms tended to have bowl-shaped, open flowers accessible to a variety of small unspecialized insects (Friis *et al.*, 1987). These would not have facilitated reproductive isolation. After another 45 My, as angiosperms began to dominate the flora, reproductive isolating mechanisms became evident. The fossil record of this time suggests pollinator specialization, including the earliest fossil evidence of bees and other long-proboscid insects, and floral traits such as tube-shaped flowers, which restrict floral rewards to specialized pollinators (Crepet, 2008). Molecular and fossil data also suggest the concomitant rise of new angiosperm groups characterized by rapid diversification of lineages with specialized pollinator associations (Lavin *et al.*, 2005).

Pollination Syndromes

Many plant species that share animal pollen vectors also share similar suites of floral traits, such as color, shape, symmetry, and scent, which appear to attract those kinds of animals. For example, hummingbirds fly by day, have excellent color vision, possess long beaks, and visit flowers while in flight. Correspondingly, hummingbird-pollinated flowers are day-blooming and tend to be brightly colored, often red, bilaterally symmetrical and tubular in shape, and frequently pendant. Floral traits also may correspond to the physiological needs of pollinators, as with hummingbird-pollinated flowers, which tend to produce copious sucrose-rich nectar that helps fuel the notoriously high metabolic rates of their pollinators. A number of these trait combinations, which are called pollination syndromes, are summarized in **Table 3** (Howe and Westley, 1988).

Implicit in the concept of pollination syndromes is the assumption that floral evolution has been strongly entrained by interactions with specific classes of pollinators, leading to strong specialization. Specialization, however, may be an evolutionary dead end. A plant highly specialized and dependent on one or a few species of pollinator is seriously vulnerable to extinction of its pollinators. Dependence on specialized pollinators can also lead to "Allee effects," which arise when populations become too sparse to persist. For example, individuals in small, isolated populations of the annual plant farewell-to-spring (*Clarkia cocinna*) were visited by so few pollinators that they could not produce enough seeds to replace themselves (Groom, 1998). If individual reproductive rates dip below replacement levels for very long, a population can dwindle to extinction.

Figs may provide a particularly impressive example of this problem. Fig trees are ecologically important components of tropical forests because the fruits they produce support frugivorous animals that are important seed dispersers for many other forest plants. Figs do not have large showy flowers to attract pollinators. Instead, fig reproduction is exquisitely dependent on minute wasps that pollinate small flowers held tightly inside the closed fig. The wasp's end of the bargain is met when the eggs it lays inside the fig hatch and its larvae feed on a few of the many developing seeds. Individual fig trees flower for a relatively short time but fig wasp populations cycle constantly. Maintenance of the wasp population requires that female wasps emerging from a mature fig find and lay eggs in figs of other trees of the same species that are flowering at times other than their natal tree. Genetic evidence shows that, despite their minute size and short lifespan, these wasps routinely move pollen between trees 5–14 km apart (Nason *et al.*, 1998). This interdependence argues that viable fig populations may require a higher density of trees than is generally assumed necessary for plants not involved in such

Table 3 Pollination syndromes: putative characteristics of flowers associated with particular groups of pollinators

Animal	*Opening time*	*Color*	*Flower Odor*	*Shape*	*Symmetry*	*Nectar*
Entomophilous						
Beetles	Day/night	Dull or white	Fruity or aminoid	Flat or bowl-shaped	Radial	Often absent
Carrion or dung flies	Day/night	Brownish or greenish	Fetid	Flat or deep; often traps	Radial	Rich in amino acids, if present
Bee flies	Day/night	Variable	Variable	Moderately deep	Radial	Hexose-rich
Bees	Day/night	Variable, but not pure red	Sweet	Flat or broad tube	Radial or bilateral	Sucrose-rich for long-tongued bees; hexose-rich for short-tongued bees
Hawk moths	Night	White or pale green	Sweet	Deep, often with spur	Radial	Ample and sucrose-rich
Butterflies	Day/night	Variable; often pink	Sweet	Deep or with spur	Radial	Often sucrose-rich
Vertebrate-pollinated						
Bats	Night	Drab, pale; often green	Musty	Flat "shaving brush" or deep tube; often on branch or trunk; hanging; abundant pollen	Radial	Ample and hexose-rich
Birds	Day	Vivid; often red	None	Tube; often hanging	Radial or bilateral	Ample and sucrose-rich

Source: Adapted from Howe HF and Westley LC (1988) *Ecological Relationships of Plants and Animals*. New York: Oxford University Press.

a tight mutualism (Anstett *et al.*, 1997). The only study to test this contention, however, found that despite heavy forest fragmentation by humans, banyan trees (a type of fig population on the Cook Islands) have not yet suffered from Allee affects (Compton and McCormack, 1999).

Plants can sometimes escape the grips of their pollinator addiction. In one group of Hawaiian shrubs, species derived from insect-pollinated ancestors have evolved wind pollination, apparently by passing through a transitional stage characterized by a relatively generalized pollination system (Weller *et al.*, 1998).

Several researchers have questioned the assumption that most plant–pollinator relationships are highly specific (Jordano, 1987; Waser *et al.*, 1996). Even specialization in insect pollinators is usually defined at the level of plant genus or family, not species. Further, insect species confined to one plant species in one geographical area often visit other plant species in other parts of their range (Thompson, 1994). In fact, specialization by pollinators may be largely a function of which plants are available. Short-lived insects might appear to be specialized, because they visit only one or a few plant species with which they temporally co-occur, whereas social insects, which have long-lived colonies and thus the opportunity to overlap with many plant species, often exhibit serial specialization on a large variety of plant species. For example, the only bee species on the Galapagos Islands is colonial and has been recorded visiting flowers of at least 60 plant species in 28 families. Finally, clustering of flowers into certain categories thought to be canalized through selection by pollinators may actually reflect physiological or morphological constraints in plants. For example, Chittka *et al.* (1992) argued that clustering of flower colors into particular narrow ranges may be a function of the physical and chemical constraints imposed by plant pigments rather than constraints imposed by the vision systems of different pollinators. If these arguments are true, the pollinator syndrome concept may be clouding our thinking about and study of plant–pollinator interactions. Researchers may be oblivious to, or fail to record, flower visitation by the "wrong" pollinators (Waser *et al.*, 1996; Ollerton *et al.*, 2009).

Competition for Pollinators

Pollinators are an important resource that plants may compete over. Whether plants compete for pollinators is determined by the factors that limit seed production. In many situations, plant reproduction is limited by mineral resources and plants are unable to increase fruit set with increased pollinator availability. However, there are many examples of plants in which reproduction is limited by pollinator availability. Pollinator limitation has important implications for plant conservation. For example, prolific flowering by invading plant species may negatively impact native plant communities by depriving native species of pollination (Grabas and Laverty, 1999; Takakura *et al.*, 2009).

Fruit and Seed Dispersal: Childhood Mobility in Plants

Many plants depend on animal dispersal of seeds and fruits (diaspores). Unlike pollen dispersal, in which insects are

major players, diaspore dispersal is dominated by vertebrates. Dispersal is important because a plant often cannot develop successfully in the shade of its mother. Dispersal may also be an important mechanism by which progeny escape the predator or pathogen populations that are well adapted to exploit them, having built up on and become adapted to their parents (Augspurger and Kelly, 1984; Mangan *et al.*, 2010). Finally, dispersal allows plants to colonize newly opened habitat.

Types of Dispersers

Vertebrates

Fish Perhaps the most unexpected and amazing diaspore dispersers are fish. This phenomenon has been most extensively studied along the Amazon River (Kubitzki and Ziburski, 1994). However, it is likely to be important in many areas with extensive seasonal flooding (e.g., Horn, 1997). Floodplain plants with heavy fruits or seeds embedded within hard shells might require dispersal by fish. Catfish are effective dispersers whereas characins are destructive and act largely as predators of all but the smallest seeds.

Mammals Many mammals are important seed-dispersal agents. Primates and bats are the most important mammalian dispersers in tropical areas. Both types of animals can move quickly across the landscape, thereby dispersing diaspores to long distances. Diaspore dispersal by bats is particularly important for forest regeneration after land abandonment in the Neotropics (Parrotta *et al.*, 1997; Medellin and Gaona, 1999). In temperate regions, diaspores may be dispersed by ungulates (e.g., antelopes, elephants, and zebras) and by many supposed carnivores. For example, black bears consume prodigious quantities of fruit, sometimes competing with humans for delectable berries (McCloskey, 1948).

Birds and Reptiles The earliest known examples of animal dispersed plant propagules include the fleshy seeds of cycad progenitors, which might have been consumed by ancient reptiles (Burbidge and Whelan, 1982), although more data are needed to support this conjecture (Butler *et al.*, 2010; Manchester *et al.*, 2010). Many dinosaurs were certainly important fruit and seed eaters and may have functioned as dispersal agents. Their modern descendents, the birds, are arguably the most important class of fruit and seed dispersers today. Otherwise, modern reptiles are only rarely important diaspore dispersers.

Birds can be hard on seeds. Beaks may break up the seed coat, rendering the embryo vulnerable to digestive acids and enzymes. Seeds may be ground up in the gizzard. However, the guts of frugivorous birds tend to be short and gentle (Levey and Del Rio, 2001) and many seeds require a trip through a bird's digestive tract to germinate successfully (Samuels and Levey, 2005; Traveset *et al.*, 2008).

Although large numbers of a broad range of birds feed on fruits, few depend solely on fruits (Izhaki and Safriel, 1989). Even waxwings, perhaps the most specialized frugivores in the temperate region, also feed on insects when they are available. Nevertheless, fruits are an important resource with which many birds produce body fat prior to migration (Bairlein and Gwinner, 1994). Moreover, in the tropics, where seasonal constraints on fruit production may be weaker, several groups of birds depend almost exclusively on fruits (e.g., quetzals, toucans, and barbets).

Invertebrates

Ants The only major insect seed dispersers are ants. Myrmechorous seeds, those adapted to ant dispersal, often possess a starch- or lipid-rich body called an elaiosome attached to a tough and smooth seed coat that is difficult for ants to crack. Seed size is constrained by selection by ants – large ants tend to carry larger seeds than small ants.

In comparison with vertebrates, ants do not carry seeds very far (Kalisz *et al.*, 1999), but they can be important dispersers for many plants. Ants may either store seeds in the nest or remove the elaiosome and then discard seeds at the nest entrance or colony waste-pile (Servigne and Detrain; Beattie and Culver, 1982). Both locations tend to have well-aerated, nutrient-rich soil that can improve plant growth. Seeds collected by ants may also gain some protection from other seed predators, which generally avoid active ant nests (Heithaus, 1981).

An important conservation issue in many areas has been the loss of native seed-collecting ants to competition from invading ants. Argentine ants (*Linepithema humile*), which have invaded Mediterranean climates globally, collect fewer seeds than the native ants they outcompete, and are predicted to cause changes in plant diversity and ecosystem function (Rodriguez-Cabal *et al.*, 2009). In other situations, introduced ants seem to function as effective seed dispersers (Stuble *et al.*, 2010). However, they sometimes produce subtle differences. For example, invading fire ants (*Solenopsis invicta*) introduced into the Southeastern USA collect seeds of bloodroot (*Sanguinarea canadensis*) but alter the distance and direction that seeds are dispersed (Ness, 2004).

Other insects Occasionally seeds are dispersed by insects other than ants. For example, scarab beetles may bury seeds with dung (D'Hondt *et al.*, 2008). Grassland termites might also aid in dispersal of grass seeds (Vigni and Melati, 1999).

Plant Adaptations to Diaspore Dispersers

Dispersal syndromes Animal dispersers impose selection on fruit and seed characters and biologists have described "dispersal syndromes," suites of characters that appear to be shared among propagules sharing certain groups of animal dispersers or particular modes of transportation (**Table 4**). For example, most dog owners can describe the common characteristics of propagules dispersed on animal fur and feathers, which include barbs, hooks, and barbed hairs that cause these annoying passengers to attach firmly to socks as well as fur. However, as with pollination syndromes, controversy rages over whether these suites of traits result from selection by particular groups of animal species or instead reflect evolutionary constraints imposed by the morphology and physiology of plant ancestors (Jordano, 1987).

Plant Adaptations to Frugivores

A ripe fleshy fruit both attracts and rewards a potential seed disperser. During ripening, fruits often change color, from an

Table 4 Characteristics of fruits and seeds dispersed by different animals

Animal	*Color*	*Fruit Odor*	*Form*	*Reward*
Vertebrate dispersers				
Hoarding mammals	Brown	Weak or aromatic	Indehiscent thick-walled nuts	Seed itself
Hoarding birds	Green or brown	None	Rounded seeds or nuts	Seed itself
Arboreal mammals	Yellow, white, green, or brown	Aromatic	Arillate seeds or drupes; often compound and dehiscent	Pulp protein, sugar, or starch
Bats	Pale yellow or green	Musky	Various; often hanging	Pulp lipid- or starch-rich
Terrestrial mammals	Often green or brown	None	Indehiscent nuts, pods, or capsules	Pulp lipid- or starch-rich
Highly frugivorous birds	Black, blue, red, green, or purple	None	Large drupes or arillate seeds; often dehiscent; seeds >10 mm long	Pulp lipid- or starch-rich
Partly frugivorous birds	Black, blue, red, orange, or white	None	Small or medium-sized drupes, arillate seeds, or berries; seeds <10 mm long	Pulp often sugar- or starch-rich
Feathers or fur	Undistinguished	None	Barbs, hooks, or sticky hairs	None
Insect dispersers				
Ants	Undistinguished	None to humans	Elaiosome on seed coat; seed <3 mm long	Oil or starch elaiosome with chemical attractant

Source: Adapted from Howe HF and Westley LC (1988). *Ecological Relationships of Plants and Animals*. New York: Oxford University Press.

inconspicuous green that is poorly discernible amid the foliage to a contrasting color such as red, blue, yellow, or black, which is conspicuous to visually searching frugivores with color vision. Fruit color may reflect the types of animals being attracted, which suggests that a disperser's visual system imposes selection on fruit color. For example, primate-dispersed fruits tend to contrast with surrounding foliage less than do bird-dispersed fruits (Lomascolo and Schaefer, 2010).

The ripe fruit may contain various nutritious goodies that attract prospective dispersers, including sugars, minerals, water, lipids, and proteins. These resources can be costly for plants to provide, and some plants produce fruits or seeds that mimic more nutritious fruits but are much less nutritious than their model fruits. A plant might get away with this strategy, particularly if it produces large numbers of fruits, but animals with a choice generally choose the most nutritious fruits (Schaefer *et al*., 2003).

An important problem for the plant is to ensure survival of at least some of its seeds. The fruit must remain on the plant long enough for the seed to develop and be provisioned by the mother. Unripe fruits often contain toxins and palatability-inhibiting compounds (Stiles, 1989). Perhaps for some readers the most memorable example of this phenomenon will be the inadvertent bite into an unripe persimmon.

However, even ripe fruits must be protected from pathogens and predators. How to selectively protect the fruit? The chili pepper plant has an ingenious method. It produces capsaicin, which deters fungal pathogens and mammals, both of which destroy seeds, but has no effect on bird dispersers, which lack capsaicin receptors and hence do not experience chilies as "hot" (Tewksbury and Nabhan, 2001).

Occupying an attractive and nutritious fruit, seeds must possess traits that promote survival during fruit consumption and digestion. Sometimes seeds can be easily separated from the pulp and are discarded by frugivores prior to digestion (remember watermelon seed spitting contests). Some seeds require passage through a vertebrate gut before they will germinate (Traveset, 1998). Such seeds are relatively indigestible, due to either hard seed coats, toxins, or simply the incompetence of the frugivore gut, and pass through unharmed. Fruits may also contain chemicals that manipulate the animal's behavior or gut activity (Izhaki, 2002). A fast ride through the gut reduces the opportunity for digestive breakdown of the seed, but too rapid passage could land the seed right under the maternal plant.

Frugivores vary in where and how they deposit seeds. Those that discard seeds as they are eating often deposit seeds individually, but may also fail to move seeds any distance from the parent. Dispersers that regurgitate seeds may move them further and often deposit seeds individually. When seeds are defecated, however, they may be concentrated at high densities. Such dense concentrations of seeds may attract seed predators or secondary seed dispersers. Ants, for example, often remove seeds from dung piles (Levey and Byrne, 1993). If no secondary dispersal process spreads these seeds, they may germinate under extremely competitive growing conditions.

Frugivores function as seed dispersers only if they move away from the parent plant before defecating, regurgitating, or otherwise discarding the seed. Some fruit compounds sensitively manipulate the guts of dispersers in ways that might maximize or optimize dispersal distances. For example, glycoalkaloids of solanum fruits induce constipation in birds, function as laxatives in mammals, and deter seed predators (Izhaki, 2002). Mild toxins might also improve dispersal by deterring foraging frugivores from finishing all the fruits at a bonanza tree. Specialization on particular fruits is rare among frugivores, perhaps because no one species of fruit provides a complete diet. This, too, may be a mechanism that benefits plants because it encourages frugivores to move about to achieve a complete diet (Wahaj *et al*., 1998).

Seed Predators as Dispersers

Many diaspores lack a fleshy fruit or other enticement to attract frugivores. These seeds may nonetheless benefit from dispersal by their predators. Many rodent and bird seed

predators hoard seeds for future use. When they forget or lose buried seeds, they may be quite effective dispersal agents. Many members of the crow and jay family are important seed hoarders and dispersers. For example, Clark's nutcrackers carry pinon pine seeds up to 22 km (Vanderwall and Balda, 1977) and bury them in small clumps (scatter-hoards). Behavioral studies indicate that these birds are highly effective at finding hoarded seeds, and will even dig into a meter of snow to find known hoards. However, individual birds will commonly store more than twice as many seeds as needed, perhaps to protect themselves against theft. Theft may be common, especially among social seed predators, such as pinyon jays. Pinyon jays can find and exploit hoards that they observed other birds burying (Bednekoff and Balda, 1996).

Caching may have significant effects on the local ecosystem. Clark's Nutcracker initiates forest succession after large fires by moving in limber pine seeds from long distances (Coop and Schoettle, 2009). These birds can also influence the genetic structure of tree populations (Rogers *et al.*, 1999).

Seeds dispersed by predators sometimes have interesting adaptations that promote offspring survival. Oaks, for example, are dispersed by their seed predators, which include birds and rodents. Both types of animals crack open acorns and feed on the cotyledons of the embryo. During germination, the embryo root (radical) emerges from the apical (pointy) end of the acorn. Several studies by Steele *et al.* (1993) have discovered that digestibility-reducing astringent tannins are concentrated at the apical end of the acorn. Further, all acorn consumers studied, including insects, birds, and rodents, preferentially consume the basal (cap) end of the acorn, which leaves the embryo intact and viable. Partially eaten acorns can germinate, sometimes at higher rates than those of intact acorns.

Evolutionary Dead Ends?

As with other forms of specialized dependence, specialization to a particular propagule disperser can be an evolutionary dead end. Perhaps the most famous example involves the fruits of the *Calvaria major* tree on the island of Mauritius in the Indian Ocean. This tree had not been observed to recruit young seedlings for more than 300 years when Temple (1977) surmised that it was lacking its essential fruit disperser, which was the dodo, a large flightless pigeon that had gone extinct in the late 1600s. Temple was able to mimic dodo digestion by feeding the *Calvaria* seeds to turkeys.

Turkeys were not available to rescue *Calvaria* on Mauritius, but Janzen (1981) has argued that another tree has survived the loss of its dispersal agent in Central America through such a substitution. Huge, rare guanacaste (*Enterolobium cyclocarpum*) trees in Costa Rica produce fruits that are readily eaten by domesticated horses and cattle. Janzen speculated that these fruits were once dispersed by the Pleistocene horse (*Equus fraternus*), which went extinct in the New World some 10,000 years ago. He demonstrated that most seeds pass unharmed through the guts of domesticated horses, and went on to argue that these animals, which were introduced by Spanish conquistadors 500 years ago, have replaced the lost dispersal agent. Thus, adaptation to one seed dispersal agent preadapted the fruit to other similar agents.

Summary

Plant–animal interactions are ubiquitous and important. A common theme throughout the study of plant–animal interactions is the enormous effect that these interactions have on plant and animal evolution. There is strong evidence that the interaction between plants and insect pollinators is the primary driver of diversity in flowering plants and the groups of insects most involved in pollination. Selection by animal consumers has driven the evolution of numerous plant defense traits. These traits form the basis of many of the uses that we make of plants today. Plant-based fibers, pharmaceuticals, and flavorings all derive from plant evolutionary responses to consumers. The plant fitness trade-offs between these defensive traits and competitive ability also play an important role in determining the composition of biotic communities.

The primary benefit that plants obtain from animals is mobility. Many, perhaps most, plants depend on animals to transport pollen and propagules. In many cases, the interactions between plants and their animal transportation providers are highly specialized and mutually beneficial. These specialized mutualisms can be quite vulnerable to extinction of either party, which is an important issue in both plant and animal conservation.

Finally, plant photosynthesis converts solar energy into chemical energy and thereby provides the energetic basis for most of the world's terrestrial life. Plants are therefore the foundation of the global terrestrial ecosystem. Aside from the decomposition of plant litter through the microbial food chain, this energy flows into the global ecosystem via animal consumers of plants. Understanding how plant–animal interactions influence this process is crucial to understanding how intact ecosystems provide the goods and services on which human endeavor, and indeed all life, depends.

See also: Cooperation, Evolution of. Edible Plants. Evolution, Theory of. Megaherbivores. Parasitism. Plant Sources of Drugs and Chemicals. Pollinators, Role of

References

Abrahamson WG and Weis AE (1997) *Evolutionary Ecology across Three Trophic Levels: Goldenrods, Gallmakers, and Natural Enemies*. Princeton, NJ: Princeton University Press.

Adamec L (1997) Mineral nutrition of carnivorous plants: A review. *Botanical Review* 63: 273–299.

Anstett MC, Hossaert-McKey M, and McKey D (1997) Modeling the persistence of small populations of strongly interdependent species: Fig and fig wasps. *Conservation Biology* 11: 204–213.

Armbruster WS (1993) Evolution of plant pollination systems – hypotheses and tests with the neotropical vine *dalechampia*. *Evolution* 47: 1480–1505.

Augspurger CK and Kelly CK (1984) Pathogen mortality of tropical tree seedlings: Experimental studies of the effects of dispersal distance, seedling density, and light conditions. *Oecologia* 61: 211–217.

Bairlein F and Gwinner E (1994) Nutritional mechanisms and temporal control of migratory energy accumulation in birds. *Annual Review of Nutrition* 14: 187–215.

Baldwin BG (1997) Adaptive radiation of the Hawaiian silversword alliance: Congruence and conflict of phylogenetic evidence from molecular and non-molecular investigations. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 103–128. Cambridge, UK: Cambridge University Press.
Beattie AJ and Culver DC (1982) Inhumation: How ants and other invertebrates help seeds. *Nature* 297: 627.
Bednekoff PA and Balda RP (1996) Observational spatial memory in Clark's Nutcrackers and Mexican Jays. *Animal Behaviour* 52: 833–839.
Bernays E and Graham M (1988) On the evolution of host specificity in phytophagous arthropods. *Ecology* 69: 886–892.
Bertin RI (1989) Pollination biology. In: Abrahamson WG (ed.) *Plant–Animal Interactions*, pp. 23–86. New York: McGraw-Hill.
Brewer JS (1998) Effects of competition and litter on a carnivorous plant, drosera capillaris (droseraceae). *American Journal of Botany* 85: 1592–1596.
Brewer JS (1999a) Effects of competition, litter, and disturbance on an annual carnivorous plant (utricularia juncea). *Plant Ecology* 140: 159–165.
Brewer JS (1999b) Short-term effects of fire and competition on growth and plasticity of the yellow pitcher plant, sarracenia alata (sarraceniaceae). *American Journal of Botany* 86: 1264–1271.
Brooks DR (1979) Testing the context and extent of host-parasite coevolution. *Systematic Zoology* 28: 299–307.
Burbidge AH and Whelan RJ (1982) Seed dispersal in a cycad, macrozamia riedlei. *Australian Journal of Ecology* 7: 63–67.
Butler RJ, Barrett PM, Penn MG, and Kenrick P (2010) Testing coevolutionary hypotheses over geological timescales: Interactions between cretaceous dinosaurs and plants. *Biological Journal of the Linnean Society* 100: 1–15.
Chase JM, Leibold MA, Downing AL, and Shurin JB (2000) The effects of productivity, herbivory, and plant species turnover in grassland food webs. *Ecology* 81: 2485–2497.
Chittka L, Beier W, Hertel H, *et al.* (1992) The colour hexagon: A chromaticity diagram based on photoreceptor excitations as a generalized representation of colour opponency. *Journal of Comparative Physiology A* 170: 533–543.
Clark L and Mason JR (1988) Effect of biologically active plants used as nest material and the derived benefit to starling nestlings. *Oecologia* 77: 174–180.
Coley PD (1983) Herbivory and defensive characteristics of tree species in a lowland tropical forest. *Ecological Monographs* 53: 209–233.
Coley PD, Bryant JP, and Chapin III FS (1985) Resource availability and plant antiherbivore defense. *Science* 230: 895–899.
Compton SG and McCormack G (1999) The bacific banyan in the cook islands: Has its pollination and seed dispersal mutualistsm been disrupted, and does it matter. *Biodiversity and Conservation* 8: 1707–1715.
Connell JH (1971) On the role of natural enemies in preventing competitive exclusion in some marine animals and rain forest trees. In: Boer PJD and Gradwell G (eds.) *Dynamics of Populations*. Wageningen, The Netherlands: Centre for Agricultural Publishing and Documentation.
Coop JD and Schoettle AW (2009) Regeneration of rocky mountain bristlecone pine (*Pinus aristata*) and limber pine (pinus flexilis) three decades after stand-replacing fires. *Forest Ecology and Management* 257: 893–903.
Cornell HV (1983) The secondary chemistry and complex morphology of galls formed by the cynipinae (hymenoptera): Why and how. *American Midland Naturalist* 110: 225–234.
Crepet WL (2008) The fossil record of angiosperms: Requiem or renaissance. *Annals of the Missouri Botanical Garden* 95: 3–33.
van Dam NM and Hare JD (1998) Differences in distribution and performance of two sap-sucking herbivores on glandular and non-glandular datura wrightii. *Ecological Entomology* 23: 22–32.
Darwin C (1875) *Insectivorous Plants*. London: John Murray.
Darwin C (1877a) *The Various Contrivances by Which British and Foreign Orchids are Fertilised by Insects* (revised version in 1984). Chicago, IL: University of Chicago Press.
Darwin F (1877b) Experiments on the nutrition of *Drosera rotundifolia*. *Journal of the Linnean Society – Botany* 17: 17–32.
Davidson DW, Longino JT, and Snelling RR (1988) Pruning of host plant neighbors by ants: An experimental approach. *Ecology* 69: 801–808.
Davidson DW and McKey D (1993) The evolutionary ecology of symbiotic ant–plant relationships. *Journal of Hymenoptera Research* 2: 13–83.
Dearing MD (1997) The manipulation of plant toxins by a food-hoarding herbivore, *Ochotona princeps*. *Ecology* 78: 774–781.
D'Hondt B, Bossuyt B, Hoffmann M, and Bonte D (2008) Dung beetles as secondary seed dispersers in a temperate grassland. *Basic and Applied Ecology* 9: 542–549.
Dicke M, Sabelis MW, Takabayashi J, *et al.* (1990) Plant strategies of manipulating predator–prey interactions through allelochemicals – prospects for application in pest-control. *Journal of Chemical Ecology* 16: 3091–3118.
Douglas AE (1994) *Symbiotic Interactions*. New York: Oxford University Press.
Duffy JE and Hay ME (1991) Food and shelter as determinants of food choice by an herbivorous marine amphipod. *Ecology* 72: 1286–1298.
Edmunds GF and Alstad DN (1978) Coevolution in insect herbivores and conifers. *Science* 199: 941–945.
Ehrlich PR and Raven PH (1964) Butterflies and plants: A study in coevolution. *Evolution* 18: 586–608.
Ellis AG and Midgley JJ (1996) A new plant–animal mutualism involving a plant with sticky leaves and a resident hemipteran insect. *Oecologia* 106: 478–481.
Endress PK (2011) Evolutionary diversification of the flowers in angiosperms. *American Journal of Botany* 98: 370–396.
Farnsworth NR, Akerele O, Bingel AS, *et al.* (1985) Medicinal plants in therapy. *Bulletin of the World Health Organization* 63: 965–981.
Folgarait PJ, Johnson HL, and Davidson DW (1994) Responses of cecropia to experimental removal of mullerian bodies. *Functional Ecology* 8: 22–28.
Fonseca CR (1994) Herbivory and the long-lived leaves of an amazonian ant-tree. *Journal of Ecology* 82: 833–842.
Friis EM, Chaloner WG, and Crane PR (1987) *The Origins of Angiosperms and Their Biological Consequences*. Cambridge, UK: Cambridge University Press.
Gallie DR and Chang SC (1997) Signal transduction in the carnivorous plant *Sarracenia purpurea* – regulation of secretory hydrolase expression during development and in response to resources. *Plant Physiology* 115: 1461–1471.
Gaume L, McKey D, and Terrin S (1998) Ant–plant–homopteran mutualism: How the third partner affects the interaction between a plant-specialist ant and its myrmecophyte host. *Proceedings of the Royal Society of London Series B-Biological Sciences* 265: 569–575.
Givnish TJ (1989) Ecology and evolution of carnivorous plants. In: Abrahamson WG (ed.) *Plant–Animal Interactions*, pp. 243–290. New York: McGraw-Hill.
Givnish TJ, Burkhardt EL, Happel RE, and Weintraub JD (1984) Carnivory in the Bromliad *Brocchinia reducta*, with a cost–benefit model for the general restriction of carnivorous plants to sunny, moist, nutrient-poor habitats. *American Naturalist* 124: 479–497.
Grabas GP and Laverty TM (1999) The effect of purple loosestrife (*lythrum salicaria* L., lythraceae) on the pollination and reproductive success of sympatric co-flowering wetland plants. *Ecoscience* 6: 230–242.
Grant V and Grant KA (1965) *Flower Pollination in the Phlox Family*. New York: Columbia University Press.
Groom MJ (1998) Allee effects limit population viability of an annual plant. *American Naturalist* 151: 487–496.
Grover JP (1994) Assembly rules for communities of nutrient-limited plants and specialist herbivores. *American Naturalist* 143: 258–282.
Grover JP (1995) Competition, herbivory, and enrichment: Nutrient-based models for edible and inedible plants. *American Naturalist* 145: 746–774.
Gündüz EA and Douglas AE (2009) Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proceedings of the Royal Society B-Biological Sciences* 276: 987–991.
Hairston NG, Smith FE, and Slobodkin LB (1960) Community structure, population control, and competition. *American Naturalist* 94: 421–425.
Hanslin HM and Karlsson PS (1996) Nitrogen uptake from prey and substrate as affected by prey capture level and plant reproductive status in four carnivorous plant species. *Oecologia* 106: 370–375.
Harvey AL (2008) Natural products in drug discovery. *Drug Discovery Today* 13: 894–901.
Heithaus ER (1981) Seed predation by rodents on three ant-dispersed plants. *Ecology* 62: 136–145.
Hodges SA (1997) Floral nectar spurs and diversification. *International Journal of Plant Science* 158(6 Supplement): S81–S88.
Hodges SA and Arnold ML (1994) Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *Proceedings of the National Academy of Sciences of the United States of America* 91: 2493–2496.
Holt RD, Grover J, and Tilman D (1994) Simple rules for interspecific dominance in systems with exploitative and apparent competition. *American Naturalist* 144: 741–771.
Horn MH (1997) Evidence for dispersal of fig seeds by the fruit-eating characid fish *brycon guatemaliensis* Regan in a Costa Rican tropical rain forest. *Oecologia (Berlin)* 109: 259–264.
Howe HF and Westley LC (1988) *Ecological Relationships of Plants and Animals*. New York: Oxford University Press.
Inouye DW (1980) The terminology of floral larceny. *Ecology* 61: 1251–1253.

Izhaki I (2002) Emodin – a secondary metabolite with multiple ecological functions in higher plants. *New Phytologist* 155: 205–217.
Izhaki I and Safriel UN (1989) Why are there so few exclusively frugivorous birds. *Experiments on Fruit Digestibility Oikos* 54: 23–32.
Janzen DH (1966) Coevolution of mutualism between ants and acacias in central America. *Evolution* 20: 249–275.
Janzen DH (1970) Herbivores and number of tree species in tropical forests. *American Naturalist* 104: 501–528.
Janzen DH (1980) When is it coevolution? *Evolution* 34: 611–612.
Janzen DH (1981) Guanacaste tree seed-swallowing by Costa Rican range horses. *Ecology* 62: 587–592.
Jermy T (1976) Insect–host–plant relationship – co-evolution or sequential evolution. *Symposia Biolologica Hungarica* 16: 109–113.
Johnson SD and Steiner KE (1997) Long-tongued fly pollination and evolution of floral spur length in the *Disa draconis* complex (orchidaceae). *Evolution* 51: 45–53.
Jordano P (1987) Patterns of mutualistic interactions in pollination and seed dispersal: Connectance, dependence asymmetries, and coevolution. *American Naturalist* 129: 657–677.
Kalisz S, Hanzawa FM, Tonsor SJ, *et al.* (1999) Ant-mediated seed dispersal alters pattern of relatedness in a population of trillium grandiflorum. *Ecology* 80: 2620–2634.
Karban R and EnglishLoeb G (1997) Tachinid parasitoids affect host plant choice by caterpillars to increase caterpillar survival. *Ecology* 78: 603–611.
Kubitzki K and Ziburski A (1994) Seed dispersal in flood plains of Amazonia. *Biotropica* 26: 30–43.
Labandeira CC (1998) Early history of arthropod and vascular plant associations. *Annual Review of Earth and Planetary Sciences* 26: 329–377.
Labandeira CC (2006) The four phases of plant–arthropod associations in deep time. *Geological Acta* 4: 409–438.
Labandeira CC (2010) The pollination of mid mesozoic seed plants and the early history of long-proboscid insects. *Annals of the Missouri Botanical Garden* 97: 469–513.
Labandeira CC, Dilcher DL, Davis DR, and Wagner DL (1994) 97-million years of angiosperm–insect association – paleobiological insights into the meaning of coevolution. *Proceedings of the National Academy of Sciences of the United States of America* 91: 12278–12282.
Labandeira CC and Sepkoski JJ (1993) Insect diversity in the fossil record. *Science* 261: 310–315.
Lavin M, Herendeen PS, and Wojciechowski MF (2005) Evolutionary rates analysis of leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology* 54: 575–594.
Leibold MA (1996) A graphical model of keystone predators in food webs: Trophic regulation of abundance, incidence, and diversity patterns in communities. *American Naturalist* 147: 784–812.
Leibold MA (1999) Biodiversity and nutrient enrichment in pond plankton communities. *Evolutionary Ecology Research* 1: 73–95.
Leibold MA, Chase JM, Shurin JB, and Downing AL (1997) Species turnover and the regulation of trophic structure. *Annual Review of Ecology and Systematics* 28: 467–494.
Levey DJ and Byrne MM (1993) Complex ant–plant interactions: Rain-forest ants as secondary dispersers and posdispersal seed predators. *Ecology* 74: 1802–1812.
Levey DJ and Del Rio CM (2001) It takes guts (and more) to eat fruit: Lessons from avian nutritional ecology. *Auk* 118: 819–831.
Lindeman RL (1942) The trophic-dynamic aspect of ecology. *Ecology* 23: 399–418.
Lomascolo SB and Schaefer HM (2010) Signal convergence in fruits: A result of selection by frugivores. *Journal of Evolutionary Biology* 23: 614–624.
Manchester SR, Lehman TM, and Wheeler EA (2010) Fossil palms (arecaceae, coryphoideae) associated with juvenile herbivorous dinosaurs in the upper cretaceous aguja formation, big bend national park, texas. *International Journal of Plant Sciences* 171: 679–689.
Mangan SA, Schnitzer SA, Herre EA, *et al.* (2010) Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* 466: 752–755.
McCloskey R (1948) *Blueberries for Sal.* New York: Viking Press.
Medellin RA and Gaona O (1999) Seed dispersal by bats and birds in forest and disturbed habitats of Chiapas, Mexico. *Biotropica* 31: 478–485.
Midgley JJ and Stock WD (1998) Natural abundance of delta *n*-15 confirms insectivorous habit of *Roridula gorgonias*, despite it having no proteolytic enzymes. *Annals of Botany* 82: 387–388.
Nason JD, Herre EA, and Hamrick JL (1998) The breeding structure of a tropical keystone plant resource. *Nature* 391: 685–687.
Ness JH (2004) Forest edges and fire ants alter the seed shadow of an ant-dispersed plant. *Oecologia* 138: 448–454.
Nilsson LA (1988) The evolution of flowers with deep corolla tubes. *Nature* 334: 147–149.
Oksanen L, Fretwell SD, Arruda J, and Niemela P (1981) Exploitation ecosystems in gradients of primary productivity. *American Naturalist* 118: 240–261.
Ollerton J, Alarcon R, Waser NM, *et al.* (2009) A global test of the pollination syndrome hypothesis. *Annals of Botany* 103: 1471–1480.
Owen TP, Lennon KA, Santo MJ, and Anderson AN (1999) Pathways for nutrient transport in the pitchers of the carnivorous plant nepenthes alata. *Annals of Botany* 84: 459–466.
Paine RT (1966) Food web complexity and species diversity. *American Naturalist* 100: 65–75.
Paine TD, Raffa KF, and Harrington TC (1997) Interactions among Scolytid bark beetles, their associated fungi, and live host conifers. *Annual Review of Entomology* 42: 179–206.
Parrotta JA, Knowles OH, and Wunderle JM (1997) Development of floristic diversity in 10-year-old restoration forests on a bauxite mined site in Amazonia. *Forest Ecology and Management* 99: 21–42.
Peakall R and Beattie AJ (1996) Ecological and genetic consequences of pollination by sexual deception in the orchid C*aladenia tentaculata*. *Evolution* 50: 2207–2220.
Roderick GK (1997) Herbivorous insects and the Hawiaiian silversword alliance: Coevolution or cospeciation. *Pacific Science* 51: 440–449.
Rodriguez-Cabal MA, Stuble KL, Nunez MA, and Sanders NJ (2009) Quantitative analysis of the effects of the exotic argentine ant on seed-dispersal mutualisms. *Biology Letters* 5: 499–502.
Rodriquez E and Wrangham R (1993) Zoopharmacognosy: The use of medicinal plants by animals. In: Downum KR, Romeo JT, and Stafford HA (eds.) *Phytochemical Potential of Tropical Plants*. New York: Plenum.
Rogers DL, Millar CI, and Westfall RD (1999) Fine-scale genetic structure of whitebark pine (*Pinus albicaulis*): Associations with watershed and growth form. *Evolution* 53: 74–90.
Root RB (1973) Organization of a plant-arthropod association in simple and diverse habitats: Fauna of collards (*Brassica oleracea*). *Ecological Monographs* 43: 95–120.
Samuels IA and Levey DJ (2005) Effects of gut passage on seed germination: Do experiments answer the questions they ask. *Functional Ecology* 19: 365–368.
Schaefer HM, Schmidt V, and Winkler H (2003) Testing the defence trade-off hypothesis: How contents of nutrients and secondary compounds affect fruit removal. *Oikos* 102: 318–328.
Schmitz OJ (1994) Resource edibility and trophic exploitation in an old-field food-web. *Proceedings of the National Academy of Sciences of the United States of America* 91: 5364–5367.
Schmitz OJ (1997) Press perturbations and the predictability of ecological interactions in a food web. *Ecology* 78: 55–69.
Servigne P and Detrain C (2010) Opening myrmecochory's black box: What happens inside the ant nest. *Ecological Research* 25: 663–672.
Shorthouse JD, Wool D, and Raman A (2005) Gall-inducing insects – nature's most sophisticated herbivores. *Basic and Applied Ecology* 6: 407–411.
Simms EL (1992) Costs of plant resistance to herbivory. In: Fritz RS and Simms EL (eds.) *Plant Resistance to Herbivores and Pathogens: Ecology, Evolution, and Genetics*, pp. 392–425. Chicago, IL: University of Chicago Press.
Simms EL and Triplett JK (1994) Costs and benefits of plant responses to disease: Resistance and tolerance. *Evolution* 48: 1973–1985.
Specht CD and Bartlett ME (2009) Flower evolution: The origin and subsequent diversification of the angiosperm flower. *Annual Review of Ecology Evolution and Systematics* 40: 217–243.
Stanton ML, Palmer TM, Young TP, *et al.* (1999) Sterilization and canopy modification of a swollen thorn acacia tree by a plant-ant. *Nature* 401: 578–581.
Stapley L (1998) The interaction of thorns and symbiotic ants as an effective defence mechanism of swollen-thorn acacias. *Oecologia* 115: 401–405.
Steele MA, Knowles T, Bridle K, and Simms EL (1993) Tannins and partial consumption of acorns: Implications for dispersal of oaks by seed predators. *American Midland Naturalist* 130: 229–238.
Stiles EW (1989) Fruits, seeds, and dispersal agents. In: Abrahamson WG (ed.) *Plant–Animal Interactions*, pp. 87–122. New York: McGraw-Hill.
Stowe KA, Marquis RJ, Hochwender CG, and Simms EL (2000) The evolutionary ecology of tolerance to consumer damage. *Annual Review of Ecology and Systematics* 31: 565–595.
Stuble KL, Kirkman LK, and Carroll CR (2010) Are red imported fire ants facilitators of native seed dispersal. *Biological Invasions* 12: 1661–1669.

Tack AJM and Roslin T (2010) Overrun by the neighbours: Landscape context affects strength and sign of local adaptation. *Ecology* 91: 2253–2260.
Takakura K, Nishida T, Matsumoto T, and Nishida S (2009) Alien dandelion reduces the seed-set of a native congener through frequency-dependent and one-sided effects. *Biological Invasions* 11: 973–981.
Taylor TN and Scott AC (1983) Interaction of plants and animals during the carboniferous. *Bioscience* 23: 488–493.
Temple SA (1977) Plant–animal mutualism: Coevolution with dod leads to near extinction of plant. *Science* 197: 885–886.
Tewksbury JJ and Nabhan GP (2001) Seed dispersal – directed deterrence by capsaicin in chillies. *Nature* 412: 403–404.
Thompson JN (1994) *The Coevolutionary Process*. Chicago, IL: University of Chicago Press.
Traveset A (1998) Effects of seed passage through vertebrate frugivores' guts on germination: A review. *Perspectives in Plant Ecology Evolution, and Systematics* 1: 151–190.
Traveset A, Rodriguez-Perez J, and Pias B (2008) Seed trait changes in dispersers' guts and consequences for germination and seedling growth. *Ecology* 89: 95–106.
Vanderwall SB and Balda RP (1977) Co-adaptations of the Clarks' Nutcracker and the pinon pine for efficient seed harvest and dispersal. *Ecological Monographs* 47: 89–111.
Vigni IL and Melati MR (1999) Examples of seed dispersal by entomochory. *Acta Botanica Gallica* 146: 145–156.
Wahaj SA, Levey DJ, Sanders AK, and Cipollini ML (1998) Control of gut retention time by secondary metabolites in ripe solanum fruits. *Ecology* 79: 2309–2319.
Waser NM, Chittka L, Price MV, *et al.* (1996) Generalization in pollination systems, and why it matters. *Ecology* 77: 1043–1060.
Weller SG, Sakai AK, Rankin AE, *et al.* (1998) Dioecy and the evolution of pollination systems in Scheideai and Alsinidendron (caryophyllaceae: Alsinoidea) in the Hawaiian islands. *American Journal of Botany* 85: 1377–1388.
Whittall JB and Hodges SA (2007) Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature* 447: 706–709.
Williams NH and Whitten WM (1983) Orchid floral fragrances and male euglossine bees: Methods and advances in the last sesquidecade. *Biological Bulletin* 164: 355–395.
Zamora R (1999) Conditional outcomes of interactions: The pollinator–prey conflict of an insectivorous plant. *Ecology* 80: 786–795.

Plant Biodiversity, Overview

Jeannette Whitton, University of British Columbia, Vancouver, Canada

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Jeannette Whitton, Nishanta Rajakaruna, volume 4, pp 621–630, © 2001, Elsevier Inc.

Glossary

Alternation of generations In plants, a reproductive cycle in which a multicellular gametophytic phase alternates with a multicellular sporophytic phase to complete the life cycle.

Autotroph An organism that obtains its nutrition by synthesizing organic substances from inorganic substances obtained from its environment.

Gametophyte In an organism with an alternation of generations, the haploid, gamete-producing phase.

Self-incompatibility A genetically based system that prevents fertilization of an egg by pollen that shares one or more alleles at a self-incompatibility locus. A mechanism to prevent self-fertilization in plants.

Sporophyte In an organism with an alternation of generations, the diploid, spore-producing phase.

Introduction

Defining Plant Biodiversity

Land Plants, also known as Embryophytes, include mosses, ferns, conifers, flowering plants, and related lineages. According to this circumscription, plants are mostly autotrophs that rely on photosynthesis, but some lineages include a small number of derived heterotrophs. Their life cycles include an alternation of generations, involving multicellular diploid and haploid phases of varying anatomical complexity, the sporophyte and gametophyte, respectively. The relative duration of these phases and the degree to which each phase is physiologically independent from the other differs among and, to a lesser degree, within the major plant lineages.

The land plants are variously divided into roughly a dozen major groups, and while consensus is building about the boundaries and membership of these groups and their relationships to one another, a consensus about the taxonomic rank of each group has yet to be achieved, and informal names remain in use. Informally, plant diversity is divided into four groups that include 10–12 extant lineages (Raven *et al.*, 2005). These are the nonvascular plants or bryophytes (mosses, liverworts, and hornworts), the seedless vascular plants (clubmosses and ferns including, horsetails, club mosses, and whisk ferns), gymnosperms (conifers, cycads, Ginkgo, and gnetophytes), and angiosperms, or flowering plants. Although some groupings are known or suspected to be nonnatural assemblages (i.e., the set of phyla described lineages that comprise an informal group do not necessarily share a unique common ancestor), life history and ecological characteristics unite their component members, and thus, we use these designations to structure the presentation of characteristics of plant biodiversity outlined below (*see* Plant Biodiversity Described).

Estimates of total plant diversity, including described and undescribed species vary by more than 50% because of two challenges: (1) Figuring out how many of the published names correspond to recognized, "good" species and (2) Estimating how much plant diversity remains undescribed or undiscovered. Recent estimates of named plant species include roughly 380,000 species (Paton *et al.*, 2008), and an additional 10–20% of flowering plant species are thought to remain undiscovered (Joppa *et al.*, 2011). Because flowering plants represent more than 90% of all plants, it seems reasonable to extrapolate and suggest that a full catalog of current global land plant diversity would include perhaps 418,000–456,000 species. As noted, the great majority of this diversity occurs within flowering plants (approximately 352,000 species), followed by about 16,000 species of bryophytes, and 12,000 species of ferns and allies. Taxonomic knowledge of various groups is not equal, so while the catalog of conifer species may be nearly complete, the estimate for mosses may require substantial revision, as their diversity becomes better understood.

Plants as Units of Biodiversity, as Resources, and as Habitat

In nearly all terrestrial habitats, plants form the dominant features of the landscape. Satellite images reveal that most of what covers the land is plant biomass. Their dominance hints at the critical functional role that plants play in the global ecosystem; yet as living organisms, plant species also constitute units of biodiversity in their own right, and thus consideration of their unique biological characteristics, habitat requirements, and conservation status should be an integral component of research, discussion, and policy affecting local, regional, and global biodiversity issues. In the context of a reference on biodiversity, this may seem an unnecessary statement, however, vertebrate species continue to be the dominant focus of much of our conservation effort, with plants often relegated to consideration as habitat, food, or fixed carbon.

The complex web of interactions that constitutes our global ecosystem critically depends on marine and terrestrial photosynthetic organisms, including the members of the plant kingdom. Photosynthetic organisms are the planet's primary

producers, harvesting light energy for their own growth, fixing carbon from the atmosphere into organic molecules, releasing oxygen to the environment, and ultimately sustaining life on earth. Globally, marine and terrestrial photosynthetic organisms fix approximately 100 billion metric tons of carbon annually and play a key role in maintaining the balance between fixed and atmospheric carbon, with roughly half of this amount being fixed by marine, and half by terrestrial photosynthetic organisms. Over the last decade, the impact of human activities on global carbon balance, and the cascade of impacts of ongoing climate change on ecosystems have become recognized as the leading environmental crisis facing our planet. In light of this, it is increasingly important that the potential for carbon fixation by photosynthetic organisms be maintained at the very least. Many human activities that threaten plant biodiversity, such as deforestation and habitat destruction or deterioration, also may decrease plant biomass, further exacerbating the carbon crisis. In consideration of global carbon cycles, vegetation is a central focus, and the emphasis tends to be on biomass and carbon sequestration capacity rather than biodiversity. While this perspective is reasonable given the critical role that plants play in this context, it is important to encompass all major plant groups within the framework of biodiversity policy, and practice. Therefore, in the presentation that follows, the biological properties of each group of plants are presented, with emphasis on those features that may impact issues relevant to discussions of biodiversity.

Plants and People

Human life and indeed all life on Earth ultimately depend on photosynthetic life forms. We depend on plants as sources of food for ourselves and for livestock, as sources of medicine, as raw materials for textiles and building materials, and as heating and cooking fuel.

The history of human civilization is inextricably linked to the history of agriculture, including the development of crops and livestock as food sources. It is estimated that only about 3000 flowering plant species have been used by people as food, and only 150 or so of these have been cultivated to any extent. Today, just six crops (wheat, rice, corn, potatoes, sweet potatoes, and manioc) directly or indirectly provide more than 80% of the human population's calories. The world's more highly industrialized regions depend on high energy input agriculture, in which crops are largely grown as genetically uniform monocultures. Such cultivation practices render these pure stands susceptible to pest and pathogen attack, and it is estimated that one-third of all crops are destroyed by pests.

Most of the global human population, however, lives in tropical and subtropical regions where relatively low input agriculture is most important. Further improvements in agricultural yields in these regions will have the greatest impact on global human welfare, where improving yields, nutritional quality, and disease resistance could greatly increase overall productivity.

Sources of wild germplasm remain of critical importance for future crop development and improvements required to feed the growing human population. Conserving the wild relatives of our crop plants has become a primary concern of plant biologists the world over. Most plants are capable of interbreeding to some degree with wild relatives, and thus, traditional breeding can be used to introduce beneficial traits into cultivars. With genetic engineering tools, the ability to interbreed is no longer required to introduce traits into cultivars, thus removing limitations on species that can potentially be used as sources of genetic material. In addition, given the small fraction of plants that have been exploited as food sources by humans, the totality of plant biodiversity can be viewed as a potential source of new crops.

In addition to the major food crops described above, plants also make their way into our diets and our cultures as sources of beverages such as coffee and tea and in the form of spices, condiments, and oils.

Another use of plants that has and will likely continue to affect human lives is as sources of medicinal compounds. It is estimated that between 35,000 and 70,000 different species have been used as medicines by various peoples of the world and 80% of the global population derive almost all of their medicines exclusively from plants. Plants are extraordinary chemical factories: they produce a wide array of alkaloids, glycosides, and saponins that are the principal source of drugs used to prevent and cure illness in both traditional and Western medicines. Modern Western pharmacopoeia have derived approximately 7000 different medical compounds from botanical sources, and 120 or so of these are presently widely prescribed. Included among these are antimicrobial agents such as quinine, used in the treatment of malaria, and analgesics, including the alkaloids morphine and codeine, both derived from the opium poppy, *Papaver somniferum*. Glycosides produced by species of *Digitalis* have enormous value for the treatment of heart ailments, both in improving the action of a failing heart and in reducing a dangerously fast heartbeat. Plants are also sources of anticancer agents. Among these is the well-known rosy periwinkle (*Catharanthus roseus*) of Madagascar, which is the source of alkaloids used in the treatment of both Hodgkin's disease and acute lymphocytic leukemia. More recently, taxol, a drug originally derived from extracts of the Pacific yew (*Taxus brevifolius*), has become a key tool in the treatment of ovarian and breast cancers. These are just a few of the thousands of plant products that have been or are presently used to cure human diseases. There is no doubt that more species have immense untapped potential. Since tests for determining pharmaceutical potential are both time-consuming and expensive, and synthetic drugs continue to be unaffordable to a large proportion of the global population, it is likely that the dependence on local plant-based cures will continue.

The qualities of flexibility, durability, and strength, along with ready availability, make plant material such as wood an ideal source of building material. Wood and related products, including paper, are of enormous commercial importance. In industrialized nations these commodities account for a significant proportion of the total national consumption of goods. Plant species such as pine, spruce, hemlock, fir, redwoods, birch, beech, and oak are used as building materials in the temperate regions while mahogany, species of the family Dipterocarpaceae, bamboo, and rattan are used in the tropics, but are increasingly available in temperate regions as well.

A wide variety of other materials are also derived from plants, including the fibers used to weave cotton, linen, and jute, as well as latexes, resins, perfumes, and dyes. Finally, plants are also used by humans for their esthetic qualities. Ornamental plants in all their color and shape add beauty to our surroundings and are used in both public and private gardens, along roadsides in our cities, and as houseplants. In an increasingly urban environment, ornamental plants often constitute a rare link with the natural world.

In recent years, interest in the use of plants as sources of biofuels has risen, because such sources are renewable. However, production of biofuels requires high energy inputs and is likely to decrease the land area available for crop food production (and carbon sequestration), and thus the potential for these products to contribute to decreasing dependency on fossil fuels is mitigated by concerns about environmental impacts and net energy gains.

It is critical that while we exploit plants to satisfy our needs we continue to conserve and restore their natural habitats, if for no other reason than to preserve the homes that undoubtedly hold a wealth of other species with the potential to serve the needs of a growing human population.

Plant Biodiversity Described

Nonvascular Plants

Nonvascular plants (often referred to collectively as the bryophytes) include three groups: the mosses (Bryophyta), approximately 15,000 species; liverworts (Hepaticophyta), approximately 7500 species; and hornworts (Anthocerophyta), approximately 250 species (**Table 1**). These three groups are characterized by their small stature, by the absence of specialized conducting tissues (i.e., xylem and phloem), which occur in other groups of plants, and by their life cycle. The conspicuous, dominant phase of the life cycle of these plants is the gametophyte (haploid) generation. Gametes are formed in specialized structures on the gametophyte called antheridia (containing male gametes) and archegonia (containing female gametes). Antheridia and archegonia may occur on the same gametophyte or on separate male and female gametophytes. Male gametes in the nonvascular plants have motile sperm and are dependent on water for dispersal to the female gametes, which are nonmotile eggs retained in the archegonia on the gametophyte. Dispersal of gametes is thought to be extremely localized (less than 1 m). After fertilization, the diploid phase of the life cycle, the sporophyte, grows out of the archegonium, where it remains attached throughout its development, dependent on the gametophyte for most or all of its nutrition. Meiosis occurs in the sporangia, specialized structures of the sporophyte, resulting in the formation of haploid spores that are dispersed primarily by air currents. Spores are extremely small, tolerant of environmental extremes, and thus have the capacity to disperse over much greater distances than gametes. Many nonvascular plants also have the capacity to reproduce asexually as gametophytes, either by fragmentation or via specialized propagules called gemmae.

Ecologically, the three groups of nonvascular plants share their dependency on water for sperm dispersal, and thus bryophytes tend to occur where water is at least seasonally available and/or tends to accumulate. Among the richest habitats for bryophyte diversity are temperate and tropical cool, moist forests, and arctic and alpine habitats. Nonvascular plants grow on a variety of substrates, including soil and bare rock, and on other plants as epiphytes. On bare rock, they play an important role as initial colonizers. Peatlands are dominated by mosses in the genus *Sphagnum*. These unique and important ecosystems, while covering about 3% of the Earth's surface, represent a significant global carbon sink, estimated to contain 25–30% of the organic carbon stored in terrestrial soils. Although peatlands tend to have low net primary production (NPP) relative to many other terrestrial ecosystems, their success in carbon sequestration results from their slow rates of biomass loss through decomposition.

More than 60% of bryophyte families have worldwide distributions, occurring on all continents. Bryophyte species also tend to have wide geographic distributions: species that occur on more than one continent are common. Two processes have contributed to establishing present-day distributions. First, much of the diversity that we see today may be ancient, and present-day distributions may be the result of vicariance, or the splitting of a larger, once continuous range into smaller isolates. Second, especially in taxa that have originated more recently, present distributions are likely the result of long-distance spore dispersal. The role of ongoing dispersal in bryophytes is relevant to biodiversity research, because it has implications for structuring of genetic variation in these species. As suitable habitats become fragmented or

Table 1 The major groups of plants

Group	*Estimated number of species*	*Percentage of globally threatened species*
Hornworts	250	Not available
Liverworts	7500	Not available
Mosses	15,000	Not available
Club mosses (Lycophytes)	1000	30.8
Ferns	14,000	7.5
Cycads	283	82.8
Gnetophytes	65	2.5
Ginkgo	1	100
Conifers	630	55.8
Flowering plants (Angiosperms)	352,000	13.9

patchily distributed as a result, for example, of clear-cut logging on small and large scales, the potential for recolonization by extirpated bryophytes depends critically on dispersal and establishment, which are poorly characterized in these groups. Although we know that spores are easily carried by air currents, we know little about the factors that would limit recolonization, and only a handful of studies have examined patterns of gene flow in bryophytes. Another consideration in bryophytes that may be relevant to biodiversity is the relationship of taxonomic species described primarily on morphological and ecological criteria to biological species – units that are united by gene flow. The reduced morphology and absence of variation in traits related to gene flow are reasons to suspect greater decoupling of taxonomic species and biological species among bryophytes. As a consequence of our lack of understanding of evolutionary processes in these groups, a particular taxonomic species that may be considered common could comprise several cryptic, genetically isolated units, some of which may be critically endangered. Compounding this problem is the relative rarity of expertise in identification of bryophytes. While both academic and nonacademic experts contribute to our knowledge of flowering plant biodiversity, in many areas, no local experts in bryophyte identification, professional or amateur, are found. Although the bryoflora of a number of regions is well characterized, the bryophytes are often far less thoroughly studied than vascular plants in the same region. One estimate suggests that nearly 30% of the global bryophyte flora remains undescribed (compared to the 10–20% unknown for flowering plants noted above). This shortage of expertise is reflected in the status of conservation initiatives in the bryophytes. As of 2011, the Red List for bryophytes consists of a list of 80 globally threatened bryophytes, of the 101 bryophytes assessed to date. Therefore, it is difficult to say what percentage of the world's bryophytes is at risk, or to make broad statements about global priorities for bryophyte conservation.

Seedless Vascular Plants

Seedless vascular plants include two major groups of plants: the Lycophytes (about 1000 species, including the club mosses and the genera *Selaginella* and *Isoetes*), and the ferns and allies (about 14,000 species; **Table 1**), including the horsetails (*Equisetum*, represented by about 15 species), whisk ferns, and other ferns. These lineages share a number of characteristics. They are the earliest set of lineages to have evolved vascular tissue (xylem and phloem). Vascular tissues provide structural stability that allows vascular plants (seedless vascular plants, gymnosperms, and angiosperms) to grow much larger than the bryophytes. The seedless vascular plants also lack seed formation, which distinguishes them from gymnosperms and angiosperms considered below.

Like the nonvascular plants, the seedless vascular plants have a free-living gametophyte generation that forms antheridia and archegonia, in which gametes are produced. Gametophytes in some groups are subterranean, nonphotosynthetic, and associated with mycchorizae (fungal symbionts that occupy the roots). The gametophytes of most seedless vascular plants are inconspicuous and ephemeral relative to the sporophytes, although in some lycophytes, for example, it may take 6–15 years for gametophytes to become fertile and produce a sporophyte. Male gametes are flagellate sperm, and therefore, as with the nonvascular plants, they are dependent on water for transport to the stationary egg. After fertilization, the diploid sporophyte grows out of the archegonium on the gametophyte, and the sporophyte eventually grows much larger than the gametophyte and becomes the free-living, dominant form of the life cycle. Meiosis occurs in sporangia on the sporophyte, producing large numbers of wind-dispersed spores that settle and germinate to form the next generation of gametophytes. This life history, in which the sporophyte is the dominant stage is a characteristic of all vascular plants.

The seedless vascular plants display a diversity of ecological tolerances, and occur in a range of habitat types, but still share some aspects of their ecology. The motile sperm remain dependent on moisture for transport to the egg, but the dominance and perennial habit of the sporophytes appear to have somewhat lessened the strength of selection for regular water availability, which appears to limit the range of habitats suitable for bryophytes. Living representatives of the Lycophytes are all herbaceous and mostly tropical. The two living genera of whisk ferns (*Psilotum* and *Tmestipteris*) are tropical and subtropical. Approximately 75% of other fern species are tropical and about a third of these are epiphytes, which grow on other vegetation. A few species of ferns are known only as gametophytes, while others do not appear to form sporophytes near their range limits. It may be that when ecological conditions are no longer favorable for some ferns, sexual reproduction is not possible. This suggests that rates of sporophyte production could act as an indication of habitat suitability in some species.

Among the seedless vascular plants, a number of families contain globally threatened species, according to the 2011 International Union for Conservation of Nature (IUCN) Red List for vascular plants. Twenty-six of 45 species of Lycophytes assessed to date are considered critically endangered, endangered or vulnerable. These numbers include 14 of 28 assessed species of the quilwort family, Isoetaceae. These small, inconspicuous plants most often occur in aquatic or wet terrestrial habitats in tropical and subtropical regions. They represent the nearest living relatives of ancient tree lycophytes that dominated the coal-forming swamps of the Carboniferous.

Among fern families, the Cyatheaceae, Ophioglossaceae, Loxomataceae, and Parkeriaceae all have in excess of 25% of their species listed as globally threatened. The Cyatheaceae is the family that includes the tree fern genus *Cyathea*. Tree ferns are conspicuous elements of tropical and subtropical montane forests. The Parkeriaceae includes only four species in the genus *Ceratopteris*, an aquatic genus of mostly tropical, edible floating ferns, some of which are cultivated for human consumption. The Ophioglossaceae includes three genera, among them the genus *Botrychium*. Members of *Botrychium* have unusual population structures, typically with very low numbers of mature sporophytes observed as scattered individuals. These ferns are perennials and may not produce aboveground structures every year. Gametophytes, which are subterranean in *Botrychium*, are rarely found, and thus the dynamics of sexual reproduction are not well-known.

Gymnosperms

The gymnosperms are conspicuous and important components of many terrestrial ecosystems. They include four lineages with living representatives, including the conifers (Coniferophyta; about 630 species), cycads (283 species in 11 genera), the maidenhair tree, *Gingko biloba* (the only species in the Gingkophyta), and the gnetophytes (including three genera, *Gnetum*, *Welwitchia*, and *Ephedra*, totaling about 65 species; **Table 1**). Together, the gymnosperms and angiosperms constitute a uniquely derived group, the seed plants. The four lineages of gymnosperms (as well as the angiosperms) all form seeds, defined as mature ovules that contain embryos. As such, the seed plants represent a further shift in the importance of the sporophyte relative to the gametophyte generation. In all seed plants, the gametophyte is enclosed within sporophytic tissue and, thus, is no longer free-living at any time during its existence. Male gametophytes are packaged and dispersed as pollen grains, while female gametophytes are retained on sporophytes in ovaries. Pollen may be transferred to the vicinity of the ovules by wind or insects (e.g., some gnetophytes and cycads are likely beetle pollinated). Pollen tubes are formed, through which male nuclei are transported to the egg cells. After fertilization, the embryo develops inside the megasporangium wall, resulting in a seed that may be dispersed by wind or animal vectors. For example, the seeds of pinyon pine are dispersed by nutcrackers.

Although the cycads and *Ginkgo* have flagellate sperm, these are released from pollen grains after they reach the vicinity of the ovule, and thus, the seed plants do not require water for gamete movement. This opens up yet another set of habitats for these plants. Among gymnosperm lineages, ecological requirements are highly varied. Conifers, the most species-rich group of gymnosperms, are also the most diverse ecologically. Conifers play an especially important role in temperate and boreal forest ecosystems, where they are often the dominant tree species. In these regions, conifer species are also of significant economic importance and are managed for timber and paper pulp production. Overexploitation of this resource has impacted these ecosystems in numerous ways, impacting not only narrow endemics such as the burrowing owl of the Pacific northwestern USA but also other resources such as native fish habitats, including the spawning grounds of Pacific salmon.

The 1997 IUCN Red List for vascular plants indicates that all four phyla of gymnosperms include families with globally threatened species. The conifers include seven families with more than 49% globally threatened species. The single species in the Gingkophyta, the maidenhair tree, *G. biloba*, is a relictual Asian species that was preserved in temple grounds in China and Japan and is not known to occur in the wild. A very large percentage of cycads (82.8%) are globally threatened (**Table 1**). The cycads occur in a variety of tropical and subtropical habitats. These relictual plants are cultivated as ornamentals and are threatened in part due to overcollecting.

Angiosperms

With more than 350,000 extant species, flowering plants represent more than 80% of plant biodiversity. They essentially share the life cycle common to all seed plants, with the innovation of having reproductive structures contained in flowers and with further reduction in the gametophytic phase of the life cycle. The flowering plants have evolved a number of features that have been described as important in their radiation, including relationships with pollinators, which opened the way for diversification of floral structures that likely contributed to reproductive isolation among diverging lineages.

Flowering plants display a wide array of variation in ecology, life history, growth form, breeding system, and other traits that make them impossible to characterize as a group. Ecologically, flowering plants are more diverse than all other groups of plants described. They occur in virtually all terrestrial habitats, and in many of these, they are the dominant elements in the landscape. A number of flowering plant species have become adapted to aquatic habitats, while others have evolved tolerance to arctic, extreme desert, and even marine conditions.

The 1997 IUCN Red List for vascular plants lists nearly 300 families that include globally threatened species. Some generalizations about broad-scale patterns of diversity in flowering plants can be made, though regional, local, and taxon-specific patterns are also important in discussions of biodiversity. Because discussions of patterns of diversity most often consider all vascular plants (including seedless vascular plants and seed plants), these patterns will be considered in subsequent sections.

A number of features of flowering plants impact the way in which genetic variation is apportioned in nature, and thus these features impact not only the susceptibility of plant species to becoming threatened but also their potential for recovery. For example, the breeding systems of plants range from highly self-fertilizing to obligately outcrossing. Among outcrossing species, a variety of mechanisms exist to minimize the potential for self-fertilization. Chief among these are genetic self-incompatibility mechanisms, which are estimated to occur in as many as 50% of angiosperms. These mechanisms operate via the action of alleles at a single locus that must differ in paternal and maternal parents in order for matings to be compatible. Highly selfing species tend to display much higher levels of differentiation among populations than many outcrossing species and thus may show a greater tendency for adaptation to locally varying conditions. While inbreeding is generally viewed as having a negative impact on threatened species, plants that habitually inbreed are well adapted to do so. Thus, it is important to understand breeding systems in order to make well-informed conservation and management decisions.

Patterns of Plant Biodiversity

Measures of Biodiversity

Various measures are used to determine plant diversity, and these measures in turn contribute to prioritizing regions for conservation efforts. A simple measure of diversity is species richness. This measure refers to the number of different species in a given region. Although there are other more robust

measures of diversity, plant diversity studies are often based on the measure of species richness. The most species-rich regions of the world are in the tropics and subtropics. Species richness, while measuring the total diversity of a region, does not take into account the degree to which species occur only within a certain region, but these are likely to be correlated. For example, not only are the floras of the tropics the most species rich regions of the world, but also approximately two-thirds of the world's known vascular plants are restricted to tropical regions.

Species endemism measures the number or percentage of species that occur only within a particular defined region. Endemism is often used in assessing the relative contribution of particular ecologically or geopolitically defined areas to broader scale patterns of biodiversity (e.g., ranking the percent endemism of each ecoregion within a continent, or of each state or province within a country). Species that are endemic in a restricted ecological or geographic area (sometimes called narrow endemics, especially if they occur in low numbers within this restricted area) best fit the colloquial notion of rarity. However, while rarity suggests a high susceptibility to threats, and thus a need for conservation assessment, not all rare or narrow endemic species face imminent threats. As a tool for conservation efforts, descriptions of patterns of endemism are more useful when combined with an assessment of threats to persistence.

The description of biodiversity hotspots is meant to capture the combined effect of high endemism and high threat level on the biodiversity of a region, and serves as a tool for prioritizing conservation efforts (Myers, 1990). To qualify as a hotspot, a region must contain at least 1500 endemic vascular plant species, and have lost 70% or more of its primary vegetation type. Hotspots are priority areas for conservation of biodiversity beyond vascular plants, although the degree to which levels of endemism can be generalized to other taxonomic groups is not fully known.

Patterns of Endemism

Based on estimates from Myers (1990), as updated in Myers *et al.* (2000) and Mittermeier (2005) as many as 150,000 vascular plant species (more than 50% of the estimated total flora based on the author's estimates) are confined to 34 hot spots comprising only 15.7% of Earth's land surface. However, because 86% of the habitat in these hotspots has been destroyed (as of 2005), the hotspot flora is now confined to 2.3% of the Earth's surface. Seven hotspots contain more than 5000 endemic vascular plants each, and together these account for roughly one quarter of the world's flora (74,060 species). Note that nonvascular plants are not included in these determinations because of the lack of understanding we have of their taxonomy and global patterns of distribution.

The 34 hotspots can be grouped into several habitat types. Predominant are tropical forests, appearing in 15 hot spots, and Mediterranean-type zones, in five. Nine comprise islands or groups of islands, and almost all tropical islands fall into one or another hotspot. The seven "hottest hot spots" are predominantly tropical – Sundaland (Indonesia, Malaysia, and Brunei), the tropical Andes, the Mediterranean Basin, Madagascar and associated islands, Brazil's Atlantic forest, the Caribbean Islands, and Cape floristic region.

The species-level analyses are complemented by an assessment of endemism among higher taxa such as plant families and genera. Madagascar and the surrounding islands possess eight endemic families and 310 endemic genera, Cape Floristic Province has five endemic families and 160 endemic genera, and New Caledonia has five endemic families and 107 endemic genera. Other hotspots containing few or no endemic families, with large numbers of endemic genera include the Caribbean with 205 endemic genera and Tropical Andes with 330. While the number of endemic higher taxa may rise or fall as our knowledge of relationships and boundaries is revised, these patterns of species richness and endemism, together with the broad destruction of habitat make these areas a high priority for conservation.

Threats to Plant Biodiversity

Intrinsic Features of Plant Species and their Interactions with Threats to Species Persistence

The effects of human activities now operate at such a global scale that it is likely that human activities directly or indirectly impact virtually all plant species to some degree. However, not all plant species are equally at risk, nor are all species equally susceptible to the action of a particular threat, even on a local scale. A variety of intrinsic factors have been suggested as affecting the degree of susceptibility of a species to threats, including life history features (e.g., annual vs. perennial, selfing vs. outcrossing, the presence and extent of seed dormancy, etc.), distribution and abundance traits, including local population size, range size, habitat specificity, and the extent of dependence on other organisms, for example, pollinators.

Population size can be an important determinant of the fate of a species, especially when considered in combination with the extent of a species geographic range. A species with a narrow geographic range and small population sizes will have an increased chance of all of its populations experiencing the same threat, and each small population will be at greater risk of extinction than in a species with a broader range and larger populations.

Specific biotic interactions can also be important for species persistence. Plants do not exist in isolation: in fact their survival often depends on other organisms. Many species of plants have evolved intimate relationships with a diverse array of organisms, without which they are unable to complete their life cycles. Symbiotic or mutualistic relationships with fungi, bacteria, insects, and birds are important for successful germination, growth, reproduction, and dispersal of propagules in many species of plants. For example, flowering plants that depend on one or a few species of animals for pollination and seed dispersal may become threatened with extinction in the face of a rapid decline in the population of these animals. Reduced seed set is the most direct effect of pollinator decline, yet indirect effects may also contribute to a decline in the plant population. For example, in the absence of pollinators, a higher percentage of seed may be set through self-pollination, increasing the expression of inbreeding depression.

Ultimately, loss of pollinators or disruption of pollination systems can cause plant extinction. A number of factors, including habitat alteration or fragmentation, grazing, introduction of alien pollinators, and the use of pesticides, can have drastic effects on pollinator populations. The plants most at risk from loss of pollinators are diecious (species where male and female flowers occur on separate plants) or self-incompatible, those that have a single pollinator, and those that propagate only by seed. Once again, a plant species with a restricted range can be especially vulnerable to population fluctuations among its pollinators and seed dispersers.

It should be noted, however, that the effects of intrinsic features of plant species on extinction risk might be overshadowed by the interaction of such traits with extrinsic threats. For example, a threatened annual species with a persistent seed bank might be expected to have a lower extinction risk than a similar species that lacks a seed bank, because the seed-banking species can avoid unfavorable growing conditions by remaining dormant as seed. However, in the presence of an invasive species that can quickly colonize open patches, the longevity of the seed bank may not aid species persistence, and indeed may facilitate the establishment of invasive vegetation.

The Most Important Anthropogenic Threats to Plant Biodiversity

The greatest threat to plant biodiversity is habitat destruction, conversion, or degradation, which arises through the combined direct and indirect effects of agriculture (including crop production and cattle grazing), logging, and development. Overharvesting of individual species, invasion by exotic species and the effects of climate change round out the list of top threats to plant species persistence (Freville *et al.*, 2007).

Agriculture has had wide ranging and profound impacts on global ecosystems. The pressure for increased food production has done more to decrease plant diversity and to physically alter our surroundings than any other human activity. In addition to crop production, livestock grazing has also imposed drastic impacts on plant communities. The selective nature of grazing, combined with the limited tolerance of some plant species to grazing, has produced substantial shifts in regional species composition.

One of the most visible aspects of habitat degradation in natural landscapes is the spread of exotic, or occasionally native, invasive vegetation that alters or displaces native communities. Disturbances to plant habitats such as those described above increase opportunities for these invasions. The introduction and often widespread dissemination of alien species, such as Norway maple (*Acer platanoides*), kudzu (*Pueraria lobata*), purple loosestrife (*Lythrum salicaria*), and Japanese honeysuckle (*Lonicera japonica*), planted in environments where there are no natural controls or defenses, have been devastating.

The eventual impact of global climate change, the product of anthropogenically driven increases in atmospheric CO_2, is as yet unknown for most plant species and communities. Research on the effects of climate change on plants focuses on understanding the direct and indirect effects of changes in temperature and precipitation patterns on species persistence. Temperature and precipitation shifts may directly affect species persistence via their physiological tolerances, but will also act indirectly by affecting interactions with competitors, pollinators, pathogens and pests, and all of these effects are the subject of ongoing research.

The current wave of extinctions resulting from systematic pressures imposed by human activities is eliminating 27,000 species each year, making this the sixth greatest mass extinction in the history of our planet. The principal factor determining the rate at which species are becoming endangered and extinct is habitat destruction and deterioration, especially in the species-rich tropics. The rain forests, which harbor this amazing diversity of species, are subject to destruction by exploitation and land clearing on an unprecedented scale. It is suggested that 60,000 tropical plant species will be at risk of extinction within the next 50 years. On a similar scale, of the 80,000 species of the temperate zone, about 8000 are threatened and several hundred endangered. At least 217 species of vascular plants and undoubtedly many species of nonvascular plants have gone extinct in North America over the past 500 years. Many extinctions have gone unnoticed because the species were not known to science. In the USA, old-growth forests of the Pacific Northwest, longleaf pine forests of the southeastern coastal plains, range-lands, grasslands and savannas, and wetlands – 33% of the species listed under the Endangered Species Act are dependent on wetlands – are all considered endangered or threatened ecosystems due to one or more reasons ranging from logging, to conversion to agriculture, and secondarily, to fire suppression or overgrazing and subsequent invasion by exotics.

Eleven hotspots have already lost at least 90% of their primary habitats while three have lost 95%, making conservation a high and immediate priority. Some hotspots have their endemic species concentrated in exceptionally small areas, making these species highly susceptible to extinction. The Eastern arc, New Caledonia, and the Philippines are especially significant in this regard. The criteria used to define hotspots exclude some of the most species-rich areas of New Guinea, Amazonia, and the Congo Basin because even though these regions contain rich endemic floras they also retain at least 75% of their primary vegetation, disqualifying them for hotspot status. Further, regions such as the Ethiopian Highlands, southeastern China, and northern Rwanda, where exceptionally rich floras face exceptional threats, are not sufficiently documented to meet the hotspot criteria. Thus, while it can be argued that hotspot analysis may be used to prioritize conservation efforts to areas where efforts will have the greatest impact, other poorly characterized regions and areas of high endemism must also be considered.

Natural Threats to Plant Biodiversity

In addition to the devastating effects on diversity by anthropogenic activities, natural catastrophes can also have negative impacts on plant diversity. Natural catastrophes such as natural fires, floods, hurricanes, landslides, and droughts occur at various times and scales in all parts of the world, but many of these events are projected or have already been observed to

occur with greater frequency or intensity as a result of rising CO_2. Thus, even these natural events are linked to human activity.

Many plant species and communities are adapted to disturbances such as fires and floods, and indeed, may depend on such periodic events for successful growth and reproduction. Floods, fires, and hurricanes and other natural events may provide the clearings or eliminate cover that otherwise prevent the establishment of species that are intolerant of shading or competition. In many landscapes, intermittent disturbance over variable spatial and temporal scales results in patchwork of habitat types or plant communities that sustain different species of plants, and therefore provide a mosaic of habitat types for nonplant species. However, the presence of human populations and infrastructure often leads to management practices that alter or eliminate natural disturbances, with unintended consequences for natural communities. For example, prairies and other grasslands, ponderosa and longleaf pine forests often depend on frequent, low-intensity ground fires. Without these fires, these communities gradually change into other community types (e.g., grasslands into forest scrub) that are often less diverse, and can result in the need to recover the original ecosystem type through active management.

Diseases also pose threats to plant diversity. The recent rapid decline in *Chamaecyparis lawsoniana* (Port Orford/Oregon Cedar) across its range in northern California and southern Oregon has been attributed to a root rot disease caused by a fungus. Apparently, the spores of this fungus are transported mainly in the mud on tires of logging trucks. Other pests and pathogens that disperse along road systems include black stain root disease fungus and gypsy moth, all having had drastic impacts on plant diversity. Once again we see that a naturally occurring threat to a species can have an increased impact as a result of human activities.

Conservation of Plant Biodiversity

The consequences of loss of plant biodiversity for human welfare and global ecology are unpredictable and wholly irreversible. Hence, identification and protection of sites of high conservation value (i.e., biodiversity hot spots) is of high priority. A key limiting factor in our ability to identify targets for conservation is a lack of knowledge of the biodiversity present. While this gap is narrowing for vascular plants, nonvascular plants and other more cryptic forms of biodiversity remain understudied, adding value to novel approaches such as deoxyribonucleic acid (DNA) barcoding for estimating taxonomic diversity, and advancing the development of priorities that reflect the diversity of less known groups (Lahaye *et al.*, 2008).

Species are neither randomly nor uniformly distributed across the planet. Rather, they respond to environmental gradients, including climate, topography, and substrate. They further reflect historical contingencies resulting from patterns of colonization, phylogenetic history, speciation and extinction, plate tectonics, and other global processes. By paying attention to these processes, conservationists can identify areas and focal taxa of greatest importance for protection. There are several criteria used in assessing the conservation value of natural areas. Species richness (or diversity), endemism, naturalness, rarity (extent of habitat), threat of human interference, amenity value, educational value, scientific value, and representativeness are some of the criteria considered when prioritizing conservation areas.

Destruction and fragmentation of natural areas are certain to lead to substantial increases in extinction rates. In some cases, *ex situ* conservation measures are warranted as a supplement or replacement for conservation on site. Although it is universally agreed that the most effective and efficient mechanism for the conservation of plant diversity is habitat protection or *in situ* conservation, it is also acknowledged that *ex situ* facilities can be critical in a comprehensive conservation program. *Ex situ* conservation programs supplement *in situ* conservation by providing for the long-term storage, analysis, testing, and propagation of threatened and rare species of plants and their propagules. They are particularly important for wild species whose populations are highly reduced in numbers, serving as an alternative to *in situ* conservation, as a source of material for reintroductions, and as a major repository of genetic material for future breeding programs of domesticated species. Methods of *ex situ* conservation can be classified according to the part of plant that is conserved – the whole plant, seed, tissues, or genetic material in culture. All these methods of *ex situ* preservation of live plant material require periodic regeneration and sexual reproduction of the stock. The latter requires knowledge of the breeding system and pattern of genetic variability of the species concerned.

Botanic gardens and arboreta have always played key roles in *ex situ* conservation, but the role of seed banks, including such ambitious endeavors as the Millennium Seed Bank, which targets seed banking 25% of the world's plant species by 2020 represent a more recent targeted global initiative. The most widely known function of arboreta and botanic gardens is to assemble and maintain a diversity of plant species. They also conduct and facilitate botanical research, especially in plant taxonomy and systematics. Arguably, botanic gardens make their most important contribution to the conservation of plant diversity through education and outreach, encouraging appreciation, and awareness of issues affecting plant diversity and of the contributions of plants to human well-being.

The Need for Further Study

Representatives of all groups of plants contribute to the health of the global ecosystem and to human welfare. As such, from a practical standpoint, any loss of biodiversity has the potential to negatively impact the future of the human population. A key limitation in the study of plant diversity is the training of professional and nonprofessional taxonomic experts in taxonomy of poorly known plant groups, including nonvascular plants, and diverse and challenging groups like the grasses and sedges, as well as taxa that have restricted distributions. As taxonomic expertise shifts away from academic institutions, our ability to train the next generation of plant taxonomists and expand our understanding of poorly known groups

may be hampered. Special emphasis should be placed on characterizing species that are thought to have potential for medicinal use. In addition, regions of the world that are highest in diversity, especially those regions that are most critically threatened, must be the target of a significant proportion of resources available for conservation. This can best be accomplished by training of additional highly qualified personnel and increasing funding to conservation agencies, including *ex situ* conservation programs. All of these initiatives will be most effective if they are coordinated by an international body that relies heavily on recommendations made by scientists and regional experts. It is only through concerted efforts such as these that we will be able to slow the loss of plant diversity around the world. As plants play such a vital role in the health of our planet and our species, it is of the utmost importance that we value and protect this diversity for future generations.

Appendix

List of Courses

1. Introductory Biology
2. Introductory Botany
3. Plant Diversity

See also: Hotspots. Plant Conservation. Taxonomy, Methods of

References

Fréville H, McConway KC, Dodd M, and Silvertown J (2007) Prediction of extinction risk in plants: Interaction of extrinsic threats and life history traits. *Ecology* 88: 2662–2672.

Joppa LN, Roberts DL, Myers N, and Pimm SL (2011) Biodiversity hotspots house most undiscovered plant species. *Proceedings of the National Academy of Sciences (USA)* 108: 13171–13176.

Lahaye R, van der Bank M, Bogarin D, *et al.* (2008) DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences (USA)* 105: 2923–2928.

Mittermeier RA, Robles Gil P, Hoffman M, *et al.* (2005) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Arlington, Virginia: Conservation International.

Myers N (1990) The biodiversity challenge: expanded hot spots analysis. *Environmentalist* 10: 243–256.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Patton AJ, Brummit N, Govaerts R, *et al.* (2008) Towards target 1 of the global strategy for plant conservation: a working list of all known plant species - progress and prospects. *Taxon* 57: 602–611.

Raven PH, Evert RF, and Eichorn SE (2005) *Biology of Plants*. 7th edn., 944 pp. New York: W.H. Freeman and Company.

Plant Communities, Evolution of

Brian J Enquist and Bruce H Tiffney, University of California at Santa Barbara, CA, USA
John Haskell, University of New Mexico, NM, USA
Karl J Niklas, Cornell University, NY, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 631–644, © 2001, Elsevier Inc.

Glossary

Community A group of coexisting species characterized by taxonomic composition and/or structural type, possessing repeatable associations of dominance, diversity, and trophic interaction.
Competition Interactions between two or more organisms for biotic or abiotic resources, resulting in decreased fitness of one or more of the participants.
Diversity A quantitative measure of the taxonomic or physiognomic composition of coexisting organisms at one of three scales: alpha (local diversity), beta (diversity change between adjacent local sites), and gamma (regional diversity).
Niche An organism's biotic and abiotic resource requirements, collectively enabling the organism's trophic and behavioral role in the community.
Succession Changes within ecological time spans in the taxonomic or physiognomic composition of a specific community in response to biotic or abiotic perturbation.
Turnover Changes in the taxonomic and possibly physiognomic composition of a community over evolutionary timescales in response to environmental or evolutionary change.

Introduction

In the past few decades, questions about the creation, maintenance, and role of biological diversity have stimulated social, political, and scientific inquiry. However, the fundamental mechanisms responsible for biodiversity remain poorly understood. We review the patterns in contemporary and historical plant communities that we believe are critical to achieving a better understanding of the processes that regulate diversity. In doing so, we take an actualist (or pure uniformitarian) stance: We assume that the basic physical and biological processes that regulate diversity today are manifestations of natural "laws" that have been in operation throughout the existence of the earth.

Clearly, any realistic juxtaposition of the concepts derived from neo- and paleoecological studies requires a clear exposition of the limitations and evidence on which they rest. Paleo- and neontological research programs each bring different strengths to bear on the question of how biological diversity originates and is maintained. The greatest strength of neoecology is the ability to achieve fine resolution. Much of what is known about the ecological dynamics of plant and animal communities is based on detailed, short-term, small-scale, neontological studies. These studies emphasize the role of local processes such as succession, predation, competition, mutualism, abiotic stress, and a host of other ecological phenomena in the regulation of biodiversity. On the other hand, the paleontological research program tends to emphasize the influence of evolutionary and physical changes on the species composition of communities. Despite their different strengths, the two disciplines often reach similar or complementary conclusions. For example, both paleontology and neontology concur that community composition is highly variable over space and time (Ricklefs and Schluter, 1993; Brown, 1995). Communities do not consist of a "tight-knit" assemblage of co-occurring species; instead, species respond individualistically to changing biotic and abiotic conditions.

Differences between the disciplines in the level of taxonomic, spatial, and temporal resolution have important consequences on how phenomena are defined and perceived. Within neontology, the definition of coexisting species can vary from very small sets to regional assemblages. In most cases, neo- and paleoecologists use the same vocabulary (e.g., community, taxonomic diversity, and stability), but the meanings of these words change as a function of the spatial and temporal scales used in the two fields. Despite these potential pitfalls, we believe that we can improve our knowledge of the important community processes by focusing on the complementary strengths of each discipline.

Brief Overview of the Evolution of Terrestrial Plants

The fossil record clearly shows that the number of plant species has increased over geological time (**Figure 1**; Niklas, 1997). The first land plants were diminutive, nonvascular, bryophyte-like organisms that inhabited the margins of the terrestrial landscape and provided resources and shelter to the first land-dwelling animals. These plants were restricted in their geographic coverage owing to their dependence on liquid water for vegetative survival and reproductive success (Bateman *et al.*, 1998).

The earliest vascular plants possessed better vegetative adaptations to life on land but were also tied to moist habitats for reproduction. They quickly diversified in morphological complexity, primarily during the Devonian, a period that saw the first plants with leaves, roots, lateral meristems, and an erect habit. These plants were larger than their nonvascular

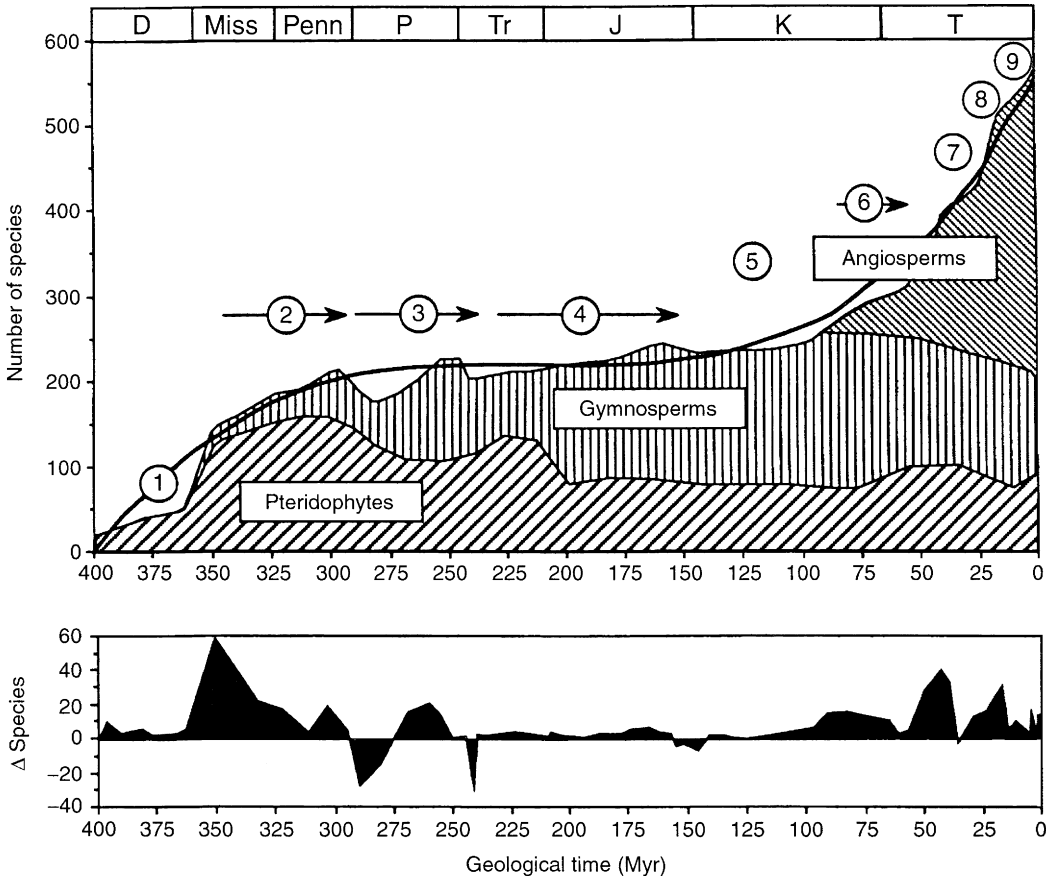


Figure 1 Total number of species and species turnover. Number of vascular land plants and change in species number are plotted as a function of geological time. Data for vascular land plant species are segregated into three reproductive groups (pteridophytes, gymnosperms, and angiosperms). Solid curved line denotes ordinary least squares regression curve for the cumulative data. Change in species number (turnover) is computed on the basis of the appearance and disappearance of species in each geological stage. 1, Origin of seed, origin of arborescent habit; 2, diverse lowland swamp communities, high atmospheric oxygen; 3, sequential decline in diversity and occurrence of pteridophyte-dominated communities, coalescence of Pangea; 4, gymnosperm-dominated communities common, Pangea slowly breaks up; 5, angiosperms first appear in the fossil record and begin to spread, becoming the dominant floristic element; 6, Cretaceous–Tertiary boundary, angiosperms diversify and become the dominant vegetational element, including closed forest communities; 7, Eocene–Oligocene climatic shift leads to dramatic Earth cooling; 8, herbaceous angiosperm communities radiate, including grasslands; 9, taiga and tundra biomes fully developed.

progenitors. By Late Devonian times, vascular plants evolved the capacity for secondary growth and with it the ability to form forests. These changes increased the complexity of the physical structure of plants and the diversity of resources available within terrestrial communities, creating opportunities that were subsequently exploited by the evolution of plants with vining, epiphytic, and understory lifestyles. Initial stages of the seed habit also evolved by the Late Devonian, allowing new ecological solutions within the swamp community and laying the groundwork for the subsequent establishment of communities in drier habitats in the Carboniferous. The basic physiognomies of modern plant communities were well established approximately 300 million years ago.

From the late Carboniferous through the early Mesozoic, gradual environmental changes associated with the formation of Pangea resulted in extinctions of lowland taxa and their replacement by increasingly drought-tolerant taxa migrating from more mesic sites. This established an early mid-Mesozoic flora dominated by a host of seed-bearing clades collectively known as the gymnosperms. Pteridophytes survived but were a secondary element of Mesozoic vegetation.

Angiosperms became an important ecological element approximately 115 million years ago, in the Early Cretaceous. They initially evolved as disturbance-tolerant herbaceous and shrubby plants occupying early successional environments, but they quickly radiated. By the Late Cretaceous, angiosperms started to appear as forest trees in increasingly more stable sites, and they were the most diverse terrestrial plant clade. However, it is not clear if they formed the dominant vegetation at this time. By the early Tertiary, angiosperms were represented by many living families and genera forming closed-canopy forests in temperate and tropical climates, and they dominated Earth's vegetation. Cooling and drying climates in the later Tertiary resulted in the diversification of mid-continental grassland communities and the rearrangement of

individual distribution patterns that created the biomes of the present. The rise of angiosperms established the full complement of structural and reproductive features (e.g., complex canopies, diverse life history, and reproductive strategies) observed today as well as creating the most diverse terrestrial communities in Earth's history. In their radiation, angiosperms did not displace equally diverse nonangiospermous seed plant communities. Rather, angiosperms both inserted themselves into existing communities and created totally new communities, often creating new opportunities for existing clades. The result has been the most speciose plant communities yet seen on Earth (**Figure 1**).

Niche Diversification and Competition

Understanding the evolution of plants in terms of niches provides the basis by which one can begin to assess the evolution of diversity within plant communities. The concept of the niche is useful because it provides a framework linking individual traits, such as morphology, physiology, and behavior, with their interactions with the biotic and abiotic environment. These features in turn influence population level traits such as growth, abundance, and distribution (Brown, 1995).

Both ecological theory and empirical evidence stress that competition is one of the important forces driving niche diversification and thus community diversity (Knoll (1986) as cited in Rosenzweig, 1995). The increase in plant species diversity through time is paralleled by evolutionary shifts in resource requirements, body size, allocation strategies, and life history traits suggestive of niche diversification and thus of the action of competition. Indeed, the degree of difference or similarity that is required for coexistence has not been quantified, and it has been hypothesized that two or more plant species with very similar resource requirements can coexist indefinitely in the same community as long as their reproductive units (e.g., spores, seeds, or fruits) are dispersed randomly such that adults do not consistently grow in close proximity (Hubbell and Foster as cited in Ricklefs and Schluter, 1993; Brown, 1995; Rosenzweig, 1995; Crawley, 1997).

Competition, however, is difficult to effectively demonstrate in the fossil record. At large temporal scales, such as those encountered in the fossil record, the evidence suggests that competition may be supplanted by other features such as shifts in the physical environment (Niklas *et al.* (1985) as cited in Ricklefs and Schluter, 1993). The shift from a Pteridophyte-dominated late Paleozoic flora to a seed plant-dominated Mesozoic flora appears to have been caused by global climate change that selected against the reproductive modes of many pteridophyte lineages, resulting in extinctions in existing lowland communities. Opportunistically, seed plants migrated in from more mesic sites to replace the lowland pteridophytes. Similarly, gymnosperm diversity was not drastically affected by the evolutionary appearance and subsequent diversification of angiosperms, despite the observation that contemporary gymnosperms are ecologically displaced by the more rapid vegetative growth and reproduction of flowering plant species (Niklas *et al.* (1985) as cited in Ricklefs and Schluter, 1993). The evidence suggests that although niche diversification is driven largely by competition, major changes in community composition at evolutionary timescales have been initiated by larger scale physical changes.

Body Size and Plant Form

Body size influences nearly all the characteristics of organisms and ultimately the composition of the communities they inhabit (Brown, 1995). The ecological and evolutionary ramifications of body size appear to have been an important basis for differentiation in vascular plants. Neoecological observations indicate that mass-specific metabolism, age to reproductive maturity, relative growth rate, and population density are all inversely correlated with body size (Niklas, 1997; Enquist *et al.*, 1999). Thus, although the total resource needs of larger individuals increase, their use per-unit-volume or per-unit-mass decreases. Body size also affects biological timing. For example, larger species have longer life spans and longer times until reproduction than do smaller species. Increases in size allow species to partition resources over time: Small species tend to be more sensitive to short-term environmental variation, whereas large species apparently cue to longer term fluctuations.

Individual clades within pteridophytes, gymnosperms, and angiosperms often display Cope's rule: Derived taxa tend to become larger. An increase in size confers a competitive advantage in the struggle for the resources, such as light and nutrients, that are necessary for growth and reproduction (Harper (1977) as cited in Crawley, 1997). Paleoecological data indicate that Cope's rule may also reflect the tendency for new clades to make their evolutionary debut in disturbed habitats (DiMichele *et al.*, 1987; Niklas *et al.* (1985) as cited in Ricklefs and Schluter, 1993). Such species often possess a small body size and its associated rapid vegetative growth and onset of sexual maturity (i.e., the classic r-selected species). Conversely, larger species within many plant clades appear to occupy more ecologically stable habitats.

Over long time spans there may be selection against larger sizes. Larger species tend to be more prone to extinction during periods of environmental change than their smaller counterparts (Bakker (1977), Stanley (1979), and Tiffney and Niklas (1985) as cited in Niklas, 1997; Brown, 1995). Presumably this is because larger species take more time to reach reproductive maturity, tend to produce fewer progeny per reproductive cycle, and have lower population densities per unit area. This suggests that evolutionary processes operating on microevolutionary and macroevolutionary timescales may have conflicting outcomes. Under the limitations of physical and biological constraints, organisms benefit from competitive advantage by becoming larger until an unpredictable environmental change forces them to extinction.

A plant's form, or shape, is as important as its size in partitioning ecological opportunities. Many aspects of plant form and architecture can be interpreted as the result of natural selection acting to optimize biomechanical stability and the ability to procure resources from the environment while minimizing hydrodynamic costs of resource transport (Niklas

and Kerchner (1984) as cited in Niklas, 1997; West *et al.* (1997, 1999) as cited in Enquist *et al.*, 1999). Only a relatively small number of "plant designs" appear to be functionally viable over the range of sizes within which vascular plants operate. It can be shown via mathematical modeling and computer simulation that most of these viable options have been exploited (Niklas and Kerchner (1984) as cited in Niklas, 1997) and that the evolution of plant size has been strongly limited by biomechanical constraints (Niklas, 1997; Bateman *et al.*, 1998). In summary, although selection has apparently operated to increase plant size within clades, physical and biomechanical constraints have restricted maximum size (**Figure 2**).

Allocation and Reproductive Strategies

Biomechanical limitations restrict niche partitioning involving body form. However, plants with different forms may vary the allocation or timing of energetic output, thus increasing the potential for coexistence. Separate species attain their "adult" size at different times by altering growth rates. Unless the total amount of energy available to the individual plant changes, such changes in growth rates must be accompanied by a trade-off in other activities requiring energy. For some species, changes in growth rate are compensated for by changes in wood density: Fast-growing arborescent species generally produce wood of a much lower density than do slower growing species (e.g., balsa versus mahogany; Enquist *et al.*, 1999). This has obvious implications for relative mechanical stability, susceptibility to disease, and species longevity. Similarly, an energetic trade-off exists with regard to the ability to tolerate herbivory, parasitism, and disease (Crawley, 1997). By investing energy in the formation of secondary chemicals, plants can often avoid enemies, albeit at the cost of slower growth and delayed sexual maturity.

Perhaps the area in which different allocation strategies have the strongest effect is within reproduction. Plants can alter both the proportion of resources used in reproduction and the amount of resources allocated to individual diseminules. Such adaptations are often linked to community structure. Large seeds, for example, tend to occur in forest communities, whereas smaller seeds are more common in open or disturbed habitats (Crawley, 1997). Variation in reproductive strategy provides plants with additional ways to partition environmental opportunities and coexist.

The fossil record suggests that reproductive adaptations have permitted the ecological exploration of previously poorly colonized terrestrial environments, starting with the invasion of land, the evolution of the seed, and finally the suite of adaptations involved in the origin of angiosperms (Niklas, 1997). Early land plants possessed pteridophytic reproduction, requiring freestanding water. With evolution of advanced forms of gymnospermy, this dependence was lost, but at the cost of vegetative reproduction, because all nonangiospermous seed-bearing plants were apparently either shrubs or trees capable of only limited vegetative reproduction.

The evolution of the seed habit, specialized dispersal mechanisms, and the capacity for seed dormancy opened up many new habitats in the late Paleozoic and Mesozoic. Together, specialized dispersal and seed dormancy permitted further niche diversification within plant communities. The morphology of dispersal units through the Late Cretaceous suggests that seed dispersal was primarily abiotic. The evolution of large seeds in many angiosperm lineages during the early Cenozoic indicates the increasing importance of animal dispersal. Additionally, the ability of seeds to remain dormant has many important ecological implications, potentially enhancing local coexistence by enabling potential competitors to partition niches in time and space. Furthermore, dormancy enables plants to "wait" until specific environmental

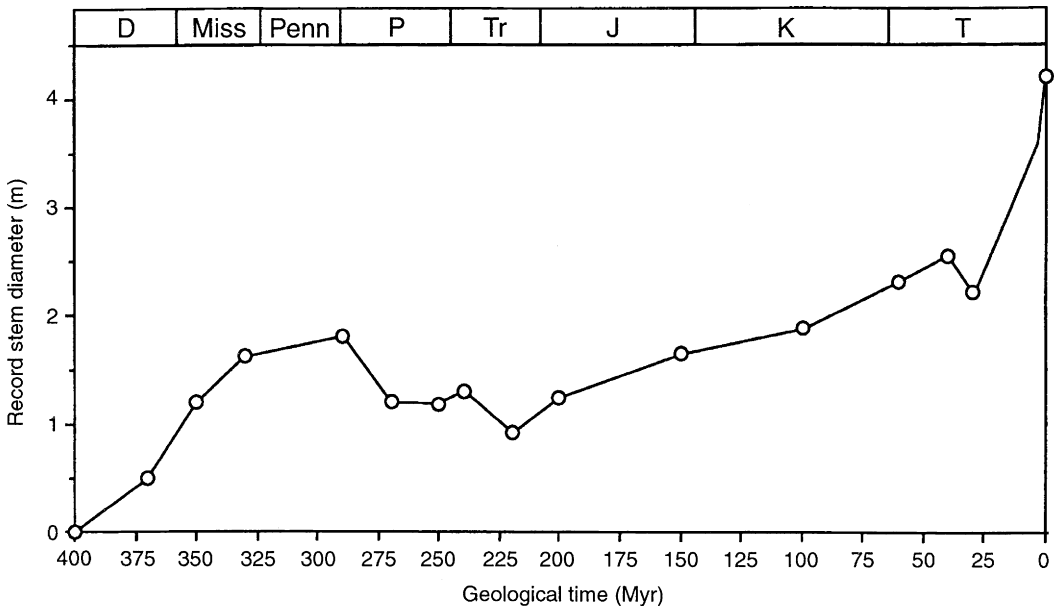


Figure 2 Record stem diameters. Largest reported stem diameters plotted as a function of geological time. Data from Niklas KJ (1997) *The Evolutionary Biology of Plants*. Chicago: Univ. of Chicago Press.

conditions are present, such as a rare rain in the desert or a high nutrient defecation following passage through an animal's digestive tract. Dormancy holds the potential for greater dispersal distances and the formation of seed banks, enhancing the geographic ranges of species and their capacity to deal with potentially protracted inclement environmental conditions. Dormancy, however, may also increase the probability of seed predation.

It is only with the angiosperms that asexual vegetative reproduction and the seed habit are found to commonly cooccur within an individual with either an herbaceous or woody growth habit. The possession of rapid sexual life cycles, the potential for an herbaceous habit, and in many cases the potential for vegetative reproduction predisposed angiosperms to outdiversify and potentially outcompete other lineages in multiple ways in response to environmental changes especially in the Cenozoic. These traits allowed for further invasion of new habitats, such as desert and arctic/alpine environments.

Plants and Climate: Physiognomy and Physiology

The patterns of plant diversity and distribution are influenced by climate. Most biomes show a remarkable degree of convergence in plant architectures, size structure, and physiology, despite drastically different biogeographic and evolutionary histories of their component taxa (Leigh (1975) and Mooney (1977) as cited in Brown, 1995; Milewski (1983) as cited in Ricklefs and Schluter, 1993). Many of the physiognomic features associated with plants found in communities within each of these biomes, such as the relative proportion of compound and simple leaves, root to shoot ratios, plant heights, leaf cuticle thickness, and leaf indument, all show consistent trends with variation in temperature, precipitation, humidity, and rainfall (Crawley, 1997; Raunkiaer (1934) as cited in Crawley, 1997; Wolfe (1978, 1979) as cited in Ricklefs and Schluter, 1993). Indeed, because of the tight association between vegetative physiognomy and climate, physiognomy is often used as a proxy for reconstructing past climates.

The tight link between climate and physiognomy is due mainly to physiological and stoichiometric tradeoffs (Crawley, 1997). One of the most important is the trade-off between water and carbon balance. In order to assimilate atmospheric CO_2, plants must be able to exchange gasses through their stomata, but open stomata cause water loss via transpiration. In many terrestrial environments water stress can limit rates of production by reducing photosynthesis. The evolutionary history of terrestrial plants is marked by the evolution of adaptations that enhance the ability to increase productivity in the face of water stress. The initial radiation of vascular plants was most likely due to the evolution of a cuticle pierced by stomata whose aperture could be regulated (Niklas, 1997). These adaptations were honed through the subsequent evolution of sunken stomata, leaf pubescence, physiologically dormancy, and the ability to mechanically position leaves in response to environmental change. Such morphological adaptations allowed for the colonization of more seasonal environments (Hutton *et al.* (1998) as cited in Bateman *et al.*, 1998).

Although it appears that changes in morphology allowed for the continued exploitation of new resources and habitats, physiological processes appear to be highly conserved over geologic time (Raven (1995) as cited in Bateman *et al.*, 1998). However, interpreting these features is difficult from fossils because the process of preservation alters chemical fossils. When physiological processes are mirrored in morphology, we infer physiological evolution. For example, changes in the density of leaf-borne stomata appear to correlate with geochemically predicted changes in atmospheric CO_2 content, at least back through the Mesozoic. Other morphological features suggest that basic plant physiology has changed little in 400 million years. C3 photosynthesis appears the norm; CAM photosynthesis may have evolved by the Pennsylvanian. C4 photosynthesis is, at the oldest, Late Cretaceous and probably younger, matching the radiation of angiosperms into hot and dry environments. The timing of observed increases in plant diversity does not appear to be associated with physiological changes.

Coevolution

The diversification of vascular plants is associated with an increasing number of coevolutionary relationships. Coevolution can be subdivided into two extreme forms: "tight" coevolution, in which the interactions between two closely linked organisms are intimately interdependent resulting in mutual evolutionary influences, and "loose" coevolution, in which the interdependence of two or more species is facultative and transient. Coevolution includes interactions in which all parties benefit and interactions in which one party benefits at the expense of the other.

The fossil record suggests tight coevolution is rare or short lived. For example, Abert's squirrel currently feeds almost exclusively on ponderosa pine, which currently has a very large geographic range; the modern distributions of Abert's squirrel and ponderosa pine closely overlap. However, Betancourt and Van Devender (as cited in Brown, 1995) suggest that this is a very recent feature since the geographic range of ponderosa pine was reduced to a handful of relict populations in the Pleistocene that probably could not have supported a viable squirrel population. Likewise, there are many angiosperm genera whose fruits or seeds are currently dispersed by specific mammals or birds. Although many of these genera had similar seed and fruit morphologies more than 30 million years ago, none of their modern dispersal agents existed during this time. In fact, these plant lineages passed through two major turnovers of mammals during the Tertiary, indicating that dispersal coevolution must have been either loose or capable of rapid adaptation. Additionally, recent analyses indicate that the radiation of the angiosperms had little apparent influence on the diversification of insect families, although it is possible that coevolutionary response occurred below the family level. In fact, insects began radiating more than 100 million years before the ascendancy of angiosperms (Labandeira and Sepkoski (1993) as cited in Rosenzweig, 1995). These observations suggest that many of the tight coevolutionary couplings seen in modern communities may be relatively transient phenomena.

Nonetheless, both neo- and paleontologists have shown that loose coevolutionary relationships have been an important part of plant evolution. Insect pollination is a basal feature among angiosperms and their closest relatives. Many floral features, including bilateral symmetry, prolonged calyx tubes, nectar and resin rewards, and sympetaly, appeared by the Turonian and suggest specialization for bee pollination (Crepet, 1996). The obligate mutualisms between yuccas and yucca moths, which have been well documented by molecular data, indicate that the general relationship was likely established as long as 40 million years ago. Similarly, fossils indicate that the close association between figs and wasps is at least as old (Pellmyr, 1999). Coevolutionary relationships with pollinators and seed dispersers probably enabled the angiosperms to maintain broadly dispersed but rare or patchy populations by promoting genetic outcrossing and minimizing local inbreeding. This created an opportunity to colonize habitats that were inaccessible to older plant clades that were more reliant on wind and water for the dispersal of seeds and gametes. Further, the presence of mycorrhizal fungi in the earliest plant communities supports the hypothesis that fungi played a key role in allowing plants to spread over terrestrial habitats (Selose and LeTacon (1998) as cited in Bateman *et al.*, 1998).

As suggested in the Red Queen hypothesis, coevolution can result in an apparent "evolutionary arms race" in which both participants are evolving "as fast as they can" only to maintain their relationship relative to each other (Van Valen (1973) as cited in Brown, 1995). Ehrlich and Raven (as cited in Ricklefs and Schluter, 1993), for example, suggested that advances in antiherbivory compounds by plants are matched by the evolution of more efficient detoxification mechanisms by herbivores. Fossil evidence for this kind of coevolution is difficult to obtain. One possible example involves seeds of *Zanthoxylon* (Rutaceae), which possess oils commonly interpreted as deterrents. Such seeds are commonly found pierced by a hole of consistent morphology from approximately 40 million years ago. In extant *Zanthoxylon* seeds, bruchid beetle larvae create similar holes. This suggests a pest–host relationship that has persisted for approximately 40 million years. Nevertheless, it is unknown if such an arms race of evolving strategy and counterstrategy occured during this period.

The paleoecological perspective also provides evidence for a very broad coevolution between vertebrate herbivores and plants. Vertebrate herbivory appeared in the Pennsylvanian and became well established by the Permian. During this time a terrestrial food pyramid of plants and synapsid herbivores was established, lasting to the Early to mid-Triassic time. These herbivores became extinct during the Middle Triassic and, after a brief hiatus, were replaced by dinosaurs. From the Late Triassic through to the end of the Cretaceous, dinosaurian herbivores dominated. Some of these herbivores were small (on the order of 30 kg), but the dominant ones were nearly two orders of magnitude larger. At the Cretaceous–Tertiary boundary, average herbivore size plummeted with the extinction of terrestrial dinosaurs and the radiation of birds and mammals.

Although the forces driving turnovers in vertebrates are hotly debated, it is evident that changes in the composition and structure of floras accompanied changes in the composition of the vertebrate faunas. The Permo–Triassic marked a very broad global transition from floras dominated by pteridophytes and seed ferns to floras dominated by cycadophytes and conifers. This transition resulted in less digestible forage for herbivores, which the fossil record suggests generally favors the evolution of larger herbivores. The radiation of herbaceous and shrubby angiosperms and low-feeding ornithischian dinosaurs during the later Cretaceous again suggests a very broad-scale coevolutionary process. Dense angiosperm communities, spatially well suited to exploitation by small mammalian and avian herbivores, apparently became common only after the extinction of the terrestrial dinosaurs at the Cretaceous–Tertiary boundary. Together, these observations suggest a coupling between community composition and structure and the evolutionary history of large herbivores.

Succession

Succession is one of the most conspicuous processes in the change of community composition and structure over ecological time scales and is one of the most widely written about topics in plant ecology. Because succession involves local immigration and extinction coupled with changes in species relative abundance, it is conceptually tied to the notion of community stability and equilibrium.

As a neoecological process, succession occurs during timescales defined by the lifetimes of individuals in populations, which vary across species and the communities they form. It is also dependent on the frequency, duration, predictability, and magnitude of environmental change, the proximity of source areas, and the diversity of life history traits present in the species pool. These features are believed to dictate the apparent orderliness of succession and thus the abundance and diversity of species within a community at any given time (Crawley, 1997).

Succession was minimally important in the earliest land plant communities because they were patchy and composed of taxa with similar growth habits and physiological requirements. By the Late Devonian a variety of clades had evolved shrubs and trees, creating multilevel communities and ushering in light-controlled succession. Other features also affected succession, including nutrient availability, and physical disturbance. Fire-driven succession has been demonstrated for the arborescent lycopod *Sigillaria*, which occurs immediately above charcoal-rich layers in Pennsylvanian swamps.

Succession in Mesozoic gymnosperm communities is more difficult to demonstrate. Retallack and Dilcher (as cited in Niklas, 1981) note that angiosperms and cycadeoids in the Late Cretaceous Dakota Formation tend to be associated with sedimentary features indicative of disturbance, whereas conifers and cycads are generally absent from such environments. Similarly, Hickey and Doyle draw attention to the fact that angiosperms in the Early Cretaceous Potomac Formation appear in otherwise conifer-dominated environments either directly above layers of charcoal (created by fire) or in sediments indicative of disturbance by water (Niklas *et al.* (1985) as cited in Ricklefs and Schluter, 1993). Such angiosperm-dominated localities subsequently revert to conifer-dominated

communities, suggesting that some or all of the earliest flowering plant species were early successional specialists.

Stability and resilience

The stability and long-term diversity of living communities are features of considerable contemporary interest and debate. Discussion of these features, however, is often confused by matters of precise definition. Although recent neoecological debate has focused on the effects of diversity on ecosystem processes, such as nutrient cycling and biomass production, some studies have examined the persistence of taxonomic composition. With reference to taxonomic stability, some ecologists maintain that the more species that are in the community, the more interactions will occur and the more resilient the community becomes to perturbation. A similar argument is made for ecosystem processes. However, a similar number of interactions could occur within a community in which continuous species turnover was occurring, changing the species composition but maintaining species richness and ecosystem processes. Thus, two different measures of the community, taxonomic composition and ecosystem properties such as species richness, could demonstrate conflicting patterns (Haskell as cited in Brown *et al.*, 2000). These features of ecosystems are complex and require careful identification of the appropriate models, definitions, and scales of investigation necessary to define stability and the methods used to measure it.

Within communities, species diversity ultimately reflects the dynamics of local colonization and extinction as influenced by the prevailing physical and climatic conditions, species life histories, and evolutionary clades present. As such, diversity (or species richness) is an emergent property of both ecosystems and evolution. It is the result of a myriad of complex interactions of abiotic and biotic factors. Species diversity per se does not dictate whether or not a community displays stability. However, increased biological diversity may confer an increased array of potential responses to varying environmental conditions. Thus, as diversity increases through increased niche diversification, the probability of including species with tolerances for new sets of environmental conditions also increases. During periods of drastic environmental change, the different capacities for productivity, mortality, reproduction, etc. may allow some community members to survive, or change their relative abundance within the community under the new environmental regime, and thus to produce a similar community under the new regime.

Community stability in the fossil record is primarily gauged in terms of taxonomic composition. However, as data are summed over increasingly long time periods, the level of taxonomic resolution generally decreases. Still, stability may be estimated by rates of extinction and origination, with high taxonomic turnover rates suggesting community instability. Using this gauge, late Paleozoic swamp communities appear to be very stable, changing slowly in the face of increasing global cooling leading to the threshold of the Permo–Triassic boundary (DiMichele and Phillips (1990) as cited in Ricklefs and Schluter, 1993). These communities contained large numbers of pteridophytic and early seed plant species, each of which offered a different "solution" to environmental change. With the transition to the generally continental climates of the early mid-Mesozoic, conifers, cycadophytes, and several other less diverse seed plant clades characterized by slow reproduction and growth rates replaced the Carboniferous swamp communities, but at a lower level of species diversity. However, despite their comparatively low diversities, these floras never experienced substantial extinction events, suggesting that taxonomic diversity and stability are not inevitably linked to one another. Similarly, even though the evolution of angiosperms led to arguably the most diverse communities in Earth history, there is no indication from origination/extinction rates that angiosperms form more or less stable communities than their predecessors.

Instead, it is possible that rates of compositional turnover may be more heavily influenced by rates of environmental or evolutionary change and independent of ecosystem function, taxonomic composition, or species richness. For example, climatically driven migrations observed in the late Pleistocene raise the possibility that species are constantly shifting in response to stimuli and may repeatedly reassort themselves into similar associations. The players may be the same over long periods of time, but the communities are not stable in the sense that we would discuss them in a neoecological context. The relative stability of communities observed from the Paleozoic to the present may suggest that most communities are relatively stable in the face of "normal" Earth change. It may take catastrophic change to destabilize the community, regardless of the diversity.

From the perspective of individual lineages, the fossil record suggests that those clades including species with a range of growth habits appear to be more long lived than clades possessed of but one growth habit. Additionally, large organisms tend to be more insulated from rapid changes in the environment and to respond to longer temporal oscillations in the abiotic environment that may be unavailable to smaller plants with shorter life spans. However, small organisms can respond rapidly to short-term environmental changes. If a clade is represented in a community by species with large and small body sizes, then that clade could persist in a community, despite short- or long-term changes in the environment, thus giving the community the appearance of taxonomic stability when gauged at the higher taxonomic levels.

Diversity Equilibria

Many experiments have apparently demonstrated shortcomings in the original theory of island biogeography, especially regarding the idea that island species richness is maintained at consistent equilibrium values by colonization and extinction (MacArthur and Wilson, 1967; Rosenzweig, 1995). However, there is increasing evidence that colonization and extinction processes may be important in maintaining an equilibrium in systems in which species composition is not severely limited by low colonization probabilities (Brown *et al.*, 2000). For example, European pollen data show that there is no overriding directional trend in plant family diversity throughout the past 10,000 years at 24 European sites located between 40 and 70°N latitude. Eighteen of the 24 sites

examined show remarkably consistent diversity despite high rates of familial turnover at each site of 30–50% per 1000-year interval (Haskell as cited in Brown *et al.*, 2000).

Testing theories about diversity equilibria over the longer time spans of the pre-Quaternary record is fraught with potential sampling and taphonomic pitfalls (Alroy (1998) as in McKinney and Drake, 1998). Nonetheless, diversity equilibria models have been extended to evolutionary timescales (Sepkoski, 1991). In these models, diversity is hypothesized to be ultimately limited by rates of speciation and extinction. Equilibrium in this case is hypothesized to be due to "niche saturation" and occurs when a clade has exploited all possible opportunities.

For several animal clades, at intermediate timescales, species origination rates appear to be richness dependent, whereas extinction rates are not. Thus, as diversity increases, origination rates decline and diversity reaches an equilibrial number (Rosenzweig, 1995; Alroy (1998), and McKinney (1998) as in McKinney and Drake, 1998). A similar pattern is also observed in the plant fossil record. During the initial debut of each major land plant clade, speciation rates tend to be high and species longevity is short, but subsequent rates of origination tend to decrease while species longevities increase (**Figure 3**). Furthermore, extinction rates have remained essentially independent of standing diversity (**Figure 4**). Species diversity equilibria, however, are attained only for short and intermediate intervals of geological time. Over longer periods, diversity has progressively increased (**Figure 3**). It is not difficult to understand why: Species equilibria require stable environmental conditions and an unchanging field of players. Over geologic time, neither occurs for a variety of abiotic and biotic reasons, such as plate tectonics evolution.

Latitudinal Gradients

An important macroecological pattern of diversity is the latitudinal diversity gradient, which is the increase in alpha and beta diversity from polar to tropical regions. This gradient is reflected by both the number of higher taxa and the number of species nested within those taxa (Ricklefs (1989) as cited in Ricklefs and Schluter, 1993). For example, typically there are approximately 1.4–2.3 species per plant family and 12 plant families per 0.1-ha area plots within temperate tree communities. In contrast, tropical tree communities of the same size have between 2 and 4 species per family and as many as 58 families (Gentry (1988) as cited in Ricklefs and Schluter, 1993). Evidence from the fossil record suggests that a latitudinal gradient in angiosperm diversity has existed since the origin of this clade. The Eocene floras of 40–50° N latitude were more diverse than those at 70° N latitude during this relatively warm geological interval. The same pattern is seen when comparisons are made among 20- to 16-million-year-old floras found in eastern or western North America and the Canadian Arctic.

Many hypotheses have been advanced to explain the latitudinal gradient in species diversity. The most promising propose multiple processes and consider both local- and global-scale patterns, invoking gradients of productivity and energy availability that correlate with changes in latitude

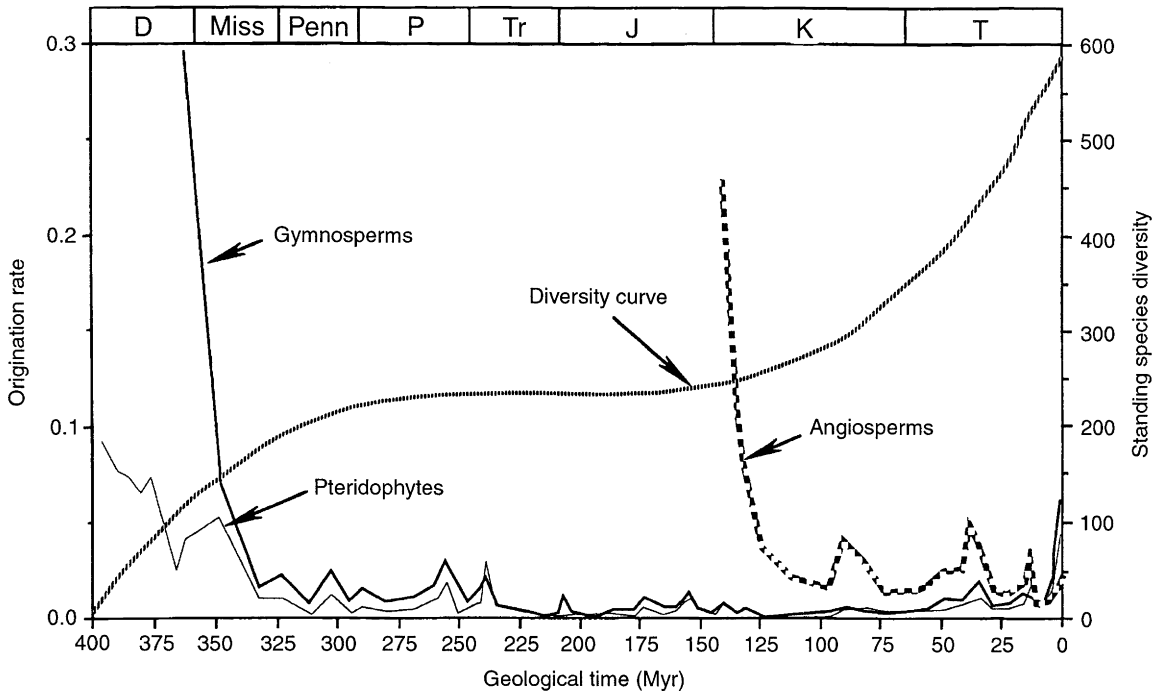


Figure 3 Origination rates of vascular plants. Origination rates of pteridophytes, gymnosperms, and angiosperms are plotted as a function of geological time. Curved line denotes ordinary least squares regression for total standing diversity. Rates are computed as the appearance of new species per unit time (relevant geological stage) per standing local diversity. Exceptionally high rates at the appearance of each group reflect low standing diversity of the group at first appearance. Data from Niklas KJ (1997) *The Evolutionary Biology of Plants*. Chicago: Univ. of Chicago Press.

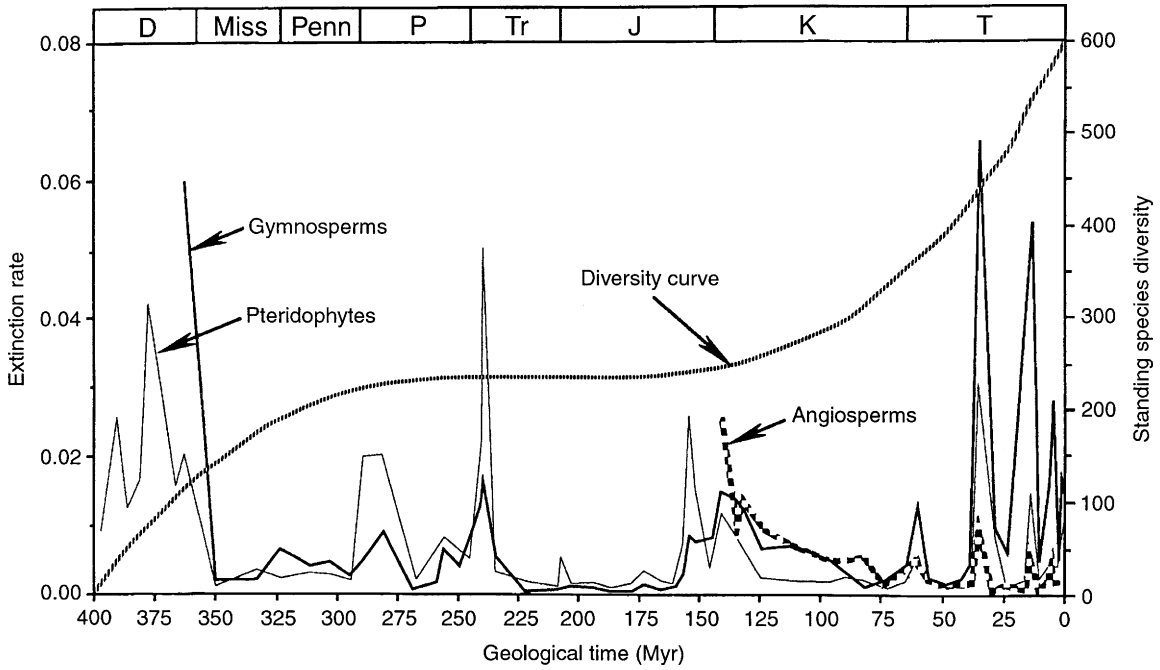


Figure 4 Extinction rates of vascular plants. Extinction rates of pteridophytes, gymnosperms, and angiosperms are plotted as a function of geological time. Curved line denotes ordinary least squares regression for total standing diversity. Rates are computed as the disappearance of new species per unit time (relevant geological stage) per standing local diversity. Exceptionally high rates at the appearance of each group reflect low standing diversity of the group at first appearance. Data from Niklas KJ (1997) *The Evolutionary Biology of Plants*. Chicago: Univ. of Chicago Press.

(Brown, 1995). Productivity ultimately controls biodiversity, and conditions that favor high productivity, such as warm temperatures and ample precipitation, are often associated with high diversity (Currie and Paquin (1987) as cited in Rosenzweig, 1995).

Differential origination rates from the tropics to the poles may also play an important role in creating and maintaining the latitudinal diversity gradient, although a fundamental mechanism for this pattern has yet to be identified. In general, it appears that major evolutionary innovations and speciation events occur more frequently in lower rather than higher latitudes and that these innovations subsequently spread toward the poles over comparatively long periods of time. The hypothesis of the equatorial region as a source of novelty is supported by trends in the invasion of land, the origin of seeds and angiosperms, and even the invasion of land by tetrapods, all of which appear to comply with a "tropics to the poles" trend (Jablonski, 1993).

Abundance

As ubiquitous as the latitudinal diversity gradient but possibly more mysterious are the patterns of abundance that are found in both modern and historical communities. Although it is generally agreed that increasing resource availability can increase diversity, very little is understood about how these resources are divided among members of a community. Abundance patterns in extant communities tend to conform to a "canonical" pattern: Most species are rare, whereas very few are common (Preston (1948, 1962) as cited in Brown, 1995; Rosenzweig, 1995; Rosenzweig (1998) as cited in McKinney and Drake, 1998).

Unfortunately, examination of past abundance patterns is confounded by preservational features. The plant fossil record consists of plant parts (dispersed leaves, fruits, wood, etc.) that are preserved in a manner that imperfectly reflects the abundance of species in communities. A deciduous tree, for example, within an otherwise evergreen forest might appear to dominate due to excessive production and shedding of organs. However, several Tertiary floras and a few Carboniferous ones are both geographically extensive and widely sampled, and they suggest that canonical abundance distributions may be a common feature in historical communities.

Dominance within a community, as reflected by abundance, appears to be a transient feature for an individual species. Indeed, the fossil record indicates that most species that were once common typically become rare with time (Knoll (1986) as cited in Rosenzweig, 1995). At a higher taxonomic level, however, certain groups do exhibit long-term dominance (e.g., the Mesozoic is the age of conifers and cycadophytes). This dominance likely reflects the increased genetic, developmental, morphological, and reproductive diversity inherent in pooling larger numbers of species within higher taxa. When members of a clade breach a developmental barrier, radically new solutions to life may arise that may allow for increased abundance via the colonization of new environments and/or the ability to outcompete with older clades. This dynamic may account for major transitions in dominance such as that between gymnosperms and angiosperms (**Figure 5**).

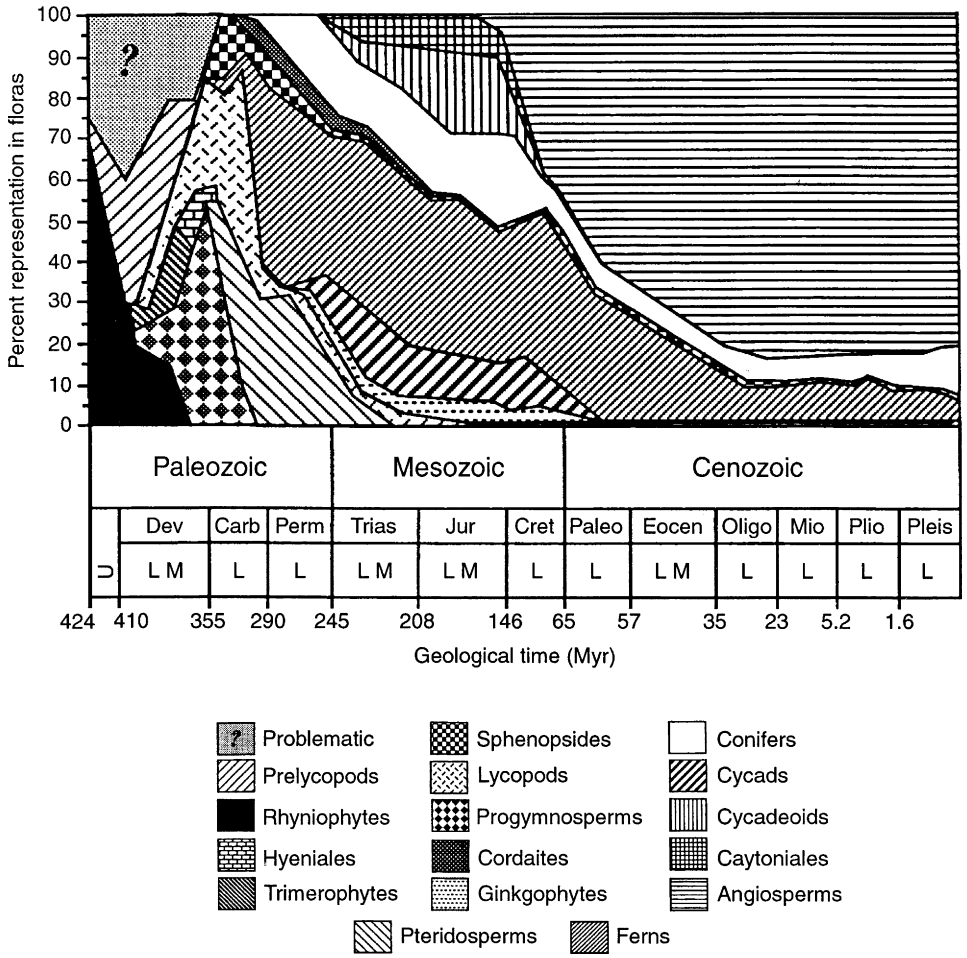


Figure 5 Changes in floristic composition. Taxonomic composition of representative floras (expressed as the percentage of each major group of vascular land plants) are plotted against geological time. Rhyniophytes and taxonomically "problematic" groups of plants dominate the first land plant floras; angiosperms dominate the most recent fossil plant floras.

Summary

Examination of the major patterns of plant diversity from both contemporary and historical communities holds the promise of illuminating the fundamental mechanisms responsible for the origin and maintenance of biological diversity. Here we have outlined several "rules" and patterns that appear to have been important in the evolution of plant communities. In many situations, the contemporary and historical evidence clearly complement each other. In other cases, such as the relative importance of competition, they appear to conflict. The latter cases may reflect the inherent research limitations of paleontology and neoecology, the real effects of temporal scale, or the operation of different processes.

Of particular note to ecologists interested in diversity should be identification of those factors that cause diversity to change at both short and long timescales. Dramatic climate change and evolutionary innovation appear to be the most important features in the historical record. It is interesting to note that the current diversity–stability debate has considered these variables not as cause but as effect, examining the influence of diversity on community and ecosystem processes rather than vice versa. We live in the most diverse terrestrial communities in the history of Earth, but they function as the result of the same basic rules of interaction, energetics, and stoichiometry as did their Paleozoic predecessors. When events, human or otherwise, create an extraordinary disturbance, biodiversity may plummet and modern communities, as we recognize them, may cease to exist. However, the evidence of 400 million years strongly suggests that the survivors will reorganize into communities and continue to function following the same basic physical, chemical, and ecological rules that created the world around us. They have done so before, and there is good reason to assume that they will continue to do so as long as this planet harbors life.

Acknowledgments

This article was completed as part of the Body-size Working Group supported by the National Center for Ecological Analysis and Synthesis, a center funded by NSF (Grant DEB-94–21535), the University of California at Santa Barbara (UCSB), and the state of California. We thank J. Damuth, W.

DiMichele, F. A. Smith, and J. H. Brown for commenting on initial versions of this article. BJE was supported by a NSF Postdoctoral Fellowship. JPH was supported by NSF Grant DEB-9707406 to J. H. Brown and by the National Center for Ecological Analysis and Synthesis. BHT acknowledges research support provided by the College of Creative Studies, UCSB.

See also: Coevolution. Diversity, Community/Regional Level. Habitat and Niche, Concept of. Latitudinal Gradients of Biodiversity. Paleoecology. Plant Biodiversity, Overview

References

Bateman RM, Crane PR, DiMichele WA, Kenrick PR, Rowe NP, Speck T, and Stein WE (1998) Early evolution of land plants: Phylogeny, Physiology, and Ecology of the primary terrestrial radiation. *Annu. Rev. Ecol. Syst.* 29: 263–292.

Brown JH (1995) *Macroecology.* Chicago: Univ. of Chicago Press.

Brown JH, Ernest SKM, Parody JM, and Haskell JP (2000) Regulation of diversity: Maintenance of species richness in changing environments. *Oecologia* 126: 321–332.

Crawley MJ (ed.) (1997) *Plant Ecology.* Cambridge, MA: Blackwell.

Crepet WL (1996) Timing in the evolution of derived floral characteristics: Upper Cretaceous (Turonian) taxa with Tricolpate and Tricolpate-derived pollen. *Rev. Paleobot. Palynol.* 90: 339–359.

DiMichele WA, Phillips TL, and Olmstead RG (1987) Opportunistic evolution, abiotic environmental stress, and the fossil record of plants. *Rev. Palaeobot. Palynol.* 50: 151–178.

Enquist BJ, West GB, Charnov EL, and Brown JH (1999) Allometric scaling of production and life-history variation in vascular plants. *Nature* 401: 907–911.

Jablonski D (1993) The tropics as a source of evolutionary novelty through geological time. *Nature* 364: 142–144.

McKinney ML and Drake JA (eds.) (1998) *Biodiversity Dynamics: Turnover of Populations, Taxa, and Communities.* New York: Columbia Univ. Press.

Niklas KJ (ed.) (1981) *Paleobotany, Paleoecology and Evolution,* Vol. 2. New York: Praeger.

Niklas KJ (1997) *The Evolutionary Biology of Plants.* Chicago: Univ. of Chicago Press.

Pellmyr O (1999) Systematic revision of the yucca moths in the *Tegeticula yuccasella* complex (Lepidoptera: Prodoxidae) north of Mexico. *Syst. Entomol.* 24: 243–271.

Ricklefs RE and Schluter D (eds.) (1993) *Species Diversity in Ecological Communities.* Chicago: Univ. of Chicago Press.

Rosenzweig ML (1995) *Species Diversity in Space and Time.* New York: Cambridge Univ. Press.

Sepkoski JJ (1991) A model of onshore–offshore change in faunal diversity. *Paleobiology* 17: 58–77.

Plant Conservation

Mike Maunder, Florida International University, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ecological restoration The Society for Ecological Restoration defines ecological restoration as an "intentional activity that initiates or accelerates the recovery of an ecosystem with respect to its health, integrity and sustainability." The term "ecological restoration" refers to the practice of the discipline of "restoration ecology."

Ecosystem services The benefits delivered or supplied to natural ecosystems. There are four main categories: provisioning, such as the production of food and water; regulating, such as the control of climate and disease; supporting, such as nutrient cycles and crop pollination; and cultural, such as spiritual and recreational benefits.

***Ex situ* conservation** "Off-site conservation," where the species or population is removed from the natural habitat and managed in a protected environment whether botanic garden, nursery, or stored as dormant seed or gametes for future propagation and use.

Extinction The end of an organism or of a group of organisms (taxon), normally a species. The moment of extinction is generally considered to be the death of the last individual of the species, although functional extinction may occur earlier as a species survives as a nonviable wild population.

***In situ* conservation** "On-site conservation" of a species or population, in the natural habitat and subject to prevailing pressures of natural selection and interacting with other species (for instance, symbiotic fungi or pollinating insects) and local cultural influences. For agricultural germplasm, *in situ* conservation would include farmed areas.

Threatened species Any of those species that are facing a risk of extinction in the near future. The IUCN-SSC categories of Threat recognize three categories, depending on the degree to which they are threatened: Vulnerable, Endangered, and Critically Endangered.

Introduction

Plants are essential to life on earth – they are a fundamental component of global biodiversity, the key structural element to terrestrial ecosystems, the basis of all terrestrial food webs, and are essential to the functioning of both terrestrial and marine ecosystems. Plants provide a wide range of essential ecosystem services, including the production of oxygen, sequestration of carbon dioxide, the creation and protection of soils and watersheds (Isbell *et al.*, 2011; Kumar, 2011), and the provision of timber, fibers, food, forage, and natural medicines. Importantly, they are also of immense cultural value to our species; through art, medicine, and horticulture there is an ancient and intimate link between plants and people. Plant diversity, as part of a global trend for biodiversity, is declining (Butchart *et al.*, 2010). The impact of a growing human population, as measured through habitat loss and degradation, the overharvesting of species, the impacts of invasive species, pollution, and climate change, is reducing global plant diversity and accelerating natural rates of extinction. Accordingly, our options for establishing a sustainable future reduced are eroding.

Historically, plant conservation focused on the plight of high-profile threatened plant species, frequently "odd" endemics with unusual histories; now, the challenge is not only to retain an increasing number of threatened species but also to retain habitats and the ecosystem services they deliver. This work encompasses not only the conservation of flowering plants (angiosperms) but also the rapidly declining gymnosperms and the many species of fern, algae, and moss (**Box 1**). Plant conservation is largely terrestrial in its focus but also encompasses the conservation of freshwater plants and the increasingly urgent conservation needs of coastal and marine plants, for instance mangroves and sea grasses (**Box 2**).

Plant conservation employs a wide variety of tools and resources to ensure the retention of current levels of plant diversity, the continuation of evolutionary processes, the provision of essential ecosystem services, and the avoidance of extinctions. This entails the deployment of a range of scientific, cultural, political, and economic tools. Importantly, plant conservation, while increasingly influenced by international and national policy, needs to be rooted in the cultures and customs of the locality, ultimately biodiversity will be retained because a community decides it is worth retaining, whether it is a Somerset meadow, a last patch of valley woodland in Yemen, or an expanse of Sahelian rangeland. This article provides a historical review of plant conservation issues and approaches and activities.

A History of Plant Conservation

Human societies have managed plant resources to ensure dependable access to timber, resins and gums, fruits, fibers, roots, and medicinals. This management has been driven largely by utility rather than any specific concern about the future of a species or a habitat. In many cultures, the retention of plant resources, whether grazing areas or sacred groves, also has a strong cultural and spiritual element. However, overharvesting is not a new problem; it is probably as old as our species. The cedars of Lebanon, *Cedrus libani*, now surviving as scattered and isolated groves in Lebanon, Syria, and Turkey, and covering probably less than 5% of their historical area, have been subject to a long history of management and

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00280-X

Box 1 The Conservation of salt water plant diversity

It is very easy to forget that the world's salt water ecosystems, the estuaries, coastal shelves, and oceans contain plants. Along the narrow coastal shelves human expansion has resulted in profound changes to habitats. Mangrove forests once covered more than 200,000 km^2 of tropical and subtropical coastlines; between 35% and 86% of this area has been lost over the past 25 years (Duke *et al*., 2007). Mangrove forests provide a number of essential ecosystem services – they provide nursing grounds for commercial fisheries, they provide timber for construction and charcoal, buffer other coastal habitats from river-borne siltation, sea level rise, storm surges, and tsunamis (Mangrove Futures Project: http://www.iucn.org/tsunami/).

Over half (55%) of the biological carbon (or green carbon) captured in the world is captured by marine living organisms, the so-called blue carbon. The ocean's vegetated habitats, in particular mangroves, salt marshes, and sea grasses, while covering less than 0.5% of the sea bed are vital carbon sinks, assimilating more than 50%, perhaps as much as 71%, of all carbon storage in ocean sediments. Coastal waters account for just 7% of the total area of the ocean. However, the extraordinary productivity of coastal ecosystems, such as coral reefs and sea grass pastures, means that these small habitat areas form the basis of the world's primary fishing grounds, supplying an estimated 50% of the world's fisheries. (Blue Carbon Report: http://www.grida.no/files/publications/blue-carbon/BlueCarbon_screen.pdf)

A study of extinction risks for sea grass species (Short *et al*., 2011) revealed that a third of the world's 72 species are in decline, with three species qualifying as Endangered (*sensu* IUCN). Sea grasses, like mangroves, provide invaluable ecosystem services, including providing habitat for coastal sea life, reducing shoreline erosion, and absorbing excess nutrients.

Marine phytoplankton are rarely featured in discussions about plant conservation; however, recent research indicates that the conservation of these marine organisms should be included in the agenda. Phytoplankton account for half of the planet's photosynthetic activity and drives marine ecosystem functioning. One study indicates that global phytoplankton is in decline (Boyce *et al*., 2010), with potential implications for geochemical cycling and fisheries.

Box 2 Crop wild relatives

The relatively few crops that provide most of the food we eat are the product of centuries of selection by breeders and the interaction of domestic and wild populations. Crop wild relatives (CWR) are the wild plant species that are genetically related to those crops, but unlike them, have not been domesticated. CWR are increasingly threatened by habitat loss and fragmentation, with the risk that invaluable genes could be lost to agriculture. However, these habitats are often ephemeral and not always incorporated into traditional protected areas portfolios.

CWR have provided novel genetic diversity that can improve productivity and crop quality, increase pest and disease resistance, and improve the tolerance of the crop to adverse conditions (Maxted and Kell, 2009). The need for these resources will only increase as agriculture spreads to new and more marginal territories and climate change impacts existing production lands and systems. The value of CWR resources has been proven. For instance, in the 1970s, the US maize crop was severely threatened by corn blight that destroyed almost US$ 1000 million worth of maize and reduced yields by as much as 50% in 1978. The problem was quickly resolved through the use of blight-resistant genes from wild varieties of Mexican maize.

There are between 50,000 and 60,000 CWR in the wild. The conservation of these invaluable genetic resources will be delivered partly through gene banks; however, it is essential that the evolutionary dynamics of the populations are maintained through *in situ* conservation both in traditional nature reserves and on farmed landscapes. The Crop Wild Relatives Specialist Group is coordinating a global program for CWR conservation, working with a variety of partners including Bioversity International.

overexploitation for over 2000 years; indeed, the marker stones to Roman forest reserves have been found on Lebanese hillsides.

Although classical authors recognized the human-caused degradation of landscapes – for instance Plato in the Critias, describes the landscape of Attica as resembling "the wasted skeleton of a sick man" – the linking of human-caused habitat degradation and species loss was established only during the eighteenth century. This was largely as a reaction to catastrophic ecological changes following colonial occupation of oceanic islands such as Mauritius (Grove, 1996). For instance, one of the earliest *ex situ* conservation attempts can be attributed to Governor Byfield of St Helena, who, in the early eighteenth century, collected two seedlings of the then very scarce St. Helena redwood, *Trochetiopsis erythroxylon*, for cultivation in his garden on the island.

The extinction of the monotypic endemic *Franklinia alatamaha* in the USA helped to establish the concept of plant extinctions. Similarly, during the late nineteenth century and early twentieth century, a small number of species were assumed to have become extinct in the wild, notably *Ginkgo biloba* and *Amherstia nobilis*; both species were considered to be extinct in the wild and only surviving in Asian temple gardens. However, these appear to have been treated as isolated novelties and did not prompt any general concern or action from the botanical community about a growing extinction problem.

The broader environmental movement can be traced, in part, to the late nineteenth century protected area movement in the USA (the Yellowstone National Park Act of 1872) and the wilderness-inspired literature of Aldo Leopold. The 1960s saw the start of a public concern for the environment and the beginning of international and national structures for conservation. This concern led to a public and political environmental awareness of the 1970s as manifested through the establishment of Earth Day on 22 April 1970 in the USA and the Endangered Species Act of 1973. The Russian agricultural botanist N.I. Vavilov, during the 1920s and 1930s, promoted the value of crop relatives and wild species, now referred to as plant genetic resources, in supporting agriculture. The present concerns about plant conservation, both as a political and as a scientific issue, can be traced to a series of international conferences on plant genetic resources during the late 1960s and 1970s. The erosion of these resources was recognized as an urgent problem at a Food and Agriculture Organization (FAO) conference in 1961, and subsequently discussed by a joint FAO and International Biological programme (IBP) meeting in 1967, where the term "genetic resources" was introduced. The FAO Panel of Experts on Plant Exploration and Introduction led the debate on plant resources from the mid 1960s to 1970s, with the creation in 1974 of the International Board for Plant Genetic Resources (now Bioversity International). Crop wild relatives remain a global priority for

Box 3 The South Atlantic Island of Ascension – too late for habitat restoration?

On discovery in 1501, the island was a bare volcanic landscape with vast sea bird colonies supporting only 25 vascular plant species; of these, 11 are endemic. However, as a result of nineteenth-century garrison settlement, a large number of exotic plants were introduced for both agricultural and horticultural purposes. The effects of exotic species have transformed the original bare volcanic landscape of Ascension into a semiforested one. Successive importations of plants from British and South African botanic gardens represented massive introductions of both plant species and associated invertebrate and fungal diversity. Consequently, the original upland *Marattia* fern thicket has been transformed into exotic woodland dominated by an introduced bamboo. In the middle altitudes, introduced *Casuarina*, *Araucaria*, and *Juniperus bermudiana* are forming open woodlands. The xeric lowland lava fields have been transformed into open woodland of *Prosopis juliiflora* with introduced African grass species. The *Prosopis* is now encroaching on the turtle (*Chelonia mydas*) nesting beaches and threatening to expand onto the unique breeding colonies of Sooty Tern (*Sterna fusca*). The transformation of terrestrial ecology has been massive and habitat restoration (*sensu strictu*) is no longer feasible; the priority should be to maintain and promote specific species and communities, for example, sea bird colonies, through the control of introduced predators (cats and rats) and herbivores (sheep and donkeys). The net effect has been an increase in total botanical diversity and primary productivity and the creation of a new synthetic ecosystem. However, in 2010, a specimen of the endemic fern, *Anogramma ascensionis*, commonly known as the parsley fern, was discovered, thought extinct since the 1950s, demonstrating the remarkable ability of some species to persist in highly modified landscapes. Ascension Island is developing as a novel ecosystem and as such indicates the future of many landscapes and ecologies.

conservation – their conservation and use being an essential foundation for sustainable agriculture (**Box 3**).

The United Nations Conference on the Human Environment held at Stockholm in 1972 is recognized as a significant starting point in global plant conservation. The Threatened Plants Committee (TPC) of the International Union for the Conservation of Nature (IUCN) was established in 1974 as a direct product of the initial "Red Data Book for Threatened Plant Species." This network was designed to deal specifically with plant species outside of the remit of FAO. The TPC acted as the template for the subsequent evolution of the IUCNs Species Survival Commission (SSC).

The discussions in the early and mid 1970s set the stage for botanic gardens and herbaria to take up leadership in plant conservation. The "Ark Paradigm," the idea that *ex situ* facilities would hold stocks of threatened plants during a period of habitat degradation, was established as a working objective by botanic gardens in the 1970s. This attitude is manifest in the 1978 Red Data Book (Lucas and Synge, 1978); for instance in the data sheet for *Dracaena ombet*, it was stated that "it seems too late for such a proposal (*in situ* conservation) to be worthwhile. Great efforts must now be made to bring the ombet into cultivation and maintain it safely in the botanic gardens of the world." A notable conference in the late 1970s was "Extinction is Forever" (Prance and Elias, 1977); this conference attempted to assess levels of species decline and loss in the Americas and subsequently set the scene for many later developments in tropical forest conservation in the Americas.

Two conferences hosted by the Royal Botanic Gardens, Kew, in 1976 and 1978, established the agenda for botanic garden activities over the next two decades (Simmons *et al.*, 1976; Synge and Townsend, 1979). As a result of the 1976 conference, the Threatened Plants Committee of the SSC of the IUCN was requested to establish a Botanic Gardens Conservation Co-ordinating Body: "to find out which threatened plants are in cultivation and where, and to keep botanic gardens informed of current conservation activities." This need was reiterated at the First International Botanic Gardens Conservation Congress in Las Palmas, Canary Islands, in 1985 with the recommendation of the establishment by IUCN of the Botanic Gardens Conservation Secretariat (BGCS) later to become Botanic Gardens Conservation International (BGCI). In parallel to discussions on botanic gardens, a number of broader conservation references were published in the early 1980s that paved the way for the 1992 UNCED meeting – The "Earth Summit." These included "The Brandt Report" (Independent Commission on International Development Issues, 1980), the Global 2000 Report (US Council on Environmental Quality and US Department of State, 1980), and the World Conservation Strategy (IUCN/UNEP/WWF, 1980). All three reports emphasized the linkages between the ongoing loss of species and habitats to human welfare and development needs, and subsequently paved the way for the Convention on Biological Diversity (CBD).

The CBD, ratified in 1992 at the Rio "Earth Summit," is the major legislative influence on the conservation of biodiversity, directly influencing national activities through the requirement to produce and implement Biodiversity Action Plans. The main objectives of the CBD are:

- the conservation of biological diversity (Articles 6–9, 11, and 14);
- the sustainable use of its components (Articles 6, 10, and 14); and
- the fair and equitable sharing of the benefits arising from the use of genetic resources, including by appropriate (1) access to genetic resources (Article 15), taking into account all rights over those resources; (2) transfer of relevant technologies (Articles 16 and 19), taking into account all rights to technologies; and (3) funding (Articles 20 and 21).

The CBD created an opportunity for plant conservation to become increasingly linked with global biodiversity and sustainable development imperatives. Accordingly, The Global Strategy for Plant Conservation (GSPC) was established as a program of the Convention on Biological Diversity (Wyse Jackson and Kennedy, 2009). The GSPC was adopted by the Parties of the CBD in April 2002. The GSPC has had a profound impact on both the strategic and the practical implementation of plant conservation. The GSPCs goal is to halt the continuing loss of plant diversity through five key themes:

1. Understanding and documenting plant diversity;
2. Conserving plant diversity;
3. Using plant diversity sustainably;
4. Promoting education and awareness about plant diversity;
5. Building capacity for the conservation of plant diversity.

Importantly, the GSPC established a direct link with the Millennium Development Goals, particularly in the areas of poverty reduction (Goal 1), combating disease (Goal 6), and promoting environmental sustainability (Goal 7).

The GSPC has been revised and new targets have been established for 2011–2020 that link with and support the 20 Aichi targets of the Strategic Plan for Biodiversity (2011–2020) (**Table 1**).

Although significant steps have been taken to establish plant conservation as a global priority, it could be argued that the practical implementation of plant conservation has not received sufficient attention. Although there have been significant and very important advances in global and national policy and strategy, new red listing initiatives, and the identification of important plant areas (IPA), practical "on the ground" implementation has been limited. The GSPC has been an extraordinary breakthrough for plant conservation; however, its implementation now needs to focus at a practical level.

The Distribution and Loss of Plant Diversity

The planet supports around 310,000 described plant species, with a potential total number of species estimated at possibly as high as 400,000 species (Govaerts, 2003). This diversity of species is concentrated into areas of unusual richness and exhibits variation at both a global and a regional scale. The taxonomic diversity of plants broadly correlates to latitude, and is generally the highest in the tropical and subtropical regions with high rainfall and mountains. Species richness and endemism were commonly used attributes for the identification of conservation priorities since they presumably reflected both the complexity and the uniqueness of ecosystems. Accordingly, early discussions identified tropical rainforests and oceanic islands as priorities.

In 1986, the International Union for the Conservation of Nature (IUCN) and the Worldwide Fund for Nature (WWF) initiated a project to identify major centers of plant diversity. The objectives of the project were: (1) to identify which areas around the world, if conserved, would safeguard the greatest number of plant species; (2) to document the many benefits, economic, and scientific that conservation of those areas would bring to society and to outline the potential value of each for sustainable development; and (3) to outline a strategy for the conservation of the areas selected. The project (Davis *et al.*, 1994, 1996, 1997) identified 234 areas of especially high species diversity, and correlates strongly with the listed 12 Megadiversity countries (*sensu* Mittermeier *et al.*, 1998). The "megadiversity countries" is based on three premises: (1) biodiversity is not evenly distributed among the world's countries; (2) although international conservation priorities should be based on scientific information on biodiversity, it is governments that develop conservation policies and programs; and (3) a very small number of countries, lying wholly or partly within the tropics, contain a high percentage of the world's species.

Myers (1988, 1990) identified conservation "hot-spot" areas, defined as areas of exceptional concentration of plant species, high endemism, and threatened with destruction. Myers initially identified 10 hot spots, all in tropical rainforest areas. The analysis was expanded to 18, and subsequently, 34 biodiversity hotspots have been identified (Mittermeier *et al.*, 2004). This expanded portfolio of hot spots holds an estimated 150,000 endemic plant species, 50% of the world's total. An analysis of projected habitat loss in 25 hotspots (Brooks *et al.*, 2002, 2006) indicates a very high risk of large-scale extinctions of endemic species. The Caribbean Islands illustrate many of the challenges facing plant conservationists (**Box 4**).

To identify priority areas for plant conservation at a national or a regional level, the concept of IPAs was developed. They provide a method for identifying and maintaining the richest sites for botanical diversity (algae, fungi, lichens, liverworts, mosses, and wild vascular plants). IPAs are selected with the intention of focusing on the conservation of the wild plant populations in these areas, and act as a subset in the broader context of Key Biodiversity Areas. The identification of IPAs is based on three criteria:

1. Presence of threatened plant species: the site holds significant populations of one or more species that are of global or regional conservation concern,
2. Presence of botanical richness: the site has an exceptionally rich flora in a regional context in relation to its biogeographic zone, and
3. Presence of threatened habitats: the site is an outstanding example of a habitat or vegetation type of global or regional plant conservation and botanical importance. IPAs can be identified within high-diversity hot spot regions and in the less diverse temperate regions with far fewer endemics, for instance the UK has over 150 recognized IPAs.

Barthlott *et al.* (2007) identified 20 global centers of vascular plant diversity where species richness exceeds 3000 species per 10,000 km^2, the majority being tropical, with the exception of the Mediterranean Basin, the Caucasus, the Cape, SW Australia, and the Himalayas. Within this group are five global centers where diversity exceeds 5000 species per 10,000 km^2 (**Tables 2** and **3**).

Human Influences on Plant Diversity

Habitat loss and degradation are the primary drivers for the loss of plant diversity. Historically, about half of the ice-free surface of the globe has been transformed, managed, or utilized by humans (Turner *et al.*, 1990), with an increasing proportion of the world's primary productivity used by mankind (Haberl *et al.*, 2007). The Global Analysis of Land Degradation and Improvement (Bai *et al.*, 2010) estimates that nearly 24% of the world's land area underwent degradation (as measured by declines in primary productivity) over the period 1980–2003. The Millenium Ecosystem Assessment estimated that about 60% of evaluated ecosystem services are being degraded or used unsustainably. In some regions, the loss of habitat has been near complete, for instance 90% loss of natural forest in Madagascar (Harper *et al.*, 2007).

The impacts of habitat degradation in the wet tropics have been successfully relayed to the general public; however, the

Table 1 The GSPC 2011–2020 Targets and linked targets for the Strategic Plan for Biodiversity (2011–2020)

GSPC target 2011–2020	*Relevant Target from Aichi targets of Strategic Plan for Biodiversity (2011–2020)*
1. An online flora of all known plants	T19. By 2020, knowledge, the science base and technologies relating to biodiversity, its values, functioning, status and trends, and the consequences of its loss, are improved, widely shared and transferred, and applied
2. An assessment of the conservation status of all known plant species, as far as possible, to guide plant conservation action	T19. By 2020, knowledge, the science base and technologies relating to biodiversity, its values, functioning, status and trends, and the consequences of its loss, are improved, widely shared and transferred, and applied
3. Information, research, and associated outputs, and methods necessary to implement the Strategy developed and shared	T19. By 2020, knowledge, the science base, and technologies relating to biodiversity, its values, functioning, status and trends, and the consequences of its loss, are improved, widely shared and transferred, and applied
4. At least 15% of each ecological region or vegetation type secured through active management and/or restoration	T5. By 2020, the rate of loss of all natural habitats, including forests, is at least halved and where feasible brought close to zero, and degradation and fragmentation is significantly reduced T11. By 2020, at least 17% of terrestrial and inland water areas, and 10% of coastal and marine areas… are conserved through effectively and equitably managed, ecologically representative and well-connected systems of protected areas….and integrated into the wider landscapes and seascapes T15. By 2020, ecosystem resilience and the contribution of biodiversity to carbon stocks has been enhanced, through conservation and restoration, including restoration of at least 15% of degraded ecosystems, thereby contributing to climate change mitigation and adaptation and to combating desertification
5. At least 75% of the most important areas for plant diversity of each ecological region protected with effective management in place for conserving plants and their genetic diversity	T11. By 2020, at least 17% of terrestrial and inland water areas, and 10% of coastal and marine areas….are conserved through effectively and equitably managed, ecologically representative and well-connected systems of protected areas….and integrated into the wider landscapes and seascapes
6. At least 75% of production lands in each sector managed sustainably, consistent with the conservation of plant diversity	T7. By 2020 areas under agriculture, aquaculture and forestry are managed sustainably, ensuring conservation of biodiversity
7. At least 75% of known threatened plant species conserved *in situ*	T12. By 2020 the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained
8. At least 75% of threatened plant species in *ex situ* collections, preferably in the country of origin, and at least 20% available for recovery and restoration programs	T12. By 2020 the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained
9. 70% of the genetic diversity of crops including their wild relatives and other socio-economically valuable plant species conserved, while respecting, preserving, and maintaining associated indigenous and local knowledge	T13. By 2020, the genetic diversity of cultivated plants and farmed and domesticated animals and of wild relatives, including socio-economically as well as culturally valuable species, is maintained, and strategies have been developed and implemented for minimizing genetic erosion and safeguarding their genetic diversity
10. Effective management plans in place to prevent new biological invasions and to manage important areas for plant diversity that are invaded	T9. By 2020, invasive alien species and pathways are identified and prioritized, priority species are controlled or eradicated, and measures are in place to manage pathways to prevent their introduction and establishment
11. No species of wild flora endangered by international trade	T4. By 2020, at the latest, Governments, business and stakeholders at all levels have taken steps to achieve or have implemented plans for sustainable production and consumption and have kept the impacts of use of natural resources well within safe ecological limits T6. By 2020, all fish and invertebrate stocks and aquatic plants are managed and harvested sustainably, legally and applying ecosystem based approaches, so that overfishing is avoided, recovery plans and measures are in place for all depleted species, fisheries have no significant adverse impacts on threatened species and vulnerable ecosystems and the impacts of fisheries on stocks, species, and ecosystems are within safe ecological limits
12. All wild harvested plant-based products sourced sustainably	T4. By 2020, at the latest, Governments, business and stakeholders at all levels have taken steps to achieve or have implemented plans for sustainable production and consumption and have kept the impacts of use of natural resources well within safe ecological limits T6. By 2020 all fish and invertebrate stocks and aquatic plants are managed and harvested sustainably, legally, and applying ecosystem based approaches, so that overfishing is avoided, recovery plans and measures are in place for all depleted species, fisheries have no significant adverse impacts on threatened

(*Continued*)

Table 1 Continued

GSPC target 2011–2020	*Relevant Target from Aichi targets of Strategic Plan for Biodiversity (2011–2020)*
	species and vulnerable ecosystems and the impacts of fisheries on stocks, species, and ecosystems are within safe ecological limits
13. Indigenous and local knowledge innovations and practices associated with plant resources maintained or increased, as appropriate, to support customary use, sustainable livelihoods, local food security, and healthcare	T18. By 2020, the traditional knowledge, innovations and practices of indigenous and local communities relevant for the conservation and sustainable use of biodiversity, and their customary use of biological resources, are respected, subject to national legislation and relevant international obligations, and fully integrated and reflected in the implementation of the Convention with the full and effective participation of indigenous and local communities, at all relevant levels
14. The importance of plant diversity and the need for its conservation incorporated into communication, education and public awareness programs	T1. By 2020, at the latest, people are aware of the values of biodiversity, and the steps they can take to conserve and use it sustainably
15. The number of trained people working with appropriate facilities sufficient according to national needs, to achieve the targets of this strategy	T20. By 2020, at the latest, the mobilization of financial resources for effectively implementing the Strategic Plan for Biodiversity 2011–2020 from all sources, and in accordance with the consolidated and agreed process in the Strategy for Resource Mobilization, should increase substantially from the current levels. This target will be subject to changes contingent to resource needs assessments to be developed and reported by Parties
T16. Institutions, networks and partnerships for plant conservation established or strengthened at national, regional, and international levels to achieve the targets of this strategy	T17. By 2015 each Party has developed, adopted as a policy instrument, and has commenced implementing an effective, participatory and updated national biodiversity strategy, and action plan

Box 4 The living dead

Plant conservationists face a challenge when dealing with the "living dead," those species that survive as a few scattered, nonreproductive individuals, often surviving in an exotic dominated landscape without their pollinators or seed dispersal agents. Oceanic islands, having suffered massive levels of habitat loss, contain many examples of the living dead. For example both *Kokia cookei* from Hawaii and *Ramosmania rodriguesiana* from Rodrigues survive only as sterile single clones. Indeed, *K. cookei* survives as a grafted single clone. The endemic palm, *Hyophorbe amaricaulis*, survives only as a single tree in the Curepipe Botanic Gardens, southwest Mauritius. The single specimen has persisted within the botanic garden for over 50 years, with no regeneration recorded.

The Malvaceae illustrate many of the problems facing critically endangered island plants. In Hawaii, the two endemic Malvaceae genera *Kokia* and *Hibiscadelphus* are highly threatened. Of the seven described species of *Hibiscadelphus*, four species are known only from the original type specimen, three are Extinct, whereas *Hibiscadelphus woodii* survives with only four wild plants. Similarly, with the genus *Kokia*, two species are Critically Endangered, *K. cookei* survives only in cultivation, whereas *Kokia lanceolata* is Extinct. The endemic Hawaiian species of *Hibiscus* and *Abutilon* are all threatened.

As a result of intensive management, some "living dead species" are recovering; a good example is the case of the Round Island Bottle Palm, *Hyophorbe lagenicaulis*. As a result of invasive species control, specifically goat and rabbit, the wild population is recovering (Maunder *et al.*, 2002), while the potential exists to utilize cultivated stocks to compensate for the population bottle neck suffered by the wild population (Asmussen-Lange *et al.*, 2011).

impact on arid lands and rangelands is less widely recognized (Reynolds *et al.*, 2007). Dry or arid areas cover 41% of the Earth's surface and are home to over 1.7 billion people, including a large proportion of the world's poorest peoples. The Millennium Ecosystem Assessment Report (2005) indicates that desertification threatens over 41% of the Earth's land area; between 20% and 70% of dry lands are already degraded, resulting in a decline in agricultural productivity, loss of biodiversity, and the breakdown of ecosystems. Environmental goods and services such as fuel wood, harvested gums and resins, medicinal and food plants, habitat for wild animals and grazing for livestock, soil fertility, and soil moisture for agricultural production are all being lost. Desertification currently impacts the livelihoods of some 250 million people in the developing world. These changes not only influence the livelihoods of millions of people but also have a direct impact on biodiversity – the loss of pasture or charcoal species, the loss of wildlife (e.g., antelope populations in the Sahel), or the near extinction of endemic passerines in Ethiopia (Spottiswoode *et al.*, 2009). In addition to the degradation of extensive rangelands, the arid regions are also suffering the loss of unique and often very localized habitats; these include the valley forests of Yemen (Arabia's only closed canopy evergreen forest), the fog forests of Oman and Chile (the lomas of the Atacama), and the spiny forests of Madagascar. The arid areas contain some of the world's most iconic plant species, such as the tree cacti of the neotropics, the baobabs of Africa and Madagascar, and important economic resources such as gum Arabic and frankincense.

Deforestation continues to be a major cause of decline for global plant diversity. Around 13 million ha of forest were converted to other uses or lost through natural causes each year in the decade 2000–2010, compared with 16 million ha per year in the 1990s. The net change in forest area in the period 2000–2010 is estimated at 5.2 million ha per year (an area about the size of Costa Rica), down from 8.3 million ha per year in the period 1990–2000 (FAO, 2010). South America and Africa continue to have the largest net loss of forest.

Table 2 Numbers of plant species occurring in and endemic to each of the 34 hotspots (CI, http://www.conservation.org/where/priority_areas/hotspots/hotspots_revisited/key_findings/Pages/endemic_plant_species.aspx)

Hotspot	*Plant species*	*Endemic plant species*	*Endemics as a percentage of world total*
Atlantic Forest[a]	20,000	8000	2.7
California Floristic Province	3488	2124	0.7
Cape Floristic Region	9000	6210	2.1
Caribbean Islands[a]	13,000	6550	2.2
Caucasus[a]	6400	1600	0.5
Cerrado	10,000	4400	1.5
Chilean Winter Rainfall – Valdivian Forests	3892	1957	0.7
Coastal Forests of Eastern Africa	4000	1750	0.6
East Melanesian Islands	8000	3000	1.0
Eastern Afromontane	7598	2356	0.8
Guinean Forests of West Africa[a]	9000	1800	0.6
Himalaya[a]	10,000	3160	1.1
Horn of Africa	5000	2750	0.9
Indo – Burma	13,500	7000	2.3
Irano – Anatolian	6000	2500	0.8
Japan	5600	1950	0.7
Madagascar and the Indian Ocean Islands[a]	13,000	11,600	3.9
Madrean Pine – Oak Woodlands	5300	3975	1.3
Maputaland – Pondoland – Albany[a]	8100	1900	0.6
Mediterranean Basin[a]	22,500	11700	3.9
Mesoamerica[a]	17,000	2941	1.0
Mountains of Central Asia	5500	1500	0.5
Mountains of Southwest China[a]	12,000	3500	1.2
New Caledonia	3270	2432	0.8
New Zealand	2300	1865	0.6
Philippines	9253	6091	2.0
Polynesia – Micronesia	5330	3074	1.0
Southwest Australia[a]	5571	2948	1.0
Succulent Karoo	6356	2439	0.8
Sundaland	25,000	15,000	5.0
Tropical Andes[a]	30,000	15,000	5.0
Tumbes – Choco[a]	11,000	2750	0.9
Wallacea	10,000	1500	0.5
Western Ghats and Sri Lanka[a]	5916	3049	1.0

[a]Hotspots overlapping, at least in part, with centers for vascular plant diversity by Barthlott *et al.* (2007).

Table 3 The five global centers for vascular plant diversity (*sensu* Barthlott *et al.*, 2007)

Center	*Area (km)*	*Total spp.*	*Endemism spp.*	*% endemism*	*% protected*
Costa Rica-Choco	78,000	12,500	5500	44	18.8
Tropical Eastern Andes	62,000	10,000	3000	30	19.1
Atlantic Brazil	50,000	6000	4500	75	6.3%
Northern Borneo	57,000	9000	3500	39	7.7
New Guinea	87,000	6000	2000	33	1.8

South America suffered the largest net loss of forests between 2000 and 2010, at about 4.0 million ha per year, followed by Africa, which lost 3.4 million ha annually. Asia, which had a net loss of forest of some 600,000 ha annually in the 1990s, reported a net gain of forest of more than 2.2 million ha per year in the period 2000–2010, primarily due to the large-scale afforestation reported by China and despite continued high rates of net loss in SE Asia driven by industrial oil palm plantation expansion.

The extensive conversion of wild habitats has resulted in changes to both the ecological and the taxonomic composition of these areas. In many parts of the world, the change is almost complete; taxonomically and ecologically diverse wildlands have been converted to agricultural landscapes dominated by a small range of domesticates, for example, cereal monocultures replacing species rich steppes and prairie, oil palm replacing high-diversity lowland rainforest.

The Mediterranean biomes, as exemplified by the Mediterranean Basin and the California Floristic Province, are subject to increasing human population densities and urban spread and high levels of invasion by exotic species (Schwartz *et al.*, 2006; Underwood *et al.*, 2009). The Mediterranean Basin

has been subject to the near complete transformation of its original ecology, a process that can be traced back to the Neolithic. The contemporary pressures of intensive agriculture, urbanization, water extraction, tourism developments, changing fire ecology, and invasive species are accelerating the rate of change.

Spectacular patterns of degradation have occurred on oceanic islands. Oceanic islands are rich in unique endemic lineages and particularly susceptible to damage by invasive species (Sax and Gaines, 2008; Caujape-Castells *et al.*, 2009). Colonized Pacific islands were subject to profound environmental changes and high levels of avian, and presumably plant, extinctions following Polynesian settlement. One example, the Pacific island of Rapa Nui (Easter Island), underwent massive environmental degradation as long ago as 1200–800 BP. In the Caribbean, Atlantic islands, and Mascarenes, environmental degradation can be attributed to European colonial administration and the development of unstable and fragile economies based on plantation products. Colonial settlement would often focus on the clearance and agricultural development of the lowlands, with the retention, planned or otherwise, of upland habitats, albeit in a modified and fragmented form. It could be argued that the remote South Atlantic island of Ascension exemplifies the future of many island ecosystems changed beyond recognition as a result of invasive species (**Box 5**).

Some islands have suffered the virtual complete loss of original habitats. This can result from periods of overexploitation over centuries. On Rapa Nui, no original plant communities have survived the Polynesian settlement and later colonial sheep ranching, the only surviving endemic tree, the toromiro or *Sophora toromiro*, is now Extinct in the Wild (Maunder *et al.*, 1999). On St Helena, only fragments of gumwood forest and montane tree fern thickets have survived, covering less than 1% of the island's land area (Maunder *et al.*, 1995). Plant conservation on oceanic islands often focuses on retaining and restoring fragments of the original biodiversity, sometimes the last individuals of a species, the "living dead" (**Box 6**).

As a result of deliberate introductions and as artifacts of the agricultural and horticultural industries, an increasing number of exotic species are growing in new territories and habitats. Although such introductions may in the short term increase local biodiversity, in the long term, such introductions are having disastrous impacts on both species diversity and ecological processes. Disastrous invasions include the impact of guava in Mauritius, prosopis in Ethiopia, miconia on the Polynesian Islands, and water hyacinth in the East African Rift Valley Lakes. Such invasions can have profound impacts on ecosystems for instance the invasion of the arid southwest of the USA by Mediterranean grasses has increased the impact of fires on succulent vegetations, with disastrous results. In addition, introduced diseases, such as chestnut blight in the USA, and pests, such as the New Zealand flat worm in the UK or the hemlock woody adelgid in North America, are impacting on plant diversity. Oceanic islands have in particular suffered from catastrophic declines in plant diversity as a result of invasive species, including the impacts of rats, goats, pigs, and ants, and diseases such as avian malaria killing bird pollinators.

Box 5 The Conservation of mosses

Mosses have developed during the past 20 years as a key component of plant conservation (Hallingback and Tan, 2010). Mosses and liverworts (Bryophytes) play a fundamental role in ecosystem services. Peat moss (*Sphagnum*) has a direct economic value as peat and plays a huge role in carbon sequestration as peat bogs by holding billions of tons of carbon as peat. Mosses play a key role in water cycles (particularly in temperate rainforests) and their value in medicine is being investigated. The Bryophyte Specialist Group is leading a global red list initiative and promoting their conservation. The Bryophyte Specialist Group has produced harvesting guidelines for commercially harvested species and is identifying centers of bryophyte endemism. A particular concern is the lack of trained moss taxonomists and ecologists in nations with rich bryophyte floras.

Recent studies of bryophyte floras have revealed unexpected concentrations of diversity that require protection. In Western Europe, the British and Irish Atlantic oak woods hold one of the richest bryophyte floras in the world. There are about 1770 bryophyte species in Europe; of these, 1150 are found in Britain. This equates to about 65% of the European bryophyte flora.

Box 6 Conserving the plant diversity of the Caribbean

The land area of the Caribbean hotspot is relatively small, a total land area of approximately 229,550 km^2; the islands support a vascular flora of 205 native plant families, with 1447 genera and 10,950 species, with 183 endemic genera and approximately 7800 species of endemic vascular plants. A large proportion of the 183 endemic genera are represented by a single species (94), with 105 of the 183 endemic taxa restricted to a single island, Cuba, holding 65 endemic genera (Francisco-Ortega *et al.*, 2007; Maunder *et al.*, 2008).

Many of the Caribbean's endemic species are restricted to areas of unique geology, for instance the extensive areas of serpentine (nickel-rich) soil in Cuba; although these serpentine soils account for only 7% of Cuba's land area, around 920 of Cuba's 6375 plant species are endemic to that habitat, as are 18 monospecific endemic genera.

The Caribbean daisies, Asteraceae, are an overlooked and extraordinary flagship family for the Caribbean. The Caribbean holds almost one-third of the world's insular endemic Asteraceae genera (41 of the 136 genera, approximately 30%). A total of 32 of the 41 Caribbean endemic genera are single-island endemics, with Cuba holding 17, Hispaniola 11, and four endemic to Jamaica; of the 41 endemic genera, a total of 26 are unispecific. The endemic Asteraceae genus *Feddea* was first collected in Eastern Cuba in 1915 and was subsequently recorded from 16 historic locations. Recent fieldwork discovered only one location with a total of five plants. *Feddea* has been placed in its own monotypic tribe, the only tribe in the Asteraceae that is endemic to an island.

Found through Cuba, Hispaniola, Jamaica, and Puerto Rico are spectacular karst uplands or mogotes; these limestone hills are rich in endemic vegetation and often represent the last wild habitats in an increasingly agricultural landscape. The conservation of these karst landscapes is essential both for retaining endemic biodiversity and importantly for maintaining water supplies to major cities and towns in those islands.

Human activities, notably the combustion of fossil fuels and the subsequent release of carbon dioxide, are altering the composition of the atmosphere. The Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) stated that: "continued GHG emissions at or above

current rates would cause further warming and induce changes in the global climate system during the twenty-first century that would very likely be larger than those observed during the twentieth century''. Global warming is expected to continue at ever-growing rates during this century and according to some scenarios the global temperature may rise by up to 6 °C (Richardson *et al.*, 2009). Understanding and preparing for the effects of climate change on plant species and communities is a major conservation challenge. The vast majority of conservation biologists accept the scientific evidence that climate change is proceeding and that we can expect profound changes to the planet's ecology and the delivery of plant conservation. However, there is still significant uncertainty on how species and ecosystems will respond to changing temperatures and importantly the changes to an ecosystem resulting from warmer temperatures. However, an increasing number of studies are indicating potentially large losses to biodiversity (Maclean and Wilson, 2011), the establishment of novel or new habitats or communities, and associated changes in ecosystem function (McClean *et al.*, 2005; Sommer *et al.*, 2010). Increasingly, adaptation to climate change is being built into conservation planning, for instance in Madagascar (Hannah *et al.*, 2008).

All of these damaging impacts are driven by one fundamental force: the increase in human population and the environmental demands of that population. The world population is currently estimated to be 7 billion. Current projections show a continued increase of population (but a steady decline in the population growth rate), with the global population expected to reach between 8.0 and 9.0 billion people by 2050. As a result of a globalized economy, volatile politics, and a fluctuating economy, it is difficult to predict where the next conservation challenge will arise. For instance, conservationists watched in alarm as illegal timber traders took advantage of the 2009 political upheaval in Madagascar to extract huge volumes of high-value *Dalbergi*a and *Diospyros* timbers for the luxury furniture trade in China. In the past decade, big agro-business has become the main driver for forest clearance in SE Asia and Amazonia in response to the huge global demands for oil palm, beef, and soya bean.

Traditional medicine is based largely on plants and supports the primary healthcare of more people worldwide than 'conventional' or western medicine. Up to 80% of the population in Africa, for example, uses traditional medicine for their primary healthcare according to the World Health Organisation. The majority of species of plants in traditional or herbal medical treatments are harvested in the wild rather than cultivated. About 15,000 species of medicinal plants are globally threatened, the causes including loss of habitat and commercial overharvesting (BGCI – http://www.bgci.org/files/Worldwide/Publications/PDFs/medicinal.pdf).

Threatened Plant Species – How Many and Where?

We have no comprehensive list of the world's threatened plant species. However, steady progress is being made in assessing the conservation status of the world's plant species, through individual species assessments (for instance using the IUCN Categories of Threat) or by regional and global estimates. What is clear is that the number of threatened species is increasing and the conservation challenge of retaining species diversity is becoming ever more demanding.

The most comprehensive survey of species decline was undertaken on behalf of the IUCN, using the old non-quantitative red listing system; this study indicated that about 33,400 plant species were threatened with extinction (Walter and Gillett, 1998), equating to about 10% of the world's approximately 300,000 described plant species. The current IUCN Red List, using the quantitative and universally applicable IUCN Categories of Threat, holds approximately 14,190 plant assessments (2011); of these, 9098 are classified as Critically Endangered/Endangered or Vulnerable. However, a far larger number of species have been assessed using a variety of other approaches. To expand the number of IUCN plant red list assessments for the proposed Barometer of Life project (Stuart *et al.*, 2010) to 38,500 records would cost an estimated 17,000,000 USD. The Sampled Red List initiative is an ongoing approach to gage global trends in plant diversity through assessments of 7000 plant species.

Brummitt *et al.* (2008) estimates that about 20% of the world's known plant species are threatened with extinction, whereas Joppa *et al.* (2010), incorporating estimates of potentially new species, estimate that about 30% of all plant species are probably threatened.

Although a comprehensive global data set does not exist as yet, a range of studies indicates the scale of the conservation challenge. The South African National Biodiversity Institute (SANBI) has completed a red list for all 20,456 plant taxa found in South Africa. The Cycad Specialist Group of the Species Survival Commission has identified that 62% of all cycad species are threatened with extinction, largely as a result of overharvesting for plant collectors and the medicinal plant trade, habitat loss, and invasive pest species. *Encephalartos whitelockii* from Uganda is threatened as a result of ongoing construction of a hydroelectric scheme in its sole location, the Mpanga Gorge, whereas *Cycas micronesica*, native to Micronesia, is threatened by two exotic pests (the butterfly *Chilades pandava* and the scale insect *Aulacaspis yasumatsui*). Four cycad species are Extinct in the Wild, including three species of *Encephalartos* (*Encephalartos brevifoliolatus*, *Encephalartos nubimontanus*, and *Encephalartos woodii*) from South Africa. Similarly, a study of *Magnolia* species by Botanic Gardens Conservation International indicates that 131 species of the total 245 species are threatened with extinction (http://www.bgci.org/files/Media_Kit/magnolia_red_list_.pdf).

An assessment of potential extinction rates for the Amazon based on ongoing habitat loss projects a loss of available habitat of between 12% and 24% with between 5% and 9% of the region's 50,000 plant species ''committed to extinction,'' equating to approximately 2500–4500 species (Feeley and Silman, 2009). A study of species conservation issues on oceanic islands (Caujape-Castells *et al.*, 2009) estimates that of the 70,000 island endemic plant species, between 3500 and 6800 may be highly threatened with extinction (qualifying for IUCN categories Endangered or Critically Endangered).

Apart from the well-publicized estimates of loss from the tropics, it is evident that the European nations are losing species as a result of changes in habitat quality, land use, invasive species, and climate change. In contrast to areas in the

tropics, European plant conservation frequently focuses on the retention of traditional agricultural practices such as hay meadow management or transhumance. In Europe, although taxonomic diversity has increased as a result of new invasive plant species, the region has suffered an associated decline in phylogenetic diversity and an increased taxonomic and phylogenetic homogenization across the European regions (Winter *et al.*, 2010).

Responses to Loss of Plant Diversity

Conservation actions are implemented to reduce or reverse the damaging impacts of human-driven change on plant diversity. Management actions aimed at the conservation of biodiversity will vary in the scale of spatial, capital, and scientific investment, vary in the extent of impact (local, regional, or global) encompassing single species management through to continental scale wilderness retention. Increasingly, effective plant conservation will include economic and social focused actions; conservation will increasingly focus on the issues of land conversion and economic exploitation of plant products. The traditional schism between on-site activities (*in situ* conservation) and off-site activities with a species or genetic focus (*ex situ* conservation) is rapidly eroding and an integrated approach to plant conservation is dominating (*sensu* Falk, Millar and Olwell, 1996). A parallel development has been an increasing trend toward ecosystem-level conservation (Weeks, 1996) in both the temperate and the tropical regions. In addition, the linkages between biodiversity management and sustainable development have been given due recognition, particularly with regard to maintaining key ecosystem services for rural communities, whether it be pasture, access to medicinal plants, or the protection of watersheds.

The majority of the world's species will be retained through the "coarse filter" approach of habitat conservation; this could potentially conserve all levels in the biodiversity hierarchy and their interactions. However, many protected areas will require increasing management because of external influences impacting on ecological processes and resulting in changes in both community structure and composition. Protected area borders are permeable to disease, invasive species, poaching, civil unrest, and climate change. Protected areas have been established with the assumption that environmental conditions and community patterns/composition have been relatively stable for long periods in the past and will continue to be stable into the future. However, climate change studies indicate that habitats are loosely organized collections of species whose coexistence is dependent on their individual physiological limits. Accordingly, protected areas must be managed in the expectation of ecological change and may well require a supporting "fine filter" approach to catch those species knocked out by climate change or invasive pathogens. The need to manage species in the light of changing climate is being actively discussed, one controversial option, assisted migration, being the movement of species out of recorded or historical ranges with the assumption that climate change will render the existing populations nonviable.

The surviving major wilderness areas (Mittermeier *et al.*, 1998), large relatively undisturbed natural areas, offer the best opportunities for retaining ecosystem and evolutionary processes. Beyond these areas, plant conservation will depend on a number of core skills: (1) the protection and active management of habitats to maintain plant diversity and ecological processes, (2) the management of individual plant populations to retain viable populations, and (3) managing the human context of plant conservation including sustainable use, cultural value, and public access.

Species recovery, the management activities required to halt or reverse the decline in a threatened species' population, can be best achieved through cross-disciplinary recovery teams applying the recommendations of a Species Recovery Plan. Such a program should have clear numerically based objectives utilizing, where appropriate, information from ecological, taxonomic, genetic, and social studies. A Species Recovery Plan provides a forum to bring all required expertise together to ensure a balanced integrated approach to species conservation.

Ecological restoration is a powerful tool for plant conservation; it has the potential to reinstate biological diversity and ecosystem processes and to provide a potent forum for community leadership in conservation. Restored lands, whether small habitat fragments or landscape-scale ecosystems, can support livelihoods, provide wildlife habitats, and provide a wide range of products and services. The potential is immense, with opportunities for post industrial and agricultural land use in the temperate regions and post forest clearance lands in the tropics. More than 2 billion ha worldwide are available for restoration and these lands represent an extraordinary opportunity to sustain ecosystem services and biodiversity, create corridors between surviving habitat areas, support local employment and agriculture and importantly, sequester carbon (WRI; http://www.wri.org/map/global-map-forest-landscape-restoration-opportunities). Restoration can take place at a local scale with strong community identity (for instance the St Helena Millenium Forest Project on the remote South Atlantic island of St Helena that is establishing a lost habitat the endemic Gumwood forest). At a larger continental scale is the Six State Restoration Project aiming to restore 1.1 million acres of habitat in Tennessee, Alabama, Missouri, Arkansas, Kentucky, and Illinois. In the wet tropics, the Harapan Rainforest restoration project, a partnership between the RSPB, Birdlife International and Burung Indonesia, which will restore 98,000 ha, shows the potential for large-scale habitat restoration.

Increasingly *ex situ* collections are an integral component of conservation programs for both wild taxa and crop genetic resources (Havens *et al.*, 2006). The world's *ex situ* plant resources encompass the traditional garden plots of the tropics, the intensive allotments, farms and gardens of Europe, and the global networks of seedbanks and botanic gardens. The worlds' network of plant genetic resource collections (e.g., seedbanks and field gene banks) contains relatively few species, but with high levels of infraspecific sampling. The CGIAR Eleven Centres together maintain over 650,000 samples of crop, forage, and agroforestry genetic resources in the public domain.

The most important 150 major crop gene banks are maintained, as part of the global network of CGIAR collections (http://www.cgiar.org/impact/accessions.html). These include the International Potato Center in Lima, Peru, with

over 5000 potato accessions, the International Rice Research Institute at Los Banos, the Philippines, with over 80,000 rice accessions, and International Maize and Wheat Improvement Centre in Mexico, with over 20,000 maize accessions.

Two recent developments illustrate a new sense of innovation and urgency for seed banking. The Svalbard Global Seed Vault (http://www.regjeringen.no/en/dep/lmd/campain/svalbard-global-seed-vault.htmlid=462220), situated in the arctic island of Svalbard, was opened in 2008, a converted old mine shaft in permafrost adapted as a global safety-storage for the storage of duplicate collections of seeds on behalf of other gene banks. The Seed Vault has the capacity to store 4.5 million different seed samples. The Millenium Seed Bank of the Royal Botanic Gardens, Kew, is the largest repository of wild collected seed for noncrop species (http://www.kew.org/science-conservation/save-seed-prosper/millennium-seed-bank/index.htm). Launched in 1996 and occupying a custom-built complex of seed vaults, laboratories, and public exhibition galleries, the MSB now holds over 30,855 species (including 12 Extinct in the Wild) and a total of 1.86 billion seeds. It currently holds approximately 10% of the world's flora and is working toward a 2020 target of holding 20% of the world's flora.

The largest collections of noncrop wild species are maintained by the world's botanic garden networks. There are approximately 2500 botanic gardens and arboreta in 148 countries worldwide and they maintain more than 6 million living plant accessions, with representatives of more than 80,000 species, almost one-third of the known vascular plant species of the world. In addition, the world's botanic gardens and botanical institutions hold a total of 142 million herbarium specimens (http://www.bgci.org/resources/1528/). Botanic gardens are increasingly being recognized as vitally important resources of horticultural skills and conservation biology research, which support the imperatives of *in situ* management of wild populations, and as readily accessible venues for public education. Botanic gardens have a fundamental role in public education as accessible venues for introducing the public, via compelling exhibits, to the imperatives of plant conservation and sustainable living (Maunder, 2008).

Plant conservation will be working with new tools, for instance the Reducing Emissions from Deforestation and Forest Degradation (REDD)+ mechanisms (REDD "plus" conservation), allowing the sustainable management of forests and enhancement of forest carbon stocks, presents a key opportunity to protect forests while combating climate change and improving human well-being in developing nations. REDD+ provides a mechanism of monetary incentives for developing countries to halt or reduce forest loss, or restore forest. Although the implementation of REDD+ is not fully tested and remains controversial, some early initiatives show promise. Many of the pioneer projects encompass 1000 ha of forest and often support populations of iconic species (Harvey *et al.*, 2010, http://www.conservation.org/documents/redd/CI_REDD_Lessons_Learned.PDF), for instance the 13 million ha Xingu Basin project in Brazil. In Kenya, the Wildlife Works flagship "Kasigau Corridor REDD project" has secured over 500,000 acres of forest and protects the wildlife migration corridor between Tsavo East and Tsavo West National Parks. These early case studies indicate great promise but the long-term application of REDD+ remains to be tested.

Plant conservationists, faced with limited resources and a huge challenge of conserving many plant species, have started to identify the areas of greatest botanical diversity as a means of focusing resources. The IPA initiative of Plant Life has been remarkably successful (Plantlife, 2010, http://www.plantlife.org.uk/wild_plants/important_plant_areas/), identifying internationally significant sites for wild plants and threatened habitats. To date, teams have identified IPA in 66 countries, examples include 200 preliminary IPAs that were identified for North Africa and the Middle East and 1771 IPAs identified in 16 European countries.

Accelerating trends for both habitat loss and degradation have promoted discussion on the most effective responses to retaining both the taxonomic and the ecological components of biodiversity. Plant conservation strategies need to utilize a variety of complementary techniques, "integrated strategies" *sensu* Falk (1990). Recovery planning for threatened plant species is a relatively recent development, with the first US plant recovery plan initiated in 1979. Species recovery work for threatened plants is now widely adopted by national and regional plant conservation authorities.

The establishment of national plant conservation networks is playing an important role by promoting (1) national strategies and action plans for plant conservation in alignment with the GSPC and associated support and funding mechanisms, (2) assessing and monitoring national needs and priorities, (3) establishing an integrated approach to plant conservation, utilizing and promoting the available professional skills, (4) developing collaborative relationships and network with protected area networks, government agencies, parastatals, and NGOs, (5) initiating a national system for monitoring progress against GSPC targets and National Biodiversity Action Plans, and (6) institutional strengthening through locally focused professional training. National strategies using the GSPC framework have been produced for China, Ireland, Philippines, Seychelles, UK, and South Africa. Europe, through Planta Europa, has produced a regional plant conservation strategy. The delivery of these strategies is coordinated through national agencies and the national representatives for the Global Strategy for Plant Conservation. Examples of national plant conservation networks include the Center for Plant Conservation (CPC) in the USA, PlantLife in the UK, New Zealand Plant Network, Canadian Botanical Conservation Network, and the Australian Network for Plant Conservation. Some networks are regional for Planta Europa for Europe.

The CPC in the USA can be viewed as one model for a national network, based at the Missouri Botanical Garden, St. Louis, MO, USA, established in 1984 as a national network of collaborating botanic gardens with the clear focus on conserving threatened native flora. The CPC has promoted the utilization of both *in situ* and *ex situ* techniques. Fundamental to the effectiveness of CPC has been the availability of quality data on the distribution and status of botanical diversity; this has been heavily dependent on the activities of the Nature Conservancy's Heritage Programs. In addition, CPC has produced three keystone references for plant conservation, namely "Genetics and Conservation of Rare Plants" (Falk and Holsinger, 1991); "Restoring Diversity: Strategies for Reintroduction of Endangered Plants" (Falk *et al.*, 1996); and "Ex Situ Plant Conservation: Supporting Species Survival in

the Wild" (Guerrant *et al.*, 2004), and a new book on plant reintroductions (Maschinski and Haskins, in press). This has resulted in the development of working collaborations between protected area authorities, government agencies, and the adoption of population genetics as a working tool for botanic garden conservation activities. CPC works with 37 leading botanical institutions in the USA and through the National Collection of Endangered Plants holds over 700 threatened species and over 9,000,000 seeds.

A number of international networks have been established for wild plant conservation. The Species Survival Commission of the World Conservation Union (SSC/IUCN) has a number of international specialist groups for threatened plant taxa (e.g., orchids, palms, crop wild relatives, sea grasses, trees, and bryophytes), national or regional specialist groups (e.g., Korea, China, and Cuba), and regional red list authorities (e.g., East Africa).

As the human population grows and more land is converted to agriculture, the big challenge for plant conservationists will be retaining landscapes and habitats for wild plants. The future for plant diversity and its role in providing essential ecosystem services will depend to a large extent on negotiations over the use of land; these negotiations will take place in village assemblies, corporate board rooms, and government offices. How much wild land we retain and how much biodiversity we can maintain in working landscapes with ever-increasing human population densities will depend on our ability as plant conservationists to demonstrate and effectively communicate the value of plants and wild habitats.

Plant conservation stands at an interesting point in history; we are completing the botanical inventory of the planet, we are learning more about plants, their diversity, distribution, and status than we could have imagined 20 years ago, and we have an increasingly sophisticated, although still not complete, understanding of how plants support life and society. Yet this diversity is eroding at a rate unimaginable to previous generations – for instance we have witnessed the effectively "instant" loss of lowland SE Asian rainforests and the Brazilian cerrado to commercial agriculture.

Plant conservation is competing for resources and land in a world of increasing agricultural production (including expanding vegetable oil and biofuel production), the overharvesting of high-value plant products (e.g., rainforest timber or medicinals), rampant urbanization, and accelerating environmental change from invasive species and climate change. Plant conservation could easily become focused on salvaging the lost. Future investment in scientific and practical activity should focus on the retention of viable habitat areas and the management of socially and economically important plant resources. Although a proportion of the world's plant diversity can be saved as "threatened species," the vast majority of plant diversity will depend on conserving viable habitat areas. A key strategy will be retaining extensive areas of wild habitats in a matrix of human-dominated landscapes; these surviving habitats will not only allow the retention of natural ecological processes but should also allow for the potential for migration in response to climate change.

There will need to be a continued investment in new tools for plant conservation. Key new research areas include advancing the ability to restore habitats and ecosystem services; although extensive restoration research has been undertaken for temperate ecosytems, it is encouraging to see more work on tropical habitats. The management of habitat fragments and isolated plant populations will become increasingly important, specifically the techniques to maintain biodiversity in fragments and how to link isolated fragments in increasingly human-dominated landscapes. As plant diversity fades, so will traditional knowledge, as cultures become homogenized; hence, we are losing tools and approaches that will be valuable for the sustainable use of plant diversity.

Above all, the global investment in plant conservation needs to be scaled up; we have many of the tools we need but we need to dramatically expand the scope and impact of our work. Key to this is building both stronger institutions and professionals. Importantly, plant conservationists will need to play an influential role in the debates on land use, agricultural production, and habitat conversion, and importantly, effectively promoting plant conservation as a service for the public good. Although effective plant conservation will, in part be dependent on establishing effective policies, monitoring, and management tools, it will ultimately depend on society placing a real value on wild plants and landscapes.

Appendix

List of Courses

1. MSc Botanical Conservation, University of Plymouth, UK
2. Diploma in Plant Conservation Strategy, Royal Botanic Gardens, Kew
3. MSc Taxonomy and Biodiversity, Imperial College, UK
4. MSc and Diploma Biodiversity and Taxonomy of Plants, University of Edinburgh with Royal Botanic Gardens, Edinburgh, UK
5. MSc Plant Diversity, University of Reading, UK
6. MSc Ethnobotany, University of Kent, UK

See also: Comparing Extinction Rates: Past, Present, and Future. Deforestation and Land Clearing. Endangered Plants. Global Species Richness. Hotspots. Human Impact on Biodiversity, Overview. *In Situ*, *Ex Situ* Conservation. Introduced Species, Impacts and Distribution of. Mangrove Ecosystems. Plant Biodiversity, Overview. Restoration of Biodiversity, Overview. Role of Botanic Gardens. Valuing Ecosystem Services

References

Asmussen-Lange C, Maunder M, and Fay M (2011) Conservation genetics of the critically endangered Round Island bottle palm *Hyophorbe lagenicaulis*. *Botanical Journal of the Linnean Society* 167: 301–310.

Bai ZG, Jong de R, and van Lynden GWJ (2010) An update of GLADA – Global assessment of land degradation and improvement, pp. 58. *ISRIC report 2010/08*, Wageningen: ISRIC – World Soil Information.

Barthlott W, Hostert A, Kier G, *et al.* (2007) Geographic patterns of vascular plant diversity at continental to global scales. *Erkunde* 61(4): 305–316.

Bowles ML and Whelan CJ (1994) Restoration of Endangered Species: Conceptual Issues. *Planning and Implementation*. Cambridge, UK: Cambridge University Press.

Boyce DG, Lewis MR, and Worm B (2010) Global phytoplankton decline over the past century. *Nature* 466: 591–596.

Brooks TM, Mittermeier RA, Da Fonseca GAB, *et al.* (2006) Global biodiversity conservation priorities. *Science* 313: 58–61.

Brooks TM, Mittermeier RA, Mittermeier CG, *et al.* (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology* 16(1): 909–923.

Brummitt N, Bachman S, and Moat J (2008) Applications of the IUCN Red List: Towards a global barometer for plant diversity. *Endangered Species Research* 6: 127–135.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: Indicators of recent declines. *Science* 328: 1164–1168.

Caujape-Castells J, Tye A, Crawford DJ, *et al.* (2009) Conservation of oceanic island floras: Present and future global scenarios. *Perspectives in Plant Ecology, Evolution and Systematics* 12: 107–129.

Davis SD, Heywood VH, and Hamilton AC (eds.) (1994) *Centres of Plant Diversity: A Strategy for their Conservation*, Vol. 1. Europe, Africa, South West Asia and the Middle East: IUCN/WWF.

Davis SD, Heywood VH, and Hamilton AC (eds.) (1996) *Centres of Plant Diversity: A Strategy for their Conservation*, Vol. 2. Asia, Australia and the Pacific: IUCN/WWF.

Davis SD, Heywood VH, and Hamilton AC (eds.) (1997) *Centres of Plant Diversity: A Strategy for their Conservation*, Vol. 3. The Americas: IUCN/WWF.

Duke NC, Meynecke JO, Dittmann S, *et al.* (2007) A world without mangroves? *Science* 317: 41–42.

Falk DA (1990) Integrated strategies for conserving plant genetic diversity. *Annals of the Missouri Botanical Garden* 77: 38–47.

Falk DA and Holsinger KE (eds.) (1991) *Genetics and Conservation of Rare Plants*. New York: Oxford University Press.

Falk D, Millar CI, and Olwell M (eds.) (1996) *Restoring Diversity: Strategies for Reintroduction of Endangered Plants*. Washington, DC: Island Press.

FAO (2010) Global Forest Resources Assessment 2010. FAO forestry paper # 163. Rome: FAO. http://www.fao.org/docrep/013/i1757e/i1757e.pdf

Feely KJ and Silman MR (2009) Extinction risks of Amazonian plant species. *Proceedings of the Natational Academy of Sciences* 106(30): 12382–12387.

Francisco-Ortega J, Santiago-Valentín E, Acevedo-Rodríguez P, *et al.* (2007) Seed plant genera endemic to the Caribbean Island biodiversity hotspot: A review and a molecular phylogenetic perspective. *Botanical Review* 73: 183–234.

Francisco-Ortega J, Ventosa I, Onviedo R, *et al.* (2008) Caribbean Island Asteraceae: Systematics, molecules, and conservation on a biodiversity hotspot. *Botanical Review* 74: 112–131.

Govaerts R (2003) How many species of seed plant are there? A response. *Taxon* 52: 583–584.

Grove R (1996) *Green Imperialism: Colonial Expansion, Tropical Island Edens and the Origins of Environmentalism, 1600–1860*. Cambridge: Cambridge University Press.

Guerrant EO, Havens K, and Maunder M (eds.) (2004) *Ex Situ Plant Conservation: Supporting Species Survival in the Wild*. Covelo, CA: Center for Plant Conservation with Society for Ecological Restoration. Island Press.

Haberl H, Erb KH, Krausmann F, *et al.* (2007) Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *Proceedings of the National Academy of Sciences* 104: 12942–12947.

Hallingback T and Tan BC (2010) Past and present activities and future strategy of bryophyte conservation. *Phytotaxa* 9: 266–274.

Hannah L, Lohse D, Hutchinson C, Carr JL, and Lankerani A (1994) A preliminary inventory of human disturbance of world ecosystems. *Ambio* 23: 246–250.

Hannah L, Radhika D, Lowry II P, *et al.* (2008) Climate change adaptation for conservation in Madagascar. *Biology Letters* 4: 590–594.

Harper G, Steininger M, Tucker C, Juhn D, and Hawkins F (2007) Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation* 34: 325–333.

Harvey CA, Zerbock O, Papageorgiou S, and Parra A (2010) *What is Needed to Make REDD+ Work on the Ground? Lessons Learned from Pilot Forest Carbon Initiatives*, pp. 121. Arlington, VA, USA: Conservation International.

Havens K, Vitt P, Maunder M, Guerrant EO, and Dixon K (2006) Ex situ *plant conservation and beyond. Bioscience* 56(6): 525–531.

Isbell F, Calcagno V, Hector A, *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature* 477: 199–202.

Maclean IMD and Wilson RJ (2011) Recent ecological responses to climate change support predictions of high extinction risk. *Proceedings of the National Academy of Sciences* 108(30): 12337–12342.

Joppa LN, Roberts DL, and Pimm SL (2010) How many species of flowering plant are there? *Proceedings of the Royal Society B: Biological Sciences* 2011(278): 554–559.

Lucas GLI and Synge H (1978) *The IUCN Plant Red Data Book*. Morges, Switzerland: IUCN.

Kumar P (ed.) (2011) *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations (TEEB: The Economics of Ecosystems and Biodiversity)*. UNEP

Maschinski J and Haskins KE (eds) (in press) *Plant Reintroduction in a Changing Climate: Promises and Perils*. Center for Plant Conservation with Society for Ecological Restoration. Island Press.

Maunder M (2008) Beyond the greenhouse. *Nature* 455: 596–597.

Maunder M, Culham A, Bordeu A, Allanguillame J, and Wilkinson M (1999) Genetic diversity and pedigree for *Sophora toromiro* (Leguminosae): A tree extinct in the wild. *Molecular Ecology* 8: 725–738.

Maunder M, Leiva A, Santiago-Valentin E, *et al.* (2008) Plant conservation in the Caribbean biodiversity hotspot. *Botanical Review* 74(1): 197–207.

Maunder M, Page W, Mauremootoo J, *et al.* (2002) The decline and conservation management of the highly threatened endemic palms of the Mascarene Islands. *Oryx* 36: 56–65.

Maunder M, Upson T, Spooner B, and Kendle T (1995) Saint Helena: Sustainable development and conservation of a highly degraded island ecosystem. In: Vitousek PM, Loope LL, and Adsersen H (eds.) *Islands: Biological Diversity and Ecosystem Function*, pp. 205–217. Berlin: Springer and Verlag.

Maxted N and Kell SP (2009) *Establishment of a Global Network for the In Situ Conservation of Crop Wild Relatives: Status and Needs*. Rome, Italy: FAO Commission on Genetic Resources for Food and Agriculture.

McClean C, Lovett JC, Kuper W, *et al.* (2005) African plant diversity and climate change. *Annals of the Missouri Botanical Garden* 92: 139–152.

Mittermeier RA, Myers N, Thomsen JB, da Fonseca GAB, and Oilivieri S (1998) Biodiversity hotspots and major tropical wilderness areas: Approaches to setting conservation priorities. *Conservation Biology* 12(3): 516–519.

Mittermeier RA, Robles Gil P, Hoffmann M, *et al.* (2004) *Hotspots Revisited*. Mexico: CEMEX.

Myers N (1988) Threatened biotas: Hotspots in tropical forests. *The Environmentalist* 8: 1–20.

Myers N (1990) The biodiversity challenge: Expanded hot spot analysis. *The Environmentalist* 10: 243–255.

Plantlife (2010) *Important Plant Areas Around the World: Target 5 of the CBD Global Strategy for Plant Conservation*. Salisbury, UK: Plantlife International. www.plantlife.org.uk/

Prance GT and Elias TS (eds.) (1977) *Extinction is Forever*. New York, USA: New York Botanical Garden.

Reynolds JF, Smith DM, Lambin EF, *et al.* (2007) Global desertification: Building a science for dryland development. *Science* 316(5826): 847–851.

Richardson K, *et al.* (2009) Climate change: Global risks, challenges and decisions. *Synthesis Report from the International Scientific Congress*. Copenhagen, Denmark.

Sax DF and Gaines SD (2008) Species invasions and extinction: The future of native biodiversity on islands. *Proceedings of the National Academy of Sciences* 105(1): 11490–11497.

Schwartz MW, Thorne JH, and Viers JH (2006) Biotic homogenization of the California flora in urban and urbanizing regions. *Biological Conservation* 16: 1338–1350.

Short FT, Polidoro B, Livingstone SR, Carpenter KE, *et al.* (2011) Extinction risk assessment of the world's seagrass species. *Biological Conservation* 144(7): 1961–1971.

Simmons JB, Beyer RI, Brandham PE, GLI Lucas, and Parry V (eds.) (1976) *Conservation of Threatened Plants*. London: Plenum Press.

Sommer JH, Kreft H, Kier G, Jetz W, Mutke J, and Barthlott W (2010) Projected impacts of climate change on regional capacities for global plant species richness. *Proceedings of the Royal Society B: Biological Sciences* 277: 2271–2280.

Spottiswoode CN, Wondafrash M, Gabremichael MN, *et al.* (2009) Rangeland degradation is poised to cause Africa's first recorded avian extinction. *Animal Conservation* 2009: 1–9.

Stuart SN, Wilson EO, McNeely JA, Mittermeier RA, and Rodriguez JP (2010) The Barometer of Life. *Science* 328: 177.

Synge H and Townsend H (eds.) (1979) *Survival or Extinction: The Practical Role of Botanic Gardens in the Conservation of Rare and Threatened Plants*. Royal Botanic Gardens, Kew: The Bentham-Moxon Trust.

Turner BL, Clark WC, Kates RW, Richards JF, Mathews JT, and Meyer WB (eds.) (1990) *The Earth as Transformed by Human Action*. Cambridge, UK: Cambridge University Press.

Underwood EC, Viers JH, Klausmeyer KR, Cox RL, and Shaw MR (2009) Threats and biodiversity in the Mediterranean biome. *Diversity and Distribution* 15: 188–197.

Walter KS and Gillett HJ (1998) *1997 IUCN Red List of Threatened Plants*. Gland, Switzerland: World Conservation Monitoring Centre and IUCN-World Conservation Union.

Weeks WW (1996) *Beyond the Ark: Tools for an Ecosystem Approach to Conservation*. Washington, DC, USA: Island Press.

Winter M, Schweiger O, Klotz S, *et al*. (2010) Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proceedings of the National Academy of Sciences* 106(51): 21721–21725.

Wyse Jackson PS and Kennedy K (2009) The Global Strategy for Plant Conservation*:* A challenge and opportunity for the international community. *Trends in Plant Science* 14: 578–580.

Plant Invasions

David M Richardson, Stellenbosch University, Matieland, South Africa
Petr Pyšek, Academy of Sciences and Charles University, Prague, Czech Republic

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by David. M. Richardson, volume 4, pp 677–688, © 2001, Elsevier Inc.

Glossary

Alien plants Plant taxa in a given area whose presence there is because of intentional or accidental introduction as a result of human activities. Synonyms include exotic plants, nonindigenous plants, and nonnative plants.
Enemy Release Hypothesis The idea that alien species have a better chance of establishing and becoming dominant when released from the negative effects of natural enemies that, in their native range, lead to high mortality rates and reduced productivity.
Evolution of Increased Competitive Ability Hypothesis (EICA) A notion that posits that plants introduced to an environment that lacks their usual herbivores or disease agents will experience selection favoring individuals that allocate less energy to defense and more to growth and reproduction.
Fluctuating Resources Theory of Invasibility A theory, formulated for plants, that predicts that pronounced fluctuations in resource availability enhances community invasibility if coinciding with the availability of sufficient propagules to initiate an invasion.
Homogenization (Biotic) The addition to, and often the partial if not extensive replacement of, local biotas by alien species, which can result in decreased compositional turnover (β-diversity) of species between distant areas, in terms of both taxonomic and phylogenetic similarity.
Invasibility The properties of a community, habitat, or ecosystem that determine its inherent vulnerability to invasion.
Invasive alien plants Naturalized plants that produce reproductive offspring, often in very large numbers, at considerable distances from parent plants, and thus have the potential to spread over a considerable area.
Invasiveness The features of an alien organism, such as their life-history traits and modes of reproduction that define their capacity to invade, that is, to overcome various barriers to invasion.
Naturalized plants Alien plants that reproduce consistently and sustain populations over many life cycles without the input of resources or direct intervention by humans; they often recruit offspring freely, but mainly near adult plants, and do not necessarily invade natural ecosystems (cf. invasive alien plants).
Propagule pressure A concept that encompasses variation in the quantity, quality, composition, and rate of supply of alien organisms resulting from the transport conditions and pathways between source and recipient regions.
Range expansion The process whereby a species spreads into new areas (usually new regions, rather than local-scale movements) due to natural dispersal, though the expansion may be assisted by human-mediated changes to the environment. It differs from invasion in that human-mediated extra-range dispersal (i.e. across a biogeographical barrier) is not implicated; the concept can be applied to both native and alien species.

Technological innovations, driven by rapid increases in the global human population and requirements for diverse products and services from nonnative plants, have facilitated the large-scale (in terms of numbers of individuals, taxa, and foci of introduction) movement of plants to regions far beyond their natural dispersal range. This, together with widespread changes to disturbance regimes and the development of novel anthropogenic plant communities, has led to many plants becoming invasive in foreign environments. Plants carried to new habitats far from their natural ranges are very often introduced, intentionally or accidentally, as seeds; they arrive without key natural enemies that limit their performance in their natural range. For many reasons, most introduced species do not spread from planting sites, and cause no damage in the receiving environment. Many alien plants provide us with food, fuel, and timber, beautify our cities, and serve many other useful purposes without invading. However, some plant species that are able to spread in new environments have profound influences in invaded ecosystems, and some are important contributors to habitat transformation and biodiversity loss in many parts of the world. This article focuses primarily on the ecology of those species that have invaded natural and seminatural systems. It explores how plants have been moved around the world, the factors that determine invasiveness and invasibility, how invasive plants affect invaded ecosystems, and the range of management interventions for dealing with invasive plants.

Concepts and Terminology

Terms such as *colonization* (or *colonizer*) and *invasion* (or *invader*) are sometimes applied with reference to species that reclaim their distributions rapidly following disturbance, or that quickly occupy new sites adjoining their original range following changes to factors that previously limited their

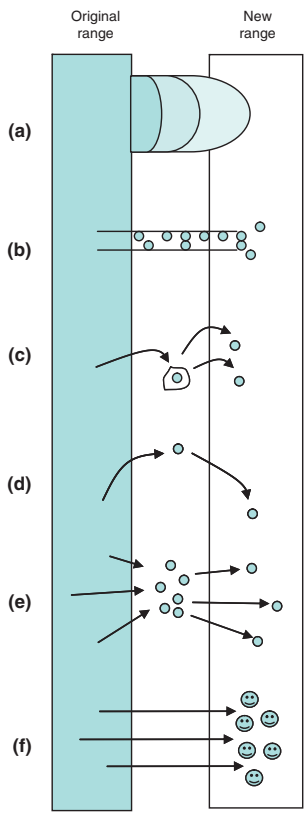
 http://dx.doi.org/10.1016/B978-0-12-384719-5.00319-1

ranges; the former is used in an ecological context, whereas the latter conveys biogeographical context. Within any assemblage, species vary in their ability to persist under conditions of environmental change, and with respect to their ability to invade new sites. Most species occupy a reasonably fixed position on a continuum from colonizers or pioneers to late-seral species. The former typically undergo large fluctuations in population size whereas the latter are specialized for more stable environments where rapid population growth is unimportant. Such strategies are associated with distinctive suites of life-history traits.

Episodes of range expansion and contraction mediated by natural disturbances, interactions with co-occurring organisms, and restricted by natural barriers have driven the diversification of floras and the delimitation of the world's phytogeographic zones. Such events have also molded the biology of species in many important ways. Consequently, some plant species can respond rapidly to disturbances, and increase in numbers from populations comprising small numbers of individuals. This capacity is influenced by a suite of interacting traits such as aspects of the reproductive biology of the species. Importantly, human activities in the last few centuries have radically altered the opportunities and challenges for plant species with regard to dispersal, colonization, and persistence.

Humans have influenced the distribution of plants for millennia by altering opportunities for persistence, regeneration, and spread. Such influences were originally mostly at local and regional scales. However, growing human populations and technological advances mean that humans now exert an influence on plant population dynamics not just locally, but also at a global scale. An important reason for this is that human actions have moved thousands of plant species to areas considerably beyond their normal dispersal ranges. Such changes have not only brought about quantitative changes to dispersal pathways but have also forged entirely new invasion pathways whereby plants are moved to new environments in new ways (**Figure 1**).

The study of weeds (plants, not necessarily alien, that grow in sites where they are not wanted and that have detectable economic or environmental impacts) has for centuries been directed at solving problems (how can weeds be killed or how can their effects on crops be reduced?) Colonization dynamics have been systematically explored by plant ecologists for over a century. The detailed study of the dynamics of plants introduced to regions far from their natural ranges is a more recent field of investigation (although Charles Darwin and other eighteenth century naturalists made important observations in this regard). Milestones in the study of such biogeographical events, termed invasions, have been Charles Elton's (1958) book *The Ecology of Invasion by Animals and Plants*, the international program on *The Ecology of Biological Invasions* (1982–1988) under the auspices of SCOPE (Scientific Committee on the Problems of the Environment) (Drake *et al.*, 1989), and, for Europe, the Delivering



	Description	Examples	Frequency of dispersal events	Number of propagules per event	From	To	Enemy release	
Leading-edge dispersal	Gradual changes in range initiated by individuals colonising new areas from the edge of the range (diffusive spread)	Recolonisation after glaciation, response to human-mediated habitat change	Continuous	Depends on the local abundance at the range edge	One region	One region	No	Natural dispersal and/or human-mediated dispersal
Corridor	A physical connection of suitable habitat forms linking a portion of the range to another area (geodispersal or mass coherent dispersal)	The Great American Exchange, the Panama, and Suez Canals	Continuous	Many	One region	One region	Unlikely	
Jump dispersal	Long-distance dispersal over sometimes substantial distances, but there is still a connection between the new and original ranges	A few propagules disperse over mountain ranges or water-bodies	Infrequent	Few, but not necessarily from range edge	Few regions	Few regions	Partial or temporary	
Extreme long-distance dispersal	Propagules move far beyond the dispersal range seen over ecological time-scales (sweep-stakes dispersal)	Oceanic rafting, aerial plankton, dispersal by birds	Once or very infrequent	Very few	Only one or a few regions	Only one or a few regions	Very likely	
Mass transit	A dispersal route is established such that many individuals can move from different parts of the range to many new sites	Stowaways, for example rats on ships and hull fouling organisms	Rare to often	Several	Few regions to the whole range	Many regions	Likely	Hum-mediated dispersal
Cultivation	Propagules, sometimes mature individuals, are actively moved (directed dispersal) and receive resources to establish and persist	Ornamentals, Agricultural species, their contaminants and pests	Can be one event or continuous	Few to many	Few regions to whole range	Many regions	No to yes depending on quarantine	

Figure 1 Extra-range dispersal categorized into six types of dispersal pathways. Reproduced from Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, and Richardson DM (2009) Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology & Evolution* 24: 136–144, with permission from Elsevier. The cultivation pathway is the dominant pathway for most modern plant introductions.

Alien Invasive Species Inventories for Europe initiative (DAISIE, 2005–2008; http://www.europe-aliens.org). The primary international forum for research on plant invasions is the Ecology and Management of Alien Plant Invasions (EMAPI) conference series (Richardson *et al.*, 2010).

Humans have influenced the geographic ranges of plants for millennia, by shuffling their distributions and altering their abundances, both within and beyond their natural ranges. The extent to which such movements prior to the age of European colonialism shaped current distributions in Europe and Asia is often overlooked. Whether these range changes represent invasions is debatable since many changes were within or adjoining the natural ranges of such species. This is one reason for the blurred distinction between the categories colonizers and invaders in the older ecological literature. Examples of human-induced changes in the distribution of pines (*Pinus* spp.; Pinaceae) serve to illustrate the different categories of range expansion. The distributions of species such as *Pinus brutia*, *P. halepensis*, *P. pinaster*, and *P. pinea* have changed dramatically due to human activities over many centuries. *P. brutia* subsp. *eldarica* currently occurs in scattered locations from Azerbaijan to Pakistan. Genetic studies have shown that it occurs naturally only on a single mountain in Azerbaijan; the rest of its current range is due to humans who spread and managed this pine as a semi-domesticated landrace. Most of its known locations follow ancient trade routes along the path taken by Alexander the Great (356–323 BC). In North America and other parts of the natural range of *Pinus*, many native species have invaded meadows, abandoned lands, grazed grasslands, and other habitats where they did not occur before humans arrived. This is the case with many invasive species in their native ranges. Even within native ranges, species colonize human-modified habitats in which they did not evolve. Giant hogweed (*Heracleum mantegazzianum*) is an informative example. This species occurs naturally in forest fringes and tall-herb communities of a high altitudinal belt of the Western Caucasus where it grows at low densities among highly competitive stout herbs. It never occurs at levels of abundance that exist in the invaded range. However, in human-disturbed habitats in the lowlands surrounding these mountains hogweed creates large dominant stands with high cover and abundance, similar to those in many European countries where it invades. Thus, both patterns can be found in the native geographical range of the species, but they are strongly habitat-dependent (Pyšek *et al.*, 2007).

In other cases, range expansions have been augmented by human-mediated dispersal of seeds beyond their normal dispersal ranges. Most pines are equipped to be colonizers; when moved by humans to areas outside their natural range, many of these species become invaders, as shown, for example, by the invasive behavior of *Pinus radiata* and other pines in Australia, New Zealand, and South Africa.

Given the above-mentioned and other difficulties with classifying species that are alien to a given region, there has been much debate about terminology in the scientific literature. Only in the last decade has some standardization been achieved (*see* Glossary for a definition of the terms that are now most often used), although some concepts are still being actively debated.

The History of Alien Plant Invasions

Humans moved plant species beyond their natural ranges long before the dawn of the Age of Discovery and European colonialism. The frequency and pervasiveness of such early translocations were generally too limited to have resulted in widespread, damaging invasions. The voyages of exploration and the gradual settlement of the East and the New World by European powers (Britain, France, Germany, the Netherlands, Portugal, and Spain) heralded the start of a huge wave of translocations of plants and animals. Early colonizers transported mainly organisms from their homelands and actively assisted them in adapting (acclimatizing) to new regions. Many early botanical gardens established in colonies were essentially centers of introduction and acclimatization. This phase of ecological imperialism shaped the alien floras of many parts of the world that currently face the most severe problems with invasive alien plants. To cite but one instructive example: Plant introductions from Europe and Asia to South Africa between 1650 and 1806 reflected the need of Dutch colonists to cultivate a wide variety of agricultural and horticultural crops, mainly from their homeland in Europe and from the Dutch possessions in the East. Fifty or more crop plants were introduced within the first few years of European settlement, including many present-day weeds of agriculture. Subsequent waves of human immigrants, especially between 1800 and 1870, led to a surge in plant and animal introductions. Most of the plant species that now cause the greatest problems as invaders in South Africa arrived between 1825 and 1860, including hundreds of species of acacias, eucalypts and other plants from Australia. Other regions, such as North America or Australia, experienced a similar influx of alien plant invaders for the same reasons.

The magnitude and composition of the influx of alien plants to a given region during the colonial era was determined by many factors, including the geographical location and climate of the area, it's cultural and economic links with colonial powers and other colonies, and special features of the region that influenced decisions on what types of plants to introduce. For example, the four regions of the world with a Mediterranean-type climate outside the Mediterranean Basin (California, Central Chile, southern and southwestern Australia, and the southwestern tip of South Africa) all received the bulk of their alien floras from the Mediterranean Basin. Smaller components came from other regions with which they had colonial ties, and even fewer from other regions. Current problems with plant invasions in these regions still reflect the dynamics of biological exchange during the colonial era, and the species composition of alien biota can still be, to a large extent, related to the time of discovery and the provenance of early human colonizers. More recent introductions from a broader range of regions are, however, becoming more prominent in alien floras; this is starting to blur the colonial stamp on invasions.

The rapid globalization of European trade routes starting in the eighteenth century led to a dramatic increase in the magnitude and diversity of the transcontinental movement in plants. Trends affecting plant invasions in the nineteenth and twentieth centuries are complex and region specific. There are important differences in patterns between Northern and

Southern Hemispheres, between eastern and western parts of the Northern Hemisphere, and between what are now known as industrialized and developing countries. The numbers of alien plant species that are becoming established are still increasing in most regions. Although new arrivals originate from a wider source area than the early introductions, they tend to arrive *via* more direct routes than in the past.

Among the fundamental factors that have affected invasions globally, especially in the twentieth century, have been (1) the improvement and expansion of transportation systems, both at an intercontinental scale, with the emergence of aircraft transportation and massive growth in shipping, and within regions (roads, railways, internal navigation canals, interbasin-transfer conveyance schemes, etc.); (2) massive habitat transformations and changes to prevailing disturbance regimes, considerably increasing the extent of ruderal habitats that are favorable for alien species and serve as foci for their further spread; (3) the introduction of large numbers of livestock to temperate grasslands of the New World dominated by caespitose grasses that lack the ability to withstand grazing pressure and were replaced by alien annual grasses preadapted to high grazing pressure; (4) widespread afforestation using alien tree species (the area of tree plantations – comprising mainly alien species of *Pinus*, *Eucalyptus*, *Tectona*, and *Gmelina* – doubled between 1980 and 1995 to cover 180 million ha globally); and (5) the increasing demand for alien plants for use in horticulture (North American seed and nursery catalogs offer more than 59,000 plant taxa for sale to national and international markets, including many known invasive species), and for other human uses. Developments in the twentieth century (especially the last quarter) that are particularly pertinent to plant invasions have included the increasing attention paid to scouring floras of previously underexplored areas for new useful plants; and the growth of agencies to facilitate the rapid transfer of germplasm of many multipurpose tree species for agroforestry and other purposes. Recent interest in the use of alien plants for biofuel production is likely to exacerbate problems. Most of the world's worst invasive plants reached the stage where they cause major damage only in the past four decades. It is now well established that plant invasions are a significant component of human-caused global change.

The Current Extent of Alien Plant Invasions

Plants are the best studied of all invasive organisms. Our understanding of invasion mechanisms and processes is based to a large extent on research on plant invasions. An analysis of published research on biological invasions showed that plants are the taxonomic group that is studied most frequently; almost half of all published case studies on invading species worldwide focus on plants. The two most frequently studied species are animals (zebra mussel *Dreissena polymorpha*, and Argentine ant, *Linepithema humile*), but among the 10 invasive species receiving most research attention, six are vascular plants (*Centaurea maculosa*, *Lythrum salicaria*, *Phragmites australis*, *Tamarix ramosissima*, *Spartina alterniflora*, and *Alliaria petiolata*) (Pyšek *et al.*, 2008).

Invasive alien floras include representatives of all main growth forms of higher plants (annuals, perennials, aquatic plants, epiphytes, geophytes, grasses, herbs, shrubs, trees, etc.). Alien plant invasions have affected aquatic ecosystems (estuarine, freshwater, marine, and wetlands) and most types of terrestrial ecosystems throughout the world, including arid systems, coastal systems, forests (boreal, temperate, and tropical), grasslands, savannas and woodlands (including mangroves), and shrublands. There are very few habitats globally that are not seriously affected by plant invasions; basically only habitats under extreme ecological conditions where plants are exposed to a strong stress, such as alpine habitats or acidic mires, etc., are rarely invaded (Chytrý *et al.*, 2008).

Though alien plant invasions are now a global phenomenon, they have been much better studied in some parts of the world than others. This is partly because of the history of introductions and dissemination within regions, and partly because of the global distribution of ecologists and the recent history of international scientific programs. At the end of the twentieth century most notable contributions to the literature on plant invasion ecology came from Australia, the Mediterranean Basin, New Zealand, the Pacific islands (especially the Galapagos and Hawaiian archipelagos), South Africa, the United Kingdom, and the mainland United States (especially California and Florida). Research from other regions, notably Europe, is now contributing key insights into plant invasion ecology (Pyšek *et al.*, 2008).

Some prominent examples of alien plant invasions worldwide (including those that have had severe impacts and those that have been particularly well studied) are:

1. The rapid spread of alien trees and shrubs of the genera *Acacia*, *Hakea*, and *Pinus* over large areas in fynbos shrublands in South Africa (**Figure 2**). These invasions threaten hundreds of native plant species with extinction, change fire and nutrient-cycling regimes, and considerably reduce streamflow from watersheds.
2. The rapid invasion of mesic and wet montane forests on the Society and Hawaiian Islands by the alien tree *Miconia calvescens* (Melastomataceae), which transforms the forests, threatening many species with extinction.
3. The invasion of caespitose grasslands in North America's Great Basin by the European annual grass *Bromus tectorum* (cheatgrass; Poaceae) which is highly flammable and has increased the fire frequency from 60–110 years on average to 3–5 years. Cheatgrass invaded about 200,000 km^2 between 1890 and 1930, representing one of the fastest rates of aerial spread reported in the plant invasion literature (Mack, 1989).
4. The invasion of Europe by *Heracleum mantegazzianum* (giant hogweed). This species, introduced as a garden ornamental to the UK in 1817, escaped from cultivation very quickly, and has spread in at least 19 countries. It forms large dense stands in which few other species can survive (Pyšek *et al.*, 2007).
5. Widespread invasions of African grasses in the Neotropics and Australia, introduced to improve pasture yields. Prominent examples are *Andropogon gayanus* (Gamba grass) in tropical Australia, *Hyparrhenia hirta* (Coolatai grass) in South America, and *Pennisetum ciliare* (buffel grass) in the USA.

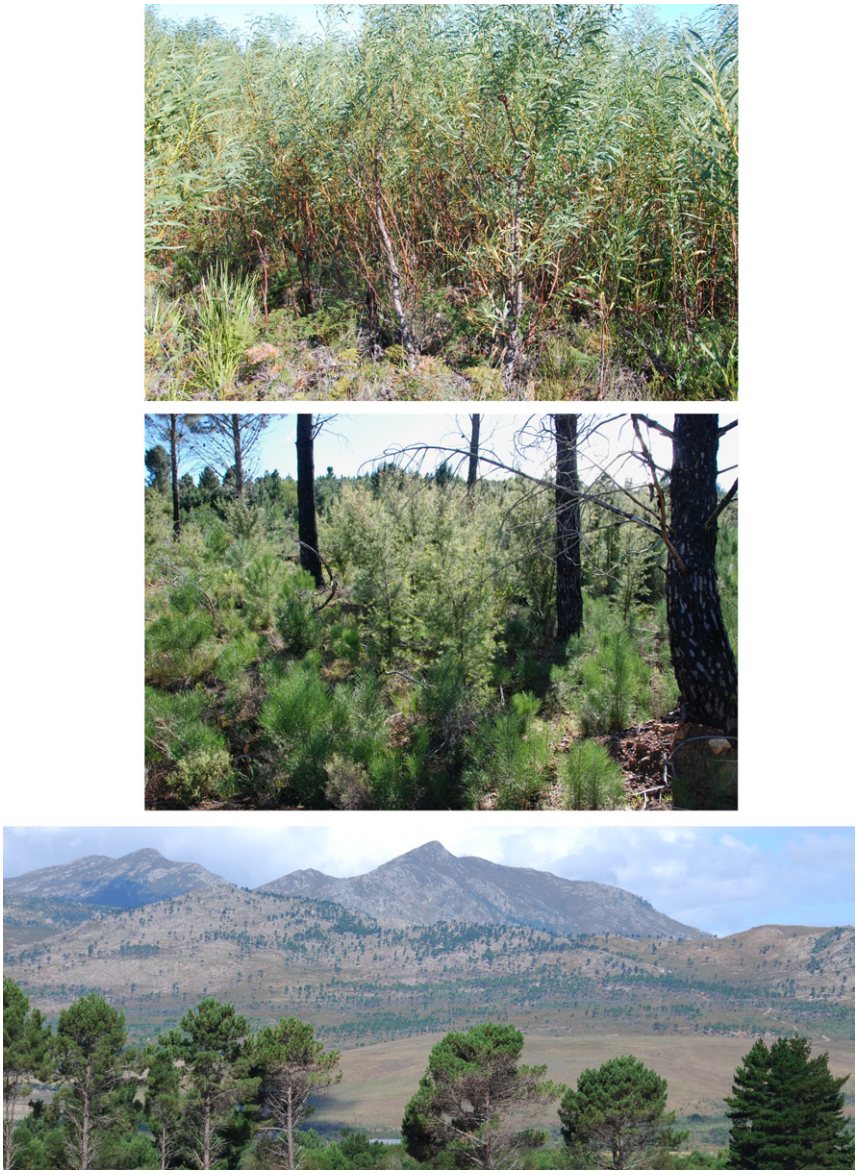


Figure 2 One of the most dramatic ecosystem transformations caused by invasive alien plants anywhere is that brought about by alien trees in South African fynbos vegetation where trees, especially those in the genera *Acacia* (*Acacia saligna*; top), *Hakea* (*Hakea sericea*; middle), and *Pinus* (*Pinus pinaster*; bottom) have changed species-rich shrublands into species-poor forests over tens of thousands of hectares. Photos, permission from David. M. Richardson.

6. The rapid spread of *Mimosa pigra* (mimosa; Fabaceae), a South American tree, to form dense thickets over about a million hectares of wetlands in Australia's Northern Territory. These thickets transform sedgelands into monotonous tall shrublands, and form an impenetrable understory in swamp forests, with profound consequences for biodiversity and for traditional Aboriginal food gathering.
7. The explosive spread of *Eichhornia crassipes* (water hyacinth; Pontederiaceae) in river and lagoons of West Africa and the Great Lakes of East Africa, especially in the past two decades, causing major disruption of fishing and navigation.

Which alien plant species are the most widespread or cause the most damage? There are numerous regional or national lists of invasive alien plants, some of which provide data on areas invaded or rankings of the severity of invasions. However, since there are no generally accepted ways of quantifying when an area is invaded or when a species is invasive (as opposed to naturalized), attempts at compiling global lists of major plant invaders have generally been less than satisfactory. A global review was recently conducted of invasive trees and shrubs around the world. This study listed 622 species (357 trees, 265 shrubs in 29 plant orders, 78 families, and 286 genera) that are currently clearly invasive. Regions with the largest number of woody invasive alien species are: Australia (183); southern Africa (170); North America (163); Pacific Islands (147); and New Zealand (107) (Richardson and Rejmánek, 2011). For plants, the DAISIE initiative recently collated lists of plant species alien to Europe (DAISIE, 2009).

At least 5789 plant species are currently considered alien to at least one region in Europe, of which 2843 arrived from other continents. Considering only naturalized species, there are 3749 species of which 1780 are native to other continents (Lambdon *et al.*, 2008).

Invasion Processes

Plant invasions involve the following fundamental phases: introduction to the region by humans (intentionally or accidentally), establishment, population growth (sometimes accompanied by genetic adjustment), spread to new areas within the region (and often also outside the region *via* further dispersal by humans); interaction with the local biota and disturbance regime, and displacement of native elements. There are many ways of conceptualizing the various processes involved in invasion and interactions with biotic and abiotic features of the new environment. One may depict the various potentially limiting factors as a series of barriers (Richardson *et al.*, 2000). The simplest representation of such a model shows (1) *a geographic barrier*, which must be overcome by dispersal; (2) a *habitat barrier*, which requires preadaptation or genetic adjustment to the conditions of the new environment; and (3) a *biotic barrier*, which integrates the forces of predation, herbivory, competition, and interference that must be overcome in the new habitat, or the new mutualistic relationships that must develop. Additional complexity can be added by, for example, splitting the geographic barrier into two components (to account for factors limiting introduction to the region and dispersal within the region, respectively), by adding a *reproductive barrier* (to account specifically for factors that potentially limit seed set), or by splitting the biotic barrier into components (e.g., to isolate the role of mutualisms). **Figure 3** shows a refinement of the basic barrier for plants model to include insights from other taxa.

Stages of Invasion

Dispersal to a New Area

Important reasons for the intentional widespread translocation of plants include agriculture, forestry and agroforestry, botanical gardens, horticulture (including the commercial trade in seeds, bulbs, and cuttings, and urban gardeners using seed exchanges), and soil stabilization. Many plants have been moved around the world inadvertently, notably in ship ballast, with military transport, and as contaminants in fertilizers, hay and straw, grains, wool, and cotton. For the 622 woody invasive plants listed in a recent global review, those introduced for horticulture were most numerous (62% of species: 196 trees, 187 shrubs), followed by species introduced for forestry (13%), food (10%), and agroforestry (7%) (Richardson and Rejmánek, 2011). In Europe, 62.8% of all currently naturalized aliens were intentionally introduced. Ornamentals and escapees from horticulture account for the highest number of species (52.2%). Among accidental introductions, contaminants of seed, mineral materials and other commodities are responsible for 1091 alien species introductions to Europe (76.6% of all species introduced unintentionally) and 363 species are assumed to have arrived as stowaways (directly

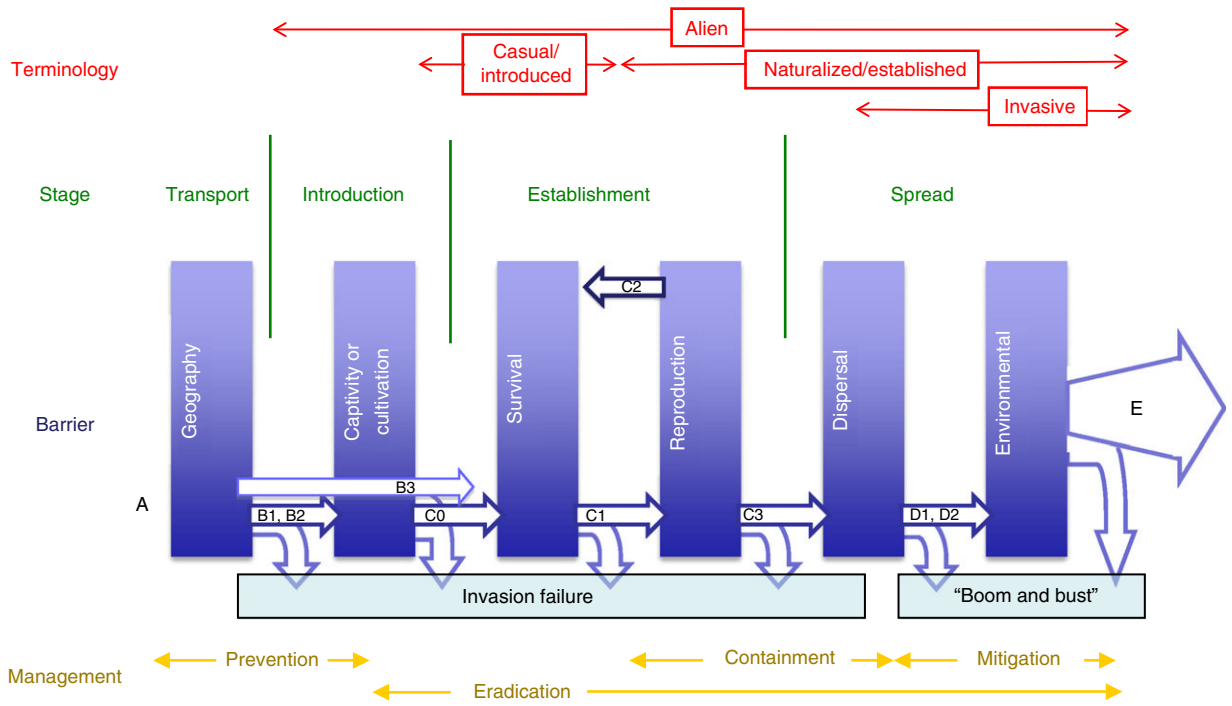


Figure 3 A proposed unified framework for biological invasions, recognizing that the invasion process can be divided into a series of stages, that in each stage there are barriers that need to be overcome for a species or population to pass on to the next stage. Different parts of the framework emphasize views of invasions that focus on individual, population, process, or species. Reproduced from Blackburn TM, Pyšek P, Bacher S, *et al.* (2011), with permission from Elsevier. A proposed unified framework for biological invasions. *Trends in Ecology and Evolution* 26: 333–339.

associated with human transport but arriving independently of commodity) (Lambdon *et al.*, 2008).

Naturalization

Empirical evidence shows that the chance of becoming naturalized, and later invasive, increases markedly with an increase in the number of propagules (any part of a plant that can give rise to a new individual, most often seeds, but also parts of vegetative reproductions such as bulbs and rhizomes) introduced, and with multiple introductions (including introductions at different times and from different source populations). More propagules reduce the likelihood of extinction and increase the chance of long-distance dispersal. Multiple introductions allow the incipient invader to sample a greater range of sites over space and time in the new environment and (since different introductions often originate from different source populations) improve the likelihood of introducing a genotype closely suited to local conditions. Also, multiple introductions increase the likelihood of forming novel genotypes that facilitate invasion, by hybridization of genotypes that may not meet in the native range; for this reason, some introduced populations may exhibit a greater genetic variation than is the case in the native range, although the opposite is more often true as a result of a genetic bottleneck. Successful establishment entails dealing with numerous physical, chemical, and biotic barriers. Many introduced plants are initially grown in small populations that are inherently susceptible to extinction due to chance events. To establish and persist, a population must exhibit $dN/dt > 0$ when N is small (the invasion criterion). If an introduced plant can deal with various reproductive barriers, and can reproduce in the wild without human assistance for a sufficiently long period of time to cope with local conditions, it becomes naturalized.

Spread

Invasion involves dispersal within the new area and population growth. Species in invasive alien floras show a wide range of adaptations for dispersal. Many species are dispersed by passive agents such as water or wind. A large proportion of the world's most widespread and damaging invaders, especially woody species, are dispersed by birds and mammals. More than 40% of invasive tree species and over 60% of invasive shrub species are dispersed by birds (Richardson and Rejmánek, 2011). The rapidity with which mutualistic seed-dispersal interactions establish suggests that vertebrate-dispersed plants have converged into generalized dispersal syndromes regardless of phylogenetic and geographical origins. Epizoochorous dispersal, mainly by cattle and sheep, facilitates the spread of many (mainly herbaceous) invaders. Vegetative reproduction is also important, especially in aquatic environments. For example, within the water hyacinth genus, the most invasive species (*E. crassipes*) is the one with the highest capacity for vegetative multiplication.

The dynamics of range expansion and population growth of invasive alien plants typically follow a standard pattern. There is frequently a time lag between the arrival of an alien plant in a new habitat and the start of widespread invasion. Examples include *Thlaspi caerulescens* (Brassicaceae), which was cultivated at Oslo Botanical Garden in Norway since 1814 and was first collected as an escapee in 1874, spread slowly until 1900, then expanded rapidly until it reached most of its present range in about 1945; *Mimosa pigra* (Fabaceae), which was virtually confined to small areas around Darwin in Australia for 80 years before exploding; *Melaleuca quinquenervia* (Myrtaceae), which showed no invasive tendencies for its first 50 years in Florida (United States); and *H. mantegazzianum* which was introduced to the Czech Republic in 1862, first recorded outside cultivation in 1877, but only starting to spread substantially some 70 years later (Pyšek *et al.*, 2009). Such lags are usually due to one or more of the following (non exclusive) factors: (1) the founder population may maintain a stable, small population until genetic adjustment occurs or essential mutualists (seed dispersers, pollinators, and mycorrhizal fungi) arrive (the extent of genetic alteration preceding or accompanying invasion has been studied for a few widespread invaders, for example, *Ailanthus altissima* in the United States); (2) some lags are probably explained by an initial shortage of safe sites, which have become more abundant as human-induced disruption of ecosystems has increased (increasing time also improves the chance of a potential invader encountering a safe site formed by a rare event such as a flood); and (3) populations spread slowly at their periphery but only show accelerated growth rates when there are many foci of growth. In some regions, the sudden start of spread may also be related to landscape transformation. In the Czech Republic (Central Europe), many species entered the exponential phase of invasion after World War II when political change resulted in large-scale transformations of landscapes that created many disturbed, eutrophicated habitats.

The lag phase is followed by a phase of sudden growth during which populations increase at an exponential rate (and generally become noticed as invaders). One reason for the increased growth rate (besides those mentioned earlier that may prevent its realization) is the typical two-phase pattern of spatial expansion. This involves the densest recruitment of offspring close to founder populations (neighborhood diffusion) and the establishment of isolated colonists through long-distance dispersal. As satellite foci grow in size through diffusion, often coalescing with each other and the founder metapopulation, more propagules become available for additional jump dispersal. With increasing numbers of growth foci, population growth and range increase rapidly. The rate of spread is frequently augmented by intentional or accidental movement of plants within the invasion arena by humans, thus creating additional nascent foci. This process, termed "stratified diffusion," has been documented for many plant invasions involving disparate plant taxa and environments and is evident at scales ranging from global, to regional, to landscape. The end of the exponential phase usually occurs when most optimum sites for invasion are occupied (or when successful control is instituted). Alien plant invasions may also follow linear trajectories (e.g., in the case of spread along rivers or coastal foredunes). Invasions in these habitats usually proceed as a wave front.

Spread rates reported for invasive alien plants vary considerably. Some examples of the fastest aerial spread are 5000 $m^2\ y^{-1}$ for *B. tectorum* (Poaceae) in North America, and 4000–13,000 $m\ y^{-1}$ for *Heterotheca latifolia* (Asteraceae) in the Georgia piedmont. For linear spread over large areas, the

fastest reported invasion is *Wedelia trilobata* (Asteraceae), which invaded 2500 km of Australian coast in 15 years, yielding the average rate of 167 km y^{-1}. At the local scales, the rates of linear spread vary from 970 m y^{-1} for *Fraxinus ornus* (Oleaceae) along a river in France; 76 m y^{-1} for *M. pigra* (Fabaceae) in wetlands in northern Australia; 21–31 m y^{-1} for *Acacia cyclops* (Fabaceae) and *P. pinaster* (Pinaceae) in South African fynbos; and 14 m y^{-1} for *Ammophila arenaria* (marram grass; Poaceae) on coastal dunes in California (Pyšek and Hulme, 2005).

Invasiveness and Invasibility

A major thrust of studies on plant invasions in recent decades has been the search for traits that separate invasive species from noninvaders and features that distinguish invaded systems from those that are apparently able to resist invasion. Although the two concepts are intimately linked, the vast majority of studies have addressed the issues separately.

Invasiveness

Numerous studies have described the features of individual alien plant species that have enabled them to invade given areas. For example, the recipe for success for *Lantana camara* (Verbenaceae) includes the following ingredients: (1) dispersal of seeds for long distances by many native and introduced birds; (2) toxicity of its fruit for many mammals, which limits damage by herbivory; (3) its ability to resprout vigorously following damage (e.g., by trampling); (4) its ability to invade a wide range of habitats, including sites disturbed by alien mammals on islands; (5) production of allelopathic substances, which improves its competitive ability; and (6) its ability to flower profusely for long periods, thus attracting pollinators and ensuring copious seed set.

Attempts have been made to define invaders for particular regions (usually biomes or countries) or the entire world and to divide groups of plants into invaders and noninvaders. A notable early attempt to delimit traits of plants that enable them to become weeds or invaders was Herbert Baker's description of the ideal weed. Briefly stated, Baker's hypothetical superweed is a plastic perennial that germinates in a wide range of physical conditions, grows quickly, flowers early, is self-compatible, produces many seeds that disperse widely, reproduces vegetatively, and is a good competitor. Many of the world's most widespread invaders and weeds possess many "Baker characteristics," but none has all, and many notable invaders possess very few, if any, of these traits. Is there an optimum number of Baker's traits that confers invasiveness? The work of Mark Williamson and coworkers on British annual plants suggests that optimum weediness is realized in plants with an *intermediate* number of Baker characters (4 or 5 out of 7).

Plants that have performed best as alien invaders are clearly not evenly distributed among taxonomic groups. Among the angiosperms, two families are clearly overrepresented among the most widespread and damaging invaders of natural communities: Fabaceae and Poaceae. Other families that seem particularly prone to contributing invaders of natural areas are Myrtaceae, Rosaceae, Salicaceae, and Tamaricaceae. The Asteraceae, the largest family of flowering plants, contributes many species to alien floras in most regions, but the family is not significantly overrepresented. Among invasive trees, the Fabaceae and in particular the genus *Acacia* sensu lato, and Pinaceae, particularly the genus *Pinus* are exceptional. For shrubs, the family Rosaceae (particularly species in the genera *Pyracantha*, *Rosa*, and *Rubus*) is also noteworthy.

Biological traits most often reported to be associated with invasiveness are those related to size, vigorous spatial growth, high fecundity, efficient dispersal, small genome size, and some physiological traits such as a high relative growth rate or a high specific leaf area. For example, differences in invasiveness among pine species (*Pinus*) can be explained using only three traits (seed mass, length of juvenile period, and interval between seed mast years), and if dispersal by vertebrates and characteristics of fruits are included, invasions of woody species can be reasonably predicted using this suite of traits (Rejmánek and Richardson, 1996). However, it has been shown recently that traits interact with other factors, namely the propagule pressure and the residence time of a species in the region (by comparing species with largely different residence time, we are likely to obtain biased results as some alien species are still in the process of filling their ranges in the new area), and that they directly affect invasion success rather at the more advanced stages of the invasion process. This points to the context-dependence of the role of traits in plant invasions; their importance changes with the stage of invasion, and those that confer invasiveness at some stage may not do so at other stages.

Many studies have attempted to gain an understanding of the life-history attributes of plants that makes them successful invaders by comparing traits of native and alien species or invasive and less invasive alien plants. Rejmánek (1996) considered perspectives from many of these studies and added many new insights to arrive at a general model of seed plant invasiveness based on recognized and assumed causal and correlative relationships between various traits, including those related to the native geographical ranges. This model is probably close to the limit of what can be achieved in defining invasiveness at a global scale. The most practical predictor of invasiveness is still whether the species has become invasive elsewhere (discussed later).

Two hypotheses that have received considerable attention recently in the quest to explain invasiveness in plants are the Enemy Release Hypothesis and the Evolution of Increased Competitive Ability Hypothesis. Both deal with the performance of alien plants in the absence of their natural enemies (*see* Glossary).

Invasibility

Charles Elton's (1958) *The Ecology of Invasions by Animals and Plants* (which, it should be noted, is heavily biased in favor of evidence on animal invasions) put forward the view that undisturbed native communities are not susceptible to invasion by introduced species. Recent debate on this topic has been rather futile since so few communities on earth are unaffected

by human activities. What has become abundantly clear over the past few decades is that disturbance, be it naturally occurring or human induced, is a fundamental driver of plant invasions. Irrespective of other factors that facilitate or limit invasions, the susceptibility of communities to invasion by alien plants increases with increasing disturbance up to a threshold, after which disturbance acts as a barrier. Useful insights into the types and magnitude of disturbance required to initiate and sustain invasions emerge from examining cases of invasion of pine trees (*Pinus* species) in many parts of the Southern Hemisphere. Most records document pine invasions in areas where the natural disturbance regime has been noticeably altered by humans. Changes to disturbance regimes have involved mainly increased or decreased herbivore pressure (grazing, browsing, and trampling), altered fire frequency, or mechanical clearing of vegetation. Different vegetation types differ in the type and magnitude of disturbance required to facilitate invasions. All records of pine invasions into native forests showed that severe disturbance is required to facilitate seedling recruitment and population growth. Invaded grasslands in Australia, New Zealand, and South Africa invariably have markedly changed disturbance regimes. The same applies for most invaded shrublands. The ultimate cause of increased invasibility due to disturbance in these cases is clearly the reduced competition from resident plants through the reduction in ground cover. Several pine species have spread into South African fynbos where the only disturbance is naturally occurring fire; intense fires and the dearth of vigorous plants immediately postfire combine to create opportunities for invasion. In general, intermediate levels of disturbance clearly make sites more vulnerable to invasion by pines, whereas severe disturbance usually prevents invasion. Naturally occurring disturbances where vegetation cover is removed are caused by factors such as volcanicity, slope instability, wind, and flooding; these are also important determinants of invasibility for pines.

Empirical evidence shows that certain communities are more susceptible to invasion by a wide range of alien plants than others. For example, the greater susceptibility to invasion of island communities (the more isolated, the more invasible) compared with mainland ones has long been noted. Islands where established alien species make up 40% or more of the flora include Bermuda, Hawaii, Ascension, Rodrigues, Tristan da Cunha, Lord Howe, and New Zealand (the figures for continental areas are much lower). Among the many reasons advanced to explain this are: (1) the species richness of islands is low (theoretically reducing the likelihood of an introduced species encountering a close competitor); (2) island floras have evolved in isolation, often without adaptation to high levels of competition, grazing, trampling, or regular burning; (3) the small size of islands means that the history of human intervention is concentrated (more intense disturbance); and (4) many islands were colonized very early and have been the crossroads of intercontinental trade (and thus were exposed to greater numbers of potential invaders).

Among other systems that seem particularly susceptible to alien plant invasions are riparian zones. In virtually every region, these zones are among the most severely invaded habitats. An important reason for this is the frequent disturbance from water-level fluctuation, including floods, which disperses propagules and provides favorable sites for establishment.

Variation in how much a community, ecosystem or a region is invaded could be simply due to differences in the number of arriving aliens. Therefore, we need to ask not only whether the region has more alien species, but whether it is intrinsically more susceptible to invasions (Lonsdale, 1999). Because most invading species fail to establish, such intrinsic invasibility can only be determined if processes of immigration and extinction are considered. It is important to distinguish between two measures. Invasibility is the inherent vulnerability of a community to invasion and it is ideally measured as the survival rate of alien species introduced to the system, thus accounting for losses due to competition with resident biota, effects of enemies, chance events and other factors (Lonsdale, 1999). Invasibility differs from the level of invasion, which integrates the effects of invasibility, propagule pressure and climate and is defined as the actual number or proportion of alien species in a community, habitat, or region (Chytrý *et al.*, 2008). Therefore, relatively resistant communities can be invaded to a high level if exposed to high propagule pressure. Even relatively vulnerable communities will experience low-level invasions if propagule pressure is low.

Recent studies on invasions in European habitats (Chytrý *et al.*, 2008) showed that: (1) habitats differ considerably in their invasibility. The differences in the level of invasion between Central European habitats are mainly caused by inherent habitat properties and to a lesser extent by propagule pressure and climatic differences between regions. (2) Patterns of habitat invasion are consistent across different regions of Europe. The same habitats usually have either a high or a low level of invasion despite their geographical location. The most invasible habitats are those with fluctuating availability of resources, especially nutrients; most of these habitats are frequently or strongly disturbed.

These studies serve as an example of how information collected at a fine scale of invasions (proportion of alien species in the total number of species in a vegetation plot of tens to hundreds of square meters) can be translated to a regional and continental scale and provide information for risk-assessment. A European map of risk from alien plant invasions (Chytrý *et al.*, 2009) was constructed based on the levels of invasions in particular habitats and justified by previous work showing that the habitat identity is a suitable descriptor of the level of invasion (**Figure 4**).

Linking Plant Traits to Environmental Features

There have been important recent advances in our understanding of the determinants of invasiveness and invasibility. There is now a reasonable understanding, at least for some severely invaded regions, of the plant traits that are correlated with success as invaders and also the environmental features that mediate invasions. Further progress, certainly at scales of resolution that have value in management, will rely on the development of models that link plant traits with critical environmental features. Such models show that interactions between the various determinants of invasive success are

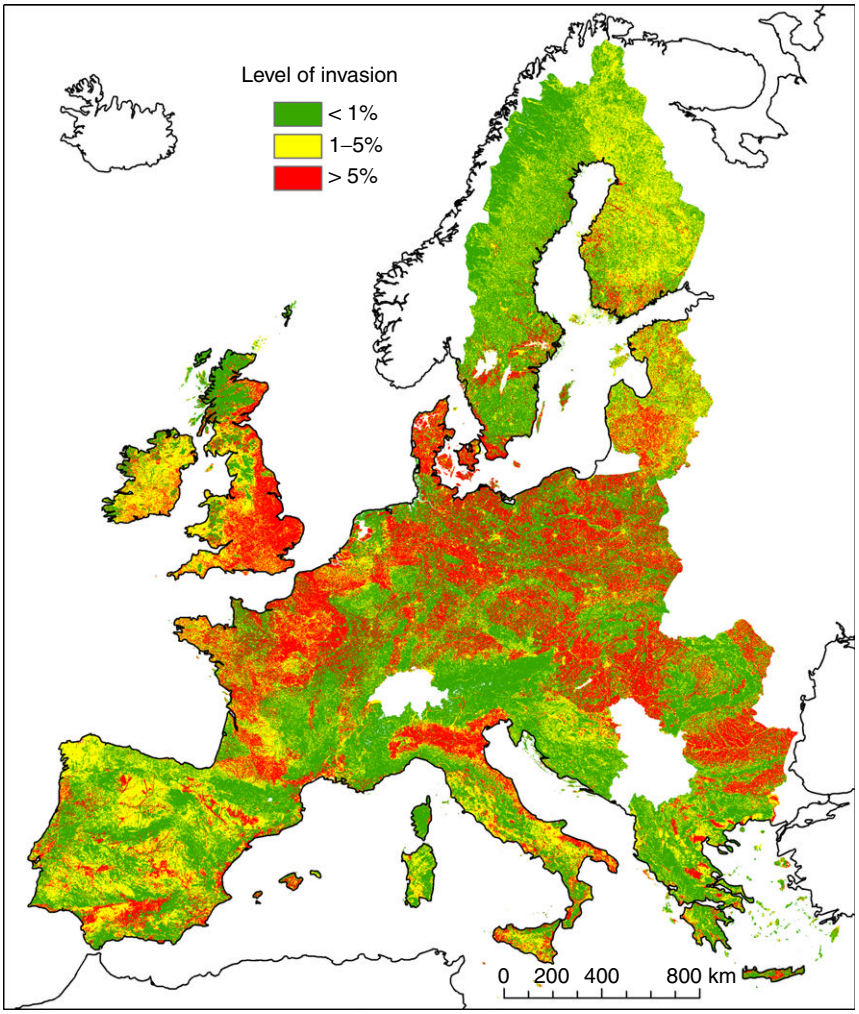


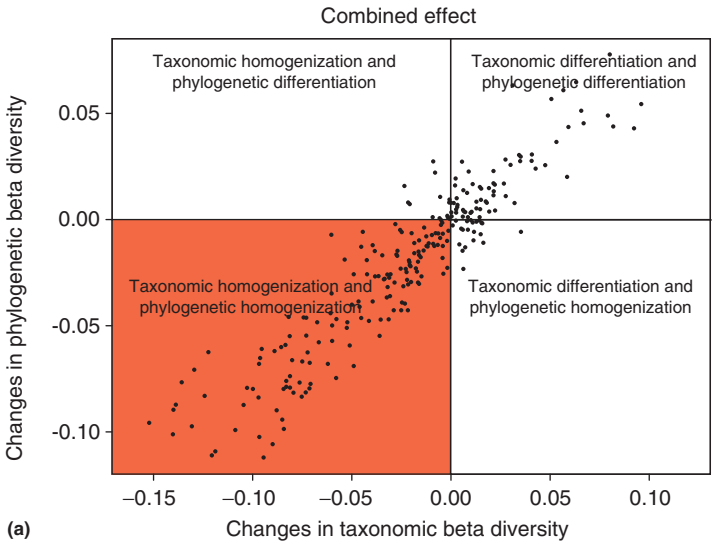
Figure 4 European map estimating the level of invasion by alien plants, based on the mean percentage of neophytes (species introduced after AD 1500) in vegetation plots. Within the mapping limits, areas with nonavailable land-cover data or insufficient vegetation-plot data are blank. Reproduced from Chytrý M, Pyšek P, Wild J, Pino J, Maskell LC, and Vilà M (2009) European map of alien plant invasions based on the quantitative assessment across habitats. *Diversity and Distributions* 15: 98–107, with permission from Wiley.

sometimes at least as important as the main effects. For example, in the case of pines invading forests, shrublands, and grasslands in the Southern Hemisphere, interactions between basic features of the environment, the features of the disturbance regime, and plant traits explain the different spread rates observed in different areas.

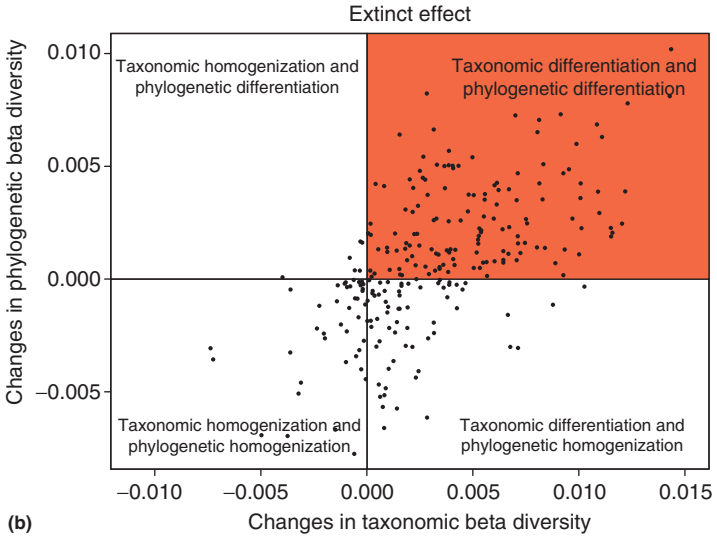
Modeling Plant Invasions

The use of models has considerably improved our understanding of invasions and how to deal with them. The many types of models that have been applied to plant invasions may be crudely grouped into three categories. *Simple demographic models* include exponential, logistic, logistic-difference, and stochastic models; they predict the future number of individuals in a population by making assumptions about the nature of population growth and by estimating the demographic parameters regarded as being important in determining population dynamics. *Spatial-phenomenological models* describe plant–environment interactions using empirical data (without invoking ecological mechanisms); they include regression models, geometric models, and Markov models. *Spatial-mechanistic models* are based on independent estimates of ecological parameters affecting invasions; they include reaction–diffusion models, population dynamic metapopulation models, and individual-based cellular automata models.

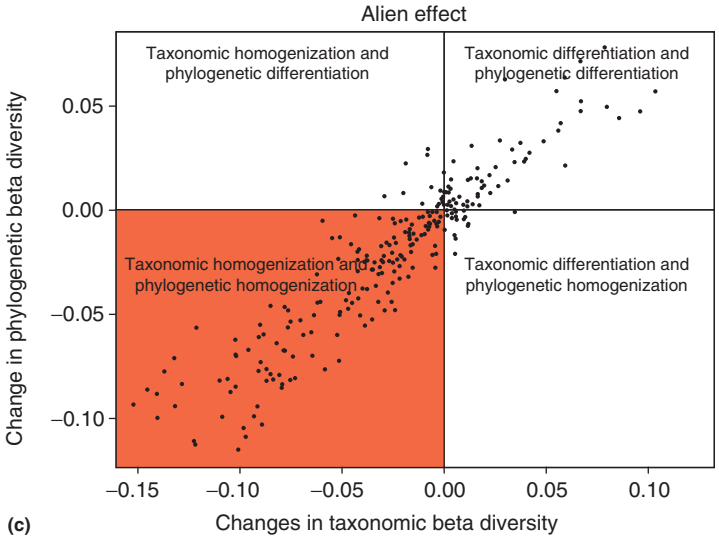
A particularly influential model in the study of invasions was that published by Skellam in 1951. He used a diffusion equation, combined with estimates of population growth, to show that invasion fronts advance at a constant velocity. Skellam's model, on which most subsequent invasion models have been based, did not incorporate long-distance dispersal, now known to be critically important for modeling the rate and variability of an expanding population front. Recent advances in mathematical modeling have facilitated the development of spatial-mechanical models (notably individual-based,


Combined effect
Taxonomic homogenization and phylogenetic differentiation
Taxonomic differentiation and phylogenetic differentiation
Taxonomic homogenization and phylogenetic homogenization
Taxonomic differentiation and phylogenetic homogenization
Changes in phylogenetic beta diversity
0.05
0.00
-0.05
-0.10
−0.15
−0.10
−0.05
0.00
0.05
0.10
Changes in taxonomic beta diversity

(a)


Extinct effect
Taxonomic homogenization and phylogenetic differentiation
Taxonomic differentiation and phylogenetic differentiation
Taxonomic homogenization and phylogenetic homogenization
Taxonomic differentiation and phylogenetic homogenization
Changes in phylogenetic beta diversity
0.010
0.005
0.000
−0.005
−0.010
−0.005
0.000
0.005
0.010
0.015
Changes in taxonomic beta diversity

(b)


Alien effect
Taxonomic homogenization and phylogenetic differentiation
Taxonomic differentiation and phylogenetic differentiation
Taxonomic homogenization and phylogenetic homogenization
Taxonomic differentiation and phylogenetic homogenization
Change in phylogenetic beta diversity
0.05
0.00
−0.05
−0.10
−0.15
−0.10
−0.05
0.00
0.05
0.10
Changes in taxonomic beta diversity

(c)

cellular automata mixture models), which allow much more accurate modeling of all stages of plant invasions. Besides offering exciting opportunities for exploring critical processes in invasion, thus contributing to invasion theory, these models have important applications in management (Hui *et al.*, 2011).

Managing Plant Invasions

Plant invasions cause many types of impacts that justify management intervention. Impacts may manifest at the scale of populations, communities, or ecosystems. Some impacts of invasive plants are rapid and dramatic, such as those that transform entire ecosystems over a decade. Others are slow and cryptic, such as those involving the infiltration of pollination or seed-dispersal networks that may cause no dramatic effects but which have profound influences at longer time scales. Impacts may be measured using ecological or economic metrics (Pyšek and Richardson, 2010). An important implication of biological invasions is biotic homogenization. Winter *et al.* (2009) have demonstrated how the European flora is being homogenized (both phylogenetically and taxonomically) through a combination of plant extinctions and introductions/invasions (**Figure 5**).

There are three fundamental options for managing invasive alien plants: (1) prevention, (2) removal (eradication), and (3) ecological management. The appropriateness of different control approaches depends on the ecology of the invader and the invaded system, and on many socioeconomic issues.

Prevention

Despite pessimistic prognoses in the 1980s regarding the ability of ecology to predict plant invasiveness and invasibility at scales useful for management, important advances have been made. Various methods have been developed for screening introduced plants to assess the risk of them invading particular habitats. As mentioned earlier, a good predictor of invasiveness is whether a species has invaded other (similar) areas where it has a longer history as an alien. For example, 90% of invasive plant species in Australia are also invasive in other locations. A major problem with applying this concept to management or regulation is the time lag inherent in many (most) invasions – at what stage should the performance of a species be scored with regard to invasiveness? Also, many species are currently being introduced directly from their native ranges to many different areas simultaneously.

Removal

The options for controlling an invasive plant decline rapidly when the phase of exponential increase is reached. Management is clearly most effective when directed at removing invaders while they still occupy a small range and occur in small populations. Many potentially widespread invaders are effectively kept in check by various means of control at this stage. Removal is much less feasible once the invaders have spread over large areas, but is still attempted in some cases, usually when the damage the invaders cause overrides economic constraints and when other control measures are unavailable or impractical. For example, expensive, large-scale clearing of invasive pines and other alien trees in South African fynbos is warranted by estimated costs of reduced water production in the absence of control.

Ecological Management

Experience has shown that satisfactory control of plant invaders is usually only achieved when several complementary methods, including biological control, improved land management practices (e.g., through prescribed burning and modified stocking densities), herbicides, and mechanical methods are carefully integrated.

Confounding Factors – New Trajectories and Conflicting Human Perceptions

Invasive plants are being effectively managed in many parts of the world, albeit at a huge, and increasing, cost. The dimensions of problems associated with plant invasions are changing rapidly, and management strategies need to be flexible enough to deal with exigencies. Among the main reasons for changing trajectories of invasions are the following:

- The increasing magnitude and pervasiveness of global trade and free trade agreements means that more species are being introduced and in greater number (increased propagule pressure).
- The increasing availability and demand for alien plants for a wide range of purposes, especially for agroforestry and horticulture.
- Global change (including climate change, changed nutrient regimes, elevated CO_2, habitat fragmentation); although little is known of how all these factors affect invasions, especially in concert, most ecologists agree that these changes will exacerbate problems with invasive alien plants.

Figure 5 Changes in taxonomic and phylogenetic beta diversity, a measure of how species composition has changed in the past centuries, among European regions, based on comparison between current and past floras, and plotted for the effects of species extinctions (b) invasions (c) and combination of both (a). Red areas indicate the significant effects of both measures. Extinctions of native species lead to taxonomic and phylogenetic differentiation (current regional floras are less similar to each other than were original floras consisting of native species only) because they were local and have not occurred across the whole of Europe. Invasions lead to taxonomic and phylogenetic homogenization, with current floras being more similar to each other than was the case with original floras, because many alien species are widespread and occur in many regions. Since invasions considerably exceeded extinctions in numbers, they have an overwhelming effect if both processes are assessed in concert. Reproduced from Winter M, Schweiger O, Klotz S, *et al.* (2009) Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proceedings of the National Academy of Sciences of the USA* 106: 21721–21725, with permission from PNAS.

- Genetic engineering (GMOs, transgenic plants); genetically modified plants could acquire novel traits that may make them (better) invaders.
- Conflicts of interest; many alien plants are essential crops in some parts of the landscape/region/country and damaging invaders in other parts. Good examples are many tree species used in forestry and agroforestry, and *Echium plantagineum* (Boraginaceae), an important dry-season forage and a honey plant in South Australia, but an important weed of pastures in New South Wales. A major challenge is to develop objective methods for assessing the costs and benefits associated with our increasing dependence on alien species.

See also: Agricultural Invasions. Economic Control of Invasive Species. Introduced Species, Impacts and Distribution of. Land-Use Issues

References

Chytrý M, Maskell LC, Pino J, *et al.* (2008) Habitat invasions by alien plants: A quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology* 45: 448–458.

Chytrý M, Pyšek P, Wild J, *et al.* (2009) European map of alien plant invasions based on the quantitative assessment across habitats. *Diversity and Distributions* 15: 98–107.

DAISIE (ed.) (2009) *Handbook of Alien Species in Europe*. Berlin: Springer.

Drake JA, Mooney HA, and di Castri F (eds.) (1989) *Biological Invasions. A Global Perspective*. Chichester: Wiley.

Elton CS (1958) *The Ecology of Invasions by Animals and Plants*. London: Methuen.

Hui C, Krug RM, and Richardson DM (2011) Spread models in invasion ecology. In: Richardson DM (ed.) *Fifty Years of Invasion Ecology. The Legacy of Charles Elton*, pp. 329–343. Oxford: Wiley-Blackwell

Lambdon PW, Pyšek P, Basnou C, *et al.* (2008) Alien flora of Europe: Species diversity, temporal trends, geographical patterns and research needs. *Preslia* 80: 101–149.

Lonsdale M (1999) Global patterns of plant invasions and the concept of invasibility. *Ecology* 80: 1522–1536.

Mack RN (1989) Temperate grasslands vulnerable to plant invasion: Characteristics and consequences. In: Drake J, di Castri F, Groves R, *et al.* (eds.) *Biological Invasions: A Global Perspective*, pp. 151–179. Chichester: Wiley.

Pyšek P and Hulme PE (2005) Spatio-temporal dynamics of plant invasions: Linking pattern to process. *Ecoscience* 12: 302–315.

Pyšek P, Cock MJW, Nentwig W, and Ravn HP (eds.) (2007) *Ecology and Management of Giant Hogweed (*Heracleum mantegazzianum*)*. Wallingford, UK: CAB International.

Pyšek P, Jarošík V, Pergl J, *et al.* (2009) The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. *Diversity and Distributions* 15: 891–903.

Pyšek P and Richardson DM (2010) Invasive species, environmental change and management, and ecosystem health. *Annual Review of Environment and Resources* 35: 25–55.

Pyšek P, Richardson DM, Pergl J, Jarošík V, Sixtová Z, and Weber E (2008) Geographical and taxonomical biases in invasion ecology. *Trends in Ecology & Evolution* 23: 237–244.

Rejmánek M (1996) A theory of seed plant invasiveness: The first sketch. *Biological Conservation* 78: 171–181.

Rejmánek M and Richardson DM (1996) What attributes make some plant species more invasive? *Ecology* 77: 1655–1661.

Richardson DM and Rejmánek M (2011) Trees and shrubs as invasive alien species – A global review. *Diversity and Distributions* 17: 788–809.

Richardson DM, Daehler CC, Leishman MR, Pauchard A, and Pyšek P (2010) Plant invasions: Theoretical and practical challenges. *Biological Invasions* 12: 3907–3911.

Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta DF, and West CJ (2000) Naturalization and invasion of alien plants – concepts and definitions. *Diversity and Distributions* 6: 93–107.

Winter M, Schweiger O, Klotz S, *et al.* (2009) Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proceedings of the National Academy of Sciences of the USA* 106: 21721–21725.

Plant Phenology Changes and Climate Change

Tim Sparks, Technische Universität München, Germany; Poznań University of Life Sciences, Poland, and University of Cambridge, UK
Annette Menzel, Technische Universität München, Germany

© 2013 Elsevier Inc. All rights reserved.

Glossary

Maximum temperature The highest temperature recorded in a 24-h period (usually daytime), often then averaged across a calendar month.
Mean temperature The average of minimum and maximum temperatures.
Minimum temperature The lowest temperature recorded in a 24-h period (usually night-time), often then averaged across a calendar month.
Phenology The study of the timing of natural events, such as leafing or flowering.
Phenophase The phenological event that is being recorded, such as first leafing.
Synchrony The temporal matching of phenological events.

Phenological Data

The changing of the seasons in temperate zones has been the subject of human fascination for many centuries. Alongside prevailing weather, it forms a staple topic of conversation. It has been, and continues to be, economically important by dictating farming and forestry operations. This interest has resulted in various levels of organization of recording observations in the changing seasons.

Records of the timing of events, known as phenophases, typically concern the beginning, peak, or end of a particular botanical event. In the case of farming, these will also include operations such as the drilling and harvesting of a crop. It is simpler and less subjective to record the timing of a first or last event (such as leaf bud burst, first leaf, first flower, end of flowering, and trees bare) than it would be to record (by estimation or measurement) the date of the peak. In broad terms, the first and last events are more typical in networks involving many recorders while more specialized phenological recording, often using growth stages such as biologische bundesanstalt, bundessortenamt, and chemical industry (BBCH) codes (Meier, 1997), is more likely within research organizations.

Even for the simple phases, such as first flower, there is a clear need to have an accurate description of the event. Many will record first flower when the center of the flower is visible, but undoubtedly there are records made when the petals are still closed. In some species, flowering may be defined as when pollen is being released, for example, in grasses or in catkin species such as Hazel *Corylus avellana*. For plant species with complex leaves, such as Ash *Fraxinus excelsior* or Horse Chestnut *Aesculus hippocastanum*, the definition of first leaf needs careful attention. Autumn events associated with leaf color changes or leaf fall will also need clear definition. For perennial species, it is possible to record the same individual plant year after year, but this is obviously not possible for annual species. The more advanced networks of recorders, such as that operated by the Deutscher Wetterdienst (German Meteorological Service), provide both a written description and a photographic example.

The bulk of plant phenological data concerns events in spring, partly because these events are easier to record and partly because spring is psychologically more important: a time of rebirth and renewal. The Old Testament quotes "For, lo, the winter is past, the rain is over and gone; the flowers appear on the earth; the time of the singing of birds is come …." Within spring events, leafing and flowering phases are clearly the most popular. The timing of spring must have been kept by Man since he could first keep records, but the oldest written records we are aware of come from the former Japanese capital of Kyoto where the dates of Cherry blossom festivals were recorded. In broad terms, national networks of recorders developed in three phases: from the 1870s under the auspices of learned societies, after the Second World War under the auspices of government hydrometeorological institutes, and in the 21st century in response to climate impacts monitoring (**Figure 1**). In addition to organized networks there have been many valuable records collected by

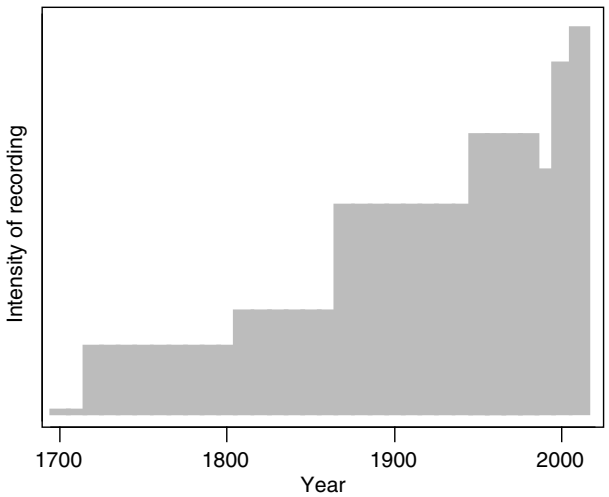


Figure 1 A diagrammatic representation of the intensity of phenological recording over the last three centuries.

individuals (e.g., Fitter and Fitter, 2002) and local groups (e.g., Abu-Asab *et al.*, 2001).

More recently, new technology has started to influence how we record the changing of the seasons. From 1981, it has been possible to assess the greenness of the Earth *via* satellite remote sensing. Data can be examined to see when spring arrives (when vegetation becomes photosynthetically active) over a large scale by assessing the "green wave" (Schwartz, 2003). This technology is unlikely to take over from traditional observations because there are still some problems associated with satellite images being obscured by cloud, the relatively low frequency of satellite recording (8 days), and the inability of a satellite to detect events such as flowering. More recently, options provided by remote image collection by digital cameras and webcams have started to be investigated. These overcome many of the frequency and cloud problems experienced by satellites, but still may fail to detect events such as flowering.

Phenological Response to Temperature and Other Environmental Drivers

Plant growth primarily requires light, water, mineral nutrients, and suitable temperatures. These apparently simple demands involve a large number of environmental factors and physiological processes, such as meteorological and edaphic as well as biotic factors, all of which may theoretically influence phenology. However, in temperate and boreal zones, the annual course of plant life is clearly split up: (1) a physiologically active growing season beginning with the recovery of photosynthesis and growth of evergreens (visible in phenology as "May shoots") or bud burst of deciduous plants ("leaf unfolding") and (2) a dormancy period during winter with cessation of growth and higher frost resistance. Prior to this, in autumn, the chlorophyll of deciduous leaves is decomposed ("leaf coloring") and finally leaves are shed to reduce the evaporative surface area in winter and to avoid frost damage ("leaf fall").

This synchronization of the morphological and physiological status of plants to the seasons is primarily triggered by temperature and day length. The photoperiod roughly restricts growth to suitable time periods, within which the timing of spring events, such as bud burst, leaf unfolding, and flowering, is mainly regulated by temperature: chilling temperatures break winter dormancy and subsequent warm temperatures induce bud burst.

Thus, as long as the chilling requirements are fulfilled, the phenological onset of spring correlates very well with air temperature of the preceding months. Many studies have shown strong, significant correlations between winter/spring temperatures and phenological spring phases, such as bud burst, leaf unfolding, or flowering in mid and higher latitudes (see Menzel, 2003; Root *et al.*, 2003; Sparks and Menzel, 2002; Walther *et al.*, 2002), which seems to be nearly linear over a broad temperature range (**Figure 2**, also **Figures 3** and **4**). Menzel *et al.* (2005) demonstrated that this response of spring events to temperature exists both for temperatures at the larger scale and for local climate station data. Even for time series covering the last century, no renunciation of the linearity at the warmer temperature tail is apparent.

An extensive survey of several European countries comprising 254 mean national records revealed that most phases correlated significantly with mean monthly temperatures of the month of onset and the two preceding months (Menzel *et al.*, 2006a). For 19% of the phenophases the highest correlation was with the month of onset, 63% with the preceding month, and 18% with temperature 2 months earlier. For spring and summer, and most of the autumn fruit-ripening phases, the mean correlation coefficient was negative. Thus, higher temperatures are related to earlier onset dates; for example, in flowering and leaf unfolding the average correlation coefficient was -0.69. Often the temperature response is quantified by means of linear regressions. The comprehensive European study further showed that spring and summer phases advanced by up to 4.6 days per °C warming, with an average advance of 2.5 days per °C and autumn leaf coloring was delayed by up to 2.4 days per °C (average 1.0 days per °C).

In addition, circulation patterns, such as the North Atlantic Oscillation which controls winter and spring temperature in the western parts of Europe, has been shown to explain up to 50% of the variability in spring phenological phases in Central and Northern Europe (e.g., Menzel, 2003). Consequently, spring phenological phases can serve as proxy data for spring air temperatures.

Autumn phenology of deciduous plants, such as leaf coloring or leaf fall, is generally more complex and in consequence less understood. Various environmental factors may drive autumn phenology, even the relationship between simple meteorological factors and leaf coloring/leaf fall is still obscure. Numerous hypotheses about weather or climate conditions modifying the onset of leaf coloring in temperate regions comprise occurrence of minimum temperatures, dropping of daily mean temperatures (inducing autumn dormancy), warm autumn weather and sufficient water supply (lengthening the growing season), high autumn temperatures (accelerating leaf coloring), warm June (promoting early leaf coloring), high radiation in autumn (inducing early leaf coloring), dry periods in summer (promoting leaf coloring), and some sort of constant length of the growing season or constant temperature sums (see Estrella and Menzel, 2006). Only at the regional level can it be deduced that warm August and September delay leaf coloring, whereas warm May and June advance it.

Other possible triggering factors for spring phenology discussed are soil moisture and atmospheric CO_2 concentrations. In many areas of the mid and higher latitudes, soils are mostly water saturated at the end of winter and the start of spring, thus soil moisture is not a limiting factor and has not been shown to influence phenological onset dates. However, there is some experimental and historical evidence for the Mediterranean climate zone that geographical changes in rainfall produced significant, complex, and strongly species-specific, as well as spatially and temporally variable, phenological effects (Penuelas *et al.*, 2004). The effect of elevated atmospheric CO_2 concentration, although it increased from a preindustrial value of 280 to 380 ppm nowadays, cannot be deduced from observational data. Experimental evidence suggests that there are only limited effects. For example, in

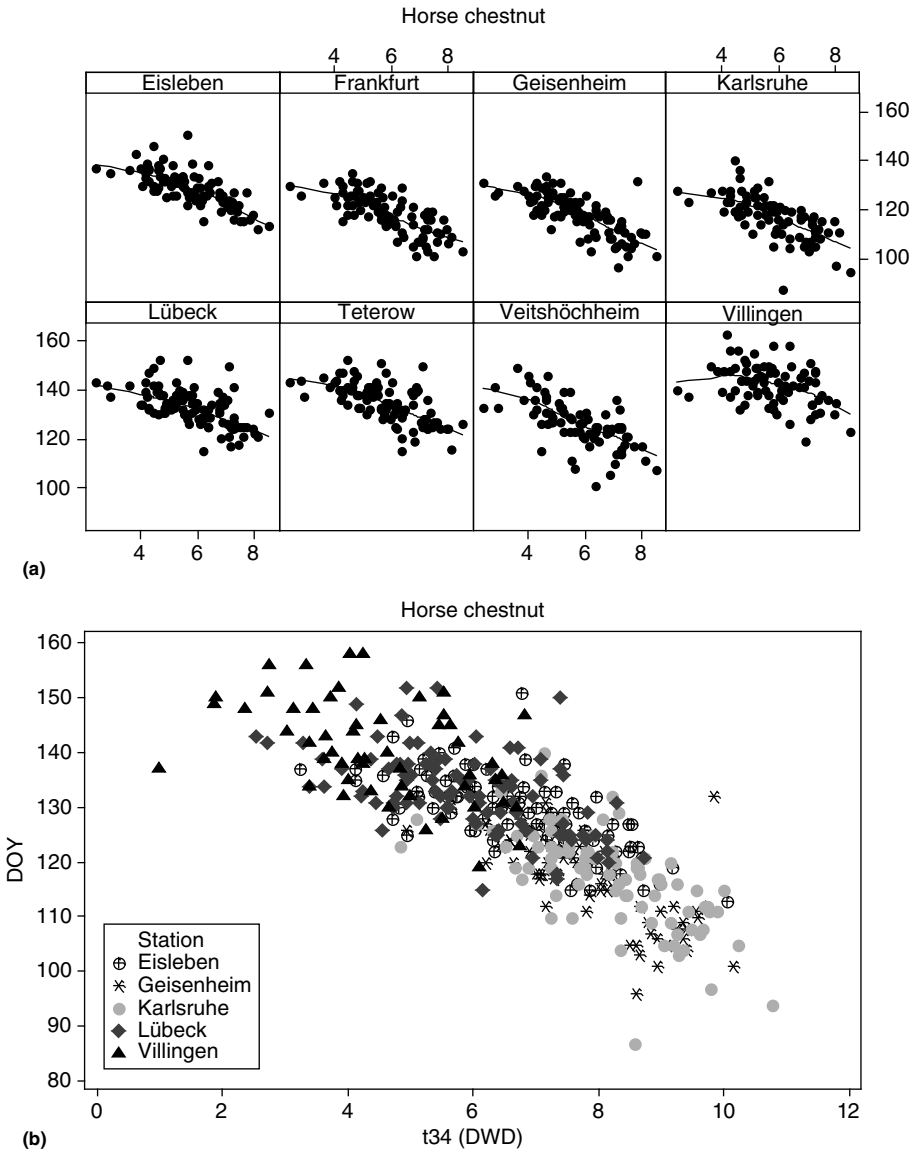


Figure 2 Beginning of flowering of *Aesculus hippocastanum* at stations in Germany vs. monthly mean air temperatures in March/April (a) using mean temperature for Germany TYN CY 1.1 dataset (Documentation for the TYN CY 1.1 data-set, 2006) and (b) using temperature data from adjacent climate stations of the German Meteorological Service. Reproduced from Menzel A, Estrella N, and Testka A (2005) Temperature response rates from long-term phenological records. *Climate Research* 30: 21–28, with permission from Inter-Research.

100-year-old temperate forest trees at the Swiss Canopy Crane site near Basel, Switzerland with 4 years of 540 ppm CO_2 atmosphere, phenological variables, such as bud break, leaf fall, and leaf duration, were highly species-specific and were not affected by elevated CO_2 in any consistent way (Asshoff *et al.*, 2006).

Evidence for Phenological Change

The fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC, 2007) concluded that temperature changes have already affected many physical and biological systems. Among the biological systems affected, phenology is one of the prominent bioindicators of climate change as their timing is strongly related to weather and climate. Since 2001, more and numerous studies have also demonstrated an earlier onset of spring events for mid- and higher latitudes and a lengthening of the growing season. Reviews and meta-analyses of this documented literature reveal a consistent pattern of change across the Northern Hemisphere (Parmesan and Yohe, 2003; Root *et al.*, 2003; Sparks and Menzel, 2002; Walther *et al.*, 2002; IPCC, 2007). The average advance of spring events amounted to 2.3 days per decade (Parmesan and Yohe, 2003) derived from 677 multispecies studies from any location that reported neutral, negative, and positive results. The second meta-analysis by Root *et al.* (2003) reported a stronger advance of 5.1 days per decade, because of restricting the analysis to publications reporting significant changes of one or many species.

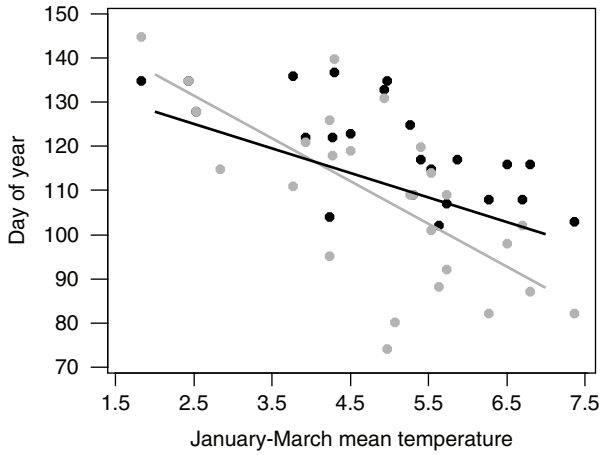


Figure 3 The relationship between Oak leafing (black symbols and line), Bluebell flowering (gray symbols and line), and January–March mean temperature.

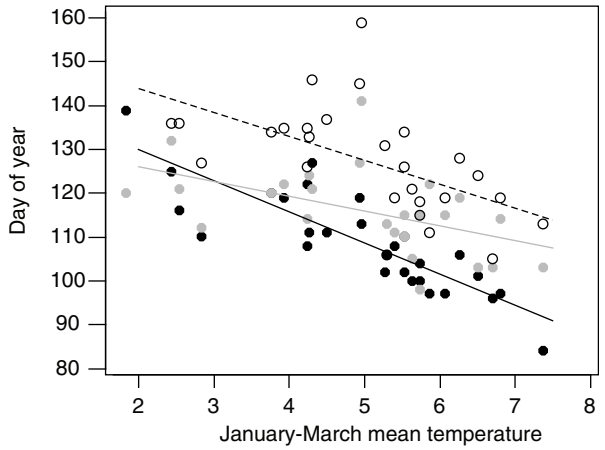


Figure 4 The relationship between Garlic Mustard flowering (black symbols and line), Lady's Smock flowering (gray symbols and line), Orange Tip butterfly appearance (open symbols, dashed line), and January–March mean temperature.

However, published single-site or single-species studies are particularly open to suspicion of being biased toward predominantly reporting climate change-induced impacts. Therefore, the European COST725 study analyzed an enormous systematic phenological network dataset of more than 125,000 observational series of 542 plant and 19 animal species in 21 European countries (1971–2000) (Menzel *et al.*, 2006a). Across Europe, 78% of all leafing, flowering, and fruiting records advanced (30% significantly) and only 3% were significantly delayed, whereas the signal of leaf coloring/fall was ambiguous. This shows that the previously published results of phenological changes are not biased by reporting or publication predisposition; the average advance of spring/summer in Europe almost exactly matched the result by Parmesan and Yohe (2003) of 2.5 days per decade.

In addition to these apparent changes in spring, a lengthening of the growing season has been observed in the mid- and higher latitudes of the Northern Hemisphere. Here, the different approaches based on phenological, satellite, and climatological studies reveal comparable results: a lengthening of the growing season of *ca.* 10–20 days in the last few decades, mostly owing to an earlier start (e.g., Menzel and Fabian, 1999; Walther *et al.*, 2002).

A coherent pattern of observed change does not mean equal changes across species and regions. Nearly all studies analyzing multispecies datasets found species-specific differences (see reviews by Sparks and Menzel, 2002; Walther *et al.*, 2002). In many cases, species resuming growth and activity earlier in spring exhibit the strongest advances in recent decades (Sparks and Menzel, 2002; Menzel *et al.*, 2006a; Menzel *et al.*, 2006b). In flowering, Fitter and Fitter (2002) deduced that annuals had advanced flowering more than congeneric perennials, and insect-pollinated species more than wind-pollinated ones. In a UK based meta-analysis, Thackeray *et al.* (2010) concluded that advances in terrestrial plant phenology were greater than those for terrestrial invertebrates or vertebrates. Furthermore, several studies have pointed out regional differences, for example, observed changes in spring are more pronounced in Western than in Eastern Europe. Quite interestingly, an earlier start of the growing season is linked to a higher spatial variability, which may imply further consequences of these observed changes (Menzel *et al.*, 2006b).

Consequences of Change

Plant species occupy a niche, environmentally and temporally. They persist only if conditions are suitable for their growth and reproduction. At some point in their life cycle they compete for resources with other plant species (competitors), they may be dependent on plant or animal species (allies), and they will invariably be subject to damage by animals (grazers). In earlier sections, it was very clear that plant phenology has changed as a consequence of a changing climate. Change has also been observed in the phenology of many invertebrate and vertebrate species.

If the phenology of the plants, their competitors, their allies, and their grazers all change in synchrony then it is obvious that the *status quo* will be maintained. However, this is clearly not the case (e.g., Fitter and Fitter, 2002) and we have to consider what the consequences may be of differential change in phenology. Some species are highly adaptable to changing environments (including climate) but as a consequence of this success they are typically a pest or a weed problem, and these species may do very well in a changing climate. Those that show an inability to respond may suffer from local extinction if competition or grazing increases, or if they lose synchrony with their allies. If they have low powers of dispersal, that is, the inability to move to an analog environment, then extinctions could be more widespread.

The consequences of phenological desynchrony are, as yet, less studied. However, this is a key area of research needed to inform conservationists where effort should be focused if biological diversity is to be maintained. Fitter and Fitter (2002) found clear differences in the changes in flowering phenology in British plants. Annual species had changed more than perennials, insect-pollinated species more than wind-pollinated species, and species close to the center of their

geographical range more than those further away. They also detected examples where opportunities for hybridization between related species would both increase and decrease. If a plant species is specialized to the insect species that pollinates it then phenological synchrony is likely to be more critical than if it can be pollinated by a range of insect species. Synchrony issues with grazers may be more critical for the herbivorous insects than for the plant species themselves, although marked changes in herbivory may affect plant growth. Specialized herbivorous insects may depend on one or a few plant species, but, in addition, at the correct developmental stage and hence nutritional value. A commonly used example uses Oak trees *Quercus* spp. and Winter Moth *Operophtera brumata* (e.g., Buse *et al.*, 1998) where phenological development of both leaves and caterpillars is very responsive to temperatures, but caterpillars must feed on the young leaves before tannin levels become too high.

In the recent past, there was considerable concern that Bluebells *Hyancinthoides non-scripta*, a characteristic flower of ancient woodlands in the UK, would suffer from earlier canopy closure, that is, exclusion of sunlight. However, both this herb and the dominant canopy species appear to be temperature responsive. In **Figure 3** are shown the response of Bluebell flowering and Oak *Quercus robur* leafing to January–March temperature. Data are taken from a single observer (A. Phillips) in the English West Midlands between 1976 and 2004. Both have highly significant ($p<0.001$) regression lines; Bluebell -9.7 ± 2.0 days per °C and Oak -5.6 ± 1.3 days per °C. Although the regression slopes are not significantly different ($p=0.103$), there is some suggestion that, at higher temperatures, Bluebells may gain extra days of direct sunshine before the canopy leaves. Such a suggestion needs further examination from more data; ideally, from both more years and more locations. Broad statements like this have to be tempered by the known greater phenological variability in warmer, and hence earlier, springs (Menzel *et al.*, 2006b).

Data from the same location can also be used to examine the relationship of flowering of Garlic Mustard *Alliaria petiolata*, Lady's Smock *Cardamine pratensis*, and the emergence of the Orange Tip butterfly *Anthocharis cardamines* whose caterpillars feed on the seed heads of these plants (**Figure 4**). The regression coefficient on January–March mean temperature of Garlic Mustard flowering is -7.1 ± 0.9 days per °C, that of Lady's Smock -3.4 ± 1.2 days per °C, and the Orange Tip -5.4 ± 1.6 days per °C. All coefficients are very significant ($p\leq0.01$) and those of the two flowers are significantly different from one another ($p=0.015$). This graph suggests that the butterfly may lose synchrony with Lady's Smock at higher temperatures, but maintain it with Garlic Mustard and that the latter may become a more important larval foodplant as the climate warms.

Unresolved Questions

It is clear from the previous section that the consequences of differentially changed phenology are not yet known. The conservation community has serious concerns that a breaking down in synchrony could be very serious for some species (Root *et al.*, 2003). The most susceptible species are likely to be those that exist at low population levels, in isolated communities, have specific ecological niches, have low powers of dispersal, and have low responsiveness to temperature. In other words, adaptable species are likely to benefit, and specialists suffer, from rising temperatures. These are, however, broad statements and should be treated with caution. Plant species do not exist in isolation from one another but in balance with their competitors, allies, and grazers. A change in status in any of these will affect the balance and there are too many permutations to make prediction a simple task.

Another issue to further muddy the waters is the introduction of new species. These could be either species expanding their geographical range in response to environmental change or could be alien species (such as garden escapes), which may have "jumped the garden fence" as the local climate becomes more suitable for their reproduction and persistence. There are numerous examples of exotic species, for example, Rhododendron *Rhododendron ponticum* and Japanese Knotweed *Fallopia japonica*, becoming serious environmental problems. Prediction of the effect of new species into existing communities is also at a very early stage.

Analysis, like that summarized in **Figures 3** and **4**, will typically fit a linear relationship between phenology and temperature. This straight line can only be an approximation to the real relationship and only be valid over a relatively narrow range of temperatures. There is no doubt that, in seasonal climates, there is a limit to how far a plant species can advance its leafing, flowering, or fruiting, or delay its leaf fall. Future warming is predicted to push our environments into new, nonanalog, situations outside of our experience. What will happen to these species at these new temperatures? It may be possible that we can gauge their reaction by looking at their situation in warmer parts of their distribution, although this is likely to be within a different plant community. Can species adapt, for example, through genetic selection? If so, how much can they change and over what timescale? Species with high-generation turnover and high reproduction rates are likely to be more adaptable. Many of these species are called weeds!

We have hardly touched in this article on autumn phenology because the changes so far experienced are more mixed and the drivers of change much more complex; probably involving temperature, sunshine, soil moisture, diurnal temperature range, and windiness. This is another area requiring more attention.

Finally, most research has involved a comparison with temperature, and typically with either mean temperatures or accumulated temperature sums. What has been lacking is an examination of the effects of extreme events (drought, frost, etc.) and the consequences of increased frequency in these extremes.

Demonstrating the responsiveness of plant species and differential response is relatively easy. Prediction of the consequences of change in face of the huge number of permutations of climatic change and plant-community structure is much more complex.

See also: Climate Change and Extinctions. Climate Change and Wild Species. Climate, Effects of

References

Abu-Asab MS, Peterson PM, Shetler SG, and Orli SS (2001) Earlier plant flowering in spring as a response to global warming in the Washington, DC, area. *Biodiversity and Conservation* 10: 597–612.

Asshoff R, Zotz G, and Korner C (2006) Growth and phenology of mature temperate forest trees in elevated CO_2. *Global Change Biology* 12: 848–861.

Buse A, Good JEG, Dury S, and Perrins CM (1998) Effects of elevated temperature and carbon dioxide on the nutritional quality of leaves of oak (*Quercus robur* L.) as food for the Winter Moth (*Operophtera brumata* L.). *Functional Ecology* 12: 742–749.

Documentation for the TYN CY 1.1 data-set, http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1_text.html (accessed November 2006).

Estrella N and Menzel A (2006) Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. *Climate Research* 32: 247–252.

Fitter AH and Fitter RSR (2002) Rapid changes in the flowering time in British plants. *Science* 296: 1689–1691.

IPCC (2007) Climate Change 2007: Impacts, Adaptation and Vulnerability. In: Change ML, Parry OF, Canziani JP, Palutikof PJ, van der Linden, and Hanson CE (eds.) *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate*. Cambridge, UK: Cambridge University Press.

Meier U (1997) *BBCH-Monograph. Growth stages of plants—Entwicklungsstadien von Pflanzen-Estadios de las plantas-Développement des Plantes*. Berlin und Wien: Blackwell Wissenschaftsverlag.

Menzel A (2003) Phenological anomalies in Germany and their relation to air temperature and NAO. *Climate Change* 57: 243–263.

Menzel A, Estrella N, and Testka A (2005) Temperature response rates from long-term phenological records. *Climate Research* 30: 21–28.

Menzel A and Fabian P (1999) Growing season extended in Europe. *Nature* 397: 659.

Menzel A, Sparks TH, Estrella N, *et al.* (2006a) European phenological response to climate change matches the warming pattern. *Global Change Biology* 12: 1969–1976.

Menzel A, Sparks TH, Estrella N, and Roy DB (2006b) Altered geographic and temporal variability in phenology in response to climate change. *Global Ecology and Biogeography* 15: 498–504.

Parmesan C and Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Penuelas J, Filella I, Zhang XY, *et al.* (2004) Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist* 161: 837–846.

Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, and Pounds JA (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421: 57–60.

Schwartz MD (ed.) (2003) *Phenology: An Integrative Environmental Science*. Dordrecht: Kluwer.

Sparks TH and Menzel A (2002) Observed changes in the seasons: An overview. *International Journal of Climatology* 22: 1715–1725.

Thackeray SJ, Sparks TH, Frederiksen M, *et al.* (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology* 16: 3304–3313.

Walther GR, Post E, Convey P, *et al.* (2002) Ecological responses to recent climate change. *Nature* 416: 389–395.

Plant–Soil Interactions

Joan G Ehrenfeld[†], Rutgers University, NJ, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Aerenchyma Porous root tissue, especially well developed in wetland plants, that allows diffusive flux of oxygen from above-ground tissues to root tips. This tissue supports the respiratory demand of the root tissues and allows oxygen to leak into the surrounding soil.

Mycorrhizae A root tip that is infected with fungi in a mutually beneficial partnership. The fungal hyphae explore large volumes of bulk soil, absorbing nutrients, and transferring them to the plant; the plant supplies the organic carbon necessary for growth and energy production to the fungus. Different groups of fungi form vesicular–arbuscular mycorrhizae (fungal hyphae invaginate into the plant root cells) and ectomycorrhizae (fungal hyphae grow between plant root cells and form a thick sheath over the root tip, but they do not invaginate). Several other forms are specific to particular plant families (Ericaceae, Orchidaceae).

Net mineralization/immobilization The net result of microbial decomposition of organic matter is either the incorporation of nutrient elements (particularly nitrogen) into the microbial biomass, rendering it unavailable for plant uptake (immobilization), or their release into the soil solution (mineralization) after microbial demand for each element has been satisfied.

Nutrient uptake capacity The instantaneous rate of nutrient acquisition, usually measured in brief (1–2 h) incubations. Uptake capacity reflects the abundance of transport sites on the root cell membranes and their affinity for nutrient ions.

Plasticity Ability of a plant to respond to temporal changes or spatial variation in environmental conditions by altering the size or the distribution of plant parts. These are phenotypic, rather than genetic changes.

Rhizodeposition The mixture of sloughed cells, mucilages, and low-molecular-weight sugars, amino acids, and other compounds leaked from root cells, which are deposited in the soil adjacent to the surface of fine roots. Exudation takes place from the root tip back to the zone of suberization. The chemical quality and quantity of the exudate is altered by the presence of mycorrhizae.

Rhizosphere Volume of soil adjacent to, and strongly influenced by a plant root. The rhizosphere is usually considered to extend about 2 mm from the root surface, and includes the "rhizoplane," or soil directly in contact with the root surface.

Siderophore Chemicals secreted by roots (primarily non-protein-forming amino acids), which complex with insoluble metal ions bringing them into solution and permitting their transport to and uptake into the root.

Soil aggregate A crumb-sized unit of soil, composed of aggregated soil minerals, microbes, and soil microfauna, which are cemented together by a combination of biological materials such as polysaccharide secretions, fungal hyphae, and chemical substances such as precipitated carbonates or silicates. Aggregates are classified by size and stability in water (disintegrating versus retaining their structure and integrity).

Soil organic matter Organic substances, including a wide variety of carbohydrates, proteins, lipids, waxes, phenolic, and humic compounds, which accumulate in soil as a result of both plant and microbial growth. These compounds include low-molecular-weight materials, which are rapidly decomposed to carbon dioxide; larger compounds, which may be slowly decomposed over years to decades; and large, complex, aromatic substances, which may be stable within the soil for millennia. Soil organic matter affects all aspects of the soil's biology, chemistry, and physics.

Soil texture The relative abundance of sand ($50\ \mu m < \phi < 2$ mm), silt ($2\ \mu m < \phi < 50\ \mu m$), and clay ($\phi < 0.2\ \mu m$) particles in the soil (USDA criteria). Particle size distribution determines the distribution of pore sizes, which in turn strongly affects the behavior of water in the soil.

Introduction

Plants are often assumed to be passive recipients of the environmental conditions that happen to exist at the location in which they grow. Soils supply mineral nutrients, and water, and they provide anchorage, but they also set limits on plant growth. Plants, however, provide the organic substrate that forms the basis for all biological activity in the soil and that forms the soil organic matter; their growth drives the formation of soil from bedrock and alters its chemistry and physics. This chapter reviews the two-way street that constitutes the plant–soil system and examines the implications of these mutual effects for the management and conservation of biodiversity.

Soil–plant interactions take place over a wide range of spatial and temporal scales (**Figure 1**). At a scale of microns, roots etch microscopic channels of weathering within mineral grains and provide a variety of different habitats for microbial growth. These microscale phenomena cause millimeter-scale differences in the abundance and diversity of microbes and the soil micro- and mesofauna. Nutrients flow toward roots over distances of millimeters to centimeters; active uptake causes sharp concentration gradients of nutrient elements, moisture, and acidity. These modifications spread out from the root surface to form a cylinder of differentiated, rhizosphere

[†]Deceased.

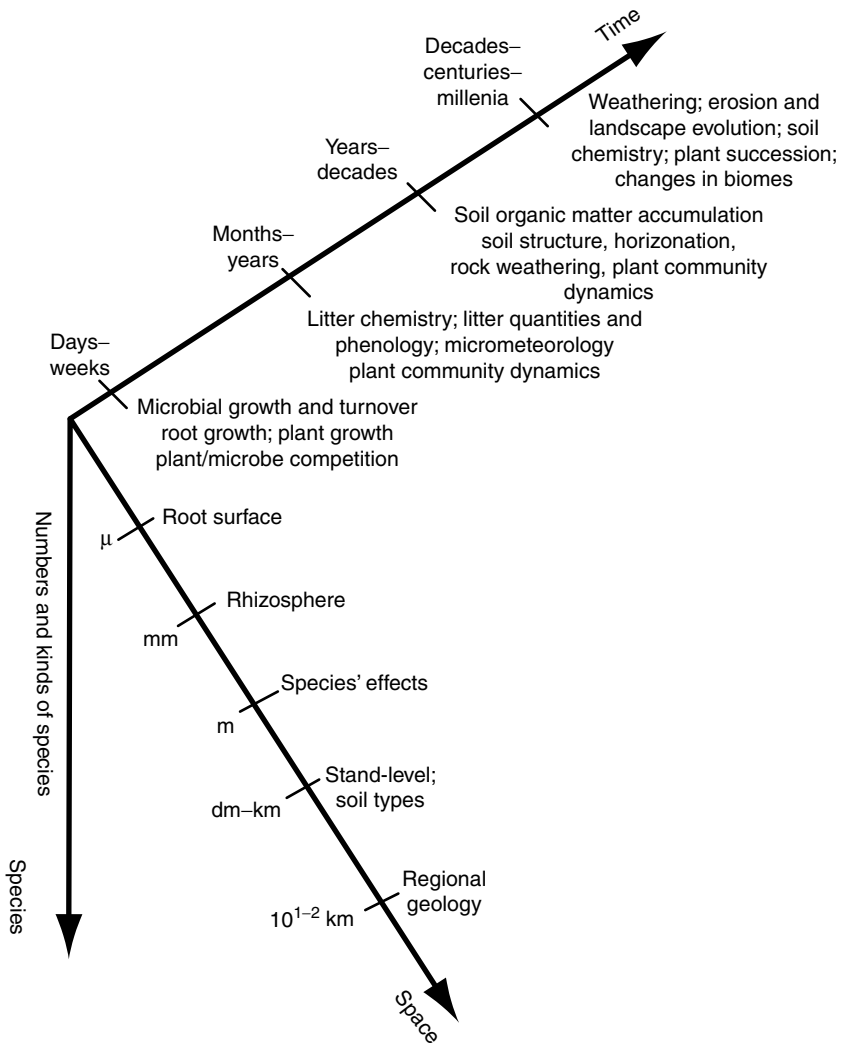


Figure 1 Concept of three-dimensional structure of plant–soil interactions. Processes occurring across dimensions of time (minutes to millennia) and space (microns to hundreds of kilometers) occur with respect to the species pool – both the numbers of species and the kinds of species.

soil that may be millimeters thick. Mycorrhizal hyphae extend from the root surface into the soil over distances from centimeters to meters; they bind soil particles together into aggregates and contribute to the long-term accumulation of soil organic matter. The quantity and quality of such millimeter-scale effects vary among plant species, and so patches of soil filled by roots of different species may become chemically and physically different at scales of centimeters to tens of meters, depending on plant size. These species-level effects may translate into stand-level effects. At these larger spatial scales, however, local and regional differentiation of soils due to differences in topography, geological substrate, and geological history may override the effects of plant species and constrain the nature of the plant–soil interactions. But at these scales, plant communities also exert control over the long-term accumulation of organic matter.

Plant–soil interactions occur over an equally great range of temporal scales. At the level of the individual plants, uptake of nutrients and water occurs more or less continuously over hours, but rates vary over seconds or minutes, depending on environmental conditions. The aggregate effects of this uptake over periods of months (seasons) or years reflects the continuity of plant community composition. Similarly, the exudation of substances used by the microbiota, and the response in microbial growth and metabolic activity, occurs over timescales of minutes to hours, but the aggregate effect on soil properties and on plant growth extend over seasons and years. Over the long term (decades to centuries or millennia), different plant communities cause different types of soil to develop.

Finally, these patterns of interaction themselves depend on the size and diversity of the available pool of plant species. In highly diverse communities, the effects of individual species may be masked by the variability among those effects, or by the overlapping nature of root and shoot systems. In communities that are species-poor, either because of human manipulation (e.g., forest plantations of a single species) or because of biogeographic and climatic constraints (e.g., boreal

forests with few species), the individual species-specific effects seen at small scales may translate into effects seen over hectares or square kilometers.

Within each of the temporal and spatial scales of interaction, the mechanisms creating the interaction are subjected to feedback, both positive and negative (**Figure 2**). Indeed, the presence of feedback in the soil–plant system is one of its salient properties: The results of a particular mechanistic effect alters the environment so as to modify (increase or decrease) the continued activity of that mechanism. Feedbacks have been recognized as crucial in mediating such disparate processes as the supply of mineral nutrients through decomposition, and the activities of soil pathogens and mycorrhizae.

The complex interplay of soil and plants has commanded attention for a long time. Gilbert White, in "The Natural History of Selbourne" in 1778, observed that "In a district so diversified with such a variety of hill and dale, aspects, and soils, it is no wonder that great choice of plants should be found." Darwin pointed out that plant growth is modified by the presence of soil fauna, especially earthworms; he noted that Gilbert White had also observed this a century earlier. The German scientist Justus Liebig (1876) founded the scientific basis for agriculture by emphasizing the role of soil in controlling vegetation and crops through the supply of mineral nutrients. Indeed, the perception of soil as the ultimate control on the development of plant communities has been a central paradigm of plant ecology for at least a century.

In a parallel development, the role of plants in shaping soils was recognized with the beginnings of soil science (Dokuchaev, 1879; cited in van Breeman, 1995). Hans Jenny, in his classic work "Factors of Soil Formation" in 1941, identified plants and animals as one of the primary forces in the genesis of soil from bedrock, and acknowledged that soil scientists extending back to the nineteenth century had perceived plants and soils to be a "coupled system." Conversely, the critically important role of plants in causing soils to form from weathered and unweathered rock was elaborated in detailed studies of primary succession on newly exposed substrate by Crocker and Major in 1955. In the past several decades, the multiple forms of interaction between plants and soil have become a major focus for research.

Plant–soil interactions play a critical part in human interactions with the biosphere. Humans create plant communities ranging in complexity from single species of crop plants in agriculture to attempts at the restoration of complex communities; the choice of plants is conditioned by soil qualities and alters those qualities. Human-caused changes in atmospheric chemistry – the addition of nutrients such as nitrogen and sulfur as pollutants, the increase in CO_2 concentration – directly and indirectly affect plants, soils, and the interplay between them. Changes in climate, driven by these atmospheric changes, have similar effects. Human-caused extinction of species, the introduction of species into novel habitats, and the spread of weedy species change plant community composition, thus driving changes in soil. Conversely, the widespread degradation of soils from erosion and overgrazing constrains the kinds of plant communities that can develop, often to the detriment of human society. These are but a few examples of the profound importance of plant–soil interactions in the management, exploitation, and conservation of biodiversity.

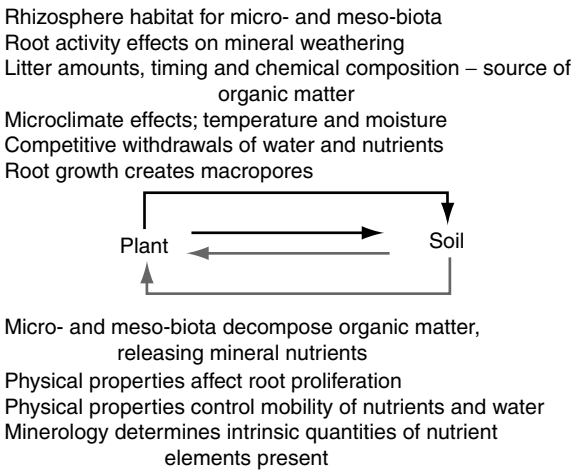


Figure 2 Feedback processes in plant–soil interactions. Changes in plant growth, physiology, and so forth, and changes in plant community composition can alter soil properties, which in turn will drive further changes to the plants. Conversely, soil properties that constrain the patterns of plant growth, physiology, and community composition will be affected by the plants that grow under those conditions. Positive feedbacks will cause the soil–plant system to evolve in a particular direction, unless deflected by a disturbance external to the system. Conversely, negative feedbacks will maintain the system in a constant state.

Small-Scale Processes: Micron to Millimeter Scales

The Rhizosphere: Structure of the Root/Soil Interface

Plants contact soil most intimately along the surfaces of their roots. The soil in direct contact with the surface of nonwoody roots, the "rhizoplane," differs in biological, chemical, and physical characteristics from the "bulk soil" (**Figure 3(a)**). The chemical and physical influence of the root surface on the soils decreases rapidly with distance from the root surface; thus, there is a cylinder of "rhizosphere" soil that is influenced by the root but not directly in contact with it. Its width varies with the texture and mineralogy of the soil and the diameter and physiological activity of the roots, being from <1 to several millimeters. The rhizosphere changes longitudinally along the root with distance from the root tip, root age, physiological stress to the plant, and degree and type of mycorrhizal infection (**Figure 3(a)**).

The rhizosphere is most strongly affected by rhizodeposition, the transfer of organic substances from the root into the rhizosphere. This material includes whole cells sloughed from the root cap and the epidermis, secretions of mucilage from these cells, and soluble materials leaked from intact cells and cracks caused by emerging new lateral roots ("exudate"). These materials mix with polysaccharide secretions of the adhering bacteria. The mucilage, or "mucigel," lubricates the root surface as it pushes past soil particles and helps to maintain the physical continuity between the root and the bulk soil.

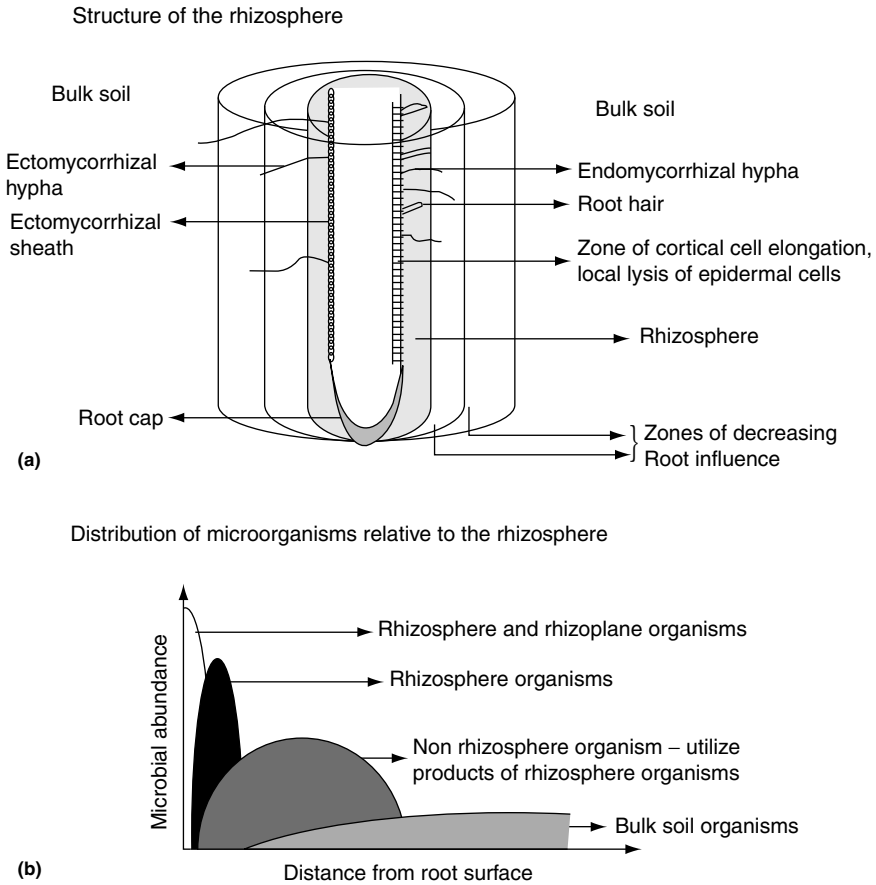


Figure 3 (a) A schematic diagram of the rhizosphere, illustrating different zones of the root tip (root cap, zones of cell elongation, and epidermal cell lysis), different types of mycorrhizal fungal growth (endomycorrhizal and ectomycorrhizal), and rhizosphere soil. The diagram also illustrates the zone of decreasing root influence with distance from the soil. (b) Distribution of microorganisms with distance from the roots, reproduced from Bazin *et al.* (1990) in Lynch JM (ed.) *The Rhizosphere*. New York: John Wiley and Sons; different species of microorganisms utilize substrates directly derived from roots or secondary materials produced by other microbes in adjacent zones.

In addition, the mucigel may adsorb nutrient ions, especially cations, and prevent their removal from the root surface by leaching. **Table 1** illustrates the diversity of compounds released from various species. The relative proportions of each class of compound may vary with position along the root; for example, amino acids are released in higher amounts near the root tips than farther back.

Studies in a variety of plants have shown that approximately 10–20% of the carbon fixed during photosynthesis is released to the soil through these mechanisms. This rhizodeposition accounts for 40–90% of the total amount of carbon transferred to the roots from the shoots; the remainder is respired by the root tissues or is used to construct new root biomass. Commonly cited values of 10–100 mg g^{-1} root dry weight for soluble substances and 100–250 mg g^{-1} root for insoluble materials give some indication of the amounts released. The amounts of carbon transferred from the shoot to the root system and the fraction of that carbon that is released to the soil vary greatly with plant species and soil conditions (**Table 2**), mycorrhizal infection, soil microbial abundance, and soil physical conditions. The total amounts of carbon involved in the transfer can be quite large (an estimation is 1.2 t C ha^{-1} $year^{-1}$ to 7.5 t C ha^{-1} $year^{-1}$).

The effects of this input of carbon and nitrogen to the rhizosphere is a large increase in the density, diversity, and biomass of the microbiota in the rhizosphere, compared to the bulk soil. Cell counts of microbes in the rhizosphere are usually in the range of $10^{10} - 10^{12}$ cells cm^{-3}, compared to 10^9 cells cm^{-3} in bulk soil. **Figure 3(b)** illustrates a typical distribution of microbes with respect to the root surface. Rhizosphere microbial cell densities: bulk soil cell densities range from <1 for microalgae to >24 for many bacteria; for most microbial groups, ratios are >5. Successions of different microbial species occur (a) over time, as newly produced root tips extend into the soil and are colonized, (b) over the length of the root (from root cap to zones of suberization and secondary thickening), and (c) with distance out from the root surface. Not only is the population of microbes in the rhizosphere larger than in the bulk soil, but it differs in composition. Bacteria are more abundant than fungi, whereas the reverse is often true in the bulk soil, and more species of bacteria may be found in the rhizosphere. Several genera of bacteria capable of fixing atmospheric N_2, including *Enterobacter, Klebsiella, Azotobacter, Azospirillum, Bacillus, Pseudomonas*, and *Clostridium*, are only found in the rhizosphere; these nonsymbiotic bacteria may (but not always) contribute to

Table 1 Compounds identified in root exudates, classified by general type[a]

Type	*Compounds*
Simple sugars	Hexoses (glucose, fructose, galactose), pentoses (arabinose, xylose, rhamnose), di- and oligosaccharides (wide variety), uronic acids, amino sugars
Amino acids	$\alpha\gamma$-Alanine, aspartic acid, phenylalanine, glutamic acid, leucine, serine, threonine α- and γ-Aminobutyric acids, α-aminoadipic acid, many others
Aliphatic acids	Formic, acetic, oxalic, citric, malonic, tartaric, malic, succinic, lactic, palmitic, stearic, oleic, linoleic, linolenic, others
Aromatic acids	p-Hydroxybenzoic, ferulic, o-cumaric, gallic, vanillic, sinapic, shikimic, trans-cinnamic
Complex carbohydrates	Polygalacturonic acid, other hydroxy-, keto-, di- and tri-carboxylic acids
Enzymes	Several phosphatases, peroxidase, invertase, urease, β-glucosidase
Vitamins	B group, biotin, pantothenic, nicotinic acids, thiamin, pyridoxine, riboflavin
Proteins and peptides	Unspecified
Plant hormones	Auxins, gibberellins, kinetins, ethylene
Other	Isoflavonoids, thiophenes, benzofurans, terpenes, lectins, phytoallexins, alcohols

[a]Commonly occurring and representative specific compounds are listed; in all cases, a very large number of specific compounds have been identified in each chemical class. *Source*: Reproduced from Vancura V and Kunc F (eds.) (1988) *Soil Microbial Associations – Control of Structures and Functions*. New York, NY: Elsevier.

Table 2 The range of values observed for the fate of photosynthetically fixed carbon in the root system

Plant type	*Percentage of total C fixed transferred to root*	*Percentage of root C in root transferred to soil*	*Percentage of root C respired*
Wheat	27–59	4–29	4–79
Barley	27–54	21–31	6–60
Maize	28–36	18–52	33–50
Pea	44–65	29–43	55–74
Yellow poplar trees	40	–	–
Douglas fir trees	73		

Source: Reproduced from Lynch JM (ed.) (1990) *The Rhizosphere*. New York: John Wiley and Sons.

greater plant growth. Species of *Pseudomonas* are particularly important, as many have been found to benefit plant growth. The composition of the rhizosphere biota changes rapidly (within months) in response to changes in the species composition of the plant community, demonstrating the close coupling of the rhizosphere community to the species-specific morphology and physiology of root systems. Variation in the abundance and diversity of plant pathogens also reflects the differences in rhizodeposition among plant species.

Structural complexity in the zone of root/soil contact is provided by extensions of the root surface into the soil matrix. Most plants produce extensive root hairs from cells just distal to the root cap and zone of elongating cells. The zone of root hair development extends only a short distance along the root (1–4 cm), as the epidermal cells producing the hairs are sloughed off in more basal areas (**Figure 3(a)**). Production of root hairs is itself affected by soil chemical conditions (e.g., pH, ion concentrations), texture, and water content. This zone provides increased structural complexity (e.g., a greatly expanded root surface area that is in direct contact with soil particles). Structural complexity also results from mycorrhizal infection, particularly by ectomycorrhizae, as fungal hyphae spread out from the root surface into the soil (**Figure 3(a)**). Ectomycorrhizal fungi, in particular, produce large quantities of extraradical mycelium; hyphae may extend for several meters from the host plant. These hyphae greatly expand the contact zone between the root and the soil.

Nutrient and Water Uptake: The Function of the Root/Soil Interface

Nutrient supply to plants, at the scale of the root surface and the immediately adjacent soil, is a complex function of the chemistry of individual nutrient ions. The availability of nutrients other than nitrogen is ultimately determined by the chemistry of the soil minerals present, which in turn reflects the geology of the parent material. Weathering releases these nutrients from the crystalline lattice of the minerals into solution, making them available for uptake; however, both physical and chemical weathering rates are themselves strongly affected by vegetation. Roots penetrate cracks and fractures in bedrock, widening them, allowing water to penetrate, and promoting chemical weathering. Chemical weathering depends on the supply of protons for the dissolution and transformation of soil minerals. Plant-derived sources supply most of the acidity involved in weathering; these include CO_2 from root and microbial respiration, which forms H_2CO_3 in the soil solution, organic acids in root exudates and the products of microbial decomposition of plant tissues, and hydrogen ions released from roots in exchange for nutrient cations. Soil genesis also results from the chelation and mobilization of iron and aluminum by the organic acids. As different species of plant generate different types and amounts of organic acid formation, the composition of the plant community strongly affects the nature of the weathering and pedogenic processes that take place. Thus, processes at the scale of the root surface result in broad-scale patterns of correlated vegetation and soil development.

Plant-mediated weathering and pedogenic reactions generate a supply of nutrient cations; their availability for uptake depends on a different set of intersecting plant–soil processes. This includes the diffusion rates, exchange, adsorption, and precipitation chemistry of each element, patterns of water movement through the soil, the morphology of the root system, the amount and diversity of mycorrhizal fungi, and the physiology of uptake by the roots. Nutrients are brought to the

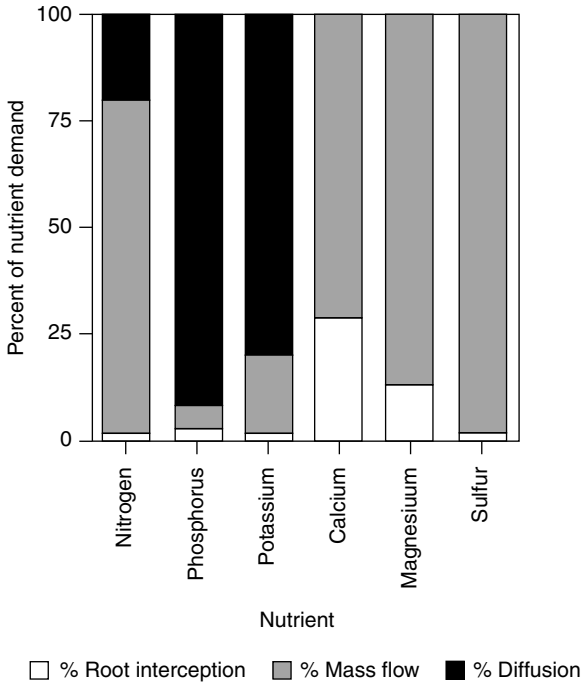


Figure 4 The relative importance of different aspects of root activity in the uptake of several nutrients. Reproduced from Jungk in Waisel Y, Eshel A, and Kafkafi U (1991) *Plant Roots – The Hidden Half*. New York: Marcel Dekker Inc., with permission from Marcel Dekker Inc.

root surface in the mass flow of water and through diffusion; roots encounter nutrients as they grow through the soil. The relative importance of each pathway varies with the different elements (**Figure 4**). The availability of sparingly soluble elements, such as iron and manganese, may be enhanced by siderophores in the exudate. These are nonprotein-forming amino acids, which complex with these metals, facilitating their mobility and uptake. Organic acids help to mobilize metal nutrients by lowering pH, and help to solubilize P by preferentially binding to the metals. For diffusion-limited nutrients such as P, zones of depletion develop within the rhizosphere (**Figure 5(a)**); this does not occur with highly mobile ions such as NO_3^- (**Figure 5(b)**). The size of the depletion zone reflects the balance of supply (diffusion and mass flow) and demand (uptake rate).

Unlike the nutrient cations derived from soil minerals, the availability of nitrogen depends on the microbially mediated decomposition of plant-derived organic matter. It is often assumed that the supply of readily metabolizable carbon from root exudate will promote increased nitrogen mineralization because it relieves carbon limitation, this is not always the case. Theoretical studies have suggested that the net effect of exudates on nitrogen dynamics will depend on the C:N ratio of the exudates relative to that of the indigenous organic matter of the soil, so that balance between the release of N through mineralization versus immobilization in new microbial biomass will vary with the quality of the exudates. Several studies with herbaceous plants have found that the presence of rhizosphere microbes can increase availability of N to the plants, but the effect varies with the availability of N from the soil and the kind of plant grown. Alternatively, N

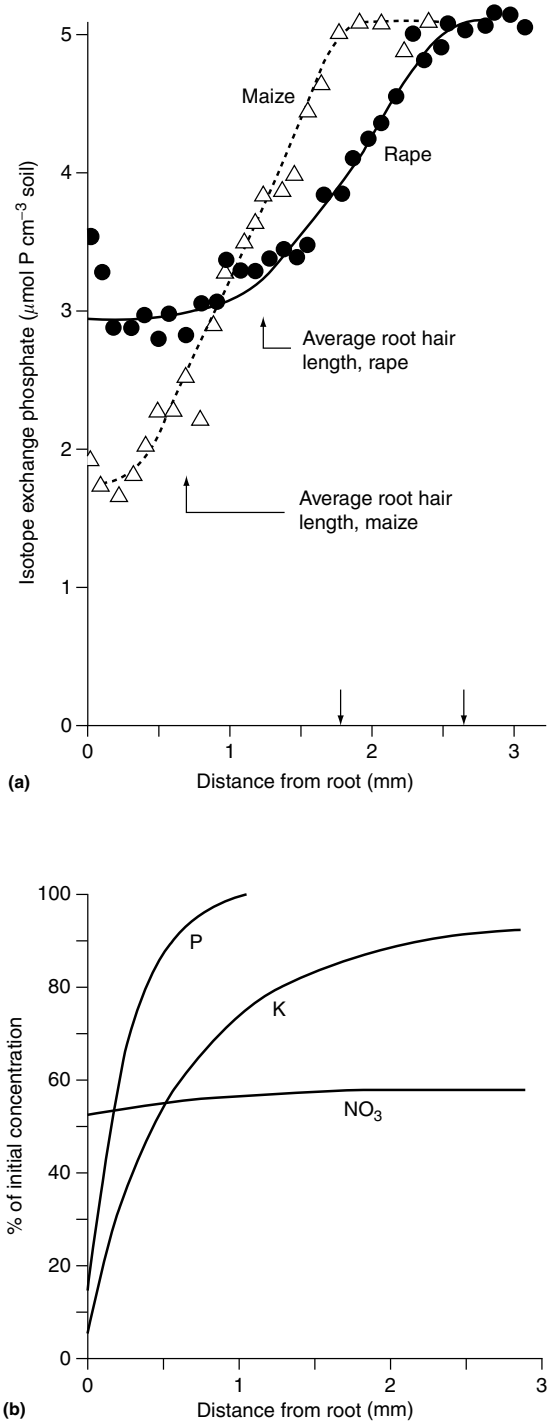


Figure 5 (a) Observed patterns of P depletion with distance from the surface of roots of maize and rape; the zone of low P concentrations adjacent to the root is an indicator of the distance that root hairs grow outward from the root surface. (b) Calculated depletion curves (as percent of bulk soil concentration) of three nutrients; the depletion rate out from the root surface depends strongly on the chemical mobility (combination of diffusivity and likelihood of adsorption, precipitation, or other immobilizing chemical processes). Reproduced from Jungk in Waisel Y, Eshel A, and Kafkafi U (1991) *Plant Roots – The Hidden Half*. New York: Marcel Dekker Inc., with permission from Marcel Dekker Inc.

may be immobilized within the microbial communities in response to root exudation. Thus, the effects of root exudation on the supply of mineral N to plants is a complex function of the kind of plant, the nature of the organic matter in the soil, and the amounts and composition of the exudate. Root exudate may decrease or increase N availability, even in different horizons of the same soil.

Not least among the effects of roots on the rhizosphere soil is the alteration of soil pH. In addition to the excretion of organic acids and CO_2, roots exchange protons and hydroxyl ions for nutrient ions, particularly ammonium (NH_4^+) and nitrate (NO_3^-). Because nitrate is rapidly taken up when it is present, a net imbalance of anion and cation uptake occurs. Maintenance of electroneutrality during this requires that an anion be extruded in exchange for the nitrate ion; this anion is usually hydroxyl (OH^-). Conversely, hydrogen ions (H^+) are excreted during uptake of ammonium. Soil pH can rise as much as two units within the rhizosphere when nitrate uptake is rapid, whereas it may decrease by up to two units when all uptake is in the form of ammonium.

Physical Structure of Soil

Plants modify the physical arrangement of soil particles in a variety of ways. In most soils, the grains of mineral material and the soil organic matter are bound together in aggregates. In the rhizosphere, the mucigel produced jointly by the plant and the resident microbes bind clay minerals together into microparticles of about 50 μm diameter. Fungal hyphae can bind together these microparticles into larger aggregates of 1–2 mm diameter. The products of root growth (and root-supported mycorrhizal fungal growth) are especially important in stabilizing the aggregates over time. Moreover, different plant communities create different quantities and sizes of soil aggregates, depending in part on the morphology of the root system and the extension and mass of mycorrhizal hyphae.

Pore size distributions are also affected by roots through several other mechanisms. Alternate wetting and drying of the soil induced by transpiration can affect cracking patterns, especially in clay-rich soils. Roots also preferentially grow into larger pores, or zones of lower penetration resistance, and may thereby enlarge them. The death and decay of roots also creates macropores or channels of highly porous organic matter; these large voids provide preferential pathways for the flow of water. Finally, the degree of contact between the root surface and the soil varies with root age, amount of exudation, and root physiological status (e.g., hydration).

Roots also lose molecular oxygen to the soil, as air diffuses from leaves and stems through lacunae in the cortex tissues and through cracks in the surface of the root epidermis. This "radial oxygen loss" is particularly large in wetland plants, in which the root cortex develops extensive air spaces ("aerenchyma") through the degradation of cell walls. Root oxygen loss can be sufficiently high to alter the reduction–oxidation potential of the surrounding bulk soil (**Table 3**). As a result, a zone of 2–3 mm of oxidized soil (often apparent as a coating of orange–red iron oxide grains) may develop around the root surface under anoxic, waterlogged conditions. This zone of oxidized iron minerals acts as a trap for other metal ions in the soil, such as cadmium, copper, lead, and zinc, which become enriched in the surface concretions as much as 10-fold greater than the bulk soil concentrations. Moreover, by maintaining a higher redox potential in the soil, denitrification and methanogenesis may be inhibited; this both decreases the presence of toxins (methane) adjacent to the root surface and also alters the net carbon and nitrogen balance of the soil.

Several species of wetland plants, including the yellow water lily (*Nuphar lutea*) and common reed (*Phragmites australis*) take advantage of a thermal diffusion mechanism to develop positive air pressures in young leaves and drive a mass flow of air into the roots. This air can effectively oxygenate long rhizomes (>2 m long in *Phragmites*) plus the surrounding soil. Only a few wetland plants employ mass flow to move air through the roots and back to the leaves; the rest rely on diffusion of oxygen through the aerenchyma. In most plants, however, the aerenchyma also provides a pathway for

Table 3 Effects of roots on reduction–oxidation potential (Eh, mv) of wetland soils, mediated by root oxygen loss

Plant species	*Growth form*	*No plant*	*With plant*
[a] *Acer rubrum*	Tree	+280	+320
[a] *Rosa palustris*	Shrub	+220	+240
[b] *Salix viminalis*	Shrub	+200	+330
[a] *Saururus cernuus*	Herb	+295	+310
[c] *Myriophyllum verticillatum*	Herb	+170 to +200	+270 to +280
[d] *Isoetes lacustris*	Herb	+600	+50 to +200
[e] *Ranunculus circinatus*	Herb	−200	+400
[f] *Myriophyllum tenellum* + *Isoetes braunii*	Herbs	−100	+190
[g] *Spartina alterniflora*	Herb	−50	+150

For each study, the maximum observed difference between planted and unplanted soils are reported.

[a] Havens (1997) *Wetlands* 17: 237–242.

[b] Grosse W (1997) In: Rennenberg, *et al.* (eds.) *Trees.* The Netherlands: Backhuys Publication.

[c] Carpenter S, Elser J, and Olson K (1983) *Aquatic Botany* 17: 243–249.

[d] Wium-Andersen S, and Andersen JM (1972) *Limnology and Oceanography* 17: 948–952.

[e] Flessa H (1994) *Aquatic Botany* 47: 119–129.

[f] Jaynes ML, and Carpenter SR (1986) *Ecology* 67: 875–882.

[g] Howes BL, Howarth RW, Teal JM, and Valiela I (1981) *Limnology and Oceanology* 26: 350–360.

venting methane produced in anoxic soil to the surface; plant stems account for >90% of the methane release to the atmosphere from waterlogged soils.

Soil physical properties, including texture, bulk density, temperature, and moisture-holding ability, affect the number, size, diameter, and branching patterns of roots in a great diversity of ways. In many plants, clayey soils inhibit root growth compared to sandy soils. This reflects the ability of the roots to grow into pores between soil particles and to penetrate aggregates of particles. Mechanical impedance, reflected in the soil's bulk density, affects all aspects of root growth, root system morphology, and uptake functions. Temperature is another factor affecting root growth; optimal temperatures vary among species and are usually lower than the optimal temperatures for shoot growth. Aeration and moisture status, which are inversely related to each other, are jointly controlled by climate and by the transpiration rate of the plant community. Temperature, aeration, and moisture content also affects the morphology of the root system (root diameters, density of root hairs, and frequency of branching). Not surprisingly, texture, mechanical impedance, temperature, moisture, and aeration are interrelated (as they all are affected by particle size distribution), and their effects on root growth are interactive.

Mesoscale Processes: Centimeters to Meters

The microscale interactions between root surfaces and the adjacent soil are part of a larger set of interactions that occur over scales of centimeters to meters. Strategies of resource allocation by whole plants, including changes in growth rate, differential growth of above-ground and below-ground tissues, changes in root longevity and turnover, and physiology (i.e., in changing metabolic capacity for nutrient uptake) affect the vertical and horizontal extent of root penetration through the soil, thereby affecting the spatial extent over which microscale plant–soil interactions occur. At this larger scale, the deposition of litter from the above-ground tissues becomes as important as the patterns of root growth within the soil; indeed, litter input may dominate the biological processes within the soil. However, spatial variation in soil properties, driven by topography, bedrock geology, and other exogenous factors, also affects root system morphology and growth, and condition the fate of the litter.

Nutrient Acquisition and Root Growth: Individual Plants

The acquisition of nutrients and water is strongly affected by the architecture and size of the root system, a function of the allocation of photosynthate below ground. Conversely, the availability of nutrients and water molds the growth of the root system, thus setting up a feedback interaction. The metabolic capacity to absorb nutrients not only varies among plant species, but also varies within an individual as the availability of nutrient elements changes. Nutrient uptake capacity is often inversely related to availability, because transport systems in the root cell membranes increase in activity as nutrients become more limited in supply. For example, barley plants growing in nitrogen-limited cultures have uptake rates of both ammonium and nitrate more than 200% greater than plants growing with excess N; similarly, phosphate uptake rates increase by 400% in phosphorus-limited cultures, and sulfate uptake rates increase by almost 900% in sulfur-limited culture. Similar inverse relationships between availability, as measured in soil solutions or extracts, and short-term uptake rates have been found for a wide range of wild plants, including both herbs and trees. However, while relative uptake capacity varies within a species inversely with the availability of nutrient elements, absolute rates of uptake of plant species from low-nutrient environments may be less than, equal to, or greater than plant species found in high-nutrient environments. Thus, the effects of nutrient availability on root physiology depend on both the species of plant and the variation present in the soil.

Nutrient acquisition is also affected by the abundance and distribution of the roots. Different kinds of plants are characterized by differences in the relative allocation of biomass to roots and shoots (**Figure 6**). In general, grasses tend to have higher relative allocation to roots than do woody plants, although trees and shrubs in highly nutrient-limited environments, such as alpine tundra or bogs, have higher root:shoot ratios than do woody plants from more nutrient-rich environments such as tropical or temperate deciduous forests.

Root systems also adjust to varying nutrient availability by changes in morphology (length, diameter, branch density, etc.). Root length density (the length of roots per unit volume of soil) commonly increases in zones within the soil that have higher concentrations of limiting nutrients, as has been frequently demonstrated in experiments (**Figure 7**). However, such responses are not seen in all species, and the response to variations in nutrient availability depends on a plant's tendency to produce thick versus thin and fast-growing versus slow-growing roots. Increases in root length in one area, in response to a patch of high nutrient availability, are often compensated for by decreases in other parts of the root system. Furthermore, nutrient concentrations in plant tissues do not necessarily increase, and plant growth also often does not show an overall increase, suggesting that plants compensate for increased growth in one part of the root system in response to increased nutrient availability by decreasing growth elsewhere.

The response also depends on the plant's tendency to produce herringbone versus dichotomous patterns (**Figure 8**). The optimal response depends in part on the plant's ability to allocate carbon to the root system (versus shoots, leaves, or reproductive structures), the size of an area of increased nutrients, and the chemistry (i.e., the diffusivity) of the nutrient at stake. Research by A. Fitter has shown that herringbone patterns are most efficient for mobile nutrient ions, but herringbone and dichotomous patterns are equivalent for immobile ions. Others have shown that herringbone patterns are more common in roots growing in infertile soils; the lower amount of branching allows the root system to explore a larger volume of soil. In fertile soils and in nutrient-rich microsites, branching patterns become more dichotomous; this pattern presumably permits more intensive and complete exploitation of the soil resource.

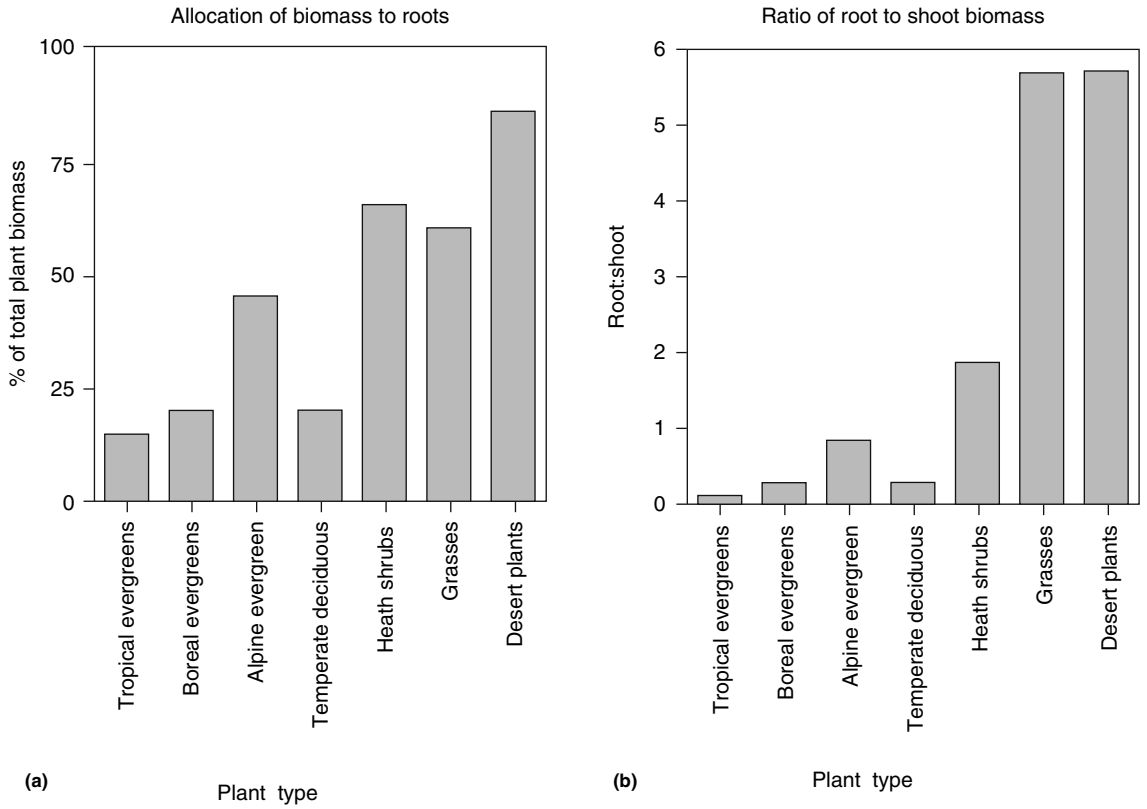


Figure 6 The relative allocation of biomass to roots, as (a) the percentage of total plant biomass and (b) with respect to shoot biomass, for plants from various habitats. Reproduced from Klepper in Waisel Y, Eshel A, and Kafkafi U (1991) *Plant Roots – The Hidden Half*. New York: Marcel Dekker Inc.

The heterogeneity of root systems in response to soil conditions reflects the high degree of heterogeneity in soil conditions at scales of centimeters to tens of meters. In agricultural fields, for example, chemical and physical soil properties vary significantly between points that are only tens of centimeters apart, despite the apparent uniformity that might be expected from the nearly flat topography, homogenization of the soil from plowing, and planting of a single species of crop plant (**Figure 9**). Extreme patchiness in one soil resource may not be paralleled by similar patchiness in other resources in the same site; the distribution of mobile, limiting elements like nitrate can vary greatly among points only 3 cm apart whereas immobile, nonlimiting elements (P, K) do not vary, even over large distances.

In natural communities, significant variation in nutrient concentrations and nutrient supply rates (e.g., nitrogen mineralization rates) have been documented in soils beneath different co-occurring species of grasses and trees. Spatial variation in soil properties is also generated by topographic variation. Even small changes in elevation or geomorphic position (e.g., hill crest, upper and lower slope positions, valley bottoms) are associated with differences in soil profile development, accumulation of organic matter, soil texture, the mineralization rates of nutrient elements, and gaseous losses through microbial processes such as denitrification. At scales of tens to hundreds of meters, differences in plant community composition are clearly correlated with differences in soil properties related to topography.

Spatial heterogeneity of soils is particularly strongly developed in plant communities in which either plants are patchily distributed over otherwise barren ground (typical in many arid ecosystems) or in which differing life-forms (e.g., trees and grasses) are interspersed (e.g., savanna ecosystems). In deserts, the scattered plants form "islands of fertility" because of their ability to trap both plant litter and wind-blown fine soil particles. The soils within and beneath the scattered shrubs or tussock grasses in these systems have higher amounts of nitrogen, phosphorus, organic matter, silt or clay-sized particles, and moisture than the intervening bare areas. Similarly, patches of woody vegetation within grasslands accumulate higher amounts of nutrients and organic matter, again leading to patches of higher fertility within the grassland landscape. Plasticity of root system morphology and physiological capacity allows plants and communities to respond to this spatial variation in soil properties.

Nutrient Acquisition and Root Growth: Stand Scale

The response of individual plants to variation in soil properties manifests itself at the level of the community as changes in root biomass and the production and mortality rates of the roots. Many studies have shown that root biomass tends to be lower in more nutrient-rich sites, and experimental studies in which nutrients (particularly nitrogen) are added often confirm these observations. Similarly, it has been noted that

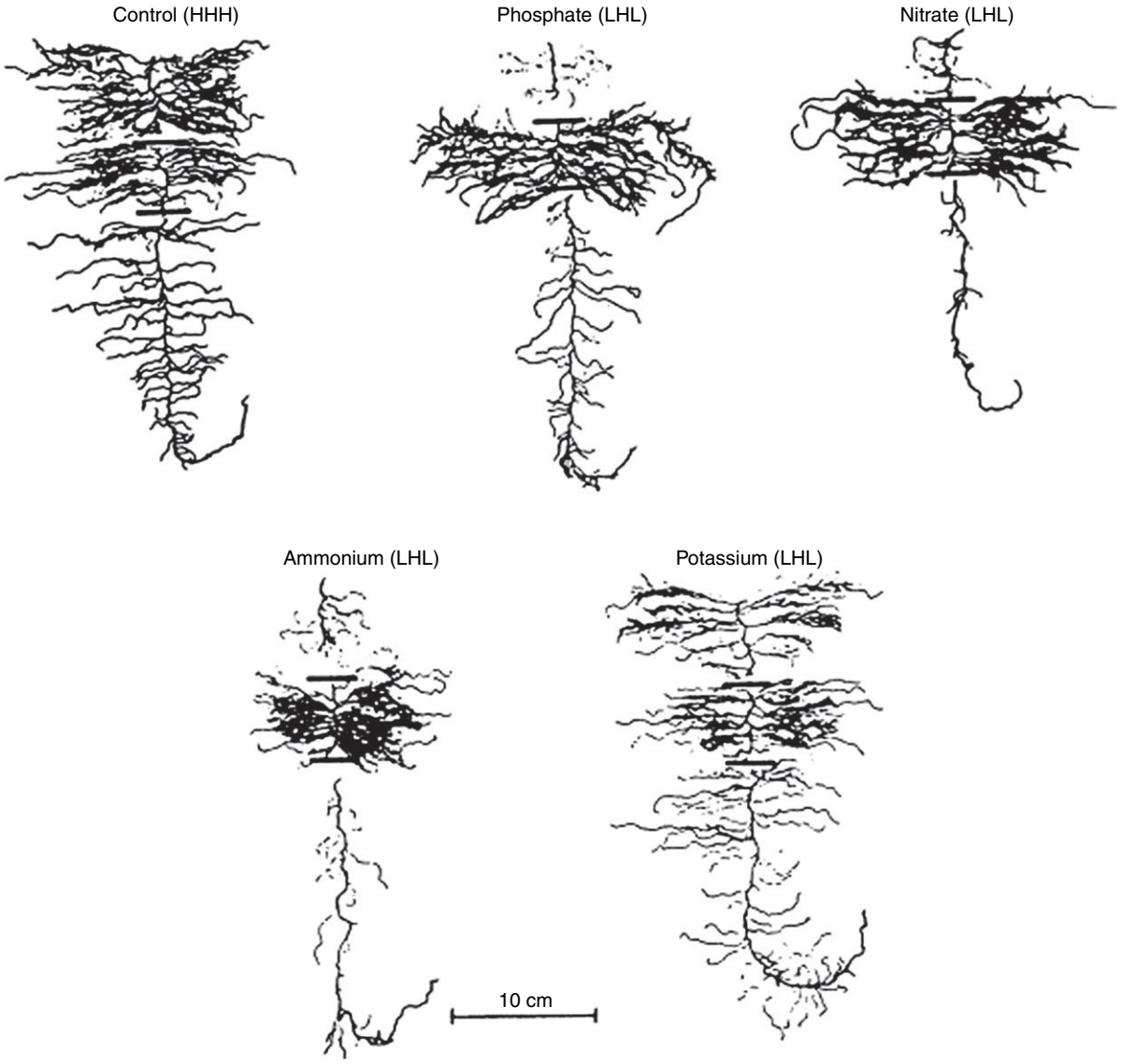


Figure 7 The results of a classic experiment by Drew in which the roots of barley plants grew through zones of nutrient-enriched soil. Control plants received all nutrients throughout the root system; the other plants received high concentrations of the target nutrient only in the middle of the root system. Reproduced from Drew MC (1975) Comparison of the effects of a localized supply of phosphate, nitrate, ammonium and potassium on the growth of the seminal root system, and the shoot, in barley. *New Phytologist* 75: 479–490.

plants typical of infertile sites have a higher ratio of root biomass to above-ground biomass than do plants from fertile sites. In a comprehensive review of available literature on nutrient cycling and root production, Vogt *et al.* (1986) found that the mass of fine roots in forest ecosystems is higher in forests with larger amounts of nitrogen in the forest floor, implying that the biomass of roots increases when nitrogen is sequestered in compounds that decompose slowly and accumulate in thick forest floors. This supports the idea that root biomass is greater in infertile than in fertile soils.

Rates of production and mortality of fine roots also respond to nutrient status, although experimental studies have documented both increases and decreases in response to nutrient additions. For example, Vogt *et al.* (1986) suggested that the rate of root turnover (input of new roots per year) is higher in forests with smaller amounts of nitrogen input in litterfall (an index of nitrogen limitation), but experimental addition to N to northern hardwood forests caused a decrease in mortality rates of fine roots. The contradictory data may reflect differences between the "normal" growth patterns of roots relative to the inherent fertility of the soil and the response of existing root systems to large, sudden changes in fertility induced by experimental nutrient additions.

Competitive Interactions between Plants and Microbes

As the plant community and the soil microbiota require the same resources (nutrients and water), there may be competition between the microbiota and plants. Microbial

demand for nitrogen is closely tied to the availability of degradable carbon, as carbon is often considered to be their most limiting resource. In this context, the likelihood of competitive interactions with plants depends on the C:N ratio of the available carbon resources; root exudates tend to have low C:N ratios (about 15:1), litter normally has high C:N ratios (»30:1, and >100:1 for woody plant tissues), and soil organic matter ratios are in the range of 10–20. Theoretical analyses have suggested that the relative mineralization versus immobilization of N derived from root exudates varies in relation to the relative availability and quality of the soil organic matter, where "quality" refers to the metabolic capacity of the microflora to degrade the components of the organic matter. When N is limiting (i.e., when C:N ratios are high), microbes may take up mineral N from the same sources (soil solution, exchange surfaces) that plants use, resulting in competition.

Experimental additions of mineral N (NH_4^+, NO_3^-), to forest soils have shown that the microbiota captures a substantial proportion, often a majority, of the added N, but that nitrification rates are decreased when plants are present. These results suggest that while the microbiota compete successfully for mineral N in solution, plants can limit the availability of $NH4^+$ to the ammonia oxidizers (nitrifiers). However, long-term studies suggest that although microbes may outcompete plants for mineral N in the short term, the long-term activity of root systems, compared to the shorter, more localized bursts of activity of the microbes in response to changing temperature and moisture regimes, may allow plants to compete successfully. In this regard, the influence of plants on the soil moisture regime, through uptake of water and loss in transpiration, may indirectly affect competitive interactions for N by causing limitation of microbial activity due to soil drying. Furthermore, competition for N may be modified by differential use of the inorganic and organic forms of N in the soil.

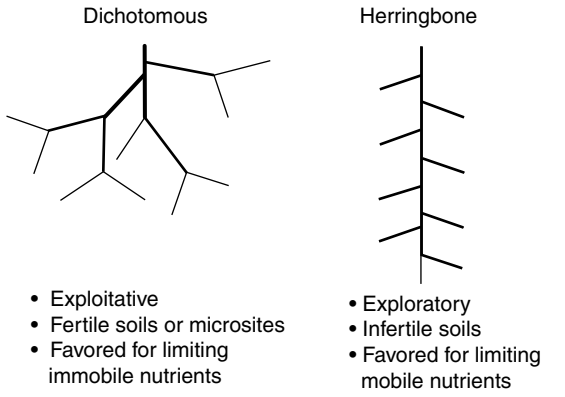


Figure 8 Alternative patterns of root growth. Dichotomous branching systems have more internal segments of the root system, compared to herringbone branching systems. The internal segments, or links, tend to be larger in diameter and therefore are more "expensive" to construct and maintain (in units of carbon fixed during photosynthesis). Plastic responses of roots to heterogeneous soil conditions involve changes in branching pattern, branch number, and length, branching angle, root diameter and tissue density (observed as changes in specific root length and area, or length and area per unit root mass), and density of root hairs. Reproduced from Fitter A in Waisel Y, Eshel A, and Kafkafi U (1991) *Plant Roots – The Hidden Half*. New York: Marcel Dekker Inc. and other papers.

Litter Inputs to the Soil: Effects of Plant Tissue Chemistry

Perhaps the most widely studied and best-known interaction of plants and soil is the supply of organic matter to the soil through the deposition of litter. Litter deposition includes "fine litter," the leaves, herbaceous stems, and smallest woody twigs, "coarse woody debris," the large twigs, branches, and boles of shrubs and trees, "greenfall," of leaves and leafy twigs clipped by herbivores such as squirrels, and root litter, the

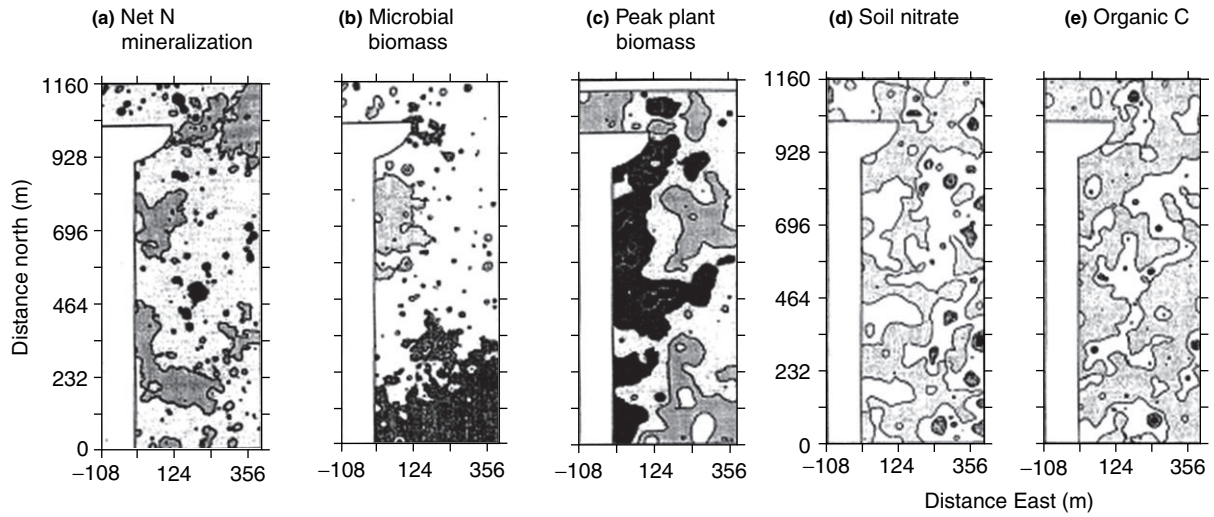


Figure 9 Spatial patterning of physical and biological soil properties in a soybean field in Michigan (US). Units: (a) net N mineralization, mg N g^{-1} d^{-1}; (b) microbial biomass, mg C g^{-1} soil; (c) peak plant biomass, g m^{-2} (d) nitrate, mg N g^{-1} soil, (e) elevation, m. Reproduced from Robertson GP, Klingensmith KM, Klug MJ, Paul EA, Crum JC, and Ellis BG (1997) Soil resources, microbial activity, and primary production across an agricultural ecosystem. *Ecological Applications* 7: 158–171, with permission form ESA.

dead roots of all size classes. However, most measurements of "litter" refer only to fine litter, as this is most easily measured. Litter is the largest source of organic matter supplied to the microbiota of the soil, and thus it fuels all of the microbial processes responsible for decomposition, nutrient cycling, and reduction–oxidation reactions of soil minerals and also is the ultimate source of the soil organic matter.

Litter inputs vary widely in quantity and timing. Tropical rain forests may have >2000 g m^{-2} year^{-1} litterfall, whereas tundra ecosystems may have <200 g m^{-2} year^{-1}. Deposition of fine litter is seasonal in temperate zones and continuous in the tropics, but the formation of coarse woody debris may be sporadic (i.e., large inputs following major storms, or infrequent inputs at any time of year). Amounts of coarse woody debris are comparable (approximately 120–3000 Mg m^{-2} year^{-1}) but much more patchily distributed. Greenfall production is also sporadic, as it reflects the population dynamics of the herbivorous insects or mammals that cause it. The timing of root death is quite variable among ecosystems; in some forests, root growth and death are pulsed, but in other sites, production and mortality may occur simultaneously and not in synchrony with seasons. Assuming that root populations are at steady state, so that root production rates can be used as an estimate of root litter deposition rates, the amounts of root litter production similarly range from roughly 100 g m^{-2} year^{-1} to 1,700 g m^{-2} year^{-1}.

The ratio of above-ground to below-ground litter inputs is also quite variable among ecosystems (**Table 4**). Root litter inputs can be very high in grassland ecosystems as well, with ratios of below-ground:above-ground litter mass much greater than 1. The variability of the relationship between above-ground and below-ground litter production was clearly demonstrated by Raich and Nadelhoffer (1992), who found no statistically significant relationships between the amount of fine root production and the amount of above-ground litter production over a wide geographic range of locations and forest types.

As important as the quantity and timing of input of litter is, the chemical composition of the plant tissues is perhaps more important as the ultimate control on the activity of the soil microbiota. The ability of microbes to decompose the litter material depends on (a) the ratio of carbon to nitrogen and (b) the chemical forms in which the carbon occurs. Carbon occurs in plant tissues in several forms, including "non-structural carbohydrates," which include soluble and membrane-bound contents of the cytoplasm and vacuoles. Cell walls contain cellulose (an unbranched chain of glucose residues), hemicelluloses (branched chains of glucose residues), and pectins; these materials are variably bound to lignin, a complex phenolic material that provides strength to the tissues. Because of its molecular structure, lignin is extremely resistant to decomposition; relatively few soil microorganisms have the requisite enzymes to degrade it. Other important carbon-containing compounds include polyphenols and tannins; these compounds act both as deterrents to herbivores and as inhibitors of decomposition. The nitrogen in plant tissues is largely found in protein; approximately 50% of the total leaf nitrogen may be in enzymes and structures associated with photosynthesis.

A large number of studies over many years have shown that the ability of the soil microbiota to decompose plant material depends on both the C:N ratio and the lignin:N ratio. The C:N ratio describes the quality of the substrate (e.g., litter material) relative to the demand by the microbiota for both energy and nitrogen. As the C:N ratio of the microbiota is lower (approximately 3–5 for bacteria and 10–12 for fungi) than any plant material (range from about 12 for leguminous herbs such as vetch and alfalfa to 400–600 for hardwood and coniferous sawdust, respectively), decomposition will be slowed in proportion to the energy/nutrient imbalance between the decomposers in a given soil and the particular substrate available. However, decomposition rates are also affected by both the absolute amount of lignin in the plant tissues and by the lignin:N ratio. **Table 5** illustrates typical results of a

Table 5 An example of the role of plant tissue chemistry in regulating decomposition rates

Plant	*%N*	*%Lignin*	*%C*	*C:N*	*k*
White spruce	0.52	13.1	47.0	89.4	−0.757
Douglas fir	0.61	20.5	46.7	76.6	−0.850
Balsam poplar	0.58	13.6	44.6	76.9	−0.907
Aspen	0.64	14.1	45.5	71.1	−1.351
Grass	0.81	15.2	45.8	56.5	−1.391
Dogwood	0.78	7.5	41.5	53.2	−1.706
Rose	1.15	3.4	44.5	38.7	−1.841
Cow-parsnip	1.31	7.0	39.6	30.3	−2.550

Tissues were allowed to decompose in laboratory microcosms for 4 months. Regression analyses of the data showed that the best predictors of mass loss were the C:N ratio, %N and lignin:N, in that order. Nitrogen content was more important early in decomposition; lignin became more important as decomposition proceeded.
Source: Reproduced from Taylor BR, Parkinson D, and Parsons WFJ (1989) Nitrogen and lignin content as predictors of litter decay rates: A microcosm test. *Ecology* 70: 97–104.

Table 4 Above- and below-ground inputs of organic matter to the soil in forest ecosystems

Ecosystem	*Root litter input (kg ha^{-1} year^{-1})*	*Above-ground input (kg ha^{-1} year^{-1})*	*Ratio*
Tropical broad leaf evergreen	n.d.	9438 ± 1104	–
Warm temperate deciduous	5732 ± 1970	9369 ± 790	0.61
Warm temperate evergreen	9053 ± 453	4432 ± 234	2.04
Cold temperate deciduous	2280 ± 920	3854 ± 213	0.59
Cold temperate evergreen	6152 ± 1007	3144 ± 194	2.91
Boreal evergreen	998 ± 208	2428 ± 204	0.41

Source: Reproduced from Vogt K, Grier CC, and Vogt DJ (1986) Production, turnover and nutrient dynamics of above and below ground detritus of world forests. *Advances in Ecological Research* 15: 303–377; averages of values reported for several studies for each community type.

comparative study of decomposition rates (k) of different plant materials, showing the close relationship between tissue chemistry and decomposition. Although studies differ in identifying C:N, lignin: N, nitrogen concentration or lignin concentration as the primary factor regulating decomposition rates, and in describing the relationship between decay rate and lignin:N as linear or nonlinear, almost all identify some subset of these characteristics as critical.

Decomposition not only produces CO_2, but also results in (a) the production of microbial metabolic products that are stable in the soil, and form the passive or long-term component of soil organic matter, and (b) the mineralization of nutrients such as nitrogen, sulfur, and phosphorus, which are bonded to carbon. Release of these nutrients into the soil solution is, again, a function of the availability of the nutrients in the plant litter substrate relative to the demand by the microbiota, as modified by the efficiency of assimilation of the microbes. Thus, the chemical composition of the plants that compose a community will affect the rates of accumulation of organic matter in the soil, as well as the rates of mineralization of limiting nutrients for plant growth. Alterations of plant community structure, through disturbance, succession, or management, can precipitate changes in nutrient cycling rates and carbon sequestration rates that create feedbacks to the plant community. For example, extensive browsing by moose on birch trees in boreal forests allows white spruce to grow up in their place; the switch from readily decomposable birch leaves to recalcitrant spruce needles causes decreases in the availability of nitrogen, which in turn favors the slow-growing spruce more than nitrogen-demanding birch. This positive feedback maintains spruce-dominated forests until stand-destroying fires permit birch to become re-established.

Interactions with the Soil Biota

Plant interactions with the soil biota are a crucial component of the processes that control litter decomposition, nutrient cycling, and plant community composition. Over the past decade, the understanding of plant–soil biota interactions has increased dramatically. The soil food web involves a large number of organisms, ranging from bacteria and fungi, to consumers over a wide range of sizes (nematodes at the microscale of microns to insect larvae at the macro scale of millimeters to centimetres.) The soil fauna includes organisms that not only are feeding on each other, but also herbivores and pathogens that are directly affecting plants. Plant inputs to the soil food web – root and shoot litter inputs – are often categorized by their chemical characteristics (*see* Litter Inputs to the Soil: Effects of Plant Tissue Chemistry), with more readily decomposed materials providing the base for a bacterially mediated food web, whereas more recalcitrant materials provide the base for a fungal-based food web. Each food web involves numerous kinds of organisms, and the food webs tend to be longer (more steps) than above-ground food webs, and highly multitrophic. Direct and indirect interactions between plants and the soil fauna characterize the wealth of interactions. Direct interactions involve the herbivores, pathogens, parasites that directly consume plant tissue, infect the roots, and thereby affect plant survival and growth. These organisms include pathogenic nematodes and fungi and insect larval root feeders as the two most prominent groups.

Plant pathogens exert important effects on the plant community through differential survival and growth of infected and uninfected plants. Numerous experiments, and agricultural practice, have shown that multiple generations of the same plant species in a given soil allows pathogen populations to build up, exerting a negative feedback on the plant community (or significant loss of plants in an agricultural planting of a single crop species). The pathogen load induced by training a soil with a given plant species has been widely explored as the mechanistic basis for many plant community changes. For example, nonnative plant species often become established in a region without their associated community of soil pathogens. Several studies of invasive herbs and forbs in grasslands, and one study of a woody plant invader, show that the reduction in negative plant–soil feedback resulting from the lack of pathogens and other soil faunal enemies can contribute to invasive success of the nonnative species. For example, the dune grass *Ammophila arenaria* supports generalist pathogens in its invaded ranges, but not the species-specific pathogens found in its native range and responsible for its decrease in abundance during normal dune succession. Similarly, as species from southern Europe invade communities in northwestern Europe, they experience less negative soil feedback (negative effects on growth after a first generation of plants has "trained" the soil by promoting associated plant pathogens) than do species native to northern Europe. Black cherry (*Prunus serotina*), an invasive species in Europe that is native to the US, also experiences greater effects of soil pathogens in its native soil than it does in soils in the invaded range, suggesting that the lack of negative feedback from the soil pathogen population facilitates the invasion.

Plant interactions with mutualists in the soil biota are as important in altering plant communities and providing plant–soil feedbacks as interactions with pathogens and parasites. Nitrogen-fixing bacteria are well-known symbionts, and nonsymbiotic nitrogen-fixing bacteria that are associated with plant roots can also be important. Both symbionts and associative bacteria increase the amount of available nitrogen both to the specific plants with which they are associated, and overall in the community, over time. The classic example of a dramatic change in plant communities due to nitrogen fixation is the invasion of *Morella faya*, a leguminous tree, which supports nitrogen fixation, unlike the native vegetation. The presence of the invasive tree and its symbiont results in an increase in N inputs from fixation from 0.2 kg ha^{-1} $year^{-1}$ to 18 kg ha^{-1} $year^{-1}$.

Mycorrhizal fungal symbionts also have complex interactions with plant communities. More than 80% of plant species have associated mycorrhizal fungi, which may be symbionts, nonobligate mutualists, or even function as parasites. Although most mycorrhizal fungi are thought to be nonspecific in their associations with particular plant species, the identity of particular fungal species can affect both plant performance, and the effects of the mycorrhizal infection on community dynamics. A number of studies have shown that the identity and diversity of arbuscular–vesicular mycorrhizae affects the composition and diversity of the plant community, with increases in plant diversity being associated with the

presence and diversity of mycorrhizal fungi. However, not all studies find such results; the effects of mycorrhizal fungi on plant community structure can vary with soil conditions and fertility, and the identity of both plant species and fungal species.

Plants also interact indirectly with soil organisms that are not parasites, pathogens, or mutualists. These organisms include the vast bulk of bacteria, fungi, nematodes, protozoa, mites, collembolan, enchytraids, earthworms, and other organisms that make up the soil biota. The primary pathway of interaction is the supply of decomposable substrate to this biota, both through dead root inputs and above-ground litter inputs. Both the quantity and the chemical quality of this organic input have a profound influence on the function, biomass, and composition of this fauna, as has been shown in numerous studies. Conversely, the mineralization of nutrients by this biota provides the major feedback to the plant community. The supply of mineral nutrients, and the form in which they are supplied, determine plant community composition, plant productivity, and interspecific interactions among the plant species. Plants that produce recalcitrant litters (high C:N ratios, high lignin content, high lignin:N ratios, high polyphenol content) tend to promote fungal-dominated soil microbial communities, with fungal:bacterial ratios approaching 1.0, slow rates of mineralization, organic forms of nutrients (e.g., amino acids for nitrogen, organophosphorus compounds for phosphorus) present, and a soil fauna that is dominated by fungal-feeding organisms. These food webs are associated with slow plant growth, low net primary production, and species that are poor competitors. In contrast, plants producing labile litter (low C:N, low lignin, low polyphenols) stimulate a bacterial-based food web, with species that are primarily bacterivores as the primary consumers. The macrofauna also tends to differ, with earthworms less common in acidic, fungal-dominated humus materials of slowly-decomposing litters and more common in the less acidic, bacterial-dominated richer soils of rapidly decomposing species. Indeed, the chemical and physical characteristics of plant litter has proven to be more important than climate or any other source of influence on the activity of the soil biota that carry out decomposition and nutrient cycling. The differential in nutrient availability in turn drives differences in plant species composition, as described above in the discussion of litter inputs. In brief, plants and the soil biota are tightly linked through the reciprocal influences on litter qualities and quantities: as these ecosystem properties change, nutrient supply changes, with differential effects on plant species and therefore on the composition of the plant community. Changes to the plant community in turn affect the quantity and quality of litter returned to the soil biota.

In sum, plants and the soil biota interact through both direct and indirect pathways (**Figure 10**). Although there are numerous variants on the specific mechanisms and pathways, they all involve a process by which the soil biota can influence competitive interactions among plant species, and therefore plant community composition. This may be through direct subsidies from mutualists, or indirect supply of nutrients from the saprotroph web. Because plant traits, especially the chemistry and physical structure of litter, have such a dominating effect on the activity and composition of the soil biota, any change in the species composition, or relative abundance of plant species in a community can result in a cascade of effects on the soil biota. These effects in turn affect the cycling and accumulation of carbon, nitrogen, phosphorus, and other nutrients within the soil, thus affecting ecosystem cycles. Recent work has shown, for example, that above-ground food webs and trophic cascades result in such changes, mediated by changes in plant composition and the ensuing litter inputs to the soil. Herbivores alone affect the soil biota through their differential effects on plant success, as well as the physical effects of trampling and the chemical inputs from their waste. But predators can affect the soil biota also: predation-driven

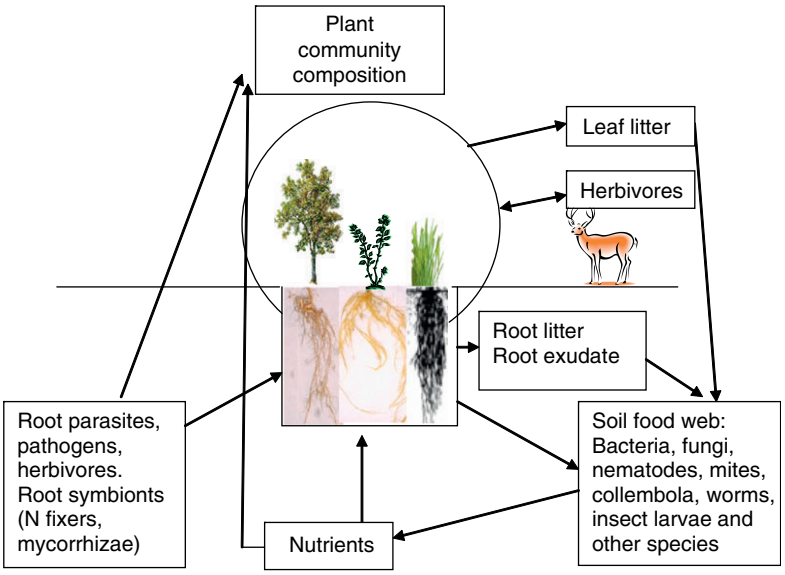


Figure 10 Pathways of interactions between the plant community and the soil biota. Both direct (symbiont, predators and parasites) and indirect (nutrient availability mediated by the saprotrophic soil biota) pathways are shown. Reproduced from Bardgett RD and Wardle DA (2010) *Aboveground-Belowground Linkages*. Oxford: Oxford University Press.

changes in the herbivore community can bring about change in plant community, and the resulting changes in plant inputs to the soil drive changes in decomposition and nutrient cycling resulting from changes to the soil biota.

Large-Scale Interactions: Stands and Biomes

The small- and medium-scale interactions described earlier contribute to larger-scale patterns that reflect, on one hand, the structure and composition of the whole plant community and, on the other hand, geographical effects on soil structure. These patterns may change over time, in response to successional changes in the plant community and in response to the geomorphological evolution of the landscape.

Geographic Patterns of Soils in Relation to Vegetation

In the classic formulation of Hans Jenny, the formation of soils reflects not only climate, bedrock geology, time, and local topography, but also vegetation. Over the broad geographic scale of biomes, regular patterns of soil characteristics are associated with patterns of plant community composition (**Table 6**). In polar regions, extreme climatic conditions constrain the development of vegetation; without organic matter inputs, microbial activity is low, weathering processes are extremely slow, and soils are barely more than bedrock material. In boreal and temperate forests, differences in the kinds of plant species result in differences in the rates of decomposition and accumulation of soil organic matter and the production of organic acids; these processes in turn affect the kinds of chemical weathering and pedogenic processes that occur. In grasslands, the high production rate of fine, nonwoody roots introduces both organic matter and the root-associated processes described earlier to depths of several meters; rapid decomposition and incorporation of surface litter combined with the pattern of root production create soils with very different properties than those of forest soils. Indeed, many studies have examined soils beneath forests that developed following long-term management of an area as grassland or cropland; rapid changes in the amounts, spatial distribution, and chemical quality of the soil organic matter cause changes in cation and metal chemistry associated with the changed chemistry of organic acid production and correlated changes in nitrogen availability and microbial activity.

Variations in Soils within Biomes: Effects of Species Composition

Within biomes, soil properties vary geographically with the species composition of the vegetation. The differential effects of particular tree species on soil properties result in parallel differences among stands when different species are dominant. These effects are clearly shown when single-species plantations are compared (**Figure 11**). The effects of different plant species on soil acidity is particularly important, as changes in pH and acidity affect the mobility of cationic nutrients and can result in large losses through leaching. The differences are due to both differences in the amount of cations stored in the vegetation and also to differences in the loss of cations in water leaching through the profile. Cations lost through leaching may cause a permanent change in the quality of the soil, unless weathering rates are equal to loss rates.

Successional changes in plant communities may in part be driven by plant-caused changes in soil structure and chemistry. In many communities, high rates of nitrogen cycling

Table 6 Regional characteristics of soils and vegetation

Vegetation	*Climate*	*Soil order*	*Speed of pedogenesis*	*Weathering processes*	*Horizons present*
Cold desert	Extremely dry, cold	Entisols	Extremely slow	Precipitation of carbonates	Only C
Tundra	Dry, cold	Inceptisols	Slow	Organic acid formation clay weathering Fe-hydroxides	Thin O, A, B
Boreal forest	Very moist, cold	Spodosols	Slow to moderate	Organic acid leaching Fe-hydroxides clay weathering	Thick O, A, leached E differentiated B
Temperate forest	Moist, mesic	Inceptisols alfisols	Moderate	Organic acid formation clay weathering Fe-hydroxides	Thin O, A, variable E, deep, differentiated B
Grassland	Moist to dry mesic	Mollisols	Moderate	Nonmobile organic acids; $CaCO_3$ precipitation, Fe-hydroxides	O absent, deep A, thin B, little differentiation
Savanna	Moist to dry warm	Vertisols, Utisols	Fast	Nonmobile organic acids, clay weathering, variable $CaCO_3$ and Fe–OH formation	No O, deep A, differentiated, very deep B
Desert	Dry warm	Aridisols	Slow	$CaCO_3$ precipitation Fe–OH formation	No O, A, shallow B
Tropical rainforest	Wet, very warm	Oxisols	Very fast	Mobile organic acids Fe, Al–OH formation strong clay weathering	Thin O, A thin E, very deep, differentiated B

Source: Reproduced from Ugolini and Spaltenstein (1992) Pedosphere. In: Butcher SS, Charlson RJ, Orions GH, and Wolfe GV (eds.) *Global Biogeochemical Cycles*, pp. 123–145. San Diego: Academic Press.

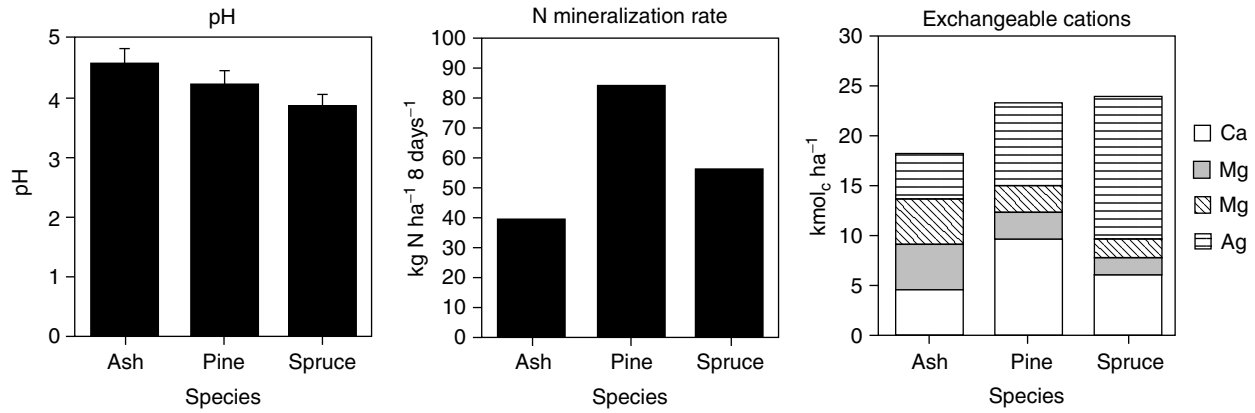


Figure 11 Variations in soil properties among different tree plantations. Data from Binkley D and Valentine D (1991) Fifty-year biogeochemical effects of green ash, white pine, and Norway spruce in a replicated experiment. *Forest Ecology and Management* 40: 13–25, with permission from Elsevier.

(mineralization and nitrification) are observed early in succession, in association with herb- or shrub-dominated communities, but these rates decrease markedly as trees come to dominate the community. These decreases in nitrogen cycling rates are associated with changes in litter chemistry (more recalcitrant material, lower nutrient concentrations) and in root quantities and distributions. A good example of this process is demonstrated by boreal ecosystems. Early in succession, N-fixing alders (*Alnus* spp.) in shrubby communities support high rates of nitrogen mineralization and nitrification; these rates decline rapidly when the alders are replaced by poplars (*Populus balsamifera*) and white spruce (*Picea glauca*). The decline in nitrogen cycling rates that accompanies the successional shift from alders to poplars is driven not only by differences in the N content of the litter, but also by the presence of high concentrations of tannins and phenolics in the spruce needles. The tannins inhibit microbial activity, whereas the phenolic compounds are used as a carbon source by the microbiota; however, both compounds result in N immobilization.

Invasions of trees into grasslands are associated with a wide range of changes in soil properties. For example, water content of the soils changes due to (a) differences in inputs (stem flow and drip from canopy leaves accentuate inputs, thick litter layers promote infiltration, but high rates of interception by the canopy can reduce total precipitation reaching the surface, and interception within the litter layer can prevent moisture from reaching mineral soil), (b) differences in soil porosity (large pores and channels associated with the growth of woody roots, and the growth of such roots to greater depths than the grasses), and (c) differences in loss (higher transpiration rates by trees, but shading decreases soil temperature and therefore evaporation rates, as does the accumulation of thick litter layers). Large withdrawals of cationic nutrients occur, which are sequestered in the larger amounts of perennial biomass of woody plants. Nitrogen becomes increasingly unavailable, as recalcitrant, lignin-rich litter accumulates. Finally, greater acidity of leaf litter tissue, greater production of organic acids from litter, and protons released through cation and ammonium uptake all drive acidification of the soil profile, with concomitant changes in both leaching loss of cations and weathering rates.

Just as changes in species composition may drive changes in soil properties, soil properties can affect successional patterns. Fertile soils are thought to support faster rates of succession (i.e., rates of species replacement), but with lower plant diversity than infertile soils. However, a wide variety of patterns have been observed, and many of the studies involve applications of nutrient fertilizer to experimental plots. Additions of nutrients have promoted the growth of early successional annual species, or the persistence of perennial herbs at the expense of woody pioneers, but opposite results have also been observed.

Successional processes have also been related to the mechanism of competition among plant species for soil resources. Tilman (1988) has shown that competitive ability is related to the capacity of a plant to obtain nutrients from the soil solution and thereby reduce concentrations below the minimum needed for competitors; this ability is balanced by the need to compete for light, so that the outcome of competition (and thereby the composition of a community) reflects the tradeoffs in competing for different resources. Plant growing on infertile soils tend to have lower nutrient requirements (low uptake rates but compensating low growth rates) and higher rates of nutrient conservation (higher efficiency of nutrient use) compared to plants typically found in fertile soils. The tradeoffs among characteristics promoting competitive success in infertile soils (strong competition for nutrients) versus success in fertile soils (competition for light being more important) ultimately determine the successional process in a given site.

Successional change can also be driven by negative plant–soil feedback resulting from the promotion of parasites and pathogens. In grasslands, for example, it has been shown that early successional species can create dense populations of their soil-based enemies, resulting in a competitive advantage for later-successional species that are not affected by these populations of parasites and pathogens. Changes in vegetation towards mid-successional stages of herbaceous plants result from this feedback. However, it is not well known whether similar negative feedback processes operate with the woody species that make up later-successional stages.

Clearly, as soil properties themselves are dynamic, and change with changes in species composition, interplant

interactions such as competition will change over successional time as the soil–plant system itself evolves. Such a situation has been observed in the Netherlands, where high rates of atmospheric nitrogen deposition have driven a successional change from heath-dominated (*Erica tetralix*) communities to grass-dominated communities (*Molinia caerula*). *E. tetralis* is characterized by low growth rates, low rates of biomass loss to shed tissues, herbivory, and so on, and poorly decomposable litter, which supports low rates of mineralization but high rates of sequestration in thick organic horizons. In contrast, *Molinia* is characterized by high growth rates, high rates of tissue loss to shedding and herbivory, and rapidly decomposable litter. As expected, in experiments specifically testing competitive ability, *Erica* is a superior competitor when nutrient supplies are low, whereas *Molinia* is the superior competitor when nutrient supplies are high. Thus, the feedback relationships between plants and soils can drive successional changes in community composition.

Implications for the Management and Conservation of Ecosystems

Clearly, the complex, multiscale interactions of plants and soil affect all aspects of both the biology of plants and the properties of soils. This implies that both purposeful and inadvertent human effects on plant–soil systems will have diverse, perhaps unexpected ramifications, as changes in soils affect plants and *vice versa*. Following are brief discussions of four major areas of current concern, in which plant–soil interactions are likely to play a major role.

Forestry, Agriculture, and Ecosystem Restoration: Human Creation of Ecosystems

Humans create plant communities for a wide variety of purposes, from the production of food, fiber, and fuels to the restoration of "natural" ecosystems and the creation of beauty for its own sake (in gardens and parks). The choice of plant species is often conditioned by soil properties; certain crops cannot be grown in certain soils, or conversely particular soils may be ideally suited to particular crops. Conversely, the choice of crop plant or tree can alter soil properties, in some cases perhaps irreversibly. As described earlier, the substitution of coniferous trees for deciduous trees, a widespread practice because of the rapid growth and high economic value of many conifers, results in acidification, the loss of nutrient cations through leaching, and the creation of thick litter and humus layers. In some cases, foresters use mixtures of tree species that are specifically designed to take advantage of differences in plant–soil interactions among the species. For example, larch (*Larix* spp.) and lodgepole pine (*Pinus contorta*) are often interplanted with Sitka spruce (*Picea sitchensis*) in plantations in the British Isles, because larch and pine promote higher rates of nitrogen cycling and therefore faster growth of the spruce. The increased nitrogen availability results, in part, from more widespread roots of the interplanted pines or larches, which reduce waterlogging in the organic horizon and thereby stimulate mineralization. Nitrogen-fixing plants are often introduced into both agricultural and forest plantations in order to alter the cycling of N through the soil.

Ecosystem restoration similarly involves the deliberate establishment of particular plant species and species mixtures. The choice of species is most commonly made on the basis of replicating the "natural" community, rather than explicitly with respect to soil properties. However, plant communities are sometimes specifically designed to alter soil properties: the introduction of fast-growing grasses to introduce organic matter in barren soils, and the parallel introduction of N-fixing species into such soils, the use of nutrient-demanding, fast-growing species to remove nutrients (through repeated harvest) from overfertilized fields being restored to native grasslands, or from wetlands being used for waste-water treatment are examples. Conversely, soil properties may unexpectedly affect the course and success of restorations. The potential for such effects is clearly illustrated in efforts to restore coastal wetlands in San Diego, California. Coarse-textured sediments were used to construct the wetland, which were unlike the fine-textured sediments found in naturally occurring marshes. Although the introduced salt marsh plants survived and spread, they only attained about half the height of plants in undisturbed marshes; the cause of the problem was the low storage rate of organic matter in the coarse sediments and the concomitant low intrinsic supply rate of nitrogen from the sediments.

Exotic Species Invasions – Inadvertent Human-Caused Changes in Species Composition

The spread of exotic plant species around the world is recognized as one of the major threats to biodiversity. Plant communities are being altered through invasions of species that in some cases become part of the community, but in other cases exclude and eliminate the native species. These invasions often cause economic problems, as well as challenges for conservation of native communities, as the invading species degrade grazing lands, choke stream channels, or promote wildfires.

These changes in plant community composition are likely to alter soil properties, especially when the exotic species is of a different growth form than the natives or when its litter chemistry is different. Recent studies have shown that, for example, invasions of cheat grass (*Bromus tectorum*) into desert grasslands alters the abundance and species composition of fungi, protozoa, nematodes, and soil invertebrate communities, decreases the abundance of N-fixing lichens, and thereby decreases the amount of nitrogen in the soil. Invasions of exotic shrubs (*Berberis thunbergii*) and grasses (*Microstegium vimineum*) into deciduous forests stimulates increases in nitrification and soil pH. Invasions of the hawkweeds (*Hieracium* spp.) into pastures in New Zealand causes the accumulation of nitrogen, but decreases soil pH. Invasions of grasses (*Melinus minutiflora*) into Hawaiian shrub lands stimulates increases in net nitrogen mineralization rates. It is likely that such changes are widespread and may affect the ability of resource managers to remove the exotics and restore the native plant communities.

Nutrient Pollution – Inadvertent Human-Caused Changes in Soil Properties

The profound and far-reaching changes in atmospheric chemistry that are accompanying the spread of industrial society around the world include greatly increased inputs of several nutrients to soil–plant systems, through the deposition of particles and the dissolution of soluble ions in precipitation. The most important changes include greatly elevated deposition of nitrogen and sulfur. As these are both important plant nutrients and as their chemistry is intimately linked with other chemical and biological processes within the soil, a cascade of interactive changes to the soil–plant system are taking place in many places around the world.

Deposition of sulfates at rates below the level that directly injures plants still causes problems through soil acidification. Leaching of sulfate in soil water carries with it nutrient cations, and in soils with low base saturation and a parent material that supplies few cations through weathering, cations can become seriously depleted. Calcium, an essential element for the integrity of root function, is particularly strongly affected. Hardwood forests in New Hampshire (US) that have received large inputs of sulfate for several decades have prematurely stopped growing, and their tissues show abnormally low concentrations of calcium. Depletion of nutrient cations is accompanied by increases in the concentration of aluminum ions in the soil solution and on exchange sites, and the aluminum can be directly toxic to plants. Depletion of calcium also leads to increased sensitivity to frost damage of needles, due to change in the structure of cell membranes, as has been demonstrated for declining red spruce in the eastern US.

Nitrogen deposition to ecosystems is now occurring at double the estimated rates prior to the industrial revolution. Nitrogen originates from agriculture (applications of fertilizers and manures, and the planting of N-fixing crops), combustion (automobiles, power plants, and also forest burning), and waste disposal (e.g., sewage). Although nitrogen is considered as the most important limiting resource in most terrestrial ecosystems, an excess does not simply alleviate stress on plants. As was discussed earlier, plants growing in nutrient-poor soils have a variety of characteristics, including patterns of both root and above-ground morphology and physiology, characteristics of life history, and intrinsic growth rates, which adapt them to such habitats and, which contrast strongly with the characteristics of plants found in nutrient-rich soils. Added nutrients often cannot be absorbed or assimilated by such plants, nor can they respond with increased growth. Rather, species adapted to nutrient-rich soils become capable of competitively displacing these species. Thus, many ecosystems that had been subjected to limiting supplies of nitrogen have undergone profound changes in species composition. The substitution of grass for heather in European heathlands, as discussed earlier, is a case in point. Similar changes in native plant communities are occurring elsewhere in the industrialized world. And because communities found in infertile soils tend to have high diversity of plants, including many rare species, the increasing dominance of nitrophilous species may threaten many of them.

Nitrogen additions to soil not only affect the plant community through competitive displacements and substitutions, but also affect the internal nitrogen dynamics of the soil. Evidence from a variety of ecosystems has shown that N additions stimulate intrinsic mineralization processes, thus accentuating the effects of added N. Moreover, changes in the microbiota under conditions of high N inputs can result in the development of acid-tolerant nitrifiers; the increased levels of NO_3^- not only cause changes in plant community composition, but cause leaching of cations and thus changes in soil chemistry. As plant community composition changes, the chemistry, amount, and timing of litter inputs will change, as will all other components of the plant–soil system. The long-term effects of this chain of linked changes can only be guessed.

Changing CO_2 in the Atmosphere – Inadvertent Human-Caused Changes to the Soil–Plant System

The rising concentration of CO_2 in the atmosphere is expected to affect virtually all aspects of the plant–soil system. The topic can be touched on only briefly here, simply to note the major processes and interactions that are anticipated. Plants are expected to respond to increasing CO_2 by increasing their overall rate of production (particularly C_3 plants), increasing the allocation of carbon to below ground (including root growth, root turnover rates, root respiration rates, and rhizodeposition rates). Water use efficiency and nutrient uptake rates also often increase in experimentally enriched plants. Changes in litter chemistry (increased C:N ratios, increases in lignin, phenolics, or tannins) are anticipated but have yet to be clearly demonstrated. However, individual species responses have proven to be highly variable; studies of the growth response of a wide range of species has shown that changes in growth rate can range from a 25% decrease to a 200% increase. However, it is clear that virtually all aspects of plant morphology and chemistry that have been shown to affect soil properties are likely to change, precipitating changes in the soil in accordance with the processes discussed in this chapter. **Figure 12** illustrates the complexity of the possible responses. The research done over the past decade has demonstrated that climate change and the linked changes in atmospheric chemistry have both direct effects, through changes in temperature and precipitation, and indirect effects, through changes to plant growth from both these climatic changes and the increase in CO_2 concentrations in the atmosphere on plant–soil interactions. Moreover, changes in the distribution of both above-ground and below-ground herbivores and predators in response to changing climate will have ramifying effects on soil food webs and nutrient cycles, largely mediated by changes in the plant community. These types of interactions have now been demonstrated in many studies, but there is much still to be learned about how climate change and CO_2 changes affect plant–soil interactions.

Although a large amount of research is currently being conducted on these linkages, the complexity of the problem and the variability in response among different plants and in different soils make the discovery of general patterns and predictability difficult and yet to be achieved.

Biodiversity and Ecosystem Function

The effects of changing plant community diversity on the structure and function of soil is among the most important

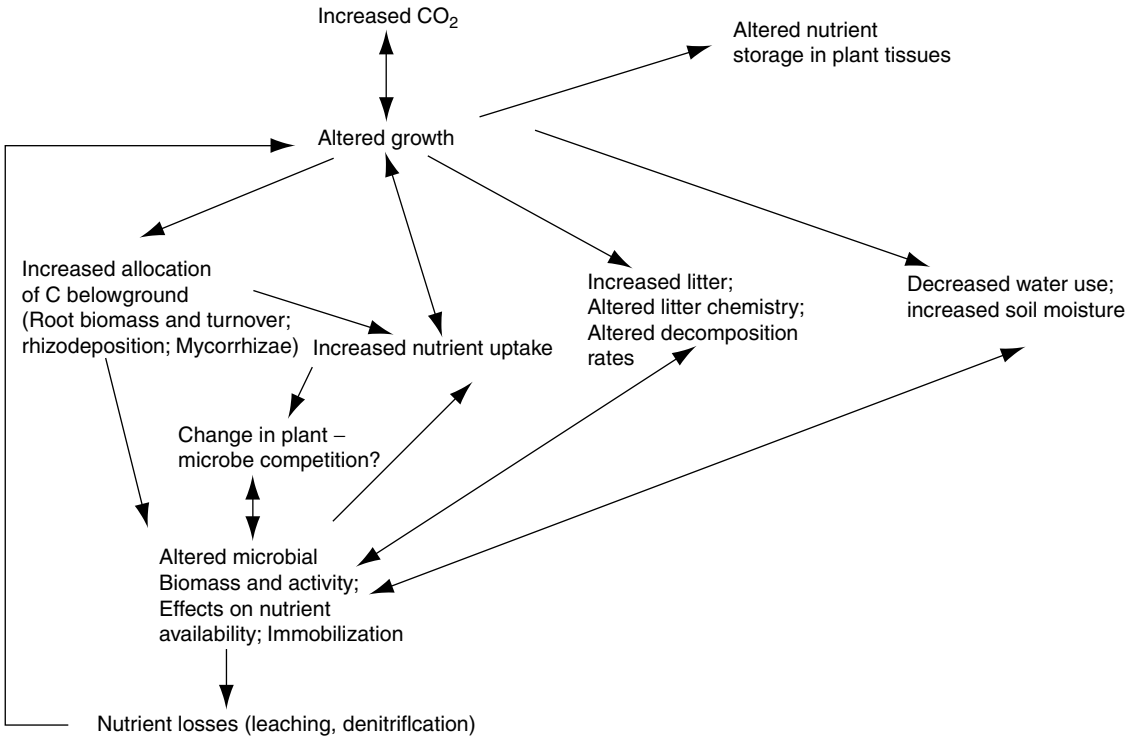


Figure 12 A schematic outline of the potential changes in the soil–plant system that may be induced by changes in atmospheric CO_2 concentrations.

current topics of research. Several studies utilizing experimental grasslands have shown that such ecosystem processes as nitrogen mineralization, carbon accumulation and mineralization, nitrogen concentrations, soil respiration, and litter decomposition rates are altered (often decreased) by decreases in plant diversity, but other studies, including studies in forested ecosystems, have presented contradictory results. In many cases, particular species or particular classes of species (e.g., nitrogen-fixers) have large effects. In addition, responses of soil processes to changes in plant diversity often are largest in comparing small numbers of species; once the plant community reaches a moderate number of species (three or four), there is no further change as plant species diversity increases. Reciprocal effects of diversity within the soil microbiota and soil fauna on the plant community are largely unknown.

The effects of changes in diversity within the plant community to the biology, chemistry and physics of soils, and *vice versa*, will be among the most important questions to be addressed in the future.

See also: Biogeochemical Cycles. Carbon Cycle. Climate Change and Ecology, Synergism of. Ecology of Agriculture. Food Webs. Forest Ecology. Greenhouse Effect. Nitrogen, Nitrogen Cycle. Plant Invasions. Restoration of Biodiversity, Overview. Soil Biota, Soil Systems, and Processes. Subterranean Ecosystems. Succession, Phenomenon of. Temperate Forests. Terrestrial Ecosystems

References

Atkinson D (ed.) (1991) *Plant Root Growth. An Ecological Perspective*. Oxford: Blackwell Science Publications.

Bardgett RD and Wardle DA (2010) *Aboveground-Belowground Linkages*. Oxford: Oxford University Press.

van Breeman N (1995) Nutrient cycling strategies. *Plant and Soil* 168–169: 321–326.

Chapin III FS, Sala OE, Burke IC, *et al.* (1998) Ecosystem consequences of changing biodiversity. *BioScience* 48: 45–52.

Ehrenfeld J, Ravit B, and Elgersma K (2005) Feedback in the plant-soil system. *Annual Review of Environment and Resources* 30: 75–115.

Eissenstat DM and Yanai RD (1997) The ecology of root lifespan. *Advances in Ecological Research* 27: 2–60.

Fitter AH, Atkinson D, Read DJ, *et al.* (eds.) (1985) *Ecological Interactions in Soil. Plants, Microbes, and Animals*. Oxford: Blackwell Scientific Publishers.

Foster RC (1988) Microenvironments of soil microorganisms. *Biology and Fertility of Soils* 6: 189–203.

Glinski J and Lipiec J (1990) *Soil Physical Conditions and Plant Roots*. Boca Raton, FL: CRC Press.

Harley JL and Russell RS (eds.) (1979) *The Soil–Root interface*. London: Academic Press.

Hobbie S (1992) Effects of plant species on nutrient cycling. *Trends in Ecology and Evolution* 7: 336–339.

Hu S, Firestone MK, and Chapin III FS (1999) Soil microbial feedbacks to atmospheric CO_2 enrichment. *Trends in Ecology and Evolution* 14: 433–437.

Jenny H (1980) *The Soil Resource*. New York: Springer-Verlag.

Kaye JP and Hart SC (1997) Competition for nitrogen between plants and soil microorganisms. *Trends in Ecology and Evolution* 12: 139–143.

Lynch JM (ed.) (1990) *The Rhizosphere*. New York: John Wiley and Sons.

Metting Jr. FB (ed.) (1993) *Soil Microbial Ecology*. New York: Marcel Dekker.

Robinson D (1994) The responses of plants to non-uniform supplies of nutrients. *New Phytologist* 127: 635–674.

Tilman D (1988) *Plant Strategies and the Structure and Dynamics of Plant Communities*. Princeton, NJ: Princeton University Press.

Vancura V and Kunc F (eds.) (1988) *Soil Microbial Associations – Control of Structures and Functions*. New York, NY: Elsevier.

Vogt K, Grier CC, and Vogt DJ (1986) Production, turnover and nutrient dynamics of above and belowground detritus of world forests. *Advances in Ecological Research* 15: 303–377.

Waisel Y, Eshel A, and Kafkafi U (eds.) (2002) *Plant Roots – The Hidden Half*, 3rd ed. Boca Raton, FI: CRC Press.

Plant Sources of Drugs and Chemicals

William H Gerwick, University of California, San Diego, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Algae A group of aquatic plant-like organisms which obtain energy through photosynthesis.
Anticancer agent A chemical substance that is capable of bringing about a remission or cure of the family of diseases known as cancer.
Biosynthesis The biochemical process by which a plant produces molecules of utility or adaptation.
Cyanobacteria A group of photosynthetic prokaryotes, largely aquatic, which have an ancient evolutionary origin and exceptional capacity to produce bioactive natural products.
Cyanotoxin A group of toxins produced by various fresh water, marine and terrestrial cyanobacteria.
Cytotoxicity The ability of a chemical compound to kill cells in an experimental system, such as in a petri dish.
Drug development The process by which an early stage lead compound is developed using medicinal chemistry, pharmacology, toxicology and human clinical trials to become an approved pharmaceutical.
Natural product A compound, organic in nature (e.g., composed of carbon, hydrogen, oxygen and nitrogen), that is produced by a living organism for a purpose other than basic life functions, and provides some adaptive value to the producing organism.
Peptide A compound composed of several amino acids that are linked together by amide bonds.
Secondary metabolite A chemical substance produced for a reason other than basic life requirements and with some adaptive or defensive value.

Introduction

The intent of this article is to illustrate by example the enormous diversity of plant-derived natural products and their tremendous importance to human society. It is precisely this diversity of structure that has been, and continues to be, of such incredible value to the pharmaceutical and agricultural industries. Ultimately, this chemical diversity is rooted in the inherent biodiversity of our planet. Sadly, space limitations for this article limit this treatment to only a sampling of some of the most interesting examples. These have been chosen to exemplify several concepts: various chemical adaptations typical of particular plant groups, classes of compounds of utility to society, older examples deriving from ethnobotanical information, and modern examples resulting from methodical broad-based natural product screening programs.

It cannot be debated that the premier importance of plants to human existence is as a source of food and oxygen. Related to their serving a food function is the use of plant-derived materials as food additives to impart desirable taste or textural properties. Examples include the use of a class of biopolymer from red marine algae, the carrageenans, to stabilize emulsions in such diverse products as ice cream and beer, and the use of the sugar alcohol sorbitol, a product of the mountain ash *Sorbus aucuparia*, as a noncariogenic sweetener with humectant (wetting) properties in toothpaste. However, this article will focus on another property of plants; their ability to produce unique molecular entities with potent pharmacological effects in mammalian systems. Why do plants make such compounds? Given their complicated, often exotic, structures, they certainly expend considerable biochemical energy to produce these compounds. Although some hold that such compounds are vestigial in nature (previously served a function that is no longer in evidence), there is excellent experimental evidence to support the idea that most, if not all, are produced with powerful adaptive functions, such as defense against potential predators, prevention of microbial infections, and inhibition of the growth of competing species.

However, complicating the above scenario somewhat is the quite recent recognition that a number of "plant-derived" compounds, including several of importance to human society (e.g., the camptothecins and vinca alkaloids, two classes of anticancer alkaloids), are in fact produced by endophytic fungi that live in specific association with their plant hosts. Indeed, in a number of other contexts, there is a growing recognition that microbial life forms are the actual producers of natural products which have been isolated from and ascribed to their macroscopic host plants and invertebrates (e.g., insects, sponges, and tunicates). Nevertheless, many of these adaptive metabolites still possess beneficial properties to the host macroscopic organisms, and by extension, to their microscopic symbionts, but serve to illustrate some of the complexity of relationships in the natural world.

Primitive peoples throughout the world make use of their indigenous flora as a source of medicines. Through the process of trial and error, these cultures have examined and discovered many plants that produce unique molecular entities with valuable biological properties. The fields of "ethnopharmacology" or "ethnomedicine" seek to obtain new pharmaceutical leads from a study of the native medicines of these primitive peoples, an approach that has been highly successful and resulted in such compounds as aspirin from willow bark (*Salix*), the anticancer alkaloid vincristine from the Madagascar periwinkle (*Catharanthus*), and the antimalarial sesquiterpene artemsinin (qinhaosu) from the Chinese herb *Artemisia annua*.

Beginning in the 1950s, comprehensive evaluations of the unique constituents of plants were undertaken. In the USA, these earliest efforts were coordinated largely by the National Cancer Institute (NCI), and so necessarily had a focus on anticancer properties. Assays for potential anticancer activity at that time were run in intact animals (*in vivo* evaluations). Although even the most modern drug-development efforts require evaluation of a drug for efficacy and toxicity in animal systems, this usually occurs only after a candidate pharmaceutical has shown the requisite property in isolated biochemical and cell-based screens. Consequently, current *in vivo* evaluations occur only on a very small fraction of the most promising candidate pharmaceuticals. In contrast, the early screening efforts of plant extracts used *in vivo* techniques; consequently, they were very slow, expensive, and, perhaps worst, used test animals in large numbers.

In the late 1970s, and gaining broad acceptance in the 1980s, a major shift away from whole animal primary screening was achieved with the introduction of cell-based screening. In the cancer effort, this largely had the endpoint of cytotoxicity to cancer cells grown in small petri dishes or, later on, 96-well trays. Expansion of this idea was undertaken by the NCI wherein extracts and compounds have been evaluated for their level and profile of toxicity to 60 different cancer cell lines. The cell lines were chosen so as to represent cancers affecting nine different organ types and have been quite useful in detecting new cytotoxins as well as giving information about their molecular mechanism of action.

The past 15 years has seen a growing departure from even cell-based assays. Over the past 40 years, there have been tremendous advances in understanding at a molecular level the causes of many diseases, including cancer. This has translated into assays that screen for compounds or extracts that interfere with specific enzymatic reactions or protein–protein interactions that have been shown to underlie a particular human disease. With the reduction in screening format from whole animal to cell to isolated protein, it is now possible to screen hundreds of thousands of compounds or extracts in a few weeks time using high-throughput screening (HTS) technologies. Here, a target protein is produced in large quantities by molecular biological means in a surrogate organism, placed in 96-, 384-, or even 1536-well plate formats, and compounds or extracts evaluated for inhibitory effects using robotics that work around the clock. It is hoped that screens of this design and scope will uncover new generations of pharmaceuticals that will have highly specific actions, targeting the underlying causes of disease, with very high potency, and leaving normal or undiseased cells untouched (Folmer *et al.*, 2010).

Unfortunately, this approach has not fully lived up to this expectation, and over the past 5 years, there is a growing recognition that cell-based screening systems have several advantages over these purely *in vitro* biochemical target screens. In cell-based screening, the compounds found active are necessarily active in a biological context, that is they are cell penetrant and stable to cellular enzymes, in addition to being efficacious at desired biochemical targets. As a result, the molecular leads found from cell-based screening are of higher quality and already have been selected for several desirable pharmaceutical properties (Appendino and Pollastro, 2010).

Examples from Marine Plants Illustrating Plant Sources of Drugs and Chemicals

Anticancer Agents from Marine Cyanophyta (Cyanobacteria)

The marine "plants" generally include a number of groups, including microscopic as well as macroscopic forms. The major subdivisions of the microscopic forms include dinoflagellates, cryptophytes, chrysophytes, cyanophyta (=cyanobacteria), and prochlorophyta. The major groups of macroscopic marine plants include the Rhodophyta (=red algae), Chlorophyta (=green algae), Phaeophyta (=brown algae), and a number of marine angiosperms (sea grasses); the latter represent terrestrial species that have "readapted" to life in the ocean. Although interesting and significant molecules have been isolated from all of these groups, secondary metabolites of special note have been obtained from marine cyanobacteria, dinoflagellates, and red and green macrophytes. Examples from these four groups follow below (Blunt *et al.*, 2010).

Curacin A and the "chemotype" concept

Cyanobacteria take on many different physical forms in the oceans, from gelatinous encrustations and tufts of 20 cm hairlike filaments, to microscopic unicells free in the water column. One species of pan-tropical distribution, *Lyngbya majuscula* (family Oscillatoriaceae), also known as "Mermaid's Hair," has been particularly plentiful in its production of structurally unique and biologically important secondary metabolites (Jones *et al.*, 2009). Such an example is given by a collection made in Curaçao, in the southern Caribbean, which was reported to possess an extract with fish, snail, brine shrimp, and cancer cell toxicity. Bioassay guided isolation efforts led to the isolation of several structurally unrelated substances which were each responsible for only some of the observed biological properties. The cancer cell and brine shrimp toxicity was principally due to a unique thiazoline-containing lipid, named curacin A (**Figure 1**). Its mechanism of cytotoxicity was shown to involve inhibition of microtubule formation, crippling the ability of cells to properly segregate chromosomes at mitosis. Several very valuable anticancer agents that are in common use in the clinic today work by this same essential mechanism (e.g., *see* Paclitaxel). Curacin A has been shown to interact with microtubules at the same site as the antigout drug, colchicine (an alkaloid from the plant *Colchicum autumnale* = autumn crocus). Unfortunately, curacin A was found difficult to work with *in vivo* because of solubility and stability problems. Recent synthetic medicinal chemistry efforts have overcome at least some of these

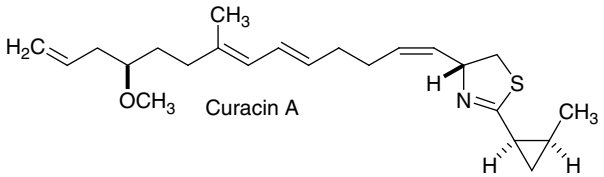


Figure 1 Structure of the antimitotic lipid "curacin A," obtained from a Caribbean collection of the blue–green alga (cyanobacterium) *Lyngbya majuscula*.

problems, and work continues to develop an anticancer drug patterned on this "structural concept."

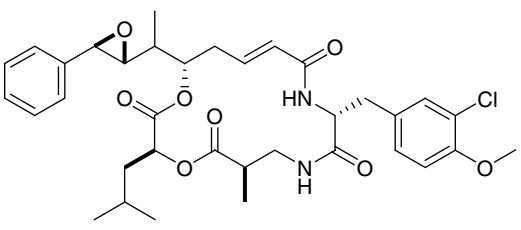
Sources of curacin A for continued drug testing have largely been from materials freshly collected in Curaçao, although the living alga does produce curacin A in laboratory culture. Interestingly, when collections of *L. majuscula* were made from approximately 20 sites along the leeward coast of Curaçao, only two chemotypes, both quite limited in extent, were found to make curacin A. Continued investigations of this phenomenon have substantiated these findings, and thus a very intriguing conclusion must be drawn, which is of enormous significance to "biodiversity preservation." It is widely held that in order to preserve the valuable genomic potential of a particular "species," it is sufficient to preserve a few individuals or a limited population. This work with Curaçao collections of *L. majuscula* suggests that in order to accomplish the larger goal, it is actually necessary to consider preservation of the many, almost innumerable, "chemotypes" of a given species (Gerwick *et al.*, 2008).

Cryptophycin

Another very promising anticancer treatment derives from the freshwater cyanobacterium *Nostoc* sp. Although one of the active compounds, cryptophycin A, was first isolated as an antifungal substance, it was only after several more years that, on re-isolation, its anticancer potential was realized (**Figure 2**). In common with curacin A, cryptophycin also blocks microtubule assembly processes, although at a different drug-binding site (the vinca alkaloid site). In animal testing, cryptophycin has brought about actual "cures" of some tumor types. An enormous effort has been devoted to the total chemical synthesis of cryptophycin and several hundred analogs with the intent of ensuring that (1) a ready supply of the drug is available for clinical testing and (2) that advanced cancer treatment evaluations utilize the most effective molecule in this drug series. Testing of the drug in humans advanced to phase II trials, but was terminated due to lack of efficacy and toxicity. Sadly, it seems that possibly the wrong cryptophycin analog was taken forward into human clinical trial; however, there are continuing rumors that the drug class will be explored for its therapeutic utility by another major pharmaceutical company.

Toxins of Impact and Utility from Dinoflagellates

Dinoflagellates represent a very diverse group of organisms, made up of both freshwater and marine varieties. Dinoflagellates are protists, which are not strictly considered plants or animals although they share characteristics of both, and appear to be the result of ancient tertiary endosymbiosis between animal cells, a unicellular red algal cell, and additional cell types. Some dinoflagellates are photosynthetic whereas others are parasitic to fish or to other protists. They can be very small or as big as 2 mm in diameter, such as in the case of *Noctiluca* sp. A curious feature of some dinoflagellates is that they are capable of bioluminescence.



Cryptophycin A

Figure 2 Cyclic peptide structure of the anticancer metabolite "cryptophycin."

Dinoflagellate blooms in the world's oceans have had tremendous impact on human society; despite this importance, many aspects of these organism's genetics, biology, and biochemistry are poorly understood. Hundreds of people suffer from the effects of seafood poisoning and millions of fish and other marine life are killed each year as a direct consequence of large blooms of dinoflagellates. Some of the coastal species can "bloom" during the warmer months and can even make the water appear red or golden colored; hence, the term "red tide." For many years, seafood poisonings and fish kills have been associated with these dinoflagellate blooms, but it is only within the past decade that the chemistry behind these effects has been unraveled.

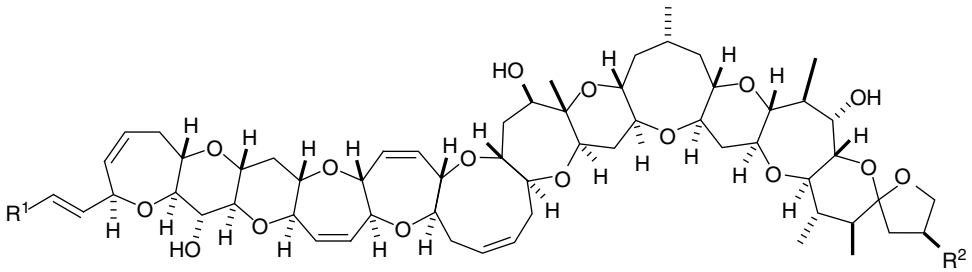
Ciguatoxin and maitotoxin

A prime example of a dinoflagellate toxin with huge impact on society is given by Ciguatera Seafood Poisoning (CSP), which is quite prevalent in tropical areas around the globe. CSP is more common than any other illness associated with consumption of tainted seafood. Generally, the poisoning results from ingestion of coral reef fish that have accumulated the toxins through their diet. Symptoms include, but are not limited to, memory loss, joint pain, miosis, erethism, cyanosis, and prostration. One of the more interesting symptoms associated with ciguatera is a neurological disturbance leading to reversal of hot and cold sensations. Most of the toxic effects of CSP are attributed to ciguatoxin, which is a metabolically oxidized derivative of a substance produced by the dinoflagellate, *Gambierdiscus toxicus*. The structure of ciguatoxin was deduced after a monumental 15-year effort by the laboratories of Dr Paul Scheuer at the University of Hawaii, USA, and Drs Takeshi Yasumoto and Michio Murata at Tohoku University, Japan (**Figure 3**). The structures of ciguatoxin and the probable precursor produced by the dinoflagellate are shown below as well. These compounds illustrate the common polyether motif found in many dinoflagellate toxins.

Another toxin that should be noted here is maitotoxin (**Figure 4**). Probably produced by dinoflagellates related to those that produce the ciguatoxin precursor, this compound is even more toxic than ciguatoxin. It is by far the largest compound ever isolated that is not a biopolymer (e.g., protein or carbohydrate). Its structure was also elucidated by the laboratories of Yasumoto and Murata, and represents a milestone in organic structure analysis.

Gymnodinium breve

For many years it was thought that seafood poisonings caused by dinoflagellates were limited to tropical regions of the world. However, in 1987, 14 dead humpbacked whales were discovered in Cape Cod Bay, Massachusetts, USA. Examination



Ciguatoxin R^1 = -CH(OH)-CH_2OH; R^2 = OH
Pre-ciguation (CTX-4B) R^1 = -CH=CH2; R^2 = H

Figure 3 Polyether structure of ciguatoxin, the active poison in ciguatera shellfish poisoning.

of the animals showed that they had been well until just before their deaths, appearing to be well nourished and showing no signs of ill health. A few months later, fishermen and beach-goers along the North Carolina coast began to complain of respiratory problems and eye irritation. Soon after, seafood consumers began to complain of diarrhea and dizziness. In later years and on a periodic basis, these same scenarios have been reported all along the east coast and the Gulf of Mexico. Most of these outbreaks have been connected with the dinoflagellate *Gymnodinium breve*. This dinoflagellate frequently forms large blooms, or "red tides," off the Florida coast, which result in many tons of dead finfish washing up on the beaches. This, in turn, repels many tourists and has forced many fisheries to be closed. In the early 1980s, the toxic constituents produced by this organism were found to be brevetoxin A and B, along with other structurally related compounds (**Figure 5**). Additionally, during especially heavy blooms and unfavorable weather patterns, the dinoflagellate brevetoxins are aerosolized and negatively affect pulmonary function in beachside populations, especially in the elderly. Red tide toxins have many and diverse consequences to human society.

Kainic Acid and Domoic Acid, Neurotoxic Anthelminthics from Red Algae

Two Japanese red algae, *Digenea simplex* and *Chondria armata*, have been employed for more than 1000 years in Japan for their potent anthelminthic properties; that is, eliminating intestinal worms, such as parasitic roundworms (*Ascaris lumbricoides*), whip worms (*Trichuris trichura*), and tape worms (*Tenia* spp.). Two closely related compounds, domoic acid and kainic acid (**Figure 6**), have been isolated from these red algae and are responsible for these therapeutic effects. These compounds were traditionally ingested by consumption of the producing algae, their names deriving from the Japanese names for these seaweeds, *domoi* and *kaininso*. Kainic acid has become an important tool in neurobiological research because of its potent neurotoxic effects on neurons, working through the glutamate receptor. However, domoic acid has been identified as the active poison in amnesic shellfish poisoning (ASP), which has recently struck both east and west coasts of North America. The levels of the drug that are ingested for anthelminthic properties are relatively small when compared with the amount of compound observed in the recent outbreaks of ASP. However, the macrophytic red algae were determined to not be the culprit in the recent ASP outbreaks; rather, a planktonic pennate diatom, *Nitzschia pungens f. multiseries*, now renamed *Pseudonitzschia australis*, is the producer in this latter case.

The mechanism of action of these compounds is very similar, as is suggested by their closely related chemical structures. Both are analogs of glutamic acid, a well-known neurotransmitter found in the brain of mammals. The algae-derived substances are known as excitotoxic amino acids as they can mimic glutamate so as to cause massive calcium ion influxes into cells, a lethal event in various types of brain neurons. Domoic acid is two to three times more potent than kainic acid and 100 times more potent than L-glutamate at inducing these neurotoxic events. Although these compounds (domoic and kainic acid) are environmental neurotoxins to be avoided, they have shown considerable utility in biomedical research. For example, the resulting extensive neuronal loss seen with the appropriate dose of kainic acid is very similar to that observed for Huntington's disease, and hence, provides an excellent model for the study of this disease in humans. Kainic acid has also been used in the study of epilepsy and Alzheimer's disease.

A problem of epidemic proportion that occasionally arises from these excitotoxic amino acids is known as amnesic shellfish poisoning (ASP). It has been shown that blooms of the planktonic diatom *P. australis* result in a build-up of domoic acid in shellfish. Three highly publicized ASP outbreaks have occurred in the past 15 years in North America. The first was caused by the ingestion of mussels containing high levels of domoic acid from Prince Edward Island, Canada. The second was a large pelican kill in Monterey Bay, California. The domoic acid was found to accumulate through the food chain from diatoms to small fish to anchovies to pelicans. The last reported outbreak was in razor clams and Dungeness crabs in Oregon and Washington. The discovery of high concentrations of domoic acid led to the shutdown of these seafood harvests for several years, causing substantial economic hardship for afflicted communities. As a result of these ASP outbreaks, strict guidelines have been developed for the maximum allowed domoic acid content in harvested shellfish (20 ppm).

Green Algal Metabolites

Coral reef environments are habitats of high speciation and biodiversity ranging from microorganisms to top predators.

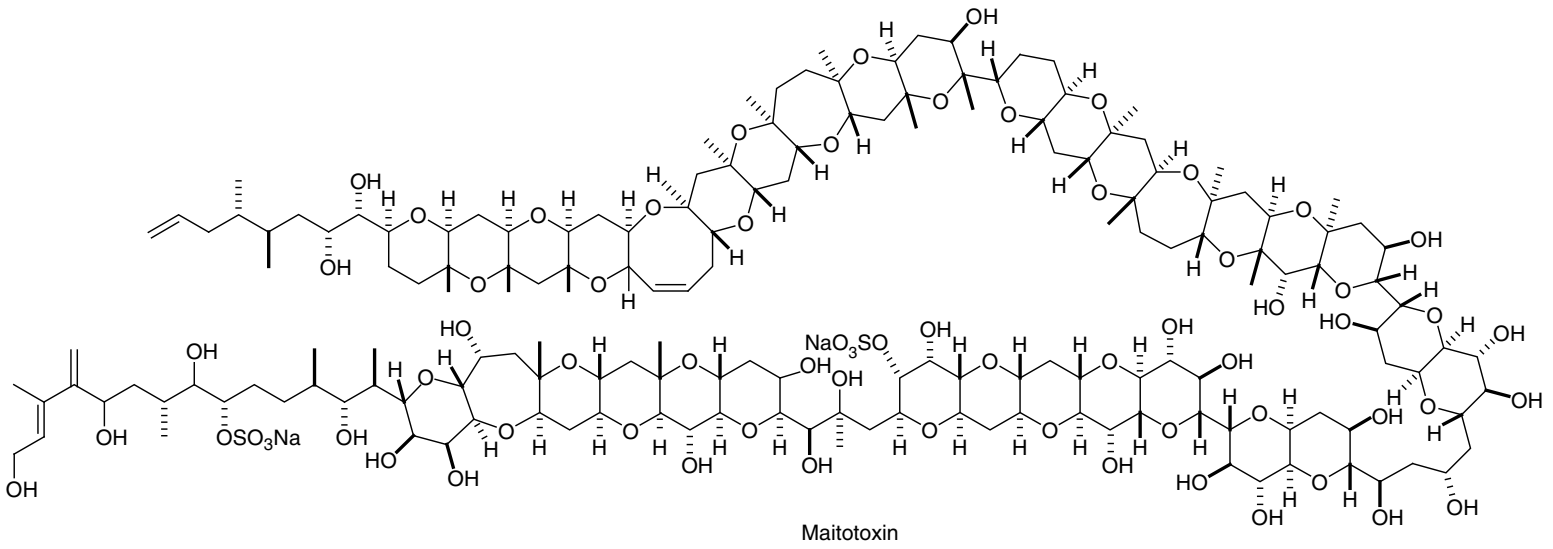


Figure 4 Structure of maitotoxin, the most complex nonpeptide structure ever determined.

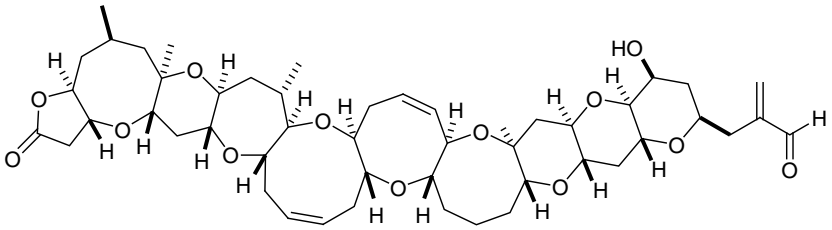


Figure 5 Brevetoxin A, the causative fish poison produced by the red tide organism *Gymnodinium breve*.

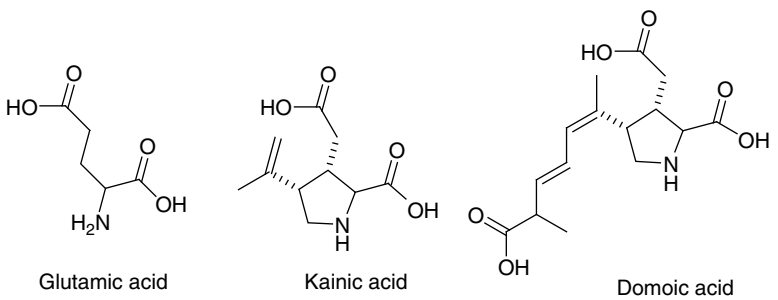


Figure 6 Structures of the neurotoxic amino acid, glutamic acid, present in mammalian cells and two neurotoxic congeners, kainic acid from red algae, and domoic acid from red algae and diatoms.

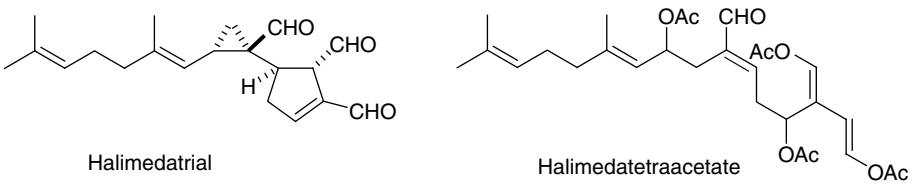


Figure 7 Structure of the green algal predator defense compound halimedatrial and the "protoxin" halimedatetraacetate.

This biodiversity inherently creates intense competition between species, including predator–prey relationships such as between herbivores and their algal diets. Methods that algae use to protect against predation include calcification, production of tough spiny appendages, and the biosynthesis of chemically noxious deterrent molecules that are either toxic or deter feeding.

The green alga *Halimeda* sp. employs both calcification and chemical defense strategies. As a result, *Halimeda* species constitute the largest macroalgal biomass in many tropical reef systems. In feeding preference studies, and analyses of reef fish stomach contents, *Halimeda* spp. have been shown to not be a food source for generalist herbivores. Although calcification can provide algae with a strong structural protection against potential predators, some herbivores have coevolved specialized feeding structures that allow them to macerate even very tough materials. However, the dominance of *Halimeda* spp. in reef systems is enhanced through its production of a potent defensive and antifouling substance, a diterpenoid natural product named halimedatrial (**Figure 7**). Halimedatrial is a trialdehydic secondary metabolite that is structurally similar to the iridoid aldehydes and the insect antifeedant warburganal. Halimedatrial displays a broad spectrum of activity against marine bacteria, a marine fungus, the division of fertilized sea urchin eggs, and the motility of sea urchin sperm. Halimedatrial was also toxic to common reef fish as well as being inhibitory of reef fish feeding at ecologically relevant doses (Paul *et al.*, 2007).

In the alga, halimedatrial exists in a "protoxin" form, a protected tetraacetate named halimedatetraacetate. Field experiments with *Halimeda* sp. have shown that halimedatetraacetate is converted to halimedatrial on wounding of the alga. This conversion is rapid and specific to the area of herbivory or wounding and does not take place simply on introduction to air but is an enzymatic conversion of a "protoxin" secondary metabolite to a highly potent antiprediatory substance. This activated defense of *Halimeda* represents the first to be described in the marine environment. Interestingly, a similar activated defense system that produces feeding deterrents has previously been described in Russulacean mushrooms. Additionally, it has recently been shown that dialdehydes such as halimedatrial can crosslink proteins, and thus serve a dual function in coenocytic green algae (the entire thallus of *Halimeda*, like many other green algae, is constructed from a single

connected cell!). Specifically, on mechanical damage to the tissue, conversion of halimedatetraacetate to halimedatrial occurs, thereby cross-linking proteins and providing a clotting mechanism so that cellular contents are retained.

Examples from Terrestrial Plants Illustrating Plant Sources of Drugs and Chemicals

Anticancer Agents from Higher Plants

Paclitaxel

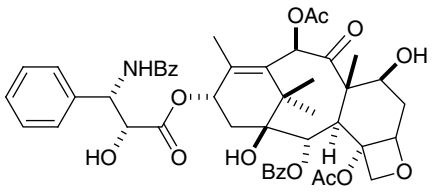
The story of the discovery and development of paclitaxel as an anticancer agent is a superb example of the opportunities and problems inherent in plant-derived medicines (Grothaus *et al.*, 2010). An NCI-sponsored plant collection program in the Pacific Northwest region in 1960 gathered 650 plant species for screening through various cell-based and *in vivo* anticancer evaluations. The twigs and bark of the Pacific Yew, *Taxus brevifolia*, showed very good activity in a cell assay and was subsequently subjected to bioassay-directed isolation of the active compound by Wall and Wani at the Research Triangle Institute. *Taxus* species have a rich history of use by native peoples for treatment of various illnesses, including cancer. Painstaking efforts were involved in the isolation of the active component, paclitaxel, which was shown to be a new taxane derivative by X-ray crystallography (**Figure 8**). Despite its intriguing structure and impressive activity in cell assays (but not in antileukemic *in vivo* assays!), research on paclitaxel languished for nearly a decade due to a combination of short supply of the drug, difficulty in working with this lipid soluble substance, and lack of research funds.

Interest in the drug was rekindled when Susan Horwitz at the Albert Einstein Institute showed that paclitaxel had the opposite effect on microtubules compared with all other known classes of antimitotic anticancer agents. By binding to a unique drug-binding site on the β-subunit of tubulin, paclitaxel promotes the stabilization of microtubules whereas other classes of antitubulin agents (e.g., vinca alkaloids, colchicines, and curacin A) promote their destabilization (Alfermann, 2010). This mechanistic difference fueled additional evaluations of paclitaxel, and once formulation problems were overcome, paclitaxel emerged as the most significant new anticancer drug of the past 25 years. Currently, paclitaxel has shown good efficacy to ovarian, breast, and lung cancers, as well as several others. However, its development was greatly impaired from severe supply shortages, a common issue for natural product-derived drugs. It has been estimated that it takes the recoverable paclitaxel content of four full grown 70-year-old trees to treat a patient for 1 year. At its peak, natural collections of 60,000 pounds of Pacific Yew bark were made in 1988 and 1989, which threatened the survival of this relatively common tree. Subsequently, supply of the drug was made plentiful by harvest of the needles from cultured trees (*Taxus baccata*), extraction of the essential core of the paclitaxel molecule, and its conversion by chemical synthesis into paclitaxel. More recently, plant cell culture techniques have been used to provide a steady and economical supply of paclitaxel (Alfermann, 2010). This latter process involves growing cells derived from *Taxus* in fermentation, followed by extraction and purification of the natural product. This fermentation process is attractive because the actual natural product is isolated, rather than an intermediate, thus eliminating most of the chemical reagents and solvents used in the semisynthetic approach. The 25-year development period from isolation to clinical use is typical of modern pharmaceutical development wherein many basic scientific, regulatory, and safety issues must be answered before drugs reach the market. The paclitaxel supply problem has subsequently initiated a national program to explore how to resupply a natural product drug in large scale during its development as an anticancer agent (Marasco and Schmidt-Dannert, 2007).

Camptothecin

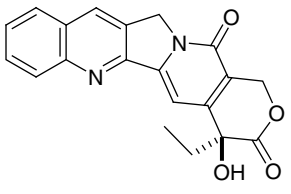
The same group that discovered paclitaxel made a second major contribution in the fight against cancer. This latter compound, called camptothecin, comes from the wood and bark of the tree *Camptotheca acuminata*. Monroe Wall, who was working at the eastern regional Research laboratory of the USDA in Philadelphia, Pennsylvania, discovered it in the 1950s. Previously, this group had been screening plant extracts for steroids that could be used as cortisone precursors. In 1958, samples from this entire collection of extracts were sent to the NCI for testing against certain cancer cell lines. Only one extract, that from *C. acuminata*, showed activity in this particular biological assay (Gunatilaka, 2009).

Camptothecin was eventually isolated through the use of the Craig Countercurrent Distribution method, which takes advantage of different compounds having varying solubilities in nonmiscible solvents. In 1966, the structure of camptothecin and its corresponding sodium salt were published on the basis of a thorough spectroscopic and chemical degradation study (**Figure 9**). In early animal trials, camptothecin showed significant antileukemic activity to the L1210 cell line at concentrations as low as 0.2 mg kg^{-1} body weight of mice. By 1970, camptothecin had progressed to phase I clinical trials at NCI and soon progressed to phase II clinical trials.



Taxol

Figure 8 Structure of the anticancer natural product "paclitaxel" from the tree *Taxus brevifolia*.



Camptothecin

Figure 9 Structure of the anticancer natural product "camptothecin" from the tree *Camptotheca acuminata*.

Encouraged by these results, the NCI decided to introduce the water-soluble sodium salt of camptothecin into phase II clinical trials. The results of these trials were not very promising and camptothecin research lost much of the interest it had attracted in its earlier phase I trials. It was later discovered that the sodium salt of camptothecin was only one-tenth as active as the parent compound against the P388 cancer line.

In 1985, it was found that camptothecin kills cancer cells by a novel mechanism and interest was rekindled. These later studies showed that camptothecin and various water-soluble analogs worked by inhibiting the action of mammalian topoisomerase I (T-I, an enzyme in mammalian tissues responsible for unwinding DNA during the replication process) by a mechanism never before observed. This coincided with later discoveries that T-I is over-expressed in a variety of advanced cancerous tumors.

In 1991, two water-soluble analogs of camptothecin, topotecan and irinotecan, were introduced into advanced phase II clinical trials in the USA Within the next few years topotecan was also approved for use in Japan and other analogs were approved for phase I clinical trials in Europe. Subsequently, topotecan and irinotecan were approved for use in certain types of ovarian and breast cancers in the USA.

Bioactive Alkaloids: Tropane Alkaloids

Alkaloids are a very diverse and abundant class of natural product. Their chemical structures range from relatively simple to very complex; but all contain one important theme, a basic nitrogen atom, usually heterocyclic, within their structural formula. Alkaloids are not only diverse in chemical structure, but are also diverse in the organisms that produce them. Alkaloids are categorized into different groups depending on the different moieties within their chemical structures. One of the best-known subclasses is the tropane alkaloids. The name is derived from the inclusion of a nitrogen-containing bicyclic appendage, tropane. Historically, two of the most important drugs in this class are atropine and cocaine (**Figure 10**). The tropane alkaloids continue to be vigorously investigated by many laboratories throughout the world.

Atropine was originally characterized from the poisonous plant *Atropa belladonna*. Aside from its poisonous properties when ingested, early usage of this plant involved placing the juice from the berry into the eye, causing dilation of the pupil. Women of the sixteenth century thought this to be a beauty-enhancing quality, hence the plant name "belladonna" ("beautiful lady" in Italian). In modern times, atropine has been used to counteract the effects of cholinesterase inhibitors. These inhibitors, such as physostigmine and organophosphate insecticides, act as competitive inhibitors of acetylcholine at the muscarinic receptor. Numerous synthetic routes have been developed to produce atropine; however, none are economically feasible. At the present time, the most viable economic means by which to obtain this useful alkaloid is still by isolation from the raw plant material. Atropine can be isolated from a number of different plant sources (e.g., *Hysoscyamus muticus*).

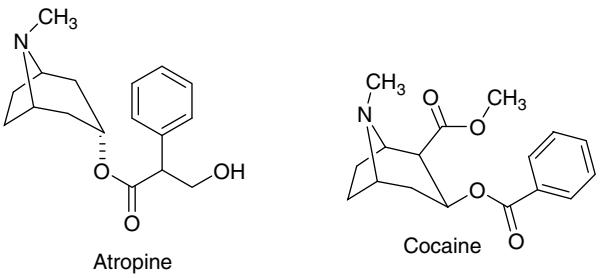


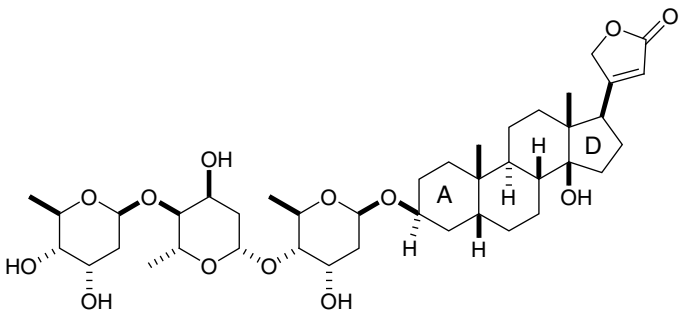
Figure 10 Structures of the tropane alkaloids atropine and cocaine.

Another tropane alkaloid, cocaine, was first isolated from *Truxillo coca* in 1860 with its intentional use in western medicine as an anesthetic. In Incan civilization this shrub was known as the "The Divine Plant." Natives would chew the leaves (mixed with lime so as to solubilize the alkaloidal substituents) in order to endure long journeys or extensive manual labor without experiencing the inherent fatigue associated with great exertion. In present-day times, cocaine's highly addictive stimulatory effect has given rise to many societal problems, in addition to retaining its importance as an anesthetic. Chemically, cocaine is a tropane alkaloid containing a methyl ester appendage. The primary mechanism of action of cocaine is that of a dopamine reuptake inhibitor. In essence, this leaves dopamine at the synapse for a longer time, giving rise to the euphoric sensation. Research on cocaine is still very actively pursued across the globe, from synthetic derivatives to effects on prenatal exposure to the drug.

Other Bioactive Classes: Steroidal Glycosides – *Digitalis*

Original therapeutic usage of *Digitalis* sp., commonly known as foxglove, can be found dating back to tenth century England, Wales, and Europe where it was used as an expectorant and treatment for epilepsy, swelling, and sore throats. Officially entering the London Pharmacopeia in the mid-1600s, and subsequently the pharmacopeia of most other countries by the eighteenth century, the late 1700s saw the emergence of *Digitalis* as a remedy for "dropsy." This disease condition is characterized by a swelling that results from inadequate heart function and is known today as congestive heart failure. The utility of this plant was introduced to Western medicine by William Withering in 1775, who was led to the medicinal value of *Digitalis* through conversations with a Welsh woman known for healing dropsy using herbal remedies containing foxglove. The active components of foxglove are classified as cardiac glycosides, two of which give rise to the drugs digoxin and digitoxin, the favored treatment for rapid atrial fibrillation (**Figure 11**).

Cardiac glycosides are naturally occurring steroids that are functionalized by sugar and lactone moieties. The most active of these possess an unsaturated lactone ring connected to C-17 of the steroid nucleus. Optimal cardiac activity also requires *cis* stereochemistry of the A/B and C/D ring junctures. Structure–activity studies on the cardiac glycosides have demonstrated that the sugar moiety aids in bioavailability and solubility, whereas functionalization of the steroid nucleus with hydroxyl, methyl, and lactone ring groups is essential for pharmacological activity. The unsubstituted aglycone, or steroidal portion, is less active than the glycoside (sugar attached).



Digitoxin

Figure 11 Structure of the cardiac glycoside digitoxin obtained from the foxglove *Digitalis* sp.

In treating heart failure, the cardiac glycosides of *Digitalis* function by increasing the strength and efficiency of ventricular contraction, thus shortening the length of contraction and allowing the heart muscle a longer relaxation time between contractions. This results in recovery of the myocardium, decreased heart rate, and improved renal function through enhanced circulation.

Nearly 30 different cardiac glycosides have been identified in *Digitalis purpurea* and more than 70 from *Digitalis lanata*. Cardiac glycosides have been observed in nine other *Digitalis* species. Interestingly, none of the major glycosides of *D. lanata* are identical to those isolated from *D. purpurea*, with the most common alteration in the molecule being acetylation of hydroxyl groups in the *D. lanata* representatives. This fact indicates great chemical, biochemical, and metabolic diversity among *Digitalis* species.

Historically, a drawback of *Digitalis* treatment is its toxicity. This toxicity led to its being eschewed by most medical practitioners until its reappearance in relatively modern treatments. The fact that no fewer than nine prescription products from *Digitalis* species are used to treat congestive heart failure, atrial fibrillation, and atrial and ventricular tachycardia demonstrates the human need and utilization of not only the innocuous plants, but also the noxious varieties, such as *Digitalis*. Moreover, the story of *Digitalis* may well have been the origin of the famous adage in the field of pharmacy, namely that – The difference between a poison and a medicine is simply a matter of dose.

Herbal Remedies

Ginseng

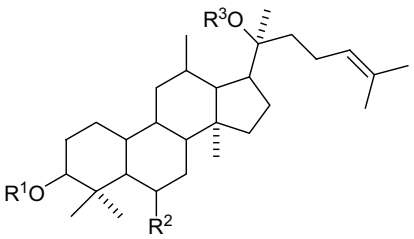
Perhaps the most venerable of traditional Chinese herbs is "ginseng" or "Korean ginseng," also known by its Latin name *Panax ginseng*. Usage of *P. ginseng* has a long and rich history in China dating back to the Han dynasty some 2000 years ago. The earliest description of its application appeared in the oldest Chinese pharmacopoeia, believed to have been written in the first century during the late Han dynasty. It is the root of *P. ginseng* that is used extensively in Chinese medicine as an effective tonic for enhancing stamina as well as the capacity to endure fatigue and physical stress.

Panax ginseng was originally found in northeastern China, Korea, and eastern Siberia. Other congeners of ginseng, such as the *Panax quinquefolius* (American ginseng) are also used widely as medicines. *Panax quinquefolius* was initially reported growing wild in the northeastern USA. Another common member is the Siberian ginseng (*Eleutherococcus senticosus*), which belongs to the Ariliaceae family and contains distinctly different glycosides. The main source of *P. ginseng* used in the USA and Europe comes from Korea where it is cultivated extensively. Two main forms of ginseng are currently being used; white ginseng, which is processed from the air-dried roots of five to seven year old plants, and red ginseng, which is white ginseng that has been steam treated for 2–4 h.

Research on the chemistry of *P. ginseng* suggests that the major active components of the roots are the ginsenosides (glycosides), which are derivatives of the triterpene "dammarane." To date, more than 20 ginsenosides have been reported from the roots, leaves, and flower buds of ginseng (**Figure 12**). Common sugars found attached to the protopanaxadiol core of these ginsenosides include glucose, maltose, fructose, and saccharose. In spite of the discoveries of these bioactive glycosides in ginseng, the precise pharmacological mechanisms of ginseng's actions are unclear. However, there are extensive reports of the effects of ginseng on the function of the neuroendocrine system, the central nervous system (CNS; memory, learning, and behavior), carbohydrate and lipid metabolism, the immune system, and the cardiovascular system. For instance, ginsenoside Rbl was demonstrated to have CNS-sedative, tranquilizing, and hypotensive actions whereas ginsenoside Rgl was shown to be CNS-stimulating, hypertensive, and possess antifatigue properties. As this example demonstrates, many of the reports on the biological activities of ginseng's metabolites are contradictory; this may be due to the difference of ginsenoside content in ginseng root or root extracts.

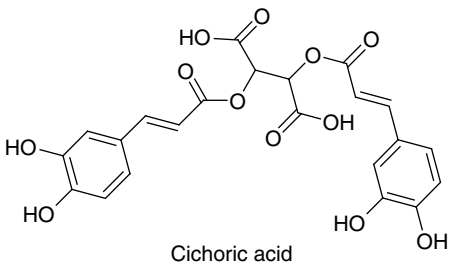
Echinacea

The application of *Echinacea* as a medicinal plant is derived from its use by Native North Americans. Traditional use of *Echinacea* involved the treatment of numerous conditions, such as external applications to wounds, burns, inflammation of lymph nodes (mumps), and insect bites. Internal uses of the plant included stomach cramps, headache, coughs, chills, measles, and gonorrhea. The root of *Echinacea* was the most important part of the plant used in these native treatments. There are also reports of the use of the juice or a paste of macerated fresh plant by Native North Americans. In view of



	R^1	R^2	R^3
20(s)-protopanaxadiol	Glc^2- Glc	H	Glc^6-Glc
20(s)-protopanaxatriol	H	O- Glc^2- Rha	Glc

Figure 12 Two of a number of ginsenosides from *Panax ginseng*, a traditional herbal remedy from China.



Cichoric acid

Figure 13 Structure of cichoric acid, one of the biologically active components of *Echinacea*.

the multitude of beneficial uses of *Echinacea* by these people, it was likely one of their most important medicinal plants.

The first pharmacological studies on *Echinacea* extracts began in the early 1950s and were mainly concerned with its nonimmunological actions. However, we now know that *Echinacea* acts mainly *via* stimulation of the nonspecific immune system. The main immunostimulatory components of *Echinacea* extracts were found to be lipophilic alkylamides as well as cichoric acid, which is a derivative of caffeic acid. Cichoric acid (2,3-*O*-dicaffeoyltartaric acid) was first isolated from *Echinacea purpurea* and found to cause significant stimulation of phagocytotic activity in an *in vitro* granulocyte bioassay (**Figure 13**). In addition, polysaccharides have also been suggested to be active components in *Echinacea* juice and aqueous extracts. A purified mixture of polysaccharides from the roots of *E. purpurea* were found to increase *in vitro* phagocytosis by macrophages 23–32% at concentrations of 0.01 and 0.001 mg ml^{-1}. It appears that the complete immunostimulatory properties of *Echinacea* extracts are due to the combination of activities from several different classes of compounds.

Conclusions

The examples given here clearly demonstrate the pivotal role that plants have played in the development of much of our current pharmacopeia. Indeed, it has been estimated that as much as 37% of all pharmaceutical sales are for compounds that derive, either wholly or in part, from natural products. For anticancer drugs, the percentage deriving from natural products is even higher (approximately 60%). Nevertheless, the enormous task that we require of pharmaceuticals, to have potent and selective action for a particular disease state, to be orally active and distribute with high efficiency to the target organ, and to be nontoxic to normal cells and tissues of the body, makes the development of any new pharmaceutical a long-term affair. On average, it is estimated that it requires nearly 15 years from the time of discovery to the time of clinical application of a new drug entity, a process estimated to cost approximately 1 billion US dollars in 2010.

Recognition of the enormous long-range value of plant-derived natural products, and the genetic sequences encoding for them, is one of the most compelling pragmatic arguments for biodiversity preservation. Indeed, the International Cooperative Biodiversity Group program in the USA has the simultaneous objectives of natural product drug discovery, biodiversity preservation, and developing short- and long-term economic benefit from ecologically preserved habitats, mainly in the tropics. Indeed, perhaps one of the strongest arguments in favor of protecting and promoting a rich biodiversity is provided by the direct socioeconomic benefit that is realized from plant species and their rich natural products chemistries.

As our understanding of the molecular basis for various diseases increases and defines new drug targets, we will need an ever-increasing source of molecular diversity to act as starting points for developing new generations of more potent and more selective drug agents. Nature's storehouse is still the largest and most chemically diverse source of unique molecular architectures. Preservation of this resource, and of the genes that encode for it, should be an overarching societal priority.

See also: Biodiversity as a Commodity. Bioprospecting. Ecosystem Services. Pharmacology, Biodiversity And. The Value of Biodiversity

References

Alfermann A-W (2010) Production of natural products by plant cell and organ cultures. *Annual Plant Reviews* 39: 381–399.

Appendino G and Pollastro F (2010) Plants: Revamping the oldest source of medicines with modern science. In: Buss AD and Butler MS (eds.) *Natural Product Chemistry for Drug Discovery*, pp. 140–173. Cambridge, UK: Royal Society of Chemistry.

Blunt JW, Copp BR, Munro MHG, Northcote PT, and Prinsep MR (2010) Marine natural products. *Natural Product Reports* 27: 165–237.

Folmer F, Jaspars M, Dicato M, and Diederich M (2010) Photosynthetic marine organisms as a source of anticancer compounds. *Phytochemistry Reviews* 9: 557–579.

Gerwick WH, Coates RC, Engene N, *et al.* (2008) Giant marine cyanobacteria offer exciting pharmaceutical potential. *Microbe* 3: 277–284.

Grothaus PG, Cragg GM, and Newman DJ (2010) Plant natural products in anticancer drug discovery. *Current Organic Chemistry* 14: 1781–1791.

Gunatilaka AL (2009) Natural products in plants: Chemical diversity. In: Begley TP, (ed.) *Wiley Encyclopedia of Chemical Biology,* vol. 3, pp. 261–277. Hoboken, NJ: John Wiley and Sons, Inc.

Jones AC, Gu L, Sorrels CM, Sherman DH, and Gerwick WH (2009) New tricks from ancient algae: Natural products biosynthesis in marine cyanobacteria. *Current Opinions in Chemical Biology* 13: 216–223.

Marasco EK and Schmidt-Dannert C (2007) Exploring and accessing plant natural product biosynthesis in engineered microbial hosts. In: Kayser O and Quax W, (eds.) *Medicinal Plant Biotechnology,* vol. 2, pp. 287–317. Weinheim, Germany: Wiley-VCH Verlag.

Paul VJ, Arthur KE, Ritson-Williams R, Ross C, and Sharp K (2007) Chemical defenses: From compounds to communities. *Biological Bulletin* 213: 226–251.

Pollinators, Role of

David W Inouye, University of Maryland, MA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Phenology The timing of recurring plant and animal life cycle stages, such as flowering or emergence of pollinators.

Poricidal anthers Anthers that function similarly to salt shakers; they must be vibrated at the right frequency (by buzz pollinators) to dispense pollen grains through the pores at the end of the anthers.

What is Pollination?

Pollination is the transfer of a pollen grain (male gametophyte) to a flower's stigma (receptive surface of the female reproductive organ), where it may germinate, grow through the style, and fertilize an ovule to produce a seed. This transfer can be accomplished by abiotic agents such as wind and water, but the majority of pollination is effected by animal pollinators seeking nutritional rewards such as pollen and nectar, or sometimes other resources such as warmth, floral perfumes, oils, or resins, or even reproductive opportunities through mimicry of potential mates by ornate flowers. The relationship between the plants and pollinators is commonly assumed to be mutualistic, with the plants benefiting from the transfer of pollen and the pollinators receiving a nutritional or other reward, but there are also instances in which the plants appear to provide no reward and cases in which pollinators may also be seed predators. This rich diversity of relationships has been fertile ground for investigation by pollination biologists for hundreds of years. The importance of pollination in agriculture has also been a powerful stimulus for the study of plants and pollinators.

There is an important distinction to be made between "flower visitor" and "pollinator." Visitors to flowers (anthophilous animals) cannot be assumed to be pollinators, as in reality they may be nectar or pollen thieves which, owing to a mismatch in morphology or an unusual behavior, do not pollinate while removing floral resources. For example, "base workers," or insects that remove nectar from between the petals of a flower with an unfused corolla, or insects too small to contact the reproductive parts of a flower, would not be pollinators despite the fact that they may spend a lot of time in harvesting nectar. However, their activities may be detrimental to the flowers' pollination, as pollinators may be deterred by finding no nectar in flowers already visited by nectar thieves. Similarly, a small insect that collects pollen from anthers but never contacts stigmas is an example of a pollen thief that does not pollinate. Simple experiments, such as collecting stigmas of flowers to confirm pollen deposition after a visit, or the observation of transfer of a pollinium (a packet of pollen, characteristic of flowers in the Orchidaceae or Asclepiadaceae), can help to distinguish between visitors and pollinators. Unfortunately, many animals observed on flowers are probably best categorized as flower visitors because this kind of confirmation of pollination has not been conducted. It is surprising that much remains to be learned about the pollinators of many important agricultural species, in addition to other plants.

Although the origins of much of what we know about pollination biology lie in the realm of natural history, pollinators are now studied from many scientific perspectives, including foraging behavior, physiological ecology, coevolutionary interactions, and the perspective of trophic network theory. The significance of native pollinators in agriculture and the conservation biology of endangered pollinator species are now also commonly studied. Techniques such as telemetry, harmonic radar, and even GPS tags are now providing insights into pollinators that were not previously possible. For all these reasons pollinators are appreciated for the economic value of their ecosystem service of pollination, their esthetic value, and tractability for experimental studies.

The Diversity of Pollinators

The list of species of animals that serve as pollinators is long and diverse. The largest group of pollinators is insects, but both flying and nonflying mammals, birds, and at least one reptile have been recorded as pollinators. **Table 1** shows a list of pollinator classes for the world's wild flowering plants (the angiosperms; ~240,000 species), and estimates of the number of species in each. Much work remains to be done to confirm the activities of these taxa as pollinators, but these numbers are the first estimates available.

Invertebrate Pollinators

Although flower visitation has been observed by species from at least 16 orders of insects, only four of them include many species that regularly pollinate flowers and that seem to have been involved in coevolutionary interactions with plants. These are the beetles (Coleoptera), flies (Diptera), butterflies and moths (Lepidoptera), and ants, bees, and wasps (Hymenoptera). In addition, the thrips (Thysanoptera) include some pollen-eating species that may be pollinators, some stoneflies (Plecoptera) and true bugs (Hemiptera) will visit flowers and eat pollen and nectar, and some lacewing (Neuroptera), scorpionfly (Mecoptera), and caddisfly (Trichoptera) species eat nectar. Additional work is needed to confirm whether most of these lesser-known flower visitors are indeed pollinators.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00112-X

Table 1 Pollinator classes for the world's wild flowering plants (~240,000 angiosperm species)

Pollination categories	*Estimated pollinator taxa*
Wind (abiotic)	20,000
Water	150
All insects	289,166
Bees	40,000
Hymenoptera (bees and wasps)	43,295
Lepidoptera (butterflies and moths)	19,310
Diptera (flies)	14,126
Coleoptera (beetles)	211,935
Thrips	500
All vertebrates	1221
Birds	923
Bats	165
Mammals other than bats	133

The numbers are "the number of species comprising the invertebrate and vertebrate genera, families, and orders in which there are more known effective pollinators than there are pollen or nectar cheaters, robbers, or avoiders".

Source: Reproduced with permission from Nabhan GP and Buchmann SL (1997) Services provided by pollinators. In: Daily GC (ed.) *Nature's Services. Societal Dependence on Natural Ecosystems*, ch. 8, pp. 133–150. Washington, DC: Island Press.

Beetles

Beetles have been abundant since at least the Mesozoic, and it is likely that some of them have been flower visitors since the origin of the earliest angiosperms. For example, the current association of beetle pollination (cantharophily) with primitive woody angiosperms (*Magnolia* is an example) probably dates back to the evolutionary origins of both groups. Beetle pollination is considered to be the most primitive type of pollination by animals, and is not very important in cool temperate regions. It is most common in the moist tropics, and to a lesser degree in arid areas. Beetles constitute the largest order of insects, and some of this diversity is thought to have arisen through the same evolutionary radiation of flowers and insects during the Tertiary that led to the origin of the other major orders of flower-visiting insects.

Flies

Flies are probably the second most common order of flower visitors (after Hymenoptera). Fly visitors from at least 71 families of Diptera to flowers in 137 plant families have been recorded in the literature (compilation by Inouye). The families Syrphidae (hoverflies), Bombyliidae (bee flies), and the Muscidae are especially common as flower visitors. In tropical areas, the diversity of flower-visiting Diptera can rival or exceed that of the Hymenoptera. For example, 4856 species of Diptera are recorded from Australasia from flower-visiting families, compared with ~2570 bees (superfamily Apoidea), whereas for the Neotropics the estimates are >2940 species for Diptera and 5630 species for bees (Roubik, 1995). In some parts of the world, including the Faroe Islands, New Zealand, and Australia, where bees are nonexistent or relatively rare, flies have filled that ecological niche of pollinators.

Pollination by flies (myophily) is economically important; in tropical areas flies are the primary pollinators of cacao and also pollinate mango, cashew, and tea. Roubik (1995) lists pollinators of 785 species of cultivated plants in the tropics, and 26–31 of these plants are apparently pollinated only by flies, 32–33 by flies as the primary pollinators, and 87–101 more by flies as secondary pollinators. In Europe, flies are used commercially to pollinate protected crops of onion, chive, carrot, strawberry, and blackberry. In both tropical and temperate areas, flies appear to be more important as pollinators than has been generally recognized, and there is a need for additional research on them.

In addition to the more specialized flower-visiting flies, carrion, and dung flies are pollinators of some plants, in a category of pollination called sapromyophily. The basis for attraction of the flies is floral fragrances that mimic rotting flesh or dung, and typically the flies do not receive any nutritional reward for their visits. Some of the most spectacular flowers in the world are fly pollinated, such as the inflorescence of the titan lily (*Amorphophallus*, not really a lily), which can reach 10 ft tall and is called the corpse lily for its fearsome odor; it attracts and is pollinated by flies. Another example is the largest flower in the world, *Rafflesia*, found in Borneo, which can reach a meter in diameter and smells like rotting meat. Blood-feeding Diptera such as mosquitoes and biting flies (Tabanidae; horseflies, deerflies) are also sometimes pollinators, attracted to flowers by the nectar they offer.

Some flies have remarkably long tongues that permit them to collect nectar from correspondingly long corolla tubes. A tabanid fly in South Africa that visits flowers with long corolla tubes has a proboscis length of up to 35 mm, whereas another species in the Nemestrinidae has a proboscis up to 57 mm (and flies with the proboscis folded under its body and sticking out behind). Some bee flies (Bombyliidae) also have relatively long tongues.

Flies provide one of the few examples of a system in which a flower species (globeflower, *Trollius*) has an obligatory pollinator species (several species of *Chiastocheta* flies) whose larvae are also seed predators, consuming a proportion of the seeds. There are some interesting evolutionary dynamics that lead to a stable mutualism between globeflowers and the globeflower flies.

Butterflies and Moths

The Lepidoptera (butterflies and moths), as a consequence of their two distinct lifestyles as primarily herbivorous larvae and commonly flower-visiting adults, have dichotomous relationships with plants. In at least one case, the pollination of yucca, the act of pollination by yucca moths is accompanied by oviposition and the larval moths are seed predators. Although not all Lepidoptera feed as adults, most, if not all, do depend on nectar as a source of both sugars and amino acids. Some of the longest lived of all butterflies, the neotropical *Heliconius* species, also feed regularly on pollen by collecting it on their proboscis and regurgitating nectar on the pollen mass to leach out amino acids. Female butterflies of some species are dependent on the amino acids and sugars they collect in nectar to provide nutrients needed for making eggs, so the more nectar they can collect the more eggs they can lay. The floral rewards sought by Lepidoptera are not always nutritional; some species are attracted to flowers by the pyrrolizidine alkaloids that they produce, which the male butterflies require for the production of pheromones, for mating success, and for the establishment of multispecies leks. Lepidoptera, as

agents of natural selection, may be responsible for much of the diversity of defensive chemicals in plants (to deter herbivorous caterpillars) as well as diversity in flowers and the rewards they offer.

Two families of moths stand out as pollinators. One is the hawk moths (Sphingidae; also called sphinx moths), and the other is the noctuid moths (Noctuidae; the largest family in the order Lepidoptera). The hawk moths are strong and agile fliers that maintain very high body temperatures (e.g., as high as 46 °C) while in flight. Some diurnal species are easily mistaken for hummingbirds (e.g., the European hummingbird hawk moth, *Macroglossum stellatarum*). Some hawk moths are migratory, and (in successive generations) may move hundreds of kilometers during a summer in North America. Hawk moths include the pollinators with the longest known tongues; most of these extraordinary species are found in Madagascar. Darwin measured nectar spurs in some orchids from Madagascar that were 29 cm long, and concluded that there must be moths there that have proboscis lengths of the same length. Forty years after his prediction, a hawk moth with a proboscis that reaches greater than 24 cm in some individuals was found and given the name *Xanthopan morgani praedicta*. There is actually a large guild of long-tongued hawk moths and flowers with matching corolla lengths in Madagascar.

Some noctuid moths are long-distance migrants, flying distances as far as 1600 km. One species is reported to travel as much as 1000 km in a day when migrating from the Baltic States to Britain. This presents the opportunity for some very long-distance gene flow on the part of their food plants! Some species migrate vertically, moving up in altitude some time during the year to avoid unfavorable seasonal weather. Before these migrations, the moths may store substantial quantities of fat, and at least one such species was an important part of human diets; the bogong moth (*Agrostis infusa*) was collected while estivating in the Snowy Mountains by aboriginal Australians.

Ants, Sawflies, Wasps, and Bees

The Hymenoptera (ants, sawflies, wasps, and bees) are the largest and most diverse group of pollinators. Ants are only rarely recorded as pollinators, although they are not uncommon as flower visitors collecting nectar. One suggestion for why there are not more species of plants adapted for pollination by ants is that the antibiotic and other secretions that ants have on their exoskeletons are detrimental to pollen survivorship and growth. Sawflies are more common as pollinators. They have herbivorous larvae, and in some cases the adults restrict their flower visits to the larval host plant. The adults may also eat flower parts in addition to nectar and pollen. Flower visitation by wasps has been studied in part because of interest in attracting and sustaining populations of parasitoid wasps (which prey on pest insects) in agricultural situations. But predatory and social wasps have also been recorded as flower visitors and a few species of plants are specialized for pollination by pseudocopulation by male wasps (*see* Pollination by Deception). Some of the most specialized relationships between plants and pollinators are those between species of figs and the wasps that pollinate them; many of those relationships are reported to involve single species of wasps pollinating each fig species. The wasp larvae develop in some of the seeds produced by the pollinating activity of their mothers.

There are fewer species of bees than beetles, but a much higher proportion of bees are pollinators than are beetles. Almost all species of bees, both adults and larvae, are dependent on flowers for pollen and nectar as nutritional resources, and of all the insects bees are the most highly adapted for flower visitation and pollination. Their behavior, morphology, and senses of vision and smell all appear to be adapted for finding and collecting floral resources. For example, the trichromatic color vision of bumble bees and honey bees includes ultraviolet, which allows them to see the ultraviolet reflectance patterns that many flowers use as nectar guides or a petal color. Humans cannot see these patterns without the help of cameras.

Although pollen (primarily for feeding larval bees) and nectar (both a larval and adult resource) are the most common resources collected by bees, some flowers offer oil as a reward. The oil is collected by bees in the family Anthophoridae; females of these species have absorbent brushes for holding the oil and sharp edges on their legs that are used to squeeze the oil out so that it can be mixed with pollen as food for larvae. Some bees collect resin as a floral reward, which they then use in nest construction.

Pollination by bees (melittophily) is very important for agriculture (*see* Pollinators in Agriculture and their Economic Value). The social bees are the best-known crop pollinators, and are the species most commonly managed for pollination or honey production. These include honey bees (*Apis*), bumble bees (*Bombus*), and stingless bees (*Trigona*). The presence of large quantities of honey (concentrated nectar), pollen, and larvae makes social bee nests an attractive target for insects and vertebrates that prey on nests and adult bees. Bee-eaters and honey guides are two examples of Old World birds that specialize on such resources, and honey and wax are also attractive resources collected from wild colonies by humans in some parts of the world.

Vertebrate Pollinators

Mammals

Although invertebrates are the most common pollinators, mammals, birds, and lizards have also been documented as pollinators. Mammalian pollinators include both bats, some of which primarily eat nectar and pollen, and nonflying species such as some Australian marsupials. Most bat pollination occurs in the tropics, but some migratory bats are important pollinators in temperate North America to about 30° N. In the Old World species of the family Pteropidae, fruit bats, are important pollinators, whereas in the New World flower-visiting bats are confined to the family Phyllostomidae. Although pollination by bats (chiropterophily) is geographically widespread, the other mammalian pollinators are more restricted and less common as pollinators, including lemurs in Madagascar and several species of Australian marsupials such as sugar gliders, honey possums, and some marsupial mice. There are scattered reports of other mammals, including mice and giraffes, which may serve as pollinators in unusual cases.

Vertebrate taxa other than mammals or birds are only rarely reported as pollinators; there are a few reports of lizard pollinators on islands.

Birds

Bird pollination (ornithophily) is common in many parts of the world. Several families of birds are primarily nectarivorous and undoubtedly include important pollinators. These include the hummingbirds (Trochilidae; New World), honeyeaters (Meliphagidae; Australia, New Zealand, and parts of Asia), sunbirds (Nectariniidae; Africa and Southwest Asia to the Philippines), sugarbirds (Promeropidae; South Africa), flowerpeckers (Dicaeidae; Asia and Australasia), and Hawaiian honeycreepers (Drepanididae; Hawaii). Other families that include some nectarivores or scattered records of flower visitation include the Thraupidae (honeycreepers and some tanagers; New World), Icteridae (orioles; New World), Zosteropidae (white-eyes; Africa, Asia, and Australia), and Psittacidae (lorikeets and hanging parrots; Southeast Asia and Australasia). A number of flower-visiting birds are recently extinct in the South Pacific islands, but in at least one case an introduced species has taken over the role of a pollinator for some of the endemic plants that were pollinated by the extinct birds.

Coevolutionary History of Plants and Pollinators

The "abominable mystery" that Darwin referred to, the origin and relationships among flowering plants, is still being revealed. Beetle pollination is thought to be the primitive condition in flowering plants, but by 100 Ma butterflies and moths had joined beetles as important pollinators. Recently discovered fossil evidence indicates that the origin of the angiosperms (flowering plants) may date back as far as the Late Jurassic, suggesting that the origin of the coevolutionary relationship between flowers and pollinators may also be older than previously thought. Much of the tremendous biodiversity found in the flowering plants is thought to have arisen through evolutionary interactions with pollinators, and certainly much of it is maintained through their flower-visiting activities that result in the production of seeds. The remarkable diversity of flower size, shape, odor, rewards, color, and pollination mechanisms is the result of these coevolutionary relationships.

Pollinators in Natural Systems

Obligate and Facultative Mutualisms

The relationships between plants and pollinators range from obligate to facultative in nature. Although one study in Brazil found that 43% of the plants studied were visited by only a single species of pollinator, other studies have found tens of visitor species to a single plant species. In only a few cases have researchers quantified the relative importance of different species in cases with multiple pollinators of a single plant species. A common finding of multiyear studies is variation, both temporal and spatial, in abundance of these different groups of pollinators. The visitors to a single plant species can include diverse taxonomic groups, for example, both vertebrate and invertebrate pollinators (e.g., hummingbirds and bumble bees). By having several groups of potential pollinators, it is more likely that these plants will not suffer from pollen limitation, or a shortage of pollination; if one group of pollinators is unusually low in abundance at a particular site or in a particular year, it is likely that another group will not be. Thus it can be important to have a diversity of pollinators available.

The obligate nature of some plant–pollinator relationships extends also to pollinators, as some insects appear to visit flowers of only a single species of plant. In such cases, there must be a good match in the phenology of life cycles of both plant and insect to ensure success for both partners, and one concern about global climate change is the potential for such phenological relationships to become mismatched. There are similar cases of plant species visited by only a single species of pollinator. Although a plant or pollinator may thus have an obligate relationship with a single partner, the partner may be involved in relationships with multiple species.

Although we assume that most plant–pollinator relationships are mutualistic in nature, some examples appear to involve a mixture of positive and negative interactions from at least the plant's perspective. Several studies have demonstrated seed predation or other forms of herbivory by insects that may also serve as pollinators. These examples include beetles and palms, moths and yuccas, wasps and figs, as well as flies and moths and some flowers that they pollinate. Some of these relationships (e.g., yucca flowers and their moths) can be farther complicated by the presence of closely related species of insects that are herbivores but not pollinators. Much detailed study is needed to dissect the complicated interactions in such relationships.

There is no doubt that long-lived or highly mobile pollinators must interact with a large number of plants, as their life spans or travels extend beyond the temporal or spatial availability of single plant species. Thus a group of plant species, even if they are not sympatric, could be considered mutualists if in concert they sustain and share a particular pollinator. This kind of multispecies relationship has significant conservation implications, as it may be necessary to conserve plant species flowering widely separated in time and space to conserve a (e.g., migratory) pollinator. This idea has led to the concept of a "nectar corridor" to sustain long-distance migratory pollinators. Hummingbirds are an example of a species that requires nectar resources along a long migratory route; many North American species overwinter in Mexico but reproduce during the summer as far north as Alaska. Monarch butterflies are another example, with individual butterflies traveling thousands of miles from summer breeding grounds to wintering areas in Mexico, and then partway back the next spring.

Cheaters in Pollination

Pollen and Nectar Robbers

The nutritional rewards that pollinators find in flowers, pollen, and nectar are sometimes harvested by flower visitors that

may pierce or bite holes in the flowers to obtain the resources "illegitimately." Some species of birds (*Diglossa*, the flowerpiercers) and bees (*Xylocopa*, carpenter bees; some short-tongued *Bombus*, or bumble bees, and some stingless bees) may obtain most of their nectar in this fashion. Once these primary nectar robbers have created holes, other species may learn to use them, as secondary robbers. Typically, these flower species that are robbed have long corollas, and the nectar robbers do not have mouthparts long enough to obtain the nectar without robbing. Although biologists at least as far back as Darwin have assumed that nectar robbing would have a significant negative influence on the robbed plants, more recent evidence suggests that in fact most insect nectar robbers that have been studied have either a beneficial or neutral effect on the plants. In the process of moving around on the flowers, or perhaps while collecting pollen from the same flowers, robbing insects may effect pollination. Even in the absence of such a direct effect, robbers may indirectly increase the fitness of plants they rob by influencing the behavior of legitimate pollinators. The reduced quantity of nectar in a robbed flower may induce a legitimate pollinator to visit more flowers or to fly greater distances between flowers, thereby increasing gene flow.

Pollination by Deception

Cheating in pollination biology is not restricted to the pollinators. Plants also provide fascinating examples of deception, attracting some pollinators by falsely advertising opportunities for feeding or sex. The orchid family is perhaps the best known in this regard, as many species of orchids have flowers with no nectar, and pollen that is not accessible or attractive to pollinators. In such species, it appears that the pollinators gain nothing from visiting (and pollinating), and quickly learn to avoid the flowers. Other orchids have flowers that, in both morphology and odor (production of insect sex pheromones), resemble female insects well enough that males attempt to copulate with them and in the process pollinate (pseudocopulation).

Pollinators in Agriculture and their Economic Value

Our knowledge of the pollination biology of crop species is surprisingly sparse. For example, the pollination requirements of about one-third of the crop species grown in the European Union are still unknown. These plants include at least 264 species that are cultivated as crops, or gathered from the wild, for human or livestock food, or for the essential oils they contain. The best-known pollinators of cultivated food plants are honey bees, which have been widely introduced around the world (*see* Conservation Biology and Pollination). However, honey bee populations in the US have been decimated in recent years by the introduction of two species of parasitic mites, and are also being threatened by introduction of a hive beetle. Most feral colonies of honey bees in the US have disappeared since the arrival of the *Varroa* and tracheal mites in the 1980s, and the number of managed colonies has declined to less than a half of its peak. This decline has resulted in a growing reliance on and appreciation of the roles of other, native, species in the US and elsewhere as pollinators of crops.

The rapid growth of the almond industry in California around the beginning of the twenty-first century has had a significant impact on commercial beekeeping of honey bees. Almonds require pollination to produce fruit, and the value of the crop is so high that almond growers are willing to pay significant fees (e.g., ~$140 per colony in 2010) for use of the bees during the few weeks of flowering. Beekeepers from as far away as Florida have moved bees to California to take advantage of these fees, creating shortages for some other parts of the country.

These problems with honey bee availability have also prompted some attempts to quantify the economic value of pollination services provided by both managed and wild pollinators. A recent model estimates that US consumers realize $1.6–5.7 billion in annual social gains that would be lost if honey bee services for 62 crops were reduced. The potential value of nonhoney bee pollinators in the US agricultural economy can be estimated at $4.1–6.7 billion yr^{-1}. An estimate of the global value of pollination services has been placed at $117 billion yr^{-1}.

Introduced Species of Pollinators

Undoubtedly, the largest-scale introduction of pollinators has been the global traffic in the European honey bee (*Apis mellifera*). Such continental-scale introductions have a long history. For example, European colonists in North America imported honey bees as early as 1641. This species went on to establish feral colonies over much of North America and, probably through subsequent introduction, to much of the neotropics. They were also introduced to Australia. Probably, the second largest introduction of pollinators has been the spread of bumble bees (*Bombus*). The first introduction of bees anywhere specifically for pollination was the introduction of European bumble bees to New Zealand in 1885 (to pollinate introduced clover for the introduced cattle), and subsequently they were also introduced to North America. The second successful introduction of a pollinator was a fig-pollinating wasp (*Blastophaga psenes*) into California in 1899.

In these cases, the incentive for introducing bees or wasps was for the pollination of introduced species of plants that were not attractive to or could not be pollinated by native pollinators (e.g., clover in New Zealand and figs in California). More recently, international transport of bumble bees has been the result of their use in pollination of greenhouse crops, especially tomatoes. In the 1980s, techniques were developed for the year-round commercial-scale propagation of bumble bee colonies (which normally have an annual life cycle of a few months). Commercially produced colonies of bumble bees have now largely replaced the tedious hand pollination of tomato flowers with electronic vibrators (tomato flowers produce no nectar, but are visited by bees that "buzz" the flowers to release pollen from poricidal anthers to collect it for feeding to their larvae, and in this process they pollinate the flowers) and the use of chemical sprays to induce fruit set.

In the early 1990s, a European species of bumble bee (*Bombus terrestris*) was introduced to Japan for use in greenhouse pollination; it has subsequently escaped and is now established in the wild. Although an effort is being made to

exterminate the introduced species in the wild (and to develop native species for use in greenhouse pollination), it is likely to be unsuccessful. The phenomenon of "pathogen spillover" has been documented recently for bumble bees, with wild bees found close to greenhouses having a higher incidence of disease than those further away; presumably, this results from contact with infected commercial bees that are foraging outside the greenhouse. There is some suspicion that the loss of a native North American bumble bee species (*Bombus occidentalis*) in much of its former range along the west coast may be a consequence of disease or parasites introduced by the cultivation of colonies in Europe from wild-caught queens that were then reintroduced to the US (having acquired the parasites or diseases during rearing in Europe). Two other species of North American bumble bees may have gone extinct in the past few years, perhaps because of introduced diseases.

Other species of pollinators have also been introduced. The wasp pollinators of figs have been introduced to New Zealand and the US, where they pollinate naturalized figs. The Japanese white-eye, an introduced passerine in Hawaii, has taken over the role of a pollinator of two native plants that lost their original pollinator to extinction. An Asian bee is now the primary pollinator of alfalfa plants grown for seed production in the US and a Japanese bee is used in the US and Canada for pollination of apples. The alfalfa leaf-cutting bee (*Megachile rotundata*) was introduced to North America to pollinate introduced alfalfa, and has also been introduced recently to Australia. Every year, millions of these bees are imported to the US from Canada, where they are raised and then shipped as pupae.

Pollinators as Vectors and Victims of Engineered Genes

As genetic engineering of crop plants has become more common, and as these crops are approved for cultivation in the field, concerns have arisen about the potential for gene flow through pollen transport between the engineered genomes and wild plants. One management tool to prevent, or at least minimize, this potential is to plant a buffer zone of unengineered plants around a field of engineered plants in hopes of intercepting pollinators carrying pollen before they can visit related native plants.

There is also the potential for pollinators to be victims of engineered genes. For example, laboratory studies of monarch butterfly caterpillars showed that if they eat significant quantities of corn pollen from plants engineered to have the *Bacillus thuringiensis* toxin in their leaves, it can kill them. Corn is wind pollinated, and a likely scenario is that large quantities of corn pollen could be distributed onto leaves of milkweed plants at the margins of cornfields. The potential for this to have detrimental effects on the monarch butterfly and other herbivores remains to be studied but may not be as serious as was first feared.

Conservation Biology and Pollination

Threats to Plant–Pollinator Mutualisms

Pollinators now face a large variety of threats of anthropogenic origin (Kearns *et al.*, 1998). These include habitat fragmentation, a variety of effects of agriculture, pesticides, and herbicides, introduction of parasites and diseases, and the introduction of both pollinators and plants. Fragmentation of habitats is likely to affect nonflying pollinators most strongly, as it may be difficult for them to locate and visit patches of flowers that are isolated by areas of different, perhaps nonpreferred, habitat. But even winged pollinators may be reluctant to leave a preferred habitat, such as undisturbed forest, to fly across a newly created pasture to another patch of forest. If the fragments are small enough to affect the size of pollinator populations, there may be negative genetic consequences for the smaller populations.

Modern agriculture can have a variety of negative impacts on pollinators. The creation of large fields that are disturbed regularly by plowing can prevent ground-nesting bees from establishing populations, and if only a single plant species is available there may not be enough pollen and nectar resources to support the life cycle of a pollinator species. The use of pesticides to control agricultural pests can have a negative impact on pollinators, and herbicides may remove species that could provide resources for the pollinators. Pesticide use in nonagricultural areas, such as spraying of large tracts of forest to control lepidopteran herbivores, has also been shown to have strong negative impacts on nontarget pollinator populations.

The introduction of exotic plants, such as weedy plants, could have potential implications for the pollination of native plants. If the exotics are close relatives to native species, there is the potential for hybridization. And if the exotics are prolific producers of nectar or pollen, they may draw pollinators away from native species that may then suffer a deficit in seed production. The potential for these kinds of consequences is not yet well understood, but this area is beginning to attract attention from researchers.

The ecological consequences of widespread introductions of *Apis mellifera* and *Bombus* for native plants and pollinators are not well known because of a lack of data from before the introductions. We can speculate that the introduction of honey bees to North America may have had significant consequences for species of bumble bees that have the same proboscis length, because the perennial honey bee colonies have many more workers than the annual bumble bee colonies. The spread of *Apis mellifera scutellata*, the African subspecies ("Africanized honey bees"), from its point of release in Brazil to as far north as the US was much more recent and has provided opportunity for some studies of their impact on native neotropical pollinators (relatively minor, it appears).

Species from every major group of pollinators, invertebrate and vertebrate, have been classified as endangered, and recent extinctions have been documented for others. Thus, pollinators are just as susceptible to the current human-induced mass extinction as any other group of organisms.

Potential Management Solutions

A first step in solving problems in the conservation of plant–pollinator relationships is gathering information, both about the nature of the relationships and the problems they face. Probably, the best way to conserve pollinators is to

preserve habitat that includes their food plants and nest sites. Given that the growing human population is resulting in rapid habitat destruction, the establishment of protected areas is an important conservation tool. In agricultural situations, management techniques as simple as changing the timing of pesticide application, leaving buffer zones around fields where bees can nest and food plants can grow, or providing suitable artificial nesting sites can make a significant difference. Domestication of wild bees by providing and managing nest sites may also play a role in their maintenance in agricultural situations.

If pollinators are locally extinct, it may be possible to reintroduce them once the factors that caused the original extinction are addressed. For example, a species of bumble bee that has disappeared from its native range in the UK may be reintroduced there from populations thriving in New Zealand, where the species was introduced. If the species is globally extinct, plants that were dependent on that pollinator could be maintained by hand pollination, or by the introduction of a suitable replacement. Although the bird was not introduced for that purpose, the presence of the Japanese white-eye in Hawaii has had the effect of replacing an extinct endemic avian pollinator. Another species of plant in Hawaii that lost its native pollinator is being maintained by hand pollination. Removal of introduced pollinators, such as the honey bee, may also help to preserve populations of native pollinators.

We have taken some important steps toward conservation of pollinators and thus the relationships they have with plants. But much basic knowledge about pollination and pollinators remains to be learned and it remains to be seen how many of these species can be conserved.

Appendix

List of Courses

1. Plant–Animal Interactions
2. Pollination Biology
3. Plant Ecology
4. Mutualism

See also: Extinction, Causes of. Forest Canopies, Plant Diversity. Mass Extinctions, Notable Examples of

References

Committee on the Status of Pollinators in North America, N.R.C. (2007) *Status of Pollinators in North America*. Washington, DC: The National Academies Press.

Dafni A (1992) *Pollination Ecology: A Practical Approach*. Oxford: IRL/OUP.

Faegri K and van der Pijl L (1979) *The Principles of Pollination Ecology*, 3rd edn. Oxford: Pergamon Press.

Grant V and Grant KA (1965) *Flower Pollination in the Phlox Family*. New York: Columbia University Press.

Jones CE and Little RJ (eds.) (1983) *Handbook of Experimental Pollination Biology*, Scientific and Academic Editions, New York.

Kearns CA and Inouye DW (1993) *Techniques for Pollination Biologists*. Niwot, CO: University Press of Colorado.

Kearns CA, Inouye DW, and Waser NM (1998) Endangered mutualisms: The conservation biology of plant–pollinator interactions. *Annual Review of Ecology and Systematics* 29: 83–112.

Matheson A, Buchmann SL, O'Toole C, Westrich P, and Williams IH (eds.) (1996) *The Conservation of Bees*. New York: Academic Press.

Nabhan GP and Buchmann SL (1997) Services provided by pollinators. In: Daily GC (ed.) *Nature's Services. Societal Dependence on Natural Ecosystems*, ch. 8, pp. 133–150. Washington, DC: Island Press.

Proctor M, Yeo P, and Lack A (1996) *The Natural History of Pollination*. Portland: Timber Press.

Real L (1983) *Pollination Biology*. Orlando: Academic Press.

Roubik DW (ed.) (1995) *Pollination of Cultivated Plants in the Tropics*, Agricultural Services Bulletin #118. Food and Agricultural Organization of the United Nations. Rome: Food and Agriculture Organization of the United Nations.

Willmer P (2011) *Pollination and Floral Ecology*. Princeton, NJ: Princeton University Press.

Pollution, Overview

William H Smith, Yale University, CT, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 731–743, © 2001, Elsevier Inc.

Glossary

Acid deposition Precipitation or dry fall-out with a pH of less than 5.0.
Anthropogenic Related to the activities of human beings.
Biodiversity Total inventory of genes, organisms, species, populations, communities, and their habitats.
Biogeochemical cycles Pathways and storage of chemical elements through the biota and geologic resources, including the atmosphere, hydrosphere, and lithosphere.
Biomagnifications The increase in body burden of persistent pollutants in organisms at higher trophic levels in food webs.
Climate change Variability of weather over time.
Ecotoxicology The study of chemical stressors on ecological resources.
Food web The pathways of energy and nutrient movement between organisms in an ecosystem.
Hydrologic cycle The movement of water between the surface of the earth and the atmosphere.
Nitrogen saturation Nitrogen levels in excess of the biological nitrogen needs of an ecosystem potentially causing adverse ecological effects.
Ozone Tri-atomic oxygen naturally present (and beneficial) in the stratosphere but present in the troposphere (due to human activities) in amounts potentially toxic to the biota.
Pollution Materials or chemicals in excess of natural levels caused by the activities of humans.

Introduction

Biodiversity includes the total inventory of genes, organisms, species, populations, communities, and their environments, and it integrates all of their complex interactions. This chapter focuses on the approximately 1.4 million species of living organisms known to presently exist on earth while recognizing that this total may represent fewer than 15% of the actual number (Raven and Wilson, 1992). Over geologic timescales, species numbers have been very dynamic. Massive species disappearances have occurred as follows: dinosaurs and marine reptiles in the Cretaceous (66 million years ago), reptiles and conodonts in the Triassic (210 million years ago), trilobites and marine animals in the Permian (250 million years ago), marine reef plants and animals in the Devonian (365 million years ago), and nautiloids and small ocean animals in the Ordovician (440 million years ago). Their mass extinctions are presumed to have resulted from earth impacts (asteroid collision?), climatic change, or large-scale alteration of biogeochemical cycles.

Changes in major element biogeochemical cycles over geologic time have been responsible for the appearance and disappearance of life forms. Changes in the oxygen cycle, for example, which allowed increased concentrations of oxygen gas in the troposphere and ozone in the stratosphere, allowed the evolution of life on earth to proceed. Major episodes of volcanic activity, resulting in marine eruptions and the release of sulfur containing and particulate pollutants, has periodically destroyed life.

Contemporary losses of species are recognized to result from physical stressors (e.g., habitat loss, harvesting), biological stressors (e.g., exotic organism introductions), and chemical stressors (changes in biogeochemical cycles, pollution). This chapter focuses on the latter group of stressors. Presently humans are recognized to have the ability to alter global and regional biogeochemical cycles via air, water, and soil pollution. As a result, the impact of pollution on biological diversity is a contemporary concern of major dimension (Barker and Tingey, 1992; Peters and Lovejoy, 1992).

Perturbation of the Carbon Biogeochemical Cycle

Mineral carbon is located in a limestone ($CaCO_3$) reservoir from which it may be leached into a mineral solution as dissolved hydrogen carbonate (HCO_3^-) formed when dissolved CO_2 reacts with $CaCO_3$. In the atmosphere carbon exists as carbon dioxide (CO_2). Carbon dioxide in the atmosphere is transferred to the biosphere via photosynthesis. Organic (fixed) carbon in the biosphere is ultimately returned to the atmosphere as CO_2 via microbial decomposition of organic matter. Another fraction of carbon is ultimately fixed as coal, petroleum, and natural gas. This fraction may return to the atmosphere as CO_2 via combustion.

In recent decades, human activities have accelerated the return of CO_2 to the atmosphere via increased deforestation, biomass burning (Nepstad *et al.*, 1999), and fossil fuel combustion (IPCC, 1996).

The accelerating input of CO_2 to the atmosphere from human deforestation and fossil fuel combustion (very approximately 8 billion metric tons annually) (**Figure 1**) is a significant driver of global climate change. In addition, the incomplete combustion of fossil fuels generates a variety of toxic and precursor compounds that contaminate the atmosphere and ultimately may cause adverse effects on plant and animal populations. Both of these stressors (i.e., climate change and air pollutants) have significant potential to reduce biological diversity.

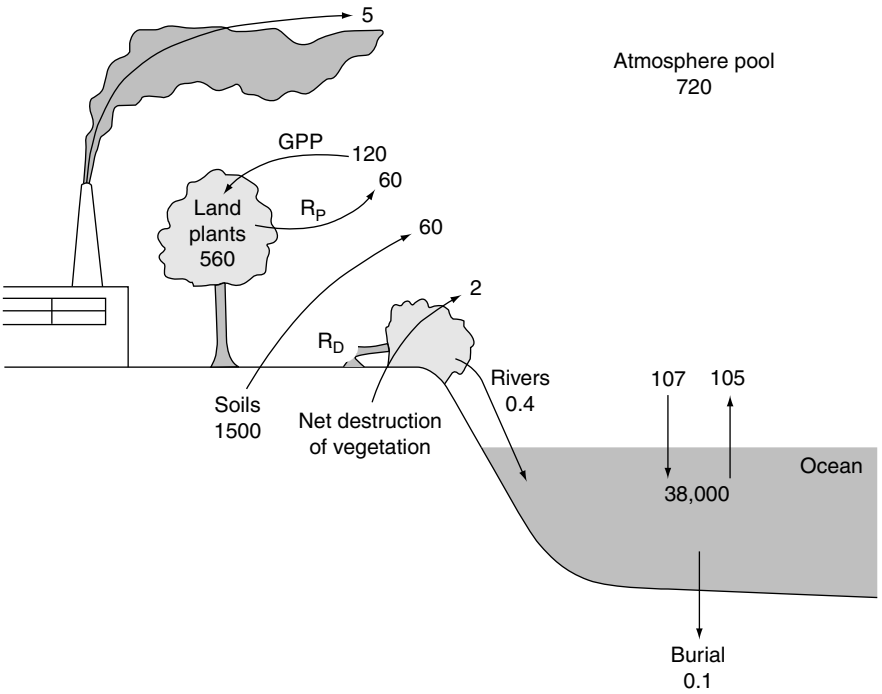


Figure 1 The global carbon cycle. The storage (numbers without arrows) and the annual flows (numbers with arrows) are shown for the global system. The units are pentagrams of carbon ($= 10^{15}$ g C = one gigation = 1 billion metric tons).

Climate Change

Climate is defined as the time averaged value of meteorological quantities. Over time, climate, like an ecosystem, is characterized by change not constancy. In geological terms, the climate of the earth is most typically characterized by extended "moderate periods" with equable weather the year round, lack of ice caps, and generally warm seas. Humans evolved after the last "moderate period" and our development has been in a period of climatic revolution. This period of revolutionary climatic alteration has been characterized by a complex of "cycles within cycles." Over the past 3000 years, for example, the general evidence suggests that the northeastern portion of North America has become cooler and more moist. Over the past several hundred years, however, particularly during the first half of the present century, there is evidence for a moderating trend. There is general agreement that there has been a systematic fluctuation in recent global climate characterized by a net worldwide warming of approximately 0.5 °C between the 1880s and the early 1940s. Warming has generally continued over the past three decades.

The regulation of global climate is complex and incompletely appreciated. Numerous hypotheses have been proposed to explain the forces responsible for the variability of climate. The most plausible of these include variations of the solar constant, changes in solar activity, passage of the solar system through an interstellar gas-dust cloud, variation in the velocity of the earth's rotation, gigantic surges of the Antarctic ice sheet, changes in the earth's orbital parameters, and alterations in the interactions between glaciers and oceans. Climate may respond rapidly and dramatically to small changes in these independent variables.

Added to this complexity and uncertainty is the suggestion that the activities of human beings, particularly land use activities and atmospheric contamination, are currently influencing global and regional climates. Numerous trace gases of the atmosphere, including water vapor, carbon dioxide, methane, nitrous oxide, halocarbons, and ozone (greenhouse gases), have strong infrared absorption bands. As a result, these gases can have a significant effect on the thermal structure of the atmosphere because they absorb within the 7- to 14-μm atmospheric window, which transmits most of the thermal radiation from the surface of the earth and troposphere to space. Presumably a primary result of more carbon dioxide and other "greenhouse" gases in the atmosphere will be warming. While incoming solar radiation is not absorbed by carbon dioxide and these trace gases, portions of infrared radiation from the earth to space are. Over time, the earth will become warmer. While the forces controlling global temperature are varied and complex, as suggested, the increase of 0.5 °C since the mid-1800s is generally agreed to be at least partially caused by increased carbon dioxide.

General circulation models, widely used to predict future climates, are numerical models of the earth atmosphere system that solve the basic equations for atmospheric motion and provide boundary conditions of the earth and ocean. General circulation models divide the landscape into a series of regions called *grid boxes*. Each grid box consists of a series of layers, which represent land, ocean, and layers of the atmosphere. Global weather is simulated using difference equations, which detail the physics and dynamics of energy and material movement among the boxes. To evaluate the climatic change hypotheses, models are initially run with current atmospheric conditions. They are then run with altered atmospheric trace

gas concentrations (for example, carbon dioxide) to simulate a new climate. The difference between these two equilibrium climate simulations provides an estimate of how the climate system may change. Most general circulation models predict warming and overall intensification of evaporation and precipitation in the warmer climate regimes of doubled atmospheric carbon dioxide.

Using a mid-range scenario of driving variables, an increase in global mean surface air temperature relative to 1990 of about 2 °C by 2100 has been estimated. Under any scenario, the average rate of warming will probably be greater than any seen in the past 10,000 years (IPCC, 1996). General circulation model predictions of the change in the hydrologic cycle with doubling of atmospheric carbon dioxide are not clear. While precipitation is generally estimated to increase, warming will intensify evaporation. As a result, soil moisture may not increase with warming; it may even decrease. The movement of water into and out of the soil-vegetation system has a large influence on the hydrologic cycle and is very variable regionally. Models have predicted that in the United States, portions of the Northwest could become wetter, portions of the Southern Great Plains could become somewhat drier, and portions of the Northeast could become considerably drier.

Ecosystems and Climate Change

The consequences of a warmer global climate, with even very modest changes in regional temperature and hydrologic patterns, can have profound effects on ecosystems. Natural ecosystems, as opposed to artificial ecosystems designed and intensively managed by humans, may be the most vulnerable to climate change. We are least able to intervene and facilitate adaptations of natural ecosystems because of their expansiveness, our limited knowledge about fundamental ecological processes, and the absence of effective global institutions to assume responsibility for adaptations.

A shift in climate zones may exceed the ability of vegetation, for example grassland, savanna, or forest systems, to adapt through migration. If climate zones shift hundreds of miles over the next century, rates of vegetative dispersal and colonization, which are on the order of tens of miles in a century, may be inadequate to maintain present-day ecoregion patterns. With regard to terrestrial ecosystems, climate change may also adversely influence habitats by changing patterns of major stressor groups including fire regime, drought regime, and insect and disease relationships.

Alteration of ecosystems could lead to a change in species diversity, including a loss of species. The potential for migration of plants and animals to new suitable habitats is not known, but barriers such as water bodies, roads, or development could impede migration. Isolated (“island”) species may find themselves in climate zones that are no longer suitable for their survival.

Montane species, unless they can migrate to higher elevations, may face extirpation or extinction. A study of a 3 °C warming in the Great Basin National Park in eastern Nevada concluded that it would cause 20 to 50 percent of species in individual mountain ranges to go extinct (Murphy and Weiss, 1992).

Forests

Tropical montane cloud forests are unique systems due to their dependence on regular cycles of cloud formation for their existence. Still *et al.* (1999) have simulated environmental conditions for such forests at CO_2 doubling and found that a combination of reduced cloud contact and increased evapo-transpiration could threaten the persistence of these systems. Pounds *et al.*, 1999 have, in fact, provided evidence that alterations in bird, reptile, and amphibian communities in the highland forests at Monteverde, Costa Rico, have occurred and may be related to recent increases in air temperature.

The vulnerability of forest ecosystems to climate change is a function of the forest location, biology, and human management activity. Forests at greatest risk will generally be those located in the northern portions of temperate latitudes. Altitude, proximity to continental margins, and distance from large water bodies will be important influencing factors. Plant species that are able to use increased CO_2 efficiently will have an advantage over other species. Similarly, species that distribute seeds widely may fare better under climate change. A mixed-species forest may tolerate a wider variety of changes than would a single-species forest. Individual plants with a low tolerance for climate fluctuations would be least adaptable to ongoing climate change. Individual populations with little genetic diversity among plants might prove to be at greatest risk of long-term decline. The most intensively managed industry and private forest land may be least at risk of catastrophic loss or long-term decline because mitigation efforts to reduce such effects will be undertaken. Many private forest managers have both the financial incentive (and resources) and the latitude to protect against extensive loss from climate-related threats. They can use several available techniques: short rotations to reduce the length of time a tree is exposed to an unsuitable climate; planting better-adapted varieties developed through selection, breeding, or genetic engineering programs to reduce vulnerability; and thinning, weeding, pest management, irrigating, improving drainage, or fertilizing to improve general health (**Table 1**).

The first observable effects of climate change on ecosystems will not be so much climate related as weather related. The near-term effects of climate change will be driven by changes in weather extremes and mediated through those stressors that have always been the primary controllers of ecosystem structure and health: insects, disease, wind, and fire. Even in regions where productivity may be ultimately improved, the transition period could be extended and punctuated by sudden dieback and decline. Forest ecosystems are complex, long-lived systems that can only slowly adjust to climate but that can suddenly be threatened by weather-related stresses.

The near-term response of forest systems to climate change will involve complex reactions to new averages, new patterns, and new extremes in weather variables. Some forest species that are specialized to current climate conditions may not thrive. Altered patterns of exposure to high and low temperatures could mean that winter chilling requirements will not be met. Flowering, seed-formation, and seed-dispersal processes could be disrupted, especially if pollinators do not

Table 1 Forest ecosystem vulnerability to climate change caused by alteration of the global carbon cycle

Forest location	*Forest biology*	*Forest management*
Higher latitude	Small, fragmented range	Fragmented forests
Higher elevation	No or few migration corridors	Less-diverse forests
Continental interior	Low genetic variance	High stand density
Maritime sites	Low species diversity	Inappropriate species
Forest-range boundaries	Genetically specialized to site	
Low-productivity sites	History of widespread dieback	
High-fire-risk sites	Heavy seed	
Drier sites		

Table 2 Forest vulnerabilities of selected United States forest systems to changes resulting from climate alteration

Region	*Tree*	*Potential stressors/key climate factor and vulnerability*
Northeast/North Central	Maple	Insect defoliators/warm, dry weather *Armillaria*(root decay)/drought Deep root freezing/lack of winter snow cover **Vulnerability:** High potential for damage with warmer temperatures; drier conditions
Southeast	Loblolly pine	Southern pine beetle/prolonged hot, dry weather Fungus (fusiform rust)/warm, moist weather Fire (favors longleaf or shrub)/warm, dry weather Storm damage/increase in coastal storms **Vulnerability:** Potential for much warmer weather (with increase or decrease in precipitation) to reduce productivity.
Rocky Mountain/Pacific Southwest	Ponderosa pine	Borers, bark beetles/drought Fire/drought or lightning **Vulnerability:** Resistant to near-term climate change, though productivity may decrease
Pacific Northwest (Coastal)	Douglas fir	Most stressors not strongly weather related **Vulnerability:** Resistant to near-term climate change, though productivity may decrease
Alaska	Spruce	Spruce beetle/warm weather (speeds insect development), moisture problems, erratic freezes **Vulnerability:** High potential for rapid effects because climate plays pervasive role.

adjust to changing conditions. With longer growing seasons, trees might add more "light" earlywood relative to the "dense" latewood that forms at the end of the growing season. This would mean a lower-quality wood for structural lumber and higher costs for pulp mills. Changes in early growing season weather conditions, particularly moisture and frosts, may affect the establishment of seedlings. Warmer and moister weather might favor the spread and boost the significance of certain fungal diseases and insect pests. Elsewhere, the drying effects of higher temperatures could be especially damaging, especially where the frequency of drought is increased. Associated with droughts would be higher risks of secondary threats from forest fires and insects. Insect damage may increase, for example, if insect pests produce more generations or persist longer during the tree-growing season.

The initial effects of climate change will not at first be easily recognized as distinct from the effects of the normal regulators of forest growth and development. The potential initial effects of climate change can be illustrated by current weather-related stresses on selected highly valued tree species in several regions of the United States. These potential vulnerabilities, by region, are presented in **Table 2**.

Coastal-Marine

Coastal zones are very productive ecosystems, very rich in species, and very perturbed by human activities. In addition, they may be at substantial risk from global climate change. These risks stem from changes in water column temperature or other water quality parameters, change in ocean current dynamics, and alterations of sea level.

Widespread extinctions are not likely, but expansive alteration of community distributions and composition are possible. Most marine organisms, in contrast to long lag times discussed for forest ecosystems, have high mobility, large ranges, high fecundity, and rapid growth rates. As a result, they may be able to adapt to climate change more efficiently than terrestrial systems (Ray *et al.*, 1992).

Coral reefs are the most species rich of marine ecosystems. These reefs are largely restricted to tropical waters, but generally do not occur where temperatures exceed 30 °C for extended periods. As a result, even small increases of 1 to 2 °C for the surface waters of tropical oceans could have important implications for reef ecosystems.

Temperate latitude estuarine and tidal wetland and marsh ecosystems will be at substantial risk if sea level increases

over the next several decades. For each cm of rise in sea level, beaches may erode 1 m landward. Storm surges, a major eroding force, will increase, especially in areas of modest shore slopes. For every 10 cm rise, saltwater wedges in estuaries and tidal rivers may advance 1 km. Sea-level rise will also increase salinity intrusion into coastal freshwater aquifers (NAS, 1987).

Fresh Water

Freshwater ecosystems will present challenges for native biota if global climate continues to warm. Changes in precipitation and evapotranspiration can combine to offset surface-water and groundwater quantity and quality, low and high flow conditions, and drought and flood frequency. In addition, water column temperature is a key determinant of habitat quality. Stream temperature, for example, is a fundamental habitat criterion for all salmonids. For example, the habitat of brook trout in the southern Appalachian Mountains is already fragmented by natural and anthropogenic forces. Climate warming would directly and indirectly fragment and limit brook trout range even more via temperature change and hydrologic cycle change. Flebbe (1997) used geographic information system techniques to evaluate warming effects on brook trout distributions. Both suitable area and stream length for these fish were seen to decrease as suitable habitat is increasingly restricted to mountaintops.

Perturbation of the Nitrogen Biogeochemical Cycle

The atmosphere is 78 percent by volume elemental nitrogen (N_2). The N_2 molecule is extremely stable and only very specialized pathways allow this nitrogen to be made available to the biota. Elemental nitrogen is incorporated into chemically bound forms (fixed) by biochemical processes mediated by microbes. High-energy lightning discharges can produce nitrogen oxides. The nitrogen incorporated in the biota is mineralized to the inorganic form during the decomposition of biomass. The production of gaseous N_2 and N_2O by microorganisms and the evolution of these gases to the atmosphere completes the nitrogen cycle through a process called denitrification (**Figure 2**). While elemental nitrogen and elemental oxygen (O_2) do not react at ambient temperature and pressure conditions in the troposphere, they do react at the elevated temperatures of combustion processes. As a result, all anthropogenic activities involving combustion, from industrial and power generation combustion to automobile and truck engine combustion, function to inject large amounts of nitrogen oxides (NO and NO_2) to the lower atmosphere. These oxides are extremely important pollutants and have substantial potential significance for biological diversity. Nitrogen oxides present risks to the biota directly in the form of acid rain and nitrogen saturation and indirectly as they act as precursors for the formation, along with hydrocarbons, of ozone in the troposphere.

Acid Rain

Acid rain is appropriately described as an old environmental problem with a new image. Acid rain, more than any other environmental contaminant, has focused societal concern on ecosystem toxicology. Natural rain, including precipitation in

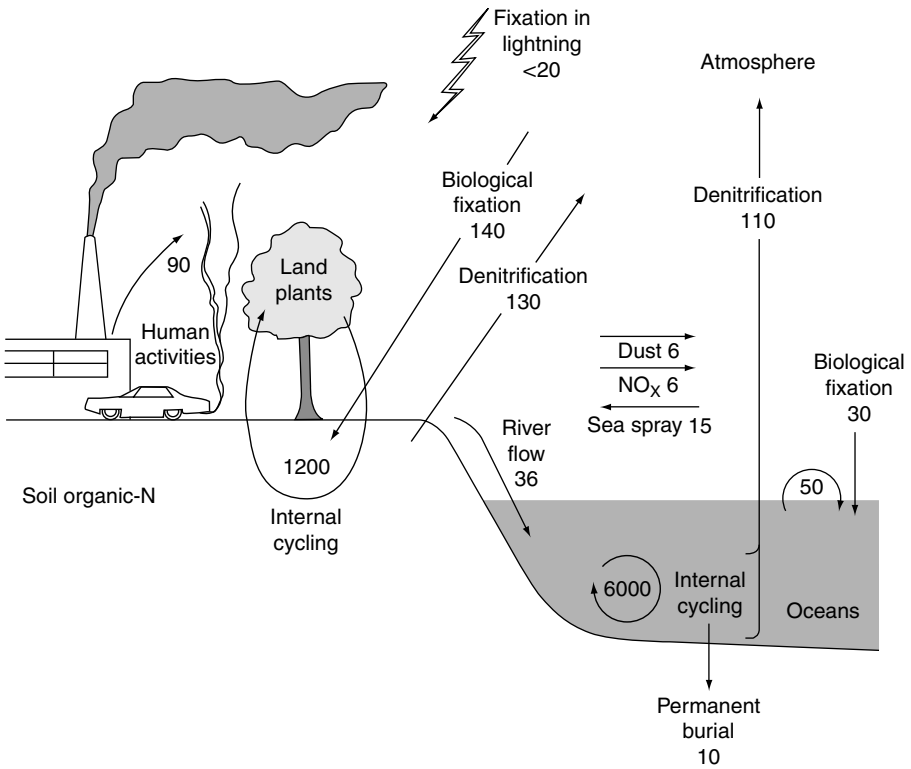


Figure 2 The global nitrogen cycle. The storage (numbers without arrows) and the annual flows (numbers with arrows) are shown for the global system. The units are 10^{12} g nitrogen.

relatively clean or unpolluted regions, is naturally acidic, with a pH in the range of 5.0 to 6.0. This natural acidity results from the oxidation of carbon oxides and the subsequent formation of carbonic acid. Formic and acetic acids, originating primarily from natural sources, may also contribute minor amounts of acidity to precipitation.

In regions downwind from electric-generating power stations employing fossil fuels, industrial regions, or major urban centers, precipitation can be acidified below pH 5.0. Precipitation with a pH less than 5.0 is designated acid rain. This human-caused acidification of precipitation results primarily from the release of sulfur dioxide (SO_2) and nitrogen oxides (NO and NO_2) from smokestacks and tailpipes. The sulfur and nitrogen oxides are subsequently oxidized to sulfate (SO_4^{2-}) and nitrate (NO_3^-), hydrolized, and returned to earth as sulfuric (H_2SO_4) and nitric (HNO_3) acids. This additional acidification of precipitation by human activities can readily reduce the pH of precipitation in downwind regions to between 4.0 and 5.0 on an annual average basis. Individual precipitation events that have a pH in the 3.0 to 4.0 range are not uncommon.

The atmosphere deposits acidity onto the landscape both during and in between precipitation events. In the latter case, termed dry deposition in contrast to wet deposition, the acids are delivered in the gas phase or in association with fine particles (aerosols). Acid deposition is a term that includes acid delivery in the form of precipitation (rain, snow, fog, and cloud moisture) plus dry deposition. In view of the importance of both wet and dry deposition in acid transfer from the atmosphere to the biosphere, acid deposition is a much more appropriate descriptor than acid rain.

Aquatic Ecosystems

The pH range for undisturbed lakes and streams with full compliments of native biota, especially selected finfish species important to humans, typically is within the range of 6 to 8. Lakes and streams with a pH < 6 do occur naturally. Land use practices, for example, conversion from agriculture to forest in the watershed of a lake can, over time, also cause lake acidity to increase. Aquatic resources acidified from the input of acid rain are also well documented, particularly in the northern United States, eastern Canada, and Scandinavia.

Fortunately, most surface waters in the United States have sufficient buffering (acid-neutralizing) capacity to be resistant to acidification via atmospheric deposition. In the instance of high acid deposition inputs and poorly buffered aquatic systems, however, declines in water column pH can be caused by acid rain. In the mid-1980s a National Surface Water Survey was conducted in areas of the United States known to contain lakes and streams with little capacity to neutralize acids. This survey identified the Adirondaks and mid-Atlantic Highlands region as having both sensitive aquatic resources and high levels of acid deposition.

The adverse response of aquatic biota to acidification is very well documented (NAPAP, 1991). Sensitive species may be stressed or lost at small increases in acidity.

Acid-sensitive mayfly and stonefly species can be impacted at pH levels near 6.0 and sensitive finfish species, for example, fathead minnows, can be lost at pH 5.6 to 5.9. Acid sensitive phytoplankton, zooplankton, and benthic macroinvertibrate species may also be lost. Acid-tolerant species may increase significantly with acidification. Survival of fish in acidic waters is primarily related to pH level, inorganic chemistry of aluminum, and the concentration of calcium. Other relevant considerations include food web effects, spatial and temporal variation in exposure to acidity and aluminum at various life stages, and behavioral avoidance. Critical pH levels for selected aquatic organisms are presented in **Figure 3**.

Forest Ecosystems

The adverse impact of acid deposition on forest systems is focused on the montane forests of eastern North America. Prevailing wind patterns and the importance of cloud moisture delivery of acidic materials places these high-elevation forests at special risk. Both morbidity and mortality of large numbers of forest trees are easily documented in montane forests. Acid deposition may be directly or indirectly involved in this symptomology. The primary mechanisms of proposed effects are foliar weathering, foliar and soil nutrient leaching, aluminum and heavy metal soil toxicity, and predisposition to greater microbial disease and insect activity (Smith 1989, 1990).

Nitrogen Saturation

For forests in temperate latitudes, the input of nitrogen via atmospheric deposition may approximate 0.5 kg *N* ha/yr in clean atmospheric conditions. This low input combined with acidic organic soils in cool climates with high lignin content and slow rates of decomposition and, therefore, slow internal nitrogen turnover cause most forest systems to be nitrogen deficient. Numerous forests provided with supplemental nitrogen exhibit enhanced growth rates indicating that they were nitrogen limited prior to fertilization. In forest locations with less than clean atmospheric conditions (acid rain), forests may receive up to 27 kg *N* ha/yr if both dry and wet deposition sources are inventoried. At elevated levels of atmospheric input and normal dynamics of internal nitrogen cycling, forests can be fertilized by nitrogen in acid rain and exhibit increased growth and productivity.

If high levels of nitrogen deposition persist, however, nitrogen availability may reach a level that exceeds the vegetative and microbial nitrogen demand and cause the system to enter a condition designated nitrogen saturated. If sustained, evidence has been presented to indicate that nitrogen saturation can be detrimental to forest systems and lead to decreases in foliar biomass, imbalances in foliar nutrient concentrations, enhanced nitrification, and possibly tree morbidity and mortality. Nitrogen saturation of forest soils may lead to high levels of nitrate in stream waters of forest catchments. Excess nitrogen loading of surface waters can impose stresses on aquatic species.

Nitrogen Oxides as Tropospheric Ozone Precursor

In the presence of sunlight nitrogen dioxide is dissociated and forms equal numbers of nitric oxide molecules and oxygen

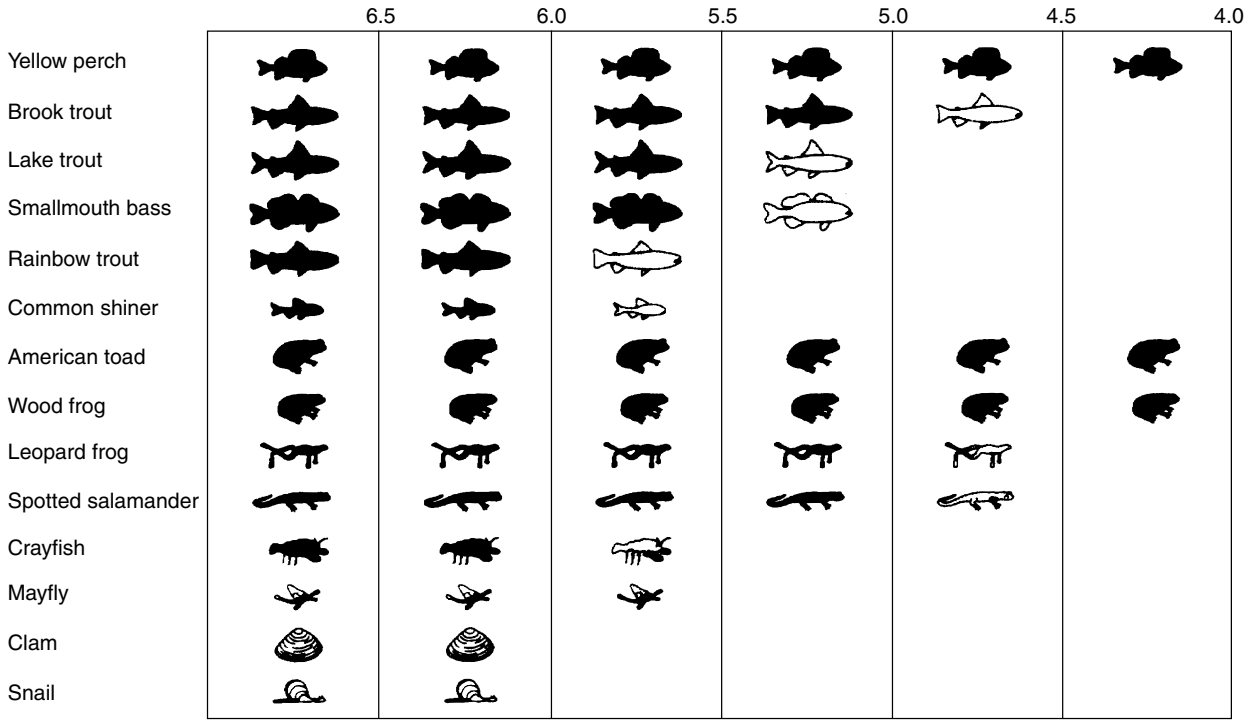


Figure 3 Critical pH levels for selected aquatic organisms where the American toad, wood frog, leopard frog, and spotted salamander are shown in the embryonic stage, and the crayfish, mayfly, clam and snail are selected species.

atoms. The oxygen atoms rapidly combine with molecular oxygen to form ozone. This ozone then reacts with the nitric oxide, on a one-to-one basis, to reform nitrogen dioxide. The steady-state concentration of ozone that is produced by this cycle is very small. When hydrocarbons, aldehydes, or other reactive atmospheric constituents are present, however, they can form peroxy radicals that oxidize the nitric oxide back to nitrogen dioxide. With reduced nitric oxide available to react with the ozone, the latter may accumulate to relatively high concentrations.

High concentrations of ozone are toxic to vegetation. Morbidity and mortality of agricultural and forest plants caused by ozone are well documented. California provides numerous examples of ozone stress on wildland ecosystems. In southern California the predominant native shrub land vegetation consists of chaparral and coastal sage scrub. The former occupies upper elevations of the coastal mountains, extending into the North Coast ranges, east to central Arizona, and south to Baja California; the latter occupies lower elevations on the coastal and interior sides of the coast ranges from San Francisco to Baja California. Westman (1979) applied standard plant ordination techniques to these shrub communities to examine the influence of air pollution. The reduced cover of native species of coastal sage scrub documented on some sites was statistically indicated to be caused by elevated atmospheric ozone. Sites of high ambient ozone were also characterized by declining species richness.

Ponderosa pine is one of five major species of the "mixed conifer type" that covers wide areas of the western Sierra Nevada and the mountain ranges, including the San Bernardino Mountains, in southern California from 1000 to 2000 m elevation. Other species represented include sugar pine, white fir, incense cedar, and California black oak. The response of these five major tree species to ozone in the San Bernardino National Forest is variable as shown by field surveys and seedling exposures. Ponderosa pine exhibits the most severe foliar response to elevated ambient ozone. An aerial survey conducted by the USDA Forest Service indicated 1.3 million ponderosa (or Jeffrey) pines on more than 405 km^2 (100,000 acres) were stressed to some degree. Mortality of ponderosa pine has been extensive. White fir has suffered slight damage, but scattered trees have exhibited severe symptoms. Sugar pine, incense cedar, and black oak have exhibited only slight foliar damage from oxidant exposure. A 233 ha study block was delineated in the northwest section of the San Bernardino National Forest in order to conduct an intensive inventory of vegetation present in various size classes and to evaluate the healthfulness of the forest. Ponderosa pines in the 30 cm diameter class or larger were more numerous than other species of comparable size in the study area. These pines were most abundant on the more exposed ridge crest sites of the sample area. Mortality of ponderosa pine ranged from 8 to 10 percent. The loss of a dominant species in a forest ecosystem clearly exerts profound change in that system. Miller (1989) concluded from his investigations that the lower two-thirds of the study area will probably shift to a greater proportion of white fir. It was judged that incense cedar will probably remain secondary to white fir. Sugar pine was presumed to be restricted by lesser competitive ability and dwarf mistletoe infection. The rate of composition change was deemed dependent on the rate of ponderosa pine mortality. The upper one-third of the study area, characterized as more

environmentally severe due to climatic and edaphic stress, supports less vigorous white fir growth. Following the loss of ponderosa pine in this area, sugar and incense cedar may assume greater importance. Miller judged, however, that the natural regeneration of the latter species may be restricted in the more barren, dry sites characteristic of the upper ridge area. California black oak and shrub species may become more abundant in these disturbed areas. Generally, the data support the hypothesis that forest succession toward more tolerant species such as white fir and incense cedar occurs in the absence of fire. In the presence of fire, pine may be favored by seedbed preparation and elimination of competing species.

The changes in forest composition caused by ozone in this southern California forest have created a management concern, as well as ecological change, because the forest is intensively used as a recreational resource and the loss of ponderosa pine is judged to reduce the aesthetic qualities of the forest.

Perturbation of the Sulfur Biogeochemical Cycle

The natural sulfur cycle is complex in that it involves several gaseous species, poorly soluble minerals, and several species in solution (**Figure 4**). Sulfur enters the atmosphere from natural sources as hydrogen sulfide (H_2S) from active volcanoes and the decay of organic matter in anaerobic environments (swamps, tidal flats), sulfur dioxide (SO_2) from active volcanoes, and particles of sulfate salts (e.g. ammonium sulfate) from sea spray. Approximately one-third of all sulfur compounds and 99 percent of the SO_2 reaching the troposphere comes from human combustion activities. The combustion of coal and oil for the production of electricity and the smelting of metal-bearing ores have historically been major sources of SO_2 to the atmosphere.

The industrial and energy sector release of SO_2 over multiple decade timescales has reduced biological diversity in local environments surrounding point source facilities in many parts of the world.

In North America, for example, we can cite the cases of Ducktown, Tennessee, and Sudbury and Wawa, Ontario. The Ducktown, Tennessee, area consists of approximately 243 km^2 (60,000 acres) and was originally covered with southern deciduous forest. Mining operations were initiated in the basin in 1850 and smelting operations were most active between 1890 and 1895. By 1910 gross forest simplification resulting from excessive sulfur dioxide had created three new vegetative zones surrounding Ducktown. In a 27 km^2 (10.5 $mile^2$) area closest to the source, vegetation was devastated and largely eliminated. All trees and shrubs were destroyed, and only a few, isolated islands of sedge grass occurred in the outer portions of this zone. A belt of grassland ecosystem, 68 km^2 (17,000 acres) in size, surrounded the barren zone. The principal grassland species was broomsedge. A transition zone of somewhat indefinite boundary and consisting of approximately 120 km^2 (30,000 acres) was located beyond the grassland. Few trees were located along the inner edge of the transition zone. Sassafras, red maple, sourwood, and post oak were common in the middle of the transition forest. The uninfluenced forest beyond the impact of the smelter consisted principally of mixed oaks, hickory, dogwood, sourwood, black tupelo, and some eastern white pine. The distance of

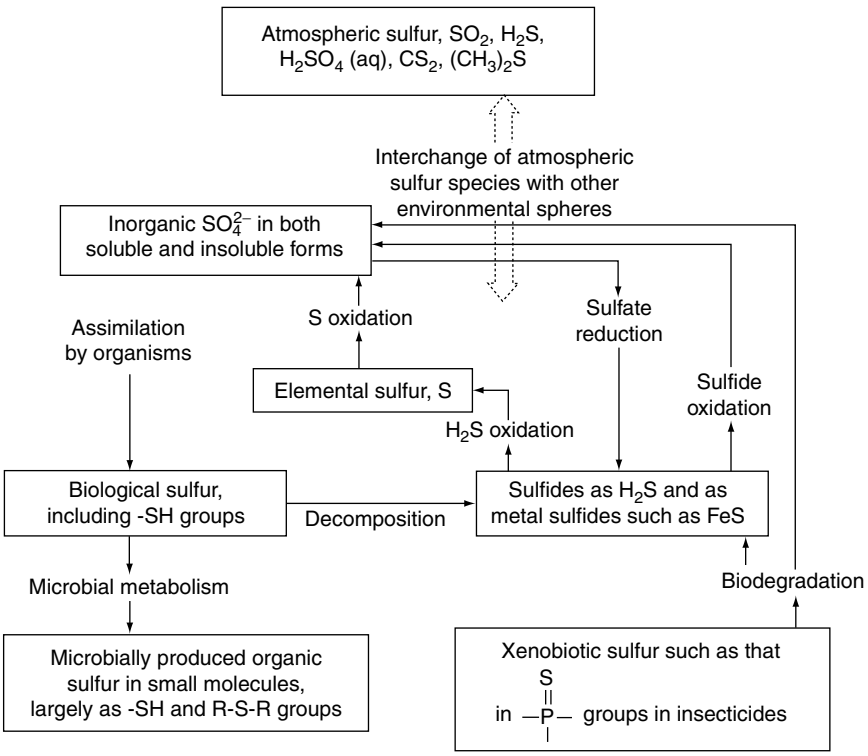


Figure 4 The global sulfur cycle.

vegetative impact extended 19 to 24 km (12 to 15 miles) to the north and approximately 16 km (10 miles) to the west of the smelter. Eastern white pine damage was recorded 32 km (20 miles) from the industry.

Sheet and gully soil erosion has been excessive in the acutely damaged inner zone. Micrometeorological changes in the inner zone relative to the surrounding forest have been substantial; summer air temperature averages are 1 to 2 °C higher, while the winter air temperature averages 0.3 to 1 °C lower, the soil temperature is 11 °C higher in the summer; the wind velocities are five to 15 times higher and rainfall is consistently lower.

Reforestation of Copper Basin has not been rapid due to severe erosion, low soil nutrients and moisture, and high winds. Recent reforestation efforts have employed both pine and hardwood species (Smith, 1990).

In the Sudbury area, a century of sulfur dioxide fumigation, copper and nickel particulate deposition, fire, soil erosion, and increased frost action have interacted to create 10,000 ha (25,000 acres) of barren land and 36,000 ha (89,000 acres) of stunted, open birchmaple woodland. Three large nickel and copper smelters have historically discharged several thousand tons of sulfur dioxide daily into the surrounding atmosphere. Sulfur dioxide emissions from this area have historically approximated 10% of the North American sulfur dioxide total and 25% of the smelter total. Extensive simplification of the mixed boreal forest ecosystem surrounding this region has occurred primarily via the mortality of eastern white pine throughout a 1865 km^2 (720 square mile) area to the northeast of the Falconbridge, Copper Cliff, and Coniston smelters. Acute impact on balsam has been recorded in excess of 40 km (25 miles) from the source of sulfur dioxide. Red oak, red maple, and red-berried elder are more tolerant and may exist in disturbed forests as close as 1.6 km (1 mile) to the smelters. Morbidity and mortality of forest trees in the Sudbury region continued to spread at least until the construction of the world's tallest smokestack of 403 m (1250 feet) at Copper Cliff in 1972 and the closing of the Consiton smelter also in 1972.

Soil erosion has followed the destruction of surrounding forest ecosystems. Rainfall has been made highly acidic, commonly less than pH 3.0 in the early 1970s. Elevated nickel and copper concentrations in soils have been recorded to distances of 50 km (31 miles). Eroded sediment, contaminated with trace metals, has contaminated area lakes and water resources. This contamination, along with acidification of surface waters, has resulted in numerous aquatic species extirpations.

Forest destruction in the vicinity of an iron smelter in Wawa, northern Ontario, exhibits a more discrete pattern, one smelter relative to three, and a shorter history of impact resulting in less equilibration of vegetative response than in Sudbury. The Wawa smelter initiated operations in 1939 (significantly expanded in 1949) and has released as much as 100,000 tons of sulfur dioxide annually to the surrounding atmosphere. Negative impact is primarily confined to a strip northeast from the plant in the direction of the prevailing wind. Symptoms of sulfur dioxide damage may be observed for at least 32 km (20 miles) to the northeast. The mixed boreal forest in the Wawa area consists mainly of white spruce, black spruce, balsam fir, jack pine, white cedar, larch, and white pine in the dominant layer. Mountain maple and *Pyrus decora* are common in the understory.

Trace Pollutants of Global Importance

Persistent organochlorine compounds, including the polychlorinated biphenyls (PCBs) and organochlorine pesticides, are among the most important anthropogenic contaminants with global status. These pollutants biomagnify in food webs and have resulted in acute and subtle ecotoxicological effects in aquatic and terrestrial organisms. Despite increased regulation and reduced use, urban and industrial areas remain a significant source of PCB pollutants. The insecticide DDT has seen reduced or terminated use in Canada, the United States, and western Europe, but increased use in Mexico, Central America, India, and eastern Europe. Both PCBs and DDT (including its breakdown products DDD and DDE) can enter the atmosphere via volatilization, or in combination with aerosols, and be transferred from the atmosphere to the biosphere via both dry and wet deposition. Long-distance atmospheric transport and subsequent deposition have been observed to contaminate numerous ecosystems remote from primary sources including estuarine and lacustrine systems. Less research has been directed to documenting contamination of terrestrial systems, especially forests.

Northern and boreal forests develop a thick organic forest floor and in recognition of the tendency of trace organic pollutants to bind to the surface (adsorb) or penetrate (absorb) into soil particles, it is especially important to examine these organic soils for trace pollutants. Such an examination of forest floors of remote northern New England hardwood and montane forest soils has revealed the presence of numerous PCBs and DDT and its breakdown products (Smith *et al.*, 1993).

While the soil and water column concentrations of these pollutants are relatively low, their low water solubility but high lipid solubility allow the processes of bioaccumulation and bioconcentration to elevate these potential toxics to levels sufficient for adverse biological effect (Allen-Gil *et al.*, 1997a, b). These elevated concentrations can cause change in community composition (Ferraro and Cole, 1997).

Conclusion

Human beings have–through their efforts to generate electricity, manufacture things, and transport themselves and their goods–managed to alter major biogeochemical cycles of the earth. In addition, the manufacture and use of a variety of environmentally persistent chemicals have resulted in the global circulation of materials with potentially biologically toxic properties.

Perturbation of the carbon cycle via increased combustion and deforestation and the resulting global warming presents a high risk to biological diversity. Continued global warming will result in species extirpations and possible extinctions as we progress to the middle of the twenty-first century. Changes in the nitrogen cycle caused by human activities are less impressive than the alteration of the carbon cycle but still have

the extended-term potential to change both terrestrial and aquatic ecosystems and to result in species losses. Changes in the sulfur cycle, while substantially less widespread than alterations in the carbon and nitrogen cycles, have proven historically to have been sufficient to cause numerous local extirpations in the vicinity of point sources.

The atmosphere, while highly structured (layered) horizontally, has no vertical partitions. As a result, the injection of chemically persistent toxins anywhere in the world means they may begin a journey of global circulation and eventual deposition. Accumulation of these toxins via environmental media or food-web exposure can place certain species at risk of loss from receiving environments.

Pollutants of global importance and the ability of humans to alter fundamental biogeochemical cycles must be recognized as important mechanisms for the loss of biological diversity. Efforts to monitor the trends of both pollutants and major chemical cycles along with species at risk must be made a high priority for contemporary environmental management. The loss of species and communities must be recognized as a primary reason for increased international efforts to reduce the production and release of pollutants.

See also: Carbon Cycle. Climate Change and Ecology, Synergism of. Greenhouse Effect. Nitrogen, Nitrogen Cycle

References

Allen-Gil SM, Gubala GP, Wilson R, *et al.* (1997a) Organochlorine pesticides and polychlorinated biphenyls (DCBS) in sediments and biota from four U.S. Arctic lakes. *Archives of Environmental Contamination and Toxicology* 33: 378–387.

Allen-Gil SM, Landers DH, Wade TL, *et al.* (1997b) Heavy metal, organochlorine pesticide and polychlorinated biphenyl contamination in Arctic ground squirrels (*Spermophilus parryi*) in northern Alaska. *Arctic* 50: 323–333.

Barker JR and Tingey DT (eds.) (1992) *Air Pollution Effects on Biodiversity.* New York: Van Nostrand Reinhold.

Ferraro SP and Cole FA (1997) Effects of DDT sediment-contamination on macrofaunal community structure and composition in San Francisco Bay. *Marine Biology* 130: 323–334.

Flebbe PA (1997) Global climate change and fragmentation of native brook trout distribution in the southern Appalachian Mountains. In: Gresswell RE, Dwyer P, and Hamre RH (eds.) *Wild Trout VI: Putting the Native Back in Wild Trout* Proceedings of the 6th Wild Trout Conference. August 17–20, 1997. Bozeman, MT.

Intergovernmental Panel on Climate Change (IPCC) (1996) *Climate Change 1995. Technical Summary of the Working Group I Report.* Cambridge: Cambridge University Press.

Manahan SE (1994) *Environmental Chemistry*, 6th ed Boca Raton, FL: Lewis Publishers.

Miller PR (1989) Concept of forest decline in relation to western U.S. forests. In: MacKenzie JJ and El-Asbury MT (eds.) *Air Pollution's Toll on Forests and Crops.* New Haven, CT: Yale University Press.

Murphy DD and Weiss SB (1992) Effects of climate change on biological diversity in western North America: Species losses and mechanisms. In: Peters RL and Lovejoy TE (eds.) *Global Warming and Biological Diversity*, pp. 355–368. New Haven, CT: Yale University Press.

National Academy of Sciences (NAS) (1987) *Responding to Changes in Sea Levels: Engineering Implications.* Marine Board, National Research Council. Washington, DC: National Academy Press.

National Acid Precipitation Assessment Program (NAPAP) (1991) *1990 Integrated Assessment Report.* Washington, DC.

Nepstad DC, Verissimo A, Alencar A, *et al.* (1999) Large-scale impoverishment of Amazonian forests by logging and fire. *Nature* 398: 505–508.

Peters RL and Lovejoy TE (eds.) (1992) *Global Warming and Biological Diversity.* New Haven, CT: Yale University Press.

Pounds JA, Fogden MPL, and Campbell JH (1999) Biological response to climate change on a tropical mountain. *Nature* 398: 611–615.

Raven PH and Wilson EO (1992) A fifty-year plan for biodiversity surveys. *Science* 258: 1099–1100.

Ray GC, Hayden BP, Bulger Jr. AJ, and McCormick-Ray MG (1992) Effects of global warming on the biodiversity of coastal-marine zones. In: Peters RL and Lovejoy TE (eds.) *Global Warming and Biological Diversity*, pp. 91–104. New Haven, CT: Yale University Press.

Schelesinger WH (1992) *Biogeochemistry – An Analysis of Global Change.* New York: Academic Press.

Smith WH (1989) Effects of acidic precipitation on forest ecosystems in North America. In: Adriano DC and Johnson AH (eds.) *Acidic Precipitation,* vol. II. pp. 165–188. New York: Springer-Verlag.

Smith WH (1990) *Air Pollution and Forests*, 2nd ed. New York: Springer-Verlag.

Smith WH (1993) Forests. In: *Preparing for an Uncertain Climate*, vol. II, Chapter 6. Publication No. OTA-0-567. US Congress, Office of Technology Assessment. Washington, DC: US Government Printing Office.

Smith WH, Hale RC, Greaves J, and Huggett RJ (1993) Trace organochlorine contamination of the forest floor of the White Mountain National Forest, New Hampshire. *Environmental Science Technology* 27: 2244–2246.

Still CJ, Foster PN, and Schneider SH (1999) Simulating the effects of climate change on tropical montane cloud forests. *Nature* 398: 608–610.

Westman WE (1979) Oxidant effects on Californian coastal sage scrub. *Science* 205: 1001–1003.

Wuebbles DJ and Edwards J (1991) *Primer on Greenhouse Gases.* Chelsea, MI: Lewis Publishers.

Population Density

Brian H McArdle, University of Auckland, Auckland, New Zealand

© 2013 Elsevier Inc. All rights reserved.

Glossary

Autocorrelation Densities from adjacent samples are likely to be more similar than are ones far apart.

Compound Poisson distribution A family of probability distributions of numbers per area where some areas have a higher expected density than others.

Poisson distribution The probability distribution of numbers per area if the organisms are distributed completely at random in space.

Measuring Population Density

Generally (although inevitably not exclusively), population density of a species of organism is measured either as a feature of a natural (or manipulated) system that responds to the environment or as a predictor (natural or manipulated) or causal agent of some response of the ecosystem.

As a Response Variable

Population density is often used as a simple relative measure of how an organism responds to local conditions. If conditions are not good for the species, the density will be low (organisms will have died or moved out of the sampled area), whereas if conditions are good the density will be high (organisms will have reproduced and/or immigrated into the area). In this way, changes in density can provide insight into the natural history of the preferences and tolerances of individuals of the species. Of course, if the species is regulated by density-dependent processes (e.g., mortality or emigration) then the relationship of density with the attractiveness of the environment can be obscured. Even though the environment changes in a positive way, there may be no increase in density.

Sometimes, density can be used as an explicit proxy for population size, which of course is what many ecologists want to know about. This is particularly true in applied ecology (e.g., conservation and fisheries science). Unfortunately, the link between population density and population size is not always direct (Gaston and McArdle, 1994). Therefore, definitions of rarity that use either population density or species range are likely to be misleading compared with a definition that uses the product of the two.

The main problem lies in defining the area to be sampled. If it includes the entire population of interest, then the density multiplied by the area gives total population size. However, if the area does not include the whole population then this simple calculation does not work. However, it will perhaps give a relative measure of the population so that changes in the population size will be reflected in changes in the density. Unfortunately, density-dependent processes can weaken this link. If, as the population increased in size, the population was unable to expand the area it inhabited, then the density would increase in proportion to the size. However, if the population was able (or driven by density-dependent migration) to expand its range, then the density could remain constant whereas the population grew. Since range expansion appears to be common, density should only be used as a proxy for population size when the range is constrained, as on islands. In most studies, therefore, density simply gives the number of organisms present in some defined study area. Even today when habitat is increasingly fragmented, this will seldom correspond to a biological population.

As a Causal Agent

In many studies, changes in density are seen as causing changes in the population of the same or other species. For example, many experiments are performed each year in which the density of organisms is manipulated (or simply observed) to investigate the response of individuals in the same species (e.g., intraspecific competition) or other species (e.g., predation or interspecific competition). In such studies, the biological effect of density is generally determined by the number of direct or indirect interactions between individuals (of the same or different species). Density in this case is often acting as a proxy for the number or probability of interactions. There are problems associated with this use of density. In particular, the area has to be defined carefully. If the area for which the density is being estimated includes habitat unsuitable to the study species, then the zeros measured from such sites are not due to sampling error (sampling zeros) but reflect a genuine inability of the species to live there (structural zeros). Including structural zeros in the density estimate would lead to an underestimation of the density that individuals of the same species actually experience. Of course, if it is the response of other species that is of interest, then these areas of habitat may be relevant (other areas may not). Again, the purpose for which the density is to be used determines how one defines the area over which it is to be measured.

There are additional problems. It is well known that even if one defines the extent of the study area, estimating the number of organisms can be difficult. Perhaps less well appreciated is that natural surfaces can be generally considered as fractal (or at least approximately so). The effective area of a given area of the earth's surface experienced by an elephant is considerably less than that experienced by a mite. The area used to calculate the density of the organism is nearly always

anthropocentrically biased. The more similar the organism is to humans with regard to size and habit, the more accurate the estimate of effective area is likely to be, and therefore the more relevant the estimate of density.

Of course, a simple estimate of density can still be used as a relative measure of the effective density. However, it must be remembered that organisms are seldom (if ever) uniformly distributed over a study site. Typically, some parts of the site have a high density and others a low density. If most of the organisms are located in one small part of the site, the average density for the site will reveal little about the density most experienced. The marked nonlinearity of most ecological dynamics means that inferences based on the average density are unlikely to be particularly relevant. The situation is similar though more complex when one is interested in the density of one species experienced by another (e.g., in predation or interspecific competition studies).

Population density is therefore nearly always used as an approximate proxy for the real abundance measure of interest. The degree to which it can be regarded as relevant in a particular study depends to a large extent on issues of scale.

Spatial Patterns

Within a species' range, we can expect the density to vary with the habitat from high (optimal) to low (suboptimal). The effect of this on the ground depends on the scale at which one is sampling. For example, at the largest scale it has often been noted that the density of a species tends to be highest in the core of its range and lowest at the periphery. However, much local variation can still exist within the species range at a number of scales down to that of the individual organism (spatial pattern often depends on spatial scale). Similarly, the range and regions of optimal conditions may be temporally variable. For small animals such as insects, the vagaries of wind and weather can lead to shortlived increases in different geographic regions at different times (**Figure 1**). Other species (e.g., habitat specialists) tend to be far more predictable. Of course, temporal stability can be dependent on the scale: The spatial pattern might be relatively stable at one scale (e.g., geographic region) but not at another (e.g., locally within habitat patches).

It is worth noting that although we may often refer to the density as patchy, the pattern of variation is often more continuous. The pattern is better shown with contour maps rather than with a tiling or mosaic. The edges of the so-called patches may not be real discontinuities but rather areas of lower density. The density within the patches may not be constant. Therefore, although we may often talk in terms of patches, we must think in terms of contours. Of course, thinking in terms of patches is an implicit acknowledgment that the presence/absence data are often more reliable and easily interpreted than abundance or density data. This is because the organism counts produced by most sampling programs are subject to considerable statistical sampling error.

Statistical Patterns in Spatial Variation

One of the most common ways of investigating population density is to count organisms in small representative portions (e.g., quadrats) of a larger study area. Then, one estimates the mean density or total numbers in the study site. One can also deduce many characteristics by examining the frequency distribution of the counts per sampling unit.

The most obvious feature of real data is that the frequency distribution of these counts is long tailed and heavily skewed (**Figure 2**). Low densities are common, but very high values sometimes occur. As a general rule, a relatively small number of sampling units contain most of the organisms, and many of the sample units are empty. This indicates a fundamental difference between total number of organisms and the typical density experienced by members of another species (e.g., a predator or competitor) or the typical density experienced by members of the population.

Thus, the total number of animals in the whole area is estimated by the arithmetic mean $\times$ total area. However, the typical number of animals found in a quadrat, i.e., the density of the species most commonly experienced by other species (e.g., predators) searching at the same scale as the quadrats under study, is given by the median number per quadrat. This is approximated by the geometric mean, antilog (mean (log(count))), or more commonly by the pseudogeometric mean, antilog(mean(log(count + 1))). The number 1 is added to each count to overcome problems with zeros. However, the density most commonly experienced by members of the target species is neither of these. Since most of the animals are in a few quadrats, most experience very high densities. Thus, the best estimate of the effective conspecific density, the average density experienced by a member of the population at the spatial scale of the quadrat, is very different. Simply, there are N_i animals in the ith quadrat that experienced a density of N_i, and the density experienced by the average animal is therefore given by $\sum N_i^R / \sum N_i$, which gives a density that is very different from the other two average densities.

Statistically, there are many distributional models that could fit these kinds of quadrat count data. The most commonly used are members of the compound Poisson family (e.g., negative binomial and Poisson log-normal). That these long-tailed distributions fit in no way suggests that the mathematical processes that generate them have anything to do with the biological processes that generated the numbers. However, there are a few interesting ideas that can be derived from these distributions that may have biological relevance in some situations, and which when combined provide some understanding of one of the few ecological laws discovered to date (*see* Taylor's Power Law).

The simplest way of viewing these models suggests that the expected density of organisms at any point in space derives from some long-tailed continuous distribution (a gamma for the negative binomial, and a log-normal for the Poisson log-normal), but that the number actually recorded is from a Poisson with that expected value. There are many ways of parameterizing these distributions, but perhaps the most useful is with the mean (m) and shape parameter (k). The shape parameter defines the relationship between the variance and the mean. The relationship is often displayed as: $\sigma^2 = \mu + (1/k)\mu^2$ indicating that the variance of the counts increases approximately as the square of the mean, with the rate determined by the shape parameter. A small $1/k$ leads to small variances for a given mean, whereas a large $1/k$ leads to large

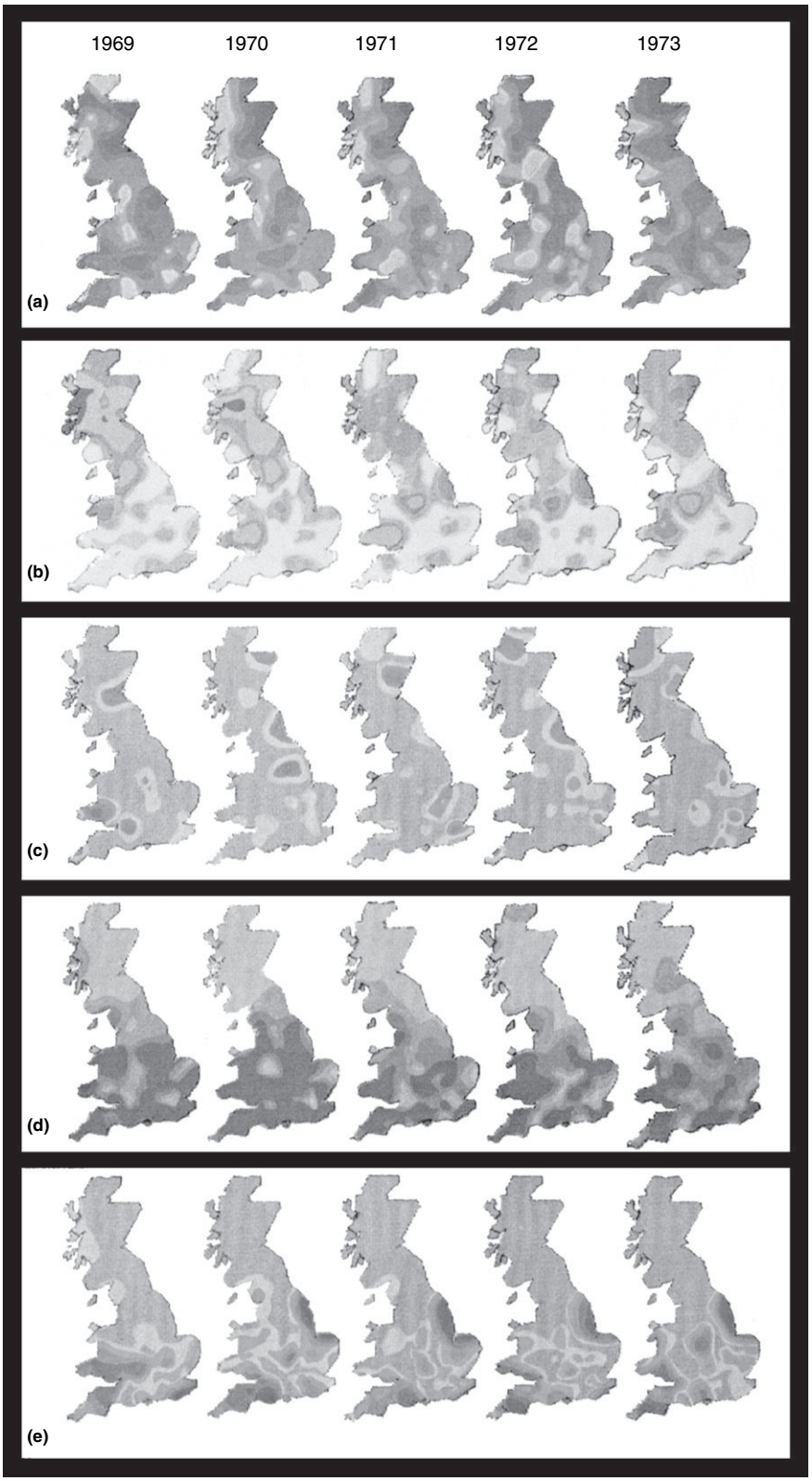


Figure 1 Maps showing the variability of regional relative density in five British moth species with very different dynamics: (a) *Xanthorhoë fluctuata* (golden carpet moth) is a widespread species, numerically fairly stable but with mobile holes in its distribution, (b) *Cerapterix graminis* (antler moth) has a distribution limited by its host plants (fine upland grasses), but within these limits the distribution of density is constantly changing, (c) *Euxoa nigricans* (golden dart moth) has a spotted distribution that is spatially and temporally chaotic (consider how difficult it is to define the range of species such as this), (d) *Spilosoma lutea* (buff ermine moth) has a numerically very stable population with an unstable northern limit, and (e) *Callimorpha jacobea* (cinnabar moth) is highly concentrated in suitable areas but has a highly mobile distribution. Reproduced from Taylor RAJ and Taylor LR (1979) Population dynamics. In: Anderson RM, Turner BD, and Taylor LR (eds.) *The 20th Symposium of the British Ecological Society*, pp. 1–27. Oxford: Blackwell, with permission from Wiley.

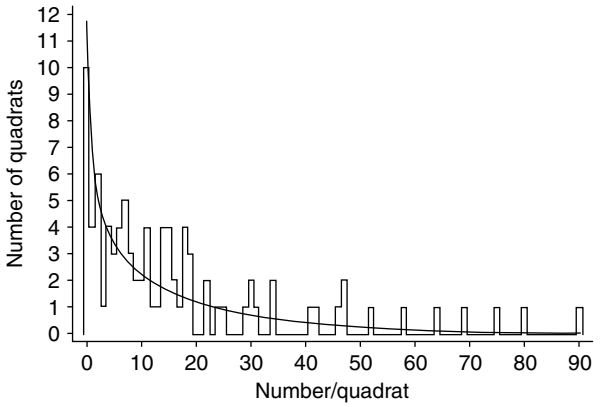


Figure 2 Frequency distribution of density (per 0.25 m^2) of the New Zealand cockle (*Austrovenus stutchburyi*). A compound Poisson (negative binomial) distribution is fitted. Note the extremely crowded quadrats; more than 50% of the animals are in quadrats with densities of more than 30 per 0.25 m^2, although the mean density is only 16.9 (median = 11). The estimated shape parameter k of the negative binomial is 0.61.

variances. Thus, large $1/k$ values suggest high aggregation (some quadrats containing high numbers of organisms and some quadrats containing low numbers), and low $1/k$ implies low aggregation. ($1/k = 0$ suggests a Poisson distribution for organisms whose dispersion is indistinguishable from random.) Therefore, $1/k$ and related statistics are commonly used as the aggregation indices. If the statistical basis of the compound Poisson was in fact correct (i.e., expected density at a point varying spatially, and the number actually collected a Poisson variable), then $1/k$ separates the Poisson sampling error from the underlying distribution of expected density. Indeed, it is the coefficient of variation squared of the underlying distribution of expected density, i.e., the ratio of the variance of the gamma or log-normal with the squared mean:

$$\frac{1}{k} = \frac{\sigma_X^2 - \mu}{\mu^2}$$

Compound Poisson distributions have a biologically relevant property. If a population of organisms is distributed in an aggregated fashion and if some of them randomly die, the mean density changes but the shape or aggregation parameter of a compound Poisson distribution does not. This is of course as it should be: Random mortality leaves the degree of aggregation unchanged. Therefore, if the compound Poisson family of distributions is appropriate, the basic relationship between mean and variance of animal densities (counts) should be $\sigma^2 = \mu + (1/k)\mu^2$ if mortality is operating randomly in space. Of course, there is no need for this to be true (indeed, it might be considered unlikely). Real data might be expected to show a general relationship similar to this, but the degree of aggregation will change with the mean density (possibly due to the density-dependent processes). This is the case for virtually all sets of field data. Therefore, the compound Poisson models are not necessarily always appropriate, but they often fit the available data.

In recent years other distributions have been applied. For example, the Zero-inflated Poisson distribution. The model here again assumes two separate processes: some sample units are from places where the organism is unlikely to be found (modeled by a simple Binomial distribution), others are from places where it can be found at a constant expected density, the number actually collected will be a random count from a Poisson distribution. Apart from the constant expected density in areas where the animal can be found, this version of a compound Poisson (a Binomial and a Poisson) is eminently plausible to ecologists, particularly when the area to be sampled is large enough to include areas the organisms cannot inhabit. It also acknowledges the important distinction between structural zeros (the organisms do not live here) and sampling zeros (they can live here, your sample, by chance, failed to catch any). It is generally more frequent to use an extension of this distribution, the Zero-inflated Negative Binomial, which accepts that in those areas where the organisms can exist they are unlikely to have a uniform expected density. Interestingly, both these distributions also have variance–mean relationships of $\sigma^2 = \mu + c\mu^2$. It is worth pointing out that due to the extreme flexibility of the negative binomial and Poisson log-normal it is usually extremely difficult to convincingly demonstrate, which of the four distributions described above is most appropriate for general use (though the Zero-inflated Poisson is generally the least adequate). Generally they all fit pretty well.

Taylor's Power Law

One of the few ecological laws that have some universality was discovered in 1961 by L.R. Taylor. He worked with populations that had been sampled at many sites on many occasions. He showed that the variance of the densities over the sites on an occasion when plotted against the mean density over the sites on that occasion resulted in a very tight power curve. A plot of log (variance) against log(mean) gave a generally straight line (Taylor's power plot (TPP)). This implies that the formula is $\sigma^2 = \alpha\mu^\beta$ (Taylor's power law). As might be expected from the compound Poisson model, the slope of this line is often close to 2.

This plot and the formula are not just applicable to organism abundances. They are commonly used by statisticians searching for transformations to remove the variance–mean relationship and therefore variance heterogeneity, a problem for many traditional methods of statistical inference. They try to find a transformation that gives a slope of zero to the log(variance)–log(mean) relationship.

Real data are seldom quite so obligingly simple as Taylor's original model implies. Indeed, the dominant pattern is linear in the log–log plot but curved at low densities. Alternative formulae have been suggested to accommodate this. The modified compound Poisson (or Nelder) model $\sigma^2 = \mu + \alpha\mu^\beta$ generally performs well, as does a split straight line, although the threshold that the latter implies seems unlikely to some biologists. Since some populations are demonstrably not Poisson even at the lowest densities, alternatives such as $\sigma^2 = \gamma\mu + \alpha\mu^\beta$ and $\sigma^2 = \gamma + \mu + \alpha\mu^\beta$ have also been used.

Although there may be a variety of models, the basic pattern in the data is clear, but the interpretation is not. In all these models, the parameters α and β have a simple statistical interpretation, but what they imply about the population

dynamics is less clear. The parameter α clearly represents the variance (and therefore the level of aggregation) when the mean density is 1. In the modified compound Poisson (or Nelder) model it represents the value of $1/k$ when the mean is 1. If $\beta = 2$, then this suggests that the level of aggregation is constant at α at all densities. Thus, contrary to much that has been written about the use of the TPP, it is α that tells us about aggregation and not β. The slope of the relationship (β) indicates how rapidly aggregation changes with the mean and not the level of aggregation. To make this clear, we can rewrite the modified compound Poisson model as $1/k = \sigma^2 - \mu/\mu^2 = \alpha\mu\beta^{-2}$. Aggregation is therefore related to density by a power relationship (i.e., a linear relationship in a log–log plot). Thus, $\beta - 2$ indicates how aggregation varies with the mean. Of course, if $\beta \approx 2$ the level of aggregation arguably does not change with density. If $\beta < 2$ then aggregation decreases as density increases. This is consistent with density-dependent migration out of higher concentrations into lower ones, evening out density variation as the overall mean density increases. If $\beta > 2$ then the organisms are becoming more concentrated – for example, if population growth were restricted to areas of higher density with no diffusion. Of course, it would be very unwise to link any value of α or β to a particular dynamic process. One reason is that α and β depend on the particular context (habitat, sampling method and scale, and even the age structure of the population) of the study and not, as suggested by Taylor, on a characteristic slope β for a species and an intercept α depending on the sampling situation. This makes it difficult to interpret TPPs with any confidence. However, with the spatial variance × multiple times data sets originally studied by Taylor, the patterns of variation implied by β can help us understand the biology of the species at least for the site and scale studied.

Interpretation is much more difficult if the TPP is generated by a spatial variance × multiple sites sampling scheme. Here, the variance is over samples (e.g., quadrats) taken within a site, but each point on the graph is given by a different site. Since habitat is likely to vary over sites, we are likely to be looking mainly at patterns of habitat variation and utilization. A $\beta > 2$ might imply that sites with higher means had greater local concentrations of resources than sites with lower means. A $\beta < 2$ might suggest that where the resources were abundant, rather than being locally concentrated they were distributed more uniformly. Of course, the concept of resources is a broad one that includes predator- or competitor-free space and other dynamic features of the environment.

Within-Species Occupancy-Abundance Patterns

In recent years a number of workers have noted that there is a clear relation between the abundance of a species and the number of sampling units in which it is recorded. A number of models have been suggested but they are mainly variants on $p = 1 - (\mu/k)^{-k}$, where p is the probability of a sample unit being occupied, i.e., having an individual recorded on that sampling occasion, and k is the shape, "aggregation," parameter of the negative binomial distribution. This model is simply the proportion of nonempty samples predicted by the negative binomial with those parameters. It has been combined with Taylors' power law to provide a better fit to the data (since k varies with the mean μ if the slope β of the TPP is not 2) – Gaston *et al.* (2006). An alternative approach to describing the relationship has been tried using Zero-inflated Negative Binomial models, but there is little evidence that either of the models fit the data better than the other. Both seem to do adequately.

Density–Biomass Thinning Laws

In this article, population density has been defined as the number of organisms per unit area. However, there is no reason why density should not be measured using other properties of the population. Probably the most common alternative to abundance is biomass. It should therefore not be surprising that there is a relationship between the two measures of density. However, nonbiologists might be surprised by the relationship that seems to exist in plants.

Plants: − 3/2 Thinning Rule (Yoda's Law)

In 1963, Yoda and coworkers noted that if a cohort of plants were monitored over the course of a growth season, there was a clear negative relationship between the biomass and numerical density as the plants "selfthinned." This relationship seems to be well modeled by a power relationship:

$$B = cN^{-b}$$

where B is the biomass density, N is the density, and c and b are constants. The value of $-b$ has generally been thought to be between -0.3 and -0.8 with an ideal value at -0.5. It is generally assumed to be the result of density-dependent mortality and allometric patterns of growth interacting with the way in which trees pack together in the canopy. It is called the $-3/2$ thinning rule because it is commonly formulated in terms of w, the biomass per individual, as

$$w = cN^{-a}$$

where $a = b - 1$. The exponent is now between -1.3 and -1.8 with an ideal value of -1.5.

In recent years, its generality has come into question. There does seem to be a generally negative relationship between density and biomass in most investigated species, but there is little evidence that there is any ideal value of a, let alone one of -1.5.

Animals: − 4/3 Thinning Law

In contrast to the empirical origin of the plants' thinning law, zoologists derived an analogous rule for animals from theory. The postulated relationship is again a power function $w = cN^a$, where $-a = -4/3$ rather than $-3/2$ as in plants. To date, it has been most convincingly shown in salmonid fish but even there the exponent seems to be spatially variable.

Determinants of Spatial Density Variation

Although the existence of spatial variation in the population density of a species is almost trivially selfevident, the processes that can produce it are many and varied. This reinforces the

problems of inferring process from pattern. The mere existence of spatial variation at any given scale is seldom enough to suggest the processes that generated it. There are three main classes of processes that generate spatial variation in density.

External Forcing

The most obvious (and at many scales probably the most common) cause of spatial variability in population density is variation in the quality of the available habitat. If the environment varies spatially at any scale, so will the density of many species. A patchy environment usually produces a patchy distribution of organisms. Of course, some species will be affected more than others. There is enormous variation in the tolerance of organisms to variation in environmental factors.

Another external influence (which might be viewed as a special case of environmental variation) might be the distribution or spatial response of predators or competitors. Predator- or competitor-free space can be as significant a resource as food or space with suitable environmental conditions. Other things being equal, higher density can be expected in such areas.

Lifestyle Characteristics

Of course, interactions among conspecifics can also lead to spatial variation in density. Social aggregations of many kinds are common in animals and at some scales can lead to variation in density that is independent of the environmental conditions. Conversely, random variation in density at local scales can be reduced by spacing behaviors (such as territoriality).

Some of the possible conspecific interactions are not strictly social. Many species modify the local environment, which makes it more attractive to recruits (e.g., larval settlement). This can also result in the formation of aggregations.

Other biological processes result in strong aggregations. Individuals from some species have little choice regarding where they will live (e.g., most plants). Those that start life close to their parent (e.g., many forms of vegetative reproduction, nondispersing seeds, and colonial animals) will stay there. Aggregations will often be the result.

Population Dynamics

Perhaps one of the most topical areas of research concerns the possibility that spatial patterns of density variation could be the result of intrinsic dynamic processes – an emergent property of the dynamic system of which the population is a part. Arguments are often taken from complexity or chaos theory and often refer to the behavior of cellular automata models (such as the Game of Life referred to in most popular texts on complexity and chaos theory).

In fact, as one might expect, the tendency of spatially distributed dynamic systems to produce a spatial pattern is related to a very simple property of the system, which in ecology has a simple biological interpretation. Phipps (1992) postulated that spatial pattern will only be produced in cellular automata models if areas neighboring high-density sites are more likely to achieve high density than other areas. He called this the neighborhood coherence principle. In ecology, this predicts that only species that attract conspecifics either by recruitment (e.g., vegetative growth or limited dispersal from place of birth) or by movement (e.g., flocking) or that somehow reduce the death rate in areas near high densities are likely to show intrinsically generated spatial patterning that is independent of environmental forces. All the sophisticated theory results in a trivial prediction that social and other aggregative behaviors are the main reason why spatial pattern emerges in populations for which the environmental factors have little influence. Generally, the more important question is the following: Given that the biology of the animal predisposes it to aggregation, why is it not more obvious? Density-dependent forces such as predation and emigration that tend to reduce high densities will tend to prevent major concentrations. It is the balance of these ecological forces that produces the observed patterns. This is not to say that there may not be other system dynamics that can produce spatial patterning, but they will almost certainly be similarly spatially nonrandom in their action or somehow be relatable to the forcing processes described previously (*see* External Forcing) and, above all, they will be biologically obvious. There is nothing mysterious about the action of nonlinear dynamics in producing spatial patterns. It depends on the properties of the system and in a biological context seems generally to agree with common sense.

Consequences of Spatial Density Variation

Population Dynamic Consequences

The existence of spatial variation in population density implies consequences at both the individual and the population levels (and other scales in between). The most obvious is that the state of the individual is frequently influenced directly and indirectly by the surrounding density of conspecifics (intraspecific density-dependent processes). The direct effects may include behavioral changes (e.g., increased propensity for emigration), physiological changes (e.g., increased stress hormones and reduced or enhanced sexual activity), or even morphological changes (e.g., the social phase of locusts). It must be appreciated, however, that the density perceived by an individual will seldom be that measured by researchers. Again, the issue of scale is crucial. Measuring density averages counts over areas and reduces variation at smaller scales. These may, however, be the scales at which the individual's perception operates. Local high densities determine the perceived density per individual and the resulting consequences.

Similarly, the indirect consequences of local high density can be very different than expected if one simply examines the average density at a larger scale. The impact of predators, parasites, and pathogens can be very different if the organisms are concentrated rather than uniformly distributed. Spatial variation in density can be as important as differences in overall average density.

These considerations have led to an increasing appreciation that population dynamic models should be spatially explicit. A move in this direction started with optimal foraging models,

but recently entire populations have been modeled in this way. One class of such models is those that describe metapopulations. A metapopulation is a population of subpopulations. Although this is usually an oversimplified spatial pattern – all organisms are in discrete, fixed clusters or in transit between them – the behaviors are quite different from those in a simple uniform population model. Rate constants (e.g., birth or death rates) averaged over a spatially distributed population will seldom adequately reproduce the behavior of a population. It is worth noting that models that use discrete partitions of space do not behave in the same way as those based on continuous space (i.e., on differential calculus). Therefore, although ecologists are currently attempting to produce more realistic models by incorporating space, the way it is incorporated may influence the conclusions.

Genetic/Evolutionary Consequences

Although the population dynamic effects of density variation may be important at a variety of scales, the evolutionary consequences have been considered only at larger scales: a subpopulation of organisms becomes spatially isolated from the rest and allopatric speciation may then occur. Work by D.S. Wilson on multilevel selection theory has incorporated the spatial organization of organisms into groups at different scales to expand the range of evolutionary forces.

One genetic consequence of low density and high aggregation in less mobile organisms is the possibility of genetic drift and inbreeding. Such populations are possibly more vulnerable to disease and can show depression of reproduction rates if the aggregations are stable and isolated.

Temporal Patterns in Density

When population density is plotted against time, a logarithmic scale is often used because populations can fluctuate so widely, over many orders of magnitude, that the arithmetic scale is inadequate. Locusts are an obvious example. However, for many species the magnitude of the fluctuations depends on the temporal and spatial scale over which the density was determined. For example, insect species with a single synchronized generation per year will fluctuate enormously within the year – huge numbers of eggs and first-instar larvae decline throughout the year to much smaller numbers of adults – but might not fluctuate much between years. If one averaged or totalled density over the year, much of the variation would be smoothed out. Similarly, sampling at the same time each year would also not provide information about within-year variability. The spatial scale over which the density is averaged is also crucial. Variation in the density of starlings through time will appear very different when measured at the scale of the bird table, the garden, or the suburb.

Shape of Distribution

Although the nature of the distribution of organism counts sampled in space is well known, far less effort has gone into characterizing the distribution of counts in time. Part of the reason, of course, is the difficulty in obtaining long time series of suitable data. Also, often the data are not independent; rather, they are autocorrelated: The value at one time will be more similar to recent values than those from later times. This makes characterizing the distribution difficult. At this stage all that can be sensibly stated is that, like the spatial distributions, they are discrete and long tailed, and the members of the compound Poisson family could be used as an initial approximation. However, the autocorrelation often leads to multimodality in real, short time series.

Red-Shifted Variance

One observed characteristic of such time series of density data is that, generally, the variability (measured, e.g., as SD(log(N))) increases with the length of the series. There are at least three phenomena, which can produce this effect, and they are not mutually exclusive: (1) *Long-term trends*: If density is consistently increasing or decreasing over time (as commonly occurs due to human impact and climate change), then its variance will naturally increase as long as the trend continues. (2) Internally generated autocorrelation in the series: autocorrelation (or serial correlation) is present when the density at the next sample occasion is similar to the current value, perhaps because some of the same organisms are still living (interval between samples $<$ generation time) or because the number living at the next time period depends on the current breeding population. In this situation, short time series cannot vary widely, but as the series becomes longer the variance gradually increases to some asymptotic level. This is of course particularly common with long-lived species that, naturally, we sample at a frequency convenient to us but not necessarily relevant to the species. (3) Externally forced autocorrelation: If the physical environment fluctuated in an autocorrelated fashion then the population's density might vary in response. Such autocorrelated time series are characteristic of "red-shifted" spectra of environmental variables (characteristic of climatic factors). It may be worth mentioning that Power law distributions that describe scale-free properties of certain systems have, for most realistic values of their parameters, variances that increase with sample size.

Taylor's Power Law

If time series of densities are produced for many sites, then a plot of the variance over time against the mean over time for the sites will provide a close relationship, as occurs for spatial variances and means. Essentially the same models describe this relationship as those used for the spatial TPP. Temporal variance increases with the mean as does spatial variance. Again, the slope is usually approximately 2. The intercept α measures the variability when the average density at a site is 1, and therefore it describes a sort of baseline temporal variability at this spatial and temporal scales. The slope β describes how this variance changes for sites that have higher or lower densities. A slope of 2 suggests that there is essentially no change in variability at higher densities. Values higher than 2 indicate that the species is more variable at denser sites. This is a behavior characteristic of species subject to local

explosions in numbers at certain sites. A good site is sometimes good and sometimes bad. A bad site is always bad. A slope of less than 2 means that sites with generally high densities tend to be less variable than poor sites. This is consistent with density dependence, although other habitat-driven, or demographic stochasticity explanations are possible. Slopes of more than 2 appear to be relatively rare; less than 2, common.

Determinants of Temporal Density Variation

Although the processes that can produce temporal variability in density are similar to those that produce spatial variability, they are not identical. The effects of choices of scale tend to be greater. However, the same three classes of factors are discussed.

External Forces

Fluctuations in the environment will drive temporal variability of density in many species. Invertebrates in particular are at the mercy of the elements, although vertebrates are demonstrably not immune. The extreme winter of 1962 and 1963 in the British Isles resulted in a marked decrease in the densities of many bird species that took years to recover. The densities of many marine organisms are driven by El Niño-like phenomena (as are many terrestrial organisms). The threshold-like effects of many environmental variables can lead to large fluctuations in organism densities. For example, a species may not breed until the temperature is above some value (which may only happen every few years) or unless rain falls (especially in desert species).

Of course, the consequences for a species of such fluctuations in the physical environment may be indirect. Food, competitors, disease, and predation may also be driven variables, in many cases amplifying or extending the direct effects. Clearly, an organism's lifestyle (food/nutrient requirements, vulnerability to disease or predation, etc.) can also influence its susceptibility to such forces.

Life History Characteristics

Life history characteristics, such as synchronized generations or life stages (perhaps dormant) that are not available for sampling (e.g., the seed bank or the nymphs of the 21-year cicada), can lead to apparently major changes in density. Arguably, the variation is the result of a particular definition of density. Whether such a definition is useful will depend on the question under investigation.

Some species have reproductive strategies that amplify fluctuations in density. Physiological and/or morphological changes (usually triggered by habitat characteristics) can lead to dramatic increases in population growth, for example, the shift from sexual to asexual reproduction when conditions are good.

An organism's mobility (or its ability to disperse propagules) can also influence its temporal variability. If it has an aggregated, patchy, spatial distribution (e.g., metapopulation structure) the degree of connectedness between the parts of the population and the degree to which they fluctuate independently can influence the extent of the variation of the whole population.

Population Dynamics

By definition, density-dependent processes should be major determinants of the degree of temporal variation. In fact, the extent to which such processes influence density is largely unknown. The problem is two-fold: (1) It has proved easier to study the apparent consequences rather than the process. That is, it is easier to show that changes in density over a period depend on densities prior to that period (the statistical detection of density dependence) than it is to show the ecological process that produced the relationship. However, there is little doubt that internal density-dependent processes cause many species to fluctuate less than they might otherwise. (2) The second process is more subtle. Complexity and chaos theory predict that certain combinations of high growth rate and strong density-dependent dynamics can lead to chaotic behavior in the resulting population densities. Since this can lead to apparently random fluctuating densities, it is difficult to separate the effects of density dependence from the results of external forces. Indeed, since these driving variables may be the result of chaotic processes (e.g., the weather or market forces), it is difficult to determine how such density-dependent effects can be distinguished without much more data about environmental factors than is generally available.

Consequences of Temporal Density Variation

Increased Probability of Extinction

The most widely discussed consequence of low density is the increased probability of extinction. Although it is difficult to demonstrate empirically, it seems obvious that a population whose size (not necessarily density) varies widely is more likely to go extinct (all things being equal) than a population whose size varies less. Once a population reaches low densities within a restricted spatial range, chance events can determine whether it has a future. If the numbers are sufficiently low a single death or failure to breed can lead to eventual extinction. It is worth reiterating that such stochastic extinctions are the result of low population sizes and not densities.

The probability of extinction can (occasionally) be influenced by low density *per se*. For example, regardless of the population size, at very low densities Allee effects may increase the chances of extinction. That is, if the density of a sexually reproducing species becomes so low as to affect the probability of locating a mate, then the population can start a slide to extinction from which it cannot escape. Of course, if the population responds by aggregating, then the distinction between local and regional density becomes crucial. Ultimately, it will be the fate of the aggregations, and therefore the total number of organisms, that determines the destiny of the population.

Cascade Effects

Of course, fluctuating densities of one species can have knock-on effects on other species. Patterns of density variation affect man's exploitation of the natural world. Outbreaks of pests or

periods of low fish density have economic consequences, and assessing the risk of such events is a lucrative job for well-paid consultants.

Such an anthropocentric view ignores the cascade of effects that are likely to propagate up and down the food chain, for example, the changes consequent on fluctuations in the density of kelp on the west coast of North America. Such effects are usually difficult to demonstrate unequivocally.

Evolutionary Consequences

Temporal variation in density can produce genetic bottlenecks with consequent loss of genetic variability. However, if the population rebounds rapidly the reduction in heterozygosity may be minimal.

Trans-specific Patterns

Comparing Densities across Species

It should be apparent that both the measure and the scale of a density must be appropriate for the processes or patterns of interest. In particular, the appropriate area might not be the same for different species. If perceived density of conspecifics is required, the appropriate area would probably depend on the movement patterns of individuals of the different species and would likely differ among species. If perceived density by a predator species is required, then the appropriate area would be defined for the predator and would likely be the same for most prey species. Since the biological relevance of a density depends on the processes being investigated, it is likely to vary between species. As a result, it is extremely difficult to interpret patterns of simple arithmetic density measured across species at some arbitrary scale. This is not to say that such patterns are not interesting (indeed they are sometimes tantalizingly so); however, interpreting the causes and consequences of these patterns will be more difficult because the measure and scale chosen will be to a greater or lesser extent irrelevant to the dynamics of most of the species.

In addition, many other factors complicate such trans-specific comparisons, for example, phylogenetic relatedness, which leads to statistical nonindependence and possible artifactual detection of pattern. Trans-specific patterns of population density must therefore not be taken at face value. Their meaning is seldom as simple as it appears.

Distribution of Density over Species

It has long been observed that organism abundances are approximately log-normally distributed over species, regardless of the size of the sample area. Preston (1962) suggested that this distribution, which he named the canonical log-normal, had a remarkably constant property, $(2(\mathrm{SD}\ \log_2)^{-0.5}) \approx 0.2$ (where SD $\log_2$ is the standard deviation of the abundances after transforming using log base 2). This translates into a SD $\log_{10} \approx 1$. In fact, subsequent work has demonstrated that nature is not so kind, but the SD $\log_{10}$ does appear reasonably consistent. An explanation based on the law of large numbers was given by R. May in 1975, who suggested that this phenomenon is inevitable when large numbers of species are involved. Alternative models have been suggested, notably MacArthur's broken-stick model and the geometric log series. Although the log-normal model does not imply any ecological generating process, the other models usually suggest such a process. These (and other more complex niche partitioning models) often fit as well as a log-normal, but there is a danger that workers will infer a process from an adequate fit. The amount of information in a single sample frequency distribution is clearly not sufficient to distinguish between generating processes, especially given the difficulties of defining and comparing densities across species. It is worth noting that, like the Normal distribution, the Log-normal is a Stable distribution; meaning that it can be generated by a linear combination of an arbitrarily large number of other distributions. In the Log-normal case their values must be logged before combination. It can therefore be the result of the combination of a large number of multiplicative processes. No single process can be inferred from any data that could plausibly be log-normal.

Trans-species $-3/2$ Thinning Law

The $-3/2$ thinning law has a trans-species extension (although the word thinning seems inappropriate). If we plot the maximum recorded density of species against their biomass density we should get a classic allometric power curve $w = cN^{\alpha}$ with an exponent α of close to -0.5. In 1998, Enquist *et al.* suggested that for the curve α is closer to -0.333 (a $-4/3$ thinning law like that suggested for the intraspecific relationship in animals). They suggest a theoretical support for this value.

The existence of a negative relationship between body size and density in animals has long been accepted. However, research suggests that there are two main shapes to the relationship on a log–log plot: linear negative (**Figure 3(a)**) and polygonal (**Figure 3(b)**). Other forms, usually nonlinear with a generally negative slope, have occasionally been suggested for other groups, but currently it is difficult to assess their generality.

The way in which the densities are estimated clearly influences the form of the relationship; in more than 500 studies investigated by Blackburn and Gaston (1997), the scale of the area considered (local vs. regional) was the most important factor. This reinforces the idea that until we know and understand the nature of the density that is being measured we cannot interpret any relationships observed with which it is involved.

The linear negative relationships had highly variable slopes and there was no tendency for them to cluster around -0.333. Indeed, all the slopes (including manifestly polygonal ones) had a mean of -0.51 and a mode near -0.75. Other workers have also found that an exponent of around -0.75 appears quite common, though by no means universal, and that no one theoretical explanation for the trend can be supported.

Evidence for the polygonal relationship appears to be declining as more systems are investigated.

Relationships with Other Traits

Breeding Systems

Characteristically low-density species (especially plants) rely less on outcrossing and sexual reproduction in general than do more common species.

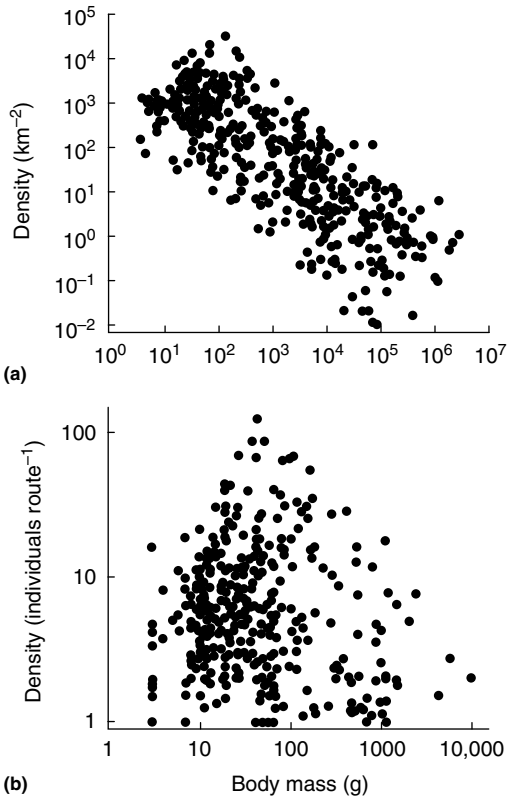


Figure 3 Density–body size relationships showing the two main types. Note that despite having generally negative slopes, both have peak densities at intermediate body masses. This is a common phenomenon of unknown significance: (a) selection of mammals of the world (reproduced from Damuth (1987) as cited in Blackburn TM and Gaston KJ (1997) A critical assessment of the form of the interspecific relationship between abundance and body size in animals. *Journal of Animal Ecology* 66: 233–249) and (b) density (individuals per BBS route, $n=380$) for North American land birds (reproduced from Brown and Maurer (1987) as cited in Blackburn TM and Gaston KJ (1997) A critical assessment of the form of the interspecific relationship between abundance and body size in animals. *Journal of Animal Ecology* 66: 233–249), with permission from Wiley.

Reproductive Investment

It has been suggested that rare species (restricted range or low density) produce smaller numbers of propagules per unit time, although they often have longer life spans.

Dispersal Ability

In general, studies have shown that rare species (with restricted ranges and lower regional density) have poorer dispersal abilities than more common ones.

Homozygosity

As expected from theory, there is ample evidence that less dense species have lower genetic diversity than more common species.

Resource Usage

Since density, as previously noted, is often driven by environmental factors, it is not surprising that low-density species are frequently associated with restricted or uncommon conditions and resources.

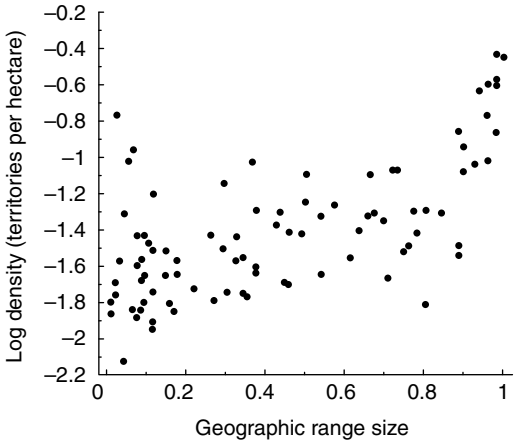


Figure 4 The relationship between local density ($\log_{10}$ territories ha^{-1}) and the proportion of sites occupied on farmland Common Bird Census plots in Britain in 1975. Reproduced from Gaston KJ, Blackburn TM, and Lawton JH (1997) Interspecific abundance–range size relationships: An appraisal of mechanisms. *Journal of Animal Ecology* 66: 579–601, with permission from British Ecological Society.

Trophic Status

One of the oldest ecological concepts, the ecological or Eltonian pyramid (Elton, 1927), predicts that the higher a species is in the food chain, the lower its density. There is ample evidence that this is usually true, but there are enough exceptions to remind us that few ecological rules have universal generality.

Latitude

For many years it has been suggested that density for a species tends to be higher in higher latitudes. (This surely has to be true for blood-sucking insects.) In part, this is a prediction of the density compensation hypothesis, which predicts that in communities of low diversity (e.g., islands) density will be higher. Although there is evidence to support this prediction for islands, the evidence for latitude is more equivocal.

Geographic Range

The relationship between abundance and range was first noted (publicly) by Darwin and has occupied many workers since. That it exists is not in doubt (**Figure 4**), but whether it implies anything of biological interest is debatable. The following is the central question: Does such a relationship imply anything about differences in life history between rare and common species? It is easy to show that in an environment that varies spatially in its attractiveness or ability to support species (as in the compound Poisson model described earlier) but in which all species have similar properties besides mean density (or at least whose properties are independent of their average density), there must be a positive relationship between density and the proportion of sites at which they are found. Such a relationship therefore does not imply anything about the life history or lifestyle of common compared to rare species (Kunin and Gaston, 1997).

To demonstrate that such differences do exist, it is necessary to compare the observed distribution of organisms with an appropriate null such as that discussed in the previous paragraph. This model is sometimes called the "sampling" or "artifact" model and has been shown to fit some data very well. However, it is not simply a sampling phenomenon, disappearing if a sufficient number of organisms have been collected or a sufficient number of areas are sampled. It is predicted by many population dynamic processes in which rate parameters (e.g., birth rate and death rate per individual) are independent of local density. Efforts to investigate the density–range relationship in the past have been made more difficult by only using the density at sites at which the species is actually present – local density (treating all zeros as structural and therefore ignoring them).

In fact, there is an easier approach to the problem. Since the probability of being found at a site depends on the degree of aggregation as well as on the mean, the null (the artifact model) assumes that aggregation at the chosen scale is constant (or at least that common species have the same distribution, and therefore mean, of aggregation as rare species). A trend showing that common species were less aggregated (lower $1/k$) than rare ones would show that common species were found at more sites than the null suggests, i.e., they have wider ranges than would be expected if they had the same life history characteristics as the rare ones. Such a trend might be considered interesting (although there might be many explanations – some trivial – for such a pattern). A trend showing that they were more aggregated would be even more interesting. It is worth noting that most of the studies performed to date use either snapshots of the spatial pattern of species or maps accumulated over some extended period of time. Both of these ignore the temporal variation of range (and mean density) that is typical of many groups of organisms. If such temporal information is available, examining each species separately and demonstrating life history phenomena for each and then searching for a trend for those over mean density would be more productive. The author discusses this in more detail in the section Relationships with Other Traits.

Trans-specific TPPs

Although Taylor's power plots were originally defined for single species, it is possible to plot variance–mean relationships across species, with each species contributing a single point. Such plots are clearly difficult to justify if the area over which the data are collected is not equally relevant to all species and the means and variances used are not truly representative of the species. However, if these problems are ignored, then the trans-specific TPP can address the density–range problem by plotting the level of aggregation against mean density. The null described in the section Relationships with Other Traits now becomes equivalent to a TPP slope (β) of 2 or a horizontal line (slope $= \beta - 2$) on a log($1/k$) versus log(mean) plot. If $\beta < 2$ then aggregation appears to decline as mean density increases; the range of common species is greater than the null would suggest. However, before accepting an alternative of differences in the life history characteristics of rare and common species, another null must be considered and rejected. If all the species have individual slopes $\beta < 2$ (a very unlikely event in most groups), then a trans-specific TPP will tend to have a slope of less than 2 as well, even if the rare and common species have the same life history characteristics. In most situations, it will be impossible to check this since if there were enough information to produce spatial variance TPPs for each species, it would be unnecessary to produce a trans-specific plot in which each species was represented by a single point. It would be more informative to fit separate TPPs and search for trends in the parameters.

Conclusions

The concept of density is an amorphous one. It has to be specifically determined for a particular purpose. The area used for both the sample unit and the study area is crucial and can only be determined after considering what aspect of density is relevant to the current study, and it may even then be difficult to determine since it depends on information about the species' biology that may not be available (e.g., degree of aggregation and movement patterns). This makes comparisons across species especially difficult to interpret (particularly from published studies). That there is significance in all these patterns seems obvious, but interpreting them rigorously is a considerable (possibly insuperable) challenge.

See also: Conservation Genetics. Measurement and Analysis of Biodiversity. Population Dynamics

References

Blackburn TM and Gaston KJ (1997) A critical assessment of the form of the interspecific relationship between abundance and body size in animals. *Journal of Animal Ecology* 66: 233–249.

Elton CS (1927) *Animal Ecology*. London: Sidgwick and Jackson.

Enquist BJ, Brown JH, and West GB (1998) Allometric scaling of plant energetics and population density. *Nature* 395: 163–165.

Gaston KJ, Borges PAV, He F, and Gaspar C (2006) Abundance, spatial variance and occupancy: Arthropod species distribution in the Azores. *Journal of Animal Ecology* 75: 646–656.

Gaston KJ and McArdle BH (1994) The temporal variability of animal abundances: Measures, methods, and patterns. *Philosophical Transactions of the Royal Society London* 345: 335–358.

Kunin WE and Gaston KJ (1997) *The Biology of Rarity. Causes and Consequences of Rare–Common Differences*. London: Chapman & Hall.

May RM (1975) Patterns of species abundance and diversity. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*. Cambridge: MA: Harvard University Press.

Phipps MJ (1992) From local to global: The lesson from cellular automata. In: DeAngelis DL and Gross LJ (eds.) *Individual Based Models and Approaches in Ecology: Populations, Communities and Ecosystems*, pp. 165–187. New York: Chapman & Hall.

Preston FW (1962) The canonical distribution of commonness and rarity. *Ecology* 43: 185–215, 410–432.

Population Diversity, Overview

Jennifer BH Martiny, University of California, Irvine, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 759–767, © 2001, Elsevier Inc.

Glossary

Allele One of several forms of a gene.
Chromosome A long, threadlike structure in the nucleus of a cell composed of DNA and protein.
Conspecific Of the same species.
Gene flow The incorporation of genes from one group of individuals into the gene pool of another group.
Gene pool The total of all genes in a group of individuals.

Origins and Detection

Populations are usually defined ecologically or genetically. The ecological entity is called a demographic unit, a group of individuals whose population dynamics are largely independent of those of other groups. The genetic entity is called a Mendelian population, a genetically distinguishable group of individuals evolving somewhat independently of other groups. Demographic units may be Mendelian populations and vice versa, but the two are not necessarily the same. Moreover, as is the case when defining any taxonomic boundaries, populations are not always clear, discrete units. Although the majority of individuals in a population breed with members of the same population, some individuals may move and interbreed with individuals from other populations, making a clear cutoff between most populations impossible. Because data about genetic populations are more readily available than that about demographic populations, here a genetic definition is used.

Just as populations can be defined in multiple ways, so can population diversity. Population diversity can refer to the amount of genetic diversity among a collection of populations. Populations that exchange genetic material frequently will not be as diverse as populations that are isolated, perhaps because of a geographic barrier, from one another. This article defines population diversity as the number of populations within a defined area and considers both the average number of populations per area for a species as well as total population diversity across the globe.

The following sections review the processes by which conspecific groups of individuals diverge into genetically distinct populations and how these genetic differences are detected.

Variability within Species

Much biodiversity exists within a species. This variation occurs at three different levels. First, there is a genetic variation within individuals of sexually reproducing species. Individuals carry multiple copies of most genes (two copies in diploid species such as humans), and these copies (alleles) may be different. Second, there is genetic variation between individuals. Except in the case of identical twins, no two individuals will have the same alleles for every gene. Finally, there are genetic differences among populations. A population may contain alleles that other populations do not have and vice versa. In addition, two populations may differ in the proportion of individuals that carry alleles shared by both the groups.

Processes of Divergence

Individuals are not usually spread out evenly within a species' range but are clustered into localized groups. When these groups have limited genetic exchange, they may diverge into genetically distinguishable populations. (Methods to detect genetic divergence between populations are discussed later, *see* Importance.) Populations can diverge through drift and natural selection. These agents of microevolution result in genetic differentiation by acting on populations at different rates or in different directions.

Genetic Drift

Gene frequencies may change from one generation to the next simply because of chance. Particularly in a small population, random sampling error in the reproduction of alleles may contribute to changes in allele frequencies from one generation to the next. For instance, an allele that is only represented in a few individuals may be completely lost in the next generation if those few individuals by chance fail to reproduce or pass down that particular allele to their offspring. In addition, new alleles may be produced by mutation and increase in frequency in one group such that a unique allele is present in some groups and not others.

Natural Selection

Natural selection is the differential reproductive success of genetically different individuals resulting from the interaction of the individuals' inherited traits and their environment. Thus, different environmental conditions or biotic communities in a habitat can result in different selection pressures on populations. Some alleles may be beneficial to the reproduction of the individual carrying them in one population, whereas in another population the same allele may be relatively neutral or harmful to the carrier. In other words, selection in these two groups operates at different rates or directions, contributing to their genetic divergence.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00114-3

Factors Influencing Divergence

Two factors, population size and dispersal, greatly influence the rate and importance of the previously mentioned agents of divergence. The smaller the population size, or number of individuals in a group, the greater the effect of genetic drift. The probability of gene frequencies in the next generation being representative of the previous generation decreases as the number of reproducing individuals decreases.

Dispersal, or the migration of individuals from one group to another, affects the amount of gene flow between groups. When an individual successfully reproduces in another group, the genetic composition of the new group becomes more similar to the group from which the immigrant originated. Thus, gene flow has the effect of homogenizing genetic composition, thereby decreasing the rate of divergence resulting from the three agents discussed previously.

Detection of Divergence

In general, the degree of genetic divergence between groups of organisms is measured by comparing the amount of variability within a group to the variability between groups. Thus, the genetic diversity embodied by the individuals in each group must be assessed in order to detect genetic divergence. Genetic diversity may be estimated indirectly by measuring morphological characteristics and migration rates or directly by examining the genetic material. A survey of some of these methods is provided in the following sections.

Morphological Variation

The variation that is observed among individuals is the result of their genetic composition, their environment, and the interaction between these factors. Thus, morphological variation will in some part represent underlying genetic variability, and it will also be intertwined with variation caused by environmental factors. Nevertheless, in the absence of direct genetic methods, morphological variation of a suite of traits is often used to resolve the population structure of a species. There have been attempts to tease apart genetic and environmental factors by using breeding experiments. In these studies, the environment is kept constant and morphological traits of parents and offspring are measured. One can then estimate the heritability, or the variation resulting from the genetic background, of a trait in the experimental conditions.

Migration Rates and Population Size

The migration rate of reproductive individuals into a group gives an estimate of the amount of gene flow, or exchange of genetic material, with outside groups. The lower the gene flow into a group, the greater the impact of genetic drift. The other critical factor affecting the degree of genetic drift in a population is its size: the smaller the population, the greater the expected drift. Thus, a combination of migration rate and population size estimates gives an indirect estimate of the amount of genetic divergence that occurs resulting from genetic drift.

Protein Variation

Currently, most studies on population differentiation examine DNA variation directly. However, before these methods became cost-effective, researchers assayed protein (allozyme) variation among individuals between groups. The composition and structure of a protein are determined by the sequence of nucleotides (the subunits that make up the DNA molecule) comprising that gene. Alleles of the same gene have differences in their sequences, and each allele gives instructions for making a particular protein. The proteins can be distinguished by their rate of movement through a gel medium that is exposed to an electrical field. The movement rates of proteins differ because of different sizes and net charges of the molecules. By sampling many individuals in a group, one can assess the number of unique alleles and their frequencies. Thus, the genetic divergence between groups can be estimated from the variation of a random sample of proteins. The greater the differences in the proteins and their frequencies among groups, the greater the genetic divergence.

DNA Variation

Comparing the nucleotide sequence of particular genetic regions offers the greatest resolution of genetic differentiation. The genetic divergence between groups can then be estimated by comparing the similarity of the sequences among individuals in various groups. Usually, specific sections of an organism's DNA are targeted using the polymerase chain reaction (PCR) method. This section of DNA can then be examined by a variety of methods that vary in their resolution from sequencing to fingerprinting methods.

Fingerprinting methods have been commonly used when direct nucleotide sequencing was not cost-effective. For instance, restriction fragment length polymorphism analysis of a PCR-amplified gene region offers a broad examination of DNA differences among samples. The isolated DNA is exposed to proteins that break the DNA wherever a particular sequence of nucleotides occurs. The length of these fragments can be distinguished by their rate of movement through a gel medium that is exposed to an electrical field (similar to the method used to detect protein variation). The greater the difference in the pattern of fragment lengths, the greater the difference in the piece of DNA.

As the cost of sequencing continues to fall, met genomics approaches are also likely to become standard for population differentiation studies. This method does not target a particular gene (and therefore eliminates the problem of PCR biases) and instead involves sequencing random fragments from a pool of environmental DNA. The approach requires a large amount of computational power to organize and compare sequences among samples; however, this obstacle is quickly being reduced and computing power and storage becomes widely accessible to the researchers.

Indices of Divergence

A variety of indices are used to evaluate the population structure or the genetic divergence within a species. One of the most common indices reported is Sewall Wright's F_{ST}. This statistic can be calculated from the allele frequencies of a gene or genes and describes the variation among groups relative to the variation within the groups. The greater the variation that can be attributed to differences between groups rather than

within groups, the greater the F_{ST} value. Other measures of divergence describe the genetic distance between groups. These measures are indices of similarity or dissimilarity and are calculated using any of the types of data described previously, including protein frequencies and DNA sequences. Groups that are more isolated and for longer periods of time are expected to be more genetically distant.

Importance

Many of the benefits that biodiversity confers on humanity are delivered through population diversity. These benefits include esthetic enjoyment, species conservation, discovery and improvement of pharmaceuticals and agricultural crops, replenishment of stocks of economically valued species, and delivery of ecosystem services.

Again, population diversity is defined here as the number of populations in an area. Certainly, characteristics of a population, such as the area it occupies and the number of individuals it includes, will affect the benefits provided by a single population. In general, however, as the number of populations of a species increases, its area occupied and number of individuals increase. Thus, the following discussion focuses on numbers of populations and ignores population characteristics.

Esthetic Value

Populations of organisms, their physical environments, and the interactions between them make up natural ecosystems. People value the esthetic benefits of the populations in these habitats. For instance, bird watchers enjoy bird populations, hikers appreciate shade trees, and divers seek out reef fishes. Others simply enjoy the scenery created by the sum total of populations in an area.

Species Conservation Value

By definition, populations are essential to the conservation of species. Specifically, the number and size of populations influence the probability of species extinction; a species with many large populations is less susceptible to the extinction than a species with a few small populations. Migrants from other populations can prevent the extinction of a population by contributing individuals when numbers are low or supplying genetic variation needed to adapt to changing environmental conditions. If local extinction does occur, individuals from other populations can find a new population in the area. The threat of rapid global climate change makes population diversity within a species especially critical; a species with many populations is more likely to include the variation necessary to adapt to new conditions.

Genetic Value

Populations of the same species may produce different types or quantities of defensive chemicals, compounds that may be of medicinal or agricultural value to humans. The story of the development of penicillin provides a clear example of the importance of population diversity to pharmaceuticals. The widespread use of penicillin as a therapeutic drug did not occur until 15 years after Alexander Fleming's discovery of the compound in common bread mold. One reason for this delay was a search to find a population of the mold that produced greater quantities of penicillin than did the original population.

Population diversity among wild crop relatives supplies genetic material to agricultural strains. The world's three main crops (rice, wheat, and maize) are widely planted in genetically uniform strains. The emergence of a new disease or pest can therefore threaten large fractions of the harvest at a time. Wild crop relatives contain genetic variation that does not exist in the agricultural populations, including genetic resistance to a variety of diseases and pests. Thus, when an agricultural strain is susceptible to a disease or pest, researchers search for genetic resistance in wild crop relatives. Thousands of populations may have to be tested until one is found that carries the genetic resistance that can be used to protect the crop. For example, when the grassy stunt virus emerged as a serious threat to the rice crop in Southeast Asia in the late 1960s and 1970s, a search for resistant rice varieties was conducted. Only one strain of a wild rice at the gene bank of the International Rice Research Institute was found to resist the virus. Genetic variation in wild populations will also be crucial in providing genetic material to sustain crop yields in the face of human-induced changes in growing conditions.

Direct Economic Value

The reduction of economically important species often has direct consequences for local peoples. For instance, the decline of fish harvests leads to the loss of income for fishermen and the loss of an important source of protein for much of the human population. Populations of economically valued species not only act as a safety net against species extinction but also increase the species' harvest level. If more populations exist, then more populations can be harvested. Populations may also increase harvests indirectly. A collection of populations that are connected by dispersal is called a metapopulation. In a metapopulation, dispersal of individuals from other populations can rescue a population from extinction when its numbers are critically low. Over the long term, a population will be larger if it is a component of a larger metapopulation than if it is part of a smaller metapopulation or completely isolated.

Ecosystem Service Value

Ecosystem services – services that natural ecosystems provide to human civilization for free – are perhaps the most important benefits that populations confer on humanity. They include natural processes such as purification of air and water, detoxification and decomposition of waste, generation and maintenance of soil fertility, pollination of crops and natural vegetation, and pest control. Populations deliver these services, and therefore population diversity at global, regional, and local scales affects the quality and quantity of the services

provided. (Except where stated below, this article focuses on how the number of populations affects ecosystem services, although the size, density, and genetic composition of populations are also important factors.)

Global Population Diversity

Higher global population diversity probably enhances the delivery of global ecosystem services such as regulation of biogeochemical cycles and stabilization of climate. For example, the larger the area remaining under natural tree cover in the Canadian taiga, the greater the amount of carbon stored there. Although deforestation in this region may not result in the extinction of any tree species, a large-scale loss of tree populations would affect the global balance of greenhouse gases in the atmosphere.

Regional Population Diversity

For many ecosystem services, global population numbers will not be as important as the population diversity in the region of interest. For these services, it is necessary not only that there be many populations somewhere in the world but also that they exist in the region where the services are to be provided. These services include water purification by forests and wetlands and mitigation of floods and droughts by forests. Loss of these services occurs when the populations in forest and wetland habitats are destroyed in one area, regardless of the continued existence of these populations elsewhere.

New York City provides an excellent example of the value of regional population diversity for water purification. The city was renowned for its clean water, which came from the Catskill Mountains 100 miles to the north. Natural purification processes, performed by populations of plants and soil organisms, were sufficient to cleanse the water for most of the city's history, but in recent years land development and related human activities reduced the effectiveness of these processes. In 1996, city water officials floated an environmental bond issue to purchase land, freeze development on other lands, and subsidize the improvement of septic tanks. It is expected that these actions will restore and safeguard the local populations of the organisms that filter and purify the water. If so, an investment of $1 billion in natural purification services will have saved city taxpayers $6–8 billion, the additional avoided cost (more than 10 years) of building a water treatment plant.

Regional population diversity is also essential for pest control. This function of populations is easy to take for granted, but it is dramatically illustrated when one kind of organism is transplanted to a new environment that lacks populations of predators that usually keep its numbers in check. A classic case is the importation of prickly pear cacti (*Opuntia*) into Australia by early settlers. In the absence of their normal predators, the cacti spread more than approximately 25 million ha in New South Wales and Queensland. About half of the area was so densely covered with the cactus that the land could not be used for farming or ranching, and the costs of poisoning or removing the *Opuntia* were more than the value of the land. The problem was solved by importing a cactus-eating moth, *Cactoblastiscactorum*, from the South American homeland of the *Opuntia*. Once regional populations of the moth were established, the *Opuntia* populations were decimated. Although the cactus can still be found in Australia, it now occurs only as scattered clumps.

The importance of regional populations of native pollinators for agriculture is made clear by their decline (chiefly a result of pesticides and habitat destruction). For example, for more than 60 crops grown in the US, farmers have to pay honeybee keepers to bring hives to the fields or orchards to be pollinated. This service is estimated to cost farmers more than $60 million a year and the federal government more than $80 million in subsidies. Problems within the bee-keeping industry (disease and hybridization with the Africanized honeybee) are increasing these costs.

Local Population Diversity

The number of populations at a particular location – the local diversity of species – affects local ecosystem function. In many greenhouse and field experiments, plant productivity has been found to increase with the diversity of plants. Higher diversity also seems to be associated with greater stability of plant productivity. For instance, more diverse grasslands appear to be more resistant to drought and grazing disturbance than less diverse areas. Thus, local population diversity likely influences the amount and variation of services provided by an ecosystem. Also, because regional and global services are performed by an assemblage of local ecosystems, this scale of population diversity will affect the delivery of larger scale services as well. For example, global biogeochemical cycles are probably influenced not only by the total number of tree populations on the planet but also by the diversity of trees within a habitat.

Furthermore, traits that affect local ecosystem functioning may differ among populations. Treseder and Vitousek (2001) studied multiple populations of two ecologically important tree species in Hawaii. They found genetically based differences among populations of *Metrosideros polymorpha* in the decomposability of leaf litter and among populations of *Acacia koa* in the potential to fix nitrogen. Thus, these populations are not interchangeable; if they were replaced with other populations of the same species, the functioning of the native ecosystem would be altered. Indeed, there is a growing literature on the importance of genotypic diversity within and among populations for ecosystem functioning and the delivery of ecosystem services. For instance, Schindler *et al.* (2010) found that the presence of genetically distinct sockeye salmon populations in Alaska stabilizes (reduces the variability of) total fishery abundance within and among watersheds. Thus, high population diversity translates into fewer fishery closures and greater economic value.

Extent

Although populations are critical to humanity, little is known about the extent of population diversity. Hughes *et al.* (1997) made a rough estimate of the total number of populations on the planet. They used a Mendelian population definition and restricted the estimate to eukaryotes (although the diversity of bacteria and viruses is probably enormous, information on their diversity is scarce).

The evaluation of global population diversity involved three steps. First, they estimated the average number of populations per unit area from literature on population differentiation. Second, they calculated the average range size of a species from published range maps. The product of these two numbers yields an approximation of the average number of populations per species. Finally, they multiplied this product by the number of species on Earth to arrive at an estimate of global population diversity.

Populations per Area

To quantify the number of populations per area for an average population, Hughes *et al.* (1997) searched 15 journals for genetic studies on population differentiation. They found 81 articles that provided appropriate data to quantify the number of populations per area. Most of the species were vertebrates (35), followed by plants (23), arthropods (19), mollusks (4), and a flatworm. For each species, they determined whether the sampling locations in the studies were for separate populations or a single population. If the genetic differentiation between localities was statistically significant, then all the localities were called separate populations. The number of populations per area was calculated as the number of sampling points divided by the area that was sampled. If the samples were not significantly differentiated, then it was assumed that all of the samples were taken from one population and that the size of a population was the size of the sampling area. For many species, an intermediate amount of differentiation was found. For example, although not all the samples were significantly different, some clusters of sampling locations were significantly different from others. In this case, the number of significantly diverged clusters was used as the number of populations.

Average Range Size

The average range size of a species was calculated from more than 2400 species range maps for birds, mammals, fish, and butterflies. Weighting each of the taxonomic groups equally, the mean range size of a species is 2.6 million km^2. The average range size of the most species-diverse group, the butterflies, is 2.2 million $km^2$and this was conservatively used as the estimate of the average range size of a species.

The shaded areas of species distribution maps almost always include habitats in which populations do not occur. Therefore, estimating species' ranges from these maps will probably inflate the population diversity estimate. In addition, the majority of species included occurred in temperate regions; however, it is estimated that two-thirds of species diversity exists in the tropics. Because species' range sizes in some taxa tend to increase toward the poles, the bias toward temperate areas may inflate the estimate of average range size. In contrast, the range maps that were used were restricted to one continent; therefore, the range sizes of intercontinental species were probably underestimated.

Species Diversity

The product of the average number of populations per area and the average range size of a species gives an average of 220 populations per species. Multiplying this number by three estimates of global species richness (5, 14, and 30 million) yields estimates of 1.1, 3.1, and 6.6 billion populations globally. This calculation assumes that a species' range size and its number of populations per area are independent. If these two factors are strongly correlated, however, the estimate may be quite inaccurate. However, it is not known whether a correlation exists, let alone whether the correlation is positive or negative.

Human-Induced Changes

With the expansion of the human enterprise, human beings are having a significant impact on population diversity. Populations are being both created and destroyed on a large scale as a direct result of human activities. As species are introduced to nonnative habitats, new populations are formed. At the same time, habitat disturbance and destruction are driving the extinction of populations. Although most public and scientific attention to biodiversity is centered on species, a few recent studies have attempted to evaluate the extent to which humans are modifying population diversity.

Introductions and Extinctions in Western Australia and California

Hobbs and Mooney (1998) collected data on the deletions and additions of populations in Western Australia and California. They gathered records of species that are threatened, extinct, or introduced in these areas. The number of regional extinctions gives a very conservative estimate of population extinction since most, if not all, of the species originally had more than one population in the region. Similarly, it is reasonable to assume that the threatened species have already lost populations. These species are endangered in large part because of range contractions – that is, the extinction of populations, thereby reducing the geographic distribution of the species. The number of introduced species is also a lower bound estimate of the number of new populations. To be recorded as an introduced species, at least one population must be established, but many species may have already established multiple populations.

For California, Hobbs and Mooney (1998) gathered data for butterflies, amphibians, reptiles, birds, mammals, and plants. Of 8274 total species, 1109 are introduced (13.4%), 71 endangered, and 49 extinct. The endangered and extinct categories include subspecies, which are most certainly genetically distinct populations. For Western Australia, the study reported information for plants, mammals, birds, reptiles, amphibians, and freshwater fishes. Of more than 12,000 plants, 1032 have been introduced, 232 are threatened, and 52 have gone extinct. The percentage of all introduced species is 7.3%, which is lower than that in California. In California, no extinctions have been reported for reptiles and amphibians, and in Western Australia no extinctions have been recorded for birds, reptiles, amphibians, or freshwater fishes. As mentioned previously, however, these results mask most of the population extinctions. In the groups for which no extinctions were

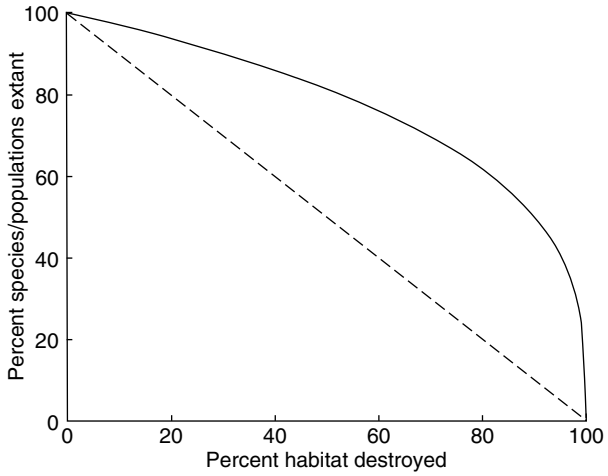


Figure 1 Predicted relationships between the area of habitat destroyed and the extant number of species and populations. The species curve (solid line) is $S=cA^z$ (see text), where z is 0.30. The population curve (dashed line) is linear.

recorded, other studies report significant species' range contractions. For instance, the loss of native vegetation in Western Australia has resulted in range contractions and population extinctions of species that rely on this habitat.

Extrapolating Population Extinctions from Tropical Deforestation Rates

Most estimates of species extinction rates are based on species–area relationships and the rate of habitat loss due to deforestation. The most commonly used model of the species–area relationship is shown in **Figure 1**. This model is $S=cA^z$, where S is the number of species, A is the area (or habitat size) in which the species are found, and c and z are constants estimated from empirical studies. This model is in agreement with a multitude of empirical studies on a variety of taxa and scales. With this formula, the rate of species extinction in an area can be approximated from the rate of habitat loss. **Figure 1** also illustrates a widely cited rule of thumb: A decrease in habitat area by 90% should result in a decrease in species diversity by approximately 50%.

Population–Area Relationship

There is no comparable work describing a population–area relationship, but most studies assume that population numbers and area probably correspond approximately in a one-to-one manner. The basis for the difference between the species– and population–area relationships is their size; a population occupies a small area relative to a species. When a substantial amount of habitat is destroyed, populations that were encompassed by the area will be lost. In contrast, many fewer species will go extinct because other populations of the species exist elsewhere. Thus, the rate of population extinction will be faster than that of species extinction. If the one-to-one relationship does hold, then when 90% of an area is destroyed about 90% of the populations in the original area will be lost, in contrast to only 50% of all species (**Figure 1**).

Extinction Rate

The rate of population extinction in tropical forest regions can be estimated with the population–area relationship in **Figure 1**. Assuming that tropical deforestation is occurring at 0.8% per year, global population diversity is 3 billion, and two-thirds of all populations exist in tropical forest regions, Hughes *et al.* (1997) calculate that 16 million populations per year, or approximately 1800 per hour, are being exterminated in tropical forests alone. This rate is three times higher than conservative estimates of species extinction.

These high extinction rates are supported by more current studies that use different approaches. Ceballos and Ehrlich (2002) investigated mammalian population extinction using historic (nineteenth century) and present-day range distribution of 173 declining mammal species from six continents. They found that together these species have lost more than 50% of their range area. Assuming that loss of range area is a result of population extinctions, these losses are much higher than species extinctions.

Loss of Genetic Diversity within Populations

These above studies focus on the extirpation of populations, rather than the total amount of genetic diversity of the remaining populations. Further work suggests that even populations that do remain may contain less genetic diversity. For instance, Garner *et al.* (2005) compared heterozygosity of healthy versus demographically challenged mammal populations and found that species that had experienced a demographic threat had significantly reduced mean heterozygosity (26% in marsupials and 22% in placentals).

Ameliorating Changes in Population Diversity

There are a variety of ways in which human modification of population diversity can be slowed and its impacts alleviated.

Controlling Invasions

In general, the impact of newly introduced populations on an ecosystem is highly unpredictable. In fact, most invasions do not produce noticeable effects. The invasive species that do have significant impacts, however, can cause widespread ecosystem disruption and economic loss. In theory, removing invasive populations can restore the native ecosystem (unless species have gone extinct). In practice, however, complete removal or even partial control of invasives is rarely successful. One of the few examples of a successful program to control invasive populations is the *Opuntia* cactus case described previously (*see* Importance).

There are three broad methods to control invasive populations: biological, chemical, and manual. One reason that introduced species can be very harmful to an ecosystem is that the species are free from their natural enemies and therefore overabundant. Biological control involves reducing the invasive species' numbers by introducing a predator or parasite from its native habitat. Chemical controls are poisons (usually synthetic) such as pesticides and herbicides. Manual control entails the physical removal or killing of individuals. All three of these fixes are usually expensive and in the case of

biological and chemical controls risk doing more harm than good to the ecosystem. An introduced predator might attack native species, and chemicals may harm native species or human beings. Thus, the best control is to prevent invasive populations from establishing in the first place. Unfortunately, this is an extremely difficult challenge – most introductions are accidental. Furthermore, with the accelerating movement of products and people around the world, the rate of introductions will certainly increase in the future.

Preservation

Protecting intact habitat from destruction or degradation is the most straightforward method to prevent the extinction of populations. If population diversity and habitat area are indeed related in a linear manner, then the more land that is preserved, the more population diversity is protected. In most cases, however, limiting almost all human access to a piece of land is usually not feasible because of competing uses, such as agriculture and urbanization. Nevertheless, these competing uses do not eliminate all prospects of conserving population diversity and the benefits they provide. In fact, in some parts of the world, many populations appear to be surviving in the countryside – that is, in human-dominated areas (e.g., Sekercioglu *et al*., 2007). Thus, a promising approach involves managing the countryside to prevent further degradation (and perhaps some restoration) of population diversity in these areas.

Restoration

Unlike species, populations can be restored and, along with them, at least some of the benefits that they once provided. Although genetic variation unique to the extinct populations will be lost, groups of individuals can be reintroduced to an area and in a relatively short time may evolve significant genetic differences from other populations. Reestablishing populations of a native species is not a simple task, however. Often, the native ecosystem is degraded and little is known about the biological requirements of the species.

See also: Diversity, Community/Regional Level. Ecosystem Services. Human Impact on Biodiversity, Overview. Modern Examples of Extinctions. The Value of Biodiversity

References

Ceballos G and Ehrlich PR (2002) Mammal population losses and the extinction crisis. *Science* 296: 904-9-7.

Garner A, Rachlow JL, and Hicks JF (2005) Patterns of genetic diversity and its loss in mammalian populations. *Conservation Biology* 19: 1215–1221.

Hobbs RJ and Mooney HA (1998) Broadening the extinction debate: Population deletions and additions in California and Western Australia. *Conservation Biology* 12: 271–283.

Hughes JB, Daily GC, and Ehrlich PR (1997) Population diversity: Its extent and extinction. *Science* 278: 689–692.

Schindler DE, Hilborn R, Chasco B, Boatright CP, Quinn TP, Rogers LA, and Webster MS (2010) Population diversity and the portfolio effect in an exploited species. *Nature* 465: 609–612.

Sekercioglu CH, Loarie SR, Brenes FO, Ehrlich PR, and Daily GC (2007) Persistence of forest birds in the Costa Rican agricultural countryside. *Conservation Biology* 21: 482–494.

Treseder KK and Vitousek PM (2001) Potential ecosystem-level effects of genetic variation among populations of Metrosideros polymorpha from a soil fertility gradient in Hawaii. *Oecologia* 126: 266–275.

Population Dynamics

Alan Hastings, University of California, Davis, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allee effect Reduction in growth rate of a population at low densities.
Competition The effect of the members of one species reducing the growth rate of the population of another species through interference or exploitation of common resources.
Density dependence Effects on the growth rate of a population due to interactions among the members of that population.
Endogenous forces Forces within a population that affect the dynamics of the population.
Exogenous forces Forces outside a population that affect the dynamics of the population.
Population Members of a single species that can potentially interact with each other.
Predation An interaction between two species in which one species consumes the other.

Introduction

The term *population* refers to the members of a single species that can interact with each other. Thus, the fish in a lake, or the moose on an island, are clear examples of a population. In other cases, such as trees in a forest, it may not be nearly so clear what a population is, but the concept of population is still very useful.

Population dynamics is essentially the study of the changes in the numbers through time of a single species. This is clearly a case where a quantitative description is essential, since the numbers of individuals in the population will be counted. One could begin by looking at a series of measurements of the numbers of particular species through time. However, it would still be necessary to decide which changes in numbers through time are significant, and how to determine what causes the changes in numbers. Thus, it is more sensible to begin with models that relate changes in population numbers through time to underlying assumptions. The models will provide indications of what features of changes in numbers are important and what measurements are critical to make, and they will help determine what the cause of changes in population levels might be.

To understand the dynamics of biological populations, the study starts with the simplest possibility and determines what the dynamics of the population would be in that case. Then, deviations in observed populations from the predictions of that simplest case would provide information about the kinds of forces shaping the dynamics of populations. Therefore, in describing the dynamics in this simplest case it is essential to be explicit and clear about the assumptions made. It would not be argued that the idealized population described here would ever be found, but that focusing on the idealized population would provide insight into real populations, just as the study of Newtonian mechanics provides understanding of more realistic situations in physics.

Density-Independent Dynamics of a Single Species without Age Structure

The very simplest description of population dynamics looks only at a single population of a single species. Thus, complications due to the effect of one species on the growth or dynamics of another are initially ignored. Any differences among individuals within the species are assumed insignificant. Thus, differences in age, sex, genetics, spatial location, and other individual characteristics are all ignored. Moreover, a single population is focused on, so any effects of immigration or emigration are ignored as well. Any effect of randomness is also ignored, which might arise either from the effects of small population sizes or environmental variability.

However, one more fundamental assumption has to be made when describing the dynamics of a population in order to consider what effect members of the population have on each other. The simplest case is density independence – in which the demography of the population does not depend on the size (in numbers) of the population. It is assumed that the per capita rates or numbers of births and of deaths do not depend on the size of the population. This is not the same as assuming that the overall birth and death rates do not depend on the population size, but it is equivalent to assuming that the overall birth and death rates are proportional to the population size.

To describe these dynamics under these assumptions in detail, the discussion needs to be split into two cases, depending on the way reproduction occurs. For some species, such as humans or many microorganisms, births and deaths occur essentially continuously through time. For other species, such as many butterflies or annual plants, organisms reproduce once each year and then die. These two possibilities are treated separately and the author then turns to more complex cases.

For the case of overlapping generations and continuous reproduction, the description of population dynamics can be written using the verbal equation

$$\text{Rate of change of population} = \text{Rate of births} - \text{Rate of deaths}$$

With the assumption of density independence, the rate of births can be written as bN, where b is the per capita birthrate and N is the population size. Similarly, the rate of deaths is mN, where m is the per capita death (mortality) rate. The verbal equation can then be written as the mathematical

equation

$$dN/dt = bN - mN$$

or, letting $b-m=r$, the intrinsic rate of increase (or "little r") as

$$dN/dt = rN \quad [1]$$

An explicit solution to this model, expressing the population size at any future time t in terms of population size at time 0 and the intrinsic rate of increase, r, can be determined. The solution of the model [1] is exponential growth, namely

$$N(t) = N(0)e^{rt} \quad [2]$$

where e is the base of the natural logarithms, $N(0)$ is the population at time 0, and $N(t)$ is the population at time t. The ecological implications of this formula will be focused on after the alternate case of growth in discrete time is discussed.

For the case of nonoverlapping generations, appropriate for annual plants, a verbal model can be written as follows:

Numbers next year = Per capita reproduction times numbers this year

In mathematical terms, this equation becomes

$$N(t+1) = RN(t) \quad [3]$$

where $N(t)$ is the population size in year t, and R is the number of individuals in the population next year produced by each individual in the population this year. This model, too, can be solved explicitly,

$$N(t) = R^t N(0) \quad [4]$$

which is geometric growth. In fact, if the growth rates in the two cases were matched so that $R=e^r$, or $\ln R=r$, the two population descriptions [2] and [4] would provide exactly the same numerical values (at the discrete times $t=1$, 2, 3, etc.).

The implications of solutions [2] and [4] are not that it would be expected that any real population grows at these rates, but, as emphasized perhaps first by Malthus several centuries ago, that exponential or geometric growth is very rapid. If r is positive or R is greater than 1, after relatively short times, the predictions of exponential or geometric growth are that populations will become very large. Thus, exponential growth cannot last, and the fundamental question of population dynamics becomes what prevents exponential growth and what are the consequences for population dynamics of these forces that prevent exponential growth.

Long time density-independent growth in natural systems obviously does not occur, with the closest examples following introduction of nonnative species. For example, the number of collared doves in Great Britain increased exponentially from less than 10 to more than 10^4 from 1955 to 1963, but the population essentially stopped growing shortly thereafter. Other introduced species have behaved similarly.

Density-Independent Age or Stage Structure

There are many ways to extend our understanding of the dynamics of a single population from the simplest cases of density-independent growth that have just been covered, and to explore what prevents exponential growth. There are two basic ways to extend our model for a single species (and remain within the context of looking at a single species). The effects of members of the population on each other and how they affect the dynamics could be looked upon. Also, the effect of differences among members of the population on the dynamics of the population could be previewed. The effect of differences among individuals, focusing on differences due to age, will be reviewed first. Other differences are also important and age is also focused on because it is a critical factor, has been well studied, and is easy to understand and measure. Does including the effect of age prevent exponential growth?

It is clear that within many populations some individuals are too young to reproduce, whereas others are too old. Thus, just looking at the numbers of individuals in the population is not enough to predict the future numbers. When looking at the effects of age structure, the proportion of females and males affecting the growth of the population is also an issue. Since the role of density dependence will still be ignored, the approach of only looking at the female members of the population will be considered, and therefore only looking at births of females to females. Thus it is implicitly assumed that males do not affect the dynamics of the population, which may be valid for some natural populations, such as moose, but not all, such as birds with parental care.

The information the author wants to include is the probability of births (or rate of births in continuous time) to a female of a given age, and the probability (or rate) of death at each age. From this information, and from the current distribution of females of different ages in the population, it should be possible to predict the future population. Once again, there are different formulations for this model, depending on whether the model is phrased in continuous or discrete time.

In discrete time, one way to write this model is known as the Leslie matrix model. A census interval, for example 5 years in humans, is first decided on. The population is described by a vector (which is just a vertical list of numbers) in which each entry is the number of females in each age class, 0–5, 6–10, 11–15, 16–20, and so on. The length of the age classes and the census times have to be the same. Then from the data, the average number of births to, and probability of survival of, a female in each age class is determined, over a 5-year interval. These data are presented in a life table, which summarizes the relevant demographic information. Life tables have been constructed from numerous species, ranging from insects, to small mammals, to long lived reptiles, to plants, and to humans. The data in a life table can be used as the entries in a matrix, known as the Leslie matrix, in which the first row are the birthrates, and the entries under the diagonal are the survival probabilities.

Analysis of this model reveals information about both the rate of growth of the population and the distribution of individuals of different ages. The key concept is the stable age distribution – the unique proportions of individuals of different ages that are preserved through time in the model. Thus, if the population starts at the stable age distribution, it will remain in the stable age distribution. Also, for a population in the stable age distribution, the proportional change in the numbers in each age class, and therefore the population as a

whole, is the same at each time step. This is simply geometric growth once again.

What happens if the population is not in the stable age distribution? It is clear that then the growth rate may change through time and will depend on the ages of the individuals – for example, a human population of all 0–5-year-olds would have no births, whereas one composed only of 20–25-year-olds could have many births. Yet, under some simple assumptions (essentially that reproduction is not confined to a single age class), if the population does not start in the stable age distribution, it will approach the stable age distribution through time. The result is that geometric or exponential growth occurs again, but only in the limit.

The analogous model in continuous time provides similar results, though there is a difference in the conceptual development. In this case, which can be traced back several centuries to the work of the Swiss mathematician Euler, the focus is typically on the rate of births through time rather than the individuals of all ages at a single time, through the use of a renewal equation. In words, the dynamics are described as

Rate of births at time t = Sum over all ages j of the birthrate at time $t-j$ × the probability of surviving to age j × the birthrate at age j.

Using the assumption of exponential growth this equation can be reduced to an equation for the growth rate of the population, the Euler–Lotka equation. This equation (or its discrete time approximation) is the basis for most estimates of the growth rate of populations.

The demography of populations and the ways in which population growth rates are calculated are worth emphasizing because they play a central role in conservation biology. A calculation of the growth rate of a population is often used to determine if the population is endangered. By looking at the effect of changes in different demographic parameters on the growth rate, one can determine which conservation efforts may have the largest effect on the growth rate and therefore the survival of the species in question.

Often, age is not a sufficient or appropriate descriptor of differences among individuals, and a classification based on stage (which could be weight, or some other characteristic) is more appropriate. Models incorporating stage can be more realistic, but care is needed to ensure that the dynamics are appropriately described. Recent advances have focused on using a continuous description of stage, particularly in discrete time using integro-difference equation models. This approach goes by the name of the integral projection model.

Density Dependence

Continuous Time

We now need to continue our search for processes that will prevent populations from growing exponentially and what the consequences of these limiting processes are for population dynamics. The simplest answer is that as the density of a species increases, either the birthrate goes down or the death rate goes up, because the members of the species compete with each other (due to the fact that the food supply or some other resource is limited). This kind of intraspecific competition is one of the fundamental reasons why populations may not grow exponentially, and its consequences will be explored, both for species with overlapping generations and those that reproduce once per year. In both of these cases and all that follows the age structure will be ignored, for simplicity.

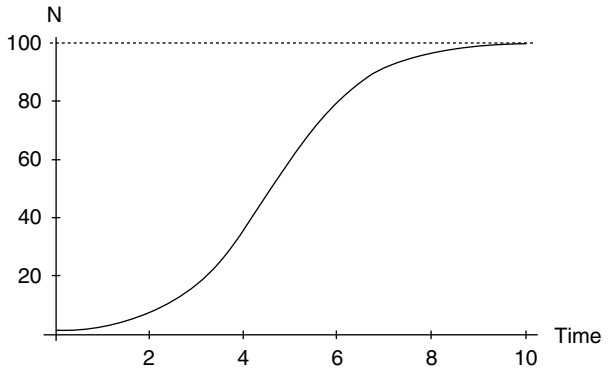


Figure 1 Explicit solution of the logistic equation with $r=1$, $K=100$, and $N(0)=1$.

One of the earliest formal descriptions of these dynamics was given by Verhulst, who produced what is now known as the logistic equation. In this equation, the growth rate of the population is described by the product of the exponential growth term as in model [1] and the term representing reduction in per capita reproduction due to interactions with other individuals. In modern terms, the model is written as follows:

$$dN/dt = rN(1 - N/K)$$

where the term K is called the carrying capacity of the population.

The logistic equation can be solved explicitly (**Figure 1**), but the most important conclusion is that any population reaches the carrying capacity, K. This qualitative conclusion also holds for similar assumptions. The search for the mechanisms leading to density dependence, and even direct evidence for density dependence in natural populations, is difficult.

Another aspect of density dependence is to reduce population growth rate at low population levels, perhaps through difficulties in finding a mate. This concept, the Allee effect, provides a lower limit to population size. If the population is too small, it may actually decline and thus can go extinct. The importance of Allee effects has been increasingly recognized, especially as a concept that underlies extinction.

Discrete Time

An alternate formulation based on the discrete time leads to very different conclusions. Here the idea is that the number of individuals next year will simply be a function of the number of individuals this year, with density dependence. The history of models like this began with examples from fisheries, such as the Ricker model:

$$N_{t+1} = N_t e^{r(1-N_t/K)}$$

The dynamics and behavior of this model are captured in the behavior of a much simpler model, the quadratic model, usually written as follows:

$$x_{t+1} = rx_t(1 - x_t)$$

where x_t is a scaled measure of population size and r is a measure of the rate of population growth. The density-dependent or carrying capacity parameter is absorbed in the scaling of x_t. Beginning with work of Robert May in the early 1970s, much of the interesting behavior of this model and its implications for population dynamics were elucidated.

The dynamics of density dependence in discrete time differ greatly from those in continuous time because of the presence of the implicit time delay imposed by the discrete time framework. As the population tends toward equilibrium, it can overshoot the equilibrium, with striking consequences. The overall dynamics of the quadratic model as a function of the parameters is as follows. If the growth rate is too low, $r<1$, the population will die out. If $1<r<2$, the population montonically approaches an equilibrium.

For larger values of the growth rate, the dynamics differ from the continuous time model. If $2<r<3$, the population approaches an equilibrium, but in an oscillatory fashion. For $r>3$, different dynamics emerge. At first the population exhibits stable cyclic behavior of alternating low and high years. For still larger values of r, the population has cycles of longer period, until above about $r=3.57$. For these large values of r, the population is typically chaotic, meaning that the population fluctuates and exhibits very sensitive dependence on initial conditions.

The search for this kind of behavior in natural populations has proven very difficult, but the attempts have greatly increased our knowledge of population dynamics. Laboratory populations including flour beetles have been shown to behave chaotically. Further study of this phenomenon is discussed later in the context of disease dynamics.

Role of Stochasticity

The discussion so far has assumed that no chance, or stochastic, factors play a role in the dynamics of the population. In many cases, this is far from the truth. For very small populations, demographic stochasticity, resulting from the chance that many births or deaths in a row might occur, has a large effect on the dynamics. In small populations, extinction may occur by chance even if the expected outcome would be for the population to grow.

A second important source of variability is the influence of the environment. Random fluctuations including year-to-year variability in weather, can cause large changes in the growth rate of a population.

A third kind of stochasticity is demographic heterogeneity, which is fixed differences in vital rates between individuals. This can be difficult to distinguish from environmental variability since both typically lead to a negative binomial description in the distribution of population sizes. However, demographic heterogeneity leads to greater variances at small population sizes, and therefore much greater extinction risk.

All of these stochastic effects need to be taken into account to understand the dynamics of natural populations. Historically, ecologists phrased this as a dichotomy, asking whether density-independent stochastic (e.g., weather) or density-dependent deterministic factors were responsible for controlling populations. The more modern version of this question is discussed later in the context of disease dynamics.

The Simple Two-Species Interactions

Overview and Definitions

Thus far, the focus is only on effects of the environment or within a single species, intraspecific interactions, on population dynamics. A fundamental question is, what is the role of interspecific interactions on population dynamics? The consequences of the interactions among many species fall within the realm of community ecology, but many of the simple effects of species interactions can be understood by taking a population dynamic approach, starting with the interactions between two species. This study will begin to highlight the important issue of which interspecific interactions act to prevent exponential growth of natural populations. The study of the interactions between two species has a long history in ecology, beginning with the seminal theoretical work of the American Lotka and the Italian Volterra, and the experimental work of the Russian microbiologist Gause.

The first step in a study of two-species interactions is to classify the interactions, primarily according to the effects each species involved has on the other. The dynamical effects of each interaction are then outlined. An interaction where one species eats another, such as wolves preying on moose, is called a predator–prey interaction. The predator gains a benefit and the prey is harmed. Interactions between herbivores and plants also are similar, so these interactions can more generally be called exploiter–victim or consumer-resource. The interactions in which one species is harmed and one gains will be treated separately because their dynamics are different or special. One of these is the interaction between disease-causing organisms and their hosts, and the other is the interaction involving insects (parasitoids) that lay their eggs inside the developing stage of other insects (hosts).

A second important two-species interaction is between two species that share a resource, such as two species that eat the same food items. This is a competitive interaction, where each species reduces the growth rate of the other. The resource, that is, the object of competition can also be nonbiological, such as light or water for plants or nutrients. Other competitive interactions might be more active and involve the production of chemicals, which would harm the competitor.

A third kind of two-species interaction is one in which both species benefit, as might occur between a pollinator and the species it pollinates. The pollinator gains, typically, a food resource, and the species being pollinated is aided in its reproduction. This is known as mutualism.

This classification does not exhaust possible two-species interactions, but these are the major interactions that have been the subject of most study. Many of the dynamical

consequences of other interactions can be deduced by considering the cases that have been outlined.

Predation

The discussion now turns to the dynamical effects of each interaction, beginning with the predator–prey interaction. The description of these dynamics began with the work of Lotka and Volterra, and this discussion starts with the simplest case they considered. It is assumed that in the absence of the predator, the prey population grows exponentially, in accordance with our examination of the single species case earlier. In the absence of the prey, it is assumed that the predator population declines exponentially. To complete the description of the dynamics, only the interaction between the predator and the prey and the consequences for the dynamics of both populations have to be described. The simplest case is to assume that encounters between predator and prey are random, so they occur at a rate proportional to the size of each population. It is also assumed that there is a fixed probability that the prey is killed in such an interaction, and that each prey individual that is killed contributes a fixed amount to the rate of growth of the predator population. These assumptions lead to the classic equations

$$dH/dt = rH - bHP$$

and

$$dP/dt = cbHP - kP$$

where H is the size of the prey population, P is the size of the predator population, r is the intrinsic rate of increase of the prey in the absence of the predator, k is the death rate of the predator population in the absence of the prey, the encounter rate is proportional to b, and c measures the conversion of prey deaths to increases in the predator population.

The analysis of the model shows that there is an equilibrium, but that the equilibrium is neutrally stable. Thus, all other solutions cycle. From this analysis, the important conclusion that the predator–prey interactions lead to cycles is depicted, but the model is unsatisfying because the solution depends completely on the initial conditions. The cycles essentially result from delayed feedbacks.

Modifying the model changes the conclusions from the unsatisfying one of neutrally stable cycles. If the equations are modified so that there is density dependence in the prey population, the system approaches equilibrium in an oscillatory fashion. Changing the description of the consumption of prey by predators can then produce stable cycles. This presence of cyclic behavior is the most important conclusion from the study of predator–prey interactions.

There are numerous examples of interactions between predator and prey in natural systems, which at least exhibit a tendency toward cyclic behavior, ranging from the classic example of hare and lynx in Canada to wolves and moose on Isle Royale. The observations in many of these cases are somewhat equivocal because it is often difficult to determine what is really causing the cycles, partly because it is difficult to isolate a single predator–prey pair within a larger community.

Host–Parasitoid Dynamics

By contrast, there is another kind of exploiter–victim relationship that does have more tightly coupled species interaction. The overwhelming majority of animal species are insects, and approximately one-fifth to one-sixth of all insect species are parasitoids: they have a life cycle where they lay their eggs in the developing stages of other insects, the hosts, and consume them from within, so the host is killed and one or more adult parasitoids emerge. Many insect pests of crops have important parasitoids. The interaction between a host and a parasitoid is essentially another version of the interaction between a prey and a predator, but the approach and the questions are somewhat different, so this case is treated separately. Here, the typical approach is discrete in time, which is appropriate for host insect species with single generations per year. The host species is assumed to grow exponentially (or geometrically) in the absence of the parasitoid. Unlike the predator–prey interaction, the parasitoid life cycle is tightly coordinated with the host because the parasitoid lays its eggs in the developing host. Thus the dynamics of the interaction are described by the probability that an individual host is parasitized, which could depend on the population density of the host and of the parasitoid.

The equations describing this interaction were first developed by Nicholson and Bailey, but these initial equations predicted that the populations would be unstable with cycles of ever-increasing amplitude. Thus, efforts since then have been focused on determining what would allow host and parasitoid to persist in a stable fashion. Suggestions like aggregation of host or parasitoid, or density dependence have been studied in models. Yet, as recent careful work by Murdoch and his colleagues on red scale and its parasitoids in California shows, the search for the stabilizing feature in natural systems is quite difficult. This is a system where a host and its parasitoids coexist stably, yet the search for the mechanisms that stabilize this interaction has been able to reject many of the conventional explanations.

Disease Dynamics

Another very important interaction that is similar to the predator–prey system is the interaction between diseases caused by bacteria or viruses and their hosts. However, in this case the focus is very different, since the number of disease organisms is not counted and instead only the number of diseased hosts is counted. These systems have become very important subjects of study in population biology partly because of very good records of the number of diseased individuals, in particular for human diseases.

The classic models in this area, which go back to the early work of Kermack and McKendrick in the 1920s, break up a population into three categories: susceptible, those individuals that can become infected; infectives, those individuals which are infected and infect others; and removed, those individuals which no longer can infect others. The simplest case is to look over relatively short times, so that overall population sizes are constant, and look at epidemics. The models, called SIR models, are framed in terms of differential equations where the key parameters are the contact rate, the recovery rate, and

the total population size. The rate of infection is assumed to be equal to the product of the contact rate, the number of susceptibles, and the number of recessives. The recovery rate for individuals is assumed to be just a constant. Then the dynamics depend on a single nondimensional parameter, the reproductive number, R_0, which is the number of new infectives produced by each infective, when the population is essentially composed of all susceptible individuals. If R_0 is less than one, the epidemic dies out, and if R_0 is greater than one, the epidemic will grow. This simple model fits many epidemics very well, as shown by the original work of Kermack and McKendrick looking at the 1905–1906 Bombay plague epidemic.

Disease dynamics have become central to some of the most important current questions in population ecology – namely, the relative importance of endogenous (within the system) versus exogenous (outside the system, often called stochastic) forces in determining the dynamics of population. This question, which is central to understanding the importance of any of the interactions outlined in this article, has been dealt with most clearly within the realm of childhood disease (chicken pox, measles, whooping cough, etc.) dynamics because the quality of the data is so good and because the models are more mechanistic.

The data, such as that for measles, have several important features, exhibiting both seasonal behavior driven in part by the school cycle and more complex multi-year cyclic behavior. There have been many different attempts to use extensions of the SIR model to understand the dynamics in a quantitative fashion. The approach taken is essentially to explicitly include the stochastic forces as an integral part of the model and to find the best fit (in a statistical sense) of the model to the data. This procedure illustrates the importance of using these parameter estimation approaches since the standard SIR model with constant infectivity does not provide a very good fit to the data, implying that the model is inadequate in some way. Further investigation shows that making the infectivity seasonally variable, corresponding to the role of school vacations in reducing the likelihood that children infect each other, greatly improves the fit of the model. Moreover, the rather complex dynamical patterns in the disease dynamics are shown to be generated by a combination of exogenous and endogenous forces: both the stochasticity and the nonlinearity (and seasonality) in the SIR model are critical for producing the observed dynamics.

Competition

The simplest models of competition are essentially phenomenologically based and start with the logistic model for a single species as their basis. The effect of each species on the other is to reduce the *per capita* growth rate. From this simple model of competition, several important ecological paradigms have emerged. The basic question to ask is what allows two species to coexist, rather than having a superior competitor eliminating an inferior one. The outcome of competitive interactions ends up being determined, in part, by the relative importance of interspecific effects as compared to intraspecific effects. Coexistence requires that the interspecific effects be weaker than the intraspecific effects.

In contrast to the predator–prey interaction, the dynamics of competition view is a monotonic approach to equilibrium without any cycles. However, if interspecific effects are strong enough, there can be a kind of priority effect. The species that wins in competition may depend on the initial densities in that case.

A more mechanistic approach to competition sheds some light on the issue of coexistence. With interference competition, in which the two species might, for example, affect each other by the production of allelopathic chemicals, interspecific effects may be stronger than the intraspecific ones. With exploitation competition, in which the two species use common resources, the relative strength of interspecific effects and intraspecific effects depends on the differences in resource use – for example, differences in the size of seeds used. From this view comes what was an extraordinarily important principle, the competitive exclusion principle of Gause, which stated that competing species have to differ in order to coexist.

The search for evidence of competition in natural populations is also difficult, but several large-scale surveys have demonstrated that there is relatively good evidence for competition. Classic work in this area includes careful studies of lizards in the Caribbean, which has shown competition for food and avoidance of competition by different species perching at different heights in trees. Competition for space has also been well documented in intertidal systems, in which removal of one species allows another to expand the range of habitat it uses. Recent work has also focused on the apparent competition, in which one species "competes" with another by allowing a common predator to increase its density, which then suppresses the competitor.

Conclusions

Population dynamics is one of the fundamental areas of ecology, forming both the basis for the study of more complex communities and of many applied questions. Understanding population dynamics is the key to understanding the relative importance of competition for resources and predation in structuring ecological communities, which is a central question in ecology.

Population dynamics plays a central role in many approaches to preserving biodiversity, which until now have been primarily focused on a single species approach. The calculation of the intrinsic growth rate of a species from a life table is often the central piece of conservation plans. Similarly, management of natural resources, such as fisheries, depends on population dynamics as a way to determine appropriate management actions.

Appendix

List of Courses

1. Ecology
2. Population Biology

See also: Diseases, Conservation and. Parasitoids. Population Density. Population Genetics. Predators, Ecological Role of. Species Interactions

References

Begon M, Mortimer M, and Thompson DJ (1996) *Population Ecology*, 3rd edn. Oxford: Blackwell Science.

Crawley MJ (ed.) (1992) *Natural Enemies*. Oxford: Blackwell Scientific Publications.

Hanski I and Gilpin ME (eds.) (1997) *Metapopulation Biology*. San Diego, CA: Academic Press.

Hastings A (1997) *Population Biology. Concepts and Models*. New York: Springer-Verlag.

Quinn III TJ and Deriso RB (1999) *Quantitative Fish Dynamics*. Oxford: Oxford University Press.

Population Genetics

Brian Charlesworth, University of Edinburgh, Edinburgh, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allele frequency The frequency of a variant form of a genetic locus within a population.
Genetic drift Evolutionary changes caused by random sampling of genotype frequencies in a finite population.
Genotype The state of an individual with respect to a defined genetic locus or set of loci.
Heritability The proportion of the variance in a trait that is due to additive genetic effects.
Inbreeding Matings between close relatives.
Mutation rate The frequency with which new mutations arise per generation.
Neutral mutations Mutations whose effects on fitness are either nonexistent or so small that their fate is controlled by genetic drift rather than by selection.
Phenotype The state of an individual with respect to a trait of interest.
Polymorphism The existence at intermediate frequencies of two or more variants at a locus within a population.
Selection The differential survival or reproductive success of individuals, associated with differences in phenotype or genotype.

Variation Within Populations

Types of Phenotypic Variation

Because evolutionary change depends on the existence of genetic variation within populations, measurement of the extent of such variation is crucial (Lewontin, 1974; Charlesworth and Charlesworth, 2010). Variation at the level of externally visible phenotypes can be divided into three categories.

Discrete Variation

This involves traits that can be divided into a small number of discrete categories, such as eye color in humans or shell color and pattern in the land snail Cepaea (Ford, 1975). It is often controlled by one or a few genes, and usually involves relatively superficial traits such as color patterns. Only a relatively small proportion of the phenotypic variation of interest to evolutionists is of this kind.

Quantitative Variation

Quantitative variation is all-pervasive. This can involve either meristic traits, such as bristle number in Drosophila, in which there is a large number of discrete categories, or continuously varying metrical traits such as body size. Variation in typical quantitative traits is known to be under the joint control of environmental effects, accidents of development, and sets of genetic variants whose individual effects are small relative to the total range of variation in the traits (Falconer and Mackay, 1996; Flint and Mackay, 2009). Statistical methods that utilize the degree of resemblance between close relatives enable the determination of the proportion of the total phenotypic variation that is contributed by additive genetic causes – the heritability, which controls the rate of response to the selection on a trait (*see* Selection on Quantitative Traits). Heritabilities for quantitative traits are typically between 20% and 80%, corresponding to the fact that artificial selection is highly effective in changing the mean value of almost every trait that has been examined (Falconer and Mackay, 1996).

Concealed Variability

A more subtle form of phenotypic variation is concealed variability, that is, variability that is only exposed when homozygous genotypes are produced by close inbreeding. This is responsible for the increased variation among inbred lines when a set of such lines is made from a random-bred base population and for inbreeding depression, which is the decline with inbreeding in the mean values of fitness-related traits such as viability and fecundity (Falconer and Mackay, 1996). Both of these phenomena reflect, at least in part, the widespread occurrence of recessive or partially recessive rare alleles in random-mating populations (*see* Mutation–Selection Balance), whose phenotypic effects are only fully exposed when they are made homozygous and are therefore not evident in randomly mating populations.

In Drosophila, special breeding methods, involving the use of genetically marked chromosomes with inversions that suppress crossing over, have shown that most haploid genomes carry at least one recessive lethal gene. Lethals contribute about half of the inbreeding depression in egg-to-adult survival that is manifested when fully homozygous genotypes are produced; a similar reduction in viability is caused by genes of small effect; the net fitness of fully inbred Drosophila is only a few percent of that of outbred flies (Crow, 1993). Even more extreme effects of complete inbreeding are likely for vertebrates, which have much larger genomes. The deleterious fitness effects of inbreeding have probably played a major role in promoting the evolution of mechanisms of inbreeding avoidance, such as the self-incompatibility genes of flowering plants (Charlesworth and Charlesworth, 2010).

Interpreting Phenotypic Variation

These basic facts were established by the early 1950s and led to an active debate concerning the causes of natural variation (Lewontin, 1974; Charlesworth and Charlesworth, 2010). From one point of view, championed by H. J. Muller, variation is mostly due to rare deleterious alleles maintained by mutation pressure at a large number of sites in the genome;

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00116-7

the coexistence of alleles at intermediate frequencies (polymorphism) is a rare event. From another point of view, advocated by Theodosius Dobzhansky, polymorphism is the norm, and reflects variation that is actively maintained by selection. In the absence of any means of identifying genes without the prior existence of genetic variability, no unbiased survey of the extent of genetic variation at individual genes was possible with the methods of classical genetics, and so this question could not be answered.

Molecular Variation

Protein Electrophoresis

This situation was transformed by the development of molecular genetics. Gel electrophoresis of soluble enzymes and proteins has provided a rapid and simple method for surveying populations for variants affecting the structure of a large number of different proteins and hence genes (Lewontin, 1974; Charlesworth and Charlesworth, 2010). The results of such surveys reveal that a large fraction of genes coding for soluble proteins are polymorphic, in the sense of having at least one rare variant whose frequency exceeds 5%; the average individual from a randomly mating population is typically heterozygous for a significant fraction (several percent) of such genes. Despite some biases in the methodology, particularly the inability of electrophoresis to detect several types of amino acid sequence changes and the restriction of the method to soluble proteins, it is clear that protein polymorphism is not an exceptional situation.

Measurement of Variation in Nucleotide Sequences

The introduction of sequencing technology has meant that population geneticists can now study variation at the level of the nucleotide sequencing itself. Surveys of within-species DNA sequence variation in nuclear genes of eukaryotes have been intensively carried out in Drosophila and humans, but comparable results are emerging from other species (Charlesworth and Charlesworth, 2010; 1000 Genomes Project Consortium, 2010). The basic conclusion is that changes from one base to another at a nucleotide site (single nucleotide polymorphisms (SNPs)) are the most abundant source of variation in natural populations. For nucleotide substitutions in coding sequences that do not change the amino acid sequence ("synonymous substitutions"), and for changes in introns and intergenic sequences, the frequency with which two randomly chosen allelic sequences from a typical Drosophila species differ at a given site ("nucleotide site diversity") is approximately 2%. The value is similar for the bacterium *Escherichia coli* and approximately 0.1% in humans (Charlesworth and Charlesworth, 2010). Literally millions of SNPs are present in human and Drosophila populations. For most genes, diversity is much lower for nonsynonymous changes, which alter the amino acid sequence. In addition to SNPs, DNA variability is contributed by insertions (including insertions of transposable elements) and deletions of sets of nucleotides; these mostly involve noncoding regions, and some can be many thousands of bases long (Charlesworth and Charlesworth, 2010). Other types of variability include variation in the sizes of microsatellite and minisatellite repeated sequences, which provide useful genetic markers (Goldstein and Schloetterer, 1999). The density of such repeats, however, is low in relation to the total size of the genome. The total level of variability at the level of DNA sequences is approximately two orders of magnitude greater than that revealed by electrophoresis, because of the high degree of synonymous and noncoding variability compared with nonsynonymous variability.

Interpreting Nucleotide Sequence Variation

Although variation at the level of nucleotide sequences must underly most heritable phenotypic variation, it has been difficult to relate the two, except for some intensively studied quantitative traits, especially human diseases (Flint and Mackay, 2009). The abundance of variation in both protein and DNA sequences might seem to vindicate Dobzhansky's view of the causes of natural variation. However, it is likely that a good deal of synonymous and noncoding variability is close to neutrality with respect to its effects on fitness (*see* Molecular Evolution and Variation). Nevertheless, there is now abundant evidence that selection also influences variation and evolution in both coding sequences and regulatory noncoding sequences; many methods for detecting selection from its effects on patterns of DNA sequence variation and evolution have been developed (Charlesworth and Charlesworth, 2010), but space does not permit them to be described in detail here. In addition, selection is the major force controlling phenotypic evolution. The effects of deterministic forces such as selection on variation and evolution within populations will thus be considered next.

Deterministic Population Genetics

Allele and Genotype Frequencies

If the focus is on a given nucleotide site, the basic descriptor of the state of a population is the set of frequencies of the four alternative states, A, T, G, and C (ignoring variants that involve insertions or deletions). If recombination within a gene is ignored, we can consider the set of nucleotide sequences observed for a defined region of the genome ("locus") as alternative alleles, whose frequencies characterize the state of the population. Mendelian inheritance implies that this state is not changed in the absence of evolutionary forces (the occurrence of intragenic crossing over and gene conversion at low frequencies means that this is a good approximation rather than an exact description). This is highlighted by the Hardy–Weinberg principle, which states that the frequencies of diploid genotypes in a random mating population at a locus with a set of n alleles with frequencies $p_1, p_2, \ldots, p_n$ rapidly reach equilibrium values given by the multinomial expansion of $(p_1 + p_2 + \ldots, p_n)^n$. The importance of this result is that it means that existing natural variation is preserved by Mendelian inheritance. This removes Darwin's difficulties over the rapid loss of variation under blending inheritance, which led him to adopt a theory of the inheritance of acquired characteristics for which there is no empirical foundation (Fisher, 1930).

Mutation

Types of Mutations and their Rates

The ultimate source of natural variation is spontaneous mutation, defined as a heritable change in the genetic material. Mutations arise without reference to the adaptive utility of the phenotypic consequences of the change in question. The most abundant mutations are single nucleotide substitutions, but small deletions and insertions due to slippage during DNA replication are relatively common as well (Drake *et al.*, 1998; Charlesworth and Charlesworth, 2010; Lynch, 2010). Insertions of transposable elements, large deletions and duplications, duplications and deletions of entire chromosomes or haploid genomes (aneuploidy and polyploidy), and chromosome rearrangements, such as inversions and translocations, also occur and contribute to the evolutionary changes in genome structure. Rates of spontaneous mutation in organisms with DNA genomes are extremely low, due to the operation of complex enzymatic systems which repair lesions in DNA; the rates of nucleotide changes per site per cell generation in DNA-based microbes are between 1×10^{-10} and 5×10^{-10}. Similar values apply to mutation rates per cell division in multicellular organisms, but the rate per organism generation is over 100 times higher because of the many cell divisions that occur during the production of the germ cells. This yields a mutation rate per nucleotide site per generation of the order of 3×10^{-9} in Drosophila and 10^{-8} in humans (Lynch, 2010; 1000 Genomes Project Consortium, 2010). Per locus rates of mutation to alleles with major phenotypic effects are substantially higher – of the order of 10^{-5} per generation – reflecting the large number of nucleotides in the coding and regulatory regions of typical genes. Rates of change in copy numbers in microsatellite and minisatellite loci are generally much higher, however: as much as 10^{-3} per locus per generation in mammals (Goldstein and Schloetterer, 1999). Mutation rates in viruses with RNA genomes are also extremely high, due to their lack of repair mechanisms (Drake *et al.*, 1998).

Evolution under Mutation Pressure

Given the low rate of mutation at the nucleotide level in DNA-based organisms, the time scale of mutational change in the frequencies of the four alternative states at a nucleotide site is very large – of the order of 100 million generations. Mutation pressure at this level is an extremely weak force, easily opposed by other evolutionary factors. For most purposes, therefore, mutation can be regarded simply as a source of new variation, and unimportant for directional evolutionary change (Fisher, 1930). This statement needs to be qualified when the aggregate effects of mutations affecting a particular phenotype are considered; the numbers of loci affecting a single quantitative trait are sufficiently large that increases in variability due to mutation are detectable in stocks that are initially genetically uniform. The rate of increase in variance per generation is typically of the order of 10^{-3} relative to the nongenetic variance for the trait (Falconer and Mackay, 1996). Given that fitness-related traits are affected by many genes, and that most phenotypic changes caused by mutations are harmful to the organism, the mean values of fitness components decline under mutation pressure when selection is relaxed by experimental manipulations (Crow, 1993; Drake *et al.*, 1998; Charlesworth and Charlesworth, 2010).

Mutation and Selection

These considerations imply that a relatively weak force of selection, far smaller than is measurable experimentally, can prevent the spread of deleterious mutations. This explains the fact that amino acid polymorphisms are usually much less frequent than silent or noncoding polymorphisms. The same observation applies to comparisons of sequences between different species (Kimura, 1983). A very important form of selection is thus purifying selection, whereby deleterious mutations are constantly being eliminated from the population (*see* Mutation–Selection Balance). Similarly, quantitative traits are often subject to stabilizing selection, such that individuals with extreme trait values are less fit than individuals with intermediate values (Falconer and Mackay, 1996). Variability in quantitative traits is maintained in part by a balance between the input of new variation by mutation and its elimination by stabilizing selection (Falconer and Mackay, 1996; Charlesworth and Charlesworth, 2010).

Selection at a Single Locus

The Basic Model

A large body of theory has been developed to describe the action of natural selection on Mendelian variation. In the simplest case, a single locus with a pair of alternative alleles, A and a, is considered. A randomly mating, infinitely large population is assumed. Ignoring sex differences, the relative values of the fitnesses of the three possible genotypes AA, Aa, and aa in a diploid can be written as 1, $1-hs$, and $1-s$, respectively, where s is the selection coefficient and h is the dominance coefficient. The parameter s measures the strength of selection ($s>0$, when A is favored by selection; $s<0$, when A is disfavored), and h measures the extent to which the fitness of the heterozygote is affected by the presence of a. In this section, the fitnesses are assumed to be constant over time.

Fitness in this context is most easily understood in terms of viability selection in a population with discrete generations, where the relative probabilities of the three genotypes surviving from egg to adult are equivalent to the three fitnesses. In general, selection may involve many different aspects of the life history, especially female fecundity and male mating success. More elaborate models have been developed to study these cases, including extensions of the theory to populations with overlapping rather than discrete generations, but the basic conclusions are similar (Ewens, 2004; Charlesworth and Charlesworth, 2010).

If the frequencies of the two alleles in one generation are p and q, respectively, the change in frequency of A over one generation is

$$\Delta p = spq\{q + h(1-2q)\}/\overline{w} \qquad [1a]$$

where $\overline{w} = 1 - 2\,pqhs - q^2 s$ is the population's mean fitness.

An interesting equivalent form, given by Wright, is

$$\Delta p = \frac{1}{2} pq \frac{d \ln \overline{w}}{dp} \quad [1b]$$

Equation [lb] implies that allele frequency change is in the direction of the gradient of mean fitness with respect to allele frequency. A more detailed analysis of the dynamics of a single locus with an arbitrary number of alleles, and arbitrary but constant fitnesses, shows that mean fitness increases monotonically as allele frequencies change, so that stable equilibria in allele frequency space correspond to the local maxima in mean fitness. The graph of mean fitness against allele frequency is known as an "adaptive landscape" (Wright, 1977; Charlesworth and Charlesworth, 2010).

Directional Selection

For the case of a pair of alleles, when $0 \leq h \leq 1$ (so that the fitness of the heterozygote is bounded by the fitnesses of the homozygotes), A is favored by selection when $s > 0$ and will progress to fixation. This case is referred to as directional selection. If A is initially rare so that second-order terms in p can be neglected. Equation [1a] can be approximated by $\Delta p \approx s(1-h)p$. The initial rate of change of the natural logarithm of allele frequency ($\approx \Delta p/p$) is therefore proportional to the fitness difference between the genotypes Aa and aa. This reflects the fact that, with random mating, rare alleles are overwhelmingly carried in heterozygotes (frequency $2pq$) since the frequency of homozygotes (p^2) is negligible. If A is completely recessive, so that $h = 1$, eqn [1a] for rare A is approximated by $\Delta p \approx sp^2$, implying that the logarithmic rate of increase in frequency of A is proportional to sp, which tends to zero with decreasing p. This reflects the extreme rarity of the favored homozygotes, and implies that rare recessive alleles are only weakly selected in randomly mating populations (Haldane, 1932; Charlesworth and Charlesworth, 2010).

Survival of Favorable Mutations

The previous conclusion is reinforced by calculations of the probabilities of survival of new mutations. Even in very large populations, a new mutation is likely to be represented initially in only one or a few individuals. Since reproduction is subject to random variation, described by the probability distribution of the numbers of surviving offspring per mated individual, there is a finite chance in each generation that all carriers of a rare mutant gene will fail to transmit it (Fisher, 1930; Haldane, 1932). The chance that a single copy of a new favorable mutation ultimately survives random loss from a large population of constant size is approximately $2s\ (1-h)$, assuming a Poisson distribution of offspring number (Haldane, 1932; Charlesworth and Charlesworth, 2010). Most favorable mutations are thus likely to be lost from the population of a few generations after they arise; on average, 148 occurrences of a mutation with a heterozygous selective advantage of 1% are needed for a 95% chance that one will survive.

This implies that there is a considerable random element to adaptive evolution at the genetic level. If there are several different loci at which mutations that provide adaptations to a given pressure of selection can occur, it may be a matter of chance which locus actually responds to the selection in a given population. The same pressure of selection can therefore result in the divergence of isolated populations at the level of the genotype, even if the same phenotype is evolving. The numerous types of genetic variant involved in the adaptation of different human populations to malarial parasites illustrate this principle (Kwiatkowski, 2005; Charlesworth and Charlesworth, 2010).

The formula for survival probability implies that a recessive favorable mutation has a zero chance of survival in an infinite population; calculations based on diffusion theory (*see* Fixation Probabilities) show that the probability for a randomly mating population of size N in this case is approximately $0.8\sqrt{s/N}$ (Charlesworth and Charlesworth, 2010). This means that recessive autosomal mutations are unlikely to become established by selection in randomly mating populations of even moderate size. As first pointed out by Haldane, it is not an accident that nonrecessive alleles have been established by selection in cases of recent adaptation involving genes of major effect, such as industrial melanism and pesticide resistance, despite the fact that most spontaneous mutations are recessive with respect to their effects on the phenotype (Ford, 1975; Charlesworth and Charlesworth, 2010).

Time Course of Gene Frequency Change

Once a favorable allele rises to a sufficiently high frequency that random loss is unlikely, its progress in a large population can be calculated by integration of the differential equation which approximates eqn [1a] when selection is weak. This procedure yields expressions for the time needed to change allele frequencies by a given amount (Haldane, 1932; Charlesworth and Charlesworth, 2010). These can be used to estimate selection intensities in experimental and natural populations, by comparing observed trajectories of gene frequency change with the theoretical predictions. This method is particularly useful for microbes, which have short generation times and can be grown in very large artificial populations. Selection coefficients of the order of 0.1% can be measured in such microbial experiments (Charlesworth and Charlesworth, 2010).

The results of these calculations show that the time required to cause a given amount of change in gene frequency is inversely proportional to s and depends only logarithmically on the initial frequency of the favorable allele, except when it is recessive. This implies that selection on nonrecessive alleles can cause evolutionary change on a timescale of the order of a few multiples of $1/s$, almost regardless of the initial conditions. Even a very weak force of selection can thus transform a population in a period of time that is negligible in geological terms. This conclusion is historically very important, because it removes many of the objections to natural selection as a potent force in evolution (Fisher, 1930; Haldane, 1932).

Selection on Quantitative Traits

Predicting the Response to Selection

This question can also be considered in relation to the quantitative traits, which are more relevant to phenotypic evolution than the traits controlled by single genes. The standard model of quantitative genetics (which involves

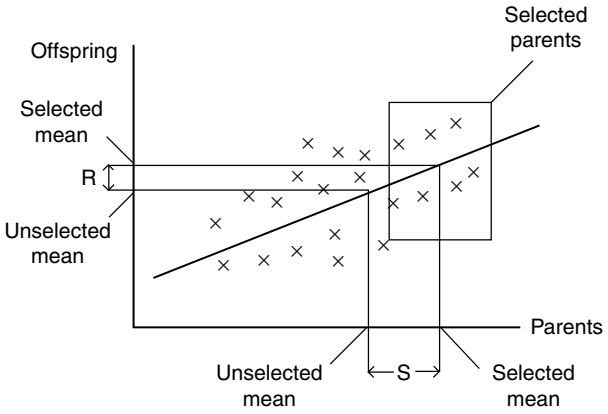


Figure 1 Each cross indicates the location of a point determined by the mean values of the phenotypic values of a pair of potential parents (x axis) and of their progeny (y axis). The straight line is the regression of y on x, whose slope is equal to the heritability. Families that are successful in reproducing are enclosed in the rectangular box.

several simplifying assumptions) implies that the response to selection on a single quantitative trait is governed by the equation

$$R = h^2 S \qquad [2]$$

where the selection response, R, is the difference in the mean value between one generation and the next; h^2 is the heritability of the trait (*see* Quantitative Variation) (there is no connection with the dominance coefficient introduced previously); S is the selection differential, which measures the intensity of directional selection applied to the trait, and is equal to the difference in mean between the parental individuals who have survived selection and the population mean before selection (**Figure 1**). Given the fact that h^2 is usually substantially greater than zero, this implies that populations can respond rapidly to directional selection on quantitative traits, as is indeed usually found to be the case (Falconer and Mackay, 1996).

Rates of Change Under Selection

In a sexually reproducing population, the ability of genetic recombination to put together new combinations of alleles means that selection can eventually produce genotypes with trait values that far exceed the range of variation that existed in the original population, simply by causing the fixation of alleles that increase trait value and were segregating in the original population. In the absence of new variability created by mutation, the selection response will eventually stop when all favorable alleles have been fixed by selection. Given the evidence for a significant input of new variation by mutation for typical quantitative traits, a relatively gentle pressure of selection in a large population will not fully deplete variation, so that sustained responses to selection are possible (Falconer and Mackay, 1996). The breeder can produce substantial changes in a single trait over less than 100 generations, comparable in magnitude to geologically rapid changes that have taken thousands of generations in nature. It is therefore unlikely that lack of genetic variation is often a serious constraint on the ability of populations to undergo phenotypic evolution, unless there is intense selection for a novel combination of traits, which cannot immediately be produced by recombination among existing genotypes.

Selection on Multiple Characters

In general, natural selection acts on suites of characters and not on a single trait in isolation. A multivariate generalization of eqn [2] has been derived (Lande, 2000; Charlesworth and Charlesworth, 2010), which is useful for the interpretation of data on selection on multiple traits:

$$\Delta\bar{z} = G\nabla \ln \bar{w} \qquad [3]$$

where $\Delta\bar{z}$ is the change per generation in the vector of mean values for the set of traits, $\nabla \ln \bar{w}$ is the gradient vector of log mean fitness with respect to the set of trait means, and G is the matrix of additive genetic variances and covariances among the set of traits. The similarity to eqn [1b] is evident.

Equation [3] shows that, in general, we can only predict the effect of selection on a trait if we know the extent to which it is genetically correlated with other traits that are also the target of selection. Short-term evolutionary changes in a given trait may be due, at least in part, simply to the fact that it covaries with another trait that is the target of selection. However, provided that the G matrix is nonsingular and fitnesses are constant in time, a stable equilibrium state corresponds to a maximum in the surface of mean fitness with respect to the mean trait values, indicating that selection will ultimately bring the population close to the optima with respect to each individual trait, unless there are strong constraints on the correlational structure of genetic variation. This implies that the concept of an adaptive landscape can be applied to quantitative traits (Lande, 2000).

Maintenance of Variation by Selection

The role of selection in preserving variation rather than destroying it, called balancing selection, is now discussed. This form of selection was unknown to Darwin, and its discovery is one of the most important contributions of population genetics.

Heterozygote Advantage

The simplest case is when the heterozygote at a locus with two alleles has a higher fitness than the two homozygotes, which was first investigated by Fisher in 1922. Let the fitnesses of AA and aa relative to Aa be written as $1–s$ and $1–t$; eqn [1b] now yields the result that there is a unique stable equilibrium at which the frequency of A is $p^* = t/(s+t)$, to which the population converges from any starting point other than fixation for A or a (Ewens, 2004; Charlesworth and Charlesworth, 2010). The classic example of heterozygote advantage is the polymorphism for the β-globin variant in humans that causes sickle-cell anemia when homozygous, a disease which is nearly lethal under natural conditions. The maintenance of this allele in human populations subject to severe malarial infections is due to the selective advantage of heterozygous carriers, conferred by their resistance to malaria. Several other human globin gene polymorphisms, as well as many other

human polymorphisms, are maintained by resistance to malaria (Kwiatkowski, 2005).

Frequency-Dependent Selection

Balancing selection can also be caused by negative frequency-dependent selection, in which the relative fitness of a genotype decreases with its frequency in the population. This obviously acts to inhibit the fixation of an allele that may initially have a selective advantage over other alleles when introduced into a population. Genetically controlled resistance to parasitic disease is subject to this form of selection, since the abundance of a parasite decreases as the number of susceptible hosts diminishes, thereby reducing the selective advantage to a host allele that causes resistance to a particular parasite genotype (Charlesworth and Charlesworth, 2010). A similar frequency dependence affects alleles controlling virulence in the parasite population. The genes thought to be involved in disease resistance, such as the major histocompatibility complex (MHC) loci of vertebrates, are often highly polymorphic and segregate for large numbers of alleles; molecular analyses show clear evidence for balancing selection on these loci (*see* Effects of Linkage to Balanced Polymorphisms and Targets of Local Selection). The role of frequency-dependent selection in this case remains to be established, however. Frequency dependence is also inherent in the dynamics of the self-incompatibility loci of flowering plants, which have very high allele numbers (Charlesworth and Charlesworth, 2010). Similarly, Batesian mimicry, in which an edible species mimics a distasteful model, exhibits frequency dependence since a predator is more likely to mistake a rare mimetic form for the model than a common one (Fisher 1930; Ford, 1975; Charlesworth and Charlesworth, 2010). It should be noted that frequency-dependent selection does not necessarily lead to polymorphism; its outcome depends on a delicate balance of the selective parameters, and dynamic complexities such as limit cycles or even chaos are possible. With frequency-dependent selection, maxima in mean fitness do not necessarily correspond to stable equilibria.

Temporal Variation in Fitnesses

Temporal fluctuations in relative fitnesses can also lead to the maintenance of polymorphisms by selection; in the case of two alleles at a locus with random mating, a sufficient condition is that the geometric mean fitness of the heterozygote over generations exceeds that of both homozygotes (Charlesworth and Charlesworth, 2010). Similarly, spatial variation in the direction of selection can maintain variation, even in the absence of heterozygote advantage. This can happen in two ways. The first requires strong density-dependent regulation within different environmental patches, so that the number of adults emerging from a patch is largely independent of the genetic composition of the eggs laid in that patch. In this case, opposing directions of selection in different patches can maintain polymorphism even if there is completely random mating among the patches (Charlesworth and Charlesworth, 2010). Directional selection in opposite directions on alleles in males and females (sexual antagonism) is a special case of this. Second, restricted migration among populations subject to different directions of selection can result in polymorphism within each population, accompanied by genetic differentiation among populations. Clinal variation in allele frequencies or quantitative trait values results from geographic gradients in selection pressures, coupled with the smoothing effect of migration, as in the case of Bergmann's rule, which states that the mean body sizes of mammal populations increase with higher latitudes (Charlesworth and Charlesworth, 2010).

Meiotic Drive

Antagonism between the effects of selection at different levels may also maintain variation. The best studied cases involve the phenomenon of meiotic drive or segregation distortion, in which one allele (D) at a locus when heterozygous causes the destruction of gametes carrying the alternative allele (d). In animals, this is usually found to occur only in males. Provided that the fertility of Dd males is affected less than linearly by the destruction of the d sperm, D will gain a transmission advantage. It will spread through the population unless there is a countervailing selective disadvantage at the level of individuals. In the best studied cases of this kind, the SD system of *Drosophila melanogaster* and the t-haplotype system of house mice, the primary disadvantage seems to come from sterility of the DD males (Charlesworth and Charlesworth, 2010).

Mutation–Selection Balance

The balance between recurrent mutations to variants that impair the functions of gene products and selection against them is probably a major factor in maintaining genetic variation, given the fact that higher organisms have tens of thousands of genes (Drake *et al.*, 1998; Lynch, 2010). The large number of changes in a coding sequence that can impair the function of a gene product implies that the process of mutation to deleterious alleles at a locus can be regarded as effectively irreversible, provided that the wild-type allele predominates in the population. If the deleterious alleles at a locus are completely recessive, with selection coefficient s and a rate of origination by mutation from wild type of u, their equilibrium frequency in a randomly mating population (assuming that $s \gg u$) is $q^* = \sqrt{u/s}$. Even lethal alleles ($s=1$) can thus reach appreciable frequencies if they are completely recessive; with $u=10^{-5}$, for example, $q^*=3\times10^{-3}$. However, experimental studies of lethal mutations in Drosophila show that they usually impair the viability of heterozygotes with wild type by as much as 2%; detrimental alleles with more minor effects (s on the order of a few percent or less) appear to be much less recessive, with mean h values of approximately 0.25 (Crow, 1993). With random mating, the much greater frequency of heterozygotes than mutant homozygotes means that selection on heterozygotes controls the frequencies of mutations; in this case, $q^*=u/(hs)$. The total rate of mutation to lethal mutations per haploid genome in Drosophila is approximately 0.01; assuming $hs=0.02$, the mean number of heterozygous lethals per diploid individual is of the order of 1, in agreement with the direct estimate mentioned previously. Evidence from estimates of selection on deleterious nonsynonymous mutations found segregating in populations suggests that the mean number of slightly deleterious mutations per individual in humans is of the order of 800, and 10

times higher in Drosophila (Charlesworth and Charlesworth, 2010).

Genetic Load

General Considerations

The previous discussion leads to the consideration of the effect of selection on the fitness of the population as a whole; if there is a large amount of variability with respect to loci under selection, it is obvious that the mean fitness of the population must be much less than that of the best genotype. If fitnesses are measured relative to the value for the optimal genotype in the system under consideration, the reduction in fitness below the maximum value of 1 can be conveniently measured by the genetic load, defined as $L = 1 - \overline{w}$. In the case of viability selection, no genotype can have a survival probability greater than 1, so that L provides an upper bound to the probability that a zygote survives to maturity. More generally, L measures the proportion of the population that dies or fails to reproduce as a result of selective differences among genotypes. The effects of multiple loci on mean fitness can be calculated by assuming that different loci act independently, so that the fitness of a multilocus genotype is given by the product of the fitnesses of all the single-locus genotypes which contribute to it. If L_i is the load contributed by the ith locus, the mean fitness with respect to m independent loci, relative to the value for the optimum genotype, is

$$\overline{w} = \prod_{i=1}^{m}(1 - L_i) \approx \exp - \sum_{i=1}^{m} L_i \qquad [4]$$

Mutational Load

With nonrecessive deleterious alleles maintained by mutation, the load for a single locus at equilibrium is approximated by $2q^*hs = 2u$. The total load is $1-\exp-U$, where U is the mean number of new deleterious mutations arising in each generation in a diploid individual. Lethal mutations contribute relatively little to this since their total mutation rate is very low, but detrimentals have a major effect: Assuming 20,000 genes with a deleterious nonsynonymous mutation rate of 10^{-5} per coding sequence, the mutational load for a mammalian population would be 0.33, a considerable burden of selective loss. An even larger load is probably caused by mutations that affect noncoding sequences of functional importance (Charlesworth and Charlesworth, 2010; Lynch, 2010).

The existence of a large mutational load suggests that there is an adaptive advantage to a reduction in the mutation rate; this can be studied theoretically by calculating the rate of spread of a rare modifier gene that reduces U by a small amount δU. If the modifier recombines freely with autosomal loci subject to mutation and selection, it has a selective advantage of $hs\ \delta U$. This raises the question of why mutation rates are not closer to zero; this probably reflects the fact that there is a fitness cost to the necessary repair systems so that U is adjusted to a level at which the costs and benefits of increased repair balance each other (Drake *et al.*, 1998).

Segregational Load

Similar calculations can be performed for models of balancing selection, yielding estimates of the segregational load. In the case of heterozygote advantage, the load due to a single locus is $st/(s+t)$ (Charlesworth and Charlesworth, 2010). Equation [4] can be used to determine the segregational load contributed by a large number of polymorphic loci with independent effects. This can be considerable, even if selection is weak. For example, 10,000 loci each with $s = t = 0.001$ would yield a mean fitness of only 0.0067. This is so low that only a very high fecundity species would be able to produce the two surviving offspring per adult needed to maintain itself. This implies that either most molecular variation has very slight or no effects on fitness, or that the assumption of multiplicative fitnesses is unrealistic.

An extreme alternative to multiplicative fitnesses is truncation selection. Genotypes at a set of loci are assumed to be ordered with respect to their fitnesses as determined by the multiplicative fitness model; a fixed proportion of the population, containing the set of genotypes with the highest fitnesses, is allowed to survive. This is equivalent to assuming that individuals compete for a limiting resource, and that only the fittest succeed. Under these conditions, a much larger number of loci can be exposed to selection for a given total L than with multiplicative fitnesses, for the same selection intensity per locus. Less extreme forms of departure from multiplicativity can have similar but smaller effects on the total load.

Substitutional Load

Genetic loads also apply to adaptive evolution. Consider the case of a biallelic locus, where A is initially deleterious and held at a low frequency. If there is a change in the environment so that A becomes favored by selection, it will start to increase. However, in any generation before it reaches fixation, the mean fitness of the population will be reduced below its final value of 1 because of the presence of the unfavored allele. There is thus a load associated with the substitution of a by A, reflecting the fact that natural selection cannot instantly transform a population that contains allele a into one which is fixed for A. The sum of the loads for each generation over the course of a gene substitution is Haldane's "cost of selection," C; if the population size is N, the total number of individuals eliminated by selection is CN. Provided that selection is not too strong, C is proportional to minus the logarithm of the initial frequency of A; a typical value is 30 (Charlesworth and Charlesworth, 2010).

The effect of changes at multiple loci can be derived as follows: Assume that there is a steady rate of change in the environment so that in each generation K loci start to experience gene substitutions of this kind. K is the rate of gene substitution in the genome as a whole such that after a long period of evolutionary time, T, the population will differ from its ancestral state by KT substitutions. If a gene substitution takes t generations to complete, Kt loci will be segregating in any given generation, each associated with an average load of C/t relative to a population that is fixed for the favorable alleles. The mean fitness under multiplicative fitness, relative to a population that is fixed for favorable alleles at all currently segregating loci, is then given by $(1 - C/t)^{Kt} \approx \exp - CK$.

Data on rates of protein sequence evolution show that K for an average amino acid site in a protein is approximately $1.5 \times 10^{-9}\,\mathrm{yr}^{-1}$ in mammals (Kimura, 1983). With 20,000 loci coding for proteins with an average size of 500 amino acids, K for the genome is 0.015. With $C=30$ and assuming one generation per year, the mean fitness is 0.64, implying a load of 0.36. A much higher load would be found if changes in functional noncoding sequences are also taken into account. This finding of a high substitutional load associated with molecular evolution was one of the main motivations for the development of the neutral theory of molecular evolution, which asserts that most evolution at the molecular level is caused by the random sampling of alleles in finite population size, genetic drift, and not by natural selection (Kimura, 1983). The substitutional load can be considerably reduced by modifications to the assumption of multiplicative fitnesses (*see* this section, above), so that this argument has lost much of its cogency (Ewens, 2004; Charlesworth and Charlesworth, 2010). Nevertheless, there are other reasons to take the neutral theory seriously (*see* Molecular Evolution and Variation).

Multiple Loci

No Selection

So far, it has tacitly been assumed that evolutionary change involving more than one locus can be modeled by assuming that alleles at different loci are distributed independently of each other in the population and have independent effects on the phenotypes and fitness. Although this may be a good approximation for many purposes, it is necessary to examine the consequences of relaxing these assumptions. Deviations from independence among loci in randomly mating populations can be described by linkage disequilibrium parameters, which measure the extent to which the frequencies of the different multilocus gamete types or "haplotypes" depart from the frequencies expected by randomly combining alleles at different loci. In the simplest case of a pair of loci, each with two alleles (A and a; B and b), there are four haplotypes: AB, Ab, aB, and ab. Let the frequencies of these be x_1, x_2, x_3, and x_4. If the allele frequencies at the two loci are p_A and p_B, we can write the haplotype frequencies as $p_A p_B + D$, $p_A(1-p_B)-D$, $(1-p_A)p_B-D$, and $(1-p_A)(1-p_B)+D$, respectively, where D is the coefficient of linkage disequilibrium. It is easily shown that $D=x_1x_4-x_2x_3$. If the frequency of recombination between the two loci is c $(0 \le c \le 0.5)$, the value of D in the next generation in an infinitely large, randomly mating population with no selection is $D(1-c)$ (Charlesworth and Charlesworth, 2010).

In the absence of evolutionary forces other than recombination, this result implies that the extent of nonrandom association between a pair of loci decays exponentially at a rate that is determined by the frequency of recombination. This result can be generalized to nonrandomly mating populations and to associations among multiple loci (Charlesworth and Charlesworth, 2010). Unless populations are far from equilibrium, departures from linkage equilibrium at a set of loci require the operation of forces tending to generate nonrandom associations between alleles at the different loci. To be effective, their magnitude must be at least of the order of the recombination frequencies among them. One such force is random genetic drift (*see* Random Genetic Drift), which can cause randomly generated linkage disequilibrium among loci for which c is of the order of the reciprocal of the population size (Ewens, 2004; Charlesworth and Charlesworth, 2010). DNA sequence surveys of populations of organisms such as Drosophila and humans reveal patterns of dependence between D values for pairs of sites and their distance apart on the chromosome that suggest that drift plays an important role in generating nonrandom associations between variants. These associations can be used to locate variants that affect quantitative traits, including human diseases (Flint and Mackay, 2009; 1000 Genomes Project Consortium, 2010).

Selection on Several Loci

Another possible force causing linkage disequilibrium is "epistatic selection," in which the difference in fitness between genotypes at one locus varies according to the genotypes at the other loci in the system. If the fitness effects of different loci combine additively, there is no epistasis, and it can be shown that polymorphic equilibria under random mating exhibit no linkage disequilibrium (Ewens, 2004; Charlesworth and Charlesworth, 2010). With epistasis, selection tends to preserve haplotypes that contain favorable combinations of alleles, whereas recombination breaks them down. Linkage disequilibrium is not necessarily present if linkage is sufficiently loose in relation to the strength of epistasis. Multiple alternative stable equilibria may occur in multilocus systems, so that the fate of a population can be affected by the initial conditions from which evolution starts. Stable equilibria do not necessarily correspond to the maxima in mean fitness in the space of haplotype frequencies, another violation of the adaptive landscape principle. However, if epistatic selection is weak in relation to the frequency of recombination, populations tend to converge to trajectories where linkage disequilibrium is nearly constant (quasi-linkage equilibrium), and mean fitness increases monotonically at a rate approximately equal to the additive genetic variance in fitness, according to Fisher's fundamental theorem of natural selection (Ewens, 2004; Charlesworth and Charlesworth, 2010). This has provided a very useful tool for the analysis of the dynamics of multilocus systems (Barton and Turelli, 1991).

Evolution of Close Linkage

There are two biologically important features of systems with strong epistatic selection. First of all, such selection may impose strong constraints on the degree of linkage between polymorphic loci. Suppose that the population is initially segregating for alleles A and a at one locus but is initially fixed for b at a second locus. If a mutation B arises at this locus, which interacts with the alleles at the first locus in such a way that AB is selectively favored but aB is disfavored, B may be unable to invade the population unless c is below some threshold value. Only mutations at loci that are sufficiently closely linked to the first polymorphism in the system will be able to establish subsequent polymorphisms. This process has probably been important in the evolution of some of the classic examples of "supergenes" (systems of very closely linked loci held in strong linkage disequilibrium by selection), such as Batesian mimicry in butterflies (Ford, 1975) and sex

chromosomes (Charlesworth and Charlesworth, 2010). Similarly, if ab and AB are both fitter than Ab and aB, a population fixed for ab can only evolve a two-locus polymorphism if there is a double mutation to AB and if c is sufficiently small. This has probably occurred in the evolution of meiotic drive systems, which require combinations of alleles at several loci where each allele on its own is disfavored (Charlesworth and Charlesworth, 2010).

Second, there is a selective advantage to modifier alleles that reduce the frequency of genetic recombination between the two loci, once a two-locus polymorphism has been established. If suitable genetic variation in recombination rates is available, this will eventually lead to very close linkage (Fisher, 1930). This principle has wide generality; analysis of the conditions for spread of rare modifiers of recombination rates has shown that randomly mating populations under epistatic selection generating linkage disequilibrium will always tend to evolve a closer linkage (Charlesworth and Charlesworth, 2010). Since genetic recombination is a near-universal feature of living organisms, this has led to the search for situations that promote rather than repress recombination; these mostly involve forces such as mutation, genetic drift, and environmental change that perturb populations away from their equilibria under selection alone (Charlesworth and Charlesworth, 2010)

Random Genetic Drift

The discovery that random sampling of allele frequencies in finite populations may be a significant factor in evolution is another major contribution of population genetics. This process has two aspects: The first is the tendency for a population of finite size to become genetically uniform, owing to the fact that there is an increasing chance as time passes that all the copies of a gene are descended from a single ancestral allele (**Figure 2**). The second is the tendency of isolated populations to diverge in allele frequencies over time, since independent trials of a population with the same initial state will arrive at different allele frequencies by chance.

The first process is closely related to the increase in homozygosity that accompanies the inbreeding of close relatives; both are conveniently studied by means of the concept of identity by descent. Two alleles at the same locus drawn from a population are said to be "identical by descent" if they trace their ancestry back to a single ancestral allele. The extent to which a population has progressed toward genetic uniformity can be measured by its inbreeding coefficient, defined as the probability that a pair of randomly sampled alleles are identical by descent from a common ancestral allele (Charlesworth and Charlesworth, 2010). This is always measured with respect to an initial generation, in which all the alleles at a locus are arbitrarily decreed to be nonidentical (**Figure 2**).

Increase in Homozygosity

The simplest "Wright–Fisher" model of genetic drift assumes a discrete-generation, randomly mating population of N hermaphroditic individuals with no selective differences among genotypes at the locus under consideration. New individuals are formed by random sampling (with replacement) of gametes produced by the parents. With diploid inheritance, such that there are $2N$ gene copies at a locus among the breeding adults, the following recursion relation for f_t, the inbreeeding coefficient in generation t, is obtained:

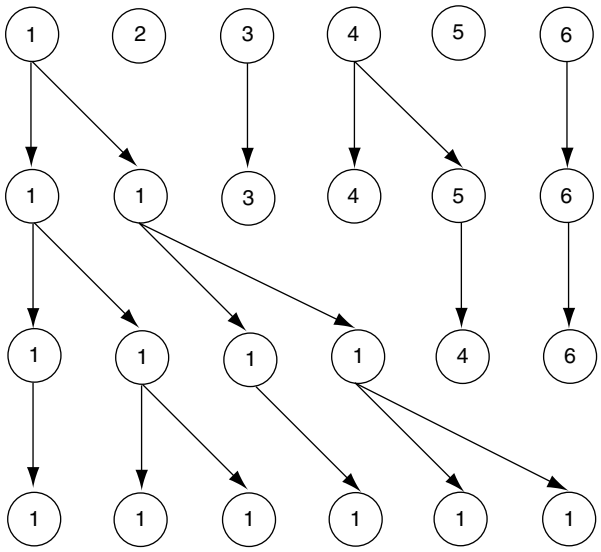


Figure 2 Each row of circles indicates the alleles at a locus present in a given generation. The distinct copies present in the initial generation are labeled 1–6. The arrows indicate the successful transmission of an allele to the next generation.

$$1 - f_t = \left(1 - \frac{1}{2N}\right)(1 - f_{t-1}) \qquad [5]$$

This follows from the fact that the probability that a pair of randomly sampled alleles are both derived from the same allele in the previous generation is $1/(2N)$, in which case they are identical by descent. The probability that they come from two different alleles is $1–1/(2N)$; their probability of identity is then f_{t-1}.

Equation [5] shows that f_t tends asymptotically to 1; if N is moderately large, $1–f_t$ decays exponentially with a half-life of $1.4N$ generations. This result can be generalized to more realistic types of breeding system by means of the concept of the inbreeding effective population size (N_e) (Charlesworth and Charlesworth, 2010). This utilizes the fact that genetic drift in these more general cases can be described by matrix equations or higher-order difference equations, such that a constant rate of decay of $1–f_t$ is reached asymptotically, yielding an expression of the same form as eqn [5]. The asymptotic decay constant can be equated to $1–l/(2N_e)$. For example, a population with N_f breeding females and N_m males has $N_e = 4N_f\, N_m/(N_{f+}N_m)$ if there is purely random variation among individuals in their offspring number. In general, N_e is less than the census number of breeding individuals, often considerably so (Charlesworth and Charlesworth, 2010).

Differentiation of Populations

The effect of drift in causing genetic differentiation among populations is described by the variance in allele frequencies among the set of populations in question. In the Wright–Fisher model, the sampling of allele frequencies at a neutral biallelic locus with alleles A and a and frequencies p and q will generate a binomial distribution of the new frequency of A, with mean p and variance $pq/(2N)$. Repetition of this sampling process results in a probability distribution of allele frequencies with steadily increasing variance (Charlesworth and Charlesworth, 2010). The variance of allele frequency, σ_p^2, can be related to the inbreeding coefficient as follows: Assume that we have a large set of completely isolated populations, all founded with the same initial frequency, p_0. At some arbitrary time, the set of populations will have a mean gene frequency p_0 (since drift does not change the mean) and variance σ_p^2. If N is reasonably large, the genotype frequencies within each given population will be in Hardy–Weinberg equilibrium; the mean frequencies of AA, Aa, and aa over the set of populations are equal to $p_0^2+\sigma_p^2$, $2(p_0q_0-\sigma_p^2)$, and $q_0^2+\sigma_p^2$, respectively.

This is the "Wahlund effect": Genotype frequencies averaged over a set of populations that are each in Hardy–Weinberg proportions show an excess of homozygotes and a deficiency of heterozygotes compared with Hardy–Weinberg expectation, the magnitude of which is determined by the variance in gene frequencies among the populations. This is a purely algebraic result, independent of the causes of the variation. It can be related as follows to the effects of genetic drift in causing increased homozygosity. Assume that a diploid individual is formed by sampling two random alleles from the same population. The probability that the alleles are identical by descent is f; they are then both A in state with probability p_0. The probability that they are nonidentical is $1-f$, in which case the probability that the two alleles are both A is p_0^2. The net probability that the individual is AA is thus $fp_0+(1-f)p_0^2=p_0^2+fp_0q_0$. Comparison with the previous expression shows that

$$\sigma_p^2 = fp_0q_0 \qquad [6]$$

This establishes that the variance in allele frequency is proportional to the inbreeding coefficient under the Wright–Fisher model, and so its change over time is governed by eqn [5]. This is often but not always true under more general models of population structure, with N_e replacing N in the binomial formula for the variance conditional on the current gene frequency. In some circumstances, particularly when the population size changes with time, the conditional variance in allele frequency requires a different denominator in order to be represented by the binomial formula. In this case, a "variance effective population size" is computed (Ewens, 2004; Charlesworth and Charlesworth, 2010).

Molecular Evolution and Variation

The Neutral Theory

These simple models of genetic drift can readily be applied to the study of molecular evolution and variation, assuming selective neutrality at the loci in question. Neutral theory allows for the possibility that many mutations are subject to purifying selection and are rapidly eliminated from the population (*see* Fixation Probabilities); it is claimed, however, that the fate of most mutations that are not removed by purifying selection is determined by drift rather than by selection (Kimura, 1983). This theory constitutes a useful null hypothesis, which can be tested against data on molecular evolution and variation by means of predictions of several different types.

Rate of Neutral Evolution

Consider first the rate of molecular evolution as measured by the rate of substitution of mutations, K (*see* Substitutional Load). In a population of N breeding adults, there are $2N$ allele copies at an autosomal locus. As shown in the section Increase in Homozygosity, eqn [5] implies that the population tends to homozygosity with probability of 1. This means that the remote descendants of the current population will all trace their ancestry back to just one of these $2N$ alleles. Under neutrality, the probability that a given allele is the ancestor is thus $1/(2N)$. It follows that the probability of fixation of a new neutral mutation in a population of size N is $1/(2N)$; the probability that it is lost is $1-1/(2N)$. If the rate of mutation to neutral variants is u per generation, the expected number of new mutations that enter the population is $2Nu$, of which only $1/(2N)$ are destined for ultimate fixation; the expected number of mutations that ultimately become fixed is $2Nu/(2N)=u$.

If the process of mutation and drift has reached a stationary state, so that the expected number of new substitutions is equal to the number of substitutions that reach completion in each generation, we obtain the fundamental equation of neutral molecular evolution:

$$K = u \qquad [7]$$

This relation holds for any level in the genetic hierarchy, from the nucleotide site through the locus to the genome as a whole, provided that K and u are defined appropriately. Since K can be determined by comparisons of DNA sequences among species with known divergence time (*see* Substitutional Load), and u values are known from genetic studies, eqn [7] can be used to test the neutral theory.

One prediction is that genomic regions whose sequences are essentially functionless, such as pseudogenes, should evolve at the mutation rate, since they are necessarily unconstrained by selection. If adaptive Darwinian evolution is a minor factor in molecular evolution, the rate for sites in these regions should be much higher than that for functionally significant regions, where selection is expected to eliminate most of the mutations (see eqn [10] below). These regions do indeed evolve at similar rates to those expected from mutation rate estimates, and other regions evolve more slowly (Kimura, 1983; Charlesworth and Charlesworth, 2010). This does not, however, rule out a role for selection in fixing advantageous variants in selectively constrained regions; it could still be true, for example, that most nonsynonymous mutations (*see* Measurement of Variation in Nucleotide Sequences) are deleterious or neutral, but a minority are advantageous rather than neutral, so that the changes that are fixed in evolution are adaptive rather than neutral. There are also some exceptional cases of higher rates of nonsynonymous

versus synonymous substitutions in coding regions; this is a strong evidence for a positive role of selection on the amino acid sequences concerned (Charlesworth and Charlesworth, 2010).

Another prediction of eqn [7] is that the rate of molecular evolution should be constant over long periods of time, since it depends only on the mutation rate. In contrast, the rate of evolution under natural selection is expected to be highly variable, because the theory described earlier (*see* Selection on Quantitative Traits) suggests that populations will tend to adapt quickly to a new environment (or go extinct), after which change will be slow or nonexistent. This is borne out by the observations of comparative biology and paleontology, which show that evolution at the external phenotypic level is generally highly episodic and triggered by ecological opportunities such as the occupation of vacant niches. Studies of DNA and protein sequence evolution suggest that the rate of evolution of a given molecule is much less variable among different lineages, or within the same lineage at different times, than is true for the external phenotype, especially when noncoding or silent substitutions are considered (Kimura, 1983; Charlesworth and Charlesworth, 2010). This has generated the concept of a "molecular clock," which is used to estimate the times of divergence of taxa when paleontological data are absent. However, there is evidence for more variability in rates of evolution than predicted by the simplest form of the neutral theory, especially for amino acid sequences. This suggests a role for selection, although the interpretation of such rate variability is controversial (Charlesworth and Charlesworth, 2010)

Neutral Polymorphism

The process of fixation of neutral variants is a slow one; calculations based on diffusion theory (*see* Diffusion Equations) show that the mean time for fixation of a new neutral mutation in a random mating population (conditioned on ultimate fixation) is approximately $4N_e$ generations, during which the variant on its way to fixation causes a polymorphism. Variability is also contributed by the large fraction of new mutations that are destined for ultimate loss. In the neutral theory, polymorphism is simply a phase of molecular evolution. Although there is a constantly shifting set of variants at any one locus, the mean amount of variability can be determined by assuming a statistical equilibrium between drift and mutation, using the following argument (Kimura, 1983; Charlesworth and Charlesworth, 2010).

The simplest version of this applies to a single locus, which is assumed to have no recombination. Thus, there are many possible sequences; new neutral mutations are assumed to occur with probability u per generation, and each mutation represents a sequence that has not been observed before. This is the "infinite alleles" model. Let h_t be the probability that two randomly sampled alleles are distinct in sequence in generation t. An argument similar to that leading to eqn [5] shows that, after neglecting terms in $u/(2N)$, h_t for a Wright–Fisher population obeys the equation

$$h_t \approx (1 - \frac{1}{2N})h_{t-1} + 2u(1 - h_{t-1}) \qquad [8a]$$

At equilibrium, rearrangement of this equation yields the relation

$$h \approx \frac{4Nu}{4Nu + 1} \qquad [8b]$$

More generally, N_e can be substituted for N in eqn [8b]. The parameter $\theta = 4N_e u$ thus controls the equilibrium level of variability under the neutral model. A large amount of neutral variability can be maintained, provided θ is sufficiently large, for example, with $N_e = 10^5$ and $u = 2 \times 10^{-7}$, $\theta = 0.08$, and $h = 0.074$. This value of h is similar to the mean per locus heterozygosity for electrophoretic alleles in mammalian populations (Lewontin, 1974; Kimura, 1983).

This model can be modified to predict the equilibrium level of diversity per nucleotide site by assuming that the units of observation are the individual sites, not the entire locus. If it is assumed that θ is much less than 1, so that at most one variant is segregating in the population at each site, eqn [8b] can be applied to yield the equilibrium value of π, the nucleotide site diversity (*see* Measurement of Variation in Nucleotide Sequences), such that π is approximately equal to θ. This is Kimura's "infinite sites" model, which is widely used in the interpretation of data on molecular variation.

The Coalescent Process

General Considerations

The growing body of data on DNA sequence variation within populations has stimulated interest in the development of statistical tests for the agreement of the observed patterns of variation with the predictions of the neutral theory. To conduct such tests, it is essential to have predictions concerning the properties of statistics describing samples of alleles from populations and not merely of properties of the populations from which the samples are drawn, since these cannot be observed directly. A powerful method for deriving the properties of samples has been developed, known as coalescent theory (Ewens, 2004; Wakeley, 2008; Charlesworth and Charlesworth, 2010). It is based on the following principle. Consider a pair of alleles at an autosomal locus sampled from a Wright–Fisher population. As shown in the section Increase in Homozygosity, there is a probability $1/(2N)$ that they are derived from a common ancestral allele in the previous generation, that is, they "coalesce." If they fail to coalesce, which has probability $1–1/(2N)$, they have a probability $1/(2N)$ of coalescing in the next generation back, and so on.

There is thus a geometric distribution of the time back to the common ancestral allele, such that the probability of time t is $(1/2N)(1–1/[2N])^{t-1}$. From the well-known properties of this distribution, the mean time back to the common ancestor is close to $2N$ and the variance is approximately $(2N)^2$. Assume that mutations occur at the rate u per site in the gene, such that each mutation that arises in the line back to descent connecting the two sampled alleles is at a different site (the infinite sites assumption, *see* Neutral Polymorphism). If there are m nucleotide sites in the sequence in question, the number of mutations conditioned on t has a mean and variance of $2tmu$ since the total time separating the alleles is $2t$, and the conditional number of mutations follows a Poisson

distribution. The mean and variance of the number of differences between the two alleles are then $m\theta$ and $m\theta + (m\theta)^2$, respectively. The result for the mean corresponds to that derived from eqn [8b] since the nucleotide site diversity is simply the number of differences between the pair of alleles divided by m, and the mean of t is equal to $2N$.

This can be extended to a sample of n alleles from a population. If N is sufficiently large, the chance of more than one coalescent event per generation can be neglected. There are $n(n–1)/2$ possible pairwise allelic combinations, so that the probability of a coalescent event is $n(n–1)/(4N)$. The time to the first coalescent event is therefore distributed geometrically, with a mean of approximately $4N/(n[n–1])$, and the variance is approximately equal to the square of the mean. The time from this to the next coalescent is also geometrically distributed, replacing n with $n–$ 1, and so on. The process can be represented by a genealogical tree, whose nodes with k and $k–1$ alleles are separated by a time tk that is geometrically distributed with mean of $2/(k[k–1])$ on the coalescent timescale of $2N$ generations (**Figure 3**). As previously mentioned (*see* Increase in Homozygosity), N_e replaces N for more general models of breeding structure. The rate of coalescence evidently diminishes as the number of nodes decreases.

This representation can be used to derive many important results, and it also provides a rapid means of simulating genetic processes, since time can be rescaled to the coalescent timescale and the properties of a sample of alleles represented by generating genealogical trees from samples drawn from the relevant distributions, with mutations scattered randomly over the branches of the tree. The expected pairwise difference between all $n(n–1)/2$ pairs of alleles on the infinite sites assumption is readily seen to be $m\theta$. Its variance is also known (Ewens 2004; Wakeley, 2008; Charlesworth and Charlesworth, 2010). This statistic provides the obvious means of estimating the nucleotide site diversity in the population (π) by equating the observed mean pairwise difference to its expectation. However, another statistic, the total number of segregating sites in a sample of n alleles (S_n), has better statistical properties under the infinite sites model. These can be obtained as follows: The total size of a tree is the sum of kt_k over the entire tree. Application of the above formulae for mean coalescent times with different numbers of alleles shows that the expectation of the tree size in units of coalescent time is simply $A = 2\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots \frac{1}{(n-1)}\right)$. The number of sites in the sample that are segregating for nucleotide variants is the number of mutations that appear on the tree. Its expectation is $2NmuA = m\theta A/2$; θ can therefore be estimated by dividing S_n by $mA/2$. The variance of this estimator is substantially lower than that based on pairwise diversity.

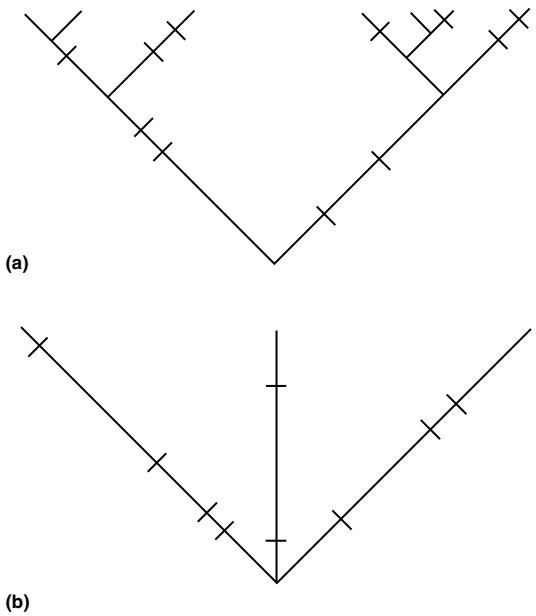


Figure 3 Two types of gene genealogies. (a) The genealogy of a set of alleles in a population of constant size. The tips of the tree represent six alleles sampled from the population. Each node corresponds to a coalescent event. Each cross line represents a new mutation in the DNA sequence of the gene in question, which has arisen since the common ancestor of the sample; mutations in internal branches of the tree give rise to at least two variants in the sample. (b) The genealogy of a set of three alleles in a population that experienced a recent reduction in size to one allele, followed by rapid expansion with no opportunity for coalescent events.

Tests of Neutrality

Several tests for departures from neutrality have been devised based on the properties of the coalescent process (Ewens 2004; Wakeley, 2008; Charlesworth and Charlesworth, 2010). One widely used method is Tajima's D test, which uses the fact that departures from neutral equilibrium have different effects on the pairwise difference estimate of θ and on the estimate of S_n. For example, a population bottleneck of reduced size that completely removes variability at a locus will be followed by a long period of recovery, during which new mutations are all at low frequencies. This means that the pairwise diversity, which is weighted by the frequencies of variants, will be much lower in relation to the observed number of segregating sites than in an equilibrium population, reflecting the fact that the genealogical tree is like a "star phylogeny" in this case (**Figure 3**). Purifying selection on the variants in question will have a similar but much smaller effect. The opposite pattern is expected if there has been a partial bottleneck, in which rare variants have been lost from the population, or if variation is affected by balancing selection. The Tajima test computes the ratio of the difference between the two estimates of θ to its standard deviation (calculated using the neutral equilibrium model), and compares the result with critical values obtained from coalescent simulations. Other tests have been devised using similar principles (Ewens, 2004; Wakeley, 2008; Charlesworth and Charlesworth, 2010).

Interaction of Drift with Deterministic Forces

Population Subdivision

General Considerations

One important complication of the neutral theory is population subdivision; species in nature are not simple randomly mating populations but instead are usually distributed over wide geographic areas, with limited migration between localities. Random genetic drift can thus produce significant genetic

differentiation among local populations if migration is sufficiently restricted. This creates two problems: How to describe data on allele or nucleotide site variant frequencies when there is differentiation among populations, and how to relate models of the underlying evolutionary processes to the data?

Partitioning Variability

The basic method used to summarize data collected from a set of populations sampled from the same species is to partition variability into a within-population component that describes the mean level of variability within a sample, and a between-population component that measures the mean difference between alleles sampled from different populations. This can be done either using data on allele frequencies, as in the case of electrophoretic loci or microsatellite loci, or using nucleotide site variant frequencies. In many cases, it is also possible to organize the populations sampled into a hierarchy of decreasing geographic scale, such as subspecies within a species, races within subspecies, and local populations within races. Variability can then be partitioned into variation between and within each level of the hierarchy (Charlesworth and Charlesworth, 2010). A frequent practice is to normalize each lower level measure of between-population variability relative to variability at the next higher level of the hierarchy, generating a set of Wright's F statistics or the related G statistics of Nei (Charlesworth and Charlesworth, 2010).

The procedures will be illustrated here for the case of DNA sequence variation in a population divided into a set of equal-ranking subpopulations. The total nucleotide site diversity, π_T, is defined as the fraction of nucleotides that differ between a pair of alleles drawn randomly from the set of samples as a whole. The within-population nucleotide site diversity, π_S, is the mean fraction that differs between a pair of alleles drawn from the same subpopulation. The amount of between-population differentiation can be measured by the between-population component of nucleotide site diversity, $\pi_{T-S}=\pi_T-\pi_S$. A normalized measure of differentiation is defined as $K_{ST}=\pi_{T-S}/\pi_S$.

The Island Model

Although the previously mentioned measures of population differentiation are useful as descriptive tools, they can also be used to estimate the evolutionary parameters that determine the extent of population differentiation. This requires the development of models of the joint effects of genetic drift, mutation, and migration, one of the most complex problems in theoretical population genetics. The simplest and most widely used model is the island model, which assumes that the species is divided into d distinct subpopulations or "demes," which each behaves according to the Wright–Fisher model with population size N. After reproduction has occured within each deme, a fraction m of each deme's genes are replaced with genes drawn randomly from the other $d-1$ demes. Coalescent theory can be used to determine the mean coalescence times of pairs of alleles sampled from the same population (t_0) or from different populations (t_1); it is found that $t_0=2Nd$ and $t_1=2Nd(1+[d-1]/[4Ndnm])$ (Wakeley, 2008; Charlesworth and Charlesworth, 2010). In the infinite sites model, the mean fraction of nucleotides that differ between a pair of alleles is equal to the product of the mutation rate per site and twice their coalescence time (*see* The Coalescent Process) so that the expected number of nucleotide site differences between alleles can be derived directly from the corresponding coalescence times.

An important and somewhat counterintuitive conclusion is that the within-population nucleotide site diversity, π_S, is equal to $4Ndu$, that is, it depends on the total number of individuals in the set of populations in the same way as the diversity in a panmictic population under the infinite sites model, and it is independent of the migration rate (with the proviso that $m>0$). As expected, the other diversity measures depend inversely on Nm, with large between-population divergence being possible only when $Nm<1$; for large d, K_{ST} is approximately equal to $1/(1+4Nm)$. Values of these statistics that are close to zero are generally taken to indicate relatively little population differentiation, whereas values close to one imply considerable differences among local populations relative to the within-population variability.

Other Models of Population Structure

Although simple to analyze, the island model is not very realistic biologically since dispersal is limited in most species, so that local populations are most likely to acquire immigrants from nearby. Attempts to generate useful results from more realistic models have taken two directions. The first involves maintaining the assumption of a set of discrete demes but allowing for differences in deme sizes. Migration is described by a migration matrix M, such that m_{ij} is the probability that an allele currently in deme i originated in deme j in the previous generation. However, it is difficult to obtain transparent general results for such a model. Under the infinite sites model, however, the result that $\pi_S=4Ndu$ still holds if migration is conservative, that is, migration does not change the sizes of local populations (Wakeley, 2008; Charlesworth and Charlesworth, 2010). Simple analytical results on genetic diversity within and between populations are only available for special cases, such as the stepping-stone model, in which demes of size N in a linear or planar array receive migrants with probability m only from immediately adjacent populations (Wakeley, 2008; Charlesworth and Charlesworth, 2010). This model also permits the analysis of the dependence of degree of genetic differentiation between populations on the distance between them. The results show that extensive differentiation between populations is possible with a one-dimensional array of populations, even if there is a considerable amount of migration, whereas the results for a two-dimensional array are similar to those for the island model.

The second approach assumes that individuals are distributed over a one- or two-dimensional spatial continuum, with density ρ. Migration is represented by the probability density that an individual moves a given distance between birth and reproduction. If migration follows a random walk, the variance of the migration function, σ^2, is sufficient to describe the process. This is Wright's isolation-by-distance model. A mathematically correct formulation of this model presents formidable difficulties, but it is possible to obtain useful approximate results (Charlesworth and Charlesworth, 2010).

Effects of Directional Selection at Linked Loci

Another complication in interpreting data on DNA sequence variation is that even neutral variation may be affected by selection at the linked sites. The classic example is hitchhiking, where a new favorable mutation arises as a unique event and spreads throughout the population. In the absence of recombination, variants at linked loci will be dragged to fixation as the favorable allele sweeps through the population so that variability at these loci is eliminated. With recombination, the magnitude of this hitchhiking effect decreases with the ratio of the frequency of recombination between the selected and neutral locus (c) to the selection coefficient (s) at the selected locus, and it is negligible when this ratio is of the order of 1 or more (Charlesworth and Charlesworth, 2010). The effects of hitchhiking by favorable mutations are thus only likely to be manifest at sites that are closely linked to the target of selection. There are many examples of the detection of such hitchhiking events in natural and human populations (Charlesworth and Charlesworth, 2010).

However, hitchhiking effects can also be caused by deleterious mutations, a process termed background selection (Charlesworth and Charlesworth, 2010). A neutral variant which is tightly linked to a deleterious mutation that is in the process of elimination from the population will have a high chance of being eliminated before it can unhitch itself by recombination. As discussed previously (*see* Mutation–Selection Balance), the genomes of higher organisms contain many loci subject to mutation to deleterious alleles; therefore, relatively small individual effects of selection against mutations at single loci may have a large cumulative effect on neutral variability.

Studies of genetic variation in Drosophila and some other species have shown a strong relationship between the amount of recombination in the region where a gene is located and the level of genetic variation that it exhibits (Charlesworth and Charlesworth, 2010). Similarly, species or populations of hermaphroditic organisms with high rates of self-fertilization, in which the absence of heterozygotes means that genetic recombination is effectively absent, also seem to show reduced levels of molecular variation. In human populations, variability at the noncoding sites that are closer to genes is lower than sites situated far away from genes. Both hitchhiking by favorable mutations and background selection can account for these patterns, and current research is attempting to distinguish between them.

Effects of Linkage to Balanced Polymorphisms and Targets of Local Selection

Selection at linked sites can also cause increased variability at the neutral sites if selection is balancing rather than directional in nature. This can be understood in terms of a single locus with two alleles, A and a, maintained by strong balancing selection in a randomly mating population. This means that the population is effectively divided into two subpopulations represented by the two allelic classes. If the two alleles are equally frequent, a neutral site linked to this locus with recombination frequency c has a probability $0.5c$ of crossing over from one subpopulation to the other; this is equivalent to the migration rate m in the case of an island model with just two islands. This leads to an expansion of the coalescent time at the neutral site by a factor of $1 + 1/(4N_e c)$, where N_e is the effective population size. This implies that the maintenance of variability by selection is accompanied by a corresponding increase in nucleotide site diversity at the sites that are very closely linked to the target of selection, as seen at the MHC loci in mammals (Charlesworth and Charlesworth, 2010).

Increases in neutral variation among subpopulations of a species may also occur at sites linked to loci subject to local selection, causing large among-population differences in allele frequencies. Variation in the behavior of neutral variability along a chromosomal region can thus provide valuable evidence on the action of spatial differences in selection pressure, and genome-wide scans for signatures of these effects in human populations have been carried out with considerable success (Charlesworth and Charlesworth, 2010; Hancock *et al.*, 2010).

Diffusion Equations

General Considerations

A full treatment of genetic drift must deal with properties other than summary statistics of allele frequency distributions of the type considered so far, particularly if selection is to be studied. The difficulties involved in exact analytical treatments of the properties of genotype frequency distributions are great, so resort is usually made to approximations involving the use of diffusion equations, which treat genotype frequencies and time as continuous variables and assume that all the evolutionary forces are sufficiently weak that second-order terms in their effects on frequencies are negligible. The standard approach is to assume a single biallelic locus. Using the continuity assumptions, the state of the population can be described by either of the following two partial differential equations, writing $\phi(x, p, t)$ for the probability density of the frequency, x, of allele A at time t, given an initial frequency p:

$$\frac{\partial \phi}{\partial t} = \frac{1}{2}\frac{\partial^2(\phi V_{\delta x})}{\partial x^2} - \frac{\partial(\phi M_{\delta x})}{\partial x} \qquad [9a]$$

and

$$\frac{\partial \phi}{\partial t} = \frac{V_{\delta p}}{2}\frac{\partial^2 \phi}{\partial p^2} + M_{\delta p}\frac{\partial \phi}{\partial p} \qquad [9b]$$

where $M_{\delta x}$ is the expected change in gene frequency and $V_{\delta x}$ is the variance of the change in gene frequency, both conditioned on x. $M_{\delta x}$ can be equated to the change in gene frequency in an infinite population; under random sampling of allele frequencies, $V_{\delta x}$ is approximated by $x(1-x)/(2N_e)$.

Equation [9a] is the forward Kolmogorov equation and eqn [9b] is the backward Kolmogorov equation (Ewens, 2004; Charlesworth and Charlesworth, 2010). Multidimensional versions of these equations describe systems with multiple alleles and multiple loci but are usually difficult to analyze.

Stationary Distributions

Equation [9a] is useful for describing the probability density function for an allele frequency. However, even for the simple case of a biallelic locus, a full general solution of this equation

to yield the density as a function of time has been obtained only for some special cases. It is most useful for studying the properties of stationary distributions of gene frequency, when drift comes into statistical equilibrium with mutation, migration, and selection (Ewens, 2004; Charlesworth and Charlesworth, 2010).

The study of such distributions has led to some important conclusions, most notably that selection is highly effective at countering the effects of drift in a randomly mating population when $N_e|s|$ is >1 but is ineffective when $N_e|s|$ is $\ll 1$, where $|s|$ is the magnitude of the selection coefficient defined in the section The Basic Model. When the first condition holds, there is little scatter around the mode of the gene frequency distribution; when the second condition holds, there is a high probability that the population is close to fixation, or fixed, for the disfavored allele. If $N_e|s|$ is of the order of 1, both drift and selection are considered to be significant forces.

Fixation Probabilities

The previous conclusion can also be derived by consideration of the probability of fixation of an allele A with intial frequency p, $P(p)$, for which a general formula was found by Kimura using the backward equation (Kimura, 1983; Ewens, 2004; Charlesworth and Charlesworth, 2010). With intermediate dominance, such that $h=0.5$ in eqn [1a], and assuming that a single copy of the mutation is present initially ($p=1/[2N]$), this formula simplifies as follows:

$$P\left(\frac{1}{2N}\right)=\frac{1-\exp\left(-\frac{N}{N_e}s\right)}{1-\exp(-2N_es)} \qquad [10]$$

For a selectively favorable allele, with $s>0$, the fixation probability tends to $(N/N_e)s$ as N_e tends to infinity. When $N_e=N$ (as in the Wright–Fisher model), this is equivalent to the branching process result for the survival probability of a favorable mutation in a very large population (*see* Survival of Favorable Mutations). The asymptote is approached when $N_es>1$, so that a selection coefficient larger than $1/N_e$ is sufficient to ensure that a new favorable mutation behaves as though the population is infinite. Conversely, a selection coefficient of this size against a deleterious mutation ($s<0$) means that it is almost certain to be eliminated from the population.

These considerations led Fisher (1930) to conclude that random genetic drift is unlikely to be effective as an evolutionary force on the grounds that (1) most species have numbers of breeding individuals at least in the tens of thousands, and usually in the hundreds of thousands or even millions, and (2) it is unlikely that a gene would have such a small effect on the phenotype that its average effect on fitness over evolutionary time would be of the order of 10^{-4} or less. Although a compelling argument with regard to genetic changes that affect the phenotype, this view has been challenged by the neutral theory of molecular evolution.

As already noted, the causes of protein sequence evolution are still controversial, but studies of the statistical properties of between- and within-species patterns of silent substitutions at the third coding positions support the idea that these are affected by both drift and selection in bacteria, yeast, and Drosophila. There is not only a tendency for certain triplets that code for the same amino acid to be favored by selection, but there is also evidence that this selection is so weak that disfavored nucleotide changes can drift to fixation within a species. Estimates of the intensity of selection can be obtained by comparing the observed distributions of frequencies of synonymous nucleotide variants with those predicted by solutions of the diffusion equations, and suggest that N_es for preferred codons is typically approximately 0.5 in Drosophila (Charlesworth and Charlesworth, 2010).

Hill–Robertson Interference

Selection may become ineffective if recombination is restricted among the loci subject to selection, because of linkage disequilibrium generated by random genetic drift (Hill–Robertson interference: Charlesworth and Charlesworth, 2010). There are several different processes of this kind, but all involve the key feature that selection at one site in the genome impedes the action of selection at another genetically linked site. This is because a finite population does not contain all possible combinations of variants that can be produced by mutation. If mutations (either favorable or deleterious) arise in different individuals, a favorable variant at one site will generally be present in a genotype with a deleterious variant at other sites, that is, there is negative linkage disequilibrium between selectively favorable variants. As recombination breaks down linkage disequilibrium, it enhances the population's ability to respond to selection. The selective sweep and background selection processes mentioned in the section Effects of Directional Selection at Linked Loci above are in fact examples of Hill–Robertson interference where the variants concerned are not all under some form of selection.

The background selection model assumes the continual occurrence of mutations that are sufficiently deleterious that their expected frequencies remain close to their low equilibrium values in an infinite population. In nonrecombining genomes or genomic regions, however, deleterious mutations will not necessarily behave like this, but instead may accumulate; this process is called Muller's ratchet. Consider, for example, the case of a haploid asexual population in which mutations occur exclusively from wild-type to deleterious alleles but not in the opposite direction (*see* Mutation–Selection Balance), and where $N_es\gg 1$. If the selective effects of mutations at different loci are identical, a population can be characterized by the frequencies of genomes containing 0, 1, 2 … mutations. If the frequency of the mutational class containing the lowest number of mutations (the least-loaded class) is sufficiently small, it will be lost from the population after a finite number of generations. Given the assumed irreversible nature of mutation and the lack of opportunity for genetic recombination, the least-loaded class cannot be reconstituted and will be replaced by the class with one more mutation. This class is now vulnerable to stochastic loss in the same way. There is thus a repetitive process of loss of successive least-loaded classes, in which the loss of each class can be regarded as a turn of Muller's ratchet. The speed of the ratchet depends in a complex way on the effective population size, the number of individuals in the least-loaded class in an equilibrium population, and the strength of selection

(Charlesworth and Charlesworth, 2010). If it does operate, there will be an approximately exponential decline in the mean fitness of the population.

In the results previously discussed, which suggest that individuals in equilibrium sexual populations carry many hundreds or thousands of deleterious mutations, even a very large asexual population of a higher organism is likely to be vulnerable to the operation of the ratchet and other forms of Hill–Robertson interference, probably leading to its eventual extinction. Asexual prokaryotes, with their much smaller genomes and enormous population sizes, are less likely to suffer this fate. This may account for the fact that very few species of higher organisms, with their large genomes, are asexual or lack recombination, whereas prokaryotes and mitochondria often have very low levels of recombination. It may also contribute to the loss of most ancestrally functional genes from Y chromosomes (Charlesworth and Charlesworth, 2010).

Group Selection

Another situation where the supremacy of individual selection can be overcome occurs when species are subdivided into local populations, among which migration is so restricted in relation to the reciprocal of local population size that drift can overcome its homogenizing influence (*see* Population Subdivision). An allele that is deleterious within its local population may drift to fixation locally in opposition to selection, since N_e for the local population is much smaller than for the species as a whole. This creates the opportunity for group or interdeme selection. If the allele in question causes its carriers to be altruistic, in the sense of conferring increased fitness on the members of their deme at the expense of suffering a loss in fitness to themselves, demes in which the allele reaches high frequency or fixation will achieve a higher mean fitness. This may render them less susceptible to extinction, or more able to contribute to the pool of migrants or to find new demes. Selection among demes can therefore result in the spread of an allele that causes a reduction in individual fitness but benefits the population as a whole (Haldane, 1932; Charlesworth and Charlesworth, 2010).

Although this is a theoretically viable mechanism, there are reasons to doubt that it is widely applicable to evolution in nature, since it requires high levels of genetic differentiation between local, as well as the appropriate balance of selective effects at the individual and group level. Studies of molecular genetic variation within and among populations indicate that many species lack extensive differentiation among local populations (*see* Partitioning Variability), suggesting that migration is often so effective that the necessary conditions for group selection are not met.

Kin Selection

Most examples of altruistic behavior, such as the sterility of the worker castes in social insects, are generally believed to be the result of kin selection rather than group selection. An altruistic genotype that aids relatives may experience a selective advantage, even in a large, randomly mating population, if the fitness benefit b to the recipients of the altruism is sufficiently large in relation to the cost c to the altruists (Haldane, 1932; Charlesworth and Charlesworth, 2010). This is because the genetic similarity of related individuals means that the relatives of an altruist have a greater chance than average of carrying an allele that promotes altruism; according to "Hamilton's rule," there will be an increase in the frequency of an altruistic allele when $b\ r > c$, where r is a measure of the relatedness of the altruists to the recipients. There is an extensive theoretical literature on the correct way to calculate r under different conditions, and the theory of kin selection has played an important role in the evolutionary interpretation of social behavior and the sex ratio.

The Shifting Balance Theory

A process that is very similar to group selection for altruism is the shifting balance theory of evolution (Wright, 1977; Charlesworth and Charlesworth, 2010). Wright postulated that epistatic interactions in fitness among alleles at different loci are widespread, resulting in multiple stable equilibria under selection. The simplest case is a haploid two-locus, two-allele model, with ab and AB both fitter than Ab and aB. Fixation for ab or AB is stable against introduction of Ab and aB, whereas fixation for Ab or aB is unstable to the introduction of ab and AB. With constant fitnesses and loose linkage, locally stable equilibria are approximated by the peaks in the surface of mean fitness as a function of the allele frequencies at the loci concerned (*see* The Basic Model). A population will approach the peak that is the nearest attractor rather than the highest peak in the landscape. Genetic drift can cause a local population to travel down the adaptive valley separating the current equilibrium to the zone of attraction of a neighboring peak, and selection can then bring it to the new equilibrium. If this is associated with a higher mean fitness than the surrounding peaks, the process of interdeme selection can cause the species as a whole to acquire the genotype associated with this peak, and hence improve in mean fitness.

In contrast to the group selection model of altruism, the new equilibrium is dynamically stable under individual selection. This process has the attractive feature that it allows the species to acquire an adaptively superior genotype or character combination that would otherwise require multiple simultaneous mutations. Its drawback is that it requires a delicate balance among restricted migration, local population size, and the nature and intensity of selection to operate with any frequency. It is difficult to distinguish the end products of the shifting balance process from ordinary individual selection, which does not have such stringent requirements, and it remains unclear whether it has played an important role in adaptive evolution (Charlesworth and Charlesworth, 2010).

Conclusions

This article has necessarily omitted many important topics. It has focused on the basic general principles governing evolutionary change in populations; space did not permit more than a brief mention of the application of these principles to

wider biological problems, including life history evolution, the evolution of genetic and sexual systems, the evolution of social behavior, speciation, and the interpretation of macro-evolution. Very little useful can be said about natural variation and evolution at almost any level without taking population genetic concepts into account.

Appendix

List of Courses

1. Evolution
2. Genetics
3. Human Genetics

See also: Diversity, Molecular Level. Evolution, Theory of. Inbreeding and Outbreeding. Phenotype, a Historical Perspective

References

1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467: 1061–1073.

Barton NH and Turelli M (1991) Natural and sexual selection on many loci. *Genetics* 127: 229–255.

Charlesworth B and Charlesworth D (2010) *Elements of Evolutionary Genetics*. Greenwood Village, CO: Roberts and Co.

Crow JF (1993) Mutation, mean fitness, and genetic load. *Oxford Surveys in Evolutionary Biology* 9: 3–42.

Drake JW, Charlesworth B, Charlesworth D, and Crow JF (1998) Rates of spontaneous mutation. *Genetics* 148: 1667–1686.

Ewens WJ (2004) *Mathematical Population Genetics, 1, Theoretical Introduction*. Berlin: Springer-Verlag.

Falconer DS and Mackay TFC (1996) *An Introduction to Quantitative Genetics*, 4th edn. London: Longman.

Fisher RA (1930) *The Genetical Theory of Natural Selection*, 1st edn. Oxford: Oxford University Press. (Variorum edn., Oxford University Press, Oxford, 1999).

Flint J and Mackay TFC (2009) Genetic architecture of complex traits in mice, flies and humans. *Genome Research* 19: 723–733.

Ford EB (1975) *Ecological Genetics*, 4th edn. London: Chapman & Hall.

Goldstein DB and Schloetterer C (1999) Microsatellites. *Evolution and Applications*. Oxford: Oxford University Press.

Haldane JBS (1932) *The Causes of Evolution*. London: Longmans Green.

Hancock AM, Alkorta-Aranbaru G, Witonsky DB, and Di Rienzo A (2010) Adaptation to new environments in humans: the role of subtle allele frequency shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 2459–2468.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge, UK: Cambridge University Press.

Kwiatkowski DP (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. *American Journal of Human Genetics* 77: 171–192.

Lande R (2000) Quantitative genetics and phenotypic evolution. In: Singh RS and Krimbas CB (eds.) *Evolutionary Genetics: From Molecules to Morphology*, pp. 335–350. Cambridge, UK: Cambridge University Press.

Lewontin RC (1974) *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Lynch M (2010) Rate, molecular spectrum and consequences of human mutation. *Proceedings of the National Academy of Sciences of the United States of America* 107: 961–968.

Wakeley J (2008) *Coalescent Theory, An Introduction*. Greenwood Village, CO: Roberts and Co.

Wright S (1977) *Evolution and the Genetics of Populations Experimental Results and Evolutionary Deductions* 3. Chicago: University of Chicago Press.

Population Stabilization, Human

Samir KC, International Institute for Applied Systems Analysis, Laxenburg, Austria

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Alene Gelbard, volume 4, pp 799–810, © 2001, Elsevier Inc.

Glossary

Family planning The conscious effort of couples to regulate the number and spacing of births through artificial and natural methods of contraception. Family planning connotes conception prevention to avoid pregnancy and abortion, but it can also refer to the efforts of couples to induce pregnancy.

Fertility The actual reproductive performance of an individual, a couple, a group, or a population measured by a variety of rates, including the crude birth rate (number of births per 1000 population) and total fertility rate (TFR).

Life expectancy The average number of additional years a person could expect to live if current mortality trends were to continue for the rest of that person's life, most commonly cited as life expectancy at birth.

Mortality Deaths as a component of population change. Measures include the crude death rate (number of deaths per 1000 population).

Population change An increase or decrease in the size, composition, or distribution of a population resulting from the interaction of births, deaths, and migration in a population in a given period of time.

Population composition The distribution of a population, usually by age and sex, measured by the number and proportion of males and females in different age groups.

Population momentum The tendency for population growth to continue beyond the time that replacement-level fertility has been achieved because of the relatively young age structure of the population.

Replacement-level fertility The level of fertility at which a couple replaces itself in the population. In low-mortality countries, a TFR of 2.1 is considered replacement because some women will die before the end of their childbearing years.

Stable population A population with an unchanging rate of growth and an unchanging age composition as a result of age-specific birth and death rates that have remained constant over a sufficient period of time. Zero population growth is a special case of population stabilization in which no growth occurs because the number of births and deaths are the same.

TFR The average number of children who would be born to a woman during her lifetime if she were to pass through her childbearing years conforming to the age-specific fertility rates of a given year.

History of Human Population Growth

For much of our history, humans have struggled to survive. By AD 1, perhaps 300 million people lived on the Earth, a paltry total after millions of years of human existence. For most of the next 2000 years, population growth was exceedingly slow. High birth rates were often offset by frightful mortality from wars, famines, and epidemics. The bubonic plague, for example, reduced the populations of China and Europe by one-third in the fourteenth century.

Despite dramatic spikes in mortality rates, the number of births exceeded the number of deaths during the seventeenth and eighteenth centuries and population growth proceeded at a slightly faster pace. As shown in **Figure 1**, world population was about 790 million in 1750 and reached 1 billion in approximately 1800.

During the next century, in Europe and in a few other areas throughout the world, better hygiene and public sanitation reduced the incidence of disease. Expanded commerce made food supplies more widely available and improved nutrition. The wild fluctuations in mortality of previous centuries began to recede, and life expectancy began a slow rise. Population grew more quickly and more steadily. Total world population was nearly 1.7 billion by the beginning of the twentieth century and reached 2 billion within the next 30 years.

The nineteenth-century surge of population growth occurred primarily in the more-developed countries. The population of Europe more than doubled between 1800 and 1900, whereas the population of North America increased nearly 12 times – fueled by immigration from Africa and Europe. In 1800, approximately one-fourth of the world population lived in the now more-developed regions of Europe (including Russia), Japan, and North America, but this share increased to about one-third by 1900.

Less-developed countries grew more slowly than the more-developed countries in the nineteenth century, but they already held the bulk of the world's inhabitants. Asia, dominated by China, had 62% of the world population in 1800, and Africa had 11%. Latin America and the Caribbean accounted for only approximately 2% of the world's population. Like North America, Latin America witnessed most of its population growth in the twentieth century.

Demographic Transition

The improvement in human survival and the consequent explosion of population growth marked the beginning of the shift from high to low mortality and from high to low fertility, which is known as the "demographic transition." This shift occurred throughout Europe, North America, and many other

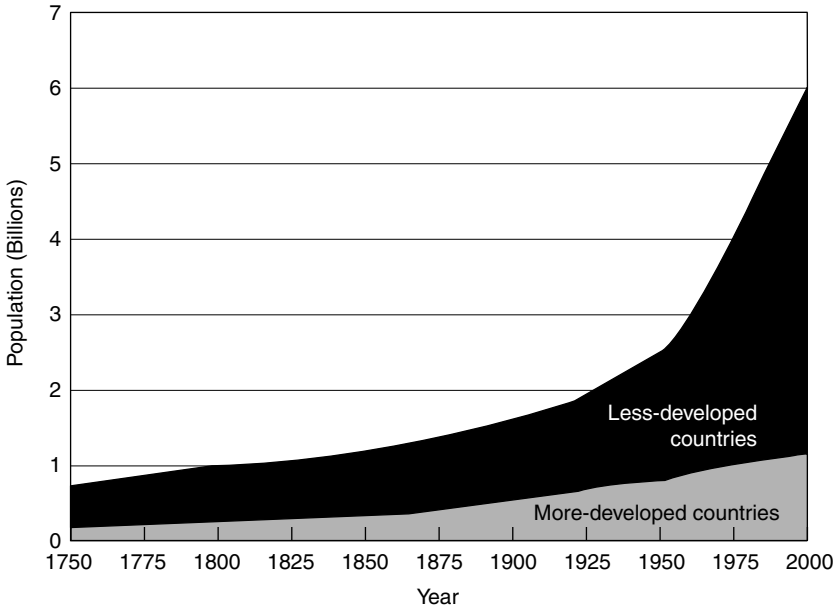


Figure 1 Population growth in more-developed and less-developed countries, 1750–2000.

areas in the nineteenth and early twentieth centuries. It gave rise to the dominant model of demographic change, which most demographers assume applies to all countries. In the classic demographic transition, the trend of high birth and death rates (and minimal population growth) is disrupted by a long-term decline in mortality. Mortality rates eventually stabilize at low levels. Birth rates also begin a long-term decline and decrease to about the same level as mortality rates. With birth and death rates at similar low levels, the equilibrium of slow population growth is regained.

The pace of change in a country will vary depending on its culture, level of economic development, and other factors. As countries pass through the various stages of the transition, population growth from natural increase (birth rate minus death rate) accelerates or declines depending on the gap between the birth rates and death rates. Not all countries will follow the same path to low fertility and low mortality as did European countries. Also, there may be additional stages of transition that have not been identified, for example, long-term population decline. However, the demographic transition theory provides a useful framework for assessing demographic trends and projecting future population size.

Population Change: 1900–1950

When the twentieth century began, more-developed countries were entering a new stage of the demographic transition. In 1900, life expectancy at birth was 47 years in the USA and between 45 and 50 years in Europe, Japan, and Australia, which was a slight increase from an average of about 40 years during the nineteenth century. A revolution in health had already begun and life expectancy would reach unimaginably high levels by midcentury. These improvements in health reflected scientific advances of the previous century. Louis Pasteur, Robert Koch, and others had identified disease-causing "germs," and Joseph Lister introduced antiseptic practices that were eventually adopted by hospitals. Mortality was also declining because of better personal hygiene and public sanitation projects that removed garbage and sewage from city streets and provided safer drinking water. Death rates for infectious diseases began to decrease well before vaccines and antibiotics were widely available.

Infants and young children benefited most from this health revolution. In the more-developed countries, the infant mortality rate (IMR; number of deaths to infants younger than 1 year of age per 1000 births) was approximately 200 in the 1800s – approximately two of every 10 babies died before their first birthday. In the early 1900s, the IMR decreased below 100 in the USA and many European countries. It was lower than 50 in nearly all these countries by the 1950s. The life expectancy at birth in the USA increased to 56 years by 1920 and to 68 years by 1950. Average life expectancy was even higher in some European countries by 1950.

Although birth rates had decreased during the latter part of the nineteenth century, women were still having relatively large families in 1900. According to the United Nations (UN), an American woman had four to five children on average; a European woman had somewhat fewer. Fertility decline quickened after 1900.

During the same period, most of Africa, Asia, and Latin America were still in the predemographic transition stage of high mortality and high fertility.

Population Change: 1950–2010

The second half of the century and the first decade of the twenty-first century brought many new demographic trends and patterns. The more-developed countries completed their transition to low mortality and low fertility and to the lowest low fertility and reversal; for example, in Bulgaria, the fertility

went down to 1.1 children per woman in 1997, lowest low among 27 European countries, and later increased (reversal) to 1.6 children per women in 2009. Population growth slowed and even turned negative in a few countries. Populations grew older. The more-developed countries also experienced sometimes disruptive changes associated with baby booms and baby busts, crises in health, and waves of immigrants and refugees.

In less-developed countries, the second half of the century brought decades of rapid population growth and swelling streams of migrants from rural to urban areas. Some countries appeared to be rushing through the various stages of the demographic transition, whereas others appeared to be following a new path of demographic change.

Mortality, Fertility, and Natural Increase

In Europe, population growth accelerated as countries recovered from the devastating effects of World War II. The rapid decline in death rates of the early part of the century slowed considerably in part because infant and childhood mortality had already decreased to very low levels. By 1975, the IMR was 10 in Japan, 16 in USA, and 15 in much of Europe. According to the US Census Bureau, US life expectancy increased by less than 10 years in the second half of the century, from 68 years to 76 years, after increasing by more than 20 years during the first half.

After World War II, "baby booms" were commonplace in Europe, although they were more modest than the baby boom that occurred in the USA between 1946 and 1964. However, by the mid-1970s, total fertility rates (TFRs; which reflect the average number of children per woman) in many European countries had decreased below two, the level at which a couple replaces itself in the population. A TFR must be slightly higher than 2.0 (about 2.1 in low-mortality countries) to reach replacement level because some women will die before the end of their childbearing years. When the TFR remains lower than two for a prolonged period, populations may experience a natural decrease because deaths will outnumber births.

European fertility had decreased during the 1930s Great Depression, but in the mid-1980s, TFRs sank to record low levels and showed little sign of recovery. By the late 1990s, the TFR was 1.2 or less in Belarus, Bulgaria, the Czech Republic, Estonia, Italy, Latvia, and Spain.

The fertility decline began in Western Europe during a period that saw delayed marriage, more divorce, high inflation, and an increase in the percentage of women attending college and working outside the home. These same social and economic factors favored lower fertility in the USA, in which the TFR reached an all-time low in 1976 at 1.7 children per woman. Below-replacement fertility also occurred in Eastern Europe and the former Soviet Union after 1990.

Two decades of low fertility has halted population growth in nearly all of Europe and Japan. In many cases, a decline in population was avoided only by the flow of immigrants from abroad. In the late 1990s, 14 European countries experienced natural decrease or fewer births than deaths each year.

Natural decrease will spread to other countries as low birth rates drastically reduce the number of people entering the childbearing ages. Although some countries have a net population gain from immigration, this is not expected to generate enough growth to stave off eventual population decline if fertility levels continue to remain low. It is possible that TFRs will increase again in some countries as recent trends toward later marriage and childbearing stabilize and couples who delayed childbearing complete their families. However, at the end of the twentieth century, not one major industrialized country had fertility above replacement level.

Among the more-developed countries, only a few traditional immigration countries (Australia, Canada, New Zealand, and the USA) can expect significant long-term population growth. They have younger age structures and more immigration than do Europe and Japan, which contributes to momentum for continued growth.

Fertility and mortality patterns have been very different among less-developed countries in the last 61 years. Gains in life expectancy accelerated after 1950. The average life expectancy at birth in less-developed countries increased from 42 to 66 years between 1950 and 2011 according to the UN estimates. The IMR decreased from 151 to 50 deaths per 1000 births during the same period.

Average life expectancy increased to more than 60 years in East Asia and Latin America by the early 1970s and to more than 70 years by the late 2000s. The IMR decreased to about 20 in East Asia and 22 in Latin America by 2008.

Progress has been much slower in sub-Saharan Africa and South-Central Asia. In the 1950s, more than 170 infants died per 1000 births in these regions. By the 2000s, the IMR was still close to 85 in sub-Saharan Africa and was nearly 56 in South-Central Asia.

The pace of mortality decline in some areas has been slowed down by the spread of human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), and dramatic declines in life expectancy were observed in some countries of sub-Saharan Africa; however, various national and international efforts have made it possible for life expectancy to rise in the late 2000s. Worldwide, more than 25 million people have died from HIV/AIDS since the beginning of the pandemic in the 1980s. An additional 34 million were living with HIV at the end of 2010. Most will die within the next decade. The UN agency that tracks the HIV/AIDS pandemic, UNAIDS, estimates that there are more than 7000 (down from 16,000 around 2000) new infections daily, 1100 (1600 around 2000) of which are in children (UNAIDS report on the Global AIDS epidemic, 2010)

Population Growth

The general reduction in death rates after 1950 led to explosive population growth in many less-developed countries. For the world as a whole, growth rates peaked during the 1960s and early 1970s at approximately 2% annually. The total population for less-developed countries increased from 1.7 to 5.7 billion between 1950 and 2010. Population growth would have been even higher if fertility rates had not started to decrease in less-developed countries. The pattern and pace of decline varied tremendously, depending on economic and social development, government policies, family planning use, and other factors (see **Box 1**; **Figure 2**).

During the 1950s and 2000s, childbearing levels declined significantly in many regions. Figure 3 illustrates that declines were greatest in Asia and Latin America and much smaller in sub-Saharan Africa. Despite these declines, all developing regions had TFRs above replacement level in 2011 and more than 76 countries had TFRs of 3.0 or higher.

Changing Age Profiles

Fertility, mortality, and migration trends are reflected in the age and sex profiles of the world's countries. The decades of high fertility rates in the less-developed countries meant ever-increasing numbers of young people, illustrated by the broad base of the age–sex pyramid shown in Figure 4. Improvements in infant mortality have also contributed to the expanding youth population. Children aged younger than 15 years comprised three-tenths of the population in the less-developed countries in 2010 and even greater proportions in some regions. In sub-Saharan Africa, children comprised nearly two-fifths (42%) of the population. Elderly people aged 65 years or older comprised only 6% of the population in all less-developed countries and 3% of the population in sub-Saharan Africa.

The base of the population pyramid for less-developed countries shows some narrowing – the result of declining fertility in many countries beginning in the 1980s. However, even with declining fertility rates, the young age structure creates considerable momentum for future growth because the population reaching childbearing age continues to expand. Women currently have fewer children than in the past, but today there are more women having children.

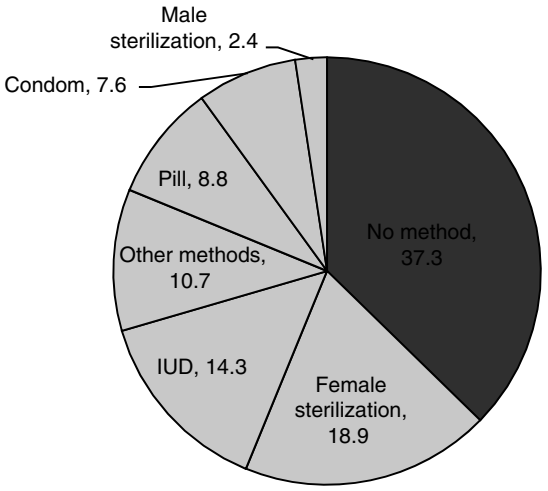


Figure 2 Contraceptive methods percentages used by currently married women, aged 15–49 years, throughout the world, 2009.

Box 1 The reproductive revolution

The "reproductive revolution" was one of the most remarkable events of the second half of the twentieth century. The development of family planning methods, such as the pill and the intrauterine device (IUD), simpler sterilization techniques, and contraceptives that can be injected or implanted under the skin made it easier and safer for women to avoid unintended pregnancies. Increased access to these methods and socioeconomic changes that motivated couples to limit their family size drove the fertility declines in the last few decades. Family planning use increased from less than 10% for married women of childbearing age in the 1960s to more than 62% of this group of women in the late 2000s (**Figure 2**).

Before 1960, women's choices of family planning methods were limited to methods, such as withdrawal, rhythm, diaphragms, foams or jellies, or ineffective methods such as herbal medicines or douching. Women's options improved immensely when the pill and the modern IUD became available after 1960. In the late 2000s, approximately 23% of women worldwide relied on one of these two methods. New contraceptives, including injectables and implants, became available in many countries in the 1980s. They have become popular methods in some African countries. Female sterilization has been widely adopted in Asia and Latin America and is the most popular method worldwide. An estimated 19% of married women (or in a union) aged 15–49 years rely on female sterilization to prevent pregnancy.

The dramatic increase in family planning use caused fertility to decline much more rapidly in the less-developed countries than it had during the fertility transition in the more-developed countries. Organized family planning programs and government promotion of family planning use were important components of this phenomenon. Some experts credit family planning programs for 40–50% of the fertility decline in less-developed countries since the 1960s.

An estimated 120 million couples worldwide want to delay or prevent another pregnancy but are not using family planning. If unmarried sexually active women were included, the number would be much higher, according to survey data.

Family planning use varies widely throughout the world. For example, less than 10% of women use family planning in Mali and less than 20% use it in Afghanistan. However, more than 60% of married women use family planning in Morocco, Argentina, Indonesia, and many other less-developed countries.

The expansion of family planning services has been controversial in some countries. There have also been many obstacles to their use. Many women report that they fear adverse health effects from specific methods. Others want to practice family planning but are dissuaded by their husband's disapproval, their limited decision-making powers, or family pressures to have more children. Some methods are opposed for religious reasons. Difficulties in obtaining and transporting supplies and a shortage of trained medical personnel have also restricted access to family planning services.

Political and cultural barriers have limited access to family planning, especially for young people. In some countries, unmarried adolescents are denied access to family planning services on the assumption that such access would promote promiscuity. However, approximately 40% of girls in less-developed countries give birth before the age of 20 years. The pace of fertility decline in Africa, South Asia, and other high-fertility regions will be affected by whether young couples delay their first birth until they are in their 20s. This delay lengthens the interval between generations and lowers the average fertility. Health analysts estimate that women aged 15–19 years face twice the risk of dying from pregnancy and childbirth as do women in their 20s. In many countries, children born to mothers aged younger than 20 years are 1.5 times more likely to die before their first birthdays than are children born to mothers in their 20s.

Most less-developed countries provide family planning services. In many countries, family planning methods are also widely available from pharmacies and private health clinics. Not all women have easy access to family planning, but the expansion in the choices of methods and availability of services throughout the world during the last 50 years has been truly revolutionary.

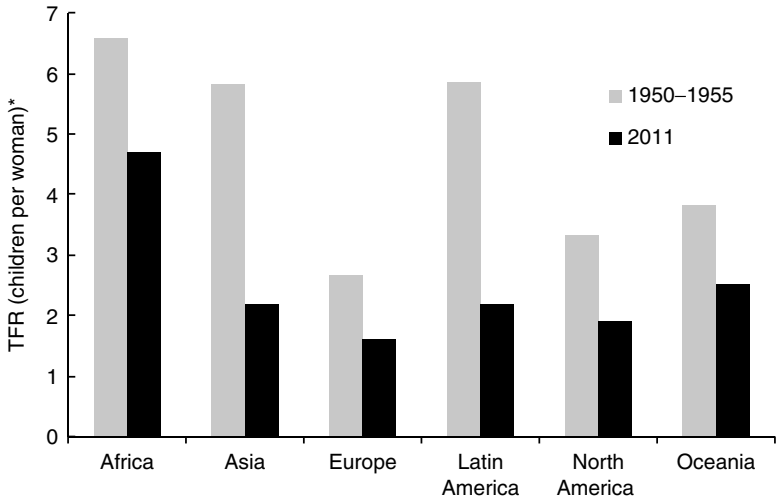


Figure 3 Fertility decline in world regions, 1950–2011. *TFR is the average number of children a woman will have under prevailing birth rates.

Causes and Effects of Human Population Growth

The demographic processes of fertility and mortality are influenced by biological, cultural, economic, geographic, political, and social factors. These factors affect demographic processes directly and indirectly through a web of interdependent variables. Cultural traditions that encourage girls to marry at a young age, for example, can contribute to high fertility rates because women will spend more years exposed to the risk of becoming pregnant. Early marriage can also lead to higher mortality because health risks to the infant and mother are greater when childbearing starts in adolescence.

In the 1980s, demographer John Bongaarts identified four variables that account for most differences in fertility rates. These "proximate determinants" of fertility are (1) the proportion of women married or in a sexual union; (2) the percentage of women using contraception; (3) the proportion of women who cannot conceive a pregnancy, especially during the infertile period following childbirth (postpartum infecundity); and (4) the level of abortion.

The importance of each proximate determinant depends on cultural, economic, health, and social factors within a population. The proportion of women in a sexual union is partly determined, for example, by the age at marriage, the proportion of women who never marry, and levels of divorce. Cultural mores about sexual activity and childbearing outside marriage also play a role.

In societies where women marry young, and in which nearly all childbearing takes place within marriage, changes in the age at marriage can significantly affect fertility. For example, in the Arab countries of the Middle East, an increase in the average marriage age for women led to significant fertility declines in some countries.

The length of postpartum infecundity usually depends on how long women breast-feed their babies. Breast-feeding releases hormones in the nursing mother that can prevent her from becoming pregnant. Postpartum infecundity is not a significant factor in such countries as the USA, in which women usually breast-feed their babies only for a few months, but it is important in sub-Saharan Africa and other traditional societies in which women commonly breast-feed their babies for 2 years. In most populations, contraceptive use and abortion are the primary determinants of fertility levels.

Education and poverty are among the most important influences on the proximate determinants and consequently have a strong indirect effect on fertility. Low levels of education and poverty are related, and they are also related to health and to levels of economic development, urbanization, and environmental conditions.

Education

Although researchers cannot determine all the reasons, education is associated with lower fertility and mortality. A formal education may act as a catalyst for changes in values and behavior. Education may make people more receptive to new ideas, such as family planning. This has been reflected in a recent review, which reports that almost universally, women with a higher level of education have fewer children and that better education is associated with lower mortality (Lutz and Samir, 2011). Social scientists note that education does not have the same effect in all cultural settings and that many other factors (such as women's status) may explain much of the association.

More educated women have higher rates of family planning use, smaller families, and healthier children than other women in the same society. Where educational levels are high, women are likely to postpone marriage until they finish their secondary schooling or college.

Married women are more likely to use family planning if they have some formal education. A 2008 Demographic and Health Survey in the Philippines indicated a contraceptive use rate of 53.2% for married women of reproductive age who had at least some secondary education. Only 18.5% of their counterparts with no formal education used a contraceptive method. The difference persists in many countries.

In most societies, total family size declines as education increases. In the mid-2000s, Honduran women aged 40–49 years

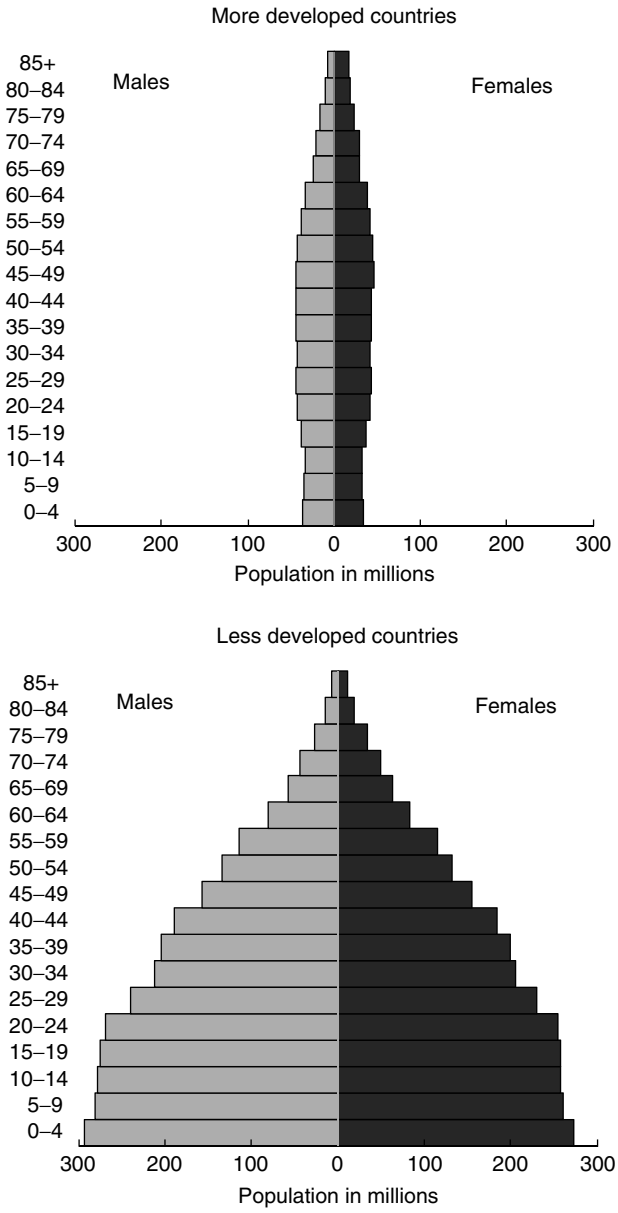


Figure 4 Population by age and sex in more-developed and less-developed countries, 2010.

with at least some secondary education had on average nearly four fewer children than women with no formal education. Similar gaps were recorded around the same time in Haiti, Nicaragua, and Guatemala and slightly less in Senegal, West Africa. Senegalese women with a secondary or higher education had 3.8 children on average, whereas women with no education had an average of 6.9 children.

Education usually expands employment options, and educated women may delay marriage and childbearing to earn income. Also, schooling may introduce young women to new ideas or values that could influence the number of children they want and also their use of family planning.

Women's education is also associated with better child health. Children of mothers with some education have fewer risk factors for infant mortality. Infants are at a higher risk of dying if they are born to adolescents or to mothers older than 40 years, if they are born into large families, or if they are born less than 2 years after an older sibling.

By delaying marriage and childbearing, education reduces high-risk births to teenage mothers. In Colombia, for example, 55% of women aged 19 years with no formal education had a baby or were pregnant at the time of interview, whereas only 17% of those who had at least some secondary education had a baby or were pregnant. The gap was less pronounced in Bangladesh and the Philippines, as shown in **Figure 5**, but greater in Nicaragua and Angola in the mid-2000s. Women who have completed some formal education tend to wait longer between pregnancies and births and to stop childbearing at a younger age than do less-educated women. Consequently, they have smaller families and fewer births after the age of 40 years.

In most societies, children of mothers with some education have a lower risk of dying than children of uneducated mothers. As shown in **Figure 6**, in Cambodia, the IMR was 111 for the children of uneducated mothers, whereas it was 45 for children of women with a secondary or higher education. The twentieth century brought enormous improvements in literacy and educational levels. The recent improvements in literacy rates and educational attainments reflect the expansion of educational services throughout the world. The United Nations Educational, Scientific, and Cultural Organization (UNESCO) reported in Education for All – Global Monitoring Report, 2011 that 83% of people aged older than 15 years were literate in a year of survey or census between 2005 and 2008 compared with only 56% in the 1950s. Basic literacy is nearly universal among populations of Europe, North America, and other industrialized regions, but the range is substantial throughout the rest of the world. In this period, an estimated 62% of the populations of South and West Asia and sub-Saharan Africa were literate, as were 72% of the populations of the Middle Eastern Arab states. More than 91% of the population are literate in Latin America and the Caribbean and about 94% in East Asia.

Rapid population growth in some countries is undermining improvements in educational attainment. For example, in the sub-Saharan African countries of Angola, Benin, and Togo, economic difficulties and burgeoning numbers of young people caused school enrollment ratios to level or decrease in the 1980s and 1990s; however, almost all countries have seen increase in the school enrollment ratios in the 2000s. In the year 1999, about 54% of girls and 62% of boys of primary school age in sub-Saharan Africa were enrolled in primary school, according to UNESCO estimates, which has increased to 74% and 78%, respectively in the year 2008. The effect of low level of school enrollment ratios will have long-term effect on the human capital of the population as measured by educational attainment of the working-age population and can only be observed in the future. KC *et al.* (2010) have shown the future based on different educational scenarios in 120 countries of the world.

Economic Development and Environment

In most societies, poor families have higher mortality and fertility than affluent families. Some of the association

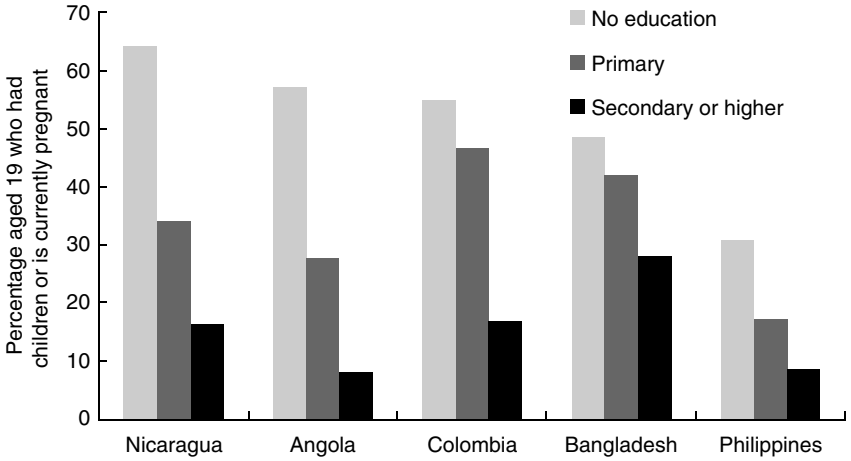


Figure 5 Mother's education and teenage childbearing in selected countries, mid-2000s. *Percentage of women aged 19 years who had a child or were pregnant by the age of 19 years.

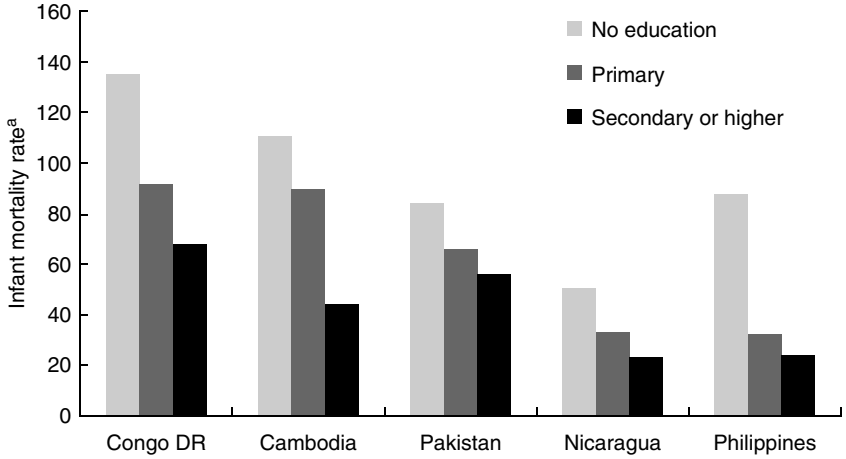


Figure 6 Mother's education and infant mortality in selected countries, mid-2000s. [b]The infant mortality rate is the number of deaths to children less than 1 year of age per 1000 live births in a given year.

between poverty and population reflects the lower educational levels and rural residence of poor households. However, the relationship among demographic variables, poverty, and affluence is highly complex, and it is tied to the broader question of how population size and the pace of population growth are linked to economic development. The issue is further complicated by questions about whether economic growth and human activity are causing irreversible damage to the natural environment.

The research into these questions has yielded contradictory results. The extremes of these differences are characterized by two opposing camps: "pessimists" and "cornucopians". The theoretical foundation of the pessimist view can be found in writings published in 1798 by the economist Thomas Malthus. He suggested that the potential population size is limited by the amount of cropland, and therefore food, available for human consumption. Malthus assumed (based on his observations of eighteenth-century English society) that if population growth continued unchecked, population would outstrip the food available and cause widespread famine and death. He also described a natural feedback mechanism: When the population grew too large for the available food supply, elevated mortality would reduce the population to the level that could be sustained by the amount of food produced.

A neo-Malthusian view of the relationship between population, economic growth, and resources gained credence between the 1940s and the 1960s, a period of unprecedented population growth and economic development. In a landmark study in the 1950s, Ansley Coale and Edgar Hoover found that the population growth slowed economic development and held down *per capita* incomes. These researchers also assumed that the supply of some natural resources and capital was fixed or that the supply would grow more slowly than population. Other researchers during this period expanded the idea that rapid population growth will eventually reach some absolute limit on resources and that population growth accompanied by the environmental stresses associated with economic development could cause irreversible damage to the basic natural systems that sustain life. Such concerns

were popularized by books such as "The Population Bomb" by Paul Ehrlich, "The Limits to Growth" by Dennis Meadows, and also by other authors.

The ideological basis of the "cornucopian" approach owes much to the writings of Ester Boserup in the 1960s and 1970s, who argued that the need for more food coupled with the synergy created by the concentration of intellects and flow of ideas in dense settlements can stimulate the adoption of better farming techniques or sharing of higher yield plant varieties. Economist Julian Simon, in "The Ultimate Resource" and other writings, also rejected the idea that population growth was a threat to the welfare of humans or the environment.

The scientific evidence concerning the effects of population size and growth on economic development was still inconclusive in the early 1990s at the national and regional level, but it was less ambiguous at the individual level. An extensive review of research completed in 1994 by the Overseas Development Council observed that "the clearest evidence of negative effects of population growth under high fertility is at the individual and household levels."

In the late 1990s, several studies provided a clearer picture of the relationship between population and development at the national level and the links between poverty and demographic factors at the household level. Several studies suggested that a rapid transition from high to low fertility contributed to the economic miracles in South Korea and other East Asian countries. The rapid fertility decline increased the share of working-age people in the population relative to young and old dependents, which created a "demographic bonus." The working-age population adds more to the economy than it consumes in services and generates taxes and savings that can be invested in education and further economic growth. This demographic bonus may last several decades; it recedes as the bulge of working-age men and women reach retirement age and the proportion of the dependent population relative to the working-age group population increases again.

Research shows that countries can benefit from this bonus only if they increase the value of their human capital – especially the youth entering the labor force – through education, and if governments adopt policies favoring international trade and industrialization. The newly industrializing Asian countries capitalized on their demographic bonus by making these investments. As researcher Andrew Mason noted, these Asian countries "raised millions of people from abject poverty and transformed some of the poorest economies in the world to some of the richest." Research on the relationship between population change, economic development, and environmental systems remains much less clear, plagued by the complexity of the relationships and the difficulty of measuring such factors as environmental quality. At a symposium on population change and economic development in 1998, scientists assigned highest priority to research on population– development–environment linkages.

Population Prospects: 2010–2050

In the last century, the world's population has undergone a sweeping change in both its total numbers and its distribution across regions. The twenty-first century is likely to see the second phase of this transformation – lower fertility and an even more dramatic redistribution of population among the more-developed and less-developed countries. Nearly all future world population growth will take place in less-developed countries. In other words, the Earth is reinventing itself demographically.

Because mortality is relatively low, fertility levels and trends will determine future population size. A common issue (and a common assumption) is when, or whether, a country will reach the "magic" replacement-level TFR of about 2.1 children per woman. With fertility at replacement level, a population eventually will cease growing and "stabilize" at a given size.

Every 2 years, the UN Population Division produces a set of population projections for every country. By 2050, the UN suggests that the total world population will increase to between 8.1 and 10.6 billion. In the high- and medium-projection series, world population will still be increasing in 2050; under the low-projection series, it will have begun a gradual decline. All the three projections assume continuing declines in fertility. They differ in terms of fertility levels attained in 2050, from 1.71 for the low projection to 2.64 for the high, illustrating the major differences in total population size that can result from seemingly very small differences in childbearing during the next 40 years.

Regardless of the projection used, the UN indicates that at least 1.1 billion people will be added to the world's population during the next 25 years. There are three reasons for this inevitable growth. First, fertility in less-developed countries is on average 60% higher than in the more-developed countries. Second, the young age structure of less-developed countries constitutes momentum for population growth for several decades, no matter what future fertility trends may be. Third, continuing improvements in mortality will contribute to additional growth, particularly in countries in which life expectancy remains comparatively low.

Recently, in addition to the projection of population by age and sex, efforts were made at the International Institute for Applied Systems Analysis on the reconstruction and projection of the population by age, sex, and educational attainment in many countries of the world for the period of 1970–2050 (Lutz *et al.*, 2007 and Samir *et al.*, 2010). The projections show that everything remaining the same and applying different education scenarios in the future, population in the year 2050 will be between 8.8 and 9.9 billion (review in Lutz and Samir, 2011).

Responses to Human Population Growth

Anxiety about the negative effects of rapid population growth and excessive population numbers has a long history. Long before Malthus, ancient Greeks and Egyptians had voiced concern about "overpopulation" in lean times. They also promoted population growth in times of plenty.

In the 1930s and 1940s, scientists and intellectuals in some less-developed countries, such as Egypt, India, and Mexico, began to express concern that rapid population growth would hinder development in their countries. Widely publicized

food shortages and famines in certain less-developed areas in the 1960s were also linked to rapid population growth.

These concerns sparked many actions throughout the world directed at lowering fertility and slowing population growth. India initiated a national policy to slow population growth in 1952. The International Planned Parenthood Federation, the largest private sector organization devoted to family planning, was founded in the same year. UN involvement in population issues also increased. The first UN meeting on global population was convened in 1954, in collaboration with the International Union for the Scientific Study of Population. UN agencies, including UNICEF and the World Health Organization, incorporated reproductive health services into their missions. In 1969, the UN Fund for Population Activities became a separate entity. Beginning in the 1960s, governments of some wealthier countries, most notably the USA, supported efforts to strengthen family planning programs in less-developed countries.

Population Policies

The idea that couples should limit their family size went against cultural mores in many societies, and some governments were loathe to support a potentially unpopular policy. Many governments embraced the more acceptable idea that fertility would decrease and that population growth would slow as living standards increased through economic development. This view was expressed at the 1974 UN World Population Conference when an Indian delegate declared that "development is the best contraceptive."

During the late 1970s and 1980s, concern about the negative effects of population growth on economic development broadened. Increasing numbers of countries accepted the idea that government actions could slow population growth.

National efforts to influence population growth include incentives to have more or fewer children and disincentives for having more than a given number of children. These efforts have met with mixed success. Some argue that China's population policies initiated in the 1970s were a success from a demographic perspective. China's TFR decreased from approximately 6.0 in the 1960s to less than 2.0 in the 1990s, in part because of government policies and programs. However, China's stringent "one-child family" policy introduced in 1979 was widely criticized for violating human rights. Between 1975 and 1977, Indira Gandhi's government in India promoted male sterilization campaigns that sometimes led to coercion. Public outrage about the reported abuses contributed to the downfall of Gandhi's government and created a backlash against family planning programs in India that took years to overcome.

Women's rights activists and others have generally opposed a demographic rationale for family planning as an infringement on individual rights. They have argued that women's rights and well-being should take precedence over national interests. Many have criticized the family planning programs' lack of integration with other health services.

During the 1970s and 1980s, women throughout the world began forming small nongovernmental organizations (NGOs) to lobby for improvements in their social, economic, and political circumstances. By the 1990s and in 2000s, women's NGOs in less-developed countries were advocating improvements in family planning programs by better informing clients about various contraceptive methods, expanding the range of methods available, and encouraging service providers to treat clients with greater respect.

By 1997, 155 countries subsidized family planning services, and 68 stated explicitly that they wanted to slow their population growth. In Africa, the world's fastest growing region, 40 countries viewed their fertility levels as too high and 36 had policies to lower fertility. A few countries, in contrast, viewed their fertility rates as too low and welcomed faster population growth. In 1997, 23 countries reported to the UN that they had explicit policies to increase birth rates. Many governments in Europe and the former Soviet Union are concerned that continued low levels of fertility will cause rapid population aging and an eventual decline in population size. Some small oil-rich countries in the Persian Gulf also want to increase, or at least maintain, current levels of population growth. They view population growth as a way to spur socioeconomic development and reduce their reliance on foreign labor.

A New Vision: The 1994 International Conference on Population and Development

In 1994, the world community redefined the world's view of population growth and the best way to address this growth. The Program of Action adopted at the International Conference on Population and Development (ICPD) in Cairo, Egypt, agreed to the need to stabilize global population growth. One hundred and eighty countries agreed to a strategy placing population within the context of sustainable development and calling for investments in human development, especially improvements in women's status, as key to stabilizing population growth. It rejected the use of demographic targets by family planning programs and it integrated family planning into a broader women's health agenda.

Participation in the meeting by NGOs was a critical factor in achieving such widespread consensus. Twelve hundred NGOs participated as delegates and observers; and for the first time, conference deliberations on global population were informed by a wide range of interests, from the grassroots to the highest levels of government. Women's groups were a driving force behind the strong emphasis on women's empowerment as part of human development. However, this focus was also driven by research from the last 30 years that linked fertility declines with reductions in infant mortality, increased use of family planning, and improvements in women's education and other aspects of women's status contributing to their empowerment.

Despite the consensus, the ICPD engendered dissent and debate. Ideological and religious tensions characterized discussions before the conference, deliberations during the conference, and follow-up after the conference. Abortion generated the most highly publicized ideological differences. There was also a debate regarding definitions of reproductive health and family and adolescent reproductive rights and responsibilities. None of the 179 nations rejected the central

premises and goals of the ICPD, despite the range of political structures, cultures, and religions they represented. This marked the first time in the history of UN conferences that no official delegation rejected the entire document.

The final ICPD document defined reproductive health to encompass a broad range of services, including family planning, prenatal and postnatal care, medical attention at birth, cancer screening, and protection from sexually transmitted diseases. It also supported access to safe abortion where it is legal, but it stated that abortion should not be used as a method of family planning.

The ICPD Program of Action specified five goals for 2015 to improve individual and family well-being and enhance women's status, including universal access to family planning and other reproductive health services; universal access to primary school education; increased access by girls and women to secondary and higher education; and reductions in infant, child, and maternal mortality. The ICPD document also called for government and private sector actions to alleviate poverty, protect the environment, encourage greater male involvement in the family, and to address the specific health and education needs of adolescents.

A unique feature of the agreement was that it called for specific financial commitments to achieve these goals. Countries agreed that less-developed countries would pay for two-thirds of the costs of reproductive health programs, including family planning, and that more-developed countries would pay for the remaining costs, estimated to be $17 billion per year by the year 2000. The historic agreements reached at the ICPD were reaffirmed at subsequent UN conferences in the 1990s, including the World Summit for Social Development, in Copenhagen in 1995; the Fourth World Conference on Women, in Beijing in 1995; the UN Conference on Human Settlements (or Habitat II), in Ankara in 1996; and the World Food Summit, in Rome in 1996. In 2004 and 2009, meetings were organized between the United Nations Population Fund (UNFPA), governments, and development partners to assess the progress and to identify gaps and challenges. In the 2004 meeting, it was reported by many that the progress was not as expected in certain regions of the World mainly due to the underfunding, destabilization caused by warfare and civil unrest, and the ravages of HIV/AIDS epidemic (Germaine and Kidwell, 2005). However, all stakeholders reaffirmed the agreement to continue to achieve the set targets.

In addition to ICPD's declaration, Millennium Development Goals (MDGs) are an important initiative that will effect the future development of population. In 2000, 189 countries of the world came together to set eight MDGs with a target to free people from extreme poverty and multiple deprivations by year 2015. Out of the eight goals, three specific goals (reduce child mortality; improve maternal health; and combat HIV/AIDS, malaria, and other diseases) are directly related to future population development and the rest (eradicate extreme poverty and hunger, achieve universal primary, promote gender equality and empower women, ensure environmental sustainability, and develop a global partnership for development) have an indirect effect on the future development of the population. So far, some countries have shown progress toward the goals and few countries are unlikely to reach the targets by 2015.

Conclusion

Population change has been one of the most significant events of the twentieth century. Since 1900, the world population has almost quadrupled in size and average life expectancy has increased by more than two-thirds. Declines in childbearing and shifts in population distribution were more striking than at any other time in history. Along with these population changes, the world has witnessed extraordinary improvements in technology, communication, education, and agriculture. These changes have undermined the dire predictions of Thomas Malthus and his successors that population growth would lead to worldwide famine and disease. However, these predictions may come true for some areas of the world. According to the World Bank, approximately one-half of the population of sub-Saharan Africa and two-fifth of the population of South Asia lives in poverty, subsisting on less than $1.25 a day. The HIV/AIDS pandemic threatens the health and well-being – and the very survival – of large portions of the population in many countries.

Under all likely scenarios, the twenty-first century will see continued population increases, at least during the first few decades; that was confirmed by the increases seen in the 2000s. This is because of the built-in momentum of growth associated with the very young age structures of most less-developed countries. The growth will also be fueled by childbearing levels that are still higher than replacement levels in most areas of the world. Not all countries will experience this growth, but they will all be affected by it.

At the end of the twentieth century, the world community made financial and program commitments to continue investments in family planning and other health programs, in education, and in greater social and economic opportunities, especially for women, arguing that these are key to future population stabilization.

See also: Energy Use, Human. Human Impacts on Ecosystems: An Overview

References

Ashford L and Makinson C (1999) *Reproductive Health in Policy and Practice: Case Studies from Brazil, India, Morocco, and Uganda.* Washington, DC: Population Reference Bureau.

Baudot BS and Moomaw WR (eds.) (1997) *The Population, Environment, Security Equation.* New York, NY: Macmillan.

Birdsall N and Sinding S (1999) *Chairman's Report on the Symposium on Population Change and Economic Development*, November 2–6, Bellagio, Italy Washington DC and New York: Carnegie Endowment for International Peace and Rockefeller Foundation.

Bledsoe CH, Casterline JB, Johnson-Kuhn JA, and Haaga JG (eds.) (1999) *Critical Perspectives on Schooling and Fertility.* Washington, DC: National Academy Press.

Castles S and Miller MJ (1998) *The Age of Migration*, 2nd ed. New York, NY: Guilford.

Cohen JE (1995) *How Many People Can the Earth Support?* New York, NY: Norton.

Cohen MA, Ruble BA, Tulchin JS, and Garland AM (1996) *Preparing for the Urban Future, Global Pressures and Local Forces.* Washington, DC: The Woodrow Wilson Center Press.

Gelbard A, Haub C, and Kent MM (1999) World population beyond six billion. *Population Bulletin* 54. March.
Germaine A and Kidwell J (2005) The unfinished agenda for reproductive health: Priorities for next 10 years. *International Family Planning Perspective* 31(2): 90–93.
Haub C and Cornelius D (1998) *1998 World Population Data Sheet*. Washington, DC: Population Reference Bureau.
Jain A (ed.) (1998) *Do Population Policies Matter?* New York, NY: Population Council.
Samir KC, Barakat B, Goujon A, Skirbekk V, Sanderson W, and Lutz W (2010) Projection of populations by level of educational attainment, age and sex for 120 countries for 2005–2050. *Demographic Research* 22: 383–472.
Livi-Bacci M (1992) *A Concise History of World Population*. Cambridge, MA: Blackwell.
Lutz W (ed.) (1996) *The Future Population of the World. What Can We Assume Today?*. Laxenburg, Austria: International Institute for Applied Systems Analysis.
Lutz W, Goujon A, KC Samir, and Sanderson W (2007) Reconstruction of population by age, sex and level of educational attainment of 120 countries for 1970-2000. *Vienna Yearbook of Population Research* 2007: 193–235.
Lutz W and KC Samir (2011) Global human capital: Integrating education and population. *Science* 333: 587.
Malthus TR (1960) In: Himmelfarb G (ed.) *An Essay on the Principle of Population as It Affects the Future Improvement of Society. With Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers*. New York, NY: Modern Library.
Population Reference Bureau (1997) *Improving Reproductive Health in Developing Countries*. Washington, DC: Population Reference Bureau.
Population Reference Bureau (2011) *World Population Data Sheet*. Washington, DC: Population Reference Bureau.
Sen G, Germain A, and Chen LC (eds.) (1994) *Population Policies Reconsidered: Health, Empowerment, and Rights*. Cambridge, MA: Harvard University Press.
United Nations (1996) *Levels and Trends of Contraceptive Use as Assessed in 1994*. New York, NY: United Nations.
United Nations (1998) *World Population Prospects: The 1998 Revision,* vol. 1. New York, NY: United Nations.
United Nations (2011a) *World Contraceptive Use 2011*. New York: United Nations.
United Nations (2011b) *World Population Prospects: The 2010 Revision*. New York, NY: United Nations.
UNAIDS (2010) *Report on the Global AIDS Epidemic*. Geneva: UNAIDS.
UNESCO (2011) *Education for All Global Monitoring Report 2011*. Paris: UNESCO.
UNDP, *The Millennium Development Goals: Eight Goals for 2015*. New York: United Nations Development program. http://www.beta.undp.org/content/undp/en/home/mdgoverview.html
The World Bank (1994) *Averting the Old Age Crisis*. New York, NY: Oxford University Press.

Population Viability Analysis

Hugh P Possingham, The University of Queensland, Queensland, Australia
Michael A McCarthy, University of Melbourne, Melbourne, Australia
David B Lindenmayer, The Australian National University, Canberra, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Stochasticity A stochastic process is one in which the state of the system cannot be precisely predicted given its current state, even with a full knowledge of all the factors affecting that process. In a population context, we may know the detailed life history parameters of a species. However, various unpredictable (stochastic) processes, such as the chance nature of birth and death (demographic stochasticity), year-to-year variation in climate (environmental stochasticity), and catastrophes, means that accurately predicting the precise size of a population in the future is not possible. Despite this, we can use probability theory and simulation models to make probabilistic predictions. A stochastic population model is one in which each possible future population size has an associated probability. This approach to prediction is the same as stating that the chance of getting a head with the next toss of a fair coin is 50%. While our prediction is accurate, we cannot say if the outcome will be a head or a tail.

Viability (and related concepts) A viable population is one deemed to have a reasonable chance of long-term persistence. Such a definition begs two questions: What is reasonable and what is long-term? It has been stated that a viable population is one that has a 99% chance of persisting 1000 years. Different managers, researchers, and policymakers use different measures that range from a 99% chance of persisting 1000 years to a 95% chance of persisting 100 years. The population size that just meets the definition of being viable is called a minimum viable population (MVP). One of the key questions of reserve design is how big does a reserve, or network of reserves, need to be to contain an MVP. This size is called a minimum viable habitat area for a particular species.

What is Population Viability Analysis (PVA)?

Traditionally, PVA is a process in which a stochastic population model is used to assess the viability of a population. The population may be a species, subspecies, metapopulation, or an isolated subpopulation of a single species. Classical PVA uses a demographic population model to derive a result such as: "given current circumstances the probability of extinction of species *X* over the next *Y* years is *Z*%." For example, in his seminal 1981 paper, Shaffer used a demographic population model to determine the probability of extinction of grizzly bears, *Ursus arctos*, in Yellowstone National Park within the next 100 years (Shaffer, 1981).

Most PVAs use a demographic population simulation model that incorporates all the processes likely to affect the dynamics of the population. PVA differs from other population modeling exercises in its focus on predicting the probability of extinction and the inclusion, typically, of stochastic processes that are believed to have a strong impact on the probability of extinction. For convenience, the authors categorize these stochastic processes into four types; each can be thought of as a way in which variability influences the dynamics of a species. Given that these processes are central to our thinking about PVA, they are described in detail. In a broader context PVA can be thought of as just one branch of the larger field of population modeling, but with a primary focus on an assessment of risk. Risk is the probability of an unfavorable event; in this case the unfavorable event is the extinction of a species or a population falling below a threshold level.

PVA is unique in the context of conservation biology because it assesses the adequacy of conservation efforts. Given an acceptable risk of extinction for any species, then, in theory we should be able to manage every species so that its extinction risk does not exceed this acceptable level. If we can do this, our conservation strategy would be deemed "adequate."

During the past decade, the authors have realized that the simple view of PVA as a process that uses a demographic simulation model to assess extinction risk is too narrow. First, the restriction of PVA to the use of demographic population modeling software ignores other tools for assessing viability. We now view PVA as a broader process that includes various sorts of theory, analytic and computer modeling, and forms of data analysis (Boyce, 1992; Beissinger and Westphal, 1998). Second, despite our inability to assess accurately the extinction probability of most populations, PVA can be used to make robust management decisions. The broader roles of PVA are explored later. It is this management-based approach to PVA that is now broadly accepted and considered practically useful.

Stochastic Population Processes and Extinction

Types of Demographic Population Fluctuation

Demographic Stochasticity

Demographic stochasticity is the chance nature of birth and death. It causes populations to fluctuate because populations are composed of individuals that are units (Kendall and Fox, 2003). Each unit counts as one and is born and dies as a unit, so populations can only increase and decrease on the set of integers. Each individual has a probability of dying and a probability, each year, of finding a mate and producing

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00173-8

offspring. The dynamics of five separate populations, fluctuating as a result of demographic stochasticity alone, are shown in **Figure 1**.

Each population has its own trajectory even though they all have the same life history parameters and started from the same population size. In a sense, demographic stochasticity is the most fundamental phenomenon that causes populations to fluctuate because it occurs in all populations. In **Figure 1** each population should, on average, increase 5% per year, but one becomes extinct. The importance of demographic stochasticity for extinction and PVA diminishes in larger populations because the minor "wobbles" caused by the unitary nature of birth and death are overshadowed by the more dramatic environmental stochasticity and catastrophes (**Figure 2**).

A rarely discussed type of demographic stochasticity is the chance nature of sexual determination in organisms with more than one sex. For example, a population may be known to have six individuals, which may appear to have some chance of persisting if conditions are good. However, there is a 1 in 32 chance that all the individuals are the same sex and the population is doomed if each has an equal chance of being male or female. When populations are small, a skewed sex ratio can have a major impact on the chance of extinction.

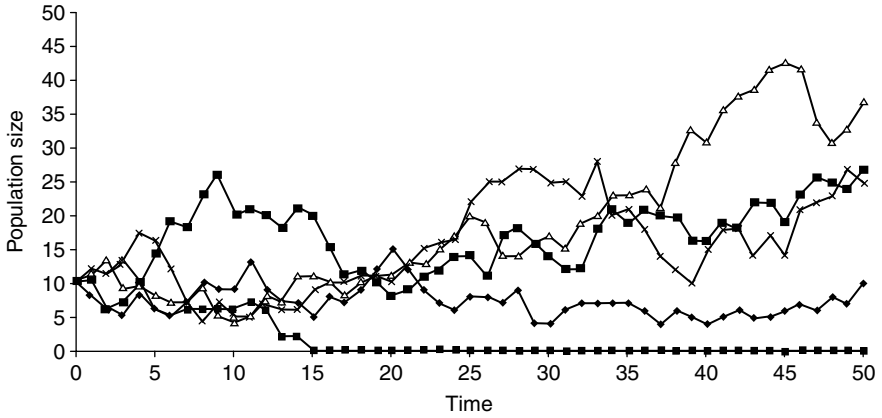


Figure 1 Five sample trajectories for a population in which demographic stochasticity is operating. For each individual, each year there is a 30% chance of dying. Those that survive have a 50% chance of giving birth to one more individual. Despite identical conditions the populations diverge rapidly, with one becoming extinct after 15 years and another increasing to more than 40 individuals. With these parameters, in the absence of stochasticity, the population should on average increase at a rate of 5% (0.7×1.5) per year.

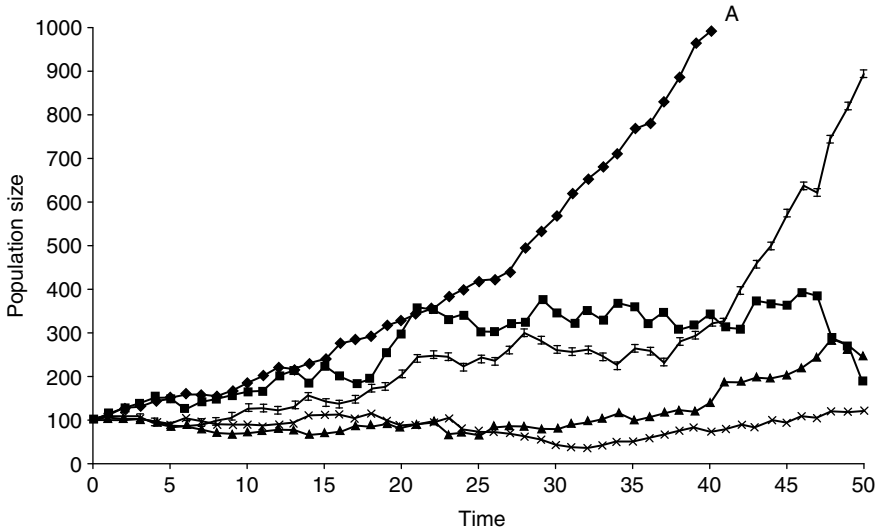


Figure 2 The impact of environmental stochasticity on the dynamics of five populations. For trajectory A, the parameters are the same as those used for **Figure 1** with just demographic stochasticity, except the population started with 100 animals. Note the fairly consistent increase of about 5% per year with minor fluctuations caused by demographic stochasticity. For the other four populations, the death rate is not fixed at 30% but varies randomly between 20 and 40% from year to year. Also, the probability of producing an offspring is not fixed at 50% but varies from 40 to 60% from year to year for the whole population. Note the much wider fluctuations in population size, even when the populations are larger than 100. Also note the differences between the trajectories and dynamics even though the parameters are the same on average – the consequence of stochasticity.

Environmental Stochasticity

As biotic and abiotic conditions vary from year to year, there will be variation from year to year in vital rates. This sort of variation tends to affect the entire population; therefore, there are years in which breeding success may be good or bad. Because this variation tends to operate on the entire population, it causes more dramatic fluctuations in population size (**Figure 2**).

Catastrophes

Some natural events cause precipitous declines in the sizes of populations, such as fire, flood, disease, and frost. Although these events can be seen as an extreme form of environmental variation, they are often considered separately as catastrophes. The authors define a catastrophe as being a single event that causes the death of a substantial fraction of a population. Recent analysis implicated catastrophes in the extinction, or near extinction, of many species. The Puerto Rican parrot was almost wiped out by a hurricane, the Guadeloupe oriole by a volcanic eruption, black-footed ferrets by canine distemper, and a population of heath hen by fire. Species with populations that are spatially restricted, such as those on islands, are particularly vulnerable to catastrophes. However, disease appears to be able to decimate continental populations over huge areas. Disease is implicated as a factor contributing to declines of once common species, such as the passenger pigeon and many Australian carnivorous marsupials such as the Tasmanian tiger. **Figure 3** shows the trajectory of four populations for which there is a 2% annual probability of 90% of the individuals being killed in a single catastrophe.

Including catastrophes is an important part of any PVA (Mangel and Tier, 1993) but a difficult component to include in models because of limited data on the frequency and impact of catastrophes. Although we often have good data on catastrophes such as fire, and these have been successfully included in models of Australian forest fauna (Lindenmayer and Possingham, 1996), we rarely have good data on catastrophes such as disease that are more difficult to detect but probably just as important.

Genetic Stochastic Processes

Genetic stochastic processes change the frequencies of genes within a population. When a population is small and isolated, genetic variation is typically lost from generation to generation. This process is called genetic drift. As frequencies of alleles change at random with each successive generation, genetic variability is reduced because rarer alleles are lost by chance through the stochastic nature of birth and death. Inbreeding is defined as mating between individuals that are related by ancestry and is more likely in populations that are, or have been, small. Inbred animals can have reduced survival, fecundity and growth (inbreeding depression). Inbreeding expression can be expressed in particular environmental conditions such as harsh winters. Inbreeding depression results in the selective removal of inbred animals and the genes carried by such animals (Lacy, 1993). Morphological defects in threatened species such as the Florida panther (*Felis concolor coryi*) and lack of breeding success in the Puerto Rican parrot (*Amazona vittata*) appear to be a result of inbreeding depression. The restoration of genetic variability is slow and believed to be between 1 and 10×10^{-6} genes per generation per 100 years. Thus, genetic mutations have only a minor effect in reversing the loss of genetic variation during the time frame typical for PVA. A longer term consequence of loss of genetic variability is a reduced ability of a population to adapt to environmental changes.

Overview of the Role of Stochasticity on Time to Extinction

Each kind of stochasticity has quite distinct effects on the relationship between extinction probability and population size (or carrying capacity). Analytic and simulation studies have shown that although the mean time to extinction, and hence the viability of a population, grows nonlinearly with

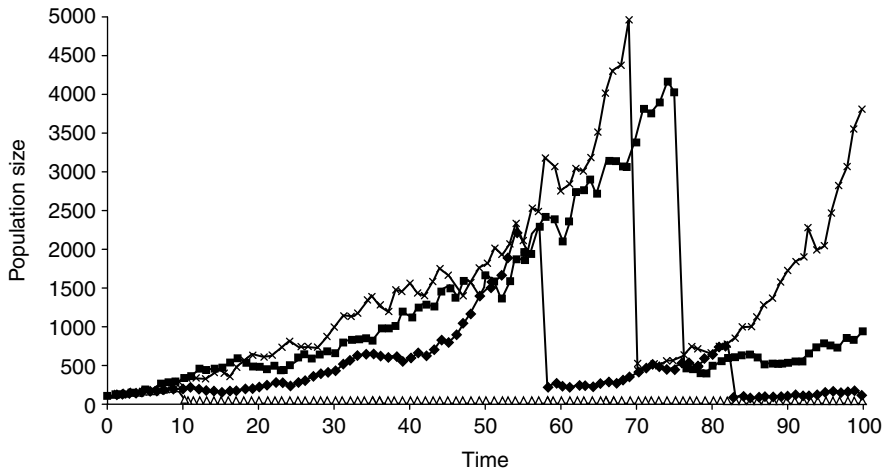


Figure 3 The impact of catastrophes on the dynamics of four populations. The same parameters as in **Figure 2** are used and include both demographic and environmental stochasticity. However, there is a 2% chance per year that 90% of the population dies. These catastrophes cause enormous fluctuations in the population sizes that can take very healthy large populations to the brink of extinction whereupon demographic and environmental stochasticity are sufficient to cause extinction.

population carrying capacity when only demographic stochasticity is operating, the other two forms of stochasticity can make even large single populations quite vulnerable to extinction (**Figure 4**). With only demographic stochasticity the mean time to extinction increases very rapidly above species-specific thresholds (usually approximately 50–100 individuals). Therefore, in the absence of environmental variability, catastrophes, and genetic stochastic processes, quite small populations could persist for very long times. Adding environmental variation means that the benefits of increasing population size are diminished and there are no simple thresholds above which populations are likely to be completely secure. Because catastrophes can reduce a population to a small size very quickly, a larger population size brings diminishing returns in terms of viability (Lande, 1993). When this occurs, the best way to improve population viability may be to ensure that a single catastrophe cannot affect the whole population – a risk-spreading approach. This issue will be discussed later.

Deterministic Extinction Processes and Extinction Vortices

A major contribution of PVA to conservation biology is that it focuses attention on stochastic extinction processes in small populations. This should not distract our attention from equally important, more deterministic processes that lead to extinction (Caughley, 1994). Habitat loss and associated fragmentation, increases in mortality rates from introduced predators, and reduced reproduction through habitat modification all have significant impacts on the carrying capacity and rates of increase of populations. Although the distinction between deterministic and stochastic extinction processes is useful conceptually, it is worth noting most important impacts on species have deterministic and stochastic components.

The deterministic and stochastic extinction forces need to be integrated. We can think of the interaction of some of these processes as extinction vortices, in which they nonadditively contribute to the demise of a species. For example, a catastrophe may take a population close to extinction, resulting in a period of inbreeding depression, which keeps the population sufficiently low so that demographic stochasticity finally causes extinction. Another example is the way in which habitat loss causes population fragmentation where each isolated population is small enough to fall victim to the interaction of genetic and demographic stochastic processes (Gilpin and Soule, 1986).

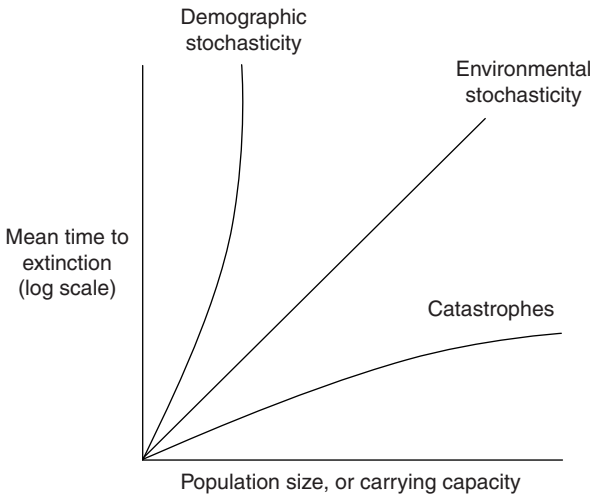


Figure 4 The three different types of population stochasticity have different impacts on the relationship between the mean time to extinction and the size (or carrying capacity) of a population.

Understanding the process of extinction is fundamental to our ability to model it. The obvious question arises: What use is an extinction model?

Why do Population Viability Analysis?

Applications of PVA

Originally, PVA had two general purposes. Shaffer's (1981) original intent was to assess the viability of a population. This assessment of extinction risk is useful in that it informs us of the effectiveness of conservation strategies. In the case of the grizzly bear in the Yellowstone region of the United States, the assessment implied that the species was unlikely to persist in the long term. In the following years, PVAs were tied very closely to the idea of a MVP. Scientists and managers used PVA to determine if the strategies for conserving a population were adequate. The goal was to use PVA to derive a target population that was large enough to be viable and then to recover the species to that level as quickly as possible.

When one of the primary tools for managing a threatened species includes setting aside an area of suitable habitat for that species, the idea of an MVP extends to that of a minimum viable habitat area (MVHA). An MVHA is a useful concept in which the main threatening process is habitat destruction. However, it has limited value when threats to a species are not simply removed by reservation, such as the continued activities of a predator (e.g., the western swamp tortoise in Western Australia). An MVHA can be predicted using a PVA by predicting extinction probabilities as a function of reserve area. Typically, the relationship between extinction probability and area conserved is sigmoidal (**Figure 5**). The MVHA is simply derived by choosing an acceptable level of extinction probability and then conserving the corresponding area. In some cases, the possibility that a single habitat patch ensures viability is unlikely because of catastrophic events, and PVA models that allow for multiple populations (metapopulations) are necessary to determine an adequate reservation strategy. Such a strategy depends on the number, size, and spatial distribution of reserves and the nature of the catastrophic events.

Another application of PVA that relies on the accurate assessment of extinction probability for a species is assembling lists of threatened species. Many countries and states are required, for legislative and resource allocation purposes, to maintain and update lists of threatened species. The International Union for the Conservation of Nature has explicit definitions of the different categories of threatened species: critically endangered, endangered, vulnerable, rare, insufficiently known, and extinct (Mace and Lande, 1991). These criteria are becoming increasingly complex and they vary from

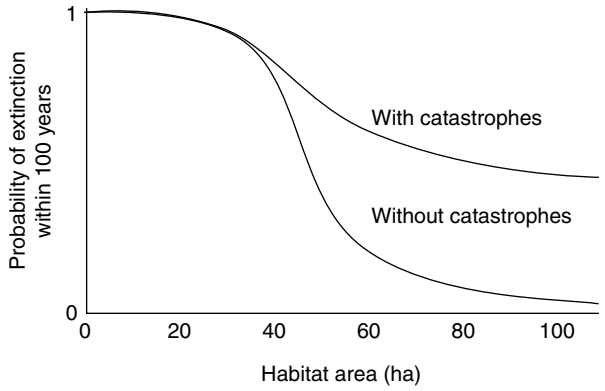


Figure 5 The relationship between old-growth patch size and extinction probability of Leadbeater's possum (Lindenmayer and Possingham, 1996) for high-quality habitat in the absence of any catastrophes. Reserves less than 20 ha are of almost no value in isolation, whereas reserves of 100 ha have considerable resilience to demographic and environmental stochasticity. When there are fires that destroy the entire patch, there is no single patch size that gives an acceptably low extinction probability.

country to country. One means of classifying species according to their level of threat would be to do a PVA on every species in the list. Although this has the merits of apparent objectivity, PVAs are rarely used for classifying threatened species because there are insufficient data to parameterize PVAs for so many species. One approach to dealing with the uncertainty associated with extinction probabilities is to take a decision analysis, or fuzzy set, approach in which each species is allocated a probability of being in a particular threatened species category (Akcakaya *et al.*, 2000).

In the United States, the ideas of MVPs, MVHAs, and species classifications are given a litigious dimension because of the federal Endangered Species Act. The Endangered Species Act and associated recovery plans and habitat conservation plans state that any development should not compromise the viability of an endangered species. Some other countries have followed suit. The idea of habitat conservation planning is to allow development where parties agree to the level of mitigation necessary to ensure viability. Because the notion of viability is central to this law and the planning processes that have arisen from the law, PVA has entered the courts in land-management and planning disputes (Bingham and Noon, 1997) .

Population Viability Analysis for Making Management Decisions

During the 1990s, scientists questioned the traditional uses of PVA and also questioned the ability of PVA to accurately predict extinction risk for most species (Possingham *et al.*, 1993; Taylor, 1995; Beissinger and Westphal, 1998). Workers have found that the probability of extinction of an organism is very sensitive to small variations in key life history parameters, especially adult mortality rates. There appears to be adequate data to make very accurate estimates of extinction risk for only a very small fraction of threatened species.

An alternative role for PVA is as a decision support tool to help make management decisions. This role focuses on the robustness of different management decisions to deliver acceptable conservation strategies rather than the accuracy of the predictions alone (McCarthy *et al.*, 2003; Baxter *et al.*, 2006). PVA can be used to make decisions regarding the following: (1) Determining if, and how much of, a population should be brought in to captivity; (2) Translocation strategies between populations or to new sites; (3) Reintroduction strategies (Armstrong and Seddon, 2008); (4) Deciding how to set priorities between species; (5) Determining optimal habitat management, including disturbance regimes such as fire and logging; (6) Control strategies for the predators, competitors, and prey of threatened species; (7) Optimal size, distribution, and spatial pattern of reserve systems.

Probably the most widespread and frequent application of PVA is through the activities of the Captive Breeding Specialist Group (CBSG). This group aids zoos and conservation agencies throughout the world to set conservation priorities and actions by work-shopping species groups region by region. Some argue that PVA has been most useful in providing a framework within which experts can operate. The task of running a PVA draws the experts together and focuses their attention on estimating population dynamics and identifying threats. This is certainly true for the CBSG workshops and users must view PVA as a decision support tool and not as the decision maker.

Other Measures of Viability

The probability of extinction of a population is not the only measure of population viability. Other common measures are the mean and median of the time to extinction. When the extinction probability for a species is fairly constant, the expected life span of the species can be thought of as being negatively exponentially distributed so the rate of extinction is simply the inverse of the mean time to extinction (Grimm and Wissel, 2004). All these measures are fairly closely related.

Because of the uncertainties associated with modeling the dynamics of very small populations (Allee effects, genetic effects, etc.), one prudent approach to providing a measure of risk is to introduce the notion of quasiextinction, which is the risk of the population decreasing in number to some preset size. For example, although the probability of extinction is the chance of decreasing to zero individuals, a manager may argue that the conservation of a species has failed long before this and prefer to set a threshold population size of, for example, 20, below which the population is quasiextinct. This approach eliminates some of the problems associated with modeling very small populations.

An even more general approach to assessing extinction risk that extends the idea of quasiextinction is to develop risk curves. A risk curve (**Figure 6**) shows the probability of a population decreasing below any population size, from zero (extinction) to the carrying capacity of the population. With a risk curve, the manager is free to pick whichever quasiextinction threshold he or she wishes, or to assess alternative management strategies by investigating changes in the whole risk curve. A quasiextinction risk curve can be summarized by the average distance from extinction, which is equal to the expected minimum population size.

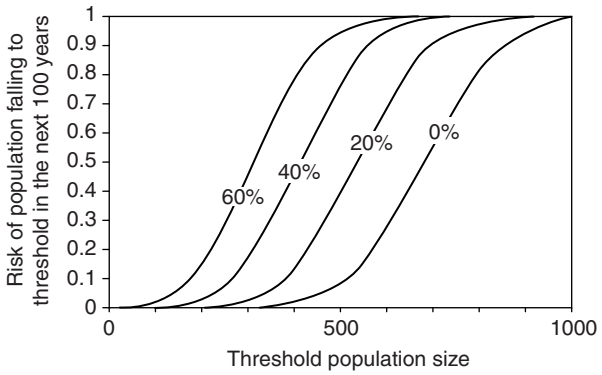


Figure 6 Hypothetical risk curves. The user can choose whatever combination of threshold population size and level of risk he or she wishes. The four lines are quasiextinction risk curves for different levels of population harvesting.

Methods and Tools for Population Viability Analysis

Demographic Models for Population Viability Analysis

Generally, a PVA is a demographic model that includes some or all of the extinction processes discussed previously. There are several ways of constructing these models, and there are several publicly available software packages that people can use to do PVA. The general categories of PVA model are discussed in the following sections.

Matrix Population Models

Basic demographic projection is one of the most powerful and important parts of PVA, although sometimes it is bypassed. By constructing an age- or stage-structured population projection matrix, one can determine both the overall growth rate of the population and the sensitivity of that growth rate to variation in life history parameters by finding the dominant eigenvalue of the matrix. More simply, the population can be projected forward in time using a simple spreadsheet. If the dominant eigenvalue of the population is less than one then we know the population is doomed to extinction (deterministic extinction). If there is uncertainty about the value of key life history parameters, then the risk of deterministic extinction can be calculated by carrying out a sensitivity analysis on the dominant eigenvalue. Although this approach does not allow for any of the stochastic forces described previously, it gives the manager an idea of how the population is faring in the absence of stochasticity. If the dominant eigenvalue is likely to be less than one, then our sensitivity analysis allows us to target those life history attributes, such as birth rate or adult death rate, most likely to generate a positive long-term growth rate. Matrix population models can include stochasticity (indeed, RAMAS Age is based on the application of projection matrices for its predictions), but without a computer package the mathematics is quite complex.

Analytic Methods and Approximations

Although including stochasticity normally requires that we proceed to a computer simulation, there are analytic methods for calculating the probability of extinction, and in particular the mean time to extinction, of populations. One approach uses Markov chain theory to model a basic birth and death process. This model includes only birth and death and can be used to generate analytic expressions for the mean time to extinction.

Another analytic approach involves approximating the stochastic population dynamics with differential equations called "diffusion" equations. Diffusion equations can be analyzed to generate explicit expressions for the mean time to extinction of a population under different sorts of stochasticity; indeed, this approach was used to generate the relationships shown in **Figure 4**.

Although analytic approaches have elegance and generality, they are restricted in the processes which can be included and they can be quite difficult to parameterize from field data. They have made significant contributions to our general appreciation of the role of stochasticity in extinction and underpin a general theory of extinction, but they are rarely used in practice.

Time Series Analysis Methods

A relatively new method for exploring extinction risks utilizes analysis of time series of population abundance. During the past decade, a considerable amount of time series data on a wide variety of organisms have been obtained. If these time series of population size are sufficiently long (at least 15 years), then it is possible to statistically "fit" different models of stochastic population growth (such as a Ricker model or logistic model) to these time series. Such models are necessarily simple but have the advantage that they do not require any life history data – only a series of counts. Given the best fit model, it is a straightforward task to simulate the dynamics into the future and calculate extinction risks (Morris and Doak, 2002).

Although this approach will prove increasingly useful for making extinction risk predictions, and hence helping with tasks such as prioritizing species, the lack of mechanisms in these models limits their application for management purposes.

Numerical Methods

The approach of using a Markov chain (or stochastic population projection matrix) can be expanded to include other forms of stochasticity besides simply demographic. If transitions from any population state to any other population state can be estimated as a function of life history attributes and biotic and abiotic influences, then increasing these matrices to increasingly higher powers allows us to project the probability distribution of possible population sizes into the future. This method provides the foundation of the RAMAS extinction risk software.

The numerical method has the advantage of some mathematical elegance, and it is certainly powerful when coupled to front-end software such as that available in RAMAS. However, most people who build models from scratch find the mathematics involved too complex and tend to use the Monte Carlo simulation method.

Monte Carlo Simulation Models

The most common way of doing PVA is to use a Monte Carlo simulation model. These simulation models are either written

from scratch or the user uses one of the available packages (Lindenmayer *et al.*, 1995). The most frequently used packages are VORTEX and RAMAS (several versions/types); however, others are available. The basic structures of all these models are fairly consistent and are best visualized in a flow chart (**Figure 7**). However, the details of how each component is modeled and which processes can be included in the models vary considerably from model to model and species to species. Each has its strengths and weaknesses. For detailed application to a particular species it is advisable, although expensive and time-consuming, to construct models specific to that species that focus on relevant processes and management options.

For example, several different kinds of PVA models were built to determine efficient and adequate conservation strategies for the northern spotted owl (*Strix occidentalis caurina*). These ranged from simple demographic models to very detailed landscape models with habitat dynamics and very specific biological factors such as predation by another owl. The diversity and number of models for this species largely reflect the controversy regarding it: The spotted owl requires large tracts of forested land that are very valuable for timber production. For species that have far less economic importance, there is neither the time nor the money to embark on detailed modeling exercises and so one of the existing packages is used.

Monte Carlo simulation models run on a PC have the advantage that they can incorporate many different parameters and processes; their complexity is limited only by the imaginations of the biologists and managers. However, the inclusion of detail and complexity can give the illusion of accuracy, whereas some consider that adding more poorly understood processes will make the results less accurate. The factors driving the extinction estimates from complex models are often very difficult to isolate.

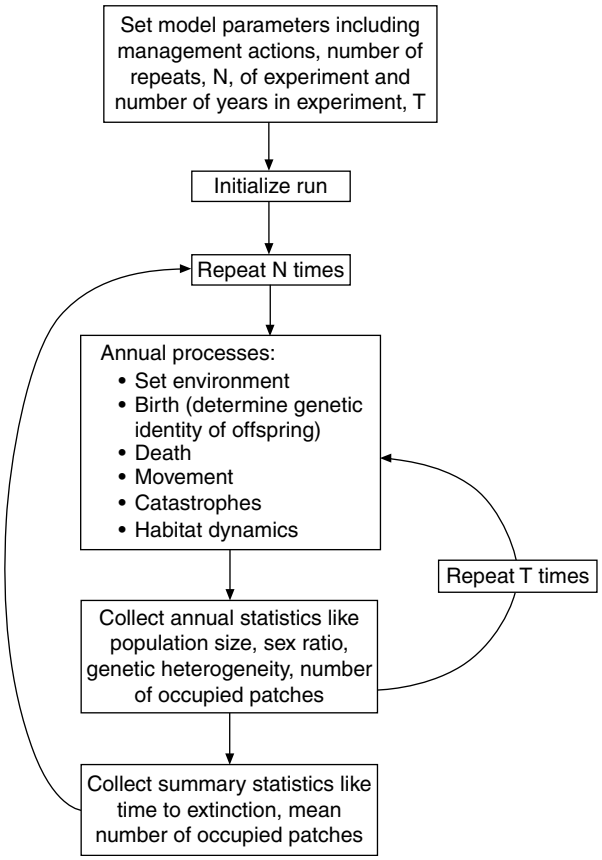


Figure 7 A flow chart showing the typical processes included in a Monte Carlo simulation PVA.

Sensitivity Analysis of Population Viability Analysis Models

Sensitivity analysis is an important component of modeling because one can use it to systematically investigate the complex interactions of a model. Sensitivity is usually measured by varying a parameter by a small amount from its estimated value. The resulting change in the state variable (e.g., the risk of extinction) provides an index of the sensitivity of the model to that parameter. Sensitivity analysis provides practical information for model builders and users by highlighting parameters that have the greatest influence on the results of the model. It can highlight model parameters that should be most accurately measured so as to maximize the precision of the model, give a general indication of the reliability of the model predictions, and highlight parameters and interactions that have the largest influence on the population to help determine effective management strategies.

The sensitivity of deterministic matrix population models can be determined analytically by eigen-analysis. Extension of these techniques to PVA is not appropriate for most PVA models except for matrix models for at least three reasons: (1) PVA models are often complex, so obtaining solutions analytically is difficult if not impossible; (2) in PVA, the result of interest is the risk of population decline and not the deterministic growth rate; and (3) interactions between variables are largely ignored.

Any method of sensitivity analysis should be clearly defined, interactions between parameters should be distinguishable from single parameter effects, and the method should account for variability associated with parameter estimates. The simplest approach to sensitivity analysis of PVA models is to vary the model parameters in turn and investigate the effect on the risk of population decline. Alternatively, such a sensitivity analysis can determine whether the relative efficacy of different management decisions changes with different parameter values. PVA models may be complex with numerous parameters, particularly when individuals are modeled, and understanding the relative importance of different parameters and interactions between parameters may be difficult. For example, if there are 10 parameters and three levels for each parameter are to be investigated, then 21 different parameter combinations would be necessary to assess each parameter independently. However, $3^{10}(=59{,}049)$ different combinations are required to test all possible interactions.

When the risk of population decline is the important state variable, logistic regression may be useful for summarizing the effects of different parameters and interactions. Data for the regression are generated by using numerous parameter combinations. For each parameter combination, the PVA model is simulated to obtain a limited number of predictions of the incidence of decline. The regression analysis uses the model parameters as explanatory variables and the incidence of

decline as the dependent variable. The regression equation provides a simple expression to approximate how the probability of decline is influenced by the model parameters.

Problematic Issues in Population Viability Analysis Models

Although all models of populations are necessarily false because they cannot include all the factors and processes that influence population dynamics, we believe that some processes and factors in PVA modeling are important but are difficult to incorporate.

Genetics in Population Viablity Analysis Models

Early work on PVA paid considerable attention to the role of genetics, although the significance of genetics in short-term threatened species management has declined (Lande, 1988). Two genetic issues are of particular concern: inbreeding depression and loss of heterozygosity. Inbreeding depression is of particular concern for very small populations, especially captive populations. Estimating the likelihood and effect of inbreeding depression on vital rates is problematic. The loss of genetic heterozygosity is a long-term concern. Populations with an effective size greater than 50, are generally thought to experience low losses of heterozygosity. VORTEX is a generally available package that allows the user to include genetic factors, hence its popularity with zoos and reintroduction programs.

Genetic data can be difficult and expensive to obtain and one needs to consider how the data will influence one's decision making. A useful role of molecular genetic data in PVA is to provide estimates and plausible bounds for demographic attributes that are otherwise impossible to obtain (such as social structure and dispersal rates).

Density Dependence in Population Viability Analysis Models

Density dependence at both low and high densities is important for PVA. At low densities the possible consequences of positive (or negative) density dependence are critical to the ability of a species to recover. For populations with limited habitat, the way in which vital rates change as populations approach their carrying capacity, and whether the carrying capacity fluctuates greatly from year to year, is known to influence estimates of viability. Unfortunately, information on density-dependent processes for most threatened species is extremely limited. Different PVA models use different forms of density dependence, from logistic growth and a population "cap" to no density dependence, whereas other models allow the user to choose the appropriate form.

Space in Population Viability Analysis models: Metapopulation Models and Risk Spreading

Space is an essential component of any PVA model for which the intention is to make statements about what management actions should take place at which particular geographic locations. For example, in determining habitat reconstruction plans for threatened species (linking corridors and new habitat patches), it is generally essential to represent space explicitly in the models.

Space can be incorporated into a PVA model either implicitly or explicitly. The implicit use of space is typical of many metapopulation models in which the interest is in the number of patches that are occupied, with no particular regard to exactly which patches are occupied. These models can be used in a general sense to report on the values of different reserve sizes or corridors. For detailed management for which the question of exactly where a corridor needs to be located, or which patch needs to be protected, might need to be answered, we normally require models in which real landscapes are explicitly represented.

Habitat Dynamics

Early PVA models, and almost all of the analytic approaches, assume a static habitat. Given that habitat loss is one of the primary causes of species extinction, the introduction of habitat dynamics in PVA models is of increasing interest. Even in the absence of habitat loss, most habitats are dynamics and some recent PVAs have focused as much on the dynamics of the landscape as on the dynamics of the population. This reflects an increasing interest in population ecology – not just in the role of space but also in the role of a changing habitat mosaic in population dynamics.

Combining habitat dynamics with spatial complexity has been facilitated by faster computers and increasing use of geographic information system (GIS) software. Future wildlife managers will certainly have access to GIS systems that enable them to interactively explore the consequences of different habitat management actions on extinction probabilities. These complex new-age GIS/PVA tools will include the tools inherent in GIS systems, habitat suitability models, management models, stochastic catastrophes, and detailed population dynamics (Akcakaya *et al.*, 2004). To some this will be a powerful and compelling tool and to others a dangerous cascade of complex untested models interacting in ways that the user will never understand.

Other Methods of Assessing Viability

There are many measures of viability and many ways of assessing viability – from a wild guess to a complex computer model, and from a simple analytic model to extrapolation from empirical data. So far, the authors have discussed more traditional model-based methods. Here, the authors briefly mention two other ways of assessing the viability of a population.

Historical Data

The occurrence of populations in isolated habitat, especially on oceanic islands, gives us an idea of how large patches of habitat need to be for the long-term persistence of certain species. For example, in the Caribbean reptiles of certain species rarely occur on islands less than critical sizes, but they are consistently present on larger islands. This sort of island data can help us set MVHAs for particular taxa. As is known from island biogeography theory, islands are not isolated and their species compliment is a consequence of local extinction and colonization processes. In this context, the presence of a species on an island may not reflect its long-term persistence,

only its recent arrival. However, where species have persisted on islands long enough to show variation from mainland populations, we can assume they have persisted for long enough to be termed viable. Although this general approach to PVA has been fruitful, the species that occur on these islands are usually not the same species for which we are trying to design reserve systems on continents. The question arises, how relevant is it to use data on MVHAs from one species for another species?

Extrapolation from Similar Species

Although we are unlikely to have data on the viability of, or MVHAs for, threatened species, it is possible that such information on similar, more common species can be used to assess viability. For example, we may know much about the fate of many small populations of a common species and use this information to determine the MVHA of a similar threatened species.

Expanding this line of thinking, some workers have begun to develop rules of thumb that relate extinction probability of MVHA to key life history attributes of a species, such as body size or home range size. We do know enough to say that predators can persist at much lower densities than herbivores of a certain size, and that larger animals and plants appear to be able to persist for longer periods in small populations than can smaller animals or plants. These fundamental patterns may enable us to draw broad conclusions about the viability of species and strategies for their protection with limited data and no predictive modeling. This crude approach may only be appropriate when data, time, and money are limited.

Testing Population Viability Analysis

Uncertainty about parameter estimates and the structure of PVA models means that their predictions may be inaccurate. As with any model, PVA is an imperfect description of reality, and model testing with real field data helps to identify the limits of its accuracy. The primary prediction of many PVA models is the risk of extinction of an endangered species over several decades or centuries. In such circumstances, we do not have the luxury to test the models by monitoring numerous populations until they become extinct. A sufficient number of populations rarely exist, and endangerment demands immediate management decisions.

Instead, we can test PVA models by predicting population attributes prior to extinction. The kinds of predictions that can be tested include the annual population growth rate, the population size in a series of years, the local population sizes in patches, and the probability of local extinction of patches (Brook *et al.*, 2000; Lindenmayer *et al.*, 2003). It is not sufficient to test the models by comparing mean predictions with observations because this ignores variability, which often is a critical factor influencing the risk of population decline. Tests of PVA models should investigate the distribution of the predictions – not only the average predictions but also factors such as the variance and predicted extremes such as local extinction. Replicates are necessary for such tests and may be obtained from annual observations over several years or from observations from several different locations. These tests rely on appropriate data obtained from the field, which may require a substantial commitment to field surveys. As noted previously, we may need to test PVA models on more common species and then, by analogy, assume that they will work for threatened species. This requires a certain leap of faith.

Correspondence between the predictions of the model and observed data does not imply that the model is correct. Tests of PVAs may not be sufficiently powerful to identify the errors, or multiple compensating errors may exist. Regardless of how many times a PVA model has been tested, new data or a new test will expose the model to a chance of being proved wrong. This is similar to the scientific process of hypothesis testing.

Discrepancies between the predictions of a PVA model and observed data should not lead to complete rejection of the model. Instead, efforts should be made to identify the source of errors, which may be rectified to provide improved predictions. Identification of model errors is a useful part of model development. Such an approach to model testing is consistent with the use of PVA models as decision support tools. Improved PVA models should contribute to improved management decisions.

Criticisms of Population Viability Analysis: What Makes a Good Population Viability Analysis?

There is a healthy skepticism about the value of PVA. Although almost all authors acknowledge that PVA is valuable and a cornerstone of the science of conservation biology, PVA, like any tool, has been abused.

The following are major criticisms of PVA: (1) We don't have enough data to deliver reliable predictions; (2) PVA software packages are models that no one understands; (3) Results are very sensitive to uncertain parameters; (4) The models do not include important ecological processes. Because of the "black box" nature of PVA, there is concern that PVAs should be more explicit and detailed in their reporting of how the model works and exactly how the parameters were chosen. This is just part of good science – that is, whatever we do should be repeatable and transparent. Despite the diversity of PVA models, there is consensus on what constitutes a good PVA. Attributes of a good PVA include: (1) A description of why the PVA is being done (a motivation); (2) A fundamental understanding of the species' ecology, including what constitutes suitable habitat and the ability of the species to disperse between patches of habitat; (3) Detailed demographic data and the way in which demographic rates change with time and habitat: Care needs to be taken in describing exactly what parameters in the model do and exactly how they were estimated from data; (4) An understanding of the response of the species to threats; (5) An assessment of the current state of the population, acknowledging inaccuracies in the estimate (6) A statement about how the PVA is likely to help managers make decisions; (7) A description of constraints on decision variables and management options; (8) An understanding of the environmental disturbances that directly threaten (or indirectly threaten through habitat alteration) a species; (9) A detailed description of the model and data so that the PVA process can be repeated and extended; (10) An honest assessment of uncertainties in the data, processes that have been

ignored, and possible flaws in model construction and parameter estimation; (11) A sensitivity analysis.

Acknowledgments

This work was conducted initially as part of the Biological Diversity Working Group supported by the National Center for Ecological Synthesis and Analysis, a center funded by NSF (Grant No. DEB-94-21535), the University of California at Santa Barbara, and the State of California. Work by all three authors was also supported by several Australian Research Council grants and a Commonwealth Environmental Research Facility grant. We thank our numerous colleagues who have worked on PVA with us, especially Mark Burgman, Sandy Andelman, Bob Lacy, Drew Tyre, and Ian Ball.

See also: Captive Breeding and Reintroduction. Carrying Capacity, Concept of. Computer Systems and Models, Use of. Conservation Genetics. Extinction, Causes of. Framework for Assessment and Monitoring of Biodiversity. Identifying Conservation Priorities Using a Return on Investment Analysis. Inbreeding and Outbreeding. Island Biogeography. Mammals, Conservation Efforts for. Measurement and Analysis of Biodiversity. Metapopulations. Population Density. Population Dynamics. Wildlife Management

References

Akçakaya HR, Ferson S, Burgman MA, Keith DA, Mace GM, and Todd CR (2000) Making consistent IUCN classifications under uncertainty. *Conservation Biology* 14: 1001–1013.

Akçakaya HR, Radeloff VC, Mladenoff DJ, and He HS (2004) Integrating landscape and metapopulation modelling: Viability of the sharp-tailed grouse in a dynamic landscape. *Conservation Biology* 18: 526–537.

Armstrong DP and Seddon PJ (2008) Directions in reintroduction biology. *Trends in Ecology and Evolution* 23: 20–25.

Baxter PWJ, McCarthy MA, Possingham HP, Menkhorst P, and McLean N (2006) Accounting for management costs in sensitivity analyses of matrix population models. *Conservation Biology* 20: 893–905.

Bingham BB and Noon BR (1997) Mitigation of habitat "take": Application to habitat conservation planning. *Conservation Biology* 11: 127–139.

Beissinger SR and Westphal MI (1998) On the use of demographic models of population viability in endangered species management. *Journal of Wildlife Management* 62: 821–841.

Boyce MS (1992) Population viability analysis. *Annual Review of Ecology and Systematics* 23: 481–506.

Brook BW, O'Grady JJ, Chapman AP, Burgman MA, Akçakaya HR, and Frankham R (2000) Predictive accuracy of population viability analysis in conservation biology. *Nature* 404: 385–387.

Burgman MA, Ferson S, and Akçakaya HR (1993) *Risk Assessment in Conservation Biology*. London: Chapman and Hall.

Caughley G (1994) Directions in conservation biology. *Journal of Animal Ecology* 63: 215–244.

Gilpin ME and Soule ME (1986) Minimum viable populations: Processes of species extinction. *In conservation biology: The science of scarcity and diversity*, pp. 19–34. Sunderland, MA: Sinauer.

Grimm V and Wissel C (2004) The intrinsic mean time to extinction: A unifying approach to analysing persistence and viability of populations. *Oikos* 105: 501–511.

Kendall BE and Fox GA (2003) Unstructured individual variation and demographic stochasticity. *Conservation Biology* 17: 1170–1172.

Lacy RC (1993) VORTEX – A computer-simulation model for population viability analysis. *Wildlife Research* 20: 45–65.

Lande R (1988) Genetics and demography in biological conservation. *Science* 241: 1455–1460.

Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist* 142: 911–927.

Lindenmayer DB and Possingham HP (1996) Ranking conservation and timber management options for Leadbeater's possum in southeastern Australia using population viability analysis. *Conservation Biology* 10: 235–251.

Lindenmayer DB, Possingham HP, Lacy RC, McCarthy MA, and Pope ML (2003) How accurate are population models? Lessons from landscape-scale tests in a fragmented system. *Ecology Letters* 6: 41–47.

Lindenmayer DW, Burgman M, Akçakaya R, Lacy R, and Possingham HP (1995) A review of three generic computer models - ALEX, RAMAS/Space and VORTEX, for modelling the viability of metapopulations. *Ecological Modelling* 82: 161–174.

Mace GM and Lander R (1991) Assessing extinction threats – towards a re-evaulation of IUCN threatened species categories. *Conservation Biology* 5: 148–157.

Mangel M and Tier C (1993) A simple and direct method for finding persistence times of populations and application to conservation problems. *Proceedings of the National Academy of Sciences USA* 90: 1083–1086.

McCarthy MA, Andelman SJ, and Possingham HP (2003) Reliability of relative predictions in population viability analysis. *Conservation Biology* 17: 1–9.

Morris WF and Doak DF (2002) *Quantitative Conservation Biology: The Theory and Practice of Population Viability Analysis*. Sunderland, MA: Sinauer Associates.

Possingham HP, Lindenmayer DB, and Norton TW (1993) A framework for the improved management of threatened species based on population viability analysis. *Pacific Conservation Biology* 1: 39–45.

Shaffer ML (1981) Minimum population sizes for species conservation. *Bioscience* 31: 131–134.

Taylor BL (1995) The reliability of using population viability analysis for risk classification of species. *Conservation Biology* 9: 551–558.

Poverty and Biodiversity

Madhav Gadgil, Indian Institute of Science, Bangalore, India

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 845–856, © 2001, Elsevier Inc.

Glossary

Artificial ecosystems Ecosystems dominated by human artifacts and human induced material and energy flows.
Biosphere people People enjoying access to resources garnered from the entire biosphere and made available through markets.
Biosphere reserve An international conservation program aiming to protect representative ecosystems throughout the world in ways compatible with development efforts.
Ecological footprints Spatial coverage of ecological impacts consequent on the activities of a given group of people.
Ecological refugees People that have lost access to their traditional base of natural resources yet have very limited access to resources through markets.
Ecosystem people People meeting the bulk of their resource requirements from a limited area near their habitation through gathering or low input agriculture and animal husbandry.
Resource catchment Locality from which the resources consumed by a group of people are derived.

Wealth and its obverse, poverty, is today reckoned in terms of the capacity of a person or groups of persons to obtain goods and services through exchanges in the marketplace. The bulk of such goods and services are partly or wholly products of intensively managed or artificial ecosystems; crop fields, plantations, shrimp ponds, mechanical fishing fleets, factories, towns, and cities. Most of these intensively managed or artificial ecosystems tend to harbor low levels of biodiversity. Rich people are then those with extensive access to produce of managed or artificial ecosystems supporting low levels of biodiversity; poor people have very limited access to such produce. Ecosystems harboring high levels of biodiversity are, on the other hand, natural or semi-natural ecosystems with low levels of human demands for their produce. Rich people rarely live in close vicinity of such ecosystems, though they may visit them for recreational purposes. Groups of poor people may, on the other hand, permanently live in their vicinity. In some cases, the poor control and serve as the stewards of such biodiversity rich ecosystems. More often though many such biodiversity rich ecosystems are subjected to overexploitation, primarily to meet the large resource demands of the rich, often living far away, with the local poor serving as agents of the resultant destruction of biodiversity. Large numbers of poor also live permanently in the vicinity of biodiversity poor, intensively managed, or artificial ecosystems. There is then no simple relationship between poverty and biodiversity; the equations vary greatly from context to context.

Of Ecosystems and People

To appreciate these complexities, one must tease out the many contexts. These may be viewed along two axes; one of the extent to which ecosystems have been transformed through human interventions, and the second of the manner in which people relate to the world of nature, to the world of manufactured artifacts, and to each other. The terrestrial and aquatic ecosystems may be classified into four major categories: (a) natural ecosystems, which are subject to very low levels of human demands because of their inaccessibility; (b) natural and semi-natural ecosystems, including low-input agricultural systems subject to higher levels of human demands; (c) ecosystems managed intensively for biological production; and (d) largely artificial ecosystems dedicated to industrial production and organized services. Building on Dasmann's (1988) pioneering work, we may, for our purpose, classify people into four major, largely inclusive categories; (a) autonomous ecosystem people, (b) subjugated ecosystem people, (c) biosphere people, and (d) ecological refugees. Ecosystem people have limited access to sources of energy other than human and livestock muscle power and to the more sophisticated artifacts. They gather or produce most of the resources they consume from their immediate surroundings, from the forest, scrub, rivers, or seas and from low-input cultivation. In some of the more inaccessible corners of the world, the ecosystem people are still autonomous. However, over most of the earth they have been subjugated by the biosphere people and have very limited control over their own resource base of natural and semi-natural ecosystems. They gather and produce little that can fetch value in markets and therefore have very limited access to produce of intensively managed and artificial ecosystems. The biosphere people owe their dominant position to their extensive control over artifacts and additional sources of energy. They engage in energy-intensive agriculture, animal husbandry, aquaculture, or in organized services or industrial production and generate much of value in the markets. They have large ecological footprints thanks to their substantial purchasing power; their resource catchments are vast encompassing all of the biosphere, bringing to them goods and services from all over the earth. This confers on them the ability to take over

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00373-7

resources in demand by the ecosystem people, catalyzing transformation of natural and semi-natural ecosystems into those managed intensively to meet their own demands or, in some special cases, conserved as natural or semi-natural ecosystems for recreational purposes. In the process, they often deprive ecosystem people of access to their traditional resources; this may convert them into ecological refugees. Ecological refugees are then people with attenuated access to resources of natural and semi-natural ecosystems, with little purchasing power to access products of intensively managed and artificial ecosystems. They often end up constituting the unorganized labor force in tracts of intensive agriculture and urban settlements (Gadgil, 1995; Gadgil and Guha, 1995).

Autonomous People

I propose to explore the relation between poverty—or wealth—and biodiversity in this framework of different categories of ecosystems and of people. **Table 1** summarizes the relationship and notes the examples that are discussed later at some length. All of these derive from my own fieldwork in India; I use them purely because I know them well. However, they do represent broader patterns encountered in other parts of the world. Not all 16 cells of the matrix are occupied; moreover our focus is on poverty, and we shall therefore not elaborate on the relationship of the biosphere people with biodiversity, except as they impact on the ecosystem people or ecological refugees. There are relatively few examples of truly autonomous ecosystem people, people fully in control of largely natural ecosystems with very light human demands, and in consequence with high levels of biodiversity. Inhabitants of Sentinelese island (ll°30′ N lat. 92°15′ E long.) in the Andaman-Nicobar chain in the Bay of Bengal provide one such example. These relatively inaccessible islands harbor tropical rain forest biota with high levels of endemism (i.e., of species restricted to this chain of islands). They were inhabited by a number of hunter-gatherer tribal groups, without knowledge of metal tools, with the exception of the Shompens of Nicobars who had a more advanced fishing economy. The British colonized the islands in the mid-nineteenth century, an attempt that was strongly resisted by the indigenous people. By the 1870s the resistance was overcome and as a consequence many of the tribal populations either drastically declined or were exterminated. During British times the islands primarily served as a convict colony, with many of the released convicts settling down to agriculture (Superintendent of Government Printing, 1909). After independence in 1947 the islands were used to create agriculture based settlements of many people displaced from then East Pakistan (now Bangladesh), as well as to supply forest resources for rapidly growing forest based industries, especially plywood (Saldanha, 1989). As these were progressively overused and exhausted leading to substantial erosion of biodiversity, pressure has built up to overcome the resistance of the tribal groups still holding out to open up their territories to commercial forest exploitation. Two of the tribal groups, however, do continue to retain hold over their territories; these are Jarwas and Sentinelese. Jarwas live on the larger South

Table 1 Variety of people-ecosystem contexts, with specific examples

Ecosystems	*People*			
	Autonomous ecosystem people	*Subjugated ecosystem people*	*Biosphere people*	*Ecological refugees*
Inaccessible natural/ seminatural ecosystems	Poor in modern economic sense, with access to high levels of biodiversity, e.g., Sentinelese islanders	—	—	—
Accessible natural/ seminatural ecosystems	—	Poor in modern economic sense, serve as agents for destruction of biodiversity, e.g., Gangtes of Manipur, sometimes as stewards, e.g., Village Forest Committees, protectors of Saranas	Visitors for recreational/ commercial use	Poor in modern economic sense, serve as agents of destruction of biodiversity, e.g., Panshet peasants, Maldharis of Gir
Ecosystems managed intensively for biological production	—	—	Well off in modern economic sense, promote low biodiversity production systems	Poor in modern economic sense, alienated from biodiversity, e.g., Jharkhand tribals
Artificial ecosystems dedicated to industrial production/ organized services	—	—	Well off in modern economic sense, promote high biodiversity systems for recreational purposes	Poor in modern economic sense, alienated from biodiversity, e.g., Jharkhand tribals

Andaman and Baratung islands (12°0′ N–12°20′ N lat. 92°45′ E–92°55′ E long), about two-thirds of which has been colonized by immigrants and subjected to forest exploitation. Their territory is therefore easily accessible; it is, however, stoutly defended by Jarwas with their bows and arrows. There are ongoing attempts to overcome this with the aid of a major road passing through their territory. This Jarwa territory remains much richer in biodiversity than the rest of the islands; Jarwas are, however, poorer in a modern economic sense with access to very few and simple artifacts. The Sentinelese live on a smaller island, which remains entirely under their control. This island too is very rich in biodiversity; the Sentinelese are also very poor in a modern economic sense. Presumably Jarwas and Sentinese do not themselves have concepts of wealth and poverty or biodiversity (Bhattacharyya, 1993; Gadgil, 1998). They apparently do have concepts of members of their own groups as opposed to aliens and desirability of excluding aliens from their territory. They remain poor in the modern sense because they have successfully maintained themselves in isolation; for the same reason their territory retains high levels of biodiversity.

Subjugation: Political and Economic

The vast majority of the ecosystem people of the world are, however, no longer isolated in this fashion. In no case do they seem to have voluntarily sought to integrate with the larger society dominated by the biosphere people. Rather the biosphere people have sought to integrate them in order to access resources of their territories, as well as to utilize their labor. The process of such integration tends to proceed through several stages involving political subjugation, followed by economic subjugation. Such subjugation is accompanied by a reduction in levels of access of the people to the biodiversity resources of their own localities; though simultaneously they may have enhanced access to other resources through market channels, thereby implying a reduction in the levels of their poverty in modern economic terms. The process is also generally accompanied by an erosion of biodiversity driven by overexploitation of natural resources to meet outside demands with the subjugated ecosystem people often serving as agents of such erosion.

These processes are the dominant form of interplay of poverty and biodiversity in the world today, and we propose to illustrate them through a series of case studies. The most detailed case to be presented would involve Gangtes, one of the Kuki tribes of Churchandpur district (23°4′–25°50′ N lat. and 93°–94°47′ E long.) of the northeastern state of Manipur on the India-Myanmar-China border, since this illustrates the complete sequence from autonomy through political and then economic subjugation (Gadgil, Hemam, and Reddy, 1997; Hemam, 1997). Like many other Kuki and Naga groups, Gangtes continued head hunting well into the nineteenth century. At that time, there were no roads so that people had to walk for several days to reach markets where they could exchange hand-woven cloth or honey for iron tools or rice. These journeys were hazardous as they could involve passage through territories of alien tribes. Each settlement was in consequence a largely self-sufficient and self-governing entity. Their base of subsistence was shifting cultivation with long fallows of about 15 years. There were no fixed village boundaries and individual social groups presumably shifted around from time to time. The settlements were located on hill-tops; the valleys were considered unsafe and more vulnerable to raids from alien groups. Among Gangtes, hereditary village chiefs determined through primogeniture made all group-level decisions in consultations with a council of elders. They assigned lands for cultivation and annually received about 16 kg of grain from each family in the group as a tribute for use in entertaining any visitors. In the course of cultivation the densest, tallest forest tracts were left alone as requiring too much labor for conversion to fields; the more recent fallows were also avoided. The people worshipped many natural elements including mountain peaks, streams, plants, and animals. They strictly protected patches of sacred groves called gamkhal as also other spirit possessed lands called nungens. Also protected were bamboo groves called mavuhak from which bamboo may be extracted only for house construction, but the shoots, much relished as food, were left alone. This preserved a luxuriant vegetation, which led Captain Pemberton to remark in 1835: "I know no spot in India, in which the products of the Forests are more varied and magnificent but their utility is entirely local" (Pemberton, 1835).

The process of political subjugation of Gangtes began in the late nineteenth and early twentieth centuries with the development of incipient links with mainstream India, then ruled by the British. Initially, the British did not so much want access to resources of Churchandpur as of the much richer Myanmar. Hence they wanted to ensure safe movements of British troops in the hill tracts of Manipur bordering Myanmar. A people without fixed, well-defined villages and with traditions of killing aliens coming into their territories were an obvious threat. The British therefore concentrated on fixing village boundaries and assigning land ownership. Given the Gangte system of hereditary village chiefs having a major role in all decisions including assigning land for shifting fields, they decided to confer on the chiefs all ownership over land, converting others into tenant farmers from a legal perspective. This move then formally introduced wealth and poverty to the Gangte society. However, during the British reign this made little operational difference in most areas, with no acceptance of private property in land by Gangtes, except in the Churchandpur town area beginning in the 1930s. Elsewhere Gangte community members continued to pay the village chief a tribute of 16 kg of grain as before. The late nineteenth and early twentieth centuries also witnessed the gradual spread of Christianity. Christianity rejects the attribution of sacred qualities to elements of nature and, consequently, taboos against killing of certain animals or felling of certain patches of forests like gamkhal. Conversion to Christianity therefore began to slowly erode the traditional belief system underlying these conservation practices.

In this phase of political subjugation there was little change in the economy. The road network remained extremely limited and the Gangte communities remained almost totally self-sufficient in terms of resource use. The fallow cycles for shifting cultivation were still long, and substantial areas of forest retained protection in forms of sacred groves. However, by introducing the legal concept of private ownership and a religion that questioned attribution of sacred qualities to

nature, this phase did set the stage for the radical changes that followed independence; but unlike on the mainland, this political subjugation had limited impact in the absence of an access to markets and therefore of any attempt at economic subjugation.

Market Forces

The British were prompted by the march of the 'Free India Army' fighting in collaboration with the Japanese into Manipur through Myanmar to strengthen the road network of Manipur on a war footing, laying the foundation for the rapid development of transport, communication, and commodification that began on independence in the 1950s. This was consistent with the policies of economic development and national integration adopted at the time of independence. These development policies also encouraged forest based industries by offering them many resources, including wood, at highly subsidized rates (Gadgil and Guha, 1992, 1995). As the production of these industries grew at a rapid pace, the demands outstripped supplies leading to severe overexploitation. Such overexploitation was facilitated by the fact that there was no social group sufficiently motivated to ensure a more sustainable pattern of resource use with a secure enough control over the resource stocks. Access to markets meant that a community no longer suffered an inevitable shortage of resources if those from their own vicinity were overused and exhausted. Neither would such a difficulty be experienced by an industry drawing resources from a larger spatial scale. Market access therefore reduced the motivation of the various parties concerned, local communities, such as those of Gangtes, as well as other consumers such as the plywood industry for restraints on levels of harvests. At the same time conversion to Christianity eroded the conservation ethos of local communities grounded in attribution of sacred qualities to various natural elements. Linkages to the larger national society also diluted the control of any one agency over any given resource base by bringing a variety of actors into play. Thus in the autonomous phase where a person might even be killed while passing through alien territory, there was a clear-cut control by a local community. In the phase of political subjugation this was legally recognized as the control in terms of ownership by the local chieftains. But in the post-independence phase of economic subjugation the state Forest Department has stepped in to claim control over forest lands and forest resources. How this claim is to be reconciled with the claim over land ownership by the Gangte chiefs has not been clarified. This uncertainty has affected the security of resource control by both parties. At the same time, the traditional authority structure of a Gangte village community headed by a hereditary chief with a council of elders has also been affected by the institution of a democratically elected village council. There are continuing contradictions in this system as well since the hereditary chief remains at the head of the elected village council. Under these circumstances, forest resources of Gangte villages have been affected by demands greatly exceeding sustainable yield levels. These demands are no longer local, but reflect the markets not only of remaining parts of the state of Manipur, but the rest of the country as well. Indeed market pulls come from even beyond national borders as Indian plywood was over some periods exported in substantial quantities to the mid-east.

The traditional Gangte society is largely egalitarian. This situation has radically changed with the linkages to the market economy as the chiefs are now motivated to take advantage of the individual ownership of vast tracts of lands assigned to them. Their assertion of such ownership rights over the land or the tree resources standing on the land is against customary norms. The chiefs therefore press home such assertions to a variable degree, depending on the level of the expected monetary gain. The higher the level of expected gain, the more firmly is ownership asserted by the chiefs. In general, such gains are highest close to the market town of Churchandpur, they decline with distance from the town and from the roadhead. As **Table 2** shows, this is reflected in the chiefs forcing all cultivators to buy plots of land close to the town, while in remote areas all that the cultivators pay is the traditional annual tribute of 16 kg of grain. Similarly close to the town community members pay a royalty to the chief even for collection of fuelwood, while further afield they have free access to even more valuable timber.

Land Use Changes

This process of over-exploitation of forest resources has also shaped patterns of use of land. Since the forest resources bring in the highest levels of net profits closest to the market town of Churchandpur, all tree stocks have been exhausted in this area. In response, settled cultivation on terraced fields with very low levels of crop diversity such as monocultures of pineapple has replaced the much more biologically diverse shifting cultivation. In more remote areas too, commercial monoculture plantations have been taken up of species like Agor (*Aquilaria aquatocha*), the highly valuable wood that was the first to be exhausted even from areas far from market.

The traditional land-use pattern of Gangtes included leaving aside substantial areas as gamkhals, or so-called village forest reserves (VFR), as well as bamboo reserves and other spirit lands. These practices have been completely abolished in villages close to market towns where land and its produce have acquired high commercial value. In some of the more remote villages the practice of protection of a VFR encircling the settlement has been revived following recognition of its value as a firebreak preventing the spread of fire to the settlement during the slash and burn operations (Gadgil, Hemam, and Reddy, 1997). The traditions of protection of such forest patches still continue in other more remote areas.

Links to markets has meant access to many new goods such as soaps and transistor radios for the Gangtes; goods that need cash for purchase. Sale of timber, fuelwood, and other forest produce such as cane is their only source of cash. But generation of this cash has led to a reduction in the return of nutrients to the fields in slash and burn cycles with a depletion of the standing tree stocks. This has reduced the levels of productivity of shifting cultivation (Ramakrishnan, 1992). The spurt in population growth, favored by introduction of modern health care systems, especially control of malaria, with a rate as high as 4% in some of the north eastern hill

Table 2 Resource ownership, sharing, and utilization patterns among different sections of Gangte tribe at different levels of accessibility and modernization

	Shifting cultivation area			
	Less accessible	*Easily accessible*	*Settled agriculture area*	*Town area*
1. Ownership	Village chief	Village chief	Private	Private
2. Commercialization of forest and current forest condition	Little commercial exploitation and less degradation	Under heavy commercial exploitation and degraded condition	No forest	No forest
3. Onset of commercial forest exploitation	Still untouched, except for domestic use	Early 1970s	Early 1950s	Early 1930s
4. Current status of exploitation	Little exploited, good forest cover	Heavily exploited, degraded forest	Completely exhausted	Completely exhausted
5. Collection of timber	Free access	Pay royalty to the chief	No collection	No collection
6. Collection of fuelwood	Free access	Free access for domestic use	Pay royalty to the chief	Pay royalty to the chief
7. Maintenance of VFR[a] and Bamboo reserve	Still intact	No more	No more	No more
8. Traditional rights of resource sharing and use	Still followed	All have disappeared except for free use of shifting fields and fuelwood collection for domestic use	No more	No more
9. Collection from forest	Fuelwood, wild vegetables, timber and other NTFP[b]	Fuelwood, wild vegetables, timber, and other NTFP	Fuelwood collected from forests of other villages	Occasional collection of fuelwood from nearby forests
10. Land use	Shifting cultivation	Shifting cultivation, terraced cultivation	Wet rice cultivation, pineapple plantation	Habitation
11. Commercial crop plantation	Chili and other vegetables	Banana, chili and vegetables	Pineapple, ginger, and other vegetables	Habitation
12. Energy use	Human muscle power, fuel-wood, and little use of fossil fuel	Human muscle power, fuelwood, and little use of fossil fuel	Human muscle power, animal power, fuel- wood, fossil fuel, and electricity	Fuelwood, electricity, fossil fuel, and less use of human and animal muscle power

[a]Village Forest Reserve (VRF).
[b]Nontimber Forest Produce (NTFP).

areas has also increased the pressure on land and led to a shortening of fallow cycle and a reduction in productivity of shifting cultivation. The net result is that today Gangtes depend on market for about 30% of their food requirements and over 50% of other needs (Hemam, 1997).

Value Appropriation

These cash requirements are met from harvests of forest produce, either through sale or through wages earned as laborers in tree-felling operations taken up by contractors. With the assertion of land ownership and demand for royalty by chiefs, part of the profits go to them. However, the chiefs are unfamiliar with the skills of development of infrastructure such as roads, organizing transport, and marketing of forest produce. As a result outside contractors carry out most of the timber harvesting operations, retaining a bulk of the profits for themselves. Thus in the early 1990s in Nalwan village the contractor paid Rs. 50,000 (US $ 1600) for rights to harvest timber from an area of over 1000 ha. A conservative estimate of the value of timber thus accessed is Rs. 60 million given a standing stock of 100 tonnes/ha and price of timber at Rs. 600/tonne. The chief was then paid less than 1.0% of the timber value. Of course, the contractor would have to invest in labor charges, making roads, and transport. Nevertheless, it is clear that bulk of the cash income flows to the contractor, a part to the chief, and a very small fraction to members of Gangte community. The Gangte use this cash to meet their routine requirements. None of it is saved or invested productively in ventures such as developing horticultural plantations to replace shifting cultivation fields. In the long run then the whole process of forest resource use taking place today is leading to exhaustion of these resources with little gains for the Gangte community members.

Today the Gangtes have greater access than they had earlier to goods on the markets and are in that sense richer, though this access is still extremely limited and by all criteria they are quite poor. However, at an earlier time when they were poorer in terms of market access, they had more secure and more equitable rights over land and substantial access to biodiversity and other natural resources; this has drastically reduced. The erosion of biodiversity this transformation has accompanied has in part been driven by the desire of Gangtes to acquire cash; however, the real driving force behind the loss of biodiversity lies in the demand for forest produce by the

biosphere people. Much of the wood of Manipur has been used for manufacture of plywood, and all the value added in this process has been appropriated by those already rather well off, controlling trade and industry of mainland India. The transformation has also witnessed a loss of the traditions of conservation of biodiversity, such as those of protection of sacred groves, by relatively poor people like Gangtes.

Gangtes thus provide an excellent case study of the different stages and processes involved in the transformation of autonomous societies of ecosystem people with access to high levels of biodiversity, but with little access to manufactured artifacts into societies that have lost control over their base of living resources, often turning into instruments of depletion of these resources. Thus subjugated, the ecosystem people may end up with some, though often insignificant, enhancement in their access to artifacts, but with significant losses in their access to biodiversity.

Ecological Refugees

Gangtes have come under the influence of the state and the market economy relatively recently. True, most of them have lost all legal claims over land or forest resources, and the claims of Gangte chiefs over large tracts of land are being contested by the Forest Department; nevertheless, in fact, Gangtes still control extensive tracts of forested land. This level of autonomy is unusual on the Indian subcontinent where state authority has been holding sway for several centuries. One such unusual area is that of Maharashtra Western Ghats, where the local peasants had retained a considerable measure of autonomy by virtue of their active role in the Maratha armies of the seventeenth and eighteenth centuries. The British were reluctant to offend them on the defeat of Marathas in the early nineteenth century. As a consequence they were permitted to continue their shifting cultivation, which was banned over much of the remainder of the subcontinent. Individual peasants, rather than any chiefs were assigned land ownership among the Marathas. So while poor in modern economic terms because their hilly terrain permitted only very restricted access to the markets, these peasants lived in tracts rich in forests and biodiversity with full rights over the resources till the time of independence. The tract was also rich in water resources with annual precipitation in excess of 3000 mm.

On independence the Indian nation state geared itself to tap these rich water resources to promote economic growth with the help of irrigation and hydroelectric projects. One such irrigation project was launched in the Panshet valley to the west of Pune, the old capital of Marathas (18°30′ N lat.–73°40′ E long.). The Indian state has assumed extensive powers of acquisition of land for public purposes and all the more fertile low lying paddy land of the Panshet valley was taken over at low rates of compensation. The peasants still retained ownership over their hill lands above the submersion level. This had been traditionally under long fallow shifting cultivation during which mango and *Terminalia chebula* trees, both producing fruit of considerable value, were left intact. However, as roads reached Panshet valley to facilitate the dam construction work, these trees caught the attention of charcoal merchants. These merchants worked in collusion with the government officials in charge of land acquisition to persuade the peasants to sell the trees at exceedingly low prices with the assurance that the peasants would soon be resettled with alternative farmland far from Panshet valley. Most peasants sold the trees, and there was a massive spurt of deforestation in the 1950s during the construction of the dam, the trees being primarily used to produce wood charcoal for consumption in Pune. The peasants, however, were never provided alternative agricultural land elsewhere and have continued staying on hill slopes above the dam for the past 40 years, while some of them have migrated to Pune, primarily to work as unskilled labor. The entire catchment has been devastated by the initial deforestation in the 1950s followed by continual, ever more intensive cultivation of hill slopes by the impoverished peasants. Today the peasants are poor in every sense and live in an environment depleted of biodiversity. They have served as agents of this destruction of biodiversity that has only benefited the charcoal merchants of Pune and provided citizens of Pune with a relatively inexpensive fuel for a period of few years. Thus subjugated, these ecosystem people may be said to have been turned into ecological refugees (Gadgil, 1979).

Our next case study is that of the tribal populations of the so-called Jharkhand tracts of the states of Bihar and West Bengal (22°0′–24°0′ N lat. 84°0′–87°40′ E long.). This region adjoins the old, thickly settled Gangetic plains that were the nucleus of the 2000 year old Magadha empire extending over much of the Indian subcontinent. Emperor Ashoka of Magadha is famous for his conquest of and subjugation of the tribes in the extensive hilly tracts that border the Gangetic plains (Gadgil and Guha, 1992). This part of the country also fell rather early to British colonialists, who quickly established a hold over the region during the second half of the eighteenth century, far more firm than that over the Maratha Kingdom discussed earlier. The forests of Jharkhand are rich in sal (*Shorea robusta*), an important timber for the production of railway sleepers. This was a resource much valued by the British who therefore acted firmly to ban all shifting cultivation by the tribal people in the early nineteenth century. The forests were also taken over either as a property of the British government or that of private landlords. Where the tribal villages were a part of the government-controlled forest lands, the tribes were totally deprived of any rights to forest resources. Unlike with Gangtes, the landlords too did not come from among the tribes themselves. They came from other castes at a higher level in the hierarchy. These landlords then exploited the lower caste or tribal tenants on their lands quite ruthlessly. The "Jharkhand" tribals were therefore subjected to extreme political as well as economic subjugation as early as the first half of the nineteenth century. Their only earnings came from work as poorly paid tenant farmers or as forest laborers along with sale of some minor forest produce at extremely low rates. They then have a long history as poor people serving as agents of destruction of biodiversity.

These Jharkhand tribals are among the most striking example of ecological refugees of the country. They have served as the most mobile source of very inexpensive, unskilled labor in many contexts. One of the major economic enterprises under the British rule was the development of tea estates replacing the rain forest of Brahmaputra valley in northeastern India. This involved wholesale destruction of biodiversity.

While the plantation owners were British, the laborers responsible for actual destruction of biodiversity were mostly Jharkhand tribals working under conditions that have been described as being close to slavery (Gadgil, 1942).

Elsewhere, these tribals were resettled in so-called forest villages. The primary focus of the plantations they worked on was to replace the natural sal-dominated, fairly diverse humid forest with monocultures of teak (*Tectona grandis*). Teak is excellent timber resistant to termite and fungal attacks, highly valued for ship building and gun carriages in the nineteenth and for furniture and house construction in the twentieth century. But teak is hard wood, little used traditionally; it also does not provide any other product of local utility. On the other hand, sal leaves are used to make plates and sal seed has value as an oil seed. Many associates of sal, such as mahua (*Madhuca indica*) and tendu (*Diospyros melanoxylon*) have great local utility. So replacement of natural sal forest by teak plantations deprives local people of access to a diversity of biological resources of value. Uprooted impoverished tribals have served over years as primary agents of such an erosion of biodiversity (Gadgil and Guha, 1995).

Transformations such as these are particularly common today in tropical developing countries, transformations that are actively encouraged by nation states as well as transnational development agencies as development efforts. These development efforts then have adverse impacts both on biodiversity and on the ability of the weaker sections of the population to access biodiversity. Indeed, they also have adverse impacts on the biodiversity conservation practices of the ecosystem people, practices such as protection to sacred plants and animals and sacred sites such as sacred groves and ponds.

Costs of Conservation

Conservation practices of the ecosystem people with their limited resource catchments are organized on limited spatial scales. Thus, the sacred groves have sizes ranging from a fraction of a hectare, to tens, at best hundreds of hectares. The biosphere people have their own conservation practices. In keeping with their larger resource catchments, they tend to protect much larger areas as nature reserves (Gadgil, 1996). Agrarian states had their tradition of hunting preserves of aristocracy, ranging in size from few hundreds to thousands of hectares. Such hunting preserves of nobles are described in Kautilya's Arthashastra, the 2000- year-old Indian manual of statescraft (Kangle, 1969). They are also known from medieval Asia and Europe. The hunting preserves were aimed at protecting major game animals such as lions or antelopes and their habitat for the entertainment of aristocracy. Their main focus was therefore on exclusion of access, especially to hunting in these areas, by all the commoners. In this, they would have tended to accentuate the poverty of people living within and adjacent to the hunting preserves.

The modern-day nature reserves are even larger in spatial scale, in keeping with the even larger resource catchments of modern-day societies of biosphere people. Thus the largest national parks of the day extend over hundreds of thousands of hectares. Many of them, as in India, have their roots in the hunting preserves of the aristocracy. Two notable examples of these are the national parks at Gir, the last stronghold of the lion in India, and at Keoladev, Ghana, a wetland that shelters enormous nesting colonies of herons, storks, pelicans, and cormorants during monsoons and huge populations of migrating waterfowl in the winter. The focus at both these reserves is on exclusion of resource access by local peasants and herders, leading to major controversies.

At Gir (21°40′ N lat. 70°30′ E long.), the controversy centers on relocation of traditional colonies of buffalo keepers, the Maldharis outside the national park. Their free-ranging buffaloes were earlier an important prey for the lions, especially the males. On being located outside, the quality of grazing for the buffaloes has declined, and despite some reduction in the mortality of buffaloes from lions, the Maldharis are worse off. Within the national park the herbivore populations have increased with reduction of competition for grazing by buffaloes, but the lion population has not gone up. In fact, male lions continue to hunt Maldhari buffaloes who are now outside the limits of the national park. While Maldharis were earlier willing to accept some buffalo kills by lions as a compensation for access to grazing within the national park, they are no longer willing to do so. So they poison carcasses of buffaloes killed by lions; as a consequence the lions suffer (Johnsingh, *pers. comm.*)

In Keoledev national park (27°15′ N lat.–77°35′ E long.), famous for its water birds too, modern conservation attempts have involved exclusion of grazing in the extensive shallow wetlands. When grazing was thus banned in early 1980s without any alternatives being made available, there were protests leading to police firing, with some deaths. But the ban was implemented forcing local peasants to substantially reduce their livestock holdings, impoverishing them to a significant extent. But this cessation of grazing has also had an adverse impact on wetlands as a water bird habitat. This is because the wetlands have now become choked by an excessive growth of a grass, Paspalum. This particular conservation measure has thus tended to impoverish people as well as the bird life, which was at the center of the conservation concern (Vijayan, 1987).

Jharkhand tribals too have suffered from being excluded in the interest of nature conservation. One of the major recreation areas for the wealthy from Calcutta city is the Betla National Park (23°40′ N lat.–84°40′ E long.) in the Daltonganj district of Bihar. The tribals who were earlier settled in may forest villages in what now constitutes this national park are facing serious difficulties. With the constitution of the national park, the forestry operations employing them have been halted. Further restrictions are imposed on their collection of forest produce. Moreover, the tribals are now being asked to move out of the forest villages altogether, without any provisions for alternative livelihoods. So the tribals who were earlier instruments of destruction of biodiversity to help meet economic demands of the wealthy are now being further impoverished to meet their recreational demands for access to biodiversity (Gadgil, 1998).

Ecodevelopment

But there are other positive developments as well. The difficulties experienced by Jharkhand tribals have triggered a

variety of political protests, some non-violent, others violent. Among their demands is the creation of a separate Jharkhand state; since they see states of Bihar and West Bengal as being dominated by interests responsible for the political and economic subjugation of tribals. One of the slogans of this Jharkhand movement has an ecological content: Teak is Bihar, sal is Jharkhand. Their demand is for the retention of natural sal-dominated forest.

A notable development accompanying these protests is the program of participatory management of forest resources. Under this program villagers are encouraged to organize village forest committees (VFC) to promote natural regeneration of forests. The VFCs are given certain authority to control forests, an authority that was earlier monopolized by the Forest Department, and are expected to assume responsibility for forest protection. In return they are given full rights over non-timber forest produce such as sal leaves and oil seeds, as well as a share of any timber that may be harvested. This program, which was initiated in 1973 in the tribal villages of Midnapore district (22°20′ N lat.–87°20′ E long.) of West Bengal, has now spread over many other parts of this tribal belt, as well as elsewhere in the country. The program has by and large been a success. It is an important instance of the poor getting organized to promote ecological restoration, and in the process enhancing their own standard of living. Such ecological restoration also tends to enhance levels of biodiversity in conjunction with production of a variety of forest produce of value to local people (Poffenberger and McGean, 1996).

The Jharkhand political struggle also incorporates an issue with significant implications for biodiversity conservation. The tribal religious practices throughout the country included protection of sacred plants, animals, ponds, and forests, as was mentioned earlier in the Gangte case study. The system of sacred groves, called "saranas," has been an important ingredient of the traditional tribal religion of Jharkhand. These saranas are variable in size—mostly small, about 0.1 ha—but very numerous and serve to protect samples of original natural plant life. These were treated with contempt by Christian missionaries who have a significant influence in this tribal region, as also by proponents of high-caste Hinduism. Now protection of these saranas has emerged as an element of the move to restore tribal self-respect and to empower them (Gokhale *et al.*, 1998). Indirectly then, conservation of indigenous biodiversity is beginning to emerge as a part of the agenda of the poor to pull themselves out of the poverty trap.

UNESCO's Biosphere Reserve program is another endeavor to combine efforts at conservation of biodiversity with alleviation of poverty. However, it has met with limited success largely because of inadequate understanding of local ecology and people's needs. Thus, Nandadevi (–30°30′–30°40′ N lat. and 79°44′–79°58′ E long.) in Western Himalayas is one of India's eight Biosphere Reserves. In this region, the traditional livelihoods depended on summer grazing in these pastures supplemented by collection of medicinal herbs. This grazing has been banned in the core zone of the biosphere reserve, along with mountaineering expeditions, another important source of cash incomes. As a result, local people feel significantly impoverished. The few ecodevelopment programs such as fuelwood plantations that have been brought in as a part of the Biosphere Reserve activities have failed to create sufficient levels of additional incomes. At the same time local people believe that in the absence of grazing the diversity of medicinal herbs in the pastures are now being replaced by extensive growth of a few species such as Rumex (Negi, 1999).

The philosophy behind biosphere reserves, that of combining conservation with development efforts, also motivates a number of development programs that go by the name of "Integrated Conservation and Development Programs" (ICDP) and the many projects funded by Global Environment Facility (GEF). However, there has been little genuine progress so far in actually realizing these goals. As a review of a GEF program in New Guinea documents, the programs remain ineffective as they continue to be designed by outsiders committed to conservation but with little understanding of what would motivate local people, especially the poor, to participate in conservation efforts (Global Environment Facility, 1998).

There are also attempts to bring to local people financial benefits either through ecotourism or through sustainable harvesting of biodiversity resources. There are limited success stories of combining the two in parts of Zimbabwe where substantial incomes are generated though high-priced hunting licenses for large mammalian game animals, with levels of such harvests kept well within sustainable limits (McNeely, 1995). These incomes are shared with local, relatively poor people. Other attempts at combining poverty alleviation with biodiversity conservation have focused on promoting enterprises such as processing of medicinal plants or wild fruit. Programs in India involving processing of fruit of *Phyllanthus emblica*, a moist deciduous forest tree species by local Solliga tribals in B. R. hills of Karnataka or processing and marketing of wild mango by local women's groups in the Kangra district of Himachal Pradesh have met with varying degrees of success (Chopra, 1998). Providing alternative means of livelihood to communities depending on a threatened biological resource has also been attempted with success in the Indian ocean island of Seychelles where turtle shell artisans were provided financial and technical support to shift to occupations like coconut souvenir carving (Global Environment Facility, 1997).

The Challenge Ahead

Poverty thus relates to biodiversity in many, complex ways. But, by and large, the poor are today instruments of decimation of biodiversity in their attempts to earn livelihoods. There has been little success in addressing this problem through designing programs that would combine conservation and sustainable utilization of biodiversity with alleviation of poverty. This then is a major challenge that must be addressed in years to come.

See also: Biodiversity as a Commodity. Biodiversity-Rich Countries. Deforestation and Land Clearing. Ethnobiology and Ethnoecology. Market Economy and Biodiversity. Social and Cultural Factors

References

Bhattacharya S (1993) *Ecological Organization of Indian Rural Populations*. Ph.D. thesis. Bangalore: Indian Institute of Science.

Chopra K (1998) Economic aspects of biodiversity conservation. *Economic and Political Weekly* XXXIII(52): 3336–3340.

Dasmann RF (1988) Towards a biosphere consciousness. In: Worster DW (ed.) *The Ends of the Earth: Perspectives on Modern Environmental History*, pp. 277–288. Cambridge: Cambridge University Press.

Gadgil DR (1942) *Industrial Evolution of India*. New Delhi: Oxford University Press.

Gadgil M (1979) Hills, dams and forests: Some field observations from Karnataka Western Ghats. *Proceedings of the Indian Academy of Sciences* 2: 291–303.

Gadgil M (1995) Prudence and profligacy: A human ecological perspective. In: Swanson TM (ed.) *The Economics and Ecology of Biodiversity Decline*. Cambridge: Cambridge University Press.

Gadgil M (1996) Managing biodiversity. In: Gaston KJ (ed.) *Biodiversity: A Biology of Numbers and Difference*, pp. 345–365. Oxford: Blackwell Science.

Gadgil M (1998) Conservation: Where are the people? *The Hindu: Survey of the Environment* 98: 102–137.

Gadgil M and Guha R (1992) *This Fissured Land: An Ecological History of India*. Berkeley, CA: Oxford University Press, New Delhi and University of California Press.

Gadgil M and Guha R (1995) *Ecology and Equity: Use and Abuse of Nature in Contemporary India*. London: Routledge.

Gadgil M, Hemam NS, and Reddy BM (1997) People, refugia and resilience. In: Folke C and Berkes F (eds.) *Linking Social and Ecological System*, pp. 30–47. Cambridge: Cambridge University Press.

Global Environment Facility (1997) *GEF Project Implementation Review*. Washington, DC: GEF.

Global Environment Facility (1998) *Study of GEF Project Lessons*. Washington, DC: GEF.

Gokhale Y, Velankar R, Chandran MDS, and Gadgil M (1998) Sacred woods, grasslands and water bodies as self-organized systems of conservation. In: Ramakrishnan PS, Saxena KG, and Chandrashekara UM (eds.) *Conserving the Sacred for Biodiversity Management*, pp. 366–396. New Delhi: Oxford and IBH Publishing.

Hemam NS (1997) *The Changing Patterns of Resource Use and Its Bio-social Implications: An Ecological Study amongst the Gangtes of Manipur*. Ph.D thesis. Calcutta: Calcutta University.

Kangle RP (1969) *Arthasasthra*. Bombay: University of Bombay.

McNeely JA (1995) Economic incentives for conserving biodiversity: Lessons from Africa. In: Bennum LA, Aman RA, and Crafter SA (eds.) *Conservation of Biodiversity in Africa. Local Initiatives and Institutional Roles*, pp. 199–215. Nairobi: National Museums of Kenya.

Negi NS (1999) *Co-variation in Diversity and Conservation Value Across Taxa: A Case Study from Garhwal Himalaya*. Ph.D thesis. Bangalore: Indian Institute of Science.

Pemberton RB (1835) *Report on the Eastern Frontier of India*, Calcutta.

Poffenberger M and McGean B (eds.) (1996) *Village Voices, Forest Choices: Joint Forest Management in India*. Delhi: Oxford University Press.

Ramakrishnan PS (1992) Shifting agriculture and sustainable development: An interdisciplinary study from North-Eastern India. In: Jeffers JNR (ed.) *Man and Biosphere Series*, vol. 10. Paris: The Parthenon Publishing Group.

Saldanha CJ (1989) *Andaman, Nicobar and Lakshadweep: An Environmental Impact Assessment*. New Delhi: Oxford and IBH Publishing.

Superintendent of Government Printing, Calcutta. (1909) Andaman and Nicobar Islands. *Imperial Gazetteer of India*, Provincial Series.

Vijayan VS (1987) *Keoladeo National Park Ecology Study*. Bombay: Bombay Natural History Society, Annual Report.

Predators, Ecological Role of

James Estes, University of California, CA, USA
Kevin Crooks, Colorado State University, CO, USA
Robert D Holt, University of Florida, Gainesville, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Apex predator An organism that occupies a food web's highest trophic level.
Bottom-up forces Population-regulating processes based on the availability of food, nutrients, and energy.
Carnivore An organism that consumes other animals.
Competitionism The view that competition regulates populations.
Food web The interconnections among organisms based on diet.
Herbivores Organisms that feed on plants and photosynthesizers.
Keystone species A strong interactor that is relatively rare.
Lotka–Volterra model An early equation relating rate of population change to the interplay between competition and predation.
Mesopredator A small to midsized predator.
Nutritionalism The view that bottom-up forces regulate populations.
Phytoplankton Microscopic primary producers that live in water column habitats.
Piscivores Predators that consume fish in aquatic habitats.
Planktivores Predators that consume zooplankton in aquatic habitats.
Top-down forces Population-regulating processes that originate from consumer limitation.
Trophic cascades A chain reaction of top-down interactions across multiple trophic levels.

Introduction

The science of ecology has undergone a succession of paradigms on the nature and importance of species interactions, including those between predators and their prey. The earliest view (henceforth termed nutritionalism) was that bottom-up forces (i.e., primary production and the efficiency of energy and material transport upward across trophic levels) regulate populations. Ecosystem ecology was built around this view of nature, which implicitly holds that apex predators, as the end points of energy and material flux, are of minor consequence to ecosystem function. Beginning in the late 1950s and early 1960s, the focus on species interactions changed to competition. In contrast with nutritionalism, competitionism holds that lateral forces within trophic levels regulate population abundance. By this view, predators are no more or less important than any other species. Top-down forces have begun to capture the attention of ecology, thereby legitimizing predators as important ecological entities. Ecologists now recognize that vital species interactions follow all three pathways (Figure 1), often simultaneously and sometimes interactively. Thus, although our focus in this article is on top-down forces generated by apex predators, understanding the ways in which predators influence biodiversity requires a more eclectic view of food webs and species interactions than simply "bottom-up versus top-down."

The authors begin with a discussion of who the predators are and how they affect populations, communities, and ecosystems. They then present a series of case studies demonstrating the wide range of systems in which predation is an important organizing process, including examples of the unifying concepts and explanations of how they were discovered. This discussion is followed by a theoretical exploration of predation and biodiversity. Next, the authors discuss the levels of biological organization at which predation can influence biodiversity and develop a conceptual

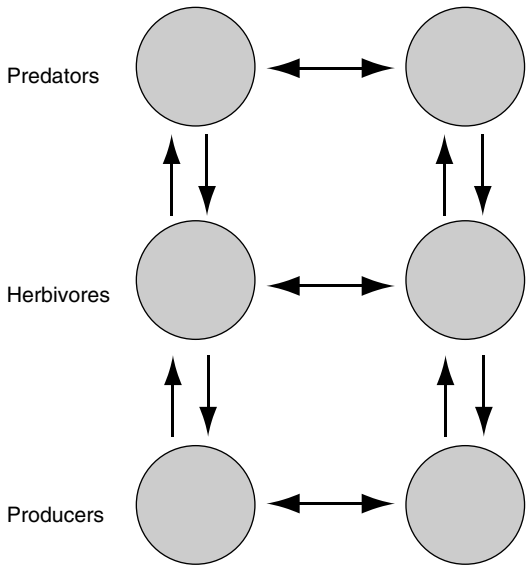


Figure 1 A simple stylized food web showing potential interaction pathways. The circles represent species, the downward-pointing arrows represent top-down forces, the upward-pointing arrows represent bottom-up forces, and the double-headed arrows represent competitive interactions. Although real food webs are far more complex, this figure shows the three main ways by which species interact with one another in nature.

model for how apex predators might influence the location and strength of bottom-up and competitive forces in systems under top-down control. The authors conclude by considering the needs and opportunities for further research on predators and how predators are likely to figure into the future of conservation biology.

Background and Definitions

What are Predators?

Most broadly defined, all consumers that utilize living tissues to the harm of another organism are predators, thus including a wide swathe of life, except for photo- and chemosynthesizers (plants and many microbes), decomposers, and detritivores. Even some plants – carnivorous species such as pitcher plants and the Venus flytrap – make part of their living as predators. Under this definition, predators would be represented by such diverse functional groups as herbivores, parasites (including microbes and parasitoids), and the immense diversity of invertebrate and vertebrate consumers that hunt and kill their prey. It can be difficult to pigeonhole predators based on simple assessments of trophic status, consumer–prey size relationships, or just about any other scheme. For example, herbivores such as elephants exist at one end of the consumer–prey size spectrum and microbes at the other, but it is uncommon for consumers in either group to kill their prey outright. Plants may even benefit from being eaten and a single act of predation by most microbes is of virtually no consequence to their prey because of the prey's immensely greater relative body size (hosts are killed not by single microbes but by enormous swarms of them). Although herbivores, microbes, and parasites are predators in this broad sense, our focus is on predators more in the sense of ordinary English, as a near-synonym for carnivore (an animal that feeds on the flesh of another animal), and in particular those that kill their prey outright and immediately on attack. Even this restricted definition includes a vast array of species and pertains to many fascinating ecological stories.

Species Interactions

The influence of predators on biodiversity depends first and foremost on direct predator–prey interactions. There is a huge amount of literature on these direct interactions. Almost any predator can, in some circumstances, directly limit the numbers or activities of its prey, and such limitation is of great practical importance in a wide range of applied ecology, such as pest control, conservation, and wildlife management. But the consequence of predation for communities and ecosystems goes well beyond these direct interactions, because prey species themselves interact with other species. Predation thus initiates proliferating chains of indirect effects, and these indirect effects are almost limitless, potentially causing populations to increase or decline anywhere in the interaction web. Because species are mobile, indirect interactions can even spill over from one ecosystem to another, at times over great distances. As will be shown below in several case studies, such indirect effects may involve long interaction chains, with broad impacts on biodiversity and ecosystem structure and function.

Indirect Effects of Predators

An awareness of the indirect effects of predators can be traced back at least to the writings of Charles Darwin, who described an interaction chain leading from cats to mice to bumble bees to clover. Similar early examples were provided by such well-known ecologists as Charles Elton and G. Evelyn Hutchinson. Hairston *et al.*'s (1960) now-classic paper (hereafter HSS) was perhaps the first effort to mold the indirect effects of predation into a conceptual model of trophic interactions and population regulation. HSS recognized four trophic groups – producers, decomposers, herbivores, and predators – and argued that although herbivores are commonly limited by predators, plants, decomposers, and predators are ordinarily limited by resources. The HSS model has weathered the test of time, along the way setting the stage for several conceptual advances, including the importance of top-down forces in population regulation and community organization, the ideas of keystone species and trophic cascades, and a generalized theory of food chain dynamics. Each of these ideas or theories is briefly explained in the following sections.

Top-Down Forces

Bottom-up forces are those passing from producers to consumers, whereas top-down forces are those passing from consumers to producers. A Special Features section in the journal Ecology, published in 1992, stimulated further interest in the relative importance of these processes, in part by pointing out that top-down and bottom-up forces need not be mutually exclusive, even though bottom-up forces are necessary for the function of all ecosystems. This realization freed ecologists to imagine diverse influences of predation in nature. A recent volume (Terborgh and Estes, 2010) has synthesized many of these pervasive and richly diverse impacts of predation on natural and human-dominated ecosystems.

Keystone Species

HSS was followed in the mid-1960s by Robert Paine's highly influential paper on food web complexity and species diversity. Paine argued that predators often selectively consume and thus limit competitively dominant species, hence enhancing species diversity by releasing their subordinates from competitive exclusion. This argument was based on three essential premises: (1) predators selectively consume the competitively dominant prey; (2) in so doing, populations of the competitively dominant species are reduced; and (3) in the absence of predation, the prey guild is limited by interspecific competition. Paine's work captured the interest of community ecologists because it linked the influence of predators to species diversity, and (perhaps most important at the time) it was supported by results from field experiments. His studies of predation by sea stars on mussel bed assemblages were conducted in the temperate rocky intertidal zone where competition for space can be extreme. This also led to two

important developments in ecology: The idea of keystone species and the intermediate disturbance model of species diversity.

The intermediate disturbance model, further refined and generalized by Joseph Connell, holds that species diversity is influenced by the intensity of disturbance (either physical or biological; **Figure 2**). The basic idea is that there is often a tradeoff between ability to withstand or recover from disturbance factors, and competitive ability. When the intensity of disturbance is very high, those species that are vulnerable to disturbance are eliminated. When the intensity of disturbance is low, competitive subordinates are excluded by dominants. These limiting conditions are relaxed at intermediate levels of disturbance, thereby elevating species diversity. The notion of keystone species, as envisioned by Paine, applied to cases in which predators were the agents of disturbance, and species were competing intensely for space. Mortality factors which clear out some competitive dominants can free up space for species which are good at rapidly colonizing open areas. Although the definition of keystone species has broadened, on the one hand, to include other kinds of interactions and grown more restrictive, on the other hand, to exclude the effects of common species, this idea is rooted historically with the indirect effects of predators. Paine's intertidal study emphasized the role of predators in regulating the diversity of their prey, but keystone species also can influence the biomass of entire trophic levels, as is seen in trophic cascades.

Trophic Cascades

A trophic cascade is the progression of indirect effects by predators across successively lower trophic levels. HSS's proposed relationship between predators, herbivores, and producers was a generalized trophic cascade. Stephen Carpenter and James Kitchell popularized this idea based on the striking influences of predatory fishes on the essential components of lake food chains – from minnows (the predatory fishes' prey) to zooplankton (prey of the minnows) and finally, to phytoplankton (prey of the zooplankton). There is growing evidence that trophic cascades occur broadly across diverse ecosystems.

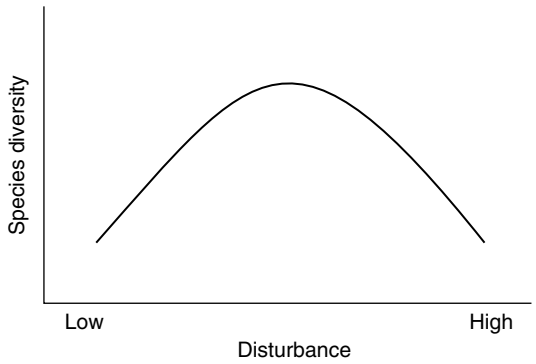


Figure 2 The intermediate disturbance model of species diversity. This model is based on the assumption that competitive exclusion occurs in benign systems. Thus, when the intensity of disturbance (including predation) is low, strong competitive interactions by the dominant species reduce species diversity. When the intensity of disturbance is high, species diversity is again low because those species that cannot cope are eliminated. Maximum species diversity occurs at intermediate levels of disturbance—strong enough to prevent competitive exclusion but not so strong as to directly eliminate species.

Generalized Food Web Theory

A generalized food web theory was developed by Stephen Fretwell to show how predation, trophic cascades, and food chain length combine to predict the strength of plant–herbivore interactions (**Figure 3**). To understand this theory, first imagine an ecosystem with producers but no consumers. Lacking consumers, the producers are limited by competition for resources. Adding herbivores creates a two-trophic level system in which the plant populations become limited by herbivory. Adding predators limits herbivore populations, thus releasing the producers from limitation by herbivory and returning them to limitation by resource competition. The progressive increase of trophic complexity cascades downward through the food chain such that plant–herbivore interactions switch from being weak to strong as the respective number of trophic levels alternates between odd and even. This story can become more complicated and the expectations are fuzzier if consumers cut across trophic levels in their diets (e.g., the fourth trophic level may also consume the second level), but its essential features do show up in many ecosystems.

Next, the authors summarize some of the case studies that provide empirical evidence for these theories and concepts. Their presentation is organized around the various major biomes in which the work was done.

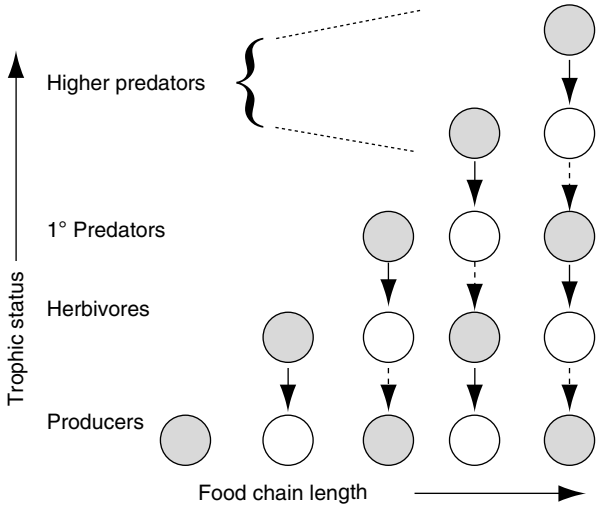


Figure 3 A graphical synopsis of Stephen Fretwell's theory of food chain length in systems under top-down control. The circles represent species or groups of species within particular trophic levels:•, resource limitation; ○, consumer limitation. The solid arrows represent strong interactions, and the dashed arrows represent weak interactions. As food chain length becomes progressively longer, the plant–herbivore interactions alternate between being weak in odd-numbered systems and strong in even-numbered systems.

Case Studies

Rocky Shores

Studies of rocky seashores furnish the earliest and some of the most compelling evidence for the effects of predation on communities and ecosystems. The first well-known experimental studies were done by Connell in Scotland. Connell's work focused mainly on competition between the two barnacles (*Cthamalus stellatus* and *Balanus balanoides*) and predation on these species by the whelk (*Thais lapillus*). This research showed that the upper shore limit of *Cthamalus* was set by physical factors (weather) and the lower limit by competition for space with *Balanus* and predation by *Thais*.

Shortly thereafter, Paine began his studies of predation by the sea star (*Pisaster ochraceous*) in mussel beds along the outer coast of Washington. Paine hypothesized that sea star predation limited the lower distribution of mussels in the midlittoral zone. This was subsequently confirmed by downward expansion of the mussel bed when the stars were removed. California mussels (*Mytilus californianus*) are the competitive dominants in this system. Predation by sea stars prevents mussels from controlling primary space (the rock surface), the principal limiting resource in this system, thus permitting other species to coexist within mussel beds. In the absence of predation by sea stars, mussels dominate space, thereby excluding the competitive subordinates (made up of many species, including barnacles, etc.) and reducing species diversity to a musseline monoculture. Subsequent research has confirmed a similar role for *Pisaster* elsewhere in western North America and for other species of mussels and sea stars elsewhere in the world.

Other predators also influence rocky intertidal communities. Work by Philip Hockey and colleagues demonstrated a trophic cascade among African black oystercatchers (*Haematopus moquini*), herbivorous limpets, and intertidal algae. Further studies of oystercatchers and limpets have confirmed similar interactions in South America, Australia, and western North America. Research in central and southern California further demonstrated how humans perturb the trophic cascade by exploiting owl limpets (*Lottia gigantea*, a large, territorial species) and by causing oystercatchers to abandon their breeding territories. The former effect causes a competitively subordinate guild of small limpets to replace owl limpets as the principal herbivore. The latter effect, induced simply by large numbers of humans beings present along rocky shores, transforms the intertidal community from a three- to two-trophic level system. Small limpets come to dominate such areas, in turn reducing the algal cover. These human-caused perturbations probably are responsible for much of the modern-day character of rocky shores in central and southern California.

A final example of predation on rocky shores concerns the loco (*Concholepas concholepas*), a large muricid gastropod that consumes intertidal mussels and is exploited by humans in central and southern Chile. Juan Carlos Castilla excluded humans from a small stretch of shoreline at the Las Cruces Marine Laboratory near Santiago in order to better understand their influence on this system, and as expected loco abundance increased. The more surprising result was a shift in the intertidal landscape, from one dominated by extensive mussel beds to one largely devoid of mussels. This particular example is noteworthy because it demonstrates how humans can perturb predator-mediated interaction chains with landscape-level consequences, and that reserves can be used effectively both to demonstrate and to mitigate such effects.

Kelp Forests

Kelp forest communities provide numerous examples of the ecological role of predators. Perhaps the most well known of these is that of the sea otter, which was hunted to near extinction in the Pacific maritime fur trade. Following protection in the early 1900s, the process of recovery created a fragmented population distribution within what had been a continuously occupied range. Contrasts between areas with and without sea otters revealed striking differences in kelp forest communities. Areas with sea otters supported lush kelp forests, whereas those without otters were extensively overgrazed by sea urchins, the otter's principal prey. These patterns result from a trophic cascade, driven by sea otter predation on sea urchins, thus releasing kelp beds from sea urchin grazing.

In addition to affording early empirical support for HSS, the sea otter–kelp forest system provides evidence for a wide range of predator-driven effects beyond those expected from simple trophic cascades. Sea otters influence numerous species by enhancing kelp abundance, thereby adding three-dimensional habitat and fueling increased primary production. This process is especially noteworthy because it shows how bottom-up processes can be altered by the top-down forces of apex predators. Other known or suspected consequences of sea otter predation in kelp forests are summarized in **Figure 4**.

The influence of sea otter predation has several interesting historical dimensions. Faunal remains in Aleut kitchen middens show that sea urchin size distributions during most of Aleut prehistory were similar to those of modern systems lacking sea otters, thus suggesting that aboriginal humans, by limiting sea otters, influenced coastal ecosystems long before modern humans arrived on the scene. Evolutionary influences of sea otter predation appear over even longer time scales. The control of herbivores by sea otter predation has decoupled a coevolutionary arms race between kelps and their herbivores, thus resulting in the development of poorly defended plants and unresistent herbivores in the North Pacific Ocean. This hypothesis has been tested through a comparison with Australasian kelp forests, which lacked a predator of comparable influence to the sea otter and which contain well-defended macroalgae and herbivores that are more resistant to those defenses. This coevolutionary scenario may explain why Northern Hemisphere kelp forests have been so devastated by sea urchin grazing following decimation of their predators. It also may explain why Steller sea cows radiated from the tropics into the North Pacific Ocean, but not elsewhere, and why the world's most diverse and largest abalones occur in the North Pacific. The sea cows were sustained by their consumption of kelp, and abalone more diffusely benefited from the rich resources provided by the kelp environment.

The sea otter–kelp forest system has changed remarkably as killer whales entered the coastal ecosystem and began preying

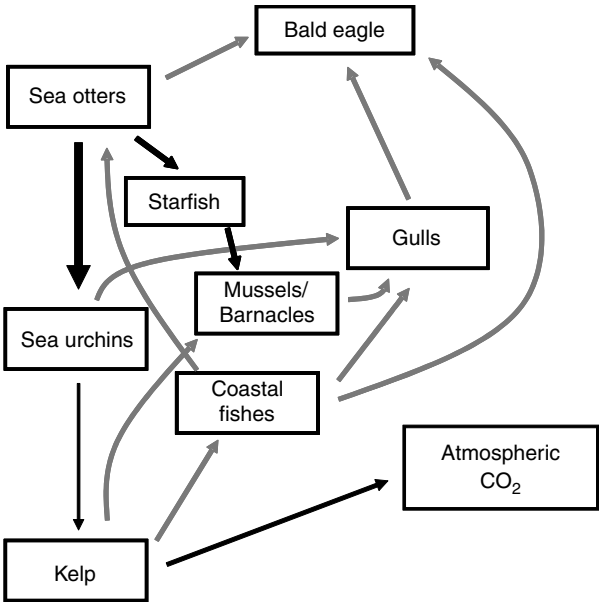


Figure 4 A conceptual representation of the direct and indirect effects of trophic cascades in the sea otter–kelp forest system. The cascade is shown on the left, with a strong limiting influence by sea otters on sea urchins (heavy arrow) and a resulting weak limiting influence of the sea urchins on kelp (lighter arrow). Some of the known or suspected indirect effects of this trophic cascade on other species and ecosystem processes are shown on the right, with the black arrows indicating top-down forcing and the gray arrows indicating bottom-up forcing.

intensively on sea otters after their normal prey populations (seals and sea lions) declined. Since the early 1990s, killer whale predation has driven otter numbers downward by approximately an order of magnitude across large areas of western Alaska. The loss of sea otters has caused sea urchin numbers to increase and kelps to decline (**Figure 5**). This example illustrates that predator–prey interactions, acting through trophic cascades, influence herbivore–plant interactions in a manner consistent with the predictions described earlier for odd- versus even-numbered food chains (**Figure 4**). It further indicates a role for predators in linking ecosystems over large areas. This ecological chain reaction may have been set off by post World War II industrial whaling, which deprived killer whales of an important food resource and may thereby have forced them to broaden their diet to include these smaller species of marine mammals.

Other species of predators initiate trophic cascades in kelp forest systems. Robert Steneck and colleagues have argued that Atlantic cod (*Gadus morhua*) in the Gulf of Maine play a similar role to that of sea otters in the North Pacific. Predation by sheephead (*Semicossyphus pulcher*, a benthic feeding fish) and spiny lobsters (*Panulirus interruptus*) in a similar manner can limit sea urchins and enhance the warm-temperate kelp forests of southern California. Kevin Lafferty has argued that the loss of predators has facilitated disease spread in sea urchins in southern California kelp forests. This involves two distinct effects. An increase in urchin density permits disease transmission to occur more readily. Moreover, ecological theory predicts that predators which eliminate individual urchins carrying a disease, thereby reduce the time over which it can distribute infective propagules. If m is the mortality rate of prey, the average lifespan of prey is 1 m^{-1}. Predators increase m, therefore reduce prey lifespan, and incidentally make it harder for infectious diseases to be sustained in the prey population. In the Southern Hemisphere, predation by rock lobsters (*Jasus lalandii*) in South Africa limits predatory whelks, in turn releasing subtidal mussel beds from limitation by whelk predation. Amos Barkai and Christopher McQuaid have argued that the loss of lobsters from these systems can lead to a predator–prey role reversal in which the hyperabundant whelks turn to feeding on lobsters, thus leading to an alternative community state that is difficult or even impossible for lobsters to reinvade. A final example from kelp forest systems comes from the work by Davenport and Anderson, who have shown that planktivorous fishes protect kelp forests in southern California from overgrazing by mesoherbivorous zooplankton.

These several examples from kelp forest systems have shown that predators can shape populations and communities on ecological timescales and life history characters on evolutionary timescales. Comparative studies of sea urchins in tropical and temperate reef systems also suggest that predators influence their prey's behavior in complex ways. In warm temperate/tropical systems, sea urchins commonly retreat to protective cracks and crevices within the reef during daylight hours in order to avoid being eaten by diurnally active benthic predatory fishes. Fishes quickly attack urchins that are experimentally removed from their refuges during the day and placed on exposed habitats. Similar patterns have been shown for a variety of warm temperate and tropical systems in which benthic predatory fishes occur. However, the nature of urchin behavior appears to differ between species whose evolutionary histories are rooted in tropical versus temperate environments. Tropical species tend to display diel sheltering as a fixed behavior, regardless of ecological context, whereas the sheltering behavior is plastic in temperate species depending on whether predatory fishes are present or absent. The explanation for this difference in plasticity may lie in the fact that tropical urchins have long been subject to predation by diurnally active fishes, whereas temperate urchins have come into contact with benthic predatory fishes more recently in their evolutionary history, and only then at the warm margins of their geographical ranges.

Lakes

Numerous reports from various lake systems throughout the world show that altered populations of apex predators result in altered food webs. There are two main reasons for the number of studies and high quality of this evidence. Lakes, as discrete and recurrent entities, are well suited for comparative and experimental studies. Furthermore, the producers and herbivores (especially phytoplankton and zooplankton) have very short generation times, thereby making population-level responses to perturbations rapid enough for scientists to observe and document.

The essential players in lake systems include four main groups of organisms: phytoplankton, herbivores, planktivores,

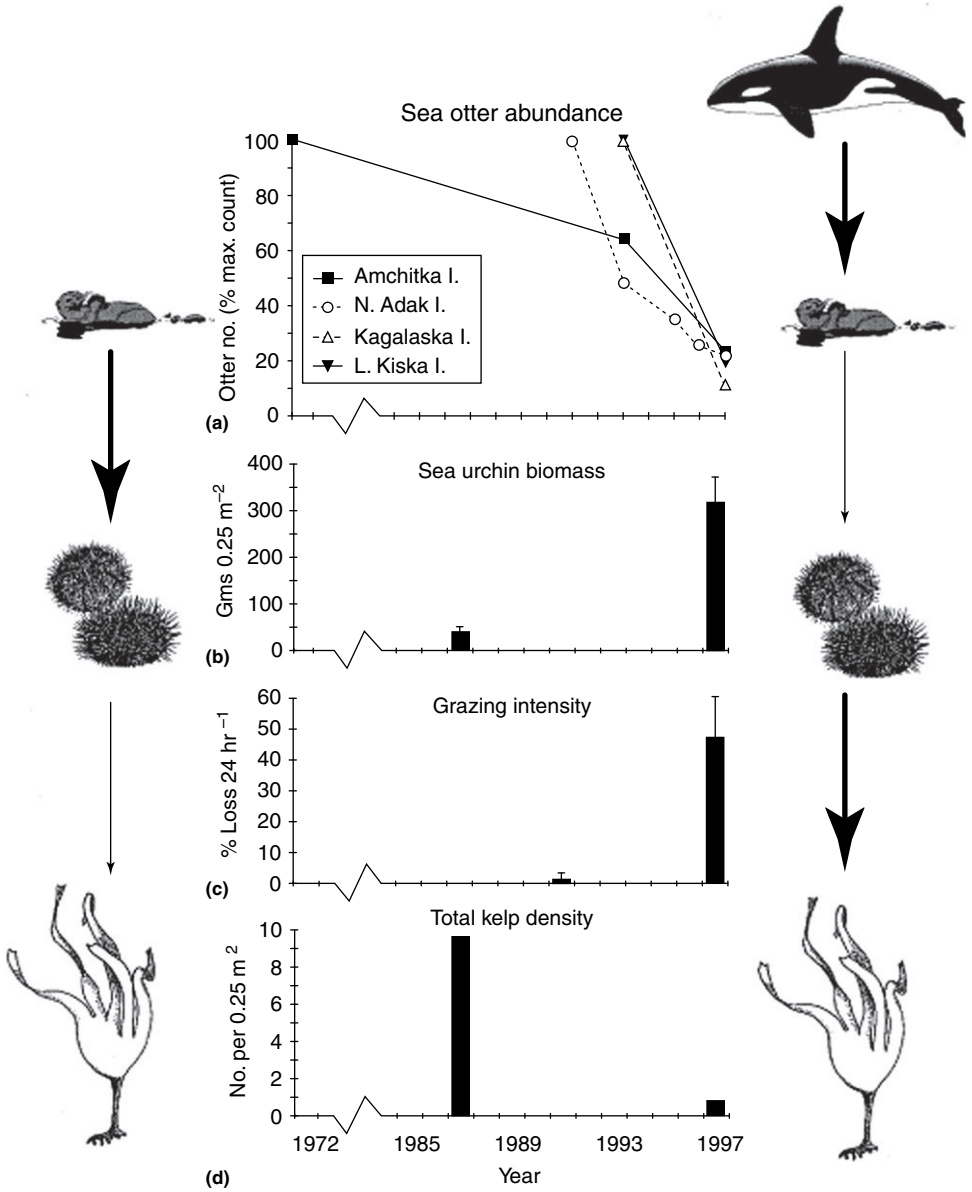


Figure 5 (a) Changes in sea otter abundance over time at several islands in the Aleutian archipelago and concurrent changes in (b) sea urchin biomass, (c) grazing intensity, and (d) kelp density measured from kelp forests at Adak Island, Alaska. Error bars in (b) and (c) indicate 1 SE. The proposed mechanisms of change are portrayed in the marginal cartoons: The one on the left shows how the kelp forest ecosystem was organized before the sea otter's decline and the one on the right shows how this ecosystem changed with the addition of killer whales as an apex predator. Thick arrows represent strong trophic interactions, and thin arrows represent weak interactions. Reproduced from Estes JA, Tinker MT, Williams TM, and Doak DF (1998) Killer whale predation on sea otters linking coastal with oceanic ecosystems. *Science* 282: 473–476, with permission from AAAS.

and piscivores. The relationship of the first three of these to piscivore abundance, explained by cascading trophic interactions, is shown in **Figure 6**. The evidence for these interactions comes from a variety of areas and approaches. Early insights were provided by contrasts among lakes in which piscivore populations varied serendipitously but for unknown reasons. There are many such examples from tropical and temperate lake systems in both the New and Old Worlds. Additional evidence that these patterns are caused by trophic cascades has come from the results of mesocosm experiments, by tracking changes associated with the fortuitous extinction or reintroduction of piscivores into particular lakes through time, and by whole-lake experiments in which the piscivores were purposely added or removed. Although the details vary depending on such factors as food chain length and the nature of particular species, the overall picture of food web dynamics in lake ecosystems is remarkably consistent, and especially highlights the ubiquitous importance of apex predators and trophic cascades in these systems.

Perhaps the earliest evidence for the influence of predation in lake systems comes from John Brooks and Stanley Dodson's analysis of New England lakes. These researchers showed that

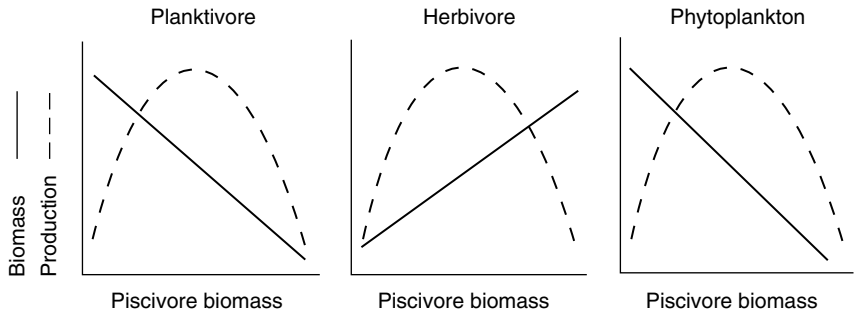


Figure 6 Piscivore biomass in relation to biomass (solid line) and production (dashed line) of vertebrate zooplanktivores, large herbivores, and phytoplankton in lake systems Reproduced from Carpenter SR and Kitchell JF (eds.) (1993) *The Trophic Cascade in Lakes*. New York: Cambridge University Press, with permission from CUP.

in the absence of planktivorous fishes, zooplankton assemblages were dominated by species with large body size because of their increased foraging efficiency and competitive superiority over small species. In lakes with planktivorous fishes, the composition of the plankton shifted toward small body size due to the influence of size-selective predation. This example was followed by Thomas Zaret and Robert Paine's report on the cascading influences of introduced peacock bass (*Cichla ocellaris*) to Lake Gatun, Panama. Peacock bass, a cichlid native to the Amazon River, were first introduced to Lake Gatun in 1965 for sport fishing and consumption. These introduced predators are voracious piscivores and they caused a remarkable series of food web effects as the bass population grew and spread across the lake. The immediate influence was a rapid and extreme reduction of planktivorous minnows, thus causing zooplankton populations, including that of mosquito larvae, to increase. This example added two interesting dimensions to the understanding of lake systems. One is the strength of influence by an exotic predator on naïve prey, with broad-ranging indirect effects across the lake and surrounding terrestrial systems. In addition to the top-down effects described previously in this section, the reduced populations of planktivorous minnows negatively impacted other apex predators, including several species of aquatic birds and predatory fishes. Another dimension is the potential impact on human health, in this case resulting from an increased threat of malaria because of increased mosquito populations. Similar examples of broad-ranging influences by exotic predators are known for many other lake and river systems throughout the world.

Lakes also provide intriguing examples of interecosystem linkages. For example, food subsidies from land to water, or in the reverse direction, can maintain populations of apex predators, thus reinforcing and even strengthening the effects of trophic cascades. Food web dynamics within lake systems can influence the surrounding terrestrial ecosystems, as demonstrated by Tiffany Knight and colleagues in Florida ponds wherein predation by fish on dragonfly larvae indirectly affects the numbers of adult dragonflies near pond margins. Adult dragonflies are voracious predators on insects, including pollinators such as bees and butterflies. The abundance of pollinators influences reproductive success in terrestrial flowering plants, so plants near ponds without fish have lower seedset than plants near otherwise similar ponds with fish. This example also shows that trophic cascades can have impacts on mutualisms, not just herbivory. Trophic cascades have even been shown to influence the net flux of carbon between lakes and the atmosphere.

Rivers and Streams

Experimental work in rivers and streams has demonstrated influences of predation on food web structure and the life history of prey populations. Both fishes and birds are important apex predators in river food webs, and like lakes, many river food webs are strongly influenced by trophic cascades. The experimental exclusion of these predators from a variety of tropical and temperate river systems by Mary Power and colleagues provided evidence for top-down forces and has revealed depth-related gradients in the outcomes of trophic cascades. These findings show that the influence of predators on food webs can be strongly influenced by prey refuges, which in turn can vary across habitat gradients. Another important contribution of the riverine studies is that they have been done in systems that deviate in food chain length, thus providing the first experimental evidence that the strength of plant–herbivore interactions varies predictably between odd- and even-numbered food chains. Work by Andrew Sih and colleagues in streams of the eastern US also shows how the risk of predation can influence the interplay between feeding and reproductive behavior in several prey species. Barbara Peckarsky and colleagues have demonstrated similar phenomena in high mountain streams in the western US.

As described above for lake systems (*see* Case Studies: Lakes), there is growing evidence for diverse linkages between streams and adjacent terrestrial ecosystems. Both terrestrial and aquatic insects have been shown to affect food web dynamics in their respective adjacent aquatic or terrestrial habitats. Fausch and colleagues have experimentally documented that invasion of nonnative salmonids interrupts reciprocal flows of aquatic and riparian invertebrates, with effects that propagate within and across stream and terrestrial food webs. Nutrient subsidies from migratory salmon strongly influence the productivity of riverine and adjacent terrestrial ecosystems. And finally, terrestrial trophic cascades have major influences on both the hydrology and community composition of riverine systems in the western US, as exemplified by the work of Robert Beschta and colleagues.

The pioneering work of John Endler and David Reznick on Trinidadian guppies provides some of the strongest and most comprehensive evidence for the selective effects of predation. The risk of predation to guppies by various larger fish species varies within and among streams, and a variety of life history characters, color patterns, and features of guppy mating systems co-vary accordingly. By manipulating predator populations and translocating guppies among habitats, these researchers demonstrated rapid selective responses to altered risks of predation.

Tropical Reefs

Numerous recent studies document strong and far-reaching influences of predatory fishes in tropical reef systems. The exclusion of various predatory fish species has repeatedly been shown to cause population increases, changes in behavior, or changes in species composition of their prey guild, which in most cases are smaller fishes. Surveys of reefs around the world have shown that all but a few remote places now lack normal abundances of sharks and other large consumer species (such as groupers and wrasses) because of human exploitation. These reduced predator populations have led in turn to significant indirect effects that in some cases have led to phase shifts, including the collapse of corals. For example, the progressive exploitation of predatory and herbivorous fishes caused a large-scale phase shift from coral to algal-dominated systems around the island of Jamaica. In the tropical Pacific, outbreaks of the crown of thorns starfish (a coral predator), which in turn have devastated coral reefs, are now thought to be, at least in part, the result of reduced predation by benthic feeding fishes on the larval and juvenile stages of the sea stars.

Oceanic Systems

Although the oceans dominate our biosphere and provide critical ecosystem services in such diverse forms as food production and climate control, relatively little is known about the role of apex predators in the open sea. One reason is that the open sea and its associated seafloor habitat present serious logistical challenges to studies of any kind. Furthermore, many ocean ecologists still embrace the view that bottom-up processes are the main drivers of biological pattern in ocean ecosystems. Although bottom-up forcing in the sea is clearly important, this does not preclude top-down effects, which might be expected for several reasons. One is that strong predator-induced effects occur broadly in lakes and the general structures of ocean food webs (from phytoplankton to zooplankton to planktivores–piscivores) are similar to those of lakes. A second is that nowhere else on the planet are predators so abundant, as witnessed by the vast schools of marine mammals, seabirds, and predatory fishes, which still can be seen in places despite millennia of depredations by humans, and which historically were even more abundant. Despite these a priori reasons to expect cascades in the oceans, there are few examples of oceanic trophic cascades. One of these involves pink salmon (*Oncorhynchus gorbuscha*) populations in the North Pacific that fluctuate on a 2-year cycle. During years when pink salmon are abundant, zooplankton are depressed and phytoplankton are abundant, whereas during years when pink salmon are rare, zooplankton are abundant and phytoplankton are relatively rare. Similar patterns have been documented following the reduction of cod in the North Atlantic Ocean and North Sea, and by the wholesale reduction of various predatory fishes in the Black Sea. In the Southern Ocean, the reduction of blue, fin, sei, and minke whales by industrial whaling may have released krill populations from limitation by predation, in turn elevating the carry capacities of other krill-feeders – pinnipeds, penguins, and perhaps additional groups of consumers. Increased growth rates and reduced age of first reproduction of seals and whales after the depletion of great whales from Antarctica have been interpreted as evidence for such effects. Changes in the isotopic composition in adelie penguin egg shell remains indicate that these birds were piscivores during most of the past 8000 years, shifting to a krill diet only after the onset of industrial whaling in the Southern Ocean early in the twentieth century.

Predation by gray whales (*Eschrictius robustus*) and walruses (*Odobenus rosmarus*) has important effects on seafloor systems. Gray whales influence these systems by resuspending sediments and consuming ampeliscid amphipods. Furrows formed by the whales in the soft benthos are colonized by scavenging lysianassid amphipods, serve to accumulate detritus, and thus facilitate a local detritus-based food web. Walruses further impact these systems by consuming clams and other large infauna, in turn attracting predatory and detritivorous sea stars.

It is also becoming evident that sharks play important roles in many marine systems, especially at mid to lower latitudes. Ransom Myers and colleagues chronicled an irruption of small elasmobranches and a resulting collapse of their bivalve prey following the reduction of great sharks from overfishing in the eastern US. Changes in the behavior of dugongs, sea turtles, and small cetaceans have been attributed to seasonal migrations of large sharks in Western Australia.

Behavioral patterns of various marine prey species also suggest strong predator–prey interactions. For example, krill and other large zooplankters typically undergo diel vertical migrations that take them beyond the foraging range of marine birds and mammals during daylight hours. Many species of forage fish and zooplankton form dense swarms, which probably reduce their likelihood of being consumed by predators that must search for and capture individual prey. Pagophillic (ice-loving) pinnipeds in the Arctic and Antarctica behave in fundamentally different ways in apparent response to whether the risk of predation is from above or below the ice. In the Arctic, where polar bears and humans are important predators, pinnipeds flee from the ice to water at signs of danger. In Antarctica, where the threat of predation is much greater in the water (from killer whales and leopard seals) than it is on the ice, the pinnipeds do not display such extreme flight behavior and often are nearly oblivious to potential disturbances when hauled out.

The authors suspect that more evidence of trophic cascades in oceans will emerge as oceanographers begin to incorporate it into their world view, particularly as the historical record is mined for evidence about ocean conditions prior to human impacts.

Boreal/Temperate Forests

Although terrestrial biotas of the New World once contained numerous large mammalian carnivores, the potential ecological significance of these predators was unknown until recently. There are many reasons for this prolonged state of ignorance. One is that the largest of these creatures – gray wolves and grizzly bears – were all but exterminated in the US and Mexico well before modern ecological research had taken form. Second, even if large carnivores have been able to persist in the face of direct persecution, they are extremely difficult animals to study due to their low densities, nocturnality, secretive habits, aggressive behavior, and wariness of humans. Third, just as ocean ecologists have downplayed the importance of predators in the open sea, many wildlife ecologists have tended to be skeptical about the importance of predation in population regulation, and this topic has been hotly debated in the wildlife literature. In a famous example, D. Irvin Rasmussen, Aldo Leopold, and then HSS attributed an irruption of mule deer on the Kaibab Plateau, subsequent overgrazing, and the eventual mass starvation of the deer herd to the extermination of gray wolves and other large predators, but Graeme Caughley later attempted to debunk this explanation. Fourth, the long generation times of key players (decades to centuries for trees; multiple years to decades for ungulates and carnivores) and the large areas required for the measurement or manipulation of their representative populations make rigorous study of the top-down effect of apex predators very challenging. Finally, political, social, ethical, and legal issues have dissuaded many scientists from studying large mammals.

Despite these difficulties, there is growing evidence for top-down effects by large predators in temperate and boreal forests. Brian McLaren and Rolf Peterson used historical information on wolf and moose abundance, together with growth ring measurements from balsam fir, as evidence for a trophic cascade at Isle Royale in Lake Superior (**Figure 7**). Wolf numbers have fluctuated substantially throughout the twentieth century, apparently in large measure because of demographic factors related to their small population size. Inverse changes in moose numbers followed wolf population fluctuations, thus suggesting regulation by wolf predation. Direct measures of changes in herbivory were unavailable. However, the distance between annual tree rings in balsam fir showed that sapling growth rates were lower when moose were abundant than when moose were rare.

William Ripple and colleagues have shown similar patterns elsewhere in North America, especially in the intermountain west. Tree ring analyses have demonstrated long-term recruitment failures of various tree species following the loss of wolves and other large predators to control programs in the early 1900s. Subsequent experiments in which deer and elk were excluded by fencing have established herbivory by large ungulates as a major cause of the tree recruitment failures. The reintroduction of wolves to Yellowstone National Park has contributed to the reversal of these patterns in some areas. Ripple and colleagues suggest that wolves have allowed for aspen recovery through a behaviorally mediated trophic cascade, with elk avoiding foraging in risky places, especially in riparian zones where they may be especially vulnerable to wolf predation. Recently, Kauffman and colleagues conducted a landscape-level test of this hypothesis, and concluded that the cascading effect of wolves is likely mediated by wolf-induced reductions in elk populations and not by behavioral mechanisms of risk avoidance by elk. Wolf-induced trophic interactions have resulted in other effects on the ecosystem. For example, the loss of riparian vegetation has caused streams to channelize, in turn reducing water table levels. Reductions in the diversity of small vertebrate species, especially passerine birds, have accompanied these changes. The influences of wolves occur via other pathways as well. In the greater Yellowstone ecosystem, the reintroduction of wolves appears to have led to the reduction of coyotes and a resulting increase in fawn survival of pronghorn antelope.

Fragmented Coastal Scrub Habitats

Although experimental manipulation of terrestrial carnivores is exceedingly difficult, fragmented habitats can provide valuable, large-scale, ecological experiments to rigorously explore the top-down effects of mammalian predators. Large carnivores are particularly vulnerable to local extinction in fragmented landscapes due to their large ranges and resource requirements, low population densities, and direct persecution by humans. Larger carnivores can depress populations of smaller mammalian carnivores, or "mesopredators," through direct predation, resource competition, and interference competition, including spatial and temporal avoidance. Consequently, the decline and disappearance of dominant carnivores in fragmented systems may lead to the ecological release of smaller predators that in turn threaten birds and other vertebrates. This process has been labeled "mesopredator release" and the phenomenon has been documented across a wide range of ecological communities, as chronicled in recent reviews by Laura Prugh, Justin Brashares, and colleagues.

In coastal southern California, intensive urbanization during the past century has destroyed most of the native coastal sage scrub and chaparral habitats, creating a "hot spot" of endangerment and extinction in the region and leaving undeveloped canyons dissecting coastal mesas to function as habitat islands immersed within a matrix of inhospitable urban habitat. Michael Soulé and colleagues proposed the mesopredator release hypothesis as a possible mechanism to explain the rapid disappearance of scrub-breeding birds from the habitat fragments. They predicted that the decline of the most common large predator (coyote) would result in the ecological release of native (striped skunk, raccoon, and gray fox) and exotic (domestic cat and opossum) mesopredators, and that increased predation by these particularly effective avian mesopredators would result in higher mortality and extinction rates of scrub-breeding birds.

To test this prediction, Crooks and Soulé exploited a serendipitous ecological experiment – spatial and temporal variation in the distribution and abundance of coyotes among these urban habitat fragments – to investigate direct and indirect effects of this top predator on community structure. In accordance with the mesopredator release hypothesis, lower visitation rates of coyotes in small, isolated remnants resulted in elevated numbers and activity of urban mesopredators.

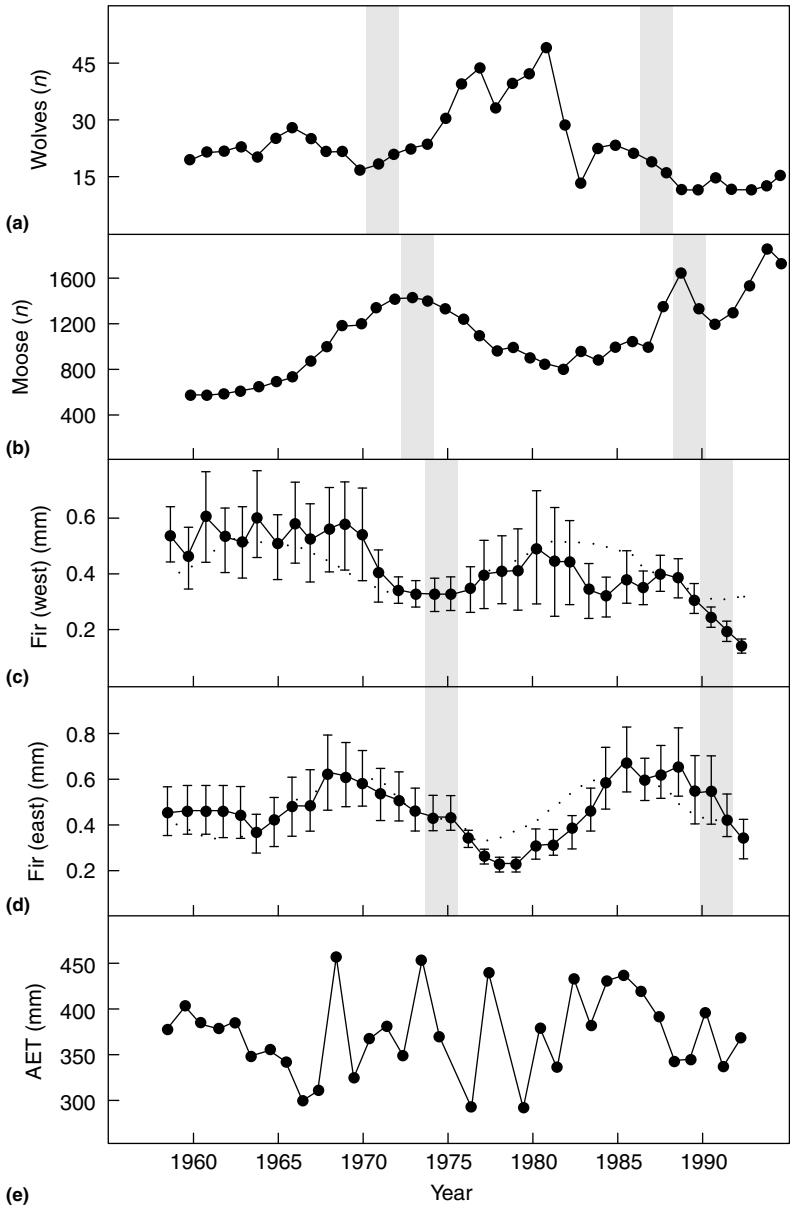


Figure 7 The trophic system on Isle Royale, reconstructed for 1958–1994. (a) Wolf abundance calculated from aerial surveys; (b) moose abundance calculated from skeletal remains and aerial surveys; mean ring-width indices for balsam fir from the west (c) and east (d) ends of Isle Royale; and (e) actual evapotranspiration from April to October, a measure of water availability during the growing season. The shaded areas highlight intervals of forage suppression that the authors believe are closely tied to periods of elevated moose density, which in turn follow periods of low wolf density. Reproduced from McLaren BE and Peterson RO (1994) Wolves, moose, and tree rings on Isle Royale. *Science* 266: 1555–1558, with permission from AAAS.

Coyotes directly preyed on some mesopredator species, particularly domestic cats. Mesopredators temporally avoided coyotes as well. In fragments that coyotes visited episodically during the course of the study, mesopredator activity increased when coyotes were absent. As predicted, scrub bird diversity was lower in fragments with fewer coyotes and more mesopredators, even after accounting for the positive effect of fragment area and the negative effect of fragment age on bird persistence.

The top-down effect of coyotes on cats seems to have had the strongest impact on the decline and local extinction of scrub-breeding birds in the urban fragments. Unlike wild predators, domestic cats are recreational hunters. Maintained well above carrying capacity by nutritional subsidies from their owners, they continue to kill even when prey populations are low. Using data on cat densities and predation rates, Crooks and Soulé estimated that cats surrounding a moderately sized fragment return approximately 840 rodents, 525 birds, and 595 lizards to residences per year. Such high levels of predation are likely unsustainable for many small vertebrate populations. For example, existing population sizes of some birds do not exceed 10 individuals in small to

moderately sized fragments, so even modest increases in predation pressure from mesopredators may quickly drive native prey species to extinction. Local extinctions of scrub-breeding birds are frequent and rapid; at least 75 local extinctions may have occurred in these fragments during the past century.

Trophic cascades generated by the decline of coyotes were embedded within a network of complex fragmentation effects that also exerted considerable bottom-up constraints on both predator and prey populations. Research in the San Diego canyons have demonstrated that a variety of taxa, including carnivores, rodents, birds, and invertebrates, require native habitat to persist, highlighting the importance of bottom-up limitations, including resources available for movement, foraging, cover, and shelter. In a study of reproductive success of scrub-breeding birds along a fragmentation gradient in San Diego, Douglas Bolger and colleagues concluded that the strength of top-down forces varied with characteristics of predator and prey and their species-specific responses along the urban gradient. Overall, this system again emphasizes how top-down and bottom-up forces work in concert to influence species diversity in terrestrial systems.

Midcontinent North American Prairies

The top-down effect of coyotes is evident in other systems as well. Historically, coyotes were generally confined to open plains and arid regions of western North America. The eradication of wolves from most of the continental US in the late 1800s and early 1900s likely facilitated the expansion of coyote populations. Currently, predator control efforts in the US are directed primarily at coyotes, and lethal and nonlethal measures have resulted in at least temporary coyote declines in some areas. In the Prairie Pothole Region of the North American midcontinent, coyote population reduction is considered one of the principal causes of increases in red fox populations starting in the 1930s; other carnivore species have been shown to increase in areas of coyote control as well. Although coyotes will kill red foxes, spatial and temporal avoidance, behavioral exclusion, and territorial shifting are likely the primary mechanisms by which coyotes reduce fox populations.

Red foxes are the most important predators on nesting ducks and their eggs and offspring in the prairie region. The ecological release of red foxes following coyote control has therefore resulted in increased predation on ground-nesting dabbling ducks, most notably mallards. Predation accounts for more than 70% of nest failures in these duck species, and in some areas intense predation on eggs, ducklings, and hens has been sufficiently intense to depress recruitment below replacement levels. Predation has resulted not only in population declines of duck species in the Prairie Pothole Region but also in altered population composition and skewed sex ratios. Similar processes are now thought to influence populations of certain upland game birds, including sage grouse.

Interestingly, since the mid-1970s, coyote populations have begun to rebound in parts of the Prairie Pothole Region due to restrictions in control and fur harvest methods and to reduced commercial value of fur. In these areas, expanding coyote populations have contributed to reduced red fox activity and higher duck nest success; overall nest success in sites where coyotes are the principal canid is nearly twice as high as that in areas where red foxes dominate. Overall, the excellent series of studies by Alan Sargeant and colleagues in the Prairie Pothole Region highlight the changes in the canid predator assemblage in North America during the past century and emphasize the top-down community-level effects generated by these dynamic canid–canid interactions.

Tropical Forests

Some of the most dramatic evidence for top-down control by apex predators comes from research by John Terborgh and colleagues in New World tropical forests. Terborgh's vision of top-down control in this system stems from a contrast between Barro Colorado Island in the Panama Canal, and Cocha Cashu Biological Station in Peru's Manu National Park. Although the two sites are similar in climate and native biota, Barro Colorado Island, because of its small size and isolation from other forest habitat, lost its apex predators (jaguars, pumas, and harpy eagles) shortly after construction of the Panama Canal. Barro Colorado Island currently supports notably higher densities of herbivorous mammals, such as agoutis, coatimundis, sloths, and howler monkeys, than does Cocha Cashu – differences attributed to the loss of predators from Barro Colorado Island. These ideas were put to a more rigorous test by using recently formed habitat fragments – the islands of Lago Guri in Venezuela – as a large-scale ecological experiment. The Caroni Valley of east-central Venezuela, once a vast, unbroken forest, was substantially altered by the 1986 creation of a hydroelectric impoundment. Within this 120-km-long by 70-km-wide reservoir, Lago Guri, the emergent hilltops became islands – isolated fragments of tropical forest that varied in size and distance from the shoreline border of unbroken forest. Although the larger islands retain nearly complete vertebrate faunas, the smaller islands lost up to 90% of the native vertebrate species, including all of the large vertebrate predators. Resulting changes in the forest system have been swift and sensational. Populations of herbivore species such as leaf-cutter ants, howler monkeys, iguanas, and rodents (all seed predators or herbivores) have increased from one to three orders of magnitude. Indirect impacts on producers have been equally dramatic. Fewer than five of approximately 6070 native tree species are continuing to successfully reproduce, thus suggesting that highly impoverished floras will result from the loss of predators. Trophic shunts – the redirection of energy and material flux through the interaction web – have influenced populations of arboreal insects and insectivorous birds.

Justin Brashares and colleagues are discovering equally dramatic and wide-ranging effects of predators in Gabon. Long-term monitoring records by game rangers in several parks establish a gradual erosion of large predators, in this case lions, leopards, spotted hyaenas, and wild dogs. Olive baboon populations have, in response, grown in abundance and expanded their diet (from one that was largely herbivorous and frugivorous to one that is now omnivorous and includes many smaller vertebrates). These changes appear to be causing not only a reduction in the abundance and diversity of

smaller birds and mammals, but also significant effects on the welfare of local humans. The increasing contact between baboons and people is facilitating the spread of disease. The baboons are also creating conflicts over agricultural resources, in turn causing villagers to hold their children out of schools in an effort to help defend their crops.

These examples illustrate two important points. First, predators often exert crucial roles in maintaining local species diversity. Terborgh's data indicate that the majority of tree species will eventually be lost from these islands systems, largely because of the loss of predators; Brashares' findings show similar losses in diversity of small vertebrates. These examples, together with the Crooks–Soulé study of fragmented coastal scrub habitats, also serves to remind us of the power of ecological experiments and the importance of scale (space and time) in designing studies on the role of apex predators in nature. It is highly unlikely that any amount of study of unperturbed terrestrial systems could have demonstrated the magnitude and breadth of effect by relatively rare apex predators. However, anthropogenically disturbed systems, such as fragmented landscapes, offer unique opportunities to understand the complex trophic interactions generated by large carnivores.

Exotic Predators on Islands

Perhaps nowhere is the top-down effect of predation on biodiversity so apparent as with the introduction of nonnative predators onto islands. Islands typically support few large predators and grazers, and apex predators such as mammalian carnivores are often absent. Consequently, insular endemic species regularly evolve in the absence of predation and thus lack adequate antipredator defenses. For example, many insular animals exhibit tame or fearless behavior that increases their vulnerability to introduced predators, and some island birds are flightless ground nesters. Similarly, many island plants do not produce the same noxious chemicals or physical defenses found in their closest mainland relatives – features that discourage herbivory.

Consequently, the introduction of nonnative predators can be catastrophic for sensitive island communities, to the point of driving insular prey species to extinction. The examples are numerous. Worldwide, of all the species that have gone extinct since 1600, 90% of the 30 species of reptiles and amphibians, 81% of the 65 mammal species, and 93% of the 176 species and subspecies of birds have been insular forms. Predation by introduced animals has been a primary cause for approximately 40% of the extinctions of birds on islands, and alien predators are endangering approximately 40% of currently threatened insular bird species. Introduced rats and domestic cats are notorious killers. Introduced rats have successfully invaded at least 80% of the world's 123 major island groups and are known to prey on a variety of insular vertebrates, including amphibians, reptiles, mammals, and birds. Indeed, predation by Pacific rats (*Rattus exulans*), black rats (*Rattus rattus*), and Norwegian rats (*Rattus norvegicus*) has been documented on at least 15, 39, and 53 insular bird species, respectively, and introduced rats are thought to be responsible for 54% of insular bird extinctions caused by predators. For example, the introduction of black rats on Big South Cape Island, New Zealand, in 1964 resulted in the rapid extinction of five species of land bird, one species of bat, and an unknown number of invertebrates, including a species of large flightless weevil. On the Galapagos Islands, introduced black rats have reduced populations of the giant tortoise and the dark-rumped petrel by preying on eggs.

Mongoose and domestic cats have also been introduced to islands, at times deliberately to control nonnative rodents. Unfortunately, they instead often eradicate native prey species. On Hawaii, introduced mongoose had little impact on rodents but decimated flightless rail populations. Domestic cats have been accidentally or deliberately introduced to at least 65 island groups and are thought to be responsible for 26% of insular bird extinctions caused by predators. Incredibly, 375 cats on Macquarie Island near Australia killed an estimated 56,000 rabbits and 58,000 ground-nesting sea-birds each year. On sub-Antarctic Marion Island, 5 cats were introduced as pets in 1949 and by 1975 approximately 2000 cats were killing 450,000 burrowing petrels annually and were suspected of driving another petrel species to local extinction. In the mid-1900s, approximately five cats were introduced to Kerguelen Island in the sub-Antarctic and their descendants have killed more than 3 million petrels per year and are responsible for the extinction of several bird populations. In the most infamous and perhaps most extreme example known, the lighthouse-keeper's pet cat on Stephen Island, a tiny island off New Zealand, arrived in 1894 and within a single year it is thought that this one cat exterminated the flightless Stephen Island wren.

The effects of introduced predators extend far beyond their prey species and can include modification of ecosystem-level processes. For example, the New Zealand flatworm (*Artioposthia triangulata*) was accidentally introduced to the British Isles in the early 1960s and, with no natural predators, has spread rapidly. The flatworm is a voracious predator of native earthworm species. Earthworms, through their burrow excavation and casting activity, provide an invaluable ecosystem service by shaping the structure and hydrological patterns of soils. Flatworm infestations and the consequent depletion of earthworms alter both soil structure and hydrology. The ramifications are far-reaching and directly impact human welfare. By depleting native earthworms, the introduced New Zealand flatworm increases the risk of surface runoff and therefore the potential of soil erosion, agrochemical pollution, and flooding.

There are a growing number of examples of indirect effects of exotic predators on islands. Introduced arctic foxes (*Alopex lagopus*) in the Aleutian archipelago have decimated ground-nesting seabirds, in turn reducing the transport of nutrients from sea to land and causing the terrestrial plant community to shift from a grassland to maritime tundra. Introduced cats have similar effects on sub-Antarctic islands. In the northern Channel Islands of southern California, predation by golden eagles has caused a decline of the endemic and endangered island fox. Golden eagles, a terrestrial predator, colonized these islands from the mainland following the loss of the native bald eagles to DDT toxicity and the establishment of feral pigs on the island as a food resource.

Threatened prey species often recover, sometimes rapidly, when alien predators are controlled or eradicated on islands.

However, the removal of alien predators can also yield unexpected results. For instance, a recent eradication of introduced black rats on Bird Island in the Seychelles has resulted in a population explosion of exotic crazy ants (*Anoplolepis longipes*). Ironically, these ants are now threatening those bird colonies that the rat eradication was intended to protect. Furthermore, simulation models predict that on islands colonized by both cats and rats, elimination of cats may release rat populations, and that increased numbers of rats may actually increase predation pressures on island birds. In essence, this cascade represents another example of the mesopredator release phenomenon, with cats on islands as top predators and rats as the mesopredators. On a variation of the theme, controlling cats on islands that also support exotic rabbits may result in more rabbits, excessive grazing by these prolific herbivores, and severe impacts to insular vegetation and associated animal species; this example, therefore, directly follows the predictions from the models of HSS and Fretwell. Clearly, alien predators on islands represent a complex, unpredictable, and occasionally dramatic example of the relationship between predators and biodiversity.

Summary and Synthesis of Case Studies

It is abundantly clear from the preceding examples that the manifestations of predation in nature are dramatic and diverse, occurring at organizational levels ranging from the behavior of individuals to the dynamics of ecosystems, and on timescales ranging from ecological to evolutionary. Numerous studies show or suggest that predators influence the abundance, distribution, and population structure of their prey. Indirect effects of predation are less appreciated by scientists and the public, despite the fact that they occur broadly in nature, are important to ecosystem function, and often result in processes that benefit human welfare. Trophic cascades are the most common of known indirect effects. These may be nearly as ubiquitous in nature as the transfer of material and energy upward through food webs. In any case, ecologists should be more surprised by the absence of such top-down effects than by discoveries of new ones.

Despite extensive evidence for trophic cascades from many of the planet's major ecosystems, we know little else about their influences on overall food webs. The lake studies provide an important exception because here a relationship between the top-down effects of predation and the bottom-up effects of production has been shown. These findings indicate that predators may fuel production in odd-numbered food chains, and that maximum production across all trophic levels should be realized at intermediate intensities of predation. A corollary to this hypothesis is that the length of food chains under top-down control is necessarily limited not by production and the efficiency of energy transfer, but by the population constrictions that occur at consumer-regulated trophic levels. Such constrictions may be common in nature because intermediate intensities of predation probably occur rarely, except in highly managed ecosystems.

The theory of trophic cascades also provides guidance on where in a food web one might expect to find strong competitive interactions. Competition should be most strongly manifested within trophic levels in which populations are resource rather than consumer limited. These occur at the odd trophic levels in odd-numbered systems and even trophic levels in even-numbered systems. Although this prediction requires further analysis, it provides hope for an integrated theory of what previously has been viewed and treated as the largely unrelated processes of top-down, lateral, and bottom-up species interactions. Moreover, apex predators can be paramount factors that govern the maintenance of species diversity in many systems, as modulators of competitive interaction as well as via direct impacts on prey.

Local extinctions and invasions of predators are increasingly common and can transform ecological communities, but unless these changes are observed or known from historical records, their significance in extant ecosystems may be difficult to understand. Often, extinct interactions are difficult to infer from historical records. As seen in some of the preceding examples, predation frequently leaves a mark on species-level characters, especially behavior, life history, and morphology. Geerat Vermeij's analysis of shell damage and morphological change (ostensibly from crushing predators) in Mesozoic marine gastropods provides a good example of one such record. The sudden increase in crushing predators was responsible for what Vermeij termed the "Mesozoic marine revolution." Using this perspective, Richard Palmer confirmed that shell structures indeed reduce the incidence of attack from crushing predators by experimentally removing spires from gastropod shells. Behavioral patterns also provide clues about the role of predators, as described in the case studies for sea urchins and pagophillic pinnipeds. Other examples could be cited, but those listed are sufficient to make the point that morphology and behavior, when thoughtfully and cautiously interpreted, frequently reflect evolutionary response to predation.

Although these examples provide insight into species-level responses to predation on historical timescales, they afford little insight into the food web effects of predators. For this purpose, one might profitably examine the characteristics of producers which have emerged over evolutionary history. Because of lags in evolutionary responses, many traits of plants today may reflect adaptations to trophic regimes that have vanished. Plants can deter herbivores by modifying their morphology, demography, and chemistry. The degree to which these defensive characters exist among plant species and populations sometimes indicates the intensity of herbivory on historical timescales, even if the herbivores are no longer present. Examples of this approach are provided by studies of differences in meristem location in steppe vegetation between the eastern and western slopes of the northern Rocky Mountains, variation in the resistance of birch trees to insect herbivores in boreal forests, the susceptibility to grazing damage in marine algae across coral reef habitats, and the evolution of reduced chemical and physical defenses in insular plants in the absence of herbivory. Many broad-leaved trees in New Zealand have a remarkable pattern in their development called "heterophylly." Up to a couple of meters in height, saplings have leaves that are thin, spiny, and have other traits that botanists believe would have made them poor forage for moas. Above the height of the tallest moa, the tree leaf forms radically change, becoming broad and soft. Moas went extinct

after the Maori colonized, but the plant traits today indicate the evolutionary importance of intense herbivory in these ecosystems. (As on many islands, predators were scarce in the fauna, so moas could reach very high abundance and penetrate most habitats.) Mismatches in extant communities between the intensity of herbivory and the degree of plant resistance sometimes can be taken as evidence for recent changes in top-down regulation, as suggested from the previously described contrast between Northern and Southern Hemisphere kelp forests. Similar approaches might be taken to discern the evolutionary importance of apex predators in other ecosystems.

What humans perceive as "dysfunctional ecosystems" are often consequences of the recently altered roles of predators (e.g., losses of native species or introductions of exotics). The authors are now beginning to see the ecological manifestations of predators and top-down forcing in places that a decade ago ecologists would not even think to have looked. These include effects on biodiversity, the success of invasive species, disease dynamics, fire ecology, carbon and nutrient sequestration, cross-ecosystem linkages, and even the composition and quality of water, soil, and the atmosphere.

Theoretical Studies of Predation and Biodiversity

The topic of predation and biodiversity involves interactions among multiple species at different trophic levels that interact in nonlinear, complex ways. A full understanding of this topic requires the analysis of mathematical and computer models, which permit one to keep track of multiple forces influencing the dynamics of interacting populations. This is particularly the case if one wishes to understand the factors governing the maintenance of species diversity. A noteworthy feature of the above case studies is that they often involve generalist predators which by feeding on a multiplicity of prey, link the dynamics of many species. So a particularly important area of inquiry is comparing specialist and generalist predators. A rich theoretical literature exists, exploring impacts of predation on the dynamics and structure of ecological communities. Here, the authors summarize highlights of this work, emphasizing conceptual insights rather than mathematical details.

A predator can influence whether or not a particular species is present in a community either by facilitating its persistence (i.e., predators can enrich species composition) or by preventing it from invading (i.e., predators can constrain species composition). If removing an apex predator greatly increases the abundance of a particular prey, and this prey is a predator on species at lower trophic levels, predator removal could indirectly lead to shifts in competitive interactions and thus persistence of species several trophic levels removed (as in a trophic cascade). Even if a predator does not dramatically affect composition, it may strongly influence relative abundances of resident community members. Finally, predators can influence the existence and magnitude of temporal fluctuations in abundance. Mathematical models help one understand all these effects.

To analyze mechanisms influencing species composition, we consider the growth of species when rare. If each species in a community increases when rare, diversity is maintained in the face of perturbations. In theoretical studies, one writes equations describing the dynamics of each species and then analyzes this set of equations (e.g., with and without a top predator). To illustrate the basic approach, the authors discuss a simple model in detail and then briefly discuss results from other models. Theoretical studies suggest that there is no single relationship between predation and biodiversity but instead many relationships, depending on numerous contingent details of systems.

Predation as a Density-Independent Mortality Factor: Effects on Biodiversity

As discussed previously, predator removal can unleash competition among prey and induce a wave of additional extinctions. To understand this effect, the authors express the growth rate of each species as a function of three factors: $dN/dt =$ (inherent growth) – (effect of resident competitor) – (mortality from resident predator), where N is abundance. Predation may both facilitate coexistence (e.g., by reducing the abundance and impact of competitors) and hamper coexistence (e.g., by direct mortality).

The most basic effect of predation is increased prey mortality. The simplest form of predation is that described by a fixed, density-independent mortality term. Generalist predators can occasionally act in this manner. Assume that two species experience strong, direct competition, described by the classic Lotka–Volterra competition model with added mortality due to predation:

$$\frac{dN_1}{dt} = N_1 r_1 \left(1 - \frac{N_1}{K_1} - \frac{\alpha_{12} N_2}{K_1}\right) - m_1 N_1 \qquad [1]$$

where N_1 is the density of species 1, N_2 is that of species 2, r_1 and K_1 are the intrinsic growth rate and carrying capacity of species 1, respectively, and the competition coefficient α_{12} is the effect of an individual of species 2 on species 1 (compared to the effect of 1 on itself). The second term expresses predation as density-independent mortality at a constant per capita rate m_1. (A comparable equation for species 2, where we reverse the subscripts 1 and 2, completes the model.)

Using the criteria that each species should increase when rare, after some algebra the following condition for species coexistence emerges:

$$\frac{1}{\alpha_{21}} > \frac{K_1(1 - m_1/r_1)}{K_2(1 - m_2/r_2)} > \alpha_{12} \qquad [2]$$

where the term $K_i(1 - m_i/r_i)$ is the effective carrying capacity of species i ($i = 1,2$) in the face of predation. This inequality implies several interesting conclusions. If $\alpha_{12}\alpha_{21} > 1$, no pattern of imposed, density-independent mortality leads to coexistence. This in effect says that interspecific competition is always stronger than intraspecific competition, and no fixed pattern of mortality will reverse this and permit coexistence. Likewise, if competition coefficients are unity, one will not observe coexistence, regardless of the pattern of mortality. If the two competitors have the same intrinsic growth rate, and predation is uniform (or more generally, $r_1/m_1 = r_2/m_2$), mortality drops out, so there is no effect of predation on coexistence. So clearly there will be some circumstances where predators will not facilitate coexistence of competing prey.

However, if $\alpha_{12}\alpha_{21}<1$, coexistence is possible in Lotka–Volterra competition. Predation can now facilitate such coexistence, if the two competitors differ in the ratio r/m. Consider a case of a competitive hierarchy, such that $\alpha_{12}=0$ but $\alpha_{21}>0$, and assume that species 2 is competitively excluded in the absence of predation. This happens if $K_2<K_1\alpha_{21}$. With the predator present, coexistence is permitted if $K_2(1-m_2/r_2)>K_1(1-m_1/r_1)\alpha_{21}$

Comparing this inequality to the previous one, after a bit of manipulation it can be seen that a necessary ingredient for predation to facilitate coexistence is that

$$r_1/m_1<r_2/m_2$$

In other words, the dominant competitor must either experience higher mortality from predation or have a lower intrinsic growth rate, than does the subordinate. However, if predation is too low (low m's for both species), there will still be competitive exclusion. Moreover, if predation is too intense, there will not be coexistence because one (or both) species will be directly eliminated by predation.

This model illustrates several general points applicable to a much broader range of models. First, predation does not always facilitate coexistence. Second, if a keystone predator effect is possible, the effect occurs only within a particular range of population parameter values, so will vary contingently with environmental conditions. Typically, predator-mediated coexistence requires an intermediate level of predation. Very intense, generalized predation will almost always reduce species richness. This is particularly likely in systems in which prey species have not had a shared evolutionary history with predators (e.g., the brown rat snake as a predator on birds on Guam), the physical structure of the environment makes it easy for predators to encounter prey (e.g., no refuges, as in open lakes), or the species in question have low intrinsic growth rates and so cannot readily replenish losses to predation. Third, predator-mediated coexistence requires a tradeoff: The species that is the superior competitor needs to be more vulnerable to predation. When this occurs, predator removal will endanger the persistence of inferior competitors (as in Paine's *Pisaster*). Another quite general point is that the appropriate measure of a species' vulnerability to predation combines mortality rates (m) and the ability to replenish losses (r). High, uniform rates of mortality tilt the balance of competitive interactions toward species with high intrinsic growth rates. Growth rates will usually depend on resource availability. Thus, bottom-up factors are involved in determining the outcome of a top-down process. Finally, along a gradient in predation pressure, provided there is a tradeoff between competitive ability and vulnerability to predation, the model predicts that species richness should be maximal at intermediate predation rates. We have seen this idea earlier, in the intermediate disturbance hypothesis of Connell (albeit without being dressed up in algebra).

This model structure assumes that predation influences competition entirely via changes in prey abundance. However, prey can also show behavioral shifts when faced with predators – for instance, spending more time in refuges and less on foraging. Such "higher order interactions" (often termed risk aversive or trait mediated effects) can either make coexistence more difficult or weaken competition, depending on the detailed nature of the behavioral changes.

The Lotka–Volterra model most literally applies to systems with strong, direct interference interactions, in which the only dynamical variables are each competing species density. In models of exploitative competition for a single limiting resource, predation that leads to density-independent mortality can influence which species wins, but it will not as readily lead to coexistence. Predation is more likely to promote (or occasionally to destroy) coexistence when mortality rates are dynamical variables responding to prey abundance, for instance because predator numbers or behavior shifts in accord with prey availability. The authors next explore several modifications of this model which illustrate the rich repertoire of dynamical behaviors made possible when predation varies dynamically in response to prey abundances.

Numerical and Functional Responses

The rate of mortality imposed by a predator reflects both predator numbers and the attack rate per predator. The total rate of mortality in the above model is $mN=fP$, where P is predator density and the parameter f is the number of prey (of the focal species) captured per individual predator. Because predators consume prey, the rates of the demographic parameters (birth, death, and movement) of predators will often vary as a function of the abundance of prey. Thus, the number of predators should depend on prey numbers. This is the numerical response. The rate at which an individual predator captures prey of a given species should also depend on the number of prey that are available, typically (although not always) increasing with prey abundance but saturating at high prey numbers. This is the functional response. It is useful to express the functional response as $f=aN$, where a is the attack rate per predator per prey – the risk of mortality an individual prey faces from an individual predator. Holling in 1957 introduced the distinction between predator functional and numerical responses as an essential step toward understanding the stability of direct predator–prey interactions, but these concepts also are significant in understanding the role of predation in maintaining prey species diversity. Different predators have distinct numerical and functional responses and therefore will have different impacts on species coexistence.

If predators strongly interfere with one another, this can weaken both functional and numerical responses to their prey. If we consider predators at intermediate trophic levels, higher order predators can also strongly dampen functional and numerical responses. For instance, active foraging by intermediate consumers may attract the attention of higher order predators. If so, adaptive responses by the intermediate consumers may keep in check their own attacks on prey, thus weakening both functional and numerical responses to increased prey abundance. Understanding the consequences of such flexible behaviors for food web dynamics is a very active area of current research. The case studies of strong top down effects sketched above can be viewed abstractly as emerging from the interlocking of functional and numerical responses by consumers across trophic levels.

Specialist Predators and Biodiversity

Specialist predators (whose diets are restricted to single prey species) typically reduce abundance of their favored species, freeing up resources for nontarget species. This can facilitate coexistence if dominant competitors attract more specialist predators or parasites than do subordinates. This diversifying effect of specialized predation can be mimicked in (1) by letting the attack rate on each species be a function of the abundance of a specialist predator, $m_iN_i=a_{ij}P_jN_i$, where P_j is the abundance of predator species j, which is specialized in its foraging just to prey species i. To complete the model, we need an equation for the dynamics of each predator: $dP_j/dt=P_j[g_j(N_i,P_j)]$, where g is a function which increases with N_i (e.g., because predators convert prey consumption into births) but may decrease with P_j (e.g., because predators interfere with each other). For any biologically reasonable system, at low numbers of their required prey, specialist predators must decrease (i.e., $g<0$). Hence, the growth rate of prey species i when it is rare and other prey are resident, will involve only the intrinsic growth of prey i, discounted by competition with the resident species. Because the residents continue to be attacked by their own specialist predators, their numbers will be depressed below carrying capacity, reducing competition imposed on the rare focal species. Thus, specialized predators quite generally are expected to facilitate prey coexistence. Moreover, a specialist predator that is effective at limiting a prey species that otherwise would be a dominant species, therefore has a strong indirect effect on the rest of its community and ecosystem.

An interesting twist is that specialist predator–prey interactions are often unstable when the predator is very effective at limiting prey numbers because of time lags in the numerical response of predators to prey, including crashes by predators that follow declines in their prey. This has two consequences for species coexistence. First, as shown by Peter Abrams and others, with saturating functional or numerical responses, instability tends to depress average predator numbers and thereby increase average prey numbers. This increases competition, making coexistence of different prey species more difficult. Second, with unstable dynamics between a resident specialist predator and its prey, there will be times when that prey is rare and an inferior competitor can invade and temporarily persist, only later to be competitively excluded when the predator is rare and the dominant competitor has rebounded in numbers. Thus, unstable specialist predator–prey dynamics induces instability in prey community composition as well.

These effects are believed to be particularly important when considering impacts of insect herbivores on plant communities, and pathogen dynamics can lead to unstable dynamics as well. For instance, forest insects and pathogens can outbreak and devastate large swaths of tree populations, and thus drive pulses in detritus and major changes in vegetation structure, impacting most species in those communities. However, most predators in the case studies discussed in this article have generalized diets and seem less prone to boom-and-bust dynamics (though data are scant on this point).

Generalist Predators and Biodiversity

Switching and ''Enemy-Free Space''

The expectation that specialist predation helps competing species to coexist depends on the very general and reasonable assumption that specialist predators will have strong numerical responses to their prey. Generalist predators can have a wide range of effects on prey communities. A single, effective generalist can act like a whole suite of specialists in promoting prey coexistence if the predator ignores whichever prey species is temporarily the rarest, concentrating attacks on common prey – the mode of foraging behavior called switching. As shown by Roughgarden and Feldman, switching predators can readily prevent competitive exclusion. In our formulation for mortality due to predation, for example, on prey species 1, $m_1N_1=a_1P_1N_1$, we can represent switching with an attack rate expressed as $a_1(N_1,N_2,\ldots)$, where a_1 declines toward zero as N_1 approaches zero but increases if the other N_i decrease. In effect, a prey species may persist because of a refuge in relative rarity, as defined by the predator's behavioral responses. Because the predator is also reducing the abundance of potential competitive dominants, this is a potent mechanism for predator-mediated coexistence. When such a predator is removed, numerous prey species may suddenly risk extinction if they compete, because such competition will become more intense as prey numbers rise after release from strong predation.

This idea is intuitively appealing, but there are surprisingly few unequivocal demonstrations of it in natural systems. Moreover, those that do, seem to involve prey species found in different habitats. A behavioral rule such that predators leave patches with few prey, and aggregate in patches with numerous prey, leads to switching and will tend to foster prey coexistence. If prey species are spatially segregated, they are not likely to be strongly competing in any case. Other potential cases of switching seem to involve prey species with very different strategies for blending into the background, different activity times, or different behavioral tactics for escaping predation. John Lawton and Michael Jefferies suggested that such species differences promoting coexistence be viewed as partitioning of enemy-free space, an aspect of niche differentiation.

In any case, several examples of species exclusion caused by predation discussed elsewhere in this article show that predators often do not tend to ignore rare prey species but rather continue attacking rare prey, even to the point of extinction.

Saturating Functional Responses

All predators have a maximal capacity for attacking prey that is set by limited time or gut capacity. Time or effort expended in attacking one prey will be unavailable for attacking other prey. If predator numbers are fixed, an increase in abundance of one prey species may reduce attacks on another. As with switching, we can represent this as $a_i=a_i(N_1,N_2,\ldots)$, but now attack rates decline with the abundance of each species. This inverse density dependence has two consequences. First, considering the effect of a prey species on itself, its mortality diminishes with increasing abundance. This inverse density dependence can lead to alternative stable states for a given

prey species – one at low and another at high densities. The low-density equilibrium may even be at zero density. More generally, at low densities a predator may have an accelerating functional response, but saturate at higher prey numbers (Holling's Type III functional response). It is believed that this effect is responsible in part for prey outbreaks, where prey may be regulated at low numbers by predators, but when perturbed they can sometimes escape to explode to much higher numbers. Second, in a community context, with a saturating response, an increase in any prey species can have a short-term effect of reducing attacks on other prey species; this in effect makes alternative prey indirect mutualists, at least over short time scales. Such mutualism is most likely when considering predators that are constrained and sluggish in their numerical responses to increasing numbers of prey (e.g., due to long generation lengths relative to those of their prey, and low in mobility, and at intermediate trophic levels).

Unlike switching, saturating functional responses are universal. The existence of indirect mutualisms arising because predator functional responses can be swamped may explain many natural phenomena, such as herding, mixed-species flocks of birds and schools of fish, and synchronized mass emergences and migrations. One consequence of importance to biodiversity is that if some prey species are reduced in abundance, there is an immediate increase in predation on other alternative prey, which thus risks extinction.

Numerical Responses and Apparent Competition

The potential for indirect mutualism among these taxa via the functional response is often offset by a kind of indirect competition via the numerical response. The case studies above show that apex predators can have dramatic effects on prey, leading to chain effects of indirect interactions across entire ecosystems. In most of these examples, the predators are generalists and have different magnitudes of impacts on different prey species. The sea otter for instance eats bivalves, snails, other invetebrates, and fish, in addition to urchins, and its ability to reduce urchins to very low levels depends on the availability of a range of alternative prey types which can sustain healthy otter populations. This implies the existence of another kind of indirect interaction through species on a given trophic level, but mediated through consumers, rather than resources. Just as consumer species can reduce each other's abundance by depleting a shared limiting resource, alternative prey species can indirectly depress each other's abundance by increasing the abundance of a shared predator. This indirect interaction is known as apparent competition. It is particularly likely if predator population growth rates increase strongly with the abundance of each prey type in the predator's diet and predator numbers are not strongly limited by other factors (e.g., territoriality or higher order predators). It is also somewhat more likely for predators with short generation lengths, not greatly exceeding those of their prey (e.g., parasitoids attacking herbivorous insects), or for predators which are highly mobile and can quickly aggregate into habitats with unusually high prey densities.

When any given prey species is rare, an increase in predator abundance will usually lead to an increase in its mortality. Predator abundance is expected to increase with the productivity and availability of alternative prey. The rate of predation on a focal prey is thus determined by the indirect, cumulative impact of alternative prey, sustaining the predator population at densities higher than permitted by the focal prey, considered alone. Particularly when prey species are not strongly competing, the negative indirect interaction of apparent competition can limit prey species diversity.

When a prey species is rare, its rate of population growth can be represented as $r - aP$. As noted previously, often the presence of alternative prey reduces the attack rate on a rare species. However, these same prey determine the magnitude of P. The net effect of prey on each other can only be determined by analyzing specific models. However, some general points are worth making. A given prey species is excluded by predation if $0 < r < aP$. Prey species with low r are particularly vulnerable to exclusion by shared predation, as are species which have high attack rates. Another way of stating the exclusion criterion is that the maximal predator density which this prey can tolerate is r/a. All else being equal, prey with low values for r/a are vulnerable to exclusion. A comparable observation was made below for the above model for predation as a density-independent mortality factor, but the difference in this case is that now the level of predation is itself a dependent variable in the system, and in particular depends on the availability and productivity of the entire suite of prey that sustain the predator. Note that a prey species with a high value for r/a can support a high abundance of a generalist predator, which can then with impunity overexploit alternative prey with lower r/a. The upshot of these observations is that there is a tendency toward the exclusion of prey species in the diet of polyphagous predators. The maintenance of high diversity requires mechanisms that offset this effect. Such mechanisms might include predator switching, constraints on the predator numerical response, or prey adaptations that reduce predation at low density (e.g., spatial refuges, with different prey species having different refuges).

Often, apparent competition will be strongly asymmetrical. A productive prey, which also can only be successfully attacked by predators when it is young, ill, or aged, may not be strongly limited by predation. However, this species can sustain a high population of the predator, which can then severely depress species more vulnerable through their life history. If different prey species occupy different habitats, and predators have limited mobility, this too can prevent exclusion via shared predation. However, mobile predators can be sustained by productive prey in one habitat and with impunity can overexploit prey in low-productivity habitats. A serious effect of habitat fragmentation is that it exposes species in habitat remnants to predation from generalist predators sustained by alternative prey in the surrounding landscape. Most examples of dramatic prey limitation by predation seem to depend on the availability of alternative prey, which permit predator numbers to remain high. The brown rat snake on Guam can persist on a diet of rats and lizards, which permits it to eat out of existence native bird species. Subsidies for predators (e.g., flows of prey from one habitat to another; a lighthouse keeper feeding her cat) in particular can permit a predator to exploit vulnerable prey to the point of local extinction.

Previously, the authors discussed the cascading effects of top predator removal. Such removals shift the factors regulating intermediate mesopredators or herbivores, which will

increase and become more regulated by food availability than they were in the past. This can unleash strong apparent competition effects at lower trophic levels, often on prey species which have evolved with relatively low levels of predation in their environment. Prey species harmed by polyphagous predators via "mesopredator release" are victims of apparent competition.

Generalist Predators and Community Stability

Generalist predators can have many different effects on the overall stability of communities. Here, the authors discuss a few interesting effects.

Richard Vance examined a Lotka–Volterra model akin to (1) but with a predator showing numerical responses to each prey species. Even if the pairwise interactions were all stable, the entire ensemble could show large-amplitude cycles, or even chaotic dynamics.

If generalist predators are mobile and seek out patches with high prey abundance, this can lead to switching. If these responses are rapid, then generalist predators help stabilize prey dynamics. However, there can be time lags in these responses, which in turn can be destabilizing. Recent theoretical explorations suggest that switching behavior in patchy environments leads to systems which persist but which have bounded oscillations. Moreover, in changing landscapes, mobile generalist predators can concentrate in habitat remnants, leading to transient spikes of high predation and extinction risk for prey species residing in these remnants.

Generalist predators usually have a mixture of a few strong interactions and many weak interactions with the species of prey in their diet. Theoretical studies have recently demonstrated that weak interactions can help reduce inherent instabilities in strong predator–prey interactions. However, the effect depends on detailed assumptions made about predator–prey feedbacks; for instance, if prey flow into a given habitat and contribute to the diet of a resident predator, this predation does not feed back to influence prey numbers in the source habitat. If the predator instead is highly mobile, its ability to feed in multiple habitats may permit severe overexploitation of prey in some habitats. Despite such cautionary tales, many ecologists are convinced that flexible foraging by generalist predators is broadly an important factor stabilizing complex food web interactions.

Given saturating functional responses, apparent competition can be reduced in unstable systems, relative to stable systems, because the nonlinearity in the functional response means that predators are harmed more by times of low prey abundance, than benefited by times of high prey abundance. This reduces the number of predators which can be sustained, thereby weakening apparent competition effects. The dominant effect may then be competitive exclusion (if the prey are strong competitors) or indirect mutualism (with noncompeting prey and saturating functional responses).

If a prey species is dominant in apparent competition, this may permit it to persist with a superior competitor, in effect being maintained by generalist predation. As noted above, selective predation, particularly by specialist predators, can help offset competitive exclusion. Peter Chesson and Jennifer Kuang have recently developed, in some detail, the insight that predation and competition are likely of comparable importance in explaining coexistence, and that a multitrophic perspective needs to replace the past focus on competitive interactions and niche partitioning to have a full understanding of the factors that maintain the rich diversity of natural communities.

The theoretical studies summarized here suggest that there is no universal effect on predation on biodiversity but rather many effects. More important, they help emphasize potential surprising effects of predator removal and highlight the range of information that a conservationist needs to gauge the likely impacts of management alternatives. Most theoretical studies of how predation impacts species diversity have been limited to interactions across two-trophic levels. More complex theoretical studies of trophic interactions are needed to investigate such phenomena as trophic cascades and their indirect effects on other species and ecosystem processes in more complex, multilevel food webs.

Generality

There is much evidence for the influences of predators on species, populations, communities, and ecosystems. However, how predictable and widely occurring are these effects? To what degree do predators regulate the structure and function of the planet's ecosystems relative to other biological interactions and physical forces?

These important questions can be asked at two levels – within and between ecosystems. Within systems, the question of generality usually concerns the geographical extent of an effect that has been demonstrated at one or several sites. Michael Foster raised this question for the influences of sea stars in mussel beds and sea otters in kelp forests, claiming that the interactions were less common than generally believed. Not surprisingly, sea star predation is important at many sites but not everywhere. For sea otters and kelp forests, Estes and Duggins approached the question by evaluating how consistently predictions of the otter–kelp forest trophic cascade played out at many sites with and without sea otters. In this case, the predictions held up (i.e., otter-dominated sites supported kelp forests and otter-free sites were deforested by sea urchin grazing) nearly everywhere they examined, from southeast Alaska to the western Aleutian archipelago (**Figure 8**), and similar results have been obtained from British Columbia.

Across ecosystems, the question of generality for top-down influences of apex predators becomes one of both relative importance and variation in process. This question is exemplified by the indirect effects of wolves on aspen recruitment in the greater Yellowstone ecosystem. Herbivory by elk has been established as a key limiting factor to aspen recruitment in many areas across the intermountain west. The reintroduction of wolves to Yellowstone has resulted in a striking flush of tree recruitment in the Lamar Valley region, thus suggesting that the repatriation of wolves would reverse the long-standing decline of aspen over much larger areas. However, Matthew Kauffman and colleagues have shown that aspen recruitment has not increased over most of Yellowstone's northern range, thus calling into question the generality of the wolf-elk-aspen trophic cascade. Whether the cascade is truly of only limited

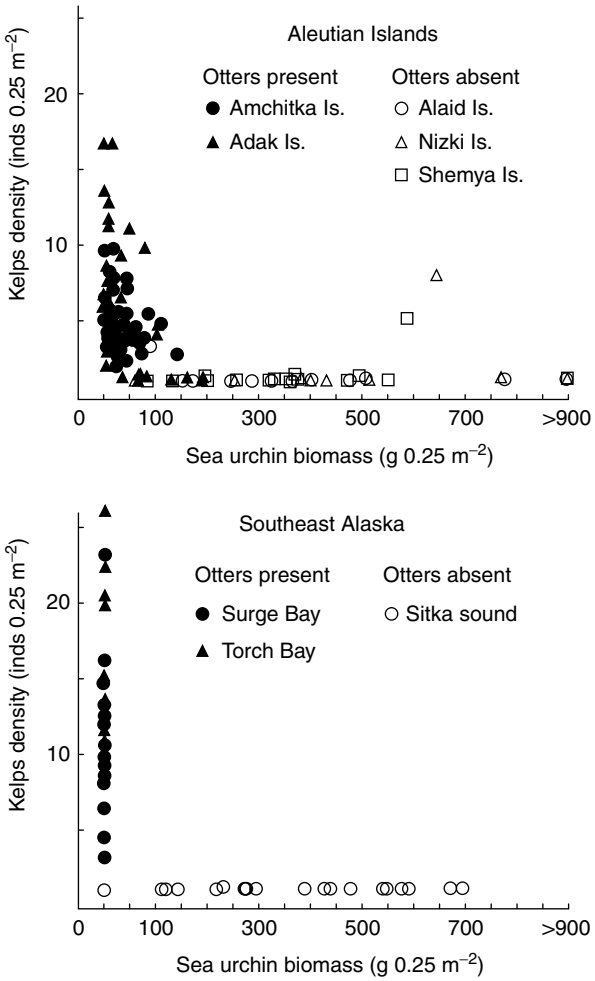


Figure 8 Epibenthic kelp density plotted against estimated sea urchin biomass from locations with and without sea otters in the Aleutian Islands and southeast Alaska. Points represent averages of 20 randomly selected plots from each location. These data show that kelp forest phase states vary predictably depending on the presence or absence of sea otters. Reproduced from Estes JA and Duggins DO (1995) Sea otters and kelp forests in Alaska: Generality and variation in a community ecology paradigm. *Ecological Monograph* 65: 75–100, with permission from ESA.

occurrence or there has simply been insufficient time since the wolf introduction for the system to respond broadly remains uncertain. We know that trophic cascades occur broadly in nature. Nonetheless, understanding the breadth of influences by predators remains one of ecology's most daunting challenges. The theoretical perspectives sketched above show that many effects in principle can happen, but being able to predict which occurs where is beyond what ecologists can presently do. But if we hope to devise strategies for the wise long-term management and preservation of natural systems, this challenge must be tackled.

The Future

Although predators have long concerned conservationists and resource managers, the future will – or should – bring heightened attention to this group. As a general rule, apex predators are more vulnerable to local extinctions than are lower trophic-level species. The rapid demise of predators has occurred in part because they have been (and in many areas still are) treated as competitors of humans for fish, wildlife, and agricultural resources; thus, they have been persecuted rather than conserved. Habitat fragmentation has also hastened local extinction rates of large apex predators because their typically low densities and large home ranges render the smaller fragments incapable of maintaining viable populations. Large-bodied predators typically have long generation lengths and low reproductive rates, so populations cannot readily withstand mortality inflicted on them by, for example, hunting and toxins. These facts, together with the growing realization that predators are commonly essential for maintaining native ecosystems, are leading to a paradigm shift in conservation biology, especially in the area of reserve design. Earlier approaches to conservation planning focused on preserving representative habitats. Many ecologists now believe that this approach, although necessary, is not sufficient. Large reserves, or a series of smaller connected reserves, are necessary to maintain viable populations of key species of predators, and in particular apex predators (typically vertebrates), which in turn are essential for maintaining the functional integrity of ecosystems. This view has important implications to restoration ecology, which until recently has focused mainly on the reintroduction of native plants and the elimination of exotic species, with scant attention given to building viable entire food webs.

If the maintenance or restoration of native predators is important to conservation biology, so is the elimination of exotic predators. Exotic predators have devastated many natural biotas, both because of their ability to reduce or exterminate native prey species and because of the many indirect effects of these prey within the ecosystem's interaction web. As discussed previously, island biotas provide the most obvious and poignant examples of such effects. The removal of exotic species is usually expensive, time-consuming, and fraught with technical challenges. Nonetheless, exotic predators, because of their great mobility and low population density, are often easier to remove than other invasive species.

The conservation and management of predators require more and better information than is currently available. As noted by George Schaller, field studies have been conducted on less than 15% of the species of mammalian carnivores. And even of these, there are few predators for which enough information is available that one could write down credible expressions for functional and numerical responses of the sort that are essential for management and predictive models. Although this is indeed a feeble record, the authors suggest that there are even fewer successful efforts to understand the ecological importance of this group.

Are all or most of the predators important players in prey population regulation and ecosystem dynamics? The fact that so many species already have been depleted or eliminated hampers the pursuit of answers. The depleted status of so many predators raises the additional question of how their influences vary with population density. Most apex predator populations probably were once regulated by competition for food, which in turn must have strengthened predator–prey

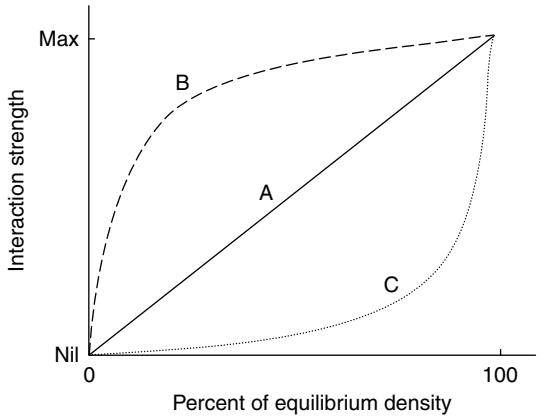


Figure 9 Several possible functional relationships between predator-induced interaction strengths and predator population abundance relative to equilibrium density. In A (solid line), the relationship is linear. In B (dashed line), the main influences of predators occur at a wide range of population densities, whereas in C (dotted line) these influences occur only at high population densities.

interactions imposed on lower trophic levels. However, humans have changed this, so we must now wonder how the strength of these interactions varies with the degree to which predators have been pressed below historical average levels. The several possible functional relationships, shown in Figure 9, have very different management implications. Some people might hope for a relationship like that depicted by B in Figure 9, because this would mean that the maintenance or restoration of the ecological roles of predators requires only that they be present in low numbers. In contrast, other people with different values or goals might hope for a relation more like that depicted by C, in which the influence of predators is of little consequence at population levels below carrying capacity.

A final need is for greater understanding of the full range of food web effects by predators. Most studies of predator effects have focused on pairwise predator–prey interactions, thus creating the overly simplistic mind-set still held by many wildlife and fisheries managers. Most studies of indirect effects take this perspective only a step further by focusing on trophic cascades. However, as discussed previously, (*see* Trophic Cascades *and* Generalized Food Web Theory) trophic cascades are expected to mediate competitive interactions and influence the strength of bottom-up forces by altering production levels at the base of the food web. And as shown in the earlier examples, interactions of this nature can spread to seemingly unlikely species or areas of ecosystem structure and function. Little is known about such interactions, although several examples of long and complex chain reactions among species show that they can be quite important.

For the most part, food web and ecosystem dynamics have been studied by one of two approaches. The oldest of these involves descriptions of food web structure and estimating the transfer of materials and energy through their various linkages, the critical assumption being that flux rate reflects interaction strength. The fallacy of this assumption is evident in the fact that strong top-down forces necessarily reduce prey abundance to such low levels that the interaction may no longer be apparent in a static food web. The second approach is to observe these dynamics directly by perturbing the system. Unfortunately, most experimental studies have been limited to processes acting on small spatial and temporal scales (for obvious logistical reasons) and to invertebrates and heterothermic fishes (for social and political reasons). Our general lack of understanding of the larger apex predators (especially large mammals) is due in large part to the almost complete absence of experimental evidence.

Creative approaches are needed to put the ecological influences of large apex predators into a dynamical context. The use of historical information and opportunities provided by "natural experiments" holds promise in this regard. The historical record is a rich source of information that has barely been utilized. Assessments of the wolf/moose/boreal forest system and the wolf/elk/temperate forest from tree ring analysis, and the sea otter–kelp forest system from faunal remains in Aleut kitchen middens illustrate the utility of this approach. Fortuitous change is another potential means of assessing the ecological roles of large apex predators. Studies by John Terborgh on the recently formed islands of Lago Guri, by Kevin Crooks and Michael Soulé on mesopredator release in urban habitat fragments, and by James Estes and John Palmisano on sea otters and kelp forests demonstrate the value of this approach. As these and other examples show, ecologists and resource managers must remain mindful of appropriate scales of time and space for both learning about predators and applying that knowledge to the conservation of biodiversity.

Appendix

List of Courses

1. General Ecology
2. Population Biology
3. Conservation Biology
4. Wildlife Ecology
5. Fisheries Biology

See also: Carnivores. Competition, Interspecific. Food Webs. Keystone Species. Parasitism. Population Dynamics. Species Interactions. Trophic Levels

References

Byers JA (1997) *American PronghornSocial Adaptations and the Ghosts of Predators Past*. Chicago: University of Chicago Press.

Carpenter SR, Kitchell JF, and Hodgson JR (1985) Cascading trophic interactions and lake productivity. *Bioscience* 35: 634–639.

Castilla JC and Duran LR (1985) Human exclusion from the rocky intertidal zone of central Chile: The effects on *Concholepas concholepas* (gastropoda). *Oikos* 45: 391–399.

Crooks K and Soulé ME (1999) Mesopredator release and avifaunal extinctions in a fragmented system. *Nature* 400: 563–566.

Fretwell SD (1987) Food chain dynamic: The central theory of ecology? *Oikos* 50: 291–301.

Hairston NG, Smith FE, and Slobodkin LB (1960) Community structure, population control and competition. *American Naturalist* 94: 421–425.
Holt RD (1977) Predation, apparent competition, and the structure of prey communities. *Theoretical Population Biology* 12: 197–229.
Myers RA, Baum JK, Shepherd TD, Powers SP, and Peterson CH (2007) Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* 315: 1846–1850.
Paine RT (1966) Food web complexity and species diversity. *American Naturalist* 100: 65–75.
Power ME (1990) Effect of fish in river food webs. *Science* 250: 411–415.
Prugh LR, Stoner CJ, Epps CW, *et al*. (2009) The rise of the mesopredator. *BioScience* 59: 779–791.
Reznick D and Endler JA (1982) The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). *Evolution* 36: 160–177.
Terborgh J and Estes JA (eds.) (2010) *Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature*, 464 p. Washington, DC: Island Press.
Vermeij GJ (1977) The Mesozoic marine revolution: Evidence from snails, predators and grazers. *Paleobiology* 3: 245–258.

Primate Populations, Conservation of

Russell A Mittermeier and Matthew C Richardson, Conservation International, VA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Russell. A. Mittermeier, William. R. Konstant, volume 4, pp 879–889, © 2001, Elsevier Inc.

Glossary

Bushmeat Wildlife such as primates not traditionally considered edible that is hunted and used for food, usually illegally.

Diversity The variety of species within a given taxonomic group.

Endemic A species or taxonomic group unique to or occurring only in a specified geographic location or region.

Flagship A popular, charismatic species that serves as a symbol and rallying point to stimulate conservation awareness and action.

Haplorrhini One of the two recognized primate suborders; includes the tarsiers, New World monkeys, Old World monkeys, and apes.

Megadiversity countries Countries which harbor a disproportionately large share of the world's fauna. In terms of primates, the four megadiversity countries are Brazil, Madagascar, the Democratic Republic of Congo, and Indonesia.

Prosimians A rather scientifically imprecise grouping that includes such "Lower" primates such as galagos, lorises, lemurs, and tarsiers.

Strepsirrhini The other of two recognized primate suborders; includes the galagos, lorises, and lemurs.

Zoogeographical regions One of six geographical divisions of the world devised for the study of the distribution of land mammals. They consist of the Palearctic, the Nearctic (North America), the Neotropical (South America), the Afrotropical (Africa south of the Sahara), the Oriental (Asia south of the Himalayas), and the Australasian (Australia, New Zealand and nearby islands), the latter two featuring a zone of overlap known as Wallacia. Antarctica forms an additional region.

Introduction

During the past few decades, interest in nonhuman primates has increased significantly on many fronts and has helped to build support for conservation. Pioneering and long-term field studies of the great apes (Schaller, 1963; Goodall, 1968; Fossey, 1983; Galdikas, 1995) have dispelled age-old myths about mankind's closest living relatives, greatly narrowing the gaps between these species and our own, and provided new insights into human origins and behavior. The continuing search for new drugs to treat global maladies such as malaria, cancer, and AIDS has required large numbers of nonhuman primates as experimental subjects (Mack and Mittermeier, 1984). For some species, this use has contributed to serious declines in wild populations and ultimately forced issues of conservation and captive breeding as part of the long-term strategy for biomedical research. In other cases, little-known and formerly obscure primate species of the Neotropics, Africa, and Asia have emerged as prominent "flagships" for conserving their tropical forest habitats, which biologists agree are the richest natural terrestrial ecosystems on the planet.

Although interest in nonhuman primates is increasing, the threats to their survival persist. Varying combinations of habitat destruction, hunting, and live capture have driven dozens of primate species and subspecies to the brink of extinction, to the point where several taxa now number only in the low thousands or hundreds, and for some species less than a hundred now remain. Such populations are doomed without long-term protection, monitoring, and a heightened understanding of their plight by local human populations.

In the face of continuing threats, however, primate conservationists can also reflect on the past century and realize with some degree of pride that, to the best of anyone's knowledge, not a single primate taxon went extinct during that period. Certainly it is true that, during the first decade of this new century, at least two primates have been feared lost: Miss Waldron's red colobus (*Piliocolobus waldronae*) and the Yunnan white-handed gibbon (*Hylobates lar yunnanensis*). However, in the case of the former, at least there remains some reasonable hope that it may yet survive, while the latter is still officially considered Data Deficient by the International Union for Conservation of Nature and Natural Resources (IUCN). The near future could conceivably witness the loss of many others, as for instance the Hainan gibbon (*Nomascus hainanus*) of China, the Tonkin snub-nosed monkey (*Rhinopithecus avunculus*) of Vietnam or the Sahafary sportive lemur (*Lepilemur septentrionalis*) of Madagascar, all Critically Endangered. The survival of these and other threatened taxa will depend on the continuation of existing conservation programs, and the establishment of new ones according to a global strategy for the most threatened. Fortunately, support for global primate conservation increased in the latter half of the 1990s after having suffered something of a dry spell, and the expertise is at hand to direct available resources to the highest priority species, habitats, and projects.

Overview of Global Primate Diversity and Conservation

The order Primates is one of 29 mammalian orders that together total more than 5500 species. It includes 16 families,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00370-1

78 genera, and at last count, 438 species and 673 taxa worldwide. Two suborders of primates are recognized, the Strepsirrhini (galagos, lorises, and lemurs) with seven families, and the Haplorrhini (tarsiers, New World monkeys, Old World Monkeys, and apes) with nine (**Table 1**).

Living strepsirrhines occur only in the Old World, despite the fact that North America was once a major center of their early evolutionary history. Of the seven extant strepsirrhine families, five (Cheirogaleidae, Lepilemuridae, Lemuridae, Indriidae, and Daubentoniidae) occur naturally only in Madagascar, where they are represented by 15 genera, 97 species, and 101 taxa. The five genera, 19 species, and 30 taxa of the Galagidae are confined to sub-Saharan Africa, while the four genera, 12 species, and 17 taxa of the Lorisidae are found in mainland Africa, India, and Southeast Asia.

The haplorhine primates are much more diverse than the strepsirhines, having three times the number of genera and three-and-a-half times the number of species and total taxa. Of the nine haplorhine families, five (Callitrichidae, Cebidae, Aotidae, Pitheciidae, and Atelidae) occur only in the New World tropics, two (Cercopithecidae and Hominidae) occur throughout much of Africa and Asia, and two (Hylobatidae and Tarsiidae) are restricted to Asia. The family Tarsiidae is currently represented in Indonesia and the Philippines by three genera, 11 species, and 16 taxa; however, a major taxonomic revision of the group is now underway, which is expected to significantly increase these numbers. New World nonhuman primates comprise 20 genera, 148 species, and 204 taxa. By comparison, the Old World monkeys (subfamilies Colobinae and Cercopithecinae), although much more widely distributed, are only somewhat more diverse, comprising 23 genera, 125 species, and 264 taxa overall. The gibbons (Hylobatidae) are a relatively small group that includes four genera, 17 species, and 27 taxa, while the great apes (Hominidae) are even less diverse with four genera, seven species, and only 14 taxa overall. It should be noted that the human primate is considered a morphologically variable but monotypic member of the latter family. For the remainder of this article, the authors use "primate" only in reference to nonhuman taxa unless otherwise specified.

Examining global primate diversity from a regional perspective, it can be seen that it is by no means evenly distributed. For example, the remaining tropical forests of Madagascar are dwarfed by the extensive tropical forests of Africa, Asia, and Central and South America, but Madagascar has by far the densest concentration of primate diversity anywhere on Earth within an area of 587,045 km^2. Although usually considered part of the Afrotropic zoogeographical region, Madagascar is in many respects a zoogeographic region by itself, and especially so with respect to its unique primate fauna. All of Madagascar's primates are endemic, with the exception of the mongoose lemur (*Eulemur mongoz*) and the brown lemur (*Eulemur fulvus*), which also occur on the nearby Comoros Islands but were almost certainly introduced there by humans.

With the exception of the Barbary macaque (*Macaca sylvanus*), all primates of sub-Saharan Africa and nearby islands (e.g., Zanzibar and Bioko) occur within the Afrotropic zoogeographical region. They include 25 genera, 101 species, and

Table 1 Primate diversity by family

Family	*Common names*	*Distribution*	*Genera*	*Species*	*Taxa*
Suborder Strepsirrhini					
Galagidae	Dwarf galagos, lesser galagos, squirrel galagos, needle-clawed galagos, and thick-tailed galagos	Africa	5	19	30
Lorisidae	Slender lorises, slow lorises, angwantibos, and pottos	Asia, Africa	4	12	17
Cheirogaleidae	Mouse lemurs, giant mouse lemurs, hairy-eared mouse lemur, dwarf lemurs, and fork-marked lemurs	Madagascar	5	30	30
Lepilemuridae	Sportive lemurs	Madagascar	1	26	26
Lemuridae	Bamboo lemurs, ring-tailed lemur, true lemurs, and ruffed lemurs	Madagascar	5	23	25
Indriidae	Woolly lemurs, sifakas, simponas, and indri	Madagascar	3	19	19
Daubentoniidae	Aye-aye	Madagascar	1	1	1
Total Strepsirrhini			24	130	148
Suborder Haplorrhini					
Tarsiidae	Tarsiers	Asia	3	11	16
Callitrichidae	Marmosets, Goeldi's monkey, tamarins, and lion tamarins	Neotropics	7	43	62
Cebidae	Squirrel monkeys and capuchin monkeys	Neotropics	3	22	36
Aotidae	Night monkeys	Neotropics	1	11	13
Pitheciidae	Titi monkeys, sakis, bearded sakis, and uakaris	Neotropics	4	44	51
Atelidae	Howler, woolly monkeys, spider monkeys, and muriquis	Neotropics	5	28	42
Cercopithecidae	Guenons, macaques, crested mangabeys, baboons, drills, white-eyelid mangabeys, kipunji, colobus monkeys, langurs, surilis, and odd-nosed monkeys	Africa, Asia	23	125	264
Hylobatidae	Gibbons and siamang	Asia	4	17	27
Hominidae	Orangutans, gorillas, chimpanzees, and humans	Africa, Asia	4	7	14
Total Haplorrhini			54	308	525
Total			78	438	673

177 taxa overall. If we consider Madagascar in combination with the African continent, the entire region harbors a significant 45% of global primate diversity at the species level, and an even more impressive 52% of higher order (beta) diversity at the generic level.

In the Neotropics, primates occur from southern Mexico through Central America and northern South America and as far as southern Brazil, northern Argentina, and Paraguay (but not Chile). Native primate populations also occur in Trinidad; there are several introduced populations of African primates on the islands of St Kitts, Nevis, Barbados, and Grenada; and both New World and Old World species have been introduced to the island of Puerto Rico. Twenty genera and 204 taxa are found in the Neotropics – amounts roughly comparable to those for Africa and Asia. However, the 148 primate species of the Neotropics are the most for any single major region, and account for 34% of global primate diversity.

Asian primates are found mainly in the Oriental zoogeographical region and in the southeastern portion of the Palearctic, as well as in the Wallacia transition zone between the Oriental and Australian regions. In southern Asia, primates are widely distributed from the Indian subcontinent and the island of Sri Lanka, throughout Southeast Asia as far as the Philippines and the Indonesian islands of Halmahera and Sulawesi, to Central and North Asia from Afghanistan through southern China (including the islands of Hainan and Taiwan) to Japan. In contrast to Africa (excluding Madagascar) and the Neotropics, where primates are basically continental, Asian primates are found in large numbers on islands as well. The region is home to 18 genera, 91 species, and 190 taxa, a somewhat lower level of generic diversity by comparison to the Neotropics and Africa but a level of species diversity roughly equal to that of Africa.

Ninety-one of the world's 193 sovereign nations have wild primate populations; the 10 countries with the highest primate diversity are listed in **Table 2**. Brazil, with 114 species, is the leader and accounts for 77% of all Neotropical primate species. Together, the top four countries (Brazil, Madagascar, Indonesia, and Democratic Republic of the Congo), which also represent the world's four major primate habitat regions, account for 301 species and 387 taxa, or nearly 60% of all living primates.

Furthermore, the top four countries for primate diversity (Brazil, Madagascar, Indonesia, and Democratic Republic of Congo) are also at the top of the list of the world's top countries for primate endemism (**Table 3**). Madagascar is by far the international superstar with 97 unique species, 15 unique genera, and five unique families, all representing 100% levels of endemism.

Endemic primates in Brazil and Indonesia account for approximately 58% and 72% of the species and subspecies that occur there. Brazil has fewer endemic species (71) than does Madagascar, despite the fact that it covers an area almost 15 times as large, and only four (*Callibella*, *Callithrix*, *Leontopithecus*, and *Brachyteles*) of its 19 primate genera (21%) are endemic. Indonesia is a distant third on the list with 28 endemic species and one endemic genus (*Simias*), after which the number of endemic species decreases precipitously. Only three other countries in the world – Tanzania (*Rungwecebus*), Peru (*Oreonax*), and Ethiopia (*Theropithecus*) – can claim an endemic primate species.

The key point is that there are a handful of countries, within the broad geographic regions that provide critical habitat for wild populations of nonhuman primates, which harbor a disproportionately large share of the world's primate fauna. The four megadiversity countries – Madagascar, Brazil, Indonesia, and Democratic Republic of the Congo – must therefore rank as the highest global priorities for conserving primates. Furthermore, if we extend the analysis to consider subspecies, we find that other countries will rise to the top of the priority list. Consider India for example. With 22 primate species and six endemics, it is well behind many other countries on the previous lists. However, of India's 31 primate taxa, 11 are endemic, and the additional six endemic subspecies represent important wild populations that should be considered when establishing conservation priority rankings.

Threats to Primates

Man-made threats to the survival of nonhuman primates can be divided into three major categories: habitat destruction (elimination, degradation, and fragmentation), hunting for food and a variety of other purposes, and capture for export or local trade (living, dead, and body parts) (Mittermeier *et al.*, 1986; Nijman, 2005). The effects of these vary significantly from species to species and from region to region, and are influenced by the extent and distribution pattern of remaining habitat, human population densities, the presence of roads and large rivers, the nature, and degree of human activities in

Table 2 World's top 10 countries for primate diversity

Country	*Genera*	*Species*	*Taxa*
Brazil	19	114	135
Madagascar	15	97	101
Indonesia	10	47	77
Democratic Republic of Congo	17	45	70
Peru	14	37	51
Colombia	13	40	51
Tanzania	13	26	46
Cameroon	20	34	44
Kenya	11	20	39
Uganda	15	24	37

Table 3 World's top 10 countries for primate endemism

Country	*Genus*	*Species*	*Taxa*
Madagascar	15	97	101
Brazil	4	71	78
Indonesia	1	29	56
Colombia	0	8	17
Democratic Republic of Congo	0	9	16
China (including Taiwan)	0	5	13
Tanzania	1	6	11
India	0	6	11
Sri Lanka	0	3	11
Kenya	0	3	9

the range of a particular species, local hunting traditions, the size and desirability of different species as food items or as sources of other products valued by humans, the demand for a given species in research or the pet trade, enforcement of existing wildlife laws, and regulation of commercial animal dealers. However, one or more of the three major threats affect almost all primates. It should be noted that, in addition to man-made threats, there are also certain other natural ones that can seriously impact populations. These include stochastic events such as cyclones, which have the potential to destroy large areas of habitat, and pandemics such as Ebola, which in recent years have decimated great ape numbers in Central Africa.

Habitat Destruction

On a global scale, habitat destruction is the principal factor contributing to the disappearance of wild primate populations. The continuing growth of the human population and its ever-expanding need for natural resources have contributed greatly to the destruction or alteration of natural habitats on an almost unimaginable magnitude, and nowhere has this problem been more acute than in the tropical regions of the world. The overwhelming majority of all nonhuman primates inhabit the tropical forests of Africa, Asia, and South and Central America. Although there are signs that the rate of destruction has been slowing somewhat since the 1990s, these forests are still being cut at an alarming rate. Just as troubling is the situation in Madagascar, which had become something of a conservation success story after years of rampant deforestation but, in the wake of recent political discord, is now suffering from a resurgence of burning, logging, and hunting, and where certain national parks are now being looted for their valuable hardwoods.

The immediate effects of habitat destruction on nonhuman primates vary significantly from one region to the next. For example, in Madagascar and the Atlantic forest region of eastern Brazil, so little suitable forest habitat remains that any further loss constitutes a grave threat to primates and other wildlife. In contrast, in the vast forest regions of Amazonia and the Congo Basin, which, along with the island of New Guinea, represent the three remaining major tropical wilderness areas of the planet, the effects of habitat destruction are only starting to be felt.

Hunting for Food and Other Purposes

The hunting of primates by human populations occurs for a variety of reasons, but by far the most important is to acquire food. Although primate hunting is prohibited by law in most countries, enforcement of such protective legislation is typically rare and sometimes nonexistent in the remote areas where this activity almost always occurs.

Hunting of primates as a source of food is a significant threat in at least three parts of the world: West Africa, Central Africa, and the Amazon region of South America. In each region, primates are among the animals most frequently hunted and they are regularly sold in markets, except where this is prohibited by law. However, even in areas in which primate hunting is common, it by no means affects all species equally. In Amazonia, for example, the larger monkeys such as *Cebus*, *Lagothrix*, *Alouatta*, and *Ateles* are heavily hunted because they are among the most desirable food species, whereas smaller monkeys such as *Saguinus* and *Saimiri* are rarely shot for food because they barely provide enough meat to recompense the hunter for the cost of his or her shotgun shell. As a result of better law enforcement in recent decades in countries such as Brazil, "bushmeat" hunting is no longer considered a serious threat to wild primate populations. The situation in Central and West Africa, in contrast, has quickly reached a critical stage because bushmeat hunters in these regions are being paid by logging companies to shoot large quantities of primates and other forest wildlife to feed work crews.

In areas where the hunting of primates for food is common, it can sometimes comprise a threat even more severe than forest destruction. In some parts of Amazonia there are large tracts of primary forest remaining in which populations of *Cebus*, *Alouatta*, *Lagothrix*, and *Ateles* have effectively been exterminated by excessive hunting (Mittermeier and Coimbra-Filho, 1977; Peres, 2000). This is sometimes referred to as the "empty forest syndrome" (Redford, 1992), and is particularly evident in Central and West Africa, where widespread hunting for the bushmeat trade (as opposed to subsistence) is devastating monkey and ape populations (Ammann, 2000). In areas where food hunting and deforestation are both prevalent, populations of forest primates and other game species can disappear quickly.

In some parts of the world, religious restrictions or other cultural factors prohibit (or inhibit) the killing and eating of primates. In many parts of India, for example, primates are rarely hunted for food because they are linked to the monkey god Hanuman, of the Hindu religion, whereas in strictly Muslim countries primates are not eaten because their flesh is considered unclean and unfit for human consumption. Indeed, in India, Hindus refuse to kill rhesus macaques (*Macaca mulatta*) and resist translocating them even when populations become so high that they constitute a menace to humans. In other countries, such as Madagascar, local taboos may exist against eating certain primates (e.g., *Indri*), whereas other species (e.g., *Eulemur* and *Varecia*) may be the most popular food items for a given tribe or village. Sadly, in recent years such taboos have begun to break down, with the result that hunting has begun to seriously impact many species that were formerly protected in this way.

Primates are also hunted to supply many other products in addition to food: traditional medicines, bait, body parts for ornamentation, and trophies. Primate hunting to provide medicinal products may be nothing more than a by-product of food hunting in most cases, and it usually involves the use of specific body parts for their supposed medicinal value. Illegal and unsustainable trade in wildlife, in particular primates, is a major conservation challenge in Asia. Trade in primates for body parts and tissues for traditional medicine, folklore, and food is particularly prevalent, cruel and enormously detrimental to populations of numerous Asian primates (Nekaris *et al.*, 2010). Trade in lorises (*Loris* and especially *Nycticebus*), for instance, is resulting in fast-declining populations and local extinctions (Nekaris and Nijman, 2007; Nekaris *et al.*, 2010). In south India, the meat of the Nilgiri

langur (*Trachypithecus johnii*) and the lion-tailed macaque (*Macaca silenus*) is regarded as an aphrodisiac and thought by local people to contain other therapeutic properties. The blood of Asian colobine monkeys, such as Phayre's langur (*Trachypithecus phayrei*) in Thailand, is widely believed to impart vigor to the drinker, especially when mixed with local whisky. So too, in various South American countries, drinking from the cup-shaped hyoid of an adult male howler monkey (*Alouatta*) is thought to cure goiters and stuttering as well as ease a mother's labor pains during childbirth. Although the hunting of primates for supposed medicinal purposes is considered a relatively minor factor overall in the global decline of wild primate populations, when it involves threatened species, such as in the case of India's lion-tailed macaque or some of the Southeast Asian colobines (*Trachypithecus*, *Presbytis*, and *Pygathrix*), it can be a serious problem.

Primates are also shot to provide bait for capturing and killing other animals, mainly in remote areas of the Amazon region. There, spotted cat hunters preferentially shoot larger monkeys such as *Lagothrix* and *Ateles* to bait crude wooden traps set for jaguars and ocelots. Any number of Amazonian primates may also be shot for fish or turtle bait, and in Sri Lanka monkeys often serve as bait for crocodiles. Although the use of primates for bait is a relatively minor threat, it can and does add to the pressures faced by overexploited large taxa such as *Lagothrix* and *Ateles*.

In some countries, primates may be killed for their skins or to provide other body parts used in ornamentation. Perhaps the most striking case of this is in Africa, where the skins of guereza (*Colobus guereza*) and related species have been used to fashion cloaks and headdresses for native African peoples, but have also figured significantly in the international fur trade. For example, in 1899 a reported 223,599 monkey skins were auctioned in London, and at least 2.5 million were probably exported to Europe between 1880 and 1900, especially to Germany, where they were used to make capes, muffs, and rugs. As recently as the early 1970s, colobus monkey rugs were common in East African tourist shops, and colobus coats were being sold in Europe and Japan.

Throughout much of Amazonia, tourist shops still offer stuffed monkeys, monkey skulls, monkey-skin hats, monkey-tail dusters, and necklaces fashioned from monkey teeth, bones, hands, feet, and tails. However, these products are typically available on a small scale, and almost always as a by-product of hunting for food.

Nonetheless, the demand for primate body parts for sale to tourists can be a serious matter if it involves threatened species. The most striking example of this was the slaughter of mountain gorillas (*Gorilla beringei beringei*) in Rwanda and the Democratic Republic of Congo, in order to provide hands and skulls for sale to European tourists (Fossey, 1983). Although rare, this practice still occurs on occasion despite effective, long-term conservation programs in this region.

Hunting primates for sport is fortunately rare and a minor threat to wild populations. It appears to be most prevalent near logging camps and within military zones in remote areas of developing countries, where arms are plentiful and law enforcement is basically nonexistent. Children armed with slingshots and air rifles are often among the worst offenders. More prestigious trophy hunting has also played a role (albeit a minor one) in primate decline. Species such as gorilla were especially desirable quarry for nineteenth-century and early twentieth-century trophy hunters, and the tales of their exploits are recounted in many books. On the whole, however, sport hunting must be considered a very minor factor unless a threatened species is involved, in which case the activity is almost always illegal as well.

A final reason for hunting primates considered here is because they are sometimes considered agricultural pests; for some African and Asian species, this can represent a significant drain on wild populations. The most striking example is that of government-sponsored "monkey-drives" that were common in Sierra Leone several decades ago. Eleven of the country's 15 primate species were routinely shot or driven into nets and clubbed to death during such drives; only three species were considered harmless to farm crops. According to government records, approximately 250,000 monkeys were destroyed in such drives between 1949 and 1952, and these were only the ones actually counted. Bounties were paid for primate heads or tails, and there was no control over the species killed.

The major crop-raiders are usually the more adaptable and widespread species, such as the savanna-dwelling baboons (*Papio*) in Africa and the macaques (*Macaca* spp.) in Asia, but there are also instances of orangutans (*Pongo* spp.) being killed for raiding fruit trees and gorillas (*Gorilla*) being killed for destroying crops. The only Neotropical species regarded as agricultural pests are the capuchins (*Cebus* and *Sapajus*), whose common names sometimes reflect their crop-raiding habits. For example, the common name for the large-headed capuchin (*Cebus macrocephalus*) in Colombia is *maicero* and one of the Surinamese names for the weeper capuchin (*Cebus olivaceus*) is *nyan-karu mongi*, both of which translate as "corn-eater."

It is difficult to assess how much damage primates actually do to crops in different parts of the world. It is equally difficult to determine how effective pest control efforts have been or to what degree they have contributed to the decline of wild primate populations. However, as primate habitats continue to be encroached on, resulting in shortages of other food sources, it is likely that the more adaptable primate species will continue to raid crops and perhaps become more dependent on them as a food source. This, unfortunately, will likely result in an increased conflict between man and nonhuman primates.

Live Capture of Primates

Primates have routinely been captured alive for export (the international trade to supply zoos and for biomedical research and pharmaceutical testing) or to serve local pet trades. The height of the international primate trade began at the end of the 1950s and continued through the early 1960s, during which time hundreds of thousands of monkeys were taken from the wild each year (Mack and Mittermeier, 1984). The trade consisted largely of rhesus macaques (*M. mulatta*), exported from India and used in laboratory tests as part of the effort to develop a vaccine for polio, and squirrel monkeys (*Saimiri*) imported by the United States from several

Amazonian countries. Subsequently, the imposition of export bans by habitat countries, import restrictions by user countries, and a decreased demand from biomedical research and zoological parks contributed to a significant decline in the international traffic in primates. Currently, the capture and exploitation of long-tailed macaques (*Macaca fascicularis*) from Indochinese countries is of particular concern in this respect.

In 1982, in recognition of the serious effect that live capture for export can have on wild primate populations, the IUCN Species Survival Commission (SSC) Primate Specialist Group (PSG) prepared a *Policy Statement on the Use of Primates for Biomedical Purposes*. The latter included the recommendation that the most threatened species be considered for use in biomedical research projects only if they are obtained from existing, self-sustaining captive breeding colonies. This policy statement was subsequently adopted by the World Health Organization and the Ecosystem Conservation Group of the United Nations, which includes UNESCO, the Food and Agricultural Organization (FAO), UNEP, and IUCN. It is still valid today.

Conservation Status of Primates

The IUCN Red List of Threatened Species provides a regularly updated, comprehensive status assessment of the world's primates. All primate taxa have been identified in this assessment as Extinct, Extinct in the Wild, Threatened (including three categories Critically Endangered, Endangered, and Vulnerable), Near Threatened, Least Concern, Data Deficient, or Not Evaluated. A taxon is defined as:

1. *Critically Endangered* if the rate of population reduction is estimated to be greater than 80% for more than 10 years or three generations, if the extent of its occurrence is estimated to be less than 100 km^2, if its population is estimated to be less than 250 individuals, and quantitative analysis indicates the probability of extinction in the wild is at least 50% within 10 years or three generations.
2. *Endangered* if the rate of population reduction is estimated to be greater than 50% for more than 10 years or three generations, if the extent of its occurrence is estimated to be less than 5000 km^2, if its population is estimated to number less than 2500 individuals, and if quantitative analysis shows the probability of extinction in the wild is at least 20% within 20 years or five generations.
3. *Vulnerable* if the rate of population reduction is estimated to be greater than 30% for more than 10 years or three generations, if the extent of its occurrence is estimated to be less than 20,000 km^2, if its population is estimated to number less than 10,000 individuals, and if quantitative analysis shows the probability of extinction in the wild is at least 10% within 100 years.

As a result of this assessment, 309 (46%) of the world's 672 primate taxa are considered Critically Endangered, Endangered, or Vulnerable. Of these, 204 taxa (30%) are listed as Critically Endangered or Endangered – 49 in the Neotropics, 42 in Africa, 26 in Madagascar, and 87 in Asia (**Table 4** and **Table 5**). At least 23 genera (*Nycticebus*, *Hapalemur*, *Prolemur*, *Varecia*, *Propithecus*, *Indri*, *Callibella*, *Leontopithecus*, *Oreonax*, *Brachyteles*, *Nasalis*, *Simias*, *Rhinopithecus*, *Pygathrix*, *Allochrocebus*, *Mandrillus*, *Rungwecebus*, *Hoolock*, *Nomascus*, *Symphalangus*, *Pongo*, *Gorilla*, and *Pan*) are considered threatened.

Primate Conservation in the New Millennium

The IUCN was established in 1948 to promote and carry out scientifically based action for the conservation and sustainable use of living natural resources. IUCN enrolls sovereign states, governmental agencies, research institutions, and nongovernmental organizations to conserve the world's natural heritage. The SSC was founded in 1949 and is the largest of the IUCN's six commissions, with some 8000 volunteer member scientists, field researchers, government officials, and conservation leaders from almost every country in the world. SSC works principally through its more than 120 specialist groups and Task Forces, of which the PSG is one of the largest. There is a section of about 115 members which deals specifically with great apes, and an additional nine sections for regions in which primates occur (West, Central and East Africa, Madagascar, the Andes, Mesoamerica, Brazil and the Guianas, South Asia, Southeast Asia, and China), each of them with at least one regional coordinator. The total number of members exceeds 400.

The founding mission of the PSG is to maintain the current diversity of the order Primates, with dual emphasis on (1) ensuring the survival of Critically Endangered, Endangered, and Vulnerable species wherever they occur, and (2) providing effective protection for large numbers of primates in areas of high primate diversity and abundance. Much of the work of the PSG over the past few decades has been made possible by Conservation International, an Arlington, VA-based NGO which has housed the Chair and the Deputy Chair over that period and has provided and facilitated many activities on behalf of primates. These include those mentioned above, and also large-scale protection of many areas of primate rainforest habitat through the Global Conservation Fund, the Critical Ecosystem Partnership Fund, and the Centers for Biodiversity Conservation in South America and Madagascar. Other key institutions that contributed significantly to primate conservation during this period included the New York Zoological Society (now Wildlife Conservation Society), the Fauna and Flora Preservation Society (now Fauna & Flora International), the Jersey Wildlife Preservation Trust (now Durrell Wildlife Conservation Trust), Wildlife Preservation Trust International (now EcoHealth Alliance), and the National Geographic Society.

In late 1977, the chairman of the PSG, in collaboration with group members, wrote *Global Strategy for Primate Conservation* (Mittermeier, 1978). This document was an attempt to organize primate conservation activities based on the highest international priorities, and to ensure that limited funds available for primate conservation were put to the best possible use. The first draft of the *Global Strategy* included 65 projects in Africa, Asia, and South and Central America. Each project was categorized as highest priority, high priority, priority, and desirable based mainly on the status of the focal

Table 4 Critically Endangered and Endangered primates

Critically Endangered	*Endangered*
Rondo dwarf galago (*Galagoides rondoensis*)	Zanzibar dwarf galago (*Galagoides zanzibaricus zanzibaricus*)
Sahafary sportive lemur (*Lepilemur septentrionalis*)	Bioko squirrel galago (*Sciurocheirus alleni alleni*)
Lac Alaotra bamboo lemur (*Hapalemur alaotrensis*)	Bioko needle-clawed galago (*Euoticus pallidus pallidus*)
Greater bamboo lemur (*Prolemur simus*)	Lowland red slender loris (*Loris tardigradus tardigradus*)
Variegated black-and-white ruffed Lemur (*Varecia variegata variegata*)	Horton Plains red slender loris (*Loris tardigradus nycticeboides*)
Southern black-and-white ruffed Lemur (*Varecia variegata editorum*)	Highland gray slender loris (*Loris lydekkerianus grandis*)
Northern black-and-white ruffed Lemur (*Varecia variegata subcincta*)	Northern gray slender loris (*Loris lydekkerianus nordicus*)
Silky sifaka (*Propithecus candidus*)	Javan slow loris (*Nycticebus javanicus*)
Perrier's sifaka (*Propithecus perrieri*)	Madame Berthe's mouse lemur (*Microcebus berthae*)
Indri (*Indri indri*)	Golden-brown mouse lemur (*Microcebus ravelobensis*)
Natuna tarsier (*Tarsius bancanus natunensis*)	Tavarata mouse lemur (*Microcebus tavaratra*)
Cotton-top tamarin (*Saguinus oedipus*)	Sambirano mouse lemur (*Microcebus sambiranensis*)
Black-faced lion tamarin (*Leontopithecus caissara*)	Ankarana sportive lemur (*Lepilemur ankaranensis*)
Blond masked titi (*Callicebus barbarabrownae*)	Golden bamboo lemur (*Hapalemur aureus*)
Black bearded saki (*Chiropotes satanas*)	Blue-eyed black lemur (*Eulemur flavifrons*)
Margarita Island tufted capuchin (*Sapajus apella margaritae*)	Sanford's lemur (*Eulemur sanfordi*)
Blond capuchin (*Sapajus flavius*)	White-collared lemur (*Eulemur cinereiceps*)
Yellow-breasted capuchin (*Sapajus xanthosternos*)	Red ruffed lemur (*Varecia rubra*)
Trinidadian white-fronted capuchin (*Cebus albifrons trinitatis*)	Western woolly lemur (*Avahi occidentalis*)
Equatorial capuchin (*Cebus albifrons aequatorialis*)	Bemaraha woolly lemur (*Avahi cleesei*)
Ka'apor capuchin (*Cebus kaapori*)	Coquerel's sifaka (*Propithecus coquereli*)
Northern brown howler (*Alouatta guariba guariba*)	Crowned sifaka (*Propithecus coronatus*)
Mexican mantled howler (*Alouatta palliata mexicana*)	Tattersall's sifaka (*Propithecus tattersalli*)
Azuero howler (*Alouatta palliata trabeata*)	Diademed sifaka (*Propithecus diadema*)
Colombian woolly monkey (*Lagothrix lugens*)	Milne-Edwards' sifaka (*Propithecus edwardsi*)
Yellow-tailed woolly monkey (*Oreonax flavicauda*)	Horsfield's tarsier (*Cephalopachus bancanus bancanus*)
Variegated spider monkey (*Ateles hybridus hybridus*)	Belitung tarsier (*Cepahalopachus bancanus saltator*)
Brown spider monkey (*Ateles hybridus brunneus*)	Peleng tarsier (*Tarsius pelengensis*)
Geoffroy's spider monkey (*Ateles geoffroyi geoffroyi*)	Sangihe tarsier (*Tarsius sangirensis*)
Mexican spider monkey (*Ateles geoffroyi vellerosus*)	Buffy-headed marmoset (*Callithrix flaviceps*)
Azuero spider monkey (*Ateles geoffroyi azuerensis*)	Pied bare-faced tamarin (*Saguinus bicolor*)
Brown-headed spider monkey (*Ateles fusciceps fusciceps*)	White-footed tamarin (*Saguinus leucopus*)
Colombian black spider monkey (*Ateles fusciceps rufiventris*)	Golden lion tamarin (*Leontopithecus rosalia*)
Northern muriqui (*Brachyteles hypoxanthus*)	Golden-headed lion tamarin (*Leontopithecus chrysomelas*)
Miss Waldron's red colobus (*Piliocolobus waldronae*)	Black lion tamarin (*Leontopithecus chrysopygus*)
Preuss' red colobus (*Piliocolobus preussi*)	Beni titi (*Callicebus modestus*)
Niger Delta red colobus (*Piliocolobus epieni*)	Olalla's titi (*Callicebus olallae*)
Bouvier's red colobus (*Piliocolobus bouvieri*)	Andean titi (*Callicebus oenanthe*)
Western purple-faced langur (*Trachypithecus vetulus nestor*)	Coimbra-Filho's titi (*Callicebus coimbrai*)
Raffles' silvered langur (*Trachypithecus cristatus vigilans*)	Uta Hick's bearded saki (*Chiropotes utahickae*)
Delacour's langur (*Trachypithecus delacouri*)	Red-nosed bearded saki (*Chiropotes albinasus*)
Cat Ba langur (*Trachypithecus poliocephalus poliocephalus*)	Black-crowned Central American squirrel monkey (*Saimiri oerstedii oerstedii*)
White-headed langur (*Trachypithecus poliocephalus leucocephalus*)	Gray-crowned Central American squirrel monkey (*Saimiri oerstedii citrinellus*)
Golden-bellied Mentawai langur (*Presbytis potenziani potenziani*)	Robust tufted capuchin (*Sapajus robustus*)
Western cross-marked surili (*Presbytis chrysomelas chrysomelas*)	Varied white-fronted capuchin (*Cebus albifrons versicolor*)
Eastern cross-marked surili (*Presbytis chrysomelas cruciger*)	Santa Marta white-fronted capuchin (*Cebus albifrons malitiosus*)
Pagai pig-tailed langur (*Simias concolor concolor*)	Maranhão red-handed howler monkey (*Alouatta ululata*)
Siberut pig-tailed langur (*Simias concolor siberu*)	Yucatán black howler monkey (*Alouatta pigra*)
Yunnan snub-nosed monkey (*Rhinopithecus bieti*)	Geoffroy's woolly monkey (*Lagothrix cana cana*)
Guizhou snub-nosed monkey (*Rhinopithecus brelichi*)	Black-faced black spider monkey (*Ateles chamek*)
Tonkin snub-nosed monkey (*Rhinopithecus avunculus*)	White-bellied spider monkey (*Ateles belzebuth*)
Red-shanked douc (*Pygathrix nemaeus*)	White-whiskered spider Monkey (*Ateles marginatus*)
Gray-shanked douc (*Pygathrix cinerea*)	Yucatán spider monkey (*Ateles geoffroyi yucatanensis*)
Schouteden's blue monkey (*Cercopithecus mitis schoutedeni*)	Ornate spider monkey (*Ateles geoffroyi ornatus*)
Somalian white-collared guenon (*Cercopithecus albogularis zammaranoi*)	Southern muriqui (*Brachyteles arachnoides*)
Dryas monkey (*Cercopithecus dryas*)	Bioko black colobus (*Colobus satanas satanas*)
Kipunji (*Rungwecebus kipunji*)	Prigogine's Angolan colobus (*Colobus angolensis prigoginei*)
Pagai macaque (*Macaca pagensis*)	Mt, Uaraguess guereza (*Colobus guereza percivali*)
Sulawesi black macaque (*Macaca nigra*)	Upper Guinea red colobus (*Piliocolobus badius badius*)
Tonkin black crested gibbon (*Nomascus concolor concolor*)	Temminck's red colobus (*Piliocolobus badius temminckii*)

(*Continued*)

Table 4 Continued

Critically Endangered	*Endangered*
Laotian black crested gibbon (*Nomascus concolor lu*)	Pennant's red colobus (*Piliocolobus pennantii*)
Wuliang black crested gibbon (*Nomascus concolor jingdongensis*)	Kirk's red colobus (*Piliocolobus kirkii*)
West Yunnan black crested gibbon (*Nomascus concolor furvogaster*)	Tana River red colobus (*Piliocolobus rufomitratus*)
Cao-Vit crested gibbon (*Nomascus nasutus*)	Ashy red colobus (*Piliocolobus tephrosceles*)
Hainan crested gibbon (*Nomascus hainanus*)	Udzungwa red colobus (*Piliocolobus gordonorum*)
Northern white-cheeked crested gibbon(*Nomascus leucogenys*)	Kashmir langur (*Semnopithecus ajax*)
Sumatran orangutan (*Pongo abelii*)	Gray-handed crested langur (*Semnopithecus priam thersites*)
Western lowland gorilla (*Gorilla gorilla gorilla*)	Southern purple-faced langur (*Trachypithecus vetulus vetulus*)
Cross River gorilla (*Gorilla gorilla diehli*)	Northern purple-faced langur (*Trachypithecus vetulus philbricki*)
Mountain gorilla (*Gorilla beringei beringei*)	Highland purple-faced langur (*Trachypithecus vetulus monticola*)
	Western Phayre's langur (*Trachypithecus phayrei phayrei*)
	Shan States Phayre'slLangur (*Trachypithecus phayrei shanicus*)
	Indochinese gray langur (*Trachypithecus phayrei crepusculus*)
	Germain's langur (*Trachypithecus germaini*)
	Capped langur (*Trachypithecus pileatus pileatus*)
	Lowland capped langur (*Trachypithecus pileatus durga*)
	Bhutan capped langur (*Trachypithecus pileatus tenebricus*)
	Shortridge's langur (*Trachypithecus shortridgei*)
	Golden langur (*Trachypithecus geei*)
	Francois' langur (*Trachypithecus francoisi*)
	Hatinh langur (*Trachypithecus hatinhensis*)
	Javan surili (*Presbytis comata*)
	East Sabah Hose's surili (*Presbytis hosei sabana*)
	East Kalimantan Hose's surili (*Presbytis hosei canicrus*)
	Somber-bellied Mentawai surili (*Presbytis potenziani siberu*)
	Northern mitred surili (*Presbytis melalophos sumatrana*)
	Eschscholtz's mitred surili (*Presbytis melalophos mitrata*)
	Stripe-naped proboscis monkey (*Nasalis larvatus larvatus*)
	Plain-naped proboscis monkey (*Nasalis larvatus orientalis*)
	Moupin golden snub-nosed monkey (*Rhinopithecus roxellana roxellana*)
	Qinling golden snub-nosed monkey (*Rhinopithecus roxellana qinlingensis*)
	Hubei golden snub-nosed monkey (*Rhinopithecus roxellana hubeiensis*)
	Black-shanked douc (*Pygathrix nigripes*)
	Golden monkey (*Cercopithecus kandti*)
	Red-bellied monkey (*Cercopithecus erythrogaster erythrogaster*)
	Roloway monkey (*Cercopithecus roloway*)
	Cameroon Preuss's monkey (*Allochrocebus preussi preussi*)
	Bioko Preuss's monkey (*Allohcrocebus preussi insularis*)
	White-naped mangabey (*Cercocebus lunulatus*)
	Tana River mangabey (*Cercocebus galeritus*)
	Sanje River mangabey (*Cercocebus sanjei*)
	Mainland drill (*Mandrillus leucophaeus leucophaeus*)
	Bioko drill (*Mandrillus leucophaeus poensis*)
	Barbary macaque (*Macaca sylvanus*)
	Lion-tailed macaque (*Macaca silenus*)
	Moor macaque (*Macaca maura*)
	Dry-zone Toque macaque (*Macaca sinica sinica*)
	Wet-zone Toque macaque (*Macaca sínica aurifrons*)
	Hill-zone Toque macaque (*Macaca sinica opisthomelas*)
	Arunachal macaque (*Macaca munzala*)
	Malayan white-handed gibbon (*Hylobates lar lar*)
	Sumatran white-handed gibbon (*Hylobates lar vestitus*)
	Carpenter's white-handed gibbon (*Hylobates carpenteri*)
	Mountain agile gibbon (*Hylobates agili agilis*)
	Müller's gray gibbon (*Hylobates muelleri muelleri*)
	Northern gray gibbon (*Hylobates muelleriri funereus*)
	Abbott's gray gibbon (*Hylobates muelleri abbotti*)
	Bornean agile gibbon (*Hylobates albibarbis*)
	Silvery gibbon (*Hylobates moloch*)
	Capped gibbon (*Hylobates pileatus*)
	Kloss' gibbon (*Hylobates klossii*)

(*Continued*)

Table 4 Continued

Critically Endangered	*Endangered*
	Western hoolock gibbon (*Hoolock hoolock*) Yellow-cheeked crested gibbon (*Nomascus gabriellae*) Southern white-cheeked crested gibbon (*Nomascus siki*) Siamang (*Symphalangus syndactylus*) Western Bornean orangutan (*Pongo pygmaeus pygmaeus*) Southern Bornean orangutan (*Pongo pygmaeu wurmbii*) Eastern Bornean orangutan (*Pongo pygmaeus morio*) Grauer's or eastern lowland gorilla (*Gorilla beringei graueri*) Central chimpanzee (*Pan troglodytes troglodytes*) Eastern chimpanzee (*Pan troglodytes schweinfurthii*) Elliot's chimpanzee (*Pan troglodytes ellioti*) Western chimpanzee (*Pan troglodytes verus*) Bonobo (*Pan paniscus*)

Table 5 Threatened primates by region

Region	*Taxa*	*Threatened (%)*	*Critically Endangered (%)*	*Endangered (%)*
Asia	182	124 (68%)	26 (14%)	61 (34%)
Madagascar	101	40 (40%)	8 (8%)	18 (18%)
Neotropics	201	79 (39%)	23 (11%)	26 (13%)
Africa	186	66 (35%)	12 (6%)	30 (21%)

species and how likely the project would be to bring about the desired conservation action. The *Global Strategy* quickly resulted in a substantial increase in funding for primate conservation activities, and in 1979 it led to the establishment of a special Primate Program and Primate Action Fund by the World Wildlife Fund-US. In addition to major endeavors supported as a result of this program, the Primate Action Fund provided rapid support for small primate conservation projects, workshops, and relevant events (ranging from $500 to $5000). It has now functioned for more than 30 years, contributing, in the process, several million dollars to hundreds of worthy projects.

Almost a decade after the *Global Strategy* was launched, the first regional primate conservation action plans were prepared by the IUCN/SSC PSG. First to be published was the *Action Plan for African Primate Conservation: 1986–1990* (Oates, 1986), which was quickly followed by the *Action Plan for Asian Primate Conservation: 1987–1991* (Eudey, 1987), and several years later by *Lemurs of Madagascar, An Action Plan for Their Conservation: 1993–1999* (Mittermeier *et al.*, 1992) and by *African Primates: Status Survey and Conservation Action Plan* (Oates, 1996), the latter being an update of the 1986 document. The PSG has also produced regional action plans, shorter publications which put primate conservation into a comprehensive regional perspective: among these are *Chimpanzees in West Africa* (Kormos *et al.*, 2003), *Chimpanzees and Gorillas in Western Equatorial Africa* (Tutin *et al.*, 2005), *Sumatran Orangutan Conservation Action Plan* (Ellis *et al.*, 2006), and *Regional Action Plan for the Conservation of the Cross River Gorilla* (Oates *et al.*, 2007). Together, these action plans have effectively focused conservation activities in three or four major regions in which primates occur, and served as useful measures with regard to the success of proposed strategies.

The first vehicle for regular and effective communication among the world's primate conservationists was the *IUCN/SSC Primate Specialist Group Newsletter*, which was launched in 1981. Changed to *Primate Conservation* in 1985, it has gradually evolved into a journal that continues to be published annually. In addition, the four regional sections of the PSG subsequently began publishing their own periodic newsletters to meet the increasing need for more timely information. *Asian Primates* appeared in 1991 and was published until 2002; after a lengthy hiatus, it has recently been replaced by an annual *Asian Primates* journal. *African Primates* debuted in 1995 and has had a similar history, having been relaunched in 2010 as a yearly journal. *Neotropical Primates* and *Lemur News* appeared in 1993, and are both still published regularly as journals. Taken in combination, these periodicals have significantly increased the amount, quantity, and timeliness of information available to primate conservationists throughout the world. Other important aspects of the PSG involve the organization of workshops and meetings, in order to discuss program activities in addition to aspects such as taxonomy and the regional coordination of conservation measures. The group also maintains databases and a website that serves as a valuable resource for members, researchers and the public alike, providing information on primate diversity, threat status, and current conservation efforts. A general media outreach on the importance of primate conservation, to ensure that it remains regularly in the news, is also a major priority. Last but not least, there is the continual search for additional resources to fund primate conservation in the field.

Outlook for the Future

In the future, there is a need to sustain conservation activities based on recommendations of the original *Global Strategy for Primate Conservation* and subsequent publications, as well as to increase the focus on those primate taxa deemed closest to extinction. The twentieth century ended without a single primate taxon having being lost – an enviable record indeed considering the number of reptiles, birds, and other mammals known to have disappeared during this period. However, at least one form has already been officially declared extinct in the new millennium, and many more species and subspecies are in serious jeopardy. Much of the groundwork for developing a more focused conservation strategy has already been done. The IUCN Red List of Threatened Animals provides a good starting point for identifying the highest priority taxa. To those that are considered Critically Endangered and Endangered, we expect that the PSG will continue to add more species and subspecies, while at the same time considering factors such as taxonomic uniqueness in order to establish priority rankings for conservation action. The results of much of this analysis are nowadays presented in *The World's 25 Most Endangered Primates*, a biennial publication that serves to draw attention to the most critically threatened forms. This synthesis and publication of information is vital for fomenting and stimulating measures for the conservation of primates and their forests worldwide.

In addition to the above activities, there are also other important priorities that need to be maintained in the coming years. These include an increase in primate-focused ecotourism, coupled with the training of nationals from primate habitat countries, the latter by way of the provision of small- to medium-sized grants for primate research and primate conservation measures in the field. New national parks and reserves need to be set aside for threatened species as well as in areas of high primate diversity and abundance, and those that already exist need to be strengthened in terms of their infrastructure as well as in the enforcement of protective legislation within them. On a broader scale, there is a need to determine ways in which humans and nonhuman primates can coexist better in multiple-use areas. Zoos must also continue to do their part, mainly by establishing conservation-oriented captive breeding programs for threatened taxa and by helping to end the illegal and otherwise destructive traffic in primates. Those research institutions using primates need to be made more aware of conservation issues and the status of the species that they use, ensure that they use primates as prudently as possible, and make every attempt to breed most or all of the primates that they require. Perhaps most important of all, public awareness needs to be created regarding the need for primate conservation and the importance of primates as a natural heritage in the countries in which they occur.

The IUCN, together with its partners, has long years of experience in assessing conservation status, developing action plans, and implementing conservation action using both local expertise in habitat countries (which is growing rapidly), and expat specialists working in close collaboration. This differs from many other less studied groups of organisms where either the manpower or the expertise, or both, are still in short supply. We know where the highest priority areas are located, and, with the exception of new species that have been discovered very recently – such as the extraordinary Stryker's snub-nosed monkey (*Rhinopithecus strykeri*) from Myanmar – we have a very good handle on conservation status. With the help of all these plans and publications, the task of amassing both the human and the financial resources needed for implementation is well underway. Although the World Wildlife Fund-US Primate Program no longer exists, many other traditional nongovernmental sources still offer grants for field, captive, and laboratory programs; academic institutions continue to provide funds for primate field studies that have significant conservation impact; government-supported efforts such as the Indo-US Primate Project provide excellent models for international cooperation; and an increasing number of zoos have joined forces to focus on regional primate faunas, generating funds not only for captive breeding programs but also for *in situ* projects. In addition, at least two new significant sources of philanthropic support dedicated to primates have been established in recent years: Primate Conservation Inc. (PCI), and the Margot Marsh Biodiversity Foundation. Several other new endowments also provide considerable annual funding for primate projects, among them the Arcus Foundation and the Mohamed bin Zayed Species Conservation Fund. Together, these organizations and agencies represent the core of funding necessary to move ahead with a global strategy for the most seriously threatened primates, and perhaps to help uncover new sources of support as well.

What we have so far is sufficient funding to deploy these experts into the highest priority places with the greatest extinction risk. Now, however, the time has come to begin ramping up our efforts, specifically in the greater deployment of specialists to carry out all the conservation projects needed. At this point of time, direct action for primates is on the order of \$10 My^{-1}, with more than 90% of that going to the great apes (which constitute only about 2% of the known species and subspecies of primates). To be truly effective, and to ensure that no primate species go extinct in the coming years, we need to increase that by an order of magnitude, to \$100 My^{-1}. Only in this way, we will be able to guarantee that the full diversity of the order Primates is maintained into the new millennium.

The more we learn, the more we see how important primates are – as flagship species for education and awareness efforts for rainforests, as pollinators, and as seed dispersers that may even be instrumental in maintaining high carbon sequestration capacity in the forests of Amazonia, Central Africa, and Southeast Asia. Also, as the need grows to evaluate, verify, and monitor the long-term health of tropical forests, a number of primate species will serve as some of the best indicators of forest health. Although activities under way in many parts of the world make it inevitable that a portion of the world's forests and primates living in them will disappear, our challenge is to try and minimize this loss wherever possible.

See also: Biodiversity-Rich Countries. Captive Breeding and Reintroduction. Conservation Efforts, Contemporary. Diversity and Conservation of Neotropical Mammals. Endangered Mammals.

Endemism. Framework for Assessment and Monitoring of Biodiversity. Human Impact on Biodiversity, Overview. Mammals, Biodiversity of. Mammals, Conservation Efforts for. Natural Reserves and Preserves. Rainforest Ecosystems, Animal Diversity. Rainforest Loss and Change. Tourism, Role of. Zoos and Zoological Parks

References

Amman K (2000) Exploring the bushmeat trade. In: Amman K (ed.) *Bushmeat: Africa's Conservation Crisis*, pp. 16–27. London: World Society for the Protection of Animals.

Ellis S, Singleton I, Andayani N, Traylor-Holzer K, and Supriatna J (eds.) (2006) *Sumatran Orangutan Conservation Action Plan*. Washington, DC/Jakarta, Indonesia: Conservation International.

Eudey AA (1987) *Action Plan for Asian Primate Conservation: 1987–1989*. Gland, Switzerland: IUCN/SSC Primate Specialist Group.

Fossey D (1983) *Gorillas in the Mist*. New York: Houghton Mifflin.

Galdikas BMF (1995) *Reflections of Eden: My Years with the Orangutans of Borneo*. Boston: Little, Brown and Co.

Goodall J (1968) The behaviour of free-living chimpanzees in the Gombe Stream Reserve. *Animal Behavior Monographs* 1(3): 161–311.

IUCN (2011) *IUCN Red List of Threatened Animals*. Available online at: http://www.iucnredlist.org

Kormos R, Boesch C, Bakarr MI, and Butynski TM (eds.) (2003) *West African Chimpanzees: Status Survey and Conservation Action Plan*. Gland, Switzerland: IUCN – The World Conservation Union.

Mack D and Mittermeier RA (1984) *The International Primate Trade*, vol. 1. Washington, DC: TRAFFIC.

Mittermeier RA (1978) *A Global Strategy for Primate Conservation*. Washington, DC: IUCN/SSC Primate Specialist Group.

Mittermeier RA and Coimbra-Filho AF (1977) Primate Conservation in Brazilian Amazonia. In: Prince Rainier of Monaco and Bourne GH (eds.) *Primate Conservation*, pp. 117–166. New York: Academic Press.

Mittermeier RA, Oates JF, Eudey AE, and Thornback J (1986) Primate conservation. In: Mitchell G and Erwin J (eds.) *Comparative Primate Biology, Volume 2A: Behavior, Conservation and Ecology*, pp. 3–72. New York: A. R. Liss.

Mittermeier RA, Konstant WR, Nicoll ME, and Langrand O (1992) *Lemurs of Madagascar: An Action Plan For Their Conservation (1993–1999)*. Gland, Switzerland: IUCN/SSC Primate Specialist Group.

Nekaris KAI, Shepherd CR, Starr CR, and Nijman V (2010) Exploring cultural drivers for wildlife trade via an ethnoprimatological approach: a case study of slender and slow lorises (*Loris* and *Nycticebus*) in South and Southeast Asia. *American Journal of Primatology* 71: 1–10.

Nekaris KAI and Nijman V (2007) CITES proposal highlights rarity of Asian nocturnal primates (Lorisidae: *Nycticebus*). *Folia Primatologica* 78: 211–214.

Nijman V (2005) *In Full Swing: An Assessment of trade in Orang-Utans and Gibbons on Java and Bali, Indonesia*. Petaling, Selangor, Malaysia: TRAFFIC Southeast Asia.

Oates JF (1986) *Action Plan for African Primate Conservation: 1986–1990*. Gland, Switzerland: IUCN/SSC Primate Specialist Group.

Oates JF (1996) *African Primates: Status Survey and Conservation Action Plan*. Gland, Switzerland: IUCN/SSC Primate Specialist Group, IUCN – The World Conservation Union.

Oates JF, Sunderland-Groves J, Bergl R, *et al.* (2007) *Regional Action Plan for The Conservation of the Cross River Gorilla (Gorilla gorilla diehli)*. Arlington, VA: IUCN/SSC Primate Specialist Group and Conservation International.

Peres CA (2000) Evaluating the impact and sustainability of subsistence hunting at multiple Amazonian forest sites. In: Robinson JG and Bennett EL (eds.) *Hunting for Sustainability in Tropical Forests*, pp. 31–36. New York: Columbia University Press.

Redford KH (1992) The empty forest. *BioScience* 42: 412–422.

Schaller G (1963) *The Mountain Gorilla, Ecology and Behavior*. Chicago: University of Chicago Press.

Tutin C, Stokes E, Boesch C, *et al.* (2005) *Regional Action Plan for the Conservation of Chimpanzees and Gorillas in Western Equatorial Africa*. Washington, DC: IUCN/SSC Primate Specialist Group and Conservation International.

Priority Setting for Biodiversity and Ecosystem Services

Derric Pennington, University of Minnesota, St Paul, MN, USA, and World Wildlife Fund – US, Washington, DC, USA
Taylor Ricketts, University of Vermont, Burlington, VT, USA
Robin Naidoo, World Wildlife Fund – US, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acquisition costs Costs of acquiring property rights to a parcel of land. Acquisition of property rights can be total (i.e., the land and title are sold to a conservation agent) or partial. Partial transfers of property rights include short-term land rental, conservation easements, and contracts between conservation agents and landowners who exchange money for land management that enhances conservation value.

Biodiversity Contraction of "biological diversity." Refers to the variety of life on Earth, including the numbers of, and differences among alleles, genes, individuals, populations, species, communities, and ecosystems. Because this term is very broad, measures of biodiversity typically focus on one aspect of biodiversity at a time (e.g., the number of species).

Conservation planning The science of identifying spatially explicit priorities and actions for the conservation of a region's biodiversity or natural capital.

Damage costs Costs associated with damages to economic activities arising from conservation programs. For example, damages to crops and livestock from wild animals living in protected areas adjacent to human settlements can result in significant losses in income. In other cases, direct wildlife attacks might physically harm or kill humans, resulting in further economic losses.

Ecosystem services The benefits people obtain (actively or passively) from ecosystems (MEA, 2005).

Efficiency frontier By combining biophysical and/or economic models with optimization methods, one can determine the maximum feasible combinations of values that can be achieved for a landscape. The results can be presented as a frontier, which is defined as the outcomes where it is not possible to improve on any particular objective without decreasing some other objective.

Management costs Costs associated with the management of a conservation program, such as those associated with establishing and maintaining a network of protected areas. Management costs can be fixed and therefore independent of the amount of conservation activities pursued (e.g., regardless of how much land is protected in an area, an office will need to be opened and a minimal number of staff hired) or variable, such as when a function of the amount and type of conservation intervention.

Opportunity costs Costs of foregone opportunities; that is, a measure of what could have been gained via the next-best use of a resource had it not been put to the current use. In terrestrial protected areas where extractive uses are forbidden, the opportunity cost represents the highest-value extractive use for that land. When purchasing land or conservation easements from private land owners, payments will reflect the value of lost opportunities. With public land or with regulation, direct financial obligations might be divorced from the value of lost opportunities. From a social perspective, it is important to include opportunity costs to track the full set of consequences of conservation planning (Naidoo *et al.*, 2006).

Spatially explicit ecological production function A mathematical expression that translates spatially explicit information on an ecosystem feature or processes to some ecological outcome or condition in the ecosystem. Estimation of production functions isolates the effect of ecosystem services as inputs to the production process. Once an ecological production function is specified, researchers can quantify the impact of landscape change on the level of ecosystem service output.

Transaction costs Costs associated with negotiating an economic exchange. In a terrestrial conservation context, the costs over and above the price of a transfer of property rights to a given parcel of land. These include the costs of searching for properties, negotiating with individual landholders, and obtaining approval for title transfer. Transaction costs can be substantial; for example, carbon sequestration projects involving afforestation or reforestation can be beneficial for conservation, but high transaction costs often limit their viability (Naidoo *et al.*, 2006).

Introduction

Conservation has been called a "crisis discipline" (Pullin, 2002). Indeed, there are several conservation crises – loss and degradation of ecosystems, climate change, to name a few – that impact biodiversity and people's well-being. Instead of responding haphazardly to one crisis or threat after another, it may be more effective to have a plan that considers not only the present concern (e.g., deforestation) but also develops a clear vision to avoid future emergencies (e.g., climate change). For these plans to be implemented effectively, it requires acknowledging that conservation does not take place in a sociopolitical vacuum (Sekhar, 2005). A key challenge of modern conservation planning is to anticipate future trends and threats, and develop conservation plans that are as efficient as possible so that they can win the support of stakeholders and provide enough flexibility to adapt to ever-changing conditions (Noss *et al.*, 2002).

Systematic conservation planning is the process of locating, configuring, implementing, and maintaining areas to promote the persistence of biodiversity and other natural values. Moreover, it is an activity whereby social, economic, and political imperatives (i.e., the real world) modify, sometimes drastically, conventional scientific recommendations (see Margules and Pressey, 2000). Historically, the practice of conservation planning was not very systematic and as a result many areas designated for conservation are located on lands, which at the time of establishment were either too remote or unproductive to be economically important (e.g., Pressey *et al.*, 1996). In the case of biodiversity, this means that many species occurring in economically productive landscapes are not protected, although threats (e.g., habitat loss and fragmentation) continue. Indeed, the variety of social, economic, and political goals factoring into conservation planning has led to a diversity of reasons why reserves are established – different proponents see different places differently. Consequently, there has been a perceived competition among stakeholders on how best to target limited conservation resources. Systematic conservation planning has been slowly evolving over the past 25 years to better incorporate these challenges into the planning process (Pressey and Bottrill, 2008).

Most of the early examples of systematic conservation planning set priorities based on biological goals, such as the citing of parks and marine protected areas that maximize biodiversity significance, often measured as the number of endemic or rare and/or threatened and endangered species, and guided by three key concepts: complementarity, irreplaceability, and vulnerability (e.g., Brooks *et al.*, 2006; Spalding *et al.*, 2007; Abell *et al.*, 2008). For instance, Conservation International (CI) has identified 25 biodiversity hotspots that comprise 1.4% of the world's land surface but contain representatives of 44% of known vascular plant species (Myers *et al.*, 2000). Other global conservation plans target areas that maintain ecological and evolutionary processes important for biodiversity persistence (e.g., interspecific interactions, regular and nomadic faunal movements, and disturbance regimes; Cowling *et al.*, 1999; Rouget *et al.*, 2003). For example, World Wildlife Fund – US has divided the terrestrial world into 867 different ecoregions, which are characterized by relatively large areas, each comprising a unique mix of environmental conditions, flora, and fauna. Of these, 233 regions ("the Global 200 (G200)") represent biodiversity that is globally significant (Olson *et al.*, 2001). These approaches provide a conceptual framework for identifying large-scale conservation area networks but fail to consider the on-the-ground realities that often impede implementation.

As systematic conservation planning evolved, planners of conservation networks also consider the financial and opportunity cost of network conservation and implementation (e.g., Ando *et al.*, 1998). Designating land for biodiversity conservation competes with other societal objectives, such as housing developments, recreation, agricultural or industrial development, and resource extraction. Gaps between analyses of biological data and real-world conservation have become less distinct through time. An initial framework for the process of systematic conservation planning (Margules and Pressey, 2000) showed how this aspect of implementation related to other tasks. An improved version of this framework made explicit the need to address socioeconomic considerations at the outset (Cowling and Pressey, 2003). Moreover, planning requires consideration of the "enabling process" – how does the foundational science actually get implemented on the ground? Ensuring success requires appropriate institutions and governance structures to exist to enable the effective use of targeted instrumental interventions to address biodiversity loss (Rands *et al.*, 2010). In other words, the revised fundamental goal of systematic conservation planning is to pick a subset of sites for conservation that have high biodiversity or other natural value, are threatened, are cheap, and have broad political and institutional support.

Conservationists have recently begun considering another motivation for conservation: ecosystem services, or the goods and services from ecological systems that benefit people (Daily, 1997). These benefits include the provision of abundant and clean water, climate regulation, food production, and recreation opportunities (MEA, 2005). Changing land use or land management affects not only biodiversity but also the provision and value of a variety of ecosystem services. For example, conversion of forests to croplands results not only in increased food production but also loss of habitat for forest dwelling species, release of carbon to the atmosphere, changes in hydrologic and nutrient cycles, and changes in local weather patterns. Unintended negative consequences for ecosystem services threaten to undermine future human well-being and biodiversity, calling into question whether the current trajectories are sustainable.

Conservation plans that focus on biodiversity alone may direct resources away from ecosystems that harbor few rare or endemic species but which support vital ecosystem services. For instance, a typical *Spartina* tidal marsh has no endemic plants and no more than 30 total plant species, yet this ecosystem provides flood regulation, waste treatment, carbon storage, and fisheries production to human communities (Gedan *et al.*, 2009). How does one best balance these various values when making conservation decisions?

If ecosystem services represent the benefits provided to people by conserved ecosystems, then the costs side of the equation is also important. Communities, governments, and donors are increasingly concerned about the effectiveness of the conservation interventions that they fund (Ferraro and Pattanayak, 2006), and consideration of costs can substantially increase the efficiency of their investments. Incorporating spatial distributions of biophysical benefits and economic costs in conservation planning have shown that limited budgets can achieve substantially larger biological gains than when planning ignores costs (e.g., Naidoo *et al.*, 2006). Although there are potential trade-offs between conservation for biodiversity, for ecosystem services, and considering different economic cost constraints, both in terms of conservation financing and to society at large, a systemic planning framework can help to identify synergies.

The number of potentially competing objectives can complicate conservation planning decisions. In rare cases, the best choice among management alternatives is obvious, such as when one alternative delivers higher levels of all ecosystem services and biodiversity compared with other alternatives (i.e., "win-win" solutions). However, in most cases, comparing

alternative management options will require evaluating trade-offs among various conservation objectives, ecosystem services, and biodiversity conservation. The authors, herein, describe recent research that is advancing these frontiers. They cover: Why it is important to consider ecosystem services and costs, along with biodiversity, in the planning process; examples of how ecosystem services can be estimated or modeled in a spatially explicit manner; examples of how plans differ when ecosystem services and costs are formally considered; how costs can fit into a dynamic framework for conservation planning; and end by discussing limitations and future directions for research that integrates biodiversity, ecosystem services, and cost effectiveness in conservation planning.

Broadening Conservation Planning with Ecosystem Services

Efforts to jointly conserve biodiversity and ecosystem services will likely be inefficient unless objectives for both are explicitly considered. Maps of species occurrences and ecosystem ranges provide important data; however, it is also important to consider the distribution and associated pressures of human populations. For example, as of 1995, more than 1 billion people lived within CI's 25 biodiversity hotspots at an annual rate of population growth 40% higher than the global average (Cincotta *et al.*, 2000). The reality of expanding human populations within biodiversity hotspots has led to recent calls from conservationists to include among their targets ecosystem services as well as biodiversity (Balvanera *et al.*, 2001).

The Millennium Ecosystem Assessment (MEA, 2005) documented the importance of ecosystem services to human well-being and showed that continued supply of these services is threatened by unsustainable human activity. Several conservationists have rallied around the idea that the public is more likely to support land protection if the land harbors not only a long list of species but also provides recreational opportunities, flood protection, clean water, and climate regulation – ecosystem services that improve people's well-being (e.g., The Nature Conservancy, 2011). However, more work is needed to better understand the scope for alignment between conservation of biodiversity and areas most important for delivering ecosystem services. Systemic conservation planning provides a good framework for transparent reporting of the trade-offs across targets – where are the potential areas where biodiversity and ecosystem services align and where they do not?

Researchers are newly able to quantify and map ecosystem services in ways analogous to the biodiversity maps used by conservation planners (e.g., Kareiva *et al.*, 2011). For instance, recent work developing and integrating ecological production functions to quantify ecosystem services into an economic framework has made several important advances. First, models have become spatially explicit. In a series of articles, Sanchirico and Wilen (1999, 2001, 2005) developed a bioeconomic model that considered distinct habitat patches and dispersal of a harvested or game species. Others have developed spatially explicit models of ecosystem services as a function of land use and land cover (e.g., Lewis and Plantinga, 2007; Nelson *et al.*, 2008, 2009; Polasky *et al.*, 2008).

Second, models have expanded beyond a single-service focus to consider provision of the bundle of ecosystem services jointly produced by the ecosystem (e.g., Antle and Stoorvogel, 2006; Boody *et al.*, 2005; Coiner *et al.*, 2001; Naidoo and Ricketts, 2006; and Nelson *et al.*, 2008, 2009). Such work allows an analysis of the trade-offs in provision among services, which provides information on the optimal spatial arrangement of land use and land management to meet multiple conservation objectives (e.g., Nelson *et al.*, 2008, 2009; and Polasky *et al.*, 2008).

Consideration of Biodiversity and Ecosystem Service Targets

Important insights on the spatial concordance of biodiversity and the provision of ecosystem services have been revealed from recent studies. At the global scale, studies have shown that regions identified as priorities for traditional conservation reasons, such as their inclusion of endemic species, high biodiversity, or threatened habitats, are not necessarily those that provide the greatest provision of ecosystem services (e.g., Turner *et al.*, 2007; Naidoo *et al.*, 2008; and Larsen *et al.*, 2011).

Naidoo *et al.* (2008) used available global datasets on species distributions within ecoregions and four ecosystem services (carbon storage and sequestration, livestock production, and water yield) to illustrate potential synergies and trade-offs between these targets. They found little spatial concordance among individual services and species, which is in accordance with findings for global range-restricted species (Larsen *et al.*, 2011) and, at a more regional scale (Anderson *et al.*, 2009). When they asked whether existing priority areas (biodiversity hotspots, high-biodiversity wilderness areas (HBWAs), and G200) effectively conserve ecosystem services, they found that service provision in the three schemes varied depending on whether these services were locally versus globally consumed (**Figure 1**). For example, carbon services, which are globally dispersed, were highest in low human-density HBWAs; conversely, locally dispersed services, water yield, and livestock production were highest in areas near densely populated areas, such as biodiversity hotspots. Similarly, Larsen *et al.* (2011) found that the level of concordance with range-restricted species could vary across different services and suggest that particular regions and service combinations can be used to reveal synergies and tradeoffs.

Although measures of biodiversity and ecosystem services do not tend to correlate well, it does not mean that this approach to conservation planning is futile. Indeed, Naidoo *et al.* (2008) not only identified several global priority areas as potential locations where both biodiversity importance and ecosystem services importance were relatively high (those ecoregions in the upper right hand corner of **Figure 2**) but also, and just as informative, the areas where there are trade-offs Naidoo *et al.*, 2008).

Global analysis relying on such coarse data may not be relevant to on-the-ground conservation efforts. However, the same poor correlation between biodiversity and ecosystem

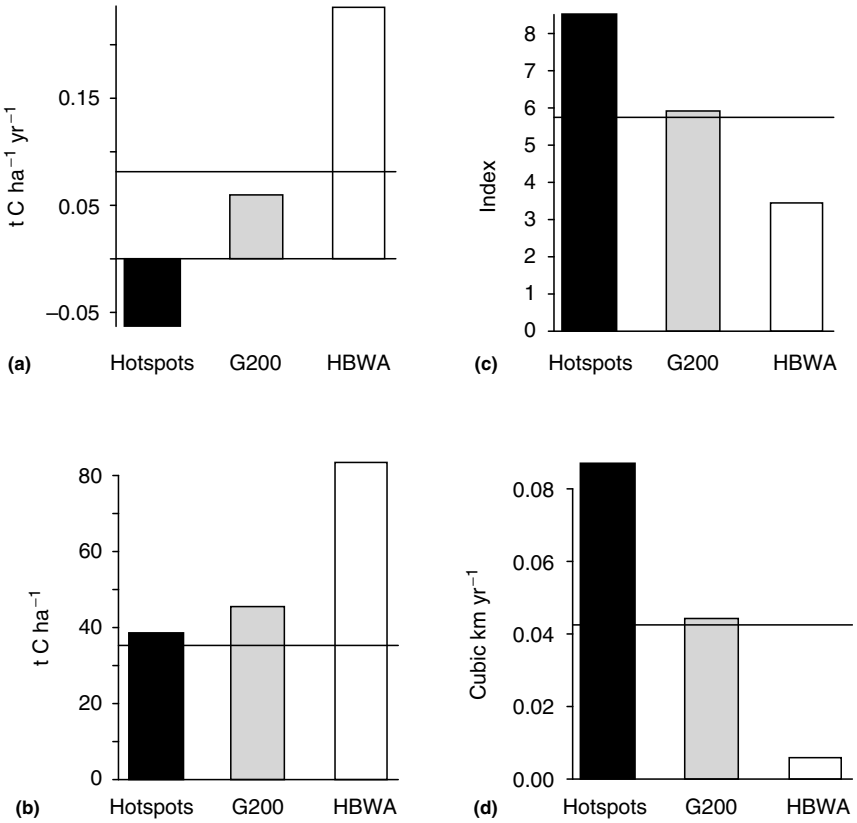


Figure 1 Mean amount of ecosystem services per unit area provided by three global conservation priority schemes. (a) Carbon sequestration. (b) Carbon storage. (c) Grassland production of livestock. (d) Water provision. Horizontal lines represent the global mean for each service. Three global biodiversity priority schemes are biodiversity hotspots, G200, and HBWA. Reproduced from Naidoo R, Balmford A, Costanza R, *et al.* (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9495, with permission from PNAS.

services described above (Naidoo *et al.*, 2008) also appears at finer scales. For example, in California's Central Coast ecoregion, biodiversity was mapped with a variety of ecosystem services: pollination, carbon storage, recreation, water storage, flood control, and forage production (Chan *et al.*, 2007). By applying the standard site-selection algorithms in MARXAN (Ball and Possingham, 2000), Chan *et al.* (2007) identified a network of protected areas capturing both biodiversity targets and ecosystem services. This network required 10–20% increase in the area of protected land beyond what was needed for biodiversity alone. Thus, just as seen at the global scale, individual sites can be identified that are relatively high in both biodiversity and ecosystem service value, regardless of a poor overall correlation between the two. Other regional and local studies across geographies have shown similar findings (e.g., Anderson *et al.*, 2009; Egoh *et al.*, 2009; Nelson *et al.*, 2008; and Bai *et al.*, 2011).

Indeed, in some cases a conservation policy directed toward only one target can actually reduce the ability to attain the other goal. For instance, in the Willamette Basin, Nelson *et al.* (2008) showed that conservation payment policies aimed at increasing the provision of carbon sequestration do not necessarily increase species conservation and that highly targeted policies do not necessarily perform as well as other general conservation payment policies. For example, at-risk vertebrate species were maximized when landowners accepted

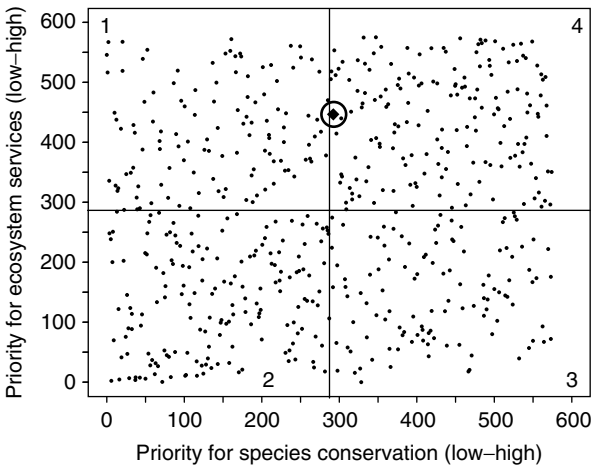


Figure 2 Relationship between ecosystem service importance and conservation importance. Each axis is the rank order (low to high) of 574 terrestrial ecoregions in terms of per-area carbon storage (*y*-axis) or area-corrected number of endemic species (*x*-axis; calculated as Number of endemic species per area$^{0.25}$). Lines indicate median values for each variable. Diamond in quadrant 4 represents the California Central Coast ecoregion. Reproduced from Naidoo R, Balmford A, Costanza R, *et al.* (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9495, with permission from PNAS.

payments to restore rare natural habitats, including oak savanna, prairie, and emergent marsh. Carbon sequestration, however, was maximized when landowners accepted payments to restore or conserve forests, including old growth, mixed, and riparian forest. Indeed, maximizing forest cover did benefit some species (e.g., the spotted owl); however, it provided little benefit for most of the 37 rare species analyzed. Thus, if conservation programs that pay for ecosystem services are not designed carefully, they may yield minimal gains in services of interest and may well result in harm to other services or biodiversity conservation.

Importance of Considering Cost-Effectiveness and Cost-Benefits

Unfortunately, conservation plans cannot be implemented for free. By ignoring the cost side of conservation planning, we are missing the opportunity to more efficiently implement conservation objectives given the reality of limited global conservation funds (Naidoo *et al.*, 2006). Just as biodiversity is not evenly distributed over landscapes (e.g., Orme *et al.*, 2005), the same is true for economic costs and this spatial variability should be considered in conservation planning (e.g., Ando *et al.*, 1998; Polasky *et al.*, 2008). Ignoring the spatial variation in costs has led to conservation plans that are at best far from optimal and at worst economically indefensible.

Conservation costs can be thought of in several major categories: acquisition costs, management costs, transaction costs, damage costs, and opportunity costs. A high cost of one type does not necessarily indicate a high cost of other types. For example, a forested parcel adjacent to a road may have low management costs, due to its accessibility, but high acquisition costs because of its proximity to urban development. Such parcels, because of their proximity to urban areas, are typically prime targets for land-use conversion, in this case forest to development, because the opportunity for keeping the land forested are much too high, as represented by the high value of using the land for development instead of forestry. From a social perspective, it is important to include opportunity costs, or the lost opportunities, to account for the full cost of conservation planning.

Most of the systematic conservation planning literature has focused on efficient representation of biodiversity in reserves (see Margules and Sarkar, 2006). For example, a retrospective analysis of a $2.8 billion land investment for conservation lands and easements between 1990 and 2006 in California, USA, showed that four times as many species and five times as much land could have been conserved if a return-on-investment algorithm had been used (Underwood *et al.*, 2008). In terms of how conservation benefit per dollar spent could affect global funding priorities, priority funding for CI's terrestrial biodiversity hotspots based solely on different biodiversity values (i.e., single taxonomic groups and endemics) did not overlap well, whereas factoring in economic costs of working in different regions revealed plans with much higher concordance (Bode *et al.*, 2008). However, it is worth noting that, typically, conservation funds cannot be spent just anywhere on anything.

Another study of global priority areas sought to maximize mammalian species per area and identified 11% of the Earth's surface for conservation attention (Carwardine *et al.*, 2008). However, these areas directly overlapped high value areas of agricultural production and dense human settlements – a classic example of conflict between people and conservation goals. Yet, a reanalysis of the distribution of the world's mammals, with a focus this time on cost-effectiveness, revealed a different picture. When planning algorithms sought to minimize conflicts between protected areas and crop production, an ecosystem service, the area of conflict was halved, whereas the total area needed to protect the world's mammals increased by less than 2% (Carwardine *et al.*, 2008).

Although conservation-planning efforts targeting protected areas are important for conserving many rare species, this approach sometimes ignores the potential conservation benefits of less "pristine" lands (e.g., agricultural or developed lands). There are relatively few examples of systematic conservation planning that explicitly consider the entire range of land cover and land management practices in a region, yet many species can coexist with some level of human activity (Pennington and Blair, 2011). The broader conservation question, beyond the question on where are the best places to locate reserves, is whether conservation objectives can be met on a landscape that includes both working lands and protected lands. Land-use decisions on working lands are based primarily on economic criteria, whether it is the local people using land to make a living or the corporations using land to maximize profits. Although land-use decisions based solely on economic returns are often detrimental to biodiversity, securing some economic return from land need not be mutually exclusive with biodiversity conservation. By thinking carefully about the pattern, extent, and intensity of human activities across the landscape, it may be possible to achieve important biodiversity conservation objectives while also generating reasonable economic returns.

For example, Polasky *et al.* (2008) developed a spatially explicit landscape-level model to analyze the biological and economic consequences of alternative land-use patterns for the Willamette Basin in Oregon, USA. In that study, researchers combined biological and economic model results, total number of species, and value of marketed commodities (including agriculture and timber output, both are provisioning services but for this study are not explicitly considered as so, and urban development), respectively, to search for efficient basin-wide land-use patterns using optimization methods from operations research. An efficient land-use pattern is one that generates the maximum number of species conserved for a given economic return (and vice versa). By maximizing the total number of species over the entire range of possible economic returns, one can trace out an efficiency frontier for the landscape (**Figure 3**). The efficiency frontier illustrates what can be achieved in terms of biological and economic objectives by carefully arranging the spatial allocation of activities across the landscape. In this case, Polasky *et al.* (2008) found land-use patterns that sustain an expected 249 species, 97% of the total number of species found for the landscape, and $25.4 billion in economic returns, 92% of the maximum economic value (point D, **Figure 3**). The efficiency frontier also demonstrates the degree of inefficiency of

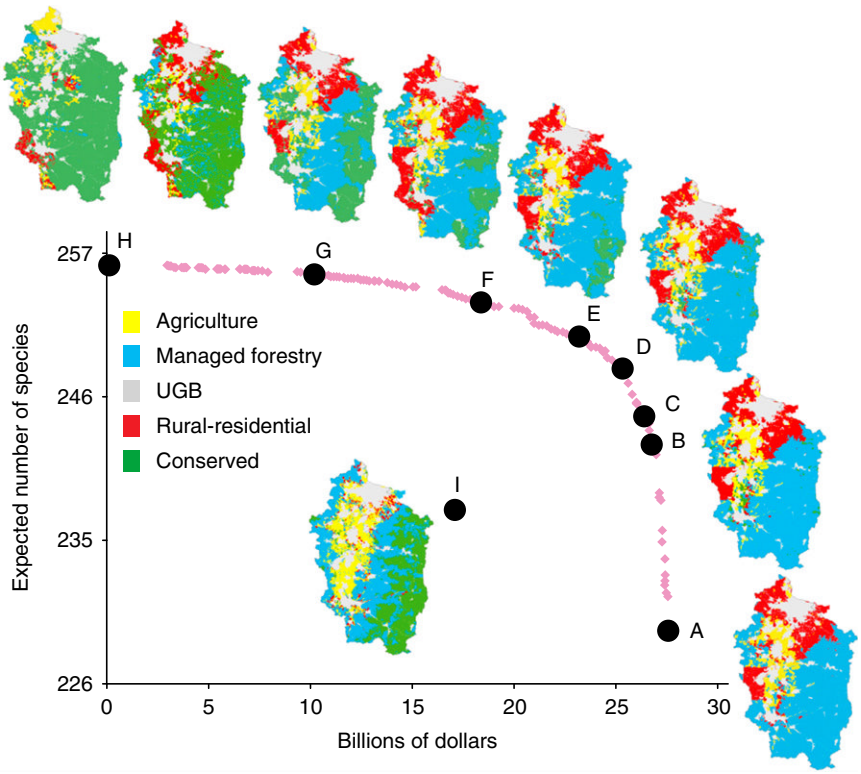


Figure 3 Land-use patterns associated with specific points along the efficiency frontier and the current landscape. Each land-use pattern shown outside of the efficiency frontier corresponds to a lettered point on the frontier. The current land-use pattern is also shown (I). Compared with the current landscape, points on the efficiency frontier have less agriculture and more residential use. There is a shift from predominantly managed forest toward conservation land as the biological objective, number of species, is emphasized more relative to the economic objective, returns from agriculture, forestry, and development. *Abbreviation:* UGB, urban growth boundary. Reproduced from Polasky S, Nelson E, Camm J, *et al.* (2008) Where to put things? Spatial land management to sustain biodiversity and economic returns. *Biological Conservation* 141: 1505–1524, with permission from Elsevier.

other land-use patterns not on the frontier. For example, the business-as-usual land-use pattern for the basin only sustains 238.6 species and generates $17.1 billion in economic returns (point I, **Figure 3**). These results indicate how proper attention to spatial management can limit trade-offs between biodiversity conservation and economic returns.

Incorporating Biodiversity, Ecosystem Services, and Economic Considerations

In its ideal form, a conservation planning effort should consider biodiversity as well as multiple ecosystem services, and estimate monetary values of both benefits and costs to ease comparisons and decisions. Including biodiversity and multiple benefits is important because it allows for the explicit evaluation of potential tradeoffs and broadens the planning conversation by considering conservation goals for biodiversity and for human well-being. Monetary valuation is useful because it can help to assess tradeoffs by making the costs of environmentally damaging activities more apparent. Estimating the monetary value of ecosystem services has been a focus of much recent research. For provisioning services where established markets exist, such as agriculture and forestry, determining economic value is relatively straightforward (e.g., Polasky *et al.*, 2008; Nelson *et al.*, 2009). For those goods and services where no current market exists, such as water filtration, economic valuation can be much more difficult (Freeman, 2003).

The studies described so far in this article strongly incorporate some of these components. But only a very few studies explicitly combine them all into full assessments. For example, although there have been many studies that quantify the opportunity cost of providing a single ecosystem service (e.g., Kindermann *et al.*, 2008 for carbon sequestration; Swallow *et al.*, 2009 for sediment delivery) and similarly for biodiversity conservation (e.g., Naidoo and Ricketts, 2006; Polasky *et al.*, 2008), there have been relatively few studies on the joint provision and value of multiple outcomes (multiple ecosystem services, biodiversity, and returns to landowners) from ecosystems. Several studies have characterized the spatial overlap among ecosystem service provision and biodiversity conservation (e.g., Chan *et al.*, 2006; Egoh *et al.*, 2008; Naidoo *et al.*, 2008; and Nelson *et al.*, 2009), but these studies have not analyzed how alternative land-use or land-management decisions affect the provision and value of ecosystem services or biodiversity conservation (Polasky *et al.*, 2011). Integrating biodiversity, ecosystem services, and economic costs into systematic conservation planning is a new area of research that seeks to broaden the cost considerations beyond those of conservation donors to society and policy makers at large.

In this final section, some of the recent studies that have moved conservation toward doing the "full monty" are described.

Naidoo and Ricketts (2006) served as a landmark example of incorporating multiple ecosystem service values into the traditional context of conservation site selection. They mapped, the monetary value of five ecosystem services (sustainable bushmeat harvest, sustainable timber harvest, bioprospecting, existence value, and carbon storage) in the Mbaracyau Forest

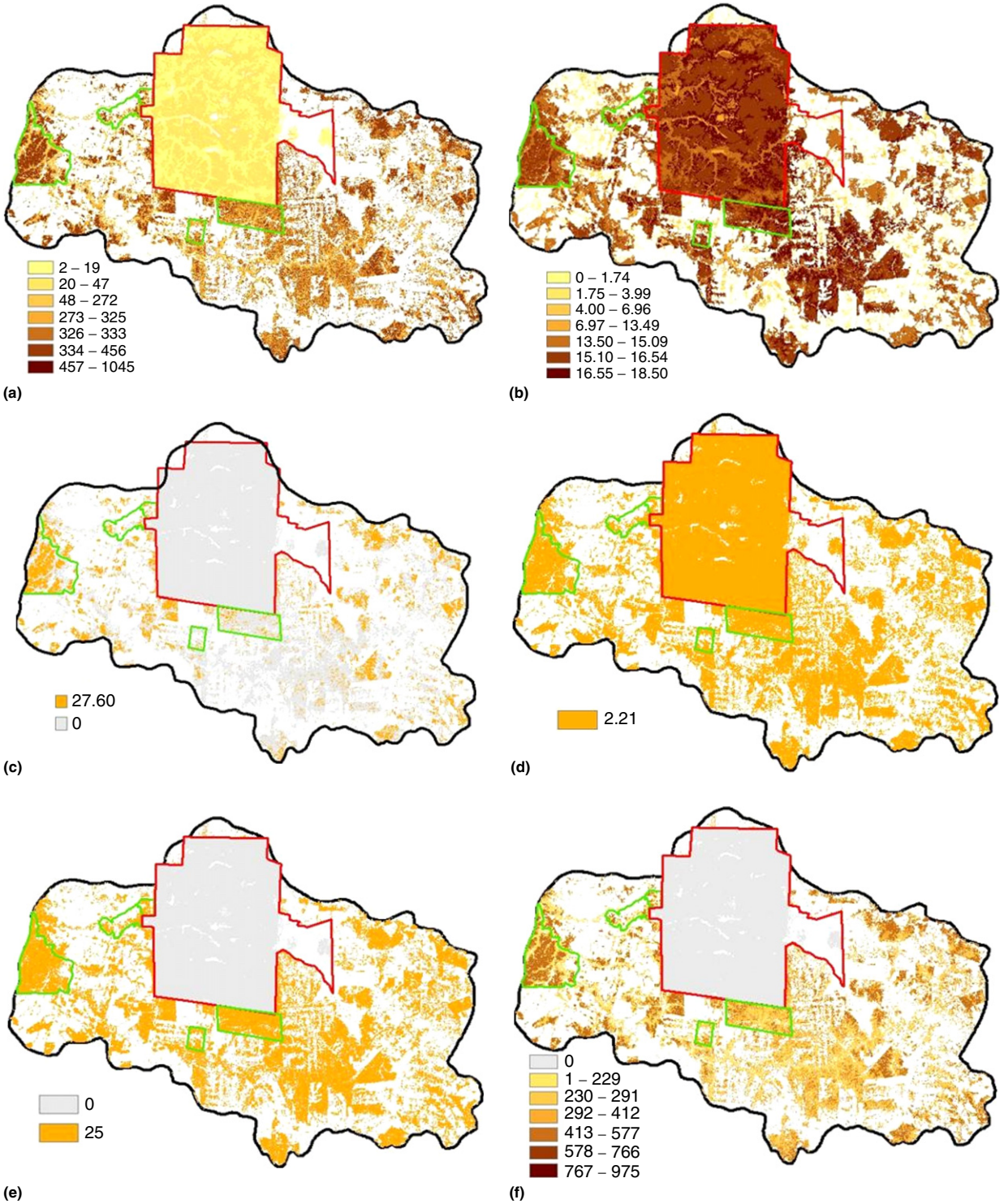


Figure 4 Economic values (net present values in US$ per hectare) in the Mbaracyau Forest Biosphere Reserve. (a) Sum of all 5 services. (b) Sustainable bushmeat harvest. (c) Sustainable timber harvest. (d) Bioprospecting. (e) Existence value. (f) Carbon storage. Reproduced from Naidoo R and Ricketts TH (2006) Mapping the economic costs and benefits of conservation. *Public Library of Science Biology* 4: e360, with permission from Public Library of Science.

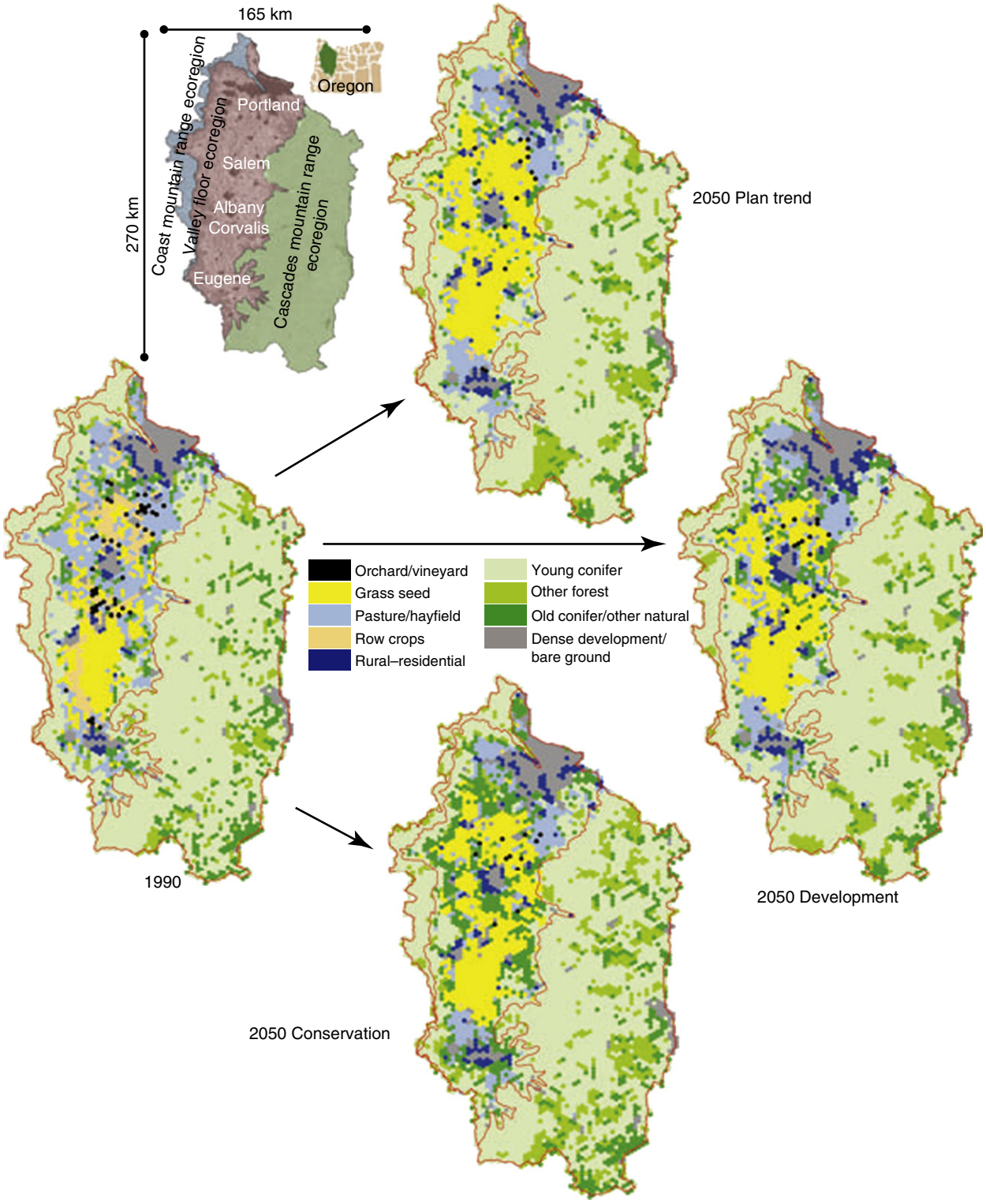


Figure 5 Maps of the Willamette Basin and the land-use pattern for 1990 and under the three land-use change scenarios for 2050. Reproduced from Nelson E, Mendoza G, Regetz J, *et al.* (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment* 7: 4–11, with permission from ESA.

Biosphere Reserve in Paraguay (see **Figure 4**). By also estimating opportunity costs, they were able to show areas where conservation values would more than cover the costs of conservation and other areas where the converse is true. Naidoo and Ricketts (2006) used these maps to evaluate three alternative locations for a proposed corridor linking two protected areas, and they found that one corridor provides much higher benefits relative to costs than the other two. This is an example of how these approaches of mapping and valuing ecosystem services can help direct conservation efforts.

Moving toward the "full monty" approach to systematic conservation planning, Nelson *et al.* (2009) compared biodiversity conservation, multiple ecosystem services, and economic outcomes under three alternative land-use trajectories for the period 1990–2050 for the Willamette Basin (**Figure 5**). They found that all biodiversity and ecosystem service measures were highest under a conservation scenario, but economic returns to landowners were higher under more development oriented scenarios. However, including payments for carbon sequestration alleviated this trade-off (**Figure 6**).

For the state of Minnesota, USA, Polasky *et al.* (2011) modeled biodiversity conservation, the provision of multiple ecosystem services, including carbon sequestration, water quality, agricultural production, and urban development. They compared historical land-use change during 1992–2001 in addition to alternative scenarios of future land-use change: no urban expansion; no agriculture expansion; agricultural expansion; forest expansion; and a conservation scenario where riparian areas and marginal agricultural lands are restored to presettlement vegetation. Not surprisingly, they did not find a single land-use scenario that maximized the levels of all ecosystem services as well as habitat for biodiversity. A conservation scenario was best for the reduction of phosphorus exported into the Mississippi River, but a no agricultural expansion scenario was best for carbon sequestration. The best scenario for habitat provision depends on the group of species being considered.

Like Nelson *et al.* (2009), they found a lack of concordance in the ranking of scenarios between net social benefits (nonmarket and market values, including carbon sequestration, water quality, agriculture and forestry production, and urban development) and the market value of returns to landowners (agriculture and forestry production, and urban development), which suggests that there is a large role for public policy in land-use decisions. For example, the agriculture expansion scenario, where all highly productive lands were put into agriculture, delivers the lowest total economic value across scenarios despite the fact that it delivers the highest financial returns to landowners in terms of crop sales.

Another example of simultaneously assessing multiple ecosystem services uses a bioeconomic model of a coral reef–mangrove–seagrass ecosystem to investigate the contribution of mangroves toward providing storm protection and as a nursery for coral reef fish, and also incorporate market values of converting mangroves into aquaculture (Sanchirico and Springborn, 2011). The authors' results show that considering multiple services can mitigate the level of payments for ecosystem services needed to ensure the delivery of other services, in this example fishing, thus, highlighting the importance of including multiple services in order to minimize financial payments. By including multiple ecosystem services, each stakeholder group that benefits from an ecosystem process with multiple services is likely to have lower costs relative to a situation where fewer services are considered, everything else being equal. Even in the case of degraded coral reef–mangrove–seagrass ecosystems, services thereof exceeded the market returns from aquaculture, highlighting the importance of expanding conservation considerations to include marginal and degraded systems as well as the pristine.

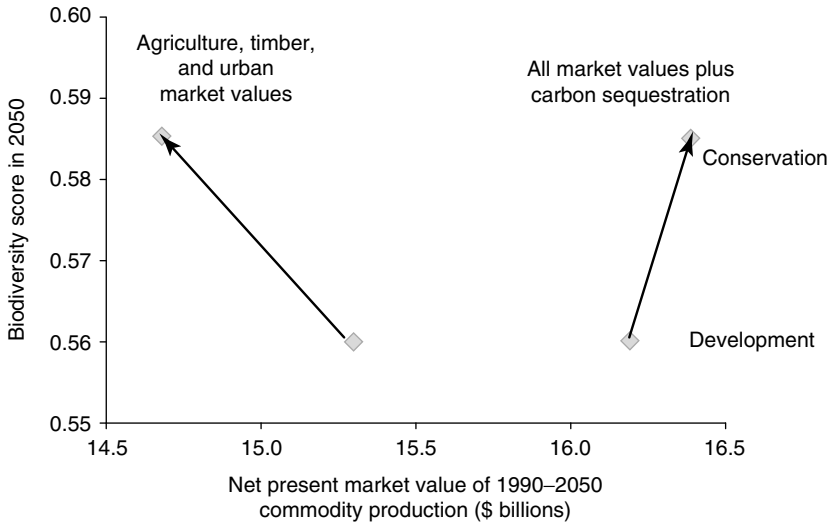


Figure 6 Tradeoffs between market values of commodity production and biodiversity conservation on the landscape between 1990 and 2050, excluding (circles) and including (triangles) the market value of carbon sequestration. The *x*-axis measures the total discounted value of commodities, whereas the *y*-axis measures the biodiversity score. Reproduced from Nelson E, Mendoza G, Regetz J, *et al.* (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment* 7: 4–11, with permission from ESA.

Conclusions

Even in the well-developed field of systematic conservation planning, spatially explicit models of ecosystem services can broaden the information available to planners and improve decision making. As shown above, mapping and quantifying ecosystem services and biodiversity conservation in a spatially explicit manner, and analyzing trade-offs between them, can help to make natural resource decisions more effective, efficient, and defensible. Integrated landscape-level analysis that tracks change across a number of dimensions of ecosystem services and biodiversity conservation is still a relatively young discipline. Application of such models, however, can generate information for decision makers showing the consequences of choices for a range of important ecosystem services and biodiversity conservation objectives (Groves, 2003). In principle, putting this information in the hands of decision makers should lead to improved landscape planning and management (Daily *et al.*, 2009).

Remaining Challenges

Many challenges remain before reaching this Promised Land. These include improving and validating the component models for ecosystem services (e.g., De Groot *et al.*, 2010; Balmford *et al.*, 2011). Researchers applying these models should also conduct sensitivity analyses to determine whether errors in the biodiversity, ecosystem services, or economic data could affect conservation outcomes. For example, Bode *et al.* (2008) found that conservation outcomes were most sensitive to uncertainty in the land cost data, followed by the habitat loss rates, and finally the biodiversity data. Indeed, these errors are likely largest in the cost or economic data, especially for nonmarket valuation. Indeed, Bode *et al.* (2008) found that these errors could result in inefficient funding strategies causing 20 times as many extinctions as errors of the same proportional size in the biodiversity data, and 6 times as many extinctions as errors of the same proportional size in the habitat loss rates.

Translating from biophysical units to monetary value units is problematic for biodiversity targets and some types of ecosystem services (e.g., Freeman, 2003). For example, trying to estimate the values of cultural and spiritual values is difficult and controversial. For some ecosystem services, estimating monetary values using market prices or applying nonmarket valuation techniques may be relatively straightforward (e.g., provisioning services (Naidoo and Ricketts, 2006; Polasky *et al.*, 2008, 2011; Nelson *et al.*, 2008, 2009) and crop pollination, which is an input to a priced commodity (Ricketts *et al.*, 2004)).

Even with our growing knowledge of biophysical systems, understanding the provision of ecosystem services also requires detailed understanding of what is of value to people. It is important to recognize that estimates of value used to identify spatial priorities, to evaluate trade-offs, and for most other results reviewed here, depend on the choice of ecosystem services to include. Moreover, evaluating multiple services can elicit unexpected outcomes and trade-offs (e.g., Polasky *et al.*, 2011).

Another important set of issues relates to land-use change and dynamics. Most conservation planning research has been static. In the examples presented above, the authors analyzed "steady-state" consequences of a landscape in terms of biophysical and socioeconomic objectives. However, the on-the-ground reality is that there is an existing landscape and changes will occur over time. For example, if the current land use is soybeans but the desired land use is a managed forest, it will take several years for trees to grow to maturity for harvest, thus resulting in a significant delay between land-use change and the actual start of forestry activities or other natural capital benefits. Also, through this period of forest growth there could be unforeseen events, such as fire or pest outbreaks or long-term impacts caused by climate change that could lead to unexpected transitions. In turn, species or ecological processes would take time to adjust. Furthermore, human population changes and shifts in markets or societal values will result in changes in the benefit of different land uses over time. All of these will affect the relative costs and benefits of investing in conservation and mean that at any given time decisions on what and where to conserve are changing.

Stochastic Dynamic Programming (SDP) is a framework for modeling optimization problems that involve uncertainty (Clark and Mangel, 2000). In conservation planning this approach has been used to find the optimal allocation of investment at each time step given the current state of the system and possible events in the future. Currently, applying SDP to problems over large areas is computationally intractable. However, there have been a few articles that analyze an optimal sequence of conservation decisions using SDP, such as schedule investments over time (e.g., Costello and Polasky (2004); Wilson *et al.*, 2006). For example, Wilson *et al.* (2006) applied SDP using information on how investment returns change through time and found that simple ranking schemes yield fewer benefits per dollar than a return-on-investment approach, suggesting that conservation actions are not independent of each other. The protection of a particular parcel of land because it is biologically valuable and also threatened by conversion may simply displace the threat to nearby areas.

Costello and Polasky (2004) showed that conservation budgets available upfront yield significantly greater biodiversity protection and highlight the importance of the timing of protected area selection. In addition, the purchase of properties for a new protected area may drive land prices up due to increased demand for land for other purposes. This can have perverse effects on local communities and also make it more expensive to establish additional protected areas. This shows how investment decisions by conservation organizations themselves change the state of play and resulting land market prices. There are approaches for predicting and incorporating negative feedbacks in dynamic conservation planning situations. Using the knowledge of the dynamics of the local market, Armsworth *et al.* (2006) showed that whenever a parcel of land is purchased, the relative price of other parcels of land is altered. Similarly, the likelihood of success can be updated in a similar manner. Future work in conservation planning should consider these dynamics and transitions.

Finally, our understanding of these techniques to model ecosystem services and trade-offs far exceeds our ability to

apply them effectively to pragmatic conservation problems (Knight *et al.*, 2006). Improving our understanding of meaningful ways to embed this approach into larger policy decisions is paramount. Importantly, what is attractive about this approach in terms of planning decisions is that it does not require agreement on (1) the motivations underlying the tradeoffs that individual people are willing to make (i.e., the reasons that people want to protect ecosystems) or (2) what considerations beyond an assessment of trade-offs should enter into public policy decisions (Polasky and Segerson, 2009). By recasting the conservation planning discussion to focus on the relevance and assessment of trade-offs, it increases the likelihood that all participants will be able to find greater common ground for collaboration, even if they hold very disparate views on other issues. In turn, this should allow greater progress toward designing and implementing policies or plans designed to protect not only the biodiversity but also the ecosystems and the services they provide.

Appendix

List of Courses

1. Conservation Biology
2. Topics in Ecology
3. Environmental Economics
4. Resource Economics
5. Environmental Studies
6. Sustainability Studies

See also: Countryside Biogeography. Modeling Biodiversity Dynamics in Countryside and Native Habitats. Modeling Terrestrial Ecosystem Services

References

Abell R, Thieme ML, Revenga C, *et al.* (2008) Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience* 58: 403–414.

Anderson BJ, Armsworth PR, Eigenbrod F, *et al.* (2009) Spatial covariance between biodiversity and other ecosystem service priorities. *Journal of Applied Ecology* 46: 888–896.

Ando A, Camm J, Polasky S, and Solow A (1998) Species distributions, land values, and efficient conservation. *Science* 279: 2126.

Antle JM and Stoorvogel JJ (2006) Predicting the supply of ecosystem services from agriculture. *American Journal of Agricultural Economics* 88: 1174.

Armsworth PR, Daily GC, Kareiva P, and Sanchirico JN (2006) Land market feedbacks can undermine biodiversity conservation. *Proceedings of the National Academy of Sciences of the United States of America* 103: 5403–5408.

Bai Y, Zhuang C, Ouyang Z, Zheng H, and Jiang B (2011) Spatial characteristics between biodiversity and ecosystem services in a human-dominated watershed. *Ecological Complexity* 8: 177–183.

Ball IR and Possingham HP (2000) Marxan (v. 1.8.6): Marine Reserve Design Using Spatially Explicit Annealing. User Manual: http://www.uq.edu.au/marxan.

Balmford A, Fisher B, Green RE, *et al.* (2011) Bringing ecosystem services into the real world: An operational framework for assessing the economic consequences of losing wild nature. *Environmental and Resource Economics* 48: 161–175.

Balvanera P, Daily GC, Ehrlich PR, *et al.* (2001) Conserving biodiversity and ecosystem services. *Science* 291: 2047.

Bode M, Wilson KA, Brooks TM, *et al.* (2008) Cost-effective global conservation spending is robust to taxonomic group. *Proceedings of the National Academy of Sciences of the United States of America* 105: 6498.

Boody G, Vondracek B, Andow DA, *et al.* (2005) Multifunctional agriculture in the United States. *BioScience* 55: 27–38.

Brooks TM, Mittermeier RA, da Fonseca GAB, *et al.* (2006) Global biodiversity conservation priorities. *Science* 313: 58.

Carwardine J, Wilson KA, Watts M, Etter A, Klein CJ, and Possingham HP (2008) Avoiding costly conservation mistakes: The importance of defining actions and costs in spatial priority setting. *Public Library of Science One* 3: e2586.

Chan K, Pringle RM, Ranganathan J, *et al.* (2007) When agendas collide: Human welfare and biological conservation. *Conservation Biology* 21: 59–68.

Chan KMA, Shaw MR, Cameron DR, Underwood EC, and Daily GC (2006) Conservation planning for ecosystem services. *Public Library of Science Biology* 4(11): e379.http://dx.doi.org/10.1371/journal.pbio.0040379.

Cincotta RP, Wisnewski J, and Engelman R (2000) Human population in the biodiversity hotspots. *Nature* 404: 990–992.

Clark CW and Mangel M (2000) *Dynamic State Variable Models in Ecology.* New York, NY: Oxford University Press.

Coiner C, Wu JJ, and Polasky S (2001) Economic and environmental implications of alternative landscape designs in the Walnut Creek Watershed of Iowa. *Ecological Economics* 38: 119–139.

Costello C and Polasky S (2004) Dynamic reserve site selection. *Resource and Energy Economics* 26: 157–174.

Cowling R and Pressey R (2003) Introduction to systematic conservation planning in the Cape Floristic Region. *Biological Conservation* 112: 1–13.

Cowling RM, Pressey RL, Lombard AT, Desmet PG, and Ellis. AG (1999) From representation to persistence: Requirements for a sustainable system of conservation areas in the species-rich Mediterranean climate deserts of southern Africa. *Diversity and Distributions* 5: 51–71.

Daily GC (ed.) (1997) *Nature's Services: Societal Dependence on Natural Ecosystems.* Washington, DC: Island Press.

Daily GC, Polasky S, Goldstein J, *et al.* (2009) Ecosystem services in decision making: Time to deliver. *Frontiers in Ecology and the Environment* 7: 21–28.

De Groot RS, Alkemade R, Braat L, Hein L, and Willemen L (2010) Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. *Ecological Complexity* 7: 260–272.

Egoh B, Reyers B, Rouget M, Bode M, and Richardson DM (2009) Spatial congruence between biodiversity and ecosystem services in South Africa. *Biological Conservation* 142: 553–562.

Egoh B, Reyers B, Rouget M, Richardson DM, Le Maitre DC, and van Jaarsveld AS (2008) Mapping ecosystem services for planning and management. *Agriculture, Ecosystems & Environment* 127: 135–140.

Ferraro PJ and Pattanayak S (2006) Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *Public Library of Science Biology* 4: e105.

Freeman AM (2003) *The Measurement of Environmental and Resource Values: Theory and Methods.* Washington, DC: RFF Press.

Gedan KB, Silliman B, and Bertness M (2009) Centuries of human-driven change in salt marsh ecosystems. *Annual Review of Marine Science* 1: 117–141.

Groves C (2003) *Drafting a Conservation Blueprint: A Practitioner's Guide to Planning for Biodiversity.* Washington, DC: Island Press.

Kareiva P, Tallis H, Ricketts TH, Daily GC, and Polasky S (2011) *Natural Capital: Theory and Practice of Mapping Ecosystem Services.* Oxford: Oxford University Press.

Kindermann G, Obersteiner M, Sohngen B, *et al.* (2008) Global cost estimates of reducing carbon emissions through avoided deforestation. *Proceedings of the National Academy of Sciences of United States of America* 105: 10302.

Knight AT, Cowling RM, and Campbell BM (2006) An operational model for implementing conservation action. *Conservation Biology* 20: 408–419.

Larsen FW, Londoño-Murcia MC, and Turner WR (2011) Global priorities for conservation of threatened species, carbon storage, and freshwater services: Scope for synergy? *Conservation Letters* 4: 355–363.

Lewis DJ and Plantinga AJ (2007) Policies for habitat fragmentation: Combining econometrics with GIS-based landscape simulations. *Land Economics* 83: 109–127.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.

Margules CR and Sarkar S (2006) *Systematic Conservation Planning.* Cambridge, UK: Cambridge University Press.

Millennium Ecosystem Assessment (MEA) (2005) *Ecosystems and Human Well-being: Synthesis.* Washington, DC: World Resources Institute Island Press.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Naidoo R, Balmford A, Costanza R, *et al.* (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9495.

Naidoo R, Balmford A, Ferraro PJ, Polasky S, Ricketts TH, and Rouget M (2006) Integrating economic costs into conservation planning. *Trends in Ecology and Evolution* 21: 681–687.

Naidoo R and Ricketts TH (2006) Mapping the economic costs and benefits of conservation. *Public Library of Science Biology* 4: e360.

Nelson E, Mendoza G, Regetz J, *et al.* (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and the Environment* 7: 4–11.

Nelson E, Polasky S, Lewis DJ, *et al.* (2008) Efficiency of incentives to jointly increase carbon sequestration and species conservation on a landscape. *Proceedings of the National Academy of Sciences of the United States of America* 105: 9471.

Noss RF, Carroll C, Vance-Borland K, and Wuerthner G (2002) A multicriteria assessment of the irreplaceability and vulnerability of sites in the Greater Yellowstone Ecosystem. *Conservation Biology* 16: 895–908.

Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on earth. *BioScience* 51: 933–938.

Orme CDL, Davies RG, Burgess M, *et al.* (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature* 436: 1016–1019.

Pennington DN and Blair RB (2011) Habitat selection of breeding riparian birds in an urban environment: Untangling the relative importance of biophysical elements and spatial scale. *Diversity and Distributions* 17: 506–518.

Polasky S, Nelson E, Camm J, *et al.* (2008) Where to put things? Spatial land management to sustain biodiversity and economic returns. *Biological Conservation* 141: 1505–1524.

Polasky S, Nelson E, Pennington D, and Johnson KA (2011) The impact of land-use change on ecosystem services, biodiversity and returns to landowners: A case study in the State of Minnesota. *Environmental and Resource Economics* 51: 219–242.

Polasky S and Segerson K (2009) Integrating ecology and economics in the study of ecosystem services: Some lessons learned. *Annual Review of Resource Economics* 1: 409–434.

Pressey R, Ferrier S, Hager T, Woods C, Tully S, and Weinman K (1996) How well protected are the forests of north-eastern New South Wales? – Analyses of forest environments in relation to tenure, formal protection measures and vulnerability to clearing. *Forestry, Ecology and Management* 85: 311–333.

Pressey RL and Bottrill MC (2008) Opportunism, threats, and the evolution of systematic conservation planning. *Conservation Biology* 22: 1340–1345.

Pullin AS (2002) *Conservation Biology.* Cambridge, UK: Cambridge University Press.

Rands MRW, Adams WM, Bennun L, *et al.* (2010) Biodiversity conservation: Challenges beyond 2010. *Science* 329: 1298–1303.

Ricketts TH, Daily GC, Ehrlich PR, and Michener CD (2004) Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences of the United States of America* 101: 12579.

Rouget M, Cowling RM, Pressey RL, and Richardson DM (2003) Identifying spatial components of ecological and evolutionary processes for regional conservation planning in the Cape Floristic Region, South Africa. *Diversity and Distributions* 9: 191–210.

Sanchirico JN and Springborn M (2011) How to get there from here: Ecological and economic dynamics of ecosystem service provision? *Environmental and Resource Economics* 51: 243–267.

Sanchirico JN and Wilen JE (1999) Bioeconomics of spatial exploitation in a patchy environment. *Journal of Environmental Economics and Management* 37: 129–150.

Sanchirico JN and Wilen JE (2001) A bioeconomic model of marine reserve creation. *Journal of Environmental Economics and Management* 42: 257–276.

Sanchirico JN and Wilen JE (2005) Optimal spatial management of renewable resources: Matching policy scope to ecosystem scale. *Journal of Environmental Economics and Management* 50: 23–46.

Sekhar NU (2005) Integrated coastal zone management in Vietnam: Present potentials and future challenges. *Ocean & Coastal Management* 48: 813–827.

Spalding MD, Fox HE, Allen GR, *et al.* (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* 57: 573–583.

Swallow BM, Kallesoe MF, Iftikhar UA, Van Noordwijk M, and Bracer C (2009) Compensation and rewards for environmental services in the developing world: Framing pan-tropical analysis and comparison. *Ecology and Society* 14. http://www.ecologyandsociety.org/vol14/iss2/art26/

The Nature Conservancy (2011) *Conservation by Design: A Strategic Framework for Mission Success.* Washington, DC: The Nature Conservancy.

Turner WR, Brandon K, Brooks TM, Costanza R, Da Fonseca GAB, and Portela R (2007) Global conservation of biodiversity and ecosystem services. *BioScience* 57: 868–873.

Underwood EC, Shaw MR, Wilson KA, *et al.* (2008) Protecting biodiversity when money matters: Maximizing return on investment. *Public Library of Science One* 3: e1515.

Wilson KA, McBride MR, Bode M, and Possingham HP (2006) Prioritizing global conservation efforts. *Nature* 440: 337–340.

Property Rights and Biodiversity

Susan S Hanna, State University Corvallis, OR, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Economic rent Payment to resources in excess of costs.
Ecosystem services Contributions of ecosystem components through genetics, information, and reproduction.
Externality An effect – either a cost or a benefit – external to the generating activity.
Free rider One who enjoys the benefit of goods or service without paying the cost.
Institutions Organizational structures that shape human interactions.
Prisoner's dilemma Uncertainty about others' behavior leading to choices that are individually rational but collectively irrational.
Public goods Goods and services which if supplied to one person can be supplied to additional people at no extra cost.
Time horizon Time period over which future benefits and costs are considered.
Tragedy of the commons Overuse of resources resulting from open access and the incomplete accounting for costs.
Transactions costs Costs other than price associated with the trade of environmental goods and services: information, negotiation, decisions, and enforcement.

The Context of Property Rights

Property rights define the conditions that guide and control human use of the natural environment. They establish the terms under which people use and sustain the capacity of the environment to generate a continuing flow of goods and services. Property rights are a means by which people interact with their environment and control their behavior toward one another. They embody the expectations people have about natural resources and influence the way people make resource use decisions. Property rights link humans to each other and to natural systems through these expectations and decisions (Bromley, 1991; North, 1992; Schmid, 1995).

Scope

Property rights to natural resources control both use and conservation. Their scope may be individual species, genetic information contained in species, or areas of land and water in which species live. The scope of property rights also includes different types of use. At present, most systems of property rights are designed for direct uses, such as catching fish for food, but they may also include indirect uses, such as the right of the public to enjoy populations of whales or to protect endangered species like bald eagles. Property rights are almost never defined for unused species or for communities of species. The idea of using property rights for the protection of not only single species but also biodiversity and its genetic foundation is a newer and broader application of their accustomed use.

Values

If resources are used, they have value. The values of ecosystems that derive from direct uses are most easily recognized – for example, the harvest of shellfish and seaweed from intertidal marine ecosystems. Their value is expressed in markets or in the subsistence they provide. But indirect use also reflects values that, although not well quantified, are important and well articulated as economic concepts. Resources have option values for the future as potential goods and services – for example, the future food or pharmaceutical uses of marine species. These potential uses are valued as insurance against the unknown. Resources also have quasi-option values in future uses that may exist through preservation or be lost by irreversible actions, such as the localized genetic information in subraces of salmon that are lost as habitat is destroyed. Finally, resources have existence values in their aesthetic or spiritual attributes, such as the enjoyment of watching an osprey catch a fish or the satisfaction of just knowing that ospreys are nesting nearby. Existence values can also be thought of as bequest values, in that the unused goods and services are available to be passed on to future generations.

An important contributor to the option value of ecosystems is ecological resilience. Resilience is the elasticity of the ecosystem; the ability of an ecosystem to absorb perturbations and continue its essential functions (Perrings *et al.*, 1995). Although much remains unknown about the necessary conditions for resilience, what is known is that there are keystone species that require protection and conservation to enhance both species diversity and critical ecosystem functions. The option values of direct uses of ecosystems are relatively easy to understand and measure, because they are linked to tangible goods like seafood and timber that are traded through markets. Considerable uncertainty exists, however, about the value of services or goods that are not currently used. Measuring and accounting for the option values of ecosystem services and unknown goods is difficult because there is no market in which people can express their willingness to pay for them or be compensated for their loss (Teeb, 2010).

People and Biodiversity

Biodiversity can be viewed in a number of dimensions, including the diversity of species, the diversity of genetic material, and the diversity of functional roles in the ecosystem.

Many of the benefits of biodiversity are public goods, in which value accrues to everyone. The issue of people and biodiversity is one of managing the direct and indirect effects of activities in ecosystems in order to maintain the diversity of genetic, species, and functional components. One challenge is how to provide appropriate incentives so that people find it in their interest to promote and maintain the public good of biodiversity. A second challenge is how the rights, rules, and responsibilities that constrain resource use can be expanded from single species or physical areas to multiple species that may be distributed over wide areas. A third challenge is how to make the transition from traditional single-species commodity production types of use to new types of use that accommodate the protection of species diversity through the maintenance of ecosystem services.

Ownership

The idea of ownership is well understood for goods produced by ecosystems. Ownership is better defined for land-based resources than for marine resources but for both, the full range of property rights described in **Table 1** applies. Property rights to goods from land ecosystems are usually clearly defined and range from private to public. Property rights to goods from marine ecosystems are similarly mixed and often still include open access. Property rights to goods from marine ecosystems vary from private/community/state in nearshore areas, to state property in offshore zones, to open access in international waters.

The idea of owning ecosystem services, in contrast, is less familiar but is gaining in experimental application. At present, ecosystem services are either considered the property of the public or of the owner of the area in which they exist. Ecosystem services are public goods, but in neither land-based nor marine ecosystems are rights to services specifically defined. One particular ecosystem service – the genetic information embodied in species – is particularly important to the question of biodiversity. Property rights to the informational properties of resources are in only the very early stages of development and application (Swanson, 2003).

Forms of Property Rights

Property rights take many forms. Differences in form derive from the scope of the rights, the type of rights owner, and the privileges and responsibilities of the rights owner. The most common categories are private, public, state, and common (**Table 1**). The types of property rights are ordered loosely along a spectrum, where the owner ranges from an individual person to no one (Hanna *et al.*, 1996).

Open Access

Open access, *res nullius*, is the absence of property rights and has no ownership assigned. It is open to all, with resources becoming owned only at the point of capture. The absence of rights is often, but mistakenly, referred to as "common property." The "tragedy of the commons" metaphor is a description of the outcome of open access. It describes the use of a common-pool resource from which it is too costly to exclude people from use. In the tragedy of the commons, individuals make choices about resource use based on their own private costs and benefits. And while the benefits of their actions accrue to them alone, the costs, in terms of the effect on the resource, are social costs spread among all users. Eventually and inevitably without some control over access, the failure to account for social costs locks people into behavior that leads to resource degradation and overuse (Hardin, 1968; Ciriacy-Wantrup and Bishop, 1975).

Common Property

Although the "tragedy of the commons" is a metaphor that is widely used, it does not really describe common property. Common property, *res communes*, is owned by an identified group of people who have the right to exclude nonowners, the duty to participate in decisions about use, and the responsibility to act as resource stewards. For example, nearshore fishing territories may be owned and managed by the residents of the adjoining coastal community. Community members as owners decide how many people have rights to use the resource, what kind of rights these may be, and what objectives they have for resource productivity.

State Property

State property, *res publicae*, is owned by citizens who assign management authority and stewardship responsibility to an agency of the state. For example, U.S. fish and wildlife resources are the property of the citizens of individual states or of the country as a whole. The citizens designate the fish and wildlife agencies of individual states and the federal government to manage these resources in their name. The agencies may grant rights of use to individual users or communities of users, but the resources remain owned by the citizens at large.

Table 1 A classification of pure forms of property rights

Property rights	*Rights holder*	*Privileges*	*Responsibilities*
Open access	None	Capture	None
State property	Citizens	Designate management authority	Stewardship
Common property	Collective	Exclude nonmembers	Participation, maintenance
Private property	Individual	Socially acceptable uses; control of access; transfer	Compliance with laws

Private Property

Private property, *res privatae*, assigns ownership to named individuals including legal individuals, such as corporations, guaranteeing to those owners a bundle of rights about access and use. Although individuals have the greatest autonomy under private property, this form of property rights is also stinted by prohibitions against unacceptable uses, such as activities that pollute. For example, individuals may own land and the fish and wildlife on that land, but the types of use to which they can put either the land or the wildlife may be constrained by law.

Property Rights over Goods

The property rights listed in **Table 1** have one attribute in common: they have, in practice, been applied almost exclusively to environmental goods. We think of resources as natural capital that create value through the size of standing stock and through the flow of resources from that stock. We are familiar with property rights over the stock and flow components of natural capital. For example, in marine ecosystems rights can be assigned to the resource only on capture (competitive fishing), to the standing stock itself (territories for sedentary species), or to the flow of goods from the standing stock (individual fishing quotas). These are the tangible components of ecosystems.

Property Rights over Services

There are other intangible components of ecosystems that do not fall under these definitions of rights. For example, species provide services to the ecosystem and to humans through genetic information, reproduction, and contribution to critical ecosystem functions. Species and groups of species also provide additional value to humans through the genetic information they contain and the pleasure they provide. Property rights are poorly defined for ecosystem services such as genetic information. This is in spite of the fact that the idea of intellectual property rights over discoveries leading from information and ideas is a familiar one. Nations have patent and copyright systems that protect rights of ownership over goods and services that flow from the application of intellectual capital; these include designs, published works, and medicines. But the focus of intellectual property rights is on the discovery or creation that results from information, not on the information itself. The lack of systems of property rights for genetic information is a critical issue for its protection. The information stored in biodiversity could lead to discoveries and patented products, but for the potential investor, the necessary protections to encourage investment are not in place. Since rights to the use of genetic information are unspecified, incentives for its protection are also absent. Biodiversity is treated as an open access resource, with the familiar "tragedy of the commons" result. The coexistence of intellectual property rights and biodiversity protection has been an area of great controversy and the focus of international efforts to reconcile interests of the developing and developed states (Kothari and Anuradha, 1999).

Optimal Forms of Property Rights

Biodiversity is a public good at both national and international scales. Developing the means to protect its value is a critical challenge to property rights, and it is widely acknowledged that strengthening property rights over ecosystem goods and services is a necessary component of biodiversity protection. Some of the same conclusions that apply to property rights and the protection of single resources apply as well to multiple resources and biodiversity. The evidence of the applied property rights literature indicates that no single type of property rights can be a remedy for all needs of resource protection; the appropriate type depends on the resource management objectives and on the context. People often advocate a particular form of property rights as being best suited to manage and sustain natural resources. For example, arguments have been made for private property rights to natural resources to provide incentives for maintaining the flow of resource goods and services into the future. The idea is that private property rights would keep people from the "prisoner's dilemma," in which ignorance over the behavior of others leads to the collectively irrational outcome of resource overuse. The argument for universal private property ignores the differences in context in which they might be applied. It also disregards possibilities for cooperation and collective action that provide assurance about others' behavior and that exist under other types of rights systems (Ostrom, 1990; Bromley, 1991; Hanna *et al.*, 1996).

In recent years counter-arguments have arisen for the superiority of collective ownership over private ownership. These arguments are often framed in the context of community-based resource management. The idea behind this argument is that collective resource ownership and management is superior in terms of providing incentives for stewardship, providing collective access to resource benefits, and fairness in resource outcomes. Disputing each of these arguments for the superiority of a single type of property right is a large body of research that has demonstrated that there is no particular form of property right that is superior in all cases. Each type of property right – with the exception of open access, which is the absence of rights – can either succeed or fail in sustaining resources. What matters is how well the property right system fits within the ecological, economic, and social context and how well the form of the right reflects the type of use.

The fundamental problem for biodiversity protection is that much of the biodiversity value falls outside the realm of direct use. Direct uses are easiest to monitor and measure, control and trade. Indirect uses such as genetic information and aesthetic appreciation are less visible because their contribution is more diffuse over space and time. Because these resource values are seldom owned, bought, or sold, they are often at a competitive disadvantage in a market economy, leaving them undervalued and overused. This leads to the common approach of protecting biodiversity not through property rights but by removing them from human influence in sanctuaries, refuges, or reserves.

Functions of Property Rights

Property rights have several functions. They delineate the population of legitimate owners. They specify the allowable

actions of these owners and their associated responsibilities so that expectations are consistent and enforcement of rules is possible. And by setting consistent expectations they reduce uncertainty about others' behavior. In the larger sense, they connect the pieces of the natural system to the pieces of the human system. The basic functions of managing resources – coordinating users, enforcing rules, and adapting to changing environmental conditions – cannot be met without a system of property rights. The way that property rights function in any particular context determines whether the natural and human system will interact in compatible or conflicting ways (Hanna *et al.*, 1996). To be compatible with long-term biodiversity protection, systems of property rights must function in a way that deals with uncertainty, externalities, transactions costs, and scale.

Uncertainty

Uncertainty is endemic in natural systems. We lack knowledge about ecosystem structure and ecosystem condition. We often lack assurance about how others will behave, or we lack confidence about the future.

For example, there are large gaps in basic biodiversity knowledge; about threshold levels of protection, the measurement of ecosystem function, and the definition of biodiversity goals. In marine ecosystems relatively little is known about ecosystem composition, links between species, food, and reproductive requirements, critical ecological processes, and tradeoffs among species. These uncertainties affect the performance of property rights. Uncertainty about ecosystems may limit the ability to define goals and objectives for biodiversity. Uncertainty about the behavior of others will encourage intensified use to capture as many benefits as possible while they are available. Uncertainty about the future will shorten the time horizons over which decisions are made, removing the incentive to invest in long-term protection. All of these forms of uncertainty work against the protection of biodiversity.

In an uncertain environment, decision making takes place through trial and error. The uncertainty creates a natural tension between the individual and the group, and between people and ecosystems. Property rights address some, but not all, of the components of uncertainty. They define and sanction ownership, resolving the question of future access. They provide assurance about the behavior of others, because they define appropriate types of use. They do not in themselves increase knowledge about the ecosystem, although they may provide an incentive for owners of property rights to produce this information (Hanna, 1998).

Externalities

Another function of property rights is to resolve the problem of externalities, in which one action affects another. When property rights to resources do not exist or are incomplete, people do not take full account of the costs of their actions because there is no corresponding "owner" to lay counter claims. In open access fisheries, for example, people compete for resource benefits by fishing as fast as they can and catching as much as they can. The result is that one person's behavior affects the amount of fish available to others over time, unless their rights are correspondingly protected. Similarly, destructive fishing practices can eventually affect the functioning of the ecosystem if rights to habitat or nonfished species are not in place. In some cases property rights may be defined but unenforceable, and the lack of enforcement then becomes equivalent to removing the right. For example, if fishing rights are expressed in terms of areas and people without rights are not excluded from those areas, their encroachment will render the right meaningless and externalities will continue.

The outcome of missing or unenforceable property rights is biodiversity loss, as only some components of the ecosystem are protected, leaving others vulnerable to external effects. The idea behind using property rights to protect biodiversity is to assign claimants to the full spectrum of ecosystem components and services, so that external effects are accounted for, or "internalized." If the owners of biodiversity are citizens, the state can then act in their behalf to protect biodiversity from various damages over space and time. The ability of the state to represent claims over biodiversity depends on a clear definition of the goods and services of ecosystems, an articulation of the relation between ecosystem components and ecosystem function, and an active constituency for protection (Perrings *et al.*, 1995; Sedjo and Simpson, 1995).

Transactions Costs

How well a system of property rights functions both affects and is affected by transactions costs. Transactions costs are the costs of doing business, which in the resource management context include costs of gathering information, coordinating users, organizing decision making, and enforcing rules. Some transactions costs remain fixed regardless of the type of process used to make decisions. Others vary with the way decisions are made – the amount of data collected, analyses done, and the process used to make decisions. Transactions costs are also influenced by the condition of the ecological system. As resources become depleted, a system of property rights must account for more and more externalities that increase the costs of management program design and enforcement. It is possible to create a system so costly to design or enforce that potential benefits are outweighed by the costs. This cost effect, particularly in consideration of actions to change property rights, is an important factor in the consideration of the transition to property rights for protecting biodiversity (Eggertsson, 1990; North, 1992).

Scale

The functioning of property rights is also affected by the scale of the area over which they apply. There is a disconnect between the geographic scale over which species are distributed and the scale at which species habitat is owned. Property rights to goods are relatively easy to define and enforce because they are associated with geographic space and their value is embodied in the goods themselves. But property rights to services are much more elusive. For example, rights to the information

value of species would involve several owners, even nations, over wide geographic space.

For rights to resources to have meaning in terms of providing incentives for their conservation, those rights must have value to the owners. The value rests in part on their uniqueness and exclusivity. When the resource is the genetic information encoded in a plant and the plant species is distributed over a range larger than that areas encompassed by the property right, the information embodied in that plant is also available to others, leaving the rights owner without the power to exclude or to capture the potential financial value from the information. This scale mismatch limits the potential for property rights to provide the appropriate incentives for investing in conserving the genetic information.

The scale question also means that biodiversity protection is often an international issue. There are global benefits to biodiversity that are unaccounted for in national systems of property rights. Species distributions are independent of national boundaries. In addition, the loss of biodiversity is often an international externality, where the impact of one nation's actions is felt by another (UNEP, 1992; Swanson, 1995).

Evolution of Property Rights

Property rights to resources tend to evolve incrementally over time in response to changes that alter the costs and benefits of particular forms. The current impetus for considering property rights for the protection of biodiversity is the scarcity resulting from declines in the number, range, and diversity of species that enhances biodiversity's value. The evolutionary path to this point, marked by a gradual expansion of property rights over ecological goods and services, is one that has reflected the relative changes in the benefits and costs of property rights as conditions of resource use change.

Single-Species Use

When resources exist in surplus to human needs, conservation actions are unnecessary and so rights, rules, and responsibilities have little meaning. In an environment of surplus, the costs of developing protective mechanisms exceed the benefits gained. As resource abundance declines, either from human pressures or natural events, the increasing scarcity raises the benefit of protection. When the benefits of protection exceed their costs, property rights in some form develop. The first controls are on how the resource is taken or how much is taken. For example, an open access, no-rules fishery becomes restricted by how people can fish – for example, by seasons or by limits on gear – and how many fish they can take. Eventually, property rights of access to the fishery are developed to restrict who can fish.

Ecosystem Use

As ecosystems are further exploited they reach a point where the focus on extracting and conserving individual species erodes the functioning of the ecosystem as a whole. Exploitation of species is interconnected, the sources of impacts are diffuse, and effects of exploitation are system-wide. Species richness declines, community interactions are altered, and resilience erodes. It is at this point that the focus of managing human effects shifts from single species to ecosystems. This shift in focus takes place against a background of using resources for the production of goods: lumber from trees and seafood from fish. The rights, rules, and responsibilities are all directed toward this end. They are not designed to control effects on unused ecosystem components or to protect the flow of ecosystem services; the system of property rights is poorly adapted to these interactive effects.

Biodiversity decline creates scarcity in the services provided by species richness and genetic information. But the same scarcity in ecosystem services that provides an incentive to create property rights also presents a difficulty to developing ecosystem-level property rights. Ecosystem management is typically proposed when exploitation levels of single species are too high and showing signs of stress. Human competition and conflict over access to individual species complicate the shift to a broader focus and cause users to challenge the legitimacy of an ecosystem approach.

Expectations About Use

Property rights systems, once in place, create expectations about what is normal. They start in motion a new path of development. At each stage in the path, the condition of the resource, the definition of rights, and the expectations about those rights influence action. The further along the path the more embedded are the property rights and the more vested are people in their continuance. When faced with the necessity of expanding those rights to the protection of biodiversity, it must be done against the legacy of the path to this point. A property rights transition to the protection and conservation of biodiversity is affected by resource conditions at the intervention point. Protecting biodiversity means losing access to some of the goods and services of the ecosystem, which, under the old property rights rules, were at the disposal of the owner. Many resources have competing uses with known market values. Unknown future uses of these resources will be hard pressed to compete against known present uses.

Property Rights and Biodiversity Protection

A number of emerging issues illustrate the challenges to the use of property rights to protect biodiversity. Four are particularly relevant: uncertainty, exclusivity, distribution of benefits, and the alignment of private and social goals.

Uncertainty

Protecting biodiversity requires more than a set of rules and responsibilities for property rights set within institutions that accommodate the attributes of the ecosystem and the people who use it. How to design property rights and institutions for this accommodation is what is at issue. Although there is general understanding of the importance of biodiversity to the stability, function, and sustainability of ecosystems, there is

poor understanding of specifically what to protect. For example, knowledge is generally lacking of the thresholds at which biodiversity loss irreversibly changes ecosystems. Knowledge of the role played by individual species in contributing to critical ecosystem functions is also absent. These uncertainties create a corresponding uncertainty in the objectives and design of property rights to reflect the full range of values involved (Hanna, 1998).

Exclusivity

A problem with many property rights that apply to natural resources is that they do not specify claims to the full range of goods and services provided by an ecosystem. In failing to fully specify property rights claims, they fail to protect exclusive use. If property rights were defined for all components of an ecosystem, users and decision makers would take all the consequences of their actions into account. This would be the first step in making biodiversity conservation profitable. But under current systems of property rights, this is rarely the case. The history is to apply property rights to natural resources as commodities but not to the services they provide or to their existence value. The lack of full specification means that it is unclear who can claim and control rights of use.

Distribution of Benefits and Costs

How benefits and costs are distributed is the core debate of natural resource policy. The same debate is at the heart of biodiversity protection. Biodiversity values have the potential to be protected through a number of different types of property right. But whatever type of right is used will create winners and losers. It is difficult to design an effective system of property rights without addressing the distributional questions as well as questions of what the objectives are, how progress toward those objectives will be measured, and the time frame over which they will be met.

How to promote equitable distributions is a difficult question in all resource policy, but it is particularly so for the complex issue of biodiversity. Equity questions extend into the international arena, where the distribution of benefits between rich and poor nations is at stake. Those who receive the benefits of biodiversity protections may be distant interests, while those who incur the costs are local resource owners. Interest in protecting ecosystem services may exist among people outside a local area, but unless these interests coincide with the objectives of local owners, protection efforts are likely to be unsuccessful. Effective resource protection depends on legitimacy – on the acceptance of rules and procedures by participants. Many natural resource systems are difficult to monitor, and the possibilities for circumvention of rules are many. When people doubt the legitimacy of the system of property rights because they cannot accept its distributional outcomes, their incentives are to undermine rather than promote its evolution to a new form. Scarcity compounds the erosion of legitimacy by creating greater incentives and opportunity for rent seeking that is characteristic of resource competition.

Alignment of Private and Social Goals

For property rights to bring private preferences into line with public preferences for biodiversity conservation, the tensions between private and social goals must be resolved. Individuals have private goals for productivity of ecosystems goods that may not be compatible with social goals for biological production. Property rights can be an effective mechanism to create the appropriate incentives to conserve biodiversity, but the existence of property rights is not necessarily a sufficient condition for conservation. To promote biodiversity conservation property rights must be combined with resource management objectives that are consistent with conservation.

In addition, biodiversity is a public good, so it is subject to the potential for free riders to enjoy the benefits without paying the costs. Biodiversity protection must be accomplished in concert with existing rights and rules that favor direct uses of ecosystem goods. How to realign private and social incentives to conserve biodiversity is an important and to a large extent unresolved question. Most current systems of rights, because they are still unspecified for ecosystem services, favor the conversion of ecosystems into goods. Options for change include the expansion of the scope of property rights, payment of compensation to owners for conserving rather than using resources, or developing prohibitions against certain uses.

Conclusions

Property rights outline the conditions under which resources can be used and the interactions of people in using them. They exist in particular contexts that are combinations of ecosystems and people. There are many forms and permutations, none of which is superior to others. What determines the effectiveness of property rights in protecting biodiversity is the extent to which they are able to perform the functions of reducing uncertainty, removing externalities, containing transactions costs, accommodating scale and reflecting biodiversity goals.

Property rights evolve in response to external conditions. Scarcity is the driving force for their existence, and their contribution of benefits in excess of costs is the means of their continuance in a particular form. Despite the increasing scarcity of biodiversity, particular challenges face the evolution of systems of property rights to meet these conservation needs. To protect biodiversity, a transition must be made between current systems of rights that protect direct uses of ecosystems to expanded systems of rights that also protect indirect ecosystem services such as genetic information and aesthetic enjoyment.

The application of property rights to protect biodiversity is challenged by forces that are both internal and external to ecosystems. Internally, the increasing scarcity brought on by overexploited resources means that the range of options has declined. Fewer adaptations are possible. Formidable information needs for keystone species and critical functions of ecosystems exist. Both internally and externally, forces are competing to shape the time horizon of resource management. Biodiversity protection and notions of sustainable

use – values held by the public at large – are long-term concepts, requiring a long time horizon for management decision making. At the same time, the overexploited levels of many resources combined with uncertainty about their sustainability leads to internal pressures to shorten the time horizon and make decisions for the short run.

As these external and internal pressures illustrate, the very conditions in ecosystems that require biodiversity protection are also those which are causing difficulties in its implementation. Increased knowledge of ecosystems is needed as are innovative designs for property rights that work within and across national boundaries.

See also: Biodiversity as a Commodity. Commons, Concept and Theory of. Environmental Ethics. Land-Use Issues. The Value of Biodiversity

References

Bromley DW (1991) *Environment and Economy: Property Rights and Public Policy*. Oxford, UK: Basil Blackwell.

Ciriacy-Wantrup SV and Bishop R (1975) Common property as a concept in natural resources policy. *Natural Resources Journal* 15: 713–728.

Eggertsson T (1990) *Economic Behavior and Institutions*. Cambridge: Cambridge University Press.

Hanna SS (1998) Institutions for marine ecosystems: Economic incentives and fishery management. *Ecological Applications* 8(1): S170–S174.

Hanna S, Folke C, and Mäler K-G (1996) *Rights to Nature*. Washington, DC: Island Press.

Hardin G (1968) The tragedy of the commons. *Science* 162: 1243–1247.

Kothari A and Anuradha RV (1999) Biodiversity and intellectual property rights: Can the two co-exist? *Journal of International Wildlife and Policy* 2(2): 204–223.

North DC (1992) *Transactions Costs, Institutions, and Economic Performance*. San Francisco: Institute for Contemporary Studies.

Ostrom E (1990) *Governing the Commons*. Cambridge: Cambridge University Press.

Perrings C, Maler K-G, Folke C, Holling CS, and Jansson B-O (1995) *Biodiversity Loss: Economic and Ecological Issues*. Cambridge: Cambridge University Press.

Schmid AA (1995) The environment and property rights issues. In: Bromley D (ed.) *The Handbook of Environmental Economics*, pp. 45–60. Cambridge, MA: Blackwell.

Sedjo RA and Simpson RD (1995) Property rights, externalities and biodiversity. In: Swanson TM (ed.) *The Economics and Ecology of Biodiversity Decline*, pp. 79–88. Cambridge: Cambridge University Press.

Swanson T (2003) Introduction to property rights and biodiversity conservation: Convergence or conflict? *Land Economics* 79(4): 457–459.

Swanson T (1995) The international regulation of biodiversity decline: Optional policy and evolutionary product. In: Perrings C, Maler K-G, Folke C, Holling CS, and Jansson B-O (eds.) *Biodiversity Loss: Economic and Ecological Issues*, pp. 225–259. Cambridge: Cambridge University Press.

TEEB. *The economics of ecosystems and biodiversity: Mainstreaming the economics of nature: A synthesis of the approach, conclusions and recommendations of TEEB*. United Nations Environmental Programme (2010). Available at www.teebweb.org.

United Nations Environmental Programme (UNEP) *Convention on biological diversity* (1992).

Proteorhodopsins: Widespread Microbial Light-Driven Proton Pumps

Oded Béjà, Technion – Israel Institute of Technology, Haifa, Israel
Jarone Pinhassi, Linnaeus University, Kalmar, Sweden
John L Spudich, The University of Texas Medical School, Houston, TX, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Chromophore A part of a molecule that absorbs light.
Metagenomics Analysis of genomic deoxyribonucleic acid from complex environmental populations or communities (as opposed to specific cultured isolates) using molecular biology techniques. Keystone approach to determine the genetically encoded information of natural microbial populations and thereby of the great majority of microorganisms that cannot easily be cultured.
Photocycle In proteorhodopsin (PR), the photochemical reaction cycle ("photocycle") denotes the process in which light energy absorbed by the retinal chromophore provokes a series of transient chemical changes in the PR molecule, which then returns to its dark state in times of less than 1 s. This cyclic process results in the transport of a proton from the inside of the cell across the cell membrane.
Phototrophy In a wide sense, it signifies the capacity to obtain food or nutrition from light. In microbiology, it denotes the process whereby light is used as an energy source for metabolic processes. We currently know organisms that carry out oxygenic or anoxygenic photosynthesis using chlorophylls to harvest light, and organisms that contain microbial rhodopsins, which use retinal for the same purpose.
Retinal The rhodopsin chromophore, also called vitamin A aldehyde or retinaldehyde. Retinal is responsible for the color of rhodopsin pigments.

Proteorhodopsins (PRs) are microbial membrane proteins that use the retinal molecule (vitamin A aldehyde) as a chromophore and function as light-driven proton pumps (Béjà *et al.*, 2000a, 2001; Spudich and Jung, 2005). These microbial rhodopsins were first discovered in 2000 using metagenomic data from the uncultured SAR86 bacterial group, and their detection has been instrumental in the development of the metagenomics research field (Béjà, 2004; Kowalchuk *et al.*, 2007; O'Malley, 2007; Grote and O'Malley, 2011). Since their discovery in the SAR86 group, PRs have been detected in different abundant and cosmopolitan oceanic microbial groups belonging to Proteobacteria, Flavobacteria, Gram-positive bacteria, Cyanobacteria, Archaea, and even picoeukaryotes. Although most PRs have been identified in uncultured marine bacteria and archaea, the relative ease of working with PRs expressed in *Escherichia coli* bacteria has been central for dissecting their light-dependent proton pump activity (using advanced biophysical assays (Friedrich *et al.*, 2002; Wang *et al.*, 2003; Man-Aharonovich *et al.*, 2004; Sineshchekov and Spudich, 2004; Bergo *et al.*, 2009) and in suggesting a role in phototrophy (Martinez *et al.*, 2007; Walter *et al.*, 2007)). However, because all these works relied on the heterologous *E. coli* system, some concern has arisen regarding PR phototrophy relevance in the natural environments (Schwalbach *et al.*, 2005). Ecophysiological studies of native marine PR-containing bacteria are currently resolving these concerns (Gómez-Consarnau *et al.*, 2007, 2010; González *et al.*, 2008; Kimura *et al.*, 2011; Steindler *et al.*, 2011). PRs are estimated by different methods (metagenomics and real-time polymerase chain reaction (PCR)) to be found in approximately 10% (Sabehi *et al.*, 2005; Fuhrman *et al.*, 2008) and as much as 85% (Fuhrman *et al.*, 2008) of marine microbes in the photic zone. The fact that surface seawater holds nearly 1 billion microbes per liter thus implies that microbial PR phototrophy directly influences oceanic energy and carbon fluxes.

PR Discovery and Mode of Phototrophy

The PR gene was discovered on a deoxyribonucleic acid (DNA) fragment belonging to a member of the uncultured bacterial SAR86 group isolated from Monterey Bay, CA, USA. This discovery became possible due to the development of methodologies to generate large-insert genomic libraries (bacterial artificial chromosome (BAC) libraries) of environmental DNA (Béjà *et al.*, 2000b). During analysis of the BAC libraries for genes involved in metabolism of particular members of the natural bacterioplankton community, the startling presence of the PR gene was noticed based on its weak homology (approximately 30% identity on the protein level) to haloarchaeal bacteriorhodopsins (Béjà *et al.*, 2000a; **Figure 1**). At the time of PR discovery, bacteriorhodopsins were already well characterized (Oesterhelt and Stoeckenius, 1971; Danon and Stoeckenius, 1974; Oesterhelt and Schuhmann, 1974), and their role as retinal-binding membrane proteins that function as proton pumps was well established (Lanyi, 1998; Haupts *et al.*, 1999; Spudich *et al.*, 2000). PRs were characterized first as proton pumps using heterologous expression in *E. coli*. *In vivo*, PR proteins fold correctly in *E. coli* membranes to form active colored proton pumps (**Figure 2**), an ability not shared by their well-characterized close relatives, bacteriorhodopsins. The first experiments were performed using retinal molecules added externally because *E. coli* does not possess the enzymatic capacity to synthesize retinal. In later experiments, *E. coli* was engineered to produce retinal via different operons found in different cultured and uncultured bacteria (Sabehi *et al.*, 2005; Kim *et al.*, 2008). Similar to

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00406-8

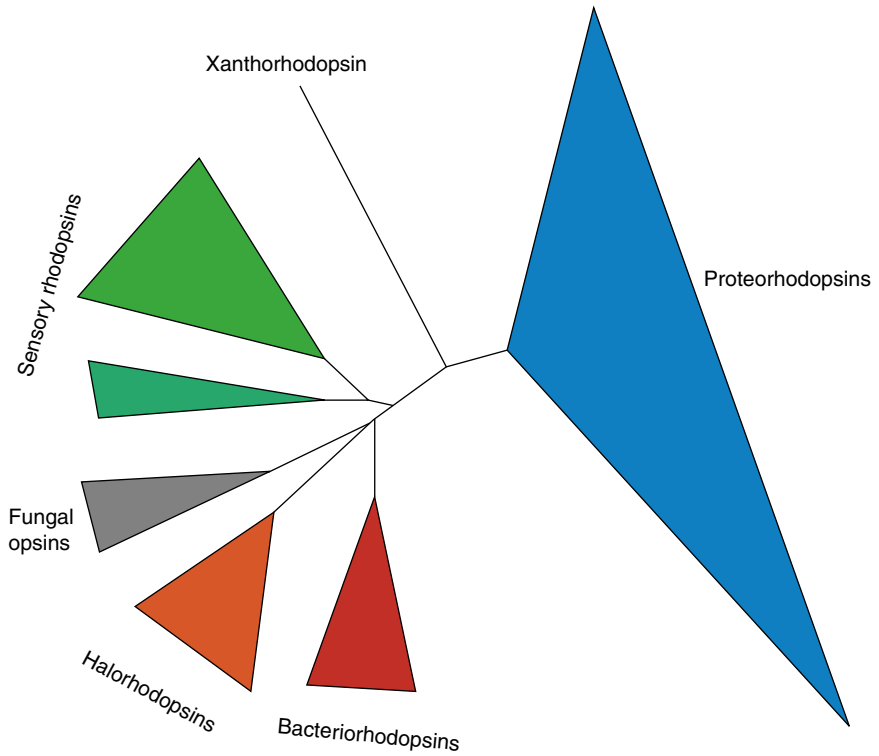


Figure 1 Schematic representation of a microbial rhodopsin tree. Modified from McCarren J and DeLong EF (2007) Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environmental Microbiology* 9: 846–858.



Figure 2 Pigmented *E. coli* cells expressing different PR variants from the Red Sea. Modified from Man D, Wang W, Sabehi G, *et al.* (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO Journal* 22: 1725–1731.

bacteriorhodopsins, PRs are believed to work using *all-trans*-retinal as chromophore. This expectation is mainly based on their high degree of conservation of retinal-binding pocket residues with other microbial rhodopsins, which use *all-trans*-retinal (Spudich *et al.*, 2000), the generation of photoactive functional pigments in *E. coli* membranes by exogenous *all-trans*-retinal (Béjà *et al.*, 2001), analysis of native planktonic membrane preparations (Béjà *et al.*, 2001), and on analyses of known retinal biosynthetic operons identified in DNA fragments coming from uncultured bacteria containing PR genes (Sabehi *et al.*, 2005). However, it has not been ruled out that, in the environment, PRs may interact with other chromophores as well (e.g., 3-hydroxyretinal; Sabehi *et al.*, 2004).

Using the *E. coli* system, it is now well established that PRs are indeed retinal-based light-driven proton pumps (Béjà *et al.*, 2000a, 2001; Man-Aharonovich *et al.*, 2004; Sabehi *et al.*, 2005; Martinez *et al.*, 2007; Walter *et al.*, 2007; Kralj *et al.*, 2011). Furthermore, using the same system, it was demonstrated that the PR light-driven proton pumping is coupled to adenosine triphosphate (ATP) biosynthesis (Martinez *et al.*, 2007; see **Figure 3** for a simple view of the process). Moreover, the PR proton pumping is also likely to provide the proton motive force to drive energy-requiring processes, such as flagellar motility or nutrient transport across the cell membrane.

Although the work with heterologous expression in *E. coli* enabled the biochemical and physiological characterization of the PR pump, establishing an authentic ecological role for the PRs has largely depended on the study of native bacteria containing PRs. The first ecophysiological experiments with marine bacteria containing PR was carried out with cultured members of the SAR11 group (Giovannoni *et al.*, 2005a) and the SAR92 group (Stingl *et al.*, 2007). SAR11 bacteria (Morris *et al.*, 2002). *Candidatus Pelagibacter ubique* (Rappé *et al.*, 2002) are abundant and cosmopolitan Alphaproteobacteria that contain PRs (Giovannoni *et al.*, 2005a, b). Notably, the first attempts to show light-stimulated growth in both SAR11 and

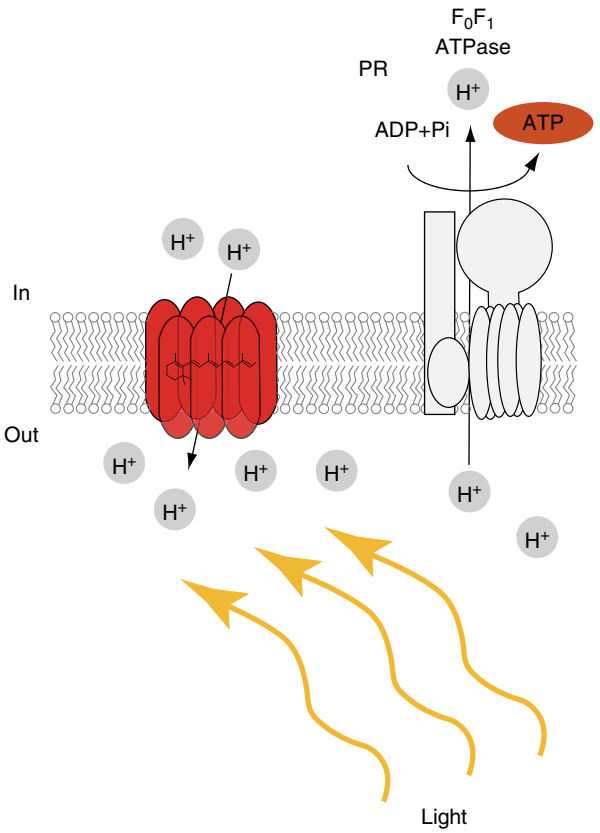


Figure 3 A simple view of a PR energy circuit. PR uses light energy to translocate protons across the cell membrane. The excess extracellular protons create a proton motive force, which can energetically drive flagellar motility, transport processes, or ATP synthesis in the cell. For the latter, a proton-translocating ATPase, a multiprotein membrane-bound complex, utilizes the proton motive force to synthesize ATP.

SAR92 failed (Giovannoni *et al.*, 2005a; Stingl *et al.*, 2007). This greatly heated up the discussion over the true relevance of PRs in the environment (Giovannoni *et al.*, 2005b; Schwalbach *et al.*, 2005; Bryant and Frigaard, 2006; Fuhrman *et al.*, 2008; Miyake and Stingl, 2011).

Light-stimulated growth mediated by PR phototrophy, coupled with biophysical characterizations of the PR molecule, was first observed in a series of experiments with a PR-containing marine Flavobacteria, *Dokdonia* sp. strain MED134 (Gómez-Consarnau *et al.*, 2007). Subsequent independent experiments with the same strain recently confirmed these observations and further showed that inhibition of retinal biosynthesis in strain MED134 abolished the light response, supporting a direct role for PR in stimulating growth in the light (Kimura *et al.*, 2011). In another PR-containing Flavobacteria, *Polaribacter* sp. strain MED152, experiments indicate that the energy provided by PR phototrophy could be used to fuel CO_2 fixation (González *et al.*, 2008). Direct evidence for the role of PR phototrophy in a marine bacterium first came with the mutational analysis of the PR gene in *Vibrio* sp. AND4 (Gómez-Consarnau *et al.*, 2010), where it was also shown that the energy provided through PR promotes survival during extended starvation conditions. In a similar manner, although initially no effect of light was seen during growth in members of the SAR11 group, ATP production through PR phototrophy was recently shown to be beneficial for SAR11 during starvation (Steindler *et al.*, 2011). Collectively, these experimental results from native marine bacteria strongly substantiate the original suggestion that phototrophy mediated by PR may influence global oceanic carbon and energy fluxes (Béjà *et al.*, 2000a).

PR Abundance and Diversity

The uncultured gammaproteobacterial group SAR86 in which PRs were first detected is widespread in different oceanic regions (Eilers *et al.*, 2000; Suzuki *et al.*, 2001). PRs are also found in the cosmopolitan alphaproteobacterial SAR11 group (Giovannoni *et al.*, 2005a; Steindler *et al.*, 2011) and in the SAR116 group (Oh *et al.*, 2010). These three bacterial groups alone comprise as much as 70% of the bacterial community in most surface waters around the world. PRs are also readily detected in other microbial groups, including other Alphaproteobacteria (see table in DeLong and Béjà, 2011); Betaproteobacteria (Giovannoni *et al.*, 2008); Gammaproteobacteria (Gómez-Consarnau *et al.*, 2010); Flavobacteria (Gómez-Consarnau *et al.*, 2007; González *et al.*, 2008); Planctomycetes (McCarren and DeLong, 2007); Cyanobacteria; Actinobacteria (Sharma *et al.*, 2008, 2009); marine Archaea (Frigaard *et al.*, 2006); and different eukaryotic groups, including fungi (Spudich and Jung, 2005) and dinoflagellates (Okamoto and Hastings, 2003; Slamovits *et al.*, 2011). The widespread distribution of different PRs in different microbial groups could be explained by frequent horizontal gene transfer of the PR gene (Frigaard *et al.*, 2006) and the adjacent retinal biosynthetic operon (McCarren and DeLong, 2007).

PRs have been reported in a variety of aquatic habitats, including the marine environment (Béjà *et al.*, 2000a, 2001; de la Torre *et al.*, 2003; Balashov *et al.*, 2005; Giovannoni *et al.*, 2005a; Sabehi *et al.*, 2005; Frigaard *et al.*, 2006; Gómez-Consarnau *et al.*, 2007, 2010; Campbell *et al.*, 2008; Cottrell and Kirchman, 2009; Lami *et al.*, 2009; Oh *et al.*, 2010), sea ice (Koh *et al.*, 2010), brackish environments (Rusch *et al.*, 2007; Atamna-Ismaeel *et al.*, 2008), and freshwater lakes (Atamna-Ismaeel *et al.*, 2008; Sharma *et al.*, 2008, 2009). Recently, PRs have also been detected on leaf surfaces of plants, i.e., phyllospheres (Atamna-Ismaeel *et al.*, 2012). Leaf surfaces of terrestrial plants cover an estimated surface area of 6.4×10^8 km^2 and comprise the main interface between terrestrial biomass and solar photon flux. This habitat harbors an immensely diverse microbial community of as much as 10^6–10^7 cells per cm^2 leaf surface (Lindow and Brandl, 2003). A mode of phototrophy that is compatible with the plant's photosynthesis would offer a significant ecological advantage to microbes inhabiting this environment.

PR genes have been found in different environmental genomic studies on large environmental DNA fragments and in different metagenomics studies that use different sequencing techniques and also using single-cell genomics (Stepanauskas and Sieracki, 2007; Woyke *et al.*, 2009; Martinez-Garcia *et al.*, 2011). In addition to the detection of specific genes, PRs have been detected in different metatranscriptomics studies and in

high numbers (Frias-Lopez *et al.*, 2008; Poretsky *et al.*, 2009; Shi *et al.*, 2010). In addition, PRs were also detected as proteins (peptides) in different environmental metaproteomics studies (Giovannoni *et al.*, 2005a; Morris *et al.*, 2010), and their proton-pumping photocycles detected spectroscopically in ocean plankton (Béjà *et al.*, 2001). Taken together, the results from studies with PRs on the DNA, ribonucleic acid, and protein levels confirm the importance of this photoprotein in light-exposed environments.

PR Spectral Tuning

PR spectral tuning was first observed among genes amplified directly from the environment by using highly specific PCR primers (Béjà *et al.*, 2001) to PRs from the SAR86 group. These genes produced pigments of two distinct absorption spectra – "green-absorbing" and "blue-absorbing" PR types – which shared >78% of their amino acid residues. Moreover, a single residue change at position 105 (leucine for green PRs and glutamine for blue PRs) functions as a spectral tuning switch to account for most of the spectral difference (Man *et al.*, 2003). In fact, the switch of a single residue from leucine to glutamine in PRs requires a switch of only one nucleotide (a change of a CTA or a CTG codon to a CAA or CAG codon, respectively). Initial work performed in the Pacific Ocean (Béjà *et al.*, 2001) identified green PRs mostly in shallow water, whereas blue PRs were identified in deeper waters (below 70 m). A similar trend was also observed in the Mediterranean Sea. However, in the Sargasso Sea, despite the green light available at the surface, blue PRs were observed on the surface and green PRs were almost absent.

Future Exploration

Antenna-assisted PRs

More recently, a new type of microbial rhodopsin was found: xanthorhodopsin (Balashov *et al.*, 2005). Xanthorhodopsin is a light-driven retinylidene proton pump complexed with a carotenoid molecule, salinixanthin, which functions as a light-harvesting antenna and transfers energy to the rhodopsin protein (Lanyi and Balashov, 2008). It was recently suggested that the ability of rhodopsins to bind salinixanthin depends on a single glycine residue (Imasheva *et al.*, 2009). It was recently suggested (DeLong and Béjà, 2011) that a number of PRs could potentially bind carotenoid antennae because some environmental PRs carry the homologous glycine residue in the correct position. Such a complex would mean that PRs in nature would be able to harvest a larger part of the light spectrum and would not be restricted to blue or green light.

Photoreactions and Mechanism

Light absorbed by the retinal chromophore in PR initiates a series of transient chemical changes in the molecule, which then returns to its dark state in times of less than 1 s. This cyclic process, called a photochemical reaction cycle ("photocycle" for short), results in transport of a proton across the membrane. Key events in the PR photocycle are the transfer of a proton from the chromophore site to a negatively charged aspartate in communication with the outside environment, followed by a conformational change that switches the accessibility of the chromophore site to the cytoplasmic side of the protein, and then transfer of the proton from a neutral glutamate residue on the cytoplasmic side to the chromophore site (Dioumaev *et al.*, 2002). These proton transfers are homologous to those occurring in bacteriorhodopsin (Lanyi, 1998), but other residues involved in proton transport in bacteriorhodopsin, including an important proton release group, are not conserved in PR. Many aspects of the PR-pumping mechanism remain unresolved, and an important goal for the future is to obtain an atomic resolution structure of PR to enable its molecular mechanism to be elucidated.

Comparison of ocean sequences shows the presence of the aspartate proton acceptor and glutamic acid proton donor residues in 97% of the PR-encoding genes, consistent with their expected proton-pumping function. However, the remaining 3% lack the proton donor residue, which is characteristic of microbial sensory rhodopsins rather than proton pumps. Further evidence from adjacent genes provides strong support for the existence of "sensory PRs" in the ocean (Spudich, 2006). In the recent report of PRs in phyllospheres, the percentage of putative sensory PRs was found to be much larger (25–70%). Compared with marine bacteria, phyllosphere microbes may have a greater need for light sensing than energy capture by microbial rhodopsins because the leaf surface is rich in nutrients but adaptations are needed to counteract the effects of fluctuations in light quality, intensity, and ultraviolet radiation (Atamna-Ismaeel *et al.*, 2012). The physiological functions of the sensory PRs are an interesting mystery for future exploration.

See also: Microbial Biogeography. Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program

References

Atamna-Ismaeel N, Finkel EM, Glaser F, *et al.* (2012) Microbial rhodopsins on leaf surfaces of terrestrial plants. *Environmental Microbiology* 14: 140–146.

Atamna-Ismaeel N, Sabehi G, Sharon I, *et al.* (2008) Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *International Society for Microbial Ecology Journal* 2: 656–662.

Balashov SP, Imasheva ES, Boichenko VA, Anton J, Wang JM, and Lanyi JK (2005) Xanthorhodopsin: A proton pump with a light-harvesting carotenoid antenna. *Science* 309: 2061–2064.

Béjà O (2004) To BAC or not to BAC: Marine ecogenomics. *Current Opinion in Biotechnology* 15: 187–190.

Béjà O, Aravind L, Koonin EV, *et al.* (2000a) Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science* 289: 1902–1906.

Béjà O, Spudich EN, Spudich JL, Leclerc M, and DeLong EF (2001) Proteorhodopsin phototrophy in the ocean. *Nature* 411: 786–789.

Béjà O, Suzuki MT, Koonin EV, *et al.* (2000b) Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environmental Microbiology* 2: 516–529.

Bergo VB, Sineshchekov OA, Kralj JM, *et al.* (2009) His-75 in proteorhodopsin, a novel component in light-driven proton translocation by primary pumps. *The Journal of Biological Chemistry* 284: 2836–2843.

Bryant DA and Frigaard N-U (2006) Prokaryotic photosynthesis and phototrophy illuminated. *Trends in Microbiology* 14: 488–496.
Campbell BJ, Waidner LA, Cottrell MT, and Kirchman DL (2008) Abundant proteorhodopsin genes in the North Atlantic Ocean. *Environmental Microbiology* 10: 99–109.
Cottrell MT and Kirchman DL (2009) Photoheterotrophic microbes in the Arctic Ocean in summer and winter. *Applied and Environmental Microbiology* 75: 4958–4966.
Danon A and Stoeckenius W (1974) Photophosphorylation in *Halobacterium halobium*. *Proceedings of the National Academy of Sciences of the United States of America* 71: 1234–1238.
DeLong EF and Béjà O (2011) The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *Public Library of Science Biology* 8: e1000359.
Dioumaev AK, Brown LS, Shih J, Spudich EN, Spudich JL, and Lanyi JK (2002) Proton transfers in the photochemical reaction cycle of proteorhodopsin. *Biochemistry* 41: 5348–5358.
Eilers H, Pernthaler J, Glockner FO, and Amann R (2000) Culturability and *in situ* abundance of pelagic bacteria from the North Sea. *Applied and Environmental Microbiology* 66: 3044–3051.
Frias-Lopez J, Shi Y, Tyson GW, *et al.* (2008) Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences of the United States of America* 105: 3805–3810.
Friedrich T, Geibel S, Kalmbach R, *et al.* (2002) Proteorhodopsin is a light-driven proton pump with variable vectoriality. *Journal of Molecular Biology* 321: 821.
Frigaard N-U, Martinez A, Mincer TJ, and DeLong EF (2006) Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* 439: 847–850.
Fuhrman JA, Schwalbach MS, and Stingl U (2008) Proteorhodopsins: An array of physiological roles? *Nature Reviews Microbiology* 6: 488–494.
Giovannoni SJ, Bibbs L, Cho J-C, *et al.* (2005a) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* 438: 82–85.
Giovannoni SJ, Hayakawa DH, Tripp HJ, *et al.* (2008) The small genome of an abundant coastal ocean methylotroph. *Environmental Microbiology* 10: 1771–1782.
Giovannoni SJ, Tripp HJ, Givan S, *et al.* (2005b) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309: 1242–1245.
Gómez-Consarnau L, Akram N, Lindell K, *et al.* (2010) Proteorhodopsin phototrophy confers enhanced survival of marine bacteria during starvation. *Public Library of Science Biology* 8: e1000358.
Gómez-Consarnau L, González JM, Coll-Lladó M, *et al.* (2007) Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature* 445: 210–213.
González JM, Fernandez-Gomez B, Fernandez-Guerra A, *et al.* (2008) Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proceedings of the National Academy of Sciences of the United States of America* 105: 8724–8729.
Grote M and O'Malley MA (2011) Enlightening the life sciences: The history of halobacterial and microbial rhodopsin research. *FEMS Microbiology Reviews* 35: 1082–1099.http://dx.doi.org/10.1111/j.1574-6976.2011.00281.x.
Haupts U, Tittor J, and Oesterhelt D (1999) Closing in on bacteriorhodopsin: Progress in understanding the molecule. *Annual Review of Biophysics and Biomolecular Structure* 28: 367–399.
Imasheva ES, Balashov SP, Choi AR, Jung KH, and Lanyi JK (2009) Reconstitution of *Gloeobacter violaceus* rhodopsin with a light-harvesting carotenoid antenna. *Biochemistry* 48: 10948–10955.
Kim SY, Waschuk SA, Brown LS, and Jung KH (2008) Screening and characterization of proteorhodopsin color-tuning mutations in *Escherichia coli* with endogenous retinal synthesis. *Biochimica et Biophysica Acta* 1777: 504–513.
Kimura H, Young CR, Martinez A, and Delong EF (2011) Light-induced transcriptional responses associated with proteorhodopsin-enhanced growth in a marine flavobacterium. *International Society for Microbial Ecology Journal* 5: 1641–1651.http://dx.doi.org/10.1038/ismej.2011.36.
Koh EY, Atamna-Ismaeel N, Martin A, *et al.* (2010) Proteorhodopsin-bearing bacteria in Antarctic sea ice. *Applied and Environmental Microbiology* 76: 5918–5925.
Kowalchuk GA, Speksnijder AG, Zhang K, Goodman RM, and van Veen JA (2007) Finding the needles in the metagenome haystack. *Microbial Ecology* 53: 475–485.
Kralj JM, Hochbaum DR, Douglass AD, and Cohen AE (2011) Electrical spiking in *Escherichia coli* probed with a fluorescent voltage-indicating protein. *Science* 333: 345–348.
Lami R, Cottrell MT, Campbell BJ, and Kirchman DL (2009) Light-dependent growth and proteorhodopsin expression by *Flavobacteria* and SAR11 in experiments with Delaware coastal waters. *Environmental Microbiology* 11: 3201–3209.
Lanyi JK (1998) Understanding structure and function in the light-driven proton pump bacteriorhodopsin. *Journal of Structural Biology* 124: 164–178.
Lanyi JK and Balashov SP (2008) Xanthorhodopsin: A bacteriorhodopsin-like proton pump with a carotenoid antenna. *Biochimica et Biophysica Acta* 1777: 684–688.
Lindow SE and Brandl MT (2003) Microbiology of the phyllosphere. *Applied and Environmental Microbiology* 69: 1875–1883.
Man D, Wang W, Sabehi G, *et al.* (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO Journal* 22: 1725–1731.
Man-Aharonovich D, Sabehi G, Sineshchekov OA, Spudich EN, Spudich JL, and Béjà O (2004) Characterization of RS29, a blue-green proteorhodopsin variant from the Red Sea. *Photochemical and Photobiological Sciences* 3: 459–462.
Martinez A, Bradley AS, Waldbauer J, Summons RE, and DeLong EF (2007) Proteorhodopsin photosystem gene expression enables photophosphorylation in heterologous host. *Proceedings of the National Academy of Sciences of the United States of America* 104: 5590–5595.
Martinez-Garcia M, Swan BK, Poulton NJ, *et al.* (2012) High throughput single cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *International Society for Microbial Ecology Journal* 6: 113–123.http://dx.doi.org/10.1038/ismej.2011.84.
McCarren J and DeLong EF (2007) Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environmental Microbiology* 9: 846–858.
Miyake S and Stingl U (2011) Proteorhodopsin. *Encyclopedia of Life Sciences (ELS)*. Chichester, UK: John Wiley & Sons, Ltd.http://dx.doi.org/10.1002/9780470015902.a0022837.
Morris RM, Nunn BL, Frazar C, Goodlett DR, Ting YS, and Rocap G (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *International Society for Microbial Ecology Journal* 4: 673–685.
Morris RM, Rappé MS, Connon SA, *et al.* (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420: 806–810.
O'Malley MA (2007) Exploratory experimentation and scientific practice: metagenomics and the proteorhodopsin case. *History and Philosophy of the Life Sciences* 29: 337–360.
Oesterhelt D and Schuhmann L (1974) Reconstitution of bacteriorhodopsin. *Federation of European Biochemical Societies Letter* 44: 262–265.
Oesterhelt D and Stoeckenius W (1971) Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature: New Biology* 233: 149–152.
Oh HM, Kwon KK, Kang I, *et al.* (2010) Complete genome sequence of "*Candidatus Puniceispirillum marinum*" IMCC1322, a representative of the SAR116 clade in the *Alphaproteobacteria*. *Journal of Bacteriology* 192: 3240–3241.
Okamoto OK and Hastings aJW (2003) Novel dinoflagellate clock-related genes identified through microarray analysis. *Journal of Phycology* 39: 519–526.
Poretsky RS, Hewson I, Sun S, Allen AE, Zehr JP, and Moran MA (2009) Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environmental Microbiology* 11: 1358–1375.
Rappé MS, Connon SA, Vergin KL, and Giovannoni SJ (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418: 630–633.
Rusch DB, Halpern AL, Heidelberg KB, *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: I, The northwest Atlantic through the eastern tropical Pacific. *Public Library of Science Biology* 5: e77.
Sabehi G, Béjà O, Suzuki MT, Preston CM, and DeLong EF (2004) Different SAR86 subgroups harbour divergent proteorhodopsins. *Environmental Microbiology* 6: 903–910.
Sabehi G, Loy A, Jung KH, *et al.* (2005) New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *Public Library of Science Biology* 3: e173.
Schwalbach MS, Brown M, and Fuhrman JA (2005) Impact of light on marine bacterioplankton community structure. *Aquatic Microbial Ecology: International Journal* 39: 235–245.
Sharma AK, Sommerfeld K, Bullerjahn GS, *et al.* (2009) Actinorhodopsin genes discovered in diverse freshwater habitats and among cultivated freshwater *Actinobacteria*. *International Society for Microbial Ecology Journal* 3: 726–737.
Sharma AK, Zhaxybayeva O, Papke RT, and Doolittle WF (2008) Actinorhodopsins: Proteorhodopsin-like gene sequences found predominantly in non-marine environments. *Environmental Microbiology* 10: 1039–1056.

Shi Y, Tyson GW, Eppley JM, and Delong EF (2010) Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *International Society for Microbial Ecology Journal* 5: 999–1013.

Sineshchekov OA and Spudich JL (2004) Light-induced intramolecular charge movements in microbial rhodopsins in intact *E. coli*. *Photochemical and Photobiological Sciences* 3: 548–554.

Slamovits CH, Okamoto N, Burri L, James ER, and Keeling PJ (2011) A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nature Communications* 2: 183.

Spudich JL (2006) The multitalented microbial sensory rhodopsins. *Trends in Microbiology* 14: 480–487.

Spudich JL and Jung KH (2005) Microbial rhodopsins: Phylogenetic and functional diversity. In: Briggs WR and Spudich JL (eds.) *Handbook of Photosensory Receptors*, pp. 1–24. Weinheim, Germany: WILEY-VCH Verlag GmbH & Co. KGaA.

Spudich JL, Yang CS, Jung KH, and Spudich EN (2000) Retinylidene proteins: Structures and functions from Archaea to humans. *Annual Review of Cell and Developmental Biology* 16: 365–392.

Steindler L, Schwalbach MS, Smith DP, Chan F, and Giovannoni SJ (2011) Energy starved *Candidatus Pelagibacter ubique* substitutes light-mediated ATP production for endogenous carbon respiration. *Public Library of Science One* 6: e19725.

Stepanauskas R and Sieracki ME (2007) Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proceedings of the National Academy of Sciences of the United States of America* 104: 9052–9057.

Stingl U, Desiderio RA, Cho JC, Vergin KL, and Giovannoni SJ (2007) The SAR92 clade: An abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Applied and Environmental Microbiology* 73: 2290–2296.

Suzuki MT, Preston CM, Chavez FP, and DeLong EF (2001) Quantitative mapping of bacterioplankton populations in seawater: Field tests across an upwelling plume in Monterey Bay. *Aquatic Microbiolial Ecology* 24: 117–127.

de la Torre JR, Christianson L, Béjà O, *et al.* (2003) Proteorhodopsin genes are widely distributed among divergent bacterial taxa. *Proceedings of the National Academy of Sciences of the United States of America* 100: 12830–12835.

Walter JM, Greenfield D, Bustamante C, and Liphardt J (2007) Light-powering *Escherichia coli* with proteorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* 104: 2408–2412.

Wang WW, Sineshchekov OA, Spudich EN, and Spudich JL (2003) Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. *Journal of Biological Chemistry* 278: 33985–33991.

Woyke T, Xie G, Copeland A, *et al.* (2009) Assembling the marine metagenome, one cell at a time. *Public Library of Science One* 4: e5299.

Protozoa

Bland J Finlay, Queen Mary University of London, Wareham, UK
Genoveva F Esteban, Bournemouth University, Poole, UK

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Bland J. Finlay, volume 4, pp. 901–915, © 2001, Elsevier Inc.

Glossary

Benthos The substrate at the bottom of the sea or fresh waters. Most benthic protozoa live in the spaces between sediment particles.
Commensal Of a protozoon – one that is loosely but not obligately associated with another organism (e.g., the ciliates attached to the external surfaces of crustacean zooplankton).
Endosymbiont An organism living in a long-term association inside a host organism for their mutual benefit (e.g., the endosymbiotic methanogenic bacteria that live inside anaerobic protozoa and utilize waste H_2 produced by the host).
Eukaryote An organism with membrane-bounded nuclei in its cells. Protozoa are unicellular eukaryotes.
Heterotrophic Mode of nutrition in which carbon is obtained from the organic compounds made by autotrophic organisms.
Phagotroph An organism that ingests solid food particles (e.g., bacteria).
Planktonic Living in the water column of lakes or the sea.

Protozoa are microscopic organisms with animal-like features. Each protozoon typically exists as a single, independent cell, and all free-living protozoa fit to the definition of phagotropic microbial eukaryotes. In some species, independent cells unite to form colonies (e.g., colonial choanoflagellates) (Lee *et al.*, 2000). There has never been unanimous agreement as to where the boundaries of the protozoa should lie, largely because of the extraordinary diversity of their lifestyles (Margulis *et al.*, 1990). Many species cause disease (e.g., malaria): Others thrive as commensals in the digestive tracts of ruminants and wood-eating insects, and many plant-like flagellates (e.g., *Euglena*) have at various times been referred to as protozoa, algae, or plants. Here, the authors are concerned principally with free-living protozoa, and the definition of this group is not taxonomic but one that is based on its key function in the natural environment: Protozoa are capable of phagotrophy – the ability to capture and ingest food particles (Fenchel, 1987). Many also have functional chloroplasts or endosymbiotic algae, and the resulting consortia are known as "mixotrophs" because they are capable of both phagotrophy and phototrophy (Esteban et al., 2010). Like most microorganisms, protozoa have very large population sizes, and they are the most abundant group of phagotrophic organisms in the biosphere. Biodiversity at the level of protozoa has characteristics that are not shared by macroscopic animals and plants. Most protozoan species are probably globally ubiquitous, so the global number of species is relatively small. A significant proportion of local protozoan species richness, at any moment in time, is rare or cryptic and awaiting the arrival of environmental conditions suitable for growth and reproduction.

Functional Roles

All of the important functional roles of free-living protozoa derive from their small size. The smallest flagellates are 2–4 μm and most are <20 μm, most amebae are <50 μm, and most ciliates are <200 μm. Exceptionally, some ameboid protozoa, such as the radiolarians and foraminiferans, and the agglutinated foraminiferans of the deep-sea benthos, may reach 2 mm or more, especially if they have spiny extensions. Because protozoa are so small, most suitable prey items are other, smaller microbes. Protozoa are the principal consumers of the immense natural resource of bacteria and other microorganisms, and because they have population growth rates that are similar to those of the microbes on which they feed (doubling times on the order of one to several days), they are usually able to control microbial abundance. Flagellated protozoa can probably consume all bacterial production in the plankton. In the benthos, protozoa overlap in their niche requirements with nematodes, rotifers, tardigrades, turbellarians, and gastrotrichs, but because of their great abundance protozoa are indeed quantitatively the most important grazers in the freshwater and marine (including deep-sea) benthos. Also, just as microbes achieve astronomical abundance on a global scale, so too are the protozoa that graze on them represented by species populations with proportionately smaller but still unimaginably large global abundances.

Protozoan grazing on microorganisms also stimulates activity of the whole microbial community, in both oxic and anoxic environments. The process involved is not fully understood, although it may operate by increasing the rate of turnover of essential nutrients that would otherwise remain locked up in the bacterial biomass. The net effect is that grazing by protozoa stimulates the rate of decomposition of organic matter (Fenchel, 1985; Finlay et al., 1997).

The variety of shapes, sizes, and relative abundances of microbial food items has driven the evolution of a comprehensive suite of methods to capture them and a considerable diversification of protozoan morphologies. In general, the size of a protozoon relative to its prey dictates the most efficient food-capturing mechanism. Where the predator is

typically much larger than its prey, filter-feeding prevails, and where the size difference is less the protozoon is more likely to be a raptorial feeder – one that seeks out relatively large individual food items. Thus, the planktonic choanoflagellate feeding on a dilute suspension of tiny bacteria 20 times smaller than itself does so with a very fine filter, and the ciliate feeding on dinoflagellates half of its size seeks out and captures each one individually. There are, of course, many exceptions, such as the marine heterotrophic dinoflagellates that use a feeding veil to trap and digest diatoms much larger than themselves.

The third main feeding type in protozoa is termed "diffusion feeding." This is particularly common in planktonic ameboid protozoa (radiolarians, foraminiferans, and heliozoans) and in the suctorian ciliates. It works when prey items collide with the sticky spines, tentacles, or axopods that radiate from the protozoon. Unlike the other two main modes of feeding, the protozoon simply waits for the arrival of its prey, much as a spider waits for an insect to be snared in its web.

Thus, there is a close link between protozoan morphology (especially of the food-capturing organelles) and the way in which a protozoon functions as a grazer. Therefore, when the authors classify the free-living protozoa into broad morphological groups, they are simultaneously allocated to broad functional groups. The three broadest morphological–functional groups are the ameboid, the flagellated, and the ciliated protozoa, and each has its own strengths as a phagotroph. Representatives of all three may feed on the same type of microbes in the same place (e.g., in an aquatic sediment), but they will differ in the mechanics and efficiency of capture of any particular food particle. A filter-feeding flagellate will have a relatively large filter area, a high volume-specific clearance and competitive superiority over filter-feeding ciliates when grazing on planktonic bacteria. A helioflagellate and a suctorian ciliate will both practice diffusion feeding, but the former will be adapted for snaring bacteria, whereas a diffusion-feeding suctorian will specialize in trapping flagellates and ciliates.

Many protozoa are microaerophilic: They seek out habitats with a low level of dissolved oxygen that is just sufficient to drive their aerobic respiration and low enough to exclude metazoan competitors and predators. Microaerobic habitats are common in aquatic sediments and in oceanic oxygen minimum zones. These are zones in which the raw materials for microbial growth arrive from opposite directions (e.g., where oxygen and light arriving from above meet carbon dioxide and sulfide from below) and in which there is therefore an elevated abundance of microbial food. Therefore, microaerophily is an adaptive behavior: It brings protozoa into contact with high abundances of microbial food. It also stimulates the growth of nutritional symbionts such as sulfide-oxidizing bacteria and endosymbiotic algae. Many microaerophilic protozoa are also temporary anaerobes, but unlike the true anaerobes that live permanently in the absence of oxygen, their metabolism is fundamentally aerobic. The true anaerobes – those that complete their entire life cycle in the absence of oxygen – live principally in aquatic sediments. There are many species, but none is ever abundant. Most use hydrogen-evolving fermentations for energy generation, and the hydrogen is used by anaerobic bacteria, especially endosymbiotic methanogens. Thus, methane is released from these protozoan consortia. The anaerobic protozoa are probably the only phagotrophic organisms capable of living permanently in the absence of dissolved oxygen (Fenchel and Finlay, 1995).

The real diversity of symbiotic associations involving protozoa is poorly known. In some cases, complex interactive behaviors have evolved between the partners. In the marine, sand-dwelling ciliate *Kentrophoros*, the entire dorsal surface of the ciliate is a coat of sulfide-oxidizing bacteria that can grow only in the narrow layer within the sediment where oxygen and sulfide overlap. The ciliate host's innate microaerobic behavior enables it to seek out the habitat that the symbiotic bacteria need for growth. The ciliate then invaginates its dorsal surface and digests the bacteria because it does not have a mouth, and this is its only source of nutrition.

It is clear that many of these symbiotic consortia involving protozoa represent tightly integrated functional units; indeed, the symbionts may be almost as deeply embedded functionally in the consortium as the protozoon's other organelles. Two points must be noted. First, it is the combined phenotype of the consortium, rather than that of any individual consortium partners, on which natural selection will operate, and there are examples of how the fitness of a protozoon in a particular habitat can be improved by the acquisition of endosymbionts (e.g., the algal symbionts of ciliates living in the metalimnia of freshwater lakes). Second, the biodiversity of protozoa, when quantified simply in terms of protozoan species richness, will fail to take account of the large supplementary microbial diversity with which the protozoa are necessarily associated.

Therefore, the diversity of free-living protozoa may be classified into broad morphological–functional groups: ameboid, flagellated, and ciliated protozoa. Almost any free-living protozoon can be placed without difficulty in one of these groups. It must be stressed, however, that these groups are not concordant with any system of classification of protozoa published in recent years; nor are they in most cases aligned with the independent lineages that are emerging in the molecular phylogenies (e.g., those based on sequence variation in ribosomal RNAs) which reflect the main episodes in the history of eukaryotic evolution. Heterotrophic flagellate groups, such as the diplomonads and trichomonads, appear in the early emerging lineages, but other flagellates (e.g., choanoflagellates, chrysomonads, dinoflagellates, and haptomonads) are classified within recently diverging lineages. The ameboid protozoa too are scattered across many lineages. The naked amebae without mitochondria (the pelobionts) diverge early and close to the diplomonads, whereas the vahlkampfiids, slime molds, and various other groups of naked and testate amebae appear in other independent lineages that are evolutionarily quite distant from each other. The morphological–functional group of the ciliates is the only one which remains intact, as a monophyletic group, in current molecular phylogenies.

The process of distilling the vast quantity of molecular information generated in recent years has generated some entirely new phyletic assemblages, including the stramenopiles, a group containing organisms as morphologically and functionally dissimilar as chrysomonad flagellates and diatoms, and the alveolates – a group that embraces the dinoflagellates, the ciliates, and a large group of exclusively intracellular parasites

(the apicomplexans) (Adl *et al.*, 2005). In the next section, the authors focus on the broad morphological–functional groups of free-living protozoa, define what is meant by species, and quantify the species within these groups.

The Nature of Protozoan Species

There is no universal agreement on what constitutes a protozoan species. The most widely used concept is the morphospecies because it is relatively easy to discriminate a great diversity of protozoa using body form alone. This concept is especially useful because morphology is closely related to the ecological function. In many free-living protozoa, the structure of the feeding apparatus and the size and shape of the cell determine how the protozoon functions in the natural environment; therefore, the cell form largely determines the ecological niche that the protozoon occupies. Discriminating species on the basis of form might then be equivalent to discriminating them according to the ecological niches they occupy. We could say that if a protozoon looks the same in different places, then it is the same in different places. However, there are problems with such a simple concept. It is known that a morphospecies can be composed of ecologically distinct populations, such as those adapted for maximum growth rate at different temperatures. Therefore, although it is possible to collect apparently identical representatives of a morphospecies in different corners of the world, and one might be encouraged by the belief that they were filling exactly the same niche in these different places, in reality they could be phenotypically quite different. The same problem applies with respect to genetic differences. It is possible to find genetic differences, especially sequence differences in ribosomal RNAs, in morphologically identical isolates collected from different places. The significance of these genetic differences is not known; nor is it known if genetic and phenotypic divergence are correlated. It appears, however, that there is no correlation between genetic divergence and geographic distance.

Protozoa spend most of their time as asexual organisms, but some do reproduce sexually, if only periodically, and in some well-studied cases (e.g., the ciliate genera *Tetrahymena* and *Paramecium*) morphologically indistinguishable biological species (reproductively isolated gene pools, also known as sibling species) have been studied thoroughly. Different sibling species within a morphospecies may be genetically identical to each other or extremely divergent (at least with respect to ribosomal RNAs), and there is no apparent correlation between genetic isolation and genetic divergence. Again, there is no good evidence for biogeography of protozoan sibling species, and members of the same sibling species can be found on different continents. It is possible that sibling species carry unique phenotypic traits that equip each species for a particular niche, but this has not been demonstrated. The real problem in using a biological species concept for protozoa is that it is not practical for all but a few easily cultivated protozoa. The majority of protozoan species have never been cultivated, and neither the frequency nor the character of their sexual behavior (if any) are known or are ever likely to be known. Most people who study protozoa in the natural environment use the morphospecies concept because it is practical, it embodies the close link between form and function, and it is the morphospecies that fills the niche, if not always the minutest niches, that can be discerned by the human observer. The morphospecies may contain much genetic variation and be capable of expressing a wealth of phenotypic variation, but it is the best tool that is available for ordering the diversity that lies within the protozoa.

One particular problem that is particularly acute when dealing with organisms the size of protozoa is that the quality of one's perception of morphology is very definitely a function of the tools that are available (Mitchell and Meisterfeld, 2005). Thus, better tools (e.g., quality microscopes) enable the discrimination of finer detail. The logical extension of this is that even individual protozoa in a population might be discriminated one from the other and could therefore be referred to as separate morphospecies. Modern optics have undoubtedly lent impetus to the business of describing many new species, some of which are undoubtedly legitimate, whereas many have been established on morphological criteria that are trivial or have no conceivable functional significance.

An additional problem with the morphospecies is that it is often difficult to decide exactly where a species begins and ends. Although most observations are made of the vast number of individuals that cluster around the central tendency of any population, the individuals at the tail ends of the distributions abut and often overlap those of other species. Morphological variation then appears to be continuous across many species. Examples of this have been described: In some of the large spongiose spumellarian radiolarians, the major features (e.g., skeletal and cytoplasmic morphology) that are used to discriminate species intergrade to such an extent that it is not possible to unambiguously ascribe individual radiolarians to nominal species.

However, the evolutionary process in radiolarians and other protozoa probably works in the same way as it does for other organisms such that it maintains phenotypically discrete species in niche space that is essentially continuously variable. The phenotypic traits that sustain these discrete species in protozoa are presumably quite diverse in character and possibly not totally comprehensible. Among the more accessible of these traits are the morphological characters that the authors use to separate species. However, the likely limitation of these is that they enable us to perceive only the fairly coarsely resolved, sometimes overlapping entities that we call morphospecies.

Protozoa and Ecosystem Function

Protozoa are abundant. One gram of soil typically contains 103–107 naked amebae, 105 planktonic foraminiferans can often exist beneath 1 m^2 of oceanic water, and almost every milliliter of fresh water or sea-water on the planet supports at least 100 heterotrophic flagellates. When expressed in global terms, these numbers are very large, and an inevitable consequence of the persistence of such a large number of very small organisms is that migration rates will be relatively high. It follows that rates of speciation and extinction must be low,

as will the consequent global number of species. It also follows that protozoa are unlikely to have biogeographies, and endemic species probably do not exist. The authors might expect that the local diversity of protozoa would account for a significant proportion of global diversity, even if at any moment in time much of this diversity is represented by rare or inactive individuals (e.g., cysts awaiting the arrival of suitable conditions). There is indeed good evidence for the global distribution of protozoan species, including the morphologically distinctive flagellate *Rhynchomonas nasuta* that has been found in most aquatic and terrestrial environments worldwide; marine foraminiferans and ciliates found living in slightly salty water of desert oases, hundreds of kilometers from marine coasts; the same radiolarian species living in high northern and southern oceanic latitudes; the same pond-dwelling ciliates living in Australia and northern Europe; and the cosmopolitan distribution of the same species of agglutinated foraminiferans in the deep-sea benthos. In general, protozoan morphospecies are ubiquitous and apparently cosmopolitan if the habitats to which they are adapted are distributed in different parts of the world (Finlay, 1998, 2002). In accordance with this, the global number of protozoan species is indeed relatively modest (**Table 1**), and the number of species that can be retrieved from a local area (e.g., a pond), in both active form and from a passive state is a significant proportion (usually at least 10% for various morphological-functional groups) of the global total. This fact may not be obvious from short-term ecological sampling programs because only a limited number of microbial niches are available at any moment in time.

Protozoa and other microorganisms have other special properties. Microbial activities interact strongly with physical and chemical factors in the natural aquatic environment (e.g., light transmission or the concentrations of essential nutrients) to create a continuous turnover of microbial niches. These niches are quickly filled from the locally available diversity of rare and dormant microbes, and the activities of the latter create further reciprocal interactions. Therefore, the diversity of active protozoan species in a pond, at any moment in time, is the result of preceding reciprocal interactions involving many biological and nonbiological factors, and the biodiversity of protozoa and other microbes is an integral part of ecosystem functions such as carbon fixation and nutrient cycling (Finlay *et al.*, 1997).

Biological Diversity and Global Species Richness

Ameboid Protozoa

Slime molds may be regarded as protozoa because they spend most of their active lives as amebae or as naked ameboid (and often macroscopic) plasmodia (**Figure 1**). They have been known by many names, including mycetozoa (fungus animals), because although they never develop hyphae they produce fruiting bodies (sporangia) supported on cellulose-rich stalks. There are two large groups – the cellular slime molds (the dictyostelids, e.g., *Dictyostelium*) and the acellular slime molds (myxomycetes, e.g., *Physarum*). Two smaller groups are the acrasids (e.g., *Acrasis*) and the protostelids

Table 1 Estimates of global species richness of extant free-living protozoa[a]

			Marine	*Nonmarine*	*Total*
Ameboid protozoa	Slime molds	Dictyostelids	–	60	60
		Myxomycetes	–	550	550
	Rhizopod amebae	Naked	180	220	400
		Testate	–	200	200
	Foraminiferans	Planktonic	40	–	40
		Benthic, inshore	4000[b]	–	4000
		Benthic, deep sea	250[b]	–	250
	Actinopod amebae	Acantharians	150	–	150
		Radiolarians, solitary	750	–	750
		Radiolarians, colonial	50	–	50
		Heliozonas	–	120	120
Flagellated protozoa	Excluding heterotrophic dinoflagellates[c]	Marine plankton	420	–	420
		Marine benthos	330	–	330
		Freshwater and soil	–	350	350
	Heterotrophic dinoflagellates		900	110	1010[d]
	Other mixotrophic flagellates		–	–	150[e]
Ciliated protozoa			1400	1660	3060
Total					11890

[a]Compiled from numerous published and unpublished sources. The more problematic estimates are highlighted.

[b]There is considerable uncertainty attached to these estimates. The figure of 4000 is generally accepted as a working figure for extant species but is probably inflated by synonyms, especially those of shallow-water benthic species. There is no firm information for species richness of deep-sea benthic foraminiferans (Gooday *et al.*, 1998). The estimate of 250 is approximately double the typical figure for local richness of species, most of which may be cosmopolitan, but there are probably many undescribed deep-sea soft-shelled foraminiferans that have in the past been ignored by geologists.

[c]The total given here for heterotrophic flagellates is 1100 species. This includes some synonyms and mixotrophs. It is believed by some that the real global total is closer to 3000 species.

[d]Assuming there are 1800 marine and 220 freshwater species, and that 50% of these are heterotrophs.

[e]This estimate is simply derived by doubling some recent estimates. Note that these species may be only temporarily phagotrophic.

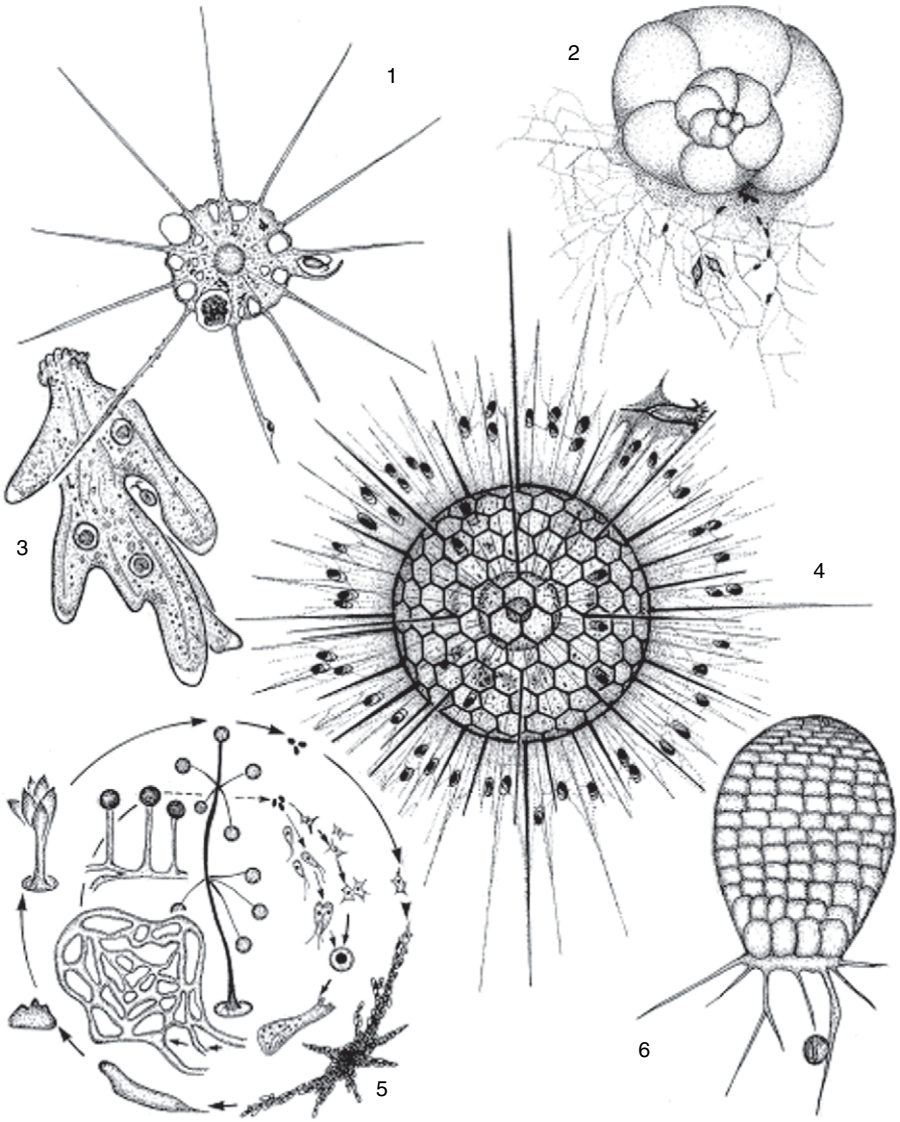


Figure 1 A selection from the variety of form and function in ameboid protozoa and slime molds. (1) an actinophryid heliozoon (*Actinophrys*; diameter of about 0.1 mm) ingesting a flagellate; (2) a benthic foraminiferan (*Rotalia*) trapping diatoms and bacteria in its reticulopodial net; (3) a naked ameba (Ameba) using pseudopodia to trap a flagellate; (4) a polycystine radiolarian (*Heliosphaera*; spherical body ~0.3 mm) with symbiotic dinoflagellates and an entrapped tintinnid ciliate; (5) polymorphic life cycles of dictyostelid slime molds (e.g., *Polysphondylium*; outer circle) and myxomycete slime molds (e.g., *Physarum*; inner circle); (6) testate ameba (*Assulina*; ~0.08 mm) with its prey (an algal cell) caught on a sticky filopod.

(e.g., *Cavostelium*). The dictyostelids are typically phagotrophic amebae, but when starved they aggregate, form a migrating slug, and then produce a fruiting body from which spores are released. These later germinate to produce amebae. In the myxomycetes, individual amebae (sometimes thousands of them) with or without flagella coalesce to form a distinctive (often brightly colored) multinucleate plasmodium which gives rise to fruiting bodies. Slime molds are common in damp forest soils, tree bark, and dead or dying wood, and abundance and local species richness are greater in deciduous than in coniferous forests. There is extreme patchiness in the distribution of dictyostelid species in forest soils, and coexistence of multiple species in the same small soil samples is rare. Most myxomycete species are believed to be cosmopolitan. The phagotrophic ameboid stages of slime molds may be quantitatively important grazers of bacteria, fungi, and the other primary decomposers of organic matter in soil. Other ecological interactions of slime molds may be at least as complex as their life cycles (e.g., the migrating slugs of dictyostelids appear capable of repelling grazing nematodes).

Rhizopod amebae use pseudopodia for locomotion and feeding. There are two large groups (**Figure 1**) – the "naked amebae" (e.g., *Acanthamoeba*, *Vannella*, *Ameba*, and *Vampyrella*) and the shelled "testate amebae" (e.g., *Arcella*, *Nebela*, and *Euglypha*). Most rhizopod amebae feed nonselectively by engulfing diatoms and other algae, unicellular and filamentous

cyanobacteria, detritus, and bacteria. However, there are notable variants: *Vampyrella* dissolves a hole in the cell wall of a green alga or desmid and then enters through the hole to digest the cytoplasm of the prey. Some are opportunistic pathogens of man (e.g., in the genera *Acanthamoeba* and *Naegleria*).

Naked amebae occur in the plankton, but these species are usually very small ($<10\ \mu m$) and typically associated with suspended flocs (e.g., marine snow). However, it is likely that all rhizopod amebae need to be attached to a surface if they are to feed, and much larger numbers are found in aquatic sediments. Amebae are particularly abundant in soils (103–107 per gram dry weight), where together with heterotrophic flagellates they probably control bacterial abundance. A few testate amebae can be planktonic (e.g., *Difflugia*), but most live in soils and in *Sphagnum* bogs.The prey size of testates must be limited by the shell aperture size.

Naked and testate amebae offer different challenges when attempting to identify species. The naked forms are distinguished by the forms of their bodies and pseudopods in locomotion, but many species constantly change shape and the absence of fixed characters often makes it difficult to identify them without resorting to an analysis of ultrastructure. The testates construct tests which have species-specific architectures, but these are often modified according to the nutritional status and local environmental factors, such that the variety of tests described is extraordinarily large and the task of deciding where one nominal species begins and the other ends is at least as difficult as it is for the naked amebae. One practical solution to the problem of identifying living naked amebae in natural samples is to allocate them to one of the several broad morphotypes, e.g., those that extend lobose pseudopodia, slug-like forms with eruptive or noneruptive locomotion, and fan-shaped forms. No clear relationship has been established between morphology and ecological function in naked amebae, although it is likely that the limax (slug-like) amebae are predominantly benthic species.

Some naked amebae (especially those that form extensive anastomosing plasmodial networks) resemble hybrids of shell-free foraminiferans and slime molds. Some can reach several centimeters in diameter. They are typically very slow moving and apparently perfectly adapted for grazing extensive areas of soil surface (*Leptomyxa*) or diatom mats in the marine benthos (*Corallomyxa*). Some free-living naked amebae are anaerobic and do not have mitochondria. As a group, these are referred to as pelobionts, rhizomastigids, archamebae, or karyoblasteans (e.g., *Mastigamoeba*, *Mastigella*, and *Pelomyxa*). They are relatively common in organically enriched anoxic freshwater and marine sediments. The very large (up to several millimeters in diameter) *Pelomyxa* supports several types of endosymbiotic bacteria, including methanogens, and probably lives exclusively in freshwater sediments. It appears to ingest any organic particles with which it makes contact. It can be locally abundant and the dominant grazer in some anaerobic lake sediments. Most species of naked amebae have been described from fresh waters and soil, and most marine species have been collected from the inshore areas. It is not clear to what extent freshwater species can also live in seawater. Some genera (e.g., *Mayorella* and *Acanthamoeba*) are well represented in both environments.

Foraminiferans live in marine or brackish water environments, and most have calcareous shells (**Figure 1**). Cytoplasm emerges through the aperture and pores in the shell to form a repeatedly branching and anastomosing reticulopodial net in which prey are trapped (Lee *et al.*, 2000). The spinose planktonic species (e.g., *Globigerinoides ruber*) have calcareous spines, and familiar Cretaceous chalk deposits (e.g., the white cliffs of Dover, UK) consist largely of the sedimented shells of these species together with the coccoliths of haptomonad flagellates. Some naked amebae that produce similar pseudopodial networks live in fresh waters (e.g., *Biomyxa*) or soil (e.g., *Reticulomyxa*).

The shells of planktonic foraminiferans are generally 0.1–1 mm in diameter, but the pseudopodial network can reach 10 mm and the spines may increase the size to more than 20 mm, making them larger than many marine metazoans. Planktonic species are open-water organisms. They are most abundant in tropical and subtropical waters, in which there is typically one per liter (or $\sim 10\ 5\ m^{-2}$, depth integrated to 150 m). Each species tolerates only small variations in temperature and salinity, and distinct assemblages of species are associated with five biogeographical provinces – the surface water masses of equivalent temperature on either side of the equator (polar, subpolar, temperate, subtropical, and tropical). However, species in these assemblages are not isolated from each other: Polar species have been recorded in upwelling zones at 10° N, and species characteristic of high latitudes are also found in the tropics, but in very deep and cool waters.

A typical planktonic foraminiferan produces a mass of reticulopodia (the "halo") that radiates outward from the shell. This is used to snare prey and to support algal symbionts. Spinose species appear to prefer zooplankton prey, and nonspinose species (e.g., *Globorotalia menardii*) prefer phytoplankton. These preferences may affect the local distribution of species: In the Red Sea north of 20° N, the dominating spinose species appear to reflect the greater local abundance of zooplankton.

Many planktonic foraminiferans, especially the spinose species (e.g., *G. ruber*), contain algal symbionts that are carried along the spines by cytoplasmic streaming. There are probably only two types of symbionts in planktonic species: the dinoflagellate *Gymnodinium beii*, which is morphologically identical in many host species, and a small yellow–green (chrysomonad-like) symbiont. The symbionts are protected from digestion by the host. The carnivorous foraminiferan *Hastigerina pelagica* envelopes itself in a cytoplasmic bubble capsule which is probably a flotation device but is also used for trapping and digesting prey. *H. pelagica* does not have symbionts but commensal dino flagellates (*Pyrocystis* spp.) that live in the capsule. The role of symbionts and commensals is unclear. High rates of photosynthesis are maintained within the consortium, the symbionts probably transfer significant amounts of fixed carbon to the host, their photosynthetic activity may be necessary for calcification of the host's shell, and the symbionts may use the host's (nitrogenous) waste.

Benthic foraminifera have a greater diversity of symbiotic partners, including diatoms (*Nitzschia*), dinoflagellates (*Symbiodinium microadriaticum*), red algae, and chlorophytes (*Chlamydomonas*). Some also sequester chloroplasts from

ingested algae. Although many benthic foraminiferan species have been described, information on their life cycles and ecology is available for only about 20 shallow-water species (e.g., *Allogromia*, *Rosalina*, *Spiroloculina*, and *Sorites*). These typically drag themselves slowly over surfaces, continually extending reticulopodia to trap prey (diatoms, algae, and bacteria). The abundance of benthic species is correlated with local biological productivity. In the fertile Mississippi Delta, numbers may reach several thousands in an area of 10 cm^2. Close to sewage outfalls, they are typically very abundant but low in species diversity. The spatial distribution of all benthic species is very patchy. This may be related to asexual reproduction (known only for benthic species) and the consequential limited range of migration of offspring from the parent. Benthic foraminiferans become relatively more important with increasing water depth, possibly due to a competitive advantage they have over metazoans in their ability to rapidly exploit seasonally deposited phytodetritus. In oxygen minimum zones along the continental margins of tropical oceans, they can, because of their relatively small size, tolerate low oxygen conditions better than can metazoans. Also, certain taxa (e.g., *Bolivina* and *Uvigerina*) seem to be adapted for life in oxygen-depleted waters.

Foraminiferans in the deep-sea benthos are often abundant (typically hundreds per 10 cm^2) and diverse (species richness may exceed that of nematodes), and they are now regarded as major components of the fauna in this zone (Gooday *et al.*, 1998). Because they consume bacteria and detritus, they probably act as a link in the food chain to higher trophic levels in the benthos. Very small species (30–63 μm) are common, but the dominant fauna >500 μm in central oceanic regions (and particularly at depths greater than the carbonate dissolution depth of ~4000 m) are large agglutinated foraminiferans, especially the komokiaceans, which have complex anastomosing networks of tubules. In the benthos underlying the oligotrophic waters of the central North Pacific, the biovolume of komokiacean tests may greatly exceed the biovolume of metazoans. However, the amount of protozoan biomass in komokiacean tests is relatively small, and it may be even less than that of the metazoans that also occupy the test. *Edgertonia floccula* consists of a network of branching agglutinated chambers buried in a large (2–8 mm) inorganic mudball that it shares with cohabiting metazoans. This foraminiferan is one of the most abundant meiobenthic animals in some areas of the abyssal northeast Atlantic. Other deep-sea agglutinated foraminifera (e.g., *Bathysiphon* and *Hyperammina*) may even reach several centimeters in length, and some species stand vertically in the sediment and project into the overlying water, in which they presumably feed. Many of the common deep-sea foraminiferan species are probably cosmopolitan in their distribution.

The xenophyophoreans are another deep-sea group about which relatively little is known. They live in agglutinated, often fragile tests (typically 1–10 cm in length) and they too probably feed on sedimented detritus. Abundances of hundreds per square meter have been reported. The multinucleate protoplasm is enclosed within a labyrinthine organic tube encrusted with foreign particles. The organism also contains numerous crystals of barium sulfate, which have no obvious function. Growth is episodic, with intervening periods of months of inactivity. They have no confirmed fossil record, although they often use the sedimented shells of the planktonic foraminiferan *Globigerina* to strengthen their tests (as in *Homogammina maculosa*).

Foraminiferan fossils are useful for biostratigraphy because (1) their sedimented remains are abundant (e.g., in the "*Globigerina* ooze" which covers much of the abyssal plain beyond the continental shelves), (2) individual shells are identifiable to species level, and (3) their remains allow determination of sea surface paleotemperature. Seawater and calcite differ in their $^{18}O/^{16}O$ ratio. This difference increases with temperature, so the ratio is used to estimate the water temperature at the time the calcite was deposited in the fossil foraminiferan shell. This estimate can then be compared with the established correlation between the temperature of water mass and the extant species it typically supports. Much of this work is supported financially by the oil exploration industry.

Actinopod amebae all have axopods, which are stiffened ray-like pseudopodia that radiate from the cell body (**Figure 1**). They also have a peripheral network of filopodia, which are long, thin, sticky pseudopodia used for trapping food organisms that make chance contact. There are three principal groups: (1) radiolarians, consisting of the polycystines with perforated spiny lattice skeletons of silica and the phaeodareans with skeletons of hollow tubes and spines; (2) acantharians with skeletal spines of strontium sulfate radiating from the center of the organism; and (3) heliozoans, which often lack an internal mineral skeleton (although some species produce skeletal spicules which are either organic or siliceous) and superficially resemble polycystine radiolarians. All actinopod protozoa, apart from a large number of heliozoans, are exclusively planktonic in deep oceanic water.

Radiolarians (**Figure 1**) and acantharians range in size from ~30 μm to several millimeters in diameter. Colonial radiolarians (e.g., *Collozoum longiforme*), which occur as gelatinous cylinders with diameters of several centimeters, can reach 3 m in length. Typical food items include diatoms, tintinnids and other ciliates, crustacean larvae, and other zooplankton, all of which are trapped in the peripheral network of axopods and filopodia. The smallest species may eat bacteria (Anderson, 1983).

Dinoflagellate (*Symbiodinium*), chrysomonad, and prasinomonad symbionts are associated with many polycystine radiolaria. They probably contribute to the host carbon metabolism because the consortium can last many weeks without exogenous food. The symbionts are carried about in the host's cytoplasmic streaming, gathering in the peripheral filopodia at dawn and withdrawing inside the host shell at sunset. The phaeodareans that live near the ocean floor do not have algal symbionts. Radiolarians are found at all water depths down to ~8000 m, although they are most abundant and diverse between 200 and 2000 m (Anderson, 1983). No benthic species are known. Like planktonic foraminiferans, each radiolarian species inhabits an oceanic water mass that lies within a particular temperature range. Moreover, because of the immense depths to which living radiolarians sink, those species living in surface waters at high latitudes are also found in the tropics but in very deep and cool waters. As a consequence, the species assemblages found in surface waters give a misleading impression of radiolarian biogeography. It is believed that

radiolarians may be transported in very deep oceanic waters between high northern and high southern latitudes.

Radiolarian abundance and productivity increase in warmer waters. Abundance may be almost nil in the surface waters of the Antarctic, up to 50 m^{-3} in the Caribbean and Gulf Stream, and up to several thousand per m^3 in the Gulf of Mexico. Colonial radiolarians are much less abundant. In the Sargasso they are considered abundant if represented by more than one colony per 100 m^3. Acantharians can reach relatively high abundances (up to 25–35 l^{-1}) and they are often the most abundant of the planktonic shelled amebae, but relatively little is known about living organisms because as soon as they are collected the fragile cell membrane breaks and the strontium sulfate skeleton dissolves. Radiolarians are important stratigraphic tools, and skeletal morphology is the main characteristic used to identify species; hence, much information is available on the diversity of radiolarian skeletons.

The heliozoans are predominantly a freshwater group. They resemble functionally the radiolarians, especially with respect to their diffusion feeding. They span a wide size range. Some (e.g., *Actinosphaerium*) can be >1 mm, but most species are ~ 40 μm (*Actinophrys*). Depending on their size, different species feed with little apparent selectivity on algae, flagellates, ciliates, and rotifers. Small heliozoan species account for most of the biomass of ameboid protozoa in the freshwater plankton. Most, however, are attached to or loosely associated with sediment and other submerged surfaces (Page and Siemensma, 1991).

Flagellated Protozoa

There is a little consensus on how to classify the flagellates, and they are here divided into broad functional groups (**Figure 2**). Heterotrophic flagellates are fundamentally important because they are abundant (there are seldom less than 1000 ml^{-1}, even in the plankton) and because their grazing activities are largely responsible for controlling the abundance of bacteria in aquatic environments. In some taxonomic groups, all species are exclusively heterotrophic (e.g., choanoflagellates and bodonids); others contain many mixotrophs (e.g., the euglenids and chrysomonads), whereas the haptomonads and cryptomonads are dominated by phototrophs and only a minority are capable of phagotrophy. In the past 25 years, a large diversity of heterotrophic flagellates has been discovered. Most of these are choanoflagellates, chrysomonads, euglenids, or bodonids. Some of the more easily recognized species (e.g., *R. nasuta*) have been recorded from a wide range of habitat types in marine, freshwater, and terrestrial environments (Patterson and Larsen, 1991; Lee and Patterson, 1998).

Many heterotrophic flagellates are anaerobes, including intestinal parasites of man (e.g., *Giardia intestinalis*), but free-living diplomonads (*Hexamita* and *Trepomonas*) and *Retortamonas* are relatively common bacteria feeders in organically enriched anoxic waters and sediments. Some anaerobic flagellates have hydrogenosomes, the anaerobic derivatives of mitochondria. Almost all of these are endosymbionts or internal parasites of animals. They are also known as parabasalians and include many of medical or economic importance, e.g., the human parasite *Trichomonas vaginalis* and *Trichomitopsis*, which degrades cellulose in the hindgut of termites and other wood-eating insects and also supports endosymbiotic methane-producing bacteria (termites may be responsible for a globally significant flux of methane to the atmosphere) (Fenchel and Finlay, 1995).

Most choanoflagellates are small (<10 μm), with a feeding filter that forms a collar around the single anterior flagellum. The flagellum creates the water current that flows through the filter. Choanoflagellates are exclusively phagotrophic and either solitary (e.g., *Monosiga*) or colonial (e.g., *Sphaeroeca*). They are important grazers of the smallest suspended bacteria, especially in the marine plankton (e.g., *Salpingoeca*, *Stephanoeca*, and *Parvicorbicula*).

Euglenids usually have two flagella which emerge from a small anterior pocket. This is a large group of planktonic and benthic protozoa, common in both fresh and marine waters. Species associated with surfaces typically keep one trailing flagellum in contact with the substrate, whereas the other pulls the cell forward. Many species are green and phototrophic (e.g., *Euglena*), but there is a great diversity of phagotrophs (e.g., *Astasia*, *Petalomonas*, *Entosiphon*, and *Notosolenus*). These are especially common in marine shallow-water sediments, in which they graze on bacteria. All bodonids are small flagellates with characteristic heterodynamic flagella (e.g., *Bodo* and *Rhynchomonas*). They are almost always found in organically polluted (even anoxic) waters or soils that are rich in bacteria, on which they feed. Many closely related ("kinetoplastid") flagellates are parasites: *Ichthyobodo* causes costiasis in freshwater fish and the "trypanosomatids" are important parasites of man (*Trypanosoma brucei* causes sleeping sickness). In heterokont flagellates, each organism has two flagella – one hairy and one smooth, with the former creating the water current that brings mainly bacterial food to the (filter-feeding) flagellate. They include the bicosoecids that live in a lorica (*Bicosoeca*) or are naked (*Cafeteria* and *Pseudobodo*), that may be stalked or colonial, and which resemble the mixotrophic loricate chrysomonads (e.g., *Dinobryon*).

In the phagotrophic chrysomonads (e.g., *Paraphysomonas* and *Spumella*), the chloroplast is essentially vestigial (and the cell is colorless). Functional types range from heterotrophs (e.g., *Paraphysomonas*) to mixotrophs (e.g., *Dinobryon* and *Uroglena*) and those that are predominantly phototrophs (e.g., *Synura* and *Mallomonas*). Chrysomonads are the most abundant heterotrophic flagellates in the freshwater plankton. The helioflagellates (also referred to as "pedinellid stramenopiles," e.g., *Actinomonas*, *Pteridomonas*, and *Ciliophrys*) are filter or diffusion feeders that superficially resemble heliozoa or choanoflagellates.

The haptomonads are typically photosynthetic, biflagellated organisms. They include the coccolithophorids and the prymnesiophytes, and some (e.g., *Emiliania*) are often abundant in the marine euphotic zone. They have an additional flagellar appendage known as the haptonema. In the marine *Chrysochromulina*, this is used in food uptake, but in *Prymnesium*, which has a very short haptonema, ingestion of food (dinoflagellates, chrysomonads, and green algae) is by means of a posterior pseudopodium. Prey organisms may be immobilized by toxins prior to ingestion (toxic bloom-forming species of *Prymnesium* are responsible for fish kills). The scale

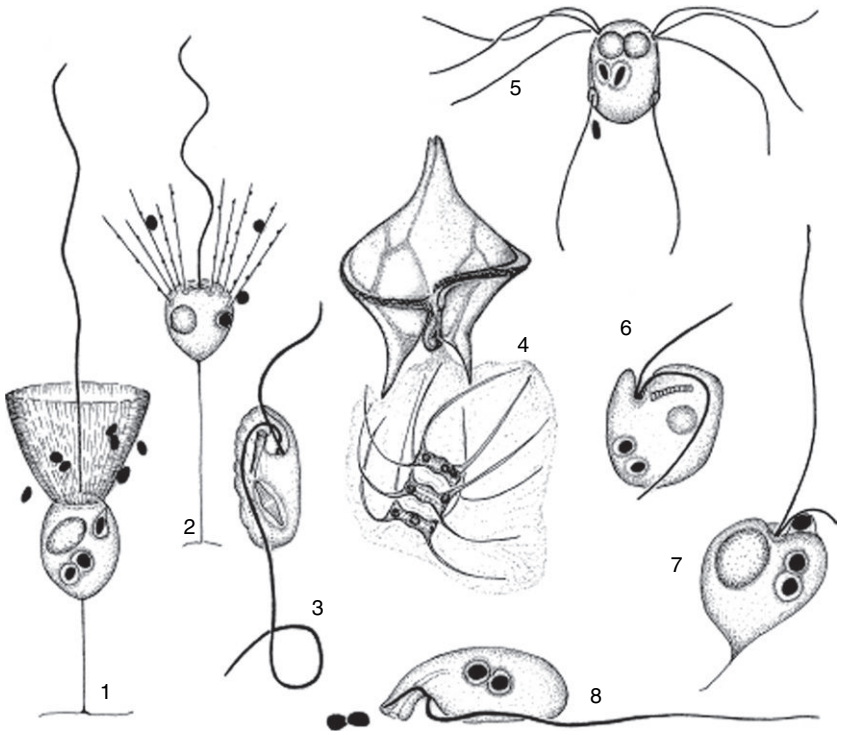


Figure 2 A selection from the variety of form and function in flagellated protozoa. (1) a choanoflagellate with bacterial food trapped on the external surface of the collar filter (*Monosiga*; body + collar ~0.015 mm); (2) a bacteria-feeding helioflagellate (*Pteridomonas*; body 0.007 mm); (3) a phagotrophic euglenid (*Dinema*; ~0.03 mm) with an ingested diatom; (4) a marine heterotrophic dinoflagellate (*Protoperidinium*; ~0.06 mm) with a diatom trapped in its feeding veil; (5) a free-living, bacteria-feeding, anaerobic diplomonad (*Hexamita*; body ~0.008 mm); (6) a small heterotrophic (bacteria-feeding) cryptomonad flagellate (*Goniomonas*; ~0.008 mm); (7) a heterotrophic, typically bacteria-feeding, chrysomonad flagellate (*Spumella*: body ~0.008 mm); (8) a bacteria-feeding bodonid (*Rhynchomonas*; body ~0.005 mm).

and importance of phagotrophy in freshwater haptomonads is unclear.

Most cryptomonads are yellow–brown and phototrophic, with two flagella of similar length (e.g., *Cryptomonas*). They are especially common in fresh waters. Heterotrophic species either lack plastids (*Goniomonas*) or have a plastid that is devoid of photosynthetic pigments (*Chilomonas*). Colorless species can be relatively abundant, especially when associated with benthic detritus.

The heteromitids and cyathobodonids represent a varied assemblage of small, poorly studied flagellates (e.g., *Apusomonas*, *Heteromita*, *Cercomonas*, *Cyathobodo*, and *Kathablepharis*), some of which (e.g., *Cercomonas*) produce pseudopodia to ingest bacteria. Others secrete a stalk (*Cyathobodo*) to attach the cell to the substrate. Many are common in soil. Others feed mainly on bacteria in the freshwater plankton (e.g., *Kathablepharis*, which also feeds on small algae) and on sediment surfaces. The percolozoans are a diverse collection of bacteria-feeding organisms, including ameboflagellates (e.g., *Naegleria*), the filter-feeding flagellate *Percolomonas*, and anaerobic flagellates with hydrogenosomes and endosymbiotic methanogens (*Psalteriomonas*).

Dinoflagellates are best known as a large group of photosynthetic flagellates and for the ability of some species to cause "toxic blooms" and "red tides" (*Protogonyaulax* causes paralytic shellfish poisoning). However, about half of all known marine species lack chloroplasts. Some of these (in the genera *Gymnodinium* and *Amphidinium*) feed on cryptomonads and sequester their chloroplasts, thus transforming themselves into mixotrophs. Moreover, many typically photosynthetic species can become mixotrophs through their ability to ingest particulate food. The nonphotosynthetic and mixotrophic species are referred to collectively as heterotrophic dinoflagellates. These can be quantitatively important in marine food webs, especially as consumers of diatoms. Some species are benthic, but little is known about these. The diversity of heterotrophic species in fresh waters is much less, and it is known mainly from the genera *Katodinium*, *Peridinium*, *Gymnodinium*, and *Ceratium*. This may be due in part to the relative rarity in fresh waters of large, chain-forming diatoms – a typical food item of many marine dinoflagellates. These large food items are trapped and digested externally in a pseudopodial feeding veil or pallium. This explains why it has often been difficult to recognize ingested organisms within dinoflagellates. Other heterotrophic dinoflagellates use a feeding tube known as a peduncle. The common freshwater dinoflagellate *Peridiniopsis berolinensis* uses such a tube to ingest the fluid contents from injured and dying protists and small metazoans.

The number of marine dinoflagellate species with a recorded capacity for phagotrophy is increasing rapidly. These include thecate species (e.g., the peridinioids) previously considered to be incapable of phagotrophy. It is widely believed that most dinoflagellates – at least in marine environments – may be capable of phagotrophy. Heterotrophic

dinoflagellates (e.g., *Oblea*, *Polykrikos*, *Heterocapsa*, and *Protoperidinium*) can dominate the protozoan biomass in coastal and oceanic waters, in which they feed on bacteria, flagellates, diatoms, other dinoflagellates, and ciliates. Their biomass can be approximately the same as that of ciliates, and the two groups may compete with each other for food (although the size range of dinoflagellates is slightly greater – from 3 or 4 μm in *Gymnodinium simplex* to 2 mm in *Noctiluca*, with the majority in the size range 20–200 μm). Also, most planktonic ciliates are incapable of ingesting the large diatoms that are consumed by dinoflagellates.

Almost all stony and stinging corals (Cnidaria) in shallow tropical waters harbor photosynthetic dinoflagellates (especially *Symbiodinium*) as symbionts. These consume carbon dioxide and bicarbonate. This promotes calcium deposition in the external skeletons of their hosts and may enhance the rate with which coral reefs are built. In the marine plankton, outbursts of parasitic dinoflagellates are sometimes associated with the collapse of zooplankton populations. The parasite *Ichthyodinium chabelardi* destroys the eggs of sardines, *Blastodinium* sp. in the copepod gut castrates its host, and *Gonyaulux catenella* (a dinoflagellate that can cause red tides) is parasitized by another dinoflagellate (*Amoebophrya ceratii*). The fossil record of dinoflagellate cysts extends over the past 220 million years, and they have a role as biostratigraphic tools.

Dinoflagellates in the marine plankton exhibit latitudinal cosmopolitanism, such that a morphospecies occurs in a circumglobal belt within fairly broad latitudinal limits which correspond to a specific temperature range in the surface waters. The recorded species richness is greatest in tropical marine waters.

Ciliated Protozoa

This diverse and distinctive group uses cilia for locomotion and feeding (**Figure 3**). They demonstrate a considerable adaptive radiation of feeding mechanisms and cell morphologies (e.g., the ribbon-shaped forms *Trachelonaphis* and *Geleia* adapted for life in the marine interstitial) (Fenchel, 1987). The smallest species tend to feed on bacteria-sized particles and the larger species on unicellular algae, filamentous cyanobacteria, other protozoa, and even rotifers and other microzooplankton. The raptorial feeders (e.g., *Prorodon*) use a simple mouth to catch diatoms, dinoflagellates, and other large food items individually; some (e.g., *Lacrymaria*) kill motile prey, whereas others (e.g., *Chilodonella*) "hoover" diatoms and other elongate food particles from surfaces. The filter feeders (e.g., *Cyclidium*) use fine-mesh filters to sieve suspended bacteria, and some of these ciliates (e.g., *Paramecium* and *Tetrahymena*) thrive in habitats with very high bacterial concentrations. Other ciliates (e.g., *Pleuronema* and *Tintinnopsis*) use coarser filters to collect small algae. In many ciliates (e.g., *Oxytricha*, *Aspidisca*, and *Strombidium*), a row of membranelles generates a water current and acts as a relatively coarse feeding filter, and some of these ciliates (e.g., *Euplotes*) feed most efficiently if they are raised on cirri (fused cilia). Many ciliates are typically sessile and aligned perpendicular to the substrate (e.g., *Stentor* and *Vorticella*). Diffusion feeders (e.g., *Podophrya* and other Suctoria) catch swimming prey (usually other protozoa) that collide with their sticky tentacles.

More is probably known about the biodiversity and ecology of free-living ciliates than any other protozoan group. There are several reasons for this, including the distinctive and immediately recognizable ciliate morphology and swimming behavior and the relative ease with which many species can be cultured. They are a group of predominantly free-living protozoa. A few are parasites (e.g., *Ichthyophthirius* infects fish and there is one human endoparasite of minor importance – *Balantidium coli*), but most species are either free-living or harmless commensals of aquatic invertebrates (the ciliate epifauna of marine crustaceans is particularly diverse). Free-living species are known from all natural aquatic habitats in which temperatures are $<45\,^{\circ}C$, including oceanic sinking detritus, freshwater and marine sediments, anaerobic municipal landfill sites, and sewage treatment plants. They are abundant (up to $\sim 106\ ml^{-1}$) in activated sludge plants, in which they consume bacteria and also flocculate bacteria and other suspended particulate matter. These activities aid the clarification of the effluent and the formation of sludge. About a dozen ciliate species are frequently recorded in used-water treatment plants, although more than 200 ciliate species have been recorded. Ciliates also have a role as indicators of the level of organic pollution in river water (Corliss, 1979, 2000; Lynn, 2008).

Most ciliates are in the size range 0.02–2 mm, so they are generally larger than the hetero trophic flagellates and other nanoplankton (0.002–0.02 mm) on which many of them feed. Planktonic ciliates are relatively abundant ($1–100\ ml^{-1}$) and important grazers of nanoplankton in marine and fresh waters. They are probably key grazers within the microbial loop which is responsible for the rapid remineralization of organic matter in the water column. The diet of planktonic metazoans includes ciliates, although the quantitative significance of this link is unclear. Benthic ciliates are often abundant ($>1000\ ml^{-1}$) and usually the most important grazers in freshwater sediments (especially in lakes) and marine sandy sediments of inshore waters.

Many ciliate species harbor prokaryotic or eukaryotic symbionts. At the oxic–anoxic boundary in the water column of lakes, most ciliates may harbor sufficient endosymbiotic algae (*Chlorella*) to render the consortia capable of net photosynthesis in dim light. The marine interstitial ciliate *Kentrophoros* carries ectosymbiotic chemolitho trophic sulfide-oxidizing bacteria, and most anaerobic ciliate species harbor methanogenic bacteria that act as a sink for (potentially inhibitory) H_2 produced by the ciliate. The endosymbiotic nonsulfur purple bacteria in the anaerobic ciliate *Strombidium purpureum* use waste H_2 from the ciliate as a reductant for photosynthesis. Most marine anaerobic ciliates carry ectosymbiotic sulfate-reducing bacteria. (Fenchel and Finlay, 1995)

Many ciliate species are microaerobic, and they seek out the (microbe-rich) oxic–anoxic boundary in sediment or the water column. Some of these species are facultative anaerobes (in the genera *Cyclidium*, *Euplotes*, *Strombidium*, and *Paranophrys*). There are approximately 70 known species of anaerobic free-living ciliate species, mainly in the genera *Metopus*, *Caenomorpha*, *Saprodinium*, *Epalxella*, *Trimyema*, and *Plagiopyla*. It is likely that all these contain hydrogenosomes (Fenchel and Finlay, 1995).

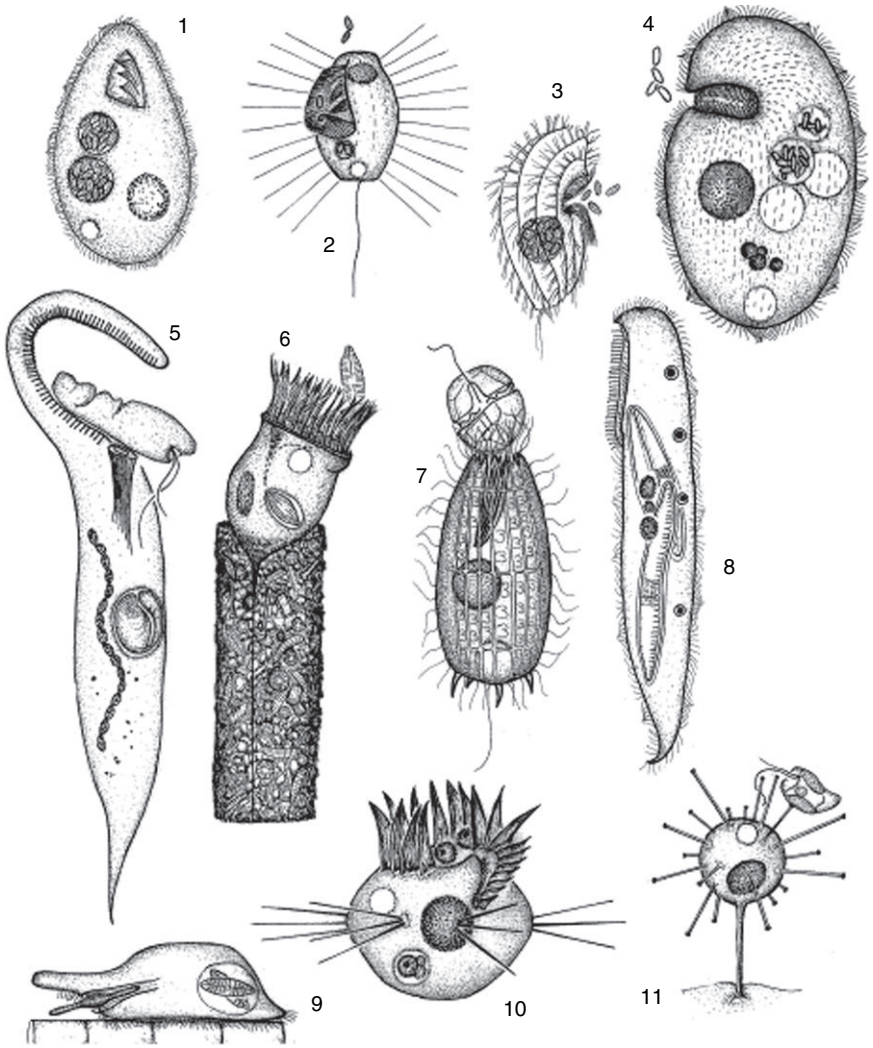


Figure 3 A selection from the variety of form and function in ciliated protozoa. (1) a bacteria-feeding tetrahymenid (*Tetrahymena*; length ~0.05 mm); (2) a small scuticociliate (*Cyclidium*; length ~0.02 mm) filter feeding on bacteria; (3) a bacteria-feeding colpodid ciliate (*Colpoda*; length ~0.02 mm) from soil; (4) an anaerobic, typically benthic plagiopylid feeding on bacteria (*Plagiopyla*; length ~0.1 mm); (5) a large haptorid (*Dileptus*; length ~0.4 mm) with toxicysts used to kill motile prey (flagellates and other ciliates) prior to ingestion; (6) a planktonic, loricate, diatom-feeding tintinnid (*Tintinnopsis*; length ~0.1 mm); (7) a gymnostome ciliate (*Coleps*; length ~0.07 mm) about to ingest a dinoflagellate; (8) a diatom-feeding karyorelictid (*Remanella*; length ~0.12 mm) typical of the marine interstitial; (9) a "hoover-feeding" cyrtophorid (*Chilodonella*; length ~0.04 mm) ingesting a diatom; (10) a planktonic oligotrich (*Halteria*; diameter ~0.04 mm) filter feeding on small algae; (11) a "diffusion-feeding" suctorian (*Podophrya*; diameter ~0.04 mm) with a flagellate trapped on a feeding tentacle.

Anaerobic ciliates also live as endocommensals in the enlarged forestomach (rumen) of ruminants and in the cecum of other mammals with postgastric fermentation. It is unlikely that any of these ciliates (e.g., in the genera *Dasytricha*, *Entodinium*, and *Polyplaston*) are capable of a free-living existence. Rumen ciliates are typically extremely abundant (105–106 ml^{-1} of rumen liquor). They consume bacteria and microscopic fragments of grass and other plant material. Some species are capable of degrading cellulose and other structural carbohydrates, and the endosymbiotic methanogenic bacteria in some species contribute to the methane emitted from the ruminants. The economic significance of such large numbers of ciliates living in the rumen is unclear.

Ciliates also live in soil, in which all species are probably capable of producing desiccation-resistant cysts (e.g., *Colpoda*). The abundance of active ciliates in soil is extremely variable and reflects repeated cycles of cyst formation and excystment in response to fluctuating physical factors such as the level of soil moisture. Many species found in soil are frequently found in other (predominantly freshwater) habitats, although only about 200 ciliate species have been described from soil.

Others

The following are often referred to as protozoa. Because they are endocommensals or parasites, or incapable of phagotrophy,

they do not fall within the definition of protozoa used here (i.e., free-living unicellular phagotrophs). They are mentioned briefly for the sake of completeness but are excluded from **Table 1**.

Chytridiomycetes, oomycetes, and hyphochytriomycetes have the typical fungal characteristics of saprotrophic (absorptive) nutrition and cell walls in the vegetative state. Some are economically important (the oomycete *Phytophthora infestans* causes potato blight). The plasmodiophorids are cell wall-free endoparasitic slime molds (*Plasmodiophora brassicae* causes club root disease in cabbages). The labyrinthulids (e.g., the "slime net" *Labyrinthula*) feed saprotropically, especially on marine algae, and form complex branching colonies of spindle-shape cells that move through slime channels. All five groups reproduce by means of unicellular flagellated zoospores (hence, they are also known as "zoosporic fungi").

The opalinids superficially resemble ciliates, but all are mouthless endocommensals living in the hindgut of amphibians and some fish (Lee et al., 2000). The microsporidians are a large group (~800 species) of intracellular parasites of animals and other protozoa. They extrude a tubular filament through which the spore is injected into the host cytoplasm. They are believed to be highly derived fungi rather than early diverging eukaryotes. The haplosporidians are a small group of parasites of aquatic invertebrates, notably marine mollusks, and the parabasalans (anaerobic flagellates with hydrogenosomes, e.g., *Trichomonas*) are all parasites, with one or two possible exceptions (*Pseudotrichomonas*).

The organisms now referred to as apicomplexans used to be an important component of the sporozoans. All are intracellular parasites of animals, and many cause life-threatening diseases of man. They share the distinctive morphological feature of an apical complex at certain stages in the polymorphic life cycle. The degree of host specificity of many apicomplexans is not known. They include the gregarines that infect insects, marine polychaetes, and other invertebrates (e.g., *Monocystis* in earthworms), and also the "coccidian" parasites including *Toxoplasma gondii* (the domestic cat is the final host, but intermediate infection in the human can cross the placenta) and about 1050 species of *Eimeria* infecting mammals, chickens, and other farmed birds. *Cryptosporidium* is an apicomplexan, and eight species are known that infect the digestive and respiratory systems of vertebrates. *Cryptosporidium parvum* infects humans and cattle, and waterborne transmission of oocysts is thought to complete the fecal–oral cycle. Infection can cause life-threatening diarrhea in immunocompromised humans. *Cryptosporidium* may be sufficiently different from other coccidians to warrant creation of the new higher taxon "cryptosporidia." *Plasmodium*, the causative agent of malaria, is an apicomplexan, as is *Pneumocystis*, which is an enigmatic parasite that can persist in the human lung for long periods. *Pneumocystis carinii* (currently renamed as *Pneumocystis jiroveci*) causes *Pneumocystis* pneumonia in HIV patients. In the early 1970s, the organism was considered to be a fungus; in the early 1980s it was considered to be an apicomplexan, and by the late 1980s it was again considered to be a fungus, possibly an ascomycete. Confusingly, it can resemble *Cryptosporidium* when viewed with the light microscope. There are approximately 4000 nominal species of apicomplexans.

The Myxozoans are a large group (~1200 species) of parasites, of which the myxosporidians are important parasites of farmed fish (*Myxobolus cerebralis* causes "torsion disease" in trout). During their life cycle they show both protozoan and metazoan characters. Recent gene sequence data (18S rDNA) indicate that myxozoans are grossly modified and simplified metazoans of the phylum Cnidaria (Siddall *et al.*, 1995).

See also: Eukaryotes, Origin of. Microbial Biodiversity. Microorganisms (Microbes), Role of. Plankton, Status and Role of. Predators, Ecological Role of

References

Adl SM, Simpson AGB, Farmer MA, *et al.* (2005) The new higher classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* 52: 399–451.

Anderson OR (1983) *Radiolaria.* New York: Springer-Verlag.

Corliss JO (1979) *The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature*, 2nd edn. New York: Pergamon.

Corliss JO (2000) Biodiversity, classification, and numbers of species of protists. In: Raven PH and Williams T (eds.) *Nature and Human society: The Quest for a Sustainable World*, pp. 130–155. Washington, DC: National Academy Press.

Esteban GF, Fenchel T, and Finlay BJ (2010) Mixotrophy in ciliates. *Protist* 161: 621–641.

Fenchel T (1987) *The Ecology of Protozoa.* Madison, WI: Science Tech Publishers.

Fenchel T and Finlay BJ (1995) *Ecology and Evolution in Anoxic Worlds. Oxford Series in Ecology and Evolution.* Oxford: Oxford University Press.

Finlay BJ (1998) The global diversity of protozoa and other small species. *International Journal of Parasitology* 28: 29–48.

Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061–1063.

Finlay BJ, Maberly SC, and Cooper JI (1997) Microbial diversity and ecosystem function. *Oikos* 80: 209–213.

Gooday AJ, Bett BJ, Shires R, and Lambshead PJD (1998) Deep-sea foraminiferal species diversity in the NE Atlantic and NW Arabian Sea: A synthesis. *Deep-Sea Research. Part II* 45: 165–201.

Lee JJ, Leedale GF, and Bradbury PC (eds.) (2000) *An illustrated guide to the protozoa*, 2nd edn. Lawrence, KS: Society of Protozoologists/Allen Press.

Lee WJ and Patterson DJ (1998) Diversity and geographic distribution of free-living heterotrophic flagellates – analysis by PRIMER. *Protist* 149: 229–244.

Lynn DH (2008) *The ciliated Protozoa: Characterization, Classification, and Guide to the Literature*, 3rd edn. New York: Springer.

Margulis L, Corliss JO, Melkonian M, and Chapman DJ (eds.) (1990) *Handbook of Protoctista.* Boston: Jones & Bartlett.

Mitchell EAD and Meisterfeld R (2005) Taxonomic confusion blurs the debate on cosmopolitanism versus local endemism of free-living protists. *Protist* 156: 263–267.

Page FC and Siemensma FJ (1991) *Nackte Rhizopoda und Heliozoea. Protozoenfauna Band 2.* Stuttgart: Fischer.

Patterson DJ and Larsen J (eds.) (1991) *The Biology of Free-Living Heterotrophic Flagellates*, Systematics Association Special Vol. 45. Oxford: Clarendon Press.

Siddall ME, Martin DS, Bridge D, Desser SS, and Cone DK (1995) The demise of a phylum of protists: phylogeny of Myxozoa and other parasitic cnidaria. *J Parasitol* 81: 961–967.

Psychrophiles

Richard Y Morita, Oregon State University, OR, USA
Craig L Moyer, Western Washington University, WA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 4, pp 917–924, © 2001, Elsevier Inc.

Glossary

Barophile Pressure-loving bacteria.
Cryobiosis Anabiosis (latent life) due to freezing.
Endolithotrophic Living inside rocks, usually sandstone.
Homeophasic adaption The adaptation of the membrane to maintain the bilayer phase.
Meltwater Ice or snow melted by radiant energy in the polar regions.
Permafrost Ground (ice, bedrock, and soil) that remains frozen below 0 °C for more than 2 years.
Thermocline In the stratification of warm surface water over cold, deeper water, the transition zone of rapid temperature decline between the two layers.
Upwelling Transport of water from the deep ocean to the surface, replacing the surface water that has moved offshore.

Definitions and Historical Background

Psychrophilic bacteria are defined as cold-loving bacteria. Specifically, their cardinal temperatures are 20 °C for maximal growth, 15 °C or lower for optimal growth, and 0 °C or lower for minimum growth (Morita, 1975), and this definition is accepted by most microbiologists. The old definition of psychrophiles applied to those organisms that produced a visible colony in 1 week at 0 °C. However, it is important to recognize that there is a continuum of cardinal temperatures for the diverse microbes in nature. From an ecological standpoint, psychrotrophs and psychrophiles are both found in cold environments, but psychrophiles are not found at temperatures higher than 20 °C. This temperature was selected as the maximum temperature for growth based on the fact that laboratory temperatures in the United States are approximately 21 or 22 °C. In higher organisms, the cold-loving organisms are known as cryophiles.

Microorganisms capable of growing at 5 °C or lower are psychrotrophs, regardless of the optimum temperature for growth. The psychrotrophs are cold-tolerant bacteria, but their maximal growth temperature ranges above 20 °C and in many cases their optimal growth temperature is also above 20 °C. A better term for these organisms that withstand cold temperatures is psychrotolerant. However, because of common usage this term should be retained. For example, there are a few texts employing the term "psychrotrophs" in their title and this term is widely used by industry (mainly dairy and food), e.g., the normal souring of milk is due to psychrotrophs. Furthermore, precedence should be adhered to in keeping the name.

Bacteria capable of growing at 0 °C were first reported by Foster in 1887 and 1892. The source material for these bacteria were from fish, natural waters, foods, wastes, rubbish, soil surface, and intestines of fish. The lowest temperature at which bacteria can grow remains to be determined definitely and −12 °C is the lowest temperature reported. The term "psychrophile" was first used in 1902 by Schmidt-Nielsen. In 1903, Müller objected to this term because the organisms described actually grow well at higher temperatures. As a result, cold-tolerant bacteria were called cryophile, Glaciale Bakterien, rhigophile, psychrotolerant, psychrocartericus, psychrobe, thermophobic bacteria, facultative psychrophile, obligate psychrophile, and psychrotrophic. A review of the literature indicates that the organisms described were actually psychrotrophs (as defined previously) with the possibility of one psychrophilic exception. For many years, it was thought that there were no bacteria that could be termed psychrophiles, only yeasts and certain algae. The lack of refrigeration equipment in the early days added much to this confusion. Thus, the term psychrophile was considered a misnomer until true psychrophiles were isolated by three different laboratories in 1964. The confusion was ended when Morita (1975) defined the term psychrophile (as noted previously). For further details, Morita (1975) should be consulted.

The Environment and its Microflora

Eighty percent of the earth's biosphere is permanently cold and the average temperature of the earth is 15 °C. Permafrost occupies 20% of the earth's surface: 80% of Alaska, 50% of Canada, 20% of China, and 50% of Russia are covered by permafrost. However, the amount of microbiological research done on psychrophiles is extremely low compared to that done on thermophilic bacteria.

The polar regions comprise about 14% of the earth's surface. Approximately 71% of the earth's surface is ocean, and more than 90% (by volume) of the oceans are 5 °C or colder. Other cold environments include caves, the tops of mountains, certain rivers and streams, the upper atmosphere (10 °C or less at 1000 m, decreasing as the altitude increases), snow and ice, and the water below the thermocline of freshwater lakes; each has its own microbial flora (Baross and Morita, 1978).

Upper Atmosphere

Air samples collected near the earth's surface up into the stratosphere exceeding 27,000 m have shown the presence of viable bacteria, viable fungi, pollen, and other microscopic

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00176-3

particles. A high incidence of viable bacteria has been reported within the troposphere at approximately 10,000 m, where the temperature may be lower than – 40 °C. The highest incidence of bacteria appears to be at altitudes slightly higher than 10,000 m, where the temperature is lower than 10 °C. Although this temperature is well within the cardinal temperature range of psychrophile and psychrotrophs, there are no reports of psychrophiles in the upper atmosphere. Furthermore, there is organic matter (cobalamin, biotin, and niacin) in the air but the concentration is very low. Thus, these organisms have the ability to resist starvation, drying, freezing, and radiation. These microbes can be the nucleation site for the formation of ice and snow.

However, in the lower atmosphere of the polar regions, nutrient agar plates were media exposed to the air and showed the presence of viable bacteria. Unfortunately, there are no reports of seawater-requiring psychrophilic bacteria from air samples. In all probability, the atmosphere, especially above the polar regions, does contain psychrophiles and psychrotrophs.

Caves

There are many glaciated and subterranean caves in which the permanent temperature is about 10 °C to below freezing. In addition to the low temperature, there is an absence of light, low levels of organic material, and relatively high moisture. Generally, there is an absence of psychrophiles in these caves, but psychrotrophs are found.

Arctic

The similarities of the Arctic and Antarctic polar regions are the continuous sunlight during the summer and the total lack of it during the winter, the presence of sea ice during much or all of the year, and the cold temperature. On the other hand, they differ in that the Arctic has major riverine inputs, but the Antarctic does not. They also differ in their landmasses, topographical features, large-scale water transport features, and the magnitude of nutrient supply. Relative microbial activities, as measured by the incorporation and respiration of ^{14}C-labeled glucose and glutamic acid (called the heterotrophic potential method by marine microbiologists and substrate-induced respiration by soil microbiologists), are comparable in both the Arctic and Antarctic near-shore water. These measurements were made by the same laboratory working at different times in both regions.

Antarctic

About 98% of the surface of Antarctica is ice, leaving 2% of the continent ice free. The lowest temperatures on Earth occur on this continent and it has the lowest precipitation and relative humidity levels, making it the driest area on Earth. As a result, dry valleys and ice-free areas occur in addition to many other types of environments (e.g., temporarily and permanently ice-covered lakes, transient rivers and streams, oligotrophic lakes, and moss and peat bogs), with all sharing low temperatures. Freezing, like drying, will produce hypersaline lakes, both thallosshaline (seawater derived) and athallosshaline (brine derived from geological and geochemical sources). For a few weeks of the austrial summer flowing liquid water can occur, producing ice melts around the margins of frozen lakes. Study of the microbiology of these meltwaters has just begun. In all probability, psychrophilic bacteria, which are also alkalophilic and/or halophilic, will be found. Even in the dry valleys of Antarctica, numerous psychrotrophs and mesophiles have been isolated. Because freezing is equivalent to drying, the indigenous microflora have the ability to tolerate low temperatures in addition to being xerotolerant (i.e., able to live in very dry environments). Within sandstones of the dry valleys, the presence of cryptoendolithotrophic bacteria has been shown along with that of algae (Friedmann, 1993). These bacteria, mainly cyanobacteria, only grow during the austral summer of the Antarctic when the temperature, due to solar radiation, rises sufficiently high to produce liquid water. It should also be mentioned that there is a greater predominance of cyanobacteria and psychrophiles in the Antarctic than in the Arctic.

Permafrost

The discovery of microorganisms in permafrost was initiated in the 1930s in the Trans-Baikal and North Ural regions, Central Yakutia, and Arctic islands (Gilichinsky, 1995). Permafrost cores yielded numerous microbes. Microbes have been reported in all the Arctic and Antarctic permafrost environments except in the lower strata of permafrost ice of Lake Vostok in Antarctica. This exception may be due to the techniques employed. Unfortunately, the early microbiologists did not recognize the abnormal thermolability of psychrophiles, and as a result psychrotrophs and mesophiles were mainly demonstrated. In the early literature, thermophiles (as high as 100 cells/g) were reported in Arctic permafrost (peat cores 10,525 years old). Viable microorganisms were reported from Holocene sedimentary deposits and Pliocene deposits (3–5 million years old). Viable fungi, spore-forming bacteria, and non-spore-forming bacteria were in the Antarctic ice sheet up to a depth of 300 m (11,600 years old) but not at 320 m (13,000 years old), whereas in a 107,000- to 200,000-year-old layer of glacier ice, viable organisms were found (Friedmann, 1993). In permafrost material millions of years old, thermophiles, mesophiles, and psychrophiles have been found (Vorobyova *et al.*, 1997). This ancient material does not have sufficient energy in it to keep the organism metabolizing for long periods of time; hence, one has to ask where the energy is coming from. In most cases, the energy supplied has been utilized since any ecological niche has a large number and consortium of microbes that utilize the energy source leaving the more recalcitrant material. This recalcitrant material, if used metabolically, will require exoenzymes to degrade it to its monomers so that it can be transported into the cell. In addition, many substrates that the cell utilizes require an active transport mechanism which also consumes energy. Because of this situation, microbes in ancient material, such as the deeper strata of permafrost, are in an anabiotic stage. Specially, the anabiotic stage is cryobiosis, brought about by the low temperature. However, the cells are also in a starvation-survival mode. Both processes generate stress proteins, some of which are shared. The Russian microbiologist

Omelyanski (1911) attributed the existence of microbes in ancient cold environments to anabiosis, a term not employed by most microbiologists. Nevertheless, one always must ask "where is the energy?"

For further information concerning viable microorganisms in permafrost, Gilichinsky (1994) should be consulted.

Oceans

The area below the thermocline is approximately 5 °C or colder. Thus, the deep sea is permanently cold. Gradually, as the higher latitudes are approached, the depth of the thermocline decreases until it is at the surface of the ocean. In the deepest portions of the oceans (trenches and deeps), barophiles can be isolated that are also psychrotrophic or psychrophilic. At the edge of the ice where ice is melting or being formed, there is a sea ice environment with an associated microbial community. All types of microorganisms can be demonstrated in waters below the thermocline, mainly because there are many different types of habitats available.

Sea Ice

Where the seawater is being frozen, a sea ice microbial community resides. The freezing process helps concentrate the dissolved nutrients, and when the ice melts the concentrated nutrients in the ice are release so that sufficient energy is present to permit the growth of this microbial community. Thus, this community is influenced mainly by the onset of spring and winter. Within this sea ice community a variety of bacteria are found displaying numerous morphological types. Seventy percent of the bacteria in the sea ice community are free living, and ice nucleation bacteria can also be found. Various pigmented and gas vacuolate bacteria are also present and 65% of the epiphytic bacteria are found associated with the diatom *Amphiprora*. The best growth response of the sea ice microbial community occurs between 5 and 10 °C.

Biodiversity of Psychrophiles

The first known species of psychrophiles described taxonomically are *Vibrio* (*Moritella* gen. nov.) *marinus* MP-1 and *Vibrio* (*Colwellia* gen. nov.) *psychroerythrus*, both isolated in 1964. The biodiversity among psychrophiles in the various cold environments has yet to be studied extensively. Nevertheless, the various species within the genera *Achromobacteria, Alcaligenes, Altermonas, Aquaspirillum, Arthrobacter, Bacillus, Bacteroides, Brevibacterium, Clostridium, Colwellia, Cytophaga, Flavobacterium, Gelidibacter, Methanococcoides, Methanogenium, Methanosarcina, Microbacterium, Micrococcus, Moritella, Octadecabacter, Phormidium, Photobacterium, Polaribacter, Polaromonas, Pseudomonas, Psychroserpens, Shewanella*, and *Vibrio* have been found to be psychrophilic. Even a psychrophilic methanotroph (resembling *Methylococcus*) has been isolated from the tundra. Both the domains *Bacteria* and the *Archaea* are represented in the various genera listed. To date, only the genus *Moritella* appears to be composed of psychrophiles only. However, it should be noted that many species of reported psychrophiles do not meet the definition given in this article. Because of this situation, the cardinal temperatures of many psychrophiles need to be determined.

All psychrophilic and barophilic bacteria that have been cultivated belong to the subdivision γ-Proteobacteria: *Shewanella*, *Photobacterium*, *Colwellia*, *Moritella*, and a new group designate containing strain CNPT3 and *Alteromonas haloplanktis*. The indication is that the combined barophilic and psychrophilic phenotype evolved independently in the different γ-Proteobacteria genera (DeLong *et al.*, 1997).

Several psychrophilic, gas vacuolate bacteria have been isolated from the sea ice and water from both the Antarctic and the Arctic. These include representatives belonging to the α-, β-, and γ-Proteobacteria subdivisions as well as the Cytophaga–Flavobacterium–Bacteroides phylogenetic group (Gosink *et al.*, 1998). These results demonstrate a wide range of phyolgenetic diversity capable of psychrophily across the domain Bacteria.

Incredibly, as much as 30% of the marine picoplankton (planktonic organisms with an average diameter of 0.2–2.0 μm) from both polar and temperate coastal waters are also Archaea (i.e., archaeoplankton) and the majority of these are associated with the Crenarchaeota. Despite this cosmopolitan distribution in the world's oceans, little is known about the physiological properties of these archaeoplankton, with the exception that they are hypothesized to potentially be psychrophilic based on the marine sponge symbiont *Cenarchaeum symbiosum* (Preston *et al.*, 1996). However, similar archaeal phylotypes have been located at a deep-sea hydrothermal vent system (Moyer *et al.*, 1998) in waters with an environmental temperature of 15–30 °C. Many hyperthermophiles are members of the Archaea that can utilize H_2 as an energy source, and recently *Methanogenium frigidum*, a psychrophilic, slightly halophilic, H_2-using methanogen, was isolated from the perennially cold, anoxic hypolimnion of Ace Lake, Antarctica (Franzmann *et al.*, 1997). Although it may be too early to state that this may be an indiction of the evolutionary processes affecting both hyperthermophiles and psychrophiles, this does demonstrate that members of the domain Archaea are also capable of the psychrophilic lifestyle and that further research is needed. The main deterrent is the lack of isolation of psychrophiles, mainly because interest in this thermal group was lacking for many years. Yet to be isolated from the meltwater in the Antarctic are psychrophilic bacteria that are also alkaliphiles. Each body of water in the Antarctic has its own cations and anions and the salinity may be very high due to freezing of the water.

In any cold environment in which microbes have been isolated, many of the isolates are psychrotrophs. When all things are equal, the psychrophiles will outgrow the psychrotrophs at low temperature.

Evolution of Psychrophiles

The common approach used in microbiology is to apply phylogenetic analysis to establish evolutionary relationships among organisms and to use this as a framework for making inferences about community structure, making inferences about genetic and thereby inferred organismal diversity, and

(to a lesser degree) to infer physiological adaptation when applicable. This approach is possible due to the detailed theory of evolutionary relationships among the domains Bacteria, Archaea, and Eucarya that has emerged from comparisons of ribosomal RNA "signature" sequences. Most researchers currently infer that the earliest common ancestor to all three domains was a H_2-metabolizing hyperthermophile based on the deepest branching lineages found in the domains Bacteria and Archaea (Pace, 1991). If we accept this premise, then the thermophiles must have evolved from the hyperthermophiles, followed by the mesophiles and finally the psychrophiles. This hypothesis has been challenged by criticism that the hot origin of life scenario is not compatible with the RNA world hypothesis (because of the instability of RNA at near water boiling temperatures) and/or by the argument that hyperthermophiles are not primitive but have instead successfully adapted to temperatures higher than the limits imposed on other life-forms (Forterre, 1996). In addition, dibiphytanyl ether lipids used in the formation of lipid "monolayer" cell membranes in hyperthermophilic Archaea were found in nonthermophilic members of the kingdom Crenarchaeota, which appear to be a major source of tetraether lipids in the marine planktonic environment (DeLong *et al.*, 1998). Recently, a nonhyperthermophilic earliest common ancestor was hypothesized based on the correlation between the G + C nucleotide content of ribosomal RNA sequences with the optimal growth rate temperature found in prokaryotes (Galtier *et al.*, 1999). If the nonhyperthermophilic hypothesis is presumed to be correct, it then implies that mesophiles were most likely the first to evolve with thermophily and psychrophily adaptations followed thereafter. Psychrophily undoubtedly has evolved multiple times, thereby giving rise to polyphyletic origins (i.e., psychrophilic members are contained within lineages spanning the divergence of divisions and domains).

A phylogenetic tree generated exclusively from SUU rRNA sequences of cultivated obligate psychrophiles demonstrates five major lineages spanning the prokaryotic domains of Bacteria and Archaea (**Figure 1**): the Crenarchaeota, the Euryarchaeota, the Flexibacter–Cytophaga–Bacterioides group, the gram-positive Bacteria group, and Proteobacteria. The majority of individual isolates are located in lineages contained either by the Flexibacter–Cytophaga–Bacteroides group or by the *Moritella*, *Colwellia*, and *Shewanella* subgroups of the γ-Proteobacteria.

Physiology of Psychrophiles (Metabolic Activities)

Psychrophiles were not isolated until 1964 mainly because of the fact that previous investigators did not recognize the abnormal sensitivity of this thermal class. Although it was known that most marine bacteria would not withstand the plating temperature of agar, microbiologists working in the Arctic and Antarctic did not take the precaution of keeping their sample material, media, pipettes, loops, and needles cold before use. Even for some psychrophiles, 10 °C is too warm and the organisms will expire.

Many assume that growth at low temperature is slow. This assumption was shown to be incorrect when *Moritella marinus* was reported to have a doubling times of 80.7 and 226 min at 15 and 3 °C, respectively. The lowest temperature recorded for growth of microorganisms is −12 °C, but there are unconfirmed reports of growth at −18, −20, and −34 °C for yeasts and fungi. Metabolic activity depends on the freezing properties of the cellular aqueous solutions and the liquid immediately adjacent to the cell. Aqueous solutions can be cooled to −10 to −20 °C before freezing occurs. Franzmann *et al.* (1987) recorded liquid water at −14.1 °C in Organic Lake, Antarctica. This liquid water naturally would be very saline since much of the water is freezing out. It is interesting to note that although *Halobacterium* and *Halomonas* are both halophilic as well as psychrophilic, no growth could be demonstrated in the liquid water from either Organic and Deep Lakes, Antarctica; hence, temperature was not the only limiting factor for growth. However, like most environments in the biosphere, the lack of other environmental or chemical conditions (mainly energy) usually prevents rapid growth of the microorganisms.

Enzymes

Unlike thermophiles, research on the physiology of psychrophiles has been neglected over the years. The few studies dealing with enzymes (and cytoplasmic proteins) from psychrophiles indicate that they are more thermolabile compared to their counterparts in mesophilic bacteria–their maximum, minimum, and optimum temperatures for activity are lower. However, most of these enzymes will operate several degrees above the maximum growth temperature of the psychrophiles. Malic dehydrogenase, the first enzyme from a psychrophile to be studied, obtained from cells of *M. marinus* was found to be stable between 0 and 15 °C and inactivation occurred between 15 and 20 °C. This organism's optimum growth temperature is 15 °C and its maximum is 20 °C. If the enzyme is partially purified, inactivation at higher than 20 °C becomes very pronounced. Phosphofructokinase/glyceraldehye-3-phosphate dehydrogenase complex, lactic dehydrogenase, hexokinase, and aldolase from this same organism lost activity when incubated for 1h to 35–40 °C. Surprisingly, glyceraldehyde-6-phosphate dehydrogenase was found to be stable at 26 °C; however, at 36 °C for 1h, it lost 90% of its activity. Triosephosphate isomerase, a glycolytic enzyme, isolated from *Moritella* sp. strain ANT-300 (maximum growth temperature of 12 °C) is extremely heat sensitive and has a half-life of heat inactivation of 520s at 25 °C. Other enzymes that have been examined from psychrophiles include α-amylase, lipase, protease, alkaline phosphatase, and β-lactamase. As a general rule, the enzymes isolated from psychrophiles have not adapted entirely to correspond with the cardinal temperatures of the respective psychrophile cell. The abnormal thermoability of enzymes from psychrophiles is not the primary target of death from heat. Nevertheless, enzymes from psychrophiles are receiving more attention due to their possible industrial and biotechnological uses.

Membranes

In order for the cell to maintain its bilayer structure of membrane lipids, homeophasic adaptation has been proposed. This

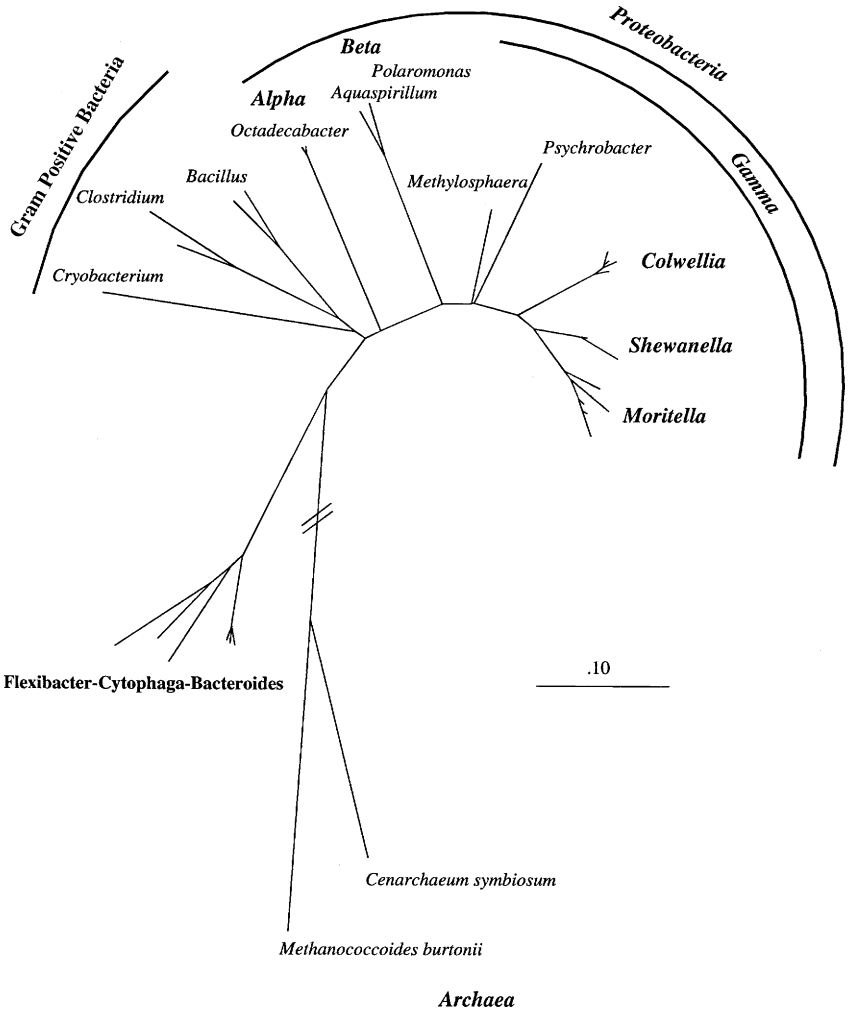


Figure 1 Radial phylogenetic tree using the neighbor-joining distance method demonstrating the evolutionary relationships among cultivated psychrophiles. The tree was constructed using complete SSU rRNA sequences contained in the Ribosomal Database Project from the 99/99 release of Version 7.1 (http://www.cme.msu.edu/RDP/html/index.html), with the additions of *Cenarchaeum symbiosum* and *Moritella* sp. ANT-300. The scale bar represents 0.10 fixed mutations per nucleotide position.

adaptation permits the cell to function properly in terms of uptake, excretion, and regulation of intracellular ionic composition.

When psychrophiles are subject to temperatures above their maximum growth temperature, the cell leaks amino acids, DNA, and RNA, indicating that the membrane can no longer control its transport mechanisms. Furthermore, as the temperature reaches the maximum growth temperature or higher, the ability to respire is impaired. As a result, the cell expires.

Mesophiles and thermophiles cannot grow at low temperatures mainly due to the loss of membrane fluidity. This is why mesophiles (e.g., *Escherichia coli*) that have lost their membrane fluidity can no longer transport substrates into the cell. The fatty acid composition of the membrane changes in response to temperature. These homeophasic adaptations are often represented by changes in polyunsaturation, chain length, branching, cyclization, and often a combination thereof (Russell, 1992). Membranes of psychrophiles and psychrotrophs contain more polyunsaturated short chains and branch and/or cyclic fatty acids than do mesophiles and thermophiles. In general, the membrane lipids are composed of an increasing amount of unsaturated fatty acids as the temperature decreases. Membrane fluidity can also be changed by *cis*/*trans* isomerization of the double bonds of fatty acids. Depending on the species in question, psychrophiles are generally endowed with a higher proportion of unsaturated fatty acids, especially hexadecenoic (16:1) and octadecenoic (18:1) acids, than mesophilic bacteria. The amounts of myristic acid (14:0) and palmitic acid (16:0) are higher in cells that are grown at higher temperatures. When grown at 0 °C, there is an increase in docosahexaenoic acid (22:6) and/or eicosapentaenoic acid (20:5) in some strains of psychrophiles and barophiles.

In order to make the membrane lipids more unsaturated when the temperature is lowered, a desaturase enzyme is used which acts on the acyl chain of the membrane lipids ((Russell, 1992). This desaturase activity is followed by temperature-dependent changes in fatty acid chain length

and branching mediated by additional synthesis. These metabolic activities all serve to maintain the fluidity of the cell membrane so that it can function properly. The foregoing data were obtained by using a psychrotroph; therefore, the ability to alter the fatty acid composition may be restricted to psychrotrophs and not psychrophiles.

It should be noted that when the psychrophile *Moritella* sp. strain ANT-300 was starved, there was induced qualitative and quantitative changes in fatty acids (e.g., fatty acid 16:1 increased from 42 to 62.5%). When starved, cells of *Moritella* sp. strain ANT-300 (as with many other bacteria) underwent fragmentation (i.e., reductive division), resulting in ultramicrocells. The starvation-survival state is the normal state of most microbes since most of the biosphere is oligotrophic. The oligotrophic nature of the psychrophile's environment is predominant since organic matter is lacking for the growth of heterotrophic psychrophilic bacteria. Thus, over most ($\sim$80%) of the earth's biosphere and in both polar environments, most of the microbes are in a starvation-survival mode most of the time.

Thus, it appears that various species of psychropiles evolved from their mesophilic component and they inhabit permanently cold environments. Their upper and lower temperature limits are 20 and $-12\,^{\circ}C$, respectively.

See also: Antarctic Ecosystems. Archaea, Origin of. Arctic Terrestrial Ecosystems. Microbial Biodiversity. Thermophiles, Origin of

References

Baross JA and Morita RY (1978) Life at low temperatures: Ecological aspects. In: Kushner DJ (ed.) *Microbial Life in Extreme Environments*. London: Academic Press.

DeLong EF, Franks DG, and Yayanos AA (1997) Evolutionary relationships of cultivated psychrophilic and barophilic deep-sea bacteria. *Appl. Environ. Microbiol.* 63: 2105.

DeLong EF, King LL, Massana R, *et al.* (1998) Dibiphytanyl ether lipids in nonthermphilic crenarchaeotes. *Appl. Environ. Microbiol.* 64: 1133.

Forterre P (1996) A hot topic: The origin of hyperthermophiles. *Cell* 85: 789.

Franzmann PD, Deprez PP, Burton HR, and McMeekin TA (1987) Limnology of Organic Lake, Antarctica, a meromictic lake that contains high concentrations of dimethyl sulfide. *Aust. J. Mar. Freshwater Res.* 38: 409.

Franzmann PD, Liu Y, Balkwill DL, Aldrich HC, Conway de Macario E, and Boone DR (1997) *Methanogenium frigidum* sp. nov., a psychrophilic, H_2-using methanogen from Ace Lake, Antarctica. *Int. J. Syst. Bacteriol.* 47: 1068.

Friedmann EI (ed.) (1993) *Antarctic Microbiology*. New York: Wiley-Liss.

Galtier N, Tourasse N, and Gouy M (1999) A nonhyperthermophilic common ancestor to extant life forms. *Science* 283: 220.

Gilichinsky D (ed.) (1994) *Viable Microorganisms in Permafrost*. Pushchino: Russian Academy of Sciences.

Gilichinsky D (1995) Permafrost microbiology. *Permafrost Periglacial Processes* 6: 281.

Gosink JJ, Woese CR, and Staley JT (1998) *Polaribacter* gen. nov., with three new species, *P. irgensii* sp. nov., *P. franzmannii* sp. nov. and *P. filamentus* sp. nov., gas vacuolate polar marine bacteria of the Cytophaga–Flavobacterium–Bacteroides group and reclassification of "*Flectobacillus glomeratus*" as *Polaribacter glomeratus* comb. nov. *Int. J. Syst. Bacteriol.* 48: 223.

Morita RY (1975) Psychrophilic bacteria. *Bacteriol. Rev.* 39: 144.

Moyer CL, Tiedje JM, Dobbs FC, and Karl DM (1998) Diversity of deep-sea hydrothermal vent *Archaea*. *Deep-Sea Res. II* 45: 303.

Omelyansky VL (1911) Bakteriologicheskoe issledovanie Sanga mamonta Prilegayushchei pochvy. *Arkhiv Biologicheskikh Nauk* 16: 335.

Pace NR (1991) Origin of life—Facing up to the physical setting. *Cell* 65: 531.

Preston CM, Wu KY, Molinski TF, and DeLong EF (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc. Natl. Acad. Sci. USA* 93: 6241.

Russell NJ (1992) Physiology and molecular biology of psychrophilic microorganisms. In: Herbert RA and Sharp RJ (eds.) *Molecular Biology and Biotechnology of Extremophiles*. Glasgow: Blackie.

Vorobyova E, Soina V, Gorlenko M, *et al.* (1997) The deep cold biosphere: Facts and hypothesis. *FEMS Microbiol. Rev.* 20: 277.

R

Rainforest Ecosystems, Animal Diversity

Gregory H Adler, University of Wisconsin at Oshkosh, Oshkosh, WI, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allopatric speciation The evolutionary development of new species in the presence of a geographical barrier which reduces gene flow and promotes genetic divergence.

Neotropics A biogeographic region that includes the New World tropics, extending from southern Mexico through the Southern Cone of South America to Tierra del Fuego. Many different ecosystems are found here, including tropical rainforest.

Species diversity This has two connotations. In a broad sense, it simply refers to the number of species of a particular taxonomic group living within a given area and is used synonymously with species richness. In a narrow sense, it refers to the number of species within a given area while simultaneously taking into account their relative abundances. In this article, species diversity is used in a broad sense and is used interchangeably with species richness.

Species richness The number of species of a particular taxonomic group living within a given area.

Tropical rainforest A forest that occurs below 1000 m in elevation, experiences high, relatively constant temperatures, and receives at least 200 mm of rainfall per year.

Introduction

Tropical rainforests contain more species of animals than any other ecosystem on Earth. This article reviews the distributions and types of rainforests and the diversity of selected groups of animals in tropical rainforests. The article concludes with a discussion of the global and regional patterns of rainforest animal diversity, some hypotheses that have been developed to explain the high diversity, and the future of rainforest animal diversity.

It is indeed ironic that Linneaus, cataloging specimens collected by early naturalists for the museums of Europe, considered the tropics to be of low diversity. However, the first naturalists collected solely in villages that contained second-growth vegetation and where hunting was intense. It was not until naturalists penetrated the dense border of vegetation between villages and undisturbed forest that the high diversity became evident. Single collecting expeditions routinely returned to Europe with thousands of undescribed species of plants and animals. Discoveries of new species of animals continue into the twenty-first century. It has been estimated that several million species of insects alone are yet to be discovered from tropical rainforests. Even larger and presumably more conspicuous animals, such as birds and mammals, are being discovered every year in the world's tropical rainforests. For instance, in the 1990s new species of monkeys were found in Brazil and two new species of deer and a relative of the goat were discovered in Laos and Vietnam. New locality records are also being established for many species of animals. For instance, the Andean spectacled bear has recently been found in lowland rainforests, although it has long been considered to occupy only higher elevation cloud forests. Alston's mouse opossum, formerly known only from parts of Central America to western-most Panama and again in western Colombia, recently has been found in lowland rainforest in central Panama.

Definition and Geographical Context

Tropical forests cover much of the land area that lies between the Tropics of Cancer and Capricorn (23°30′ N and S latitude). Tropical forests also extend outside of the two tropics in southeastern Brazil and northeastern Australia. The lowland tropics have a relatively high constant temperature, where the mean temperature of the coldest month is at least 18 °C. However, not all warm areas in the tropics support tropical forest. Rainfall, in concert with high stable temperature, is a crucial factor determining the development of forest. In many areas, closed-canopy forest begins to develop where there is at least 800–1000 mm of rain per year (Mabberly, 1992; Kricher, 1997; Whitmore, 1998; Leigh, 1999). Although the term rainforest has been applied loosely and sometimes inappropriately to virtually any type of tropical forest, a more precise definition of rainforest has been developed. Tropical rainforests occur in areas of high, relatively constant temperature (below approximately 1000 m in elevation) and where rainfall exceeds 2000 mm per year. A dry season may occur but generally does not exceed 4 months in duration. Forests growing under such conditions are composed

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00159-3

primarily of evergreen trees (<15% of the trees are deciduous) and have a closed canopy of at least 25 m in height. Larger trees, called emergents, often protrude above the canopy and may reach heights of up to 50 m. Thus, rainforests are defined based on the temperature and precipitation regimes under which tropical forests develop and also on the structural characteristics of these forests.

An understanding of the geographical context of rainforests is necessary to a discussion of animal diversity in these forests. For historical reasons (i.e., evolutionary origins of taxa in specific geographical regions and subsequent inability to disperse to other regions), many groups of animals are geographically restricted to single tropical regions. Vertebrates illustrate this point. The species-rich frog family Leptodactylidae is restricted to the New World and reaches its greatest diversity in rainforests. Iguanas (Iguanidae) and anoles (Polychrotidae), two species-rich families of lizards with greatest diversity in rainforests, are restricted to the New World, except for iguanas in Fiji that stem from New World colonizers. Families of birds such as toucans (Ramphastidae), motmots (Momotidae), jacamars (Galbulidae), puffbirds (Bucconidae), woodcreepers (Dendrocolaptidae), ovenbirds (Furnariidae), antbirds (Thamnophilidae), antthrushes (Formicariidae), and cotingas (Cotingidae) are restricted to the New World tropics, with many species occurring in rainforests. Birds such as hornbills (Bucerotidae), honeyguides (Indicatoridae), broadbills (Eurylaimidae), sunbirds (Nectariniidae), white-eyes (Zosteropidae), and pittas (Pittidae) are restricted largely to the tropics of the Old World (with a few species found also in adjacent subtropical regions). Birds of paradise (Paradisaeidae) and bowerbirds (Ptilonorhynchidae) are restricted to the tropics of Australia and New Guinea. Marsupials, which reach their greatest diversity in rainforests and surrounding areas, are restricted to the New World and the Australopapuan region. Among primates, two families of monkeys (Callitrichidae and Cebidae) are restricted to the New World tropics, whereas Cercopithecidae is restricted to the Old World and Hylobatidae is restricted to Southeast Asia. The lemurs, a group of primates composed of five families, are found solely in Madagascar.

Rainforests are found in three major regions of the world (Mabberly, 1992; Whitmore, 1998). The rainforests of the neotropics cover the greatest area (approximately 4 million km^2) and extend from southern Mexico to southeastern Brazil. The neotropics contain four fairly distinct regions that are isolated from each other by mountain ranges, savannas, or scrub vegetation. Trans-Andean rainforests are distributed from Chiapas and Veracruz, Mexico, to the Pacific slope of Colombia and Ecuador. Venezuelan coastal rainforests are found in northern Colombia, Venezuela, and the Guianas. Amazonian rainforests occupy most of the Amazon basin and constitute the largest contiguous area of tropical rainforest in the world. Atlantic rainforests formerly covered a long narrow swath of coastal southeastern Brazil but have been reduced to remnant patches by human activity. Only about 10% of the Atlantic rainforest remains. The Asia-Pacific tropics, also known as Malesia, extend from China's Yunnan Province southward to Australia. Included in this region are large parts of Southeast Asia (Burma, Thailand, Malaysia, Indonesia, Singapore, Cambodia, Laos, and Vietnam) and the Australopapuan region (New Guinea, northeastern Australia, and tropical Pacific islands such as New Caledonia, the Bismarck Archipelago, Solomon Islands, Fiji, and Vanuatu). Asia-Pacific tropical rainforest also occurs far to the west in the Western Ghats of India. This discontinuous and largely insular block of rainforest is second in areal extent only to the neotropics, covering approximately 2.5 million km^2. The Australopapuan region harbors faunal elements that are sufficiently distinct to sometimes warrant separate treatment. For instance, monotremes (egg-laying mammals), marsupials, and several families of birds are found there but not in the remainder of the Asia-Pacific tropics. Conversely, many families of birds and mammals are not found in the Australopapuan region but occur in the remainder of the Asia-Pacific tropics. Africa contains a large block of rainforest in the Congo River basin in central Africa and a smaller block of rainforest in West Africa. An outlying area of rainforest occurs in eastern Madagascar. The African rainforests cover approximately 1.8 million km^2.

Types of Rainforests

Researchers frequently delimit lowland tropical forests based on annual amounts of rainfall such forests receive and the temporal distribution of rainfall. These delimitations are somewhat arbitrary, and it is important to recognize that contiguous forest often occurs along a rainfall gradient where one type of forest grades into another with little obvious demarcation. Tropical scrub forest and tropical dry forest typically receive less than 2000 mm of rain per year, most of which falls in a few months. Severe dry seasons of up to 8 or 10 months punctuate the short rainy seasons, and many trees are deciduous and shed their leaves during the dry season to conserve water. Such forests, particularly scrub forests, often do not have a closed canopy. If the authors adhere to their definition of rainforests given previously, then these tropical forests do not constitute true rainforests (although they frequently abut rainforests and share many species of animals), and this article will not discuss their animal diversity. The remaining forest types that will be discussed are sometimes referred to as semievergreen and evergreen forests and conform to the definition of tropical rainforests. Tropical moist forests receive from about 2000 to 4000 mm of rain per year. These forests frequently have a pronounced dry season of up to 4 months, but heavy rains may occur during the dry season. Deciduous trees may occur in such forests, but they are much less frequent than in the preceding two types of forests. Tropical moist forests occupy the greatest areal extent of rainforests and occur throughout much of the three major geographical rainforest regions. Tropical wet forests typically receive more than 4000 mm of rain per year and frequently have no extended dry season. The wettest of these forests (those forests receiving more than 8000 mm of rain per year) sometimes are designated as pluvial forests and are essentially a seasonal. The driest months in these forests typically receive more rain than the wettest months in tropical moist forests.

Tropical forests may also be defined based on the elevations at which they occur. In mountainous areas, contiguous forest may be found from near sea level to the tree line. However, forest characteristics change dramatically with elevation. The upper elevational limit of lowland rainforests is frequently considered to be 1000 m, where premontane forest often occurs,

although this limit is arbitrarily chosen. Lower montane and upper montane forests, which do not fall within the operational definition of lowland rainforests, occur at increasingly higher elevations. In some areas, montane forests may occur at elevations lower than 1000 m, depending on latitude and stature of mountain ranges. For instance, small mountain ranges and lower spurs from larger ranges often support montane forests at lower elevations, which is termed the Massenerhebung effect. This article will focus on animal diversity in lowland rainforests (i.e., tropical moist and wet forests lower than 1000 m).

Overview of Animal Diversity

Tropical–temperate comparisons of diversity are commonly made to highlight the high diversity of the tropics, and this latitudinal diversity gradient has attracted considerable attention from ecologists and evolutionary biologists. These comparisons show that many major groups of animals reach their greatest diversity in tropical rainforests. Conspicuous among groups of animals that reach their greatest diversity in such forests are beetles, butterflies, frogs, birds, and mammals. This article will focus on these groups because they best exemplify the diversity of rainforest animals, and tropical–temperate comparisons will be included to highlight rainforest diversity.

It should be noted, however, that other major groups of animals, particularly among invertebrates, have been poorly sampled in tropical rainforests. It is evident, however, that other such groups also exhibit high rainforest diversity. For example, 498 species of spiders in 33 families were collected during 6 weeks at Pakitza in Manu National Park, Peru. Many of these specimens represented new species. Several weeks of sampling at the same site yielded 224 species of caddisflies (order Trichoptera), 66% of which could not be identified confidently. Sampling of dragonflies and damselflies (order Odonata) in the same area yielded 117 lowland rainforest species. At Tambopata Reserve in Peru, 135 species of ants were found, and one individual tree harbored 43 species, which is approximately the total number of ant species found in the entire British Isles or in all of Canada. The 1500-ha La Selva reserve in Costa Rica contains an estimated 4000 species of moths. Thus, when invertebrates have been sampled in tropical rainforests, diversity is usually very high, and new species invariably are found.

Diversity of Beetles

Approximately 1.4 million species of organisms have been described, about 400,000 of which are beetles. Beetles therefore constitute greater than one-fourth of all living species that have been cataloged by taxonomists. Beetles have been poorly sampled in rainforests, and any discussion of their diversity will be incomplete. However, it is clear from even preliminary sampling that beetle diversity is extremely high and that beetles no doubt are the most species-rich group of rainforest animals, comprising an estimated 40% of all arthropods. Indeed, the majority of rainforest animal species may be still the undescribed beetles.

There has been considerable conjecture over how many species of beetles occur in rainforests. Rainforest beetles reach their greatest richness in the canopy, in which standardized sampling efforts have been concentrated. Such sampling involves the deployment of insecticidal sprays into the canopy, a process known among "coleopterists" (scientists who study beetles) as fogging. Sampling of beetles (excluding weevils) by fogging 19 individuals of a single species of rainforest tree (*Lueehea seemanni*) in Panama produced 955 species, 163 species of which were found only on this one species of tree. This preliminary sampling was used to estimate the total number of arthropod species in tropical rainforests (Erwin, 1982). The estimate was based on the number of beetle species restricted to a single species of tree, the approximate number of tropical tree species, the proportion of arthropods that are beetles, and the contribution of canopy species to the overall arthropod species pool. Based on this reasoning, estimates of 30–50 million species of insects and >7 million species of beetles have been derived, which indicate that insects may constitute about 97% of all species of living organisms on Earth and that tropical rainforest beetles contribute the greatest proportion of species. These estimates of global biodiversity based on rainforest beetles stimulated biologists to reassess their more traditional estimates of global biodiversity (Wilson and Sandoval, 1996).

Sampling of ground beetles (Carabidae) at Pakitza in Manu National Park, Peru (not all species of which are confined to the ground), has revealed this 4000-ha site to have the greatest species richness of this family ever recorded. Indeed, nearly as many species have been found in this small area as have been found in all of New Guinea. More than 600 species have been recorded from Pakitza, and not only many new species but also new genera have been discovered. It should be noted that sampling was incomplete and that dozens of additional species are expected.

Comparison of beetle species richness among sites is difficult because of differences in sampling methodology and intensity and the paucity of sufficiently rigorous studies. However, several studies using the standardized fogging methodology allow researchers to calculate species richness on equal footing as the number of species per m^3/m. Among canopy and subcanopy beetles, estimates range from 0.02 species per m^3m^{-1} in an Australian rainforest tree and in rainforest in Sulawesi to 1.5 species per m^3m^{-1} in a Peruvian rainforest. Other estimates for rainforest sites are 0.29 species per m^3m^{-1} in New Guinea, 0.32 species per m^3m^{-1} in Brunei, and 1.17 species per m^3m^{-1} in Panama. The estimate for Panama was derived from the same data set from which global biodiversity estimates were made.

Diversity of Butterflies

Unlike most groups of insects, butterflies have been studied so extensively that about 90% of all living species presumably have been described. An estimated 13,750 species of butterflies, excluding skippers (family Hesperiidae), occur in the world. With skippers included, the number is approximately 17,500 species. Because of our thorough understanding of butterfly systematics, they are perhaps the most appropriate insects for examining rainforest diversity.

A tropical–temperate comparison of butterfly diversity highlights the great richness of tropical rainforest butterflies.

Table 1 Approximate numbers of butterfly species in four families in lowland rainforests of Costa Rica compared with lowland temperate forests of West Virginia

Family	*West Virginia*	*Costa Rica*
Papilionidae	5	25
Pieridae	4	31
Riodinidae	1	208
Nymphalidae	16	234
Total	26	498

Table 1 shows the numbers of butterfly species in four families in temperate forests of West Virginia and tropical rainforests of Costa Rica. These two areas are of comparable size, and the butterfly faunas in the four families are well-known (De Vries, 1987, 1997). Butterflies in two additional families (Lycaenidae and Hesperiidae) are too poorly known to make a valid comparison with those in West Virginia. Included in the species tally for West Virginia are those species known to occur in all forest types below 1000 m, whereas the species tally for Costa Rica includes only those species that are known to occur in lowland rainforests. Thus, butterflies occurring in both dry and wet temperate deciduous forests are included in the tally for West Virginia, but habitats such as tropical scrub and dry forests and mangrove forests are excluded from the tally for Costa Rica.

It is important to note that species tallies will likely vary depending on definitions of forest types and the compilation of results from additional sampling and natural history studies. Many species of butterflies live in disturbed areas adjacent to rainforest but are not included in the species tally because they are not true denizens of the rainforest.

Regardless of such qualifications, it is clear that rainforest butterfly diversity is much higher than in temperate forests. Costa Rican rainforests harbor nearly 20 times as many species of butterflies in the four families. Of particular note is the extremely high relative diversity in two Costa Rican families, Riodinidae and Nymphalidae, which together constitute nearly 90% of Costa Rica's rainforest butterfly fauna (De Vries 1987, 1997). Although the majority of West Virginia forest species are nymphalids, only one species of riodinid is found in West Virginia's lowland forests, and the two families together constitute only 65% of the total butterfly fauna.

Sampling at single sites further reveals the high species richness of Costa Rica's butterfly fauna. At the 1500-ha La Selva reserve in Costa Rica, 204 species have been recorded. However, Costa Rica's rainforest butterfly fauna is not particularly rich relative to that of other rainforest sites, especially in the neotropics. On a nineteenth-century expedition to the Amazon, Bates found 550 species of butterflies at a single site. This number is small compared with the total butterfly list for sites in the Peruvian Amazon. Within a 200-ha study site at the Explorer's Inn Reserve in southern Peru, 1234 species have been recorded, and approximately 1300 species have been recorded from Pakitza in Manu National Park. Thus, these single sites contain approximately 10% of all butterfly species in the world. The total numbers of species at these sites and other sites in the neotropics are not final; virtually every collecting expedition to these sites finds additional species. Indeed, the 1300 species from Pakitza were collected during only five sampling trips that averaged <3 weeks in duration. Nonetheless, the butterfly lists for these single sites exceed the lists for virtually all countries in the African and Asia-Pacific tropics.

Diversity of Frogs

Frogs are conspicuously species-rich in tropical rainforests, and their richness is particularly evident when breeding. At a single site in the Amazon rainforest in Santa Cecilia, Ecuador, 81 species of frogs have been recorded, which to date is the highest tally of frog species for any single site in the world. The number of species at this site equals the total number of frog species in the entire United States. High frog species richness has also been recorded from other tropical rainforest sites (Inger, 1980). In Panama, 59 species of frogs have been recorded on the 1500-ha Barro Colorado Island. More than 40 species have been recorded from La Selva, Costa Rica, 38 species from Makokou, Gabon, 25 species from Pasoh, Malaysia, and 23 species from Gogol, Papua New Guinea. On the island of Borneo, 33, 51, and 46 species of frogs have been recorded at three different lowland rainforest sites, respectively, and 29 species of larval frogs have been recorded from rainforest streams at Nanga Tekalit in Sarawak. At Sakaerat, northern Thailand, 19 species of frogs have been recorded within rainforest, with an additional five species found in adjacent dry forest.

Diversity of Birds

Tropical rainforests typically harbor the greatest diversity of birds compared with other ecosystems (Stotz *et al.*, 1996). Costa Rica's bird fauna is well-known, and a comparison of the number of resident bird species in that country's lowland rainforests with resident birds in all forest types in West Virginia illustrates this point. **Table 2** gives the approximate number of bird species that are resident in Costa Rica's lowland rainforests compared with the number that are resident in lowland forests of all types in West Virginia. Excluded from this list are migratory species that breed in West Virginia or that overwinter in Costa Rica. Also excluded are species that are found in disturbed areas adjacent to rainforests, in clearings within rainforests, or along major watercourses that may pass through rainforests. Although the same cautionary note applies to this comparison as it did with butterflies, it is again evident that Costa Rica's resident rainforest avifauna is much more diverse than that of West Virginia's lowland forests.

Costa Rican rainforests contain more than eight times as many species of birds as do all lowland forest types in West Virginia. Much of this greater diversity stems from two taxonomic sources. First, Costa Rican rainforests harbor more families of birds with resident species (42 compared with 18 in West Virginia). There are 29 families of birds that have representatives that are resident in Costa Rican rainforests but not in West Virginia's forests, and these families contribute 218 species to the total rainforest avifauna. By contrast, only seven families have resident species in West Virginia's forests but not in Costa Rican rainforests, and these families contribute only

Table 2 Approximate numbers of resident breeding bird species in lowland rainforests of Costa Rica compared with lowland temperate forests of West Virginia

Order and family	*West Virginia*	*Costa Rica*
Tinamiformes		
Tinamidae	0	3
Ciconiiformes		
Ardeidae	0	3
Threskiornithidae	0	1
Falconiformes		
Cathartidae	2	3
Accipitridae	4	21
Falconidae	0	5
Galliformes		
Cracidae	0	2
Phasianidae	3	4
Gruiformes		
Rallidae	0	1
Eurypygidae	0	1
Columbiformes		
Columbidae	1	12
Psittaciformes		
Psittacidae	0	11
Cuculiformes		
Cuculidae	0	3
Strigiformes		
Strigidae	4	9
Caprimulgiformes		
Nyctibiidae	0	2
Caprimulgidae	0	2
Apodiformes		
Apodidae	0	4
Trochilidae	0	28
Trogoniformes		
Trogonidae	0	8
Coraciiformes		
Alcedinidae	0	5
Momotidae	0	6
Piciformes		
Galbulidae	0	2
Bucconidae	0	5
Capitonidae	0	2
Ramphastidae	0	6
Picidae	6	11
Passeriformes		
Dendrocolaptidae	0	13
Furnariidae	0	13
Thamnophilidae	0	19
Formicariidae	0	8
Cotingidae	0	8
Tyrannidae	1	44
Pipridae	0	9
Corvidae	3	2
Paridae	3	0
Sittidae	2	0
Certhiidae	1	0
Regulidae	1	0
Troglodytidae	2	14
Turdidae	1	5
Mimidae	1	0
Bombycillidae	1	0
Sylviidae	0	
Vireonidae	0	4
Parulidae	1	6
Icteridae	0	3
Thraupidae	0	40
Emberizidae		11
Fringillidae	1	0
Total	43	362

10 species to West Virginia's resident forest avifauna. Second, several families that have resident representatives in both areas are much more diverse in Costa Rican rainforests. For example, approximately 21 species of hawks are rainforest residents, compared with only four resident forest species in West Virginia; 12 species of pigeons and doves are rainforest residents, but only one species is resident in West Virginia's forests; 44 species of tyrant flycatchers are rainforest residents but only one is a resident in West Virginia (although many species are migratory); and 14 species of wrens are rainforest residents but only two species are resident in West Virginia. This comparison shows that bird diversity is higher not only at the species level but also at higher taxonomic levels and highlights a common theme in tropical–temperate comparisons of diversity: There is often greater representation at higher taxonomic levels such as genus and family in tropical rainforests, and there are generally several higher level groups that are particularly rich in species in rainforests.

High bird species richness is typical of tropical rainforests, and Costa Rica's rainforests may be considered about average with respect to number of species. Approximately 410 species of birds have been found at La Selva in Costa Rica, an area of approximately 1500 ha. However, not all of these species are truly rainforest birds. If only rainforest birds are included (i.e., birds frequenting open water, marshes, cleared areas such as pastures, etc. are excluded), the total list is approximately 338 species. Approximately 314 rainforest species

have been found on Barro Colorado Island and adjacent mainland areas, with a list of all species nearing 400. By contrast, about 554 species of birds have been recorded from Cocha Cashu, Peru, with nearly 400 truly rainforest species. To date, 575 species of birds have been found adjacent to the Explorer's Inn Reserve, an area of 5500 ha in the Amazon rainforest of southern Peru. The majority of these birds are rainforest species.

Other rainforest regions also have high bird diversity. Khao Yai National Park, Thailand, harbors 318 species, approximately 216 of which are true rainforest species. A single plot of <50 ha in southern Vietnam contains 164 species of forest birds. Makokou, Gabon (2000 km^2), has 342 species of birds; 212 species have been recorded from Pasoh, Malaysia (800 ha), and 162 species from Gogol, Papua New Guinea (1000 ha). Although not all species recorded from these last three sites are true rainforest residents, the majority can be considered rainforest species.

Diversity of Mammals

Table 3 highlights the high diversity of rainforest mammals by comparing the approximate number of species that are found in Costa Rica's lowland rainforests with that of lowland forests of West Virginia (including species that were extirpated following European colonization). Costa Rican rainforests contain nearly three times as many species of mammals as do forests in West Virginia. A comparison of families shows the same trend with birds. There are 31 families represented in Costa Rica's rainforest mammal fauna and only 17 represented in West Virginia's forests. The 20 families that have rainforest representatives in Costa Rica but not in West Virginia's forests contribute 97 species. By contrast, six families contributing only eight species have lowland forest species in West Virginia but not in Costa Rican forests. Particularly noteworthy are bats (order Chiroptera). Only 11 species of bats, all from a single family, are found in West Virginia's forests, but approximately 89 species of bats in nine families are found in Costa Rican rainforests. A single family, Phyllostomidae, contains 56 rainforest species. However, not all higher level taxonomic groups are more species rich in rainforests. For instance, the order Insectivora contains nine lowland forest species in West Virginia but only two in Costa Rican rainforests. These two species are confined to the upper elevational limits of rainforest in Costa Rica and reach greater richness in montane forests. Similarly, there are 18 species of carnivores in West Virginia forests (four of which have been extirpated from the state) and only 16 in Costa Rican rainforests.

Mammal diversity at single sites within rainforests is concomitantly high (Voss and Emmons, 1996). However, species lists for most sites are certainly low because of inadequate sampling. Within a 3 km radius in lowland rainforest at Paracou, northern French Guiana, 143 species of mammals have been recorded, and more species, particularly bats, are expected. In La Selva, Costa Rica, the mammal list comprises 117 species, with as many as 138 species expected to occur there based on geographical distributions of species that overlap the area but have not been recorded. As with birds,

Table 3 Approximate numbers of mammal species in lowland rainforests of Costa Rica compared with lowland temperate forests of West Virginia

Order and family	*West Virginia*	*Costa Rica*
Didelphimorphia		
Didelphidae	1	8
Xenarthra		
Bradypodidae	0	1
Megalonychidae	0	1
Dasypodidae	0	2
Myrmecophagidae	0	3
Insectivora		
Soricidae	7	2
Talpidae	2	0
Chiroptera		
Emballonuridae	0	10
Noctilionidae	0	2
Mormoopidae	0	4
Phyllostomidae	0	56
Natalidae	0	1
Furipteridae	0	1
Thyropteridae	0	1
Vespertilionidae	11	9
Molossidae	0	5
Primates		
Cebidae	0	4
Rodentia		
Sciuridae	5	3
Castoridae	1	0
Heteromyidae	0	1
Zapodidae	1	0
Muridae	6	13
Erethizontidae	0	1
Dasyproctidae	0	1
Agoutidae	0	1
Echimyidae	0	2
Lagomorpha		
Leporidae	3	1
Carnivora		
Canidae	5	1
Ursidae	1	0
Procyonidae	1	6
Mustelidae	6	4
Mephitidae	2	0
Felidae	3	5
Perissodactyla		
Tapiridae	0	1
Artiodactyla		
Tayassuidae	0	2
Bovidae	1	0
Cervidae	2	2
Total	58	154

Costa Rica's rainforest mammal diversity can be considered average. Barro Colorado Island's list comprises 113 species, with as many as 144 species expected. Several study sites located in Amazonian rainforests have species lists of >130 species, with nearly 200 species expected from these sites. Makokou's list of mammals is 199 species, Pasoh's is 89 species, and Gogol's is 27 species.

By far the most species-rich order of mammals in these rainforests are bats (Chiroptera). Even in well-studied sites, however, bat communities are incompletely known because of difficulties in sampling species such as those that forage above the canopy. New techniques that allow researchers to identify bats based on ultrasonic recordings are adding substantially to bat inventories at sites that have been sampled for decades by more traditional means such as mist netting.

Geographical Patterns of Animal Diversity

Researchers have searched for both global and regional patterns of animal diversity. In particular, researchers have attempted to identify the major rainforest region that harbors the greatest diversity. The answer depends to some extent on the taxonomic group under consideration. At higher taxonomic levels (e.g., class and order), however, it is clear that the neotropics are generally richest in species. As noted for beetle species richness, Peru and Panama have the greatest numbers of recorded species per volume of canopy and subcanopy foliage, with the Asia-Pacific tropics, particularly Australia, having the lowest richness. Among butterflies, the neotropics again exceed other tropical regions with respect to numbers of species. Comparing species lists of countries or regions of approximately comparable size that occur in the tropics and that have substantial rainforest highlights this point. Panama in the neotropics has approximately 1550 species of butterflies, Liberia in the African tropics has about 720 species, and the Malay Peninsula in the Asia-Pacific tropics has 1031 species. Although not all species in these lists are true rainforest species, the trend of higher butterfly species richness in the neotropics is clear.

Among frogs, a similar trend is evident, with the Neotropical rainforests having the highest species richness. The greatest single-site frog species count is in Ecuador, with other Neotropical sites in Peru and Central America having nearly as many species. Sites in the African and Asia-Pacific tropics have considerably lower frog species richness, with the lowest being in insular sites in the Australopapuan region and rainforests of mainland Asia such as Thailand. However, some sites on the island of Borneo have nearly the same levels of frog diversity seen in some parts of the neotropics.

Ornithologists have long referred South America as the bird continent because of its high bird diversity; approximately one-third of all species of birds occur in this continent. Within single rainforest sites, the highest bird diversity clearly occurs in the neotropics of Amazonia in Peru and surrounding countries on the eastern slope of the Andes. Neotropical rainforests are followed by the African and finally the Asia-Pacific rainforests. Lowest rainforest bird diversity on a per site basis is found in the Australopapuan region, particularly on smaller islands.

Panglobal comparisons of mammal diversity provide a somewhat different view. On a per site basis, mammal diversity is comparable in the neotropics and African tropics. However, bats are better represented in the neotropics, whereas primates are considerably more species rich in the African tropics.

On a regional scale (i.e., within one of the three major rainforest regions), researchers have noted substantial variability in species richness. For instance, with respect to butterflies, a belt of high species richness apparently extends across the neotropics from southern Colombia and the border of Peru and Bolivia at the base of the Andes eastward into the Brazilian states of Rondonia and Acre. However, even within this belt of high butterfly diversity lie areas with apparently lower diversity.

Within the Asia-Pacific tropics, lowest diversity of most taxa is found in the Australopapuan region, particularly on smaller and more isolated islands such as New Caledonia and those of the Bismarck Archipelago, Solomon Islands, Fiji, and Vanuatu. Many of these islands, particularly those in Fiji, New Caledonia, and Vanuatu, lack entire groups such as native rodents and amphibians that have been unable to disperse long distances over open ocean. Other taxa such as butterflies, lizards, and birds, although represented, are depauperate because of island isolation and small island area and a subsequent lack of opportunity for substantial allopatric speciation within an island. Only New Guinea, by far the largest island in the region, harbors taxa that have undergone substantial allopatric speciation.

Soils have a major effect on forest characteristics and productivity and consequently on animal diversity. Although soils of tropical rainforests are notoriously poor in nutrients, there is substantial variability. Where soils change abruptly, differences in forest structure and animal species richness are often evident. For example, the most species-rich sites in the world, those at the eastern base of the Andes, are situated on relatively richer soils. In contrast, the soils of the Guyana Shield in northern South America are particularly poor, and faunas are notably depauperate. Similarly, the sandy soils in parts of the Orinoco drainage of Venezuela are also very poor in nutrients, and animal diversity is often lower than that in surrounding areas with more fertile soils.

Hypotheses of High Species Richness

Explaining the high species richness of tropical rainforests has preoccupied ecologists and evolutionary biologists since naturalists first recognized the tropical-to-polar latitudinal richness gradient and began to appreciate such richness. Many hypotheses have been advanced, but it is important to realize that none of these hypotheses has been adequately tested. It is also important to note that the great richness may be due to a combination of factors, and invoking a single-factor explanation ignores the complexities and intricacies of evolutionary and ecological processes and the long history of geological and climatalogical changes that certainly influenced diversity. Although the high species richness of animal consumers in tropical rainforests no doubt is linked with the high species richness of the plants that function as primary producers, invoking high plant diversity to explain high animal diversity is

largely inadequate because it does not explain the high plant diversity. In other words, the root cause of higher tropical diversity is left unanswered by such circular reasoning.

Researchers have sought explanations that would account for high diversity of all taxa. Several of the more important hypotheses are briefly reviewed here and may be divided into biotic and abiotic hypotheses. Biotic hypotheses emphasize interactions among species to account for high species richness in the tropics relative to temperate and polar regions. Abiotic hypotheses of species richness rely on climatological and geological processes that may promote greater speciation in the tropics. Several hypotheses are specific to terrestrial ecosystems, but latitudinal richness gradients are also evident in freshwater and marine organisms. Ecologists and evolutionary biologists must therefore grapple simultaneously with the phenomena of high species richness in both terrestrial and aquatic systems to develop explanations that will apply to all systems.

The first of the biotic hypotheses is the spatial heterogeneity hypothesis, which states that the tropics are more heterogeneous in space and structurally more complex, thereby providing more niches for animals to exploit and thus a greater number of species. This hypothesis relies on the development of greater vegetational complexity in tropical forests but does not explain adequately how this vegetational complexity arose.

The competition hypothesis states that because of a more benign climate in the tropics, organisms live closer to their respective carrying capacities. Because organisms are closer to the carrying capacity, interspecific competition for limiting resources is more intense, and such competition promotes morphological and ecological divergence and specialization on a narrower range of resources. Specialization then allows more species to coexist, thereby promoting greater diversity. This hypothesis is based on the presumption of a more benign climate, but quantifying how benign a climate is from each organism's perspective is not possible. It is also not possible to measure whether each organism in the tropics lives closer to its respective carrying capacity.

The predation hypothesis is contrary to the competition hypothesis and states that predation and parasitism are more intense in the tropics, which keep organisms below their carrying capacities. Because organisms are kept below their carrying capacities, more resources are available to support a greater number of species. It has not been determined whether or not predation rates are indeed higher in the tropics (at least not for the vast majority of animal species), and again it is not possible to determine whether or not most animals are below their carrying capacities.

The animal pollinators hypothesis states that there is less wind in the tropics, and therefore plants have evolved a greater reliance on animals to pollinate their flowers. This greater reliance promotes a closer relationship with animals and greater specialization to enhance the mutualistic relationship, which therefore promotes diversity. Establishing that there is a greater reliance on animal pollinators because there is less wind in the tropics is not possible, and this hypothesis applies only to terrestrial systems.

The first of the abiotic hypotheses is the ecological time hypothesis, which states that the tropics are older because they have not been subjected to the devastating effects of glaciation that occurred in north temperate regions and that periodically covered entire ecosystems under the great ice sheets. A corollary of this hypothesis is that tropical forests underwent a series of contractions and expansions with cyclic changes in precipitation as water was tied up in the expanding ice sheets and subsequently was released with global warming and contraction of the ice sheets. As the tropical forests contracted, they became isolated as refugia within a matrix of drier savanna and scrub ecosystems. Isolation into refugia promoted allopatric speciation and greater species richness. This corollary was an attractive explanation for pockets of endemism found in organisms such as birds in Amazonia but has become controversial.

The area hypothesis states that the tropics occupy a greater area and that more species occur there because of the species–area relationship. This hypothesis also states that there are more geographical barriers, such as mountain ranges and large rivers, in the tropics, which isolate populations, thereby reducing gene flow and promoting genetic divergence and allopatric speciation.

The energy hypothesis states that there is greater incident of solar radiation striking the tropics and that this solar radiation is more equitably distributed throughout the year. Primary production is consequently greater, which provides more resources for consumers, and more species can consequently coexist. This hypothesis does not address the origin of the greater diversity of species.

The evolutionary speed hypothesis states that generation times are shorter in the tropics and more mutations occur within a given time period. More mutations give natural selection a greater genetic base on which to operate, and speciation rates are consequently higher, which leads to an accumulation of species. This hypothesis and most of the preceding abiotic hypotheses do not explain how more species can coexist in the tropics and consequently do not explain the maintenance of high species richness in the tropics.

More recently, ecologists have focused on the importance of species-specific enemies (seed predators, herbivores, and pathogens) in promoting high tropical plant species richness. Negative feedback, which is an integral component to explain the importance of such enemies in promoting richness, has recently been demonstrated in both shade-house and field experiments (Mangan *et al.*, 2010). Although this hypothesis is specific to plants, such negative feedback may indeed lie at the root of animal diversity as well because all animals ultimately rely on the primary producers, either directly or indirectly.

The Future of Rainforest Animal Diversity

Much has been written about the rapid rate of tropical rainforest clearance by humans and subsequent impacts on animal diversity. Indeed, virtually every volume that has been written in the past two decades on rainforests addresses the potential for a catastrophic decline in global biodiversity. Several biologists have argued that the earth is on the verge of another mass extinction event of the magnitude of the Cretaceous event that eradicated the dinosaurs. Although such arguments are persuasive, it is not possible to predict

accurately the magnitude of an impending extinction episode, particularly because current knowledge of rainforest biodiversity is very limited and no precise estimate of numbers of species is available. However, with approximately 10% of the earth's entire butterfly fauna and untold thousands of species of other animals found within a single 200-ha tract of tropical rainforest, it is not difficult to surmise that the clearing of large tracts of such rainforests indeed will lead to a major decline in global animal diversity.

See also: Amazon Ecosystems. Biodiversity-Rich Countries. Forest Canopies, Animal Diversity. Rainforest Ecosystems, Plant Diversity. Rainforest Loss and Change. Tropical Forest Ecosystems

References

DeVries PJ (1987) *The Butterflies of Costa Rica and Their Natural History: Papilionidae, Pieridae, Nymphalidae,* Vol. I. Princeton, NJ: Princeton University Press.

DeVries PJ (1997) *The Butterflies of Costa Rica and Their Natural History: Riodinidae,* Vol. II. Princeton, NJ: Princeton University Press.

Erwin TL (1982) Tropical forests: Their richness in Coleoptera and other arthropod species. *Coleopterist's Bulletin* 36: 74–75.

Inger RF (1980) Relative abundances of frogs and lizards in forests of Southeast Asia. *Biotropica* 12: 14–22.

Kricher J (1997) *A Neotropical Companion*. Princeton, NJ: Princeton University Press.

Leigh Jr. EG (1999) *Tropical Forest Ecology: A View from Barro Colorado Island*. Oxford: Oxford University Press.

Mabberley DJ (1992) *Tropical Rain Forest Ecology*, 2nd edn. Glasgow: Blackie.

Mangan SA, Schnitzer SA, Herre EA, *et al.* (2010) Negative plant-soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* 466: 752–756.

Stotz DP, Fitzpatrick JW, Parker III TA, and Moskovits DK (1996) *Neotropical Birds: Ecology and Conservation*. Chicago, IL: University of Chicago Press.

Voss RS and Emmons LH (1996) Mammalian diversity in Neotropical lowland rainforests: A preliminary assessment. *Bulletin of the American Museum of Natural History* 230: 1–15.

Whitmore TC (1998) *An Introduction to Tropical Rain Forests*, 2nd edn. New York: Oxford University Press.

Wilson DE and Sandoval A (1996) *Manu: The Biodiversity of Southeastern Peru*. Washington, DC: Smithsonian Institute.

Rainforest Ecosystems, Plant Diversity

Ian M Turner, Singapore Botanic Gardens, Singapore

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 13–23, © 2001, Elsevier Inc.

Introduction

The tropical rainforests grow in conditions of high rainfall (usually more than 2000 mm per year) with few, if any, dry months, where a month would be considered dry if less than 100 mm of rain fell. The temperatures in the lowlands generally average approximately 27 °C, with little variation throughout the year. The day course of temperature with peaks in the afternoon and lows before dawn is frequently the major signal in long-term records—the so-called tropical diurnal climate. Tropical rainforest will grow under climatic conditions of relatively short dry seasons, or even quite long ones if groundwater is available to the trees. As the length and severity of the seasonal drought increases, the proportion of deciduous species in the canopy of the forest tends to increase. When more than about one-third of the canopy trees are drought-deciduous, the forest is no longer tropical rainforest by many definitions.

The plant species of the lowland tropical rainforest are generally intolerant of freezing and frequently sensitive to chilling temperatures. Many major tropical plant families have few members outside the tropics (e.g., the nutmegs, Myristicaceae), possibly because they have not developed cold tolerance.

The tropical rainforests constitute a belt of evergreen vegetation that until relatively recently almost provided a continuous link across the landmasses near the equator. Only in East Africa is rainforest naturally rare. The largest contiguous areas of rainforest were the basins of two major rivers, the Amazon in South America and the Congo in Africa. These essentially continental rainforests can be contrasted with the largely insular forests of the Asia-Pacific region.

Overview of Tropical Rainforest Biodiversity

Biodiversity can be assessed at many different levels ranging from the genetic diversity within populations to landscape or regional heterogeneity. The scale of assessment most widely studied and best understood (though still poorly known) is that of species diversity, which is the focus of this review. The meaning of the term species remains a contentious issue in biology, but in practice in the tropics species are defined by the judgment of taxonomists based largely on the gross morphology of specimens.

There can be little doubt that tropical rainforests are the most species rich of terrestrial ecosystems. Naturally they would cover about 6% of the earth's land surface. Therefore, the high species diversity is not merely a reflection of the tropical rainforests covering an enormous area. The high diversity, relative inaccessibility, and remoteness from major centers of taxonomic research (mostly situated in Europe and North America) have compounded to slow progress in completing the basic inventory of the vascular plants of the humid tropics. Most tropical areas are studied by ongoing projects to provide floras, but progress on most of these is very slow. The *Flora of Tropical West Africa* is the only major tropical flora to have been completed. Others, such as *Flora Neotropica* and *Flora Malesiana*, have published accounts for relatively small proportions of the floras they cover despite decades of work. At this pace, it may take centuries to complete the flora production. However, the state of affairs may not be as bad as it seems because much information is available in monographs, local floras, and checklists. Nevertheless, some tropical areas, notably Central Amazonia, are still poorly known and estimates of plant diversity in terms of named species must be considered as provisional and tentative. Of extant living organisms, vascular plants are among the most completely described, but there remain many undescribed taxa. Exactly how many species remain unnamed cannot be estimated readily, but I suspect that the floristic inventory of the tropical forests has reached a stage at which the number of recognized species will change relatively little because descriptions of new species are heavily compensated for by synonymization among overdescribed taxa.

In **Table 1**, I have attempted to estimate the total plant diversity (in terms of species richness) of the tropical rainforest regions of the world. It is difficult to assess the accuracy of this estimate, and until a species list is compiled it will remain uncertain. However, the derived data provide some interesting results. First, the vascular plant species richness of the tropical rainforests amounts to more than half of the estimated global total of 240,000. This clearly indicates the important contribution of tropical rainforests to global biodiversity. Second, on the broad geographic scale the data allow comparison of the three main rainforest blocks: the neotropics, tropical Africa (including Madagasacar), and the Far Eastern tropics from India into the Pacific. Africa evidently has far fewer species than the other two areas, and the neotropics are the most diverse. To a first approximation, a 3 : 2 : 1 ratio of species richness between tropical America, tropical Asia-Pacific, and tropical Africa can be considered, with an approximately equal proportion of diversity between the Old and New Worlds. At higher taxonomic levels the Neotropical dominance tends to disappear. For instance, it has been estimated that 292 higher plant families are represented in the neotropics, but Malesia, with fewer than half the number of species, has 310 families in its flora (Prance, 1994). Southern Indo-China is the region of the world most diverse in angiosperm families (Williams *et al.*, 1994).

The Rainforest Community

Any visitor to a tropical rainforest can readily appreciate the botanical diversity they contain because of the many forms of

Table 1 Estimates of the number of vascular plant species in the tropical rainforests of the world[a]

Area	*No. of species*	*Notes*
Neotropics	**93,500**	This includes some areas of vegetation other than tropical rainforest, but these are less species rich than the predominant tropical rainforest, so the overestimate is not too large.
African tropics	**20,000**	
Mainland Africa	16,000	Guineo–Congolian and Afromontane phytochoria totals combined. The overlap in species may be compensated by the failure to include species from the small rainforest areas of East Africa.
Madagascar	4,000	Assuming 5000 of the approximately 9500 species are from the island's tropical rainforest and that 80% are endemic (i.e., no overlap with mainland Africa).
Asia-Pacific tropics	**61,700**	
Malesia	45,000	I suspect this to be an overestimate, but New Guinea is still poorly known botanically and recent studies indicate high levels of diversity and endemism.
Indo-China and adjacent areas	10,000	This is an educated guess on my part. There have been few attempts to consider the flora of this region as a single entity.
Southern India	4,000	Largely the flora of the Western Ghats rainforests.
Sri Lanka	1,000	The endemic rainforest element.
Australia	700	The endemics of the Queensland rainforests that are largely Malesian floristically
Pacific Islands	1,000	There is true rainforest on some of the Pacific Islands but because of their isolation they are relatively poor in species.
Total	**175,200**	

[a]Data mostly from WWF/IUCN (1994–1997) and Johns (1995), with permission from IUCN.

plants to be observed ranging from tiny herbs to gigantic trees. There have been numerous attempts to classify plant species by their life-forms. Although there is no universally accepted or universally applicable system, most of these classifications use a similar set of criteria for determining groups within the structure. This set, at its simplest, generally includes plant mature stature and degree of woodiness, need for mechanical support from other plants, and information on trophic requirements. A classification of plant life-forms found in the tropical rainforest, based on these criteria, is given in **Table 2**.

An area of lowland forest in Brazil would have relatively few plant species in common with a site in Costa Rica and very few indeed in common with sites in Congo or Indonesia, but all would be classified as lowland tropical rainforest. Despite this strong geographic dissimilarity in species composition, there are floristic affinities throughout the world. These occur at higher taxonomic levels, particularly at the rank of family, but also include some strong generic similarities in makeup. Tropical floras tend to be more similar to each other, in terms of species abundance in common families, than to subtropical and temperate floras (Turner, 1997). **Table 3** summarizes floristic data by life-form and geographic area. Gentry (1988, 1993), in particular, emphasized the consistent patterns of family abundances, in terms of numbers of species represented in inventories of forest areas, across the tropics. Among woody stems, families such as Leguminosae, Rubiaceae, Euphorbiaceae, and Myristicaceae are nearly always among the most speciose at a lowland tropical rainforest site. Pteridophytes,[1] Araceae, and Zingiberaceae are frequently the most common ground herbs, and Orchidaceae and Pteridophytes almost invariably dominate the epiphytes. **Table 3** lists 17 very large pantropical genera that are well represented in all regions, again emphasizing the floristic similarity of diverse tropical areas. However, there are also differences between regions.

The "dipterocarp substitution" has been noted as a major anomaly in the Asia-Pacific region (Gentry, 1988). That is, in this region the Dipterocarpaceae substitute for Leguminosae as the dominant woody family. The Dipterocarpaceae are a pantropical family, but the vast majority of its species occur in the region from India east to Wallace's line. In this area, dipterocarps form a majority of the large trees in lowland forests. However, this does not translate into dominance in terms of species abundance in the complete flora, or even the tree flora, unless a very large stem diameter limit is chosen. Legumes are not as numerically important as in American or African floras, but nevertheless they are still well represented, so the dipterocarp substitution is not as important a phenomenon as might be thought. Additionally, East Malesian rainforests have only a small representation of Dipterocarpaceae. **Table 3** indicates other regional variations in representation of more important families. Perhaps the most notable of which are the virtual restriction of the important ground herb and epiphyte family Bromeliaceae to the neotropics; the importance of Bignoniaceae and Sapindaceae as climbers in the same region; the paucity of palms, particularly understory species, in Mainland Africa (although not Madagascar); and the abundance of Dichapetalaceae. There are many other smaller families with restricted ranges in the tropics.

Because of the range of spatial scales that can be considered and the relative inclusivity of the sampling, many approaches are available for assessing plant species richness of tropical forests. These include plot, transect, and plotless sampling at small spatial scales that may include all plant taxa present, larger plot sampling that considers only trees and climbers exceeding a certain lower stem diameter limit, and local or regional floristic accounts based on intensive collection within the area concerned. Relatively few efforts have been made to identify all the individual plants present within an area of rainforest largely because few people possess the botanical expertise to identify the enormous amount of juvenile and sterile plant material that such total inventories generate. Recently, an attempt was made to enumerate all the plant species occurring in 1 ha of forest in lowland Amazonian Ecuador.

[1]Pteridophyta are frequently accorded "family status" in these analyses.

Table 2 Classification of vascular plant life-forms in the tropical rainforest and a summary of their taxonomic composition

Life-form group	*Definition*	*Examples from tropical rainforest*	*Subdivisions of the group*
Trees	Autotrophic, woody, and mechanically independent (self-supporting) plants	Many taxa within the dicotyledons, palms among the monocots, a few conifers and gnetums, some cycads and tree ferns.	Mature size and form have often been used to subdivide the tree category. Unfortunately, no consistent definition has been employed.
Herbs	Autotrophic, mechanically independent plants with little or no lignification and secondary thickening of stems or roots	Many dicot and monocot taxa, many ferns; some characteristically epiphytic species are facultatively terrestrial when they fall from the trees.	The giant herbs, largely from the Zingiberales, can be distinguished by size, and in ecological terms may have more in common with trees than the other herbs.
Climbers	Autotrophic, woody, or herbaceous plants, rooted in the soil and dependent on other plants for support, at least in later stages	A wide range of taxa including dicots, monocots (e.g., palms and aroids), gymnosperms (*Gnetum* spp.), and ferns.	Woody climbers (lianas) are sometimes distinguished from herbaceous ones; means of climbing can also be employed.
Epiphytes	Autotrophic, woody, or herbaceous species habitually growing on other plants and not rooting in the soil	Epiphytes are dominated by three groups–orchids, pteridophytes, and aroids; bromeliads are an almost exclusively Neotropical epiphyte group.	Division into herbaceous and woody forms might be possible. Often, a distinction is made between species on their relative preference for crown and bole locations on their host trees.
Hemiepiphytes	Plant, mostly woody, that begin life as epiphytes but may eventually grow roots down to the ground	Strangling figs (*Ficus* spp.) are the largest group of hemiepiphytes.	
Plant parasites	Plants that directly parasitise other plants	Direct plant parasitism is confined to the dicots, and with the exception of *Cassytha* (Lauraceae) is restricted to probably advanced groups. Mistletoes (Loranthaceae and Viscaceae) and related root-parasitic groups (Santalaceae and Olacaceae) and the mostly herbaceous Scrophulariaceae form the two largest groups.	The parasites can be divided into hemiparasites that are chlorophyllous but possess haustorial connection to host xylem (e.g., mistletoes), holoparasites that are achlorophyllous and entirely heterotrophic (e.g., dodders, broomrapes, and Balanophoraceae, and endoparasites that live inside the host and are only visible externally as reproductive structures (Rafflesiaceae).
Mycotrophs	Heterophic plants that derive carbohydrate from a fungus; the fungus may derive its carbon from dead material or from another living plant	All orchids are mycotrophic as seedlings, but most become autotrophic as they develop. The small proportion that remain entirely heterotrophic represent the largest group of mycotrophic species. Other monocot mycotrophs include members of the families Petrosaviaceae, Triuridaceae, Iridaceae, Burmanniaceae, and Corsiaceae. Ericaceae, Gentianaceae, and Polygalaceae are the only dicot families with mycotrophic species.	Detailed investigations can distinguish the "parasitic" from the "saprophytic" species but these have been conducted for relatively few of the tropical species.

A total of 942 species were encountered in the 100×100-m plot (Balslev *et al.*, 1998). Lower montane forest at 1500 m in Java was estimated to contain 333 species in 1 ha (Meijer, 1959). A single plot of 10×10 m in Costa Rica contained 233 species (Whitmore *et al.*, 1985), more than any among eight plots of the same size recorded in Gabon (74–130 spp.; Reitsma, 1988). Duivenvoorden (1994) enumerated all the vascular plant species in a series of ten 0.1-ha plots in Colombia. The most species-rich plot contained 313 species. Gentry and Dodson (1987a) recorded 365 species from 0.1 ha at Rio Palenque, Ecuador, but this was from 10 noncontiguous subplots. Two other sites in Ecuador, sampled with the same methods, were considerably less diverse, with 169 and 173 species recorded (Gentry and Dodson, 1987b).

Most estimates of plant diversity in tropical forests involve the enumeration of stems above a certain diameter (or girth) at breast height (1.3 m) or above obvious outgrowths, such as buttresses or stilt roots. Many different sampling methods and size limits have been employed in such investigations, but fortunately a few approaches have been applied relatively

Table 3 Summary of the floristic composition of tropical rainforest, with a breakdown by the major geographic regions

	All tropics	*Regional breakdown*[a]		
		Neotropics	*Africa (including Madagascar)*	*Asia-Pacific*
Important tree families	Leguminosae	**Vochysiaceae**	Dichapetalaceae	**Dipterocarpaceae**
	Lauraceae	Bignoniaceae	Olacaceae	Myrtaceae
	Annonaceae	**Cyclanthaceae**		
	Rubiaceae	Lecythidaceae	Moraceae	
	Moraceae		Palmae	
	Myristicaceae	Ebenaceae	Fagaceae	
	Sapotaceae	**Pandanaceae**	Bombacaceae	
	Meliaceae			
	Palmae			
	Euphorbiaceae			
Important herb families	Pteridophyta	Bromeliaceae		
	Araceae	Gramineae		
	Zingiberaceae	Heliconiaceae		
	Cyperaceae			
	Rubiaceae			
	Gesneriaceae			
	Orchidaceae			
Important climber families	Asclepiadaceae	Compositae	Annonaceae	Annonaceae
	Convolvulaceae	Bignoniaceae	Dichapetalaceae	Palmae
	Leguminosae	Sapindaceae		**Nepenthaceae**
	Araceae	Malpighiaceae		
	Apocynaceae	Passifloraceae		
	Cucurbitaceae			
	Rubiaceae			
Important epiphyte families	Orchidaceae	**Bromeliaceae**		Asclepiadaceae
	Pteridophyta	**Cyclanthaceae**		Rubiaceae
	Araceae	**Cactaceae**		
	Piperaceae			
	Melastomataceae			
	Gesneriaceae			
Important genera (>500 spp.)	*Asplenium*	*Anthurium*		*Dendrobium*
	Begonia	*Epidendrum*		*Eria*
	Bulbophyllum	*Eugenia*		*Rhododendron*
	Clerodendrum	*Lepanthes*		*Syzygium*
	Croton	*Maxillaria*		
	Cyathea	*Miconia*		
	Dioscorea	*Oncidium*		
	Ficus	*Peperomia*		
	Habenaria	*Pleurothallis*		
	Impatiens	*Stelis*		
	Justicia			
	Phyllanthus			
	Piper			
	Psychotria			
	Schefflera			
	Selaginella			
	Solanum			

[a]Names in bold indicate families entirely restricted to a particular geographic region, or nearly so. Names underlined indicate families notable for their scarcity or absence from a particular region. Names in bold and underlined are entirely absent from a particular region, or almost so.

consistently and allow intersite comparisons of diversity. Three commonly employed methods of assessing plant diversity in tropical forests produce scales of species-richness values that are comparable, at least in the broad view (**Table 4**). I used these data to define categories of relative diversity for lowland tropical forests. Forests of very high diversity have thus far only been reported from the New World. They are known from the lowland forests of the Pacific coast of Colombia in Chocó Province and the Upper Amazon Basin stretching in a belt below the eastern slopes of the Andes from Ecuador to Peru. The occurrence of 250 species in a 1-ha plot near Manaus in the center of the Brazilian Amazon (Ferreira and Rankin-de-Mérona, 1998) countered the impression that the forests of the middle and lower Amazon were of uniformly

Table 4 Categories of plant diversity in lowland tropical rainforests

Species-richness class	*No. of species*[a]	*Examples from different geographic regions*[a]		
		Neotropics	*Africa*	*Asia-Pacific*
Low	Less than 100	Barro Colorado Island, Panama;[f] San Carlos, Venezuela;[e] Belém, Brazil[k]	Kibale, Uganda;[e] Kade, Ghana;[g] Lopé, Gabon[f]	Mulu (limestone), Malaysia[o]
Intermediate	100–199	Many, including La Selva, Costa Rica[e]	Makoukou, Gabon;[h] Korup, Cameroon;[h] Oveng, Gabon[f]	Davies River, Queensland;[j] Xishuangbanna, China;[l] Danum Valley, Malaysia[m]
High	200–249	Coca, Ecuador;[g] Jatun Sacha, Ecuador;[g] Mishana, Peru[i]	Nosy Mangabe, Madagascar[h]	Pasoh, Malaysia;[f] Lambir, Malaysia;[e] CMBRS. Papua New Guinea[n]
Very high	250 or more	Cuyabeno, Ecuador;[b] Yanamono Peru;[e] Bajo Calima, Colombia;[c] Tutunendo, Colombia;[c] Manaus, Brazil[d]	None recorded	None recorded

[p]Reitsma (1988) (*x*).

[a]Species richness can be measured in one of three ways: (*x*) Species represented among the trees and lianas of 10 cm dbh (or 30 cm gbh) or greater on a 100 × 100-m (1 ha) plot; (*y*) species represented among the first 500 tree and liana individuals of 10 cm dbh (or 30 cm gbh) or greater on a contiguous plot; and, (*z*) species represented among the trees and lianas of 2.5 cm dbh or greater on a 0.1-ha sample consisting of ten 50 × 2-m transects.

[b]Valencia, *et al.* (1994) (*x*).

[c]Faber-Langendoen and Gentry (1991) (*x*), Gentry (1986) (*z*).

[d]Ferreira and Rankin-de-Mérona 1998 (*x*).

[e]Phillips, *et al.* (1994) (*y*).

[f]Condit, *et al.* (1996) (*x*).

[g]Valencia, *et al.* (1998) (*x*).

[h]Gentry (1993) (*z*).

[i]Gentry and Emmons (1987) (*z*).

[j]Gentry (1988) (*z*).

[k]Black, *et al.* (1950) (*x*).

[l]Cao, *et al.* (1996) (*x*).

[m]Newbery, *et al.* (1992) (*x*).

[n]Wright, *et al.* (1997) (*x*).

[o]Proctor, *et al.* (1983) (*x*).

lower diversity than the uppermost reaches. High-diversity forests are well represented in the South American tropics and are of similar species richness as the lowland dipterocarp forests of West Malesia and at least some of the lowland forests of New Guinea. To date, only Madagascan rainforests have been shown to attain high-diversity status among those from Africa. The most diverse mainland African rainforests occur in Cameroon and Gabon in West Africa, but these are substantially less rich than many forest sites in the neotropics or the Asia-Pacific region and on a global scale could only be classed as of intermediate diversity. Many sites where detailed ecological research has been conducted, such as Barro Colorado Island, Panama, San Carlos de Rio Negro, Venezuela, and Kibale, Uganda, are low-diversity rainforests. These geographic patterns in local species richness appear to mirror quite closely the tropical regions of the map of plant species diversity per 10,000 km^2 (Barthlott *et al.*, 1996; see http://www.botanik.uni-bonn.de/biodiv/phytodiv.htm).

An important point is that the range of values for diversity within a locale, i.e., an area of approximately identical climate, can cover much of the general range exhibited across wide geographic and climatic gradients. For example, at Caquetá in Colombia total vascular plant diversity in ten 0.1-ha plots was 40–313 species (Duivenvoorden, 1994), at Tambopata in Peru five 1-ha plots had 60–173 species for stems of 10 cm dbh or larger (Phillips *et al.*, 1994), and at Gunung Mulu National Park in Sarawak four plots varied between 73 and 223 species for the same plot and tree sizes as those of the previous example (Proctor *et al.*, 1983). In all these cases the lowest diversity plots were ones on predictably harsher substrates, such as permanently waterlogged soils (Caquetá and Tambopata) or limestone karst (Gunung Mulu). Soil physicochemical properties are highly influential on the composition of the forest community and clearly affect species richness. However, there are also other environmental factors that can lead to gradients in plant diversity (**Table 5**). Rainfall appears to be

Table 5 Environmental trends in tree species richness per unit area in tropical rainforest

Number of species increases with total precipitation.
Number of species decreases with increasing seasonality of rainfall.
Number of species increases with soil fertility, although diversity may level off, or even decline, on the most fertile sites.
Number of species declines with altitude.

Table 6 Representation of major life-form groups in the floras of different tropical rainforest sites[a]

Site	*Mean annual rainfall (mm)*	*Total No. spp.*	*Percentage of species in major life-form groups*			
			Trees	*Herbs*	*Climbers*	*Epiphytes*
La Selva, Costa Rica[b]	3950	1450	39	27	7	27
Rio Palenque, Ecuador[c]	3000	1055	25	37	16	23
Barro Colorado Island, Panama[d]	2650	966	43	21	16	20
Manu floodplain, Peru[e]	2050	1217	57	13	18	12
Singapore (forest flora)[f]	2000	1673	56	13	18	13
Jauneche, Ecuador[c]	1850	608	28	38	22	12

[a]Sites are listed in descending order of mean total annual rainfall.
[b]Hammel (1990).
[c]Gentry and Dodson (1987b).
[d]Foster and Hubbell (1990).
[e]Foster (1990).
[f]Turner (1994).

the major determinant of diversity when old-growth sites on freely draining, nonextreme soils are compared. A relatively strong positive correlation between average annual precipitation and diversity has been observed in many tropical data sets (Givnish, 1999), although good data for total plant diversity is lacking so that only estimates for large woody stems can be used. This correlation is evident over the range of true lowland tropical rainforest and is not merely a comparison of diversity between drier forest formations and proper rainforest. Rainfall seasonality (the presence and severity of a dry season) is of key importance, but it is difficult to quantify this from standard meteorologial data, and total annual rainfall is a good predictor of diversity in most cases. It is not immediately evident why further increasing the wetness of a wet climate should continue to increase the local species diversity, although highly ombrogenous sites may favor epiphytes. As can be seen from the data presented in **Table 6**, the proportion of the flora that epiphytes represent increases quite sharply with average annual rainfall.

Why Are Tropical Rainforests so Diverse?

Lowland tropical rainforests undoubtedly represent the zenith of terrestrial biodiversity, probably with more than half the world's biota on less than a twentieth of the land surface. High plant diversity may well be a major contributory factor to the overall abundance of species in the rainforest through the specialization of a myriad of invertebrate species to each plant. However, why are there so many plant species in the tropics? In evolutionary terms, high species diversity must imply many speciation events and relatively few extinctions of species because diversity is the balance between these two processes. Therefore, the tropics may be richer in species than other areas of the world because there has been more rapid speciation here and/or a slower rate of extinction. It is possible that the benign and equable climate of the tropics makes it less likely that a species will become extinct than in the harsher climates elsewhere. The positive correlation between forest diversity and annual precipitation within the tropics also indicates a link between environmental adversity and reduced diversity.

Switching from an evolutionary to an ecological perspective, tropical rainforest diversity maintains a perplexing fascination. A commonly held view of natural communities is that they consist of populations of species in which intraspecific competition is stronger than interspecific competition. In other words, species occupy separate niches within the community and have reached a competetive equilibrium. An equilibrium model of the structure of a lowland tropical rainforest plant community requires that all the species present, including the many species of similar life-form and basic ecology such as canopy trees, or understory treelets, occupy separate niches. Given the simple requirements of plants—light, water, and nutrients from the soil—it is not immediately obvious where so many niches could come from to support the observed diversity. Differences in requirements for reproduction and regeneration may allow more axes from which to partition niches. These niches could be extremely narrow because of the constancy of the tropical environment allowing persistence of small or sparse populations. I suspect, however, that few tropical forest ecologists believe that all the plant species in a forest are ecologically distinct. A more plausible hypothesis is that most species have ecologically equivalent or similar species with which they co-occur. It is worth noting that a high proportion of the species present in a rainforest exist as populations intermingled with those of congenerics. For instance, of the 814 tree species recorded from 50 ha at Pasoh, Malaysia, 82% had a congeneric present in the plot and 70% had a congeneric present in the same broad stature class (Manokaran *et al.*, 1992). For descriptive convenience, ecologists may group species into guilds of ecologically similar entities such as pioneer trees or bole climbers, but it must be remembered that these are not necessarily discrete groupings in terms of niche space in the community. They may represent arbitrary divisions of whatever axes of differentiation are being considered.

If competetive equilibium has not been obtained then something must be preventing competitive exclusion within the community from occurring. One possible process by which this could happen is through frequency-dependent mortality operating within the community setting an upper limit on the abundance of a species and thus making it unlikely that it can totally exclude other species from the

landscape. This compensatory mortality, particularly among tropical trees, is often referred to as the Janzen–Connell effect after the two researchers who independently put forward the hypothesis that species-specific predators could produce the density-dependent mortality required to promote the coexistence of many plant species in a forest. Invertebrate seed and seedling predators, and plant pathogens such as fungi, have frequently been observed to cause severe mortality of germinants and juveniles occurring in high densities near parent trees. Detailed analyses of spatial patterns of sapling mortality on Barro Colorado Island, Panama, showed evidence of compensatory mortality in recruitment for 67 of 84 of the most common tree species (Wills *et al.*, 1997). Rare species, which make up the bulk of the diversity in the forest, may not reach sufficient adult densities to suffer compensatory mortality. However, they may just be too rare to provide statistically significant results with the available data sets. Givnish (1999) hypothesized that compensatory mortality may increase in intensity with rainfall because pest and pathogen populations are less likely to be checked by a dry season in an ever-wet forest. This may explain, at least partially, the observed increase in forest diversity with precipitation.

An alternative nonequilibrium view of the community is that chance is the major determinant of community composition and relative abundance. The majority of species are so infrequent that they very rarely come into direct competition with other ecologically equivalent, similarly sparse species. Chance, and not competitive ability, is an important determinant of which species establishes an adult successfully at a given site because most species are unable to produce enough seeds, disperse them well enough, or survive the early onslaughts of pests and pathogens to ensure that they have a candidate for any vacancy in the community (Hubbell *et al.*, 1999). Many potential episodes of interspecific competition may be won by default, and by competitively inferior species, because no adversaries were present at the right time and place.

At present we have no simple answer to the question of why there are so many species in the rainforests and how they manage to coexist. Nor does it seem likely that there is a simple answer. Probably all the factors mentioned previously, and possibly others not yet considered, are operating within any rainforest but probably not affecting all members of the community simultaneously.

The Value of Tropical Rainforest Plant Diversity

If tropical rainforests are to be conserved, conservationists need to convince a wide range of people that tropical rainforests have sufficient value to merit conservation. In attempting to do so, conservationists face a series of dilemmas. The first is whether to consider value as an economic parameter or not. The global economy hardly has a good record for preventing extinction or supporting sustainable exploitation, but the reality is that people in the tropics exploit the rainforest so that they can improve their standard of living, even if only in the short term. Conservationists can promote the development of a truly sustainable global economy, but in the meantime they must address the immediate problem of impending extinctions. Purely economic analyses generally favor rapid liquidation of forest resources, although a notable exception was a comparison of exploitation of minor forest products, such as fruits and latexes, with logging and timber tree plantations in Peru (Peters *et al.*, 1989) which showed that the relatively sustainable nontimber production generated greater revenues and thus put the highest net present value on the resource. However, the study has been criticized as being atypical in having a forest with a very high density of fruit and latex trees relatively close to a major population center providing a ready market for the products.

Are there any economic benefits to be gained from leaving rainforest undisturbed? A group of people for which the answer to this question is generally in the affirmative are traditional forest dwellers. Few of the tropical rainforest regions are without people who live in the forest, either making their living by hunting and gathering in the forest or by farming on a small scale, usually as some form of traditional shifting agriculture. Large-scale forest clearance dispossesses the traditional forest dwellers of their entire resource base and often much of their social fabric.

Tourism potential is perhaps the most economically lucrative opportunity offered by virgin tropical rainforest, but fortunately we are not yet in a situation in which every remaining area of forest can attract enough paying visitors to make conservation profitable. However, at a local level, ecotourism probably offers the best hope of generating an income from unexploited forest. Plant biodiversity, in addition to providing the glorious backcloth, is rarely likely to be the focus of ecotourism, although *Rafflesia* may be a notable exception.

The value of the tropical rainforest as a storehouse of potentially useful genetic diversity is probably the best claim for a high intrinsic value to be placed on biodiversity. For vascular plants the potential uses are largely of three main types: medicinal, novel crops, and novel genes for improvement of existing crop varieties. A new drug that is superior to other available medicines can be worth millions of dollars, not to mention that it can have tremendous humanitarian value in the reduction in suffering. New or improved crops also have substantial economic value. Given the massive concentration of biodiversity in the tropics, it is almost certain that superior and economically rewarding new products await discovery in the rainforest. However, searching the haystack for the few golden needles is a difficult, expensive, and often fruitless procedure. Island biogeography theory predicts that a 90% loss of habitat equates to a 10% reduction in diversity. Therefore, the potential economic returns from undiscovered rainforest products seem unlikely to provide strictly economic arguments for preventing the destruction of tropical forests until the area left is very small.

Forests are important for their so-called "ecosystem services"—the amelioration of atmospheric composition, water quality, and local climatic regime among a variety of beneficial emergent properties of the rainforest community. Despite considerable speculation and some very simplistic experimentation, there is no solid evidence that species richness per se has a major influence on these community-level processes. Most species in the rainforest are very rare, so it seems unlikely that the random removal of a certain proportion of

the flora would consistently influence oxygen production, carbon sequestration, nutrient cycling, and transpiration.

In the prevailing international financial framework it is difficult to find a sound, strictly economic, argument for conserving tropical rainforest. When environmental and ecological sustainability become the key goals of the global economy, conservation will make economic sense. The biodiversity of the tropics has a unique and irreplaceable intrinsic value as well as cultural, spiritual, educational, and humanitarian worth to millions of people, many of whom have never been in a tropical rainforest.

The Fate of the Tropical Rainforest

The human population of the tropics has already destroyed the rainforest occupying half the area it would be expected to cover in today's climatic regime. Much of the remaining forest is fragmented into small remnant patches, damaged, burnt, or defaunated. Extensive, relatively pristine forests are largely confined to Amazonia, the Congo Basin, and New Guinea. Most tropical rainforest nations have little remaining forest, and although they generally have designated national parks, forest reserves, or other protected area status, the actual protection that these sites receive is often inadequate to prevent deterioration. There are of course exceptions, and excellent tropical rainforest nature reserves do exist.

There is no sign that the global rates of deforestation in the tropics are declining. Tropical forests are cleared for a variety of purposes and the relative importance of these varies between regions (Bawa and Dayanandan, 1997). In Africa and parts of South America, the demand for land by a large population of poor people who need to subsist on what they can grow leads to deforestation. Clearance for cattle ranching is also common in Latin America. Establishment of plantations for commercial tree crops, such as oil palm (*Elaeis guineensis*) and rubber (*Hevea brasiliensis*), is a more destructive force in the Asian tropics. Human populations continue to grow rapidly throughout the tropics, making it unlikely that the pressure for land for food production or cash crops will decline.

There is increasing evidence that the tropical rainforest community, particularly its botanical component, is more resistant to human disturbance and fragmentation than has been feared. Studies have shown that selective logging does not reduce tree diversity very much (Cannon *et al.*, 1998) and long-isolated forest fragments maintain relatively high proportions of their original floras (Turner and Corlett, 1996). This information is met with some ambivalence by conservationists, who have tended to stress the fragility of tropical forest ecosystems. It is positive and hopeful in that it provides some scientific support to the aspirations for sustainable exploitation of tropical forests and the persistence of the native biota in nations in which only fragments remain. However, there is a real concern that knowledge of this relative robustness will be used as an excuse by those with political power to allow greater levels of exploitation and deforestation, possibly including inside currently protected areas.

From a pragmatic viewpoint, it is obvious that a multiplicity of approaches will be required to conserve the tropical rainforest biota. Where there are still large areas of relatively undisturbed forest with low human population densities, efforts to establish large inviolate parks are warranted, but in areas with a landscape dominated by human activity the remnant patches of the original forest have an important conservation value in their own context. These fragments may not contain all the original species, but they still support a rich diversity that otherwise might be completely lost. In highly degraded tropical landscapes, which are very common, forest fragments may also play a vital role in the reforestation of these derelict lands by acting as foci for tree establishment and sources of planting material.

See also: Amazon Ecosystems. Biodiversity-Rich Countries. Deforestation and Land Clearing. Forest Canopies, Plant Diversity. Plant Biodiversity, Overview. Rainforest Ecosystems, Animal Diversity. Rainforest Loss and Change. The Value of Biodiversity. Tropical Forest Ecosystems

References

Balslev H, Valencia R, Paz y Miño G, Christensen H, and Nielsen I (1998) Species count of vascular plants in one hectare of humid lowland forest in Amazonian Ecuador. In: Dallmeier F and Comiskey JA (eds.) *Forest Biodiversity in North, Central and South America, and the Caribbean*, pp. 585–594. Paris: UNESCO.

Barthlott W, Lauer W, and Placke A (1996) Global distribution of species diversity in vascular plants: Towards a world map of phytodiversity. *Erdkunde* 50: 317–327.

Bawa KS and Dayanandan S (1997) Socioeconomic factors and tropical deforestation. *Nature* 386: 562–563.

Black GA, Dobzhansky T, and Pavan C (1950) Some attempts to estimate species diversity and population density of trees in Amazonian forests. *Bot. Gaz.* 111: 413–425.

Cannon CH, Peart DR, and Leighton M (1998) Tree species diversity in commercially logged Bornean rainforest. *Science* 281: 1366–1368.

Cao M, Zhang J, Feng Z, Deng J, and Deng X (1996) Tree species composition of a seasonal rain forest in Xishuangbanna, Southwest China. *Trop. Ecol.* 37: 183–192.

Condit R, Hubbell SP, LaFrankie JV, *et al.* (1996) Species–area and species individual relationships for tropical trees: A comparison of three 50-ha plots. *J. Ecol.* 84: 549–562.

Duivenvoorden JF (1994) Vascular plant species counts in rain forests of the middle Caquetá area, Colombian Amazonia. *Biodivers. Conserv.* 3: 685–715.

Faber-Langendoen D and Gentry AH (1991) The structure and diversity of rain forests at Bajo Calima, Chocó Region, Western Colombia. *Biotropica* 23: 2–11.

Ferreira LV and Rankin-de-Mérona JM (1998) Floristic composition and structure of a one-hectare plot in terra firme forest in Central Amazonia. In: Dallmeier F and Comiskey JA (eds.) *Forest Biodiversity in North, Central and South America, and the Caribbean*, pp. 649–662. Paris: UNESCO.

Foster RB (1990) The floristic composition of the Rio Manu floodplain forest. In: Gentry AH (ed.) *Four Neotropical Forests*, pp. 99–111. New Haven, CT: Yale Univ. Press.

Foster RB and Hubbell SP (1990) The floristic composition of the Barro Colorado Island forest. In: Gentry AH (ed.) *Four Neotropical Forests*, pp. 85–98. New Haven, CT: Yale Univ. Press.

Gentry AH (1986) Species richness and floristic composition of Choco Region plant communities. *Caldasia* 15: 71–91.

Gentry AH (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann. Missouri Bot. Gardens* 75: 1–34.

Gentry AH (1993) Diversity and floristic composition of lowland tropical forest in Africa and South America. In: Goldblatt P (ed.) *Biological Relationships between Africa and South America*, pp. 500–547. New Haven, CT: Yale Univ. Press.

Gentry AH and Dodson C (1987a) Contribution of nontrees to species richness of a tropical rain forest. *Biotropica* 19: 149–156.
Gentry AH and Dodson CH (1987b) Diversity and biogeography of neotropical vascular epiphytes. *Ann. Mis. Bot. Gardens* 74: 205–233.
Givnish TJ (1999) On the causes of gradients in tropical tree diversity. *J. Ecol.* 87: 193–210.
Hammel B (1990) The distribution of diversity among families, genera, and habit types in the La Selva flora. In: Gentry AH (ed.) *Four Neotropical Forests*, pp. 75–84. New Haven, CT: Yale Univ. Press.
Hubbell SP, Foster RB, O'Brien ST, *et al.* (1999) Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* 283: 554–557.
Johns RJ (1995) Endemism in the Malesian flora. *Curtis's Bot. Mag.* 12: 95–110.
Monokaran N, LaFrankie JV, Koehummen KM, *et al.* (1992) Stand table and distribution of species in the 50-ha research plot at Pasoh Forest Reserve. Forest Research Institute of Malaysia, Research Data 1: 1–454.
Meijer W (1959) Phytosociological analysis of montane rainforest near Tjibodas, West Java. *Acta Bot. Neerl.* 8: 277–291.
Newbery DM, Campbell EJF, Lee YF, Ridsdale CE, and Still MJ (1992) Primary lowland dipterocarp forest at Danum Valley, Sabah, Malaysia: Structure, relative abundance and family composition. *Philos. Trans. R. Soc. London B* 335: 341–356.
Peters CM, Gentry AH, and Mendelsohn RO (1989) Valuation of an Amazonian rain forest. *Nature* 339: 655–656.
Phillips OL, Hall P, Gentry AH, Sawyer SA, and Vázquez R (1994) Dynamics and species richness of tropical rain forests. *Proc. Natl. Acad. Sci. USA* 91: 2805–2809.
Prance GT (1994) A comparison of the efficacy of higher taxa and species numbers in the assessment of biodiversity in the neotropics. *Philos. Trans. R. Soc. London B* 345: 89–99.
Proctor J, Anderson JM, Chai P, and Vallack HW (1983) Ecological studies in four contrasting lowland rain forests in Gunung Mulu National Park, Sarawak I. Forest environment, structure and floristics. *J. Ecol.* 71: 237–260.
Reitsma JM (1988) Forest vegetation of Gabon. *Tropenbos Technol. Ser.* 1: 1–103.
Turner IM (1994) The taxonomy and ecology of the vascular plant flora of Singapore: A statistical analysis. *Bot. J. Linnean Soc.* 114: 215–227.
Turner IM (1997) A tropical flora summarized—A statistical analysis of the vascular plant diversity of Malaya. *Flora* 192: 157–163.
Turner IM and Corlett RT (1996) The conservation value of small, isolated fragments of lowland tropical rain forest. *TREE* 11: 330–333.
Valencia R, Balslev H, and Paz y Miño G (1994) High tree alpha diversity in Amazonian Ecuador. *Biodiversity Conserv.* 3: 21–28.
Valencia R, Balslev H, Palacios W, *et al.* (1998) Diversity and family composition of trees in different regions of Ecuador: A sample of 18 one-hectare plots. In: Dallmeier F and Comiskey JA (eds.) *Forest Biodiversity in North, Central and South America, and the Caribbean*, pp. 569–584. Paris: UNESCO.
Whitmore TC, Peralta R, and Brown K (1985) Total species count in a Costa Rican tropical rain forest. *J. Trop. Ecol.* 1: 375–378.
Wills C, Condit R, Foster RB, and Hubbell SP (1997) Strong density- and diversity-related effects help to maintain species diversity in a neotropical forest. *Proc. Natl. Acad. Sci. USA* 94: 1252–1257.
World Wildlife Fund/International Union for the Conservation of Nature (WWF/IUCN) (1994–1997). *Centres of Plant Diversity*, vol. 3. Cambridge, UK: IUCN.
Wright DD, Jessen JH, Burke P, and Gomez de Silva Garza H (1997) Tree and liana enumeration and diversity on a one-hectare plot in Papua New Guinea. *Biotropica* 29: 250–260.

Rainforest Loss and Change

Karndeo D Singh, Academy of Forests and Environmental Sciences, New Delhi, India

© 2013 Elsevier Inc. All rights reserved.

Glossary

Change matrices A way to relate area of land-use classes at the beginning with that at the end of a period. From such matrices, it is possible to follow the exact path of change viz. how much land was transferred from a given use to other uses, how much came into that use, and what remained unchanged during the period.

Deforestation The change from forest to other land uses such as agriculture or ranching; or depletion of forest crown cover to less than 10% density. Because shifting cultivation involves a change of land use, it is considered deforestation.

Forest change processes A collective term for all activities on the forest land which affect the stand or site and, in particular, the biological diversity. These processes include forest exploitation, forest fragmentation, and establishment of new plantations.

Historic rainforest All land areas with a potential to support the tropical rainforest as determined by the climatic and physiographic conditions whether currently forested or not. The concept is interesting because it provides a baseline to calculate the rainforest loss in a country or region over a historic period.

Rainforest formation In the present context, the tropical rainforest formation covers forests areas occurring approximately within a latitudinal belt of 10° on either side of the equator. It may be noted that temperate and subtropical zones have forest formations structurally similar, but biologically not as diverse as the tropical rainforest.

Assessment Techniques

The methods for global forest assessment, of which Rain Forest Assessment forms a part, can be grouped into two broad classes: (1) those based on aggregation of existing country data; and (2) new surveys using remote sensing images. The United Nations Food and Agriculture Organization (FAO) is the Agency responsible for reporting on the subject. In addition, the European Union TREES Programme at ISPRA, Italy, undertakes research studies with a focus on the tropical rainforest worldwide (Achard *et al.*, 2002).

The FAO started global forest assessment in 1948 using a questionnaire approach. After publishing four reports on a 5-year cycle, the global assessment was discontinued in 1963, apparently due to poor response and inconsistent information received from the developing countries. The growing concern about the fate of forests in general and tropical forests in particular, expressed at the Stockholm Conference on Human Environment in 1972, led to initiation of the FAO/UNEP Tropical Forest Resources Assessment with the objective of providing reliable forest area and changes information for the tropical countries, with 1980 as the reference year. The assessment was conducted centrally at FAO HQ Rome by a team of experts under the technical guidance of FAO Forestry Department using reliable country information. The project estimated tropical forest area at the end of 1980 to be 1935 million ha and the annual rate of deforestation during 1980–1985 at 11.3 million ha (FAO, 1982). The news served as a wake-up call to the international community to start action for conservation and sustainable management of tropical forests.

Responding to international concerns, the FAO strengthened its programme on forest resources assessment during the eighties and carried out a number of benchmark activities including: (1) Global Forest Resources Assessment with 1990 as the reference year (FRA 1990); (2) a statistically planned multidate survey of tropical forests using remote sensing; (3) eco-floristic zoning of the tropical regions; and (4) tropical forest cover mapping. During the nineties, the FAO continued and improved on the FRA 1990 approaches and extended the coverage of the ecological zone and forest cover maps worldwide. A brief description of the FRA 1990 methods, with special reference to tropical countries will be given as they form the main source of information presented in this report (FAO, 1993, 1996).

Model-Based Assessment

The main steps were as follows:

1. The existing forest resources information on countries (both historic and current at a subnational level) was reviewed and entered in a database called Forest Resources Information System (FORIS). All subsequent observations for the area were to be added to the database as a part of time series information, which would include, besides area, volume/biomass, forest management status, forest utilization, and socio-economic data. FORIS also included subnational level demographic data since 1960 and the ecological zone information.
2. A model was developed for estimating forest cover on standard dates and making projections using time series of forest cover and demographic data and ecological zone information, all at a subnational level.
3. Based on the model equations, forest area and rate of change for various time periods were estimated at the subnational level; estimates were aggregated at the national, regional, and global levels.

Multidate Remote Sensing Survey

The sample survey of forests in the tropics was designed using multidate high-resolution satellite images, close to years 1980

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00300-2

and 1990, to estimate forest changes on a reliable basis. The tropical region was divided into 10 geographic and three forest cover strata using the forest cover and ecological zone maps. In all 117 satellite images, each covering approximately 3 million ha, were randomly selected and distributed to various strata. The study, covering nearly 10% of the area of the tropics, was conducted in a very systematic manner under a controlled environment. The two images at each sample site were of comparable quality, taken during the same season and interpreted by the same person under a central quality control. The interdependent image-interpretation procedure enabled assessment of changes with a much lower standard error compared to the independent interpretation of the two images.

The Historic Rainforest Area

Rainforest is the natural vegetation of the equatorial zone. It is associated with a very uniform temperature throughout the year (26–28 °C) and wet conditions with precipitation exceeding evaporation at least for 10 months. The historic distribution of rainforests, as described in Singh (1974), is as follows:

South and Central America: The Amazon basin extending west to the lower slopes of the Andes, east to the coast of Guyanas, south to the Gran Chaco, and north along the eastern side of Central America. Africa: The Congo basin (mainly Zaire and Congo), Gabon, and Cameroon. It extends westward into Nigeria, Ghana, Ivory Coast, Liberia, and Guinea. On the eastern side it extends to Uganda. Asia: mainly Southeast Asia, including the Malay Peninsula, the islands of Sumatra, Borneo, Sulavesi and New Guinea, part of India, Sri Lanka, Burma, Thailand, and the Philippines.

Although climate in the rainforest regions is relatively uniform, there are local variations caused by changes in soil, topography, and geology. These site variations give rise to a corresponding change in vegetation forms. Some typical forest formations in the equatorial zone are the following: The rainforest proper (also called climax rainforest) usually occurs on well-drained soils and below elevations of 500–800 m. This is generally what the common person means when he or she talks of the rainforest. Other types are mangrove forest, swamp forest, periodically flooded forest, riparian forest, hill forest (usually at altitudes higher than 800–1000 m), forest on tropical podsols (a soil type), and forests on coastal sand.

Ecological studies show evidence of striking similarities of vegetation characteristics among the rainforest formations across continents even though they occur in widely separated areas. According to Richards (1952), they are composed of plant species with comparable characteristics and great similarity in their spatial arrangement, life form, and physiognomy. The fundamental pattern of structure is thus the same through the whole extent of the rainforest.

FRA 1990, in close cooperation with the Laboratoire d'ecologie Terrestre in Toulouse, France, and member countries compiled the geo-referenced Eco-floristic Zone Map of the tropics, showing boundaries of the potential vegetation types occurring in the tropics (including the rainforest) to provide a basis for estimating the associated areas (FAO, 1989). According to this map, the potential area of the lowland rainforest worldwide is 935 million ha, with 127 million ha (13.5%) in Africa, 308 million ha (33.0%) in Asia, and 500 million ha (53.5%) in South America. South China and Australia also have small areas of tropical rainforests that are not included in the previous total.

Eco-Floristic Zone Concept

The following model describes the eco-floristic zone mapping procedure used by FAO:

$$Z = f(V,\ C,\ P,\ S|L)$$

where f is the zoning procedure (a rule, classification system, look up Table, etc.) to assign Ecological Zone Values (Z) to land units (L) of a country, region or the world, based on characteristics of natural vegetation (V), climate (C), physiography (P), and soil (S). Climatic criteria are used at a global level of classification; physiographic variables at the regional level; and soil characteristics at the country level of classification.

The Current Rainforest Area and the Rate of Loss

Model-Based Assessment

According to model assessment, the pan-tropical rainforest area at the turn of the Millennium was 680 million ha with the following distribution: Africa 81 million ha, Asia 159 million ha, and South America 440 million ha (see **Table 1**). Compared to the potential, the still remaining rainforest was 73% of the total and the rate of loss during 1990–2000 was 39 million ha with the following distribution: Africa 5 million ha, Asia 19 million ha, and South America 15 million ha, showing that Asia was losing rainforest at the fastest rate.

Using multidate data available in FORIS, a time series analysis of rainforest losses since 1960 was carried out. The area loss on a pan-tropical basis during 1960–2000 was 160 million ha, compared to 220 million ha prior to 1960, bringing out the fact that most of the rainforest loss had occurred during the post-1960 period (see **Table 2**).

Findings of FRA 1990 Remote Sensing Survey

FRA 1990 made a pioneering effort to provide information on the state of forests in the tropics to international community on a continuing and reliable basis. **Table 3** summarizes the main findings of the survey on rainforest changes during 1980–1990 grouped into two broad categories: (1) deforestation; and (2) forest degradation. The former refers to outright change from forests to nonforest land uses; and the latter is concerned with changes within the areas classified as forests. A more detailed analysis of changes will be presented later.

Two conclusions are obvious from the findings presented: (1) forest degradation was more wide spread and affecting 36.5 million ha forest area than deforestation, viz. 23.4 million ha. Added together they had altered natural vegetation forms in 60 million ha of the rainforest area. (2) Biological diversity in rainforests of Asia is most threatened on account of a very high rate of deforestation (in % terms) and an equally very high rate of forest degradation.

Table 1 Distribution of historic and current rainforest areas by region

Region	Historic rainforest area Million ha	Current rainforest		
		Area at end 2000		Loss during (1990–2000) Million ha
		Million ha	(%) of the potential	
Africa	127	81	64	5
Asia	308	159	52	19
S. America	500	440	88	15
Total	935	680	73	39

Table 2 Results of model studies using country data (million ha)

Region	The rainforest area at end of successive decades, 1960–2000					Total loss 1960–2000
	1960	1970	1980	1990	2000	
Africa	100	96	91	86	81	19
Asia	245	224	200	178	159	86
S. America	495	485	471	455	440	55
Total	840	805	763	719	680	160

Table 3 Rainforest areas and changes by Region (in million ha), 1980–1990

Region	FRA 1990 remote sensing results				Model results	
	Rainforest areas		Forest loss 1980–1990	Degradation 1980–1990	Rainforest area and loss	
	1980	1990			1990	1980–1990
Africa	91	89	2.0	3.76	86	5
Asia	161	151	9.9	16.9	178	19
S. America	500	488	11.5	15.9	455	15
Total	752	728	23.4	36.5	719	39

Source: With permission from FAO.

Findings of FRA 2000 Remote Sensing Survey

The FRA 2000 Remote Sensing Survey interpreted images of 1980, 1990, and 2000 on the same sites as FRA 1990 and used an identical statistical design. This makes the initiative of great value in two different ways: (1) to make comparison with results of FRA 1990; and (2) to provide an indication of changes in the rate of deforestation. With regards to the first point, it was found that the results of the two surveys are very close. The pan-tropical estimate of the total forest area by FRA 1990 and FRA 2000 are, respectively, 1656 and 1671 million ha (difference of only 1%); the regional estimates differ by less than 5%; the pan-tropical deforestation rates differ by 10%; and regional deforestation rates by 20% (FAO 1993, 2002).

The rainforest statistics produced by the two survey rounds, however, cannot be compared in a straight forward manner on account of changes in the definition. FRA 1990 defined rainforest to include areas with rainfall $>$ 2000 mm and dry period in a year not to exceed 2 months; while FRA 2000 included all forests in the class with length of dry period in a year not exceeding 3 months. A careful study of the two definitions leads to the conclusion that the FRA 2000 rainforest definition was broader and included all of the FRA 1990 rainforest (viz. 758 million ha) and some of the erstwhile moist forest areas (viz. 244 million ha), which made a total of 1002 million ha observed in 1980 by FRA 2000. The fact that the two surveys had made use of an identical design, surveyed the same sites on two dates, viz. 1980 and 1990, makes it possible to integrate the two datasets as one to study many interesting questions about on-going changes in the three ecological zones on two dates. **Table 4** gives observed values of FRA 2000 and FRA 1990 without brackets and inferred values in brackets.

The following interesting facts emerge from the analysis:

1. The observed deforestation during 1980–1990 in the strictly defined rainforest was 3.1% compared to a very high rate of loss (10.6%) in the adjoining moist forest areas;
2. Deforestation seems to be accelerating in (the strictly defined) rainforest areas as well as the moist forest areas included in FRA 2000.

Interpreting FRA 2000 survey findings, Drigo (2005) observes that the most active fronts of deforestation during 1980–1990 were in the wetter parts of the Moist Zone, characterized by a short dry season; and these hot fronts are getting more intense and they are fast moving toward the wetter cores

Table 4 FRA 1990 and FRA 2000 estimates of rainforest areas (all areas in million ha)

Survey rounds	*Current rainforest*			*Rainforest loss*			
	1980	*1990*	*2000*	*1980–1990*		*1990–2000*	
	Area	*Area*	*Area*	*Area*	*%*	*Area*	*%*
FRA 2000	1002	953	896	49.4	4.9	57	6.0
FRA 1990	758	734	(707)	23.5	3.1	(27.1)	(3.7)
Difference	244	(219)	(180)	25.9	10.6	(29.9)	(13.6)

Source: With permission from FAO.

of the remaining rain forest areas. This acceleration is accompanied by the indication of a possible deceleration of the deforestation rate in the (remaining) Moist Deciduous Forest. Arguably, this shift of the deforestation front toward the wetter zones is the most significant trend revealed by FRA 2000. This inference is collaborated by the fact that the rate of rainforest deforestation estimated by FRA 2000 for 1990–2000 shows an increase from 49.4 million ha in 1980–1990 to 56.6 million ha in 1990–2000 (FAO, 2002).

Taking into account all information sources, the remaining rainforest area at the end of 2000 is estimated at 900 million ha following the broader FRA 2000 definition; and at 700 million ha following the stricter FRA 1990 classification. The corresponding decadal losses during 1990–2000 are, respectively, 57 and 27 million ha. Notice also needs to be taken of the acceleration observed in the deforestation rates in the rainforest and the adjoining moist forests during 1990–2000.

Evaluation of the Model and Remote Sensing Survey Methods

It may be noted that data inputs for the model in most cases are country interpreted forest area information based on wall-to-wall remote sensing, local knowledge and ground control, more intensive than possible in global studies. In contrast, the pan-tropical sample survey is based on a statistically planned design meticulously implemented under a central supervision, but it has inherent sampling errors. For the pan-tropical area at 953 million ha in 1990, the standard error was 70 million ha (i.e., about 7.3%); and for the net area change during 1990–2000 at 60 million ha, the standard error was close to 12 million ha (i.e., about 20%). The model derived estimates at the pan-tropical and regional levels, presented in **Table 3**, are well within the confidence intervals of the remote sensing survey, but forest change estimates tend to lie on the higher side.

Errors are inherent in estimating forest area changes based on country data, which are subject to additional variations arising from the technology used, interpretation quality, timeliness of data, and classification system. FORIS has built in procedures to minimize the effect of such variations inherent in country data by using expert knowledge on country classification and data collection method; and by estimating forest area changes using time series of information at a subnational level. The model also makes use of additional natural and cultural variables, correlated with forest changes, to reduce error. But the fact remains that the precision of model estimates can be improved only through improvements in the quality of country data through country capacity building and harmonization of methods used by countries, by FAO.

Information on forest degradation is a very important plus point in favor of the pan-tropical sample survey method, not presently possible in model-based assessments. The precision of sample survey techniques for estimating changes could further be improved by incorporating natural and cultural information at the sample sites in the analytical procedure, as done by the model. Further gains are possible by integrating wall-to-wall information obtained by using coarse resolution satellite data in the sample survey based on high-resolution data.

Types of Changes within the Forest Land

So far the transfer of land from forest to nonforest class has been described. Changes within forest land are of equal if not greater interest from an ecological as well as an economic perspective. However, representative country data on these changes, on a global basis, are lacking. The only source of consistent information at a global/regional level are FAO Remote Sensing Surveys, which provide within forest change information in a matrix form (**Table 5**). The matrices are ideal for tracking changes but not very easy to read. To give a bird's-eye view of the change information, FRA 1990 developed a pictorial presentation called a flux diagram (see **Figure 1**), which shows area changes on the *x*-axis and associated average biomass changes on the *y*-axis. Thus, the figure as a whole shows the total biomass change in the form of a rectangle. Such a presentation is easy to read and informative for studies on climate change.

The FAO change matrices have nine rows and nine columns that represent the full range of land use and forest cover changes important for the study of biological diversity and climate change. The transition among the first four classes relates to changes within the forest domain and represents various grades of density or forest cover changes (closed and open), spatial disturbance (fragmented), and temporal disturbance (long fallow). The transition among the next four classes relates to transfers from forest to nonforest domain including shrub, short fallow, other land cover, and water. The last class, plantation, could be in the forest or nonforest domain (e.g., forest or agricultural plantation). The two domains are grouped as one because it is not easy to separate them with certainty on satellite images.

A distinction needs to be made between the long fallow and short fallow categories, both of which are affected by shifting cultivation. The distinction between long and short has been

Table 5 Area transition matrix for the period 1980–1990 for the wet and very moist regions of the tropical world

Classes at year 1980	*Classes at year 1990 (million ha)*									*Total 1980*	
	Closed forest	*Open forest*	*Long fallow*	*Fragmented forest*	*Shrub*	*Short fallow*	*Other land cover*	*Water*	*Plantations*	*Million ha*	*%*
Closed forest	685.36	0.96	5.35	2.32	0.42	9.16	10.11	0.36	3.70	717.74	80.5
Open forest	0.10	11.54	0.02	0.08	0.01	0.36	0.20	0.01	0.02	12.32	1.4
Long fallow	0.37	0.06	21.63	0.06	0.13	1.31	0.47	0.01	ε	24.02	2.7
Fragmented forest	0.07	0.03	0.33	7.78	0.19	1.22	1.44	0.05	0.14	11.24	1.3
Shrubs	ε		0.04	0.01	3.52	0.03	0.20	ε	0.01	3.81	0.4
Short fallow	0.27	0.04	0.27	0.22	0.02	55.78	5.04	0.13	0.14	61.92	6.9
Other land	0.13	0.01	0.03	0.02	0.64	0.62	47.32	0.09	0.05	48.93	5.5
Water	0.11	ε	0.01	0.03		0.02	0.13		0.01	0.29	ε
Plantation	ε	ε	ε			0.01	0.02		11.45	11.48	1.3
Total Million ha	686.41	12.65	27.68	10.52	4.93	68.50	64.92	0.66	15.50	891.76	
%	77.0	1.4	3.1	1.2	0.6	7.7	7.3	0.1	1.7		100

The number of sampling units was 26; land area studied was 45.54 million ha. Sampling results have been expanded to the total land area covering 891.76 million ha.
Source: With permission from FAO.

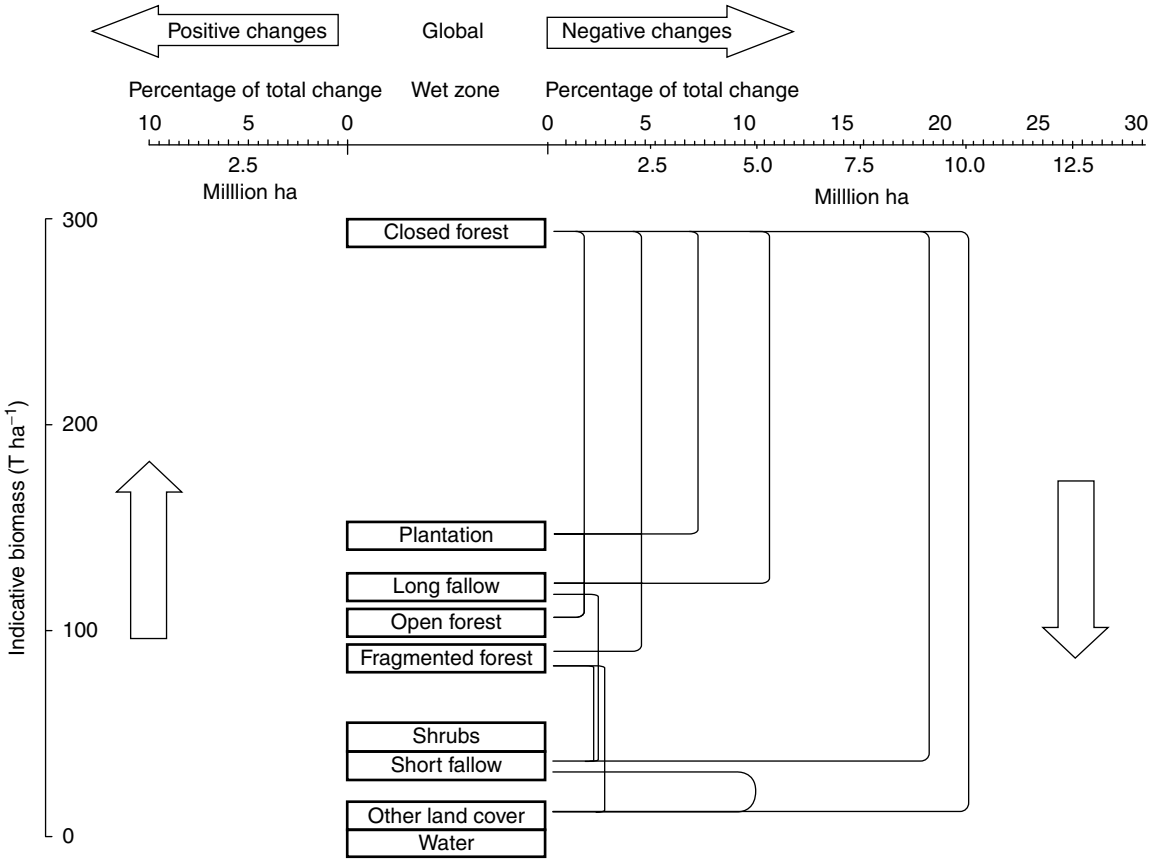


Figure 1 Flux diagram showing area transition on the *x*-axis and biomass transition on the *y*-axis. Each rectangle represents total biomass change associated with the transition.

made with the purpose of dividing the total area under shifting cultivation into predominantly forested, no matter how much degradation has occurred, and predominantly nonforested (but woody) on the basis of estimated intensity of cultivation or, specifically, of the rotation cycle. This distinction brings some clarity to the controversy about shifting cultivation which, without further subdivision, is sometimes considered non-forest (i.e., classified as 'other wooded land' according to the FAO–FORIS definition) and sometimes forest (i.e., according to the definition used by some countries).

The information appearing in the flux diagram (**Figure 1**) clearly shows that most of the changes resulted in outright loss of the rainforest. Of the total loss, about 31% went to permanent agriculture, 45% to short and long fallow (viz. shifting cultivation), 7% to fragmented forest, 3% to open forest, 11% to plantations, and 3% to other land uses. A brief account of the process of land-use changes in the three regions follows (FAO, 1996; Drigo, 2005).

Africa: The dominant form of change is *closed forest → short fallow*, which is on account of small-scale subsistence farming, whereby the fertility of the soil is regenerated during fallow periods of a few years, as agronomic additives and fertilizers are inaccessible to poor farmers;

- The sequence *closed forest → open forest → fragmented forest → other land cover* clearly represents the various stages of forest depletion, and is an effect of rural and urban population land use and energy needs.

Asia: Characteristic changes in the region are closed forest → other land cover; and closed forest → Plantations (particularly in Indonesia and Malaysia); closed forest → long fallow shifting cultivation; and long fallow shifting cultivation→ short fallow shifting cultivation (intensification of permanent agriculture on traditional shifting cultivation areas (South and South east continental Asia). The process long fallow forest class → closed forest is a direct effect of the abandoned cultivation.

South America: The dominant typology is: *closed forest → other land cover*, which represents deforestation with the highest possible biodiversity loss resulting mainly from the direct conversion of the original forest to cattle ranching and permanent agriculture. The second most frequent transition, *shrubs → other land cover*, represents the large areas of Brazilian *caatinga* (steppe typical of the northeast regions of Brazil) and *cerrado* formations (tree and shrub savanna of south-west and central Brazil) that are being converted to cattle ranching. Some positive changes such as *other land cover → short fallow*, *other land cover → shrubs* or *other land cover → fragmented forests*, which represent regrowth of previously cleared forests.

The outcome of the land-use change, viz. what happened to the deforested land, is a very important determining factor of the biodiversity loss, the impact of change on the functioning of the ecosystem proper and its affect on the adjoining ecosystems including the indigenous population. In particular, changes involving large scale land clearing, tilling of land and introduction of exotic species might have adverse local, regional, and global implications. It may even lead to ecosystem instability in extreme cases. Keeping in view such a likelihood, the driving forces and socio-economic impact of deforestation are briefly presented by region based on a recent report (UNEP, 2009).

In the Congo Basin, the forest loss and degradation is mainly caused by the displacement of people due to war and conflict and the impact of mining activities in previously untouched forest areas particularly along the tributaries of the Congo River, most likely due to population growth and settlement. In the densely populated mountainous regions and high plateaus of western Cameroon and the east of Democratic Republic of Congo, forest regeneration is reported as getting deficient due to the shortening of fallow periods, combined with the conversion of abandoned fields into pasture. The productive forest areas in the entire Basin, barring protected areas, seem to be under logging contracts.

A main driving force for the rainforest loss in Asia region is cultivation of oil palms. In 2006, two countries, viz. Indonesia and Malaysia, accounted for 85% of the total world production and 88% of global exports. Over the past decade, the area covered by oil palms in Indonesia has quadrupled, covering 4.1 million ha in 2006.

In Papua New Guinea, logging in approximately 80% of the productive forests would be completed by 2020. In Kalimantan, Indonesia, there are reports of protected area being logged with the loss of more than 56% area between 1985 and 2001. The interests of commercial activities such as palm oil plantations or mining often clash with the interests of communities, small farmers, and indigenous people.

In the Amazon region, cattle ranching, Soya cultivation, and commercial logging operations were the main contributing factor to rainforest loss. In Brazil, ranches covered an area of at least 8.4 million ha and accounted for an estimated 70% of deforestation in 2007 with an average size of 24,000 ha, some ranches could be as large as 560,000 ha. The rising demand for Soya in the world trade has been also an important factor. Brazil is presently the world's second biggest Soya bean producer, accounting for 23% of the global total and about 22 million ha of land under the crop, mostly in the south-eastern part of the country.

Implications for Biological Diversity

The large scale land clearing and human settlements, ongoing in all the rainforest regions, are bound to cause radical and even irreversible changes in the natural (primeval) ecosystem including the indigenous people, whose cultural identity is strongly correlated with the biological diversity. Efforts need to be made to arrest the process of biological diversity loss by careful conservation planning ensuring that the number of individuals per species in the ecosystem does not fall below a critical limit in all forest types, where the species occurs (Singh, 2005). There is an urgent need for nationwide continuous inventory of forests and wild animals. This is a very challenging task to monitor, as not only the number of species in rainforests is large, but also their occurrence is rare, calling for a special effort for getting reliable species identification.

Deforestation and Species Richness Loss

FRA 1990 provides a method for estimating risks of species richness loss associated with deforestation in the form of an index, using the well-known phenomenon of increase of species number with increasing forest area and conversely, loss of species richness with decrease in forest area. Details of concept and methodology for estimating the index are presented in FAO (1993).

Table 6 provides values of indices by region. The risk of species richness loss is the highest in Asia region. This fact is confirmed by findings of FAO model studies and Remote Sensing Surveys presented earlier. This region by the end of the Millennium had lost about 48% of its historic rainforest area

Table 6 Risk of species richness loss associated with deforestation (refers to angiosperm species only)

Region	*Forest area (million ha)*		*Deforestation (million ha)*	*Species richness (number)*	*Risk of species richness loss (%)*
	1980	*1990*	*1981–1990*	*1990*	*1980–1990*
Africa	91	86	5	30,700	2.0
Asia	200	178	22	40,400	4.3
South America	471	455	16	57,900	1.6

Source: With permission from FAO.

and had the highest rate of deforestation (in % terms) and highest rate of forest degradation (in absolute terms). The area logged and the intensity of logging was also high compared to other two regions. Only Indonesia still had some large tracts of rainforests left; the other countries had only small areas in a highly fragmented forms. Immediate and effective conservation measures are needed in this region.

Africa initially had a small share of the world's rainforest and lost nearly 36% of the area by the Millennium end. In particular, the West Africa had undergone a very high loss of the historic rainforest area, with the remaining forests occurring in a highly fragmented manner like islands in a sea of forest fallow (FAO, 1993). The ongoing changes in the central African region were observed to be relatively slow during 1980–1990. However, the recent flux of population at the border of Zaire and Rwanda must have affected the rate of loss and biological diversity, but this effect is not yet known. Second, the ongoing forest area changes in Africa are from closed forest to short fallow, which is an extensive and destructive form of land-use change from the biological diversity perspective. The pattern of transfer is different from that of South America or Asia, where the land-use change is more planned and intensive. With a high rate of population growth, such land-use changes are a cause of concern for forestry and biological diversity conservation perspective.

The land-use changes in South America and the Amazon, in particular, are of more recent origin compared to those in Asia and Africa. The total area of the rainforest is still very high (viz. nearly 88%) and the total area lost is relatively small. However, a state-wise distribution of the remaining forest cover and the rate of deforestation provide a different view. Only the states of Amazonas, Roraima, and Ampa have large tracts of inaccessible forests; other states, such as Maranhao, Tocantins, Para, Mato Grosso, Rondonia, and Acre, are undergoing larger than average forest area changes per capita compared to other regions of the world.

Loss of Cultural Diversity Associated with Deforestation

According to Dobzhansky (1964), there is intricate mutual relationship among inhabitants (of an ecosystem) which makes them attuned (coadapted) to the manifold reciprocal dependencies. Natural selection becomes a creative process and leads to emergence of new modes of life with more advanced types of organization. An ecosystem could thus be seen, as a higher level, interspecies organization opening possibilities for progressive evolution and responding to the pressure of natural selection in an effective manner. This bond or web of life falls apart with increasing human intervention in the ecosystem from outside.

Rainforests in all the three regions abound in cultural diversity. Any form of forest changes affects them directly, deforestation in particular. The following account of cultural richness in the three rainforest regions is based on UNEP (2009).

The Central African forests are the home of 30 million people, coming from 150 different ethnic groups. A majority of these groups are indigenous and include the Pygmies, as well as many groups of Bantu origin that have been forest dwellers for more than 1000 years and have intermixed with Pygmy, Ubangi, and Sudanic populations.

Indonesia, with rainforest area of 94 million ha, is home of 500 ethnic groups speaking more than 600 languages. This diversity is accepted as a cultural asset, which is reflected in the national slogan, *Bhinneka Tunggal Ika*, unity in diversity. These groups are widely distributed in islands with difficult access and need to be empowered to protect, manage, and share equitable benefits from harvesting operations and land-use changes.

In the Amazon Rainforests, there are more than 200 major indigenous groups, each with their own cultural heritage and language. This shows that like the flora and fauna, the cultural diversity in all the three regions is also very high, which needs to be taken into account in planning of land-use changes.

Loss of Ecosystem Resilience

There is much empirical evidence, and some theoretical support, for the notion that as diversity increases in forests, so does their robustness or resilience in terms of their ability to withstand external stresses. Agents that can cause these changes include wildfire, severe storms, or outbreaks of pests and disease. These events are an integral part of the dynamics referred to as disturbance events (Singh, 2005). The maintenance and monitoring of biodiversity at stand, species and ecosystem levels should be an integral part of forest management and conservation. The strategy should encompass protected as well as sustainable managed forest areas over the entire geographic range of a country/region.

According to Wilson (2000), if habitat destruction were to continue in forests and coral reef alone, half of the species of plant and animal diversity would be gone by the end of the twenty-first century.

The time series data on forest cover and population changes, collected by the FRA 1990 project from all regions of the tropics, show a more rapid forest area loss with an increase of the population pressure in the rainforest region compared to other ecological zones (**Figure 2**). The rainforest, per unit area, is also the most biodiverse. The two facts imply a larger loss of diversity by human settlement and hence the need for a

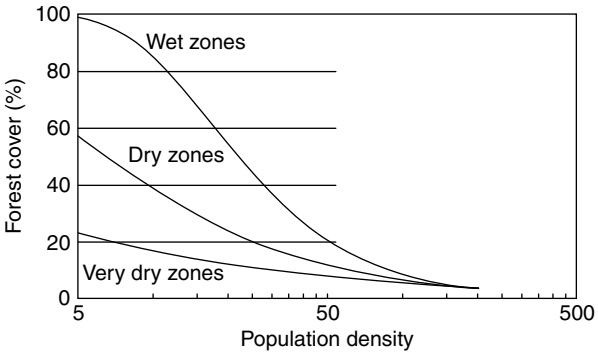


Figure 2 Illustration of model curves for different ecological zones.

long-term perspective and care in planning of land use and forestry changes in the tropical rainforest zone.

See also: Deforestation and Land Clearing. Rainforest Ecosystems, Animal Diversity. Rainforest Ecosystems, Plant Diversity. Tropical Forest Ecosystems

References

Achard F, Eva HD, Stibig HJ, *et al.* (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297: 999–1002.

Dobzhansky Theodosius (1964) *Genetics and the Origin of Species*. New Delhi, India: Oxford Book Company.

Drigo R (2005) Trends and patterns of tropical land use change. In: Bonell M and Bruijnzeel LA (eds.) *Forests–Water–People in the Humid Tropics*, pp 1–39. Cambridge, UK: Cambridge University Press.

FAO (1982) Tropical forest resources. *Forestry Paper No. 30*. Rome: FAO.

FAO (1989) *Classification and Mapping of Vegetation Types in Tropical Asia*. Rome: FAO.

FAO (1993) Forest resources assessment 1990: Tropical countries. *Forestry Paper No. 112*. Rome: FAO.

FAO (1996) Forest resources assessment 1990: Survey of tropical forest cover and study of change processes. *Forestry Paper No. 130*. Rome: FAO.

FAO (2002) Pan-tropical survey of forest cover changes, 1980–2000, results and findings. *Working Paper 49 (b)*. Rome: FAO.

Richards PW (1952) *The Tropical Rain Forest*. Cambridge, UK: Cambridge University Press.

Singh KD (1974) Spatial variation patterns in the tropical rain forest. *Unasylva* 26(106): 18–23.

Singh KD (2005) *Forest Biological Diversity, Assessment and Conservation Planning*. New Delhi, India: ITTO and WWF India.

United Nations Environment Programme (UNEP) (2009) *Vital Forest Graphics*. Nairobi, Kenya: UNEP GRID Arendal.

Wilson EO (2000) Biodiversity. *Time* April/May 2000.

Range Ecology, Global Livestock Influences

J Boone Kauffman, Oregon State University, OR, USA
David A Pyke, USGS Forest and Rangeland Ecosystems Science Center, OR, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 2, pp 33–52, © 2001, Elsevier Inc.

Glossary

Animal unit month (AUM) Quantity of forage needed to sustain one 1000-lb cow or five sheep for one month.
Biological soil crusts Collective term referring to the combination of nonvascular and non-seed-forming plants that commonly occur on the soil surface of many arid and semiarid ecosystems. These crusts may consist of mosses, lichens, liverworts, green algae, and cyanobacteria. Synonyms: cryptogams, microfloral, cryptobiotic, cryptogamic, microbiotic, organogenic, biogenic, biotic, and microphytic soil crusts.
Browse That part of leaf, twig, and reproductive growth of shrubs, woody vines, and trees available for animal consumption. The act of consuming leaves, twigs, or reproductive parts of shrubs, woody vines, and trees.
Degradation Decrease in ecosystem productivity or structure, and/or declines in native species diversity (sometimes with concomitant increases in exotic species) due to land use practices. Directly related to declines in biodiversity, ecosystem degradation encompasses soil impoverishment (e.g., compaction, erosion, salinization, loss of biological soil crusts) and hydrological alterations that diminish water availability. Similar to desertification.
Desertification Process by which an area becomes more arid through the decline in primary productivity, elimination of biological cover, shifts in plant diversity, or degradation of soils due to human uses whose impacts exceed the sustainability of the landscape. Similar to degradation.
Forbs Non-graminoid herbaceous plants.
Graminoids Grasses (family Poaceae) and grass-like plants, mostly sedges (family Cyperaceae) and rushes (family Juncaceae).
Graze Act of consuming herbaceous plants, including grass, grass-like plants, and forbs.
Overgraze Consumption of forage to the extent that declines in productivity or desirable species composition are probable (i.e., degradation). Overgrazing is excessive use in terms of the amount, duration, and frequency of plant utilization.
Pasture Lands in which the native vegetation has been removed in favor of cultivated grasses or other forage plants.
Range condition Current productivity and composition of plants on a rangeland relative to its ecological potential. Often range condition is classified into arbitrary classes (i.e., excellent, good, fair, and poor).
Rangeland Land on which the potential native vegetation is predominately grasses, grass-like plants, forbs, or shrubs. Rangelands include prairies, marshes, tundra, wet meadows, savannas, shrubland steppe, chaparral, desert grasslands, and woodlands.
Riparian zone Interface between streams and terrestrial uplands. These zones are three-dimensional areas of direct interaction between the terrestrial and aquatic ecosystems. They often contain unique species and high levels of biological diversity.
Ruminant Grazing mammal with a complex four-chambered stomach (including a rumen). Common ruminants include cattle, sheep, goats, deer, antelopes, and giraffes.
Savanna Ecosystem with a more or less continuous herbaceous layer dominated by graminoids and broad-leaved herbs with an overstory of trees or shrubs that covers less than 10% of the area. In contrast, grasslands are typified by a pure graminoid or herbaceous layer with no (or very few) tree or shrub elements that rise above the grass layer.
Woodland Ecosystem in which the mature overstory tree or tall shrub cover makes up 10 to 60% of the area. Typically, a well-developed herbaceous understory is present. Note: forests typically have a mature tree cover $>60\%$.

Rangelands of the World

Rangelands are defined as those lands where the native vegetation is predominantly grasses, shrubs, or open woodlands. Rangelands include grasslands, shrublands. savannas, open woodlands, and most desert, tundra (arctic and alpine), meadow, wetland, and riparian ecosystems. Occurring from the tropics to polar regions, rangelands cover more land area than any other type on Earth. Because of climatic or edaphic limitations, they tend to be lands that are incapable of growing marketable timber or agronomic crops without irrigation or some other modification (i.e., they are over poor-quality soils and/or are too wet. too dry, or too cold for cultivation). The composition, structure, productivity, and diversity of these ecosystems are governed by a combination of climate, geography, topography, and geology, including soil development. In addition, rangelands are used by a large number of vertebrate and invertebrate herbivores, including a diverse combination of domestic or native ungulates.

Tundra

Tundra ecosystems are dominated by perennial grasses, forbs, shrubs, and biological soil crusts consisting of cyanobacteria,

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00052-6



Figure 1 The arctic tundra landscape of the Brooks Range, Arctic National Wildlife Refuge, Alaska, with permission from Boone Kauffman.

lichens, and mosses (**Figure 1**). These species can tolerate the climatic conditions of high latitudes or altitudes. Tundra ecosystems are generally set apart from forested ecosystems by a climatic tree line defined by the 10 °C isotherm for the mean temperature of the warmest month. Tundra plants are low in stature with growth buds near the soil surface, where temperature is less variable. The morphology of herbaceous plants are often cushion or rosette growth forms and lichens are usually a foliose form. Shrubs are low-growing with short internodes between leaves. Temperate alpine tundra communities have short growing seasons similar to those in arctic tundra. In contrast, tropical alpine tundra communities of Africa, South America, and Oceania exist in conditions of freezing or near-freezing temperatures each night, with daytime temperatures allowing for active plant growth throughout the year.

The important herbivores of the Arctic are rodents and ungulates. Lemmings (*Dicrostonyx* spp. and *Lemmus* spp.) can have great impacts on arctic ecosystems during years of high population densities. Lemmings exhibit different seasonal patterns of use within various parts of their habitats. During the winter, they use wetlands and graze the soil-level parts of plants while discarding the upper portions of the plants. This grazing pattern results in piles of graminoid litter known as lemming hay. During spring, lemmings move to south-facing slopes, where solar insolation results in the earliest available green biomass. During peak population cycles when densities are high, lemmings can consume between 20 and 70% of the available plant mass.

The dominant native ungulates in the Arctic include caribou or reindeer (*Rangifer tarandus*) and musk ox (*Ovibos moschatos*). Domestication of reindeer and musk ox has led to their introduction throughout the arctic tundra and to reindeer introduction into northern Antarctica. Alpine tundra provides seasonal habitat for wild ungulates, including elk, deer, red deer, mountain sheep, and mountain goats. Domestic livestock breeds adapted to cooler environments are also present in various arctic or alpine tundra communities of the world (e.g., cattle, sheep, goats, horses, llamas, and yak).

Grasslands and Savannas

Natural grasslands, also known as campos, llanos, pampas, plains, prairie, steppe, and veld, are frequently associated with semi-permanent high air pressure systems and high amounts of solar radiation. Annual precipitation typically ranges from 160 to 1700 mm with distinct wet and dry seasons. Grasses or graminoids dominate the vegetation composition, with forbs being a secondary component and short-stature shrubs contributing a lesser amount. The grass (or graminoid) dominance of grasslands is the result of current climatic conditions and/or natural disturbance processes such as fire. With human degradation such as overgrazing or fire suppression, shrubs often increase in abundance. In the tropics and in temperate locations with summer precipitation, C_4 grasses (warm-season grasses) will predominate because of their optimal growth at warmer temperatures. Northern temperate locations, where moisture is most available when temperatures are cooler, favor C_3 grasses (cool-season grasses). Biological soil crusts frequently occupy the soil surface below or between the plant cover, particularly in drier desert grasslands.

In equatorial regions, tropical savannas are a dominant rangeland ecosystem. They form a transitional zone along a precipitation gradient from tropical rain forests to arid deserts. Dense grasslands with occasional trees and shrubs characterize tropical savannas. As in grasslands, distinct wet and dry seasons typify the climate of tropical savannas. Disturbances are a prevalent feature of savannas; during the dry season, lightning-caused fires may occur. Human-ignited fires are very common throughout the world and maintain the open nature of many tropical savannas. Fire-return intervals in savannas are among the most frequent of any landscape on Earth. It is estimated



Figure 2 Overview of the Brazilian cerrado. This is a mosaic of grasslands, savannas, and woodlands that have a high level of biotic richness. Note the fire in the background, which is a common disturbance in this ecosystem. Land conversion to cattle pastures, soybeans, and other croplands is occurring at alarming rates in this ecosystem, with permission from Boone Kauffman.

that 50–75% of the humid savannas of Africa and South America burn each year.

Tropical savannas are highly productive. Grass biomass in humid savannas (>700 mm precipitation yr^{-1}) averages about 6.6 Mg ha^{-1} in Africa and South America, and 4.9 Mg ha^{-1} in tropical Asia. In savanna-woodlands, aboveground biomass may be as high as 25–61 Mg ha^{-1}. Roots and other subterranean tissues account for the majority of the plant biomass in savannas and grasslands (de Castro and Kauffman, 1998).

Tropical savannas are also very species-rich. For example, the largest savanna type in South America is the Brazilian cerrado, which often contains over 400 species ha^{-1} (**Figure 2**). Except for tropical rain forest, this is among the richest vascular plant assemblages on Earth. For example, about 3500 plant species and 400 bird species have been identified in the cerrado of the 5800 km^2 Federal District surrounding Brasilia, the capital of Brazil.

Savannas can experience rapid changes in species composition and structure following the introduction of an anthropogenic perturbation. Changes in the patterns of fire are one example; fire suppression results in an increase in the dominance of tree and shrub vegetation. Conversely, increasing the occurrence of fire enhances grass abundance. Linkages between fire, grazing animals, disease, and vegetation structure were described by Sinclair (1979). He related how the elimination of a ruminant disease, rinderpest, was a perturbation that changed the structure and composition of landscapes within the Serengeti of Africa. The immediate effect of rinderpest elimination was an increase in wildebeest numbers. On the plains, increased grazing pressures reduced grasses and increased forbs. Grant's gazelle increased along with the forbs, thus leading to population increases in cheetah. In woodlands, grazing reduced the grasses, which in turn reduced the fuel to support fires. The reduction in grasses resulted in a reduction in buffalo. Tree numbers then began to increase, which favored an increase in the giraffe populations.

Plants within natural temperate grasslands and tropical savannas vary in their tolerance to large numbers of grazing animals. The images of grasslands and savannas for many people include large herds of ungulates, such as bison (*Bison bison*) in the Great Plains of North America and wildebeest in East Africa. In these ecosystems, large herds of ungulates were a keystone feature that strongly influenced composition and structure. Plants are well adapted to herbivory in these grass-dominated ecosystems. However, large herds of ungulates are not a natural feature in all grasslands and savannas (e.g., the Intermountain and Palouse grasslands of the United States and the Brazilian cerrado). In these grasslands, rodents and insects are the dominant native herbivores.

The most widely distributed wild ungulates in North American grasslands and woodlands are pronghorn (*Antilocapra americana*), elk (*Cervus elaphus*), and deer (*Odocoileus* spp.). However, only bison comprised large herds within the Great Plains until they were decimated in the late nineteenth century by hunters. Bison also occurred in other North American grasslands, but not in large herds. Rodents and lagomorphs are also common grazers in North America. Population eruptions can result in years when the vegetation is heavily grazed by these animals in local areas, but within a few years disease, predators, competition, or other density-dependent factors reduce their numbers.

The common ungulates in South American grasslands and savannas include tapirs (*Tapirus* spp.), peccaries (*Catagonus*, *Dicotyles*, and *Tayassu* spp.), camelids (*Lama* spp.), and deer (*Odocoileus* and *Ozotoceros* spp.). Some of these are primarily forest or gallery forest (riparian) dwellers, but they also use savannas or grasslands. Only one native lagomorph (*Sylvilagus brasiliensis*) occurs in South America. The dominant vertebrate grazers in South American savannas are a diverse group of



Figure 3 An African woodland/savanna in South Africa. (Photo by Ron Shea.), with permission from Ron Shea.

rodents. The largest rodent, the capybara (*Hydrochoerus hydrochaeris*), is widely distributed, but prefers wetland savannas such as the Brazilian Pantanal. Capybara graze in herds as large as 1500 animals.

The African grasslands and savannas maintain the greatest species richness of wild ungulates of any rangeland ecosystem in the world (**Figure 3**). Rodents comprise a lower herbivore biomass than ungulates in most regions of Africa. The African ungulates consist of three taxonomic orders: Proboscidea (1 species; elephant); Perissodactyla (6 species; rhinoceros and zebras); and Artiodactyla, with five families. The families include Suidae (3 species; pigs), Hippopotamidae (2 species; hippopotamus), Tragulidae (1 species; water chevrotain), Giraffidae (2 species; giraffes), and Bovidae (79 species; buffalo and antelope). Although a few of the African ungulates are forest-specific species, most have distributions that include grasslands, savannas, and woodlands. Riverine and wetland locations with high rainfall in East Africa maintain populations of hippopotamus, waterbuck, and buffalo, whereas midheight grass regions with intermediate moisture support zebra, giraffe, and hartebeest. Native ungulates have been extirpated or driven to extinction in many regions of Africa. In the Fété Olé savanna of the Sahel, 11 species of native ungulates are now extinct.

Northern Asia has few native ungulates within its grassland/savanna ecosystems. Common ungulates include gazelle (*Procarpa* or *Gazella* spp.), wild horse (*Equus przewalskii*), and the Bactrian camel (*Camelus bactrianus*), and none of these form large herds. Rather, rodents are the primary vertebrate grazers within the northern Asian grasslands. They influence these ecosystems by creating local areas of high disturbance through digging and grazing. Voles, marmots, and ground squirrels often undergo population cycles that may result in the consumption of large amounts of vegetation when populations are high, thus reducing forage for local livestock herds. Such situations can exacerbate overgrazing since livestock densities are not typically reduced in years of high rodent numbers. Southern Asia is more diverse in numbers of native ungulate species with the inclusion of several species of wild ass, gazelle, and antelope. Asian elephants and rhinoceros were also present, but are now endangered and extirpated from practically all of their former range.

In the Australian grasslands, large grazing animals are represented by marsupials (euros and kangaroos, *Macropus* spp.; wallabies, *Petrogale* spp.). The red kangaroo (*Macropus rufus*) grazes in the arid and semi-arid zones, whereas the antilopine kangaroo is restricted to tropical grasslands of the north. Introduced herbivores, such as the European rabbit (*Oryctolagus cunculus*) and the Timor water buffalo (*Bubalus bubalis*), compete for herbage with native grazers in southern Australia. In Australia as throughout the world, the high productivity of grasslands and savannas has led to their widespread use by domestic cattle, sheep, and goats.

Temperate or Cold-Desert Shrublands and Semi-Deserts

Deserts or desert shrublands can be broadly classified into two different types—hot deserts of subtropical and tropical latitudes and cold deserts of temperate latitudes. Temperate deserts and semi-deserts occur at higher latitudes than hot deserts (between 30° and 50°), which allows for moisture from oceans to move into these regions by way of prevailing circulation patterns. Most of the temperate deserts and semi-deserts are characterized by sparse vegetation that is usually dominated by low shrubs and herbaceous plants. Total vascular plant cover rarely exceeds 50%, and biological soil crusts commonly occupy interspaces between vascular plants.

Cold deserts are characterized by cold winters and hot summers. Most occur in North and South America (Great Basin, Intermountain West, and Patagonia) and Eurasia (Kazakho-Dzungarian deserts and semi-deserts). In North America and western Eurasia, most of the annual precipitation

arrives as snow in the winter, while summers are dry. However, eastern Eurasia receives most of its moisture during the summer from frontal or cyclonal storms that move west from East Asia, while the winter remains dry. Patagonian semi-deserts receive low amounts of precipitation throughout the year, with strong winds exacerbating arid conditions during the summer from elevated evaporation.

The importance of native ungulates in cold deserts varies widely among these geographic regions. Large herds of ungulates, largely saiga antelope (*Saiga tatarica*), zheiran gazelle (*Gazella subgutturosa*), and wild ass (*Equus hemionus*), once roamed Eurasian ecosystems (until 100–200 years ago), but are now largely restricted to reserves. Wild sheep are common to both Northern Hemisphere ecosystems, with *Ovis ammon cyclocero* present in Afghanistan and Iran and *Ovis canadensis* in North America. Common native ungulates in North American cold deserts include pronghorn, deer, and elk. Feral horses and burros are now locally common in areas. Although bison were once found in this region, they never attained large numbers as found on the Great Plains and were largely absent by the time European exploration began. Only one native Patagonian ungulate exists, the guanaco (*Lama guanicoe*). Today domestic herbivores are predominate throughout all of these regions.

Hot Deserts

Hot deserts (or hot desert shrublands) are the most arid form of rangeland in the world. The hot deserts generally occur between latitudes of 15° and 30° N and S, the belt of the subtropical anticyclones that create near permanent high-pressure systems. Most hot deserts receive less than 120 mm of rain annually. The position within the high-pressure belts also results in a wide variation in the amount, location, and timing of rainfall. A wide spacing of drought- and heat-adapted shrubs is characteristic of hot desert shrubland vegetation. Perennial vascular plants normally cover less than 10% of the land area within hot deserts. Hot desert vegetation must be adapted to the uncertainty in the timing of adequate growing conditions that are characteristic of this ecosystem. The vascular plants persist in the variable and unpredictable growing conditions by using a host of adaptations to avoid dry seasons while quickly responding to moisture when it becomes available. Perennial plants often have persistent belowground organs that facilitate survival during the dry season. When moisture is available, some perennial plants can respond with a rapid production of adventitious roots and leaves. Many desert species (the Cactaceae, some Euphorbiaceae, and others) have photosynthetic stems that may allow growth during sporadic storm events without the cost of leaf production. Plant reproduction is limited to years with sufficient moisture. Ephemeral annuals survive in deserts by germinating and completing their life cycles during wet periods and then persisting in the soil as a dormant seed during dry periods. In locations where fog is the major source of water (coastal deserts of Chile, Peru, and Namibia), fog desert vegetation forms, with biological soil crust species often being the dominant ground cover.

Native ungulates are virtually missing from most hot deserts except along major rivers and wetlands. They may occupy the fringes of the hot deserts and migrate into the deserts during favorable conditions. An example is the desert big-horn sheep in the North American Mojave Desert. Though populations are found in adjacent semi-deserts where greater food and water are available, use of the Mojave Desert by bighorn only occurs during wet periods. Australian marsupials follow a similar pattern, with most only using the arid desert areas during favorable times. A few wallabies are found in rocky arid areas, but never in large numbers.

Herbivores are found in low densities in the North African (Sahel and Sahara) and Middle Eastern deserts. Native ungulates of these deserts include gazelle, ibex, oryx, wild asses, and wild boar. The exception to a low density of ungulates in hot deserts is found in the Karoo and Kalahari Deserts of southern Africa. Here, springbok, gemsbok, and hartebeest disperse in small herds or treks during dry periods. Once moisture produces aboveground growth of grass, these animals congregate into larger herds. Populations of these animals 200 years ago were larger than those today. Although forage is scarce and unpredictable, domestic livestock (cattle, sheep, and goats) grazing remains common in hot deserts of the world.

Biological Impoverishment Associated with Livestock Grazing

Domesticated animals (livestock) have played prominent and largely beneficial roles in human society for thousands of years, providing food, fuel, fertilizer, transport, and clothing. Yet livestock have had a dramatic negative impact on global biodiversity. The influence of livestock on global biodiversity is of concern simply because cattle and all other ruminant livestock graze about one-third to one-half of the planet's total land area. Along with pigs and poultry, they eat feed and fodder raised on one-fourth of the cropland. Livestock grazing is the most ubiquitous human activity on Earth and occurs on more area than any other land use. The impacts of livestock are far-reaching and often not immediately apparent (**Table 1**). In many regions of the world, grazing has reduced the density and biomass of many plant and animal species, reduced biodiversity, aided in the spread of exotic species and disease, altered ecological succession and landscape heterogeneity, altered nutrient cycles and distribution, accelerated erosion, and decreased both the productivity and land use options for future generations.

Given the widespread degradation caused by livestock on the diversity and structure of native range-lands, one would assume that rangelands are the source of a large proportion of the feed utilized by the grazing animals on Earth. Yet it is estimated that three-quarters of the world's 3 billion domestic ruminants are raised in conjunction with farming and are fed crop residues, hay, and other forages such as alfalfa. Roughly 38% of the world's grain, especially corn, barley, sorghum, and oats, is fed to livestock. In the United States, livestock accounts for 70% of the domestic grain consumption; in contrast, India and sub-Saharan Africa provide about 2% of their cereal harvest to livestock. A 1983 U.S. Department of Agriculture report stated that ≈103 million ha of fertilized pastures and forage crops provided 84% of the U.S. livestock forage. The ≈312–319 million ha of U.S. rangelands provided

Table 1 Some direct and indirect influences of domestic livestock on biological diversity

Desertification or degradation—the impoverishment of native plant and animal ecosystems
Deforestation/forest conversion to pasture
Contribution of greenhouse gases to the atmosphere—directly from methane emissions and carbon losses due to forest conversion
Conversion of native rangeland to cropland
Dispersal of exotic species and/or the creation of conditions suitable for their increase
Spread of disease to native herbivores
Geomorphic influences on stream channels (widening and incision) that disrupt floodplain/stream interactions
Water diversions for livestock forage production
Development of permanent water sources in areas where they traditionally did not occur
Erosion/sedimentation or fecal inputs into aquatic ecosystems
Influences on biogeochemical cycling
Control of wildlife species perceived as competitors or predators
Applications of herbicides and pesticides to control undesirable plants and insects
Altering natural fire regimes—elimination of fires in natural glasslands and shrublands or frequent introduction in pastures converted from tropical forests
Social causations—overpopulation, government incentives, culture, and customs

only 16% of the forage. Two-thirds of these rangelands are privately owned, and the rest are publicly owned. In 1984, permits issued by the U.S. Bureau of Land Management accounted for less than 4% of the forage consumed by the U.S. herds.

Rangelands of the world are important reservoirs of biological diversity as well as water, natural beauty, and other plant, animal, and mineral resources. For example, native rangelands provide essential habitat for a wide variety of wildlife and plant species. An estimated 84% of U.S. mammal species and 74% of bird species are associated with rangeland ecosystems (particularly using riparian zones within these semi-arid landscapes). Public rangelands in the United States are home to more than 3000 wildlife species, including about one-third of the nation's threatened or endangered species. Compared to forests, rangelands have received comparatively less attention with respect to environmental degradation in general, and losses in biodiversity in particular.

Because of the great differences in composition, soils, climate, and historical and current land uses among rangelands of the world, grazing effects will vary (see **Table 1**). Degradation occurs when the frequency, intensity, or season of use exceeds the biological thresholds for ecosystem persistence or recovery. Plant communities that evolved with large herds of ungulates may have adaptations that allow persistence with high levels of defoliation by domestic livestock. Conversely, those rangelands that evolved with low densities of large herbivores may be susceptible to alterations following the introduction of even low levels of domestic livestock impacts.

Unsustainable grazing and ranching practices result in the loss of forests, grasslands, shrublands, savannas, and the indigenous species that inhabit them. As the most pervasive land use, livestock grazing has had widespread and dramatic ecological impacts, including loss of native species, changes in species composition, soil deterioration, degradation of fish and wildlife habitat, and changes in ecosystem structure and function. The loss of biological diversity and productivity of range-lands due to overgrazing is a concern because it diminishes both human and ecological welfare. However, livestock influences on biodiversity are not limited to rangelands. The range of environmental effects of livestock encompasses such varied land uses as conversion of tropical rain forest to pasture in the Amazon Basin, water pollution from manure associated with large livestock factory farms in Europe and the United States, and the desertification of semiarid rangelands in the African Sahel, western United States, and interior Australia.

Effects of Livestock on Terrestrial and Aquatic Ecosystems

Figure 4 presents a simple model depicting how livestock influence biological diversity and ecosystem integrity. The direct or primary livestock influences are immediate and readily observable in the field. These direct influences include: (1) removal of vegetation through grazing; (2) the trampling of soils, vegetation, and biological soil crusts; (3) the redistribution of nutrients via forage removal, defecation, urination, gaseous loss, and animal gain; and (4) the dispersal of exotic plant species and pathogens. The cumulative effects of these four primary influences lead to a suite of physical and biotic responses or adjustments within the ecosystem or landscape. These are the secondary influences and they include alteration in disturbance regimes (fire cycles), accelerated erosion, altered hydrology (e.g., runoff, infiltration rates, and water-holding capacity), altered competitive relationships among organisms, and changes in plant and animal reproductive success and/or establishment patterns of plant seedlings. Tertiary influences are the long-term cumulative effects of domestic livestock on rangelands. These are the changes in structure, composition, and productivity of plant and animal communities that occur at community, ecosystem, and landscape scales. Tertiary influences are the overall declines in the biotic richness or diversity of affected aquatic and terrestrial areas as a result of long-term unsustainable grazing practices.

Alterations in aquatic and riparian ecosystems may include structural changes in stream channels, changes in water quality (temperature, chemistry, and sediments), and the loss of native aquatic biota. Tertiary influences are the manifestations of desertification and landscape degradation.

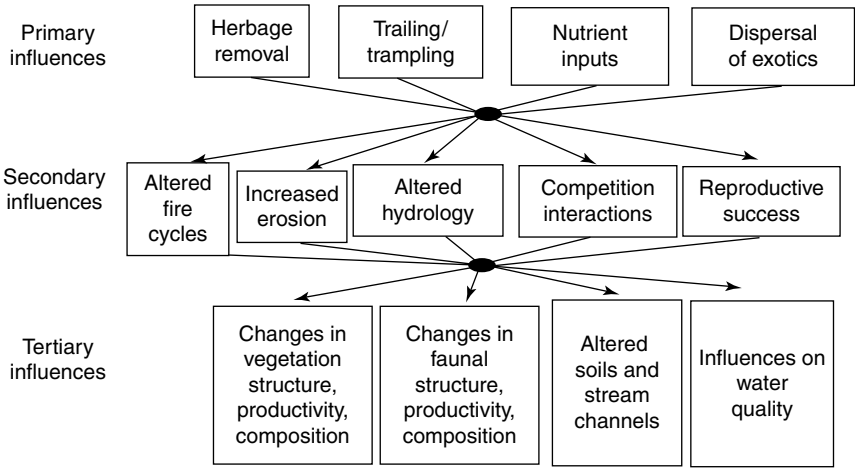


Figure 4 A conceptual model of the influences of livestock on ecosystem structure and function. Direct influences are those that are immediate and readily observable in the field. The synergistic effects of these influences lead to the secondary influences, the physical and biological adjustments in ecosystems. The tertiary influences are the long-term alterations in landscapes due to the introduction of livestock. Tertiary influences are the cumulative effects of the primary and secondary influences of livestock and are reflective of varying degrees of desertification or degradation. The degree of alteration would be affected by unique characteristics of the ecosystem (i.e., soils, topography, climate), as well as the intensity and frequency of livestock use.

Direct or Primary Influences

Forage removal by livestock can decrease the quality and quantity of forage for wild grazers as well as decrease nesting or hiding cover for other wildlife species. Vegetation removal by livestock can also reduce the protective litter layer, thus decreasing the quantity of carbon and nutrients incorporated into the soil as organic matter. Along streams, vegetation removal may affect stream shading as well as the amount of vegetation that enters into the stream. Vegetation inputs into streams are important because they are often the primary energy and nutrient source for headwater aquatic ecosystems in forests and rangelands.

The effects of herbage removal depend on a number of environmental, animal, and land management factors. The quantity of plant mass consumed and season, duration, and frequency of grazing are important determinants of how herbivory affects ecosystems. Overgrazing and degradation of ecosystems are a combination of excesses in all of these factors. The season of use refers to the phenological stage during which the plants are grazed. Typically, grazing when plants are dormant is less detrimental than grazing during active growth. Season of use may also influence the reproductive success of plants if flowers and seed heads are not allowed to form or are eaten. In addition, the degree of harm is dependent on the duration and amount of grazing. Declines in plant vigor will occur if grazing is of such duration or intensity that plants are not allowed to regrow and replenish carbon and nutrient reserves. For shrubs and trees, grazing more than the current year's growth will reduce the size of the plant and, depending on the plant's morphology, can result in the death of the plant. Finally, the frequency of grazing is an important determinant. Plants, other organisms, or soils in ecosystems in which grazing is interspersed with periods of rest have a greater opportunity for recovery than in ecosystems that are grazed continuously or annually (**Box 1**).

Trampling damage occurs via the mechanical compression and compaction of soils or the physical destruction of biological soil crusts or vegetation as animals move across the land. The break-up of soil crusts (both physical and biological) can increase soil loss via wind and water erosion. The susceptibility of soils to compaction is related to soil texture (sand, silt, and clay content), as well as the season of grazing. Clay soils are more susceptible to compaction than sandy soils and wet soils are more susceptible than drier soils. Soil compaction results in reductions in soil porosity, water-holding capacity, and rates of infiltration. All of these result in less water available for plant growth and more water lost via surface runoff.

Another direct impact of livestock is the influence on nutrient inputs and redistribution. The effects of herbivory on litter reduction, trampling, and soil compaction include direct influences on nutrient cycles. Trampling of biological soil crusts can reduce nitrogen inputs while increasing nitrogen losses via erosion. Urination and defecation are forms of nutrient redistribution. These inputs can enhance soil fertility, but can also have serious effects on water quality and aquatic organisms. Concentrated nutrient inputs can increase bacterial and protozoan pathogens, promote algal growth, and alter water chemistry—particularly the depletion of dissolved oxygen.

The final direct influence of livestock is the dispersal of undesirable exotic organisms (as well as native weeds). Seeds of exotic species may be transported on the fur or hooves of animals or deposited in the feces. As livestock move through an area, they may deposit exotic plant seeds while simultaneously preparing a seedbed via trampling. A grazing preference for desirable native species may further improve the competitive advantage of the less desired weeds, thus increasing their probability of establishment and dominance. Animals do not graze indiscriminately. Different animals have

Box 1 Riparian zones: Hot spots of biodiversity on rangelands

Riparian zones are the unique environments adjacent to rivers and streams and comprise assemblages of plant and animal communities whose presence can be either directly or indirectly attributed to factors that are stream-induced or stream-related (Kauffman and Krueger, 1984). Functionally, they are three-dimensional zones of direct interaction between terrestrial and aquatic ecosystems. Boundaries of riparian zones extend outward to the limits of flooding and upward into the canopy of streamside vegetation (Gregory *et al.*, 1991). These zones perform a variety of ecosystem functions, including influences on water quality, flood attenuation, and provision of valuable fish and wildlife habitats.

Riparian areas are hot spots of biodiversity. This is particularly true in arid and semiarid environments, where riparian zones may be the only tree-dominated ecosystem in the landscape. The presence of water, increased productivity, favorable microclimate, and unique disturbance (flooding) regimes combine to create a disproportionately higher biological diversity than the surrounding uplands. In the Intermountain West and Great Basin of the United States, about 85% of the wildlife species are dependent on riparian zones for all or part of their life cycles. In these same riparian zones, well over 100 plant species commonly can be found on a single gravel bar (<50 m in length along the streambank).

Livestock grazing is among the most significant land uses affecting the structure and productivity of riparian zones. Because of the high plant productivity, close proximity to water, favorable microclimate, and level ground, these areas are preferred habitat for livestock, and overgrazing commonly occurs. Grazing changes the composition and structure of the riparian zones, which lead to dramatic effects on the large numbers of wildlife that use the areas for feeding, breeding, or hiding cover. Of particular importance are the effects of livestock grazing on streamside forests of cottonwood, aspen, and willows. Long-term overgrazing can eliminate the woody vegetation that provides essential breeding bird habitat. In addition, trampling in riparian zones decreases infiltration and breaks down streambanks, resulting in channel incision or widening. These influences on streambank structure sever the linkages between the riparian and aquatic zones through the elimination of overbank flooding. The removal or loss of riparian vegetation results in the decline of organic inputs into the stream and the loss of overstory vegetation shade, which can dramatically affect the aquatic biota, such as salmon and trout.

Because of the high level of biodiversity and other ecosystem values of riparian zones, particularly in semiarid and arid environments, their restoration yields many positive benefits. Riparian zones are naturally an environment with frequent disturbances (floods), and because of this, the adapted biota often exhibit considerable resilience following the cessation of land uses that cause degradation or that prevent recovery. In rangelands, this most often includes abusive grazing practices. Thus, the cessation of overgrazing is the first important step in riparian restoration and should be implemented prior to determining if other active restoration measures need to be taken. For lasting riparian restoration, a larger effort should be aimed at the recovery of the entire watershed (Kauffman *et al.*, 1997).

Sources: Gregory SV, Swanson FJ, McKee WA, and Cummins KW (1991) An ecosystem perspective of riparian zones. *BioScience* 41: 540–550.
Kauffman JB and Krueger WC (1984) Livestock impacts on riparian ecosystems and streamside management implications … A review. *J. Range Management* 5: 430–437.
Kauffman JB, Beschta RL, Otting N, and Lytjen D (1997) An ecological perspective of riparian and stream restoration in the western United States. *Fisheries* 22: 12–24.

preferences and select some plant species over others. This will confer a selective advantage on less palatable species over those preferred by livestock. Often exotic species are less palatable and more grazing tolerant than native species. Livestock grazing in rangelands where native species are not well adapted to large herbivores will confer a selective advantage to exotic plant species.

Livestock also spread exotic diseases to native herbivores. Numerous herds of bighorn sheep in the United States have been eliminated following disease transmission from domestic sheep.

Indirect Effects: Secondary and Tertiary Influences

The direct or primary influences of livestock elicit a number of feedback responses that affect additional species, functions, and processes. The outcomes of these far-reaching effects are described as secondary influences and include changes in landscape disturbance cycles (e.g., fire regimes), accelerated rates of erosion, alterations in hydrology and plant available water, and alterations in successional patterns due to changes in competition and reproduction of both the native and exotic species (see **Figure 4**).

The cumulative impacts of livestock grazing are apparent in many desertified or degraded ecosystems throughout the world. The tertiary influences of livestock overgrazing on ecosystems are the characteristic endpoint of desertification or land degradation (**Figure 4**). The tertiary effects include dramatic alterations to the biotic structure, composition, and productivity of ecosystems. These are closely linked to simultaneous degradation of hydrological and soil properties. Declines in the cover of native perennial plants and biological soil crusts and concomitant increases in bare ground, unpalatable shrubs, or noxious weeds and annuals are examples of the long-term degradation of ecosystems. Tertiary effects of livestock are often difficult to separate from other significant influences on ecosystems. Degraded landscapes are often the cumulative result of inappropriate livestock management in concert with other poor land management practices, including logging, road building, inappropriate agronomy, fuel-wood harvesting, mining, water diversions, and over-harvesting of plant and/or animal resources (i.e., hunting, poaching).

Livestock Grazing and the Desertification of Native Rangelands

Desertification of rangelands has had strong negative effects on Earth's biodiversity and the continued rate of desertification represents a major global threat. Desertification is manifested in a combination of declines in native species diversity, ecosystem productivity, loss of topsoil, changes in nutrient cycles, and alterations in hydrology and microclimate. For example, Flather *et al.* (1994) reported that 40% of all U.S. federally listed Threatened and Endangered (T&E) species were associated with rangelands. More T&E plant species occur on rangelands in the United States than any

other land cover type. Trends in biodiversity loss are likely similar in other rangelands of the world. This is particularly true for large native herbivores, which are now largely restricted to small preserves or parks. Nearly all of Earth's rangelands have been degraded by human activities, including livestock grazing, the introduction of exotic species, fuelwood harvesting, alteration of natural fire cycles, wildlife depredation, and conversion to cropland or urbanization.

Estimates by the United Nations Environmental Programme (UNEP, 1990) indicated that 73% of the world's 3.3 billion ha of dry land is at least moderately desertified, having lost more than 25% of its productive capacity. It has been estimated that ≈84% of the world's rangelands are at least moderately desertified (i.e., conditions in which plant productivity has declined by ≥25%; or indications that accelerated wind or water erosion has occurred due to land use; or soil salinity has reduced crop yields 10–50%). Of the 3.1 billion ha of rangeland in the world, Mabbutt (1984) estimated that as many as 1.3 billion ha are severely degraded (i.e., productivity loss exceeds 50%).

The UNEP study on soil degradation reported that over the past 45 years about 11% (1.2 billion ha) of Earth's vegetated soils have become degraded to the point that their original biotic functions were damaged; reclamation would be costly or in some cases impossible. Most of these degraded soils are suffering from "moderate" degradation, that is, productivity is greatly reduced but can still be used for agriculture. The soil's original biotic functions (its capacity to process nutrients into a form usable by plants) have been partially destroyed and only with major improvements can productivity be restored. A smaller proportion (≈3% of Earth's surface) showed severe degradation, where original biotic functions are largely destroyed and may be reclaimable only with major financial and technical assistance. Nine million hectares of land were classified as extremely degraded, that is, unreclaimable and beyond restoration. On this land the original biotic functions are fully destroyed. At continental scales, land areas with soil degradation classified as moderate to severe ranged from 4.4% in North America, to 14.4% in Africa, to 24.1% in Central America and Mexico.

Throughout the world, accelerated wind and water erosion is the process that causes the majority of soil degradation, and overgrazing is the most ubiquitous cause of the acceleration in the erosion process. Globally 35% of the soil degradation that has occurred since 1945 was attributed to overgrazing (UNEP, 1990). Thus overgrazing is the most pervasive cause of soil degradation, affecting 679 million ha. In Africa and North America, grazing was the principal cause for 49 and 30% of the soil degradation, respectively. Soil erosion resulting from the decrease in plant cover and trampling may require thousands of years for recovery, particularly in semiarid and arid ecosystems.

Perhaps the largest data set showing trends in the ecological condition and biological impoverishment of rangelands at large scales (country or continental scales) are those of public lands in the western United States. Determining the pristine state of rangelands is difficult, and current approaches are being challenged as to their usefulness for truly ascertaining ecological condition. This is because few areas of undisturbed or pristine rangeland exist (i.e., areas with no history of livestock grazing), unlike the case with old growth or primary forest (i.e., never deforested). In addition, comparison of the degree of departure from a single hypothetical climax composition is specious at best. Keeping in mind these uncertainties associated with ecological interpretations of rangeland condition, only about 4% of nonarctic rangeland in the United States is considered to be in excellent condition (i.e., covered with native vegetation and suffering no loss in productivity). Most U.S. rangeland is in fair or good condition (i.e., slight to moderate degrees of degradation), but a significant portion (15%) is categorized as poor or severely degraded. In 1990, the U.S. Bureau of Land Management reported that only 33% of its holdings were in good or excellent condition.

Range trends, whether the rangelands are improving, degrading, or static, are a measure of how well rangelands are being managed. The results vary depending on the parameters examined. The U.S. land-managing agencies suggest that about 86% of rangelands are either improving or static (i.e., exhibit no discernable trend), and that 14% are degrading. Yet portions considered static or stable may be in such a degraded ecological state that further decline would be difficult to discern. For example, many western rangelands are completely dominated by exotic annuals or weeds, and others have experienced extensive shrub encroachments during the latter half of the twentieth century. These patterns make it difficult to interpret landscape-scale changes in ecological condition.

Though many scientists believe that many U.S. rangelands are currently in the best ecological condition of the last 60 years, others point out that more than 50% of the rangelands contain less than half of their potential natural plant composition after more than 50 years of "modern range management" (i.e., directed toward recovery and sustained use). Because of the loss of soil and native species, coupled with ongoing invasions of exotic species, many suggest that desertified North American rangelands will not likely regain their former diversity in time frames of less than a century. Research and innovative approaches for the restoration of biological diversity of rangelands are needed.

Conversion of Tropical Forests to Livestock Pasture

Livestock impacts are not limited to Earth's rangelands. Rain forests are among the most biologically diverse ecosystems in the world. Tropical forests cover about 720 million ha and contain 40–90% of the world's species. The dramatic rate of conversion of tropical forests to cattle pastures is one of the most deleterious environmental impacts on global biodiversity in the latter half of the twentieth century. Conversion of tropical forest to pasture not only results in the extirpation of native rain forest plants and animals, but also alters the microclimate, hydrology, and fire cycles, and is a significant source of greenhouse gases (**Box 2**).

Most of the deforestation in the tropical rain forests, moist forests, and dry forests of Latin America is for conversion to cattle pasture. In addition to soil, climatic, and vegetative characteristics, socioeconomic status of the landowner is a determinant of the patterns of land use. For example, on large ranches in the Brazilian Amazon, tropical forests are converted

Box 2 Biological soil crusts: Their role and susceptibility in rangeland ecosystems

Biological or microphytic soil crusts consist of a combination of nonvascular plants (also known as cryptogams; these include cyanobacteria, mosses, lichens, liverworts, and green algae). Biological soil crusts serve a number of ecosystem functions that make them an important component of the biological diversity of many of the semiarid and arid environments of the world. These soil crusts exist in nearly all ecosystems where vascular plant material (live or dead) covers less than 100% of the ground surface; they occupy the interspaces between vascular plants. Spatial coverage within some ecosystems may be greater than that of vascular plants, ranging from 10 to 100% coverage in some undisturbed temperate and arctic ecosystems. Biological soil crusts are adapted to extreme environmental conditions of extended drought and extreme high and low temperatures. These species are metabolically active only after hydration. Some are able to hydrate under conditions of saturated air, such as fog, dew point, or high vapor pressure; others require liquid water.

The functional role of biological soil crusts in semiarid and arid ecosystems is multifaceted. Some species are involved in the breakdown of humus and in the release of nutrients. However, a more significant role in some ecosystems may be their direct contribution of increasing nitrogen availability to other plants and organisms within the ecosystem through biological nitrogen fixation. Nitrogen is often an important limiting factor for ecosystem productivity. Biological soil crusts (cyanobacteria and cyanolichens) can be important sources of nitrogen for many semiarid ecosystems, such as the North American Great Basin and Colorado Plateau. Fixed nitrogen is immediately released by these organisms into the surrounding soils; however, the amount taken up by vascular plants has not been widely studied.

Biological soil crusts can diminish the rate of wind and water erosion from rangelands. Soil stability is enhanced by these soil crusts through the formation of exudates of polysaccharides that bind soil particles into aggregates. They also protect the soil from raindrop erosion (the dislodging of soil particles by individual raindrops, also known as splash erosion) by absorbing the kinetic energy of raindrops. The potential to diminish soil erosion varies with biological crust composition, increasing from algal, to lichen, to moss-covered crusts. Crusts create a surface roughness that slows the surface runoff of water and may trap soil particles arriving via wind deposition. The phototropic nature of the crust species is an adaptation that contributes to their upward growth and soil entrapment when soil buries the crusts. They are important contributors to dune stability when dunes consist of sufficient amounts of fine soil particles (silts and clays).

Complex direct and indirect mutualistic relationships exist between biological soil crust and vascular plants. In addition to nitrogen fixation and soil stability, the presence of soil crusts is positively correlated with floristic diversity in several ecosystems around the world. This positive correlation may be related to the increased seedling establishment and survival of vascular plants when grown in association with biological soil crusts. The surface variation that crust species provide may also contribute to greater variation in suitable sites for successful seedling germination and establishment.

Throughout the rangelands of the world, the introduction of large numbers of domestic ungulates has damaged biological soil crusts via herbivory and trampling. Herbivory by ungulates is generally restricted to locations where domestic reindeer have overgrazed lichen-dominated winter habitat in arctic tundra ecosystems. Globally, the most prevalent ungulate impacts are related to trampling, which breaks the dry, brittle lichens into small pieces that are then blown from the site. In many rangelands, intense grazing by livestock around watering places results in the elimination of lichens near the water; as distance from water increases, livestock impacts on lichen abundance decrease. Biological soil crusts growing on soils with low aggregate stability (i.e., sandy soils) are more susceptible to trampling damage during dry periods. The disruption of the dry soil surface breaks the bonds between the biological soil crust and the soil. Both the soil particles and the crust species are then susceptible to wind and water erosion.

Following disturbance or loss, biological crust recovery time depends on the severity of the disturbance and the environmental characteristics of the ecosystem. Primary recovery begins with cyanobacteria within one year after disturbance. Belnap (1993) estimated that some ecosystems in the U.S. Southwest may require 30–40 years for the replacement of sheath material in the soil, 45–85 years for lichen diversity to recover, and over 250 years for moss cover to return if left undisturbed. Recovery of crust species is more rapid in locations with higher effective precipitation and finer-textured soils, and if inoculating material is present. Clearly, ecosystems in which biological soil crusts are a major component need appropriate grazing management or protection to maintain or restore biodiversity and to prevent degradation, erosion, and desertification.

Source: Belnap J (1993) Recovery rates of cryptobiotic crusts: Inoculate use and assessment methods. *Great Basin Naturalist* 53: 89–95.

directly to cattle pasture. In contrast, small-scale subsistence farmers may use the land for one or two shifting cultivation cycles before ultimately converting to cattle pasture.

Slashing and burning tropical forest for cattle production converts some of the most structurally and biologically diverse ecosystems on Earth to simple pastures dominated by exotic grasses (**Figure 5**). Rain forests frequently have over 100 canopy tree species per hectare and millions of highly interdependent organisms living within the forest and soil. In contrast, the only trees, if any, in pastures are cultivated fruit and nut trees or highly invasive second-growth trees. Rather than the diverse array of birds, mammals, insects, reptiles, amphibians, and other native tropical fauna, the pastures are dominated by the cattle and the few other animals that thrive in human-modified environments.

Deforestation affects biological diversity in three ways: (1) destruction of habitat; (2) fragmentation of formerly contiguous forest habitat; and (3) edge effects within the forest that is immediately adjacent to pastures. In fragments and edge forests, the microclimate is altered such that temperatures and wind speeds are higher, and relative humidity is lower. Higher wind speeds and drier conditions result in increased tree losses due to windthrow and maybe increased water stress. The drier conditions, in concert with increased quantities of dead wood, increase the susceptibility of fragmented and edge forests to wildfires. Ignition sources for wildfire are widespread in a tropical landscape mosaic of fragmented forest and pastures because fire is frequently used to maintain the pastures. Because native flora and fauna are poorly adapted to fires in tropical rain forests, even light fires result in high rates of mortality. Low-intensity fires that spread from cattle pastures to adjacent edge forest have been found to completely kill the overstory canopy, while only $\approx$40% of the species had the capacity to sprout after burning (Kauffman, 1991).

Data from the Brazilian space agency (INPE) and the U.S. Space Agency (NASA), as well as that from other Brazilian and U.S. scientists, indicate that the Amazon Basin is experiencing the world's highest absolute rate of forest destruction. The rate of deforestation averaged 15,000 $km^2\ yr^{-1}$ from 1978 to 1988, while the rate of habitat degradation (i.e., the sum of cleared and fragmented lands) was 38,000 $km^2\ yr^{-1}$. The rate of deforestation has accelerated in recent years from about



Figure 5 Tropical rain forests are among the most biologically diverse ecosystems on Earth. The principal cause of deforestation in the Neotropics (and elsewhere) is for cattle pasture conversion. In this photo, the biologically diverse rain forest has been replaced by a pasture dominated by a few exotic grasses and invasive plants, with permission from Boone Kauffman.

11,000 $km^2\ yr^{-1}$ in 1991 to just over 20,000 $km^2\ yr^{-1}$ from 1995 to 1997. At the rates experienced in the late 1990s, 5479 ha of forest are lost every day; this is equivalent to about 3.8 ha destroyed every minute.

It is difficult to determine the total area of tropical rain forest that has been converted to cattle pasture. Fearnside (1993) reported that by 1991 the area of forest cleared in the Amazon Basin had reached 426,000 km^2. This estimate did not take into account forest areas affected by fragmentation and edge effects. Cumulatively, areas in which biodiversity has been affected by forest clearing (primarily for cattle pasture) may be over twice that of deforestation alone (Skole and Tucker, 1993). Because about 13% of the region has now been deforested (INPE, 1996, 1998), the total area affected by deforestation and fragmentation could comprise a third of the entire Amazon Basin.

Rain forests tend to receive the most attention, but tropical dry forests have been deforested to an even greater degree. These are also the most abundant type of tropical forest. Few areas of intact (i.e., uncut or primary) tropical dry forest still exist in Latin America today (e.g., the Brazilian caatinga and tropical deciduous forests of Central America and Mexico).

Although scientists and economists have questioned both the sustainability and economic viability of cattle pastures in areas of former tropical forest, the deleterious consequences to biological diversity are evident. If cattle pasture production is neither sustainable nor economically viable, then why are rates of deforestation so high? Contrary to popular belief, deforestation in the Americas has not been a response to international demands for beef nor a response to increasing human populations (hoof-and-mouth disease is prevalent in Brazilian cattle, which precludes importation into the United States, and the Amazon Basin supports only $\approx$5% of all cattle production in Brazil). Deforestation largely occurs because it is a way to obtain title to the land, or to profit from land speculation, government financial incentives, and subsidies. In addition, labor costs on ranches are quite low compared to intensive agriculture or agroforestry systems. The environmental limitations in many tropical landscapes, coupled with governmental and social forces, combine to promote a land use that produces little food and little direct monetary return, but tremendous environmental degradation (Hecht, 1983).

Livestock Grazing and the Greenhouse Effect

Livestock production and related land uses are significant sources of many greenhouse gases. In addition, changes in climate due to the greenhouse effect are predicted to have strong feedbacks on the environments that provide feed and water for livestock. Among the most significant influences of the livestock industry in relation to the greenhouse effect is slash burning, which occurs during the conversion of tropical rain forest to pasture. When slashed tropical rain forest burns, as much as one-half of the aboveground carbon (76–112 Mg ha^{-1}) may be released as CO_2, CO, and other radiatively active gases. As much as 800–1600 kg nitrogen ha^{-1} are also released by burning forest (Guild *et al.*, 1998). When tropical dry forest is slashed and burned, the proportion of biomass consumed is even greater—as much as 95% of the aboveground carbon pool may be released as greenhouse gases during slash fires. The burning of the African savannas, most often initiated by herders for forage enhancement, is another important source of CO_2. Savanna fires contribute as much as 16% of the total annual CO_2 arising from agricultural sources. However, much of this will be absorbed the following year by regrowing grasses. The same is not true for forest that is converted to pasture; the biomass of pasture grasses will be about 5% of the forests that it replaced.

Methane is a powerful greenhouse gas with a radiative absorption capacity that is 21 times that of CO_2. The global increase in methane results from human activities such as livestock production, manure management, rice production, landfills, and the production and use of oil, gas, and coal. Livestock and manure management contribute about 16% of the total annual methane production (about 87 million Mg out of 550 million Mg produced each year; de Haan *et al.*, 1996). Methane emissions from ruminants (cattle, sheep, and goats) result from their capacity to utilize large amounts of fibrous grasses. Low productivity and poor feed quality are characteristic for most land-based ruminant production systems in arid regions. Poor feed quality also typifies pastures in the tropics and subtropics. In these scenarios, emissions per animal are higher than in scenarios where animals are on higher-quality feeds. With the conversion of tropical forest to cattle pasture, not only are large carbon sinks shifted to significant global atmospheric carbon sources, but the resulting land use becomes a significant global methane source.

Other Factors

Conversion of Native Rangelands to Croplands or Pasture

In addition to the effects previously mentioned, there are many other livestock-related land use activities and industries that affect global biological diversity (see **Table 1**). Grazing on native grasslands and shrublands supports a dwindling share of the livestock forage supply of the world. As intensive livestock production has expanded and the importance of grazing native rangeland has declined, there have been large conversions of rangelands to seemingly more lucrative land uses. The highly diverse Brazilian cerrado, a mosaic of savanna, grasslands, and open evergreen woodlands, covers about 2 million km^2 in Brazil. The cerrado is undergoing a rapid conversion from species-rich savanna/woodland to planted pastures, soybeans, and other annual crops. This region has experienced more land cover change than the Amazon forests (i.e., about 600,000 km^2 of the cerrado had been cleared as of 1991, compared to about 400,000 km^2 for the Brazilian Amazon; Klink *et al.*, 1994).

Many rangelands of the U.S. Southwest and southern Great Plains were plowed into croplands when government subsidies made it cost-effective to utilize groundwater for irrigation (i.e., the Ogalala aquifer). Often, these water sources were rapidly depleted and groundwater became prohibitively costly to pump. This resulted in the abandonment of lands that are now in a desertified state—a depauperate cover of native vegetation coupled with high rates of soil erosion.

Exotic Species and Pathogens

Invasions of exotic species are among the greatest threats to global biodiversity and livestock play significant direct and indirect roles in facilitating their establishment. Livestock influence exotic plant distribution through seed dispersal in fur and dung. Establishment of exotics is enhanced via creation of seedbeds through trampling and forage removal. In those regions where large herbivores were not a strong selective pressure on the native flora (i.e., much of the Pacific Northwest, Intermountain West, and California, in the United States), exotic annuals now dominate areas formerly dominated by perennial bunchgrasses. Indirectly, the livestock industries also facilitated introductions and spread of exotic plant species through cropland plantings (fodder and feed grains) that contained weed seeds.

In contrast to unanticipated shifts of native ecosystems to exotic species dominance, many native rangelands and forests have been converted to exotic grass dominance through purposeful seeding. The goals of these so-called "range improvements" were to increase forage for livestock, decrease soil erosion, and provide a desired alternative to weeds. However, there was scant consideration for influences on biological diversity. For example, in the United States millions of hectares formerly occupied by shrub/bunchgrass ecosystems have been seeded with grasses of exotic origin (e.g., crested wheatgrass, *Agropyron cristatum*, and Lehmann lovegrass, *Eragrostis lehmanniana*).

Livestock, Water Developments, and Aquatic Ecosystems

Livestock and related management activities have strong effects on the biodiversity of aquatic ecosystems. These not only include direct effects such as consumption of streamside or wetland vegetation, trampling of streambanks, and fecal inputs, but also indirect effects such as irrigation withdrawals, water developments, and wetland draining. These activities can alter aquatic diversity through degradation of aquatic habitats as well as water quality (temperature, chemistry, and microbial composition). Numerous studies have documented accelerated streambank erosion and losses of streambank structure due to livestock grazing. The changes and the accrual of sediment in the channel degrade spawning and reproductive habitats for fishes and aquatic insects. Overgrazing results in the simplification of stream channels, which may include loss of channel sinuosity, increases in channel widths, increased channel incision, and decreases in deep pools. Incised and simplified channels result in the elimination of important linkages between the floodplain and stream channel that positively influence biodiversity.

Water diversions for pasture or forage crop irrigation are a widespread influence on the biodiversity of rivers streams and their associated riparian zones. Diversions range from the dewatering of small streams for the irrigation of a single pasture or farm field to huge dams that span major rivers. Regardless of the scale, when water is diverted or dammed, the natural hydroperiod of the river or stream is altered (i.e., changes in the timing and magnitude of peak and low flows throughout the year). This can have dramatic influences on channel-forming processes and riparian community development. Because many aquatic and riparian organisms are adapted to cycles of floods and flood effects, these hydrological alterations can completely alter their environment, leading to species losses. Throughout the western United States, livestock-related agriculture is a dominant user of irrigation water stored in federally-constructed dams. Significant detrimental effects on native fishes, as well as on biologically diverse riparian plant and animal communities, occur both upstream and downstream of the dams or diversions.

Land rehabilitation and forage enhancement efforts have been attempted throughout degraded rangelands of the world. However, these have often had disappointing results. Because

of misinterpretations of ecosystem needs or ignorance of ecological consequences, rehabilitation efforts have sometimes exacerbated the deterioration and degradation that they were intended to reverse. A striking example is the establishment of permanent water sources in dry season grazing areas in Africa (similar developments have also been implemented in the U.S. West). Water sources were developed in areas so distant from surface water that they were not traditionally used by livestock during the dry seasons. Following developments such as the drilling of wells or construction of small catchment basins, lands around these African water sources were often severely overgrazed and trampled. In addition to losses in biodiversity, these water sources have fostered the growth of herds beyond the carrying capacity of the land, resulting in range deterioration and livestock deaths during droughts. During droughts, livestock did not die from lack of water but from starvation. The ultimate consequences include degradation of the native ecosystem coupled with increased human suffering and hardship.

Water pollution is a common problem associated with intensive livestock production in relatively confined spaces. Manure is a valuable organic fertilizer and soil builder in modest amounts, but it is a dangerous environmental hazard when waste production exceeds the absorptive capacity of land and water. In fecal-polluted waterways, pathogens detrimental to both the native biota and humans are present. In addition, increased nitrogen and phosphorus concentrations can result in the eutrophication of waterways, in which algae blooms rapidly consume oxygen to the point where fish kills occur.

Biogeochemical Cycling

Livestock grazing can affect nutrient cycling in numerous ways. The most obvious would arise from the dramatic effects associated with accelerated erosion, trampling of soils, and the replacement of perennial plants (and often biological soil crusts) with annuals or shrubs and bare ground. The direct consumption of vegetation changes the patterns and distribution of litter, which affects decomposition and nutrient cycling. Consumption of herbage resulting in less litter also increases the proportion of bare ground, thereby increasing susceptibility to erosion. Accelerated erosion results in the loss of nutrients, soil organic matter, and the capacity of the soil to store water and retain nutrients. In riparian zones, diminished litterfall can also affect the aquatic biota. For many headwater streams, the nutrients and energy that drive in-stream aquatic ecosystems are derived from terrestrial inputs. It is particularly deleterious to aquatic ecosystems when long-term grazing results in the elimination of shrub and tree overstories in streamside environments.

The synergistic effects of forage removal and trampling damage by livestock have been described in Oregon riparian meadows that were grazed for over a century. Grazed meadows had lower rates of nitrogen mineralization and lower soil organic matter than areas where grazing had been halted for 9–12 years. These indicators of decreased productivity were associated with dramatically lower water infiltration rates, higher soil bulk density, lower levels of residual litter, and lower root biomass in the grazed sites.

Desertification often results in the alteration of natural desert grasslands and savannas into shrublands with bare soil in the intershrub spaces. Under these scenarios, nutrient and water resources become concentrated under shrubs. Livestock grazing can also concentrate nutrients in areas where livestock may congregate through the concentration of urine and feces (i.e., resting places, water holes, or streamsides).

Attempts to Control Undesirable Species

Other activities associated with livestock management that affect biodiversity include predator control (e.g., eliminating carnivores via trapping, poisoning, or shooting) and other forms of wildlife control for species that are perceived as competitors with the livestock industry. These activities, called "animal damage control," are often carried out without consideration of their effects on biodiversity, ecological processes, or ecosystem functions. Techniques of animal damage control are often not specific to the "target species" and other "nontarget" species can be eliminated as well. Poison bait traps and many pesticides often kill nontarget animals (and insects). The black-footed ferret (*Mustela nigripes*) was almost driven to extinction because of the decline of the prairie dog (*Cynomys* spp.), its major prey. Prairie dog populations have precipitously fallen due to habitat loss by agricultural development and from purposeful elimination because they eat the same forage as livestock. Many large carnivores have been eliminated in regions because of their real or perceived threats to livestock. The livestock industry interest groups provide the most vehement opposition to the reintroductions of large carnivores to extirpated ranges (e.g., wolves to Yellowstone National Park).

As with predator control, the use of petrochemicals (herbicides and pesticides) to control unwanted species of plants and insects has sometimes resulted in undesirable effects on non target species. In animal, plant, and insect control programs, scant attention has been paid to the functional role that these organisms may play in ecosystems and possible feedback responses to their removal. In addition, questions regarding the underlying reasons for the increase in pest or weed damage frequently were not asked or left unanswered. In the western United States, practices such as mechanical and chemical removal of shrubs and trees were implemented for the benefit of increased cattle forage, but at the expense of other human and ecological values. For example, widespread rangeland herbicide applications eliminated traditional food and medicinal plants used by indigenous peoples on many western U.S. rangelands. This is because herbicides applied to control unwanted plant species also kill a suite of other native broad-leaved species. Species with limited capacities for reinvasion may be locally extirpated from sites after a single herbicide application. Similarly, insecticides may lack selectivity, such that a suite of species performing a multitude of ecosystem functions are eliminated. Many of these chemicals act as environmental estrogens and their elevated concentrations within secondary consumers (predators and scavengers) have resulted in reproductive and physiological maladies.

Altering Fire Regimes

In many grasslands and savannas of the world, fire is a dominant disturbance feature of the landscape. The high productivity of flammable grasses, climatic conditions with

long dry periods, and a prevalent ignition source (lightning or humans) result in the frequent occurrence of fires. Fires in grasslands and savannas are important in cycling nutrients and in influencing biotic structure and composition. Both the flora and fauna of grasslands and savannas are adapted to, and often dependent on, fire for their continued existence. Altering the fire regime can lead to a vastly different structure and composition. Livestock alter fire regimes by removing the fuels (i.e., the grasses) that carry the fire in rangelands and savannas. For example, throughout the U.S. West, livestock grazing has resulted in the diminution of fires in rangelands with a concomitant increase in shrub species that are not as fire-adapted. Increases in the abundance of juniper (*Juniperus* spp.) and mesquite (*Prosopis* spp.), as well as declines in aspen (*Populus tremuloides*), are related to the synergism of livestock and the decline of fire occurrence. Long-term changes in ecosystem structure due to grazing and fire suppression have resulted in a scenario in which simple livestock removal would not completely remedy the ecosystem decline; vegetation manipulations coupled with the reintroduction of fire would also be necessary components of ecological restoration.

Conclusions: Human Institutions and Livestock-Caused Degradation

Livestock create an array of environmental problems, not because cows, sheep, goats, and other grazers are hazards in themselves, but because human institutions have forced animal farming out of alignment with the ecosystems in which they are practiced. The root causes of land degradation, including deforestation and desertification, lay not in the animals, but in the social and political institutions and population pressures placed upon the environment. Sociopolitical causations must be addressed if degradation of biodiversity by livestock is to be halted and ecosystems restored.

In many countries there persists government-based subsidies and economic incentives that result in the expansion of livestock production that accelerates environmental degradation. Government policies on pricing, taxing, and land titling incentives affect resource use by influencing decisions about type of use, inputs, technology adoption, and investments in development. Policies have misguided livestock development through the subsidized pricing of inputs and products that induced the non-sustainable use and degradation of natural resources. Resources have been priced too low to reflect their true environmental costs, thus passing the debt of environmental degradation along to future generations. In the United States, subsidized grazing on public lands and federal dams that provide water for irrigated pasture and fodder production are prominent examples. In the Brazilian Amazon, land speculation, titling procedures, and government financial incentives have been the main reasons for the conversion of tropical rain forests into cattle pastures. In many less-developed countries, strong human population growth is fueling demand for livestock products while at the same time limiting the traditional sources for livestock production. Increases in per capita income and urbanization are raising the demand for livestock products. The growing social inequalities among the rich and the poor result in different motivations for degradation: the greed of the rich and the desperation of the poor.

Desertification is the result of inappropriate land use activities, including overgrazing, overcultivation, salinization due to irrigation practices, and deforestation. Although differences exist among different rangelands, cultures, and countries of the developing world, the underlying forces of degradation include population densities that are greater than the land can sustain and, more fundamentally, social and economic inequities that push people into marginal environments and vulnerable livelihoods. Particularly important in the developing world is the inequitable distribution of land among the rich and poor. When a disproportionate share of the land is kept in the hands of a few, the ability of the impoverished majority to sustainably manage the marginal lands that they control is severely compromised.

The first step in the ecological restoration of degraded rangelands is the cessation of those activities that are causing degradation or preventing recovery (Kauffman *et al.*, 1997). Changing the social and institutional incentives that result in the degradation of ecosystems may be more difficult to achieve than the biological and physical aspects of ecosystem restoration. It must be recognized that sustained productivity and livelihoods will most likely occur when lands are managed in a manner that closely approximates their evolutionary development and where ecosystem processes are allowed to occur with minimal disruptions or alterations. In rangelands, management that mimics the natural grazing patterns of large wild herbivores may be the most sustainable and successful in maintaining biological diversity. Practices that maintain functional ecosystems rather than attempt to maximize short-term meat production are an integral feature of sustained land management. Similarly, if functional features and biodiversity are to be maintained in forests, the utilization of standing rain forest should be encouraged rather than replacing them with cattle pastures.

See also: Agriculture, Sustainable. Agriculture, Traditional. Biogeochemical Cycles. Cattle, Sheep, and Goats, Ecological Role of. Desertification. Fires, Ecological Effects of. Plant–Animal Interactions

References

de Castro EA and Kauffman JB (1998) Ecosystem structure in the Brazilian Cerrado: A vegetation gradient of aboveground biomass, root mass and consumption by fire. *J Tropical Ecol* 14: 263–283.

Dodd JL (1994) Desertification and degradation in sub-Saharan Africa: The role of livestock. *Bioscience* 44: 28–34.

Fearnside PM (1993) Deforestation in Brazilian Amazonia: The effect of population and land tenure. *Ambio* 22: 537–545.

Flather CH, Joyce LA, and Bloomgarden CA (1994) *Species Endangerment Patterns in the United States.* USDA Forest Service General Technical Report RM 241. Ft. Collins, Colorado: Rocky Mountain Forest and Range Experiment Station.

Guild LS, Kauffman JB, Ellingson LJ, *et al.* (1998) Dynamics associated with total aboveground biomass, C, nutrient pools and biomass burning of primary forest and pasture in Rondonia during SCAR-B. *J. Geophys. Res* 103: 32,091–32,100.

de Haan C, Henning H, and Blackburn H (1996) *Livestock and the Environment. Finding a Balance.* European Commission Directorate-General for Development.

Hecht SB (1993) The logic of livestock and deforestation in Amazonia. *Bioscience* 43: 687–695.

INPE (1996) *Deforestation Estimates in the Brazilian Amazon, 1992–1994*. Brasilia, Brasil: Instituto Nacional de Pesquisas Espacais, Diario Official da Uniao.

INPE (1998) *Deforestation Estimates in the Brazilian Amazon, 1995–1997*. Brasilia, Brasil: Instituto Nacional de Pesquisas Espacais, Diario Official da Uniao.

Kauffman JB, Beschta RL, Otting N, and Lytjen D (1997) An ecological perspective of riparian and stream restoration in the western United States. *Fisheries* 22: 12–24.

Klink CA, Macedo RH, and Mueller CC (1994) *Cerrado: Processo de Ocupacao e Implicacoes para a Conservacao e Utilizacao Sustenavel de Sua Diversidade Biologica*. World Wildlife Fund. *Brasil Report*. 104.

Mabbutt JA (1984) A new global assessment of the status and trends of desertification. *Environ. Conservation* 11: 103–113.

Oldeman LR, Hakkeling RTA, and Sombroek WG (1990) *World Map of the Status of Human-Induced Soil Degradation: An Explanatory Note,* revised 2nd ed. International Soil Reference and Information Center/United Nations Environmental Programme. Netherlands: Wageningen.

Sinclair ARE (1979) The eruption of the ruminants. In: Sinclair ARE and Norton-Griffiths M (eds.) *Serengeti: Dynamics of an Ecosystem*. Chicago: University of Chicago Press. 82–103.

Skole D and Tucker C (1993) Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science* 260: 1905–1910.

Steinfeld H, de Haan C, and Blackburn H (1996) *Livestock–Environment Interactions—Issues and Options*. European Commission Directorate-General for Development.

United Nations Environmental Programme (1990) *Global Assessment of Soil Degradation*. Kenya: UNEP, Nairobi.

Wolf EC (1986) Managing rangelands. In *State of the World 1986*. Washington, D.C.: Worldwatch Institute.

World Resources Institute (1994) *World Resources 1994–1995*. New York: Oxford University Press.

Rare Biosphere

Carlos Pedrós-Alió, Institut de Ciències del Mar, CSIC, Passeig Marítim de la Barceloneta, Barcelona, Spain

© 2013 Elsevier Inc. All rights reserved.

Glossary

Collectors curve Plot of the number of species found versus the number of individuals examined. The number of species generally rises very quickly at first and then levels off toward an asymptote as fewer new species are found per unit of individuals collected.

HTS High-throughput sequencing. A number of techniques that are able to automatically process many sequences in parallel, thus obtaining hundreds of thousands of sequences simultaneously.

LSU rDNA The gene in DNA that codes for the ribosomal RNA in the large subunit (LSU) of the ribosome. This is 23S rRNA in prokaryotes and 28S rRNA in eukaryotes.

Metagenomics An approach to collect and sequence all the genomes in a natural environment.

Morphospecies A species defined on the basis of morphology alone. This criterion has been traditionally used with many eukaryotic organisms, when information about interbreeding or molecular composition was not available. In microorganisms it clearly underestimates real diversity.

OTU Operational taxonomic unit. As the description implies, any pragmatically defined (but objective) way of classifying the individuals into mutually exclusive bins will generate OTUs. For example, those sequences sharing a similarity of the LSU rDNA larger than 97%.

PCR Polymerase chain reaction. A method for amplifying a target piece of DNA producing between thousands and millions of identical copies.

Rank-abundance curve A plot of the number of individuals in each species ordered by their rank, starting with the most abundant species and ending with the most rare.

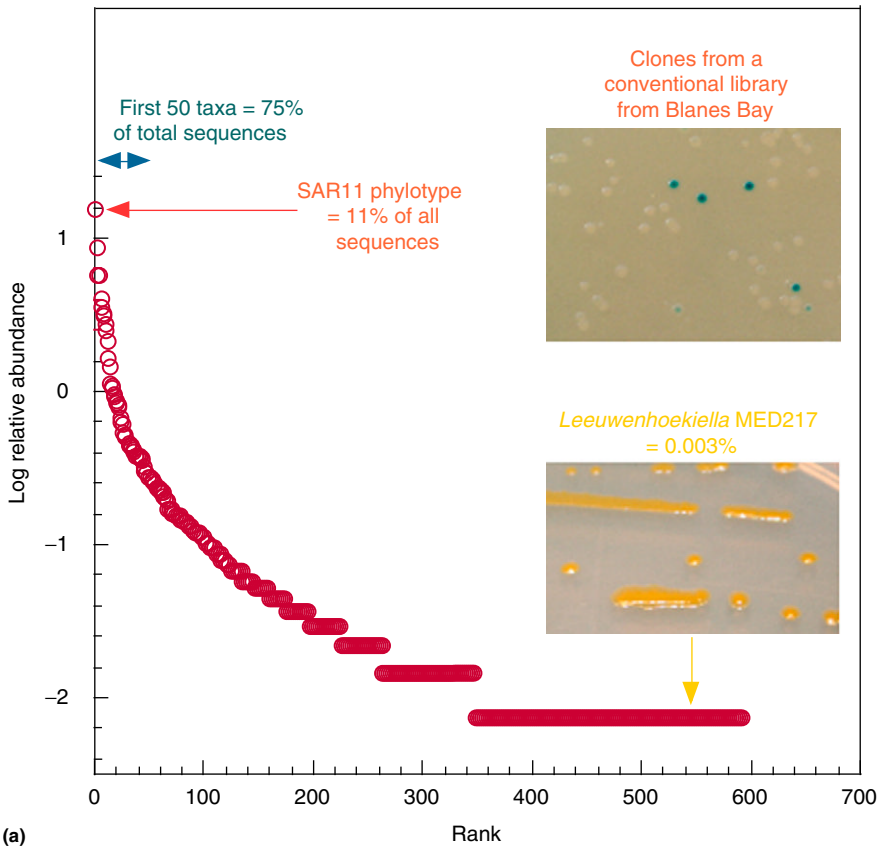
SSU rDNA The gene in DNA that codes for the ribosomal RNA in the small subunit (SSU) of the ribosome. This is 16S rRNA in prokaryotes and 18S rRNA in eukaryotes.

What is the Rare Biosphere?

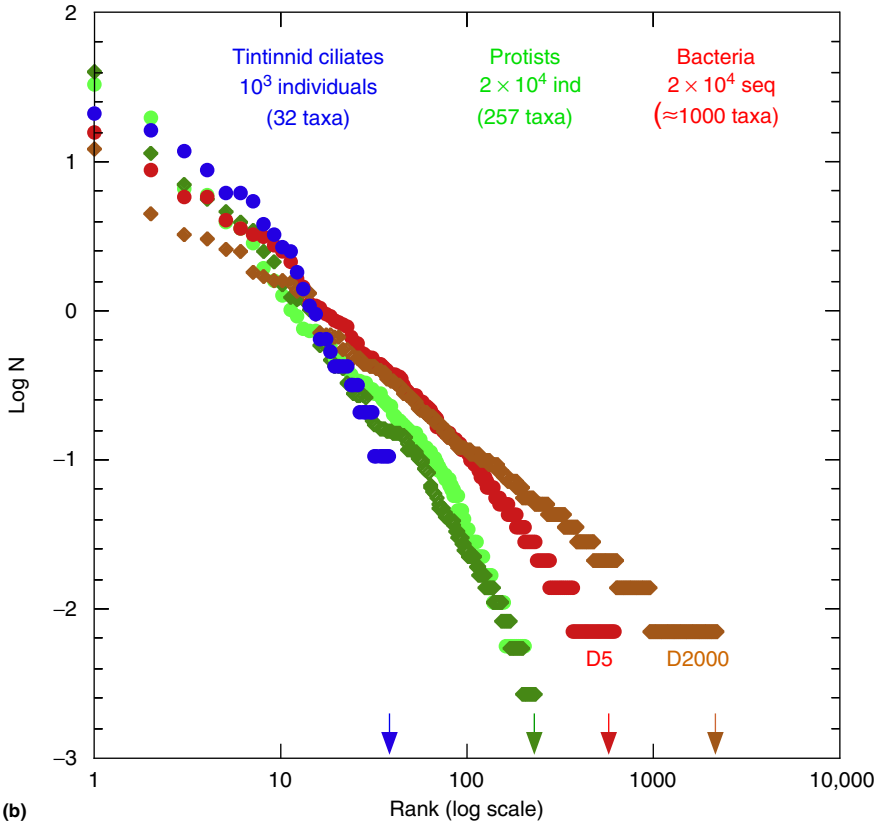
Sogin *et al.* (2006) defined the rare biosphere with the following sentences: "The large number of highly diverse, low-abundance OTUs constitutes a 'rare biosphere' that is largely unexplored… This 'rare biosphere' is very ancient and may represent a nearly inexhaustible source of genomic innovation. Members of the rare biosphere are highly divergent from each other and, at different times in earth's history, may have had a profound impact on shaping planetary processes."

The concept is easy to understand by looking at a rank-abundance curve (**Figure 1**). As stated by Sogin *et al.* (2006): "A relatively small number of different populations dominate all samples, but thousands of low-abundance populations account for most of the observed phylogenetic diversity." This is the case for most living beings in most ecosystems. What makes the case of microorganisms special is that the "depth" of this rare biosphere is at least one order of magnitude (and likely two or three) larger than that of animals and plants. As a consequence, the difference in abundance between the most abundant and the most rare species ranges several orders of magnitude. This effectively "hides" members of the rare biosphere from their detection with conventional molecular techniques. Hundreds or even thousands of individuals of the more abundant species have to be examined before a single representative of a rare one is encountered. Thanks to current developments in high-throughput sequencing (HTS) technology, exploring this rare biosphere is now beginning to be possible.

In fact, the paper by Sogin *et al.* (2006) not only introduced the term, but also a method to study it based on massive parallel sequencing. The procedure included one initial polymerase chain reaction (PCR) step to amplify all the small subunit (SSU) rDNA in the sample and then a round of 454-pyrosequencing. The latter technique provides around a million simultaneous sequences from a sample. Since the previous PCR step insured that only SSU rDNA molecules were present, all the sequences obtained in the second step were relevant for the study of diversity. Thus, for the first time, hundreds of thousands of sequences could be analyzed, compared to the few hundreds that were practical with conventional cloning and sequencing. Metagenomics can also provide many sequences and, moreover, without the biases introduced by the PCR step. However, precisely because this PCR step is not there, the sequences obtained correspond to all the genes of all the microorganisms in the sample. As a result, only a very small fraction of the sequences is relevant for diversity studies. Venter *et al.* (2004), for example, obtained over one million genes from the Sargasso Sea, but only about 4000 rDNA molecules. The two approaches are useful for different purposes. The second step, the pyrosequencing, also has potential problems. Particularly, even if sequencing errors are a very small percentage of the total (as is the case), when one has to deal with hundreds of thousands of sequences, there will be a significant number that will be erroneous. Distinguishing between the good and the bad sequences is a delicate methodological issue that has received considerable attention (Huse *et al.*, 2010; Quince *et al.*, 2011, and references therein). For the purpose of the present article, we do not need to worry. It is sufficient to realize that current sequencing technologies provide from one to two orders of magnitude more sequences than previous conventional molecular techniques and this fact now allows the exploration of the rare biosphere.


First 50 taxa = 75%
of total sequences
SAR11 phylotype
= 11% of all
sequences
Clones from a
conventional library
from Blanes Bay
Leeuwenhoekiella MED217
= 0.003%
Log relative abundance
1
0
−1
−2
0
100
200
300
400
500
600
700
Rank

(a)


Tintinnid ciliates
10^3 individuals
(32 taxa)
Protists
2×10^4 ind
(257 taxa)
Bacteria
2×10^4 seq
(≈1000 taxa)
D5
D2000
Log N
2
1
0
−1
−2
−3
1
10
100
1000
10,000
Rank (log scale)

(b)

The Number of Microbial Species

Naturalists, environmentalists, and ecosystem managers like to have a checklist of the species present in a given province, or a given nature reserve, or a national park. These checklists are very useful for decision making. For example, a park with several endangered species present will require a higher priority for conservation efforts than a park where none of the species is in danger. The checklists are also useful for visitors, who can familiarize themselves with the organisms they can expect to find in the park when visiting it. Obviously, checklists are at the base of both conservation biology and ecotourism. In most cases, checklists are available for mammals and birds. In some cases there are checklists also available for amphibians, reptiles, and fishes. Usually, checklists of insects only include butterflies. In some very well-studied environments, there will be a long list of bees, ants, beetles, and all the rest. But, what about the microbes? Is there a checklist of the microbes of any place? Most of the time, managers and environmentalists don't even think about including microbes in the inventory of diversity of a place. But, sometimes, conscientious biologists realize the microbes are missing and then they ask a microbial ecologist the dreadful question: "Do you know how many species of microbes are there in this ecosystem?" The answer, of course, is "no." But the conscientious biologist will not stop there. "This is your chance to make the microbes visible, to put them on a par with the larger animals and plants. Are you going to miss this opportunity?" "Give an educated guess, one hundred? One thousand? A million?"

A substantial part of this paper is a long answer to this question and the reasons why despite microbes making up most (and by most, the author means more than 95%) of the biodiversity on Earth, microbiologists cannot come up with an estimate of the number of species in any single place; much less the number of microbial species on the Planet, particularly the hundreds, or thousands, or perhaps millions that make up the rare biosphere.

The Biological Species Concept and the Phylo-Phenetic Concept in Microorganisms

First of all we should determine what a species is. Despite the fact that there are more than 20 concepts for the definition of species, the most widely used is the biological species concept. Two populations belong to the same species if their members can reproduce sexually and their descendants are fertile. Bacteria, however, do not have sex. They do have several mechanisms to exchange genetic material, but not sex. Therefore, the biological species concept is not applicable to bacteria. The most commonly used definition of a bacterial species is a pragmatic one, the phylo-phenetic species concept: A species is "a monophyletic and genomically coherent cluster of individual organisms that show a high degree of overall similarity in many independent characteristics, and is diagnosable by a discriminative phenotypic property." (Rosselló-Mora and Amann, 2001). The basic requirement to use this concept is to have the bacterium in pure culture, because the definition requires a series of genetic and physiological measurements that can only be done in culture. Once isolated, the taxonomist can determine the sequence of the SSU rDNA, whether the organism is a strict aerobe or whether it can grow by fermentation or not. Moreover, its genome can be sequenced and DNA–DNA hybridization experiments can be carried out to determine the relationship to its closest relatives.

It would be desirable to compare the phylo-phenetic concept with the biological species concept. This would tell us whether we are treating all living beings in a coherent way or whether we have a bias to be splitters in one domain of life while being lumpers in a different domain. Fortunately this can be done, not with bacteria, but with some Protists that do have sex. One particularly good example is that of the diatoms of the genus *Pseudo-nitzschia*. Many diatoms do have plus (+) and minus (−) strains. A+ strain can mate with a – strain of the same species and carry out sexual reproduction, with a mechanism equivalent to that of animals and plants.

Figure 1 (a) Rank-abundance curve for the bacteria in a sample from the Northwestern Mediterranean Sea. The curve plots the number of individuals of each species vs. the rank of that species. In this case, operational taxonomic units (OTUs) are used as proxies for species. The plot follows the convention of showing relative abundance in a log scale in the *Y*-axis and rank in a linear scale in the *X*-axis. There is a 4-log difference in the abundance of the most abundant OTU (a SAR11 phylotype, in orange) that accounted for 11% of all the sequences, and one of the least abundant ones such as *Leeuwenhoekiella blandensis* MED217 that was found to comprise only 0.003% of the sequences (in yellow). The former was the most abundant in conventional clone libraries but has not been isolated in pure culture. The latter could be isolated in pure culture but was not found in conventional clone libraries. The 50 most abundant OTUs (blue arrow) made up 75% of all the sequences, but most of the OTUs were very rare. The latter form the rare biosphere. (b) Rank-abundance curves for different microorganisms in samples taken in the Northwestern Mediterranean Sea. The plot retains the conventional log scale for the *Y*-axis (as in (a)), but shows the *X*-axis also in a log scale to facilitate comparisons across the whole range of abundances. Tintinnid ciliates and protists were binned into morphospecies using inverted microscopy. Bacteria were binned into OTUs based on 454-pyrosequencing of the V6 region of the SSU rDNA. D5 and D2000 are samples from the surface and the bottom of an open sea station. The total number of taxa (arrows on the *X*-axis) was 10^1 for Tintinnids (blue arrow), 10^2 for protists (green), and 10^3 for bacteria (at least in the deep sample, brown arrow). Only 10^3 individuals were examined for Tintinnids and, therefore, their curve is truncated for lack of data. However, the same number of individuals (or sequences) was examined for protists and bacteria. Morphospecies, however, are known to lump together many species. Thus, if true species could have been determined for protists, their curve would have been closer to the curves for bacteria. Tintinnid data from Dolan JR and Marrasé C (1995) Planktonic ciliate distribution relative to a deep chlorophyll maximum: Catalan Sea, N.W. Mediterranean, June 1993. *Deco-Sea Research I* 42(11112): 1965–1987; protist data from Margalef R (1969) Estudios sobre la distribución a pequeña escala del plancton marino (in Spanish). *Memorias de la Real Academia de Ciencias y Artes de Barcelona* 60: 1–22; Estrada M (2008) Explorations on phytoplankton diversity. An appreciation of Ramon Margalef's Contributions. In: Valladares F, Camacho A, Elosegui A, *et al.* (eds.) *Unity in Diversity. Reflections on Ecology after the Legacy of Ramon Margalef*, pp. 191–222. Bilbao: Fundación BBVA, and bacterial data from Pommier T, Neal PR, Gasol JM, Coll-Lladó M, Acinas SG, and Pedrós-Alió C (2010) Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquatic Microbial Ecology* 61: 221–233.

The descendants are fertile. The + and − strains of different species cannot have a successful sexual reproduction. This, therefore, fulfills the biological species concept. How does this compare to the phylo-phenetic definition of species in diatoms? Amato *et al.* (2007) studied 95 clones of *Pseudo-nitzschia* isolated from the Bay of Naples. Traditional taxonomy of these diatoms relies on light microscopy. Morphologically, the clones studied fit into two different morphospecies: *Pseudo-nitzschia pseudodelicatisima*-like and *Pseudo-nitzschia delicatissima*-like. To use the phylo-penetic concept, the researchers analyzed the *rbc*L gene, the large subunit (LSU) rDNA gene, and the interspacer regions (ITS1) and (ITS2). The *rbc*L gene is a functional gene (located in the chloroplast) that codes for the LSU of ribulose biphosphate carboxylase, the central enzyme for carbon fixation in photosynthesis. The LSU rDNA is similar to the commonly used molecular marker for all microorganisms (SSU rDNA). The ITS1 and ITS2 are noncoding regions separating the three ribosomal RNA genes in the RNA operon. Because they do not code for anything, these ITS regions are extremely variable and allow separation of very closely related organisms. In parallel, Amato *et al.* (2007) tried all possible sexual crossings among the different clones. According to the biological species concept there were eight species hidden in the two morphospecies (**Figure 2**). The ITS regions clustered the clones into the same eight species, while the LSU rDNA produced only seven; that is, two of the species had identical LSU rDNA. Once they had this information, Amato *et al.* (2007) did a careful study of the morphology of the different species by electron microscopy and, effectively, they could find small, but clear differences among the eight species. This example shows that when it is possible to compare them carefully, the biological species and the phylo-phenetic concepts are quite similar in the delineation of species. It also demonstrates that the rDNA underestimates real diversity.

Pure Cultures Versus Sequences

Thus, if we could isolate in pure culture all the bacteria inhabiting the oceans, we could proceed to define species with the phylo-phenetic concept and come up with an estimate of the total bacterial diversity that could be compared reasonably well with the diversity of animals and plants. Unfortunately, many microorganisms, some would say most, have not been isolated in pure culture. Despite recent efforts to design better culturing techniques, it is estimated that perhaps only 1% of the bacteria in the oceans have been isolated in pure culture. Worst, some of the most abundant bacteria have not been isolated and many of those in pure culture are not abundant in nature. The microbial ecologist, therefore, has to resort to something less desirable, but pragmatic. The choice that has been made is to use the SSU rDNA as a probe for diversity studies. One good thing about this choice is that, as was evidenced in the *Pseudo-nitzschia* example presented above, the biological species concept is more of a splitter than SSU rDNA. If we use SSU rDNA to determine

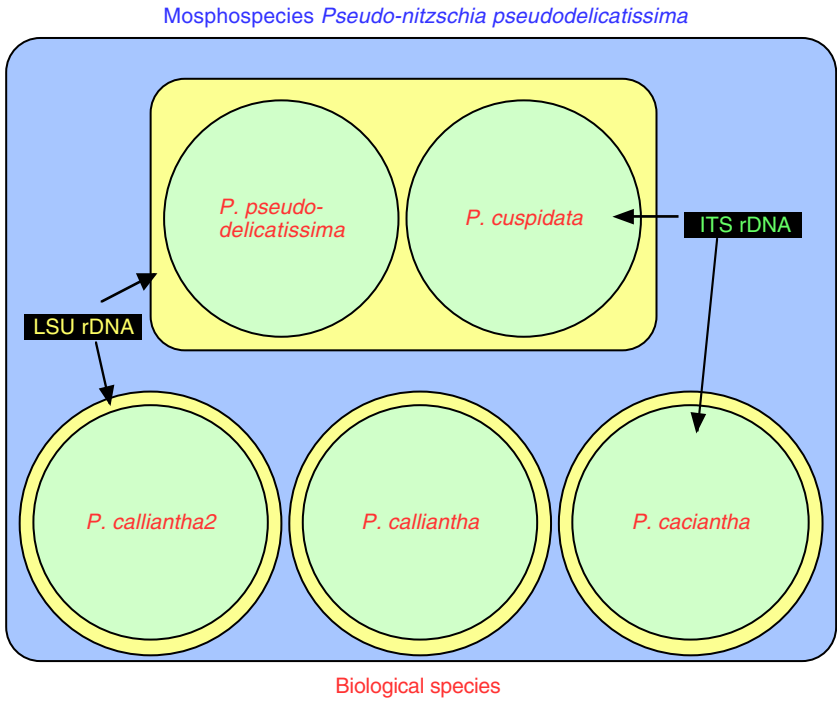


Figure 2 A comparison of different approaches to determine protist species. The blue rectangle encompasses all the clones that belong to the morphospecies *Pseudo-nitzschia pseudodelicatissima*. The use of LSU rDNA identifies four clusters within this morphospecies (yellow). The use of ITS regions further subdivides one of the clusters in two (green). These finer divisions coincide exactly with the biological species concept tested by crossing + and − clones of each groups of species (red). Only one of the two morphospecies considered by the authors is shown as an example. Reproduced from Amato A, Kooistraa WHCF, Levialdi Ghiron JH, Mann DG, Pröschold T, and Montresor M (2007) Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* 158(2): 193–207, with permission from Elsevier.

species pragmatically, therefore, we will be underestimating the diversity of microorganisms, not overestimating it. Our estimate will be conservative.

Well, then, why don't we just sequence the oceans and count the number of different rDNAs? In fact, this has been attempted several times with different techniques. The first step is to concentrate microbial cells on a filter. Despite the fact that there are almost one million bacterial cells per ml of seawater, in order for the techniques to work, we need to have a certain amount of biomass on a filter. Next, we need to extract the nucleic acids (NAs). The biochemical procedures to do this have been well established for decades. However, some cells will not break up and some others may have their NAs destroyed by the procedure. Therefore, we will have a sample that may be biased with respect to the natural environment. Nevertheless, the resulting NAs will have the DNA and the RNA of most of the microorganisms that were collected on the filter. We need to use a method to select only the NAs of interest, that is, the SSU rDNA. The usual procedure is to use the PCR. Using primers that will match conserved sections of the rDNA, this technique will generate many copies of the rDNAs in the sample. Again, the PCR may introduce some bias: some NAs, by virtue of their secondary structure, may be less efficiently amplified, others may have sequences that do not exactly match the primers, and so on. But we will have a mixture with many copies of the SSU rDNAs of many, probably most, of the members of the original assemblage without any other NAs.

Now we need to separate the molecules corresponding to each living being so that we can sequence them. Usually, this is done constructing a clone library. The amplified SSU rDNAs are mixed with a culture of *Escherichia coli* cells prepared to take up pieces of DNA. The procedure is done so that most cells will have incorporated only one molecule of DNA from the mixture. When these cells are grown in a Petri dish, each cell will reproduce producing a clone of identical cells, each one carrying one copy of the 16S rDNA of one of the organisms in the sample. Now we can easily sequence this fragment. In principle, if an organism was very abundant in the sample, there should be many clones with its sequence. And the opposite should be true for the rare ones. Thus, despite the biases mentioned above, the clone library will provide an estimate of how many different species were there and an indication of their relative abundances.

The Failure of Conventional Molecular Techniques

So, how many species are there? Using this approach, Hagström *et al.* (2002) sampled several areas of the ocean and constructed clone libraries consisting of around 600 clones each. At the time these were considered to be very large libraries. Comparing the sequences from the different places these authors came up with an estimate of a few thousand species of bacteria for the whole ocean. Unfortunately, this still does not resolve the issue, because there is a question that needs to be answered first: How many clones need to be sequenced to have all the species in a sample? Let us look at the example in **Table 1**. The table shows the results from a clone library constructed with 18S rDNA from an Antarctic sample, where a total of 67 clones were sequenced (Díez *et al.*, 2001). At the time this was a reasonable number of clones. The most abundant rDNA

Table 1 Distribution of clones in a clone library constructed from a surface sample collected in the Scotia Sea (Antarctica) in January 1998

# clones	*Clone name*	*Closest match (% similarity)*	*Group*
16	ANT37-5	*Phaeocystis antarctica* (99.8)	Prymnesiophytes
6	ANT12-7		Novel stramenopiles
6	ME1-27	*Pelagomonas calceolata* (100)	Pelagophytes
5	ANT12-6		Novel stramenopiles
5	ANT12-24		Novel stramenopiles
5	ME1-2	*Mantoniella squamata* (96.8)	Prasinophytes
3	ANT12-13		Fungi
3	ME1-3	*Ostreococcus tauri* (94.1)	Prasinophytes
2	ME1-13		Ciliates
2	ANT12-10		Novel stramenopiles
2	ME1-4	*Prymnesium patelliferum* (94.3)	Prymnesiophytes
1	ANT12-14		Cercomonads
1	ME1-11	*Strombidium purpureum* (93.3)	Ciliates
1	ANT12-3	*Pseudo-nitzschia multiseries* (98.6)	Diatoms
1	ME1-16	*Skeletonema costatum* (96.8)	Diatoms
1	ANT37-6	*Gymnodinium mikimotoi* (96.7)	Dinophytes
1	ANT12-4		Glaucocystophytes
1	ANT12-5		Novel stramenopiles
1	ANT12-9		Novel stramenopiles
1	ANT12-8		Novel stramenopiles
1	ME1-24	*Hyphochytrium catenoides* (90.2)	Novel stramenopiles
1	ANT12-11	*Hyphochytrium catenoides* (93.7)	Novel stramenopiles
1	ANT12-2	*Phaeocystis antarctica* (97.6)	Prymnesiophytes

The clones without a closest match are those where the similarity to the closest match was lower than 90%.

Source: Data from Díez B, Pedrós-Alió C, and Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Applied and Environmental Microbiology* 67(7): 2932–2941.

in the sample corresponded to *Phaeocystis antarctica* and the cloned sequence was 99.8% similar to that of the organism known from pure cultures. This was a very logical result, since *Phaeocystis* is known to be a very abundant alga in some Antarctic waters and the high similarity insured that the organism in the sample was essentially the same as the one in culture. Next, there were several organisms found at intermediate abundances (say between six and two clones each), and finally, only one clone was found for the last 12 organisms. The difference in abundance between *Phaeocystis* and these rare organisms was one order of magnitude. Now, what would happen if instead of 67 clones we had sequenced 670? Likely, we would have found 160 clones of *Phaeocystis* and 10 clones of each one of the rare organisms. No doubt we would have found many new organisms represented by only one clone. These new organisms were present in the sample from the beginning, but in concentrations two orders of magnitude lower than *Phaeocystis*. Obviously, the species richness of our library depends on how many clones we obtain and sequence. This is similar to the problem that plant and animal ecologists face with the rare species. Since nobody can count and identify all the individuals in a community, ecologists have resorted to collector curves. These curves show the increase in species as more individuals are identified. Normally, when a sufficient number of individuals have been identified the rate of increase of new species found abates and an asymptote is eventually reached. This is a clear sign that we have a proper estimate of the species richness in our community.

Unfortunately, most of the time this does not happen in the case of microorganisms. The collector curves very rarely reach an asymptote (**Figure 3**). Why don't the curves for microorganisms behave like those for animals and plants? The reason is sex. Most microorganisms do not need sex for reproduction. As mentioned, bacteria (and archaea) do not have sex, and most known protists can reproduce for long periods of time without sex. Animals and plants, however, need sex to reproduce. This means that individuals from the two sexes have to find each other. When a species becomes rare, the chances of a male and a female meeting decrease and, below certain population abundance, they will not meet and the species will become extinct. A dramatic example of this was the recent extinction of the Po'ouli (*Melamprosops phaeosoma*), a Hawaiian honeycreeper whose last individual died in the hands of conservationists in 2004. The individual had been captured as a last effort to breed the species in captivity. At that time, the wild population consisted of only three birds (a male and two females), but their territories were so far away from each other that they could not possibly meet without human intervention. Unfortunately, the captured bird died and the other two were never seen again (Powell, 2008).

Microorganisms will not have this fate. In principle, a single cell, under appropriate conditions, can reproduce and become several orders of magnitude more abundant in a matter of hours. Thus, the number of rare species in microbial communities can potentially be huge. Think of identifying all the bacteria in 1 ml of seawater. Since there are approximately one million prokaryotic cells in 1 ml of seawater, this would provide one million sequences. For sure there would be several prokaryotic species represented by only one sequence. So, let us sequence all the bacteria in 1 l of seawater. Now we would have 100 million sequences and many more species represented by only one clone. What would happen if we could sequence the whole ocean? There are 10^{29} prokaryotic cells in the ocean. To place the magnitude in context, there are only 10^{11} stars in our galaxy or neurons in our brain. How many species can there be in such a huge number of cells?

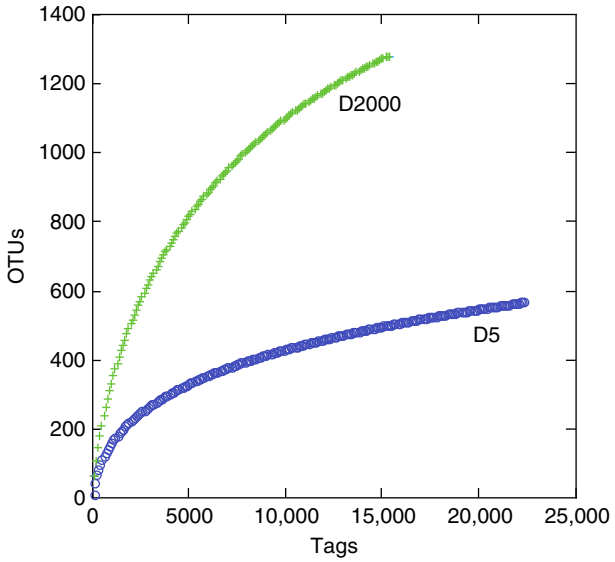


Figure 3 Collector curves for bacteria in two Northwestern Mediterranean Sea samples, coinciding with those shown in **Figure 1(b)**. Data from Pommier T, Neal PR, Gasol JM, Coll-Lladó M, Acinas SG, and Pedrós-Alió C (2010) Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquatic Microbial Ecology* 61: 221–233.

HTS

It is impractical to sequence more than 1000 clones with conventional techniques, so just to begin to cope with these dimensions requires a different approach: HTS. In the past decade several technologies have been developed that use massive parallelism to obtain hundreds of thousands of sequences simultaneously. Two of the most popular ones are 454-pyrosequencing and Illumina. So, let us see what we get when we increase the number of sequences from a sample by one or two orders of magnitude. **Figure 3** shows an example of data obtained by pyrosequencing from the Mediterranean Sea. After examination of 20,000 sequences, the collector curves were not reaching an asymptote. The number of species found was 1577 and, moreover, about 60% of these were represented by only one sequence (Pommier *et al.*, 2010). Obviously, 20,000 sequences are still not enough to saturate the collector curve. We need to increase our resolution at least another order of magnitude or, preferably, two orders of magnitude. Using this approach, Amaral-Zettler *et al.* (2010) estimated that 1 l of seawater must contain about 20,000 bacterial species. However, this cannot be extrapolated to the whole ocean, because the 20,000 species in two different liters of water may share some, many, or perhaps all species. How many species are shared is not known.

Here we need to examine the next characteristic of the rare biosphere: its "depth." Up to this point, we have established that, because most microorganisms do not need sex to reproduce, they can be extremely rare in a community and yet not become extinct. And we have seen that HTS has confirmed that the number of rare species is very high.

Is Everything Everywhere?

Given that microorganisms are very small and that they are found in very large numbers, it could be expected that they would be easily dispersed worldwide (Finlay, 2002). Thus, all microbial species would arrive at all ecosystems. Once there, they would only be able to grow and become abundant if environmental conditions were adequate. Otherwise, these species coming from far away would remain in very low concentrations, that is, they would be members of the rare biosphere. This, in fact, is what the well-known Baas-Becking expression "everything is everywhere, but the environment selects" tried to capture. For example, bacteria living in the sands of the Sahara desert are taken away by storms and deposited in the Atlantic and the Mediterranean periodically. It is unlikely that these soil bacteria will find conditions favoring their growth in the ocean and, therefore, they will remain in low concentrations. If they are not able to grow, they should eventually disappear because there are several mechanisms that kill bacteria. The question is whether the loss mechanisms are faster or slower than the colonization events. If colonization is faster than losses, all bacteria will be present in all ecosystems in the world. In this case, the rare biosphere in any ecosystem will contain absolutely all the biodiversity on Earth, only many or most of the species will be found in concentrations several orders of magnitude lower than the abundant ones. If the loss mechanisms are faster than colonization, however, the rare biosphere will be different in different parts of the ocean, and not all the species will be present in any one ecosystem.

One interesting property of the rare biosphere is that the main loss mechanisms for bacteria are either not active or reduced. The two most important loss mechanisms are viral lysis and predation by protists. Viruses must find their prey cells by chance, since they only move by Brownian motion. If the concentration of susceptible cells is below a certain threshold, the probability of a virion encountering a susceptible cell is too low to be of any significance. Thus, the species in the rare biosphere are protected from viral attack, because they are found in concentrations that are too low for their viruses to find them. In the case of predation the issue is a bit more complicated. Predators, if not selective, will cull down all bacteria in proportion to their abundance. Therefore, being rare would not be a huge advantage. However, if a bacterium is rare because it cannot grow, most predators will tend to avoid eating it in favor of bigger, more active, cells and being rare will also protect from predation at least partially. Thus, the only loss mechanisms affecting bacteria in the rare biosphere will be starvation and, if they are close to the surface of the ocean, ultraviolet radiation. This means that losses for rare bacteria are very slow and, therefore, the colonization events do not need to be very frequent for the rare biosphere to be very deep.

At the present time we cannot go any further with this reasoning. The available techniques do not allow a sequencing deep enough to see whether all the species on Earth are present in every place or not. But we can certainly state that the number of rare species in any ecosystems will be very large. The experimental evidence available cannot address this question. For example, Galand *et al.* (2009) carried out one of the so far more detailed studies examining several Arctic Ocean samples with 454-pyrosequencing. They divided the OTUs obtained from abundant (those comprising more than 1% of the total) and rare (those accounting for less that 0.01% of the total). They compared community composition among the samples by constructing dendrograms, either with the abundant or with the rare OTUs and they found very similar clustering of samples in the dendrograms. This suggested that the rare biospheres were different in every sample and the results were, thus, against the notion that everything is everywhere. However, the number of sequences examined for every sample was between 10,000 and 30,000, and the collector curve for the samples was not saturated (see their Figure S7). Therefore, the possibility remains that increasing the number of sequences by one (or two, or several…) orders of magnitude would have found all the OTUs in all the samples. The question therefore, is not resolved and, precisely because of this, we cannot say how many species of bacteria are there in the Ocean. The constant improvements in sequencing technology will allow a much better estimate in the near future.

Reasons to be Rare

One more question is why are all these organisms rare? Obviously, all organisms need to have grown in another place or at some time in the past. Otherwise, the loss mechanisms, even if reduced for rare organisms, would eventually cause their extinction. the author will discuss a few possibilities briefly. A more detail discussion of this topic can be found in Pedrós-Alió (2012).

One possibility is that the bacteria currently found in low abundance were abundant in the past but conditions changed and they began to decline in abundance. A typical example would be the seasonal cycle. Some bacteria will find better conditions in spring, for example, after a phytoplankton bloom, they will be very abundant at this time. As the bloom decays, however, conditions become unfavorable for this particular bacterium and its growth rate decreases or even stops. Both viruses and predators will cull down its numbers, until the species crosses the threshold below which it is protected from viruses and predation is reduced. In the absence of growth, this species will in effect have become part of the rare biosphere. The rank-abundance curve shown in **Figure 1** is a dynamic curve. At any one time, some of the abundant species are probably on their way down (their rank will eventually move to the right) whereas some rare bacteria that are beginning to find appropriate conditions for growth but are still rare are on their way up (moving their rank towards the left in the curve).

Evidence for this dynamic situation was presented by Jones and Lennon (2010) and Campbell *et al.* (2011). These authors

compared the ranks of bacteria obtained directly from DNA or indirectly from RNA. In the latter case, the ribosomal RNA is retro-transcribed to DNA and, then, the regular sequencing is carried out. The idea behind this approach is that very active bacteria tend to have many ribosomes, while inactive bacteria generally have very few ribosomes. The rank of a given species in the DNA-based estimate indicates the abundance of the species. The RNA-based estimate, in turn, is influenced by the abundance but it will also reflect the activity. Comparing the two estimates, the former authors could find many OTUs that had higher ranks in the RNA-based than in the DNA-based estimates. These OTUs were likely starting to grow and would eventually become abundant members of the community (and eventually increase their rank in the DNA-based estimate as well). However, many OTUs had lower ranks in the RNA-based estimate. These OTUs were likely the ones being culled down by viruses and predators and, eventually, they would find this reflected in lower DNA-based rank. This shifting of OTUs between rare and abundant is probably happening constantly to many of the members of the community. In fact, this is what gives bacterial communities the tremendous capacity to adapt to different situations, from degradation of oil spills to growth in the presence of antibiotics or heavy metals.

The other option is that the rare bacteria are actively growing and abundant at a different place in space and have been transported from elsewhere. The bacteria reaching the sea with Sahara dust storms are an example of this. Another possibility is that of bacteria forming symbiosis with marine animals. These bacteria will be extremely abundant in the specialized organs of the animal, for example, *Vibrio fischerii* is found in abundances of around 10^{11} cells in the light organ of the squid (McFall-Ngai, 2008). The animals culture the bacteria so they grow in the organs during the day, when the squid rests hidden in the sediments. At night the squid rises into the water column to feed and the light emitted by the bacteria helps to camouflage it from both predators and prey. In the morning the squid vents 90% of the cells to avoid becoming overloaded with bacteria. These vented *V. fischerii* are diluted out and transported by mixing and currents and, effectively, become part of the rare biosphere. It is essential for both the bacterium and the squid that they be part of the rare biosphere, because the young squids need to acquire their symbiotic bacteria from the water. If the bacteria were not there, even if in very low concentrations, the new generation of squid would not be able to inoculate their light organs. A similar situation must be true for all the symbiotic bacteria and, since essentially all marine animals have some kind of symbiotic microbiota, these symbiotic bacteria will account for a significant portion of the rare biosphere.

Conclusions and Perspectives

The rare biosphere has proven to be a useful conceptual framework for many different studies of diversity, taxonomy, and ecology. The continuous development in HTS technologies is already providing massive amounts of information. In the next few years, we can expect to be able to estimate, for the first time, the total number of bacterial taxa in the oceans. Aside from its basic scientific interest, knowledge of these extremely large numbers of bacterial taxa promises access to many genes with interesting potential applications.

Appendix

List of Courses

1. Microbial Ecology
2. Microbial Oceanography
3. Biodiversity
4. General Ecology
5. General Microbiology

References

Amaral-Zettler L, Artigas F, Baross J, *et al.* (2010) A global census of marine microbes. In: McIntyre AD (ed.) *The International Census of Marine Life*, pp. 223–245. Oxford: Wiley-Blackwell.

Amato A, Kooistraa WHCF, Levialdi Ghiron JH, Mann DG, Pröschold T, and Montresor M (2007) Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* 158(2): 193–207.

Campbell BJ, Yua L, Heidelberg JF, and Kirchman DL (2011) Activity of abundant and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences of United States of America* 108: 12776–12781. http://www.pnas.org/cgi/doi/10.1073/pnas.1101405108.

Díez B, Pedrós-Alió C, and Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Applied and Environmental Microbiology* 67(7): 2932–2941.

Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061–1063.

Galand PE, Casamayor EO, Kirchman DL, and Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. *Proceedings of the National Academy of Sciences of United States of America* 106: 22427–22432.

Hagström A, Pommier T, Rohwer F, *et al.* (2002) Use of 16S ribosomal DNA for delineation of marine bacterioplankton species. *Applied Environmental Microbiology* 68: 3628–3633.

Huse SM, Dethlefsen L, Huber JA, Mark Welch D, Relman DA, and Sogin ML (2010) Ironing the wrinkles in the rare biosphere. *Environmental Microbiology* 12: 1889–1898.

Jones SE and Lennon JT (2010) Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences of United States of America* 107: 5881–5886. http://dx.doi.org/10.1073/pnas.0912765107.

McFall-Ngai M (2008) Hawaiian Bobtail Squid. *Current Biology* 18: R1043–R1044.

Pedrós-Alió C (2012) The rare bacterial biosphere. *Annual Review in Marine Science* 4: 449–466. doi: 10.11446/annurev-marine-120710-100948.

Pommier T, Neal PR, Gasol JM, Coll-Lladó M, Acinas SG, and Pedrós-Alió C (2010) Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquatic Microbial Ecology* 61: 221–233.

Powell A (2008) *The Race to Save the World's Rarest Bird. The Discovery and Death of the Po'ouli. Stackpole Books*. Mechanicsburg, PA: EE.UU.

Quince C, Anders Lanzen A, Davenport RJ, and Turnbaugh PJ (2011) Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* 12: 38.

Rosselló-Mora R and Amann R (2001) The species concept for prokaryotes. *FEMS Microbiology Reviews* 25: 39–67.

Sogin ML, Morrison HG, Huber JA, *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of United States of America* 103: 12115–12120.

Venter JC, Remington K, Heidelberg JF, *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66–74.

Recombination

Abraham B Korol, University of Haifa, Haifa, Israel

© 2013 Elsevier Inc. All rights reserved.

Glossary

Chiasmata χ-Like configurations in meiosis.
Crossing-over Complex interaction of homologs within bivalents at pachytene resulting in reciprocal exchange of genetic material between nonsister chromatids.
Gene conversion A process of nonreciprocal (unidirectional) transfer of genetic information.
Genetic interference Interdependence of crossover events that includes chiasma interference (when an exchange in one segment affects the probability of exchange in an adjacent segment) and chromatid interference (which determines the probabilities of chromatid configurations in successive pairs of chiasmata, i.e., the proportions of 2-, 3-, and 4-strand double exchanges).
Synaptonemal complex (SC) A three-part proteinaceous structure that affects the number and distribution of crossovers and converts crossovers into functional chiasmata.

Types of Recombination

There are various ways to classify different types of recombination. No matter how logical a classification system may seem, the simplest way is to compare alternative aspects (modes) of the general phenomenon of recombination. These modes may be based on types of participating cells (meiotic or mitotic and male or female), cell compartments (nuclear or mitochondrial), chromosomes and DNA sequences (homologous or nonhomologous), and types of exchanges (reciprocal or nonreciprocal).

Crossing-Over and Chiasmata

Two types of cell division are known for eukaryotic organisms: mitotic and meiotic. Recombination takes place in both. In cells that enter meiosis, the chromosomes have already undergone DNA replication so that each chromosome consists of two sister chromatids connected by a shared centromere. During the meiotic prophase, pairing of homologs brings about bivalent formation; each bivalent combines four chromatids. Complex interaction of homologs within bivalents at pachytene results in reciprocal exchange of genetic material between nonsister chromatids, a process referred to as crossing-over (**Figure 1**). Observation of corresponding meiotic stages (diplotene–diakinesis) reveals characteristic χ-like configurations called chiasmata. These configurations are interpreted as cytological manifestation of crossing-over. Chiasmata were observed and interpreted as a result of an unknown process of rejoining genetic material at the end of the nineteenth century, a decade before the discovery of partial linkage between genes in plants and two decades before Morgan's studies on crossing-over on *Drosophila*.

Nonhomologous Chromosome Segregation

After crossing-over occurs, chromosomes of each bivalent segregate independently relative to chromosomes of other bivalents in two consequent meiotic divisions. These two processes, crossing-over and recombination of nonhomologs, are characteristic of all sexual eukaryotes. With independent segregation, the ratio of parental and nonparental combinations of corresponding alleles is 1 : 1, although a departure from independent assortment of unlinked genes, termed quasilinkage or pseudolinkage, was observed in both plants and animals, especially in interspecific hybrids (Korol *et al.*, 1994). Pseudolinkage is widespread in fish species, presumably reflecting the genome paleopolyploid nature (Danzmann *et al.*, 2008).

Mitotic Recombination

Exchange of genetic information can also occur during mitotic cell division, but it does so with a frequency of a few orders of magnitude lower than that of meiotic ones. The major exclusion from this rule is mitotic recombination resulting in high variation of somatic cells of the vertebrate immune system. Mitotic recombination may occur in the germline cells, for example, in *Drosophila* males where meiotic recombination is normally suppressed. Some genetic factors increase the rate of mitotic recombination manifold. For example, homozygosity for an inserted mobile element may result in high premeiotic recombination rates in *Drosophila* males comparable to that in female meiosis (Sved *et al.*, 1991). Mitotic crossing-over between the centromere and any locus results in two daughter cells homozygous for the entire segment distal to the site of exchange (**Figure 2**). This may have a strong negative impact if the initial cell was heterozygous for some detrimental alleles (*see* Applied Aspects).

Nonreciprocal Exchange

Normally, meiotic and mitotic crossovers are based on reciprocal physical exchange of genetic material between the parental chromosomes. Studies of recombination in fungi (*Sordaria*, *Neurospora*, and *Ascobolus*) have revealed a process of nonreciprocal (unidirectional) transfer of genetic information called gene conversion that accompanies reciprocal exchange (crossing-over). Gene conversion is based on DNA repair of mismatched strands along the "recombination tract." The demonstration of this process was due to the possibility of classifying marker segregation (spore color) of ordered spores in asci resulting from individual meiotic cells in hybrids between colored and wild-type strains. In addition to the expected regular 4 : 4 (colored : noncolored) octads, aberrant

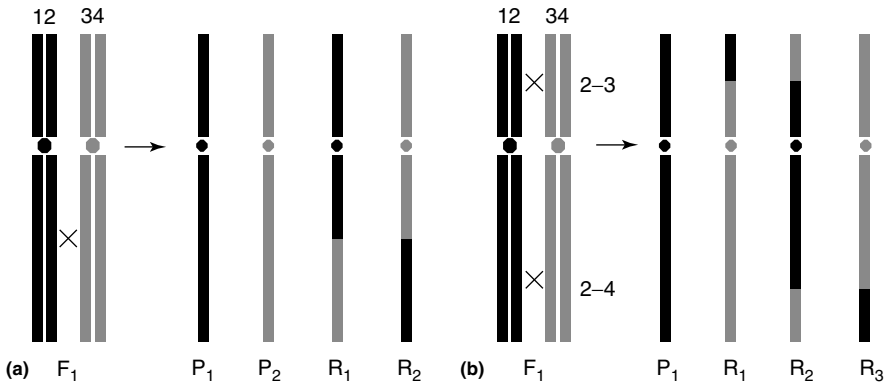


Figure 1 Crossing-over at the four-chromatid stage: (a) Single exchange resulting in two parental (P1 and P2) and two recombinant (R1 and R2) chromosomes. (b) Double exchange involving nonsister chromatids, 2 and 3 at the crossover point in the short arm and 2–4 in the long arm – the result is one nonrecombinant (P1), two single-recombinant (R1 and R2), and one double-recombinant (R2) chromosomes. Note that if the same two nonsister chromatids participate in the exchange at both positions (e.g., 1–3 and 1–3) then two parental and two double-recombinant products will appear. Meiotic configurations involving all four chromatids in the double exchange (e.g., 1–4 at the first site and 2–3 at the second) will result in four single-recombination products.

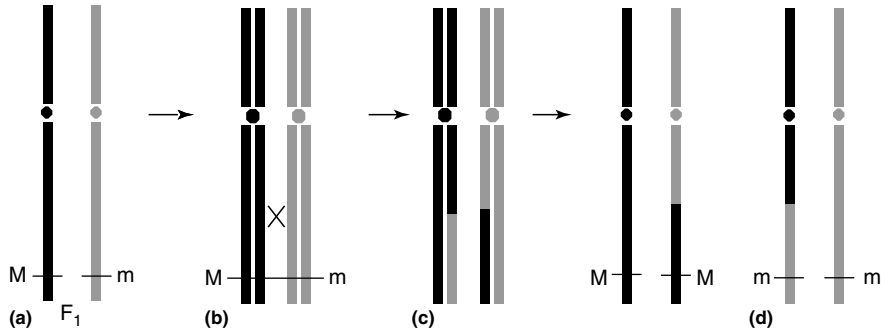


Figure 2 Segmental homozygosity resulting from mitotic crossing-over: (a) A diploid heterozygote for a marker locus (M/m); (b) after DNA duplication each homolog is represented by two sister chromatids (the position of mitotic crossing-over between nonsister chromatids is marked by $\times$); (c) the structure of the homologs after crossing-over; and (d) the daughter cells resulting from mitosis are homozygous for the entire segment distal to the point of exchange.

ratios 2 : 6, 6 : 2, 3 : 5, and 5 : 3 are also observed, reflecting nonreciprocal exchange (*see* Genetic Control and Mechanisms of Recombination). The availability of sequence data for genes of several higher eukaryotic organisms (wild barley, maize, and *Drosophila*) allowed estimating population genetic parameters characterizing the rate of reciprocal (crossing-over) and nonreciprocal (gene conversion) recombination events, as well as the contribution of recombination and mutation to sequence variation at the population level. Both in outcrossing and self-fertilizing species, the contribution of gene conversion to total recombination may be twice as much as compared to crossing-over. Recombination, in total, contributes to sequence diversity no less than the mutation process (Morrell *et al.*, 2006).

Ectopic Recombination

Crossing-over occurring between homologous sequences of homologous chromosomes generates new genetic variation without changing genome structure. However, crossing-over may also involve homologous sequences of nonhomologous parts of the genome. This process is called ectopic recombination. In particular, numerous families of repeated elements that are spread in eukaryotic genomes may participate in ectopic recombination. Retrotransposons can be especially active, either from nonhomologous chromosomes or from different sites of the same chromosome (giving rise to inter- and intrachromosomal ectopic recombination), resulting in chromosome rearrangements. In fact, the frequency of such events is very low, and it is unclear which mechanisms reduce this danger. One possibility may be a strong bias of recombination interaction in favor of gene conversion if such ectopic contacts occur. Ectopic recombination may also occur between such partners as chromosome and a plasmid, such as a 2-μm plasmid in yeast harboring the gene for site-specific recombinase Flp and the tumor-inducing T-DNA segment of the Ti plasmid of *Agrobacterium tumefacience* interacting with the host plant chromosome.

Illegitimate Recombination and Horizontal Transfer

Unlike homologous and site-specific ectopic recombination, illegitimate recombination is a process of joining DNA

molecules that have only a small sequence homology. Illegitimate recombination is common in nature, especially in interactions between retroelements and host genome DNA and in some host–parasite systems. One of the most common examples is the previously mentioned interaction between bacteria, such as *A. tumefasciens*, and their host plants mediated by the bacterial plasmid transformation (*see* Ectopic Recombination). Although this process can be considered as a "physiological" rather than a genetic transfer, it may also have evolutionary consequences as one of the mediators of horizontal gene transfer. Illegitimate recombination is involved in many situations of foreign DNA integration in direct transformation experiments aimed at obtaining genetically engineered (transgenic) organisms.

Extranuclear Recombination, Exchanges between Cell Compartments

Recombination was proven to affect the organization of mitochondrial DNA. Mitochondrial recombination was detected in *Drosophila*, mussels, and various vertebrate species including fishes, birds, and mammals (Tsaousis *et al.*, 2005). Convincing evidence of widespread recombination in mitochondrial genomes has been obtained for plants. In fact, a specific substrate for frequent mitochondrial recombination was discovered and sequenced in a few plant species. Especially interesting from the evolutionary point of view are putative DNA exchanges between different cell compartments: nuclear–mitochondrial, nuclear–chloroplast, and mitochondrial–chloroplast. Many such sequences demonstrating "chloroplast → mitochondrial" transfer were obtained for cereals. The existence of mitochondrial recombination limits the validity of evolutionary reconstructions based on a certain part rather than on the entire mitochondrial genome.

Basic Features of Reciprocal Homologous Recombination (Crossing-Over)

Recombination Rate and Map Distance

The average number of crossovers in a given segment represents its genetic length. Correspondingly, genetic distance between two loci can be defined as the average number of recombination events occurring in the segment. Therefore, for loci flanking a unit length segment one exchange per meiosis is expected on average, resulting in 50% recombinant gametes. This notion of genetic distance serves as a basis for gene mapping. Two fundamental facts make genetic mapping possible: (1) linear organization of genetic material in chromosomes and (2) increasing recombination rate between loci with the increase in their physical distance. Map distance is of additive nature: For any subdivision of a segment, the segment length is the sum of lengths of its subintervals. A limitation of map distance as a measure is that it cannot be estimated directly: an odd number of exchanges in a segment marked by two loci leads to the appearance, in the progeny, of the same recombinant genotypes as those resulting from one exchange, whereas an even number results in no observable recombinants. Recombination frequencies (*rf*s) estimated from the observed proportions of recombinant and nonrecombinant individuals coincide with map distances only for short segments when multiple exchanges can be ignored. Therefore, transformation of experimentally estimated *rf* values into map distances (x) is needed. Relations of the form $rf = rf(x)$ are referred to as mapping functions.

Genetic Interference and Mapping Functions

Multiple exchanges are relatively rare among eukaryotes. These vary between one and three for the majority of higher organisms, with five or six rarely occurring, although values up to 12 have been reported (Korol *et al.*, 1994). In *Drosophila*, the frequency of tetrads with four or more crossovers is extremely low (usually much less than 1%).

Interference: Its Measurement and Basic Properties

Consider two adjacent marked segments m1–m2 and m2–m3 with recombination rates rf_1 and rf_2. If exchanges in adjacent segments occur independently, the probability of a double exchange within m1–m3 can be determined as $rf_{12} = rf_1 rf_2$. Under the same assumption, recombination frequency rf_3 between loci m1 and m3 is given by

$$rf_3 = rf_1(1 - rf_2) + rf_2(1 - rf_1) = rf_1 + rf_2 - 2rf_{12}$$

Much evidence indicates that the observed frequency of double crossovers usually differs from the expected one, a phenomenon termed as genetic interference. The degree of interference is measured by the coefficient of coincidence (c):

$$c = \frac{rf_{12}(\text{observed})}{rf_{12}(\text{expected})} = \frac{rf_{12}(\text{observed})}{rf_1 rf_2}$$

The existence of interference was shown both genetically and cytologically. It is ubiquitous in eukaryotes. To date, only in a few species (*Shizosaccharomyces pombe*, *Aspergillus nidulans*, *Tetrahymena thermophila*, and possibly *Ustilago maydis*) has recombination been found to proceed with no interference (Loidl and Scherthan, 2004, and references therein). If the occurrence of crossing-over in one segment reduces the probability of exchange in an adjacent segment ($c < 1$), then it is positive interference; when one exchange increases the frequency of another ($c > 1$), it is negative interference. The nematode *Caenorhabditis elegans* is an example of extreme positive interference: a single crossover formed per chromosome suppresses formation of additional exchanges.

Two types of interference have been distinguished: chiasma interference, in which an exchange occurring in a segment affects the probability of another one occurring in an adjacent segment, and chromatid interference, which determines the pattern of chromatid configurations in successive pairs of chiasmata in a bivalent (i.e., the proportions of two-, three-, and four-strand double exchanges) (**Figure 1**). In the absence of chromatid interference, the ratio should be 1 : 2 : 1. Evidence of chromatid interference is scarce. Chiasma interference determines the pattern of distribution between successive exchanges along a chromosome. The amount of interference within a segment depends on its distance from the centromere. Near the centromere, interference between exchanges in the adjacent segments of the same chromosome

arm is maximal. It is generally believed that no interference exists across the centromere, although many exclusions are known.

Mapping Functions

Positive interference ($c<1$) is a widespread phenomenon. It was found that in many cases the approximation $c(rf)=2rf$ (Kosambi interference) fits the data well (i.e., $c\approx 0$ at small distances and $c\approx 1$ at large distances). The corresponding mapping function

$$rf = 0.5\tanh(2x) \quad \text{or} \quad x = 0.25\ln((1+2rf)/(1-2rf))$$

appeared to be a good approximation for *Drosophila*, rice, mouse, and other eukaryotes. The formula for *rf* values in adjacent intervals, corresponding to Kosambi's function, takes the following form:

$$rf_3 = \frac{rf_1 + rf_2}{1 + 4rf_1rf_2}$$

Situations with no interference can be represented by the mapping function proposed by J. Haldane:

$$rf(x) = 0.5\left(1 + e^{-2x}\right)$$

Some results indicate that negative interference (i.e., an excess of double exchanges over the level expected with the assumption of independence), may also occur. Such data were observed for *Drosophila* and some plants, fungi, and viruses. One should take into account that values of $c>1$ can also result from marker scoring errors and data pooling, owing to heterogeneity of individual *rf* values.

Distribution of Crossovers within Eukaryotic Chromosomes

The efficiency of crossing-over strongly depends on between- and within-chromosome patterns of exchange distribution in the nucleus. Here, two different (albeit related) problems are generally distinguished: (1) crossover distribution in relation to major components of the chromosome, that is, the centromere, telomeres, heterochromatin (intensively stained cytological segments of chromosomes – gene-poor regions), and euchromatin (faintly stained gene-rich regions), and (2) relative positions of crossovers with respect to one another. These questions were thoroughly studied based on scoring recombination events between multiple markers and complemented by cytological observations. A wide range of spatial variation in exchange frequency was revealed – from long stretches of heterochromatin in which crossing-over is almost blocked to short segments containing recombination hotspots (*see* Recombination Signals, Hotspots of Recombination, and Transcription). Heterochromatic blocks affect the distribution of crossovers in the entire genome. Recombination non uniformity of the chromosomes was first established in *Drosophila* and then confirmed in a wide range of organisms. Thus, in the B genome of wheat the physical length of DNA per exchange varies more than 150-fold along the chromosome (Lukaszewski and Curtis, 1993). In the centromeric region of the human X chromosome, the recombination rate is at least eight-fold lower than the average rate of female recombination on the X chromosome.

The Effects of Sex

By the Haldane–Huxley rule, if crossing-over is strongly suppressed or lacking in meiosis of one of the sexes, then that sex is invariably the heterogametic one. The effect of sex on *rf* has been established in a variety of animal species, with this rule holding in all cases of alternative (all-or-none type) sex differences. Provided crossing-over occurs normally during the formation of both male and female gametes, then *rf* in the heterogametic sex may be either lower or higher than that in the homogametic one, with the sign and magnitude of sex difference in *rf* being segment specific. Sex differences in recombination have also been revealed by cytological analysis. Higher female recombination was observed in many species. In humans, the total length of female maps is nearly twice as long as that from male maps; in some regions (mainly subtelomeric ones), the opposite tendency is observed. Most of the examined genomic regions in maize exhibit higher male recombination or no difference; the extent of the difference (if it exists) may strongly depend on the genotype. A substantially higher crossing-over rate in male compared to female meiosis has been found in *Arabidopsis* (Korol *et al.*, 1994). This is also true for some animals exhibiting a higher degree of linkage and/or more localized pattern of exchange distribution in the homogametic sex than in the heterogametic one, such as marsupials (Trivers, 1988).

The mechanisms responsible for differences in *rf* between male and female meiosis are unknown. In most animals, sex is determined by a certain chromosome. Thus, it can be suggested that the difference between male and female meiosis is controlled by genes located on this chromosome. Alternatively, sex chromosome can canalize the development of general physiological (biochemical) sex differences (e.g., in the hormonal status) in a toggle-like manner, with specific recombination features of male and female meiosis being a secondary effect of the main regulator. The latter suggestion is supported by data from the fish *Oryzias latipes* in which the heterogametic sex is represented by XY males (Yamamoto, 1961). When males XY were transformed into functional XY females by a hormonal treatment, *rf* increased by a factor of 5 compared to the level in normal males. In humans, it was found that the distribution of sex difference in *rf* is related to genomic imprinting manifested in region-specific differential modification (methylation) of parental alleles. We speculated that gametic selection may be an important contributing factor to the evolution of sex differences in recombination (Korol *et al.*, 1994). Lenormand and Dutheil (2005) supported this hypothesis by a detailed analysis of available evidence on many species.

The Effects of Environmental Conditions

If diversity is essential to population survival, it would be highly adaptive for organisms to increase genetic variation in challenging (stressful) environments and decrease it under optimal conditions (Korol *et al.*, 1994). Such regulation should increase the chances of population survival without dissipation of the accumulated hidden genetic variation in favorable times. An abundance of evidence on environmental modulation was found for the main sources of variation:

mutation, chromosomal rearrangements, transposition of mobile elements, and recombination. The environment can affect recombination in two ways. Firstly, it can affect recombination indirectly when changes in the frequency of alleles at recombination controlling loci depend, through a feedback mechanism, on changes in the frequency of the selected (fitness-related) loci (*see* Evolution of Recombination). Secondly, it can affect recombination directly when in response to environmental factors, the organism modifies (within genetically determined limits) the value of *rf*, that is, when not *rf* but the norm of reaction with respect to *rf* is genetically determined. The fact of direct influence of the environment on recombination has long been known in genetics. As early as 1917, H Plough observed in *Drosophila* a U-shaped dependence of *rf* on temperature: Recombination was minimal at optimum temperature and increased with the departure of temperature from the optimum. Similar data indicating an increase in recombination level with deteriorating environment have been reported in many works. Extreme conditions also tend to reduce interference.

Ontogenetic Memory and Transgeneration Effect

A common belief is that environmental factors may affect recombination only when the treatment occurs during a relatively short period between premeiotic S-period and prophase of the first meiotic division. Therefore, in many organisms the sensitive stage may comprise a very small part of the life cycle. Does it mean that environmental stresses on earlier stages of the development cannot affect recombination? Some limited evidence available for a few organisms indicate that just the opposite may be true. Neel (1941) studied the effect of severe starvation of *Drosophila* larvae on meiotic recombination. An increased rate of recombination in chromosome 3 was observed during the whole lifetime of the females, representing some memory of starvation-induced changes in germline cells during many mitotic cell divisions until meiosis. Zhuchenko *et al.* (1983) showed that temperature stress applied to *Drosophila* F1 hybrids at the embryonic stages can affect the rate of meiotic recombination during the entire lifetime of F1 females, implying that stress-induced changes survived many divisions of germline cells until meiosis. In *S. pombe*, X-irradiation caused an increase in the recombination rate which persisted for 8–10 cell generations after irradiation (Takeda *et al.*, 2008). Limited data indicate that the memory of stress-induced changes may extend over generations. In experiments of Landner (1970) with *Neurospora*, temperature treatments applied as early as 3–4 days before fertilization, affected the rate of meiotic recombination, that is, in the next generation. A similar transgeneration effect was also found in the aforementioned study of Zhuchenko *et al.* (1983) with *Drosophila*: changes in meiotic recombination rate were recorded in F1 females whose wild-type mothers were subjected to heat treatment of 36 °C for 12 h.

Somatic recombination is considered a general plant response to environmental stresses like UV, high temperature, herbicides and other factors, but the usual expectation is that stress-induced stimulation should disappear together with the inducing factor (Pecinka *et al.*, 2009). Experiments with *Arabidopsis* have shown that this may not be the case: treatment with UV light or with a chemical mimicking pathogen attack caused increased somatic recombination in leaves of the treated plants and of several subsequent generations (Molinier *et al.*, 2006). These findings were interpreted by the authors as a manifestation of the epigenetic effect. Using ten different physical and chemical stress treatments, Pecinka *et al.* (2009) tried to repeat these results on several lines of *Arabidopsis* by scoring for somatic recombination in the treated (S0) and subsequent generations of selfing (S1 and S2). The results were negative: despite highly significant induction observed in S0, only small and nonconsistent effects were observed in S1 and S2, suggesting that the transgeneration effect in *Arabidopsis* is not a rule and occurs only under very special conditions. Obviously, a similar question of reproducibility should also be addressed for stress memory related to meiotic effects.

Feedback Regulation of Genetic Variation

If stress may be recombinogenic, it is natural to assume that the degree of the effect should depend on the intensity of the stress, that is, the degree to which an individual is stressed. In other words, induced changes in recombination in maladapted genotypes should be higher than in well adapted ones (Zhuchenko and Korol, 1983). Experimental evidence for the last principle has been obtained in the author's lab for *Drosophila* and tomato (reviewed by Korol *et al.*, 1994). In *Drosophila*, the increase in male recombination induced by heat treatment was negatively correlated with flies' resistance to high temperatures. In tests with tomato varieties and their F1 hybrids, low temperature did not increase the chiasma frequency in cold-resistant varieties and in their F1 hybrids, whereas the reverse was true for cold-sensitive genotypes. By contrast, elevated temperatures caused an increase in recombination only in heat-sensitive genotypes. Populations may benefit from such a feedback mechanism by generating additional variation to meet new, challenging conditions while preserving previously evolved adaptive gene combinations, when there is no need for additional variation, thereby reducing the recombination load. In the last decade, the concept of plastic- and fitness-associated recombination has become the subject of theoretical studies (Hadany and Beker, 2003). Theoretical models and growing evidence indicate that similar strategies may also be relevant to other systems affecting genetic variation: transposition of mobile genetic elements, DNA repair, mutation rate, and sexual versus asexual reproduction (e.g., Lupu *et al.*, 2006; Agrawal and Wang, 2008; Schoustra *et al.*, 2010).

Genetic Control and Mechanisms of Recombination

Selection for Changed Recombination, Genetic Modifiers of Recombination

Selection for altered recombination proved an important tool in analyzing genetic control of recombination. Such experiments have been conducted on dozens of organisms, including *Drosophila*, flour beetle *Tribolium castaneum*, grasshopper *Schistocerca gregaria*, silkworm *Bombix mori*, lima bean, and fungi *Neurospora* and *Schizophyllum* (Korol *et al.*, 1994). These experiments complemented the numerous findings on the existence of a considerable amount of genetic variation in

rf in natural and laboratory populations (*see* Intraspecific Variation for the Rate of Recombination) available for selection to act on. The effectiveness of directional selection for altered *rf* has been shown to depend on the segment under study, the size, structure, and origin of the start population; the breeding system; the estimation procedure; and the intensity and duration of selection. Genetic analysis of the accumulated differences between selected lines enabled the question of the genetic basis of variation in recombination to be addressed. In as few as 10 generations, divergent selection for *rf* in the p–Y region of silkworm chromosome 2 succeeded to produce lines with *rf*=37–39% and 5–7%, starting from 25.6%. The obtained evidence suggests that recombination (frequency and distribution) in eukaryotic genomes is under complex control of polymorphic modifiers with small to moderate effect. The rate of recombination in a given region may depend on both linked modifiers and genes located on other chromosomes.

Meiotic Mutants as an Analytical Tool in Recombination Studies: Overlapping of DNA Repair, Recombination, and Segregation Systems

The first discovered point mutation (in 1922) affecting recombination was *Drosophila* gene c(3)G, which almost completely blocks crossing-over in females. During subsequent years, many other genes with strong effect on different meiotic steps and recombination were isolated in eukaryotes (mainly in fungi, *Drosophila*, maize, and recently *Arabidopsis* and the nematode *Caenorhabditis elegans*). The availability of diverse mutants affecting the coordinated steps of meiosis and recombination allowed a detailed cytogenetic characterization of meiosis (reviewed by Heyer *et al.*, 2010). These include commitment to meiosis, regulation of major steps in chromosome conjugation and formation of the synaptonemal complex (SC) (*see* Ultrastructure of Recombination: SC and Recombination Nodules (RN)), DNA breaks formation and strand exchange, chiasma maintenance, and chromosome disjunction. The broad network of recombination–meiosis mutants identified in yeast and other fungi, *Drosophila*, *C. elegans*, and *Arabidopsis* serves as a basis for cloning recombination genes and studying the molecular mechanisms of recombination. Numerous results were obtained on this rich material showing a deep overlapping of pathways related to DNA recombination and DNA repair. Thus, mutations for loci controlling meiosis and recombination also tend to exhibit disturbed repair functions. Such a duality seems to be an ancient phenomenon. In prokaryotes, the best example is the RecA protein of *Escherichia coli* that is involved in early stages of recombination and, simultaneously, in induced SOS repair. Recombination is also strongly affected by mutations in genes controlling other enzymatic steps in DNA metabolism, for example, replication. Genes for DNA repair and recombination exhibit a high degree of sequence, structural, or functional similarity throughout life. This conservation was established using a combination of genetic and cytological methods with molecular cloning, sequencing, and genetic transformation. Thus, defects in the recombination–repair system of one species may be complemented by a gene transferred from another species, sometimes a distant one (e.g., from a prokaryote to eukaryote). However, this does not exclude high interspecific differences in many important features of recombination (Korol *et al.*, 1994).

Comparison of recombination in meiotic mutants and in corresponding wild-type genotypes provides an interesting conclusion. It appears that genomic distribution of exchanges in normal meiosis is more restricted, less random, than in meiosis altered by mutations, indicating that these restrictions are largely a result of evolutionary adjustment of recombination. Meiotic mutants tend to lessen the restrictions, and regions normally excluded from exchanges become involved in crossing-over. Simultaneously, a reduction of interference is also observed despite a general tightening of linkage (Bhagat *et al.*, 2004). In meiotic mutants, the correspondence between physical and recombination length of chromosome segments becomes closer. In many cases, mei mutations referred to as hyper-rec can significantly enhance crossing-over (and not only in specific regions) or increase the rate of multiple exchanges, intragenic recombination, and/or mitotic crossing-over manyfold.

The meiotic system manifests a close connection between homologous conjugation, crossing-over, and regulation of chromosome segregation. This association is invariable for normal meiosis and served as a basis for the view that regulation of chromosome segregation is one of the main functions of crossing-over. Meiotic mutants were isolated where the two phenomena, segregation and recombination, are not coupled. For instance, crossing-over is normally almost completely suppressed in chromosome 4 of *Drosophila melanogaster*. Some mei-mutants with reduced recombination and/or altered exchange distribution within large chromosomes exhibit crossing-over within chromosome 4 and disturbed segregation of this chromosome (Sandler and Szauter, 1978).

Ultrastructure of Recombination: SC and RNs

Electron microscopy reveals meiosis-specific organization of chromosome associations at the meiotic prophase I. A three-part proteinaceous structure, called SC, is initiated at the zygotene stage, maintains the paired homologs together during pachytene, and disappears at the diplotene. Short stretches of DNA of the homologs are involved in the central part of SC where intimate molecular pairing takes place. SCs are considered as structures that affect the number and distribution of crossovers and convert crossovers into functional chiasmata. In many species (e.g., human, mouse, zebrafish, and Arabidopsis), the difference in total lengths of SCs in male and female meiocytes is correlated with the difference between male and female recombination map lengths. Small electron-dense bodies, RNs, were detected within the central part of the SC. Early RNs that appear before pachytene are more or less evenly distributed along the bivalent and are supposed to participate in synapsis. Late, or pachytene, RNs seem to represent recombination enzymes associated with exchanging DNA molecules. The distribution of late RNs corresponds well to chiasma distribution and is believed to represent the sites for recombination. Numerous meiotic mutants were tested with respect to their ultrastructural and molecular phenotype – the morphology and biochemistry of SC. An interesting fact

is that only in a few organisms was meiotic recombination normally found without SC (*S. pombe*, *A. nidulans*, *T. thermophila*, and possibly *U. maydis*). However, in many cases SC was found in achiamatic meiosis (with no recombination). In *S. pombe* and *A. nidulans* with recombination occurring without SC, the frequency of exchanges per bivalent is high, and especially interesting is the absence of interference. Thus, the absence of SC may allow for more reshuffling of the homologs. Analysis of SC appeared to be an important tool in determining the causes of altered fertility in farm mammals.

Molecular Mechanisms of Recombination

When considering molecular mechanisms of recombination in eukaryotes, a broader spectrum of basic processes involved may also be addressed, including chromosome conjugation and segregation. This is partly due to the prevailing experimental methodology based on analyzing cytogenetic and molecular effects of numerous mutants defective for meiosis–recombination and DNA repair. Historically, the studies of molecular mechanisms of recombination in prokaryotes served as the main source of models to be tested in eukaryotes. The basic enzymatic activities involved in prokaryote recombination include those of endo- and exonucleases, strand transfer, DNA unwinding, ligation, and resolving Holliday's heteroduplex structure. Eukaryotic analogs were found and (are to be) characterized in the best-studied eukaryotes, such as yeast, *Drosophila*, Arabidopsis, and mammals. Clearly, eukaryotic recombination systems are supposed to be more complex due to additional problems related to meiosis, such as homologous pairing and segregation.

The classical concept of molecular recombination included the following steps: (1) intimate pairing of homologs; (2) formation of single-stand DNA nicks at homologous sites and local denaturation of double-stranded DNA; (3) formation of recombination intermediate (Holliday junction) by reciprocal strand transfer and renaturation, its migration from the initiation site, and heteroduplex formation (with local mismatched DNA due to heterozygosity of the parental chromosomes); and (4) resolution of the Holliday structure, either after or before isomerization. In the first case, the content from both sides of the considered regions remains unchanged, whereas in the second case it results in reciprocal exchange. In both cases, the mismatched positions should undergo correction. Depending on the strand used as a template in the repair process, two alternatives will be observed: recovering the initial state and nonreciprocal information transfer (gene conversion).

A new understanding of eukaryotic recombination mechanisms resulted from the discovery of the leading role of double-strand DNA breaks (DSBs) in yeast (Szostak *et al.*, 1983) that became the leading concept (Szostak model) of crossover formation for more than two decades. Details on different pathways related to this process were also studied in mammalian cells, *Drosophila*, and plants based mainly on recombination–repair mutants. Another strategy is to study specially engineered genetic models with a modified recombination system (Shalev *et al.*, 1999). Homology of recombination genes in one species to genes controlling DSB-dependent recombination in yeast may be indicative for a shared DSB mechanism of recombination initiation, as exemplified by the *Drosophila* mei-W68 gene and yeast Spo11 (McKim and Hayashi-Hagihara, 1998). However, new evidence obtained on different organisms points to limitations of the Szostak model that focused only on the crossover-formation pathway of recombination, whereas new models take into account the noncrossover pathway (potentially leading to gene conversion) and explain details of DSB initiation in both pathways and their interference-sensitive and -insensitive classes (reviewed by Cromie and Smith, 2007; Heyer *et al.*, 2010). In particular, recent studies on the fission yeast *S. pombe* point to the existence of a novel mechanism for the regulation of crossover distribution along the chromosome – controlled repair of DSBs based on partner choice model (differential interaction with the sister chromatid or with the homologous chromosome) (Hyppa and Smith, 2010). This mechanism explains the rather high stability of the crossover rate per DNA physical unit length despite high variation of the DSB rate along the genome. The molecular mechanisms of DSB repair involved in meiotic and mitotic recombination differ in many aspects (for a recent review see Andersen and Sekelsky, 2010).

Recombination Signals, Hotspots of Recombination, and Transcription

Although recombination occurs all over the genome, the accumulated data points to the existence of hotspots and cold spots of recombination. In certain cases, highly specific interactions were found between recombination–repair mutations and "signal" DNA sequences resulting in a sharp increase in *rf* within defined genomic segments. The best-known recombination signal is the octamere 5′GCTGGTGG-3′ of *E. coli* referred to as χ element. Heptamer 5′ATGACGT-3′A is known as a meiotic recombination hotspot of fission yeast *S. pombe*. A few recombination hotspots have been characterized in mammals, with the major histocompatibility complex being one of the best-known examples. In humans, 8-oxoguanine, resulting from endogenous oxidative stress, is a hotspot signal spread over thousands of positions in the genome. In humans and *Drosophila*, a 13-bp sequence motif CCNCCNTNNCCNC is associated with hotspots of fine-scale recombination (Stevison and Noor, 2010). Recent studies on humans and mice showed that a major player for hotspot specification is PRDM9 gene (Baudat *et al.*, 2010). PRDM9 is predicted to recognize the mentioned 13-mer motif. Allelic variants of PRDM9 zinc finger region are significantly associated with variability in genome-wide hotspot usage among humans, indicating that the binding of PRDM9 protein to specific DNA sequences controls the initiation of recombination at specific locations in the genome. The minisatellite structure of the PRDM9 zinc finger region points to its potential to increase variability leading to new PRDM9 variants with new DNA binding specificities, hence new genome-wide distributed hotspots.

In eukaryotes, mobile elements may act as recombination enhancers. Thus, with identical locations on homologous chromosomes, a pair of P elements of *D. melanogaster* can induce recombination at a frequency of 20% in males in

which recombination is normally absent (Sved *et al.*, 1991). This effect serves as a basis for an efficient method of fine genetic mapping, in which P elements are used to promote recombination at their insertion sites. However, usually *Drosophila* genome regions with increased density of transposable elements display reduced recombination. In general, both positive and negative correlations between the abundance of transposable elements and the recombination rate were found in eukaryotic organisms. The accumulating evidence indicates that recombination hotspots are associated with DSB sites, depend on chromatin organization, and are influenced by epigenetic changes.

In the 1960s and 1970s, it was speculated that initiation of recombination is attached to transcription promoters (Whitehouse, 1966). In moving from lower to higher organisms, the recombination density over the genome is reduced by three or four orders of magnitude with no marked trend in structural genes. Thuriaux (1977) attributed this paradox to preferential localization of recombination in structural genes. It was suggested that the requirement for common specific changes in chromatin organization may cause an overlapping in control of transcription, recombination, repair, and mutation (Whitehouse, 1966; Thuriaux, 1977; Korol *et al.*, 1994; Laurencon *et al.*, 1997). The evidence from many eukaryotes indicates that regions with high *rf* correspond to gene-rich regions (Lichten and Goldman, 1995), although an opposite trend was found in *C. elegans*. The results obtained during the past decades for yeast (*Saccharomyces cerevisiae* and *S. pombe*) support the idea of recombination–transcription relationships, but also show their complicated nature. In particular, genes near hotspots display a higher expression level, but not in meiosis when their expression is lower. It appeared that upstream gene sequences acting as transcriptional control sites may also play a role of inducible recombination hotspots. This implies that environmental effects on recombination reflect partially overlapping environmental regulation of transcription and recombination (Koren *et al.*, 2002). In humans, increased meiotic recombination coincides with high GC and gene-rich regions, but recombination rate within genes is lower compared to their upstream and downstream regions (Coop *et al.*, 2008). Surprisingly, a negative correlation was found between crossover rate and gene expression in meiosis (McVicker and Green, 2010).

In general, the available information shows high intragenomic heterogeneity and nonrandomness of recombination. These are caused by a hierarchical control system that involves interaction of a complex web of recombination–repair genes with peculiarities of macro- and micro-organization of the genetic material (**Table 1**). The important role of gene distribution in this regulation indicates its adaptive role in the evolution of genetic systems.

Biodiversity of Recombination Systems

Evolutionary Trends in Genomewise Recombination Rate

A considerable (by several orders of magnitude) decrease in *rf* per unit physical length of DNA is observed when moving from prokaryotes and lower eukaryotes to higher organisms (Thuriaux, 1977). Even among eukaryotes, the density of recombination events within the genome varies widely (**Table 2**). The genetic length of the genome varies less and strongly depends on haploid chromosome number. A relatively low correlation is displayed by fine-scale recombination rates in genomic regions with parallel syntenic blocks of different species pointing to fast evolution of recombination hot spots. This is reflected in human-chimpanzee comparisons and even among different human populations. But at larger scales, including the genome level, recombination rates are

Table 1 Hierarchical control of recombination events

System effects:
- Control of the nucleus as a whole, manifested in the classical *interchromosomal effect* of rearrangements, as well as effects of eu- and aneyploidy and supernumerary (*B*) chromosomes
- Effect of chromosome size
- Regulation based on cytonuclear interactions
- Sex differences in recombination frequency and distribution
- Effect of environmental conditions (stress) and age
- Position of euchromatin (as main target for recombination) relative to the centromere, telomeres, and heterochromatic blocks

DNA sequence organization of the target region:
- Heterozygosity for micro- and macroinversions, deletions, and translocations
- Microsite organization of DNA sequences including distribution of specific regulatory sequences like recombination hotspots, micro-, and minisatellites
- Distribution of epigenetically modified (e.g., methylated) DNA stretches
- Effects of mobile genetic elements
- Distribution of gene-rich islands

Genes controlling DNA metabolism:
- Genome-, chromosome-, and segment-specific effects of major *rec*-genes of the *coarse control* system affecting the basic steps in recombination mechanics (pairing, DNA strand exchanges, recombination repair)
- Segment-specific regulation of crossover frequency by genes of the *fine control* system with relatively small effects of individual components
- Genes encoding active transposases

Interaction of factors both within and between the foregoing groups plays an important role in observed patterns of recombination frequency and distribution within genome and its variation within and between species.

Table 2 Recombination characteristics of some eukaryotic genomes

Species	*Genome length (cM)*	*Haploid number (n)*	*Chromosome length (cM)*	*kb:cM ratio*
Magnaporthe grisea	600	7	85	80
Pinus taeda	1700	12	140	12,000
Arabidopsis thaliana	600	5	120	230
Brassica rapa	1900	10	190	280
Fhaseolus vulgaris	1200	11	110	530
Hordeum spontaneum	1200	7	170	4000
Lycopersicum esculentum	1400	12	120	900
Drosophila melanogaster	300	4	75	570
Bombix mori	3400	28	120	130
Tribolium castaneum	600	10	60	350
Apis mellifera	3700	16	230	45
Danio rerio	2900	25	120	590
Gallus gallus	3200	29	110	310
Monodelphis domestica	650	9	70	5400
Mus musculus	1400	20	70	1800
Canis familiaris	2100	39	54	1200
Homo sapiens	3600	23	160	800

well correlated. Higher "regularity" of recombination due to frequent outcrossing, the obligatory occurrence of at least one exchange per bivalent, and increasing genome size are considered to determine the evolutionary importance of recombination in higher sexual eukaryotes. In lower organisms, recombination was supposed to play a less important role due to a slight chance of sexual contacts between individuals (Maynard Smith *et al.*, 1991). Consequently, natural populations of such species should manifest a clonal structure, with strong correlation (linkage disequilibrium (LD)) between alleles at different loci. However, this may not be the case in some protozoan species, such as malaria *Plasmodium falciparum* (Conway *et al.*, 1999). The revealed population pattern of *P. falciparum* includes high polymorphism and a rapid decay of LD with distance (reaching independence between polymorphic sites of 0.3–1.0 kb) that looks like that of higher sexually reproducing eukaryotes.

Intraspecific Variation for the Rate of Recombination

The analysis of genetic variation of recombination parameters allows a deeper insight into patterns, mechanisms, and the evolutionary role of recombination. The genetic differences in crossing-over rate seem to have been originally established by Sturtevant in 1913 who discovered modifiers sharply reducing crossing-over during oogenesis in heterozygous females of *Drosophila*. Later, he suggested these to be inversions, a prediction that was subsequently confirmed by cytological analysis of salivary gland chromosomes. Among point mutations affecting recombination, the first to be discovered was c(3)G, which almost completely blocks crossing-over in *Drosophila* females. Additional studies demonstrated high genetic variation in *rf* in *D. melanogaster* and other *Drosophila* species (Korol *et al.*, 1994; *see* Meiotic Mutants as an Analytical Tool in Recombination Studies: Overlapping of DNA Repair, Recombination, and Segregation Systems). Recombination studies in *Drosophila ananassae* males revealed high variation in crossing-over rate in the second and third chromosomes; individual *rf* values in the ml-ru segment of chromosome 3 varied between 2.2% and 42.8% (i.e., 20-fold). Differences between maize plants in *rf* in the sh-su-wx segment increased as the degree of relatedness decreased and there was correspondence between micro- and macrosporogenesis. Polymorphism for B chromosomes (also known as supernumerary chromosomes) and heterochromatic knobs and their interaction play a significant role in the variation of *rf* in maize. B chromosomes affect pairing and crossing-over in many plants and animals.

Abundant evidence for intraspecific genetic variation in crossing-over rate and/or chiasma frequency is available for many eukaryotes, both lower (*Coprinus*, *Schizophyllum*, *Neurospora*, *Saccharomyces*, and *Sordaria*) and higher (pearl millet, pea, jute, ryegrass, barley, wheat, tomato, snails, silkworm, grasshoppers, locusts, flour beetle, lizards, mouse, rat, and cattle) (Korol *et al.*, 1994). Human variation in fine-scale recombination patterns among individuals was examined based on dense genome-wide single-nucleotide polymorphism data in nuclear families (Coop *et al.*, 2008). Similar recombination hot spot usage in male and female meiosis was found combined with high heritable variation in the proportion of crossovers detected in these hotspots. Intraspecific variation in recombination rate and genomic distribution may be of primary importance in a better understanding of the evolution of recombination, genetic basis of biodiversity in natural systems, optimizing genetic conservation and breeding strategies, designing experiments on genetic mapping of complex traits, and map-based cloning.

Recombination and Domestication Evolution

Two questions are of major interest when discussing the interface domestication–recombination: (1) How have recombination features of the progenitors affected the pattern of selection response and selection advance? (2) How has selection for domestication traits affected the recombination system? Despite many theoretical studies on the interaction

between selection and recombination, the very limited available evidence concerns only a few organisms out of a few hundred domesticated plant and animal species. Burt and Bell (1987) studied the relationship between male chiasma frequency and features of reproductive strategy, longevity, body size, etc. of more than 30 mammalian species. They found that domesticated animal species exhibit higher rates of exchanges than might be expected based on a general relationship among these parameters. The authors attributed this difference to the advantage of increased recombination that facilitated to overcome unwanted correlations between characters in the course of selection. Similar conclusions were achieved in comparative analysis of chiasma frequency of domesticated and wild plant species (Ross-Ibarra, 2004).

At the stage of primitive selection, recombination was a major factor of genetic variation allowing for a deep and very fast (on the evolutionary scale) reorganization of domesticated organisms. Consequently, this selection pressure could become a driving force of evolution of increased recombination, at least during the first phase of domestication (thousands of generations). However, at the second phase, after regular agriculture had become established with more stable environmental conditions, stabilizing selection had to play an ever-increasing role. This might reverse the direction of recombination evolution, at least in some species. For example, Williams *et al.* (1995) established a higher chiasma frequency in primitive forms of maize compared to modern industrial lines.

Polyploid formation is considered a unique evolutionary pathway exploited by both animal and plant organisms. A few of the most important crops, such as wheat, cotton, or potato, are allopolyploids. The recombination–domestication interface may have very interesting aspects when related to polyploids. These include the evolution of diploid-like chromosome behavior in meiotic pairing, recombination and segregation, the effect of various forms of recombination on the coevolution and sequence polymorphism of genomes sharing one nucleus, and the effect of selection on the foregoing processes. Detailed mapping and DNA sequencing efforts on the polyploid crops combined with comparative evidence on their diploid relatives and/or ancestors may be considered an important way of highlighting the role of recombination in evolution and domestication.

Evolutionary Effects of Recombination

Recombination and Genetic Variation for Adaptation to Varying Environment

Producing variation is considered one of the major evolutionary functions of recombination that is supposed to assist in adaptation to both spatial and temporal environmental changes. Abundant genetic variation manifested by natural population is a result of a complex interaction between a few factors: mutation, selection, population subdivision, breeding system, and recombination in its various forms. All other things being equal, genetic variance for a sexual population with free recombination may be several times that for a population with no recombination (Charlesworth, 1993). Classical estimates from *Drosophila* species indicated that up to 25–40% of variation in fitness in natural populations is regenerated by crossing-over in one or two generations from a randomly drawn chromosome pair. Thus, a temporary arrest of the mutation process would not result in a significant decrease in variation over many generations. But does this mean that populations always harbor enough variation to react to external challenges? The author believes that the answer is negative; otherwise, it would not be possible to change recombination by direct selection for quantitative traits unrelated to recombination (*see* Experimental Evolution of Recombination).

An important question is how recombination affects genetic variation. In a population at linkage equilibrium ($LD = 0$), changes in recombination rate do not affect the amount of variation. However, this does not mean that the level of variation is independent of *rf*. Indeed, the stability of polymorphism varies with *rf*: The general trend is a reduction of polymorphism stability with increased recombination. This will occur when one models the dynamics of a diploid population subjected to stabilizing or directional selection with all selected loci linked within a block. Stability of polymorphism for these loci inversely depends on recombination within the block. However, the effect of linkage may be opposite for configurations with multiple blocks of selected loci (Nevo *et al.*, 1997). Another example is selection (either diploid or haploid) in a fluctuating environment (Korol *et al.*, 1996). The effect of recombination on the amount of variation may also result from a dependence of the steady-state LD on *rf*: $LD \neq 0$ is achievable (if at all) only at tight linkage and/or strong selection.

Recombination as a Source of Evolutionary Novelties

Through gene duplications (by unequal exchanges), recombination creates prerequisites for increasing variability. In fact, it is a general belief that the evolutionary formation of a new gene by mutation–selection interaction is preceded by the duplication of the genetic material. Local duplications may result from unequal crossing-over, unequal sister chromatid exchange, polymerase slippage, or replicative transposition. Positive association between recombination and copy number variation is well documented for various eukaryotes. There is abundant evidence (from both pro- and eukaryotes) suggesting that structurally (and in many cases functionally) related genes are linked to one another. These may include tandemly repeated gene families, gene clusters (with arbitrary spacing and orientation of the individual genes), and dispersed distribution of the family members across the genome. Different mechanisms were proposed for allele diversification after duplication. For example, a new allele may be produced by intrachromosomal homologous recombination mediated by DNA sequences within the duplicated elements. The creative role of different forms of recombination (unequal crossing-over, gene conversion, reciprocal exchange, and ectopic and illegitimate recombination) is well documented for different organisms.

Recombination was recognized to play an important role in the evolution of gene families that are usually organized in the genome as arrays of tandem repeats residing in one or a few chromosomes. Coordinated change of individual

elements of such families is referred to as concerted evolution. The diversification of the elements by mutations is opposed to homogenization processes caused by unequal crossing-over and gene conversion. This process is opposite to the scenario of evolution of new functions through "recombination (unequal crossing-over) → duplication → diversification (via mutation)". In the case of multigene families, selection presumably favors the most homogeneous arrays. Different variants of homogenized arrays may coexist within a population, as shown on the rDNA cluster in *Drosophila*. This implies a low recombination between different arrays combined with high intrachromosomal recombination within arrays.

Recombination and Sequence Polymorphism

During the past decades, new evidence became available on the association between recombination and DNA sequence polymorphism in natural populations. The first group of data came from *Drosophila*: Begun and Aquadro (1992) compared variation in *rf* per physical DNA length in different regions of *D. melanogaster* genome with polymorphism of DNA sequences of these regions. Positive correlation between recombination and sequence polymorphism was established. Similar results were obtained in different *Drosophila* species, humans, and some plants. Recombination also seems to play an important role in sequence variation of natural populations for the mitochondrial genome.

Two opposite though not mutually exclusive hypotheses were proposed as explanations for the association between recombination and sequence polymorphism. The first, referred to as selective sweeping, is based on selection of new favorable mutations. In genomic regions with a low recombination rate, the process of fixation of positively selected mutations will result in fixation of closely linked neutral or even slightly deleterious sequence variations. The second explanation is based on the theory of background selection (Nordborg *et al.*, 1996) that considers the consequences of purifying selection against deleterious mutations. This process can also reduce sequence variation at closely linked sequences. Combination of these two mechanisms seems to explain the main pattern of polymorphism–recombination association, although the contribution of additional factors cannot be excluded.

A simple and universal measure of compositional organization of nucleotide sequences is GC content. GC content displays wide interspecific variation as well as high within-genome heterogeneity, especially in mammals, where it is presented in the form of GC isochors. GC content is strongly correlated with biological features of genome organization such as distribution of various classes of repeated elements, gene density, level and tissue-specificity of transcription, and mutation rate. The available data on humans and other organisms point to a strong correlation between GC content and recombination. This connection may be caused by increased recombination in GC-rich regions. An alternative assumption is that recombination machinery is responsible for the evolution of the genomic GC distribution due to the bias of mismatch repair within the gene conversion tracts in favor of GC, which preferentially converts A/T into G/C at sites that are heterozygous for AT and GC (the biased gene conversion hypothesis). Also, we cannot exclude the possibility that another factor affecting both the evolution of intragenomic heterogeneity in GC content and recombination is responsible for their positive association. An important aspect of the recombination impact on genome organization is related to the amount and distribution of repetitive elements. Amplification of retrotransposons is considered as an important source of repetitive DNA leading to an increase in genome size ("genomic obesity") in many plant taxa. Recombination was proposed as the mechanism counterbalancing this trend, in particular, unequal crossing-over and intrachromosomal recombination between the repeated elements (Shirasu *et al.*, 2000). Recombination is an important mechanism affecting the formation of microsatellites. Owing to their abundance in the genomes and high variability, microsatellites are considered as accelerators of evolutionary adaptation (Kashi and King, 2006).

Horizontal Gene Transfer

Numerous mechanisms exist that normally prevent the exchange of genetic material among distant species. Nevertheless, if this process is possible even at a low frequency, it may have important consequences on the long-term evolutionary scale. Only 30 years ago, such a possibility was considered an unrealistic one. Now, rich evidence is available showing that the opposite is true. Illegitimate recombination seems to be the main responsible mechanism. Extensive similarities have been revealed between house-keeping genes of Archaea and Eubacteria, although for the majority of informational genes (involved in transcription and translation) the difference is very large. An interesting fact is that some of the foregoing genes showing high similarity between the kingdoms are clustered in the genome that may result from horizontal transfer events.

If horizontal transfer is indeed an evolutionary significant form of genetic recombination, one could consider ecological interaction between species a replacement of sexual interaction (vertical transfer). In other words, ecologically interacting species are the most probable exchanging partners. Viral DNA integration in the host nuclear genome is considered a relatively frequent event (on the evolutionary scale) for animal hosts and recently also for plants. Likewise, there is evidence of an opposite, host → parasite flow of genetic information: Many transfers of signaling domains of pathogenesis-related plant genes were detected in bacterial genomes.

Evolution of Recombination

Theoretical Models of Recombination Evolution

Artificial selection experiments suggest that almost every population has enough stored genetic variability to ensure response to direct selection for changed recombination frequency (*see* Selection for Changed Recombination, Genetic Modifiers of Recombination, *and* Intraspecific Variation for the Rate of Recombination). The observed polymorphism at loci affecting recombination could be either balanced (selected) or

transient. Theoretical analysis shows that under a stable environment, a panmictic population should evolve toward a minimum possible level of recombination. This can be formulated in terms of the fate of a selectively neutral modifier locus affecting recombination.

Understanding the forces maintaining sex and recombination is considered one of the most challenging problems in evolutionary theory (Michod and Levin, 1988). Shared genetic control and molecular mechanisms of DNA recombination and DNA repair across life (*see* Meiotic Mutants as an Analytical Tool in Recombination Studies: Overlapping of DNA Repair, Recombination, *and* Segregation Systems and Molecular Mechanisms of Recombination) indicate their common evolutionary origin and functional overlap in extant organisms. It could probably be supposed that repair functions played the leading role at early evolutionary stages, having provided opportunities for a large increase in the genome size and the transition from haploidy to diploidy. The latter offered the possibility of recombination repair of two-strand DNA lesions that is impossible in haploid systems. Some authors hypothesize that the subsequent stages in the evolution of recombination and sex were associated with repair alone. These explanations of repair functions were called "physiological" (Maynard Smith, 1978). Another physiological explanation of recombination in sexuals is its association with chromosome segregation (*see* Meiotic Mutants as an Analytical Tool in Recombination Studies: Overlapping of DNA Repair, Recombination, *and* Segregation Systems), although numerous examples are known, such as male *Drosophila* and female silkworm, in which normal segregation is associated with achiasmatic (without crossing-over) meiosis.

Combinative, or generative, hypotheses consider the main function of sex and recombination in shuffling genes. This removes negative correlation between favorable alleles at different loci, thereby increasing the efficiency of natural selection. Generative models can be classified according to the source of linkage disequilibria between selected loci: stochastic or deterministic, caused by new mutations or variation of external conditions (Kondrashov, 1993). In the 1930s, Fisher and Muller proposed that sex may be advantageous by combining beneficial mutations randomly occurring in different individuals (Hartfield *et al.*, 2010). A complementary version of the stochastic–mutation explanation considers the role of recombination in selection against deleterious mutations (Muller, 1964). According to Muller, deleterious mutations tend to be fixed in a finite asexual population due to random drift despite purifying selection (Muller's ratchet), whereas recombination helps to stop this process. In the 1980s, a few deterministic models of selection against deleterious alleles were proposed, with the evolutionary advantage of recombination being dependent on linkage disequilibria resulted from synergistic interaction between harmful mutations produced at a high rate (Kondrashov, 1988). Other models of deterministic sources of linkage disequilibria include (1) adaptation to temporarily or spatially varying environment, (2) antagonistic species interaction (the Red Queen hypothesis), and (3) instraspecific competition (sib competition or tangled bank hypothesis) (Hamilton *et al.*, 1990; Kondrashov, 1993; Korol *et al.*, 1994; Feldman *et al.*, 1997; Barton and Charlesworth, 1998; Otto and Michalakis, 1998). In most of these models, the conditions favoring increased recombination are associated with negative linkage disequilibria among selected loci owing to stringent conditions for epistasis. Some mechanisms (e.g., selection of beneficial mutations) may favor recombination in the absence of epistasis, with linkage disequilibrium caused by finite population size. The diversity of models proposed to explain the evolution of sex and recombination and their rather complicated "relationships" with the evidence from studies of different organisms, lead to the idea of pluralistic explanation (Michod and Levin, 1988; Korol *et al.*, 1994; West *et al.*, 1999). It suggests that different mechanisms could have played different roles in the emergence, evolution, and, especially, maintenance of recombination and sex in diverse phyletic lines of eukaryotes.

Testing the Theoretical Assumptions

Despite voluminous literature, the problem has been studied little experimentally or by observations in natural populations. In the classical period of genetics, it was proposed that sex and recombination may facilitate selection for advantageous and against deleterious mutations and assist in adaptation to changing environment. The major question is whether these processes in turn are able to promote evolution of the recombination system. Are the resulting changed patterns of recombination detectable experimentally? Is it possible to find corresponding changes in nature? Theoretical models answer positively to these questions. However, do these models (and/or their parameter values) fit the real world? Unfortunately, there is not enough evidence to answer this question. Two different and complementary approaches can be considered (Korol *et al.*, 1994; Barton and Charlesworth, 1998; Korol, 1999).

Testing the Initial Assumptions

The aim is to test whether the real estimates of the main parameters underlying the proposed mechanism(s) fit the expectations. These include mutation rate, effective population size, mode of interaction of deleterious or beneficial mutations, relative role of stabilizing and directional selection, and cost of sex and meiosis. The major advantage of this approach is that it allows the accumulation of data on basic parameters affecting the evolution of recombination. The drawback is that it is difficult to believe that even the main factors were taken into account so that the predicted direction of recombination evolution is determined correctly. As an example of a successful application of this approach, competition experiments with yeast *S. cerevisiae* (Greig *et al.*, 1998) are worth mentioning. Using sexual and asexual strains it was found, contrary to the expectations following from the concept of costs of sex or meiosis, that sex can confer significant selective advantage to its carriers. Direct tests have also been conducted with the nematode *C. elegans* demonstrating the possible role of recombination in purifying selection (Zetka *et al.*, 1987). It was shown that increased *rf* owing to enhancer Rec-1 results in a higher fitness of genotypes carrying mutator gene mut-6 compared to analogous genotypes with a normal recombination rate. Therefore, the presence of a gene for increased recombination, which presumably has no effect

on mutation, reduces the genetic load caused by the mutator system. However, studies of normal (wild-type) genotypes of *C. elegans* give very low estimates of mutation rate, which is inconsistent with the hypothesis of selection against deleterious mutations.

Searching for Changed Recombination

The second approach is to test the "final effect" by comparing rates and genomic distribution of recombination in experimental and natural populations subjected to contrasted selection regimes during many generations. The advantage of such an analysis is that the target effect (changed pattern of recombination) is estimated directly (*see* Experimental Evolution of Recombination), although one cannot be sure that it is causally connected only with the presumed process.

Recombination System and Life History Traits

Rasmusson (1927) was possibly the first to discuss the genetic basis of putative differences in recombination systems of outbreeding and inbreeding species. On the basis of extremely high variation in the recombination rate in peas that is usually not found in *Drosophila*, he advanced a hypothesis of possible differences in the system of genetic control of recombination between outbreeders and selfers. In plants, elevated chiasma frequencies were indeed found in selfers compared to related cross-pollinating species (Ross-Ibarra, 2006). The species origin and life conditions determine the strategy for adaptation and hence the "optimum" amount of recombination. Evidence of variation of chiasma frequency in nature and its correlation with ecological and life historical traits is available for different organisms and may be employed to compare alternative theoretical models. In mammals with low reproduction rates, longer life cycles, and small progeny sizes, recombination is higher than in species with short generation times and large progeny sizes (Burt and Bell, 1987), fitting the Red Queen model. Koella (1993) found that in plants with animal-dispersed seeds, recombination is higher than in other plants, and perennials show a higher recombination than annuals when compared among genera but an opposite tendency is displayed on the species level. Similarly, on the whole-genome level, higher cM/Mb ratios were found in trees compared to short-lived herbs and shrubs (Jaramillo-Correa *et al.*, 2010).

The rapidly increasing number of genetically mapped organisms simultaneously demonstrates two opposite patterns. The first is a rather high similarity of some genome-wide characteristics of the recombination system among different species, including the narrow spectrum of the crossover number per chromosome, positive interference, the centromeric effect, and a positive correlation with GC content. The second is the discovery of high variation of recombination features within and between species, with extreme outliers, like honey bees (and other social insects), with their extremely low Mb/cM ratio compared to other insects and some species, for example, *Monodelphis* or *Pinus* and *Hordeum*, with considerably higher cM/Mb compared to correspondingly other vertebrates and plants. In most of the cases, we still do not know the reasons for both, high similarity and high variation. Especially intriguing is the high stability of number of exchanges per chromosome among higher eukaryotes (4–6 sixfold) despite huge variation of Mb/cM ratio (more than two orders of magnitude – see **Table 2**).

Experimental Evolution of Recombination

Changes in Recombination Caused by Selection for Adaptive Traits

Theoretical models indicate that directional or variable selection for multilocus traits may promote evolution toward increased recombination if the selected population is polymorphic for recombination-modifying loci (Charlesworth, 1993; Korol *et al.*, 1994; Otto and Barton, 1997). Only a few experiments seem to have a direct bearing on this question. The following are examples with regard to *Drosophila* (Korol, 1999). Clegg *et al.* (1979) attempted to test Fisher's (1930) prediction that stable favorable conditions may promote evolution toward reduced recombination. The experiment started with a large cage population of *D. melanogaster* with maximum LD between allozyme loci of chromosome 3. No general trend in *rf* for 20 generations was revealed, which was explained as an indication of a low epistatic component of selection within the marked interval. Flexon and Rodell (1982) established a synthetic population of *D. melanogaster* and subjected it to selection for DDT resistance for 300 days. By the time of estimation, resistance increased from 10- to 100-fold relative to the control level. The control and treatment populations were compared with respect to recombination rate across the genome. Increased *rf* between marker loci on chromosomes 2 and 3 but not X was found in the selected population. Most experimental studies on recombination evolution have employed one-way selection for fitness traits. The limitation of this scheme is the danger of a change in *rf* caused by initial linkage disequilibria between selected genes and recombination modifiers. Therefore, to assess the effect of selection for geotaxis on recombination in *D. melanogaster*, Korol and Iliadi (1994) employed two-way selection. Forty generations of selection for altered geotaxis resulted in an increment in recombination for the genome of 78 cM for geo+ and 66 cM for geo− lines compared to nonselected control. Selection for negative geotaxis did not affect *rf* in chromosome 2, whereas selection in the opposite direction caused a 4-fold increase in *rf* in the b-cn segment spanning the centromere of chromosome 2. The observed change in *rf*, caused by two-way selection for geotaxis, was attributed to the advantage conferred by selection on recombinants. Similar results were obtained by Rodell *et al.* (2004) in divergent selection for the sternopleural bristle number in *D. melanogaster*. Significantly, increased rates of recombination in chromosomes 2 and 3 were found in both directions of selection, and a tendency for reducing crossover rates was observed in populations subjected to stabilizing selection.

Experimental Evolution of Recombination Caused by Adaptation to Stress

The author reviewed a series of studies with *D. melanogaster* conducted in his laboratory and aimed to demonstrate that population adaptation to adverse conditions may

promote evolution for increased recombination (Korol *et al.*, 1994; Korol, 1999). The experiments were performed on four large heterogeneous populations with the objective of simulating the process of population adaptation to stress and evaluating its effect on recombination. At the start, each population was subdivided randomly into two parts, control (C) and treatment (T); T variants were then subjected to daily fluctuating temperature with amplitude increasing with generations (from 23 to 27 °C at the beginning to 12–32 °C at the end of the experiment). C variants were maintained at 25 °C.

Increased Recombination Resulting from Adaptation to Stressful Conditions

The control and treatment populations appeared to diverge genetically with respect to thermotolerance. Changes in recombination were evaluated using multiple marked lines. All four populations showed one pattern – a segment-specific increase in *rf*, resulting from selection for thermoadaptation. Those that reacted most were the near-centromeric regions of chromosomes 2 and 3 and the left arm of chromosome 3 (**Figure 3**). The T and C variants were also compared with respect to the frequency of double exchanges (interference). Significant differences in the values of coefficients of coincidence have been found for adjacent and nonadjacent intervals of large autosomes. The general tendency was a reduced interference in T variants, corroborating the results for recombination rate.

Reduction of Recombination Rate under Optimal Stable Conditions

The described divergence over time with respect to recombination rate resulted from two independent trends. An increase in *rf* was observed in populations adapting to adverse conditions, and simultaneously tightening of linkage in control populations that were maintained under optimal temperature. The proportion of these two effects varied in different "treatment–control" pairs corresponding to the structure of the starting population. These results can be considered as the first experimental verification of Fisher's (1930) hypothesis that in a constant environment selection favors tighter linkage. Indirectly, these results also indicate that selection against harmful mutations alone seems to be insufficient for maintenance of the normal level of recombination, at least in the genetic material that was tested.

Applied Aspects

Recombination as a Tool of Genetic Mapping

Recombination-based mapping as a method of studying the genetic material topography has been a major analytical approach in genetics for nearly a century. However, until recently this technique was of little applied interest because of a lack of easily scorable markers. The situation changed dramatically with the appearance of DNA markers. Establishing high-quality maps was considered one of the most important objectives in the Human Genome Project, which also included a few model organisms. Mapping efforts allow a better understanding of the organization of genetic material and allow one to conduct far-reaching comparative mapping. This approach is a power tool of ever-increasing importance for studies of genome structure, functioning, and evolution, and for diverse medical and commercial applications of modern genomics. Fine genetic mapping is a basis of positional cloning, an approach proved successful in many organisms. Its efficiency strongly depends on the Mb/cM ratio in the region harboring the target gene. If the target region is very rarely involved in recombination, then map-based cloning is impractical. Induced recombination may be a nontraditional approach for such situations. The availability of dense maps facilitates mapping of genetic disease genes in humans with subsequent cloning, sequencing, and medical applications. An important application of mapping is genetic dissection of complex traits, or quantitative trait loci mapping, especially in the framework of new marker-assisted (or "genomic selection") strategies in plant and animal breeding.

Recombination as a Major Source of Genetic Variation in Breeding

When viewed at the level of phenotypic traits, genetic innovations produced by recombination on artificial hybridization

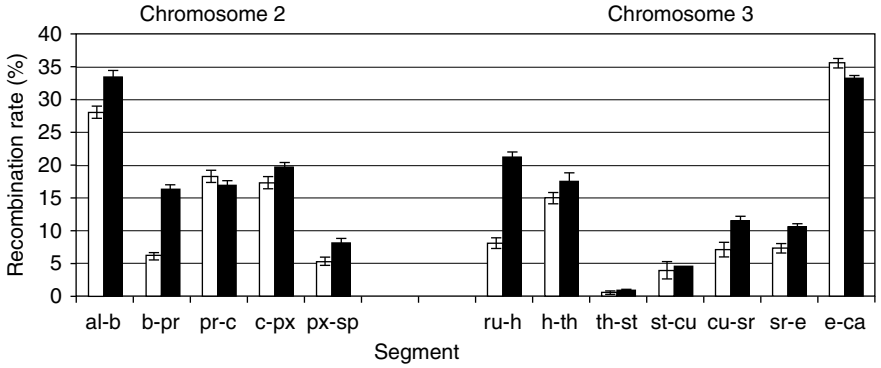


Figure 3 Adaptation to daily fluctuating temperature results in a segment-specific increase in recombination rate in large autosomes of *D. melanogaster* (based on data from Korol AB, Preygel IA, and Preygel SI (1994) *Recombination Variability and Evolution*. London: Chapman & Hall).

can arbitrarily be divided into three main types: (1) transgression for individual traits with the range of trait values in the segregating progeny exceeding the respective parental ranges; (2) formation of new trait combinations of the crossed components; and (3) the appearance of new traits ("anomalous" variation) as a result of recombination in genetic complexes with strong nonallelic interactions. These forms of recombination variability play a significant role in breeding programs providing the raw material for selection. The importance of controlling recombination increases with the necessity of involving new genetic resources for breeding purposes, especially in light of ever-decreasing homeostasis and tolerance to abiotic and biotic stresses of elite animal breeds and plant cultivars. New genes are supposed to be introgressed from exotic germplasm based on recombination. This problem is complicated by reduced recombination on distant crosses and linkage drag (close linkage between the target gene and undesirable genes). Hence, manipulating recombination may become an efficient tool of modern breeding.

Recombination and Cancer

There are many aspects of anomalous chromosome and cell behavior and cancer related to DNA metabolism and recombination–repair. A few are noted here.

Mitotic Recombination

A surprisingly large proportion of human neoplasms appear to result from loss of function of tumor-suppressing genes, for example, due to loss of heterozygosity of tumor suppressors. In retinoblastomas caused by loss of heterozygosity, about half of the homozygosity cases appear to be the result of mitotic recombination.

Specificity of Chromosome Translocations

Many types of human cancers are associated with nonhomologous recombination events producing specific translocations. This results in an enhanced expression of oncogenes residing in the vicinity of the breakpoint. Some of these oncogenes are transcription factors participating in normal development. It seems reasonable to assume that interaction between recombination and transcription (*see* Recombination Signals, Hotspots of Recombination, and Transcription) may also involve ectopic recombination resulting in the specificity of some tumorigenic translocations (Barr, 1998).

Defects in DNA Repair–Recombination

A few cancer-related genetic disorders of this type have been characterized, including Bloom syndrome, which manifests mutator and hyper-recombination phenotypes; Ataxia telangiectasia mutation, which increases V(D)J-mediated chromosome rearrangement in T lymphocytes; and Omenn syndrome, a disorder caused by mutations of recombination–activation genes Rag1 and Rag2 (related to V(D)J recombination) resulting in severe immune deficiency combined with a complete lack of circulating T and B cells. Recent studies in *Drosophila* and mice show that one of the most important tumor-suppressing genes controlling cell apoptosis, p53, is also involved in meiotic recombination, suggesting that tumor-suppressive functions may have been co-opted from primordial activities linked to recombination (Lu *et al.*, 2010).

Recombination Shuffling of Genes: New Technologies of Artificial Sequence Optimization

Recombination combined with selection has become the basis for new technologies of "molecular breeding" aimed at improving protein functions via directed evolution *in vitro*. It includes reshuffling of a few copies of the coding DNA by homologous or site-specific recombination with subsequent screening of the resulting sequences for functional (enzymatic) activity. The reshuffling can be provided by template switching in selfpriming polymerase chain reaction (PCR) referred to as sexual PCR. The efficiency of such artificial evolution may be enhanced by increasing the initial genetic variation involved in recombination, for example, by mutagenesis of the target sequences based on error-prone PCR. Another possibility is to include a wider spectrum of parental sequences using the natural diversity of homologous genes of the same or related species or even distant taxa (Crameri *et al.*, 1998). This approach is especially promising for the improvement of molecules that provide resistance against toxic agents or increase survival on minimal media. Experiments including transformation of the engineered sequences into single-cell organisms (*E. coli*) and selection of the transformants showed dozens-to-hundreds- and even thousands-fold increases in efficiency of the target molecules obtained after recombination–selection cycles. DNA shuffling can improve the targeted pathway by complex mechanisms, even though it may be impossible to predict the optimized sequence by a rational design. Of special interest is the use of recombination for engineering novel therapeutic drugs. Namely, homology-independent recombination of human genes with genes from other species may be used for construction of "humanized" proteins (with a part of nonhuman sequence) possessing the desired clinical activity (Griswold *et al.*, 2005). The success of using recombination as a strategy for protein engineering corroborate the findings that recombination causes more tolerable changes in protein coding sequences compared to random mutation, implying a higher importance of recombination as a source of meaningful evolutionary genetic variation (Drummond *et al.*, 2005).

Targeted Gene Replacement

Gene targeting is an experimental approach aimed at producing site-specific changes in the genome based on genetic recombination. Predetermined changes can be introduced in the target site, in both somatic and germline cells. Usually, the targeting system involves a gene mediating site-specific recombination and a specific target locus. The best-known examples are Cre and Flp recombinases (of bacteriophage P1 and yeast) with IoxP and FRT targets, respectively. The basic idea is to first transform the target sites flanking a selectable marker into the recipient genome. Subsequent introduction of the recombinase activity leads to the excision

of the marker. Many important applications have been proposed and many others are expected using the powerful strategies of targeted gene replacement. These include controlled expression of transformed genes, disruption of a chosen gene *in vivo* for a better understanding of its role in physiological and pathological processes, disruption of genes to rechannel the biosynthesis and achieve overproduction of certain molecules, site-specific targeting of DNA into embryonic stem cells for production of gene-modified model organisms, targeted delivery of specific constructs aimed at producing antitumor effects or complementing a mutated gene, and testing of gene expression in donor versus recipient tissues in biomedical transplantation studies.

Conclusions

Recombination plays a key role in a remarkably broad spectrum of genetic processes related to DNA repair; sexual reproduction; immunity; genetic diversity; evolution of genes, gene families, chromosomes and the entire genome; and evolutionary adaptation and speciation. Recombination is the basis for traditional breeding, genetic mapping, recombinant DNA technologies and genetic engineering, gene targeting and gene therapy, developmental biology, and cancer genetics. Genetic studies of various organisms have accumulated abundant evidence on common features and peculiarities in genetic control and molecular mechanisms of the main types of genetic recombination, including reciprocal exchange between homologous chromosomes (crossing-over), nonreciprocal exchange (gene conversion), site-specific recombination occurring between homologous sequences, and illegitimate recombination between nonhomologous sequences. However, despite a long history, thousands of studies, and great progress during the last two decades, recombination continues to be a puzzle in the aspects of its mechanisms, evolutionary effects, contribution of biodiversity, and the factors determining its own evolution.

See also: Adaptation. Coevolution. Diversity, Molecular Level. Genetic Diversity. Life History, Evolution of

References

Agrawal AF and Wang AD (2008) Increased transmission of mutations by low-condition females: Evidence for condition-dependent DNA repair. *PLoS Biology* 6: e30.

Andersen SL and Sekelsky J (2010) Meiotic versus mitotic recombination: Two different routes for double-strand break repair. *Bioessays* 32: 1058–1066.

Barr FG (1998) Translocations, cancer and the puzzle of specificity. *Nature Genetics* 19: 121–124.

Barton NH and Charlesworth B (1998) Why sex and recombination? *Science* 281: 1986–1990.

Baudat F, Buard J, Grey C, *et al.* (2010) PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327: 836–840.

Begun DJ and Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* 356: 519–520.

Bhagat R, Manheim EA, Sherizen DE, and McKim KS (2004) Studies on crossover-specific mutants and the distribution of crossing over in *Drosophila* females. *Cytogenetic and Genome Research* 107: 160–171.

Burt A and Bell B (1987) Mammalian chiasma frequencies as a test of two theories of recombination. *Nature* 326: 803–805.

Charlesworth B (1993) Directional selection and evolution of sex and recombination. *Genetical Research* 61: 205–224.

Clegg MT, Horch CR, and Kidwell JF (1979) Dynamics of correlated genetic systems. VI. Variation in recombination rates in experimental population of *Drosophila melanogaster*. *Journal of Heredity* 70: 297–300.

Conway DJ, Roper C, Oduola AM, *et al.* (1999) High recombination rate in natural populations of *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences USA* 96: 4506–4511.

Coop GC, Wen X, Ober C, Pritchard JK, and Przeworski M (2008) High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* 319: 1395–1398.

Crameri A, Raillard SA, Bermudez E, and Stemmer WP (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 391: 288–291.

Cromie GA and Smith GR (2007) Branching out: Meiotic recombination and its regulation. *Trends in Cell Biology* 17: 448–455.

Danzmann RG, Davidson EA, Ferguson MM, *et al.* (2008) Distribution of ancestral proto-Actinopterygian chromosome arms within the genomes of 4R-derivative salmonid fishes (Rainbow trout and Atlantic salmon). *BMC Genomics* 9: 557.

Drummond DA, Silberg JJ, Meyer MM, Wilke CO, and Arnold FH (2005) On the conservative nature of intragenic recombination. *Proceedings of the National Academy of Sciences USA* 102: 5380–5385.

Feldman MW, OttoS P, and Christiansen FB (1997) Population genetic perspectives on the evolution of recombination. *Annual Review of Genetics* 30: 261–295.

Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford: Clarendon.

Flexon PB and Rodell CF (1982) Genetic recombination and directional selection for DDT resistance in *Drosophila melanogaster*. *Nature* 298: 672–675.

Greig D, Borts RH, and Louis EJ (1998) The effect of sex on adaptation to high temperature in heterozygous and homozygous yeast. *Proceedings of the Royal Society London B* 265: 1017–1023.

Griswold KE, Kawarasaki Y, Ghoneim N, Benkovic SJ, Iverson BL, and Georgiou G (2005) Evolution of highly active enzymes by homology-independent recombination. *Proceedings of the National Academy of Sciences USA* 102: 10082–10087.

Hadany L and Beker T (2003) On the evolutionary advantage of fitness-associated recombination. *Genetics* 165: 2167–2179.

Hamilton WD, Axelrod RA, and Tanese R (1990) Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences USA* 87: 3566–3573.

Hartfield M, Otto SP, and Keightley PD (2010) The role of advantageous mutations in enhancing the evolution of a recombination modifier. *Genetics* 184: 1153–1164.

Heyer WD, Ehmsen KT, and Liu J (2010) Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics* 44: 113–139.

Hyppa RW and Smith GR (2010) Crossover invariance determined by partner choice for meiotic DNA break repair. *Cell* 142: 243–255.

Jaramillo-Correa JP, Verdú M, and González-Martínez SC (2010) The contribution of recombination to heterozygosity differs among plant evolutionary lineages and life-forms. *BMC Evolutionary Biology* 10: 22.

Kashi Y and King DG (2006) Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* 22: 253–259.

Koella JC (1993) Ecological correlates of chiasma frequency and recombination index of plants. *Biological Journal of the Linnean Society* 48: 227–238.

Kondrashov AS (1988) Deleterious mutations and the evolution of sexual reproduction. *Nature* 336: 435–440.

Kondrashov AS (1993) Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity* 84: 372–387.

Koren A, Ben-Aroya S, and Kupiec M (2002) Control of meiotic recombination initiation: A role for the environment? *Current Genetics* 42: 129–139.

Korol AB (1999) Selection for adaptive traits as a factor of recombination evolution: Evidence from natural and experimental populations. In: Wasser SP (ed.) *Evolutionary Theory and Processes: Modern Perspective*, pp. 31–53. Dordrecht: Kluwer.

Korol AB and Iliadi KG (1994) Recombination increase resulting from directional selection for geotaxis in *Drosophila*. *Heredity* 72: 64–68.

Korol AB, Kirzhner VM, Ronin YI, and Nevo E (1996) Cyclical environmental changes as a factor maintaining genetic polymorphism. II. Two-locus diploid selection. *Evolution* 50: 1432–1441.

Korol AB, Preygel IA, and Preygel SI (1994) *Recombination Variability and Evolution*. London: Chapman & Hall.

Landner L (1970) Variation of recombination frequency in *Neurospora crassa* following meiosis and evidence for a premeiotic sensitive stage. *Molecular and General Genetics* 27: 385–392.
Laurencon A, Gay F, Ducau J, and Bregliano J-C (1997) Evidence for an inducible repair-recombination system in the female germ line of *Drosophila melanogaster*. III. Correlation between reactivity levels, crossover frequency and repair efficiency. *Genetics* 146: 1333–1344.
Lenormand T and Dutheil J (2005) Recombination difference between sexes: A role for haploid selection. *PLoS Biology* 3: e63.
Lichten M and Goldman ASH (1995) Meiotic recombination hotspots. *Annual Review of Genetics* 29: 423–444.
Loidl J and Scherthan H (2004) Organization and pairing of meiotic chromosomes in the ciliate *Tetrahymena thermophila*. *Journal of Cell Science* 117: 5791–5801.
Lu W-J, Chapo J, Roig I, and Abrams JM (2010) Meiotic recombination provokes functional activation of the p53 regulatory network. *Science* 328: 1278–1281.
Lukaszewski AJ and Curtis CA (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theoretical and Applied Genetics* 86: 121–127.
Lupu A, Nevo E, Zamorzaeva I, and Korol AB (2006) Ecological-genetic feedback in DNA repair in wild barley, *Hordeum spontaneum*. *Genetica* 127: 121–132.
Maynard Smith J (1978) *The Evolution of Sex*. Cambridge, UK: Cambridge University Press.
Maynard Smith J, Dowson CG, and Spratt BG (1991) Localized sex in bacteria. *Nature* 349: 29–31.
McKim KS and Hayashi-Hagihara A (1998) mei-W68 in *Drosophila melanogaster* encodes a Spoll homolog: Evidence that the mechanism for initiating meiotic recombination is conserved. *Genes and Development* 12: 2932–2942.
McVicker G and Green P (2010) Genomic signatures of germline gene expression. *Genome Research* 20: 1503–1511.
Michod RE and Levin BR (eds.) (1988) *The Evolution of Sex: An Examination of Current Ideas*. Sunderland, MA: Sinauer.
Molinier J, Ries G, Zipfel C, and Hohn B (2006) Transgeneration memory of stress in plants. *Nature* 442: 1046–1049.
Morrell PL, Toleno DM, Lundy KE, and Clegg MT (2006) Estimating the contribution of mutation, recombination and gene conversion in the generation of haplotypic diversity. *Genetics* 173: 1705–1723.
Muller HJ (1964) The relation of recombination to mutational advance. *Mutation Research* 1: 2–9.
Neel JV (1941) A relation between larval nutrition and the frequency of crossing over in the third chromosome of *Drosophila melanogaster*. *Genetics* 26: 506–516.
Nevo E, Kirzhner V, Beiles A, and Korol A (1997) Selection versus random drift: Long-term polymorphism persistence in small populations (evidence and modelling). *Philosophical Transactions of the Royal Society B* 352: 381–389.
Nordborg M, Charlesworth B, and Charlesworth D (1996) The effect of recombination on background selection. *Genetics Research* 67: 159–174.
Otto SP and Barton NH (1997) The evolution of recombination: Removing the limits to natural selection. *Genetics* 147: 879–906.
Otto SP and Michalakis Y (1998) The evolution of recombination in changing environments. *Trends in Ecology and Evolution* 13: 145–151.
Pecinka A, Rosa M, Schikora A, *et al.* (2009) Transgenerational stress memory is not a general response in *Arabidopsis*. *PLoS One* 4: e5202.
Rasmusson J (1927) Genetically changed linkage values in Pisum. *Hereditas* 10: 1–52.
Rodell CF, Schipper MR, and Keenan DK (2004) Modes of selection and recombination response in *Drosophila melanogaster*. *Journal of Heredity* 95: 70–75.
Ross-Ibarra J (2004) The evolution of recombination under domestication: A test of two hypotheses. *American Naturalist* 163: 105–112.
Ross-Ibarra J (2006) Genome size and recombination in angiosperms: a second look. *Journal of Evolutionary Biology* 20: 800–803.
Sandler L and Szauter P (1978) The effect of recombination-defective meiotic mutants on fourth-chromosome crossing-over in *Drosophila melanogaster*. *Genetics* 90: 699–712.
Schoustra S, Rundle HD, Dali R, and Kassen R (2010) Fitness-associated sexual reproduction in a filamentous fungus. *Current Biology* 20: 1350–1355.
Shalev G, Sitrit Y, Avivi-Ragolski N, Lichlenstein C, and Levy AA (1999) Stimulation of homologous recombination in plants by expression of the bacterial resolvase ruvC. *Proceedings of the National Academy of Sciences USA* 96: 7398–7402.
Shirasu K, Schulman AH, Lahaye T, and Schulze-Lefert P (2000) A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Research* 10: 908–915.
Stevison LS and Noor MAF (2010) Genetic and evolutionary correlates of fine-scale recombination rate variation in *Drosophila persimilis*. *Journal of Molecular Evolution* 71: 332–345.
Sved JA, Blackman LM, Gilchrist AS, and Engels WR (1991) High level of recombination induced by homologous P elements in *Drosophila melanogaster*. *Molecular and General Genetics* 225: 443–447.
Szostak JW, Orr-Weaver TL, Rothstein RJ, and Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* 33: 25–35.
Takeda J, Uematsu N, Shiraishic S, Toyoshima M, Matsumotoaand T, and Niwae O (2008) Radiation induction of delayed recombination in *Schizosaccharomyces pombe*. *DNA Repair* 7: 1250–1261.
Thuriaux P (1977) Is recombination confined to structural genes of the eukaryotic chromosome? *Nature* 268: 460–462.
Trivers R (1988) Sex differences in rates of recombination and sexual selection. In: Michod RE and Levin BR (eds.) *The Evolution of Sex: An Examination of Current Ideas*, pp. 270–286. Sunderland, MA: Sinauer.
Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, and Zouros E (2005) Widespread recombination in published animal mtDNA sequences. *Molecular Biology and Evolution* 22: 925–933.
West SA, Lively CM, and Read AF (1999) A pluralist approach to sex and recombination. *Journal of Evolutionary Biology* 12: 1003–1012.
Whitehouse HLK (1966) An operator model of crossing over. *Nature* 211: 708–713.
Williams CG, Goodman MM, and Stuber CW (1995) Comparative recombination distances among *Zea mays* L. inbreds, wide crosses and interspecific hybrids. *Genetics* 141: 1573–1581.
Yamamoto T (1961) Progenies of sex-reversal females mated with sex-reversal males in the Medaka, Oryzias latipes. *Journal of Experimental Zoology* 146: 163–179.
Zetka MC, Rose AM, and Baillie DL (1987) A test of the Muller Ratchet Hypothesis using the REC-1 strain of *Caenorhabditis elegans*. *Genetics* 116(supplement): 30.
Zhuchenko AA and Korol AB (1983) Ecological aspects of the recombination problem. *Theoretical and Applied Genetics* 64: 177–185.
Zhuchenko AA, Kovtyukh LP, and Korol AB (1983) Testing the "ontogenetic memory" hypothesis on *Drosophila*. *Doklady Academii Nauk USSR* 269: 1493–1495 (in Russian).

Reforestation

David Lamb, University of Queensland, Brisbane, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Degradation A temporary or permanent loss of forest structure, productivity, and native species diversity. A degraded site might still contain trees (i.e., a degraded site is not necessarily deforested) but it has lost at least some of its former ecological integrity and has a lowered capacity to supply goods or services.

Ecological restoration To re-establish the presumed structure, productivity, and species diversity of the forest originally present at a site. The ecological processes and functions of the restored forest will closely match those of the original forest.

Plantation monoculture A plantation forest involving a single species. The original biodiversity at a site is not recovered although the protective function and many of the original ecological services may be re-established.

Reforestation The re-establishment of trees and understory plants at a site previously occupied by forest cover. This can occur by natural means or by some form of human intervention including planting seedlings or sowing seeds.

Rehabilitation To re-establish the productivity and some, but not necessarily all, of the plant and animal species thought to be originally present at a site. For ecological or economic reasons, the new forest might also include species not originally present at the site. The protective function and many of the ecological services of the original forest may be re-established.

Introduction

Biodiversity and productivity are lost when natural forest ecosystems are fragmented or cleared. Reforestation is used to reverse this process but this can take a variety of forms. In some cases natural regeneration may occur and be sufficient to restore forest cover but, in other cases, direct seeding or planting seedlings will be necessary if reforestation is to occur within a reasonable time. Reforestation is often used to simply restore productivity but significant proportions of plant and animal biodiversity can also be reestablished if appropriate silvicultural techniques are used. The extent to which biodiversity is restored depends on (1) the type of reforestation, (2) the area covered, and (3) its location within the landscape. True ecosystem restoration is very difficult and more modest goals are often all that is possible, particularly over large areas. Successful reforestation usually requires solving social and economic problems as well as just ecological problems. These socioeconomic issues are often the most difficult to resolve because the benefits of reforestation may take some years to develop whereas the costs to landholders are immediate.

Need for Reforestation

Large areas of the world's forested lands have been cleared to enable agricultural development. Estimates by the Food and Agricultural Organisation of the United Nations (FAO) (2010) suggest the annual net rate of loss in forest cover was 8.3 million hectare between 1990 and 2000 and 5.2 million hectare between 2000 and 2010. Even if the rate of the world's population growth is reduced and human populations are stabilized, it is possible that much more of the world's remaining natural ecosystems will be cleared. Many of these cleared lands have been used for food or fiber production but large areas have simply been degraded through misuse and then abandoned. In the face of this process of biological simplification, much effort has gone into developing a comprehensive, representative, and adequate series of nature reserves to protect the remaining biodiversity. There are doubts about the effectiveness of this network because most reserve areas are isolated and often cover only a small proportion of the landscape. More biodiversity might be protected if some of the under-used degraded lands could be reforested and different components of the reserve network could be linked together.

Methods of Overcoming Degradation

Degradation is a subjective term. A newly cleared area of forest might be regarded as prime agricultural land by a farmer but as degraded wildlife habitat by a bird enthusiast. To some extent, degradation is in the eye of the beholder. However, many would agree that once-forested land that has lost both its structure and diversity and is not used in any productive way is now, in some sense, degraded. This view would be reenforced if economic productivity had also been reduced. The two main ecological components of degradation are shown in **Figure 1**. One component is the proportion of the original biodiversity still present and the other is the proportion of the original structure or productivity remaining. The original undisturbed ecosystem is at point A and point B which represents the same site in a degraded state. The degraded area now has less biodiversity and is at a lower level of productivity or structure. The degraded site might degrade even further, perhaps because of recurrent fires (point C) or, if not disturbed further, might gradually recover, unaided, back to point A. The circumstances under which this recovery could occur are described further below.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00121-0

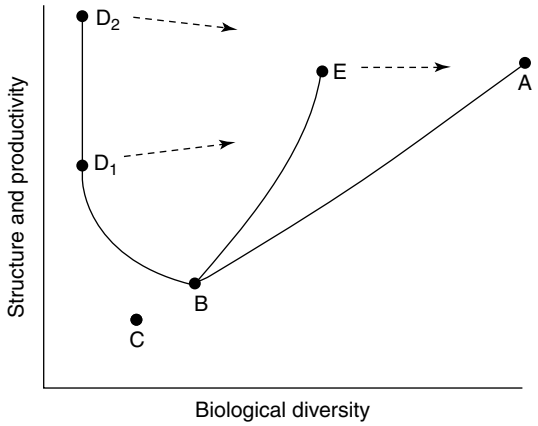


Figure 1 Various methods of reforestation after degradation. Degradation has moved the system from its original state at point A to a degraded state at point B. Over time, the site can degrade further to C or recover naturally back to A. Reforestation via plantation monocultures can lead the system to D1 or D2, depending on the productivity achieved. Rehabilitation may recover most of the structure and productivity but only part of the former biodiversity (E). Over time, some of the original species may recolonize the forests at D and E from intact forest, causing them to drift back toward point A. (After Bradshaw, 1995. Reprinted from Cairns J (ed.) (1995) *Rehabilitating Damaged Ecosystems*. Lewis Publishers.), with permission from CRC Press.

The process of recovery can be accelerated by reforestation but there are several ways in which this might be done. Each leads to a different outcome. *Restoration* (or, more properly *Ecological Restoration*) describes the process in which the aim of reforestation is to restore the ecosystem back to its former condition (point A) containing the original complement of plant and animal species. At this point the ecosystem will have regained its original structure and productivity and the ecological processes that sustained the system will have been reestablished. Reforestation using *plantation monocultures*, however, attempts to restore structure and productivity but not the original biodiversity (point D). Such reforestation is commonly undertaken for commercial reasons and often uses exotic tree species. The structure and productivity may be partially regained (point D1) or even exceed that of the original ecosystem because of the use of fertilizers or other management inputs (point D2).

Between these two approaches lies a midway position that might best be described as *rehabilitation*. This may recover the original structure and productivity but may not recover all of the original biodiversity (point E). The new ecosystem may contain a mix of native and exotic species. Over time, the new systems at D and E may gradually drift toward point A if some of the original species are able to recolonize from intact forest nearby or they may remain in a partially restored state if no such recolonization takes place.

These three forms of reforestation are a necessary simplification of the variety of methods of reforestation that might take place on degraded lands. They differ in the degree to which biodiversity is recovered but they also share certain common attributes. These include the fact that a new stable and productive land use is achieved and that at least some of the ecological services and protective functions of the original forest have been recovered.

Can Restoration Ever be Achieved?

Some have questioned if reforestation can ever achieve complete ecological restoration and recover all the biodiversity once present at a degraded site (Lamb, 2011). There are several problems. One is that the term restoration implies that the original species complement at a site is known and thus that there is a clear target. The second is the assumption that this community would have persisted over time when, in fact, the target may have drifted because of the changes in climatic or other environmental conditions. The third problem is that the process of degradation may have caused certain of the original species to go extinct whereas some exotic species became naturalized. These difficulties mean that restoration is sometimes an uncertain target that might be difficult to achieve in practice. For these and other reasons, the Society for Ecological Restoration defines ecological restoration as

> the process of assisting the recovery and management of ecological integrity. Ecological integrity includes a critical range of variability in biodiversity, ecological processes and structures, regional and historical context, and sustainable cultural practices.

Ecological restoration may also be difficult for other reasons. In some severely degraded sites the numbers of species to be restored may be too high or the magnitude of the environmental changes (e.g., to soil fertility or salinity) may be so immense that restoration is too difficult to achieve even if the technical means were available. In such cases the financial costs of attempting to fully restore the original system might simply be too high. In some situations social constraints may also apply. For example, some landowners or managers may be unwilling to agree with reforestation of degraded sites because their opportunity costs are too high and the restoration of biodiversity conservation is not a goal they share. However, such persons may be quite willing to undertake other forms of reforestation that generate economic benefits and, coincidentally, some conservation benefits.

Methods of Reforestation

Reforestation can take place in a variety of ways. In some situations degraded forest ecosystems can regenerate without human intervention. A well-known example is the reforestation that has occurred over the past 100 years in much of the north-eastern US following the abandoning of some of the land previously cleared for agriculture. The former farm sites are now occupied by deciduous woodlands with a structure and diversity probably similar to that which was originally present. This same natural recovery process can be found in many other temperate and tropical forests (Stanturf and Madsen, 2005; Lamb, 2011). The rate at which such successions take place varies widely but most are usually slow. Natural recovery after degradation is not an inevitable process and it only takes place under certain conditions. In all cases

the disturbing agents that were the causes of degradation must have been removed and the original topsoils must have remained more or less intact. Further, remnant communities of the original forest species must have persisted nearby in the landscape to act as a source of propagules and colonists for the new succession.

A particular problem with such slow recovery processes is that it prolongs the risk that sites might get damaged again by disturbing agents such as fire or grazing. If this occurs the site can be degraded once more (taking it back once more to point B or C on **Figure 1**). Another problem with natural recovery is that it might not be recognized as being underway. In these circumstances the risk is that the site may be seen as abandoned wasteland and open to use by newcomers. For these reasons, more active reforestation programs may sometimes be needed to accelerate the process and overcome these risks. These direct reforestation programs can take a variety of forms but all require that the causes of degradation (fire, grazing, invasive species, firewood collection, etc.) be removed and that sites be actively protected. The subsequent reforestation process can then be carried out by direct seeding or by planting seedlings.

Direct Seeding

The advantages of direct seeding are that a costly nursery stage is avoided and seeds can be broadcasted in the field comparatively cheaply. Seeds can be sown by hand or from a vehicle or can be distributed from the air and large areas can be covered relatively quickly (**Figure 2**). The disadvantage of direct seeding is that only a small proportion of the seed distributed is likely to reach the seedling stage as many seeds are removed by animals or insects, fail to germinate under field conditions or are out-competed by weeds. This means that direct seeding usually requires large volumes of seeds to achieve a desired tree density and it may be impossible to acquire such volumes for many less-common species. Direct seeding has often been used at mine sites where topsoil has been respread after mining has ceased and where competing weeds have not yet re-established. Direct seeding is also possible in areas where competing weeds can be removed by plowing or the use of weedicides. In such circumstances a cocktail of species might be reintroduced as seeds although the diversity of species in such a mix can be limited by the availability of seeds of less-common species and these might have to be re-established at the site as seedlings.

Planting Seedlings

Planting of seedlings raised from seed (or cuttings) in a nursery is a more expensive operation but it may be necessary for species with very small seeds or particular germination or mycorrhizal requirements. It also necessitates other management inputs such as site preparation (ripping or plowing), weed control, or fertilizing which improve survival and foster more rapid growth of the planted seedlings. In plantation monocultures being established for commercial reasons, it is common to plant trees in regularly spaced rows. This also permits the use of planting machines pulled behind tractors which can lead to very rapid planting rates. In a restoration operation a more irregular planting design might be preferred because it leads to a more random distribution of species.



Figure 2 Forest establishment after direct sowing of seeds from the air at a bauxite mine in northern Australia. Some supplementary plantings of certain species can also be done where seeds are difficult to collect or where germination and establishment of these species in the field is poor, with permission from P. Erskine.

The Choice of Species

A key issue in all reforestation is the identity of the species to use. If the goal of reforestation is commercial timber production, then a fast-growing, exotic species might be chosen. Such species usually have readily available seed, well-established nursery and planting methods, and defined silvicultural prescriptions to govern their management. If, however, ecological restoration of a former ecosystem is the goal, then only native species will be used. But this can be difficult to achieve. In many cases it may not be possible to acquire the seed of all of the original species or to raise large numbers of seedlings of those needed. In many cases environmental or socio-economic constraints may mean that ecological restoration is not possible and that some form of rehabilitation must be adopted. In that case it may be useful to use species that provide both economic and conservation benefits. These may include native and exotic species. Some of the types of species that might be used are shown in **Table 1**. These include economically useful species, those able to tolerate or modify the new site conditions, those attractive to particular wildlife or those not able to reach the site through normal dispersal mechanisms.

Table 1 Plant species that might be chosen for reforestation programs

Type of species	*Reason*
Native species	To restore the original communities once present at the site; where possibly the seed of these species should be collected from local sources
Exotic species	Since only exotics can tolerate the conditions now present at the degraded site or because they are fast growing and provide a commercial return
Commercially useful species	For example, providing timber, fruit, oils, and nuts; a financial return is needed to justify funding of reforestation
Traditionally or culturally useful species	For example, those providing fuel, food, or medicines used by local communities
Tolerant species	Able to tolerate the adverse conditions likely to be found at degraded sites such as low soil fertility, exposed windy conditions, soil salinity, or occasional fires; in some cases, only exotic species may have the required attributes
Species with fleshy fruit	To attract seed-dispersing frugivores to the site
Poorly dispersed species	Species with large fruit or those with high specialized animal dispersers unable to tolerate the habitats now present at the degraded site
Nitrogen-fixing species	To improve soil nitrogen levels
Fast-growing species	To rapidly occupy the site, shade out weeds, and improve microclimatic conditions
Understory species	To provide the structural complexity and habitat conditions required by some wildlife species
Multipurpose species	Trees that fulfill more than one requirement such as providing timber and food
Special species	Species required for some special purpose; e.g., rare or endangered trees requiring enlarged populations, ''keystone'' species producing food required by wildlife at certain times of the year, etc.

Forestry is a long-term business and management objectives for a particular forest can change over time. In some areas plantations established for timber production have acquired a greater value as recreation or conservation reserves. In these circumstances it may be useful to manipulate the composition of the forest to favor these new management objectives (*see* Case Study 1).

Planting Densities and Species Richness

The choice of planting density and the number of species planted will be determined by the objectives of reforestation and the resources available.

Planting Densities

Commercial tree growing in well-watered temperate or tropical locations usually requires moderately close initial spacings (perhaps 1100 trees per hectare) to exclude weeds and ensure rapid site occupancy. Close spacings also reduce side branching and promote rapid height growth. Where ecological restoration is the objective, close spacings might also be used. Again, the dense planting excludes weeds and ensures rapid site occupancy. But high planting densities are costly and many of these seedlings may die as competition develops. In drier areas it may be preferable to use lower density plantings to ensure adequate water is available for each tree as it develops.

Low-density plantings may also be useful if large areas of degraded land must be restored but resources are limited. Observations of isolated trees in vacant fields suggest these trees often attract seed-dispersing birds. The birds frequently deposit seeds around the base of the tree that germinate and grow into a cluster of seedlings around each tree. Planting isolated trees or tree clusters to act as bird perches could simulate this process and foster successional development over large areas, albeit at the expense of much slower rates of recovery.

Species Richness

The numbers of species that should be included in these plantings can vary. As noted above, commercial plantings established by industrial plantation owners generally use a single species. Reforestation carried out to support biodiversity conservation usually involves the use of a much large number of species. The actual number depends on circumstances. In landscapes where large numbers of forest remnants remain close by (i.e., the dispersal distances are short), only a modest number of species might need to be planted to initiate a succession because natural colonization by other species will take place over time and enrich the planted area. In these circumstances a dense planting of mostly short-lived early colonizer species may be sufficient to accelerate the successional process and only a small number of species might need to be used. Likewise, if the number of canopy species in the natural forest is low, it may be possible to replant most of these species and rely on successional development to restore understory diversity (*see* Case Study 2). In both cases the recovery process is initiated by planting but is completed by natural successional processes. Interestingly, the same colonization process can occur within simple plantation monocultures (**Figure 3**). The value of such plantations as wildlife habitats is likely to be substantially enhanced by this understory.

A different approach is needed where remnant vegetation is not close by or seed dispersal might be limited due to the absence of appropriate wildlife species able to carry seed to the deforested site. In such cases a much higher proportion of the total plant diversity may need to be brought to the site to achieve any degree of biodiversity within a useful time period. Under these circumstances ecological restoration might be difficult to achieve.



Figure 3 Native tree seedling colonization in the understory of a tropical hardwood plantation monoculture in northern Australia. The plantation is within 200 m of natural forest and its presence has catalyzed the recovery of substantial plant biodiversity at the site, with permission from J. Frin.

A common difficulty in all of these circumstances is that seed dispersers such as birds may distribute weed species as well as native species. All restoration programs, but particularly those relying on successional development to achieve ecosystem diversity, may therefore need substantial weed control before full site occupancy by the original species is achieved. Even then, tending may be needed for some years until canopy cover is achieved and the new community is less prone to colonization by weeds.

Ecosystem Recovery After Reforestation

Random plantings of a variety of trees will not necessarily restore the ecosystem that was once present at a site. These species must be able to reproduce and regenerate and the wildlife community must be reintroduced as well to re-establish trophic networks. Some plant species may require particular soil or microclimatic conditions to germinate and grow and will not become established from seed dispersed into a newly forested area until these conditions are met. Other species may need particular regeneration niches to reproduce after they have been planted at a site. For Species A to successfully colonize or reproduce at a site, Species B may need to be established first. In most ecosystems such rules of assembly are only poorly understood. What is clear, however, is that many animal species (including insects) are part of this development process because of the mutualistic partnerships they often form with plants that enable processes such as pollination and seed dispersal. Species A (a plant) may require Species B (a bird) for pollination and that, in turn, also needs Species C and D (plants flowering and supplying food at other times of the year). Ecosystem restoration is dependent on the formation of these complex trophic networks and interactions.

It is commonly assumed that animals will colonize new plant communities once successional processes have created appropriate habitat conditions. Like plants, animals may be classified as early colonists, tolerant of a broad range of habitat conditions, or later colonists that have more specialized habitat requirements that only develop when the succession has reached a relatively mature phase. A commercial plantation established using a single exotic tree species is likely to be much less attractive to many wildlife than a spatially and structurally complex new forest of native tree and undergrowth species. Likewise, the structural features of an old-growth forest such as thick litter layers, hollow trees, or rotting logs on the forest floor may take years to develop following restoration. These might need to be artificially simulated to foster the re-entry of particular wildlife (e.g., by dragging logs into a restored area or using nest boxes).

Wildlife from early successional stages is often able to recolonize restored forests relatively quickly. But two factors can limit the extent to which recolonization occurs more generally. Some wildlife, especially those with habitats in mature or old-growth forests, are unable to cross degraded landscapes lacking trees. Restored forests not linked with natural forest remnants may therefore be permanently deprived of such animal species unless some means can be found to overcome the blockage. A second limiting factor is that small forest remnants in degraded landscapes often contain only a subset of the original species formerly present. Vertebrates appear to disappear more rapidly from fragments than do plants whereas top-carnivores and large-bodied animal species in particular have large spatial requirements, making them especially sensitive to forest fragmentation. This means that recolonization may be impossible unless these species are brought in from more distant geographic locations. Wildlife recolonization may also be limited by predatory behavior or competitive relations between various animal species. Naturalized exotic predators such as foxes or cats represent a particular problem.

Many threatened or endangered wildlife species will require special habitat conditions. Where the threat to these species is from naturalized exotic predators, these new habitats might be most easily restored on offshore islands, which ensures the exotics are excluded (that is the isolation effect noted above becomes an advantage). This is considered in Case Study 3. However, a dilemma can then emerge over whether the specialized habits required by the endangered species means other species are discriminated against. Is ecological restoration the goal or is a more specialized and restrictive objective such as maximizing the population of a particular species through some form of rehabilitation more appropriate? Answers to such questions obviously depend on local circumstances.

All natural ecosystems are subject to disturbances and the well-known intermediate disturbance hypothesis suggests

ecosystem diversity reaches an optimum when the frequency of disturbances reaches some intermediate level. This means that the initial protection offered to reforestation projects to prevent further degradation must be relaxed at some stage to enable normal ecosystem processes to occur. There are few guidelines to describe how this should be carried out. The disturbance regime that produced degradation is different from that fostering diversity but the nature of these differences is rarely understood. For example, at what point can fire be allowed to occur? Indeed, when should fire be actively encouraged? And how intense should this fire be? These issues will not be important in reforestation projects where enhanced productivity is the objective but might be crucial to the ultimate success of many restoration projects.

Choosing Areas to Reforest

Reforestation is expensive and takes time. It is rarely possible to quickly reforest large areas at once and strategic planning is needed to identify priority areas where the early benefits from reforestation can be maximized. When reforestation is being carried out for commercial reasons, the location of a reforestation area will be dictated by markets and soils. The best locations will be those where transport conditions are shortest and where productivity is highest.

When the purpose of reforestation is to restore biodiversity to a degraded site, the choice is more complex. Possible locations are outlined in **Table 2**. Common targets are degraded areas within forested nature reserves or the buffer strips around such reserves. In these areas the task is to restore the ecological integrity of a particular site and hence that of the larger reserve area as well. Since colonization distances are likely to be short, there is a high probability of success if disturbing agents can be removed and weeds can be eradicated.

Another common target is to establish corridors linking forest fragments still remaining in a formerly forested landscape. The purpose of this is to foster connectivity between forest fragments as well as to increase the area of habitats for species in these landscapes. Some have argued that narrow corridors are more likely to be the areas of increased predation rather than a means by which wildlife species movement across the landscape and genetic exchange are enhanced. But others have argued that any increase in spatial heterogeneity should be beneficial and the advantages would seem to outweigh any disadvantages provided corridor widths are not too narrow.

Key targets for many restoration projects are the habitats of particular species such as those whose populations have declined or are endangered by other predatory species. This presumes, of course, that these habitat requirements are known. In the latter case, offshore islands have been especially useful areas to use because they are of sufficient size to contemplate eradicating such predators and to be assured that recolonization will not occur. Encouraging results have been reported from New Zealand (*see* Case Study 3).

Many reforestation programs are carried out because of the improvements in ecological services rather than because they restore biodiversity. For example, stream-sides are often degraded by grazing cattle that cause bank erosion and slumping. Reforestation can stabilize these areas and the vegetation can act as a filter by trapping topsoil erosion. Such reforestation can also have conservation benefits, of course, since these areas are often the focal points within drier landscapes for much wildlife activity. Reforestation of upper or lower watershed areas may be necessary to improve water logging or salinity and, again, indirectly improve landscape biodiversity. Some of the most innovative and impressive restoration has been carried out at mine sites. These areas are intensively degraded and can cause serious pollution problems (e.g., heavy metals, acid drainage, salinity) in large areas of the catchment below the mine site even though the size of the actual mine itself may be small. Mine site restoration usually requires careful stockpiling and respreading of topsoils but the large expenses of doing this are commonly only a small percentage of total mine expenses (*see* Case Study 4 *and* Case Study 5).

Specialized Conditions or Sites with Particular Problems

Some sites have particular problems. This might be because of the degree of degradation suffered (compacted or infertile soils, saline areas, waterlogged sites) or because of the stressful environmental conditions present (e.g., unstable soils, windswept, or cold mountain areas, desert areas). A variety of techniques

Table 2 Priority areas for reforestation

Location	*Reason*
Degraded areas in or around nature reserves	To remove weeds and restore ecological integrity to the reserve
Corridors between forest fragments or reserves	To improve connectivity between habitats and metapopulations
Habitats of endangered or threatened species	To increase the populations of these species
Offshore islands	Because islands may be areas free from naturalized weeds or pests that would otherwise preclude restoration on a mainland
Riparian areas	To protect streamsides and because riparian areas are often centers of wildlife activity in many drier landscapes
Upper watersheds	To stabilize catchments, prevent erosion, and improve downstream water quality
Lower watersheds	To lower water tables or reduce salinity
Mine sites or other severely degraded areas	To overcome the extreme degradation that may have occurred and to prevent these areas from becoming erosion sources or sources of toxic leachates
General landscape matrix between forest fragments	To increase landscape heterogeneity and to provide "stepping stones" between natural forest remnants

have been developed to deal with many of these conditions, including ripping and plowing, fertilizing, using drains or mounds to allow more tolerant species to become established at waterlogged or saline sites, etc. In some cases mycorrhiza and soil fauna such as earthworms must be reintroduced as well as nitrogen-fixing plants to promote nutrient cycling. Sites in extreme environments such as mountainous or exposed areas are particularly difficult. Sometimes tolerant exotic species can be planted to facilitate the later establishment of native species. But the survival rates of planted seedlings in especially severe environments can be very low. This raises the question of whether it is worth even attempting to restore severely degraded ecosystems or whether a better return on the resources to be invested might come from restoring only moderately degraded landscapes. In such cases a balance must be struck between what is biologically possible and what is economically feasible.

Socioeconomic Issues

Ecological restoration can be difficult because it requires detailed knowledge of the autecology of a number of species and of the processes of successional development. But the economic and social issues confronting many reforestation projects are equally complex. One of the most difficult to overcome is the financial cost of reforestation. Not only is reforestation expensive but it has an opportunity cost. Some landholders will be willing to undertake reforestation to conserve biodiversity but most will need the financial benefits of reforestation to exceed the direct and indirect costs of carrying it out (*see* Case Study 6). An additional problem is that, unlike most agricultural crops, tree growing takes many years to produce a financial return whereas most of the costs occur in the first few years. These cost issues will dissuade many landowners from undertaking reforestation.

A second issue concerns the type of reforestation carried out. Many landholders will use simple monocultures because these are the easiest to manage but monocultures usually offer only modest biodiversity conservation benefits. The task is to find methods of reforestation that generate financial benefits for landholders as well as conservation benefits. There are, in fact, several possibilities. These include surrounding monocultural plantations with species-rich buffer strips or planting different species in different parts of the landscape according to the environmental conditions at each site thereby creating a heterogeneous landscape mosaic. Another possibility is to establish polycultures that include a number of tree or understory species (*see* Case Study 2). These can be financially profitable as well as creating structurally complex habitats suitable for many wildlife species (Lamb, 1998, Lamb, 2011).

Socioeconomic problems increase when reforestation is attempted at a landscape scale. Most conservation benefits, including biodiversity conservation, are only generated when reforestation is carried out at this scale. But reforestation at a landscape scale usually requires dealing with a large number of stakeholders. The most obvious are landowners, leaseholders, or people who have some traditional claim to the land (although that claim might not be recognized by the state). Others may be downstream water-users, conservationists or simply the general public. As noted earlier, these various groups may view degradation and the desirability of reforestation in quite different ways. This means those proposing to undertake reforestation on a large scale must resolve (1) where reforestation is to be done, (2) how much is to be carried out, and (3) what types of reforestation are to be used at particular sites. Experience in many places suggests this is best done by involving all stakeholders in the planning and management process, including those in a weak bargaining position and those with little political power. In this way, all interests can be protected from the beginning. Some form of educational program may also be necessary to allow the various stakeholders to see the reasons for the project and the benefits it may have for themselves (e.g., perhaps in employment opportunities) as well as for the community as a whole. Likewise, some form of incentive or compensation may be needed to persuade the participation of landholders controlling in critical areas within the landscape. One way in which this might be done is to pay for the environmental services generated by reforestation (e.g., watershed protection, carbon sequestration, biodiversity protection). If this collaborative planning or various forms of compensation do not occur there may be a high risk of continued disturbances and further degradation.

Conclusion

Although, deforestation continues globally, reforestation of degraded lands and forests can help protect biodiversity and restore it to landscapes from where certain species have disappeared. There is a wider variety of silvicultural techniques available than is generally supposed although more work is needed to develop techniques that generate both financial and conservation outcomes. One of the key tasks is to find ways of making these options attractive to landholders at particular sites and to coordinate reforestation activities across degraded landscapes to foster biodiversity conservation. Although the survival of many wildlife species will be assisted by reforestation some species may not benefit, at least in the short term. These include forest-dependent fauna with specialized habitat requirements and large home ranges. Such species will continue to depend largely on natural forests although reforestation within and around such areas will be beneficial.

Case Study 1: Forest Restoration by Transforming Monospecific Conifer Plantations in the UK

Not all forms of reforestation involve re-establishing trees on largely deforested land. In parts of the UK reforestation now also involves converting conifer forests containing a single dominant tree species to more diverse species-rich forests. Much of the reforestation effort during the twentieth century in the UK was aimed at establishing timber plantations of fast-growing commercially valuable pine or spruce species. Many of these species were exotics such as Sitka spruce or Norway spruce. In recent years managers are seeking to transform these structurally simple forests to more complex mixtures of softwood and native broad-leaved species that have a greater biodiversity and conservation value. Where landholders are interested, the first step, in the process is to identify the

appropriate type of native forest that might be grown at the site. This is usually based on surveys of any natural woodlands in the surrounding landscape together with historical records. The second step is to identify those sites with the greatest opportunities for conversion. These might be those where the existing forest already has a good representation of specialist woodland ground flora, where there are existing patches of seminatural forest nearby able to act as a source of colonists or sites where the location and size of the new forest will enable it to make a contribution to regional conservation goals.

The preferred method of conversion is to rely on natural regeneration by native species and this has been favored by manipulating the existing tree canopy. In some cases this has involved line thinnings and in other cases by clear-felling small patches. Thinning is expensive and the extent to which this is being carried out often depends on the availability of markets for the timber being produced. The native species able to take advantage of the improved canopy gaps have been either present in the forest understory or have colonized from nearby woodlands.

The objective of reforestation in the original commercial timber plantations in the UK was to create a harvestable timber resource. By contrast, the objective of this particular reforestation program is to gradually convert the plantations of largely exotic conifer species to deciduous broadleaved or mixtures of broadleaved and native conifer forests that conserve both local and regional plant and animal biodiversity. Results so far have been encouraging (Harmer *et al.*, 2005).

Case Study 2: Mixed-Species Plantings in North Queensland, Australia

Several types of reforestation have been carried out on cleared land formerly occupied by rainforest in northern Australia. The government Forestry Department initially established monocultures of various species including native high-value timber trees. The normal rotation of these was around 40 years although some have been retained for up to 60 years. Over time many of these have been colonized by species from nearby natural forests transforming these simple plantations into very diverse forest communities (**Figure 3**). In other cases farmers have established mixed-species plantings of mainly native high-value timber trees with the objective of generating a future cash income while also contributing to region conservation goals. Most of these have used less than ten tree species. A third approach has been to reforest certain areas with the objective of restoring former forest ecosystems without attempting to create a timber resource. These have used an average of 50 tree species. Most of these plantings, with the exception of the old plantation monocultures, are less than 20years old and are yet to develop habitats suitable for wildlife with specialized habitat requirements. Nonetheless, all have acquired high levels of plant and animal diversity. The extent of this diversity and the types of species present has depended on the types of plantings used, their age, and their location in the landscape relative to the remaining areas of natural forest. Not surprisingly, the species-rich plantings established to restore the natural forest have generated faster rates of recovery than the mixed-species timber plantations. Perhaps more surprising has been the very high levels of diversity found in the older monocultures with species-area curves showing more than 30 plant species present within 300 m^2 (Keenan *et al.*, 1997; Erskine *et al.*, 2007).

Case Study 3: Restoration of Habitats of Endangered Biota in New Zealand

Sometimes specific forms of reforestation are needed to conserve particular biota. This is the case on the island (or minicontinent) of New Zealand which has a unique biota but many of which have become endangered since the arrival of humans. In the case of birds, for example, 48% of species have been lost in the past 200 years because of the introduction of predatory and browsing mammals. Even in protected areas, attempts to maintain populations of some presently endangered species are difficult because of the difficulty in controlling predatory species over large areas. New Zealand has many offshore islands that were once part of the mainland (and consequently contain many mainland species). Although a number have been deforested or degraded, they are potential safe havens for endangered species since it may be feasible to eradicate predators and browsers over these smaller and more isolated locations. This has fostered a shift in attitude from trying to simply preserve species to restoring their preferred habitats.

The first step has been to mount a vigorous campaign to eradicate exotic species such as cattle, goats, rodents, and cats from the islands. Although the larger grazing animals are comparatively easy to remove, thereby enabling forest regeneration, it has only been in more recent years that techniques for eradicating rodents have become successful. In many situations this alone has led to improvements in habitat quality. The second step, on islands where deforestation has been severe, has been to replant trees and shrubs to re-establish appropriate habitat conditions. Where populations of target species are already present, these two steps have often been sufficient to enable them to survive and flourish. In other cases it has been necessary to translocate animals from the mainland to establish new breeding populations. The success of the program has been followed by similar attempts to restore biodiversity to islands of forest on the New Zealand mainland. These are the patches of degraded forest land within largely agricultural or otherwise severely disturbed landscapes.

The objective of these programs has been quite specific; it is not to necessarily restore the original forest ecosystems but rather to seek to re-establish sufficient of the original ecosystems to enable the survival of nominated species or communities. This goal acknowledges that restoration may never be complete because some exotic organism may always be present or because knowledge of the original system may always be incomplete. In the New Zealand context, however, the goal is a way of ensuring the protection of a significant proportion of the country's biological heritage (Towns and Ballantine, 1993; Saunders and Norton, 2001).

Case Study 4: Reforestation after Mining Tropical Forest Lands in Brazil

Mining causes complete degradation of an ecosystem and perhaps represents one of the greatest challenges to those interested in restoration, particularly if the site was previously occupied by a biologically rich tropical rain forest. The Trombetas bauxite mine in Amazonia has an annual reforestation program of 100 ha. Following mining, the sites are leveled and stockpiled topsoil is spread back over the area to a depth of 15 cm. The site is then ripped and replanted using seed, stumps, and seedlings. About 70 species are planted at a density of around 2500 trees per hectare.

Studies carried out in an area that had been replanted 10 years earlier found forest development had been rapid. A number of plant species had colonized the site from the adjoining rain forest. After 10 years these represented 60% of the total number of woody species present and the site contained around 50% of the species present in the adjoining, undisturbed forest. Some areas had rather sparser forest cover than others and the colonization process was less vigorous in these, most probably because of the presence of grasses in the understory. The density and composition of the understory developing beneath the planted trees were related to the overstory basal area and topsoil depth but negatively related to the distance from the undisturbed forest. Primary forest species were able to colonize up to 640 m although most colonists had smaller diaspores. Most of the colonists in the more open sites with grass were early pioneer species.

Surveys of wildlife found low densities of birds but high densities of bats. Most of the bird species were representative of secondary forests and are unlikely to have carried many seeds from the undisturbed forest nearby. However, the bats included several important seed dispersers which feed on a variety of tree species and are able to carry seed over large distances. A number of other mammal species (e.g., opossums, deer, agouti, and armadillo) were also observed within the study site.

It appears that a dynamic and sustainable successional environment has been established in a much shorter time than it would otherwise take to occur. This was due to the respreading of topsoil which contained a number of important tree species and the choice of trees planted in the original planting program. The program was less successful where tree survival or growth allowed grasses to become established and at these sites the successional process may be rather longer (Parrotta *et al.*, 1997).

Case Study 5: Reforestation after Mining in Temperate Forest Lands in Australia

The *Eucalyptus marginata* forests in the south west of Western Australia are extremely diverse and the region contains around 780 plant species. Much of this diversity is contained within the understory plants rather than in the tree overstory. The area has a Mediterranean climate with hot dry summers and cool wet winters. Mining started in the area in the 1960s. As in Brazil, bauxite is extracted using open cut mining. Before mining, the topsoil and overburden are removed in two layers (0–15 cm and 15–80 cm) and stored. After mining, the topsoil is replaced in the same sequence and the site is deep-ripped to break up the compacted mine pit floor. Logs and rocks are also added to provide fauna habitat. Replacing the topsoil reintroduces some species having dormant seed stored in the topsoil. Other plant species are re-established using a combination of direct seeding and by planting seedlings. Those reintroduced as seedlings are species with scarce seed or with seed that is difficult to germinate and which need special propagation methods. Most of these species are introduced at more or less the same time. Of the 140 species usually found at any particular site about 28 species are reintroduced in topsoils, 78–113 species are brought back using seed and the remainder are planted as seedlings. Sufficient seed or seedlings are used to produce a tree density of around 1300 trees per hectare. No attempt is made to reintroduce wildlife since these are able to colonize from nearby intact forest that was not disturbed by mining.

These reforestation methods have been developed and refined over a long period and there is now evidence that they generate productive and structurally diverse forests containing most of the original plant and animal diversity. The species restored include herbivores, detritivores, nectivores, insectivores, and carnivores although some wildlife species are still limited by the absence of habitat features such as hollow-bearing trees or rotting logs that will take some time to develop. The new communities are adapted to the wildfires that normally occur in these forests and fire-adapted species are able to regenerate naturally (Koch and Hobbs, 2007).

Case Study 6: Reforestation of Degraded Lands in Nepal

Any form of reforestation must be compatible with the needs of landowners. In Nepal, forests are an integral part of the farming system with leaves and grass being used for fodder and animal bedding. Such material is also used in compost to maintain productivity of agricultural lands. Timber is regularly harvested for fuel.

Some parts of the middle hills of Nepal have been highly degraded over time because of intensive harvesting. However, externally sponsored attempts to encourage land users to reforest have not always been successful. If there is forest near villages, there is commonly little interest in forest protection or tree planting although indigenous management systems (primarily to define user access rights) for these forests may exist. By contrast, where there is a severe shortage of forest products and any remaining forests are more distant, there is likely to be a genuine interest in forest development activities and well-developed indigenous management systems for the remaining forests. These systems are likely to have biological objectives (e.g., rotational cutting) as well as simply defining user access rights. That is, interest in reforestation depends on the perceived need of land users and not on an external perception that forest degradation or a loss of biodiversity has occurred.

Indigenous management systems can foster reforestation under certain conditions and some areas of private farmland in Nepal now have more forest cover than in earlier years.

However, this does not invariably occur. For example, larger trees may be protected but grazing beneath these may lead to a failure of regeneration. Likewise cutting of smaller trees for firewood may be tolerated whereas the larger trees are saved. A particularly difficult situation can be where a cash market for forest products develops. In such cases poor people can succumb to the temptation to cut firewood, often at night, to sell the next day. Indigenous management systems must be very strong to resist such pressures.

Tree planting has been carried out at particularly degraded sites. The most successful reforestation has been with pines, although broad-leaved species are more valued by villagers as leaf fodder. Manipulative trials with young pine plantations have found that many broad-leaved species can also then regenerate from suppressed stumps and root stocks or seed once the sites are protected. This natural successional development increases the silvicultural options available, giving local communities greater flexibility in determining their preferred forest management strategy. It is also possible to thin the young pine stand and enrich it using other broad-leaved species that might otherwise be difficult to establish. This, too, increases the options available. Reforestation of degraded sites and the development of indigenous management systems may not always take place without some outside intervention. Once benefits start to accrue to land users, however, neighbors are often quick to pick up the ideas and follow the example. At this point it is important that government agencies who might legally own the land are cautious in imposing rigid bureaucratic controls within which the new forests must fit. For example, allowing someone from outside the user group to harvest resources from the new forest could easily unleash a surge of unregulated harvesting, leading to a renewed degradation (Gilmour, 1990; Gilmour *et al.*, 1990).

See also: Agriculture, Sustainable. Agrobiodiversity. Biodiversity in Logged and Managed Forests. Biodiversity-Friendly Farming. Biofuels and Biodiversity, Wildlife Habitat Restoration. Crop Mixtures and the Mechanisms of Overyielding. Economics of Agrobiodiversity. Feeding the World and Protecting Biodiversity. Forest Ecology. Impact of Ecological Restoration on Ecosystem Services. Landscape Diversity. Poverty and Biodiversity. Restoration of Biodiversity, Overview. Timber Industry. Tropical Forest Regeneration

References

Erskine PD, Catterall CP, Lamb D, and Kanowski J (2007) Patterns and processes of old field reforestation in Australian rain forest Landscapes. In: Cramer V and Hobbs RJ (eds.) *Old Fields: Dynamics and Restoration of Abandoned Farmland*, pp. 119–144. Washington: Island Press.

FAO (2010) *Global Forest resources Assessment 2010: Main Report*. Food and Agriculture Organization of the United Nations FAO Forestry Paper 163.

Gilmour DA (1990) Resource availability and indigenous management systems in Nepal. *Society and Natural Resources* 3: 145–158.

Gilmour DA, King GC, Applegate GB, and Mohns B (1990) Silviculture of plantation forest in central Nepal to maximize community benefits. *Forest Ecology and Management* 32: 173–186.

Harmer R, Thompson R, and Humphrey J (2005) Great Britain – conifers to broadleaves. In: Stanturf JA and Madsen P (eds.) *Restoration of Boreal and Temperate Forests*, pp. 319–338. Boca Raton, FL: CRC Press.

Keenan R, Lamb D, Woldring O, Irvine T, and Jensen R (1997) Restoration of plant biodiversity beneath tropical tree plantations in northern Australia. *Forest Ecology and Management* 99: 117–131.

Koch JM and Hobbs R (2007) Synthesis: Is Alcoa successfully restoring a Jarrah forest ecosystem after bauxite mining in Western Australia? *Restoration Ecology* 15: S137–S144.

Lamb D (1998) Large-scale ecological restoration of degraded tropical forested lands: The potential role of timber plantations. *Restoration Ecology* 6: 271–279.

Lamb D (2011) *Regreening the Bare Hills: Tropical Forest Restoration in the Asia-Pacific Region*. Dordrecht: Springer.

Parrotta JA, Knowles OH, and Wunderlie JM (1997) Development of floristic diversity in 10 year old restoration forests on a bauxite mined site in Amazonia. *Forest Ecology and Management* 99: 21–42.

Saunders A and Norton DA (2001) Ecological restoration in Mainland Islands in New Zealand. *Biological Conservation* 99: 109–119.

Society for Ecological Restoration Home page http://www.ser.org.

Stanturf JA and Madsen P (eds.) (2005) *Restoration of Boreal and Temperate Forests*. Boca Raton, FL: CRC Press.

Towns DR and Ballantine WJ (1993) Conservation and restoration of New Zealand island ecosystems. *Trends in Ecology and Evolution* 8: 452–457.

Religious Traditions and Biodiversity

Fikret Berkes, University of Manitoba, Winnipeg, Canada

© 2013 Elsevier Inc. All rights reserved.

Glossary

Animism Belief in spiritual beings. The term is associated with the anthropology of E. B. Tylor, who described the origin of religion and primitive beliefs in terms of animism in *Primitive Culture* (1871). Tylor considered animism a minimum definition of religion and asserted that all religions, from the simplest to the most complex, involve some form of animism. From Latin *anima*, 'breath' or 'soul'.

Biocultural diversity The diversity of life in any of its manifestations, biological and cultural, which are interrelated (and likely coevolved) within a complex social–ecological adaptive system.

Ethics Codes that exert a palpable influence on human behavior. Embedded in worldviews, ethics provide models to emulate, goals to strive for, and norms by which to evaluate actual behavior.

Monotheism Belief in the unity of the Godhead, or in one God, as opposed to pantheism and polytheism. Monotheism is a firm tenet of Islam and Judaism. Christianity, with its concept of Trinity, alone among the three monotheistic religions, dilutes monotheism. From Greek *mono*, 'one', and Greek *theos*, 'god'.

Pantheism The doctrine that identifies the universe with God. In Western thought, the term is associated with the Dutch philosopher Baruch Spinoza. His view represents an important criticism of the 'orthodox' view of a god whose reality is somehow external to the reality of the world. From Greek *pan*, 'all', and Greek *theos*, 'god'.

Religion Human recognition of superhuman controlling power, and especially of a personal God or gods entitled to obedience and worship; the effect of such recognition on conduct or mental attitude.

Sacred natural sites Areas of land and water having special spiritual significance to people and communities.

Stewardship To hold something in trust for another, as in the Biblical (human) responsibility for husbanding God's gifts. In the present context, examples include Australian aboriginal 'looking after country', Andean Quechua 'caring for Mother Earth', and Canadian Ojibwa 'keeping the land'.

Traditional ecological knowledge A cumulative body of knowledge, practice, and belief, evolving by adaptive processes and handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and with their environment. It is a subset of indigenous knowledge, which is local knowledge held by indigenous people or local knowledge unique to a given culture or society.

Traditional societies Groups in which knowledge, practice, and belief are handed down through generations largely by cultural transmission. Tradition itself evolves by adaptive processes, but not all tradition is necessarily adaptive.

Worldview The larger conceptual complex in which ethics are embedded. A.N. Whitehead called it the conceptual order, or one's general way of conceiving the universe, which supplies the concepts by which one's observations of nature are invariably interpreted. In general, worldviews limit and inspire human behavior, shape observations, and perceptions. Arnold Toynbee's *Weltanschauung*.

The Context: Religion for Encoding Ethics

Different groups of people in different parts of the world perceive and value sacredness differently. Religion is a general term that has become established in the Western world for a whole class of cultural codes and rituals. In the West, there often is little connection between a person's religion and environmental ethics. In many other societies, religion is part of a way of life as well as a worldview. For example, in Chinese there is no one word that translates as 'religion' in the Western sense. In some traditional societies, including many Amerindian groups, religion, worldview, and environmental ethics and practice are inseparable. It is appropriate, therefore, to define religion in its broader sense of 'human recognition of superhuman controlling power', inclusive of both monotheistic and pantheistic traditions.

Religions provide a central organizing myth and include cultural symbols for a moral code. Conceived that way, religions can be thought to include a wider variety of beliefs, within the definition of a superhuman controlling power. For example, one can refer to the 'religion of the market'. Peter Timmerman uses the term 'econotheism' as the adoption of market economics "to bring the rest of the earth under new management". Econotheism, argues Timmerman, "contains within it all the elements of a religion: a priesthood, scriptures, catechisms (Econ 100 textbooks), an explanation of happiness and misery, a way to get to heaven, and a core model of the human". Eugene Anderson (1996) holds that religion is best regarded as something providing an emotionally powerful way to 'sell' a moral code. The content of the moral code is negotiable and variable; it is not in itself a part of the religion. One can have a moral code without a religion, as in secular humanism, but at the risk of losing the emotional content supplied by religion.

Religion has not been a particularly popular topic among those involved in biodiversity conservation. The index of the 1140-page *Global Biodiversity Assessment* of 1997 contains a grand total of only two minor entries on religion. One problem with religion is that it has often been bent to serve all possible ends, including those destructive of biodiversity and cultural diversity. The role of religion in biodiversity conservation, and more generally in environmental ethics, is controversial. However, in more recent years, the profile of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00122-2

religion, as relevant to environmental sustainability, has increased. Two volumes (Taylor, 2005; Jenkins, 2010) provide details on the roles of various religious traditions, along with a journal devoted to this area (*Journal of the Study of Religion, Nature & Culture*). As well, a number of volumes examine spiritual traditions and sacredness in various cultures (Posey, 1999; Turner, 2005; Shaw and Francis, 2008; Berkes, 2012; Verschuuren et al., 2010).

There are several cautions in interpreting the discussion on religion and biodiversity. First, generalizations are always risky. As Donald Worster (1988) put it in *The Ends of the Earth*, "not a few scholars have fallen into the trap of speaking of 'the Buddhist view of nature' or 'the Christian view' or 'the American Indian view', as though people in those cultures were all simple-minded, uncomplicated, unanimous, and totally lacking in ambivalence. Every culture, we should assume, has within it a range of perceptions and values, and no culture has ever really wanted to live in total harmony with its surroundings".

Second, there is invariably a gap between 'the ideal' and 'the actual' in making sense of how societies deal with biodiversity. Ethics do not describe how people actually behave but rather set out how people *ought* to behave. There often is a discrepancy between belief and practice. Some scholars argue that one could spend too much time on values and beliefs, neglecting to examine behavior or the actual effects of these values on human practices regarding the environment. The field of environmental history studies the historical record of changes in the landscape as a way of evaluating the actual effects of the belief-and-practice complex of a culture.

A third caution is the intangible nature of traditions and their record in the literature. Not only do traditions change and adapt all the time, but the literature is based largely on outsiders' accounts of various religious traditions, especially those that do not have a written record. Since traditional societies are, by definition, those cultures in which much knowledge is transmitted orally, available documentation on such groups is almost always suspect. This is because insiders do not normally record their teachings in writing, and outsiders do not always interpret these teachings correctly and tend to add their own interpretations.

Monotheistic Traditions

Debates Over the Judeo–Christian Tradition

The controversy about environmental values in Christianity and Judaism has centered on the appropriate interpretation of the relationship between God, 'man', and nature as set out in the Book of Genesis 1:26–28:

> "26 And God said, Let us make man in our image, after our likeness: and let them have dominion over the fish of the sea, and over the fowl of the air, and over the cattle, and over all the earth, and over every creeping thing that creepeth upon the earth.
>
> 27 So God created man in His own image, in the image of God created He him; male and female created He them.
>
> 28 And God blessed them, and God said unto them, Be fruitful, and multiply, and replenish the earth, and subdue it: and have dominion over the fish of the sea, and over the fowl of the air, and over every living thing that moveth upon the earth".

There are two widely held interpretations of the relationships implied by these quotations. The 'mastery' interpretation is that Genesis gives humans a unique status among the species by virtue of creation in the image of God, and awards humans a God-given right to exploit nature without moral restraint, except where it affects human welfare. God, according to this view, seems to have intended humans to be his viceroy on earth. Humans are to the rest of earth's species as God is to humans. Humans are given dominion over the earth, and are expressly asked to subdue (Hebrew *kabas*, stamp down) the earth, as if nature needed humans to put it in order.

The 'stewardship' interpretation also relies on the verses above, as well as others before and after it, and emphasizes responsibility. Since humans are created in the image of God, they have not only rights and privileges but also special duties and responsibilities. Foremost among these is the human duty to rule the dominion of nature wisely. To degrade nature and destroy God's other species would violate the trust God has placed in his viceroy. Thus, far from giving a free hand to exploit and degrade nature, according to this view, God expects humans to exercise stewardship in the wise use of nature.

The Assisi Declarations (1986) reproduced here are excerpts from five of the addresses by religious leaders to their own faithful in the Buddhist, Christian, Hindu, Jewish, and Moslem worlds. The Declarations were the highlight of an interfaith ceremony at the World Wildlife Fund's 25th anniversary celebrations. The three monotheistic traditions represented in **Box 1**, **Box 2**, and **Box 3** (Christian, Jewish, and Moslem, respectively) make it clear that all three strongly support the stewardship interpretation of Genesis. It should be noted that Islam's holy book, the Qur'an, is similar in content to the Old Testament, and gives humans dominion over nature, but at the same time, as the declaration notes, vests in them trusteeship and accountability.

Box 1 Excerpts from the five declarations of assisi: The christian declaration by Father Lanfranco serrini

Because of the responsibilities which flow from his dual citizenship, man's dominion cannot be understood as license to abuse, spoil, squander or destroy what God has made to manifest his glory. That dominion cannot be anything else than a stewardship in symbiosis with all creatures. On the other hand, his self-mastery in symbiosis with creation must manifest the Lord's exclusive and absolute dominion over everything, over man and over his stewardship. At the risk of destroying himself, man may not reduce to chaos or disorder, or, worse still, destroy God's bountiful treasures. For St. Francis, work was a God-given grace to be exercised in that spirit of faith and devotion to which every temporal consideration must be subordinate: uncontrolled use of technology for immediate economic growth, with little or no consideration for the planet's resources and their possible renewal; disregard for just and peaceful relations among peoples; destruction of cultures and environments during war; ill-considered exploitation of natural resources by consumer-oriented societies; unmastered and unregulated urbanization; and, the exclusive preoccupation with the present without any regard for the future quality of life.

Box 2 Excerpts from the five declarations of assisi: The jewish declaration by Rabbi Arthur Hertzberg

The encounter of God and man in nature is conceived in Judaism as a seamless web, with man as the leader and custodian of the natural world. Even in the many centuries when Jews were most involved in their own immediate dangers and destiny, this universalist concern never withered… Now, when the whole world is in peril, when the environment is in danger of being poisoned, and various species, both plant and animal are becoming extinct, it is our Jewish responsibility to put the defence of the whole of nature at the very centre of our concern.

… Man was given dominion over nature, but he was commanded to behave towards the rest of creation with justice and compassion. Man lives, always, in tension between his power and the limits set by conscience.

Our ancestor Abraham inherited his passion for nature from Adam. The later rabbis never forgot it. Some 20 centuries ago they told the story of two men who were out on the water in a rowboat. Suddenly, one of them started to saw under his feet. He maintained that it was his right to do whatever he wished with the place which belonged to him. The other answered him that they were in the rowboat together – the hole that he was making would sink both of them (Vayikra Rabbah 4:6).

Box 3 Excerpts from the five declarations of assisi: The moslem declaration by Dr. Abdullah Omar Nasseef

Unity, trusteeship and accountability, that is *tawheed*, *khalifa* and *akhrah*, the three central concepts of Islam, are also the pillars of the environmental ethics of Islam. They constitute the basic values taught by the Qur'an. It is these values which led Muhamad, the Prophet of Islam to say: "Whoever plants a tree and diligently looks after it until it matures and bears fruit is rewarded", and "If a Moslem plants a tree or sows a field and men and beasts and birds eat from it, all of it is charity on his part", and again "The world is green and beautiful and God has appointed you his stewards over it". Environmental consciousness is born when such values are adopted and become an intrinsic part of our mental and physical makeup. Moslems need to return to this nexus of values, this way of understanding themselves and their environment. The notions of unity, trusteeship and accountability should not be reduced to matters of personal piety; they must guide all aspects of their life and work. Shariah should not be relegated just to issues of crime and punishment, it must also become the vanguard for environmental legislation. We often say that Islam is a complete way of life, by which it is meant that our ethical systems provides the bearings for all our actions.

Searching for a New Religious Base for Environmental Ethics

The debates over the interpretations have been going on since the 1970s. Many environmentalists do not see monotheistic stewardship as a sufficient solution. The critics include the historian Lynn White Jr., who started the debate over Judeo–-Christian environmental ethics with his classic 1967 paper in *Science*, 'The historical roots of our ecologic crisis'. According to White, Christianity is the most anthropocentric (human-centered) religion the world has ever seen, especially in its Western form. In pre-Christian times every tree, every spring, every stream, every hill had its own *genius loci*, its guardian spirit. These spirits were accessible to men, but were very unlike men; centaurs, fauns, and mermaids show their ambivalence. Before one cut a tree, mined a mountain, or dammed a brook, it was important to placate the spirit in charge of that particular situation, and to keep it placated. By destroying pagan animism, Christianity made it possible to exploit nature in a mood of indifference to the feelings of nature.

Since the roots of the environmental crisis "are so largely religious", White continues, "the remedy must also be essentially religious". He briefly entertains and then rejects Zen Buddhism as a solution: "Zen is as deeply conditioned by Asian history as Christianity is by the experience of the west, and I am dubious of its viability among us". Seeking instead an alternative Christian view, White proposes Saint Francis of Assisi as a patron saint for ecologists. "The key to an understanding of Francis is his belief in the virtue of humility – not merely for the individual but for man as a species. Francis tried to depose man from his monarchy over creation and set up a democracy of all of God's creatures".

White may be taken to recommend a Franciscan theology as the basis for a Christian environmental ethic. Others have provided other Christian theologies. For example, Rene Dubos in *A God Within* (1972) recommends St. Benedict, as the Benedictines were the original environmental managers of Europe who drained swamps and made the countryside both productive and beautiful. The theologian John B. Cobb Jr., casts his net wider, beyond European cultures, in looking for a religious solution in *Is It Too Late*? *A Theology for Ecology*, and others have done the same (Levine, 1994; Wirzba, 2003).

Accepting White's (1967) view that Christianity is largely responsible for the environmental crisis, Cobb (1995) examines in particular American Indian and Chinese worldviews. He rejects the American Indian view because, he thinks, Indians did not respect human life sufficiently and because the Indian way of life cannot support the existing North American population. Turning to the Chinese tradition (represented by Taoism), Cobb argues that it was unable to prevent deforestation and other ecological ills. He concludes that it is more prudent for the West to fix the Western tradition than to find a non-Western alternative. Rejecting Francis as too radical (Saint Francis had also preached poverty to the Catholic Church!), Cobb finally opts for Albert Schweitzer's Christianity and his reverence-for-life ethic.

The historian Arnold Toynbee differs from the above-mentioned commentators in that he sees no fundamental problem in developing a new environmental ethic that can be integrated with Asian traditions. As a classically educated scholar, he sees a historical continuity from pantheistic traditions ('the original religion of all mankind') to the Greco–Roman religion and to Asian religions. At the western end of the Old World, this nature-worship, which is the original religion of all mankind, has been overlaid by an opaque veneer of Christianity and Islam. However, when a native of the monotheistic portion of the present-day World travels eastward beyond the easternmost limits of Islam, he finds himself in a living premonotheistic World, and, if he has had a Greek and Latin education, the religion of present-day East Asia will be more familiar to him.

Monotheism is exceptional among the religions of the world in its doctrine about subduing nature, says Toynbee. As enunciated in the Book of Genesis, monotheism has removed the sense of awe and, with it, the age-old restraint that was once an effective check on human greed. His analysis does not so much negate the stewardship interpretation as makes it irrelevant. The real solution to the crisis, in Toynbee's thinking, has to do with restoring a sense of sacred respect for the earth and its creatures; the remedy lies in reverting from a worldview of monotheism to a worldview of pantheism.

Of course, the historical roots of the ecological crisis are not only found in the area of religion. A complex of social changes related to the Age of Enlightenment, the Industrial Era, and the commodification of nature has been implicated, as part of what Karl Polanyi (1944) calls the *Great Transformation*. A major factor, according to C.J. Glacken (1967) and others, is the Age of Enlightenment concept of an external environment analytically separate from humans. Many scholars, including Gregory Bateson, hold that this separation is the basis of the Cartesian dualism of mind versus matter, and hence humans versus the environment. Searching for a way to bridge the dualism, Bateson became 'aware gradually that the unity of nature … might only be comprehensible through the kind of metaphors familiar from religion … an integrative dimension of experience that he called the sacred' (Bateson and Bateson, 1987, p. 2).

There is no agreement among experts regarding the relative importance of the emergence of monotheistic traditions in the West versus that of the Enlightenment philosophy and the great transformation many centuries later. The change in values accompanying the former was no doubt complemented and reinforced by the latter. Together, these two waves of change left the Western world with a fundamentally altered relationship with nature. The 'disenchantment of the world' in modern society may be seen as a loss of spiritual enchantment with life. For premonotheistic people, by contrast, nature was not just a treasure trove of natural resources; nature was a goddess and the whole of the environment was sacred.

Asian Traditions

The religious traditions of South Asia and East Asia, with their many gods and nondominant relationship to nature, are similar in many ways to the premonotheistic traditions of Europe, and fundamentally different from the monotheistic tradition, as Toynbee would argue. "Confucianism and Taoism and Shinto, like the pre-Christian Greek cults of the corn-goddess Demeter (Ceres in Latin) and the wine-god Dionysus, counsel man to respect nature even when he is applying his human science to coax nature into bestowing her bounty on man". Elements of this thinking can be seen in the Buddhist and Hindu declarations in Assisi. The statements "we should be wary of justifying the right of any species to survive solely on the basis of its usefulness to human beings" (**Box 4**) and "the divine is not exterior to creation, but expresses itself through natural phenomena" (**Box 5**) signal a worldview that is fundamentally different from the monotheistic one.

The influence of Hindu teachings can be traced in the environmental literature regarding species preservation and animal rights. The philosopher Arne Naess (1989) adopted the concepts of "identification" and "self-realization" from Hindu thought and used them as central ideas in Deep Ecology. Zen Buddhism (a form of Buddhism) is one of the few major Asian

Box 4 Excerpts from the five declarations of assisi: The buddhist declaration by the Venerable Lungrig Nmgyal

Buddhism is a religion of love, understanding and compassion and is committed towards the ideal of non-violence. As such it also attaches great importance towards wildlife and the protection of the environment on which every being in this world depends for survival. The underlying reason why beings other than humans need to be taken into account is that, like human beings, they too are sensitive to happiness and suffering. We should therefore be wary of justifying the right of any species to survive solely on the basis of its usefulness to human beings. We are told that history is a record of human society in the past. From existing sources there is evidence to suggest that for all their limitations, people in the past were aware of this need for harmony between human beings and nature. They loved the environment. They revered it as the source of life and well-being in the world. We regard our survival as an undeniable right; as coinhabitants of this planet, other species too have this right for survival. And since human beings as well as other non-human sentient beings depend upon the environment as the ultimate source of life and well-being, let us share the conviction that the conservation of the environment, the restoration of the imbalance caused by our negligence in the past, be implemented with courage and determination.

Box 5 Excerpts from the five declarations of assisi: The hindu declaration by Dr. Karan Singh

Not only in the Vedas, but in later scriptures such as the Upanishads, the Puranas and subsequent texts, the Hindu view-point on nature has been clearly enunciated. It is permeated by a reverence for life, and an awareness that the great forces of nature – the Earth, the sky, the air, the water and fire – as well as various orders of life including plants and trees, forests and animals, are all bound to each other within the great rhythms of nature. The divine is not exterior to creation, but expresses itself through natural phenomena.

The Yajurveda lays down that "no person should kill animals helpful to all. Rather, by serving them, one should attain happiness". This view was later developed by the great Jain Tirthankara, Lord Mahmavira, who regenerated the ancient Jain faith that lives down to the present day. For the Jains, Ahimsa, or non-violence, is the greatest good, and on no account should life be taken. This philosophy was emphasized more recently by Mahatma Gandhi, who always spoke of the importance of Ahimsa and looked upon the cow as a symbol of the benign element in animal life. All this strengthens the attitude of reverence for all life including animals and insects. The Hindu tradition of reverence for nature and all forms of life, vegetable or animal, represents a powerful tradition which needs to be renurtured and reapplied in our contemporary context. India, the population of which is over 80% Hindu, has in recent years taken a special interest in conservation.

religions to put firm roots in contemporary Western culture through the writings of D. T. Suzuki, Alan Watts, Peter Timmerman, and the poet Gary Snyder. However, "pop Zen" or "West Coast Zen", as Callicott and Ames (1989) call it, has been used to reinforce selfish individualism already endemic in Western culture, instead of the self-discipline, simplicity, and enlightenment of true Zen.

Buddhism has also influenced some environmental writing. The most radical chapter in E. F. Schumacher's (1973) *Small Is Beautiful* is entitled "Buddhist economics". The West is rich in technological and economic means, Schumacher points out, but it has little clue about the ends that are most appropriate for the use of those means. Buddhist economics proposes the solution that appropriate means should be used to achieve appropriate ends. The discovery of these appropriate ends, in turn, can be guided through our place as caretakers of the world around us, and the appropriate means will then become obvious.

Taoism, one of the Chinese religious traditions, considers nature through the concept of Tao, the way, and emphasizes living in harmony with nature, of "flowing with nature". Taoism has influenced Western culture and environmental thinking through such books as *The Tao of Physics* by Fritjof Capra (1975). In a Taoist story told by Timmerman, a philosopher falls asleep and dreams that he is a butterfly. From that moment on, he is never sure if he is a philosopher dreaming that he is a butterfly, or a butterfly dreaming that he is a philosopher. In Taoist thinking, the crucial point is that there is no right answer: the moment an answer is attempted, the original symmetry is broken and one crystallizes out as a philosopher or a butterfly—but not both. Capra, a physicist, argues that quantum physics works in much the same way as electrons move out of a probability sphere into determined existence.

Romantic notions of Asian religions are debunked by the geographer Yi-Fu Tuan (1974), who pointed out the discrepancies between attitude and behavior. "Western humanists commonly show bias in favor of Taoist and Buddhist traditions", says Tuan. But the people of China, through their long history, have transformed and degraded the landscape in a major way. Specifically, Taoism, with its vaguely anti-urban and anti-humanistic stance, according to Tuan, represents little more than a hermit's point of view in the larger scheme of environmental destruction. Tuan's exposition of episode after episode of land abuse by the ancient Chinese, long before they were influenced by the West, leads to the conclusion that Chinese religions are unlikely to provide a solution for all. However, this is not to deny that there existed a tradition of natural philosophy in China consistent with contemporary ecological ideals. Rather, Tuan's critique points out that there often exist "glaring contradictions of professed ideal and actual practice".

Pantheistic Traditions

Traditional Peoples, Indigenous Knowledge

Some traditional societies retain elements of pantheism, the 'original religion' of all humankind. In these societies, religion, worldview, and environmental practice are often intertwined (Brown, 1953; Callicott, 1994). Religious sanctions may be invoked in two ways in direct support of biodiversity conservation: through the prohibition of areas or of species. 'Sacred groves' or sacred forests occur throughout the world, especially in India, Indonesia, South America, and parts of Africa (Ramakrishnan *et al.*, 1998). **Box 6** illustrates the nature of sacred groves in one African country. There has been increased attention given to examining the role of sacred natural areas in conservation, and specifically in biocultural diversity conservation (Maffi and Woodley, 2010). Even small sacred groves may be surprisingly effective in conserving biodiversity (Verschuuren *et al.*, 2010). In the case of species taboos (the word is borrowed from the Polynesian *tabu*), a study by Colding and Folke (1997) showed that about 30% of taboos identified worldwide prohibited the use of species that also happened to be listed as threatened by the IUCN *Red Data Book*.

Traditional worldviews may also help conserve biodiversity indirectly through the emotional involvement of people with their land and living things. This may entail, as in some American Indian traditions, a community-of-beings worldview in which animals and plants are considered persons—not human persons but persons nonetheless. Land may also provide a strong sense of place or sense of identity for a social group, as seen with American Indians and Inuit, Southeast Asian indigenous peoples, and cultures in Oceania.

The report *Our Common Future* identified indigenous peoples as "the repositories of vast accumulations of traditional

Box 6 Sacred groves and traditional forest conservation in Nigeria

Among the Yoruba of southwestern Nigeria, there were a number of categories of traditional sacred forests, although these have fallen into disuse in recent years. According to the study, *Nigeria's Threatened Environment* (1991), land was regularly set aside as hunting forests, religious groves, and isolation and quarantine forests, and to serve as the abode of fairies and spirits.

Igbo ode (hunting forest) was lands located at some distance from settlements and devoted to game. *Igbo egan* (high forest) was abandoned secondary forest. When put to use for game-hunting activities, it became *Igbo ode*. Lands were often named according to their wildlife inhabitants, e.g., *Igbo erin* (elephant forest) *and Igbo efon* (buffalo forest). Only brave hunters dared use such specialized forests.

Igbo oro (religious groves) were places set aside for religious worship of many of the elements of the physical environment. They were not extensive (usually less than a quarter of a hectare) and were uncultivated forests located on the borders of a settlement, in as many separate locations as there were families of the deities.

Igbo egbee (religious groves "land of sorrows") were reserved forests for the burial of people whose deaths were considered mysterious. Such lands, isolated further away from settlements, were never put under cultivation in the past when diseases were rampant and sudden deaths were attributed to the anger of the gods. Only brave hunters dared enter such lands.

knowledge and experience that link humanity with its ancient origins" (World Commission on Environment and Development 1987, p. 12) and encouraged the use of this indigenous knowledge to develop new adaptive strategies for conservation. Who are these traditional societies and indigenous peoples? According to international criteria, four characteristics distinguish indigenous peoples from others. They are descendants of groups inhabiting an area prior to the arrival of other populations; they are politically not dominant; they are culturally different from the dominant population; and they identify themselves as indigenous.

Indigenous peoples are found in many parts of the world. Callicott's (1994) book on environmental ethics and religious traditions provides a representative survey of indgenous ethics from various parts of the world, with major sections entitled "Polynesian Paganism", "American Indian Land Wisdom" (with major cases of Lakota shamanism and Ojibwa totemism), "South American Eco-eroticism" (cases of Tukano systems theory and Kayapo agro-ecology), "African Biocommunitarianism" (cases of Yoruba anthropo-theology and San etiquette of freedom), and "Australian Aboriginal Conservators".

The globalization of Western culture has meant, among other things, the globalization of Western modes of production (e.g., monocultures) and resource conservation (expert-knows-best positivist science). Has ancient knowledge become irrelevant, or perhaps simply been swamped by Western science and practice? Conversely, are there useful lessons that can be learned from indigenous knowledge and practice? These are some of the central questions in a growing body of literature on traditional ecological knowledge.

The Context of Traditional Ecological Knowledge (Indigenous Knowledge)

Berkes (2012, p. 7) defines traditional ecological knowledge as a cumulative body of knowledge, practice, and belief, evolving by adaptive processes and handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and with their environment. As a knowledge–practice–belief complex, traditional ecological knowledge includes the religious traditions of a society. It is both cumulative and dynamic, building on experience and adapting to changes. It is an attribute of societies with historical continuity in resource use on a particular land. By and large, these are nonindustrial or less technologically oriented societies, many of them indigenous or tribal, but not exclusively so. Traditional ecological knowledge is a subset of indigenous knowledge, generally defined as local knowledge held by indigenous peoples or local knowledge unique to a given culture or society.

Any discussion of traditional ecological knowledge and indigenous conservation needs to be qualified. The question of the relevance of traditional societies for biodiversity conservation is confounded by what has been called the 'noble savage' syndrome. On the one hand, there is the romantic, Rousseauian view of 'the primitive', intrinsically attuned to nature and somehow living 'in balance' with the environment. Such noble savages tend to become 'fallen angels' when they come into contact with the dominant society. Lest they become a threat to the very ecosystem in which they live, they should continue to live as primitives, according to this view. On the other hand, there is the view that primitive people are not noble savages but tend to be ignorant and superstitious. Historically, they lived as biological populations at the mercy of natural forces and supernatural beliefs, not as organized communities with their own knowledge systems and ecologically adaptive practices. They had a tendency, even as primitive hunters, to trigger massive species extinctions, as in New Zealand, Polynesia, Madagascar, and perhaps the Americas and Australia.

As with all myths, there are elements of truth at the basis of these two views, and many people, including conservationists and indigenous people themselves, no doubt believe parts of one or more of these views. An alternative third view might start with the recognition that traditional ecological knowledge is not mere tradition but a set of adaptive responses that have evolved over time. All societies, prescientific and scientific, strive to make sense of how the world behaves, and to apply this knowledge to guide practice. Because people were dependent for their survival on resources in their immediate environments, there were strong incentives for them to use resources sustainably. They could not mask this dependence with fossil fuel subsidies or capital markets in a globalized economy. Thus, the ability to nurture and sustain biodiversity was a selective pressure on these societies. That is, there was a survival value in conserving and augmenting local biodiversity.

Prescientific, traditional systems of management have been the main ways by which societies have managed their natural resources for millennia. In many cases, the main reason we still have any biodiversity to speak about is because of these systems of management. Within this context, biodiversity conservation is the indirect outcome, rather than the objective, of traditional practices, but the practices themselves had adaptive value – biodiversity conservation was often a matter of survival.

Such a probabilistic and evolutionary view clearly rejects the notion that all traditional people always manage resources well. What the literature suggests is that groups can make mistakes but can also learn from their mistakes. They are capable of responding to resource collapses and to environmental change, along the lines suggested in the books *Linking Social and Ecological Systems* (Berkes and Folke, 1998) and *Sacred Ecology* (Berkes, 2012). Detailed accounts of traditional ecological knowledge do not support the noble savage. What they do is something very different: They provide a documentation of extremely detailed, pragmatic, empirical knowledge, represented in religious traditions and rituals, bound up with an emotional tie to the land and living beings.

Conserving Biodiversity: Traditional 'Rules of Thumb'

How indigenous knowledge of nature and human social behavior is translated into resource use practices that tend to promote sustainable use of the biota and conservation of biodiversity is little known. Better known, many traditional people exhibit resource use restraints that promote

conservation. But knowledge, belief, and practice leading to restraint tend to be intermingled, making it difficult to trace linkages among them. Conservation researchers often ignore social restraints on grounds that they are not considered rational in our worldview, often involving beliefs in supernatural forces, for example, as in the case of taboos. To arrive at an appropriate set of social restraints is an order of magnitude more complicated than to employ knowledge of habitat preference and behavior for efficient hunting strategies. To implement a set of social restraints is complicated also because it requires continued cooperation of a large number of individuals.

The dealings of traditional societies with nature are often hedged by prescriptions as to what, when, and how much is to be left undisturbed. These prescriptions become part of a culture and are mediated by religious traditions. Indian ecologist Madhav Gadgil and colleagues (1993) identified four kinds of widely used 'rules of thumb' as social restraints leading to indigenous biodiversity conservation practice:

Provide Total Protection to some Biological Communities or Habitat Patches

These may include pools along river courses, sacred ponds, sacred mountains, meadows, and forests. For example, sacred groves were once widely protected from Africa to China, and in fact, throughout the Old World. In the tribal state of Mizoram in northeastern India, they continue to be protected even after conversion to Christianity. It is now called a 'safety forest', while the village woodlot from which regulated harvests are made is called the 'supply forest'. Ecological theory suggests that providing such absolute protection in 'refugia' can be a very effective way of ensuring persistence of biological populations.

Provide Total Protection to Certain Long-Lived (K-selected) Species

Trees of all species of the genus *Ficus* are protected in many parts of the Old World. It is notable that *Ficus* is considered a keystone genus significant to the conservation of overall biodiversity. Local people seem to be often aware of the importance of *Ficus* as affording food and shelter for a wide range of birds, bats, and primates, and it is not difficult to imagine that such understanding was converted into widespread protection of *Ficus* trees at some point in the distant past. Taboos with apparent functional significance may also be placed on some less obvious species within the ecological community. For example, some Amazon fish species considered important for folk medicine are taboo and are avoided as food.

Protect Critical Life History Stages

In south India, fruit bats may be hunted when foraging, but not at daytime roosts on trees that may be in the midst of villages. Many waders are hunted outside the breeding season, but not at heronaries, which may again be on trees lining village streets. Cree Indians of James Bay in the subarctic hunt Canada goose, a major subsistence resource, but rarely kill or even disturb nesting geese. The danger of overharvest and depletion of a population is clearly far greater if these vulnerable stages are hunted and the protection afforded to them seems to be a clear case of ecological prudence.

Mandate Local Stewards to Monitor and Oversee Resource Use

Traditional resource harvesting systems in diverse parts of the world rely on the leadership of a traditional expert to organize the harvest, control access, supervise local rules, and generally act as a 'steward'. This practice also ensures the proper use and transmission of knowledge. Further, in some societies, major events of resource harvest are carried out as a short-term, prescribed group effort. Thus, many tribal groups engage once a year in a large-scale communal hunt. Such a group exercise may also serve the purpose of assessing the status of prey populations and their habitat and may help to adjust resource harvest practices to sustain yields and conserve biodiversity.

Considering that contemporary scientific prescriptions for biodiversity conservation are, in effect, also 'rules of thumb', these traditional prescriptions add up to a reasonably good set of practices. Thus, traditional knowledge, with its time depth of observations, and an accumulated experience of learning-by-doing, can help us arrive at practical, locally tailored prescriptions for resource use relevant to conservation needs.

Ecosystem People, Local Knowledge, and Sacred Respect

Many practices used by indigenous people contribute to the conservation of biodiversity through the use of more varieties, species, and landscape patches than do modern agricultural food production systems (which tend to rely heavily on the monoculture of a few varieties). Several volumes illustrate this notion (Posey, 1999; Grim, 2001; Turner, 2005; Berkes, 2012). Gary Nabhan's (2000) work in Arizona and Mexico provides an illustration of biodiversity enhancement due to traditional multispecies agriculture, in this case, as compared to a protected area. Investigating two similar oases in the Sonoran Desert, Nabhan found 32 species of birds in the US Organ Pipe Cactus National Monument on the Arizona side of the desert. The other oasis, 50 km away on the Mexican side of the border and still being farmed in the traditional Papago style by a group of Indian villagers, supported 65 species. A possible explanation was that the farmed oasis provided greater diversity and amounts of food available for the birds, as compared to the protected oasis.

There is evidence that indigenous practices based on good local observations and natural history can provide the capability to respond to feedbacks from the environment and to ecosystem change. The book *Linking Social and Ecological Systems* documents examples of multiple-species management, resource rotation, succession management, landscape patchiness management, and other ways of responding to and managing pulses and ecological surprises (Berkes *et al.*, 2000). Even assuming that biodiversity conservation is the *indirect outcome* of these practices, for the most part, the amount of ecological understanding suggested by these practices is considerable. Social mechanisms behind these practices include adaptations for the generation, accumulation, and transmission of knowledge and the use of local institutions and rules for social regulation. Religious traditions are important for the cultural internalization of traditional practices and for the development of worldviews and cultural values appropriate for them.

Some traditional societies have ecosystem-like concepts. Ancient conceptualizations of ecosystems exist in several American Indian, Asia–Pacific, European, and African cultures. Among some indigenous peoples of the North American Subarctic, *land* is more than a physical landscape; it encompasses the living environment, including humans. For example, the term used by the Dene groups of the Canadian Western Subarctic, *ndé* (*ndeh*), is usually translated as *land* but its meaning is closer to *ecosystem* because it conveys a sense of relations of living and nonliving things on the land. However, it differs from the scientific concept of ecosystem in that *ndé* is based on the idea that everything in the environment has life and spirit (Berkes, 2012).

Table 1 provides a summary of traditional ecosystem-like concepts in which the unit of nature is often defined in terms of a physical boundary (such as a watershed) and living and nonliving elements are considered interlinked (references in Berkes, 2012). For example, among the Gitksan people of the Pacific Northwest, tribal chiefs describe their land boundaries as 'from mountain top to mountain top' and orient themselves by two directional axes within this watershed framework: vertically up and down from valley bottom to mountain top, and horizontally, upstream and downstream. Detailed land use maps of the kinship-based *house* groups (*wilps*) of the Gitksan show that there is a close correspondence between watershed areas and *wilps* or clusters of *wilps*. Clearly, these are not merely political boundaries but watershed-ecosystems-as-territories.

In Oceania there was a wealth of ecosystem-like concepts. Examples include the ancient Hawaiian *ahupua'a*, wedge-shaped land units controlled by local chiefs, encompassing entire valleys stretching from the top of mountains to the coast and shallow waters. The variations of the Hawaiian system may be found in the Yap *tabinau*, the Fijian *vanua*, and the Solomon Islands *puava*. The common point in each is that the term refers to an intimate association of a group of people with land, reef, and lagoon and all that grows on or in them. It is the 'personal ecosystem' of a specific group of people. In the Solomons, for example, a *puava* is a named territory consisting of land and sea, and it includes all lands and resources associated with a kinship group.

Table 1 Examples of traditional applications of the ecosystem view

System	*Country/region*
Watershed management of salmon rivers and associated hunting and gathering areas by tribal groups	American Indians of the Pacific Northwest
Delta and lagoon management for fish culture (*tambak* in Java), and the integrated cultivation of rice and fish	South and Southeast Asia
Vanua (in Fiji), a named area of land and sea, seen as an integrated whole with its human occupants	Oceania, including Fiji, Solomon Islands, and ancient Hawaii
Family groups claiming individual watersheds (*iworu*) as their domain for hunting, fishing, and gathering	The Ainu of northern Japan
Integrated floodplain management (*dina*) in which resource areas are shared by social groups through reciprocal access arrangements	Mali, Africa

These examples provide insights into what Raymond Dasmann (1988) calls ecosystem people – people who depend heavily on natural resources of their own localities, and hence develop a very detailed understanding of their environment, as well as a spiritual intimacy with the land. Such indigenous ecological understanding is different from that of the scientist. These traditional views tend to depict ecosystems not as lifeless, mechanical, and distinct from people, but as fully alive and encompassing humans. In some cases, traditional concepts of land also incorporate spirits of animals and other natural objects (as among the Dene of Northern Canada and Alaska), and spirits of human ancestors, as among some African groups and among the Australian aborigines (the concept of Dreamtime).

Conclusions

In monotheistic traditions, humans have dominion over the earth. They have a God-given right to subdue the earth, but they are also the stewards of God with responsibility to take care of the earth. By contrast, pantheistic traditions do not have a dominant relationship with nature. Asian religious traditions and pre-Christian European traditions also have a nondominant view of nature. The record of different religious traditions in conserving biodiversity is mixed. This is because human behavior is conditioned not only by religious beliefs but also by many other forces and because both culture and science evolve over time (Capra, 1996).

Traditional knowledge–practice–belief systems, tapping the wisdom of many traditional cultures with pantheistic traditions, offers a number of lessons. These systems are characterized by a similarity of concepts of nature, in which humans are part of nature. As well, these knowledge systems are characterized by similarity in design. An example is shifting cultivation, developed apparently independently, by tropical forest peoples in Africa, South America, South Asia, and New Guinea. At the same time, traditional knowledge systems are characterized by a remarkable diversity in practice, even in adjacent areas. For example, taboos used in the social control of resources in Oceania can vary greatly from one island to another, and even from one part of an island to another. Such diversity in resource use practice contrasts with the relatively uniform conservation prescriptions of government agencies, and highlights the need for conceptual pluralism.

One important lesson from traditional ecological knowledge is that values and beliefs are important in encoding ethics, including the ethics of conservation. As Rappoport (1984) and Anderson (1996) point out, the use of emotionally powerful cultural symbols is important to develop and enforce a moral code. If this is true, the incorporation of values and beliefs into biodiversity conservation efforts is more likely to succeed than the use of purely scientific arguments or purely economic incentives.

A number of contemporary concepts appear to be exploring the combination of emotions and ideas that may help restore a sense of sacred respect. They include deep ecology (Naess, 1989), the biophilia hypothesis (love of living beings)

(Kellert and Wilson, 1993), bioregionalism (with its combination of local self-reliance and sense of belonging), the related notion of 'sense of place' (Tuan, 1974), and the Gaia hypothesis which is the contemporary version of the Mother Earth idea. Each provides an approach to the understanding of reciprocal ties that bind humans with the natural world, and these ties invariably have a spiritual or religious aspect.

See also: Conservation Movement, Historical. Ecosystem, Concept of. Environmental Ethics. Historical Awareness of Biodiversity. Indigenous Peoples and Biodiversity. Literary Perspectives on Biodiversity. Social and Cultural Factors. Stewardship, Concept of

References

Anderson EN (1996) *Ecologies of the Heart: Emotion, Belief and the Environment*. New York and Oxford: Oxford University Press.

Assisi Declarations (1986) Reproduced in World Wildlife Fund *WWF News*, November/December 1986. http://www.nyo.unep.org/eaf/eafadec.pdf (last checked: 21 April 2012).

Bateson G and Bateson MC (1987) *Angels Fear: Towards an Epistemology of the Sacred*. New York: Bantam Books.

Berkes F (2012) *Sacred Ecology*, 3rd edn. New York: Routledge.

Berkes F, Colding J, and Folke C (2000) Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications* 10: 1251–1262.

Berkes F and Folke C (eds.) (1998) *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge University Press.

Brown JE (ed.) (1953) *The Sacred Pipe: Black Elk's Account of the Seven Rites of the Oglala Sioux*. Norman: University of Oklahoma Press.

Callicott JB (1994) *Earth's Insights: A Multicultural Survey of Ecological Ethics from the Mediterranean Basin to the Australian Outback*. Berkeley: University of California Press.

Callicott JB and Ames RT (eds.) (1989) *Nature in Asian Traditions of Thought: Essays in Environmental Philosophy*. Albany, NY: State University of New York Press.

Capra F (1975) *The Tao of Physics*. Boston: Shambhala.

Capra F (1996) *The Web of Life*. New York: Anchor Books/Doubleday.

Cobb Jr. JB (1995) *Is It Too Late? A Theology for Ecology*. Denton, TX: Environmental Ethics Books.

Colding J, and Folke C (1997) The relation between threatened species, their protection, and taboos. *Conservation Ecology* 1: article 6, 19 pp. http://www.consecol.org/vol1/iss 1/art6

Dasmann RF (1988) Towards a biosphere consciousness. In: Worster D (ed.) *The Ends of the Earth*, pp. 277–288. Cambridge: Cambridge University Press.

Dubos R (1972) *A God Within*. Scribner.

Gadgil M, Berkes F, and Folke C (1993) Indigenous knowledge for biodiversity conservation. *Ambio* 22: 151–156.

Glacken C (1967) *Traces on the Rhodian Shore: Nature and Culture in Western Thought from Ancient Times to the End of the Eighteenth Century*. Berkeley: University of California Press.

Grim JA (ed.) (2001) *Indigenous Traditions and Ecology: The Interbeing of Cosmology and Community*. Cambridge, MA: Harvard University Press.

Jenkins W (ed.) (2010) The spirit of sustainability. *Encyclopedia of Sustainability*, vol. 1. Great Barrington, MA: Berkshire.

Kellert SR and Wilson EO (eds.) (1993) *The Biophilia Hypothesis*. Washington, DC: Island Press.

Levine MP (1994) *Pantheism*. London: Routledge.

Maffi L and Woodley E (2010) *Biocultural Diversity Conservation: A Global Sourcebook*. London and Washington, DC: Earthscan.

Nabhan GP (2000) Native American management and conservation of biodiversity in the Sonoran Desert bioregion. In: Minnis PE and Elisens WJ (eds.) *Biodiversity and Native America*, pp. 29–43. Norman OK: University of Oklahoma Press.

Naess A (1989) *Ecology, Community and Lifestyle: Outline of an Ecosophy*. Translated and edited by D. Rothenberg. Cambridge: Cambridge University Press.

Polanyi K (1944) *The Great Transformation*. New York: Rinehart.

Posey DA (ed.) (1999) *Cultural and Spiritual Values of Biodiversity*. Nairobi: UNEP and Intermediate Technology Publications.

Ramakrishnan PS, Saxena KG, and Chandrashekara UM (eds.) (1998) *Conserving the Sacred for Biodiversity Management*. New Delhi: Oxford and IBH Publishing.

Rappaport RA (1984) *Pigs for the Ancestors*, 2nd edn. New Haven, CT: Yale University Press.

Schumacher EF (1973) *Small Is Beautiful*. London: Blond & Briggs.

Shaw S and Francis A (eds.) (2008) *Deep Blue: Critical Reflections on Nature, Religion and Water*. London: Equinox.

Taylor BR (ed.) (2005) *Encyclopedia of Religion and Nature*. London and New York: Thoemmes Continuum.

Tuan Y (1974) *Topophilia*. Englewood Cliffs, NJ: Prentice-Hall.

Turner NJ (2005) *The Earth's Blanket: Traditional Teachings for Sustainable Living*. Vancouver/Seattle: Douglas & McIntyre/University of Washington Press.

Verschuuren B, Wild R, McNeely JA, and Oviedo G (eds.) (2010) *Sacred Natural Sites: Conserving Nature & Culture*. London and Washington, DC: Earthscan.

White L (1967) The historical roots of our ecologic crisis. *Science* 155: 1203–1207.

Wirzba N (2003) *The Paradise of God: Renewing Religion in the Ecological Age*. New York: Columbia University Press.

World Commission on Environment and Development (1987) *Our Common Future*. Oxford and New York: Oxford University Press.

Worster D (ed.) (1988) *The Ends of the Earth: Perspectives on Modern Environmental History*. Cambridge, UK: Cambridge University Press.

Remote Sensing and Image Processing

Ruth DeFries, Columbia University, NY, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Ruth DeFries volume 5, pp 121–143, © 2001, Elsevier Inc.

Glossary

Classification The process of assigning the pixels of an image to discrete categories.
Electromagnetic radiation The energy transmitted through space in the form of electric and magnetic waves that can be detected by various types of sensors.
Geographic information systems (GISs) Computer systems that store, enhance, analyze, and display layers of geographic data and connect these to alphanumeric databases.
Image processing The manipulation of digital image data, including operations such as image display, restoration, enhancement, and classification.
Pixel Abbreviated from "picture element," the smallest part of a picture (image).
Sensor A device that gathers energy, converts it to a digital value, and presents it in a form suitable for obtaining information about the environment.

Introduction

It is now widely recognized that biodiversity is a multidimensional and multiscale phenomenon encompassing different organization levels (populations, species, functional groups, communities, ecosystems, and landscapes) and a wide range of spatial scales (from microhabitat heterogeneity to global-scale patterns). Identifying patterns of biodiversity and their causal factors is therefore an enormous task that requires (1) mapping and monitoring of biological patterns across different spatial and temporal scales and (2) analysis of such patterns with respect to diverse aspects of the physical and human environment. Remote sensing is one of the best tools available for coping with this challenge. Information obtained from remote sensing is intrinsically multidimensional (horizontally, vertically, temporally, and spectrally) and may cover spatial scales ranging from a few centimeters to entire continents (**Figure 1**). Data derived from remote sensing provide information on all levels of biological organization, including species, functional groups, ecosystems, and biomes. These capabilities make remote sensing an extremely valuable tool for studies of biodiversity.

The aim of this article is to draw attention to the potential and actual contribution of remote sensing to studies of biodiversity. The article consists of two main sections. The first section introduces fundamental concepts of remote sensing. The second section discusses applications of how remote sensing is being used for mapping, monitoring, and modeling patterns of biodiversity. The article is concluded by a brief summary that points to some current and expected future developments in this field.

Fundamentals of Remote Sensing

An Overview of the Remote-Sensing Process

Remote sensing can be broadly defined as "the technique of obtaining information about objects through the analysis of data collected by special instruments that are not in physical contact with the objects of investigation" (Avery and Berlin, 1992). A refined definition of remote sensing that better fits the scope of this article is: "The practice of obtaining information about the Earth's surface, using images acquired from airborne or spaceborne vehicles by measuring reflected, emitted, or returned electromagnetic radiation."

This narrower definition restricts our discussion to applications in which the general target of the observation is the Earth's surface, the measured energy is electromagnetic radiation, and the sensors are positioned on airborne or spaceborne platforms. Although some techniques of remote sensing do not fit this formalization (e.g., acoustical or sonar technologies), the above definition covers most of the methods that are currently applied for the documentation, monitoring, and mapping of biodiversity.

Figure 2 schematically illustrates the basic elements of the remote-sensing process. Any remote-sensing application consists of two distinct processes: data acquisition (detection and recording of electromagnetic radiation) and data analysis (extraction of information from the recorded data). The first process can be performed either photographically (i.e., by a photographic camera) or electronically (by electronic sensors). A photographic camera uses optical lenses and a light-sensitive film to detect electromagnetic radiation. The film acts as both the detector of the radiation and the recording medium. Electronic sensors convert electromagnetic radiation into electronic signals that can be stored as digital images. The end product of this stage of remote-sensing is therefore either a photograph or a digital image. Most current remote-sensing systems produce digital images.

The second step of the remote-sensing process is data analysis (**Figure 2**). Remotely sensed data, either photographs or digital images, must be analyzed in order to provide useful information about the observed features. Analytical techniques available for the analysis of remotely sensed data are diverse, ranging from traditional methods of visual interpretation to sophisticated computer-based image processing. The final product of both visual interpretation and computer-based

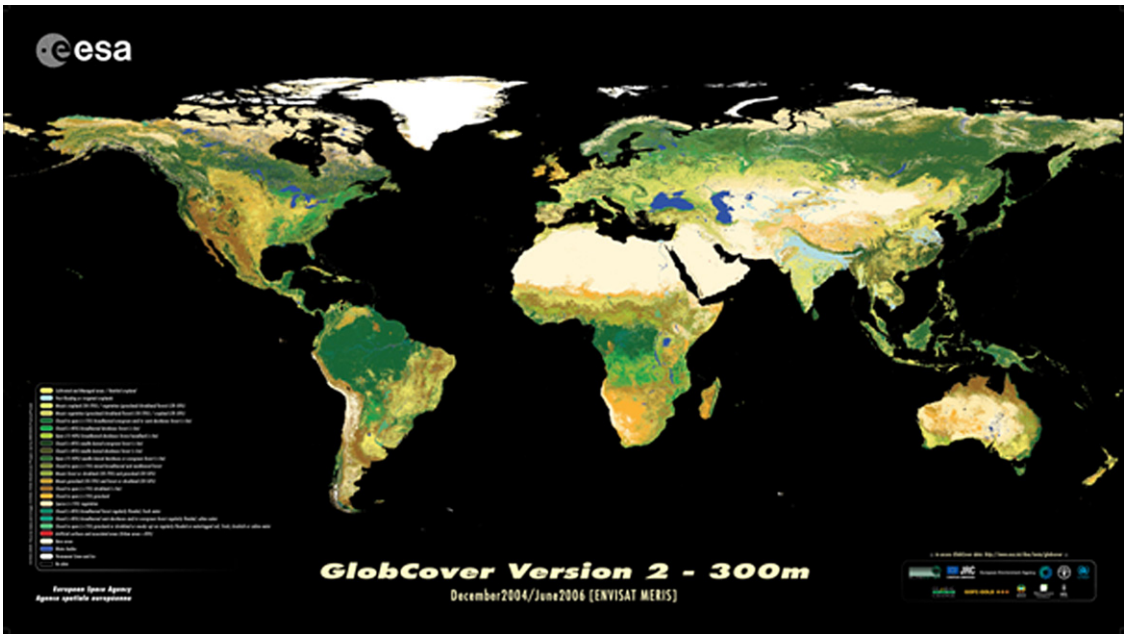


Figure 1 A global map of land cover derived from data acquired by MERIS.

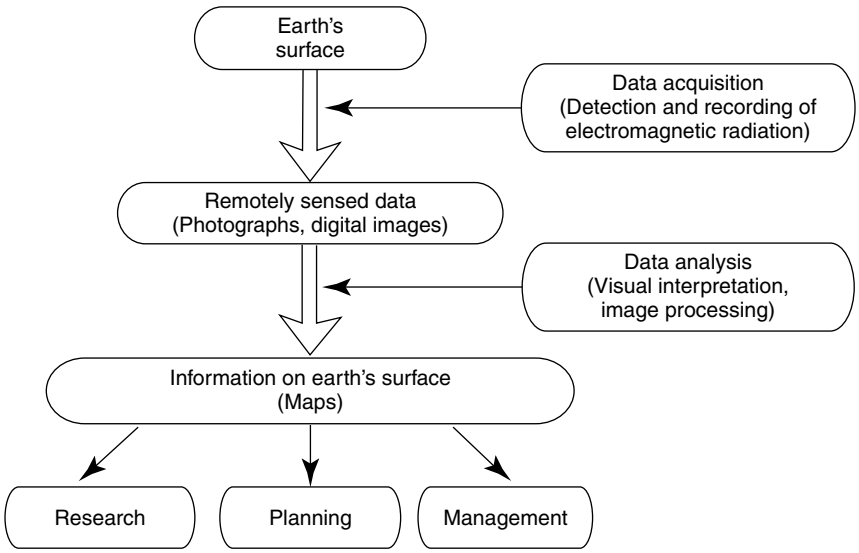


Figure 2 Basic elements of the remote-sensing process.

image processing is usually a map showing the spatial distribution of the variables of interest. Maps generated from remotely sensed data can be used directly, or in combination with other data sources, for a variety of applications, including research, decision making, planning, and management.

Electromagnetic Radiation and Its Measurement

As already defined, remote sensing is concerned with the measurement of electromagnetic radiation reflected, emitted, or returned from objects and features on the Earth's surface. In passive remote sensing, the sensor measures radiation reflected or emitted from the surface. The main source of radiation reaching the Earth is the Sun. Some of the solar radiation is reflected from the Earth's surface and can therefore be detected by cameras or electronic sensors. Other components of the solar radiation are absorbed at the Earth's surface and are reradiated as thermal energy. This form of electromagnetic radiation can also be detected and recorded by appropriate sensors. As different objects emit and reflect different types and amounts of radiation, images capturing such differences contain much information on the nature of the Earth's surface. In active remote sensing, the sensor sends its own electromagnetic energy to a target and measures the returned radiation.

The Electromagnetic Spectrum

A basic characteristic of electromagnetic radiation is its wavelength. Solar radiation consists of a wide and continuous range of wavelengths, stretching from $<10^{-6}$ to $>10^{7}$ μm ($1\ \mu m = 10^{-6}$ m). This continuum is subdivided into several divisions called spectral bands, which share similar characteristics (**Figure 3(a)**). The main bands used in remote sensing are ultraviolet (UV), visible, infrared (IR), and microwave. The UV band lies between the X-rays and visible light with wavelength limits of 0.01 and 0.4 μm (**Figure 3(a)**). Wavelengths shorter than 0.3 μm are unable to pass through the atmosphere and therefore only the 0.3–0.4 μm wavelength interval, or near-UV, is available for remote sensing (**Figure 3(b)**).

The visible band stretches from 0.4 to 0.7 μm (**Figure 3(a)**). This spectral range constitutes an extremely small portion of the full electromagnetic spectrum and its boundaries are defined by the wavelength limits of human vision (**Figure 3(b)**). The visible spectrum is further subdivided into three equal segments: blue (0.4–0.5 μm), green (0.5–0.6 μm), and red (0.6–0.7 μm).

The IR band extends from 0.7 to 1000 μm (**Figure 3(a)**). In remote sensing, this wide band is divided into two major components: the reflected IR band (from 0.7 to 3.0 μm) and the emitted, or thermal, IR band (from 3.0 to 1000 μm). The first component is essentially solar radiation reflected from the Earth's surface. This component is further subdivided into near-IR (<1.3 μm) and mid-IR (>1.3 μm). Radiation in the near-IR behaves, with respect to optical systems, in a manner analogous to radiation in the visible spectrum. It can therefore be detected and recorded by films and cameras similar to those used for the visible spectrum. The thermal component of the IR spectrum represents heat energy that is continuously emitted by all objects on the Earth's surface. It ranges from 3 to 1000 μm, but wavelengths beyond 14 μm are largely absorbed by the atmosphere and are therefore not available for remote sensing (**Figure 3(b)**).

The microwave band falls between the IR and the radio bands (**Figure 3(a)**). It ranges from approximately 1 mm to 1 m and contains the longest wavelengths used in remote sensing. An important feature of microwave radiation is its ability to pass through clouds, precipitation, and tree canopies.

Types of Remote Sensors

The sensors used in remote sensing can be divided into four main groups:

1. Photographic camera sensors.
2. Electro-optical sensors.
3. Passive microwave sensors.
4. Active sensors.

Photographic cameras can detect wavelengths ranging from approximately 0.3–0.9 μm (**Figure 3(b)**). This includes the near-UV band (0.3–0.4 μm), all the visible light (0.4–0.7 μm), and the relatively short wavelengths of the near-IR (0.7–0.9 μm). This spectral range, which is nearly twice as wide as the human vision, can be recorded directly into film and is therefore referred to as the photographic spectrum.

Electro-optical sensors measure radiation in narrow wavelength ranges that can be located at various points along the near-UV, visible, reflected IR, and thermal IR spectral bands

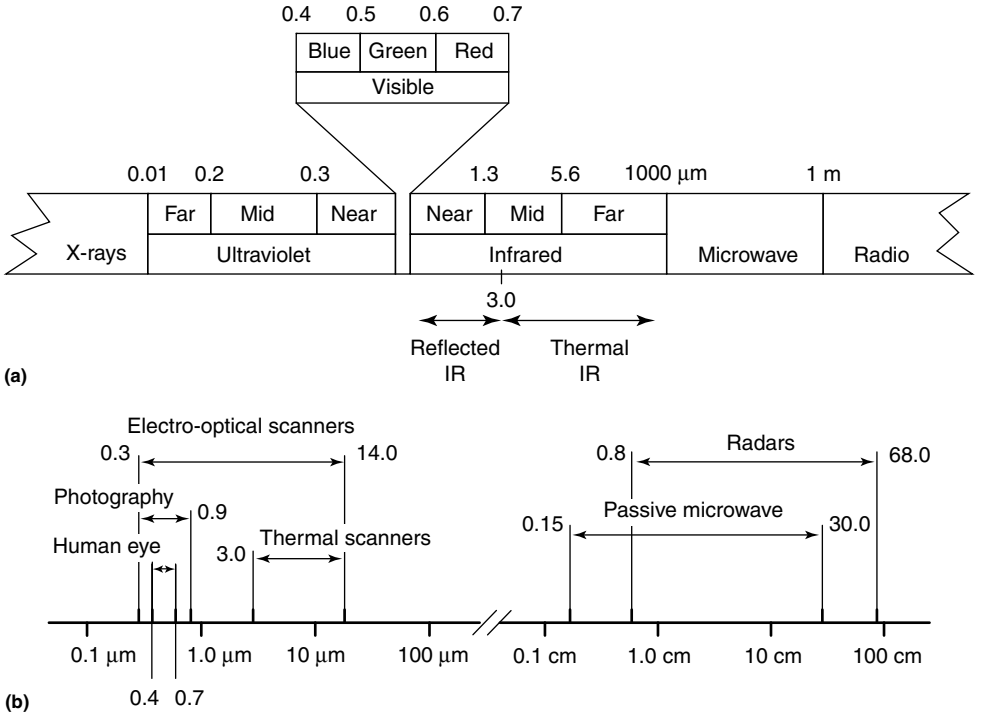


Figure 3 Characteristics of electromagnetic radiation. (a) Major divisions of the electromagnetic spectrum (not to scale). (b) Spectral ranges of common remote-sensing systems.

(Figure 3(b)). This spectral range (0.3–14 μm) is termed the optical spectrum because wavelengths in this range can be reflected and refracted using optical devices such as lenses and mirrors. An important characteristic of electro-optical sensors operating at the thermal IR band is their ability to collect data day and night without being sensitive to dust and haze.

Passive microwave sensors are designed to measure microwave radiation that is naturally emitted from the Earth's surface. These sensors operate at a spectral range of 0.15–30 cm (Figure 3(b)).

Active sensors emit radiation and measure the amount of radiation returned to the sensor. Radar systems are active sensors that transmit artificially produced energy to the Earth's surface and record the reflected component of this radiation. Radar can penetrate through clouds in contrast to optical systems that are severely challenged by cloud cover. Light detection and ranging laser (LIDAR) is another type of active sensor that emits radiation in optical wavelengths and can measure the distance to a target on the surface. LIDAR systems cannot penetrate through clouds.

Three forms of remote sensing can therefore be distinguished: the detection of solar radiation reflected from the Earth's surface, the detection of radiation emitted from the Earth's surface, and the detection of radiation transmitted from the sensor itself.

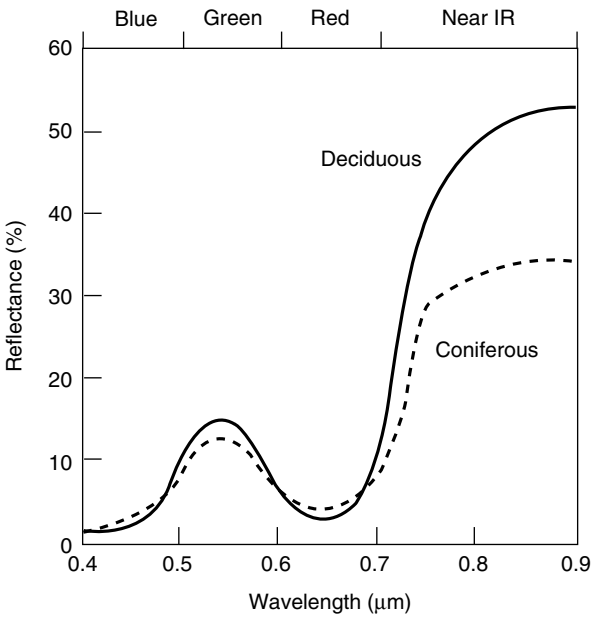


Figure 4 Generalized spectral reflectance curves for deciduous (broad-leaved) and coniferous (needle-bearing) trees. Note that both reflectance intensity and differentiation ability are highest in the near-IR region.

Sources of Remotely Sensed Data

Very High-Resolution Images

Aerial photography is the oldest form of remote sensing. Despite the increasing availability of more sophisticated imaging systems, aerial photographs remain one of the most reliable and widely used sources of remotely sensed data. High spatial resolution enables one to detect and analyze patterns at very fine spatial scales. The availability of historical aerial photographs makes it possible to measure and analyze long-term ecological processes.

Both black-and-white and color photographs are used in remote sensing. Historically, black-and-white photographs were taken with either panchromatic film or IR-sensitive film. Panchromatic film is characterized by an aggregate spectral sensitivity that includes the near-UV, blue, green, and red wavelengths (0.25–0.7 μm). It provides a tonal rendition that closely approximates the brightness of the scene being photographed and can be used to distinguish between objects of truly different color. IR-sensitive film has spectral sensitivity that extends from 0.25 to approximately 0.9 μm. This range encompasses the near-IR band in addition to the near-UV and visible wavelengths. Thus, IR-sensitive film operates beyond the confines of the visible spectrum and can be used to detect information that is unavailable for human vision. IR-sensitive films are especially valuable for vegetation classification because differences between vegetation classes are often more distinct in the near-IR than in the visible range (Figure (4)).

Although the information contained in black-and-white photographs is expressed by differences in gray levels, color films offer the additional qualities of hue (dominant wavelength), chroma (color strength), and value (color intensity). This added information greatly improves interpretation ability. Two kinds of color films historically used in aerial photography are normal color films and color IR films. Normal color films are sensitive primarily to wavelengths of the visible spectrum and provide a color rendition that approximates the original scene as it would be viewed by the human eye. This facilitates photo interpretation because objects appear in their natural colors. Color IR films are sensitive to green, red, and near-IR radiation (0.5–0.9 μm). Because maximum reflectance from vegetation occurs in the near-IR spectral region, such films are highly valuable for studies of vegetated landscapes. At present, aerial photos generally produce digital images with the digital values representing the electromagnetic radiation in different wavelengths (*see* Sources of Remotely Sensed Data: Satellite Imagery). Many types of sensors can be mounted on airplanes, including hyperspectral sensors to measure reflectance in many wavelengths and active LIDAR sensors.

A number of commercial satellites have been launched to obtain data at very high spatial resolutions (5 m or less) with panchromatic and multispectral bands, such as IKONOS (named after the Greek term for "image") and Quickbird. These levels of spatial resolution approximate the resolution obtained by standard aerial photography.

Satellite Imagery

Data acquired by spaceborne imaging systems are usually recorded in digital form and are therefore available as digital images. A digital image is composed of a two-dimensional array of picture elements (pixels), with each pixel corresponding to a particular area on the ground. The radiance measured over the ground area represented by each pixel is translated into a digital number (DN). This number (also referred to as brightness) is a positive integer expressing the

average intensity of radiance measured from the particular ground area. Two important characteristics of digital images are spatial resolution and radiometric resolution. Spatial resolution refers to the size of the ground area represented by each pixel in the image. For example, a spatial resolution of 10 m × 10 m means that each pixel in the image corresponds to a ground unit of 10 m × 10 m. The term radiometric resolution refers to the number of DN values, or brightness levels, used for recording differences in radiation intensity. Thus, a higher radiometric resolution permits the detection of finer differences in radiation intensity. Both spatial and radiometric resolution has important consequences for the amount of information available in remotely sensed data.

Satellite images suitable for mapping, monitoring, and modeling patterns of biodiversity have usually been derived from multispectral sensors measuring reflected or emitted radiation in the optical spectrum (0.3–14 μm). Imaging systems operating at this spectral range have been used in a variety of national and international Earth observation programs. Three of these programs, namely, Landsat, Systeme Pour l'Observation de la Terre (SPOT), and National Oceanic and Atmospheric Administration (NOAA) Advanced Very High-Resolution Radiometer (AVHRR), provide multidecadal observations and are of particular importance for studies of biodiversity. The main characteristics of these programs are outlined below. Many other sensors are useful for biodiversity, including multispectral optical sensors MODIS (Moderate Resolution Imaging Spectroradiometer) and MERIS (Medium Resolution Imaging Spectrometer), radar, hyperspectral, and LIDAR (Light Detection and Ranging). The three programs discussed below (Landsat, SPOT, and coarse resolution sensors) provide the longest time series, which is often important in biodiversity studies to assess change in habitat. Emerging technologies that are promising for biodiversity studies are also discussed.

Landsat Program

The Landsat program is the longest running project for acquisition of repetitive multispectral data of the Earth's surface. The first Landsat satellite was launched in 1972 and the most recent one (Landsat 7) in 1999. Landsats 1–3 covered the Earth every 18 days. Landsats 4, 5, and 7 have a coverage cycle of 16 days (Landsat 6 has failed to achieve orbit). At present (2011), the only active satellites in this program are Landsats 5 and 7. However, data recorded by earlier Landsat missions are still being used in many remote-sensing applications.

Table 1 compares the characteristics of data acquired by the two main sensors used in the Landsat program, the Multispectral Scanner (MSS) and the Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM). The MSS collects reflected solar radiation in four contiguous spectral bands, two in the visible spectrum at 0.5–0.6 μm (green) and 0.6–0.7 μm (red), and two in the near-IR at 0.7–0.8 and 0.8–1.1 μm. This imaging system was used by Landsats 1–5 (note that the same spectral bands had different numbering systems in Landsats 1–3 and Landsats 4 and 5). Data collected by the MSS are framed into individual scenes that cover a nominal ground area of 185 km × 185 km at a spatial resolution of 79 m in Landsats 1–3 and 82 m in Landsats 4 and 5 (**Table 1**).

Table 1 Characteristics of Landsat MSS and TM Data

Band	*Wavelength range (μm)*	*Nominal spectral location*	*Spatial resolution (m)*
MSS[a]			
4 (1)	0.5–0.6	Green	79 (82)
5 (2)	0.6–0.7	Red	79 (82)
6 (3)	0.7–0.8	Near-IR	79 (82)
7 (4)	0.8–1.1	Near-IR	79 (82)
TM and ETM			
1	0.45–0.52	Blue-green	30
2	0.52–0.60	Green	30
3	0.63–0.69	Red	30
4	0.76–0.90	Near-IR	30
5	1.55–1.75	Mid-IR	30
6[b]	10.4–12.5	Thermal IR	60,120
7	2.08–2.35	Mid-IR	30
8[c]	0.52–0.90	Panchromatic	15

[a]MSS bands of Landsats 1, 2, and 3 are numbered 4–7. The same bands are numbered 1–4 in Landsats 4 and 5. Spatial resolution is 82 m in Landsats 1–3 and 30 m in Landsats 4, 5, and 7 (thermal band 60 m).
[b]The thermal band is 120-m resolution on TM and 60-m resolution on ETM.
[c]Panchromatic band 8 was added to ETM on Landsat 7.

The TM was designed to improve the imaging capabilities of Landsat satellites. A major difference between the TM and the MSS is the acquisition of data in seven bands instead of four, with new bands in the visible (blue-green), mid-IR, and thermal portions of the spectrum (**Table 1**). Also, based on experience with MSS data and extensive research, the wavelength range and location of the TM scanner have been modified to facilitate the spectral differentiability of major Earth surface features. The radiometric resolution of the TM is higher than that of the MSS (256 vs 64 DNs) and its spatial resolution is also higher (30 m for all bands except the thermal band, as compared with approximately 80 m in the case of the MSS). These overall capabilities significantly improved the spectral coverage, radiometric resolution, and spatial resolution of Landsat images. The ETM has a panchromatic band at a 15-m resolution.

Since 2008, all images in the Landsat archive are available for download at no charge.

SPOT Program

SPOT is an international Earth observation program initiated by the French government in the late 1970s. The first SPOT satellite was launched in 1986, and the most recent one (SPOT 5) in 2002. All satellites have an Earth-coverage cycle of 26 days. The sensors of SPOT 1, 2, and 3 consist of two identical high-resolution visible (HRV) imaging systems, each designed to operate in either of two modes: a black-and-white panchromatic mode over the range 0.51–0.73 μm (green–red) and a multispectral mode over the ranges 0.50–0.59 μm (green), 0.61–0.68 μm (red), and 0.79–0.89 μm (near-IR). Spatial resolution of the panchromatic mode is 10 m, whereas that of the multispectral mode is 20 m. SPOT 5 has a higher resolution of 2.5–5 m in the panchromatic mode, 10 m in multispectral mode, and 20 m for short-wave IR band. The data are encoded over a range of 256 DNs, and individual

images are formatted to cover a 60 km × 60 km area. An important advantage of the HRV imaging system is the ability to adjust its viewing angle by ground commands, and thus to observe the same location several times within the normal 26-day coverage cycle. This "revisit" capability facilitates the documentation and monitoring of short-term dynamic processes (e.g., vegetation responses to episodic rainfall or drought events). It also increases the chance of obtaining cloud-free images in areas where cloud cover is a problem.

Coarse-Resolution Sensors

The NOAA of the USA operates national weather programs that rely to a large extent on data collected by meteorological satellites. Imaging systems of such satellites are usually designed to have a relatively low spatial resolution, but very high temporal resolution. Satellites of the NOAA series, one of the longest weather satellite programs, observe the complete Earth twice a day at a ground resolution of 1.1 km × 1.1 km at nadir. These satellites carry a multispectral scanner known as the AVHRR. The AVHRR was designed primarily for meteorological applications, but it has been applied successfully for many land-oriented applications. Details of the specific spectral coverage vary with the specific mission, but in general, AVHRR sensors collect data in four or five bands falling in the visible, near-IR, and thermal IR spectral regions. Images acquired by AVHRR are available at a full 1.1-km resolution (local area coverage) and at a subsampled resolution of 4 km (global area coverage). In 2000, the US National Aeronautics and Space Administration (NASA) launched the MODIS sensor as part of Earth Observations System platform. MODIS acquires daily images at 250 m in the red and IR bands, 500 m at other spectral bands, and 1 km in thermal bands. MERIS is a European coarse-resolution sensor launched in 2002, but data are less easily available than from AVHRR or MODIS.

Emerging Technologies

Several remote-sensing technologies are not as well developed as the long-term time series (Landsat, SPOT, and coarse resolution sensors) but research is demonstrating the utility of these emerging technologies for biodiversity studies. Hyperspectral scanners are imaging systems that acquire multispectral images in many, very narrow, contiguous spectral bands throughout the visible, near-IR, and mid-IR portions of the spectrum. This spectral capability permits discrimination among objects that have fine diagnostic reflection characteristics that are "lost" within the bands of conventional electro-optical scanners. For example, Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), one of the main hyperspectral scanners, collects data in 224 contiguous bands between 0.4 and 2.45 μm.

Other emerging technologies are active LIDAR and radar sensors. These sensors provide information on three-dimensional vegetation structure that can indicate characteristics of habitats. Airborne sensors have tested these technologies. To date (2011) hyperspectral and LIDAR sensors are not flown on spaceborne satellites. Some satellites obtain radar data, such as the Advanced Land Observation Satellite (ALOS), which was launched in 2006 and lost power in 2011.

Imaging radars have been used for remote sensing since the late 1970s. However, most of these applications were of experimental nature. For example, the Shuttle Imaging Radar carried by the space shuttle Columbia in its second flight was used to collect data from different continents in an attempt to construct a library of images representing a wide variety of environments differing with respect to climate, geology, vegetation, and other qualities. More recent studies have evaluated the potential of radar data for vegetation mapping, detection of deforestation, determination of vegetation biomass, and estimation of various forest parameters. Such applications have led to a better understanding of the capabilities of radar imagery and demonstrated that this form of remote sensing may complement many characteristics of images based on the optical spectrum.

Analysis of Remotely Sensed Data

Data obtained from remote sensing must be interpreted in order to become usable information. The outcome of the interpretation process is usually a map showing the spatial distribution of the objects or variables of interest. Such maps can be analyzed with respect to other sources of information using geographical information systems (GIS) (*see* Integration of Digital Images within a GIS Environment).

Interpretation of Remotely Sensed Data

In the case of aerial photographs, the interpretation process is often based on visual examination: The interpreter systematically examines the image (and possibly supporting materials such as maps or field reports) and tries to identify the nature of the objects and phenomena seen in the image. This form of image interpretation relies on the ability of the human mind to qualitatively evaluate spatial patterns in an image. In most cases, it involves the identification and delineation of discrete areal units throughout the image. For example, in studies of vegetation variation, the interpreter outlines the boundaries between areas covered by different vegetation types.

Digital images are usually interpreted using computer-based image processing. Recall that the brightness of each pixel in a digital image is expressed by a numerical integer (its DN value); the general objective of image processing is to "translate" the original image into a more informative image by mathematically manipulating the recorded DN values. Practically this is done by subjecting the original image, pixel by pixel, to appropriate mathematical operations and storing the results of the computation as a new image. The mathematical algorithms used in image processing are literally infinite, but most kinds of algorithms can be categorized into one of three main operations: image restoration, image enhancement, and thematic classification.

Image restoration involves various operations of spectral and geometric corrections. Digital images are noisy and suffer from different kinds of spectral and geometric distortions that reduce their quality and limit their use. For example, the spectral characteristics of objects detected in a digital image are distorted by differences in sun angle, topographic variation, atmospheric scattering, and sensor characteristics. All of these factors may influence the intensity of electromagnetic radiation measured by the sensor. Other factors may introduce

positional errors to digital images. Radiometric corrections are mathematical operations that compensate for various sources of spectral distortion in the data. Geometric corrections attempt to correct for positional errors and to transform the original image into a new image that has the geometric characteristics of a map.

Image-enhancement operations attempt to improve the detectability of objects or patterns in the image. For example, contrast-stretching algorithms increase the contrast between features of interest and thus improve interpretation ability. Edge enhancement algorithms are designed to automatically detect boundaries between different kinds of features. Other algorithms operate on multispectral data. For example, vegetation indices (VIs) represent a group of spectral operations that attempt to enhance vegetation mapping and monitoring. The most widely used VI is the normalized difference vegetation index (NDVI) defined by the general equation

$$\text{NDVI} = \frac{\text{near-IR band} - \text{red band}}{\text{near-IR band} + \text{red band}}$$

This index can be computed using data acquired by a variety of imaging systems (e.g., Landsat, SPOT, NOAA AVHRR, and MODIS) and has proved to be a sensitive indicator of the presence and condition of green vegetation.

Thematic classification is a process of assigning all pixels of an image into land-cover classes (themes) based on their DN values. In applications such as land cover or vegetation mapping, thematic classification is the main object of the analysis. The outcome of the classification process is a new image where each pixel represents a particular land-cover type. Such an image can be interpreted as a map depicting the spatial distribution of the selected land-cover types within the scene (**Figure 1**).

Two standard approaches for image classification are unsupervised classification and supervised classification. In unsupervised classification the pixels of the image are numerically classified into a number of spectral classes based on the likeness of their DN values. The land-cover identity of each spectral class is then determined using reference information such as field data, aerial photographs, or published maps. In supervised classification, the analyst first selects areas of known cover type in the image. These "training sites" are analyzed statistically to identify the spectral characteristics of each cover type. With appropriate algorithms, the DN values of each unknown pixel in the image are compared with those of the training sites, and the pixel is assigned to the land-cover type to which it is most similar. It is important to emphasize that classification of digital images is never perfect. The success of image classification depends on many factors and may vary considerably from one case to another. Therefore, any classification of remotely sensed data must be followed by some form of accuracy assessment.

Traditional classification approaches are based on the spectral values of a pixel. The pixel-based approach classifies each pixel independent of the neighboring pixels. In contrast, object-oriented classification considers the spectral characteristics of a pixel along with spatial information form texture and context information in the image. The image is segmented into objects that form the classification units. Objects rather than single pixels are considered in the classification process.

Integration of Digital Images within a GIS Environment

GISs are computer-based systems capable of storing, manipulating, and visualizing geographically referenced data. Basically, any kind of information that can be presented by standard maps can be converted into digital form and handled by GIS. An important benefit of GIS is the ability to spatially integrate multiple types of data stemming from different sources of information. A typical GIS operation is overlay analysis. Many other kinds of spatial operations can be performed with GIS. For example, aggregation is an operation by which detailed map categories are aggregated into less detailed categories (e.g., aggregation of detailed vegetation units into a few broad vegetation types). Buffering is an operation by which a zone of user-defined width is created around map objects. Data merging is an operation by which different grid maps of the same area are coregistered into a uniform, multidimensional database. This operation is commonly applied to combine digital images representing different wavelengths, different times of observation, or different sensors. For example, monitoring seasonal variation in NDVI requires the coregistration of images taken at different times. Low-resolution images (e.g., AVHRR) can be merged with images of higher resolution (e.g., Landsat TM) for calibration or interpretation purposes. Using the merging capability of GIS, one is able to coregister digital images with "ancillary" data (e.g., elevation and slope) that correspond to each pixel in the image. Once such data are spatially registered, information available in the GIS can be used to improve image classification. Thus, the interaction between remote sensing and GIS is two way in nature; remote sensing is used to generate digital maps that can be fed into GIS, whereas GIS data can be used to interpret and classify remotely sensed data.

GIS can also be used for modeling purposes. For example, by overlaying vegetation maps derived from remote sensing with corresponding maps of climatic and topographic variables, one is able to develop regression models that predict vegetation characteristics from information on climate and topography.

Applications of Remote Sensing for Studying Biodiversity

The previous section (Analysis of Remotely Sensed Data) has focused on general concepts and elements of remote sensing. Its main purpose was to provide the scientific and technical background required to understand how remote-sensing technology can be used for collecting, analyzing, and visualizing images of Earth's surface. This section presents three major applications to illustrate the wide spectrum of biodiversity issues that can be studied with remote sensing.

Three main kinds of applications are: (1) mapping of land cover and vegetation patterns, (2) measurement of temporal changes in land cover and vegetation characteristics, and (3) modeling of species distribution patterns. The applications can occur at global, regional, or local scales with increasing amounts of spatial detail. This trade-off between spatial scale and spatial resolution is a common characteristic of remote-sensing applications. Temporal intervals between cloud-free successive satellite images of the same system range from daily to annually and spatial resolutions range from meters to

kilometers. Generally an application can cover large areas at coarse spatial resolution or small areas at fine spatial resolution. These differences point to the diversity of spatial and temporal scales that can be investigated using remotely sensed data.

Remote sensing provides proxies for patterns and changes in biodiversity. Species are not directly observable with remote sensing. Remote sensing of habitat distribution and quality provide indirect measures of biodiversity patterns (Hansen *et al.*, 2010).

Mapping Vegetation and Land-Cover Patterns

Mapping vegetation and land-cover patterns has been the most common application of remote sensing. It may contribute to studies of biodiversity both directly (by documenting the distribution of different components of biodiversity) and indirectly (by providing knowledge on landscape and habitat characteristics that influence patterns of biodiversity). Land-cover maps categorize each grid cell according to a predefined

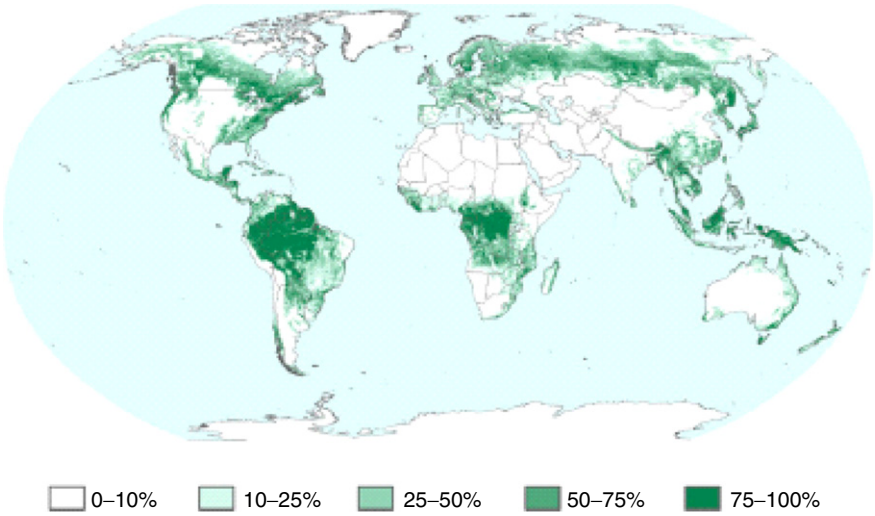


Figure 5 Global-scale estimated percent tree cover (reproduced from Hansen M, Stehman S, and Potapov P (2010) Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences of the United States of America* 107: 8650–8655).



Figure 6 Four-meter multispectral IKONOS data for La Selva Biological Station, Costa Rica, for mapping tree crowns (reproduced from Clark D, Read J, Clark M, Cruz A, Dotti M, and Clark D (2004) Application of 1-m and 4-m resolution satellite data to ecological studies of tropical rain forests. *Ecological Applications* 14: 61–74).

classification system, either as a discrete cover type or as a mixture of vegetation characteristics such as woody, grass, or bare ground.

Global land-cover maps are generally derived from data with relatively coarse spatial resolution and high temporal resolution. Discrimination among cover types is based on spectral signatures of different cover types and the temporal differences in phenology. Global land-cover maps are useful for indicating biome-level diversity and broad-scale patterns of where different species might occur. At a finer spatial resolution, data from the Landsat sensor have been the workhorse for mapping vegetation patterns. Individual scenes cover 185 km × 185 km. Land-cover maps over larger areas can be constructed by mosaics of multiple images although technical difficulties arise if the images are acquired on different dates. At an even finer spatial resolution, very high-resolution satellite or aerial data can be used to map land-cover patterns in more detail but over relatively small areas. The appropriate data and scale for land-cover maps depend on the application for broad-scale biodiversity studies (global and regional scale maps) or fine-scale, detailed maps for biodiversity patterns in a specific place. **Figures 5** and **6** illustrate the range of scales from global distribution of forest cover using MODIS data (**Figure 5**) to mapping tree crowns from very high-resolution IKONOS data (**Figure 6**).

Land-cover maps produced from remote sensing can be combined in a GIS with other layers such as precipitation, temperature, soil type, and human settlements to infer species distributions.

Measuring Vegetation Changes

Identifying patterns, scales, and trends of vegetation changes is important for understanding the dynamics and structure

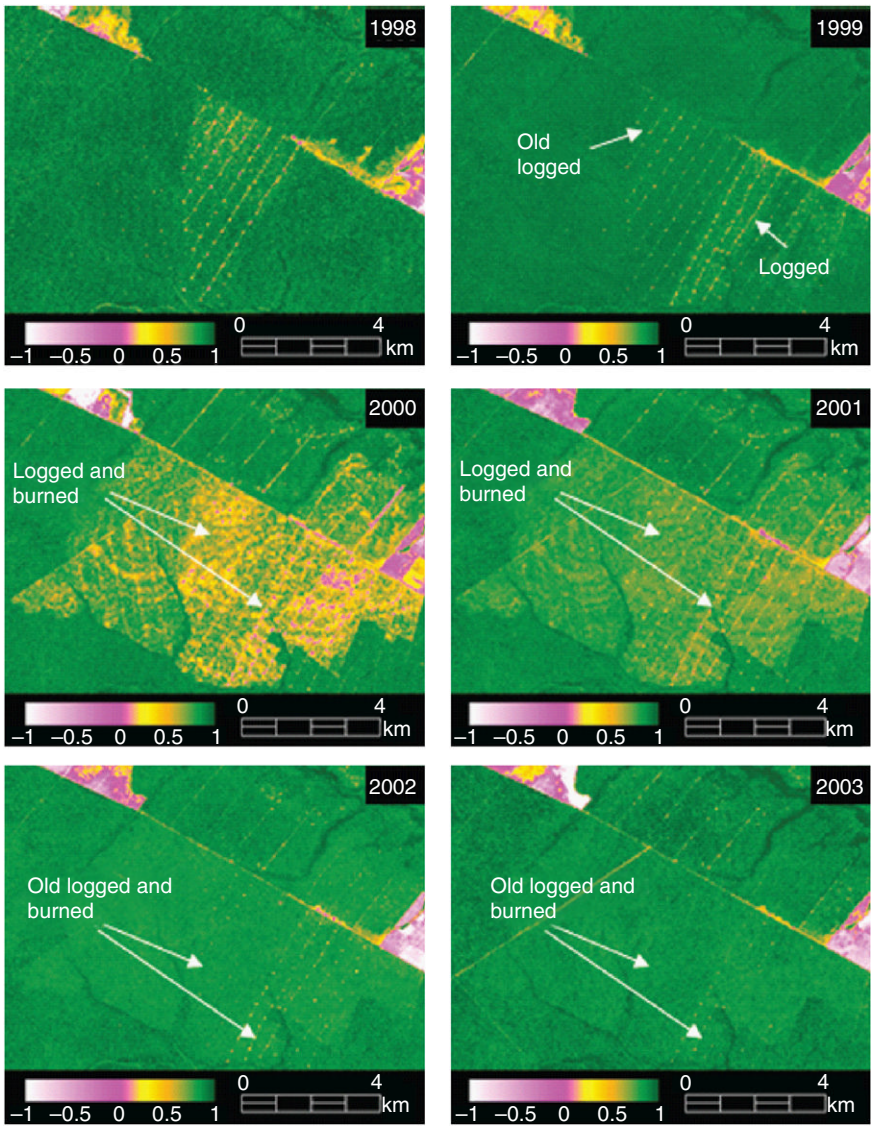


Figure 7 Annual change from 1998 to 2003 in forest degradation from selective logging and burning in Sinop region, Mato Grosso state, Brazil, depicted from Landsat data. Graphic courtesy of Carlos Souza, Jr. (DeFries, 2008).

of both plant and animal communities. As with mapping applications, measurements of vegetation changes based on remotely sensed data may provide direct evidence for changes in biodiversity (e.g., changes in the relative frequency of different vegetation types) as well as information on environmental changes that may influence biodiversity patterns (e.g., deforestation and forest fragmentation). Mapping deforestation is a common application for remote sensing. Other more subtle changes in vegetation are also observable, such as logging and burn scars (**Figure 7**).

Measuring vegetation changes requires a time series of remote-sensing data that is calibrated to eliminate artifacts from changes in the sensor or atmospheric conditions. The long time series of Landsat data are commonly used to map change. At coarser resolutions, AVHRR provides a multidecadal time series and can be used to identify interannual variability and trends in vegetation from climate variability.

Modeling Species Distribution and Species Richness

Although most plant and animal species cannot be detected directly by remote sensing, the integration of data obtained from remote sensing with field-based observations of plants and animals enables one to develop statistical models that predict distribution patterns of individual species from information on the environment. A similar approach can be applied to model and predict patterns of species diversity.

Data from hyperspectral sensors are potentially useful for mapping species distributions based on the notion that species vary in upper-canopy pigments, water, and nitrogen content. In a study in Hawaiian lowland rainforests, species richness was significantly correlated with spatial variation in reflectances at wavelengths for measuring pigments (530 and 720 nm), canopy water (1201 nm), and leaf water/nitrogen (1523 nm) acquired by the airborne hyperspectral AVIRIS sensor. The results suggest that increasing spectral diversity is linked to increasing biochemical diversity which in turn is linked to plant species richness (Carlson *et al.*, 2007), demonstrating the utility of remote sensing for mapping diversity at high spatial resolution (**Table 2**).

LIDAR and radar data provide information about the three-dimensional characteristics of habitats, making it possible to model species distributions. Canopy structure is determined from the backscatter of these active sensors, making it possible to map structure related to habitats for different animal species (**Figure 8**).

Table 2 Pearson correlation coefficients (r) between woody species richness (RICH) and the AVIRIS reflectance derivatives at 530 and 720 nm (pigments), 1201 nm (canopy water), and 1523 nm (leaf water/nitrogen)

Variable	*RICH*	R_{d530}	R_{d720}	R_{d1201}	R_{d1523}
RICH	–				
R_{d530}	0.84	–			
R_{d720}	0.87	0.87	–		
R_{d1201}	0.77	0.74	0.70	–	
R_{d1523}	0.91	0.81	0.90	0.79	–

Relationships are significant at the $p < .01$ level.

Source: Reproduced from Carlson K, Asner G, Hughes R, Ostertag R, and Martin R (2007) Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems* 10: 536–549, with permission from ESA.

Other Applications

The applications presented above exemplify various aspects of biodiversity that can be investigated using remote sensing. Three additional aspects, not covered by these case studies, are mentioned here to provide a more complete picture of the subject. These are fire, grazing, and canopy gap dynamics.

Fire is a major source of disturbance in many kinds of ecosystems. Studies of fire distribution in space and time may therefore contribute to our understanding of the dynamics and structure of ecological communities. Active fires are detected using thermal wavelengths from the MODIS sensor and made publicly available (**Figure 9**). Burn scars are also mapped from optical sensors such as Landsat and MODIS.

Another aspect of biodiversity that has been studied extensively with remote sensing is grazing. Data obtained from remote sensing have been used to investigate vegetation responses to both natural and domestic grazing. Some studies have focused on spatial patterns (e.g., differences between fenced and nonfenced areas), whereas other studies were designed to quantify temporal vegetation changes caused by grazing (e.g., biomass removal, replacement of herbaceous vegetation by woody species, or shrub encroachment).

Canopy gaps caused by the death of one or more trees are a major disturbance mode in many temperate and tropical ecosystems. Such disturbances may have considerable effects on the diversity of plant and animal communities. Traditionally, studies of canopy gap dynamics have focused on simple "static" parameters such as gap size distribution and total gap area. An alternative (and more direct) approach for studying gap dynamics is to analyze time sequences of remotely sensed data.

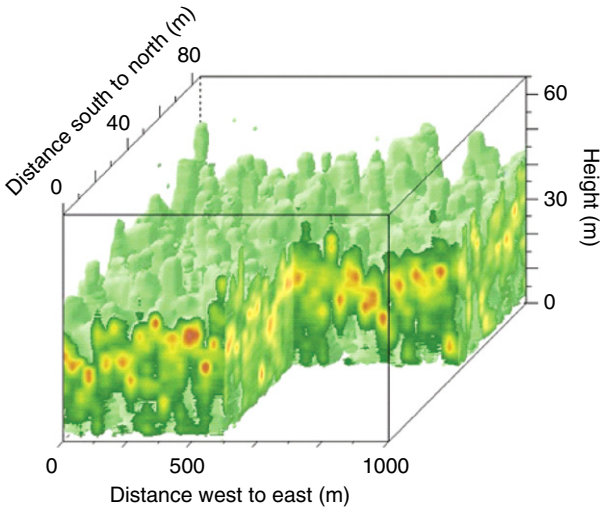


Figure 8 Canopy structure of an approximately 1 km^2 area of La Selva Biological Research Station, Costa Rica, derived from airborne LIDAR sensor in 1998. Color is intensity of the LIDAR return at each height (green is lowest, red is highest). Graphic courtesy JE Weishampele (DeFries, 2008).

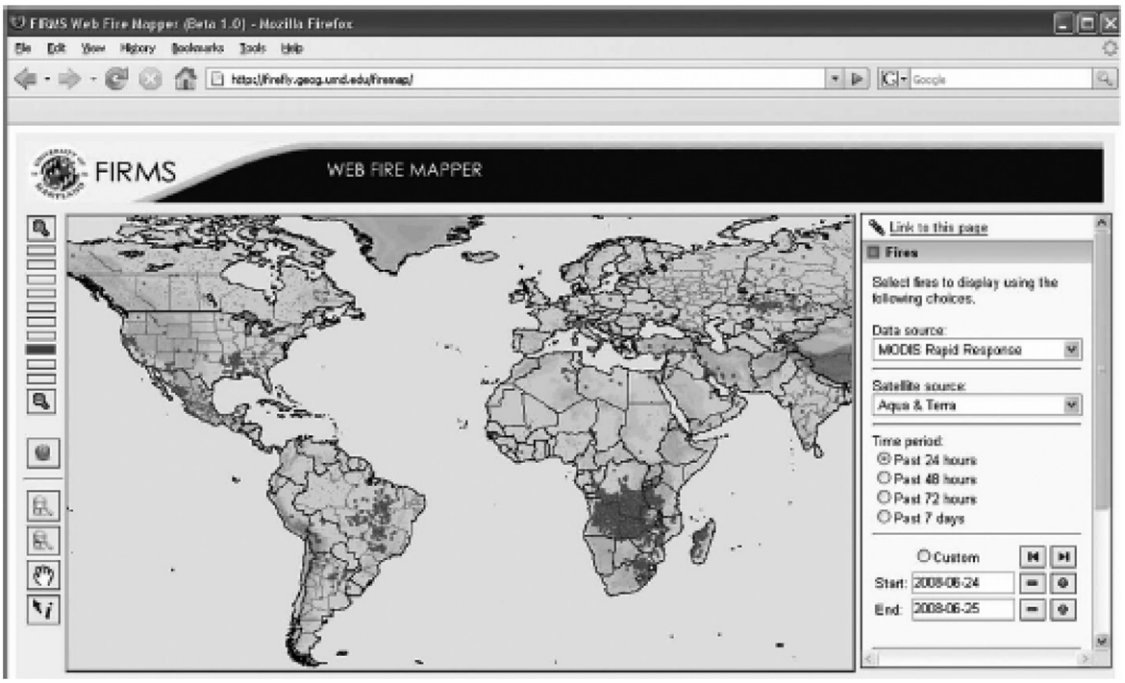


Figure 9 A screenshot of the open source version of Web Fire Mapper that distributes data on active fires detected by MODIS (reproduced from Davies D, Ilavajhala S, Wong M, and Justice C (2009) Fire information and resource management system: Archiving and distributing MODIS active fire data. *IEEE Transactions on Geoscience and Remote Sensing* 47: 72–79).

Summary

Remote sensing in general, and Earth observation satellites in particular, have brought a new dimension to understanding of the Earth's surface. Since the launch of the first Landsat satellite in 1972, millions of images of the Earth have been acquired from spaceborne imaging systems. Earth observation satellites have provided images over spatial scales ranging from local regions to global cover, at spatial resolution of a few meters to tenths of a kilometer, and at spectral wavelengths ranging from near-UV to microwave radiation. Some of these data have been acquired systematically and repetitively over many years, with temporal resolution ranging from hours to a few weeks. In addition, most of the Earth observation satellites were designed to provide multispectral images of the Earth's surface. This enormous amount of information has been utilized extensively for studies of biodiversity. As demonstrated in this article, data obtained from remote sensing have been used for mapping various components of biodiversity, measuring different ecosystem characteristics (e.g., forest structure), analyzing factors affecting the distribution of plant and animal species, monitoring natural and human-induced vegetation changes (e.g., succession and deforestation), and evaluating the impact of various disturbances (e.g., fire) and ecological interactions (e.g., grazing).

Recent advances in remote-sensing technologies and computing power have added to the toolkit from aerial photography and multispectral optical data. Hyperspectral and very high-resolution optical sensors as well as active LIDAR and radar sensors are increasing the contributions of remote sensing to the study of biodiversity. There are many reasons to believe that in the near future much more information on biodiversity will be obtained from remote sensing. In parallel to these advances in space technology, new analytical methodologies are being developed for the analysis and interpretation of data from very high-resolution, hyperspectral, and active sensors. One can therefore expect a considerable growth in the contribution of remote sensing to studies of biodiversity during the next decade.

See also: Computer Systems and Models, Use of. Measurement and Analysis of Biodiversity. Natural Reserves and Preserves

References

Avery TE and Berlin GL (1992) *Fundamentals of Remote Sensing and Airphoto Interpretation*, 5th edn. Englewood Cliffs, NJ: Prentice Hall.

Carlson K, Asner G, Hughes R, Ostertag R, and Martin R (2007) Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems* 10: 536–549.

Clark D, Read J, Clark M, Cruz A, Dotti M, and Clark D (2004) Application of 1-m and 4-m resolution satellite data to ecological studies of tropical rain forests. *Ecological Applications* 14: 61–74.

Davies D, Ilavajhala S, Wong M, and Justice C (2009) Fire information and resource management system: Archiving and distributing MODIS active fire data. *IEEE Transactions on Geoscience and Remote Sensing* 47: 72–79.

DeFries R (2008) Terrestrial vegetation in the coupled human–earth system: Contributions of remote sensing. *Annual Review of Environment and Resources* 33: 369–390.

Hansen M, Stehman S, and Potapov P (2010) Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences of the United States of America* 107: 8650–8655.

Reptiles, Biodiversity of

F Harvey Pough, School of Life Sciences, Rochester Institute of Technology, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Ancestral Describes a character or character state of the organism being considered that retains the primitive condition for its evolutionary lineage.

Derived Describes a character or character state of the organism being considered that has changed from the ancestral condition for its evolutionary lineage.

Ectothermy Deriving the energy needed to raise body temperature from sources outside the body.

Endothermy Deriving the energy needed to raise body temperature from within the body – i.e., from metabolic heat production.

Heliothermic Regulating body temperature primarily by moving between sun and shade.

Operative temperature A measure of environmental temperature that combines the effects of heat exchange via radiation, convection, with metabolic heat production, and evaporative heat loss.

Paraphyletic A taxonomic grouping of animals that does not meet the cladistic criterion of including the most recent common ancestor and all its descendants.

Sister group The evolutionary lineage most closely related to the one being discussed.

The widespread adoption of phylogenetic systematics (cladistics) as the basis for taxonomic designations has had more impact on our understanding of reptiles than on any other group of terrestrial vertebrates. As our knowledge of ancestral and derived characters of reptiles has grown and been applied over the past decade, our perception of evolutionary relationships of lineages within the Reptilia has changed, and some lineages have even moved in and out of the Reptilia. Our current understanding of the relationships of extant reptiles recognizes three major monophyletic lineages (Pough *et al.*, 2004): Turtles, crocodilians and birds, and lepidosaurs (with tuatara, lizards, and snakes as the extant representatives). Birds are derived from the lineage that includes crocodilians as well as dinosaurs, pterosaurs, and several other groups of extinct reptiles (Chiappe, 2007), but the anatomical and physiological specializations seen at the avian grade (most notably endothermy, insulation, and flight) make birds functionally different from other reptiles. As a result, the precladistic allocation of nonavian reptiles to the field of herpetology and avian reptiles to ornithology is useful, and will be followed here. Nonetheless, the evolutionary origins of the characters of birds provide a useful context for considering the present and past diversity of other reptiles.

Extant Evolutionary Lineages

Reptiles, exclusive of birds, include about 9300 living species. Our understanding of the phylogenetic relationships of the group is in flux. The Reptile Database (http://www.reptile-database.org/) is a convenient reference and except where other sources are cited, the author has followed its January 2011 taxonomic divisions. The vast majority of nonavian reptiles (about 9000 species) are lepidosaurs (tuatara, lizards, and snakes). There are approximately 315 living species of turtles and 23 crocodilians (Florida Museum of Natural History http://www.flmnh.ufl.edu/cnhc/). Species diversity is greatest in the tropics, but reptiles are not limited to warm regions – temperate deserts in the Holarctic, Africa, and Australia have rich lizard faunas, and temperate aquatic habitats in North America and Asia are home to many species of turtles. Reptiles occur on every continent except Antarctica, occupy habitats ranging from freshwater swamps to desert salt flats, and exploit ecological zones from subterranean to arboreal.

Turtles (Testudines or Chelonia)

The shell that makes turtles instantly recognizable is formed by bone overlain by horny epidermal scales called scutes. The ribs, vertebrae, and parts of the pectoral girdle are fused to the dorsal shell (carapace), which is connected to the lower shell (plastron) by a bony bridge. The scutes or bone are secondarily reduced in some lineages of turtles, most notably the freshwater soft-shell turtles (Trionychidae) and the marine leatherback turtle (*Dermochelys coriacea*).

The earliest fossils of turtles are from the late Triassic (Joyce and Gauthier, 2004), and two major lineages, Pleurodira and Cryptodira, can be distinguished by the early Jurassic. Pleurodires have the common name side-necked turtles because they bend the neck horizontally when they retract their heads, whereas cryptodires bend their necks vertically. The geographic occurrence of pleurodires is currently restricted to the Southern Hemisphere (although fossil pleurodires are known from the Northern Hemisphere), and all of the approximately 70 species are aquatic. Pleurodires are the only turtles native to Australia and New Guinea and the only aquatic turtles in sub-Saharan Africa.

The approximately 250 species of cryptodires have a worldwide distribution and include specialized marine and terrestrial forms as well as aquatic and semiaquatic species. Tortoises (Testudinidae, about 50 species with a worldwide distribution in temperate and tropical regions) are the most terrestrial turtles. Most tortoises have domed carapaces, and many have forelimbs that are modified for digging. Semiaquatic and aquatic turtles generally have low streamlined

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00162-3

shells and webbed feet. Most of the 125 species of emydid turtles fit this description. Emydids are found primarily in the New World, with one European species, the pond turtle *Emys*. The paraphyletic group "Bataguridae" is the Old World equivalent of the emydids, with about 60 species in Europe and Asia and one genus (*Rhinoclemmys*, 9 species) in the Americas. The sea turtles (Cheloniidae (6 species) and Dermochelyidae (1 species)) are still more specialized for swimming, with forelimbs that are modified as flippers and short necks that cannot be retracted. Sea turtles have a worldwide distribution, primarily in tropical and subtropical regions although the range of the leatherback sea turtle extends well into the North Atlantic and South Pacific.

Crocodilians (Crocodylia)

Crocodilians are part of the archosaur lineage that includes dinosaurs and birds as well as a number of other groups. Most crocodilians are large (at least 1 m in total length) and have bony plates in the skin. These characteristics have contributed to an extensive fossil record that extends back to the middle of the Triassic. Crocodilians reached the height of their diversity in the Mesozoic (Clark, 1997). The history of the group includes terrestrial species, ranging in size from the meter-long *Araripesuchus rattoides* that ate plants and soil-dwelling invertebrates to *Sarcosuchus imperator*, which, at 12 m and tipping the scales at 8000 kg, was as large as tyrannosaur dinosaurs. Metriorhynchids, a group known from the Jurassic and Cretaceous, were highly specialized for marine life with the forelimbs modified as paddles and a lobed tail. They may have been viviparous like some other Mesozoic marine reptiles.

In contrast to their rich Mesozoic history, crocodilians are the smallest group of extant reptiles, with only 23 species. Crocodilians are primarily inhabitants of tropical and subtropical regions, but the American and Chinese alligators (*Alligator missippiensis* and *A. sinensis*) are found in temperate areas. All living crocodilians are aquatic; they swim with lateral undulations of the tail, which is somewhat flattened laterally. The nostrils of crocodilians are at the tip of the snout and are slightly elevated, as are the eyes. Alone among nonavian reptiles, crocodilians have developed a secondary palate that separates the passageway for air from the oral cavity. This combination of characters allows crocodilians to float with only the eyes and nostrils exposed.

Three lineages of extant crocodilians are recognized. The Alligatoridae includes eight species of freshwater crocodilians. All except the Chinese alligator occur in the New World. *Paleosuchus* (dwarf caimans) are among the smallest crocodilians at 1.7 m total length, and *Melanosuchus niger* (6 m) is one of the largest. Alligatorids have relatively broad snouts and the teeth of the lower jaw fit into pits in the upper jaw and cannot be seen when the mouth is closed. Crocodylids have narrower snouts than alligators, and the fourth tooth of the lower jaw of a crocodile is visible when the mouth is closed. The 14 species of crocodiles have a worldwide distribution, primarily in tropical and subtropical regions. Several species of crocodiles, especially the American crocodile (*Crocodylus acutus*) and the saltwater crocodile (*C. porosus*), regularly make long sea journeys, and the saltwater crocodile has colonized islands 1000 km from the nearest mainland. The saltwater crocodile is the largest species, growing to a length of more than 7 m. The African dwarf crocodile, *Osteolaemus tetraspis* (2 m), is the smallest member of the family. The Gavialidae includes just one species, the gharial (*Gavialis gangeticus*), found in northern India and Nepal. It is a specialized fish eater with a long, extremely narrow snout. A similar but less specialized long-snouted fish eater, the false gharial *Tomistoma schlegeli* from the Indonesian Archipelago, is placed in the Crocodylidae by some authorities and in the Gavialidae by others.

Lepidosaurs (Lepidosauria)

More than 95% of living reptiles are lepidosaurs, and unlike turtles and crocodilians, this lineage displays enormous diversity in body form and ecology. Lepidosaurs are covered by scales (the name means "scaly reptile") and they shed the skin at regular intervals. They have a transverse cloacal slit (rather than the longitudinal slit seen in crocodilians and turtles), and many species exhibit caudal autotomy that is facilitated by fracture planes within the vertebrae of the tail. Lepidosaurs include two lineages of extant forms, the Rhynchocephalia (the tuatara of New Zealand) and the Squamata (snakes and lizards) (Evans and Jones, 2010). Like their sister group, the Archosauria, lepidosaurs are diapsid (i.e., primitively they have two fenestrae defined by bony arches on the sides of the skull). Modifications of the diapsid condition help to define subgroups within lepidosaurs and contribute to functional differences in prey handling.

Rhynchocephalia

The Rhynchocephalia, as it is currently understood (Gauthier *et al.*, 1988), includes several extinct forms dating back to the Triassic and a single extant lineage, the Sphenodontidae, represented by a single, genetically variable species of tuatara (*Sphenodon punctatus*) that is found only on islands off the coast of New Zealand (Hay *et al.*, 2010).

Tuatara are lizardlike in body form and reach a total length of about 50 cm. (Tuatara is a Maori word meaning "spines on the back" and does not add an s to form the plural.) Unlike their sister group, the Squamata, tuatara retain the fully diapsid skull condition, with a lateral opening that is bounded dorsally by the postorbital and supratemporal bones and ventrally by the supratemporal and jugal. These bony connections impart rigidity to the skull that allows tuatara to deliver a forceful bite. The single row of teeth on the lower jaw fits between two rows of teeth in the upper jaw – an outer row on the maxilla and an inner row on the palatines – and prey is chewed to a pulp before it is swallowed. Tuatara lack the paired hemipenes (intromittent organs) that characterize squamates.

Squamata

A large number of anatomical features characterize squamates. Prominent characteristics are reductions in the bony connections forming the arches in the skull, allowing a degree of kinesis unknown in other nonavian reptiles, and the presence

in male squamates of paired hemipenes that develop as outgrowths of the posterior wall of the cloaca and are housed in the base of the tail. The hemipenes, which are everted in use, are held in the female's cloaca by spines, and sperm passes along a lateral groove.

Reduction or loss of limbs, usually accompanied by elongation of the trunk and tail, is a recurrent theme among squamates that reaches its zenith in snakes. Phylogenetically, snakes are included within lizards, but the abundance and diversity of snakes combined with their morphological, ecological, and behavioral specializations make it helpful to discuss lizards and snakes separately.

Lizards

Fossils of lizards are known from the middle Jurassic, and the major lineages were distinct by the end of the Mesozoic. Lizards are found on every continent except Antarctica and are well represented on oceanic islands because large species, such as iguanas, can survive long journeys from the mainland on floating vegetation, and small species (most notably skinks and geckos) often travel as stowaways with humans.

The more than 3000 species of extant lizards can be assigned to 22 families (Pough *et al.*, 2004), three of which are very large: Scincidae (more than 1000 species), Iguanidae (more than 900 species), and Gekkonidae (870 species). Skinks and geckos have worldwide distributions, whereas only three genera of iguanids are found outside the New World (*Brachylophus* on Fiji, and *Oplurus* and *Chalarodon* on Madagascar).

Most species of iguanids are diurnal, visually oriented lizards whose activities and social behaviors can easily be observed, and they occur in areas accessible to biologists. As a result, iguanids figure prominently in ecological and behavioral studies of lizards. The family as it has traditionally been defined included several genera of large herbivorous lizards (the Iguaninae) and a variety of smaller, primarily insectivorous species that were placed in other subfamilies (Etheridge and de Queiroz, 1988). A review of anatomical characters elevated the subfamilies of Iguanidae to family status, leaving only the large herbivores in the Iguanidae (Frost and Etheridge, 1989). This reclassification had a broad impact on the biological literature, because the newly elevated families include species that have formed the basis for much of the research on lizards. More recently, an analysis based on molecular characters has supported the traditional interpretation of Iguanidae (Macey *et al.*, 1997), and that classification is used here.

Snakes

Phylogenetically snakes are included within lizards, and the most primitive snakes, aniliids and uropeltids, lack many of the derived characters that are responsible for the distinctive ecology and behavior of advanced snakes. Even primitive snakes have greatly increased the number of precloacal vertebrae, however, and have lost the bony connection between the postorbital and supratemporal bones, freeing the supratemporal to rotate on its flat articulation with the parietal. This anatomical change allows the supratemporal and quadrate bones to act as part of the lower jaw, and these bones show varying patterns of elongation in advanced lineages of snakes. The brain of snakes is enclosed in a rigid box formed by ventral extension of the frontal and parietal bones and their articulation with the sphenoid bone.

Snakes are clearly derived lizards, but the sister group of snakes and the conditions that led to their evolution are still controversial topics. Snakes focus images on the retina by moving the lens, whereas lizards focus images by changing the shape of the lens. This difference and differences in the photosensitive cells in the retinas of lizards and snakes have long been interpreted as indicating that snakes are derived from a subterranean lineage of lizards with degenerate eyes. Other hypotheses of the origins of snakes are plausible, although they do not specifically account for differences in the eye. For example, legs are not particularly useful to small squamates that live in dense vegetation such as clumps of grass, and this is the habitat occupied by many legless lizards. This epigean adaptive zone might provide an intermediate stage in the evolution of snakes. Alternatively, aquatic habits could have been an intermediate stage in the evolution of snakes. *Pachyrachis problematicus*, a recently described elongate aquatic lepidosaur from Cretaceous marine deposits, was more than a meter long and had more than 100 presacral vertebrae and a skull that shows many of the derived features of snakes, including envelopment of the brain by bone (Caldwell and Lee, 1997). Two recent molecular studies, however, support the hypothesis of a terrestrial origin of snakes (Townsend *et al.*, 2004; Vidal and Hedges, 2004).

Snakes were apparently widespread by the late Cretaceous; fossils of that age are known from India, Africa, and South America. Most fossils consist of isolated vertebrae, but vertebral characters play a minor role in classification of living snakes and it is difficult to associate these fossils with extant lineages. Snakes radiated extensively in the Cenozoic, and about a dozen families are usually recognized, containing more than 2500 species. Primitive snakes, including boas and pythons, can be compared to more advanced forms, the Caenophidia, which include elapids (cobras, coral snakes, sea snakes, and their relatives), viperids (pit vipers and true vipers), and the Colubroidea, a category that contains more than two-thirds of the total number of extant species of snakes (Zaher *et al.*, 2009).

Ancestral and Derived Characters

The functional ecology of reptiles is shaped by a combination of ancestral and derived characters, and an appreciation of these features and the ways in which they interact is central to understanding how reptiles function as organisms. Important derived characters include reproduction via a shelled egg (with the secondary development of viviparity in many lineages), excretion of nitrogenous wastes as urate salts, and a dry skin that is relatively impermeable to the passage of water. Retained ancestral characters include ectothermal thermoregulation (with the accompanying low metabolic rate and the absence of an insulating layer of hair or feathers) and an anatomically three-chambered heart (except in crocodilians) that permits intracardiac blood shunting.

Sex Determination

Both genetic sex determination (GSD) and environmental (temperature-dependent) sex determination (TSD) occur in reptiles (reviewed by Sarre *et al.*, 2004). Heteromorphic sex chromosomes have been described in some species of turtles (male heterogamety: XY or XXY), snakes (female heterogamety: ZW, ZZW, or ZWW) and lizards (both male and female heterogamety), but other species in those groups are homogametic. Crocodilians and the tuatara lack heteromorphic sex chromosomes (Huchzermeyer, 2003; Norris *et al.*, 2004).

GSD occurs in most species of reptiles with heteromorphic sex chromosomes as well as in some species with homomorphic sex chromosomes. TSD is characteristic of the tuatara, all crocodilians, and many species of turtles and lizards, including some species with heteromorphic sex chromosomes. Sarre *et al.* (2004) proposed that GSD and TSD are the ends of a continuum. This view is supported by the observation that in some species, temperature can override GSD to produce individuals that are phenotypically male and genotypically female or the reverse (Quinn *et al.*, 2007).

Temperature acts on gonadal development during a critical period in embryogenesis. For turtles, the tuatara, and especially crocodilians, the critical period occurs after the eggs have been deposited in the nest, but for lizards the critical period occurs while the eggs are still within the oviducts of the female (Shine *et al.*, 2007). Pieau and Dorizzi (2004) argued that TSD is produced by the effect of temperature on the gonads alone, and that no extragonadal tissues are involved.

The ecological and evolutionary significance of TSD have been debated since the phenomenon was described (Shine, 1999). The Charnov–Bull model proposes that TSD maximizes the fitness of parents by adjusting the sex of the offspring to the nest temperature, and that relationship has been demonstrated experimentally for an Australian agamid lizard (Warner and Shine, 1998).

Temperature and Water Relations

With the exception of a few very large species, nonavian reptiles are ectotherms. That is, they obtain the energy they need to raise body temperatures to activity levels from external sources, either directly from the sun or via contact with surfaces that have been heated by the sun. Reptiles employ a mixture of behavioral and physiological mechanisms of thermoregulation: They move between sun and shade, orient the body to maximize or minimize interception of solar radiation, and concentrate or disperse melanin to change reflectivity. These mechanisms are supplemented by intracardiac shunting of blood and adjustments of the peripheral circulation to accelerate or retard exchange of heat with the environment. In contrast, birds and mammals are endotherms, changing the insulation provided by hair or feathers and adjusting metabolic rate to match rates of heat production and heat loss.

Ectothermy imposes some limits on reptiles. When heat sources are not available – at night for example, in water, or beneath the closed canopy of a tropical forest – reptiles must either accommodate their temperature requirements to the ambient temperature or cease activity. In the absence of environmental constraints, however, ectothermal thermoregulation can produce body temperatures that are as high and as stable as those of birds or mammals. Lizards and snakes, especially species that live in open habitats, generally control their body temperatures within activity temperature ranges of 2–3 °C. Terrestrial turtles, too, control their body temperatures by moving between sun and shade during periods of activity. The large body sizes of some turtles and crocodilians stabilize their body temperatures. Leatherback sea turtles, which weigh up to 1000 kg, are found off the coasts of New England and Nova Scotia in the summer in water temperatures as low as 8 °C. A countercurrent exchange mechanism in the flippers conserves the heat produced by muscular activity as the turtles swim, and body temperatures can exceed water temperature by more than 15 °C (James and Mrosovsky, 2004).

The dry skin of reptiles is an essential component of their ectothermal thermoregulation. In contrast, the high rate of evaporation from the moist skins of extant amphibians limits their capacity to maintain body temperatures above ambient levels as well as their ability to be active during the day when evaporative stress is high. Reptilian water balance is additionally facilitated by enzymatic pathways that convert nitrogenous waste to uric acid, which is excreted as a complex mixture of urate salts. The shelled eggs of reptiles absorb water from the substrate during development, and the availability of water in the nest environment influences the body size and, probably, the viability of newly hatched reptiles.

Metabolism and Energetics

Because ectotherms do not depend on metabolic heat production for thermoregulation, their resting metabolic rates are low – one-seventh to one-tenth those of endotherms of the same body size. The influence of low resting metabolic rates can be seen in the energy use, body size, and the exercise physiology and foraging behavior of reptiles.

Energy Requirements and Conversion Efficiency

The daily energy budget for a lizard is only 3% of the energy requirement for a mammal of the same body size (Bennett and Nagy, 1977). Low energy use by the lizard results partly from its low resting metabolic rate during the day when it is active and partly from a reduction in body temperature (with a consequent further reduction in metabolic rate) at night when it is inactive. In contrast, a mammal maintains a higher metabolic rate during the day and may increase its metabolic rate above resting levels at night to counteract the effect of falling ambient temperature.

Low metabolic rates translate to low energy requirements for reptiles. As a result, reptiles can cope with environments in which food is limited, either episodically or permanently. Areas of low primary production, such as deserts, and even regions where there is no primary production, such as desert salt flats and mobile sand dunes, are home to a variety of lizards.

Efficient conversion of food to biomass is a second consequence of the low metabolic rates of reptiles. Endotherms use most of the energy they ingest to stay warm and be active – 19 species of birds and mammals have an average net

conversion efficiency of 1.4% (range 0.5–3%), whereas nine species of reptiles have an average net conversion efficiency of 46.6% (range 18–86%) (data from Pough, 1980).

Body Size

Mass-specific metabolic rates ($W\,g^{-1}$) are related to body mass by an exponent that approximates -0.25, although statistically significant variation from that value is associated with both phylogeny and ecology (Hayssen and Lacey, 1985; Andrews and Pough, 1985). The consequence of that negative allometry is a mass-specific energy requirement that increases steeply as body size decreases (**Figure 1**).

The low metabolic rate of reptiles makes extremely small body size energetically feasible, and lizards, in particular, include very small species. The modal body mass of lizards falls in the range 1–10 g, and nearly 20% of lizards have adult body masses less than 1 g. In contrast, there are no mammals with adult body masses less than 1 g, and fewer than 5% of mammals have body masses below 10 g (see Pough *et al.*, 2009, p. 386). More than half the living species of lizards occupy a zone of body sizes in which there are practically no mammals, and the smallest lizards have their major competitive and predatory interactions with invertebrates. In contrast to lizards, snakes are relatively large reptiles – nearly 80% of snakes are larger than 20 g. Thus, snakes and lizards illustrate two different ways to succeed as a squamate in a world that is largely dominated by endotherms: Small body size is a morphological specialization of lizards, just as elongation and skull kinesis are morphological specializations of snakes.

Exercise Physiology and Behavior

The low resting metabolic rates of reptiles limit their aerobic metabolic capacity – i.e., the maximum rates of oxygen consumption they can achieve during strenuous activity. Both reptiles and mammals increase resting rates of oxygen consumption about 10-fold when they run at maximum speed, so the differential in metabolic rates of ectotherms and endotherms persists during activity. Most reptiles have limited aerobic metabolic capacity, and they derive most of the ATP used during high levels of activity from anaerobic pathways.

Glycogen, the metabolic substrate for anaerobic metabolism, is stored in the muscle cells and a brief period of activity can exhaust the muscles' glycogen supply. As a result, anaerobic metabolism is an effective way to support bouts of intense muscular activity, but it is not suitable for sustained behaviors such as long-distance locomotion. The activity patterns of many reptiles are built on this relationship – they employ sit-and-wait modes of foraging, they escape predators by fleeing to a nearby shelter, and their social behaviors do not include sustained high levels of activity.

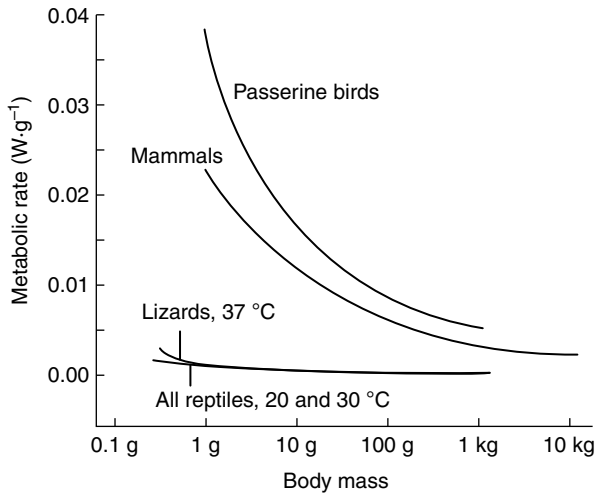


Figure 1 The mass-specific energy requirements of animals increase at small body sizes. The metabolic rates of reptiles are one-seventh to one-tenth those of birds or mammals of the same body size, and the smallest reptiles have adult body masses less than 1 g, whereas very few birds and mammals are smaller than 10 g as adults. Modified from Pough FH (1980) The advantages of ectothermy for tetrapods. *American Naturalist* 115: 92–112, with permission from University of Chicago Press.

Active Reptiles and the Evolution of Endothermy

In an ecological context, the generalization that reptiles are ectotherms with low metabolic rates and low levels of activity applies to extant species, although some modifications of that characterization will be addressed in subsequent sections. In an evolutionary context, however, the generalization is patently false because birds are reptiles and they are endotherms with high metabolic rates. Clearly the reptilian lineage has a capacity for endothermy that is barely expressed in nonavian reptiles. An examination of the evolution of endothermy explains that dichotomy and emphasizes how tightly anatomical and physiological characteristics are linked to thermal ecology.

Ectothermy is the ancestral condition for vertebrates, and the derived condition of whole-body endothermy has evolved at least twice, in mammals and in birds. In a broader view, regional endothermy has evolved independently in sharks, tunas, and billfishes, and whole-body endothermy may have been characteristic of pterosaurs (flying archosaurian reptiles of the Mesozoic) and some lineages of dinosaurs.

Ectotherms and endotherms have very different relationships to their physical environments: Ectotherms rely primarily on behavioral thermoregulation to raise their body temperatures because they have low metabolic rates, and the absence of insulation facilitates uptake of heat from the environment. In contrast, endotherms use internal heat production from high metabolic rates to regulate body temperature, and they require insulation to retain metabolic heat in their bodies.

The evolutionary transition from ectothermy to endothermy is complicated by a Catch-22 – adding insulation to an ectotherm impairs its behavioral thermoregulation, but in the absence of insulation, any heat produced by increasing its resting metabolic rate is lost to the environment. The solution to this paradox lies in finding a basis for the evolution of insulation or the evolution of an increased metabolic rate that does not depend on the pre-existing occurrence of the other character.

Mammals are the sister group of reptiles (including birds), and the common ancestor of mammals and reptiles was ectothermal. Thus, the evolution of endothermy in the mammalian lineage may provide a model for the evolution of endothermy among reptiles. Anatomical changes seen in the

fossil record of predatory synapsids, the sister group of mammals, strongly support the hypothesis that the initial step in the evolution of mammalian endothermy was the selection of increased locomotor capacity. These changes include the evolution of a cursorial body form, changes in the rib cage that suggest the presence of a diaphragm, and increased surface area in the nasal passages to warm and humidify large volumes of air. Increasing levels of locomotor activity require an increase in metabolic rate, and internal heat production would create a selective value for insulation (see Pough *et al.*, 2009, Chapters 11 and 18, for details and references).

Some features of the fossil record of birds suggest that a similar scenario can be applied to the evolution of avian endothermy, but others appear to contradict that interpretation. Like the predatory synapsids, the small maniraptoran dinosaurs that form the sister group of birds appear to have been fleet-footed predators that pursued their prey. This interpretation suggests that these dinosaurs may have evolved the metabolic capacity for endothermy just as synapsids did, and the impressions of feathers in fossil dinosaurs supports this interpretation. However, two lines of evidence cast doubt on the hypothesis that the dinosaurian precursors of birds had high metabolic rates. Examination of an excellently preserved specimen of the small dinosaur *Sinosauropteryx* suggests that it had simple septate lungs that were ventilated by a piston-like movement of the liver like those of living crocodilians. Lungs of this sort would not support high rates of oxygen consumption. This interpretation is supported by CAT scans of the nasal passages of dinosaurs that reveal no trace of modifications of the nasal passages to warm and humidify large volumes of air (Ruben *et al.*, 1996).

This evolutionary perspective emphasizes intricate interconnections among the anatomical and physiological characters of extant reptiles and their ecology and behavior, as well as the evolutionary capacity for breaking those links. An examination of living reptiles reveals additional connections among anatomy, physiology, ecology, and behavior.

Characteristics of Groups

Evolution proceeds in a mosaic fashion; all descendant lineages inherit ancestral characteristics, and each lineage develops a set of derived characters that interact with ancestral characters and with each other to shape the way animals live. The following sections focus on combinations of ancestral and derived characters found in specific groups of reptiles that are important in shaping their ecology and behavior. TSD is widespread among turtles, for example, and crocodilians show extensive parental care. To avoid repetition, examples from other lineages are included when a topic is discussed. Additional details, examples, and references can be found in Pough *et al.* (2009); Pough *et al.* (2004), and the sources cited.

Turtles

The Nest Environment and the Size and Sex of Hatchlings

All turtles lay eggs in holes which they excavate with their hind legs. Physical conditions within the nest determine both the size and the sex of hatchlings of many species of turtles (Packard and Packard, 2000). In general, embryos in nests constructed in moist substrates metabolize most of their yolk before they hatch and grow larger than embryos in drier nests. Hatchlings of snapping turtles (*Chelydra serpentina*) from moist nests run and swim faster than the smaller hatchlings that emerge from dry nests. This difference in locomotor performance, which persists for at least a year after hatching, might result in higher first-year survival of hatchlings from moist nests.

TSD is widespread among both pleurodires and cryptodires. For most species of turtles, low temperatures during incubation produce males (Pattern Ia), but in a few species males are produced at high temperatures (Pattern Ib), and some species show a biphasic response in which females develop at both low and high temperatures and males are produced at intermediate temperatures (Pattern II). Environmental sex determination can produce both male and female hatchlings from the same nest, because the shift between sexes typically occurs within a range of only 1–2 °C. Thus, eggs from the top of a nest can develop into females while eggs deeper in the same nest produce males, or clutches laid early in the season when nest sites are exposed to the sun can produce females, whereas clutches laid in the same sites later in the summer produce males because the nests are now shaded by vegetation.

Gigantism

Giant tortoises are among the most spectacular reptiles. The largest species of living tortoises are found on the Galápagos and Aldabra Islands and reach carapace lengths of more than a meter and body weights exceeding 100 kg, but large tortoises are not limited to islands. Extant members of the genus *Geochelone* have a pantropical distribution, with mainland representatives in Africa, India, Asia, and South America as well as additional island species on Madagascar, Sri Lanka, and Indonesia (Ernst and Barbour, 1989; Pritchard, 1996). The largest living mainland species, the African spur-thighed tortoise *G. sulcata*, reaches a carapace length of 76 cm, and fossil tortoises with carapace lengths from 60 cm to 1.5 m are known from Pleistocene deposits on all of the continents except Australia and Antarctica. Giant tortoises were present in North America when humans arrived, and a subfossil carapace from Florida contains a spear point and is charred as if it had been placed upside down and cooked on fire.

Crocodilians

Reproduction and Parental Care

All crocodilians are oviparous, laying eggs either in nests that the female excavates in the ground or in mounds of vegetation that the female builds. Crocodilians in some areas are reported to lay eggs on floating mats of vegetation. Parental care is probably universal among extant crocodilians, although the reproductive behavior of some species is unknown, and parental care appears to have been reduced in populations that have been hunted extensively (Lang, 1989). Females are the primary caregivers, but males often participate. A female crocodilian remains near her nest while the eggs are incubating. As the young approach hatching, they start to vocalize within the eggshells, and these vocalizations signal the female

to begin opening the nest. The male may participate in opening the nest, and both parents pick up hatchlings in their mouths and transport them to water. Adult crocodilians show remarkable dexterity in this process, sometimes picking up an intact egg and cracking the shell to release the hatchling.

Parental care is virtually unknown among turtles, but about 100 species of squamates exhibit some degree of parental care (Shine, 1985; Gans, 1996). These behaviors range from defense of the nest site to brooding the eggs. In an exception to the generalization that reptiles are ectotherms, females of some species of python gather their eggs into a clump encircled by coils of the body and use muscular contractions to keep the temperature of the eggs close to 32 °C. The frequency of contractions, and thus the python's rates of oxygen consumption and heat production, increases as ambient temperature falls. A female Indian python (*Python molurus*) can maintain her egg mass at 31–32 °C at ambient temperatures between 23 and 32 °C, and functions as an endotherm during this period.

Lepidosaurs

Tuatara

Tuatara have a number of unusual ecological and physiological characteristics. They are crepuscular and forage on cool, foggy nights, with body temperatures as low as 6 °C. During the day, tuatara bask in the sun and raise their body temperatures to 28 °C. Tuatara feed on invertebrates and nestling shearwaters. The prey is crushed as the lower jaw closes with an initial vertical movement and is then sheared by the interdigitating rows of triangular teeth as the lower jaw slides forward.

Tuatara, unlike other lepidosaurs, lack hemipenes. Sperm is transferred by apposition of the cloacas of the male and female. In late spring, female tuatara excavate shallow depressions in which they deposit 6–15 eggs. Embryonic development requires 12–15 months (Saint Girons, 1985).

Tuatara are the sister group of squamates, and their anatomy shows characters that are probably ancestral for squamates, including the connection between the jugal and supratemporal bone forming the lower border of the lateral arch of the skull. Some authors have used the phylogenetic position and primitive anatomical characters of tuatara as a basis to suggest that their ecology and physiology also reveal ancestral conditions for squamates, but this is probably not the case. New Zealand is a cold place, and the native geckos and skinks are active at body temperatures well below those of geckos and skinks in warmer parts of the world. It seems most likely that the low activity temperatures of tuatara and the slow development of their eggs are derived characters that reflect current ecological conditions.

Squamata

Lizards have proven to be ideal study-animals. Many lizards are diurnal and have limited home ranges, so their activities can readily be observed. As a result of their ease of study, species of *Anolis*, *Sceloporus*, *Pogona*, *Urosaurus*, and *Uta* have played seminal roles in the fields of animal behavior, ecology, and evolutionary morphology and physiology (Greer, 1989; Vitt and Pianka, 1994; Wainwright and Reilly, 1994; Pianka and Vitt, 2003). Most snakes, in contrast, are secretive in their behavior and direct observation is much more difficult. Nonetheless, the combination of perseverance with radiotelemetry has produced several important field studies (Shine, 1991; Seigel and Collins, 1993; Greene, 1997).

Lizards

Classifying species of lizards in terms of the foraging modes they employ provides a framework that organizes diverse aspects of their biology into a series of useful generalizations (**Table 1**). Like other mobile animals, lizards display a spectrum of foraging behaviors which extends from species that perch motionless on lookouts and make short dashes to capture prey to species that are in nearly continuous motion as they forage, searching for hidden prey under leaf litter and in holes and cavities. Species at the inactive end of this spectrum are called sit-and-wait predators, those in the middle are cruising foragers, and the species that move continuously are called active foragers.

Sit-and-wait lizards rely on vision to detect moving prey and may be selective in deciding on which prey to attack. Widely foraging lizards examine places where prey may be hiding; olfaction is an important sensory mode for these species, and they usually take any prey items they encounter.

Sit-and-wait lizards are usually cryptic and are most likely to be detected by widely foraging predators, whereas widely foraging lizards are conspicuous because of their movement and are consequently vulnerable to sit-and-wait predators. Thus, there may be an alternation of predatory modes at successive levels in a food chain – widely foraging predators find sedentary prey and are eaten by sit-and-wait predators. In the American Southwest, for example, whiptail lizards (*Cnemidophorus*) forage widely and locate sedentary insects such as termites. Leopard lizards (*Gambelia wislizenii*) wait motionless beneath bushes and capture whiptail lizards when they come within range. In turn, leopard lizards are captured by coachwhip snakes (*Masticophis flagellum*), which move from bush to bush in search of prey.

The quick dash that a sit-and-wait lizard makes to capture prey requires high sprint speed, whereas the continuous movement of a widely foraging predator requires endurance. These characteristics are reflected in the physiology of lizards – sit-and-wait lizards can sprint rapidly, but they rely primarily on anaerobic metabolism and have limited endurance. Widely foraging lizards cannot sprint so fast, but they have relatively larger hearts and higher hematocrits than sit-and-wait species, so they have greater aerobic metabolic capacities and greater endurance.

The colors of sit-and-wait lizards often match the backgrounds on which they rest, and their patterns obscure the outlines of the body, making them cryptic when they are motionless. Widely foraging lizards move across a variety of backgrounds and do not necessarily match any particular one. They are often solid-colored or striped, and these patterns may confuse predators when a lizard is moving. Sit-and-wait lizards are exhausted by brief periods of rapid locomotion, and when they are attacked, they sprint to shelter. Widely foraging lizards often run from bush to bush, changing direction when they are out of sight and travelling a long distance before they stop. The generally long tails of widely foraging

Table 1 Ecological and behavioral characteristics associated with the foraging modes of lizards[a]

	Foraging mode		
Character	*Sit-and-wait*	*Cruising forager*	*Widely foraging*
Foraging behavior			
Movements/hour	Few	Intermediate	Many
Distance travelled/hour	Small	Intermediate	Large
Primary sensory mode	Vision	Vision and olfaction	Olfaction
Exploratory behavior	Low	Intermediate	High
Predators and prey			
Types of prey	Mobile, large	Intermediate	Sedentary, often small
Types of predators	Widely foraging	?	Sit-and-wait and widely foraging
Risk of predation	Low	?	High
Mode of escape	Cryptic, flee a short distance to shelter when attacked	?	Conspicuous, may flee a long distance when attacked
Physiological characteristics			
Sprint speed	High	?	Low
Endurance	Low	?	High
Aerobic metabolic capacity	Low	?	High
Anaerobic metabolic capacity	High	?	Low
Heart mass	Low	?	High
Hematocrit	Low	?	High
Energetics			
Daily energy expenditure	Low	?	High
Daily energy intake	Low	?	High
Social behavior			
Social system	Territorial	?	Not territorial
Home range	Small	Intermediate	Large
Body form			
Trunk	Stocky	Intermediate	Slim, elongate
Tail	Often short	?	Often long
Color and pattern	Often blotched, may match background, rarely colorful	?	Often solid or striped, sometimes colorful
Reproduction			
Mass of clutch relative to mass of female	High	?	Low
Viviparity	Often	Sometimes	Rare

[a]Foraging modes are presented as a continuum from sit-and-wait to widely foraging species. In most cases, data are available only for species at the extremes of the spectrum. *Source*: Modified from Pough FH, Andrews RM, Cadle JE, Crump ML, Savitzky AH, and Wells KD (2004) *Herpetology*, 3rd edn. Saddle River, NJ: Prentice Hall.

lizards may facilitate caudal autotomy as a last-ditch means of escape.

Sit-and-wait lizards can survey their surroundings from their perches whereas widely foraging lizards can see little beyond their immediate horizon. Sit-and-wait species have small home ranges and are often territorial, whereas widely foraging species have larger home ranges and are rarely territorial. The bulk of eggs within the oviducts are probably a handicap for a widely foraging lizard, and widely foraging species have smaller ratios of clutch mass to maternal mass than the sit-and-wait species. Viviparity is more common among sit-and-wait lizards than among widely foraging species.

These generalizations organize a large number of empirical observations into a coherent whole, but assigning evolutionary cause-and-effect interpretations to the correlations may confound phylogeny with ecology. Most of the sit-and-wait species that have been studied are in the iguanian lineage, whereas the widely foraging lizards are scincomorphs. Thus, the effects of phylogeny (iguanian vs. scincomorph) cannot be separated from those of ecology (sit-and-wait vs. widely foraging). However, a study of several closely related species of lacertid lizards identified sit-and-wait and widely foraging species, and detected many of the same behavioral and physiological correlations that are found in the broader comparison (Huey *et al.*, 1984).

Because many species of lizards use behavioral mechanisms to regulate their body temperatures within narrow limits while they are active, the thermal environment plays a central role in defining the niche of a species. The thermoregulatory characteristics of habitats can be defined as high cost or low cost, depending on the behavioral and ecological trade-offs a lizard must make to maintain a body temperature that is different from ambient temperatures. Lizards that live in habitats in which basking sites, perches, and escape sites are abundant and

close to each other have low ecological and behavioral costs of thermoregulation and are likely to maintain body temperatures within a narrow range during activity. In contrast, lizards that live in habitats where basking sites are scarce or ephemeral, such as the understory of a tropical forest, have high costs of thermoregulation. These lizards often have body-temperatures indistinguishable from ambient temperature.

The costs of thermoregulation in a particular habitat are not the same for all lizards because body size profoundly affects the thermoregulatory options available to a lizard. An example is provided by three species of *Ameiva* that live in microsympatry on the Osa Peninsula of Costa Rica. The three species partition the structural habitat on the basis of the radiant environment. Adults of the smallest species (*A. quadrilineata*) weigh about 10 g and forage in short vegetation in full sun, whereas the middle species (*A. festiva*) weighs about 32 g and forages in areas where taller vegetation provides broken shade. The largest species (*A. leptophrys*) weighs 83 g and forages in deep shade beneath the forest canopy.

Ameiva are widely foraging lizards, and the three species display the same alternation of thermoregulatory and foraging behaviors: A lizard basks in the sun until its body temperature reaches 39 or 40 °C, and then forages until its body temperature cools to 35 °C when it ceases foraging and basks again. The rate of cooling is inversely proportional to body size, and relative rates of cooling appear to explain the differences in foraging sites of the three species. *Ameiva quadrilineata* cools from 40 to 35 °C in 4 min, compared to 7 min for *A. festiva* and 12 min for *A. leptophrys*. The slow cooling of *A. leptophrys* allows it to move deep under the canopy before it has to return to the forest edge to bask, whereas the smaller species cool too rapidly to move far from basking sites.

A different physical relationship apparently prevents the two larger species from foraging in the sun-drenched habitat occupied by *A. quadrilineata*. The equilibrium temperature reached by an object in the sun is sensitive to its size. The equilibrium temperatures of small objects are strongly affected by convective heat exchange, whereas the equilibrium temperatures of larger objects are closely tied to the radiant environment. As a result of this physical relationship, under the conditions on the Osa Peninsula, an *A. quadrilineata* in full sun reaches equilibrium at 39 °C, which is within its activity temperature range. In contrast, if an *A. festiva* or *A. leptophrys* remained in the sun, its body temperature would rise to a lethally high level.

Thus, the effectiveness with which each species can use a specific thermal environment appears to contribute to habitat partitioning by these species – the forest floor is a high-cost environment for the small and medium species and only the largest species can forage effectively in that habitat. Conversely, the absence of shade creates high costs of thermoregulation for the large and medium species in the open habitat, and only the smallest species can exploit this area.

Snakes

Elongation of the body is the immediate distinctive feature of snakes, and specializations associated with feeding form the core of their diversity. The body form of snakes results primarily from an increase in the number of vertebrae in the neck and trunk, and the longest species have more than 400 precaudal vertebrae. Even generalized snakes – rat snakes (*Elaphe*), gopher snakes (*Pituophis*), Australian whipsnakes (*Demansia*), and blacksnakes (*Pseudechis*) – have head-plus-trunk lengths which are 10–14 times the maximum trunk circumference. In contrast, elongate lizards (*Cnemidophorus, Egernia, Lacerta*, and *Eumeces*) have length/circumference ratios of 2.0–2.5.

In a functional sense, repackaging the mass of a lizard into a long, thin body form means that a snake must feed a large body through a small mouth. Lizards and snakes almost always swallow their prey whole; many lizards manipulate prey items in their mouths until they have been softened by repeated bites, whereas nearly all snakes swallow prey without any oral processing. Most legless lizards eat large numbers of small prey items; they retain the connection between the postorbital and supratemporal bones on the side of the skull, and this structure gives the skull sufficient rigidity to crush the exoskeletons of insects. In contrast, snakes have lost the postorbital–supratemporal connection and have replaced the body symphysis at the front of the lower jaw with a flexible ligament. These changes, which increase skull kinesis and allow the lower jaws to move independently, also reduce the ability of snakes to exert pressure on food in the mouth. A few snakes have secondarily evolved modifications of the skull and jaws that allow them to eat hard-bodied prey, but most snakes feed on soft prey.

Lizards normally consume several small prey items (or a large item ingested in several pieces) on a daily basis, whereas most snakes can swallow prey items with diameters larger than the snake's head, and eat less frequently.

The prey of snakes is large relative to the predator, and many snakes have specializations (such as constriction or venom) that allow them to immobilize prey before it can injure the snake. Prey size must be evaluated relative to the size of a snake in two respects – mass and diameter. That is, a prey item can be light or heavy compared to the snake that eats it and it can also be slim or bulky. The first comparison can be called relative mass and the second relative girth.

These two variables create four categories of prey (Greene, 1997):

1. Type I prey have both small mass and small girth ratios. Insects are Type I prey, and this is the prey type that is characteristic of most lizards and probably of the earliest snakes.
2. Type II prey are elongate and combine a large mass ratio with a small girth ratio. Earthworms, caecilians (elongate, burrowing amphibians), eels, and other snakes are examples of Type II prey. These food items require specializations to immobilize the prey (constriction or venom), but do not require large gapes. Most primitive snakes (e.g., the pipe snakes, Cylindrophis and Anilius) are limited to Type II prey.
3. Type III prey are heavy and bulky and have large mass and girth ratios. Frogs, mammals, and some lizards are Type III prey. This prey type requires both immobilization and large gape.
4. Type IV prey are bulky and light, with small mass ratios and high girth ratios. Birds (which are bulky because of

their feathers) are Type IV prey and require gape specializations but not immobilization. Boas and pythons take Type III and IV prey, as do most of the advanced snakes (elapids, vipers, and colubroids).

The foraging behavior of snakes extends from active searchers that use visual or olfactory cues to find prey to mobile ambushers that move to places where they are likely to encounter prey. Rattlesnakes (*Crotalus*), for example, coil beside trails followed by rodents and wait for prey, and tropical racers (*Mastigodryas*) wait beside sunny spots on the forest floor to capture lizards (*Ameiva*) that pause to bask.

The exercise physiology of snakes appears to parallel differences in their foraging behavior. Widely foraging species, such as North American whipsnakes (*Masticophis*), have high aerobic metabolic capacities and high endurance, whereas species that rely on ambush (*Lichanura*) have lower aerobic capacities and endurance. Widely foraging species probably eat more often than ambushing species and take smaller prey, but even a frequent feeder such as the coachwhip snake (*Masticophis flagellum*) eats at 10-day intervals (Secor and Diamond, 1998), whereas most lizards probably eat daily.

Type III prey (heavy and bulky) can pose a risk of injury to a snake, especially to those species that swallow prey while it is still struggling. Constriction and venom are specializations that allow snakes to subdue a prey item before it is swallowed. Constriction is characteristic of primitive snake lineages and of boas and pythons. Apparently it evolved as a method of controlling elongate prey and subsequently proved effective with bulky prey as well. Constriction requires enclosing a prey item within coils of the snake's body. Little pressure is exerted on the prey, and friction between adjacent coils holds them in place. Death may result from asphyxiation (because the snake takes up the slack each time the prey exhales) or because increased internal pressure interferes and ultimately stops the action of the heart. The short-radius curves required for constriction require short vertebrae and trunk muscles that span only a few vertebrae. This morphology is not compatible with rapid locomotion, which requires large-radius curves formed by long vertebrae and muscles that span many vertebrae. The evolutionary appearance of advanced snakes in the Miocene coincides with the spread of grasslands. Constrictors (largely small boas) dominated the snake fauna at the beginning of the Miocene, but by the end of that epoch the snake fauna was composed largely of colubroids (in the sense of Zaher *et al.*, 2009). These were fast-moving snakes, but their locomotor specializations precluded constriction. Although some lineages of colubroids secondarily evolved constriction, venom is the distinctively colubroid method of immobilizing prey with minimum risk to the snake.

Venom evolved early in the history of colubroids as part of the feeding apparatus. The defensive use of venom is a secondary development, enhanced in some cases by morphological specializations. The venom channel in the fangs of spitting cobras, for example, makes an abrupt right-angle turn near the tip of the fang that enables the cobra to emit a spray of venom that travels a meter or more. The snake aims for the predator's eyes, and accounts by people who have been sprayed confirm the powerful deterrent effect of this defense. Attacking a venomous snake can have severe consequences for a predator, and the warning mechanisms of venomous snakes (such as the rattle of rattlesnakes, the hood of cobras, and the distinctive colors of coral snakes) foster both learned and innate avoidance behavior by potential predators (Greene, 1988; Pough, 1988).

Duvernoy's gland, which lies at the rear of the maxilla and produces venom that immobilizes prey, is a derived character of colubroids. The embryonic development of Duvernoy's gland is linked to development of the posterior pair of maxillary teeth. These rearmost teeth are enlarged and grooved in many colubroids. Although colubroids are often called "the nonvenomous snakes" in popular literature, many colubroids use venom to subdue prey, and a few species known as rear-fanged snakes (notably the African boomslang, *Dispholidus typus*) can deliver defensive bites that are lethal to animals as large as humans. In the two groups popularly known as venomous snakes, elapids (cobras, coral snakes, sea snakes, and their relatives) and viperids (vipers and pit vipers), the maxilla has shortened, thereby moving the enlarged fang on the maxilla to the front of the mouth. Elapids have short fangs in a relatively immobile maxilla, whereas viperids have long fangs and the maxilla rotates so that the fang lies horizontally when the mouth is closed.

The venoms of elapids attack the nervous system and rapidly immobilize prey. In contrast, the venoms of vipers and pit vipers have high proteolytic activity and cause tissue destruction. The proteolytic activity of venom speeds digestion because the venom hydrolyzes the body tissues of the prey from the inside while the snake's digestive enzymes attack from the outside. Some vipers, especially species of the African genus *Bitis*, appear to combine proteolytic venom, injected deep into the prey by very long fangs, with modifications of the body-form to feed on very bulky prey. Large species of *Bitis* are remarkably heavy-bodied – length/circumference ratios of the puff adder (*B. arietans*) and Gaboon viper (*B. gabonica*) are 4 : 5. The thick bodies of these snakes accommodate bulky prey without disrupting locomotion or interfering with the function of internal organs.

The physiology of the digestive system of snakes is as specialized as their trophic anatomy (Secor, 2008). Large meals taken at infrequent intervals pose a challenge – the physiological cost of digesting a meal is high, but occasional periods of high activity are separated by long periods during which the digestive system is inactive. Within 24 h of feeding, snakes display remarkable factorial increases in rates of oxygen consumption; in the masses of organs in the cardiovascular, respiratory, digestive, and excretory systems; and in activities of digestive enzymes. Between feedings, the sizes of the organs and the activities of the enzymes regress to a resting state. This cyclic hypertrophy and regression, which is interpreted as a mechanism that conserves energy compared to the alternative of maintaining the physiological systems continuously at high levels, emphasizes the extent to which snakes display trophic adjustments that are shared by other vertebrates (Starck, 2003; Secor, 2005; Cox and Secor, 2008; Castoe *et al.*, 2008).

Conservation Issues

Many of the causes of declining populations of reptiles are depressingly familiar and affect other kinds of animals and plants as well. Habitat loss as a result of human activities is probably the major reason for shrinking populations,

although mortality on roads may be more significant than is generally recognized. Road kills are considered the major form of human-induced mortality for Australian reptiles and amphibians, exceeding the effect of habitat loss by nearly 70% (Ehmann and Cogger, 1985). Some European countries have built drift fences and constructed tunnels beneath roads to reduce mortality of reptiles and amphibians. Introductions of exotic species (especially on islands) and pollution (potentially including global warming (Dunham, 1993)) are additional factors that affect animals and plants on a large scale. Other threats are limited to particular groups and result from combinations of human activities and specific biological characteristics. For example, high-speed watercraft are a risk factor for crocodilians (Grant and Lewis, 2010).

Climate Change

Climate change affects the long-term persistence of populations of reptiles, as it does in other organisms, via changes in their biological and physical environments (Walther *et al.*, 2002). Most evaluations of climate change have focused on global warming, but climate change also encompasses decreased temperatures in some areas as well as changes in rainfall and in the intensity of solar radiation that reaches the Earth's surface, and global warming does not mean a uniform temperature increase everywhere or for every species.

For example, a change in cloud cover could lower the operative temperature of heliothermic reptiles while warming the average air temperature. An increase or decrease in the average ambient temperature could disrupt the match between seasonal cycles that are cued by day length, such as the timing of mating and egg laying, and cycles that depend on temperature, such as the appearance of plant growth stages or insect prey that provide the food needed by hatchlings. Similarly, changes in the amount or timing of rainfall could result in soil being too wet or too dry for successful egg development.

The ability of a species to resist the effects of climate change depends on a variety of factors (Packard and Packard, 2000; Morjan, 2003; Beeson and Cree, 2010, 2011; Doody and Moore, 2010). For example, changes in solar radiation or ambient temperature could exceed the ability of adult animals to alter daily or seasonal activity patterns to keep their body temperature within the range that permits normal activity. Nest temperatures or water potentials could exceed the limits for normal embryonic development, even if adults modify the seasonality of mating and egg laying and their selection of nest sites. Species with geographic ranges that do not permit migration, such as species living on islands, coral reefs, and near the tops of mountains are especially vulnerable to climate warming (Hamman *et al.*, 2007; Hawkes *et al.*, 2007; Bickford *et al.*, 2010; Fordham and Brook, 2010; Foufopoulos *et al.*, 2011).

Although climate change is expected to decrease biodiversity, some species may benefit from the changed conditions. For example, the average body size of lizards, *Lacerta vivipara*, in four populations in southern France has increased over an 18 year period, apparently in response to a 2.2 °C increase in the daily maximum temperature during August, which is the first month after the lizards hatch (Chamaillé-Jammes *et al.*, 2006).

TSD is another aspect of reptilian biology that could be affected by climate change (Mitchell *et al.*, 2008; Kallimanis, 2009; Mitchell and Janzen, 2010). Long-term changes in nest temperature could bias sex ratios, and species in which the switch from one sex to the other occurs within a narrow range of temperatures might be especially vulnerable to the effect of global warming (Hulin *et al.*, 2009). A warming climate should increase the temperatures of nests, but the relationship between climate and sex determination may be complex. For example, the proportion of male hatchlings in a population of red-eared slider turtles increased from 2001 to 2004 despite a warming trend that resulted in earlier nesting by females (Tucker *et al.*, 2008). Apparently the low soil temperature of the early nests favored the production of male hatchlings, despite an overall warming trend.

The ability of a species to change its geographic range by migrating or changing elevation as climate changes is a critical factor in predicting the effect of climate change on species with temperature dependent sex determination (Araújo *et al.*, 2006). Some species might be able to mitigate the effect of climate change by altering their nesting season or nest-site selection (Doody *et al.*, 2006; Hawkes *et al.*, 2007; Schwanz and Janzen, 2008).

The effects of global warming on species with TSD may not be universally detrimental. For sea turtles, an increase in nest temperature by 2 °C would shift the proportion of hatchlings substantially toward females, and this change might be beneficial for some populations (Poloczanska *et al.*, 2009).

Turtles

Many species of turtles are long-lived and display a suite of coevolved traits, including delayed sexual maturity and high annual survivorship of adults (Congdon *et al.*, 1994). Adult turtles are often nearly immune from nonhuman predators, and a turtle that survives to adulthood can expect a long reproductive period. When adult mortality is increased by human activity, as is the case for sea turtles which are drowned at sea in the nets of shrimp trawlers and killed for food when they come ashore to nest, a crucial component of the life history strategy is destroyed. Sea turtles are the most visible examples of excessive adult mortality resulting from human activities, but automobiles kill adult turtles and tortoises all over the world.

Well-intentioned efforts at management can founder on aspects of the biology of a species. Sea turtles, for example, exhibit TSD – males are produced at low temperatures and females at higher temperatures. Conservation activities at the nesting beaches of sea turtles usually include moving the eggs from natural nests to a protected site where they are buried in the sand to complete incubation. Under these circumstances, all of the eggs are exposed to similar temperatures, and as a result the hatchlings will be predominantly one sex. Even worse, in some cases the incubation conditions used by management programs produced high proportions of male hatchlings.

The value of headstarting programs for turtles is unclear – do they increase survival of hatchlings (the most vulnerable life history stage) and do headstarted turtles become

reproductively competent adults? Some evidence for success can be found among the more than 20,000 tagged hatchling or yearling green turtles (*Chelonia mydas*) that were released from the Cayman Turtle Farm on Grand Cayman Island in the Caribbean between 1980 and 2001. A total of 392 of those turtles were recaptured between 1984 and 2001, and some of them had returned to nest on beaches near their sites of release (Bell *et al.*, 2005).

Crocodilians

The skins of crocodilians have long been used for expensive leather goods, and hunting has depleted many populations. The recovery of the American alligator, *Alligator missippiensis*, following passage of protective legislation shows the potential for managing and restoring crocodilians. Success requires enforcement of laws and control of commerce, and these are difficult goals in many areas where crocodilians occur.

Some of the threats to crocodilians result from their semiaquatic life style. Like other large, slow-moving aquatic vertebrates, crocodilians are injured by collisions with watercraft (Grant and Lewis, 2010).The position of crocodilians near the top of aquatic food chains may make them susceptible to the effects of pollution. The pesticides dicofol, DDT, and their metabolites have reduced the hatching success and hatchling viability of alligators in Lake Apopka and have interfered with normal sexual development (Guillette *et al.*, 1994; Milnes *et al.*, 2008). Dicofol and DDT have estrogen-analogue properties and have affected the development of both sexes. Six-month-old female alligators from Lake Apopka had abnormal ovarian morphology and higher plasma levels of estradiol-17β than females from the control site. Male hatchlings from Lake Apopka had abnormally small penises and plasma testosterone levels, only one-quarter of those of males from the control site.

Lepidosaurs

Most extant species of lepidosaurs are relatively small. Of course, small size does not protect a species, but it does tend to make its problems less conspicuous – and hence less appreciated – than those of animals like sea turtles and crocodilians. With a few exceptions, which have earned special status by being unique (e.g., tuatara) or colorful (the San Francisco garter snake, *Thamnophis sirtalis tetrataenia*), the status of populations of lepidosaurs is unknown.

Larger species of snakes and lizards are killed for food and for their skins. Professional hunters in Central and South America ship iguanas (*Iguana*) to food markets in the cities and collect tegu lizards (*Tupinambis*) for their skins. Monitor lizards (*Varanus*) are hunted for their skins in Africa and Asia. In 2000, almost 1.25 million live lizards and snakes and nearly 3 million lizard and snake skins were traded legally. Indonesia, Malaysia, Singapore, Vietnam, Sudan, El Salvador, Argentina, Columbia, and Paraguay were the largest exporters, and Japan, Germany, Spain, and the United States were the largest importers (World Resources Institute, 2007). The lizards and snakes hunted for the leather trade are large animals with long adult life spans and are high in the food chain. There is little information about the status of wild populations of most of these species, but it seems unlikely that animals with these biological characteristics can withstand high rates of human predation indefinitely.

See also: Amphibians, Biodiversity of. Birds, Biodiversity of. Endangered Amphibians. Endangered Reptiles. Fishes, Biodiversity of. Mammals, Biodiversity of. Vertebrates, Overview

References

Andrews RM and Pough FH (1985) Metabolism of squamate reptiles: Allometric and ecological relationships. *Physiological Zoology.* 58: 214–231.

Araújo MB, Thuiller W, and Pearson RG (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography* 33: 1712–1728.

Beeson AA and Cree A (2010) A cold-adapted reptile becomes a more effective thermoregulator in a thermally challenging environment. *Oecologia* 163: 571–581.

Beeson AA and Cree A (2011) Integrating physiology into conservation: An approach to help guide translocations of a rare reptile in a warming environment. *Animal Conservation* 14: 28–37.

Bell CDL, Parsons J, Austin TJ, Broderick AC, Ebanks-Petrie G, and Godley BJ (2005) Some of them came home: The Cayman Turtle Farm headstarting project for the green turtle Chelonia mydas. *Oikos* 39: 137–148.

Bennett AF and Nagy KA (1977) Energy expenditure in free-ranging lizards. *Ecology* 58: 697–700.

Bickford D, Howard SD, Ng DJJ, and Sheridan JA (2010) Impacts of climate change on the amphibians and reptiles of Southeast Asia. *Biodiversity and Conservation* 19: 1043–1062.

Caldwell MW and Lee MSY (1997) A snake with legs from the marine Cretaceous of the Middle East. *Nature* 386: 705–709.

Castoe TA, Jiang ZJ, Gu W, Wang ZO, and Pollock DA (2008) Adaptive evolution and functional redesign in core metabolic proteins in snakes. *PloS One* 3(5): e2201.http://dx.doi.org/10.1371/journal.pone.0002201.

Chamaillé-Jammes S, Massot M, Aragón P, and Clobert J (2006) Global warming and positive fitness response in mountain populations of common lizards *Lacerta vivipara*. *Global Change Biology* 12: 392–402.

Chiappe LM (2007) *Glorified Dinosaurs: The Origin and Early Evolution of Birds.* New York: John Wiley & Sons.

Clark JM (1997) Patterns of evolution in Mesozoic Crocodyliformes. In: Fraser NC and Sues H-D (eds.) *In the Shadow of Dinosaurs: Early Mesozoic Tetrapods*, pp. 84–97. Cambridge, UK: Cambridge University Press.

Congdon JD, Dunham AE, and van Loben Sels RC (1994) Demographics of common snapping turtles (Chelydra serpentina): Implications for conservation and management of long-lived organisms. *American Zoologist* 34: 397–408.

Cox CL and Secor SM (2008) Matched regulation of gastrointestinal performance in the Burmese python, *Python molurus*. *The Journal of Experimental Biology* 211: 1131–1140.

Doody JS, Guarino E, Georges A, Corey B, Murray G, and Ewert M (2006) Nest site choice compensates for climate effects on sex ratios in a lizard with environmental sex determination. *Evolutionary Ecology* 20: 307–330.

Doody JS and Moore JA (2010) Conceptual model for thermal limits on the distribution of reptiles. *Herpetological Conservation and Biology* 5: 283–289.

Dunham AE (1993) Population responses to environmental change. In: Kareiva PM, Kingsolver JG, and Huey RB (eds.) *Biotic Interactions and Global Changes*, pp. 95–119. Sunderland, MA: Sinauer.

Ehmann H and Cogger H (1985) Australia's endangered herpetofauna: A review of criteria and policies. In: Grigg G, Shine R, and Ehmann H (eds.) *The Biology of Australasian Frogs and Reptiles*, pp. 435–461. Chipping Norton, NSW, Australia: Surrey Beatty.

Ernst CH and Barbour RW (1989) *Turtles of the World.* Washington, DC: Smithsonian Institution Press.

Etheridge R and de Queiroz K (1988) A phylogeny of the Iguanidae. In: Estes R and Pregill G (eds.) *Phylogenetic Relationships of the Lizard Families, Essays*

Commemorating Charles L. Camp, pp. 283–367. Stanford, CA: Stanford University Press.
Evans SE and Jones MEH (2010) The origin, early history, and diversification of the lepidosauromorph reptiles. In: Bandyopadhyay S (ed.) *New Aspects of Mesozoic Biodiversity. Lecture Notes in Earth Sciences* 132. pp. 27–44.
Fordham DA and Brook BW (2010) Why tropical island endemics are acutely susceptible to global change. *Biodiversity and Conservation* 19: 239–342.
Foufopoulos J, Kilpatrick AM, and Ives AR (2011) Climate change and elevated extinction rates of reptiles from Mediterranean Islands. *The American Naturalist* 177: 119–129.
Frost DR and Etheridge R (1989) A phylogenetic analysis and taxonomy of iguanian lizards (Reptilia: Squamata). *Miscellaneous Publications of the Museum of Natural History, University of Kansas* 81: 1–65.
Gans C (1996) An overview of parental care among the Reptilia. *Advances in the Study of Behavior* 25: 145–157.
Gauthier JA, Kluge AG, and Rowe T (1998) Amniote phylogeny and the importance of fossils. *Cladistics* 4: 105–209.
Grant PBC and Lewis TR (2010) High speed boat traffic: A risk to crocodilian populations. *Herpetological Conservation and Biology* 5: 456–460.
Greene HW (1988) Antipredator mechanisms in reptiles. In: Gans C and Huey RB (eds.) *Biology of the Reptilia*, vol. 16. pp. 1–152. New York: A. R. Liss.
Greene HW (1997) *Snakes, The Evolution of Mystery in Nature*. Berkeley, CA: University of California Press.
Greer AE (1989) The Biology and Evolution of Australian Lizards. Chipping Norton, NSW, Australia: Surrey Beatty.
Guillette Jr. LJ, Gross TS, Masson GR, Matter JM, Percival HF, and Woodward AR (1994) Developmental abnormalities of the gonad and abnormal sex hormone concentrations in juvenile alligators from contaminated and control lakes in Florida. *Environmental Health Perspectives* 102: 680–688.
Hamman M, Limpus CJ, and Read MA (2007) Vulnerability of marine reptiles in the Great Barrier Reef to climate change. In: Johnson JE and Marshall PA (eds.) *Climate Change and the Great Barrier Reef*, pp. 445–496. Townsville, Australia: Great Barrier Reef Marine Park Authority and Australian Greenhouse Office.
Hawkes LA, Broderick AC, Godfrey MH, and Godley BJ (2007) Investigating the potential impacts of climate change on a marine turtle population. *Global Change Biology* 13: 923–932.
Hay JM, Sarre SD, Lambert DM, Allendorf FW, and Daugherty CH (2010) Genetic diversity and taxonomy: A reassessment of species designation in tuatara (Sphenodon: Reptilia). *Conservation Genetics* 11: 1063–1081.
Hayssen V and Lacey RC (1985) Basal metabolic rate in mammals: Taxonomic differences in the allometry of BMR and body mass. *Comparative Biochemistry and Physiology* 81A: 741–754.
Huchzermeyer FW (2003) *Crocodiles: Biology, Husbandry and Disease*. Cambridge, MA: CABI Publishing.
Huey RB, John-Alder H, and Nagy KA (1984) Locomotor capacity and foraging behavior of Kalahari lizards. *Animal Behaviour* 32: 41–50.
Hulin V, Delmas V, Girondot M, Godfrey MH, and Guillon J-M (2009) Temperature-dependent sex determination and global change: Are some species at greater risk? *Oecologia* 160: 493–506.
James MC and Mrosovsky N (2004) Body temperatures of leatherback turtles (*Dermochelys coriacea*) in temperate waters off Nova Scotia, Canada. *Canadian Journal of Zoology* 82: 1302–1306.
Joyce WG and Gauthier JA (2004) Paleoecology of Triassic stem turtles sheds new light on turtle origins. *Proceedings of the Royal Society* 271: 1–5. London B.
Kallimanis AS (2009) Temperature dependent sex determination and climate change. *Oikos* 119: 197–200.
Lang JW (1989) Social Behavior in Crocodiles and Alligators. In: Ross CA (ed.) *Crocodiles and Alligators*, pp. 102–117. New York: Facts on File.
Macey JR, Larson A, Ananjeva NB, and Papenfuss TJ (1997) Evolutionary shifts in three major structural features of the mitochondrial genome among iguanian lizards. *Journal of Molecular Evolution* 44: 660–674.
Milnes MR, Bryan TR, Katsu Y, *et al.* (2008) Increased posthatching mortality and loss of sexually dimorphic gene expression in alligators (*Alligator mississippiensis*) from a contaminated environment. *Biology of Reproduction* 78: 932–938.
Mitchell NJ and Janzen FJ (2010) Temperature-dependent sex determination and contemporary climate change. *Sexual Development* 4: 129–140.
Mitchell NJ, Kearney MR, Nelson NJ, and Porter WP (2008) Predicting the fate of a living fossil: How will global warming affect sex determination and hatching phenology in tuatara? *Proceedings of the Royal Society B* 275: 2185–2193.
Morjan CL (2003) Variation in nesting patterns affecting nest temperatures in two populations of painted turtles (*Chrysemys picta*) with temperature-dependent sex determination. *Behavioral Ecology and Sociobiology* 53: 254–261.
Norris TB, Rickards GK, and Daugherty CH (2004) Chromosomes of tuatara, Sphenodon, a chromosome heteromorphism and an archaic reptilian karyotype. *Cytogenetic and Genome Research* 105: 93–99.
Packard GC and Packard MJ (2000) Developmental plasticity in painted turtles. *Chrysemys picta. Functional Ecology* 14: 474–483.
Pianka ER and Vitt LJ (2003) *Lizards: Windows to the Evolution of Diversity*. Berkeley, CA: University of California Press.
Pieau C and Dorizzi M (2004) Oestrogens and temperature-dependent sex determination in reptiles; all is in the gonads. *Journal of Endocrinology* 181: 367–377.
Poloczanska ES, Limpus CJ, and Hays GC (2009) Vulnerability of marine turtles to climate change. *Advances in Marine Biology* 56: 151–211.
Pough FH (1980) The advantages of ectothermy for tetrapods. *American Naturalist* 115: 92–112.
Pough FH (1988) Mimicry and related phenomena. In: Gans C and Huey RB (eds.) *Biology of the Reptilia*, vol. 16. pp. 153–234. New York, NY: A. R. Liss.
Pough FH, Andrews RM, Cadle JE, Crump ML, Savitzky AH, and Wells KD (2004) *Herpetology*, 3rd edn. Saddle River, NJ: Prentice Hall.
Pough FH, Janis CM, and Heiser JB (2009) *Vertebrate Life*, 8th edn. San Francisco, CA: Benjamin Cummings.
Pritchard PCH (1996) The Galápagos tortoises: Nomenclatural and survival status. *Chelonian Research Monographs No. 1*. Lunenburg, MA: Chelonia Research Foundation.
Quinn AE, Georges A, Sarre SD, Guarino F, Ezaz T, and Graves JAM (2007) Temperature sex reversal implies sex gene dosage in a reptile. *Science* 316: 411.
Ruben JA, Hillenius WJ, Geist NR, *et al.* (1996) The metabolic status of some Late Cretaceous dinosaurs. *Science* 273: 1204–1207.
Saint Girons H (1985) Comparative data on lepidosaurian reproduction and some time tables. In: Gans C and Billett F (eds.) *Biology of the Reptilia 15*, pp. 59–148. New York, NY: John Wiley & Sons.
Sarre SD, Georges A, and Quinn A (2004) The ends of a continuum: Genetic and temperature-dependent sex determination in reptiles. *Bioessays* 26: 639–645.
Schwanz LE and Janzen FJ (2008) Climate change and temperature-dependent sex determination: Can individual plasticity in nesting phenology prevent extreme sex ratios? *Physiological and Biochemical Zoology* 81: 826–834.
Secor SM (2005) Evolutionary and cellular mechanisms regulating intestinal performance of amphibians and reptiles. *Integrative and Comparative Biology* 45: 282–294.
Secor SM (2008) Digestive physiology of the Burmese python: Broad regulation of integrated performance. *The Journal of Experimental Biology* 211: 3767–3774.
Secor SM and Diamond J (1998) A vertebrate model of extreme physiological regulation. *Nature* 395: 659–662.
Seigel RA and Collins JT (1993) *Snakes, Ecology and Behavior*. New York: McGraw-Hill.
Shine R (1985) Parental care in reptiles. In: Gans C and Huey RB (eds.) *Biology of the Reptilia*, vol. 16. pp. 275–329. New York, NY: A. R. Liss.
Shine R (1991) *Australian Snakes, A Natural History*. Balgowlah, NSW, Australia: Reed Books.
Shine R (1999) Why is sex determined by nest temperature in many reptiles? *TREE* 14: 185–189.
Shine R, Warner DA, and Radder R (2007) Windows of embryonic sexual lability in two lizard species with environmental sex determination. *Ecology* 88: 1781–1788.
Starck JM (2003) Shaping up: How vertebrates adjust their digestive system to changing environmental conditions. *Animal Biology* 53: 245–257.
Townsend TM, Larson A, Louis E, and Macey JR (2004) Molecular phylogenetics of Squamata: The position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. *Systematic Biology* 53: 735–757.
Tucker JK, Dolan CR, Lamer JT, and Dustman EA (2008) Climatic warming, sex ratios, and red-eared sliders (*Trachemys scripta elegans*) in Illinois. *Chelonian Conservation and Biology* 7: 60–69.
Vida IN and Hedges SB (2004) Molecular evidence for a terrestrial origin of snakes. *Proceedings of the Royal Society, London B* 271: S226–S229.
Vitt LJ and Pianka ER (1994) *Lizard Ecology, Historical and Experimental Perspectives*. Princeton, NJ: Princeton University Press.
Wainwright PC and Reilly SM (1994) *Ecological Morphology, Integrative Organismal Biology*. Chicago IL: University of Chicago Press.

Walther G-R, Post E, Convey P, *et al.* (2002) *Ecological Responses to Recent Climate Change. Nature* 416: 389–395.
Warner DA and Shine R (1998) The adaptive significance of temperature-dependent sex determination in a reptile. *Nature* 451: 566–568.
World Resources Institute (2007) *EarthTrends: Environmental Information.* Washington, DC: World Resources Institute. http://earthtrends.wri.org
Zaher H, Grazziotin FB, Cadle JE, Murphy RW, de Maura-Leite JC, and Bonatto SL (2009) Molecular phylogeny of advanced snakes (Serpentes, Caenophidia) with an emphasis on South American Xenodontines: A revised classification and descriptions of new taxa. *Papéia Avulsos de Zoologia* 49: 115–153.

Residents and Transients in the Fossil Record

Albert J van der Meulen, University of Utrecht, Netherlands
Pablo Peláez-Campomanes, National Museum of Natural History, CSIC, Spain

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, pp 1–7, © 2007, Elsevier Inc.

Explanation of Concepts and Terminology

Fossil Community—Membership Times

Calculation of residence times is only feasible under rare conditions, of which the two main ones are (i) the availability of a rich and dense fossil record from a restricted area allowing to reconstruct the community history in great detail and (ii) accurate age control of the rock succession that yielded the record. The Miocene in the Calatayud-Daroca basin in North Central Spain (Daams *et al.*, 1999a) meets the two requirements. Thus, van der Meulen *et al.* (2005) disposed of a well-dated, rich record of fossil rodents (~50,000 teeth) from 102 localities ranging from 17 to 10 Ma. Eighty-seven assemblages derived from four sections situated in an area of about 2 km^2 near Villafeliche (Province Zaragoza). The sections have been firmly correlated lithologically, biostratigraphically, and magnetostratigraphically, and form the framework of the locality succession. Seventeen additional localities are situated in an area less than 12 km away, and have been biostratigraphically interpolated to the Villafeliche succession. The sections in the studied area have been successfully correlated to the geomagnetic polarity timescale allowing the determination of sediment accumulation rates and the estimation of ages for the localities (Krijgsman *et al.*, 1994, 1996; Daams *et al.*, 1999b). The 102 localities form a discontinuous record, because the fossiliferous beds alternate with unfossiliferous and/or unwashable (limestone) ones. The studied assemblages represent the time interval from 16.8 to 10 Ma; the mean temporal resolution (number of localities divided by length of time interval) is about 66,000 years. The densest record is between 16.8 and 12 Ma giving a resolution of 46,000 years; the resolution of the later part (12–10 Ma) is comparatively poor (200,000 years).

van der Meulen *et al.* (2005) reconstructed membership times in two steps:

1. They calculated inferred ranges of the species and species lineages, that is, "modifying the observed taxon ranges by estimating the ends of ranges after the method of Barry *et al.* (2002), to statistically avoid taxon absences owing to insufficient sampling and chance" (p. E110):

$$P_i = 1 - (1 - p_i)^r$$

where P_i is the probability of finding a single specimen of taxon i in the localities adjacent to those containing the observed first or last occurrence (FO and LO); p_i the probability of finding a specimen of the species in the locality containing the last or first occurrence of the taxon; r the number of specimens in the localities for which the probability of taxon i being there is estimated; and r is cumulative. In contrast to Barry *et al.* (2002) who employed the mean abundance of a taxon over its range, van der Meulen *et al.* (2005) used the frequencies in the localities containing the FO and LO. "In the cases of a gap inside the encountered range of the taxon, the lower frequency found in the localities adjacent to the gap was used. … Following Barry *et al.* (2002), the absence in a locality is assumed to indicate true local extinction, when the calculated P_i equals 0.8 or is higher (The original text of van der Meulen *et al.* (2005) erroneously states "lower"). The calculated ends of ranges are referred to as inferred first and last appearance (IFA and ILA) to distinguish them from FO and LO yielded by the raw data. The differences between IFA and FO, and ILA and LO are quite variable. Some do not differ at all, and some differ very much. The greatest differences are caused by finds of single elements in very rich localities" (p. E118).
2. All remaining "gaps caused by absence in a single intermediate locality are closed (as they may also result from chance), and only then [they] closed remaining gaps ≤200,000 years" (p. E110). The 200 ky was chosen to average over the extremes of intermediate and small astronomical periodicities and their possible climatic and lithologic influence on the presence or absence of certain taxa. Thus, they assumed to have excluded those temporary absences of taxa, which resulted from the influence of the main astronomical cycles on deposition. Assumingly, the astronomical cycles lead also to habitat shifts; however, the study focused on general trends in community history.

Fossil Communities

The resulting ranges (**Figure 1**) are interpreted as the residence times of community members, and co-occurring inferred ranges at locality age as the composition of the community existing at that time. The resulting curve of number of community members (N_{mem}, **Figure 2**) through time represents a signal of community change, because it is not correlated to sample size, and is averaged over climatic and lithologic oscillations.

An assemblage from a single locality represents a past community from which it derived, but the existence of assemblage zones in biostratigraphy indicates that successive assemblages are usefully grouped in higher units, which are recognized spatially to a certain extent. Olson (1952) introduced the term chronofauna, "defined as a geographically restricted, natural assemblage of interacting animal populations that has maintained its basic structure over a geologically significant period of time" (p. 181). Today, however, extant communities are generally described as dynamic rather

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00238-0

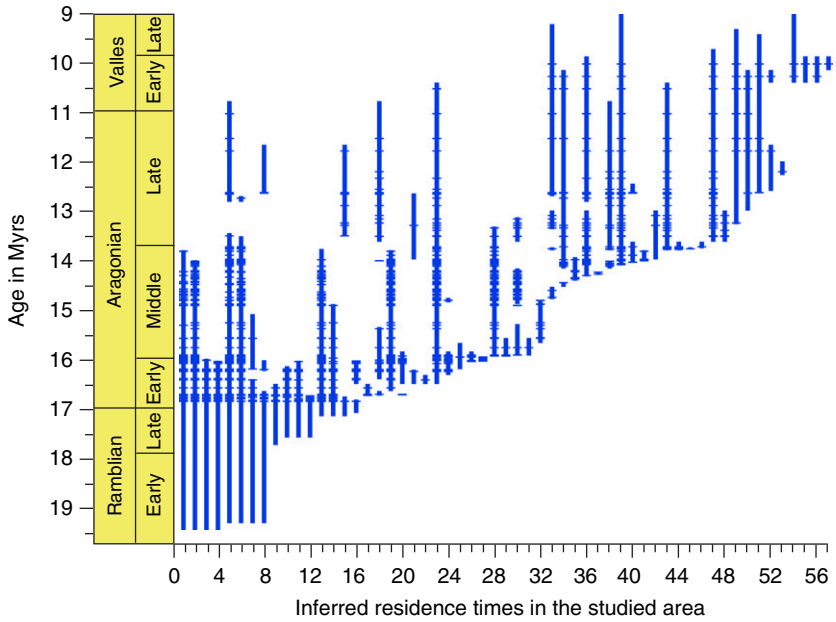


Figure 1 Observed occurrences (horizontal lines) and inferred ranges (vertical lines) of the 57 Miocene rodent taxa studied by van der Meulen *et al.* (2005). Miocene continental stages in Europe, and their correlation to the timescale (after Daams *et al.*, 1999b) are shown at the left. The entries before 17 Ma and exits after 10 Ma have been taken from the literature (Reproduced from Figure 2 in van der Meulen AJ, Peláez-Campomanes P, and Levin SA (2005) Age structure, residents, and transients of Miocene rodent communities. *Am. Natl.* 165: E110–E125.), with permission from Chicago University Press.

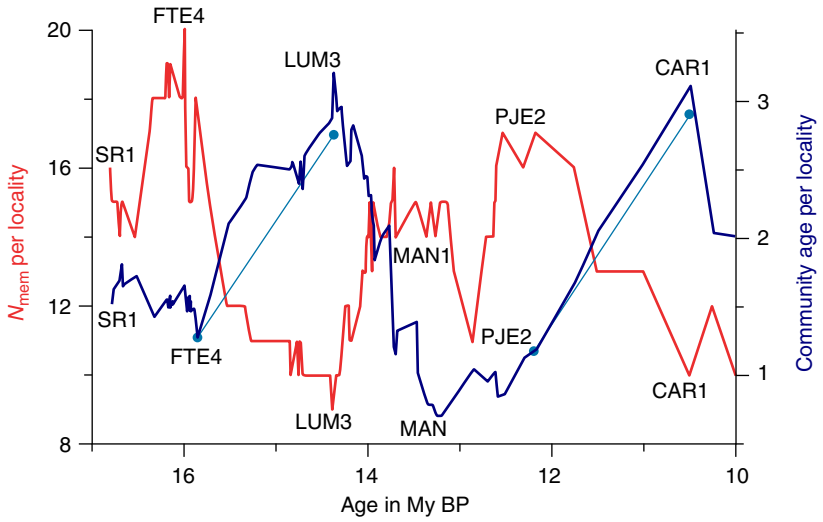


Figure 2 The temporal changes in the numbers of the rodent-community members (N_{mem}, red line) and in CA (dark blue line) based on the record of van der Meulen *et al.* (2005). The two lines mirror each other, illustrating the negative correlation between them, because entries increase N_{mem} and decrease CA, exits may cause the opposite effect (see text). The turnover between LUM3 and MAN is better represented by the CA curve than the N_{mem} curve, because after ~13.8 Ma entries and exits are more or less balanced. The two straight lines between FTE and LUM3, and between PJE2 and CAR1, are drawn to compare the CA increase and linear increase in time in these trajects with N_{mem} loss. The difference between them indicates preferential loss of later arrivals and retaining early arrivals (seniority rule, see text). Abbreviations refer to locality names (Reproduced from Figure 3 in van der Meulen AJ, Peláez-Campomanes P, and Levin SA (2005) Age structure, residents, and transients of Miocene rodent communities. *Am. Natl.* 165: E110–E125.), with permission from Chicago University Press.

than static entities, and the need to use a special term (chronofauna, paleocommunity, and faunal complex) for paleontologically recognized communities seems now less necessary than before. We use Tokeshi's (1999) shorthand for a community, as the continued assemblage of coexisting species, taking "continued" to refer both to ecological and evolutionary timescales. However, Olson's chronofauna and the Early-Middle and Late Aragonian–Early Vallesian communities discussed by van der Meulen *et al.* (2005) display considerably more compositional variation than an extant community even to the extent th'at a number of members never coexisted. Therefore, they are perhaps better related to

the metacommunity concept (Wilson, 1992), "a set of local communities that are linked by dispersal of multiple potentially interacting species" (Leibold *et al.*, 2004). The fossil metacommunity may be defined as a temporal set of communities linked by residents, long-term components of the fauna or flora of a certain area (see below). This raises the question: is it possible to delineate the temporal extent of a fossil metacommunity?

Turnover

Turnover refers to intervals of increased rates of taxon gains and/or losses, as a result of which the fauna or flora after the interval significantly differs from the ones before. Immigrations and/or speciation are responsible for the gains, emigration, and extinction for losses. There exists a vast amount of literature focusing on relating taxon turnover with environmental (climatic, tectonic, and volcanic) events of varying magnitudes (see e.g., Barnosky, 2001). Here, we are concerned with the question which turnover events cause the change from one metacommunity to another. Paraphrasing the question "Where does a community begin and end?" (Samuels and Drake, 1997, p. 427), the paleoecologist has to answer the question "When does a community begin and end?" van der Meulen *et al.* (2005) proposed the residents to define their fossil-rodent metacommunities, as these are the most continuous, coexisting members forming a core group of species while accompanying members (transients) come and go (**Figure 3**). Boundaries between successive metacommunities based on replacement of residents have to be drawn within turnover intervals, and are arbitrary. The turnover between the Early-Middle and Late Aragonian—Early Vallesian metacommunities lasted about 1 My between 14 and 13 Ma (**Figure 3**).

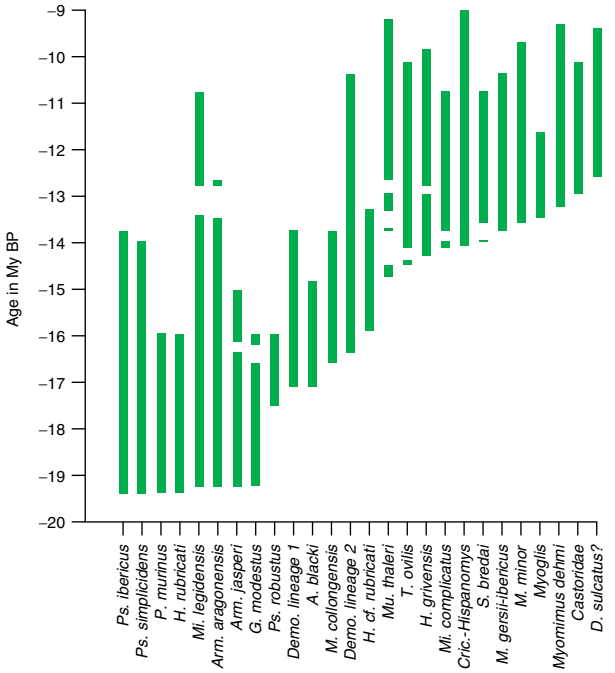


Figure 3 The ranges of the Aragonian residents showing the two communities which they define (van der Meulen *et al.*, 2005). The turnover phase which gives rise to a new community takes place between ~14 and ~13 Ma. One out of 14 residents of the earlier community, and six out of 14 residents of the later community, are associated with more forested habitats (Reproduced from Figure G1 in van der Meulen AJ, Peláez-Campomanes P, and Levin SA (2005) Age structure, residents, and transients of Miocene rodent communities. *Am. Natl.* 165: E110–E125.), with permission from Chicago University Press.

Community Age

Community age (CA) (**Figure 2**) refers to the mean of the residence times of the community members at the time of the locality age (van der Meulen *et al.*, 2005). "CA increases linearly when plotted against the decreasing ages of the localities during intervals without entries and exits. Entries inherently cause CA to decrease relative to the linearly expected CA, because newcomers contribute a low value to the sum of the RT_{mem} and simultaneously increase N_{mem}. In terms of its CA, an expanding community is rejuvenating. In contrast, the effects of exits vary depending on whether the RT_{exit}s' of disappearing taxa are below, above, or approximately equal to CA. In the latter case, there is little or no effect on the linear increase with time; when RT_{exit} is well below (or above) CA, the latter value will increase (or decrease) relative to the linearly expected one. So, disappearances of transient taxa cause true aging (i.e., more than linear increase) of the community, while all entries and the disappearances of long-time residents cause rejuvenation of the community" (p. E113).

Figure 2 shows significant negative correlation ($P<0.1$) between the number of community members and CA. The preferential disappearances of transients while residents remain during the time of taxon loss, has been termed seniority rule (van der Meulen *et al.*, 2005). In the paleontological literature, temporal species-richness variation is usually illustrated with curves representing the number of taxa against time (see the N_{mem} curve in **Figure 2**). However, these curves fail to illustrate intervals of maximal change of community composition, that is, phases of turnover, during which taxon losses are compensated by new entries, the net number of taxa varying very little (see the interval between 13.9 and 13.0 Ma in **Figure 2**). In contrast, the CA curve keeps dropping because both exits of long residents and entries lower CAs.

Residents and Transients

The distinction of the two components is based in the mean (1.54 My) of all residence times. Out of the 57 taxa, 26 of them with ranges $\geq$1.54 My were called residents, the remaining form the transient group (**Figure 4**). The histogram of all residence times indicates the existence of three groups, because of a group of six residents with very high persistence (further discussed below).

van der Meulen *et al.* (2005) analyzed the habitat preferences and demographic categories of the taxa recorded in their fossil communities. Residents appeared to be associated with early stages of succession, that is, they are supposed to have lived in open habitats, while transients are associated with later stages of succession, that is, living in more wooded

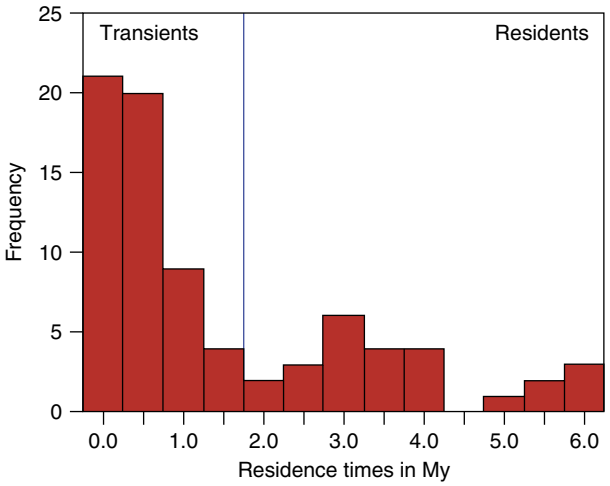


Figure 4 Distribution of the residence times based on the inferred ranges in **Figure 1**, to show the differences between residents and transients indicated by the low class of 2.0 My just above the mean (blue line), which has been taken as the divide. Additionally, a gap at the class of 4.5 My separates two groups of residents. The one to the right includes those residents of the earlier community, which continue into the later one (see **Figure 1**).

habitats. In ecological succession under stable conditions, early successional species are "transient" and are replaced by competitive later arrivals. To explain the reversed roles in the described fossil record, van der Meulen *et al.* (2005) assumed high degrees of disturbance causing the "competitively superior species to become transients because of their dependence on ephemeral late successional habitats" (p. E118).

Related Concepts in the Literature

Magurran and Henderson (2003) studied a 21-year dataset from an estuarine fish community, and demonstrated that commonness and rarity of the species is related to their permanence in the assemblage. The members could be clearly separated into two abundance–persistence groups: core species (abundant and persistent) and occasional species (rare and infrequent). This distinction of core and occasional species was facilitated by a gap in the distribution of persistence around 10 years. Ecologically, the core species are associated with the estuarine habitat and the occasional species with different habitats. van der Meulen *et al.* (2005) suggested that their residents and transients are long-term equivalents of the core and occasional species. Ongoing research of the present authors shows that relative abundance and residence times of the Miocene rodents of Spain are correlated, and the residence-time distributions shows gaps (**Figure 4**), similar to the results of Magurran and Henderson (2003). Furthermore, as expected by Magurran and Henderson (2003), the modes of the fossil residents and transients are much farther apart than their core and occasional species. The latter authors showed that the arrival of occasional species can be modeled as a stochastic event "often related to unusual weather conditions such as storms or unusually settled and warm periods" (p. 715). The fossil rodent transients have been related to ephemeral late successional habitats, while the residents are adapted to the prevalent conditions during the Miocene in Spain. The ephemeral conditions have to be taken in the context of evolutionary time: the transient residence times vary between 0.02 and 1.55 My. Furthermore, habitat preferences of the fossil rodents are not known in any detail.

However, perhaps the most significant commonality between the ecological and paleoecological study is their recognition of components among community members on the basis of persistence, abundance, and ecological requirements. Magurran and Henderson (2003) argue that "Explanations of species abundance distributions cannot … be divorced from biology," and neither can residence time of community members, as we will argue further in the next section.

Residents seem also to be comparable with Sheldon's long-term generalists, "species with properties that enable them to survive throughout wide environmental fluctuations over geological timescales" (Sheldon, 1996, p. 209; van der Meulen *et al.*, 2005). In the same paper, Sheldon (1996) introduces his Plus ça change model, which "proposes that morphological stasis is the usual response to widely fluctuating physical environments on geological timescales (until thresholds are reached)" (p. 209). Thus, the model predicts a tendency for generalists to show little or no evolutionary change. However, the rodent residents yield another, more complex pattern. Four of the six most persistent lineages (>5.0 in **Figure 4**) belong to the Gliridae (dormice), of which only one (*Armantomys aragonensis*—*A. tricristatus*) underwent modest anagenetic change in size and morphology permitting the distinction of two successive species (Daams, 1990). The remaining three glirids (*Microdyromys legidensis*, *Armantomys jasperi*, and *Pseudodryomys ibericus*) show the morphological stasis expected from the Plus ça change model. However, the two cricetid lineages (*Democricetodon* lineage 2 and *Cricetodon-Hispanomys* lineage) do show considerable evolutionary change. For example, the first lineage, recently revised by van der Meulen *et al.* (2003), comprises four distinct species. The second cricetid lineage needs further revision. These, admittedly few, examples do illustrate that the Plus ça change model, even though it is worded cautiously, needs to be refined.

It is also noteworthy, that the mentioned glirids and cricetids differ in their abundances. While the last often dominate the assemblages, the first do not in spite of their long persistence. So, it may be that phylogeny plays a role in explaining abundance adding further weight to the above quoted conclusion of Magurran and Henderson (2003).

Van Dam *et al.* (2006) studied a long (20 My) small mammal dataset including the record used by van der Meulen *et al.* (2005). The former authors estimate that some 80% of the FO and LO of lineages may be close to the ages of speciation and true extinction. They demonstrate that origination, extinction, and turnover rates constitute turnover cycles with periods of 2.4–2.5 and 1.0 My, which are linked to low-frequency modulations of Milankovitch oscillations (Laskar *et al.*, 2004). Pulses of turnover occur at minima of the 2.37 My eccentricity cycle and nodes of the 1.2 My obliquity cycle, which are both associated with expansion and cooling and affect regional precipitation. This "correlation between

orbital configurations, climatic events and rodent turnover strongly suggests that the latter is controlled by astronomically forced climate change" (Van Dam *et al.*, 2006, p. 690). This new astronomical hypothesis is consistent with the Court Jester (Barnosky, 2001) and Turnover Pulse hypothesis (Vrba, 1985) "in the sense that lineage events tend to cluster and that the clusters correlate to abiotic events" (Van Dam *et al.*, 2006, p. 690). A finding of interest to community residence times is that most turnover maxima are associated with an increase in the proportion in the short-lived species as indicated by the simultaneous occurrence of peaks of standing diversity and troughs in the mean lifespan per 0.1 My. Standing diversity is based on counts of those taxa that cross the boundaries of the intervals used (0.1 My in this case); taxa restricted to the interval are ignored. Mean lifespan is the mean of the total ranges of the lineages present at each 0.1 My interval. This relationship may be explained by the seniority rule (van der Meulen *et al.*, 2005). Seniority refers to the tendency that during episodes of community member loss later arrivals (transients) are the first to disappear, while residents are preferentially retained and continue the community. Thus, peaks of diversity are caused by transients and troughs represent the remaining residents.

Jackson and Erwin (2006) surveyed the literature for analyses of the fossil record that provide rigorously documented patterns to explore "the dynamics and underlying processes of ecological and evolutionary change in deep time" (p. 322). As to community membership they found that it is "commonly more stable and persistent than expected by chance, even in the face of the extreme environmental changes of the Ice Ages, …," taken from appropriate examples from reef studies (Pandolfi, 1996, 2002; Pandolfi and Jackson, 2001; Aronson *et al.*, 2004) and small mammals (McGill *et al.*, 2005). These studies are based on comparisons of Pleistocene community compositions with a far greater resolution than the fossil rodent study from which the residence groups were derived. Tokeshi (1999) correctly notes that community persistence (and resistance and resilience) depends on the definition of a community one adopts, and that, therefore, there is "no universal formula for assessing community persistence" (p. 354). However, in the reef studies community membership was compared with random sampling of the habitat-specific species pool, and McGill *et al.* (2005) adopted the neutral theory (Hubbell, 2001) as the null hypothesis to evaluate community similarities. The use of residents and transients is an alternative approach to these studies.

Concluding Remarks

Special conditions of Spain, being a peninsula with a generally dry climate (also during the Miocene, e.g., Daams *et al.*, 1999), have enabled the distinction between residents and transients, and make it an interesting area to study community dynamics. Astronomically forced humidity changes are reflected in changes of species numbers, which are higher during the more humid and lower during the more arid phases (Daams *et al.*, 1999, and literature therein; van der Meulen *et al.*, 2005; Van Dam *et al.*, 2006). The early successional residents are adapted to the typical Spanish conditions. During the more exceptional humid episodes, succession may advance to later stages and transients enter the communities, in which they generally do not replace but are added to the residents present.

Acknowledgments

We greatly appreciate the careful reviews by the Editor-in-Chief of the Encyclopedia of Biodiversity and by an anonymous. P.P-C acknowledges financial support by the MEC project number CGL2004-0294/BTE.

See also: Climate Change and Ecology, Synergism of. Climate Change and Extinctions. Modern Examples of Extinctions

References

Aronson RB, Macintyre IG, Wapnick CM, and O'Neill MW (2004) Phase shifts, alternative states, and the unprecedented convergence of two reef systems. *Ecology* 85(7): 1876–1891.

Barnosky AD (2001) Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the Northern Rocky Mountains. *J. Vertebr. Paleontol* 21: 172–185.

Barry JC, Morgan ME, Flynn LJ, *et al.* (2002) Faunal and environmental change in the Late Miocene Siwaliks of Northern Pakistan. *Paleobiol. Mem* 3(Supplement to *Palaeobiology* 28): 1–71.

Daams R (1990) Hypsodont Myomiminae (Gliridae, Rodentia) from the Miocene and the Oligocene–Miocene boundary interval of Spain. *Scripta Geol.* 95: 1–62.

Daams R, van der Meulen AJ, Alvarez Sierra MA, *et al.* (1999a) Stratigraphy and sedimentology of the Aragonian (Early to Middle Miocene) in its type area (North-Central Spain), *Newslett. Stratigraphy* 37(3): 103–139.

Daams R, van der Meulen AJ, Alvarez Sierra MA, Peláez-Campomanes P, and Krijgsman W (1999b) Aragonian stratigraphy reconsidered, and a re-evaluation of the Middle Miocene mammal biochronology in Europe. *Earth Planet. Sci. Lett.* 165: 287–294.

Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography.* Princeton, NJ: Princeton University Press.

Jackson JBC and Erwin DH (2006) What can we learn about ecology and evolution from the fossil record? *Trends Ecol. Evol.* 21(8): 322–328.

Krijgsman W, Garcès M, Langereis CG, *et al.* (1996) A new chronology for the middle to late Miocene continental record in Spain. *Earth Planet. Sci. Lett.* 142: 367–380.

Krijgsman W, Langereis CG, Daams R, and van der Meulen AJ (1994) Magnetostratigraphic dating of the Middle Miocene climate change in the continental deposits of the Aragonian type area in the Calatayud-Teruel basin (Central Spain). *Earth Planet. Sci. Lett.* 128: 513–526.

Laskar J, Robutel P, Joutel F, Gastineau M, Correira ACM, and Levra B (2004) A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys* 428: 261–285.

Leibold MA, Holyoak M, Mouquet N, *et al.* (2004) The metacommunity concept: A framework for multi-scale community ecology. *Ecol. Lett.* 7(7): 601–613.

Magurran AE and Henderson PA (2003) Explaining the excess of rare species in natural species abundance distributions. *Nature* 422: 714–716.

McGill BJ, Hadley EA, and Maurer BA (2005) Community inertia of Quaternary small mammal assemblages in North America. *Proc. Natl. Acad. Sci. USA* 201: 16701–16706.

Olson EC (1952) The evolution of a Permian vertebrate chronofauna. *Evolution* 6: 181–196.

Pandolfi JM (1996) Limited membership in Pleistocene reef coral assemblages from the Huon Peninsula, Papua New Guinea: Constance during global change. *Paleobiology* 22: 152–176.

Pandolfi JM (2002) Coral dynamics at multiple scales. *Coral Reefs* 21: 13–23.

Pandolfi JM and Jackson JBC (2001) Community structure of Pleistocene coral reefs of Curacao, Netherlands Antilles. *Ecol. Monogr.* 71: 49–67.

Samuels CL and Drake JA (1997) Divergent perspectives on community convergence. *Trends Ecol. Evol.* 12: 427–432.

Sheldon PR (1996) Plus ça change—a model for stasis and evolution in different environments. *Palaeogeogr. Palaeoclimatol. Palaeoecol* 127: 209–227.

Tokeshi M (1999) *Species Coexistence*, *Ecological and Evolutionary Perspectives*. Oxford, UK: Blackwell.

Van Dam JA, Abdul Aziz H, Alvarez Sierra MA, *et al.* (2006) Long-period astronomical forcing of mammal turnover. *Nature* 443: 687–691.

van der Meulen AJ, Peláez-Campomanes P, and Daams R (2003) Revision of medium-sized Cricetidae from the Miocene of the Daroca-Villafeliche area in the Calatayud-Teruel basin (Zaragoza, Spain). In: Martínez NL, Peláez-Campomanes P, and Fernández M (eds.) *En torno a Fósiles de Mamíferos: Datación, Evolución y Paleoambiente Colloquios de Paleobiología* Vol. Ext. 1, pp. 385–441. Madrid: Universidad Complutense.

van der Meulen AJ, Peláez-Campomanes P, and Levin SA (2005) Age structure, residents, and transients of Miocene rodent communities. *Am. Natl.* 165: E110–E125.

Vrba ES (1985) Environment and evolution: Alternative causes of the temporal distribution of evolutionary events. *Suid-Afrikaanse Tydskrif vir Wetenskap* 81: 1–22.

Wilson DS (1992) Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. *Ecology* 73: 1984–2000.

Resource Exploitation, Fisheries

John Beddington, Imperial College of Science, Technology and Medicine, London, UK
Ray Hilborn, University of Washington, WA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by John Beddington, volume 2, pp 161–172, © 2001, Elsevier Inc.

Glossary

Fishing effort Level of fishing activity quantified in terms of the number and power of vessels and duration of fishing.
Maximum sustainable yield Maximum average level of catch that can be removed from fish population over a very long time period.
Mesh size Size of the holes in a fishing net.
Recruits Young fish that are first caught by a fishery.
Stock recruitment relationship Relationship between the adult breeding stock and the number of recruits produced by that stock.
Total allowable catch Level of permitted catch established by fishery regulations.

History

Fishing as a commercial activity has been conducted since the very earliest of times. Early records indicate that it was occurring in both Egyptian and Indian societies as long ago as 5000 BC, and fisheries in the Mediterranean were already significant as long ago as 1000 BC.

Since the early years, fishing has increased largely as a result of the changes that have occurred in the human population and its activities. In **Figure 1(a)**, the estimates for world population from 1800 are shown. This shows an increase from around 1 billion in 1800 to 7 billion at present. The major increase in population has occurred in the past 100 years or so, and particularly since 1950 the rate of increase has been very high.

In **Figure 1(b)**, estimates of fish catch and aquaculture production since around 1800 are shown. The figure indicates a catch of around 2 million tonnes in 1800, increasing to approximately 20 million tonnes in 1950. Since that time, catches have increased to a 2006 level of around 144 million tonnes. The striking similarity between **Figure 1(a)**, that shows the world population, and **Figure 1(b)**, that shows the global fish catch, results from the fact that fisheries to a large extent have mimicked the change in population and the changes in human activity reflected by increasing industrialization. Capture fisheries production since 1990 has been roughly level and the growth since then is completely due to increasing aquaculture.

Fishing as a commercial activity can usefully be divided into two elements: that involved with catching and that involved with preserving and processing the catch.

Methods of Catching Fish

There are five main methods.

- Hooks and lines: This is perhaps the simplest method of fishing from the earliest times; handlines have been used, but current commercial longlines can be up to 100 km in length.
- Gill nets: In this operation, nets are hung from vessels or floating buoys that are largely invisible to the fish, which are caught as they try to pass through the net. Such fishing methods can catch many species other than the main target species and large high seas gill nets are now banned under international treaties.
- Trawls: This type of fishing involves a vessel dragging a large net (or trawl). This can be done either through the

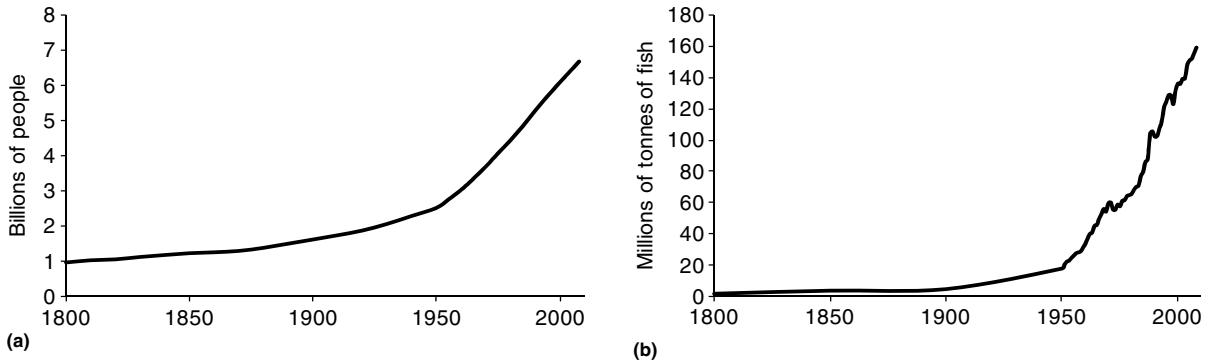


Figure 1 (a) World population (Klein Goldewijk and van Drecht, 2006), with permission from Netherlands Environmental Assessment Agency. (b) World fish catch, 1800–2008, including catches from inland fisheries and aquaculture from FAO and Hilborn *et al.*, 2003. Capture fisheries alone were around 92 million tonnes in 2006 (Anon, 2009), with permission from FAO.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00055-1

water, close to or along the bottom of the sea (demersal trawling), or in the water column (pelagic trawling).

- Seines: The principle of this operation is that a school of fish is encircled by a net, which is then tightened, reducing the circumference of the net until the school is caught. In modern tuna purse seining, a fast motor-boat holds the end of the net steady while the main vessel encircles a shoal of tuna releasing the net as it goes. The net is then pulled in from the main vessel.
- Traps: Fish are usually attracted to a baited trap that allows easy entry but difficult exit. Traps are most commonly used for invertebrates like crabs and lobsters, but many fish can also be caught in traps.

Although the basic methods of fishing are little changed since earliest times, the development of motor vessels to succeed sailing or manpower and the increasing power of engines and gear to locate fish has dramatically increased the fishing power of vessels over the years.

Processing the Catch

Original methods for processing or preserving fish involved sun drying, smoking, salting, and pickling. Subsequent developments as a result of industrialization have involved canning, preserving fish on ice, freezing, and processing on board the vessel, and reduction to meal. Modern factory vessels are highly industrialized operations that catch, process, and package fish, ready for delivery directly to markets.

Changes Since 1945

The industrialization prior to World War II increased the pressure on fish stocks, but the war provided a brief respite as commercial fishing activity was reduced, for obvious reasons. Analyses of this period indicate significant recoveries in major fish stocks that had been heavily exploited prior to the start of the war.

Since then, total fish catch has increased by almost a factor of 5. Much of the increase has come from the coastal regions, where fishing activity has expanded with increase in population, creating more demand for fish products. However, in this period, an activity of global significance has been the development of distant-water fishing by fleets primarily, but not exclusively, of the USSR and Japan. These fleets consisted of large vessels capable of operating for many months away from the home port and with substantial industrial freezing capacities. In the maritime regimes applying in the 1950s and 1960s, these vessels fished the rich continental shelf waters off the coast of many countries. In this period, the catch of distant-water fleets increased from around 500,000 tonnes in 1950 to a peak of 8.5 million tonnes in 1972. The change in the maritime regime brought about by the UN Conference on the Law of the Sea and consequent UN Convention on the Law of the Sea has altered this in a dramatic way. Coastal states have set up Exclusive Economic Zones (EEZs) and restricted fishing on their continental shelf waters. This has resulted in a decline in the proportion of the world fish catch coming from distant-water fleets from 15% in 1972 to 5% by the late 1990s. The most important change has been the stagnation of catch; there was no increase in capture fisheries production from 1990 and there appeared to be no new significant fish resources to exploit. Any increase in fish production will have to come from aquaculture.

Fisheries Management

The need to manage fisheries as an activity has been recognized since early times, and as long ago as the fifteenth century, Dutch vessels operating in the North Atlantic had a closed season and restricted gear types in order to reduce pressure on fish stocks. Probably with an analogy to farming, it was also recognized that in addition to restrictions on season and types of gear, there was merit in protecting young fish that had yet to breed or fish that were currently breeding. However, regulations were comparatively few and were largely unenforceable except via peer pressure at a local level. The need for management was recognized at an international level after World War II, and a number of international fishery commissions were set up. These varied from the International Whaling Commission, with its focus on whaling worldwide, to commissions specifically aimed at dealing with fisheries in a particular area, which were the most typical. These commissions were set up via international conventions agreed between the coastal states of the region with fishing interests and states with distant-water fishing fleets operating in the area.

It is fair to say that with rare exceptions these international fishery commissions were largely ineffective in directly regulating the level of fishing pressure on fish stocks either via restricting the amount of fishing (effort) or by regulating the catch. The main difficulty was that the commissions typically found the necessary degree of consensus difficult to achieve among states with differing economic imperatives and the commissions relied on countries to enforce regulations on their own fishermen. Most agreements reached concerned readily enforceable regulations such as restrictions on gear types and on fishing seasons. Despite these difficulties, the fisheries management in these bodies was dependent on proper scientific assessment (discussed below) and routinely collected statistics on the catches of species under exploitation. Such data have proved invaluable in assessing the history of exploitation of particular fish stocks and are still of use today. A major exception is the International Pacific Halibut Commission, founded in 1923 which has been very effective. Unlike other commissions, it has only two member nations, the U.S. and Canada.

Post-UN Convention on the Law of the Sea

Since the Law of the Sea convention came into force, the need for international regulation of fisheries by international commissions has been reduced. Most states now have an Exclusive Economic Zone (EEZ) which extends 200 miles from their coast. The reason for 200 miles is that this is an approximation to the extent of the continental shelf where, with very rare exceptions, the vast majority of fish are caught, the nutrient-rich environment providing food for plankton and the fish communities dependent on it. Management in these EEZs has had mixed success, but they do in principle provide

the opportunity for a single entity, the coastal state, to directly regulate the fisheries in the area.

The Current Situation

The UN Food and Agriculture Organization (FAO) has been collecting comprehensive data on fisheries since around 1950. Their figures indicate that the total world catch of fish is dominated by a small group of very large states, or those states that have particularly rich fish resources or a high level of distant-water fleet activities. In fact, 12 countries take around 70% of the total world catch (**Table 1**).

In the context of biodiversity, the world fish catch is dominated by a relatively small number of species: some 30% of the total world catch consists of only 11 species (**Table 2**), and the many hundreds of other species caught only form a very modest proportion of this total catch. Nevertheless, fisheries for certain highly valued species such as squid, shrimp, and tuna are extremely important in global terms despite the volume of catch being small.

The development of commercial fisheries in the last three decades of the twentieth century has led to such pressure on fish resources that the paramount question is whether the stocks can continue to sustain these levels of exploitation. Such questions have been posed for a number of decades and it has been recognized for several centuries that overexploitation can occur unless proper management action is taken. However, the current situation at a global level is clearly of serious concern. In a 2008 study by FAO, it has been calculated that of those important commercial fish stocks, where sufficient information was available to allow judgment to be made, 52% are fully exploited, 19% are overfished, 8% are depleted with no evidence of any recovery, and 1% are depleted with some evidence of recovery. In other words, almost 70% of the key commercial fish stocks in the world are fully exploited or overexploited and 28% have been depleted to an extent that their capacity to deliver a sustainable yield has been seriously eroded.

The total catch of commercial fish stocks appears to be stabilizing around some 90 million tonnes. It is important to recognize that this apparent stabilization is not an intrinsic characteristic of the biological productivity of the oceans but rather a combination of that productivity and the plethora of individual fisheries that exploit it. Indeed, productivity varies quite significantly from area to area in the world. It is estimated that perhaps 15–20 million tonnes of additional catch could be obtained if all stocks were well managed.

Table 1 Fishery catches in 2006 by country (aquaculture not included)

Country	*Catch ($\times 10^6$ tonnes)*
China	17.1
Peru	7.0
United States	4.9
Indonesia	4.8
Japan	4.2
Chile	4.2
India	3.9
Russian Federation	3.3
Thailand	2.8
Philippines	2.3

Table 2 Fishery catches in 2006 by main species

Species	*Catch ($\times 10^6$ tonnes)*
Anchoveta	7.0
Alaska pollock	2.9
Skipjack tuna	2.5
Atlantic herring	2.2
Blue whiting	2.0
Chub mackerel	2.0
Chilean jack mackerel	1.8
Japanese anchovy	1.7
Largehead hairtail	1.6
Yellowfin tuna	1.1

Economic Issues

The historical changes that have led to the current level of overexploitation of the world's fisheries have been paralleled by a somewhat unsatisfactory economic situation. Studies on changes in fishing capacity are relatively few as data are often difficult to come by and changes in overall levels of fishing are difficult to quantify, particularly where diverse regimes and types of fishing vessel operate. However, a recent attempt at synthesis has been made by FAO by referring to the number of decked vessels, which are reported as being licensed by states. The distinction between a decked and an undecked vessel is a reasonable one between primarily commercial vessels, which have a reasonable level of construction, and simple undecked vessels, canoes, etc. Such data that are available indicate that the number of undecked vessels primarily operating in subsistence fisheries in Africa and Asia have remained constant over the past two decades. By contrast, since 1970, the number of decked vessels has doubled. It is worth noting that the vast majority of these increases in vessel numbers have come from one state alone: China. In the period between 1980 and 2005, FAO reported that the number of decked vessels in China increased from 60,000 to 500,000.

This significant increase in the capacity of fleets as measured by the number of vessels also conceals the phenomenon that, on the whole, vessels are getting larger and more technologically sophisticated. Their costs are also correspondingly higher and thus the value and operating costs of the vessels of the world fleet in recent times are likely to be significantly more than those of a few decades ago–certainly more than the simple doubling indicated by the statistics.

In a 1997 study, Garcia and Newton analyzed the behavior of the world fleet in a very simple but informative way. They calculated that the total revenue obtained by the world's fishing fleets (essentially the total first-sale value of the world catch) was less than the costs of operating the fleets by a very large amount. They quoted a figure of some $54 billion for the operating loss in the year 1989.

How such operating losses are being sustained is an obvious question, to which there is an equally simple answer. These losses have got to be sustainable by either public, i.e., state, or private subsidies of the fishing operations. Given that most fishery operators are small, owning at most a few vessels, it is quite clear that such losses could not be sustained by the private sector and hence it is the state subsidies that are underwriting the continuing losses of the global fleet. In a 2009 study, the World Bank and FAO estimated that there was $50 billion in lost profits in world fisheries; lost because of excess fleet capacity and fishing effort (World Bank, 2009).

This excess capacity and fishing effort has two unfortunate consequences. First, it means that substantial world resources are being devoted to the continuation of a failing economic activity. Second, vessels are being subsidized to continue in a fishery when normal economic forces would have driven them out much earlier. Hence the fishing pressure on already heavily exploited resources has been continued artificially.

This issue of subsidy in world fisheries has provoked major concern since the work of Garcia and Newton. A recent examination by Milazzo for the World Bank indicated the scale of the problem and showed how substantial and ubiquitous these subsidies are (Milazzo, 1998). Clearly, an immediate removal of subsidies from the fishing fleets around the world would lead to quite unacceptable levels of social deprivation and loss of food supplies. However, the overcapacity of fishing fleets has serious consequences for the continuation of sustainable fishing. States have recognized this and in some areas a reverse process is happening. For example, within the European Union there is a program, the Multi Annual Guidance Program, which is aimed at ensuring, by fiscal means, a reduction in the European fleet. This and other mechanisms for reducing fishing effort are going to be essential if fisheries are to be managed sustainably in the future.

Ecosystem Considerations

Ecosystem impacts of fishing

Much of the discussion in the review so far has focused on the general characteristics of fisheries and it has treated individual fish resources as if they were in isolation from the ecosystem in which they are embedded. However, the level of exploitation that has been occurring has been accompanied, as might be expected from simple ecological principles, by a change in the species composition of the ecosystems that are being exploited. Even in well regulated fisheries the abundance of target species is often reduced by 50–80%, and when fisheries are not well regulated the decline in ecosystem abundance may be much greater. Excess fishing can also cause major changes in trophic structure. Studies have shown that overfishing cod and other demersal species can lead to major increases in shrimp, crabs, and lobster, which are called "trophic cascades." Destructive fishing practices such as the use of explosives or bottom trawls and dredges on sensitive habitats can eliminate much of the native biota from an area. Many efforts are underway to ban such destructive practices but it must be recognized that the impacts of bottom trawls and dredges are very habitat dependent and on soft bottom or habitats subjected to frequent natural disturbance the impact on native biota is much less than on hard bottoms and rarely disturbed habitats.

The Issue of discards

It is well known that certain commercial fisheries operate inefficiently, in that the species targeted are not the only ones to be removed, damaged, or killed. In 1994, Alverson and coworkers (Alverson *et al.*, 1994) estimated that the annual level of discards in commercial fisheries is on average some 27 million tonnes (a very significant portion of the overall total world catch). This occurs primarily in fisheries where the gear, almost by definition, is unselective and unwanted species are caught. The elimination of such discards by gear change and regulation, and the utilization of the catch for human consumption offers a significant potential for increasing overall productivity. Later estimates suggest the level of discarding has declined, in large part to utilization of less valuable species in aquaculture feed, and some estimates are now that the level may be less than 10 million tonnes per year.

Fisheries Science

Although work in the late nineteenth century anticipated some developments, the main pioneers of fishery science worked in the first two decades after World War II. Of the many scientists involved in the development of fisheries as a science, five individuals stand out: Schaefer, an American working at the Inter American Tropical Tuna Commission; Ricker, a Canadian working at the Nanaimo Laboratory in Canada; and three British scientists, Beverton, Gulland, and Holt, all working at the Fisheries Laboratory in Lowestoft. It is their work that set the scene for fisheries science as it is today. All of this work was developed before the availability of modern computing power and the scale of their achievements, set in this context, is impressive.

Maximum Sustainable Yield (MSY)

The basic question of fisheries science is how exploitation and removal of the individuals from a population, affects the dynamics of that population: the processes of birth, growth, and death. A central concept in this exploration is sustainable yield. The idea is simple but it requires an explanation in the context of population ecology. In its unexploited state, a population will on average be in balance with its environment, the processes of birth and growth will be offset by death, and the population will be at equilibrium. A simple differential equation, commonly used in other branches of ecology, expresses this idea well. The rate of change of the biomass B of a population over time is given by

$$dB/dt = rB(1 - B/B_0) \quad [1]$$

When the biomass is equal to B_0, the rate of change is zero and the population stays at its equilibrium (unexploited) level B_0. When B is very small compared to B_0, the population will grow exponentially at a rate r. At intermediate levels, the

rate of change will decline from its maximum at $B = 0$, as B increases to zero when $B = B_0$. It is possible to plot the right-hand side of eqn [1] to show how the increase in biomass of a population is affected by its size.

This is done in **Figure 2(a)**. The figure shows that there is a maximum increase in biomass, which occurs at a level where the biomass is equal to $B_0/2$. Some simple calculus can be used to show that this maximum is given by the expression $rB_0/4$. If exploitation is removing biomass from the population, then as long as removals are less than the maximum, they will be replenished by the population. These removals (the yield) are thus sustainable so long as they do not exceed this maximum, which is called the maximum sustainable yield (MSY). The MSY will vary with the abundance of the fish stock concerned (B_0) and its capacity to increase (r).

The exploitation of the population can be modeled by a simple modification of eqn [1]. Let p be the proportion of the population that is removed by exploitation. The equation for exploitation is then

$$dB/dt = rB(1 - B/B_0) - pB \qquad [2]$$

This is shown by a simple modification of **Figure 2(a)** in **Figure 2(b)**.

The expression for the removals pB is a simple straight line with slope p. The larger the proportion removed, the steeper the line. The removals are replenished by the increase in biomass exactly where the line intersects the curve showing the increase in biomass. Simple calculations can show that where $p = r/2$, the maximum sustainable yield is reached. Where p is greater than this level, there will be a sustainable yield, but it will be less than the maximum and the biomass of the population will be reduced below the level where it produces the maximum sustainable yield. Increasing the intensity of harvesting then will actually reduce rather than increase the yield. By contrast, if p is less than $r/2$, increasing the intensity of harvesting will increase the yield.

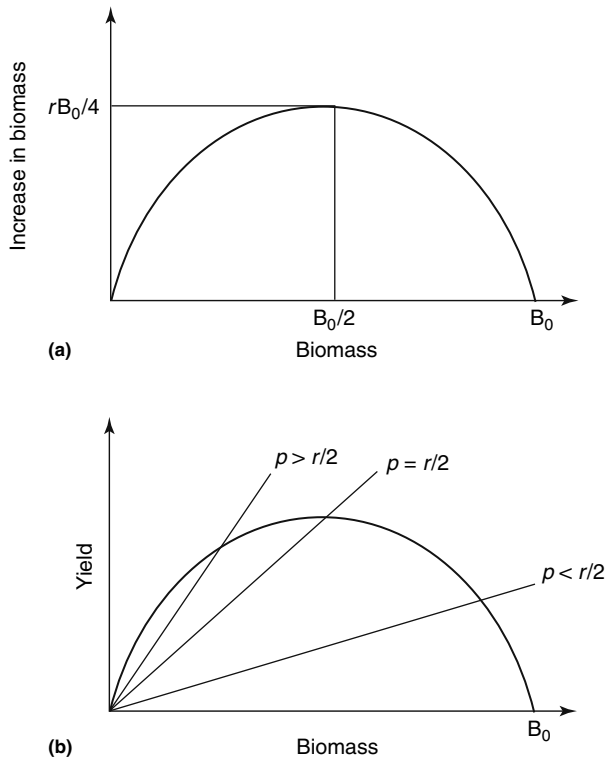


Figure 2 (a) Rate of change (increase) of biomass related to biomass: The position of stock size when MSY taken is $B_0/2$ and the MSY level is $rB_0/4$. (b) Catch (yield) levels and the rate of change of biomass. Where intersection occurs, an equilibrium yield is being taken.

In the practical situation, the proportion of the fish stock removed will be determined by the level of fishing effort: the number of boats, their size, the type of fishing gear, the duration of fishing, etc. If the level of fishing effort is less than that required to take the maximum sustainable yield, increasing the effort will result in an increased yield. If, however, the level of fishing effort is higher than that required to take the maximum sustainable yield, further increases in that effort will result in a decrease in sustainable yield. Such a situation is often termed as overcapacity, as removing fishing effort will actually increase yield in the longer term. It should be emphasized that the idea of MSY developed above is an oversimplification and that environmental variability and other factors render its simple application problematic.

The simple models used to illustrate the ideas of sustainable yield and fishing capacity do not allow proper exploration of some of the other key areas of fisheries science. To explore these requires a deeper understanding of the way in which fish populations behave. The idea can be developed simply using an idealized fish population as a subject; here young fish develop from their larval form and become part of the population at a particular time of the year. A year later, they would have grown and some of the cohort would have died, either through predation or other natural causes. The process continues and a typical temperate fish population will consist of a number of age groups, with fewer fish surviving to greater and greater ages. This is illustrated in **Figure 3(a)**.

Exploitation of this population can, in principle, start at any age, but in practice will be determined by fishing methods. For example, the size of the net will determine what age groups are caught; smaller and younger fish can escape through the mesh. **Figure 3(b)** shows how this will affect the age structure of the population. In the exploited population, fewer fish survive to greater ages and the heavier the exploitation, the more extreme will be the reduction in older fish. However, the biomass of the population is affected not just by natural deaths and fishing but also by growth. As fish get older, they increase in size according to some relationship. **Figure 4** illustrates this for a typical fish.

In a comprehensive work first published in 1957, Beverton and Holt explored how the yield could be affected by changing the age at which fishing commenced and the intensity of fishing. Their results show how the interplay between these two parameters, which can be changed by fishing practice, can alter the level of sustainable yield. A centrally important result is that in many cases, delaying the age at which fishing starts to operate can improve yields. This was in part the basis for many national regulations and international agreements to regulate the mesh sizes of nets and the type of fishing gear permitted.

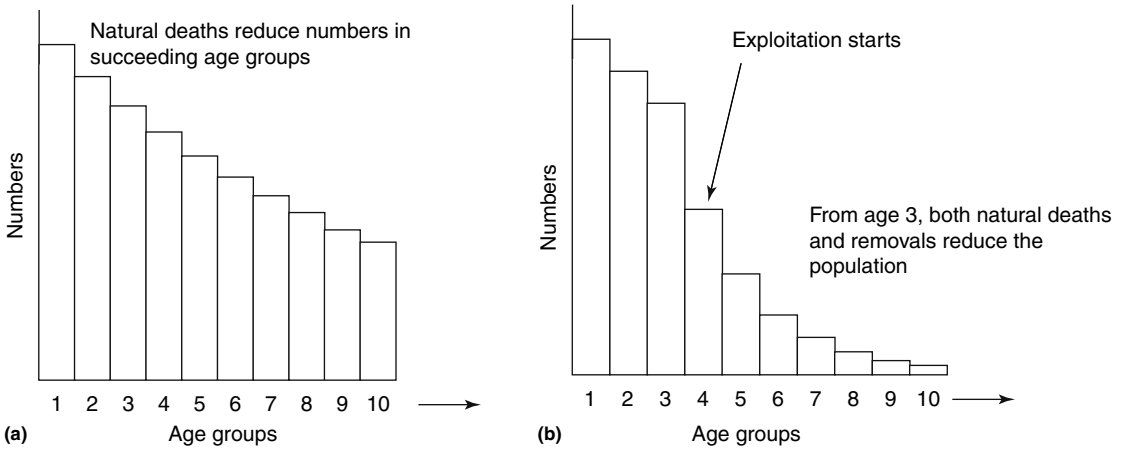


Figure 3 (a) Numbers of fish in each age group in an unexploited population. (b) Numbers of fish in each age group where fishing starts from age 3.

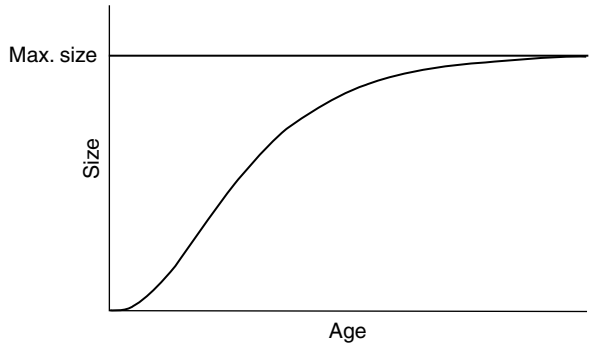


Figure 4 Schematic growth curve for typical fish growth.

Stock and Recruitment

One further element is missing from the description of the dynamics of an exploited fish population: reproduction. Typical fish are enormously fecund, laying as many as several hundred thousand eggs at a time. The eggs and the larvae into which they develop become part of the plankton, where they are subjected to high levels of mortality from predators and from starvation. To illustrate the enormous mortality involved, a male and female fish might produce 100,000 eggs. For the population to replenish itself, neither grow nor decline, two eggs need to survive to adulthood. There are two issues here. The first is conservation, for clearly sufficient adult fish must be permitted to survive to breed to allow population renewal. The second is variability, for it is hard to conceive of a natural process that can provide such control: if on average fewer than two individuals survive, the population is failing to replenish itself; if more than two survive, it will increase. Unsurprisingly, observations on fish recruitment, as it is termed, show high variation. Nevertheless, as more and more fish stocks were heavily fished in the 1980s and 1990s, it became clear that fishing frequently did reduce the number of spawning fish sufficiently to reduce the resultant recruitment (Hilborn and Walters, 1992). Fisheries management agencies now routinely set lower limits on spawning stock and design their harvest strategies to keep stocks above these levels.

Management

The management implications of this analysis are important. If fishing starts on age groups that are not yet mature, then fishing intensity must be controlled to ensure sufficient adults survive to breed so that declines in recruitment do not occur. If fishing does not start on age groups until after breeding has occurred, the conservation issues become less intense, although care is still needed. For example, larger, older fish produce more eggs.

An overexploited fish population will thus exhibit characteristic symptoms. First, there will be few age classes in the population and hence the average age/size of the catch will be lower than in less intensively exploited times. Second, the recruitment of young fish will be in decline, reflecting overexploitation of the adult stock. In such a situation, the management action needed is straightforward, a reduction in fishing intensity and an increase in the age of first fishing. However, depending on the severity of overexploitation, fishing may need to cease altogether.

Although the management action needed to address overexploitation is well understood, once severe overexploitation has occurred, there is no guarantee that reducing or ceasing fishing will reverse the process and that the stock will recover. There are plausible mechanisms that can prevent this, of which the most important are likely to be increased competition or predation by other species. Accordingly, the best fisheries management practice is aimed at preventing overexploitation from occurring.

Monitoring

Clearly, a key to managing fish stocks is the ability to monitor what is happening to the stock under exploitation. There are direct methods of surveying and monitoring stocks using research vessels, which can work well in many cases. Most fisheries management agencies now operate routine surveys of fish abundance for the commercially most important stocks. However, in other situations, such methods are either inefficient or economically unjustifiable. Regular monitoring of

information from the catches of the commercial fleets permits a number of different methods to be used. These methods use this statistical information to reconstruct the behavior of the stock. One common method involves returning to the model of eqn [2]. The yield or catch from the stock was given by the equation

$$\text{Catch} = pB \quad [3]$$

As noted earlier, p depends on the level of fishing effort. The definition of fishing effort (F) is dependent on circumstances, hence for a particular definition of fishing effort, for example, hours fished,

$$p = qF \quad [4]$$

where q is a scaling factor (known in the literature as the catchability coefficient)

$$\text{Catch} = qFB \quad [5]$$

Statistics can provide information on catch for known effort and hence catch per unit of effort (CPUE) is proportional to stock biomass, with the constant of proportionality being the catchability coefficient. A variety of statistical methods exist to estimate the catchability coefficient and as a consequence the stock size. Time series of information on catch and effort thus provides direct insight into the way in which the stock has responded to exploitation.

A second widely used method uses information not just on the catch but on the age composition of that catch. This extra information is used in a method known as virtual population analysis (initially developed by Gulland) to reconstruct the changes in the population under exploitation. The method is simple in conception, although computationally quite laborious. In essence, a simple accounting procedure is used: a catch of 1000 eight-year-olds this year implies that at least 1000 seven-year-olds were alive in the previous year. Adjustments for natural deaths and fish catches then permit the changes in population to be estimated.

Fisheries Economics

It is reasonable to enquire why fisheries are in such a poor state when the science of fisheries is reasonably well developed, methods of monitoring stocks exist, and management actions needed to address overexploitation are understood. The answer lies primarily in the domain of social science.

Fisheries as an economic activity have one fundamental difference from other industrial activities. In manufacturing industry, increases in supply develop in response to demand by increasing inputs of labor, raw materials, capital, etc. In fisheries (and in other industries that exploit renewable resources) there is a limit to supply, which is determined by the maximum sustainable yield. Increasing the inputs to a fishery – more fishermen, better gear, more vessels, etc. – actually reduces the yield if the stock is already being exploited at or above the level of effort that produces the maximum sustainable yield. In this situation, increases in demand reflected in higher prices can lead to increases in fishing effort, lower catches, and hence higher prices. Depending on the situation, vessels may still be profitable or covering costs even though the effort being expended is well above that required to take the catch and the price paid by consumers well above that, which would prevail if efforts were lower and the catch higher. In a situation where there is no control of the fishery, this process can continue until resources are severely overexploited and individuals in the industry barely cover their operating costs. The issue is one of ownership; many fisheries are still (and most were) common property resources. In such cases, access is unrestricted and additional effort enters the fishery as long as the return is better than in other activities. Because fishing is a specialized activity (fishing gear and vessels have few alternative uses), other alternative activities offer very low returns. Hence most fisheries, without management tend to be overcapitalized and economic returns are poor.

Most government actions in the period 1970–2000 have been the reverse of what would be economically efficient. Fishing operators have been subsidized and have stayed with the industry rather than leave, thus accentuating the overexploitation problem. Clearly, social issues play an important role here; it is obviously problematic to let fishing communities face economic collapse, but the scale of subsidy involved is substantial (*see* Economic Issues, above). There have been moves by governments to reduce fishing fleets, either by direct purchase of vessels or by providing incentives to fishermen to use fewer vessels. Fleet sizes in most developing countries declined during the 2000s.

Fishery Management in Practice

The management of fisheries involves some simple principles based on sustainable yield ideas yet its practice is often fraught with difficulty. Management measures typically control fishing by regulations that impose limits on catch (quotas, or total allowable catches (TACs)) or limits on access or effort (licensing, closed areas, and seasons) or by indirect controls such as minimum size of catch or restrictions on types of gear.

It is fair to say that such methods have had mixed success, for a variety of reasons. Control and policing of regulations are difficult and can be extremely costly. Some regulations, TACs in particular, provide an economic motivation for misreporting of catches and hence undermine the scientific basis of the process. Political issues, particularly when there is significant overcapacity, can override fisheries management concerns. Often in the face of uncertainty, regulations are set with an optimistic view of the stock, which subsequent scientific information reveals to be unfounded. Even when regulation has succeeded in its conservation goals, the fishing industry remains overcapitalized and continues to need assistance. By 2010 success was more evident as a number of prominent stocks were rebuilding in the US and Europe. In 2011 the former chief fishery scientist for the US government announced that overfishing in the US had ended.

One development where there appears ground for some optimism is the construction by certain states of regimes based on the allocation of property rights to operators: Australia, Iceland, and New Zealand have been particularly active. In these regimes, operators are allocated the rights to a particular level of catch (or proportion of some fixed catch). These rights,

known as Individual Transferable Quotas, can be bought and sold. In this way the inefficient operations can be bought out by the more efficient ones, which can adjust their capacity to the appropriate level for exploiting the fish stock. The attraction of such a regime is that it is in the interests of the operators (owners) to ensure that the stock is exploited sustainably and that the capacity of the fishing fleets matches the production of the resource. Allocation of these individual quotas is highly controversial in some countries because they may lead to concentration of quota ownership in a few companies, and those granted with the initial quotas often become very wealthy on what many view as a public resource. Other methods of allocating exclusive access to fishing based on communities or fishing sectors are being explored as more socially equitable ways of stopping the competitive nature of fisheries.

New International Measures

Although most states now operate 200 mile zones, in which they control fishing, in certain areas the continental shelf extends beyond 200 miles and exploitation of the same stock can be possible both within the 200 mile zone and beyond it on the high seas. The UN Agreement on Straddling Stocks, offers a new international framework for addressing problems of this sort. Its aim is to facilitate the creation of regional bodies composed of both coastal states and those with distant-water fleets operating in the area. It is too early to assess whether this will be successful.

Certain types of species, tuna in particular, are highly migratory and are exploited over very large areas of ocean. Management regimes exist under International Treaty to regulate these fisheries. There are now international organizations for the Atlantic, Eastern Pacific, Western Pacific, and Indian Ocean tuna stocks in addition to a specific organization for southern bluefin tuna. The status of bluefin tuna is of particular concern, and there is a long history of overcatching of agreed quotas and severe depletion of bluefin stocks. In 2010, Atlantic bluefin was proposed to the Convention on International Trade in Endangered Species to have all international trade stopped. The proposal was rejected but was indicative of international concern about the inability of the member countries to set scientifically based quotas and enforce them.

Ecosystem Management

Fish stocks exist in aquatic ecosystems, forming the prey of some species and being predators of others, yet most fisheries science and almost all fisheries management regimes tend to focus on individual species. The regime that was set up to deal with the Southern Ocean, the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR), had a different focus. This regime was specifically orientated to the Southern Ocean ecosystem and regulations were aimed at ensuring that each element of the ecosystem was protected from the effects of exploitation. The primary reason for this markedly different approach was that the large baleen whales of the Southern Ocean had been grossly overexploited by distant-water fleets this century. A significant development of a fishery for their common prey, krill, could hinder, or indeed stop the recovery of the populations following the cessation of whaling. To date, CCAMLR has been reasonably successful in its operation, setting precautionary levels of catch for krill as well as regulating other fisheries.

However, better scientific understanding of the implications of exploiting groups of interacting species is needed. Current understanding is really little more than educated commonsense – for example, a predator cannot be exploited at its MSY level if its prey is simultaneously harvested at a high level.

The Future of Commercial Fisheries

The increase in the world population over the next two decades will clearly produce an increase in basic demand for food. This is likely to be enhanced in the case of fish as it is a preferred protein source and increasing wealth tends to be associated with increasing fish consumption. Additionally, much of the population increase is estimated to be occurring in Asia, where fish is a preferred source of animal protein.

Statistical analyses of trends in capture fisheries point to a maximum harvest level in the region of 95 million tonnes, a value close to recent average landings. However, more recent analysis by FAO indicates that there is a potential for increasing these catch levels if the full potential of underexploited areas and species was to be realized. Similarly, there is scope, as discussed earlier, for significant improvement in the reduction of unwanted bycatch.

Improved fishery management could, in principle, add significantly to the global production – recall the high proportion of overexploited stocks. However, the extent to which species interactions may mean that the recovery of one stock may lead to lower catches from another is not well understood.

FAO has concluded that a substantial increase in catch levels of around 20 million tonnes could be achieved if (1) degraded resources are rehabilitated, (2) underexploited resources are successfully managed to increase yields, (3) fully exploited resources are not degraded, and (4) discarding and wastage are reduced. This is the challenge for fisheries science and management in the new century.

See also: Aquaculture. Fishes, Biodiversity of. Fish Conservation. Fish Stocks

References

Alverson DL, Freeberg MH, Pope JG, and Murawski SA (1994) *A Global Assessment of Fisheries Bycatch and Discards*. FAO Fisheries Technical Paper 339.

Anon (2009) *State of World Fisheries and Aquaculture 2008*. FAO Fisheries Technical Paper No. 457.

Beverton RJH and Holt SJ (1957) *On the Dynamics of Exploited Fish Populations* vol. 2. London, UK: Her Majesty's Stationery Office.

Garcia SM and Newton C (1997) Current situation, trends and prospects in world capture fisheries. In: Pikitch E, Huppert DD, and Sissenwine M (eds.) *Global Trends in Fisheries Management*, pp. 3–27. Bethesda, MD: American Fisheries Society Symposium 20.

Hilborn R and Walters CJ (1992) *Quantitative Fisheries Stock Assessment. Choice Dynamics and Uncertainty*. New York: Chapman & Hall.

Hilborn R, Branch TA, Ernst B, *et al*. (2003) State of the world's fisheries. *Annual Review of Environment and Resources* 28: 359–399.

Klein Goldewijk K and van Drecht G (2006) "HYDE 3.0: Current and historical population and land cover," In: Bouwman AF, Kram T, and Goldewijk KK (eds.) *"Integrated modelling of global environmental change. An overview of IMAGE 2.4," Netherlands Environmental Assessment Agency (MNP)*. The Netherlands: Bilthoven.

Milazzo M (1998) *Subsidies in World Fisheries, a Reexamination*. Washington, DC: World Bank Technical Paper No. 406.

World Bank (2009) *The Sunken Billions: Economic Justification for Fisheries Reform*. Washington, DC: The World Bank.

Resource Partitioning

FA Bazzaz and S Catovsky, Harvard University, MA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 173–184, © 2001, Elsevier Inc.

Glossary

Coexistence Ability of species to persist together in an ecological community indefinitely in the absence of any major environmental change.
Ecological community An assemblage of species occurring together in a given space and time.
Heterogeneity Variation in environmental conditions through space (spatial) or time (temporal).
Niche Pattern of response of an individual, a population, or a species to the physical and biological gradients of its environment.
Niche differentiation Differential resource use or response that results from long-term, consistent reciprocal selection.
Reciprocal selection Evolutionary change as a result of species acting as selective agents on each other through interspecific interactions.
Resource A consumable or depletable substance, such as a nutrient, water, or light, that is required by plants for maintenance, growth, and reproduction.
Trade-off Constraint on expression of a species' trait/behavior as a result of expression of a correlated trait/behavior.

Introduction

The scientific community and the public at large are expressing great concerns about the global loss of biological diversity. Currently, biological diversity is threatened worldwide by the widespread destruction of forests and by impacts of global environmental change, such as nitrogen deposition and global warming. In addition, ease of travel, the dramatic increase in the number of people who travel, and deliberate introductions of organisms into various habitats can all lead to the modification of ecosystems and the loss of biological diversity. Invasions are now rampant and may lead to a great deal of homogenization of the Earth's biota. This situation has led the renowned biologist Gordon Orians to call this epoch the "Homogecene."

If we are to understand and predict the consequences of loss of species diversity for the functioning of the biosphere, we need to ascertain what factors control patterns of species richness both within and across ecological communities. The notion of community diversity has its roots in the writings of Henry Gleason and Frederic Clements. These American ecologists differed sharply in their interpretation of the plant community. The Clementsian conception of the community assumed it to be a cohesive active unit. The notions of reciprocal selection and niche differentiation are both imbedded in this idea, although not explicitly. In contrast, Gleason thought of the community as a chance assemblage of species with overlapping physiological requirements. This view may assume that there is no interaction between species or that these interactions are incidental. More recent discussion of why there are so many species was rekindled by G. E. Hutchinson in 1959 when he asked this question in his now very famous "Homage to Santa Rosalia" paper.

Currently, there are many hypotheses about the generation and maintenance of plant diversity (see discussions in both Tilman and Pacala (1993) and Huston (1994). Much like the Clements–Gleason debate, discussion of the maintenance of species diversity often focuses on the distinction between theories that assume evolution has driven species to partition resources so as to prevent competitive exclusion (equilibrium) and those that explain diversity through the action of demographic dynamics and do not assume a past evolutionary history between species (nonequilbrium). Ecologists commonly favor one mechanism for maintenance of diversity over others, but it is more likely that the strength of any cause differs under different locations and circumstances. It is now critical that we ascertain what characteristics of species and ecosystems lead to the dominance of any one diversity mechanism, so that we can develop a predictive framework for determining causes for maintenance of species diversity across all ecological communities.

To move the science of biodiversity from a descriptive to a predictive discipline, we need to ascertain what mechanisms underlie global patterns of species diversity. This requires clearly defining what we mean by diversity and recognizing its importance (*see* Assessment of Diversity and its Functional Significance). We must then determine what biological processes lead to the development of diversity within ecological communities. In this paper, we take a predominantly resource-based approach to understanding the causes of biological diversity. We highlight why resources play a central role in the maintenance of species diversity (*see* Resources and Niche Differentiation: Theoretical Considerations and Resources *and* Resources and Niche Differentiation: Empirical Evidence) and then discuss how this resource-focused view relates to other hypotheses that have been proposed to explain global patterns of species diversity (*see* Additional Mechanisms for Maintenance of Species Diversity).

Assessment of Diversity and its Functional Significance

Presently, there is no agreed quantitative way to assess diversity, except to use the taxonomic species concept. However, the

use of species as the primary unit of biological diversity is not always appropriate for the situation. We have previously argued that we should focus on different components of ecological diversity depending on the patterns and processes being studied. Indeed, there is now an increasing awareness that biological diversity should encompass the full range of organizational units in ecology, from genes to ecosystems. If we wish to incorporate this broader definition of diversity into our study and stewardship of ecological communities, we must present a clear framework for the accurate assessment of diversity. Developing such a framework involves explicitly outlining the different kinds of biological diversity that should be recognized and how the level of diversity can be quantified.

In many cases, the confusion arises from not recognizing the appropriate scale for the ecological phenomenon of interest. The need for clarification of scale was first recognized by Robert H. Whittaker in 1972, when he addressed the importance of distinguishing within-and between-patch diversity (known as α-diversity and β-diversity, respectively). The temporal scale of ecological dynamics may be as relevant to our discussion of biodiversity assessment as the spatial scale. This notion is well illustrated by the need in certain systems to distinguish between apparent diversity and total diversity, which may be quite different from one other. Many ecological communities depend on an intermittent availability of resources. For example, in deserts, when water occasionally becomes available in large quantities, the number of plant species (and perhaps microbes and the animals that depend on these plants) dramatically increases. In such systems, apparent diversity, which usually involves the enumeration of the species present at a given time, can vastly underestimate total diversity. Other kinds of diversity measurement may be appropriate, depending on the ultimate goals of the study. For example, aesthetic diversity for humans may depend on the number of flower colors present in a community rather than species diversity assessed by a professional (**Figure 1**).

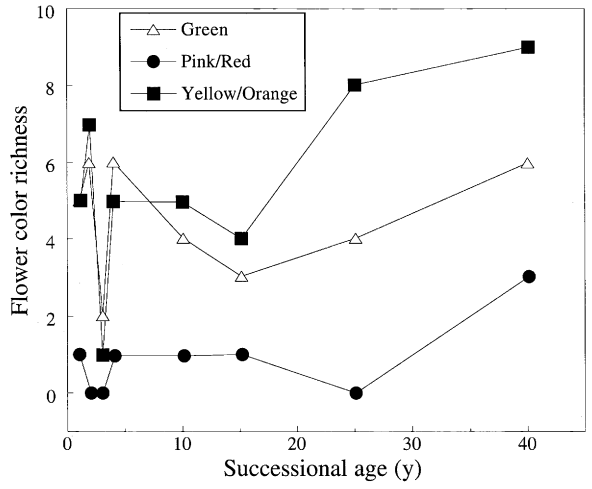


Figure 1 Changes in floral color diversity in midwestern old-field successions. Reproduced from Bazzaz FA (1996) *Plants in Changing Environments: Linking Physiological, Population, and Community Ecology.* Cambridge, UK: Cambridge Univ. Press, with permission from CUP.

Specifying what components of ecological communities should be measured has become increasingly important since expanding our definition of biodiversity. As we begin to recognize the central role that ecosystems play in the global cycling of elements, functional diversity has been identified as an important characteristic in ecological assessment. Functional diversity measures the range of energy and matter processing that a particular community can undertake, e.g., carbon uptake and nitrogen fixation. Explicit in this concept is the idea of functional redundancy, which states that often more than one species in a community can play a role in a particular ecosystem process (see chapters in Schulze and Mooney, 1993). Functional diversity may be very important in the discussion of biological diversity because the removal of a given species may be much more critical to a particular system function than the removal of another species, even one with more biomass.

Although plant communities are only theoretical constructs (see Bazzaz, 1996), much discussion has gone into the notion of biological diversity and its impact on the functioning of ecosystems. A number of recent, high-profile experiments have addressed how many species must be present for a given ecosystem to function appropriately, e.g., the Ecotron experiment (Shahid Naeem and others), field studies on California serpentine grassland communities (David Hooper and Peter Vitousek) and on midwestern tallgrass prairie communities (David Tilman and others), and the BIO-DEPTH experiment carried out at eight grassland sites across Europe. Current data in these grassland ecosystems seem to suggest that the number of species for greatest ecosystem productivity hovers close to nine. Unfortunately, there is still considerable uncertainty about what mechanisms are driving these diversity–productivity patterns (see discussion in recent BIO-DEPTH paper: Hector *et al.*, 1999). Some ecologists currently believe that complementarity in species' resource use could make a multispecies community more effective at utilizing available resources (to produce higher net productivity), while others argue that more diverse communities are merely more likely to contain more productive species (known as the sampling effect). As we currently lack a complete mechanistic understanding of the relationship between species diversity and ecosystem function and as relevant data exist for few other ecosystem types, we cannot yet know how general these results will turn out to be. In these particular experiments, the species are all somewhat similar in their behavior and no "keystone" species exists. In communities containing a species that performs a special function (such as nitrogen fixation), the removal or addition of such a species may severely impact ecosystem function, e.g., *Myrica faya* invasion in Hawaii. We must also pay close attention to species that supply unique goods and services to humans.

Resources and Niche Differentiation: Theoretical Considerations

Maintenance of species diversity is often thought to be a result of niche differentiation. This idea was implicit in much early writing on community structure that recognized that species could not coexist with one another unless they had sufficiently different patterns of resource use. As plants are predominantly sessile organisms, it is assumed that plant species that are

components of a community have coexisted and competed for a long time. In this way, they are assumed to have shaped each other's responses to the multitude of critical physical and biological resources in their local environment through reciprocal selection, which could result in niche differentiation in the long term (Connell, 1980). This idea is similar to character displacement in animals. However, requirements for niche differentiation are stringent and may not always be met. Long-term species co-occurrence is needed and the frequency of encounter between plant species will depend on dispersal patterns (*see* Additional Mechanisms for Maintenance of Species Diversity: Spatial Structure). In addition, reciprocal shaping of species' responses to their environment is not easy to accomplish, unless initially traits related to fitness are somewhat independent of each other. To avoid assuming that niche differentiation has taken place in the past, Bazzaz (1996) suggested previously that it is more tractable to consider copresence rather than coexistence, especially in communities with a poorly known history. Studies on the postglacial migrations of temperate tree species during the Holocene in North America have clearly demonstrated how species have moved individualistically across the landscape and have not necessarily coexisted in the same community for long periods of time (see papers by Margaret Davis). Here, we consider communities where species are found together without regard to whether or not they co-occurred over evolutionary time and therefore have shaped each other's responses to the environment.

While niche differentiation has been touted as a major cause of the generation and maintenance of biodiversity, particularly in late successional stable ecosystems, it was unclear for many years how the concept applied to plant communities. The notion of resource partitioning was originally developed for animal communities, where species can potentially utilize a wide range of food resources and thus can easily partition up their resource environment. This idea was harder to apply to plant communities, as plants all have relatively similar resource requirements (Connell, 1978). As a result, much ecological research has focused on trying to determine how species partition themselves along resource axes. Progress on understanding factors controlling plant community diversity has been achieved by considering species' responses to multiple resources in the environment. David Tilman (1982) tackled this issue by using a mathematical model to show that species coexistence was feasible if there were trade-offs in species' requirements for different ratios of resources. This approach clearly demonstrated the potential for a resource-based approach to contribute to explaining plant species diversity. Often, however, species' responses to environmental factors are not independent of each other. Thus, it may be more informative to develop a multidimensional response surface that approximates the response of the species to its environment (**Figure 2**) (Bazzaz, 1996). Gradient techniques from vector calculus now allow us to quantify these response surfaces.

In this multidimensional view of the plant niche, maintenance of diversity depends on species separating themselves along multiple resource axes, such that they can coexist for a long time. It is assumed that there must be a minimal overlap between members of a community. In 1973, Sir Robert May

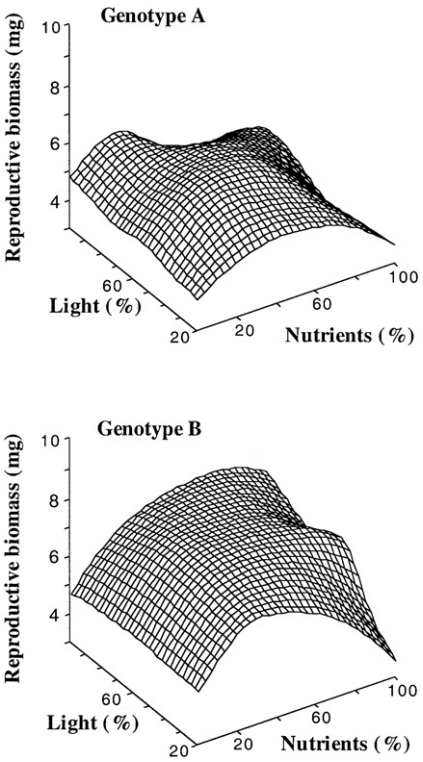


Figure 2 Differential responses of two genotypes of *Polygonum* to varying light and nutrient gradients. Reproduced from Bazzaz FA (1996) *Plants in Changing Environments: Linking Physiological, Population, and Community Ecology.* Cambridge, UK: Cambridge Univ. Press, with permission from CUP.

attempted to quantify the minimum allowable overlap of the species for their stable coexistence and concluded that, if the difference between the means of their responses is more than twice that of the spread of each individual curve, stable coexistence should be possible. If species are highly specialized (narrow response curves), they do not need to position themselves very far apart on one particular resource axis to prevent competitive exclusion, whereas species that are more generalist in their resource use (wide curves) must be further separated in niche space. This kind of analysis has to be extended to multidimensional response surfaces. We still do not know, even theoretically, what the minimum overlap to allow coexistence should be when we consider species' responses to more than one critical environmental gradient.

These minimum overlap conditions for coexistence may differ for contrasting community types. For example, it is reasonably well established that plants of early successional communities have broader responses to resource gradients than do plants of late successional communities. This pattern alters the length and intensity with which species interact with one another, and thus the degree to which species partition the environment. However, it has not been determined whether this pattern results from or leads to the lower species diversity of early successional communities relative to late successional ones. Early successional plants have a near-geometric type of distribution (each species takes approximately the same proportion of resources to produce the same amount of biomass). As succession proceeds, there is a shift toward a

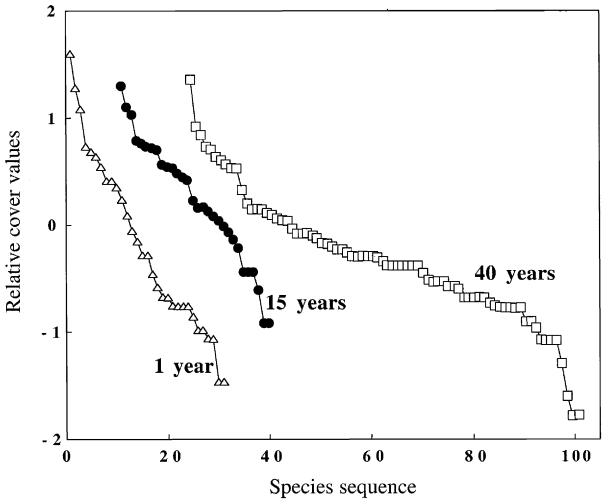


Figure 3 Changes in community dominance–diversity relationships during 40 years of old-field succession. Modified from Bazzaz (1975) *Ecology* 56, 485–488, with permission from ESA.

more log-normal distribution (**Figure 3**) and then finally, without major disturbance, a return to a near-geometric distribution. This variation in patterns of resource use between species of different successional position creates a hump-backed relationship between diversity and successional time, as predicted by the intermediate disturbance hypothesis (Connell, 1978).

We have seen that mathematical models of community structure predict that diversity is largely a function of how species partition themselves along multiple environmental resource axes. Theoretical considerations suggest that trade-offs in plant species' responses along multiple environmental gradients could permit multi-species coexistence. But, do we see such trade-offs in nature and are they sufficient to explain long-term species coexistence? In the next section, we discuss the evidence for niche differentiation in plant communities and consider whether or not the degree of resource partitioning in these communities is adequate to explain species diversity in different plant communities. Resources may be partitioning both spatially and temporally and we present evidence for both kinds of niche differentiation among plant species.

Resources and Niche Differentiation: Empirical Evidence

Spatial Heterogeneity

Most resource partitioning arguments proposed to explain species diversity assume that the environment is heterogeneous. Coexistence is possible because species have different resource requirements and are specialized to succeed on particular patch types. Diversity is maintained through the presence of an array of patch types. As our ability to measure and quantify environmental variation has improved, we have developed clear evidence that the environment is heterogeneous with regard to resources that are critical to plants. To date, however, we have not critically assessed if the degree of heterogeneity recorded by our instruments is representative of the resource environment perceived by the plants themselves (Bazzaz, 1996). Despite this uncertainty, there is now clear evidence that, at least on the broad scale, species do partition themselves along resource axes. Studies that relate species' distributions to multiple environmental factors using either direct or indirect gradient analysis have identified numerous soil-related axes of specialization, e.g., R. H. Whittaker's 1956 work on the vegetation of the Great Smoky Mountains. Research on experimentally produced plant communities also supports the notion that species may be differentiated along soil resource axes (see studies discussed in Bazzaz, 1996).

Although plants themselves can influence their local environment and thus produce intrinsic heterogeneity within the community, much spatial resource heterogeneity is generated by extrinsic factors, the most common of which is disturbance. In fact, disturbance is judged both by change in availability of resources and by perception of that change. Disturbance events create heterogeneity in resource availability at multiple spatial and temporal scales, and these different scales can allow species to partition the environment in more ways. Awareness of the importance of disturbance events for understanding the maintenance of plant species diversity was heightened when Peter Grubb (1977) proposed that regeneration traits of species should provide a critical axis for differentiation, the so-called regeneration niche. Much of the research on partitioning along this regeneration axis has focused on the broad-scale effects of canopy disturbance, addressing differences in conditions and plant responses between gaps and closed forest. Gap creation generates variability in the light environment and has been advanced as a major factor in increasing diversity at the stand level. Broad guilds of plant species that respond similarly to gap formation (light demanding vs shade tolerant) have been identified in numerous forest communities where vertical canopy stratification creates a wide range of light microenvironments (Whitmore, 1989). More recently, sophisticated mathematical techniques have been used to classify both physiological and demographic responses of tree seedlings to light availability more precisely (see SORTIE papers of Steve Pacala, Richard Kobe, and others). Much of this research into species' responses to total light quantity has found considerable overlap in response, and species' partitioning of the light microenvironment is not usually sufficient to explain maintenance of species diversity, particular in communities with high numbers of co-occurring species, such as lowland tropical rain forest.

An additional mechanism for the maintenance of diversity arose from the gap partitioning hypothesis, originally developed by Ricklefs (1977) and applied specifically to the tropical rain forest by Julie Denslow (1980). Here, finer-scale variation in resource availability across the gap–understory continuum was suggested as a possible mechanism for coexistence in multispecies communities. Unfortunately, this hypothesis has not yet been validated in tropical rain forests, where we lack a good understanding of the maintenance of species diversity. Experimental work by Tim Sipe and F. A. Bazzaz in temperate forests in eastern North America, in contrast, provides some support for the gap partitioning hypothesis. In experimentally created gaps at Harvard Forest in central Massachusetts, we found clear micro-environmental

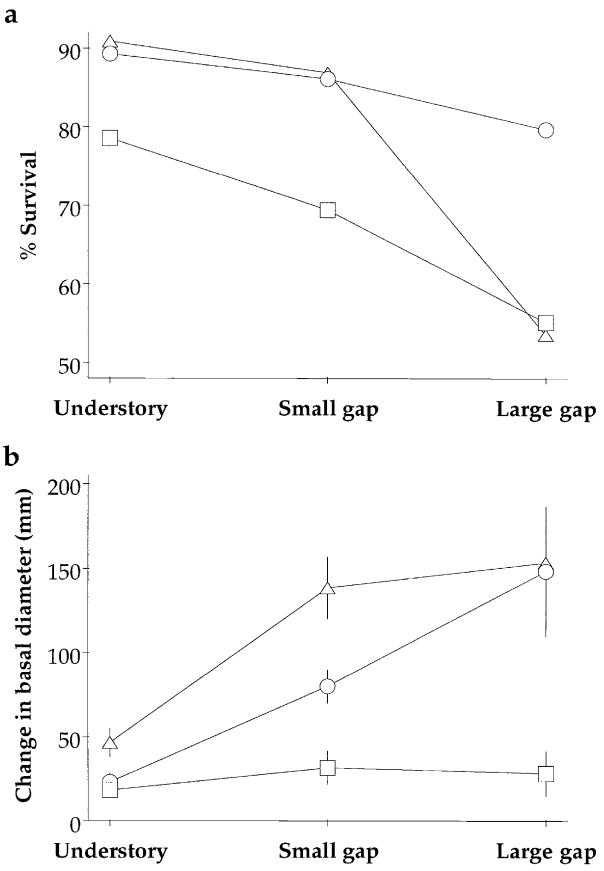


Figure 4 Survival (a) and growth (b) responses of striped maple (*Acer pensylvanicum*, triangles), red maple (*Acer rubrum*, circles), and sugar maple (*Acer saccharum*, squares) seedlings to canopy gap size. Data redrawn from Sipe and Bazzaz (1994) *Ecology* 75, 2318–2332, and Sipe and Bazzaz (1995) *Ecology* 76, 1587–1602.

differences between gaps of different sizes (especially with regard to light). We compared morphological and physiological responses of three co-occurring maple (*Acer*) species to this light variation and found clear evidence that the species differed in their response to gap size. Many response variables showed significant differences between large gaps, small gaps, and understory plots, and these differences often varied between species, creating distinct species' preferences for canopy gap environment (**Figure 4**). We also observed substantial variation in light availability between different parts of canopy gaps, due to seasonal and diurnal trends in solar patterns. This variation, however, was rarely reflected in differences in species' responses to position within the gap.

Temporal Heterogeneity

As resources vary in both space and time, species could partition their resource environment temporally as well as spatially. Temporal heterogeneity in resource availability is often not considered in discussions of maintenance of diversity, but could prove central to our understanding of species coexistence. Differences in the timing of plant developmental events, such as germination, flowering, and fruit drop, could play a critical role in mechanisms of species coexistence. In some communities that have existed for a long time, there is evidence for clear separation in flowering between groups, e.g., Peter Ashton's work on staggered mast-flowering in the Dipterocarpaceae in the tropical rain forests of Southeast Asia. In the tallgrass prairie of the American Midwest (communities that have a long evolutionary history and great species diversity), we have demonstrated clearly how these species separate their life history events along the entire growing season, i.e., phenological separation. One can recognize three fairly distinct groups of flowering species (**Figure 5**). Also, in these grasslands one can observe the separation of the phenology of the introduced species, *Poa pratensis*. Although it is now considered by many investigators to be part of the tallgrass prairie, it is an introduced species (from European grasslands) that grows well in the cool portion of the season. Thus, it is able to grow at two distinct times of the year: before and after the main growing season of its competitors, the native, warm-season grasses of the community (**Figure 6**). In this way, this species separates its growth activity from that of the native species and reduces competition for critical resources.

This kind of phenological separation is also evident in some early successional communities where species generally have broad niches (*see earlier discussion in* Resources and Niche Differentiation: Empirical Evidence: Spatial Heterogeneity). In old-field communities in the American Midwest, *Chenopodium album* delays its growth and reproduction much later than all other species in the community (Bazzaz, 1996), thereby reducing competition during the main portion of the growing season. In this case, it is not clear that competition has shaped these phenological responses, as species in this early successional community have not necessarily been present together over evolutionary time. Many arguments about coexistence assume that competition has shaped the niches of co-occurring species in the past (see earlier discussion). However, in communities of unknown history, the question remains whether a species has modified its phenology by competition with the native species or was preadapted to function in these communities. An example of such preadaptation is found in the shifted daily flowering behavior of the native *Erigeron annuus* and the introduced *Lactuca scariola*. In winter annual communities of the midwestern United States, *Lactuca* opens its flowers in the morning hours 7:00 to 10:00 a.m. and requires a large number of pollinators. After 10:00 a.m., however, *Erigeron* starts to open and *Lactuca* closes its flowers (**Figure 7**). As a result, the pollinators shift their foraging to *Erigeron*. Thus, while both species use the same pollinators, the daily separation in the opening of the flowers ensures that both are pollinated. At first glance, this may appear to be a convincing example of niche differentiation. However, when examined carefully, one finds that the introduced *Lactuca* developed this flowering behavior in its native habitat in Europe and not by competition for pollinators in the successional fields of the Midwest.

Temporal differences in species' traits may extend to very early stages of the life cycle. The formation of seed banks that can persist in the soil for a long time may also lead to maintenance of species diversity. A persistent seed bank may buffer species' responses to short-term variation in environmental

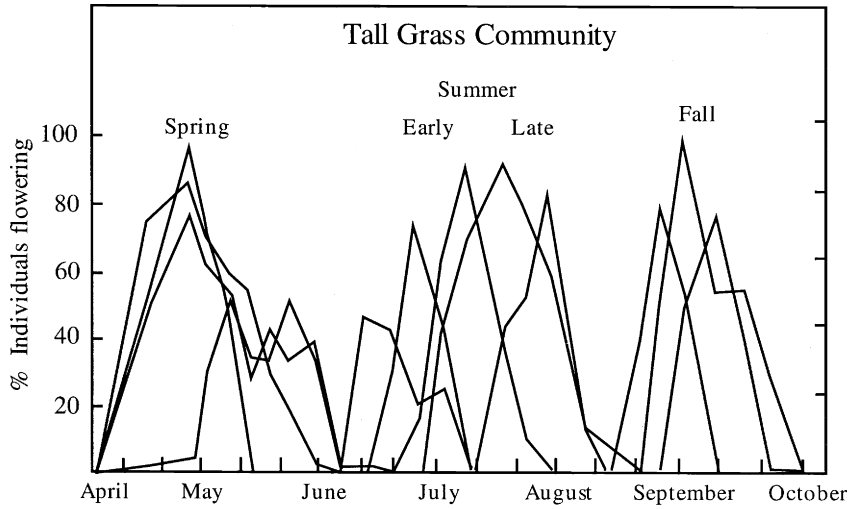


Figure 5 Seasonal flowering patterns of a midwestern prairie community with the major species shown as separate lines. Redrawn from Parrish and Bazzaz (1979) *Ecology* 60, 597–610, with permission from ESA.

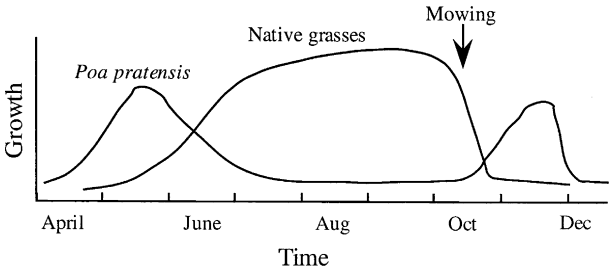


Figure 6 Separation of growth phenology of the introduced grass, *Poa pratensis* (bluegrass), from that of the predominant native prairie grasslands. Redrawn from Bazzaz and Parrish (1982) in *Grasses and grasslands*.

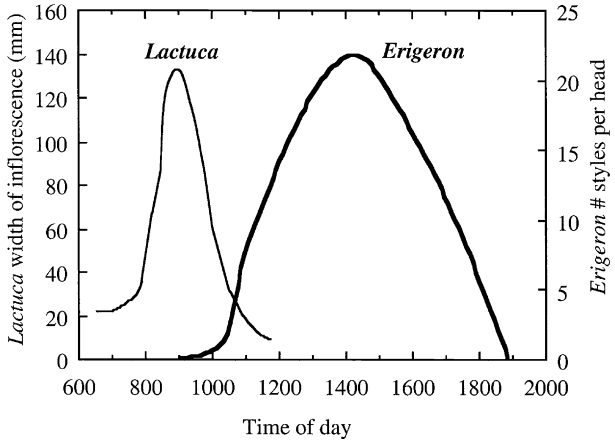


Figure 7 Importance of the daily timing of flowering for two members (*Erigeron annuus* and *Lactuca scariola*) of a midwestern winter annual plant community. Redrawn from Parrish and Bazzaz (1978) *Oecologia* (*Berl.*) 35, 133–140, with permission from Springer.

conditions and prevent irreversible population crashes. Weedy species are known to form persistent seed banks. In good years for a particular species, many seeds are produced and enter the soil seed bank. In bad years, in contrast, few seeds are produced. The species can, therefore, partition the time axis to contribute the most to future generations in years when conditions are favorable for growth and reproduction. The development of this sort of persistent seed bank has been mathematically modeled and is often called the storage effect (see work by Peter Chesson). Coexistence is promoted if species have differing sensitivities to temporal variation in environmental conditions.

Further niche axes may potentially be created through the combination of temporal heterogeneity in multiple resources (*see discussion in* Resources and Niche Differentiation: Theoretical Considerations). Because plant responses to a particular resource are contingent on prevailing local environmental conditions, e.g., temperature and relative humidity, the time course of resource availability patterns could critically determine a plant's performance. If resource availability is high when other conditions are favorable (i.e., resources are congruent), then plants may be able to make good use of that resource. If, however, conditions are unfavorable and plant activity is inhibited to some extent, then plants may not be able to take full advantage of the resource in ample supply. Species might have different levels of tolerance for the extent of resource congruency in a given plant community (Bazzaz, 1996). Additional research by our laboratory at Harvard Forest has demonstrated that, for seedlings in a canopy gap where environmental conditions show distinct temporal patterns, resource congruency could play an important role in permitting species coexistence (see papers by Peter Wayne and Gary Carlton).

Resource Partitioning: Is This Sufficient to Understand Species Diversity?

There is now good empirical evidence for substantial resource partitioning between species in the same community.

Underlying variation in soil resources and disturbance-mediated changes in the aboveground microenvironment provide primary axes for species' specialization. Further axes may be produced by responses to multiple environmental gradients, e.g., resource congruency, and by substantial temporal separation of developmental events, such as germination and flowering. It is not clear, however, that these differences are adequate to prevent competitive exclusion in the long term, as the theoretical considerations for community-level consequences of niche differentiation in multidimensional space have not been adequately explored (*see* Resources and Niche Differentiation: Theoretical Considerations). In transient, early successional communities, this discussion may not be relevant, but in communities with a long-term evolutionary history, we certainly need to consider whether resource partitioning plays a central role in community diversity. Communities characterized by long-term species' interactions, such as the tropical rain forest and the tallgrass prairie, typically contain a high number of species, and so it is critical that we unequivocally establish the role of niche differentiation in maintenance of species diversity. To date, experimental evidence suggests that resource partitioning is not sufficient to fully explain the high species richness of some such communities, e.g., tropical rain forest. Work by Hubbell *et al.* (1999) on Barro Colorado Island in Panama demonstrated that gaps did not explain variation in species richness. Gap formation did increase seedling establishment, but this effect was nonspecific and unpredictable.

Additional Mechanisms for Maintenance of Species Diversity

Resource partitioning is clearly not the only diversity mechanism operating in plant communities containing a large number of species. species' specialization on different portions of multidimensional resource space prevents competitive exclusion by reducing the degree of interspecific interactions. However, other processes within plant communities might also reduce the extent or slow the rate of competitive exclusion. In this section, we discuss three additional mechanisms that have been suggested as causes of maintaining plant diversity.

Frequency-Dependent Effects

Mathematical models of competition clearly demonstrate that stable coexistence is possible if intraspecific interactions are stronger than interspecific ones. Weak interspecific interactions commonly arise from reciprocal selection acting over evolutionary time to produce niche differentiation. However, other intraspecific density-dependent effects could act to effectively reduce interspecific interactions. A mechanism for such density-dependent effects in species-rich plant communities like tropical rain forest was first suggested by Joseph Connell and Daniel Janzen. They both argued that the most common and dominant species in the community would be most commonly attacked by herbivores and pathogens, resulting in community-level compensatory mortality. In this way, no single species would be able to dominate the community and reduce overall species diversity. Host-specific herbivores or pathogens could act in a similar compensatory manner, increasing the average distance between parent plants and effective offspring recruitment. Currently, however, empirical support for frequency-dependent mortality in tropical rain forests has been variable. Broad-scale examination of size-specific population dynamics of tropical trees has demonstrated negative density-dependent effects in some cases, e.g., the work of Cam Webb and David Peart in Borneo, but not others (Hubbell and Foster, 1986).

Chance

The strongest alternative theory to the hypothesis that resource partitioning determines community diversity proposes that plant communities are dynamic entities containing a random assemblage of species not organized by reciprocal selection (see Clements vs Gleason dichotomy earlier). According to this view, higher diversity communities result from larger regional species pools rather than from species dividing up the resource environment more finely. Under these so-called lottery model scenarios, coexistence is possible as all individuals have an equal chance of reaching a particular location and whichever individual reaches the spot first will successfully establish there.

Finding little evidence for compensatory mortality in tropical rain forests (see above), Hubbell and Foster (1986) have argued that these lottery models may well represent tropical rain forest dynamics and explain their high species diversity. In these communities, beyond the basic early vs late successional dichotomy, many tropical forest species are essentially ecologically identical. Fluctuations in numbers result simply from chance events, and these fluctuations lead to a random walk in population densities with no stabilizing tendencies at all. Species coexistence is possible because the time for competitive elimination is even longer than the rate of new species formation.

Can we reconcile the niche differentiation vs chance perspectives on tropical rain forest diversity? It might be that the relative importance of deterministic and stochastic processes in community structure depends on the characteristics of the particular community. Chance could play a more important role in community diversity when communities are dynamic and have a fast rate of change, e.g., the tropical forest on Barro Colorado Island. In contrast, in older communities, where the rate of change is much slower and where there is likely to have been considerable sympatric speciation, e.g., the forests of Southeast Asia, niche differentiation could be the more dominant process maintaining species diversity. Until we develop a broad conceptual model of plant communities that incorporates both deterministic and stochastic processes, it will be difficult to test these ideas explicitly. Currently, the relative importance of niche differentiation vs chance effects in structuring plant communities is not well resolved.

Spatial Structure

Implicit in the notion of the regeneration niche is the idea that trade-offs between different components of a species' life cycle

may contribute to species diversity. For plant species, different strategies can be employed during the recruitment and establishment phases of the life cycle, and these strategies could be subject to selection pressure. The importance of life cycle tradeoffs for maintenance of species diversity was effectively demonstrated following awareness of the importance of spatial structure in population and community dynamics. Incorporation of spatial structure into our conceptual models of species diversity reduces the requirements for long-term stable coexistence of species, as dominance of a community by one species is commonly prevented by restrictions on a species' dispersal capabilities. By assuming that there is indeed a trade-off between a species' competitive and colonization ability, i.e., no "superspecies" exists, Tilman (1994) was able to demonstrate that coexistence of an unlimited number of species on a single resource was possible. Although we have much anecdotal evidence that trade-offs in species' competitive and colonization abilities do exist, e.g., during old-field succession (Bazzaz, 1996), we currently do not know if the exact nature of the trade-offs is sufficient to explain multispecies coexistence. It is clear, however, that recruitment limitation is a major factor controlling the structure and diversity of both temperate and tropical forests.

As plants only interact with close neighbors, spatial structure (position of individuals within the community) is a critical consideration for plant diversity as it determines the extent and intensity of intra- and interspecific interactions. Spatial heterogeneity generated by the plant community itself, through differential species' effects on their local environment, could reduce interspecific competition between dominant species if contrasting patch types favor different species. For example, it is now becoming clear that, in many multistory plant communities, the understory herb/shrub layer may play a role in the spatial structure and diversity of the community. In a number of forests worldwide, the understory is dominated by a specific species, such as bamboo in Japan and ferns in low-nutrient forests in the eastern United States. These dense understory layers act as filters for seedlings trying to ascend to the canopy. As the density of the understory is spatially variable and dependent on resource conditions (soil moisture, light availability), the strength of this filtering capability will also be spatially heterogeneous. In experimental work at Harvard Forest, Lisa George and F. A. Bazzaz have found that ferns substantially influence conditions at the forest floor, in terms of both microenvironmental characteristics and the activity of seed/seedling predators. Examination of seedling demography patterns revealed the strong species selectivity of the fern layer (**Figure 8**). Red maple seedlings dominated in fern areas, while the contribution of yellow birch and red oak to seedling bank composition increased in fern-free areas. These fern-induced seedling spatial dynamics have important implications for neighborhood competitive interactions between co-occurring tree species and ultimately for species coexistence.

Spatial structure in forest dynamics may also be produced by the distribution of tree species themselves. Canopy trees can alter light and soil characteristics in the understory and thus influence patterns of seedling regeneration. When a species influences conditions at a site to promote recruitment of conspecifics, this dynamic will set up a stand-level positive

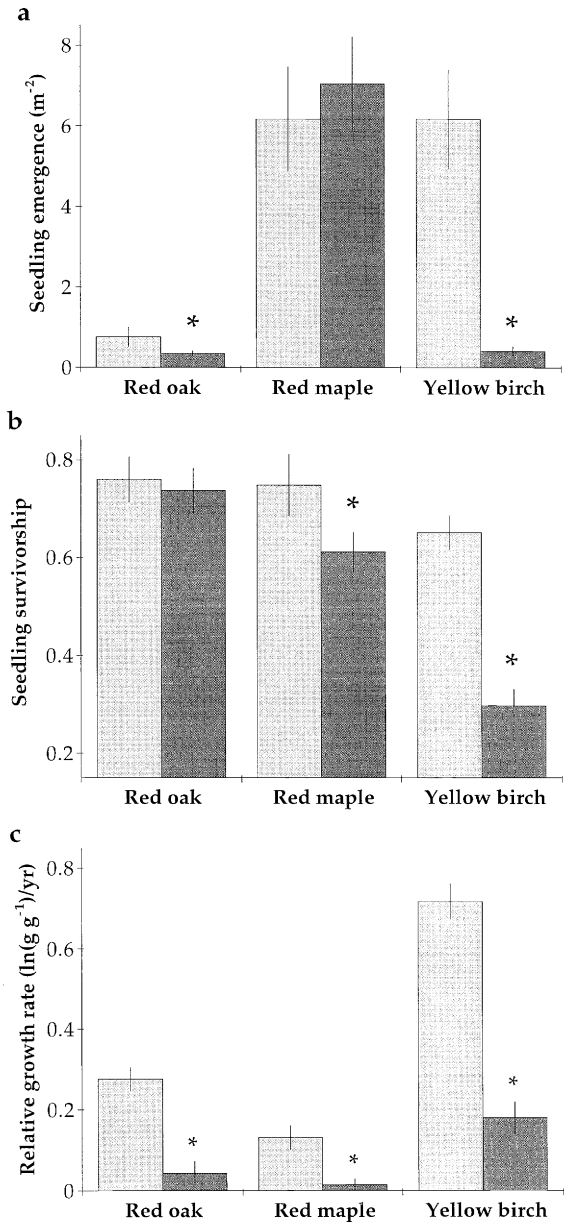


Figure 8 Variation in seedling demography in experimental fern plots: ferns removed (light shading) and ferns present (dark shading). Redrawn from George and Bazzaz (1999) *Ecology*, 80, 833–845, 846–856, with permission from ESA.

feedback with a result that this species may dominate a stand until the next major disturbance. The outcome of positive feedbacks for diversity will depend on the characteristics of the particular community. In relatively species poor temperate forests, these positive feedback effects may promote species coexistence by creating stands dominated by different species and thus reducing interspecific interactions. Sugar maple and hemlock are able to coexist in eastern United States forests as a result of each stand type favoring regeneration of its own species. In contrast, in species-rich tropical forests, these spatially structured positive feedback effects may act to encourage development of single-species stands and thus reduce overall forest species diversity.

Conclusion

Biological diversity and its maintenance remain central to the science of ecology. As we lose more and more species as a result of human activities, the need to document, understand, and evaluate biological diversity becomes increasingly pressing. In this article, we have attempted to clarify our understanding of resource partitioning, as it may influence local or regional diversity. Precisely defining what we mean by diversity is the first step in any scientific endeavor on the subject, and the way that biodiversity is interpreted will depend on the goals of the study. Understanding the functional significance of biological diversity for ecosystem processes has become particularly relevant for assessing the full impact of species loss on the functioning of the biosphere.

Any attempts to determine the underlying causes for global diversity patterns rely on a mechanistic understanding of the processes driving species diversity in ecological communities. We have evaluated the extent to which we can understand species diversity from a resource-based approach. Theory suggests that long-term species coexistence is possible if species' responses to their environment are sufficiently separated in multidimensional resource space. Empirical evidence from a wide range of plant communities supports the notion that species separate themselves along spatial and temporal resource axes. As yet, however, we have not been able to combine this information with theoretical models to establish if the degree of separation is sufficient to favor long-term coexistence and maintenance of high species diversity. Certainly, in communities with very high species diversity such as tropical rain forest, it becomes more challenging to understand coexistence in terms of resource partitioning. Other mechanisms for maintenance of diversity such as frequency-dependent mortality, chance, and spatial structure may play more significant roles in these systems. Now that we have established likely contenders for mechanisms of species coexistence, it is important that we develop a comprehensive, integrative conceptual model that can incorporate these distinct factors into a cohesive framework of biological diversity.

See also: Diversity, Taxonomic versus Functional. Habitat and Niche, Concept of. Species Coexistence. Species Diversity, Overview. Stability, Concept of

References

Bazzaz FA (1996) *Plants in Changing Environments: Linking Physiological, Population, and Community Ecology.* Cambridge, UK: Cambridge Univ. Press.

Connell JH (1978) Diversity in tropical rain forests and coral reefs. *Science* 199: 1302–1310.

Connell JH (1980) Diversity and the coevolution of competitors, or the ghost of competition past. *Oikos* 35: 131–138.

Denslow JS (1980) Gap partitioning among tropical rainforest trees. *Biotropica* 12: S47–S55.

Grubb PJ (1977) The maintenance of species-richness in plant communities: The importance of the regeneration niche. *Biol. Rev.* 52: 107–145.

Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

Hubbell SP and Foster RB (1986) Biology, chance, and history and the structure of tropical rain forest tree communities. In: Diamond J and Case TJ (eds.) *Community Ecology*, pp. 314–329. New York: Harper & Row.

Hubbell SP, Foster RB, O'Brien ST, *et al.* (1999) Light gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* 283: 554–557.

Huston MA (1994) *Biological Diversity.* Cambridge, UK: Cambridge Univ. Press.

Ricklefs RE (1977) Environmental heterogeneity and plant species diversity: An hypothesis. *Am. Nat.* 111: 376–381.

Schulze E-D and Mooney HA (eds.) (1993) *Biodiversity and Ecosystem Function.* Berlin: Springer-Verlag.

Tilman D (1982) *Resource Competition and Community Structure.* Princeton, NJ: Princeton Univ. Press.

Tilman D (1994) Competition and biodiversity in spatially structured habitats. *Ecology* 75: 2–16.

Tilman D and Pacala S (1993) The maintenance of species richness in plant communities. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 13–25. Chicago: Univ. of Chicago Press.

Whitmore TC (1989) Canopy gaps and the major groups of forest trees. *Ecology* 70: 536–538.

Restoration of Animal, Plant, and Microbial Diversity

Edith B Allen, University of California, CA, USA
Joel S Brown, University of Illinois, IL, USA
Michael F Allen, Center for Conservation Biology, University of California, CA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 185–202, © 2001, Elsevier Inc.

Glossary

Active restoration Requiring manipulation by humans for successful colonization and/or establishment of organisms and ecosystem functioning.
Functional diversity Having all of the functions required for maintenance of ecosystem processes, but not necessarily all of the species richness.
Functional redundancy, functional similarity Having species that may be substituted because their contributions to ecosystem processes are similar or overlapping.
Passive restoration Relying on natural successional processes for restoration after the stresses that caused the disturbance have been removed.
Reclamation A revegetation or land management goal that includes a lower diversity of species and may include substitutions by introduced species.
Reference area An undisturbed or natural area chosen to compare with a restored site to determine the success of restoration.
Rehabilitation Creation of an alternative ecosystem following a disturbance, different from the original and having utilitarian rather than conservation values.
Restoration The manipulation of organisms and ecological processes to create self-organizing ecosystems that resemble predisturbance structure and functioning and promote conservation of biodiversity.

Restoration Goals: Conservation and Biodiversity

Conservation of biodiversity is the central goal of most restoration efforts and ranges from reintroductions of individual species of rare plants and animals to efforts to reintroduce a high diversity of species (Jordan *et al.*, 1987, 1988; Bowles and Whelan, 1994; Falk *et al.*, 1996). Restoration may be defined as the manipulation of organisms and ecological processes to create self-organizing, sustainable, native ecosystems as integral parts of the landscape, as much as possible as they existed before disruptive human disturbances. Where propagules of native organisms are remnant, restoration may require reintroducing ecosystem functions such as fire or hydrologic regime to enable natural recolonization and recovery processes. Re-creation of a close replica of a previously existing ecosystem type is increasingly difficult in a world with a growing human population, considering fragmentation and limits to dispersal of organisms; global invasions of exotic species that cause large-scale replacement and even extinction of native species; air, water, and soil eutrophication; and habitat loss by land conversion to urbanization and agriculture. Even nature reserves and parks suffer from visitor overutilization and other impacts, and restoration coupled with more careful management is required to preserve the original flora and fauna.

There are limits to restoration whether biodiversity or functioning of the ecosystem is concerned, but restoring all of the species richness that was originally present on a site is usually more difficult than restoring similar functioning. The reasons are that many rare species are difficult to propagate, their basic biology has often not been studied, they have lost genetic diversity, and their propagules are limited due to habitat loss. They are also less likely to be chosen for many kinds of restoration because they are more expensive to reintroduce and they contribute relatively little as individual species to ecosystem functioning. On the other hand, abundant species are generally better known ecologically, and in some cases, individual abundant species will regulate ecosystem functioning. Restorationists more often focus on reintroducing the abundant plant species or the "matrix" species to initiate succession and recovery of disturbed lands. Reintroducing dominant species enables a recovery of major functions and most of the vegetation vertical and horizontal structure. For instance, John Ewel showed that low-diversity replanted (but not restored) tropical forest in Costa Rica may have similar levels of soil nutrients, organic matter, and nutrient loss, compared to natural forest, but will not have the same conservation value for species preservation.

Additional limits to restoration of diversity are economic, political, and social. Restoration for the intrinsic value of nature, with its complement of biodiversity, often includes participation by laypersons who wish to see nature returned to a state they remember from years ago or that was documented in historical texts or anecdotally. Lesser goals than restoration include reclamation or rehabilitation (National Academy of Sciences, 1974), which have utilitarian values and less emphasis on conservation and biodiversity values. Reclamation may include a less diverse mix of species and exotic substitutions, while rehabilitation is simply to make the land useful again and may produce an alternative ecosystem, such as turning a forest into a pasture.

In this article, we will examine the possibilities and the limits to restoration of biodiversity. Different species have different limits and require different approaches for restoration. We divide the discussion into restoration of plants, animals, and soil microorganisms. This is admittedly an overly simplistic approach, as they all interact. However, points of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00419-6

overlap among the groups are brought out numerous times within each section. The division is logical in that the three groups require different degrees of active or passive restoration, depending on the level of disturbance of a site (**Figure 1**). The kinds of disturbances that require restoration are varied. They range from drastic alteration of the ecosystem, such as surface mining or other construction projects that remove the topsoil; to abandonment from agriculture that leaves soil nutrients and a largely weedy seed bank; to invasion of natural vegetation by exotic weeds that leaves the soil intact but depletes native biodiversity; to alteration of certain ecosystem functions such as fire cycle or hydrologic regime; to extirpation of individual species with no other associated impacts to the physical habitat. Plant introduction and management are central to restoration projects where disturbance has caused vegetation removal or weed invasion. Where plant propagules remain or are readily dispersed, as after abandonment from agriculture in the northeastern United States, passive restoration may suffice. For animals, habitat may be created by manipulating the structure of the plant community, with the hope that the animals will recolonize. This has come to be called the "build it and they will come" hypothesis (see article by Palmer *et al.* in *Restoration Ecology*, Vol. 5, 1997). However, recolonization may be restricted by fragmentation, lack of corridors, or lack of propagules, so reintroduction of the animal into restored or intact vegetation may be needed.

For microorganisms, passive restoration is generally the rule in contrast to higher organisms that may be purposefully reintroduced. Very few species of microorganisms that are members of natural ecosystems are cultured, so they are not even available for restoration purposes. The exceptions are symbionts (N-fixing bacteria, mycorrhizae) that are used routinely in agriculture and reforestation, but less often in restoration. All the other taxa that contribute to belowground functioning of native ecosystems are never or rarely cultured, including soil saprotrophic fungi, bacteria, nematodes, and other soil micro- and mesofauna. For all of these groups, we must learn how to manage the soil to promote their recolonization and to understand the distances from which they may be able to recolonize. In addition, there may be 1000s of species of microorganisms even in 1 cm^3 of soil. So the argument must be made whether all these species can or should be reintroduced. Many ecologists argue that there is functional redundancy among so many microbial species and that a minimum number of functional groups rather than a minimum number of species are required for restoration of ecosystem functioning.

The goals for restoration may vary for these three groups of organisms. For plants and animals, species diversity concerns are high, but for soil microorganisms, functional diversity is more often expressed as the concern. The limits and possibilities for restoration of diversity vary for the three groups, and the discussion below will expand upon these topics.

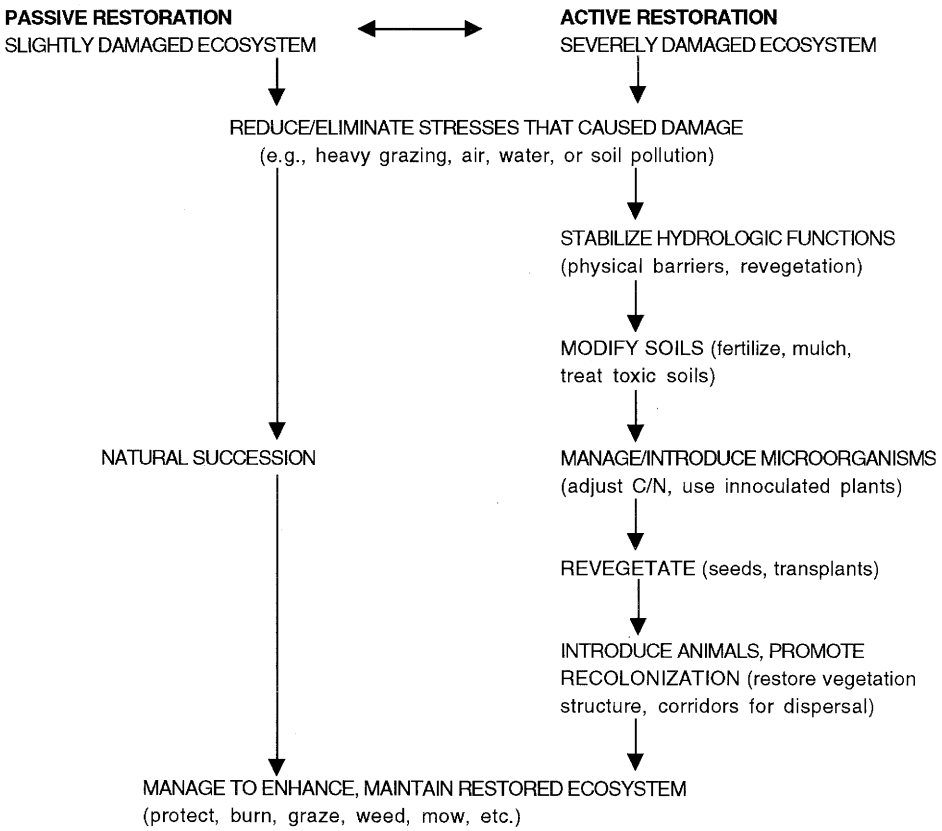


Figure 1 A comparison of active and passive restoration. Active restoration requires various degrees of human intervention, depending on the degree of disturbance, while passive restoration only requires removal of the stresses that caused the initial disturbance, such as grazing or pollution, followed by natural recovery processes.

Restoring a Diverse Plant Community

To improve plant species diversity, the kind of revegetation or vegetation management is determined by the degree of disturbance. These form a gradient of disturbance types and recovery possibilities that require the entire range of active to passive restoration. Heavily disturbed sites such as surface mines need to have all species reintroduced, whereas slightly disturbed sites may be missing certain species that are not adapted to the new postdisturbance conditions, such as alteration of fire regime. Other sites may require weed control to increase native species diversity. There are no examples of severely disturbed sites that have been restored to their original complement of species. This may be related to the high costs involved, the changed environmental conditions, and the sheer impossibility of artificially reintroducing a large number of species. As richness increases in more productive habitats, the probability of reintroducing even a majority of the species diminishes. One of the best examples of attempts to restore a diversity of species comes from bauxite mining in southwestern Australia, where 80 species or more may be included in the seed mix. However, the adjacent undisturbed jarrah (*Eucalyptus marginata*) forest may contain more than 200 species. In Wyoming sagebrush–grasslands, mining reclamation regulations are not as strict, and typically only 5–10 native species are planted in an area that may have 50 or more naturally occurring (**Figure 2**). The dominant rather than the rare species are typically chosen for revegetation, and the long "tail" of rare species that typifies a dominance–diversity curve for natural areas is missing (**Figure 2**). Reclamation and restoration both generally create communities of a few abundant species but few rare species, whereas undisturbed communities may have the same few abundant species, but in addition will have many rare species.

Limits to Restoring a Diverse Plant Community

The reasons for omitting the many species that form the rare species tail in **Figure 2** are many. Seeds or other propagules must be collected locally for the restoration to reflect the local

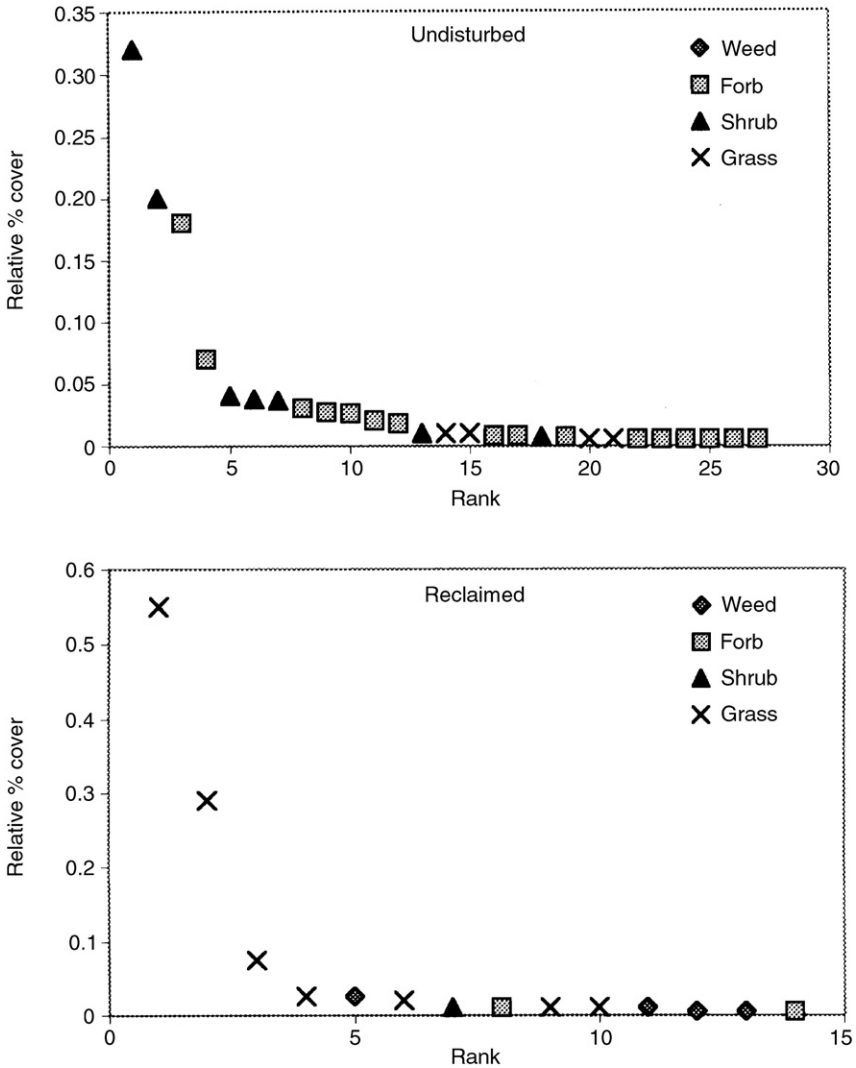


Figure 2 Dominance–diversity curves for reclaimed and natural vegetation in Wyoming sagebrush–grassland. The long tail of inabundant species found in natural vegetation is seldom reestablished in revegetated areas.

genetic populations, but these are typically not available unless the project is planned well in advance, usually two growing seasons, and the seeds are collected specifically for that project. Collecting native seed is becoming a large industry but it is unusual to have all the species available on the open market at the time they are needed, and even rarer to have the seed collections from local populations (see chapters in Falk *et al.*, 1996).

Loss of genetic diversity and lack of local ecotypes are also a limitation to restoration. Where local extirpations have occurred, the nearest populations may no longer have all the genetic diversity of the original, and restoration in the true sense of restoring genetic as well as species diversity is no longer possible. Locally selected ecotypes have the "home team advantage" of being better adapted to local conditions and also avoid problems of outcrossing and hybrid depression with remnant native individuals (see article by Montalvo *et al.* in *Restoration Ecology*, Vol. 5, 1997). Additional discussion of genetic issues is in the section on animal restoration.

Once the appropriate local seeds or propagules have been collected, there is little information on seed dormancy and propagation of most species. The emphasis on conservation biology in the past decade has resulted in increased concern for research on rare species, and information on propagation, microenvironment requirements, reasons for disappearance, interactions with pollinators and other species, and so forth have enabled restorationists to include rare species in revegetation plans. More often, restorationists work with relatively unknown species and must begin research anew for each species, as, for example, the work on restoring the endangered *Amsinckia grandiflora* (see chapter by Pavlik in Falk *et al.*, 1996). Lack of biological information on species translates into practical economic limitations. Most often, only the species that are best understood with the most available propagules are used.

Once germination and propagation requirements are understood, the plants must go into a field setting that presents a whole new set of problems. Different species germinate at different times and have different growth rates, so some will never emerge from certain seed mixes. The northern Great Plains of the United States are dominated by *Bouteloua gracilis* in the coal mining regions, but this has proven to be one of the most difficult species to reestablish for mining reclamation. It is a slow-growing, late-germinating, warm-season grass, but when it is seeded as part of a mix of native species, the cool-season grasses germinate first, grow quickly, and dominate. Reclamationists have devised numerous methods to reestablish short-grass prairie, such as alternating seed drill rows or planting *Bouteloua* seeds a year earlier. However, the most important consideration for companies that have spent millions of dollars on earth-moving is to stabilize the soil, rather than to establish high levels of diversity that are not required by law in any case. Preventing soil erosion is the first goal, and the mines of this region use a fast-growing native plant mix that reduces the establishment of slow-growing species. The goals of soil stabilization using productive plant species and establishing a diverse mixture are often at odds (**Figure 3**).

Reestablishing the full complement of species may require reintroduction in stages, as shown in examples from the tallgrass prairie. When Robert Betz began prairie restoration in the late 1960s at the Fermi Lab prairie in Illinois, he quickly learned that the dominant native grasses could be readily established from seed. He used the grasses to form the matrix of vegetation, followed by later introductions of the less common forbs. These forbs could not colonize naturally into the restored grassland because of the high density of grasses, but could be hand planted. Thus Betz was able to restore 116 plant species to the Fermi Lab prairie, but only by expensive hand labor. Similarly, the Curtis Prairie in Wisconsin, the oldest restored prairie, was initially planted during the 1930s, but plants have been continually introduced since then (see chapter by Cottam in Jordan *et al.*, 1987). Their survival has been monitored, and they now contribute to the most diverse tallgrass prairie anywhere. The Curtis Prairie has over 300 species in 10 acres, more than any remnant natural prairie.

Competition from invasive plant species is another important limit to restoration of diversity. The invasives may be

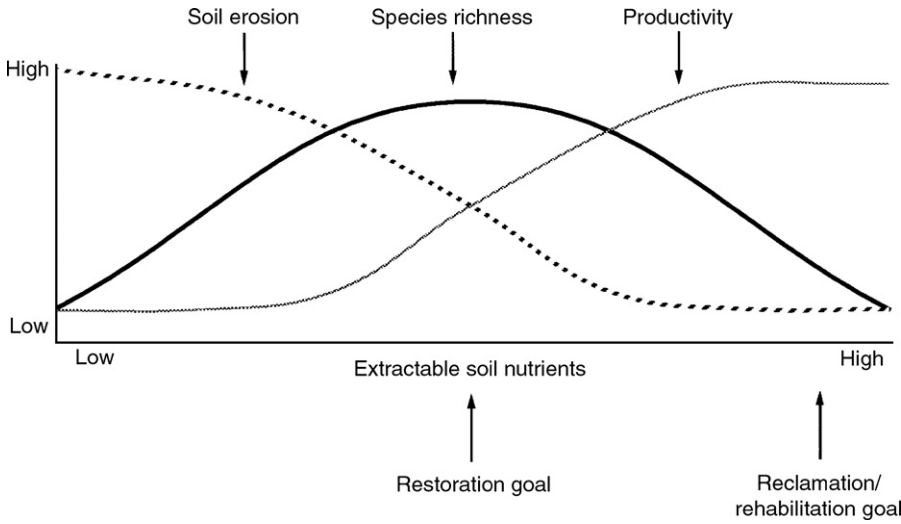


Figure 3 The relationships among soil fertilizer level, plant species richness, plant productivity, and soil erosion. The trade-off for using fertilizer to promote high initial rates of plant productivity to reduce soil erosion is a loss in diversity.

weeds that are part of the successional process, but a more difficult problem than early seral weeds are those invasive plant species that persist and cause vegetation type conversions. In this case, the major restoration activities involve mowing, fire, or selective weeding to remove the offensive species, planting or managing for regrowth of natives, and then continual management to keep the invasives from becoming dominant once again (see chapters in Allen, 1988; Falk *et al.*, 1996). Such activities have been undertaken in Florida, where *Schinus terebinthifolium* and *Melaleuca quinquenervia* dominate wetlands, Hawaii, where exotic perennial grasses and *Myrica faya* have replaced native vegetation, California, where Mediterranean annual grasses have replaced native perennial grasslands, and in many other sites worldwide (Lee, Richardson in *Encyclopedia of Biodiversity*).

Poor, toxic, eutrophied, or erosive soils also impair the conservation value of restored communities. Any practice that changes the soil from its original physical or chemical state will also change the plant species composition. An example comes from the china clay mines in England, where the mining refuse is high in sand. Local plants are poorly adapted to the droughty and nutrient-poor sand, so strand plants are imported for revegetation (Bradshaw and Chadwick, 1980). As expected, these artificial strand communities are poor in diversity. However, Bradshaw has documented some relatively high species diversity following natural colonization onto abandoned waste sites, although not as high as surrounding native areas. There is a tendency during revegetation to overfertilize soils to agricultural levels when fertilizer is applied, with expected poor results for native plant diversity. The reason for overfertilizing is usually economic, to promote vegetation establishment and therefore prevent soil erosion, but may also be aesthetic, to please the paying public as rapidly as possible. As shown by the experiments of Tilman, (Tilman in *Encyclopedia of Biodiversity*) overfertilization, especially by nitrogen, will promote the productivity of a few species to the detriment of stand richness (**Figure 3**). For restoration to promote the natural levels of species richness, the soils must be in their natural fertility condition as much as possible. If erosion is a problem, then other means to control erosion, such as mulch or temporary mechanical barriers, could be used to stabilize the soil and allow the slow-growing plant species to establish within a mixture of fast-growing species.

Improving Plant Species Diversity

Restorationists have used many techniques to increase richness of plant communities when the seed and propagule sources are limiting. Taking advantage of the existing seedbank is a primary one, where the seedbank still exists. Mined land reclamation laws in developed nations often require retopsoiling with fresh topsoil, and many studies document the importance of this source. Certain species still need to be supplemented, as in the case of late seral species absent from the seedbank of jarrah forest in southwestern Australia (see chapter by Bell in Allen, 1988). Jarrah seed is never found in the seedbank, and it reproduces vegetatively after fire in natural communities. It must be reintroduced as nursery transplants for revegetation to be successful. While the seedbank is an important source of diversity in many kinds of restorations, it is also notorious for harboring a large complement of early successional and weedy species. Bradshaw and Chadwick (1980) have recommended against using topsoil just for this reason, if appropriate weed-free subsoil can be found and amended to support plant growth.

Dispersal of propagules into the restored site is another way to increase diversity that depends upon landscape structure such as proximity of source populations or existence of corridors. In a study of 2- to 18-year-old revegetated, untopsoiled roadsides in San Diego County, dispersal from adjacent mediterranean shrubland more than doubled the richness from 12 planted species to a maximum of 16 colonizing native species. However, dispersal of native species only occurred when the shrublands were adjacent to the highway, and occurred rarely in urban areas of the highway. Reclaimed surface mines in Wyoming that are surrounded by native sagebrush–grassland have a higher richness of colonizing species than, for instance, a revegetated landfill in Staten Island that is surrounded by suburbia (see articles by Bell and by Ehrenfeld in *Restoration Ecology*, Vol. 5, 1997). While dispersal is often limited by vectors in terrestrial habitats, aquatic and riparian habitats typically are recolonized rapidly (National Academy of Sciences, 1992). Water is a very effective medium for propagule dispersal, and these habitats are often not even revegetated, unless the soil needs to be stabilized rapidly in riparian edges or the dispersing propagules include unwanted exotic species.

Vegetation may be replanted to attract animals that are seed dispersers or pollinators and thus may create a positive feedback on the future reproduction and diversity of a site. When a limited suite of species is chosen, the selected species become especially critical. Synthetic grasslands that simulated tallgrass prairie had sufficient diversity and especially structure to attract birds and small mammals (Howe and Brown, 1999). Shrub islands were planted on a surface mine in Wyoming to increase the movement of animals and microorganisms onto a site that was otherwise dominated by grasses (Allen, 1988). The feedback of these shrub species and patterns determined both their mortality and recolonization patterns of additional plant species. In Costa Rica, Karen Holl and Daniel Janzen have shown that tree and shrub "islands" within pastures are critical for attracting animals that disperse seeds and then increase the diversity of the plant community.

Single-Species Plant Introductions

Individual plant species have been reintroduced where they were extirpated, largely because of laws protecting endangered species. Mitigation laws often require transplantation of a rare species from a site that is about to be destroyed to a safe site. For instance, Howald (see chapter in Falk *et al.*, 1996) documents 40 instances where rare species were transplanted for habitat mitigation purposes in California. Of these, only 5 were considered successful, 7 had limited success, 13 were not successful, and others were unknown or ongoing and too recent to evaluate. The reasons for lack of success were varied but included moving plants to sites where they did not exist

previously, having different environmental conditions in the transplant site, or using poor horticultural techniques. Primack (chapter in Falk *et al.*, 1996) planted seeds of 41 species into sites in Massachusetts where they were once known to occur. Of these, only 10 produced seedlings, and only three produced a second generation. In this study, the reasons for failure may have been due to exotic species, changes in natural disturbance regime, and other changes such as anthropogenic nitrogen deposition that causes inappropriately high levels of soil fertility. The relatively few successful transplantations indicate that mitigation transplantation in general is poor policy because the kind of planning and postintroduction management that are required is seldom done in practice. Often little consideration is given to the quality of the habitat into which organisms are being reintroduced. Overall, the reintroduction of one species does little to improve the diversity of a site, but it may be the only option to avoid extinction, provided restorationists and land managers use their best practices.

While many single-species introductions focus on rare species, they may alternatively focus on abundant species. The reintroduction of a keystone species or ecosystem engineer (see Palmer *et al*. in *Restoration Ecology*, Vol. 5, 1997) is one way to "jump start" natural successional processes in a community. Shrubs and trees were removed from rangelands for many years in an effort to increase forage production for domestic animals, but this was short-lived and in the long run promoted degradation of the community because the deep-rooted shrubs changed the nutrient and moisture balance of the entire stand. The reintroductions of *Pistacia lentiscus* in Israel, oak trees in California, or *Artemisia tridentata* in mined lands in Wyoming enable the shrub or tree island effect to reinitiate, building up soil and biotic resources around the transplanted shrub (Allen, 1988). This could potentially result in further increases in diversity as the shrubs enable recolonization of additional species.

Succession on Restored Lands

Restoration depends upon succession to complete the process that was started by revegetation, but succession alone does not always bring us to the goal we wish to achieve. Revegetated lands are not static and can be expected to undergo change in species density, richness, and relative composition. The degree to which they undergo succession varies, depending upon the amount of colonization, initial species complement, and site conditions. Little change in species composition occurred after bauxite mining in the Australian jarrah forest. The site typified Egler's initial floristics model of succession, where late seral dominants colonize in the earliest stages of succession, because very few species that did not occur in the original seed mix became dominant later on. Sites that were 16 years old appeared structurally different from sites that were 1–2 years old in that woody plants were larger and more evident in the older sites, but most of the same species were present in sites of all age classes. A few surprising species did colonize the older sites, such as a native orchid. In Anthony Bradshaw's experiments on plants colonizing industrial waste heaps in England, most plants that did not establish appeared to be limited by dispersal and establishment, rather than poor substrate. The sites were too distant from propagule sources or had dense stands of initially colonizing species that prevented establishment of a diversity of species. The initial floristics model can also be applied to Wyoming, Pennsylvania, and Czech surface mined lands, where the native grasses that dominated early on allowed little colonization of native trees, shrubs, or forbs after 10–30 years.

Even when soil conditions are optimal and sources of propagules for local colonization are available, the initial revegetation treatments will affect long-term successional processes, as shown by recent studies in the journal *Restoration Ecology*. A 30-year-old planted forest that had been severely disturbed by hydroelectric development in New Zealand showed little change over time when 46 native species were planted. The species composition was chosen to resemble the dominants during secondary succession after fire in natural forest, but had only a fraction of the total diversity. However, an even less diverse stand developed when no vegetation was planted, consisting of a stand of exotic *Citisus scoparius* and European grasses. A comparison of planted forest stands on bauxite mines in Brazilian tropical forest showed that planting a low-diversity commercial mix resulted in a low-diversity, but high-productivity forest, while seeding and planting mixes of 70 or more species promoted a diverse, but less productive forest. In all cases, colonization of certain late successional forest species that were not part of the original mix was poor. An interesting contrast comes from the Arctic on the North Slope in Alaska, where more than 100 species of native plants colonized 20 years after planting two native and one exotic grass species over a dozen sites. Thus the productivity and species richness of the initial planted stand will make a large difference in the ability of other species to colonize later. From these examples, and many others in the literature, it is apparent that succession may promote an increase in richness if the planted stand is not so aggressive that it precludes establishment of colonizers. Most often, additional interventions are needed if the desired diversity is to be achieved.

Restoring Animal Diversity

An Introduction to Animal Restoration

Like plant restoration, animal restoration involves reintroducing or encouraging the return of native species to an area or region from which they have been lost. An animal restoration project may involve just a single species, as in the reintroduction of wolves into Yellowstone National Park, or an entire community, as in Augrubies Falls National Park in South Africa, where over 10 species of large mammals have been reintroduced (including black rhinoceros, eland, giraffe, and springbok).

Animal and plant restorations will generally involve the same ecological considerations, and most differences between the two concern a matter of degree and emphasis (Jordan *et al.*, 1988). Save for insects (which represent a challenging form of restoration more in line with promoting microbial diversity), plant restorations generally call for a larger number of species. In terms of ecosystem function, animals have significant roles in restoration as pollinators and herbivores of

the plants. Animal roles in ecosystem function also occur when herbivores alter nutrient cycles and where animals may be "ecological engineers" of the physical environment, such as the contributions to soil function from the digging and foraging activities of earthworms or burrowing mammals (Jordan *et al.*, 1987). In general, animals are more mobile than plants and while plants may-disperse widely and unexpectedly as seeds and pollen, a single animal may range over scales that are much larger than that of the restoration project and site. Some of John Laundre's radiocollared mountain lions from south central Idaho have appeared as far away as Yellowstone National Park in Wyoming and mountain ranges in central Nevada. As a consequence of mobility, restoration sites from the perspective of animals will be open rather than closed systems. In animal restoration, the environs of the restoration site may deserve equal attention. Much attention has been given by all stakeholders to the possible fates and consequences of wolves wandering in and out of their Yellowstone restoration site.

The ecology of animal restoration draws less heavily from theories of ecological succession and more from the ecology of invasions and the ecology of small population sizes (Major, 1989). Plant restorations, as succession, start with an inoculum of plants that initiate a process of changes in species composition that eventually leads to the desired natural community in terms of structure, function, and appearance. This aspect of succession is absent from most animal restorations except in the case of the mostly passive development of an insect and soil invertebrate community that fits and complements the plant community. With most animal restorations, each species enters the community at a very small population size either as an invader or as a reintroduction. Small and alien describes the starting population. Most random introductions of animals to a community fail. Naive animals fall to predators more easily, compete less successfully for resources, often fail to develop or find suitable denning or nesting sites, and sometimes attempt to emigrate from the site. In the absence of preexisting burrows, reintroduced prairie dogs simply roam far and wide in search of colonies. Small animal populations face twin genetic and demographic threats. Small, sexually reproducing populations may suffer inbreeding depression and produce a higher proportion of young with genetic defects. Demographically, small populations may teeter on the edge of extinction. An accidental death, a missed breeding opportunity, or a chance skewing of sex ratios or age distributions may compromise the population irrecoverably.

Animal restoration begins with knowledge of why the species or animal community is currently absent or threatened at the site. Next comes an assessment of whether the site is currently suitable for the animals and an evaluation of what site preparations are necessary to promote success. Then the current status of the animals receives attention to determine the best and likeliest sources of individuals. The animal(s) may be present on site but at low numbers, they may be present near the site or far from the site, or they may only exist as captive populations. Finally, the passive or active restoration of the animals begins with appropriate considerations of how to manage and monitor the project's success. We will briefly examine each of these steps.

Reasons for the Absence of an Animal or Animal Community from a Site or Region

Habitat fragmentation or changes in land use may reduce or eliminate particular animal species, subsets of species, or entire taxonomic groups of animals from an area. For instance, most animals associate with particular types and structures of vegetation. When altered, by livestock grazing or agriculture for instance, some of the original animals disappear as they cannot find food, nesting, or denning sites or they may be excluded by changing intensities of competition or predation from other animals more suited to the new circumstances. In rivers and lakes, pollution, erosion, and sedimentation drastically change species composition. Everything from invertebrate larvae to fish may die off or be replaced by other species more tolerant of the polluted or modified waterways.

Hunting, poaching, and commercial harvesting may be so intense as to extirpate an animal or group of animals from an area or region. The African gray parrot and South American macaws face threats from those supplying the demand for pets. Incidental mortality in fishing nets threatens many of the world's populations and species of sea turtle. Overharvest of commercially valuable species has created a litany of crashes such as California sardine, North Sea herring, blue whales, kaluga sturgeon, and southern bluefin tuna. As a valuable source of meat and a denizen of potential rangeland and farmland, the pronghorn antelope of the western half of the United States, like the American bison, faced extirpation by the 1900s.

Exotic species, those species accidentally or intentionally transported by humans into new places, often eliminate native animals via predation, competition, or keystone effects on structure and function of the ecosystem. The introduction of mosquitoes into Hawaii brought bird malaria for which the birds were no more prepared for than the peoples of the Americas were for the introduction of smallpox from Europe. An inoculum of larvae in ship ballast introduced zebra mussels from Europe into the Great Lakes of North America. In the early 1990s, the mussel's spread was spectacular, and with it has come dramatic declines in the abundances of phytoplankton and the fish and invertebrates relying on the phytoplankton. Competition from introduced North American gray squirrels has eliminated the native European red squirrel from much of its former range in England and Wales.

With respect to restoring their wildlife, Australia and Israel provide interesting contrasts. Both have seen their native mammals ravaged. In Israel this has gone on dramatically for millennia whereas it is more recent in Australia. In Israel, overexploitation and habitat modification have been the bane of most animals like lions and crocodiles, both extirpated long ago. Deforestation, reaching its nadir under the Turks just prior to World War I, nearly extirpated the subspecies of European red squirrel and jay from Israel's wooded habitats. In Australia, wildlife habitat is much more available and overexploitation generally more benign than in the Middle East. However, introduced species (red foxes, feral house cats, and European rabbit) of animals and plants have dictated the decline in numbers and range of comparable Australian mammals. Having contributed to their extirpation on the mainland, cats now occupy Kangaroo Island off of southern Australia and threaten the last remaining population of Tamar

wallaby. Feral goats compete with and threaten the dwindling population of yellow-footed rock wallabies in the Flinder's Range. Red foxes prey heavily upon Australia's native small mammals and consume the eggs of ground-nesting birds.

The reasons for the absence of animal species from their former range are manifold but fall roughly into the categories of habitat change, overexploitation, or exclusion by animal species nonnative to the particular location. Understanding the reason for the absence of an animal or community of animals is the essential starting point for animal restoration.

Site Assessment and Habitat Restoration

All species have the capacity to grow exponentially under ideal conditions, but no population can grow exponentially forever. There are limits to growth, and population interactions such as competition and predation from other species can exclude a species from a community. The goal of site assessment and habitat restoration is to ensure near ideal conditions for the single species or the community of animals under restoration. Ideal conditions means ample food, space, shelter, and safety from predation. Limits to growth invites a consideration of specific factors in the environment that might currently or ultimately limit the growth and size of the animal's population. Considerations of competition and predation focus on which current species at the site, possibly exotic species, might preclude a successful introduction of the desired species to the site.

Site assessment evaluates whether the area offers an appropriate combination of food, safety, shelter, and space for a small population of animals to grow exponentially and prosper. If conditions are not appropriate, then habitat restoration must precede animal restoration. If the vegetation of a site has been heavily degraded, animal restoration may await plant restoration. At Midewin National Grassland south of Chicago, long-term hopes include first restoring the tallgrass prairie and then reintroducing elk and bison. The site is currently a mix of abandoned munitions facilities, farm fields, pastures and assorted oldfields, woodlots, and groves. Wetland restorations often require establishing the aquatic and surrounding vegetation and then passively or actively reintroducing the associated aquatic invertebrates (usually all passive), vertebrates (usually passive but nest boxes or nesting platforms may be used to encourage the return of waterfowl, egrets, or herons), and fish (sometimes passive or active). Different species have different needs, and reconstructing those habitats that fill an animal's needs is central to attracting wildlife or to introducing wildlife.

The possible reasons for the animals' absence from a site focuses site assessment. Often several factors combine to explain an animal's absence. When habitat alteration or fragmentation changes and reduces the diversity of animals, the first step is to determine whether the site has been, or needs to be, revegetated to its former state. This condition of the site does not have to be exact with respect to its former state, but rather it need only include the salient environmental factors that favor the animal's ecological requirements and aptitudes. For instance, many bird species are most responsive to the structure rather than the exact composition of the vegetation. Hence, bird community restoration can commence with promoting the appropriate vegetation structure rather than being particular to its composition (which may be the goal of a corresponding plant restoration). Habitat fragmentation offers unique challenges because the remaining area may simply be too small to successfully support the desired animal or animal communities. The site assessment may examine the need for more space or wildlife easements, the need for corridors to connect habitat fragments, or the need to enhance the quality of a habitat fragment beyond its natural state to permit the successful persistence of animals in a smaller, more confined space. At Jackson Hole, Wyoming, a wintering elk population is maintained on supplemental hay within a fenced pasture surrounded by the extensive human developments that have robbed the elk of most of their wintering habitat. In spring, the elk return to their natural mountain pastures. In Israel and elsewhere, feeding stations stocked with animal carcasses or offal have been used to facilitate the maintenance and restoration of such animals as wolves, griffon vultures, and eagles.

The California condor provides an illustration of a complex site assessment. Reasons for the decline of condors that culminated in their removal from the wild in the late 1980s included too small a population (demographic threats), habitat fragmentation, lack of carcasses, and lead poisoning from carcasses containing lead shot or bullets. The restoration plan called for captive breeding, gazetting of habitat and habitat corridors, restrictions on development, and for reasons beyond just the condor, the banning of lead shot for hunting. The actual restoration took several approaches to evaluating the complexities of habitat suitability and the ability to introduce a small population of naive animals. First, less rare Andean condors were introduced as a surrogate species. They permitted tests of the reintroduction techniques and indicators of ranging patterns and habitat suitability. The actual reintroductions have used the original site north of Los Angeles but also additional sites, including one near the Grand Canyon. The California release site has more recently had condors but suffers much greater degrees of habitat fragmentation and development. The alternative sites have seen condors less recently, but have the advantage of offering extensive tracts of original and natural habitat.

In response to habitat degradation or fragmentation, habitat restoration may be as simple as adding nest boxes, as in the case of bluebirds at some sites in northeastern Illinois, or as involved as adding proper soil, microbes, and restoring an entire vegetation community, as in the case of toxic waste sites or former mining operations. Ceasing the use of pesticides or release of pollutants may be all that is necessary to restore a habitat. Peregrine falcons in North America have benefited from the ban on DDT. Portions of fish communities and whole invertebrate communities can recover just from preventing sewage and pollution discharges into waterways. The structure of the environment with respect to shrub cover in arid lands, the mixture of ages and types of trees in forests, and the availability of salt licks and waterholes for wildlife all become considerations for particular habitat restorations. Koala in Australia require particular species of eucalyptus trees as food, whereas gravel beds at particular depths provide spawning grounds for trout of the Great Lakes.

When a species absence is the consequence of overexploitation, success often depends less on the availability of

suitable habitat and more on the cessation of hunting, poisoning (including mortality from hazardous chemicals and pollutants), harvesting, or poaching. Wolves were exterminated throughout the western half of the United States through individual hunting and trapping and by explicit eradication programs. Their reintroductions into Yellowstone National Park, Wyoming, and the White Mountains of Arizona involve presumably moderate to high quality habitat. The success of the programs probably depends most on the cessation of poaching, which at present is a minor problem for Yellowstone and a major threat to the Arizona reintroduction. In 1989, it was discovered that stock assessments of cod within the Canadian North Atlantic were wildly optimistic. The stock had crashed. Subsequent quotas follow from the premise that habitat quality remains high and that reduced harvest will be sufficient to promote recovery. In Southeast Asia, ediblenest swiftlets face decline and extirpation. The birds nest in caves, use saliva for nest building, and face true "nest predation." Humans gather the nests even as the nests may have eggs or nestlings. Even with complete bans on nest harvesting, places like Sarawak, Malaysia, still face intense nest destruction from local people, pirates from countries such as Indonesia, and organized collecting groups. Policing thousands of cave entrances in often remote places throughout the country is impractical. But, until poaching relaxes or ceases, restoration and recovery cannot be achieved. Having reduced poaching of black rhinoceros throughout the 1980s, the Kenyan Wildlife Service has begun the recovery and restoration of current and former populations.

Habitat restoration for overexploited species requires the first step of controlling the harvesting. However, the absence of wildlife from an area may coincide with other changes to the habitat, some which at first may seem subtle. "Nature abhors a vacuum" is a saying recognizing that unused food or opportunities in an environment often become filled by alternative species or by exotics. Habitat restoration may require controlling the abundance of exotic species or those species that compete or prey. Factors contributing to the decline of red squirrels in England include an increase in oaks and a decline in conifers. Both factors assist the exotic gray squirrel in outcompeting the red squirrel. Habitat restoration may require both changes in forest composition and active control measures of gray squirrels. The draining of reservoirs, netting, or poisoning of lakes infested with exotic fish such as carp has preceded several fish restorations in North America.

Sources of Animals for Restorations

Small remnant populations, dispersal from other areas, or active reintroductions provide the sources for animal restorations. Ideally, animal restoration should begin while a remnant population still occupies the site. Such a population, while often small, has the advantages of already being established and acclimated to the site. This avoids problems associated with naive animals unfamiliar with the site. In general, once habitat restoration has assured a quality site, resident populations recover faster than immigrants from other areas, which recover faster than reintroduced populations from captive stock. When remnant populations exist, restoration activities can immediately focus on improving vegetation, improving food availability, improving removal of exotics or competitor species, and/or ceasing harvesting and pollution. Success is measured by the recovery of the species or the expansion of the species into previously unoccupied parts of the site.

When a restoration relies on immigrants, there are the twin concerns of: "Is the site suitable for the establishment of immigrants?" and "Will the immigrants find the site?" When a site such as an oldfield, mine tailings, or forest clearcut is restored, there will usually be successful invasions of insects, soil invertebrates, aquatic invertebrates, mammals, and birds from the environs. For instance, in an oldfield subject to small experimental prairie restorations (Howe and Brown, 1999), there was an invasion of two small mammals (prairie vole and white-footed mouse) and some red-wing blackbirds have shifted their breeding from a nearby pond to the prairie plots. Mourning doves preferentially feed on rather than off of the prairies. Invasions of insects, birds, and mammals ease the restoration effort but sacrifice complete control of the resulting composition of animals. Species that are already well established and abundant in the region of a restoration are the most likely invaders. Rare and remote species are the least likely. Target species may require coaxing in the form of habitat corridors connecting the restoration site with existing populations. Temporary augmentation with food or nesting boxes can encourage reestablishment. In Tsavo West National Park, Kenya, rhinos have been corralled into a sanctuary (ca. 70 km^2) that is much smaller than the whole park (ca. 9000 km^2). The sanctuary is fenced to keep the rhinos in, patrolled to protect the animals from poaching, and supplied with piped water to maintain permanent water holes.

Animal restoration may require the active reintroduction of animals (Bowles *et al.*, 1994). This is best done by translocating individuals from wild populations, but may require captive breeding. Wild populations provide experienced individuals, whereas captive animals may be particularly naive and unfamiliar with the wild. The reintroduction population should have the appropriate balance of sexes and age classes and sufficient genetic diversity to preclude serious consequences of inbreeding. When the reintroduction is very expensive, is politically sensitive, or is from a very small source population, great care goes into ensuring that the site offers close to ideal conditions and that the animals are prepared for the new environment. When animals are relatively cheap and available, the reintroductions can be more numerous and there is room for experimentation. Wild asses (onagers) in the Negev Desert of Israel, pronghorn antelope in Arizona, lynx in Switzerland, white rhinoceros in Lake Nakuru of Kenya, and peregrine falcons in various midwestern cities all provide examples of reintroductions from captive or wild source populations. In the case of the Guam rail, this bird had been driven extinct on Guam by the introduction of the brown tree snake. Rather than attempting the impossible task of eradicating the snakes on Guam, the rail has been reintroduced on a neighboring, snake-free island.

Promoting Microbial Diversity

Microbial Diversity and Functional Redundancy

Compared to plants and animals, microorganisms are highly diverse and offer a special challenge to understanding

biodiversity and to assuring successful restoration. Microbes should be the underpinnings of any discussion of biodiversity as they constitute the vast majority of the diversity of any ecosystem at any location, yet are rarely even mentioned in terms of maintaining diversity. Microbes should be a focus, not an afterthought, for restoring disturbed lands. Without animals and most species of plants, ecosystems would stabilize and most ecosystem functions would be performed. Without microbes, the ecosystem would cease to function.

Estimates of microbial diversity range up to 10^5 prokaryotic taxa per gram of soil, although 10^4 is generally a more accepted value. In a study of restoration of a Wyoming coal surface mine, 57 different fungal taxa were found out of 135 colonies randomly chosen from approximately 12,500 colony sources (spores, hyphal fragments, bits of organic matter). These taxa were evaluated within a 4-cm^2 area (using 5 g of soil). Given the numbers, it is not surprising that a large fraction of these organisms remain undescribed. Recent estimates suggest that less than 10% of soil organisms ranging from bacteria to spiders have been described. And, a few efforts that extracted DNA directly from soils have even found new kingdoms of prokaryotes. Despite the difficulty in estimating diversity, every study published has shown a loss in the richness of taxa with severe disturbance. This ranges from soils in the Pumice Plain of Mount St. Helens, which was sterilized, to burned areas in which only the aboveground material and soil surface were directly affected. Because of the extremely high microbial diversity, the concept of microbial functional redundancy has been raised. Among those thousands of species, how many need to be restored to maintain ecosystem functions? This is the single largest challenge for studying soil microbes and restoration.

There are three critical issues for evaluating microbial diversity and restoration. The first issue is defining the spatial and temporal distribution of species and functional groups and their relationship to ecosystem processes. The second concern is assessing the richness of organisms within the different functional groups. The third is the system-level capacity for dispersal and natural reestablishment versus the need for artificial introduction of microbial inoculum.

Spatial and Temporal Arrangement

Just as important as the richness of organisms are the changes in their spatial and temporal distributions. Unfortunately, few studies have evaluated microbial communities using species increment curves or overlap estimates. Several types of analyses are critical to understanding biodiversity and restoration. These include species × area, species × time, and dispersion relative to ecosystem processes. Unfortunately, few data sets are available to evaluate microbial recolonization on this basis.

Microbial richness estimates tend to be taken on a per-sample basis. However, plants and animals tend to be analyzed on an area basis. This makes comparative studies difficult but opens an important area of research. Nevertheless, evaluating the spatial array is absolutely critical. Fungi, for example, exist as a network of hyphae (a mycelium) extending from a few millimeters (such as a *Trichoderma* colony occupying a single fern petiole) to tens of meters in diameter (fairy rings, or the giant Wisconsin *Armillaria*, for example). For the same Wyoming data set, the species increment rate was the same up to the size of a 400-cm^2 patch of disturbed as well as reference area (**Figure 4**). However, in the disturbed site, as one expanded outward, the species increment rate declined whereas it continued to increase in the reference area. In the reference area, new species were added as the habitat changed. In the disturbed area, the habitat was rather uniform across the site (in this case, mixed, respread topsoil on a surface mine). Thus, one conclusion is that microbial activity and composition become more diffuse and repetitive across scales in severely disturbed areas, and overall landscape diversity is lower than in native undisturbed areas.

Frequency of sampling over time is also crucial for describing microbial biodiversity. Many microorganisms are only identified based on sporulating structures. However, these organisms may be present continuously but can only be found periodically. For example, macrofungi are spread widely in the mycelial stage and live for many years. Several continuous years of observations are needed for the right conditions to occur before a sporocarp forms. In many cases, fruiting times may occur over successional time. For example, on Mount St. Helens, establishing the first ectomycorrhiza found on the Pumice Plain took 5 years. We never saw a sporocarp to identify the fungus that formed on the ectomycorrhizal root tips of conifers. Development of techniques for DNA fingerprinting will eventually allow us to identify more of these organisms even when they do not sporulate.

Samples are often taken at the wrong time, leading to erroneous conclusions about microbial diversity. For example,

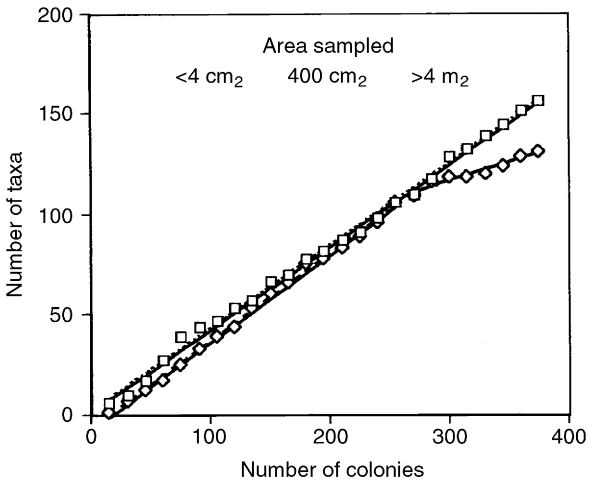


Figure 4 Species increment curve for microfungi at a Wyoming surface mined site. Shown are the taxa in the reference area (squares) versus the disturbed area (diamonds). The increment curve for the reference area is linear through the 375 colonies across the entire area ($y = 1.88 + 0.413x$, $r^2 = 0.997$). The curve for the disturbed area for a small area (4 m^2) is not significantly different from the reference area ($y = -6.743 + 0.435x$, $r^2 = 0.998$) but the slope (0.187) is significantly lower ($r^2 = 0.94$) across the larger areas. These results indicate that at small scales, microbial diversity remains very high. However, at the landscape level, microbial diversity is lost in disturbed communities.

soil animals migrate vertically in response to soil moisture conditions. Soil samples are normally taken from surface soils. Thus, the mesofauna may only be detected when they return to the surface. Usually soil organisms are sampled at the convenience of the investigator, but microbial populations are often event driven (e.g., precipitation). Thus, frequent or a deterministic sampling regime is needed to detect their presence.

Spatial relationships are just as crucial as temporal ones. Ecosystem processes (e.g., decomposition, mineralization, and immobilization) are not uniform across an undisturbed area. Microbial-regulated processes tend to be highly patchy and organized to optimize production. However, possibly the greatest impact of disturbance on ecosystem dynamics is spatially mixing soils and creating relatively uniform conditions across a site. In fact, this led to an oldfield view of succession that still largely dominates restoration practices, where a relatively uniform aboveground community is planted. However, succession may be a patchwork of starts and stops, with a few initial colonists acting as islands that become the nuclei for future colonists. Succession and microbial composition and activity are tightly coupled to the developing patchwork. In restored or recovering ecosystems, these patch recovery patterns are evident. In many abandoned, disturbed sites, no spatial recovery is detectable.

Diversity and Functional Groups

The biodiversity of types of microbes in ecosystems is daunting. However, to a certain degree, maintaining or recovering the functional groups of microbes is the first critical task in restoration efforts. Microbes play every ecosystem role at every site. In fact, microbes alone can, and do, form fully functioning ecosystems without higher plants or larger animals. In the most extreme environments of the Sahara desert and the uplands of the Antarctic Dry Valleys, microbes are the only living organisms, existing on aeolian-deposited or ancient carbon inputs. In many extreme environments, microbes make up the entire ecosystem. These range from the simple endolithic (inside rocks) communities of the Dry Valleys of Antarctica to the thermal pools of New Zealand and Yellowstone geysers. As one proceeds to more favorable environments, more and more types of microbes emerge, subdividing the processes of primary production and decomposition. At all sites, microbes undertake primary production (bacteria, cyanobacteria, algae). The relative contribution of the microbes to the overall proportion of net primary production tends to range from high in more extreme environments (such as deserts and tundra) or situations with dispersed nutrients at low concentrations (open oceans) to low in conditions highly favorable such as tropical rain forests.

In addition to directly fixing C, microbes also catalyze the nutrient cycling processes that transfer elements directly to plants or convert unavailable nutrients into forms that can be taken up and utilized. Thus, they are indirectly linked to carbon fixation by providing limiting resources. Although these organisms are generally modeled as "microbial mass," they often live symbiotically with plants. Mycorrhizal fungi probably have the largest biomass within this group. These fungi form mutualisms with plants and transfer from soil to plant a range of soil resources, from water to N to P. Importantly, they can also make unavailable soil resources, such as bound P, available, by producing organic acids and phosphatases, and Fe with siderophores. Other prokaryotes fix atmospheric N_2, ranging from free-living forms such as cyanobacteria and *Arthrobacter*, to symbionts such as *Frankia* and *Rhizobium*. Other microbes catalyze almost every other nutrient transformation that is biotically important to the sustainability of ecosystems, from N and S transformations, to Fe state transitions, to immobilization of heavy metals and bioremediation of toxic organics. In the case of mutualistic symbionts, many studies have demonstrated that an increasing diversity of species and genotypes can be critical to establishing and maintaining a diversity of plants.

Microbes are the dominant decomposers. Higher animals only take a small fraction (1–10%) of the NPP; the remainder of the energy goes to microbes. The animals themselves constitute a source of a slightly different C source from plant material, making a new type of C resource. Microbes then utilize almost all of the remaining plant material, thereby releasing the nutrients immobilized in plant tissues. Only a small fraction of C remains, as highly complex plant constituents or recalcitrant microbial compounds. These are critical in that this forms the organic matter essential to recovery of all sites.

Free-living Saprobes

In every study, microbial diversity even of disturbed lands continues to increase with increasing sampling. While the actual slope of species increment curves may be lower than for undisturbed areas, it still remains very high. It is not clear if the reduced diversity of microbes is a factor in these detrimental responses. However, it is clear that if the environmental conditions for free-living microorganisms are present, a high diversity of species and a high density of individual cells will reestablish. Thus, restoration of free-living microbes is largely a matter of management of the soils, rarely by inoculation with bags of "beneficial" microorganisms. To date, there is no evidence that biodiversity per se of free-living microorganisms limits saprobic microbial activity in restored lands.

Free-living saprobes form the bulk of the microbial diversity in both functional pathways and the diversity of taxa. As we look at the known studies, those processes catalyzed by free-living microbes always occur, sometimes in detrimental levels. For example, *Thiobacillus ferrooxidans* uses Fe^{2+} in pyrite, which results in the release of sulfuric acid, detrimentally reducing the pH of streams. Immediately following disturbance, there is a rapid mineralization of N, resulting in N leaching and denitrification. *Reduction* in some of these microbial-catalyzed processes often is an important restoration task.

Soil Animals and Food Webs

Microbes consist of prokaryotes and fungi. These are capable of immobilizing nutrients such as N in the presence of excess C. Soil formation is also dependent on mixing of surface organic matter down through the horizon. These activities are undertaken by a food web of enormous complexity. Soil

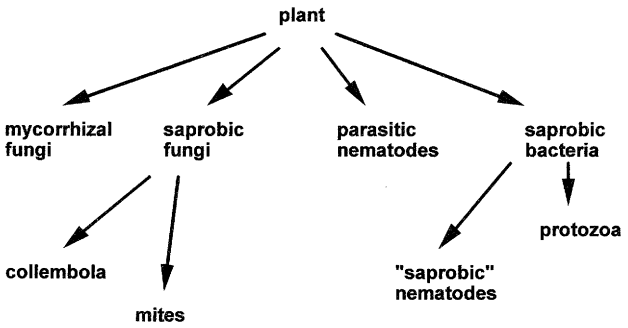


Figure 5 Food web channels. Carbon is allocated from plants initially to four critical microbial functional groups. Parasites, especially nematodes, take a large fraction of the net primary production and largely respire it away. This provides little benefit to soil organic matter and should be discouraged, probably largely by weed reduction and initiating a diverse plant mix. Mycorrhizal fungi provide resources in exchange for their carbon and also appear to be critical for producing stable soil organic matter. Soil bacteria and fungi are two separate webs that are grazed by different animals. While the fungi and bacteria mineralize C, they can simultaneously immobilize critical nutrients. The animals are important to retain the nutrient mineralization process as well as mixing in soil organic matter.

animals generally invade rapidly, either dispersing directly or by moving with soils or other materials. Soil food webs are generally characterized using functional groups as the richness of species is simply too high to characterize in detail. Food web analyses indicate that there are distinct channels (**Figure 5**) that can be affected by the soil conditions and the composition of the microbes. Undertaking detailed studies of the role of biodiversity in these food webs is a critical future task. We currently do not know if species changes really matter to the recovery process.

Symbionts

Symbiotic microorganisms are much less diverse and clearly play critical roles in the establishment and persistence of vascular plants and plant composition. These roles are basically of two types, pathogens that inhibit plant growth, and mutualists that extract resources and exchange those resources with plants for energy or provide protection in some form, again in exchange for energy.

Plant pathogens are of two basic types for our purposes, specialists and generalists. Specialist pathogens are those that are associated with only a single species or group of host species. They tend to be highly diverse. Generalist pathogens tend to be widely spread across plant groups. Specialist pathogens are known to be devastating in agricultural ecosystems. However, they tend to be much less of a problem in restoration efforts. This probably results from the efforts made to restore a diversity of plants, making it more difficult for a pathogen to find a host and build up adequate inoculum densities. Exceptions exist when there is a high prevalence by a single species coupled with an exotic introduction.

Generalist pathogens may be another matter. These are highly diverse organisms that often live as saprobes except when conditions prove favorable to a parasitic lifestyle. For example, there are a wide variety of fusaria and rhizoctonia fungi found in virtually all soils. These can destroy a wide variety of plants under appropriate conditions. *Phytophthora cinnamomi* is responsible for loss of plants ranging such as the eucalyptus in the Jarrah forest in Australia. Often, these are almost undetectable except for very short times. In Wyoming, snow mould reduced sagebrush densities up to 60% and reduced growth in the survivors. This "mould" was a complex mix of fungi, not *Typhula sp.* found in the snow mould diseases of wheat. The disease was opportunistic and only found during El Niño years of high autumn rains and locations of high snowfall accumulation. It was found only one year and only in locations of high snowfall. Plant parasitic nematodes are always found in soils. They are responsible for high levels grazing, but, remarkably, rarely can nematode damage be observed in a restoration project.

Thus, despite the examples where disease was present, there are remarkably few demonstrations where diseases were highly diverse or markedly changed the outcomes of a restoration effort. Even under rather optimum conditions, such as tropical seasonal forests, we have observed few instances of root or shoot disease and then, it tended to be single root tips or individual leaves, but not widespread across a site. This supports the need to establish a diverse plant community. Clearly, it is generally not desirable to restore pathogens to a restoration site.

Mutualistic symbionts are relatively diverse but that diversity may play unique roles in restoration. Probably the best known are N-fixing prokaryotes. Legumes tend to be important early colonizers as N often limits primary production. In croplands, nodulation tends to be highly specific. This led to the all too common practice of using commercial inoculum. However, at sites ranging from glacial outwash in Alaska to Mount St. Helens to a seasonal tropical forest, we have planted or observed invading legumes. In no case were legumes present and functioning nodules absent. We know little about dispersal of rhizobia, but they do appear to be dispersed readily. Further, we now know that host specificity, at least for *Rhizobium*, appears to be largely associated with plasmids and not a nuclear genomic component. Thus, limitation in nodulation is likely not a function of the presence of rhizobia, but of the conditions of the site.

The effectiveness of the nodulation, however, could be an important question. In well-established hot deserts, *Bradyrhizobium* was an efficient bacterium stimulating high rates of N fixation. However, it was slow-growing and deep in the soil profile. *Rhizobium* was fast-growing and found near the surface. It rapidly colonized plants but was not an effective fixer.

In pasture soils, different rhizobia are distributed in patches scattered across a site. Thus, while the presence of rhizobia is likely not limiting to restoration, having a diversity of populations capable of acting with a range of plants under a range of conditions may be important.

We know far less about *Frankia*, although there is a wide diversity of associated plants. In bioassays of respread cold-desert soils, we found that the plants failed to become nodulated. However, the soils had high N concentrations, which may have restricted activity. Alternatively, invading species such as Russian Olive has nodules even in areas where it has previously not been found. Unfortunately, beyond just a few observations of groups and N fixation rates, there are no studies of the diversity of N-fixing species or genotypes in restored areas of these critical groups.

Mycorrhizal fungi have been studied in much greater detail. Their diversity is highly variable. In desert sites, we have found as few as two or three species. Alternatively, in forests, there can be hundreds of species and thousands of genotypes. These fungi are often eliminated by the disturbance event. However, even when they are not completely eliminated, the diversity of species is often radically altered.

Moreover, many species depend on a mycelial network that can extend up to many meters across. This spatial structure is always broken up, providing opportunities for new taxa to invade. The resulting pattern is an increase in the intraspecies diversity with more, smaller clones. As these clones expand, some die and disappear while others continue expanding into the open habitats. Thus, intraspecific diversity initially increases as many propagules arrive and then declines as fewer colonies come to predominate.

Recovery of symbionts is a critical limiting step in restoration. There are two limiting steps: first, invasion of propagules, and second, establishment on site. Invasion is by physical or biological vectors. The most notable physical vector is wind. Wind has been shown to move organisms as large as mycorrhizal fungal spores up to 2 km. However, there are important limitations. Spores larger than 70—100 μm in diameter are rarely wind-dispersed. In those cases, animals are the vectors for microbes. Many animals feed on microbial spores. This can occur directly. For example, the diet of many rodents can be predominantly fungal sporocarps. This was the major means for mycorrhizal recolonization on Mount St. Helens following the eruption. Other propagules are transported unintentionally. Ungulates and rabbits feed on forbs and grasses, but in doing so, they tear plants from the soil, bringing fungal hyphae and internal spores and vesicles. Animals such as gophers preferentially feed on the nodules of legumes in addition to mycorrhizal fungi. The microbes are adapted to pass through the guts and are deposited across restored sites. Thus, just as for plants and animals, a key factor in restoration is the proximity of the source areas (**Figure 6**).

Microbial Establishment

Different microbial species have different abilities to reestablish on a disturbed site. Because of their remarkable diversity, we are unable to artificially return even a small fraction of the microbes necessary for successful restoration. Thus, dispersal from surrounding areas is critical. In all of these invasions, two factors emerge as critical: distance and directionality for the appropriate vectors, and a suitable site. Adjacent source areas are important for reinvasion. At Mount St. Helens, for example, disturbed areas within or adjacent to surviving patches were rapidly recolonized. It took several years for sites at a greater distance to recover (**Figure 6**). This pattern can be found in many other areas. Both physical and biotic vectors travel along specific pathways. These pathways must be present. In areas without appropriate connections allowing vectors (wind flows or animal immigration), the microbes will not invade.

Just because a microbe reaches a site does not mean it establishes. An appropriate host must be present and the soil conditions must be favorable. This can be facilitated in the planting regimes. For example, planting in patches provides windbreaks, increasing harvest of wind-dispersed microbes in addition to organic matter and other seeds. These patches also are lures to animals that are dispersing microbes. Thus, restoration of the microbes on a site often rests on reconstructing a spatially complex structure that provides protection from wind and for animals.

Another mechanism for restoring microorganisms is simply to respread the original, preferably fresh, topsoil. When topsoil is limiting, it can be stored for a short while and respread. However, mining studies have repeatedly demonstrated that as a mine moves, the newly stripped topsoil can be immediately replaced into the newly restored area. This facilitates microbial recovery

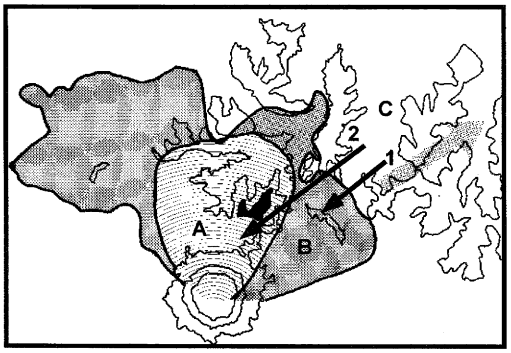


Figure 6 Invasion pathways and source areas. Recovery of microbes on Mount St. Helens can serve as a model of reestablishment. There were three critical types of disturbance. In area A, virtually everything was sterilized. Area B was the blast zone, where everything aboveground was eliminated and ash was deposited up to a meter in depth. Area C was the area of high ashfall. Area C recovered rapidly as most organisms survived in protected areas, establishing on the ash within the first year. This included mycorrhizae, nematodes, nitrogen-fixing bacteria, basidiomycetous fungi, mites, and collembola. Some small patches, where pocket gophers survived and emerged, served as islands. Invasion via animal dispersal also occurred from area C (pathway 1). Recovery across the site was well underway within 2–5 years. Area A was largely invaded from vectors moving from area C (pathway 2). This process was quite slow. Windborne saprophytic fungi (primarily imperfect fungi and bacteria) came in almost immediately. Symbiotic nitrogen-fixing bacteria established in 2 years, from legumes probably dispersed by birds. Other symbiotic microbes took 5–10 years to reestablish.

rapidly. Salvaged transplants would have the same effect of introducing microbes into a disturbed soil.

In circumstances where soil microorganisms are completely lost, it may be necessary to inoculate. We know of no cases where inoculation of bacteria or saprobes has been done directly to facilitate microbial processes, although use of fresh topsoil or salvaged transplants is commonly done. However, mycorrhizal fungal inoculation can improve establishment, particularly of trees and shrubs, in areas where dispersal is limited due to long distances to source areas or where toxic conditions eliminate the native microbes. A few ectomycorrhizal fungi can be cultivated and, in some cases, limited endomycorrhizal inoculum is available through a few companies.

There are no standards for microbes and restoration. Some protocols require the addition of symbiotic mutualists, but all others assume that microbes recover just fine and will "do their jobs." In fact, soil microbes probably never "stabilize." Their short individual life spans, coupled with the ability of some members of each functional group to invade and establish, makes assessment of composition and activity difficult. Inoculation can be an important practice in conditions where little or no inoculum for an entire functional group remains and has little chance for reinvasion. However, restoration requires management of soils, plants, and animals to encourage natural migration, patch structure for concentration of resources, and a complex structure that facilitates spatial and temporal diversity in ecosystem processes. If these conditions are met, it is likely that microbes will be capable of taking care of themselves quite well.

Conclusions

Assessing Restoration Success

These examples of restoration of plants, animals, and soil microorganisms all show the difficulties and limitations of restoration of the entire richness of a prior existing community. While dominant plants and animals may be reintroduced, microorganisms are all expected to recolonize naturally. The resultant lower diversity restored communities indicated that if preservation of biodiversity is the goal, then conservation prior to disturbance is the preferred alternative, rather than restoration after disturbance. In focusing on species richness, we have placed little emphasis on ecosystem functioning, even though restoration of functioning is one of the major goals of restoration. Natural ecosystems provide ecosystem services, such as water supply, oxygen, soil stability, natural products, and so forth for free. Reclamation or rehabilitation is usually sufficient to provide these basic services, without the necessity of reintroducing all of the original biodiversity.

Measurements of both structure and functioning are used to assess restoration success. Restoration success is usually assessed by comparing the restoration site to a reference area, a native site with structure and functioning that are predetermined as the restoration goal. Measurements are compared between the reference and restoration sites. Structural measurements, such as the richness, density, and relative composition of species, are easier to measure than functional measures such as decomposition, nutrient cycling, erosion rate, or biological functions such as species reproduction and mortality or food web energy throughput. Yet it is the functional measurements that we need to determine whether the restored land has really stabilized, not simply the relatively easy measurements that require species counts. Measurements of restoration success are often not legally required, so many restoration/reclamation efforts receive no assessment at all. When they are, a species count, density, or percent cover is often all that is required to declare success.

Designer Ecosystems

Preservation of certain rare species may require manipulation of the ecosystem to stabilize their populations, possibly to the detriment of associated species. Such actions are already taken in numerous situations. For instance, wetland parks for shore birds have been diked, dredged, and dammed to create aquatic habitat for bird species with different water depth requirements. Pastures in Europe receiving high anthropogenic nitrogen deposition are mowed at critical times of the year to reduce the growth of nitrophilous-dominant plants and promote survival of rare plant species. James MacMahon has termed these "designer ecosystems" because they are highly managed ecosystems that have a specific conservation goal, compared to the ecosystem where the species in question may occur naturally. Biodiversity has become highly manipulated in many areas where human populations are dense and where the remnant landscape is managed to promote as high a diversity of species as possible. Virtually all restored communities are missing species, so in one sense restoration may be considered an unintended experiment to determine the impacts of rare or other missing species on community and ecosystem functioning. This will require more research and monitoring than has been done in the past. One aspect of ecological restoration that has not been emphasized in this article is the general lack of data. Many sites are restored that have never received any kind of monitoring or research, or the data are simply not available. The generalization that restoration will not return the original diversity holds for the limited number of sites that have been studied, but as more data become available, we will understand more about how to manipulate ecosystems to maximize diversity.

Acknowledgments

We thank Karen Holl and Zev Naveh for reviewing the manuscript. The research reported here was supported by grants from the Division of Environmental Biology, National Science Foundation, and the National Research Initiative, U.S. Department of Agriculture.

See also: Captive Breeding and Reintroduction. Conservation Biology, Discipline of. Conservation Efforts, Contemporary. Microbial

Biodiversity. Plant Conservation. Reforestation. Restoration of Biodiversity, Overview

References

Allen EB (ed.) (1988) *The Reconstruction of Disturbed Arid Lands*. Boulder, CO.: Westview Press.

Allen EB (ed.) (1997) *Restor. Ecol.* 5: 275–354.

Allen MF (1988) Re-establishment of VA mycorrhizas following severe disturbance: Comparative patch dynamics of a shrub desert and a subalpine volcano. *Proc. R. Soc. Edinburgh, Sect. B: Biol. Sci.* 94: 63–72.

Allen MF (1991) *The Ecology of Mycorrhizae*. Cambridge, UK: Cambridge Univ. Press.

Bowles ML and Whelan CJ (eds.) (1994) *Restoration of Endangered Species: Conceptual Issues, Planning and Implementation*. Cambridge, UK: Cambridge Univ. Press.

Bradshaw AD and Chadwick MJ (1980) *The Restoration of Land*. Berkeley: Univ. of California Press.

Buckley GP (ed.) (1989) *Biological Habitat Reconstruction*. London: Belhaven Press.

Collins HP, Robertson GP, and Klug MJ (eds.) (1995) *The Significance and Regulation of Soil Biodiversity*. Dordrecht/Norwell: Kluwer Academic.

Falk DA, Millar CI, and Olwell M (eds.) (1996) *Restoring Diversity: Strategies for Reintroduction of Endangered Plants*. Washington, D.C: Island Press.

Howe HF and Brown JS (1999) Effects of birds and rodents on synthetic tallgrass communities. *Ecology* 80: 1776–1781.

Jordan WR, Gilpin ME, and Aber JD (eds.) (1987) *Restoration Ecology: A Synthetic Approach to Ecological Research*. Cambridge, UK: Cambridge Univ. Press.

Jordan WR, Peters RL, and Allen EB (1988) Ecological restoration as a strategy for conserving biological diversity. *Environ. Manage.* 12: 55–72.

Major J (1989) *Animals in Primary Succession*. Cambridge, UK: Cambridge Univ. Press.

National Academy of Sciences (1974) *Rehabilitation of Western Coal Lands*. Cambridge, MA: Ballinger.

National Academy of Sciences (1992) *Restoration of Aquatic Ecosystems*. Washington, DC: National Academy Press.

Restoration of Biodiversity, Overview

Joy B Zedler, University of Wisconsin, WI, USA
Roberto Lindig-Cisneros, Universidad National Autónomo de México, México

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Zedler, J.B., Linding-Cisneros, R., Bonilla-Warford, C., Woo, I, volume 5, pp 203–212, © 2001, Elsevier Inc.

Glossary

Biodiversity restoration Attempts to remove abiotic, biotic, and human constraints on any and all native plant, animal, and microbial species.
Exotic species A species occurring beyond its natural range or potential natural dispersal range.
Habitat remnant A small area that is relatively undisturbed and represents a much larger historical ecosystem.
Nurse plant A plant that shelters and facilitates the growth of others.
Propagule Reproductive unit (spore, seed, bulb, cyst, egg, bud, larva, etc.) that gives rise to new individuals.

Introduction

The desire to restore biodiversity stems from a long history of declining species richness. In the USA, 1175 native plant and animal species have become threatened or endangered and 371 native plant communities are rare (Grossman *et al.*, 1994). Major, widespread losses of diverse gene pools, species, and communities of organisms have challenged conservationists to reverse the trend, recognizing that today's conditions are greatly altered from those under which native species formed communities and ecosystems. Efforts to sustain and restore diversity are widespread, involving private landowners, governmental agencies, and nonprofit organizations such as The Nature Conservancy in USA.

Restoration practitioners who seek to restore biodiversity need to understand a restoration site's historical and current conditions in order to predict whether it can support and sustain more taxa. This requires determining environmental constraints as well as biological and management constraints. Here, we focus on factors that constrain biodiversity restoration by considering genetic issues, single-species and multispecies efforts, and landscape-scale approaches. The effectiveness of biodiversity restoration efforts relates to the degree of degradation of the site and to the effort expended in restoration and long-term follow-up (Zedler, 1999). Where damages are extensive, as in restoring mine-spoils and landfills, major efforts are required to reinstate water flows, rebuild soil, and reintroduce soil microorganisms, vegetation, and fauna. Where much of the native ecosystem is still intact but propagule sources are limiting, restoration of diverse plants and animals might only require species reintroductions.

Judging when biodiversity has been restored is complicated, because every natural community has some common species and many uncommon and rare species. There are few explanations for these natural patterns or for differential loss of species, as a natural area becomes degraded. Differential restorability is also puzzling. That is, the question why some diverse communities are readily restored and others are not. Plants and animals have many individual traits that make broad generalization difficult. Thus, practical, region-specific handbooks are needed for restoring biodiversity to sites that differ in degree of degradation and the amount of effort that is feasible.

In a global review of restoration outcomes that were compared to degraded ecosystems (89 projects), Rey Benayas *et al.* (2009) found a 44% increase in biodiversity accompanied by a 25% increase in ecosystem services. However, restoration did not recover all of the biodiversity or all of the combined ecosystem services of those documented for their respective reference ecosystems. It is often the case that efforts to restore species richness produce communities that are less diverse than reference systems. Thus, the first priority for sustaining biodiversity is to prevent the loss of natural habitat, including eradication of invasive species.

The second priority for sustaining biodiversity is to remove the abiotic and biotic constraints on native species populations and communities, working at both local and landscape scales. Such efforts require restorationists to sustain genetic diversity, reintroduce extirpated species, and restore assemblages found in reference ecosystems. In addition, restorationists can make greater use of novel opportunities, such as public rights-of-way or sustainable production areas, to keep native species in their historical landscapes. The restoration of biodiversity at the regional scale will likely need to involve both restoration of historically diverse sites and the use of novel approaches in novel sites. Both efforts benefit from strategic landscape planning and ecosystem-specific guidance for restoring biodiversity.

Genetic Considerations

Populations need genetic variation to respond to environmental stresses and to evolve as the environmental conditions change. Restoration ecology had its beginnings with early research on population genetics when Anthony Bradshaw found that some genotypes of the grass *Agrostis* could tolerate heavy metals while others were unsuitable for revegetating contaminated mine-spoils. Today, ecological restorationists try to select appropriate genotypes for species reintroductions,

choosing propagules from locally adapted populations. In at least one USA case (giant reed, *Phragmites australis)*, they also need to be able to distinguish the native strain from the aggressively invasive exotic strain. Restorationists would aim to eradicate the exotic strain, while retaining the native to aid in restoring brackish wetlands. Genetic composition is thus an early concern in deciding how to restore biodiversity.

Issues

Restoration plans can aim to restock depleted populations, reintroduce extirpated species, or, in special cases, introduce new species to a site (e.g., an endangered species that is native to the region but not the site that becomes available for restoration). The gene pool introduced initially can determine the long-term outcome, *via* founder effects. That is, genetic bottlenecks can develop if populations are restored with low numbers of propagules, ramets come from very few genotypes, or transplants are not locally adapted. For example, inbreeding depression is known to impair sperm production and population fitness of the rare Florida panther (Lacy, 1994). But even if populations are locally adapted through the evolution of coadapted gene complexes, outbreeding and the mixing of alleles can decrease fitness. Alternatively, outbreeding depression might be outweighed by hybridization.

Inbreeding, outbreeding, and hybridization are all considered in selecting sources and numbers of propagules. Practitioners consider both within-donor variability and proximity of the donor site to the reintroduction site. Because clonal plants are easily propagated, they are readily used in revegetation efforts. But collecting a few ramets from a single donor clone to produce thousands of plugs will not restore genetic diversity. In Mission Bay, California, restored eelgrass (*Zostera marina)* was genetically less variable than natural eelgrass in nearby San Diego Bay (Williams and Davis, 1996). In contrast, sowing seeds to Indiana's large Kankakee Sands restoration site produced forb populations that were as genetically variable as remnant populations (Dolan *et al.*, 2007).

Genotypic composition and variation have ramifications beyond individual populations, as shown by plantings of three morphologically distinct genotypes of *Spartina alterniflora* (from three USA states) to a salt marsh restoration site in Delaware (Seliskar *et al.*, 2002). Each genotype expressed unique attributes that determine ecosystem structure and functioning, including differential support of animal populations.

Guidance

Existing guidance for restoring genetic diversity is often species specific, due to differences in dispersal, rates of gene flow, and genetic differentiation (Bowles and Whelan, 1994). Guidance for collecting and translocating propagules includes: (1) Donor site variables, for example, the maximum geographic distance from each donor site to the restoration site; the number of donor populations to tap; which donors to tap, considering how collection might deplete the donor population; and the number of individuals to select from each donor population. (2) Restoration site variables, for example, which microsites offer the greatest potential for establishment; whether to reestablish one large patch or many small; whether to mix genotypes from multiple donors and allow the site to select suitable genotypes. (3) Initial numbers to plant/release; this depends on demography, sex ratio, nonrandom production of offspring, and population fluctuations across generations. A rule of thumb is to establish a population of at least 500 individuals (Lacy, 1994).

Because few species have received adequate study, reintroduction efforts often need research to develop such guidance. In Western Australia, Sinclair and Hobbs (2009) recommended collecting seeds from at least 30 individual shrubs of an outcrossing seeder, *Daviesia divaricata* ssp. *divaricata*, in order to capture its genetic diversity in subsequent plantings.

Captive Breeding and Greenhouse Propagation

Animal populations can be increased through captive breeding before release to the wild, but genetic changes during captivity can compromise survival. For example, captive-bred tamarins that were released in Brazil moved slowly, were confused by scarce food, and unwary of predators, while reconditioned and trained tamarins fared better (Lacy, 1994). Captive populations benefit from continual introduction of wild individuals to help maintain diversity.

Plants that are propagated for multiple generations in warm, humid greenhouses can experience high mortality when transplanted to harsh restoration sites. Hardening (i.e., the gradual shift to harsher conditions) helps acclimate plants that can adjust to outdoor conditions. Like captive breeding, propagation of plants over multiple generations can lead to selection for indoor growth, with loss of traits needed to survive outdoors.

Manipulating Gene Pools

Some authors propose manipulating gene pools to assist evolution and help plants or animals survive extreme conditions. One way is to introduce nonlocal genotypes of the target species. Another approach is to design propagules for extreme restoration sites that are hypersaline, waterlogged, or contaminated. Morphologically varied clones of *Distichlis spicata* were tissue-cultured under various stressors to select for different characteristics, then propagated and transplanted to salt marsh restoration sites. In the field, genotypes differed in canopy structure, belowground production, detritus production, and decomposition rate (Seliskar and Gallagher, 2000). A third suggestion is to create hybrids between local and nonlocal parents (Jones and Monaco, 2009). Manipulated gene pools could enhance survival and overall genetic diversity, although the resulting gene pool would be novel, rather than restored. Several restoration ecologists argue that restoration of novel landscapes requires novel approaches.

Restoring Single Species to Habitat Remnants

The reintroduction of a single species is usually an attempt to reverse its extirpation or restock a native game species, such as

quail in Georgia. Targets also include species that perform valued functions, such as restoring wolf populations to sustain naturally functioning food webs (i.e., control herbivores). Where historical distributions of the target species are uncertain, some plantings or animal releases can be restorations at the regional scale but range expansions at the site scale. Among Southern California's two dozen estuaries and lagoons, only a few still support viable populations of multiple endangered species. Those with continual tidal flushing provide opportunities to establish a grass (*Spartina foliosa*) used for nesting by the federally endangered light-footed clapper rail (*Rallus longirostrus levipes*) (Zedler, 2001). Even if neither species occurred there historically, introductions would enhance biodiversity in both the site and region.

Species have been extirpated for various reasons but especially habitat loss. If a species' habitat no longer exists, reintroduction will be limited to the whole-ecosystem restoration projects (*see* Restoring Communities). If suitable habitat remains, attempts to reintroduce a single species begin with addressing the abiotic and biotic constraints that led to its imperilment.

Abiotic Constraints

Many plant species require specific germination sites, such as open patches with high light penetration. A former habitat that is no longer sufficiently disturbed (e.g., by flood scouring, temporary standing water, treefall/tip-up mounds, mammal burrowing) would lack sufficient open space for plant establishment. Artificial clearings and seed additions can compensate, if seed banks are depleted or dispersal is limiting.

Changes in the frequency, intensity, and timing of fire can reduce rare species populations, and reintroducing fire can be beneficial to ecosystems that burned naturally. Slash pine (*Pinus elliottii*) plantations in Georgia are being restored to long-leaf pine (*Pinus palustris*) forests using controlled burning, which favors an understory wiregrass vegetation and helps sustain the first that facilitates the establishment and persistence of the target tree canopy (Kirkman *et al.*, 2007).

Other abiotic constraints are related to soil conditions. For example, establishment of an annual salt marsh species in California depends on microtopographic conditions, for example, shallow tidal pools that reduce dominance by perennial plants and provide canopy openings (Varty and Zedler, 2008). In USA grasslands, high soil nutrient levels often hinder the restoration of species-rich prairies by helping the invasives outcompete native grasses and forbs.

Biotic Constraints

Habitat loss and competition from nonindigenous species are thought to be the most important causes of the imperilment of rare and endangered species in the US (Wilcove *et al.*, 1998). In California grasslands, exotic annual grasses reduce germination, growth, and survival of native plants, such as the historically dominant perennial grass (*Nasella pulchra*). Efforts to restore target species are also constrained by invasive species, in part because restoration actions typically disturb soils and facilitate seedling establishment.

Some habitat specialists depend on a nurse plant to facilitate germination or establishment. As an example of the latter, many cacti in the Sonoran desert depend on desert ironwood (*Olneya tesota*) to improve the early survival (Suzán *et al.*, 2009). In such cases, restorationists first need to establish populations of the nurse plant.

Parasitic plants need suitable host plants, and the effects of exotics can be quite complex. In San Diego Bay, California, an exotic annual grass (*Parapholis incurva*) of the high-intertidal marsh acted as a pseudohost for a reintroduced population of the endangered hemiparasite (*Cordylanthus maritimus* ssp. *maritimus*). The hemiparasite could attach its haustoria to the grass's roots, but the annual host died before the hemiparasite could mature and set seed (Fellows and Zedler, 2005). In contrast, the natural hosts (perennial plants that provide moisture and nutrients through summer droughts) sustain the parasite through summer, when it produces seed.

Herbivory and predation are further threats. Outbreaks of herbivorous insects can impair reintroductions if vulnerable species are planted too densely. Invasive predators, particularly foxes and cats, challenge reintroduction efforts of many macropod species in Australia (Short *et al.*, 1992). Exotic foxes preyed on a reintroduced marsupial, the quokka (*Setonix brachyurus*), near Perth, Western Australia. And in French Polynesia, the *Partula* snail, a native to Moorea, was driven to extinction after a predatory snail (*Euglandina rosea)* was introduced to control another introduced snail (*Achatina fulica*), which was raised for escargot (Mace *et al.*, 1998).

Animal life histories and trophic levels can help predict potential for reestablishment; for example, late breeders with small clutches are less likely to establish in new surroundings, and carnivores and omnivores are less readily translocated than herbivores (Griffith *et al.*, 1989). Behavioral rigidity is also a predictor; animals that can tolerate human influences flourish in human-dominated landscapes. For example, laboratory-reared and released Mauritius kestrel (*Falco punctatus*) changed their behavior to make use of the agricultural areas and to prey on exotic species, and their numbers increased 10-fold in 25 years (Crade and Jones, 1993).

Management Solutions

Reintroduction efforts are most efficient when the reasons for a species decline or extirpation are known. If unknown, hypothesized causes can be tested experimentally. For example, Pavlic (1996) found that both fire and weed control were needed to reestablish *Amsinckia grandiflora* to northern California grasslands.

Repeated reintroductions are needed when a single trial does not coincide with optimal conditions. In San Diego Bay, sparse pollinators on a small intertidal island limited seed production of *Cordylanthus maritimus* ssp. *maritimus*. After the reintroduction effort was moved to a larger salt marsh where upland patches supported ground-nesting bees, pollination, and seed production improved (Parsons and Zedler, 1997).

Hall (1987, in Primack, 1996) reviewed 15 plant reintroduction projects and provided this advice: (1) Use appropriate planting techniques and match microsite conditions for each species. (2) Select sites carefully; include safe sites.

(3) Document the project in detail. (4) Maintain good growing conditions (eliminate competitors). (5) Monitor over the long term. Griffith *et al.* (1989) reviewed animal reintroduction attempts, focusing on efforts in the United States following passage of the Endangered Species Act. Only 7% of the 93 species of birds and mammals were rare species; most were game species. Resulting guidelines were to: (1) Select quality habitat. (2) Consider life history features. (3) Introduce where competitors are lacking. (4) Use wild-caught animals. (5) Transplant large numbers of individuals.

Restoring Communities

Efforts to restore entire communities often focus on the vegetation, rather than the entire biota (vascular and nonvascular plants, animals, microorganisms) and typically produce assemblages with lower species richness and fewer rare species than that occur in natural communities. As an example, reference forests in southern Mexico supported 30 species of birds, compared to 15 in restored forests (MacGregor-Fors *et al.*, 2010). In a broad survey of restoration outcomes for coastal ecosystems, Borja *et al.* (2010) estimated that full recovery of historical species would require 15–25 years or much longer, even though short-lived species might recover in less than 5 years.

Assembly rules (constraints that a community imposes) are often sought to understand natural successions of species, as well as to guide the sequence of species needed in restoration efforts (Temperton *et al.*, 2004). Tentative rules based on analysis of seminatural grasslands in Germany were to seed annual forbs (generalists) to provide high quality food for herbivores, then add diverse perennial producers and selected herbivores, aiming ultimately to attract additional herbivores and carnivores. A related concept, alternative states, emphasizes that communities shift from their native condition to an altered state, often invaded, once a threshold is exceeded. Both states have internal feedback mechanisms; thus, an invaded state is difficult to restore to the more natural state (Hobbs and Suding, 2009).

Another approach is to plant species judged by experts to have a low coefficient of conservatism (C = ≤5), because those with C = 6–10 are less disturbance-tolerant. Use of species with low C recognizes that restoration involves disturbance, and in fact, the floristic quality (an index based on C values) for Midwestern USA prairies is, on average, lower for remnant than the restored prairies (Packard and Mutel, 1997). As a corollary, species with high habitat specificity (e.g., host-specific insects that rely on rare plants) are more difficult to restore. Panzer *et al.* (1995) identified 256 insects from the Chicago area that were dependent on habitat remnants and absent in more disturbed habitats. An exhaustive study of this same region's vegetation showed that more than two-thirds of the native plant species were remnant-dependent and only one-third had broad distributions. Thus, a minority of species occupied a majority of the landscape.

Restoration sites that are near or adjacent to the remnant ecosystems offer the greatest potential for achieving high diversity. The natural dispersal of epibenthic invertebrates from adjacent salt marshes into restored wetlands rapidly restored diverse assemblages along the North Carolina coast (Sacco *et al.*, 1994) and in San Diego Bay (Scatolini and Zedler, 1996). In both cases, tidal water aided larval dispersal. Still, densities were lower than in reference sites. In the same San Diego Bay wetland, fish species composition matched that of reference sites and densities were higher, apparently because the constructed channels were deeper and broader than natural channels (Williams and Zedler, 1999).

Combining multiple restoration actions with long-term management improves chances of restoring diversity. In Wisconsin, hillside prairies are restored by cutting trees, but burning is necessary thereafter to sustain diverse native grasses and forbs. In New Jersey pine forests, where military exercises eliminated the woody vegetation of dwarf forests (native pygmy pine, *Pinus rigida*), species richness was restored within 2 years using both plantings and mulch (compost to augment soil organic matter and retain moisture; Fimbel and Kuser, 1993). And in Florida, wetlands were restored in phosphate-mining sites within an agricultural landscape by contouring depressions, importing stockpiled sediments, and mulching with substrate from nearby wetlands. These multiple efforts provided dipteran species similar to those in natural wetlands, and the 20 most common species also had similar densities (Streever *et al.*, 1996).

Attempts to restore diverse vegetation rely on species-rich seed mixtures and plantings of seedlings. Animals, however, are usually expected to disperse unaided into restoration sites. Likewise, algal and microbial components are either ignored or limited to the introduction of mycorrhizal fungi where nutrients are a constraint. Reliance on natural dispersal takes longer and favors species with high mobility. Among the most mobile species will be invasive exotics, and those that establish early will hinder restoration of native biodiversity. Plantings are strongly recommended. The constraints on restoring multiple species include abiotic and biotic factors, as for single-species reintroductions.

Abiotic Constraints

The substrate of restoration sites is often inadequate for supporting diverse native plants. In many sites, high nutrient levels reduce plant diversity. In Northumberland, UK, seeding experiments to establish species-rich grasslands showed that fertilizers decreased diversity while grazing increased diversity (Chapman and Younger, 1995). In Saskatchewan, Canada, experiments to control the exotic grasses *Agropyron cristatum* and *Bromus inermis* demonstrated that high nutrient levels, especially nitrogen, favored the weeds and impaired restoration (Wilson and Gerry, 1995).

Insufficient nutrients can also constrain restoration. In the Appalachians, open-pit coal mine reclamation sites are low in nitrogen and phosphorus, because the iron-rich mineral soils weather, oxidize, and form complexes that remove phosphorus (*via* precipitation) from solution (Daniels and Zipper, 1995). In San Diego Bay, dredge spoils limited salt marsh restoration because the sandy substrate had low nitrogen-holding capacity. *Spartina foliosa* required continual additions of nitrogen to produce tall canopies, which were needed to attract light-footed clapper rails (*Rallus longirostris levipes*) to

nest (Zedler, 2001). Soil that is low in phosphorus can also limit plant growth, especially in uplands. There, mycorrhizae can be added to help plants procure nutrients and moisture.

Unfavorable pH levels develop when fen and sulfate-rich soils are exposed to oxygen. In The Netherlands, acidic soils caused the loss of calciphilic species, and liming did not raise pH sufficiently for native species to reestablish, even after 5 years of treatment (van Duren *et al.*, 1998). Also in The Netherlands, atmospheric nitrogen deposition was found to deplete native species and degrade acidic heath. Further restoration efforts involved liming (to raise pH) followed by sod removal (exposing low-nutrient mineral soil); native plant species were then able to reestablish (De Graaf *et al.*, 1998).

Substrate texture can also limit the restoration of animals. Field experiments showed that ground-nesting Caspian terns (*Sterna caspia*) preferred sand to compacted sediments, so a large nesting site near Lake Ontario was constructed with sandy substrate, which attracted the target birds (Quinn and Sirdevan, 1998). Even the specific texture of sand can determine which species of rare tiger beetles (*Cicindela* spp.) can be restored. Both aquatic and terrestrial burrowing animals have strong preferences for substrate texture.

Insufficient topographic heterogeneity limits the variety of microsites necessary for diverse communities (Vivian-Smith, 1997). In salt marshes, tidal creeks provide conduits for fish and invertebrates to access shallow pools on the marsh plain, and pools develop into feeding oases. In restoring desert wasteland, the excavation of depressions and mounding the spoils to funnel water allowed water, organic matter, and nutrients to collect (Boeken and Shachak, 1994). Such depressions support more plant species.

At the landscape scale, loss and fragmentation of habitats have reduced connectivity among the remaining patches. Both dispersal and recruitment limited species abundances in experimental additions of up to 54 species to 2×2 m grassland plots; species richness increased 83% following one-time sowing of seeds (Tilman, 1997). Processes that overcome natural dispersal barriers can allow novel species to invade habitats and consequently cause marked changes in community composition, diversity, and functioning.

Biotic Constraints

Propagules are often limiting to restoration efforts. Efforts to restore prairie pothole wetlands in North America by eliminating drainage structures (tiles, ditches) were constrained by insufficient seed banks and limited dispersal. After 3 years, restored wetlands averaged 27 species compared to 46 in reference wetlands (Galatowitsch and van der Valk, 1996). On the other hand, propagules of nonindigenous species are often abundant in restoration sites, especially small sites with a high proportion of edge (e.g., many prairie and savanna restoration sites).

In mesic to dry ecosystems, insufficient mycorrhizae can slow restoration by reducing nutrient uptake by target plants. Adding one fungal inoculum might not suffice, since a diversity of mycorrhizal species led greater vascular plant species richness in USA prairies, European grasslands, and Canadian old-field restoration (van der Heijden *et al.*, 1998).

Exotic invaders are especially problematic in restoration sites. Invasions prompt restoration efforts, and restoration efforts often create conditions that favor weedy species. In a semidesert area of Arizona, attempts to establish native grasses in mesquite (*Prosopis juliflora* var. *velutina*) woodland were impaired by the dominance of the understory by an exotic grass, *Eragrostis lehmannian*. Reintroduced natives were sparse relative to the exotic plant, depending on the interaction of several factors, such as fire (a treatment variable) and rain intensity (introductions were done in three successive years) (Biedenbender and Roundy, 1996).

Plant invasions are especially problematic in wetlands. Perennial clonal forbs (e.g., purple loosestrife, *Lythrum salicaria)* and grasses (e.g., *Phragmites australis, Phalaris arundinacea*) invade and dominate many types of wetlands. In microcosms with varying levels of water, nutrients, and leaf litter cover, *Bidens cernua*, a native annual, dominated all 24 treatments in the first year. After 5 years, most of the microcosms were becoming dominated by purple loosestrife, with only a few of the original species remaining. Flooding regimes and soil fertility were the main factors influencing establishment and growth of *Lythrum*, with high water and low fertility limiting its dominance (Weiher *et al.*, 1996).

Exotic animals that constrain restoration in San Diego Bay wetlands include sessile invertebrates (e.g., mussels, *Musculista senhousii)* that usurp space on various substrates and fish (e.g., the carnivorous goby, *Acanthogobius flavimanus*) that prey on native fishes (Williams *et al.*, 1998). In some cases, creating a novel ecosystem in a region (e.g., ponds with permanent water in Oregon) provides habitat for exotic predators (e.g., bullfrogs, *Lithobates catesbeiana*), which prey on native species. In other cases, the presence of one exotic species facilitates the arrival of others, resulting in biodiversity loss, not gain. Simberloff (1999) calls this process an invasional meltdown.

Management Solutions

Where restoration begins by removing invasive plants, the characteristics of the invader dictate site preparation; for example, in Midwestern USA grasslands, *Bromus inermis* monotypes are plowed before native species are planted, but *Phalaris arundinacea* needs repeated herbiciding to kill belowground dormant buds. Restoration often needs to be delayed until herbicides are no longer needed. Fire is commonly reintroduced to combat exotic woody plants. Where habitats are small (e.g., remnant prairies along railroad rights-of-way) or where controlled burns are a hazard to people or structures, the inability to reintroduce fire constrains biodiversity restoration.

The more harsh the conditions, the more effort is needed to establish vegetation. Planting seedlings in clusters can facilitate the accumulation of organic matter, soil water, and mycorrhizae (MacMahon, 1998). Planting seeds in safe sites aids germination and revegetation, especially in severe environments. In alpine meadows, disturbed areas around ski runs are ameliorated by covering seeds with fibrous mats (Urbanska, 1995). Mulch is effective where it does not blow or wash away. Industrial byproducts can be used as organic soil amendments, potentially improving soil water-holding

capacity. In San Diego, kelp byproducts are composted with perlite to provide a soil conditioner that accelerated restoration of salt marsh vegetation (O'Brien and Zedler, 2006). Temporary irrigation systems are also used to water seeds and plugs, often for multiple years, until deep roots form and sustain drought-tolerant plants.

Some species need to be reestablished during brief opportunities, such as seasonal rainfall, or following unusual events, such as forest fires, flooding, or winds that uproot trees and expose soil (Primack, 1996). Following bauxite mining in southern Australia, the ripping and cultivation of substrate and sowing needs to be timed to coincide with seasonal rainfall to improve restoration of *Eucalyptus marginata* forests. In Southern California, restoring tidal influence to salt marshes in December or January allows high tides to disperse seeds to newly excavated areas and winter rainfall to assist seedling establishment. The sequence of planting can also be important (Eriksson and Eriksson, 1998). Introducing nurse plants aided restoration of limestone grasslands of northern Switzerland; *Arabis hirsuta* and *Primula veris* established best in the shelter of neighboring plants; elsewhere, seedlings succumbed to drought and substrate-heaving (Ryser, 1993). In a Mexican forest restoration project, the addition of a legume as a nurse plant benefited the restoration of birds by providing a food source (MacGregor-Fors *et al.*, 2010).

Restoring Biologically Diverse Landscapes

Occasionally, the restoration of diversity to large landscapes is undertaken with strategic planning that matches species needs to opportunities, including willing landowners and funding for public improvements. The Southern California Coastal Wetland Recovery Project (www.scwrp.org) is an example. This regional effort involves state agency personnel who aim to reverse habitat loss, remove threats to endangered species, and mitigate impacts of highway, ports, and other developments. Extensive involvement of stakeholders and years of planning have led to several projects that have restored tidal flows to nontidal and semitidal lagoons.

Several watershed planning efforts are also underway, especially within the context of wetland mitigation. Efforts in Wisconsin, Tennessee, and Georgia combine wetland restoration (the focus of the Environmental Law Institute) with biodiversity restoration (the focus of The Nature Conservancy). These nongovernmental organizations are working with multiple stakeholders to develop watershed plans to serve as national models for prioritizing wetland restoration sites and targets to improve water quality, reduce flooding, and sustain diversity within watersheds and downstream waters.

Landscape Objectives

At the landscape (or watershed) scale (several square kilometers), the restoration of biodiversity often has multiple functional goals of value to people (ecosystem services). Such services include managing water quality, treating waste, providing social and cultural functions (visual quality, education, heritage), and offering products (e.g., producing fiber or food). Multifunctionality is facilitated by spatial and temporal heterogeneity of the restored landscape, with feedbacks from biodiversity. Spatial heterogeneity fosters fluxes between habitats, builds ecological "networks," and fosters species interactions (e.g., a facilitator of a desired species or a predator that feeds on a pest species). Collectively, heterogeneity enhances a landscape's chances of supporting and sustaining biodiversity. To these three landscape-restoration objectives, we add adaptive restoration or "learning while restoring." For all four objectives, long-term study and monitoring are needed to understand and quantify outcomes and thus improve biodiversity restoration.

The ability of a landscape to support increased biodiversity is greater if it includes many large, appropriately located restoration sites. Similarly, restoring biodiversity at a specific site is facilitated where the surrounding landscape has many large natural habitats that still support the target community. Along Delaware Bay, 5040 hectares of tidal wetlands were restored, enhanced, and preserved to mitigate impacts to fish by cooling water discharged by a major power company (Weinstein *et al.*, 2001). Selfrestoration was facilitated by breaching dikes around marsh hay fields and herbiciding invasive *Phragmites* in degraded brackish marshes. Tidal water could then transport propagules of microorganisms, algae, plants, and animals from the adjacent 32,000 ha of fully tidal wetlands, allowing the native grass, *Spartina alterniflora*, and diverse aquatic animals to reestablish. Proximity to reference sites and tidal dispersal were keys to selfrestoration.

Novel Opportunities

Most watersheds are humanized and unsuitable for large-scale restoration. Still, it is possible to restore and retain many of the species that were historically widespread, although not necessarily exactly where they once occurred. Novel sites could be suitable places to introduce native plants that once occurred within the region. Using novel sites for native vegetation would not necessarily restore native assemblages, but it would help sustain some gene pools, augment some populations, and reduce the area of nonnative vegetation, potentially reducing exotic species invasions.

Many European farmers have been given incentives to plant wildflowers to restore biodiversity to arable lands. In Switzerland, landscapes with multiple wildflower strips supported greater diversity of butterflies (and more specialist species) where the area of strips was greater and plant species were more diverse (Aviron *et al.*, 2010). Near Decorah, Iowa, remnant prairie vegetation between a highway and a railroad track was expanded by plantings. In Wisconsin, many private landowners use native plants for landscaping, and children of several dozen schools have restored prairies that teachers use in nature instruction. Presumably, these regions benefit from the cumulative effects of many small sources of propagules.

Plantations differ from natural forests in being monocultures, but they can be managed to provide a diversity of understory plants. In fir (*Abies*) and pine (*Pinus*) forests in western México, seeding of lupine (*Lupinus elegans*) in reforestation sites created conditions that favored the establishment of native forbs and shrubs, considerably increasing

forest species richness (MacGregor-Fors *et al.*, 2010). In New Zealand, indigenous plants and animals occurred in forest plantations that were close to natural habitat (Norton, 1998). In northern Australia, plantations of native *Flindersia brayleyana* and *Araucaria cunninghamii* supported more diverse understory vegetation than did the exotic *Pinus caribea* (Keenan *et al.*, 1997). Planting diverse trees (six conifers and four hardwoods) in Minnesota reduced infection of fungi (*Armillaria* spp.) and increased wood production (Gerlach *et al.*, 1998).

Power line rights-of-way provide novel opportunities to establish habitat for wildlife. Such rights-of-way are typically cleared of trees and shrubs by cutting and use of herbicide, which reduces native plant and wildlife diversity and creates corridors for species invasions. Brown (1995) suggested using competitive cover crops to inhibit the establishment of trees. *Dactylis glomerata* (a nonindigenous grass) reduced survival of two tree species (*Fraxinus pennsylvanica* and *Acer saccharum*), but biomass and diversity of forbs were also reduced. Field experimentation could identify additional suitable native species.

Stormwater management ponds could support greater biodiversity if topographic heterogeneity were increased and shorelines were more convoluted. In Wisconsin, such ponds attract many invasive aquatic species (Johnson *et al.*, 2008). Thus, shoreline plants and aquatic animals need to be planted. Stormwater conveyances can employ native willows (*Salix* spp.) to stabilize banks and floodplains in preference to concrete flumes. Frequently disturbed sites could reduce physical and chemical maintenance and attract wildlife.

Biofuel production also involves frequent disturbance, for example, annual harvesting of grass crops. Biodiversity would be enhanced if native species and locally adapted genotypes were used, if biofuels were grown using diversity-friendly agricultural practices (e.g., avoiding mowing during bird-nesting times), and if the ecological footprint were minimized (low land area per biomass) (Groom *et al.*, 2008). Frequent mowing and removal of all aboveground biomass would likely select for a few grass species and limited animal diversity. Landscape-scale guidelines are needed to phase plantings and harvests and to combine herbaceous and woody biofuel crops to support more biodiversity than in widespread monocultures.

There are concerns that the species that are most adaptable to novel restoration sites will dominate landscapes and restoration efforts will result in "generic communities" (Zedler, 1999). Because habitat restoration involves disturbances, such as contouring and grading, the species that thrive in restored habitats are likely to be adaptable, disturbance-tolerant generalists. While such species have utility in establishing cover quickly and in resisting invasion by less desirable species, biodiversity will not be restored without further efforts to introduce and attract the full spectrum of native plants and animals. Continual experimentation (e.g., seed additions to grassland, as in Tilman, 1997) can serve this purpose while learning more about the abiotic and biotic constraints that limit restoration of various species and assemblages

Large restoration sites are amenable to experimentation. Most of the constraints on biodiversity restoration can be tested experimentally, for example, varying genetic donors and translocation procedures, varying microsites for single-species reintroductions, and varying assemblages and sequential introductions in multispecies approaches. Restoration sites are underused by ecologists, but, when approached within an adaptive framework, they afford many opportunities to "learn while restoring." Adaptive restoration begins by identifying the unknowns that need to be known to restore landscapes or sites. It proceeds in phases, with sequential field experiments that answer the most urgent questions first. Researchers or managers monitor outcomes and then use the results to select and implement the most promising approaches across larger areas. At Tijuana Estuary, California, the first experiment determined that five of eight native halophytes needed to be transplanted as seedlings (the other three could selfrestore, given a seed source). A subsequent, larger experiment used that guidance and also determined that amending the soil, planting in tight clusters, and adding tidal creek networks could increase plant growth and overall ecosystem functioning. Adaptive restoration led to a handbook on how to restore the region's tidal wetlands (Zedler, 2001).

See also: Agrobiodiversity. Conservation Efforts, Contemporary. Genetic Diversity. Introduced Plants, Negative Effects of. Metapopulations

References

Aviron S, Herzog F, Klaus I, Schupbach B, and Jeanneret P (2010) Effects of wildflower strip quality, quantity, and connectivity on butterfly diversity in a Swiss arable landscape. *Restoration Ecology Early View* 19: 500–508.

Biedenbender SH and Roundy BA (1996) Establishment of semidesert grasses into existing stands of Eragriostis lehmanniana in southern Arizona. *Restoration Ecology* 4: 155–162.

Boeken B and Shachak M (1994) Desert plant communities in human-made patches – Implications for management. *Ecological Applications* 4: 702–716.

Borja A, Dauer D, Elliott M, and Simenstad C (2010) Medium and long-term recovery of estuarine and coastal ecosystems: Patterns, rates and restoration effectiveness. *Estuarine and Coasts* 33: 1249–1260.

Bowles ML and Whelan CJ (eds.) (1994) *Restoration of Endangered Species*. Cambridge, UK: Cambridge University Press.

Brown D (1995) The impact of species introduced to control tree invasion on the vegetation of an electrical utility right-of-way. *Canadian Journal of Botany* 73: 1217–1228.

Chapman R and Younger A (1995) The establishment of a species-rich grassland on a reclaimed opencast coal site. *Restoration Ecology* 3: 39–50.

Crade TJ and Jones CG (1993) Progress in restoration of the mauritius kestrel. *Conservation Biology* 7: 169–175.

Daniels W and Zipper C (1995) Improving coal surface mine reclamation in the Central Appalachian Region. In: Cairns Jr. J (ed.) *Rehabilitating Damaged Ecosystems*, pp. 187–217. Boca Raton, FL: Lewis Publishers.

De Graaf MCC, Verbeek PJM, Bobbink R, and Roelofs JGM (1998) Restoration of species-rich dry heaths: The importance of appropriate soil conditions. *Acta Botanica Neerlandica* 47: 89–111.

van Duren IC, Strykstra RJ, Grootjans AP, ter Heerdt GNJ, and Pegtel DM (1998) A multidisciplinary evaluation of restoration measures in a degraded Cirsio-molinetum fen meadow. *Applied Vegetation Science* 1: 115–130.

Eriksson O and Eriksson A (1998) Effects of arrival order and seed size on germination of grassland plants: Are there assembly rules during recruitment? *Ecological Research* 13: 229–239.

Fellows M and Zedler JB (2005) Effects of the non-native grass, Parapholis incurva (Poaceae), on the rare and endangered hemiparasite, Cordylanthus maritimus subsp. maritimus (Scrophulariaceae). *Madroño* 52: 91–98.
Fimbel RA and Kuser JE (1993) Restoring the pygmy pine forests of New Jersey's Pine Barrens. *Restoration Ecology* 2: 117–129.
Galatowitsch SM and van der Valk AG (1996) The vegetation of restored and natural prairie wetlands. *Ecology and Applications* 6: 102–112.
Gerlach JP, Reich PB, Puettmann K, and Baker T (1998) Species, diversity, and density affect tree seedling mortality from Armillaria root rot. *Canadian Journal of Forest Research* 27: 1509–1512.
Griffith B, Scott JM, Carpenter JW, and Reed C (1989) Translocation as a species conservation tool: Status and strategy. *Science* 245: 477–480.
Groom M, Gray EM, and Townsend PA (2008) Biofuels and biodiversity: Principles for creating better policies for biofuel production. *Conservation Biology* 22: 602–609.
Grossman DH, Goodin K, and Reuss C (1994) *Rare Plant Communities of the Conterminous United States: An Initial Survey*. Arlington, VA: The Nature Conservancy.
van der Heijden MGA, Klironomos JN, Ursic M, *et al.* (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 36: 69–72.
Johnson PT, Olden JD, and Vander Zanden MJ (2008) Dam invaders: Impoundments facilitate biological invasions in freshwaters. *Frontiers in Ecology and the Environment* 6: 357–363.
Jones TA and Monaco TA (2009) A role for assisted evolution in designing native plant materials for domesticated landscapes. *Frontiers in Ecology and the Environment* 7: 541–547.
Keenan R, Lamb D, Woldring O, Irvine T, and Jensen R (1997) Restoration of plant biodiversity beneath tropical tree plantations in Northern Australia. *Forest Ecology and Management* 99: 117–131.
Kirkman K, Mitchell RJ, Kaeser MJ, Pecot SD, and Coffey KL (2007) The perpetual forest: Using undesirable species to bridge restoration. *Journal of Applied Ecology* 44: 604–614.
Lacy RC (1994) Managing genetic diversity in captive populations of animals. In: Bowles ML and Whelan CJ (eds.) *Restoration of Endangered Species*, pp. 63–89. Cambridge, UK: Cambridge University Press.
Mace GM, Pearce-Kelly P and Clarke D (1998). An integrated conservation programme for the tree snails (Partulidae) of polynesia: A review of captive and wild elements. *Journal of Conchology* (supplement 2): 89–96.
MacGregor-Fors I, Blanco-García A, and Lindig-Cisneros R (2010) Bird community responsesto differentforest restoratinefforts. *Ecological Engineering* 36: 1492–1496.
MacMahon JA (1998) Ecology as a basis for restoration. In: Pace ML and Groffman PM (eds.) *Successes, Limitations and Frontiers in Ecosystem Science*. New York: Springer-Verlag.
Norton DA (1998) Indigenous biodiversity conservation and plantation forestry: Options for the future. *New Zealand Forestry* 43: 34–39.
O'Brien E and Zedler JB (2006) Accelerating the restoration of vegetation in a southern California salt marsh. *Wetlands Ecology and Management* 14: 269–286.
Packard S and Mutel, CF (eds.) (1997) *The Tallgrass Restoration Handbook*. Washington, DC: Island Press.
Panzer R, Stillwaugh D, Gnaedinger R, and Derkovitz G (1995) Prevalence of remnant dependence among the prairie and savanna-inhabiting insects of the Chicago region. *Natural Areas Journal* 15: 101–116.
Parsons L and Zedler JB (1997) Factors affecting reestablishment of an endangered annual plant at a California salt marsh. *Ecological Applications* 7: 253–267.
Pavlik B (1996) Defining and measuring success. In: Falk DA, Millar CI, and Olwell M (eds.) *Restoring Diversity: Strategies for Reintroduction of Endangered Plants*, pp. 127–155. Washington, DC: Island Press.
Primack RB (1996) Lessons from ecological theory: Dispersal, establishment, and population structure. In: Falk DA, Millar CI, and Olwell M (eds.) *Restoring Diversity: Strategies for Reintroduction of Endangered Plants*. Washington, DC: Island Press.
Quinn JS and Sirdevan J (1998) Experimental measurement of nesting substrate preference in Caspian terns, Sterna caspia, and the successful colonisation of human constructed islands. *Biological Conservation* 85: 63–68.
Rey Benayas JMR, Newton AC, Diaz A, and Bullock JM (2009) Enhancement of biodiversity and ecosystem services by ecological restoration: A meta-analysis. *Science* 325: 1121–1124.
Ryser P (1993) Influences of neighbouring plants on seedling establishment in limestone grassland. *Jounal of Vegetaion Science* 4: 195–202.
Sacco JN, Seneca ED, and Wentworth TR (1994) Infaunal community development of artificially established salt marshes in North Carolina. *Estuaries* 17: 489–500.
Scatolini SR and Zedler JB (1996) Epibenthic invertebrates of natural and constructed marshes of San Diego Bay. *Wetlands* 16: 24–37.
Seliskar DM and Gallagher JL (2000) Exploiting wild population diversity and somaclonal variation in the salt marsh grass Distichlis spicata (Poaceae) for marsh creation and restoration. *American Journal of Botany* 8: 141–146.
Seliskar DM, Gallagher JL, Burdick DM, and Mutz LM (2002) The regulation of ecosystem functions by ecotypic variation in the dominant plant: A Spartina alterniflora salt marsh case study. *Journal of Ecology* 90: 1–11.
Short J, Bradshaw SD, Giles J, Prince RIT, and Wilson GR (1992) Reintroduction of macropods (Marsupalia: Macropodoidea) in Australia: A review. *Biologcal Conservation* 62: 189–204.
Simberloff D and Von Holle B (1999) Positiver interactions of nonindigenous species: Invasional meltdown? *Biological Invasions* 1: 21–32.
Sinclair EA and Hobbs RJ (2009) Sample size effects on estimates of population genetic structure: Implications for ecological restoration. *Restoration Ecology* 17: 837–844.
Streever WJ, Portier KM, and Crisman TL (1996) A comparison of dipterans from ten created and ten natural wetlands. *Wetlands* 16: 416–428.
Suzán H, Nabhan GP, and Patten DT (1996) The importance of Olneya tesota as a nurse plant in the Sonoran Desert. *Journal of Vegetation Science* 7: 635–644.
Temperton VM, Hobbs RJ, Nuttle T, and Halle S (eds.) (2004) *Assembly Rules and Restoration Ecology*. Washington, DC: Island Press.
Tilman D (1997) Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* 78: 81–92.
Urbanska KM (1995) Biodiversity assessment in ecological restoration above the timberline. *Biodiversity and Conservation* 4: 679–695.
Varty A and Zedler JB (2008) How waterlogged microsites help an annual plant persist among salt marsh perennials. *Estuaries and Coasts* 31: 300–312.
Vivian-Smith G (1997) Microtopographic heterogeneity and floristic diversity in experimental wetland communities. *Journal of Ecology* 85: 71–82.
Weiher E, Wisheu IC, Keddy PA, and Moore DR (1996) Establishment, persistence and management implications of experimental wetland plant communities. *Wetlands* 16: 208–218.
Weinstein MP, Teal JM, Balletto JH, and Strait KA (2001) Restoration principles emerging from one of the world's largest tidal marsh restoration projects. *Wetlands Ecology and Management* 9: 387–407.
Wheeler BD, Shaw SC, Fojt WJ, and Robertson RA (eds.) (1995) *Restoration of Temperate Wetlands*. Chichester: Wiley.
Wilcove DS, Rothstein D, Dubow J, Phillips A, and Losos E (1998) Quantifying threats to imperiled species in the United States. *BioScience* 48: 607–615.
Williams GD, Desmond JS, and Zedler JB (1998) Extension of two nonindigenous fishes, Acanthogobius flavimanus and Poecilia latipinna, into San Diego Bay marsh habitats. *California Fish and Game* 84(1): 1–17.
Williams GD and Zedler JB (1999) Fish assemblage composition in constructed and natural tidal marshes of San Diego Bay: Relative influence of channel morphology and restoration history. *Estuaries* 22: 702–716.
Williams SL and Davis CA (1996) Population genetic analysis of transplanted eelgrass (Zostera marina) beds reveals reduced genetic diversity in Southern California. *Restoration Ecology* 4: 163–180.
Wilson SD and Gerry AK (1995) Strategies for mixed-grass prairie restoration: Herbicide, tilling, and nitrogen manipulation. *Restoration Ecology* 3: 290–298.
Zedler JB (1999) The ecological restoration spectrum. In: Streever W (ed.) *An International Perspective on Wetland Rehabilitation*. Dordrecht/Norwell, MA: Kluwer Academic.
Zedler JB (ed.) (2001) *Handbook for Restoring Tidal Wetlands*. Boca Raton, Florida: CRC Press.

Riparian Landscapes

Robert J Naiman, University of Washington, WA, USA
Henri Décamps, Université Paul Sabatier, Toulouse, France
Michael E McClain, UNESCO-Institute for Water Education (IHE), The Netherlands

© 2013 Elsevier Inc. All rights reserved.

Introduction

Transitional semiterrestrial/semiaquatic areas regularly influenced by fresh water, and usually extending from the edges of water bodies to the edges of upland communities, are often referred to as riparian zones or floodplains (Malanson, 1993; Naiman *et al.*, 2005). Owing to their spatial position, they integrate interactions between the landscape's aquatic and terrestrial components. They are dynamic environments characterized by variable energy regimes, substantial habitat heterogeneity, a diversity of ecological processes, and multidimensional biophysical gradients. They are often locations of concentrated biodiversity at landscape, regional, and continental scales (Naiman and Décamps, 1997).

Riparian zones associated with running waters are three-dimensional biophysical structures. They are set in complex ecological and cultural matrices – from headwaters to the sea – that have both surface and subsurface components that are changing over time. Expansion and contraction of riparian zones along rivers occur in relation to precipitation, river flow, and geomorphology. One segment of the channel may be fed largely by upwelling groundwater, whereas at other locations surface runoff may penetrate into bed sediments (alluvium) accumulated over millennia by cut-and-fill alluviation. During flooding, when surface flow may recharge groundwater aquifers and spill out over the floodplains, eroding or depositing sediment in accordance with the energy dynamics of water interacting with geomorphic features, riparian characteristics aid in sediment deposition and the dissipation of erosive energy. During dry periods, when flow in the channel may be maintained only by alluvial and karstic aquifers, riparian characteristics quickly adjust to the environmental regime. Thus, rivers may be viewed as a collection of dynamic multidimensional pathways along which aquatic–terrestrial linkages vary spatially and temporally (Rood *et al.*, 2003). Anthropogenic influences contribute greatly to this variation, as river valleys have been foci for human settlements and commerce for millennia.

Fires, drought, flooding, landslides, wind throw, grazing, and other natural disturbances, coupled with human interventions such as logging, urbanization, farming, and damming, alter vegetative patterns and soil–plant nutrient exchange at a variety of scales. This has direct consequences for ecological processes in riparian zones – such as biodiversity, productivity, sediment transport, and live and dead wood recruitment.

In an increasingly human-dominated world, the biodiversity riparian zones must be viewed in a landscape context – that is, as natural-cultural systems (Décamps, 1996; Arthington *et al.*, 2010). Although surface and subsurface patterns and processes act as key drivers for sustaining riparian goods and services, it is human perceptions and cultural representations of landscapes that shape the biodiversity of contemporary riparian zones. Here, we focus on the biodiversity and unique ecological functions of riparian zones and how these functions are linked to dynamic biophysical processes and interactions across multiple spatial and temporal scales along natural rivers.

Riparian Zones as Focal Points for Biodiversity

River corridors normally possess highly diverse floral and faunal communities and attendant biological processes. These attributes emerge from the unique spatial organization – a mosaic of habitats continually changing in response to variable water flows above and below ground – and biotic responses to the locally variable topography and climate. The net effect is substantial small-scale heterogeneity. Further, the inherent heterogeneity introduced by local topography affects the frequency and duration of inundation, and the local microclimate influences the ability of individual species to flourish, thereby adding even more physical as well as genetic heterogeneity. Collectively, heterogeneity at local scales offers a great variety of conditions for life, making riparian corridors focal points for biodiversity. Riparian diversity is further magnified at the catchment scale since riparian corridors extend from the highest to the lowest elevations above sea level, and the higher elevation riparian zones are numerous, which tends to augment regional diversity.

Studies throughout the world indicate that riparian areas generally have high levels of both of plant and animal diversity. In the Pacific Northwest of the US, 74% of all plant species within one catchment are found in the riparian corridor, which corresponds to only ~3–8% of the catchment depending on topography. According to other reports, all periodically flooded forests in the Amazon basin may have about 20% of the 4000 to 5000 estimated Amazonian tree species, and about 1400 vascular plant species occur along the Adour River riparian corridor in France, representing 30% of the French flora. However, this information may be slightly misleading because comparative studies are seldom undertaken in surrounding uplands. It has been suggested that riparian zones increase regional diversity >50% by harboring different species assemblages, rather than more species, when compared to surrounding uplands.

Many riparian corridors vary in species richness along the river's course, but some do not. There appear to be three types of riverine reaches that fundamentally differ in terms of biodiversity: Naturally constrained river reaches increase in biodiversity downstream, whereas braided reaches have relatively low diversity and meandering reaches have high biodiversity. In general, longitudinal studies show that plant diversity generally increases downstream, with the peak reached in the piedmont (transition zone between mountain and lowland domains) as the riparian zone widens, or at major tributary

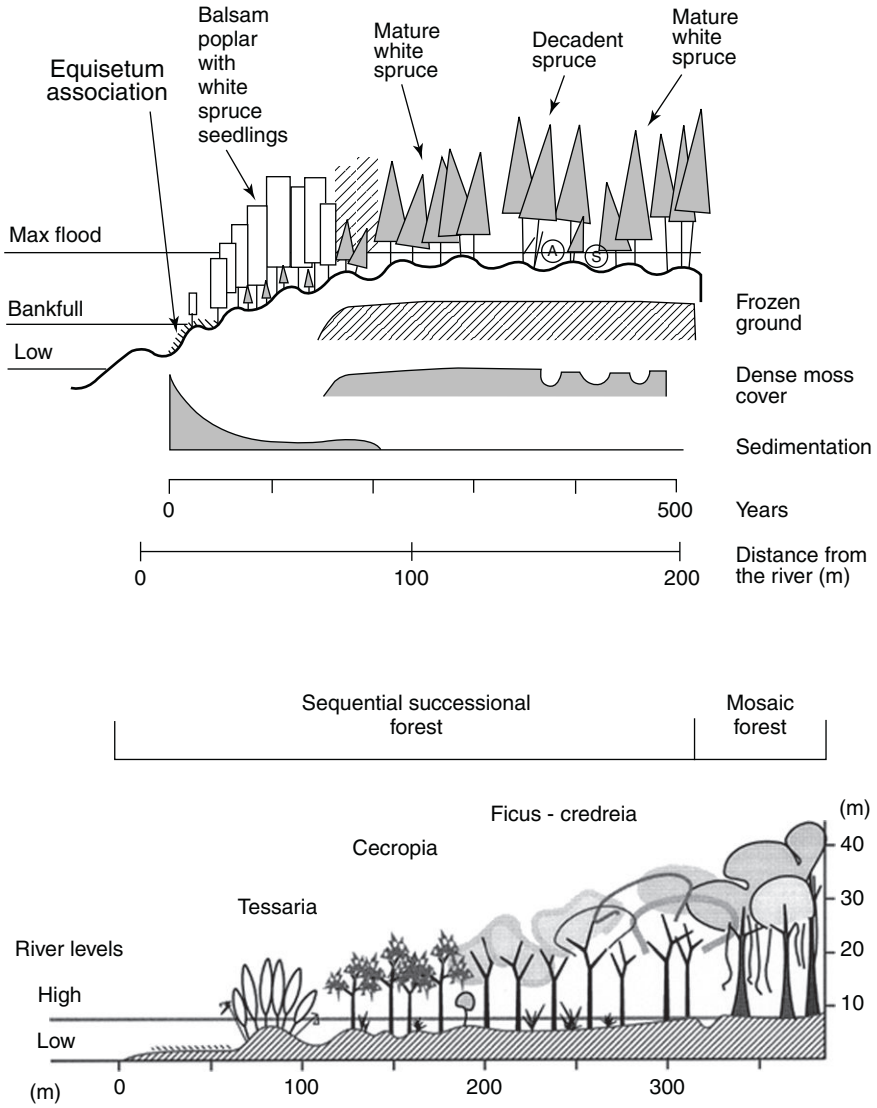


Figure 1 Floodplain profiles of two rivers in natural conditions: The river Beaton, British Columbia, Canada, above (after Nanson and Beach, 1977) and the Río Manu, Peru in the upper Amazon basin, below (after Salo *et al.*, 1986). Such relationships between landforms and forest succession characterize riparian forests along rivers under natural conditions all over the world (Décamps, 1996).

deltas. Additionally, natural gradients in nutrient subsidies between other ecosystems and riparian zones can affect biodiversity. In coastal British Columbia and Alaska nutrient subsidies from spawning Pacific salmon (*Oncorhynchus* sp.) increase nutrient loading to riparian soils shifting plant communities toward nutrient-rich species, which in turn decreases plant diversity (Hocking and Reynolds, 2011). These effects are mediated by interactions between salmon density and the physical characteristics of watersheds.

These catchment-scale patterns suggest a maximum diversity at an intermediate level of natural disturbance, which induces considerable spatial heterogeneity. However, other physical attributes (e.g., channel gradient, lithology, level of confinement) and local climate also modify site-specific species richness and can result in no net trend in longitudinal species richness. Lateral patterns in plant diversity also generally peak at intermediate levels of natural disturbance and moisture availability. Diversity is often low immediately adjacent to the river, increasing with local elevation and geomorphic complexity, and then declining slightly upslope of the riparian-upland interface, although there are variations to this generalization (Van Pelt *et al.*, 2006). Factors controlling species expression are also keys to understanding riparian plant diversity: Flood frequency, site productivity, and spatial complexity have been related to plant species richness (**Figure 1**). In contrast, human induced disturbances like deforestation and farming tend to reduce both levels of biodiversity and the processes promoting diversification.

Overall, riparian soils appear to have an impressive diversity. There may be as many as 10,000 species of ectomycorrhizal fungi, >3000 species of bacteria, >5000 species of nematodes, and tens of thousands of species of mesofauna (mites, collembolans) and macrofauna (ants, termites, earthworms). Aboveground faunal diversity is better known but

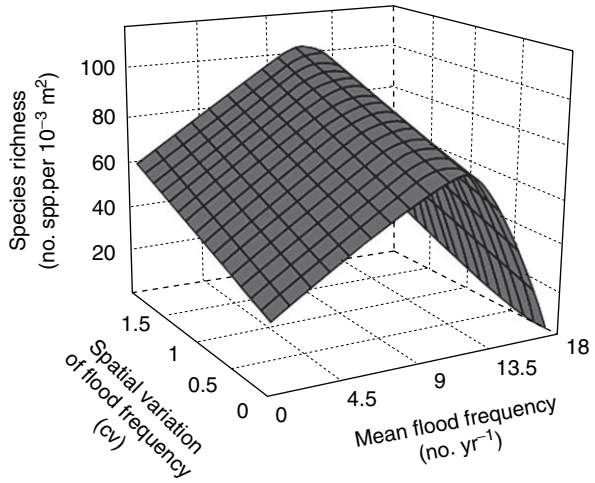


Figure 2 Three-dimensional representation showing relations among species richness, flood frequency, and topographic variation for riparian areas in southeast Alaska (Pollock *et al.*, 1998).

it is far from being completely understood. Reasonably comprehensive studies exist for some groups, especially birds, whereas knowledge about other groups is nearly nonexistent (e.g., most invertebrates). Flooding certainly affects invertebrate community dynamics – distribution, abundance, and diversity – with species richness often increasing with elevation above the river even though flood duration will affect groups differentially (**Figure 2**). Among the invertebrates, spiders (Arachnida) and beetles (Coleoptera) are the better-investigated groups.

In many areas, a large proportion of vertebrates utilize the riparian zone, but it is not always apparent whether individual groups are more or less diverse there (Sabo *et al.*, 2005). For example, in western Oregon and Washington (USA), 87% of the 414 resident species of amphibians, reptiles, birds, and mammals use riparian zones or wetlands, but only 10% use specialized habitat within riparian zones. More is known about species depending on the streams for foraging or reproduction than about species using riparian zones only occasionally. Certainly, amphibians are more diverse in riparian zones because of their reproductive needs. For reptiles, birds, and mammals, it depends on the local environment and specific life history needs.

The best information available is for birds (Gilles and St. Clair, 2008). In the drier western US, bird diversity at the site scale (alpha) is generally high relative to the number of potential species. Between habitats, diversity (beta) varies across a catchment, with riparian assemblages differing most from upland assemblages at the highest and lowest elevations. This pattern is attributed to enhanced avian movements within the riparian corridors. The corridors for bird movements, in turn, facilitate faunal mixing on a broader scale, increasing regional diversity (gamma) within landscapes.

Mammal diversity in riparian zones appears to be great, but this may be deceptive. Mammals are highly mobile, and many, such as moose and elk in winter, make frequent use of riparian zones for feeding and cover. Certainly bats feeding on emerging stream insects are abundant at times, and other medium-sized predators (badgers, foxes, cats) come to feed and drink. Browsers and grazers, as well as rodents and shrews, are often abundant, but most can be found in the uplands too.

Are riparian zones more diverse than the surrounding uplands? The answer is that it depends on the environmental setting and the taxon being considered. Ecological theory would suggest that, in general, riparian areas are highly diverse because of their mosaic structure, presence of refugia, broad ranges of environmental settings and species assemblages.

Riparian Zones as Buffers Against Nutrient Pollution from Upland Runoff

The nutrient buffering characteristics of riparian zones are well known, and may stem from their biodiversity (Lowrance *et al.*, 1995; McClain *et al.*, 2003). As a result, the effectiveness of riparian zones as nutrient filters for water flowing from agricultural catchments to streams has led to major government-supported programs in North America and Europe to conserve and restore riparian buffers. The filtering capability of riparian zones is due to their position in the landscape, and to their geomorphic, hydrologic, and biotic processes. Since riparian zones lie at the interface of terrestrial and aquatic ecosystems, virtually all surface and shallow subsurface runoff in catchments must pass through them in order to reach the stream channel. It has been hypothesized that the diversity of plants contributes to the efficient uptake of nutrients.

Riparian zones are generally planar, with lower gradients than surrounding uplands. This, plus the baffling effect of riparian vegetation, dissipates the kinetic energy of surface flows during storms and causes entrained sediments to be deposited. This is an especially effective mechanism for retaining particulate phosphorus and the chemical pollutants associated with sediment particles. The capacity of riparian zones to retain dissolved nutrients like N, P, Si, Ca, and Mg is controlled by hydrologic characteristics (e.g., water table depth, water

residence time, and degree of contact between soil and groundwater) and by biotic processes (e.g., uptake and biogeochemical transformation by diverse plants and microbes). The relative influence of these factors depends on soil characteristics, nutrient input rates, and vegetative and microbial diversity.

Elucidating the volume and pathway of water moving through riparia is fundamental for understanding nutrient removal and retention (Junk *et al.*, 1989; Tockner *et al.*, 2000). If local groundwater passes beneath the rooting zone or if extensive piping occurs, the roots of riparian vegetation and associated microbes cannot access the nutrients. Riparian vegetation along small streams normally has good access to the water table, and many studies have shown that in those situations riparian vegetation can act as buffers for nonpoint-sources of N and P. However, hydrologic pathways are often more complex along larger rivers, especially where deposits of coarse alluvium are extensive.

Riparian zones are known to be especially effective at protecting surface waters from nitrate runoff (Lowrance *et al.*, 1995). Early studies documented total nitrogen retention by riparian zones ranging from 67% to 89% of total up-slope inputs. Wetlands and soils are also sites of N retention, but a plethora of studies have shown focused removal of nitrate in groundwater moving laterally through riparian zones. In a large-scale study of N budgets in 16 catchments covering 250,000 km^2 of the northeastern US, river exports of N accounted for only 25% of the total N inputs to the catchments, implying that most of the N was retained in some manner.

Candidate mechanisms preventing nitrogen flowing through riparia from entering streams include denitrification, assimilation and retention by the vegetation with the uptake by biota followed by storage in organic detritus. Denitrification is invoked most often as the primary mechanism of nitrate retention and the most important given that N is removed permanently from the system and returned to the atmosphere as N_2 and N_2O. However the extreme spatial and temporal variability of denitrification rates in riparian zones makes it difficult to determine accurate fluxes and to extrapolate these to wider areas. In addition to nitrate, saturated conditions, available carbon, topography, and soil grain size (i.e., water logging potential) are pertinent environmental factors for denitrification.

Plant uptake results in a short-term accumulation of nutrients in nonwoody biomass and a long-term accumulation in woody biomass. Riparian forests are especially important sites for biotic accumulations of nutrients because transpiration may be quite high, increasing the mass flow of nutrient solutes toward root systems, and because morphological and physiological adaptations of many flood-tolerant species facilitate nutrient uptake under low-oxygen conditions. Thus, assimilation by the diverse riparian forest community could be the primary mechanism of nitrate removal from groundwater during the growing season. Nitrogen removal efficiency by riparian buffers is mainly positive across a wide range of climate conditions, nitrate inputs, soil characteristics, and vegetative community types.

The current consensus that most riparian zones effectively remove nitrate from subsurface water is based largely on studies where groundwater inputs are restricted to shallow subsurface flow paths by impermeable layers that force maximum interaction with riparian soils and vegetation. Limited research suggests that there is less effective nitrate retention in riparian areas connected to large upland aquifers where riparian hydrology is often dominated by surface transport or ground water transport below rooting zones. Additionally, it has been discovered that considerable nitrogen can be released from riparian zones to streams as dissolved organic nitrogen. The importance of this pathway, and the forms and subsequent fates of the organic nitrogen are unknown. Collectively, the evidence assembled so far raises doubts about all riparian zones being efficient filters for nutrients – some may be leaky whereas others release nutrients after they have been transformed to organic matter – and if biodiversity might have a role.

Riparian Zones as Buffers Against High In-Stream Nutrient Levels

In much the same way that riparian zones intercept and retain nutrients from upland overland and subsurface runoff, they may also intercept and retain nutrients flowing in adjacent streams. The same mechanisms of N and P removal are at work, including biotic uptake, physical adsorption, and microbial denitrification, and the effectiveness of riparian buffering depends on the characteristics and biodiversity of riparian soils, vegetation, and connectivity between the river and riparian zone along surface–subsurface exchange pathways. Water in streams exchanges continually with the interstitial waters of bed and bank sediments in what may be viewed as a mosaic of surface–subsurface exchange patches. This zone of mixed surface and subsurface waters is known as the hyporheic zone and may extend from a few centimeters to a few kilometers from the channel margin. The volume of surface water moving along subsurface flow paths can be equal to or greater than that moving in the channel.

Hyporheic flow paths act as both sources and sinks of nutrients to streams and the degree of hyporheic influence on river nutrient levels depends on the extent of hyporheic zones, the fraction of river flow diverted through them, the residence time of individual flow paths, and the specific biological and biogeochemical processes operating along flow paths. In general, streams with greater hyporheic exchange – and possibly greater plant diversity – tend to retain and process nutrients more efficiently. Additionally, because the amount of surface-hyporheic water exchange relative to channel volume decreases exponentially with increasing channel size, the efficiency of nutrient removal linked to hyporheic exchanges tends to decrease with increasing stream size. The most important sites of nutrient retention therefore lie in the low-order stream networks of any river basin. For example, first order N-retention rates in headwater streams of the Mississippi River basin average 0.45 day^{-1}, whereas the rate of removal in the mainstem river is two orders of magnitude less, or 0.005 day^{-1}.

Plant cover also influences the efficiency of riparian zones in filtering nutrients and pesticides (Lowrance *et al.*, 1995). A riparian zone vegetated with poplar is more effective for winter nitrate retention than one vegetated with grass. Some trees are better than others in filtering nitrate: *Populus x*

canadensis effectively removes nitrate from saturated soils with a subsequent accumulation of nitrogen in root biomass. Roots of alder, willow, and poplar seem to favor colonization by proteolytic and ammonifying microorganisms and, particularly for alder roots, to inhibit nitrifying microorganisms. It follows that changing plant cover may affect water quality.

Riparian Zones as Buffers Against Suspended Sediments

Riparian vegetation also facilitates the removal of suspended sediments – along with their nutrient, carbon or pollutant contents – from overland flow whether from the uplands or from the adjacent river. A diversity of structural sizes and resilience to flow aids this process. Sediments and sediment-bound pollutants carried in surface runoff are effectively deposited in mature – and diverse – riparian forests and in streamside grasses. Riparian areas remove 80–90% of the sediments leaving agricultural fields in North Carolina. Sediment deposition may be substantial in the long term, with coarser material deposited within a few meters of the field–forest boundary, and finer material deposited further into the riparian forest. Grassy areas are especially effective as they often transform channelized flows into expanded shallow flows, which are more likely to deposit sediment. However, the performance of grassy vegetation seems to be highly variable and to be of short duration when several floods occur within a limited period.

Removal of fine sediment from runoff by riparian zones occurs as a consequence of the interactive processes of deposition and erosion, infiltration, and dilution. This is important because fine sediments carry higher concentrations of labile nutrients and adsorbed pollutants. In forested catchments, with relatively low nutrient concentrations, fine sediments in riparian zones can be sources or sinks for nutrients, depending on how oxidation–reduction conditions affect absorption/desorption to fine particles. For example, the riparian zone of a small deciduous forest stream in eastern Tennessee (USA) was a net source of inorganic phosphorus when dissolved oxygen concentrations in riparian groundwater were low, but was a sink when dissolved oxygen concentrations were high.

In contrast to nitrogen, phosphorus adheres strongly to particles and considerable movement is normally associated with sediment flux. Reductions of 50–85% in total P are observed, with the greatest removal occurring in the first few meters of the riparian zone. Results for soluble P in surface runoff are less consistent due to variability in water flow (i.e., sheet flow, channel flow, or percolation). In general, significant amounts of phosphorus may first accumulate in riparian zones but then be transported to aquatic ecosystems in a different form via shallow groundwater flow, possibly as a result of increased decomposition of organic matter.

Only limited research has been conducted on subsurface P transport in riparian zones. However, the same environmental conditions leading to redox gradients in soils that are responsible for nitrogen removal by denitrification also mediate desorption and release of phosphorus. Although riparian zones may act as effective physical traps (sinks) for incoming particulate phosphorus, they may enrich runoff waters in available soluble phosphorus. Finally, unlike nitrate, phosphorus removal by soil retention and biotic uptake results in accumulation within the system. Consequently, the long-term performance of riparian zones receiving high P inputs remains unclear.

Riparian Zones as Sources of Energy for Adjoining Aquatic Systems

By far the most thoroughly investigated and best understood connection between riparian and stream food webs is via the transfer of riparian plant litter to streams (Wallace *et al.*, 1997). Riparian organic matter inputs represent highly diverse allochthonous sources of energy as opposed to the autochthonous organic matter contributed by aquatic primary producers. In low-order streams beneath closed-canopy riparian forests the influx of carbon from riparian plant sources, both surface and subsurface, may amount to 80–95% of total organic carbon influx to streams. Although the area-normalized flux of riparian litter to river systems decreases downstream, the total amount and diversity of riparian litter input to the river continues to rise and varies as a function of channel morphology and riparian forest structure and composition.

Once in the stream, riparian organic matter is decomposed by a variety of specially adapted microbial and invertebrate fauna (Dang *et al.*, 2005). When litter (mainly leaves and needles) first enters streams there is a brief period (a few days) of rapid leaching in which 25% or more of the initial dry weight can be lost. Biotic decomposition is initiated by hyphomycete fungi that break up the litter's structural integrity by secreting enzymes to hydrolyze cellulose, pectin, chitin, and other difficult-to-digest compounds. Fungal community composition is closely tied to the riparian forest, and fungal species diversity has been positively correlated with riparian tree richness. With time fungi give way to bacteria as the dominant microorganism in the decay process. Decomposition rates are driven by substrate quality, stream nutrient concentrations, redox conditions, and temperature. Fungal decomposition alone can fragment leaves into flakes of finer particulate organic matter within weeks. This fragmentation process is critical to energy dispersion in streams and rivers because finer fragments tend to be more mobile and to therefore fuel metabolism in downstream river sections. Microbially colonized litter has higher nutrient concentrations than noncolonized litter, and is therefore the preferred choice of macroinvertebrate consumers – shredders and collectors such as caddis flies and black flies – that make up the next link in aquatic food webs.

Riparian arthropods are also important energy sources to stream consumers (Nakano *et al.*, 1999). Arthropods fall into streams from overhanging foliage by accident and the flux is proportional to arthropod abundance in the canopy. It is also postulated that the diversity of the arthropod flux is related to plant diversity. Arthropods may also wash into streams and rivers during overland flow events. The normalized flux (per square meter of channel area) is higher in smaller streams flowing beneath a closed riparian canopy but even in larger streams and rivers the flux may remain substantial at the channel margins. Once in the aquatic system, riparian arthropods are consumed by drift foraging fish and may

constitute a major proportion of their diet. Riparian-derived arthropods are higher quality food than riparian litter and are directly available to top consumers such as fish. Experimental evidence shows that withholding this energy input from streams has consequences that reverberate through aquatic food webs and ultimately upset the basic composition of the stream community.

Energy flows in both directions across the terrestrial–aquatic interface, and riparian food webs are also subsidized by aquatic resources. The emerging adults of aquatic insects are an important energy source to a variety of riparian arthropods, and this energy subsidy is passed to higher trophic levels by the lizards, bats, shrews, and birds that consume riparian arthropods. Riparian arthropods inhabiting resource-scarce habitats such as exposed gravel bars and desert riparian environments appear to rely almost exclusively on aquatic prey. For example, emerging aquatic insects composed 80–100% of the diet of certain staphylinid and carabid beetles and about 50% of the diet of lycosid spiders inhabiting gravel bars of the Tagliamento River in NE Italy.

Reciprocal energy subsidies such as these are especially important over the course of the year in temperate regions due to strong seasonal variability in the emergence and abundance of different insects. For example, aquatic arthropod abundance peaks following 'leaf-out' of riparian forests in spring and defoliation in autumn whereas riparian arthropod abundance peaks during the summer when forest productivity is maximal.

Large Animal Influences on Riparian Zones

Large animals influence nutrient and energy flows by consuming and redistributing energy and nutrients within riparian zones as well as across adjacent system boundaries, and maintaining their natural population dynamics are important for riparian biodiversity (Helfield and Naiman, 2006). Most importantly, large animals may alter the hydrologic and geomorphic characteristics of riparian zones, causing fundamental changes in energy and nutrient cycles, and altering plant community composition and structure.

Animals that pond water, dig holes, trample plants, or move materials cause fundamental geomorphic changes. For example, hippopotamus (*Hippopotamus amphibious*) increase the ponding of water in African stream networks by creating and maintaining deep pools and forming trails between channels and adjacent terrestrial feeding areas. Beaver (*Castor canadensis* and *C. fiber*) profoundly influence the short- and long-term structure, function, and diversity of riparian zones of drainage networks in the boreal forests of northern latitudes by cutting wood and building dams. In catchments where beaver are abundant there may be 2–16 dams per kilometer of stream length, and each dam may retain between 2000 and 6500 m^3 of sediment. Ponds are eventually abandoned as they fill with sediment or as local food resources are depleted and, once abandoned, dams fail and ponds drain to produce nutrient-rich wetland meadows.

Animals browsing riparian and aquatic vegetation strongly influence riparian community structure, soil development, and propagule dispersal. Animals that browse selectively keep preferred plant species from dominating the plant assemblage and thereby provide an advantage to species not browsed. For example, moose (*Alces alces*) prefer willow (*Salix*) and poplar (*Populus*), thus giving a competitive advantage to white spruce (*Picea glauca*), which is not browsed.

The actions of all large herbivores influence the belowground components of riparian systems as well, with consequent effects on interactions that determine long-term ecosystem function. For example, grazing activities that stimulate the growth of early successional species, and therefore retard succession, help to maintain higher belowground productivity.

Pacific Salmon Influences on Riparian Zones

A remarkable example of the consequences of animal-mediated nutrient and energy flows in riparian zones is the migration of salmon (*Oncorhynchus* spp.) from the North Pacific Ocean to spawning areas in fresh water. Migrating Pacific salmon transport marine-derived (MD) carbon and nutrients upstream and, on death after spawning, hydrologic and animal pathways distribute these elements throughout aquatic and riparian systems (Bartz and Naiman, 2005; Naiman *et al.*, 2009). In an important system-scale feedback, fertilization of riparian plant communities with MD-nutrients enhances the growth of some riparian plants. This may lead to reduced diversity among riparian plants, but it positively influences salmon over the longer term by supplying stream organisms with an increased supply of nutritious litter and by improving salmon habitat via an influx of large diameter riparian-derived wood.

Historically, spawning salmon represented a flux of nearly 7000 Mt of nitrogen and more than 800 Mt of phosphorus to rivers in California, Idaho, Oregon, and Washington. Although fluxes have been reduced by $>90\%$ during the past century as populations have declined, salmon still are an important source of nutrients to many river and riparian systems of Canada, Alaska, Russia, and Japan.

Maintaining Riparian Biodiversity

There are four principal categories of threat to the biodiversity of riparian zones – flow regulation, pollution, climate change, and land use alterations (**Table 1**); all are linked and exacerbated by the modification of river flows and wetland inundation regimes (Rood *et al.*, 2010). One of the most promising strategies for integrating freshwater management into the broader scope of ecological sustainability is the provision of 'environmental flows' – the flows left in rivers, or restored to developed rivers, to sustain key ecological and societal values. Other terms, such as 'in-stream flows,' 'ecological flows,' 'environmental water allocations,' and the 'normative flow regime,' convey much the same concept. Environmental flows describe the quantity, timing, and quality of water flows required to sustain freshwater and estuarine ecosystems and the human livelihoods and well-being that depend on these ecosystems, including biodiversity. There is now wide recognition that a dynamic, variable water regime is required to maintain

Table 1 Summary of the major types of anthropogenic environmental change and their principal effects on riparian zone biodiversity

Environmental change	*Principal effects on riparian zones*
Flow regulation	
Flow regime	Alters community composition and successional processes; loss of life history cues
Dams	Lotic to lentic; inundation above dam; altered flow, nutrient, sediment, and temperature regimes below dam
Withdrawals	Lowers water table; alters flow regime; decreases alluvial aquifer recharge; system simplification
Channelization and Dredging	Lowers water table; desiccates riparian forest causing terrestrialization and change in community composition; possible decline in biodiversity
Levees	Isolate rivers from floodplains, thereby reducing hydraulic connectivity laterally and vertically. Constrain channel migration; alter riparian successional trajectories
Pollution	
Nutrients	Increases productivity; shifts community composition toward tolerant species; organic loading leads to redox changes
Toxic materials and acid rain	Decreases productivity; declining biodiversity; system simplification; shifts community composition toward few tolerant species
Climate	
Precipitation	Modifies entire flow regime, groundwater-surface water exchanges, and channel morphology and stability; loss of life history cues
Temperature	Spatial patterns and phenology of riparian species are changed
Land use	
Vegetative cover	Modifies albedo and feedbacks to climate; changes local microclimate and successional trajectories
Invasive species	Introgression and hybridization; increased competition for space and resources; may reduce biodiversity
Resource management	Usually alters successional trajectories and community composition

the native biodiversity and ecological processes characteristic of every river and wetland ecosystem. Yet, it remains a challenge to translate this 'natural flow regime' paradigm into quantitative environmental flow prescriptions for individual river reaches from source to sea (Mahoney and Rood, 1998). These challenges are receiving wide recognition and, in many cases, are being met; thereby maintaining biodiversity and other riparian characteristics.

Riparian areas are often prioritized for protection to preserve specific services described in the preceding sections and also to preserve ecological and evolutionary processes important in conserving and generating biodiversity in altered landscapes. In their unique position along waterways connecting other landscape features, dynamic riparian areas provide refugia and corridors of movement for wildlife and plants and transfer pathways for energy and nutrients. The dynamic disturbance regimes of riparian areas also impose a continual stimulus for plant diversification. Riparian buffers have been shown to preserve and stimulate heightened biodiversity in agricultural and other modified landscapes. These functions may be threatened, however, by invasions of alien plant species which can reduce biodiversity and modify natural processes. Maintaining key components of the natural flow regime is effective in reducing alien invasions.

Conclusions

Riparian zones are highly complex physical and biological systems that both harbor significant biodiversity and facilitate biodiversity conservation across landscapes. Their complexity, biodiversity, and distinctive ecological functions are maintained through strong spatial and temporal biophysical connectivity with adjacent riverine and upland systems. Water, sediments, and nutrients enter riparian zones from adjacent uplands and streams, mixing and reacting along dynamic surface and subsurface flow paths (Van Pelt *et al.*, 2006). Under normal flow conditions the biodiverse riparian zone retains a significant portion of these materials and generally returns chemically purer water to streams and rivers. At the same time riparian zones are important sources of energy to both upland and aquatic systems in the form of plant and insect tissues. Many stream food webs fundamentally depend on these resources and many upland animals depend on them as important subsidies to their diets. At the same time, riparian communities benefit from the enhanced productivity of adjoining ecosystems (especially aquatic systems) through physical and biotic feedbacks that return a portion of that productivity to riparian zones in the form of organic matter and nutrients.

The unique biodiversity and ecological functions of riparian zones are linked to dynamic biophysical processes and interactions across multiple spatial and temporal scales. Maintaining these interactions and the connectivity driving them is a fundamental requirement for maintaining healthy riparian zones and the many services they provide; effective management requires maintaining connectivity, both in the timing and extent of flows as well as in the movements and types of animals.

See also: Agriculture, Nutrient Management, and Water Quality. Biogeochemical Cycles. Birds, Biodiversity of. Diversity, Taxonomic versus Functional. Ecosystem Function, Principles of. Food Webs. Freshwater Ecosystems. Freshwater Ecosystems, Human Impact on. Functional Groups. Hotspots. Keystone Species. Landscape Corridors. Nitrogen, Nitrogen Cycle. Plant–Animal Interactions. River Ecosystems. Salmon. Subterranean Ecosystems. Wetlands Ecosystems

References

Arthington AH, Naiman RR, McClain ME, and Nilsson C (2010) Preserving the biodiversity and ecological services of rivers: New challenges and research opportunities. *Freshwater Biology* 55: 1–16.

Bartz KK and Naiman RJ (2005) Impacts of salmon-borne nutrients on riparian soils and vegetation in southeast Alaska. *Ecosystems* 8: 529–545.

Dang CK, Chauvet E, and Gessner MO (2005) Magnitude and variability of process rates in fungal diversity litter relationships. *Ecology Letters* 8: 1129–1137.

Décamps H (1996) The renewal of floodplain forests along rivers: A landscape perspective. *Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie* 26: 35–59.

Gillies CS and Clair St. CC (2008) Riparian corridors enhance movement of a forest specialist bird in fragmented tropical forest. *Proceedings of the National Academy of Sciences of the United States of America* 105: 19774–19779.

Helfield JM and Naiman RJ (2006) Keystone interactions: Salmon and bear in riparian forests of Alaska. *Ecosystems* 9: 167–180.

Hocking MD and Reynolds JD (2011) Impacts of salmon on riparian plant diversity. *Science* 331: 1609–1612.

Junk WJ, Bayley PB, and Sparks RE (1989) The flood pulse concept in river-floodplain systems. In: Dodge DP (ed.) *Proceedings of the International Large River Symposium. Canadian Special Publication of Fisheries and Aquatic Sciences*, vol. 106, pp. 110–127. Ottawa, Canada: NRC Research Press.

Lowrance RR, Altier LS, Newbold JD, *et al.* (1995) Water quality functions of riparian forest buffer systems in the Chesapeake Bay Watershed. *U.S. Environmental Protection Agency Chesapeake Bay Program, E.P.A. Publication 903-R-95-004 CBP/TRS 134/95.* Annapolis, MD.

Mahoney JM and Rood SB (1998) Streamflow requirements for cottonwood seedling recruitment: An integrative model. *Wetlands* 18: 634–645.

Malanson GP (1993) *Riparian Landscapes*. Cambridge: Cambridge University Press.

McClain ME, Boyer EW, Dent CL, *et al.* (2003) Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems* 6: 301–312.

Naiman RJ and Décamps H (1997) The ecology of interfaces: Riparian zones. *Annual Review of Ecology and Systematics* 28: 621–658.

Naiman RJ, Décamps H, and McClain ME (2005) *Riparia: Ecology, Conservation and Management of Streamside Communities*. San Diego: Elsevier/Academic Press.

Naiman RJ, Helfield JM, Bartz KK, Drake DC, and Honea JM (2009) Pacific salmon, marine-derived nutrients and the characterstics of aquatic and riparian ecosystems. In: Haro AJ, Smith KL, Rulifson RA, *et al.* (eds.) *Challenges for Diadromous Fishes in a Dynamic Global Environment*, pp. 395–425. Bethesda, Maryland: American Fisheries Society Symposium 69.

Nakano S, Miyasaka H, and Kuhara N (1999) Terrestrial-aquatic linkages: Riparian arthropod inputs alter trophic cascades in a stream food web. *Ecology* 80: 2435–2441.

Nanson GC and Beach HF (1977) Forest succession and sedimentation on a meandering-river floodplain, northeast British Columbia, Canada. *Journal of Biogeography* 4: 229–251.

Pollock MM, Naiman RJ, and Hanley TA (1998) Plant species richness in riparian wetlands: A test of biodiversity theory. *Ecology* 79: 94–105.

Rood SB, Braatne JH, and Goater LA (2010) Favorable fragmentation: River reservoirs can impede downstream expansion of riparian weeds. *Ecological Applications* 20: 1664–1677.

Rood SB, Gourley C, Ammon EM, *et al.* (2003) Flows for floodplain forests: Successful riparian restoration along the lower Truckee River, Nevada, USA. *BioScience* 7: 647–656.

Sabo JL, Sponseller R, Dixon M, *et al.* (2005) Riparian zones increase regional species diversity by harboring different, not more species. *Ecology* 86: 56–62.

Salo J, Kalliola R, Häkkinen I, *et al.* (1986) River dynamics and the diversity of Amazon lowland forest. *Nature* 322: 254–258.

Tockner K, Malard F, and Ward JV (2000) An extension of the flood pulse concept. *Hydrological Processes* 14: 2861–2883.

Van Pelt R, O'Keefe TC, Latterell JJ, and Naiman RJ (2006) Structural development and stand evolution of riparian forests along the Queets River, Washington. *Ecological Monographs* 76: 277–298.

Wallace JB, Eggert SL, Meyer JL, and Webster JR (1997) Multiple trophic levels of a forest stream linked to terrestrial litter inputs. *Science* 277: 102–104.

River Ecosystems

KE Limburg, State University of New York, NY, USA
DP Swaney, Cornell University, NY, USA
DL Strayer, Cary Institute of Ecosystem Studies, Millbrook, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Allochthonous Referring to the production of organic matter outside the ecosystem; in streams and rivers, this would be inputs from the upstream watershed.

Autochthonous Referring to the production of organic matter within the ecosystem, for example, primary production by aquatic plants.

Cultural eutrophication The acceleration of nutrient overenrichment caused by humans; it can be caused by direct, point-source pollution (sewers) or by diffuse, nonpoint-source pollution (such as fertilizer runoff from farm fields).

Hyporheic zone The area of saturated soils beneath a stream or a river channel; it interacts with the stream through the processes of hydraulic upwelling and downwelling.

Riparian zone The vegetated areas on either side of a stream or river, with saturated soils usually underlying aerated soils.

River continuum concept The first of a series of conceptual, unifying models to explain and predict the structure and function of river ecosystems.

Stream order A numbering system to denote stream size within a network of streams. There are several stream order systems. Stream order is also dependent on the scale of observation.

Watershed or catchment A unit of landscape in which precipitation drains to a common stream or lake; alternatively, a watershed is the divide between catchments.

Introduction

Rivers are the sinuous conduits of overland flow that drain the world's continents. They displace very little of the world's land area and their standing stock of water is small relative to other hydrologic storages. Of the estimated 1,385,984,610 km^3 of global water, the average volume of rivers composes about two ten-thousandths of 1%, or 0.006% of the freshwater supply (Maidment, 1993). Nevertheless, rivers are of disproportionate importance in terms of geomorphology, ecology, and biodiversity, for they serve as transport vectors for sediments, organic matter, nutrients, and species. In contrast to the image of a river as a pipeline from the source to the mouth, river networks form bifurcating, fractal patterns through the landscape that are among the most complex in nature. Along any reach of the river channel, a fully three-dimensional world lies between the riverbanks; therein are found habitats in which organisms form ecological communities and process energy and matter in various ways, depending on where the reach is located. Because of the intimate connections of rivers with the landscape, the burgeoning pressures of human-accelerated environmental change are becoming particularly acute in these systems.

Anatomy of Rivers

To a hydrologist, the basic unit of the landscape is the catchment or watershed, in which water organizes itself into flowpaths. Watersheds are formed by the various forces of continental uplift and wearing down (weathering) of the bedrock by water and wind. Geomorphologically, vegetation within watersheds plays a dual role: It aids in weathering and conversion of rock to soils and it retards surface runoff, diverting some of it to gaseous water vapor by evapotranspiration. Watershed boundaries, or divides, are those more resistant zones where water flows either one way or the other into alternative drainages. The Continental Divide in the US separates the western drainages (which flow into the Pacific Ocean) from those (principally the Mississippi) that drain east or south. Divides at the continental scale can often limit the flow of species, but the process of species flow is affected at all scales of watersheds.

Watersheds and their accompanying flowing waters are organized hierarchically, that is, smaller watersheds usually form into larger watersheds (unless a physical boundary prevents this), which in turn may form parts of regional drainages. Hydrologists in the 1940s and 1950s recognized that streams and rivers form drainage networks and defined topological metrics to describe these. The most common metric is stream order. Although there are various definitions of order, Strahler (1952) considers that a first-order stream has no tributaries, a second-order stream is formed at the meeting of two first-order tributaries, a third-order stream at the meeting of two second-order tributaries, etc. Shreve's link magnitude system (Shreve, 1966) defines the magnitude of a stream, downstream of the meeting of two streams, as the sum of the magnitudes of those streams. It should be noted that order is inevitably related to the scale of the map on which the stream and its tributaries are plotted (typically, in the US, the smallest streams delineated at 1:24,000 scale USGS topographic maps are considered first-order streams, but higher resolution maps produce higher order streams). This is a

characteristic of the fractal nature of drainage networks (Rodriguez-Iturbe and Rinaldo, 1997). Another metric is the bifurcation ratio, or the ratio of first- to second-order, second- to third-order, etc. streams within a drainage network. The drainage density, or the ratio of cumulative length of streams (L) to watershed area (L^2), is a third measure of network organization. High bifurcation ratios might indicate the presence of many small streams, and in conjunction with a high drainage density, also describe a highly dissected landscape. More complex models, which take into account differences in elevation, geology, and soils, provide even more detail for water flow and habitat creation.

Early models of watershed hydrology assumed that precipitation events saturated the soils, producing overland flow, beginning from the base of a drainage and moving uphill, much like filling a bathtub. However, in the 1970s, the variable source area model, in which the flow originates from source areas that expand and contract in different parts of the watershed in response to the dynamics of precipitation and drainage characteristics, gained acceptance. Considerable overland flow during a storm is actually water that had been stored in the soil, which is displaced by the new water coming in.

From an ecological perspective, the multitude of geomorphologic forms and flowpaths gives rise to a similar multitude of habitats and ways for matter to be transformed, at scales ranging from the microscopic to the entire basin. A typical cross-sectional view of a stream is given in **Figure 1**.

In this figure, it is clear that the stream is in contact with several areas: A zone of saturated sediments below, called the hyporheic zone; an area with emergent mud or gravel bars overlying saturated soils, called the parafluvial; and beyond that a zone of bankside vegetation, also in contact with saturated soils, the riparian zone. In upland streams, where the terrain is steep and rugged, these zones may be compressed around the stream channel, but farther downstream the zones may spread out across the floodplain that was carved out over geological time by the flowing water.

Hynes (1975) was among the first stream ecologists to point out the close connection between rivers and their watersheds. As systems, rivers and streams are inherently open and serve both as connectors, transporters, and processors of material. Among the biological connections of streams and watersheds are the inputs of organic matter from vegetation: leaves, insect frass, partial, or entire trees (which form debris dams). Water does much of the physical work in carving channels and carrying sediment, but organisms mediate this by regulating the rates of mineral use and transport, fixing organic matter and passing it along the food chain, and even altering landforms so that hydrology is affected (e.g., the work of beavers).

Energy and Matter Transformation (Biogeochemistry)

From a biogeochemical standpoint, rivers differ from other ecosystems in at least three important ways: First they are flowing systems that provide a source of kinetic energy, as well as externally produced nutrients, to resident organisms. Second they are spatially complex and extensive systems, with a relatively large interfacial connection to contiguous terrestrial ecosystems. Such a large buffer zone has significant implications for processing nutrients and other fluxes into the river. Finally, as discussed in the previous section, they are transitional systems, changing in nature from small, shallow, flashy first-order streams that are closely connected to the landscape, to large, relatively deep and steady rivers that may be tidally influenced as they become estuarine in nature near their mouths. The interplay between light, nutrients, and turbulent energy and circulation structures the biotic community, and this structure changes along the river from the sources to the mouth.

Riverine ecosystems are fueled by inputs of nutrients, including carbon, nitrogen (N), phosphorus (P), and silicon (Si). Nutrients that are available to primary producers must generally exist in inorganic (mineral) forms. In pristine watersheds, rock weathering and decomposition of organic material in forests, grasslands, and soils provide the sources of these nutrients. In most watersheds today, nutrient loading is dominated by anthropogenic inputs, including fluxes from the

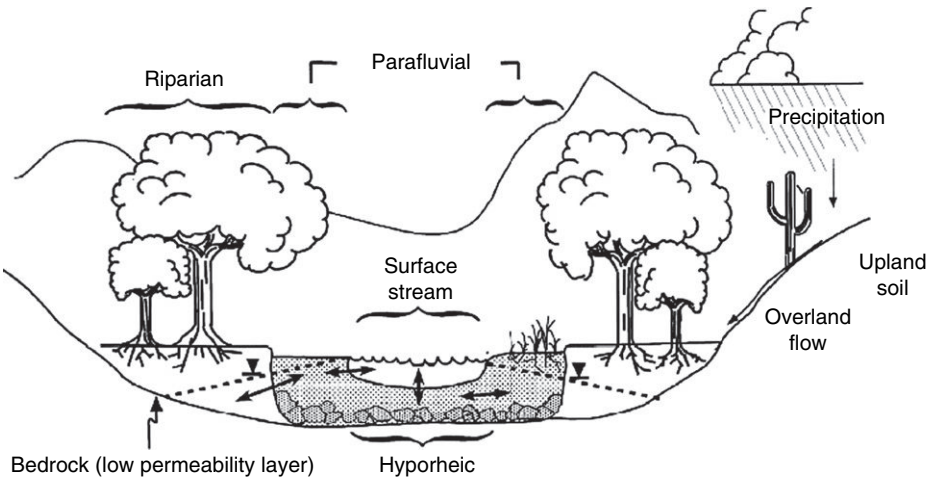


Figure 1 Cross-sectional view of the stream channel and zones that interact with it. Reproduced from Fisher SG, Grimm NB, Marti E, Holmes RM, and Jones Jr. JB (1998) Material spiraling in stream corridors: A telescoping ecosystem model. *Ecosystems* 1: 19–34, with permission from Springer.

atmosphere, agricultural and urban fertilizer applications, N fixation of crops, and net food and feed transfers across watershed boundaries. These loads can be collectively termed Net Anthropogenic Nutrient Inputs (NANI) and simple accounting methods are used to estimate them, especially for N and P (Howarth *et al.*, 1996, 2011; Boyer *et al.*, 2002; Russell *et al.*, 2008; Hong *et al.*, 2011). Atmospheric deposition of nutrients from local or distant agricultural, industrial, and automotive sources occurs via rainfall and dust. Within the watershed, fertilizer from agricultural and urban applications enters rivers in surface runoff and groundwater; discharges of sewage treatment plants and industry provide other sources. Frequently, nutrients also enter rivers in various organic forms (i.e., as byproducts of other organisms, such as soil particles, plant and animal remains, and feces). Sources of organic nutrients associated with human activity include untreated sewage and manure runoff from farms. Organic nutrients entering a river segment from the watershed upstream augment the local primary production of biomass within it to fuel heterotrophic organisms. Such external sources are termed allocthonous as opposed to the autochthonous production occurring within the river. Regardless of whether nutrients enter rivers in organic or inorganic forms, excessive loading can result in accelerated riverine metabolism in which organic matter is consumed in respiration, with a consequent increased biological oxygen demand (BOD). When BOD exceeds the local oxygen supply in a reach of the river, available oxygen levels may decline precipitously (resulting in the oxygen sag curve downstream of nutrient sources familiar to environmental engineers), with catastrophic results to riverine organisms. Such anoxic zones typically resemble the popular image of polluted waters: smelly masses of decaying algae, fish kills, and greatly reduced biodiversity.

Because organisms require nutrients in particular ratios (in oceanography, these are termed Redfield ratios after their discoverer Redfield (1890–1983)) to maintain their metabolism, deviations of environmental concentrations from these ratios can result in species shifts. For example, the shift of the available N:P ratio below the Redfield ratio of 16:1 (moles N per mole P) in fresh waters is often associated with the onset of a bloom of cyanobacteria (blue–green algae) that are capable of fixing N from the atmosphere. Depletion of silicate below the Si:N molar ratio of 1:1 has been observed to shift the dominant species of algae from diatoms (which require silicate to synthesize their siliceous skeletons) to species that do not utilize silicate. Such species shifts are sometimes associated with nuisance algal blooms.

It must be emphasized that the dominant physical controls of riverine ecosystems change drastically as stream order increases. The ecosystem of the shallow, well-mixed, intermittent stream is controlled almost completely by weather and the nature of the contiguous terrestrial ecosystem – geology and soils, nutrient processing in the riparian zone, light levels established by the degree of canopy closure, etc. Near its mouth, the river is deeper and broader and influenced by the sea. Relatively dense seawater meets the lighter freshwater, sometimes resulting in stratification, which, together with the effects of the tide, yields complicated patterns of circulation. The cumulative sediment load of the watershed results in decreased water clarity and increased sediment deposition (the deltas of many large rivers). The combined workings of internal circulation and turbidity govern the light and nutrient fields of the water column as well as the distribution of biological communities in ways that are not completely understood at present.

Different Conceptual Models of Riverine Ecosystems

Beginning in the 1970s, synthetic conceptual models of rivers began to be developed as part of the ongoing movement to understand the world's ecosystems, under the auspices of the International Biological Programme (IBP). One of the fundamental insights that came to the fore was the necessity of integrating stream ecosystem dynamics over a range of spatial and temporal scales.

River Continuum Concept (RCC)

The RCC (Vannote *et al.*, 1980) originated as a theoretical concept to organize and predict the structure and function of riverine ecosystems from source to sink. The premise is that river ecosystems are dominated by the physical factors that create the drainage network and that ecological communities tend to be selected in response to fluvial geomorphic processes, which, along with hydrology, force a more or less predictable template onto the biotic components of the system. According to the RCC, the kinds of producers, consumers, magnitude of organic inputs, and degree of heterotrophy or autotrophy should be predictable along the physical gradient. Ecological communities should be in dynamic equilibrium with the stream, and the openness of streams, connecting upstream and downstream processes, is manifested in the RCC.

Figure 2 depicts the RCC. The first thing that can be seen is that the stream is organized as a gradient, from an idealized headwater stream to a large river (stream order is on the left of the diagram, and it suggests a logarithmic scale). The figure also attempts to unify biological communities with functional and ecosystem-level properties.

The RCC authors divide streams into headwaters (orders 1–3), medium-sized streams (4–6th order), and large (>6th order). They point out that many headwater streams are dominated, in terms of the energy coming in, by the allochthonous inputs from out of channel, because the tree canopy tends to close off the stream from direct sunlight. In this case, the *P:R* ratio (the ratio of primary production to ecosystem respiration) is less than 1, and most of what enters the stream is coarse particulate organic matter (CPOM, matter >1 mm), which, in the case of tree boles, can be very coarse indeed. The aquatic food chain here is dominated by organisms that can deal with CPOM, mediated by microbial processing – the so-called guilds of shredder, grazers, collectors, and their predators.

Moving downstream, the relative contribution from allochthonous inputs declines, and autochthonous production (from submersed macrophytes, phytoplankton, and epiphytes) and processed inputs from upstream (now fine particulate organic matter (FPOM), 0.005–1 mm grain size) become more important. Somewhere in this mid-zone, *P:R*

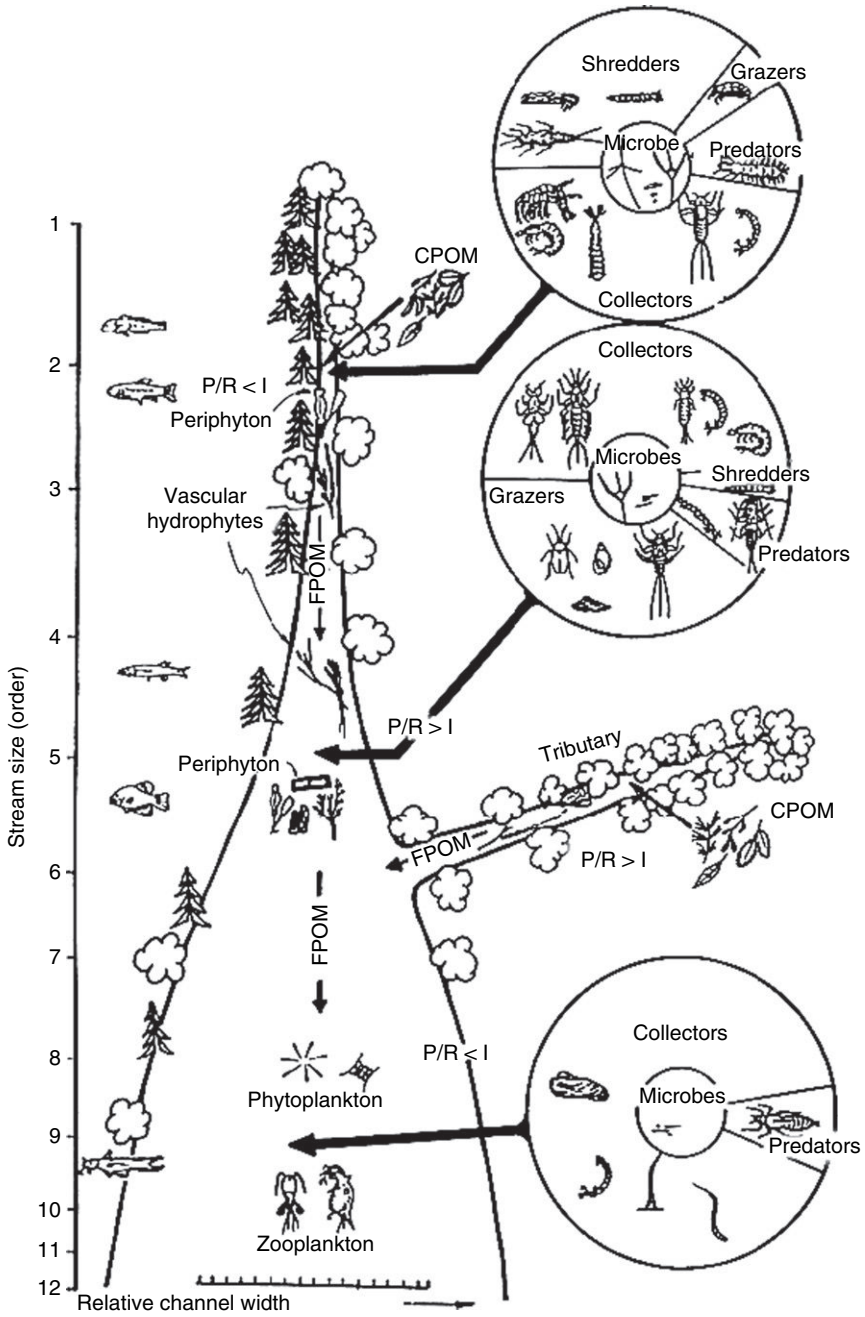


Figure 2 The River Continuum Concept. Reproduced from Cummins KW, Cushing CE, and Minshall GW (1995) Introduction: An overview of stream communities. In: Cushing CE, Cummins KW, and Minshall GW (eds.) *Ecosystems of the World, Vol. 22: River and Stream Ecosystems*, pp. 1–8. Amsterdam, The Netherlands: Elsevier.

reaches a maximum and is greater than 1, indicating that this is the most productive (*in situ*) part of the continuum.

The large river end of the gradient is dominated by FPOM inputs and primary production, but the latter can be limited due to turbidity (due in part to inorganic sediment, not discussed explicitly in the RCC). Turbidity attenuates the light in the water column; thus, phytoplankton production may be limited and *P*:*R* is predicted to decrease. Because large rivers increase in depth, they become more lake-like, with a deep water column and a muddy benthic zone, so that many of the stream insects drop out of the system.

The RCC also makes predictions about where in the system the maximum diversity of organisms should occur (in the medium-sized reaches, or areas with high physical variation, because these are arguably the most diverse of habitats), and that the biotic components of the system will serve to enhance the stability of the ecosystem (referring to a long-ongoing debate as to the importance of biodiversity to ecosystem stability).

One final prediction, or observation, in the RCC is that rivers have both spatially and temporally redundant groups of organisms that tend to perform the same kinds of ecosystem functions, that is, breaking down the organic matter inputs and remineralizing nutrients; they also argue that one should not expect to see the development of successional stages of biotic communities in rivers, "…because the communities in each reach have a continuous heritage rather than an isolated temporal composition within a sequence of discrete successional stages (Vannote *et al.*, 1980, p. 135)." We shall see later on, that scientists dispute this idea.

Resource (Nutrient) Spiraling

> *All land represents a downhill flow of nutrients from the hills to the sea. This flow has a rolling motion. Plants and animals suck nutrients out of the soil and air and pump them upward through the food chains; the gravity of death spills them back into the soil and air. Mineral nutrients, between their successive trips through this circuit, tend to be washed downhill. Without the impounding action of soils and lakes, plants and animals would have to follow their salts to the coast line.* (Leopold, 1941.)

The famous conservationist, Leopold, makes important points in his elegant prose: First, there is a tendency for elements to run downhill from the uplifted land to the sea, and second, that biota play a key role in slowing down this transfer. Third, he speaks of the rolling motion of mineral transfer across the land and waterscape. Webster (1975) used the term spiraling to describe the combined downhill movement and cycling aspects of nutrients and organic matter, and Elwood *et al.* (1983) derive a specific model to describe the dynamics. The concept describes how an atom of a resource (a nutrient like N, P, or C) moves into a stream and how it is retained, processed, and released downstream. The atom may travel downstream, be picked up in a part of the system, cycled there many times, buried for some time, or released relatively quickly. The spiraling model provides an analytic framework to quantify the relationships between biogeochemical processing of materials and hydrodynamic transport, proposing the spiraling length as a measurable index of this interaction.

Spiraling of nutrients in streams is slowed by biotic retention (through sorption and ingestion) of dissolved and particulate forms. Recycling of nutrients *in situ* further slows

	Mechanism		Effect on nutrient cycling		Ecosystem response to nutrient addition	Ecosystem stability	Categorization of study streams
	Retention	Biological activity	Rate of recycling	Distance between spiral loops			
(a)	High	High	Fast / Short; Cycling rate -1; Stream flow; Distance between loops		Conservative ($I > E$)	High	MI 2,3 PA 1,2,3
(b)	High	Low	Slow / Short		Storing ($I > E$)	High	OR 1,2 ID 1 MI 1
(c)	Low	High	Fast / Long		Intermediately conservative < A but > D	Low	ID 3 MI 4 PA 4
(d)	Low	Low	Slow / Long		Exporting ($I = E$)	Low	OR 3,4 ID 2,4

Figure 3 Predicted effects of interactions of downstream transport and measures of biological activity, and postulated responses of the river ecosystem to disturbance, here in the form of nutrient addition. I, import; E, export. The last column refers to streams in a study (by state) and their approximate stream order. From Minshall (1983), reproduced in Cummins KW, Cushing CE, and Minshall GW (1995) Introduction: An overview of stream communities. In: Cushing CE, Cummins KW, and Minshall GW (eds.) *Ecosystems of the World, Vol. 22: River and Stream Ecosystems*, pp. 1–8. Amsterdam, The Netherlands: Elsevier.

down their downstream transport. Adopting a steady-state, averaged approach for a first derivation, the total downstream flux of a nutrient (F_T) is related to the standing stock of nutrient available per unit length of stream (N) and the average downstream travel speed (velocity) of the nutrient (v):

$$F_T\left(\mathrm{g\ s^{-1}}\right) = N\left(\mathrm{g\ m^{-1}}\right) \cdot v\left(\mathrm{m\ s^{-1}}\right)$$

As the biota take up the nutrient (postulated to occur mostly on the stream bottom), the average downstream velocity of the nutrient slows. The channel waters are assumed to be highly advective and the channel bottom retentive.

A second aspect of spiraling is the rate at which nutrients are cycled in the stream. The use rate of a nutrient N is defined as

$$U\left(\text{use rate, g m}^{-1}\text{s}^{-1}\right) = k\left(\text{s}^{-1}\right) \cdot N\left(\text{g m}^{-1}\right)$$

where k is a constant of proportionality (the fraction of the standing stock of N that is recycled per unit time, time being seconds in this case). A nutrient atom on average completes a cycle in $T = 1/k$.

By dividing the first equation by the second one, a ratio of retention to cycling is obtained:

$$Nv/Nk = v/k = vT = F_T/U = S$$

$$\left(m\ s^{-1}\right)/s^{-1} = \left(m\ s^{-1}\right)s = \left(g\ s^{-1}\right)\left(g\ m^{-1}\ s^{-1}\right) = (\mathrm{m})$$

The parameter S is the average downstream distance traveled by a nutrient atom during a cycle (between bottom and flowing water) and is called the spiraling length. If the value of k is large (high cycling rate within the stream section) and v is low (resistance to stream transport), then S will be short and a greater retention of the nutrient is implied.

Figure 3 illustrates several predictions for how the spiraling length will be affected by different retention and biological processing rates. Ecosystem stability and response to nutrient loading (a disturbance) are also predicted as a consequence of spiraling.

The nutrient spiraling model can be elaborated for any number of compartments, and different nutrient spirals can be measured simultaneously as well. Carbon is one exception that is treated separately, because carbon moves so readily in and out of the dissolved and gaseous phases.

Another important point about spiraling is the potential to identify fast and slow components of the system. Slow components (with tight nutrient spirals and short S) are postulated to be relatively resistant to perturbation, implying that these help to stabilize a system in the face of disturbance.

Serial Discontinuity Concept

This heuristic model is a corollary to the RCC, developed by two river scientists (Ward and Stanford, 1983) from the Western US, where many rivers do not flow freely. Whereas the RCC is intended to describe more or less pristine systems, Ward and Stanford were concerned about rivers that contain dams and impoundments. They argued that such regulating structures reset the river continuum, although not always in the low-order to high-order direction. The key feature of serial discontinuity is that it can cause a stream reach to behave or function like reaches that the RCC would predict should occur in a different stream order. **Figure 4** illustrates the general concept.

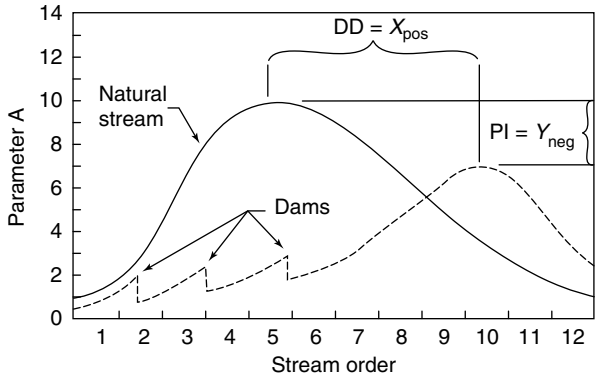


Figure 4 Main features of the Serial Discontinuity Concept, describing the effect of an impoundment on a river ecosystem. DD, discontinuity distance is the distance that a parameter is displaced; PI, parameter intensity, or degree to which a parameter is altered (Ward and Stanford, 1983).

The serial discontinuity concept proposes that in regulated rivers, two measurable parameters of interest are the discontinuity distance (i.e., how far up or downstream a given stream parameter becomes displaced) and the parameter intensity (how much has the maximum value, duration, etc. of the parameter been changed).

Flood Pulse Concept

In contrast to the RCC, which predicts a downstream decreasing influence of the riparian zone, the flood pulse concept (Junk *et al.*, 1989; Junk and Wantzen, 2004) deals with rivers that interact strongly with the floodplain by rising out of the channel bed. This theory describes the effect of floods on river channel and floodplain in large, unregulated rivers of this type, and thereby incorporates a lateral dimension in the consideration of ecosystem structure and function.

During nonflood periods (see **Figure 5**), the floodplain (which is functionally similar to a large wetland) develops and maintains its own nutrient cycles, and depending on the geomorphic configuration of the floodplain, matter may accumulate in place or enter the stream as described in other models. The pulse of a seasonal flood causes the stream to push out, sending nutrients and river biota over the floodplain. If such a pulse is predictable, then communities of organisms will evolve to take advantage of the pulse because it amplifies resource availability. The flood pulse is postulated to enhance diversity and productivity by structuring the dynamics of plants, nutrients, detritus, and sediments.

During floods, matter that was remineralized in the floodplain is quickly dissolved and made available to aquatic primary producers and production in many of these systems (including the Amazon; Bayley, 1995) exceeds decomposition. As the floodwaters stop rising, decomposition in deeper, slower moving parts of the stream increases, often becoming

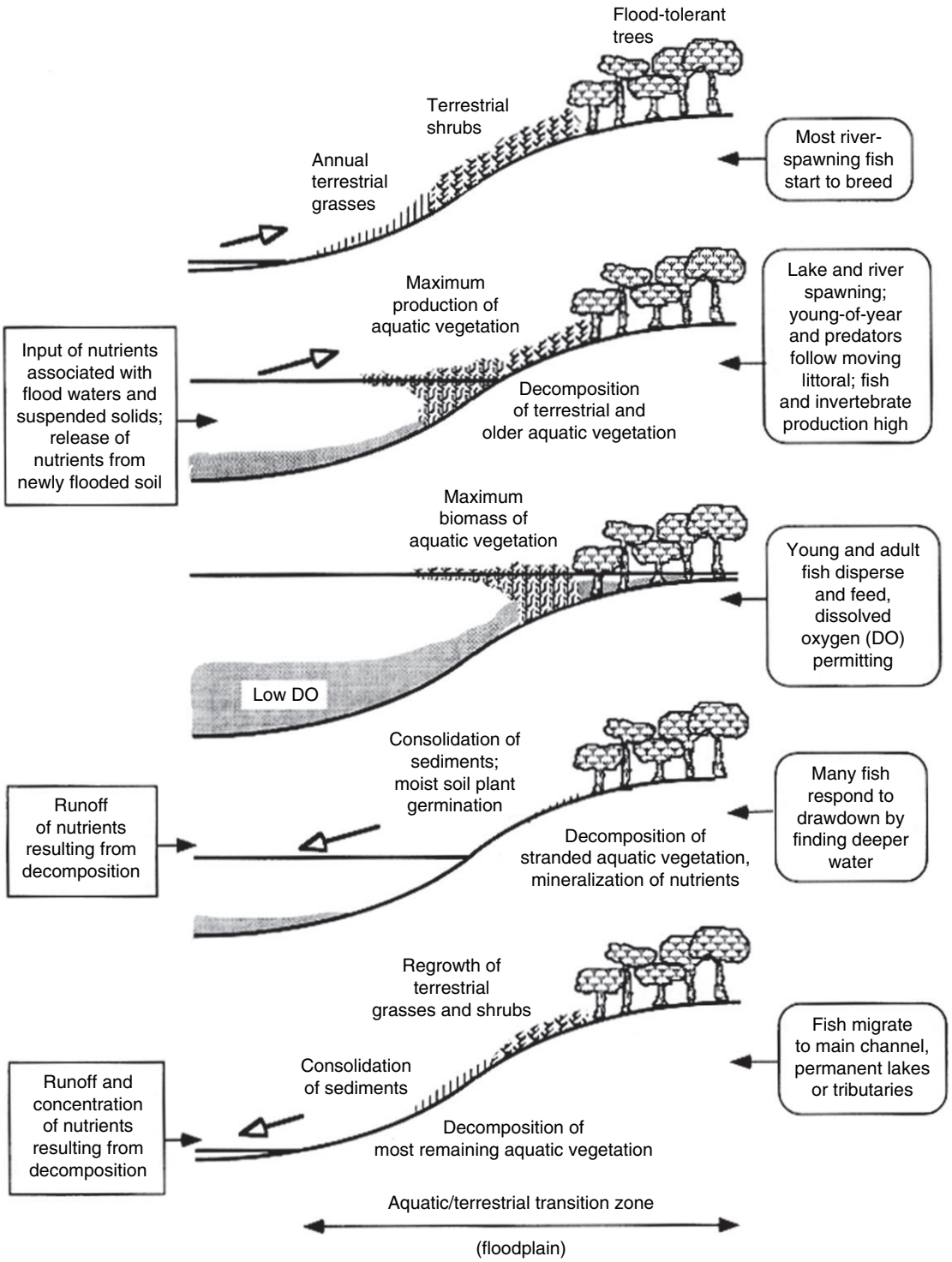


Figure 5 Schematic of the Flood-Pulse Concept. Reproduced from Bayley PB (1995) Understanding large river-floodplain ecosystems. *BioScience* 45: 153–158, with permission from AIBS.

anoxic. Terrestrial plants, in the meantime, shift into the sites of retreating waters and begin to build up terrestrial biomass again.

The moving littoral zone is used by many fish as a nursery for young and by aquatic invertebrates that can utilize shallow environments. Many fish have adapted to the pulse by spawning just before the floods so that their offspring can be on hand when the pulse triggers off the high primary production with associated high secondary production. A useful paradigm in the tropics, it is thought that this kind of floodplain–river dynamic was also common in temperate ecosystems before extensive human alterations occurred.

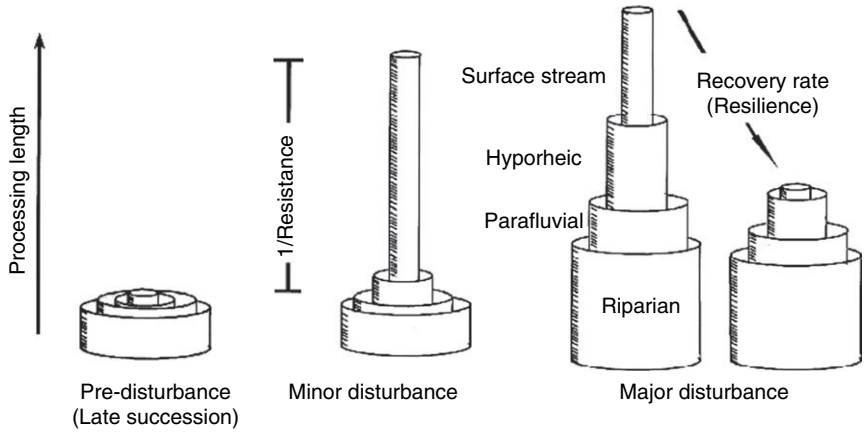


Figure 6 Diagram of the Telescoping Ecosystem Model. Reproduced from Fisher SG, Grimm NB, Marti E, Holmes RM, and Jones Jr. JB (1998) Material spiraling in stream corridors: A telescoping ecosystem model. *Ecosystems* 1: 19–34. The cylinders represent different zones of the system. See text for explanation.

Telescoping Ecosystem Model (TEM)

Fisher *et al.* (1998) presents another way of looking at the holistic, biogeochemical functioning of river systems and their environs, and bears some discussion. In contrast to the RCC that visualizes riverine systems as continua, ;and in contrast to models such as nutrient spiraling, which assumes rivers are simple, uniform chemostats, these authors emphasize heterogeneities in space and time:

> *The landscape is a patchwork of many patch sizes and shapes. Materials follow a large number of routes from uplands to oceans. The path is stochastic and often downhill because the predominant transporting vehicle is water. Because these hydrologic linkages are driven by weather, the path is episodic and jerky. Within patches, characteristic rates of uptake, transformation, and release occur: thus, as the vehicle (water) travels, its load (elements) is adjusted in response to patch-specific material dynamics.* (Fisher *et al.*, 1998, p. 20.)

This perspective corresponds to the understanding of variable source flow in watershed hydrology and takes into account the patchiness and position of patches in the landscape.

As connectors of landforms with lakes and oceans, rivers with relatively fast throughput link land and water units with considerably larger residence times; thus, the connection of land–river–lake or land–river–ocean is (to some extent) slow–fast–slow in terms of biogeochemical movements or high–low–high in terms of retention. Because of the difference in time constants among these units of the landscape, Fisher *et al.* (1998) focus on the landscape factors that influence the ability of rivers to retain matter.

River systems are composed not only of the surface flow channel, but also of several adjacent parts (see **Figure 6**). The stream is embedded in a saturated zone called the hyporheic (underneath the flow) zone and around this is a zone often seen as gravel bars at the surface, with underlying saturated sediments termed the parafluvial. The riparian (bank side) zone beyond this forms a larger buffer between the stream and upland areas and also overlies saturated soils. These zones are often readily distinguishable, particularly in arid landscapes. Fisher *et al.* (1998) term the spatially explicit arrangement of these units as the configuration of the system, and this can affect the rates of transfer, uptake, processing, and release of materials.

Similar to the spiraling length index developed by Webster (1975) and elaborated on by Elwood *et al.* (1983), Fisher *et al.* define a parameter they call the processing length or the length of subsystem required for biogeochemical transformation of some substance (note that this is left intentionally vague). The key element that is different about processing length versus spiraling length is that the former applies to all these zones (stream channel, hyporheic, parafluvial, and riparian) and that these characteristic processing lengths are expected to be quite different. For example, the NO_3 concentration in the authors' study stream, Sycamore Creek in Arizona, is controlled by algal uptake rate (the predominating *N*-transforming process), with a processing length of 42 m. In the parafluvial zone, nitrate is controlled by nitrification (due to microbial decomposition) and has a shorter processing length of 13 m.

The RCC posits that classical ecological succession does not occur in river systems, because the physical constraints dominate the development of ecological communities. Fisher and colleagues, who work in desert ecosystems where rains are infrequent and episodic, view river systems as going through succession, with processing lengths becoming shorter over time, until an episodic disturbance (flood, fire, dumping toxicants) resets the system.

Fisher *et al.* (1998) liken their observation of stream system zonation, and the dynamic differences in processing lengths within each, to the cylinders of a telescope (hence the model name). In particular, with respect to disturbance, they hypothesize that

1. When a stream system is disturbed, the different subsystems will alter their processing lengths, but differently. The resistance of the system, or how little it changes when

perturbed, is inversely proportional to processing length change and will be greatest away from the center of the telescope; and
2. The resilience of the stream telescope (how fast the system returns to the previous state) decreases away from the center.

These two emergent properties of the TEM are shown in Figure 6. However, cross-links (interactions between the subsystems) are proposed to enhance the resilience of the system. Thus, the TEM is a model that can not only predict the rates of system change, but also maintains a hierarchical structure.

Riverine Ecosystem Synthesis (RES)

Similar to the TEM is the RES proposed by Thorp *et al.* (2006). An attempt to synthesize and generalize the multitude of lotic paradigms, the RES identifies hydrogeomorphic patches – places where flows and geomorphology produce particular characteristics – that are arrayed throughout catchments and drive ecosystem dynamics. The RES emphasizes the importance of connectivity to terrestrial parts of catchments as well as temporal influences.

Geomorphic–Trophic Hypothesis and Other Geomorphological Controls on Ecology

Recent work in relatively undisturbed ecosystems has demonstrated how topography and landscape position can influence not only the biogeochemistry of waters within a catchment, but also the invasibility of aquatic habitats by organisms, which sets the stage for subsequent ecological interactions. At a Long-Term Ecological Research (LTER) site in northern Alaska, researchers found that the fish communities of lakes, connected by streams, are determined by such factors as basin steepness (which determines waterfalls that may not be passable by certain fish species), lake depth and size, etc. (Hershey *et al.*, 1999). This results in a limited number of fish communities, and the species present appear to regulate the presence and abundance of other organisms (invertebrates and phytoplankton). Within this simple, Arctic ecosystem, the geomorphic–trophic hypothesis provides an insight into how ecological communities develop, and in particular, how basin geomorphology can constrain the mix of species. These mechanisms likely are at play in many high-elevation systems that are relatively undisturbed. Even if natural manifestations of this hypothesis may be difficult to observe in systems subject to human alteration, the fact that rivers are engineered in various ways does, in fact, promote some of the predictions of this hypothesis.

Riverine Biota

Microbiota

Microbiota are widely recognized as being the most abundant and diverse organisms on earth. One cubic centimeter of water can hold more than 10^9 bacteria. Microbes are responsible for most of the biogeochemical processing of nutrients and organic matter in rivers. Some important functional groups of bacteria are nitrifiers (which convert organic N into nitrate and nitrite), the denitrifiers (which convert nitrate and nitrite back into atmospheric N), sulfate reducers, and methanogens (which convert carbohydrates into methane). Microbes can be free-swimming or attached to particles; they are found in the hyporheic and riparian zones as well as in the water column itself. Microbial diversity, although recognized as vast, is difficult to assess, but new techniques involving DNA are improving the process of identification and classification.

Plants

Riverine plants fall into two general groups: the littoral, mostly rooted vegetation (macrophytes or literally, large plants), and the microscopic, free-living phytoplankton as well as attached microalgae. Macrophytes can be further divided into those that send their leaves up above the water surface (emergent macrophytes) and those that photosynthesize within the water column (submersed macrophytes). A number of macrophytes are also free-floating, such as the water hyacinth (*Eichornia crassipes*), which has enlarged, buoyant leaf stems. A number of algal species are also macrophytic.

Aquatic vascular plants have a number of special adaptations for dealing with the conditions imposed by the aquatic environment. Many plants have roots that extend into the water-saturated sediments, and as a result, are faced with an anoxic environment. To avoid root anoxia, many species have evolved air spaces (aerenchyma) in their roots and shoots that conduct oxygen down from the leaves. An example is the American water celery (*Vallisneria americana*), which translocates so much oxygen to its roots that it actually oxidizes the surrounding sediments, which in turn causes iron to precipitate and form concretions of iron oxides on the roots.

Macrophytes that grow in the brackish or salty parts of rivers have more fresh water within their cells than in the surrounding water; thus, the osmotic concentration gradient tends to force water out of the plant into the water. To avoid dehydration in this manner, vascular plants have a number of adaptations for maintaining their osmotic balance, including specialized cells that either serve as barriers to salt intrusion or glands that actively excrete salt. In mangroves (*Rhizophora* spp., *Avicennia* spp.), for example, the root endodermis is highly salt resistant and serves as a salt filter, and leaf cell membranes also serve as filters by excluding sodium and selectively transporting potassium. The emergent salt-marsh plant, *Spartina*, excretes salt crystals from specialized glands in its leaves.

Phytoplankton, the microscopic plants that float in the water column, rely on turbulent flow to remain high enough in the water column to photosynthesize. In this photic zone, phytoplankton can fix carbon and reproduce, but below this zone, they can only respire away their carbon. In shallow rivers, this is usually not a problem, as the entire water column may be photic, but larger rivers usually become turbid due to the downstream transport of sediments, organic detritus, etc.

The main groups of phytoplankton are the green algae (colonial greens, diatoms, desmids, and others), blue–green algae or cyanobacteria, and dinoflagellates (which are actually both autotrophic and heterotrophic). Diatoms have hard

exteriors, called frustules, which are composed of silicate; thus, they cannot persist if this mineral is in short supply. They are an important taxonomic group in rivers, and often, the riverine populations are maintained by inputs from the tributaries. Algae in small streams often grow directly on rocks (epilithic algae) and can be an important source of organic matter in open (not covered by tree canopy) river segments. A number of microalgae, including diatoms, can grow on the surfaces of attached plants. These may be disproportionately important in riverine food webs. Blue–green algae, such as *Anabaena* or *Oscillatoria*, can fix N from the atmosphere by means of specialized cells called *heterocysts*; they will do this if there is an excess of available P, such as from cultural eutrophication.

Macrophytic algae are attached to hard substrates and are common in streams or along the littoral zones of larger rivers. Some, such as *Cladophora*, grow as filaments, whereas others (*Ulothrix*, *Oedogonium*) grow in tufts; *Chara* and *Nitella* are branched and their cell walls are reinforced with calcium carbonate. The filamentous green alga *Cladophora glomerata* is a cosmopolitan species that does particularly well in the presence of elevated nutrients, and therefore, large colonies serve as bellwethers of cultural eutrophication.

In a review of phytoplankton diversity from 67 rivers around the world, Rojo *et al.* (1994) found that the average species richness was 126, and about half of the species were reported as sporadic. The average species richness was higher in temperate rivers (130 species) than in the tropics (75 species). The total numbers of species enumerated, by family and zone, are given in **Table 1.**

Although the tropical rivers may have been underrepresented in the survey, there are some clear trends. Diatoms are the most diverse taxon and predominate in temperate systems, whereas desmids are of greater importance in tropical rivers.

Invertebrates

The invertebrates that live in rivers can be divided into two groups: the zooplankton (animals that live suspended in the water) and the zoobenthos (animals that live in or on surfaces such as the river bottom, plants, etc.). River currents usually are much faster than the swimming speeds of the zooplankton; thus, washout downstream is a major problem for river zooplankton. Consequently, zooplankton are common in rivers only where their growth rates (or inputs from nearby lakes and wetlands) exceed loss rates from washout and other losses (e.g., predation). Thus, river zooplankton is usually dominated by rotifers and small cladoceran crustaceans, which have high growth rates. Zooplankton densities often are the highest when the water is warm and low, when the growth rates are high and the washout rates are low. The biomass of zooplankton in the main channel of rivers is usually much lower than that in lakes, but off-channel habitats such as oxbow lakes may support rich zooplankton populations. River zooplankton often includes a few dozen species of rotifers, cladocerans, and copepods, fewer than 10 of which are abundant. At least some of these species probably are sustained by inputs from nearby habitats that are more hospitable to zooplankton, such as backwaters, impoundments, or lakes. Most riverine zooplankton are neither specialized nor endemic: the same species and genera of zooplankton that live in rivers are also common in lakes, and occur in rivers worldwide. Despite its low density and richness, river zooplankton plays an important role as food for young of many species of riverine fish, as well as adults of a few specialized zooplanktivorous fish such as paddlefish, *Polyodon spathula*.

By comparison with the zooplankton, the riverine zoobenthos is often rich and dense. A typical large river contains several hundred species of benthic invertebrates representing a biomass of 1–100 g m^{-2} (ash-free dry matter). Dominant groups include insects (chironomid midges and many others), bivalves, prosobranch (gill-breathing) snails, tubificid oligochaetes, and (in tropical rivers) crabs and prawns. As is the case with zooplankton, most of these species are not associated with the open waters of the main channel, but with shorelines, backwaters, vegetation, and other structurally complex habitats along the river margin (**Table 2**). Unlike the zooplankton, many zoobenthic species are found only in rivers and are endemic to particular river systems. Thus, the rivers of southeastern North America contain several hundred species of mollusks and crayfish that are found neither in American lakes nor in rivers elsewhere in the world. The local density and species composition of riverine zoobenthos are governed by the

Table 1 Species richness of riverine phytoplankton. Temperate and tropical percentages indicate the taxon representation in temperate and tropical planktonic assemblages, respectively

Family	*Total number of species*	*Temperate (%)*	*Tropical (%)*
Chlorophyceae	385	28	27
Chrysophyceae	43	3	3
Cryptophyceae	25	—	—
Cyanobacteria	141	10	9
Diatomophyceae	408	40	25
Dinophyceae	27	—	—
Euglenophyceae	112	7	4
Xanthophyceae	11	—	—
Zygophyceae (desmids)	221	9	29

Table 2 Number of species of zoobenthos living in the main channel and backwaters of the River Danube in Austria

Taxonomic group	*Main channel*	*Dammed backwaters*	*Forested backwaters*	*Total*
Porifera (sponges)	0	1	2	2
Hydrozoa (hydras)	0	0	2	2
Turbellaria (flatworms)	9	8	15	23
Mollusca (snails and mussels)	32	42	53	62
Annelida (worms)	28	57	40	69
Crustacea	11	10	27	28
Insecta	224	234	536	703
Other	2	2	8	8
Total	306	354	683	897

nature of the bottom (particle size, organic content, stability), flow conditions (current speed, frequency, and size of floods and droughts), the availability of food supplied by phytoplankton, rooted vegetation, attached algae, and the watershed, and biological interactions such as predation and competition. According to the River Continuum Concept, large rivers should be dominated by filter-feeding animals (both benthic and planktonic) and burrowing deposit-feeders. Data to test this idea in large rivers have not been collected, but it appears that large rivers actually contain a rich variety of feeding types, including filter-feeders, deposit-feeders, predators, algal scrapers, and herbivores that eat rooted plants.

Like the zooplankton, the zoobenthos is an important source of food for riverine fishes. Insects and crustaceans appear to be especially important as fish food, perhaps because many species are active and live above the sediment–water interface, where they are accessible to fish. Benthic filter-feeders (especially bivalves) are sometimes so abundant that they control the number and kind of phytoplankton and other suspended particles in rivers. Through their burrowing and feeding activities, benthic animals may also affect sediment mixing and the exchange of substances (e.g., nutrients, toxins) across the sediment–water interface. Also, humans use riverine invertebrates as food (e.g., prawns, crabs, crayfish, mussels, oysters), bait (crayfish, large insects), or ornament (e.g., pearly mussels, whose shells and pearls have been collected for thousands of years).

Human alteration of the hydrology of large rivers, introduction of toxins and exotic species, and destruction and isolation of the lateral habitats that are important to both zooplankton and zoobenthos have had strong effects on the biodiversity and function of riverine invertebrates. Because of their considerably greater degree of habitat specialization and endemism, the large-river zoobenthos has been affected more severely than the zooplankton, with hundreds to thousands of species now extinct or endangered.

Fish

Fish compose more than half the known species of vertebrates. At present, around 29,000 valid species of fish have been described, but the actual number is likely closer to 32,000 (Nelson, 2006; Froese and Pauly, 2011). Of these, approximately 12,000 species live strictly in fresh water and another 500 species use freshwater habitats over some part of the life cycle (recall that fresh water composes only 1% of the world's water!). Many of the freshwater resident species, and all of the occasionals, use rivers for part or all of their lives.

Riverine fish faunas can be very depauperate, as in small, oligotrophic headwater streams, or they can be highly diverse as in many large, tropical rivers. Fish occupy virtually every heterotrophic ecological niche, from parasite to detritus-feeder, to herbivore, planktivore, insectivore, and piscivore (fish-eater). Large, predatory fish can also consume other organisms, such as small mammals, herps, and birds. In the tropics, where seasonal floods cause rivers to spread out over the floodplain (*see* Flood Pulse Concept, above), many fish species time their reproduction just before the floods, so that their offspring use the flooded forests as nursery habitat. Many tropical species also directly consume leaves, fruits, and seeds while foraging in the flooded forest.

Riverine fish species are adapted to flow regimes that are considerably more turbulent than lakes or oceans. In the riffles and rapids of lower order streams, many fish spend much of their time in patches of quiet waters, such as under rocks, behind boulders or tree roots, or along the riparian banks, emerging briefly to feed. The larvae of freshwater lamprey (ammocoetes) actually live under the stream, in the hyporheic zone, for up to six years. In higher order streams, current velocities are slower, and fish may swim more freely in the water column. An exception is in tidal estuaries, where currents at peak flood and ebb may be as high as a meter per second or more. Some fish apparently ride the tides selectively, entering a tidal current to move either up- or downstream, and then moving out of the flow to remain in place. This may be a common strategy for the fragile larvae to move upstream in tidal rivers (cf. Morgan, 2006).

A fairly small, but important, group of fish undertakes large-scale migrations between fresh and marine waters. These so-called diadromous fish number about 200 species, most of which are anadromous or freshwater-spawning. This group includes many of the salmons, the southern-hemisphere galaxiids, and many herrings. These species use the freshwater environment to spawn; the young rear in the relatively protected riverine or estuarine environment and subsequently migrate to sea to feed, grow, and mature. A few species, notably the freshwater eels (Anguillidae), have the opposite strategy (catadromy), spawning at sea and then moving into fresh water to mature. Diadromous fish are important in many fisheries, as their spawning migrations tend to concentrate their numbers and make them relatively easy to catch. Unfortunately, this group of fish is under a multitude of threats from overfishing, habitat loss or alteration, and pollution. In the North Atlantic basin, of the 22 major species, all but one (striped bass, *Morone saxatilis*) have declined severely and one (the houting, *Coregonus oxyrhinchus*) was recently declared extinct (Limburg and Waldman, 2009).

Studies of riverine fish diversity on a global scale reveal that species richness increases with the size of the drainage basin and with river discharge (note – discharge and drainage size are closely correlated in unimpounded rivers). Tropical rivers also tend to be more speciose, and this is in part due to warmer water temperatures, which stimulate primary productivity (**Figure 7**). The Amazon River, being both the largest river and tropical, supports close to 4000 species. Some of the high-diversity groups in tropical rivers include the catfishes (Siluriformes), the minnows (Cypriniformes), and characins (Characiformes).

Other large animals

Although the numbers of other animal species that use rivers are far smaller, they can be important components of these ecosystems. Among the amphibians and reptiles are frogs and water snakes, many turtles, and the Crocodylia (alligators, crocodiles, and allies). Shoreline bird species are numerous, including herons, ibises, cranes, and top predators such as eagles and ospreys. Mammals, such as beaver (*Castor* spp.),

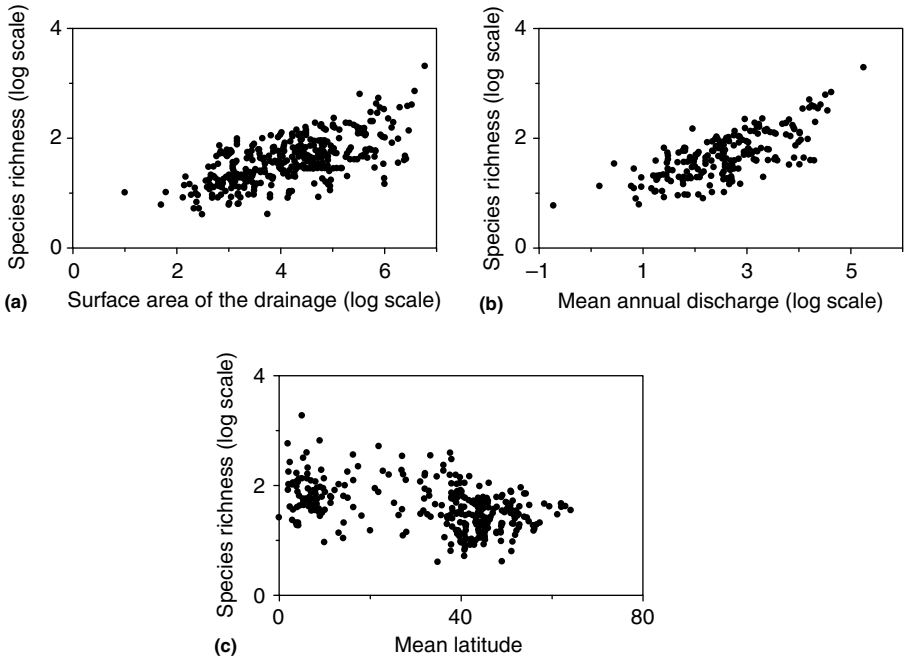


Figure 7 Patterns of riverine fish species richness, as functions of (a) drainage basin size, (b) mean annual river discharge, and (c) latitude. Note the log scales. Reproduced from Oberdorff T, Guõgan J-F, and Hugueny B (1995) Global scale patterns of fish species richness in rivers. *Ecography* 18: 345–352.

otter (*Lutra* spp.), and hippopotamus (*Hippopotamus amphibius*), function as keystone species in many systems, either by virtue of their physical alteration of the habitat (beavers, hippos) or trophic influence (otters). Beavers build dams, which create ponds that alter the flow characteristics, and hence the processing of organic matter, in small- to medium-sized (first to fifth order) streams in North America and Europe. Hippos, by creating aquatic highways between their daytime resting areas and night-time foraging grounds, can cause drainage alterations that can be seen in satellite imagery (McCarthy *et al.*, 1998). Some tropical Asian rivers support freshwater dolphins and sharks. Further, many large mammals, including many of the threatened or endangered "charismatic megafauna," occur in riparian zones of rivers; for example, flooded tropical grasslands are an important riparian habitat for such animals.

Threats to Biodiversity

Humans have always modified rivers and their watersheds, both deliberately and inadvertently. Human settlement and development transforms the landscape from its natural patterns of vegetative cover to networks of roads, drainage works, farms, and urban areas aimed at efficient transportation, food production, and sanitation. In the last century, technological advances have enabled the transformation of the landscape on an unprecedented scale. The effects of human activities now even threaten to modify climatic factors. The effects of these transformations on the resident species of the watershed have been underestimated in the past, largely due to ignorance. **Table 3** reviews some examples of the effects of human activities on riverine ecosystems.

Dams

In many areas, dams are constructed for flood control, water supply, and hydroelectricity. To keep up with world demand for water and energy, dams are being constructed at an unprecedented rate. Approximately 40,000 large dams have been constructed in the past 50 years, and the total area of land flooded by large dams exceeds 4,00,000 km^2 – an area the size of California. In the US alone, there are nearly 20,00,000 small dams and 68,000 dams two stories (6 m) or higher. Additionally, many other small structures (e.g., culverts) disrupt biological connectivity as well.

The most obvious ecological effect of a dam is to act as a barrier to water flowing downstream. Sediments are no longer transported downstream, but rather accumulate behind the dam. For example, before the construction of the Aswan Dam in the 1960s, the Nile River in Egypt transported more than 120 million tons of sediment to the Mediterranean annually, and 10 million tons were dispersed in the floodplain and delta, providing fertile soils that were farmed for millennia. The dam now retains 98% of the sediments, which cut off the vital mechanism of soil renewal, causing a decline in floodplain agriculture and also resulting in coastal erosion.

In ecological terms, a dam is a most effective barrier for migrating fish and other species, and the presence of dams is known to eliminate or drastically reduce the populations of these species unless mitigating structures (e.g., fish ladders) are put in place. Even these are only partially successful: in many rivers with hydroelectric dams, the passage of fish to upstream spawning grounds is a one-way ride, for most of these fish are killed as they attempt to move downstream and pass through the turbines. In large rivers with multiple mainstem dams, fishways may also fail because fish cannot negotiate several

Table 3 Threats to biodiversity in rivers

Human impact	*Proximate effects*	*Consequent short-term change in*		
		Productivity	*Dominant species*	*Species richness*
Damming	Reduction and change in temporal patterns of flow, sediment flux, turbidity	Decreases downstream associated with nutrient reduction	Shifts associated with changes in nutrient ratios (e.g., N : P, Si : N), flow reduction, and barriers to migration	Probable decline, particularly in migratory species
Stream channelization	Increase in flow velocity; disruption of benthos, hyporheic zone	?	Species shifts associated with altered flow regime	?
Nutrient loading	Fertilization of productivity; fueling of increased respiration	Increase in GPP, respiration; BOD	Species shifts associated with changes in nutrient ratios, oxygen levels	Declines associated with anoxia; potential collapse of some communities
Toxic substance loading (heavy metals, pesticides, toxic organics)	Increased mortality of resident species; reproductive failure	Probable reduction in productivity in response to toxicity	Shifts to tolerant species	Probable short-term decline, but possible long-term increase
Introductions of alien species	Colonization of ecosystem by new competitors	?	Potentially dramatic shifts if the introduction is successful	Potentially dramatic decreases as the new species become established
Land use change (logging, agriculture, urbanization)	Typically, increased sediment and nutrient load; hydrological alterations; possible increases in loading of toxics	Probable shift toward heterotrophy in response to nutrient additions and light reductions	Shifts associated with the first four impacts above	Potential decreases associated with the first four impacts above
Overharvesting of species	Depletion of target species; perturbations in predator and prey populations	?	Shifts to nontarget species; other food-chain-dependent shifts	?
Climate change	Changes in temperature, precipitation, evaporation, and atmospheric CO_2	Increases or decreases, depending on the location	?	?

BOD, Biological oxygen demand. GPP, Gross primary production.

dams to reach spawning grounds. Many fishways are also inappropriately designed to pass certain species, such as sturgeons and eels.

Less obvious is the effect of dams on flow regime and biogeochemistry. Dams alter the temporal pattern of river discharge as well as the flow pattern of groundwater. Nutrient and sediment fluxes are typically diminished downstream because of settling or biological uptake in the impoundment. The resulting changes in elemental ratios (e.g., the decrease in available silicate compared with N and P; Humborg *et al.*, 2000) downstream have been associated with species shifts, reductions in productivity, and decreased yields in the downstream fishery.

Stream Channelization and Interbasin Connection

Stream channelization is designed to improve the navigability of rivers or to reduce the flooding potential in streams. In the former case, the river channel is deepened and widened by dredging, which destroys the benthic habitat. In the latter, the streambed is sometimes straightened and paved to increase the capacity of the stream to transport water downstream. These processes alter the flow regime of the stream, favoring species that tolerate faster, turbulent currents, and excluding others. Changes to the streambed can affect the conductivity of water through the hyporheic zone, which affects nutrient processing (*see* Different Conceptual Models of Riverine Ecosystems).

Additionally, many waterways have been engineered to enhance transportation and trade. Such activities have been undertaken for millennia, such that many alterations have become part of the cultural landscape and heritage. The Mississippi River is a case example, where locks permit barges to change elevation and canalways connect across basins. Similarly, the Saint Lawrence Seaway raised water levels above the natural rapids so that commercial shipping vessels could travel

from the ocean into the Great Lakes. The effects in both systems have included habitat loss, ecosystem alteration, major changes in the hydrograph, and species invasions (*see* Nutrient Loading).

Damming for water supply purposes (*see* Dams) can also result in interbasin transfers of water, nutrients, and biota. In periods of high water demand, which often coincide with periods of low water in rivers, riverine water withdrawal or detention and transfer to municipal water supply networks or irrigation systems can significantly modify flow regimes. For example, the New York City water reservoir system relies in part on dams in the Upper Delaware basin, effectively transferring water from the Delaware to the Lower Hudson River basin via the City's distribution network. In China, the South–North Water Transfer Project is the largest water-works project ever developed there. Plans are to transfer nearly 45×10^9 m^3 of water annually from the Yangtze River in the south to the Yellow River via several routes. This will result in massive changes in water supply, habitat, and fisheries, and will affect communities both downstream and in headwater areas (i.e., across international borders).

Nutrient Loading

As discussed above (*see* Energy and Matter Transformation (Biogeochemistry)), nutrient loading fuels the ecosystem. Excessive loading results in a catastrophic collapse when the resulting increase in metabolism exceeds the available oxygen supply of the river. Cultural eutrophication of rivers is severe to dire in some parts of the developing world, in particular in India, Southeast Asia, and Africa, as well as around many urban centers worldwide. Eutrophication of major watersheds such as the Mississippi draining into the Gulf of Mexico, the Scheldt and the Rhine draining to the North Sea, California's Sacramento River draining into Suisun Bay, or the many eutrophied rivers draining toward the Baltic Sea produce or exacerbate hypoxic or anoxic episodes with increasing frequency.

Toxic Substances

Associated with pesticide application in agriculture (or suburban gardens), industrial effluents, and chemical spills, have a spectrum of effects, all deleterious to organisms. Some herbicides can drastically reduce productivity by direct mortality of aquatic plants. Other chemicals kill sensitive species higher in the food chain. Subtler effects include behavioral changes (due to endocrine-disrupting compounds) in predators or prey species, which can shift the structure of food chains, increase susceptibility to disease, and cause infertility. What complicates this story is that different organisms do not respond uniformly to exposure to a particular toxic substance, so that the effect on a particular riverine system is determined largely by the mix of species present. Chronic exposure to a stream of toxic material will gradually shift the species present to those that can tolerate it, generally resulting in a decline in species richness.

Overharvesting Species

In addition to humanity's indirect impacts on biota, humans threaten biodiversity directly through the overharvesting of species. In the US and Canada, a number of river-associated fisheries are currently on the decline, including those for several species of Pacific salmon, Atlantic salmon, American shad, sturgeons, eels, and oysters. Worldwide, fishing pressures in rivers are enormous. Many river fisheries in developing countries are artisanal, providing food for a local community, and little is known about harvesting pressures. However, responses of fish communities to overharvesting are becoming understood. For instance, in one West African river, catch per unit of fishing effort remained relatively constant from the 1950s through the 1990s, but the trophic structure of the harvested fish altered from large piscivores to small planktivores, with loss of food web complexity (**Figure 8**; Welcomme, 1999). This phenomenon is called fishing down food webs, and was first noted in marine fisheries (Pauly *et al.*, 1998). At the population level, overharvesting often selectively removes older age classes, pushing the age of first reproduction down, and often skews the sex ratios in the population (e.g., female sturgeon are selectively sought for their roe from which caviar is produced).

Introduced Species

Humans are also introducing new species into ecosystems at an unprecedented rate. Some of these introductions are intentional: for example, salmon, trout, and bass are stocked into rivers and lakes all around the world for sport fishing, and grass carp are introduced to control aquatic weeds. Many introductions are unintentional, arising from the pumping out of ballast water in large tankers, overland transfer of boats with attached organisms from one drainage to another, or accidental spills. Although most organisms that are introduced into a novel environment perish, the increase in commercial boat traffic particularly increases the likelihood of a successful introduction.

The rates of exotic species introductions are difficult to quantify, in part because many of the species are too small to notice. However, Mills *et al.* (1996) documented the rate of species introductions into the Hudson River, in New York State; since 1840, the rate has been about one successful introduction per year – much higher than natural rates of species flow.

Direct perturbations of the mix of species, either by introduction of new species, or removal of extant species by overharvesting, may cause drastic changes in the relations of the food web, thereby also affecting predators and prey. Some species additions affect their physical and chemical environment as well: in North America, zebra mussels, which were introduced into the Great Lakes in 1988 and spread rapidly through the Mississippi and eastern drainages, have been documented to increase the clarity of the water column by virtue of their filter feeding, and also carpet large areas of the benthos, where they have successfully colonized.

Assessment and Management

River ecosystems are clearly important and are at risk. Their direct economic importance to societies includes their use in transportation, water supply, energy, provision of harvestable products, and they are of cultural and spiritual importance. Their indirect importance, sometimes termed

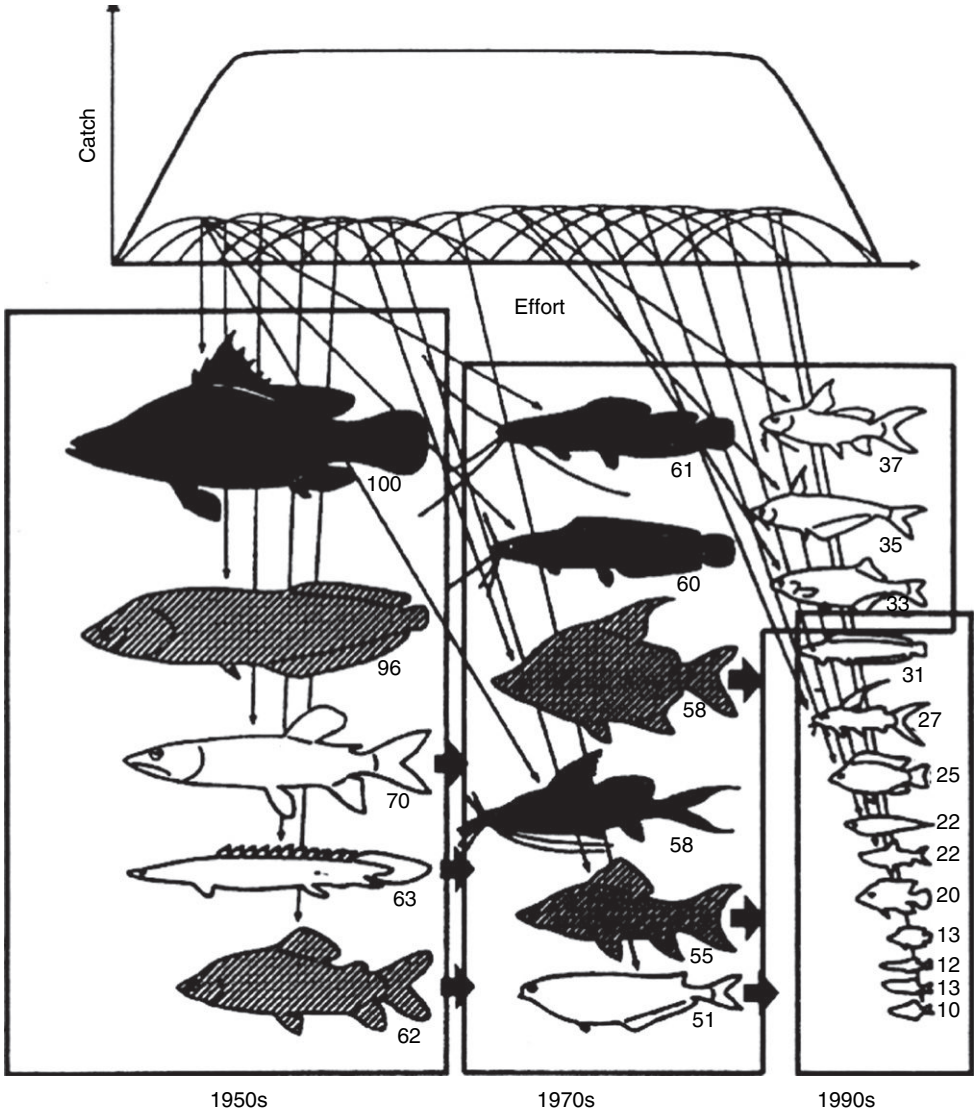


Figure 8 Example of a fished-down riverine food web, Oueme River, West Africa. Diagram shows how fish catches changed from larger fish to smaller and smaller target species over time. Fish colored in black disappeared before 1965, grey are fish whose numbers were seriously reduced. Arrows indicate species which reduced the age at first reproduction. Note that overall catch per unit of effort remained largely constant. Reproduced from Welcomme RL (1999) A review of a model for qualitative evaluation of exploitation levels in multi-species fisheries. *Fisheries Management and Ecology* 6: 1–19, with permission from Wiley.

essential ecosystem services, includes their fundamental role as biogeochemical transformers of energy and matter, physical transformers of the landscape (shaping the land through fluvial processes), and the provision of a wide variety of ecological habitats along the river continuum.

Historically, the assessment of river ecosystems was hindered by two factors. First, on a practical basis, it is difficult to adopt a single methodology to assess the state of a large river system from headwaters to the mouth. Many of the methods that were developed for small stream study cannot be easily adapted for the higher-order parts of the system, and at the other end of the continuum, oceanographic methodologies may be designed for scales of study that are larger than rivers. Second, until the development of scale-spanning river paradigms such as the River Continuum Concept, ecological studies of rivers tended to consider a river in isolation from its watershed, but as is now widely recognized, upstream processes in the watershed (including human-accelerated land-use change) have major downstream effects on rivers. (A third difficulty may be due to the somewhat arbitrary tendency of researchers to specialize on broadly different parts of the system: there are many more small-stream and estuarine scientists than there are river scientists.)

The revival of the concept of ecosystem health has proven to be a useful point of departure for the current study and management of river ecosystems (see, e.g., Karr and Chu, 1999 and Simon, 2003). Metrics of ecological communities, such as those that compose indices of biotic integrity (IBI, see Karr and Chu, 1999), can be combined with in-stream biogeochemical measurements (e.g., processing lengths of critical

chemical cycles, ecosystem metabolism, or food web relationships as identified by tracers), in-stream habitat assessments, and whole-watershed metrics (primarily land use, soils, geology, topography, etc.) to provide an integrated assessment of the status and threats to rivers. An emerging consensus is that assessment and management no longer fall solely in the domain of scientists and managers, but must involve those who share a stake in the management process. This is evident in the growth of watershed management councils, which struggle to deal fairly with the tradeoffs between economic prosperity and sustainability of rivers and lakes in watersheds. The difficulty of finding equitable solutions increases as the scales of political boundaries increase: rivers that share or cross international jurisdictions require a higher level of political cooperation. Maintaining river ecosystems as healthy components of the landscape will be one of the great challenges of the twenty-first century.

See also: Endangered Freshwater Invertebrates

References

Bayley PB (1995) Understanding large river-floodplain ecosystems. *BioScience* 45: 153–158.

Boyer EW, Goodale CL, Jaworski NA, and Howarth RW (2002) Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern U.S.A. *Biogeochemistry* 57/58: 137–169.

Cummins KW, Cushing CE, and Minshall GW (1995) Introduction: An overview of stream communities. In: Cushing CE, Cummins KW, and Minshall GW (eds.) *Ecosystems of the World, Vol. 22: River and Stream Ecosystems*, pp. 1–8. Amsterdam, The Netherlands: Elsevier.

Elwood JW, Newbold JD, O'Neill RV, and Van Winkle W (1983) Resource spiraling: An operational paradigm for analyzing lotic ecosystems. In: Fontaine TD and Bartell SM (eds.) *Dynamics of Lotic Ecosystems*, pp. 3–27. Ann Arbor, MI: Ann Arbor Science Publishers.

Fisher SG, Grimm NB, Marti E, Holmes RM, and Jones Jr. JB (1998) Material spiraling in stream corridors: A telescoping ecosystem model. *Ecosystems* 1: 19–34.

Froese R and Pauly D (eds.) (2011) *FishBase*. Penang, Malaysia: World Wide Web Electronic Publication. www.fishbase.org, version (06/2011).

Hershey AE, Gettel GM, McDonald ME, *et al.* (1999) A geomorphic-trophic model for landscape control of Arctic lake food webs. *BioScience* 49: 887–897.

Hong B, Swaney DP, and Howarth RW (2011) A toolbox for calculating net anthropogenic nitrogen inputs (NANI). *Environmental Modelling & Software* 26(5): 623–633.

Howarth RW, Billen G, Swaney D, *et al.* (1996) Riverine inputs of nitrogen to the North Atlantic Ocean: Fluxes and human influences. *Biogeochemistry* 35: 75–139.

Howarth RW, Swaney DP, Billen G, *et al.* (2011) Nitrogen fluxes from the landscape are controlled by net anthropogenic nitrogen inputs and by climate. *Frontiers in Ecology & Environment*. doi: 10/1890/100178.

Humborg C, Conley DJ, Rahm L, Wulff F, Cociasu A, and Ittekkot V (2000) Silicon retention in river basins: Far-reaching effects on biogeochemistry and aquatic food webs in coastal marine environments. *Ambio* 29: 45–50.

Hynes HBN (1975) The stream and its valley. *Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie* 19: 1–15.

Junk WJ, Bayley PB, and Sparks RB (1989) The flood pulse concept in river-floodplain systems. *Canadian Special Publication in Fisheries and Aquatic Sciences* 106: 110–127.

Junk WJ and Wantzen KM (2004) The flood pulse concept: New aspects, approaches, and applications – an update. In: Welcomme R and Petr T (eds.) *Proceedings of the Second Large River Symposium (LARS)*, pp. 117–149. Pnom Penh, Cambodia. Bangkok: RAP Publication.

Karr JR and Chu EW (1999) *Restoring Life in Running Waters: Better Biological Monitoring*. Washington, DC: Island Press.

Leopold A (1941) Lakes in relation to terrestrial life patterns. *A Symposium on Hydrobiology*, pp. 17–22. Madison, WI: University of Wisconsin Press.

Limburg KE and Waldman JR (2009) Dramatic decline in North Atlantic diadromous fishes. *BioScience* 59: 955–965.

Maidment DR (ed.) (1993) *Handbook of Hydrology*. New York, NY: McGraw-Hill Inc.

McCarthy TS, Ellery WN, and Bloem A (1998) Some observations on the geomorphological impact of hippopotamus (*Hippopotamus amphibius* L.) in the Okavango Delta, Botswana. *African Journal of Ecology* 36: 44–56.

Mills EL, Strayer DL, Scheuerell MD, and Carlton JT (1996) Exotic species in the Hudson River basin – a history of invasions and introductions. *Estuaries* 19: 814–823.

Minshall GW, Petersen RW, Cummins KW, *et al.* (1983) Interbiome comparison of stream dynamics. *Ecological Monographs* 53: 1–25.

Morgan SG (2006) Larval migrations between the Hudson River estuary and New York Bight. In: Levinton JS and Waldman JR (eds.) *The Hudson River Estuary*, pp. 157–170. Cambridge, UK: Cambridge University Press.

Nelson JS (2006) *Fishes of the World*, 4th edn. New York, NY: John Wiley and Sons, Inc.

Oberdorff T, Guégan J-F, and Hugueny B (1995) Global scale patterns of fish species richness in rivers. *Ecography* 18: 345–352.

Pauly D, Christensen V, Dalsgaard J, Froese R, and Torres F (1998) Fishing down marine food webs. *Science* 279: 860–863.

Rodriguez-Iturbe I and Rinaldo A (1997) *Fractal River Basins: Chance and Self-Organization*. Cambridge, UK: Cambridge University Press.

Rojo C, Alvarez Cobelas M, and Arauzo M (1994) An elementary, structural analysis of river phytoplankton. *Hydrobiologia* 289: 43–55.

Russell MJ, Weller DE, Jordan TE, Sigwart KJ, and Sullivan KJ (2008) Net anthropogenic phosphorus inputs: Spatial and temporal variation in the Chesapeake region. *Biogeochemistry* 88: 285–304.

Shreve RL (1966) Statistical law of stream numbers. *Journal of Geology* 74: 17–37.

Simon TP (2003) *Biological Response Signatures: Indicator Patterns Using Aquatic Communities*. Boca Raton, FL: CRC Press.

Strahler AN (1952) Dynamic basis of geomorphology. *Geological Society of America Bulletin* 63: 923–938.

Thorp JH, Thoms MC, and Delong MD (2006) The riverine ecosystem synthesis: Biocomplexlity in river networks across space and time. *River Research and Applications* 22: 123–147.

Vannote RL, Minshall GW, Cummins KW, Sedell JR, and Cushing CE (1980) The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37: 130–137.

Vörösmarty CJ, Sharma KP, Fekete BM, *et al.* (1997) The storage and aging of continental runoff in large reservoir systems of the world. *Ambio* 26: 210–219.

Ward JV and Stanford JA (1983) The serial discontinuity concept of lotic ecosystems. In: Fontaine TD and Bartell SM (eds.) *Dynamics of Lotic Ecosystems*, pp. 29–42. Ann Arbor, MI: Ann Arbor Science Publishers.

Webster JR (1975) Analysis of potassium and calcium dynamics in stream ecosystems on three southern Appalachian watersheds of contrasting vegetation. PhD Thesis, University of Georgia, Athens, GA.

Welcomme RL (1999) A review of a model for qualitative evaluation of exploitation levels in multi-species fisheries. *Fisheries Management and Ecology* 6: 1–19.

Role and Trends of Protected Areas in Conservation

Timothy M Boucher, The Nature Conservancy, VA, USA
Mark Spalding, The Nature Conservancy, Newmarket, UK
Carmen Revenga, The Nature Conservancy, VA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biome Can be defined as a "major regional community of plants and animals with similar life forms and environmental conditions. It is the largest geographic biotic unit, and is named after the dominant type of life form, such as tropical rain forest, grassland, or coral reef." A single biome can be widely scattered about the planet.

Biogeographic realms Describe the largest division of the historic and evolutionary patterns of terrestrial plants and animals. In these regions, plants and animals evolved in the isolation created by barriers such as oceans, mountain ranges, or deserts. A realm can be a single continent or cut across continents.

Ecoregions Distinct ecological units nested within biomes and realms. They are defined as relatively large units of land or water containing a distinct assemblage of natural communities sharing a large majority of species, dynamics, and environmental conditions. Ecoregions represent the original distribution of distinct assemblages of species and communities. Ecoregions tend to have flora and fauna quite different from others. They are the smallest ecological unit analyzed in most protected area studies.

Overview: How Protected Areas Are Defined, Selected, and Created

Definitions

Iconic examples of protected areas, recognized worldwide, include the Grand Canyon National Park, the Serengeti National Park, and the Great Barrier Reef Marine Park. These vast natural areas are set aside primarily for nature conservation. In reality, the term 'protected area' encompasses a far broader array of places, from tiny, locally managed fishing reserves to large landscapes where people and nature live in a tight balance.

The most widely accepted formal definition of a protected area was developed by the International Union for Conservation of Nature (IUCN) in 2008:

> A clearly defined geographical space, recognised, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values. (Dudley, 2008)

This brief definition encompasses four critical facets:

- *Spatial:* Protected areas are tracts of land or sea with defined and fixed boundaries.
- *Set aside:* The terms 'recognized, dedicated, and managed' all point to the need for the site to be treated differently from surrounding areas. Protected areas are often established as legal entities, but this is not required. However, there must be some means of ensuring that these spaces are cared for in a distinct and effective manner.
- *Long term:* Protected areas are seen as long-term or permanent management interventions.
- *Nature:* Cultural values and human use can be central elements of protected areas, but the primary purpose must be nature conservation. Other spatial management regimes for the management of natural resources for forestry, fisheries, agriculture, or other uses, which do not give prominence to nature conservation, are often not regarded as protected areas.

In application, the term 'nature' has been broadly interpreted to include landscapes and seascapes where human activities have considerably altered the original state, but where natural values are still considered as conservation targets.

The IUCN protected area management categories classify, but do not rank, sites based on their intended management approaches (Dudley, 2008).

Ia. Strict nature reserves are strictly protected areas set aside to protect biodiversity and also possibly geological/geomorphological features, where human visitation, use, and impacts are strictly controlled and limited to ensure protection of the conservation values. Such protected areas can serve as indispensable reference areas for scientific research and monitoring.

Ib. Wilderness areas are usually large unmodified or slightly modified areas, retaining their natural character and influence, without permanent or significant human habitation, which are protected and managed so as to preserve their natural condition.

II. National parks are large natural or near-natural areas set aside to protect large-scale ecological processes, along with the complement of species and ecosystems characteristic of the area, which also provide a foundation for environmentally and culturally compatible spiritual, scientific, educational, recreational, and visitor opportunities.

III. National monuments are set aside to protect a specific natural monument, which can be a landform, sea mount, submarine cavern, geological feature such as a cave, or even a living feature such as an ancient grove. They are generally quite small protected areas and often have high visitor value.

IV. Habitat/species management areas aim to protect particular species or habitats, and management reflects this priority. Many management areas will need regular, active

interventions to address the requirements of particular species or to maintain habitats, but this is not a requirement of the category.

V. Protected landscapes/seascapes comprise areas where the interaction of people and nature over time has produced an area of distinct character with significant ecological, biological, cultural, and scenic value, and where safeguarding the integrity of this interaction is vital to protecting and sustaining the area and its associated nature conservation and other values.

VI. Protected areas with sustainable use of natural resources conserve ecosystems and habitats, together with associated cultural values and traditional natural resource management systems. They are generally large, with most of the area in a natural condition, where a proportion is under sustainable natural resource management and where low-level non-industrial use of natural resources compatible with nature conservation is seen as one of the main aims of the area.

Many protected areas do not fit neatly into these definitions. In the United States, for instance, the term 'national monument' refers as much to the manner of creation as to the characteristics of the site. National monuments are created by presidential proclamation, a far simpler process than the creation of national parks or wilderness areas, which require Congressional approval (Hardy, 1998). The Arctic National Wildlife Refuge, classified as IUCN Category IV – Habitat and Species Management – actually functions as a strict nature reserve or a wilderness area because of its remoteness and vast size. At the same time, when Congress approved the creation of this refuge, it expressly allowed the future use of 1.5 million acres for oil and gas extraction. The divergence between assigned category and use varies from one country to another. In many countries, national parks are large tracts of near-wilderness where nature runs its course with little intervention. The legal status of such areas may closely match IUCN Category II but *de facto*, the park, or some parts of it, may resemble strict nature reserves or wilderness (IUCN Category I). By contrast, many of the national parks in Europe are managed landscapes where grazing and agriculture are considered part of a system with considerable biodiversity value and ongoing human activity is integral to the management. Many of these sites are classified as IUCN Category V.

Buffer zones surrounding protected areas are intended to guard against encroachment and mitigate edge effects. Though rarely protected, they may be managed via restrictions similar to those of IUCN Category VI protected areas.

History

Throughout human history, people have set aside land considered to have special natural values. Some ancient sacred sites were protected because of their natural features. Other forms of protection likely arose, planned or otherwise, to prevent overharvesting by restricting access to certain areas for hunting or fishing.

The first historical records of protected areas come from India. Recommendations for protection are made in the Arthashastra texts written in the fourth century BC, and the Emperor Ashoka, in 252 BC, issued the first legal decree for the protection of wildlife and forests (Bhatt *et al.*, 2008). One of the oldest sites with a continuous record of conservation management may be the New Forest in southern England. This site, now a national park, was set aside in 1079 as a royal hunting reserve, first to protect against local harvesting to maintain a rich stock of game and subsequently for timber. Other early utilitarian purposes for conservation included watershed protection, such as the Main Ridge Reserve declared in Tobago in 1761, and the widespread use of closed or taboo (*tapu*) areas to manage fisheries in many Pacific Islands.

The Yosemite Grant by the U.S. government to the State of California in 1864, to protect its scenic, esthetic, and recreational values against commercial development, typifies the shift to a more modern concept. The declaration of the first national park soon followed with the establishment of Yellowstone National Park in 1872, which by law was "reserved and withdrawn from settlement, occupancy or sale … and dedicated and set apart as a public park or pleasuring-ground for the benefit and enjoyment of the people" (Yellowstone National Park Act, 1872). Other countries followed the lead: Australia's Royal National Park was established in 1879, New Zealand's Tongariro National Park in 1894, Sweden's Laponia in 1909, and the famed South Africa's Kruger National Park in 1925 (Chape *et al.*, 2008).

International listing of protected areas began in the early 1960s at the instigation of the United Nations; the first list of 'national parks' included approximately 1000 sites. By 2003 the figure was considerably more than 100,000 sites (Chape *et al.*, 2003). In the marine environment, progress was slower. In 2003, only 0.45% of the oceans were designated as marine protected areas; this figure almost trebled to 1.31% by 2010.

Selection of Protected Areas: Reasons

Justifications for the establishment of protected areas have evolved over time. Preservation of wildlife as a food source later became the interest of sports hunters, and still later, for its own sake or for the maintenance of ecosystem integrity. Similarly, the protection of timber stands for lumber has come to encompass the maintenance of forest ecosystem function. Water resources for agriculture and other human uses also came to be valued for the maintenance of biodiversity. Recreation became, and still is, an important motivation. The nineteenth-century romanticism that thought that protected areas were to remain largely free of human presence gave way to the twentieth-century concept of 'multiple use' and the inclusion of human residence within protected area boundaries.

Concern for disappearing species arose in the nineteenth century and informed the protected area theory of the twentieth century. The preservation of small numbers of animal and plant species in zoos and botanical gardens led to the realization that true survival of species necessitated the conservation of the habitat of those species. Increasing emphasis on nature conservation came from growing awareness that sites and species were increasingly at risk of despoliation or extinction. The first tracts of Wicken Fen, England's oldest nature reserve, were purchased in 1899 to safeguard one of the last remnants of a once-vast wetland that was about to be lost. Today the protection of species and habitats lies at the heart of

many protected area designations, and the desire to build more secure systems of protection has led to objective scientific approaches for selecting individual sites and networks of protection.

Esthetic values continue to play a role, in part because scenic areas are valued as a component of recreation. Utilitarian needs and arguments are undergoing resurgence, with growing desire to use natural systems to secure benefits such as freshwater supplies, coastal protection, and carbon storage and sequestration (Dudley *et al.*, 2010; Mcneely and Mainka, 2009).

Selection of Protected Areas: Methods

Generally

Ecological knowledge is a critical input variable for the design of individual sites and of networks. Understanding species' range sizes, minimum viable population sizes, interactions, behaviors such as dispersal and migration, and biogeochemical processes is critical in determining a size and design for a protected area to allow a healthy population to survive. Long-term observations are needed to understand natural variation over time, and both expected ecological succession and disturbance and the potential influence of climate change must be considered.

The edge effects of reserves that adjoin altered land have been well documented. Highly fragmented habitats such as forests with high margin–area ratios will typically have a more altered and less stable population of forest-restricted species than will single tracts of equivalent area. Together with location, edge effects will also facilitate human access, with fragmented or small patches allowing more opportunities for hunting or other anthropogenic disturbance (Gascon *et al.*, 2000). *Ex situ* threats such as water or air pollution, or the spread of invasive species, are difficult to avoid but larger, unfragmented sites may be less susceptible. Location may be critical in avoiding or minimizing the impact of threats from without. In many freshwater and marine habitats, for instance, upstream influences can greatly reduce the value of small protected areas. Considering upstream threats when selecting sites can make a considerable difference.

In the face of pressing conservation needs and the rapid conversion of natural habitat for human activity, conservation organizations and governments have engaged in numerous prioritization exercises, such as the Nature Conservancy's Conservation by Design (Anon, 2006). Prioritization exercises tend to focus on representation of all habitat types. There is also a degree of flexibility that allows for the acquisition of lower priority sites that will be lost if not acquired immediately.

Protected areas are rarely selected or designed based solely on ecological principles. Governments and private organizations rarely have a totally free hand. Political boundaries and private ownership often limit the acquisition of areas with conservation value. Other constraints include economic limitations and the real or perceived impact on stakeholder communities that may stand to gain or lose from site designation and that can be great supporters or vociferous challengers (Borrini-Feyerabend *et al.*, 2004).

Networks and Connectivity

A desire to achieve more secure conservation outcomes has spurred efforts to build comprehensive and systematic approaches to designation. Ecological processes are the true goal of these systems: capturing spatial patterns of connectivity such as migration and dispersion across land and seascapes, ensuring sufficient population sizes and genetic diversity, and securing against loss due to natural stochastic events or anthropogenic influences by replicating protection in multiple locations. These approaches seek to ensure that a full array of diversity is protected in a 'system' of protected areas.

With this in mind, considerable effort has gone into building connectivity between protected areas. Methods can include the consideration of natural dispersion routes and setting of maximum distances among sites to maximize the likelihood of successful movement among sites. Corridors to facilitate movement among sites may be continuous areas of appropriate habitat or may be discontinuous stretches where 'stepping stone' patches facilitate species movements. In the case of migratory species, the concept of connectivity can be very large scale and can traverse multiple countries and even hemispheres.

Systematic planning efforts trace back to plans for a network of *Zapovedniks* (strict nature reserves) to represent all of Russia's natural diversity in 1917 (Danilina, 2008). At global scales the desire to assess progress and encourage addition of new protected areas to global systems led to the development and use of new biogeographic classification systems (Spalding *et al.*, 2007; Udvardy, 1975) to assess global and regional protected areas coverage. Simple overlays of protected areas with such maps draw attention to deficiencies and often drive new conservation effort. More focused conservation effort, based on species richness, rarity, endemism, and restricted range, attempts to identify key areas for special attention.

International prioritization and coordination efforts are largely for guidance, although both the Ramsar Convention on Wetlands (Ramsar) and the Convention on Biological Diversity (CBD) carry at least some obligations by signatories to develop minimum coverage and to utilize standard measures of importance and of representativeness. The CBD set targets for protection coverage by member states; the Ramsar contracting parties encouraged a unified approach by adopting the Marine Ecosystems of the World classification (Spalding *et al.*, 2007) to encourage global representative systems (COP10, Resolution X.20, Changwon, Republic of Korea).

At national and regional scales, there is increasingly a legal requirement to develop at least a partially systematic protection of nature. The European Union Habitats and Birds Directives, for instance, by law require member states to protect a large network of Special Areas for Conservation and Special Protection Areas. Known as the Natura 2000 network, which as of 2011 includes approximately 18% of the land area of the 27 European Union nations, these sites encompass key habitats and species that were identified through a systematic review. Many sites were already protected, but the process served to identify gaps, whereas the multinational recognition provided under European law added another layer of legal protection, making it harder for landowners or national governments to remove or limit protected status.

There is also a need to embed sites within a broader landscape management context, to reduce conflicts and build synergies with uses and activities in adjacent land and seascapes. Systematic planning tools exist that enable users to visualize social, economic, and biodiversity information together, to build hypothetical protected area systems, and even, with sufficient input variables, to design systems of protected areas that would accommodate all societal needs and maximize human and natural benefits.

Species Protection

The objective of the CBD is to significantly reduce the rate of biodiversity loss. An analysis of the latest IUCN Red List (IUCN, 2011) shows that the number of species classified as threatened is growing, suggesting that this CBD goal is not being met. Many organizations, responding to this ongoing loss, focus on individual species or certain taxa at high risk of extinction and endeavor to create protected areas for these species.

Such is the goal of the Alliance for Zero Extinction (AZE), a consortium of conservation organizations. The AZE in 2005 analyzed the extinction threat for IUCN Red List threatened species that have a single confined range. These analyses covered birds, amphibians, mammals, reptiles, and coniferous plants. The analysis then determined the extent to which the species' ranges were protected. Of the 595 that contained 794 AZE species (updated in 2010 to 920 species), only 203 (43%) were fully contained within protected areas; another 87 (15%) were partially within a protected area (Ricketts *et al.*, 2005) (**Figure 1**). Most sites were found in the tropics, with more than 350 sites found in Tropical Moist Forests. Of the 329 sites in the Americas, only 151 (45%) have some degree of protection; of the 250 sites covering 334 species in Central and South America (including the Caribbean), 40% have some degree of protection.

Formal Creation of Protected Areas

The most common approach to protection entails legal designation, typically by national or local governments, of an area for a particular purpose. This may entail governmental ownership or, in countries such as the United States, governmental management of public lands. It is not unusual for privately owned lands to be incorporated into protected areas, particularly in protected landscapes where existing farming or other land uses, known in the United States as inholdings, continue. In a few countries, notably the United States, effective protection has also been widely achieved through the establishment of conservation easements whereby owners agree to permanent legal restrictions on land use and activities on their land in exchange for significant tax benefits. The property remains in private hands or is held by land trusts (Kiesecker *et al.*, 2007).

Protection associated with private lands and with traditional land tenure enjoys minimal legal safeguards other than land-use zoning, water rights, and prohibition on trespass. Private protected areas are widespread under ownership of nongovernmental organizations, local communities, and commercial hunting or tourism operators (Beck *et al.*, 2004; Bernstein and Mitchell, 2005; Jones *et al.*, 2005). Traditional systems remain widespread, ranging from holy mountains in Mongolia to small fishery reserves across the Pacific Islands. Increasingly, such areas are being strengthened through legal recognition, either of the individual sites (such as the Islamic hema system in many Arabian grazing ranges) or of the communities rights to manage communal lands and waters. The Locally Managed Marine Area Network expanding across the Pacific, although community driven, has benefitted significantly from the development of policy at local, regional, country, and international levels to recognize local ownership and governance of marine resources. The assessment of the

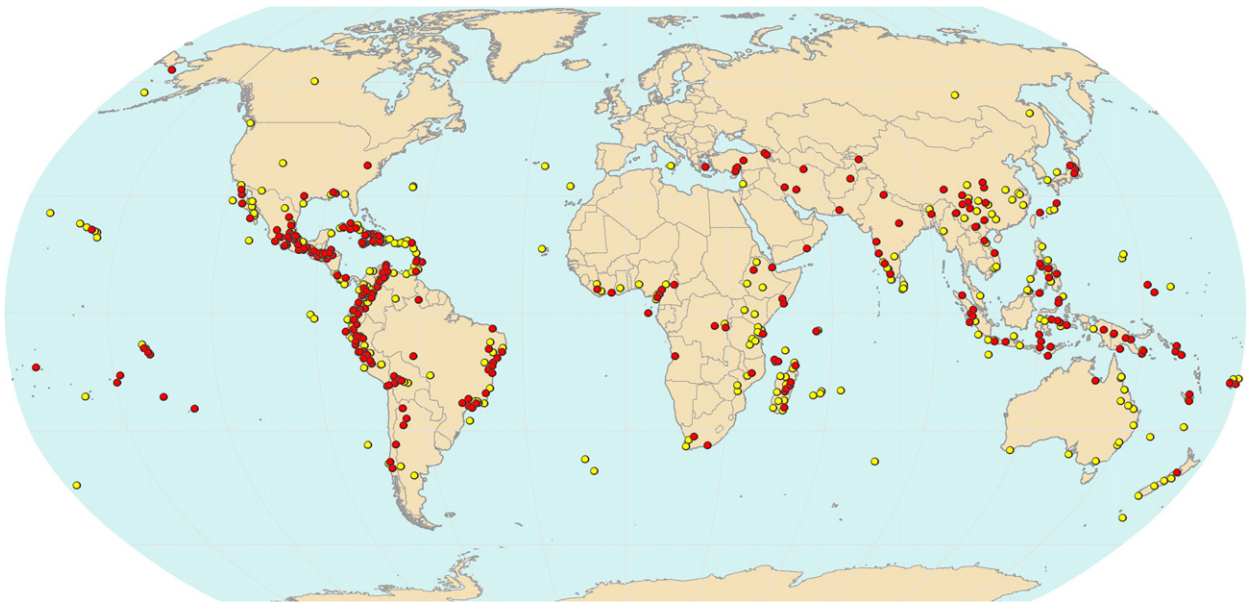


Figure 1 Map of 595 sites of imminent species extinction. Yellow sites are either fully or partially contained within declared protected areas (203 and 87, respectively), and red sites are completely unprotected or have unknown protection status (257 and 48, respectively).

coverage and effectiveness of privately protected lands and traditional systems is challenging, and there is likely considerable underreporting in global data sets, but this is gradually improving.

International Approaches

The patterns and processes of nature rarely follow political boundaries; there are numerous cases where habitats, species, or populations cannot be adequately conserved in one country alone. Such efforts include bilateral funding, comanagement agreements such as transboundary protected areas (TBPAs), and the development and implementation of international agreements intended to improve global patterns and trends in conservation and protection. Three important global mechanisms provide for the identification and designation of protected areas directly:

- The 563 Biosphere Reserves in 110 countries are sites recognized under UNESCO's Man and the Biosphere Programme. They are nominated by nations for international recognition on a voluntary basis. The original intention was to develop a systematic approach to conservation by developing a biogeographically representative network of protected areas, although this aspect has received relatively little direct attention in recent years. One unique defining feature has been an effort to support complementary human use. Many sites are categorized into core, buffer, and transitional zones with differing levels of human use and access.
- Ramsar sites are protected areas recognized under the Convention on Wetlands of International Importance. Following a number of guidelines and criteria, member nations identify and protect important wetlands. In 2011 there were nearly 2000 designated sites covering 1.9 million km^2.
- World Heritage Sites are places of 'outstanding universal value' nominated by the 187 nations that have ratified the 1972 World Heritage Convention. Chosen to protect natural or cultural attributes or both, the sites are selected in a review process that determines if the importance of a given site 'transcends national boundaries.' The 923 sites selected as of 2011 (183 selected for natural attributes and 28 for natural and cultural attributes) are then protected by the member nations.

Finally, the influence of the 1993 CBD is considerable. The original text of this convention called generally for greater efforts to protect biodiversity; a subsequent agreement set tangible targets for effective conservation of 10% of each of the world's terrestrial ecological regions by 2010 and of marine regions by 2012. In 2011 these targets were reset:

> Target 11: By 2020, at least 17 per cent of terrestrial and inland water areas, and 10 per cent of coastal and marine areas, especially areas of particular importance for biodiversity and ecosystem services, are conserved through effectively and equitably managed, ecologically representative and well connected systems of protected areas and other effective area-based conservation measures, and integrated into the wider landscapes and seascapes. (Convention on Biological Diversity, 2010)

Table 1 Total area of TBPA by region (TBPA, 2007)

Region	*TBPA area (km^2)*
North America	1,511,627.08
Central and South America	1,424,697.66
Europe	188,153.30
Africa	931,617.95
Asia	570,505.86

Although many of these international agreements and conventions have some legal framework, effective enforcement mechanisms are typically weak and cases are very rarely brought before the multinational or international courts. Even so, a number of NGOs and site managers have used the high profile of international recognition as a tool to support protection. For many countries, national pride in international recognition can also lead to associated benefits from tourism, particularly for World Heritage Sites. In some cases there are also financial and other supporting mechanisms for international sites.

A TBPA straddles one or more national or provincial boundaries. Though the area is managed cooperatively and human-made barriers are removed, each country retains sovereignty over the area within its borders. According to the most recent official inventory (2007), there are 227 TBPAs incorporating 3043 individual protected areas and more than 4.6 million km^2 (**Table 1**).

Status: Where and How Much; Measures of Effectiveness

Introduction

The World Database on Protected Areas (WDPA) is the major source of information on protected areas. Maintained by the World Conservation Monitoring Centre of the United Nations Environmental Programme and the IUCN World Commission on Protected Areas, the database contains crucial information from national governments, nongovernmental organizations, academic institutions, and others. It is used for ecological gap analysis, environmental impact analysis, and other components of conservation decision-making by governments and the private sector. The WDPA is also used to report progress toward the United Nations Millennium Development Goals, specifically Goal 7, which seeks to ensure environmental sustainability, and toward the 2020 Biodiversity Target of the CBD, specifically subsidiary Targets 1.1 and 1.2, which call for protection of at least 17% of each of the world's ecological regions.

Global, Continental, and Country-Level Coverage

Protected area coverage has been rising steadily since 1950 (**Figure 2**) and has increased sharply since 1970. As of 2010, the total protected area stands at 13.4% of Earth's terrestrial surface (Coad *et al.*, 2009a, b). At a global level, this figure exceeds the CBD 2010 target of 10%, but there is

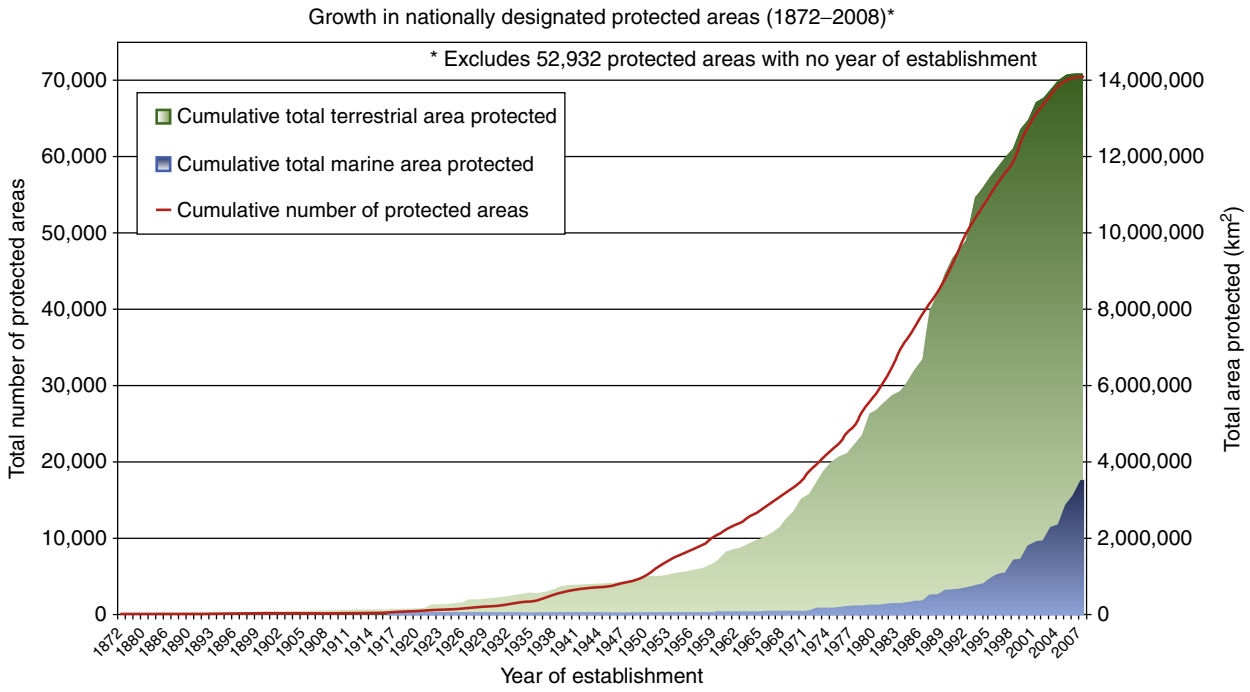


Figure 2 Global trend: cumulative global growth in the area of nationally designated protected areas (1872–2007). Total number of national sites = 113,962, of which 70,289 have establishment dates given in the WDPA.

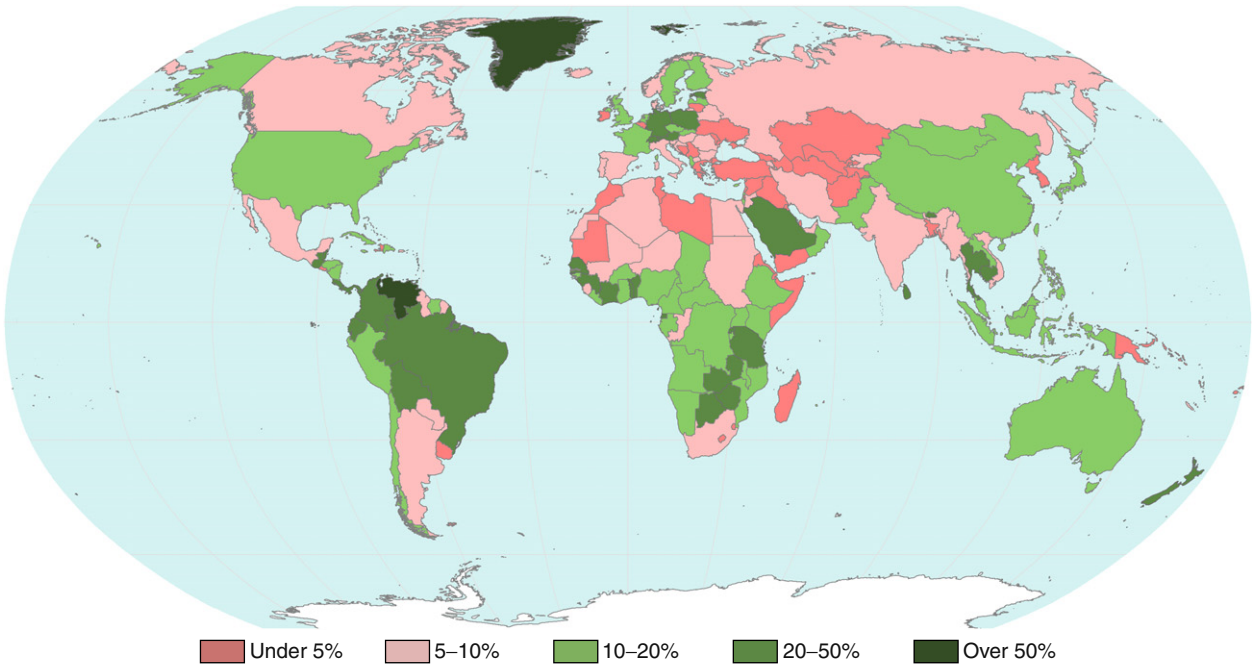


Figure 3 Land protection by country.

considerable variation among countries, ecoregions, biomes, and realms, and this measure provides no indication of effectiveness or representativeness.

The 236 nations that submit data to the WDPA have an average of 13% terrestrial protection, but only 108 countries have more than 10% protection (Figure 3). Twenty-nine countries have protected more than 25% of their land – including the very large and high biodiversity country of Brazil (28% protection) (UNEP-WCMC, 2010). Brazil also accounts for 74% of the land area added to protected status since 2003

Table 2 Regional land protection

Region	*Total area (000,000 km^2)*	*Percentage protected area coverage*
South America	19.84	21.09
Central America	0.74	23.54
North America	27.54	17.93
East Asia	12.76	15.94
Eastern and Southern Africa	12.91	14.66
Europe	6.48	13.79
Australia/New Zealand	9.11	9.95
Western and Central Africa	13.33	10.33
North Africa and Middle East	13.08	8.37
North Eurasia	23.33	7.69
Southeast Asia	10.07	13.73
South Asia	4.87	7.09
Caribbean	0.88	14.69
Pacific	3.02	9.7
Antarctic	14.79	0.01

(Jenkins and Joppa, 2009). Very little land outside of Brazil has been protected since 2003.

The UNEP-WCMC recognizes 15 large geographic regions consisting of continents and subcontinents. Nine of the 15 have protected more than 10% of their lands (**Table 2**).

Protection by IUCN Category

Of the 110,000 protected areas in the WDPA database, approximately 40% are uncategorized 19.7% of the land area in database); some have not been assigned an IUCN protected area management classification or the designation information was not made available to the WDPA. There is also disagreement as to how to apply the categories to particular sites. Of the categorized 66,000 sites, 33,070 (2.24 million km^2) falls into IUCN Category IV (habitat/species management area). Categories II (national park) and VI (protected area with sustainable use of natural resources) have a much lower number of protected areas (3731 and 3347, respectively), but they cover a greater total area – a combined total of 6.74 million km^2 (**Table 3**).

Table 3 IUCN categories: Area and percentage of nationally designated sites by IUCN management category

IUCN category	*Area protected (million km^2)*	*Percentage of total area of sites with assigned IUCN category*
I	1.78	1.35
II	3.37	2.56
III	0.22	0.17
IV	2.24	1.70
V	2.63	1.99
VI	3.37	2.56
Other	3.33	2.53
All PAs	16.94	12.85

PA, protected area.

Terrestrial Habitats and Ecosystem Protection

To assess the representativeness of protected areas, which measures the extent to which biodiversity is actually protected, numerous studies have used WDPA data to evaluate progress toward the CBD targets, all using a global biogeographic framework known as the World Wildlife Fund Ecoregions (Olson *et al.*, 2001) that divides terrestrial areas into realms, biomes, and, at the finest level, ecoregions. More than 120,000 records in the WDPA are difficult to use, interpreting results are challenging. Some countries have not fully classified or have reclassified their protected lands. The most prominent example is the United States, which initially included all the National Forests and National Grasslands, totaling some 77.7 million ha. By 2009, the Forest Service had removed all of the National Forests – more than 76 million ha – from the WDPA (Coad *et al.*, 2009b). This change affected the amount of temperate coniferous forests globally protected. Countries can refuse to allow data to be included for various reasons. Some countries provide imprecise data about location and boundaries. Location data can be inaccurate because countries often report sites as points and the boundaries are assumed to be a circle centering on that point and representing the reported size of the site. When adjacent to another protected site, some area might be lost in the analysis where overlap occurs, even though the two sites do not actually overlap. Imprecise location data can also be a problem in ecoregional analyses, because some part of the protected area may actually be in a different ecoregion than that represented by the assumed boundaries. There is also a debate about the inclusion of biosphere reserves, as, in some countries, these have no legal status and protection may be limited (**Table 4**).

Biogeographic Realms

Biogeographic realms represent the broadest scale patterns of variation in terrestrial biodiversity. The targets set under the CBD do not include realms. Nonetheless, the extent to which realms have been protected is informative in assessing broad-scale patterns of protection. At this level, 10% protection has been achieved in five of the seven realms (Coad *et al.*, 2009b).

Biomes

Although realms capture major patterns of evolutionary uniqueness, biomes represent physiognomic diversity, capturing the major habitat regions of the planet. Only three of 14 biomes have less than 10% of biome area protected (Coad *et al.*, 2009b). Using the WDPA 2009 data set, Jenkins and Joppa (2009) found that temperate grasslands were the least protected and most severely converted, with less than 4% protected (**Table 5** and **Figure 4**). In contrast, an earlier study found that flooded grasslands, montane grasslands, and tropical moist forests each had more than 20% protection (Hoekstra *et al.*, 2005). Since 2006 this disparity has increased, especially within tropical systems, reflecting a continued emphasis on protecting tropical systems. Encouragingly, protection in the Mediterranean Biome has increased from 5 to 7.5% – a sizeable increase reflecting the work of many conservation organizations and the success of prioritizing an ecosystem.

Because biomes are found in multiple realms, simple biome statistics may overlook major spatial variation in protection efforts. Given the unique evolutionary history of different realms, it is critical to consider both realms and biomes to obtain even a simple measure of representative protection. Jenkins and Joppa (2009) found that 46% of biomes within each realm remain below the 10% threshold.

Table 4 Protected area coverage of WWF realms

Realm	*Percentage protection (%)*
Neotropic (Central and South America)	27.1
Afrotropic (sub-Saharan Africa)	15.8
Nearctic (North Mexico, USA, and Canada)	12.2
Australasia (including New Guinea)	11.8
Palearctic (Northern Africa and Eurasia)	11.1
Indo-Malay (South and Southeast Asia)	9.9
Oceania (Pacific Islands)	6.1

Terrestrial Ecoregion Protection

Coad *et al.* (2009b) found that 54% (447 out of 821) of ecoregions met or exceeded the 2010 CBD 10% target (**Figure 5**).

Forests

Using WDPA 2008 data, Schmitt *et al.* (2009) calculated the protection percentages for each WWF forest ecoregion (**Table 6** and **Figure 6**). Globally, 18.8% of forest cover was protected; only Oceania did not meet the 10% threshold for forest protection. The neotropical realm had the highest level of forest protection, at 34% (**Table 6** and **Figure 7**).

Dry and Subhumid Lands

The Millennium Assessment identified four types of drylands (dry subhumid, semiarid, arid, and hyperarid) and used the following WWF biomes to assess the extent of protection: deserts and xeric shrublands, temperate grasslands, tropical and subtropical grasslands, and Mediterranean forests and shrublands (Safriel and Adeel, 2005). Although three of the four biomes have more than 10% protection (**Table 7**), 130 of the 227 dryland ecoregions (57%) are below the 10% threshold (Coad *et al.*, 2009b).

Mountains

In 2002, UNEP-WCMC defined mountain extent using criteria such as elevation and slope (Blyth *et al.*, 2002). A recent analysis using this data set found that mountain areas cover 39.3 million km^2 (26.5%) of the world's terrestrial area. By the end of 2009, nationally designated protected areas (PAs) covered 5.7 km^2 (14.3%) of the world's mountain areas outside Antarctica. Mountain PAs thus comprise 17.3 million km^2 (32.5%) of the world's total terrestrial protected area coverage. Mountain protected area coverage has increased globally by 21% since 1990. Out of 198 countries with mountain areas, 87 (43.9%) still had less than 10% of their mountain areas protected, whereas 18 (9.1%) had more than 50% protected area. In 87 of the 728 ecoregions with mountain areas, more than 50% of the mountain area was protected. In 16 (2.2%), mountain protection reached levels of 90–100%. However,

Table 5 Biome realm protection percentage

Biomes	*Global (%)*	*Australasia (%)*	*Afrotropic (%)*	*Indo-Malay (%)*	*Nearctic (%)*	*Neotropic (%)*	*Oceania (%)*	*Palearctic (%)*
Tropical and subtropical Moist broadleaf forests	21	11	14	10	–	32	3	8
Tropical and subtropical dry broadleaf forests	8	10	6	8	0	9	2	–
Tropical and subtropical coniferous forests	7	–	–	6	7	8	–	–
Temperate broadleaf and mixed forests	11	20	–	9	12	29	–	9
Temperate conifer forests	25	–	–	15	33	–		15
Boreal forests/taiga	9	–	–	–	10	–	–	8
Tropical grasslands, savannas, and shrublands	13	6	14	10	8	11	4	–
Temperate grasslands, savannas, and shrublands	4	2	0	–	3	2	–	5
Flooded grasslands and savannas	20	–	28	73	–	15	–	8
Montane grasslands and shrublands	25	46	8	34	–	14	–	32
Tundra	18	74	–	–	22	–	–	12
Mediterranean forests, woodlands, and scrub	7	12	19	–	21	1	–	5
Deserts and xeric shrublands	9	13	9	7	14	9	–	8
Mangroves	20	19	12	10	–	37	–	–

Source: Adapted from Jenkins CN and Joppa L (2009) Expansion of the global terrestrial protected area system. *Biological Conservation* 142: 2166–2174.

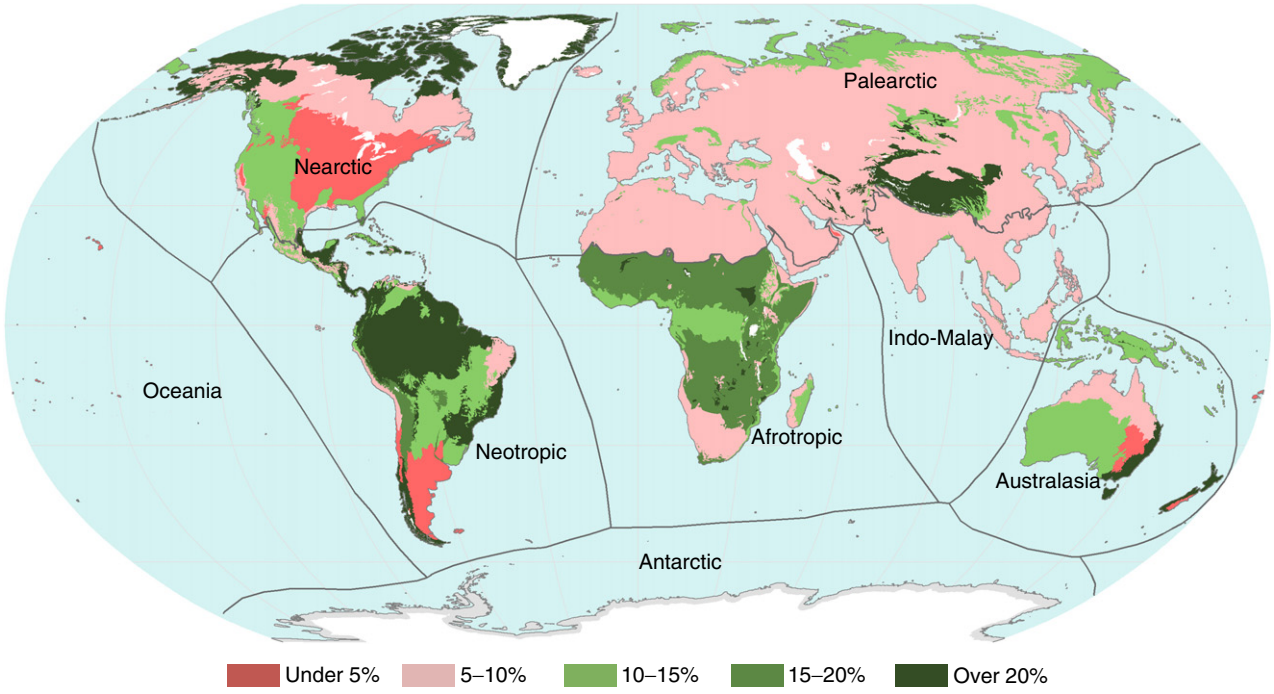


Figure 4 Realm-Biome Protection Percentage – map created from data derived.

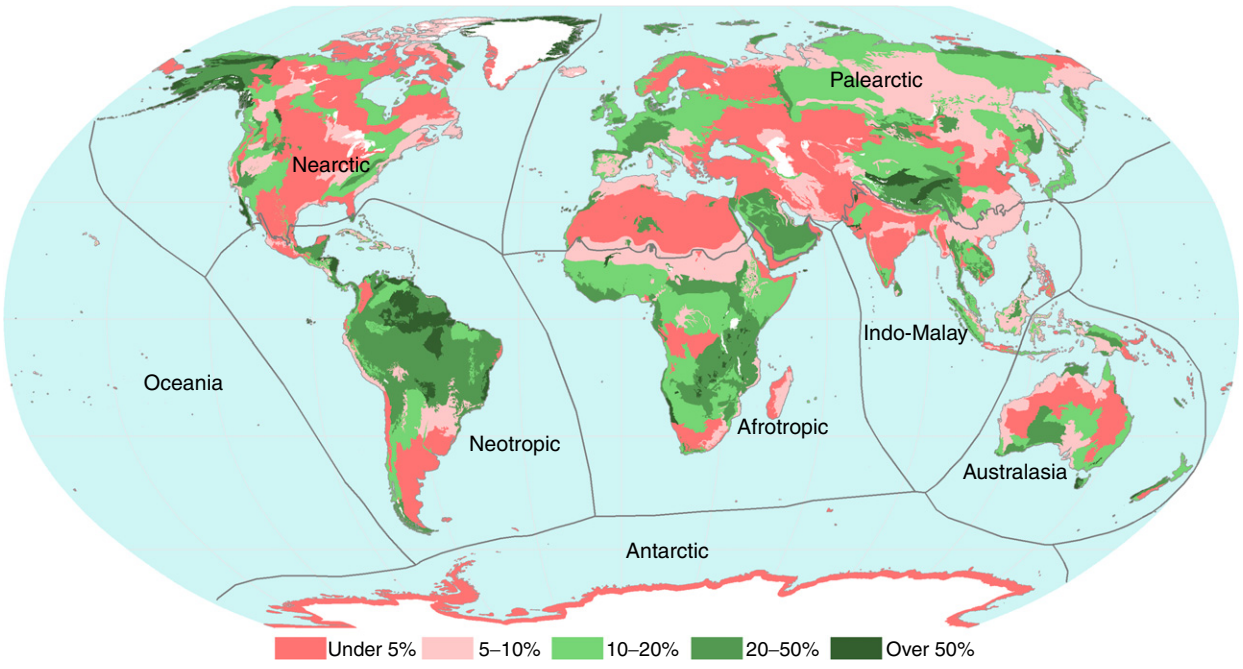


Figure 5 WWF Ecoregion protection percentage.

332 (45.6%) had less than 10% of mountain areas protected; 68 (9.3%) had no mountain areas protected at all (Rodriguez-Rodriguez *et al.*, 2010).

Islands

Very little work has been done to determine the protected status of island terrestrial area. Using the WDPA 2009 and a global country data set (ESRI, 2008), a global assessment of marine islands (171 countries have such islands) reveals a protection percentage of 29.3% of terrestrial area (Boucher, 2011, unpub.). This figure includes Greenland, which is almost 40% of the combined island area and has 51.4% protection. With the removal of Greenland, global island percentage protection drops to 15%. For tropical islands (with 118 countries

represented), protection percentage drops to 11.5%. When refined to Pacific tropical islands, the protection percentage drops to only 2.5%. Individually, 58 of the 118 tropical island countries fall below the CBD 2010 target of 10%. For some islands that host high biological diversity, protection percentages are low, such as Puerto Rico (7.7%), New Caledonia (5.5%), Madagascar (4.2%), Fiji (1%), and Haiti (under 1%). Tropical islands that exceed 10% protection include Cuba (12%), the Dominican Republic (15%), Jamaica (19.7%), the Seychelles (31%), and the Ecuadorian islands, including the Galapagos Islands (86%) (Boucher 2011, unpub.).

Marine Protection

Marine protection has always lagged far behind terrestrial protection, likely as a result both of lack of awareness and of the challenges of legal designation and implementation, particularly in the high seas, which make up some 45% of the entire planet. Through increased awareness and focused effort, both national and international, marine protection has increased rapidly (more than 280% since 2003) in recent years to 5880 sites in 2010 (Spalding *et al.*, 2010), with 4.7 million km^2 protected (1.31% of the global ocean). This protection is concentrated within 200 nmi of coastlines, where some 3.21% falls within Marine Protected Areas (MPAs).

Most conservation attention is concentrated in very nearshore waters, so although 4.32% of shallower continental shelf waters (less than 200 m deep) fall within MPAs, only 1.06% of areas beyond the continental shelf are covered. Only 12 of 190 states and territories have MPA coverage of their territorial waters at or above 10% (Spalding *et al.*, 2010).

From a biogeographic perspective, MPA coverage is also very uneven and does not adequately represent all ecoregions and habitats even in coastal areas. Marine ecoregions have been developed for continental shelf waters in a process analogous to terrestrial systems (Spalding *et al.*, 2007), and of the 232 marine ecoregions, 102 (44%) have less than 1% coverage, whereas only 44 (19%) have more than 10% protection. Some MPAs are extremely large; 11 of them exceed 100,000 km^2 in size and comprise 60% of the global coverage. The high sea areas (beyond 200 nmi) from coastlines remain almost entirely unprotected.

Table 6 Percentage forest protection by WWF realm

Realm	*Percentage forest protection*	*Percentage ecoregions below 10%*
Afrotropic	16.0	75.8
Australasia	18.1	46.8
Indo-Malay	16.6	51.5
Nearctic	17.5	70.2
Neotropic	34.2	62.1
Oceania	8.2	50.0
Palearctic	10.5	82.0

Source: Adapted from Schmitt CB, Belokurov A, Besançon C, *et al.* (2009) Global analysis of the protection status of the world's forests. *Biological Conservation* 142: 2122–2130.

Freshwater Ecosystems

Freshwater ecosystems, also called inland waters, and their dependent species are among the most imperiled ecosystems in the world (Millennium Ecosystem Assessment, 2005).

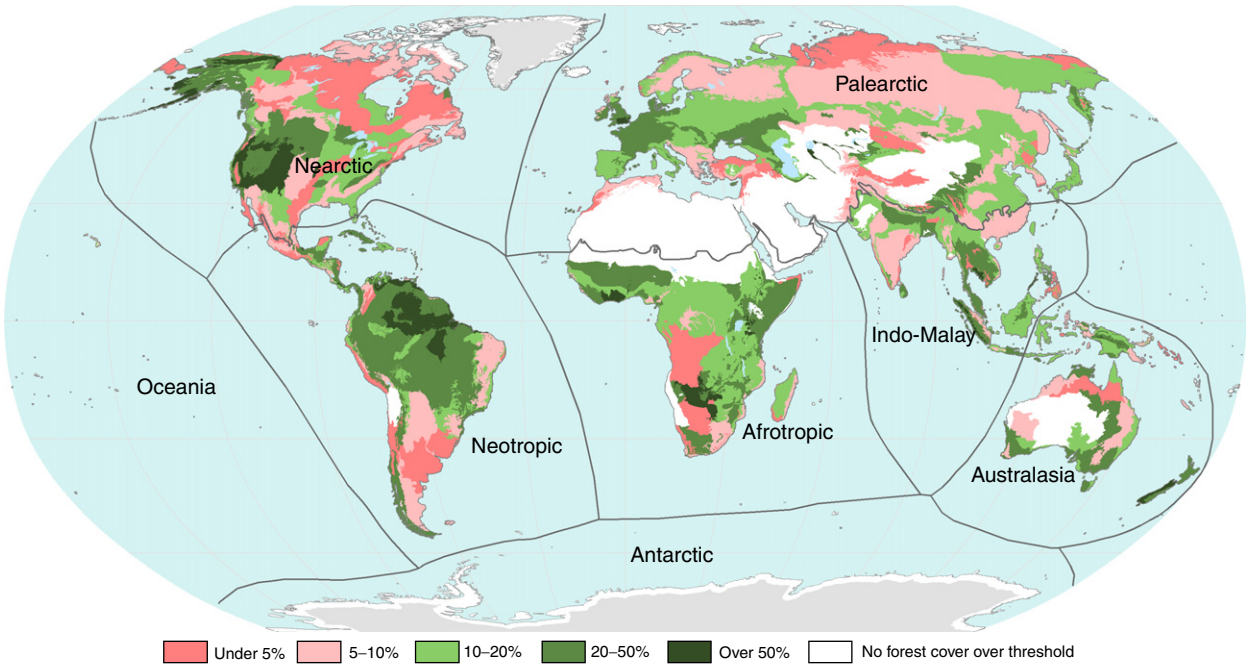


Figure 6 Forest protection – using a combination of GLC2000, MODIS05 VCF, and analyzed with WDPA 2008 data.

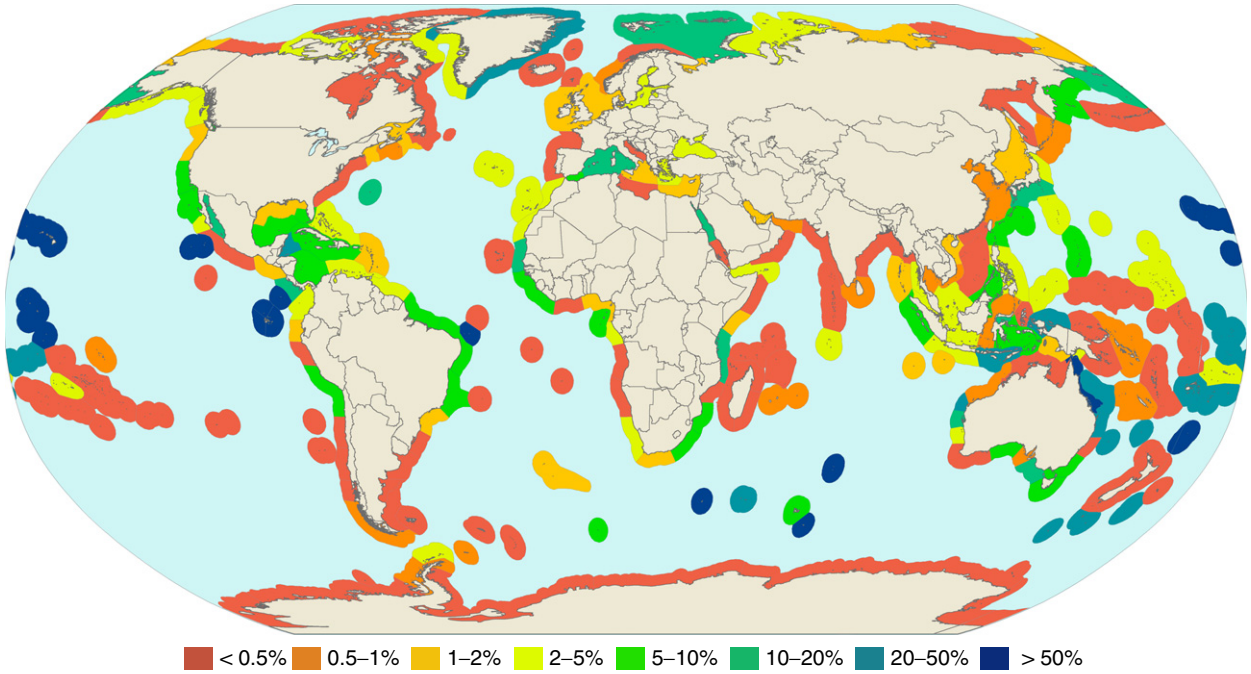


Figure 7 MPA coverage by marine ecoregion (percentages represent continental shelf area only).

Table 7 Protection of dryland biomes

WWF biome	*Area (km²)*	*Percentage of area protected*	*Number of ecoregions in biome*	*Percentage of ecoregions meeting the 10% target*
Deserts and xeric shrublands	27,984,645	10.8	96	43
Temperate grasslands, savannas, and shrublands	10,104,080	4.1	43	9
Mediterranean forests, woodlands, and scrub	3,227,266	10.2	39	54
Tropical and subtropical grasslands, savannas, and shrublands	20,295,424	15.9	49	63

According to the IUCN Red List of Threatened Species, one-third of amphibians and freshwater mammal species and one-fourth of the more than 240 freshwater turtle species are threatened with extinction. This situation is not surprising, as humans have settled in and around watercourses and relied on them for millennia. Humans have used freshwater ecosystems as sources of food and water, for transportation, and for waste disposal. As a consequence, many of the world's waterways have been heavily altered, impacting both freshwater species and the services derived from aquatic ecosystems, such as water supply, flood protection, and aquifer recharge. These nonmarketed services actually have a higher value than the marketed benefits of conversion of natural habitats to other uses (Millennium Ecosystem Assessment, 2005; Balmford *et al.*, 2002).

Many of these services are often unprotected, and their loss usually comes at a high price. There is ample evidence that the cost to restore key ecosystem services after degradation far exceeds the cost of protecting these services; in a comprehensive review of river restoration efforts in the United States, Bernhardt *et al.* (2005) point to the more than US$14 billion spent rehabilitating degraded streams in the United States since 1990.

Despite the value of services provided and the degree of imperilment of freshwater ecosystems, protection and conservation of inland waters has come at a slow pace when compared to the protection of terrestrial and, more recently, marine ecosystems. For decades, protected areas have been a key conservation strategy for terrestrial habitats, but for inland aquatic habitats such as rivers, lakes, marshes, and groundwater aquifers, this has not been the case. The traditional approach of fencing off an area for protection and management is unlikely and impractical for many aquatic habitats, particularly those with linear features, such as rivers and streams. Terrestrial protected areas do contain many and varied inland aquatic habitats, but many freshwater habitats and species

have not benefited much from surrounding or adjacent terrestrial protected areas largely because those protected areas have not been designed or managed for aquatic ecosystems and because they are especially vulnerable to pressures occurring outside the protected area (Abell *et al.*, 2007). Many terrestrial protected areas afford only partial protection to aquatic habitats, and important issues such as longitudinal and lateral connectivity and maintenance of essential hydrological processes (i.e., environmental flows) are not considered. Moreover, rivers are often used for the sole purpose of demarcating a terrestrial protected area boundary and are not integrated into the protected area's conservation and management objectives.

The only clear exception is the protection of aquatic habitats with defined surface areas, such as lakes, lagoons, marshes, and other wetlands. Under Ramsar, signatory countries have protected 1919 coastal and inland wetland sites with a total surface area of 186.7 million ha. Together these sites comprise the Ramsar List of Wetlands of International Importance, with the overarching goal to "maintain an international network of wetlands which are important for the conservation of global biological diversity and for sustaining human life through the maintenance of their ecosystem components, processes and benefits/services" (Ramsar). As of 2011, Ramsar had 160 contracting countries that could designate Ramsar sites and in doing so commit to protecting them from development, pollution, and other pressures (**Figures 8** and **9**).

For surface water bodies, establishing protected areas continues to be a customary conservation approach, but effective design and management has to include protection, or at a minimum, sound management of important features and processes that are essential to sustaining the wetlands being protected. These key features include rivers or groundwater aquifers that feed into the particular lake or wetland but are often located outside the protected area boundary.

Protecting linear features like rivers and stream by means of a protected area is not always feasible. Only a handful of countries including the United States, Canada, Australia, and New Zealand afford certain protections to entire rivers or portions of rivers. The United States designates as Wild and Scenic Rivers 12,560 miles of rivers or portions of rivers with outstanding natural, cultural, and recreational values that should be sustained in a free-flowing condition for the enjoyment of present and future generations (National Wild and Scenic Rivers Act of 1968). These rivers are federally protected from dams and other in-stream modifications. The law stipulates that each designated river be administered by either a federal or a state agency, but protection relies heavily on voluntary stewardship among stakeholders and regulatory programs by different governments agencies. Designation does not affect existing water rights or private property rights.

The Canadian Heritage Rivers System (CHRS) protects 41 rivers across Canada (CHRS, 2011). This government-supported stakeholder participatory program aims to promote, protect, and enhance Canada's river heritage. Canadian Heritage Rivers must meet established criteria regarding natural, cultural, and recreational values and be vetted through a rigorous and lengthy process. Public support is essential. Heritage River designation carries a commitment to protect its heritage values and integrity via a management plan or heritage strategy (CHRS, 2011). The CHRS relies on voluntary participation, community involvement, and collaboration between government agencies and stakeholders. The different provincial, federal, and territorial government agencies retain their traditional jurisdictional powers and management responsibilities; the CHRS does not have overall authority.

New Zealand has 13 river catchments and two coastal lakes largely protected through its Water Conservation Orders (New Zealand Ministry of the Environment, 2011). Water

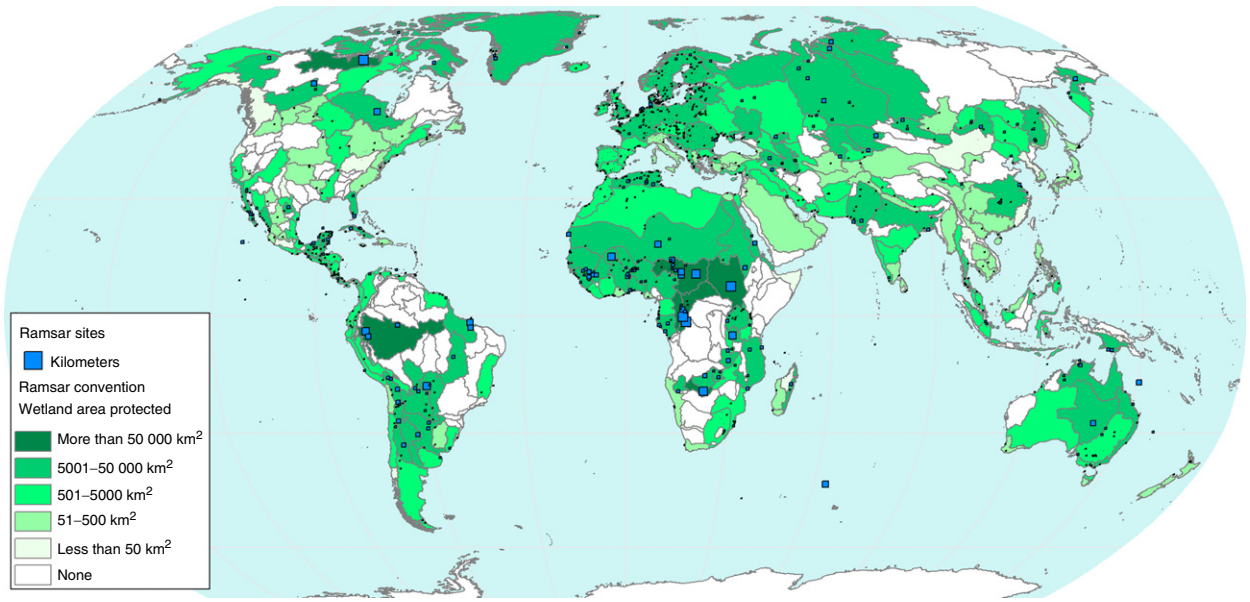


Figure 8 Ramsar wetland area protection by freshwater ecoregion (in km^2).

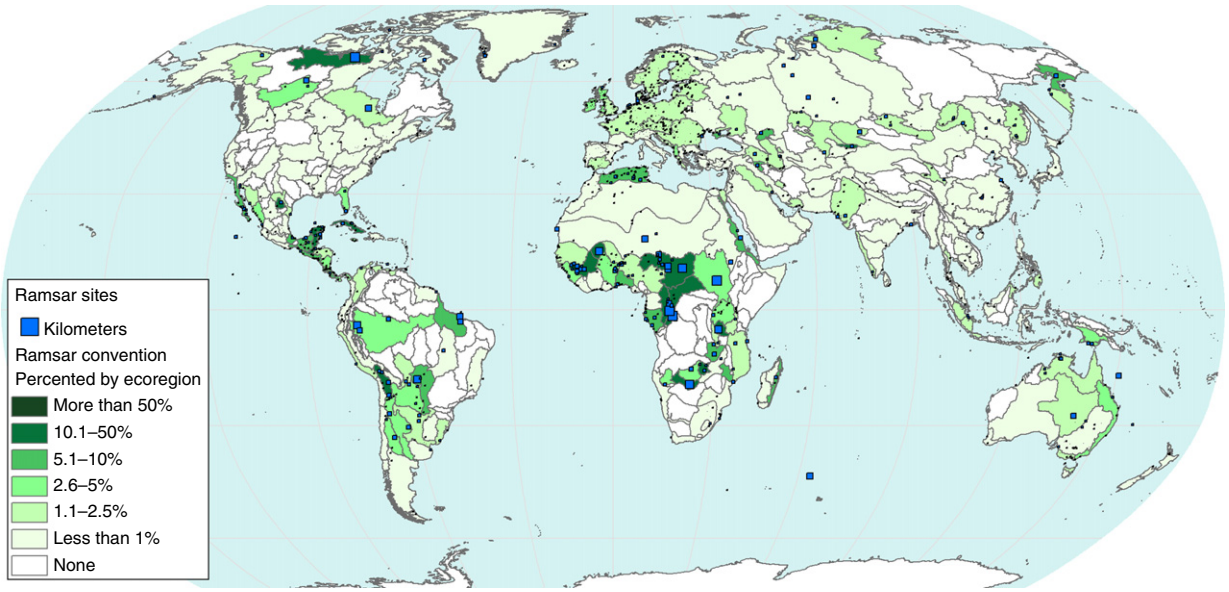


Figure 9 Ramsar wetland percentage area by freshwater ecoregion.

Conservation Orders aim to protect water bodies with outstanding natural, scenic, recreational, or cultural value (New Zealand Ministry of the Environment, 2011). The Minister of the Environment grants the order after a thorough review by an appointed special tribunal with a determined period for stakeholder comment; if granted, the order can restrict or prohibit new water abstractions, discharges, and other uses impacting the water body. Once an order is awarded, regional councils must modify their plans to comport with the order's provisions.

In Australia, the states of Victoria, Queensland, and New South Wales have similar river protection legislation. Queensland passed its Wild Rivers Act in 2005, and since then has declared 10 rivers as 'Wild Rivers.' The aim of the 2005 Act is to preserve those rivers whose natural values are intact or almost intact – that is, they retain essential riverine processes (Queensland Environment and Resource Management, 2011). A Wild River Declaration restricts and sets conditions for new development activities that may impact the river's natural values, such as water withdrawals, intensive land uses in the wild river area, and in-stream infrastructure projects. The State of Victoria, via its Heritage Rivers Act of 1992, has 18 Heritage Rivers listed for protection; however, according to a review by Nevill and Phillips (2004), 13 years later, "all river management plans remain as drafts without the required ministerial endorsement." The State of New South Wales has listed six rivers as 'Wild Rivers' because of their near-pristine condition; all lie within existing terrestrial protected areas" (NSW Office of Environment and Heritage, 2011).

Groundwater ecosystems enjoy even less protection than rivers. Many countries lack a basic inventory of groundwater aquifers, their condition, or how they feed into or connect to surface water bodies. Water quality legislation such as the U.S. Clean Water Act and the European Water Framework Directive provides some piecemeal protection to ground waters, most commonly by restricting in-stream flow uses in surface waters (Aldus and Bach, 2011). Fewer policies protect groundwater-dependent springs, wetlands, or groundwater aquifers.

Groundwater ecosystems and rivers, particularly the middle reaches and floodplains, are underrepresented in the global protected areas network, but figures showing exactly how underrepresented they are have been hard to come by. A key limitation has been the lack of inventory and assessment of freshwater ecosystems and their condition, even at the national level. This limitation has hindered the tracking of progress of individual countries toward their commitments under the CBD. The CBD in 2004 called on parties to "establish and maintain comprehensive, adequate and representative systems of protected inland water ecosystems within the framework of integrated catchment/watershed/river-basin management" (CBD Conference of the Parties, 2004a, b). Yet when the CBD 2004 Programme of Work on Protected Areas called for conservation of at least 10% of each of the world's ecological regions by 2010 for terrestrial ecosystems, it made no mention of inland waters as it was assumed that these habitats would be covered by the terrestrial target because they encompass riverine, lacustrine, wetland, and groundwater ecosystems.

The assumption that terrestrial protected areas protect the freshwater ecosystems found within them is therefore ingrained in the thinking and approach of most nations and of the CBD. Until recently, there was no global spatially explicit database of river networks at a resolution sufficient to allow for a detailed estimation of how much protection – at least in terms of coverage – terrestrial protected areas convey to waterways. In 2010, however, WWF in collaboration with McGill University and UNEP-WCMC released an assessment of the extent of coverage of rivers by protected areas by calculating both the length of rivers that lie within protected areas and the length of rivers that have their entire upstream

catchments under protection (WWF, 2010). This assessment showed that 69% of the world's rivers were without protected areas in their upstream catchment and that smaller headwater streams were more protected at catchment level than were larger rivers. The analysis therefore pointed to opportunities to increase the representation of freshwater ecosystems in the global protected area network. The analysis revealed important regional differences. Central and South America led the way in river protection, both at river length within protected areas and at upstream catchment levels. Asia, North America, and Europe/Middle East had fewer river miles and upstream catchment areas protected (WWF, 2010). This first estimate of river coverage provides a sense of how many lotic ecosystems could be protected if the management of the existing protected area network were tailored to address freshwaters but gives no indication of how protected areas actually protect freshwater processes and species.

Protected areas can and do work for freshwaters, but proper design and management are essential. Protecting freshwater ecosystems requires new approaches that incorporate the connectivity of freshwater systems and their key processes, such as spring peak flows that cue aquatic species to migrate, reproduce, and feed. Many of these approaches exist but are usually not considered within a protected area framework. Zoning rules, such as establishing riparian buffers to control runoff, reduce bank erosion, and provide shade, are in effect in many basins but are not *per se* protected areas. Water recharge areas or fishing reserves are other examples of protection that contribute to the conservation of freshwater habitats. Protected area design needs to include such measures, and protected area management needs to look beyond the terrestrial park boundary in order to reduce the rate of loss of freshwater species and the habitats upon which they depend.

Condition

The number of protected areas, the cumulative area protected, or even the representativeness of biomes or ecoregions say nothing about the effectiveness of protected areas in achieving their intended purposes. The CBD also calls for protected areas to be effective and equitably managed – metrics that cannot be assessed simply through measures of total coverage. Condition is hard to measure directly without extensive and time-consuming ground surveys, but there are indicators of the function of protected areas.

'Paper parks' are areas that have been officially gazetted as protected but have been given little or no active management (Brandon *et al.*, 1998). Some paper parks may be secure by virtue of remoteness, but in most cases, the condition of these protected areas will degrade due to encroachment for natural resource extraction or conversion to agriculture. In Brazil, for instance, 34% of protected areas have only one or two staffers and 25% of others are entirely unstaffed (Balazina, 2011).

Protected area failure differs from paper parks. Failure occurs when the protected area does not attain its specific goals for its stated establishment, even if managed as planned. The management plan may not have been capable of producing the intended outcome, it may not have been implemented properly, or enforcement against encroachment and prohibited uses may have failed.

Bruner *et al.* (2001) used questionnaires to assess the extent of land clearing, logging, hunting, fire, and grazing in protected areas. Of 93 parks in 22 tropical countries, 43% had had no net clearing after having been established. Even those heavily impacted by logging and hunting were in considerably better condition than the surrounding areas.

Bruner concludes that the claim that most parks in tropical countries are 'paper parks' – that is, parks in name only – is not substantiated, but in fact, the parks studied were not paper parks. All had guards, and effectiveness was most strongly correlated with the density of guards.

Human intervention may be needed to maintain natural conditions: artificial flooding regimes to maintain wetlands, native species planting, or invasives removal are all management interventions to maintain 'natural' systems in any site and particularly in small, isolated sites. Sites with only partial restrictions on use and activities may provide critical buffers or corridors to connect more strictly protected areas.

Conversion of Land in Protected Areas

Despite legal protection and active management, areas under protection, or some parts of those areas, are sometimes converted from their natural state to some form of human use such as agriculture or settlement. Using GLOBCOVER 2008 and WDPA 2009, the amount of protected area converted from a natural state was calculated. This global overview reveals that 2.4 million km^2 (9.2%) of the area under protection has been converted. This analysis does not evaluate fine-scale or subtle changes such as grazing or selective logging and does not exclude a small number of protected areas that may have been established on existing converted lands. No biome is untouched (**Table 8**). Conversion of protected areas ranges from below 1% in tundra to 72% in temperate grasslands.

Ongoing and Emerging Threats

Generally

Threats to protected areas abound, mostly in the form of external pressures that manifest as encroachment and result in deterioration. Protected areas often encompass viable agricultural land and rich natural and mineral resources, and so are very attractive to those living near them. Not surprisingly, the worst threats are habitat degradation and loss due to agricultural incursion, invasive species, and exploitation of natural and mineral resources, such as illegal logging, poaching of wildlife, overfishing, and mining. *Ex situ* threats such as pollution and climate change pose challenges to even the best managed PAs.

Inherent challenges include reserve design, isolation, and fragmentation. To be effective, the protected area's design and management must address the threats to the area and the biodiversity it supports. If essential management activities such as monitoring the health of habitats, enforcing the rules of the protected area, and working with local people to balance their needs and aspirations with the need to protect nature are not completely successful, threats can cause degradation even to well-designed areas.

Table 8 Percentage and area (km^2) of protected area conserved for human use (agriculture/urban), by WWF biome

WWF biome	*Percentage protection converted (%)*	*Area (km^2) protected converted*
Tropical and subtropical moist broadleaf forests	9.9	411,121
Tropical and subtropical dry broadleaf forests	44.5	147,796
Tropical and subtropical coniferous forests	18.4	8,779
Temperate broadleaf and mixed forests	27.6	341,578
Temperate coniferous forests	7.5	43,353
Boreal forests/taiga	0.4	6,282
Tropical grasslands, savannas, and shrublands	22.2	599,598
Temperate grasslands, savannas, and shrublands	72.3	272,897
Flooded grasslands and savannas	12.4	39,860
Montane grasslands and shrublands	28.9	380,595
Tundra	0.02	405
Mediterranean forests, woodlands, and scrub	18.5	43,919
Deserts and xeric shrublands	3.9	100,827
Mangroves	17.9	16,172

Encroachment

Encroachment into protected areas takes many forms, from large-scale resource extraction to fine-scale incursion such as cattle grazing and selective logging. In many countries, roads that provide entry into protected areas become pathways for both large and small encroachment (Cropper *et al.*, 2001). Human encroachment can also move invasive species into a protected area. The pressure of growing populations in the surrounding area may lead to encroachment (Birner *et al.*, 2006). Although this is especially true in developing countries, encroachment occurs in developed countries, too.

Of all the threats to protected areas in forested regions, deforestation, be it clear-cutting or selective logging, is the most pressing and widespread. Efforts to curb deforestation have had mixed results. Success is heavily tied to enforcement of the legal boundaries, which in turn requires political will and stability. The extent of deforestation in protected areas can be dramatic. An analysis of satellite imagery of the protected areas of Kalimantan in Indonesian Borneo found that the lowland PAs had been reduced by 63% (18,500 km^2), primarily by intensive logging (Curran *et al.*, 2004).

Mining

Balancing the economic needs of a country with conservation needs is always tricky. As countries grow, more and more resources, including underground minerals that can be very valuable, are required to fuel this growth. Mining entails the clearing of land for infrastructure, such as roads, housing, and in some cases, surface access to minerals. It also contaminates freshwater with toxins from the tailings. Miners may also poach wildlife.

The extent of mining in protected areas is poorly studied. The World Bank in 2010 initiated a study of artisanal mining in protected areas in Liberia and Gabon, and preliminary results include the lack of economic alternatives and risk to the environment through mercury use (PROFOR, 2011). In May 2011, the WWF's Central Africa Programme Office and World Wildlife Fund–U.S. announced the launching of a new project to address the impacts of artisanal and small-scale mining in and around the world's protected areas.

Industrial mining is conducted in some protected areas, in some cases as a preexisting and grandfathered use and in other cases by design. For instance, the Government of Canada together with the governments of the Northwest Territories and the Yukon has developed Terms of Reference for the Mineral and Energy Resource Assessment (MERA) process for proposed national parks in northern Canada. In some countries, including French Guiana, Finland, Indonesia, New Zealand, and Australia, mining companies are pushing governments to open protected areas to industrial mining activity.

Poaching and Bush Meat

Wildlife has long been taken as a protein source for humans. As growing human populations delete wildlife populations outside protected areas, people turn to poaching wildlife within protected areas. New roads, often built for timber extraction, facilitate access into remote or inaccessible areas; hunting for food has become a major problem in protected areas. Wildlife in many places is also a source of income, as bushmeat for local and even international markets or in illegal trade such as that of elephant tusks and rhino horns. Many protected areas have now become vital sanctuaries for what were wide-ranging species such as elephants and rhinos. Increased law enforcement in and around protected areas alone seems inadequate to stop poaching. Unless and until local people benefit more from the nonconsumptive value of wildlife and develop reliable sources of protein and income, even the largest, well-planned protected area may be incapable of protecting this component of its biological diversity.

Isolation

As the world develops, urban areas grow, land is cleared for agriculture, forests are logged, and myriad other human activities alter or eliminate natural areas. Protected areas that were once part of a larger landscape become islands in a sea of human development. This isolation undermines the ecological integrity of the ecosystems and the survival of the species that were to be protected. Isolation poses a threat to the long-term viability of wildlife and especially migrations of large mammals (Newmark, 2008).

With the loss of unprotected natural habitat, the protected area, in some extreme cases, remains the last fragment of natural habitat. This is especially true for tropical forests, which contain the majority of the world's terrestrial biodiversity. In an assessment of forest cover in and around 198 tropical protected areas, DeFries *et al.* (2005) found that of the protected areas studied, 69% in moist tropical forests and 54% in dry forests experienced a decline in forest habitat within 50 km of their periphery, with 28% and 29%, respectively, declining by more than 5% of the forest area. Although this loss occurred in most tropical areas, it was particularly high in southern and southeast Asia. Other studies show that isolation is occurring in the temperate and boreal regions as well. Young *et al.* (2006) found that even in the enormous boreal region of Canada, protected areas are becoming increasingly isolated as forests outside the protected areas are logged and fragmented.

Funding

Creating protected areas and managing them effectively require money. Only a few countries have managed to assure long-term, sustainable financing for protected areas. This funding gap is particularly acute in developing countries (De Oliveira, 2002) and for marine protected areas (Wells, 2008). What is needed is a way to find new and sustainable financial resources to supplement government funding for existing protected areas.

Tourism

Tourism is a mixed blessing for protected areas. Entry fees and hospitality and souvenir concessions generate income that provides much needed funds to run protected areas. Tourism to protected areas has been growing rapidly, so much so that in some countries, tourism has become an important part of their economies. In Kenya, for instance, tourism, which centers on the country's large parks, is the third highest contributor to gross domestic product (Messerli *et al.*, 2010). Tourism can also have a negative side, as it requires facilities and infrastructure to support the industry. Tourism can put considerable pressure on natural ecosystems and wildlife ('loving nature to death') (Bushell and Eagles, 2007).

Trends in Protected Area Conservation

Accretion in Number and Size of Areas

Opportunities to designate new protected areas or expand existing protected areas, or increase connectivity, will diminish due to population growth and the associated demands for natural resources, including arable land, energy, and water. Within years, or a few decades at most, the world's terrestrial surface will largely be assigned in terms of ownership and use, with marine spaces lagging only a short way behind. Protected areas will be by far the largest and most important tool for safeguarding nature. It also seems likely that the social and economic value of this estate will be more clearly established and will be far greater than currently perceived.

Some authors suggest that designation will continue at relatively rapid rates in the near future. Some countries still have relatively large areas of important biodiversity outside of protected areas. These areas are largely unmanaged. McDonald and Boucher (2010) suggest that the future of the protected area accretion appears to be positive. They regressed country-level protection rates from 1950 to 2005 against other available country-level time-series data. Variables positively correlated to protected area increase were high per capita gross domestic product, high levels of primary education, and a more stable and independent political structure. Poverty reduction is actually a more important factor in predicting the addition of protected area than is the biodiversity significance of the region, which attracts external pressure and funding to create protected areas. The analysis suggests that by 2030, global land protection could reach 15–29%, with more new land being protected in the next 20 years than in the previous 20 years.

The effectiveness of additional protected areas may be hindered by the lack of available buffer area. In other words, new protected areas may be isolated and expanded protected areas may be fragmented when created. In fact, protected areas may actually attract development (Wade and Theobald, 2010). After investigating spatially explicit data sets and 304 protected areas, Wittemyer *et al.* (2008) concluded that population is growing more rapidly (double the rural average rate) close to the protected areas than in more distant areas. Joppa *et al.* (2009) used the same data as Wittemyer *et al.* (2008) but concluded that the observed population expansion likely resulted from nearby population centers and that the growth observed near protected areas did not differ from average rural growth.

International carbon markets and the payments for the maintenance or enhancement of natural carbon stocks may facilitate large-scale accretion of protected areas. The United Nations Reducing Emissions from Deforestation and Forest Degradation (REDD) program should lead to considerable payment flows to tropical forest countries for the avoidance of deforestation. The REDD discussions have focused on areas outside current protected areas as these are the locations where extensive loss and degradation are occurring (Asner, 2011; Sandbrook *et al.*, 2010). There is also considerable interest in the development of similar mechanisms for coastal systems to protect the 'blue carbon' contained both in living biomass and in the soils of mangroves, saltmarshes, and seagrass beds; the high productivity of these systems has high value for carbon capture and storage (Crooks *et al.*, 2011; Laffoley and Grimsditch, 2009).

The marine environment still represents a conservation frontier. It has been suggested that the largest gains in protected area coverage will be in the marine realms despite the high costs and challenges of management. Some 50% of the oceans lie beyond any national jurisdiction, requiring the development of policies and approaches for management of the 'high seas.' One possible advantage for marine conservation is that issues of ownership and use rights are generally simpler than on land. This has meant that large-scale approaches, such as ecosystem-based management and ocean zoning, are perhaps more easy to implement, although even here the challenge of integrating the needs of multiple sectors remains considerable (Agardy *et al.*, 2011; Halpern *et al.*, 2008; Ruckelshaus *et al.*, 2008).

Loss of Existing Protected Areas

In a recent analysis, Mascia and Pailler (2010) identified at least 89 instances (in 27 countries) since 1900 of protected area downgrading (decrease in legal restriction), downsizing, and degazetting (PADDD). The extent of PADDD is unknown; Mascia and Pailler (2010), citing Zimmerer *et al*. (2004), report that nine countries (Botswana, Cameroon, Gabon, Ghana, Guinea-Bissau, Luxembourg, Pakistan, Somalia, and Togo) experienced a 5–60% decline in protected area coverage between 1985 and 1997. Some proposals to downgrade, downsize, or degazette have failed. Reasons for PADDD vary widely, but often center on access to and use of natural resources. In some cases, the official action merely recognized the failure of a protected area.

Shift to Multiple-Use Areas

New protected areas will tend increasingly to be multiple-use protected areas (IUCN Categories V–VI) rather than strict protection (IUCN Categories I–IV). This result is in accord with studies that have noted the rise of multiple-use PAs (Naughton-Treves *et al*., 2005). Another facet of this shift is the planned inclusion of indigenous groups or communities living in or near protected areas. In protected area theory and practice, the trend is away from the traditional concept of protecting areas by excluding people, which in its extreme form entailed the forced removal of people living in the area. Governments and conservation organizations now attempt to work with people who live in or near protected areas. Conservationists realize that failing to involve the local people often leads to failure in the long-term because those people need a form of economic activity to replace their use of the natural resources in the area.

Protected Area Management in the Context of the Surrounding Landscape

Protected areas will clearly remain a critical cornerstone of conservation efforts, but it seems likely that the boundaries between protected area approaches and other natural resource management measures will become increasingly blurred. In part, this is being driven by the growing realization that the need to manage natural resources at effective ecological scales can require natural resource management outside protected areas. Further impetus may come from the growing recognition of the value of ecosystem services. This is likely to encourage many nontraditional conservation partners, including utilities, agriculture, tourism, commercial fisheries, and others who may call for the protection of natural resources on strictly utilitarian grounds: to safeguard water supplies, reduce flood risks, or protect stocks. Sustainable management of natural resources across a matrix of protected areas and unprotected areas, beyond designated buffer zones (where they exist), would likely enhance the effectiveness of protected areas.

See also: Agricultural Invasions. Biodiversity in Logged and Managed Forests. Conservation and People. Conservation and the World's Poorest of the Poor. Conservation Efforts, Contemporary. Conservation Movement, Historical. Convention on Biological Diversity. Deforestation and Land Clearing. Development by Design: Using a Revisionist History to Guide a Sustainable Future. Ecosystem Services. Hotspots. Indigenous Peoples and Biodiversity. Marine Protected Areas: Static Boundaries in a Changing World. Oceanic Islands: Models of Diversity. Rainforest Loss and Change. Slash-and-Burn Agriculture, Effects of. Tourism, Role of. Traditional Conservation Practices. Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy. Wildlife Management

References

Abell R, Allan JD, and Lehner B (2007) Unlocking the potential of protected areas for freshwaters. *Biological Conservation* 134: 48–63.

Agardy T, di Sciara GN, and Christie P (2011) Mind the gap: Addressing the shortcomings of marine protected areas through large scale marine spatial planning. *Marine Policy* 35: 226–232.

Aldus A and Bach L (2011) Protecting Groundwater-Dependent Ecosystems: Gaps and Opportunities. Environmental Law Institute, National Wetlands Newsletter, May–June 2011, pp. 19–22.

Anon (2006) Conservation by design, tenth anniversary edition. Arlington, VA: The Nature Conservancy. Available online at http://www.nature.org/ourscience/conservationbydesign/cbd.pdf

Asner GP (2011) Painting the world REDD: addressing scientific barriers to monitoring emissions from tropical forests. *Environmental Research Letters* 6.

Balazina A (2011) Unprotected forests. Estado de Sao Paulo, 17 July 2011. Published online at http://www.estadao.com.br/estadaodehoje/20110717/not_imp745940,0.php

Balmford A, Bruner A, Cooper P, *et al*. (2002) Economic reasons for conserving wild nature. *Science* 297: 950–995.

Beck MW, Marsh TD, Reisewitz SE, and Bortman ML (2004) New tools for marine conservation: The leasing and ownership of submerged lands. *Conservation Biology* 18: 1214–1223.

Bernhardt ES, Palmer MA, Allan JD, *et al*. (2005) Synthesizing U.S. river restoration efforts. *Science* 308: 363–367.

Bernstein J and Mitchell BA (2005) Land trusts, private reserves and conservation easements in the United States. *Parks* 15: 48–60.

Bhatt S, Kothari A, and Tuladhar P (2008) South Asia. In: Chape S, Spalding M, and Jenkins M (eds.) *The World's Protected Areas. Status, Values, and Prospects in the Twenty-First Century*. Berkeley, CA: University of California Press.

Birner R, Maertens M, and Zeller M (2006) *Need, Greed or Customary Rights – which Factors Explain the Encroachment of Protected Areas?* Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference. Australia: Gold Coast.

Blyth S, Groombridge B, Lysenko I, Miles L, and Newton A (2002) *Mountain Watch: Environmental Change and Sustainable Development in Mountains*. Cambridge, UK: United Nations Environment Programme World Conservation Monitoring Centre.

Borrini-Feyerabend G, Kothari A, and Oviedo G (2004) *Indigenous and Local Communities and Protected Areas: Towards Equity and Enhanced Conservation*. Gland, Switzerland: IUCN.

Brandon K, Redford KH, and Sanderson SF (1998) *Parks in Peril*. Washington, DC, and Covelo, CA: The Nature Conservancy and Island Press.

Bruner AG, Gullison RE, Rice RE, and Da Fonseca GAB (2001) Effectiveness of parks in protecting tropical biodiversity. *Science* 291: 125–128.

Bushell R and Eagles P (2007) *Tourism and Protected Areas: Benefits Beyond Boundaries*. CAB International Oxfordshire, UK: CAB International.

Chape S, Blyth S, Fish L, Fox P, and Spalding M (compilers) (2003) United Nations List of Protected Areas. IUCN, Gland, Switzerland and Cambridge, UK and UNEP-WCMC, Cambridge, UK. lx + 44 pp.

Chape S, Spalding MD, and Jenkins MD (eds.) (2008) *The World's Protected Areas: Status, Values, and Prospects in the Twenty-first Century*. Berkeley, CA: University of California Press.

CHRS (Canadian Heritage River System) (2011) http://www.chrs.ca/About_e.htm

Coad L, Burgess ND, Bomhard B, and Besancon C (2009a) Progress on the Convention on Biological Diversity's 2010 and 2012 targets for protected area

coverage. A technical report for the IUCN international workshop "Looking to the Future of the CBD Programme of Work on Protected Areas", Jeju Island, Republic of Korea, 14–17 September 2009. Cambridge, UK: UNEP-WCMC.
Coad L, Burges ND, Loucks C, Fish, *et al.* (2009b) *The Ecological Representativeness of the Global Protected Areas Estate in 2009: Progress Towards the Convention on Biological Diversity 2010 target.* UNEP-WCMC, WWF–US and ECI, University of Oxford.
Conference of the Parties (2004a) Report from the 7th Conference of Parties to the Convention on Biological Diversity: Resolution VII/4: Biological diversity of inland water ecosystems.
Conference of the Parties (2004b) Seventh Meeting of the Conference of the Parties to the Convention on Biological Diversity – Decision VII/30. Strategic Plan: Future evaluation of progress. Goal 1 – Promote the conservation of the biological diversity of ecosystems, habitats and biomes; Target 1.1.
Crooks S, Herr D, Tamelander J, *et al.* (2011) *Mitigating Climate Change Through Restoration and Management of Coastal Wetlands and Near-Shore Marine Ecosystems: Challenges and Opportunities.* Washington, DC: The World Bank.
Cropper M, Puri J, and Griffiths C (2001) Predicting the location of deforestation: The role of roads and protected areas in North Thailand. *Land Economics* 77: 172–186.
Curran LM, Trigg SN, McDonald AK, *et al.* (2004) Lowland forest loss in protected areas of Indonesian Borneo. *Science* 303: 1000–1003.
Danilina N (2008) Northern Eurasia. In: Chape S, Spalding M, and Jenkins M (eds.) *The World's Protected Areas. Status, Values, and Prospects in the Twenty-First Century.* Berkeley, CA: University of California Press.
DeFries R, Hansen A, Newton AC, and Hansen MC (2005) Increasing isolation of protected areas in tropical forests over the past twenty years. *Ecological Applications* 15: 19–26.
De Oliveira JAP (2002) Implementing environmental policies in developing countries through decentralization: The case of protected areas in Bahia, Brazil. *World Development* 30: 1713–1736.
Dudley N (2008) *Guidelines for Applying Protected Areas Management Categories.* Gland, Switzerland: IUCN.
Dudley N, Stolton S, Belokurov A, *et al.* (eds.) (2010) *Natural Solutions: Protected Areas Helping People Cope with Climate Change.* Gland, Switzerland, Washington, DC and New York, USA: IUCN WCPA, TNC, UNDP, WCS, The World Bank and WWF.
ESRI Data & Maps [CD-ROM] (2008) Redlands, CA: Environmental Systems Research Institute.
Gascon C, Williamson GB, and da Fonseca GAB (2000) Receding forest edges and vanishing reserves. *Science* 288: 1356–1358.
GLOBCOVER (2008) GLOBCOVER Products Description Manual Report. European Space Agency, I2.2, Section 3: 18.
Halpern BS, McLeod KL, Rosenberg AA, and Crowder LB (2008) Managing for cumulative impacts in ecosystem-based management through ocean zoning. *Ocean and Coastal Management* 51: 203–211.
Hardy VC (1998) Grand Staircase-Escalante National Monument. CRS Report for Congress, ENR 98-993.
Hoekstra JM, Boucher TM, Ricketts TH, and Roberts C (2005) Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecology Letters* 8: 23–29.
IUCN (2011) IUCN Red List of Threatened Species. Version 2011.1. http://www.iucnredlist.org
Jenkins CN and Joppa L (2009) Expansion of the global terrestrial protected area system. *Biological Conservation* 142: 2166–2174.
Jones B, Stolton S, and Dudley N (2005) Private protected areas in East and southern Africa: Contributing to biodiversity conservation and rural development. *Parks* 15: 67–77.
Joppa LN, Loarie SR, and Pimm SL (2009) On population growth near protected areas. *PLoS ONE* 4: e4279.http://dx.doi.org/10.1371/journal.pone.0004279.
Kiesecker J, Amaon G, Comendant T, *et al.* (2007) Conservation easements in context: A quantitative analysis of their use by The Nature Conservancy. *Frontiers in Ecology and the Environment* 5: 125–130.
Laffoley D dA and Grimsditch G (2009) *The Management of Natural Coastal Carbon Sinks.* Gland, Switzerland: IUCN.
Mascia MB and Pailler S (2010) Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. *Conservation Letters* 4: 9–20.
McDonald RI and Boucher TM (2010) Global development and the future of the protected area strategy. *Biological Conservation* 144: 383–392.
Mcneely JA and Mainka SA (2009) *Conservation for a New Era.* Gland, Switzerland: IUCN.
Messerli H, Christie I, Konishi Y, and Surabian G (2010) *Kenya's Tourism: Polishing the Jewel.* Washington, DC: World Bank.
Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Wetlands and Water Synthesis.* Washington, DC: World Resources Institute.
National Wild and Scenic Rivers Act (1968) Public Law 90–542; 16 U.S.C. 1271 *et seq.*
Naughton-Treves L, Buck Holland M, and Brandon K (2005) The role of protected areas in conserving biodiversity and sustaining local livelihoods. *Annual Review of Environmental Resources* 30: 219–252.
Nevill J and Phillips N (2004) The Australian freshwater protected areas resourcebook. Hampton Melbourne: OnlyOnePlanet. From www.onlyoneplanet.com.au. Accessed July 2006.
Newmark WD (2008) Isolation of African protected areas. *Frontiers in Ecology and the Environment* 6: 321–328.
New Zealand Ministry for the Environment (2011) Water Conservation Orders. Available online at http://www.mfe.govt.nz/issues/water/freshwater/water-conservation/. Viewed 11 May 2011.
NSW Office of Environment and Heritage (2011) Wild Rivers page. Available online at http://www.environment.nsw.gov.au/parktypes/wildrivers.htm. Viewed 16 May.
Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* 51: 933–938.
Program on Forests, PROFOR (2011) Online at http://www.profor.info/profor/knowledge/impact-artisanal-and-small-scale-mining-protected-areas. Viewed 16 Jan 2011.
Queensland Environment and Resource Management (2011) Wild Rivers Website. Available online at http://www.derm.qld.gov.au/wildrivers/index.html. Viewed 16 May 2011.
Ramsar Convention Secretariat (2006) *The Ramsar Convention Manual: A Guide to the Convention on Wetlands (Ramsar, Iran, 1971)*, 4th edn. Gland, Switzerland: Ramsar Convention Secretariat.
Ricketts TH, Dinerstein E, Boucher T, *et al.* (2005) Pinpointing and preventing imminent extinctions. *Proceedings of the National Academy of Sciences* 51: 18497–18501.
Ruckelshaus M, Klinger T, Knowlton N, and DeMaster DP (2008) Marine ecosystem-based management in practice: Scientific and governance challenges. *BioScience* 58: 53–63.
Safriel U and Adeel Z (Coordinating Authors) (2005) Dryland systems in Millenium Ecosystem Assessment. Millenium Ecosystem Assessment Current State and Trends Assessment Report.
Sandbrook C, Nelson F, Adams WM, and Agrawal A (2010) Carbon, forests and the REDD paradox. *Oryx* 44: 330–334.
Schmitt CB, Belokurov A, Besançon C, *et al.* (2009) Global analysis of the protection status of the world's forests. *Biological Conservation* 142: 2122–2130.
Spalding M, Wood L, Fitzgerald C, and Gjerde K (2010) The 10% target: Where do we stand? In: Toropova C, Meliane I, Laffoley D, Matthews E, and Spalding M (eds.) *Global Ocean Protection: Present Status and Future Possibilities.* Gland, Switzerland: IUCN. The Nature Conservancy, UNEP-WCMC, UNEP, UNU-IAS, Agence des aires marines protégées, France. correction to statistics available on http://www.iucn.org/about/work/programmes/marine/marine_resources/marine_publications/?7040/global-ocean-protection
Spalding MD, Fox HE, Allen GR, *et al.* (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *Bioscience* 57: 573–582.
Udvardy MDF (1975) *A Classification of the Biogeographic Provinces of the World.* Morges, Switzerland: International Union for Conservation of Nature and Natural Resources. (now known as the World Conservation Union or IUCN).
United Nations Environment Programme World Conservation Monitoring Center (UNEP-WCMC) (2010) *State of the World's Protected Areas 2010: An Annual Review of Global Conservation Progress.* Cambridge: UNEP-WCMC.
Wade LA and THEOBALD DM (2010) Residential Development Encroachment on U.S. Protected Areas. *Conservation Biology* 24: 151–161.
WDPA (2009) World Database on Protected Areas Annual Release 2009 (web download version), February 2009. Available at http://www.wdpa.org accessed 26 March 2009.
Wells S (2008) Policy brief: Marine biodiversity and networks of marine protected areas. Report to the ninth meeting of the conference of the parties of the Convention on Biological Diversity, 19–30 May 2008, Bonn, Germany. Newark, Delaware: Gerard J. Mangone Center for Marine Policy.
Wittemyer G, Elsen P, Bean WT, Burton ACO, and Brashares JS (2008) Accelerated human population growth at protected area edges. *Science* 321: 123–126.
WWF (2010) Increasing Protection of River Catchments. Factsheet prepared by WWF for the CBD's 10th Conference of the Parties, Nagoya, Japan, October.

Yellowstone National Park Act (1872) Enrolled Acts and Resolutions of Congress, 1789–1996; General Records of the United States Government; Record Group 11; National Archives.

Young JE, Arturo Sánchez-Azofeifa G, Hannon SJ, and Chapman R (2006) Trends in land cover change and isolation of protected areas at the interface of the southern boreal mixed wood and aspen parkland in Alberta, Canada. *Forest Ecology and Management* 230: 151–161.

Zimmerer KS, Galt RE, and Buck MV (2004) Globalization and multi-spatial trends in the coverage of protected-area conservation (1980–2000). *Ambio* 33: 520–529.

Role of Botanic Gardens

Peter Wyse Jackson, Missouri Botanical Garden, St Louis, Missouri, USA
Lucy A Sutherland, Australian National Botanic Gardens, Canberra, ACT, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Conservation Maintenance of the diversity of living organisms, their habitats and the inter-relationships among organisms and their environment.
***Ex situ* conservation** Literally, away from the site or location; in this context means the conservation of components of biological diversity outside their natural habitats, for example, botanic gardens, seed banks.
Herbaria Collections of preserved plant and fungal specimens and their associated data. They are concerned primarily with scientific research and documenting the vast diversity of plant and fungal life.
***In situ* conservation** Literally, in or at the site or location; in this context means the conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species in their natural surroundings.
Seed banks Seeds are stored in conditions of low moisture and temperature. This technique is used for orthodox seeds of crops and wild species.

Introduction

Over the past few decades, botanic gardens throughout the world have undergone a remarkable transformation. As the greatest repositories of living plant collections worldwide, their importance and roles are better understood and appreciated by increasingly wide audiences. Many botanic gardens have been reinvigorated and rejuvenated to undertake a wide range of new tasks, particularly in education and plant conservation. The contribution of botanic gardens to cultural development, economic progress, and commercial expansion has been of great significance to many countries throughout the world over the past four centuries.

In the last 20–30 years, there has been a renaissance in botanic gardens worldwide, largely as a result of the developing concern for biodiversity loss and the need for many more institutions to become active in plant resources conservation. There has also been a corresponding rise in botanic garden involvement in research and conservation of the floras of the regions or countries in which they are situated.

Defining the Botanic Garden

There are no formal criteria or general agreement as to what constitutes a botanic garden. The lack of a very clear definition has blurred the edges between what are public parks or private collections and what are true scientifically based botanic gardens. In some instances, a garden has retained the name 'botanic' for historic reasons. Some or even most of the plant collection may survive, but often all the scientific activities have ceased and documentation has been lost. One might argue for the removal of these from a global list of botanic gardens. However, Wyse Jackson (1999) explains that it is precisely these institutions in many parts of the world that are currently being revived, redeveloped, and reestablished to become potentially important botanical centers.

An early definition of a botanic garden given by the International Association of Botanic Gardens was '…a botanic garden or arboretum is one open to the public and in which the plants are labelled' (IABG 1963 cited in Wyse Jackson, 1999). However, this definition fails to recognize the complexities of these institutions. *The Botanic Gardens Conservation Strategy* (IUCN-BGCS and WWF, 1989) introduced a wider view by presenting a list of characteristics defining a botanic garden that incorporate many of the diverse roles that these institutions now undertake. These defining characteristics included:

- adequate labeling of the plants
- an underlying scientific basis for the collections
- communication of information to other gardens, institutions, organizations, and the public
- exchange of seeds or other materials with other botanic gardens, arboreta, or research stations (within the guidelines of international conventions and national laws and customs regulations)
- Long-term commitment to, and responsibility for, the maintenance of plant collections
- maintenance of research programs in plant taxonomy in associated herbaria
- monitoring of the plants in the collection
- open to the public
- promoting conservation through extension and environmental education activities
- proper documentation of the collections, including wild origin
- undertaking scientific or technical research on plants in the collections (IUCN-BGCS and WWF, 1989, p 5).

This list does not, however, constitute a complete list of all the activities undertaken by botanic gardens. There are many institutions, which are recognized as botanic gardens, but often their work addresses only some of these criteria. The most widely used definition of a botanic garden is that adopted by Botanic Gardens Conservation International

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00392-0

(BGCI) in its *International Agenda for Botanic Gardens in Conservation* (Wyse Jackson, 1999, p 27 cited in Wyse Jackson and Sutherland, 2000, p. 12). BGCI suggests this definition encompasses the spirit of a true botanic garden:

> 'Botanic gardens are institutions holding documented collections of living plants for the purposes of scientific research, conservation, display, and education.'

Within the context of this article, the use of the term 'botanic gardens' should be interpreted to include arboreta and other specialized plant collections. In some parts of the world, the term botanical garden is used, rather than botanic garden. These can and are used interchangeably and both are correct.

Types of Botanic Gardens

Within the definition of a botanic garden given, a great diversity of institutions may be included, ranging from large gardens with several hundred staff and a wide range of activities, to small institutions with limited resources and activities.

A diverse range of organizations and administrations manage botanic gardens. Many are state administered or managed by regional or local authorities and receive public funding. About one quarter of the world's botanic gardens belong to universities and other research institutes for higher education and a relatively small proportion are private. A fast growing sector in the botanic gardens world is in the area of community botanic gardens, particularly in such countries as Australia and the United States. These gardens are designed to serve specific needs in their local communities and are often managed by those same communities. In recent years, the trend has been for botanic gardens to gain greater financial and administrative independence, often becoming trust administered and operating partly with funds gained through independent fund raising efforts.

Botanic gardens worldwide have been classified into 12 major types (Wyse Jackson and Sutherland, 2000, p. 14). Many botanic gardens have multipurpose roles and consequently, do not fit neatly into any well-defined category.

1. 'Classic' multipurpose gardens – these are often institutions with a broad range of activities in horticulture and horticultural training; research, particularly in taxonomy with associated herbaria and laboratories; and public education and amenity. They are generally state supported.
2. Ornamental gardens – these are often very beautiful establishments with diverse plant collections that are documented; they may or may not currently have research, education, or conservation roles. Some ornamental gardens are privately owned and many municipal gardens fall into this category.
3. Historical gardens – these include the earliest gardens developed for the teaching of medicine; some were established for religious purposes. A number of these gardens are still active in medicinal plant conservation and research, and today are primarily concerned with the collection and cultivation of medicinal plants and increasing public awareness about them.
4. Conservation gardens – most of these have recently been developed in response to local needs for plant conservation. Some contain, or have associated areas of, natural vegetation in addition to their cultivated collections. Included in this category are native plant gardens, which only cultivate plants from their surrounding region or national flora. Most conservation gardens play a role in public education.
5. University gardens – many universities maintain botanic gardens for teaching and research and many of these gardens are open to the public.
6. Combined botanical and zoological gardens – these are currently reassessing the roles of their botanical collections. Plants collections are being researched and developed that provide habitats for the displayed fauna, and interpretation of these habitats to the general public is an important element.
7. Agro-botanical and germplasm gardens – they function as an *ex situ* collection of plants of economic value or potential for conservation, research, plant breeding, and agriculture. Several are experimental stations associated with agricultural or forestry institutes and contain associated laboratory, plant breeding, and seed testing facilities, but many are not open to the public.
8. Alpine or mountain gardens – these are most frequently in mountain regions of Europe and some tropical countries. They are specifically designed for the cultivation of mountain and alpine flora, or in the case of tropical countries, for the cultivation of subtropical or temperate flora. Some alpine and mountain gardens are satellite gardens of larger lowland botanic gardens.
9. Natural or wild gardens – these contain an area or natural or seminatural vegetation, which is protected and managed. Most are established to play conservation and public education roles and include areas where native plants are grown.
10. Horticultural gardens – these are often owned and maintained by horticultural societies and open to the public. They exist primarily to foster the development of horticulture through the training of professional gardeners, plant breeding, registration, and conservation of garden plant varieties.
11. Thematic gardens – these specialize in growing a limited range of related or morphologically similar plants, or plants grown to illustrate a particular theme generally in support of education, science, conservation, and public display. These include orchid, rose, *Rhododendron*, bamboo and succulent gardens, or gardens established on such themes as ethnobotany, medicine, bonsai, topiary, butterfly gardens, carnivorous plants, and aquatics.
12. Community gardens – these are generally small gardens with limited resources, developed for and by a local community to fulfill its particular needs such as recreation, education, conservation, horticultural training, and the growth of medicinal, and other economic plants.

Distribution of Botanic Gardens

Worldwide there are 3026 known botanic gardens, in 175 countries or territories (Appendix 1). The most comprehensive

overview of the distribution and occurrence of botanic gardens is provided by the on-line 'Garden Search' database developed by Botanic Gardens Conservation International (www.bgci.org).

Earlier efforts to document the number, distribution, facilities, and activities of botanic gardens were represented by a series of International Directories of Botanical Gardens (i.e., Howard *et al.*, 1963; Henderson and Prentice, 1977; Henderson, 1983; Heywood *et al.*, 1990). A review of these directories highlights a remarkable increase in the number of botanic gardens. In the 1983 edition, 798 botanic garden entries were recorded. The 1990 edition recorded more than 1400 botanic gardens and arboreta and since that directory was published, the number of botanic gardens and arboreta known worldwide has more than doubled to a total of 3026 in a period of around 20 years. This represents a large number of newly established botanic gardens, as well as the revitalization and transformation of existing plant collections, parks, and other gardens into botanic gardens. BGCI's 'Garden Search' database was originally developed based on the earlier 1990 Directory and now as an on-line resource it enables individual botanic gardens to maintain and contribute data and information about their own institutions. Data on the number and distribution of botanic gardens worldwide presented in this article draws heavily on this database.

While the number of known botanic gardens worldwide has increased rapidly over the past decades, it is recognized that more than 50% of the world's botanic gardens have been created within the past 50 years. Of those for which the date of their foundation is known, only about 750 were established before 1950. The current trend for the creation of new botanic gardens in many parts of the world shows no signs of diminishing and each year many new botanic garden institutions are being proposed, planned, constructed, or opened.

Approximately one quarter of the known botanic gardens of the world occur in Europe (823 – 24%) with the largest numbers in the United Kingdom (117), France (104), and Germany (104) (**Table 1**). A similar number occur in North America (836 – 24%), in the United States (733), and Canada (103). In other regions of the world large numbers of botanic gardens occur in China (148), Australia (132), India (131), and the Russian Federation (109). Of all the regions of the world, the Middle East has the least number of botanic gardens (28), but there are also relatively few in South East Asia (49) and the Caribbean Islands (54) (**Table 1**). In Latin America (South and Central America and parts of the Caribbean), there has been a rapid rise in the number of botanic gardens in many countries in recent decades, for example, in Mexico (95), Argentina (48), Brazil (40), and Colombia (28). In the Caribbean, Cuba has 12 botanic gardens and in Africa, the largest number of botanic gardens is found in South Africa (22) and Nigeria (16) (**Figure 1**, **Table 1**, Appendix 2, **Box 1**).

There are less than 50 countries and territories without any botanic gardens (**Table 2**). Of these, about 40% are in Africa or the Middle East. There are also relatively few well-established botanic gardens in the countries and territories that make up the Pacific islands with only New Zealand having a significant number (20).

Table 1 Distribution and number of botanic gardens

Region	*Number of botanic gardens*
Africa and the Indian Ocean	157
East and South and South-west Asia	432[a]
Australasia	164[b]
Caribbean Islands	54
Central America	117[c]
Europe	823[d]
Former Soviet Union	193[e]
Middle East	28
North America	836[f]
South America	173
Southeast Asia	49
Total	3026

[a]Excludes South-East Asia.
[b]Includes the southern pacific islands and Papua New Guinea.
[c]Includes Mexico.
[d]Includes the Baltic countries that were part of the former Soviet Union (but not Belarus, the Russian Federation and Ukraine) and Greenland.
[e]Excludes Estonia, Latvia and Lithuania.
[f]Canada and the USA only.

Early History of Botanic Gardens

No one can doubt the remarkable contributions that botanic gardens have made to the development of human societies and culture over the past few centuries. We know, for example, that they became an important expression of the rebirth of Europe in the Renaissance, particularly with the establishment of a series of early botanic gardens in Italy and other European countries in the sixteenth century (Prest, 1988). These botanic gardens, regarded as the first in the world, were established as university physic gardens, whose primary function was to cultivate 'simples' for medicines and teach medical botany to physicians and apothecaries (Schiebinger, 2004, pp. 57–58). The earliest was the Pisa Botanic Garden, founded in 1543 by Luca Ghini, and followed closely by the Padua Botanic Garden in 1545. The remarkable botanic garden of Padua survives today more or less in its original form. Its heritage values were recognized by UNESCO in 1997 and it was the first botanic garden to be added to the World Heritage list for its 'outstanding universal value.' In 2003, the Royal Botanic Gardens Kew in the UK was also listed as a World Heritage Site.

Many other early European botanic gardens survive today as cherished historic monuments, important sources of pride for their founding universities and cities. These include those in Zurich (1560), Leiden (1577), Leipzig (1579), Paris (1597), Montpellier (1598), Oxford (1621), Uppsala (1655), Edinburgh (1670), Berlin (1679), and Amsterdam (1682).

In other parts of the world, earlier gardens that could claim to be botanic were created in previous civilizations, such as China, pre-Hispanic Mexico, and the Arab world. In China, for example, the Du-Le Garden flourished during the Son dynasty (AD 420–479). Within this garden, a collection of medicinal plants was cultivated, each one displayed with its own label. However, we know too little on how these early gardens functioned to be sure whether any can be classified as true botanic gardens, according to BGCI's accepted definition.

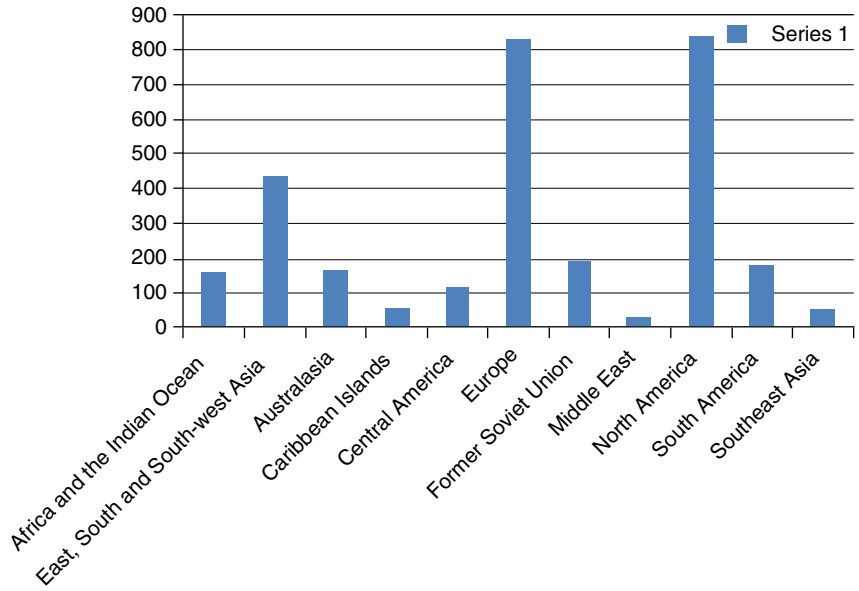


Figure 1 Distribution and number of botanic gardens. Source: BGCI database (2012), www.bgci.org.

Box 1 Case study: botanic gardens in Brazil

The development of botanic gardens in Brazil illustrates the remarkable recent growth of botanic gardens. Brazil is rich in plant species and, nationally, can be regarded as one of the world's most diverse countries for its biodiversity. Current estimates suggest that Brazil may contain in the region of 55,000 plant species, making the development of a strong base of botanic gardens an urgent priority to help manage and conserve such a rich heritage and diversity of plants. The Rio de Janeiro Botanic Gardens, founded in 1808, remains the largest, oldest and best supported botanical institution in that country. Indeed, it was the first botanic garden created in South America and is today the leading botanical institution in the continent. The second botanic garden founded in Brazil later that century was the Museu Paraense Emílio Goeldi e Parque Zoobotânico in Belem, in the State of Pará. The third major botanic garden established in Brazil was the Jardim Botânico de São Paulo, created in 1938. In the latter part of the twentieth century a remarkable explosion in the number of botanic gardens took place in Brazil. In the *International Directory of Botanical Gardens II* (1969) four botanic gardens were listed. In a subsequent International Directory published in 1990 (Heywood, Heywood and Wyse Jackson, 1990), the number of botanic gardens in Brazil had grown to 11. By 2000, the number had grown to include 26 institutions; three of those listed in 1990 are no longer mentioned (Bruni *et al.*, 2000). In a review of the world's botanic gardens prepared in 2001 (Wyse Jackson, 2001), the number listed had grown to 29 and today, the BGCI on-line 'Garden Search' database of botanic gardens of the world lists 40 institutions.

Table 2 Countries and territories without botanic gardens

Afghanistan	Liechtenstein
American Samoa	Maldives
Andorra	Mali
Aruba	Marshall Islands
Bahrain	Mauritania
British Indian Ocean Territory	Mayotte
Brunei Darussalam	Micronesia
Cambodia	Nauru
Central African Republic	New Caledonia
Chad	Niger
Cocos (Keeling) Islands	Niue
Comoros	Pitcairn
Cook Islands	Qatar
Cyprus	Saint Pierre and Miquelon
Djibouti	San Marino
Equatorial Guinea	Sao Tome and Principe
Eritrea	Somalia
Falkland Islands (Islas Malvinas)	Swaziland
French Polynesia	Syria
Guam	Timor-Leste (East Timor)
Guinea	Tonga
Guinea-Bissau	Tuvalu
Holy See (Vatican City)	Western Sahara
Kiribati	Yemen
Lebanon	

NB: Refer to Appendix 2 for a comprehensive list of countries with botanic gardens.

From these early beginnings, the major European botanic gardens have evolved to become the great institutions they are today, not only combining diverse activities in science and horticulture with public amenity and education, but also caring for a rich heritage of historic buildings, landscapes, living plants, and preserved collections. When considering the heritage importance of botanic gardens, their great buildings immediately come to mind such as the great glasshouses, their orangeries, statuary, fountains, terraces, and pergolas. Wonderful historic buildings that survive today in European botanic gardens include Kew Palace (1631) at the Royal Botanic Gardens Kew, UK; the Curvilinear Glasshouse designed by Richard Turner (1843–1869) at the National Botanic Gardens, Glasnevin, in Dublin, Ireland; the Orangery of the University of Vienna Botanic Garden (1755), the eighteenth century King's Gate by Sabatini and the Villanueva Pavilion at the Royal Botanic Garden Madrid; and the Palm

House (1872–1874) of the University of Copenhagen Botanic Garden. Nevertheless, our botanic garden heritage is much more than this. The collections of living plants, herbaria, libraries, seed banks, and works of art, manuscript archives, and other preserved collections contain tens of millions of specimens, representing an unequalled library of life and a window to the plant kingdom. Botanic garden herbaria alone contain almost 150 million preserved plant specimens, collections on which the scientific names and descriptions of most plant species have been based. Together botanic gardens grow representatives of about 100,000 plant species – perhaps a quarter of the world's plants, represented by some 6 million individual plant holdings.

In temperate regions a large number of urban municipal botanic gardens were founded in the nineteenth and twentieth centuries. In the twentieth century, such gardens were a particular feature in the United States, Australia, and New Zealand. Many of these gardens did not develop major scientific facilities, but there are notable exceptions to this, such as Missouri Botanical Garden, which was founded in 1859, the first botanic garden in the United States. Another exception was the Palmengarten in Frankfurt, Germany, founded in 1869. Most of these municipal botanic gardens developed significant activities in horticulture, building and maintaining major documented plant collections, as well as community engagement for recreation and informal learning.

In the countries of the former Soviet Union, the earliest botanic gardens are in Russia. The Apothecaries Garden of Moscow State University was created in 1706 and the botanic gardens of the Komarov Institute in St Petersburg were founded in 1713. Nevertheless, many of the 100+ botanic gardens in this region have been created within the past 50 years, mainly to act as centers for botanical research and plant introduction. They are generally associated with their national academy of science and universities. In China, more than 100 botanic gardens have been created in the past 20 years to support horticulture, research, native plant conservation, local education, and as a place for public relaxation.

The contribution of botanic gardens to the exploration of so many parts of the world is often underestimated. The interests of naturalists and explorers such as David Douglas, Augustine Henry, Joseph Banks, Archibald Menzies, and many more were stimulated and fostered by their connections with botanic gardens, helping to discover and introduce exotic plants from many parts of the world that have become a fundamental part of our familiar garden floras.

The world's botanic gardens have played a fundamental part in developing present day trade and economies of many countries and also helped shape patterns of agriculture in many parts of the world (**Box 2**).

The earliest tropical botanic gardens were motivated by considerations of trade and commerce, rather than by science for its own sake. European colonial powers, particularly Britain, the Netherlands, and to a lesser extent Germany, Belgium, Spain, and Portugal, established important tropical botanic gardens in their colonies in Africa, the Caribbean, India, South-East Asia, and South America in the eighteenth and nineteenth centuries. These tropical colonial botanic gardens became icons of modernity and had three main purposes. First, they were a place to study plants and their uses locally to support the new colonial settlements (Holttum, 1970, p. 707; Huxley, 1992, p. 376). Second, they were introduction centers for species that would be useful to the local economy (Huxley, 1992, p. 376). Third, some were created by governments as instruments of colonial expansion and commercial development (Heywood, 1987, p. 3). Consequently, the work within these tropical botanic gardens significantly influenced scientific, cultural (Huxley, 1992, p. 374), and economic development of many tropical countries by playing a major role in the establishment of agriculture (Heywood, 1987, p. 3; 1990, p. 4; Brockway, 2002), horticulture, forestry (Holttum, 1970, pp. 707–708, 710–711; Heywood, 1991; Given, 1993, p. 38) and trade routes (Hill, 1915, p. 210; Huxley, 1992, p. 376).

The first botanic garden to be founded in the tropics was the Pamplemousses Botanic Gardens in the Indian Ocean island of Mauritius, founded in 1729. It was initially established to provide fresh fruit and vegetables for the colony and to provide provisions for ships calling at the nearly port of St Louis. Through the Garden many economic plants were introduced, such as sugarcane and spice plants, which became the basis of the island's economy. Many economically useful plants were introduced through the earliest tropical botanic gardens, such as breadfruit, cinchona, cloves, cocoa, coffee, oil palm, and rubber, often associated with one of the major European gardens such as Amsterdam and Kew. Notable amongst these early tropical gardens are St Vincent (1765); Calcutta, India (1787); Rio de Janeiro, Brazil (1808); Sydney (1816) and Hobart (1818), Australia; Bogor, Indonesia (1817); Peradeniya, Sri Lanka (1821); Durban, South Africa (1849); Singapore (1859), as well as Bath (1779), Castleton (1859), Cinchona (1868), and Kingston (1872), all in Jamaica. Most of these early tropical botanic gardens still thrive today. The first oil palm (*Elaeis guineensis*), introduced into Indonesia in 1848 from tropical Africa, survived until approximately 20 years ago at the Bogor Botanic Garden (Kebun Raya Bogor) on the island of Java, the ancestor of the high-grade oil palms that are now widely cultivated in plantations in Indonesia and Malaysia.

Box 2 Case study: the contribution of botanic gardens to development

During 1560–1620, the horticultural explosion of botanic garden development is principally as a result of globalization which earlier work by Stafleu (1969, p. 35–36) attributed to the growth to the close political and economic relations between Turkey and Austria, which resulted in the introduction of flowers into Western Europe. As part of this globalization process, the flourishing trade and economy of The Netherlands, Austria and Germany saw the centre of horticultural activity shift from Italy at the end of this period (Stafleu, 1969, p. 38) and botanical institutions began to influence trade, economies and consequently, development. Arguably, this influence is still felt to the present day and the tulip industry is a good case in point. Carolus Clusius, the first director of the botanic garden at Leiden in The Netherlands, played a pivotal role in the history of its cultivation and was responsible for distributing many new varieties of bulbs throughout Europe (Pavord, 2000, p. 57–58). His work transformed the gardens of northern Europe (Pavord, 2000, pp. 57–58) and the legacy of his work is its contribution of the horticultural sector to the Dutch economy. Today, Dutch producers account for 93% of the total EU export of flower bulbs (Netherlands Turkey Trade Directory, accessed 03/02/2012).

Botanic Gardens and Research

Botanic gardens are extremely active worldwide in undertaking and promoting scientific research on plants and on biological diversity in general. Their collections and libraries provide important resources to support such research and many botanic gardens have excellent research facilities either within or associated with the institution. These include laboratories, herbaria, greenhouses, and growth chambers with controlled conditions, field experimental areas, climatic, and weather stations, data management systems, and advanced equipment for molecular, and genetic studies. Many botanic gardens are closely associated with universities and therefore, have special opportunities to undertake or develop research programs.

Plant taxonomy has traditionally been the primary focus of plant science research in many of the major botanic gardens, in particular those with major associated herbaria. This has resulted in the preparation of a number of major floristic and taxonomic works by such botanic gardens as the Royal Botanic Gardens Kew, Missouri Botanical Garden, New York Botanical Garden, and the Royal Botanic Garden Edinburgh (**Box 3**). Notable regional floras prepared either by, or with the close involvement of, botanic garden-based scientists include Floras of North America, China, the Arabian Peninsula, the Greater Antilles, Central America, Tropical East, and West Africa.

Box 3 Case study: developing a world flora on-line by 2020

Through decision X/17 of the Conference of the Parties of the U.N. Convention on Biological Diversity held in Nagoya, Japan in 2010, a consolidated update of the Global Strategy for Plant Conservation (GSPC) 2011–2020 was adopted. Target 1 of the Global Strategy for Plant Conservation calls for the achievement of '*An online Flora of all known plants*' by 2020.

A widely accessible Flora of all known plant species is a fundamental requirement for plant conservation. The terms and technical rationale for Target 1, within the Global Strategy for Plant Conservation (GSPC) 2011–2020, propose that an online World Flora be developed as a framework capable of accommodating regional floristic information (at national or lower level) that can provide answers in both regional and global contexts. It is suggested that the Flora should include accepted names and a comprehensive synonymy, building on the results of the previous (2010) Target 1 aimed to develop '*a widely accessible working list of known plant species as a step towards a complete world flora*' (Secretariat of the Convention on Biological Diversity, 2002, p.6). This target was achieved at the end of 2010, as The Plant List (www.theplantlist.org). The terms and technical rationale suggest that a World Flora should include geographic distributions to at least country level, drawing on national floras, checklists, and monographs; habitat data; identification tools (e.g., interactive keys, images, and descriptions); conservation status (with links to assessments being carried out under Target 2); and other enhancements as practicable, for example, vernacular names. Much of these data already exist in digital or printed format and can be used to populate the Flora.

When complete, the World Flora online will be expected to include information on approximately 400,000 plant species, and comprise vascular plants and bryophytes. Leadership in the development of the proposed World Flora is likely to be provided by the four largest botanic garden plant science institutions, i.e. Missouri Botanical Garden, New York Botanical Garden, Royal Botanic Garden Edinburgh, and the Royal Botanic Gardens, Kew.

The current main research areas in botanic gardens include:

- biotechnology
- climate change
- conservation biology
- conservation genetics
- data management systems & information technology
- ecology
- ecosystem conservation
- education
- ethnobotany
- exploration
- floristics
- horticulture
- invasive species biology and control
- molecular genetics
- phenology
- plant breeding
- plant population dynamics and diversity
- pollination biology
- propagation
- restoration ecology
- seed and spore biology
- species recovery
- systematics and taxonomy
- urban environments.

Not all botanic gardens have the resources (i.e., staff, facilities, and expertise) to enable them to play significant roles in botanical research. Nevertheless, most can contribute to such research by making their collections and facilities available to researchers.

Botanic Garden Collections

Over the past few decades there has been an increasing trend amongst botanic gardens to cultivate native plants from their own regions. Many botanic gardens have also assumed special responsibilities for the cultivation of rare and endangered plants and other taxa of conservation importance. In some cases, botanic gardens concentrate on the development of their collections following particular thematic, taxonomic and geographical lines, including the broad categories outlined below. It should be stressed, however, that this list is by no means complete and it should not be suggested that botanic gardens are either largely focused on such thematic collections, or that they do not maintain extensive collections of groups that fall outside of any of the categories listed.

The following broad categories are well represented worldwide in botanic garden *ex situ* collections:

Thematic collections

- Alpine and rock garden plants
- Aromatic plants
- Bonsai
- Bulbs and other geophytes
- Carnivorous plants
- Medicinal plants
- Ornamental plants (including cultivars and their wild progenitors)
- Rare and endangered plants

- Temperate fruits
- Temperate herbaceous perennials
- Temperate trees, shrubs, and climbers
- Tropical fruits (particularly fruit trees)
- Tropical timber trees
- Tropical flowering and other ornamental trees and shrubs
- Wild crop relatives
- Xerophytes

Taxonomic collections

- A wide range of temperate woody families (such as Aceraceae, Betulaceae, Fagaceae, Rosaceae).
- Aroids
- Bromeliads
- Cacti
- Conifers
- Crassulaceae
- Cycads
- Euphorbiaceae
- Ferns
- Ficus
- Grasses (including bamboos)
- Leguminosae
- Orchids
- Palms
- Rhododendrons

Geographical collections

- Native plants of temperate North America, Europe, Asia.
- Woody native plants of temperate Australia, New Zealand, and South America.
- Plants of Mediterranean regions of Europe, Southern Africa, South America, and Australia.
- Island plants, such as from Macaronesia, the Mascarene Islands, New Zealand, Japan, and Hawaii.
- Plants of the mountains of Europe, North America, and Asia.

Less well represented in botanic garden collections are the floras of tropical continental South America, South-east Asia and Africa, and other countries where there are still a limited range and number of botanic gardens (Appendix 2).

Botanic gardens *ex situ* (off-site) collections contain more than 6 million living plant accessions worldwide. The figures given in **Table 3** represent a conservative estimate based on the numbers of species and/or accessions reported by botanic gardens and included in BGCI databases. For some institutions and countries, knowledge of their collections is so scanty as to make it impossible even to estimate the likely number of accessions in a country. In those cases, an assumption has been made that no collections are held and they were excluded from the calculations. For that reason, the final total of +6 million accessions in botanic gardens worldwide must be regarded as extremely conservative estimate and it is suggested that a true figure may be as much as 20% to 50% higher.

As **Table 3** shows, the distribution of accessions in botanic gardens worldwide is uneven. The majority of collections are in Europe and North America (making up more than 50% of the total). The collections in most tropical and developing countries are significantly smaller than those in many temperate countries. This is not only partly due to an imbalance of institutional resources between such regions, but also due to the fact that many of the botanic gardens in tropical and developing countries are younger institutions without the long history of development of many European countries, that has allowed the growth of extensive collections built up over decades, and sometimes centuries.

Table 3 Botanic garden living plant accessions – estimated numbers and their distribution

North America	700,000
Russia and former Soviet Union countries	389,500
Africa and Indian Ocean Islands	216,600
Caribbean Islands	86,100
Western Europe	2,800,500
Central and Eastern Europe	404,000
Australasia and Oceania	331,500
South West Asia and the Middle East	58,500
South Asia	224,000
South America	145,500
Middle America	138,700
East and South East Asia	636,000
Total	6,130,900

Botanic Garden Herbaria

A high percentage of the botanic gardens of the world contain, include or are closely associated with herbaria. The largest herbaria are generally found in the oldest and largest botanical institutions in each country and region. Many contain diverse, historic and irreplaceable scientific resources documenting the world's plant diversity. For historic reasons, many original collections and much type material on which plant names have been based are housed in herbaria outside the natural range of the species concerned. To facilitate scientific research, most herbaria have a policy to provide free access to their collections to bonafide users for the purposes of scientific study. Although the majority of specimens preserved in botanic garden herbaria are of vascular plants, some important collections of nonvascular plants (e.g., bryophytes), algae and fungi are also maintained. The combined total number of specimens in botanic garden herbaria worldwide exceeds 150 million.

Major botanic garden herbaria include:

- Museum National d'Histoire Naturelle, Paris, France (8,000,000 specimens);
- New York Botanical Garden, USA (7,300,000 specimens);
- Komarov Botanical Institute and Botanic Garden, St Petersburg, Russia (7,160,000 specimens);
- The Royal Botanic Gardens Kew, UK (7,000,000 specimens);
- Missouri Botanical Garden, St Louis, USA (6,300,000 specimens)
- Conservatoire et Jardin Botaniques, Geneva, Switzerland (6,000,000 specimens);
- The Berlin Dahlem Botanic Garden (3,500,000 specimens);

- National Botanic Garden of Belgium (3,200,000 specimens);
- University of Helsinki Botanic Garden, Finland (3,126,000 specimens);
- Uppsala University Botanic Garden, Sweden (3,100,000 specimens);
- Royal Botanic Garden Edinburgh, UK (3,000,000 specimens);
- Jena University Botanic Garden, Germany (3,000,000 specimens);
- Copenhagen University Botanic Garden, Denmark (2,510,000 specimens);
- Beijing Botanic Garden, China (2,469,596 specimens);
- Kebun Raya Bogor, Indonesia (2,000,000 specimens);
- University of Oslo Botanic Garden, Norway (1,800,000 specimens);
- University of Tokyo Botanic Garden, Japan (1,700,000 specimens)
- Goteborg Botanic Garden, Sweden (1,600,000 specimens);
- Indian Botanic Garden, Kolkata (Calcutta), India (1,500,000 specimens);
- Botanischer Garten der Universitat Wien, Vienna, Austria (1,400,000 specimens);
- Australian National Herbarium, Australia (1,400,000 specimens);
- Pretoria National Botanical Garden, South Africa (1,200,000 specimens);
- Royal Botanic Gardens Melbourne, Australia (1,200,000 specimens);
- Real Jardin Botanico, Madrid, Spain (1,100,000 specimens);
- Royal Botanic Gardens Sydney, Australia (1,000,000 specimens);

Source: Thiers (undated).

Botanic Garden Education

The Convention on Biological Diversity (CBD) (United Nations Environment Programme, 1994) and Agenda 21 (United Nations, 1993), resulting from the Earth Summit held in Rio de Janeiro in 1992, highlighted the importance that public education and awareness raising play in promoting sustainable development, and improving the capacity of people to address environment and development issues.

The importance of botanic gardens in public education and raising awareness about the environment and sustainable development has grown in recent years. Some botanic gardens have sought to incorporate a vision for a more socially and environmentally sustainable future into their educational programming (see Dodd and Jones, 2010). As botanic gardens worldwide collectively receive in excess of 250 million visitors a year, they have considerable potential and opportunities to reach huge audiences with their educational programs. With growing urbanization, botanic gardens play an ever increasing and crucial role in public education, awareness raising, and engaging people with nature. In some cities they may represent one of the only opportunities for urban inhabitants to visit a natural or seminatural setting in their region. As the population becomes isolated from the natural environment there is a risk that people will become unaware of how their daily lives impact on the environment (Dodd and Jones, 2010). Therefore, there is a need to increase public sensitivity to environment and development problems, and foster a greater sense of personal environmental responsibility, motivation, and commitment towards sustainability.

A role of botanic gardens in education is also to act as an advocate for the maintenance of biodiversity and so educational topics addressed by botanic gardens include, but are not limited to, development issues, invasive threats, genetically modified foods, the relationship between people and plants, the role of science in plant conservation, sustainable living, and the value of biodiversity.

Formal education is a major focus for many botanic gardens and has strong historical roots, especially for such university botanic gardens as Oxford and Cambridge in the UK. The education role can involve providing materials and expertise for university teaching, as well as devising and providing formal education programs for other age groups. This often involves working with schools and contributing to schools' curricula.

Increasing more informal learning activities are being offered by botanic gardens. These are most often directed toward particular target audiences such as families, casual visitors, young people, and decision makers. Botanic gardens use a variety of techniques to encourage learning, engage a diverse community, and convey educational messages including guided tours, audio tours using latest technologies, social media, classes and workshops, exhibitions, festivals, interpretive signs, publications, and other such media as the internet, radio, television, and newspapers.

Botanic Garden Networks and Their Development

The International Association of Botanic Gardens (IABG)

The development of botanic garden networks began in the 1950s with the establishment of the International Association of Botanic Gardens (IABG) in 1954; although strong links between individual botanic gardens around the world have existed for more than a century. Such links were generally for the exchange of plant material and expertise. The International Association of Botanic Gardens (IABG) is affiliated to the International Union of Biological Sciences (IUBS) as a commission of the International Association of Botanical and Mycological Societies (IABMS). In recent years, IABG has not been active and the role in facilitating networking has been played by other international organizations, in particular Botanic Gardens Conservation International, as well as national and/or regional botanic garden network organizations.

The IABG operated by means of a series of geographically based regional divisions – in Europe, Ibero-Macaronesia, Latin America, Australasia-Oceania, east Asia, and with what was then the American Association of Botanic Gardens and Arboreta (now the American Public Garden Association) providing a corresponding service to institutions in North

America. All botanic gardens, arboreta, or other institutes and their staff were eligible for membership of IABG through the various regional divisions.

The aims of IABG are to promote:

- international cooperation between botanic gardens, arboreta, and similar institutes maintaining scientific collections of living plants;
- the study of taxonomy of plants to benefit the world community;
- documentation and exchange of information, living plants, and specimens between botanic gardens and similar institutes;
- the conservation of plants through cultivation and other means within botanic gardens and similar institutes;
- the introduction to cultivation of appropriate plants of benefit to the community;
- habitat conservation by cooperation between IABG and other relevant bodies;
- horticulture as an art and science.

Botanic Gardens Conservation International

Botanic Gardens Conservation International (BGCI) is a membership organization and the largest international network for botanic gardens. Its focus is primarily on the conservation of plant diversity, highlighting not only the importance of biodiversity conservation, but also the essential links to global issues including poverty, human well-being and climate change. It includes more than 700 members, which are mostly botanic gardens from 118 countries and the organization aims to support and empower its members, as well as the wider conservation community, so that their knowledge and expertise can be applied to reversing the threat of the extinction crisis facing plants.

BGCI was first established in 1987 as the IUCN Botanic Gardens Conservation Secretariat (BGCS). It began to build its membership of botanic gardens worldwide and develop a program of activities in support of botanic gardens. In 1989, *The Botanic Gardens Conservation Strategy* was published and the following year BGCS became independent from IUCN, and subsequently known as Botanic Gardens Conservation International (BGCI). BGCI registered as a UK charity and independence helped it to gain a greater measure of self-determination and made it possible for the organization to receive charitable donations in the UK. In addition to its head office in the UK, BGCI has national foundations in the USA and Russia and regional offices in several countries (www.bgci.org).

A primary concern of BGCI has been to provide a means for botanic gardens in all parts of the globe to share information and news about their activities, programs, and any new advances made that benefit conservation and education. Networking and capacity building for botanic gardens has been assisted through BGCI's magazines and the publication of a series of resource books, manuals, and policy handbooks on the development of botanic gardens and their roles, on such subjects as plant reintroductions, *ex situ* conservation, environmental education, education for sustainability, computer software, regional action plans, the Convention of Trade in Endangered Species of Fauna and Flora (CITES) and the Convention on Biological Diversity (CBD). Furthermore, it has supported its members by providing a number of tailor-made training and capacity-building workshops and works collaboratively with the Royal Botanic Gardens Kew to offer several International Diploma courses specific to the key roles of botanic gardens.

BGCI also organizes an international botanic gardens congress every 3 years since the first such congress in Las Palmas de Gran Canaria, Spain in 1985. BGCI also holds regular international congresses focusing on botanic gardens education.

BGCI has developed unique on-line computer databases on the botanic gardens of the world. The 'Garden Search' lists every known botanic garden, arboretum and many more similar institutions maintaining living plant collections in cultivation, with details of the facilities, collections, and work of more than 1800 botanic gardens. The 'Plant Search' lists details of collections in botanic gardens worldwide including locating rare and threatened plant species in cultivation around the world.

This database is compiled from the lists of living collections submitted to BGCI by the world's botanic gardens. The database presently includes more than 575,000 records. Information on threat status of each taxon is provided by a direct link to IUCN's 1997 and latest Red Lists. Further nomenclatural and bibliographic information is provided through the link to the TROPICOS database maintained by the Missouri Botanical Garden. The database is also linked to lists of medicinal plants, crop wild relatives and CITES listed plants. The 'Plant Search' database is recognized as an important mechanism to monitor the achievement of the *ex situ* plant conservation target of the Convention on Biological Diversity's Global Strategy for Plant Conservation (GSPC).

National and Regional Botanic Garden Networks

The past decade has also seen the establishment and growth of a wide range of national and regional organizations in all parts of the world for, or including, botanic gardens. National networks of botanic gardens help gardens work together to share information and coordinate a united role in the country. Partnerships and networking with bodies that have a complementary mission increase support for conservation initiatives. There are well-established networks in several countries including Australia, China, India, Mexico, Russia, South Africa, and the United States (refer to www.bgci.org). Furthermore, the networks in such regions as Europe, the Caribbean and sub-Saharan Africa, have had a significant role in strengthening capacity and extending conservation collaborations.

BGCI has worked to support this development and provide such organizations with assistance and support and, in addition, to collaborate closely with these sister networks. Amongst its activities in this area it provides a secretariat for the European Botanic Gardens Consortium, a body linking the national botanic garden networks in Europe.

In the United States, the Center for Plant Conservation is a network of botanic gardens that is focused primarily on

preventing the extinction of US native plants. The Center is a network of 37 botanic gardens. Founded in 1984, the Center coordinates a national program of *ex situ* conservation of rare plant material, with the aim of ensuring that material is available for restoration and recovery efforts for these species. CPC also works in research, restoration, technical assistance, education, and advocacy through the efforts of the network and its national office based in St Louis. The cooperative CPC network maintains the National Collection of Endangered Plants of more than 750 of native plants. Live plant material is collected from nature under controlled conditions and then maintained as seed, rooted cuttings, or mature plants.

Botanic Gardens and Plant Conservation

Throughout the world, the loss of biodiversity is recognized as one of the most challenging issues that we face if a sustainable planet is to be achieved in the future and if we are to safeguard the genetic resources that will be required by future generations. Despite the adoption by most of the world's countries of national biodiversity conservation strategies and plans during the past decade, biodiversity continues to rapidly disappear as natural habitats are destroyed and more and more species become endangered. Worldwide it is recognized that tens of thousands of plant species are rare or endangered and potentially face extinction in this century if current trends continue. Although the potential extinction crisis faced by plants worldwide has been recognized for several decades, only recently has a coherent plan of action for their conservation been proposed.

Ex situ facilities, such as botanic gardens, have shown an ability and vital role in managing plant resources and the evolution of their conservation role has reflected a scientific and cultural shift for plant collections from having a curatorial goal, which focuses on accumulating typological specimens, to the adoption of population genetics as a working tool to conserve threatened species (Maunder *et al.*, 2004a, p. 7 and b, p. 413).

The Botanic Gardens Conservation Strategy states that 'The purpose of *ex situ* conservation is to provide protective custody. It is justifiable only as part of an overall conservation strategy to ensure that species ultimately survive in the wild. Its role should be seen as a means to an end, not an end in itself: as a source of material for reintroduction into damaged habitats and to enhance populations as part of ecosystem management, for research and education, for selecting material for introduction into the nursery trade, local agriculture, amenity planting, and local forestry, etc. Another role is to take the pressure off wild populations for plants that are likely to be the subject of interest by scientists, commercial horticulturists, hobbyists, or local gatherers. Above all, *ex situ* conservation makes plants available for use by [humankind] (IUCN-BGCS and WWF, 1989, p. 21)'.

Ex situ conservation of wild plants has become a central and unique role of botanic gardens and has several purposes (**Box 4**). They have the appropriate facilities and staff expertise in botany and horticulture to be effective centers for *ex situ* conservation. *Ex situ* conservation can include the maintenance of samples of whole individuals, as well as seed, pollen, vegetative propagules, and tissue or cell cultures.

Box 4 The purposes of *ex situ* conservation

Ex situ conservation has several purposes:

- Rescue threatened germplasm.
- Produce material for reintroduction, translocation, reinforcement, habitat, and landscape restoration and management.
- Produce material for conservation biology research.
- Bulk up germplasm for storage in various forms of *ex situ* facility.
- Supply material for various purposes to remove or reduce pressure from wild collecting.
- Grow those species with recalcitrant seeds that cannot be maintained in a seed store.
- Generate skills and knowledge to support wider conservation aims.
- Make available material for conservation education and display.

Source: Reproduced from Smith P (2006) Wild seed bank sand taxonomy. In: Leadlay E and Jury S (eds.) *Taxonomy and Plant Conservation*, pp. 294–304. UK: Cambridge University Press.

Ex situ plant conservation activities do not themselves constitute the protection of wild species; they in fact complement the wider range of plant conservation-related activities (Smith, 2006). However, caution is needed because as a method of conservation, *ex situ* is inherently deficient in that it is not usually possible to maintain more than a limited sample of the genetic diversity in cultivation or in storage. Prance (2004, p xiii) highlights that a drawback to *ex situ* conservation is that '…in most cases it halts or distorts the natural process of evolution.' Consequently, it is often regarded as preservation rather than conservation. In addition, it may lead to unpredictable genetic change and can become in practice a form of domestication. In contrast, *in situ* conservation, at least in theory, allows plant populations to develop and evolve in, and as part of, the ecosystem of their natural habitat. In practice, both methods should form an integrated conservation response for the retention and restoration of wild plant diversity (Maunder *et al.*, 2004, p. 5).

Increasingly an important type of botanic garden is developing, particularly in some tropical countries, where gardens have been created alongside national parks and play key roles in integrated conservation, sustainable development, and public education. Examples of these botanic gardens include Limbe Botanic Garden in Cameroon and the Tam Dao and Cuc Phuong Botanic Gardens in Vietnam.

Despite the above mentioned issues, *ex situ* conservation in botanic gardens has several important benefits:

- *Ex situ* conservation may be the only option available when a natural habitat has been destroyed.
- It can be very cost-effective.
- Seeds of many species especially lend themselves to compact storage (allowing bulk samples); they are economical and can undergo long-term storage.
- Plant collections can give users ready access to a wide range of genetic variation within a species.
- Botanic gardens provide propagation and often research facilities, together with horticultural and other applied scientific skills needed in practical species conservation.

Box 5 *Ex situ* conservation priorities

Priorities identified for inclusion in *ex situ* plant conservation programs include the following:

- Species or taxa that are in immediate danger of extinction, either locally, nationally or globally.
- Species or taxa that are of local economic importance, such as minor food crops, medicinal plants and wild or cultivated plants providing the basis of local industries, agriculture, horticulture and crafts.
- Species or taxa, such as local ecotypes, that may be required for specific reintroduction or habitat restoration and management schemes.
- Local 'flagship' species or subspecies that will stimulate conservation awareness and can be incorporated into education and fund raising programs.
- Species or taxa that are of special scientific interest, such as narrow endemics or geographical relics.

Source: Reproduced from Wyse Jackson PS and Sutherland LA (2000) *The International Agenda for Botanic Gardens in Conservation*. UK: Botanic Gardens Conservation International.

- *Ex situ* conservation provides back-up for populations of threatened plants in the wild, contributing material for reintroduction, restocking, and restoration, as well as advice, and data for field management.

In 2002, the Global Strategy for Plant Conservation was agreed by the international community through the UN's Convention on Biological Diversity to address the potential loss of so much of the world's plant diversity (Secretariat of the Convention on Biological Diversity, 2002). To date, most global investment in *ex situ* conservation has focused on the retention of plant material of value for crop breeding (Maunder *et al.*, 2004b, p. 389). Although it is recognized that the conservation of plant resources is fundamental to the future survival of humanity and of many other species that rely on plants to provide the fabric of most terrestrial ecosystems, plant conservation beyond crops has barely received the attention that it needs until recent years (**Box 5**).

The role and value of *ex situ* conservation has increasingly been acknowledged in international conventions and legislation. Hand in hand with this development the botanic garden community has developed a series of international and national policies in relation to their priorities and activities in plant conservation. These include botanic garden action plans, contributions to national plant conservation strategies and individual botanic garden conservation policies and other documents.

At the international level several important botanic garden-focused policy documents have been prepared, primarily under the leadership and initiative of Botanic Gardens Conservation International. The most widely known such initiative is the International Agenda for Botanic Gardens in Conservation (Wyse Jackson and Sutherland, 2000), which is described below. It is currently being revised and updated by BGCI.

International Agenda for Botanic Gardens in Conservation

In 2000, Botanic Gardens Conservation International (BGCI) prepared and published the *International Agenda for Botanic Gardens in Conservation* to provide a global framework for botanic garden policies, programs, and priorities in biodiversity conservation (Wyse Jackson and Sutherland, 2000). The key purpose was to assist botanic gardens worldwide to set their own future agendas to address the changes that are impacting on the world and to help guide their future responses to the environmental crisis. The agenda was based on contributions from and consultations with more than 300 institutions and individuals throughout the international botanic garden, botanical, and conservation communities.

The International Agenda defines the global mission of botanic gardens worldwide in conservation as:

- Stemming the loss of plant species and their genetic diversity worldwide.
- Focusing on preventing further degradation of the world's natural environment.
- Raising public understanding of the value of plant diversity and the threats it faces.
- Implementing practical action for the benefit and improvement of the world's natural environment.
- Promoting and ensuring the sustainable use of the world's natural resources for present and future generations.

A registration system for botanic gardens was subsequently launched and botanic gardens were invited to adopt the International Agenda as their (or part of their) institutional policy on conservation. Five hundred botanic gardens worldwide registered their contributions to the achievement of the International Agenda by making a written undertaking to work for the implementation of its provisions.

The *International Agenda for Botanic Gardens* has also been formally recognized by the international community and specifically by the United Nations Convention on Biological Diversity, which noted that it represents an important component of and contribution to the achievement of the Convention's Global Strategy for Plant Conservation (Secretariat for the Convention on Biological Diversity, 2002).

Global Strategy for Plant Conservation

For botanic gardens, the adoption of the Global Strategy for Plant Conservation (GSPC) by the Convention on Biological Diversity in April 2002 was both a significant landmark, and a major achievement. Botanic gardens were central to the development and subsequently to the implementation of this strategy (Wyse Jackson, 2002). It clearly demonstrated the value of botanic gardens immersing themselves in the political world of biodiversity policy making and advocacy.

The adoption of the Global Strategy for Plant Conservation in 2002, therefore, presented significant new opportunities and challenges for the botanical community throughout the world. The Strategy was developed in response to a growing realization that up to 100,000 plant species are currently threatened worldwide and new concerted action to promote programs focused on plants is urgently needed if huge losses in plant diversity are to be averted (**Box 6**).

The first phase of the GSPC (2002–2010) included a series of 16 outcome-orientated targets that proposed what needed to be achieved for plant conservation by 2010 (Secretariat of

Box 6 Global strategy for plant conservation

The ultimate and long-term objective of the Global Strategy for Plant Conservation is to halt the current and continuing loss of plant diversity. The strategy provides a framework to facilitate harmony between existing initiatives aimed at plant conservation, to identify gaps where new initiatives are required, and to promote mobilization of the necessary resources. It also provides a tool to enhance ecosystem conservation and the sustainable use of biodiversity and to focus on the vital role of plants in the structure and functioning of ecological systems and assure the continued and future provision of the goods and services such systems provide.

The scope of the strategy is:

- Understanding and documenting plant diversity
- Conserving plant diversity
- Using plant diversity sustainably
- Promoting education & awareness about plant diversity
- Capacity building for plant diversity

the Convention on Biological Diversity, 2002). The GSPC also provided a new and innovative framework against which government programs and the initiatives undertaken by a wide range of national and international organizations could be aligned. Through the adoption of the GSPC, governments throughout the world committed themselves to the implementation of the GSPC and to the achievement of the 16 international targets. In 2009, a comprehensive report on its achievements and progress towards the targets made was published (Secretariat of the Convention on Biological Diversity, 2009). Although the first phase of the GSPC (2002–2010) had almost come to an end, an updated Strategy for the period 2010–2020 was adopted. This strategy includes revised targets and renewed objectives, with the aim not only of moving forward the international plant conservation agenda with continued urgency, but also integrating such efforts with the CBD's Strategic Plan for Biodiversity 2011–2020 (Secretariat of the Convention on Biological Diversity, 2010). Both these strategies were adopted during the CBD's Conference of the Parties meeting in Nagoya, Japan in November 2010. The updated Strategy also seeks to align plant conservation policies more effectively with each country's development agenda, helping to address the causes of decline of plant diversity and ensure that economics, sustainable development, and biodiversity can go hand in hand.

Botanic Garden Roles in Addressing Climate Change

The achievement of plant conservation targets becomes even more challenging due to climate change. There are still few scientific assessments made on the impact of climate change on plant diversity and plant communities and there has to date been little incorporation of climate change mitigation and adaptation measures into National Biodiversity Strategies and Action Plans (particularly related to plants, plant-based ecosystems, and their conservation). Botanic gardens clearly have a responsibility and real opportunity to respond effectively to the climate change challenge. In Australia, for example, the Council of Heads of Australian Botanic Gardens (CHABG) has developed and is implementing a national response from the country's capital city botanic gardens to climate change (CHABG, 2008).

At the risk of oversimplification, the key botanic garden roles in climate change can be broken down under five major headings:

1. Leadership, public awareness, education, and advocacy
2. Botanical and Ecological Research
3. *Ex situ* conservation of germplasm
4. *In situ* management of species and ecosystems, and the
5. Utilization of germplasm

Leadership, Public Awareness, Education, and Advocacy

Botanic gardens receive huge numbers of visitors annually (approximately 250 million) and have great opportunities to become centers of public awareness on the:

- realities and consequences of climate change
- importance of plants
- impact of climate change on plant diversity
- need to take immediate action
- contribution of botanic gardens to addressing climate change.

Botanical and Ecological Research

In botanical and ecological research related to climate change, many traditional and on-going botanic garden activities are highly relevant, such as plant taxonomy, exploration and monitoring of species and their ecosystems, status surveys and increasingly, the responses of plants and ecosystems to a changing environment.

The responses of plants to climate change are complex and in general, very poorly understood. These will be the drivers of plant extinction in the coming years and, if the loss of plants as a result of climate change is to be prevented, there will need to be greater knowledge of how individual plants and populations respond to the following factors, as well as the interactions between them:

- changed competition regimes
- changed seasonal temperatures and patterns
- changes in available water
- changed snow/frost/freezing regimes
- loss of pollinators
- inbreeding and population viability changes
- changed nutrient status of soils
- pests, diseases, and invasive species
- disturbance and ecological succession
- migrations
- food chains
- rising carbon dioxide (CO_2)
- trophospheric ozone (O_3)
- nitrogen cycle changes
- light levels
- methane (CH_4)

The broad fields of conservation biology and population genetics are of fundamental importance. They provide the basis for what we need to know in order to undertake species recovery, ecological restoration and increasingly the reconstruction or construction of new habitats and ecological units.

Research on invasive species and pathogens is an additional area where botanic gardens will be expected to contribute, using their diverse collections and knowledge to monitor and predict potentially damaging invasions and develop effective methodologies to manage, control, or eradicate them.

The importance of safeguarding and extending exists and new ecological corridors have often been stressed as a means to help mitigate the impact of climate change on biodiversity. In particular, these can assist the movement of pollinators and seed dispersers. For plants in most cases, however, the rates of climate change predicted will not provide sufficient time to migrate at a sufficiently rapid rate. In those cases assisted migration of species may become one of the only options available and botanic gardens will play a critical role in these activities.

Ex situ Conservation of Germplasm

The conservation of germplasm and the associated scientific research by botanic gardens is expected to represent a major component of their contribution to addressing the impacts of climate change on biodiversity.

Maintaining threatened plants in botanic gardens worldwide is already a significant component of their work, which is a well-recognized contribution towards the achievement of Target 8 of the Global Strategy for Plant Conservation. There may well be a case for formulating a long-term botanic garden action plan and target, say for 2050, to provide a basis for planning what capacity is needed for ex situ conservation through botanic gardens over the next 40 years.

The ongoing development of storage of seed through botanic garden seed banks will continue to form a major component of botanic garden contributions to *ex situ* conservation. Notable examples include the international efforts as part of the Millennium Seed Bank Partnership, an initiative of the Royal Botanic Gardens Kew, and the European ENSCONET project, as well as newly formed initiatives such as the Australian Seed Bank Partnership, which aim to safeguard Australia's flora through a national network of conservation seed banks.

In situ Management of Species and Ecosystems

Botanic gardens are also likely to be increasingly active in *in situ* management of species and ecosystems providing active interventions on species growing in the wild to try to reduce the worst impacts of climate change (e.g., clearance of invasive species and habitat enhancement).

In March 2011, leading botanic gardens came together to initiate the Ecological Restoration Alliance and use the combined wealth of their expertise, research knowledge, and unparalleled diversity of plant materials to restore degraded ecosystems. The alliance is an initiative of BGCI working with these botanic gardens to develop and test good practice guidance at pilot sites around the world. This program will build on existing efforts of individual gardens, unite botanic gardens as a leading force in ecological restoration, and guide strategic action to restore threatened habitat around the world. The program aims to demonstrate that the essential components of successful restoration are the natural plant diversity of an area, the involvement of local people in the management and design of the program, and the acceptance of the impact of climate change (BGCI, 2011).

Utilization of Germplasm

The role of botanic garden in the utilization of germplasm is also likely to become increasingly important. This will be the case as new plants will be required for new situations, such as: breeding, development, and use of alternative economic plants, varieties and crops; supporting changing livelihoods (particularly in developing countries); and providing germplasm for ecological and restoration purposes. The importance of further developing collections of crop wild relatives in botanic gardens should be highlighted.

A key area where botanic gardens are increasingly playing a lead role is in the utilization of germplasm in species recovery and habitat restoration, including managing the migrations of species and the reestablishment of ecosystems in the face of climate change.

In outlining these five key areas for botanic garden involvement in climate change it should be noted that botanic garden research contributions will play a valuable cross-cutting role in all areas, including:

- understanding public attitudes to climate change and plants;
- the plant taxonomy that is required to underpin all botanical research;
- *in situ* and *ex situ* conservation methodologies and technologies;
- understanding the biology of species, ecosystems, and the goods and services they provide.

Botanic garden responses to biodiversity conservation, and in particular to climate change, will also involve fundamental changes in their horticultural practices and in the nature, composition and management of their living collections.

Botanic Gardens and the Sustainability Movement

Achieving environmental sustainability has become a major concern and challenge for governments and most major institutions worldwide as there is increasing awareness that the Earth's natural resources are finite and need to be used and managed with care and responsibility. It is recognized that sustainability has also become a matter of individual responsibility and one of the greatest personal challenges being faced by individuals is ensuring that our actions are not detrimental to the planet or threaten its environment or biodiversity. Botanic gardens, as leading environmental

organizations, face an even greater challenge not only to improve their own environmental performance, but also to ensure that the institutions are developed as models of sustainability.

It is clear that botanic gardens also need to provide leadership and guidance for those who still remain to be convinced of the urgency of the environmental crisis and the need to stem the loss of biodiversity. However, developing an individual botanic garden to become such a model is not an easy task. Most botanic gardens are very significant producers of carbon dioxide and reducing their carbon footprint requires innovation, determination, and sacrifices. In the temperate world most major botanic gardens have large expanses of heated glasshouses and very few are currently maintained by the use of renewable energy sources. Retrofitting botanic gardens to become modern energy-efficient institutions with a low environmental impact is a particularly difficult task, especially for those gardens that have inherited a diverse range of historic buildings, facilities, and glasshouses.

Nevertheless, despite the challenges we face it is clear that much can be achieved and constant improvements can be made to the ways in which we perform environmentally. To develop a more coherent response to the need to promote sustainability, some botanic gardens have developed and adopted a range of new policies and practices to guide their actions in this regard. The key components of such sustainability programs can be categorized under the following headings:

- Developing environmentally friendly practices in horticulture and botanic gardens management
- Promoting and supporting biodiversity conservation
- Supporting sustainable development locally, nationally, and internationally
- Education for and about the environment
- Recycling and reducing waste and energy consumption

Developing Environmentally Friendly Practices in Horticulture and Botanic Gardens Management

The key activities in relation to environmentally sustainable horticulture in botanic gardens include composting programs, the reduction, and eventual elimination of the use of peat from nonrenewable sources, elimination of harmful pesticide use and a major reduction in the use of herbicides. Many botanic gardens now work to ensure that almost all organic waste materials generated are composted, including waste paper from offices, much of which is now composted. In glasshouses, integrated pest management solutions are being increasingly applied as an alternative to using pesticides. A significant development in some botanic gardens is the creation of fruit and vegetable display and demonstration gardens, sometimes managed according to organic gardening principles. These often become the focus of public education programs in horticulture and are used for composting workshops as well as for gardening classes and demonstrating specific horticultural topics.

Encouraging and conserving wildlife in botanic gardens are also recognized as a component of their programs. This has included the development of biodiversity inventories of the wild organisms found in the institution, including birds, insects, and other invertebrates, plants, and fungi (e.g., Willis and Morkel, 2007).

Promoting and Supporting Biodiversity Conservation

Support for the implementation of the UN Convention on Biological Diversity (CBD) has increasingly been mainstreamed into the activities of botanic gardens worldwide. In particular the Global Strategy for Plant Conservation has been adopted as a central component of the priorities and policies of many. The membership of Botanic Gardens Conservation International has steadily increased since its establishment in 1987, indicating enhanced interest in and involvement in biodiversity conservation issues. Several botanic gardens and their network organizations have become members of the Global Partnership for Plant Conservation (see www.plants2020.net), an informal network of leading institutions and organizations supporting the achievement of the GSPC.

Many botanic gardens maintain collections of plants that are listed by IUCN, or listed in national legislation, as threatened in the wild. Data on them are generally included in BGCI's on-line database, the Plant Search.

Supporting Sustainable Development Locally, Nationally, and Internationally

A range of activities are undertaken by botanic gardens to support sustainable development at various levels. This includes policy development, support for sustainable development initiatives, such as Fair Trade projects, the development of energy conservation, recycling, and other resource conservation initiatives. Capacity building for sustainable development is also a key function of many botanic gardens, which offer training places in horticulture, botany, conservation, and other disciplines. Such training opportunities are often made available to international participants, such at the International Diploma Courses offered by the Royal Botanic Gardens Kew in Botanic Garden Management, Botanic Garden Education, Plant Conservation Techniques, and Herbarium Management. Other botanic gardens carry out their capacity building and training initiatives in association with local universities and colleges, offering undergraduate and postgraduate degree opportunities.

Education for and about the Environment

Sustainability is a key and fundamental part of the education programs of many botanic gardens, much of which is focused on the theme of education for sustainability. Examples of specific programs in this area include an annual educational event organized by the National Botanic Gardens of Ireland, which hosts a special Sustainable Environment Fair each year. A range of environmental organizations are invited into the Gardens for the event and other activities are organized, such as an organic Farmers' Market, recycled wooden furniture sales, basketry and wood turning demonstrations, guided tours of our sustainability efforts, and sales of fair trade and

other products. The Missouri Botanical Garden in St Louis, USA undertakes a similar event, with an annual Green Homes and Great Health Festival where best practice and demonstration of up-to-date methodologies and technologies in sustainability are demonstrated. At the Missouri Garden a special department called 'Earthways' has become a leader in providing guidance, support, and mentorship for local communities, companies, and other organizations and groups in developing sustainable practices.

Recycling and Reducing Waste and Energy Consumption and Conservation

One of the greatest sustainability challenges for most botanic gardens has been in the area of waste reduction and recycling. Each year visitors to botanic gardens leave vast quantities of waste in litter bins and trashcans, all of which has to be disposed of. Some gardens have taken the approach of providing separate receptacles for recyclable and compostable materials. Others have removed waste receptacles entirely and ask visitors to carry away their own waste and dispose of it responsibly. A number of gardens operate a single-stream recycling program, where waste is separated after collection and segregated into compostable, recyclable and waste that goes to landfill. To assist this conversion of waste, some botanic gardens have introduced new compostable materials into their operations, such as compostable cups and other containers, as well as the removal of plastic water bottles from sale and the installation of new water-filling stations so that visitors can refill reusable water containers.

Some botanic gardens have adopted policies to reduce energy consumption and using energy from sustainable sources. This has involved retrofitting existing buildings in some cases to make them more energy efficient. Efforts are made to reduce minimum temperatures in the greenhouses, offices, and other buildings. A 'Turn down/switch off' campaign for staff is often adopted, sometimes promoted and guided by Green Teams, or groups of energy conservation officers chosen from the staff of the gardens and representing each area of the institution. Setting annual targets for energy consumption reductions has been found by some institutions to be the best way of achieving or monitoring progress in making energy savings.

Botanic Gardens and Human Health

Botanic gardens throughout the world have been linked to human health issues for hundreds of years. As discussed earlier, the first botanic gardens of the world, established in Italy nearly 470 years ago, were created to assist in the teaching of medical students as part of universities in Pisa (1543), Padua (1545), Florence (1545), and Bologna (1547). Elsewhere in Europe other early botanic gardens (Physic Gardens) associated with medicinal plant study include Zurich, Leiden, Leipzig, Paris, Montpellier, and Oxford.

There is a growing new awareness of the importance of plants for primary health care amongst botanic gardens, even in today's modern world. This development can be traced back to 1988 when a ground-breaking initiative on Plants for Healthcare, the Chiang Mai Declaration, was developed by the World Health Organization (WHO), IUCN, and WWF. This reminded the world community that 80% of the world's people depend on traditional medicine for their primary health care needs. This gave rise to heightened awareness that the use of medicinal plants in primary health care needs to be recognized and safeguarded. In addition, the Chiang Mai Declaration led to the publication of new *Guidelines on the Conservation of Medicinal Plants* (WHO, 1993), which helped to promote international cooperation and coordination in medicinal plant conservation. The importance of plants for healthcare should not be underestimated. For example, traditional Chinese medicine alone uses more than 5000 plant species and traditional medicines in India are based on 7000 different species.

Today many botanic gardens maintain specific interests in medicinal plants, health care and in the promotion of human well-being. Wyse Jackson and Dennis (1998) recorded more than 440 botanic gardens that are known to maintain specific collections of medicinal plants. In that regard, there are key roles in medicinal plants played by botanic gardens in many developing regions, notably in countries such as Cameroon, China, Colombia, India, Sri Lanka, and Vietnam. The roles of such gardens include research on plant use in primary healthcare, studies on the sustainable harvesting and use of wild medicinal plant resources, and public education. In developed countries, particularly in Europe and North America, botanic gardens too continue to play roles in medicinal plants, specifically in areas such as raising public awareness and education and some aspects of pharmaceutical research, including bioprospecting for new sources of plant-based drugs.

The work of botanic gardens in medicinal plants is just one aspect of their role in global health promotion. Botanic garden work and priorities are increasingly directed towards managing and conserving biodiversity and ecosystems, both of which represent part of the fundamental foundations for human health and well-being.

The Millennium Ecosystem Assessment (2005) states that biodiversity is the foundation for human well-being. The World Health Organization's (WHO) contribution to the Millennium Ecosystem Assessment (2005) suggests that human health is strongly linked to the health of ecosystems, which meet many of our most "critical needs." Furthermore, the heads of the five biodiversity-related conventions recently stated:

Biodiversity can indeed help alleviate hunger and poverty, can promote human health, and be the basis for ensuring freedom and equity for all.

A report by Botanic Gardens Conservation International (Waylen, 2006) suggests that botanic gardens, through their biodiversity focused work, can make contributions to promoting human health and well-being, which can be categorized as follows:

- Improving healthcare
- Medicinal plant conservation and research
- Improving nutrition

- Poverty alleviation
- Community and social benefits.

Specific work undertaken by botanic gardens in this regard includes:

- *Ex situ* (off-site) conservation of medicinal plants, particularly those that are rare or endangered;
- Undertaking research in taxonomy, ecology, horticulture, ethnobotany, species conservation and biology, habitat restoration and management, and other areas;
- Public education;
- *In situ* conservation of plant resources in the wild (both wild plant species as well as crop plants and their wild relatives), in protected areas and in the wider landscape;
- Community-based programs, such as medicinal plant initiatives, city greening activities, and tree planting;
- Leisure services, such as the provision of green spaces for public relaxation;
- Collaboration with other bodies, organizations, and institutions in promoting healthcare and human well-being.

A recent report by BGCI (Hawkins, 2007) on medicinal plant conservation and botanic gardens elaborates on these aspects further and provides details on a wide range of examples from across the world.

Conclusion

The work of botanic gardens has adapted to changing needs and throughout time these institutions have contributed to social and economic development; from the early gardens founded to grow medicinal plants, to the eighteenth and nineteenth century gardens created for studies of economic botany and industry and trade development and more recently, the mobilization of botanic gardens to take an active role in *ex situ* conservation and support integrated conservation management. As our understanding of the impact of climate change unfolds and plant diversity faces increasing pressures, the work of botanic gardens will continue to adapt to address global priorities and meet society's need.

Appendix 1

List of Courses

1. Plant Conservation
2. Conservation and Plant Science
3. Biodiversity and Taxonomy of Plants
4. Plant Biology and Conservation
5. International Environmental Policies

Appendix 2

Numbers of botanic gardens known in each country and territory

Albania	1	Greece	11	Panama	2
Algeria	3	Greenland	1	Papua New Guinea	4
Angola	1	Grenada	1	Paraguay	1
Anguilla	1	Guadeloupe	2	Peru	10
Antigua and Barbuda	1	Guatemala	2	Philippines	12
Argentina	48	Guyana	1	Poland	32
Armenia	3	Haiti	2	Portugal	13
Australia	132	Honduras	4	Puerto Rico	5
Austria	20	Hungary	14	Reunion	4
Azerbaijan	4	Iceland	2	Romania	15
Bahamas, The	3	India	131	Russian Federation	109
Bangladesh	4	Indonesia	5	Rwanda	2
Barbados	3	Iran	3	Saint Helena	1
Belarus	8	Iraq	1	Saint Kitts and Nevis	1
Belgium	28	Ireland	16	Saint Lucia	1
Belize	3	Israel	13	Saint Vincent and the Grenadines	1
Benin	5	Italy	106	Samoa	1
Bermuda	1	Jamaica	6	Saudi Arabia	4
Bhutan	1	Japan	65	Senegal	5
Bolivia	5	Jordan	1	Serbia and Montenegro	3
Bosnia and Herzegovina	4	Kazakhstan	11	Seychelles	1
Botswana	1	Kenya	14	Sierra Leone	2
Brazil	40	Korea, North	2	Singapore	2
Bulgaria	10	Korea, South	51	Slovakia	10
Burkina Faso	3	Kuwait	1	Slovenia	5
Burma (Myanmar)	3	Kyrgyzstan	5	Solomon Islands	1
Burundi	2	Lao	1	South Africa	22
Cameroon	4	Latvia	2	Spain	28
Canada	103	Lesotho	3	Sri Lanka	6
Cape Verde	1	Liberia	1	Sudan	1

Cayman Islands	1	Libyan Arab Jamahiriya	1	Suriname	1
Chile	12	Lithuania	9	Sweden	9
China	148	Luxembourg	1	Switzerland	25
Colombia	28	Macedonia (FYR)	7	Taiwan	5
Congo	1	Madagascar	3	Tajikistan	5
Congo, Democratic Republic of the	3	Malawi	4	Tanzania	6
Costa Rica	8	Malaysia	12	Thailand	8
Cote d'Ivoire/Ivory Coast	1	Malta	1	Togo	3
Croatia	14	Martinique	2	Trinidad and Tobago	2
Cuba	12	Mauritius	2	Tunisia	4
Czech Republic	27	Mexico	95	Turkey	8
Denmark	10	Moldova	3	Turkmenistan	1
Dominica	2	Monaco	1	Turks and Caicos Islands	1
Dominican Republic	1	Mongolia	1	Uganda	4
Ecuador	10	Montserrat	1	Ukraine	38
Egypt	8	Morocco	4	United Arab Emirates	2
El Salvador	1	Mozambique	4	United Kingdom	117
Estonia	3	Namibia	1	United States of America	733
Ethiopia	4	Nepal	2	Uruguay	2
Fiji	3	Netherlands	52	Uzbekistan	4
Finland	8	Netherlands Antilles	1	Vanuatu	1
France	104	New Zealand	20	Venezuela	13
French Guiana	2	Nicaragua	2	Vietnam	6
Gabon	1	Nigeria	16	Virgin Islands (British)	1
Gambia	1	Norfolk Island	1	Virgin Islands (U.S.)	2
Georgia	5	Norway	6	Zambia	1
Germany - Deutschland	104	Oman	3	Zimbabwe	4
Ghana	5	Pakistan	8	Total known botanic gardens	3026
Gibraltar	1	Palau	1		

See also: Biodiversity and Human Health. Climate Change and Extinctions. Convention on Biological Diversity. Endangered Plants. Gene Banks. *In Situ*, *Ex Situ* Conservation. Loss of Biodiversity, Overview. Plant Conservation. Restoration of Biodiversity, Overview. Sustainability and Biodiversity. Zoos and Zoological Parks

References

BGCI (2011) Ecological Restoration Alliance (online). Available from http://www.bgci.org/ourwork/era/ (accessed 17/02/2012).

BGCI (2012) *Ex Situ* Conservation (online). Available from http://www.bgci.org/ourwork/ex_situ/ (accessed 19/02/2012).

Brockway L (2002) *Science and Colonial Exploration: The Role of the British Royal Botanic Gardens*. New Haven & London: Yale University Press.

Bruni S, Sampaio Pereira T, Gonzalez JAP, Correia Lima ME, Ferrao M de SR, and Matos A de A (2000) *Directory of the Botanical Gardens of Brazil*. Rio de Janeiro, Brazil: The Brazilian Network of Botanic Gardens. *Expressao e Cultura*.

Council of Heads of Australian Botanic Gardens (2008) *National Strategy and Action Plan for the Role of Australia's Botanic Gardens in Adapting to Climate Change*. Commonwealth of Australia.

Dodd J and Jones C (2010) *Redefining the Roles of Botanic Gardens – Towards a New Social Purpose*. London: Research Centre for Museums and Galleries Leicester and Botanic Gardens Conservation International.

Given D (1993) Conservation of biodiversity – A fundamental for botanic gardens. In: Froggatt P and Pates M (eds.) People and Plants and Conservation: Botanic Gardens in the 21st Century, pp. 37–41. Christchurch, New Zealand: Royal New Zealand: Institute of Horticulture.

Hawkins B (2007) *Plants for Life: Medicinal Plant Conservation and Botanic Gardens*. London, UK: Botanic Gardens Conservation International.

Henderson DM (1983) *International Directory of Botanical Gardens*, 4th edn Koenigstein: Koeltz Scientific Books.

Henderson DM and Prentice HT (1977) *International Directory of Botanical Gardens*, 3rd edn. Regnum vegetabile, vol. 95.

Heywood V (1987) The changing role of the botanic garden. In: Bramwell D, Hamann O, Heywood V, and Synge H (eds.) *Botanic Gardens and the World Conservation Strategy*, pp. 3–18. London: Academic Press.

Heywood V (1990) Botanic gardens, development and the conservation of plant resources. In: He SA, Heywood VH, and Ashton P (eds.) Proceedings of the International Symposium on Botanical Gardens, *Nanjing China September 25-28 1988*, pp. 3–15. Nanjing China: Jiangsu Science and Technology Publishing House.

Heywood V (1991) Developing a strategy for germplasm conservation in botanic gardens. In: Heywood V and Wyse Jackson P (eds.) *Tropical Botanic Gardens: Their Role in Conservation and Development*, pp. 11–23. London: Academic Press.

Heywood CA, Heywood VH, and Wyse Jackson P (1990) *International Directory of Botanical Gardens*. Koenigstein, Germany: Koeltz Scientific Books.

Hill AW (1915) The history and functions of botanic gardens. *Annals of the Missouri Botanical Garden* 2: 185–240.

Holttum RE (1970) The Historical Significance of Botanic Gardens in South East Asia. *Taxon* 19(5): 707–714.

Howard R, Wagenknecht BL, and Green PS (1963) *International Directory of Botanical Gardens*, 1st edn. Regnum vegetabile. vol. 28.

Huxley A (ed.) (1992) *The Royal Horticultural Society Dictionary of Gardening. Volume 1 A-C*. London: Macmillan Press Limited.

IUCN-BGCS and WWF (1989) *The Botanic Gardens Conservation Strategy*. U K: IUCN Botanic Gardens Conservation Secretariat.

Maunder M, Havens K, Guerrant E, and Falk D (2004a) *Ex situ* methods: A vital but unused set of conservation resources. In: Guerrant E, Havens K, and Maunder M (eds.) Ex Situ *Plant Conservation*, pp. 3–20. Washington: Island Press.

Maunder M, Havens K, Guerrant E, and Falk D (2004b) Realizing the full potential of *ex situ* contributions to global plant conservation. In: Guerrant E, Havens K, and Maunder M (eds.) Ex Situ *Plant Conservation*, pp. 389–418. Washington: Island Press.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well Being: Biodiversity Synthesis* (online). Washington, DC: World Resources Institute. Available from http://www.maweb.org/documents/document.354.aspx.pdf (accessed 17/02/2012).

Netherlands Turkey Trade Directory (2012) *Agriculture/Horticulture* (online). *Available from* http://www.nttd.com.tr/pages/15-key-sectors (accessed 03/02/2012).

Pavord A (2000) *The Tulip*. London: Bloomsbury Publishing.

Prance G (2004) Introduction. In: Guerrant E, Havens K, and Maunder M (eds.) *Ex Situ Plant Conservation*, pp. xxiii–xxix. Washington: Island Press.
Prest J (1988) *The garden of eden the botanic garden and the re-Creation of paradise*. New Haven and London: Yale University Press.
Schiebinger L (2004) *Plants and Empire*. Cambridge, Massachusetts and London: Harvard University Press.
Secretariat of the Convention of Biological Diversity (2002) *Global strategy for plant conservation, brochure*. Montreal, Canada: Secretariat of the Convention on Biological Diversity.
Secretariat of the Convention on Biological Diversity (2009) *The Convention on Biological Diversity Plant Conservation Report: A Review of Progress in Implementing the Global Strategy of Plant Conservation (GSPC)*. Montreal: Secretariat for the Convention of Biological Diversity.
Secretariat of the Convention on Biological Diversity (2010) Strategic Plan for Biodiversity 2011–2020. Available online at http://www.cbd.int/sp/ (accessed 28/05/2012).
Smith P (2006) Wild seed banks and taxonomy. In: Leadlay E and Jury S (eds.) *Taxonomy and Plant Conservation*, pp. 294–304. UK: Cambridge University Press.
Stafleu F (1969) Botanical Gardens Before 1918. *Boissiera* 14: 31–46.
Thiers B [continuously updated] Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/ih
United Nations (1993) *Earth Summit Agenda 21 The United Nations Programme of Action from Rio*. USA: United Nations Department of Public Information.
United Nations Environment Programme (1994) *Convention on Biological Diversity Text and Annexes. Switzerland: The Interim Secretariat for the Convention on Biological Diversity. United Nations Environment Programme*.
Waylen K (2006) *Botanic Gardens – Using Biodiversity to Improve Human Well-Being*. Richmond UK: Botanic Gardens Conservation International.
WHO (1993) *Guidelines on the Conservation of Medicinal Plants*. Geneva, Switzerland: World Health Organization.
Willis C and Morkel A (2007) *National Botanical Gardens: Havens of Biodiversity. Veld and Flora 93(4) December*. The Botanical Society of South Africa.
Wyse Jackson P (2001) An international review of the *ex situ* plant collections of the botanic gardens of the world. *Botanic Gardens Conservation News* 3(6): 22–33.
Wyse Jackson P and Dennis F (1998) *Directory of Medicinal Plant Collections in Botanic Gardens*. London, UK: Botanic Gardens Conservation International.
Wyse Jackson PS (1999) *Experimentation on a Large Scale- An Analysis of the Holdings and Resources of Botanic Gardens. Botanic Gardens Conservation News Vol 3 (3) December 1999*. UK: Botanic Gardens Conservation International.
Wyse Jackson PS (2002) Development and adoption of the global strategy for plant conservation: an NGO's perspective. *Botanic Gardens Conservation News* 3(8): 25–32.
Wyse Jackson PS and Sutherland LA (2000) *The International Agenda for Botanic Gardens in Conservation*. UK: Botanic Gardens Conservation International.

S

Salmon

Michael H Schiewe, Anchor QEA, WA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Alvin Newly hatched salmon with unabsorbed yolk sac.
Anadromous Aquatic organisms that spawn and undergo early development in freshwater, migrate to sea to grow and mature, and return to freshwater to reproduce.
Fry Juvenile salmon that have absorbed their yolk sac and emerged from the gravel.
Grilse Atlantic salmon that spends only 1 year at sea before returning to spawn.
Iteroparous Aquatic species that can spawn more than one time.
Jack A sexually precocious male salmon that spends one winter or less at sea before returning to its natal stream to spawn.
Kelt A spawned out Atlantic salmon or steelhead that is returning to the ocean to recuperate, sexually mature, and spawn again.
Parr Older juvenile salmon that are distinguished by parr marks – prominent oval spots on their sides.
Semelparous Aquatic species that die after reproduction.
Smolt A downstream migrating salmon that has begun the physiological transition to seawater.

Introduction

The Pacific and Atlantic salmon are a diverse group of fish distributed throughout the Northern Pacific and Northern Atlantic Oceans and their adjacent freshwater coastal habitats. The Pacific Salmon include seven species in the genus *Oncorhynchus*: pink (*Oncorhynchus gorbuscha*), chum (*Oncorhynchus keta*), sockeye (*Oncorhynchus nerka*), coho (*Oncorhynchus kisutch*), and Chinook (*Oncorhynchus tshawytscha*) salmon are found in North America and throughout the Pacific rim; masu (*Oncorhynchus masou*) and amago (*Oncorhynchus amago*) salmon are found only in Asia. In addition, there are two trout members of the genus *Oncorhynchus* that have anadromous forms: steelhead (*Oncorhynchus mykiss*, the anadromous form of rainbow trout) and the sea-run cutthroat trout (*Oncorhynchus clarkii*). There is a single species of Atlantic salmon (*Salmo salar*). This article focuses on the anadromous salmonids of North America, and their long history of decline in abundance and diversity.

Classification

Over the past two centuries, fish have been the subject of extensive systematic research and countless taxonomic revisions – and among the some 24 000-plus named species of fishes, the salmonids have been no exception. Currently, the salmonids are members of the class *Osteichthyes* (the bony fishes), the subclass *Actinopterygii* (the ray-finned fishes), the division *Teleostei*, and the super order *Protacanthopterygii*. The *Protacanthopterygii*, which contains some 310-plus species, is divided into three orders: the *Salmoniformes* (the salmonids), the *Esociformes* (the pike and mudminnow), and the *Osmeriformes* (the smelts). There are about 70 species in the family *Salmonidae* (order *Salmoniformes*), which can be divided into three subfamilies: *Salmoninae* (salmon and trout), *Coregoninae* (whitefishes), and *Thymallinae* (graylings).

According to Behnke (1992), the earliest fossil record of the salmonids is *Eosalmo driftwoodensis*, which was discovered in Eocene deposits some 40–50 million years ago in British Columbia. The separation of the Atlantic Ocean group (*Salmo*) and the Pacific Ocean group (*Oncorhynchus*) likely occurred by the mid-Miocene about 15 million years ago. By the end of the Miocene (about 5 million years ago), the ancestral form of *Oncorhynchus* in North America evolved into two distinct lineages, one leading to the Pacific salmon and the other to the Western trout. The evolution of Pacific salmon is still the subject of scientific debate. Particularly intriguing is the question of whether salmon evolved first in freshwater, only later developing the ability to migrate to the ocean, or whether they evolved first as marine species that later developed the ability move into freshwater refugia to spawn and rear. No matter which the case, the currently recognized species are generally believed to have differentiated only 500 000 to 1 million years ago, which represents an extremely short time period for such extensive speciation to occur. Rapid speciation was probably favored by the geological and climatic history of North American landscape – constantly being shaped and

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00293-8

reshaped by volcanism, alternating periods of warm and cold climate, alternating arid and pluvial periods, and rearrangement of major drainage basins.

The classification of North American salmonids does not stop with the traditional Linnaean system of classification for plants and animals. In particular, salmonids have been classified into evolutionary significant units (ESUs) or distinct population segments (DPSs) to facilitate risk assessments under the federal Endangered Species Act. The ESU and DPS are for all practical purposes the same. By definition, an ESU is any group of populations that is distinct according to two criteria: (1) It must infrequently exchange genes with other conspecific population units; and (2) It must represent an important component in the evolutionary legacy of the species (Waples, 1995). The concept of the ESU and DPS as a fundamental unit of conservation under the Endangered Species Act has been reviewed and endorsed by the National Research Council (1995) and is an important landmark in evolutionary thinking applied to conservation. This is because the ESU emphasizes genetic structure and life history adaptation as a focus for protection as opposed to more typological definitions of species or subspecies based on appearance.

In Western North America, 52 distinct salmonid ESUs or DPSs have been identified, with 27 of them currently listed as threatened or endangered under the Endangered Species Act. On the East Coast of North America the Atlantic salmon Gulf of Maine DPS is listed as endangered.

Life Histories

Salmon and anadromous trout display an amazing variety of life history patterns, both among and within species. However, many features are held in common. The life cycle begins with females depositing eggs in nests or redds in gravel bottoms of rivers and lakes. The young emerge from the gravel as alevins or fry to rear in freshwater for periods varying from a few days to several years. As juveniles or parr approach the time for their seaward migration, they undergo physiological metamorphosis or smoltification during which, among other changes, they develop the ability to osmoregulate in the hypertonic marine environment. In the ocean, they often undertake extensive migrations, covering many thousands of miles, while they grow and mature. After periods that range anywhere from a few months to 4 or more years, adult salmon return with remarkable fidelity to their natal river and tributary to spawn and continue the cycle.

The key characteristic that distinguishes the Atlantic salmon and anadromous trout from the five species of Pacific salmon is the fact that Pacific salmon species are semelparous and die after spawning only once, whereas anadromous trout and Atlantic salmon are iteroparous and may spawn multiple times. As is evident in the ensuing sketches of species-specific life history traits, diversification among salmon primarily has involved permutations that alter the schedule and timing of freshwater versus ocean habitat use (Busby *et al.*, 1996; Groot and Margolis, 1991; Shearer, 1992; Trotter, 1989).

Pink Salmon (Also Known as Humpback Salmon)

Pink salmon, which are distributed throughout the North Pacific and Bering Sea north of 40° N, are the most abundant of the Pacific salmon. In North America, pink salmon spawn from August to November and generally return at smaller sizes than other salmon – ranging between 1.0 and 2.5 kg. Returning adults undergo only modest migrations, often spawning in intertidal areas. After emerging from the gravel, they swim directly to sea, where they migrate and feed extensively in the North Pacific. After 18 months at sea, they return to their river of origin to spawn and die. This fixed 2-year cycle leads to the genetic isolation of what are referred to as odd- or even-year runs. Typically, different geographic regions are dominated by odd- or even-year runs, but not both. The tendency of having only one type of run (odd or even) in each river may exist because the rivers are not productive enough to support spawning by pink salmon every year.

Chum Salmon (Also Known as Dog Salmon)

Chum salmon are second only to pink salmon in abundance, but because of their large size they comprise up to 50% of the marine biomass of salmon in the North Pacific. Historically ranging from the San Lorenzo River in California to the Mackenzie River on the Beaufort Sea to the north, returning adults can weigh as much as 20 kg. Like pink salmon, chum salmon tend to spawn low in river systems and migrate quickly to estuarine and marine waters, often spending extended periods rearing in estuarine habitats. Once offshore, chum salmon are widely distributed throughout the North Pacific, typically spending 2–5 years at sea, but occasionally as long as 7 years. There are two generally recognized races of chum salmon, which are distinguished by their time of return to freshwater: Early and late stocks, or summer- and fall-runs, with the fall-run populations far more prevalent.

Sockeye Salmon (Also Known as Red Salmon)

Sockeye salmon are the third most abundant of Pacific salmon, comprising 17% by weight and 14% by number of the salmon total. They display a wide variety of life history patterns, including a subspecies that spends its entire life in freshwater – the kokanee. The vast majority of sockeye salmon populations spawn either in tributaries associated with lakes or directly in lakes on gravel shoals. After emerging from the gravel, the juvenile sockeye salmon spend from 1 to 3 years rearing in lakes. There are, however, rare populations that spawn in the rivers not associated with lakes and populations that migrate directly to sea shortly after emergence from the gravel. Like the other salmon species with extended freshwater rearing, smolts tend to move quickly downstream, spending little time in the estuarine and near-shore habitats and are found widely distributed throughout the North Pacific Ocean. Adults typically spend 1–4 years at sea, before returning to spawn in their natal lakes and tributaries. Maturing adults enter freshwater from late spring to early summer, with spawning occurring from late July through January.

Coho Salmon (Also Known as Silver Salmon)

Coho salmon are the most widely distributed of the Pacific salmon, but they are the least abundant. The vast majority of the populations have a 3-year fixed cycle – spending about 18 months in freshwater and 18 months at sea. In North America, they range from Monterey Bay, California, in the south to the Krupuk River in Alaska. Coho salmon typically enter freshwater from August to January and quickly move upstream to spawn in coastal streams and the smaller tributaries of large river systems. After emergence from gravel, juveniles rear in the smaller tributaries, and overwinter an additional year before beginning their seaward migration during the spring of their second year of life. At sea, coho salmon tend to stay on the continental shelf, feeding and growing in marine waters at roughly the same latitude as their stream of origin.

Chinook Salmon (Also Known as King Salmon)

Chinook salmon are the largest of the Pacific salmon – returning at weights up to 45 kg. In North America they range from central California to Kotzebue Sound, Alaska, and spawn in diverse habitats ranging from tidewater to small tributaries located as far as 3200 km inland. Locally, runs tend to be identified by the time at which adults enter freshwater and begin their spawning migrations, hence the colloquial names spring-, summer, fall-, late fall-, and winter-run Chinook salmon. However, these designations do not capture the true biological diversity of this species.

Along with sockeye salmon, Chinook salmon display the greatest within-species variation in life histories of all the Pacific salmon, including variation in age at seaward migration, variation in duration of freshwater, estuarine, oceanic residency, variation in ocean distribution and oceanic migratory patterns, and variation in age and season of spawning migration. This variation is derived from two sources, one racial and the other environmental. The racial differences result from geographic and hence genetic isolation that divides Chinook salmon populations into one of two behavior forms, stream-type and ocean-type. Stream-type populations spend 1 or more years in freshwater before seaward migration; undertake extensive offshore migrations, and return to their natal streams during the spring or summer, often holding in freshwater for several months before spawning in late summer. In contrast, populations of ocean-type Chinook salmon migrate to sea before age 1, utilize estuarine and near-shore habitats for extended periods, often spending their entire ocean residence on the continental shelf, and return to their natal stream in the fall immediately before spawning. A second major source of life history variation in Chinook salmon is expressed as run-timing, age at outmigration, and age-at-spawning differences within races, within populations, and sometimes even within cohorts. For example, it has only recently been found that ocean-type fall Chinook salmon originating in the Snake River (Columbia River basin) display not only the typical pattern of subyearling migration, but also some variable proportion of a year class will also overwinter in the riverine environment and migrate to the ocean as yearlings the following year (Connor *et al.*, 2005). The adaptive value of such variation is well known in evolutionary theory – by spreading the timing of migration and spawning among years, these fish are effectively "spreading the risk" and reducing the likelihood of total failure if any single year or even week is disastrous (the argument works at a variety of temporal and spatial scales). Uncertainty in juvenile survival and marine productivity is enormous for salmon, and this is reflected by the fact that year-to-year variability in spawner counts for salmon (see **Figure 1**) have coefficients of variation typical of pest insects and other notoriously fluctuating species.

Atlantic Salmon

Atlantic salmon are the only anadromous salmonid native to the Atlantic coast of North America. Historically found in nearly every major coastal river north of the Hudson River, the extinction of several U.S. populations by the end of the nineteenth century has shifted the southern extent of the species distribution over 200 km northward. The majority of adult Atlantic salmon enter freshwater to spawn during the spring (late May to early June primarily), spending most of the summer holding in freshwater and then spawning in fall. In general, the more northern populations arrive earlier, and spawn earlier. However, a few populations do not enter freshwater until late summer or early fall, bypassing the summer holding period and consequently spawning directly. Juvenile Atlantic salmon emerge from the gravel in the spring and typically spend from 2 to 4 years rearing in freshwater (although some populations spend as little as 1 year and others as much as 8 years in freshwater). Outmigration of juveniles usually occurs from April to mid-June, with the smolts spending little time in the estuarine and near-shore zone. Atlantic salmon are widely distributed in the North Atlantic, with American and European stocks often concentrating and mixing in the waters off Greenland – a distribution which enabled the development of highly efficient commercial fisheries that nearly drove the species to extinction in the 1960s. The majority of Atlantic salmon spend two winters at sea before returning to spawn.

Steelhead

There are two major genetic groups of *O. mykiss* presently recognized in North America: the *inland group* and *coastal group*, with the distinguishing character being their use of freshwater habitat. As the name implies, the coastal form populates the wet coastal regions of western North America, whereas the inland form is found in the more arid interior regions of western North America. Both forms occur in British Columbia, Washington, and Oregon; Idaho has only the inland form, and California has only the coastal form. Both forms display a wealth of life history variation, including an anadromous form that is known as steelhead. Although steelhead can be observed entering freshwater to spawn throughout the year, they usually display striking seasonal peaks in spawning activity. This has resulted in the use of seasonal run timings to distinguish different populations (fall-run, summer-run, etc.). Steelhead can be divided into two basic reproductive ecotypes based on their state of sexual maturity at the time of river entry and the duration of

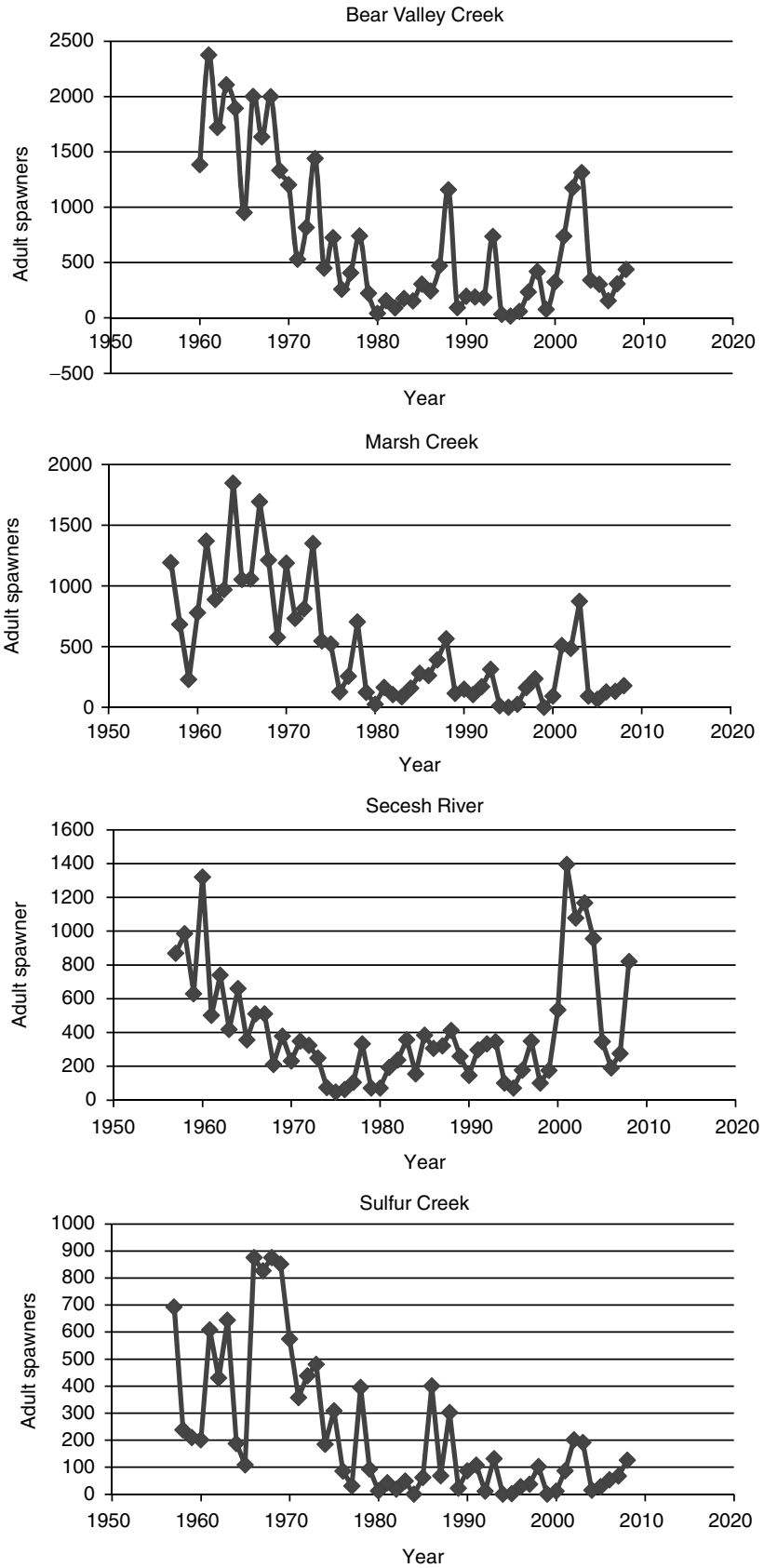


Figure 1 Counts of spring/summer Chinook salmon returning to four Salmon River tributaries, 1957–2008. Note that returns before the late 1970s were substantially higher than subsequent years. This abrupt decline coincided with completion of eight dams between the spawning grounds and the ocean; construction of several major hatcheries and the release of large numbers of hatchery fish; and a major shift in ocean conditions. In both the early 1980s and mid-1990s several Snake River populations were clearly at risk of extinction, but have since shown signs of rebuilding.

spawning migration. The stream-maturing type (commonly known as fall steelhead in Alaska, summer steelhead in the Pacific Northwest and northern California) enters freshwater in a sexually immature state and spends several months maturing before they spawn. In contrast, the ocean-maturing type (spring steelhead in Alaska and winter-run elsewhere) enters freshwater in a sexually mature state and spawns shortly thereafter. While many rivers contain populations of both ecotypes, coastal streams tend to be dominated by winter steelhead, whereas inland reaches tend to be predominately the summer steelhead.

Despite the major differences in migration timing of the adults, both ecotypes of steelhead generally spawn from December through March. After emerging from the gravel in late spring, juveniles spend up to 7 years in freshwater, with 1–3 years being the norm. Most populations in Alaska and British Columbia tend to smolt and migrate to sea at 3 years of age, whereas the more southern stocks tend to smolt at 1 or 2 years of age. Most North American steelheads spend 2 years at sea; however, some of the more southern stocks characteristically spend only 1 year at sea. An unusual variant in life history found among selected steelhead populations in Southern Oregon and Northern California is that of the half-pounder. Half-pounders are juvenile fish that migrate to sea, only to return to freshwater in an immature state after only 2–4 months. The biological basis for this unusual behavior is unknown. As an iteroparous species steelhead can spawn multiple times, with the frequency varying both within and among populations. The more northern populations tend to have a lower proportion of repeat spawners than those further south.

Sea-Run Cutthroat Trout

Cutthroat trout display a complex suite of life histories – some exclusively freshwater and others involving a period of marine residency. While it is the latter life history strategy that justifies their inclusion in this article, the reader should be aware that the interrelationships between the different life history variants are not fully understood, and sea-run forms probably cannot be treated as isolated from the freshwater forms (Trotter, 1989).

In North America, sea-run cutthroat trout are distributed along the Pacific Coast from the Eel River in California to Prince William Sound in Alaska. Adults typically spawn in the small tributaries of low-gradient streams and the lower-gradient downstream reaches of large river systems from late winter to late spring. Juveniles rear in freshwater for 2–4 years, generally utilizing slower moving waters and pools. Those populations that migrate into the protected waters of large estuaries and sheltered bays tend to smolt at age 2, whereas those populations migrating into the near-shore ocean often delay smoltification until age 3 or 4. Sea-run cutthroat trout seldom spend more than a few months in marine waters and stay well within the near-shore zone, typically returning to freshwater in late summer, fall, or winter of the same year they migrate to sea. Although fish returning to freshwater to spawn return to their natal stream with high fidelity, non-maturing fish are known to leave the ocean and overwinter in nonnatal streams. Like steelhead and Atlantic salmon, sea-run cutthroat are iteroparous, and multiple spawning is quite common.

Special Salmon Adaptations

Anadromy

Anadromous fishes are those that spawn in freshwater, migrate to the ocean to forage and mature, and return to freshwater to spawn and begin the cycle again. As anadromous species, the salmonids pass through multiple habitat transitions and face major physiological challenges – the most critical adaptation for this lifestyle is the ability to osmoregulate in both freshwater and marine environments, and to do so at precisely prescribed times. The evolution of anadromy has intrigued biologists for as long as salmon have been studied. Although many theories have been put forth to explain the evolution of anadromy, such a major life history character must, on balance, provide an overall advantage to individual fish. The potential advantages of anadromy include (a) a mechanism for dispersal and rapid recolonization of regions suddenly made available to fish (such as the retreat of a glacier); and (b) being able to get the "best of two worlds" (freshwater and marine). The "best" of the freshwater world involves reduced predation on eggs and juvenile fish because the embryos develop in the protected environment of gravel until they emerge at a large enough size to actively avoid stream-dwelling invertebrate predators. The "best" of the marine world involves higher growth rates for larger maturing fish in the oceans, made possible because in upwelling areas and years of high productivity the regions of the ocean frequented by salmon provide an abundance of larger prey items that far exceeds the prey base of most freshwater ecosystems.

Homing

The legendary ability of Pacific and Atlantic salmon to home with high fidelity to their stream of origin is one of the more spectacular migrations in nature. Beginning with the pioneering work of Hasler, Quinn, and others, it is growing increasingly clear that this ability involves a combination of magnetic detection, solar navigation, and olfactory cues. The salmon's ability to navigate the high seas appears to be based on the ability to detect inclination and declination of the earth's magnetic field, combined with celestial orientation and an "endogenous circannual rhythm, synchronized by day length or rate of daylight change to sense latitude" (Quinn, 1982). Once in the vicinity of their home stream, it appears that salmon homing is facilitated by following unique gradients of odors that are learned during the downstream migration of juveniles.

Despite uncertainty about the exact mechanism of homing, the fact that salmon return to their natal streams is ecologically quite significant. From a very practical perspective, it assures that salmon return to an area with favorable spawning and rearing conditions. In addition, although often considered a failure in homing, the small proportion of salmon in a population that stray to neighboring streams play an important role in the colonization of vacant habitat. Such a nearest neighbor strategy is more likely to be successful in matching salmon with habitat than a random pattern of straying, and hence it is a more efficient means of reestablishing sustainable populations.

Historical Status and Overview of Decline

Historically, spawning populations of Atlantic salmon were abundant in most rivers and their tributaries throughout Europe, the British Isles, and the Baltic North, and in all the major North America rivers north of the Hudson River along the Atlantic seaboard. Likewise, the thousands of rivers and streams dissecting the coastal lands surrounding the North Pacific Ocean supported major populations of Pacific salmon and anadromous trout. Today, however, these once plentiful species are greatly reduced in both abundance and distribution. Owing largely to the actions of humans and the encroachment of "civilization," populations of Atlantic salmon throughout much of Europe and in New England have all but disappeared from the landscape. Similarly, in the Pacific Northwest, only a fraction of the streams and rivers that once supported robust salmon populations continue to do so. Although there are some obvious differences, there are also some striking similarities in the causes and patterns of decline of Pacific and Atlantic salmon – in Europe the onset of the Industrial Revolution, in New England the colonization by European settlers, and in the Pacific Northwest the arrival of the Euro-American settlers. The 1991 review of Nehlsen *et al*. perhaps best captures the status of salmon in the Pacific Northwest by noting that, of 214 individual populations of native naturally spawning populations of salmon in California, Oregon, Idaho, and Washington, 101 were at high risk of extinction, 58 were at moderate risk of extinction, and 54 were of special concern. The situation has become so dire that the 27 federally listed endangered and threatened salmonid ESUs and DPSs span almost every major freshwater ecosystem from the Sacramento River northward, with vast political and social consequences for the western United States

Habitat, Harvest, Hatcheries, and Hydropower as Major Threats to Salmonid Diversity

Many factors, both natural and human caused, are major sources of mortality among populations of North American salmon. While there is a handful of examples of a single factor causing populations to go extinct (e.g., a dam completely blocking access to spawning and rearing habitat), more often than not there is a constellation of factors that place salmon at risk of going extinct. An extremely thorough review of the current status of Pacific Northwest salmon and the factors contributing to their decline was recently published by Lichatowich (1999). In *Salmon without Rivers*, Lichatowich chronicles the methodical extirpation of one salmon population after another as the Euro-American settlers and their industrial economy replaced what he refers to as the "gift economy" of the native Indian tribes. Given the widespread destruction of habitat and a "modern" harvest philosophy built on the false premise of unlimited productivity, it is indeed amazing we still have viable populations of salmon in the Pacific Northwest.

The human-caused threats to salmon are often described in a framework of the four Hs: habitat, harvest, hatcheries, and hydropower. This convenient framework is followed here, but it is also a little misleading in how it categorizes risk factors in tidy compartments. For example, while it has been traditional to view habitat threats as freshwater and affected only by human destruction, this view grossly neglects the importance of estuarine and marine habitats and the role they play in salmon productivity. Further, while there is a tendency to dismiss the ocean as a homogeneous "black box" that is unaffected by human activities, there is likely to be a strong interaction between some of the Hs such as hatcheries and ocean conditions. For example, the industrial-scale production of hatchery fish in large river systems such as the Columbia River Basin may not be problematic except during periods of low primary and secondary productivity in the near-shore ocean. Hatchery and naturally produced smolts may compete only during periods of low ocean productivity. In addition, the ill effects of individual fitness compromised by gene flow from hatchery fish may be felt only if wild fish are already stressed by poor ocean conditions.

Habitat

The destruction of freshwater habitat is perhaps the single most significant factor affecting Pacific and Atlantic salmon in North America. Logging, mining, and agriculture (including grazing, irrigation withdrawals, soil erosion, and chemical contamination) have been major causes of salmon decline. While many of the impacts of these activities are similar, some are unique. For example, the logging practices of the past two centuries has caused widespread destruction of riparian vegetation along stream corridors (producing increased in-stream temperatures) and has caused landslides due to poor road-building practices (resulting in sedimentation and destruction of spawning and rearing habitat). Other impacts, such as the creation of thousands of acres of impermeable surfaces, are unique to urbanization. While it is beyond the scope of this article to describe in detail the nature of all of the different forms of habitat destruction, it is important to note that most are clearly the product of human actions and hence controllable.

Harvest

Pacific and Atlantic salmon have been the target of harvest by the indigenous peoples of North America for many thousands of years. However, it was not until the early- to mid-1800s that the Euro-American settlers (with their advanced methods of capture and preservation) began to so overfish salmon that extinction became a possibility. Particularly destructive have been offshore fisheries that indiscriminately harvest mixtures of plentiful hatchery fish and dramatically less abundant wild fish. Although harvest is currently regulated to some extent, it still continues at substantial rates, even on endangered and threatened ESUs. For example, fall Chinook salmon from the Snake River are currently harvested at an annual rate that exceeds 20%, and some listed ESUs in the Puget Sound area are harvested at annual rates that exceed 40%. Fish taken by harvest represent individuals with maximum reproductive value and the impact of their mortality before spawning is unarguably large. It may be that the models guiding our harvest policy are themselves flawed, as suggested by recent theoretical papers which suggest that harvest models built with

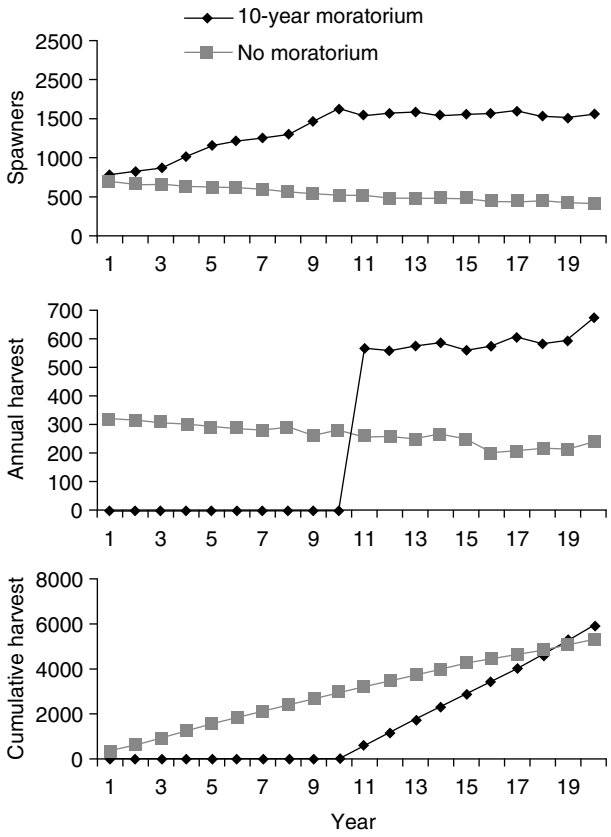


Figure 2 Analysis of the benefits of a temporary moratorium on harvest for Snake River fall Chinook salmon. A simple age-structured model iterates salmon populations forward 20 years, using estimates of variability and recruitment from 1980 onward. Two alternative simulations are run: simply continuing to harvest at the current rate versus halting harvest for 10 years and then reinitiating harvest for years 11 through 20, but at 75% of the current rate. Note that if there is a 10-year moratorium followed by harvest at 75% of the current rate, the population rebuilds and total cumulative harvest after 20 years (even including 10 years without any harvest) ends up being higher than if current practices are continued. One may be able to have fish and harvest, if harvest is temporarily halted.

the concept of an optimum percentage take and maximum sustained yield often fail in highly variable environments (Lande *et al.*, 1997; Ludwig, 1998). The good news is that simple models suggest that threatened ESUs such as Snake River fall Chinook salmon may be readily rebuilt by temporarily curtailing harvest – indeed with only a 10-year moratorium populations may recover sufficiently to then sustain a vigorous harvest that annually yields double the number of fish currently removed from the river (**Figure 2**). Certainly, the issue of harvest for these threatened stocks needs to be reevaluated – for perspective, it is hard to imagine anyone tolerating the hunting of threatened whales at a rate leading to 20% adult mortality per year, given the status of whales as listed under the Endangered Species Act.

Hatcheries

The building of hatcheries and the production and release of millions of hatchery salmonids in the Pacific Northwest closely tracked the wholesale destruction of habitat and overfishing. Indeed, hatcheries were widely viewed as the solution – and in fact a solution that facilitated the extraction of natural resources to proceed unchecked. It was not until the past several decades that biologists began recognizing that such use of hatcheries were, by themselves, a potential problem. Beyond the obvious creation of harvest opportunities and destructive mixed-stock fishing, there is now much greater recognition that the release of large numbers of hatchery fish can have detrimental ecological and genetic effects on natural populations.

Ecological impacts of the release of large numbers of hatchery fish include intra- and interspecific competition for food and space among hatchery and wild salmonids, predation by larger hatchery salmonids on their wild counterparts, transmission of infectious diseases from hatchery to wild salmonids, and altered migratory behavior of wild fish caused by the timing and magnitude of hatchery releases (reviewed by Flagg *et al.*, 2000). While many of these interactions were believed limited to the freshwater environment, a recent study suggests density dependent interactions between wild and hatchery salmonids in the marine environment as well (Levin *et al.*, 2001). The potential for density dependent competition in the marine environment was further highlighted in a recent analysis of trends in the combined abundance of adult pink, chum and sockeye salmon in the North Pacific Ocean (the three most abundant salmonid species). Ruggerone *et al.* (2010) compiled data indicating that combined abundance of these species doubled from 309 to 636 million over the past 50 years, with 22% of total attributable to hatchery releases of chum salmon.

Other recent studies indicate that some species of hatchery-origin salmonids, when they spawn in the natural environment, have reduced intrinsic productivity compared to that of their wild counterparts, and that the overall productivity of a mixed hatchery–wild population decreases as the proportion of hatchery fish increases. For example, Fleming and Gross (1993) and Araki *et al.* (2007) reported reduced reproductive fitness of coho salmon and steelhead due to genetic and environmental effects on size at maturation, run-timing, behavior, and other possible traits. Williamson *et al.* (2010) documented that hatchery-origin spring Chinook salmon in a tributary of the Columbia River produced about half the juvenile progeny per parent when spawning naturally than did natural-origin fish. To the extent that the reduced reproductive fitness is heritable, there is an obvious long-term potential for hatchery fish to eventually reduce the fitness of wild populations. Collectively these studies raise serious questions about the potential use of hatchery programs to supplement and contribute to the rebuilding of sustainable salmon populations.

Hydropower

Dams block access to spawning habitats, create reservoirs that flood spawning habitats, and may stress fish migrating through them to such a large extent that females cannot replace themselves due to depressed physiological fitness. Dams affect both upstream adult spawning migrations and downstream juvenile migrations. Primary mechanisms for the

negative impacts of dams involve increased mortality of juveniles associated with passage through turbines, atmospheric gas supersaturation, and decreased flow and turbidity. On the other hand, engineering improvements in dams can mitigate some of these effects, although certainly not to the extent that rivers return to the equivalent of their free-flowing natural state. Because dams are the most visible impediment to salmon survival, getting fish around, over, or through dams has been the focus of much research. Unfortunately, the understandable focus on hydropower as a risk factor for salmon has left the "other Hs" less well studied. Recovery of salmon is likely to often require more than simply removing dams – we may also need improvements in spawning habitats and in ocean and estuarine environments, further reductions in harvest, and perhaps major modifications of hatchery programs. Finally, it is worth pointing out that although removal of dams is often championed in terms of "returning to a normative river" (Independent Science Group (ISG), 1999), most proposed dam removals still leave many dams untouched and our rivers so altered by other factors that they remain a long way away from the ideal of a free-flowing "wild river."

Natural Environmental Variability and Climate Change

Although the tendency in recent years has been for resource managers to focus on the human-caused factors that have contributed to the decline of salmon, natural variation in environmental conditions, both freshwater and marine, are also major factors affecting the abundance and distribution of salmon. Operating on seasonal, annual, and decade-long timescales, events such as regime shifts, El Ninos, and changes in seasonal patterns of coastal circulation (particularly upwelling) all can have a profound effect on oceanic survival of salmon. Particularly noteworthy has been the research documenting correlations among broad patterns of productivity, and major physical features of the ocean–atmosphere interaction. For example, Hare *et al.* (1999) found that harvest (and presumably abundance) of Pacific salmon in Alaska has varied inversely with catches along the U.S. West Coast over the past 70 years. In exploring various explanations for this observation, they found a striking correlation between switching pattern of north–south abundance and the Pacific decadal oscillation (PDO; a recurring pattern of pan-Pacific atmosphere–ocean variability). While there are several ocean–atmosphere indices such as the PDO, they all, in concept, attempt to capture the key physical forces that drive decadal-scale basin-wide climate variability. For example, since the last major climate shift of 1976–77 in the North Pacific Ocean, oceanographers have documented decreased mixed-layer depths in the Subarctic Gyre and increased sea surface temperatures along the west coast. The biological responses to these physical changes were profound, including major changes in primary and secondary production, and micro- and mesoscale changes in zooplankton biomass and species composition. In addition to salmon, a large number of higher trophic level fish, birds, and mammals show similarly dramatic shifts in abundance and distribution coincident with these changes in ocean conditions.

Operating on an even grander scale, the impacts of long-term climate change will undoubtedly have profound consequences for all animal and plant species. Indeed, climate change is perhaps the most complex environmental and conservation challenge of the 21th century. Salmon, because of their anadromous nature, will suffer the dubious distinction of being subject to impacts in both the marine and freshwater habitats – a distinction that may well challenge their ability to adapt and survive over much of their current range. In the freshwater phase of their life cycles, salmonids in many areas will likely face major changes in temperature and precipitation patterns that will greatly alter the quantity and quality of habitat in ways that will challenge their ability to survive and adapt. This will likely profoundly affect their range and productivity over a latitudinal gradient. Climate change is also expected to depress oceanic productivity by disrupting upwelling patterns in a way that reduces the delivery of nutrients (Bakun, 1990). To the extent that salmonid population fluctuations are determined by bottom-up forces, the possibility of reduced coastal nutrient upwelling and warming conditions of offshore habitat could have dire consequences for these and many other fish species. All of these changes will require resource management and conservation policies that are robust to change, and adaptable to a changing environmental baseline (Schindler *et al.*, 2008).

Of the many studies designed to predict the consequences of climate change on aquatic species, one of the more dramatic is Welch *et al.* (1998). To explore potential marine impacts on salmon, Welch and colleagues compiled an extensive data set on distribution of salmon at sea and found that, for sockeye salmon, marine distribution was limited by sharp thermal boundaries. Using a climate model, and assuming a doubling of greenhouse gases over the next 50 years (which is consistent with the current rate of input), it was predicted that sockeye salmon would be excluded from most of the Pacific Ocean and restricted to a relatively small area of the Subarctic Pacific for much of the year. Similar analyses revealed marked reduced ocean habitat for other species of Pacific salmon as well.

What is the Value of Salmonid Biodiversity and Can We Rescue Salmon from Extinction?

Two arguments dominate the mainstream ecological literature regarding the "value of biodiversity": (1) diversity fosters enhanced productivity by allowing a more complete exploitation of the environment *via* a variety of different forms, each adapted to different habitats or resources, and (2) diversity fosters enhanced stability by spreading the risk and providing redundancy in the face of unpredictable catastrophes. Both hypotheses regarding the value of biodiversity apply well to salmon. First, different salmonid ESUs spawn in different portions of the freshwater environment (inland versus coastal, mainstream rivers versus streams, lakes versus rivers, and so forth) and migrate to different regions of the ocean. Any loss of salmonid diversity would mean that the productivity of these ecosystems would be reduced, with minimal likelihood that any other species could take the place of salmon (since the anadromous lifestyle is so unique). Second, due to

climatic fluctuations, different months of the year will be favorable for migration; in fact, recent data indicate that even a shift of a few days with respect to the time at which juvenile Chinook salmon arrive at the ocean can alter survival rates by a factor of two. If salmon diversity is reduced, then there will be less variety in run-timing and life history schedule and consequently much less insurance against the vagaries of climatic fluctuations. The precipitous decline of Pacific and Atlantic salmon throughout much of their range raises a number of important questions: Can we save the salmon? Should we save the salmon? And what do we need to do to save the salmon?

Can We Save the Salmon?

The answer to this question is an unequivocal yes. A species complex that has survived and flourished under the wide range of conditions from which salmon have evolved has almost certainly weathered numerous survival bottlenecks in the past, caused by environmental perturbations ranging from volcanic eruptions to the total loss of huge blocks of habitat to glaciers, fire, and floods. Perhaps the key attribute that has allowed the salmonids to survive these events is the tremendous diversity of life histories and patterns of habitat use. The salmonids represent a plastic species complex with a robust evolutionary potential that should facilitate survival under diverse conditions.

That said, many populations in the Pacific Northwest and Northeast North America are already extinct, and many others have never been at lower levels of abundance. Saving Pacific and Atlantic salmon will clearly require immediate action that will include economic sacrifice, societal discipline, and a commitment to a science-based recovery strategy – all at a very significant cost to the people of the affected regions.

Should We Save the Salmon?

Because of the societal implications and costs, this question perhaps extends well beyond the boundaries of biology. However, from strictly an ecological perspective, the salmonids are a vital component of the North Pacific and North Atlantic ecosystems and the associated coastal habitats. The ecological importance of salmonids as predators, prey, and nutrient recyclers is obvious to even the most casual observers of nature. While as scientists we can marvel at the complexity of an ecosystem and strive to understand species and environmental interactions in the broadest sense, we must also acknowledge that such complex and dynamic systems may never be fully understood. Nonetheless, the loss of such a key species complex as salmonids over such a broad geographic area would no doubt be a biological event without precedent.

What Do We Need to Do to Save the Salmon?

As noted earlier, the keys to solving the salmon problem are societal commitment, discipline, and a science-based recovery strategy. While societal commitment and discipline are not scientific issues, they certainly are not devoid of science. To make a decision on whether and how to proceed, *society* must have knowledge – knowledge of the complexity of the problem and the certainty (or lack thereof) associated with various courses of action. This knowledge is vital to the economic decision-making process.

In contrast, the design of recovery strategies is largely a science issue. At the core of every science-based recovery strategy should be an analytical framework that allows: (1) the methodical determination of survivorship in each of the life history stanzas of salmon, (2) the identification of actions and opportunities to increase survivorship, (3) the assessment of biological (and societal) feasibility of each action or suite of actions, and (4) a rigorous monitoring and evaluation plan. Whatever models are used should be clear and understandable to the average citizen, and data used as model inputs must be available to all.

Clearly, society cannot continue the extraction-based economy of the past 200 years and expect even remnant populations of salmon to exist over even a fraction of their former range. A major shift in environmental decision making will be required. In particular, we must abandon the habits of the past two centuries, which have involved almost universally approaching conservation through a "reduction from the status quo" perspective, and the belief that we can replace lost production with hatchery programs. For example, rather than arguing about whether we should reduce harvest by 10, 20, or even 50% from current levels, any harvest above 0% should require a scientifically sound justification. Moreover, rather than rely on artificial production to replace natural production from diverse habitats, the focus should be on restoring those lost habitats in a spatially and ecologically meaningful pattern. When populations face a high probability of extinction, modest adjustments of the status quo are not enough. Unless society seriously acts to rebuild salmon populations, the future of salmon is not assured, and moderate actions ought to be recognized as optimistic gambles (as opposed to reasonable moderation).

See also: Captive Breeding and Reintroduction. Captive Breeding and the Evolutionarily Significant Unit. Climate, Effects of. Fish Conservation. Resource Exploitation, Fisheries

References

Araki H, Cooper B, and Blouin MS (2007) Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science* 318: 100–103.

Bakun A (1990) Global climate change and intensification of coastal ocean upwelling. *Science* 247: 198–201.

Behnke RJ (1992) *Native Trout of Western North America*. Bethesda, MD: American Fisheries Society. American Fisheries Society Monograph 6.

Busby PJ, Wainwright TC, Bryant GJ, *et al.* (1996) NOAA Technical Memorandum NMFS-NWFSC-27. *Status Review of West Coast steelhead from Washington, Idaho, Oregon, and California*. Seattle, WA: US Department of Commerce.

Connor WP, Sneva JG, Tiffan KF, Steinhorst RK, and Ross D (2005) Two alternative juvenile life history types for fall Chinook salmon in the Snake River Basin. *Transactions of the American Fisheries Society* 134: 291–304.

Flagg TA, Berejikian BA, Colt JE, *et al.* (2000) NOAA Technical Memorandum NMFS-NWFSC-41. *Ecological and Behavioral Impacts of Artificial Production Strategies on the Abundance of Wild Salmon Populations*. Seattle, WA: U.S. Department of Commerce.

Fleming IA and Gross MR (1993) Breeding success of hatchery and wild coho salmon (*Oncorhynchus kisutch*) in competition. *Ecological Applications* 3: 230–245.

Groot C and Margolis L (eds.) (1991) *Pacific Salmon Life Histories*. Vancouver, BC: University of British Columbia Press.

Groot C, Margolis L, and Clarke WC (eds.) (1995) *Physiological Ecology of Salmon*. Vancouver, BC: University of British Columbia Press.

Hare SR, Mantua NJ, and Francis' RC (1999) Inverse production regimes: Alaska and west coast Pacific salmon. *Fisheries* 24: 6–14.

Independent Science Group (ISG) (1999) Scientific issues in the restoration of salmonid fishes in the Columbia River. *Fisheries* 24: 10–19.

Lande RB, Saether B, and Engen S (1997) Threshold harvesting for sustainability of fluctuating resources. *Ecology* 78: 1341–1350.

Levin PS, Zabel RW, and Williams JG (2001) The road to extinction is paved with good intentions: Negative association of fish hatcheries with threatened salmon. *Proceedings of the Royal Society London* 268: 1153–1158.

Lichatowich J (1999) *Salmon without Rivers*. Washington, DC: Island Press.

Ludwig D (1998) Management of stocks that may collapse. *Oikos* 83: 397–402.

National Research Council (1995) *Science and the Endangered Species Act*. Washington, DC: National Academy Press.

National Research Council (1996) *Upstream: Salmon and Society in the Pacific Northwest*. Washington, DC: National Academy Press.

Nehlsen W, Williams JE, and Lichatowich JA (1991) Pacific salmon at the crossroads: Stocks at risk from California, Oregon, Idaho, and Washington. *Fisheries* 16: 4–21.

Quinn TP (1982) A model for salmon navigation on the high seas. In: Brannon EL and Salo EO (eds.) *Proceedings of the Salmon and Trout Migratory Behavior Symposium*, pp. 229–237. Seattle, WA: School of Fisheries, University of Washington.

Ruggerone GT, Peterman RM, Dorner B, and Myers KW (2010) Magnitude and trends in abundance of hatchery and wild pink salmon, chum salmon, and sockeye salmon in the North Pacific Ocean. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science* 2: 306–328.

Schindler DE, Augerot X, Fleishman E, *et al*. (2008) Climate change, ecosystem impacts, and management of Pacific salmon. *Fisheries* 33: 502–506.

Shearer WM (1992) *The Atlantic Salmon*. Oxford, UK: Blackwell Scientific Publications, Fishing News Books.

Trotter PC (1989) Coastal cutthroat trout: A life history compendium. *Transactions of the American Fisheries Society* 118: 463–473.

Waples R (1995) Evolutionary significant units and the conservation of biological diversity under the Endangered Species Act. *American Fisheries Society Symposium* 17: 8–27.

Welch DW, Ishida Y, and Nagasawa K (1998) Thermal limits and ocean migrations of sockeye salmon (*Oncorhynchus nerka*) Long-term consequences of global warming. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 937–948.

Williamson KS, Murdoch AR, Pearsons TN, Ward EJ, and Ford MJ (2010) Factors influencing the relative fitness of hatchery and wild spring Chinook salmon (*Oncorhynchus twhawytscha*) in the Wenatchee River, Washington, USA. *Canadian Journal of Fisheries and Aquatic Sciences* 67: 1840–1851.

Scale, Concept and Effects of

David Clayton Schneider, Memorial University of Newfoundland, Newfoundland and Labrador, Canada

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 245–254, © 2001, Elsevier Inc.

Glossary

Biodiversity Includes taxonomic diversity along with habitat and genetic diversity.
Dimension Set of measurement units that are completely similar. Commonly used dimensions are length (inches, meters, etc.), time (seconds, years, etc.), and mass (grams, pound-mass, etc.).
Diversity distribution A frequency distribution showing the number of classes (usually species) that have one, two, three, etc. individuals per class (per species).
Diversity index A single number used to characterize a diversity distribution. The most common is s, the number of classes (species) into which a collection has been sorted.
Fractal Length dimension with exponent other than L^1 (length), or L^2 (area), or L^3 (volume).
Scale Unit measure (resolution, inner scale) relative to largest multiple (range, outer scale).
Scope Ratio of largest multiple to unit measure.

Biodiversity as a Measurable Quantity

Scalable Quantities

Biodiversity, no matter how it is measured, is a quantity with units on one of several types of measurement scale. Consequently, some care is needed in defining the measure of diversity at hand to obtain accurate scaling relations with area or time.

A well-defined measure of biodiversity has five parts: (i) a name, (ii) a symbol, (iii) a procedural statement that sets forth the method and conditions for measurement, (iv) a set of numbers generated by that procedural statement, and (v) units on one of several types of measurement scales. A scaled quantity is conveniently represented as a symbol equated to a set of numbers arranged inside brackets, multiplied by the unit of measurement. An example is the number of cichlid fish species found in each of six African lakes, as reported by Ricklefs and Schluter (1993):

Procedural statement	*Name*	*Symbol*	*Numbers*	*× Units*
Ricklefs and Schluter (1993), p. 358	Species richness	$S_j =$	$\begin{bmatrix} 200 \\ 136 \\ 200 \\ 7 \\ 9 \\ 40 \end{bmatrix}$	$\times$# lake^{-1}

In this example the lakes are listed in order from large (Lake Victoria = 69,484 km^2) to small (Lake Edward = 2150 km^2).

Measurements occur on one of one types of scale: nominal (yes/no), ordinal (ranks), interval (arbitrary zero point such as longitude in degrees), and ratio (true zero such as distance from the equator in kilometers). The mathematical rules for working with a ratio scale quantity (such as number of species per island) differ from those for working with quantities on other types of scales (such as ranking of lake area from large to small in the previous example). If our interest is in scale and its effects, then we need to work with the species-abundance distribution (on a ratio type of scale) rather than the rank-abundance distribution (on a rank type of scale).

Taxonomic Diversity Distributions

Diversity studies typically employ three basic quantities: N is the number of organisms, A is the area, and T is time or duration of measurement. A collection of N organisms from a unit area A_0 during a unit period T_0 is sorted into s groups (typically species) as shown in **Figure 1**. Any taxonomic level can be used in sorting, but for the sake of clarity species will be used to illustrate taxonomic diversity distributions. The number of organisms n_i is recorded for each species. A familiar characterization of this taxonomic diversity is the curve of abundance across species, ranked from high (largest n_i) to low (smallest n_i, usually one organism per species). The information in the rank-abundance curve can be re-expressed (see **Figure 1**) as a frequency distribution $s(n_i = n_k)$. The symbol $s(n_i = n_k)$ represents the number of species at abundances ranging from $n_k = 1$ upward to the abundance n_k of the most common species in the collection. This frequency distribution shows, in full detail, the information on species diversity in a single collection of size $\Sigma n_i = N$, taken from unit area A_0 in unit time T_0. If our interest is in the effects of scaling, then we will need to work with the frequency distribution $s(n_i = n_k)$, which is on a ratio type of scale, rather than with the rank abundance curve (**Figure 1**), which is on an ordinal type of scale.

Following practice in the literature on diversity, we can reduce the information in the species abundance distribution $s(n_i = n_k)$ to an index with a single value. The most common index is the total number of species s, obtained by counting classes or equivalently by summing the distribution $s(n_i = n_k)$ across abundance classes n_k. We might also compute Shannon

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00124-6

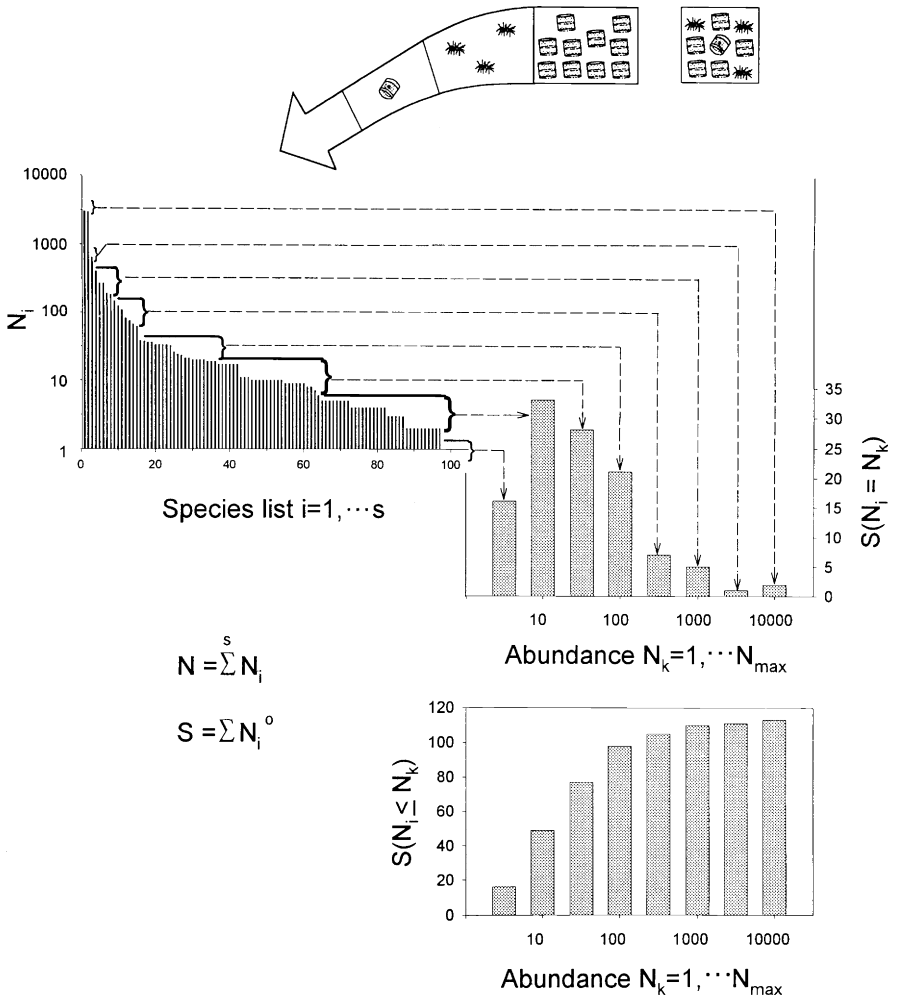


Figure 1 Collection of organisms sorted into a rank-abundance curve. Arrows show how the taxonomic diversity distribution $S(N_i = N_k)$ is constructed from the rank-abundance curve. The cumulative diversity distribution $S(N_i \leq N_k)$ is then constructed from the diversity distribution. N is the number of organisms in the collection, and S is the number of species.

Weaver diversity or the Simpson index for the collection. Both go beyond gross measures of diversity such as species number s, but fall short of describing all the information in the species diversity distribution $s(n_i = n_k)$ for the collection N.

If we obtain a second collection, again from an area A_0 during a period T_0, we expect a somewhat different number of organisms $n_j = 2$, a new set of species $s_{j=2}$ (many the same as in the first collection), and a new diversity distribution $s_2(n_i = n_k)$. With several collections $s_j(n)$, now labeled $j = 1$ through jt, we have a fuller characterization of diversity. We can construct the combined distribution $S(n_i = n_k)$ across some number of collections, jt. The total number of species S across the combined collection will exceed the number of species s_j in any one collection.

If the collections were always drawn by the same procedural statement, they can be viewed as samples from a set of all possible collections having unit area A_0 and duration T_0. The diversity distribution for this hypothetical set is $S(n_i = n_k)$. This is the species abundance distribution for the community, provided all collection sites have an equal (or at least known) chance of appearing in the sample. The total number of species S and total number of organisms N are computed respectively as $S = \Sigma s(n_i = n_k)$, $N_{tot} = \Sigma n_i s(n_i)$. Integral signs ($\int$) replace the summation signs for distributions based on continuous values of population size n_k (May, 1975). Following May (1975) we can compare an observed distribution $s(n_i = n_k)$ to a mathematical function, allowing us to compute species numbers at unobserved values of abundance n_k. An impressive number of methods (Bunge and Fitzpatrick, 1993) have been devised to estimate S from several collections $s_j(n_i = n_k)$, taken to be samples.

When making statistical estimates from taxonomic distributions, it often proves convenient to use the cumulative distribution $s(n_i \leq n_k)$ rather than the distribution $s(n_i = n_k)$. The cumulative distribution rises monotonically from left to right, as it records the number of species with n_k or fewer organisms per species, rather than recording the number of species at each value of n_k. Both distributions express exactly the same information. The cumulative distribution is less lumpy and therefore easier to use.

Other Diversity Distributions

Genetic variability within a species can be characterized in the same fashion as taxonomic diversity. For a collection of N individuals the usual single-locus measure is the proportional presence of each type of allele $q_i = n_i/2N$. This information can be expressed as a rank-abundance distribution of allelic proportions from common (q_i large) to rare (q_i small). For convenience, the information in this distribution is reduced to an index such as the number of classes (alleles) or expected heterozygosity ($H_e = 1 - \Sigma q_i^2$). The information in several collections can be expressed as a rank-abundance curve for each collection and a rank-abundance curve for the combined collection. This information can in turn be reduced to an index, such as F_{st}, which compares expected heterozygosity for the entire collection H_e to the average over k collections ($\overline{H}_e = k^{-1}\Sigma H_e$). This approach, applied to genetic diversity at the single-locus level, can be extended to coarser classifications, such as recognizable genotypes in a population (Mallet as cited in Gaston, 1996, p. 16). If we are interested in the effects of scale on genetic diversity, then the logic of working with ratio scale quantities will compel us to take the (highly unusual) step of re-expressing the rank-abundance curve as the genetic diversity distribution $q(n_i = n_k)$, which is on a ratio type of scale. This is the frequency of genotypic classes for which there was one individual ($n_k = 1$), two individuals ($n_k = 2$), etc. This can be converted to a cumulative frequency distribution $q(n_i \leq n_k)$ to again smooth away the inconvenient lumpiness of the distribution $q(n_i = n_k)$.

We can construct a diversity distribution to summarize habitat diversity. This requires defining a collection area A divided into N units of size A_0, which are sorted into habitat classes. From this we can construct the usual rank-abundance curve, running from the most common habitat (n_i large) to rarest (n_i small). The same information can be plotted on a ratio scale as a habitat diversity distribution $h(n_i = n_k)$, expressing the frequency of habitats found in one unit ($n_k = 1$), in two units ($n_k = 2$), etc.

Alpha, Beta, and Gamma Diversity

In 1960, R. H. Whittaker distinguished taxonomic diversity within a habitat (alpha diversity) from diversity within larger geographic units composed of several habitats (gamma diversity). Examples of these larger units include islands, lakes, landscapes, and coastal embayments. Whittaker defined beta diversity as the among-habitat diversity, or diversity change going from a single habitat to the larger geographic unit. Later (**Box 1**) Whittaker added two more spatial scales or "levels"—that of point diversity within a habitat and that of regional (epsilon) diversity for broader geographic units containing smaller units, each characterized by its own gamma diversity. Whittaker presented these as convenient labels, recognizing that point, alpha, gamma, and epsilon diversity were arbitrary points on a continuum, and that the size of these units might be defined differently in any given study. Starting with Whittaker, the most effective use of these labels has been in distinguishing diversity patterns at the scale of habitats from patterns at larger scale. The labels are also useful in discussing the very different processes responsible for changes in diversity at habitat, regional, and global scales. Alpha, beta, and gamma diversity have remained in wide use during the past 40 years; point and epsilon diversity did not come in to wide use. In practice, gamma diversity is sometimes dropped in favor of two labels—within-habitat (alpha) and among-habitat (beta) diversity.

It is of note that similar sets of arbitrary labels to distinguish pattern and process over a range of spatial scales have been developed in both marine (Haury *et al.*, 1978; **Box 2**) and terrestrial ecology (Delcourt and Delcourt, 1988; **Box 2**). Habitat features persist longer in terrestrial than marine systems, and hence the time-scales associated with any given spatial scale are greater for terrestrial than marine systems (Steele, 1991; **Box 2**).

Concept of Scale

Multiscale Analysis

Multiscale questions naturally arise for biodiversity: How does diversity change with collection size? How does it change with area? How does it change with duration of collection? These questions can be addressed quantitatively by examining the change in a biotic diversity distribution with change in collection size N, in collection area A, and in collection duration T. For habitat diversity similar questions of scale arise: How does the habitat diversity distribution $h(n_i = n_k)$ change when the collection area A is doubled? As with taxonomic diversity,

Box 1 Alpha, beta, and gamma diversity

In an evolutionary context, Whittaker (1977) defined inventory diversity at four spatial scales or levels: the point sample, the habitat, the landscape, and the region. Differentiation diversity is defined as the change in diversity between these four levels.

Point diversity The number of species for a small or microhabitat sample within a community regarded as homogeneous. Also called internal alpha or subsample diversity.

Pattern diversity Change going from one point to another within a habitat. Also called internal beta diversity.

Within-habitat or alpha (α) diversity The number of species in a sample representing a community regarded as homogeneous (despite its internal pattern). It measures the number of potentially interacting species.

Between-habitat or beta (β) diversity Change along an environmental gradient or among the different communities of a landscape. Developed by Whittaker (1960) as a measure of packing of competing species along a gradient, it was defined as the ratio of regional (gamma) diversity to average within-habitat (alpha) diversity. Wilson and Shmida (1984) review other measures of beta diversity.

Landscape or gamma (γ) diversity The number of species in a set of samples including more than one kind of community.

Geographic differentiation or delta (δ) diversity Change along a climatic gradient or between geographic areas.

Regional or epsilon (ε) diversity The number of species in a broad geographic area including differing landscapes.

Box 2

	Evolutionary biology	Terrestrial ecology		Marine ecology	
	Whittaker (1977, Figure 14)	Delcourt and Delcourt (1988)		Haury *et al.* (1978)	Steele (1991)
	Area	Area	Time	Area	Time
10^{15}				Mega	
		Mega	10^6+ years	Macro	
10^{12}				Meso	
		Macro	10^4–10^6 years		1–10 years
10^9	ε				
		Meso	500–10,000 years	Coarse	Days–years
	δ				
10^6	γ				
		Micro	1–500 years	Fine	1–10 days
10^3	β				
	α				
1 m^2	pattern				
	point			Micro	
10^{-3}					

we expect genetic and habitat diversity to be greater in large collections than in small collections, greater in widely separated samples than in closely placed samples, and greater in large areas than in small areas. We do not expect that a doubling in area will simply double the number of habitats or the number of alleles.

There are several ways to compare diversity at large and small scales. One of the more venerable is to rate (or plot) diversity relative to island size (**Figure 2**). Rating diversity against area is equivalent to rating against collection size N if organism density is the same on large and small islands. Another venerable method is to make a series of collections and then plot a measure of diversity against effort, measured as the accumulated number of collections (**Figure 2**). These "collector's curves" are equivalent to plots against accumulated collection size N if organism catch per unit effort remains constant across collections. **Figure 3** shows the increase in diatom species number with increase in effort (number of slides). A third method, currently popular but less venerable, is to fix the unit size, measure the separation between units (in time or space), and then plot some measure of difference in diversity between units as a function of separation or lag (**Figure 2**). An additional method is to construct a diversity distribution from a large area and then take the average diversity distribution for each half of the area, each quarter, etc. This method of coarse graining (**Figure 2**) is better known in geophysics (Rodriguez-Iturbe and Rinaldo, 1997; Barenblatt, 1996) than in biology. Coarse graining expresses the same information as lagging, an equivalence that goes unrecognized. Both lagging and coarse graining assume that, on average, organism number N does not depend on distance separating samples, and hence diversity differences are due to spatial scale (separation) rather than spatial gradients in organism number N.

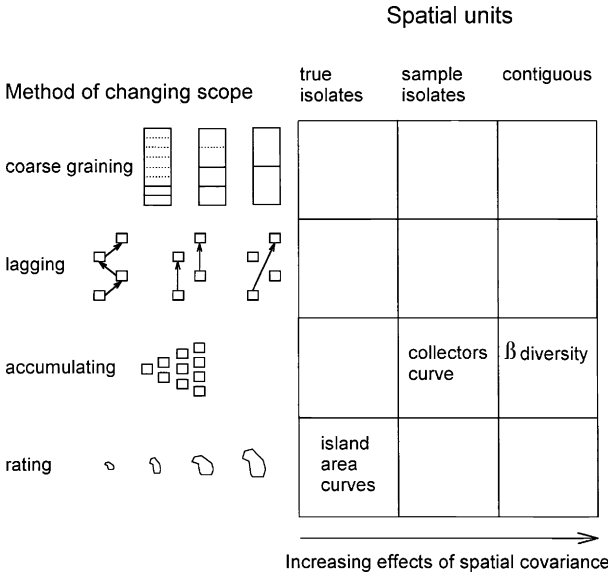


Figure 2 Scaling methods change the spatial scope by one of several different methods, explained in the text. For biodiversity, the traditional applications are collector's curves (increasing the number of samples), island area curves (islands rated from small to large), and beta (β) diversity (increase in number with the number of contiguous samples along a transect).

These various methods of multiscale comparison express the same information in principle but not in practice. Lagging and coarse graining (in space) express the same information if the spatial layout is the same. In practice, somewhat different information is obtained because coarse graining is applied to contiguous units, whereas lagging is applied to separated

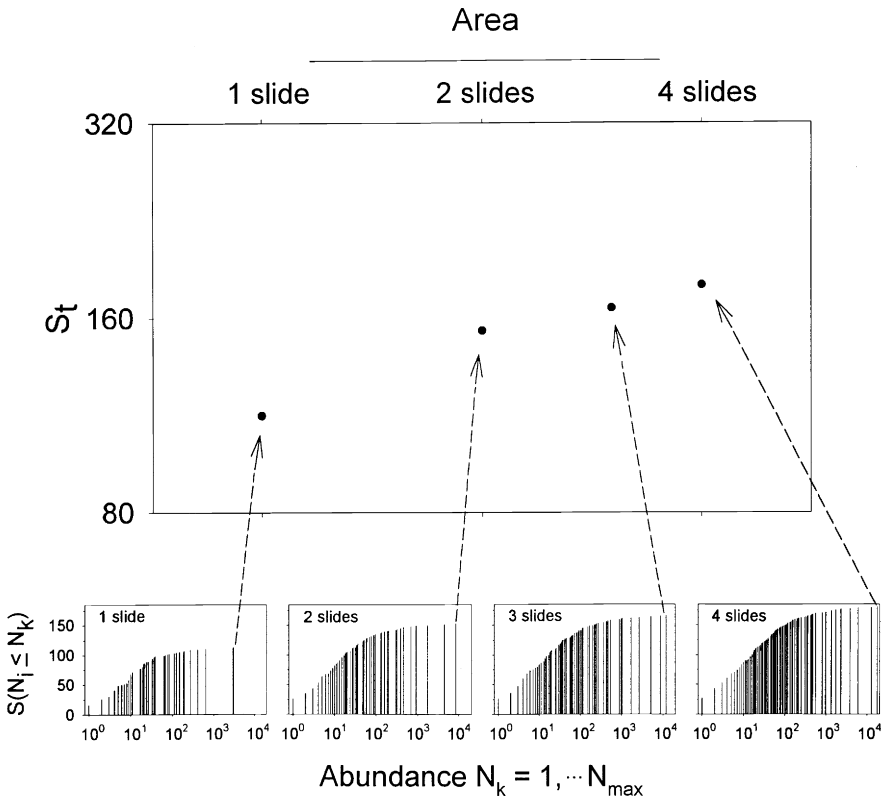


Figure 3 Cumulative diversity distributions from areas of increasing size (one, two, three, and four slides). Note the linear relation (on a logarithmic scale) between species number S and slide area. However, species number S represents only part of the information in the cumulative diversity distributions from areas of increasing size. Data from Patrick R (1968) The structure of diatom communities in similar ecological conditions, *Am. Nat.* 102, 173–183.

units. Collector's curves introduce time-scales. In a seasonal environment, collections made monthly for a year are expected to grow more rapidly in species number than collections made annually for a decade. Rating relative to area introduces another factor—the dynamics of population exchange within species. A plot or rating of diversity against area will differ in scope for evolutionary isolates (continents), ecological isolates (islands), partial isolates (islands in an archipelago), and quadrat samples from a larger area (Rosenzweig, 1995).

Scope

In comparing diversity within or among studies we would like to have a measure that does not depend on choice of units. To accomplish this, the scope of a quantity is defined as the ratio of its maximum to minimum value. For the cichlid example, the scope in species number is (200 species lake^{-1})/(7 species lake^{-1}) = 29. For non-cichlid species the scope in number per lake is different: (111 species lake^{-1})/(17 species lake^{-1}) = 6. The scope of a quantity has units that cancel and hence are expressed as dimensionless numbers, independent of units of mass, length, or time. However, the units that form the dimensionless ratio should be kept in view to avoid drawing conclusions outside the scope of measurement.

Care is required in comparing quantities with scopes that differ procedurally. One can increase the scope of collection size by increasing area from unit area A_0 to total area A_{tot}. The scope is $\Sigma A/A_0$. This generates accumulation curves relative to area: accumulation of organisms $N(A)$, of species $S(A)$, of genotypes $Q(A)$, and of habitats $H(A)$. One can increase the scope of collection size by repetitive collection at one location, resulting in accumulation curves relative to time: $N(T)$, $S(T)$, $Q(T)$, and $H(T)$. There is no reason to expect that change in diversity relative to change in spatial scope will match change in diversity relative to change in temporal scope.

Spatial scope can be expanded in several ways. Searching out a wide range of isolated units—islands in an ocean and lakes on a landscape—will increase the scope. With this method the scope is the ratio of the largest to smallest spatial unit, $A_{max}A_{min}^{-1}$. For the lake example, the spatial scope is (69,484 km^2 lake^{-1})/(2150 km^2 lake^{-1}) = 32. Accruing units will increase spatial scope if these are distributed over an area. Accrual can be by adjacency (along a transect or concentric rings around a point) or by random sampling within a larger area (in statistical jargon, within the frame of all possible units). With accrual, the scope is $A_0^{-1}\ \Sigma A$ = number of units. Lagging can apply to contiguous units; the scope is $(L_{max} - L_{min})/L_{min}$, where L_{max} is the distance across the block or along the transect, and L_{min} is the distance across the unit block. Fine graining increases scope via an increase in the

frequency of measurement. With fine graining, the scope is A_{tot}/A_{min}, where A_{min} is the smallest resolvable area within a large block A_{tot}.

Temporal scope can be varied as well. In principle, one can compare short time periods with long periods, but there are no examples of this. Temporal scope increases as units of equal duration T_0 accrue; the scope is the ratio of the time series to the minimum unit: $T_0^{-1}\ \Sigma T =$ number of units. Temporal scope also increases by fine graining down to the smallest time units T_0 within a continuous period T. The scope is T/T_0.

Complete and Incomplete Similarity

The classical solution to the problem of change in scale has been to define two systems (such as a model boat and full-scale prototype), define groups of similar measurement units (typically mass, length, and time), and then work out the scaling from model to prototype with reference to groups of similar units called dimensions mass $[M]$, time $[T]$, Euclidean lengths $[L]$, areas $[L^2]$, and volumes $[L^3]$. If the model and prototype are completely similar (Barenblatt, 1996) then we expect that a 10-fold increase in length of prototype relative to model will result in a 10^3-fold increase in volume.

$$\frac{\text{Volume}_{\text{prototype}}}{\text{Volume}_{\text{model}}} = \left(\frac{\text{Length}_{\text{prototype}}}{\text{Length}_{\text{model}}}\right)^3 \tag{1}$$

We also expect a 1000-fold increase in mass if model and prototype are constructed of the same material and have the same density. The scaling relation is mass $\cong$ volume $\cong$ length^3. Dimensional analysis, based on the principle of complete similarity, has been notably successful in engineering and several areas of physics, including geophysical fluid dynamics. It has been notably unsuccessful with natural rather than engineered objects, or with dynamics that "lurch" rather than ticking along in a clock-like fashion.

In ecology, the failure of complete similarity was addressed by comparative methods of "hierarchy theory," which recognizes that large-scale spatial structure or dynamics cannot be computed directly from local scale structure or rates. Using Herbert Simon's concept of "nearly decomposable levels," the hierarchical method describes structure or computes dynamics at two (or more) levels and then compares the results. For example, in the biodiversity literature it is customary to distinguish local or within-habitat diversity from regional (among-habitat) diversity. One widely recognized finding (Holt as cited in Ricklefs and Schluter, 1993, p. 77) is that the relation of species number to area at the habitat level (i.e., within a habitat) does not match that at the ecosystem level (i.e., among larger areas having multiple habitats).

When completely similar scaling fails, a property (e.g., coastline length) at a large scale can often be computed from the average property at a smaller scale by a power law expressing incomplete similarity (Barenblatt, 1996). The most familiar example of incomplete similarity is that of a fractal relation between coastline length and Euclidean length L. If we take large steps along a coastline and then count these as if on a straight line, we obtain a coastline perimeter of length $P(L_{large})$. If we take small steps and then count these as if in a line, we obtain a length $P(L_{small})$ that exceeds $P(L_{large})$. The relation of the two measures of coastline length is

$$\frac{P(L_{large})}{P(L_{small})} = \left(\frac{L_{large}}{L_{small}}\right)^{\alpha} \tag{2}$$

In briefer form, $P(L) = c \times L^{\alpha}$. In even briefer form, $P(L) \cong L^{\alpha}$. The exponent α is not an integer, as it would be for completely similar Euclidean objects. Mandelbrot (1977) coined the term "fractal" to describe non-Euclidean objects where α is a non-integral constant.

More generally, incomplete similarity relates some property $S(x)$ to some measure x according to a power law:

$$S(x) = cx^{\alpha} \tag{3}$$

This relation states that when x is rescaled (e.g., by a factor of 2), the property $S(x)$ is still proportional to x^{α}. In the example of African lakes, how does species number scale with area? A rough scaling is obtained by applying the principle of homogeneity of scope (Schneider, 1994), which requires that all terms in an equation have the same scope. An exponent is applied to equate the scope of diversity with that of lake area:

$$\frac{s(A_{big})}{s(A_{small})} = \left(\frac{A_{big}}{A_{small}}\right)^{\alpha} \tag{4a}$$

$$\frac{200}{40} = \left(\frac{69,484}{2150}\right)^{\alpha} \tag{4b}$$

Solving for the exponent, we obtain the scaling for incomplete similarity

$$\alpha = \log(200/40)/\log(69,484/2150) = 0.463.$$

This is a rough estimate that could be improved by using regression to include more information in the scaling. Based on the rough scaling, a doubling in lake area from, for example, 2150 km^2 to 4300 km^2 is expected to amplify cichlid species number by $2^{0.463} = 1.4$ rather than by a factor of 2.

Incomplete similarity applies to cumulative frequency distributions (**Figure 3**) as well as to single indices such as species numbers. The scaling based on incomplete similarity is

$$P[X \geq x] = c \cdot x^{\alpha} \tag{5}$$

where x is again some measure of interest, such as length, area, or time. For some property (such as species per unit area) the probability distribution relative to the measure is

$$S(x) = c \cdot x^{-(\alpha+1)} \tag{6}$$

The exponent that quantifies incomplete similarity can be estimated by selecting isolated units, as in the cichlid example. The exponent can be estimated via autocorrelation techniques, which measure correlation of a quantity with itself at increasingly large spatial separations. Scaling exponents can be estimated by fine graining (Milne, 1997) or equivalently by coarse graining (**Figure 2**). Estimates via fine or coarse graining are equivalent to those from lagging (**Figure 2**), a fact that is well recognized for time series analysis but largely overlooked in the literature on spatial structure. Fine graining (or equivalently, coarse graining) has been used to quantify the

fractal scaling exponent for a variety of landscape features (Rodriguez-Iturbe and Rinaldo, 1997). Fine graining (under the name of box counting or the dividers method) has been used to quantify the scaling exponent for habitats as diverse as bare soil, leaf perimeter, seaweed, and coral heads; it has been used to quantify coastline nesting habitat of eagles, pine nesting habitat of woodpeckers, and foraging habitat of rabbits (Milne, 1997).

The dynamics of systems that are incompletely similar at small and large scales are also examined via renormalization and coarse graining (Barenblatt, 1996). Renormalization has been used to investigate the dynamics of landscapes (Rodriguez-Iturbe and Rinaldo, 1997) and population interaction (Levin and Pacala as cited in Tilman and Kareiva, 1997, p. 271). Renormalization is an appropriate technique for examining the dynamics of biodiversity, which do not proceed by evenly paced transitions, but instead jump and lurch by episodes of invasion, local extinction, speciation, and anthropogenic extinction.

Incomplete similarity provides an alternative to hierarchical (usually two-level) analysis of structure and dynamics in ecology. Currently, the hierarchical approach is more familiar and remains the dominant approach when analyzing ecological interactions (Tilman and Kareiva, 1997) or biodiversity patterns at multiple scales (Ricklefs and Schluter, 1993; Rosenzweig, 1995).

Scaling Relations for Biodiversity

The development of scaling relations for biodiversity has been vigorous for taxonomic diversity. It has just begun for habitat diversity and does not yet exist for genetic diversity. For taxonomic diversity, the history of theoretical development can be summarized briefly as a sequence of scaling relations.

$S \cong A^{\beta}$

Species number scales incompletely with area. This is the oldest and best known scaling relation in ecology, with a rich history (Ricklefs and Schluter, 1993; Rosenzweig, 1995). The scaling coefficient was recognized to be less than 1 for most situations, but values were completely empirical and could not be generalized beyond the data used to estimate the coefficient.

$S \cong N^{1/4}$

Species number scales incompletely with numbers according to a 1/4 power law. This relation, due to Preston, was discussed in detail by May (1975), who provided a list of models. It is worth noting that the canonical value (1/4) depends on the assumption that the taxonomic diversity distribution has a lognormal form.

$S \cong N^{1/4} \cong A^{1/4}$

Species number scales with total numbers and hence with area according to a 1/4 power law. Preston's canonical theory and MacArthur and Wilson's theory for isolated communities attribute the scaling to equilibrium conditions (May, 1975), but it is now known that scaling laws such as this can arise in non-equilibrium conditions (Barenblatt, 1996). This scaling assumes that organism number N scales in a 1:1 fashion with area.

$N \cong A^{1}$

Numbers scale completely with area. This scaling will always hold for coarse graining—we can always compute the average numbers in smaller areas from numbers in larger areas. This scaling will hold on average for accumulation if we take multiple sequences of organism counts. However, this scaling is usually untrue for any one sequence of accruals because of patchiness. Often, the rate of accrual will appear to accelerate because often the densest site will be sampled late rather than early during accrual. Complete scaling of numbers N with area A^1 is easier to assume than demonstrate for islands or other isolated units.

$S \cong A^{\beta}$, where $\beta_{\text{island}} > \beta_{\text{block}}$

The scaling exponent is steeper for islands ($0.25 < \beta_{\text{island}} < 0.33$) than for isolated blocks of land on continents ($0.13 < \beta_{\text{block}} < 0.18$). Rescue effects, where small populations are maintained by frequent migration from surrounding areas on continents, readily explain the shallow scaling for continental communities (Rosenzweig, 1995). Milne (1997) describes a correction factor to account for scaling effects in comparing the area of Euclidean blocks to the area of islands with fractal shapes.

$S \cong A^{\beta}$, where $\beta_{\text{province}} > \beta_{\text{block}}$

Biogeographic provinces, isolated at evolutionary time-scales (Rosenzweig, 1995), have scaling exponents that exceed blocks within a province. Estimates of β_{province} fall closer to unity (Rosenzweig, 1995) than to typical values for β_{block}.

$S(n_i \leq n_k) \cong S(A_i \leq A_k) \cong H(A_i \leq A_k)$

Species diversity, as a distribution, scales with habitat diversity, again as a distribution. Substantial qualitative support exists as correlations of species richness with several habitat variables (Wright *et al.* as cited in Ricklefs and Schluter, 1993, p. 73; Rosenzweig, 1995), including fractal measures of habitat complexity (Milne, 1997). An analytic review, similar to that of May (1975), is needed to integrate the taxonomic diversity literature with the literature on the structure and dynamics of landscapes (Rodriguez-Iturbe and Rinaldo, 1997).

Effects of Scale

On Estimates of Diversity

Estimates of diversity and of rate of extinction are required at both local and global scales, but measurements are restricted to small scales. The principal effect of incomplete similarity on estimates of diversity and extinction rate is that it precludes intuitive scalings or extrapolations. If species number, or any other index of taxonomic diversity, scales completely with area or with number of organisms, then scaling could be handled with intuition-based 1:1 scalings. A survey could be used to estimate diversity in a larger area by knowing only the area of interest relative to the area surveyed to make the extrapolation. Extinction rates could be estimated over large areas by

knowing only what percentage of the area had been examined directly. Complete similarity of taxonomic diversity with any other quantity has yet to appear in the literature, which is vast and growing (Gaston, 1996; Ricklefs and Schluter, 1993).

Incomplete similarity poses little problem if the scaling exponent is used to interpolate within the scope at which it was estimated. Extrapolation to larger scales poses a problem because of uncertainty in choice of exponents. Extrapolation becomes possible if the scaling exponent increases in a predictable way with increase in the degree of isolation (Rosenzweig, 1995). Choosing an incomplete scaling exponent, however, remains fraught with uncertainty.

Problems of non-intuitive (incomplete) scaling and uncertainty in choice of scaling exponent attend any effort to estimate change in taxonomic diversity. For habitat diversity, the problems are less acute because remote sensing can provide maps at high resolution over large areas. Consequently, scaling exponents can be estimated over a scope wide enough to be useful, from large areas of interest down to the small scale of plots and quadrats amenable to experiment and direct measurement of population and community dynamics. The number of variables that can be sensed remotely is limited. Hence, there is need for scaling relations that connect remotely sensed variables to population variables, such as density, biomass, production, movement, recruitment, and mortality. There is also a need to connect remotely sensed variables to community variables such as extinction rate.

On Conservation Planning

Conservation planning is carried out at far larger scales than measurements of the underlying dynamics. The problem is inescapable but not solvable by intuitive modes of thinking based on complete similarity. One cannot execute a field study, and apply it with certainty to conservation issues at larger scales, using the intuitive notion that rates (e.g., per capita mortality) are independent of spatial scale (i.e., scale as A^0). Nor can one proceed on the intuitive notion that a variable will scale completely with area on a 1:1 relation. The idea of incomplete similarity has already started to take its place in conservation planning. An example is the "single large vs several small" debate over refuge size—a debate that recognizes that doubling area does not leave rates unchanged nor does it double quantities such as the number of species. Several other examples are provided by Milne (1997). The next step will be to develop incomplete scalings that are reliable and based on biologically based theory rather than empirical coefficients. One particularly exciting topic, which may soon yield to theoretical development, is the scaling of taxonomic diversity to habitat diversity. The basis in biological theory already exists. Species lists increase by invasion or speciation, for which adaptation to habitat via natural selection is important. Species lists shrink by local or global extinction, for which change in habitat is important. There is thus the biological basis for quantitative scaling of taxonomic diversity with habitat diversity. It is a promising topic for theoretical development, for which reliable computations are needed to address urgent questions of local and planet scale change in biodiversity.

See also: Measurement and Analysis of Biodiversity

References

Barenblatt GI (1996) *Scaling, Self-Similarity, and Intermediate Asymptotics*. Cambridge, UK: Cambridge Univ. Press.

Bunge J and Fitzpatrick M (1993) Estimating the number of species: A review. *J Am. Stat. Assoc.* 88: 364–373.

Delcourt and Delcourt (1988) Quaternary landscape ecology. Relevant scales in space and time. *Landscape Ecol.* 2: 45–61.

Gaston KJ (ed.) (1996) *Biodiversity*. Oxford: Blackwell.

Haury, *et al.* (1978) Patterns and processes in the time-space scales of plankton distributions. In *Spatial Pattern in Plankton Communities*, pp. 227–327. New York: Plenum.

Mandelbrot B (1977) Fractals. Form, Chance, and Dimension. San Francisco: Freeman.

May RM (1975) Patterns of species abundance and diversity. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*, pp. 81–120. Cambridge, MA: Belknap.

Milne BT (1997) Applications of fractal geometry in wildlife biology. In: Bissonnette JA (ed.) *Wildlife and Landscape Ecology*, pp. 32–69. New York: Springer-Verlag.

Ricklefs RE and Schluter D (1993) *Species Diversity in Ecological Communities*. Chicago: Univ. of Chicago Press.

Rodriguez-Iturbe I and Rinaldo A (1997) *Fractal River Basins. Chance and Self-Organization*. Cambridge, UK: Cambridge Univ. Press.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge Univ. Press.

Schneider DC (1994) *Quantitative Ecology. Spatial and Temporal Scaling*. San Diego: Academic Press.

Steele (1991) Can ecological theory cross the land-sea boundary? *J. Theor. Biol.* 153: 426–436.

Tilman D and Kareiva P (1997) *Spatial Ecology*. Princeton, NJ: Princeton Univ. Press.

Whittaker (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.* 30: 279–338.

Whittaker (1977) Evolution of species diversity in land communities. *Evol. Biol.* 10: 1–68.

Wilson and Shmida (1984) Measuring beta diversity with presence-absence data. *J. Ecol.* 72: 1055–1064.

Seagrasses

Carlos M Duarte, IMEDEA, CSIC-Universitat de les Illes Balears, Spain

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 255–268, © 2001, Elsevier Inc.

Glossary

Hydrophily Underwater pollination developed by most seagrasses and many freshwater angiosperms.

Seagrasses Marine angiosperms, all monocots, able to grow submersed in marine waters, to which they are restricted.

Introduction

Seagrasses are angiosperms restricted to marine life. They encompass a limited (approximately 50; Table 1) set of closely related species which have successfully colonized the coastal areas across the coastal ocean from intertidal areas to depths in excess of 50 m in the clearest waters (Duarte, 1991), except for polar seas, from which they are absent (den Hartog, 1970). Seagrasses often develop vast meadows (**Figure 1**), which are important seascapes covering approximately 0.6×10^6 km^2 in the ocean (Duarte and Cebrián, 1996) and are responsible for a primary production of approximately 0.60 Gt C year^{-1} (Duarte and Chiscano, 1999) and 15% of the total net CO_2 uptake by oceanic biota (Duarte and Chiscano, 1999). In addition, seagrasses provide habitat for animals, many of which are economically important or endangered, and also play an important role by estabilizing sediments in coastal areas (Hemminga and Duarte, 2000). Hence, the elucidation of their diversity, its maintenance, and its influence on the functions of seagrasses in the ecosystem are important goals for the sustainable management of these key ecosystems. This article provides a comprehensive view of seagrass diversity by examining the origin and current distribution of seagrasses, the controls on seagrass species and genetic diversity and their functions in the ecosystem, and threats to the sustainability of seagrass biodiversity.

The Origin and Evolution of Seagrasses

Seagrass originated early in the evolution of angiosperms, approximately 100 million years ago, from ancestors that have been hypothesized to be freshwater plants or mangrove- or salt-marsh-like plants. The origin of seagrasses appears to be polyphyletic, and recent analyses based on molecular phylogeny have supported the development of seagrass species from three clades of primitive plants (Les *et al.*, 1997). There are two main hypotheses on the origin of seagrasses. One identifies marsh- or mangrove-like plants as their ancestors (den Hartog 1970), which is supported by the observations that some seagrass species present lignified rhizomes and that two genera present viviparous reproduction, similar to some mangrove taxa. The second hypothesis identifies, on the basis of molecular phylogeny, freshwater plants to be the seagrass ancestors (Les *et al.*, 1997). A freshwater ancestor appears clearer for the genus *Enhalus* (*Enhalus acoroides*), the only seagrass species without submarine pollination, which has been associated with related freshwater plants (*Vallisneria* sp.), that it resembles in form and pollination mode.

The fossil record of seagrass is very scarce, and has not been very helpful in testing hypotheses on the origin or evolution of seagrasses. The oldest specimens in the record consist of, at most, three genera: *Archeozostera*, believed to be the ancestor to the extant genera *Heterozostera*, *Zostera*, and *Phyllospadix*; *Thalassocharis*, ancestor to the extant genera *Amphibolis*, *Thalassodendron*, *Halodule*, *Cymodocea*, and *Syringodium*; and *Posidonia*, the only extant genera among the three from the Cretaceous.

Evolution of Seagrasses

Although the path of seagrass evolution is not very clear due to the small size of the fossil record, there is evidence that most modern seagrass genera were already established in the Eocene, but the genera *Phyllospadix* and *Enhalus* appear to have evolved in the Palaeocene. The evolution of seagrasses has been postulated to be linked to the development of four major capacities essential for life in the marine environment: (i) the capacity to grow underwater, (ii) the development of mechanisms to tolerate the high salinity of the sea, (iii) the development of substantial anchoring capacity so as to enable seagrasses to remain attached to the substrate in the high-energy marine environment, and (iv) the capacity for underwater pollination (hydrophily). The net result of the action of these constraints is that seagrasses represent a rather homogenous plant group: All seagrasses are rhizomatous plants, with basal meristems, generally blade-like leaves, and mostly hydrophilous pollination. These adaptations may have all been present in some of the freshwater plants, particularly those resistant to high salinity, that may have acted as ancestors to the seagrasses. In contrast, the hypothesis of possible marsh or mangrove ancestors would have required the evolution of most of these adaptations. In particular, the capacity to grow underwater relies critically on the development of an extensive lacunar system able to transport gases between the different organs of the plants, thereby supplying photosynthetic oxygen to the roots and, possibly, CO_2 derived from respiration to enhance leaf photosynthesis. This lacunar system may have already been present in possible freshwater ancestors, such as that for *E. acoroides*. Hydrophily, which is

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00189-1

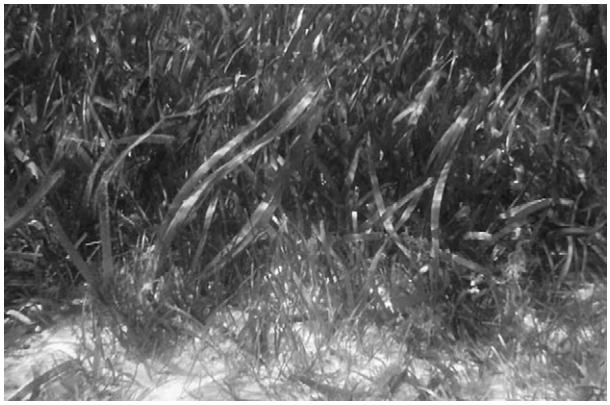


Figure 1 A mixed (seven species) seagrass meadow in Bolinao (Pangasinan, the Philippines).

believed to have played an important role in seagrass evolution, affecting their speciation, was already present in possible freshwater ancestors but must have been evolved independently under the hypothesis of mangrove- or marsh-like ancestors. Although its importance is undisputed, the capacity for underwater pollination is not an essential adaptation to life in the marine environment. For example, *Enhalus*, the most recent seagrass genus, pollinates over the water surface as did its ancestor freshwater plants, and yet it is clearly a successful plant in the Indo-Pacific region. A freshwater ancestor for seagrasses is further coherent with new evidence suggesting that the first angiosperms may have also originated in freshwater environments.

The Seagrass Flora: Species Richness

Seagrass Taxonomy: Status and Uncertainties

The seagrass flora is composed of approximately 50 species, distributed in 12 genera, belonging to only two families from two closely related orders (**Table 1**). Despite the small number of species, the taxonomy of seagrasses is far from settled, and there is considerable uncertainty not only at the level of species but also at the level of genera. Indeed, the total number of species described could, if uncritically accepted, exceed 60. Although species descriptions are often based on reproductive structures, these characters are, in practice, not always useful as diagnostic criteria to separate them. In practice, the identification of species in field studies is often dependent on vegetative characters to discriminate between genera, whereas the discrimination of species within genera is often based on the biogeographical segregation between species because congeneric species are often present in different floras. The diagnostic characters most commonly used to classify seagrasses into genera and species include differences in the venation, the presence of tannin cells, denticulation, and the tip shape of leaves (den Hartog, 1970). Although differences between reproductive structures are likely to be robust indicators of the broad genetic differences expected from species, differences in leaf anatomy are not as reliable. Leaf characteristics are remarkably plastic within seagrass species, even within adjacent shoots, rendering their diagnostic value questionable. Classification systems based on differences in leaf characteristics are therefore conducive to an ever-increasing number of species, particularly in the small and very plastic genus *Halophila*.

Table 1 Seagrass genera and accepted species indicating their membership to the different seagrass floras identified in **Figure 2**

Species	*Biogeographic membership*
Amphibolis antarctica	S. Australian flora
Amphibolis griffithii	S. Australian flora
Cymodocea angustata	Indo-Pacific flora
Cymodocea nodosa	Mediterranean flora
Cymodocea rotundata	Indo-Pacific flora
Cymodocea serrulata	Indo-Pacific flora
Enhalus acoroides	Indo-Pacific flora
Halodule pinnifolia	Indo-Pacific flora
Halodule uninervis	Indo-Pacific flora
Halodule wrightii	Caribbean flora
Halophila baillonis	Caribbean flora
Halophila beccarii	Indo-Pacific flora
Halophila capricornii	Indo-Pacific flora
Halophila decipiens	Caribbean and Indo-Pacific floras
Halophila engelmannii	Caribbean flora
Halophila minor	Indo-Pacific flora
Halophila ovalis	Indo-Pacific flora
Halophila spinulosa	Indo-Pacific flora
Halophila stipulacea	Indo-Pacific flora
Heterozostera tasmanica	S. Australian flora
Phyllospadix iwatensis	Temperate W. Pacific flora
Phyllospadix japonicus	Temperate W. Pacific flora
Phyllospadix scouleri	Temperate E. Pacific flora
Phyllospadix serrulatus	Temperate E. Pacific flora
Phyllospadix torreyi	Temperate E. Pacific flora
Posidonia australis	S. Australian flora
Posidonia oceanica	Mediterranean flora
Posidonia ostenfeldii	S. Australian flora
Posidonia sinuosa	S. Australian flora
Posidonia angustifolia	S. Australian flora
Posidonia coriacea	S. Australian flora
Posidonia denhartogii	S. Australian flora
Posidonia kirkmanii	S. Australian flora
Posidonia robertsoniae	S. Australian flora
Syringodium filiforme	Caribbean flora
Syringodium isoetifolium	Indo-Pacific flora
Thalassia hemprichii	Indo-Pacific flora
Thalassia testudinum	Caribbean flora
Thalassodendron ciliatum	Indo-Pacific flora
Thalassodendron pachyrhizum	S. Australian flora
Zostera asiatica	Temperate W. Pacific flora
Zostera capensis	S. Atlantic flora
Zostera capricorni	S. Australian flora
Zostera caulescens	Temperate W. Pacific flora
Zostera japonica	Temperate W. Pacific flora
Zostera marina	N. Atlantic, Mediterranean, W. and E. Pacific floras
Zostera mucronata	S. Australian flora
Zostera mulleri	S. Australian flora
Zostera noltii	N. Atlantic and Mediterranean floras
Zostera novazelandica	New Zealand flora

Source: Reproduced from Phillips RC and Meñez EG (1988) *Seagrasses. Smithsonian Contributions to Marine Science No. 34.* Washington, DC: Smithsonian Institute, and Kirkman H and Walker DI (1986) Regional studies—Western Australian seagrass. In: Larkum AWD, McComb AJ, and Shepherd SA (eds.) *Biology of Seagrasses*. New York: Elsevier.

It is clear that a major revision of the taxonomy of seagrass species is needed, and that this revision cannot be based on anatomical or morphological characters alone. Experiments to test for phenotypic plasticity and the more parsimonic methods of molecular taxonomy must play a critical role in this revision. The use of biogeographic separation as diagnostic criteria for species and genera is also questionable because the current distribution of seagrasses may result from local extinctions isolating previously continuous distributions. Indeed, populations of some species with a wide distribution are genetically isolated across biogeographical regions (e.g., *Zostera marina*), although there is no evidence that such isolation may have resulted in sufficient genetic changes as to render them different species. In fact, the description of some seagrasses as separate species on the basis of their biogeographic separation (e.g., *Zostera* and *Phyllospadix* species at either side of the Pacific) has yet to be validated by robust genetic analyses. Indeed, examination of seagrass phylogeny using molecular markers has led to the postulation that there is no basis to separate *Zostera* and *Heterozostera* species into different genera (Les *et al.*, 1997) and has generated a new, more parsimonic, tree of genetic relationships between the species (**Table 2**) that represents a useful departure point for the needed reassessment.

Independently of current uncertainties regarding the exact size of the seagrass flora, it is clear that seagrasses are minor contributors to the richness of plant species in the sea, where they represent $<0.5\%$ of the oxygen-evolving photosynthetic species. This minor contribution cannot be accounted for by the comparatively much longer evolutionary history of marine algae because seagrasses also represent $<0.02\%$ of the angiosperm flora, despite an early origin in the angiosperm evolution. Although extinctions of some genera are documented in the fossil record, it is unlikely that past seagrass floras ever comprised a substantially greater number of species than does the extant flora.

The reason for the paucity of speciation in seagrasses remains a subject of intense debate, involving many possible constraints which operate collectively to result in the remarkably low membership of the seagrass flora. The evolution of the land angiosperm flora is closely associated to a coevolution with insect pollinators. The fact that there are no insects in the sea has been used to argue that the very low extent of speciation in the seagrass flora is attributable to the inefficiency of hydrophilous pollination (Van der Hage, 1996). The stress associated with submerged life in a highly saline, high-energy environment may also further constrain the number of species which have been capable of undergoing the needed physiological, anatomical, and architectural (all seagrass species are rhizomatous) adjustments. No single constraint can be invoked in isolation to account for the low number of marine angiosperm species. The constraints imposed by hydrophilous pollination are not likely to differ much from those associated with the fertilization of red algae, the most species-rich phylum of marine algae. The adaptation of angiosperms to underwater life has led to approximately 500 species in fresh waters, and hundreds of angiosperms have adapted to life in saline environments. In addition to these factors, the greater possibilities for exchange of genetic material between distant populations in the sea, which reduces allopatric isolation, may have played an important role in accounting for the small number of marine angiosperms when compared with the apparent shorter-range exchange between angiosperms on land or in fresh water (Ackerman, 1998).

Table 2 Classification of seagrass families, subfamilies, and genera derived from analyses of morphological and reproductive structures (den Hartog, 1970) and the more robust analysis of parsimonic molecular phylogeny (Les *et al.*, 1997)[a]

Morphological phylogeny	*Molecular phylogeny*
Potamogetonaceae	Zosteracea
Subfamily Zosteroideae	*Heterozostera*
Heterozostera	*Phyllospadix*
Phyllospadix	*Zostera*
Zostera	
Subfamily Posidonoideae	Cymodoceaceae
Posidonia	*Amphibolis*
Subfamily Cymodoceoideae	*Cymodocea*
Amphibolis	*Halodule*
Cymodocea	*Syringodium*
Halodule	*Thalassodendron*
Syringodium	
Thalassodendeon	Posidonieaceae
	Posidonia
Hydrocharitaceae	Hydrocharitaceae
Subfamily Halophiloudeae	*Enhalus*
Halophila	*Halophila*
Subfamily Hydrocharitoideae	*Thalassia*
Enhalus	
Subfamily Thalassioideae	
Thalassia	

[a]Ruppiaceae were included in the original analysis of molecular phylogeny. Although closely related to the seagrass taxa, Ruppiaceae are not restricted to life in the sea; therefore they are often excluded from the seagrasses as an ecological category of obligated marine plants.

Source: Reproduced from den Hartog C (1970) *The Seagrasses of the World*. Amsterdam: NorthHolland, and Les DH, Cleland MA, and Waycott MA (1997) Phylogenetic studies in Alismatidae. II. Evolution of marine angiosperms (seagrasses) and hydrophily. *Syst. Bot.* 22: 443.

Seagrass Biogeography and Species Richness

The continuity of the marine environment across large spatial scales is best portrayed by the vast distributional ranges of many seagrass species, as exemplified by *Zostera marina*, which extends from subtropical to subartic waters. Basically, only nine floras, including some areas of overlap, can be identified (**Table 1** and **Figure 2**) in the shallow, coastal areas inhabited by seagrasses in between polar waters (Hemminga and Duarte, 2000). Six of these species assemblages are composed of temperate species, and only two are represented by exclusively tropical (or subtropical) species.

The biogeographic patterns of seagrass floras are parallel to those followed by hermatypic corals, tropical fish, and mangrove species, all of which reach their highest diversity in the Indo-Pacific region, centered at Malaysia, which is believed to represent the center of origin of these organisms. Indeed, the number of seagrass species declines sharply with increasing distance from the region of highest diversity along the major oceanic currents in the region (Mukai, 1993).

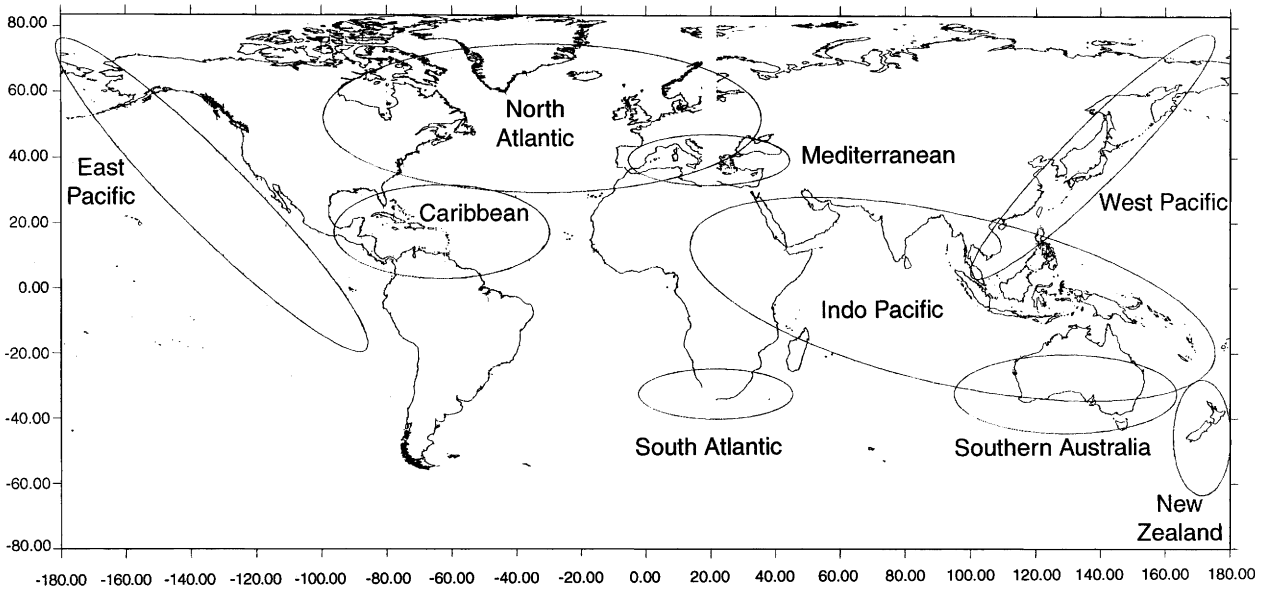


Figure 2 The distribution of seagrass floras. The species present in each flora are listed in **Table 1**.

The boundaries between seagrass floras are often sharp, corresponding to frontal areas separating different water masses, such as the Almería-Oran density front separating Atlantic from Mediterranean surface waters. Oceanographic fronts act as physical barriers for the dispersal of organisms, preventing the dispersal of seagrass propagules. The importance of major oceanographic fronts as boundaries between different seagrass floras indirectly validates the perception that the low number of seagrass floras and species result from the continuity of the marine environment over large scales.

Geographically isolated floras, however, have many common characteristics, such as the presence of congeneric species assemblages in distant floras such as the twin assemblage of *Thalassia–Halodule–Syrigingodium* in the Caribbean and the Indo-Pacific regions and *Zostera–Phyllospadix* species in both sides of the temperate Pacific zone (**Figure 2**). Single genera are also common to different floras, such as *Posidonia* present in the Mediterranean and Southern Australian floras (**Table 1** and **Figure 2**). The occurrence of twin species suggests an even wider early distribution of the species, separated by the rearrangements of continental mass and local extinctions. The high species richness in the Indo-Pacific region (**Figure 2**) suggests that this may have represented an area of particularly high speciation acting as a focus for radiative dispersal of the species or that, alternatively, extinctions have been fewer there.

Distribution and Controls of Seagrass Species Diversity

Distribution and Controls at the Global Scale

The biogeographic factor constraining the distribution of floras is the major determinant of species richness at the global scale because it sets the maximum species richness possible at any one site. This maximum, however, may never be realized. For example, in Western Australia, the flora contains many species which often grow segregated, forming mostly monospecific stands. The maximum species richness reported at any one stand is 12 species, but the majority of meadows studied are monospecific, with an exponential decline in the number of meadows with increasing number of species (**Figure 3**). The mean richness of 596 descriptions of seagrass meadows encountered in a thorough revision of the published literature is approximately two seagrass species per meadow (mean $\pm$ SE $= 1.99 \pm 0.07$; Hemminga and Duarte, 2000). The resulting diversity, as calculated from the biomass distribution of the seagrasses using the Shanon–Weaver index, is closely linked to the species richness. The majority of seagrass meadows are monospecific and therefore have the minimum diversity index of 0. The highest Shanon–Weaver diversity index for a seagrass meadow was 1.56 for a seagrass meadow containing seven seagrass species in Bolinao (Pangasinan Province, the Philippines), and indices >1 have only been reported for meadows in the Indo-Pacific region, although the available estimates for such comparisons are few. These values compare poorly with diversity indices calculated for most other marine taxa, which generally oscillate between 1 and 4 for marine algae, invertebrates, and fish assemblages (Margalef, 1980).

The low species richness of seagrass stands, with most meadows being monospecific or containing two species, extends to areas occupied by floras with high numbers of species, such as Western Australia, in which species tend to be segregated. Mixed meadows, however, are common in the Red Sea and the Indo-Pacific region (**Figure 4**), with the highest species richness found in Philippines waters. As a result of these patterns, seagrass species richness shows a clear latitudinal pattern, with the average number of seagrass species being greatest in meadows near the equator and least in meadows in cold, temperate areas (**Figure 5**). The seagrass species richness in the subtropics in the Southern Hemisphere is somewhat greater than that for similar latitudes in the

Northern Hemisphere (Figure 5), reflecting the high species richness in the Australian continent and the Indian Ocean. The pattern toward species richness declining from the tropics to the cold temperate zone is comparable to those described for terrestrial taxa.

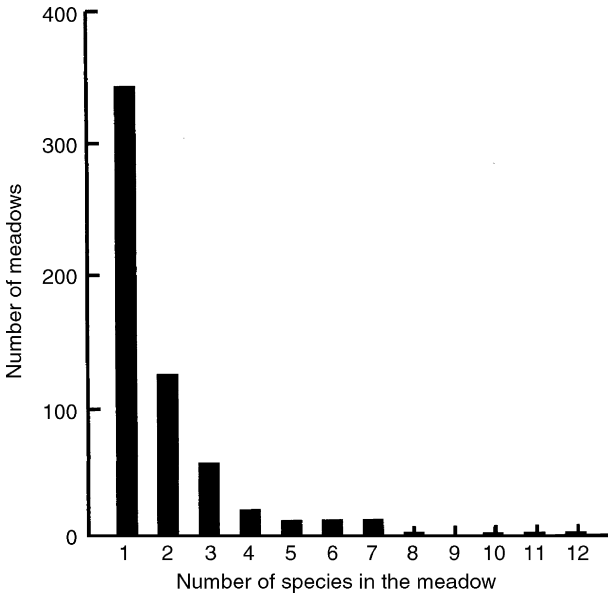


Figure 3 The number of species present in 596 seagrass meadows reported in the literature.

Distribution and Controls at the Regional Scale

The species richness encountered at any one location is constrained by the size of the flora present in the particular biogeographic region and the suitability of growth conditions in the area to support the different species. The study of the habitat requirements of seagrasses has received considerable attention and is relatively well delineated. Seagrass life is restricted to marine waters, extending to estuarine areas with low salinity, such as those encountered in the Baltic Sea. Seagrasses are further constrained by exposure to desiccation and physical disturbance by waves or, in cold seas, ice scouring in shallow waters, and the downslope limit is imposed by light, with seagrasses penetrating to a depth receiving, on average, 11% of the surface irradiance (Duarte, 1991). Within these limits, seagrasses require appropriate substratum which, for most species—except *Phyllospadix*, *Amphibolis*, and *Posidonia*, which are able to grow attached to rocks—is sandy to silty sediments. Highly silted sediments, which are also associated with reduced sediment conditions and accumulation of phytotoxins, are not appropriate to support seagrass life, which is restricted to sediments with low organic contents, redox potentials above -200 mV, and sulfide concentrations <100 μM. Although these growth requirements define habitats appropriate to support seagrass life, different species are expected to have different tolerances to these stress factors, which should lead to differences in species richness among locations at the regional scale.

Species-specific growth requirements are poorly documented for seagrasses, but some regional patterns are

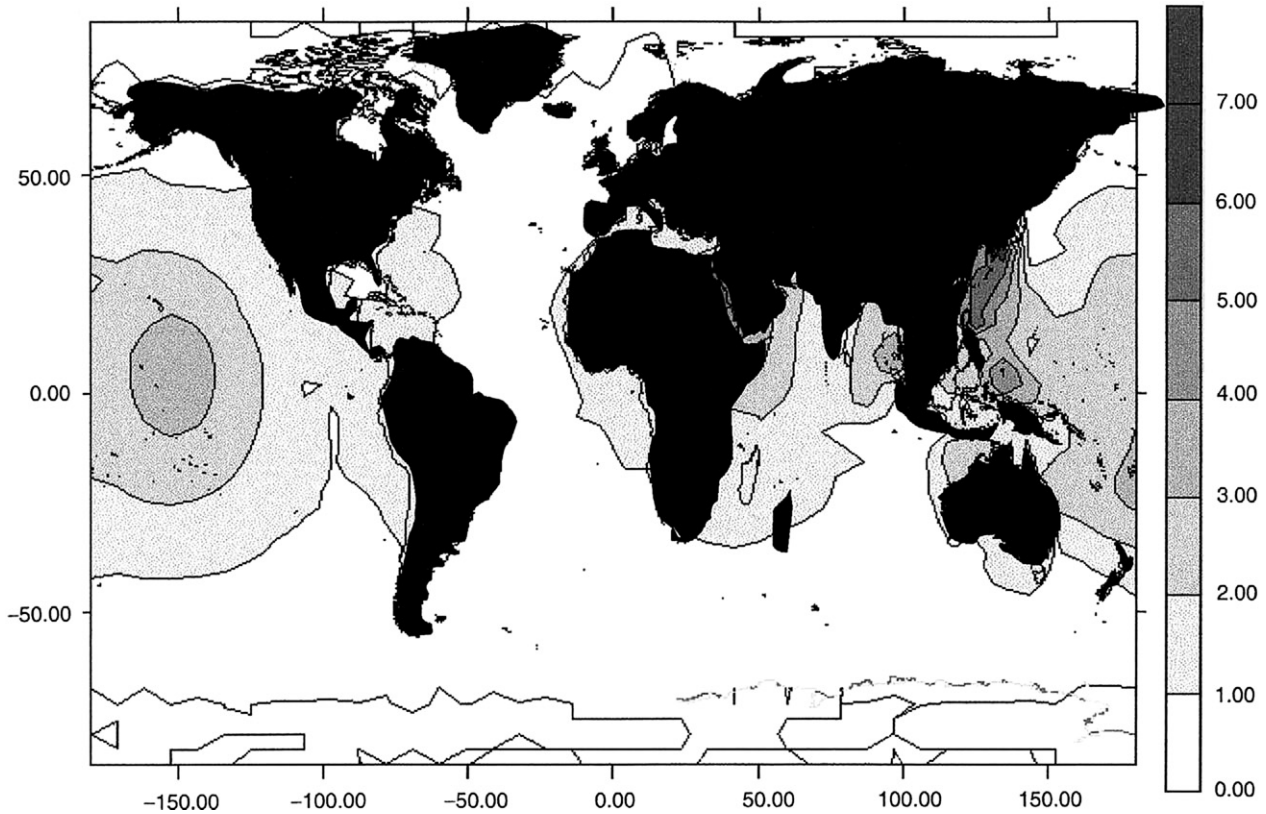


Figure 4 Contour isolines of seagrass species richness generated from data on 596 seagrass meadows reported in the literature.

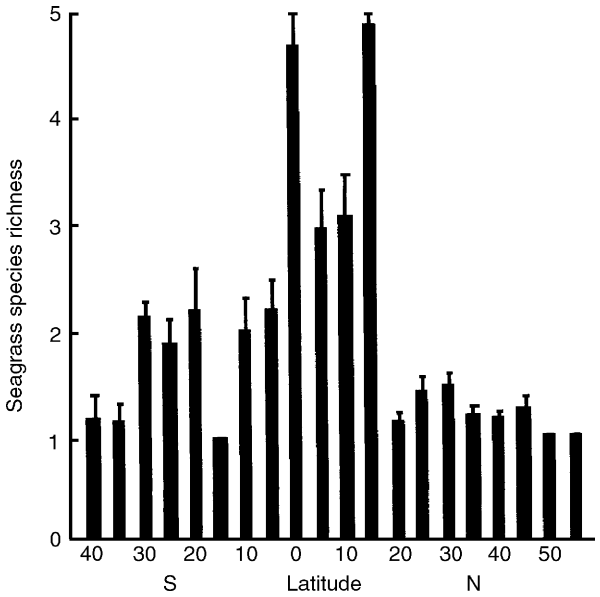


Figure 5 The mean (±SE) species richness in seagrass meadows growing in different latitudinal ranges. Compiled from data on 596 seagrass meadows encountered in a thorough revision of the literature.

emerging. Seagrass species richness declines along gradients of siltation in Southeast Asian coastal waters, with a steep reduction in species richness as the silt content of the sediments exceeds 15%, and only mono-specific (*E. acoroides*) meadows occur at highly silted locations (**Figure 6**). That these patterns derive from species-specific differences in the resistance to disturbance has been predicted by theory and verified experimentally (Duarte *et al.*, 1997).

The Maintenance of Local Species Diversity

Seagrass species are closely related and therefore often engage in competitive interactions in the resource-poor environments which they inhabit. Competitive interactions, if strong, could lead to species replacement and the dominance of one species following a period of coexistence. Indeed, disturbance experiments have documented how a period of high diversity during intermediate stages of colonization is followed by the dominance by the climax species, which is typically slow growing. That the occurrence of such monospecific stands of climax species may results from strong resource limitation has been demonstrated by long-term fertilization experiments of a monospecific *Thalassia testudinum* meadow, which was colonized by the pioneer species *Halodule wrightii* following experimental nutrient additions (Fourqurean *et al.*, 1995).

High disturbance is also expected to lead to species-poor meadows which contain only small, pioneer species capable of fast colonization and growth, often arrested in a permanent stage of colonization. This is best illustrated by the occurrence of meadows containing only *Halophila* species in areas visited by dugongs in Southeast Asia. The grazing mode of the dugong uproots the seagrass, creating small gaps (feeding tracks about 2 m long and 30 cm wide) that must be recolonized in between successive visits. Only the small *Halophila* species, able to grow

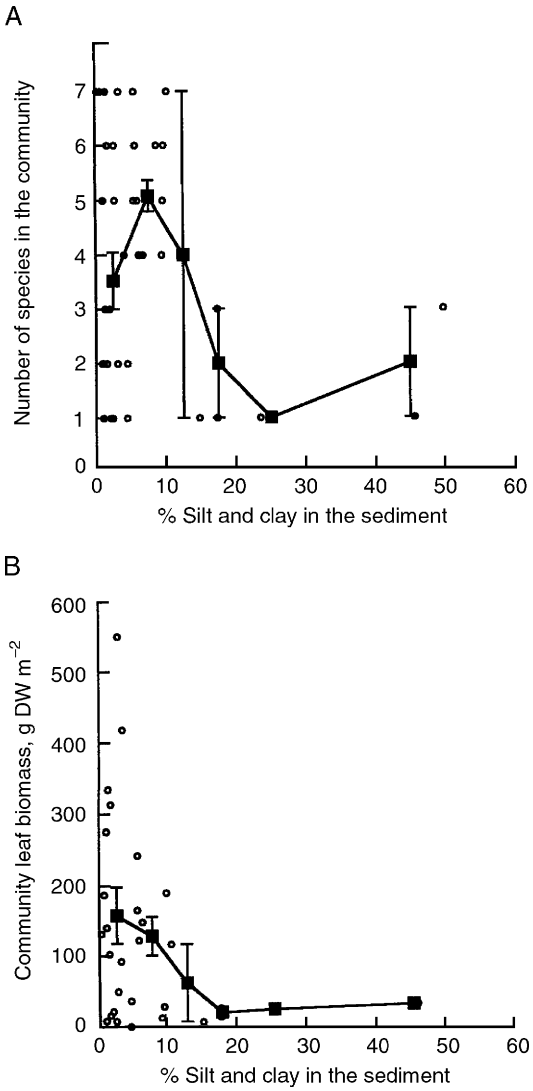


Figure 6 The reduction in the average species richness and community biomass of Southeast Asian seagrass meadows with increasing sediment silt content (Reproduced from Terrados J, Duarte CM, Fortes MD, *et al.* (1997) Changes in community structure and biomass ol seagrass communities along gradients of siltation in SE Asia. *Est. Coastal Shelf Sci.* 46: 757–768.).

9 m annually, are sufficiently fast colonizers to withstand this pressure. Because *Halophila* leaves are also tender, low in fibers, and nutrient rich, dugongs have been postulated to be actively farming these seagrass species (Preen, 1995).

Hence, seagrass species diversity is expected to comply with the intermediate disturbance hypothesis, in which the highest diversity is expected at moderate levels of disturbance. Multispecific seagrass meadows would therefore develop in the presence of moderate disturbance levels. The maintenance of the high species diversity in the most diverse meadow reported to date, located in the Philippines, has been attributed to the disturbance induced by the activity of burrowing shrimps, which maintain small-scale ($<m^{-2}$) patchiness by generating gaps that allow the proliferation of pioneer species in meadows otherwise dominated by large, climax species (Duarte *et al.*, 1997).

Although the preceding discussion highlights the role of competitive interactions, there is also evidence of positive interactions whereby the presence of some species may facilitate the occurrence of other species, thereby maintaining high seagrass diversity. This role has been assigned to "engineering species" (Duarte *et al.*, 1997) able to modify the growth conditions to improve their suitability to maintain plant growth. The occurrence of some seagrass species in Southeast Asia has been reported to be associated with the presence of the dominant species, *Thalassia hemprichii*, which is also linked to high biomass development in the meadows (Terrados *et al.*, 1997). Seagrasses are able to alter the sedimentary conditions in which they grow, injecting oxygen into the sediments through their roots, thereby avoiding the accumulation of phytotoxins produced under anoxic conditions. Through this effect, seagrass species with extensive root systems, such as *T. hemprichii*, may maintain sediment conditions tolerable for seagrass growth where the organic loading derived from the seagrass growth would otherwise render sediments highly reduced and unsuitable for seagrass growth.

Seagrass Genetic Diversity

Seagrasses are clonal organisms, so the genetic diversity present in meadows may be low. Indeed, vegetative reproduction is by far the dominant mode of propagation of seagrasses; therefore, the hypothesis that entire meadows may be composed of one or a few organisms cannot be rejected without serious consideration. Some aspects of seagrass biology, such as the occurrence of mass mortality across large spatial scales, have been ascribed to the dominance of vegetative reproduction in seagrasses and an assumed low genetic diversity to resist diseases. However, examinations of seagrass genetic diversity are still few and have lagged well behind studies of genetic diversity in other angiosperm groups.

The introduction in the mid-1990s of molecular techniques to quantify the extent of genetic diversity of seagrass populations is now providing the needed information to test the assumed low genetic diversity of the meadows. Examination of the genetic structure of seagrass populations has shown considerable variability in the extent of genetic diversity within and among species. Genetic diversity can be important in some of the species investigated (e.g., *Posidonia australis*, *T. testudinum*, *Halodule uninervis*, and *Z. marina*), whereas very low levels of genetic variability have been demonstrated for other species (e.g., *Amphibolis antarctica* and *Posidonia oceanica*). Even in the genetically variable species, most of the variation is found within local populations, and there is reduced gene flow at scales of only a few kilometers or even meters along the depth gradients.

The level of genetic uniformity of seagrasses appears to be greater than that typical of terrestrial plants, supporting the view that the low levels of speciation in seagrasses could derive from inefficient sexual reproduction. The significant within-meadow variability in some species falsifies the assumption that seagrasses are primarily clonal and challenges the explanation of high clonality as the factor responsible for large-scale decline in *Z. marina*, the populations of which show considerable genetic variability. In contrast, the widespread decline of *P. oceanica* in the Mediterranean could be associated with its high clonality. Although some genetic variation has been reported for populations along depth gradients, sizable variation has only been found near *P. oceanica's* biogeographical limit in the Mediterranean–Atlantic transition zone. Indeed, a seagrass (*Z. marina*) clone growing in the Baltic has been identified as the largest and oldest marine organism found to date (Reusch *et al.*, 1999). The results to date provide no patterns of possible regulatory factors in seagrass genetic diversity because the differences observed among species do not appear to correspond to differences in reproductive mode or dispersal properties. The study of the genetic structure of the seagrass population is an active area of research; therefore, more robust conclusions than those currently available are expected to be derived in the near future.

Biological Diversity in the Seagrass Habitat

As indicated previously, seagrass ecosystems are important reservoirs of biological diversity in the sea. Seagrass ecosystems tend to contain more species than the adjacent unvegetated bottoms, and the species common to both seagrass meadows and unvegetated areas tend to be more abundant within the meadows. The enhanced biological diversity and richness within seagrass meadows result from their importance as habitat for (i) species that use seagrasses, and the rich detritus they produce, as food sources; (ii) species that use seagrass surfaces as substratum; and (iii) species that seek refuge within seagrass meadows. These species encompass a broad taxonomic spectrum from fungi to mammals.

Microbial Diversity

The abundance of microbial communities appears to be enhanced within seagrass meadows compared to adjacent bare sediments. These microbial communities benefit from the large amounts of detritus seagrass produces as well as the efficient retention of carbon deposited from the water column in vegetated sediments. As a result, there seems to be clear links between bacterial activity and seagrass production and nutritional status, and exploratory analyses have shown that seagrass meadows host highly diverse bacterial communities. The association between seagrass meadows and fungi is even clearer because seagrass rhizomes and roots are heavily colonized by fungi, although mychorrizal-type associations do not appear to have been developed in the sea. In fact, some of these fungal species may be pathogenic; in particular, the pathogenic *Laberynthula* species have been found to be widespread in seagrass stands. *Laberynthula* species were associated with large-scale seagrass (*Z. marina*) declines in the Atlantic in the 1930s, although the role of these pathogens on the decline remains hypothetical. Recent analyses have also reported a rich ciliate community to be associated with seagrass leaves. This ciliate community encompasses a wide diversity of taxa, and their role in the seagrass ecosystems has yet to be elucidated. These microbial communities may rely directly on detritus produced by seagrasses or benefit from the trapping of particles within the quiescent waters within seagrass canopies.

The best studied microbial community associated with seagrass meadows is the epiphytic microalgal community that develops on the seagrass leaves. This community is largely represented by large benthic diatoms, which colonize newly produced seagrass leaves and are replaced, as the age of the leaf substratum increases, by macroalgae and sessile invertebrate communities. Hence, there is a gradation in epiphytic community structure from the younger leaf base toward the leaf apex, particularly in species with long leaf life spans such as the Mediterranean seagrass *P. oceanica*, the leaves of which last for about a year. The ephiphytic community is an important component of the ecosystem metabolism within seagrass meadows, often supporting a primary production comparable to that of the host plants; it also makes an important contribution to heterotrophic processes. In addition, the epiphytic community is responsible for high deposition rates of carbonates on seagrass surfaces, particularly in warm waters, and is therefore responsible for the high carbonate deposition in seagrass sediments.

Macroalgae

Macroalgae often live in association with seagrasses, prominently as epiphytes on the leaf apices of species with long leaf life spans. The number of epiphytic macroalgae species exceeds 350, including all major taxonomic divisions, in which blue-green algae are relatively rare. Competitive interactions and other negative effects are particularly intense in eutrophic environments because the abundance of opportunistic algae (*Ulva*, *Enteromorpha*, and *Chaetomorpha*) is promoted under high nutrient loading. The growth of these opportunistic algae can be so extreme that they can develop thick blankets over the seagrass meadows, which may suffocate or uplift (from the buoyancy due to gaseous release) the plants. In addition, competitive interactions may also be established with the siphonales, notably species of the genus *Caulerpa*, the rhizoids of which confer them the ability to extend horizontally over soft sediments and exploit the sediment nutrient pools. Seagrass meadows in Southeast Asia often contain many *Caulerpa* species. Introduction of *Caulerpa* species (*C. taxifolia* and, recently, *C. racemosa*) to the Mediterranean poses an important threat, where invasive expansion occurs, to the seagrass meadows (e.g., *P. oceanica*).

Invertebrates

The invertebrate fauna associated with seagrasses is very rich, including a multitude of organisms living epiphytically on their leaves and on their rhizomes and associated with seagrasses on the sediments. Epiphytic fauna comprise more than 170 species. Meiofauna include nematodes, rotifers, tardigrades, copepods (planktonic and epibenthic harpacticoid species), and ostracods. Copepods and other free-living zooplankters often swarm within seagrass beds. The leaves support a rich community of epiphytic filter-feeding sessile organisms, including hydrozans, ascideans, and bryozoans. Spongi grow attached to the seagrass rhizomes exposed aboveground and sometimes epiphytically on their leaves (e.g., *E. acoroides*). Polychaetes are abundant within seagrass sediments and also reach high densities on exposed seagrass rhizomes. Borer polychaetes have been reported to cause damage to seagrass leaves and rhizomes.

Gastropods rank among the most abundant organisms on seagrass leaves, on which they graze the epiphytic algae and sometimes erode the leaf epidermis, damaging the seagrasses. The large gastropod *Strombus gigas*, the queen conch, is an important consumer in Caribbean *T. testudinum* meadows, but heavy fishing pressure has decimated its populations. Seagrass beds are important habitat for bivalves including scallops, which show behavioral adaptations toward a preferential recruitment to seagrass beds. Bivalves may experience an increased growth rate within seagrass beds, as reported for Mercenaria clams and mussels, due to the particle trapping by seagrasses and may also be less prone to predation there. These bivalve populations engage in mutualistic relationships with the plants, such as those reported for blue mussels (*Mytilus edulis*) and temperate seagrasses, because their enhanced activity within seagrass beds provides inorganic nutrients to the surrounding waters and sediments which promote seagrass growth. An outstanding association between seagrasses and bivalves is that by *Pinna nobilis*, a very large (up to 80cm tall) bivalve that is largely restricted to *P. oceanica* beds. *Pinna nobilis* is a protected Mediterranean species and is particularly endangered because, in addition to direct threats from collection and predation, its populations are declining as a result of the widespread decline of *P. oceanica* across the western Mediterranean. Furthermore, *P. nobilis* is the host of a crustacean species (*Pontonia pinnophylax*) which lives in mutualistic association inside its large shell and is also an endangered species because of the decimated populations of the host bivalve. This nested association between *P. oceanica*, *P. nobilis*, and *P. pinnophylax* has been termed a "biodiversity Russian doll" (Richardson *et al.*, 1997) and used as a paradigm of how the threat to landscape components, such as seagrass meadows, involves a threat to the many species therein contained. Although nudibranchs are not particularly prominent components of seagrass beds, sea hares (Aplysia) have been observed to graze on seagrass leaves, particularly when gathering in large numbers during the reproductive season.

Isopods and amphipods are major consumers of both seagrasses and their epiphytes. In particular, the isopod *Idotea* plays an important role as a link between seagrasses and higher trophic levels in food webs. Amphipods rank among the most abundant organisms within seagrass meadows. Thalassinid shrimps are important components of seagrass meadows, particularly in tropical waters in which their burrowing activity generates gaps within the seagrass and affects the growth of the adjacent shoots. Thalassinid shrimps can also be major grazers in seagrass meadows, removing up to 30% of their production. Thalassinid shrimps clip leaves of seagrass species adjacent to their burrows and transport them into the burrows to decompose prior to consumption. In addition, their burrowing activity and the active and passive ventilation of their burrows appear to have a major effect on the sediment conditions, attenuating the reduced stage of the sediment while increasing nutrient availability, which may promote seagrass growth. Hence, Thalassinid shrimps are important engineering species within (particularly tropical) seagrass beds, in which their disturbance effect has been

equated to that of typhoons and is responsible for generating small-scale, highly dynamics gaps that are believed to be essential for maintaining multispecific seagrass beds (*see* Distribution and Controls of Seagrass Species Diversity). Seagrass beds are also important nursery areas for prawns such as tiger prawns (e.g., *Penaeus semisulcatus* and *P. esculentus* in Australian waters), which have been found to feed on seagrass seeds, and caridean shrimps are abundant in subtropical and tropical seagrass meadows. Decapod crustaceans such as *Callinectes* spp. can also be important predators of seagrass seeds. Fiddler crabs (*Uca* spp.) are also very abundant in tropical and subtropical intertidal beds, in which they use their developed claw to remove the epiphytes of the seagrass leaves for consumption.

Detritivore echinoderms, such as holthurids and ophiurids, and herbivores such as sea urchins are abundant within seagrass meadows and are among the dominant invertebrate components therein. Sea urchins can graze heavily on seagrass leaves—to the point that sudden population outbreaks may defoliate seagrass meadows over vast areas.

Vertebrates

The fish communities within seagrass beds typically comprise between 12 and 50 species in any one meadow. Fish communities tend to be more abundant within seagrass meadows than in adjacent bare bottoms, and they seem to seek refuge from predation there or benefit from the higher prey densities often encountered within the seagrass bed. Only approximately 20 species worldwide graze on seagrass. Some species, however, graze on seagrass and/or the epiphytes present on their leaves, such as *Boops salpa* (a herbivorous species in the Mediterranean), parrot fish, and rabbit fish (*Syganus* sp.). Cryptic species such as stonefish are often encountered within seagrass meadows, and gobids are also abundant in seagrass meadows, sometimes in association with Thalassinid shrimps. Active predators are often observed visiting seagrass meadows to prey on the abundant animal resources present therein. Other species, such as rays, can act as important disturbance agents by uprooting the plants when hiding under the sediment. Demersal fish production has been found to correlate with seagrass biomass, likely reflecting the higher prey production within these ecosystems.

The main reptile users of seagrass beds are adult sea turtles, which are important consumers of seagrass meadows. The most studied species is the green turtle, *Chelonia mydas*, which grazes on the seagrass. The association of turtles with seagrass meadows is so strong as to be reflected in some vernacular names (e.g., turtlegrass for the Caribbean species *T. testudinum*). Their importance as grazers in seagrass beds, however, has decreased due to the decline of the populations. In Southeast Asia, sea snakes are often observed foraging on seagrass beds.

In temperate climates, intertidal seagrass beds (mostly *Zostera* sp.) often support migrant goose and swan populations. They are the most important vertebrate grazers in cold, temperate waters, in which low water temperature would hamper the digestibility of seagrasses by poikilotherm grazers. Goose and swans feed on both leaves and rhizomes, and they may play an important role in the dispersal of the plant seeds. Many bird species, including waders, shore birds, gulls, pelicans, and diving birds, feed on seagrass meadows, particularly on intertidal ones.

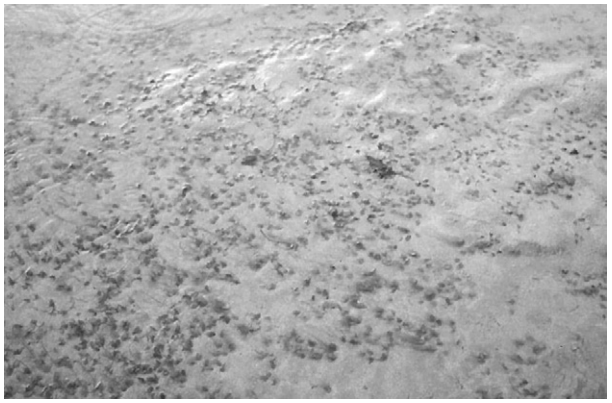


Figure 7 A *Halophila ovalis* population in a large intertidal flat visited by dugongs in Con Dao Island (South Vietnam).

The syrenidae or marine cows are large marine herbivores that feed exclusively on seagrasses and, in the case of manatees (Caribbean waters), also on freshwater plants growing in rivers. Dugongs (*Dugong dugong*) prefer fast-growing, small seagrasses, especially *Halophila*, which in Southeast Asia develop large stands in intertidal flats (**Figure 7**), have low lignin and cellulose contents, and regenerate fast. Dugongs feed by absorbing the sediments through their snouts and filtering out the seagrass while dropping the finer sediment particles. Their feeding activity leaves behind distinct feeding trails, usually approximately 20–30 cm wide and 1 or 2 m long, that are bare of seagrasses. In contrast, manatees graze only on the leaves and do not demonstrate a digging behavior. Dugongs and manatees are threatened and protected throughout their ranges. Although the manatee population is still significant, dugong (Southeast Asia to Australia) populations have disappeared from most of their original range, and sizable populations have only been reported from the Great Barrier Reef. Although active hunting of sea cows has been abandoned, their loss to the point of extinction of local populations continues due to mortality induced by fishing gear, dynamite and cyanide fishing, and accidental collisions. As a result, the dugong is now among the most endangered marine mammals at great risk of extinction.

The Function and Management of Seagrass Diversity

Seagrass meadows are one of the ecosystems that provide most valuable services for the biosphere. They produce approximately 0.60 Gt C $year^{-1}$ and are responsible for one-sixth of the total net CO_2 uptake by oceanic biota (Duarte and Chiscano, 1999). Seagrasses provide habitat for animals, many of which are economically important or endangered, and also play an important role by stabilizing sediments in coastal areas (Hemminga and Duarte, 2000). These functions are closely linked to seagrass diversity because there is evidence indicating that mixed meadows can support considerably

more biomass and be much more productive than less diverse meadows within regions (Terrados *et al.*, 1997). High seagrass diversity may also be linked to a more effective retention of sediments because the more complex canopies of mixed meadows will achieve a greater dissipation of energy by acting at different scales of the energy spectrum than the simpler canopies of monospecific meadows. They are also expected to host a larger number of animal taxa than monospecific ones because of their diversity of substrate and food types. However, the possible functional roles of seagrass diversity remain largely untested, as is unfortunately the case for most other marine taxa.

The importance of seagrass meadows as components of marine diversity does not stem primarily from the diversity of the angiosperm but rather, as indicated previously, from the high diversity that these ecosystems harbor. The role of seagrass meadows as sites of enhanced marine diversity was first addressed in the study of commercially important species, which often use seagrass meadows as nurseries. These studies were later extended to the study of endangered marine organisms, such as marine turtles and dugongs, which rely on seagrasses for food. Indeed, marine biodiversity presents a nested structure within seagrass meadows, which contain endangered species that often act as hosts for species that are necessarily rare, such as the *P. oceanica*—*P. nobilis* (mollusca)—*P. pinnophyla* (decapoda) association described in Mediterranean waters (Richardson *et al.*, 1997). The pattern toward high biological diversity within seagrass meadows has been further strengthened by the demonstration of high microbial diversity, with the leaves of seagrass species containing complex microbial assemblages (fungi, bacteria, microalgae, ciliates, and flagellates) which benefit from the substratum offered by the plants as well as their release of detrital carbon to support these active microbial food webs. The plants also benefit from the activity of epiphyte grazers that prevent an excessive accretion on seagrass leaves, alleviating the negative effects of moderate levels of eutrophication for the seagrasses. Seagrass ecosystems also support food webs distant from the meadows; exported seagrass material has been documented as the carbon basis for some deep-sea communities, and seagrass material accumulating on shore, which may form mounds up to 3 m in height, is used by insects and other terrestrial animals.

Managing Seagrass Diversity: Threats

The limited number of seagrass species in the ocean implies that threats to one or a few species may have devastating consequences for the overall diversity of these plants. Unfortunately, catastrophic seagrass decline has been reported for many seagrass species (Short and Wyllie-Echevarria, 1996), both locally and across entire basins, such as the wasting disease that decimated *Z. marina* in the Atlantic Ocean in the 1930s and the widespread decline of *P. oceanica* in the Mediterranean Sea (Marbá *et al.*, 1996). These losses often result from direct human-induced disturbance, primarily reduced water clarity derived from increased nutrient loading, but they also involve diffuse sources likely associated with changes in climate (Marbá *et al.*, 1996) which may increase the occurrence of diseases (Short and Wyllie-Echevarria, 1996). Human disturbance leading to seagrass loss includes many factors in addition to eutrophication, such as siltation derived from deforestation (Terrados *et al.*, 1997), mechanical damage from boating, dredging, filling, dynamite fishing and trawling, and altered sediment dynamics after coastal constructions (Short and Wyllie-Echevarria, 1996). Remarkably, seagrasses have considerable resistance to the release of toxic compounds (Short and Wyllie-Echevarria, 1996).

The seagrass area lost cannot be quantified due to lack of a reliable baseline information, but it is certainly large since even large meadows may be lost within a few years (Marbá *et al.*, 1996). Concern about the loss of seagrasses, and the associated loss of ecosystem functions, biological diversity, and resources, has been translated into legislation aimed at protecting seagrasses in many countries (e.g., USA, Spain and France), and particular emphasis has been placed on the importance of preserving seagrass meadows at the International Convention of Biological Diversity (i.e., the Rio Convention). The effectiveness of these measures to promote seagrass recovery, however, is uncertain because, although disturbed seagrass meadows can recover, the current capacity to predict this process is meager. Models of seagrass recovery timescales predict vast differences in recovery time among species, with small, fast-growing species being able to recover within a few years provided that adjacent reproductive populations exist, and the recovery of slow-growing, long-lived species such as *P. oceanica* involving timescales of centuries (Duarte, 1995). Therefore, restoration efforts have been designed to speed up recovery, but these efforts have a high cost and can only be implemented in the countries with the strongest economies.

Hence, the management of seagrass biodiversity must be based on prevention measures aimed at achieving sustainability targets. In particular, management practices should be based on the known habitat requirements of seagrasses to maintain water quality targets (particularly water transparency) and sedimentary fluxes (particularly to prevent erosion and reduce burial rates) compatible with seagrass life. Deterioration of these properties will lead to reduced seagrass diversity, species loss, and the loss of the associated seagrass functions and biological diversity in the meadows they form.

See also: Grazing, Effects of. Marine Ecosystems

References

Ackerman JD (1998) Is the limited diversity of higher plants in marine systems the result of biophysical limitations for reproduction or evolutionary constraints? *Functional Ecol.* 12: 975.

den Hartog C (1970) *The Seagrasses of the World.* Amsterdam: North Holland.

Duarte CM (1991) Seagrass depth limits. *Aquat. Bot.* 40: 363–377.

Duarte CM (1995) Submerged aquatic vegetation in relation to different nutrient regimes. *Ophelia* 41: 87–112.

Duarte CM and Cebrián J (1996) The fate of marine autotrophic production. *Limnol. Oceanogr.* 41: 1758–1766.

Duarte CM and Chiscano CL (1999) Seagrass biomass and production: A reassessment. *Aquat. Bot.* 65: 159–174.

Duarte CM, Terrados J, Agawin NSW, Fortes MD, Bach S, and Kenworthy WJ (1997) Response of a mixed Philippine seagrass meadow to experimental burial. *Mar. Ecol. Prog. Ser.* 147: 285–294.

Fourqurean JW, Powell GVN, Kenworthy WJ, and Zieman JC (1995) The effects of long-term manipulation of nutrient supply on competition between the seagrasses *Thalassia testudinum and Halodule wrightii* in Florida Bay. *Oikos* 72: 349–358.

Hemminga M and Duarte CM (2000) *Seagrass Ecology.* Cambridge: Cambridge Univ. Press.

Jones CG, Lawton JH, and Shachak M (1997) Positive and negative effects of organisms as physical ecosystem engineers. *Ecology* 78: 1946–1957.

Kirkman H and Walker DI (1986) Regional studies—Western Australian seagrass. In: Larkum AWD, McComb AJ, and Shepherd SA (eds.) *Biology of Seagrasses.* New York: Elsevier.

Les DH, Cleland MA, and Waycott MA (1997) Phylogenetic studies in Alismatidae. II. Evolution of marine angiosperms (seagrasses) and hydrophily. *Syst. Bot.* 22: 443.

Marbá N, Duarte CM, Cebrián J, *et al.* (1996) Growth and population dynamics of *Posidonia oceanica* on the Spanish Mediterranean coast: Elucidating seagrass decline. *Mar. Ecol. Prog. Ser.* 137: 203–213.

Margalef R (1980) *Ecología.* Barcelona: Omega.

Mukai H (1993) Biogeography of the tropical seagrasses in the western Pacific. *Aust. J. Mar. Freshwater Res.* 44: 1.

Pettitt JM (1984) Aspects of flowering and pollination in marine angiosperms. *Oceanogr. Mar. Biol. Ann Rev.* 22: 315–342.

Phillips RC and Meñez EG (1988) *Seagrasses. Smithsonian Contributions to Marine Science No. 34.* Washington, DC: Smithsonian Institute.

Preen A (1995) Impacts of dugong foraging on seagrass habitats: Observational and experimental evidence for cultivation grazing. *Mar. Ecol. Prog. Ser.* 124: 201–213.

Reusch TBH, Borström C, Stam WT, and Olsen JL (1999) An ancient eelgrass clone in the Baltic. *Mar. Ecol. Prog. Ser.* 183: 301–304.

Richardson CA, Kennedy H, Duarte CM, and Proud SV (1997) The occurrence of *Pontona pinnophylax* (Decapoda: Natanti: Pontoniinae) in the Fan mussel *Pinna nobilis* (Mollusca: Bivalvia: Pinnidae) from the Mediterranean. *J. Mar. Biol. Assoc. U.K.* 77: 1227–1230.

Short FT and Wyllie-Echevarria S (1996) Natural and human-induced disturbance of seagrasses. *Environ. Conserv.* 23: 17–27.

Terrados J, Duarte CM, Fortes MD, *et al.* (1997) Changes in community structure and biomass ol seagrass communities along gradients of siltation in SE Asia. *Est. Coastal Shelf Sci.* 46: 757–768.

Van der Hage JCH (1996) Why are there no insects and so few higher plants in the sea? New thoughts on an old problem. *Functional Ecol.* 10: 546.

Slash-and-Burn Agriculture, Effects of

Stefan Hauser, IITA Democratic Republic of Congo, Kinshasa, DR Congo
Lindsey Norgrove, CABI, Delémont, Switzerland

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bulk density The mass of the undisturbed dry soil per unit volume.
Cation exchange capacity (CEC) The capacity of a soil to adsorb cations such as calcium, magnesium, and potassium on the surfaces of soil particles and organic matter.
Dormant/Dormancy Condition under which a seed will not germinate, even if all environmental requirements for germination are met.
Fallow A phase in which no crop is on the land and the volunteer vegetation regrows.
Leaf area index Single-sided total leaf area of the vegetation per unit area of land.
Mulch Retained slash covering soil surface; this can be *in situ* slashed or imported biomass.
Pegging The penetration of the fertilized peanut flower into the soil.
Penetrometer resistance Energy required to push a probe of defined size and shape into the soil.
Transpiration Amount of water taken up by the vegetation and released through the leaves to the atmosphere.

Definition of Slash-and-Burn Agriculture

Slash-and-burn agriculture is a generic term for agricultural systems in which the fallow vegetation is manually slashed, left to dry, and cleared from the field by burning before crop cultivation. "Swidden" is an English dialect word for a burned clearing; thus, swidden agriculture is a synonym for slash-and-burn agriculture. After a cropping phase, the land is abandoned to a fallow phase. Later, the cycle is repeated. Only those systems that alternate between crop and fallow phases are included in this definition. Multistory tree gardens, home gardens, and cocoa plantations, where crops are permanently cultivated, are excluded. With the exception of labor, slash-and-burn farmers use few or no external inputs. Implements such as machetes and hoes are most commonly used. Systems where machinery is used for clearance and irrigated systems are excluded from this definition. Not considered in this context are the systems such as the *ankara* of the Western Highlands in Cameroon, the *nkule* of the Tanzanian grasslands, and the *gy* of Ethiopia, in which vegetation is slashed, gathered, covered with soil, and then burned inside the soil mounds.

Reasons for Slash-and-Burn Agriculture

Most slash-and-burn farmers are poor (*see* Geographic Distribution). Unlike many people in developed countries, they have little choice in how to live their lives. Often the only resource available to them is land. Thus, farming, whether subsistence or market oriented, might be their only option. However, even cash crop production may not provide sufficient income for farmers to change to other economic pursuits. Thus, they remain farmers. Farming methods used reflect the economic constraints of the farmers and tend to minimize labor requirements (*see* Labor Requirements in Slash and Burn Agriculture).

Removal of Biomass

After slashing the vegetation, the amounts of resultant biomass debris on the soil are extremely variable, ranging from less than 10 Mg ha^{-1} to more than 500 Mg ha^{-1} dry matter. Large amounts of biomass make planting difficult and would require too much labor to remove manually. Thus, farmers use fire to remove the biomass, giving easier access to the soil for planting.

Release and Addition of Nutrients from Biomass

Where the fallow vegetation has a high proportion of aboveground biomass as wood, the nutrients bound in it are not easily available to annual crops, as wood does not decompose rapidly. Burning the biomass releases the nutrients.

Weed Avoidance and Weed Suppression

Slash-and-burn agriculture farmers have three principal methods to suppress weed infestation: site selection, the timing of the burn, and tree retention. In forested areas, farmers might choose older secondary or primary forest for clearing because the arable weed seed bank is depleted or absent as the fallow phase has exceeded the seed viability period. If fields become too weedy, farmers may abandon them and clear a new field as the labor requirement for a new clearing may be less than weeding an old field. In short fallow rotations, where the fallow length is insufficient to deplete the weed seed bank, some farmers delay the burn, until a flush of weeds has germinated. These weeds are destroyed in the burn.

Pest and Disease Avoidance

As pesticides are not available to most slash-and-burn farmers, they attempt to avoid crop pests and diseases by appropriate site selection. This option is limited to regions where long fallow phases would withdraw the hosts of pests and diseases for a sufficiently long time to eliminate them. However, pests such as rodents and monkeys would not be affected.

Declining Crop Yields

The combination of nutrient depletion and increasing weed, pest, and disease pressures leads to declining crop yields. After 1–4 years, the field is abandoned to a fallow phase usually because of crop yield declines, or high weed biomass decreasing labor productivity to the point where clearing a new field requires less labor than maintaining an old one. In South America, farmers often abandon their fields when they cannot expect the yield of the subsequent crop to be more than half that of the first crop. However, superimposed on this is the social custom of caring less for cropped fields that are expected to yield less, which accentuates and accelerates any decline. When a cash crop is grown, reasons for abandonment can be external. Farmers in Indonesia reportedly abandoned their slash-and-burn pepper farms, before crop harvest, when the market price for pepper fell.

The Importance of Fallow Length; Long and Short Fallow Slash-and-Burn Systems

Two main types of slash-and-burn agriculture are distinguished, which differ in their effects on the environment: long fallow systems (shifting cultivation) and short fallow systems.

The fallow is the successional vegetation that follows the cropping phase. It may be dominated by trees, shrubs, or grasses, depending on the climax vegetation type, management history, and successional stage. From a utilitarian perspective, its purpose is to reverse the degradative processes of cropping. Here, "degradative" refers to soil fertility decline, weed buildup, crop pest or disease buildup, or a combination of these. The fallow can act as a weed, pest, or disease break, and fallow vegetation accumulates nutrients in the aboveground biomass, improves soil physical properties through root penetration, and usually permits the recovery of soil macrofauna. Thus, the sustainability of slash-and-burn agriculture depends primarily on the rate of degradation during the cropping phase, regeneration during the fallow phase, and their relative time allocation.

A model of this is detailed in **Figure 1** (modified after Guillemin, 1956). However, the diagram is a simplification and the actual rates of change are unknown. At low human population density, even on nutrient-poor soils, fallow phases may exceed 20 years; thus, the system is case 1 type "long fallow system." Labor is the main factor limiting the production. For a cropping phase of time a, fallow length $(b+c)$ may exceed the optimum. The optimum fallow length, b, is defined as the minimum fallow length required to maintain and maximize crop yields per unit area in the long term and thus occurs just when the system has recovered enough to permit this.

When population densities increase, more land is required and fallow phases are shortened, up to the minimum length required to restore soil fertility (case 2). However, where human population densities increase to the extent that more land needs to be cropped than in case 2, fallows are further shortened and system recovery is not possible. If no additional inputs are made, this can cause soil fertility and weed, pest, and disease problems, and lead to lower crop yields (case 3 type short fallow system).

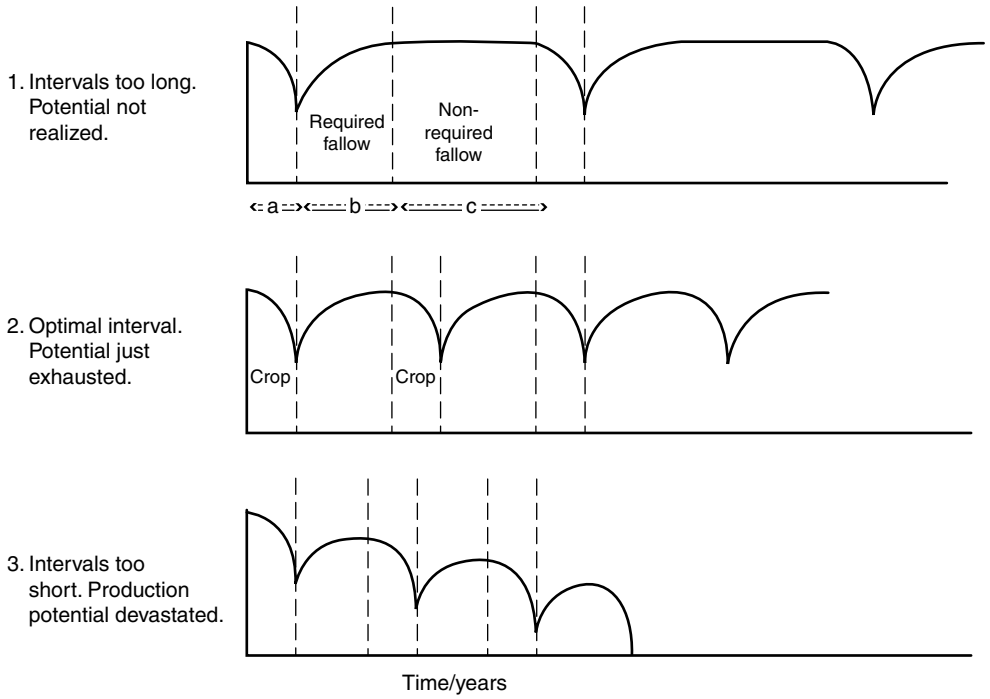


Figure 1 Degradation and recovery in slash-and-burn systems as related to the cropping interval and when no agronomic adjustments are made. Modified from Guillemin R (1956) Evolution de l'agriculture autochtone dans les savanes de l'Oubangui. *Agronomie Tropicale* 11: 143–176, with permission from Springer.

Geographic Distribution and Characteristics of Slash-and-Burn Agriculture

History

Slash-and-burn agriculture is probably the oldest method of land preparation for planting crops. It was used as a technique in China for establishing rice fields as early as the late Stone Age, 8000–10,000 years ago, and in Mexico at least 5500 years ago, concurrent with the domestication of maize. It was used in Central Europe, along the Danube River, to penetrate the postglacial European forests, for cereal production and, later, in the boreal forests of Scandinavia.

Geographic Distribution

Today, slash-and-burn agriculture is practiced in the moist savanna, the savannah–forest transitional and forest zones of the tropics and subtropics of Central and South America, Africa, Asia, and the islands of Australasia. The majority of countries in these regions have weak economies. Slash-and-burn agriculture is also practiced in other nonindustrial areas, such as Bhutan. An estimated 36 million km^2 of land is under slash-and-burn agriculture, approximately 30% of the global soil resource. Human population densities in these areas can vary tremendously. In Africa, rural population densities in some parts of the Congo basin are less than three people per square kilometer, yet exceed 400 people per square kilometer in parts of south-east Nigeria.

Labor Requirements in Slash-and-Burn Agriculture

In areas with low population density, long fallow systems predominate. Long fallow systems in forested areas have large land, yet capital and labor requirements are low. Soil fertility, weed, pest, and disease problems are avoided, rather than managed, by shifting to a new field. In the Amazon basin, only 8% of the human energy input to cultivate a cassava field, including postharvest processing, is used for slashing and burning. The energy or labor efficiency of long fallow slash-and-burn systems is the highest among all agricultural systems in terms of energy invested versus energy gained with the crop yield. This is largely due to the absence of fossil fuel use and chemical inputs. However, the destruction of the biomass and the release of the accumulated energy and carbon (*see* Consequences of Burning) in the burn are not considered in such calculations.

In areas with higher population densities and consequently shortened fallow periods, economic conditions become more suitable for market oriented, commercial farming as the higher population densities naturally create markets. Although labor is available, small-scale farmers have scarce financial resources to purchase agricultural inputs. Owing to infrastructural, economic, and soil-related problems of pesticide and fertilizer use, high-input, intensive agriculture is rarely practiced. With reduced fallow length, the labor requirement increases as additional field work, such as weeding and tillage, becomes necessary (*see* Effects During the Cultivation Phase). In short fallow systems of southern Cameroon, land preparation including slashing, burning, and cleaning the soil surface of unburned debris and weed stumps was 10–13% of the total labor required to establish and harvest a peanut/maize/cassava intercrop.

Although population densities are increasing in many tropical areas and thus fallow periods are shortened, in some situations the reverse has occurred. There were decreases in the population density of the Mayan lowlands, and this was followed by a transition from intensive agriculture to long fallow slash-and-burn systems.

Typical Fallow/Crop Sequences in Slash-and-Burn Agriculture

In West Kalimantan in the mid-1990s, fallow lengths averaged 17 years after a 1-year cropping phase; thus, less than 10% of land was cultivated at any time. Such areas are characterized by low availability of labor and a lack of infrastructure, leading to very limited exchange of products and thus predominantly subsistence farming. In 1996, in southern Cameroon, 32% of fields cultivated had a previous fallow length of 8 years or more or had not been previously cultivated. In the early 1980s, in north Sierra Leone, in the forest–savanna transition zone, most fields had been established from fallows of 30 years or more. Most farmers planted upland rice (*Oryza sativa*) for 1 year before abandoning the land. A minority of farmers followed rice by a peanut and then millet (*Setaria* spp.) sequence before abandonment (Nyerges, 1989).

In long fallow systems, crop diversity within one field can be high. A root or tuber species is often the dominant crop, which is intercropped with other roots and tubers, grains, vegetables, and herbs and spices. In the early 1980s, researchers counted an average of 10 crop species per 25 m^2 in 3-month-old fields of the Maring people of Papua New Guinea. The dominant crop was taro (*Colocasia esculenta*). Other crops included yams (*Dioscorea* spp.), sweet potato (*Ipomoea batatas*), maize (*Zea mays*), beans (*Phaseolus* spp.), sugarcane (*Saccharum officinarum*), and summer squash (*Cucurbita pepo*). Taro monocrops were also common. In the Columbian and Peruvian Amazon, fields may be dominated by cassava, constituting 80% of crop numbers. In Pará, in the Brazilian Amazon, monocrop maize (*Z. mays*) fields are common and farmers may use fallow 3–30 years old or even primary forest.

In southern Cameroon, four of the most common slash-and-burn field types are as follows:

- The *essep* or long fallow or primary forest conversion field used to gain title to the land. This is approximately 0.25 ha and a few larger forest trees are retained. Relatively shade-tolerant crops, particularly tannia (*Xanthosoma sagittifolium*), egusi melon (*Cucumeropsis mannii*), and plantain (*Musa* spp. AAB) are grown with many minor crops. After the plantain harvest, the fields are either abandoned for 20 or more years or burned and recultivated after 1–4 years to plant a peanut/maize/cassava field.
- The peanut/maize/cassava field (*Arachis hypogaea*, *Z. mays*, *Manihot esculenta*) with plantains (*Musa* spp. AAB) and leafy vegetables as minor components. Fallow length is 8–10 years near the border with Equatorial Guinea yet 2–4 years near the capital city, Yaoundé.

- Monocrop horticultural systems, particularly tomatoes (*Lycopersicum lycopersicon*) or maize for fresh consumption. These fields are most common close to the capital and external inputs, both inorganic and organic, are used.
- *Musa* spp. (banana and plantain) monocrops or perennials such as oil palm (*Elaeis guineensis*) and cocoa (*Theobroma cacao*). These are usually established after long fallow forest clearing with *Musa* often intercropped with the perennials.

Clearly, slash-and-burn systems can have high or low crop diversity. This does not depend on eco-region or population density, but rather on the preferences of the farmers.

Effects of Clearing at Field Establishment

Farmers in different parts of the world use slash-and-burn techniques, yet their goals may vary. These range from completely and irreversibly removing the initial vegetation to retaining a rich repository of indigenous and useful plant species while introducing crops. Elaborate clearing and tree-felling techniques have been developed by different indigenous cultures to facilitate the drying and burning of biomass. These techniques are summarized by Peters and Neuenschwander (1988).

The types and numbers of trees retained vary considerably, depending on the tradition of the farmers and their intentions. However, usually very large high-canopy trees are not felled because it is not worth the considerable effort involved, it would require paying for the labor of a skilled chainsaw operator, and the felled tree would take up a considerable part of the plot. Other trees retained include those producing fruits, nuts, or medicine, or those believed to maintain or enhance soil fertility or have a commercial value. Such retained trees are a source of seeds, although many such species do not necessarily establish in an environment imposed by slash and burn. Furthermore, these trees can harbor many species of animals and possibly host microsymbionts. The retention of trees in slash-and-burn fields can facilitate the return of the initial vegetation, depending on the stand density of trees retained and the size of the clearing.

Reductions in tree densities through slash and burn or other forms of clearance may affect densities and species numbers of arboreal and litter-dwelling ants. Yet these species may also have a role in pest control. In areas cleared of forest in the Amazon, it was found that isolated trees can retain considerable ant diversity, although the species composition was not identical to that of the nearby forest.

When slash-and-burn farmers in the forests of West and Central Africa clear for a peanut/maize/cassava short fallow field, they completely clear the fallow vegetation and remove all trees so that there is abundant light for these crops, which do not tolerate shade. Additionally, as any mulch layer would impede peanut growth and pegging, the farmer attempts a complete burn. The soil is tilled by hoeing after burning and weeding is frequent to reduce competition. In contrast, the root, tuber, and plantain crops of the "essep" field can be planted even if trees are retained, and if the slash is not completely burned, the soil is not tilled and weeding is neither as intensive nor as frequent as in the short fallow fields.

Partial or complete removal of vegetation has differential effects on the microclimatic and hydrological conditions after clearing. Soil temperatures are usually higher in cleared areas and even higher when the slash is burned. The absence of tall vegetation leads to a different redistribution of rain, as less or no rain will be intercepted in the canopy. Raindrops reach the soil surface with their full impact, potentially contributing to compaction and crust formation at the surface. Surface compaction can lead to temporarily ponding water at the surface and, if on sloping land, can cause erosion. Furthermore, crust formation can impede the gas exchange between soil and atmosphere, resulting in deteriorating conditions for the soil fauna and flora.

Slash-and-burn fields have lower biomass than forests, and crops do not cover the land for the entire year. The leaf area index of crops may not be as high as in a forest; thus, transpiration is lower. With reduced transpiration, water flux through the soil is increased, which can increase mineral losses due to leaching. The root systems of crops have to grow for each cropping cycle, whereas in the forest the root system of the trees remains in place and potentially adds exploited soil volume every growing season. Access to and uptake of water at the start of the rains is thus immediate in forest. Rooting depths of crops are restricted to the upper layers of the soil, whereas forest vegetation roots may penetrate deeper layers drawing on water during dry seasons, not accessible by and available to food crops. The more permanent canopy and root system of forest contributes to less strong fluctuations of most climatic and hydrological factors.

Consequences of Burning

Heat Evolved during the Burn

Few studies have measured the heat evolving from the burning of slashed biomass. The amount of heat evolved depends on the amount of biomass and its water content. Climatic conditions during the burn, such as wind speed, and topographic features, such as the aspect and slope of the field, have an impact on the intensity of the burn and the maximum temperatures attained. Slash-and-burn agriculture in savanna regions or in short fallow systems, where small amounts of biomass are burned, does not cause significant heat-related effects on the soil. Temperatures at the soil surface are relatively low, heat penetration into the soil is limited to a depth of 1–2 cm, and the exposure time to the heat is short. However, even fast burns, which do not attain high temperatures, can kill seeds lying on the soil surface. There are no reports of light burns affecting seeds buried in the soil.

Burning of cleared forest biomass, depending on its water content, can reach 800 °C near the soil surface, and these temperatures will be maintained for a longer period. If temperatures reach 400 °C, soil organic matter may be lost by combustion, resulting in decreases in soil organic C and N. Clay particles may fuse, altering the soil texture and its cation exchange capacity (CEC). The spatial variability of temperatures is higher in a cleared forest than in short fallow or savanna systems. The boles of large trees may burn for more than a week, evolving large amounts of heat and depositing

large amounts of ash in a small area. In southern Cameroon, on an Ultisol, the maximum temperatures attained during burning of 3000 Mg $ha^{-1}$of wood, equivalent to a bole of approximately 0.65 m diameter at a wood density of 0.6 Mg m^{-3}, were 788 °C at the soil surface, 225 °C at 5 cm, 172 °C at 10 cm, and 105 °C at 20 cm depths. The burn lasted for approximately 24 h. Such severe burns usually kill plant seeds even in deeper layers and reduce living microbial biomass in the soil, including mycorrhizae, rhizobia, and other microsymbionts. Further, all soil meso- and macrofauna are affected, depending on the temperature reached and the soil depth to which it penetrated.

Soil, Soil Nutrients, and Soil Physical Properties

Depending on the intensity and size of slash-and-burn clearing, local consequences range from nutrient inputs from the atmosphere such as sulfur dioxide and flying ash to changes in microclimate and increased variation in rainfall pattern, and influxes of pests, diseases, and invasive weeds.

Many soils in forested areas of the tropics have a low nutrient status, aluminum toxicity problems, high levels of phosphorus fixation by iron oxides, and a low CEC due to the dominant kaolinitic clays, which are also susceptible to structural collapse under mechanization. Where the fallow vegetation has a high proportion of aboveground biomass as wood, the nutrients bound in it are not readily available to annual crops, as wood does not decompose rapidly. Burning the biomass releases the nutrients, and, with the exception of nitrogen and sulfur, a large proportion will remain in the ash on the soil surface. Although the nitrogen is almost entirely lost to the atmosphere, the sulfur, as sulfur dioxide, can be redeposited as sulfuric acid near the burn, if the humidity of the air is high. When the fuel (biomass and debris) is very dry, the burn is fast and hot and strong upward air currents result, which carry away the ash particles. Windy conditions during the burn will increase such losses. In the Amazon, element transfer to the atmosphere due to ash particle transport and volatilization has been reported as up to 98% C, 98% N (these actually being inevitable losses), 33% P, 31% K, 24% Ca, and 43% Mg of the initial amounts in the fuel. Sulfur losses to the atmosphere were estimated between 69% and 76% of that in the original biomass.

Wood ash contains calcium, magnesium, and potassium in the form of phosphates, carbonates, silicates, and other elements. The nutrients in the ash are partially water soluble and will be released into the soil solution with the rain. Thus, slash-and-burn agriculture makes nutrients available to crops. Although farmers realize the fertilizing effect of the ash, they are more concerned about achieving complete burns to remove the biomass. The longevity of increases in calcium, magnesium, and potassium concentrations in the topsoil depends partially on the CEC. Although the increased calcium concentrations are likely to be long term because of low calcium mobility and solubility, magnesium and particularly potassium are prone to leaching. Experiments in southern Cameroon showed the fertilizing effect of ash to be very short lived and not exceeding the first two rainy seasons after a burn. Burning usually reduces soil organic matter content. On kaolinitic soils, much of the CEC is on organic matter; thus, a reduction in organic matter will reduce CEC and increase the risk of leaching. In southern Cameroon, it could be shown that increases in exch. K were measurable down to 50 cm within 6 months of burning forest biomass.

However, with the rain dissolving the ash, soil pH usually increases rapidly to above pH 8 and remains high for some months. Topsoil pH usually increases after burning. In soils where aluminum toxicity problems occur, burning reduces extractable aluminum concentrations and increases phosphorus availability.

It has been reported in the literature that the burn loosens the soil and facilitates tillage and planting by reducing bulk density and penetrometer resistance. Direct measurements in southern Cameroon could not confirm these reports. Farmers experience the forest floor to be difficult to penetrate with tools before the burn, probably because of the dense network of small but woody roots in the litter layer and close to the soil surface. The litter layer and the roots will be destroyed in the burn and will no longer be an obstacle. Furthermore, attempts to work the soil before burning are made on dry soil, which is harder, whereas planting after the burn usually happens after some rains have softened the soil.

Hot and long-lasting burns such as under a tree bole produce white ash. The soil color changes to bright red, and compaction and a collapse of its structure can be observed. This is probably due to a relative excess of monovalent cations that lead to a breakdown of cation bridges between clay minerals, the combustion of soil organic carbon and thus substances binding soil particles, and the elimination of soil macro- and microfauna mixing soil particles and organic materials to form stable aggregates. Slash-and-burn farmers recognize such areas in their fields and usually do not plant crops there.

In savanna regions with small amounts of biomass, the effects of burning on the soil are not discernible. Ash inputs do not cause a measurable increase in pH or cation concentrations. Neither soil organic carbon nor total soil nitrogen is affected by such light burns. In grass fallows, the amounts of biomass may not be an obstacle to planting. However, the quality of the slashed biomass is low. Biomass retention as mulch would have adverse effects on crop growth, because soil nitrogen might be immobilized in the decomposition process, causing nitrogen deficiency for the crop.

Most slash-and-burn agriculture systems use only the *in situ* slashed vegetation. There are, however, systems such as the *chitemene* system in southern Africa in which wood from a larger cleared area is collected and concentrated in a smaller area where it is burned. Farmers seek nutrient augmentation or concentration in the topsoil by piling large amounts of wood on the soil surface. The wood is transported to the field from the surrounding Miombo woodland. The area cleared may be up to 20 times larger than the field in which the wood is burned, adding large quantities of nutrients. Although this system appears to be designed to concentrate nutrients, its efficiency in retaining these nutrients remains low. In an experiment conducted in Zambia, 84% of the phosphorus contained initially in the vegetation was accounted for by increases in the top 50 cm of soil at 40 days after burning. However, 57% of this phosphorus was at 20–50 cm depth,

in a region already inaccessible to any growing crop. For potassium, only approximately 10% of the original input was retained by the top 50 cm of the soil at 40 days after burning. The remainder presumably had leached below 50 cm depth.

In other systems, only part of the biomass is burned while a certain portion is used for other purposes such as fencing the field against animals, as fuel wood, or for construction.

Soil Bacteria and Fungi

Burning can reduce living microbial biomass in the topsoil and potentially their decomposing activities. However, there are reports that burning increased microbial biomass and activity in tropical savanna and in a tropical plantation forest. It has been shown in Kenyan soils that total bacteria counts dropped after burning. Nitrogen-fixing bacteria are very susceptible to pH changes after burning. Such effects of the burn on microsymbionts can subsequently affect the composition and productivity of the vegetation as certain species may depend on symbionts.

Many crops have mycorrhizae and are able to access organic phosphorus using the enzyme phosphatase in a hydrolysis reaction. Thus, although burning can increase phosphate availability, if it reduces mycorrhizae, overall effects could be negated, depending on whether the crops grown are mycorrhizal and whether the planting material is infected. Studies in India comparing vesicular–arbuscular mycorrhizal (VAM) infection of plants in burned and nonburned areas of a tropical forest found that infection was reduced in burned areas. However, studies from wet tropical forests in Costa Rica found that mycorrhizae survived burning.

Invertebrates: Soil Fauna

The soil fauna are only directly affected by the burn if they are present in the topsoil during the burn. As fields are usually burned toward the end of the dry season, groups such as earthworms are usually withdrawn in deeper soil layers and thus are not directly affected by the heat. However, after burning large amounts of biomass, the ash is dissolved in the first rains. This solution is very high in pH (up to 10). Such solution infiltrating the soil might kill certain species that usually live under rather acidic conditions under the previous forest cover (pH of approximately 4.0–4.5 in most tropical forests).

Weeds

Severe burns usually kill plant seeds even in deeper layers and can reduce the weed seed bank. Although burning is believed to reduce weed pressure, it can facilitate weed invasion. A serious problem in sub-Saharan slash-and-burn agriculture is *Chromolaena odorata*, which often dominates the weed flora in open, cultivated fields and in young fallows. *C. odorata* seeds have higher germination rates when they are exposed to light on the soil surface. Thus, airborne seeds can colonize newly burned fields. On the contrary, where slash is retained, seeds fall onto the slash, which dries out quickly, preventing germination. Once established, *C. odorata* plants can survive burning as their woody stumps may remain unaffected in the upper soil layers. Generally burns have little damaging effects on species regenerating from rootstocks or rhizomes. Often such species are favored in systems exposed to fire because burning destroys seedlings of other species and thus can reduce competition from other weeds. *C. odorata* and *Imperata cylindrica* are typical examples of species being favored by regular fire.

Other Plant Species

Slash-and-burn agriculture in a tropical dry deciduous forest in Mexico eliminated 29% of resprouting species, particularly those that were present at low abundance. This reduction was exacerbated by repeated burning.

Vertebrates

The effect at field establishment and specifically at burning depends on the species and its mobility. Although mammals may escape the slash and the burn by moving into surrounding unaffected vegetation, many reptiles and amphibians are killed when forest is cleared and burned, as they cannot escape quickly. Depending on the season and the temporal pattern of activity, animals that hibernate through the dry season may escape the fire if burning happens before larger rains have fallen. Where vertebrate and particularly mammal and bird densities decrease after slash-and-burn agriculture, the potential for fallow recovery may be reduced because many vertebrates are responsible for seed dispersal.

Effects during the Cultivation Phase

In traditional shifting cultivation with long fallows, the cropping phase lasts 1–4 years. As the length of cultivation increases, the potential of land to recover to its initial vegetation is reduced. Operations such as tillage and weeding frequently disturb the soil, and the viability of seeds of the initial vegetation is reduced. Regrowth from stumps is slashed more frequently, increasing the risk of dieback. There are changes in microclimatic and edaphic conditions that may reduce the probability of survival of species of the initial vegetation. In addition, the cropping phases are windows for weed invasion, and a return of the initial vegetation may be severely impeded by species such as *I. cylindrica* and *C. odorata*.

Tillage

The traditional shifting cultivator does not use many tools to work the soil. Most crops are planted with a simple planting stick or by opening small planting holes with a machete or hoe. Most crops, such as cassava, tannia, taro, plantain, melon, and maize, do not require tillage to succeed. If slash and burn is practiced in short fallow systems in degraded land, crops such as melon and plantain are not grown as they are considered not to produce well, whereas other crops can perform well but require tillage. The soil is tilled either to mix the ash

with the soil, to bring the seed to the required planting depth, or to concentrate nutrient-rich topsoil around the crops (mounding). Most intensive soil disturbance is caused by tillage for crops such as yams (*Dioscorea* spp.), which are mounded; peanuts (*A. hypogaea*), for which the soil is tilled at planting and a second time before pegging; potatoes (*Solanum tuberosum*), and sweet potato (*I. batatas*), for which ridges or small mounds are made. Farmers throughout have made their own observations and know which crops respond positively to tillage. However, tillage is a labor-intensive operation and is more frequently used where labor availability is high and soil fertility is low.

Tillage disrupts the natural layering of the soil and breaks up soil aggregates. Naturally formed pores (earthworm channels, termite and ant galleries, and root channels) are interrupted and often are filled with loose fine material. Soil fauna dwelling in such pores will be negatively affected by tillage, as their habitats are destroyed or the number of potential niches is reduced. Tillage changes the water regime in the topsoil and might lead to a more rapid drying out of the top layer, causing stronger moisture and temperature fluctuations. Soil mites and springtails respond to such fluctuations by withdrawing into deeper layers, which may be less rich in substrate and thus do not permit the same level of activity and population densities.

When land is cleared, slashed, burned, and tilled, earthworm density, diversity, and activity are reduced. This has been shown in south-west Nigeria, where the earthworm fauna is dominated by *Hyperiodrilus africanus* and the epigeic *Eudrilus eugeniae*, in the Peruvian Amazon, in southern Cameroon, and in south-east Mexico. Earthworms are classified functionally into epigeic, anecic, and endogeic categories. Epigeics live in the litter layer. Anecics feed on litter and soil, dragging litter into their vertical burrows. Endogeics live in the soil and feed on soil organic matter and dead roots. Anecic and epigeic are the groups most immediately affected by slash, burn, and tillage due to the loss of current and future substrates and physical habitat disruption. These groups will take longer than endogeics to recolonize the soil and litter layer under the recovering fallow.

Tillage affects the seed bank in the soil. This is less important in long fallow slash-and-burn systems as seeds will normally be on the soil surface where they are most likely to be killed by the burn. In younger fallow, with light burns, not all seeds on the soil surface may be killed. Where the land has been previously cultivated with tillage or any soil-moving operation has been conducted (harvest of root and tuber crops), a seed bank consisting of the buried portion of seeds protected from the effect of subsequent burns already exists in the soil. Buried seeds can become dormant and survive for several years. Tillage after the burn will move such seeds closer to the soil surface, increasing their chance of germination, whereas other, younger ones may be buried and become dormant.

Weeding

Most weeding operations are conducted with the same or similar tools as tillage. Therefore, in addition to the tillage, weeding affects the soil, soil inhabitants, and the weed seed bank. It is difficult to separate the terms "seed bank" and "weed seed bank," as the same species may be both: a weed in the cropping phase and a desired plant in the fallow phase of a slash-and-burn rotation (cycle). Moreover, many slash-and-burn farmers tend volunteers that are considered food or vegetables, and they retain these species often until the seeds are mature, assuring the same species being present in a following cropping phase. To simplify, all nonplanted species will be treated as weeds in the cropping phase, except for the desired volunteers. The weeding of a field, established after forest clearing, will eliminate seedlings and reduce stump regrowth of forest species. Depending on the cutting height and the frequency of weeding, these species may eventually die. When the land reverts to fallow, these species will be unable to reestablish immediately. The fallow may thus consist only of weeds from the cropping phase.

Pest and Disease Suppression

Certain pests and diseases of crops may be avoided by choosing land that was under fallow for a long time. The efficiency of pest and disease avoidance depends on the distances between fields, the mobility of pests, and the vectors by which pests and diseases are spread. Soil-borne diseases might be affected by the dramatic soil chemical changes after the burn and the associated increase in soil pH and the changed microclimate. One important mechanism in pest and disease suppression, not directly related to "slash and burn," is the complex mixture of crop species in slash-and-burn fields. Up to 40 species might be planted or tended within a single plot. Pest and disease buildup is not possible or is strongly delayed if the organisms are confined to a few isolated plants. The large number of pests and diseases with their different ecological requirements does not permit a general statement of the effects of slash-and-burn agriculture on their presence and severity in crops.

Radopholus similis is a plant-parasitic nematode, cosmopolitan in the tropics. It is the greatest cause of yield loss in bananas and plantains worldwide. Peanut and maize are also hosts, as are *Commelina benghalensis* and *Fleurya aestuans*, minor components of weed and fallow communities in West and Central Africa. *R. similis* does not survive in the soil for more than 6 months when host roots and corm pieces are absent. Thus, even a single-year fallow can prevent infection of the following crop, if *C. benghalensis* and *F. aestuans* are uprooted and the new planting material is pest free. The Maring people of Papua New Guinea plant banana suckers into their field during the burn. Heat would penetrate the outer layers of the banana sucker and would be likely to kill nematodes infecting the planting material. Other peoples roll banana suckers in ash before planting, this is believed to both kill nematodes and have a fertilizing effect.

Certain crop diseases are affected by the degree of shade in the plot. For example, the fungus *Mycosphaerella fijiensis*, the causal agent of black sigatoka disease of plantains and bananas causes less damage to bananas when they are grown under shade trees than when they are in full light. *Mycosphaerella musicola* is the less virulent yellow sigatoka, dominant at higher altitudes. Workers in the Caribbean, Central America, and Africa have noted that when trees are eliminated from slash-and-burn fields, these diseases destroy larger areas of the leaves and yields are reduced by up to 40%.

Nowadays, short fallow slash-and-burn farmers have a limited choice of land, but in some areas, they have access to pesticides. With a broad range of crops and more recently intensive vegetable production, the use of pesticides has become more common in many developing countries. Slash-and-burn farmers, although as yet rarely, do increasingly resort to pesticide use. Although any pesticide may be targeting a specific pest or disease, it usually affects other organisms as well and indirectly organisms relying on the affected nontarget organisms. With often little knowledge of the effects of pesticides on slash-and-burn agriculture, cultivators may destroy organisms, important to the ecosystem's functioning. Thurston (1992) provides a very comprehensive description of pest and disease control in such low-input cropping systems.

Management of Crop and Weed Residues

Some effects of slash-and-burn land preparation are aggravated by crop management. In natural vegetation, in both long and short fallows, litter will accumulate on the soil surface, contributing to reduced soil temperature and providing nutrient and substrate inputs. During the cropping phase, some crops produce litter, and uprooted or cut weeds could be retained as mulch. Yet mulching or simple retention of such residues is not a common practice amongst slash-and-burn farmers. In most cases, litter and weeds are removed, thrown outside the field, or piled on tree stumps. In the Congo basin forest zone, tree stumps with weed residues are often covered on purpose to prevent the stump resprouting or to accelerate rotting of the stump. Retaining residues may create niches in which some of the fauna may survive the cropping phase. Yet slash-and-burn farmers are aware that mulch can also provide potential niches for pests.

Nutrient Export through Crop Harvesting

In addition to unproductive nutrient losses (combustion, wind, leaching, erosion, or transfer to unavailable forms), the harvested crops contain nutrients that are exported from the field. As slash-and-burn crop yields are usually low, crop export losses are considered marginal compared to other losses, especially at burning of old fallows. However, in short fallow systems, which produce less biomass and thus have a lower nutrient input from the fallow, with more frequent or prolonged cropping phases, the crop nutrient export losses will become more important in the overall nutrient balance of a field.

Effects of Prolonged Cropping Phases

Prolonged cropping phases permit the invasion of weeds suited for conditions in a clearing. Such weed species die out if the field is abandoned after a short cropping phase and if the previous vegetation establishes rapidly. However, if the cropping phases are prolonged, these weeds have more time to establish, produce seeds, or propagate by vegetative means such as rhizomes. Consequently, competition between crops and these invasive weeds increases, and once farmers no longer consider cropping a suitable option, these weeds may be so abundant that they reduce or even eliminate the chances of recovery of the initial vegetation. Often the ability of the initial vegetation to return is further reduced by repeated burning. Repeated burns kill tree stumps and seeds and can select for grasses (Poaceae), which may initially establish from seeds but often propagate by rhizomes in the soil and thus are less affected than other species by the burn. The result is that the plot may revert to "arrested succession" grassland and be rendered unsuitable for cropping in the medium term. This may be a goal of a farmer who wishes to create pasture or may be an unintended consequence. In south-east Asia, prolonged cropping after forest clearing has facilitated the invasion of the land by *I. cylindrica*.

Generally, climax forest species have seeds that remain viable for a short time. Thus, extended cropping phases will compromise the regeneration potential of the forest. However, where the intended purpose is to repeatedly crop after short fallow phases, the absence of viable seeds of forest species is academic, as they would have no possibility to establish on that site.

Effects during the Fallow Phase

The effects of slash and burn on biodiversity in the fallow phase are strongly related to the conditions prevailing during the cropping phase. In long fallow systems with short cropping phases, the transition into a fallow dominated by pioneer trees is fast, without a phase in which arable weeds dominate. In the highlands of Papua New Guinea, the Maring people tend the resprouting trunks of the pioneer tree rokunt (*Saurauia* sp.) in their fields near the end of the cultivation phase, even though these compete with their crops. Seedlings of *Casuarina oligodon*, a nitrogen-fixing tree, are transplanted to the field at this time to enhance soil fertility recovery during the fallow. Such reestablishment of the fallow vegetation will probably ensure the reestablishment of the soil micro-, meso-, and macrofauna. With an increased length of the cropping phase, fields tend to get weedy as species not existing in the previous fallow might have established in the cultivation phase. These are likely to form at least part of the first fallow succession community. In such a situation, the establishment of a fallow vegetation as before the slash-and-burn cycle is delayed, and in some cases impossible.

Biomass production and nutrient uptake and cycling are the primary processes of soil fertility restoration. The amounts of biomass and nutrients accumulated are a function of the fallow length. The minimum fallow length to completely restore soil fertility (**Figure 1**, case 2) will vary depending on soil type, the types of weeds, crop pests and diseases present, and the growth rates, composition, and succession of the fallow species.

Aboveground biomass accumulation rates of forest fallows vary between sites and may vary between fallow ages. Most data from 4-year-old fallows range from 24 to 56 Mg ha^{-1}, for 8-year-old fallows from 40 to 150 Mg ha^{-1}, and for older fallows up to 500 Mg ha^{-1}. Biomass accumulates for at least the first eight years. Some authors have reported that half the biomass in a 20-year-old fallow had already accumulated by the fourth year. In contrast, others show that accumulation

rates of approximately 5.5 Mg ha^{-1} year^{-1} remained constant from the 4th to the 15th year.

Generally, soil organic matter levels increase with time under fallow. During 4 years of bush fallow in south-west Nigeria, soil organic carbon content increased by 46% and total nitrogen by 59%. A 15–20-year-old fallow had soil organic matter levels of 3.8% compared with 1.1% in cultivated plots. In the drier region of north-east Brazil, a 10-year-old fallow was sufficient to restore soil carbon and nitrogen levels to those under undisturbed vegetation. Many authors have reported that soil organic matter and soil nitrogen concentrations do not increase substantially after 8–10 years of fallow.

Occasionally, farmers use indicators to judge when to clear and crop a fallow. In south-west Cameroon, farmers stated that there is a definite minimum fallow time to regain fertility, expressed in years but judged by the type and maturity of the vegetation growing there. The Turumbu of Yangambi in the Democratic Republic of Congo (formerly Zaire) judged when to break fallow by the girth of umbrella trees (*Musanga cecropioides*) and the biomass of the undergrowth. They cleared when they could walk easily beneath the trees. These secondary considerations coincided with soil fertility regeneration (Jurion and Henry, 1969). In south-west Nigeria, farmers claim to use the appearance of the vegetation cover and the presence of particular plant species and of earthworm casts to decide whether a fallow is ready for cropping.

In Papua New Guinea, bird and reptile species diversities were low in young fallows, but increased as succession proceeded. Reptiles were less affected than birds. The bird communities in young, grassy fallows were dominated by obligate granivores, whereas the forest supported specialist feeders (frugivores, nectarivores, and branch gleaners). Butterfly species richness showed a similar pattern.

In southern Cameroon, species richness of birds in *C. odorata* fallow was only a fifth of that in secondary or primary forest. Butterflies and termites had similar magnitudes of declines. Canopy ants and canopy beetles were completely absent in such fallows, given the near-absence of trees. The species richness of flying beetles, leaf-litter ants, and soil nematodes was not different between forest and young fallow. Presumably, the young fallow, located within a mosaic of recovering fallows and secondary forest, provided sufficient food and suitable conditions for their continued survival. For more details, see Lawton *et al*. (1998).

In southern Cameroon, earthworm activity, indexed by surface cast production, declined rapidly after forest clearance, burning, and cropping and did not recover during two years of fallow. At the same site, cast production decreased after clearing, burning, and cropping a 4-year-old *C. odorata* bush fallow, but did recover during the fallow phase. The change in living conditions from forest to cropland was probably too drastic for the existing earthworm community to maintain its activity level. A new community structure comprising species capable to cope with the new conditions could not establish within 2 years of fallow. On the contrary, the community in bush fallow had probably already undergone some shifts, thus recovering as soon as bush conditions were reestablished. The response of termites was different. In both forest and bush clearings, termite densities dropped to approximately 1% of densities in the retained forest. There was no recovery during the fallow phase. The number of species in the cleared area was only half as high as in the forest.

The Scale Dependence of the Effects Caused by Slash-and-Burn Agriculture

Topographic Preferences and Fragmentation

Slash-and-burn farmers have topographic preferences. Long fallow farmers often prefer gently sloping land as this facilitates the burn. They avoid slopes that are too steep as this increases the risk of soil erosion. Yet their choice is limited by the type of terrain available. In other parts of the tropics, farmers prefer the plateaus and avoid cropping even on the gentle slopes. Valleys fringes, although difficult to access and waterlogged during the rainy season, are under greater pressure from farmers for dry season cropping of high-value market crops. In contrast, riverine forests and swamps are avoided and therefore can remain undisturbed.

Such preferences result in fragmentation. Ultimately, in areas of high population density, fragments of the previously contiguous vegetation will remain only in areas unsuitable for agriculture. Not all species of plants and animals are represented in such fragments of vegetation, and thus a recovery of the surrounding to the initial condition is not possible. However, if slash and burn occurs in particular topographic positions in low population areas where systems have sufficiently long fallow to restore the forest, it does not have major effects on biodiversity.

Temporal and Spatial Scales

Traditional systems at low population densities, with small fields (0.25–3 ha) in 1-year cropping phases and long fallows of 20 years, use 5% of the land at any given time (**Figure 2**). Fields are surrounded by fallow land at various stages of recovery. Even if the fields are cropped for up to four years, not more than 20% of the land will be cultivated. Recolonization of a cleared plot by species from the surrounding fallow occurs rapidly after farmers stop weeding or abandon the land. The rate of recolonization by animals depends on the species' mobility. Plant species' invasion or recolonization of the site depends on the species' type of seed and the method of seed dispersion. Unless species are extremely immobile or, in the case of plants, if the fallow length is too short for species to reach the reproductive phase, this system can maintain the initial biodiversity.

With increasing human population densities, slash-and-burn agriculture is practiced in shortened bush or grass fallow cycles, with prolonged cropping phases. Both processes lead inevitably to a larger proportion of land being under cultivation. This occurs rapidly when fallows are shorter than 10 years in systems cropping for more than 1 year and when fallows are shorter than 5 years in systems cropping for 1 year (**Figure 2**).

Topographic preferences and an increasing proportion of land cropped at any given time lead to larger scale fragmentation, with areas of the initial or recovering vegetation isolated by larger areas of cropped land and early stages of fallow. Spatial fragmentation can compromise the following: (1) the

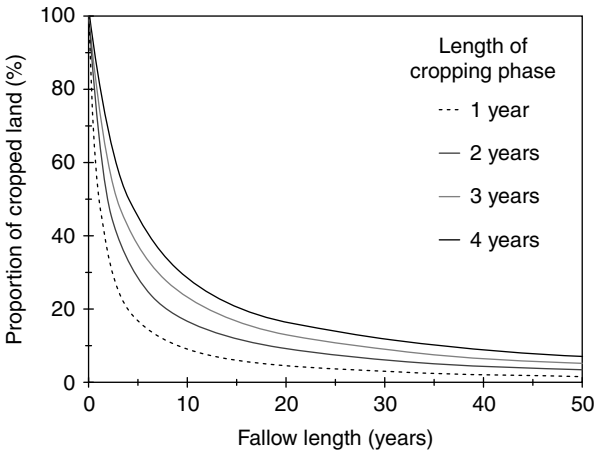


Figure 2 Proportion of cultivated land at 1–4 years of cropping phases, as a function of fallow length.

ability of the remaining undisturbed or recovering habitat fragments to maintain diversity, which determines the potential to recolonize the disturbed habitats, and (2) the ability of species to be dispersed to and establish in disturbed areas. Both aspects depend on the species investigated. Generally, with increasing size of an animal species, spatial requirements for foraging and hunting increase, and thus their chances of survival diminish with increasing fragmentation. A number of additional factors, such as mating habits and the ability to cross disturbed areas to reach other undisturbed fragments, will modify the likeliness of species' maintenance. Smaller animals, largely invertebrates, might not be affected as severely by fragmentation as larger animals. On isolated trees retained in pastures in the Amazon, a considerable diversity of ants was found, with a large proportion of species usually found in the forest. Ant diversity was positively related to the epiphyte load of trees.

Fragmentation of forest by pasture land such as in the Amazon basin negatively affects certain tree species retained in the forest fragments. Some species need cross-pollination to produce seeds. If fragments become too small or too distant from each other, the tree density might fall below a threshold required for pollen to reach other trees. Such a process, observed in mahogany (*Swietenia macrophylla*), reduces or eliminates seed production and may lead to local extinction.

The potential to regenerate degraded land to the initial vegetation and habitat depends on seed dispersion and mobility of animals. Seed dispersion by birds, bats, and monkeys has been shown to be important in disturbed and undisturbed areas. However, in Uganda, it was shown that bats and birds did not considerably contribute to the dispersal of forest species' seeds to degraded grassland areas. The vast majority of seeds moved by these animals belonged to species dominantly found in disturbed areas or to species unable to establish in the grassland. Similarly, seed dispersion by gibbons confined to small forest fragments after large-scale slash-and-burn agriculture in Kalimantan was found to be uncertain as individuals were extremely hesitant or unable to cross cleared areas to reach other fragments.

The effects of slash-and-burn agriculture on biodiversity are various. However, generally, the shorter the cropping phase and the longer the fallow phase, and the smaller the proportion of cropped land, the greater the possibility of recovery of the previous vegetation. However, these criteria limit the productivity of slash-and-burn systems. To maintain their livelihoods, the choice for many farmers whose land allocation is fixed is either to prolong the cropping phase or to shorten the fallow phase. Clearly, research should also compare and contrast the effects of these options on biodiversity, rather than focusing exclusively on comparisons with undisturbed ecosystems that are increasingly unrepresentative.

Alternative Systems to Slash-and-Burn Agriculture

Plantations of Perennial Crops and Trees

Plantation systems include monocrop oil palm, fast-growing timber species, coconut, rubber, tea, coffee, and cocoa. Most plantation systems are established by slash and burn. Timber plantations are often established in forests by selective felling or poisoning of undesired trees, underbrushing, and sometimes light burning of the understory. In such systems, the forest habitat is only slightly altered and reestablishes quickly. In southern Cameroon, timber plantations had similar levels of termite species richness to secondary forests and more species of leaf-litter ants than both secondary and near-primary forests (Lawton *et al.*, 1998). However, the dominance of one tree species may affect the underbrush regrowth and the conditions for soil fauna through a narrower range of substrate. Many plantations, such as oil palm, coconut, rubber, tea, and other species, which demand full sunlight, are planted after clear-cut felling. Depending on the intensity of cropping, these plantations may be frequently weeded and treated with pesticides. Under such conditions, biodiversity may be greatly reduced. In contrast, cocoa plantations in southern Bahia, Brazil, are established in old secondary or primary forest. Cocoa is shade tolerant, so although the lower-canopy trees and herbaceous components are slashed, many upper-canopy trees are retained. These include many valuable timber species, rosewood (*Dalbergia nigra*), pau Brasil (*Caesalapinia esplinata*), and cedro (*Cedrela odorata*), that have been nearly eliminated elsewhere by loggers. Consequently, associated fauna have been maintained, including rare primate species, for example, the sagui (*Callithrix kuhlii*) and the lion tamarin (*Leontopithecus rosalia*).

Pastures

In large parts of the Amazon basin and in Central America, forest is clear cut to establish pastures. Although in tree and palm plantation systems, a forest-like microclimate will be reestablished after some years, pastures will remain unshaded. Any forest regrowth is regularly slashed, and in some situations, pastures are regularly burned to reduce forest regrowth and promote grasses.

Improved Short Fallow Systems

Slash-and-burn agriculture is demanding on the natural resource base, irrespective of the fallow type (forest or short

fallow). It requires large areas for cultivation as yields are low. In many situations, the natural regrowth in short fallow systems is not capable of restoring soil fertility because of the shortened period of time available in which the volunteer regrowth cannot produce sufficient biomass and accumulate sufficient amounts of nutrients. Improved fallow systems seek to replace the, usually undesired, volunteer regrowth by species that produce more biomass faster, contain more nutrients, or produce a type of biomass that is of a different physical structure than the natural regrowth. Such species are expected to out-compete the weed flora quickly and not to contribute to propagation or maintenance of crop pests and diseases. Often legumes, fixing atmospheric nitrogen, are used to augment the nitrogen resources of the soil.

Improved fallow can be based on tree, shrub, or herbaceous species. Depending on the fallow type used for improved fallows, the biomass is either slashed and burned or only slashed and retained as mulch. Tree- and shrub-based systems are likely to be cleared and managed in the same way as older natural fallows because the large proportion of wood in the slash can be removed only in a labor-efficient way by burning it. The extensively researched alley cropping system in which trees were pruned and the prunings were mulched has not been accepted by farmers, mainly due to its labor requirements and poor crop yield responses in farmers' fields (Hauser *et al.*, 2006). Systems using herbaceous legumes, such as *Mucuna pruriens* or *Pueraria phaseoloides*, reduce the number of plant species surviving in the fallow, and *P. phaseoloides* has been shown to reduce plant parasitic nematode densities. Depending on the crop planted after herbaceous fallow, the slashed mulch can be retained as a mulch layer, retaining soil fertility, protecting the soil from compaction and erosion, and rendering good living conditions with minimal disturbance to soil invertebrates. Although such systems will not maintain biodiversity similar to that in forest, the increased crop productivity (yield increases up to 100% for maize) and the option to crop the land at higher frequencies than in natural fallow cycles (*Pueraria* 1–2 years of fallow; *Mucuna* often annual cropping with 6–8 months of fallow) allow farmers to remain on the same land and, theoretically, reduce their need to clear forest or older fallow, although social customs to gain title to land may dictate otherwise. With less land used for agricultural production, more land may be retained in undisturbed condition and provide for the long-term retention of biodiversity.

Slash-and-Burn Agriculture and Carbon Sequestration

The environmental credentials of slash-and-burn systems have been questioned globally as vegetation burning releases stored carbon to the atmosphere as carbon dioxide. The global average aboveground biomass of secondary humid forests is approximately 450 Mg ha^{-1}, although there is much variation; thus, assuming a complete burn, up to 200 Mg ha^{-1} C can be released to the atmosphere, equivalent to approximately 730 Mg ha^{-1} of CO_2. In addition, carbon dioxide release from the soil may occur as exposed soil organic matter and decaying roots decompose. The quantity of CO_2 released will depend on the initial forest biomass and climatic conditions governing decomposition during phases after the burn. Yet given that many slash-and-burn systems are no till or minimal till, soil carbon losses through accelerated decomposition rates of soil organic matter may be much less than in mechanized systems.

Furthermore, food crops will assimilate less carbon from the atmosphere than the initial vegetation would have assimilated, thus reducing the C-fixation in biomass and contributing to a larger release of carbon in the cropped plots during the cropping phase. Carbon sequestration during the cropping phases is low. While maize in slash-and-burn systems produces 4–6 Mg ha^{-1} of aboveground biomass within 3–4 months, 50–75% is straw, which remains on the field and will decompose in the following months, and the maize cobs are removed from the field and consumed. Cassava may produce between 20 and 40 Mg ha^{-1} total biomass (aboveground and belowground combined), yet approximately 50% of that biomass are roots that will be removed. Thus, cropping phases generally do not sequester significant amounts of carbon.

Data on biomass accumulation during fallow phases are rare and highly dependent on the fallow vegetation. In southern Cameroon, the dominant short fallow species *C. odorata* produced approximately 11 Mg ha^{-1} of aboveground carbon including litter within a 2-year fallow period. Two shrub fallow systems using the species *Flemingia macrophylla* and *Dactyladenia barteri* produced approximately 13 Mg ha^{-1} aboveground carbon including volunteer species and litter. A tree fallow system with *Senna spectabilis* produced 26.5 Mg ha^{-1} aboveground carbon including volunteer species and litter. Thus, the fallow type has a considerable effect on the amount of C that is sequestered. Even if all aboveground fallow carbon is burned, the time-average C stocks are higher in the tree system than in any other. Considering aboveground carbon accumulation in crops is 10 Mg ha^{-1} year^{-1} and carbon accumulation in the fallow phase is linear, the tree system holds, on average, 16.5 Mg ha^{-1} carbon per year versus 8.8 Mg ha^{-1} in the natural fallow system. During longer fallow phases, carbon accumulation at Yangambi (Democratic Republic of Congo) was 40 and 61 Mg ha^{-1} in 5- and 8-year-old fallows, respectively. Carbon accumulation curves tend to be sigmoidal with initially low rates in the first 2 years, followed by higher rates, and then level off as stands age.

When calculating "ecological footprints" of slash-and-burn agriculture, the carbon dynamic of all fallow stages in the slash-and-burn cycle needs to be considered. While carbon is released in the newly cut plot and minimal carbon storage occurs during the cropping phase, carbon assimilation occurs in the various fallow phases, so the net impact of the system is likely to be a small loss or in equilibrium, assuming that fields remain in the long fallow cycle. Furthermore, slash and burn is conducted predominantly by people who make negligible demands on global fossil energy resources.

Modern systems integrating short rotation forestry with agricultural production may offer reasonable C sequestration rates combined with food production and the option to replace fossil fuels. In Cote d'Ivoire, Benin, and the Democratic Republic of Congo, acacia plantations were established for timber or charcoal production. Such plantations are cut after 8–9 years at an estimated 80–100 Mg ha^{-1} standing tree biomass (not including litter and roots) or 36–45 Mg ha^{-1} of C, which is equivalent to a time-average C stock of 19–24 Mg ha^{-1} year^{-1}. The trees are either sold as building

material or converted to charcoal. The residues are burned *in situ,* and the land is cropped and replanted to trees for the next cycle. Crop yields were reported to be higher after such plantations than after similar aged natural fallow, dominated by grasses, in Benin and Congo, yet not in Cote d'Ivoire. If trees are used for construction, the C will be sequestered. Charcoal production and use will ultimately return all sequestered C to the atmosphere, yet replaces at least partially the use of fossil fuel. Thus, slash-and-burn systems can be designed to hold substantial amounts of C and to produce materials that will retain C over long phases, yet farmers' food production and income generation needs and targets need to be considered to make these systems more attractive.

See also: Agriculture, Sustainable. Agriculture, Traditional. Carbon Cycle. Deforestation and Land Clearing. Fires, Ecological Effects of. Human Impacts on Ecosystems: An Overview. Indigenous Peoples and Biodiversity. Poverty and Biodiversity. Rainforest Loss and Change. Soil Biota, Soil Systems, and Processes

References

Guillemin R (1956) Evolution de l'agriculture autochtone dans les savanes de l'Oubangui. *Agronomie Tropicale* 11: 143–176.

Hauser S, Nolte C, and Carsky RJ (2006) What role can planted fallows play in humid and sub-humid West Africa? *Nutrient Cycling in Agroecosystems* 76: 297–318.

Jurion F and Henry J (1969) *Can Primitive Agriculture be Modernised?* London, UK: Agra Europe. (translation from French).

Lawton JH, Bignell DE, Bolton B, *et al.* (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature* 391: 72–76.

Nyerges AE (1989) Coppice swidden fallows in tropical deciduous forest: Biological, technological, and sociocultural determinants of secondary forest successions. *Human Ecology* 17: 379–400.

Peters WJ and Neuenschwander LF (1988) *Slash and Burn. Farming in the Third World Forest.* Moscow, ID, USA: University of Idaho Press.

Thurston HD (1992) *Sustainable Practices for Plant Disease Management in Traditional Farming Systems*, 279 pp. Boulder, CO: Westview Press.

Social and Cultural Factors

Jeffrey A McNeely, The International Union for Conservation of Nature (IUCN), Gland, Switzerland

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adaptation Any behavioral pattern that makes a human or a cultural group more fit to survive and reproduce in comparison with other individuals or cultures.

Carrying capacity The maximum population that can be sustained indefinitely in a given area without changing the ecosystem in ways that will eventually reduce the sustainable population. This balance between population and resources is a dynamic one that involves changes in technology and other factors.

Community A social unit consisting of members who are in direct interaction with each other and who have a collective identity, however defined. Relationships in such a group are principally primary rather than secondary in nature, and conformity to group norms is achieved mainly by peer pressure.

Culture All capabilities and habits acquired by people as members of society.

Indigenous peoples Those groups of human beings who share and preserve direct, everyday connections to their distinguishable distinctive cultural roots (even though they may have willingly migrated or been forcefully moved from their homelands, including moves to cities). These indigenous peoples consciously or unconsciously draw on specific knowledge and strategies developed and tested by past generations to address current problems. Cultural traditions emerge and are maintained in a dynamic process of creative invention and reinvention, as well as borrowing and adaptation from other subgroups and cultures.

Language A system of conventional spoken or written symbols that enable human beings to communicate as members of a social group and participants in its culture.

Management The efforts of humans to select, plan, organize, and implement programs designed to achieve specified goals; activities can range from protective measures to ensure that human influences on natural resources are minimized to greater interventions required to maintain diversity, install facilities, control populations, and eliminate unwanted elements.

Myth A story, usually of unspecified origin and at least partially traditional, that ostensibly relates actual events to explain some practice, belief, institution, or natural phenomenon and that is frequently associated with religious beliefs; myths provide models for human behavior and often emphasize behaviors that support conservation of biodiversity.

Religion The relationship between people and that which they regard as holy, often in supernatural terms. Religions include a wide range of ceremonies and ritual practices aimed at supporting the moral and ethical values of a society and maintaining natural or built sacred spaces where such ceremonies and practices can take place.

Religion and Biodiversity

Religions order the spiritual and physical relationships of people with other humans and with their environment. Approaches to conserving biodiversity that are based on cultural and religious values are often much more sustainable than those based only on legislation or regulation. While religions offer important guidance on the relationship between people and the rest of biodiversity, all religions are characterized by a wide gap between their philosophy and the practices of the people who have accepted the philosophy.

Religion frequently has had the effect of preventing excess human demands from outstripping the environmental resources that are required to support them. For example, in Roman times, high priests dictated that the well-being of humanity depended on thousands of animals being sacrificed every year, thereby having the effect of keeping the size of herds within the carrying capacity of the grazing lands. More fundamentally, religions provide a complex holistic view of how to use natural resources, supporting this knowledge with an ethical perspective that is built on the implicit social contracts that enable communities to function.

All of the world's major religions are today sensitive to the importance of biodiversity, though of course their historical writings do not use today's conservation vocabulary. The following is a brief summary of some major religious belief systems and how they relate to modern biodiversity concerns.

Animism is a term applied to a wide variety of beliefs that focus on numerous spiritual beings concerned with human affairs and capable of either helping or harming people. Animism is a view of the world found among many traditional peoples who conceptualize an intrinsic spiritual connection between humans and nature. This connection has been described by Harvard University biologist Edward O. Wilson as "biophilia," an innate human need for contact with a diversity of life-forms. For example, in the North American desert, the Yaqui Indians believe that close communication exists among the plants, animals, birds, fish, rocks, and springs that inhabit the Sonoran Desert. Many traditional approaches to conservation are supported by religious beliefs, often based on various kinds of animism that have the effect of fostering respect for plants and animals. In many parts of the world, people have established sacred sites on the basis of inherent spiritual or religious significance. Such sacred sites, based ultimately on animistic beliefs, are often sanctuaries for biodiversity. They may well survive substantial cultural changes; for example, the Parliament of Kenya (a very modern institution) voted in 1990 to protect all of the country's remaining

sacred forests, known as *kaya*. Many animistic systems of belief are accompanied by the idea of *taboo*, that which is forbidden; breaking a taboo can bring sanctions such as illness, social ostracism, or even death. Taboo often applies to certain sets of natural resources that are particularly vulnerable to overexploitation. Animism is also often associated with *totemism*, a complex of ideas and practices based on the belief in a mystical relationship (often kinship) between people and certain animals or plants; these relationships often include reverential and genealogical relationships between social groups or individuals and the totems. Totems normally are associated with taboos of avoidance, so an Amazonian Indian hunter within a social group that has the peccary as a totem may be forbidden from hunting peccaries. Thus totemism also helps to restrict exploitation of harvestable resources.

Buddhism, with a total of about 300 million practitioners found in many Asian countries, teaches that a behavior has a natural relationship to its resulting consequences in the physical world. All Buddhist teachings and practice come under the heading of *Dharma*, which means truth and the path to truth. Based on the teachings of the Buddha, who lived in Nepal more than 2500 years ago, Buddhism teaches that people are responsible for their actions and go through a cycle of rebirths before finally reaching *nirvana*. Right actions lead to progress toward *nirvana* while negative actions, such as killing animals, lead to regression from that goal. Committed to the ideal of nonviolence, Buddhism also attaches great importance to wildlife and the protection of biodiversity (though of course not using such modern terms in its ancient teachings). Respect for life in the natural world is essential, and by living simply one can be in harmony with other creatures and learn to appreciate the interconnectedness of all that lives. This is not to claim that nature is unchanging; on the contrary, Buddhism recognizes change as the very essence of nature. The Buddha taught that all things are interrelated and do not have an autonomous existence, so the health of the whole is inseparably linked with the health of the parts and the health of the parts is inseparably linked with the whole.

Christianity, with some 1.6 billion members, is the dominant religion in Europe, sub-Saharan Africa, the Pacific, and the western hemisphere. It has numerous denominations, all based on the belief in a single all-powerful God who created nothing unnecessarily and omitted nothing that is necessary. Christians believe that Jesus Christ was the son of God and a great teacher whose message was that all of creation is the loving action of God, who continues to care for all aspects of existence. The very nature of biodiversity considered in itself, without regard to humanity's convenience or inconvenience, is considered to give glory to the Christian God. Christianity teaches in effect that humanity may not disorder biodiversity or destroy God's creations, at the risk of destroying itself; in the Holy Bible, Ecclesiastes 3:19 says: "For that which befalleth the sons of men befalleth beasts … as the one dieth, so dieth the other … a man hath no preeminence above a beast." Human acts of irresponsibility toward creatures are contrary to the divine wisdom, which sustains and gives purpose to the interdependent harmony of the universe. Modern Christian teachings are that humanity, both individually and collectively, should perceive the natural order as a sign and sacrament of God, recognizing that all creatures and objects have a unique place in God's creation. However, Christianity also recognizes a special role of humanity within creation. The Holy Bible contains Christianity's main teachings. It provides strong support to the protection of natural resources, including passages mandating the preservation of fruit trees (Deuteronomy 20:19, Genesis 19:23–25), agricultural lands (Leviticus 25:2–4), and wildlife (Deuteronomy 22:6–7, Genesis 9). The Bible often refers to the impressive intelligence of wild creatures, such as in Jeremiah 8:7–8, Proverbs 6:6–8 and 30:24–28, Numbers 22:22–35, and Isaiah 1:3.

Hinduism, the dominant religion in India with about 700 million followers, teaches the all-encompassing sovereignty of the divine, manifesting itself in a graded scale of evolution. While the human race is currently at the top of the evolutionary pyramid, it is not seen as something apart from the earth and its biodiversity. Hinduism is permeated by a reverence for life and an awareness that the great forces of nature – earth, sky, air, water, and fire – as well as biodiversity are all bound together within the great rhythms of nature. Hindus recognize that all lives play their fixed roles and function together so that no link in the chain of life is lost. The divine is not exterior to creation but expresses itself through natural phenomena. Hindus consider at least some forests and groves as sacred and associate various plants and animals with gods and goddesses in the Hindu pantheon. The most important teachings of Hinduism are contained in the Bhagavad Gita, a dialogue between Sri Krishna and Arjuna, which has a clear description of ecology as a cycle of life dependent on everything from bacteria to birds. Like animists, Hindus believe that all plants and animals have souls, and that people must do penance even for killing plants and animals for food. Hinduism involves sacrifice, forms of ritual worship that are designed to protect life through reinvigorating the powers that sustain the world by securing cosmic stability and social order. Hindus recognize certain rivers and mountains as sacred, because they give and sustain life.

Islam, with about a billion adherents, is the dominant religion in North Africa, the Middle East, and many Asian countries. Its teachings are very similar to those of Christianity in relation to biodiversity. The entire universe is God's creation; Allah makes the waters flow on the earth, upholds the heaven, makes the rain fall, keeps the boundaries between day and night, creates all of biodiversity, and gives it the means to multiply. *Tawheed* is the principle of Oneness of Allah, teaching that "there is no God but God and Mohammed is His messenger." This testifies to the unity of all creation and the interlocking grid of the natural order of which people intrinsically are parts. The *Holy Koran* provides a set of principles that define the relationship of man to God and of God to the environment in its totality. However, humans are considered a very special creation because they alone have the power to reason and think, giving them the potential to do great good, or great harm. For Islam, the role of people on earth is that of a *khalifa*, or trustee of God, so humans are entrusted with the safekeeping of earth and its biodiversity. People are answerable for their actions, including maintaining the unity of Allah's creation, the integrity of the earth, and its biodiversity. The Islamic relationship to biodiversity is that each generation uses and makes the best use of biological resources, without compromising the interests of future generations. The Koran

says, "With it we have produced diverse pairs of plants each separate from the others. Eat (for yourselves) and pasture your cattle; verily, in this are signs for men endowed with understanding" (Sura 20 Aya 53). While humans have the right to utilize and subjugate natural resources, this involves a commitment to conserve them both quantitatively and qualitatively. In the Koran, God says, "There is not an animal (that lives) on the Earth, nor a being that flies on its wings, but (forms part of) communities like you" (Sura 13 Aya 15). Thus Islam looks upon biodiversity as an expression of God's wisdom and omnipotence and as support for human development. This leads to the necessity of conserving and developing biodiversity for its own sake as well as for the benefit of humans. And the Prophet Mohammed said, "There is a reward in doing good to every living thing." Such beliefs can lead to biodiversity being conserved. For example, in Pakistan, representatives of original tree species still persist in old Moslem graveyards, because of a taboo against cutting these trees. And the only surviving population of a freshwater turtle, *Trionyx nigricans*, survives in a sacred pond dedicated to a Moslem saint in Bangladesh.

Jainism is one of the oldest living religions, beginning in India at least 2800 years ago. Jains believe that all living beings have an individual soul (*Jiva*), which occupies the body until it dies, then leaves the body and immediately takes birth in another. Attaining nirvana and thereby terminating this cycle of birth and death is the goal of Jain practice, much like Buddhism. Jainism is based on the principle of *ahimsa* (nonviolence) toward human beings and all of nature. Jains teach that no human quality is more subtle than nonviolence and no virtue greater than reverence for life. While biodiversity often is affected negatively by people, the intention to harm is what makes an action violent, and without violent thought no violent action is recognized. Jain cosmology recognizes the fundamental natural phenomenon of symbiosis or mutual dependence, with all aspects of nature belonging together and bound in a physical as well as a metaphysical relationship – an ancient perspective that is reflected in modern ideas about biodiversity.

Judaism, which originated in the Middle East but whose 18 million practitioners today are thinly spread around the world, teaches that God created the world, making order out of primal chaos. The sun, the moon, the stars, plants, animals, and ultimately humanity were each created with a rightful and necessary place in the universe. The encounter of God and man in nature is conceived in Judaism as a seamless web with man as the leader and custodian of the natural world. Judaism maintains that the earth is the arena that God created for man, half beast and half angel, to prove that he could behave as a moral being. Judaism views God as the divine "Giver of the Torah," which sets out a series of ethical obligations including many relevant to biodiversity. It teaches that the relationship between man and nature is one of ownership, though limited by good sense; man was commanded to behave toward the rest of biodiversity with justice and compassion. But humanity inevitably lives in tension between its power and the limits set by conscience. The Bible contains much environmental wisdom, for example, allowing agricultural fields to go fallow one year out of seven (Leviticus 25:1–5). The Bible also warns that even in time of war, it is forbidden to destroy fruit-bearing trees of one's enemies (Deuteronomy 20:19).

Shinto is the system of indigenous religious beliefs and practices of Japan, first appearing in written form around 1400 years ago, after Buddhism had been introduced into Japan. Shinto is based on beliefs concerning the nature and attributes of *Kami*, sacred power, which is found in all individual things. Shinto temples are often established on sites that have particular spiritual integrity and force, often with large groves of trees (totaling nearly 120,000 ha in Japan). Shinto is based on numerous sacred texts, collectively known as Shinten. Shinto is strongly based in rural agricultural practices, involving various ceremonies and festivals that guide the relationship between people and nature.

Sikhism, a modern religion that began in India in the late fifteenth century, builds on the message of the oneness of a universe created by an Almighty God, who is master of all forms in the universe and the source of the birth, life, and death of all beings. God has created the natural beauty, which exists and can be found in all of biodiversity. Without his *hukum* (order) nothing exists, changes, or develops, but having brought the world into being, God sustains, nourishes, and protects it. Biodiversity exists under God's command and with God's grace. Sikhism teaches against a life of conspicuous consumption, emphasizing mastery over the self rather than mastery over nature. A major value of Sikhs is to be in harmony with the earth and all creation. Obviously, people are moving farther in the opposite direction, but Sikhs argue that the current instability of the natural system of the earth is a reflection of the instability and pain within humans.

Taoism has a history of more than 2500 years and has been one of the main components of Chinese traditional culture. *Tao* means simply "the way" and is the origin of everything. The Tao took form in the being of the Grandmother Goddess, who came to earth to enlighten humanity. In Taoism, everything is composed of two opposite forces known as *yin* and *yang*. When the two forces reach harmony, the energy of life is created. Taoism judges affluence by the number of different species; thus a society with healthy levels of biodiversity is affluent, while societies with declining biodiversity are themselves in decline. Chinese Taoism, as reflected in the Tao Te Ching (*c.* 600 BC), calls for people to have a noninterventionist role in the natural world.

These short descriptions indicate that virtually all religions are very sensitive to biodiversity concerns, essentially asserting the interdependence of all life-forms and the need to treat all life with respect. But this religious perspective has not always been sufficient to conserve biodiversity. The special role of humanity within Christianity has often resulted in overexploitation of natural resources and alienation of land from people; the recognition of the Ganges as a river sacred to the Hindus has not prevented significant pollution and biodiversity loss at the hands of people; the unity maintained by balance advocated by Islam is today being replaced by environmental pressures that lead to disunity and imbalance; and the Tao noninterventionist perspective did not prevent the ancient Chinese from making profound changes to the land and water around them. Numerous indigenous cultures have been destroyed or fundamentally modified by missionaries from the major religions, who often are accompanied by technological innovations. But religions constantly evolve, and it is possible that a stronger religious response to

biodiversity loss will evolve as the concerns about such loss begin to affect people more profoundly.

Language and Biodiversity

This section draws from a contribution by Luisa Maffi. As an essential part of human cultures, languages provide the categories to conceive the natural and social world. A species of plant or animal that is important to a community gets named, and by learning that name, the people within the culture also learn what is vital to know in their natural environment. Thus language helps to organize the world and frees energy for other tasks, using words like pegs on which to hang the meanings stored in the storehouse of the human mind and providing a framework that binds together the details into a meaningful whole. Words for objects and phenomena help us learn the connotations, associations, emotions, and value judgments of a culture. Thus knowledge, beliefs, and values are linguistically encoded, thereby helping to promote a diversity of adaptive ideas and support various forms of biodiversity.

Many scientists argue that the preservation of the world's linguistic diversity (totaling more than 6000 languages) and the distinct forms of local knowledge that local languages encode must be incorporated as an essential goal in bioculturally oriented diversity conservation programs. Indeed, the diversity of human languages may be the best available indicator of human cultural diversity.

Many indigenous peoples believe that the external reality or environment is no different than the description of this reality or environment in a linguistic ecology. Variation in humans is adaptive, and language helps to support cultural diversity in humans. Human culture is a powerful adaptation tool and language both enables and conveys much cultural behavior.

Many countries with high biodiversity also have high cultural diversity, as indicated by their large number of languages. Languages have multiplied and thrived in places where natural selection has produced a rich variety of landscapes, animals, and plants. Researchers have found that linguistic diversity in sixteenth century North America was greatest in areas with the greatest diversity of habitats, irrespective of the latitude. Ten of the 12 "megadiversity countries" are among the top 25 countries for endemic languages. This correlation between biodiversity and linguistic diversity may have been fostered by a process of coevolution of small-scale human groups with their local ecosystems, in which humans interacted closely with the environment, modified it as they adapted to it, and acquired intimate knowledge of it. This knowledge was encoded and transmitted through language, supporting the contention of some linguists that life in a particular human environment depends on the ability of people to talk about it.

In many parts of the world, anthropologists have found a strong correlation between biodiversity, linguistic diversity, and ethnobiological knowledge. More generally, landscapes are anthropogenic not only in the sense that they are physically modified by human intervention, but also because they are symbolically brought into the sphere of human communication by language – by the words, expressions, stories, legends, and songs that encode and convey human relationships with the environment and that inscribe the history of those relationships onto the land. When people name places, they are identifying where things happen within the local environment, thus providing an entry in a mental encyclopedia that helps to describe the ecological niche occupied by the local peoples.

For example, certain American Indian groups in the Pacific Northwest have an extensive vocabulary for talking about salmon and other fish, vegetation, streams, and trails, along with extensive development of grammatical devices that express the direction, distances, and relative positionings of these valuable features. Thus language helps to enable these seminomadic hunter–gatherers to be very clear about where they were and where they needed to go to find food and other necessities. Many cultures in the far north have numerous names for various types of snow, and pastoral peoples have detailed terms for various types of livestock. The richness of language thereby reflects the many ways people relate to their environment.

Indigenous Peoples, Resource Management, and Biodiversity

Humans earn a living from the earth and its biodiversity in four main ways: hunting and gathering, pastoralism, shifting cultivation, and permanent agriculture. Only the latter is sufficiently productive to enable large cities to flourish, and the very high productivity enabled by irrigation may have been needed for civilization. Historical evidence indicates that traditional activities related to agriculture, fishing, and livestock husbandry sometimes have led to sustainable systems that were in the self-interest of the people involved, at least at their current level of population and technology. But all four forms of using biological resources carry with them the dangers of overexploitation. Broadly speaking, the combination of technology and population growth potential requires human cultures to develop effective ways of living in a sort of balance with the resources available with a given technology. The extinction of numerous cultures over time indicates how frequently the limits have been exceeded. However, many cultures have developed belief systems and practices that served to limit overexploitation and to foster greater biodiversity (even if these benefits sometimes were unintended).

For most of our history as a species, humans were hunters and gatherers, living from harvesting a very wide range of plants and animals; a few such groups still survive today, usually where agriculture is impossible or impractical. Hunter–gatherers were usually nomadic and fit ecologically into their habitat in much the same way as other omnivorous species. An important discovery that interrupted this presumed harmony and may have had major ecological impacts on biodiversity was the deliberate use of fire for driving game and clearing forest. This cultural activity produced major modification in vegetation patterns and fauna distribution in at least some regions. The invention and use of spears, bows and arrows, harpoons, and nets also must have affected the interrelationships between human beings and populations of other animal species. The effects of culture on humans themselves in

the hunter–gatherer phase included impacts on techniques of food gathering, especially hunting. Nevertheless, with the exception of the use of fire in some situations, it is difficult to accurately determine the impacts on biodiversity of these changes.

Given the very great variability in approaches to resource management taken by various groups of hunting peoples, it is not surprising that various hypotheses have been developed to explain why some groups of indigenous peoples appear to have been able to live in better balance with their resources than other groups. A common observation is that ethnic groups that have lived in an area for extensive periods of time and appear most dependent on their locally available resources are most likely to have developed sustainable forms of resource use. Many of the population processes that link human hunters and their prey occur over timescales that elude both ethnographic and archeological field-work. The critical factor is the distinctive features of the foraging economy and their outcomes for resource population ecology, as a matter of behavior with practical consequences.

Fishing is a specialized form of hunting and gathering. People who depend primarily on fish for their protein typically have developed both technology and management techniques to ensure sustainable yields. For example, Amazonian Indians may avoid certain parts of the river during breeding seasons; eastern Indonesian fishermen have complex systems of taboos, known as *sasi*, which control how and when the fishing grounds are to be used; and Polynesian fishermen had various taboos that served the function of managing their fisheries.

The development of domestication of plants and animals about 10,000 years ago marked the beginning of an entirely new era in the interplay between human society and biological systems. The controlled breeding of plants and animals, using a small sample of the world's biodiversity, established a more reliable and expandable resource base than subsistence hunting and gathering, enabling humans to reduce substantially the space required for sustaining each individual and allowing the human population to increase. All early farming activities consisted of redistributing plant and animal species in a given area by deliberately increasing the local concentrations of the species humans valued and decreasing concentrations of species that competed with the favored ones. In terms of biodiversity, the selective breeding of species also expanded the range of human impact from habitats and species to genes. The 40 or so species of animals that have been domesticated have often greatly increased their genetic diversity, range, and populations due to human management, as have the 100 or so main domesticated plant species and the thousands of other plants selected for use as food, spice, medicine, decoration, and construction. But the gains for these species typically have been at the expense of other species which people found less easy to mold to their vision of the desirable.

Domestication of livestock apparently arose after domestication of plants and remains the most significant use of arid and semiarid lands by people. The dangers of overgrazing have led to many beliefs about human relationships, livestock management, and range management. For example, many Middle Eastern pastoralists had complex ways of sharing grazing lands, including the idea of *hema*, a traditional grazing reserve. At least some species of trees typically are considered sacred by pastoral peoples and are not destroyed except in times of dire need; conservation of such vegetation thereby provides an emergency resource reserve (as well as conserving biodiversity).

Fallowing was the agricultural technique originally followed in most places; it remains a component of the shifting cultivation systems in many parts of Asia, Africa, and Latin America today. Such systems often use fire as a means of clearing the land, so they are sometimes called "slash-and-burn" systems. Such a system involves cultivating a plot of land – known as a "swidden" – for a few years and leaving it to develop into successional stages for a much longer period. This process creates a mosaic of vegetation, much of which is useful for foods, medicines, building materials, and other useful products (such as dyes, colorings, ceremonial objects, etc.). Many species of wildlife, including birds and terrestrial mammals, are attracted to the fallow swiddens because of their high productivity of nutritious vegetation. Furthermore, part of the land was often left out of cultivation. Thus in northeast Indian states like Manipur, as much as 10 to 30% of the land was permanently maintained under natural mature vegetation in the form of sacred groves. Historically, this would have ensured the persistence of almost all the natural elements of biodiversity, coupled with stimulation of overall productivity by favoring faster-growing early successional species in the patchwork of successional stages covering 70–90% of the land. Selection for adaptation to highly heterogeneous local environments also promoted considerable genetic variation in the cultivated species.

As an example of how shifting cultivation affects biodiversity in the Ecuadorian portion of the Amazon forest, Runa Indian swiddens resemble agroforestry systems rather than the slash and burn that merely results in temporary clearings in the forest canopy. Compared to unmanaged fallows, management actually increases species diversity in 5-year-old fallows. Between 14% and 35% of this enhanced species diversity is attributed to direct planting and production of secondary species. Thus Runa agroforestry can be seen as a low-intensity succession management system that alters forest composition and structure in the long term.

Among the most diverse of agricultural systems known are the home gardens in the humid tropics of tropical Asia, the result of long historical development of technology designed to meet the needs of local agricultural communities. In West Java, the typical home garden appears as a crowded assemblage of trees, shrubs, climbers, herbs, and creeping plants that are used for fruit, vegetables, starchy food crops, spices, ornamentals, medicines, fodder, fuel, and building material, and involve greater than a hundred species useful to people.

The Kantu of Kalimantan, Indonesia, plant at least 44 varieties in one area with an average of 17 per household. A gene pool of potatoes of some 3000 varieties representing eight species is traditionally under cultivation in the Andes. In Papua New Guinea, as many as 5000 varieties of sweet potato are under cultivation, with as many as 20 varieties being planted in a single garden. Thus indigenous peoples often maintain high biodiversity, at least among those species useful to them.

Traditional Belief Systems and Resource Management

The way of life of the world's tribal or indigenous peoples depends closely on biodiversity, often using cultural and religious beliefs to avoid disastrous overexploitation. Notions of sustainability of use are inherent in the value systems of many traditional societies, usually manifested in some notion of intergenerational equity. For example, the Iroquois tribe in North America would plan for the seventh generation when making their decisions, the life span, incidentally, of the dominant tree in their region. And the Koyukon Indians believe that future events will depend on the way people behave today and that the world can be nurtured by prudent use or harmed by unrestrained abuse; but equally important, the natural world will respond to gestures of respect given by those who recognize its sensitivity and awareness and humble themselves to its power. Plains Indians generally believed that their relations with other species were regulated by expectations and obligations similar to those that governed relations between kin or allies; such relations could vary from beneficial to harmful and could be mediated by ritual specialists.

In Colombia, the forests of the northwest Amazon basin harbor the world's most diverse array of plants and animals. This region is often considered part of the world's greatest remaining tropical wilderness. But the Tukano people who live there perceive their "wilderness" environment to be anthropogenic, transformed and structured in the past by the symbolic meaning their ancestors gave to resources and the knowledge they obtained about the plants and animals that enabled people to survive. Their forest is a system of resources in which the energy produced is directly proportional to the amount of energy it receives, a very modern perspective. They know that they cannot harvest more than the forest can produce, and they apply sophisticated knowledge of individual species and their uses. Their myths tell of animal species that were punished by the spirits for indulging in gluttony, boastfulness, improvidence, and aggressiveness. These myths serve as object lessons to human society, in which animals are metaphors for survival. By analyzing animal behavior, the Tukano find an order in the physical world within which human activities can be adjusted.

The restraints on harvesting may involve protection to keystone species that may support the persistence of a range of other species. Thus fig trees belonging to the genus *Ficus* are recognized as important sources of fleshy fruits available in seasons when no other species are producing such fruits in the tropical forest communities. *Ficus* thereby promote persistence of a number of insect, bird, bat, squirrel, and primate species for which they serve as a critical resource in a period of fruit shortage. All species of *Ficus* are even today to some extent protected as sacred trees through much of tropical Asia and Africa, where the local communities are aware of this ecological role. In other parts of the world, other species of useful trees may also be sacred, such as date palms in desert areas, baobab trees in Africa, palmyra in Brazil, and kapok among the Mayas.

Harvesting restraints also include protection of critical stages in the life history of species that are especially vulnerable to overharvest. Thus a nomadic hunting tribe of Western India has the tradition of releasing any pregnant does or fawns of antelope or deer caught in their snares, and egrets, storks, herons, pelicans, ibises, and cormorants at their colonial nesting colonies are given immunity from hunting over most of India, although these birds are hunted in the nonbreeding season. In many Asian villages, fruit bats are not hunted at their daytime roosts but may be killed at a distance from the roost during the night. Numerous such examples can be found in all parts of the world

Some traditional communities have developed detailed resource regulation systems. For example, many traditional communities – whether "indigenous" or "tribal" like the Tara'n Dayaks of West Kalimantan or the nontribal, such as the ribereños of Amazonian Peru – establish community reserves and formulate rules about how the species therein can be exploited, with the expressed intention of preserving these resources for the existing community and for future generations. Therefore, in some cases reserves may be protected by religious sanctions and in others by a more overtly "social" contract and enforced by strong economic and social sanctions; combinations of approaches are also found.

Biological diversity may well have been higher in rural areas in former times, when large numbers of farmers of many different cultures had long-term stakes in the land they farmed and had control over their own technology. These historical systems of land management were highly variable, following a range of different rules to take into account specific attributes of the physical systems within which they were found, cultural views of the world, and the economic and political relationships that existed in the setting. Despite their great diversity, such systems often had characteristics such as clearly defined boundaries, specific rules on the harvesting of different products, involvement of the affected people in these collective choices, a system of monitoring the use of resources, cultural sanctions for those who violated the operational rules, inexpensive local mechanisms for resolving conflict, and ways of organizing these activities so that different types of decisions were taken at different levels.

Urban Versus Rural Perceptions

About half the world's population today lives in urban settings far removed from the agricultural systems on which their lives depend. While rural peoples typically have a very good practical understanding of biodiversity, those living in cities increasingly have no more than a theoretical understanding, perhaps reinforced by an occasional visit to a national park or by a television program about nature. The growing numbers of nature-based films, protected areas, and forms of tourism based on idealized visions of nature indicate that urban people still feel a need to connect to biodiversity. Of course, different cultures have different perceptions, and the diversity of human societies is reflected in the diversity of perceptions urban people have of their environment. Any classification of such cultural values will necessarily mask some of this diversity, but rural people who live in close contact with the realities of nature typically have very different attitudes about biodiversity from urban people who have only a distant relationship with nature. Examples of these contrasting attitudes are presented in **Table 1**.

Table 1 Rural and urban attitudes about biodiversity

Rural attitudes	*Urban attitudes*
Humans as one part of nature's biodiversity	Humans not part of biodiversity
All living creatures considered equal	Humans superior to all creation
Culture/nature as a continuum (no such thing as wilderness)	Culture/civilization human, and nature considered wild
Natural way is right, and human activities should be molded along nature's rhythms	Human technology is superior, needs to mold nature to suit human needs
Biodiversity has an integral set of multiple values (cultural, spiritual, material)	Biodiversity is predominantly a material (economic) resource

These attitudes are ideal types, and will not necessarily be found in their pure forms in any rural or urban community. Rather, societies both over time and over space can be placed on a continuum between these opposites, as human attitudes change due to both internal social dynamics and external influences. However, it is clear that such attitudes greatly influence how humans use natural resources and therefore the impacts they have on biodiversity.

In marked contrast to the rural areas where people still live in a close relationship with biodiversity, urban dwellers are able to gain the benefits from the biodiversity of the entire globe. Their connections through trade and communications enable urban dwellers to have rapid and unprecedented access to biodiversity. Ironically, the expanding access to more biodiversity has tended to weaken the feedback between human welfare and the way biological resources are managed; the urban dwellers typically have only the very faintest notion of how their consumption might have impacts on biodiversity in the countryside of a distant country.

Conclusions

Establishing a connection between specific cultural practices and conservation or enhancement of biological diversity is by no means a simple matter, for the overtly declared purpose of a practice that seems to help conserve biological diversity may in fact be quite different. Thus, in South Asia many sacred ponds have helped conserve the indigenous fish fauna. But people may leave these ponds alone out of respect for some deities, not with an expressly declared purpose of conserving fish diversity. It is then quite possible that many cultural or religious practices that seem to promote conservation may have originated from different motivations, while others declared as promoting conservation may in reality achieve something very different.

Traditional societies have often protected parts of the natural landscape they live in, or left untouched some of its elements. Most such societies, for instance, have considered certain sites as sacred, where most or all human activities are prohibited. Most societies have also considered certain species as sacred, with elaborate myths and folk tales about how humans originated from such species, or how these species are incarnations of gods and deities, or in some way associated with them, or how they obtained magical powers. These could be wild species that are left undisturbed (and therefore have survived in even the most densely populated and radically altered human landscapes, such as rhesus and bonnet macaques in Indian cities) or domesticated plants and animals whose utilitarian value is intertwined with spiritual values (such as the cow in Hinduism). Such belief systems can have the effect (although perhaps unintentional) of regulating resource use.

A rich literature of traditional conservation practices indicates that they have been a common feature for many cultures over many years. This literature indicates that such practices serve group interests of communities and remain viable only so long as (1) local communities continue substantial levels of dependence on resources harvested from their immediate vicinity, (2) local communities have full control over the local resource base, and (3) local communities retain a sufficiently high level of internal cohesion. In most cases, these conditions are no longer fulfilled when outside state or corporate bodies establish control over natural resources.

When local people are part of a local ecosystem, their behavior directly affects their own survival. But cultural mechanisms that have been developed as adaptations to the environment over tens or hundreds of generations are quickly cast aside when trade or new technology frees people from traditional ecological constraints, changing them from "ecosystem people," who are adapted to their local ecosystem, into "biosphere people," who often live in cities and can draw from the resources of the entire world.

While changes in traditional attitudes toward biodiversity can be a result of internal dynamics (e.g., population increase), they are perhaps more often a result of outside influences: interaction with a modern culture, intrusion of the market, a technical innovation imported from another culture, and so on. Some have argued that traditional resource use patterns may be sustainable only under conditions of low population density, abundant land, simple technology, and limited involvement with a market economy. When confronted with market pressures, higher population densities, new technologies, and increased opportunities, few indigenous peoples appear able to maintain the integrity of their traditional methods, even when these are reinforced by religious beliefs.

See also: Aesthetic Factors. Agriculture, Traditional. Commons, Concept and Theory of. Ethnobiology and Ethnoecology. Historical Awareness of Biodiversity. Hunter-Gatherer Societies, Ecological Impact of. Indigenous Peoples and Biodiversity. Literary Perspectives on Biodiversity. Religious Traditions and Biodiversity

References

Bennett JW (1976) *The Ecological Transition: Cultural Anthropology and Human Adaptation*. New York: Pergamon Press.

Boyden S (1992) *Biohistory: The Interplay Between Human Society and the Biosphere, Past and Present*. London: Parthenon.

Burger and Julian (1990) *The Gaia Atlas of First People*. New York: Doubleday.

Crosby AW (1987) *Ecological Imperialism: The Biological Expansion of Europe 900–1900*. Cambridge, UK: Cambridge University Press.

Gadgil M and Guha R (1992) *This Fissured Land: An Ecological History of India*. Delhi: Oxford University Press.
Hamilton LS (1993) *Ethics, Religion and Biodiversity*. Cambridge: The Whitehorse Press.
Kinzig AP, Warren P, Martin C, Hope D, and Katti M (2005) The effects of human socioeconomic status and cultural characteristics on urban patterns of biodiversity. *Ecology and Society* 10(1): 23.
Maybury-Lewis D (1992) *Millennium: Tribal Wisdom and the Modern World*. New York: Viking Press.
Papayannis T and Mallaroch J-M (eds.) (2009) *The Sacred Dimension of Protected Areas*. Gland, Switzerland and Athens: IUCN and Meditteranean Institute for Nature and Anthropos.
Posey and Darrell A (1999) *Cultural and Spiritual Values of Biodiversity*. Kenya: United Nations Environment Programme, Nairobi.
Ruddle KE and Johannes RE (1985) *The Traditional Knowledge and Management of Coastal Systems in Asia and the Pacific*. Jakarta, Indonesia: UNESCO Regional Office for Science and Technology for Southeast Asia.
Suzuki D and Knudtson P (1992) *Wisdom of the Elders*. New York: Bantam Books.
Tomalin E (2004) Bio-divinity and biodiversity: Perspectives on religion and environmental conservation in India. *Numen* 51(3): 265–295.
Turner BL, Clark W, Kates R, Richards J, Mathews JT, and Meyer WB (1990) *The Earth as Transformed by Human Action*. Cambridge, UK: Cambridge University Press.
Verschuuren B, Wild R, McNeely J, and Oviedo G (eds.) (2010) *Sacred Natural Sites: Conserving Nature and Culture*. London: Earthscan.

Social Behavior

Daniel I Rubenstein, Princeton University, Princeton, NJ, USA
Dustin R Rubenstein, Columbia University, New York, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Affiliative behavior Behavior that is supportive and brings individuals together.
Agonistic behavior Behavior that is aggressive, threatening, combative, and submissive.
Altruism Behavior that enhances the recipient's reproduction while reducing the actor's.
Cooperative breeding Social system in which more than two individuals care for young.
Musth Annual reproductive period in male elephants characterized by extreme aggressiveness and secretions from temporal glands.
Polyandry Mating system in which a female mates with more than one male.
Polygyny Mating system in which a male mates with more than one female.
Protandry Breeding system in which individuals change sex from male to female.
Protogyny Breeding system in which individuals change sex from female to male.

Introduction

No animals are completely asocial. All must come together at some point to breed. Few, however, are highly social coming together and interacting repeatedly. For those that do interact frequently, social interactions – both affiliative and agonistic – occur sometimes with relatives, sometimes with strangers, sometimes with members of the same sex, sometimes with members of the opposite sex; and sometimes with members of the same generation while at other times with members of other generations. The interactions themselves can be aggressive, cooperative, or even altruistic and can develop into strong relationships among particular individuals. Depending on the nature of these relationships and with whom they form, a variety of social systems can develop. Some may be made up mostly of kin or mostly of nonkin, some may be based on territorial separation or on the aggregation of competitors, some will exhibit monogamous as opposed to polygamous relationships between the sexes, and some will rely on the cooperation and help of nonmates in the rearing of offspring. This seemingly bewildering range of social diversity is shaped ultimately by ecological circumstances and patterns of demography. But the very nature of a population's social structure itself will in turn affect the population's demography and its place in the biological community. It is these dual connections between social activities and the environment and the resulting feedbacks that mandates understanding both the causes and consequences of animal sociality so that effective and efficient management or conservation policies can be developed for protecting endangered species and for preserving biodiversity and ecosystem integrity.

From Social Behavior to Social Organization: Patterns and Mechanisms of Formation

Social animals form groups. While some are temporary, others are more permanent. Given that animals in groups incur automatic costs of increased disease and parasite transmission as well as intensified competition, groups will only form when there are sufficiently large benefits to offset these costs (Alexander, 1974). Benefits largely come in three forms. First, animals can develop forms of social behavior specific to stable groups that compensate them for the costs of group living. One such example is forming mutual grooming partnerships as do olive baboons (*Papio anubis*) to lower disease and parasite transmission. Second, by forming groups animals can enhance foraging by being better able to find, acquire, or defend food. Examples include colonial cliff swallows (*Hirundo pyrrhonota*) that transfer information about the locations of rich but ephemeral feeding sites or troops of monkeys that drive smaller troops away from feeding trees. And third, animals in groups can reduce their risk of being preyed on by either increasing the likelihood of detecting predators, diluting their personal risks, or by decreasing the likelihood that predators can make a kill; confusion and cooperative defense are mechanisms that provide such antipredator benefits. Examples include the scattering of fish in schools, or the gathering of young inside a ring of adult musk oxen (*Ovibos moschatus*) facing outward toward approaching predators with upturned horns.

Depending on the nature of the social relationships that develop among individuals, groups take on particular organizational forms (Emlen and Oring, 1977; Rubenstein, 1986). As **Figure 1** illustrates, these relationships are shaped by the features of an individual's ecological and social environment. Particular distributions of food or water (bottom-up factors) and predators (top-down factors), in conjunction with the physiological demands of individuals differing in body sizes or reproductive states, will determine the frequency and magnitude of competitive and cooperative interactions that occur. The outcome of these repeated social interactions will shape overall time budgets and activity patterns. Because females maximize their reproductive success relative to other females by their ability to raise offspring to the age of independence, females are forced to efficiently solve the three ecological problems: finding

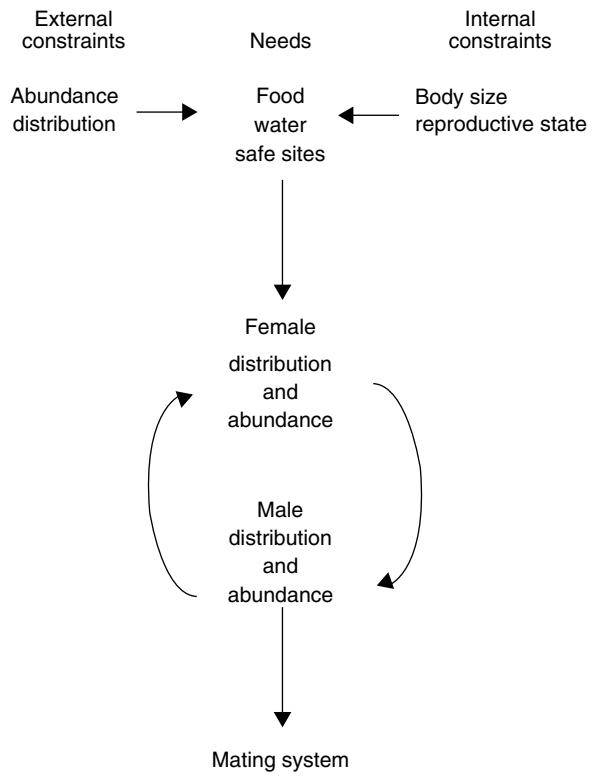


Figure 1 Model of ecological and physiological pressures influencing both intra- and intersexual relationships and showing how they shape social systems. Arrows depict social relationships. What type of society develops depends on details and perceptions of environmental conditions and interactions between males and females.

food or water; avoiding predators; or finding safe sites for young (**Figure 1**). The particular associations and distributions that develop will depend on the particular ecological and physiological circumstances that females experience. These associations and distributions in turn will shape male associations and distributions, because male fitness is mostly determined by their ability to acquire mating access to disproportionate numbers of females. Especially for mammals, the mating system that develops becomes the core of a species' social system.

Ecological Challenges and Phenotypic Constraints

Resource distribution in time and space shapes animal social systems (Emlen and Oring, 1977). In environments in which resources are abundant but evenly distributed, competition is low enough to permit females to aggregate. If sufficiently large foraging or antipredator benefits can be derived by females that aggregate, and the groups that form are not too large, then these groups can be defended by single males and so-called "harem defense polygyny" results. If resources, however, are more patchily distributed so that competition among females intensifies periodically, then female group sizes will vary and female associations become more transitory. Rather than defending unstable groups of females, males instead attempt to defend resource patches sought by females. Typically, in these systems of "resource defense polygyny" the most able males defend the best patches and thus gain access to the largest number of females for the longest periods of time. If resources are not only patchily distributed, but also the patches are large, widely separated, and fluctuate seasonally in abundance, then competition among females becomes so low that the formation of large groups is even more likely, provided that females can range widely and follow the shifting locations of peaks in food abundance. Males will thus be forced either to follow these large groups competing for, and then tending, one reproductive female at a time ("wandering") or to position themselves at the intersection of female migratory routes and wait for females to visit them ("lekking"). In the first case, intense male–male competition generates a mating system based on "male dominance polygyny," and in the latter case females are afforded the exquisite opportunity to simultaneously compare many males before choosing with which one to mate! Whenever resources are sparsely but somewhat evenly distributed, high levels of competition prevent females from forming groups. As a result, individual females defend territories thus ensuring a regular supply of a renewing resource. Since solitary individuals searching for members of the opposite sex will face increased predation risk, pairs often share territories and monogamy results.

This model accounts for the diversity of mating systems for many different groups of animals including insects, fish, reptiles, rodents, and many varieties of birds, ungulates, carnivores, marine mammals, and primates. One of the best illustrations showing how environmental forces interact with physiological constraints to shape a species' social system emerges from Peter Jarman's classic study of African antelopes (Jarman, 1974). By showing how body size affected the ways in which different species perceived, and then responded to, the distribution and abundance of forage and predators of grasslands, Jarman showed why particular social systems increased survival and reproductive prospects for particular species. He argued that the smallest bodied species, such as dik-dik (*Madoqua kirkii*), duikers (*Cephalophus* spp., *Sylvicapra* spp.), suni (*Neotragus moschatus*), and klipspringers (*Oreotragus oreotragus*) require limited amounts of high-quality vegetation. But given their small size, such food items often appear as if they are widely scattered. Faced with high levels of competition and intensified risks of predation, territoriality and monogamy appear to be the best strategies. Pairs generally live in wooded or shrub-rich areas where moisture enables vegetation to grow and renew itself well into the dry season. By signaling territorial ownership via scent rather than by means of sound or visual display these small-bodied species reduce the chances that any of a large number of carnivores will prey on them. As species increase in body size, both physiologically determined dietary needs and the way acceptable forage becomes distributed on the landscape changes. Since hiding or "crypsis" becomes an untenable antipredator strategy for larger and more widely ranging species, forming groups becomes the best strategy for larger species to lower predation risk. Fortunately, with larger size also comes an ability to subsist on more abundant, lower quality vegetation. When food is patchily distributed, as it is for impala (*Aepyceros melampus*), reedbuck (*Redunca* spp.), and some gazelles (*Gazella* spp.), males defend the best patches that females prefer. When the vegetation is more evenly distributed, which often results simply from the fact that larger

species such as eland (*Taurotragus oryx*) and Cape buffalo (*Synceros caffer*) can utilize even the lowest quality items, larger groups form. Because the biggest species view large continuous swards of a landscape as acceptable, competition is virtually eliminated and many males associate with many females. With such high levels of male–male competition, defense of a small subgroup of females becomes impossible and dominance defense systems develop.

The same sorts of connections between changing ecological circumstances shape the types of sociality exhibited in other taxa. A brief survey of mammals illustrates some of the more general patterns. For the equids the close association between food and water enable horse (*Equus caballus*) and plains zebra (*Equus burchelli*) females of different reproductive states to associate permanently (Rubenstein, 1986, 1994, 2010). Thus, males are able to defend such groups and so-called harems form. When these two resources are widely dispersed, as for Grevy's zebra (*Equus grevyi*) and the Asiatic wild ass (*Equus hemionous*), females of different states are precluded by metabolic constraints from foraging together. As a result, males compete for territories along traveling routes that take females from feeding areas to watering points.

In felids, females remain separated when food is scarce and the habitats are densely wooded because in such circumstances individual prey can be caught and consumed before competitors can intervene. In more open habitats and where both prey and competitors are large and much more numerous, coalitions of females form to help hold on to kills until they are completely devoured. If these female coalitions are themselves large, then there is pressure for males to aggregate to control reproductive access to females. Thus, the only highly social felid is the savanna-living lion (*Panthera leo*), yet even in this plentiful landscape the leopard (*Panthera pardus*) remains solitary because it can safely catch its large prey in trees (Packer, 1986).

For canids monogamy is the rule. But variations do occur with small-bodied foxes (*Vulpes* spp.) sometimes exhibiting polygyny and large-bodied hunting dogs (*Lycaon pictus*) and timber wolves (*Canis lupus*) developing polyandry (Moehlman, 1986). Typically canids need help from nonbreeding foragers to help nourish lactating mothers. For the midsized jackals, such as the silverback jackal (*Canis mesomelas*), the helpers are young from previous litters that cannot themselves find successful breeding locations. For the smallest species, however, prepartum investment in young is relatively small for a female and thus her mate provides all the help that is needed. In years when prey resources are very high, competition between mothers and their soon-to-be-fecund daughters is low and their daughters are not forced to disperse. As a result they provide neighboring males with additional mating opportunities. For the largest bodied species, however, female prepartum investment is very high and they need all the help they can get. To enlist the support of other adult males that must hunt cooperatively to capture large prey, a dominant female not only kills the offspring of other females in the group, but also she mates with their mates so that these males behave as if they are the sires of the dominant female's young. Hence, depending on size-determined metabolic investments and needs, social systems of canids can vary from polygyny to polyandry and sometimes use the services of juvenile or adult nonbreeders to help rear offspring.

Monogamy is also the rule in birds, with more than 90% of all species exhibiting social monogamy. Unlike in mammals, where mothers are typically able to take care of offspring on their own, bi-parental care is essential for producing young in most birds. However, this does not mean all pairs are faithful. Although more than 90% of birds are socially monogamous, an almost equal percentage of these seemingly faithful species are not actually sexually monogamous. Although rarer, other forms of mating systems also occur in birds. When resources are clumped, males defend those that are essential to females. For example, in many species of hummingbirds, males defend territories when flowers that females need are sufficiently clumped. Lekking is particularly common in birds compared to mammals. Some of the most recognizable avian species, like the ornamented birds-of-paradise, the bowerbirds and their decorated bower structures, and the dancing manakins lek. In these species, males cannot defend resources or females, and so they aggregate in areas where females are likely to compete against each other for female attention.

Ecology not only shapes animal mating systems, but also in some cases, the social systems themselves. Many species of vertebrates live in complex societies where individuals delay dispersal and independent breeding opportunities to help other group members raise their young. In these "cooperatively breeding groups," dispersal decisions are typically influenced by ecological factors. Limited resources, such as territories, nest sites, and even mates, prevent young individuals from dispersing to breed on their own. Although the spatial distribution of resources on the landscape can influence dispersal and helping decisions, temporal variation in resource abundance may also affect cooperative breeding behavior. In birds, cooperatively breeding species tend to be found more often in unpredictable environments where rainfall and resources vary erratically from year-to-year. Having helpers-at-the-nest may allow cooperatively breeding species to buffer these uncertain times better than those species without these nonbreeding auxiliaries (Rubenstein and Lovette, 2007, Jetz and Rubenstein, 2010).

Kinship and Demography

Although the abundance and distribution of key ecological resources are important in determining the particular pattern of cooperation and sociality exhibited by a population, other factors such as kinship and demography also play significant roles. Since W. D. Hamilton formulated the concepts of "inclusive fitness" and "kin selection," the magnitude of the costs and benefits associated with particular social behaviors that ultimately determine which strategy is the best, had to be adjusted by degree of relatedness between social partners. Altruistic or cooperative behavior between relatives should be favored whenever Benefits $>$ Costs $\times r$ where r is the degree of relatedness. For both parents and offspring and among full siblings, $r = \frac{1}{2}$ for grandparents and grand offspring and among half-siblings, $r = \frac{1}{4}$; and among first cousins, $r = \frac{1}{5}$. Hence, the stronger the degree of relatedness among relatives, the more likely altruists are to enhance the reproductive opportunities of kin while incurring costs associated with diminished personal reproduction. Thus, in the examples presented here, it is not surprising to find that the coalitions

that form among lionesses when protecting kills and among male lions when protecting mating opportunities with females are formed most often among full siblings; nor is it surprising that the helper jackals recruit to rear additional offspring are themselves the full siblings of the offspring being raised. In general, strong kinship lowers the threshold for the appearance of altruistic and cooperative behavior and may substantially affect the costs of living in groups, particularly in cooperatively breeding species.

Nowhere is this more apparent than in elephant societies (Archie *et al.*, 2006) (Archie *et al.*, 2006). As is the case in many mammals, elephants live in fission–fusion societies in which individuals, or the subgroups they inhabit, separate and merge from time to time. In African elephants (*Loxodonta africana*), up to 20 individuals and their young live in core family groups in which daughters remain after reaching sexual maturity, while males leave. Depending on ecological conditions, core groups can separate into smaller units or merge with other core groups to form bond groups. Extensive genotyping of individuals using molecular DNA markers show that not only do family groups typically consist of close genetic relatives, but also that bond groups are most likely to form when the oldest females of the core groups were genetic relatives.

But as many of these cases illustrate, high population density and the intensified competition elephants engender for finding suitable territories with sufficient food and habitable burrows are the factors ultimately responsible for favoring the establishment of coalitions of relatives or the recruitment of helpers in cooperatively breeding societies. Thus, demographic factors, such as population density as well as sex ratio and age structure that result from differences in phenotype-specific vital rates also shape patterns of social organization. And since these features of populations are often altered by human activities, understanding how they shape, and are shaped by, patterns of sociality is essential if conservation assessment and planning are to be effective.

Hypothetically, if mortality, for example, were age specific and higher for prereproductive females than males, then high breeding sex ratios (males/females) would result and mating relationships would generally become more polygynous. If the sex-specific patterns of mortality were reversed, however, then polyandry would become more common. Since some of these mortality concerns were responsible for the variation in canid social structure described earlier (*see* Ecological Challenges and Phenotypic Constraints), these mortality schedules can have important consequences. Similarly, if mortality rates were size specific and happened to be greater for larger males rather than smaller ones, then discrete size polymorphisms could arise in populations. This appears to be the case regarding the maintenance of so-called alternative male mating strategies. Often the typical strategy adopted by older and larger males of defending harems or resources yields the most reproductive gains but it often incurs the highest costs as well. Because displaying, fighting, attracting the attention of predators, and delaying reproduction while growing are all costly activities, males adopting not only less successful but also less costly tactics can also flourish (Rubenstein, 1980).

Maintenance of such alternative mating patterns is common among many species of insects, fish, amphibians, birds, and mammals. For example, in bluegill sunfish (*Lepomis macrochirus*) males typically defend nest sites where they display, attract females with which to mate, and then fertilize and guard the eggs they lay. Since only the largest males have the ability to defend such nest sites, they must delay breeding in this fashion often for more than 7 years. As a result, smaller and younger males have evolved various cuckolding strategies that may in fact be equally successful evolutionarily. In one, the so-called sneak begins breeding at 2 years of age. Although very small, such sneaks are virtually all testes and because of their inconspicuousness can dive from the surface just as a mating pair release sperm and eggs. By exuding large volumes of sperm, some make it to the eggs and fertilize a few. In the other, males delay breeding for as long as do females. By being the same size and color as females, these males join the mating pair and apparently fool the displaying male into thinking that he is courting two females. Then as the original pair releases their gametes, he does too and thus fertilizes some of the females' eggs. For these two strategies to be equally successful alternatives, the costs and benefits of each must vary inversely with frequency. Although they sometimes do, as in the case of the bluegill sunfish, in other species they do not (Gross 1991; Brennan *et al.*, 2008).

A variant on this theme is associated with sex change. For many species, especially among the fishes, individuals begin life as one sex and end life as another. In bluehead wrasse (*Thalassoma bifasciatum*), for example, most individuals begin reproducing as females while residing with one male. If he disappears, then the largest female changes sex and becomes the harem-tending male (Warner *et al.*, 1975). Alternatively, in anemonefish (*Amphiprion* spp.), some individuals change from male to female. In either protandrous (male first) or protogynous (female first) species (Fricke and Fricke, 1977), the sex that is last is the one whose reproductive success is most influenced by body size, and this ultimately depends on the nature of the social system.

Clearly, a variety of environmental features influence the patterns of sociality that populations of animals develop. In many cases these patterns are flexible and species can vary in the system of social organization they exhibit depending on environmental conditions. Thus, although female burros and asses typically live in transitory groups whose membership changes, when populations move from arid to mesic areas, social relationships can change. As was found on an island off the coast of Georgia, with food and water both more abundant and less separated, females with differing needs can and do coalesce into permanent groups. Males, which in arid areas are forced to establish territories along routes to and from water, respond by defending these groups much like males of horses, their close kin, and harem-like societies emerge. In some species of birds where pair-bonds are the norm, cooperative breeding groups form via the retention of offspring during hard times. In these occasional, or "facultative" cooperative breeders, flexible social behavior is an important part of their social lives.

Egalitarianism and Despotism

In some societies, rewards are relatively evenly distributed among members, whereas in others they are not. When animals compete for critical resources, differences in fighting

ability often lead to skewed payoffs, with dominant individuals securing a disproportionate number of rewards, which are typically related to reproduction (Hager and Jones, 2009). The degree to which dominants can control the actions of subordinates varies. If it is incomplete and the rewards to subordinates low, then subordinates will be induced to leave and dominants will suffer. Thus, dominants may be constrained to concede some resources to subordinates. For males, skew is often related to mating opportunities and recent models suggest that when control by dominants is incomplete, the degree of skew will depend mostly on ecological factors and fighting ability (Reeves and Shen, 2006). In meekats, for example, reproductive success of subordinate group members increases with increases in body condition and years with high rainfall (Clutton-Brock *et al.*, 2001). Thus, for highly social and cooperatively breeding species, the genetic diversity of populations and the diversity of behavioral strategies they exhibit are likely to change with changes in climate and environmental features.

The existence of dominance hierarchies is a fact of life for most social species, and as noted above, play an important role in structuring societies, skewing payoffs, and when those payoffs are tied to reproduction, affecting the long-term evolutionary sustainability of species. But the degree to which dominance influences these outcomes depends on the costs of achieving and maintaining dominance ranks, since high costs often shorten life span and reproductive potential. Conventional wisdom based on many studies posits that dominants only experience high levels of stress in unstable societies. But recent work on baboons shows that even in stable societies life at the top can be stressful. While stress levels in baboons generally decrease with rank, alpha males exhibit cortisol levels equal to those of the most subordinate members of the group and are much higher than those of males immediately below the top male (Gesquiere *et al.*, 2011). Given that life at the top is likely to be costly for social species in both stable and unstable societies, it is likely that as deteriorating environmental conditions exacerbate competition for resources and mates, the longevity of alpha males will be decreased, thus altering the genetic diversity of populations and perhaps their demographic potential as well.

Organization and Networks

Social organizations are often characterized by general characteristic such as the number and types of mating pairs or the stability of relationships. But such qualitative characterizations can often be misleading and hide subtle differences in the function or causation of societies. As a result, tools from network theory are being used to quantify the overall structure of societies and the degree of "centrality" and extent of connectedness (termed "betweeness") that individuals display within societies. The use of such quantitative measures is particularly important for fission–fusion societies where individuals often form many different types of relationships. In equids, for example, both Grevy's zebras and Asiatic wild asses live in societies in which males generally live apart from each other and defend territories while females move among territories singly or in groups that regularly change membership. When networks of proximity based on who has been associating with whom are constructed, the graphs they produce are very different (**Figure 2**) (Sundaresan *et al.*, 2007). Inspection shows that just about every wild ass is connected to every other wild ass and that individuals are organized into one large community. In Grevy's zebra, however, more communities (called "connected components") exist, even though within each community every individual is connected to every other individual. In general, individual Grevy's zebras show higher levels of centrality than onagers because the number of close associates (termed "degree") is greater in Grevy's zebras than in wild asses. In addition, "cliquishness" of Grevy's is greater than wild asses because the close associates of particular Grevy's zebras are more likely to be close associates themselves than in the wild asses.

By computing these and other metrics, social network analysis provides a means of quantifying a variety of different features of society. And by doing so, social network analysis can help assess how social structure can cope with, or perhaps evolve in response to, environmental factors. For example, the metrics described in the previous paragraph reveal that Grevy's zebra society is characterized by more modularity than that of wild asses and that this difference corresponds to a difference in the uncertainty of resource availability and levels of predation that each species experiences. Whereas wild assess live in environments in which humans have extirpated predators and have provided many predictable watering points, Grevy's zebras face mobile predators and large fluctuations in rainfall due to periodic and unpredictable periods of drought. Network modularity appears to help prevent the loss of important information for coping with environmental fluctuations.

From Social Organization to Ecological Impacts

Despite the fact that some species exhibit social flexibility and that there exists significant "environmental determinism" in the development of patterns of sociality, knowledge is only now emerging about how changes in social systems impact important vital rates, alter interspecific competitive abilities, and shape a population's growth rate and genetic structure. Never before have the social systems of so many species been forced to respond to human activities, most of which lead to reductions in population size, through overharvesting, fragmentation of landscapes or habitats, and changes in the earth's climate. How are we able to predict which species will respond and how will a species' response impact its long-term growth and the health of the ecosystem in which it resides? When we intervene will we do so with a sufficient understanding of what the consequences of changing social behavior and social systems are likely to be? For example, will we know when a species switches from a monogamous to a polygynous social system and whether its growth rate or genetic diversity will increase or decrease? Will we know whether it will be more or less able to withstand selective harvesting or habitat fragmentation? Such consequences of differing patterns of sociality must be understood before endangered species or their habitats can be protected. At least for free-ranging horses attempts to limit population growth using immunocontraception has led to a cascade of unintended

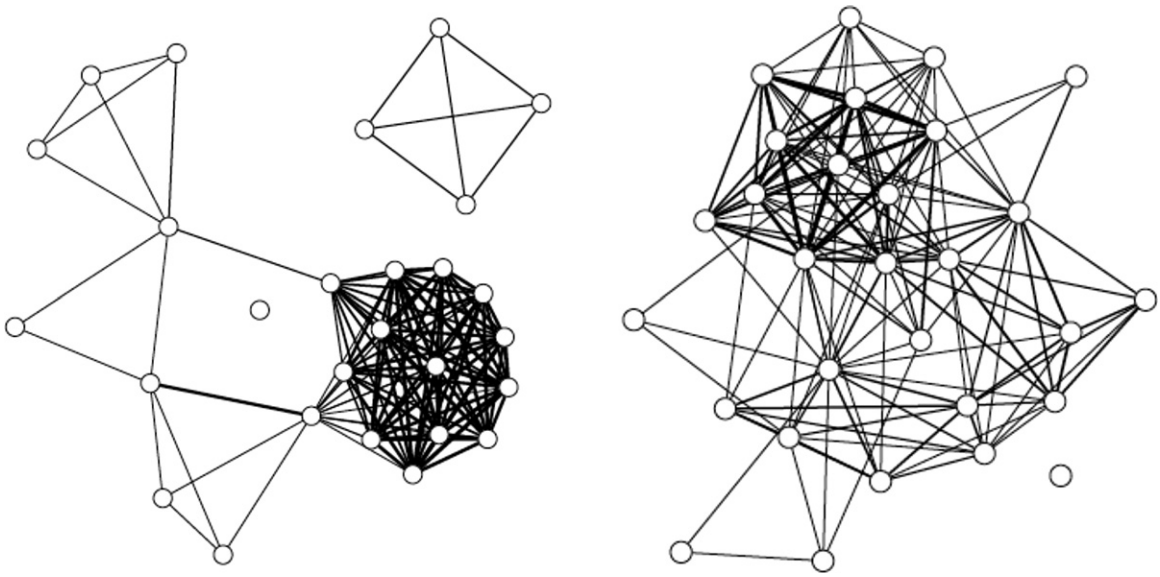


Figure 2 Network graphs of the fission–fusion societies of Grevy's zebras and Asiatic Wild asses. Both societies have one individual who is isolated from everyone. In general, Grevy's zebras exhibit more modularity than wild asses where virtually everyone individual is connected to every other individual.

consequences that have altered group stability and mating patterns (Madosky *et al.*, 2010) as well as the length of the breeding season and long-term male–females social and reproductive relationships (Nuñez *et al.*, 2010).

Consequences of Social Change

Many examples are now appearing of human induced changes in species' social systems. One population of feral horses inhabiting a barrier island off the east coast of USA has been impacted twice by human activities and both the times major changes to social behavior and social organization resulted in large-scale changes in the population's vital rates. The first instance occurred when dredging of surrounding ship channels altered currents and sediment deposition around the island. Before the emergence of hundreds of hectares of a new continuous sward of grassland, the island's horses exhibited two types of social organization. Where the vegetation was abundant and evenly distributed, females formed permanent membership groups and associated with a single male. Such harem groups were stable and males protected their females from harassment by intruders and rarely herded their own females. Where the preferred high-quality vegetation was broken up into patches of variable size by extensive ridges of sand dunes, females were not able to aggregate into permanent membership groups. As a result, males wandered in search of reproductively receptive females and, when doing so, harassed females and limited their ability to forage. Not surprisingly, the body condition and reproductive success of the harassed females was much lower than that of females bonded to males. After the emergence of the new sward of grassland in the region where females lived in the fission–fusion type of society, females were able to aggregate and males were able to establish harems and the growth rate of the entire population increased markedly. Thus, when the abundance and distribution of key ecological resources alter social systems so that they more closely resemble a species' typical system, vital rates and the overall health of the population improve.

When the perturbation moves a population away from its normal pattern, as happened during the feral horse population's second brush with human activity, vital rates decline. As the horse population approached its carrying capacity, population control measures were instituted. During the initial round of management actions, more than 85% of the harem stallions were removed from the island. The resulting social disruptions were massive; many females were separated from their males and young, inexperienced bachelor males took over. Extremely high levels of harassment led to low reproductive rates despite the fact that overall density had also been reduced and food supplies were expanding. As these two disruptions illustrate, human activities that alter existing patterns of social relationships dramatically alter reproductive rates.

Similar outcomes occurred when large-scale El Niño induced climatic changes resulted in major adjustments to time and activity budgets of Alaskan red foxes (*Canis aureus*) (Zabel and Taggart, 1989). Before the 1982–1984 El Niño, many of the foxes on an island in the Bering Sea bred polygynously and produced litters of sizes equivalent to, and often greater than, those of monogamous pairs. After the El Niño and the reproductive failure of the large seabirds, the foxes shifted their diet to smaller, less abundant, and harder to catch prey. Dietary shortages were common, fewer foxes bred and polygyny disappeared. Moreover, the litter sizes of these monogamous pairs were smaller as well. Clearly, changes in social behavior in response to altered environmental conditions enabled some individuals to make the best of a bad situation. Overall, however, the health of the population suffered.

Major changes in the breeding patterns of elephants (*Lo. africana*) could result if climate changes alter the

patterning of grasslands in East Africa. Ordinarily group sizes and compositions change seasonally in response to changing abundances of vegetation. After the rains, when grasses grow rapidly on the extensive upland plains, small family groups of elephants aggregate into large matrilineal assemblages. With the flush of new vegetation many come into estrus and become reproductively active. At this time only the most dominant bulls come into musth, a heightened sexual and aggressive state. By aggregating, females incite male–male competition thus ensuring that they mate only with the best females. During the dry season, however, after the grasses on the plains stop growing and have been consumed by elephants and other ungulates, the elephants retreat to the swamps where other grasses and browse remain. The patchy nature of the habitat, however, prevents large groups of females from remaining together and the huge herds fission into smaller family units. In addition, few females remain to be mated and the dominant bulls go out of musth. Subordinates become sexually active at this time, but reproductive opportunities are limited. Thus, under a normal environmental regime two evolutionary consequences result: most males do not mate, but those that do are the most fit; and because genetic effective population size is affected by the number of males relative to the number of females mating, the effective population size of elephants is much smaller than the census population size. If global warming increases the duration and intensity of dry seasons, then selection for the best males will be relaxed but the effective population size of the population will increase, provided that elephants can sustain their high metabolic levels and maintain their high fertility and survival rates. With global changes in climate appearing in many different habitats, other examples of social readjustments that either change a species' evolutionary potential or that act as demographic shock absorbers in the short run, but are unlikely to provide compensatory relief in the long run, are likely to become more common.

Consequences of Demographic Change

Many demographic changes leave their mark by altering critical aspects of social behavior or by disrupting the development of important social relationships. As the elephant example illustrates, changing environmental conditions can change operational sex ratios thus altering a population's ability to maintain genetic diversity. But by changing sex ratios in other ways, humans can put populations at risk by making it difficult for them to grow in size. In lions, for example, trophy hunting, if it removed pride defending males and nomads in proportion to their abundances, would not severely affect population growth and effective population sizes would be maintained at normal levels. But if such hunting focused on pride males, then increased turnover would foster increased infanticide by take over males. While such diversification of the gene pool might increase effective population sizes, it would certainly lower population growth rates because infant survival, which is already low, would plummet. Much clearly depends on whether or not the population is polygynous or monogamous and whether the perturbation will accentuate or reverse the pattern. Large-scale climate changes could reverse and ameliorate the normal pattern as is likely to be the case in elephants, but selective hunting or poaching, as in the case of lions, could exacerbate an already critical situation.

Selective poaching could even make it difficult for populations to recover once the poaching has been eliminated. Since hunting elephants for ivory meant that poachers sought the individuals with the largest tusks first, male numbers were reduced to dangerously low levels before large tusked females and smaller males were hunted. Because females prefer to mate with older males, lowered sex ratios and distorted age structures could make it difficult for all females to mate. As a result, population growth rates could remain low for historically heavily poached populations.

Harvesting in fish might be accelerating already rapid population declines because of disruptions to normal social interactions. When harvesting is size selective and reduces the number of large females in a population, fecundity might be disproportionately reduced since ovary volume scales allometrically with linear body size dimensions; larger females have disproportionately larger ovaries and high fecundity. And as sex ratios increase, male harassment of the remaining females may further reduce recruitment and population growth rate. When such size selective harvesting removes more males than females, sex ratios are reduced and females may become less choosy and more aggressive when selecting mates. While such changes in the intensity of sexual selection might not alter population growth rates, they might alter the genetic diversity of the population. Yet if males provide the majority of parental care and large males are removed from the population as if often the case in sports fisheries, then recruitment will be markedly reduced as unguarded eggs are cannibalized. The implications of overzealous size-selective harvesting in sex-changing fish or in fish that exhibit alternative male mating behaviors appear to be equally severe. For species such as anemonefish in which the largest males change to females, removal of the largest females will force males to change sex at ever smaller sizes, thus producing fewer eggs. Population growth rates will be curtailed and the species will find itself at risk all because of overzealous aquarium traders.

Migratory Atlantic salmon (*Salmo salar*) provide perhaps the clearest example of where human activities are changing the balance of male mating strategies within a population and hence the patterns of sociality. Atlantic salmon typically are born in fresh water and develop for 1–2 years before smolting and heading to sea to grow and fatten by feasting on a rich supply of marine invertebrates. Once they attain a certain size, they become sexually mature and return to rivers. There they travel long distances upstream until they reach clear and cool breeding grounds. Some individuals, however, never head to the sea. Instead, they remain in their natal streams and mature sexually at young ages and at small sizes. Such individuals are called parr and they never go through the smolting process that adapts them to a marine lifestyle. Under pristine environmental conditions, the fraction of the population that becomes parr is small since the reproductive gains of such a strategy are low. When competing for mates with larger more aggressive males, parr fare poorly. What matings they obtain are derived by "sneaking" among a mating pair and releasing milt at just the right time. Such events are rare and the mixing of milt is poor

so such "sneak" matings result in few young being sired. But the survival prospects of parr are high since they do not incur the risks of going to sea and, moreover, they begin breeding at a very early age. Thus, over a lifetime, reproductive success is moderate. But as human fishing increases and the netting of the older larger salmon intensifies, the relative lifetime reproductive success of parr is improving. Since the costs of migrating to sea and back again have increased dramatically, the long life span of parr give them a relative advantage. And given that competition with the larger males for mates is also being reduced, the chances of parr securing matings are also improving. Thus, it is not surprising that the composition of the population is changing as parr increase markedly in abundance. The impact on the long-term stability of the population is unclear. But the long-term impact on the balancing of mating morphs and sexual relationships is.

Consequences of Inappropriate Interventions

When humans intervene and develop management and conservation plans they sometimes do so without accounting for important aspects of social behavior. As a result some disastrous consequences have ensued.

Ignorance about the mating system of sperm whales (*Physeter macrocephalus*) has led to the implementation of harvesting plans that have almost led to their extinction. By using data from previous harvests in which the counts of males were much fewer than those of females, sperm whales were thought to exhibit a harem breeding system. Such a conclusion could only have been derived by assuming that sperm whales lived in permanent, closed membership groups. By making this assumption, the International Whaling Commission concluded that most males sighted would be superfluous thus disproportionate hunting of males would not limit the growth rate of the population; most of the males caught would not be necessary for fertilizing females.

Unfortunately, the social relationships of sperm whales are very different from those assumed. Females are the core of the society and do appear to live in permanent membership groups. But these groups often merge when diving for fish and divide labor between fishing and tending the young that cannot dive deeply. Males are forced to leave their natal group when reaching sexual maturity, but they cannot accompany the breeding associations to the tropics. Instead they remain at the higher latitudes foraging in the cooler more productive waters. Thus, when males are hunted on the breeding grounds it is breeders and prereproductives that are taken. With breeding males severely reduced in abundance and the younger, subadult upstarts located thousands of kilometers away, many females are going unmated. As a result, the population is failing to recover and it will be years before the next generation of males is ready to start the population, already handicapped by such a low recruitment potential, on the road to recovery. Knowledge about sperm whale social relationships could have prevented the mistaken overharvesting of mature males.

Another species in which ignorance about important features of social behavior is problematic is the black rhinoceros (*Diceros bicornis*). During the past 25 years, numbers have declined from 65,000 to fewer than 5000 in 2011 because demand for their horns is so great. In an attempt at making rhinos less desirable, some nations have implemented a strategy of dehorning. Other nations have opted to translocate rhinos to extremely well-protected areas. Unfortunately, horns, and thus the value of rhinos, regrow quickly. But what is more troubling about the dehorning treatment is that it appears to disrupt effective maternal antipredator behavior. Although dehorned rhinos were no more likely to flee from predators than intact rhinos, the disproportionate disappearance of offspring being reared by dehorned mothers as opposed to intact mothers in areas with large numbers of lions and spotted hyenas suggests that horns play a vital parental and protective function. Dehorning, although initially thought of a cure for the disappearance of adults via poaching, became a contributing factor to the disappearance of young via predation. Although dehorning might enhance adult survival in the short run, it would limit a population's growth potential in the long run.

A third example where ignorance hindered, but did not harm, a conservation program involved the reintroduction of Asiatic wild asses (*E. hemionus*) to habitats within its historical range. A goal of the Israeli government is to repopulate Judea and Samaria with biblical animals. Since the Palestinian race of the wild ass was extirpated by the Ottaman Turks at the beginning of the twentieth century, the onager was a prime candidate for translocation. The first reintroduction took place in 1982 with subsequent additions throughout the 1980s and increased the number of breeding females to 14. By the mid-1990s, however, the population had hardly grown and the number of breeding females totaled 16. What was not known at the time was that breeding success is bolstered by being reared in the wild and that onagers could facultatively adjust the sex ratio of the offspring they produced. In the Negev Desert, free-ranging onagers gave birth disproportionately to sons.

Trivers and Willard (1973) were the first to suggest that differences in the ability to invest in the successful rearing of offspring should lead to individual differences in the primary sex ratios of offspring. They argued that females with sufficient resources should invest in offspring of the sex with the higher variance in reproductive success. By producing well-endowed offspring of this sex, mothers would increase the number of grand offspring they were likely to have. In polygynous species like the onager, males are usually the sex with the greatest variance. Since onager females compete more on the basis of resource exploitation, or utilization efficiency, rather than via direct conflicts settled by social status, sex biases in offspring production were not expected. However, because age and experience affect utilization efficiency, differences in bodily condition and ability to invest in offspring were expected to exist among females in this onager population. Thus, it was heartening to find that middle-aged females – those in the best physical condition – gave birth to more sons than daughters. Conversely, those breeding for the first or second time gave birth to daughters as did older females nearing the end of their reproductive lives. The strategy the Israelis employed of translocating middle-aged females to the desert was sensible in economic terms since high risks were attached to releasing very young females and low future expectations were

associated with releasing older females. But had the Israelis released old females with yearling daughters they would have quickly accelerated the growth of the population. At the time, an understanding of how environments shape patterns of onager social behavior and the impacts that these behaviors would have on population dynamics was unknown.

The Kakapo (*Strigops habroptila*), a critically endangered ground-dwelling parrot native to New Zealand, represents another conservation story where sex-ratio manipulation may be critical to saving the species. Males, the largest parrot species in the world, are not only larger than females and better able to escape predators, but also females are particularly at risk of predation when incubating nests. In 1982, all remaining birds in the wild were brought into captivity. Of the 62 remaining birds in 2001, less than 30% were female and only six female fledglings had been produced since 1982. Owing to the supplemental feeding in captivity, most females were producing sons. After deciding not to boost female condition through supplemental feeding, the offspring ratio evened to approximately 50:50 and more daughters have since been produced.

Most animals exhibit some form of social behavior. Much of this behavior is intimately tied to the environment in which the organisms live. By understanding the relationships between environments and social behavior and between social organization and population processes, we should be in a better position to understand the problems endangered species are likely to face and what interventions to protect them are likely to be effective. In a time when global warming is a reality and human modification of the landscape is occurring at alarming rates, gaining a better understanding and appreciation for animal social behavior will help in preserving species and ecosystems across the planet.

Appendix

List of Courses

1. Behavioral Ecology
2. Animal Behavior
3. Social Behavior
4. Conservation Behavior

See also: Population Density. Salmon. Species Coexistence

References

Alexander RD (1974) The evolution of social behavior. *Annual Review of Ecology and Systematics* 5: 325–383.

Archie EA, Morrison TA, Foley CAH, Moss CJ, and Alberts SC (2006) Dominance rank relationships among wild female African elephants. *Loxodonta africana. Animal Behaviour* 71: 117–127.

Archie EA, Moss CJ, and Alberts SC (2006) The ties that bind: Genetic relatedness predicts the fission and fusion of social groups in wild African elephants. *Proceedings of the Royal Society: Biological Sciences* 273: 513–522.

Brennan BJ, Flaxman SM, and Alonzo SH (2008) Female alternative reproductive behaviors: The effect of female group size on mate assessment and copying. *Journal of Theoretical Biology* 253: 561–569.

Clutton-Brock TH, Brotherton PNM, Russell AF, *et al.* (2001) Cooperation, control and concession in meerkat groups. *Science* 291: 478–481.

Emlen ST and Oring LW (1977) Ecology, sexual selection, and the evolution of mating systems. *Science* 197: 215–223.

Fricke H and Fricke S (1977) Monogamy and sex change by aggressive dominance in coral reef fish. *Nature* 266: 830–832.

Gesquiere LR, Learn NH, Simao MCM, *et al.* (2011) Life at the top: Rank and stress in wild male baboons. *Science* 15: 357–360.

Gross MR (1991) Evolution of alternative reproductive strategies: Frequency-dependent sexual selection in male bluegill sunfish. *Philosophical Transactions of the Royal Society London: Biological Sciences* 332: 59–66.

Hager R and Jones CB (eds.) (2009) *Reproductive Skew in Vertebrates: Proximate and Ultimate Causes*. Cambridge, UK: Cambridge University Press.

Jarman PJ (1974) The social organisation of antelope in relation to their ecology. *Behaviour* 148: 215–267.

Jetz W and Rubenstein DR (2010) Environmental uncertainty and the global biogeography of cooperative breeding in birds. *Current Biology* 21: 72–78.

Madosky JM, Rubenstein DI, Howard JJ, and Stuska S (2010) The effects of immuno-contraception on harem fidelity in a feral horse (*Equus caballus*) population. *Applied Animal Behavior Science* 128: 50–56.

Moehlman PD (1986) The ecology of cooperation in canids. In: Rubenstein DI and Wrangham RW (eds.) *Ecological Aspects of Social Evolution*, pp. 64–86. Princeton, NJ: Princeton University Press.

Nuñez CMV, Adelman JS, and Rubenstein DI (2010) Immunocontraception in wild horses (*Equus caballus*) extends reproductive cycling beyond the normal breeding season. *Public Library of Science ONE*, 5(10): e13635.http://dx.doi.org/10.1371/journal.pone.0013635.

Packer C (1986) The ecology of sociality in felids. In: Rubenstein DI and Wrangham RW (eds.) *Ecological Aspects of Social Evolution*, pp. 429–451. Princeton, NJ: Princeton University Press.

Reeve HK and Shen S-F (2006) A missing model in reproductive skew theory: The bordered tug-of-war. *Proceedings of the National Academy of Sciences* 103: 8430–8434.

Rubenstein DI (1980) On the evolution of alternative mating strategies. In: Stadden JER (ed.) *Limits to Action: The Allocation of Individual Behaviour*, pp. 65–100. New York: Academic Press.

Rubenstein DI (1986) Ecology and sociality in horses and zebras. In: Rubenstein DI and Wrangham RW (eds.) *Ecological Aspects of Social Evolution: Field Studies of Birds and Mammals*, pp. 282–302. Princeton, NJ: Princeton University Press.

Rubenstein DI (1994) The ecology of female social behavior in horses, zebras and asses. In: Jarman P and Rossiter A (eds.) *Animal Societies: Individuals, Interactions and Organisations*, pp. 13–28. Kyoto: Kyoto University Press.

Rubenstein DI (2010) Ecology, social behavior, and conservation in zebras. In: Macedo R, (ed.) *Advances in the Study Behavior: Behavioral Ecology of Tropical Animals* Vol. 42. pp. 231–258. Oxford, UK: Elsevier Press.

Rubenstein DR and Lovette IJ (2007) Temporal environmental variability drives the evolution of cooperative breeding in birds. *Current Biology* 17: 1414–1419.

Sundaresan SR, Fischhoff IR, Dushoff J, and Rubenstein. DI (2007) Network metrics reveal differences in social organization between two fission–fusion species, Grevy's zebra and onager. *Oecologia* 151: 140–149.

Trivers RL and Willard DE (1973) Natural selection of parental ability to vary the sex ratio of offspring. *Science* 179: 90–92.

Warner RR, Robertson DR, and Leigh EG (1975) Sex-change and sexual selection. *Science* 190: 633–638.

Zabel CB and Taggart SJ (1989) Shift in red fox, *Vulpes vulpes*, mating system associated with El Niño in the Bering Sea. *Animal Behaviour* 38: 830–838.

Soil Biota, Soil Systems, and Processes

David C Coleman, University of Georgia, GA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Domains The major divisions of the biota on earth, namely: Bacteria, Archaea, and the Eucarya.

Ecosystem engineers The concept whereby members of the macrofauna (e.g., termites and earthworms) are actually moving parts of the soil volume for their own uses (e.g., making macropores which permit flow of large amounts of water rapidly through the soil).

Immobilization The process wherein nutrients are taken up or immobilized in litter and other organic detritus until later (usually weeks to months) in the decomposition process.

Larger mesofauna and macrofauna Live in pores in portions of soil profiles.

Microfauna Animals that have high turnover rates, and live in water films in soils. Small mesofauna, such as nematodes, are also water-film dwellers.

Mineralization The availability of inorganic nutrients in the decomposition of organic detritus, occurring after the immobilization phase.

Organizing centers Soils are the centers of history and activity in terrestrial ecosystems.

Soils as Components of Ecosystems

Soil-Forming Factors

Soils are an intriguing, relatively thin (often <1 m depth) zone of physical–chemical and biological weathering of the earth's land surface. Soils are formed by an array of factors, namely climate, organisms, parent material, the extent of slope, and aspect (relief) operating over time (**Figure 1**). These factors affect major ecosystem processes, such as primary production, decomposition, and nutrient cycling, which lead to the development of ecosystem properties unique to that soil type, as a result of its previous history. For example, a deep loess soil in Iowa, with a very fertile and deep surface or "A" horizon, containing considerable amounts of organic matter, will be very different from an "A" horizon developed in the Nebraska sandhills, with much greater porosity and lower water retention due to the nature of the sandy surface material. As noted in the soil-forming factors diagram (**Figure 1**), the array of biota – namely microbiota, vegetation, and consumers (herbivores, carnivores, detritivores) – is influenced by the soil processes and in turn has an impact on the soil system.

Polyphasic Nature of Soils, Influence on the Biota

Soils are perhaps the ultimate in interface media, located at the intersection of four principal entities: the atmosphere, biosphere, lithosphere, and pedosphere (**Figure 2**). Soils provide a wide range and variety of microhabitats, thus accommodating a very diverse biota. The microbes (bacteria and fungi) are found in numerous microsites, well–aerated or not; bacteria may thus respire either aerobically or anaerobically. There is an enormous amount of surface area (hundreds of m^2 per gram of soil) on the soil particles, which range in size classes from clays (0.1–2 μm in diameter) to silts (2–50 μm diameter) and sands (0.05–2 mm diameter). Numerous microbes and micro and mesofauna (protozoa and nematodes) exist in water films on these particles and in or on the surfaces of microaggregates formed from the primary particles. In turn, the more mobile fauna, from collembola and mites (larger mesofauna) to the macrofauna (earthworms, millipedes, ants, termites, and fossorial or earth-dwelling vertebrates), move through the macro and micropores in the soil. The macrofauna play a role in moving parts of the soil profile around and form many sorts of burrows and pores; they are often termed as "ecological engineers."

Soils as Organizing Centers in Ecosystems

Soils may be viewed as the organizing centers for terrestrial ecosystems. Major functions such as ecosystem production, respiration, and nutrient recycling are controlled by the rates at which nutrients are released by decomposition in the soil and litter horizons and transported to the photosynthetic layers of the ecosystem. This is particularly true for less heavily managed, near-natural ecosystems, many of which occur on soils having relatively poor nutrient status. In these systems, mycorrhizas are often obligate partners in the obtaining of adequate nutrients for the growing plants. Mycorrhizas (literally "fungus roots") are efficient at extracting nutrients from both mineral and organic sources, enabling plants to thrive in habitats that are considered to be poor in nutrients. We need to be aware of these and other mutualistic associations between microorganisms and roots, such as rhizobia and actinorrhizas, various root-associated symbiotic bacteria that facilitate the nitrogen fixation process on which the entire ecosystem often depends. These associations have arisen in soils over an evolutionary time and are key to an understanding of the ecosystem function.

Major Soil Processes

Decomposition: Immobilization and Mineralization

A very large proportion (greater than 90%) of the terrestrial net primary production is returned to the soil as dead organic

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00128-3

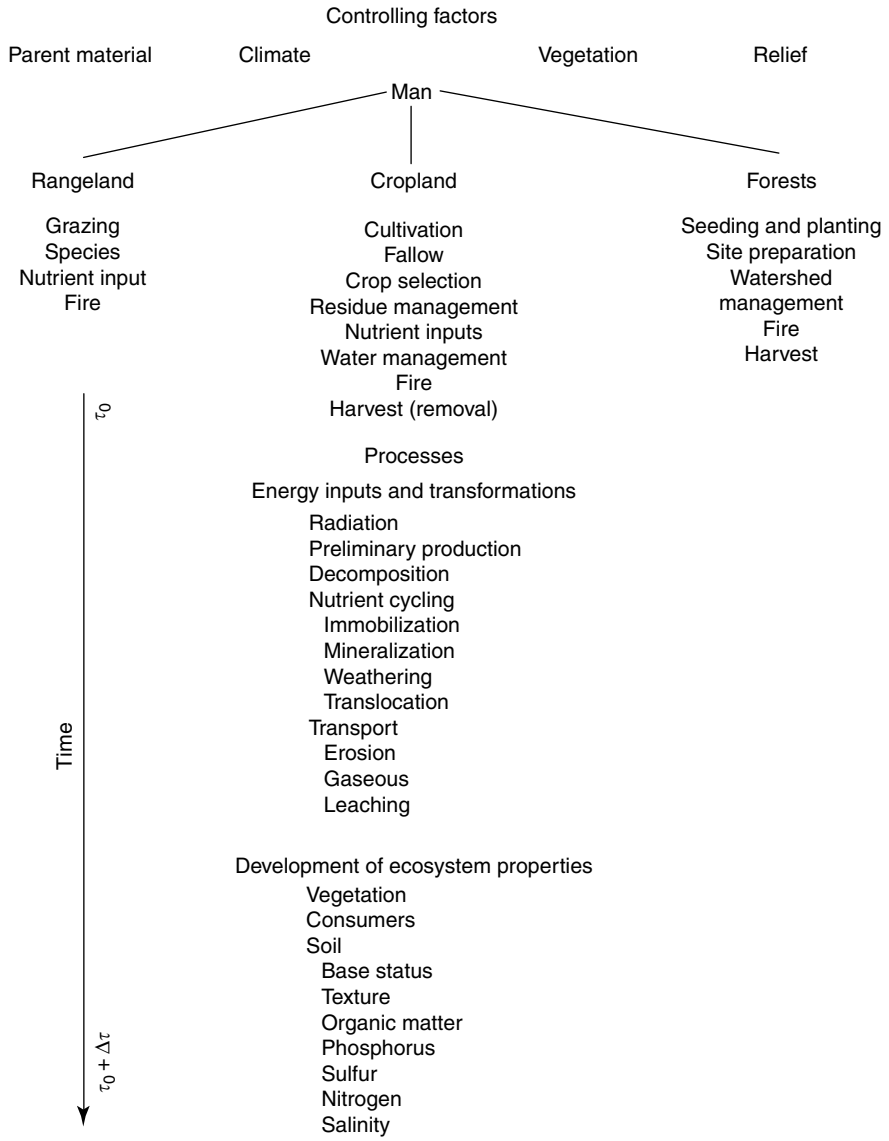


Figure 1 Soil-forming factors and processes, and the interaction over time. Reproduced from Coleman DC, Reid CPP, and Cole CV (1983) Biological strategies of nutrient cycling in soil systems. *Advances in Ecological Research* 13: 1–55; and Coleman DC, Crossley Jr. DA, and Hendrix PF (2004) *Fundamentals of Soil Ecology*. San Diego: Academic Press.

litter. This litter, consisting of leaves, roots, and wood from trees and organic residues from agricultural fields, is decomposed on or in the soil, and the nutrients contained within it are recycled for further use. The decomposition process drives complex food webs in the soil, with numerous interactions between the initial agents of decomposition, the bacteria and fungi, and the fauna which in turn feed on them.

Decomposition is the catabolism of organic compounds in plant litter and other organic detritus. Decomposition is principally the result of microbial activities; few soil animals have cellulases in their guts, which allow them to hydrolyze the celluloses in plant residues. The decomposition of organic residues involves the activities of a variety of soil biota, including both microbes and fauna, which interact conjointly in the process. For example, the initial breaking up of plant litter usually is conducted by the chewing and macerating action of both large and small animals. This initial breaking into smaller pieces, or "comminution," is a process that benefits the fauna, which derive nutritional benefit from the litter or microbes initially colonizing the plant material. The increased surface area and further inoculation of the smaller pieces enhances the microbial access to, and breakdown of, these tissues.

Nitrogen Cycle: Major Processes

Nitrogen enters the ecosystem via nitrogen fixation, in which the dinitrogen molecule (N_2) is separated into two nitrogen atoms, with considerable expenditure of energy and the assistance of the nitrogenase enzyme, to break the triple covalent bond. The atoms are ammonified and then used in the production of amino acids and proteins in the plants. Another

avenue for nitrogen entry into the soils is by lightning fixation, in which the extensive high-voltage energy in the lightning charge ruptures the dinitrogen molecule, hydrogens are attached, and then the ammonium is brought in by rainfall. As shown in **Figure 3**, nitrogen is lost from the system via harvest and erosion of organic forms of N, it can be ammonified in decomposition, and then undergoes nitrification to nitrate (NO_3^-), after which it can be taken up by biota, either through plant roots or into microbial tissues. If there is adequate energy and low amounts of oxygen present, there can be denitrification, in which the nitrogen is lost as either nitrogen gas (N_2) or nitrous oxide (N_2O). For further details, consult textbooks on ecology or ecosystem studies.

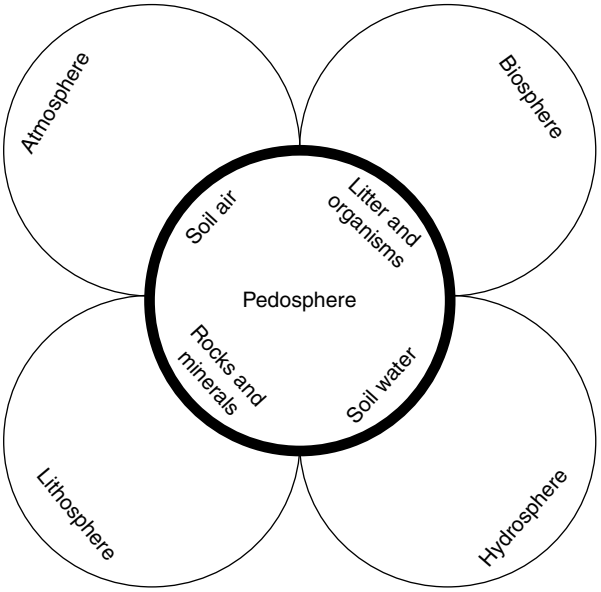


Figure 2 The pedosphere showing interactions of abiotic and biotic entities in the soil matrix. Reproduced from FitzPatrick EA (1984) *Micromorphology of Soils*. London: Chapman and Hall; and Coleman DC, Crossley Jr. DA, and Hendrix PF (2004) *Fundamentals of Soil Ecology*. San Diego: Academic Press.

The nitrogen cycle is of critical importance to biodiversity considerations, because key points in the cycle are dependent on the relatively species-poor assemblages of microbes, including the nitrogen fixation and nitrification steps. There are only a few species of nitrogen-fixing rhizobia, in the genera *Rhizobium* and *Bradyrhizobium*. The other principal nitrogen-fixing symbiont, the bacterium *Frankia* (Actinomycetaceae) forming the actinorrhiza (literally actinomycete-root), contains only a few species in the genus. However, approximately 194 plant species in eight families and four different subclasses of flowering plants have been identified as hosts. These plants share the general tendency to grow in marginal soils and play an important role as pioneer species in early successional habitats.

In the nitrification steps, noted in the previous paragraph there are only a few genera and species of nitrifiers. Most of them are autotrophic and quite sensitive to changes in soil pH. This means that these organisms may be unusually prone to getting diminished or eliminated in regions where there is considerable acid rain.

Biodiversity in Soils

Evolutionary History

Soils as we know them, with well-differentiated profiles, probably developed concurrently with the origin of a land

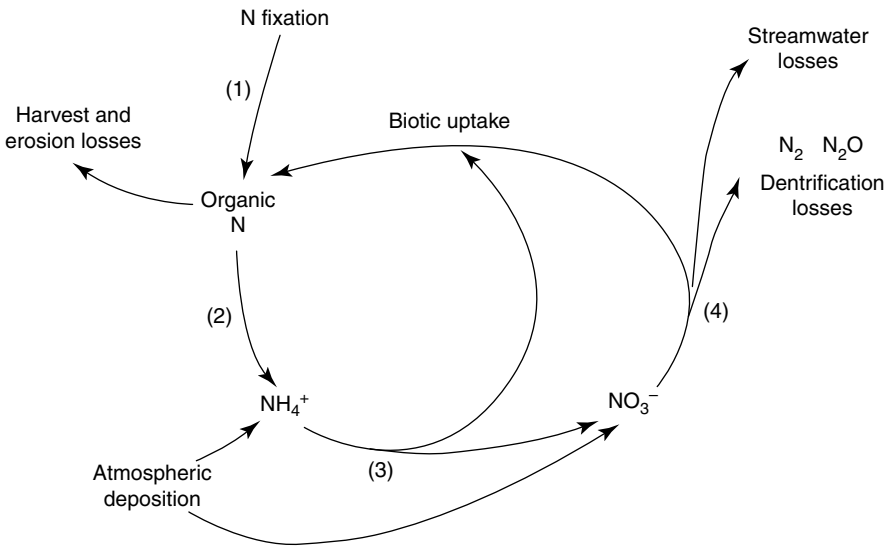


Figure 3 Inputs and outputs of N, which make up the intersystem transfers in an ecosystem. Numbers indicate groups of organisms active at a given stage of the N cycle. (1) *Rhizobia*, *Frankia*, (2) ammonifiers, (3) nitrifying bacteria, and (4) denitrifying bacteria. Denitrifying bacteria must compete with the biotic uptake of NO_3^- by other microbes and plants; thus the rate of nitrification sets an upper limit on denitrification losses. Stream water losses represent the excess of available N over that taken up by biotic processes. Modified from Waring RH and Schlesinger WH (1985) *Forest Ecosystems*. Orlando, FL: Academic Press.

flora in the early Devonian era, about 450 million years ago. The microorganisms that inhabit the soils, particularly the prokaryotic microbes such as the cyanobacteria, originated perhaps 3 billion years ago.

Diversity of Biota

Biodiversity is an inclusive concept, which encompasses a wide range of functional attributes in ecosystems in addition to being concerned with the numbers of species present in the system. This differentiates it from the concept of species diversity, which is concerned with the identity and distribution of species in a given habitat or region.

Soil biodiversity is best considered by focusing on the groups of soil organisms that play key roles in ecosystem functioning. Spheres of influence (SOI) (Coleman, 2008) of soil biota are recognized, such as the root biota, the shredders of organic matter, and the soil bioturbators. These organisms influence or control ecosystem processes, and exert further influence via their interactions with key soil biota (e.g., plants). What is the extent of redundancy within the functional groups of these SOI? Some soil organisms, such as the fungus and litter consuming microarthropods, are very speciose. For example, there are up to 170 species in one order of mites, Oribatida (members of the Arachnida, eight-legged arthropods), in the forest floor of one watershed in western North Carolina. In the soil biota, some of the species-poor functional groups, such as specialized bacteria, that is, nitrifiers and nitrogen fixers (**Figure 3**) are considered, at present, to be at most risk. Others include fungi forming mycorrhiza, a symbiotic association that benefits both plant and fungus, with the plant supplying high-quality carbon to the fungus, and the fungal hyphae exploring a greater volume of the soil, obtaining scarce mineral nutrients, particularly phosphorus. Other species-poor functional groups include macrofaunal shredders of organic matter (e.g., millipedes) and bioturbators of soils, which include various types of earthworms and termites.

Three Great ''Domains'' of Organisms on Earth

All of life exists in three great ''kingdoms,'' or domains. These domains are (1) the Bacteria (eubacteria), which are the bacteria as generally considered; (2) Archaea (archaebacteria), which include the methanogens (methane-producers), most extreme halophiles (ones living in hypersaline environments), and hyperthermophiles (ones living in volcanic hot springs, and in mid-sea ocean hot-water vents); and (3) Eucarya (eukaryotes) (**Figure 4**). The first two domains are prokaryotes, which are unicellular organisms, lacking a unit membrane-bound nucleus and other organelles, usually having their DNA in a single circular molecule. Eukaryotes, in comparison, consist of all of the organisms that have a unit membrane-bound nucleus and other organelles, such as mitochondria (Woese *et al.*, 1990). Eukaryotic organisms are often multicellular. This scheme is based on an increasing body of evidence from ribosomal RNA (rRNA) phylogenies, that the archaebacteria are worthy of the same taxonomic status as eukaryotes and bacteria. As shown in **Figure 4**, the universal rRNA tree develops from a postulated ''cenancestor,'' leading to the relative positions of the three great domains (Pace, 2009).

Number of Species of Prokaryotes

Recent estimates of the number of prokaryotic species range from 100,000 to 10 million. Interestingly, the number of species of bacteria so far described in soil amount only to about 4000. This discrepancy is largely due to the fact that only a small proportion, usually less than 1%, of the bacteria present in soil or any other medium are amenable to culturing and subsequent microscopic observation. Most taxonomic descriptions of prokaryotes are made using some variants of 16S rRNA probing, using molecular techniques.

It should be noted that, on the basis of the accepted criterion for separating taxa in microbial studies, which is greater than 70% DNA homology, a mouse and a human would be considered as being in the same species. This leads to complications, as we shall see, in discussing the total amount of genetic diversity of all organisms, including the yet largely unknown diversity of Archaea and Eubacteria. The latter now are estimated to have an array of 100 phyla, which are genetically as diverse as the phyla in Animalia, Plantae, and Fungi in older classification systems.

Biomass, Diversity, and Numbers of Bacterial Species on Earth

The Bacterial biomass and numbers are often vastly underestimated. We are currently delineating the overall genetic makeup of isolates taken from soils, which are determined by the use of molecular probes. The total numbers of bacteria on earth in all habitats are truly staggering: $4–6 \times 10^{30}$ cells, or 350–550 petagrams of carbon (Whitman *et al.*, 1998). One petagram is 10^{15} g, or 1 billion metric tons. The number of bacteria that are calculated to exist in soils is approximately 2.6×10^{29} cells, or about 5% of the total on earth. A majority of bacteria exist in oceanic and terrestrial subsurfaces, especially in the deep mantle regions, extending several kilometers below the earth's surface.

Knowledge of the diversity of archaea and bacteria is still limited, but with the use of new massive parallel sequencing devices (Roche 454 and other pyrosequencing techniques), we are beginning to approach an understanding of the very large biodiversity of these organisms. More than one full decade into the ''genomic era,'' it is necessary to consider a bacterial species as described by its pan-genome, which is comprised of a ''core genome'' containing genes common in all strains and a ''dispensable genome,'' containing genes present in two or more strains and genes unique to only one or two unique strains. In essence, the pan-genome of the bacterial domain is of nearly infinite size. For example, in eight genomes representative of the serogroup diversity among group *B Streptococcus* (GBS) strains, the pan-genome contains 2713 genes, of which 1806 belong to the core genome and 907 belong to the dispensable genome. The GBS pan-genome is predicted to grow by an average of 33 new genes every time a new strain is sequenced. However, this pattern does not hold for all bacterial species.

Metagenomics encompasses the full array of microorganisms present in an ecosystem at any point of time (Dinsdale *et al.*, 2008). Metatranscriptomics refers to the number of microbes in a given community or ecosystem that are respiring or

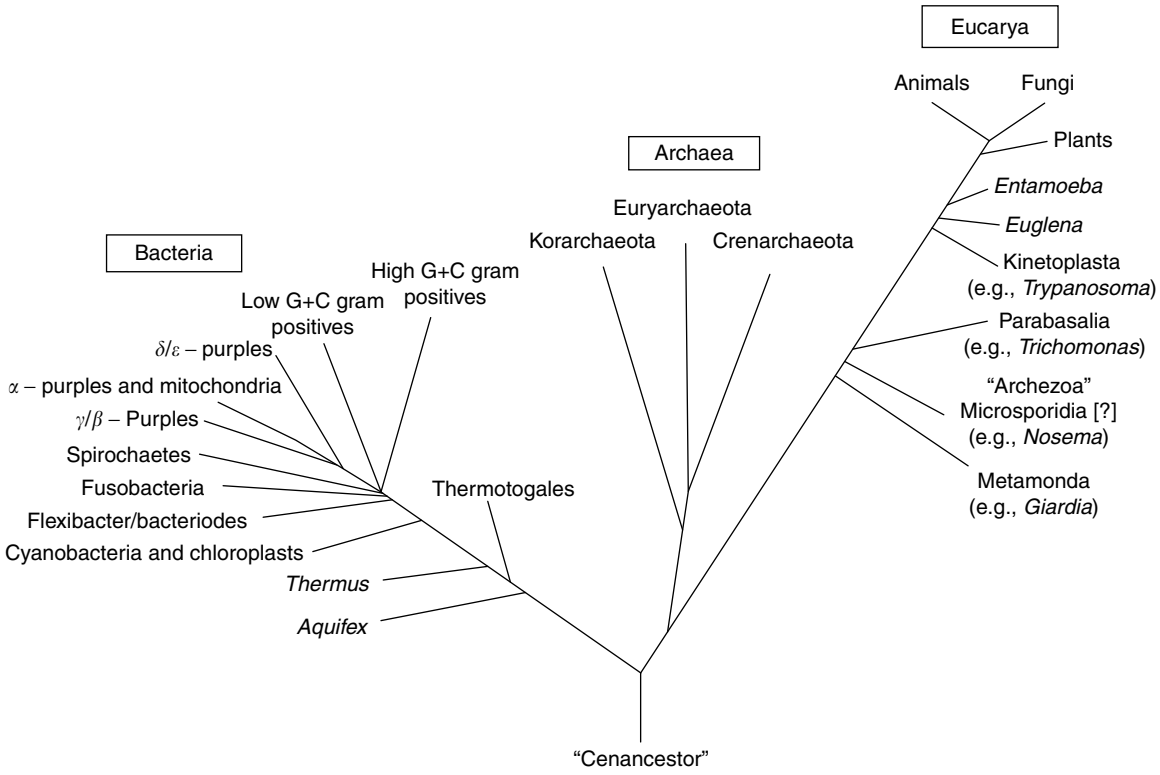


Figure 4 Schematic drawing of a universal rRNA Tree of Life showing the relative positions of groups in the domains Bacteria, Archaea, and Eucarya. The location of the root (the cenancestor) corresponds to that proposed by reciprocally rooted gene phylogenies. (Branch lengths have no meaning in this tree.)

synthesizing at a given time. At any given moment, an estimated 10^{30} bacterial and archaeal genes are mediating essential ecological processes throughout the world (Moran, 2009).

Viruses as Quasiorganisms

Viruses are quasiorganisms, not included in the three domains. Viruses are RNA or DNA molecules contained within the protein envelopes. Viral particles are metabolically inert, carrying out neither biosynthetic nor respiratory functions. They multiply only within the host cells, by inducing a living host cell to produce the necessary viral components. Once assembled, the replicated viruses escape from the cells. Viruses infect all sorts of animals, plants, and microbes. Viruses parasitizing bacterial cells are commonly called bacteriophages, or simply phages. Although little is known about the ecology of viruses, they can persist in soils for many years and decades. Some research on viruses in deserts showed that they were inactivated in soils at acid pH levels between 4.5 and 6. There is little information on the overall species diversity of viruses in soils. Current estimates are 5000 species known and perhaps 130,000 in existence.

Considerable new knowledge of virus ecology is being gained in aquatic realms: limnology and oceanography. Very little is known about viruses and their ecology in soils, despite their known effects on beneficial bacteria and soil-borne plant pathogens and horizontal gene transfer (transduction). Estimates from 5% to 25% of the carbon fixed by primary producers enter into the microbial loop via virus-induced lysis at various trophic levels in the aquatic environments (**Figure 5**). Despite the concerns about the amount of viral adsorption to soil particles, particularly clays, the greater physical and chemical diversity of soils makes it possible that viruses have an even larger potential to play roles that are similar to those in the aquatic environments.

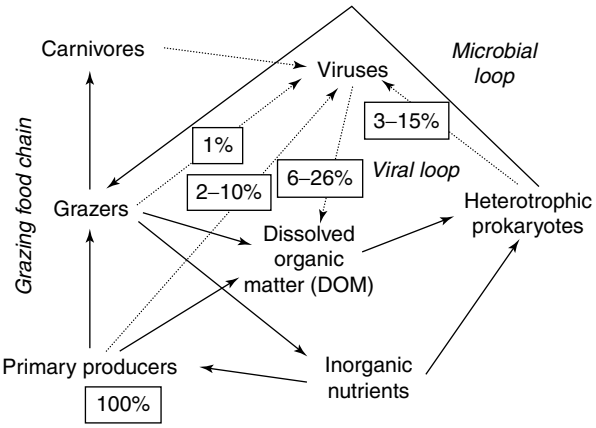


Figure 5 Pelagic food chain model and virus-mediated carbon flow (Reproduced from Kimura M, Jia Z-J, Nakayama N, and Asakawa S (2008) Review–Ecology of viruses in soils: Past, present and future perspectives. *Soil Science & Plant Nutrition* 54: 1–32, with permission from Wiley.). The dotted lines show virus-mediated pathways. All values are based on the flux of carbon fixed by primary producers (100%).

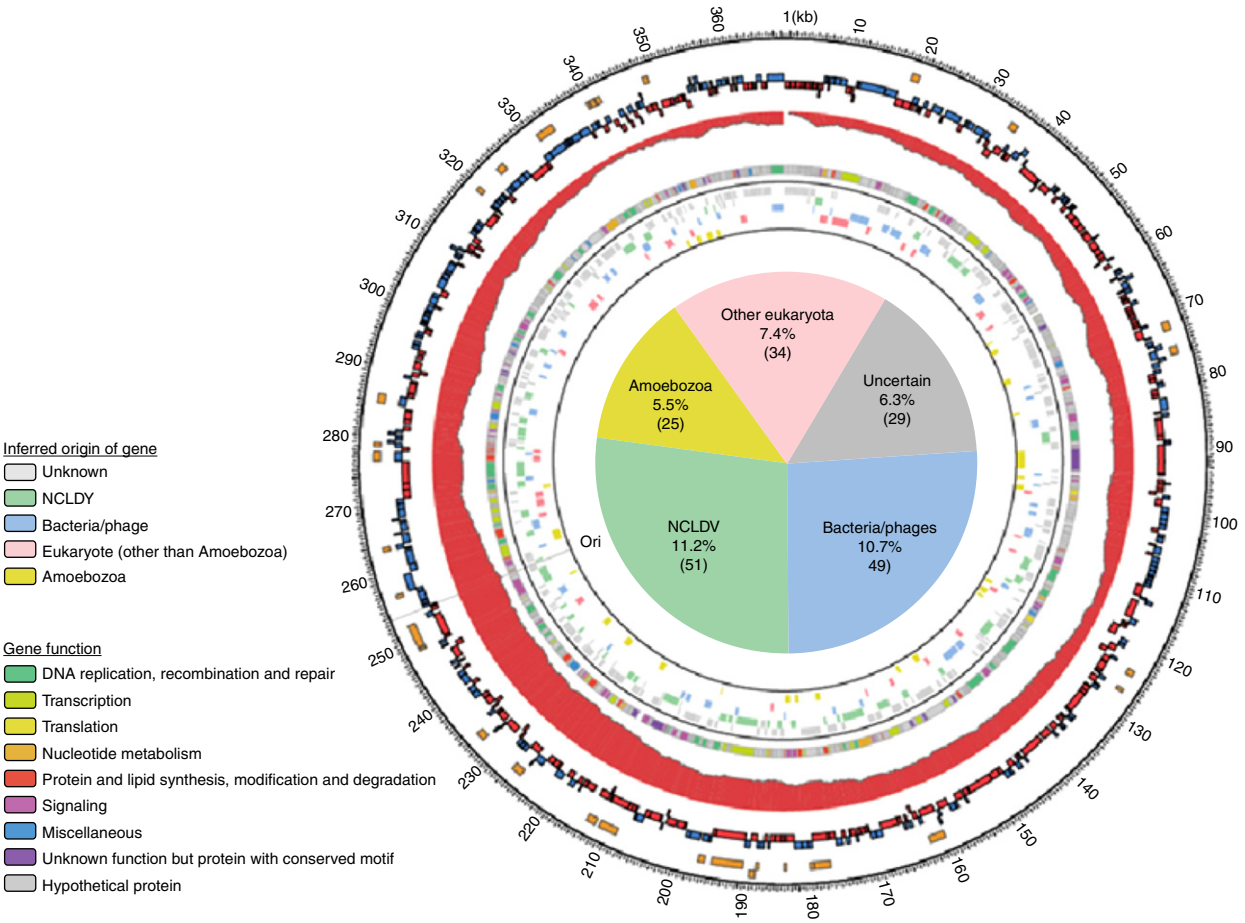


Figure 6 Map of the Marseillevirus chromosome. Rings correspond to genome coordinates in kilobases, proteins and protein-coding genes. The pie chart inside the ring represents taxonomic breakdown of Marseillevirus genes by probable origins inferred by phylogenetic analysis or sequence conservation. NCLDV, nucleocytoplasmic large DNA virus of eukaryotes. Reproduced from Kimura M, Jia Z-J, Nakayama N, and Asakawa S (2008) Review – Ecology of viruses in soils: Past, present and future perspectives. *Soil Science & Plant Nutrition* 54: 1–32.

Because there are an estimated 10^{31} phages on the planet and they can move between environments, the potential reservoir of genes that can be transferred both locally and globally by phage is enormous. There is little restriction to the types of genes carried by the viral community, suggesting that they influence a wide range of processes, including biogeochemical cycling, short-term adaptation, and long-term evolution of microbes.

Recent molecular studies of mechanisms to fend off phages and other forms of invasive DNA, including clustered regularly interspaced short palindromic repeats are increasing the ways in which we can quantify phage and prokaryotic interactions. Although the current usage of phages in soil biological control studies has focused primarily on the plant-pathogenic bacteria, the existence of consortia in which insects, bacteria, and viruses interact in a complex five-way mutualistic symbiosis is a proof that interactive complexities exist and are worth exploring. The time scale of viral evolution and the diversity of viral linkages that emerged during that extended time period would explain why the viral-specific genes probably outnumber the cellular ones in metagenomic studies, which forms a very large reservoir of molecular diversity.

As an example of virus–microbial–eucarya interactions, consider the giant Marseillevirus, a previously uncharacterized giant virus of amoebae. The genome repertoire of the virus is composed of nucleocytoplasmic large DNA viruses (NCLDV), core genes and genes apparently obtained from eukaryotic hosts (Boyer *et al.*, 2009, **Figure 6**). There is a distinct possibility that amoebae are "melting pots" of microbial evolution where diverse forms emerge, including giant viruses with complex genetic repertoires of various origins.

Numbers and Biodiversity of Eukaryotes

Fungal Diversity

Fungi are multicellular eukaryotes that are found in many habitats worldwide. They have long, ramifying strands (hyphae), which can grow into and explore many microhabitats, and are used for obtaining water and nutrients. The hyphae secrete a considerable array of enzymes, such as cellulases, and even lignases in some specialized forms, decomposing substrates *in situ*, imbibing the decomposed subunits and translocating them back through the hyphal network. Fungi are abundant, particularly in undisturbed forest floors in which literally thousands of kilometers of hyphal filaments will occur per gram of leaf litter.

Table 1 Comparison of the numbers of known and estimated total species globally of selected groups or organisms

Group	*Known species*	*Estimated total species*	*Percentage known*
Vascular plants	220,000	270,000	81
Bryophytes	17,000	25,000	68
Algae	40,000	60,000	67
Fungi	69,000	1,500,000	5
Bacteria	3000	30,000	10
Viruses	5000	130,000	4

Source: Reproduced from Hawksworth DL (1991) The fungal dimension of biodiversity: magnitude, significance and conservation. *Mycological Research* 95: 641–655.

Fungi are still little-described, with possibly less than 5% of them known to science (69,000 described; perhaps 1,500,000) in existence (**Table 1**). This is largely because of the fact that so many fungi are associated with tropical plants and animals, and these in turn have not been described.

As noted at the beginning of this section the roles of mycorrhizas in soil systems are being increasingly viewed as central to much of terrestrial ecosystem function. The total number of mycorrhizal species may be just a few thousand, but they are essential to the growth and reproduction of numerous families of plants (van der Heijden *et al.*, 1998). Recent experimental studies have noted that species richness, namely with large versus small numbers of species of Arbuscular mycorrhiza, has a positive impact on plant primary production in macrocosms of North American old fields (fields undergoing succession and not intensively managed) (Harrison, 1997).

Microfauna

The unicellular eukaryotes, or Protoctista, include a wide range of organisms, which are more often called protists, or protozoans. These include the flagellates, naked amoebae, testacea, and ciliates (**Figure 7**). These organisms range in size from a few cubic micrometers in volume to larger ciliates, which may be up to 500 μm in length and 20–30 μm in width. Protozoa are quite numerous, reaching densities of 100,000–200,000 per gram of soil. Bacteria, their principal prey, often exist in numbers up to 1 billion per gram of soil. All of these organisms are true water-film dwellers and become dormant or inactive during episodes of drying in the soil. They can exist in inactive or resting stages for literally decades at a time in very xeric environments.

About 40,000 extant protozoan species have been described, but many more undoubtedly are awaiting scientific discovery. In an extensive survey of soils from Africa, Australia, and Antarctica, in some cases nearly half of the total species described were new to science. This was particularly true in Africa, where of the 507 species identified, 240 of them, or 47%, were previously undescribed. Even in a more extensively investigated region, Australia, 43% of the total of 361 species were new to science. In Antarctica, 95 species were described, with only 14% or 15%, being unknown.

Because many habitats have not yet been investigated, and the isolation procedures are still imperfect, 70–80% of all soil ciliates may yet be unknown. This high proportion may hold true for the other protozoan groups as well.

Mesofauna

Nematodes Nematodes feed on a wide range of foods. A general trophic grouping is bacterial feeders, fungal feeders, plant feeders, and predators and omnivores. Anterior (stomal or mouth) structures can be used to differentiate general feeding or trophic groups. The feeding categories are a good introduction, but feeding habits of many genera are complex or poorly known. For example, some genera in immature phases will feed on bacteria and then become predators on other fauna once they have matured. Because of the wide range of feeding types and the fact that nematodes seem to reflect ages of the systems in which they occur (e.g., annual vs. perennial crops, or old fields and pastures and more mature forests), they have been used as indicators of the overall ecosystem condition.

As nematodes use a diverse range of food resources it is logical that their feeding, activity, and population dynamics will interact with other components of the ecosystem. For perhaps a century nematodes were almost exclusively regarded as pathogens having adverse effects on plant and animal production. Realization of their wider contributions to ecosystem processes developed in the 1970s, as did shifts toward managing rather than controlling populations of economically important parasites of plants and livestock and the use of nematodes as biocontrol agents for insects.

This is a growing area of research in soil ecology, and one in which the intersection of community analysis and ecosystem function could prove very fruitful. Current species of nematodes described total around 5000, and more than 20,000 may exist. For additional information on the activities of these diverse creatures, see Yeates (2010).

Collembola Collembolans, or "springtails," are the primitive Apterygote (wingless) insects. They are called "springtails" because many of them have a spring-like lever, or furcula, which enables them to move many body lengths away from predators by using it in a springing fashion. Collembolans are ubiquitous members of the soil fauna, often reaching abundances of 100,000 or more per square meter. They occur throughout the soil profile, where their major diet is decaying vegetation and associated microbes (usually fungi). However, like many members of the soil fauna, collembolans defy placement in exact trophic groups. Many collembolan species will eat nematodes when those are abundant. Some feed on live plants or their roots. One family (Onychiuridae) may feed in the rhizosphere and ingest mycorrhizae or even plant-pathogenic fungi.

Eight families of Collembolans occur in soils. Many Collembolans are opportunistic species, capable of rapid population growth under suitable conditions. Eggs are laid in groups. Collembolans become sexually mature with the fifth

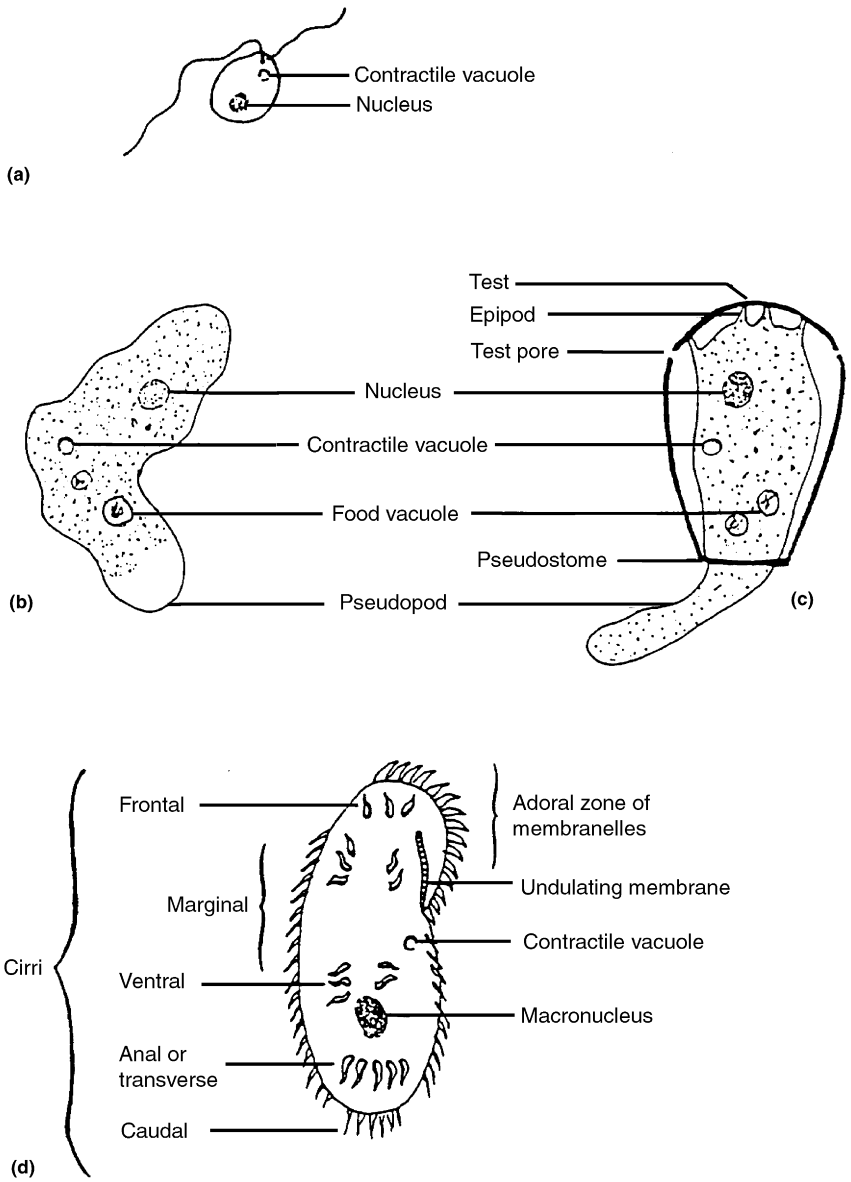


Figure 7 Morphology of four types of soil Protozoa: (a) flagellate (Bodo), (b) naked amoeba (Naegleria), (c) testacean (Hyalosphenia), and (d) ciliate (Oxytricha). Reproduced from Coleman DC, Crossley Jr. DA, and Hendrix PF (2004) *Fundamentals of Soil Ecology*. San Diego: Academic Press.

or sixth instar, but they continue to molt throughout their life. Although many species are bisexual, some of the common species are parthenogenic, consisting of females only. Collembolan "blooms" are a phenomenon of late winter or early spring, when some species may appear in large numbers on the surface of snow banks, on the surface ice of pond water, or on lichen-covered granite outcrops. There are some 6500 described species and possibly more than 10,000 in existence.

Mites (Acari) The soil mites, Acari, are chelicerate arthropods related to the spiders. They are often the most abundant microarthropods in many types of soils. A 100 g sample may contain as many as 500 mites representing nearly 100 genera. This diverse array includes participants in three or more trophic levels, with varied strategies for feeding, reproduction, and dispersal.

Four suborders of mites occur frequently in soils: the Oribatei, Prostigmata, Mesostigmata, and Astigmata. Occasionally, mites from other habitats are extracted from soil samples. These include, for example, plant mites (also called spider mites), predaceous mites normally found on green vegetation, and parasites of vertebrates or invertebrates. The most numerous ones are the true soil mites. The oribatid mites (Oribatei) are the characteristic mites of the soil and are usually fungivorous or detritivorous. Oribatid mites are very speciose, numbering 20,000 known species, with perhaps 60,000 yet to be described. A widespread assumption in the soil ecology literature has been that the oribatids feed on fungi of decomposing leaf and root litter, hence may have narrow niche overlaps in food items they consume. Recent studies in Europe have demonstrated a somewhat wider range of food items in oribatid diets that have been measured by the

stable isotope ratios ($^{15}N/^{14}N$). A total of four feeding guilds among 36 species or taxa were differentiated. In temperate deciduous and tropical forests, the Oribatids are very speciose. One notable example of high temperate zone oribatid diversity is that up to 170 species occur in just a few hundred square meters of the forest floor of a mesic, mixed deciduous forest watershed in western North Carolina. This species richness, or α-diversity (the diversity within a given site), exceeds that of some tropical forest oribatids. In a field study, species richness of oribatids declined with decreased litter species richness in experimental (1 m^2) enclosures from five down to one species of deciduous tree litter. The decline in oribatid species richness was attributed to the lower physical and chemical diversity of the available microhabitats from polyspecific to oligospecific litters, including leaf petioles, which specialized feeders such as phthiracarid ("box") mites would occupy preferentially.

Impacts of oribatid mites farther up the food chain, in ways not often considered by soil ecologists, have been demonstrated recently. In tropical America and probably in the tropics more generally, oribatids as a group contain as many as 80 alkaloids and represent a major dietary source of alkaloids in poison frogs. A conservative assumption is that they accumulate alkaloids from their food sources. To what extent do fungi or other organisms contribute alkaloids to the oribatids' diet?

Mesostigmatid mites are nearly all predators on other small fauna, although a few species are fungivores and may become numerous at times. Astigmatid mites are associated with rich, decomposing nitrogen sources and are rare except in agricultural soils. The Prostigmata contains a broad diversity of mites with several feeding habits. Very little is known of the niches or ecological requirements of most soil mite species, but some interesting information is emerging. For further details on the life-history characteristics of these interesting animals, see Coleman *et al*. (2004). About 20,000 species have been described and possibly more than 80,000 exist.

Macrofauna

Termites Termites (Isoptera) are one of the major ecosystem "engineers" particularly in tropical regions. Termites are social insects with a well-developed caste system. By their ability to digest wood, they have become economic pests of major importance in some regions of the world. Termites are arranged in five different families. The termites in a more primitive family, the Kalotermitidae, possess a gut flora of protozoans, which enables them to digest cellulose. Their normal food is wood that has come into contact with soil. Many species of termites construct runways of soil, or along root channels, and some are builders of large, spectacular mounds. Members of the phylogenetically advanced family Termitidae possess a formidable array of microbial symbionts (bacteria and fungi, but not protozoa), which enable them to process and digest the humified organic matter in tropical soils and to grow and thrive on such a diet.

Although termites are mainly tropical in distribution, they occur in temperate zones and deserts as well. Termites are often considered the tropical analogs of earthworms since they reach large abundances in the tropics and process large amounts of litter. Termites parallel earthworms in ingestive and soil turnover functions. The principal difference is that the earthworms egest much of what they ingest in an altered form (that enriches microbial action), whereas termites can transfer large amounts of soil or organic material into building nests and mounds (carbon sinks). More than 2000 species of termites have been described, and probably up to 10,000 exist.

Earthworms The earthworm fauna of North America is surprisingly poorly known, given the importance of these animals to soil processes and soil structure. Much of the evidence for earthworm effects on soil processes comes from agroecosystems and involves a small group of European lumbricids (family Lumbricidae in the order Oligochaeta). In North America, south of the southern limit of the Wisconsinan glaciation, several native genera exist. However, exotic (often peregrine European lumbricids) earthworm species have been introduced into much of this area following human population changes and colonizations. Impacts of the exotic earthworms on native species are not well understood, although there is evidence that when native habitat is destroyed and native earthworm species extirpated, exotic earthworms colonize the newly empty habitat. As more extensive studies are carried out, it is becoming clear that earthworms are present in a wide variety of tropical as well as temperate ecosystems.

Earthworms play important roles in the fragmentation, breakdown, and incorporation of soil organic matter (SOM). This affects the distribution of SOM and also its chemical and physical characteristics. Changes in any of these soil parameters may have significant effects on other soil biota, by changing their resource base (e.g., distribution and quality of SOM, microbes, or small soil fauna) or by changing the physical structure of the soil. Recent evidence indicates that earthworm activities impact the communities of other soil biota through their effects on the chemical and physical characteristics of SOM, causing changes in the oribatid species richness and microarthropod abundances. It is probable that earthworm-induced changes in the microbial and microarthropod communities will also have impacts both higher and lower in the soil food web. Some 3650 species of earthworms have been described and possibly as many as 8000 exist.

Levels of Disturbance; Effects on Biodiversity

The impact of invasive species on plants and soil organisms is a subset of a more general phenomenon: the effects of various levels of disturbance on soil biodiversity. Several recent studies have shown little support for optimization of local soil animal diversity at intermediate levels of disturbance. In fact, a wide range of disturbances of soil by tillage could strongly elevate or reduce macrofaunal group diversity, but the diversity of microfauna was little affected. Apparently the magnitude of these effects depends on soil type, climate, and tillage operation. The overall consensus seems to be that in conversion from natural vegetation to agriculture, the diversity of a wide range of microbial communities and faunal groups, ranging from nematodes to earthworms and termites, is significantly reduced in the disturbed agroecosystems. Rigorous experimental testing of the disturbance–diversity relationship is still needed to ascertain its importance as a regulator of belowground diversity. We must keep in mind the spatio-temporal

aspects of soil communities. Thus enhanced habitat complexity and diversity of available food resources will support localized "hot spots" of more diverse faunal communities. A more general aspect of disturbance studies is that related to land use and global change phenomena (Wall, 2004).

Conclusions

It is apparent that a large proportion of the biota associated with soils is as yet undescribed, with the most extreme cases being the bacteria and fungi. However, of even the somewhat more extensively studied groups, such as the oribatid mites, more than half remain unknown to science. Therefore, it is premature to give even a rough estimate of the total numbers of species that occur in many of these taxa; as such large percentages of the total number of organisms are still unknown.

When considering the diverse array of many kingdoms in all three domains of the biota, there are ample causative factors for the impressively large biodiversity observed in soils. These factors encompass many physically distinct microhabitats, organismal phenologies, numerous legacy effects, including stabilizing effects of SOM, and facilitation effects from many mutualistic interactions. In addition to the highly heterogeneous environment, a high degree of omnivory prevails in resource-use among the decomposer organisms, tending to reduce competition, and hence plays a major role in explaining the high degree of functional complementarity in decomposer communities. This is another way of considering soils as organizing centers in the terrestrial ecosystems (*see* Soils as Organizing Centers in Ecosystems).

It is incumbent on the rising generation of ecologists and biologists to develop more innovative ways to describe, catalog, and understand the myriad patterns and processes in the biosphere, which are due in large part to the actions of the biota. It is hoped that some of the observations in this article, plus the insights offered by the references listed will encourage this effort.

Appendix

List of Courses

1. Soil Ecology
2. Terrestrial Ecosystems
3. Global Biodiversity

See also: Archaea, Origin of. Eukaryotes, Origin of. Food Webs. Fungi. Nitrogen, Nitrogen Cycle. Soil Conservation

References

Bardgett RD, Usher MB, and Hopkins DW (eds.) (2005) *Biological Diversity and Function in Soils*. Cambridge: Cambridge University Press.

Boyer M, Yutin N, Pagnier I, *et al.* (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proceedings of the National Academy of Sciences of the United States of America* 106: 21848–21853.

Coleman DC (2008) From Peds to Paradoxes: Linkages between soil biota and their influences on ecological processes. *Soil Biology & Biochemistry* 40: 271–289.

Coleman DC, Crossley Jr. DA, and Hendrix PF (2004) *Fundamentals of Soil Ecology*. San Diego: Academic Press.

Dinsdale EA, Edward RA, Hall D, *et al.* (2008) Functional metagenomic profiling of nine biomes. *Nature* http://dx.doi.org/10.1038/nature06810.

Harrison MJ (1997) The arbuscular mycorrhizal symbiosis: An underground association. *Trends in Plant Science* 2: 54–60.

Kimura M, Jia Z-J, Nakayama N, and Asakawa S (2008) Review – Ecology of viruses in soils: Past, present and future perspectives. *Soil Science & Plant Nutrition* 54: 1–32.

Moran MA (2009) Metatranscriptomics: Eavesdropping on complex microbial communities. *Microbe* 4(7): 1–7.

Pace NR (2009) Mapping the tree of life: Progress and prospects. *Microbiology and Molecular Biology Reviews* 73: 565–576.

van der Heijden MGA, Klironomos JN, Ursic M, *et al.* (1998) Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 396: 69–72.

Wall DH (ed.) (2004) *Sustaining Biodiversity and Ecosystem Services in Soils and Sediments. SCOPE 64*. Washington, DC: Island Press.

Waring RH and Schlesinger WH (1985) *Forest Ecosystems*. Orlando, FL: Academic Press.

Whitman WB, Coleman DC, and Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America* 95: 6578–6583.

Woese CR, Kandler O, and Wheelis ML (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America* 87: 4576–4579.

Yeates GW (2010) *Nematodes in Ecological Webs*. Chichester: John Wiley & Sons. http://dx.doi.org/10.1002/9780470015902.a0021913.

Soil Conservation

Dorota L Porazinska, University of Florida, FL, USA
Diana H Wall, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Agroecosystem An ecosystem under agricultural management practices.
Biocontrol Control of agricultural pests by the use of predators and other beneficial organisms (e.g., control of turf grass crickets by parasitic nematodes).
Eutrophication The process by which a water body becomes enriched in dissolved nutrients (nitrogen, phosphorous) that stimulate the growth of aquatic plant life (e.g., algae), usually resulting in the depletion of oxygen from water. This frequently creates unfavorable conditions for fish and other biota.
Herbivores Animals feeding on plants only.
Mineralization Process of transformation from organic to inorganic form.
Soil food web Representation of all feeding interactions among the organisms in the soil (who eats whom).
Soil structure Spatial arrangement of soil particles.
Soil texture Percentile composition of clay, silt, and sand in soil.
Trophic Describing the feeding habits or the kind of nutrition used by a group of organisms.
Water-holding capacity Capacity of soil to hold water (e.g., sandy soils have very low water-holding capacity).

Definition of Soil

A Complex and Dynamic System

Soil is composed of living and nonliving components organized vertically, in a profile of horizontal layers or horizons. Soil is the habitat for a great abundance of diverse living organisms, many of which are microscopic (e.g., bacteria, fungi, protozoa, nematodes, and microarthropods) (**Table 1**). The nonliving component of soil consists of the solid phase, or weathered parent geological material, which contributes to the physical (e.g., soil texture, structure, density, and porosity (the space within and between aggregates)) and chemical (e.g., fertility, moisture, and acidity) properties. The formation of soil is a function of climate, parent material, biota, and topography, which through geologic time has formed a multitude of natural soil types of varying properties across geographic landscapes (Jenny, 1980). Hundreds to thousands of years may be required to form just a centimeter of soil.

Table 1 Examples of some groups of soil fauna ordered by body width and the number of described species in soil

Soil fauna	*Described species number*
Microfauna (1–130 μm body width)	
Protozoans	1500[a]
Nematodes	24,379[b]
Mesofauna (80 μm–4 mm body width)	
Mites	42,000[c]
Springtails	6284[b]
Diplurans	976[b]
Pot worms	600[a]
Termites	2955[d]
Macrofauna (1–35 mm of body width)	
Isopods	5000[a]
Centipedes	2743 [b]
Millipedes	10,000[a]
Earthworms	3600[a]
Ants	11,000[e]
Flies (larvae)	150,000[b]

[a]Wall and Virginia (2000).
[b]Hallan (2007).
[c]Halliday *et al.* (2000).
[d]Engel and Krishna, personal communication.
[e]Fernández (2003).
Only about 1–5% of soil biota have been described to species level.

Role of Soil

Fertile or productive soils provide the basis of the economic wealth of a nation by providing food, fiber, and fuel. Soil supports human civilization by supplying nutrients for growth of plants, including agricultural crops, by regulating the water flow from rainfall to groundwater, and by filtering and transfering the harmful substances that might enter the atmosphere or groundwater (Wall, 2004). Water quality is largely dependent on the filtering capacity of the living and nonliving components of the soil. Soil is a natural resource, that contributes to the foundation of our social and industrial infrastructure.

Underground systems, as fundamental constituents of all terrestrial ecosystems, influence and are influenced by the functioning (the performance) of the aboveground systems. For example, soils affect the aboveground biodiversity, ecosystem nutrient and energy cycles and fluxes, and even certain atmospheric components (e.g., global cycles of C and N). Soils are one of the largest reservoirs of global carbon and are the major terrestrial reservoir of dead organisms – plants, animals, and microorganisms. The quality (chemical composition) and quantity of dead material (organic matter) received from both the aboveground and within the soil determine the primary nutrient and energy base for the soil biota. Plant litter (dead leaves, twigs, and roots) and other dead organic matter are decayed by soil microbial and faunal assemblages (organized in detritus food webs) during decomposition, a process of transforming organic matter to inorganic compounds

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00127-1

(e.g., nitrate, ammonium, and phosphate), which supplies nutrients for plant growth. Thus, the factors affecting the quality and quantity of soil organic matter (SOM), such as climate, cultivation, invasive species, and atmospheric nitrogen (N) inputs (e.g., acid rain), can alter the diversity of above and underground organisms, the rate of decomposition, the structure of soil, and the availability of nutrients needed for plant growth.

Soil Quality

Attributes of High-Quality Soils

In the broader sense, the quality of soil refers to an ability to sustain biological productivity and the diversity of plant, animal, and human populations, and to maintain quality of water and air. High soil quality also implies an ability of soils to maintain high fertility, productivity, and resist erosion. Natural differences in the quality of soils indicate different capacities of soils to resist stress whether of natural (e.g., wind, fire, and rainstorm) or anthropogenic (e.g., plow, invasive species, and pesticides) origin. In the narrower sense, soil quality relates to the inherent combination of biotic, physical, and chemical properties that allows soils to have long-term productivity.

A number of factors are involved in defining the quality and productivity of soils: parent material (the geologic weathering of rock), climate, SOM, soil structure (aggregation of soil particles), soil stabilization, and soil biota (including roots) (Coleman *et al.*, 2004; Bardgett, 2005). One of the most important factors is SOM. The amount and type of SOM and the products of decomposition affect several soil properties. A higher content of SOM results in higher cation exchange capacity (CEC), higher water-holding capacity, higher infiltration rates, better soil aeration, and increased soil particle aggregation, all leading to improved moisture infiltration and retention, reduced runoff of nutrients, and less soil erosion. CEC refers to the ability of soil to store nutrients (e.g., calcium, magnesium, and potassium) for future plant uptake. Organic matter, soil biota abundance and diversity, plant roots, and water and air movement affect soil structure formation and aggregation (Hartel, 1998; Paul and Clark, 1996). Soil stabilization, that is the ability of soil to maintain its structural integrity when subjected to natural or anthropogenic stress, occurs when there is some degree of aggregation of soil particles. The soils then become more resistant to soil degradation (e.g., erosion, loss of fertility, reduced filtering, and buffering capacity).

Erosion

Erosion is a general term for the removal of the earth's surface by abrasive actions of wind, water, waves, and glaciers. There are two types of erosion: geological and accelerated. The Grand Canyon in the US is an example of the effects of geological erosion that occurred over millions of years ago. Geological erosion in natural ecosystems is a slow process, typically occurring at a rate slower than the rate of soil formation. This is because the soil is protected by the vegetation process. For example, aboveground vegetation can reduce the wind speed and hence the roots help in anchoring the soil. Accelerated erosion happens when the rate of soil loss is higher than the rate of soil formation. This type of erosion occurs when the lighter individual particles of soil aggregates are detached and transported by the wind or water. These particles can be blown to great distances as dust (the Dust Bowl of the 1930s in the US) and be deposited to form new soils or washed into streams, rivers, and oceans. The risk of soil erosion depends on the natural conditions (climate, slope, vegetation cover, and soil) and land use (e.g., removal of the protective cover of vegetation). Erosion results in a deterioration of soil quality and imposes hazards to humans and organisms in the terrestrial and aquatic ecosystems (e.g., loss of soil fertility, reduced plant productivity, loss of water and air quality, flooding, mudslides, and sedimentation).

Erosion is the primary factor degrading soils on a global scale (**Figure 1**). In many areas of the world, eroded land is no longer productive and is often abandoned. A recent estimate indicates the amount of arable land on earth is about 1.40 billion hectares (**Figure 2**) (Food and Agriculture Organization, 2010). In many areas, continuous intensive agricultural cultivation leads to soil erosion, salinization, and desertification and often to loss of the land from agricultural production. It is estimated that globally 25 billion metric tons of soil erodes each year from agricultural land (17 tons per cultivated hectare, or 4.5 tons per person) (Food and Agriculture Organization, 1992). To accommodate the rapidly increasing human population, global agriculture will have to increase the quantity or quality of food produced on a given area or increase cropland area into marginal areas of lower soil quality. Thus, the global efforts to conserve soils are of immediate importance for human well-being.

Water is estimated to be responsible for more than one-half of the global soil degradation, followed by wind, chemical (e.g., salinization, acidification, pollutants, atmospheric nitrogen deposition, excessive fertilizers, pesticides, and manures), and physical (soil compaction, water logging, and subsidence) degradation. However, whether a region is affected more by wind or water is dependent to a large degree on its climate. Wind erosion is a more serious problem for agricultural lands in the arid and semiarid regions of the world (e.g., North Africa, the Near East, parts of Asia, Australia, northwest China, southern South America, and North America). Salinization, the increased concentration of salts in the topsoil, can occur in the irrigated lands of the world due to increased evaporation, susceptibility of soils to salty groundwater, and invasion of seawater. These forms of soil degradation continue to decrease the availability of productive land for future food, fiber, and fuel production.

As high-quality and agriculturally productive land diminishes, other ecosystems (e.g., forests, wetlands, and meadows) become vulnerable. Over the past three centuries, 2.2 billion ha of forests (down from 4 to 6.2 billion ha worldwide) have been converted to agricultural lands. Forests generally have a high biodiversity and are a reservoir of hundreds of billions of tons of carbon stored in trees and soils. On a global scale, conversion of forests to croplands releases CO_2 into the atmosphere, contributing to the global warming. Conservation of soil is critical to the maintenance of sustainable ecosystems for the future.

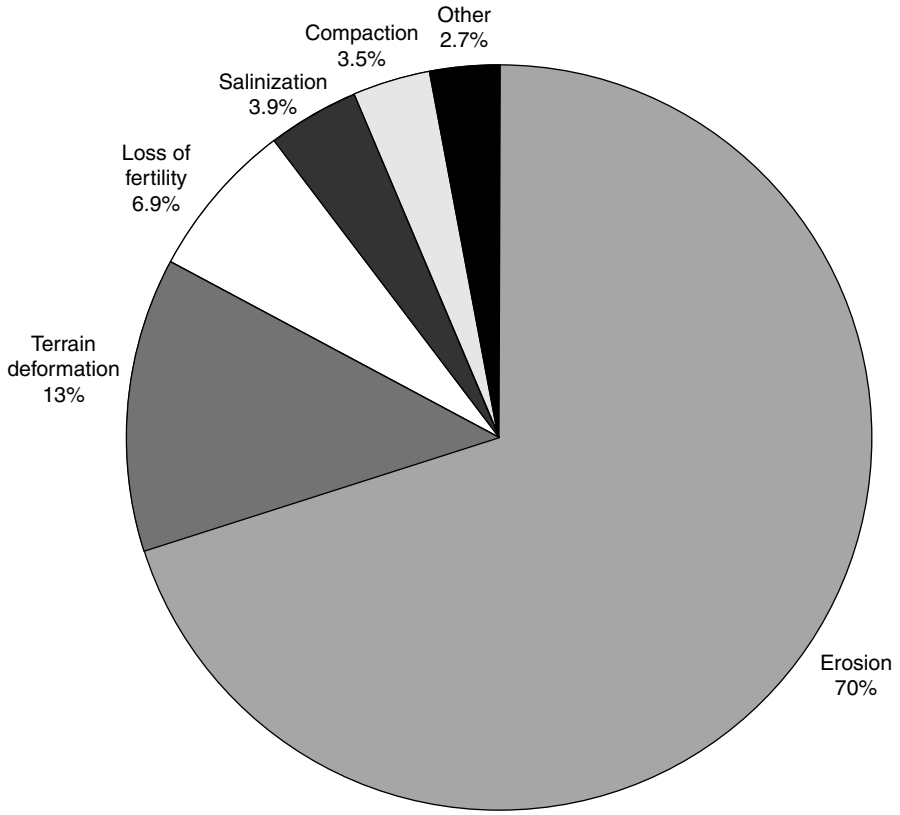


Figure 1 Causes of degradation of the world's land surface. 1995. The percentage indicates the respective impact of the various forms of degradation. The "Other" section includes pollution (1.1%), overblowing (0.6%), waterlogging (0.5%), acidification (0.3%), and subsidence (0.2%).

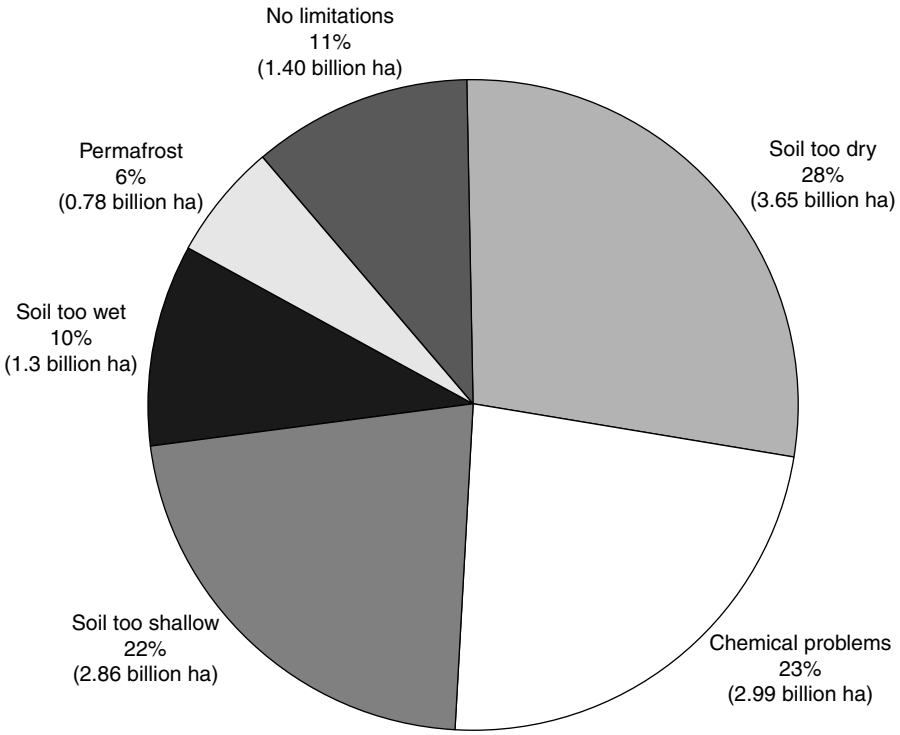


Figure 2 Total world's land surface and reasons why the entire land on earth cannot be used for the agricultural activity. The percentage indicates fraction of total land, and numbers in parenthesis indicate the actual area in hectares (ha).

Effects of Conversion of Natural Systems on Soil Quality

In natural systems, the amount of SOM is maintained by high (generally greater than 90% of primary production) litter inputs, but in agricultural and grassland grazed systems due to harvesting the yield or herbivory, only about 50% of the primary production contributes to the SOM. In agriculture, the lower amounts of organic matter input to SOM results in less substrate (energy and nutrients) for the soil food web and fewer nutrients are available to plants over time. In intensive agriculture, nutrients removed by crops or lost with erosion of topsoil have to be replaced with synthetic fertilizers in order to maintain high productivity.

Human activity changes the chemical and physical properties, biodiversity, and quality of the soil. Agriculture (including forestry and grazing), urbanization, industrialization, and climate change are the most significant modifiers of the soil quality. Tillage (plowing), one of the most successful agricultural practices, has been used for centuries to reduce weeds, aerate soils, and break up compacted soils. However, tillage along with other intensive agricultural practices, such as application of fertilizers and pesticides, use of heavy machinery, irrigation, and monocropping have resulted in degradation of soils in many regions of the world. Intensive tillage destroys soil aggregates, redistributes and enhances SOM turnover by increasing the rates of decomposition, decreases the quality and quantity of organic matter, changes soil climatic conditions, and decreases soil moisture, soil food web complexity, soil fertility, and soil stability (Paustian *et al.*, 1997). These changes increase microbial activity, organic matter oxidation, soil compaction, and water runoff. The timescale of agricultural soil degradation varies from decades to a few years (in tropical soils) to a single rainstorm. The loss of topsoil depends on the original status of the soil quality, farming practices, climatic events, and the amount of topsoil (Jenny, 1980).

Construction (e.g., building of dams, urban development) and deforestation also contribute to an increased rate of erosion and the loss of soil quality. Local effects of soil degradation are manifested on a landscape scale (watersheds, rivers, wetlands, or oceans) through surface and groundwater pollution and soil sedimentation, impacting the quality of water and air, and resulting in economic impacts on the tourism and fishing industry.

Conservation and Restoration

Methods of Soil Conservation in Agriculture

The rate of soil degradation presently exceeds the rate of soil formation in systems impacted by humans (Jenny, 1980, 1984). All the vegetative productivity depends to a great extent on the nutrient availability in the soil, but in managed systems, such as intensive agricultural systems with continuous monocultures, a disproportionate amount of nutrients are removed from the system as crop yield. Although synthetic fertilizers renew the nutrient pool necessary for plant growth, they do not regenerate the quantity or the heterogeneity in the quality of SOM. Thus, slower decomposing organic materials (e.g., fertilizers such as manures) are increasingly used to replenish SOM. Crop harvesting also leaves the soil bare and exposed to erosion by the rain and wind. Thus, the diminishing quality of the soil resources in natural and managed ecosystems, and the increasing demand for food, fiber, and fuel supply, has resulted in a reevaluation and selection of land management strategies that enhance longevity and quality of the soil ecosystems. These alternatives focus on maintenance of SOM, soil fertility, soil biodiversity, soil structure, and soil stabilization, and a reduction of soil erosion as a means of sustaining ecosystems for the long term.

The methods to reduce soil erosion are directed toward protecting the surface of the soil (topsoil) and include tillage practices (reduced tillage, conservation tillage, no till, minimum till or zero tillage – techniques that use specially designed machines and herbicides for minimal impact on the soil); diversified farming or cover crops; polycultures; adding organic composts or mulches; crop rotations; plowing techniques, such as contour farming (plowing at right angles to the land to create ridges to hold water); terracing (leveling areas on a slope to prevent water runoff); and timing of plowing.

The methods of agricultural soil conservation have proven valuable whether used alone or in combination. In the US, soil conservation methods have decreased the soil erosion rates so that only one-third of the agricultural lands are eroding faster than the average rate of soil formation. Although this is a considerable improvement over the Dust Bowl days of the 1930s in the US, soil conservation still appears to be necessary. A recent summary of the long-term experiments has evaluated the methods of conservation tillage intensity and crop management used alone or in combination on soil carbon storage. The combination of reduced tillage, bare fallow, increased inputs of crop residue, and crop rotations (with use of perennial vegetation) was more efficient at restoring the SOM and soil carbon and reducing levels of soil erosion than any of these management methods used separately. The rates of erosion were decreased in Creek Basin, Wisconsin, over the past 140-year period (historical and current data) because of the improvements in local agricultural land management from 1975 to 1993. These studies provide evidence supporting the efficacy of soil conservation practices.

Reduced Tillage

One of the most important factors decreasing the soil quality is tillage. Reduced tillage practices that incorporate crop residues retention in the soil are among the best alternatives for conventional tillage and have increased globally over the past 10 years. Although reduced tillage practices increase SOM and soil moisture and contribute in many cases to an improved soil food web, they may require a greater use of herbicides due to the establishment and spread of weeds and soil pathogens. In some situations, crop yields may decrease due to the presence of the weeds and pathogen diseases.

Conventional and reduced tillage systems appear to be economically comparable, particularly if crop rotation is used in the reduced tillage system. The reduced tillage system, however, promotes greater long-term benefits for ecosystems locally and globally. Soil biotic complexity is generally affected positively by the retention of crop residues, and the rate of decomposition generally changes from a microbial to a fungal

pathway resulting in a slower pulsed release of nutrients for plant uptake. The retention of soil residues also results in a reduction in CO_2 emissions to the atmosphere as compared to the conventional tillage, and a long-term conservation of beneficial chemical and physical properties of soil.

In conventional high-input agroecosystems, high crop productivity is achieved by the application of synthetic fertilizers (mainly nitrogen, potassium, and phosphorus), not by decomposition of SOM. As SOM diminishes in conventionally tilled systems, N and P retention is reduced and significant amounts of added fertilizers (nitrogen and phosphorus) are lost from the system. Fertilizer runoff and nutrient leaching reduces surface and groundwater quality. At a regional and global scale, nutrients from agricultural fields are lost to the atmosphere or lead to the eutrophication of coastal waters with effects on coral reefs, estuaries, and fisheries. For example, in the Mississippi River, the nutrient levels have doubled or tripled since 1950s, resulting in large algal blooms that deplete oxygen from the water, killing fish and shrimp, as well as other organisms.

In contrast, in most reduced tillage systems, the quantity of nitrogen applied as fertilizer is reduced, and there is an increase in the amount of applied nitrogen that becomes immobilized in the soil's organic matter. The nitrate leaching potential is reduced and subsequently the risk of water pollution is lowered. However, increased earthworm populations co-occurring with the reduced tillage practices can result in an increased soil porosity and higher nitrate leaching.

Diversified Cropping

Organic matter can also be conserved in soils by growing cover crops, termed diversified or multiple cropping, during fallowing. The advantages of cover crops are numerous. They not only reduce soil nitrogen leaching (nitrates become immobilized in plant biomass), but also provide a relatively effective means of weed control. They also decrease erosion since the plants and not the soil intercept the raindrops. Cover crops are selected based on several attributes, including plant species that improve soil (e.g., depth of roots, type of organic matter produced, and legumes that fix atmospheric nitrogen), their effect on plant pathogens and their predators, or for their economic benefits. Although multiple cropping requires an extensive knowledge of the type of crop, and in what sequence to grow cover crops, the benefits have been recognized for centuries. Multiple cropping systems are increasingly being used as a tool in modern agriculture. Today's research provides a scientific support for the diversified agricultural systems as they maintain soil quality (SOM, soil fertility), low insect and disease occurrence, and high plant productivity.

Organic Amendments

Organic materials or mulches in many instances offer a long-term option for maintaining soils with higher SOM in long-term agricultural production. Composts provide sources of nutrients that promote plant productivity and soil quality. Other benefits include improved soil aggregation, soil aeration, water-holding capacity, and CEC. Examples of organic amendments include green manures (herbaceous crops plowed under while green), chicken and cow manure, pig slurry, urban grass cuttings, and home garden composts of leaves, grass, and food remains. When conformed to quality standards set by national environmental agencies, composts such as those derived from the municipal and industrial wastes have the potential to improve soil quality and at the same time reduce the conversion of land to landfills (land conservation). Currently, however, the supply of compost does not meet demand, and with the higher cost of compost as compared to synthetic fertilizers, the use of organic amendments is not an economical option for many growers in the industrialized nations. An increase of consumer interest in the purchase of organically grown food products may increase future use of composted organic materials.

Bioremediation of Pollutants

SOM, soil structure, fertility, and stability – and thus the soil quality – can be affected by pollutants. Chemical may be introduced to the soil system deliberately (e.g., fertilizers and pesticides) or accidentally (via spills or failures of technological processes). Depending on the final concentration, the presence of many hazardous and toxic chemicals can have an immediate or long-term effect on soil quality. Restoration of polluted soils through bioremediation not only avoids the risk of human and animal health hazards, but also restores, to some degree, the degraded land. Bioremediation, unlike other methods, takes advantage of soil biota or plants to detoxify the contaminants in the soil.

Soil is a natural habitat to diverse microorganisms and other soil biota that assist in the reduction of the soil contaminants. Microorganisms, for example, may use the pollutant as a substrate or energy source for their metabolism and change the composition of the compound to a less harmful chemical. Pollutants can be changed from harmful to environmentally safe compounds through natural bioremediation (use of indigenous microflora), biostimulation (addition of nutrients to soil to stimulate activity of indigenous microflora), bioventing (addition of gases such as oxygen or methane to stimulate activity of microflora), bioaugmentation (inoculation of soil with exogenous microorganisms), landfarming (mixing healthy and toxic soil), and phytoremediation (use of plants) (Skipper, 1998).

Bioremediation offers an interesting and economical alternative to the conventional way of cleaning soil contaminants. The entire treatment occurs at the polluted site as opposed to the more traditional method of moving large amounts of soil to treatment facilities and then back to the field. Bioremediation is often cheaper, resulting in savings of energy used to remove and transport soil, and typically has a lower environmental impact, as it depends on the natural processes with no hazardous byproducts (Skipper, 1998). The major constraints of bioremediation apply to the nature of the pollutants, as microorganisms cannot access many new synthetic compounds. Bioremediation can be further complicated by soil contaminants that are mixtures of pollutants, degradation of which may require the presence of several types of organisms each capable of transforming a different chemical in the mixture.

Manipulating the Biotic Community

Another aspect of soil quality relates to the diversity of biotic communities. Native land transformation, agricultural

management, and pollution significantly affect, in most cases negatively, the abundance and diversity of organisms (Foissner, 1999; Bardgett, 2005). The array of multiple plant species (trees, shrubs, grasses) seen in native systems contribute to the soil stabilization because the heterogeneity of rooting distribution, rooting depth, and plant chemical composition help to maintain soil structure and a diverse biotic community. In conventional agricultural systems, monocropping or cropping of one plant species (if not one variety) provides homogeneous root morphology, root depth, and litter quality across the landscape. The differences between the underground heterogeneity of natural systems and the underground homogeneity imposed by the monoculture agriculture must be considered when developing the long-term soil conservation plans. Using polycultures or intercropping is an agricultural practice that mimics, to some degree, the underground heterogeneity of natural systems. An unresolved question is whether creating or restoring the structure of soil by this practice will actually recreate and restore the function provided by the soil biota.

The abundance and diversity of microflora and fauna in soil provides beneficial services to humans and ecosystems, including, but not limited to, nutrient mineralization, biological control, nitrogen fixation, soil carbon sequestration, soil aggregation, and soil stabilization (Wall, 2004). In natural systems, the decomposition of organic matter (plant litter, dead roots, animals) to inorganic chemicals necessary for plant growth involves many different groups of soil organisms. Microflora (bacteria and fungi) influence the amount of C stored in the soil and are responsible for the global cycling of many minerals (e.g., N, C, S, P) (Gregorich *et al.*, 1997). Microflora play a critical role by aggregating the soil particles, which helps to prevent soil erosion. For example, they produce chemicals (e.g., polysaccharides) that bind soil particles, or have morphological structures (e.g., fungal hyphae) that connect the soil particles. Microfauna (e.g., protozoa, nematodes) affect soil fertility by trophic interactions with microflora that increase available N and P for plants. Nematodes, for instance, by grazing on bacteria can alter bacterial abundance and activity and thus significantly affect the rate of organic matter turnover and nutrient availability. Invertebrates transport microflora throughout the soil and in addition to functions similar to microfauna, mesofauna (e.g. mites and springtails) feed on fungi and other small fauna (Brussaard *et al.*, 1997). Macrofauna (e.g., earthworms, termites, ants, and snails) comminute and redistribute organic matter within the soil profile and by burrowing affect soil physical properties (Lee and Foster, 1991). They have been termed "soil engineers" for their role in mixing soils and microflora.

Tillage, monocropping, and addition of fertilizers and pesticides often reduce the diversity of many components of soil biotic communities, limiting at the same time their potential positive effects on many ecosystem processes. Improvement of soil quality will depend on the integrated management practices that reduce the artificial energy inputs (fertilizers) and exploit the food webs of soil organisms. Maintenance or exploitation of soil biodiversity for beneficial human services may require a significant change in intensive conventional tillage. Reduced tillage, for instance, promotes more diverse decomposer communities (often with a higher fungal proportion) and lower mineralization rates (reduction of N losses). A long-term addition of organic matter (compost or no till) promotes microbial and faunal activities through which restoration of nutrients and SOM are possible. However, considerable research is needed to determine how these practices will affect the food web. Just as all soils differ in their composition, food webs vary with soil, climate, and the quality and quantity of organic matter. Therefore, different organic composts cannot be expected to create the same food web in all the soils.

Another way to maintain soil biodiversity and underground heterogeneity is the development of agricultural mosaics. A traditional agricultural process (still common in the tropics and many European countries) promotes a patchy landscape with the use of strip weed margins, hedgerows, and shelterbelts. The latter two types create vertical structures to reduce wind, and also provide microhabitats and refugia for many beneficial soil organisms. A huge area with a monocropping system increases the risk of pest outbreaks. Among many characters of agroecosystems with low pest potential are diversified crops in time and space, crop rotations, a structural mosaic of cultivated and uncultivated lands, the presence of perennial crops, and high crop genetic diversity. All these elements not only stimulate the development of favorable habitats for biological control agents, but they also provide a reservoir for recolonization of organisms involved in decomposition, detoxification, and other renewal processes.

Conservation of Urban Soils

Human population growth has been followed by urbanization. Conversion of the native land for housing, parking, roads, industry, landfills, mining, and so on, not only reduces potentially arable land but also affects global nutrient cycles (C and N). The additional consequences of urbanization include an increased use of the fertilizers and pesticides in home lawns and gardens. These contribute to the soil and water pollution, decreased soil biotic complexity, and increased nitrogen runoff. Alongwith the urban development it is important to adopt urbanization strategies that are environmentally safe. Proposed methods for environmental urbanization, including vertical rather than horizontal sprawl, smaller house acreage, and preservation of open spaces between developed areas are already being socially accepted and implemented. On a smaller scale, planning home gardens based on the local climate and natural vegetation is becoming more popular.

Assessment and Monitoring of Soil Quality

The goal of providing future generations with an opportunity for a high quality of life cannot be detached from the preservation or improvement of environmental quality. In order to evaluate the status or change of the environment, assessment, monitoring, and regulation programs are being developed. On the basis of monitoring data, farming strategies, urban development techniques, and industrial technologies are being modified to prevent further degradation of a particular system. Typically, to evaluate the quality of the soil system, or its state of sustainability, the system attributes are compared to the native or minimally disturbed reference sites. This procedure can provide an initial estimate of human

impact (e.g., agriculture, urbanization, and pollution) and through monitoring (an assessment of a soil ecosystem's attributes through time) give us a perspective on the positive or negative aspects of a management strategy. To be effective, monitoring has to be connected with the process of management decisions; therefore adjustments or modifications of management strategies should accompany these decisions (a process termed "adaptive management").

Indicators

Indicators are the measures that tell us about the status of the environment over time. Monitoring of, for example, air and water quality as well as CO_2 concentrations are officially regulated by many countries. There have been many indicators proposed for monitoring the soil quality status, but it has been difficult to select a single indicator that would aid policy makers.

The quality of soil can be defined in terms of many variables or combinations of variables. These variables include items such as biodiversity, levels of specific mineral elements and pollutants, levels of primary productivity, or profitability. Some aspects of high-quality soils may contradict each other. For instance, higher levels of nitrogen in the soil may stimulate primary production but may suppress biodiversity. The choice of an indicator is further complicated by the spatial heterogeneity of soils. Different soil systems may require a different suite of indicators. In addition, soil has many functions and maintaining those functions may have different impacts at different scales. For example, continuous cultivation may increase local crop productivity, but may deplete carbon stored in the soil, contributing to the globally elevated atmospheric CO_2. Theoretically, some optimal condition involving a combination of different aspects of the soil quality can be achieved.

Physical Attributes of Soil as Indicators

The ability of soil to provide mechanical support for plant growth, diversity of life forms, and agricultural activities is another measure of the soil quality. The most important physical attributes of potentially indicative functions of soil quality are soil porosity, soil stability, water storage capacity, aeration, and soil structure. Any of these measures would indicate a change in the soil quality, but might not indicate other changes that are important to the ecosystem function.

Soil porosity determines how well liquids, gases, and heat can be stored and transmitted within the soil matrix. A greater porosity often indicates greater storage and transmission ability. Soils have greater stability and resistance to natural and anthropogenic stress when the structural integrity of the soil is intact. Soil stability decreases with tillage, the use of heavy equipment, and heavy rainfall or irrigation, because soil aggregates are destroyed and greater compaction and erosion occurs. Soil water storage capacity is particularly important for the biotic component (plants, microbes, fauna) of soil ecosystems. The best plant performance occurs in soils almost saturated with water and declines as the soil dries. Many of the soil fauna are the aquatic organisms that require a film of water around soil particles for reproduction and mobility. The abundance and diversity of soil organisms can also be negatively affected by extremes in soil moisture.

Soil porosity and soil water content are directly related to soil aeration. Aeration is considered as an ability of the soil to store and transmit gases, particularly oxygen and carbon dioxide, as these are the components of roots, microbes, and fauna activity. Low porosity and waterlogged soils might induce an inadequate aeration and negatively affect the plant growth, root activity, soil biodiversity, and trace gas flux (methane).

These physical soil characteristics are inherently related to the soil structure. The extent, size, and shape of aggregates can strongly affect physical, chemical, and biological properties of the soil. Typical indicators of good soil structure are well-established aggregates with a desirable porosity, water-holding capacity, and aeration, resulting in more favorable environments for biodiversity and plant productivity as well as resistance to degradation.

Chemical Attributes of Soil as Indicators

In terms of chemical attributes, the major functions of soil are to store and supply sufficient amounts of nutrients to sustain an ecosystem's primary productivity and immobilize or detoxify hazardous compounds that are toxic to plants and animals. Among the chemical properties of soil the most useful as indicators of soil quality are mineralogy, organic matter content, CEC, salinity, and pH (acidity or alkalinity). Soil mineralogy determines bioavailability (adsorption and precipitation) and mobility of nutrients in the soil solution and is dependent on the content and type of silicate clay minerals and the content of Fe and Al oxide and hydrous oxide minerals.

SOM (or soil carbon) is usually the most important factor affecting the soil quality and productivity. SOM has a direct influence on physical and chemical soil characteristics and an indirect effect on plant production, and it is the major energy source for the abundance and diversity of soil organisms. SOM increases water retention and decreases runoff by preventing the sealing of the soil surface and promoting infiltration (and less erosion). When bound to the soil particles as aggregates, it improves soil structure. The upper part of the soil typically has the maximum SOM, generally about 1–10%, compared to 1% or less at greater depths. Mineral soils with high SOM are generally the most productive, and the carbon contained in the SOM is an important component of the global carbon cycle. Because of the positive correlation between high levels of organic matter and desirable attributes of mineral soil, it has been suggested as a reliable measure of soil quality. Even though changes in SOM are slow and vary locally and regionally, SOM is reliable as a long-term, rather than short-term, indicator of changes in soils.

CEC is another useful indicator of changes in soil quality. In general, a higher content of SOM and clay particles has a positive effect on the CEC. The soil pH can also significantly influence nutrient availability. For instance, Ca, Mg, and K deficiencies are often present in acidic soils, with Fe and Zn deficiencies appearing in alkaline soils. Trace elements and heavy metals may become more bioavailable in acidic soils. Plant productivity not only depends on the storage and supply of nutrients, but also on the ability of the soil to detoxify pollutants. Phytotoxicity and bioaccumulation may occur in areas under the strong influence of human activities (e.g., fuel

burning, mining, and industry) particularly in soils deficient in organic matter and low in CEC and pH.

Biological Attributes of Soil as Indicators

The environmental impact of human activities can be indicated, as outlined earlier, by the changes in the physical and chemical properties of the soil ecosystem. Despite their usefulness in characterization of the general status and quality of the soil, physical and chemical indicators have limitations. Do we know the biologically important thresholds of soil chemicals? Do we know how different sets of physical soil factors affect the bioavailability of soil chemicals? Are they stable enough to permit the detection of environmental change? Are they easily interpretable? How useful are the indicators across regional and global scales? These and other questions have led many scientists to examine other indicators.

As the biological attributes of the soil system directly relate and respond to the physical and chemical attributes, soil biota have been investigated as indicators of soil quality. Many organisms are sensitive to the changes in the physical and chemical properties of the soil–air or soil–water interfaces. Research shows that disturbances affect the soil biota differently depending on their physiology and life histories. Specific groups of organisms are associated with the soil at spatial (vertical vs. horizontal, rhizosphere vs. bulk soil) and temporal scales. Moreover, biota vary in their geographic distribution, therefore those groups of organisms that occur in all the soils (nematodes, protozoa, and microbes) are promising as indicators on a global scale, whereas those geographically limited groups (e.g., earthworms) may be more valuable on a local scale. The use of soil invertebrates as bioindicators might be problematic because the taxonomic and life history diversity of soil fauna requires extensive knowledge and training. As with the physical and chemical indicators of the soil, no single bioindicator has yet been proposed.

Microbes, protozoa, nematodes, mites, and earthworms are the most noted indicators for soil quality monitoring. Any changes in the community structure of soil biota can be based on a direct analysis of the dynamics of all identified taxa, or on data derived from taxonomic or ecological indices. Measures for comparisons of soil communities include traditional morphological data (species, genus, family, or functional group level), or more recently, molecular data (Porazinska *et al.*, 2010a, b; Wu *et al.*, 2009) As identification of soil organisms according to the species or genus level is difficult, a higher level of taxonomic hierarchy or functional group level is often applied. To relate the status of soil quality to the ecosystem change, the ecosystem processes driven by soil biota (e.g., decomposition, mineralization, and respiration) can be measured.

Many of the smaller organisms in soil share a number of attributes that are necessary for useful indicators. These include high abundance for ease of monitoring, occurrence in all soils (microbes, fungi, protozoa, and nematodes), sensitivity to changes in soil chemistry and pollutants (microbes, fungi, protozoa, nematodes, mites, and earthworms), sensitivity to soil physical changes (most taxa), and representation of a wide range of groups in the soil food web (nematodes, mites). Both protozoa and nematodes have been used to monitor the effects of pesticides and heavy metals. The nematode community structure has also been used to illustrate the effects of natural ecosystem succession, environmental disturbance, and land management practices. Limitations to both protozoa and nematodes as indicators include difficulties in the enumeration and identification of species and their applicability at larger geographic scales. Recent advances in diagnostics using massive sequencing molecular data, however, might remove these limitations (Porazinska *et al.*, 2010a). Although common indices for abundance, diversity, richness, or rate of a particular process for different ecosystems may not exist, the applicability of bioindicators at a local scale might provide a powerful long-term tool for the soil conservation efforts.

Policies and Regulations

Adoption of soil conservation measures by farmers may not always be an easy choice. Surveys of farmers reveal that a conservation ethic is a less effective motivation for adoption of soil conservation practices than demonstration of economic benefits. However, farmers knowledgeable about erosion and soil degradation are more likely to adopt soil conservation practices than uninformed farmers. Thus, education and access to information are important aspects of any conservation effort, as are regulations and policies that conserve or sustain the land.

When private incentives differ from societal incentives, the influence of government policies and regulations can have a major impact on managing the natural resources. In the US, 25% of arable land is regulated with the beneficial effects for soil conservation. Conservation tillage in the US was part of the 1985 Food Security Act (FSA), which encouraged farmers to take erosion-control measures by providing farm subsidy payments. The US Food, Agriculture, Conservation, and Trade Act of 1990 (amendment to FSA) established financial penalties and ineligibility for most farmer program subsidies to farmers who produced agricultural crops on wetlands that were converted after enactment. This act also established the Conservation Reserve Program, providing an opportunity for the farmers to take highly erodible lands out of production by receiving annual rental payments from their 10-year contracts with the Department of Agriculture. The Federal Agriculture Improvement and Reform Act of 1996 introduced ''planting'' flexibility, giving farmers entering commodity programs freedom to choose crops on the contracted acreage. Under the same act, soil erosion control and wetland restoration regulations were improved. These few examples illustrate that government policies can slow land degradation.

The effects of land cultivation and soil erosion expand beyond the border of an agricultural field, state, or even country. Separate agricultural ecosystems are connected via a network of groundwater, streams, and rivers. Silt, sediments, nutrients, and other agricultural pollutants that are transported to streams, rivers, and ultimately marine systems can restrict the possibilities for navigation, irrigation, food production, and fisheries, and can affect water and air quality. Thus, the local ecological impacts of agricultural practices have been recognized globally. This has resulted in national and international policy frameworks including organizations such as: the Food and Agriculture Organization (FAO), the International Geosphere and Biosphere Program – Global Change in Terrestrial Ecosystems (IGBP – GCTE), the International Union of Soil

Science (IUSS), Global Assessment of Soil Degradation (GLASOD), and International Panel on Climate Change (IPCC), all of which have been active in either research, communication, or assistance to countries requiring immediate soil conservation measures and agricultural improvement.

There are also many international agreements whose policies include or are directed at ensuring soil sustainability. Agenda 21 is the action plan signed at the 1992 United Nations Conference on Environment and Development in Rio de Janeiro by the leaders from 169 countries. Agenda 21 devotes an entire article to sustainable agriculture and rural development both at national and international levels. Among many others, it aims for policies on land reform, less environmentally destructive developments, and conservation and rehabilitation of soil resources. Another example is the Convention on Biodiversity, which encourages agricultural practices that sustain biodiversity and ecosystem functioning. As our knowledge increases about the components of soil and their complexity, we realize that soil is a rare natural resource that should be used wisely.

Summary

Soils provide the basis for the world's food, fiber, and fuel production, as well as for the functioning of global ecosystems. They provide a habitat for a diversity of species comparable to the diversity of life above ground. Human activity changes the chemical and physical properties and biodiversity of the soil. Agriculture, urbanization, and industrial development are the primary activities that are accelerating the loss of soil quality. As ecosystems are so interconnected, degradation of soils has detrimental effects on all life. Practices exist that can be selected to conserve and allow the sustainable use of soils.

The most effective methods improving the quality of agricultural soils are reduced tillage, cropping mosaics, multiple cropping, crop rotations, and the application of organic amendments. Not only do they prevent further loss of SOM and soil carbon (anWashington, DC: National Academy Press important property of soil environments), but they also prevent loss of other favorable physical and chemical soil properties. To determine the status of soil environment, assessment and monitoring programs are being developed. Soil quality can be estimated with the use of various indicators reflecting physical, chemical, or biological soil attributes. The collected information serves as a basis for evaluation and modification of current management methods. Different soil ecosystems may require different suites of indicators.

When private incentives differ from societal, government policies and regulations can have a pronounced impact on managing soil resources. Since the problem of soil degradation crosses spatial scales (local, regional, and global), development and implementation of national or international policies is essential.

See also: Agriculture, Sustainable. Agriculture, Traditional. Ecology of Agriculture. Greenhouse Effect. Soil Biota, Soil Systems, and Processes

References

Bardgett RD (2005) *The Biology of Soil: A Community and Ecosystem Approach*. Oxford: Oxford University Press.

Brussaard L, Behan-Pelletier VM, Bignell DE, *et al.* (1997) Biodiversity and ecosystem functioning in the soil. *Ambio* 26: 563–570.

Coleman DC, Crossley Jr DA, and Hendrix PF (2004) *Fundamentals of Soil Ecology*. San Diego, CA: Elsevier Academic Press.

Fernández F (2003) *Introducción a las hormigas de la región neotropical*. Bogotá: Instituto Humboldt.

Foissner W (1999) Soil protozoa as bioindicators: Pros and cons, methods diversity representative examples. *Agriculture, Ecosystems and Environment* 74: 95–112.

Food and Agriculture Organization (FAO) (1992) *Protect and Produce: Putting the Pieces Together*. Rome: FAO.

Food and Agriculture Organization (FAO) (2010) Statistics Division, FAO Resource Page, http://www.fao.org/economic/ess/ess-publications/ess-yearbook/essyearbook2010/yearbook2010-reources/en/

Gregorich EG, Carter MR, Doran JW, Pankhurst CE, and Dwyer LM (1997) Biological attributes of soil quality. In: Gregorich EC and Carter MR (eds.) *Soil Quality for Crop Production and Ecosystem Health*, pp. 81–113. New York: Elsevier.

Hallan J (2007) Synopsis of the described animalia of the world. http://insects.tamu.edu/research/collection/hallan/0ReportHi.htm

Halliday RB, O'Connor BM, and Baker AS (2000) Global Diversity of Mites. In: Raven PH and Williams T (eds.) *Nature and Human Society: The Quest for a Sustainable World : Proceedings of the 1997 Forum on Biodiversity*, pp. 192–212. Washington, DC: National Academy Press.

Hartel PG (1998) The soil habitat. In: Sylvia DM, Furman JJ, Hartel PG, and Zuberer DA (eds.) *Principles and Applications of Soil Microbiology*, pp. 21–44. Upper Saddle River, NY: Prentice Hall.

Jenny H (1980) The Soil Resource: Origin and Behavior. *Ecological Studies* Vol. 37. New York: Springer-Verlag.

Jenny H (1984) The making and unmaking of fertile soil. In: Jackson W, Berry W, and Coleman B (eds.) *Meeting the Expectation of the Land*, pp. 44–55. San Francisco, CA: North Point Press.

Lee KE and Foster RC (1991) Soil fauna and soil structure. *Australian Journal of Soil Resources* 29: 745–775.

Paul EA and Clark FE (1996) *Soil Microbiology and Biochemistry*. San Diego, CA: Academic Press.

Paustian K, Collins HP, and Paul EA (1997) Management controls on soil carbon. In: Paul EA (ed.) *Soil Organic Matter in Temperate Agro ecosystems: Long Term Experiments in North America*, pp. 15–49. Boca Raton, FL: CRC Press.

Porazinska DL, Giblin-Davis RM, Esquivel A, Powers TO, Sung W, and Thomas WK (2010b) Ecometagenetics confirms high tropical rainforest nematode diversity. *Molecular Ecology* 19: 5621–5630.

Porazinska DL, Sung W, Giblin-Davis RM, and Thomas WK (2010a) Reproducibility of read numbers in high-throughput sequencing analysis of nematode community composition and structure. *Molecular Ecology Resources* 10: 666–676.

Skipper HD (1998) Bioremediation of contaminated soils. In: Sylvia DM, Furman JJ, Hartel PG, and Zuberer DA (eds.) *Principles and Applications of Soil Microbiology*, pp. 469–481. Upper Saddle River, NY: Prentice Hall.

Wall DH (2004) *Sustaining Biodiversity and Ecosystem Services in Soil and Sediments*. Washington, DC: Island Press.

Wall DH and Virginia RA (2000) The world beneath our feet: Soil biodiversity and ecosystem functioning. In: Raven PR and Williams T (eds.) *Nature and Human Society: The Quest for a Sustainable World*, pp. 221–245. Washington, DC: National Academy Press.

Wu T, Ayres E, Li G, Bardgett RD, Wall DH, and Garey JR (2009) Molecular profiling of soil animal diversity in natural ecosystems: Incongruence of molecular and morphological results. *Soil Biology and Biochemistry* 41: 849–857.

South American Natural Ecosystems, Status of

Philip M Fearnside, National Institute for Research in the Amazon (INPA), Brazil

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bioregion One of six biogeographic divisions of South America consisting of contiguous ecoregions. Bioregions are delimited to better address the biogeographic distinctiveness of ecoregions.

Ecoregion A geographically distinct assemblage of natural communities that share a large majority of their species and ecological dynamics, share similar environmental conditions, and interact ecologically in ways that are critical for their long-term persistence.

Ecosystem A set of interacting living and nonliving components in a defined geographic space. Ecosystems include both plant and animal communities and the soil, water, and other physical elements of their environment.

Major ecosystem type Groups of ecoregions that share minimum area requirements for conservation, response characteristics to major disturbance, and similar levels of β diversity (i.e., the rate of species turnover with distance).

Major habitat type Groups of ecoregions that have similar general structure, climatic regimes, major ecological processes, β diversity, and flora and fauna with similar guild structures and life histories.

Original Extent of Terrestrial Ecosystems

Ecosystems can be classified in many ways, making the number of categories vary widely depending on the use intended. Here, the system adopted by Dinerstein *et al.* (1995) is used. This divides the continent into 95 terrestrial "ecoregions," exclusive of mangroves. These are grouped into four "major ecosystem types:" tropical broadleaf forests, conifer/temperate broadleaf forests, grasslands/savannas/shrublands, and xeric formations. Within each of these categories are varying numbers of "major habitat types," such as tropical moist broadleaf forests. These are further divided into nine "bioregions." Amazonian tropical moist forests, for example, is a bioregion.

The 95 ecoregions, with their hierarchical groupings, are presented in **Table 1**. Also included are the ratings for conservation status, biological distinctiveness, and biodiversity priority derived by Dinerstein *et al.* (1995). This study made a systematic survey of the status of natural ecosystems in Latin America and the Caribbean (LAC) and applied a uniform methodology to assigning priorities to these ecosystems for conservation efforts. The work was done for the United States Agency for International Development (USAID) by the WWF–US Biodiversity Support Program (BSP). The document is based on three workshops, plus consultations with relevant organizations and individual experts (the list of contributors contains 178 names).

The classification system is hierarchical, starting with four "major ecosystem types" (e.g., Tropical Broadleaf Forests), which are divided into 10 "major habitat types" (e.g., Tropical Moist Broadleaf Forests). These are crossed with six bioregions (e.g., Amazonia) and divided into 95 ecoregions (e.g., Rondônia/Mato Grosso moist forests). The system allows the priority of some ecoregions to be promoted upward based on uniqueness and regional representation, even if indicators of diversity and vulnerability are not so high.

The effort was unusual in emphasizing protection of areas with high β diversity (a measure of the turnover of species along ecological gradients), as well as the more commonly used α diversity (species diversity within a habitat). In the case of mangroves, the diversity assessed is ecosystem diversity, including aquatic animal life. This avoids mangroves receiving the unjustly low diversity ratings that tend to result when assessments are restrained to terrestrial organisms, especially trees.

Although the ecoregions identified in **Table 1** refer to "natural" (pre-Columbian) ecosystems, it should be emphasized that these had already been subject to millennia of influence by indigenous peoples prior to the arrival of Europeans. This influence continues today, together with much more rapid alterations from such activities as deforestation and logging done by nonindigenous residents. "South America" is taken to include the three Guianas (different from usage by the Food and Agriculture Organization of the United Nations (FAO)) and to exclude Panama (however, in the case of ecoregions that extend into Panama, the area estimates in **Table 1** include the Panamanian portions). The ecoregions are mapped in **Figure 1**. The ecoregion numbering corresponds to **Table 1** and also to the report by Dinerstein *et al.* (1995); the numbering presented here is not continuous, since the report also includes ecoregions in Mexico, Central America, and the Caribbean. Extensive bibliographic material on the delimitation of the ecoregions and on the state of knowledge about them can be found in Dinerstein *et al.* (1995).

Mangroves occur along the coasts of Brazil, the three Guianas, Venezuela, Colombia, Ecuador, and northern Peru. Dinerstein *et al.* (1995) divide them into five complexes: Pacific South America, Continental Caribbean, Amazon–Orinoco–Maranhão, Northeast Brazil, and Southeast Brazil. Each complex is further subdivided into 2–5 units, corresponding to distinct segments of coastline. Mangroves are essential to maintain populations and ecological processes in surrounding marine, freshwater, and terrestrial ecosystems.

Present Extent of Terrestrial Ecosystems

Unfortunately, information is not available on the present extent of each of the 95 ecoregions listed in **Table 1**.

Table 1 Terrestrial ecoregions of South America

Major ecosystem type	*Major habitat type*	*Bioregion*	*Ecoregion name*	*Ecoregion number*	*Countries*	*Original area (km²)*	*Conservation status*[a]	*Biological distinctiveness*[b]	*Biodiversity priority*[c]
Tropical broadleaf forests	Tropical moist broadleaf forests	Orinoco tropical moist forests	Cordillera La Costa montane forests	17	Venezuela	13,481	3	2	I
			Orinoco Delta swamp forests	18	Venezuela, Guyana	31,698	4	3	III
			Guianan Highlands moist forests	20	Venezuela, Brazil, Guyana	248,018	5	2	III
			Tepuis	21	Venezuela, Brazil, Guyana, Suriname, Colombia	49,157	5	1	II
			Napo moist forests	22	Peru, Ecuador, Colombia	369,847	4	1	I
		Amazonian tropical moist forests	Macarena montane forests	23	Colombia	2366	3	2	I
			Japurá/Negro moist forests	24	Colombia, Venezuela, Brazil	718,551	5	1	II
			Uatumã moist forests	25	Brazil, Venezuela, Guyana	288,128	4	3	III
			Amapá moist forests	26	Brazil, Suriname	195,120	4	3	III
			Guianan moist forests	27	Venezuela, Guyana, Suriname, Brazil, French Guiana	457,017	4	3	III
			Paramaribo swamp forests	28	Suriname	7760	3	3	III
			Ucayali moist forests	29	Brazil, Peru	173,527	2	1	I
			Western Amazonian swamp forests	30	Peru, Colombia	8315	4	1	I
			Southwestern Amazonian moist forests	31	Brazil, Peru, Bolivia	534,316	4	1	I
			Juruá moist forests	32	Brazil	361,055	5	2	III
			Várzea forests	33	Brazil, Peru, Colombia	193,129	3	1	I
			Purús/Madeira moist forests	34	Brazil	561,765	4	4	IV
			Rondônia/Mato Grosso moist forests	35	Brazil, Bolivia	645,089	3	2	II
			Beni swamp and gallery forests	36	Bolivia	31,329	4	4	IV
			Tapajós/Xingu moist forests	37	Brazil	630,905	3	4	IV
			Tocantins moist forests	38	Brazil	279,419	2	4	III

	Northern Andean tropical moist forests	Chocó/Darién moist forests	39	Colombia, Panama, Ecuador	82,079	3	1	I
		Eastern Panamanian montane forests	40	Panama, Colombia	2905	2	1	I
		Northwestern Andean montane forests	41	Colombia, Ecuador	52,937	2	1	I
		Western Ecuador moist forests	42	Ecuador, Colombia	40,218	1	2	I
		Cauca Valley montane forests	43	Colombia	32,412	1	1	I
		Magdalena Valley montane forests	44	Colombia	49,322	1	1	I
		Magdalena/Urabá moist forests	45	Colombia	73,660	2	3	II
		Cordillera Oriental montane forests	46	Colombia	66,712	3	1	I
		Eastern Cordillera Real montane forests	47	Ecuador, Colombia, Peru	84,442	3	1	I
		Santa Marta montane forests	48	Colombia	4707	3	2	I
		Venezuelan Andes montane forests	49	Venezuela, Colombia	16,638	2	1	I
		Catatumbo moist forests	50	Venezuela, Colombia	21,813	1	4	III
	Central Andean Tropical moist Forests	Peruvian Yungas	51	Peru	188,735	2	1	I
		Bolivian Yungas	52	Bolivia, Argentina	72,517	2	2	I
		Andean Yungas	53	Argentina, Bolivia	55,457	3	3	III
	Eastern South American tropical moist forests	Brazilian Coastal Atlantic forests	54	Brazil	233,266	1	1	I
		Brazilian interior Atlantic forests	55	Brazil	803,908	2	2	I
Tropical dry broadleaf forests	Orinoco tropical dry forests	Llanos dry forests	74	Venezuela	44,177	2	4	III
	Amazonian tropical dry forests	Bolivian Lowland dry forests	76	Bolivia, Brazil	156,814	1	1	I
	Northern Andean tropical dry forests	Cauca Valley dry forests	77	Colombia	5130	1	4	III
		Magdalena Valley dry forests	78	Colombia	13,837	1	4	III
		Patiá Valley dry forests	79	Colombia	1291	1	4	III
		Sinú Valley dry forests	80	Colombia	55,473	1	4	III
		Ecuadorian dry forests	81	Ecuador	22,271	1	1	I
		Tumbes/Piura dry forests	82	Ecuador, Peru	64,588	2	1	I
		Marañion dry forests	83	Peru	14,921	2	3	II

(*Continued*)

Table 1 Continued

Major ecosystem type	*Major habitat type*	*Bioregion*	*Ecoregion name*	*Ecoregion number*	*Countries*	*Original area (km²)*	*Conservation status*[a]	*Biological distinctiveness*[b]	*Biodiversity priority*[c]
			Maracaibo dry forests	84	Venezuela	31,471	2	4	III
			Lara/Falcón dry Forests	85	Venezuela	16,178	2	4	III
		Central Andean tropical dry forests	Bolivian montane dry forests	86	Bolivia	39,368	1	3	II
Conifer/ temperate broadleaf forests	Temperate forests	Southern South American temperate forests	Chilean winter rain Forests	87	Chile	24,937	2	2	I
			Valdivian temperate forests	88	Chile, Argentina	166,248	3	1	I
			Subpolar *Nothofagus* forests	89	Chile, Argentina	141,120	3	3	III
	Tropical and subtropical coniferous forests	Eastern South American tropical and subtropical coniferous forests	Brazilian *Araucaria* forests	105	Brazil, Argentina	206,459	1	3	II
Grasslands/ savannas/ shrublands	Grasslands, savanas, and shrublands	Orinoco grasslands, savannas and shrublands	Llanos	110	Venezuela, Colombia	355,112	4	3	III
		Amazonian grasslands, savannas, and shrublands	Guianan savannas	111	Suriname, Guyana, Brazil, Venezuela	128,375	4	3	III
			Amazonian savannas	112	Brazil, Colombia, Venezuela	120,124	4	3	III
			Beni savannas	113	Bolivia	165,445	2	3	II
		Eastern South American grasslands, savannas, and shrublands	Cerrado	114	Brazil, Paraguay, Bolivia	1,982,249	3	1	I
			Chaco savannas	115	Argentina, Paraguay, Bolivia, Brazil	611,053	3	2	I
			Humid Chaco	116	Argentina, Paraguay, Uruguay, Brazil	474,340	3	4	IV
			Córdoba montane savannas	117	Argentina	55,798	3	4	IV
		Southern South American grasslands, savannas, and shrublands	Argentina monte	118	Argentina	197,710	4	3	III
			Argentina Espinal	119	Argentina	207,054	4	3	III
			Pampas	120	Argentina	426,577	2	3	III
			Uruguayan savannas	121	Uruguay, Brazil, Argentina	336,846	3	3	III

	Flooded grasslands	Orinoco flooded grasslands	Orinoco wetlands	128	Venezuela	6403	4	3	III
		Amazonian flooded grasslands	Western Amazonian flooded grasslands	129	Peru, Bolivia	10,111	4	3	III
			Eastern Amazonian flooded grasslands	130	Brazil	69,533	3	3	III
			São Luis flooded grasslands	131	Brazil	1681	2	4	III
		Northern Andean flooded grasslands	Guayaquil flooded grasslands	132	Ecuador	3617	2	3	II
		Eastern South American flooded grasslands	Pantanal	133	Brazil, Bolivia, Paraguay	140,927	3	1	I
			Paraná flooded savannas	134	Argentina	36,452	2	3	II
	Montane grasslands	Northern Andean montane grasslands	Santa Marta paramo	137	Colombia	1329	3	1	I
			Cordillera de Mérida paramo	138	Venezuela	3518	4	1	I
			Northern Andean paramo	139	Ecuador	58,806	3	1	I
		Central Andean montane grasslands	Cordillera Central paramo	140	Peru, Ecuador	14,128	3	1	I
			Central Andean wet puna	141	Bolivia, Argentina, Peru, Chile	183,868	3	2	I
			Central Andean wet puna	142	Chile	188,911	3	2	I
			Central Andean dry puna	143	Argentina, Bolivia, Chile	232,958	3	2	I
		Southern South American montane grasslands	Southern Andean steppe	144	Argentina, Chile	198,643	4	4	IV
			Patagonian steppe	145	Argentina, Chile	474,757	3	2	I
			Patagonion grasslands	146	Argentina, Chile	59,585	3	3	III
Xeric formations	Mediterranean scrub	Central Andean Mediterranean scrub	Chilean matorral	148	Chile	141,643	2	1	I
	Deserts and xeric shrublands	Orinoco deserts and xeric shrublands	La Costa xeric shrublands	168	Venezuela	64,379	2	4	III
			Arayua and París xeric Scrub	169	Venezuela	5424	2	3	II
		Northern Andean deserts and xeric Shrublands	Galapagos Islands xeric scrub	170	Ecuador	9122	3	1	I
			Guajira/Barranquilla xeric scrub	171	Colombia, Venezuela	32,404	2	3	II
			Paraguná xeric scrub	172	Venezuela	15,987	2	3	II
		Central Andean deserts and xeric shrublands	Sechura Desert	173	Peru, Chile	189,928	3	3	III
			Atacama Desert	174	Chile	103,841	3	3	III

(*Continued*)

Table 1 Continued

Major ecosystem type	*Major habitat type*	*Bioregion*	*Ecoregion name*	*Ecoregion number*	*Countries*	*Original area (km²)*	*Conservation status*[a]	*Biological distinctiveness*[b]	*Biodiversity priority*[c]
		Eastern South American deserts and xeric shrublands	Caatinga	175	Brazil	752,606	3	3	III
	Restingas	Northern Andean restingas	Paranaguá restingas	176	Venezuela	15,987	2	2	II
		Amazonian restingas	Northeastern Brazil restingas	177	Brazil	10,248	1	1	I
		Eastern South American restingas	Brazilian Atlantic coast restinga	178	Brazil	8740	1	1	I

[a]Conservation status codes: 1, critical; 2, endangered; 3, vulnerable; 4, relatively stable; 5, relatively intact.

[b]Biological distinctiveness codes: 1, globally outstanding; 2, regionally outstanding; 3, bioregionally outstanding; 4, locally important.

[c]Biodiversity priority codes: I, highest priority at regional scale; II, high priority at regional scale; III, moderate priority at regional scale; IV, important at national scale.

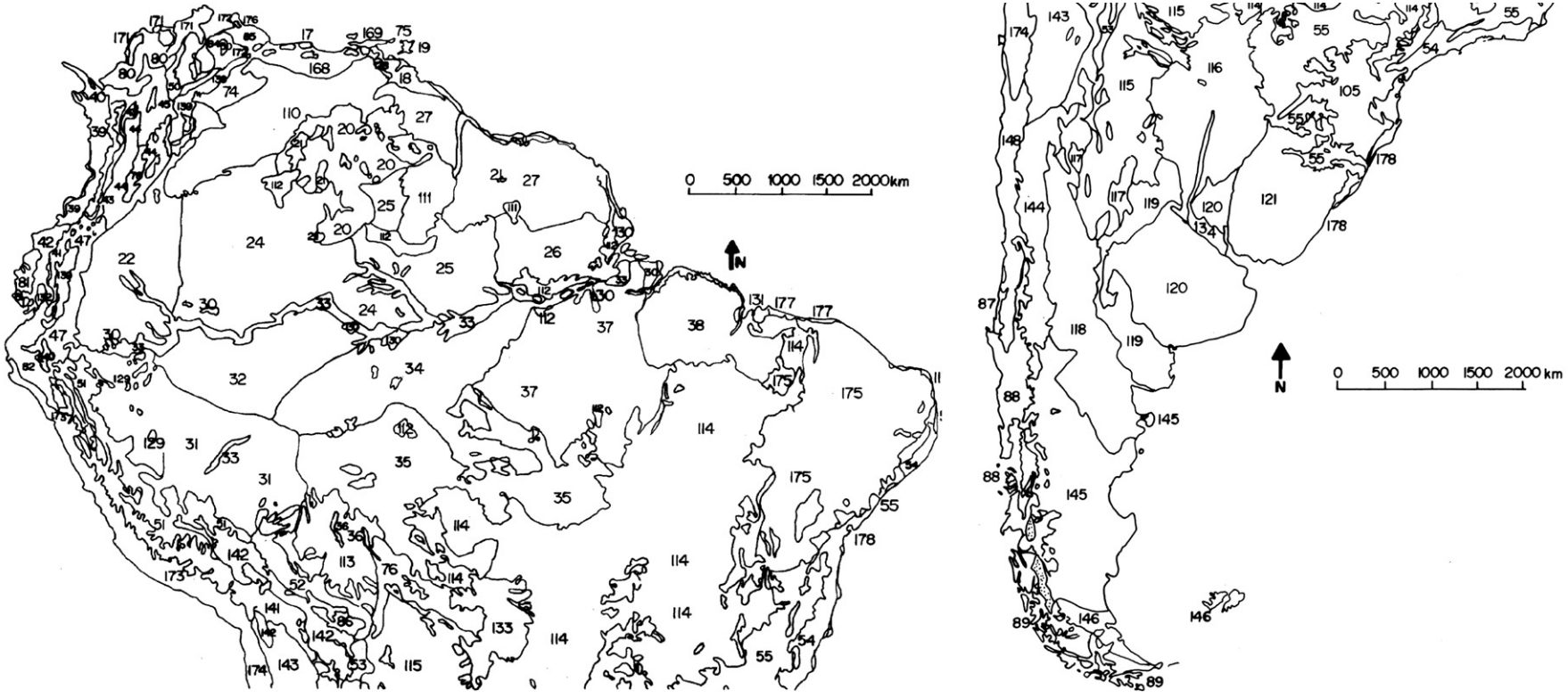


Figure 1 Ecoregions for pre-Columbian vegetation of South America. Numbers correspond to **Table 1**.

Table 2 Area of tropical forest present in 1990 (km^2)

Country	*Tropical rain forests*	*Moist deciduous forest*	*Dry deciduous forest*[a]	*Very dry forest*	*Desert Desert*	*Hill and montane forest*	*All forests*[a]
Bolivia	0	355,820	73,460	0	40	63,850	493,170
Brazil	2,915,970	1,970,820	288,630	0	0	435,650	5,611,070
Colombia	474,550	41,010	180	0	0	24,900	540,640
Ecuador	71,500	16,690	440	0	0	31,000	119,620
French Guiana	79,930	30	0	0	0	0	79,970
Guyana	133,370	31,670	0	0	0	19,120	184,160
Paraguay	0	60,370	67,940	0	0	270	128,590
Peru	403,580	122,990	190	2690	1840	147,770	679,060
Suriname	114,400	33,280	0	0	0	0	147,680
Venezuela	196,020	154,650	2220	1	0	103,900	456,910
Total	4,389,320	2,787,330	433,060	2691	1880	826,460	8,440,870

[a]Includes cerrado, caatinga, and chaco.

Information on the extent of tropical forests in approximately 1990 is available from the FAO Tropical Forest Resources Survey (FAO, 1993). These data are tabulated by country in **Table 2**. More recent FAO reports (e.g., FAO, 2010a) provide national data for forests and for woodlands, but without distinguishing between the various groups of ecosystems (such as tropical forest). Forest types are separated in national reports for some countries, including Brazil (FAO, 2010b). Nontropical areas are covered by a variety of national surveys (Harcourt and Sayer, 1996). National data are important because decisions regarding land-use policies and conservation are taken at the national level – not at the levels of bioregions or ecosystem types. Over half of the South American continent is represented by a single country: Brazil (**Figure 2**).

An idea of the extent of existing ecosystems can be gained from measurements of land cover in 1988 made using 1×1 km resolution data from the AVHRR sensor on the NOAA satellite series (Stone *et al.*, 1994). These are tabulated in **Table 3**.

It should be emphasized that many ecosystems can be heavily disturbed by logging and other activities without the change being evident on satellite imagery. This is true for Landsat TM imagery (30×30 m resolution) used for deforestation estimates in Brazil, and the limitations are much greater for 1×1 km AVHRR data.

Brazil is the country with the most extensive satellite information on forest cover and its loss. Unfortunately, information on nonforest vegetation types such as cerrado is much less complete. Considerable confusion arises between the FAO classification and others such as the one adopted here because FAO classifies cerrado, caatinga, and chaco as "forests." FAO classifies areas with only 10% crown cover as "forests" if the trees are 5 m high or if it has "trees able to reach these thresholds *in situ*" (FAO, 2010a, p. 209).

Brazil's Legal Amazon region originally had 4 million km^2 of forests, the rest being cerrado and other types of savannas. Agricultural advance was slow until recent decades because of human diseases (especially yellow fever and malaria), infertile soil, and vast distances to markets. These barriers have progressively crumbled, although a range of limiting factors restricts the extent and the duration over which

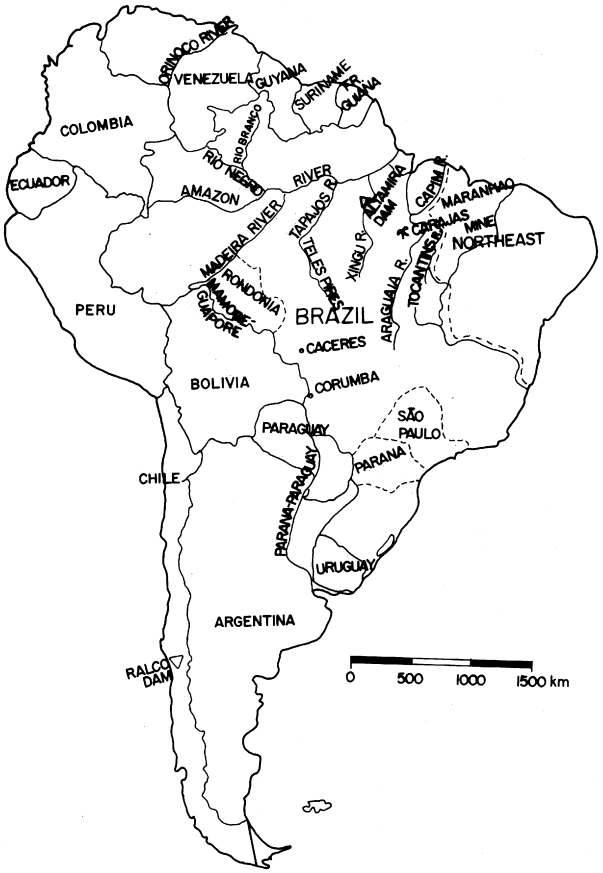


Figure 2 Locations mentioned in the text.

many uses of deforested areas can be maintained (Fearnside, 1997a). Deforestation in the region has been predominantly for cattle pasture, with critical contributions to the motivations for the transformation coming from the role of clearing as a means of establishing land tenure and in allowing land to be held and sold for speculative purposes (Fearnside, 2005, 2008).

The Atlantic forests of Brazil (ecoregions 54 and 55) have been almost completely (>95%) destroyed, mainly for

Table 3 Land cover in South America in 1988

Country	*Closed tropical moist forest*	*Recently degraded TMF*	*Closed forest*	*Degraded closed forest*	*Woodlands*	*Degraded woodlands*	*Savanna, grasslands*	*Degraded savanna, grasslands*	*Scrublands, Shrublands*	*Desert, bare soil*	*Water*	*Snow, rock, ice*	*Other*	*Total*
Argentina	1.2	0.0	96.8	0.6	645.4	15.7	755.4	232.8	894.8	37.9	34.0	31.4	35.7	2779.8
Bolivia	323.5	12.7	409.2	24.6	345.1	102.2	87.7	86.2	4.8	16.5	11.9	0.1	1.1	1089.4
Brazil	3522.3	519.7	3686.0	1692.2	1555.9	330.0	740.0	179.4	0.0	0.0	80.9	0.0	124.0	8388.5
Chile	0.0	0.0	134.1	29.1	75.2	29.8	101.1	14.0	86.9	186.8	7.0	16.6	3.8	684.5
Colombia	581.6	5.4	622.5	11.4	116.3	14.5	255.5	64.0	0.0	0.0	3.1	0.0	22.8	1110.1
Ecuador	115.5	1.7	121.0	1.7	33.7	4.3	41.9	13.3	3.2	2.5	0.6	0.0	0.8	223.1
French Guiana	78.8	0.0	79.8	2.4	0.6	0.0	0.2	0.0	0.0	0.0	0.1	0.0	1.0	84.1
Guyana	159.4	2.0	171.6	2.4	5.4	0.3	18.4	1.5	0.0	0.0	1.2	0.0	3.7	204.3
Paraguay	0.3	0.0	8.9	0.2	209.1	50.7	104.0	26.5	0.0	0.0	0.6	0.0	1.1	401.1
Peru	620.8	19.1	654.7	19.1	88.0	78.8	139.0	97.4	64.3	88.0	8.3	0.7	5.6	1244.1
Suriname	126.0	2.5	128.5	10.0	0.5	0.3	1.2	0.4	0.0	0.0	1.1	0.0	3.3	145.2
Uruguay	1.4	0.0	2.1	0.0	0.9	0.0	154.1	11.0	0.0	0.0	3.0	0.0	5.9	177.0
Venezuela	379.1	0.2	415.5	9.9	33.9	40.2	243.3	82.0	27.2	0.0	11.4	0.0	8.4	871.8
Unclassified														313.0
Total	5909.9	563.4	6530.7	1803.7	3109.8	666.9	2642.0	808.5	1080.6	331.7	163.2	48.9	217.2	17,716.1
Continent (%)	33.4	3.2	36.9	10.2	17.6	3.8	14.9	4.6	6.1	1.9	0.9	0.3	1.2	
Category (%)		8.7		21.6		17.7		23.4						100.0

Note: All values in thousands of km^2 or percent. "TMF" includes tropical moist, semideciduous, and gallery forests, "Grasslands" includes those seasonally flooded, "Closed Forest" includes TMF, montane forests, cool and temperate deciduous forests, and tropical seasonal forests, "Degraded Grasslands" includes agriculture, "Desert, Bare Soil" includes inland salt marsh communities, and "Other" includes wet vegetation and mangroves. More recent land-cover maps for South America are available, but without tabulation by country (Eva *et al.*, 2004; Arino *et al.*, 2008; Tateishia *et al.*, 2008).

agriculture, silviculture, and real estate development. Most of what remains of this extraordinarily rich ecosystem is in protected areas, but unprotected areas continue in rapid retreat. These forests are recognized as major "hot spots" of biodiversity (Heywood and Watson, 1995; Myers *et al.*, 2000; Stotz *et al.*, 1996).

In Andean countries, clearing by small farmers has predominated in driving deforestation, in contrast to the predominant role of medium and large cattle ranchers in Brazil. Migration from densely populated areas in the Andean highlands (*altiplano*) has led to settlement in lowland forests areas, with consequent upsurges in clearing (e.g., Rudel and Horowitz, 1993).

Savanna ecosystems have suffered heavy human pressure. The pampas of Argentina and the Uruguayan savannas of Uruguay and Southern Brazil (ecoregions 120 and 121) have largely been converted to agriculture. The Brazilian cerrado, originally covering 2 million km^2, is the largest ecoregion in South America, as well as holding the largest number of species of any of the world's savannas. The cerrado was largely intact until the mid-1970s. Clearing, especially for soybeans and planted pasture, reduced the cerrado to 55% of its original area by 2002 according to MODIS imagery (Klink and Machado, 2005). The advance of clearing has proceeded at an accelerating pace, speeded by infrastructure projects and an array of government subsidies.

The temperate and coniferous forests of the Southern Cone have been under severe pressure from logging. These forests are usually logged by clear-cutting in a manner similar to their counterparts in the North American temperate zone. This contrasts with the "selective" logging (highgrading for a few species) that characterizes timber extraction from the diverse forests of the tropical region.

Human Use of Converted Areas

Conversion of natural ecosystems to agroecosystems and secondary forests creates landscapes that maintain biodiversity to varying degrees. "Shifting cultivation" as practiced by indigenous peoples and by traditional nonindigenous residents (*caboclos*) in Amazonian forests maintains a substantial part of the original biodiversity. This contrasts with the effect of the vast expanses of cattle pasture that have replaced this, either directly or following a phase of use in pioneer agriculture by small farmers who have recently arrived from other places.

In densely settled areas along the coast of Brazil and in the southern portions of the country, agricultural use has gone through a series of "cycles," such as sugarcane and coffee. The productivity of many areas has been damaged by soil erosion and other forms of degradation. Cattle pasture is often the land use replacing these crops. Plantation silviculture has grown steadily since the 1970s and covered more than 74,000 km^2 by 2010. Soybeans (207,000 km^2 in 2008) have also made large advances.

In Argentina and Uruguay, cattle ranching and wheat and rice farming are major land uses. Natural vegetation is better represented in areas with little agricultural potential, such as mountain and polar areas and arid and semiarid zones.

Human Use of Remaining Natural Habitats

Areas that remain under natural vegetation cover, rather than being converted to other land uses through clearing, are also subject to human use and alteration. Selective logging in tropical forests, for example, leaves much of the basic structure of the ecosystem intact, but also can lead to significant changes that can set in motion a sequence of events leading to complete destruction of the ecosystem. Logging leaves a substantial amount of dead biomass in the forest, including the crowns and stumps of harvested trees and all of the biomass of the many additional trees that are killed by damage sustained during the logging process. Openings created in the canopy allow sunlight and heat to penetrate to the forest floor, drying out the fuel bed more quickly than in unlogged forests. Climatic variations such as those provoked by the El Niño phenomenon make logged forests especially susceptible to entry of fires. Ample opportunities for fires are provided as fields are burned to prepare land for planting and as cattle pastures are burned to control invading weeds. The fires burn slowly through the understory, charring the bases of trees as they go. Many of these trees then die, leading to a positive-feedback process whereby more dead biomass and canopy openings are provided and subsequent fires begin with greater ease, killing still more trees. This can degrade the entire forest within a few years (Nepstad *et al.*, 2001).

Tropical forests are also used for "extractivism," or the collection of nontimber forest products (NTFPs) such as rubber and Brazil nuts. This does relatively little damage to the forest, although extractivists do have an impact through hunting and through clearing for subsistence crops. The extractivist population can also play a protective role in defending the forest against encroachment by more aggressive actors such as ranchers and loggers. This is the basis of the extractive reserve system in Brazil (see Anderson, 1990).

Savannas are often grazed by cattle without cutting trees. Cerrado (ecoregion 114), "lavrado," or Guianan savannas (ecoregion 111), the Pantanal wetlands (ecoregion 133), and the llanos of Venezuela (ecoregion 110) are among the savannas often used in this way. Increasing fire frequency, virtually all a result of human-initiated burning, can lead to shifts in species composition and to a drain of nutrients.

Aquatic ecosystems are traditionally exploited by fisheries. This alters the relative abundance of the species present. Use of watercourses as recipients for sewage and other pollutants also affects aquatic life in many ways.

Threats to Remaining Natural Habitats

Terrestrial Ecosystems

Deforestation

Deforestation is the dominant transformation of forested ecosystems that threatens biodiversity. In Brazil, which holds most of the continent's remaining forests, ranching is the dominant use for land once deforested. In the 1990s, soybeans began to enter forested regions, representing a new force in this process (Fearnside, 2001). Soybeans had already been a

major factor in transformation of the cerrado since the 1970s. The most important effect of soybeans is not loss of forest directly planted to the crop, but the extensive infrastructure of waterways, railways, and highways that are built to transport soybeans and the inputs needed to grow them. The cycle of deforestation that has repeatedly occurred along Amazonian highways can be expected to accompany these new access routes (Fearnside, 2007).

Population growth is a fundamental contributor to deforestation and other forms of natural habitat loss. In recent years, however, the redistribution of population through migration has overshadowed the impact of absolute growth in population size. These include migrations from the semiarid Northeast of Brazil to Amazonia, from Paraná to Rondônia, from the highlands of Bolivia, Peru, and Ecuador to the Amazonian lowlands and, in the case of Ecuador, to the Pacific lowlands as well.

Logging and Charcoal Manufacture

Logging is an increasingly important factor in Amazonia, and the catalytic role of this activity in increasing the flammability of the logged forest gives it potential impact far beyond its direct damage. So far, logging in Brazil has been dominated by domestic demand for sawn wood, plywood, and particleboard, which is almost entirely supplied from tropical forests rather than from silvicultural plantations (which produce wood for pulp and, to a lesser extent, charcoal). However, global markets for tropical timber are presently dependent on supplies from Asian forests that will soon come to an end if current rates of exploitation continue. In the 1990s, Asian logging companies began buying land and/or obtaining concessions in such countries as Brazil, Guyana, and Suriname, and pressure from global timber markets can be expected to increase in the future. Asian loggers are also the principal forces in clear-cutting the Valdivian and *Nothofagus* forests of Chile (ecoregions 88 and 89).

In eastern Amazonia, demand for charcoal for pigiron smelting in the Carajás area is a potential threat to forests. Carajás, with the world's largest deposit of high-grade iron ore, is expected to be mined for 400 years at the present rate of exploitation. Wood from native forests is inherently cheaper as a source of biomass for charcoal production as compared to plantation-grown sources. Charcoal manufacture has an impact on the forest both through direct removal (including officially sanctioned forestry management systems) and by increasing the profitability of logging and deforestation (see Anderson, 1990).

Deforestation impacts are magnified by fragmentation and edge effects (Laurance and Bierregaard, 1997). This division of the remaining natural habitat into many small islands surrounded by cattle pastures or other highly modified land uses, together with forming edges with increased entry of light, wind, and foreign organisms, results in many changes in the remaining natural ecosystems. Most of these changes are forms of degradation, such as greatly increased mortality in the trees that provide the dominant component of forest structure. Vine loads on trees near edges also increase, leading to further increase in mortality and susceptibility to windthrow.

Other Threats

Climate change represents a major long-term threat to many South American ecosystems (Fearnside, 2010). In addition to higher temperatures, continued global warming would cause dramatic increases in the frequency and severity of droughts in Amazonia both due to the El Niño phenomenon that is triggered by warming of surface water in the Pacific Ocean (Cox *et al.*, 2004) and due to even faster increases in the frequency of sea-surface temperature anomalies in the Atlantic Ocean such as the one that caused a disastrous drought in 2005 (Cox *et al.*, 2008).

Removal of fauna through hunting is a virtually universal consequence of proximity of human settlements to natural habitats. The removal of fauna can affect seed dispersal, pollination, and other processes needed for maintaining plant and animal communities. Introduction of exotic species also represents a threat to natural ecosystems. Exotic species are a particularly severe problem in the Valdivian and *Nothofagus* forests of Chile (ecoregions 88 and 89).

Mangrove ecosystems are subject to some unique threats. Shrimp culture in mangrove areas has had severe impacts on the coast of Ecuador. Mangroves in Maranhão have been subject to pressure for charcoal manufacture. In São Paulo state mangroves have often suffered from oil spills and are also losing ground to real estate development. This has also affected restingas (ecoregions 176–178).

Aquatic Ecosystems

Dams

Hydroelectric dams have major impacts on river ecosystems by blocking fish migration, by eliminating rapids and replacing well-oxygenated running water with reservoirs that usually have anoxic water in their lower layers. The composition of fish present changes radically and undergoes a succession of changes as reservoirs age. Anoxic water released through the turbines severely reduces fish and freshwater shrimp productivity in the rivers downstream of the dams.

In Brazil, the 2010 Plan, released in 1987, listed more than 300 dams for eventual construction in Brazil, independent of the expected date of completion. Of these, 65 dams were in the Amazon region. Economic difficulties have caused projected construction dates to be successively postponed, but the ultimate number of dams has not changed. Most contentious is the Babaquara Dam (renamed the "Altamira Dam") on the Xingu River, which would flood greater than 6000 km^2 of forest, much of it in indigenous areas (Fearnside, 2006).

In Chile, the dams planned in Patagonia, together with their transmission line to Santiago, are expected to have major environmental impacts. In Uruguay, at least five major dams are planned for construction in the next few years. Brazilian-financed dams are moving forward in Peru, Bolivia, and Guyana.

Waterways

Industrial waterways, known as *hidrovias* in Brazil, greatly alter aquatic habitats. No less than seven waterways are under construction or planned for soybean transport on barges: the Paraguay–Paraná (*Hidrovia do Pantanal*), the Madeira River

waterway, the Tocantins-Araguaia waterway, the Teles Pires–Tapajós waterway, the Capim River waterway, the Mamoré–Guaporé waterway, and the Rio Branco and Rio Negro–Orinoco waterways. Waterway construction involves blasting rock obstructions, cutting sharp curves, and dredging sediment from the river beds. The Corumbá–Cáceres stretch of the *Hidrovia do Pantanal*, if built, would lower the water level in the Pantanal wetlands (ecoregion 133), threatening one of the world's most renowned concentrations of wildlife.

Other Threats

Other threats to aquatic habitats include sedimentation from soil erosion and landslides. This is severe, for example, in rivers draining steep areas of former Atlantic forest in the coastal mountains of Brazil. Mining for gold, tin, and diamonds in Amazonia can also inject large amounts of sediment into streams and rivers.

Destruction of varzea forest (ecoregion 33) in Amazonia can affect aquatic life through loss of important fish breeding areas and food sources for fruit- and seed-eating fish. Destruction of varzea lakes and overfishing represent additional threats.

Status of Protected Areas

The choice and design of reserves depend on the financial costs and biodiversity benefits of different strategies. In Brazil, rapid creation of lightly protected "paper parks" has been a means of keeping ahead of the advance of barriers to establishment of new conservation units, but emphasis must eventually shift to better protection of existing reserves (Fearnside, 1999).

Creating reserves that include human occupants has a variety of pros and cons (Kramer *et al.*, 1997). Although the effect of humans is not always benign, much larger areas can be brought under protection regimes if human occupants are included (Fearnside, 2003). Additional considerations apply to buffer zones around protected areas. A "fortress approach," whereby uninhabited reserves are guarded against encroachment by a hostile population in the surrounding area, is believed to be unworkable as a means of protecting biodiversity, in addition to causing injustices for many of the human populations involved.

Priorities for Conservation

Indigenous peoples have the best record of maintaining forest, but negotiation with these peoples is essential in order to ensure maintenance of the large areas of forest they inhabit (Fearnside and Ferraz, 1995). The benefits of environmental services provided by the forest must accrue to those who maintain these forests. Development of mechanisms to capture the value of these services will be a key factor affecting the long-term prospects of natural ecosystems.

In the case of deforestation in Amazonia, a variety of measures could be taken immediately through government action, including changing land tenure establishment procedures so as not to reward deforestation, revoking remaining incentives, restricting road building and improvement, strengthening requirements for environmental impact statements for proposed development projects, creating employment alternatives, and, in the case of Brazil, levying and collecting taxes that discourage land speculation. A key need is for a better informed process of making decisions on building roads and other infrastructure such that the full array of impacts is taken into account.

Environmental services represent a major value of natural ecosystems, and mechanisms that convert the value of these services into monetary flows that benefit the people who maintain natural habitats could significantly influence future events in the region (Fearnside, 1997b). Environmental services of tropical forests include maintenance of biodiversity, carbon stocks, and water cycling. The water cycling function, although very important for countries in the region, does not affect other continents as the first two services do. At present, avoiding global warming by keeping carbon out of the atmosphere represents a service for which monetary flows are much more likely to result from international negotiations. Activities under the United Nations Framework Convention on Climate Change (UN-FCCC) are at a much more advanced stage of negotiation than is the case either for the Biodiversity Convention or for the "Non-Binding Statement of Principles" and possible future convention on forests.

In the case of carbon, major decisions regarding credits for tropical forest maintenance are pending in ongoing negotiations. Regardless of what is decided, global warming is a permanent consideration that can be expected to receive increasing weight in decision making. The threats to natural ecosystems in South America are many, and recognition of the multiple environmental services provided by them is a key factor in ensuring that substantial areas of each of these ecosystems continue to exist, thereby maintaining their biodiversity.

Acknowledgments

I thank Eric Dinerstein and the World Bank for permission to publish **Figure 1** and **Table 1**, and Tom Stone and the American Society for Photogrammetry and Remote Sensing for permission to publish **Table 3**. Brazil's National Council of Scientific and Technological Development (CNPq305880/2007-1, 573810/2008-7, 575853/2008-5) and National Institute for Research in the Amazon (INPA) provided financial support. SV Wilson and two anonymous reviewers made helpful comments on the manuscript.

See also: Deforestation and Land Clearing. Fires, Ecological Effects of. Grazing, Effects of. Indigenous Peoples and Biodiversity. Rainforest Loss and Change

References

Anderson AB (ed.) (1990) *Alternatives to Deforestation: Towards Sustainable Use of the Amazon Rain Forest*. New York: Columbia Univ. Press.

Arino O, Bicheron P, Achard F, *et al.* (2008) GLOBCOVER: The most detailed portrait of Earth. *European Space Agency Bulletin* 136: 24–31.

Cox PM, Betts RA, Collins M, *et al.* (2004) Amazonian forest dieback under climate–carbon cycle projections for the 21st century. *Theoretical and Applied Climatology* 78: 137–156. http://dx.doi.org/10.1007/s00704-004-0049-4.

Cox PM, Harris PP, Huntingford C, *et al.* (2008) Increasing risk of Amazonian drought due to decreasing aerosol pollution. *Nature* 453: 212–215.

Dinerstein E, Olson DM, Graham DJ, *et al.* (1995) *A Conservation Assessment of the Terrestrial Ecoregions of Latin America and the Caribbean*. Washington, DC: The World Bank.

Eva. HD, Belward AS, De Miranda EE, *et al.* (2004) Land cover map of South America. *Global Change Biology* 10(5): 731–744. http://dx.doi.org/10.1111/j.1529-8817.2003.00774.x.

FAO (Food and Agriculture Organization of the United Nations) (1993) *Forest Resources Assessment 1990: Tropical Countries (FAO, Forestry Paper 112)*. Rome, Italy: FAO.

FAO (Food and Agriculture Organization of the United Nations) (2010a) *Global Forest Resources Assessment 2010*. Rome, Italy: FAO, 340 pp. http://www.fao.org/forestry/fra/fra2010/en/

FAO (Food and Agriculture Organization of the United Nations) (2010b) *Global Forest Resources Assessment 2010-Brazil Country Report*. Rome, Italy: FAO, 111 pp. http://www.fao.org/forestry/fra/fra2010/en/

Fearnside PM (1997a) Limiting factors for development of agriculture and ranching in Brazilian Amazonia. *Revista Brasileira de Biologia* 57(4): 531–549.

Fearnside PM (1997b) Environmental services as a strategy for sustainable development in rural Amazonia. *Ecological Economics* 20(1): 53–70.

Fearnside PM (1999) Biodiversity as an environmental service in Brazil's Amazonian forests: Risks, value and conservation. *Environmental Conservation* 26(4): 305–321.

Fearnside PM (2001) Soybean cultivation as a threat to the environment in Brazil. *Environmental Conservation* 28(1): 23–38. http://dx.doi.org/10.1017/S0376892901000030.

Fearnside PM (2003) Conservation policy in Brazilian Amazonia: Understanding the dilemmas. *World Development* 31(5): 757–779. http://dx.doi.org/10.1016/S0305-750X(03)00011-1.

Fearnside PM (2005) Deforestation in Brazilian Amazonia: History, rates and consequences. *Conservation Biology* 19(3): 680–688. http://dx.doi.org/10.1111/j.1523-1739.2005.00697.x.

Fearnside PM (2006) Dams in the Amazon: Belo Monte and Brazil's hydroelectric development of the Xingu River basin. *Environmental Management* 38(1): 16–27. http://dx.doi.org/10.1007/s00267-005-0113-6.

Fearnside PM (2007) Brazil's Cuiabá-Santarém (BR-163) Highway: The environmental cost of paving a soybean corridor through the Amazon. *Environmental Management* 39(5): 601–614. http://dx.doi.org/10.1007/s00267-006-0149-2.

Fearnside PM (2008) The roles and movements of actors in the deforestation of Brazilian Amazonia. *Ecology and Society* 13(1): 23. [online] URL http://www.ecologyandsociety.org/vol13/iss1/art23/

Fearnside PM (2010) Tropical forests of Amazonia. In: Schneider SH, Rosencranz A, Mastrandrea MD, and Kuntz-Duriseti K (eds.) *Climate Change Science and Policy*, pp. 104–112. Washington, DC: Island Press.

Fearnside PM and Ferraz J (1995) A conservation gap analysis of Brazil's Amazonian vegetation. *Conservation Biology* 9(5): 1134–1147.

Harcourt CS and Sayer JA (eds.) (1996) *The Conservation Atlas of Tropical Forests: The Americas*. New York: Simon & Schuster.

Heywood VH and Watson RT (eds.) (1995) *Global Biodiversity Assessment*. Cambridge, UK: Cambridge Univ. Press.

Klink CA and Machado RB (2005) Conservation of the Brazilian Cerrado. *Conservation Biology* 19(3): 707–713. http://dx.doi.org/10.1111/j.1523-1739.2005.00702.x.

Kramer R, van Schaik C, and Johnson J (1997) *Last Stand: Protected Areas and the Defense of Tropical Biodiversity*. Oxford, UK: Oxford Univ. Press.

Laurance WF and Bierregaard RO (eds.) (1997) *Tropical Forest Remnants: Ecology, Management, and Conservation of Fragmented Communities*. Chicago, IL: Univ. of Chicago Press.

Myers N, Mittermeier CG, Mittermeier RA, da Fonseca GAB, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

Nepstad DC, Carvalho G, Barros AC, *et al.* (2001) Road paving, fire regime feedbacks, and the future of Amazon forests. *Forest Ecology and Management* 154: 395–407.

Rudel TK and Horowitz B (1993) *Tropical Deforestation: Small Farmers and Land Clearing in the Ecuadorian Amazon*. New York: Columbia Univ. Press.

Stone TA, Schlesinger P, Houghton RA, and Woodwell GM (1994) A map of the vegetation of South America based on satellite imagery. *Photogrammetric Engineering and Remote Sensing* 60(5): 541–551.

Stotz DF, Fitzpatrick JW, Parker TA III, and Moskovitz DK (1996) *Neotropical Birds: Ecology and Conservation*. Chicago, IL: Univ. of Chicago Press.

Tateishia R, Bayaera MA, Ghara H, *et al.* (2008) A new global land cover map, GLCNMO. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37(Part B7): 1369–1372.

Southern (Austral) Ecosystems

Peter H Weston, University of Adelaide and National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, NSW, Australia
Robert S Hill, University of Adelaide, SA, Australia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Continental drift The movement of continents as a result of plate tectonic processes. This was especially important for Southern Hemisphere continents.
Disturbance A set of processes that lead to an ecological point crisis, that is, an event or events that are not part of the normal environment for the organisms that are influenced by it and that can be catastrophic in its impact.
Eucalyptus The dominant tree genus in Australia today.
Gondwana The southern supercontinent, named after an ancient kingdom in India. Gondwana rifted apart over many millions of years, giving rise to the current Southern Hemisphere landmasses.
Nothofagus A genus of Southern Hemisphere trees with a classic "Gondwanic" distribution in southern South America, southeastern Australia, New Zealand, New Caledonia, and New Guinea.

Introduction

A large number of plant taxa are shared by New Zealand, southern South America, and southeastern Australia (including Tasmania) but are found nowhere else. Joseph Hooker was so impressed by this pattern that in 1853 he wrote in his Introductory Essay to the Flora Novae-Zelandiae:

> …many of the peculiarities of each of the three great areas of land in the southern latitudes are representative ones, effecting a botanical relationship as strong as that which prevails throughout the lands within the Arctic and Northern hemisphere zones, and which is not to be accounted for by any theory of transport or variation, but which is agreeable to the hypothesis of all being members of a once more extensive flora, which has been broken up by geological and climatic causes.

Thus was born the concept of "Southern (austral) ecosystems." Hooker knew nothing of the geological basis for the patterns that he observed but he thought that they demanded explanation. We now know that the "once more extensive flora" about which he speculated did exist and that the "geological and climatic causes" that he hypothesized were the consequences of continental drift. "Southern (austral) ecosystems" thus comprise an historical entity rather than an ecological class as one finds in a volume of Ecosystems of the World.

The Southern Hemisphere is dominated by oceans, and the development of these oceans is critical in understanding the evolution of terrestrial biodiversity of this part of the world. All major southern landmasses, along with some that are now mostly in the Northern Hemisphere, were once part of the supercontinent Gondwana. The breakup of Gondwana, and the consequent movement of the continents that were once part of it, produced vicariant distribution patterns, as observed by Hooker, but it also had other profound effects on the Gondwanic biotas.

History

In 1859, Joseph Hooker further emphasized the close relationship between the floras of the southern landmasses in his Introductory Essay to the Flora of Tasmania and persisted in maintaining that this must be the relic of a continuous, ancestral flora. His friends and colleagues soon disagreed. In 1859, Charles Darwin in his Origin of Species cited Hooker's austral floristic pattern but not his geological speculations, explaining the pattern primarily as the result of repeated, independent episodes of long-distance dispersal across the Southern Ocean. In 1880, Alfred Russell Wallace in his Island Life, while noting that a number of animal taxa, such as the ratite birds, also showed an austral distributional pattern, chose to explain this as the result of waves of southward dispersal from the Northern Hemisphere, followed by extinction in the north. He agreed with Darwin that the austral floristic relationship must be the result of long-distance dispersal. Neither had any time for geological instability on the scale that Hooker had proposed.

These pioneering biogeographers spawned different, competing research traditions that remain alive and distinguishable today. Hooker inspired an intellectual lineage of phytogeographers to search for repeated distributional patterns and to postulate general explanations for them, invoking dispersal and vicariance of entire floras. This approach was applied in extraordinary breadth by Croizat in his "panbiogeographic" analyses of numerous plant and animal distributions. Croizat's work produced not only a general distributional pattern ("generalized track") crossing the Southern Ocean, consistent with Hooker's analysis, but plenty of other general patterns too, including tracks spanning all ocean basins. After 100 years, the Hookerian tradition was still looking for an explanatory geological theory.

A competing (and for long, dominant) research tradition, inspired by Darwin and Wallace, flourished among zoogeographers, paleontologists, and many phytogeographers. The primary aim of this approach was to construct plausible scenarios describing the centers of origin and dispersal histories for particular taxa, based chiefly on their patterns of occurrence in the fossil record, knowledge of their phylogeny, the ecology of extant species, and knowledge of geological history. An example of this tradition is the work of Darlington (1965).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00271-9

Corroboration of the theory of continental drift through the discovery of plate tectonics in the 1960s vindicated Hooker's approach by providing the former land connections that he had predicted, in the form of the ancient supercontinent Gondwana. The transformation from "austral flora" to "Gondwanan biota" did, however, involve a substantial expansion in scope. Its potential boundaries shifted northward to include the whole of South America and Africa, Madagascar, India, all of Australia plus southern New Guinea, and New Caledonia. Refinements in geological reconstruction progressively added more fragments to Gondwana, such as eastern Sulawesi, the "Outer Melanesian Arc" (Solomons to Vanuatu and Fiji), and, eventually, large parts of eastern Asia.

The immediate effect of this geological revolution on biology was to flip the conventional null hypothesis in historical biogeography from long-distance dispersal to overland dispersal. "Darwinian" biogeographers changed the paleogeographic maps on which they plotted dispersal routes but continued to use the same methods as before. A new question was added to their task list: "To what extent can extant distributions be explained by continental drift?"

Different problems exercised the minds of "Hookerian" biogeographers. While a number of Croizat's tracks were corroborated in surprising detail as vicariant patterns, others, such as his trans-Pacific tracks, remained geologically anomalous. It could be argued that all approaches to historical biogeography had been shown by continental drift to be inadequate in some way or other.

The apparent heuristic success of Croizat's track method and the emergence of new cladistic techniques for reconstructing phylogeny inspired the development of a new approach, cladistic (or vicariance) biogeography, in the 1970s and 1980s (see, Humphries, this volume). This is based on the idea that historical biogeographic relationships can be summarized in the same way as phylogenetic relationships: in the form of tree diagrams. However, cladistic biogeographic trees connect areas of endemism occupied by taxa, rather than the taxa themselves. Methods of analysis have been developed that reconcile "area cladograms" for different taxa to produce "general area cladograms," which summarize episodes of biotic dispersal and fragmentation. The optimistic expectation of this approach is that general patterns will dominate such analyses, producing a small number of general area cladograms. Several cladistic biogeographic analyses of austral biotas have been published, most of which reproduce expanded, cladistically resolved versions of Hooker's austral biota. However, the most ambitious of these (San Martin and Ronquist, 2004) produced a general area cladogram for plants that was more congruent with the predictions of long-distance dispersal than the geological area cladogram. In contrast, the general area cladograms for animal groups, and especially for noninsect animals, were more congruent with geology. Ironically, phytogeographers, who had been the main supporters of the Hookerian tradition over the previous 150 years, had provided San Martin and Ronquist with predominantly Darwinian patterns, while zoogeographers, who had been the most enthusiastic advocates of the Darwinian approach, had provided support for Hooker's hypothesis.

The latest methodological advance that has contributed new evidence on the biogeographic history of southern biotas is relaxed clock molecular dating. This approach is a development of more simplistic methods based on the assumption of a "molecular clock." If a set of homologous DNA sequences, sampled from a clade of biogeographic interest, has accumulated site mutations at a constant rate, then it can be used to calculate the relative age of speciation events. Relaxed clock molecular dating is more sophisticated in that it does not assume a constant evolutionary rate but instead attempts to model the rate of change of the rate of molecular evolution in different lineages and incorporate these estimates into the process of producing a molecular chronogram – a tree diagram in which branch lengths are proportional to estimated time since divergence. Such a chronogram can be "calibrated" by reference to the minimum age of one or more clades, as indicated by the fossil record, in combination with a maximum age constraint. A number of studies based on molecular dating analyses have tested whether a range of plant and animal taxa are old enough to have acquired their disjunct, austral distributions as a result of continental drift. These have suggested that overland dispersal, vicariance, and long-distance dispersal have all contributed to Gondwanic patterns and consequently, that the explanations of both Hooker and Darwin for these patterns were partly correct.

Patterns

In 1853, Hooker counted 77 plant species and "upwards of 100 genera, subgenera, or other well-marked groups of plants" entirely or nearly confined to New Zealand, Australia, and extratropical South America. His analysis has not been replicated but subsequent authors have added additional taxa without attempting to be comprehensive. For example, Crisci *et al.* (1991) added 10 vascular plants as well as 41 invertebrate taxa, one vertebrate family, and one genus of fungi. Some of these additions were taxa found in New Caledonia, New Guinea, and southern Africa, and hence ignored by Hooker. Some include numerous subclades showing overlapping Gondwanic distributions and therefore are more informative than other taxa. The Diamesine and Podonomine Chironomid midges are classic examples of such groups (Brundin, 1966), as is the plant family Proteaceae, for which 15 higher taxa show separate, overlapping Gondwanic distributions (Sauquet *et al.*, 2009).

Clearly, the pattern of disjunct distributions on which Hooker originally based his concept of an austral flora has not been weakened by further research, but has been substantially strengthened by the addition of more botanical examples as well as a large number of zoological ones. What other patterns are relevant to the reconstruction of Gondwanic biogeographic history? The relationships between distributional patterns and other variables such as means of dispersal, area, and climatic factors have been discussed in great detail by Darwinian biogeographers. General area cladograms that specify the putative splitting sequence of Gondwanic biotas are potentially useful, as are estimates of the age of particular taxa derived from molecular dating. Finally, biogeographers of all persuasions have sought to explain patterns of occurrence of austral taxa in the fossil record.

Distributional Patterns in Relation to Means of Dispersal, Climate, and Area

Darwinian biogeographers emphasized a number of patterns that they thought outweighed Hooker's austral pattern and these were reviewed by Darlington (1965). He pointed to the low number of vertebrate taxa that belong to the austral biota: marsupials, chelyid turtles, leptodactylid and hylid frogs, and galaxiid fishes, and of these, only the galaxiids occur in New Zealand. If these dispersed over land between Australia and South America, then why did no other vertebrate groups follow them? Why did hardly any mammals, reptiles, or amphibians walk to New Zealand? Darlington considered plants and insects, which comprise the great majority of austral taxa, to have more effective means of dispersal than most terrestrial vertebrates.

Darlington also argued that climatic patterns were highly significant in explaining austral distributions. The southern parts of South America, Tasmania, and New Zealand are cold and their western sides very wet because of the strong prevailing westerly winds. The austral biota (as circumscribed by Hooker) is largely restricted to these areas. Moreover, strong westerly winds blow directly over parts of Tasmania, New Zealand, and southern South America. The area of land in the Southern Hemisphere tapers off markedly to the south, greatly restricting the amount of land available to the austral biota. This suggested to Darlington that the "southern end of the world" would be a poor evolutionary center. Biological diversity also tapers off to the south. The extreme southern parts of Australia, New Zealand, and South America all have lower biological diversity per unit area than the less climatically extreme areas further north.

General Area Cladograms

Do Gondwanic landmasses share a common ancestral biota to the exclusion of Laurasian landmasses? What are the cladistic interrelationships between different Gondwanic biotas? These are the kinds of questions that cladistic biogeographers are inclined to ask about the Southern Hemisphere. Among the first studies to focus primarily on these questions was that of Humphries (1981), which was based on cladograms for *Nothofagus* and 24 other taxa. Most of these cladograms were derived from precladistic taxonomic treatments and all relied entirely on morphological evidence. Nevertheless, he discerned two general cladistic patterns and concluded that "the two positions for South America … are probably due to the fact that it is a huge composite area and shouldn't be treated as a single area of endemism" (Humphries, 1981, p. 205).

More recent attempts to resolve a general area cladogram for Gondwanic landmasses have tended to subdivide South America, and sometimes other continents and islands such as Australia, New Zealand, and New Caledonia, into two or more areas for purposes of analysis. They have also been able to use more rigorously produced cladograms, some of which have incorporated molecular data.

The most comprehensive cladistic biogeographic analysis of Gondwanic landmasses yet published, that of Sanmartín and Ronquist (2004), incorporated 72 area cladograms (53 for animal clades, 18 plant clades and one fungal group). The most striking feature of this analysis was the different cladistic resolutions found for plant and animal groups. The general area-cladogram for the 18 plant clades can be summarized thus:

((New Guinea, New Caledonia), (SW Pacific, SE Asia)), ((N South America, (Holarctic, (Africa, Madagascar))), (S South America, (New Zealand, Australia)))

This pattern is incongruent with what we know of the geological relationships of these continental fragments: New Caledonia clusters with New Guinea (not New Zealand as predicted by geology), Madagascar clusters with Africa (not India), and Australia clusters with New Zealand (not S South America). In all of these cases, continental landmasses cluster with present day near neighbors rather than landmasses to which they were most recently connected by dry land, suggesting that long-distance dispersal has been the dominant process in the creation of these patterns. Detailed reconstructions of biogeographic events associated with nodes in this tree also found significantly higher proportions of dispersals across the Tasman Sea, the Mozambique Channel and the South Atlantic Ocean than would be expected due to chance.

In contrast, the general area cladogram for the 53 animal clades matches geological predictions in clustering Australia with geographically distant southern South America, to which it was connected by continuous continental crust through Antarctica as recently as 35 million years ago (ma), not to its present day nearest neighbor, New Zealand. The noninsect animal area cladogram is even more congruent with geology and can be summarized thus:

(New Caledonia, ((Madagascar, (SE Asia, India)), (N South America, (Holarctic, Africa))), (New Zealand, (S South America, (Australia, New Guinea))))

This not only resolves relationships between Australia, New Guinea, South America, New Zealand, India and Madagascar as predicted by geology, suggesting that information from the oldest Gondwanan vicariance events has been preserved, but it also fails to cluster New Caledonia with any other landmass, perhaps reflecting the suggestion (e.g., Grandcolas *et al.*, 2008) that New Caledonia was completely inundated from the early Paleocene to the Middle Eocene, and was recolonized entirely by long-distance dispersers from diverse sources. Interestingly, the congruent placement of New Zealand in geological and animal area cladograms is inconsistent with recent suggestions (e.g., Waters and Craw, 2006) that New Zealand was completely inundated during the Oligocene.

Relaxed Clock Molecular Dating

Molecular dating analyses have been used to test hypotheses of Gondwanan vicariance by estimating the ages of disjunct clades: if a clade is estimated to be older than the known geological age of separation of the continental fragments on which it occurs, then vicariance cannot be rejected. A number of such studies have recently been published, the results of which have been mixed. Plant and animal taxa that are "classic members" of the austral biota mostly show patterns resulting from biotic dispersal and vicariance, often combined with

some long-distance dispersal. Examples include the plant families Podocarpaceae (Wagstaff, 2004), Atherospermataceae and Monimiaceae (Renner, 2005), Nothofagaceae (Knapp *et al.*, 2005) Araucariaceae, Proteaceae (Barker *et al.*, 2007) and Winteraceae (Marquinez *et al.*, 2009), and animal clades including marsupials chironomid midges, Onychophora (Allwood *et al.*, 2010), and ratite birds (Phillips *et al.*, 2010). In these examples, disjunctions between South America and Australia are almost all older than 35 ma, consistent with the geological age of severance of the land connection between these continents through Antarctica. No plant clades but some animal taxa (e.g., Onychophora, some chironomids) that show disjunctions involving Africa, have been estimated to be older than the age of separation of these continents from the rest of Gondwana (105 ma). However, the great majority of clades showing Australia–NewZealand and Australia–New Caledonia disjunctions are younger than the oldest (84 ma) oceanic crust separating Australia from the largely submerged continent Zealandia, consistent with long-distance dispersal across the Tasman and Coral Seas. Although some of these clades are also estimated to be younger than the postulated Cenozoic marine inundations of New Caledonia and New Zealand (Goldberg *et al.*, 2007; Grandcolas *et al.*, 2008) a sizeable minority are enigmatic in being old enough to predate the supposed Oligocene drowning of New Zealand (Renner, 2005; Knapp *et al.*, 2005; Barker *et al.*, 2007; Allwood *et al.*, 2010).

Many other taxa that show austral distribution patterns, especially subclades in the plant families Asteraceae, Aizoaceae, Apiaceae, Boraginaceae, Caryophyllaceae, Fabaceae, Gnetaceae, Orchidaceae, Poaceae, and Plantaginaceae, have been estimated to be too young to have been involved in any Gondwanan vicariance events and their trans-oceanic disjunctions should be attributed to long-distance dispersal, not vicariance (Crisp *et al.*, 2009, and references therein).

Fossil Record

The patterns of occurrence of taxa in the fossil record provide evidence of the minimum ages of taxa and often indicate that they previously had more extensive distributional ranges than today. In most cases, absence of fossils of a taxon from deposits of a particular age or geographic location cannot reasonably be construed as evidence that the taxon really did not exist at that time or place. Absence of fossil evidence can also be due to incompleteness of the fossil record. However, a few taxa fossilize so readily and abundantly and are so reliably identified as fossils that their absence from the record has to be taken seriously as an indication of either true absence or extreme rarity. The angiosperm genus *Nothofagus* falls into this class.

The fossil record has other limitations, the most restrictive of which is the relatively low proportion of taxonomic characters that are preserved. Vertebrates, for example, are usually only represented by skeletal remains. Their "soft parts" and DNA are rarely preserved. Similarly, many plant taxa are represented only by their fossilized pollen grains. While these sometimes display several distinctive synapomorphies diagnostic of particular clades (see, e.g., Sauquet *et al.*, 2009), most often they do not. Moreover, pollen grains are such simple structures that the probability of convergent evolution of a distinctive pollen type in the ancestor of an extant clade and in a different, extinct taxon, while not being high, is not negligible. Despite these limitations, the fossil record is a rich source of historical information. Rather than inadequately reviewing the paleontological literature pertaining to all Gondwanic taxa, we concentrate on examples of a couple of key angiosperms.

Probably no taxon has been as strongly associated with the concept of the austral biota as *Nothofagus*. In 1853, Hooker included two subgroups (as *Fagus*) in his list of Antarctic plant groups, Darlington (1965) came close to defining his "southern end of the world" on the basis of its distribution, and van Steenis considered it a key genus for plant geography. *Nothofagus* has an exceptionally detailed fossil record, with both pollen and macrofossils relatively common. At least five distinct types of fossil *Nothofagus* pollen are recorded and they include all the extant subgenera, *Lophozonia*, *Fuscospora*, *Nothofagus*, and *Brassospora*, respectively.

Nothofagus pollen is produced in abundance and fossil pollen is extremely common. If *Nothofagus* pollen is not present in sediments then the genus probably was not present in the immediate vicinity. Therefore, the only limitation to determining the early biogeographic history of the genus is the lack of sediments at critical times and places. To a certain extent, dating the earliest appearances of *Nothofagus* in early Campanian sediments of southern Gondwana (Dettmann *et al.*, 1990) by pollen assemblages alone could involve some circularity, since they are sometimes indicator species for the pollen stratigraphy. However, in both southern Australia and the Antarctic Peninsula, where the earliest *Nothofagus* pollen has been found, the grains are preserved in marine sediments that are independently dated using dinoflagellate and invertebrate fossil assemblages. Earliest records of *Nothofagus* pollen date from the early Campanian, making *Nothofagus* one of the oldest extant angiosperm genera in the fossil record. Records of this pollen are from (in the order they appear) Australia, West Antarctica, New Zealand, and South America (Dettmann *et al.*, 1990). By the early Maastrichtian, the *Nothofagus* clade had diversified into all four modern subgenera, first appearing in West Antarctica, shortly thereafter in South America, and subsequently in Australia some 12–15 million years later (Dettmann *et al.*, 1990). According to the pollen records, the last of the modern subgenera to appear in Australia was either the oldest subgenus, *Lophozonia*, or one of the two youngest, *Brassospora*. New Caledonia and New Guinea are devoid of fossils other than subgenus *Brassospora*, the current earliest record being from the Miocene in New Guinea (Dettmann *et al.*, 1990).

In contrast to *Nothofagus*, the eucalypts (Myrtaceae: *Eucalyptus* alliance) have only been proposed as a Gondwanan clade relatively recently. Like many other plant taxa, the pollen grains of most eucalypts are not particularly distinctive and may be confused with those of other myrtaceous groups such as *Syncarpia* and *Metrosideros* (Hill, 1994). They are pollinated by insects, birds, and mammals and produce fewer, less widely dispersed pollen grains than wind-pollinated *Nothofagus*.

Phylogenetic studies (e.g., Para-O *et al.*, 2006) have resulted in the recognition of seven genera of eucalypts:

Eucalyptus, with more than 500 species, nearly all of which are endemic to Australia (several species occur in New Guinea, one of which extends west and northwest to Sulawesi and Mindanao, several occur in the lesser Sunda Islands as far west as Flores); *Corymbia*, with 113 species, restricted to Australia and New Guinea; *Angophora*, with about 20 species restricted to eastern Australia; *Arillastrum* with one species endemic in New Caledonia; *Eucalyptopsis* with one species endemic in New Guinea; *Allosyncarpia*, with one species restricted to Arnhem Land in northern Australia; *Stockwellia*, with one species surviving as a single population in the Queensland wet tropics. Unfortunately, fossil names for eucalypts have not yet been changed to match this new classification, and so all fossils mentioned here are referred to *Eucalyptus*.

In form the species range from low shrubs to the tallest flowering plants in the world. Eucalypts have been estimated to contribute 75% of Australian vegetation, and at the wetter margins of the continent, they dominate nearly all vegetation except rain forest and allied mesic types. Only in the arid Australian interior are eucalypts generally lacking in dominance.

Given their diversity and current dominance, the eucalypts would be expected to yield a complex and challenging fossil record. That this is not the case has been well documented (Hill, 1994). The oldest reliably dated and described Australian macrofossils associated with *Eucalyptus* include tree stumps, leaves and umbels about 20–22 million years old from several sites in the southeast of Australia (Hill, 1994).

The early pollen record of *Eucalyptus* is enigmatic mostly because of the difficulty in identifying eucalypt pollen. However, probable eucalypt pollen is now recorded from the late Paleocene of the Lake Eyre Basin in inland Australia (Hill, 1994).

An intriguing aspect of the fossil record of *Eucalyptus* is its apparent presence in both South America and New Zealand. A group of three fruits from probable Eocene sediments in Patagonia have been described as *Eucalyptus patagonica*, although there is some doubt as to whether they belong to the *Eucalyptus* alliance (Hill, 1994). However, Gandolfo *et al.* (2011) recently described exquisitely preserved leaves, flower buds, and fruits of *Eucalyptus* sensu stricto, with the flower buds assigned to *Eucalyptus* subgenus *Symphyomyrtus*. The full ramifications of these fossils will take some time to appear. Leaves from early Miocene sediments in the South Island of New Zealand are very similar to many *Eucalyptus* species and occur with eucalypt-like fructifications. This record appears to be very good, especially when it is considered in conjunction with the pollen record, since *Eucalyptus* pollen in New Zealand occurs from the Miocene–early Pleistocene, suggesting a relatively recent extinction.

Processes

The complex patterns described here for the Southern Hemisphere biota suggest complex processes, and we still have much to learn about the way in which the Southern Hemisphere biota was molded into its living form. In the following section we briefly mention some of the more likely processes that have been inferred to have had a significant impact.

Continental Drift

Although the precise detail of the breakup of Gondwana is still being refined, our broad understanding is robust and very important for the history of the biota. When continent-sized landmasses rift apart there must be profound direct and/or indirect consequences for the biota. These include changes in climate as the various landmasses move into and out of different climatic zones and the altered relative position of the continents can also influence climate, for example, by altering ocean currents on a massive scale. Plate convergence resulting in subduction, volcanism and continental collision has also profoundly influenced biotas by bringing once separate communities into secondary contact, creating new environments such as mountain ranges and driving extinction.

Vicariance

The most obvious direct consequence of continental rifting is vicariance, the physical separation of once continuous populations of species, which must have often led to allopatric speciation as conditions on the resultant landmasses changed relative to one another.

Biotic Dispersal

Dispersal and intermixing of whole communities and biotas over land has clearly been an important process in the assembly of terrestrial ecosystems throughout evolutionary history. Such processes would have been responsible for the establishment of numerous sympatric distributions of widespread ancestral taxa on Gondwana, as well as the progressively smaller supercontinents that formed and were sundered during the process of Gondwanan fragmentation. This process has also been facilitated by continental collisions and tectonic accretion, most notably the collisions between the northward-drifting Indian, African, South American and Australian continents and Laurasia, or parts of it, during the Cenozoic Era. The resulting north–south biotic dispersal seems to have been predominantly southward in direction. As the Gondwanan landmasses moved close to their present positions, the opportunity for relatively simple, predominantly land-based dispersal opened in three major regions: from Southeast Asia into Australia, from North America to South America, and from Europe to Africa. There are many compelling examples of such southward dispersals and a smaller number of northward ones.

Long-Distance Dispersal

The Southern Hemisphere is dominated by oceans, which provide formidable barriers to dispersal. However, it seems likely that many organisms have successfully crossed these ocean barriers and formed successful populations at the other end. This is likely to have been in the direction of prevailing winds and associated sea-surface currents – predominantly westerly at temperate southern latitudes and easterly in the tropics. While long-distance dispersal is a fact of life, it is very difficult to provide clear evidence for it and so it is largely recorded on a narrative basis. However, the methodological

inadequacy of long-distance dispersal as a general explanation should not be taken to imply its unreality as a process.

Climate Change

The climate history of the Southern Hemisphere is relatively well known and has clearly undergone major change during the rifting of Gondwana. This has been due to a combination of effects, including altered ocean currents, changing atmospheric CO_2 levels and shifting continental positions. The most obvious place where this has had an impact is Antarctica, and this makes a useful case study to demonstrate the potential impact of climate change.

Antarctica has a long history as a continent close to the South Pole. Throughout the rifting of Gondwana, Antarctica has been one of the more stable landmasses, and although its history is complex, it can be considered as a long-standing high-latitude region. The idea that Antarctica was a source of plants and animals for surrounding high-latitude landmasses such as southern South America, southern Australia, and New Zealand is central to Southern Hemisphere biogeography. However, Antarctica today not only is geographically isolated from the other southern landmasses, but is also far too climatically hostile to support the biota concerned.

Antarctica now has only a very sparse covering of plants near the coastline, only two species of which are vascular. Most of the continent is without plant cover, and where terrestrial plants occur, they are usually lichens and mosses. The Antarctic vertebrate fauna is sea-based, using the land only temporarily. In contrast, there is abundant evidence from Antarctica and adjoining landmasses to demonstrate the presence of complex biotas near the South Pole at times in the past. The almost total lack of vascular plants on Antarctica today testifies to major extinctions on this continent. This rates as one of the most complete regional extinction events in the Earth's history, and it is important for us to understand it now more than ever, since it was indisputably a climate-based extinction.

There is now abundant fossil evidence for complex Cretaceous and Cenozoic forests on Antarctica that contain taxa (or their descendants) that often still occur in high southern latitudes in South America, Australia, and New Zealand (Hill and Scriven, 1995). A combination of plate tectonics and other less well defined factors can be used to explain a very different climate at high latitudes in the past, and physiological research on the extant flora has demonstrated that diverse plant life was possible in Antarctica in the past without other physical changes.

During the Cretaceous at high southern latitudes massive speciation occurred, and there is no doubt that lineages were produced that today occupy niches from the equator down to the furthest southern latitudes at which the biota can survive.

The Cenozoic Antarctic vegetation is known from only limited data, but it appears that it continued to show basic similarities to Paleogene floras from Australia and South America. The separation of Australia from Antarctica began in the late Cretaceous, although opportunities for floristic interchange across restricted water gaps probably persisted well into the Cenozoic. Opening of the Drake Passage began in the mid Oligocene, but details are lacking concerning the extent of dry land, probably as islands, between the Antarctic Peninsula and South America prior to that date.

The question of when the Cenozoic vegetation of Antarctica was eliminated by increasing cold and ice cover has not yet been answered by the fossil record. This is largely because of the confusion introduced into the spore and pollen record by the recycling of plant microfossils once processes of glacial erosion and sedimentation had begun. However, there is reasonable evidence for a cover of temperate forest in the Eocene, perhaps over much of Antarctica. Truswell (1991) noted that sea-surface temperatures about Antarctica were close to 0 °C after the end of the Eocene, allowing the formation of sea ice; temperatures on land were probably lower, and it is unlikely that anything but a highly specialized, restricted flora would have persisted for long in such conditions. However, sparse Oligocene palynofloras containing relatively abundant *Nothofagus* (at least three subgenera) and a low diversity of other vascular plants have been described from Antarctica (Hill and Scriven, 1995). The data indicate a coastal *Nothofagus* forest that was diverse enough to suggest an extension of temperate forest in the region beyond the Eocene.

Milder sea-surface temperatures adjacent to the lower latitude regions of the northern Antarctic Peninsula probably allowed the persistence there of some form of woody vegetation for even longer, and there are good fossil data to suggest vegetation cover into the late Oligocene at least.

The probable middle to late Pliocene Sirius Formation sediments contain leaves, wood, and pollen in an excellent state of preservation, and there can be no doubt that the plants which produced these fossils were growing in situ on the site at the time the fossils were deposited. A tundra vegetation consisting of at least 18 plant species, including seeds, fruits, flowers, leaves and wood, and some in situ plants has been reported, and this includes leaves, wood and pollen of *Nothofagus* (Ashworth and Cantrill, 2004). Therefore the evidence available suggests progressive extinction from an earlier vegetation in which *Nothofagus* was very diverse.

Disturbance

The angiosperms (flowering plants) play a critical role in the current Southern Hemisphere biota, providing a substantial component of the biodiversity and also acting as an important part of the food chain and niche requirements for many other organisms. Angiosperms probably evolved in the early Cretaceous and radiated worldwide. Early angiosperms probably had generalized pollination and dispersal strategies which made them ideal for long-range dispersal by pioneering the fresh sedimentary surfaces of coastal deltas, lagoons, and tidal flats. During the early Cretaceous, these pioneering early angiosperms or angiosperm ancestors dispersed along coastlines out of the rift valley of West Gondwana (between Africa and South America).

Angiosperms were established generally in Gondwana by Barremian–Aptian times (Dettmann, 1989), and those lineages that were widespread in southern Gondwana by the close of the Cretaceous formed the foundation for the living austral floras. Dettmann (1989) has suggested that rift valley

systems at high southern latitudes acted as migrational pathways for early angiosperms, and this may be because of the high disturbance level in these environments and the early successional nature of these angiosperms.

Some prominent angiosperm taxa from the early high southern latitude flora still exist today, and their ecology is well understood. A good example is *Nothofagus*, which has been studied extensively. South American *Nothofagus* does not usually regenerate continuously at low altitudes, especially at lower latitudes, where climatic conditions favor more complex forests and *Nothofagus* seedlings cannot establish. Instead, *Nothofagus* relies on catastrophic disturbance, a relatively common phenomenon in the highly unstable Andes mountain chain and the Southern Alps of New Zealand, to provide a fresh, cleared substrate which is readily colonized by seedlings. This regeneration behavior is probably very similar to that which occurred early in *Nothofagus* history, when it either occupied dense forests on unstable sites and was unable to regenerate in the absence of disturbance, or was in more sparse forest on less equable sites where its seedlings could survive in the understory. Hill (1994) hypothesized that the large reduction in *Nothofagus* diversity at high latitudes today, especially away from the Andes, may be due to the relative modern stability of the landscape, which, coupled with the density of the forest cover, does not provide the right environment for the continued presence of many *Nothofagus* species.

Thus, at high southern latitudes during the Cretaceous–Paleogene, unstable habitats provided both a pathway for the migration of early angiosperms and also the potential for peripheral isolation of populations of widespread species, which may have been a critical factor for evolution in the region (Hill and Scriven, 1995).

A more recent form of disturbance has been the widespread impact of fire. In Australia, fire history as a major impact on the biota can be traced back to the early Cenozoic at least, but its influence has been particularly profound only in the past 100,000 years or so. Fire is now very significant in areas of the Southern Hemisphere with hot, dry summers.

Plant Growth at High Southern Latitudes

Woody plants no longer grow at very high southern latitudes (e.g., Antarctica), and yet they were a key component of Cretaceous biotas very close to the South Pole. In the past, debate has centered on whether this was possible, since the photoperiods at such high latitudes require long, dark winter periods. The issue is more one of suitable temperatures than suitable photoperiods, but nevertheless, the question remains: How did plants grow and what form did the communities take in such an unusual photoperiod? It is a real limitation of the fossil record in general that behavioral traits such as dormancy rarely leave a clear imprint (Truswell, 1991). Nevertheless, information on physiological adaptations to life at high latitudes can be obtained, especially from fossil wood. Cretaceous gymnospermous wood from paleolatitudes probably higher than 70° S shows a consistent growth pattern, with increments, probably annual in origin, clearly delineated, and indicating a pronounced seasonal influence (Hill and Scriven, 1995). The amount of wood added annually was usually large, even when compared with modern low-latitude trees, but there is considerable variation in ring widths from year to year, suggesting that the trees were highly sensitive to environmental fluctuations (Truswell, 1991). A fossil forest at one early Cretaceous site had trees spaced 3–5 m apart and the tallest preserved trunk height is 7 m. Hill (1994) used these data to explain the reduction in wind pollination in understory plants during the very early Paleogene in southern Australia. With the sun at a low angle in the sky during summer and tracking an almost circular path around the horizon during the day, he concluded that it was likely that the forest structure was quite different to that observed in heavily forested regions at lower latitudes today. Relatively widely spaced, conical trees probably dominated and this may have continued until at least the late Cretaceous, and thus wind pollination was a viable strategy for understory plants below the open canopy. As today's forest-bearing landmasses moved to lower latitudes and the sun angle increased, a closed forest structure became possible, due to the increase in incoming solar radiation. This may have been critical for the loss of wind pollination as a viable strategy in the associated understory plants.

The presence of a relatively dense plant cover, including large and rapidly growing trees, in regions south of the Antarctic circle, was for a long time viewed as difficult to explain, given that today growth at high latitudes is considered to be constrained by characteristic day-length patterns and by the temperatures associated with a regime of long rigorous winters and short cool summers (Truswell, 1991). However, it is now generally accepted that the light energy at high latitudes was sufficient, at present values of axial obliquity, to maintain forest growth in the past provided the temperatures, particularly in winter, were higher.

Enhanced atmospheric CO_2 levels may have been influential in producing the high annual increments of growth observed in Cretaceous trees. The overall effect of the high predicted levels of CO_2 in the Cretaceous on plant growth are difficult to quantify, since experiments carried out on living plants examine the instantaneous effects of large changes in CO_2 concentration on plants which have evolved to a particular ambient level. When plants have evolved to accommodate higher levels (as in the Cretaceous) the result may have been even more spectacular (Hill and Scriven, 1995).

Conclusion

The biota of the Southern Hemisphere is extremely diverse and is the result of a complex history. However, clear biotic linkages occur among the major landmasses. These linkages have historical significance, with vicariance and long-distance dispersal contributing to the geographic spread of organisms and climate change and disturbance being among the major processes that have led to the living species that occur there. Our knowledge of the history of the Southern Hemisphere is incomplete, but we have enough information from interrelationships among living organisms and the fossil record to be able to reconstruct at least part of this history with reasonable accuracy. The most important event in the history of

the Southern Hemisphere as it influences the living biota was the rifting of Gondwana. Not only did this separate the biota on to smaller landmasses that subsequently had very different histories, but it also influenced other processes such as climate change and disturbance regime. This distinguishes the Southern Hemisphere biota very clearly from that of the Northern Hemisphere, and it is important not to presume that the processes in the two hemispheres were at all similar.

See also: Antarctic Ecosystems. Australia, Biodiversity of Ecosystems. Ecosystems of South America

References

Allwood J, Gleeson D, Mayer G, Daniels S, Beggs JR, and Buckley TR (2010) Support for vicariant origins of the New Zealand Onychophora. *Journal of Biogeography* 37: 669–681.

Ashworth AC and Cantrill DJ (2004) Neogene vegetation of the Meyer Desert Formation (Sirius Group) Transantarctic Mountains, Antarctica. *Palaeogeography Palaeoclimatology Palaeoecology* 213: 65–82.

Barker NP, Weston PH, Rutschmann F, and Sauquet H (2007) Molecular dating of the "Gondwanan" plant family Proteaceae is only partially congruent with the timing of Gondwanan break-up. *Journal of Biogeography* 34: 2012–2027.

Brundin LZ (1966) Transantarctic relationships and their significance as evidenced by chironomid midges. *Kungliga Svenska Vetenskapsakademiens Handlinger Series* 4(11): 1–472.

Crisci JV, Cigliano MM, Morrone JJ, and Roig-Junent S (1991) Historical biogeography of southern South America. *Systematic Zoology* 40: 152–171.

Crisp MD, Arroyo MTK, Cook LG, *et al.* (2009) Phylogenetic habitat conservatism on a global scale. *Nature* 458: 754–758.

Darlington PJ (1965) *Biogeography of the Southern End of the World*. Cambridge, MA: Harvard University Press.

Dettmann ME (1989) Antarctica: Cretaceous cradle of austral temperate rainforests? In: crame JA (ed.) *Origins and Evolution of the Antarctic Biota*. vol. 47, pp. 89–105. Geolical Society Special.

Dettmann ME, Pocknall DT, Romero EJ, and Zamaloa M de C (1990) *Nothofagidites* Erdtman ex Potonié 1960 – A catalogue of species with notes on the palaeogeographic distribution of *Nothofagus* Bl (Southern Beech). *New Zealand Geological Survey Paleontological Bulletin* 60: 1–79.

Gandolfo M, Gonzalez C, Zamaloa M, Cuneo N, and Wilf P (2006) *Eucalyptus* (Myrtaceae) macrofossils from the early Eocene of Patagonia, Antarctica. *In Botany Proceedings of the Botanical Society of America Annual Meeting*, p. 162. Chicago, CA: California State University, Abstract 473.

Gandolfo MA, Hermsen EJ, Zamaloa MC, *et al.* (2011) Oldest known Eucalyptus macrofossils are from South America. PLoS ONE 6(6): e21084. doi:10.1371/journal.pone.0021084.

Goldberg J, Trewick SA, and Paterson AM (2008) Evolution of New Zealand's terrestrial fauna: A review of molecular evidence. *Philosophical Transactions of the Royal Society B. Biological Sciences* 363: 3319–3334.

Grandcolas P, Murienne J, Robillard T, *et al.* (2008) New Caledonia: A very old Darwinian island? *Philosophical Transactions of the Royal Society B. Biological Sciences* 363: 3309–3317.

Hill RS (1994) The history of selected Australian taxa. In: Hill RS (ed.) *History of the Australian Vegetation: Cretaceous to Recent*. Cambridge, MA: Cambridge University Press.

Hill RS and Scriven LJ (1995) The angiosperm-dominated woody vegetation of Antarctica: A review. *Review of Palaeobotany & Palynology* 86: 175–198.

Humphries CJ (1981) Biogeographical methods and the Southern Beeches (*Fagaceae: Nothofagus*). In: Funk VA and Brooks DR (eds.) *Advances in Cladistics*, pp. 177–207. New York: The New York Botanical Garden.

Knapp M, Stöckler K, Havell D, Delsuc F, Sebastiani F, and Lockhart PJ (2005) Relaxed Molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). *PLOS Biology* 3: 38–43.

Marquinez X, Lohmann LG, Salatino ML, Salatino A, and Gonzalez F (2009) Generic relationships and dating of lineages in Winteraceae based on nuclear (ITS) and plastid (rpS16 and psbA-trnH) sequence data. *Molecular Phylogenetics and Evolution* 53: 435–449.

Parra-O C, Bayly M, Udovicic F, and Ladiges P (2006) ETS sequences support the monophyly of the eucalypt genus *Corymbia* (Myrtaceae). *Taxon* 55: 653–663.

Phillips MJ, Gibb GC, Crimp EA, and Penny D (2010) Tinamous and Moa flock together: Mitochondrial genome sequence analysis reveals independent losses of flight among Ratites. *Systematic Biology* 59: 90–107.

Renner S (2005) Variation in diversity among Laurales, early Cretaceous to present. In: Friis I and Balslev H (eds.) *Plant Diversity and Complexity Patterns: Local, Regional and Global Dimensions*, pp. 441–458. Copenhagen: Kgl. Danske Videnskabernes Selskab.

Sanmartín I and Ronquist F (2004) Southern hemisphere biogeography inferred by event-based models: Plant versus animal patterns. *Systematic Biology* 53: 216–243.

Sauquet H, Weston PH, Anderson CL, *et al.* (2009) Contrasted patterns of hyperdiversification in Mediterranean hotspots. *Proceedings of the National Academy of Sciences of the USA* 106: 221–225.

Truswell EM (1991) Antarctica – A history of terrestrial vegetation. In: Tingey RJ (ed.) *The Geology of Antarctica*, pp. 499–537. Oxford: Clarendon.

Wagstaff SJ (2004) Evolution and biogeography of the austral genus Phyllocladus (Podocarpaceae). *Journal of Biogeography* 31: 1569–1577.

Waters JM and Craw D (2006) Goodbye Gondwana: New Zealand biogeography, geology and the problem of circularity. *Systematic Biology* 55: 351–356.

Speciation, Process of

Guy L Bush, Michigan State University, MI, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 371–381, © 2001, Elsevier Inc.

Glossary

Allopatric Populations, species, or taxa occupying different and disjunct geographical areas.
Ecotone The boundary or transitional zone between adjacent ecological communities or biomes.
Epistasis The interaction of non-allelic genes in which one gene (epistatic gene) masks the expression of another gene at a different locus.
Genetic drift The occurrence of random changes in the gene frequencies of small isolated populations not due to selection, mutation, or immigration.
Isolating mechanisms Any intrinsic or extrinsic mechanism or barrier that inhibits the free exchange of genes between populations.
Pleiotropic gene A gene that has more than one independent phenotypic effect.
Polyploid Having more than one set of homologous chromosomes.
Sister species A pair of species that have arisen from a single speciation event; each is the other's closest relative.
Sympatric Populations, species, or taxa that occur together in the same geographical area within the dispersal range of one another.

Speciation and Species

Speciation

The ultimate source of the earth's biodiversity is speciation, a process that occurs when gene flow is reduced sufficiently between sister populations to allow each to become irrevocably committed to different evolutionary lineages. Speciation results in the splitting of a lineage into two or more species (cladogenesis). However, when a species is transformed over time (anagenesis) by the acquisition of phenotypic and genetic modification, there is no increase in the number of species and thus no speciation event (**Figure 1a**).

The Species Problem

Although the end product of speciation seems clear, defining when and under what conditions two populations become irrevocably committed to different evolutionary lineages has proved to be difficult and controversial. A frequently invoked species definition is the biological species concept (BSC) of Mayr (1942), who proposed that "species are groups of actually or potentially interbreeding natural populations that are reproductively isolated from other such groups" (p. 120).

However, the BSC and its later modifications have several limitations. It can be applied objectively only to sexually reproducing taxa and to natural populations that are sympatric (i.e., those that occur in the same geographic area) where they have the opportunity to interbreed. Asexually reproducing taxa and fossils must be treated subjectively. A more serious limitation of the BSC is that sympatric taxa that maintain distinct phenotypic and genotypic differences, but sustain some gene flow through interbreeding with close relatives, are generally not recognized as species. Recent molecular studies on closely related sympatric animal and plant taxa recognized as distinct, closely related sister species (the products of one speciation event) reveal that gene flow between them may still occur (Feder *et al.*, 1995; Taylor *et al.*, 1997). Such neospecies have acquired sufficient genetically based differences to maintain independent genotypes and distinct phenotypes but are still capable of sustaining some gene flow as they diverge. Several other species concepts have been proposed (Futuyma, 1998), but all suffer from problems of circularity, inconsistencies, and untestable qualities.

For this reason, the genetic-based species concept of Mallet (1995) provides a useful, less assumptionladen, operational species definition for identifying sister species. Mallet proposed that sister species be recognized only when they maintain distinctly different sympatric genotypic clusters. Defining a species based on genotypic clusters represents the level of differentiation most interesting to those studying speciation. Generally, sympatric sister taxa that maintain distinct genotypic clusters exploit different habitats or hosts. The wide use of modern DNA analysis in evolutionary studies makes this a useful approach particularly when coupled with sound biological studies. However, as with the BSC, using genotypic clusters to determine the biological status of geographically isolated sister populations without a "test of sympatry" is still a subjective although somewhat more quantitative call.

Reproductive Barriers to Gene Flow and the Evolution of Mate Recognition Systems

Reproductive Barriers

The free exchange of genes between populations adapting to different niches inhibits the evolution of independent genetic systems. As soon as adaptive gene combinations are formed, they are broken up by recombination through interbreeding. There are several biologically based reproductive barriers or

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00129-5

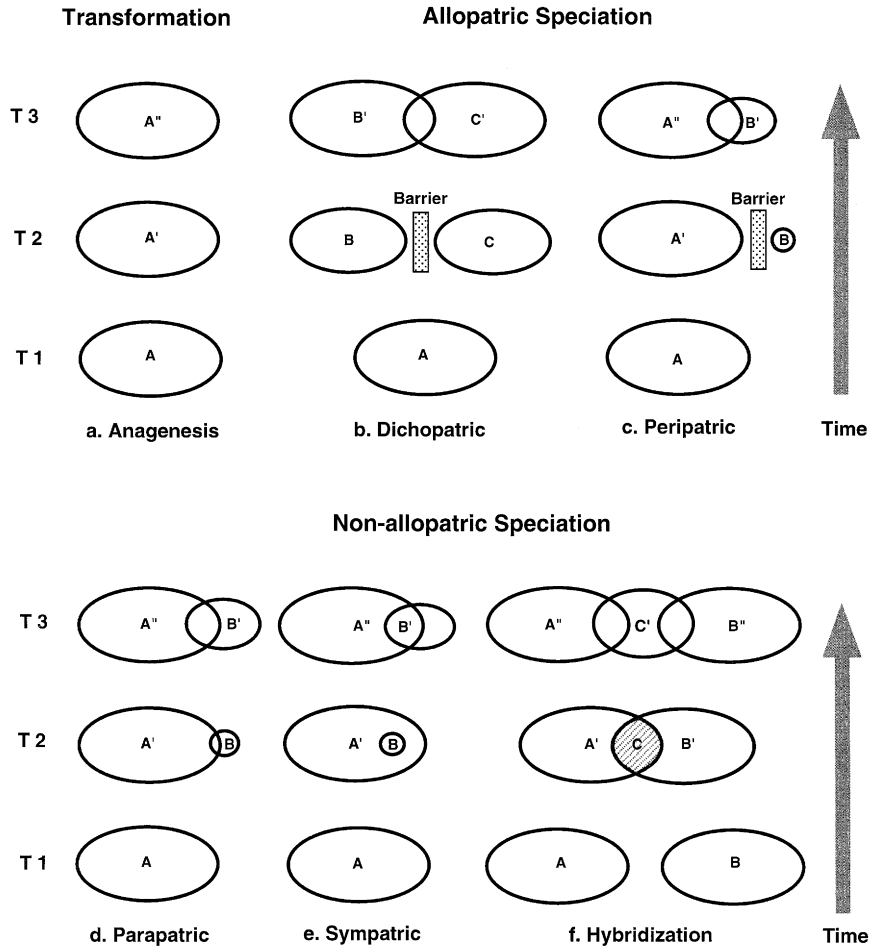


Figure 1 Major modes of speciation. See text for details on the processes involved in divergence and speciation.

"isolating mechanisms" that evolve during the course of speciation which limit gene flow between sister species (Table 1). These biological barriers constitute components of a mate recognition system that promotes assortative mating between individuals adapted or adapting to shared environmental and reproductive conditions.

Two Modes of Speciation

There are two primary modes of speciation recognized by most evolutionary biologists (Table 2). Allopatric (geographic) speciation occurs when sister populations are isolated for a period by a physical barrier such as a mountain range or an expanse of uninhabitable terrain or water. In the absence of gene flow, such isolated populations inevitably accumulate unique mutations and, over time, may diverge genetically by genetic drift and in response to divergent selection pressures. If sufficient genetic divergence occurs during isolation to establish and maintain distinct genotypic clusters, particularly when their ranges later overlap, then the taxa are regarded as species.

Non-allopatric (ecological) speciation occurs in the absence of physical barriers to gene flow when sister populations diverge genetically and become ecologically and reproductively isolated as they adapt to different habitats. New sister species thus evolve within the dispersal range of the offspring from a single deme. Several different patterns of geographic and ecological speciation are recognized based on the factors involved in promoting their divergence, such as population structure, habitat specialization, hybridization, and polyploidy.

Geography, Geographic Races, and Allopatric Speciation

The process of geographic speciation (often referred to as allopatric speciation) may happen in two ways. Dichopatric or vicariant speciation (Figure 1b) occurs when a widespread species becomes geographically divided into two or more large subpopulations by an insurmountable or impassable barrier. A second mode of geographic speciation, usually called peripatric or founder event speciation (Figure 1c), occurs when a geographically isolated population is established by a single fertilized female or a very small number of founding individuals.

However, determining the actual biological status of sister taxa that have arisen in isolation following either a vicariant or a founder event is difficult and controversial. Species can be

Table 1 Classification of reproductive isolating mechanisms

Prezygotic isolating mechanisms: mechanisms that prevent interspecific mating
Temporal isolation (potential mates have overlapping ranges but reproduction occurs at different times)
Habitat isolation (potential mates have overlapping ranges but reproduction occurs in different habitats)
Ethological isolation (potential mates meet but do not mate)
Mechanical isolation (potential mates attempt to copulate but no sperm is transferred)
Mechanism that prevents fertilization
Gametic mortality (sperm is transferred but egg is not fertilized)
Gametic incompatibility (gametes meet but fertilization is not completed)
Postzygotic isolating mechanisms: mechanisms that prevent the development of interspecific hybrids
Zygote mortality (eggs fertilized but zygote dies)
Hybrid inviability (zygote produces F_1 hybrid of reduced viability)
Hybrid sterility (F_1 hybrid zygote is fully viable but partially or completely sterile or produces deficient F_2 hybrid)
Coevolutionary or cytoplasmic interactions (individuals from a population infected by an endoparasite or with a particular cytoplasmic element are fertile with each other, but fertility and/or viability break down when matings occur between infected and uninfected individuals)

Table 2 Classification of modes of speciation

Allopatric speciation
Dichopatric speciation (vicariant division of wide-ranging species by an extrinsic barrier or extinction of intervening populations)
Peripatric speciation (by evolution in a very small isolated colony)
Nonallopatric speciation
Parapatric speciation
Sympatric speciation
Autopolyploid speciation
Spontaneous thelytokous parthenogenetic speciation
Hybrid speciation
Homoploid hybrid speciation
Introgressive speciation
Recombinational speciation
Polyploid speciation
Allopolyploid speciation
Direct and reticulate allogenous speciation
Symbiont-induced speciation

recognized when artificial hybridization of allopatric taxa in the laboratory reveals post-mating reproductive isolation. If no post-mating incompatibility is noted and mating occurs, then it is necessary to determine if there are factors important in pre-mating isolation not provided in the laboratory environment of the tests. Whether or not such populations are species or geographic races may require an unequivocal "test of sympatry" in which the two taxa coexist without fusing. Only when previously isolated taxa have the opportunity to interbreed under natural field conditions can the species status of closely related taxa lacking post-mating reproductive isolation be confirmed. If they maintain distinct sympatric genotypic clusters, they should be recognized as distinct species. Because such tests cannot often be performed, deciding on whether isolated taxa represent species or geographic races (i.e., subspecies) is a subjective decision. However, it should be stressed that when two taxa overlap naturally (i.e., are parapatric) and do not interbreed, this does not necessarily mean that they speciated sympatrically. This is because it is often impossible to establish if the overlap represents a case of secondary contact following allopatric speciation or if the taxa evolved as the result of non-allopatric divergence.

Dichopatric Speciation

When populations of geographically isolated sister taxa remain large during and following geographic subdivision, they will slowly diverge genetically by genetic drift and as they adapt to local conditions in isolation. Over time, they may accumulate sufficient genetic differences to cause negative pleiotropic gene interaction and hybrid incompatibility among genes responsible for proper mate recognition and genome integration. Reproductive isolation therefore occurs by chance as the by-product of genetic divergence in isolation rather than the outcome of natural selection acting directly to promote reproductive isolation.

Dichopatric race Formation and Speciation in Salamanders

The long-term study of evolution in the plethodontid salamander genus *Ensatina* by David Wake and colleagues (Wake and Schneider, 1998; Wake, 1997) provides an excellent example of dichopatric speciation and incipient species formation. The *Ensatina eschscholtzii* complex represents an ancient lineage that has undergone several instances of range contraction, isolation, and expansion and divergence of populations along broad ecological gradients. It is composed of seven contiguous subspecies wrapped in ring-like fashion around the Central San Joaquin Valley of California (**Figure 2**). Historically, populations of *Ensatina* slowly differentiated into highly distinct forms from the northern subspecies *E. e. oregonensis* in color and pattern as they expanded southward. Blotched forms occur in the Sierra Nevada and mountains of southern California, whereas unblotched forms are found in the coastal and northern region of their range. The animals, which are long-lived (10–15 years) and take 4 years to mature, are sedentary and disperse only short distances (home range 10–20 m).

Hybridization and intergradation occur between adjacent subspecies to varying degrees in all except one contact zone. The exception occurs in the Cuyamaca Mountains of San Diego County, in which *E. e. eschscholtzii* and *E. e. klauberi*

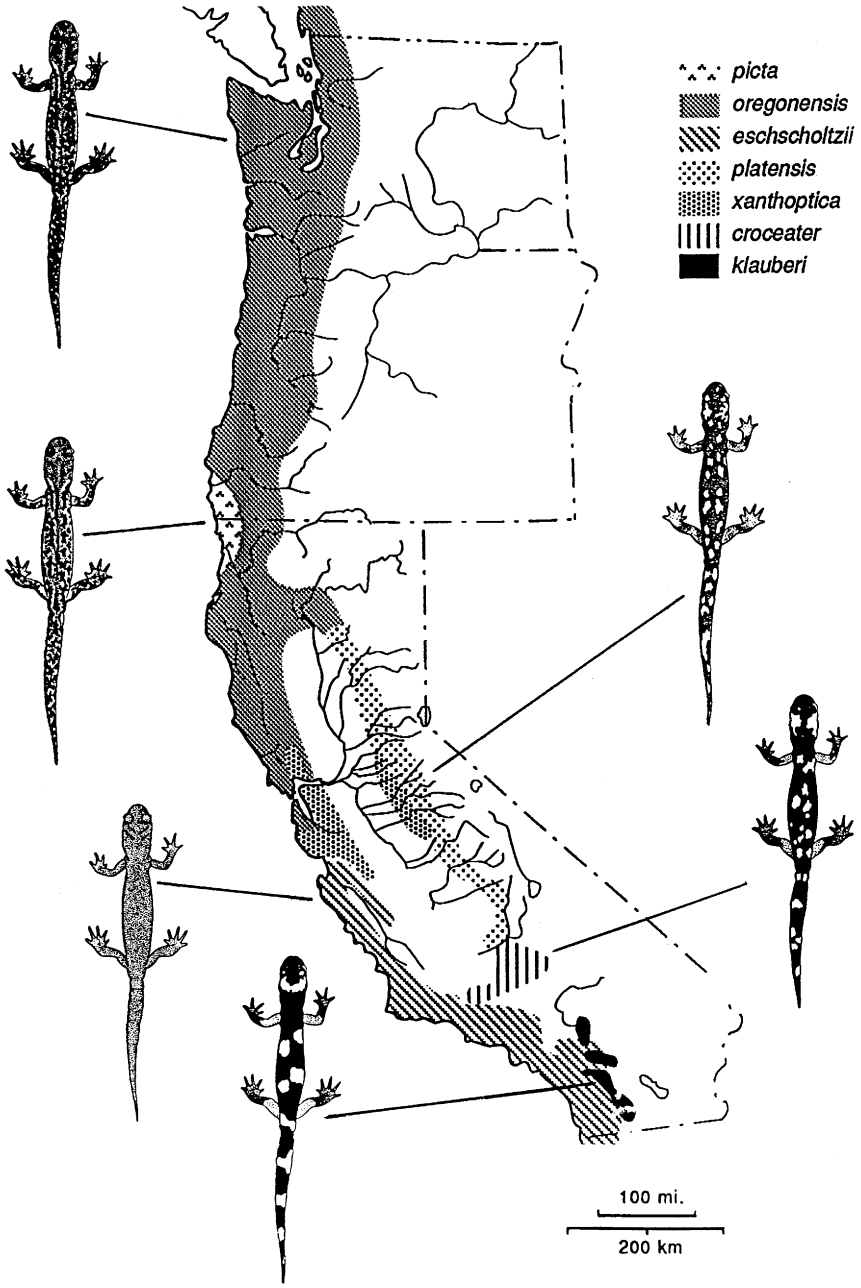


Figure 2 Geographic distribution of taxa in the salamander *Ensatina* complex based on molecular markers rather than morphological traits. Modified with permission from Wake DB (1997) Incipient species formation in salamanders of the *Ensatina* complex. *Proc. Natl. Acad. Sci. USA* 94: 7761–7767, and Futuyma DJ (1998) *Evolutionary Biology*. Sunderland, MA: Sinauer.

overlap but rarely interbreed in sympatry and are regarded by some biologists as distinct species. In other areas, the complex includes many geographically and genetically distinct entities, some of which are near the species level of differentiation.

A limited narrow secondary hybrid zone also occurs in Calaveras County between the upland *E. e. platensis* and a coastal, lowland subspecies, *E. e. xanthoptica* that established an outpost in the Sierran foothills sometime in the past. *Ensatina e. platensis* lives in the cool and moist, closed canopy forest at higher elevation than those of *E. e. xanthoptica*, which prefers warmer and drier lower elevations. Hybridization between the pure subspecies is rare, with most hybrids representing progeny from backcrosses or from crosses between hybrid individuals. Hybrid individuals appear to be at a selective disadvantage and thus subject to strong negative selection. They have reduced success in competing for preferred habitats and mates and seem to be more prone to predation. Because little or no gene flow occurs across the narrow hybrid zone, the coastal and Sierra taxons behave as incipient species or taxa in the final stages of speciation. Four other hybrid

zones in central California involving *E. e. xanthoptica* have been characterized in detail. The *Ensatina* complex presents a full array of differentiation from geographically variable populations or races to well-marked species. It provides an excellent example of how speciation has progressed slowly by adaptive divergence and stochastic processes during periods of geographic isolation.

Peripatric Speciation and the Founder Effect Principle

There are several examples in which species appear to have evolved in small, isolated populations at the periphery of the range of a sister species. Unlike the slow rate of evolution in dichopatric speciation, peripatric speciation is postulated to take a relatively short time. How and how often this divergence occurs is controversial (Barton, 1996; Hollocher, 1996).

Three models, each based on a founder event in which a population is established by only a few individuals, have been proposed to account for the rapid evolution of species observed on islands and elsewhere (**Figure 1c**). The founder effect principle was developed by Mayr (1954). It is based on the assumption that reproductive isolation from the parent species can evolve rapidly in a population established by a very small number of founding individuals (i.e., 2–10). He postulated that in such populations a genetic revolution could take place as a by-product of inbreeding, selection, drift, and genome reorganization. While the population is small, these genetic changes may promote substantial morphological and ecological shifts.

A modification of the founder effect principle was proposed by Carson (1975). In his founder-flush model of speciation, isolated populations undergo a series of population expansions and drastic contractions to a very small number of individuals. Carson believes that founder-flush speciation can occur only in certain cross fertilizing diploid organisms with "open" genetic systems. The genome in such organisms represents a clique of harmoniously collaborating or coadapted genes united by strong epistatic interactions. Their genomes also have abundant pleiotropic interacting genetic polymorphisms and share a high recombination index. Carson hypothesized that these attributes provide great genetic flexibility that predisposes such organisms to speciate by the founder-flush process.

In Carson's view, the drastic events required to reorganize the original genetic system and restore new balances that are incompatible with ancestors are accomplished during cycles of the founder-flush process. Selection, which is relaxed during the flush phase of population expansion, is greatly intensified as the population crashes. Repeated disorganization and reconstitution of the genome results in the rapid evolution of reproductive isolation.

Templeton (1980) proposed a third modification of founder-induced speciation that involves a genetic transilience (Carson and Templeton, 1984). It is similar to Carson's founder-flush speciation but requires changes in only one or a few segregating units, commonly with epistatic modifiers responsible for reproductive isolation that occurs when a population rapidly passes through an extremely unstable intermediate genetic state. As in the case of the founder-flush model, a genetic transilience involves strong inbreeding and large variance in population size. The critical trigger that initiates a transilience requires a reweighting of fitness components owing to drift-induced shifts in allele frequencies at one or more major loci that have pleiotropic effects on reproductive isolation.

It seems clear that speciation may occur rapidly following the geographic isolation of a small population. It is not clear, however, whether speciation results from a founder-flush or transilience process or solely from natural selection in response to factors such as runaway sexual selection or rapid adaptation to divergent ecological and reproductive conditions. Because conditions required for stochastic transitions are severe and well documented cases are lacking in which drastic genetic changes caused by founder effects result in speciation, it appears that the process of speciation in small, peripheral populations is the same as that which occurs in dichopatric speciation. Although certain kinds of epistasis can promote strong reproductive isolation, divergence occurs by selection and not random genetic drift (Barton, 1996).

Ecology, Ecological Races, and Nonallopatric (Ecological) Speciation

Reproductive isolation between sister populations in cases of allopatric speciation arises as a by-product of genetic changes that originate and accumulate independently in each population during periods of geographic isolation. Speciation may or may not be accompanied by an ecological transformation. In contrast, nonallopatric or ecological speciation inevitably involves a shift to a new niche at the time of speciation, and adaptation to different habitats is the driving force initiating and sustaining genetic divergence (**Figures 1d–1f**). Divergent selection on existing genetic variation results in the evolution of reproductive isolation as different suites of genetic variants are favored in each habitat. The emergence of reproductive isolation is also facilitated when mating occurs within a preferred habitat or if mate preference that is based on a phenotypic variant of a sexually selected trait becomes associated with a preferred habitat (Johnson *et al.*, 1996).

A misconception concerning nonallopatric speciation is that it requires a process called reinforcement, which involves the divergence of the mate recognition system between populations during the speciation process as an outcome of selection against hybrids (Butlin, 1987). Because theoretical models suggest that recombination will not allow reproductive reinforcement to develop, critics have argued that parapatric and sympatric speciation are unlikely to occur in nature. These models, however, fail to factor in the role of habitat preference and specialization in speciation and the effects of ecological divergence, which always accompanies sympatric speciation. When habitat choice is taken into consideration, conditions for sympatric speciation are greatly relaxed and its occurrence is probable (Johnson *et al.*, 1996).

Three modes of nonallopatric speciation are generally recognized in sexually reproducing organisms. Parapatric speciation (**Figure 1d**) takes place when sister species evolve while adapting to contiguous but spatially segregated habitats or ecotones across a narrow contact zone (Bush, 1994). Sympatric speciation (**Figure 1e**) occurs in the absence of

geographic segregation when sister species evolve within the dispersal range of the offspring of a single deme. During the course of sympatric speciation, the probability of mating between two individuals depends on their genotypes, and divergence occurs between populations adapting to alternate habitats within the "cruising range" of each other. Parapatric and sympatric speciation actually represent extremes of a continuum in the pattern and extent of habitat and geographic-imposed spatial segregation and gene flow reduction that occurs during nonallopatric divergence. Hybridization (**Figure 1f**) is a third mode of nonallopatric speciation. New species rapidly evolve from a mating between individuals of two closely related species by a variety of different processes. It is a frequent mode of speciation in plants but appears to be rare in animals.

Parapatric Speciation

This mode of speciation occurs when sister species evolve while adapting to contiguous habitats (or ecotones along a zone of contact. Examples from nature involve situations in which individuals from a species adapted to one habitat invade and colonize an adjacent habitat. Individuals bearing novel genetic recombinants capable of exploiting and reproducing in the new habitat are the first to invade and colonize the narrow zone where the two habitats meet. As adaptation proceeds, the new colony expands throughout the range of the new habitat. Only along the adjoining borders between the original and new habitat is gene exchange possible between the two populations. As the new population adapts to conditions imposed by the new habitat, divergent selection promotes the evolution of reproductive isolation and eventually parapatric species.

Parapatric Speciation in Mimulus

An example of rapid parapatric adaptation and speciation of a plant to a newly established vacant niche is that of the monkey flower *Mimulus cupriphilis* in California (Macnair and Gardner, 1998). This species grows only on relatively dry and toxic mine tailings of two small copper mines in California. It grows in close proximity and is recently derived from the widespread hydrophylic *M. guttatus*.

The two species differ in many ways. *Mimulus cupriphilis*, an obligate annual, flowers earlier and produces many small flowers that differ in shape and color from those of *M. guttatus*. *Mimulus cupriphilis* also has higher fitness on the dry tailings of the copper mines and flowers early when pollinators are rare. Because it is self-fertilizing, it is reproductively isolated from the outcrosser *M. guttatus*, which blooms later when its larger flowers are fertilized by bumblebees. Genetic studies have revealed that the species differ in a few major genes controlling flowering time, flower size, corolla spot number, and general size. Because all the genetic systems are recessive with the exception of flowering time, recessive alleles would be spread by natural selection in the original outbreeding parent species. In inbreeders, there is no difference between dominant and recessive alleles.

Mimulus cupriphilis has evolved recently since the copper mines are less than 150 years old. The shift to the new soils created by the mines involved the development of a copper-tolerant ecotype with a primary semidominant adaptive mutation that shifted flowering time earlier in the plants growing in dry habitats. In the absence of pollinators early in the season, selection favored alleles for self-fertilization. This was accomplished by flower size reduction that brought the stigma and anthers in close proximity. Reduction in size of corolla and associated structures may have also freed up resources for seed production. The result is a new selfing species that has evolved locally in a very short time and that is well adapted to the unique conditions of the mine tailing fields.

There is no evidence of a genetic revolution or reduction in genetic diversity in the new species and F_2 progeny. Nor is there evidence of significant breakdown in so-called gene complexes or major epistatic interactions. Therefore, it is concluded that speciation has been achieved by new selection pressures on normal *M. guttatus* as it colonized a new and unusual habitat. Major genes of large effect that play important roles in reproductive isolation have also been demonstrated in other sister species of *Mimulus* species (Bradshaw *et al.*, 1995). In such cases, new species have the potential to appear nearly full-blown in relatively few evolutionary steps because of changes in a few essential genes.

Sympatric Speciation

It is now apparent that sympatric speciation, once thought to be rare in animals and plants, occurs more frequently than previously realized. Because two closely related sympatric sister species cannot share the same resources, nonallopatric speciation is inevitably accompanied by the shift and exploitation of a new habitat by the daughter species. Such habitat or resource shifts reduce competition between sister species. It is for this reason that sympatric speciation is often referred to as ecological speciation.

Sympatric Speciation in Cichlid Fish

The important role of ecology in sympatric speciation is exemplified by the study of Schliewen *et al.* (1994). In a molecular phylogenetic study they discovered that two endemic cichlid fish species flocks (9 and 11 species) in two small, ecologically monotonous volcanic lakes (4.15 and 0.6 km^2) in Cameroon speciated sympatrically. Because each crater is isolated from rivers and lacks internal structure, past lake levels are not responsible for physically isolating sister populations during speciation. Field observations, stomach content analyses, and the presence of specialized morphological features related to feeding confirm that the species within each lake have different benthic or pelagic trophic and reproductive ecologies. Although all species nest on the bottom and are capable of encountering one another at moderately high frequencies along ecotones, no hybrids are found because the species mate assortatively and do not interbreed.

Phylogenetic trees were constructed based on an analyses of a 340-base pair fragment of mtDNA cytochrome b and an additional 350 base pairs from the rapidly evolving mitochondrial control region for all 20 species and all tilapine species in neighboring river systems and lakes. This analysis confirmed that the lake species are monophyletic with respect to the river species, i.e., each flock evolved within each lake after a single colonization event.

Sympatric Host Race Formation and Speciation in Rhagoletis *Fruit Flies*

The importance of habitat and resource shifts in the nonallopatric speciation process is exemplified in the case of recent host race formation and speciation in the tephritid fruit fly genus *Rhagoletis* (Bush and Smith, 1998; Feder, 1998). As in other phytophagous and parasitic insects, mating in these flies occurs on the host plant, primarily on the host fruit in which eggs are deposited and larvae develop. *Rhagoletis pomonella*, whose native hosts are fruits of several hawthorn (*Crataegus*) species, colonized apples in approximately 1860, more than 200 years after this fruit tree was introduced to North America by Europeans. Early infestations on apples were restricted to a small area of the Hudson River valley in southern New York where hawthorn is common. The apple flies later spread over the entire northeastern United States and southeastern Canada, where it became a major pest. In a brief span of 150 years, the apple and hawthorn populations diverged in several genetically based traits, such as eclosion time of the adult in summer, and host recognition and acceptance. In addition, strong frequency differences in alleles at several loci coding for proteins are maintained by strong selection and greatly reduced gene flow between the races. The apple and hawthorn races now behave as semispecies or species. All seven members of the *R. pomonella* species group are sympatric in eastern North America and speciation in this species group has always been accompanied by a shift to a new host plant family or genus.

Other Modes of Nonallopatric Speciation

Spontaneous Thelytokous Speciation

This mode of sympatric speciation occurs when a unisexually reproducing taxa arises spontaneously from an unfertilized egg of a diploid bisexual species. Subsequent reproduction in taxa originating in this way produces only females from unfertilized eggs. Several good examples are provided by White (1978).

Autopolyploid Speciation

Occasionally, polyploidization of a diploid species may occur spontaneously in one or more individuals. Because autopolyploid individuals have three or more chromosome sets, each chromosome has more than one homologous pairing partner. During meiosis, multivalents are produced leading to unbalanced gametes and zygotes, sterility, and other problems. Only rarely does autopolyploidy result in the origin of new species, such as in the common potato (*Solanum tuberosum*) and its relatives (Grant, 1981). These usually originate from crosses between races whose chromosomes differ only slightly.

Speciation by Interspecific Hybridization

A new species can arise two ways by interspecific hybridization (**Figure 1f**). Homoploid hybrid speciation results in a diploid-derived species, whereas polyploid hybrid speciation produces a species that combines a complete set of chromosomes from each hybridizing parental species. Hybrid species occupy habitats different from those of the parental species, thus reducing competition and the level of gene flow between them. Hybrid speciation, which is far more common in plants than in animals, can occur in at least four ways.

Introgressive Hybrid Speciation

Individual gene exchange among closely related species provides recombinant offspring that shift to and exploit a new habitat not utilized by either parental species. The hybrid species may be interfertile with one or both parental species, but it is reproductively isolated from them by premating barriers to gene flow. This mode of speciation has been reported in several plant species (Grant, 1981), but confirmation of this mode of speciation remains controversial and requires definitive experimental and analytical studies.

Recombinational Hybrid Speciation

A far more common mode of hybrid speciation involves the formation and establishment in the progeny of a chromosomally sterile or semisterile species hybrid of a new, structurally homozygous recombination type. Individuals are fertile within the line but isolated from other lines and from the parental species by a chromosomal sterility barrier. It is most likely to occur when the hybrid interface is long and the organisms involved are predominantly selfing, relatively fertile, and possess few structural chromosome differences between the parental species.

Hybrid Speciation in Wild Sunflowers

A molecular study of hybrid speciation in the wild sunflowers *Helianthus* by Rieseberg *et al.* (1995) revealed that F_1 hybrids of *H. annus* and *H. petiolaris* are semisterile with pollen viabilities less than 10% and seed set less than 1%. F_2 pollen viability is highly variable, ranging from 13 to 97%. The two species are distinguished by several morphological and chromosomal features, and based on chloroplast DNA and nuclear ribosomal DNA variation they occur in divergent clades. Although the species are sympatric throughout much of the western United States, they have different ecological requirements. *Helianthus annus* is restricted to heavy, clay soils, whereas *H. petiolaris* predominantly inhabits dry, sandy soils.

Helianthus anomalus is a rare endemic to xeric habitats in northern Arizona and southern Utah. It is well within range of parental species and is a recombinational hybrid resulting from a cross between *H. annus* and *H. petiolaris*. The F_1 hybrids with parental species are partially sterile because chromosomal structural differences enhance reproductive isolation. A preliminary survey of 126 loci in natural populations of the parental species indicated that *H. anomalus* has loci derived from both *H. annus* and *H. petiolaris*. Some blocks of markers, possibly protected from recombination, are transmitted intact.

Helianthus anomalus combines rDNA repeat units and allozymes of *H. annus* and *H. petiolaris* as predicted for diploid hybrid species, although individuals possess chlDNA haplotypes of *H. annus* and *H. petiolaris* rather than a unique haplotype. Genetic linkage maps generated for all three species using random amplified polymorphic DNA markers reveal loci distributed onto 17 linkage groups corresponding to the haploid chromosome number of the three species. Although levels of polymorphisms vary from 212 in *H. annus* to 400 in

H. petiolaris, map density is similar among species. By comparing genomic location and linear order of homologous markers, chromosomal structural relationships were inferred among the three species.

Even though 6 linkage groups showed no changes in all three species, the remaining 11 linkages were not conserved in gene order. The parental species differ from *H. anomalus* by at least 10 separate structural rearrangements, 3 inversions and a minimum of 7 inter-chromosomal translocations. The genome of *H. anomalus* is thus extensively rearranged relative to its parents. All 7 novel rearrangements in *H. anomalus* involve linkage groups that are structurally divergent in parental species, suggesting that structural differences qmay induce additional chromosomal rearrangements upon recombination.

Allopolyploid (Amphiploid) Speciation

Interspecific hybridization can also result in combining two or more complete chromosome sets. F_1 hybrids produced between two established related species are often sterile because chromosomes lack sufficient homology to pair well at meiosis. Fertility is restored if hybrids persist long enough by asexual reproduction until somatic doubling of the chromosomes can occur in a flower, or until there is a rare union between two unreduced gametes. A new sexually reproducing species is then established that is "instantaneously" isolated from both parental species (Grant, 1981).

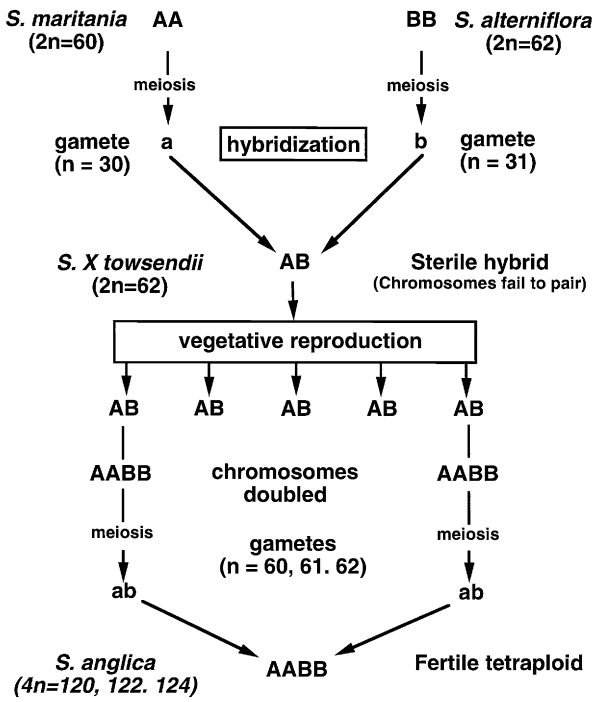


Figure 3 The origin of the allopolyploid marsh grass species *Spartina anglica* (AABB) that resulted from hybridization between *S. maritima* (AA) and *S. alternifolia* (BB). Each letter in parentheses represents a haploid genome. The sterile hybrid, *S. x townsendii* (AB), reproduced vegetatively until the genomes of individual plants doubled by autopolyploidy. This results in the establishment of a new fertile sexually reproducing tetraploid species, *S. anglica*, with a full complement of chromosomes from *S. maritima* and *S. alternifolia* (AABB) that can produce normal gametes (ab).

Allopolyploid Speciation in Spartina Anglica

The recent natural rapid evolution of the amphiploid perennial salt marsh grass, *Spartina anglica*, provides an example of allopolyploid speciation (Raybould *et al.*, 1991). This species originated on the south coast of England at the end of the nineteenth century. It arose as a result of chromosome doubling in *S. x townsendii*, a hybrid between the native British *S. maritima* and the North American *S. alterniflora*, introduced by shipping (**Figure 3**). *Spartina anglica* is now widespread along the English coast and is highly successful.

Although more than half of all plant species are directly or indirectly the by-products of allopolyploid speciation, allopolyploid speciation is relatively rare in animals (White, 1978).

Direct and Reticulate Allogenous Speciation

There are two modes of allogenous speciation (i.e., combining the genomes of two distinct species). In the case of direct allogenous speciation, hybridization between two bisexual, closely related species combines the genomes of two distinct parental species giving rise to a new, unisexual species (Bullini, 1994). The hybridization event produces either an allodiploid or an allopolyploid unisexual species that acquires clonal (parthenogenesis) or hemiclonal (hybridogenesis; i.e., it must mate with males of a bisexual parental species) modes of reproduction. In most such clonal and hemiclonal organisms, the heterozygous genetic structure of the parental species is retained.

In the case of reticulate allogenous speciation, individuals of unisexual hybrid taxon hybridize with a bisexual relative giving rise to new, unisexual species, often with a higher ploidy level than that of the parental species (Bullini, 1994). Most have highly heterozygous genomes that display heterosis. Clones produced by direct and indirect allogenous speciation exhibit heterosis and demographic advantage over both parental species and individuals.

Summary

After 140 years of speciation research, it is now clear that animal and plant species can originate in a variety of ways. Rapid autopolyploid speciation is common in plants and accounts for many plant species. In animals, this mode of speciation is relatively rare. Dichopatric speciation gives rise to many, possibly the majority, of the species in some animal groups such as the land vertebrates. In fish, both allopatric and nonallopatric speciation have been reported. The majority of living organisms, however, are insects. Some authorities estimate there are as many as 30–40 million species, and all agree that there are more than 10 million. Approximately 70–75% are highly specialized parasites that feed in or on plant and animal tissue. Although many of these insects probably originated by nonallopatric speciation, it is not clear what percentage have done so. The same is true for the many mites and nematode species that also may number in

the millions. Because the mode of speciation in only an extremely small number of animal and plant species has been established, it is impossible to estimate how often each mode of speciation occurs in any particular group of organisms. Such estimates must await research on a great many more taxa.

See also: Cladogenesis. Speciation, Theories of. Species, Concepts of. Species Diversity, Overview

References

Barton NH (1996) Natural selection and random genetic drift as causes of evolution on islands. *Philos. Trans. R. Soc. London B* 353: 785–794.

Bradshaw HD, Wilbert SM, Otto KG, and Schemske DW (1995) Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers *(Mimulus)*. *Nature* 376: 762–765.

Bullini L (1994) Origin and evolution of animal hybrid species. *Trends Ecol. Evol.* 9: 422–426.

Bush GL (1994) Sympatric speciation in animals: New wine in old bottles. *Trends Ecol. Evol.* 9: 285–288.

Bush GL and Smith JJ (1998) The genetics and ecology of sympatric speciation A case study. *Res. Population Ecol.* 40: 175–187.

Butlin R (1987) Speciation by reinforcement. *Trends Ecol. Evol.* 2: 8–13.

Carson HL (1975) The genetics of speciation at the diploid level. *Am. Nat.* 109: 83–92.

Carson HL and Templeton AR (1984) Genetic revolutions in relation to speciation phenomena: The founding of new populations. *Annu. Rev. Ecol. Syst.* 15: 97–131.

Feder JL (1998) The apple maggot fly, *Rhagoletis pomonella*, flies in the face of conventional wisdom about speciation. In: Howard DJ and Berlocher SH (eds.) *Endless Forms: Species and Speciation*, pp. 130–144. Oxford: Oxford Univ. Press.

Feder JL, Reynolds K, Go W, and Wang EC (1995) Intra- and interspecific competition and host race formation in the apple maggot fly, *Rhagoletis pomonella* (Diptera: Tephritidae). *Oecologia* 101: 416–425.

Futuyma DJ (1998) *Evolutionary Biology*. Sunderland, MA: Sinauer.

Grant V (1981) *Plant Speciation*. New York: Columbia Univ. Press.

Hollocher H (1996) Island hopping in *Drosophila*: Patterns and processes. *Philos. Trans. R. Soc. London B* 351: 735–743.

Johnson P, Hoppensteadt F, Smith J, and Bush GL (1996) Conditions for sympatric speciation: A diploid model incorporating habitat fidelity and nonhabitat assortative mating. *Evol. Ecol.* 10: 187–205.

Macnair MR and Gardner M (1998) The evolution of edaphic endemics. In: Howard DJ and Berlocher SH (eds.) *Endless Forms*, pp. 157–171. Oxford: Oxford Univ. Press.

Mallet J (1995) A species definition for the new synthesis. *Trends. Ecol. Evol.* 10: 294–299.

Mayr E (1942) *Systematics and the Origin of Species*. New York: Columbia Univ. Press.

Mayr E (1954) Change of genetic environment and evolution. In: Huxley J, Hardy AC, and Ford EB (eds.) *Evolution as a Process*, pp. 157–180. London: Allen and Unwin.

Raybould AF, Gray AJ, Lawrence MJ, and Marshall DF (1991) The evolution of *Spartina anglica*. *Biol. J. Linnean Soc.* 44: 369–380.

Rieseberg LH, Fossen CV, and Desrochers AM (1995) Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* 375: 313–316.

Schliewen U, Tautz D, and Pääbo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368: 629–632.

Taylor EB, McPhail JD, and Schluter D (1997) History of ecological selection in sticklebacks: Uniting experimental and phylogenetic approaches. In: Givnish TJ and Sytsma KJ (eds.) *Molecular Evolution and Adaptive Radiation*, pp. 513–534. Cambridge, UK: Cambridge Univ. Press.

Templeton AR (1980) The theory of speciation via the founder principle. *Genetics* 94: 1011–1038.

Wake DB (1997) Incipient species formation in salamanders of the *Ensatina* complex. *Proc. Natl. Acad. Sci. USA* 94: 7761–7767.

Wake DB and Schneider CJ (1998) Taxonomy of the Plenthodontid salamander genus. *Ensatina, Herpetologia* 54: 279–298.

White MJD (1978) *Modes of Speciation*. San Francisco: Freeman.

Speciation, Theories of

Hope Hollocher, University of Notre Dame, Notre Dame, IN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Coadapted gene complexes Genes that have been selected to work in a coordinated fashion to confer high fitness.
Epistasis An interaction between different alleles at two or more loci such that the phenotype differs from what would be expected if each locus acted independently.
Fitness The average number of offspring produced by individuals with a certain genotype relative to the number produced by individuals with other genotypes as a result of differences in survival and reproductive success.
Gene pool The collective set of genes in a population at a particular time.
Heterogamety Having two different sex chromosomes.
Hybrid inviability The phenomenon whereby one or both sexes of progeny produced in crosses between two different species is unable to survive.
Hybrid sterility The phenomenon whereby one or both sexes of progeny produced in crosses between two different species are able to survive, yet are unable to reproduce.
Pleiotropy One gene affects more than one trait.
Reproductive isolation Intrinsic barriers to the production of offspring.

Introduction

Speciation is the process by which new species are formed. Thus, its central importance to the study of evolution and issues of biodiversity cannot be overstated. The seemingly simple definition of speciation as the origin of species belies the fact that the actual process of speciation is almost as complex and varied as the diversity of life itself. No one mechanism of speciation is sufficient to describe the origination of all the diversity we see in the world and more often than not several speciation mechanisms come into play simultaneously in the formation of a single new species. The study of speciation attempts to analyze these competing mechanisms to determine the relative importance of different evolutionary processes promoting species divergence under particular circumstances. By looking for recurring patterns across a wide spectrum of individual case studies, we are devising models that help explain speciation more generally. This knowledge then serves as a framework for making policy decisions regarding the preservation of biodiversity, for the ultimate goal of conservation biology is not just to preserve current species, but also to preserve the natural processes that help generate new ones.

Theoretical Considerations

Although speciation has been studied intensively for more than 150 years, only recently have we begun to develop a clearer understanding of the general patterns of genetic change that occur during the process. This type of knowledge is absolutely essential for describing speciation using a population genetic framework. A key aspect of all population genetic models of speciation is the assumptions made about the underlying genetic architecture – the number and types of genes and their interactions – of traits that have changed during speciation. Differing underlying genetic architecture of traits can influence directly how they respond during speciation even when the same evolutionary processes are known to be at play (e.g., see Rettelbach *et al.*, 2011; Nosil and Schluter, 2011). Therefore, examination of the genetic architecture of species differences is one fruitful avenue for gaining insights into the evolutionary processes involved in speciation.

Equally important to theories of speciation is determining which traits are most influential for promoting species divergence. We recognize species because they are different from one another, be it morphologically, behaviorally, or reproductively. In all cases, a sequence of events has led to the eventual fixation of these observed differences. Knowing whether certain traits are more prone to change than others during speciation or whether certain traits have a larger impact either on past or future patterns of divergence directly informs our models of speciation. This knowledge improves our ability to target groups of organisms at the nascent stage of divergence and to focus our attention on specific traits for future studies of speciation that will allow general patterns to emerge.

Defining Species

Why is There More Than One Definition of Species?

One would think that by now evolutionary biologists would agree on what constitutes a species. However, it turns out that species concepts are almost as varied as the types of researchers that work on the question. To date there is no single, universally accepted definition of a species, although certain frameworks for thinking about species have come in and out of favor over the years (Howard and Berlocher, 1998; Coyne and Orr, 2004; Wilkins, 2009). Most definitions of species serve a useful research purpose and are a function of the types of questions that are being addressed (e.g., whether species designations should serve a strictly taxonomic purpose as in systematics or represent the active evolutionary unit of diversification as in population genetics). This is not to imply that the definition of species is purely a semantic issue, but rather that the diversity of definitions reflects very different perspectives as to what factors should be emphasized in

describing species. Beyond these purely theoretical issues, species concepts have an additional problem of not always being easy to apply in real-life situations.

As a result of these conflicting purposes of species concepts, every subdiscipline within speciation research has formulated its own particular species concept. However, the problems surrounding species definitions are more fundamental than simple differences in perspective. There is an inherent problem of trying to categorize groups of organisms that are undergoing a continual process of change. Speciation is a process, not an instantaneous event. It is also not caused by a single unified mechanism, but by a collection of different mechanisms of divergence affecting different traits at potentially different rates. The formation of natural groups does eventually result from the ongoing action of each of these processes of divergence and at some point in time these natural groups can and should be delineated as separate species. However, because of the fluidity of the process, the boundaries that circumscribe these natural groups are often at best blurry given the fact that each underlying mechanism of divergence can potentially delineate a separate natural group. In some instances, this creates no problem and these different boundaries will actually coincide with each other. In these cases, several different species concepts will correspondingly give the same species designations. In other instances, these boundaries may not correspond well at all. Hence, depending on which process is being emphasized, the actual definition of a particular species may also vary in these circumstances. In most cases, conflicting species boundaries are not so much a problem of the competing species definitions as they are a reflection of the complicated nature of the underlying biological process that governs speciation. For a population geneticist interested in the process of speciation, it is those cases where the boundaries of divergence based on different criteria do not correspond well that often prove to be most interesting to study.

Most Commonly Used Species Concepts

Because of space constraints, a thorough critique of competing species concepts is beyond the scope of this article (see the chapters by De Queiroz, Harrison, Shaw, and Templeton in Howard and Berlocher, 1998; for a more philosophical and historical treatise see Wilkins, 2009). However, a general overview of the most commonly used species concepts will be presented here, paying particular attention to how each concept relates to the elucidation of underlying speciation processes.

Morphological and Phenetic Species Concepts

Intuitively, using morphological criteria to get a first approximation of species boundaries works rather well. Different groups of organisms look different from each other. This simple "morphological species concept" was used by the early naturalists of the eighteenth and nineteenth centuries and is still often used today by field biologists and paleontologists. Most evolutionary biologists also use this concept as a starting point to which other criteria are then later added (the fact that species appear morphologically distinct is usually taken for granted by most practicing evolutionary biologists). The "phenetic species concept" embellishes on this first approximation by incorporating specific morphometric analyses to describe explicitly the phenotypic groupings. From a functional point of view, groups of organisms that are morphologically distinct most likely represent different groups genetically as well.

Using morphological criteria alone, however, carries with it the risk of categorizing groups of organisms as the same species when indeed there are reproductive barriers between them (as is the case for cryptic species, which are morphologically identical but nevertheless represent distinctly different gene pools) or categorizing groups of organisms as separate species when in fact the difference is not genetically based, but solely a phenotypic response to varying ecological conditions (such is the case for some plants that can change the morphology of their leaves in response to varying light levels). In addition, the categorization of species based solely on morphological criteria without the incorporation of an explicit evolutionary framework is generally unsatisfactory to the majority of people working in the field of speciation.

The Genotypic Species Cluster Definition

An improvement over most morphological approaches to defining species has been proposed by Mallet (1995) using genetic criteria to cluster groups of organisms rather than strictly morphological traits (termed the "genotypic species cluster" definition). In this case species are defined as genetically distinguishable groups of individuals that exhibit few or no intermediate genotypes (both for single and multiple loci) when in contact. This is clearly an advantage over strict morphological criteria because it gets right to the heart of speciation – genetic differentiation. It is also very satisfying as an operational definition of species groups, because genetic clustering is easily diagnosable for all taxa, whether sexual or asexual. However, this definition is only easily applicable to the situation where species are in direct contact with each other and cannot be generally applied. More importantly, it is still a phenetic approach to species definitions and lacks any inherent evolutionary framework for the interpretation of species groups from a mechanistic point of view.

The Biological and Recognition Species Concepts

Until recently, the most enduring species concept has been the "biological species concept," first outlined in detail by Dobzhansky (1937) and subsequently popularized by Mayr (reviewed in Mayr, 1982). The concept defines species as interbreeding groups of organisms that are reproductively isolated from other groups. In other words, species are groups of organisms that share a common gene pool. Speciation, under this concept, is the development of a biological barrier to gene flow and the subsequent differentiation of the two resulting gene pools. Biological barriers to gene flow were termed isolating mechanisms and covered the full gamut from prezygotic barriers (including ecological or habitat isolation, seasonal or temporal isolation, sexual or ethological isolation, mechanical or physiological barriers to mating, isolation by different pollinators, and gametic isolation blocking fertilization) to postzygotic barriers (including hybrid sterility, hybrid inviability, and other forms of reduced fitness in progeny

resulting from crosses between differentiated populations). The biological species concept is firmly rooted in speciation processes and encompasses a number of different mechanisms that can serve to reduce the exchange of genes between different populations and thus allow them to evolve independently.

The "recognition species concept" is very similar to the biological species concept in that species are defined by interactions among its members (Paterson, 1985). In addition, both concepts try to pinpoint traits that are biologically relevant in species divergence. In the case of the recognition concept, species are defined as groups of organisms that share a common system of fertilization or specific mate recognition system; i.e., the emphasis is placed on what holds species together as a cohesive reproducing group of organisms rather than on what may separate them. The same general concept is embedded in the biological species concept, but the emphasis on mate recognition systems versus isolating barriers does influence how questions about speciation mechanisms are formulated and places more importance on evolutionary processes that positively influence the specific mating interactions between members of a species rather than on the more negative effects associated with preventing mating from occurring between members of different species. In either case, both of these concepts are mechanistically based and represent a prospective view of species distinctions by giving an indication of how isolated groups are now and will continue to be in the future.

Phylogenetic, Evolutionary, and Genealogical Species Concepts

Other more recently derived species concepts rely more heavily on a retrospective view by defining a species in a strictly historical sense as a separate evolutionary lineage that is internally connected through time (i.e., the "phylogenetic species concept" (Cracraft article in Otte and Endler, 1989; Nixon and Wheeler, 1990), the "evolutionary species concept" (Wiley, 1978), and the "genealogical species concept" (Baum and Shaw, 1995); for a good overview, see Harrison article in Howard and Berlocher, 1998; likewise, de Queiroz, 2007). These species concepts are less oriented toward process and identifying specific biological traits that maintain cohesion within species or promote divergence between species and instead are more oriented toward the final evolutionary result – lineage divergence. That is not to say that process cannot be usefully inferred from looking at historical patterns of lineage sorting and splitting (*see* The Cohesion Species Concept and Incorporation of Genealogical Models), but the emphasis in these particular species concepts is clearly on pattern rather than on trying to incorporate a variety of biological processes into the species definitions themselves.

The Cohesion Species Concept

Perhaps the most inclusive species concept that has been proposed to date is the "cohesion species concept" (see Templeton in Otte and Endler, 1989; see also Templeton in Howard and Berlocher, 1998). This single species concept combines key elements of the biological, recognition, phylogenetic, evolutionary, and genealogical species concepts in order to form a comprehensive framework for defining species and studying speciation mechanisms. Like the recognition species concept, the emphasis in this case is on defining species based on mechanisms that provide cohesion among members of a population. Unlike the recognition concept, cohesion involves not just shared fertilization systems promoting gene flow (termed "genetic exchangeability" in this model) but also shared historical fates (i.e., belonging to a common evolutionary lineage) as well as shared evolutionary tendencies (i.e., experiencing common selective regimes), both of which will maintain species cohesion (through "demographic exchangeability") in the absence of gene flow. With this definition, the repertoire of traits important to speciation is broadened beyond traits determined primarily by their effects on reproduction and includes any traits involved in ecological adaptation, regardless of whether or not they directly affect the probability of genetic exchange. Although some would argue this more inclusive, pluralist species concept obfuscates the ability to clearly define species (if anything goes, then what should be the object of study?), in reality, the concept serves as a better starting point for speciation studies precisely because it is not narrowly focused on any one particular speciation mechanism to the exclusion of all others. The practical application of the cohesion species concept has been improved of late by incorporation of historical patterns of genetic variation to use as an explicit framework for both testing hypotheses of species boundaries and investigating speciation processes by modeling how different evolutionary forces pattern this genetic variation under different geographical scenarios (Templeton in Howard and Berlocher, 1998; Templeton, 2007; *see also* Incorporation of Genealogical Models).

Evolutionary Processes Involved in Speciation

Because genetic change lies at the heart of speciation, the formation of new species involves the action of the same evolutionary processes known to cause genetic changes within populations. The principal evolutionary processes most commonly involved in species divergence are natural selection, genetic drift, sexual selection, and mutation. All of these processes act to change allele frequencies within populations over time and there is no question that they operate during speciation as well. The challenge of speciation studies comes from understanding how these evolutionary processes operate alone and in conjunction with each other within very specific geographical, ecological, behavioral, and genetic contexts. Change any one of these contexts and the same evolutionary process can give rise to very different outcomes and also impact the resulting genetic patterns of speciation.

Natural Selection

Darwin (1872) was the first to popularize the idea that speciation could readily occur through the prolonged action of natural selection, the process by which genetic variants that are better suited to the natural environment increase in frequency while variants that are not well suited decrease. In Darwin's model of speciation, competition to survive

was thought to be extremely intense. Therefore, the modification of descendants through natural selection served to increase the adaptation of any one population to the environment while also acting to decrease the intense competition between populations by promoting diversification. Over time, speciation would result automatically under these conditions.

Because of the general acceptance of Darwin's idea that natural selection is a potent driving force in species divergence, it is often erroneously considered the only evolutionary force important in species formation. However, natural selection can act to cause change (e.g., through directional or disruptive selection as imagined originally by Darwin) as well as to prevent change (e.g., through balancing selection). Because of the dual nature of selection, speciation studies need to distinguish between those circumstances under which natural selection is promoting speciation versus those where natural selection is a force that actually maintains cohesion and needs to be overcome in order for divergence to occur.

Natural selection's role as a powerful agent for stasis rather than change comes into play in speciation studies in many instances. One particular instance that has been the focus of a prolonged debate in speciation studies surrounds the issue of coadapted gene complexes. Organisms are required not only to adapt to the external environment but also to their own internal genetic environments. Organisms develop and survive as a result of the coordinated control of a complex network of genes selected to function together properly. All possible genetic changes that may help an organism survive in a particular external environment must also function within the constraints posed by this internal genetic environment in order for these changes to be propagated through the population. Genes that have been selected to work particularly well together are called coadapted gene complexes. Once coadapted gene complexes have evolved, it can be very difficult for natural selection to deviate from the established optimum (termed a fitness peak) in order to reach a new optimum that may ultimately be better for the population, but would require the population to experience a temporary loss of fitness (through suboptimal gene interactions) during the transition.

Wright (1932) coined the phrase fitness landscape or adaptive landscape to describe this complicated behavior of fitness as a function of gene interactions. Different genotypic combinations plotted in two dimensions create a complex fitness landscape in the third dimension, much like how a mountainscape appears from an airplane. Mountain peaks correspond to regions of high fitness (for particularly good combinations of genes) separated by valleys of low fitness (for not so good combinations of genes). Natural selection, being somewhat myopic, would work to drive populations up a local peak, but would not be able to take populations across valleys to neighboring peaks of fitness even if those peaks represented greater overall fitness.

Because the true nature of adaptive landscapes is a point of contention for competing models of divergence (see Coyne and Orr, 2004; Gavrilets, 2004), much work in population genetics has been geared toward understanding what the fitness landscape really looks like under a variety of circumstances (for a review, see Wolf *et al.*, 2000). In spite of its importance, it has been relatively difficult to get a firm handle on the topology of adaptive landscapes partly because at any given moment natural populations tend to occupy a single peak. Advances have been made more recently by studying real and digital microbial populations, which can be followed for tens of hundreds of generations (e.g., see Elena and Lenski, 2003; Clune *et al.*, 2008; Melnyk and Kassen, 2011), and provide more direct information (both actual and theoretical) to help conceptualize the shape of adaptive landscapes and how they influence evolutionary trajectories. Different species themselves represent the successful transition from one optimal state to another that has been frozen in time. Characterizing the genetic transitions that distinguish closely related species gives us clues to the actual shape of the fitness landscape from a complementary perspective (Kao *et al.*, 2010; Wu and Ting, 2004). It also allows us to generate genetic models of speciation that describe how different species are able to make this shift from one fitness peak to another.

Genetic Drift

Given the possible constraints posed by selection acting on coadapted gene complexes, an additional mechanism that would facilitate the movement of a population from one fitness peak to another is often incorporated into models of speciation. Genetic drift is the random change in allele frequencies caused by sampling across generations in a finite population. The actual change in allele frequency caused by genetic drift is random for any given generation; however, the effects of drift will accumulate over time. In the absence of the introduction of new variation through mutation or immigration from another source, genetic drift will serve to decrease the amount of genetic variation present in a population. Because this process is entirely random, how one population responds can differ from how another population responds just by chance, and drift can actually serve to differentiate two populations in the complete absence of natural selection.

In Wright's original model describing adaptive landscapes, genetic drift was one mechanism by which populations could move from one adaptive peak to another. For small population sizes, drift effects can be so strong as to overcome the effects of natural selection. If genetic drift could free the population from the constraints posed by natural selection holding it on one particular fitness peak, then the population could randomly explore the adaptive landscape and possibly fall under the domain of attraction of another fitness peak better than the original one. Individually, genetic drift and natural selection can be very effective for bringing about change. Together they may be even more effective. Although speciation is not specifically addressed in Wright's original model, it does form the theoretical basis for Founder Effect Speciation, which will be discussed in more detail later in this article (*see* Peripatric or Founder Effect Speciation).

Sexual Selection

Adapting to the natural environment is only one manner in which organisms can increase their ability to leave more

offspring for future generations. Another method is to improve their ability to secure mates through the action of sexual selection. Sexual selection arises from differences in reproductive success caused by competition over mates (Darwin, 1872; Andersson, 1994). Because variance in reproductive success can often be quite high (especially for males), adaptation to the mating environment can sometimes be even more effective at promoting change than adapting to the natural environment and has resulted in the rapid divergence of traits related to reproduction, such as the morphology of male genitalia, sperm composition and abundance, and male secondary sexual characteristics and behavior (Eberhard, 1985; Grant and Grant, 1997; Edwards *et al.*, 2005).

The pressure for sexually reproducing organisms to have compatible reproductive systems ensures that male and female reproductive systems will coevolve. Reproductive systems include behavioral traits related to mate recognition, courtship, and mating as well as physiological traits related to gametogenesis, fertilization, and the production of offspring. Reproductive systems are generally subject to several counteracting evolutionary forces that drive the system in different directions; therefore, mating systems within any particular population tend to reach a stable equilibrium (Lande, 1981; Iwasa and Pomiankowski, 1995). Although the population may be evolving directionally through time, individuals who deviate too much from the norm at any given point along this evolutionary trajectory will generally be disfavored either because of mate discrimination or because of natural selection pressures. Although stabilized with respect to any one population, reproductive systems can be very fluid with respect to the number of stable equilibria that are theoretically possible. Therefore, across several populations, traits related to reproduction are particularly susceptible to changes that can eventually lead to speciation (Lande, 1981; Boake, 2005; Maan and Seehausen, 2011).

The pressure to maintain sexual compatibility (positive male–female coevolution) within sexually reproducing populations is only one important evolutionary consideration for the role of sexual selection in speciation. In addition, there is a second evolutionary process operating that involves antagonistic male–female coevolution which is also very effective at bringing about change, especially in the case of internal fertilization (see Rice in Howard and Berlocher, 1998; Parker and Partridge, 1998). The reproductive interests of males and females are not always identical, especially when there is any deviation from strict monogamy. In these cases, evolutionary conflicts can easily arise. For example, females may mate with multiple males to maximize their own reproductive success, but that creates a battleground for sperm competition among males. To combat the intensified selection for sperm to perform better under these circumstances, the chemical composition of male seminal fluid may evolve to include proteins that act to kill other sperm. These proteins in turn may also have negative pleiotropic effects on the female reproductive system, which is then selected to combat these negative effects. In this scenario, polygamy sets the stage for continual antagonistic selection to occur between male and female reproductive systems. As seen earlier, within any one particular population, antagonistic coevolution will force the male and female reproductive systems to track each other very closely. However, the effect is very localized and populations that are not in contact with each other have the potential to diverge rapidly with respect to their reproductive physiology (Wu *et al.*, 1996; Rice article in Howard and Berlocher, 1998; Hosken *et al.*, 2009; Prokupek *et al.*, 2010; Lessios 2011).

Mutation

Neither selection nor drift can operate in the absence of genetic variability. Although quite a bit of variability can be generated just through genetic recombination and the random assortment of chromosomes during reproduction, ultimately spontaneous mutation is the primary source of genetic variation. Mutation is considered random with respect to the environmental challenges faced by organisms, although that is not to say it is without bias. Because mutations are important in the context of their effects on the phenotype of organisms, certain phenotypic classes of mutations can be more common than other classes of mutations. In this respect, most spontaneous mutations tend to be deleterious to the fitness of an organism. In addition, mutations are constrained by the chemical nature of DNA or particular properties of the DNA sequence itself, resulting in certain types of changes being more common than others. This can lead to certain classes of base substitutions being more common and to recurring mutation producing the same allele several times independently as is seen for mutations involving slippage during DNA replication of tandemly repeated sequences.

An important consideration for speciation studies is whether new mutations tend to have large or small phenotypic effects (i.e., does divergence occur in leaps and bounds or through the accumulation of numerous small changes?). To model speciation, it is also important to know whether mutations tend to have dominant or recessive effects with respect to fitness – the dynamics of how allele frequencies change over time are dependent on whether or not alleles at low frequencies (as is always the case for new mutations) are readily subject to natural selection or are masked by the effects of other alleles. We also need to know whether new mutations impact more than one phenotypic trait (i.e., have pleiotropic effects) or interact with other loci (i.e., exhibit epistasis) to generate new phenotypes. All of these factors impact the interaction between new mutations and population genetic processes that can influence the change in frequency and ultimate fixation of new mutations in a population.

Common Geographical Modes of Speciation

The crux of speciation studies involves understanding the processes by which a single population splits to form two diverging lineages. Speciation involves the weakening of forces that hold populations together (such as the presence of gene flow, shared mating systems, recurring mutation, and natural selection in common environments) in favor of processes that drive them to diverge (such as the action of genetic drift, mutation, and natural and sexual selection in different

environments in the absence of gene flow). Most modes of speciation incorporate the action of all the evolutionary forces discussed earlier (natural selection, genetic drift, sexual selection, and mutation), but emphasize the operation of certain processes over that of others. Several of these modes of speciation are discussed elsewhere in this volume (see also Coyne and Orr, 2004); therefore the author will focus discussion here on a subset of all possible modes of speciation, concentrating on the specific interaction of different evolutionary forces operating within the context of different geographical and ecological situations, keeping in mind that most speciation events in actuality will be influenced by a combination of modes and evolutionary processes (see Butlin *et al.*, 2008; Fitzpatrick *et al.*, 2008).

Classical Allopatric Speciation

Perhaps the simplest conceptual framework for speciation is the allopatric model. Intuitively, it is easy to recognize that if a physical barrier prevents populations from exchanging genes, they will be free to follow independent evolutionary trajectories. Extensive geographical variation exists (reviewed in Mayr, 1963); therefore, any complete barrier to gene flow most likely means the operation of natural selection will differ to some extent on either side of the barrier. Even in the absence of varying environmental pressures, the continual action of mutation, genetic drift, and sexual selection (*see* Evolutionary Processes Involved in Speciation) can change the populations relative to each other over time, if they are completely isolated from each other and no longer actively exchanging genes.

Given enough time, speciation will most likely occur under this scenario of no gene flow between isolated populations. Therefore, the interesting issues to explore in allopatric speciation involve examining the relative rates of divergence caused by the action of natural selection, sexual selection, mutation, and drift as well as determining which traits are most easily affected. There are no general rules that can be presented for what happens during allopatric speciation; examples for all possible scenarios (where each of the evolutionary forces has been shown to play a role to a varying degree) can be cited. The specific outcome is very much dependent on the circumstances surrounding the isolation. It has also been readily shown that changes in the reproductive system (both prezygotic behavioral changes in mating propensity and postzygotic physiological changes in the ability to generate offspring) occur rather commonly when populations are separated (reviewed in Coyne and Orr, 2004).

Peripatric or Founder Effect Speciation

Peripatric speciation represents a variation on allopatric speciation. In this case, a small population forms at the periphery of a larger population. This type of geographical isolation can happen anywhere, but is most easy to visualize in the situation of founders colonizing oceanic islands or, more generally, isolated pockets of habitat. Because the founding of peripheral populations can sometimes involve the movement of only a few individuals (or even a single gravid female), this type of speciation has also become known as founder effect speciation. The emphasis here is on the interaction between genetic drift and natural or sexual selection that occurs during the early stages of speciation.

If a new population is founded by a small number of individuals, just by chance the genetic composition of the founding population may differ significantly from that of the original source population because of genetic drift (*see* Genetic Drift). The population need not remain small for very long in order for this sampling effect to influence the future evolutionary trajectory of the population. In addition to this immediate genetic change, oftentimes small populations founded in peripheral habitats or on islands also experience changed environments (both the physical environment, including such things as the quality of the habitat, the distribution of resources, or the presence of competitors or predators, as well as the mating environment, represented by a shift in the distribution of available mating types and preferences), creating new selection regimes. Even in the complete absence of new selective environments, the shift in allele frequencies alone can potentially have a profound effect on how the population will respond to selection because of the changed internal genetic environment that results from drift. The combination of shifting gene frequencies by drift and the presence of potentially new selection regimes under this scenario has led some researchers to propose that this type of speciation can occur more rapidly than the more standard form of allopatric speciation which generally involves populations of larger size and less drastic changes in the physical and mating environment on isolation (for reviews, see Hollocher article in Grant, 1998; Templeton, 2008).

The theoretical framework used to justify the conclusion that speciation would be accelerated during founder effect speciation stems directly from Wright's model of an adaptive landscape (*see* Natural Selection *and* Genetic Drift). A major underlying genetic assumption that enters into the idea that the random sampling of alleles during the founder event can have a profound effect on the evolutionary trajectory of a population is that epistasis (where interactions between alleles at different loci produce phenotypic effects that are not predicted by the action of the individual allelic effects considered alone) and pleiotropy (where a single locus can directly influence more than one phenotypic trait) are quite common. It is under the assumptions of this type of genetic architecture that fitness peaks of varying heights will exist in the adaptive landscape and where random shifts in allele frequencies can have profound effects (e.g., see Gavrilets and Hastings, 1996; Gavrilets, 2004). If allelic effects governing traits important in speciation are more additive (where interactions between alleles at different loci are minimal), then allele frequency changes will not necessarily impact the trajectory of natural selection greatly.

Much of the debate surrounding the likelihood of founder effects accelerating the process of speciation has focused on the specific influence drift alone would have on the probability of shifting from one fitness peak to another (for a reviews, see the Barton and Hollocher articles in Grant, 1998; Coyne and Orr, 2004; Gavrilets, 2004). What has emerged from these theoretical studies has been the idea that the actual

size of the founding population does not play as crucial a role in determining the probability of shifting from one fitness peak to another as does the underlying genetic architecture of fitness. On the basis of these theoretical results, researchers have begun to shift their focus to evaluating the genetic architecture underlying traits that change during speciation to see how often epistasis is an important component (*see* Genetic Patterns and Processes of Species Differentiation; see also Phillips, 2008). In addition to the genetic architecture influencing rates of change, the actual nature of the trait itself can affect the type of response that is expected under founder effect speciation. Reproductive isolation (both prezygotic and postzygotic) can be particularly susceptible to rapid change under this scenario because of the tight coevolution of male and female traits that normally occurs via sexual selection (*see* Sexual Selection). The random sampling of individuals during a founder event can easily move the population away from the stable equilibrium that characterizes the reproductive system in the original population. Reestablishment of a new equilibrium can often involve a radical shift in the mating system of the new population relative to the ancestral one. For sexually selected traits, random genetic drift coupled with sexual selection can act as a particularly powerful mechanism for driving speciation (Lande, 1981; Boake, 2005).

Parapatry and Speciation through Hybridization

Parapatric speciation occurs when new species evolve in contiguous, yet spatially segregated habitats. Unlike allopatric speciation, the populations that are diverging during parapatric speciation maintain a zone of contact and do not cease the exchange of genes completely. In this case, a balance is achieved between continual gene flow and strong natural selection to maintain divergent populations at the two ends of the contiguous habitats. The zone of contact between the two diverging populations, where hybridization between the differently adapted types takes place, is called a hybrid zone (for more general information on hybrid zones, see Harrison, 1993; Hewitt, 2001). There can be a single zone of contact along a linear environmental gradient or several points of contact between habitats that are more patchily distributed. In both cases, the basic dynamics of hybrid zones are similar. In practice, it is generally impossible to distinguish between the situation in which the two populations continually maintained contact during the process of divergence (in which case the hybrid zone would then be considered a primary zone of contact) and the scenario in which the populations were actually allopatric at some point during divergence and then more recently came back into contact (in which case the hybrid zone would be considered a secondary zone of contact).

In either event, hybrid zones are particularly compelling to population geneticists because of the opportunity they present for examining the rate of exchange of genes that may be under different selection pressures in the two habitats. For most stable hybrid zones, the point of contact between the two species represents a semipermeable barrier to gene flow. Examining the distribution of different alleles sampled along a transect through the two habitat types can provide insights into the nature of the selective forces operating during divergence. The distribution of alleles at loci that have absolutely no effect on fitness will be governed entirely by migration distances and rates of exchange between the two species. Genes that are important for maintaining adaptive differences will show entirely different dynamics relative to these more neutral genes, revealing different patterns depending on the number of genes involved in the adaptations and the strength of selection operating in the two different habitats. Examination of the fates of alleles that shows differential movement across the hybrid zone is also a powerful method for actually mapping the genes responsible for adaptation (see Barton article in Harrison, 1993), which has been expanded to include alleles responsible for early species differentiation even outside the context of hybrid zones (Wu, 2001; Via, 2009).

In addition, hybrid zones represent natural laboratories for determining whether there is speciation by reinforcement. If two species coming into contact exhibit some degree of postzygotic isolation, it is then theoretically possible for selection to act on mating behavior (increasing assortative mating) to eliminate the production of hybrids, thus reinforcing the divergence that has already occurred and completing the speciation process (Dobzhansky, 1940; Servedio and Noor, 2003; Servedio, 2004). Reinforcement is a very appealing concept because it allows postzygotic isolation to play an active role in driving speciation rather than simply being a pleiotropic consequence of divergence. Although evidence of reinforcement occurring in nature has been presented (e.g., Noor, 1995; Nosil *et al.*, 2003), theoretical arguments narrowly limit the range of circumstances under which it is likely to occur and it remains unclear how prevalent the operation of reinforcement may actually be in hybrid zones (reviewed in Coyne and Orr, 2004). The ability of reinforcement to cause complete cessation of gene flow between diverging species is highly dependent on the genetic architecture of mate preference mechanisms, potentially limiting the overall efficacy of reinforcement even further (Bank *et al.*, 2012).

Hybridization is generally perceived as creating zones of tension between the operation of natural selection and gene flow. Speciation is thought to be proceeding through the constant action of natural selection serving to increase adaptation at the two extremes of the zone in the face of a continuous influx of genes that hamper adaptation. Intermediate types that form at the zone of contact are often inferior in fitness and disfavored by natural selection. The actual situation is far more complicated than this simple scenario suggests, even for hybrid zones that generally operate in this fashion (see Harrison, 1993). More importantly, it has become increasingly clear that hybridization does not always play such a negative role in speciation. Instead, it has been shown repeatedly that hybridization can actually provide an important arena for evolutionary innovation (Gross and Rieseberg, 2005; Arnold *et al.*, 2011). In this case, hybridization presents the opportunity for the formation of unique genetic combinations through the mixing of different gene pools. In certain circumstances, these unique gene combinations end up performing better than either parental type in particular habitats and can result in the establishment of novel independent evolutionary lineages and the formation of new species (Donovan *et al.*, 2010).

Sympatric Speciation

Sympatric speciation represents the extreme opposite of allopatric speciation by being wholly independent of geographical context. New species form well within the dispersal range of the ancestral species through divergent natural selection to adapt to alternative habitats. Although ecological adaptation is thought to be the most important process in most cases, sympatric speciation can also occur through the sole operation of sexual selection (Higashi *et al.*, 1999). Because gene flow is known to have a powerful homogenizing effect on populations, it was generally thought that sympatric speciation would be impossible. More recently, modeling and several important case studies have revealed this skepticism to be unwarranted (Berlocher and Feder, 2002; Bolnick and Fitzpatrick, 2007). Most well studied examples of sympatric speciation involve shifts from one host to another that occur in the same geographical region. The genetic adaptations involved in host shifts do not appear to be simple and often involve several different loci that interact to influence a wide variety of traits associated with fitness on a particular host. Often adaptation to one host will preclude the ability to do well on the other, especially if changes in life history are needed in order to track the host species more closely. Host-specific fitness trade-offs will then result, creating the opportunity for divergent selection to cause different host-adapted genotypes to segregate in the population. This type of divergent selection alone will not necessarily result in speciation, but if it is combined with a certain level of host fidelity in which individuals who are reared on a particular host and then return to feed and mate on that same host as adults, then there is a much greater probability that different populations specifically adapted to one host or another will evolve and give rise to new species. Although the focus of sympatric speciation studies typically has been on host shifts, the same principles also apply to adaptation to specific environments and habitat selection more generally (Schluter, 2000).

Genetic Patterns and Processes of Species Differentiation

It is clear from the above discussion that the tempo and mode of speciation not only are affected by geographical and ecological considerations but also can be directly influenced by the nature and genetic architecture of the traits that are diverging. Whether genetic drift can act as a facilitator of speciation during founder effect speciation, or whether specific adaptations generally result in fitness trade-offs that lead to divergence in sympatric speciation, is partially dependent on how commonly epistasis and pleiotropy underlie traits subject to change during speciation. In addition, specific models to explain genetic patterns of postzygotic reproductive isolation (*see* The Genetics of Species Differences) predict very different outcomes depending on whether isolation involves a few or many genes and whether they act recessively or not. To begin sorting through these different possibilities, it is necessary to evaluate the genetic basis of traits that have diverged during speciation to see if any common genetic patterns begin to emerge.

Ultimately, assaying the genetic basis of traits having diverged between species could be helpful in determining whether genetic patterns of change inform us about the evolutionary processes themselves that have played a major role during divergence. If this is true, then genetic patterns of change alone could serve as genetic road maps for navigating the process of speciation. Although we are still in the early stages of research that examines speciation from a strictly genetic point of view, population genetic models investigating what the properties and distribution of these genetic changes would look like under different evolutionary scenarios and across different time scales have begun to be formulated (e.g., see Hey, 2006; Mustonen and Lassig, 2009; Wolf *et al.*, 2010).

The Genetics of Species Differences

A remarkable and repeated pattern in speciation is the tendency for interspecific hybrids to be either sterile or inviable (collectively known as postzygotic reproductive isolation). This phenomenon is so striking and common that it forms the central premise of the biological species concept (*see* Most Commonly Used Species Concepts) and has been the focus of intense genetic study over the past several decades. A central organizing principle for the study of the genetics of postzygotic reproductive isolation has been Haldane's rule. Haldane (1922) observed, "When in the F1 offspring of two different animal races one sex is absent, rare or sterile, that sex is the heterogametic sex." As Haldane's rule is obeyed when males are the heterogametic sex as well as when females are the heterogametic sex, the genetic mechanisms that explain Haldane's rule are thought to be fundamental to speciation in all taxa.

Several explanations for Haldane's rule have been posited and an almost equal number have been rejected (reviewed in Hollocher, 1998; Coyne and Orr, 2004). Part of this controversy stems from the difficulties inherent in performing genetic analyses on such traits as hybrid sterility and hybrid inviability. Part also stems from the fact that most researchers naturally sought a single, universally applicable genetic mechanism to account for all cases conforming to Haldane's rule. The general consensus today is that Haldane's rule requires separate explanations depending on whether or not hybrid inviability or hybrid sterility is being considered.

For hybrid inviability, Haldane's rule results because genetic incompatibilities causing inviability that evolve between species tend to act recessively. Given this recessivity, genes causing hybrid inviability will be expressed in the heterogametic sex while remaining masked in the homogametic sex, thus generating Haldane's rule. For hybrid sterility, the explanation is more complicated and involves the joint action of several different processes. As is the case for hybrid inviability, Haldane's rule for hybrid sterility results because genetic incompatibilities causing sterility also tend to act recessively. In addition, however, when males are the heterogametic sex, the evolution of hybrid male sterility is accelerated, most likely due to the additional action of sexual selection driving the rapid divergence of male sexual traits (*see* Sexual Selection; Wu *et al.*, 1996; Rice article in Howard and Berlocher, 1998; Coyne and Orr, 2004). Much detailed work on the

characterization of the genetic basis of postzygotic reproductive isolation (namely male hybrid sterility) has revealed that many genes of small effect are involved and that epistasis is of primary importance for the expression of this trait (reviewed in Wu and Palopoli, 1994; see Wu and Hollocher article in Howard and Berlocher, 1998; Orr, 2001; Coyne and Orr, 2004).

Postzygotic reproductive isolation is only one trait that has generally diverged between species. To formulate a broader picture of the genetic patterns of divergence, it is necessary to compare the genetic architecture of this trait with what is seen for other traits that diverge between species (reviewed in Hollocher, 1998; Orr, 2001). The traits that have been looked at in some genetic detail include interspecific mate discrimination and interspecific differences in secondary sexual characteristics as well as more general morphological traits. Overall, the pattern that emerges is that these traits, too, are governed by many genes of small effect. However, in contrast to what was found for the genetic basis of postzygotic reproductive isolation, epistasis does not play a dominant role in governing the evolution of these species differences. Because both hybrid sterility and hybrid inviability are only manifested when gene interactions are rendered incompatible, the ubiquitous and rapid evolution of these traits reflect the highly integrated nature of gene networks operating in living organisms and the internal selection pressure these networks exert on physiological and developmental systems to constantly adjust to small changes in the genome that accumulate through the normal action of mutation, selection, and drift. Changes in morphological traits, however, are more tightly circumscribed developmentally. Although morphological traits also require integrated gene networks for their proper expression, genetic changes that evolve between species must still be able to operate well within that particular developmental system, whereas hybrid sterility and hybrid inviability signal the breakdown of gene interactions more globally.

Genetic Variation Within Versus Between Species

Comparison of patterns of genetic variation within species versus patterns of genetic variation between species for the same traits can be useful for gaining insights into the evolutionary mechanisms that may have played a role during species divergence (reviewed in Hollocher, 1998; Stern, 2011). If these within-species versus between-species comparisons reveal strong similarities, then generally it can be concluded that speciation proceeded through the same general action of evolutionary processes (in terms of type, direction, and strength) normally operating on these traits within species. In contrast, strikingly different patterns of within-species versus between-species comparisons could reveal the operation of a different set of evolutionary constraints operating during divergence of species than what normally occurs within species.

Interestingly, within- and between-species genetic patterns of sterility and inviability do not show similar patterns at all. Not only is the role of epistasis drastically different between the two comparisons, but also the relative frequency of genes that affect sterility versus inviability is completely reversed depending on whether within- or between-species patterns are considered. Additionally, within- and between-species patterns of genetic variation for mate discrimination, secondary sexual characteristics, and general morphological traits are very similar with respect to the additivity of genic effects underlying these traits (see Hollocher, 1998; Orr, 2001; Wolf *et al.*, 2010), yet are very different, especially for morphological traits, with respect to whether species differences represent changes in the coding regions of proteins or changes in the upstream regulatory elements governing gene expression, where interspecific differences have been found to be disproportionately tied to changes in gene regulation. This disjunction between the genetic patterns observed within species versus those observed between species suggests that the evolutionary processes that impact standing genetic variation within populations may not be the same that operate over longer periods of time to produce successful species. In other words, time scale does matter and within-species evolutionary dynamics may be very different from the operation of evolutionary processes that ultimately govern the emergence of different species (Stern, 2011).

New Directions for Studying Speciation

Incorporation of Genealogical Models

The classic approach for determining patterns of genetic change between species has been to cross species and analyze the genetic basis of their differences. Most detailed analyses have focused on the genetic bases of postzygotic and prezygotic reproductive isolation. The emphasis of this approach is on understanding how specific biological properties of species change as they diverge from one another. Therefore, the method involves starting with groups of organisms that show interesting differences in evolutionarily important traits and working down to the genetic level to gain insights into speciation mechanisms that may involve these traits specifically.

A very different approach is to use gene genealogies to trace the evolutionary history of genetic variation without regard to the phenotypic effects (in fact, assuming that the genes being investigated are neutral). Population genetic theory involving gene genealogies was originally developed to study the patterns of genetic variation within and among populations of a species to investigate microevolutionary processes such as gene flow, genetic drift, and natural selection. This theory is now being extended from populations up to the species level to investigate how speciation affects patterns of variation within and between species (e.g., Templeton article in Howard and Berlocher, 1998; Hey, 2006; Knowles and Carstens, 2007). This approach examines the genetic structuring of neutral variation within and between species to evaluate the role such things as random genetic drift, continual gene flow, and different geographical patterns play in speciation.

The two approaches described above supply very different information about speciation. Each approach in isolation has contributed substantially to formulating hypotheses about speciation. Using a combined approach could potentially be even more powerful for investigating important issues in speciation. Future studies of speciation would benefit from

continuing to use traditional genetic methods to identify candidate genes for phenotypic traits that may have been critically important in promoting divergence and then analyzing the genealogical structure of these gene regions in the context of the structuring of genes presumed to be neutral to test specific hypotheses about the role the candidate gene regions may have played during speciation. The evolutionary history of genes that contribute directly to phenotypic differences between species is embedded in the evolutionary history of all genes. A combined approach would allow us to tease apart the population dynamics of genes directly involved in speciation from the dynamics of neutral genes. In other words, the genealogical structuring of neutral genes can serve as a baseline measure of population history (supplying information on such things as the geographical patterns of divergence) to compare to the genealogical structuring of genes thought to contribute more directly to the speciation process itself (which would supply information on selective forces important in divergence). In taking this dual approach to understanding the genetics of speciation, we will be able to move one step closer to inferring the evolutionary process of speciation based on the examination of the genetic patterns of divergence.

Incorporation of Developmental Biology

For a complete understanding of the process of speciation, it is not enough simply to understand in broad terms the number of genes involved in species divergence and the general distribution of their effects. As this article has revealed, this approach has been very helpful so far to give us new insights into genetic mechanisms of speciation. However, ultimately it will be necessary to identify the specific products of genes that have changed during speciation to try to understand how the changes in these gene products have specifically led to the changes in the phenotypes that we observe between species. To achieve understanding at all these different levels, it will be necessary to incorporate developmental biology into speciation theories in order to model how specific genetic changes are translated through the developmental program of an organism to give rise to new species. Although the very roots of evolutionary biology are firmly anchored in the field of development, it is only recently that attempts have been made to meld these two disciplines (e.g., see Carroll, 2005; Stern, 2011). The union of these two fields is important not only from a mechanistic point of view for understanding how genes are translated into phenotypes but also for investigating how the structure of developmental systems themselves can serve as direct conduits for the action of natural selection in generating phenotypic diversity.

Conclusions

The study of speciation is a rich and exciting field. Despite being the object of intense study for the past 150 years, speciation studies continue to capture the imagination of evolutionary biologists and offer a wide range of important questions ripe for investigation. The continued vibrancy of the field comes from its ability to track technological and theoretical advances being made in disciplines such as population genetics and molecular and developmental biology, and to synthesize these distinct approaches to provide fresh perspectives on important ongoing issues in speciation, which due to their innate complexity will never offer simple explanations.

See also: Biodiversity, Evolution and. Phenotype, a Historical Perspective. Population Genetics. Speciation, Process of. Species, Concepts of. Species Diversity, Overview

References

Andersson M (1994) *Sexual Selection*. Princeton, NJ: Princeton University Press.
Arnold ML, Ballerini ES, and Brothers AN (2011) Hybrid fitness, adaptation and evolutionary diversification: Lessons learned from Louisiana Irises. *Heredity* http://dx.doi.org/10.1038/hdy.2011.65.
Bank C, Hermisson J, and Kirkpatrick M (2012) Can reinforcement complete speciation? *Evolution* 66: 229–239.
Baum DA and Shaw KL (1995) Genealogical perspectives on the species problem. In: Hoch PC and Stevenson AG (eds.) *Experimental and Molecular Approaches to Plant Biosystematics*. St. Louis, MO: Missouri Botanical Garden.
Berlocher SH and Feder JL (2002) Sympatric speciation in phytophagous insects: Moving beyond controversy? *Annual Review of Entomology* 47: 773–815.
Boake CRB (2005) Sexual selection and speciation in Hawaiian drosophila. *Behavior Genetics* 35: 297–303.
Bolnick DI and Fitzpatrick BM (2007) Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology Evolution and Systematics* 38: 459–487.
Butlin RK, Galindo J, and Grahame JW (2008) Sympatric, parapatric or allopatric: The most important was to classify speciation? *Philosophical Transactions of the Royal Society B* 363: 2997–3007.
Carroll SB (2005) *Endless Forms Most Beautiful: The New Science of Evo Devo*. New York, NY: W.W. Norton and Co.
Clune J, Misevic D, Ofria C, Lenski RE, Elena SF, and Sanjuan R (2008) Natural selection fails to optimize mutation rates for long-term adaptation on rugged fitness landscapes. *PLoS Computational Biology* 4: e1000187.http://dx.doi.org/10.1371/journal.pcbi.1000187.
Coyne JA and Orr HA (2004) *Speciation*. Sunderland, MA: Sinauer.
Darwin C (1872) *The Origin of Species*, 6th edn. New York: Reprinted in 1976 by Macmillan.
de Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology* 56: 879–886.
Dobzhansky T (1937) *Genetics and the Origin of Species*. New York: Columbia University Press.
Dobzhansky T (1940) Speciation as a stage of evolutionary divergence. *The American Naturalist* 74: 312–321.
Donovan LA, Rosenthal DR, Sanchez-Velenosi M, Rieseberg LH, and Ludwig F (2010) Are hybrid species more fit than ancestral parents species in the current hybrid species habitats? *Journal of Evolutionary Biology* 23: 805–816.
Eberhard WG (1985) *Sexual Selection and Animal Genitalia*. Cambridge, MA: Harvard University Press.
Edwards SV, Kingan SB, Calkins JD, *et al.* (2005) Speciation in birds: Genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences of the United States of America* 102: 6550–6557.
Elena SF and Lenski RE (2003) Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4: 457–469.
Fitzpatrick BM, Fordyce JA, and Gavrilets S (2008) What, if anything, is sympatric speciation? *Journal of Evolutionary Biology* 21: 1452–1459.
Gavrilets S (2004) *Fitness Landscapes and the Origin of Species*. Princeton, NJ: Princeton University Press.
Gavrilets S and Hastings A (1996) Founder effect speciation: A theoretical reassessment. *The American Naturalist* 147: 466–491.
Grant PR (ed.) (1998) *Evolution on Islands*. Oxford, UK: Oxford University Press.

Grant PR and Grant BR (1997) Genetics and the origin of bird species. *Proceedings of the National Academy of Sciences of the United States of America* 94: 7768–7775.
Gross BL and Rieseberg LH (2005) The ecological genetics of homoploid hybrid speciation. *The Journal of Heredity* 96: 241–252.
Haldane JBS (1922) Sex-ratio and unisexual sterility in hybrid animals. *Journal of Genetics* 12: 101–109.
Harrison RG (ed.) (1993) *Hybrid Zones and the Evolutionary Process*. Oxford, UK: Oxford University Press.
Hewitt GM (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology* 10: 537–549.
Hey J (2006) Recent advances in accessing gene flow between diverging populations and species. *Current Opinion in Genetics & Development* 16: 592–596.
Higashi M, Takimoto G, and Yamamura N (1999) Sympatric speciation by sexual selection. *Nature* 402: 523–526.
Hollocher H (1998) Reproductive isolation in Drosophila: How close are we to untangling the genetics of speciation? *Current Opinion in Genetics & Development* 8: 709–714.
Hosken DJ, Martin OY, Wigby S, Chapman T, and Hodgson DJ (2009) Sexual conflict and reproductive isolation in flies. *Biology Letters* 5: 697–699.
Howard DJ and Berlocher SH (eds.) (1998) *Endless Forms: Species and Speciation*. Oxford, UK: Oxford University Press.
Iwasa Y and Pomiankowski A (1995) Continual change in mate preferences. *Nature* 377: 420–422.
Kao KC, Schwartz K, and Sherlock G (2010) A genome-wide analysis reveals no nuclear Dobzhansky–Muller pairs of determinants of speciation between *S. cerevisiae* and *S. paradoxus*, but suggests more complex incompatibilities. *PLoS Genetics* 6: e1001038.http://dx.doi.org/10.1371/journal.pgen.1001038.
Knowles LL and Carstens BC (2007) Estimating a geographically explicit model of population divergence. *Evolution* 61: 477–493.
Lande R (1981) Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences of the United States of America* 78: 3721–3725.
Lessios HA (2011) Speciation genes in free-spawning marine invertebrates. *Integrative and Comparative Biology* 51: 456–465.
Maan ME and Seehausen O (2011) Ecology, sexual selection and speciation. *Ecology Letters* 14: 591–602.
Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology & Evolution* 10: 294–299.
Mayr E (1963) Animal Species and Evolution. Cambridge, MA: Harvard University Press.
Mayr E (1982) *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Harvard University Press.
Melnyk AH and Kassen R (2011) Adaptive landscapes in evolving populations of *Pseudomonas fluorescens*. *Evolution* 65: 3048–3059.
Mustonen V and Lassig M (2009) From fitness landscapes to seascapes: Non-equilibrium dynamics of selection and adaptation. *Trends in Genetics* 25: 111–119.
Nixon KC and Wheeler QD (1990) An amplification of the phylogenetics species concept. *Cladistics* 6: 211–223.
Noor MAF (1995) Speciation driven by natural selection in *Drosophila*. *Nature* 375: 674–675.
Nosil P, Crespi BJ, and Sandoval CP (2003) Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *Philosophical Transactions of the Royal Society B* 270: 1911–1918.
Nosil P and Schluter D (2011) The genes underlying the process of speciation. *Trends in Ecology & Evolution* 26: 160–167.
Orr HA (2001) The genetics of species differences. *Trends in Ecology & Evolution* 16: 343–350.
Otte D and Endler JA (1989) *Speciation and Its Consequences*. Sunderland, MA: Sinauer.
Parker GA and Partridge L (1998) Sexual conflict and speciation, Phil. *Transactions of the Royal Society of London, Series B* 353: 261–274.
Paterson HEH (1985) The recognition concept of species. In: Vrba ES (ed.) *Species and Speciation*. Pretoria: Transvaal Museum.
Phillips PC (2008) Epistasis – the essential role of gene interactions in the structure and evolution of genetics systems. *Nature Reviews Genetics* 9: 855–867.
Prokupek AM, Eyun S-I, Ko L, Moriyama NE, and Harshman LG (2010) Molecular evolutionary analysis of seminal receptacle sperm storage organ genes of *Drosophila melanogaster*. *Journal of Evolutionary Biology* 23: 1386–1398.
Rettelbach A, Hermisson J, Dieckmann U, and Kopp M (2011) Effects of genetic architecture on the evolution of assortative mating under frequency-dependent disruptive selection. *Theoretical Population Biology* 79: 82–96.
Schluter D (2000) *The Ecology of Adaptive Radiations*. Oxford, UK: Oxford University Press.
Servedio MR (2004) The what and why of research on reinforcement. *PLoS Biology* 2: e420.http://dx.doi.org/10.1371/journal.pbio.0020420.
Servedio MR and Noor MAF (2003) The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology and Systematics* 34: 339–364.
Stern DL (2011) *Evolution, Development, & and Predictable Genome*. Greenwood Village, CO: Roberts and Company Publishers.
Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America* 106: 9939–9946.
Wiley EO (1978) The evolutionary species concept reconsidered. *Systematic Zoology* 27: 17–26.
Wilkins JS (2009) *Species: A History of the Idea*. Berkeley, CA: University of California Press.
Wolf JB, Brodie, III ED, and Wade MJ (eds.) (2000) *Epistasis and the Evolutionary Process*. Oxford, UK: Oxford University Press.
Wolf JBW, Lindell J, and Backstrom N (2010) Speciation genetics, current status and evolving approaches. *Philosophical Transactions of the Royal Society B* 365: 1717–1733.
Wright S (1932) The roles of inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetrics*, vol. 2, pp. 356–366.
Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology* 14: 851–865.
Wu C-I and Palopoli MF (1994) Genetics of postmating reproductive isolation in animals. *Annual Review of Genetics* 28: 283–308.
Wu C-I, Johnson NA, and Palopoli MF (1996) Haldane's rule and its legacy: Why are there so many sterile males. *Trends in Ecology & Evolution* 11: 281–284.
Wu C-I and Ting C-T (2004) Genes and speciation. *Nature Reviews Genetics* 5: 114–122.

Species–Area Relationships

EF Conor, San Francisco State University, CA, USA
ED McCoy, University of South Florida, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Relative abundance distribution The frequency distribution depicting the number of species in a community as a function of the number of individuals comprising each species.

Species–area curve A graphical depiction of the dependence of species richness on area.

Species–area model A function used to describe species–area curves.

Species–area relationship The dependence of the number of species in a sample region on the area or size of the region.

Introduction and Underlying Mechanisms

Watson first described the species–area relationship in 1835 by remarking that as the area of a county in England increases by a factor of 10 the number of plant species found in that county increases by a factor of 2 (**Figure 1**). The long history of discourse on species–area relationships has evolved from a focus prior to the 1960s on its empirical utility in determining the optimal sample size and sample number for community description, for determining the minimum area of a community, and in extrapolating predictions of species richness to areas larger than those sampled to a focus in the 1960s and 1970s on the mechanisms underlying the species–area relationships, in finding the best mathematical model for species–area relationships, and in explaining species–area relationships in the context of the equilibrium theory of island biogeography (Connor and McCoy, 1979). Since 1980, the primary focus of discussion involving species–area relationships has been their use and application in conservation biology to determine the optimal design of nature reserves and to project the expected loss of species richness from a region undergoing specified levels of area reduction (habitat loss). We review the discourse on species–area relationships beginning in the 1960s and focus our synthesis on the application of species–area relationships in conservation biology.

Three biological mechanisms have been proposed to account for the species–area relationships: (1) the habitat diversity hypothesis, (2) the area per se hypothesis, and (3) the passive sampling hypothesis.

The Habitat Diversity Hypothesis

The habitat diversity hypothesis (Williams, 1964) proposes that the increase in species richness in large areas relative to small areas arises because large areas have a greater variety of habitats than small areas. This greater variety of habitats permits species that are only found in specific habitats to occur in large areas and permits species that require multiple habitats to persist in large areas, resulting in higher species richness in large areas than in small areas and the existence of species–area relationships. The habitat diversity hypothesis views area as affecting species richness indirectly because of its association with habitat diversity rather than any direct effect of area on the ability of species to colonize or persist in larger areas. Studies that demonstrate positive correlations between the number of species and the number of habitats on an island have been viewed as supporting the operation of the habitat diversity hypothesis.

The Area Per Se Hypothesis

The area per se hypothesis (Simberloff, 1976) is based on the assumptions that the abundance of each species in a sample region varies as an increasing function of that region's area and that the probability of each species going stochastically extinct there is a decreasing function of abundance, and therefore of the area. Given these assumptions, large areas would have more species than small areas because more species would persist (not go locally extinct) in large areas. The area per se hypothesis suggests that even in a group of patches consisting of a single type of habitat, one would observe a species–area relationship. Correlations between species richness and area in studies that purport to examine a single habitat type and experimental studies that examine the

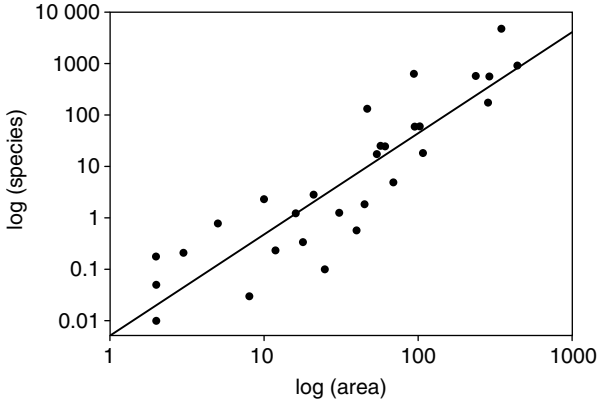


Figure 1 Species–area curve for the vascular plants of the Galapagos Archipelago. The logarithm of species richness is commonly plotted as a function of the logarithm of area, but species richness may also be plotted as a function of area or log(area).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00132-5

consequences of reducing patch area are viewed as evidence supporting the operation of the area per se hypothesis.

The Passive Sampling Hypothesis

The passive sampling hypothesis (Connor and McCoy, 1979) conjectures that larger areas are more likely to receive more colonists than smaller areas and that these colonists are likely to represent a wider array of species than the pool of colonists arriving on small areas. Therefore, purely as a result of the higher abundance of colonists expected for large areas and independent of any increase in habitat diversity or reduction in extinction probabilities, one would expect more species to arrive on large areas, leading to species–area relationships. Several authors have examined the number of colonists arriving at their study sites and concluded that passive sampling accounts partially for the species richness of invertebrates on intertidal boulders, fish in stream pools and riffles, and birds on forest habitat islands. Each of these studies shows that for habitats that are colonized seasonally, those patches receiving more colonists have higher species richness.

Two other area–dependent factors that affect species richness have recently entered discussions concerning species–area relationships, particularly for habitat patches: the resource concentration hypothesis and edge effects.

The Resource Concentration Hypothesis

The resource concentration hypothesis attempts to explain the phenomenon that habitat patches with large amounts of resources (e.g., monocultures, areas of high plant density, or large patches) have higher densities of insects. Therefore, the resource concentration hypothesis conjectures that population density should be positively correlated with patch area. In 1973, R.B. Root conjectured that the higher density of animals in larger patches might be solely a consequence of movement behavior (the movement hypothesis); herbivores are more likely to find, and remain in, large, monospecific stands of their host plant than in small or heterogeneous patches. If many species have higher population densities in large patches because of the resource concentration hypothesis, then extinction probabilities should be even lower than expected from the area per se hypothesis, and hence resource concentration could contribute to observed species–area relationships. E.F. Connor and co-workers have recently reviewed the existing literature on the relationships between animal population density and patch area and found that for insects and birds positive area–density relationships are common, but not so for mammals.

Edge Effects

Edge effects, or habitat edge dependent changes in abundance or risk of mortality, have been reported for species in a variety of taxa in habitat patches. Given that the proportion of a patch that occurs within any fixed distance from its edge is inversely related to the area, edge effects could lead to species–area relationships even within a single habitat type. Such an edge effect on species richness would be mediated by a reduction in the abundance of a species on small patches because of a larger amount of "edge habitat", leading to higher probabilities of local extinction. Therefore, part of the dependence of species richness on area that has previously been attributed to area per se may actually be caused by edge effects.

Multiple Causes of Species–Area Relationships

Habitat diversity, area per se, passive sampling, edge effects, and resource concentration are not mutually exclusive mechanisms and may operate individually or in combination to cause species–area relationships (Connor and McCoy, 1979). Experiments on invertebrate colonization on artificial substrates of varying size and habitat diversity clearly demonstrate the joint contribution of habitat diversity and either the area per se or passive sampling hypothesis, or both, to the generation of species–area relationships. Studies that subsampled islands using a constant-sized quadrat on each island in an attempt to eliminate the effect of habitat diversity found strong effects of habitat diversity, but species richness remained an increasing function of island area even for the data derived from constant-sized quadrats. Using path analysis, D.D. Kohn and D.M. Walsh concluded that area had both a direct effect on species richness (area per se and/or passive sampling) and an indirect effect mediated by the effect of area on habitat diversity.

The evidence that would allow one to clearly partition the causes of a specific species–area relationship among these underlying mechanisms is exacting. Habitat diversity is difficult to measure, and edge effects might make it impossible to separate the effects of increasing area from the effects of area on habitat diversity. The area per se, resource concentration, and passive sampling hypotheses act by affecting the abundances of species on an island. Area per se and resource concentration do so by reducing extinction probabilities via allowing larger populations (or more dense populations) to persist on large islands, and passive sampling does so by proposing that greater numbers of colonists arrive on large islands. To demonstrate an effect of area per se or resource concentration, one must ultimately show that species on average have larger population sizes (or in the case of resource concentration, higher population densities) and lower extinction probabilities on large islands or habitat patches. To demonstrate an effect of passive sampling, one must show that the arrival rate of colonists on large islands or habitat patches is greater than for the small patches and that the colonizing individuals arriving on large patches comprise a greater number of species than those arriving on small patches. To demonstrate that edge effects contribute to species–area relationships would require evidence that species absences from small patches could be uniquely attributed to edge effects and not to the alternative mechanisms discussed above.

Sampling and Statistical Practice in Description of Species–Area Relationships

Sampling Practice

Two main sampling schemes have been used to generate data on the relationship between species richness and area. The most widely used approach has been to sample physically separated areas such as islands or habitat patches or to sample

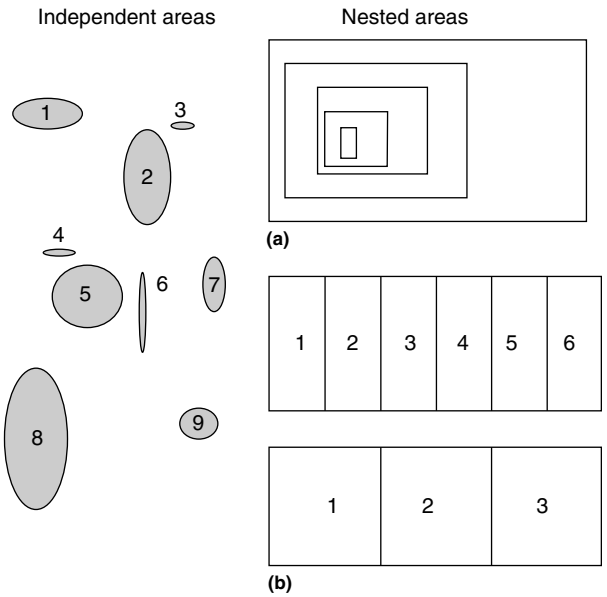


Figure 2 Independent and nested sampling designs used in species–area studies. Independent areas are physically separate. Nested areas may be serially self-contained as in A or abutting as in B. If abutting areas are used, then the sampled region is partitioned into a series of equal-sized, adjacent areas, and this process is repeated for the range of different sized areas. Only two areas are depicted in B, one with six replicates and one with three.

adjacent or abutting areas of continuous habitat as independent, nonoverlapping replicates (**Figure 2**). For physically separated areas, sampling from the observed natural range of areas generates a range of sample areas. For continuous habitat, the region is divided arbitrarily into a series of nonoverlapping subregions to generate a range of sample areas. The alternative approach, which has been used widely by plant ecologists, is to enumerate the locations of species within a larger region and generate a sample of areas by subsampling a range of areas within the larger region. These subsamples may or may not overlap, depending on the choice of the researcher (**Figure 2**).

Independent Areas

Physically independent sample areas have been widely used because species–area data could be readily generated for natural geographical units, such as islands, by combining published monographic species lists with published data on island areas. The majority of species–area curves were published in the 1960s and 1970s by the scientists gleaning existing data from the monographic literature and combining data on species richness with published estimates of area. Species–area curves based on gleaned data could be produced relatively quickly from a library. Another advantage of using physically independent sample areas is that they are arguably statistically independent. The statistical independence of separate, nonoverlapping areas allows use of the methods of statistical inference and hypothesis testing associated with ordinary least-squares (OLS) regression. The application of OLS regression permits the estimation of the parameters (and their standard errors) for any of the proposed functional forms of the species–area relationship that could be made sufficiently linear by transformation of the data and an assessment of the statistical significance of the fit of a particular model. Furthermore, the use of statistically independent areas allowed the rigorous comparison of parameter estimates between studies via analysis of covariance or other simpler techniques.

Nested Areas

Plant ecologists have generated data for many years on the relationship between species richness and a plot or quadrat area for a variety of plant communities. Much of this work was done to determine the optimal plot size to use when describing plant communities. Presumably because of the saving in sampling effort, plant ecologists used serially self-contained or nested quadrats rather than physically independent quadrats to determine optimal quadrat size. However, nested quadrats are not statistically independent so it will be inappropriate to use OLS regression to fit a linear model to such data and assess the fit of the model or estimate its parameters. Therefore, use of nested quadrats to examine species–area relationships in plant communities was largely discontinued by the end of the 1970s. Recent research programs aimed at understanding the structure and dynamics of tropical forests have once again led to the studies of species–area relationships using nonindependent data. However, Condit et al. (1996) partition their largest quadrat into as many equal-sized quadrats as possible and use the mean for all nonoverlapping quadrats to estimate species richness for that quadrat size. They repeat this process of partitioning their large quadrat (50 ha) into the maximum possible number of nonoverlapping subquadrats for a variety of quadrat sizes. Condit et al. (1996) then plot the mean and standard error of species richness for each quadrat size as a function of quadrat area to produce a species–area curve. This approach leads to appropriate estimates of species richness and its standard error for each individual quadrat size, since within a size category the quadrats used to estimate the mean and standard error of species richness are independent. However, since the same sampled area is used to estimate species richness for each quadrat size, these estimates are not statistically independent and Condit et al. (1996) appropriately refrain from using regression techniques that require independence to fit statistical models to species–area data.

Statistical Practice

OLS regression has been used to fit models to species–area data. By applying OLS regression, estimates of the model's parameters and their associated standard errors can be obtained readily. If one makes the additional assumptions that the model's errors are independent, homoscedastic, and normally distributed and that species richness or the logarithm of species richness is a linear function of area or its logarithm, then rigorous statistical inferences about the parameters may be made. In recent years, nonlinear regression procedures have become widely available so that inherently curvilinear species–area models may be directly fit to the data. Nevertheless, OLS regression continues to be used widely.

P.J. Vincent, J.M. Haworth, and M.R. Williams point out that the assumptions of normality and homogeneity of error variances are unlikely to be met by species–area data. First of all, species richness is unlikely to be normally or log-normally distributed and probably should be treated as a discrete, rather than a continuous, variable. Second, the variance in species richness is likely to be a function of the mean species richness. These properties arise partially because on very small areas both the mean and variance of species richness must be nearly zero. P.J. Vincent and J.M. Haworth suggest analyzing species–area data using a generalized linear model and treating species richness as a Poisson-distributed variable. Williams also suggests analyzing species–area data using a generalized linear model but recommends that species richness be treated as a binomially distributed variable (*see discussion in* Extreme Value Model and Random Placement). In both the cases, parameters and their standard errors can be estimated and rigorous statistical inferences can be made. Since 2000, many authors have applied generalized linear models to the analysis of species–area relationships. Williams et al. (2009) summarize this literature and make sound recommendations for model fitting and model selection.

Functional Form of the Species–Area Relationship

When species richness is plotted as a function of area, the resulting plot is a curve that may be linear, concave downward, concave upward, or sigmoid (**Figure 3**). The shape of species–area curves appears to be a function of the particular range of areas studied, and this may, in part, explain the variety of transformations used to linearize species–area curves.

Historically, much of the discussion of species–area relationships focused on the specific functional form of the relationship between species richness and area. In 1921, Arrhenius proposed that species–area relationships followed a power function (**Figure 3(b)**): $S = cA^z$ which, for statistical convenience, has commonly been approximated by the log–log transformation: $\log S = \log c + z \log A$ with c and z as constants. The double-logarithmic transformation linearizes the power-function model, so that species–area curves are linear on a double-log plot (**Figure 1**). In 1922, Gleason championed an exponential model of the species–area relationship (**Figure 3(c)**): $S = \log c + z \log A$ because he observed the power-function model to predict impossibly large numbers of species for large areas. The exponential model produces species–area curves that are linear on a semilogarithmic plot. Both of these models were supported because they seemed to fit specific data sets using independent areas reasonably well and because it was argued that they could be derived as a consequence of assuming that the distribution of individuals among species (the relative abundance distribution) was either log-normal or log-series, respectively. Other ecologists reported that when a wide range of areas were sampled, neither the power-function model nor the exponential model fit

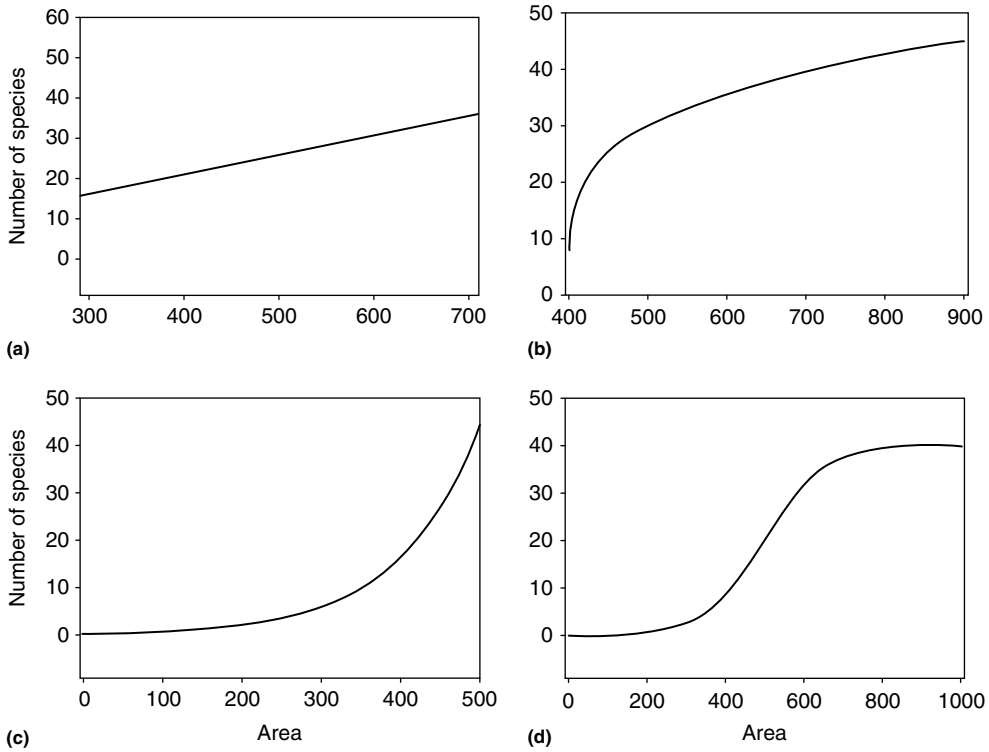


Figure 3 Shapes of species–area curves encountered in empirical studies: (a) linear species–area curve; (b) concave-downward species–area curve (linearized by the log–log transformation); (c) concave-upward species–area curve (linearized by a semilogarithmic transformation, area log-transformed); (d) sigmoid species–area curve (usually not transformed). Linear curves are encountered at intermediate spatial scales, concave-downward curves at larger spatial scales, concave-upward curves at small spatial scales, and sigmoid curves when a wide range of spatial scales is studied.

their data since species–area curves appeared sigmoid in shape (**Figure 3(d)**).

Connor and McCoy (1979) examined the fit of the power-function model, the exponential model, and two other models to data from 100 species–area studies. They found that although the power-function model often fit the data well, it was not found to be the best fit model substantially more often than other models. Based on this analysis, the observation that the log–log transformation can linearize a wide range of curves, and because of its widespread use, Connor and McCoy (1979) recommended the continued use of the power-function model of species–area relationships.

Recent discussions of the functional form of the species–area relationship have continued to promote the use of the power-function model but have continued to justify its use because of its hypothesized connection to log-normal relative abundance distributions (Rosenzweig, 1995). Given a common log-normal relative abundance distribution from which the abundances of species on all sites (islands or habitat patches) arise, it is possible to derive species–area data that are reasonably well fit by a power-function model. However, simply because a particular set of species–area data is reasonably well fit by a power-function model does not imply that this fit derives as a consequence of sampling from a common log-normal relative abundance distribution (Hanski and Gyllenberg, 1997). Many island archipelagos are colonized from multiple source regions rather than from a single source pool, and individual islands within such archipelagos may receive different proportions of their colonists from each source. However, habitat patches may be more accurately viewed as units whose species composition is determined by sampling from a source pool with a single relative abundance distribution.

Continued debate about the functional form of the species–area relationship assumes that underlying the variability we observe in empirically estimated species–area curves lies a single true function that will characterize this relationship. However, most data on species–area relationships using independent areas have small sample sizes, cover a narrow range of areas, and may be confounded by area-dependent sampling effort, among other problems. Given the quality of the data available, it will be difficult to discern the functional form of the species–area relationship, even if a single form exists. Furthermore, it is possible that differences among taxa or spatial scales could generate species–area relationships of different functional forms.

Self-Similarity and the Power-Function Model of Species–Area Relationships

Harte and co-workers have recently proposed an alternative derivation of the power-function model of the species–area relationship. This derivation comes from examining nested areas and making the assumption that successive partitions of a continuous area have the fractal property of self-similarity. That is, the distribution of species in successive partitions is independent of the spatial scale. Beginning with a large continuous area A_0 with S_0 species, bisections of A_0 such that the rectangles comprising the two halves of A_0 have an area $A_i = A_0/2^i$ each will have on average S_i species, where i indicates the ith bisection of A_0. The assumption of self-similarity requires that a species known to be in A_i will have probability a of being found in at least a specific one of the two rectangles of area A_{i+1} created by bisection, and therefore the fraction of species found in A_i that are found in a specific one of the A_{i+1} rectangles equals the same constant a for all i bisections. The constancy of a for all i bisections follows from the assumption of self-similarity since a, the probability of occurrence in a half-patch under bisection, must be independent of spatial scale. Harte et al. (1999) show that the assumption of self-similarity leads directly to the power-function model of the species–area relationship: $S_i = cA_i^z$, with the self-similarity parameter $a = 2^{-z}$. Furthermore, since under successive bisections of an initial continuous area the value of a must lie between 0.5 and 1, the relationship between a and z dictates that the slope of the power-function model, z, must lie between 0 and 1. Hence, the power-function model of the species–area relationship is derivable without recourse to any assumptions about the underlying relative abundance distribution of species.

Harte and colleagues have extended the assumption of self-similarity to generate an expectation for the relationship between the number of endemic species and area and the expected spatial turnover between two patches isolated by a known distance, d. These extensions of the assumption of self-similarity provide a basis for estimating the expected loss of species from areas undergoing habitat reduction, for calculating z from spatially separated sites, and for estimating species richness at spatial scales larger than commonly possible in previous studies of species–area relationships. However, as Harte and colleagues point out, the assumption of self-similarity is not likely to hold for a wide range of spatial scales, so care must be taken to test the assumption of self-similarity for the taxa and spatial scales of inference. Although Harte and colleagues have published a few examples where the assumption of self-similarity seems reasonable over a specific range of spatial scales, Plotkin and colleagues show that tropical forest trees are not self-similar in distribution over spatial scales from 1 to 10^5 m^2.

Extreme Value Model and Random Placement

An alternative model of species–area relationships based on random placement was first proposed by Coleman in 1981 (Coleman et al., 1982) and extended by Williams (1995). The random placement model of species–area relationships derives the expected number of species on a site, $\bar{s}(\alpha_k)$, of area A_k as a consequence of placing the n_i individuals of the ith species on sites independently and at random with probability: $\alpha_k = A_k/A_t$, where A_k is the area of the kth site, A_t, is the combined area of all sites, and S is the total number of species among all the sites:

$$\bar{s}(\alpha_k) = S - \sum_{i=1}^{s} (1 - \alpha_k)^{n_i}$$

As presented by Coleman et al. (1982), data on the total abundance of each species combined among all sites were required to estimate the expected species–area curve under the

random placement model. Because the random placement model required data that ecologists seldom have, censuses of the abundances of each species at each site, and because this model does not yield fitted parameters comparable to other species–area models, few authors attempted to fit the random placement model to their data.

Williams (1995) extended and adapted the random placement model to be approximated by an extreme value function which permits model fitting within the context of the generalized linear model and requires data only on species richness and area, not on the abundances of the individual species. The extreme value function model of species richness, $\bar{s}$, in log A is then $\bar{s} = P[1 - \exp(\gamma \log A + \log d))]]$ with P being the number of species in the species pool, and γ and log d the slope and intercept of the model, respectively. Williams (1995) outlines a method for estimating P if the species composition of the biota is unknown but suggests that the best estimate of P is the total number of species found. The extreme value function model has been fitted to a limited number of data sets but appears sigmoid when species richness is plotted as a function of log(area). The random placement or extreme value function model of species–area relationships is appealing because it is derived from a hypothesis of independence within and between species. However, as Williams (1995) points out, discriminating between the power function and the extreme value function models with most existing data sets derived from sampling independent areas will be difficult.

Interpretation of the Parameters of Species–Area Models

The widespread use of the power-function model of the species–area relationship coupled with specific numerical expectations for its slope parameter, proposed by Preston in 1960, led to numerous attempts to infer biological significance to the values of this parameter. The slope parameter measures the rate at which species are added as area increases, and the intercept parameter has been considered a function of taxon-specific attributes and environmental variation. Only a limited attempt has been made to offer biological interpretations of the intercept parameter, partly because in the power-function model the absolute value of c depends on the units in which area is measured (Connor and McCoy, 1979; Rosenzweig, 1995). The parameters of other models of the species–area relationship have generally been treated as fitted statistical constants, with no attempt to interpret them biologically.

The slope of the power-function model of the species–area relationship and patterns of variation in this parameter have been subjected to considerable analysis and interpretation. Connor and McCoy critiqued many of these interpretations in their 1979 review. We will briefly examine a few of these interpretations and touch on those which are proposed very recently.

Canonical Slope Values

Preston proposed that isolated islands in equilibrium that sample colonists from a common log-normal relative abundance distribution with parameter $\gamma = 1$ (Preston's canonical hypothesis) will have a slope of 0.262 in the power-function model. He subsequently broadened the range of slope values that he expected from isolates to values between 0.17 and 0.33. May showed in 1975 that using a wide range of log normal parameter values leads to power-function slope values in the 0.15–0.39 range. Many authors have generated slopes from the power-function model of the species–area relationship and interpreted values in the range of 0.2–0.4 to be consistent with Preston's idea that species–area curves arise because the species richness of islands results from sampling from an underlying log-normal relative abundance distribution. However, Connor and McCoy (1979) challenged Preston's idea of a canonical range of slope values by showing that slopes in the 0.2–0.4 range are expected purely as a consequence of the tendency to publish studies that show a high correlation between species richness and area and because the variance in species richness will always be less than the variance in area. Connor and McCoy (1979) concluded that slope values from the power-function model of species–area relationship in the 0.2–0.4 range could not be used as evidence for the existence of an underlying log-normal relative abundance distribution.

Island–Mainland Differences in Slope Values

Preston extended his idea that islands or isolates should have power-function slopes in the range of 0.2–0.4 to project that nonisolated or mainland areas should have lower slope values than isolated areas. His rationale was derived from the belief that nonisolated areas would sample from a truncated relative abundance distribution with a higher ratio of species to individuals. In their monograph *The Theory of Island Biogeography*, MacArthur and Wilson (1967) modified Preston's idea, suggesting a specific range to be expected for slopes derived from nonisolated areas, 0.12–0.19, and explaining the lower slopes as arising because of the transient hypothesis. The transient hypothesis suggests that more transient individuals will be encountered in small, nonisolated areas than in small, isolated areas, which, in turn, will lead to more species being encountered on small, nonisolated areas than on small, isolated areas. The greater number of species found on small, nonisolated areas would depress the slope of the species–area curve for mainland areas relative to that expected for islands. The available evidence, although limited, is consistent with the idea that power-function slopes for mainland areas are lower than those for islands.

Hanski and Gyllenberg (1997) develop dynamical models of species incidence that generate predictions about the slopes of species–area curves as a function of the moments of the relative abundance distribution and the ratio of species' colonization and extinction rates. Hanski and Gyllenberg (1997) develop two models, one in which sites are colonized from an external source, their island–mainland model, and one in which the sources of colonists are internal to the system of sites, their metapopulation model. They find that species–area slopes are lower for their metapopulation model and claim that such a model is analogous to mainland areas, whereas their island–mainland model is analogous to truly insular

situations. Hanski and Gyllenberg (1997) suggest that their models imply that the observation of lower slope values on mainland sites is a result of metapopulation dynamics rather than the transient hypothesis.

The Effect of Isolation on Slope Values

MacArthur and Wilson extended their transient hypothesis to predict that the slopes of species–area curves should be lower for distant archipelagos of islands than for islands located close to mainland areas. However, Schoener illustrated with species–area curves for birds in 23 archipelagos that exactly the opposite pattern occurs. Schoener's result suggests that the slope of species–area curves depends on the size of the source pool of colonizing species, which would be smaller for distant than near island groups. In other words, the low slopes for species–area curves observed for isolated archipelagos tell us no more than that isolated biotas are depauperate. Hanski and Gyllenberg (1997) suggest that Schoener's observation of lower slope values for isolated archipelagos arises because isolated archipelagos behave according to their metapopulation model, not as an island–mainland system.

Other Interpretations of Slope Values

Many other efforts have been made to explain variation in the slope of the power-function model of the species–area relationship. These efforts include (1) attempts to equate the slope of the species–area curve with β, or between-habitat diversity, and the intercept with α or within-habitat diversity, (2) attempts to predict patterns in the latitudinal dependence of the species–area relationship (both slope and intercept), and (3) attempts to identify taxonomic and trophic group differences in species–area curves, among others. Connor and McCoy (1979) critiqued many of the attempts to interpret the parameters of the power-function model of the species–area relationship and recommended that these parameters be viewed as fitted constants, with no specific biological interpretation.

Conclusions

Species–area relationships represent a pattern expected in nature that may arise from the colonization and development of quasi-independent biotas on islands or habitat patches (island–colonization model) or from the loss, reduction, and fragmentation of a previously continuous or widespread biota into remnant habitat patches (mainland-vicariance model). Historically, inquiry into the mechanisms underlying the species–area relationship and its functional form has been biased toward the island-colonization model. However, models based on self-similarity more closely represent the mainland–vicariance model and provide a basis to unify species–area relationships with other patterns in the geographical distribution of species (e.g., distribution and abundance relationships, compositional similarity among sites, and relative abundance distributions). Continued empirical study of the species–area relationships and inquiry into the functional forms expected under specific biological and sampling models will continue to improve our understanding of spatial patterns of species richness.

Use of Species–Area Curves in Conservation Biology

What can an understanding of the species–area relationship contribute to the preservation of biodiversity? The question has interested ecologists for more than 30 years since the pioneering work of N.W. Moore, E. Maarel, and others and has led, almost from the beginning, to a bewildering confusion of missteps and dead ends (see Shafer (1990) for an overview). Implicit in this question are the important assumptions that species richness is the primary object of preservation and that area is the primary influence on species richness. Here, we shall explore two separate, but similar, aspects of the question. We shall see what ecologists have been able to conclude about the loss of species accompanying area reduction and what they have been able to conclude about the best way to slow the loss.

Loss of Species from Area Reduction

As we have illustrated, the species–area relationship is an extraordinarily common pattern in nature, and it has been documented numerous times. It would seem, therefore, that calculating the loss of species accompanying a certain amount of area reduction would be a rather straightforward exercise. All that one would need to know to perform the calculation would be the original number of species (S_{original}), the amount of area reduction ($A_{\text{reduced}}/A_{\text{original}}$), and the slope of the species–area relationship (z). For example, if one assumes a power-function model of the species–area relationship $S = cA^z$, then $S_{\text{original}} = c(A_{\text{original}})^z$, $S_{\text{reduced}} = c(A_{\text{reduced}})^z$, and, therefore, $S_{\text{reduced}}/S_{\text{original}} = (A_{\text{reduced}}/A_{\text{original}})^z$ (**Figure 4**). Although calculations of this sort are reasonably common, they require at least five assumptions. First area reduction completely eliminates species that were originally present (but some species may survive on relict fragments or disturbed lands). Second all of the species that were originally present were distributed homogeneously (but most species are unlikely to be distributed in this manner). Third an appropriate model has been selected to describe the species–area relationship. Fourth the slope of the species–area relationship is accurate and a constant (but it is not likely to be, for a number of reasons). Fifth the loss of species is a direct consequence only of area reduction (but it is not likely to be, for a number of reasons). The last two assumptions have received the most attention from ecologists. We shall discuss both of them further. Choice of the slope of the species–area relationship is critical. Clearly, if one knew the original number of species and area and the reduced number of species and area, then the slope would be determined precisely. The point of the calculations presented in this section, however, is to estimate potential species loss, before it happens, so the slope of the species–area relationship also must be estimated, from theory and empirical study. A great deal of uncertainty accompanies such estimation. First one must decide which model of species loss is appropriate. For example, if the relationship between

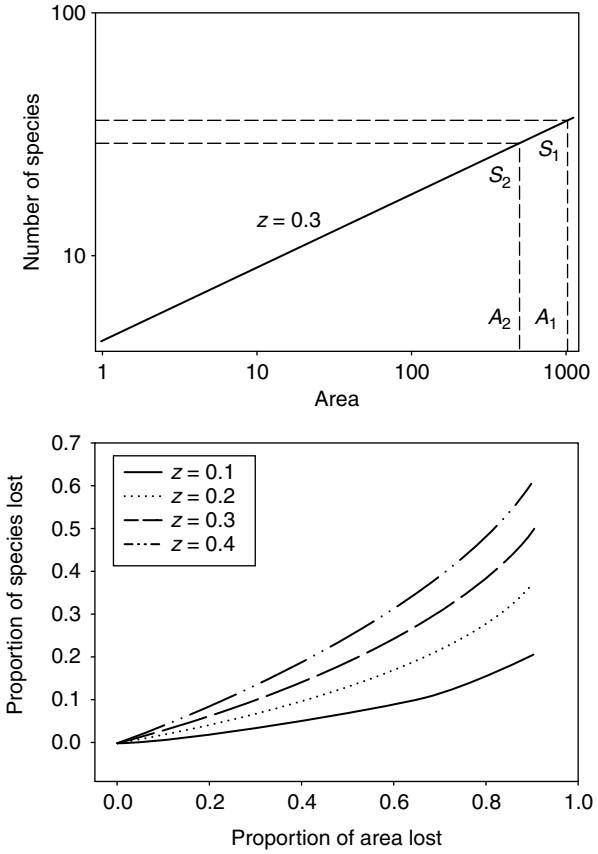


Figure 4 Loss of species richness under specified levels of habitat loss expected using the power-function model of the species–area curve. (a) Illustrates the single-large-or-several-small controversy calculation of the number of species expected for a single large reserve (SLR) of 1000 units of area (34.5 species) vs. the number of species expected for a reserve half that size (500 area units), 29 species. If the species lists from the two half-sized reserves do not overlap extensively, then the two half reserves should contain more species than the SLR. For this example, $S = 4.5A^{0.3}$. (b) The expected proportion of species lost when area is reduced by a fixed proportion as calculated using the power-function model of the species–area relationship for various values of z and $S = 4.5A^z$.

species and area were assumed to be linear (**Figure 3(a)**), then a much different rate of loss of species would be predicted to occur than if the relationship were assumed to be a power function (**Figure 3(b)**). Second one must assume that the slope of the species–area curve is constant between the spatial scales over which species loss is to be estimated. A non-constant slope value would suggest that an appropriate species–area model has not been completely specified and could lead to the substantial misestimation of species loss. Third one must decide if the reduced area better represents an isolate (a true island) or simply a subsample of the original area. Typically, slopes of species–area relationships are greater in the first case (approximately 0.20–0.40, on a log–log scale) than in the second (approximately 0.12–0.19, on a log–log scale). Further, it may be that the reduced area could come to function less like a subsample and more like a true island over time, so that the range of potential slopes could expand dramatically over time. Fourth even if one managed to choose the more accurate of the two ranges of slope values, an impressive amount of variation would still remain. For example, with the equations presented earlier in this section, if area is reduced by 80% and $z = 0.20$, then species loss is estimated at 27%, whereas if area is reduced by the same percentage but $z = 0.30$, then species loss is estimated at 38%. Fifth even if one knew precisely the correct slope of the species–area relationship in one location at a time, Connor and McCoy (1979) and others have shown that the slope is not likely to be amenable to translation across locations, latitudes, or time.

One empirical means for focusing the choice of slope values is the body of literature documenting decline in species richness accompanying area reduction. J.M. Diamond termed this process "relaxation." A number of studies, for example, have examined the relaxation in the species of mammals that may have occurred on nature reserves, both in North America and in Africa. The results of these studies clearly illustrate the difficulty in choosing a slope of the species–area relationship. For African savanna reserves, R. East has shown that estimates of species loss accompanying a 100-fold reduction in area vary from about 40% to more than 90%, depending on the required minimum viable population size (i.e., the population size that ensures persistence with a certain probability for a certain length of time). W.D. Newmark has shown that similar high variability exists for western North American national parks.

At least three potential reasons underlie the poor predictability generally seen in studies of relaxation. The first two reasons, the presence of statistical or sampling errors and the lack of sufficient time since area reduction took place to allow full relaxation, could be addressed by careful experimentation. The protocol is to document the species richness of a relatively large area, reduce the area, and then document the species richness of the reduced area. Unfortunately, most of these kinds of experiments have been carried out at such a small scale that their general value at the scale of typical nature conservation efforts is questionable. Slope values derived from these experiments are, therefore, not likely to be of much practical value. A few reasonably large scale experiments in area reduction have been undertaken, most notably The Minimum Critical Size of Ecosystem project of D.H. Janzen, T.E. Lovejoy, R.O. Bierregaard, and others. These relatively large scale experiments suggest that the species–area relationship alone does not predict species loss particularly well. Slope values derived from these experiments, therefore, also are not likely to be of much practical value. The third reason, lack of the autecological (i.e., species-specific ecological) information needed to infer that the loss of species is attributable directly to relaxation, cannot be addressed as readily by experimentation.

Recall that one of the assumptions necessary to predict loss of species accompanying area reduction was that the loss of species is a direct consequence only of area reduction. Loss of species actually may occur only as an indirect consequence of area, because of some factor that happens to be correlated with area reduction, such as level of disturbance, complexity of habitat structure, or – probably most importantly – fragmentation. It seems clear that in virtually all real-world cases, area reduction is likely to be accompanied by fragmentation. That

is, one relatively large area is not likely to be reduced to a single relatively small area, but rather to an archipelago of relatively small fragments. For example, Simberloff (1992) has shown that fragmentation effects, as well as size-dependent habitat changes, are at least as important as area reduction in predicting species loss in the relatively large scale experiments mentioned previously. It has been known for quite some time, at least since the pioneering work of N.W. Moore and others, that fragmentation can have adverse effects on organisms, independent of the effects of area reduction. The species–area relationship predicts that each of the fragments will support fewer species than the original area, but it alone cannot predict the cumulative number of species in the entire archipelago of fragments. As D. Simberloff and L.G. Abele have clearly shown, we also must know the degree of overlap in the species compositions of the fragments. If, for example, the compositions of the fragments were identical, then the archipelago would support fewer species than the original area. If, however, the species compositions of the fragments did not overlap at all, then the archipelago could support as many – or, potentially, even more – species than the original area. Assuming, realistically, that compositions overlap to an intermediate degree, calculating the loss of species by the method we have just illustrated could yield either an underestimate, if many species were confined to but a small subset of fragments, or an overestimate, if they were not. Fragmentation, therefore, adds more uncertainty – the degree of overlap in species compositions – to any prediction of the loss of species accompanying area reduction. We could predict the degree of overlap in the species compositions of fragments better if we had detailed information on the habitat requirements and minimum viable population sizes of individual species and on spatial variation in species richness, but as D. Simberloff and others have shown, such information is notoriously difficult to obtain and, therefore, is in very short supply.

Recently, a method has been suggested by Harte and coworkers whereby the slope value for species–area relationships among fragments might be obtained more easily. This method is based on the idea of self-similarity that we discussed in Self-similarity and the Power-Function Model of the Species-area Relationship. Essentially, the method requires only measurement of the degree to which species lists in separated, small, censused patches overlap as a function of the distance between patches and the area of the patches. Self-similarity, then, allows translation of the resulting slope value across scales, so that censusing of prohibitively large areas or obtaining extensive autecological information becomes, in theory, unnecessary. A.P. Kinzig and J. Harte have also developed a procedure for estimating species loss from habitat reduction that derives from the assumption of self-similarity and from examining the relationship between the number of endemic species and area, an endemics–area relationship (EAR). In the examples examined thus far, the estimated species loss from area reduction is less using Kinzig and Harte's EAR approach than from the traditional approach based on the species–area relationship outlined above. Whether these approaches for estimating species–area slopes and calculating species loss will lead to yet another dead end remains to be determined but will depend on how reasonable it is to assume self-similarity over a range of spatial scales.

Slowing the Loss of Species

In the previous section Loss of Species from Area Reduction, we explored what the species–area relationship was able to tell us about species loss accompanying area reduction. In this section, we shall explore what the relationship can tell us about slowing that loss. In doing so, we shall treat area reduction with concomitant loss of species as an accomplished fact and concern ourselves solely with maximizing the remaining species richness. We shall, also, focus on maximizing species richness by the judicious choice, or design, of nature reserves.

The serious study of the relationship between the species–area relationship and the design of nature reserves began in the mid-1970s. At that time, several papers by J. Terborgh, J.M. Diamond, E.O. Wilson, E. Willis, and others came out more or less simultaneously highlighting the overarching importance of largeness of nature reserves. Some of these papers further maintained that the species–area relationship indicates that nature reserves should be subdivided as little as possible, that is, a single intact reserve of a certain area will support more species than a series of individually smaller reserves that are cumulative of the same area. Diamond, for example, said "many species that would have a good chance of surviving in a single large reserve would have their survival chances reduced if the same area were apportioned among several smaller reserves." The basic problem with these ostensibly reasonable proposals – call them largeness and singularity – is that they may be too simplistic for real-world conservation challenges. They may be too simplistic because of the substantial set of assumptions that accompanies them. We shall address perhaps the most basic of these assumptions in detail. This assumption is that the importance of largeness and singularity follows directly from the species–area relationship.

Most of the early papers based their arguments for largeness and singularity on MacArthur and Wilson's theory of island biogeography. This theory of island biogeography takes a dynamic view of the maintenance of species richness on islands, as a balance between colonization and extinction. The resulting species richness, therefore, is considered an equilibrium number of species. If relatively large areas support relatively large populations, then relatively large areas also should have relatively low extinction rates and, consequently, support relatively high equilibrium numbers of species. Accordingly, any subdivision of a relatively large area essentially creates smaller areas, each with higher extinction rates, so that the equilibrium number of species on each of the areas necessarily will fall (relaxation). The early papers also made a variety of proposals to enhance colonization, reflecting the potential importance of clustering and corridors and of shape, but we shall not address these proposals.

Almost immediately, the arguments for the overarching importance of largeness and singularity were challenged. Simberloff and Abele showed that the species–area relationship did not unambiguously favor a single intact reserve over an archipelago of smaller reserves. For example, if one assumes a power-function model of the species–area relationship ($S = cA^z$) and that $A_{small} = A_{large}/2$, then the number of species in the large reserve is $s_{large} = c(A_{large}^{z} = c(2A_{small}^{z} = 1.2S_{small})$, (**Figure 4**). This equation could yield S_{large}

$< 2S_{small}$ in a variety of realistic circumstances. Subsequently, R.W. Rafe elaborated the mathematical underpinnings of the comparison of the relative abilities of a single large reserve (SLR) and two reserves of half the area (THR) to support species. R.W. Rafe, A.J. Higgs, M.B. Usher, and others clearly illustrated that the proportional overlap of species between the smaller reserves is a critical issue. In essence, the greater the overlap, the more an SLR is favored, and vice versa. The relevant, and critical, information necessary to predict the amount of overlap is mostly lacking, a fact mentioned earlier by Simberloff and Abele in response to criticisms of their analysis.

By the early 1980s, the controversy over choice of nature reserves had evolved into the single-large-or-several-small (SLOSS) debate. O. Järvinen, D. Simberloff, L.G. Abele, and others emphasized, once again, that the species–area relationship – or the underlying island biogeographic theory, for that matter – did not unambiguously favor a single intact reserve over an archipelago of smaller reserves. That some species might need relatively large areas for their continued well-being, whereas others could thrive in, or actually required, certain relatively small areas, had become well established in the ecological literature, beginning with the early work of J. Terborgh, R.T.T. Foreman, and others, so their proposal should not have been very controversial. This was not the case, however, and the debate continued. In 1986, M.E. Soulé and D. Simberloff gave the SLOSS debate the last rites: The SLOSS debate is no longer an issue in the discussion about the optimal size of nature reserves. Yet, rumors of the death of the SLOSS debate were greatly exaggerated, for it continued beyond the mid 1980s – briefly mutating into the single-large-or-plentifully-patchy (SLOPP) debate of M.P. Gilpin in the late 1980s – right up to the present. M.E. Soule and D. Simberloff suggested that the truly important remaining question was "the dynamics of species extinction after the reserves are set up and surrounded by habitat modified by human activities," as had A.J. Kushlan, H. Picton, and others nearly a decade earlier. Clearly, this statement implies that too much emphasis had been placed on area (the extinction side of island biogeography) and too little on isolation (the colonization side of island biogeography). It also implies that relatively small fragments may have its worth despite their small sizes. Relatively small fragments often are not simply random samples of larger habitat units, but rather may represent special places that have been left either inadvertently or purposely. For example, R.T.T. Forman, I. Hanski, E.D. McCoy, and others have shown that relatively small fragments could harbor unusually high population densities of species or disproportionate representations of rare species. Finally, this statement implies that fragmentation cannot proceed indefinitely without severe consequences. Simberloff, for example, has pointed out that "thousands of small fragments, with large aggregate area, will not be expected to allow conservation of many species." If this question – about the dynamics of species extinction – is indeed the only important one remaining, then the information needed to slow the loss of species accompanying area reduction largely is the same information that is needed to predict the loss of species in the first place. For example, J.G. Dony, M.E. Soulé, D. Simberloff, and others have recommended identifying target species, determining their minimum viable population sizes (MVPs), and then using known densities to estimate needed area. Unfortunately, even if sound MVPs could be calculated for several target species in a location, the areas required to support those population sizes are likely to be poorly known, if known at all. This lack of information also plagued earlier proposals for estimating needed area, such as E.D. McCoy's minimum refuge area, T.E. Lovejoy and D.C. Oren's minimum critical size, and S.T.A. Pickett and J.N. Thompson's minimum dynamic area.

How Incidence Functions and Nestedness Fit In

Recall that the SLOSS debate was generated largely by one of the early nature reserve design principles: a single intact reserve of a certain area will support more species than a series of individually smaller reserves that are cumulative of the same area. For this principle to apply universally, the minimal assumptions are that the habitats included on areas of different sizes are more or less uniform, that the population densities of species on areas of different sizes are similar, and that the process of relaxation largely is deterministic. Subsequent ecological research has made it clear that none of these assumptions always holds. However, do these assumptions hold in enough instances or to such a degree that it is possible to know the species composition of various-sized fragments with a high degree of certainty? We shall address two related methods developed to address this question.

J. Diamond developed incidence functions as a tool for determining minimum nature reserve area. The incidence function, I, equals the number of fragments (islands) of a certain size harboring a certain species divided by the total number of fragments of the same size. Supposedly, the higher the value of I at a particular size, the higher the probability that the species can persist in fragments of that size. E.F. Connor, D. Simberloff, M. Williamson, and others questioned the validity of these conclusions, but, if they do have any validity, even in a general way, then certain patterns of incidence functions could indicate that species disappear from increasingly fragmented habitats in a predictable manner.

B.D. Patterson and W. Atmar developed a sophisticated extension of incidence functions, the nestedness of species' geographical distribution. Species' distributions are nested when the species on the most species-poor (and, likely, smallest) fragment comprise a subset of those on the next most species-poor fragment, which, in turn, comprise a subset of those on the next most species-poor fragment, and so on. K.B. Jones and co-workers supplied the first explicit evidence for nested distributions in the mid-1980s, although implicit evidence can be found in papers by N.W. Moore, M.D. Hooper, R.T.T. Foreman, and others a decade earlier. Jones and co-workers attributed the pattern to selective extinctions in the absence of colonizations, as have most subsequent authors, although D.T. Bolger, M.V. Lomolino, and others have shown that the pattern actually can result from a number of causes. Nestedness could have important implications for nature reserve design. If the species on an archipelago of habitat fragments were a perfect subset of the species in a single large unit of habitat, then the archipelago could never contain more species than the single large unit. Furthermore, if fragments of similar size were to harbor similar suites of species, because

species loss with area reduction is deterministic, then the archipelago must contain fewer species than the single large unit. The temptation is therefore, great, to assume that simultaneously significant species–area and nested subset relationships must indicate that smaller fragments have the fewest species, but recent evidence supplied by D. Doak, E.D. McCoy, W.J. Boecklen, and others suggests that this assumption is not a good one.

Conclusions

Regardless of the reason for it, poor predictability of the effects of area reduction can stall effective decision making. A good example is the controversy over the amount of species loss accompanying tropical deforestation. For more than 30 years, ecologists have sought shortcut methods to substitute for the autecological information that is needed to improve predictability but is so difficult to obtain. To date, they have not been particularly successful. An interesting additional example that involves species–area relationships is the controversy over the amount of species loss accompanying global warming discussed by K.A. McDonald, J.H. Brown, R.W. Skaggs, W.J. Boecklen, T.E. Lawlor, and others. So, we are saddled with a seemingly unresolvable dilemma: we need more information, but we cannot afford to wait for it. Faced with this dilemma, some biologists, for example, R. East and M. Kent, have suggested that we view the species–area relationship as perhaps the best tool available for making conservation decisions and have promoted the idea of retaining large units of habitat as a matter of general course. If one is focusing solely on predicting the loss of species with area reduction and refrains from becoming overly specific, then this may be a reasonable strategy. Others, for example, C.F. Mason and E.D. McCoy, have suggested that reliance on species–area relationships may lead to undesirable conservation decisions and have promoted a "save-all-the-pieces" strategy. If one is focusing on forestalling the loss of species, then this may be a reasonable strategy (which has been termed "several little all too small" (SLATS) by E.D. McCoy and H.R. Mushinsky) for habitats suffering from extreme area reduction and fragmentation. The SLATS strategy is likely to become increasingly relevant for most habitats, because the amount of habitat needed to allow species to persist, let alone to flourish or to evolve, appears to be much larger than humans are willing to grant.

See also: Diversity, Community/Regional Level. Habitat and Niche, Concept of. Island Biogeography. Metapopulations. Population Density. Population Diversity, Overview

References

Coleman BD, Mares MA, Willig MR, and Hseih Y (1982) Randomness, area, and species richness. *Ecology* 63: 1121–1133.

Condit R, Hubbell SP, LaFrankie JV, *et al.* (1996) Species–area and species–individual relationships for tropical trees: A comparison of three 50-ha plots. *Journal of Ecology* 84: 549–562.

Connor EF and McCoy ED (1979) The statistics and biology of the species–area relationship. *American Naturalist* 113: 791–833.

Hanski I and Gyllenberg M (1997) Uniting two general patterns in the distribution of species. *Science* 275: 397–400.

Harte J, Kinzig A, and Green J (1999) Self-similarity in the distribution and abundance of species. *Science* 284: 334–336.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Princeton, NJ: Princeton University Press.

Rosenzweig ML (1995) *Species Diversity in Space and Time.* Cambridge, UK: Cambridge University Press.

Shafer CL (1990) *Nature Reserves: Island Theory and Conservation Practice.* Washington, DC: Smithsonian Institution Press.

Simberloff D (1976) Experimental zoogeography of islands: Effects of island size. *Ecology* 57: 629–648.

Simberloff D (1992) Do species–area curves predict extinction in fragmented forests? In: Whitmore TC and Sayer JA (eds.) *Tropical Deforestation and Species Extinction*, pp. 75–89. London: Chapman & Hall.

Williams CB (1964) *Patterns in the Balance of Nature.* London: Academic Press.

Williams MR (1995) An extreme-value function model of the species incidence and species–area relations. *Ecology* 76: 2607–2616.

Williams MR, Lamont BB, and Henstridge. JD (2009) Species–area functions revisited. *Journal of Biogeography* 36: 1994–2004.

Species Assemblages, Macroecology, and Global Change

Brian J McGill, University of Maine, ME, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

β diversity How two assemblages at different points in space differ (or equivalently how the composition of an assemblage changes as one moves across space).

Assemblage A group of organisms found within some specified geographic area at one point in time that are evolutionarily related and that include many species, typically spanning across a taxonomic Family, Order, or Class.

Frequency distribution A histogram or description of the distribution of values. The frequency distribution not only contains information about the mean and the variance, but also higher moments. The normal (or Gaussian distribution) is the best known frequency distribution. But many aspects of assemblages show a lognormal distribution. Empirical data may show a frequency distribution that is not a good match to any known theoretical distribution.

Scale The order of magnitude, length, area or time period relevant to the data being collected and processes being studied. Thus a small scale might be data collected within one square meter, while a large scale might be data collected spanning 100 square kilometers.

Trait Any property (usually quantitative) of an individual that can be measured. The distribution of traits in an assemblage then gives a representation of the assemblage with respect to that property. Some properties (e.g., abundance or range size) can only be measured at the species level, which technically makes them not a trait, although the author will refer to "species traits" as a shorthand.

Species Assemblage Basics

Definition

Ecologists have long been fascinated by studying systems with more than one species. Thus, community ecology is one of the older subfields in ecology. But ecologists have been surprisingly vague about what exactly is the proper unit of study when more than one species is involved. Terms such as community, guild, assemblage, etc. have all been invoked. To an insider, there are nuances associated with different words. Community emphasizes a group with strong interactions between the species. Guild implies a group that shares and competes for a common resource. Assemblage is usually recognized as a relatively looser group. The rise of macroecology (Brown and Maurer, 1987; Brown, 1995) has chosen to focus on the assemblage as the unit of study, while still not providing a precise definition. In attempting to bring clarity to this terminological thicket, Fauth *et al.* (1996) identify an assemblage as a group of organisms that share a common geographical location and a common phylogenetic history. This accords reasonably well with the sense used in macroecology. Brown (Brown, 1995) notes "Most (macroecological) hypotheses and analyses require that the organisms being considered be not precisely identical but at least comparable, in the sense that they are subject to similar evolutionary and/or functional constraints". Thus, for the purposes of this entry, an assemblage is defined as "a group of organisms found within some specified geographic area at one point in time that are evolutionarily related and that includes many species, typically spanning across a taxonomic Family, Order, or Class". Thus, the birds of Yosemite national park are an assemblage, as are the sphingid moths of North America.

Processes Acting on an Assemblage

An assemblage, like any other entity studied in ecology, is of interest because it is presumed that there is a common set of forces or processes that affects all examples of assemblages, allowing for the development of a general theory about which forces have what effects under what conditions. Broadly there are five types of processes that are presumed to drive most of the variation and pattern that we see in assemblages:

- Biotic filtering – species interactions such as competition, disease, or predation that limit which species can be part of an assemblage.
- Abiotic filtering – environmental constraints such as frost or drought that limit which species can be part of an assemblage.
- Evolution – the processes of speciation and global extinction that control which species can potentially occur in an assemblage.
- Dispersal limitation – when some species that could exist within an assemblage do not simply because they have not sent propagules to the geographic area of the assemblage. Dispersal limitation is often perceived largely as random and thus as the main way in which stochasticity is incorporated into assemblages.
- Sampling – when small areas are studied, only a few individuals can fit into the area. Because the few individuals cannot possibly span the diversity in the larger region, random sampling determines which species, traits, etc. are represented in small areas.

Thus, the study of assemblages is about determining the effects and relative importance of these five forces.

The Role of Scale

One of the most important determinants of which processes are most influential on an assemblage is the spatial scale. The definition given above of an assemblage was fairly specific about what taxonomic extent applies (Family through Class) but extremely vague about spatial scale. Under the definition given, an assemblage could be anything from the birds of Frank Preston's back yard to the globe, and indeed assemblages of birds have been studied at both extremes (Preston, 1960) and nearly every scale in between. If the one word "assemblage" is going to be used to cover such a range of scales, it is important to recognize that the relevant processes may change with scale. The current working assumption (although it is still a hypothesis needing more testing) is that biotic filtering and random dispersal limitation are most important at very small scales. As the scales get large enough to cover varying climates, abiotic filtering becomes important. And at the largest scales, speciation and extinction become important as well as the fact that dispersal events are now exceedingly rare leading to preemption effects (the first species to successfully migrate influences the long-term trajectory of evolution).

Describing an Assemblage

The simplest description of an assemblage is a list of the species present. A much more useful description gives not only the list of species but also the abundance of each species (**Table 1**). Such descriptions, however, are rather cumbersome and assume the reader is well acquainted with the biology of the species in the list. The next five sections present five different ways of describing assemblages. The first two ways are short in comparison to a list (**Table 1**) and represent ways to summarize an assemblage. One focuses on means and other univariate descriptions of assemblages, while the second summarization captures the broader distribution; but both methods lose the detailed list of species identities. The third approach involves looking how different attributes of an assemblage are related to each other. The fourth and fifth approaches involve comparison of two assemblages (the first three all involve just one assemblage). The fourth involves comparing assemblages across space, while the fifth involves comparing assemblages across time.

Table 1 The most common representation of an assemblage – a list of species, possibly with the abundance of each species included

Species	*Abundance N_i*	*Mass (g) m_i*
American Black Duck	1	1200.0
American Kestrel	1	117.0
Killdeer	4	95.0
Barred Owl	1	720.0
Chimney Swift	7	23.0
Ruby-throated Hummingbird	1	3.2
Yellow-bellied Sapsucker	11	50.0
Downy Woodpecker	1	27.0
Hairy Woodpecker	1	66.0
Yellow-shafted Flicker	5	130.0
Least Flycatcher	3	10.3
Eastern Phoebe	10	20.0
Eastern Kingbird	4	40.0
Philadelphia Vireo	3	12.0
Red-eyed Vireo	2	17.0
Blue Jay	4	85.0
American Crow	24	450.0
Tree Swallow	6	20.0
Cliff Swallow	11	21.0
Barn Swallow	16	19.0
Black-capped Chickadee	13	11.0
Red-breasted Nuthatch	7	10.0
White-breasted Nuthatch	1	21.0
Ruby-crowned Kinglet	2	6.5
Veery	31	31.0
Hermit Thrush	2	N/A
Wood Thrush	33	31.0
American Robin	47	77.0
Gray Catbird	6	37.0
European Starling	53	82.0
Northern Parula	2	8.6
Yellow Warbler	6	9.5
Chestnut-sided Warbler	1	9.7
Myrtle Warbler	1	N/A
Black-throated Green Warbler	1	8.8
Black-and-white Warbler	1	10.7
Ovenbird	21	19.5
Common Yellowthroat	9	10.0
Chipping Sparrow	12	12.0
Savannah Sparrow	2	20.0
Song Sparrow	35	20.0
White-throated Sparrow	37	26.0
Rose-breasted Grosbeak	6	45.0
Red-winged Blackbird	9	52.0
Common Grackle	14	115.0
Brown-headed Cowbird	1	44.0
Purple Finch	7	25.0
American Goldfinch	14	13.0
House Sparrow	2	28.0

Mean and Univariate Descriptions of an Assemblage

Obviously, the greatest possible simplification in describing an assemblage is in reducing **Table 1** to a single number. The oldest and simplest such measure is to assess the diversity of an assemblage by counting the number of species, S, present (i.e., species richness). Counting the number of individuals, N, in an assemblage is also common and can be revealing despite its simplicity. Because species richness increases in an inconveniently nonlinear fashion with the number of individuals sampled (Gotelli and Colwell, 2011), a number of measures have attempted to provide sample-size corrected comparators of species richness, including Margalef diversity $S_{\text{Margalef}} = (S-1)/\ln N$, the Chao estimation of total (including unobserved) species richness $S_{\text{Chao}} = S + S_1^2/(2S_2)$, where S_1 is the number of singletons (species observed once) and S_2 is the number of species observed twice, and finally rarefaction derived comparisons (Gotelli and Colwell, 2011). The free software package EstimateS provides easy, state-of-the-art access to these methods.

There are also a wide variety of methods that measure the evenness of a community – how similar in abundance the

different species in the assemblage are. As will be seen in the next section (*see* Variation Within an Assemblage), the short and universal answer is always not very even, but this still allows for potentially interesting variations in degree of evenness between different assemblages. The notion of diversity is intended to capture both richness and evenness (a community of five equally abundant species is more diverse than a community with five species, four of which only occur once). Two common measures of diversity are Shannon's $H = -\Sigma p_i \ln p_i$ and Simpson's $D = 1/\Sigma p_i$ where p_i is the relative abundance of each species i (i.e., $p_i = N_i/N_{\text{total}}$). These can be turned into evenness measures by dividing by the maximum attainable values (Shannon evenness $J = H/\ln(s)$ and Simpson evenness given by D/S). Dozens of other univariate metrics capturing some aspect of an assemblage exist (Maurer and McGill, 2011).

Treating an assemblage as measured by traits rather than species and abundances can lead to different views. If one wants to provide a simple assemblage-level description of a trait such as body mass or maximum lifespan, some form of mean is the obvious choice. But exactly which mean to report is not as obvious. Firstly, some traits such as body mass are clearly made normal in shape by a log-transformation suggesting that a geometric mean ($\mu = (\Pi m_i)^{1/S}$) is more appropriate, although if there is no clear evidence of the appropriateness of a log-transformation, it is probably best to use an arithmetic mean. Secondly, it is not clear whether a species mean, an abundance-weighted-mean, or an individual mean is best. A species mean takes a species-level trait value, $\mu_{\text{species}} = 1/S\ \Sigma m_i$ (*cf.* **Table 1**). An abundance-weighted mean weights the species trait by the abundance of that species $\mu_{\text{abund-weighted}} = 1/N_{\text{total}}\ \Sigma m_i N_i$. Finally, an individual-level mean averages across all individuals: $\mu_{\text{individual}} = 1/N_{\text{total}}\ \Sigma m_i$ (where m_i is mass of an individual *cf.* **Table 2**). The abundance-weighted mean and the individual mean should be identical only if the species mean trait values are calculated from the assemblage under study. See the legend of **Table 1** for reports of common mean and univariate statistics summarizing the assemblage in **Table 1**.

Table 2 An individual perspective on the same data as in **Table 1**. Note that each row represents an individual (the full version of **Table 2** would have 492 rows for the same data as **Table 1**). Rather than recording abundance, species with abundance greater than one show up in the table more than once. And rather than using species averages, traits (here body mass) are measured on each individual

Individual band ID	*Species*	*Mass (g)*
3719	American Black Duck	1218.2
3824	American Kestrel	109.7
3823	Killdeer	92.1
3783	Killdeer	93.6
3812	Killdeer	97.8
3725	Killdeer	99.2
3765	Barred Owl	720.5
3801	Chimney Swift	22.0
3754	Chimney Swift	24.3
…	…	…

Variation within an Assemblage (Distributions or Histograms)

The last section focused on various attempts to summarize an entire assemblage to a single number, ignoring variation. This section explores descriptions of the variation found within an assemblage but which still simplifies out the details of which species are present. Ultimately ecology and evolution are studies of variation. If there was only one species with clonal copies of an identical organism, there would be no need for the study of ecology or evolution. It is the enormous variation of life that defines both fields. To quantitatively capture this variation, some quantitative property must be identified. One can then describe this trait across the assemblage. The typical way of reporting the variation is to plot a histogram giving the relative frequencies of different trait values. This quantitative description of the variation of traits in assemblages is a central pursuit of macroecology. Before launching into the description of the variability of various quantitative properties of an assemblage, it is useful to address two technical points: the level at which the properties are measured, and introducing the lognormal distribution.

Species Versus Individual Level Distributions

There are several ways to look at properties across an assemblage: one is species centric and one is individual centric. An individual-centric property distribution measures the property on each individual in the assemblage (**Table 2** for the property body mass). In contrast, a species-centric distribution measures a property value for each species (**Table 1**) and then examines the property across all species in the assemblage. Some properties can only be examined in a species-centric fashion since the property is inherent to the species rather than the individual; examples include the abundance of a species, species age, and the size of the species range. But other properties like body mass can easily be conceptualized in either a species-centric or individual-centric approach (White *et al.*, 2007) (**Figure 1**) (compare **Tables 1** and **2**). It is a surprisingly under-examined question, which approach is most appropriate. At its heart, the choice ought to be indicative of the level (species vs. individual) at which one thinks the four processes operating on an assemblage apply. A hybrid approach, the abundance-weighted species distribution, serves as a compromise. This approach takes one trait value for each species and then multiplies it by the number of individuals (i.e., abundance) of each species. In practice, the abundance-weighted species distribution is used primarily in situations where an individual-centric distribution is desired but data at the individual level are not available (**Figure 1(c)**). Whether it is a good approximation depends on how similar the individuals within a species are. It has long been assumed that individuals within a species are more similar than differences between species, but in practice, traits can vary substantially within a species (and often more than between species) (Messier *et al.*, 2010).

Lognormal Distributions

As will be seen shortly, a large number of features of assemblages show a log-normal distribution or something

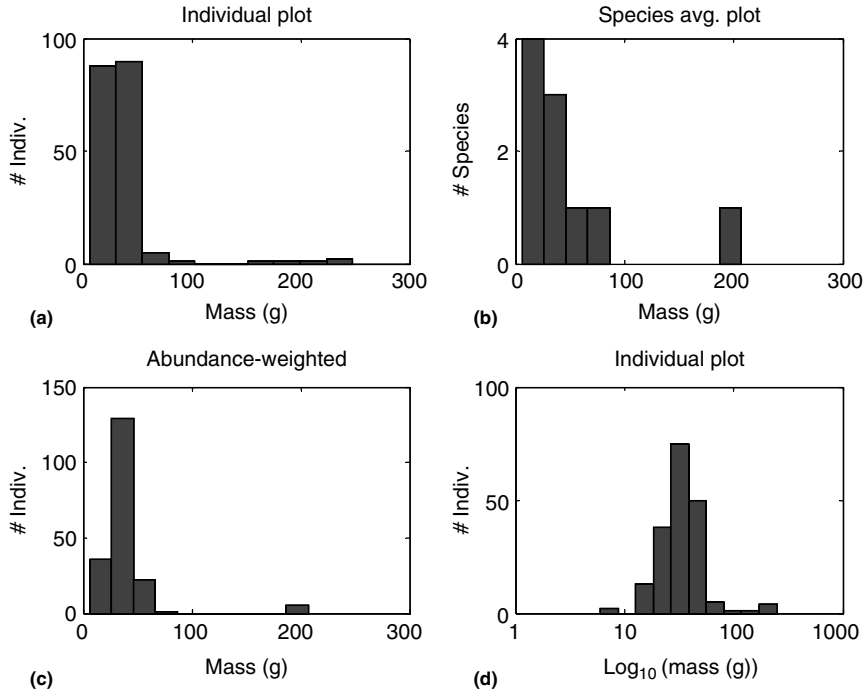


Figure 1 The distribution of body masses for rodents trapped during one trapping period at Portal, New Mexico (Ernest *et al.*, 2009). The data are the same in all four graphs, but are plotted in different ways. (a) The mass of each individual rodent is used to generate a histogram (frequency distribution). (b) The mean of each species is used to generate a histogram (note the change of interpretation of the y-axis). Plots (a) and (b) are broadly similar (most species/individuals are small with a few very large species/individuals), but some clear differences occur including a smaller proportion of the very smallest animals and more spread around 200 g in the individuals plot (a). (c) An attempt to recreate the plot in (a) using the species averages of (b) and the abundance of each species. It clearly is an inaccurate representation of (a). (d) The same plot as (a) but on a log scale for body mass. This figure looks near to a normal distribution (Gaussian Bell-curve), which suggests the distribution of mass is close to a log-normal distribution.

approximating a log-normal distribution. A variable X shows a log-normal distribution when $\log(X)$ is normally distributed. Lognormal distributions generally appear strongly skewed (asymmetric) when plotted as a histogram on an arithmetic scale. This shows there are a few instances of very large values of X and a great many instances of small values of X. Plotting a histogram on a log-scale removes this skew and reveals something close to a symmetric normal (Gaussian) distribution (**Figure 1(d)**). There may remain small amounts of skewness (asymmetry) with the left or right side being slightly more predominant. Considerable debate has gone into whether these asymmetries are real aspects of the data (and in which case what causes them) or whether they are artefacts (e.g., sampling biases). The lognormal is a ubiquitous distribution across the sciences (Limpert *et al.*, 2001) describing the distribution of sizes of computer files on a hard-drive (many small files, few large files) to the distribution of income among individual humans (again many low incomes and a few large incomes). The lognormal distribution is generated when many random variables are multiplied together. This makes it hard to identify specific processes for data showing a lognormal distribution (just as we do not attempt to explain why data has a normal distribution). There are a number of other distributions with strong right skew on an arithmetic scale (e.g., the logseries, zero-sum multinomial, zipf-mandelbrot, and power distributions are all commonly used in place of the lognormal distribution). In practice, ecological data are noisy enough that it is almost impossible to decisively adopt the lognormal in preference to one or more of these alternatives. Thus, the author uses lognormal here as a general description of the overall shape of a distribution while acknowledging that other very similar distributions might fit equally well.

Body Mass Distribution

Body mass of an individual is one of the most basic of all measurable traits. It directly influences resource requirements and it affects the scale at which the individual interacts with the world. In general, body mass is distributed lognormally across an assemblage (**Figure 1**). Thus, there are many smaller organisms and only a few larger organisms. Body mass was first studied at the species level (distribution of average body masses of all the species in an assemblage) (May, 1988), but has also been extensively studied for individuals (White *et al.*, 2007). May noticed that the distribution was asymmetrical with species apparently missing on the left (small mass) side and suggested this might be due to poorer sampling of small species. Studying individual level assemblages especially makes sense for species which are indeterminate growers (leading to wide variation in body mass within a species) or in which juveniles are predominant. Thus, studies of tree size and plankton size have been especially common at the individual level. For individual-level distributions, it is common to apply the power-distribution and

study the exponent, which often is close to -2. However, both the species and individual level distributions of body mass are roughly lognormal in nature. Several authors have also shown that the shape of the body mass distribution changes with spatial scale. In particular, as the spatial scales become small enough such that individuals in the assemblage are likely to compete with each other, competitive forces appear to dominate and the lognormal distribution transforms into a log-uniform distribution (a flat histogram on a log scale) (Brown, 1995), which is the probability distribution version of Hutchinsonian ratios of body size.

Abundance (Species Abundance Distribution or SAD)

Another very basic property, which unlike body mass is only applicable at the species level is abundance. How many individuals of each species are there? It is a somewhat surprising but extremely well-established and general fact that there is enormous variation in abundance across an assemblage. Different species of birds in North America vary by seven orders of magnitude of abundance (McGill, 2006). There are typically a few species that are very abundant and many species that are rare. This is suggestive of a skewed distribution and indeed SADs appear lognormal-like (for historical reasons $\log_2$ is typically used) (**Figure 2(a), (b),** and **(c)**). This pattern seems to be as close as ecology gets to a truly universal law. A great many attempts have been made to understand which mechanism is responsible for lognormal-like SADs (McGill *et al.*, 2007), but there is a growing consensus that the SAD does not have enough information to differentiate which mechanisms or processes drove community assembly. This is especially true in the context of our growing understanding of how small sample sizes affect the shape of the SAD. Early workers (Preston, 1948) thought small samples simply chopped off the left hand side of the lognormal distribution (a so-called veiled distribution), but recent more mathematically rigorous work shows that the effects are more subtle and varied but that analyzing small samples distorts the underlying distribution (Green and Plotkin, 2007).

The SAD appears to be lognormal-like across all spatial scales. At one extreme is a small community of strongly interacting individuals (**Figure 2(a)** and **2(b)**). As long as the sample size is large enough to be statistically useful (usually at least a few hundred individuals), then a lognormal-like distribution is found. At the other extreme, global SADs look at the distribution of the total number of individuals across the globe for each species in an assemblage (**Figure 2(c)**). Such global SADs also appear to be lognormal-like, although it has been shown several times that as the spatial scale becomes larger, SADs becoming increasingly more log-left skewed (the left tail of the lognormal distribution becomes heavier) due to the increasing number of transient species detected (McGill, 2003).

Range Size and Occupancy Distribution

Measures of the spatial extent covered by a species are another example of an attribute that makes sense only at the species but not at the individual level. Two common measures are used. One is the size (area) of the geographic range of a species. This typically conceptualizes a species as occupying a blob covering typically some portion of a continent (or continents) such as those found in field guides, and the area of this blob is the range size. Note that this is different than the home range, the area within which a single individual lives and hunts. Alternatively, one can see whether a species is present or absent at a number of different sites. Then the percentage of sites the species is present in is known as occupancy. Clearly if the sites used are distributed over a large part of a continent, then occupancy and range size would be strongly correlated to each other. However, occupancy can also be applied at much smaller spatial scales, in which case it becomes a measure of how much a species fills in its range or how many different types of habitats a species occurs in.

The distribution of range sizes across an assemblage also show a roughly lognormal (**Figure 2(d)**) distribution with a few species covering large portions of the globe and most species being much more narrowly distributed. Occupancy shows a different pattern (**Figure 3(b)**), however. Occupancy is capped at 100% and thus cannot show a long right tail of increasing values. Instead a histogram of occupancies typically shows a roughly U-shaped pattern, with many species with low occupancies, fewer species with intermediate occupancies, and then slightly more species with high occupancies. This U-shaped distribution of occupancy is one of the earliest patterns to be documented as being generally true of all assemblages (Raunkiaer, 1909).

Traits and Trait Distributions

One can imagine measuring the distribution of a number of other traits across an assemblage. For example, one could measure the distribution of life history traits such as maximum life span across an assemblage. Or one could measure functional traits such as maximum metabolic rate (in an animal) or maximum photosynthetic rate of a leaf (in a plant). The idea of a trait is increasingly gaining traction in ecology and is defined as some quantitatively measurable aspect of an individual that is thought to be of significance to the organism. Two broad categories of traits are functional traits, which are measurements of features that have a direct relevant to performance and fitness such as photosynthetic rate or leaf mass per unit area (basically leaf thickness) and life history traits, which describe the overall allocations to survival and reproduction of an organism such as age of first reproduction or maximum life span. Relatively little study has been made of the assemblage-wide distribution of such traits. One thing that seems clear is that assemblages show a great deal of variation in traits not only between species but also within species. Variation within a single assemblage captures a large fraction of all global variation in a trait (Messier *et al.*, 2010). However, general expectations about whether traits are distributed normally or lognormally or otherwise have not yet been developed (**Figure 2(e)** and **2(f)**).

Covariation (Correlation) within an Assemblage

The last two sections covered patterns addressing a single variable measured across an assemblage. This effectively

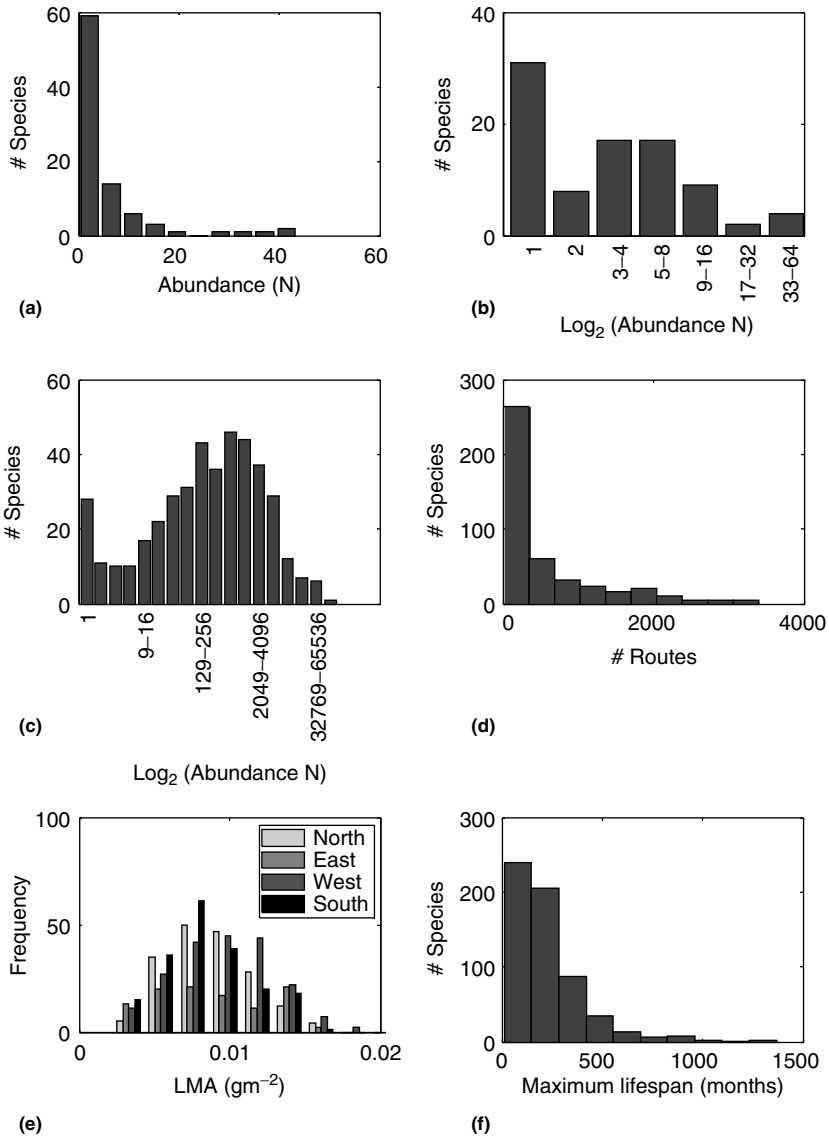


Figure 2 The frequency distribution (histogram) of several species-level properties of assemblages of birds found from the North American Breeding Bird Survey (Sauer *et al.*, 1997). (a) The distribution of species abundances for a single route at 45.05N, 69.033W near Orono, Maine. There are many rare species (species with few individuals) and only a few very abundant species. This is a local species abundance distribution. (b) The same data as (A) plotted on a logarithmic scale (plots for species abundance distributions traditionally are plotted on a log-base-2 scale). This is approximately normal (implying a lognormal distribution) but there is an excess of singletons (species with only one individual). (c) A $\log_2$ plot of total species abundances across all routes in the entire region covered by the North American Breeding Bird Survey (the continental US and southern Canada). This represents a global species abundance distribution. It is similar in shape to the local distribution (b). (d) The frequency distribution of range sizes of species on an arithmetic scale (which is similar in shape to (a) and is also approximately lognormal). (e) Frequency distribution of LMA (Leaf Mass per unit Area). The distribution at four different plots labeled by compass directions is shown. Note how the frequency distribution of this trait remains very similar across different plots (despite having very different sets of species in each plot). LMA is a commonly measured plant functional trait that measures leaf thickness. It is strongly positively correlated with leaf lifespan. Note that this functional trait is distributed nearly normally. (f) Frequency distribution of maximum life span (in months) of mammals (Ernest, 2003). Note that in contrast to (e) this trait is distributed in a fashion approaching lognormal.

assumes every property of an assemblage is independent of every other property, and more specifically that each property can be examined in isolation. In practice, trade-offs and allocation strategies often result in strong interactions between properties within an assemblage. These are most commonly studied using a correlation or regression approach. As before, these correlations sometimes occur at the species level and sometimes occur at the individual level.

Energy versus Mass

One of the oldest known linkages between two traits is between the body mass of an individual, M, and the amount

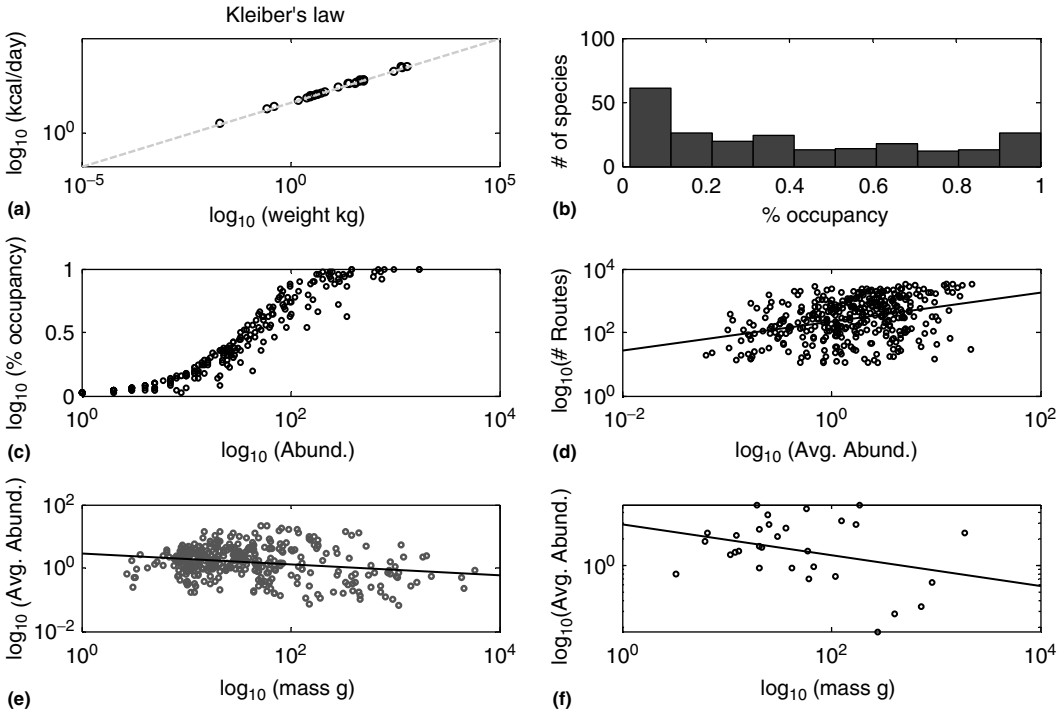


Figure 3 The relationships between various properties of individual species within an assemblage (except Figure (b)). (a) The relationship between body mass and energetic flux (metabolic rate) known as Kleiber's law. (b) A frequency distribution of occupancies in trees in a 50 ha plot on Barro Colorado Island, Panama (Condit *et al.*, 2000). Compare with **Figure 2(d)**. This plot also shows many species with low occupancy, but in contrast shows many species with high (near 100%) occupancy. This is due to the truncated scale of occupancy (occupancy cannot go higher than one). (c) The relationship between abundance and occupancy in the same data as (b). There is a positive correlation between abundance and occupancy with very low abundances having occupancies close to zero and very high abundances having occupancies of 1.0 (100%). (d) The relationship between abundance and range size in the North American Breeding Bird Survey (Sauer *et al.*, 1997) (here measured by the total number of routes a species is observed on). There is a positive relationship, but there is considerable variation. (e) The relationship between body mass and abundance of species in the same dataset as (d). The relationship is negative but very noisy. (f) The same data as (e) plotted at the taxonomic level of family rather than species.

of energy consumed, *E*. Kleiber (1947) showed that among mammals, $E = cM^{3/4}$ (**Figure 3(a)**). This is known as the mouse-to-elephant curve as it places the energy consumption of species as small as a mouse and as large as an elephant on a single standardized curve. Later work has extended the link between mass and energy by showing that this pattern appears quite general. Different groups of organisms may have a different value for the constant, *c*, with high energy requirement groups like birds having a higher *c*, and low energy requiring groups like reptiles having a smaller *c*, but the exponent is always very close to 3/4. Amazingly, it is possible to place all species from single-celled organisms to blue whales on a single curve with $E \propto M^{3/4}$. Plants also follow 3/4 scaling although it is not common to measure plant mass, leading to other exponents such as the scaling of energy with diameter at breast height (DBH) as $E = c(DBH)^2$. One of the most surprising aspects of this rule is the extremely high amount of variance explained, often having $r^2 > 0.95$. A number of recent studies have begun to show that 3/4 scaling may not hold at the margins such as for young plants (Muller–Landau *et al.*, 2006), prokaryotes (DeLong *et al.*, 2010) and the smallest mammals (McNab, 1986). These seem to be exceptions that prove the rule, however.

The 3/4 scaling of energy flux rates in turn seems to drive a number of other closely related factors (Calder, 1984). For example, production (growth in body size plus mass of offspring) appears to scale as $M^{3/4}$. Amazingly, other factors still more distantly related to energy flux such as the population growth rate, *r*, also scales as $M^{3/4}$ although the amount of variance explained goes down as one gets more removed from pure energy flux. Note that the energy flux per unit mass ($\propto M^{3/4}/M = M^{-1/4}$) scales as $M^{-1/4}$ so larger animals have less energy flux per unit volume. Thus, an elephant has more total energy fluxing through its body than a mouse, but has less energy per unit mass (and effectively per unit volume). One important implication of this is that drug dosages usually scale as $M^{3/4}$, thus a human who weighs twice as much as another human does not need twice as much drug but rather $2^{3/4} \approx 1.68$ times as much drug. Also, note that energy has units that include seconds in the denominator (s^{-1}) so one might suspect that the reciprocal of mass-specific energy (which is proportional to $M^{-1/4}$) would give the scaling of time, and indeed, many time based measures such as time between heart beats or average life span scale as $M^{1/4}$. An interesting sidenote is that this means that the total number of heartbeats is proportional to lifespan divided by time-

between-beats $\propto M^{1/4}/M^{1/4} \propto M^{\circ}$ suggesting that the total number of heartbeats in a lifespan is a constant across a group such as mammals.

There are a number of theories to explain 3/4 power scaling. The oldest theory, which noted that surface area increases as length^2 while mass increased as length^3 predicted a $M^{2/3}$ scaling. While 0.667 is close to 0.75, there seems to be clear evidence that 2/3 scaling is not correct and that while there is noise in the exponent, the results clearly center on 3/4. A recent paper (West *et al.*, 1997) suggested that 3/4 scaling is a result of fractal branching networks used to distribute the key components of energy (e.g., blood vessels, xylem tubes). This model makes a number of assumptions and a number of predictions about fractal branching networks, which have not yet been decisively tested. Several other explanations have emerged recently including a theory based on a different type of distribution network, an exploding network (Banavar *et al.*, 2010), on a general comparison of space (3-dimensions) versus lifetime (4 dimensions including time), and an argument that different systems (vascular, pulmonary, renal, etc.) each have their own scaling exponent all of which are bounded between 2/3 (for the reasons given above) and 1, but which average out to about 3/4 (Darveau *et al.*, 2002). More work is needed to decisively test which of these mechanisms are central.

Abundance Versus Occupancy/Range-Size

Another pair of variables that are often linked to each other are abundance versus measures of spatial extent (range size or occupancy depending on the scale of measurement). In general species of high abundance occupy larger areas (**Figure 3(c)** and **3(d)**). However, care must be used to avoid circularity. For example, global abundance will be trivially correlated with the number of sites a species is found at. The most careful approach is to use average abundance at all sites where the species is observed (i.e., average all nonzero abundances across sites for a species), which eliminates such statistical artifacts. This measure still tends to have a strong positive correlation with range size and occupancy. The abundance–range size relationship has a good deal of noise (**Figure 3(d)**). In contrast, the abundance occupancy relationship (**Figure 3(c)**) is fairly tight and can be made even tighter when correcting for the fact that species are clumped in space (Conlisk *et al.*, 2009). Thus, while in principle different aspects of rarity such as abundance and range size used as criteria for listing a species as endangered can vary independently there is a propensity for these aspects to be correlated.

Abundance Versus Mass

As already seen, body mass is strongly correlated with energy flux and ultimately with resource requirements. Thus, one might expect that the size of a population that could be supported might be related to body mass and that body mass could then be predictive of the widely varying abundances of different species in an assemblage. A simplistic guess would be that each species would consume the same amount of energy. This suggests that the number of individuals, N, times the energy requirements, $E = cM^{3/4}$ multiplied together to give the total energy usage by a species would be a constant – i.e., $k = NcM^{3/4}$. Solving this for N and combining constants we get $N = cM^{-3/4}$. Although this model is based on the bold assumption of equal energy per species, there is some evidence that it holds (Damuth, 1987). This pattern seems to hold most strongly when data are chosen to make sure that the assemblage studied is all from a single trophic level (e.g., mammalian herbivores) and that measures of maximum or global abundance are used. This model is usually applied at the species level, but can be applied at the individual level by using bins of different size classes (Cyr *et al.*, 1997; Enquist *et al.*, 1998). However, the current consensus (McGill, 2008) suggests that many organisms do not follow this precise law, showing a negative scaling exponent (large organisms are less abundant than small organisms) but an exponent greater than $-3/4$, suggesting that large organisms capture larger proportions of the total resources consumed by an assemblage (**Figure 3(e)**).

Abundance Versus Other Traits

The orders of magnitude variation in abundance across different species in an assemblage remains a striking pattern demanding explanation. As a result, there have been many studies seeking to find traits of a species that will predict its abundance relative to other species in the assemblage. In short what traits make a species common or rare? Factors such as dispersal ability, vegetative versus sexual reproductive modes, ability to enter dormant states, etc. have all been explored (Murray *et al.*, 2002). To date, though, this research program has had little success with no predictors that seem consistent across different studies and no predictors that explain significant proportions of the variability in abundance (beyond the already mentioned weak relationship to body size and range size). One reason for this is that most studies use local measures of abundance, but it is known that a single species will show considerable (usually orders of magnitude) variation in abundance across space. Thus, comparing abundance from a randomly chosen site (and thus essentially a randomly chosen abundance for the species) versus a trait will have considerable noise. It is possible to turn this around and use the spatial structure of abundance for a single species across space to predict abundance: distance from a species' peak abundance is one of the better predictors of relative abundance in a community (McGill and Collins, 2003). Going forward, any attempts to find species-level traits that correlate with abundance should use a species-level (i.e., global) measure of abundance. It is well known that abundance is a very labile feature in phylogenies, meaning that species that are evolutionarily close to each other can have widely varying abundances or equivalently that evolutionary proximity is a poor predictor of abundance (McGill, 2008). Understanding what factors make a species abundant or rare remains a major unsolved problem.

Traits Versus Traits

One can look at the covariation within an assemblage of any two traits. When positive correlations are found, this is often

indicative of some common underlying factor driving both traits. When negative correlations are found, this is often indicative of trade-offs. For example, the recent discovery of the leaf economic spectrum has shown that a suite of leaf functional traits are related. Leaf thickness (or leaf-mass-per-unit-area) and leaf life span are positively correlated with each other, as are maximum photosynthetic rate per unit mass, nitrogen and phosphorus concentrations (mass based) while the two groups are negatively correlated; the first principle component explains 74% of the variance (Wright *et al.*, 2004). Proximately this may well be driven by the common underlying factor of how much investment in structural material is made in the leaf. A more ultimate cause may be trade-offs in leaf investment in environments with high- or low-leaf mortality. In contrast, there are negative correlations between root biomass and shoot biomass within a single species (a plant has to choose how much energy and nitrogen to invest to each component).

Similar patterns can be found in life history traits. Maximum life span and age at first reproduction are positively correlated with each other but negatively correlated with measures of fecundity per unit time. This leads to the notion of a fast–slow spectrum. Some species live a weedy life style of growing quickly, reproducing early and often and dying young, while other species live long lives with slow reproduction and investment into high quality young. Broadly life span and fecundity are negatively related because allocation to fecundity reduces allocation to maintenance functions.

Variation between Assemblages Across Space (β Diversity)

The previous three sections explored patterns within a single assemblage (a set of individuals and species at a single location). It is also interesting to explore how similar or different assemblages are as one moves across space. By definition when comparing two locations, one is comparing two distinct assemblages. But these two assemblages could be very similar or quite different to each other. Tobler's First Law of Geography suggests that two assemblages that are closer to each other should be more similar than two assemblages that are far apart. The quantitative study of these questions is a field known as β (beta) diversity.

It is generally recognized that all five of the processes driving assemblages are involved in creating β diversity. Some difference is created by dispersal limitation – the fact that individuals' in one assemblage cannot easily disperse into a new assemblage in a new location. If there were no dispersal limitation and assemblages were connected by an infinite amount of dispersal then all assemblages would be identical (or more precisely, all differences would be due only to finite sampling effects). If there is an underlying environmental variability, possibly arranged along a gradient or possibly arranged in a patchy fashion then deterministic abiotic filtering will cause species assemblages to differ at points in space with different environments. This of course can then create differential biotic filtering across space, which can further exaggerate assemblage differences. Finally, when one moves across spatial scales so large that the assemblages have experienced different evolutionary histories and thus different regional source pools, the assemblages must necessarily differ.

Three Ways to Describe β Diversity

There are broadly three ways to characterize β diversity. On the surface they appear quite different, but they are in fact capturing the same underlying processes – the change in the make up of assemblage as one moves across space. The simplest method and the origin of the term "β diversity" comes from Whittaker (1975). Whittaker described a situation where one looks at a small plot embedded in a much larger region. Following this hierarchy, at the small end one has α diversity (e.g., species richness) of a single plot (in practice usually an average of richness across all plots) and at the large end of the hierarchy one has γ diversity as the diversity of the entire region. β diversity is what is left in the middle, specifically β diversity is the difference between the γ and α diversity (specifically $\beta = \gamma - \alpha$ or originally $\beta = \gamma/\alpha - 1$) (see **Figure 4(d)**). In practice, β diversity is the difference between the plots. If each plot has an identical set of species, then β diversity is zero. If each plot is slightly different from other plots in assemblage composition, then β diversity is a bit above zero and so forth to the extreme case where every plot has a different set of species in its assemblage giving maximum β diversity.

A second way of describing β diversity is to plot a species–area relationship (Rosenzweig, 1995) – a plot of species richness, S, versus plot size/area, A. Typically a species–area relationship is plotted on a log–log scale and is approximately a straight line, implying the species–area relationship is well approximated by a power–law relationship, $S = cA^z$ (**Figure 4(a)**). The steeper the slope of the species–area relationship (the value z on a log–log plot), the more rapidly new species are encountered as one expands the sampling area, and the higher the β diversity. When very large ranges of spatial scales are plotted, a tri-phasic (steep/shallow/steep sloped curve) is found (**Figure 4(b)**). This is often interpreted as indicative that different processes are involved in driving the species–area relationship at the three different scales (*see* Sampling Scale, *and* Habitat and Dispersal Limitation Scales, *and* Biogeographic Scales) (Rosenzweig, 1995).

The third way of describing β diversity is to use a measure of similarity between pairs of plots. First one needs a definition of the similarity between two sites. As with species diversity, there are dozens of proposed metrics. Two of the most common are the Jaccard similarity and Sorenson similarity. Jaccard similarity is the number of species in common between two assemblages divided by the total number of species across both systems. Sorenson similarity is the number of species common between the two assemblages divided by the average number of species across both plots. These measures use only information on whether a species is present or absent in an assemblage. An even more nuanced view can be used by looking at abundances, rating two assemblages as less similar if a species is very common in one assemblage and rare in the other even if it is present in both. One common metric for this is the Bray–Curtis measure. Let $N_{i,j}$ be the abundance of species i in community j. Then the Bray–Curtis index of

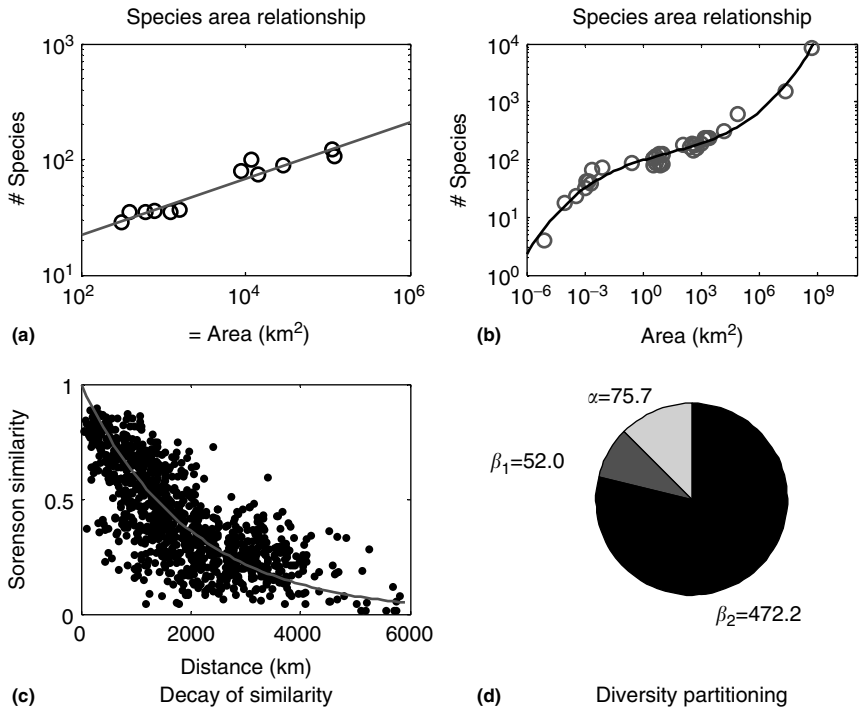


Figure 4 Different ways of reporting β diversity (change in species composition of assemblages across space). (a) Species–area relationship for birds. Each point is one island in the Carribean ranging from Grenada to Cuba (Preston, 1962). (b) A species–area relationship for birds of North America (Preston, 1960). Note that as the spatial scales covered expand, a more complex pattern than observed in (a) emerges – namely the curve becomes ''tri-phasic'' (steep slope, shallow slope, steep slope). (c) A plot of the decay of similarity of distance for the North American Breeding Bird Survey (Sauer *et al.*, 1997). (d) A plot of a hierarchical partitioning of diversity for the same data as (c). Each route has a diversity of 75.7 species (on average), which is α diversity. The total system (all of North America) has 600 species. This is reached by an average increase of diversity of 52.0 (to 125.7 species) within each stratum (the BBS classification of ecoregions). In contrast 472.2 new species are picked up by moving between strata. Thus, most of the diversity in the BBS occurs between different ecoregions and secondarily within a single route. Comparing routes (assemblages) within an ecoregion adds comparatively less diversity.

similarity is $2\ \Sigma_i\ \min(N_{i,1},N_{i,2})/\Sigma_i\Sigma_j N_{i,j}$. Given a measure of similarity, one can then plot similarity as a function of distance, which is known as a decay-of-similarity-with-distance plot (**Figure 4(c)** and **4(d)**). It typically shows something approximating exponential decay. This again describes how quickly species change as one looks at larger areas. Thus, all three approaches capture the same underlying process – different areas contain different sets of species. However, the first two methods (hierarchical and species–area relationship) look only at the addition of new species as area is added, while decay of similarity looks at differences between assemblages due to new species (as in the first two approaches) but also to species disappearing as area expands. Formal mathematical links between these three views of β diversity have been derived (Koleff *et al.*, 2003; Anderson *et al.*, 2011).

There are very few patterns known to be general about β diversity. It is likely that nowhere on earth has zero β diversity (i.e., where assemblages at different locations in space are all the same). Certain taxonomic groups tend to have lower β diversity, usually groups that have outstanding dispersal abilities such as bacteria or ferns, but even these groups do not show zero β diversity. Much β diversity is driven by underlying spatial change in the environment (Buckley and Jetz, 2008). There is growing evidence that human-built environments (cities, agriculture) tend to greatly reduce β diversity even though humans do not necessarily decrease α diversity (species richness of individual assemblages). Like most patterns about assemblages, the structure of β diversity depends on scale (Lennon *et al.*, 2001).

In the section Variation Within an Assemblage (Distributions or Histograms) we had a good understanding of the patterns, but a poor understanding of the processes. In contrast in this section on β diversity but we do not have a good understanding of the patterns in β diversity, but we do have a fairly good grasp of what the dominant driving processes are at different scales.

Sampling Scale

At the smallest scales (the area occupied by a few individuals – as small as a few centimeter for grasses to tens of meters for trees to many kilometer for some top predators), the main determinant of β diversity is just random sampling. A given location could be filled by individuals from many species, but at any one time only one individual can fill it (this is most obvious in plants). Thus, the chance event that determines which individual gets there first and establishes is the dominant factor. Microsite variation in topography, soil pH, etc. (some of which may be driven by the organisms) may play

some role, but stochasticity predominates. Thus, if we compare the "assemblage" at two sites (assemblage is in quotes here since assemblages at this scale consists of only one individual and thus one species) the differences will be stochastic, and the best predictor of the likelihood that both individuals will be from the same species is their relative abundances at regional scales (more abundant species are likely to be found at both sites). Sampling processes are clearly involved in the first steep portion of the triphasic species–area relationship (**Figure 4(b)**) and perhaps part of the second flatter portion.

Habitat and Dispersal Limitation Scales

Beyond very small scales, pure random sampling does not generate realistic levels of β diversity (McGill, 2011a). However, β diversity at intermediate scales are well-modeled if one models nonrandom sampling where individuals of a single species tend to come in clumps. Specifically to accurately predict β diversity at scales of kilometers it is necessary to know two things: not only the relative abundance of species (as in the sampling scale above) but also an understanding of the degree to which individuals of a species are likely to be sampled at the same time. If an individual of species X is sampled into assemblage A, then additional individuals of species X are more likely to be sampled again (than its relative abundance would suggest) into assemblage A. There are two main reasons for this clumping or correlation of individuals within a species. One explanation is habitat heterogeneity. As one moves across the landscape at scales of hundreds of meters to kilometers, one moves across riparian habitats, forests, meadows, etc. as well as smaller variations such as upland forest versus lowland forest. These habitats are driven by a combination of topography and disturbance history and strongly influence which species are most likely to settle there. This can cause individuals of a species to be more likely to co-occur in an assemblage than suggested by random chance. The second reason that we see this co-occurrence within an assemblage or clumping is dispersal limitation. Dispersal across space is in general a costly proposition. As a result, offspring generally are located close to their parents. This phenomenon is known as dispersal limitation and can result in amplifying chance locations of founding individuals, such that if one individual of species X happens to end up in the area covered by assemblage A, many of its offspring are also more likely to end up in the area of assemblage A. The effect of the increased probability of co-occurrence within a single assemblage A, is to make pairs of assemblages more different from each other. Thus, if species X is more common in assemblage A than expected by random chance, it is likely that in assemblage B, some other species Y is more common than expected by random chance, either because assemblage B occurs in a somewhat different habitat or because the random founder effects preceding dispersal limitation are different in assemblage B. Thus, at these intermediate scales ranging from the scale of a few dozen individuals (many meters) to large populations the two factors of habitat heterogeneity and dispersal limitation seem to make assemblages more different from each other than we would expect from random chance alone. These two factors appear to be associated with the middle or flat phase of the triphasic species–area relationship (**Figure 4(b)**).

Biogeographic (Climate and Evolution) Scale

At truly large spatial scales – hundreds to thousands of kilometers – the primary driver of β diversity becomes the geographic ranges of species. Indeed for birds (McGill and Collins, 2003) and trees (McGill, 2011a) it would appear that for any sample bigger than about a $100\ \text{km} \times 100\ \text{km}$ square, all species whose geographic range covers that square will show up somewhere in the square. Thus, the processes controlling geographic ranges become the dominant process. In this case turnover is caused by moving out of the geographic range of some species and into the geographic range of other species. The β diversity between two assemblages A and B is driven by which geographic ranges assemblage A fall into and which geographic ranges assemblage B falls into. Our current understanding of species ranges suggests that two main forces control them. One is climate – many boundaries of many species ranges are set by fairly clear climatic constraints. These may include obvious factors like freezing temperatures or total rainfall but can also include more subtle factors such as the seasonality of precipitation or on very specific details such as in the case of the saguaro where below freezing temperatures last for longer than 12 h. It also seems that deep-time evolutionary processes are important. The act of speciation creates new species. Presumably evolutionary forces drive the wide variation in size of small to large-ranged species. And vicariance events – the dispersal of a species from one continent or biogeographic province to another (a form of dispersal limitation but on a rather different temporal and spatial scale) clearly play a role. Beyond these generalities, our understanding of what creates and maintains species ranges is fairly weak and needs more development. To look at β diversity that uses abundance (e.g., Bray–Curtis index) rather than presence/absence data (e.g., Sorenson), one can replace geographic ranges with the idea of abundance surfaces (**Figure 5(a)**) (McGill and Collins, 2003). Abundance surfaces are effectively 3-dimensional geographic ranges with the third dimension (height) representing the abundance of a species at each point across its geographic range. Then abundance-based β diversity is driven by processes that affect species ranges and the abundance surfaces on them. Even less is known about what factors drive abundance surfaces than we know about what drives geographic ranges. But broadly the main drivers are likely climate and evolution, and these processes are the main drivers in the third, steep phase of the triphasic species–area relationship (**Figure 4(b)**) and possibly some of the second flat phase as well.

Variation in an Assemblage Over Time (Dynamics)

Just as one can compare two assemblages at different points in space, one can also compare the assemblages from one spatial location at two different points in time. And one finds that the

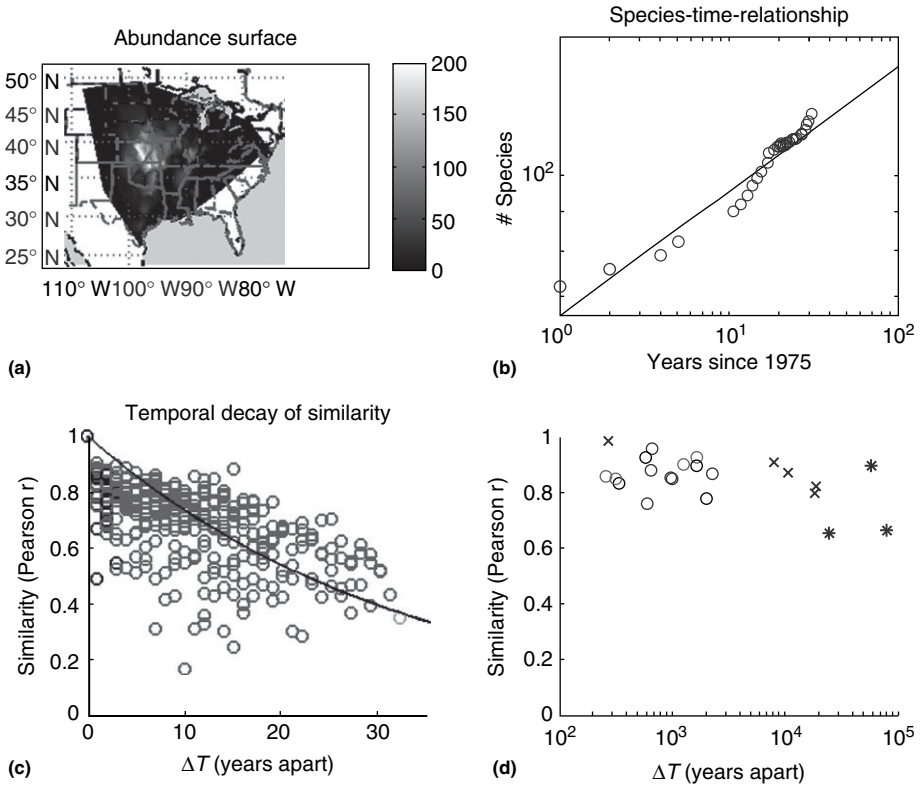


Figure 5 (a) An abundance surface for the Dickcissel derived from the North American Breeding Bird Survey data. The remaining figures look at β diversity in time instead of space. Similar plots can be used. (b) A species–time relationship (same as species–area relationship with time substituted for space) can be plotted. Compare with **Figure 4(a)** (a species–area relationship). The data source is the same as **Figure 2(c)** covering the years 1975–2007. (c) The same data as (b) but plotted as a decay of similarity with distance over time. Note that unlike space, the exponential decay curve is not a good fit – similarity drops too fast in short time periods (representing noise between years) while in long time periods similarity has not decayed as fast as exponentially, representing some long-term stability once the short-term noise has been accounted for. (d) Another decay of similarity with time but over a much longer time span requiring a log-scale on the x-axis. Data are of small mammal bones found in wood rat middens.

species composition found at a fixed spatial location changes over time with species appearing and species disappearing and the relative abundance of species changing. There is an increasing awareness that the patterns that emerge from such comparisons are very similar to the patterns found in space and that many of the same processes are involved. Thus, increasingly ecologists are talking about β diversity in time and using the same tools (species–time relationships instead of species–area relationships and decay of similarity with time curves instead of decay of similarity with distance) (**Figure 5(b)–5(d)**). There is one fundamental distinction, which is that if the time intervals are short enough, the individuals being counted in the two assemblages may in fact overlap (individual X is counted in the assemblage at time A and the assemblage at time B). At longer time scales (i.e., the time of a couple of generations) the same individuals will not be counted, but the offspring of the original individuals will be counted. Thus, temporal β diversity tends to have an inherent mechanism that ensures that change in species composition shows more inertia. But beyond this many of the same mechanisms including sampling, dispersal, habitat preferences, climate, and evolution can play similar roles as the time scales increase.

Species–Time Curves and Temporal β Diversity

As with spatial β diversity, the same three different perspectives (species–time, temporal scale partitioning, and decay of similarity with time curves) can be used to describe what is the same basic underlying phenomenon of change in species composition over time. The first tool used was the species–time relationship (**Figure 5(b)**), but all three methods have been used. The species–area–time–curve allows one to find an equivalence between space and time. For example, Alder and colleagues found that an equivalent amount of turnover occurs in 1 year or 50 m^2 for Kansas plants, 1 year or 1 ha for small mammals, and in the very long dispersing short lived plankton, 5000 km^2 is needed to match the new species found in 1 year (Adler *et al.*, 2005). Decay of similarity with time curves (**Figure 5(c)** and **5(d)**) seem similar to the equivalent spatial patterns but with a subtle difference – there seems to be a faster than exponential drop in similarity in very short time frames. This suggests a system with an innate level of noise (causing rapid drops in short time scales). As with spatial β diversity, the causes and dominant processes change with scale. The author examines four temporal scales here.

Sampling and Detection Scale

At the smallest temporal scales (less than one generation), the assemblage being sampled at time *A* and time *B* is essentially the same assemblage – any differences are due to sampling and detection errors of the survey methods. This is not a particularly biologically interesting scale or process, but it does present a baseline for the amount of turnover due to sampling noise that can then be used to demonstrate that larger amounts of turnover at longer time scales must be due to more interesting biological processes. Obviously the amount of noise at this scale depends on the organism and the survey methods. Stem-mapped tree plots will have almost no noise and turnover at this scale, while birds identified by sight and sound in quick point counts could show significant turnover due to sampling and detection errors.

Population Dynamics Scale

At the next largest time scale (one to a few dozen generations), population dynamic processes play an important role. Classic differential equations such as the logistic and Lotka–Volterra equations are used to describe how populations change over time. However, these populations assume a constant environment and a closed system. In reality population dynamics show considerably more fluctuation due to immigration and emigration from the population as well as external forcing factors that change from year to year including the weather, the abundance of competitors and predators, and the amount of food available. As an example, extreme drought or frost years relative to the norm for a given location can cause major diebacks and drastically change the species composition.

Climatic Forcing Scale

At even longer time scales (thousands to hundreds of thousands of years) there can be trends in climate. These trends in climate can drive major changes in species composition as species track their optimal climates across the continent. Using various fossil assemblages such as tree pollen and diatoms preserved at the bottom of lakes, small mammal bones and plant macrofossils (leaves and sticks) preserved in packrat middens and fossil corals scientists have been able to track the changing composition of assemblages over timescales of thousands to hundreds of thousands of years at one location in space. These data sources make it clear that between about 20,000 years ago (the last glacial maximum) and 10,000 years ago (when temperatures stabilized at approximately what they are today) there was very high turnover (change in a majority of the species) at many locations. What is less clear is what happens to an assemblage. Assuming a group of species are together in an assemblage at time *X* at location *A*, will the same assemblage of species be found at some future time *Y* at location *B*, or will the assemblage be broken up? Conventional wisdom, focused on plants suggests that the assemblage will be broken up, causing not only the movement of assemblages but also the reassembly of new assemblages never seen before. This is known as a nonanalog community or assemblage (Williams *et al.*, 2001). However, other studies have shown that assemblages remain largely intact over hundreds of thousands of years (i.e., inertia) (**Figure 5(d)**) (Pandolfi, 2002; McGill *et al.*, 2005). The relative importance of inertia versus nonanalog communities in different communities is an open question.

Evolutionary Scale

Over timescales approaching a few million years, assemblages show change due to macroevolution. Species go extinct and speciation creates new species. Ultimately the species richness of an assemblage is determined by a balance between speciation and extinction events (Rosenzweig, 1975). A number of factors are known to affect speciation and extinction events including the number of currently existing species (a form of density dependence) (Rosenzweig, 1975) and periods of climatic transition. The relative importance of sympatric versus allopatric speciation (and the relative role of geology/topography vs. climate in the latter) is also being actively debated. The evolution of the distribution of traits is also studied. For example, in body size there is good evidence that the mammals have greatly increased the variance (dispersion or width of the distribution) with some debate remaining about whether the mean has increased or remained the same as when the clade started (Smith *et al.*, 2004). Body size has also changed in response to temperature (increasing in cold temperatures and decreasing in warm temperatures).

Assemblage Properties and Global Change

The sections Mean Descriptions of an Assemblage, Variation Within an Assemblage (Distributions or Histograms), and Covariation (Correlation) Within an Assemblage, have focused on basic research questions about assemblages: what patterns are there and what processes underlie these patterns? However, this knowledge about assemblages is increasingly being applied to the modern human-caused biodiversity crisis. There are at least three broad ways in which knowledge of assemblages can be useful in this applied context: indicators, filling knowledge gaps, and baselines.

Indicators

Extreme human impacts such as clear cutting, building roads, or plowing fields have rather obvious impacts on the natural ecosystem, allowing for an informed debate about the pros and cons of such development in a particular situation. However, many development and management practices are much more subtle. Such practices include exurbanization (low-density housing development outside the suburbs), selective logging, roads in wild areas, etc. The ecological consequences of such actions can often be subtle and even counterintuitive. For example, clear cutting that leaves behind patches of original forest habitat can actually increase species diversity in the retained patches in some cases and for some taxa, but this is mostly because it mixes forest with edge

species. More nuanced measures of the impact of ecosystems are clearly needed. Advanced multivariate statistics are clearly one approach. But such approaches have limited available expertise within the scientific community and great difficulty in translation to managers and policy makers. The tools of describing assemblages given in this entry can provide a useful toolkit for measuring changes.

Figure 6 provides two examples of using assemblage descriptions as indicators as applied to a breeding bird survey route in late secondary forest. **Figure 6(a)** builds on the techniques from the section Mean Descriptions of an Assemblage, and also from a recent study, (McGill, 2011b) which showed that using three distinct univariate measures, one to represent total assemblage abundance, one to represent total assemblage diversity and one to represent assemblage evenness provided a good description of assemblages and separated multiple assemblages well in ordinations. Specific measures depend on sampling methodologies and sizes. **Figure 6(b)** builds on the techniques of the section Variation Within an Assemblage (Distributions or Histograms) by using a particular way of plotting species abundance distributions to control for total assemblage abundance and richness and which highlights the proportion of rare species in the assemblage. Both figures suggest that this assemblage initially improved and then leveled off to a noisy equilibrium by several key indicators of healthy ecosystems (higher richness, abundance, evenness, and proportions of rare species). Other aspects of assemblages can be tracked as well. For example, studies of fisheries show that the average mass of individuals being caught and the average trophic level (as measured by isotope analysis) have decreased over recent decades. The relative proportions of different size classes (basically the frequency distribution of body size) can be an important indicator of successional stage in forests and an indicator of ecosystem health. Some studies have also shown that α diversity may increase or be unchanged under human disturbances, but the β diversity may drop considerably – in short humans are homogenizing the landscape, which will reduce overall (γ) diversity even if simplistic α-diversity measures fail to show this.

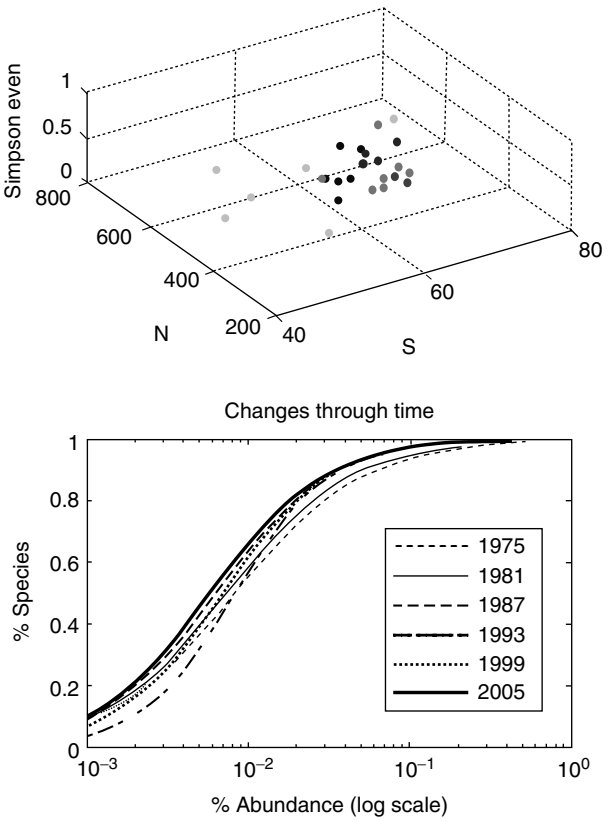


Figure 6 Two examples of how our understanding of assemblages can be used to help in assessing human-caused impacts on assemblages. Both datasets are for birds (see **Figure 5(b)**). (a) Each point represents one year (1975–2007) with darker points representing later years. A simple ordination is performed by plotting these points vs. species richness, total abundance and evenness (McGill, 2011b). Overall the system has moved from low richness and high abundance to the opposite (high richness and low abundance), then returned to settle around a region intermediate in both abundance and richness. Evenness has also increased. This suggests that this one site near Orono ME has actually shown improvement especially from the 1970s to the mid 1980s, possibly with some overshoot but appears stable at the present time. (b) The same data plotted as species abundance distributions. This plot is a new way of plotting species abundance distributions (McGill *et al.*, 2007). The *x*-axis shows abundance as a percentage of all individuals in the assemblage (i.e., relative abundance) on a log scale. The y-axis shows the percentage of species with an abundance equal or less than that shown by where a horizontal line intersects the curve. Thus, for the dark 2005 line, 80% of species have an abundance of 2% or less. As a result rare species appear on the left side and common (high relative abundance) species appear on the right, making it easy to determine the relative proportions of rare species (higher line on the left side). This graph allows for easy comparison between assemblage with different numbers of species or individuals. This data suggest that the proportion of rare species was lowest in 1975 and 1981, generally increased, with an extreme deviation in 1993 and is the highest it has ever been in the last year of 2005. Again this is suggestive of an assemblage that has become less detrimentally impacted by humans in the past 35 years.

Knowledge Gaps

Another use of our knowledge of assemblages is to fill in sparse information about endangered species. A few charismatic endangered species such as the grizzly bear or panda bear are extremely well studied but most endangered species including not just plants and insects but even vertebrates like birds in the tropics are poorly studied. Conservation action plans may be needed even before studies on the many thousands of endangered species can be performed. Our expectations about assemblages can be used to provide approximate but very quick answers to fill in these knowledge gaps. One good example is the link between body size and many other factors such as maximum rate of population increase (r) or offspring size or clutch/litter size (mediated by the link between body size and metabolic rate (*see* Covariation (Correlation) Within an Assemblage)). Two papers use such relationships to expand our knowledge about poorly known dinosaurs (Farlow, 1976) and endangered tropical birds (Pereira *et al.*, 2004).

Species–area relationships provide a quick tool to estimate the approximate effects of reducing the amount of available

habitat (if a species area curve is calibrated for an area and ecosystem, then the number of species can be read off for the projected new, smaller area). Traditional reserve design (deciding which areas of land to set aside for conservation purposes) has focused heavily on α diversity by building parks in the areas of highest α diversity (so-called hot spots of diversity). However, assessing trade-offs such as more small reserves versus fewer large reserves depends on understanding the β diversity in the system. Understanding the factors that drive β diversity (habitat variability and climate variability) can help us to predict the likely β diversity for many more taxa than what we could feasibly measure β diversity for.

Historical Range of Variability

One of the major shifts in our understanding of ecological systems in recent decades has been the growing appreciation of the continuous presence of variability. Stable unchanging systems do not occur naturally. Thus, the mere act of observing that an assemblage has changed neither proves that humans have caused the variation nor does it tell us whether the change is large. Knowing what the historical range of variability is (i.e., how much temporal β diversity there is) helps us answer these questions. It gives us an immediate scale for how much change is "normal". This especially holds true if we have records of the system from before humans were impacting the system, either from paleontological data or from long-term record keeping starting when human impacts were low. As one example, we know that the amount of global warming expected in the next couple of centuries is on the same order of magnitude as the changes from the last glacial maximum to present conditions. This tells us that human-induced climate change is large relative to historical levels of variation. However, the fossil record also tells us that in many parts of the world extinctions in response to this climate change were almost nonexistent. In North America, no small mammals went extinct and only one species of tree went extinct over two million years of over a dozen glacial cycles. This suggests that climate change alone is survivable by ecosystems (although unfortunately it also tells us that tracking climate change was the main survival mechanism, and tracking climate may be significantly harder for species to do in the modern fragmented landscape humans have created). In other systems we have so far exceeded natural historic variability (e.g., increases in CO_2 in the air relative to historic levels in the last several thousand years) that there can only be something unique in history and likely human caused going on. Knowing the historical range of natural variability can be an important step in understanding the role of humans.

Conclusions

The study of assemblages of species is one of the basic areas of study in ecology. There has been an explosion of research on assemblages in recent decades. We are increasingly able to summarize and simplify our descriptions of assemblages (*see* Mean Descriptions of an Assemblage *and* Variation Within an Assemblage (Distributions or Histograms)), understand the patterns of structure within an assemblage (*see* Covariation (Correlation) Within an Assemblage), and understand how assemblages change over space and time (*see* Variation Between Assemblages Across Space (β diversity) *and* Variation in an Assemblage Over Time (Dynamics)). The study of assemblages has led the way in helping ecologists to understand the important influence of scale. And increasingly our understanding of assemblages is providing tools to help humans to measure and understand the consequences of human-caused global change.

Appendix

List of Courses

1. Ecology for Majors
2. Macroecology
3. Biogeography

See also: Defining, Measuring, and Partitioning Species Diversity. Latitudinal and Elevational Range Shifts Under Contemporary Climate Change. Latitudinal Gradients of Biodiversity. Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data

References

Adler PB, White EP, Lauenroth WK, Kaufman DM, Rassweiler A, and Rusak JA (2005) Evidence for a general species–time–area relationship. *Ecology* 86: 2032–2039.

Anderson MJ, Crist TO, Chase JM, *et al.* (2011) Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters* 14: 19–28.

Banavar JR, Moses ME, Brown JH, *et al.* (2010) A general basis for quarter-power scaling in animals. *Proceedings of National Academy of Sciences* 107: 15816.

Brown JH (1995) *Macroecology.* Chicago: University of Chicago Press.

Brown JH and Maurer BA (1987) Evolution of species assemblages: Effects of energetic constraints and species dynamics on the diversification of the North American avifauna. *The American Naturalist* 130: 1–17.

Buckley LB and Jetz W (2008) Linking global turnover of species and environments. *Proceedings of National Academy of Sciences* 105: 17836.

Calder WAI (1984) *Size, Function, and Life History.* Mineola, NY: Dover.

Condit R, Ashton PS, Baker P, *et al.* (2000) Spatial patterns in the distribution of tropical tree species. *Science* 288: 1414–1418.

Conlisk E, Conlisk J, Enquist BJ, Thompson J, and Harte J (2009) Improved abundance prediction from presence–absence data. *Global Ecology and Biogeography* 18: 1–10.

Cyr H, Downing JA, and Peters RH (1997) Density–body size relationships in local aquatic communities. *Oikos* 79: 333–346.

Damuth J (1987) Interspecific allometry of population density in mammals and other animals; the independence of body mass and population energy use. *Biological Journal of the Linnean Society* 31: 193–246.

Darveau CA, Suarez RK, Andrews RD, and Hochachka PW (2002) Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* 417: 166–170.

DeLong JP, Okie JG, Moses ME, Sibly RM, and Brown JH (2010) Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. *Proceedings of National Academy of Sciences* 107: 12941.

Enquist BJ, Brown JH, and West GB (1998) Allometric scaling of plant energetics and population density. *Nature* 395: 163–165.

Ernest SKM (2003) Life history characteristics of placental nonvolant mammals. *Ecology* 84: 3402.
Ernest SKM, Valone TJ, and Brown JH (2009) Long-term monitoring and experimental manipulation of a Chihuahuan desert ecosystem near Portal, Arizona, USA. *Ecology* 90: 1708.
Farlow JO (1976) A consideration of the trophic dynamics of a Late Cretaceous large-dinosaur community (Oldman Formation). *Ecology* 841–857.
Fauth JE, Bernardo J, Camara M, Resetarits Jr. WJ, Buskirk JV, and McCollum SA (1996) Simplifying the jargon of community ecology: A conceptual approach. *The American Naturalist* 147: 282–286.
Gotelli NJ and Colwell RK (2011) Estimating species richness. In: Magurran AE and McGill BJ (eds.) *Biological Diversity: Frontiers in Measurement and Assessment*, pp. 39–54. Oxford: Oxford University Press.
Green JL and Plotkin JB (2007) A statistical theory for sampling species abundances. *Ecology Letters* 10: 1037–1045.
Kleiber M (1947) Body size and metabolic rate. *Physiological Reviews* 27: 511.
Koleff P, Gaston KJ, and Lennon JJ (2003) Measuring beta diversity for presence–absence data. *Journal of Animal Ecology* 72: 367–382.
Lennon JJ, Koleff P, Greenwood JJD, and Gaston KJ (2001) The geographical structure of British bird distributions: Diversity, spatial turnover and scale. *Journal of Animal Ecology* 70: 966–979.
Limpert E, Stahel WA, and Abbt M (2001) Log-normal distributions across the sciences: Keys and clues. *BioScience* 51: 341–352.
Maurer BA and McGill BJ (2011) Measurement of species diversity. In: Magurran AE and McGill BJ (eds.) *Biological Diversity: Frontiers in Measurement and Assessment*, pp. 55–65. Oxford: Oxford University Press.
May RM (1988) How many species are there on earth? *Science* 241: 1441–1448.
McGill B and Collins C (2003) A unified theory for macroecology based on spatial patterns of abundance. *Evolutionary Ecology Research* 5: 469–492.
McGill BJ (2003) Does mother nature really prefer rare species or are log-left-skewed SADs a sampling artefact? *Ecology Letters* 6: 766–773.
McGill BJ (2006) A Renaissance in the study of abundance. *Science* 314: 770–771.
McGill BJ (2008) Exploring predictions of abundance from body mass using hierarchical comparative approaches. *The American Naturalist* 172: 88–101.
McGill BJ (2011a) Linking biodiversity patterns by autocorrelated random sampling. *American Journal of Botany* 98: 481–502.
McGill BJ (2011b) Species abundance distributions. In: Magurran AE and McGill BJ (eds.) *Biological Diversity: Frontiers in Measurement and Assessment*, pp. 105–122. Oxford: Oxford University Press.
McGill BJ, Etienne RS, Gray JS, *et al.* (2007) Species abundance distributions: Moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* 10: 995–1015.
McGill BJ, Hadly EA, and Maurer BA (2005) Community inertia of Quaternary small mammal assemblages in North America. *Proceedings of National Academy of Sciences.* 102: 16701–16706.
McNab BK (1986) The influence of food habits on the energetics of eutherian mammals. *Ecological Monographs* 1–19.
Messier J, McGill BJ, and Lechowicz MJ (2010) How do traits vary across ecological scales? A case for trait–based ecology. *Ecology Letters* 13: 838–848.
Muller–Landau HC, Condit RS, Chave J, *et al.* (2006) Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecology Letters* 9: 575–588.
Murray BR, Thrall PH, Gill AM, and Nicotra AB (2002) How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. *Austral Ecology* 27: 291–310.
Pandolfi J (2002) Coral community dynamics at multiple scales. *Coral Reefs* 21: 13–23.
Pereira HM, Daily GC, and Roughgarden J (2004) A framework for assessing the relative vulnerability of species to land-use change. *Ecological Applications* 14: 730–742.
Preston FW (1948) The commonness and rarity of species. *Ecology* 29: 254–283.
Preston FW (1960) Time and space and the variation of species. *Ecology* 41: 611–627.
Preston FW (1962) Canonical distribution of commonness and rarity.1. *Ecology* 43: 185.
Raunkiaer C (1909) Formationsundersogelse og formationsstatistik. *Botanisk Tidsskrift* 30: 20–132.
Rosenzweig ML (1975) On continental steady states of species diversity. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*, pp. 121–141. Cambridge, MA: Belknap Press of Harvard University Press.
Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.
Sauer JR, Hines JE, Gough G, Thomas I, and Peterjohn BG (1997) *The North American Breeding Bird Survey Results and Analysis*. Laurel, MD: USGS Patuxent Wildlife Research Center.
Sibley DA (2000) *National Audobon Society The Sibley Guide to Birds*. New York: Alfred A. Knopf.
Smith FA, Brown JH, Haskell JP, *et al.* (2004) Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *The American Naturalist* 163: 672–691.
West GB, Brown JH, and Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. *Science* 276: 122–126.
White EP, Ernest SKM, Kerkhoff AJ, and Enquist BJ (2007) Relationships between body size and abundance in ecology. *Trends in Ecology and Evolution* 22: 323–330.
Whittaker RH (1975) *Communities and Ecosystems*, 2nd edn. New York: MacMillan Publishers.
Williams JW, Shuman BN, and Webb T (2001) Dissimilarity analyses of late-Quaternary vegetation and climate in eastern North America. *Ecology* 82: 3346–3362.
Wright IJ, Reich PB, Westoby M, *et al.* (2004) The worldwide leaf economics spectrum. *Nature* 428: 821–827.

Species Coexistence

Robert D Holt, University of Florida, Gainesville, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Coexistence The state of two or more species being found in the same place at the same time.
Community The assemblage of species found in a defined area, in which these species can interact.
Competition The most widely used definition: use or defense of a resource by an individual that reduces resource availability to other individuals. Alternative definition: a reduction in one species' growth rate because of the effects of another species.
Food web The pattern of feeding relationships among organisms.
Lotka–Volterra competition model A model describing the per-capita growth rates of two competing species as linear, declining functions of the abundances of each species.
Mechanistic models of population dynamics Models that are explicit about resource consumption, the relationship of consumption with demography, and mortality factors.
Microcosm studies Laboratory experiments involving population dynamics of small organisms with short generation lengths.
Natural enemy Any species that consumes or parasitizes another species, a general term that includes predators, herbivores, parasites, pathogens, and parasitoids.
Niche (Grinnellian/Hutchinsonian) The range of resources and conditions within which populations of a species are expected to persist. (Eltonian): The role of a species in a community.
Open communities Communities that receive immigrants from external sources and may export emigrants as well.
Permanence A property of communities that ensures the long-term coexistence of species because community trajectories (variation in numbers through time) have no species approaching very low numbers.
Zero-growth isoclines In a graph with axes describing factors important to population growth (e.g., resource abundance), a line along which a population's growth rate is zero.

Introduction

The Oxford English Dictionary states that to "coexist" is "to live together in the same place, at the same time, with another." The topic of "coexistence" focuses on how biological species are organized into communities, in space, and through time. Identifying the factors that influence coexistence is fundamental to understanding biodiversity. Important factors that influence coexistence include interspecific interactions, spatial and temporal scales, and historical contingencies. Insights into coexistence come from laboratory and field experiments, historical reconstructions and observational studies, and theoretical explorations. The topic of coexistence has had a long history in community ecology, some appreciation of which is necessary to understand contemporary perspectives. After a few general remarks about how scale, interactions, and the notion of permanence are related to coexistence, the author provides a survey of the "phenomenology" of coexistence and exclusion, and the lessons derived from field and classic lab studies. The author then discusses key ingredients in both historical and current theoretical interpretations of species coexistence, emphasizing conceptual generalities rather than model details, and concludes with suggestions regarding significant unanswered questions about coexistence.

General Issues

The Importance of Scale

In some sense, all the species in the teeming plenitude of life present on Earth coexist because they are found in one place (Earth) at one time (now). However, from the perspective that matters most to an individual organism such as a vampire bat, a tulip, or a flatfish – the perspective defined by the space that an individual occupies or moves within over its life span – populations of any given species co-occur with only a tiny fraction of all living things. The same holds for community samples, for instance those gathered in the quadrats beloved of plant ecologists or the towed plankton nets of limnologists; the number of species recorded is always much fewer than the potential number. That is, what counts as "coexistence" varies with the scale of ecological inquiry. Viewed at small spatial scales, tailored to the ecology of individual organisms, or seen through the lenses of conventional field methods, most species in the global flora and fauna do not coexist. However, even at this scale, many (and sometimes very many) species can coexist. Species coexistence is also temporally bounded. In the long term, most species face extinction, and even in the short term, coexistence may be transient rather than permanent.

The Importance of Interactions

Species (like ships passing in the night) may either coexist or fail to do so for reasons having nothing to do with each other. A null model of community organization is that communities arise from the independent responses of species to the environment. However, coexistence often reflects interactions among species. Sometimes, species may interact in such a way that species identity, *per se*, does not really matter. In recent years, a neutral theory of community organization has been developed by scientists such as Steve Hubbell and Graham

Bell, which assumes that all individuals within broadly defined groups (e.g., trees competing for space in a tropical forest) have equal probabilities of giving birth, dying, and dispersing to new locales, irrespective of taxonomic identity. This assumption seems biologically implausible, yet the neutral theory performs surprisingly well in predicting many patterns in species-rich communities (Rosindell *et al.*, 2011). Neutrality does not predict perpetual coexistence, but instead transient coexistence as species drift in abundance, some increasing and others decreasing due to chance. In a closed neutral community, in the long run, just one species persists. Given immigration, species richness can be maintained if external source pools harbor multiple species. Yet at a sufficiently large scale, there can be no immigration, and maintenance of diversity in this theory ultimately requires recurrent speciation. The neutral theory is an important – and still controversial – contribution to community ecology, highlighting the importance of speciation, dispersal, and drift (chance changes in abundance) as potential influences on coexistence and species richness.

Close analyses of particular species, however, often show that one cannot understand distribution and abundance without considering species-specific traits and deterministic processes of interspecific interactions. In the rocky marine intertidal, for instance, over broad regions there are repeatable patterns of zonation, with different species predictably dominating at different tidal zones (e.g., barnacles high up in the splash zone, and mussels crowding the lower depths). Experimental species removals often lead to expansions in the width of the zone occupied by other species, demonstrating the importance of interspecific interactions. The remainder of this article will largely focus on these more deterministic processes that influence species coexistence.

Familiar categories of interactions include competition for resources, natural enemy–victim interactions (a "natural enemy" is any species that consumes a species of interest, such as in predator–prey and host–pathogen interactions), commensalisms and mutualisms (e.g., plant–pollinator relationships), direct mechanisms of interference such as allelopathy (in effect, poisoning), and a broad range of environmental modifications (collectively called "ecological engineering"). In the study of coexistence, competition has received by far the most attention, but there is increasing emphasis as well on how natural enemy–victim interactions and mutualisms influence coexistence. Interactions are not fixed properties of species, or of species pairs or larger ensembles, but can vary in strength and pattern as a function of the physical environment and many contingent details of community structure. If species A affects species B, and species B in turn influences species C, there is an indirect interaction between A and C. Sometimes, this indirect interaction reflects altered abundances. For instance, an effective predator may reduce an herbivore's abundance, relaxing consumption on plants. Other cases are more subtle, involving shifts in species' traits. For example, Os Schmitz has shown that the presence of a spider can lead to altered habitat use by grasshoppers, thereby altering grasshopper impacts on plant communities. A complete understanding of coexistence requires one to consider indirect and direct interactions among multiple species. Interspecific interactions may be strongly asymmetrical, and indeed, interactions frequently cause species exclusion rather than coexistence. Analyzing mechanisms of coexistence and exclusion forms an essential part of the conceptual core of the discipline of community ecology.

Coexistence and Permanence

Coexistence emerges from the assembly of communities by colonization from species pools at larger spatial scales. The paleontological record shows that species' distributions are not stable but move in concert with climate change, fluctuations in sea level, and the stately movement of the continents. Chance vicissitudes of dispersal provide a small trickle of potential colonists "testing the waters" outside their normal range, but some dispersal gaps may be near-impassable. Local coexistence at the very least requires species' geographical ranges to overlap. The success of human-mediated introductions suggests that many species do not coexist together today naturally in local communities primarily because historical factors prevented them from ever encountering one another.

Within a species' geographical range, for it to be a member of a local community, at one time it must have colonized, increasing from low numbers. Given the vagaries of climate and fluctuations in the abundance of other species, resident species also experience times of low abundance and extinction risk. To persist, a species must be able to rebound from these dangerous troughs of low numbers. A robust form of coexistence is for each species to be able to increase when rare. If all trajectories describing fluctuations in a community are bounded away from zero, all species in the community will coexist. This criterion for coexistence is called "permanence" in mathematical ecology. Examining conditions for increase when rare provides a natural protocol for both experimental studies and mathematical models of coexistence. Permanence is always relative to a certain spatial and temporal scale. A given pair of species may not coexist indefinitely in any single site but nonetheless coexist at larger scales across many sites, and in the neutral theory, all coexistence is impermanent.

Constraints on Coexistence: Examples from Introduced Species

In recent times, humans have greatly accelerated movements of species into novel habitats, both as deliberate introductions (e.g., most lowland birds found in Hawaii) and as incidental "hitchhikers" tracking human transport (e.g., many benthic invertebrates in San Francisco Bay are exotics released as byproducts of ship ballast dumps). The fates of introductions suggest that species interactions can constrain coexistence.

Competition

Classical biological control involves deliberately introducing species to control pests. Between 1947 and 1952, the Hawaii Agriculture Department released parasitoid species to control the oriental fruit fly, an economically significant pest. (Parasitoids are insects, e.g., braconid wasps, whose larvae live

within and ultimately kill their hosts, such as caterpillars.) Three wasps in the genus *Opius* were considered to be potential control agents. The first species established, *Opius longicaudatus*, parasitized approximately 20% of the fruit fly hosts. The second species introduced, *Opius vandenboschi*, was more effective, parasitizing about 30%; as *O. vandenboschi* increased, *Opius longicaudatus* decreased toward extinction. The third species, *Opius oophilus*, was even more effective, parasitizing up to 80% of the hosts and in turn replacing *O. vandenboschi*. This system provides an example of competitive exclusion in exploitative competition, where species dependent on a single limiting resource (here, host insects) cannot coexist. Increasing parasitism reduces the availability of unattacked hosts. The competitive dominant in this pattern of competitive displacement is the parasitoid species that persists at the lowest host abundance – an example of a simple rule of dominance in resource competition discussed below. Understanding how multiple species persist on a limited resource base is a perennial theme in the study of coexistence.

Natural Enemies

Very effective control agents can, after introduction, eliminate the target species (and thus themselves) over broad areas. The floating fern *Salvinia molesta* from Brazil escaped from a garden in Sri Lanka and became a serious aquatic pest in much of the Old World wet tropics. The beetle *Cyrtobagous salviniae* was introduced in Australia and proved highly successful at limiting the fern, which (with the beetle) is now found in only a few scattered populations. This example illustrates another species coexistence problem: how do effective natural enemies manage to coexist with their prey without overexploiting them and driving themselves to extinction?

Community and Spatial Contexts

Many game departments routinely introduce game species. Failed introductions reveal constraints on species coexistence. Caribou were once common in the Maritime Provinces but declined during the 1800s to extinction by 1915. Attempted reintroductions failed because the caribou pick up a nematode parasite (carried by snails and incidentally consumed as the caribou forage on vegetation). The nematode, lodging in the caribou brain, is fatal within a few months. The nematode population is sustained by another host, white-tailed deer, which tolerates parasitism. Deer are abundant because land practices by European settlers in the nineteenth century expanded second-growth habitats that deer favor; this may explain the disappearance of caribou. Caribou herds wander very widely, and so readily encounter nematode-laden snails. Another ungulate species, moose, coexists with deer, even though infected moose also quickly die. Moose and deer overlap in diet and can compete; in contrast, caribou and deer do not overlap in diet. Nonetheless, it is caribou, not moose, which have been indirectly excluded by deer. Moose are relatively sedentary (for their body size) and can occupy patches of highland forests with deep winter snows, where deer do not penetrate. At the coarse scale of North America, caribou and white-tailed deer coexist but with little spatial overlap among their ranges. Moose and deer, in contrast, overlap geographically, coexisting at the level of landscapes, but at a finer scale, they show spatial segregation in habitat use.

This example illustrates several points. First, natural enemies can prevent coexistence. Second, coexistence or exclusion may arise from complex webs of multispecies interactions. The nematode directly causes exclusion of caribou, but viewed more expansively, exclusion is caused by the entire vegetation–snail–deer ensemble, which collectively governs nematode abundance and hence caribou infection rates. Third, movement patterns can influence interspecific interactions. Moose and white-tailed deer coexist at the landscape scale because moose have spatial refuges from infection, whereas caribou do not and are thus excluded.

Coexistence and Exclusion: Messages from Bottles

The previous examples strongly suggest the importance of species interactions in determining coexistence but are not conclusive. Without detailed observational studies tied to parameterized models or rigorous field experiments, it is difficult to persuade skeptics that other hypotheses are not also plausible. As an alternative, one simple but illuminating approach to the study of coexistence is to put together a few species in a confined setting and see what happens.

The study of "nature in a bottle" simplifies the world in important respects (e.g., such experiments typically focus on a few species in closed and temporally constant environments). For this reason, such experiments are viewed skeptically by some ecologists. Despite limitations, a wealth of important messages about species coexistence have emerged from bottle experiments – messages that have proven robust when applied to natural communities. Microcosm studies of ecological processes are now enjoying a renaissance in community ecology.

Competition Experiments

The term "competition" usually refers to an interaction among individuals (within or among species) that arises because they seek a resource in short supply. If this involves direct harm, the interaction is referred to as "interference competition." If instead competition involves depletion of a resource, one refers to "exploitative competition."

Protozoa

Most ecology textbooks (Hutchinson, 1978) recount famous experiments by the Russian ecologist G. F. Gause, who, as a young man in the 1930s, put mixed cultures of protozoa into vials to study species coexistence. Gause's experiments compared populations of the ciliate protozoans *Paramecium aurelia* and *Paramecium caudatum* grown separately, and together, on a nutritive medium containing their essential resource (bacterial food). Both species thrived when alone, but *P. aurelia* usually displaced its congener in joint cultures within 30–50 generations. This outcome was reversed if the medium was completely replenished with fresh nutrient on a regular basis. Gause argued that metabolic byproducts were building up in the experiments, and that part of the dominance of *P. aurelia* involves

its resistance to the chemical byproducts of metabolic activity as well as its superior ability to exploit the food base. Here, competition combines both environmental modification by a species and exploitation of a limiting food resource.

In other experiments, Gause found that *P. aurelia* could coexist with another species, *Paramecium bursaria*, even in the confines of a closed culture. *P. bursaria* contains symbiotic algae, which release oxygen in photosynthesis. In incompletely mixed cultures, bacteria accumulate on the bottom, creating a zone slightly depleted of oxygen. The protozoan with the algae in effect carries its own oxygen supply into this hypoxic habitat and so can use a food source unavailable to the other, competitively superior species. Here, coexistence depends on both the availability of different habitats and differential species' abilities to utilize those habitats.

Beetles

Many ecologists have examined the dynamics of mixtures of beetle species competing for grain in containers. L.C. Birch examined several species combinations that fed on wheat or maize under different conditions of temperature and moisture. Usually, one species won. The criterion for dominance was simple: The winner always had the highest intrinsic growth rate (the maximal rate of population growth when a species is rare and growing alone). However, the identity of the winner depended on both the resource type and the physical conditions. Moreover, competitive exclusion sometimes required many generations; in one experiment, the apparent loser survived until the experiment had to be terminated.

Experiments by T. Park with a different set of stored grain pests (*Tribolium castaneum* and *Tribolium confusum*) revealed a different possibility – for some abiotic conditions, the outcome was indeterminate. Either species could win, with the winner tending to be the species initially most abundant. The net interaction between these beetles combines several mechanisms, including exploitation of a shared resource and cannibalism (both within and between species). In one experiment, Park intriguingly observed that when a coccidian parasite, *Adelina* sp., was present, the normal dominant *T. castaneum* became a weak competitor, permitting the persistence and even the dominance of *T. confusum*.

Sometimes, coexistence in microcosms reflects very subtle differences between species. Crombie raised two grain beetles (*Rhyzopertha dominica* and *Oryzaephilus surinamensis*) in vats of cracked wheat and found indefinite persistence, despite the fact that there seemed to be just a single resource. The main difference in the two species seemed to be in larval feeding habits; larvae of *Oryzaephilus* lived and fed outside the wheat grains, whereas larvae of *Rhyzopertha* were sufficiently small to live and feed from within the grains. This slight difference sufficed for the competitors to coexist.

Key Insights

Several insights emerged from "bottle" studies of competition. First, species competing in confined quarters in a homogeneous medium for a single limiting resource typically do not coexist. Second, the winner depends on the conditions, and the outcome depends on the environment – there is no universally superior competitor. Third, exclusion is not instantaneous and may require many generations; this leads to the possibility of transient coexistence. Fourth, indefinite coexistence occurs but requires heterogeneity in the environment as well as differences in species' responses to this heterogeneity. However, species' differences permitting coexistence can be quite subtle. Fifth, pairs of species can interact in many distinct mechanistic ways (e.g., via impacts on resource levels or on levels of pollution from metabolic waste). It is difficult to generalize among studies if one has not clearly identified and characterized specific mechanisms of interaction. Finally, the outcome of competition can vary due to effects of other species, including parasites and mutualists.

Predation

In the classic literature, compared with studies of competition, less attention was paid to predator–prey and host–parasite interactions (and essentially none to mutualism). Gause did carry out experiments with the predatory protozoan *Didinium nasutum* and a prey protozoan *P. caudatum*. The predator quickly overexploited its prey and became extinct. To achieve coexistence, Gause concluded that there needed to be a spatial refuge for prey, inaccessible to the predator, or recurrent immigration. Luckinbill reexamined the unstable predator–prey pair studied by Gause but reduced the encounter rate between predators and prey by adding methyl cellulose (a thickener) to the medium so as to slow predator movement. Reducing movement rates in effect expands the size of the spatial arena of the interaction (scaled by the distance a predator moves per unit time). This led to stable coexistence. The qualitative message of microcosm experiments on predation is that patchiness, localized dispersal, and spatial heterogeneity (e.g., refuges) may facilitate the coexistence of effective specialist predators and their vulnerable prey.

Constraints on Coexistence: Rules of Dominance

The Competitive Exclusion Principle

The experimental observation that in homogeneous well-mixed lab environments it was difficult to achieve coexistence between similar species became enshrined in ecology as Gause's principle or the "competitive exclusion principle." Another way to state this principle is to note that, to coexist indefinitely, different species must have distinct ecologies.

Assume we are examining a local community defined by a spatial scale in which all individuals can reach all sites over a single generation. The basic logic of the competitive exclusion principle is impeccable (Levin, 1970). Consider two species with continuous generations, where both species respond to the same environmental factors, denoted by E. The quantity E could be many things (e.g., resource availability, predator abundance, or the competitors' own summed numbers). We make three assumptions: (i) There is a single limiting factor, (ii) the species interact in a closed habitat (i.e., no immigration), and (iii) each species when alone settles down to an equilibrium with constant densities (i.e., the environment is temporally constant). The growth rate of species i is $dN_i/dt = N_i f_i(E)$, where N_i is its density. To complete the

model, one must also characterize the effect of the species themselves on the magnitude of E, leading to either direct or indirect density dependence in demographic parameters such as birth or death rates. For example, if two species are consuming a single resource, consumption depresses resource levels, reducing growth rates.

For species i, there will be some value of the environmental factor, E^*, at which that species equilibrates. It is a biological truism that any two species will almost surely differ in some way. Hence, it is very improbable that they will have exactly the same values for E^*. In other words, there should be no long-term persistence of two species limited by a single factor in a constant, closed environment

Exploitative Competition: The R Rule*

What counts as a "limiting factor" needs to be interpreted quite broadly and can be difficult to identify in practice. Nonetheless, sometimes a single, simply characterized limiting factor defines a rule of species' dominance. For instance, the Hawaiian parasitoids fit the "R^* rule" proposed by David Tilman; that is, the dominant species, given exploitative competition for a single resource, is the species persisting at the lowest resource level. (The asterisk denotes equilibrium, and R^* is the resource level at which a given consumer species is in demographic equilibrium.) An advantage of this dominance rule is that one can measure R^* of each species when alone and then predict the outcome of competition. A species may have a low R^* and be competitively superior either because it is efficient at resource consumption or because it has low mortality (e.g., it can escape predation).

The R^* rule successfully predicts the outcome of competition among phytoplankton competing for single nutrients in microcosms and also characterizes competition in some natural systems. For example, in the nitrogen-poor soils of Cedar Creek, Minnesota, plant species compete for nitrogen. Wedin and Tilman showed that species with low R^* for nitrogen won in pairwise competition experiments. If species had similar values for R^*, the rate of competitive displacement was greatly reduced, as expected by theory.

Apparent Competition: The P Rule*

Analogous "rules of thumb" arise in other situations. For instance, Sharon Lawler carried out a microcosm experiment in which a predatory protozoan (*Euplotes patella*) coexisted fine with either of two prey protozoans (*Tetrahymena pyriformis* and *Chilomonas paramecium*) grown alone. However, when all three species were placed together, *Chilomonas* was driven extinct. In single-prey cultures, *Tetrahymena* sustained four times the number of predators as did *Chilomonas*. In mixed cultures, the latter species suffered higher predation than it could sustain. The limiting factor here is the abundance of a shared predator, which responds numerically to its prey. Let P^* be the abundance of the predator sustained (and tolerated) by a prey species. A "P^* rule" now describes dominance: The winning prey species is the one sustaining greater P^*. This form of indirect competition between species arising from shared natural enemies (including parasites) is called apparent competition (Holt and Lawton, 1994); the word "apparent" is used because the interaction has comparable consequences for coexistence as does classical exploitative competition for resources but may occur even between species with totally different resource requirements. Dominance in apparent competition may reflect different vulnerabilities to a natural enemy (as in the caribou example) or because one prey is highly productive and sustains an abundant enemy population (as in Lawler's experiment).

Mechanisms of Coexistence

The previous discussion emphasized constraints on species coexistence arising from interspecific interactions. Rules of dominance are important conceptual tools that quantify these constraints and help identify biological traits leading to dominance. However, even in simple microcosms, coexistence can occur, and most natural communities are rich in species. Species in principle may coexist when any of the assumptions leading to the competitive exclusion principle are violated. This suggests three classes of mechanisms promoting species coexistence of potentially competing species in a local community:

1. Species may coexist in a closed, temporally constant world if they experience different limiting factors; this includes classical niche partitioning of resources, as well as mechanisms involving predation and parasitism, and direct interference.
2. Species may coexist, even though they experience the same limiting factor, if the environment is temporally variable and species respond differently to this temporal variation (temporal niche partitioning).
3. Species may coexist if the environment is spatially open or interactions are localized; the implications of space for coexistence can include spatial niche partitioning at scales broader than the local community, mechanisms such as colonization–competition tradeoffs in metapopulations, and microscale habitat partitioning.

From the 1950s to the mid-1970s, stimulated largely by G.E. Hutchinson and his brilliant student Robert MacArthur, most community ecologists emphasized classical niche partitioning in studies of species coexistence. In recent years, the balance of attention has shifted markedly to a broader range of coexistence mechanisms. Ecologists now believe that maintenance of diversity – coexistence writ large – often depends on spatial dynamics in open communities, food web interactions (including predation and parasitism), and nonequilibrial dynamics reflecting either extrinsic temporal variation or the endogenous instability of complex ecological systems. Moreover, as noted above, some ecologists have concluded that in natural communities, many species cooccur, without necessarily permanently coexisting, in the sense of tending to increase when rare.

Traditional Approaches to Coexistence

Classical Niche Partitioning

Competitive exclusion is expected if the growth rates of multiple species are determined by a single limiting factor. Species may coexist, even in an unvarying and spatially confined bottle,

given multiple limiting factors, such that each species is limited more strongly by its own distinct set of factors. When species coexist, one sensible approach to begin to understand this coexistence is to map their niche requirements against the spectrum of limiting factors present in the environment. This basic methodology has been a source of significant insights for much of the history of ecology and continues to be fruitful. The classic paper in this genre is the 1958 study by Robert MacArthur of wood warblers in a New England boreal forest. Five warbler species in the genus *Dendroica* occurred in the same tract of forest; all had similar body sizes and ate the same range of insect taxa. MacArthur found that the species segregated with respect to microhabitat (with one species feeding at tree-top, another in low branches, and so on). He argued that because of this microhabitat partitioning, each species consumed an independent pool of insect prey, thereby reducing the potential for competitive exclusion. This study led to a proliferation of field studies of niche partitioning patterns.

Lotka–Volterra Competition Model

MacArthur's study was motivated in part by Gause's experiments on competitive exclusion, which in turn were stimulated by theoretical models of interacting species explored by the Italian mathematician Vito Volterra. In the usual textbook formulation (the famous Lotka–Volterra equations; see for e.g., Morin, 2011), the competitive effect of one species on another is expressed as a direct, density-dependent reduction in abundance:

$$\frac{dN_1}{dt} = r_1 N_1 (K_1 - N_1 - \alpha_{12} N_2)/K_2 \qquad [1]$$

where N_1 and N_2 are the abundances of competing species 1 and 2, respectively, K_1 is the carrying capacity of species 1, r_1 is its intrinsic rate of increase, and α_{12} is a competition coefficient (the equation for species 2 is the same, with subscripts 1 and 2 switched). The quantity α_{12} measures the effect an individual of species 2 has on reducing the per capita growth rate of species 1 relative to the effect of an individual of 1 on its own species. Both species increase when rare and hence coexist if

$$\frac{1}{\alpha_{21}} > \frac{K_1}{K_2} > \alpha_{12}. \qquad [2]$$

The outer inequalities imply $1 > \alpha_{12}\alpha_{21}$. This necessary condition for coexistence states that at least one competitor experiences stronger intraspecific than interspecific competition. If the inequalities are reversed, either species can exclude the other if it is initially sufficiently abundant (as in Park's experiments with *Tribolium*). With competition coefficients near unity (i.e., density dependence occurs uniformly within and between species), the species with higher carrying capacity wins.

The Lotka–Volterra model usefully describes competition, and multispecies extensions have been the focus of a rich body of theoretical work. However, the model is difficult to use predictively because its parameters must be measured during competition experiments. Moreover, taken literally, the model best describes competition due to direct interference. Recognition of these limitations led to the development of a wide range of mechanistic models of resource–consumer and other interspecific interactions beginning in the late 1960s, and continuing to the present.

Whither Limiting Similarity?

This development was concordant with interest in the important concept of "limiting similarity," which is the notion that there might be a quantifiable limit to how similar species can be in their utilization of resources and still coexist. To gain an understanding of the idea of limiting similarity, note that in expression 2, similar species should have competition coefficients near 1, and so coexistence will not occur unless carrying capacities are finely balanced. For a period in the history of the discipline, the goal of quantifying limiting similarity represented a "holy grail" for community ecology. Early theoretical explorations suggested that simple rules might permit predictions of the maximal number of species that could persist on a defined resource base, providing a basic tool for understanding biodiversity patterns. These studies were based on the Lotka–Volterra model coupled with assumptions about resource use. In some limiting cases, the Lotka–Volterra model emerges as a reasonable approximation of more complex resource–consumer interactions, and one can directly map niche overlap onto measures of competition. With simplifying assumptions about symmetry in resource use, environmental variation, and other factors, a limiting similarity of competing species could be calculated and compared against observed similarities. For instance, using MacArthur's warbler data, it was assumed that spatial overlap in foraging (measuring similarity) was directly related to competition coefficients. Substitution into the model showed that observed overlap was consistent with long-term coexistence. If measures of niche overlap in fact portray the strength of competition, observational studies of overlap and niche partitioning would provide a powerful link between descriptive analyses of community patterns and the dynamical forces of species interactions.

Further exploration has tempered the initial flush of enthusiasm for this approach to coexistence. After considerable grinding of mathematical gears, theoretical ecologists have concluded that limiting similarity is not likely given by a single number that pertains across all systems. Instead of a single general theory of limiting similarity and coexistence, there are many special theories. Moreover, there is a real sense in which increasing species similarity facilitates, rather than hampers, coexistence. Equation [2] describes bounds on the permissible differences in the two species' carrying capacities. If the α_{ij} are less than unity, coexistence is more likely if carrying capacities are nearly equal rather than very different. The more similar species are in their demographic responses to the environment (i.e., their basic fitness), the more similar their carrying capacities are likely to be. All else being equal, this kind of similarity promotes coexistence.

Manipulative Field Experiments

Recognizing the limitation of observational studies of niche partitioning as evidence for competition, ecologists turned to manipulative field experiments, removing species

and monitoring changes in the abundance of others. Reviews of such experiments show that when species are suspected to compete (e.g., because of overlap in resource requirements and habitat), they often do compete. For instance, Hairston noted that in the Great Smoky Mountains, two species of *Plethodon* salamanders had altitudinal ranges that were nearly mutually exclusive, with only a narrow range of overlap. In contrast, in the Balsam Mountains, altitudinal overlap was extensive. Hairston hypothesized that this was due to stronger competition in the Smokies, and using reciprocal transplants and removals, he demonstrated that competition (due to aggression) was indeed much stronger in the narrow overlap zone in the Smokies.

Unfortunately, few field experiments have been directly tied to mechanistic models of competition. Moreover, such experiments tend to focus on species that already coexist, at least to a degree; removals then assess the magnitude of competition, given coexistence, rather than coexistence or exclusion *per se*. One study that directly addressed coexistence was performed by Bengtsson, who studied three species in the zooplankton genus *Daphnia* (which compete exploitatively for algal food) on the coast of the Baltic Sea. Observational studies suggested that usually just one, or more infrequently two, species was present in any given pool. Bengtsson added all possible combinations of the three species to artificial pools and found that there were no extinctions in the single-species pools. Yet high extinction rates occurred in the two- and three-species sets. This study directly demonstrates the importance of competition in constraining coexistence in natural conditions.

Mechanisms for Local Coexistence: Current Approaches

Mechanistic Models of Multiple Limiting Factors

Most theoretical studies of competition today focus on models in which the mechanisms of interaction are clearly described. Detailed analyses of mechanistic models of competition are often mathematically challenging, but important insights can often be gleaned from simple graphical analyses. A generalization of the competitive exclusion principle is that "coexistence of n species requires n limiting factors, and n distinct feedback effects via the environment." Mathematically, one expresses the growth rate of each species as a function of a vector of limiting factors (e.g., $dN_i/dt = N_i f_i(E_1, E_2, \ldots)$), and adds equations for the environmental factors themselves, which in turn depend on N_i. For instance, an environmental factor might be resource availability, which becomes reduced by consumption. For illustrative purposes, the author discusses one example in detail (MacArthur 1972, cited in Chase and Leibold 2003, which also treats many other cases).

Consider two species competing for two resources. Assume that exploitation depresses resources and that resource–consumer dynamics tend toward a stable equilibrium. For each species, there will be some resource combination that allows equilibrium (with births matching deaths). On a graph with axes of resource abundance, assume that this combination is a straight line, as in the line marked 1 in **Figure 1(a)** (for species 1) (MacArthur, 1972). This line is the "zero growth isocline" of species 1. The linear form of the isocline implies that resources are qualitatively substitutable so that a sufficient supply of one compensates for low abundances in the other. At equilibrium, resource abundances lie along this line; if resources lie above this line, species 1 should increase, depressing resource levels (with reverse dynamics below the line). Equilibrial coexistence requires that resource levels be such that both consumers have zero growth; graphically, the isoclines must cross. The isoclines do not intersect (**Figure 1(a)**) if species 1 has a lower R^* for each resource. If species 1 is resident and at equilibrium, and a few individuals of species 2 are introduced, the introduction will fail due to competitive exclusion. If isoclines cross (**Figure 1(b)**), species 1 has a lower R^* for one resource than does species 2, and the converse is true for the other resource. Coexistence is now possible.

Crossing isoclines reflect differences in species' ecologies and the existence of distinct limiting factors (here, linear combinations of resources). However, such niche differences do not suffice for coexistence. As in the Lotka–Volterra model, there must be broad similarities in how the two species

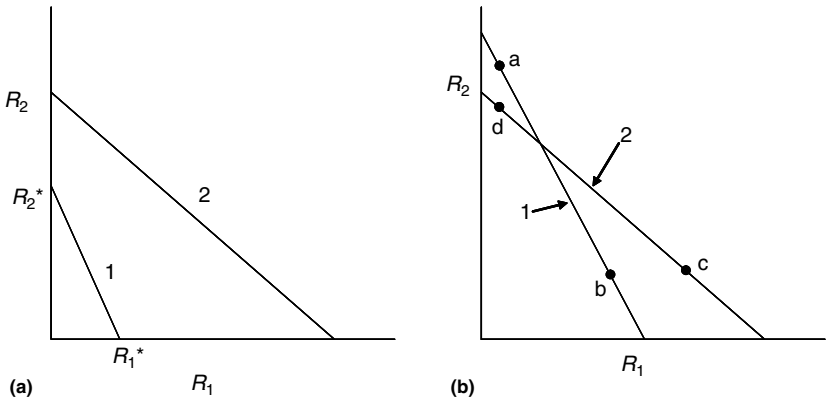


Figure 1 Graphical model of resource competition. The isocline of species i (=1,2) is a combination of resource abundances where it has zero growth. At lower resources (toward the origin), a species declines, (a) Competitive exclusion of species 2 by species 1; (b) Potential coexistence. (See text for details.)

respond to the environment. In **Figure 1(a)**, exclusion occurs because the species are too different in their overall resource requirements. Coexistence, in contrast, is possible if overall resource requirements are approximately the same for the two species (**Figure 1(b)**). Peter Chesson has suggested that one can broadly understand the requirements for coexistence in terms of the interplay of "stabilizing mechanisms" arising from different kinds of niche differentiation, and "equalizing mechanisms", which means that no species has too great a basic fitness advantage.

Having crossing isoclines does not guarantee coexistence. If species 1 is alone, and resources equilibrate at point a (in **Figure 1(b)**), species 2 invades. However, if resources instead equilibrate at point b, species 2 is excluded. If species 2 is resident, species 1 invades if resources are at point c but not at point d. Comparing these points to each species' R^* (for each resource) suggests a necessary requirement for coexistence: Given two resources, each competitor must have the greater impact on that resource for which it has the lower R^*. In effect, each species must limit itself (via resource depletion) more strongly than it limits the other species. Whether or not this occurs depends on both the intrinsic renewal rates of the resources and the rates of resource consumption. A complete analysis of conditions for coexistence requires one to track the dynamics of both consumers and each resource, and goes beyond the scope of graphical models. But this graphical model illustrates the important insight that coexistence depends on a balancing of overall similarities and differences in species' niche requirements. It does not suffice for understanding competitive coexistence, because one must also consider feedback effects mediated through the species' impacts on their environments. This depends not just on species' properties, but the structure of the entire system in which the interaction is embedded, including ecosystem processes (e.g., resource renewal rates).

Food Web Effects on Coexistence

This general approach can be extended in many ways, for instance by including interactions among species across trophic levels. With predation, herbivory, or parasitism inflicted on competing species, one greatly increases the number of potential limiting factors. This is a large and complex topic, and the author briefly discusses some highlights here.

Specialist predators and parasites typically reduce the abundance of their favored species, freeing up resources for other nontarget species. This can facilitate coexistence if dominant competitors tend to attract more specialist predators or parasites than do subordinates. In tropical forests, specialist herbivores and pathogens sustained by adult trees can hamper recruitment by juveniles of the same species, freeing up resources that can be used by other species – an explanation of high tropical tree diversity called the Janzen–Connell effect. Generalist predators can have the same effect via fixed preferences for dominant competitors or if they reduce attacks on whichever species is temporarily rarest. Pathogens can likewise promote coexistence if their impacts are strongest on dominant competitors. For instance, at Silwood Park in the United Kingdom, an experimental manipulation of foliar fungal pathogens using fungicide showed that competitively dominant grasses were reduced by parasitism, freeing resources for other species and thus enhancing diversity.

However, predators or pathogens that attack multiple species can hamper coexistence (as discussed previously). This is particularly likely if the prey species under attack do not strongly compete, and if some prey species are sufficiently productive to sustain high abundances of the predator or parasite (Holt and Lawton, 1994). Coexistence can occur, despite apparent competition under high predation intensity, if different species utilize different refuges from predation – a subtle form of niche partitioning. In general, whenever species interact via multiple mechanisms (providing different limiting factors; e.g., different specialist pathogens on different tree species), coexistence may occur. As another example, if two species compete for a single limiting resource, the species with higher R^* may nevertheless persist if it can also consume the superior competitor. This mixture of predation and competition, called intraguild predation, is particularly common in aquatic food webs, and at higher levels in terrestrial food webs.

Local Habitat Heterogeneity

Traditional ecological models of interacting species such as eqn [1] assume that populations are spatially well mixed and average over local environmental variation. Relaxing this assumption often promotes coexistence, and for several distinct reasons. As shown in Gause's experiments, even in the confines of a microcosm, there can be spatial heterogeneity in abiotic conditions or resource availability that influences coexistence. Assume that within the microcosm each species has a set of conditions in which it is superior. If competitive interactions are sufficiently localized (e.g., because of limited movement or habitat selection), and each species spends more time in the microhabitat where it is superior, one can readily generate coexistence. This simple mechanism for coexistence via local habitat partitioning is very important. A review by Schoener (1974) of resource partitioning studies following in the footsteps of MacArthur's warbler study revealed the ubiquity of habitat differences among potential competitors; subsequent years have not altered this basic message. In animals, habitat selection can be an important mechanism promoting coexistence with local spatial heterogeneity. If individuals of a species (when rare) can discriminate among local microhabitats and spend more time in those that provide the greatest fitness rewards, the rate of increase at low N will be increased, relative to that of a species that utilizes habitats at random. If each species has a habitat in which it is competitively dominant, habitat selection will sharpen habitat partitioning and make coexistence more likely. Habitat selection can permit subordinate species to withstand superior competitors. Recent studies in east Africa suggest that the fleetness of the cheetah helps it persist in the face of direct aggression by lions and hyenas (who kill cheetah cubs) by allowing it to seek out areas with low lion and hyena densities. For many organisms (particularly terrestrial plants), an individual occupies a small site and interacts with only a few

neighbors during its lifetime. Probabilistic events of individual life histories (birth, death, and dispersal) lead to spatial variance in competition, even among sites with homogeneous physical conditions. Moreover, dispersal typically occurs over small spatial scales. Combining these two general facts together in models leads to nonuniform spatial patterns in which competition is typically stronger within species (because they tend to occur in clumps) than between species (which tend to become segregated spatially). This mechanism may help explain the coexistence of large numbers of similar coexisting species in many plant communities.

Temporal Niche Partitioning

Outside the controlled confines of the lab, environments are rarely constant. Temporal variation occurs on scales ranging from diurnal cycles to climatic changes over millennia. Many ecologists have intuitively felt that temporal variability should weaken negative interspecific interactions and facilitate coexistence. Theoretical studies (Chesson, 2000) have conclusively shown that this is not the case. Instead, temporal variation can provide a rich arena for differentiation among species in responses to the environments, which can promote coexistence. The basic idea is that if species A does better when the local environment is in state A′ and species B does better when the local environment is in state B′, and the environment alternates between A′ and B′, coexistence is possible. This can happen in several different ways.

Nonlinear Consumption

Assume two consumer species exploit a single resource and have saturating, nonlinear relationships between feeding rates and resource availability. Species A has a higher rate at low resource levels, whereas species B enjoys a higher feeding rate at high levels. In a constant environment, the resource level will equilibrate at a constant level, and the species with lower R^* wins. However, given large-scale seasonal variations in resource supply rates or mortality, resource levels may fluctuate between high and low levels, permitting competitive coexistence. This mechanism hinges crucially on nonlinearity in species' responses to the shared limiting factor; if feeding rate were to increase linearly with resource availability in both species, then even in a variable environment one would observe competitive exclusion. Studies of phytoplankton by Jef Huisman have shown that given comparable nonlinear consumption rates, competition for multiple essential resources can lead to unstable dynamics, even in constant environments, and this instability in turn permits coexistence of multiple species on just a few abiotic resources.

The Storage Effect

Desert rains fall sporadically and at different times in different years, sometimes early in the spring when the air is cool and at other times later in the hot summer. After rain, plant species emerge from the seed bank and compete for the pulse of water and soil nutrients. These annual plants are short-lived as adults, often completing their entire aboveground life cycle in just a few weeks. Some species do disproportionately better following warm rains, whereas others do better following cold rains. The seeds each species produces enter the seed bank and germinate gradually during the following years (like a time-release capsule of past recruitment). This "storage" of good years of recruitment in a long-lived seed bank can facilitate coexistence. Again, it is not variability alone that allows coexistence, but variability combined with niche partitioning (different species being superior at different times) and the buffering effect of a long-lived seed bank, generating nonlinear responses of species to competition and the environment (Chesson, 2000). The storage effect exemplifies the growing recognition that the internal structure of populations (age, size, the presence of a seed bank, etc.) can have significant consequences for species coexistence.

Coexistence in Open Communities

In contrast to laboratory microcosms, many natural communities are open, coupled to the external environment via dispersal. Such coupling influences the local coexistence of species in many distinct ways.

Autecology and Population Size Effects

A consumer species that relies on a sparse or a sporadic resource, or extracts different essential resources in different habitats, may need to be mobile to persist. Or if absolute population size is small, extinction is risked even in favorable environments. Specialist consumers are likely to become extinct if their resources are rare. All these problems are aggravated, given barriers to dispersal and a small habitable area. These considerations may help explain why food chains are often short and specialist consumers absent on oceanic islands or isolated habitat patches. Moreover, as one increases the number of species coexisting deterministically on a fixed resource base, the abundance per species declines. Small population sizes have an increased risk of extinction; this can put a loose limit on coexistence, aggravated in small and isolated habitat patches.

Colonization–Extinction Dynamics

It was previously mentioned that strong specialist predator–prey interactions are prone to local extinctions. Coexistence may require dispersal among habitats, with prey dispersing sufficiently fast that they can find and reproduce within empty habitat patches before being discovered by predators. Similarly, if a guild of competitors utilizes a single resource, species with successively lower values of R^* should eventually colonize and displace species with higher R^*. If there are extinctions, asynchronous among patches, inferior competitors may have a temporary window of opportunity during which they can occupy habitat patches and reproduce sufficiently to colonize other patches. This metapopulation mechanism for coexistence is promoted by a tradeoff between competitive ability and colonization ability, and in principle, it can permit the coexistence of many competing species

(Tilman and Lehmann as cited in Tilman and Kareiva, 1997). Even if the inferior competitor is not a superior colonizer, it may be able to persist if it has a lower basal extinction rate when it occurs in patches alone compared with that of the superior competitor.

Landscape Heterogeneity

Coexistence Due to Spillover

If a species can persist in one local community, then with emigration, it can also be found in other nearby communities where otherwise it would be excluded. To model this "spillover" effect, assume that the species when rare declines at rate $r<0$ because of competitive exclusion but immigrates from an external source at rate I. Its local dynamics are described by $dN/dt=I+rN$, which implies $N^*=I/|r|$ at equilibrium. A species that should be absent, considering only local interactions, may not just persist but even be abundant if there is a large rate of immigration (e.g., from a productive source habitat into an unproductive habitat), and the rate of exclusion is slow. In this scenario, the answer to the species coexistence problem ultimately requires analyzing population dynamics at the appropriate spatial scale, larger than the local community; spatial niche partitioning at this larger scale is responsible for coexistence.

Landscape Mechanisms of Exclusion

Spatial coupling can also generate novel mechanisms for exclusion. For instance, a species that is a superior competitor in one habitat may be excluded by a species that is inferior there but sufficiently abundant elsewhere so that via dispersal it can "swamp" the local habitat. For instance, Ted Case and associates recently described impacts of an invasive ant species, the Argentine ant, on an entire community of ants in coastal southern California. In contrast to many ants, the Argentine ant shows little intercolonial hostility and so has a high carrying capacity; it competes exploitatively for food but also preys on the juvenile life stages of some ant species. The Argentine ant readily becomes established in disturbed habitats. It can then spread into fragments of natural habitat resulting from intense development. Case suggests that in these fragments, many other ant species have declined to the point of extinction because of the sheer force of numbers. These effects are particularly dramatic in small fragments, and at the edges of large fragments, because of incursions from surrounding human-disturbed habitats in which the Argentine ant is abundant.

The author now returns to the studies of competition in bottle experiments and introduced species discussed previously (*see* Constraints on Coexistence: Examples from Introduced Species and Coexistence and Exclusion: Messages from Bottles). In light of what we now know about mechanisms of coexistence, it is easy to see why competitive exclusion was often observed in the classic bottle experiments with protists and grain beetles. The design of these experiments included most of the elements assumed in the syllogism leading to the competitive exclusion principle. The lab environments were climatically controlled, and the culture bottles were spatially closed, precluding two of the major classes of coexistence mechanisms. The culture media and conditions (e.g., stirring) were set up to have a restricted number of limiting resources (ideally, one) and little within-culture spatial heterogeneity in limiting food resources; the communities are very simple, and so there is little opportunity for complex food web interactions. The most interesting result of these experiments may be that coexistence is observed at all. In the field study of introduced parasitoids, these species were specialized to the same host species, and the environment was relatively constant in time and homogeneous in space (the lowlands of tropical islands such as Hawaii have very stable climates, and the habitat is a deliberately homogenized landscape – plantations with regularly spaced fruit trees). It is likely that competitive exclusion is most readily observed in simple communities with few species in unstructured environments (e.g., islands tend to be low in species richness, particularly in heavily disturbed habitats dominated by introduced species). Most natural systems, in contrast, are temporally variable and spatially heterogeneous and harbor rich, complex communities with a multiplicity of interactions percolating through food webs; therefore, it should not be surprising that it is often difficult to document competitive exclusion in the wild. Recent analyses of time series of abundance and demographic rates in multispecies communities suggest that multiple regulatory factors are involved in species coexistence.

Future Directions

There are many important themes relevant to species coexistence that have not been addressed in-depth by ecologists but are likely to receive considerable attention in the near future.

Transient Coexistence

Ecological theory has traditionally emphasized equilibrial community states and the development of criteria for exclusion and indefinite persistence. However, exclusion takes place over some timescale. If this is long (e.g., relative to climate change), communities may exhibit transient dynamics and be far from equilibrium. This is particularly likely when considering interactions among species that are very similar in their niche requirements and environmental effects; as noted above, the abundances of similar species can vary through time in an essentially random fashion (i.e., Hubbell's community "drift"). If drift is occurring, after a sudden shift in relative abundances, species should tend to stay where they are put, rather than return to their initial abundances. An experiment conducted by Mark McPeek and Adam Siepelski with damselflies showed that species gained no demographic advantage when perturbed to low numbers, consistent with the hypothesis that the species are co-occurring at present, but not necessarily permanently coexisting. An important task for future ecological theory and empirical work is to derive a deeper understanding of transient dynamics and the drift hypothesis as explanations for community structure, complementing views based on deterministic mechanisms of coexistence.

Allee Effects, Niche Construction, and Coexistence

Most ecological theory applies literally only to clonal, asexual organisms. However, most species of concern to ecologists are sexual. At very low population sizes, outcrossed sexual species should have depressed growth rates simply because the two sexes have to get together to reproduce – an "Allee" effect (positive density dependence in growth rates at low densities). Allee effects can influence coexistence. For instance, consider two sexual species with identical ecologies in a landscape with many habitat patches. If colonization is infrequent, whichever species arrives first in a patch should be able to dominate there because the first colonist increases until births just match deaths. When the second species appears, its potential birth rate will then also just match its death rate (because its ecology is the same as that of the resident), but its realized birth rate will be depressed because it has difficulties in pair formation at low numbers. Hence, in this patch, the second species to arrive will be excluded. In the landscape as a whole, the two species may coexist if by chance each species is the first to occupy a subset of patches. In some patches, both species may by chance appear initially in sufficient numbers so as not to experience the Allee effect; in these patches, abundances should drift through time. Over the entire landscape, patterns in abundance will not be correlated with any discernible factors in the external environment. The magnitude of the Allee effect varies greatly among species, depending on details of the mating system and mate-finding strategies (e.g., selfing in plants should greatly weaken Allee effects). This influence of mating ecology on coexistence may provide an as yet poorly understood source of variation among taxa in community structure. Allee effects can also arise when organisms modify their local environments so as to enhance fitness, as when a nitrogen-fixing plant enriches the soil environment, thereby easing seedling establishment. Theoretical studies suggest that such niche construction can at times hamper coexistence and at times facilitate coexistence.

Speciation Mode and Coexistence

Over long timescales, all species originate from other species. Understanding the mode of speciation may provide an insight into coexistence. For instance, sympatric speciation occurs when a lineage diverges within a single community. Most models of sympatric speciation depend on substantial ecological differentiation being present from the beginning; as speciation unfolds, the ecological differences between the two emerging species must be sufficient to withstand cross-mating. In other words, with this mode of speciation, mechanisms of coexistence (e.g., substantial habitat or resource segregation) are built into the very branching pattern of the phylogenetic tree. In contrast, allopatric speciation requires geographical isolation. This can occur without any ecological difference arising between the daughter lineages (although speciation should occur more rapidly if correlated with ecological divergence). Dispersal can later bring the daughter species together, with essentially any degree of ecological difference being possible at this stage. Speciation may also often involve changes in sexual selection and mating systems, which can occur with no change in ecological requirements. After speciation, when these species' ranges begin to overlap, relative abundances should be particularly prone to community drift.

Evolutionary perspectives, such as phylogenetics (understanding the family tree of life), are more generally transforming studies of coexistence. Darwin in *On the Origin of Species* hypothesized that competition should be particularly strong between closely related species, which are likely to have quite similar resource requirements. A recent laboratory study by Lin Jiang built on Gause's classic experiments to examine this hypothesis. Pairwise contests among all possible pairwise combinations of 10 species of protists, combined with a phylogeny, revealed that the more closely related two protist species were, the more likely one would observe competitive exclusion, rather than coexistence, and if there was coexistence, the more likely one of the two species would be reduced greatly in its abundance. Recent studies show that the Janzen–Connell effect likewise has a phylogenetic signal, in that closely related tropical tree species inhibit each other's recruits more strongly than do distantly related species, via shared pathogens – an instance of apparent competition among close relatives. Distantly related species have different defensive strategies, minimizing apparent competition and in effect providing an axis for niche differentiation.

Conclusions

In conclusion, understanding the factors that promote or constrain the coexistence of species is an ongoing enterprise at the intellectual core of the study of biodiversity. Community ecologists have a rich smorgasbord of hypotheses to explain both species coexistence and exclusion. In most natural communities, it is likely that many mechanisms operate simultaneously. A challenging problem is to ascertain the relative contribution of these ecological mechanisms of coexistence to explain major patterns in biodiversity in space and time. There is much work to be done in analyzing mechanisms of coexistence in the context of food web dynamics and metapopulations. Moreover, little mention was made in this article about commensalism, mutualism, and nontrophic interactions, all of which can be crucial to coexistence. An improved understanding of the factors governing species coexistence is needed to address many applied problems, particularly the conservation of natural communities. Given the importance of spatial openness for coexistence, the anthropogenic alteration and simplification of landscapes pose particular dangers to the continued coexistence of many species. Indeed, the basic unstated problem of conservation is how to structure our activities and modify our impacts so that most of the world's biota can manage to coexist (at some spatial scale and, it is hoped, over long timescales) with just a single species, namely, ourselves.

See also: Coevolution. Food Webs. Introduced Species, Impacts and Distribution of. Landscape Diversity. Parasitism. Population Dynamics. Predators, Ecological Role of. Scale, Concept and Effects of

References

Chase JM and Leibold MA (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. Chicago: University of Chicago Press.

Chesson P (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31: 353–366.

Holt RD and Lawton JH (1994) The ecological consequences of shared natural enemies. *Annual Review of Ecology and Systematics* 25: 495–520.

Hutchinson GE (1978) *An Introduction to Population Ecology*. New Haven, CT: Yale University Press.

Levin SA (1970) Community equilibrium and stability, and an extension of the competitive exclusion principle. *The American Naturalist* 104: 413–423.

Morin PJ (2011) *Community Ecology*, 2nd edn. Oxford: Blackwell.

Rosindell J, Hubbell SP, and Etienne RS (2011) The unified neutral theory of biodiversity and biogeography at age ten. *Trends in Ecology and Evolution* 26: 340–348.

Schoener T (1974) Resource partitioning in ecological communities. *Science* 185: 27–39.

Tilman D and Kareiva P (eds.) (1997) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*. Princeton, NJ: Princeton University Press.

Species, Concepts of

James Mallet, University College London, London, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Cladistic A classification based entirely on monophyletic taxonomic groupings within a phylogeny; taxonomic units that are paraphyletic or polyphyletic are rejected. A cladist is one who practices cladistics, usually in the sense of using parsimony to adjudicate between data from multiple characters in the construction of a cladogram, which is an estimate of the true phylogeny.

Coalescent theory Coalescent theory is a recent development in population genetics, which investigates the merging of gene sequences traced backward in time within gene genealogies.

Cohesion The sum total of forces or systems that hold a species together. The term is used especially in the interbreeding and cohesion species concepts. Cohesion mechanisms include isolating mechanisms in sexual species as well as stabilizing ecological selection, which may cause cohesion even within asexual lineages.

Disruptive selection Selection acting to preserve extreme phenotypes in a population. Speciation usually involves disruptive selection, because intermediates (hybrids between incipient species) are disfavored (*also see* Stabilizing selection).

DNA barcoding A method of delimiting species via clustering of short stretches of DNA sequence data called "barcodes," usually from mitochondrial DNA.

Gene flow Movement of genes between populations, usually via immigration and mating of whole genotypes, but sometimes single genes may undergo horizontal gene transfer via transfection by microorganisms.

Gene pool The sum total of the genetic variation within a reproductively isolated species population; this term is mostly used by supporters of the interbreeding species concept.

Genomic cluster A synonym for genotypic cluster.

Genotypic cluster In a local area, a single genotypic cluster (or species) is recognized if there is a single group of individuals recognizable on the basis of multiple, unlinked inherited characters or genetic markers. A pair of such genotypic clusters (or species) is recognizable if the frequency distribution of genotypes is bimodal. Within each genotypic cluster in a local region, allele frequencies will conform to Hardy–Weinberg equilibrium, and the different unlinked loci will be in approximate linkage equilibrium. The presence of more than one species or genotypic cluster can then be inferred if the distribution of genotypes is bimodal or multimodal, and strong heterozygote deficits and linkage disequilibria are evident between the clusters.

Isolating mechanisms The sum total of all types of factors that prevent gene flow between species, including premating mechanisms (mate choice), and postmating mechanisms (hybrid sterility and inviability). Modern authors deny that these mechanisms have necessarily evolved to preserve the species' integrity as originally assumed, though this may sometimes be the case in reinforcement of premating isolation. Isolating mechanisms are a subset of the factors that cause cohesion of species under the interbreeding and cohesion species concepts.

Monophyletic A grouping that contains all the descendants of a particular node in a phylogeny. Monophyly is the state of such groupings. Compare paraphyletic and polyphyletic. Butterflies (Rhopalocera) and birds (Aves) are the examples of two groups thought to be monophyletic.

Paraphyletic A grouping that contains some, but not all, of the descendants of a particular node in a phylogeny. Paraphyly is the state of such groupings. Compare monophyletic and polyphyletic. Moths (Lepidoptera, excluding butterflies) and reptiles (amniotes, excluding birds and mammals) are examples of two groups thought to be paraphyletic.

Phenetic A classification or grouping based purely on overall similarity. Pheneticists use matrices of overall similarity rather than parsimony to construct a "phenogram" as an estimate of the phylogeny. Examples of phenetic methods of estimation include unweighted pair group analysis (UPGMA) and neighbor joining. Cladists reject phenetic classifications on the grounds that they may result in paraphyletic or polyphyletic groupings.

Phylogenetic Pertaining to the true (i.e., evolutionary) pattern of relationship, usually expressed in the form of a binary branching tree, or phylogeny. If hybridization produces new lineages, as is common in many plants and some animals, the phylogeny is said to be reticulate. Phylogenies may be estimated using phenetics, parsimony (cladistics), or methods based on statistical likelihood.

Polyphyletic Groupings that contain taxa with more than one ancestor. "Polyphyly" is the state of such groupings. Compare paraphyletic and monophyletic. "Winged vertebrates" (including birds and bats) are examples of a polyphyletic group.

Real, reality Two tricky words found frequently in the species concept debate. Reality is typically used to support one's own species concept, as in: "The conclusions set forth above … lead to a belief in the reality of species" (Poulton, 1904); similar examples can be found in the writings of Dobzhansky, Mayr, and many phylogenetic systematists. The term reality in this sense is similar to an Aristotelian essence, a hypothetical pure, albeit obscure, truth that underlies the messy actuality; unfortunately, in everyday language real also means actual (curiously, a reality in the first sense may be unreal under the second!). By rejecting the reality of species, one can therefore send very mixed messages: some readers will understand the author to be a

nominalist who merely believes useful terms require no theoretical underpinning; others assume the author is nonsensically using some definition that does not apply to the actual organisms. Here, when the author discusses the reality underlying a species concept, he means it in the first sense, a hypothesis. Many authors of species concepts and some philosophers of science argue that definitions must be underpinned by a theoretical justification or reality. Other philosophers such as Wittgenstein and Popper agree that terms need no such definition to be useful.

Sibling species A pair of closely related, morphologically similar species (usually sister species).

Speciation The evolutionary process of the origin of a new species.

Specific mate recognition systems (SMRSs) Fertilization and mate recognition systems in the recognition concept of species, the factors leading to premating compatibility within a species. Also see cohesion, which is similar to SMRS, but includes postmating compatibility as well.

Stabilizing selection Selection which favors intermediate phenotypes.

Taxonomic inflation The process whereby the numbers of species in the checklist of a group increases due to a change in species concept rather than due to new discoveries of previously unknown taxa.

What are Species Concepts for?

Individual organisms can usually be recognized, but the larger units we use to describe the diversity of life, such as populations, subspecies, species, or genera are not so readily discerned. Taxonomists further group species into genera, families, orders, and kingdoms, whereas ecologists group species into higher structures such as communities and ecosystems. The justification for these group terms is utility, rather than intrinsic naturalness, but as far as possible we attempt to delimit groups of organisms along natural fault lines, so that approximately the same groupings can be recovered by independent observers. However, there will be a virtually infinite number of different, albeit nested, ways of classifying the same organisms, given that life has evolved more or less hierarchically.

Darwin (1859) felt that species were similar in kind to groupings at lower and higher taxonomic levels; in contrast, many recent authors suggest that species are more objectively identifiable, and thus more real than, say, populations or genera. Today, much of ecology, evolution, and biodiversity appear to assume that the species is the fundamental taxon. It has been argued that these fields could be undermined if, say, genera, or subspecies, had the same logical status as species.

Species concepts originate in taxonomy, where the species is "the basic rank of classification" according to the International Commission of Zoological Nomenclature. The main use of species in taxonomy and derivative sciences is to order and retrieve information on individual specimens in collections or data banks. In evolutionary biology, it may also be helpful to delimit a particular kind of divergent evolution, "speciation," which produces a result qualitatively different from within-population evolution, although it may of course involve the same processes. In ecology, the species is a group of individuals within which variation can often be ignored for the purposes of studying local populations or communities, so that species can compete, for example, whereas subspecies or genera are not usually considered in this light. For measuring biodiversity, in conservation studies, and in environmental legislation, species are potentially important units, which we would like to be able to count both regionally and globally.

It would be helpful if a single definition of species could satisfy all these uses, but a generally accepted definition has yet to be found, and indeed is believed by some to be an impossibility. Yet a unitary definition should be possible if species are more real, objectively definable and fundamental than, say, genera or subspecies. Conversely, even if species have no greater objectivity than other taxa, unitary nominalistic guidelines for delimiting species might be adopted, perhaps after much diplomacy, via international agreement among biologists; after all, if we can adopt meters and kilograms, perhaps we could agree on units of biodiversity in a similar way. In either case, knowledge of the full gamut of today's competing solutions to the species concept problem will probably be necessary for a universal species definition to be found. This article reviews several major proposals currently on the table, and their usefulness in ecology, evolution, and conservation.

Statement of Bias

The author is of Darwin's opinion that the species, like genera and populations, do not have clear boundaries. In evolution and in ecological and biodiversity surveys over large areas, their reality has, in the author's view, been grossly overestimated (Mallet, 2008). In contrast, it is clear to any naturalist that species are usually somewhat objectively definable in local communities. It is the author's belief that confusion over species concepts has been caused by scientists not only attempting to extend this demonstrable local objectivity of the species over space and evolutionary time, but also arguing fruitlessly among themselves about which is the most important apparent reality (or concept) that underlies this spatiotemporal extension of local objectivity. To the author, agreement on a practical, unified species-level taxonomy is possible, but will be forthcoming only if we accept that different species may often lack a unitary biological reality over their geographic ranges and across geological time.

Just as Marxist theory may be wrong, yet remains a convenient approach for studying political history, the author hopes that his own views can provide, even for the skeptic, a useful framework on which the history of proposals for species concepts can be compared. A variety of alternative outlooks can be found by Mayr (1982), Cracraft (1989), Claridge *et al.* (1997), Howard and Berlocher (1998), Hey (2001), and Coyne and Orr (2004).

Darwinian Species Criteria

Darwin's Morphological Species Criterion

Early thinkers assumed that each species had an Aristotelian form or essence, and that variation within a species was due to imperfections in the actualization of this form. Each individual species was defined by its essence, which itself was unvarying and inherently different from all other species essences. Before 1859, it was generally agreed that this essence was descent: all individuals within a species were related by descent, whereas no individuals were descended from individuals in another species. This mode of thought of course precluded transformation of one species into another, and was associated with belief that each form was separately created by God.

Darwin's extensive travels and knowledge of taxonomy led to a realization that the distinction between intraspecific and interspecific variation was false. Individuals could be descended from members of other species: "descent with modification." His abandonment of essentialist philosophy and its species concept went hand in hand with his appreciation that variation itself was among the most important characteristics of living organisms, because it was this variation, which allowed species to evolve. To Darwin, species could be distinguished from divergent varieties if they showed few intermediates: "hereafter we shall be compelled to acknowledge that the only distinction between species and well-marked varieties is, that the latter are known, or believed, to be connected at the present day by intermediate gradations, whereas species were formerly thus connected" (Darwin, 1859, p. 485).

Darwin guessed (correctly) that essentialist species would be hard to give up: "… we shall have to treat species in the same manner as those naturalists treat genera, who admit that genera are merely artificial combinations made for convenience. This may not be a cheering prospect, but we shall at least be freed from the vain search for the undiscovered and undiscoverable essence of the term species" (Darwin, 1859, p. 485). He argued that species were little more than varieties that acquired their claim to a greater reality only when intermediates died out leaving a morphological gap: "… I believe that species come to be tolerably well-defined objects, and do not at any one period present an inextricable chaos of varying and intermediate links" (Darwin, 1859, p. 177). This morphological gap criterion, which seems to have been accepted by many early evolutionists (e.g., Wallace, 1865; Robson, 1928), has been called a "morphological species concept" because Darwin used the gaps in morphology to delimit species; however, it is easy to extend his species criterion to ecology, behavior, or genetics (*see* Genotypic Cluster Criterion and Assignment Tests).

Polytypic Species

A major revolution in zoological taxonomy occurred around 1900. As the great museum collections became more complete, it became obvious that many apparently distinct species found in different areas intergraded where they overlapped. These replacement species began to be regarded as subspecies within a polytypic species, an idea suggested for geographical varieties by early systematists and Darwinists such as Wallace (1865). The taxonomic clarification that followed, which allowed identifiable geographic varieties to be named validly below the species level as subspecies, was conceptually more or less complete by the 1920s and 1930s. At the same time, other infraspecific animal taxa such as local varieties or forms were deemed un-nameable in the Linnaean taxonomy. These changes are now incorporated into the International Code of Zoological Nomenclature. Similar ideas were promoted in botany by G.L. Stebbins (see Mayr, 1982), although in the International Botanical Code local varieties and polymorphic forms remain valid and nameable taxa.

The Rephilosophization of Species, the "Interbreeding" Concept

In January 1904, Poulton read his famous presidential address – "What is a species?" – to the Entomological Society in London (see historical analysis by Mallet, 2004). Following up on ideas raised (but immediately dismissed) by Wallace (1865), Poulton proposed "syngamy" (i.e., interbreeding) as the true meaning of species. Poulton and Wallace were both particularly knowledgeable about swallowtail butterflies (Papilionidae). In swallowtails, there exist strong sexual dimorphisms: the female color pattern often mimics unrelated unpalatable butterflies whereas the male is nonmimetic. The females themselves are sometimes polymorphic (e.g., in *Papilio memnon* and *Papilio dardanus*), each female form mimicking a different distasteful model. Under a crude morphological criterion, each divergent form could be designated as a different species, whereas mating observations in the wild showed that the forms were part of the same interbreeding group. A related idea was promoted by the botanist Lotsy, who termed the interbreeding species a "syngameon" (Lotsy, 1917). In the 1930s, Dobzhansky studied morphologically indistinguishable sibling species of *Drosophila* fruit flies and concluded that Lotsy's approach had some value (Dobzhansky, 1937). A member of a species will rarely, if ever, interbreed with a member of a sibling species; each chooses mates within its own species. Dobzhansky proposed his own interbreeding species concept, later popularized by Mayr as the "biological species concept," so named because interbreeding within species, coupled with reproductive isolation between species, was considered the single true biological meaning or reality of the term species (reviewed by Mayr, 1970, 1982).

A short definition of the biological species concept is as follows: "species are groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr, 1970). The biological concept of species was not so much new as a clarification of two distinct threads: (1) a local component, the Poulton/Dobzhansky interbreeding concept, and (2) a global component, which extended the interbreeding concept to cover geographically separated populations of actually or potentially interbreeding subspecies, as in the preexisting idea of polytypic species.

This extended interbreeding concept was, until about 30 years ago, almost universally adopted by evolutionists. The species concept problem appeared to have been solved; species were interbreeding communities, each of which formed a "gene pool" reproductively incompatible with other such

communities. The new concept answered both perceived problems of Darwin's morphological approach: (1) That mutants and polymorphic variants within populations might be considered as separate species. (2) That sibling species might be misclassified as members of the same species. The new approach was promoted in a long series of books and articles by Dobzhansky, Mayr, and their followers. Mayr, in particular, was highly influential by justifying the taxonomic application of the polytypic species criteria in terms of the new concept of "gene flow." The new species concept was considered necessary for understanding speciation in relation to the population genetics. Darwin was hereafter deemed to have failed to solve, or even address the origin of species in his 1859 book, because he supposedly had not appreciated the fundamental importance of interbreeding (Mallet, 2010a).

To take this view, it was necessary to see species in a new post-Darwinian light. Instead of species being defined simply, using man-made criteria based on demonstrable characters such as gaps in morphology, species became defined by processes important in their own maintenance, that is by means of their biological function (Mayr, 1982). Significantly, the philosophical term "concept" came into vogue along with these new ideas about species, and the term "species problem," which hitherto had referred to the problem of how species arose (Robson, 1928), became instead the problem of defining what species were. The important features of species defined by the "biological species concept" were that they were protected from gene flow by physiological isolating mechanisms (Dobzhansky, 1937) or "reproductive isolating mechanisms," including prezygotic mechanisms (ecological, mate choice, and fertilization incompatibilities) and postzygotic mechanisms (hybrid inviability and sterility caused by genomic incompatibilities). Curiously, by going beyond simple, Darwinian character-based identification of species, the "biological concept of species" became less universally applicable in biology; for example, Dobzhansky (1937) simply concluded that asexuals (between which no interbreeding is possible) could not have species.

Poulton, Mayr, and Dobzhansky emphasized that their new concept was based on the reality that underlay species, rather than being merely a criterion useful in taxonomy. In this new philosophical approach, taxonomic criteria became separated from conceptual issues, and practical delimitation took on a secondary role. The same division between concept and criteria is still evident in most discussions of species concepts today, for example in the "general lineage concept" (de Queiroz, 1998; *see* Evolutionary and Lineage Concepts). The concept was seen to be true from first principles, and became untestable: violations of the criterion of reproductive isolation, such as hybridization, intermediates, and inapplicability to many plants and asexuals caused problems in classification, but did not disprove or even challenge the underlying truth of the concept itself. These imperfect actualizations of species' true reality were expected in nature. Mayr claimed that the biological concept would do away with "typology" (his term for species definitions based on a fixed, unvarying type or Aristotelian essence), but in many ways it can be seen that the biological concept reverts to a new kind of essentialism, where evolutionary maintenance via interbreeding is the underlying reality, or essence of species.

Alternative Species Concepts

It is interesting that exactly this kind of search for the essence of species had already been criticized by Darwin (1859), as we have seen (*see* What are Species Concepts for?). In his chapter "Hybridism," Darwin specifically argued against using postzygotic isolation (hybrid sterility and zygote inviability) as cut-and-dried characteristics of species. In this discussion, he made no mention of prezygotic isolation, another component of the reproductive isolation that characterizes species under the biological concept. However, we can infer that Darwin, the discoverer of sexual as well as of natural selection, would almost certainly have argued that mate choice, along with hybrid sterility and inviability, are all found within as well as between species, although to a lesser extent. Oddly, Mayr (1982, p. 269) claimed that Darwin treated species "purely typologically (i.e., as an essentialist) as characterized by the degree of difference," and also that Darwin "had strong, even though perhaps unconscious, motivation … to demonstrate that species lack the constancy and distinctiveness claimed for them by the creationists." Whether or not it is reasonable to criticize Darwin in such a contradictory way can be debated, but it is clear that Mayr's proposition that interbreeding was the true essence or reality of species immediately laid itself open to debate. Although the interbreeding concept had a long run (and still does), proposals for different kinds of biological reality of species were eventually forthcoming. By proposing a single reality for species, Poulton, Dobzhansky, and Mayr reopened Pandora's box of alternative essences deemed more important by other biologists.

Ecological Species Concept

Asexual organisms such as the bdelloid rotifers can clearly be clustered into groups recognizable as taxonomic species, very likely because competition has led to the extinction of intermediates (Hutchinson, 1968; Fontaneto *et al.*, 2007). However, distinct sexual forms such as oaks (*Quercus*), between which there are high rates of hybridization, can remain recognizably distinct even where they cooccur. This suggested to van Valen (1976) that the true meaning of species was occupancy of an ecological niche or adaptive zone rather than interbreeding, yielding an ecological species concept. It became clear to Mayr by 1970 also (see Mayr, 1982) that gene flow could not unite every population in a polytypic, biological species' range, and that stabilization of phenotype might be effected by ecologically mediated stabilizing selection (*see* Recognition Concept of Species *and* Cohesion Concept of Species) rather than purely because of gene flow.

Recognition Concept of Species

An important attack on the biological species concept came from Hugh Paterson in the early 1980s. His claims were twofold: First, that the Dobzhansky/Mayr term "isolating mechanisms" implied that reproductive isolation was adaptive, which Paterson felt was unlikely. Second, that the true reality underlying species was prezygotic compatibility, consisting of mating signals and fertilization signals. According to

Paterson (1985), this compatibility is strongly conserved by stabilizing selection, whereas isolating mechanisms such as hybrid sterility or inviability are nonadaptive and can be argued to be a result rather than a cause of species separateness. To Paterson, the true reality of species must be adaptive, and the idea that species are held together by beneficial isolating mechanisms was a fallacious argument that relied on adaptation at the whole species level (Mallet, 2010a, b). Paterson termed his idea of species the "recognition concept" versus Mayr's "isolation concept," and its important characteristics specific mate recognition systems (SMRSs) instead of isolating mechanisms. Species were defined as "that most inclusive population of individual biparental organisms, which share a common fertilization system" (Paterson, 1985).

The idea is generally recognized as a useful critique and has gained strong currency in some circles. However, it appears also that SMRSs are more or less the inverse of prezygotic isolating mechanisms, and that the recognition concept therefore differs practically from the biological species concept mainly by focusing on the subset of traits acting before fertilization. In addition, the interbreeding concept had always stressed a common gene pool and compatibility within a species, as well as isolation between species.

Species Concepts Based on History

Monophyly

The rise of cladistic methods revolutionized systematics by proposing that all classification should be based on the idea of "monophyly." This new system formalized the principle that paraphyletic and polyphyletic taxa were unnatural groupings, which should not be used in taxonomy. It was natural to attempt to apply this idea throughout systematics, all the way down to the species level, leading to a monophyly criterion of species (Hennig, 1966), a type of phylogenetic species concept (*see* Diagnostic Species Concept). Species were seen as forming when a single interbreeding population split into two branches or lineages that did not exchange genetic material. In a somewhat different formulation, the cladistic species concept, species are branch segments in the phylogeny, with every branching event leading to a new pair of species (Ridley, 1989). Otherwise, if only one of the two branches were recognized as new, the other branch would become paraphyletic.

Perhaps the main criticism of this idea is that it could, if applied in taxonomy, cause great nomenclatural instability. Monophyly exhibits fractal self-similarity and can exist at very high or very low levels of the phylogeny, so the precise level at which species taxa exist becomes unclear. Suppose that a new monophyletic form is discovered overlapping with, but remaining distinct from, a closely related local form in the terminal branches of an existing species. Recognition of this taxon as a species would leave the remaining branches within the original species paraphyletic. Many other branch segments would then need to be recognized at the species level, even if they interbreed and have reticulate, intermingling relationships. Many phylogenetic systematists therefore adopt a different phylogenetic concept, the diagnostic concept (below), which can allow paraphyly at the species level.

Diagnostic Species Concept

The motivation for the diagnostic concept, usually called the "phylogenetic species concept" by its adherents, was again to incorporate phylogenetic thinking (Hennig, 1966) into species-level taxonomy. There are many cases of hybridization between taxa on very different branches of species-level phylogenies, which suggests that interbreeding and phylogenetic realities conflict. Cracraft (1989) also noted that many bird taxa, normally thought of as subspecies, were far more recognizable and stable nomenclaturally than the polytypic species to which they supposedly belonged (*see* Genotypic Cluster Criterion and Assignment Tests). Cracraft therefore argued that the polytypic/interbreeding species concept should be rejected, and, in its place, we should use a diagnostic criterion in the form of fixed differences at one or more inherited characters. A phylogenetic species is an irreducible (basal) cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent (Cracraft, 1989). According to Cracraft, species defined in this way are the proper basal, real taxa suitable for phylogenetic analysis and evolutionary studies.

Of course, if diagnostic criteria are applied strictly, rather small groups of individuals, or even single mutant specimens, might be defined as separate species, leading to potentially unbridled taxonomic inflation. Cracraft recognized this and argued that such trivial diagnosable groups would have no parental pattern of ancestry and descent, that is, they are not proper populations. However, this qualification appears similar to an interbreeding criterion of species, whereas the whole approach of using diagnostic characters was an attempt to get away from interbreeding.

Most evolutionary biologists balk at the idea of speciation being merely the acquisition of a new geographically diagnostic character, a DNA base pair or color pattern change perhaps (Harrison, 1998). Speciation is only a different, or special, kind of evolution if the new species is a distinct population, which can coexist locally with its sibling or parent population without losing its integrity.

Characters used to diagnose phylogenetic species may not be shared derived characters; they may be primitive (plesiomorphic) characters, or they may have evolved several times. Therefore, phylogenetic species need not be monophyletic, and could presumably be paraphyletic and perhaps polyphyletic. Cracraft appears confused on this matter: on one hand, he claims that phylogenetic species "will never be nonmonophyletic, except through error" (Cracraft, 1989, p. 35), but on the other hand he recognizes that "their historical status may (sometimes) be unresolved because relative to their sister species they are primitive in all respects. Whether they … (are) truly paraphyletic … is probably unresolvable" (Cracraft, 1989, p. 35). It seems odd to allow a phylogenetic species even to be paraphyletic (let alone polyphyletic), because paraphyly and polyphyly contravene the basic tenets of phylogenetic systematics, and because one of the main justifications for a phylogenetic species concept is that species defined via other concepts might sometimes be paraphyletic: "the biological species concept cannot be applied to the *Thomomys umbrinus* complex unless one is willing to accept paraphyletic species,

and to do so would be a de facto admission that biological species are not units of evolution" (Cracraft, 1989, p. 46; also see Davis, 1997, p. 374). The phylogeneticists' resolution of that problem, using diagnostic characters, leads to the same difficulty all over again! This rather glaring logical inconsistency considerably undercuts the argument for a diagnostic species concept.

In spite of these logical problems, Cracraft highlighted genuine and important practical problems with the polytypic application of the interbreeding concept, and as a result this phylogenetic species concept has been very influential. Recently, many molecular systematists, including botanists (Davis, 1997), have taken up Cracraft's suggestion and used diagnostic differences between geographic populations, in some cases at single DNA base pairs, as evidence that two forms are separate species even if they intergrade freely at the boundaries of their distribution. Ornithologists and primatologists in particular have used diagnostic characters to reassign many taxa long thought of as subspecies to the level of full species, resulting in rather severe taxonomic inflation (Isaac *et al.*, 2004).

Genealogy

Other phylogenetic systematists continued to be interested in monophyly-based species concepts. A problem with a monophyly concept is that a single, true phylogeny of taxa may, in a sense, rarely exist: a phylogeny of species can be seen as an abstraction of the actual genetic history, consisting of multiple organismal genealogies (Hennig, 1966), or multiple gene genealogies (Avise and Ball, 1990; Maddison, 1997), some of which may undergo genetic exchange with other taxa. There is now good evidence that occasional horizontal gene transfer and hybridization may selectively cause genetic material to flow between unrelated forms (Mallet, 2008). Furthermore, there are multiple gene lineages within any population, so that, if such a population were to become geographically or genetically split into two distinct forms, it would take some time before each population became fixed for different, reciprocally monophyletic gene lineages at any single gene. The idea of monophyly for whole genomes then becomes hard to define, especially near the species boundary. However annoying, phylogenetic methods and evolutionary theory must face up to these facts (Avise and Ball, 1990; Maddison, 1997). Instead it was suggested that monophyletic species could be defined if they formed exclusive sets, where each member was related more closely to members of its own species than to any members of another species. Species were argued to be formed at the fuzzy boundary between reticulate (within species) and divergent (between species) organismal genealogies; it was suggested that such species could be delimited when a consensus of multiple gene genealogies indicated reciprocal organismal monophyly. This was called the "genealogical species concept" (Baum and Shaw, 1995).

Critics argue that this idea has many problems in common with other monophyly concepts of species (Davis, 1997; Hudson and Coyne, 2002). Geographic forms that have become isolated in small populations or on islands, say, could rapidly become fixed for gene lineages, and become viewed as separate species without any biologically important evolution taking place. Conversely, clearly distinct sister taxa such as humans and chimpanzees still share gene genealogy polymorphisms at some genes such as the human leukocyte antigen (HLA) complex involved in immunological defense, and might therefore be classified as the same species under genealogical considerations.

Bayesian Species Delimitation

A new approach employs a Bayesian analysis to compare competing species models using multilocus sequence data (Yang and Rannala, 2010). For example, the one-species model assumes that the gene trees for different loci fluctuate according to the standard neutral (coalescent) theory within a single population, whereas the two-species model assumes two species that diverged some time ago, so that there will be more concordance among the gene trees across loci. The two-species model assumes a "biological species concept," i.e., that gene flow ceased at some time in the past. The competing species models are compared using their posterior probabilities, allowing one to assess if sequence data are compatible with the one-species model, or whether the two-species model has to be invoked. It should also be emphasized that more than two species can be readily accommodated in this hypothesis testing framework.

Concordance of gene trees across multiple loci is used as evidence to infer species-level monophyly (as in the reciprocal monophyly concept of the section Monophyly), but the method does not rely on reciprocal monophyly of each gene tree. It accounts for random fluctuations in the coalescent process in both the extant and extinct species and thus takes into account genealogical concerns of the section Genealogy. Its statistical approach, however, accommodates uncertainties in the gene trees, unlike the genealogical species concept, the suggested implementation of which ignored phylogenetic errors in the gene trees when constructing the consensus. Simulations suggest that small amounts of gene flow ($Nm < 0.1$ migrants per generation, say) has little impact on the behavior of the method. The method thus should recognize distinct species despite occasional hybridization (Zhang *et al.*, 2011).

This method represents a considerable advance in sophistication of analysis of genomic data to delimit species. In that it employs Bayesian posterior probability to infer the number of species taxa from a set of data, it is similar to "assignment tests" to detect multilocus genotypic structure (*see* Alternative Species Concepts) but instead it uses neutral coalescent modelling rather than a simpler model of free recombination and assortment within species. The method currently requires adoption of a guide tree of taxa: populations that are potential species and their relationships must be defined before statistical decisions to award species status can be made. It is arguable that Bayesian delimitation may view spatially separate populations as species simply because they happen to have separated some time ago, and not because they have speciated in the sense of their ability to overlap without fusion. It may thus exhibit the same problems (if they are viewed as problems) as the monophyly concept of species (*see* Monophyly).

Combined Species Concepts

Evolutionary and Lineage Concepts

Faced with the problem of studying the evolution of species through time, the paleontologist Simpson (1951) proposed his evolutionary species concept, in which a species is "a lineage (an ancestral–descendant sequence of populations) evolving separately from others and with its own unitary evolutionary role and tendencies." In other words, Simpson combined the idea that species were historical lineages with the concept of their evolutionary and ecological role. The key essence here appears to be "evolutionary independence." This concept appeals to phylogenetic systematists and paleontologists alike, because of its historical dimension, and to neontologists because of its acknowledgment that biological mechanisms are what make the species real. De Queiroz (1998) is one of the more recent reviewers to propose that a single concept, which he calls "the general lineage concept," under which species are segments of population-level lineages, underlies all other species concepts. According to de Queiroz, the other apparently competing species concepts merely emphasize different characters or criteria for species delimitation, but all acknowledge implicitly or explicitly that evolutionary separateness of lineage is the primary concept. This is a nice ideal, but vague terms like "evolutionary independence" have little logical force in their application to actual forms that hybridize or undergo genetic exchange. Furthermore, by failing to delimit where species are along the continuum of diverging lineages, de Queiroz sidesteps the whole issue of species definition rather than solving it. The main content of de Queiroz's (1998) idea seems to be that species emerge gradually. However, this particular finding had been clarified 140 years earlier.

Cohesion Concept of Species

In similar vein, Templeton's (1998) "cohesion concept" combined a number of competing ideas of species. He accepted criticisms of the interbreeding species concept, and attempted to correct it by infusing ideas from the evolutionary, ecological, recognition, and genealogical concepts. Templeton argued that a combination of ecological and reproductive cohesion is important for maintaining a species' evolutionary unity and integrity. As well as applying to asexual taxa (too little sex), Templeton also applied his idea to species like oaks that undergo frequent hybridization and gene flow (too much sex for the interbreeding concepts). He further argued that separateness of genealogy is another important characteristic of species.

We are perhaps nearing the apogee of the species debate with these combined concepts. By incorporating evolutionary, phylogenetic, and genealogical origins together with every possible biological means by which the distinctness of species is currently maintained, these combined concepts cover all the bases. One can acknowledge that species evolve and are maintained as cohesive wholes by all of these multifarious processes; yet at the same time species can, and perhaps should be seen as separate from their histories of origin and from current reasons for their integrity. If groups with very different and conflicting biological and evolutionary characteristics are all considered species, there should exist a simpler criterion that unites them. It can also be argued that to conflate the origin and evolutionary role of a taxon with the definition of that taxon itself may lead to circularity, particularly in conservation or ecological studies, or when investigating speciation.

Dissent: Maybe Species Are Not Real

Throughout the history of the species debate, starting with Darwin, there have been some who argue that species are not individual real objects, but should instead be considered as man-made constructs, merely useful in understanding biodiversity and its evolution. These people are not necessarily nihilists, who deny that species exist: they simply argue that morphological and genetic gaps between actual populations would be more useful for delimiting species than inferred processes underlying evolution or maintenance of these gaps. By their refusal to unite these ideas under a single named concept, this silent majority has rarely found a common voice.

Taxonomic Practice

Taxonomists are on the front line of the species battle, because it is they who ultimately decide whether to lump or split taxa, and at what level to name them as species. If the objectivity and individuality of species as the primary taxon exists, taxonomists' activities do not seem to have been made any easier. Many taxonomists have simply ignored or denied belief in the evolutionary reality of species. In any group, it is probably true to say that at least 10% of taxonomic species are subject to revision because of the practical difficulties in delimitation.

For this reason, since the rise of the polytypic/interbreeding species concept, there has been little impact of the postwar species concepts on practicing taxonomists, even while the debate raged around them, at least up until the 1980s. Procedure, at least in zoology, was more or less as follows: geographic variants, which blended (or were thought to be able to blend) together at their boundaries were united within a single, polytypic species, unless morphological or genetic differences were so great that it seemed necessary to recognize two species. However, whenever two divergent forms differing at several unrelated traits overlapped spatially, they were recognized as separate species even if a few intermediates suggested some hybridization or gene flow. Some taxonomists regarded subspecies as artificial taxa to be avoided, and may either have ignored geographic variation, or elevated subspecies of formerly polytypic species to the rank of full species. But good taxonomic practice on species remained broadly similar across most branches of systematics, and involved careful analysis of multiple, chiefly morphological character sets tested in large samples of specimens collected from as many geographic regions as possible.

This view on species and subspecies had led to a steady reduction in the numbers of recognized species in zoology, as more and more dubiously separated taxa, previously ranked as species, became inserted as subspecies or local populations into larger and larger polytypic species. Recently, however, the diagnostic version of the phylogenetic species concept (*see* Diagnostic Species Concept) has been making strong inroads into

zoological nomenclature, with the result that counts of species on continents are again climbing as former subspecies are re-elevated to the species level, in spite of intergradation at their boundaries (Isaac *et al.*, 2004). However, the situation could get much worse; many *Heliconius* butterflies, for example, have over 30 geographic subspecies per species, all of which can be diagnosed easily. The numbers of bird and butterfly species could easily increase 2–10 times in some groups if the diagnostic criteria were generally adopted, and indeed in some well-known groups, for example primates, a doubling of species numbers has already been observed. Most of these increases have come from reclassification of known subspecies or populations, rather than from discovery of new populations (Isaac *et al.*, 2004). The one reality that is clear in species-level taxonomy is that the species is not real enough to remain at the same taxonomic rank while fashions in species concept change. This provides evidence that actual species taxa have been and still are man-made groupings that do not employ objective biological or evolutionary essences, even if such essences exist.

Populations Are Evolutionary Units, Not Species

Botanists deal with geographically variable organisms with low powers of dispersal, and have tended not to be as happy with the polytypic/interbreeding concept applied with such apparent success in zoology. Meanwhile, the strong surge in experimental population genetics and evolutionary studies that followed the books by Dobzhansky, Mayr, and Stebbins has led to a greatly improved understanding of gene flow in natural populations. Gene flow, even in quite mobile animals such as birds or butterflies, may not unite distant populations into a common gene pool. If such populations only rarely exchange genes, then gene flow across the range of a continental species is clearly insufficient to explain species integrity, because it would be outweighed easily by weak local patterns of adaptation or genetic drift (Endler, 1977).

This increasing input of population biology into systematics and evolution led to the proposal by Ehrlich and Raven (1969), Levin (1979), and others that species are not real biological units at all; instead, local populations are the only real groupings united by gene flow within a common gene pool, and which adapt to local conditions, compete, and so on. Any homogeneity of ecological niche or genetics over the range of a species might be owing either to recent spread and simple evolutionary inertia or to similar stabilizing selection everywhere. To these authors, species exist and are real in local communities, but it is fallacious to treat distant populations in the same way (*see* Genotypic Cluster or Genomic Cluster Criterion).

This viewpoint is generally understood and respected by population biologists, but curiously has not been incorporated explicitly into current thinking on species in systematics and evolution. Perhaps there is a sneaking suspicion that even very weak levels of gene flow may explain species integrity over wide areas.

Phenetic Species Concept

In the 1960s and 1970s, a major systematics movement proposed numerical methods in taxonomy now usually referred to as "phenetics." Pheneticists, as they were called, argued that taxonomy and systematics should be based on multivariate statistical analysis of characters rather than on underlying evolutionary or biological process information. If taxa were defined by nonevolutionary criteria, studies of evolution would be freed from the tautology of testing hypotheses about processes, when those same processes are used as assumptions in the definitions of the taxa under study. Species, like other taxa, would be defined in numerical taxonomy on the basis of multivariate statistics, as clusters in phenotypic space (Sokal and Crovello, 1970).

Phenetics was reviled by those who believed that classifications should be phylogenetic. However, for species the approach is closely similar to the intuitive methods adopted by the most actual taxonomists, who use multiple morphological or genetic characteristics to sort individual specimens into discrete groups between which there are few intermediates (*see* Ecological Species Concept). Some large areas of practical taxonomy are based purely on this phenetic approach. Bacterial systematists, for instance, use multiple biochemical tests to assign microbes to species taxa (Claridge *et al.*, 1997). The usefulness of this taxonomic method is attested by its success in hospitals for predicting pathogeneticity and antibiotic sensitivity.

Phenetic classifications based on morphology introduce the danger that, if convergent characters are used as data, one may group unrelated forms into polyphyletic taxa. In addition, single gene polymorphisms and sexual dimorphism can affect multiple morphological characters. This could lead to a recognition of multiple species within polymorphic populations. Sibling species, however, could be lumped into the same species using a phenetic approach, unless a set of highly diagnostic characters could be found. Nonetheless, these problems are mainly due to the lack of characters found in morphological datasets. Phenetic methods have proved much more successful in distinguishing unrelated, although cryptic, taxa from polymorphic forms when coupled with molecular genetics techniques developed since the 1960s, including allozymes and DNA-based methods (Avise, 1994).

Genotypic Cluster Criterion and Assignment Tests

For morphological or genetic gaps to exist between species, gene flow (if any) between species must be balanced by an opposing force of disruptive selection. In his own work, the author had studied hybrid zones between geographic forms of butterflies, and the author attempted to show that a practical statistical definition of species versus geographic races could be constructed using morphological and genetic gaps alone, rather than employing the phylogenetic or evolutionary processes that caused the gaps to exist.

However, to define species by means of the gaps between them requires consideration of the nature of the gaps to avoid falling into the trap of defining polymorphic forms as separate species, or of lumping sibling species. Rather than merely using external morphology, in difficult cases the author proposed that we could consider genetics as well. DNA has a digital, rather than analog code, so there are genetic gaps between virtually any pair of individuals. Clearly, then, we

cannot use just any discreteness at the genetic level to define species. Separate sexes and polymorphic female forms of mimetic *Papilio* butterflies also have gaps between them in exactly this way. A genetic element, which may be a single base pair, an allele at a gene, the entire mitochondrial genome, a chromosomal rearrangement, or perhaps a sex chromosome, may determine the genetic or morphological differences between such polymorphic forms.

To be considered part of a single local population, and therefore part of the same local species, we expect that polymorphic genetic elements like mimicry genes and sex chromosomes will be approximately randomly combined with polymorphisms at genetic elements found on other chromosomes or extrachromosomal DNA. Each individual may be a distinct multilocus genotype, but we recognize a single grouping of genotypes because polymorphisms at one genetic element are independent of polymorphisms at others. Conversely, if alleles at one locus are strongly associated with alleles at other, unlinked elements (i.e., linkage disequilibrium or gametic disequilibrium), we have evidence for more than one separate population; if these two populations overlap spatially, the groupings are probably also separate species.

Many of us therefore proposed a "genotypic cluster criterion" for species (Mallet, 1995; Feder, 1998). The term "genomic cluster" would perhaps be a more appropriate synonym in today's postgenomic age. Species are recognized by morphological and genetic gaps between populations in a local area rather than by means of the phylogeny, cohesion, or reproductive isolation that are responsible for these gaps (Mallet, 1995). In a local area, separate species are recognized if there are several clusters separated by multilocus phenotypic or genotypic gaps. A single species (the null hypothesis) is recognized if there is only a single cluster in the frequency distribution of multilocus phenotypes and genotypes. The genotypic gaps may be entirely vacant, or they may contain low frequencies of intermediate genotypes, or hybrids (**Figure 1**). The definition is useful because one avoids tautological thinking: hypotheses about speciation or phylogeny of taxa become independent of assumptions about the nature of reproductive isolation or phylogeny underlying the taxa studied.

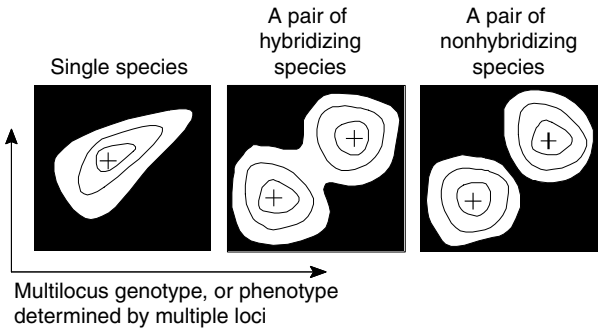


Figure 1 Genotypic or genomic cluster criterion for species. A sample of individuals is made at a single place and time. Numbers of individuals are represented by the contours in multidimensional genotypic space. Peaks in the abundance are represented by "+." Two species are detected if there are two peaks in the genotypic distribution (right, bimodal distribution; also see Jiggins and Mallet, 2000). Otherwise the null hypothesis of a single species is not rejected (left, unimodal distribution). Note: the axes represent multidimensional morphological/genotypic space, not geographic space.

Bayesian statistical procedures, called "assignment tests," are today regularly used to estimate the numbers of genotypic clusters from individual multilocus genotype data (Pritchard *et al.*, 2000; Falush *et al.*, 2003; Huelsenbeck and Andolfatto, 2007): if such genetic clusters are statistically distinguishable in sympatry, most would agree that one is dealing with separate species.

Genotypic clusters are neither profound nor original; the author traces the earliest use of the method to Darwin's (1859) morphological gap criterion of species (*see* Darwin's Morphological Species Criterion), although earlier sources probably employ similar ideas since acceptance of evolution is not required to identify species on the basis of character distribution. Many similar proposals have been made more recently (Simpson, 1937; Hutchinson, 1968; Sokal and Crovello, 1970; Avise and Ball, 1990; Cohan, 1994; Smith, 1994). The approach is essentially the same in most taxonomic decisions (*see* Taxonomic Practice), in the phenetic concept (*see* Phenetic Species Concept), or in a practical application of the biological species concept (see below). Multilocus genotypic clusters are almost universally applied as a criterion of speciation in theoretical models of sympatric speciation (e.g., Dieckmann and Doebeli, 1999; Gavrilets and Waxman, 2002; Kondrashov and Mina, 1986; Kondrashov and Kondrashov, 1999): in these models, a bimodal genotypic distribution evolves via reproductive isolation, but it is the demonstration that a pair of genetically divergent groups of individuals emerge from a single population, rather than the mere existence of hybrid inviability or mate choice, that is required for evidence of speciation.

This general use of direct morphological or genetic criteria in the definition of species, as opposed to reproductive or phylogenetic inferences made from such data, has apparently lacked widespread support due to the supposed need for a separation between concept and taxonomic criterion (Mayr, 1970). The author's intention was to justify the Darwinian and practical taxonomists' species definition statistically and in terms of genetics, rather than to enforce, as in DNA barcoding, the use of genes instead of morphology to define species. Most genotypic cluster species can be recognized morphologically; for example, minor pattern elements in *Papilio* butterflies can be used to unite the various polymorphic forms; however, with abundant molecular marker data, we could easily use the genetic version of the criterion to sort actual specimens.

There is, of course, every reason to conclude after seeing a male butterfly mating with an unlike female that they belong to the same species. However, because hybridization does occur occasionally between forms normally thought of as different species, one is not so much using the mating behavior itself to define species as inferring that such matings are common enough to cause homogenization of gene frequencies between such males and females. We infer that, if we were to analyze their genomes, the two forms would have similar genetic characteristics apart from those few genes determining the polymorphism or sexual dimorphism, that is they would belong to the same genotypic cluster. Instead of reproductive compatibility being the primary criterion of species, as in the

biological species concept, we can turn the argument on its head and infer from limited data on reproductive compatibility that a single genotypic or genomic cluster is probable.

Asexual forms, unclassifiable under the interbreeding concept, and arbitrarily definable at any level under concepts depending on phylogeny, can be clustered and classified as a genotypic clusters in exactly the same way as sexual species. The precise taxonomic level of species clustering for asexuals is somewhat arbitrary, as in the phylogenetic concepts, but at least the method acknowledges this arbitrariness rather than purporting to use some higher evolutionary principle. Many asexual forms such as bdelloid rotifers have clearly distinguishable species taxa (Hutchinson, 1968; Fontaneto *et al*., 2007), probably due to ecological selection for distinct characteristics. In bacteria, competition is thought to structure largely asexual populations with occasional promiscuous horizontal gene transfer into recognizable genetic clusters (Cohan, 1994, *also see* Ecological Species Concept). Thus, reproductive isolation may help but is not required for genotypic clustering.

Critics have argued that the genotypic cluster criterion in sexual species is nothing other than a gene flow concept of species under a different guise. This is true for one specialized interpretation of gene flow in sexual populations. If we define gene flow as successful or effective, as opposed to actual input of genes, we can see that a gene flow criterion becomes similar to the genotypic cluster criterion. To find whether a hybridization or gene flow event is successful, we must either follow the fate of every gene through all possible descendants for all time, or we may examine the genotypic state of a population and determine if genes from one form are mixed randomly with genes from another form. Looking for random associations of genes within genotypes in the genotypic cluster approach will be methodologically the same as a genotypic analysis to determine whether a population is interbreeding, but the latter requires additional assumptions. The genotypic cluster criterion in sexual species could thus be looked upon simply as a practical application of the biological species concept. However, one may prefer the genotypic cluster criterion to the interbreeding concept, if only because its name emphasizes that the definition is character-based, rather than actually based on interbreeding, and is thus applicable to asexuals as well as sexual species.

If a single geographic race, which previously intergraded at all its boundaries with other geographic races, were to split into two forms that coexist as separate genotypic clusters, we could have a situation where the original polytypic species became paraphyletic. The new species has been derived from only one of the component subspecies. Thus, paraphyly of species would have to be recognized as a possibility under this definition, as in both interbreeding and diagnostic concepts.

The Unreality of Species in Space and Time: Races Versus Species

Geographic races often form clusters differing at multiple loci from other races in the same species. The interbreeding concept or genotypic cluster criterion can be used to justify a classical polytypic species if the various geographic races meet at contact zones that contain abundant intermediates (hybrid zones). We sample multilocus genotypes or phenotypes in local areas of overlap and determine whether a single peak (one species, i.e., abundant hybrids) or two peaks (two different species, i.e., rare hybrids) are evident in the local genotypic distribution (**Figure 1**). Hybridization may occur, but if it is so rare that character and genotypic distributions remain distinctly bimodal in zones of overlap, we usually classify them as separate species, even under the interbreeding concept.

Although this spatial extension of the local species is practical to apply to any pair of forms in contact, it is unlikely to lead to general agreement. The problem is that hybrid zones can be very narrow and may have separate forms that are highly distinct at multiple characters or loci, in spite of complete unimodal blending in local areas of overlap. Even adherents of the interbreeding concept are reluctant to lump such geographic forms within the same species. Examples include North American swallowtail butterflies (*Papilio glaucus*/*Papilio canadensis*; see Hagen *et al*., 1991) and European toads (*Bombina bombina*/*Bombina variegata*; see Szymura, 1993).

An even worse problem is found in ring species, which form a continuous band of intergrading subspecies, but whose terminal taxa may be incompatible, and overlap without intergrading. Commonly cited examples are the herring gulls and lesser black-backed gulls (*Larus argentatus* complex; Mayr, 1970). Similarly, although most hybrid zones between European *Bombina* are unimodal, the same pair of taxa may have bimodal genotypic distributions in other zones of overlap (Szymura, 1993). Thus, geographic forms may be apparently conspecific in some areas, but overlap as separate species in other areas. Finally, if distinct populations are geographically isolated and there is no area of overlap, one cannot disprove the null hypothesis of same species under interbreeding or genotypic cluster criteria, but biologists are reluctant to unite such populations if they are very divergent. Laboratory hybridization could be tried, but many overlapping species are known to hybridize freely in captivity, while remaining separate in nature. There are good examples even in our closest relatives, the great apes, for instance the bonobo (*Pan paniscus*) versus the chimpanzee (*Pan troglodytes*), and among the gorillas (Uchida, 1996), but similar decisions must be made in almost any animal or plant group.

The problem of extending local species criteria spatially is due to the way in which spatially separated lineages diverge: time since population divergence is usually somewhat correlated with geographic distance. Paleontologists face a similar temporal problem when classifying fossils in different strata. Evolutionary rates may vary, but all lineages must ultimately be continuous, so there may be no very logical place to put a species boundary in time any more than there is in space. Paleontologists, like neontologists, use operational species on the basis of morphological gaps between taxa from the same and different time periods (Simpson, 1937; Smith, 1994).

These difficulties show why there is no easy way to tell whether related geographic or temporal forms belong to the same or different species. Species gaps can be verified only locally and at a single point in time. One is forced to admit that Darwin's insight is correct: any local reality or integrity of species is greatly reduced over large geographic ranges and time periods (*see* Mayr (1970) *and* Biodiversity in Space and Time).

The Importance of Species Concepts for Biodiversity and Conservation

Traditional: Species as Real Entities

Different species concepts seek to define species in mutually incompatible ways. Thus, a monophyletic species concept seems not very useful to evolutionary biologists because of difficulties with multiple gene genealogies and paraphyletic remnants. In contrast, the interbreeding concept and other concepts incorporating biological processes of species maintenance (e.g., recognition, ecological, evolutionary, and cohesion concepts) suffer in the eyes of phylogenetic systematists because they lack phylogenetic coherence and produce paraphyletic taxa, or worse. If we were to allow the basal unit of our taxonomy to incorporate paraphyly, it would be harder to justify a strict adherence to monophyly at other taxonomic levels. It is beyond the scope of this article to resolve these difficult issues, but these conceptual conflicts fuel the continued debate, and also highlight the fact that if species are indeed real, objective biological units, their unifying reality has been extremely difficult to verify.

Many ecological and biodiversity studies of actual organisms ignore these difficulties, and assume that species are objectively real basal units. Thus, in ecology, we have theories of global species diversity. In conservation, we have the Endangered Species Act in the US, which prescribes the conservation of threatened taxa we call as species. Populations not viewed as species, particularly putative hybrid taxa (like the red wolf, *Canis rufus*, of the southeast US), became seen as less valuable, even if rare. How do we recognize that a taxon is hybridized? Obviously, to be a hybrid, it must be a mere intergrade between two, real, objectively identifiable entities. The Endangered Species Act viewed species as important real conservation units and hybrids as unimportant. It did this because it incorporated the species concept in vogue at the time of its enactment, that is the biological species concept, in which hybridization is seen as an unnatural breakdown in isolating mechanisms (Mayr, 1970).

Alternatives: Genetic Differences More Valuable than Species Status

In 1990, the hybrid policy of the Endangered Species Act was rescinded. However, although some hybrid populations are eligible to be viewed as valuable in conservation, others, especially hybrids with invasive taxa can be a problem for conservation. In spite of much discussion, no replacement policy about hybrids has been forthcoming. There is undoubtedly a greater realization today that other levels in the taxonomic hierarchy are important elements of biodiversity. The diagnostic concept of species, while claiming to support the basal, objective nature of species, may be useful in that it encourages recognition of important biodiversity at a lower level than the interbreeding concept, in this case as subspecies within polytypic species. Conservation geneticists today often advocate the conservation importance of "evolutionary significant units," "management units," or "stocks" (a fisheries term) defined below the species level by means of various criteria along the continuum of genetic differentiation at molecular markers (Moritz, 1994). But the reality of spatiotemporally extended species continues to elude us, and biodiversity in terms of numbers of species, including endangered species, remains difficult to measure. The author argues that this will always be the case. Because populations evolve, cohesion over large regions of space and time will continue to break down. If this is so, then it might seems best to adopt some other measure of conservation value that relies purely on the degree of genetic differentiation, for instance, at molecular genetic markers.

Species Differences as Ecologically Important Markers

However, there are many who would oppose using molecular divergence as a measure of biodiversity. Species of sexual macroorganisms within a local area such as a nature reserve are, for the most part, easily and objectively identifiable using morphology, behavior, genetics, or phylogeny. A pair of similar species normally must be ecologically distinct to coexist. Sexual species will need some prezygotic isolation, so their mating behavior must also be different. Thus, counting species in a local area makes considerable ecological sense, and conserving species diversity in a local area would conserve actual ecological and behavioral diversity. Behavioral and morphological differences that differ among species seem more valuable evolutionarily, as well as having more interest to conservationists, than the probably mainly neutral differences at molecular genetic markers.

Biodiversity in Space and Time

As we have seen, this local view breaks down when we try to apply the term "species" over large areas or geological timescales. In some cases, there is an excellent homogeneity over large areas; for example, the painted lady butterfly (*Vanessa cardui*) and the barn owl (*Tyto alba*) have a virtually worldwide distribution and look similar everywhere. Other species are not so homogeneous: the familiar mallard (*Anas platyrhynchos*) group of ducks is as widespread, but has become highly differentiated into some 18 or so forms in far-flung outposts of the world. Exactly how many mallard populations are good species, and how many are races, or indeed, how many races there are in total, is a matter of taste. Current authorities recognize about 10 species, but there might easily be five or 15 in alternative treatments. One of the forms, the Mexican duck, *A. platyrhynchos diazi*, is threatened with hybridization by the true mallard, *A. platyrhynchos platyrhynchos*, which has been expanding from the north, and the American black duck (*Anas rubripes*) also hybridizes with the mallard, but appears to resist hybridization somewhat better than the Mexican form – hence its species status.

Faced with these difficulties, should we worry about whether or not we have a definite species when conserving endangered taxa over large areas? In the author's view, probably not, and criteria of distinctiveness, or ecological role should be more important on a case-by-case basis. Whatever one's answer to this question, it does not seem sensible to rely on a unitary spatiotemporal reality of species as a guide. We might conform to the current trend for taxonomic inflation, and

upgrade the Mexican duck to a separate species instead of a subspecies, but this should ideally have little effect on our view of its conservation value: this has led to no actual change in our knowledge of biological characteristics that affect conservation value. Most conservationists now agree that the former fetish for species-level legislation was a mistake. Conservation and legislation now generally recognize that living, evolving populations form fractal continua with species and genera over time and space, rather than attempting to rely on spurious fundamental units.

Species can still be fundamental units of local biodiversity, but they have this clarity only in a small zone of time and space, and so species counts become less and less meaningful as larger and larger areas are covered. Taxonomists might come to nominalistic agreements on a case-by-case basis, but even this shows little sign of happening yet. Ecological theory, as well as conservation and biodiversity studies must however recognize that species counts over large expanses of space and time represent only a sketchy measure of biodiversity, a measure that owes more to taxonomic and metaphysical fashion than to science. Yet conservation still depends on lists of endangered species at both local and global levels. It would seem necessary to have either a better way than species lists to estimate conservation value, or at the very least a more stable species criterion less prone to taxonomic inflation. However, it is the bleak truth that agreement on this matter has not yet been achieved.

See also: Differentiation. Speciation, Process of. Subspecies, Semispecies, Superspecies

References

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. London: Chapman & Hall.

Avise JC and Ball RM (1990) Principles of genealogical concordance in species concepts and biological taxonomy. In: Futuyma DJ and Antonovics J (eds.) *Oxford Surveys in Evolutionary Biology*, vol. 7, pp. 45–67. Oxford: Oxford University Press.

Baum DA and Shaw KL (1995) Genealogical perspectives on the species problem. In: Hoch PC and Stephenson AG (eds.) *Experimental and Molecular Approaches to Plant Biosystematics, Monographs in Systematic Botany from the Missouri Botanical Garden*, vol. 53. St Louis, MO: Missouri Botanical Garden.

Claridge MF, Dawah HA, and Wilson MR (eds.) (1997) *Species: The Units of Biodiversity*. London: Chapman & Hall.

Cohan FM (1994) The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. *American Naturalist* 143: 965–986.

Coyne JA and Orr HA (2004) *Speciation*. Sunderland, MA: Sinauer Associates.

Cracraft J (1989) Speciation and its ontology: The empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: Otte D and Endler JA (eds.) *Speciation and its Consequences*, pp. 28–59. Sunderland, MA: Sinauer Associates.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 1st edn. London: John Murray.

Davis JI (1997) Evolution, evidence, and the role of species concepts in systematics. *Systematic Botany* 22: 373–403.

de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation. A conceptual unification and terminological recommendations. In: Howard DJ and Berlocher SH (eds.) *Endless Forms. Species and Speciation*, pp. 57–75. New York: Oxford University Press.

Dieckmann U and Doebeli M (1999) On the origin of species by sympatric speciation. *Nature* 400: 354–357.

Dobzhansky T (1937) *Genetics and the Origin of Species*. New York: Columbia University Press.

Ehrlich PR and Raven PH (1969) Differentiation of populations. *Science* 165: 1228–1232.

Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton, NJ: Princeton University Press.

Falush D, Wirth T, Linz B, *et al.* (2003) Traces of human migrations in *Helicobacter pylori* populations. *Science* 299: 1582–1585.

Feder JL (1998) The apple maggot fly, *Rhagoletis pomonella*: Flies in the face of conventional wisdom. In: Howard DJ and Berlocher SH (eds.) *Endless Forms. Species and Speciation*, pp. 130–144. New York: Oxford University Press.

Fontaneto D, Herniou EA, Boschetti C, *et al.* (2007) Independently evolving species in asexual bdelloid rotifers. *PLoS Biology* 5: 914–921.

Gavrilets S and Waxman D (2002) Sympatric speciation by sexual conflict. *Proceedings of the National Academy of Sciences USA* 99: 10533–10538.

Hagen RH, Lederhouse RC, Bossart JL, and Scriber JM (1991) *Papilio canadensis* and *P. glaucus* (Papilionidae) are distinct species. *Journal of the Lepidopterists' Society* 45: 245–258.

Harrison RG (1998) Linking evolutionary pattern and process. The relevance of species concepts for the study of speciation. In: Howard DJ and Berlocher SH (eds.) *Endless Forms. Species and Speciation*, pp. 19–31. New York: Oxford University Press.

Hennig W (1966) *Phylogenetic Systematics*. Urbana, IL: University of Illinois Press.

Hey J (2001) *Genes, Categories, and Species. The Evolutionary and Cognitive Causes of the Species Problem*. New York: Oxford University Press.

Howard DJ and Berlocher SH (eds.) (1998) *Endless Forms. Species and Speciation*. New York: Oxford University Press.

Hudson RR and Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution* 56: 1557–1565.

Huelsenbeck JP and Andolfatto P (2007) Inference of population structure under a Dirichlet process model. *Genetics* 175: 1787–1802.

Hutchinson GE (1968) When are species necessary? In: Lewontin RC (ed.) *Population Biology and Evolution*, pp. 177–186. Syracuse, NY: Syracuse University Press.

Isaac NJB, Mallet J, and Mace GM (2004) Taxonomic inflation: Its influence on macroecology and conservation. *Trends in Ecology and Evolution* 19: 464–469.

Jiggins CD and Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution* 15: 250–255.

Kondrashov AS and Kondrashov FA (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature* 400: 351–354.

Kondrashov AS and Mina MV (1986) Sympatric speciation: When is it possible? *Biological Journal of the Linnean Society* 27: 201–223.

Levin DA (1979) The nature of plant species. *Science* 204: 381–384.

Lotsy JP (1917) La quintessence de la théorie du croisement. *Archives Neerlandaises des Sciences Exactes et Naturelles Séries III* 3: 351–353.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology* 46: 523–536.

Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology and Evolution* 10: 294–299.

Mallet J (2004) Poulton, Wallace and Jordan: How discoveries in *Papilio* butterflies initiated a new species concept 100 years ago. *Systematic Biodiversity* 1: 441–452.

Mallet J (2008) Hybridization, ecological races, and the nature of species: Empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B* 363: 2971–2986.

Mallet J (2010a) Why was Darwin's view of species rejected by 20th century biologists? *Biology and Philosophy* 25: 497–527.

Mallet J (2010b) Group selection and the development of the biological species concept. *Philosophical Transactions of the Royal Society B* 365: 1853–1863.

Mayr E (1970) *Populations, Species, and Evolution*. Cambridge, MA: Harvard University Press.

Mayr E (1982) *The Growth of Biological Thought. Diversity, Evolution, and Inheritance*. Cambridge, MA: Belknap.

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution* 9: 373–375.

Paterson HEH (1985) The recognition concept of species. In: Vrba ES, (ed.) *Species and Speciation Transvaal Museum Monograph*, vol. 4. Pretoria: Transvaal Museum.

Poulton EB (1904) What is a species? *Proceedings of the Entomological Society of London* 1903: lxxvii–cxvi.

Pritchard JK, Stephens M, and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
Ridley M (1989) The cladistic solution to the species problem. *Biology and Philosophy* 4: 1–16.
Robson GC (1928) *The Species Problem. An Introduction to the Study of Evolutionary Divergence in Natural Populations*. Edinburgh: Oliver and Boyd.
Simpson GG (1937) Patterns of phyletic evolution. *Bulletin of the Geological Society of America* 48: 303–314.
Simpson GG (1951) The species concept. *Evolution* 5: 285–298.
Smith AB (1994) *Systematics and the Fossil Record. Documenting Evolutionary Patterns*. Oxford: Blackwell Scientific.
Sokal RR and Crovello TJ (1970) The biological species concept: A critical evaluation. *American Naturalist* 104: 107–123.
Szymura JM (1993) Analysis of hybrid zones with *Bombina*. In: Harrison RG (ed.) *Hybrid Zones and the Evolutionary Process*, pp. 261–289. New York: Oxford University Press.
Templeton AR (1998) Species and speciation. Geography, population structure, ecology, and gene trees. In: Howard DJ and Berlocher SH (eds.) *Endless Forms. Species and Speciation*, pp. 32–43. New York: Oxford University Press.
Uchida A (1996) What we don't know about great ape variation. *Trends in Ecology and Evolution* 11: 163–167.
Van Valen L (1976) Ecological species, multispecies, and oaks. *Taxon* 25: 233–239.
Wallace AR (1865) On the phenomena of variation and geographical distribution as illustrated by the Papilionidae of the Malayan region. *Transactions of the Linnean Society of London* 25: 1–71.
Yang Z and Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Philosophical Transactions of the Royal Society USA* 107: 9264–9269.
Zhang C, Zhang D-Y, and Yang Z (2011) Evaluation of a Bayesian coalescent method of species delimitation. *Systematic Biology* 60: 747–761.

Species Distribution Modeling

Jane Elith, The University of Melbourne, Parkville, VIC, Australia
Janet Franklin, Arizona State University, Tempe, AZ, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Abiotic Nonliving components of ecosystem – the chemical and physical environment.
Biotic Living components of ecological community – organisms, species.
Covariate A variable that might predict the response of interest; also termed *predictor* or *explanatory variable.*
Extent A component of spatial scale: the size of the study domain.
Extrapolation Estimating the value of a response variable at unmeasured locations by extending a model to new places or times.
Geospatial data Digital representations of data about Earth, including geographic locations, stored in a geographic information system (GIS).
Grain A component of spatial scale: sample resolution, the size of a single observation.
Proximal Direct or causal – for example, a factor directly related to species distribution and abundance (in contrast with distal factors that may be surrogates or proxies for causal factors).

Introduction

Species distribution modeling (SDM) is a methodology – a set of procedures, definitions, and techniques – built on a foundation of core ecological (Whittaker *et al.*, 1973) and biogeographical (Holdridge, 1947) concepts about the relationship between species distributions (or other biotic response variables describing aspects of biodiversity) and the physical (abiotic) environment. SDMs are quantitative, empirical models of species–environment relationships typically developed using species location data (abundance, occurrence) and those environmental variables thought to influence species distributions. Sometimes biotic information (other species, vegetation types, etc.) is included as a predictor variable. Provided appropriate data are available, SDMs can be applied to any taxa, including marine, terrestrial, and freshwater species, and at any grain and extent. Although the models can provide insights into factors influencing patterns of biodiversity even without predicting to maps, often the models are extended to making spatial predictions using maps of the environmental predictors. The practice of SDM has flourished over the past four decades in part because geospatial environmental data have become so widely available. The spatial predictions as well as the models themselves can inform both basic and applied biodiversity research.

Some of the earliest uses of SDMs focused on the insight and explanation that the models or their mapped predictions can provide, and this role continues to be important to this day. For instance, Leathwick and Austin (2001) explored competitive interactions between tree species in New Zealand's forests in a dataset in which geographic disjunctions provided the equivalent of a natural removal experiment. Regression models relating plot-based estimates of tree density to environmental and biotic covariates indicated substantial competitive displacement (e.g., **Figure 1**). This was one of the first pieces of statistical evidence for competition effects under field conditions. A second example of this theory-oriented use of SDMs is the recent focus on examining evidence for, and methods for detecting, niche shifts or niche conservatism under processes of invasion or speciation (e.g., Peterson, 2011).

Spatial (mapped) predictions from SDMs are also used as inputs to other analyses, particularly in conservation planning and reserve design. For example, Linke *et al.* (2008) modeled the distributions of 400 freshwater benthic invertebrate taxa and predicted to rivers in southern Australia. These predictions provided the information necessary for estimating the irreplaceability or conservation value of river catchments. SDMs are also used for environmental impact assessment and land-use planning to determine suitable locations for habitat restoration and species reintroductions and to predict the risk of species establishment in new areas, including invasive species, pathogens, and disease vectors. One of the fastest-growing uses of SDM for risk analysis is to predict the potential impact of projected patterns of anthropogenic climate change on

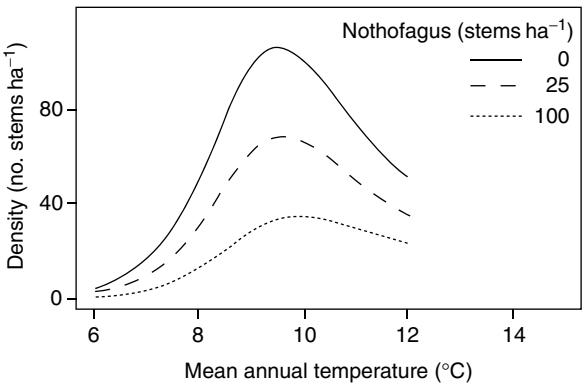


Figure 1 Density of the tree *Weinmannia racemosa* in relation to mean annual temperature, for 0, 25, and 100 stems per hectare of a competitor, *Nothofagus* (four species, combined). The values of all other environmental covariates were set to fixed levels. Note this is an example of a partial dependence plot (also called a *fitted function* or a *response curve*).

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00318-X

species distributions (Pearson and Dawson, 2003). This requires extrapolation from empirical models into novel environments, as does predicting species invasions, and this important issue will be discussed later (*see* Predicting to New Times and Places).

SDMs are applicable to varying geographic scales and levels of biological organization. They have been used to study the patterns and causes of species distributions and community properties such as species richness at large spatial scales – for example, for macroecological questions. Evolutionary biologists use SDMs to explore the geographic structure of variation within and between taxa in phylogeographic studies, and ecologists test hypotheses about geographic variation in species traits (Kozak *et al.*, 2008; Dobrowski *et al.*, 2011).

Recent works including Guisan and Zimmerman (2000), Pearson (2010), Elith and Leathwick (2009a), and Franklin (2010) have described SDM methodology (species and environmental data requirements, modeling methods, model evaluation and uncertainty) in considerable detail with numerous pointers to the burgeoning literature on all aspects of SDM. This article will focus on the role of SDMs in quantifying species–environment relationships (response curves), the ecological and geographical theory underpinning SDM model formulation, issues of sampling and data fundamental to constructing models that are well suited for strong inference and prediction, and what to do when you do not have that perfect dataset (minimum requirements, robust modeling approaches). We will briefly discuss applying the SDM framework to biodiversity variables other than species, such as diversity measures or community composition, and address the critical issues involved when extrapolating predictions from SDMs to novel environments. Despite the importance of relevant and probing evaluations of model fit and predictive performance (accuracy), we only mention these topics briefly because their details will vary with the application, and they are covered in detail in the cited works.

Some Fundamentals

Niches

An important piece of ecological theory underpinning SDMs is niche theory. Here we will rely heavily on a recent paper by Colwell and Rangel (2009) that presents the theory in ways particularly relevant to SDMs, but many other authors including Peterson, Soberón, Chase and Leibold, Pulliam, Kearney, and Ackerly have made important contributions to the published discussion as summarized and referenced in Franklin (2010). It is G.E. Hutchinson's view of the niche that provides most insight and utility for SDMs. Hutchinson contended that the niche is an attribute of a species or population, with axes defined by those conditions (environmental attributes) and resources necessary for the persistence of the species. Thus the niche exists in niche space, defined by axes corresponding to environmental attributes (Hutchinson's "hypervolume") and without reference to the geographical space (biotope) in which species live. The biotope is the physical expression of niche axes. Hence there is an interesting and conceptually important duality or interplay between niche space and the biotope (Colwell and Rangel, 2009; Elith and Leathwick, 2009a). Niche and biotope space are only partially reciprocal – that is, the correspondence of the two is not directly 1 : 1. Any one point in niche space might have many (or no) realizations in the biotope, whereas any one point in the biotope will link to only one point in niche space (**Figure 2**). This duality is applicable to any niche. The fundamental niche is that set of conditions that allow persistence (i.e., where population growth rate is not negative). A *potential distribution* is that subset of a species' fundamental niche that actually exists in the biotope at the current time. This need not be fully occupied and usually is not. Most species modelers view the realized niche as the physical expression of the fundamental niche, typically a subset of it due to biotic interactions, dispersal limitations, disturbances, and the difference between potential distributions and fundamental niches.

This all seems clear, but questions arise once we start to try to work out exactly what we are modeling when we have a set of data in hand and a model to use. We will cover some of these in following sections; also see Franklin (2010, Section 3.2.1). Suffice to say these questions have caused some modelers to decide to call their models *species distribution models*, emphasizing the origin of the data, focusing on predicting distributions, and avoiding a commitment to niches, whereas others prefer the term *ecological niche models* (ENMs) because niches are what they are most interested in. Certainly, all the applications that require predictions to new geographic regions (e.g., predicting potential distributions of invasive species) or new times (e.g., paleodistributions or under future climates) rely strongly on the conceptual links to niches.

The niche space–biotope duality has an important attribute: models fitted to niche axes are ignorant of physical distances between sites (**Figure 2**). The model uses identified relationships in niche space to predict to sites in the biotope. Any geographic patterning that occurs in the predictions (e.g., higher probabilities of occurrence being clustered together) results from the geographic patterning of the niche axes. This patterning is known as *spatial autocorrelation* (SAC; Legendre, 1993), meaning the dependencies among values of a given variable at different geographical separations. It is a common feature of environmental conditions on Earth and is the cause of many of the patterns we observe in nature. But geographic clumping of species is not always caused by SAC of environment: other factors including disturbance history, biotic interactions (positive and negative), life history traits (e.g., dispersal ability), and social behavior of individuals can play a role – these are all features of the biotope (though some could also be defined in niche space).

The topic of SAC and its role in and effect on species distributions is too complex to cover well here. There is an increasing body of relevant literature on topics, including how to measure SAC in model residuals, the role of geographic predictors, and models for jointly characterizing geographic and environmental effects (Rangel *et al.*, 2006; Dormann *et al.*, 2007; Franklin, 2010). The latter include autoregressive methods, generalized linear mixed models (GLMMs), and generalized estimating equations. A problem is that it is always difficult to distinguish between geographic and environmental effects using the observational data typically available for SDMs. Models fitted to only geographic predictors – for

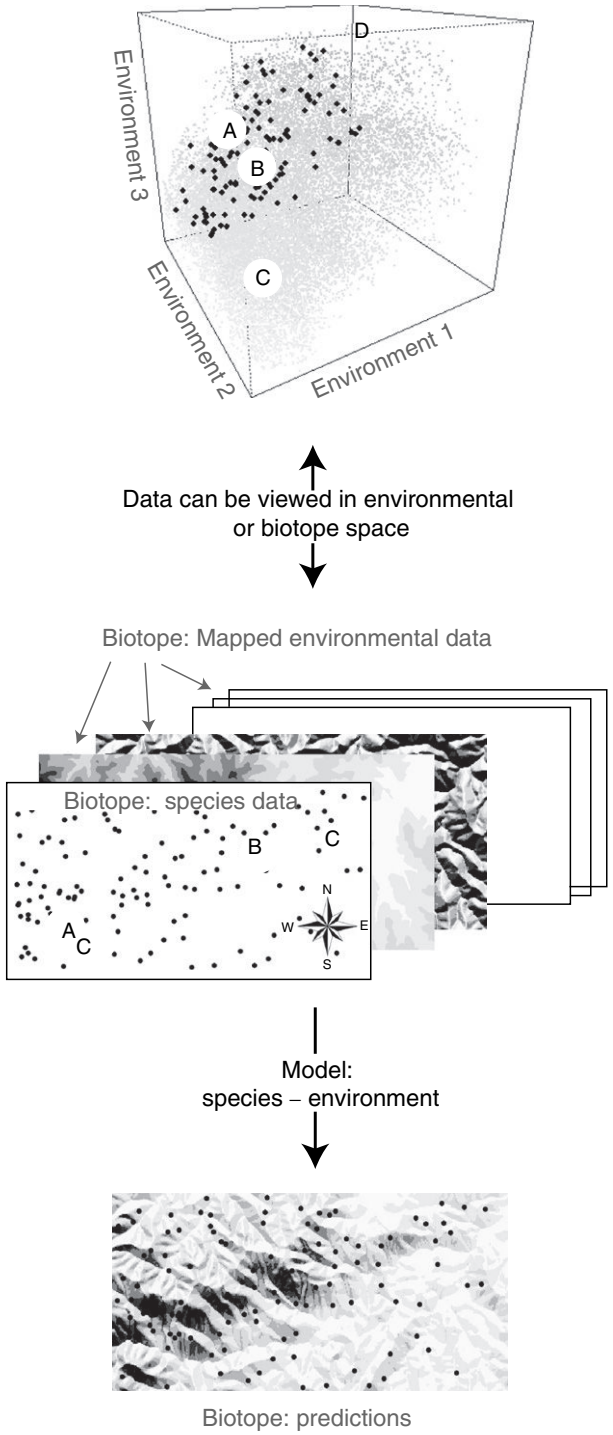


Figure 2 The relationship between niche (environmental) and biotope (geographic) space, showing (i) a point in environmental space might have more than one realization in the biotope (point C); (ii) a point in environmental space may not occur in the biotope (point D); (iii) points can be nearby in environmental space and far apart in the biotope (A and B); (iv) a model can create patterned predictions due to spatial autocorrelation (patterning) in the biotope realization of environment.

example, to two-dimensional trend surfaces with latitude and longitude as axes – have been shown in some studies to predict as accurately as environment-based models within sampled regions. This is not surprising if species have severe dispersal limitations and occur in highly fragmented landscapes, or if suitable maps of relevant environmental predictors are not available. It does not mean that environment is not important, merely that the signal is hard to determine from geographically restricted samples and can be swamped by geographic factors. A model with only geographic predictors is not useful for predicting to new times or places.

The Shapes of Species Responses to Environment

According to niche theory, the shape of the curve relating species abundance or fitness to an environmental gradient (**Figure 1**) is unimodal (hump-shaped) and describes upper and lower tolerance limits as well as an optimum (central tendency). Although early work depicted response curves of species in a community as Gaussian with breadth limited by competition, real communities do not usually show this idealized pattern. However, species in a community often do show a pattern of turnover along a gradient, and this pattern is assumed in many multivariate community analysis methods. There is empirical support for skewed or even nonunimodal response curves where species reach physiological limits of growth and survival at the harsh end of a gradient or are limited by competition at the ameliorative end (or both). Research in this area has focused on estimating response curve shape, niche breadth (specialization), and marginality (the location and density of optima).

It is also widely acknowledged that response curves can take almost any form under different circumstances. For example they might be linear if there is a strong limiting factor, or if only a portion of the gradient is sampled. They might be multimodal in the presence of a strong competitor. They may be complex if there are interactions among variables in the high-dimensional environmental hypervolume, especially if predictors used in modeling are surrogates or proxies for the true resource gradients. If complex response functions are fitted, they can be evaluated in terms of possible causes (incomplete sampling of the gradient and so forth) and implications for use of the model (especially if it is to be used for projection). Unbounded response curves – for example, a positive linear or monotonic response to a temperature variable – might give nonsensical projections – for example, the distribution of climatically suitable habitat under a scenario of global warming.

Geographical Models and Theory for Predictor Variables

Even though modern uses of SDM rely as much on conceptual models of geographical data as on ecological theories, this relationship to geographical data models is less well articulated in the SDM literature. There are two contrasting conceptual models used for representing geographical data (Goodchild, 1994). In the "field" view, the precise value of geographical variables can be measured at any location, and geography can be represented as a multivariate or vector field.

Although this is a useful framework for many types of modeling, it does not easily accommodate all geographic phenomena that interest spatial scientists. In the contrasting "entity" view, geography is represented as empty space containing various discrete objects or features (entities), such as mountain peaks, atmospheric low-pressure troughs, fire scar perimeters, or species occurrences. In SDM, one is often accomodating both an entity view of species abundance or occurrence and a field view of environmental predictors, as well as developing models that represent species distributions as a field, either continuous or discrete, of occurrence probability or habitat suitability. Because complex phenomena such as habitat suitability or species richness can never be represented precisely in digital geospatial databases, geographic representations of model error or uncertainty are particularly important for SDM.

There are also some conceptual or theoretical frameworks available to guide the choice of environmental (abiotic) factors affecting species distributions, so we know generally what kind of GIS data we should be looking for. Mackey and Lindenmayer (2001) describe the primary environmental regimes of light, water availability, heat sum, and nutrients, and they outline a spatially nested hierarchy of processes controlling their distribution for terrestrial environments – from global distributions of solar radiation (controlling climate) and lithologies of the geologic substrate, to the local topography moderating heat and moisture patterns at the mesoscale, and finally plant canopies altering those patterns at the microscale. Although this framework has been applied most effectively for plant species, Mackey and Lindenmayer posit that it applies to all types of organisms. A different tradition in wildlife ecology describes the factors determining habitat suitability (for animals) as access to food, water, and cover (nest sites, predator avoidance). The spatial hierarchy described suggests that SDMs using only broad-scale predictors, such as climate factors that vary slowly in space, are likely to capture the broad-scale features of species distributions – for example, ranges – but not the finer-scale aspects of habitat suitability as determined by variables such as terrain and land cover for terrestrial species, nutrient availability for marine species, and flow rates for freshwater species (*see* Environmental Data).

In a Perfect World

We present a hypothetical example to sketch out idealized modeling approaches and provide a reference point for dealing with typically available data (in the real world). Imagine a species of cypress pine, *Cupressus chimera*, that is easily distinguished from all other species and always found if an occupied site is visited. This species is endemic to Uruguay and Brazil, has had opportunity to disperse to all suitable habitats within South America, and tends to be competitively superior to other local trees. Surprisingly, its habitat has been undisturbed by humans or historic events, including volcanic eruptions, landslides, and fires. A team of ecologists wanting to understand its distribution and its varying density across the landscape has time and resources to gather much data. They start by talking to experts and gathering available information on the environmental correlates of related species. They familiarize themselves with the environmental gradients in the region and gather relevant GIS data. They are interested in topographic, substrate, and climatic effects on its distribution and decide to model the whole distribution with 250-m × 250-m grid cell sizes. Their climate scientist colleagues say that climate could conceivably vary at that resolution in this region. Soil scientists have unusually comprehensive data on soil depth, fertility, and pH at that resolution. Terrain experts have a 30-m digital elevation model of the area from which they can calculate relevant attributes including slope, aspect, and topographic wetness indices. They prefer to calculate these at fine scales because these best represent terrain processes but have agreed to summarize them to 250 m.

The ecologists plan their fieldwork. They choose four major gradients of temperature, rainfall, soil fertility, and hillslope position and aim to sample representatively across those while ensuring that they cover all geographic regions. They plan to visit 500 sites and, in 250-m × 250-m quadrats, record the abundance of *C. chimera*. They first conduct a pilot study to ascertain the environmental limits of the cypress. They will sample a little beyond these, but not extensively so – they do not want to waste time sampling totally unsuitable habitat. They choose their 500 sites (ensuring a minimum distance between them of 1 km), take geographic positioning systems (GPSs) to the field, and visit every site regardless of its distance from roads.

Once ecologists are back in the office, the modeling begins. First, the GIS data are prepared using models and data summaries that ensure the predictor information is relevant to the species – for instance, the rainfall data are summarized to emphasize water availability in the driest months because the species is sensitive to water stress. Collinearity between the predictors is assessed and dealt with. Checks of the species data reveal they are Poisson distributed, so the ecologists decide to use Poisson regression models. In a first exploratory phase of model building, boosted regression trees are used to reveal important predictor variables and interactions between them. No substantial interactions appear. Generalized additive models (GAMs) are used for final model fitting, and confidence intervals are estimated. The model fit is assessed, and the residuals are tested for SAC. The ecologists summarize the models, including the selected variables, fitted functions, and predictions. They discuss the results with experts and iterate the models a few times, testing the effect of new predictors and exploring the leverage of certain unusual observations. Since the final model is to be used to understand the correlates of distribution and to predict within the sampled regions, it is documented with attention to interpretation of the fitted functions and cross-validated evaluations of its predictive ability using deviance, correlation, and root mean square errors for statistical summaries. The predictions are presented as means with upper and lower 95% confidence intervals.

Reality Check: Typical Data

Planned data collection such as that just described brings benefits and is a worthwhile investment. However, most modelers are faced with less than ideal data or much more complex real-world situations, so now we move to the typical challenges of modeling with available resources. We should

note, though, that a dataset is not a fixed entity. One might start to model and use the initial results but design a process of collecting important missing data in parallel with achieving other goals. Relevant key terms for following these recently emerging research topics include "biodiversity survey costs and benefits," "adaptive surveillance," and "decision theory." More broadly, the task of the species modeler is to look at the data as realistically as possible, to understand the models and their underlying assumptions, to think through the aims and priorities of the modeling, and then to do as good a job as possible with the data at hand. Better to find that the data are inadequate to the task and report as much than ignorantly fit a model and assume all is well.

Species Data

In a SDM, observations of the species form the response – that is, the dependent variable whose variation might be described by the predictors. For a single species, these can be counts, cover-abundance estimates, or presence–absence or presence-only records, and modeling and validation methods exist for all these response types. In this section, it will become clear that there is a strong interplay between features of the species data and the modeling method because many variations on methods are developed for dealing with the realities of data.

Sample Characteristics

The hypothetical ecologists gathered 500 records across major environmental and geographic gradients. Sample size is important because the number of samples dictates whether the response to environment can be modeled in all its complexity, especially when there are many predictors. Few samples will only allow the strongest trends to be identified. Even though some algorithms might be more robust than others to small samples, no method can "invent" information that is not present in the data. If modeling rare species or small datasets (e.g., less than 30 sampled sites; or for presence–absence data, less than 30 records of the rarest class), it is worth reading the literature on what can be done with small sample sizes.

Covering the main gradients in the region achieves several things (Cawsey *et al.*, 2002). It ensures that sampling is unbiased and representative of the region. It also provides information to the model about the shape of the species response (the "fitted function," **Figure 1**) across varying environments, and for endemic species can allow upper and lower thresholds and optima to be identified. Albert *et al.* (2010) explore a range of sampling designs tailored to specific questions that SDMs might address. If a species' habitat extends well beyond the environmental limits of a sample (i.e., only a portion of its range has been sampled), the model fitted to the sample might represent the current distribution of the species in that region well enough but using it in applications requiring extrapolation is likely to be problematic. This sort of partial sampling of ranges is common – for instance, when survey data from a particular geographic region are used to model all species that occur in the region, or when fisheries data are used to model all recorded species including bycatch species. Uneven distributions of samples are also common in citizen science programs (e.g., **Figure 3**), and indications of data support for the model (or perhaps the minimum resolution at which predictions provide useful information) can be useful.

SDMs assume that each sample is independent, meaning that each brings new information to the model. This is most often violated in datasets compiled from several sources, in



Figure 3 The 107,000 unique locations (gray dots) for observations in the online citizen science project, eBird. Restricted to data collected across the conterminous US from 2004 to 2009.

which it is common to find replicate records at one location or very dense local sampling. If the available data include replicates and dense clusters, it is best to consult the literature for appropriate methods for cleaning them or modeling them. Increasingly, tools are becoming available for screening and cleaning data. Samples are also assumed to be random. One common nonrandom sampling design groups records within broader units like regional parks and without sampling any of the intervening habitat. These can be accommodated in generalized linear or additive mixed models or Bayesian hierarchical models, because the units can be treated as random effects. Samples along transects or from fishing towlines can also be modeled with specialized methods, including mixed models or autoregressive models (Aarts *et al.*, 2008).

Record Characteristics

The preceding discussion addressed the general properties of the available sample – its geographic and environmental extent and the independence of records. We now turn to think about the properties of the individual records, though admittedly these two categories overlap. Here we consider the types of data being collected and the relationship between that and what is modeled. Our hypothetical ecologists collected abundance data because they were interested in its variation across the landscape in the context of the communities in which the cypress occurred. In contrast, many datasets comprise records of the observed presence or absence of the species. Usually, modelers are intending to model probability of presence of the species from these data, though some simplify the results to a presence–absence prediction because end users want a map of suitable versus unsuitable habitat. As will be seen in the following section (*see* Presence-only Data), the advantage of presence–absence compared with presence-only samples is that presence–absence samples have clear information on what has been surveyed. However, the meaning of absence can be somewhat unclear and can vary with temporal and spatial resolution of sampling. This can make it difficult to work out what probabilities really mean. For instance, a tree is more likely found in a 1-km^2 quadrat than a 1-m^2 plot, a species might be difficult to detect amongst dense vegetation, an annual or cryptic species might only be observed in specific seasons, and a vagile species (e.g., a bird, a fish) might be present at some times and not others. These examples represent different factors (with varying impacts) that influence the record of presence or absence and likewise could affect abundance estimates.

The quadrat or plot size is important in so far as it needs to be well suited to the ecology of the species and the heterogeneity of the landscape. Further, inferences about probability of occurrence are tied to the sampling effort, and a mismatch between the area surveyed and area predicted to (e.g., surveys in 20-m $\times$ 20-m quadrats predicted to 1-km $\times$ 1-km grid cells) can influence the accuracy of predicted probabilities, particularly if grid cells are not environmentally homogeneous.

Imperfect detection is a more important issue than one might assume given its level of attention to date in the species modeling literature, so here we provide a small focus on it, anticipating that it should rise in prominence in the coming years. The species data that we have so far called presence–absence data are only true presence–absence data if detection probability is 1. There is growing evidence that detection probabilities for many species are less than 1 – in other words, false absences due to observation errors are common. This means that a model that does not address imperfect detection is modeling some unknown combination of probability of occurrence and probability of detection. Researchers including Royle, Nicholls, MacKenzie, Tyre, Wintle, and Kéry have developed relevant theory and applications. Kéry *et al.* (2010) make a case for calling models that assume perfect detection models of *apparent* distribution. In contrast, a model of *true* distribution would require that two nested processes are explicitly modeled: (1) the ecological process that generates the true distribution of the species, and (2) the observation process that governs probability of detection (**Figure 4**). Models that deal with imperfect detection are commonly called "site-occupancy" models and are special cases of generalized linear mixed models (Kéry *et al.*, 2010). The key piece of information they require is repeat visits to at least some sites and within a short enough time frame that occupancy can be considered unchanged (i.e., that population closure can be assumed). Garrard *et al.* (2008) address related survey issues for plants. Among the consequences of wrongly assuming perfect detection are that occupancy is underestimated, the estimated strengths of covariate relationships are biased toward zero, and varying detection probabilities are wrongly assigned as varying probabilities of occupancy (**Figure 4**). With all these problems, one might wonder why explicit modeling of detection problems is relatively rare. In some cases, it may be that ecologists are making an unacknowledged and mostly untested assumption that more sites are better than more visits. This is yet to be broadly assessed, but see Mackenzie and Royle (2005) for tips on allocating survey effort. Kéry suggests that a lack of awareness, the requirement for repeat visits, a current scarcity of easy-to-use software that links to GIS data, and no software implementations taking advantage of modern methods of feature (variable) selection have all contributed to slow uptake. These are changing as we write, and we envisage more attention to this issue in the coming years. We note that imperfect detection also impacts abundance estimates, but these can also be dealt with under the repeat visit–generalized linear mixed model framework.

The third example given earlier, that a species might only be observed in some seasons, can now be seen in the light of the preceding paragraph as an issue of detection. Detection probabilities can vary seasonally, so an obvious way to include seasonal effects is to model them as covariates of the detection probabilities. In the absence of repeat visit data, seasonal effects have also been included as predictors in conventional SDMs. The fact that vagile species only sometimes occupy sites reflects lack of population closure and is technically challenging to model in any framework, with or without detection. Wildlife modelers often approach it by thinking about what defines a *site*, and – rather than strictly define it – model use versus availability (resource selection functions). This puts the focus on where and how the animal uses a territory, rather than on the definition of the territory. Finally, other record characteristics related to locational and taxonomic accuracy are important regardless of the type of data, but covered in the following section.

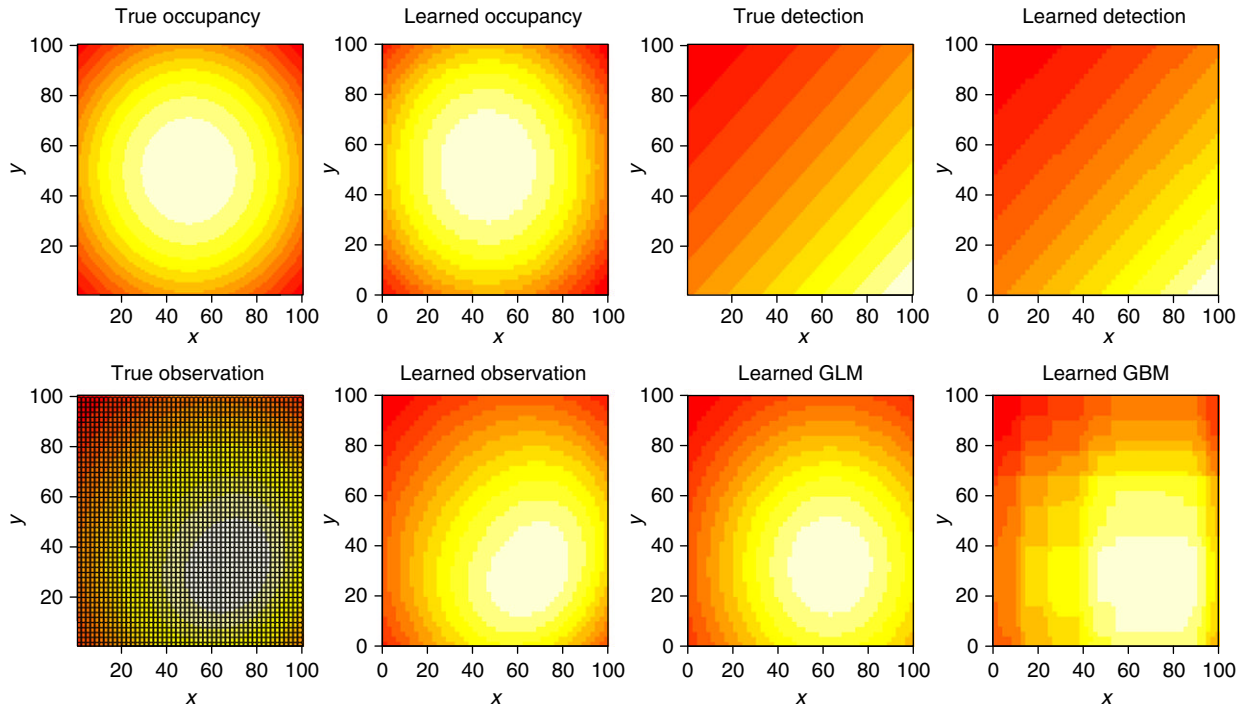


Figure 4 Simulated data modeled with site-occupancy models and models (GLM; and boosted trees, GBM) that assume perfect detection. Top row: true occupancy and detection (pale yellow high, red low) and their modeled counterparts fitted ("learned") using a site-occupancy model. Bottom row: true (simulated) observations resulting from the occupancy–detection patterns in the top row with a grid showing sampling locations, and estimated observations learned using the site-occupancy model, the GLM and the GBM. Note that only the site-occupancy model can be decomposed to predict occupancy rather than observation.

Presence-only Data

Presence-only data are now so widely used that they deserve special attention, though some of the issues covered in this section may also apply to other types of data. Two distinct types of presence-only data exist. The first, radio-telemetry data, are common to wildlife studies and addressed in another chapter. What we refer to here as presence-only data are records of presence without accompanying records of surveyed but unoccupied sites, such as records from compilations of observations, checklists, herbarium and museum specimens, and so forth. Presence-only data are increasingly used for SDM because they are often the only source of information on species locations. Their quality and particular characteristics vary according to their source. For instance, fisheries data may be strongly biased to those areas known to be productive and within reach of a port. Museum data as collated in online data repositories such as the global biodiversity information facility (GBIF) are sourced from numerous data providers throughout the world. This makes them a tremendously important source of biodiversity data, but errors, duplications, and biases are inevitable. Their use for species modeling brings particular challenges (Newbold, 2010), including errors and lack of specificity in location, errors and lack of currency in taxonomy, and biases in the sampled geographic or environmental space, the temporal spread of surveys, and the recorded taxa. These violate several assumptions outlined earlier about the types of samples needed for species modeling. Biases are a particularly challenging issue, because – without information on whether the sample is representative of the species or biased – SDMs will model an unknown combination of survey bias and suitable environments for the species.

Errors including lack of precision can sometimes be corrected, for instance, by working with experts to update taxonomies, using ancillary information to more accurately locate and model records (Niamir *et al.*, 2011), or quantifying uncertainty in location and exploring its effect on model output. Bias is much harder to deal with because it is generally difficult to quantify the biases that exist, but various methods have been suggested, including using information on survey effort for related species (Phillips *et al.*, 2009). The level of research effort being devoted to correcting errors and addressing bias indicates the importance of presence-only data sources for SDM. When using such data it is crucial to be realistic about data quality and apply methods to deal with common errors and biases.

Biotic Interactions and Disturbances

We discuss biotic interactions and disturbances here, between species and environmental data, because they affect the species records but can also be considered predictor variables. Use of SDMs for predicting effects of climate change and invasive species have reinvigorated efforts to include biotic interactions in models because they are clearly one of the ecological realities that will substantially affect the future distributions of species. There is limited scope in correlation-based species distribution

models to deal cleanly with biotic interactions because those interactions have affected the distribution of species data, and there is rarely independent evidence of what would have occurred if those effects (competition, mutualism, facilitation) were removed (but see Leathwick and Austin, 2001, for an exception). In some large, carefully selected datasets, there may be enough information to usefully explore the effects of competition on species abundance. For instance, Meier *et al.* (2011) used data from 6000 forest plots to model the abundance of European beech in relation to three major competitors. Models, including predicted distributions of the competitors, indicated strong effects of competitors in favorable growing conditions but weakened links in unfavorable environments. Biotic interactions are a fact of nature, and we anticipate ongoing efforts to understand and model these, perhaps most effectively through mixes of correlative and mechanistic models.

Disturbances – recent (fire, land clearing for agriculture) and historic (glaciations, past volcanic eruptions) – also affect distributions, with the temporal scale of the effect varying with the disturbance and the ecological characteristics of the target species. So far these are rarely included as predictors in models due to lack of mapped data, but remote sensing information is likely to promote their use. One advantage of having them as predictors is that conceptually it would then be possible to remove their effect from the modeled distribution where potential distributions are required.

Environmental Data

In our perfect world example, the modelers gathered GIS data describing the environmental gradients in the study area and then summarized them in ways that were relevant to the species. In reality, in spite of growing volumes of geospatial data available for many parts of the word, much of it derived from remote sensing (Skidmore *et al.*, 2011), these data do not necessarily describe the most proximal environmental factors that limit species distribution and abundance; nor are they always available at scales relevant to the species. Further, detailed knowledge of the proximal factors that limit species distributions is not available for most species. Instead, SDM analyses often model many species (e.g., to determine patterns of biodiversity or conduct and environmental impact analysis), and a common and convenient set of available predictors is used. These predictors frequently include descriptors of climate (and sometimes substrate, terrain, and land cover) in terrestrial settings, hydrology in aquatic ecosystems, and bathymetry, productivity, and sea surface temperature in the marine environment. The challenge in SDM is to develop a set of candidate predictors that are likely to have a proximal relationship to species distributions. This often involves applying a transformation or model, sometimes a complex one, to the mapped variables. In the following sections we discuss relevant variables over a range of scales, recognizing that many local modeling efforts might have access to, or ability to collate, a potentially broad suite of relevant predictors.

Climate

Although it is well known that climate limits species distributions at broad spatial scales, the variables that climatologists typically measure, and interpolate into maps such as monthly precipitation and minimum and maximum temperature averages (or measures of humidity or solar radiation), are not necessarily the most biologically meaningful factors. For example, in our perfect world example, precipitation during the dry season was known to be a limiting factor for the species of interest. Therefore, in the 1980s, researchers in Australia defined a set of bioclimatic variables for SDM (Busby, 1991) that are still used today (Hijmans *et al.*, 2005). These involve transforming monthly precipitation and temperature averages (average is usually defined as 30 years in climatology) into measures of seasonal averages or variation, contrasting the warm and cool or wet and dry seasons. Temperature sums such as growing degree days (above some minimum) and water-balance measures based on calculations of potential and actual evapotranspiration may be even more closely related to the primary regimes of heat and water availability, and for some species the most important parameters might be climate variability and extremes (e.g., frequency and duration of droughts and floods).

Some applications of SDMs (e.g., for exploring macroecological theories) are at coarse resolution and global extent, and these tend to focus on climate variables. For those wanting to model at finer grain, until now the main globally coherent dataset has been the Worldclim variables (Hijmans *et al.*, 2005) with 19 rainfall and temperature-based variables and a finest resolution of 1-km grid cells. Current efforts are likely to provide a broader suite of variables in the near future. In addition, with so much emphasis on the use of SDM to assess the potential impacts of global warming on ecosystems, it is important for species distribution modelers to understand the data and methods that climate scientists use to interpolate historical climate records, and to appropriately downscale coarse-resolution climate model projections of future climate scenarios.

Terrain

Elevation is either a modeled or a satellite product and is available globally. Elevation per se represents an indirect environmental gradient, correlated with but not directly measuring temperature and moisture regimes (Austin, 2002). There is rarely a good reason to include elevation as a predictor, since other more proximal variables (such as temperature) are usually available. At any given elevation, terrain strongly modifies the radiation regime and available soil moisture and these effects of slope, aspect, and hillslope position can be modeled (Wilson and Gallant, 2000). Other elevation-based measures are also useful for some species, including local variation (e.g., ruggedness calculated from standard deviations of elevation in certain windows), compound topographic index (a measure of the water flow through the landscape relying only on topography), and estimates of landscape position (ridge or valley).

Substrate, Soils, and Land Cover

Information on substrate and soils is usually one of the hardest to find in a format and at a scale useful for species modeling. Traditional approaches to soils mapping usually focus on classifications, whereas continuous variables such as soil depth, pH, fertility, and water-holding capacity are more

directly related to species distributions. When continuous soil properties are mapped, they are often available only at coarse scales. Recent focuses on harmonizing soils information globally and on modeling soils characteristics in relation to terrain are likely to produce useful products for species modeling.

Land cover or vegetation type is often an important, or even the only, predictor used in species distribution (habitat suitability) models for wildlife. Vegetation composition and structure are strongly related to habitat use for many species, from butterflies and birds to bears. Land-cover maps, based on remotely sensed and other data, are now being compiled at regional to continental scales to monitor global environmental change, and these data products often include somewhat detailed information about natural vegetation categories beyond simply life form (forest, grassland), for example, for North America and Europe. These large-area datasets, as well as subregional maps of vegetation type and vegetation attributes, have proven useful in SDM. In some cases, remotely sensed biophysical image maps or data products are used directly (without classification) in modeling distributions. These include maps of net primary productivity, leaf area index, canopy roughness, reflectance, greenness indices, and other spectral transforms. Land-use maps show the distribution of human disturbance in the form of land-use change, and fire history maps show the distribution of natural disturbances. Land-cover, land-use, and disturbance maps can be brought into SDM to address, to some extent, biotic interactions (distribution of other species) and community dynamics as affected by disturbance (*see* Biotic Interactions and Disturbances).

It is important to keep in mind that, when using predictors such as soil survey maps, large-scale soil databases, land-cover maps, or vegetation maps, these were produced using protocols, decision rules, methods, and models that are as complex as SDM (in some cases, indistinguishable from SDM – e.g., Ohmann *et al.*, 2011) and so have many assumptions built into them, for example, from the choice of classification system to the resolution or "minimum mapping unit" and the spatial homogeneity of those units. They are often developed to suit multiple needs and may have multiple attributes. In particular, when a categorical (thematic) map is used in SDM, it is necessary to reclassify or transform the categories in ways that are meaningful. If modeling habitat suitability for a forest-dependent bird using a land-cover map as a predictor, it may not make sense to retain dozens of categories of crop type or urban land cover. These use many degrees of freedom, estimating parameters for each category, and are essentially meaningless. A simple quantitative predictor such as forested area within a certain neighborhood could be derived from a land-cover map. As another example, soil or rock-type maps can be reclassified into categories of high, medium and low nutrient availability, or soil series could be grouped into simpler categories such as soil order.

Aquatic Systems

The most sophisticated sets of predictor variables for freshwater systems are based on rivers and streams represented in GIS as a network topology, with each river or stream section between adjacent confluences represented by a unique segment, and the ability to trace up- and downstream to summarize environments and trace species movements. Environmental attributes including physiographic, bioclimatic, edaphic, land-cover, and disturbance-related variables can be estimated at hierarchically nested spatial scales (e.g., the riparian zone of each segment, its immediate watershed, and its upstream contributing catchment area). These provide biologically relevant data for predicting distributions.

In marine systems, depth is an indirect proxy for several proximal predictors: temperature and its variability, salinity, light, pressure, and the availability of elements (e.g., calcium). Marine species distributions are also affected by ecosystem productivity (e.g., as quantified through measures of chlorophyll *a* or zooplankton), by tides and currents, and by dissolved oxygen. Gradually, more globally complete sets of marine data are becoming available (e.g., Tyberghein *et al.*, 2012), and various countries are developing their own regional datasets at higher resolution.

Multicollinearity of Environmental Predictors

An important issue requiring attention is collinearity between variables – that is, the statistical correlation between two or many variables (Dormann *et al.*, 2012). Collinearity is common in large sets of predictors because several variables might be representing the same underlying but unknown latent (unobserved) and most proximal variable. Collinearity may not be problematic if the main use of the model is for prediction within the range of the sampled data (interpolation). In these cases, selection from a large candidate set of variables might be successful, particularly with large numbers of species records and methods with variable selection procedures that are robust to collinearity. However, prediction to new times or places will be problematic if the correlation structure changes because the model might have selected a correlate of the true variable, whose relationship with the true variable is not maintained in the new prediction space. The fundamental problem is that statistical models identify correlation, not causation, and there is no straightforward way to decide which variables to focus on and which to ignore. Many datasets include collinear variables – for example, the many bioclimatic variables available in datasets such as Worldclim are often highly correlated. It is tempting to use all predictors at hand and "let the model decide" which are important through model selection procedures when there is not strong theory predicting the most proximal factors limiting the distribution. However, it leads to more interpretable and parsimonious models, and hopefully more transferable (useful for projections) ones if a subset of predictors can be chosen based on the literature, expert opinion, or exploratory data analysis.

The Models

Generally in SDM, the response variable is a measure of species occurrence or abundance, and there are multiple candidate predictors describing the physical environment that are thought to be important limiters of species distributions. SDM is thus a supervised classification or statistical learning problem (sensu Hastie *et al.*, 2009).

A continuum of approaches are available for supervised learning. At one end of the spectrum are those approaches that have high bias (make assumptions about the form of the relationship between response and predictors) and low variance (tend to make robust predictions when applied to new data). These are sometimes referred to as *model-driven* or *parametric* methods. At the other extreme are those approaches that have low bias but high variance, which some call *data-driven* or *nonparametric*. Although it can generally be said that machine learning methods are nonparametric and data driven, and statistical methods are parametric and model driven, this distinction is being increasingly blurred by the regularization methods used in machine learning to reduce variance, and nonparametric methods used in statistics to reduce bias, such as the curve-fitting used in GAMs.

The predominant statistical framework that has traditionally been used for SDM is that of generalized linear models (GLMs) in which the many predictors are environmental variables (X) and the single response variable (Y) is a measure of species occurrence or abundance (or another biodiversity measure). *Generalized* means that the model can cope with nonnormal error distributions of the response variable. For example, in SDM if Y is binary (species presence or absence), then the binomial distribution is used; if Y is a count, then Poisson errors might be appropriate. The GLM can be expressed as

$$g(E(Y)) = \mathrm{LP} = \hat{\beta}_0 + \sum_{j=1}^{p} X_j \hat{\beta}_j + \varepsilon$$

where the predictor variables (X's) are combined to produce a linear predictor, LP, and the expected value of Y, $E(Y)$, is related to the LP through the link function, $g()$. The link function describes how the mean of Y depends on the linear predictor, and a variance function describes how the variance of Y depends on its mean for any particular distribution from the exponential family. The *linear* part of GLM means that the models are linear functions of the parameters, not necessarily of the independent variables. GLMs have the flexibility to characterize nonlinear responses to a predictor by adding terms to the linear model such as polynomial terms (to characterize unimodal responses) or piecewise linear terms (to estimate threshold responses). Categorical predictors can be incorporated as dummy variables, and interactions can be specified. GLMs are the basis for an extended set of models, including GLMMs, geographically weighted regression, and spatial autoregressive methods that can deal with all sorts of complexities of real datasets (*see* Species Data).

GAMs are closely related to GLMs and tend to have lower bias but higher variance because the coefficients of the GLM (βs) are replaced by a function, f, typically some kind of a smoothing spline whose complexity is controlled by the degrees of freedom allocated to it. GAMs are convenient for fitting skewed response curves or any form of nonlinear response because the shape of the function in the GAM is data driven and not prespecified. For both GLMs and GAMs decisions are required about how to select the final model – including the complexity of the fitted functions and whether to select subsets of variables. Ecologists have paid much attention to model selection based on information criteria such as Akaike's Information Criterion (AIC), but there are a broad suite of possible approaches including shrinkage of parameters, L1-regularization, and so on (Hastie *et al.*, 2009).

With changing emphasis on what is required of models and rapid advances in computing power and machine learning and data-mining algorithms, methods have proliferated. Examples include classification and regression trees (decision trees) and ensembles of them, artificial neural networks, maximum entropy methods, genetic algorithms, and support vector machines. Each has particular strengths and weaknesses. For instance, decision trees use recursive binary splits (a hierarchical splitting procedure) to partition the predictor space into rectangles within which the response is most homogenous. Their tree-like structure identifies interactions and easily handles categorical predictors and threshold responses, and it represents information in a way that is intuitive and natural to visualize. They can be applied to binary, continuous and categorical responses. However, there are some drawbacks: decision trees tend to vary with the sample (i.e., they have high variance) and do not characterize linear or smooth responses very well. Ensembles of trees (including bagged trees, boosted trees, and random forests) overcome these limitations and are increasingly used. They have strong predictive capacity and reliable variable selection, and they can fit complex responses. Several can be applied to a variety of types of species data (e.g., presence–absence, count) similar to regression models. Intriguing recent applications of tree ensembles address imperfect detection (Hutchinson *et al.*, 2011) and spatial and temporal variations in bird–habitat associations (**Figure 5** and Fink *et al.*, 2010). Currently, machine learning methods tend to be viewed as fitting more complex models and therefore are useful for prediction but perhaps not so well suited for applications requiring smoother models (prediction to new times or places). This is an oversimplification, and it is likely that the breadth of machine learning methods will continue to expand and find useful applications in species modeling.

Presence-only data tend to be modeled both with specialized methods and with adaptations of methods applicable to presence–absence data. Multivariate distance and envelope methods are also sometimes called *profile* methods (Pearce and Boyce, 2006) and characterize habitat suitability in terms of descriptors of the environmental conditions where species are observed to occur. The descriptors can be ranges (minimum and maximum), medians, means, or multivariate distances in environmental measurement space. Those methods based on ranges are sometimes called *envelope* methods, and they were among the earliest approaches used in SDM. Multivariate distance measures are widely used in community ecology and other fields, and they can be used to classify new observations in terms of their distance in environmental hyperspace from a descriptor of the species occurrence locations – for example, mean centroid or nearest neighbor. Distance measures that have been used in SDM include Gower and Mahalanobis. Mahalanobis scales the difference between an observation and the mean vector for species occurrences by the covariance between predictors, thus making the distance measure independent of the scales of the predictors. However, Mahalanobis gives all predictors equal weight and cannot be used with categorical predictors. Nonetheless, distance-based methods appear to have

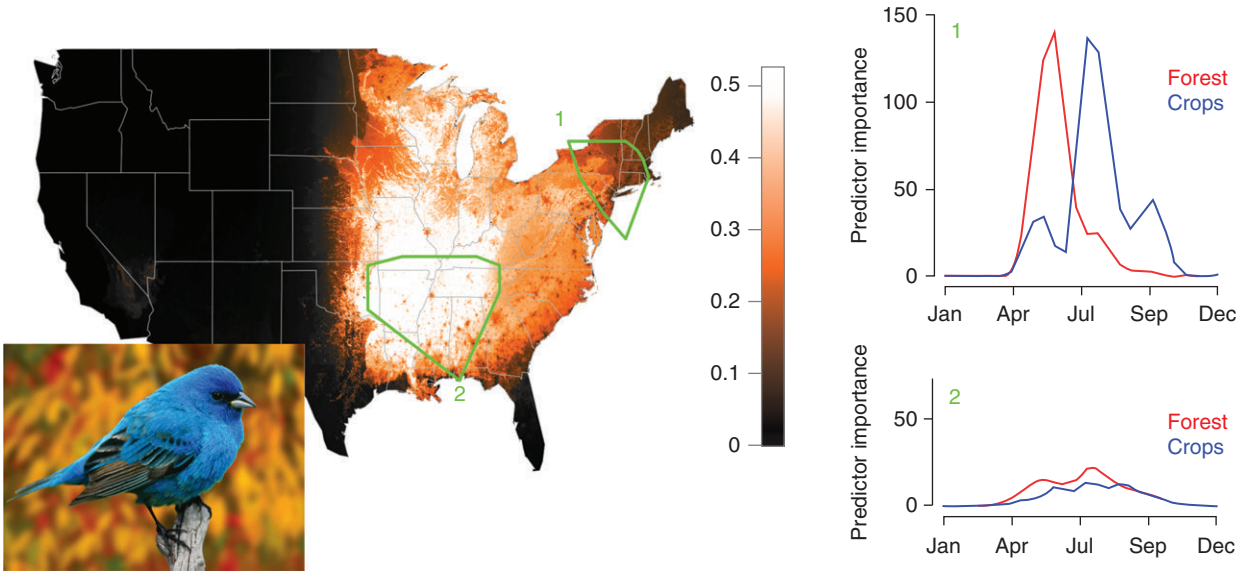


Figure 5 Example of use of a complex model for ecological inference. Here a machine learning method adapted for dealing with spatial and temporal variation in species–covariate relationships (Fink *et al.*, 2010) is used to predict the distribution of indigo bunting (*Passerina cyanea*) based on ~600,000 presence–absence records at 107,000 locations across 5 years (**Figure 3**) and covariates, including remotely sensed habitat classes. The map shows the predicted distribution on 20 June, and the plots show trajectories of predictor importance in two regions (green polygons) for two example habitats: deciduous forest (red) and cultivated crops (blue). In the northeast (region 1), foraging patterns shift from the forest (targeted on arrival and subsequently for breeding) to the crops (important during fall migration). In comparison in the core of the bunting's breeding distribution (region 2), the species commonly occurs in most habitats so there is much less variation over the season and no dominant predictors (note the smaller predictor importance values) as in the northeast.

a strong and continuing tradition of use in wildlife management studies (e.g., Hellgren *et al.*, 2007).

On the machine-learning side of things, applications of maximum entropy theory (MaxEnt) and genetic algorithms (GARP) have been widely used for modeling presence-only data, primarily because they provided interface-driven software that made it easy to deal with common data types. In common with adaptations of regression methods, these methods compare the environments occupied by the species with those in a selected region. Environments in those regions are characterized by taking background samples (also called *pseudo-absences*, though the latter term has a range of subtly different interpretations; Phillips *et al.*, 2009). Details of how each approach works and performs are described in recent publications.

Bayesian methods are also increasingly used for species modeling, particularly for more challenging modeling problems where the flexibility of the Bayesian approach allows realistic specification of the problem and useful ability to update models as new information becomes available. Recent texts by Michael McCarthy and Marc Kéry introduce these methods to ecologists.

Although the modelers in our perfect-world example evaluated predictions of species abundance using measures of model fit and predictive performance typical for continuous response variables (e.g., root mean square error), it is worth noting that most SDMs are developed using presence–absence or presence–background data and use performance measures suitable for binary predictions (reviewed in Elith and Leathwick, 2009b; Franklin, 2010). These include threshold dependence measures such as the area under the (receiver operating characteristic) curve (AUC) and threshold-dependent measures such as the kappa coefficient of agreement (where the threshold is a value applied to a probabilistic prediction of species occurrence in order to distinguish presence from absence predictions). Issues of how to set up testing and training datasets in ways relevant to the modeling problem are important, as are evaluations beyond simple "fit to data sample" measures, including whether the model and its predictions make sense to experts.

Species Distribution Modeling for Other Biodiversity Measures

The general modeling approach of SDMs can be applied to a range of biodiversity attributes. For instance, community diversity, community composition, community type, and functional group distribution have all been modeled, though species diversity (often species richness) has received most attention. Various theories of species diversity hypothesize the driving factors to be energy availability, which can be indicated at broad scales by ecosystem productivity or potential evapotranspiration (Currie, 1991) or environmental heterogeneity (Pausas *et al.*, 2003). Two distinct approaches have been used: calibrating a model using species richness as the dependent variable versus developing SDMs for each species in an area and adding their distribution maps to estimate richness per grid cell (a stacked SDM approach). Using the first approach, species richness is increasingly being predicted directly from remotely sensed variables or indices related to productivity, water and nutrient status, and vegetation structure and phenology (e.g., net primary productivity, greenness indices, surface temperature, and image texture). Essentially, these are biodiversity models with

only one or a few predictors that are exclusively derived from remote sensing data.

Predictive maps of species richness are uninformative about species composition, and often information on community or ecosystem composition is needed for biodiversity assessment and conservation planning (Ferrier and Guisan, 2006). Individual SDMs have been stacked together (predict species distributions first, then classify communities), and community or ecosystem type has been modeled as a categorical dependent variable (classify, then predict). Composition has also been modeled by estimating species turnover among sites – for example, using a form of matrix regression (generalized dissimilarity modeling) or other approaches based on multivariate distance.

Other frameworks for summarizing general species attributes or distributions are also relevant to understanding and managing biodiversity. Plant life forms or functional types form the basis of most mechanistic dynamic vegetation models, and so SDMs developed for functional types may be useful for evaluating the impacts of environmental change on biodiversity. Biogeographical regions have also been defined, based on SDM-like models, at global (Kreft and Jetz, 2010) and regional (Leathwick *et al.*, 2011) scales.

Predicting to New Times and Places

Although the original uses of SDMs were for understanding ecological data and predicting within sampled regions, predictions to new times and places have now become a strong focus. SDMs are inherently poorly suited to such applications because they are based on correlations (not causal relationships) estimated within a training region that is not necessarily related environmentally or ecologically to the new conditions of interest. Their basis in realized distributions also creates difficulties for useful projections elsewhere (Colwell and Rangel, 2009; Sinclair *et al.*, 2010). Nevertheless, researchers persist with these applications because information is required and so far there are few alternative practical methods for producing it. Two main scenarios are climate change and invasions. These are high-priority research areas because both can have immediate and long-lasting detrimental effects on biodiversity and both are costly to mitigate.

For climate change applications, SDMs fitted to current environments are projected forward to new climates predicted and downscaled from global circulation models The early simple applications of these models now seem relatively naïve, and researchers are exploring new avenues by including biotic interactions; defining the ability of species to migrate – often across fragmented landscapes – with changing climates; exploring the realities of microclimates that might provide refugia, including realistic fine-scale predictors that affect the local complexity of species distributions; controlling the shapes of the fitted functions; identifying where climates are novel; using predictors suggested by ecophysiological models as important constraints to persistence; and linking the SDM outputs to process-based models that explicitly focus on birth, death, dispersal, spread, and other population processes. Uncertainty across modeling methods and alternative GCMs and emission scenarios has also been explored. These advances may help to make predictions more realistic, but they cannot overcome the inherent problem that a model based on correlations does not and cannot identify causes, nor can one based on realized records predict fundamental niches. Perhaps the best way forward is to carefully define what ecological and data characteristics make species more likely amenable to successful prediction, to keep collecting biological information so models are as well informed as possible, and to progress understanding and visualization of uncertainties and their integration into decision making so these can be properly understood and accounted for by policy makers.

Another "across-time" application is prediction to past climates. Paleodistributions are hindcast using paleoclimate maps based on reconstructions or models typically available only for key time periods such as the last glacial maximum and mid-Holocene (e.g., from the Paleoclimate Modeling Intercomparison Project). Paleodistributions modeled from SDMs have an advantage over future distributions predicted from climate scenarios in that they can be evaluated with paleodistribution data (fossils, pollen, molecular phylogeographies), albeit these data may be sparse in time and space. This approach has recently been used to compare alternative hypotheses of allopatric diversification, identify locations of past climatic refugia, estimate postglacial migration rates, and even identify likely postglacial migration pathways, for plants, vertebrates, and invertebrates in both the tropics and temperate–boreal zones.

SDMs are also used outside their training range for predicting the potential distribution of invasive species. Also called *pest risk mapping* and *climate matching*, these models are usually of interest in biosecurity for identifying species likely to become pests, guiding monitoring programs to areas most likely to be infested, and providing a basis for sharing costs of containment or eradication between jurisdictions. Again, there is a mismatch between what is required (a prediction of the fundamental niche of the species that could be occupied in the invaded range if the species was able to spread everywhere and if biotic conditions were suitable) and what is possible with a SDM. A major issue is that the native range of a species might only represent part of its fundamental niche, so models are uninformed about the ability of the species to persist in climates outside of those encountered in the native range. Added to that, the usual problems with incomplete and biased species data, imprints of biotic interactions, and lack of knowledge about the true predictors of distribution make the task a difficult one. At the time of writing, a much stronger appreciation of the technical problems is developing make we anticipate advances in tools for diagnosing and reporting likely problems and methods for controlling models perhaps through integration with expert knowledge. Invasive species are also modeled with partially mechanistic models, and in the invaded range, SDMs have been combined with various types of spread models (Iverson *et al.*, 2011).

Conclusion

Species distribution models are statistical models of species–environment relationships based on species location

(abundance, occurrence) data and measures of environmental variables limiting species distributions. They are usually used to make spatial predictions. SDMs have been developed for marine, terrestrial, and freshwater species at a wide range of scales. In spite of the growing availability of bioinformatics databases and geospatial data, the accuracy and usefulness of SDMs are often limited by data availability, though limitations can be addressed through additional survey and sampling, as well as geospatial data processing.

There are hundreds of published SDM studies, and the number is growing. Until recently, these studies often emphasized the technical issues of modeling methods and model evaluation. Tremendous progress has been made in statistical and machine learning methods, advances that tend to blur the distinction between them. Increasingly, SDM studies now emphasize the application of SDM to problems in environmental assessment, forecasting, and macroecology. SDM is a powerful tool for biodiversity assessment and conservation planning. SDMs are used to project species distributions to novel times and places because of the overwhelming importance of questions regarding the impact of environmental change on biodiversity. Based on correlations and not causal mechanisms, SDMs are inherently poorly suited to making projections under novel conditions. Combining SDM with other data and approaches, such as spatially explicit models of population and community dynamics, may improve forecasting and impact assessment of environmental change at larger spatial scales.

The particular application of SDM, as well as the characteristics of the data, must determine what is emphasized when modeling. Species detection probability, environmental limiting factors, SAC, sample size, sample bias, scale, disturbance, biotic interactions, and demographic processes will vary in their importance among species and datasets and depend on the intended purpose of the model. Thus, understanding the data, the question, the particular features of different model types, and how to assess errors and uncertainties is part of the art of fitting and evaluating species distribution models.

See also: Biogeographical Models. Climate Change and Extinctions. Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. Climate, Effects of. Computer Systems and Models, Use of. Framework for Assessment and Monitoring of Biodiversity. Habitat and Niche, Concept of. Habitat Selection. Introduced Species, Impacts and Distribution of. Phylogeny. Plant Invasions. Species, Concepts of

References

Aarts G, Mackenzie M, McConnell B, Fedak M, and Matthiopoulos J (2008) Estimating space-use and habitat preference from wildlife telemetry data. *Ecography* 31: 140–160.

Albert CH, Yoccoz NG, Edwards Jr. TC, Graham CH, Zimmermann NE, and Thuiller W (2010) Sampling in ecology and evolution – bridging the gap between theory and practice. *Ecography* 33: 1028–1037.

Austin MP (2002) Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling* 157: 101–118.

Busby JR (1991) BIOCLIM – a bioclimate analysis and prediction system. In: Margules CR and Austin MP (eds.) *Nature Conservation: Cost Effective Biological Surveys and Data Analysis*. Canberra, Australia: CSIRO.

Cawsey EM, Austin MP, and Baker BL (2002) Regional vegetation mapping in Australia: A case study in the practical use of statistical modelling. *Biodiversity and Conservation* 11: 2239–2274.

Colwell RK and Rangel TF (2009) Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences* 106: 19651–19658.

Currie DJ (1991) Energy and large-scale patterns of animal-species and plant-species richness. *American Naturalist* 137: 27–49.

Dobrowski SZ, Thorne JH, Greenberg JA, *et al.* (2011) Modeling plant ranges over 75 years of climate change in California, USA: Temporal transferability and species traits. *Ecological Monographs* 81: 241–257.

Dormann CF, Elith J, Bacher S, *et al.* (2012) Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*.

Dormann CF, McPherson JM, Araujo MB, *et al.* (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* 30: 609–628.

Elith J and Leathwick JR (2009a) Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics* 40: 677–697.

Elith J and Leathwick JR (2009b) The contribution of species distribution modelling to conservation prioritization. In: Moilanen A, Wilson KA, and Possingham HP (eds.) *Spatial Conservation Prioritization: Quantitative Methods & Computational Tools*. Oxford University Press

Ferrier S and Guisan A (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology* 43: 393–404.

Fink D, Hochachka WM, Zuckerberg B, *et al.* (2010) Spatiotemporal exploratory models for broad-scale survey data. *Ecological Applications* 20: 2131–2147.

Franklin J (2010) *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge, UK: Cambridge University Press.

Garrard GE, Bekessy SA, McCarthy MA, and Wintle BA (2008) When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology* 33: 986–998.

Goodchild MF (1994) Integrating GIS and remote sensing for vegetation analysis and modeling: Methodological issues. *Journal of Vegetation Science* 5: 615–626.

Guisan A and Zimmerman NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.

Hastie T, Tibshirani R, and Friedman JH (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edn. New York: Springer-Verlag.

Hellgren EC, Bales SL, Gregory MS, Leslie DM, and Clark JD (2007) Testing a Mahalanobis distance model of black bear habitat use in the Ouachita Mountains of Oklahoma. *Journal of Wildlife Management* 71: 924–928.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, and Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.

Holdridge LR (1947) Determination of world formations from simple climatic data. *Science* 105: 367–368.

Hutchinson RA, Liu L-P, and Dietterich TG (2011) Incorporating boosted regression trees into ecological latent variable models. In: *Proceedings of the Twenty-Fifth Conference on Artificial Intelligence. San Francisco*. pp. 1343–1348.

Iverson L, Prasad A, Matthews S, and Peters M (2011) Lessons learned while integrating habitat, dispersal, disturbance, and life-history traits into species habitat models under climate change. *Ecosystems* 14: 1005–1020.

Kéry M, Gardner B, and Monnerat C (2010) Predicting species distributions from checklist data using site-occupancy models. *Journal of Biogeography* 37: 1851–1862.

Kozak KH, Graham CH, and Wiens JJ (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution* 23: 141–148.

Kreft H and Jetz W (2010) A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography* 37: 2027–2215.

Leathwick J, Snelder TH, Chadderton L, Julian K, Elith J, and Ferrier S (2011) Use of generalised dissimilarity modelling to improve the biological discrimination of river and stream classifications. *Freshwater Biology* 56: 21–38.

Leathwick JR and Austin MP (2001) Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology* 82: 2560–2573.

Legendre P (1993) Spatial autocorrelation: Trouble or new paradigm? *Ecology* 74: 1659–1673.

Linke S, Norris RH, and Pressey RL (2008) Irreplaceability of river networks: Towards catchment-based conservation planning. *Journal of Applied Ecology* 45: 1486–1495.

Mackenzie DI and Royle JA (2005) Designing efficient occupancy studies: General advice and tips on allocation of survey effort. *Journal of Applied Ecology* 42: 1105–1114.

Mackey BG and Lindenmayer DB (2001) Towards a hierachical framework for modelling the spatial distribution of animals. *Journal of Biogeography* 28: 1147–1166.

Meier ES, Edwards Jr. TC, Kienast F, Dobbertin M, and Zimmermann NE (2011) Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. *Journal of Biogeography* 38: 371–382.

Newbold T (2010) Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography* 34: 3–22.

Niamir A, Skidmore AK, Toxopeus AG, Muñoz AR, and Real R (2011) Finessing atlas data for species distribution models. *Diversity and Distributions* 17: 1173–1185.

Ohmann JL, Gregory MJ, Henderson EB, and Roberts HM (2011) Mapping gradients of community composition with nearest-neighbor imputation: Extending plot data for landscape analysis. *Journal of Vegetation Science* 22: 660–676.

Pausas JG, Carreras J, Ferre A, and Font X (2003) Coarse-scale plant species richness in relation to environmental heterogeneity. *Journal of Vegetation Science* 14: 661–668.

Pearce JL and Boyce MS (2006) Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology* 43: 405–412.

Pearson RG (2010) Species' distribution modeling for conservation educators and practitioners. *Lessons in Conservation* 3: 54–89.

Pearson RG and Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography* 12: 361–371.

Peterson AT (2011) Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography* 38: 817–827.

Phillips SJ, Dudík M, Elith J, *et al.* (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications* 19: 181–197.

Rangel TFLVB, Diniz-Filho JAF, and Bini LM (2006) Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography* 15: 321–327.

Sinclair SJ, White MD, and Newell GR (2010) How useful are species distribution models for managing biodiversity under future climates? *Ecology and Society* 15: 8.

Skidmore AK, Franklin J, Dawson TP, and Pilesjö P (2011) Geospatial tools address emerging issues in spatial ecology: A review and commentary on the Special Issue. *International Journal of Geographical Information Science* 25: 337–365.

Tyberghein L, Verbruggen H, Pauly K, Troupin C, Mineur F, and De Clerck O (2012) Bio-ORACLE: A global environmental dataset for marine species distribution modeling. *Global Ecology and Biogeography* 21: 272–281.

Whittaker RH, Levin SA, and Root RB (1973) Niche, habitat, and ecotope. *American Naturalist* 107: 321–338.

Wilson JP and Gallant JC (eds.) (2000) *Terrain Analysis: Principles and Applications*. New York: John Wiley and Sons.

Species Diversity, Overview

A Ross Kiester, Turtle Conservancy, NY, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Clique A group of organisms in a food web in which any pair of species in the clique shares at least one prey species. Thus, it is a group of species that have similar diets and are different from other such groups.

Gap analysis A biodiversity policy tool which compares the actual distribution of species and vegetation classes to areas managed for the long-term protection of biodiversity and other classes of land management. Species and classes that are poorly represented in protected areas are gaps in the protection of all biodiversity. Such species and classes are candidates for proactive protection.

Lognormal distribution A statistical distribution which is normal or Gaussian ("bell curve") when the logarithm of the original data is used. Species abundance data from many communities fit this distribution well. The lognormal distribution often occurs as the result of the law of large numbers in which many small factors interact multiplicatively to produce a result such as species diversity.

Simpson diversity index An index of diversity based on species abundance data. If p_i is the frequency of species i in a community of n different species, then Simpson diversity $= \left(1 - \sum_{i=1}^{n} p_i^2\right)$ and is equal to the probability that two randomly chosen individuals from a given community are of the same species. This index of diversity is statistically unbiased, making it especially useful in practice.

Species diversity The variability of a group of species classified in some way. For example, individual species can be classified by size and then a species size diversity measure can be developed which expresses the variability with respect to size. The most common classification is by abundance or population numbers. Then species diversity measures the variability of the number of individuals of each species. The Simpson diversity is an example of such an index. Many such indices have been developed for species diversity.

Species richness The number of species occupying a particular area (such as an island) or biological entity (such as a branch of a tree) without regard to any other properties of the species. Species richness may also be expressed as the list of species that generates that number.

Taxonomic rank The level in the Linnaean hierarchy to which a particular taxon or group of species is assigned. For example, mammals are at the level of the class (Mammalia) and hippopotamuses are at the level of family (Hippopotamidae).

Species diversity is one of the most fundamental aspects of biodiversity. However, there are many ways to think about it and ever more ways to measure it. The most basic idea is that of species richness, which is simply a count of the number of species inhabiting a given area or habitat. Species richness may also be thought of as the list of species that were counted. Patterns of species richness in space and time are some of the most foundational data in ecology, biogeography, and paleontology. These patterns both suggest and test theories of species richness dynamics. In addition to species richness, there are many concepts of species diversity. Usually species diversity is formed from species richness by further classifying the species by some attribute, such as abundance, size, or ecological role. When species are classified by abundance or population size, species diversity is the variability in the distribution of individuals into species. This species abundance relation and other measures of species diversity link species richness to many ecological processes such as population dynamics and competition. They also link to evolutionary processes such as adaptive radiation and the evolution of phenotypic plasticity. Species diversity is the building block for the diversity of higher taxa and for the diversity of ecological associations such as communities and biomes. Species diversity is an important policy tool because species diversity is relatively easy to measure and is understandable by many people. Conservation programs and goals are often cast in terms of the number of species that might be protected by a given action.

Introduction

The study of biodiversity often begins with species diversity because it is the most familiar aspect of biodiversity as a whole. Popular interest in biodiversity such as bird watching usually begins with learning to distinguish species, and this leads to a direct appreciation of species diversity. Later, other aspects of biodiversity, such as morphological variants or subspecies, the phylogenetic grouping of species, or the habitats and landscapes associated with particular species, are added to the initial view. Species diversity also lies at the heart of much of evolution and ecology, the sciences most concerned with biodiversity. Each discipline, in its own ways, seeks to explain as much as possible about species diversity and its patterns in time and space. Both sciences use the patterns of species diversity found in nature as tests for ideas about various ecological and evolutionary processes. The relation of species diversity to all other elements of the biodiversity hierarchies of taxa and ecological assemblages (such as communities and ecosystems) is also an important focus of these sciences.

In efforts to conserve the world's biodiversity, species diversity is a key policy concept. It is often the easiest to measure, and it is possible to frame management goals in terms of species diversity. As a practical matter, it is frequently necessary to use species diversity as a surrogate for other levels of biodiversity that may be much more difficult to measure.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00133-7

Here, we initially focus on species richness (the number of species in a given area) and then discuss other approaches to species diversity. Finally, we discuss the roles of species diversity in biodiversity policy.

Species Richness

The simplest form of species diversity is species richness, which is the number of species which occur in a given area. For example, the species richness of hippopotamuses in Africa is two. However, even within this simple definition there are many important issues. For example, species richness is thought of as a number, but corresponding to each number is also a list of species. The list (*Hippopotamus amphibius*, *Choeropsis liberiensis*) corresponds to the richness number of two for living African hippos.

Computing species richness requires choosing a concept of species and a higher taxon for which a list of species exists. There are many species concepts available, but most are grounded in the theory of evolution and have a theoretical basis. That is, they are entities for which there is a theory that provides methods of determining similarity and difference. The theory is what is used to resolve differences between various workers classifications. Species are thus in contrast to many other biodiversity entities such as higher taxonomic categories or ecological associations such as communities and ecosystems. These entities are more creations of humanity and their classifications are more artificial. Thus, a vegetation classification is created by ecologists arriving at a consensus for the purpose of accomplishing some work. Since diversity of any entities is a function of the method of classification which creates the units of diversity, theoretical kinds such as species (and genes) offer the possibility of more objective measures of diversity. As a practical matter, calculating species richness depends on the taxonomy available and hence on the taxonomic philosophy of the practitioners for a given taxon. Furthermore, species richness is a function of the rank of the taxon chosen: Species richness will depend on whether one chooses the family Hippopotamidae, the order Artiodactyla, the class Mammalia, and so on.

Species richness is most commonly expressed in terms of area. It may be expressed in terms of a political subdivision of the earth or an ecological subdivision of the earth, or it may be calculated by reference to an equal area grid placed arbitrarily with regard to political boundaries. In all cases, a process of discretization occurs. A species is counted simply as present or absent within a given area. Attributes such as the abundance of the species in the area or the percentage of the area occupied are not taken into account. For equal area grids we obtain an approximation of the underlying continuous variation in the species richness in such a way that we are able to quantify richness. **Figure 1** shows a map of species richness for the reptiles (crocodilians, turtles, lizards, and snakes) of the continental USA. The grid cells are 100 miles on a side or 10,000 square miles in area. **Figure 2** shows hand-drawn isolines of species richness which give a sense of the underlying continuous distribution.

However, species richness may also be expressed in terms of other kinds of aggregates of species. We may determine the species richness of the parasite load of a single animal or we may determine how many species of epiphytes live on a given tree.

As with any process of discretization, the patterns of species richness revealed are very much a function of the underlying scale of the discretization. **Figures 3** and **4** show species richness of terrestrial vertebrates for the state of Oregon at two scales of discretization. It is clear that the pattern of species richness is scale dependent. Any inferences made from a map of species richness must clearly take the effect of scale into account. It is worth noting that finer resolution maps are not necessarily better for all purposes. There is an uncertainty trade-off between accuracy and comparability in the choice of grid size. A smaller grid size is more accurate, but a larger grid size may produce cells that are more comparable since the effect of data of varying accuracy is smoothed out. For example, a coarser resolution map may provide better insights into large-scale biogeographic patterns because ecological noise may be averaged out. Furthermore, a minimum scale of discretization must be chosen that is appropriate given the accuracy of the individual range maps from which species richness is calculated. The data for **Figure 1** were compiled from the *Peterson Field Guides to Amphibians and Reptiles* (Conant and Stebbins). The range maps in these books show a great deal of cartographic generalization, so a relatively large-scale grid is appropriate. In general, it is a good idea to make the grid cell size several times larger than the spatial scale of the assumed errors in the individual maps.

Most work has been done with species richness as a number because the manipulation and analysis of large lists have only become possible with relatively recent advances in computer science. However, there is clearly more information in the list of species richness than the simple number. Two areas may easily have the same number of species but quite different lists. In general, the further apart two areas are, the fewer species they will have in common. Species richness thus has two components: The absolute number of species in an area and a relative measure of the uniqueness of one area compared to another. This measure is frequently computed by comparing the lists of species from the two areas and noting the number of species that are unique to each area and the number that are common to both areas.

The maps shown in **Figures 1**, **3**, and **4** are typical of species richness maps derived from the mapped ranges of individual species on which they depend. Calculation of species richness from a set of individual maps assumes that the maps are relatively consistent with regard to accuracy and underlying scale. This in turn requires that the taxonomy of the group under consideration be relatively complete and stable. This requirement is true only for certain taxa and for some parts of the world. In the USA and Europe, vertebrates, butterflies, and trees are generally sufficiently well-known to allow the construction of species richness maps. For the entire planet, only a few groups, such as birds, turtles, and milkweed butterflies, are sufficiently well-known to allow global species richness maps to be constructed. It is difficult or impossible to calculate species richness for poorly known taxa that have many undescribed species or whose taxonomy is controversial. Even for very well-known groups, it is not clear how up-to-date the range maps are since they are often necessarily based on collections made over a long period of time and

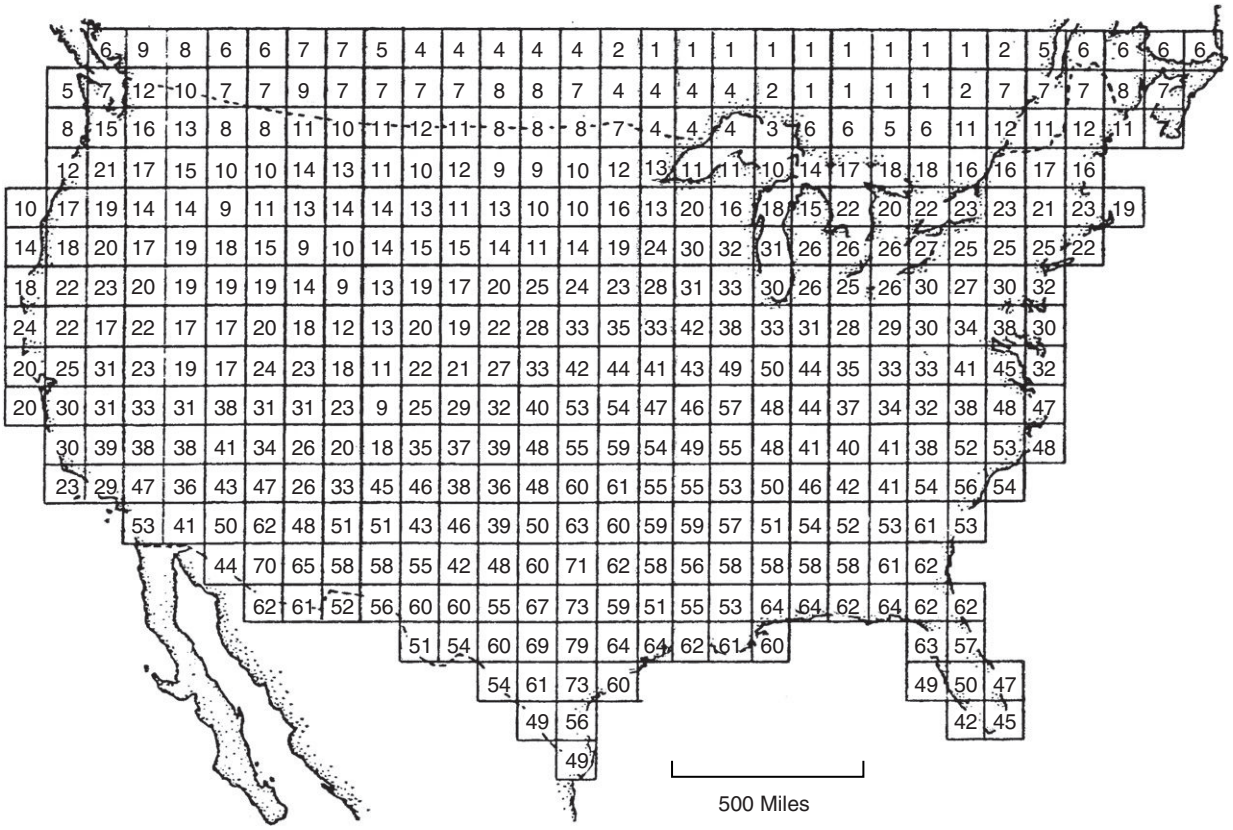


Figure 1 Species richness of North American reptiles; raw richness on a grid scale of 10,000 square miles per cell. Reproduced with permission from Kiester AR (1971) Species density of North American amphibians and reptiles. *Systematic Zoology* 20: 127–137.

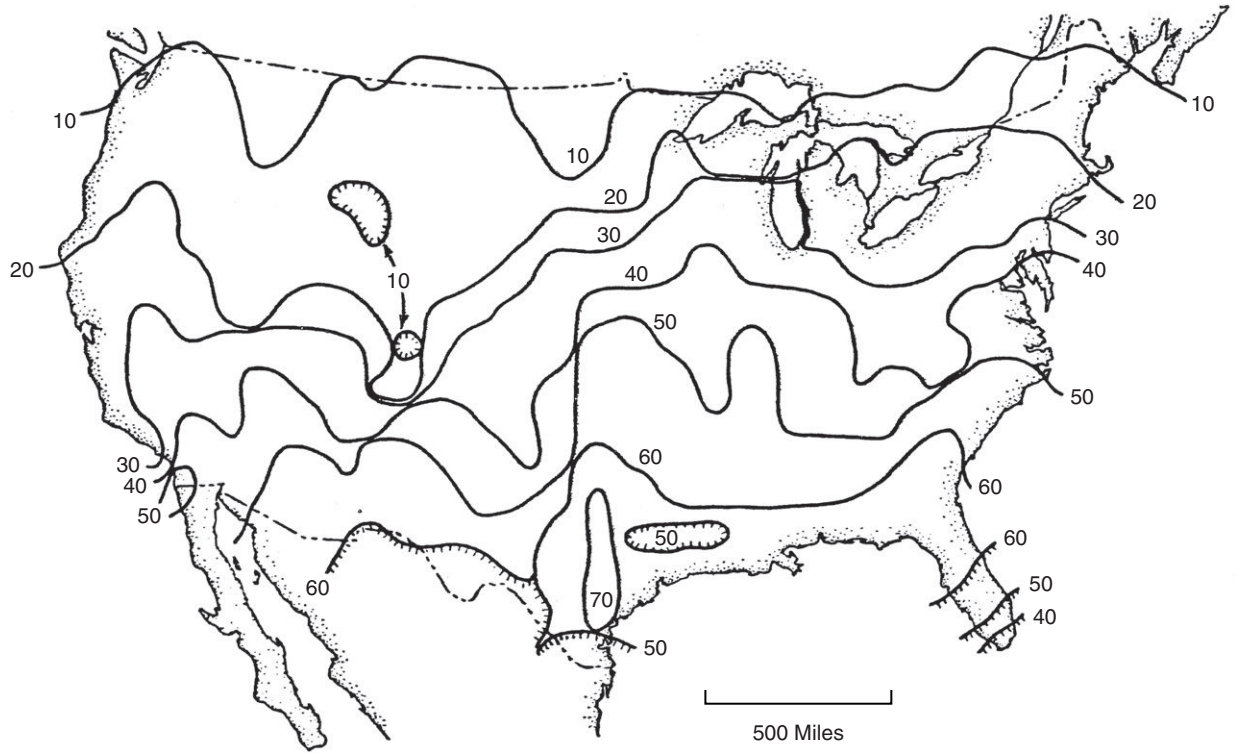


Figure 2 Isolines of North American reptile species richness. The lines were hand interpolated from the data of **Figure 1**. Reproduced with permission from Kiester AR (1971) Species density of North American amphibians and reptiles. *Systematic Zoology* 20: 127–137.

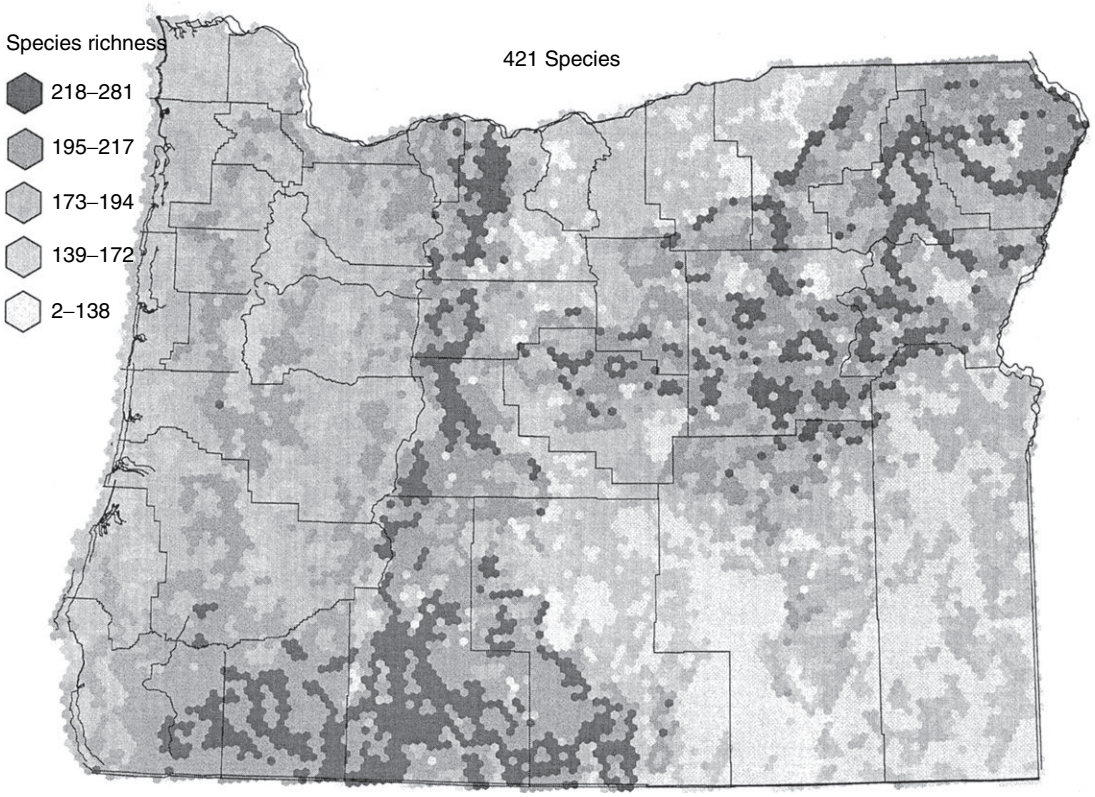


Figure 3 Terrestrial vertebrate (amphibians, reptiles, breeding birds, and mammals) richness for Oregon; raw richness on a grid scale of 71 km^2. Reproduced with permission from Kennelly, 1998.

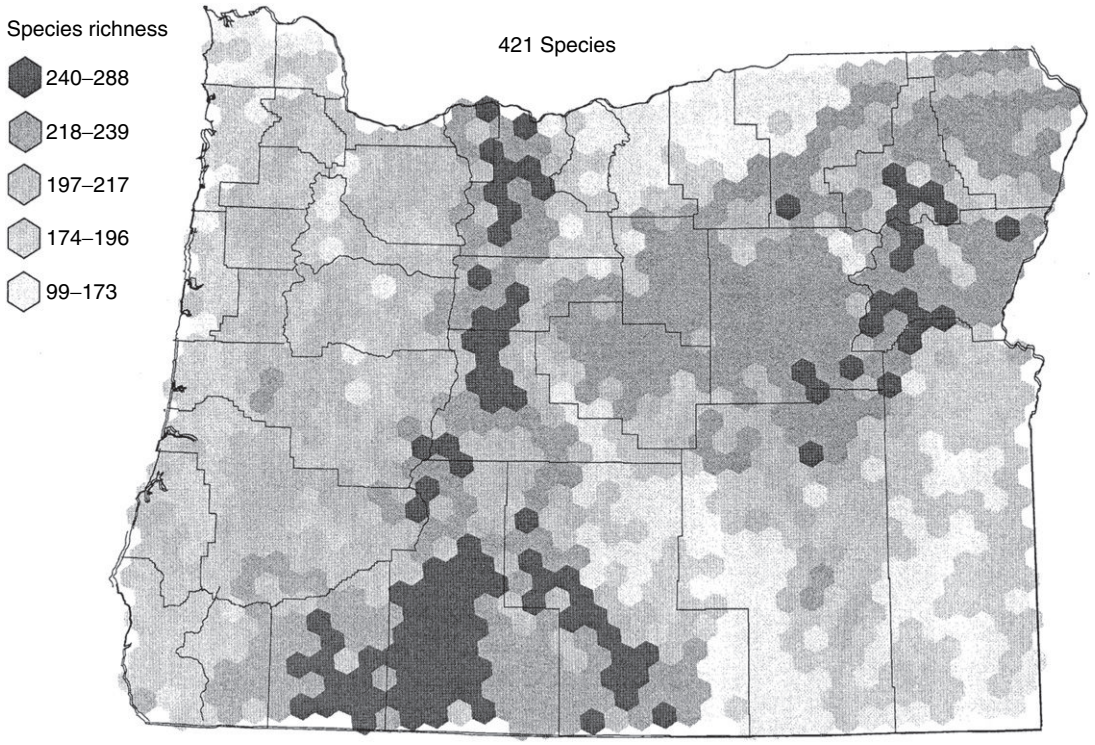


Figure 4 Terrestrial vertebrate (amphibians, reptiles, breeding birds, and mammals) richness for Oregon; raw richness on a grid scale of 213 km^2. Reproduced with permission from Kennelly, 1998.

almost never on a recent species-wide sampling program. For example, the data used to make the maps of Oregon in **Figures 3** and **4** were based on locality records during a 30-year period in order to have relatively complete coverage. Therefore, data for species whose ranges have changed drastically in the past two decades will be out-of-date and consequently the richness maps are in some sense always out-of-date.

There are two ways in which species richness data are accumulated. The first is the slow process of accumulation of the description of new species and the second is a more rapid method in which species richness of a given area is estimated from a statistically designed sampling effort.

As science progresses, our knowledge of species richness increases. This pattern, which is usually an increasing function of time, is called a species accumulation curve. We can think of species as being accumulated on several scales of effort and geography. For example, at the largest scale, that of all scientific efforts taken together for the entire world, we can see a historical accumulation of species. In 1758, Linnaeus recognized 11 species of turtles, in 1889 Boulenger listed 201, in 1992 Iverson listed 257, and in 2010 the Turtle Taxonomy Working Group listed 328 (Turtle Taxonomy Working Group, 2010). New species of turtles are still being described even though they are among the best known animals. The change from 1992 to 2010 also reflected a general change in the taxonomic philosophy used by turtle systematists. Many taxa that were listed in 1992 as subspecies were elevated to species rank adding species to the total by taxonomic reinterpretation rather than discovery. The number of turtle species has not increased due to any organized or statistically designed effort. The work has simply proceeded as a haphazard attempt at exhaustive enumeration.

In contrast, several attempts have been made to statistically estimate species diversity at a particular location. Usually these efforts determine both the number of species and the number of individuals. In such efforts the number of species caught is a biased estimator since rare species are often necessarily missed. The contrast between haphazard enumeration and statistical estimation is that between more complete knowledge without error estimates and less complete knowledge with error estimates. The best method of synthesizing these two approaches via meta-analysis is an outstanding research problem.

For many taxa and most localities throughout the world, species richness is known only for certain restricted localities such as refuges and research field stations. In some cases, the exact area of knowledge is known, but in others the site may be only generally described. These point locality data are sometimes used to extrapolate species richness over larger areas. This is a new area of research reviewed by Colwell and Coddington (1994), who found many unanswered questions. However, extrapolation may be the best hope of obtaining an estimate of global diversity in the foreseeable future.

Patterns of Species Richness

The most basic pattern of species richness is that between the number of species and the size of the area considered. In general, the larger the area, the more species will be found in it. This species–area relationship frequently takes the form $S = cA^z$ where S is the number of species, A is the area considered, and c and z are estimated constants. By taking logarithms of both sides of this equation, we find that the logarithm of species richness is a linear increasing function of area. However, many empirical studies have found that the data scatter about a wide band of values of z and that other mathematical forms of the relationship may be equally likely. The analysis of the mechanisms by which area determines species richness is a major line of inquiry for ecologists and biogeographers.

Species richness is very unevenly distributed over the entire Earth. The most important general pattern is that richness increases as one goes from the poles to the equator. **Figure 2** shows that this latitudinal gradient in species richness can be quite striking. The average number of species of reptiles increases from 10 at the Canadian border to more than 60 at the Mexican border (**Figure 2**). Of course, many taxa are exceptions (penguin species diversity is highest closer to the south pole and there is only one species at the equator), but overall the pattern is strong. Again, studies of the latitudinal gradient in species richness have been a major line of inquiry for ecologists and biogeographers. Another pattern is the peninsula effect whereby species richness decreases toward the end of a peninsula. For example, the number of species of mammals and reptiles decreases in Florida and Baja California as one moves toward the end of these peninsulas. This peninsular pattern can be seen for reptiles on the Florida peninsula in **Figure 2**.

Species richness varies strikingly in time. The paleontological record provides good evidence that overall species richness increases with time but that there may be massive extinction events that drastically reduce species richness in a relatively short time. At ecological timescales species turnover is common so that the list of species is often changing. The total number of species usually changes less and usually in response to a change in the environment.

Determinants of Species Richness

Over the continuum of scales of time and space, the sciences of ecology, biogeography, and evolution all strive to understand the processes which determine species richness. Indeed, concern for species richness unites these fields (as well as bridging fields such as macroecology and paleoecology). One can easily imagine that they are all part of a single science concerned with species richness. McKinney and Drake (1998) proposed that this field be called the study of biodiversity dynamics and that it should include all kinds of diversity, not just species diversity.

The first question to be addressed is the role of history and the question of equilibrium versus nonequilibrium dynamics. Are the patterns of diversity that we see a result of the vicissitudes of history or of the action of relatively general mechanisms? For example, is the latitudinal gradient in species richness in North America fully recovered from the end of the last ice age. For certain taxa with good dispersal abilities, such as birds or butterflies, it is very likely that they have finished responding to that episode of climate change and are near equilibrium. Other groups, such as salamanders, with very

poor dispersal capabilities may still be in the process of moving into areas opened up at the end of the glaciation and hence their latitudinal gradient may not yet be at equilibrium. Another way of viewing this question is to ask to what extent are there general mechanisms that determine species richness in particular cases and to what extent are there only particular cases. Some ecologists believe that there are only case studies because the details of any particular case overwhelm any instance of a potential general pattern. Others are much more sanguine that general rules and even laws will be able to account for species richness. In any event, for all of the biodiversity sciences the relation between the general and the particular is a difficult problem.

One useful way of dealing with the issue of the general and the particular was developed by Williams (1969), who introduced the distinction between the distant and close views in island biogeography. Island biogeography traditionally has received special attention from students of diversity because the patterns and apparent processes seem to be simpler and easier to understand than those on continents. In Williams' dichotomy the distant view is typified by MacArthur and Wilson's (1963) theory of island biogeography that accounts for the patterns of species richness seen on islands that are of different size and different distances from a source mainland. This model predicts an equilibrium number of species as a function of rates of colonization and extinction. These rates are statistical composites formed from an ensemble of many species and are the mechanisms of the distant view. Here, only the number view of species richness is used. The equilibrium number of species on an island is predicted, but the actual list of species is not. This is not viewed as a problem because the theory predicts that species are constantly turning over and so the list is always changing. The close view is taken by Williams (1969) in his work on the lizards of the genus *Anolis* of the Caribbean or by Lack (1976) in his study of the land birds of Jamaica. In this approach, the focus is on a list of species and the details of the biology of each species examined are considered important. Mechanisms such as colonizing ability, competition, ecological plasticity, and the details of the interactions of species and available habitats are studied on a species by species basis. Adaptive radiation is often considered an important mechanism for generating diversity on a given island. The distant view has the strength of numbers and interprets broad patterns well. It does not account for the actual biology of any particular species. The close view gives detailed accounts of the biology of particular species but generally is not able to predict the number of species. Thus, the distant view focuses on the number version of species richness, whereas the close view is concerned with the actual list of particular species. The close view is not contained within the distant view. Rather it is complementary to it.

A Diversity of Species Diversities

Beyond species richness are an enormous number of species diversities. Generally, these take the form of a species richness of species classified by some property of the species. Species may be classified by abundance (population size), body size, trophic position, taxa, membership in various ecological assemblages, and in many other ways. As with species richness, the actual number of species in a given class or the list of species in a given class may be analyzed. Because there may be several classes to which individual species are assigned, it is possible to develop an enormous variety of species diversity indices which summarize in various ways the pattern of distribution of species into the classes of a particular classification. The study of these diversity indices, their sampling properties, and the inferences about ecology and evolution that can be drawn from them have been the subject of major scientific activity in the past few decades.

The most common extension of species richness is to the species abundance relation. Here, a set of species is classified by the abundance or population size of each species. Some species are rare and others common; the distribution of all species abundances is called the species abundance curve. **Figure 5** shows the species abundance curve for a classic data set of macro-Lepidoptera (mostly large moths) collected by a light trap at Rothamsted in England in 1935. A total of 6814 individuals of 197 species were collected. Thirty-seven species were represented by a single individual, whereas the most common species was represented by 1799 individuals. In this example, the number of classes of the classification is just the number of different integer valued numbers of abundance. All abundances greater that 65 were represented by single individuals. As with many data sets of species abundance, this one fits a lognormal distribution well. The many attempts to

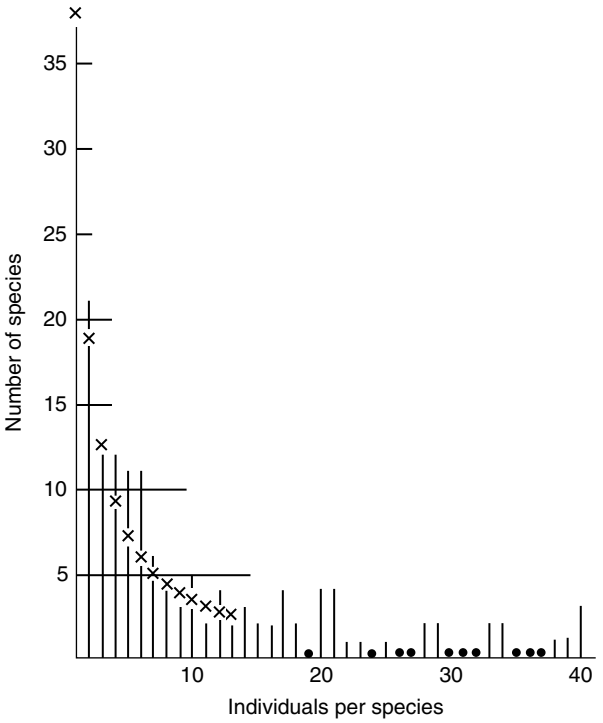


Figure 5 Species abundance curve for macro-Lepidoptera captured by light trap at Rothamsted, England, in 1935. The *x*'s represent the fit of a lognormal distribution. Beyond an abundance of 40 are 26 values ranging from 42 to 1799, each representing 1 or 2 species. Reproduced with permission from Williams CB (1964) *Patterns in the Balance of Nature and Related Problems in Quantitative Ecology*, 324 pp. London and New York: Academic Press.

understand these data both as empirical distributions and as data to test predictions made by theories of community structure have been reviewed by May (1975) and Magurran (1988). Many of the indices fall on a continuum from emphasizing rare species to emphasizing common species. Simple species richness is at one end of this continuum since it weights all species equally. At the other end is the Simpson index, which gives the most weight to the species with the most individuals. Of the many empirical indices of species diversity that combine species abundance data, Lande (1996) found that only the Simpson index is statistically unbiased and leads naturally to measures of similarity between multiple communities. Using both species richness and the Simpson index seems to provide a useful characterization of the species abundance properties of a community.

Living organisms vary enormously in size and it is commonplace that there are fewer large species than small species. Hutchinson and MacArthur (1959) pioneered an ecological approach to understanding the species diversity of body sizes. **Figure 6** shows the patterns of species diversity by body size for mammals over three scales in North America. At the scale of the continent, the distribution is unimodal centered on medium and small sizes. At the finest scale (ponderosa pine community of the Oregon Cascade Mountains) the distribution is near flat and shows distinct clumps. This area of research has many competing hypotheses which have been reviewed by Kelt and Brown (1998). Among these hypotheses, a basic idea is that there is more environmental variation at smaller size scales: The world looks more complex to a small beetle than to a larger grazing mammal. This environmental complexity in turn provides a great number of potential niches at smaller scales, allowing coexistence of more species. A provocative analysis by Holling (1993) presented evidence that species sizes of birds and mammals in individual ecological communities are clumped into a small number of discrete, separate size groups. He interpreted this pattern to be the result of a small number of structuring processes ranging from vegetative growth to disturbance dynamics and geomorphological processes which form a discontinuous temporal hierarchy and generate the discontinuous size distributions.

The familiar concept of a food web leads to a classification of species by their position in the web. Here, the basic pattern is that there are fewer species higher in the food web: There are fewer species of top predators than herbivores. Beyond these simple trophic level generalizations, the field is controversial (Yodzis, 1993). One intriguing result concerns the relationship between the species diversity of a web and the number of dominant cliques in the web. A dominant clique of a food web is a group of species in which the predator species share at least one prey species in common and in which the groups are isolated from other such groups and do not share any of their prey species. **Figure 7** shows this relatively strong relationship. This pattern may be interpreted to mean that food webs are organized into functional units (groups of interacting predators and prey) and that species diversity increases with the number of such groups.

The species diversity of higher taxa ranging from genera to phyla varies a great deal: The phylum Uniramia, which contains the myriapods and insects, has well over 1 million described species "but many millions more probably await

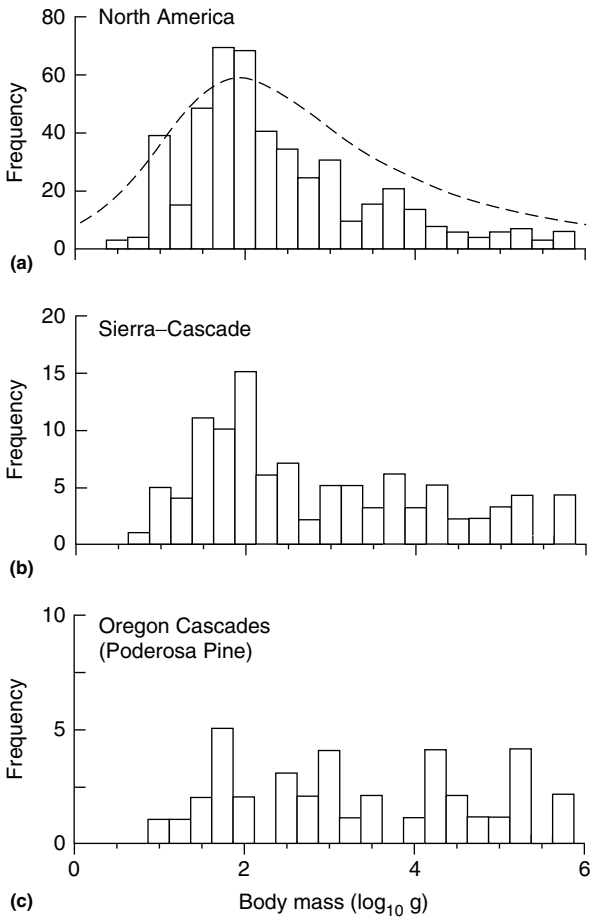


Figure 6 Species diversity of body size for North America mammals at three scales: (a) continental North America, (b) Sierra Nevada and Cascade ranges, and (c) Oregon Cascades ponderosa pine habitat. Reproduced with permission from Kelt DA and Brown JH (1998) Diversification of body sizes: Patterns and processes in the assembly of terrestrial mammal faunas. In: McKinney ML and Drake JA (eds.) *Biodiversity Dynamics*, pp. 109–131. New York: Columbia University Press.

scientific discovery" (Barnes, 1998, p. 247), whereas the phylum Placozoa apparently has only a single species – *Trichoplax adhaerens* (Barnes, 1998, p. 179). Other ranks in the taxonomic hierarchy vary similarly in their diversity. We are just beginning to understand the patterns of species diversity of higher taxa and clearly this requires a synthesis of evolutionary and ecological theories. Why orchids and beetles are more diverse than many comparable groups is a function of evolutionary plasticity and ecological opportunity. Evolutionary plasticity is a result of both genetic diversity and developmental constraints and opportunities, whereas ecological opportunity is influenced by structural diversity and energy availability. Currently, we lack a detailed understanding of both of these concepts. However, recent paleontological work has demonstrated some interesting patterns of the temporal structure of species diversity of higher taxa. In particular, as origin and extinction probabilities of higher taxa are followed from the Paleozoic to the present there appears to be an accumulation of higher taxa with lower extinction

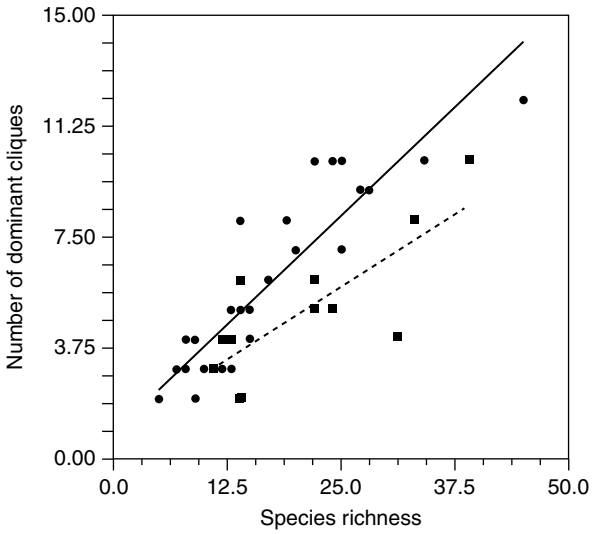


Figure 7 Species richness versus number of dominant cliques for 40 food webs. Circles represent data from food webs from fluctuating environments, and the solid line is a linear regression for these points. Squares represent data from food webs in constant environments, and the dashed line is a linear regression for these points. Reproduced with permission from Yodzis P (1993) Environment and trophodiversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 26–38. Chicago: University of Chicago Press.

probabilities. Certain groups seem to have greater staying power than others and so they are more prevalent today. These patterns offer new possibilities for testing ideas about the evolution of higher taxa species diversity.

At the other end of the taxonomic hierarchy, species may be classified by their internal variability. This variability has something of a hierarchical structure of its own: subspecies, geographic structure, population structure, and genetic structure. Subspecies represent the largest scale of variation within species and are usually thought of as being composed of a major portion of a species range within which several characters are concordant. Geographic variation may also be continuous as with clines or nonconcordant with different characters varying with different geographic patterns. Population structure ranges from continuous to highly fragmented and island-like with many varieties of metapopulation structure in between. Genetic structure follows in part from population structure and provides variability beyond that. The degree of genetic polymorphism shown by local populations provides the most basic variation shown by species. There have been few attempts to quantify species diversity of within-species diversity, although such analyses should lead to further integration of ecological and evolutionary approaches to species diversity.

Summary: Structure of Taxa and Ecological Assemblages

Almost any classification of species results in the fact that species richness of the classes is unequal and often shows interesting patterns. The description of these patterns and their interpretation in terms of evolutionary and ecological processes is a cornerstone of biodiversity science. Two of the most important kinds of classification are those that classify species by membership in higher taxonomic categories (which, it is hoped, represent phylogenetic groups as well) and into various ecological assemblages such as food webs, communities, landscapes, and biomes. These patterns of species diversity function both to generate and to test hypotheses about the phylogenetic and ecological processes. A great deal of effort has been devoted to analyzing some classifications such as abundance. However, it is clear that attempting to explain the pattern produced by a single classification will inevitably be incomplete. Abundance, body size, and trophic position are strongly related – large top predators are uncommon – and rather than try to explain these characteristics separately, work now focuses on attempting to explain the whole syndrome. The methodology of progressive synthesis outlined by Ford (2000) offers one way to counteract the usual pattern of fragmenting questions. The ultimate challenge to students of biodiversity is to devise better ways of understanding how the phylogenetic and ecological classifications of species diversity interact.

Species Diversity and Biodiversity Policy

Of the many levels of the hierarchy of biological diversity, from genes to the world, species diversity historically has played the dominant role in attempts to maintain and conserve biodiversity. Generally, this has led to a species by species approach such as is exemplified by the Endangered Species Act. This approach has usually been reactive and has been analogized to emergency room medicine. However, recent efforts such as the Gap Analysis Program have focused on sets of species and attempted to lay out proactive approaches. These approaches have in turn been analogized to public health medicine. Generally, it is less expensive to protect a species from extinction while it is relatively common rather than to wait until is it endangered. Both of these approaches use species diversity as a key level of biological diversity (gap analysis also uses vegetation communities as an important level). This focus on species diversity is due to several factors. First, the idea of conserving species is well understood by many people who have a good intuitive understanding of the concept. Second, as mentioned previously, species are scientifically well-defined entities. Finally, species diversity is apparently a good surrogate for many other levels in the hierarchy, such as genes, which are much more difficult to measure in nature. Therefore, it is likely that species diversity will remain a central focus of conservation biology even as other levels of biodiversity become better understood.

See also: Diversity, Community/Regional Level. Diversity, Molecular Level. Diversity, Organism Level. Diversity, Taxonomic Versus Functional. Landscape Diversity. Species–Area Relationships

References

Barnes RSK (ed.) (1998) *The Diversity of Living Organisms*. Oxford: Blackwell.

Colwell RK and Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London B* 345: 101–118.

Ford ED (2000) *Scientific Method for Ecological Research*. Cambridge, UK: Cambridge University Press.

Holling CS (1993) Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological monographs* 62: 447–502.

Hutchinson GE and MacArthur RH (1959) A theoretical ecological model of size distribution among species. *The American Naturalist* 93: 117–123.

Kelt DA and Brown JH (1998) Diversification of body sizes: Patterns and processes in the assembly of terrestrial mammal faunas. In: McKinney ML and Drake JA (eds.) *Biodiversity Dynamics*, pp. 109–131. New York: Columbia University Press.

Kiester AR (1971) Species density of North American amphibians and reptiles. *Systematic Zoology* 20: 127–137.

Lack D (1976) *Island Biology*. Berkeley: University of California Press.

Lande R (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5–13.

MacArthur RH and Wilson EO (1963) An equilibrium theory of insular zoogeography. *Evolution* 17(4): 373–387.

Magurran AE (1988) *Ecological Diversity and its Measurement*. Princeton, NJ: Princeton University Press.

May RM (1975) Patterns of species abundance and diversity. In: Cody ML and Diamond JM (eds.) *Ecology and Evolution of Communities*, pp. 81–120. Cambridge, MA: Harvard University Press.

McKinney ML and Drake JA (eds.) (1998) *Biodiversity Dynamics*. New York: Columbia University Press.

Ricklefs RE and Schluter D (1993) *Species Diversity in Ecological Communities*. Chicago: University of Chicago Press.

Tokeshi M (1993) Species abundance patterns and community structure. *Advances in Ecological Research* 24: 111–186.

Turtle Taxonomy Working Group (Rhodin AGJ, van Dijk PP, Iverson JB, and Shaffer HB) (2010) Turtles of the world, 2010 update: Annotated checklist of taxonomy, synonymy, distribution, and conservation status. In: Rhodin AGJ, Pritchard PCH, van Dijk PP, *et al.* (eds.) *Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group*. Chelonian Research Monographs No. 5, pp. 000.85–000.164, doi:10.3854/crm.5.000.checklist.v3.2010, http://www.iucn-tftsg.org/cbftt/

Williams CB (1964) *Patterns in the Balance of Nature and Related Problems in Quantitative Ecology*, 324 pp. London and New York: Academic Press.

Williams EE (1969) The ecology of colonization as seen in the zoogeography of anoline lizards on small islands. *Quarterly Review of Biology* 44: 345–389.

Yodzis P (1993) Environment and trophodiversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities*, pp. 26–38. Chicago: University of Chicago Press.

Species Interactions

Jessica J Hellmann, University of Notre Dame, Notre Dame, IN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Community The collection of all species that occur together in space and time.
Intrinsic rate of growth A parameter, often designated as the variable *r*, reflecting the rate of change in population size due to births and deaths when resources are not limiting.
Population A collection of individuals of the same species in a defined geographic area that interbreed.
Population dynamics The change in the number of individuals in a population over time.
Resource Any necessity that is used or consumed for organismal growth, reproduction, or survival, including food, space, shelter, and nutrients.
Species A group of individuals that is capable of interbreeding; a species is composed of one or more populations.
Trait A quality possessed by an individual organism that affects its ability to survive and reproduce.

Introduction

The world can be divided into two parts – the living and the nonliving. Scientists refer to the nonliving component of the earth as the abiotic environment, and all living organisms are labeled biotic. Organisms survive and reproduce under many abiotic and biotic influences. Abiotic nutrients and elements are needed for survival and reproduction, and nutrient limitation is a common factor regulating limits to growth. Biotic factors also influence how an organism lives and grows. When organisms come into contact, they can facilitate, control, harm, use, or even help to propagate each other. Biotic associations that involve members of two or more different species are called species interactions.

Because the world contains a tremendous diversity of species, interactions among the biota are common. Many of these interactions are conspicuous and have a strong influence on the organization of biological communities. Studies of species interactions have produced a large body of data in a great variety of ecosystems, and these studies comprise an entire discipline called community ecology. The treatment of species interactions given here will briefly discuss the types of interactions found in nature and the simple mathematical representations of their effects, discuss the role evolution plays in shaping interactions, and provide some empirical examples. Because an increasing number of community ecologists are examining the effect of human influences on species interactions, the intersection of community ecology and conservation biology is also briefly discussed.

Pairwise Interactions

Species interact in a tremendous variety of ways, and associations can involve two, several, or many species. First, the various kinds of interactions that include only two species are discussed. (A discussion of multiple species interactions is provided in the section Multispecies Interactions.) When individuals of two species interact, this association is called a pairwise interaction. Pairwise interactions, like all interactions, are described by the direction and the strength of their effects. Direction refers to whether the impact on each associating organism is positive or negative, and interaction types are classified according to the direction of their effect (Table 1). Interaction strength refers to how much an interaction controls species' dynamics.

More specifically, the strength of an interaction is a measure of how much one species affects the population size and growth rate of another species. Strength is often detected as the effect on the population of one species by removing the second species. Imagine two interacting species, species A and B. If we would like to test the effect of B on A, we could remove B and measure the resulting change in population size of A. If species A responds to the removal of species B with a large change (increase or decrease) in abundance or population growth rate, we say that B has a strong effect on A. If little effect on the population of species A is seen by removing species B, the interaction plays a small role in the population dynamics of species A, and the influence of B on A is weak. In turn, the effect of A on B can be tested by removing A and observing how the population size or growth rate of B responds.

Table 1 Pairwise interaction types

Direction of effect on species 1	*Direction of effect on species 3*		
	+	−	*None*
+	Mutualism	Predation/parasitism	Commensalism
−	Predation/parasitism	Competition	Amensalism
None	Commensalism	Amensalism	–

Many factors determine how strongly two species interact. For example, the population size of either species, the behavior and age of the interacting individuals, the abundance and availability of abiotic resources, and the patchy nature of the environment all influence how strongly two species affect one another. In other words, understanding the nature of an interaction is more complicated than simply identifying that two species come into contact. In fact, even a simple experiment of interaction strength such as that described in the previous paragraph may not provide enough information on exactly how organisms interact with one another in different settings or circumstances. For example, the strength of an interaction can change at different times through the seasons or in different places in the species' habitat. The strength of pairwise interactions can also change due to the forces of evolution, but first the types and effects of pairwise interactions in ecological time are discussed.

Competition

Competition is characterized by the common use of resources by two or more species, and the effect of competition on all participating species is negative (**Table 1**). Members of the same or different species can compete for resources. Within-species competition is known as intraspecific, and between-species competition is called interspecific. Of these, interspecific competition meets the definition of a species interaction, and intraspecific competition will not be discussed.

Interspecific competition can be further divided into two categories: interference and exploitative. Interference competition is the more direct form of competition, in which individuals of one species actively dominate a resource, preventing or decreasing the access of another species to those resources. Exploitative competition, however, is less combative. It is the diminishment of a limited resource by a species so that the other species is not able to utilize as much of the resource as it would in the absence of the exploitative competitor.

We can predict how the population dynamics of two competing species will be affected by their interaction with the use of a mathematical model. A mathematical theory of competition began with the work of Lotka and Volterra in the late 1920s and early 1930s. The Lotka–Volterra model quantifies the influence of competition by expressing the population dynamics of each species in an equation that includes both growth in the absence of competition and the influence of competition itself. In other words, Lotka–Volterra represents the change in population size through time as a function of the growth rate of each species and the detrimental effect of competition on each population. Remember that the presence of competitors causes harm. This harm reduces the total number of individuals in each population that would be possible in the absence of competitors.

In the Lotka–Volterra model, two equations are needed to represent the population dynamics of competition (one equation for each species). The pair of equations describes the rate of change in population size with time for the two species. Subscripts represent a variable's species designation; hence, N_1 refers to the population size of species 1, and N_2 refers to the population size of species 2. The relative effect of each species on the population size of the other is determined by the value of the competition parameter α. This competition parameter can also be interpreted as the amount of overlap between the two species in the resources they require.

The effect of one member of species 2 on one member of species 1 is represented by the parameter α_{12}, and the effect of species 1 on species 2 is the value α_{21}. The values of α_{12} and α_{21} can range from 0 to 1. If they take on a value close to 1, then the effect of one species on the other is very strong. If they take on a value close to 0, then the effect of one species on the other is weak. The values of α_{12} and α_{21} can be different from each other because it is not necessary that the two species affect each other equally. If the values of α_{12} and α_{21} are the same, this special case is called symmetric competition.

If both α_{12} and α_{21} are zero, then the two species do not compete at all and grow according to a simple expression known as logistic growth. In logistic growth, a population grows from zero to some carrying capacity, K, according to an intrinsic growth parameter, r, as shown in **Figure 1**. The hallmark of logistic growth is its density dependence. In density-dependent growth, the expansion of a population slows as the number of individuals in the population approaches the carrying capacity or the maximum number of individuals that a habitat can support. All of the terms in the following equation for species 1, except the component $\alpha_{12}N_2$, represent the logistic growth of species 1. The remaining $\alpha_{12}N_2$ term reflects the slowing or reduction in population growth of species 1 that results from competition with species 2. The same is true for the equation of species 2; the term $\alpha_{12}N_1$ represents the negative effect of competition with species 1

$$\text{Species 1}: \frac{dN_1}{dt} = \frac{r_1 N_1 (K_1 - N_1 - \alpha_{12} N_2)}{K_1}$$

$$\text{Species 2}: \frac{dN_2}{dt} = \frac{r_1 N_1 (K_2 - N_2 \alpha_{21} N_1)}{K_2}$$

on population 2's otherwise logistic growth.

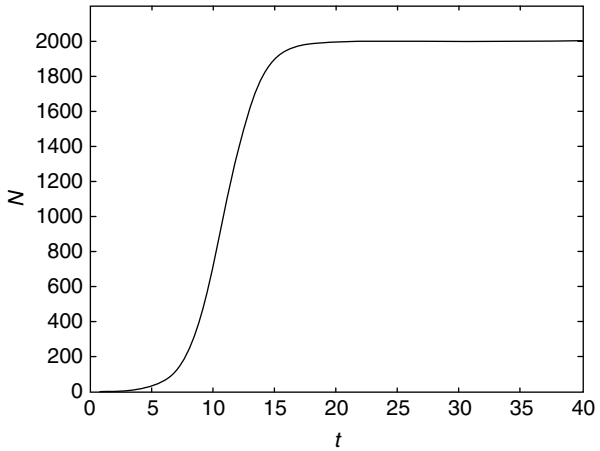


Figure 1 Sample plot of logistic growth. N is the number of individuals in a population of a single species, and t is time. In logistic growth, the number of individuals increase quickly when the population size is small, but the rate of growth decreases to zero as N approaches the carrying capacity. Here, the carrying capacity is 2000 individuals.

These equations can predict the impact on N_1 and N_2 for different values of r_1, K_1, α_{12}, r_2, K_2, and α_{21}. For some values of r, K, and α, one species will drive the other to extinction. In other cases, however, each species can maintain a positive population size in the presence of the other ($N_1>0$ and $N_2>0$). One such case is shown in **Figure 2**, in which the population size of both species is plotted against time. To produce this graph, one begins with population sizes just larger than zero for both species and uses the competition equations to predict the change in population sizes (N_1 and N_2) as time passes. Note that in **Figure 2**, both species eventually reach a stable and constant population size.

What conditions produce an outcome in which two competing species can coexist? It can be shown that coexistence depends on the relative value of the carrying capacities of the two species in comparison to the values of the competition coefficients, α. In short, coexistence is possible whenever the value of the competition coefficients is less than the ratio of the carrying capacities (if $\alpha_{12}<K_1/K_2$ and $\alpha_{21}<K_2/K_1$). This makes intuitive sense: Two species are able to persist on the same resource if they have a relatively small effect on one another. In other words, there is a constraint on the amount of resource overlap (α_{12} and α_{21}) between two competing species that is given by ratios of their carrying capacities.

The idea that two species, based on the Lotka–Volterra equations, cannot significantly overlap in resource use is known as the competitive exclusion principle. Generally stated, no two species can occupy the same niche. (A niche can be imprecisely defined as a species' role in a community or the entire set of resources that a species requires.) If the same niche was filled by two species, or if two species used exactly the same resources, the species would compete strongly, their α values would approach 1, and one of the two would be driven to extinction.

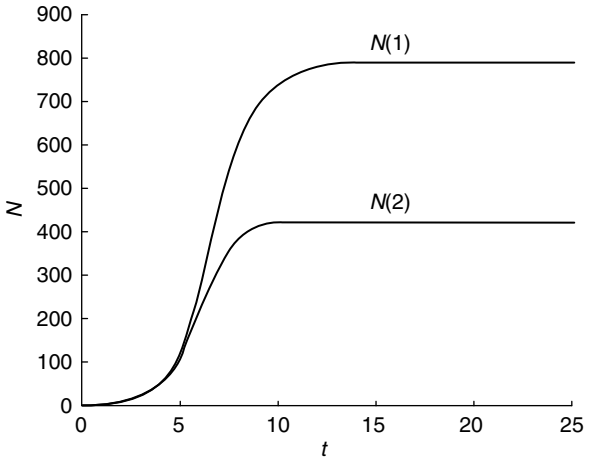


Figure 2 Sample plot of the population growth of two competing species according to the Lotka–Volterra competition equations. *N* is the number of individuals; *N*(1) refers to the number of individuals in one species, and *N*(2) refers to the number of individuals in the other species. In this diagram, species 1 and species 2 coexist because the abundance of both populations is positive through time. The growth of each population appears logistic (see **Figure 1**), but because of competition, the level at which each curve stabilizes is less than the carrying capacity for each species. In other words, the number of individuals in a population in a competitive environment is smaller than the total number of individuals that is possible in the absence of a competitor.

Additional factors not captured in the Lotka–Volterra equations influence coexistence, however, and species may coexist in situations in which classical theory would predict otherwise. The role that these factors play in coexistence reduces the general applicability of the competitive exclusion principle, and a number of more sophisticated competition models include these factors for specific situations. For example, spatial and temporal heterogeneity and variability of resources may enable a greater diversity of species to persist in an area than simple theory would predict (e.g., see Powell and Richerson, 1985; Tilman, 1994). In other words, the patchy nature of resources in a habitat may enable two species to reduce their overall competitive overlap. Alternatively, species may compete only when resources are in short supply.

Controversy regarding the role of competition in natural communities has resulted in a tumultuous debate in the ecological literature. Some authors argue that competition is a critically important factor regulating community composition. Others contend that competition is overemphasized at the expense of other types of interactions that are equally (if not more) important in determining the number and kinds of species that can persist in a community. This debate as to the importance of competition in controlling community structure is ongoing, but there is little argument that competitive interactions can be found and quantified.

Gause was the first to study competition in 1934 using experiments with two Paramecium species that shared a single resource. Gause found that one species always went extinct when both species were put together, suggesting that, as the Lotka–Volterra theory predicted, coexistence of two competitors is possible only in special cases. In 1946, however, Cromie placed two beetle competitors in a mixed environment of flour (i.e., resources) and small pieces of fine glass tubing (i.e., environmental complexity) and found that the species coexisted. Interestingly, in the absence of tubing coexistence was not possible, and one species always drove the other to extinction in those instances. Cromie's experiments suggest that some factors or influences that are not found in the simple Lotka–Volterra model may be important in facilitating coexistence. Since the time of Gause and Cromie, many other ecologists have studied the role of competition in natural ecosystems. In the early 1980s, Schoener and Connell each reviewed the literature for the prevalence of competition and found that a majority of published field studies successfully detected the impact of competition on the population dynamics of species (Connell, 1983; Schoener, 1983). A second review a decade later confirmed these findings but found a wide range of competitive effects with significant differences in the strength of competition among habitats and different types of species (Gurevitch *et al.*, 1992).

Predation

Predation is the consumption of one species by another. Consumption leads to a negative effect on the population of one of the two interacting species, known as the prey, and a

positive effect on the second population or species, known as the predator (Table 1). Also as in competition, predation has a rich history of mathematical theory, beginning with the work of Volterra. The predator–prey interaction is modeled with two equations: one for the prey (species 1) and one for the predator (species 2). Again, the components of these equations are discussed without going into great detail

$$\text{Species 1} : \frac{dN_1}{dT} = rN_1 - aN_1N_2$$

$$\text{Species 2} : \frac{dN_2}{dt} = baN_1N_2 - dN_2$$

According to this model, in the absence of a predator, species 1 would grow exponentially according to the rN_1 portion of the first equation; unlike the Lotka–Volterra competition equations, no carrying capacity limits growth in this model. This type of growth is called density independent. In other words, the rate of increase in the prey population remains constant through time and is independent of the number of individuals in the two populations. (Again, r represents the intrinsic rate of growth.) The number of prey removed from the prey population by predation is represented by the term aN_1N_2, where a describes the fraction of encounters between predator and prey that ultimately result in kills. Predators consume prey for a single purpose – to produce new predators. The number of predators born (i.e., new predators added to the N_2 population in the second equation) is a function of how efficient predators are at producing offspring than from consuming prey. The variable b describes this efficiency, and the term baN_1N_2 thus represents births into the predator population. In this model, prey die only from predation, but predators die from other causes, and this death rate is the variable d. The number of individuals removed from the predator population due to death is the term dN_2, and the fraction of predators removed is constant through time.

Several outcomes are predicted from this simple model. First, predators can drive prey to extinction by consuming them faster than prey can reproduce. Second, the predators can go extinct if their death rate exceeds their ability to produce new predators from the consumption of prey. Lastly, coexistence is possible. However, unlike the competition situation discussed in the previous section, these predator–prey equations do not reach a stable and persistent number of predators and prey in cases of coexistence; instead, the abundance of each species cycles through time. These cycles are called oscillations, with the predator abundance lagging behind that of the prey. As the number of prey increases, it is followed by an increase in the number of predators, driving a decrease in prey, and in turn driving a decrease in the number of predators. Once the predators decrease, the pressure of predation declines and prey can increase again, and the cycle repeats. The cyclic dynamics of predator–prey can be seen in Figure 3.

To determine whether the oscillations predicted by the Volterra theory reflect the true nature of predator–prey interactions, researchers have sought evidence of population cycles in nature. The most well-known example of predator–prey oscillations was shown in lynx feeding on hare in Canada. Recent evidence, however, calls into question whether or not the oscillations observed in the lynx–hare system are, in fact,

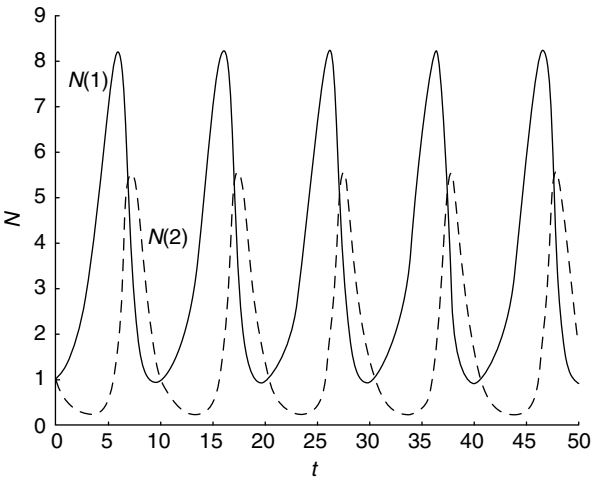


Figure 3 Sample diagram of a predator–prey system as predicted by the Volterra equations. N is the number of individuals $\times$ 103 and t is time. The solid line, $N(1)$, is the number of individuals in the prey population, and the dashed line, $N(2)$, is the number of individuals in the predator population. Both predator and prey coexist in this diagram because neither goes extinct. The abundance of predator and prey is not stable through time but goes through oscillations or cycles. Notice that increases and decreases in predator population size follow those of the prey.

caused by a predator–prey interaction or by some nonpredation factor affecting hares, such as seasonal food shortage.

There is agreement that, oscillating or not, predator–prey interactions are common and are often strong. For example, following the extinction of many important large forest predators in conjunction with human alteration of the forested landscape, eastern US deer populations expanded to extraordinarily high levels. This suggests that predators once played a large role in suppressing the numbers of deer. A second example is the introduction of the brown tree snake to the island of Guam. In this case, the snake decimated the island's bird community, driving most of its prey to local extinction.

Several factors that are not included in the Volterra model can influence the dynamics of a predator–prey interaction, increasing the likelihood of coexistence and dampening cycles. Environmental spatial heterogeneity in abiotic and biotic factors can benefit either predators or prey. For example, in some portions of a species' habitat, it may be more difficult (or easier) to find or capture prey. It is also possible that the Volterra model does not reflect the true nature of a specific predator–prey interaction. For example, it may be unrealistic to model some prey species as density independent. Researchers have modified the Volterra model in many ways to increase its realism, including versions that incorporate density-dependent growth of prey.

Parasitism and Herbivory

Several other types of interactions have effects similar to those of predation, where one species benefits to the detriment of another. Herbivory, the consumption of plants by animals, is a subtype of predation but differs in some important ways. First,

most herbivores consume only a part of their prey, and consumption infrequently results in death. Second, plants cannot escape their predators as many animal prey do, changing the nature of the predator–prey interaction. Just as predators can greatly affect prey, herbivores play an important role in the dynamics of their host plants. For example, the introduction of the gypsy moth to the eastern US dramatically reduced the population size of many forest tree species, especially oaks and aspen. Also, the effects of herbivorous agricultural pests on crops are the impetus for widespread application of pesticide.

An interaction known as parasitism is also distinct from simple predation. Parasites are usually smaller than their prey (also called a host) and often feed on their hosts without killing them. When parasitic interactions are lethal, parasites kill their host after using or feeding on it for a period of time. Many human diseases such as malaria (a protozoan), intestinal round worms, and disease-causing trematodes such as liver flukes are parasitic. Viruses can also be considered a type of parasite. One interesting form of parasitism is a phenomenon known as brood parasitism in which one species of bird lays its eggs in the nest of another, leaving the offspring to be raised by a surrogate parent. In this case, the parasite does not consume the host but steals some of its parental care. Brood parasites, like all parasites, reduce the success and reproductive output of their hosts. An often fatal form of parasitism in many insect populations is infection by parasitic wasps or flies called parasitoids. These wasps lay eggs in the body cavity of their host; the parasitic larvae then feed on the host during development, killing it on emergence. Some types of parasitic interactions may have dynamics similar to those predicted by Volterra predator–prey theory. Recent evidence shows, for example, that cyclic red grouse populations in Britain may be driven by the effects of a nematode parasite on host reproductive ability.

Mutualism and Symbiosis

When both participants in an interaction benefit from their association, the interaction is known as mutualism (**Table 1**). Many types of mutualisms can be found in nature, and increasing evidence indicates that positive interactions may be more important than many ecologists previously considered in determining who can coexist in a community. There is relatively little theory about how mutualism promotes coexistence or how mutualistic species impact the population dynamics of one another, but a growing body of experimental evidence suggests that mutualism is common and important in ecological communities and, therefore, should be incorporated into our classic theories of community ecology (Bruno *et al.*, 2003). For example, many of the common interactions among organisms, such as animal pollination of plants and some animal dispersal of seeds, are mutualisms. Fruit-eating birds, for example, get nutrition from a plant in exchange for dispersing the plant's seeds across a wide area.

Mutualistic interactions can be facultative or obligatory. Facultative interactions are those in which neither species requires the other to persist but each benefits from the interaction when the opportunity arises. The association between a fruit-bearing tree and its avian visitors is facultative. Many species of lycaenid butterflies have mutually beneficial and facultative associations with ants. The ants are thought to protect the larvae from predators whereas lycaenid larvae offer the ants a nutritious secretion in return.

Obligatory interactions are those in which each species requires the other for survival. These are typically highly specialized species associations. A classic pollination-based obligatory mutualism is that of the fig and fig wasp. Nearly every species of fig is pollinated exclusively by a specialized species of fig wasp, which lives inside the fig fruit. The fig rely on the wasps for pollination and the wasps rely on the fig as a habitat in which to execute their life cycle. Several plant species also have obligate mutualistic interactions with fungi. For example, many species of orchid require a fungal association for seed germination. The fungus, in turn, derives nutrients from the orchid. A third example is an obligatory interaction between some species of algae and fungi. The algae and fungus live together, creating an organism known as lichen.

The lichen example also illustrates a critical vocabulary point. Lichen is an example of a symbiotic mutualism. The terms "mutualism" and "symbiosis" are often used interchangeably, but they actually differ subtly. The term symbiosis stresses that two species live close to one another in prolonged association. Although many symbiotic interactions can also be mutualistic, this need not be the case. For example, you and the bacteria inhabiting your stomach which aide in your digestion are symbiotic and mutualistic, but you and the flukes feeding in your liver (if there are any) form a parasitic symbiosis.

Asymmetrical Interactions

One species can harm or help another species without any benefit or detriment in return. Commensalism refers to the benefit of one species, species A, from the presence of another species, species B, whereas B experiences no effect from the presence of A (**Table 1**). Conversely, amensalism refers to the detrimental effect of species B on A whereas B experiences no effect of A in return (**Table 1**). Far less research has focused on commensalism and amensalism than on other types of interactions, and the strength of commensalism and amensalism is generally thought to be weak. A famous example of commensalism is an association between cattle egrets and cattle. The egrets eat insects flushed by the cattle. The presence of the egrets, however, has no measurable effect on the cattle. Amensalism often occurs as the incidental damage to one species from the presence or activity of another. For example, in the cattle–egret example, some ground-dwelling insects suffer incidental mortality from the cows that step on them.

Interactions and Evolution

Previously, the direction of interaction effects and differences in interaction strength were discussed. It was also mentioned that interactions can change through time and space. In ecological time, interactions between species affect population size. In the longer term, however, interactions can lead to evolutionary change in the two participating species, either

reducing or increasing the strength of the association. For example, in competitive and predatory interactions, one species may evolve mechanisms to avoid strong interaction with others in response to the pressures of natural selection. In mutualistic interactions, on the contrary, specialized behaviors and morphology can be selected through time to promote efficiency, and species can grow to be more dependent on one another.

Evolution typically results from a pressure known as natural selection. Natural selection acts in the following way: Among a collection of organisms in a population, those individuals that are best suited for their environment will be the most successful and will produce a greater fraction of a population's offspring than will other individuals. Successful individuals will genetically pass their beneficial traits to their offspring if these traits are heritable. Through many generations of enhanced survival and reproduction of individuals with beneficial qualities, directional change in physiological, morphological, or behavioral traits at the population level can result. Natural selection can act in a variety of ways. For example, natural selection may favor traits that confer efficiency in food digestion or nutrient acquisition. Interactions between species can also be a force of natural selection for either or both species involved.

When natural selection results in reciprocal trait evolution in two interacting species, it is called coevolution. In other words, if morphological, physiological, or behavioral traits determined by genes in two interacting species evolve in response to an interspecific interaction, coevolution has taken place. The pressure of natural selection (called selection pressure) in coevolution stems from the positive or negative effects that species have on each other. Depending on the type of interaction driving coevolution (competition, predation, mutualism, etc.), many trait changes are possible.

Evolution and Competition

Remember that competition is thought to limit how similar two coexisting species can be. If, as seen in the Lotka–Volterra model of competition, resource overlap can drive one of two competitors to extinction, then competition can also be a strong selection pressure. Individuals with traits that enable them to compete less with other species can have an advantage compared to other members of their population. Over many generations of selection for minimizing overlap, populations can evolve away from interaction with their competitors.

The evolutionary change in a trait in response to competition is called character displacement. Character displacement results from selection for traits that reduce the overlap in resource use with a similar species. If two or more competing species evolve in response to one another, character displacement can be considered a type of coevolution. For example, in the Galapagos Islands, a group of coexisting finches have distinctly different body and beak sizes as well as beak shapes. Size specialization allows each finch to exploit a different food resource. The seemingly well-organized categories of finches suggest that there was once a strong influence of competition on these species and that competition drove the finches to evolve different strategies and diverge. In fact, the type of pattern seen among finches on the Galapagos is widespread. Many species of similar type in a single habitat (e.g., desert rodents or island lizards) appear to be divided into distinct body size categories. This division suggests that selective pressure to reduce resource overlap was once (and may still be) an important influence.

It should be noted that the pattern of body size classes often seen in nature need not necessarily evoke character displacement as an explanation. An alternative hypothesis is the idea of selective retention of invaders. This hypothesis states that as a community is being filled by new arrivals to the system, those invading species that are able to persist are ones whose niche is currently vacant. In other words, invaders who find their niche already occupied are not able to persist in the system and go extinct shortly following invasion. This hypothesis is distinct from the character displacement hypothesis in which no evolutionary adaptation is necessary to explain the morphological differences seen among similar species.

Evolution and Mutualism

Change in morphologies in response to competitive pressure, as in the Galapagos finches, is one type of interaction-driven evolution. A second type can occur in mutualistic interactions. In mutualism, natural selection may favor those individuals that benefit more from an interaction than do other members of the same species. If the partner species in a mutualism is under analogous selective pressure, reciprocal, genetic changes in both species may result. Over many generations, coevolution increases the benefit each species gains from its interaction with the other.

Many plant–pollinator systems appear to have experienced some coevolutionary change. For instance, it may be advantageous for plants to attract pollinators that visit only a few or a single plant species, and many plants attract pollinators with specialized flashy petals and nutritious rewards. Because the next flower visited by such a specialist pollinator is more likely to be a member of the same plant species, a plant that attracts specialists enjoys an increased rate of successful pollination and a decreased expense of excess pollen production. Many pollination associations seem to be coevolved in this way, including hummingbirds with long curved bills feeding on species with long curved flowers and orchid species dispersing fragrances that attract specialized bee visitors.

Other coevolved mutualisms between animals and plants are also known. For instance, some acacia trees have mutualistic associations with ants. An ant colony resides in the tree and uses tree secretions called nectaries for food. The acacia uses the ants for protection from other insects and plants. Ants can attack herbivores that land on the plant, and in some cases the ants even remove competing plant material from around the acacia's base. Studies have shown that mutualistic acacias perform very poorly without ant inhabitants. Similarly, several species of ants are found only on acacia, indicating that for some, acacia is an obligatory ant resource.

In both the acacia–ant and plant–pollinator examples, it is difficult to imagine how such specialized associations could evolve without multiple steps of reciprocal morphological and

behavioral change. It is thought that specialized interactions evolved from generalized interactions, where survival and reproduction benefits of increased mutualism brought two select species closer together. Such increased association ended with, in some cases, obligatory dependence.

Evolution and Predation

Predators can place selection pressure on their prey to avoid capture, and the benefits of efficient catch and consumption of prey can drive the evolution of predators. The evolution of a trait in prey that makes predation more difficult can be followed by the reciprocal evolution of a trait in predators that reduces the prey's new advantage. Conversely, traits that confer advantages to predators can lead to increased natural selection on prey to evolve predation avoidance traits. Again, if two or more predators and prey evolve in response to one another, this is called coevolution.

Through evolutionary time, many species of prey have evolved defenses against predation, including the ability to flee, avoid detection, or being difficult to eat. Predators have also evolved techniques and physical characters to improve their efficiency or probability of successful capture of prey. Predators can become less conspicuous, faster, or more skilled at capturing and killing prey. This phenomenon was first discussed by Darwin in the late 1800s; he wrote of a constant struggle between prey to escape their enemies and predators to consume their prey. Today, we refer to the constant struggle as an evolutionary arms race, with the analogy to nations that reciprocally accumulate weapons to maintain an even balance of defense.

One frequently cited example of a coevolutionary arms race exists among plants and their herbivores. Plants must defend themselves against their predators since they cannot flee, and they often do so with chemicals called secondary compounds. These compounds either have a negative impact on the physiology of herbivores or decrease the ability of herbivores to consume the plant. Studies have shown that many plants increase the tissue concentration of secondary compounds in response to herbivory, suggesting that such compounds are indeed defensive. In addition, the presence of secondary compounds has been shown to effectively reduce the number of herbivores feeding on a plant. Some herbivores, however, have overcome the negative influence of secondary compounds by adapting ways to metabolize or resist the effects of the plant's defenses. Some herbivore species even use defensive compounds as a cue for identifying the plant as a food source. The pressure to evolve chemicals to escape herbivory, as well as the pressure to overcome antiherbivory compounds in order to sequester food, may have driven the degree of specialization seen in some herbivore–plant relationships.

A second example of a predation-like coevolutionary process can occur between brood parasites and their hosts. Recall that brood parasites lay their eggs in the nests of other species; parasite offspring hatch and are raised by the surrogate parent. Brood parasites place a selective pressure on their host to identify parasitic eggs and either remove them or abandon the nest. However, as hosts become better at recognizing parasitic eggs, parasites are also selected to produce eggs that appear more similar to the host species. Parasites can become specialized on an individual host, producing eggs resembling only a single host species.

Multispecies Interactions

Limiting the treatment of interactions to two species helps us to classify associations and consider the effect an interaction might have on the evolution of participating species. In actuality, however, all organisms live in communities composed of more than two species. This reality is not reflected in the simple theories of competition and predation. Unfortunately, the addition of multiple species to simple models can quickly make them complex and difficult to manipulate. Despite these difficulties, a large body of empirical and theoretical work considers multispecies interactions and the role of multispecies interactions in driving evolutionary change.

Food Webs and Trophic Levels

One way to characterize interactions in a community is to document who eats whom. A food chain is a series of prey species and their predators (**Figure 4**). Food chains begin with plants (known as primary producers), followed by the animals that consume plants (primary consumers or herbivores), the animals that eat herbivores (secondary consumers), and so on. The collections of all food chains in a community comprise a food web. One can think of a food web as a diagram that describes how all species or groups of species in a community

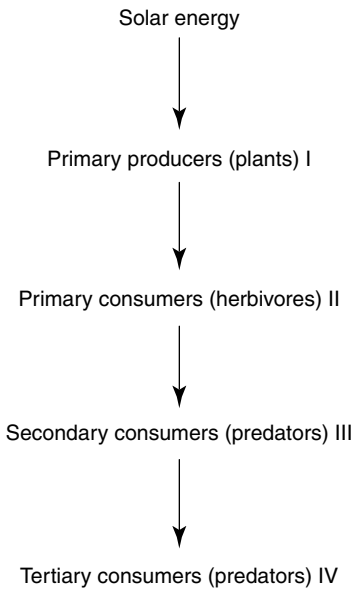


Figure 4 General diagram of a food chain. All food resources originate from the energy of the sun. Primary producers use this energy to create plant biomass. Herbivores consume plants to create herbivore biomass. Carnivores, in turn, consume herbivores and other predators. Roman numerals represent different trophic levels; multiple species can exist within a single trophic level.

use one another for food (**Figure 5**). The positions (producer, secondary consumer, etc.) within a food chain or food web are known as trophic levels. Food webs and trophic levels are useful paradigms for thinking about species interactions; not only do they show the relationships among predators and prey but also suggest when species might compete or interact in other ways. Because organisms live in a world of multispecies interactions, food webs enable us to count interactions and consider the relative importance of multiple interactions on a single species.

Food webs can be very dynamic, even varying seasonally, but others are highly stable. Ecologists measure the robustness or resistance of a food web to change, and we call this robustness, stability. Specifically, stability refers to how a food web will respond to the removal, addition, or a large change in the population size of one or more community members. If the extinction of one community member is followed by the extinction of several other members, the web is unstable. If the web adjusts only slightly in response to extinction or large population shifts, it is stable. The issue of stability in food webs is important because the environment is very dynamic, and over long periods of time events such as large changes in resources or natural disasters are likely to occur. If a community persists through time, it must be resistant to such extreme events.

In 1972, May was the first to propose that webs with many members and many links were unstable. Using a mathematical model, May showed that the more interactions in a community, the greater is the potential for a small perturbation to affect multiple web members and the structure of the web. This result, however, seems counterintuitive because many of Earth's communities are diverse, with many species and many links. Also using mathematical theory, other researchers have shown that in certain cases diversity can lead to stability. For example, food web stability may be related to the asymmetry or symmetry of its interactions or to the location or arrangement of strong interactions in the web.

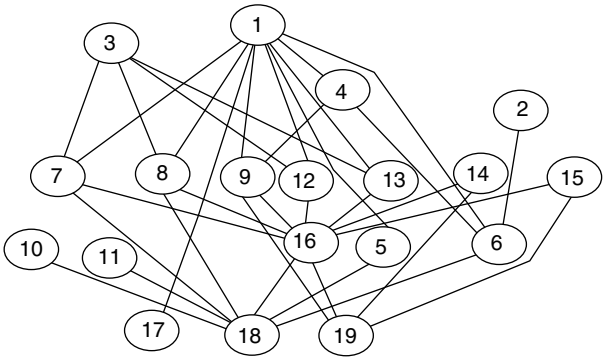


Figure 5 A food web of inhabitants in a pitcher plant in West Malaysia. Pitcher plants have modified leaves that hold water and small communities of insects and microorganisms. All numbers except 16 (bacteria and protozoans), 17 (live insects that have fallen in the pitcher), 18 (drowned pitcher residents), and 19 (other organic debris) represent an insect species. Predators are higher in the web than are prey. For example, Nos. 17–19 are at the bottom of the food web, and they are preyed on by those species that are connected by a line and are higher in the diagram. Top predators are among the highest in the figsures (e.g., 1, 3, and 4), and they prey on all species to which they are connected. Species that prey on some species but are also prey themselves can be found in the middle of the diagram (e.g., 8, 9, 12, and 13). Reproduced with permission from Pimm SL, Lawton JH and Cohen JE (1991) Food web patterns and their consequences. *Nature* 350: 669–674.

Keystone Species

A species that exerts a strong influence on the dynamics of several members of its community is called a keystone species. A keystone species' influence is one that is larger than we might expect based on its abundance alone. In other words, a keystone species must exert its influence based on the important interactions it has in the community and not simply on its dominance or high population size.

For example, the sea star, *Pisaster ochraceus*, is a classic keystone species in marine habitats of Oregon and Washington. It preys on several mussel species, including the mussel *Mytilus californianus*. Mussel species compete strongly for places to adhere on intertidal rocks. *M. californianus* is highly successful at dominating intertidal spatial resources, blocking out space needed by other mussel species. By feeding on *M. californianus*, *P. ochraceus* opens up rock space, resulting in an increase in the diversity of the community. In the absence of *P. ochraceus*, several species of mussel would be driven to extinction by the dominant competitor *M. californianus*, but in the presence of *P. ochraceus* the number of mussel species that can persist in the community is increased. **Table 2** shows some experimental results of removing *P. ochraceus* from the intertidal zone in Washington.

Other predatory keystones include the sea otter and the bass; however, not all keystone species are found at the tops of food chains. For example, nitrogen-fixing bacteria may be keystone members of forest communities if they increase diversity by reducing the effects of nitrogen limitation. Beavers influence species diversity in their communities by altering the habitat structure of streams and ponds. Because keystones maintain the biodiversity of their communities, identifying and preserving them is an important conservation goal. To date, however, there is no general theory that predicts which members of a community will be keystones. The only way to identify a keystone species is to study its community in detail and the community effect of its removal.

Trophic Cascades and Indirect Interactions

Keystone species exert their strong influence on community composition by a combination of direct and indirect effects. Each of the interactions previously discussed (*see* Pairwise Interactions) in the context of pairwise interactions were direct – two species in contact and each influences the dynamics of the other. An indirect interaction occurs when species influence one another through their interactions with other species. Many different types of indirect interactions have been identified, including apparent competition, interaction modifications, and trophic cascades (Vandermeer, 1969; Wooton, 1994).

Apparent competition is the case in which two prey species share a single predator. If one of the two prey species were to increase in abundance, an increase in predator population size

Table 2 Species composition in the presence and absence of *Pisaster ochraceus*

	Suvery date			
	July 1963	*April 1973*	*August 1966*	*June 1971*
Plot type[a]	Control	Control	Removal	Removal
% cover by species[b]				
Balanus cariosus	10	12		
Balanus glandula	~15	14		
Chthamalus tissus	~15	10		
Pollicipes polymerus	1	2	5	
Mytilus californianus	5	2	95	100
Endocladia muricata	2	3		
Corallina vancouvcricnsis	20	28		
Lithothamnium sp.	16	5		
Hedophyllum sessile				
Halichondria panacea	5	5		
Total % covered[c]	89	85	100	100

[a]Plots were of two types: those in which *P. ochraceus* was removed (removal) and those in which *P. ochraceus* remained (control).
[b]All species are sessile invertebrates or algae that live on the rocks of the intertidal zone.
[c]Numbers do not necessarily total 100%; remaining space was either unoccupied or held by a few rarer species.

might follow. An increase in predators can then lead to a decline in both prey species. For example, an increase in the number of one of two grasshopper species has been shown to increase the number of parasitoids infecting both host species. Interaction modification is the case in which one species changes the interaction of two directly interacting species. For example, the amount of phytoplankton (small green plants) in the water column of a lake may affect a predatory fish's ability to find and capture a prey species.

A trophic cascade refers to the effects on several species in a food chain in response to a change in the population size of a single chain member. Trophic cascades were first discussed by Hairston *et al.* (1960), and the effect of cascades on population sizes at each level in a food chain was thereafter named the Hairston-Smith-Slobodkin (HSS) hypothesis. HSS predicts that predators indirectly control the abundance of primary producers by reducing the numbers of herbivores through predation and releasing producers from some herbivory pressure. Predation on herbivores keeps herbivore abundance low, and a smaller number of herbivores consume fewer producers than they would in the absence of predators. Empirical data support the HSS claim that predation can limit the consumption of producers in some systems. For example, as mentioned previously, after humans caused the local extinction of large predators, some herbivore densities (such as that of eastern deer) greatly increased. The increase in herbivores has in turn increased the amount of herbivory damage to forest plant species and has changed forest plant dynamics.

Trophic cascades can also be expanded to four trophic levels. For example, in lakes with piscivorous (fish-eating) fish, there is a decrease in the number of planktivores (plankton-eating fish), an increase in the number of herbivores, and a decrease in the amount of green phytoplankton in comparison to lakes without piscivores. This alternating suppression and release of predatory pressure might explain why some lakes are green and others are clear; water clarity (i.e., the amount of phytoplankton) may be determined, in part, by the number of trophic levels in the lake food web.

Metacommunities

Many of the historic theories and examples in community ecology, including food web theory and theories about the relationship between community diversity and stability, consider a single community in a single location. In recognition that communities are often highly dynamic and interconnected across geographic space, a framework has recently emerged to describe interactions among communities themselves. These metacommunites are collections of communities that are linked by dispersal of multiple, potentially interacting species (Leibold *et al.*, 2004). The metacommunity concept has arisen, in part, to enhance predictions about species interactions at larger spatial scales.

Multispecies and Evolution

Just as direct and pairwise interactions can influence evolution, species assemblages may influence selection pressures as well. Multiple and indirect effects, however, can select traits in complicated ways (Miller and Travis, 1996). In multispecies assemblages, many interactions can simultaneously place evolutionary pressure on a species. Where these pressures counterbalance, no change in morphology or behavior may result. When the pressures coincide, evolution by natural selection may occur. In general, a trait will be selected if an organism benefits from possessing the trait over the sum total of all its interactions. There are many examples of evolution in the balance of multiple pressures. For example, a mutualistic interaction between a lycaenid butterfly and its associated ants may have arisen in response to the pressures of parasitic infection of larvae by parasitoids (Pierce and Mead, 1981).

Human Impacts on Interactions

Many important services and goods used by humans result from species interactions. Services include pollination of

plants and the maintenance of soil fertility, and the products of these services provide goods such as agricultural crops, timber, and clean water. Increasing evidence suggests that human influences on the environment are impacting the very interactions that create these useful goods and services. Influences such as climate change, alteration of the earth's biogeochemical cycles, habitat loss, overharvesting, and pollution can cause disassociations among once-interacting species or change the nature of a service-providing association.

Human stresses on the environment can drive local extinction, and via species interactions single extinctions can impact multiple community members. Stresses such as climate change are changing the abiotic environment in a systematic way, forcing local extinctions in some areas and population expansion in other areas. The responses of individual species to these changes are unique, shifts in species ranges can result, and the overlap of species can break down and new ones can form.

For example, historical data show that vegetation responds very slowly to regional temperature change, and rates of plant species migration will be slow in response to global warming. Some insects, however, have the potential for long-range dispersal and will respond to climatic changes more quickly than plants will. If insects shift faster than plants, some plant populations may be left without pollinators. Furthermore, even if interacting species, such as plants and insects, were able to shift at the same pace, other stresses such as habitat fragmentation stand in their way. If one species is able to shift over a fragmented landscape and another species is not, disassociation can result.

Humans can alter the timing of interactions as well as the ability of species to persist in the same geographic location. For example, evidence suggests that human influences on biogeochemical cycles, including increases in anthropogenic nitrogen deposition, can impact the growth rate of plants, affecting their flowering and development timing (i.e., phenology). Changes in plant phenology may affect the temporal overlap between herbivores and their food plants, possibly leading to herbivore population decline.

Lastly, human domination of the earth has greatly increased the distribution of a few, particular species. Many organisms travel with people as they disperse throughout the world, and some of these species have established themselves in locations other than where they originally evolved. These species are often dominant competitors; they spread quickly and displace other organisms. Because they drive native species extinct or harm native populations, such introductions have significantly altered native interactions in many parts of the world. For example, the introduction of European grasses to California has excluded native plants from all but very small relic patches throughout the state. The loss of native plants has in turn affected the viability of many herbivores that are specialized to forage on native California plants.

An alarming number of Earth's species are going extinct at the hand of human civilization. Identifying and conserving keystone species, as well as predicting and preventing extinction cascades, will be critical to successful mitigation and evasion of further biodiversity loss. To preserve the interactions and communities we need, community ecologists will be asked to provide predictive and restorative information on a great diversity of community and interaction types. Ecologists also will be asked to increase the generality of their findings because species are going extinct more quickly than we can identify taxa and describe communities.

Summary

Species interactions are classified by the direction of their effects and are divided into the direct pairwise categories of competition, predator–prey, mutualism, commensalism, and amensalism (**Table 1**). The strength of an interaction or the amount an interaction affects the population size of its participants, can be determined by experimentally removing one species and observing the population response of the second species and vice versa.

A simple mathematical theory for competition predicts that the amount of resource overlap between two coexisting, competing species must be small (**Figure 2**). A simple mathematical theory for predation predicts that for coexisting predators and prey, population sizes will oscillate through time (**Figure 3**). In both cases, theory may oversimplify the opportunities for species coexistence. Factors such as variation in resources, space, and time, as well as the complexity of a habitat, can minimize resource overlap among competing species or dampen predator–prey oscillations. Mutualistic or positive interactions do not have a historic theory for predicting when and where they will occur, but mutualisms are common in nature.

Pairwise interactions between species can select for changes in morphological, physiological, and behavioral traits through evolutionary time. Coevolution occurs when there is reciprocal genetic change between two interacting species. Coevolution in competitive systems can lead to character displacement. Coevolution in predator–prey systems can lead to an arms race of capture and escape. Also, coevolution in mutualistic systems can lead to highly specialized associations between mutually benefiting species.

Communities are composed of many species, and indirect as well as direct interactions control community composition (**Figure 5**). Indirect interactions include trophic cascades, apparent competition, and the modification of a direct interaction. A keystone is a species that strongly influences the biodiversity of its community via indirect and direct effects (**Table 2**). Keystones are often, but not always, found at the top of a trophic chain or food web. Both indirect and direct interactions can influence the evolution of community members.

Humans are causing widespread extinction of species, and the loss of species is leading to the decline of many goods and services provided by species interactions in natural systems. Human stresses on the environment, such as climate change, alteration of geochemical cycles, and destruction of habitat, can break down species associations because individual species, not communities, respond to changes in the environment.

See also: Coevolution. Competition, Interspecific. Nest Parasitism. Parasitism. Predators, Ecological Role of. Trophic Levels

References

Bruno JF, Stachowicz JJ, and Bertness MD (2003) Inclusion of facilitation in ecological theory. *Trends in Ecology and Evolution* 18: 119–125.

Connell JH (1983) On the prevalence and relative importance of interspecific competition: Evidence from field experiments. *American Naturalist* 122: 661–696.

Gause CF (1934) *The Struggle for Existence*. New York: Hafner.

Gurevitch J, Morrow LL, Wallace A, and Walsh JS (1992) A meta-analysis of competition in field experiments. *American Naturalist* 140: 539–572.

Hairston NG, Smith FE, and Slobodkin LB (1960) Community structure, population control, and competition. *American Naturalist* 94: 421–425.

Leibold MA, Holyoak M, Mouquet N, *et al.* (2004) The metacommunity concepts: A framework for multi-scale community ecology. *Ecology Letters* 7: 601–613.

May RM (1972) Will a large complex system be stable? *Nature* 238: 413–414.

Miller TE and Travis J (1996) The evolutionary role of indirect effects in communities. *Ecology* 77: 1329–1335.

Paine RT (1974) Intertidal community structure: Experimental studies on the relationship between a dominant competitor and its principal predator. *Oecologia* 15: 93–120.

Pierce NE and Mead PS (1981) Parasitoids as selective agents in the symbiosis between lycaenid butterfly larvae and ants. *Science* 211: 1185–1187.

Pimm SL, Lawton JH, and Cohen JE (1991) Food web patterns and their consequences. *Nature* 350: 669–674.

Power T and Richerson PJ (1985) Temporal variation, spatial heterogeneity, and competition for resources in plankton systems: A theoretical model. *American Naturalist* 125: 431–464.

Schoener TW (1983) Field experiments on interspecific competition. *American Naturalist* 122: 240–285.

Tilman D (1994) Competition and biodiversity in spatially structured habitats. *Ecology* 75: 2–16.

Vandermeer JH (1969) The competitive structure of communities: An experimental approach with protozoa. *Ecology* 50: 362–371.

Wooton TJ (1994) The nature and consequences of indirect effects in ecological communities. *Annual Review of Ecology and Systematics* 25: 443–466.

SECOND EDITION

ENCYCLOPEDIA OF BIODIVERSITY



EDITED BY
SIMON LEVIN

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

ENCYCLOPEDIA OF
BIODIVERSITY

SECOND EDITION

EDITOR-IN-CHIEF
Simon A Levin
Princeton University, Princeton, NJ, USA

VOLUME 7



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

First edition 2001

Copyright © 2013 Elsevier Inc. All rights reserved

The following articles are US government works in the public domain and are not subject to copyright:
Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Agriculture, Nutrient Management, and Water Quality
Biodiversity and Human Health
Biodiversity in Plant Breeding
Biodiversity of Harmful Marine Algae
Biodiversity State and Trends in Southeast Asia
Biogeography, Overview
Carbon Cycle
Effects of Ecotones on Biodiversity
Endangered Marine Invertebrates
Energy Use, Human
Environmental Ethics
Ethical Issues in Biodiversity Protection
Extinction, Causes of
Fishes, Biodiversity of
Land Use Issues
Landscape Ecology and Population Dynamics
Nitrogen Deposition and Terrestrial Biodiversity

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein, Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress.

ISBN (print): 978-0-12-384719-5

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Italy

12 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Kristi Gomez and Kate Miklaszewska-Gorczyca
Production: Karen East and Kirsty Halterman

DEDICATION

The Second Edition of the Encyclopedia of Biodiversity is dedicated to the memory of two of our contributors – our colleagues, friends, and mentors, Stephen Schneider and Elinor Ostrom. Both Steve and Lin were great scientists and great global citizens, who changed the world of science, and did all they could to change the world we live in, for the better. They left gaps that can't be filled.

EDITOR-IN-CHIEF BIOGRAPHY



Simon Levin's research interests have been in complexity, and in understanding how macroscopic patterns and processes are maintained at the level of ecosystems and the biosphere in terms of ecological, behavioral, and evolutionary mechanisms that operate primarily at the level of the organisms. In recent years, Levin has turned his attention to the parallels between ecological, financial, and economic systems, particularly with regard to what makes them vulnerable to collapse, and to the evolution and development of structure and organization. Of particular interest to Levin are discounting, intergenerational and intragenerational equity, and cooperation and social norms. He has especially been interested in the management of public goods and common-pool resources. Much of Levin's ecological research is concerned with the evolution of diversification, mechanisms sustaining biological diversity in natural systems, and implications for ecosystem structure and functioning. The work integrates empirical studies and mathematical modeling, with emphasis upon how to extrapolate across scales of space, time, and organizational complexity. The essential mathematical challenge is the development of macroscopic descriptions for the collective behavior of large and heterogeneous ensembles that are subject to continual evolutionary modification. Specific attention is directed to the evolution and ecology of collective behavior, from the movement of flocks of birds and schools of fish to human decision-making. Current ecological systems of study include plant communities, as well as marine open-ocean and intertidal systems. In related work, Levin has been interested in the dynamics of infectious diseases, and in particular in the self-organization of strain structure in influenza A, and in the dynamics of antibiotic resistance. In addition, he has been involved in issues of sustainable development, with emphasis on the linkages between environmental and socioeconomic systems.

SECTION EDITORS LIST

Charles Perrings
EcoSERVICES Group
School of Life Sciences
Arizona State University
PO Box 874501
Tempe, AZ 85287
USA

David M Karl
Professor of Oceanography
University of Hawaii
Honolulu, HI 96822
USA

G David Tilman
Regents Professor and McKnight Presidential Chair in Ecology
Department of Ecology, Evolution and Behavior
University of Minnesota
St. Paul, Minnesota 55108
USA

Gretchen C Daily
Department of Biology
Stanford University
377 Serra Mall
Stanford, CA 94305-5020
USA

Georgina M Mace
Division of Biology
Centre for Population Biology
Imperial College London
Silwood Park, Ascot
Berks, SL5 7PY
UK

Monica G Turner
Eugene P. Odum Professor of Ecology
Department of Zoology, Birge Hall
University of Wisconsin
Madison, WI 53706
USA

Robert K Colwell
University of Connecticut
Department of Ecology & Evolutionary Biology
75 N. Eagleville Road, Unit 3043
Storrs, CT 06269-3043
USA

Sandra Diaz
IMBIV
Casilla de Correo 495
Velez Sarsfield 299
5000 Cordoba
Argentina

Peter Kareiva
Chief Scientist & Director Science
The Nature Conservancy
4722 Latona Ave NE
Seattle, WA 98105
USA

CONTENTS

Subject Classification xv

Contributors to Volume 7 xxi

Foreword to the First Edition xxv

Preface to the First Edition xxvii

Preface to the Second Edition xxix

How to use of the Encyclopedia xxxi

VOLUME 7

Stability, Concept of 1
Clarence L Lehman

Stewardship, Concept of 12
Peter Alpert

Storage, Ecology of 23
Caroline M Pond

Stress, Environmental 39
John Cairns Jr.

Subspecies, Semispecies, Superspecies 45
James Mallet

Subterranean Ecosystems 49
David C Culver and Tanja Pipan

Succession, Phenomenon of 63
HH Shugart

Sustainability and Biodiversity 71
Jeannine Cavender-Bares, James Heffernan, Elizabeth King, Stephen Polasky, Patricia Balvanera, and William C Clark

Systematics, Overview 85
Quentin D Wheeler

T

Taxonomy, Methods of 97
Richard Irwin Vane-Wright

Temperate Forests 112
John A Silander Jr.

Terrestrial Ecosystems 128
Stephen Roxburgh and Ian Noble

The Global Environment Facility: Financing the Stewardship of Global Biodiversity 136
Mark Zimsky, Gustavo Fonseca, Jaime Cavelier, Dirk Gaul, Jean-Marc Sinnassamy, Yoko Watanabe, and Ming Yang

The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity 144
Jeffrey J Opperman, Gerald E Galloway, and Stephanie Duvail

Thermophiles, Origin of 161
Anna-Louise Reysenbach and Margaret L Rising

The Value of Biodiversity 167
Partha Dasgupta, Ann P Kinzig, and Charles Perrings

Threatened Birds 180
S Jacob Socolar and David S Wilcove

Threatened Species: Classification Systems and Their Applications 191
Rebecca M Miller

Timber Industry 212
S Kellomäki, J Kouki, P Niemelä, and H Peltola

Tourism, Role of 222
Ralf Buckley

Traditional Conservation Practices 226
Carl Folke and Johan Colding

Translocation as a Conservation Strategy 236
Jessica J Hellmann

Trends in Nature Recreation: Causes and Consequences 241
Patricia Zaradic and Oliver RW Pergams

Trophic Cascades 258
Michael L Pace

Trophic Levels 264
Peter Yodzis

Tropical Forest Ecosystems 269
Gary S Hartshorn

Tropical Forest Regeneration 277
Robin L Chazdon

True Bugs and Their Relatives, Diversity of 287
Carl W Schaefer

U

Ultraviolet Radiation 296
Andrew R Blaustein and Catherine Searle

Urban–Suburban Biodiversity 304
Elizabeth M Cook, Rebecca L Hale, Ann P Kinzig, and J Morgan Grove

V

Valuing Ecosystem Services 314
Marc N Conte

Vents 327
Cindy Lee Van Dover

Vertebrates, Overview 333
Carl Gans and Christopher J Bell

Vicariance Biogeography 342
Christopher John Humphries

W

Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy 352
Rebecca L Goldman-Benner, Silvia Benitez, Alejandro Calvache, Aurelio Ramos, and Fernando Veiga

Wetland Creation and Restoration 367
William J Mitsch

Wetlands Ecosystems 384
Barbara L Bedford, Donald J Leopold, and James P Gibbs

Wildlife Management 403
David Saltz and Gary C White

Worms, Annelida 408
Kristian Fauchald

Worms, Nematoda 420
Scott Lyell Gardner

Worms, Platyhelminthes 437
Janine N Caira and D Timothy J Littlewood

Z

Zoos and Zoological Parks 470
Anna Marie Lyles and Dan Wharton

Glossary 481

Index 547

SUBJECT CLASSIFICATION

Agriculture

Agricultural Invasions
Agriculture, Industrialized
Agriculture, Nutrient Management, and Water Quality
Agriculture, Sustainable
Agriculture, Traditional
Agrobiodiversity
Aquaculture
Biodiversity-Friendly Farming
Breeding of Animals
Cattle, Sheep, and Goats, Ecological Role of
Crop Mixtures and the Mechanisms of Overyielding
Domestication of Crop Plants
Ecology of Agriculture
Feeding the World and Protecting Biodiversity
Herbicides
Hybridization in Plants
Indigenous Strategies Used to Domesticate Plants in Brazilian Amazon
Land-Use Changes and CO_2 Emissions Due to US Corn Ethanol Production
Range Ecology, Global Livestock Influences
Resource Exploitation, Fisheries
Slash-and-Burn Agriculture, Effects of
Timber Industry

Conservation and Restoration

Captive Breeding and Reintroduction
Captive Breeding and the Evolutionarily Significant Unit
Conservation Biology, Discipline of
Conservation Efforts, Contemporary
Conservation Movement, Historical
Diseases, Conservation and
Habitat Restoration for Biofuels, Biodiversity and Wildlife
Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Implications of Urbanization for Conservation and Biodiversity Protection
In Situ, *Ex Situ* Conservation
Landscape Corridors
Marine Conservation in a Changing Climate
Marine Protected Areas: Static Boundaries in a Changing World
Nature-Based Coastal Defenses: Can Biodiversity Help?
Oil-Palm Plantations in the Context of Biodiversity Conservation
Primate Populations, Conservation of
Priority Setting for Biodiversity and Ecosystem Services
Reforestation
Restoration of Animal, Plant, and Microbial Diversity
Restoration of Biodiversity, Overview
Role and Trends of Protected Areas in Conservation
Role of Botanic Gardens
Soil Conservation
Sustainability and Biodiversity
The Global Environment Facility: Financing the Stewardship of Global Biodiversity
The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity
Traditional Conservation Practices
Translocation as a Conservation Strategy
Tropical Forest Regeneration
Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy
Wetland Creation and Restoration

Economics of Biodiversity

Biodiversity and Ecosystem Services
Biodiversity and Pest Control Services
Biodiversity as a Commodity
Biodiversity, Human Well-Being, and Markets
Biofuels and Biodiversity, Wildlife Habitat Restoration
Conservation and the World's Poorest of the Poor
Conserving Biodiversity Outside Protected Areas
Development by Design: Using a Revisionist History to Guide a Sustainable Future
Economic Control of Invasive Species
Economic Value of Biodiversity, Measurements of
Economics of Agrobiodiversity
Economics of the Regulating Services
Evaluation of Ecosystem Service Policies From Biophysical and Social Perspectives: the Case of China

Identifying Conservation Priorities Using a Return on Investment Analysis
Impact of Ecological Restoration on Ecosystem Services
Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Land-Use Issues
Market Economy and Biodiversity
Modeling Marine Ecosystem Services
Modeling Terrestrial Ecosystem Services
Pharmacology, Biodiversity and
Poverty and Biodiversity
Property Rights and Biodiversity
The Value of Biodiversity
Tourism, Role of
Valuing Ecosystem Services
Water Funds: A New Ecosystem Service and Biodiversity Conservation

Environmental Conditions and Effects

Acid and Mercury Deposition Effects on Forest and Freshwater Aquatic Ecosystems
Air Pollution
Atmospheric Gases
Carbon Cycle
Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species
Climate Change and Ecology, Synergism of
Climate Change and Extinctions
Climate Change and Wild Species
Climate, Effects of
Desertification
Disturbance, Mechanisms of
Economic Growth and the Environment
Ecotoxicology
El Niño and Biodiversity
Energy Flow and Ecosystems
Environmental Ethics
Eutrophication and Oligotrophication
Evolution in Response to Climate Change
Fires, Ecological Effects of
Grazing, Effects of
Greenhouse Effect
Impact of Past Global Warming on Biodiversity
Implications of Urbanization for Conservation and Biodiversity Protection
Insecticide Resistance
Marine and Aquatic Communities, Stress from Eutrophication
Marine Conservation in a Changing Climate
Nitrogen, Nitrogen Cycle
Pesticides, Use and Effects of
Plant Phenology Changes and Climate Change
Priority Setting for Biodiversity and Ecosystem Services
Soil Biota, Soil Systems, and Processes
Stress, Environmental
Ultraviolet Radiation

Evolution

Adaptation
Adaptive Radiation
Biodiversity, Evolution and
Biodiversity, Origin of
Complementarity
Darwin, Charles (Darwinism)
Defenses, Ecology of
Dormancy and Diapause
Eukaryotes, Origin of
Evolution, Theory of
Evolution in Response to Climate Change
Fossil Record
Geologic Time, History of Biodiversity in
Life History, Evolution of
Origin and Diversification of Angiosperms
Phylogeny
Plant Communities, Evolution of
Residents and Transients in the Fossil Record
Storage, Ecology of

Extinctions

Comparing Extinction Rates: Past, Present, and Future
Dinosaurs, Extinction Theories for
Endangered Ecosystems
Endangered Freshwater Invertebrates
Endangered Mammals
Endangered Marine Invertebrates
Endangered Plants
Extinction, Causes of
Extinction in the Fossil Record
Global Declines of Amphibians
Latent Extinctions – the Living Dead
Loss of Biodiversity, Overview
Mammals (Late Quaternary), Extinctions of
Mass Extinctions, Concept of
Mass Extinctions, Notable Examples of
Modern Examples of Extinctions
Museums and Institutions, Role of
Natural Extinctions (Not Human Influenced)
Residents and Transients in the Fossil Record

Threatened Species: Classification Systems and Their Applications

Genetics

Conservation Genetics
Ecological Genetics
Gene Banks
Genes, Description of
Genetic Diversity
Hybridization in Plants
Inbreeding and Outbreeding
Landscape Genetics
Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms
Phenotype, a Historical Perspective
Recombination

Geographic and Global Issues

African and Asian Savannahs
Biodiversity State and Trends in Southeast Asia
Biodiversity-Rich Countries
Biogeochemical Cycles
Biogeography, Overview
Dispersal Biogeography
Diversity, Community/regional Level
Endemism
Energy Use, Human
Global Species Richness
Hotspots
Island Biogeography
Landscape Corridors
Landscape Ecology and Population Dynamics
Landscape Epidemiology
Landscape Genetics
Landscape Legacies
Landscape Modeling
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Latitudinal Gradients of Biodiversity
Migration
Neotropical Seasonally Dry Forests
Rainforest Loss and Change
Vicariance Biogeography

Habitats and Ecosystems

Africa, Ecosystems of
Alpine Ecosystems
Amazon Ecosystems
Antarctic Ecosystems
Arctic Terrestrial Ecosystems
Asia, Ecosystems of
Australia, Biodiversity of Ecosystems
Biodiversity in Logged and Managed Forests
Boreal Forest Ecosystems
Census of Marine Life
Central America, Ecosystems of
Coastal Beach Ecosystems
Corals and Coral Reefs
Desert Ecosystems
Ecosystem Function, Principles of
Ecosystem, Concept of
Ecosystems of South America
Estuarine Ecosystems
Europe, Ecosystems of
Freshwater Ecosystems
Freshwater Ecosystems, Human Impact on
Grassland Ecosystems
Habitat and Niche, Concept of
High-Elevation Andean Ecosystems
High-Temperature Ecosystems
Intertidal Ecosystems
Lake and Pond Ecosystems
Land-Use Patterns, Historic
Macroscopic Patterns in Marine Plankton
Mangrove Ecosystems
Marine Ecosystems
Marine Ecosystems, Human Impacts on
Marine Sediments
Mediterranean-Climate Ecosystems
Near East Ecosystems, Animal Diversity
Near East Ecosystems, Plant Diversity
North America, Patterns of Biodiversity in
Ocean Ecosystems
Pelagic Ecosystems
Rainforest Ecosystems, Animal Diversity
Rainforest Ecosystems, Plant Diversity
Riparian Landscapes
River Ecosystems
Seagrasses
South American Natural Ecosystems, Status of
Southern (Austral) Ecosystems
Subterranean Ecosystems
Terrestrial Ecosystems
Tropical Forest Ecosystems
Urban–Suburban Biodiversity
Vents
Wetlands Ecosystems

Human Effects and Interventions

Biodiversity and Human Health
Deforestation and Land Clearing
Ethnobiology and Ethnoecology

Habitat Loss and Fragmentation
Human Impact on Biodiversity, Overview
Hunter-Gatherer Societies, Ecological Impact of
Indigenous Peoples and Biodiversity
Landscape Legacies
Natural Reserves and Preserves
Pollution, Overview
Population Stabilization, Human
Wildlife Management

Invertebrates

Arachnids
Arthropods (Terrestrial), Amazonian
Beetles
Butterflies
Crustaceans
Endangered Terrestrial Invertebrates
Flies, Gnats, and Mosquitoes
Grasshoppers and Their Relatives
Hymenoptera
Insects, Overview
Invertebrates, Freshwater, Overview
Invertebrates, Marine, Overview
Invertebrates, Terrestrial, Overview
Isoptera
Molluscs
Moths
Myriapods
Parasitoids
True Bugs and Their Relatives, Diversity of
Worms, Annelida
Worms, Nematoda
Worms, Platyhelminthes

Microbial Diversity

Archaea, Origin of
Bacterial Biodiversity
Bacterial Genetics
Ecogenomic Sensors
Fungi
Marine Symbioses: Metazoans and Microbes
Marine Viruses
Microbial Biodiversity
Microbial Biodiversity, Measurement of
Microbial Biogeography
Microbial Diversity
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Microorganisms (Microbes), Role of
Origins, Evolution and Breakdown of Bacterial Symbiosis
Plankton, Status and Role of
Proteorhodopsin: Widespread Microbial Light-Driven Proton Pumps
Protozoa
Psychrophiles
Rare Biosphere
Thermophiles, Origin of

Plant Biodiversity

African and Asian Savannahs
Biodiversity of the Succulent Karoo
Biodiversity of Harmful Marine Algae
Biodiversity in Plant Breeding
C_4 Plants
Countryside Biogeography
Effects of Ecotones on Biodiversity
Edible Plants
Elevational Trends in Biodiversity
Forest Canopies, Animal Diversity
Forest Canopies, Plant Diversity
Forest Ecology
Globalization Effects of Common Plant Species
Herbaceous Vegetation, Species Richness in
Introduced Plants, Negative Effects of
Landscape Diversity
Large-Scale Biodiversity Experiments
Latitudinal and Elevational Range Shifts Under Contemporary Climate Change
Neotropical Seasonally Dry Forests
Phenological Shifts in Animals Under Contemporary Climate Change
Photosynthesis, Mechanisms of
Plant Biodiversity, Overview
Plant Conservation
Plant Invasions
Plant Sources of Drugs and Chemicals
Plant-Animal Interactions
Plant-Soil Interactions
Pollinators, Role of
Temperate Forests

Population Issues

Biodiversity and Cultural Ecosystem Services
Biofuels and Biodiversity: the Implications of Energy Sprawl
Conservation and People
Feeding the World and Protecting Biodiversity
Human Impacts on Ecosystems: An Overview

Implications of Urbanization for Conservation and Biodiversity Protection
Introduced Species, Impacts and Distribution of
Metapopulations
Population Density
Population Diversity, Overview
Population Dynamics
Population Genetics
Population Viability Analysis
Predators, Ecological Role of
Succession, Phenomenon of
Trends in Nature Recreation: Causes and Consequences

Public Policies and Attitudes

Aesthetic Factors
Biofuels and Biodiversity, Wildlife Habitat Restoration
Convention on Biological Diversity
Education and Biodiversity
Ethical Issues in Biodiversity Protection
Feeding the World and Protecting Biodiversity
Government Legislation and Regulations in the United States
Historical Awareness of Biodiversity
International Organizations and Biodiversity
Indirect Land Use and Greenhouse Gas Impacts of Biofuels
Justice, Equity and Biodiversity
Life Cycle Analysis of Biofuels
Literary Perspectives on Biodiversity
Museums and Institutions, Role of
Religious Traditions and Biodiversity
Social and Cultural Factors
Trends in Nature Recreation: Causes and Consequences
Zoos and Zoological Parks

Species Interactions and Interrelationships

Coevolution
Competition, Interspecific
Cooperation, Evolution of
Diversity, Ecology, and Biogeochemical Influence of N_2-Fixing Microorganisms in the Sea
Food Webs
Hemiparasitism
Indicator Species
Keystone Species
Landscape Ecology and Population Dynamics
Limits to Biodiversity (Species Packing)
Marine Symbioses: Metazoans and Microbes
Nest Parasitism
Neutral Theory and Beyond
Parasitism
Social Behavior
Species Assemblages, Macroecology, and Global Change
Species Coexistence
Species Interactions
Species-Area Relationships
Trophic Cascades
Trophic Levels

Systematics and Species Concept

Cladistics
Cladogenesis
Differentiation
Diversity, Molecular Level
Diversity, Organism Level
Diversity, Taxonomic Versus Functional
Functional Groups
Nomenclature, Systems of
Speciation, Process of
Speciation, Theories of
Species Diversity, Overview
Species, Concepts of
Subspecies, Semispecies, Superspecies
Systematics, Overview
Taxonomy, Methods of

Techniques and Measurements

Biodiversity Informatics
Biogeographical Modeling
Bioprospecting
Computer Systems and Models, Use of
Defining, Measuring, and Partitioning Species Diversity
Dynamic Global Vegetation Models
Ecogenomic Sensors
Ecosystem Function Measurement, Aquatic and Marine Communities
Ecosystem Function Measurement, Terrestrial Communities
Environmental Impact, Concept and Measurement of
Framework for Assessment and Monitoring of Biodiversity
Functional Diversity Measures
Habitat Selection
Indicator Species: Computation
Landscape Modeling
Measurement and Analysis of Biodiversity

Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity From Sampling Data
Microbial Diversity of the Oceanic Surface Picoplankton: Insights from the Global Ocean Sampling (GOS) Program
Modeling Biodiversity Dynamics in Countryside and Native Habitats
Paleoecology
Remote Sensing and Image Processing
Species Distribution Modeling

Theories and Concepts of Biodiversity

Biodiversity Generation, Overview
Biodiversity, Definition of
Carrying Capacity, Concept of
Commons, Concept and Theory of
Commons, Institutional Diversity of
Complexity Versus Diversity
Disturbance Regimes and the Historical Range of Variation in Terrestrial Ecosystems
Ecological Footprint, Concept of
Ecology, Concept and Theories in
Ecosystem Services
Effects of Ecotones on Biodiversity
El Niño and Biodiversity
Functional Diversity
Guilds
Large-Scale Biodiversity Experiments
Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data
Neutral Theory and Beyond
Oceanic Islands: Models of Diversity
Origin of Life, Theories of
Resource Partitioning
Scale, Concept and Effects of
Species Distribution Modeling
Stability, Concepts of
Stewardship, Concept of

Vertebrates

Amphibians, Biodiversity of
Birds, Biodiversity of
Carnivores
Diversity and Conservation of Neotropical Mammals
Endangered Amphibians
Endangered Reptiles
Fish Conservation
Fish Stocks
Fishes, Biodiversity of
Mammals (Pre-Quaternary), Extinctions of
Mammals, Biodiversity of
Mammals, Conservation Efforts for
Marine Mammals, Extinctions of
Megaherbivores
Museums and Institutions, Role of
Primate Populations, Conservation of
Reptiles, Biodiversity of
Salmon
Threatened Birds
Vertebrates, Overview

CONTRIBUTORS TO VOLUME 7

Peter Alpert
University of Massachusetts, Amherst, MA, USA

Patricia Balvanera
Universidad Nacional Autónoma de México, Morelia, México

Barbara L Bedford
Cornell University, Ithaca, NY, USA

Christopher J Bell
University of Texas at Austin, Austin, TX, USA

Silvia Benitez
The Nature Conservancy, Quito, Ecuador

Andrew R Blaustein
Oregon State University, Corvallis, OR, USA

Ralf Buckley
Griffith University, Gold Coast, QLD, Australia

Janine N Caira
University of Connecticut, Storrs, CT, USA

John Cairns Jr.
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Alejandro Calvache
The Nature Conservancy, Cartagena de Indias, Colombia

Jaime Cavelier
Global Environment Facility, Washington, DC, USA

Jeannine Cavender-Bares
University of Minnesota, St Paul, MN, USA

Robin L Chazdon
University of Connecticut, Storrs, CT, USA

William C Clark
Harvard University, Cambridge, MA, USA

Johan Colding
Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

Marc N Conte
Stanford University, Stanford, CA, USA

Elizabeth M Cook
Arizona State University, Tempe, AZ, USA

David C Culver
American University, Washington, DC, USA

Partha Dasgupta
University of Cambridge, Cambridge and University of Manchester, Manchester, UK

Stephanie Duvail
Institute of Research for Development, Nairobi, Kenya

Kristian Fauchald
National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

Carl Folke
Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

Gustavo Fonseca
Global Environment Facility, Washington, DC, USA

Gerald E Galloway
University of Maryland, College Park, MD, USA

Carl Gans
University of Texas at Austin, Austin, TX, USA

Scott Lyell Gardner
University of Nebraska-Lincoln, Lincoln, NE, USA

Dirk Gaul
Global Environment Facility, Washington, DC, USA

James P Gibbs
State University of New York, Syracuse, NY, USA

Rebecca L Goldman-Benner
The Nature Conservancy, Durham, NC, USA

J Morgan Grove
USDA Forest Service, Baltimore, MD, USA

Rebecca L Hale
Arizona State University, Tempe, AZ, USA

Gary S Hartshorn
World Forestry Center, Portland, OR, USA

James Heffernan
Florida International University, Miami, FL, USA, and Duke University, Durham, NC, USA

Jessica J Hellmann
University of Notre Dame, Notre Dame, IN, USA

Christopher John Humphries
The Natural History Museum, London, UK

S Jacob Socolar
Princeton University, Princeton, NJ, USA

S Kellomäki
University of Eastern Finland, Finland

Elizabeth King
Princeton University, Princeton, NJ, USA, and University of Georgia, Athens, GA, USA

Ann P Kinzig
Arizona State University, Tempe, AZ, USA

J Kouki
University of Eastern Finland, Finland

Clarence L Lehman
University of Minnesota, St Paul, MN, USA

Donald J Leopold
State University of New York, Syracuse, NY, USA

D Timothy J Littlewood
The Natural History Museum, London, UK

Anna Marie Lyles
Bio-Gist Ventures, Princeton, NJ, USA

James Mallet
University College London, London, UK

Rebecca M Miller
IUCN Red List Unit, Cambridge, UK

William J Mitsch
Florida Gulf Coast University, Naples, FL, USA

P Niemelä
University of Eastern Finland, Finland

Ian Noble
Australian National University, Canberra, ACT, Australia

Jeffrey J Opperman
The Nature Conservancy, Chagrin Falls, OH, USA

Michael L Pace
University of Virginia, Charlottesville, VA, USA

H Peltola
University of Eastern Finland, Finland

Oliver RW Pergams
Olive-Harvey College, Chicago, IL, USA

Charles Perrings
Arizona State University, Tempe, AZ, USA

Tanja Pipan
Karst Research Institute at ZRC SAZU, Postojna, Slovenia

Stephen Polasky
University of Minnesota, St Paul, MN, USA

Caroline M Pond
The Open University, Milton Keynes, UK

Aurelio Ramos
The Nature Conservancy, Cartagena de Indias, Colombia

Anna-Louise Reysenbach
Portland State University, Portland, OR, USA

Margaret L Rising
Portland State University, Portland, OR, USA

Stephen Roxburgh
Australian National University, Canberra, ACT, Australia

David Saltz
Ben Gurion University of the Negev, Israel

Carl W Schaefer
University of Connecticut, Storrs, CT, USA

Catherine Searle
Georgia Institute of Technology, Atlanta, GA, USA

HH Shugart
University of Virginia, Charlottesville, VA, USA

John A Silander Jr.
University of Connecticut, Storrs, CT, USA

Jean-Marc Sinnassamy
Global Environment Facility, Washington, DC, USA

Cindy Lee Van Dover
College of William & Mary, Williamsburg, VA, USA

Richard Irwin Vane-Wright
University of Kent, Cantenbury, UK; Natural History Museum, London, UK, and Canterbury Christ Church University, Cantenbury, UK

Fernando Veiga
The Nature Conservancy, Curitiba, Brazil

Yoko Watanabe
Global Environment Facility, Washington, DC, USA

Dan Wharton
The Chicago Zoological Society, Brookfield, IL, USA

Quentin D Wheeler
Arizona State University, Tempe, AZ, USA

Gary C White
Colorado State University, Fort Collins, CO, USA

David S Wilcove
Princeton University, Princeton, NJ, USA

Ming Yang
Global Environment Facility, Washington, DC, USA

Peter Yodzis
University of Guelph, Ontario, Canada

Patricia Zaradic
Red Rock Institute, Bryn Mawr, PA, USA

Mark Zimsky
Global Environment Facility, Washington, DC, USA

FOREWORD TO THE FIRST EDITION

Biodiversity is the totality of the inherited variation of all forms of life across all levels of variation, from ecosystem to species to gene. Soon after the term was introduced at the first National Forum on Biodiversity in 1986, and after it began its rapid spread around the world, there occurred a reconfiguration in the way much of the science of biology is conceived. Where previously comparative biology had been almost entirely focused on the fundamentals of classification, evolution, behavior, and ecology, now it was augmented by a wide range of analyses from the social sciences. Where taxonomy and biogeography had been marginalized through the middle half of the twentieth century, now they moved back toward center stage. And where extinction had been little more than a phenomenon recognized and lamented, now it became a major concern of science. Much of ecology shifted toward the study of biodiversity's role in the assembly and maintenance of ecosystems. A growing number of economists, political scientists, and bioethicists took up the issue as a part of their scholarly agenda. From this mix the discipline of conservation biology was born, and the Society of Conservation Biology became one of the fastest growing organizations in modern science. In 1992, the Rio Summit catapulted biodiversity to global prominence, from which most of the nations of the world endorsed the Convention on Biodiversity and have since used it as a guideline for conservation programs.

The new biodiversity initiative gave organismic and evolutionary biologists a global mission worthy of their science. It confirmed for those who labored in the vineyards that, as medicine is to molecular and cellular biology, the environment is to organismic and evolutionary biology. The first is responsible for personal health, and the second for planetary health. The additional evidence adduced, moreover made clear that we are in the midst of an episode of massive extinction, unprecedented since that closing of the Mesozoic Era – and that scientists must lead the attempt to save the Creation.

The articles in the *Encyclopedia of Biodiversity* are unusually eclectic, yet organized by a set of easily articulated goals. They are the following: to carry the systematics and biogeography of the world fauna and flora toward completion; map the hot spots where conservation will save the most biodiversity; orient studies of natural history to understand and save threatened species; advance ecosystems studies and biogeography to create the needed principles of community assembly and maintenance; acquire the knowledge of resource use, economics, and polity to advance conservation programs based on sustainability; and enrich the ethic of global conservation in terms persuasive to all.

The road ahead, down which we must urgently travel, will be smoothed by the exponential growth of information and a growing public awareness and support. The *Encyclopedia of Biodiversity* will serve as an important knowledge base to guide this supremely important effort.

E.O. Wilson
Harvard University, USA

PREFACE TO THE FIRST EDITION

The science of biodiversity has become the science of our future. Our awareness of the disappearance of biodiversity has brought with it a long-overdue appreciation of the magnitude of our loss, and a determination to develop the tools to protect our future. This encyclopedia, for the first time in its completeness, brings together, study of the dimensions of diversity with examination of the services that biodiversity provides, and measures to protect it.

The entries in the encyclopedia have been arranged alphabetically, but the coverage is designed functionally. At the core is a comprehensive survey of biodiversity, across taxonomic groups and ecological regions. The emergence of biodiversity is then placed in an evolutionary perspective, as background for an understanding of current trends. Particular attention is given to the loss of services – for example, in fisheries, forestry and climate mediation – that are derived from natural systems. These are placed in an economic framework through a comprehensive set of papers that address problems of valuation, costs, and benefits, and develop a framework for prioritizing actions. Finally, a review is given of institutions and other mechanisms that exist and are needed for the preservation of biodiversity and, with it, the services that humans derive from nature.

The background for understanding biodiversity is to be found in the fossil record, and in the evolutionary patterns and trends that it reveals. The encyclopedia hence discusses these patterns, the origins of biodiversity, the effects of geological events, the mechanisms of evolution, and the uniqueness of the evolutionary process, with implications for conservation and restoration. The essential processes in macroevolution are those of speciation and extinction, which together govern the dynamics of diversity at higher levels of organization. These are given extensive coverage, both from mechanistic and from a historical perspectives, and provide an essential context to understand the rest of the contributions.

The classification of organisms into species and higher taxa, and the elucidation of the mechanisms of natural selection, were the essential intellectual advances that allowed the development of the science of biodiversity. Carl Linnaeus introduced a systematic framework for understanding phylogenies, which continues to provide the foundation for evolutionary studies today; and Charles Darwin's great legacy – the theory of evolution by natural selection – is the essential organizing principle for understanding the processes that gave rise to the patterns Linnaeus recognized. The encyclopedia provides unmatched taxonomic coverage of the organization of diversity into taxonomic groups and complements that with an extensive examination of ecosystems by biogeographic region and by functional type. These chapters elucidate latitudinal trends, life zones, species–area relationships, and the distribution of diversity within and among ecological communities.

Throughout the core chapters, there is a healthy balance between empirical facts and conceptual theories. Such theories help to illuminate principles that cross systems and levels of organization, and transform the study of biodiversity into a science. Basic ecological constructs, such as the habitat and the niche, are given extensive treatment, as are key ecological mechanisms such as competition, predation, herbivory, parasitism, and mutualism. These treatments are complemented by exploration of fundamental evolutionary mechanisms related to local differentiation, aspect diversity, sex and recombination, and especially theories of extinction.

With these foundational chapters in hand, one can turn to the contemporary problems in biodiversity and compare today's rapid rates of change to the historical patterns. Key chapters examine agriculture, fisheries, and forests, their importance to human needs, and their status and trends in response to changing land-use patterns, population growth, overexploitation, and climate change. Threatened and endangered species are discussed in detail, with relation to the consequences of the spread of invading species.

The utilization of nature's bounty for food, fiber, and fuel provides some of the most obvious benefits of biodiversity to humanity. Equally important, however, are the things that are less well appreciated: the potential for the discovery of new pharmaceuticals that can improve human health, the role of biodiversity in pollinating crops and wild species, and the importance of natural systems in regulating climate, mediating nutrient fluxes, and sequestering carbon as well as toxic materials. Each of these services provides humans with direct and indirect benefits, and somehow we must find ways to weigh these benefits, along with the ethical and aesthetic values we place on natural systems and biodiversity, to provide priorities for action. Only recently have economists recognized the importance of such issues as intellectual challenges essential to our survival on the planet. Much of biodiversity is exploited by humanity as part of global commons, in which one does not pay in fair measure for extracting parts, or affecting the commons otherwise through land use or pollution. Economists have come to realize, along with ecologists and others, the magnitude of the externalities involved. When such externalities are involved, the market does not function as it must, to maintain the resource, and new measures are needed if the sustainability and resilience of these resources are to be preserved. There is as yet, however, no ecological equivalent to the power of financial institutions, such as the Federal Reserve Board in the United States, to modify individual incentives sufficiently to maintain regional or global stability in the system of interest. In this encyclopedia, some of the most enlightened and thoughtful economists turn their attention to the economic challenges, and discuss the mechanisms and institutions that might be needed.

Together, the state-of-the-art entries in this encyclopedia tell an exciting story of how biodiversity arose, continues to arise, and is maintained. It is a story of a complex, self-organizing system – the biosphere – whose pieces can be examined individually, but cannot be understood outside the context of the whole. It is also a story of the coevolution of the biosphere and *Homo sapiens*, the first species whose own activities can feed back to influence the evolution of the

biosphere on time scales that could lead to its own demise. The articles in the encyclopedia can be used as material for a wide spectrum of courses, tracing the history of the emergence of biodiversity from its origins to the challenges we face today.

This has been a massive effort, but one of the most rewarding I have ever undertaken. So many people have played a role that it is difficult to know where to begin. The project began through the initiative of Scott Bentley at Academic Press, and then was managed flawlessly by Chris Morris at AP. I cannot recall ever having dealt with an editor who operated more professionally than Chris, who combined a true vision and enthusiasm for the project with a sense of economic realities and the energy and insight to make the whole project work. In this he was ably assisted by outstanding Academic Press staff, especially Naomi Henning, Nick Panissidi, and Ann Marie Martin. At every step, it has been a pleasure to work with Academic Press, and I especially single out Chris for his fantastic and scholarly efforts.

At the next stage, the Editorial Boards were terrific in generating and commenting on ideas, suggesting authors, and critiquing contributions. More than 400 authors then adopted our view of the importance of the project, accepted the task of writing, and produced timely and comprehensive articles that make this Encyclopedia like no other source available today. To all of these, I extend my thanks and congratulations.

And finally, special acknowledgment and gratitude are due my wife, Carole, and my assistant, Amy Bordvik. Carole put up with the late nights and obsessiveness that were essential to the process, and Amy worked tirelessly and without complaint through the whole long process. To them, any expression of thanks is insufficient.

S. Levin
Princeton University, Princeton, NJ, USA

References

Levin SA (1999) *Fragile Dominion*. Reading, MA: Perseus Books.
Wilson EO (1992) *The Diversity of Life*. New York: Norton.

PREFACE TO THE SECOND EDITION

Twelve years have passed since the appearance of the First Edition of the Encyclopedia of Biodiversity, and a great deal has transpired. We continue to lose biodiversity at rates and in ways that severely threaten the services that humans derive from ecosystems. Attention to biodiversity and sustainability has led to major scientific advances, though the efforts to implement the necessary changes in policy to limit our losses remain elusive. As Edward O. Wilson pointed out in his Foreword to the First Edition, reprinted in this Edition, the pathway to address the great challenge we face will be more easily traveled as we accumulate more and more information, organized in a way that is accessible to practicing scientists, to students, to the public, and especially to the decision makers in industry and governments who hold the keys to success. With this in mind, it became essential to update the Encyclopedia of Biodiversity to reflect the new information and scientific advances that have arisen in the past dozen years.

The Second Edition retains the same breadth across the basic and applied dimensions as did the First Edition. The great majority of entries from the First Edition have been revised to reflect the changes that have taken place, or in some cases replaced by entirely new contributions. In addition, the scope of the coverage has been expanded substantially, to broader treatment of emergent scientific advances in subjects like the microbial ecology of the oceans, landscape ecology, conservation biology, and ecological economics. In all, in addition to those that have been substantially revised, there are nearly 100 completely new articles to complement the solid foundation that already existed. These include articles that broaden the general scientific foundations of the subject of biodiversity, like the measurement and estimation of species richness, as well as the application of those basic principles to management problems.

New entries expand the coverage of areas that have grown in importance in the past decade; elucidation of the importance of ecosystem services, for example, has provided a way to connect biodiversity to the economic forces that impacts it, and has stimulated major scientific advances in the past decade in terms of how to measure and protect those services. The linkages to economics more generally have provided a framework for addressing the loss of biodiversity by understanding what drives anthropogenic impacts, and how we might manage the Commons we all share. These issues have been covered in much greater depth in this new Edition, along with other issues related to climate change, land use, biofuels, and the growing subject of conservation biology.

I am extremely grateful to the excellent staff that assisted in the preparation of this edition but especially to the associate editors who advised on whom to invite, aided with invitations, shepherded manuscripts through the process, and eventually read final drafts. Katarzyna Miklaszewska, Elsevier's in-house editor, was remarkable at every stage of the process, and it has been a pleasure to work with her; she minimized damage from the unavoidable glitches that developed along the way, and maintained a cheery demeanor while guiding the Encyclopedia to completion. Chris Morris, who was indispensable in the production of the First Edition, rejoined our efforts to update the excellent Glossary he produced the first time through. Also thanks are due to Carole Levin, for her patience throughout the whole process.

The urgency of addressing the loss of biodiversity is greater now than it was a decade ago, and undoubtedly will be greater yet in another decade. The science has grown to match, spilling over into disciplines like genomics, economics, sociology, psychology, and ethics. A decade from now, most likely, a new Edition will be needed. For now, however, this Edition represents the state of the art, covering the latest advances in a rapidly changing subject. I am pleased to be a part of what has been a very satisfying partnership with Elsevier, our editors, and our authors.

S. Levin
Princeton University, USA

GUIDE TO USE OF THE ENCYCLOPEDIA

Structure of the Encyclopedia

The material in the encyclopedia is arranged as a series of articles in alphabetical order.

There are four features to help you easily find the topic you're interested in: an alphabetical contents list, a subject classification index, cross references and a full subject index.

1. **Alphabetical contents list**: The alphabetical contents list, which appears at the front of each volume, lists the entries in the order that they appear in the encyclopedia.
2. **Subject classification list**: This index appears at the start of each volume and groups entries under subject headings that reflect the broad themes of this encyclopedia. This index is useful for making quick connections between entries and locating the relevant entry for a topic that is covered in more than one article.
3. **Cross-references**: All of the entries in the encyclopedia have been extensively cross-referenced. The cross-references which appear at the end of an entry, serve three different functions:

   i. To indicate if a topic is discussed in greater detail elsewhere.
   ii. To draw the reader's attention to parallel discussions in other entries.
   iii. To indicate material that broadens the discussion.

Example

See also: Biodiversity Generation, Overview. Environmental Ethics.

4. **Index**: The index includes page numbers for quick reference to the information you're looking for. The index entries differentiate between references to a whole entry, a part of an entry, and a table or figure.
5. **Contributors**: At the start of each volume there is a list of the authors who contributed to the encyclopedia.

S

Stability, Concept of

Clarence L Lehman, University of Minnesota, St Paul, MN, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 467–479, © 2001, Elsevier Inc.

Glossary

Community The collection of all species living together in a given area. Alternatively, some designated subset, such as the avian community or the vascular plant community.
Deterministic system A dynamical system whose detailed future behavior can be predicted, in principle, for all time, assuming perfect knowledge of the system at the present.
Dynamical system A set of rules defining how certain variables change with time. Ecological models are dynamical systems representing what are believed to be (vastly more complex) dynamical systems in nature.
Dynamical variable A quantity in a dynamical system that changes with time according to the rules of the system (contrast with parameter).
Ecological model A vastly simplified mathematical representation of an ecological system intended to capture the full system's essence. Qualitatively, ecological models are to natural ecological systems as line drawings are to full-color photographs.
Equilibrium A condition of stasis in some dynamical variable.
Parameter A quantity in a dynamical system that is fixed as part of the rules of the system (contrast with dynamical variable).
Perturbation A temporary change in one or more dynamical variables or parameters of a dynamical system due to external factors.
Population A collection of individuals of the same species, often interacting ecologically and genetically.
Stochastic system A dynamical system whose detailed future behavior cannot be predicted, even in principle, due to random forces inherent in the system.

> Falling down once makes a building unstable.
> —GERALD WEINBERG (1975)

Background

The Meaning of Stability

Something is said to be stable if its condition tends to remain unchanged despite external influences. The population sizes of two different species may both be relatively constant in a steady, unchanging environment, but such constancy by itself discloses nothing of the stability of the two populations. If both populations were somehow perturbed—for example, by a spring flood eliminating half of each population—and if the first population subsequently recovered to former levels while the second declined to extinction, then the first population is said to have been stable while the second was unstable. Stability is not mere constancy—it also implies an ability to recover from perturbations.

The condition that recovers from perturbations need not be as simple as a constant population value. Populations of predators and prey, for example, may have an intrinsic tendency to cycle repeatedly through high and low values, but the cycle may be stable. Suppose the populations of a predator and its prey were both maintained at fixed levels by artificial management, such as by hunting. If the cycle returned after management ceased, then the cycle could be said to be stable, even though neither population is constant during the cycle. The cycle would be an abstract condition to which the populations return after a perturbation. In fact, the condition that recovers from perturbations can be any recognizable property of the system. It could be the number of species in the system, the collective biomass of all species together, the length of the food chain, the average quantity of nutrients leached from the soil, the degree of susceptibility to disease, or many others.

In the natural world, the stability of an object is closely tied to our perception of the very existence of that object. Any property that is unstable will not retain its condition after a disturbance, by the definition of stability. Given that the world is filled with continual disturbance, any property that is unstable is therefore likely to soon change to something else and hence not be observed. For example, if food chains beyond a certain length tend to be unstable in diverse communities, then a very long food chain, should it ever appear by

chance combination of species in an ecosystem, will eventually collapse to a length that is stable.

As human domination of global biogeochemical systems increases—including eutrophication of habitats, changes in atmospheric gases, fragmentation of habitats, distribution of toxic organic compounds, translocation of species, and a general reduction of biodiversity—the resulting perturbations test the stability of ecosystems. Also, as global changes take effect, induced changes in the systems alter the stability of their parts. Stability in ecosystems is therefore not just an abstract concept but something deeply connected to the persistence of the services on which all living beings, humans included, depend.

The Diversity–Stability Question

In the 1950s and 1960s, investigators such as Charles Elton argued that ecosystems containing more species would be more stable—less subject to fluctuations due to the myriad forces acting on them—owing to such factors as the greater number of pathways for energy to flow through them. This is the diversity–stability hypothesis. The hypothesis was generally accepted in the 1960s, although not without controversy.

Theoretical studies on the topic in the 1970s shook this general acceptance. Working with mathematical models of ecological communities, Robert May and others showed that such communities with more species were less stable: Populations of individual species returned to normal after a disturbance less rapidly as the number of species in the community increased. Although May (1974) later pointed out that certain ecosystem properties, such as total biomass of the community, can be more stable than biomasses of the constituent species, and despite objections by McNaughton (1977) and others to rejecting an accepted hypothesis based solely on theory, the acceptance of the diversity–stability hypothesis waned. During the 1970s and 1980s, ecologists generally expected that greater biodiversity would reduce stability, or at least that there was no consistent connection between biodiversity and stability.

By the 1990s, experimental evidence had accumulated that seemed to support both views simultaneously, depending on the level of focus. As the number of species increased, population densities of individual species became more variable from year to year but ecosystem parameters such as total productivity and nutrient leaching became less variable. In other words, the system seemed less stable when viewed in detail, from the perspectives of individual species, but more stable when viewed in total, from the perspective of the ecosystem (Tilman, 1999) (*see* Stability Under Changing Conditions: Biodiversity and Temporal Stability).

At the end of the twentieth century, issues concerning the effect of biodiversity on stability, and of environmental stability on biodiversity, were still being resolved, but greater biodiversity could be seen to stabilize certain important properties of the ecosystem as a whole. Many of the past controversies and apparent conflicts resulted from the application of different aspects and definitions of stability in different contexts.

Aspects of Stability

The seemingly simple idea of whether a system returns to its former condition after a disturbance has several aspects. Which is relevant depends on the ecological question at hand:

1. Does the system return to its former state after a disturbance, given unlimited time without further disturbance? This is the basic idea of stability, and if the system does return from all small disturbances, that state is said to be asymptotically stable or, commonly in the ecological literature, simply stable.
2. How much does the system fluctuate under variable conditions? This is called variability and it is typically measured statistically, for example, by the standard deviation of a time series, by the standard deviation divided by the mean, or by the reciprocal of that quantity. (The standard deviation being the square root of the average squared deviation from the mean.)
3. How small does the disturbance have to be in order that the system return? The set of all possible disturbances that allow the system to return to its former condition is called the domain of attraction. As the domain of attraction becomes vanishingly small, the system becomes unstable. Small domains of attraction correspond to fragile systems, whereas large domains of attraction correspond to robust systems.
4. How much is the system changed by a given disturbance? The more a given condition is changed, the lower the resistance associated with that condition. For example, suppose population values in a given area do not change much when average temperature increases above its normal but may change considerably when rainfall does so. Then the system would be said to be resistant to temperature changes but not to rainfall changes.
5. How fast does a perturbation decay? Resilience is the amount of time needed for a perturbation to be reduced to a specified fraction of its initial size. There has been variation in the use of this term in the ecological literature. Before the 1980s, resilience often simply referred to factors that reduced the chance that one or more species would become extinct (e.g., the size of the domain of attraction and the populations densities in that domain).
6. How long does the system take to return after a given disturbance? This is related to both how far the system is perturbed by the disturbance (resistance) and how fast the perturbation decays (resilience). The lower the resistance and resilience, the longer it takes to return.
7. How long can the condition be expected to last? The longer on average it lasts, the greater its persistence. The idea of persistence applies to systems subject to random variations, where the chance of encountering dangerously large variations during a time interval increases as the interval grows longer. The longer the time interval, for example, the greater the chance that one or more species will be driven extinct by random events.
8. Does the condition remain intact as the parameters of the system change slightly? If it does, the system is structurally stable with respect to those parameters. For example,

if precise birth and death rates of the constituent species change slightly occasionally, does the condition remain?

9. Does the condition remain intact under evolutionary pressures? If it does, the condition is evolutionarily stable. A condition may appear stable over short time, ignoring evolution, but may change on longer timescales. Indeed, over the course of geological time, no ecosystem is invariant.
10. Does the condition depend on spatial scale? What is stable in a small area may be unstable over a larger region and vice versa.

Pimm (1984) enumerated five definitions of stability in the ecological literature at three levels of complexity and three levels of organization—potentially 45 different nuances to the meaning of stability. Hence, care is necessary in defining and applying the concept.

Basic Theory of Stability

Equilibrium

The basic concepts of stability become precise when considered in the abstract. The many complexities in natural ecological communities, including factors that cannot be completely understood, mean that any abstraction of an ecological system is necessarily a vast reduction. To start this reduction, suppose the population increases solely by births and decreases solely by deaths. That is, processes such as immigration are not active. Suppose the rate of birth is highest and death is lowest when conditions are least crowded. As the population increases, resources become more limited—less light, reduced nutrient levels, less food, and so forth; the birth rate decreases and the death rate increases. At some point, the population will increase enough so that births just balance deaths. This occurs at the carrying capacity, at which an individual, on average during its lifetime, just replaces itself with a single offspring. Above this carrying capacity, deaths exceed births, so an individual on average does not quite replace itself. In the simplest case, births decrease and deaths increase along straight lines as the population increases (**Figure 1A**). The net rate of change per individual (the difference between births and deaths) is largest at the far left in **Figure 1A** (r) but decreases to zero when the population size reaches the carrying capacity (K) and becomes negative when the population size is greater than the carrying capacity.

The rate of change of the entire population is simply the rate of change per individual times the number of individuals in the population. This is graphed in **Figure 1B**, which shows the growth rate of the entire population (vertical axis) as a function of the population (horizontal axis). If the population is nonexistent (zero), then of course population growth is zero; with no individuals there are no births and can be no deaths. If the population is at carrying capacity, population growth is also zero; births balance deaths. These are two population equilibria—places where the population remains constant.

Between these two equilibria, the population grows. In this simplest case, in which the birth and death rates are straight

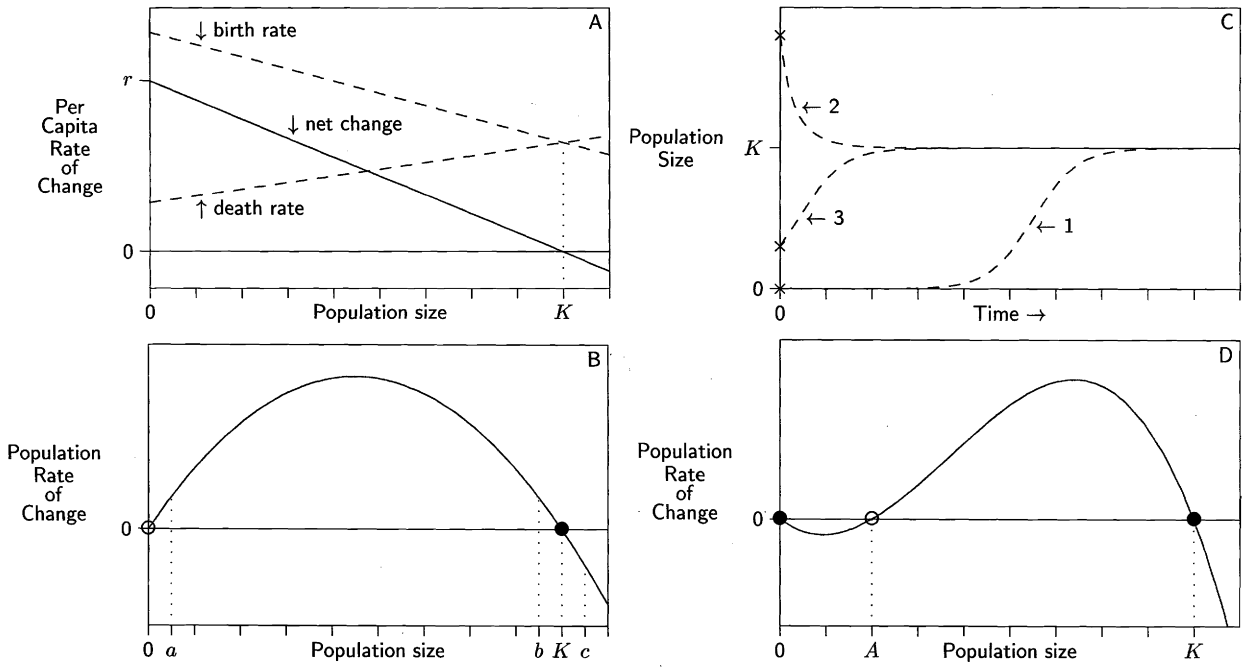


Figure 1 Geometric interpretation of stability. Solid and open circles represent stable and unstable equilibria, respectively. (A) Per capita birth and death rates are functions of population size; (B) growth of the entire population as a function of population size; (C) populations approaching carrying capacity with the passage of time, for the functions of A and B. The horizontal axis is time; the vertical axis is the population at the corresponding time. Cross symbols at the far left indicate the starting points. Curve 1 starts from a small population near zero. Curve 2 starts above the carrying capacity. Curve 3 starts below the carrying capacity but well above zero. (D) Growth of a population subject to the Allee effect, where 0 is a stable equilibrium.

lines, the population reaches the greatest rate of growth at half its carrying capacity. (In economic models, this is called the point of maximum yield.)

Geometric Interpretation

Single Species

To understand the stability of an equilibrium, the properties of the growth rate are first examined in the near neighborhood of the equilibrium. Consider the equilibrium at the carrying capacity (solid circle in **Figure 1B** at position *K*). The graph representing population change crosses the horizontal axis precisely at the carrying capacity, which means the population neither grows nor declines there. At that point the growth curve slopes downward to the right (the slope is negative). If the population for any reason falls below the carrying capacity (e.g., point *b* in **Figure 1C**), the growth function is above the axis, meaning that population growth is positive or that births exceed deaths. Thus, the population increases toward the carrying capacity. On the other hand, if the population is above the carrying capacity (e.g., point *c* in **Figure 1B**), the growth function is below the axis, meaning that population growth is negative or that deaths exceeds births. Thus, the population decreases toward the carrying capacity. In both cases, an external change moving the population away from its carrying capacity induces population growth or decline in exactly the right way to counteract the external change. The population returns to the carrying capacity. That equilibrium is stable.

Now consider the equilibrium at 0 (open circle in **Figure 1B**). As before, the graph representing population growth touches the horizontal axis precisely at 0. However, at this point the curve slopes upward toward the right (the slope is positive). Therefore, if the population ever increases, however slightly, above zero (e.g., point *a* in **Figure 1B**)—due, for example, to the arrival of propagules from some outside source—the growth function is above the axis, meaning that population growth is positive or that births exceed deaths. Thus, the population increases away from 0. The zero equilibrium is unstable.

Thus, internal dynamics carry these populations toward their carrying capacity (which is called an attractor) and away from 0 (called a repellor). If the population starts near zero, but not precisely at zero, it may linger a long time at low values and then begin a rapid growth phase before leveling off toward its carrying capacity (**Figure 1C**, curve 1). If the population starts above the carrying capacity, it rapidly decays back (**Figure 1C**, curve 2). If it starts below the carrying capacity but well above zero, it can increase to the carrying capacity without the long lag period (**Figure 1C**, curve 3).

Local Versus Global Stability

In the previous example, the carrying capacity is said to be a global attractor or globally stable because almost all replicate populations eventually arrive there (all but those starting at zero in this case). Because the equilibrium at 0 is unstable, the population has a level of permanence; if driven to low values, but not completely to zero, it can recover spontaneously.

This contrasts with populations operating under an Allee effect, wherein the equilibrium at 0 is an attractor. **Figure 1D** depicts the growth function in such a situation. Very low densities inhibit reproduction, for example, by reducing chances of encountering a mate. There are three equilibrium points—at 0, *A*, and *K*. Slopes both at 0 and at *K* are negative, meaning that these equilibria are stable, whereas the slope at *A* is positive, meaning that equilibrium is unstable. If such populations are driven to a low enough level (below *A*), they will spontaneously become extinct.

In this case, both 0 and *K* are local attractors (locally stable), but neither is a global attractor (neither is globally stable). If the system is pushed a small distance away, it will return to its former state. However, if pushed too far (beyond the domain of attraction of that equilibrium), the system will switch to another state. This is a bistable system. The region between 0 and *A* is the domain of attraction of the 0 equilibrium; the region between *A* and *K* and the entire region above *K* make up the domain of attraction of the carrying capacity. There is no domain of attraction for the equilibrium at *A* because this equilibrium is unstable (it is a repellor).

Multiple Species

The geometric example of **Figure 1** represents a single species in isolation. In this case, geometric stability arguments are direct and intuitive. The geometric arguments are similar for more than one species, but instead of curves representing growth rates, surfaces in multidimensional space are used. Each additional species requires another dimension. Unfortunately, such surfaces are difficult to visualize for more than two species.

The growth surface for a single species will in general cut through the plane of zero growth (analogous to the horizontal axis in **Figures 1B** and **1D**) along a curve. This curve is called the zero net growth isocline, and each species in general will have its own such curve. If the zero net growth isoclines for many species meet at a single point, then this is a multispecies equilibrium point. For this equilibrium to be stable, however, it is not sufficient that the growth surfaces of all species, considered separately, have negative slopes there. Interactions among the species must be accounted for, and this is best done algebraically.

Algebraic Methods

The geometric interpretation of stability, as depicted in **Figure 1B**, has a direct algebraic interpretation in terms of derivatives and eigenvalues, which correspond geometrically to slopes. Where an algebraic description of the ecological system is available (i.e., in ecological models), the algebraic method is widely used and can be followed as a recipe. The algebraic method is essential for theoretical work in ecology, but the material in this section is not essential for an intuitive understanding of the concept of stability.

Single Species

Consider an ecological system of a single species in which the symbol *N* represents the population size. This may be measured in individuals, biomass, or other units. The rate of

Table 1 Common ecological usage vs common mathematical usage

Term	*Size of perturbation with time*	
	Ecological literature	*Mathematical literature*
Asymptotically stable	Decreases	Decreases
Stable	Decreases	Does not increase
Unstable	Increases	Increases
Neutrally stable	Remains unchanged	(Typically not used)

change of this population is represented by dN/dt, which tells how large an increment (dN) occurs in the population during a small increment of time (dt). Assuming that this rate of change is a function of the population size, as in the previous discussion, then the system will be described by $dN/dt = f(N)$. For example, in **Figure 1**, $f(N) = rN(1 - N/K)$, which is the well-known logistic equation described in most introductory ecology texts.

Equilibria are commonly designated by symbols like $\hat{N}$ (pronounced "hat N"). They occur at such $\hat{N}$ where $f(\hat{N}) = 0$. If the slope of the function $f(N)$ at any such equilibrium is negative (i.e., if $df/dN\big|_{N=\hat{N}} < 0$, as shown by the solid circles in **Figures. 1B** and **1D**), then the equilibrium is stable. If the slope is positive (as shown by the open circles in **Figures. 1B** and **1D**), then the equilibrium is unstable. If the slope is zero, then further information is needed to determine stability. In the case of zero slope, the equilibrium may be stable, unstable, or neutrally stable, depending on the exact shape of the function near the equilibrium. In the ecological literature, neutral stability means that if the system is perturbed from equilibrium a small amount, it neither returns nor moves further away but rather maintains its new value. (Note that there is a difference between the way ecologists use the word "stable" and the way mathematicians do, as summarized in (**Table 1**.)

Multiple Species

Multiple species are represented by multidimensional dynamics, in general described by m equations of the form $dN_i/dt = f_i(N_1, N_2, N_3, \ldots, N_m)$, one equation for each species. Equilibria are multidimensional points $(\hat{N}_1, \hat{N}_2, \hat{N}_3, \ldots, \hat{N}_m)$ that cause all dN_i/dt to be simultaneously zero. Some of the $\hat{N}_i$ may be zero, in which case fewer than m species coexist at that equilibrium.

The slope of a system in one dimension generalizes to eigenvalues of the Jacobian matrix in multiple dimensions. (In one dimension, the slope is a special case of an eigenvalue.) To determine stability of an equilibrium, (1) construct the Jacobian matrix of the system, $J = \{\partial f_i/\partial N_j\}$, (2) substitute the equilibrium value $(\hat{N}_1, \hat{N}_3, \ldots, \hat{N}_m)$ into the matrix of step 1, and (3) determine the eigenvalues of the matrix of step 2. Neuhauser (2000) or other mathematical texts may be consulted for full details.

The resulting eigenvalues are complex numbers, and as such may contain both real and imaginary parts. (That is, an eigenvalue has the form $a + bi$, where a is the real part, bi is the imaginary part, and $i = \sqrt{-1}$). The real part determines the rate of approach to or retreat from equilibrium after perturbation. In one dimension, the imaginary part is always zero, but in higher dimensions it need not be. When it is not zero, the approach to or retreat from equilibrium follows a spiral path.

Only the real parts of the eigenvalues affect stability. If the real parts of all eigenvalues are negative, that equilibrium point is stable. If one or more real parts is positive, that equilibrium point is unstable. If none is positive but one or more real parts are zero, then further information is needed to determine stability. If some eigenvalues have negative real parts but others have positive, then the equilibrium is a saddle point (so called because of its shape in the two-dimensional case). Such saddle equilibria are attracting in some directions and repelling in others; hence, they are unstable.

Discrete-Time Systems

Many ecological models are more naturally defined on a discrete time axis. For example, insects may emerge together at a specific time each year or seeds may be set once per year in the fall. In such models, time is better represented not as a continuum but rather as a series of integers (1, 2, 3, …), for example, representing a series of years. In simplest form, discrete-time models have the structure $N_i(t+1) = f_i(N_1(t), N_2(t), N_3(t), \ldots, N_m(t))$, where $N_i(t)$ is the population level of species i at time t. To determine stability of an equilibrium in a discrete-time system, the three steps outlined previously are followed, but the resulting eigenvalues must be interpreted differently. Again, the eigenvalues are complex numbers. In the discrete-time case, it is the absolute values of the eigenvalues that determine stability. That is, it is the positive square root of the sum of the squares of the real and imaginary parts. (For the eigenvalue $\lambda = a + bi$, $|\lambda| = \sqrt{a^2 + b^2}$.) If the absolute value of all eigenvalues is less than 1, then the equilibrium is stable. If the absolute value of one or more eigenvalues is greater than 1, that equilibrium point is unstable. If none is greater than 1 but one or more absolute values are precisely equal to 1, then further information is needed to determine stability, as in the case of zero eigenvalues discussed previously.

Strength of Return

The magnitude of the eigenvalues (or, equivalently, the steepness of slope) determines the resilience, or the strength of return to equilibrium. If the eigenvalues are all very negative, the corresponding slopes will be very steep. Small deviations from equilibrium will result in large rates of growth or decline, and the system will return rapidly. On the other hand, if any of the eigenvalues is negative but close to zero, then the corresponding slopes will be shallow. Small deviations from equilibrium will not lead to rapid recovery. Of course, given a sufficiently long time, and given the absence from disturbance, the system will return regardless of the eigenvalue's magnitude, given that it is negative. However, over short times, or with repeated disturbances, the magnitude of the eigenvalues is an important part of stability.

Stability Under Changing Conditions

Biodiversity and Temporal Stability

If a system is subject to repeated disturbances, variations in a given property of that system over time will be related to the

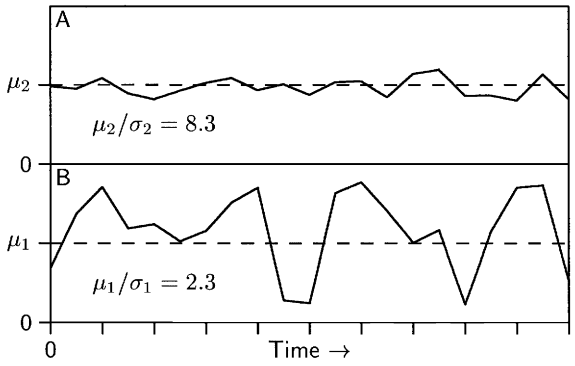


Figure 2 Temporal stability. Dashed lines indicate the long-term means, and solid lines indicate fluctuating variables. (A) Relatively small fluctuations about the mean, with a relatively high measure of temporal stability (8.3); (B) larger fluctuation about the mean, indicating lower stability, with a lower measure of temporal stability (2.3).

asymptotic stability of that property. Because asymptotic stability is an abstract concept, well defined in ecological models but difficult to determine in natural systems, various measures of fluctuations are often used instead to quantify stability of natural systems.

Suppose two populations are fluctuating through time about their individual mean abundance levels due to stochastic effects in the environment and possibly due to internal dynamics as well. Suppose fluctuations in the first population are on average relatively small compared to the first population's mean abundance, whereas fluctuations in the second population are noticeably larger compared to the second population's mean abundance (**Figure 2**). Independent of the idea of asymptotic stability, the first system can be defined to be more stable with respect to constancy in abundance. Precisely how this fluctuational stability is related to the concept of asymptotic stability discussed in the previous section is a separate theoretical issue.

The average fluctuation can be quantified in alternative ways; for example, as the average absolute value deviation from the mean or as the standard deviation of the mean. In any case, if fluctuations are large relative to the mean, the system is likely to obtain values very far from the mean. Also, if the underlying dynamics have a domain of attraction that is responsible for maintaining the mean, then large fluctuations are more likely, on rare occasions, to combine in the wrong direction and push the system out of that domain, inducing a switch to a new mode of behavior. Thus, the relative amount of fluctuation can carry information on the long-term persistence of the system.

Measures proportional to the coefficient of variation—the standard deviation over a time series divided by the mean of that time series—have been used as a measure of fluctuation. These are actually proportional to instability because larger values in the coefficient of variation correspond to lower stability. The reciprocal of the coefficient of variation (i.e., the mean divided by the standard deviation, μ/σ) has been termed temporal stability (Tilman, 1999). It carries the same information, but larger values of temporal stability correspond to greater stability of the system. In this form (μ/σ), it is reminiscent of signal-to-noise ratio in engineering. Clearly, temporal stability is greater if the mean is greater (μ, the numerator), if the standard deviation is smaller (σ, the denominator), or both. Therefore, any forces that tend to increase the mean or decrease the variation will increase temporal stability. Temporal stability is maximal (infinite) when the standard deviation is zero—when there is no variation at all.

Evidence from both experiment and theory suggests that the temporal stability of certain community characteristics of competitive plant communities tends to increase as the number of species they comprise increases (Tilman, 1999). Along a long-term nitrogen fertilization gradient, the number of species varied, changing from diverse prairie communities at low nitrogen levels to only a few agricultural grass species at high nitrogen levels. Individual species abundances varied from year to year due to many causes, both environmental and ecological. As the number of species increased, (i) the total biomass of the community tended to increase and (ii) the total biomass of the community became less variable. Simultaneously, the biomass of an individual species tended to become more variable relative to the mean biomass for that species embedded in the community. In other words, increased biodiversity appears to stabilize the community while destabilizing the individual species.

Given this pattern from experimental systems, a theoretical question immediately arises: Do standard ecological models predict similar phenomena? This is a rapidly developing area, but initial results appear to indicate that they do (Tilman, 1999). In a model of competition for a single resource, and in other more general models, temporal stability of total community biomass increased steadily and linearly as the number of species increased (**Figure 3A**). Temporal stability of individual species biomass decreased sharply as biodiversity increased from one to about five species, and then it leveled off at higher diversities (**Figure 3B**). In other words, simple ecological models appear to predict the effects observed along the nitrogen gradient—increased biodiversity will stabilize the community but destabilize individual species.

The importance of various possible causes of stabilization of community properties remain under discussion. When abundances of individual species are perturbed by complex and effectively random forces, some abundances may be perturbed to higher levels and others to lower levels. On statistical grounds, the variation in the total biomass will likely contain both positive and negative perturbations. Hence, the total is likely to fluctuate less, relative to its mean value. This has been called the portfolio effect (Tilman, 1999) by analogy to economics, wherein a portfolio diversified over many investment instruments will fluctuate less than one containing only a few. It depends on species responding differently to perturbations, if only for statistical reasons. Effects of competition allow species to expand in abundance when a competing species is suppressed by external forces. This expansion, in turn, buffers any change in total biomass. This is called the covariance effect because it is evident when there is strong negative covariance between pairs of species. Finally, increased biodiversity also often leads to increased total abundance, and this increases stability by making any given level of fluctuations smaller in proportion. This is called the overyielding effect.

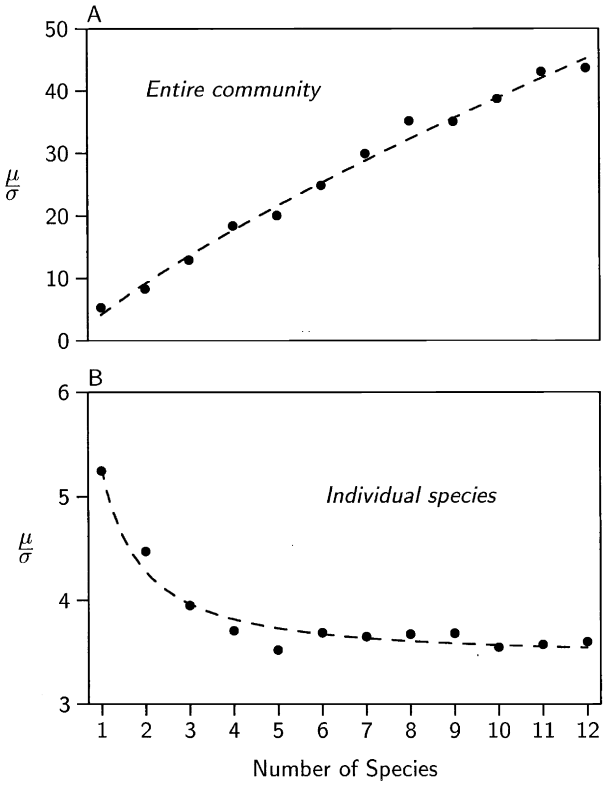


Figure 3 Biodiversity affects temporal stability. From a resource model reported by Tilman (1999) showing patterns similar to those observed in nature. Large dots indicate means of 100 replicates. (A) Increasing biodiversity stabilized total biomass; (B) increasing biodiversity destabilized individual species biomass.

Regardless of the source of the stabilization, however, it appears that certain composite community properties may be stabilized by increased biodiversity. More direct experiments that will eventually be able to clarify the effects of biodiversity on stability and other ecosystem properties have been established in many countries (Hector *et al.*, 1999; Tilman, 1999). Results of this coordinated set of experiments will emerge during the first decade of the twenty-first century and bear watching for information they will provide on the diversity–stability question.

Temporal Niches

Temporal stability, as defined previously, certainly applies to a system continually perturbed from equilibrium. However, equilibrium was not part of the definition—only the mean and the variation about the mean participated in the measure of temporal stability. Temporal stability therefore applies to cases in which equilibrium, stable or otherwise, may not even exist.

Changing environmental conditions, such as variation in average rainfall from year to year, can favor different species at different times. This, in effect, can partition resources among species, setting up temporal niches and permitting long-term coexistence of species that greatly outnumber the resources they consume (Chesson, 1994). Without changing conditions, biodiversity would be reduced—coexistence would be limited by the number of resources available.

For a given regime of environmental fluctuations, and for a given community, each species will have some pattern of fluctuations in response to the fluctuations in the environment and those of other species in the community. Again, the issue with respect to biodiversity is how the relative size of the fluctuations—the temporal stability—of each species depends on the number of species in the community and how community-level properties such as total productivity or nutrient leaching depend on the number of species.

Coexistence among species in the absence of a multi-species equilibrium can result from stochastic variations in the environment, as explained previously. However, it is not necessary that the variations be stochastic, nor that their source be the external environment. The variations could result from some regular, well-determined periodic changes in the environment, or they could result from population cycles set up by dynamics of the ecological community (Armstrong and McGehee, 1980; Huisman and Weissing, 1999). Whether the effects are external or internal, stable coexistence at increased biodiversity can exist under fluctuating conditions.

Emerging Deterministic Stability

Stable deterministic characteristics such as those described in the section on Basic Theory of Stability can emerge from stochastic systems. Individuals may live and die, populations may fluctuate in abundance, and some species may become extinct while others appear; however, amid all this complexity, simple patterns of biodiversity can emerge.

One of the early successful theoretical explanations of biodiversity concerned island biogeography (MacArthur and Wilson, 1967). On islands, simple deterministic characteristics emerge amid complex ecological change. It is a conspicuous fact that oceanic islands have fewer species than adjacent mainlands. This effect can be attributed, to a large extent, to a stable balance between local extinction of resident species and immigration of new species.

A new individual arriving on the island may be a member of a species already resident on the island, or it may represent a novel species. If the island is devoid of life, the new individual is certain to be a member of a novel species. As the island becomes more populated with mainland species, the chance that the individual is a member of a novel species decreases. Similarly, the chance that an existing species vanishes from the island increases as the number of resident species increases.

Such considerations lead to immigration and extinction curves such as those in **Figure 4**. Notice the qualitative similarity of the immigration and extinction curves in **Figure 4** to the birth and death curves, respectively, in **Figure 1**. The immigration and extinction curves intersect at an equilibrium point, where resident species becoming extinct are balanced by new species arriving. Note that at this equilibrium, neither population values nor community composition are constant. In fact, the individual species making up the community are constantly in flux. It is simply the biodiversity that remains constant.

Is this biodiversity equilibrium stable? The net rate of increase or decrease in the number of species is simply the difference between the immigration and extinction rates, which

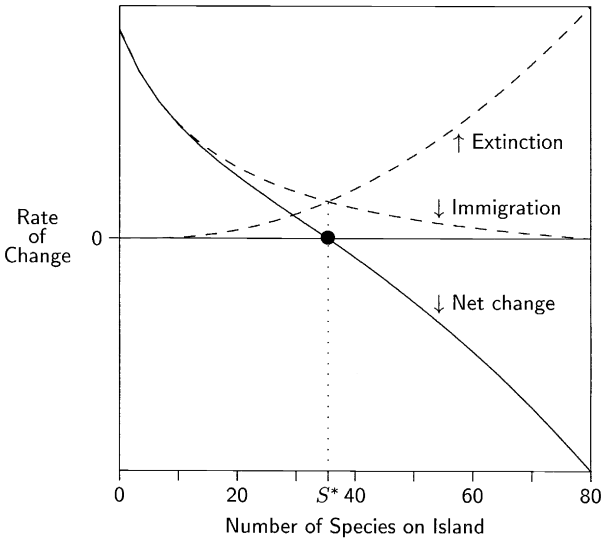


Figure 4 Stable island equilibrium. Island biogeographic immigration and extinction curves patterned after observed rates for real islands. Biodiversity reaches a stable equilibrium when immigration balances extinction.

is the solid curve in **Figure 4**. Unlike the curves of **Figure 1**A, this net rate represents the entire island; it is not a per capita or per species rate. Therefore, its slope at the equilibrium directly corresponds to stability—it need not first be multiplied by the number of species on the island. Because the slope is negative, by the arguments of the section on Basic Theory of Stability this equilibrium is stable.

Stability Amid Chaos

Consider a population that changes deterministically according to the rules of **Box 1**—a straight-line relative of the logistic equation of **Figure 1**. Populations oscillate chaotically between high and low values, but except for certain infinitely rare starting conditions they never return to a previous state (**Figure 5**). This system is chaotic in the sense that slight deviations are always magnified. In almost all cases, two slightly different starting populations, no matter how nearly identical, grow increasingly different with time.

However, amid such complete chaos can be stability. After sufficient time, population values fall into a pair of disjoint intervals, shaded in **Figure 5**. This pair of intervals is an attractor. The population then alternates regularly between the intervals, though its position within either interval cannot be predicted for very long from measurements of finite precision. This is a deterministic system with dynamics that appear superficially stochastic, but among all its instabilities arises another level of stability.

Other Kinds of Stability

Structural Stability

In the cases discussed previously, the ecological system did not change. Dynamical variables such as population levels were perturbed, but ecological parameters such as birth rates remained fixed. Ecological parameters, however, constantly change: Global temperatures increase, glaciers retreat, and spruce forests give way to pine and then hardwoods. Rainfall, soil substrates, and a host of environmental conditions alter the parameters and the structure of ecological systems. Hence, not only may the dynamical variables such as population abundances be perturbed but also the very structure of the ecological system may change.

If a given property of a system persists under small changes in the system itself, then the system is said to be structurally stable with respect to that property. The equilibrium carrying capacity in the system of **Figure 1**, for example, is structurally stable. Changes in the slopes of the birth and death rates, or small changes that make the birth and death rates curves rather than straight lines, still leave a carrying capacity that all nonzero initial populations eventually reach, provided that the net per capita growth declines smoothly with increasing population. The precise size of the carrying capacity may change, but the fact that it exists and is stable does not.

Classical Predator–Prey Systems

Some ecological models are on a razor edge of structural instability. In the simplest form of the classical Lotka–Volterra predator–prey system, included in most introductory ecology texts, the prey population is limited by predation and the predator population is limited by availability of prey (first pair of equations in **Box 2**). Predator and prey populations

Box 1

$$N_{t+1} = \begin{cases} 2N_t & \text{if } N_t < a \\ 2a(1-N_t)/(1-a) & \text{if } N_t \geq a \end{cases}$$

Piecewise linear discrete-time population growth.If $a<1/3$, there is a stable equilibrium at $N=2a/(1+a)$. If $a>1/3$, there are no stable equilibria or cycles (see **Figure 5**).

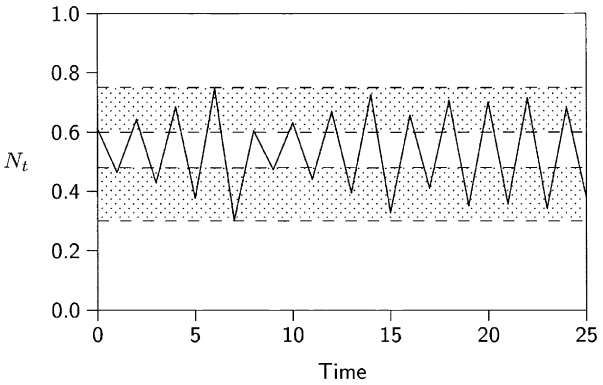


Figure 5 Stable attractor enclosing unstable population cycles. The system of **Box 1** with $a=3/8$ is chaotic, but population fluctuations are pulled into the shaded intervals (0.3, 0.48) and (0.6, 0.75). The horizontal axis is time; the vertical axis is the population at the corresponding time.

Box 2

$$\frac{dN}{dt} = r_N N - aNP$$
$$\frac{dP}{dt} = r_p NP - bP$$
$$\frac{dN}{dt} = r_N N\left(1 - \frac{N}{K}\right) - aNP$$
$$\frac{dP}{dt} = r_p NP - bP$$

Lotka–Volterra predator–prey system. Top two equations, no carrying capacity for prey; bottom two equations, carrying capacity included (see **Figure 6**).

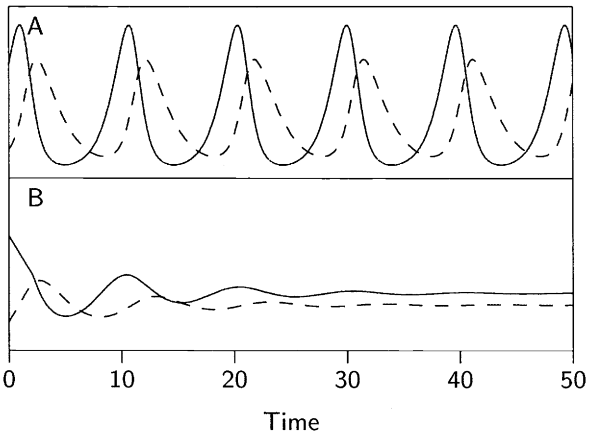


Figure 6 Structural instability in Lotka–Volterra predator–prey cycling. Solid curves represent population densities of the prey, and dashed curves represent that of the predator. Here, $r_N = 1$, $r_p = 1/2$, $a = 1$, $b = 1/2$, $K = 5$, $N(0) = 2$, and $P(0) = 1/2$. (A) The prey population is limited only by predation (top two equations in **Box 2**). The populations oscillate on neutrally stable cycles. This effect is structurally unstable. (B) The same starting conditions when the prey is also limited by its own carrying capacity in the absence of the predator (bottom two equations in **Box 2**). Initial variations damp out. This effect is structurally stable.

oscillate indefinitely about an equilibrium, with the predator population lagging behind that of the prey (**Figure 6A**). This is a common characteristic of predator–prey systems or, more generally, of producer–consumer systems. However, this particular system has a peculiar property for an ecological system: It possesses a "memory" of past events. If something perturbs either population, or both populations, to a new level, that new level will be revisited on each subsequent cycle. This system has no asymptotically stable behavior; instead, everything is neutrally stable.

Such neutral stability would not necessarily be pathological in an idealized physical system, such as a perfect harmonic oscillator, but here it is pathological. All abstractions are simplifications, and here one of the simplifications is the assumption that the prey population is limited only by the predators; in the absence of predators, prey can increase in numbers without bound. Inclusion of any carrying capacity for the prey, no matter how large or small (as in the second pair of equations in **Box 2**), changes the dynamics completely. The equilibrium becomes stable, oscillations die out, and memory of past perturbations fades with passing time (**Figure 6B**). Another simplification is that predators are never satiated: They consume all the prey they encounter. Inclusion of a more realistic predator response can make the equilibrium unstable. The oscillations converge to a stable cycle of fixed amplitude, again with memory of past perturbations fading.

Thus, this simplest Lotka–Volterra formulation is structurally unstable with respect to important ecological factors. Structurally unstable ecological systems do not commonly appear in nature, so conclusions drawn from structurally unstable models must be used with caution.

Related Effects of Eutrophication

The change in structure from stable equilibrium to stable limit cycle, as described previously, may be induced by changes to parameters of the ecosystem. Eutrophication is enrichment by high levels of nutrients, such as increased phosphorus in a lake or nitrogen in the soil. Eutrophication is one of the principal effects of human domination of ecosystems, and it may affect both their biodiversity and their stability.

Graphical analysis by Rosenzweig and others in the 1970s (Rosenzweig, 1990) showed that stable equilibria in producer–consumer systems were favored when the carrying capacity of the consumer was relatively small. If the carrying capacity of the consumer were higher, the equilibrium would be less stable (less resilient). At a sufficiently high carrying capacity, the system would pass through a structural instability and the equilibrium would lose its stability entirely. Oscillations in producers and consumers would then occur. At even higher carrying capacities, the size of the oscillations would increase, driving both producer and consumer periodically to low population levels, thereby increasing the chance of extinction.

Now, carrying capacity is directly related to the level of resources available. Eutrophication increases these resources, thereby increasing the carrying capacity of the producer. Ironically, the act of providing more food or resource to the producer can lead to its destruction as a result of induced instability. This effect has been called the paradox of enrichment. A common result in observed fertilized terrestrial systems is a reduction in the number of species. Although there can be many reasons for this, loss of stability upon enrichment has been suggested as a contributor (Rosenzweig, 1990).

Spatial Stability

A property that is stable at one spatial scale need not be stable at larger or smaller scales (Levin and Segel, 1985). This was decisively demonstrated in the mid-twentieth century by the mathematician Alan Turing. Turing's simplest example was a (nonecological) dynamical system operating in two separate cells, optionally with some migration between the cells (**Box 3**). If the two cells are completely disconnected, then there is an asymptotically stable equilibrium that both cells approach. Thus, if the density in either cell is perturbed, it will

Box 3

$$\frac{dx_1}{dt} = ax_1 - by_1 + h - m_x(x_1 - x_2)$$

$$\frac{dy_1}{dt} = bx_1 - cy_1 + h - m_y(y_1 - y_2)$$

$$\frac{dx_2}{dt} = ax_2 - by_2 + h - m_x(x_2 - x_1)$$

$$\frac{dy_2}{dt} = bx_2 - cy_2 + h - m_y(y_2 - y_1)$$

Turing's basic example of diffusive instability. Variables x_i and y_i are densities in cell i. The remaining symbols are constant coefficients, $a = 5$, $b = 6$, $c = 7$, $h = 1$, $m_x = 1/2$, and $m_y = 9/2$ (see **Figure 7**).

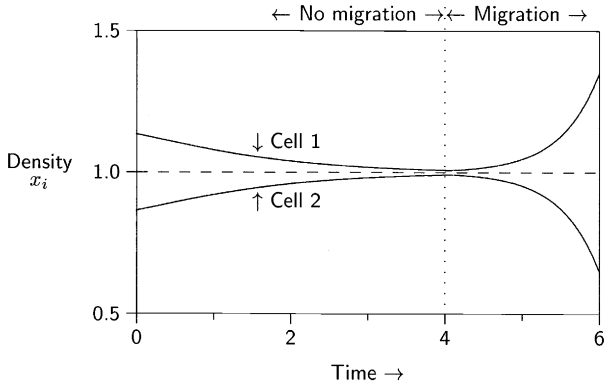


Figure 7 Spatial diffusive instability in a two-cell example. The vertical axis is density, or concentration, of two of the variables (x_1 and x_2) for the example shown in **Box 3**. The two cells are isolated until time 4, when a migration corridor between the cells is opened. Isolated, each cell approaches the same stable equilibrium (times 0–4). Connected, this equilibrium is destabilized by migration and the two cells diverge rapidly (times 4–6). Initial conditions are $x_1(0) = 1.13533$, $y_1(0) = 1.1218$, $x_2(0) = 0.864666$ and $y_2(0) = 0.8782$.

return to its equilibrial value (**Figure 7**, times 0–4). However, if there is sufficient migration between the cells, then the equilibrium, though it still exists, becomes unstable. Any random deviation that causes a difference between cells, however slight, unbalances migration and causes a sustained and accelerating net transport toward the cell with the higher concentration (**Figure 7**, times 4–6). This is called a diffusive instability or a Turing instability. An equilibrium that is stable at a point of space can become unstable when its components can diffuse to neighboring parts of the space. This resulting instability can lead to variations and patterns in biodiversity over the landscape, with the resulting patterns being stable.

The opposite can also hold. A system that is unstable in a point of space can become stable, or at least persistent, when extended over a region. Predators clearly coexist with their prey and parasites coexist with their hosts for long periods in nature. However, such coexistence has been difficult to scale down to the size of experiments. Many of the early experiments on predator–prey and host–parasite systems found them to be unstable at small scales. Confined to a small area, the predators captured all the prey, and then themselves disappeared from the experiment for lack of food (or, analogously, the parasites infected all the hosts and then themselves disappeared). However, as the spatial extent of the experiment was increased, persistence of the system increased dramatically. Ecological models of such systems show a similar behavior (Hassel and Wilson, 1997). Although each individual cell quickly runs to extinction, migration "rescues" empty cells, and the host and the parasite can have a stable average density over a large group of cells. Thus, an equilibrium that does not exist or is unstable in a point of space can be replaced by a stable equilibrium averaged over the entire region.

Something similar happens in ecological systems described by metapopulation dynamics (Hanski, 1997). Individual populations of a species may be separated by distances or barriers that inhibit movement between populations. If movement is strongly inhibited, individual populations may be driven extinct locally by stochastic effects, only to be restored later by propagules from some other population. Despite individual populations going in and out of existence, the portion of local sites occupied at any time can approach a stable equilibrium. This locally nonpersistent system can be both persistent regionally and have a stable regional equilibrium.

Evolutionary Stability

On the ecological timescale, ecological parameters such as birth and death rate are considered fixed. They are taken to be immutable characteristics of the species under consideration. However, ecological systems have many layers of complexity beyond our simple abstractions of them. On the evolutionary timescale, the parameters are malleable through the process of mutation and natural selection. Moreover, species formerly isolated may come into contact either as a result of natural causes or, now with great frequency, through the actions of humans. Both invading species and mutant phenotypes present new parameters that test the evolutionary stability of the system. A condition that is stable with ecological parameters fixed need not be stable when these parameters can change.

How the parameters might change can be seen by examining the growth rate of potential invaders, initially at negligible densities, entering the community at equilibrium. If the growth rate of any phenotype is positive, then the system can be invaded by that phenotype. The resident species, and hence the parameters of the system, are evolutionarily stable only if the growth rate of every potential invader is negative. Originally developed for behavioral systems, this idea is being applied to ecological communities (Geritz *et al.*, 1998), and such work promises to shed light on the properties of invasions and ultimately on the evolution of biodiversity.

See also: Carrying Capacity, Concept of. Disturbance, Mechanisms of. Eutrophication and Oligotrophication. Habitat and Niche, Concept of. Island Biogeography. Population Dynamics. Species, Concepts of. Stress, Environmental

References

Armstrong RA and McGehee R (1980) Competitive exclusion. *Am. Nat.* 115: 151–170.

Chesson P (1994) Multispecies competition in variable environments. *Theor. Population Biol.* 45: 227–276.

Geritz SAH, Metz JAJ, Kisdi É, and Meszéna G (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* 12: 35–37.

Hanski I (1997) Predictive and practical metapopulation models: The incidence function approach. In: Tilman D and Kareiva P (eds.) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, pp. 21–45. Princeton, NJ: Princeton Univ. Press.

Hassel MP and Wilson HB (1997) The dynamics of spatially distributed host–parasatoid systems. In: Tilman D and Kareiva P (eds.) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, pp. 75–110. Princeton, NJ: Princeton Univ. Press.

Hector A, Schmid B, Beierkuhnlein C, *et al.* (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286: 1123–1127.

Huisman J and Weissing FJ (1999) Biodiversity of plankton by species oscillation and chaos. *Nature* 402: 407–410.

Levin SA and Segel LA (1985) Pattern generation in space and aspect. *SIAM Rev.* 27: 45–67.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography.* Monographs in Population Biology 1. Princeton, NJ: Princeton Univ. Press.

May RM (1974) *Stability and Complexity in Model Ecosystems.* Monographs in Population Biology 6. Princeton, NJ: Princeton Univ. Press.

McNaughton SJ (1977) Diversity and stability of ecological communities: A comment on the role of empiricism in ecology. *Am. Nat.* 111: 515–525.

Neuhauser C (2000) *Calculus for Biology and Medicine.* New York: Prentice Hall.

Pimm SL (1991) *The Balance of Nature? Ecological Issues in the Conservation of Species and Communities.* Chicago: Univ. of Chicago Press.

Rosenzweig ML (1995) *Species Diversity in Space and Time.* Cambridge, UK: Cambridge Univ. Press.

Tilman D (1999) The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80: 1455–1474.

Stewardship, Concept of

Peter Alpert, University of Massachusetts, Amherst, MA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp. 481–494, © 2001, Elsevier Inc.

Glossary

Community-based conservation Conservation-oriented management of communal or public property by local residents.

Ecosystem management Management of an ecologically defined geographical area as an integrated whole for both sustainable yields of natural resources and maintenance of ecological processes and species populations.

Integrated conservation and development project Project that links biological conservation with human development in a local area or set of areas.

Land ethic Philosophy of proper treatment of lands and the natural communities on them.

Natural resources management Management of natural systems to maintain or provide yields of wood, water, and other resources for people.

Nongovernmental organization An organization that is not an agency or part of a government, typically one that is not-for-profit and not owned by individuals.

Stewardship Management of a thing for someone or something else.

Etymology

The term "steward" has existed for approximately 1000 years. According to the *Oxford English Dictionary*, its early form "stigweard" first appeared in manuscripts in the eleventh century. The prefix "stig-" is probably the Old English word for "house or some part of a house." "Weard" is an Old English word for keeper. A steward is literally a "household-keeper." This gives the word an interesting kinship with the word "ecology," coined eight centuries later from Greek words for "house" or "household" and "study." The steward manages what the ecologist studies.

The essence of the idea of a keeper is one who manages something for someone else. For example, the Old Teutonic root "wardo," presumably related to "weard," gave rise to the word "warden." This word was used in a conservation sense during the Renaissance. A royal act under Edward I in 1543 proclaimed that "The wardeyns shall kepe and susteyne the lands without makynge dystruction of any thynge." In thirteenth-century France, wardo became "garde" and then "guardian" in fifteenth-century England, meaning "one to whom the care and preservation of any thing is committed." Like his Latinate equivalent, a "custodian," a "guardian" could manage but not rule: "From that yere … were al custodyes & gardeyns and no mayres" (Robert Fabyan, *The Newe Chronycles of Englande and of Fraunce*, 1516).

The "household" in "steward" may at first have meant exactly that. The first meaning of "steward" in the *Oxford English Dictionary* is "an official who controls the domestic affairs of a household, supervising the service of his master's table, directing the domestics, and regulating household expenditures." However, usage of "stewardship" soon extended beyond the kitchen and pantry. As the "royal households" in England and Ireland became more exalted, "steward" became a title of high office. In early Scotland, the Lord High Steward was the first officer of the king. One such official, Robert the Steward, claimed the throne. As the new ruler of the royal household, he was no longer its steward, and his epithet changed, founding the royal house of Stuart. On the other hand, a king threatened with the loss of his crown might find it expedient to claim to be in the employ of a still higher authority and therefore to be a steward. Shakespeare had Richard II make this case: "Show us the hand of God, That hath dismiss'd us from our stewardship." The formal usages of stewardship now include the management of any property owned by someone else.

Some contemporary common usages of "steward" adhere to these formal meanings. People today probably encounter the word steward most often in the person of an airline or ship's steward, someone in charge of food service on an airplane or ship owned by someone else. Organizations such as The Nature Conservancy employ "land" or "reserve stewards" to manage natural areas. These stewards are paid to manage part of an estate for the owner, which is the organization.

Other common usages of "steward" include the management of intangible things, such as rights or beliefs. A "shop steward" is a union official charged with protecting the rights of a group of employees; a "religious steward" may be a volunteer charged with helping to maintain a faith or church. Stewardship can also refer to the management of things owned by the steward, as long as he or she is acting on behalf of interests greater than his or her own. Currently, the most common use of "steward" in relation to the conservation of biodiversity is for private individuals or companies that agree to abide by special practices designed to protect or restore natural species or ecology on their own lands. A wide variety of national, state, and local governmental and nongovernmental agencies and organizations sponsor programs that recruit private landowners to volunteer as such "ecological," "environmental," or "habitat stewards."

Despite these extensions of the original meaning of "steward," the concept of stewardship remains distinct from the concept of belonging, on the one hand, and from the concept of domination, on the other hand. Belonging implies

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00136-2

being an equal or subordinate part of a thing rather than managing it. The first premise of stewardship is the competence and right to act on the thing in one's charge so as to produce certain results. Domination, or proprietorship, implies controlling or managing a thing in one's own interests. The second premise of stewardship is that the desired outcomes of management are designed to satisfy interests greater than those of the steward alone. Stewardship is a three-way relationship between a manager, a thing being managed, and the principle or party in whose interests the manager acts.

Rationales

Stewardship can have a variety of rationales. These rationales can be classed as instrumental, if management is viewed as a means to an end, or as intrinsic if management is viewed as an end in itself. For example, the guardian of a child can be considered to be a steward of the child if he or she is taking care of the child for the sake of someone besides the guardian himself or herself. The guardian might be doing so for instrumental reasons, such as because his or her community needed more adults. Alternatively, a guardian might be taking care of a child for intrinsic reasons, such as because he or she believed it to be his or her duty to a deceased parent or because raising a child to realize his or her full potential is intrinsically good. The following sections summarize some of the instrumental and intrinsic arguments for the stewardship of biodiversity and some of the counterarguments to the use of stewardship models as the basis for conserving biodiversity.

Instrumental Arguments for Stewardship of Biodiversity

Ecological stewardship has been justified largely on instrumental or practical grounds. George Sessions (as cited in Oelschlaeger, 1992, pp. 90–130) engagingly categorized types of instrumental arguments for preserving wilderness as a "silo" for storing organisms that might have eventual use; as a "lab" for long-term, large-scale natural experiments and baseline data; as a "gym" with spectacular recreational opportunities; as a "cathedral" for the appreciation of beauty or one's natural heritage; as a "minding animals" zone needed for normal human development; as an "antitotalitarian" object lesson in how to be free and make decisions for oneself; and as a "life-support system" to furnish water, nutrients, and the right mix of gases.

Much of the history of conservation can be viewed as a succession of instrumental schools of ecological stewardship, referred to collectively by R. Edward Grumbine as "resourcism." Natural forest management has followed instrumentally motivated stewardship models for at least two centuries. Georg Ludwig Hartig, organizer of the Prussian Forest Service, wrote in 1795 that, "All wise forest management must endeavor to utilize as much as possible, but in such a way that later generations will be able to derive at least as much benefit from them as the present generation claims for itself." This view was embodied in the European concept of a "dauerwald," a permanent or perpetual forest, generally characterized by uneven-age forestry and selective cutting. Gifford Pinchot, head of the U.S. Forest Service at its inception in 1905, called conservation "the one great central problem of the use of the earth for the good of man."

The set of practices referred to as "sustainable development" is based on an instrumental argument for stewardship of ecological systems and natural resources on behalf of all people and their posterity. Herman E. Daly succinctly defines sustainable development as "development without growth." Its objective, as stated in the 1987 "Brundtland Report" of the World Commission on Environment and Development, is "meeting the needs of the present without compromising the ability of future generations to meet their own needs." On a more humble scale, this is like passing on the family farm in better shape then you found it, and one of the most resonant arguments for sustainable development is that is it essential for the welfare of one's children. Author Peter Mattheisen has appealed for stewardship of biodiversity because "indifference to the loss of species is, in effect, indifference to the future, and therefore a shameful carelessness about our children." When Margaret Jacobson and Garth Owen Smith asked Himba pastoralists in northwestern Namibia in the 1980s and 1990s why they wanted to conserve local wildlife, a common reply was so that their children could see them.

Ecosystem management focuses on natural resources management rather than on human development but makes the same argument for environmental stewardship as does sustainable development—that it is beneficial to people in the present and in the future (Johnson *et al.*, 1999). Steven L. Yaffee of the University of Michigan noted that the term "ecosystem management" covers a range of approaches to natural resources management, from "anthropocentric" plans for multiple use of natural resources through "biocentric," whole-ecosystem management to "ecocentric," natural regionwide plans. The first approach was codified as early as 1906 in the Multiple Use Sustained Yield Act of the U.S. Congress under Theodore Roosevelt. Whole-system management was in principle implemented by the National Park Service in 1963 when it accepted the recommendation in the "Leopold Report" of its Advisory Board on Wildlife Management, chaired by Starker Leopold, to manage its lands as "biotic wholes." The statutes that govern the Oregon Board of Forest Lands prescribe a regionwide plan to manage for "greatest permanent value," defined in a 1998 administrative rule as "healthy, productive and sustainable forest ecosystems that over time and across the landscape provide a full range of social, economic, and environmental benefits to the people of Oregon."

There is evidence that people in the United States find instrumental rationales for ecological stewardship more convincing than intrinsic ones. As part of a study funded by the U.S. Forest Service Office of Communication, 30,000 on-line news stories about national forests published from 1992 to 1996 were searched by computer for references to benefits in four categories: "recreational," "commodity," "ecological," and "moral/spiritual/aesthetic." Recreational benefits accounted for about 40% of the citations, commodity benefits for 30%, ecology for 20%, and moral benefits for only about 10%. Recreational benefits tend to translate into local economic benefits and are important not only to people who recreate but also to those who sell to them or tax them. A 1999 report

from the U.S. Department of Agriculture on population and economic growth in nonmetropolitan counties in the contiguous United States from 1970 to 1996 found that growth was strongly dependent on local natural amenities, as measured by an index based on weather, topographic variation, and water surface area. Populations grew by 120% on average in the counties that scored high on the amenity index and only by 1% in the counties that scored low. High-scoring counties also had a threefold higher increase in number of jobs compared to low-scoring ones.

Intrinsic Arguments for Stewardship of Biodiversity

Moral philosophy has provided many intrinsic arguments for the stewardship of species. Concepts of "moral status" have been held to require varying degrees of solicitude for different species depending on their ability to make decisions, be self-aware, feel desire or pleasure or pain, and strive to exist. A being has moral status when it has "moral duties directed to it" by "moral agents," those "able to deliberate and act in a responsible and answerable way" (Wetlesen, 1999). Some philosophers, such as Plato and Kant, assign moral status only to beings with free will, meaning normal humans. Thomas Regan has argued for moral status for beings that experience emotions and have a sense of their own individuality, meaning normal mammals more than 1year old and possibly other animals such as birds. Animal rights advocate Peter Singer has proposed sentience, the ability to have conscious feelings of pleasure or pain, as the qualification for moral status.

In addition to the issue of who or what should be accorded moral status, there is the issue of how equal this status should be. One might have equal moral duties toward all things with moral status or greater duties toward some of these things than toward others. Many ethical philosophers accept that persons have equal moral status. As signatories to the Universal Declaration of Human Rights of the United Nations, many nations accept this as well: "All human beings are born free and equal in dignity and rights. They are endowed with reason and consciousness and should act towards one another in a spirit of brotherhood." Ethical philosophers generally assign nonpersons a lesser moral status. For example, Mary Anne Weaver reasoned that people have a moral responsibility to accept the equal rights of all persons, to avoid cruelty to sentient beings, and to respect all life. Wetlesen (1999) proposed "equal value for moral persons and agents and lesser values for non-persons, in proportion to their similarity with moral persons."

Stewardship of things that have no moral status can be morally justified on the basis of their importance to things that do have moral status. For instance, it could be held to be wrong to shatter a stalactite because it will deprive other people of the enjoyment of seeing it or wrong to divert water from a pool because it will endanger a fish. In most minds, this provides the only intrinsic justification for stewardship of nonliving things. However, thinkers such as Baird Caldicott elaborate reasons for attributing intrinsic values to wholes, which could apply to ecological systems.

In addition to its instrumental schools, the history of conservation contains a tradition of intrinsic arguments for stewardship, documented in the writings of John Muir, Henry David Thoreau, David Brower, and others. In 1864, George Perkins Marsh wrote that "Man has too long forgotten that the earth was given him for usufruct alone, not for consumption." Toward the end of his career in 1949, Aldo Leopold proposed that "A thing is right when it tends to preserve the integrity, stability and beauty of the biotic community. It is wrong when it tends otherwise" and that "The last word in ignorance is the man who says of an animal or plant, 'What good is it?' "Ecologist Charles S. Elton ventured in 1958 that the real reason for conservation is that "animals have a right to exist and be left alone, or at any rate that they should not be persecuted or made extinct as species." Robert H. Nelson summed up how environmentalists who justify biological conservation on moral grounds might view things: "To argue for building a dam would be like arguing for the institution of slavery because it was economically efficient."

Intrinsic arguments for the stewardship of biodiversity have found their way into legal philosophy and animal rights legislation. Roderick Nash (1989) made a case for regarding legal rights for nonhumans as a logical continuation of the progressive enfranchisement of different economic classes, races, and genders of people. Christopher D. Stone raised the question of legal rights for nonanimals in his 1972 essay, "Should Trees Have Standing?" (Southern California Law Review 45(2), republished in 1974 by William Kaufmann, Los Altos). Some natural lands have been set aside for use by wild animals only. Several areas in the Sierra Nevada of California that have bighorn sheep are closed to human entry during most of the year.

Does it make a difference whether the rationale of stewardship of biodiversity is instrumental or intrinsic? Two ways that it might are if one type of rationale is more likely to withstand arguments against conservation or if the two types lead to different ways of managing. It has been widely argued that only instrumental justifications for biological conservation will lead to political action or stand up to conflicting economic pressures. The plant most often cited as an example of why it is important to conserve biodiversity is the rosy periwinkle of Madagascar. Chemicals extracted from this wild plant were discovered to cure childhood leukemia. Others argue that only a belief in the right of species to exist will motivate stewardship of biodiversity since most species have no demonstrable human use or known ecological significance. One interesting possibility is that it takes intrinsic arguments to generate instrumental ones. Officials of the United Nations Commission on Economic Development interviewed in the early 1990s by P. P. Craig and colleagues confided that although they made only instrumental arguments for environmental conservation in their professional work, they privately believed that the real arguments were intrinsic.

Counterarguments against Stewardship as a Model for Conserving Biodiversity

Without necessarily rejecting the idea that it is important to conserve biodiversity, many groups and individuals have rejected the concept of stewardship of biodiversity. Objections have been made to each of the two basic premises of

stewardship: that management of biodiversity is needed and justified and that people should care for biodiversity on behalf of interests greater than their own.

One objection to the idea of managing biodiversity is that it is self-defeating because the essence of the biological world is its wildness. In an essay on "the etiquette of freedom," Gary Snyder (as cited in Oelschlaeger, 1992, pp. 21–39) contrasts the meanings of "natural" and "wild." He argues that managed things may be natural but not wild and that wildness is the world itself. Paul Taylor distinguished the realm of "bioculture," in which husbandry and agriculture may take place, from that of wilderness, in which "noninterference ethics" should prevail. George Sessions (as cited in Oelschlaeger, 1992, pp. 90–130) rejects the model of stewardship for wilderness and admits only that "perhaps some ecologically enlightened version of the 'stewardship' model is appropriate for the bioculture."

A second objection to the idea of managing biodiversity is that it is foolhardy because we do not know enough about wild species and ecological systems to know which actions to take. This objection is supported by past management errors. Human suppression of fires in western North American parks and national forests in which natural fires occur regularly has inhibited reproduction of plant species and reduced the diversity of some forests. Alston Chase argued that ecosystem management has altered wildlife processes in Yellowstone National Park.

These objections to the concept of stewardship of biodiversity can be viewed as cautions about what stewards should do rather than reasons to abandon stewardship altogether. Those who argue against managing wildness do not necessarily preclude protecting wilderness areas from development. Those who argue that humans need more knowledge to manage ecological systems successfully may also argue that the best course is to try to acquire this knowledge through "adaptive management," in which the results from each management prescription are analyzed and used to plan the next. Jack Ward Thomas said that "an ecosystem is not more complicated than we think, it is more complicated than we can think." However, as then head of the U.S. Forest Service, he probably did not mean by this that ecosystem management should be abolished but rather approached with humility. Richard L. Knight has likewise proposed that an appreciation of the complexity of nature and a willingness to learn from mistakes and modify techniques are essential components of good stewardship of biodiversity.

A counterobjection to the notion that human management of biodiversity is foolhardy is that humans now have no choice but to manage natural systems because we already dominate them, foolishly or not. Globally, humans are said to consume two-fifths of the earth's photosynthetic productivity and to have doubled the rate at which nitrogen is being supplied to plants (Vitousek *et al.*, 1997). On some islands and some continents, human influences have been judged to now determine the distributions of species and the risks of extinction of animal populations more than natural factors. Many remaining areas of habitat are too small to maintain their biodiversity naturally. In southern Africa, where elephant populations have been largely restricted to movement within reserves, elephants can multiply to the point at which they deforest large areas, leading managers to kill elephants to conserve trees. In other systems, humans have eliminated key elements and may have to replace their actions artificially. Short periods of grazing by livestock have been found to promote biodiversity in interior North American grasslands, possibly because the grazing simulates the effects of the great herds of buffalo that formerly wandered the grasslands.

A second counterobjection to the charge of foolhardiness is that humans are an integral part of some ecological systems. Aboriginal burning of savanna in what is now Kakadu National Park in Australia may have been part of the fire regime for 50,000 years. This is clearly long enough for native species to have evolved in response to human-caused fire frequency.

A third objection to human management of biodiversity is that we have no right to be managers, that we stand in relation to other species, not as stewards, but as siblings. The native American writer Black Elk represented his people's belief that "We are of Earth and belong to You. Every step that we take upon You should be done in a sacred manner." In his book *The Green Fuse*, biologist John Harte explicitly rejected the idea of humans as stewards of ecological systems; it is more accurate, he said, to think of nature as the steward of humanity and to call upon humans to be stewards of their own impulses. "Self-Stewardship" is the title of the chapter that deals with stewardship in Vice President Al Gore's book *Earth in the Balance*.

The second premise for the stewardship of biodiversity is that persons should manage it for the sake of interests greater than their own. Many people reject this in favor of the premise that humans rightfully dominate nature and are the rulers and not the stewards of biodiversity. Captain Vancouver is said to have remarked upon seeing the virgin temperate rainforests of his eponymous island that here were enough masts to supply the British Navy forever. The belief that man should rule over nature figures prominently in Judeo-Christian tradition. The Book of Genesis in the Bible records that humans shall "have dominion over the fish of the sea, and over the fowl of the air, and over every living thing that moveth upon the earth." Confusingly, some recent Christian models called stewardship models still portray humans as dominant over the earth (Sessions as cited in Oelschlaeger, 1992). It should be noted that individual Judeo-Christian theologists have long argued the opposite—that humans do not have dominion over other living things. The medieval Jewish philosopher Moses Maimonides wrote, "It should not be believed that all the beings exist for humanity. On the contrary, all the other beings too have been intended for their own sake and not for the sake of something else." According to the contemporary Catholic theologian Thomas Berry, "The earth belongs to itself and to all the component members of its community." In *The Ecology of Eden*, Evan Eisenberg traces these contradictory theological viewpoints back to the biblical injunction to Adam and Eve to both work and protect the Garden of Eden, that is, to have dominion over the organisms but safeguard the divine creation.

A recent legal embodiment of the belief that people can own other life is the extension of intellectual property rights to cover species or aspects of the genetics or metabolism of a species. This principle, upheld by U.S. courts and by international governmental trade organizations, has been touted as an essential economic incentive for the development of biological

and medical technology and as a way to ensure that developing countries and local residents derive benefits from biodiversity. In 1971, the first patent on a genetically engineered organism was granted in the United States to General Electric and its employee, Anand Mohan Chakravarty, for a *Pseudomonas* bacterium into which plasmids from three other bacterial strains had been introduced. The General Agreement on Trade and Tariffs that governs much of world trade contains a Trade Related Intellectual Property Rights agreement that can cover wild species and their products. Even aspects of human biodiversity can become private property. The government of Iceland has debated transferring the intellectual rights to a national database of citizens' genetic information to a corporation.

Throughout history, many societies have held that humans can own individual living things as property, including other humans. Real property rights often include the rights to plants, animals, and minerals on the property. However, it is also usual to recognize that there are multiple interests in real property, that public and private interests are commingled on almost all lands, and that property owners have responsibilities to avoid using their property in ways that harm their neighbors or society. The interests of landowners in the natural resources on their property can be legally balanced against the recreational and environmental interests of the public. On a more moral plane, Aldo Leopold wrote in "The Ecological Conscience" in 1947 that "when a farmer owns a rarity he should feel some obligation as its custodian and a community should feel some obligation to help him carry the economic cost of custodianship."

A final counterargument against stewardship as a model for the conservation of biodiversity is that the most effective basis for conservation is love, which does not involve acting in interests greater than one's own. To explain this argument, it is useful to return to the example of a guardian and a child. Instead of taking care of a child to produce an adult or out of duty or principle, a guardian might do so out of love for the child. Love is neither an instrumental nor an intrinsic rationale; it is not rational and involves no third party or principle in whose interests one acts. Just as many will argue that love of a child leads to the best child care, Wendell Berry has proposed that "love of place" leads people to care for species and habitats. The land ethic propounded by Aldo Leopold (1966) also contains an element of love: "When we see land as a community to which we belong, we may begin to use it with love and respect" and "affection based on utility alone leads to the same pitfalls and contradictions in land as in people."

One practical implication of this view is that the best way to promote the conservation of biodiversity is to take people to see it. One cannot love without at least first sight. In southern Africa, many programs are dedicated to taking children to see wildlife. Courses for "decision makers" run by the Organization of Tropical Studies in Costa Rica discuss arguments for conservation but also take their prominent students to watch sea turtles lay their eggs in the moonlight on a tropical beach.

Contexts

The concept of stewardship has been applied to the conservation of biodiversity in a multitude of forms and with varying success. This is partly due to the ecological, economic, cultural, political, and historic contexts in which stewardship of biodiversity has been undertaken. These contexts include the mobility of animal populations, disequilibrium in ecological systems, economic incentives and disincentives for conservation, poverty and industrialization, the scale of trade, traditional views of the place of humans in nature, legal systems of land and resource tenure, and the displacement or migration of peoples. Some of these contexts for stewardship are briefly discussed in the following sections.

Ecological Factors

Stewardship requires that the scale of management match the scale of the thing being managed. In the phrase of Kai Lee, there is often a "mismatch of scales" between ecological processes and human societies. For example, many species and ecological processes cross the jurisdictional boundaries that constrain managers. More than 100 species of Neotropical migratory birds breed in North America and winter in Central or South America. Declines in North American breeding populations have been associated with forest losses in Central and South America. Marine organisms with planktonic larvae often disperse widely on ocean currents and depend on long-distance dispersal to maintain local populations. The establishment of new juveniles in marine reserves may depend almost entirely on dispersal from outside reserves. Aerial nitrogen deposition, which occurs on a global scale (Vitousek *et al.*, 1997), is thought to be responsible for replacement of heathland by grassland in The Netherlands. Due to the "openness" of ecological systems, reserves may be, as Mark H. Carr and colleagues suggest, "necessary but not sufficient for conservation."

Ecological fluxes and connectedness complicate the notions of ownership and management. Many species and the resources on which they depend fall into the category of "common pool resources," which have been defined as "those from which multiple independent users cannot be excluded and in which consumption by one user detracts from another's consumption." These include species outside national boundaries, as in international waters, and species on communally owned lands. Models for the stewardship of biodiversity in commons range from international legal agreements such as the United Nations Convention on Marine Law, which assigns states bordering oceans and seas specific rights and responsibilities for their use, to the customary land tenure systems of African pastoralists.

Land stewardship may be more difficult in habitats in which resources are relatively scarce or where natural conditions are unpredictable or rapidly changing. E. Steen observed that cultures in the relatively dry habitats around the Mediterranean have tended to exhaust soil fertility except where there are reliable pulses of water with low salt and high alluvial content, such as along the Nile. Frequent erosion and deposition of soil maintain plant species diversity in highly dynamic habitats, for example, along rivers and on coastal sand dunes, making it necessary to manage for land instability. The prevailing view of ecological systems has shifted from a notion of dynamic equilibrium, in which systems are expected

to tend to reach and return to one steady state, to one of disequilibrium, in which systems are expected to change directionally over time. This makes biodiversity a moving target for management. Some ecological systems switch between alternative states: They may show a nonlinear response to human use, remaining relatively stable until use reaches a certain intensity or cumulative level and then change relatively rapidly, or they may not change back even when use discontinues. This makes it difficult for managers to learn from experience. In habitats in which rare events strongly affect the distribution and abundance of species, managers may need to be "stewards of catastrophe," preserving or mimicking the potential for major disturbance.

Economic Factors

Among the many economic factors that impinge on stewardship are disincentives and incentives for conservation, degree of industrialization, and trade across ecological boundaries. When land is taxed based on its most lucrative potential use instead of its current use, landholders are more likely to sell or develop their land. Many laws, such as the Williamson Act in California, have been passed to remove this economic disincentive for farmers. Often, farmers agree not to develop their land in exchange for current use valuation of the land for tax purposes. Laws have also been passed to create economic incentives for stewardship of biodiversity. One widely used instrument is the conservation easement, an agreement under which a landowner agrees not to use his or her lands in certain ways in exchange for relief from taxation. The Federal Agriculture Improvement and Reform Act of 1996 created or redirected a set of U.S. Department of Agriculture programs that provide economic encouragement to private landowners to protect wildlife habitats. The Wildlife Habitat Incentives Program provides cost-sharing for habitat improvement, the Conservation Reserve Program pays farmers to fallow cropland under a vegetation cover for a decade, the Wetlands Reserve Program provides conservation easements and cost-sharing for restoration and protection of wetlands for 30 years or more, and the Environmental Quality Incentive Program combines education and technical and financial assistance to help landowners adopt practices that reduce environmental problems.

Economic conditions in developing countries have been linked to poor stewardship of biodiversity in several ways. Poor people in rural areas are likely to depend heavily on local natural resources, and a large proportion of land area is likely to be settled by farmers. Over time and with population growth, use tends to become unsustainable. Governmental agencies are likely to lack the resources to enforce protection of species in nature reserves. Wood and charcoal are a main source of fuel in cities, leading to deforestation for hundreds of kilometers into the countryside, as around Kinshasa in Congo. On the other hand, it is by no means clear that economic development improves stewardship. In central Africa, "de facto" conservation of biodiverse areas by their remoteness and local economic poverty has probably been as effective as intentional stewardship in the region. One hypothesis is that development will tend to improve stewardship of biodiversity where subsistence resource use is the main cause of its loss, as is probably the case for forests in Madagascar, but not where commercial exploitation is the main cause, as in the forests of Cameroon.

The differences between economic conditions in developing and industrialized countries are often held to require different approaches to stewardship. Conservationists in developing countries largely argue that local rural communities must receive economic benefits from protected areas. This is the basis for the spread of integrated conservation and development projects, which generally link biological conservation within a protected area to human development in the surrounding communities (Wells and Brandon, 1997). These projects are mainly limited to developing countries.

The globalization of commerce has to some extent reversed the mismatch of scales between ecology and society. It is increasingly the case that trade occurs across ecological boundaries. Global trade is often held to threaten stewardship of biodiversity because it makes it more likely that parties using or profiting from resources will live outside the region from which the resources come, and that they will therefore be less concerned about stewardship or constrained by environmental regulations. This is one basis for "bioregionalism," advocated by Sale Kirkpatrick and others, which counsels that people use only the natural resources from within their ecological region.

Cultural and Political Factors

Human cultures probably differ in their attitude toward stewardship. For example, D. R. Given argued that the Maori have an environmental ethic that emphasizes guardianship and stewardship, in contrast to a European tradition of dominance over nature. He predicted that laws giving the Maori greater control over resource management would help develop a multicultural, ecocentric biodiversity ethic in New Zealand. Others contend that economic factors tend to supersede cultural traditions. Rodrigo Sierra found no recent differences in impacts between three indigenous and three nonindigenous populations in northwest Ecuador, although there had been differences in the past when economic conditions were different. He concluded that the degree of impact of different peoples on natural systems in this region was primarily associated with economic conditions.

Cultural and political factors can result in "incidental stewardship" of biodiversity, achieved as a by-product of management for other purposes. Some of the most effective stewardship of natural habitat in both the United States and developing countries has been incidental. Lands reserved by the U.S. Army for military bases have become Nature Conservancy preserves and parts of national wildlife refuges and National Park Service units. In the western United States, the military was politically able to completely exclude resource extraction over large tracts of land, some of which now constitute the most pristine examples of semiarid habitats in the region. In some areas of Africa and Asia, small groves held sacred and reserved for religious purposes are among the least disturbed forests.

Local political authority over resource and land use has been associated with good stewardship. Al Gore has proposed that "freedom is a necessary condition for an effective stewardship of the environment" and called to witness the environmental degradation in eastern Europe and hazardous waste sites in poor communities with relatively little political power. Secure land tenure is widely regarded as a precondition for effective land stewardship by individuals and communities. In African countries, governmental ownership of wildlife on private or communal land has been blamed for lack of interest in wildlife conservation.

However, a given political arrangement may have opposite effects on stewardship in different economic contexts. The establishment of wildlife management agencies or funds as "parastatals"—governmental bodies that are economically independent of the rest of the government—has promoted wildlife conservation in Kenya and Zambia. In these countries, tourism in some parks and game management areas was generating more revenues for the central treasury than the treasury was providing to the parks largely because the tourists were almost all foreign and could afford to pay large amounts relative to the local economies. Once given political authority to raise and disburse its own revenues, the Kenya Wildlife Service was able to retain gate receipts at Amboseli National Park and share them with the Maasai around the park. During the wet season, the large ungulates in the Amboseli system rely largely on the Maasai lands for grazing, making the cooperation of the Maasai essential for conservation of biodiversity in the park. In Zambia, governmental revenues from European and North American sports hunters were channeled to the park and wildlife service for management, providing hundreds of jobs for local residents in game management areas. In contrast, requiring parks to be economically self-sufficient in the United States, in which the national park service receives more governmental revenues than it generates, would probably lead to a decrease in resources for park management.

Types of Stewards

Almost every possible type of social or political group has undertaken, or purported to undertake, the management of species and habitats on behalf of interests greater than those of the group. The possible stewards of biodiversity include governmental agencies, nongovernmental organizations (NGOs), corporations, local communities, and private individuals. Governmental agencies and NGOs include international, national, and local bodies. They have worked separately and in partnerships of different types of groups. Through legal and political action, classes of citizens have also acted as stewards. The following sections provide examples of these types of stewards and discuss some of their apparent strengths and weaknesses: Are certain types of stewards more effective or trustworthy as guardians of biodiversity?

Governmental Stewards

International governmental agencies exercise stewardship of nature in the form of conventions, declarations, and moral and financial support for conservation efforts. Through trade agreements and support for development, international bodies have both negative and positive effects on biodiversity. International governmental bodies are uniquely able to enforce stewardship of global ecology such as upper atmospheric conditions and of common pool resources such as marine fisheries and the Antarctic continent. They can finance and lobby for the protection of local sites that are internationally prized.

The most visible international governmental steward of biodiversity has been the United Nations (UN). In 1982, its General Assembly adopted a Charter for Nature that recognized an intrinsic basis for the stewardship of biodiversity:

> Every form of life is unique, warranting respect regardless of its worth to man, and to accord other organisms such recognition, man must be guided by a moral code of action. … Nature shall be respected and its essential processes shall not be disrupted.

The vote was 114 for, 17 abstaining, and 1 (the United States), opposed. UNESCO, an organization within the UN, designates areas of special cultural and conservation interest as Man and the Biosphere Reserves and World Heritage Sites. Most Biosphere Reserves center on existing protected areas, and designation as a reserve is largely symbolic. However, it may have an important educational value by affirming the preciousness of natural places and help inspire support from other sources. The National Natural Landmark program of the U.S. National Park Service is an example of a similar program on a national scale. The United Nations Environmental Programme directly funds conservation efforts and projects.

The "Montreal Protocol" is often cited as a successful example of international governmental cooperation to protect the environment. This agreement banned most release of chlorofluorocarbons into the air. These chemicals can lower the concentration of ozone in the upper atmosphere and increase the amount of cell-damaging ultraviolet radiation that reaches the earth's surface. Important scientific analysis to support this political agreement was supplied by an international nongovernmental body, the Scientific Committee on Problems of the Environment of the International Council of Scientific Unions. This is an association of professional scientific associations such as the American Association for the Advancement of Science.

International governmental trade and financial institutions such as the World Bank do not have ecological stewardship as their primary goal. The World Bank has supported projects agreed to have been environmentally harmful, notably the Aswan Dam in Egypt. However, World Bank policies formulated during the 1980s explicitly called for conservation of tropical forests. The bank set up an internal environmental review process for its projects and funded the Global Environmental Facility, which gave unusually large amounts of money to conservation projects throughout the world. Some international trade agreements, such as the Convention on International Trade in Endangered Species, have been directed at conservation of biodiversity. On the other hand, international trade agreements such as the North American Free Trade Agreement and organizations such as the World Trade

Organization have been severely criticized for failing to include adequate environmental protections.

Although dedicated primarily to promoting human development, national overseas aid agencies such as the Department for International Development of the United Kingdom or the Gesellschaft für Technische Zusammenarbeiten of Germany have been funding conservation efforts for more than a decade. In the mid-1980s, the U.S. Agency for International Development (USAID) received a specific mandate from the U.S. Congress to conserve biodiversity and tropical forests. As stewards, foreign aid agencies have the advantages of considerable independence from local pressures and enormous buying power in foreign countries. Some of their disadvantages are political inability to provide long-term support for projects or to manage them directly. For example, USAID funds rather than manages projects and expects them to become self-sufficient within 5–10 years.

National governmental agencies that manage natural resources include park, wildlife, forest, fisheries, land management, environmental, and tourism ministries and services. These agencies are stewards inasmuch as they manage the resources on behalf of groups outside the agency. The U.S. Fish and Wildlife Service refers to the National Wildlife Refuges it administers as "stewardship lands." However, the primary charge of these agencies is often to maximize the economic returns from resource use rather than to conserve biodiversity. Georgia M. Mace characterized the role of natural resource managers in governmental agencies as generally aiming to maximize continuing yields and collect extensive data on a few species rather than assessing management risks to a wide diversity of species. This would tend to make them poor stewards of biodiversity. For example, maximum productivity is generally negatively associated with maximum species richness. On the other hand, governmental agencies have significant advantages as land stewards. They often command relatively extensive resources and manage a large proportion of the natural lands in a country, including most of the large natural areas. Depending on how their professional incentives are structured, managers within agencies may be largely free from personal conflicts of interest with conservation.

Many governmental resource agencies emphasize resource use over biodiversity conservation for both economic and political reasons. In many countries, including the United States, the operating budgets of national forests are tied to the fees paid by companies for logging concessions. This provides a powerful incentive to manage forests for wood production and to focus on a few economically valuable species. The charter of the U.S. National Park Service calls upon it to protect parks for the enjoyment of visitors. Parks experience pressure to manage for maximum continuing yields of visitors and therefore to develop lodging and shops and to focus on species that are appealing and easy to see, the "charismatic megafauna."

Research institutions and research departments within resource management agencies operate natural reserves throughout the world. The approximately 30 reserves of the University of California comprise the largest system of any single university. The U.S. Forest Service has designated hundreds of Natural Research Areas, often prime examples of distinctive forest types, that are open to research use only. One goal of the sites funded under the Long-Term Ecological Research program of the U.S. National Science Foundation is to provide baseline data on ecological function and community structure in diverse habitat types that are as little disturbed by humans as possible. An advantage of research reserves is that they are generally completely protected from consumptive use and largely protected from tourism. However, the reserves are mostly small and subject to experimental manipulations such as animal exclosures and vegetation removal and nutrient addition in plots.

Nongovernmental Organizations and Communities as Stewards

Conservation land trusts, NGOs that purchase land or the development rights to land for the express purpose of maintaining it in an undeveloped state, are likely to have the conservation of species or habitats as their primary goal. Another advantage of land trusts relative to governmental agencies seems to be that they are able to act quickly to acquire habitat in immediate threat of development. The Nature Conservancy is the largest private land trust for natural lands in the United States, with more than 10 million acres of reserves throughout the country. One of its mottos, "the last of the least and the best of the rest," epitomizes its dual objectives to prevent the extinction of rare species and to conserve whole natural communities and ecological systems. The first objective requires a large number of reserves, some of which may be very small, for example, if the species of concern is a grass. The second requires large reserves, which have to be fewer in number for reasons of cost and availability. Local land trusts have also played important stewardship roles. Starting in the early 1900s, the Save-the-Redwoods League purchased a significant proportion of the remaining old-growth forests of the coast redwood, *Sequoia sempervirens*, which the trust eventually deeded to the government for parks. A disadvantage of land trusts as stewards is that they are rarely able to purchase large areas and generally control only a small proportion of the natural lands in an area. According to the Land Trust Alliance, local land trusts owned about 0.002% of land in the United States in 1999.

A highly controversial type of steward of biodiversity is the local community, the set of human residents in a place. "Community-based conservation" is an important option in societies in which a large proportion of lands are owned or customarily used by communities as a group. It is viewed as an alternative both to the establishment of protected areas that exclude consumptive uses of natural resources by local residents and to natural resources management by governmental agencies. During the 1980s and 1990s, many projects were established to encourage local communities to assume or resume the major responsibility for local environmental stewardship, especially in the developing countries of Africa and Asia. Sets of these projects have been reported on in books edited by Western *et al.* (1994), Mark Poffenburger, and others and in networking organs such as the Forest, Trees and People Newsletter (http://www-trees.slu.se). Two of the best known examples in Africa, the CAMPFIRE project in Zimbabwe and the ADMADE project in Zambia, have given varying degrees of

authority for local wildlife management to communities. Community-based forest management has become particularly widespread in India. In the late 1990s, the Van Gujjars, a forest-dwelling pastoral tribe in Uttar Pradesh, proposed to resolve their objections to the establishment of the proposed Rajaji National Park by taking responsibility for forest management in the park.

Reasons cited for turning to community-based conservation are that it is morally right, that local people have expert, "indigenous knowledge" of local species and habitats, that local people have lived in harmony with local nature for centuries, that traditional indigenous populations use a wide variety of wild species and thus have a practical incentive to conserve biodiversity, and that management by governmental agencies or in nature preserves in the area has failed to prevent habitat degradation. The Dayak groups of central Kalimantan in Borneo are reported by W. de Jong and others to provide successful examples of ethnoconservation; the Dayaks have manipulated the forest to increase productivity of species they use without reducing species diversity or simplifying structure.

One of the main reasons for doubting the efficacy of communities as stewards is the world history of environmental degradation. O. M'Hirit concluded that peoples around the Mediterranean have generally harmed their local forests and that only in the twentieth century has there been effective forest conservation in the regions. In 1957, A. Starker Leopold took a dark view:

> It is surprising that in the long history of man's conquest of the earth there is no evidence of sustained effort on the part of any people to preserve native landscape for its own sake, until our national park system began to take form late in the nineteenth century.

The consensus at the end of the 1990s was that local communities sometimes do and sometimes do not make effective stewards of biodiversity. Communities of recent immigrants in tropical rainforest tend to cause extensive deforestation because they want cleared land for crops and livestock. Even long-settled, traditional societies do not always use sustainable practices; a recent study concluded that hunters from the Piro people in the Peruvian Amazon make no attempt to conserve species vulnerable to overhunting. Practices may be sustainable only as long as human population density remains low and may change when economic and social conditions change. In a 1999 essay on community-based conservation projects in Africa, J. D. Hackel concluded that they tended to lack clear rules for the protection of local ecology and plans for what to do if conservation goals were not met. In both developing and industrialized countries, local communities may have a limited capacity to ask key questions about conservation values or to provide potential answers and limited time or will to bring issues before the public and decision makers. Others believe that community-based conservation has rarely been properly tested because governments have failed to devolve the authority over resources to communities at the same time that they have decentralized the responsibilities for resource conservation.

A different definition of a community is the set of people that have interests in a thing. In this sense, class actions by citizen groups and lobbying and public education by conservation organizations on behalf of their members are also forms of community-based conservation. The political and legal systems in the United States provide a direct route for stewardship by citizen groups through referenda and class action suits. An example of stewardship by class action is the work in the 1990s of the Mono Lake Committee in California to prevent Los Angeles County from diverting water from the streams that flow into this large desert lake. Diversion threatened, among other things, to allow coyotes to walk out to the nesting sites of birds on islands in the lake. The Sierra Club is credited with having dissuaded the U.S. federal government from creating a reservoir in the Grand Canyon in Arizona. The club's public advertising campaign asked "Would you flood the Sistine Chapel to get a better view of the ceiling?" Duck hunters are said to be largely responsible for motivating governmental protection of wetlands, which are now, under the RAMSAR convention, the first globally protected habitat type.

Such "community-of-interest-based conservation" can have the advantage of a longer economic view but the disadvantage of pitching legal and sometimes corporeal battles. The controversy over protection of old-growth Douglas fir forests in the Pacific Northwest, in the name of the northern spotted owl, pitted local logging interests against national conservation ones. According to one estimate, the protection plan that was adopted should result in economic gains of about $110 billion versus losses of only $32.5 billion. However, the losses are concentrated within the region, whereas the gains are spread throughout the country.

To avoid such battles, a broader community-based approach to conservation has been advocated based on the notion of "stakeholder rights." Stakeholders are all parties with interests of any sort in a given thing, be they economic, political, cultural, moral, or personal. One formulation of this "community-of-different-interests-based conservation" is to provide a forum for stakeholders to produce a common management plan for an area. For instance, the town library in Quincy, California, became a neutral meeting place for local environmentalists and loggers, who produced a plan for an adjacent area of national forest. To the surprise of many outside the "Quincy Library Group," the plan called for relatively intensive harvests, and its adoption by the Forest Service was challenged in court. Truman Young argued that many major conservation successes of recent decades, such as the international bans on whaling and ivory, have essentially rejected current economic stakeholder rights.

Private Companies and Individuals as Stewards

Private companies have also been engaged as stewards. One strategy has been to set up certification programs for commercial products or services. Companies or their taxes pay an independent body to certify that a product has been produced in an environmentally sound way. The expected benefit to the company is that environmentally concerned consumers will buy the product. A well-known certification program is run by the Forest Stewardship Council, an international NGO founded in 1993 "to support environmentally appropriate, socially

beneficial, and economically viable management of the world's forests." The council accredits and monitors certification bodies for forest products. Certified products may carry the council's logo. A newer, Marine Stewardship Council aims to perform an analogous function for marine fisheries.

A second strategy has been to encourage private companies to support in-house conservation projects. The nonprofit Wildlife Habitat Council coordinates a program that enlists corporations to voluntarily manage for wildlife and biodiversity protection. Participants account for about 120,000 ha in the United States. In a survey of benefits to corporations at 164 sites, as reported by H. Cardskadden and D. J. Lober, most respondents believed that the program had improved employee morale (95% of respondents) and relations with environmental groups (72%) and the local community (64%). Half of the respondents reported that programs had resulted in annual cost savings, such as through waste reduction. In the absence of scientific monitoring, there was no indication of whether there had been any benefits to wildlife or biodiversity. This is a common problem in many nongovernmental stewardship programs.

By far the most common current use of stewardship in relation to conservation is for programs in which individual private landowners serve as stewards on their own lands, often with economic incentives, technical assistance, or just moral support provided by government or NGOs. The Saskatchewan Wetland Conservation Corporation provides land management advice and a certificate of appreciation to owners of native prairies who make a Voluntary Stewardship Agreement to protect their land from development. The Massachusetts Forest Stewardship Program, run by the state government, shares the costs of habitat improvement for wildlife and rare native fish with private owners of nonindustrial woodlands of 10–1000 acres. The Georgia Endangered Plant Stewardship Network of the State Botanical Garden of Georgia trains school classes to propagate rare and threatened wild plants. The Pesticide Environmental Stewardship Program of the U.S. Environmental Protection Agency helps pesticide users to write and implement strategies for pesticide risk reduction.

One reason for appealing to private landowners to serve as stewards of their property in the United States is that development of private land is a significant cause of species decline. In 1998, the Wilderness Society estimated that conversion of private land for commercial and residential development was at least partly responsible for population losses in 35% of the 1880 imperiled plants and animals in the United States. According to Richard L. Knight, "unchecked conversion of U.S. private, open lands to human-dominated development is causing a simplification of our native biodiversity." In some regions, such as northern New England in the United States, most of the lands with natural or seminatural vegetation are privately owned.

An objection to relying on private landowners as stewards is that private landholdings are generally small compared to public ones and cannot accommodate large-scale ecological needs. Remote imaging of land use has documented that habitat patches may tend to be smaller on private than on public lands. This objective can in some cases be partly met by cooperation between landholders. In about 1986, local landowners in southern Vermont formed the Newfane Wildlife Habitat Improvement Group to plan together for wildlife management. In 1999, the group included more than 50 properties and 7000 acres in three towns. However, the total amount of land in the United States under the protection of voluntary agreements by private landowners is also small—just 0.02% of total land according to a 1999 article by Keith Wiebe and colleagues.

As the previous examples of stewardship on private lands suggest, stewardship is often undertaken through partnerships between public agencies and private parties. This is just one sort of "cross-boundary stewardship"—coordinated land management by multiple stewards, often of different types (Knight and Landres, 1998). Many integrated conservation and development projects are organized as partnerships among governmental agencies, NGOs, and local communities (Wells and Brandon, 1997).

A major advantage of cross-boundary stewardship is that is allows for land use to be "not separate but not equal," to have a diversified portfolio of stewardship types arranged side by side within ecological regions. Arguments for zoning as an alternative to multiple use are not new (Sessions as cited in Oelschlaeger, 1992, pp. 90–130). In 1971, Eugene Odum and John Phillips proposed a landscape of urban-industrial, production, and protection zones. The Biosphere Reserve concept promoted by UNESCO since 1980 envisions protected areas surrounded by an inner zone of human use unlikely to impact the protected core and an outer zone of unrestricted use. Some species do require protected areas in which there is virtually no human impact. A recent analysis of conservation versus human land use by Kent Redford and B. Richter indicated that only extremely limited and largely nonextractive use will protect all components of biodiversity. However, as Sara Vickerman argued, biodiversity cannot be conserved on reserves and with regulation alone (Defenders of Wildlife, 1998). A managed landscape can support important elements of biodiversity while meeting human needs. Significant components of biodiversity can and must be under the stewardship of private landowners.

The relative effectiveness of different types of stewards in conserving biodiversity seems to increase with the total amounts and individual sizes of the land areas they manage, the economic and technical resources they command, their independence from conflicting economic pressures and political constraints, the degree of authority they can exercise, and how close their primary objective is to conserving biodiversity. No one type of steward has greatest effectiveness in all these regards. Since it is also true in many landscapes that no one type of steward is responsible for enough land to ensure the conservation of biodiversity, it appears that a mix of stewardship types, arranged side by side in a landscape, is necessary and desirable.

See also: Commons, Concept and Theory of. Ecosystem Services. Ethical Issues in Biodiversity Protection. Historical Awareness of Biodiversity. Literary Perspectives on Biodiversity. Religious Traditions and Biodiversity. Social and Cultural Factors. Sustainability and Biodiversity. Wildlife Management

References

Defenders of Wildlife (1998) *National Stewardship Incentives: Conservation Strategies for U.S. Landowners*. Washington, D.C.: Defenders of Wildlife.

Johnson NC, Malk AJ, Sexton WT, and Szaro RC (eds.) (1999) *Ecological Stewardship: A Common Reference for Ecosystem Management*. New York: Elsevier.

Knight R and Landres P (eds.) (1998) *Stewardship Across Boundaries*. Washington, D.C.: Island Press.

Leopold A (1966) *A Sand County Almanac with Essays on Conservation from Round River*. New York: Oxford Univ. Press.

Nash RF (1989) *The Rights of Nature: A History of Environmental Ethics*. Madison: Univ. of Wisconsin Press.

Oelschlaeger M (ed.) (1992) *The Wilderness Condition: Essays on Environment and Civilization*. San Francisco: Sierra Club Books.

Vitousek PK, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of earth's ecosystems. *Science* 277: 494–499.

Wells M and Brandon K (1997) *People and Parks: Linking Protected Area Management with Local Communities*. Washington, D.C.: World Bank, World Wildlife Fund, and Agency for International Development.

Western D, Wright M, and Strum SC (eds.) (1994) *Natural Connections: Perspectives in Community-Based Conservation*. Washington, D.C.: Island Press.

Wetlesen J (1999) The moral status of beings who are not persons: A casuistic argument. *Environ. Values* 8: 287–323.

Storage, Ecology of

Caroline M Pond, The Open University, Milton Keynes, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Adipocytes, brown Specialized thermogenic cells unique to mammals and probably derived from skeletal muscle that contain numerous mitochondria and small droplets of triacylglycerols. Under neural and endocrine control, brown adipocytes can switch rapidly from inactivity to the highest known metabolic rate, generating heat from oxidation of fatty acids and/or glucose in mitochondria uncoupled from ATP synthesis by uncoupling protein 1. Thermogenesis can be very high in neonates and during emergence from hibernation or torpor.

Adipocytes, white Large cells unique to vertebrates that store triacylglycerols, take up and release fatty acids, produce various signal molecules and receptors that enable them to respond to circulating and locally acting agonists.

Adipose tissue, white Storage tissue unique to vertebrates and best developed in tetrapods that consist almost entirely of expandable adipocytes (q.v.) that contain a large droplet of triacylglycerols, plus vascular and neural tissue, and often contiguous or interchelated lymphoid tissues.

Blubber Specialized superficial adipose tissue found only in pinnipeds (seals), cetaceans (whales and dolphins), and sirenians (sea-cows) that serves as thermal insulation as well as for energy storage.

Fat body Storage tissue of arthropods that stores lipids and has endocrine and immunological functions.

Fatty acid Any of hundreds of different aliphatic hydrocarbons with an acid group. The main differences are the number and position of one or more double bonds, and the presence of substitutions and side-chains. Most fatty acids are incorporated into larger molecules with ester bonds.

Glycogen An insoluble polymer of glucose. The main storage carbohydrate in animals.

Polyunsaturated fatty acid Fatty acid with two or more double bonds in the chain of carbon atoms. They occur in distinct families defined by the positions of the double bonds, of which the commonest are *n*-3 and *n*-6.

Saturated fatty acid A fatty acid with the maximum complement of hydrogen atoms and no double bonds between carbon atoms.

Triacylglycerol Ester of glycerol and three fatty acids. The commonest lipid storage molecule in animals and green plants.

Introduction: What can be Stored?

Few organisms have continuous access to nutrients derived from feeding or generated by photosynthesis, so some storage is necessary. Many organisms have tissues, cells or molecules adapted to this role.

Specialized storage molecules are insoluble, compact, and nontoxic to the holding cells, even in large quantities. Thus oxygen, which is weakly soluble in water and can be toxic at high concentrations, is stored in substantial quantities only when bound to special pigments in blood or muscle. Vertebrates and some invertebrates and microorganisms can hold oxygen bound to the iron-containing protein, hemoglobin in body fluids. Vertebrate cardiac and skeletal muscles contain the closely related oxygen-binding molecule, myoglobin. However, even specialized diving animals such as seals and whales, whose muscles are dark red with abundant myoglobin, can store enough oxygen to support strenuous activity for about an hour at most. Stores held by sedentary animals are used more slowly so may last longer, but storage capacity for oxygen is still very minor compared to that of other nutrients.

Many vitamins and essential minerals, including iron, are too toxic to be stored in more than small quantities, and then only when bound to specialized proteins, usually in the liver. Calcified skeletons can serve as stores of calcium and phosphate in many arthropods and vertebrates. In female mammals, bone is laid down before breeding and withdrawn during lactation, as the minerals pass to the offspring in the milk. However, the quantities so stored are limited by the extra weight of dense minerals and the need to maintain normal functioning of the mother's skeleton.

The materials stored by the widest variety of organisms and in the largest quantities are those that can be broken down to provide metabolic energy. Energy-providing molecules that animals can safely store in large quantities are the complex carbohydrate, glycogen, and di- or triacylglycerols, esters of glycerol, and two or three fatty acids. Green plants store carbohydrates as insoluble starch especially in roots and other underground structures and especially in seeds and certain fruits, lipids as triacylglycerols. The complex carbohydrates generally require less energy for synthesis but their associated water makes them denser and less compact than lipids. Glycogen, a common short-term energy store in invertebrate tissues and in vertebrate muscle and liver, is always associated with large quantities of water.

Lipid molecules have such high affinity for each other that they form compact, almost water-free droplets of significantly lower density than other cell components. Comparative studies of angiosperm seeds and fruit, yeasts and various animal cells show that the intracellular mechanisms of the formation of lipid droplets are essentially similar in all eukaryotic cells (Murphy and Vance, 1999); the process is just greatly exaggerated in adipocytes. Lipid droplets arise from specialized regions of the endoplasmic reticulum and are controlled by specific enzymes and other associated proteins, called oleosin in seeds and perilipin in adipocytes.

Lipid synthesis and reclamation are biochemically quite complex and usually slower than the corresponding process of carbohydrates. While carbohydrates (and the glycerol component of lipids) can be broken down by aerobic or anaerobic metabolism, long-chain fatty acids must be oxidized aerobically, so storage lipids can only fuel prolonged, strenuous activities in animals in which the uptake and transport of oxygen are highly efficient.

Most animals can synthesize glucose from amino acids released by the breakdown of proteins but the process is slow and wasteful (because the amino groups have to be excreted). Small amounts of protein and free amino acids, particularly glutamine, may exist in the liver "only" as a storage material for use as precursors for the synthesis of proteins and nucleic acids as well as for glucose production, but most protein utilization entails "wasting" of lean tissue. The associated loss of function is particularly obvious when muscles are broken down to make glucose but protein depletion may also impair the immune and digestive systems. These deleterious effects mean that, especially in animals that have dedicated storage tissues for carbohydrate or lipid, proteins are usually only reclaimed for energy production when the other reserves are nearly exhausted.

Storage in Invertebrates

Many invertebrates are capable of surviving long periods without food. Many unicellular organisms and lower invertebrates have intracellular storage molecules and organelles that contain sufficient energy-generating material to sustain them for hours or days. In prolonged starvation, the body fabric, including structural proteins, is broken down to produce metabolic energy. The first to be reabsorbed are usually reproductive tissues, early embryos as well as gametes, but at least in less complex invertebrates, skeletal, locomotory, and digestive tissues can all be utilized if necessary. Some animals can shrink to an impressive extent while remaining capable of normal life functions and rebuild their bodies when food becomes available. For example, after 4 months of starvation, the medusae of the hydrozoan jellyfish *Aurelia aurita* shrink to less than a quarter of their original diameter but they continue to swim and can regrow if adequate feeding is restored. The ability to retain more or less normal functioning after prolonged starvation may be one of the advantages of the simple cnidarian body plan. More complex animals that have hard skeletons cannot lose so much body fabric without serious impairment of function.

Like vertebrates, higher invertebrates have specialized storage tissues that sequester energy-providing molecules at much higher concentrations than is normally possible when materials are stored in cells with other primary functions. Annelids, especially oligochaetes, have variable quantities of chloragogenous tissue (so-called because cell inclusions confer a green or yellow color) in the coelom around the intestinal wall. It has several functions, including that of storing lipids and glycogen and may be ancestral to the arthropod "fat body" or "hepatopancreas" as it is known in crustaceans.

The insect fat body is the most thoroughly studied invertebrate storage organ. It is present in larval as well as adult insects and, although originating in the abdomen, it can grow to be massive, extending into the thorax where it surrounds the major locomotory muscles. The stored lipids confer a white or yellow color when replete, as in locusts before migration and many other species before metamorphosis, but the proportion of lipid is rarely as high as it can be in adipose tissue of laboratory rodents. Storage lipids reach the muscles and other tissues as diacylglycerols or triacylglycerols in lipoproteins called lipophorins (Canavoso *et al.*, 2001).

Several hormones involved in energy metabolism and activity were shown to be secreted from the fat body in the 1950s and 1960s, leading to the suggestion that the insect fat body had gland-like functions comparable to those of the vertebrate liver. In *Drosophila* (Arthropoda, Insecta, Diptera), the fat body secretes several peptide regulators of metabolism and appetite that function remarkably like insulin-like growth factors in vertebrates (DiAngelo and Birnbaum, 2009). The fat body also contributes to innate immunity: when stimulated by bacterial extracts, *Drosophila* fat body cells produce at least seven antimicrobial peptides (Manfruelli *et al.*, 1999). Like vertebrate adipose tissue, invertebrate storage organs are much more than just repositories; they play an active role in energy metabolism and defense against disease.

Modern genetic techniques demonstrate remarkable similarities in the signal molecules regulating anatomically diverse storage tissues in invertebrates. Insulin is an ancient signal molecule known in *Caenorhabditis elegans* (Nematoda), arthropods, and all vertebrates (Watts, 2009). Most of the signals and receptors shown to be regulators of appetite and energy storage in mammals are known in the sea squirt *Ciona* (Ascidia, Chordata), an invertebrate chordate. Neuropeptide Y belongs to an ancient family of peptides that mediate signals between storage cells and the nervous system in various invertebrates.

Plants

The aerial structures of many temperate-zone and arctic plants die back in winter when cold or lack of light makes photosynthesis inefficient. Regrowth the following spring is supported by storage materials distributed through roots and surviving parts of the stem, which may be enlarged to form corms or tubers in some species. These structures are not severely limited in volume and density, so the storage material is nearly always starch with much associated water. Only a few higher plants, notably sugar beet (*Beta vulgaris*), store large quantities of disaccharides or monosaccharides. Such concentrated nutrients are often protected from herbivores by stout mechanical casings or by the presence of toxic secondary compounds. Removing or inactivating these plant defenses by leaching (e.g., cassava, *Manihot esculenta*) and/or cooking (e.g., potatoes, carrots, onions) were among the major technological advances of prehistoric humans, and greatly increased the range and nutritional quality of plant foods available to them.

Energy Storage in Seeds

The chief function of energy stores in seeds is to supply the germinating seedling until it grows its own leaves and can

photosynthesize for itself. The huge range of sizes of seeds is due mainly to the chemical composition and abundance of their storage material. The storage materials of the seeds of trees native to cool or cold climates, such as acorns, beechnuts, and hazelnuts, are usually starches, as are those of grasses such as oats, rye, wheat, and rice. Seed starches generally contain less water, and so are more compact, than those found in vegetative storage tissues, and like them are often protected from herbivory by toxins. The need to compress a lot of storage material into a small space makes synthesizing triacylglycerols from primary photosynthetic products worthwhile, and the storage materials of many large, longer-lasting seeds that are dispersed by animals are lipids, or mixtures of carbohydrates and lipids.

Green plants have a special biochemical mechanism, not known to be present in animals other than a few kinds of parasitic worms, that permits the incorporation of fatty acid carbon atoms into precursors for the synthesis of glucose, which is the starting point for the formation of the structural carbohydrates. Thus plants can convert the fatty acid components of seed lipids into stems and leaves that synthesize sugars in sunlight as well as oxidizing them in mitochondria or incorporating them into structural or storage lipids. Probably because of this additional role, seed lipids contain a much wider variety of fatty acids than animal storage fats (Murphy, 1994).

The triacylglycerols in seeds are always dispersed into tiny membrane-bound compartments called oil bodies, each only a few micrometers (thousandths of a millimeter) across at least some of which contain lipases that enable plant storage lipids to be mobilized quickly. Such structures contrast with the (relatively) huge droplet of triacylglycerols in vertebrate adipocytes that may be greater than 100 μm (0.1 mm) in diameter, or up to a million times the volume of oil bodies in seeds, and explain why nuts do not feel or taste as "greasy" as vertebrate adipose tissue and purified oils do, even though many are greater than 50% by weight lipid. Mobilization of seed storage materials begins at germination with the uptake of water that swells the proteins and starches. Preformed enzymes are activated and new ones synthesized, but, as in the case of animal systems, water-based enzymes cannot easily attack large droplets of lipid.

Chemical Composition of Plant Lipids

The temperature at which major physiological processes like seed germination take place seems to be the most important determinant of the composition of plant lipids. Most biochemical reactions take place in solution so the reactant molecules must be dissolved in liquid for metabolic processes to proceed. Unsaturated fatty acids make phospholipid membranes more fluid, so other molecules can move within and across them more readily and lower the freezing temperature of triacylglycerols. Complete breakdown of unsaturated fatty acids releases slightly less metabolic energy (about 1–2% less for each double bond) than the saturated equivalents with the same numbers of carbon atoms. Triacylglycerols containing saturated fatty acids also pack more neatly (i.e., they have higher melting points), than those containing mono- and polyunsaturates, and homogeneous assemblages solidify at higher temperatures than mixtures. So in general, more energy can be stored in a small space as saturated triacylglycerols, the more homogeneous, the better (Pond, 1998).

All the plants in which saturated fatty acids are abundant as storage lipids are native to the warm, equable climates of tropical lowlands. Plants such as coconut, oil-palm, nutmeg, and cacao that live only in hot climates do not need polyunsaturated fatty acids to keep their storage lipids fluid: they can maximize the energy obtainable from the smallest volume of storage organ with triacylglycerols containing mostly saturated fatty acids. Those of the seeds of the cinnamon tree, for example, are up to 95% lauric acid (C12:0), although the fatty acids in the fruit lipids of the same species, and of its relative, the avocado pear, are around 50% oleic acid and 25% palmitic acid, with very few medium-chain saturates, a typical composition for an oily, rather than a fatty, fruit. Large seeds such as oil-palms that contain abundant saturated triacylglycerols provide their seedlings with energy supplies that last them weeks or months.

Plant oils rich in polyunsaturated fatty acids are found in species adapted to temperate climates. Olive trees, with fruits and seeds containing mostly monounsaturated fatty acids, and nuts such as walnuts and pistachios which contain both storage lipids and starch, grow in cooler climates than that required by oil-palm or cacao, but are still not worth cultivating for their oil outside areas that have long, hot summers. Lipids containing mixtures of different kinds of fatty acids remain liquid at lower temperatures than those in which all the fatty acids are similar, so those of nearly all temperate-zone and polar species are mixtures of many different fatty acids. However, several possible mixtures of fatty acids in triacylglycerols may have similar physical properties and each plant may be adapted to germinate under slightly different conditions, so the fatty acid composition of seed lipids may differ between species growing in similar climates. Some species can make major alterations to the composition of their seed lipids according to the climate under which they are growing. The fatty acids of the seeds of the sunflower (*Helianthus annuus*) are 44–72% linoleic acid (18:2*n*-6), the remainder being mostly oleic acid (C18:1). Those growing in the coolest conditions contain the most linoleic acid, but their total yield is smaller and the crop takes longer to mature (Pond, 1998).

The seeds of the jojoba bush (*Simmondsia chinensis*) are almost unique in that the storage material is a wax, not a triacylglycerol. Jojoba oil is liquid at normal temperatures (in contrast to bees' wax, which is used by bees as a solid) because both the fatty acid and the alcohol components of the wax are long-chain molecules. The fatty acids are mostly (around 74%) C20:1, with smaller quantities of C22:1, and oleic acid (C18:1), and the alcohols also contain 20 or 22 carbons. Although the proportion of lipid in the seeds varies widely between strains and according to the conditions under which the plant is grown, the fatty acid composition of the wax, in contrast to that of triacylglycerols, is remarkably constant. Jojoba oil, like other waxes, is indigestible to most animals but the plant is cultivated widely in California and other warm, semidesert areas of the USA where the extracts are used as

lubricants and in cosmetics. Jojoba oil would also be suitable as fuel, so the crop has been proposed as a renewable substitute for gasoline.

Protection from Seed Predators

The lipids and starches concentrated in grains, nuts, and seeds provide a rich source of food for animals including ourselves. The partitioning of seed lipids into tiny oil bodies makes them very digestible for animals because less mechanical emulsification is required. Seed-eating mammals such as rodents gnaw or crush seeds with their hard, continuously growing teeth while parrots, finches, and other seed-eating birds use their powerful, finely controlled beaks to shred nuts and seeds. Fragmented nuts are digestible, highly nutritious food that is especially suitable for flying birds, which cannot afford to carry around heavy gut contents for long periods of slow digestion.

Seeds such as walnuts, almonds, cacao (chocolate) beans, and coconuts protect their nutritious contents with a hard shell (in many species, the nutshell is much harder than the tree's wood) that only the most powerful beaks or persistent gnawing teeth can open. Other nuts, such as peanuts, form underground, avoiding seed predators that cannot dig to reach them. Small seeds such as those of wild strains of flax, sesame, and sunflower scatter as they fall, making it hard work for animals to collect them, and their coats are very tough relative to their size (Pond, 1998).

Many plants have evolved chemical as well as mechanical means of deterring seed predation. Many seeds and/or fruits are rendered distasteful and/or poisonous by the addition of "secondary compounds" some of which are used as sources of flavorings and/or drugs, including caffeine (coffee beans *Coffea arabica*), opium (poppy seeds, *Papaver somniferum*), L-dopa (velvet bean *Mucuna pruriens*), betel nut (*Areca catechu*), and pepper (*Piper nigrum*). Some seed-predators, notably weevils (Insecta, Coleoptera) have evolved specific biochemical mechanisms for detoxifying the secondary compounds of their preferred food plants.

The storage materials of other plants are themselves poisonous to animals. The beans of the castor oil plant, *Ricinus communis*, are protected from herbivory by the poisonous protein, ricin, and by the fatty acid, ricinoleic acid, which can account for over 90% of the storage triacylglycerol fatty acids. Animal lipases cannot easily hydrolyze triacylglycerols containing ricinoleic acid, so most pass undigested through the gut. Eaten in small quantities, castor oil acts as a lubricant and slight irritant, thus relieving constipation in people and dogs and cats, but it is toxic in large quantities. Oil from unimproved oil-seed rape (and turnips, swedes and cabbages, and other Brassicaceae) contains over 40% erucic acid (C22:1) which renders the seeds toxic to weevils and other seed-eating insects as well as to vertebrate seed predators such as voles, mice, finches, and humans. Toxicity is not reduced by domestic cooking. In the West, rape-seed oil (known as canola oil) is produced from plants selectively bred and, more recently, genetically engineered, to form seed triacylglycerols that contain as little as 5% erucic acid. Insects as well as people find the artificially modified plants much more appetizing than the wild forms, and more pesticides are needed to protect such crops from herbivore damage.

The caterpillars of the cabbage white butterfly feed exclusively on the leaves and flowers of cabbage and related brassicas, including cauliflower, broccoli, and oil-seed rape but they have not acquired more of a taste for erucic acid than mammals have. Even when present in quite high concentrations in the diet, the unusual fatty acid is absent from phospholipids and triacylglycerols in the insects' tissues. Either it is not absorbed through the gut or not esterified into triacylglycerols, or, less likely, it is all used as fuel at once, before it has a chance to become incorporated into structural or storage lipids. Erucic acid even seems to be toxic to the plants themselves: highly specific enzymes direct it into seed triacylglycerols only. None is found in phospholipids or any other nonstorage lipid, suggesting that its presence in such molecules would disrupt membrane function (Pond, 1998).

Vertebrate Adipose Tissue

Proper adipose tissue is unique to vertebrates. It occurs sporadically among fish, often in a greatly modified form, but is most extensive in mammals, birds, reptiles, and amphibians. Its distinctive feature is white adipocytes, which can accommodate much larger quantities of triacylglycerols than any other kind of cell. The formation of single large lipid droplets is mediated by adipose-specific protein 27 (FSP27) (known in humans as cell death-inducing DFF45-like effector C) that is unique to vertebrates. In conjunction with perilipin, it forms an interface between lipids and proteins, promotes lipid uptake and coalescence of droplets while reducing the maximum rate of lipolysis (Puri *et al.*, 2007).

White adipocytes are richly innervated by the sympathetic nervous system and its transmitter, epinephrine (noradrenalin) stimulates lipolysis. Almost all the adipocytes' stores of lipid are not for their own use, but for export to other tissues as required. Not all adipocytes serve solely, or even mainly as whole-body energy reserves supplying tissues via the blood. Most adipocytes secrete and respond to regulators of appetite and metabolism and many are specialized to paracrine relationships with contiguous tissues including the immune and vascular systems (Poulos *et al.*, 2010; Pond, 2011).

The compartmental arrangement of storage lipids found in seeds, brown adipose tissue and many invertebrate storage tissues is not necessary in white adipocytes because such cells do not metabolize significant amounts of their own triacylglycerols. But brown adipose tissue oxidizes its own lipids to generate heat, although when such supplies run low, it can, like most other animal cells, take up metabolites from the blood circulation. The similarity in size and general appearance between plant oil bodies and brown adipocytes may arise from the fact that in both cases, storage lipids are utilized in the cells in which they are sequestered.

As animal cells go, replete adipocytes are huge, thousands of times larger than red blood cells, most brain cells, and the cells of the immune system that protect the body from disease. In **Figure 1**, only the top layer of adipocytes is in the plane of focus and the nucleus, cytoplasm, and lipid droplet cannot be distinguished. Unless severely depleted of lipid, adipocytes are

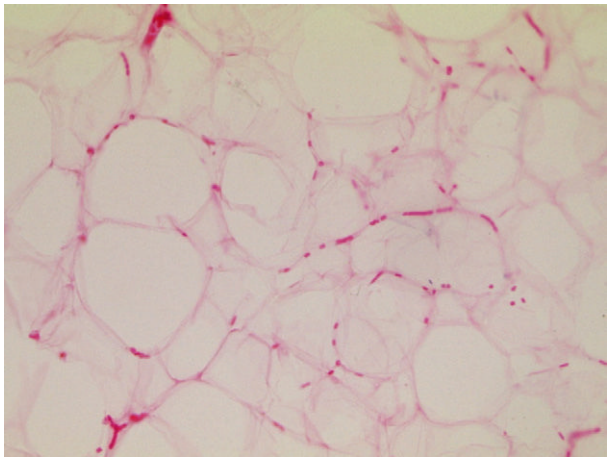


Figure 1 A thick section of fixed white adipose tissue from an obese adult rabbit. The storage triacylglycerols have dissolved away in the preparation process so most adipocytes appear "empty." The thin layer of cytoplasm near the cell membrane and extracellular mesh of collagen are stained pink, and none of the small nuclei are visible. The small bright red cells are erythrocytes (red blood cells) in blood capillaries that pemeate the tissue. A larger blood vessel is visible at the top left. The adipocytes are roughly spherical but only a few are sectioned through the center thus revealing their actual diameter of about 100 μm (0.1 mm).

spherical, or very nearly so, which is an unusual shape for functionally mature animal cells (other than eggs). Adipocytes are packed in a tight mesh of collagen that also supports numerous blood vessels and nerves. Adipocytes occupy most of the volume of replete, metabolically active white adipose tissue, but extracellular material can dominate lipid-depleted depleted tissue.

Size and Number of Adipocytes

Adipocytes mature from preadipocytes, small cells capable of rapid proliferation and of differentiation into several cell types. Once fully formed and functional, adipocytes are long-lived cells but, like most other cells, they can die by apoptosis, an orderly, often energy-consuming process in which the cells synthesize enzymes that quickly destroy their fabric. In people and sedentary domestic animals, the proportion of lipid in whole adipose tissue is rarely less than 40% and it increases with fattening, reaching as much as 85% of the total mass of the tissue in middle-aged men who are 44% by weight adipose tissue. The rest of the tissue is mostly protein and water. The proportion of lipid is always lower in wild animals, and does not change much with fatness. Dwarf hamsters (*Phodopus*) naturally become very fat during the summer, but chemical analysis shows that their adipose tissue never exceeds 46% lipid, even when it amounts to 35% of the body mass. The proportion of lipid in adipose tissue is almost constant (at 42–66% depending on the depot) in arctic foxes living wild in the high Arctic, whose body composition ranges from 3% to 33% by weight adipose tissue (Pond *et al.*, 1995).

Among mammals and birds (there are no data for other classes), the adipose tissue of larger species consists of fewer, larger adipocytes. Adipocyte number scales to (body mass)$^{0.75}$, and the adipocytes range in volume from 0.01 nl in bats and shrews, to up to 4 nl in well-fed baleen whales. Carnivorous mammals and ruminants (whose energy metabolism utilizes mainly fatty acids rather than glucose) have about four times more adipocytes than nonruminant herbivores (whose metabolism uses more glucose) of the same size and fatness but they are not fatter because the adipocytes are proportionately smaller (Pond and Mattacks, 1985). By coincidence, the adipocytes of rats and mice, small, nonruminant herbivores, are about the same size (0.1–1 nl) as those of humans, large omnivores on a high-fat diet.

Thorough studies of many specimens of the same species in the wild always reveal much interindividual variation in the total number of adipocytes in the body as a whole, much of which cannot be attributed to age, sex or any obvious feature of dietary history. But adaptations to higher lipid storage that supports prolonged, frequent fasts can be identified. For example, Svalbard reindeer (*Rangifer tarandus platyrhynchus*) live at latitudes 74–81° N and accumulate enough fat during summer and autumn to sustain them during the long winter when only low-quality dry vegetation can be reached by digging through the snow. At the beginning of winter, adipose tissue can exceed 25% of the total body mass, and up to 8 cm thick over the rump but it contains only two to three times more adipocytes than would be expected in temperate zone and tropical deer of similar size.

The number of adipocytes (relative to total body mass) proved to be even more variable in carnivores (Pond *et al.*, 1995). Some arctic foxes (*Alopex lagopus*) and wolverines (*Gulo gulo*) have up to five times as many adipocytes as expected from their overall size, but a significant minority actually have fewer than the expected number. At the time the specimens were collected, there was no consistent relationship between the total number of adipocytes and fatness, measured as the relative mass of adipose tissue in the body. The number of adipocytes in the adipose tissue is not a major determinant of the capacity for fattening. If fewer adipocytes are present, each must expand more to accommodate as much storage lipid as specimens that, for whatever reason, have proportionately more adipocytes.

In humans, and laboratory animals kept under highly artificial regimes of diet and exercise, the number of adipocytes increases with age after sexual maturity, but all these effects are slight, generating at most a doubling of the adipocyte population. They cannot be induced in naturally obese wild mammals such as dormice (*Glis glis*) and there is almost no evidence for an increase in adipocyte number after growth of lean tissue stops in any wild mammal so far investigated.

Anatomical Arrangement of Adipose Tissue

Adipose tissue in most reptiles and amphibians is concentrated in paired depots inside the abdomen, or in the case of lizards on the tail. Adipose tissue in tortoises and turtles is more widely distributed, filling spaces between muscles and in corners of the shell.

Birds have both superficial and intra-abdominal depots. The main superficial depots are at the base of the neck around the furcula ("wishbone"), on the anterior surface of the thigh,

over the breast muscle and on the back near the tail. Biochemical differences between depots in birds have not been as thoroughly studied as those of mammals, but as in mammals, the superficial depots usually enlarge more than the internal ones. They expand laterally as well as thicken, forming an almost continuous layer that is thickest over the back, breast and the anterior surface of the thigh. Over half the body mass can be adipose tissue in birds that fatten before migration or prolonged fasts associated with molting or breeding. A well-studied example is the bar-tailed godwit (*Limosa lapponica baueri*) that flies 11,000 km between New Zealand from Alaska. This small wader starts its journeys with fat reserves amounting to 55% of fresh body mass, occupying space created by selective atrophy of the gut and other nonessential organs (Landys-Ciannelli *et al.*, 2003). Measurements on such "adaptively obese" animals reveal surprisingly low, sometimes undetectable energetic costs of the additional body mass, both in flight and walking (Kvist *et al.*, 2001). Locomotory efficiency is unimpaired by adipose tissue that reaches 32% body mass in Svalbard rock ptarmigans (*Lagopus muta hyperborea*) (Lees *et al.*, 2010).

Mammalian adipose tissue is always partitioned into a few large and many small depots. Its comparative anatomy is based on measurements of the size and number of the adipocytes. Figures 2 and 3 show the patterns of distribution of larger and smaller cells that emerge for various mammals. Larger adipocytes do not always make larger depots: Some of the smallest depots, such as the popliteal behind the knee, consist of relatively large cells. Nor do adjacent depots, such as those either side of the forelimb, necessarily consist of cells of similar size.

In the adult camel (**Figure 3a**), over a third of all adipose tissue is in the humps, compared with 4–8% in the corresponding site of hares (**Figure 2b**), up to 4% in squirrels and a maximum of 1% in stoats (**Figure 2c**). But less than 1% of the camel's adipose tissue is in the inguinal depot, on the front of the thighs and the sides of the abdomen. This depot contains over a third of the total adipose tissue in stoats and squirrels, more than a fifth of that of the lioness (**Figure 3c**), and an eighth of that of horses (**Figure 3b**). There are also contrasts in the proportions of adipose tissue located on the dorsal wall of the abdomen, from nearly half in squirrels and hares (**Figure 2a** and **b**), to an eighth or less in the large animals (**Figure 3**). Although the range of sizes of adipocytes differs greatly between species and between specimens, partly because some specimens were fatter than others, the anatomical pattern of relative sizes of adipocytes is similar in all terrestrial mammals. What differs between species is the relative abundance of adipocytes in each depot: the relative size of their adipocytes forms the same general pattern. Taxonomically related species have similar patterns of distribution of adipose tissue (Pond, 1998).

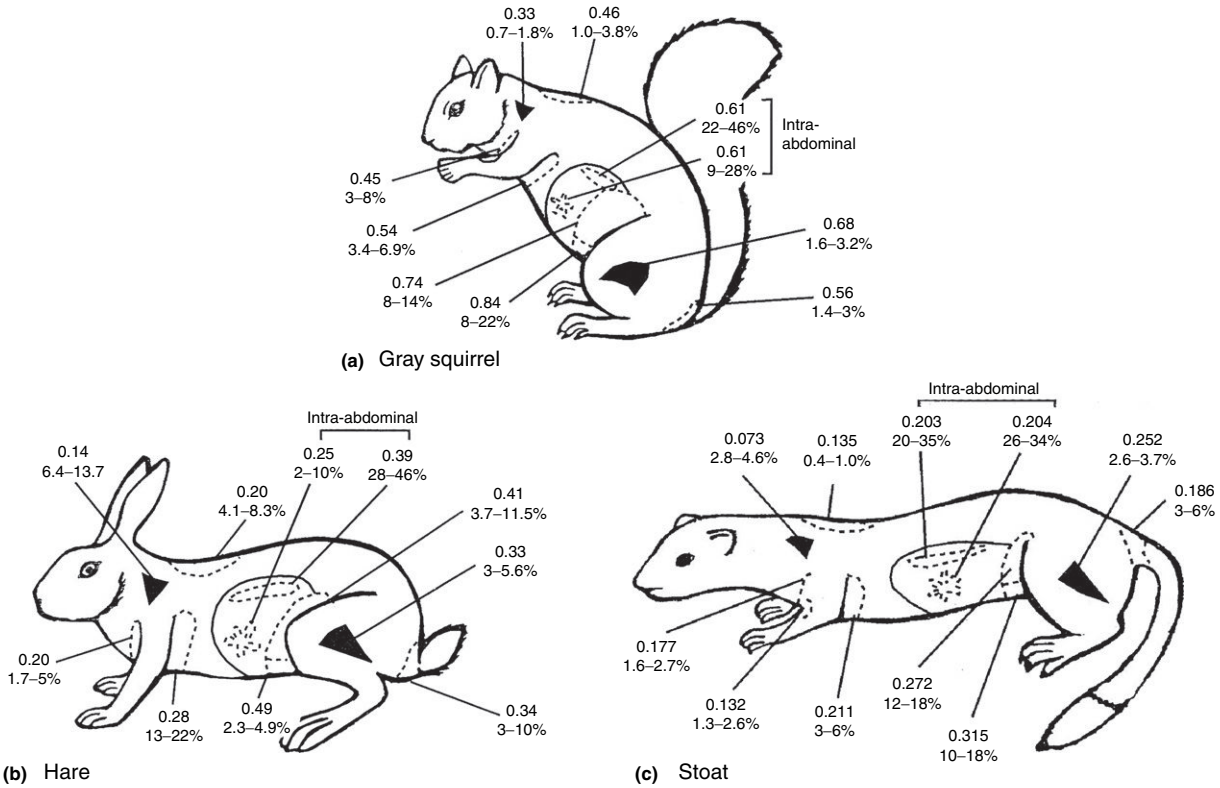


Figure 2 The organization of adipose tissue in various wild mammals. (a) The gray squirrel (*Sciurus carolinensis*), figures are the means of four specimens, body mass 0.38–0.67 kg, 3.1–14.4% dissectible adipose tissue. (b) European hare (*Lepus timidus scotius*), figures are the means of five specimens, body mass 2.4–3.5 kg, 0.3–3.2% fat. (c) Stoat (*Mustela erminea*), figures are the means of five specimens, body mass 0.19–0.4 kg, 3.2–6.3%. Upper sets of figures are the mean volume of adipocytes in nl (10^{-9} l); lower figures are the proportions of the total dissectible adipose tissue to be found in the depot. The intermuscular depots are shaded. Data and illustration from Pond CM (1998) *The Fats of Life*, 344 pp. Cambridge: Cambridge University Press.

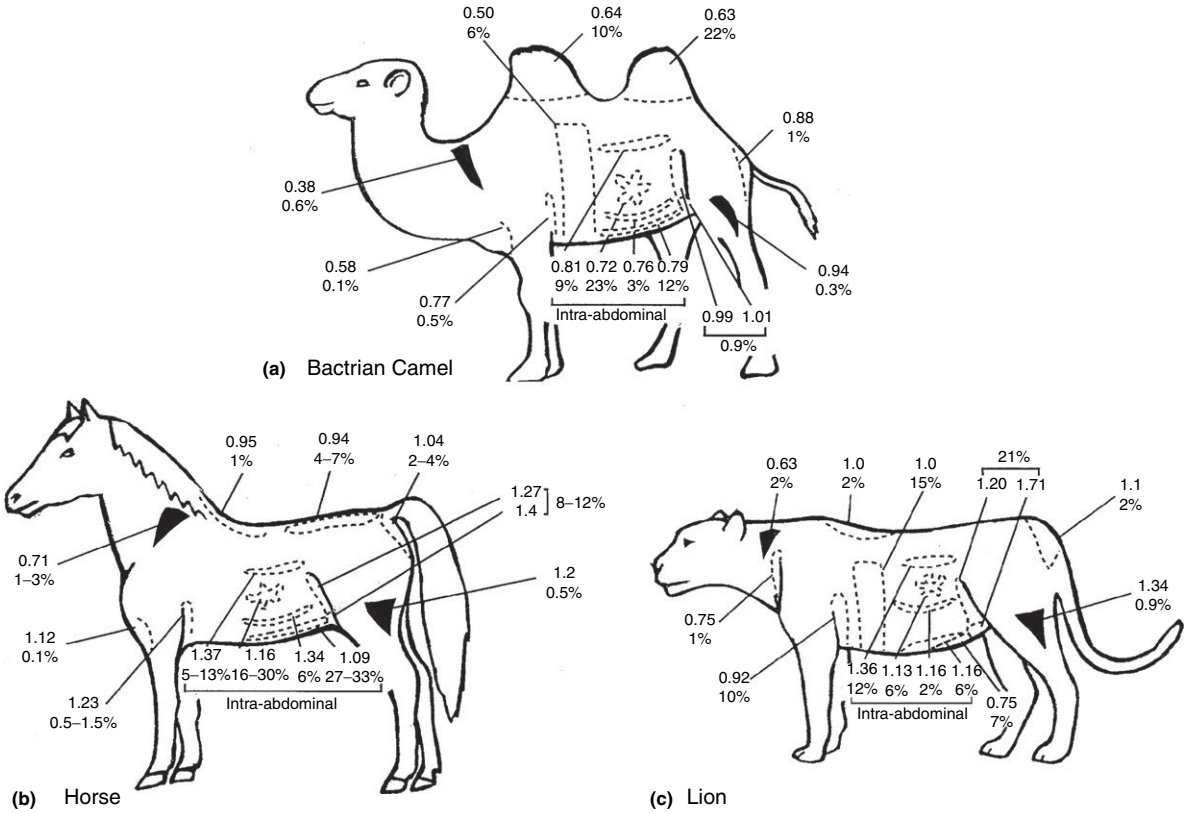


Figure 3 The organization of adipose tissue in some large zoo-bred or domesticated mammals. Data are presented in the same way as in **Figure 2**. (a) Two-humped camel (*Camelus bactrianus*), a pregnant female, body mass 650 kg, 6.1% fat. (b) Domesticated horse (*Equus caballus*); figures are means of measurements from two geldings, body mass, 444 and 570 kg, 5.2%, and 4.0% fat. (c) Pregnant female lion (*Panthera leo*), body mass 167 kg, 13.3% fat. Data and illustration from Pond CM (1998) *The Fats of Life*, 344 pp. Cambridge: Cambridge University Press.

Although a great deal is now known about the biochemical steps by which adipocytes differentiate and mature, almost all of the information comes from the study of cells in culture, so we cannot explain in detail why adipose tissue forms where it does, or what determines the species differences in the relative sizes of the depots. The gross anatomy must have implications for the animals' habits and capabilities because there are site-specific differences between adipocytes in many physiological properties. Some depots seem to be equipped mainly for taking up lipids from the blood after a meal, others for quick responses to the onset of strenuous exercise. Many of the minor depots that enclose lymph nodes may be mainly or entirely devoted to supporting immune processes and contribute little to whole-body energy storage (Pond, 2009, 2011). They are among the first to form during ontogeny and the last to be depleted in starvation. In very lean wild animals, the only replete adipocytes may be those associated with lymph nodes and heart.

Relatively minor depots may be almost undetectable in small animals, especially if the specimens are lean, so perhaps it is not surprising that larger animals (**Figure 3**) should appear to have more "extra" adipose depots for which there are no exact equivalents in the smaller species (**Figure 2**). Some are just too big to be thus explained away: up to a third of all adipose tissue of horses (**Figure 3b**) and donkeys forms a sheet, sometimes several centimeters thick, on the inner ventral wall of the abdomen. There is a small quantity of adipose tissue in the corresponding depot of large carnivores and camels (**Figure 3a** and **c**), but it is absent in most other species.

In primates, the most conspicuous such depot is the "paunch" that arises from the midline on the outer wall of the abdomen and expands laterally as it thickens, sometimes becoming very thick at the midline, and extending as far as the chest and the hip bones. In humans, the paunch adipose tissue is centered around the navel, usually somewhat thicker above it in men, and below it in women. The paunch depots seems to accumulate fat selectively in primates: It may be minimal in lean (but not emaciated) specimens, but often becomes very massive in humans and in monkeys and apes that become obese in captivity, while the other depots expand more slowly with increasing fatness, or, in the case of the intermuscular depots, hardly at all. An abdominal paunch with these properties seems to be a unique and distinctive feature of this group and is as conspicuous in lemurs (primitive strepsirrhine primates) as it is in apes, indicating that it appeared early in their evolutionary history. In spite of having this "extra" depot, with the exception of modern humans, primates are generally no fatter than other mammals when living under natural conditions (though some may become obese in zoos).

There is no trace of this depot in hares, squirrels or any related small mammal, but Carnivora have a "paunch" depot

on the outer ventral wall of the abdomen. The depot can become thick in cats and their relatives but it is relatively small in canids, ursids, and mustelids. It does not expand disproportionately any more than any other superficial depot, even in very obese wild carnivores such as arctic foxes. Very little is known about the physiological basis of the paunch depot's special properties because its complete absence in rodents precludes convenient experiment study (Pond, 1998).

Measuring Fatness in Wild Animals

Because adipose tissue is the most variable component of the body, the simplest and most widely used measure of fatness is body mass. Body mass index is this measurement adjusted for body length or height. In small animals, the guts represent such a large and variable fraction of the total body mass that such methods are not very accurate or reproducible. If samples of adipose tissue can be obtained by biopsy or post mortem, adipocyte volume is also used to estimate fatness, but such measurements from most wild animals as not as easily interpreted or reliable as experiments on laboratory rodents might suggest. Such data are modulated by the effects of body size and natural diet on the cellular structure of adipose tissue, site-specific differences in adipocyte volume and variability in cellular structure of the adipose tissue.

Figure 4 shows some data on adipocyte volume from wild polar bears. Females (but not males) stop growing when they become sexually mature and fatten enormously during pregnancy, producing a predictable relationship between body mass and adipocyte volume, but the correlation is weak for cubs and males that are still growing because adipocyte number is also increasing. However, these data do illustrate some important aspects of bear ecology: cubs still with the mother are generally fatter than those of similar size that have been abandoned to forage alone, and there is great variation in the fatness of cubs and adult males, especially among the larger specimens, probably reflecting differences in hunting success.

For children and young people, measuring skinfold thickness with calipers is widely used to assess fatness, but the skin of most mammals is too thick and tough, and the underlying adipose tissue too thin and patchy, for skinfold thickness measurements to be of much use. As wild mammals get fatter, superficial adipose tissue always expands more rapidly than that inside the abdomen. The depots shown in Figures 2 and 3 become thicker and expand laterally to cover a greater area. In very fat specimens, including many humans, adjacent superficial depots overlap and merge so that they become difficult to distinguish, but measurements of adipocyte volume can reveal their identity. Since expanding depots spread out as well as thicken, thickness alone is only a crude measure of the total adipose tissue, even if the site from which the measurements are taken is kept strictly constant. The fatness of small birds is often assessed by inspection and/or palpation of the superficial depots around the neck or leg, a simple method that takes account, at least in a qualitative way, of changes in both the thickness and the area of the adipose tissue.

A high-tech version of skinfold thickness measurements is the ultrasound scanner, which can measure fat thickness fairly accurately in large mammals, especially if, as in pigs, the hair is sparse. The method depends on the fact that sound vibrations travel at different speeds through chemically different materials, and are reflected, producing an echo, at the interfaces between the two tissues. A pulse of high frequency sound is directed into the body from a source placed on the skin, and the timing and intensity of the echo is recorded. Shaving hair

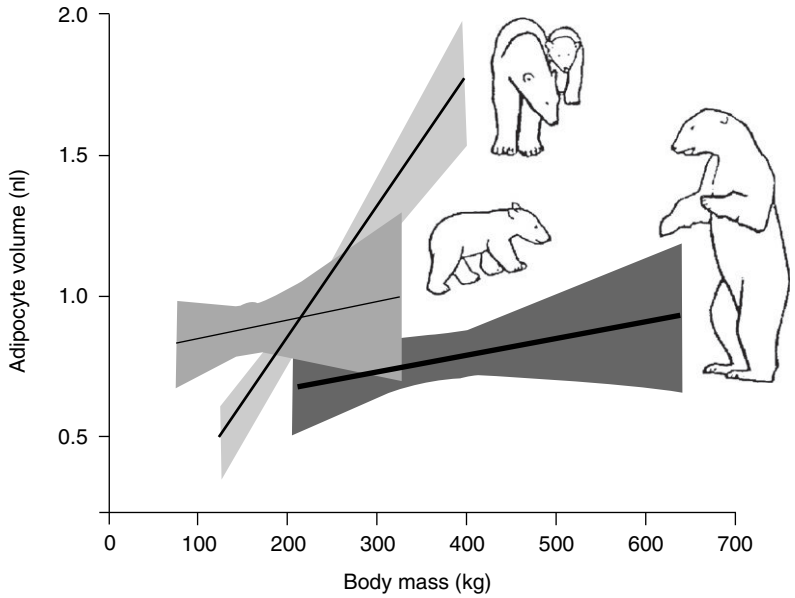


Figure 4 Mean and 95% confidence intervals of the average volume of adipocytes in a sample of adipose tissue from wild polar bears in northern Canada, measured in cubs accompanied by their mothers, weaned females, including pregnant and lactating mothers and solitary males. Data from Ramsay MA, Mattacks CA, and Pond CM (1992) Seasonal and sex differences in the cellular structure and chemical composition of adipose tissue in wild polar bears (*Ursus maritimus*). *Journal of Zoology (London)* 228: 533–544. Illustration from Pond CM (1998) *The Fats of Life*, 344 pp. Cambridge: Cambridge University Press.

is usually necessary and the probe's anatomical position must be consistent. With a resolution of only about 1 mm, ultrasound scanners are not accurate for animals smaller than dogs and for any specimens in which the superficial adipose tissue is very thin.

Lipids are much poorer conductors of electricity than the watery components of cells and body fluids. This contrast is exploited to measure whole-body conductance of electric current as an indicator of the proportion of lipid present. More fat produces lower whole-body conductance. The apparatus has to be used carefully, because hair, feathers, and the pockets of air they enclose also have high electrical resistance, but it is a useful technique, especially for small birds and mammals, and works equally well on living or freshly dead specimens. Fat appears pale in magnetic resonance imaging so this apparatus has been used to study the abundance and anatomical relations of white adipose tissue in humans and, more recently, in sedated domestic livestock and laboratory animals.

With these new techniques impractical for most large animals, dissection is still widely used. For more than 30 years, assessment of lipid reserves in wild vertebrates *in vivo* and post mortem has depended mainly on empirically derived "indices" which are based on the dimensions of the whole body, or that of the thickness or mass of one or a few adipose depots. The sample sites are chosen at random or for some technical reason such as simplicity of dissection. The enormous diversity of sample sites and indices is cumbersome and makes comparisons between studies very difficult and imprecise. One of the favorites is the kidney fat index, the mass of the perirenal adipose tissue as a percentage of that of the kidneys. This index works fairly well for lean deer, antelope and similar ungulates, unless the mass of the kidneys changes seasonally, which can happen in some drought-adapted species. It is less accurate in carnivores (see Figures 2 and 3) in which the perirenal depot represents only a small fraction of the total adipose tissue, and is indistinguishable from that on the rest of the dorsal wall of the abdomen.

Wildlife biologists also try to estimate total fatness of mammals from the lipid content of the marrow in the long bones, thereby eliminating the need for a complete dissection. However, "marrow fat indices" only work on species such as ruminants in which the adipocytes are depleted in starvation. In many other mammals, these adipocytes fail to respond to fasting, probably because they are specialized to support the immune functions of the red marrow.

A further problem with using the mass of one or a few depots as an "index" of fatness is the inherent variability of the partitioning of adipose tissue between depots. In all species for which there are sufficient data, substantial interindividual variation in the relative masses of depots that cannot be attributed to age, sex, or lean body mass is always found, even in genetically homogeneous wild populations and in animals bred and maintained under carefully controlled laboratory conditions. Of the intra-abdominal depots, the dorsal wall of abdomen depot (that includes the perirenal) undergoes the largest changes, especially in deer, sheep, cattle, antelopes, and other ruminant animals: it expands greatly in extreme obesity and shrinks to almost nothing after prolonged starvation. Its anatomical position means that, like the fat bodies of lower vertebrates, such large, and sometimes rapid, changes in size have minimal impact on the shape or size of any adjacent organs. Depots that contain lymph nodes undergo the least change, so they are relatively massive (compared with those elsewhere in the body) in lean specimens, and relatively small in obese ones.

Adipose Tissue Function

Adipocyte Metabolism

The largest component of replete adipocytes, the lipid droplet, is almost pure liquid triacylglycerols. The thin rind of cytoplasm between the outer membrane and the lipid droplet contains the relatively small nucleus displaced from its usual central position, biochemical apparatus for building and storing receptors, messenger molecules, and enzymes plus a few mitochondria where the fuel that powers their synthesis is produced. Molecular traffic into and out of adipocytes is regulated by a variety of receptors, carrier molecules, enzymes and other substances, most of which are produced by the cells themselves. Until about 20 years ago, adipocytes were known to produce only a few kinds of proteins, but techniques to separate, identify, and synthesize proteins and other biological molecules are rapidly becoming cheaper and more efficient, so biochemists are constantly adding to the list of messenger molecules secreted from adipocytes, receptors that enable them to respond to such signals from each other and more remote tissues, and transport proteins and enzymes that regulate uptake of materials that at least some adipocytes can produce.

Leptin (formerly known as *Ob*-protein) is among the best known secretions of adipose tissue, identified in the early 1990s from the study of a spontaneous mutation that causes severe obesity in mice. This small protein passes from adipocytes (and several other tissues especially those involved in reproduction) into the blood, where it is found in concentrations of around 10^{-8}–10^{-9} M. It binds to specific receptors on many neurons in the hypothalamus, the region of the brain that controls appetite (and many other drives and emotions). As well as regulating appetite, leptin adjusts energy expenditure and the physiological response to cold. Obesity could be caused by disruption to these chemical signals from the adipose tissue to the brain, leading to chronic overeating and/or very low rates of energy expenditure. Most of the signals and receptors shown to be regulators of appetite and energy storage in mammals are known in the sea squirt *Ciona* (Ascidia, Chordata), an invertebrate chordate. The appetite-suppressing hormone leptin itself seems to be specific to vertebrates, probably appearing early in the evolution of fish, but insects have analogous peptides that signal peripheral energy stores to the nervous system (Gorissen *et al.*, 2009).

Fatty acids cannot be excreted, so any excess has to be deposited in the storage tissues. Most tissues oxidize lipids, with the largest consumers being cardiac and skeletal muscle, the liver, and, especially in neonatal mammals and hibernating species, brown adipose tissue. ATP production in mitochondria is never 100% efficient, some of the energy used becomes heat but "uncoupling" of fuel utilization from the

formation of ATP greatly increases thermogenesis in brown adipose tissue (Cannon and Nedergaard, 2004). Much of the volume of brown adipocytes is occupied by mitochondria so brown adipose tissue is only about 10% by weight storage lipids, which form numerous tiny droplets instead of a single large droplet in white adipose tissue. The fatty acids released by lipolysis are consumed within brown adipocytes for heat production, while those of white adipocytes are released into the blood, where other tissues can take them up and use them.

When fully activated, for example, during rapid warming of the body at emergence from hibernation, brown adipose tissue uses fuels and oxygen at up to 10 times the rate of active muscles, producing enough heat to raise a small mammal's body temperature by more than $2\ ^{\circ}C\ h^{-1}$ (Cannon and Nedergaard, 2004). In the newly born and in awakening hibernators, much of the brown adipocytes' fuel comes from white adipose tissue. The latter's stores therefore determine the animal's ability to maintain its temperature until it can find food: mother's milk or a rich, digestible meal.

Brown "fat" was long believed to be alternative forms of white adipose tissue until the recent demonstration that the pattern of gene transcription in stem cells differentiating into brown adipocytes resembles that of muscle more closely than that of white adipocytes. Brown adipocyte precursors can be detected in skeletal muscle (Crisan *et al.*, 2008) and muscle-specific micro-RNAs can be found in such cells in tissue culture. Nonetheless under artificial conditions, brown adipose tissue can be transformed into white and *vice versa*, and at least in certain species, does so naturally. Many of the sites in which it actively generates heat in newborn mammals contain normal-looking white adipose tissue in the adult. In small mammals that live in cold climates, such as dwarf hamsters, at least some depots consist of mixtures of brown and white adipocytes throughout life, with the thermogenic adipocytes increasing when in animals living continuously in a cold environment, especially if the light/dark cycle resembles the long, dark nights of winter.

Fatty Acid Composition of Storage Lipids

Animals can synthesize the common saturated (palmitic, stearic, and myristic) and monounsaturated (oleic) fatty acids and many can insert an additional double bond into long-chain fatty acids, though this capacity is minimal in higher vertebrates including mammals. Animal desaturases work slightly differently from the corresponding enzymes in plants and microbes, so fatty acids that originate in animals are distinct from those synthesized by primary producers. Except in ruminant animals (in which the rumen microbes alter the fatty acids before they can be absorbed), most fatty acids pass unaltered from microbes or plants to herbivore and from prey to predator.

Recent studies have suggested that the lipid composition of the diet may determine the habits and habitats of poikilothermic animals. The blue-tongued skink or shingle-backed lizard, *Tiliqua rugosa*, is quite common in the deserts of Western Australia. Its natural diet is mainly plants (at least when adult) but also includes carrion and small prey, so it can be maintained on an artificial diet in captivity. After a few weeks on experimental diets, those given food containing sunflower oil (rich in polyunsaturated fatty acids) chose to spend their days (and nights) in places that were up to 5 °C cooler, than those fed on mutton fat (containing mostly saturated fatty acids) (Geiser and Learmonth, 1994). These experiments show that diet can affect important aspects of their ecology, including the times of day at which they are active, and where they forage.

Long-distance migration in birds, especially small species, is among the most metabolically demanding of all activities, fuelled almost entirely by fast, sustained mobilization of storage lipids. Sandpiper (*Calidris pusilla*) demonstrated selective incorporation of dietary fatty acids into structural or storage lipids and evidence for adaptive desaturation that maximizes energy density and efficient mobilization of the storage lipids during prolonged flight (Weber, 2009). However, studies of another species of sandpiper (*Philomachus pugnax*) produced no evidence for similar selectivity of fatty acids mobilized during shivering elicited by prolonged exposure to cold. This comparison suggests that active lipid management entails some physiological cost: the process is essential preparation for migration which requires on precise coordination between muscles during flight but is dispensable for shivering, a more chaotic activity.

Similar investigations on mammals have not yet been performed.

The lipids of cold-blooded animals generally contain a higher proportion of unsaturated fatty acids than those of warm-blooded animals, with species living in colder regions having the most polyunsaturates. Marine algae generally have more highly unsaturated fatty acids than terrestrial plants, and the animals that eat them both incorporate the algal fatty acids unchanged into their own tissues, and elongate and/or desaturate them further. Thus most plants and algae cannot build fatty acids with more than 18 carbons, but many marine animals can elongate oleic acid to form gadoleic acid (C20:1*n*-9) and various 22- and 24-carbon fatty acids. Gadoleic acid is common in the lipids of many kinds of fish, especially those living at high latitudes. Other fatty acids are also readily elongated, but those of the *n*-6 series derived from linoleic acid are even more unstable (in the sense that they readily participate in chemical reactions, especially oxidation) than those of the *n*-3 series, so the latter are more abundant in the lipids of most marine organisms.

Algae are the basis of nearly all marine food chains, so the storage and structural lipids of almost all animals that feed in or from the sea – ragworms, mussels, oysters, shrimps, crabs, fish, seals, whales, polar bears, gannets, albatrosses, and penguins – are rich in *n*-3 polyunsaturated fatty acids. Such mixtures of triacylglycerols rich polyunsaturates are often called "fish oils," though they are not unique to fish, and the basic fatty acids from which they are derived are rarely produced by fish. The lipids originate in algae but may be modified by animals as they pass up the food chain, eventually reaching top predators such as large fish, sea-birds, and marine mammals. Among the fish used as human food, cool-water species such as herring, sprats, capelin, mackerel, and halibut usually store more lipid than their tropical relatives, because at high latitudes, the supply of most kinds of food changes seasonal and many species routinely live on their fat reserves for weeks

or months. The triacylglycerol fatty acids of large cold-water fish include more monounsaturates than polyunsaturates because they come from waxes that were in the planktonic organisms that they eat. Fish such as sardines and anchovies live in warmer waters where the planktonic invertebrates rarely contain waxes, so their lipids are a rich source of *n*-3 polyunsaturated fatty acids for human consumers.

Some of the food of animals living in lakes and rivers is derived from aquatic algae, but detritus from terrestrial plants and the animals that feed on them also make a substantial contribution, so freshwater fish usually have more fatty acids of the *n*-6 series than oceanic species. The fatty acid composition of the predatory fish follows that of their diet: "farmed" salmon fed on chow made from scraps of slaughtered livestock and other materials originating from the land acquire a correspondingly "terrestrial" fatty acid composition.

As well as being chemically distinct, lipids of marine origin differ in the isotopic composition of the carbon atoms. Slightly different proportions of ^{13}C and ^{12}C are taken up by marine algae that obtain their carbon in solution as bicarbonate, and terrestrial plants that take in carbon dioxide straight from the atmosphere, so organic materials of marine origin contain about 5–7 parts per thousand more of the heavier isotope (^{13}C) than similar molecules from animals that have eaten terrestrial foods.

This effect can be used to investigate the dietary habits of individual animals and populations. Storage fatty acids and cholesterol are particularly suitable because they pass intact up the food chain, are chemically quite stable and only slowly degraded by microbes, so can be detected in excreta and animal remains as well as in biopsy samples. The isotopic and chemical composition of triacylglycerol fatty acids can indicate how much of the diet is derived from marine or terrestrial sources. This technique is particularly useful for widely ranging species with broad, locally variable diets, such as arctic foxes, that are not easily observed in the wild (Pond and Gilmour, 1997). Material from the marrow of buried human and other animal bones, and in coral fragments has been used to study natural diets and food chains.

Mammalian adipocytes can selectively take up and release certain fatty acids (there are no data on those of other vertebrates), making possible seasonal accumulation of triacylglycerols with physical properties appropriate to hibernation (Frank *et al.*, 2008). There are site-specific differences in fatty acid composition of storage triacylglycerols: except in ruminants, the adipocytes around lymph nodes, and to a lesser extent elsewhere in node-containing depots, sequester more polyunsaturates that are essential to the nutrition of the immune system than large nodeless depots. Such site-specific differences must be considered when selecting tissue samples for the study of diet.

Adipose Tissue as Thermal Insulation

The notion that the superficial white adipose tissue insulates the body against heat loss is one of the most firmly established of all dogmas in twentieth century biology, but there is very little evidence that supports it. The hypothesis can be tested by determining whether its distribution and anatomical arrangement have evolved to carry out such a role more efficiently in animals that need more thermal insulation, because they remain active all winter in cold climates, or habitually swim in very cold water. The mammalian order Carnivora occur in a wide variety of habitats from hot tropics (e.g., desert foxes, jackals) to the Arctic (e.g., polar bears, wolverines, arctic foxes) and almost all are predators, at least to some extent. Carnivora occur in an enormous range of sizes, from stoats and weasels to bears and tigers that can be several thousand times larger, and are therefore suitable for investigating whether species differences in the relative abundance of superficial adipose tissue are determined by habits and habitat or by body size (Pond and Ramsay, 1992).

Figure 5 shows some measurements of the masses of all the intra-abdominal and superficial adipose tissue from various temperate, tropical and arctic carnivores. To ensure comparability, the sample included only specimens that were of similar fatness to the bears. There is a fair amount of inter-individual variation in the partitioning of adipose tissue between internal and superficial depots but overall the mass of superficial depots increases in simple proportion to body mass. In contrast, the intra-abdominal depots become proportionately smaller in larger mammals, whether the animals are native to the Arctic or to warmer habitats.

The surface area over which this larger amount of superficial adipose tissue is spread is smaller in larger animals relative to their volume or mass. The body surface area scales as body mass$^{0.66}$, that is, it becomes proportionately smaller with increasing body size. With a smaller area to cover, the superficial depots are thicker in larger animals, even if the specimens are not fatter (i.e., if the ratios of the masses of lean and fat tissues are unchanged). The superficial adipose tissue is 10 times thicker in a 500-kg bear than in a 500-g rat or weasel from this effect alone, quite apart from the fact that in larger animals, a greater proportion of the adipose tissue is superficial, and bears are often fatter than rat-sized animals (Pond and Ramsay, 1992).

The measurements from the wild polar bears fall very near to the lines fitted to the data obtained from tropical and temperate-zone carnivores, indicating that the partitioning of adipose tissue between internal and superficial sites arises entirely from the fact that polar bears are very large and lay down large quantities of fat, and has nothing to do with adaptation to thermal insulation. There is no evidence for reorganization of adipose tissue in the smaller species of mammals that have evolved to become aquatic, for example, otters (family Mustelidae) have many adaptations of the trunk, tail, and sense organs to aquatic life, but the arrangement of their adipose tissue is not different from that terrestrial mustelids, stoats, weasels, wolverines, and badgers. Otter fur is oily and less compressible than that of the terrestrial species, so it retains its insulating properties when in water.

These principles are also true of birds. Just before migration or breeding fasts, the superficial depots enlarge greatly, and may become several millimeters thick. They also expand laterally so the depot around the wishbone extends over the anterior part of the flight muscles and along the neck. In swimming animals, indefinite expansion of the abdomen is not compatible with streamlining while in water, but a sleek profile can be maintained if the additional fat is spread out over much of the body.

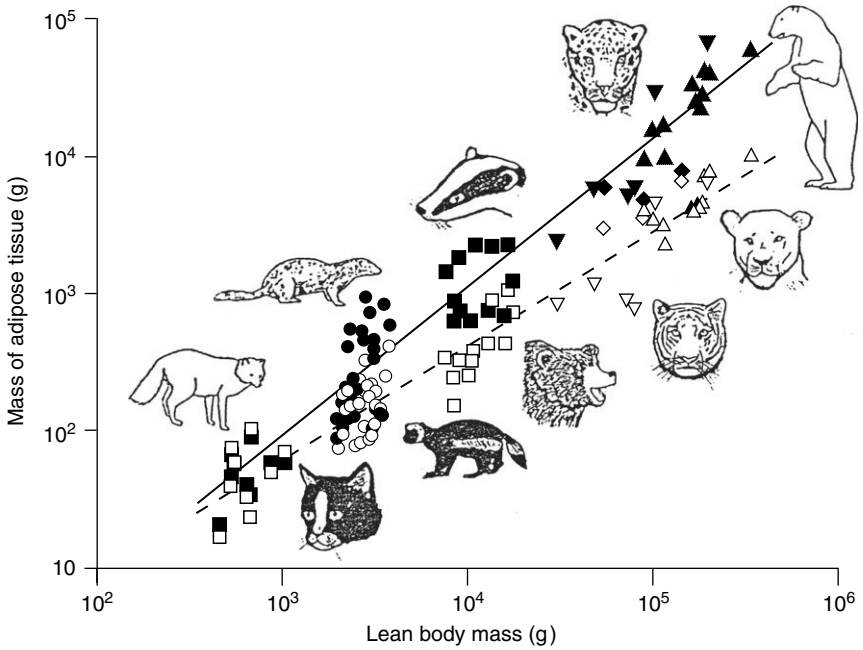


Figure 5 Allometric comparison of the masses of superficial (solid symbols) and intra-abdominal (open symbols) adipose depots in various species of the order Carnivora. Each point represents a single individual. Squares: Mustelidae (mink, badgers, and wolverines); diamonds: Felidae (cats, jaguars, tigers, and a lion); circles: Canidae (arctic foxes); and triangles: Ursidae (inverted triangles: brown bears). The lines are fitted to all the data *except* those from the polar bears (upright triangles) and fit the equations: superficial depots α (lean body mass) 1.05, $r^2 = 0.867$; intra-abdominal depots α (lean body mass) 0.81, $r^2 = 0.873$. Data from Pond CM and Ramsay MA (1992) Allometry of the distribution of adipose tissue in Carnivora. *Canadian Journal of Zoology* 70: 342–347 and additional sources. Illustration from Pond CM (1998) *The Fats of Life*, 344 pp. Cambridge: Cambridge University Press.

This analysis highlights some pitfalls of using measurements of the thickness of superficial adipose tissue as a means of estimating fatness. Fatness is a ratio of masses, the mass of energy stores as a percentage of total body mass, but adipose tissue thickness is a linear measurement. With increasing body mass, the superficial adipose tissue becomes thicker for animals of the same total body composition because the ratio of their surface area to mass decreases. The superficial adipose tissue becomes thicker in larger animals and depots that are adjacent but not contiguous in small species and the juveniles of large species merge in large adults to form a continuous layer, albeit of variable thickness. An almost continuous layer of superficial adipose tissue can create the impression that the specimen is obese, even when the layer represents only a small fraction of the total body mass.

The lactation habit rather than thermoregulation may be the main reason for the massive development of the superficial adipose depots in mammals. Reptiles are usually fattest just before maturation of the eggs begins, and the intra-abdominal fat bodies are much depleted by the time the clutch is ready to be laid. Heavily gravid females are often fasting, so the empty, inactive gut frees up yet more space inside the abdomen. In contrast to egg-lay vertebrates, the demands of lactation immediately follow those of gestation (Thompson and Speake, 2006; Pond, 2011). Most mammals feed normally while they are pregnant and accumulate lipid in preparation for lactation. The reproductive organs and the gut cannot be excluded from the abdomen, but adipose tissue can. The subcutaneous depots may simply be the most convenient place to put a lot of adipose tissue (Klein *et al.*, 2007).

In pinnipeds, cetaceans, and sirenians, the cellular and gross anatomy of adipose tissue is substantially altered to form blubber. Heat conducts through adipose tissue at only a little less than half the rate of muscle or stagnant water and much faster than through fur or feathers. So blubber has to be thick to insulate thoroughly and is satisfactory as the sole form of insulation only for large animals with habits and habits for which weight is not a major problem. Consequently, most small- and medium-sized seals retain fur as additional insulation and extant sirenians and the smallest cetaceans, porpoises, and dolphins, are restricted to temperate and tropical regions. Many of the great whales, for example, humpbacks and fin whales, which feed in polar seas, migrate to warmer waters to give birth to calves that are only a twentieth of their size. Many such places provide little or no food for adult whales, so the mothers fast for several weeks while lactating, drawing on their lipid reserves. The whales do not return to cooler, food-rich waters until their calves have grown large enough and fat enough for blubber insulation to be effective.

Seals are born on beaches or ice floes and most species do not go to sea until they are several weeks old. Neonates have thick coats of fluffy fur but little adipose tissue anywhere, although they fatten rapidly while suckling their mothers' rich, creamy milk. At about the time the seal pups are weaned, the neonatal fur is replaced by shorter, stiffer fur. But fur of some sort continues to make an essential contribution to insulation in all seals except the adults of the largest species, the elephant seals (*Mirounga* spp.). In walruses (*Odobenus rosmarus*), the hair is very sparse throughout life and at birth, the single pup is very large compared with its mother, up to twice the size of

neonatal elephant seals, although as adults, the latter are much larger.

Both these very large pinnipeds occur in cold seas, *Odobenus* in the Arctic Ocean, and one species of *Mirounga* in the northern Pacific including around Alaska, and the other in the Southern Ocean. The skin of walruses hauled out on beaches basking in sunshine is noticeably pink, due to the profuse flow of blood through it and the underlying blubber, which can be up to 15 cm thick. But walruses are a ghostly pale gray when in the sea or exposed to cold weather because the blood is withdrawn from the skin and blubber to the muscles and other warm internal tissues. Blubber can work as an adjustable insulator in this way because blood flow through it can be shut down almost completely for long periods without ill-effects.

If necessary, marine mammals can be very fat. As in the case of penguins, body bulk is only a slight impediment to swimming, and seals and whales can maintain near neutral buoyancy over a wide range of body compositions by adjusting the volume of air in the lungs. The maximum amount of blubber that whales or seals can have without risking overheating during vigorous swimming depends on body shape, the temperature of the water they are living in and how much heat they generate. Marine mammals dissipate excess heat by increasing blood flow through the blubber-free skin of the flippers and tail. On hot days, seals on land sometimes hold up a flipper to the breeze to accelerate cooling. A thickness of about 20 cm seems to be the upper limit. Blubber of that thickness apparently does not hold enough lipid to meet the energy storage needs of the very largest whales, fin whales, blue whales and their relations. They have adipose tissue in the mesentery that supports the gut and around the kidneys, and intermuscular adipose tissue in certain muscles, just as terrestrial mammals do.

Energy Storage in the Life History

Optimizing Energy Stores

Healthy animals living wild usually have around 4–8% dissectible adipose tissue under ideal nutritional conditions but may transiently be much fatter just before some special situation such as breeding, migration or hibernation (Pond, 1998). Many animals that are below 4% dissectible adipose tissue are probably having trouble finding enough food, though not yet in danger of starvation, but for others, having much less adipose tissue seems to be normal. Some common wild animals, such as rabbits, hares, moles, and foxes, seem to live and breed satisfactorily with less than 1% of the body mass being adipose tissue.

The concept of costs and benefits to accumulating fat stores has led to much theorizing about the ecological conditions that might define optimum fatness, and how they change during the animal's life history. Some computer models incorporate data about the relationship between body composition and energy utilization and simulate real situations quite accurately, predicting how much adipose tissue is appropriate for birds engaged in certain activities under certain conditions (Witter and Cuthill, 1993; Pennycuick and Battley, 2003). But direct measurements of many species in the wild show that there is often quite a bit of variation in fatness among members of the same species at the same times of year, which cannot easily be linked to habits or the ecological situation in which they live.

Polar bears are top carnivores living in the Arctic, where the climate is not only harsh, but extremely variable from time to time and place to place. These bears are among the very few large carnivores to have huge home ranges but no proper territories. They travel long distances between areas where ice conditions make seal hunting feasible, eating large but highly irregular meals. Moderate levels of obesity seem to be essential to such habits. A sample of young adult bears studied in the Canadian Arctic in early November, when the sea-ice was just beginning to freeze over and seal hunting was again becoming efficient, were between 10 and 21% dissectible fat, which amounted to 73 kg of adipose tissue for the largest bear whose total body mass was over 400 kg. Every year a few individuals, adults as well as newly weaned juveniles, are found in a severely emaciated condition, sometimes attributable to injury, but often from failure to catch seals or find carrion. Their superficial adipose tissue, which over the rump might be more than 10 cm thick in a successful hunter, has shrunk to a flimsy sheet of watery tissue. Their muscles are wasted, so the hips are narrow and the legs thin, and at first glance, they look more like a dog than a bear.

Average fatness can be very different in different localities (and probably also from year to year, though there is less information on this topic), even for wild animals of the same species. For example, the common red fox is usually lean where food is available regularly but predators abound, even when, as now often happens, it frequents towns and eats the remains of people's fatty meals. Populations living in Scandinavia and parts of Canada, where the food supply is very unpredictable, and predators and potential competitors are scarce, are substantially fatter than those elsewhere in Europe.

Birds that are fat for only part of the year may have habits that reduce the risks associated with being heavy. They may live and feed in flocks or herds where the "safety in numbers" principle applies, or spend time only in places where predators are few or absent. Small birds such as great tits that feed only by daylight become measurably heavier toward evening and lose weight during the hours of darkness, especially on long, cold nights. How much fat they carry from day to day correlates with the likelihood of obtaining food, which depends on both the presence of suitable prey, and being senior enough in the "pecking" order to get sufficient access to it, and on the prevalence of predators. Great tits in Britain became significantly heavier when their principal predator, the sparrow-hawk, disappeared in the late 1950s and 1960s, as a result of poisoning from agricultural insecticides. When improved farming practices allowed sparrow-hawks to reestablish themselves in western Britain in the 1970s, the mean body mass of great tits declined, though that of wrens, a species rarely caught by hawks, remained unchanged.

Migration

In contrast to walking, running and swimming, the muscle power required to stay airborne by active flapping flight

increases disproportionately with increasing body mass. Other factors being equal, smaller birds can generate substantially more power than is needed to keep airborne, so they can carry proportionately more fuel, up to 50% of the body mass if necessary. Larger wings make flying energetically more efficient, so the longest nonstop journeys are undertaken by medium-sized birds, such as sandpipers, knots, turnstones, curlews, and godwits of body mass around 100–800 g (Pennycuick and Battley, 2003).

Small birds of body mass around 100 g that set out with adipose tissue triacylglycerols amounting to 100% of lean tissue mass (i.e., half the body mass is fuel) can fly for 3–4 days nonstop or 3000–4000 km (depending greatly on wind direction and other weather conditions). Flying time can be prolonged by slowing down, or shortened by speeding up or by flying into a headwind. In general, larger birds can fly faster, and so get there sooner, but because they can carry less fuel, they often cannot go as far as medium-sized birds. Birds with a lean body mass of around 100 g lose about 0.7% of the body mass per hour while migrating, but larger birds of around 500 g flying similarly loaded lose 1–1.5% h^{-1}.

The rate of energy expenditure decreases as the birds become lighter through oxidization of fat. As they set out fully loaded, each gram of triacylglycerols takes them less than half as far as at the end of their journey, when their fuel is almost exhausted, because flying with half the body mass as fat uses three times as much muscle power as flying on "empty." A bird that sets out with half its total body mass as fuel and flies to exhaustion uses 40% more energy for the whole journey than one that sets out with a fuel load of 10% of the body mass, flies as far as it can, stops to feed, and continues on the next lap, repeating the cycle until it reaches its destination. Such a journey is slower because even with unlimited food, birds cannot fatten faster than about 10% of the body mass per day, usually nearer 3–6% day^{-1}. So increasing the body mass by 25% takes at least a week, 3–4 weeks to accumulate maximal fuel reserves.

The amount of fuel a large bird can carry is usually limiting, and most journeys begin with a full load and end with little to spare. Whooper swans (*Cygnus cygnus*) that fly between northwest Scotland and their breeding grounds in central Iceland are among the largest of all migratory birds. The swans set out with adipose tissue triacylglycerols amounting to about 20–25% of their lean body mass, the largest load they can take off with. If they encounter bad weather or adverse winds, they settle on the sea and wait, sometimes for 30 h or more, until travelling conditions improve enough to continue their journey.

Hibernation

Hibernation and its hot weather equivalent, aestivation, are periods of inactivity and seclusion, usually accompanied by cooler than normal body temperature, which enable animals to pass through seasons when food is scarce or inaccessible. Hibernation was among the first physiological states in which adipose tissue of wild mammals, reptiles, and amphibians was studied thoroughly. Hibernators depend on their fat reserves while cold and to fuel rewarming but reclaiming the stores is not as straightforward as it first seemed.

The enzymatic processes involved in fasting and starvation are essentially similar to those of slow exercise, but there is a critical difference: during exercise, the body is warm, often slightly warmer than when sedentary, but in hibernation, the body temperature is low, sometimes falling by 35 °C to close to 0 °C. The fatty acid composition of triacylglycerols is largely irrelevant to their role as fuel when animals are warm, but it is crucial for their use during hibernation because enzymes do not work on solidified fats.

Experiments on captive chipmunks (*Eutamias amoenus*) and golden-mantle ground squirrels (*Spermophilus lateralis*) show that they enter hibernation more readily, remain cooler for longer and are better able to survive long winters when plenty of polyunsaturated lipids are included in their diet during the weeks preceding hibernation than when they are fed on saturated fats of similar calorific value. Unsaturated fatty acids may lower the melting point of the triacylglycerols and the membrane phospholipids, enabling them to remain more fluid, and hence retain their proper affinity for carrier molecules and enzymes, at cooler temperatures. The chemical composition of the storage lipids, together with other aspects of adipose tissue, is thus adapted to the physical conditions and its role in whole-body metabolism (Frank and Storey, 1995; Frank *et al.*, 2008).

Squirrels obtain many such unsaturated fatty acids from the seeds and other plant parts that they eat. Hibernation is an active, physiologically controlled process, and especially in mammals, metabolic preparations can be identified days or weeks before the animal actually allows its body to cool. As the weather becomes cooler and the days shorten in autumn, squirrels, and other hibernatory rodents actively seek nuts and other foods that contain these lipids. The woodchuck or marmot (*Marmota flaviventris*) selectively retains linoleic acid (C18:2) before hibernation: the saturated fatty acids are released and oxidized by the muscles, liver, etc. while the animal is warm and active, but the polyunsaturates remain in the adipose tissue, for use when the body is cold. Like other mammals, the squirrels cannot add more than one double bond to most kinds of long chain fatty acids, so they depend on the increased availability in autumn of seeds rich in polyunsaturated lipids. Failure of a seed crop could prevent successful hibernation and thus lead to death from cold or starvation, even if plenty of other foods were available. Analysis of biopsies of gonadal and inguinal adipose tissue from alpine marmots (*Marmota marmota*) throughout the year indicate that selective release of certain fatty acids allows active regulation of the composition of storage triacylglycerols (Cochet *et al.*, 1999). The storage lipids remain fluid during deep hibernation and in the euthermic state by maintaining a high proportion of monunsaturates, both by retention of those from the diet and, if necessary, by synthesis.

Sleeping undisturbed in a cool, secluded place uses storage materials only very slowly: small tortoises may lose less than 5%, and rarely lose over 15%, of the body mass during 5 months of hibernation at about 5 °C. Many reptiles and amphibians emerge from hibernation with a surprisingly large proportion of their lipid reserves still remaining and regain what they lost over the winter in a few weeks of feeding. These stores are often important for fuelling mating and

egg production. The less lipid a hibernator uses during the winter, the more it has left to fuel breeding the following spring.

Reproduction

The formation of large, yolky eggs that nourish the embryo until it is at an advanced stage of development is a very ancient means of reproduction among vertebrates. Most sharks, skates and rays, whose ancestors were abundant in Devonian seas 400 million years ago, produce such eggs, which hatch into miniatures of the adults, capable of feeding for themselves.

One third of the fresh mass of the yolk of birds' eggs is storage fats, a proportion that may represent the maximum that can form a stable emulsion, holding the lipids in a state that the embryo or its mother can handle. The relative size of the yolk and the "white" part of eggs differs greatly between species, being lowest in species that feed their young on the nest for some time after hatching, such as pelicans (17% of the mass of the egg), gannets (18%), and crows (19%). Yolk content is highest in species whose relatively large eggs hatch to produce large, mature chicks that can walk well and feed themselves. The yolkiest eggs are those of kiwis (65%), megapodes, also known as brush turkeys or mound-builders (over 51%), ducks (40%) and, of course, species of the order Galliformes, which includes turkeys, peafowl, pheasants, guinea-fowl, quail, and domestic poultry.

In developing domestic chickens, the cells lining the gut start "eating" droplets of yolk around the twelfth day of incubation, and pass its lipids into the blood as lipoproteins. At the same time, the adipose tissue matures and starts to produce lipoprotein lipase in large quantities, enabling it to take up the yolk lipids. Both the number and the mean size of adipocytes increase until at least the nineteenth day of incubation, and continue after hatching on the twenty-first day. The adipose tissue serves as more than just a temporary store: its adipocytes and the lipoprotein lipase they produce discriminate between triacylglycerols of different fatty acid composition, and thereby "ration" the embryo's irreplaceable lipid provisions to ensure normal development (Speake *et al.*, 1998; Speake and Wood, 2005).

Lipids are the main agents of transfer of energy from mother to offspring in most vertebrates and many other kinds of animals, including insects. Mammals are somewhat unusual in that their eggs are very small and contain almost no yolk because soon after fertilization they attach themselves to the lining of the uterus, forming a placenta, through which the embryo obtains all the nutrients it needs. Lipids cross the placenta and enter the fetal blood but most seem to be deployed in structural roles, particularly the formation of membranes of the brain and nervous system, rather than being broken down for energy production. The situation changes abruptly at birth, when the neonate starts to suckle. The milk of most mammals is rich in triacylglycerols but contains only small quantities of carbohydrate as lactose. During suckling, carbohydrates and lipids almost swap the roles they had in the fetus: suckling mammals convert lactose into glucose, most of which is incorporated into proteins or other structural materials and lipids become the main source of fuel.

The lactation habit has the enormous advantage that it emancipates mammals from the limitation of breeding only where and when suitable food is available for the young, and permits rapid growth of the neonate (Pond, 2011). Lactation combined with storage is the main reason for the rarity of long-distance migration in relation to reproduction among mammals compared to birds and other vertebrates.

For some species, food supplies are too erratic, or obtaining them requires too many long absences from the nest, for lactation to be fuelled just by eating more. Milk production has to draw on the body's reserves of lipid from adipose tissue, calcium and other minerals from the skeleton and protein from it is not clear where, probably mainly muscle and liver. So many female mammals fatten during pregnancy, and deplete their adipose tissue lipids during lactation. Soon after birth, changes to hormones, their receptors and enzymes give the mammary glands preferential access to circulating lipids and act on adipose tissue, disabling its uptake of triacylglycerols from the blood and prompting fatty acid release. These changes direct fatty acids from the adipose tissue to the mammary gland, where they are incorporated into milk triacylglycerols. Although storage lipids contain only long-chain fatty acids, the enzymes of the mammary gland readily esterify medium-chain and even short-chain fatty acids into triacylglycerols. The precursors of milk of adequately fed mammals come straight from the gut without going via the adipose tissue so may include fatty acids that are never appear in storage triacylglycerols. But during fasting, the adipose tissue becomes the major source of fatty acids, so the milk triacylglycerols from starving animals always contain more long-chain fatty acids.

The composition of fatty acids in milk triacylglycerols is thus very variable in many mammals, depending on what the mother has just eaten, and how much is derived from her adipose tissue, or obtained directly from the diet. The contribution of storage lipids is important for wild animals where there are lipid-soluble pollutants such as dichlorodiphenyltrichloroethane (widely used as an insecticide from the 1940s to the 1970s) and polychlorinated biphenyls (formerly used as components of large batteries and other electrical apparatus). Such contaminants eaten during lactation may go straight into the milk, and in fasting mothers such as bears, those that have accumulated in adipose tissue during the previous months get released with the fatty acids and so transferred into the milk. These effects can produce alarmingly high concentrations of toxins in the offsprings' tissues, permanently impairing growth and development. Lipid storage in adipose tissue is central to this breeding strategy, enabling mothers to store enough materials to synthesize milk at a high rate, and the young to take full advantage of the nutrients their mother provides during the often brief period during which they are together. Its role is enhanced in large species that live in highly variable habitats. Polar bears (Figures 4 and 5) can carry so much storage tissue around because they have no natural predators (other than humans) and hunt by stealth and skill rather than by speed and agility. The bears' massive adipose tissue enables females to separate feeding and mothering to a greater extent than is possible for almost all other warm-blooded animals, thereby equipping them to breed successfully in spite of the arctic vagaries of food supplies.

Marine mammals can afford to be fatter, at least for short periods, than terrestrial species because they do not attempt to move far or fast on land, and so are less constrained by the

mechanical limitations to fatness. There are no large terrestrial predators in the Antarctic, so seals breeding on beaches or on sea-ice can afford to suckle their young for several weeks. But those living in and around the Arctic are obliged to share their habitat with polar bears and have evolved ways of transferring lipids from mother to young at a spectacular rate. Harp seals (*Phoca groenlandica*) that breed on pack-ice throughout the Canadian and Russian Arctic are only 3% by weight lipid at birth but reach 47% lipid at weaning 13 days later. The single pup drinks an average of 3.7 kg of milk each day, gaining weight at an average rate of 2.3 kg day^{-1}. Transfer of lipid from mother to offspring gathers pace as the young matures: the lipid content of the milk increases from 36% at the start of lactation, to 57% just before weaning.

Lactation in the hooded seal (*Cystophora cristata*) which breeds on temporary ice-floes in the North Atlantic is even more compressed. The single pup weighs about 22 kg at birth, large relative to adult size, but is only slightly fatter than harp seal pups. It almost doubles its body mass to around 43 kg in only 3–5 days of suckling. Over 70% of the increase in body mass is blubber, which holds enough lipid to sustain the pup during a postweaning fast that can last several weeks. During the 4 days of lactation, the mother's body mass declines from 179 to 150 kg, over 80% of the loss being from the blubber as she synthesizes milk at 2.5–6 times the rate of other seals that have been studied, probably a record for any mammal. Such haste may be necessary to minimize the risk that mother and pup become separated by storms and currents that move ice-floes around, as well as that of predators.

Most mammals are weaned with storage tissues replete, especially carnivores and other species for which skill and experience are essential for finding food. Much of the triacylglycerols transferred to suckling seals is stored in their own blubber, and utilized during the long period between weaning and the pups becoming proficient at finding food for themselves.

See also: Arctic Terrestrial Ecosystems. Birds, Biodiversity of. Defenses, Ecology of. Impact of Extreme and Infrequent Events on Terrestrial Ecosystems and Biodiversity. Insects, Overview. Mammals, Biodiversity of. Migration. Plant–Animal Interactions. Plant Biodiversity, Overview. Plant Sources of Drugs and Chemicals. Vertebrates, Overview

References

Canavoso LE, Jouni ZE, Karnas KJ, Pennington JE, and Wells MA (2001) Fat metabolism in insects. *Annual Review of Nutrition* 21: 23–46.

Cannon B and Nedergaard J (2004) Brown adipose tissue: Function and physiological significance. *Physiological Reviews* 84: 277–359.

Cochet N, Georges B, Meister R, Florant GL, and Barre H (1999) White adipose tissue fatty acids of alpine marmots during their yearly cycle. *Lipids* 34: 275–281.

Crisan M, Casteilla L, Lehr L, *et al.* (2008) A reservoir of brown adipocyte progenitors in human skeletal muscle. *Stem Cells* 26: 2425–2433.

DiAngelo JR and Birnbaum MJ (2009) Regulation of fat cell mass by insulin in *Drosophila melanogaster*. *Molecular and Cellular Biology* 29: 6341–6352.

Frank CL, Karpovich S, and Barnes BM (2008) Dietary fatty acid composition and the hibernation patterns in free-ranging arctic ground squirrels. *Physiological and Biochemical Zoology* 81: 486–495.

Frank CL and Storey KB (1995) The optimal depot fat composition for hibernation by golden-mantled ground squirrels (*Spermophilus lateralis*). *Journal of Comparative Physiology B* 164: 536–542.

Geiser F and Learmonth RP (1994) Dietary fats, selected body temperature and tissue fatty acid composition of agamid lizards (*Amphibolurus nuchalis*). *Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology* 164: 55–61.

Gorissen M, Bernier NJ, Nabuurs SB, Flik G, and Huising MO (2009) Two divergent leptin paralogues in zebrafish (*Danio rerio*) that originate early in teleostean evolution. *Journal of Endocrinology* 201: 329–339.

Klein J, Permana PA, Owecki M, *et al.* (2007) What are subcutaneous adipocytes *really* good for…? *Experimental Dermatology* 16: 43–70.

Kvist A, Lindstrom A, Green M, Piersma T, and Visser GH (2001) Carrying large fuel loads during sustained bird flight is cheaper than expected. *Nature* 413: 730–732.

Landys-Ciannelli MM, Piersma T, and Jukema J (2003) Strategic size changes of internal organs and muscle tissue in the Bar-tailed Godwit during fat storage on a spring stopover site. *Functional Ecology* 17: 151–159.

Lees J, Nudds R, Stokkan KA, Folkow L, and Codd J (2010) Reduced metabolic cost of locomotion in Svalbard rock ptarmigan (*Lagopus muta hyperborea*) during winter. *Plos One* 5: e15490.

Manfruelli P, Reichhart JM, Steward R, Hoffmann JA, and Lemaitre B (1999) A mosaic analysis in *Drosophila* fat body cells of the control of anti-microbial peptide genes by the Rel proteins Dorsal and DIF. *EMBO Journal* 18: 3380–3391.

Murphy DJ (1994) Biogenesis, function and biotechnology of plant storage lipids. *Progress in Lipid Research* 33: 71–85.

Murphy DJ and Vance J (1999) Mechanism of lipid-body formation. *Trends in Biochemical Sciences* 24: 109–115.

Pennycuick CJ and Battley PF (2003) Burning the engine: a time-marching computation of fat and protein consumption in a 5420-km non-stop flight by great knots, *Calidris tenuirostris*. *Oikos* 103: 323–332.

Pond CM (1998) *The Fats of Life*, 344 pp. Cambridge: Cambridge University Press.;

Pond CM (2009) Paracrine provision of lipids in the immune system. *Current Immunology Reviews* 5: 150–160.

Pond CM (2011) The evolution of mammalian adipose tissue. In: Symonds ME, (ed.) *Adipose Tissue Biology.* Heidelberg: Springer Science.

Pond CM and Gilmour I (1997) Stable isotopes in adipose tissue fatty acids as indicators of diet in arctic foxes (*Alopex lagopus*). *Proceedings of the Nutrition Society* 56: 1067–1081.

Pond CM and Mattacks CA (1985) Body mass and natural diet as determinants of the number and volume of adipocytes in eutherian mammals. *Journal of Morphology* 185: 183–193.

Pond CM, Mattacks CA, and Prestrud P (1995) Variability in the distribution and composition of adipose tissue in arctic foxes (*Alopex lagopus*) on Svalbard. *Journal of Zoology, London* 236: 593–610.

Pond CM and Ramsay MA (1992) Allometry of the distribution of adipose tissue in Carnivora. *Canadian Journal of Zoology* 70: 342–347.

Poulos SP, Hausman DB, and Hausman GJ (2010) The development and endocrine functions of adipose tissue. *Molecular and Cellular Endocrinology* 323: 20–34.

Puri V, Konda S, Ranjit S, *et al.* (2007) Fat-specific protein 27, a novel lipid droplet protein that enhances triglyceride storage. *Journal of Biological Chemistry* 282: 34213–34218.

Speake BK, Murray AMB, and Noble RC (1998) Transport and transformations of yolk lipids during development of the avian embryo. *Progress in Lipid Research* 37: 1–32.

Speake BK and Wood NAR (2005) Timing of incorporation of docosahexaenoic acid into brain and muscle phospholipids during precocial and altricial modes of avian development. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* 141: 147–158.

Thompson MB and Speake BK (2006) A review of the evolution of viviparity in lizards: structure, function and physiology of the placenta. *Journal of Comparative Physiology B - Biochemical Systemic and Environmental Physiology* 176: 179–189.

Watts JL (2009) Fat synthesis and adiposity regulation in *Caenorhabditis elegans*. *Trends in Endocrinology and Metabolism* 20: 58–65.

Weber JM (2009) The physiology of long-distance migration: extending the limits of endurance metabolism. *Journal of Experimental Biology* 212: 593–597.

Witter MS and Cuthill IC (1993) The ecological costs of avian fat storage. *Philosophical Transactions of the Royal Society of London B* 340: 73–92.

Stress, Environmental

John Cairns Jr., Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acute An exposure to an environmental stress that is brief in relation to the temporal scale of the biological system exposed.
Biosphere All the plants and animals on Earth, together with the habitat they need for long-term survival.
Biotic impoverishment The decline of a species to the point of having little or no ecological significance.
Chronic An exposure to an environmental stress that is comparable in duration to the temporal scale of the biological system exposed.
Ecosystem services The structures and functions of natural biological systems which directly or indirectly support human life.
Environmental stress An action, agent, or condition that impairs the structure or function of a biological system.
Function The performance of a biological system as a rate.
Structure The number, kinds, and arrangement of component parts at one point in time.
Tipping point When a complex system goes into disequilibrium and is replaced by a new system in evolutionary time.
Uncertainty Imperfect knowledge concerning the current or future state of a system under consideration; a component of risk resulting from imperfect knowledge of the degree of hazard or of its spatial and temporal pattern of expression.

What is Environmental Stress?

Three components are involved in the relationship that defines environmental stress. First, there is the environmental stress itself as defined previously. However, the environmental stress can only be defined in reference to its interaction with some biological system. Therefore, there must be a receptor – a biological system that is exposed to the environmental stress. Finally, there must be an adverse response – a particular structure or function of the receptor that is changed by exposure to the environmental stress to the detriment of that system. If the survival of that biological system (or another) is not threatened by the change, then there is no environmental stress.

Types of Environmental Stress

Natural versus Anthropogenic

Environmental stress can be either natural or anthropogenic (i.e., resulting from human actions) in origin. Many environmental stresses, such as most hurricanes, droughts, floods, and fires are a periodic feature of life on Earth. In contrast, environmental stresses such as the production and release of new chemical compounds and large-scale land-use changes result directly from human actions. Ironically, the suppression of natural environmental stresses such as fires and floods can also be a source of stress to biological systems resulting from human actions. Some species have adapted to periodic disturbance and cannot continue without them. For example, some seeds will germinate only after exposure to the high temperatures of a fire. However, such adaptations take time to evolve. Other stresses, such as the meteor that probably wiped out 70% of earth's species 65 million years ago, occur too quickly and intensely to result in adaptation.

Natural and anthropogenic stresses often have common components. For example, both hurricanes and wood harvesting result in downed trees. However, as a result of recent major increases in human population, technological capabilities, and standard of living globally, the amount of anthropogenic environmental stress has increased greatly. Anthropogenic environmental stress existed even 50,000 years ago, when fires set to aid hunters are thought to have altered the landscape in central Australia (Flannery, 1999). Since that time, human population increased slowly to 1 billion people in 1804 and then rapidly to 6 billion people in 1999, and the population is expected to increase by another 1 billion people every 12–15 years. The cumulative environmental stresses resulting from these exponential increases in human population are exacerbated by technologies that expand the character and scope of changes humans can make to their environments. Both the agricultural revolution (about 10,000 years ago) and the industrial revolution (about 200 years ago) expanded the types of anthropogenic stresses on the environment. In addition, increased affluence for people throughout the world increases the environmental stress on natural systems by increasing the per capita human use of natural systems. Humankind's collective ecological footprint (i.e., the amount of earth's surface required to produce the resources used and to assimilate the wastes produced) is rapidly increasing (Rees, 1996) at the same time that productive land is decreasing through erosion, salinization, and unsustainable land-use changes. The ecological footprint for an average person can range from 0.4 ha required to provide for the lifestyle of one person in India, where the level of affluence is quite modest, to 5.1 ha for a person in the US. If the lifestyle of every person living in 1996 was elevated to that of a typical North American, an additional two earths would be needed to provide the surface area required (Rees, 1996).

Characterization of Environmental Stress

Environmental stresses of both natural and anthropogenic origin can be characterized on the basis of their spatial distribution, temporal distribution, intensity, and novelty (Kelly and Harwell, 1989). Spatial distribution of a stress describes its geographic extent and pattern. One basic concern is the size of the area affected by the stress, i.e., is the stress local, regional, or global in extent? Some kinds of environmental stress have an intrinsic spatial scale: A poorly dispersed chemical spill may be quite local in its effects, whereas air pollution can affect entire regions. Some stresses will be consistently spread over an area, whereas others will occur irregularly in patches. Cumulative environmental stresses, in which many individual small patches can join together to have larger impacts at a larger scale, have been documented. In one example at the local level, the cumulative loss of small wetland areas within a watershed had demonstrable adverse effects on water quality (Johnston *et al.*, 1988). The magnitude of the biological responses to a stress will often be modified by the spatial distribution of environmental stress, for example, whether key features in the environment such as riverbanks or fencerows are affected or whether similar habitat patches nearby are left unaffected. The field of landscape ecology deals with these factors.

The temporal distribution of a stress describes its frequency and duration. Some stresses, such as chemical spills, are one-time occurrences. Others, such as the winter season or wildfires, can be expected to reoccur either at predictable or at unpredictable intervals. Some stresses rapidly ameliorate; others remain for long periods of time. The terms ''acute'' and ''chronic'' are used to describe the duration of a stress. Acute refers to stresses of short duration, whereas chronic refers to stresses that last longer in relationship to the duration of the biological system they affect. In some cases, the timing of the stress in relationship to other biological events may modify the magnitude of the biological response. For example, the spawning season for amphibians in northern North America in spring corresponds to the greatest thinning of the ozone layer and ultraviolet (UV) light penetrations. The developing eggs of some of these species have been shown to be affected adversely by UV light (Blaustein and Wake, 1995). This temporal factor may be contributing to observed declines in the numbers and kinds of some amphibian species present in those areas.

The intensity of an environmental stress describes its relative ability to evoke a response from the receptor. With increasing intensity, the impact may progress from a few, slightly affected, particularly sensitive components of the biological system to most components being grossly affected. Small changes in the histology or physiological state of individual species can be expected to occur at lower stress intensity and chronologically before changes in survival at a similar temporal and spatial scale. Similarly, at the community level, changes in species composition can be expected to occur at lower stress intensity and chronologically before changes in the community functions. The intensity of hurricanes is routinely ranked from category 1 to category 5. In the case of chemical pollution, the intensity of the environmental stress can be described by the concentration of the chemical in the environment. Thus, copper is a natural, background constituent of water in a river or stream. Intake of copper is essential for both animal and plant life. However, concentrations higher than $10\ \mu g\ l^{-o}$ in river water can be expected to eliminate a few sensitive species and change the age class distributions in others by changing reproductive success. Higher concentrations can be expected to affect more components in more obvious ways. For example, in Shayler Run, Ohio, $120\ \mu g\ l^{-1}$ of copper caused fish kills at some times of the year, avoidance of the stream reach by other fish, declines in macroinvertebrate community species richness, and other gross responses (Geckler *et al.*, 1976).

The novelty of an environmental stress will determine whether or not biological systems will have mechanisms in place to deal with it. Environmental stresses that resemble naturally occurring stresses in their mode of action will be dealt with by the system in the same ways. Novel stresses may be more devastating because no mechanisms have evolved to cope with them. For example, when human harvests of wood products mimic relatively frequent natural events such as treefall or windthrow in their spatial extent, mechanisms are in place in the biological system to recover from this event. When human-created gaps are larger and more intense than the historical disturbances, these same mechanisms may not help. Similarly, some natural systems can break down, render biologically unavailable, or disperse low levels of some chemical materials that are naturally occurring or that resemble naturally occurring substances without detectable disruption. The ability of a natural system to receive materials at some concentration, including anthropogenic wastes, without being degraded is its assimilative capacity. However, overloading a system with too much waste destroys both the structure and function of the ecosystem and its future assimilative capacity.

Receptors and Responses

An action, agent, or condition can be a stress to one biological system whereas simultaneously not affecting many others. Because environmental stress is defined by the observation of an impaired biological system, the probability of identifying an environmental stress is related to how thorough the search for impairment has been. Although there are an almost unlimited number of receptors and responses that could be affected by any particular environmental stress, it is generally impractical to monitor more than a small sample. Also, experience has shown that it is easy to overlook a response that may be important. For example, DDT caused eggshell thinning in some birds, although routine toxicity tests failed to identify this response.

The most useful responses to examine for studies that aim to influence decisions about environmental management tend to be those that are clearly related to stated environmental goals. These tend to be responses that are both biologically and socially relevant and that can be measured reliably. Sometimes, the responses that society cares most about cannot be measured directly. Other, presumably related, responses can be measured. In addition, responses that occur earlier in a chain of events, and which lead to an ecologically relevant

Table 1 Examples of receptors and responses used in studies of environmental stress

Level of biological organization	*Structural responses*	*Function responses*
Individual	Condition	Growth
	Fat stores	Fecundity
	Histopathology	Physiological function
Population	Occurrence	Yield
	Abundance	Gross morbidity
	Age structure	
Community	Species richness	Production
	Trophic structure	Respiration
	Proportion of exotics	Extinction rate
Ecosystem	Nutrient pool size	Materials cycling
	Biomass	Materials export
Landscape	Habitat proportions	Regional production
	Patch size	Materials export
	Perimeter-to-area ratio	Resistance to stress

event, may be useful as early warnings of conditions that have the potential to cause unacceptable damages. **Table 1** summarizes some responses that have been used to evaluate environmental stress and guide environmental management.

As the goals in environmental protection have changed over time, so have the responses that are monitored. Most early tests of environmental stress were designed to protect one species – humans. This objective spurred tests that monitored the physiology of species used as human surrogates. Gradually this protection was extended, first to domesticated animals and plants and then to commercially valuable wild species. These additional tests monitored the survival of populations of important species. Currently, goals extend beyond the protection of individual species and include the protection of biodiversity and ecosystem services (i.e., those structures and functions of natural biological systems that directly or indirectly support human life). Assessments should reflect these new goals because millions of species in the environment must be protected. Each species cannot be examined individually. In practice, a few species must serve as surrogates for many others, and a few systems serve as surrogates for many others.

Environmental responses may be characterized by type and scale. Responses are either structural (e.g., describing the number and kinds of components, such as the macroinvertebrate community structure) or functional (e.g., describing performance or flux, such as biological oxygen demand or primary production). Also, unique responses may occur at many distinct spatial and temporal scales and levels of biological organization (e.g., cells, tissues, organs, organisms, populations, communities, ecosystems, landscapes, biomes, and the world). Some attributes at higher levels of biological organization are not present at lower levels; for example, energy flow and nutrient spiraling are properties of ecosystems but not of organisms. Other attributes are present in some form at many levels; for example, one can measure the diversity of phenotypes at the population level and the diversity of species at the community level. Environmental goals can be stated on many of these levels, but tests of environmental stress are largely limited to those levels that are more accessible to human observation.

An awareness of scale provides two contrasting approaches to study environmental stress. Top-down methods start with observed damage to a biological system of interest and investigations move down through hierarchical levels. Component structures and functions are examined in order to diagnose the causative agent and plan remedial actions. At the outset, the damage has already been done, so the relevance of the changes is known. However, the causative agent and the chain of events leading to unacceptable damage are not known. Bottom-up methods start with an environmental stress, and the effects of that stress on biological systems are determined through designed experiments. Because experiments on small and quick biological systems at lower scales are generally less expensive, these experiments are most common. In bottom-up assessments, the causative agent is known at the outset, but the importance of ultimate changes at any ecologically relevant higher scale is not known.

Microcosms and mesocosms are attempts to increase the spatial and temporal scales and level of complexity in biological systems that can be used in designed experiments of environmental stress. Microcosms and mesocosms simulate important attributes of natural systems in laboratory or outdoor conditions. As the names indicate, the main difference is in size. Microcosms are sometimes small enough to hold in one's hand; mesocosms may cover one acre. Neither is an exact reproduction of any real ecosystem, but they do enable studies of environmental stress in ways that avoid damaging natural systems. On rare occasions, environmental stress may be studied in designed experiments using entire ecosystems. For example, one whole system manipulation was carried out in the Hubbard Brook drainage basin in New York State (Bormann and Likens, 1979). Such efforts are of great value in calibrating models.

A General Environmental Stress Syndrome

A threshold is defined in Webster's *Third International Dictionary* as "the point at which a physiological or psychological effect begins to be produced." Moving upwards in biological systems, this effect can be generalized to the point at which a response begins to be produced. Woodwell (1974) asked the question

> Is it reasonable to assume that thresholds for effects of disturbance exist in natural ecosystems or are all disturbances effective, cumulative, and detrimental to the normal functioning of natural ecosystems?

Thresholds may be artifacts of testing procedures, reflecting the power of particular test designs rather than a feature of the system being studied. However, perhaps the more important question is "Can humans detect those environmental changes that are important to their own quality of life?" As is the case with human health, the gradient in environmental systems may be extensive between robust health and collapse in some cases, but an abrupt transition from health to collapse may occur in others. By reviewing information available about the behavior of ecosystems under stress,

Table 2 Responses expected in stressed ecosystems

Energetics
Community respiration increases
Gross production/community respiration (*P*/*R*) becomes unbalanced
Maintenance cost increase; gross production/standing crop biomass (*P*/*B*) and community respiration/standing crop biomass (*R*/*B*) ratios increase
Importance of auxiliary energy increases
Exported or unused primary production increases
Nutrient cycling
Nutrient turnover increases
Horizontal transport increases, vertical cycling of nutrients decreases
Nutrient loss increases
Community structure
Proportion of *r*-strategist increases
Size of organisms decreases
Life spans of organisms or parts decrease
Food chains shorten
Species diversity decreases, dominance increases, redundance declines
General system-level trends
Ecosystems become more open
Autogenic successional trends reverse
Efficiency of resource use decreases
Parasitism increases, mutualism decreases
Functional properties more robust than structural properties

Source: Reproduced from Odum EP (1985) Trends expected in stressed ecosystems. *Bio Science* 35: 419–422, with permission from Jstor.

several researchers have tried to outline general ways in which ecosystems respond to various types of stress (Barrett *et al.*, 1976; Odum, 1985; Rapport *et al.*, 1985; Schindler, 1990). An environmental general stress syndrome at the ecosystem level may include the features listed in **Table 2**; however, experience is continually modifying this list. These efforts to derive an environmental stress syndrome are important because they define a progression of impact in which some minor changes precede other more serious ones. By recognizing changes early in the progression of impact, remediation could begin and crises could be averted. However, the challenge of finding one general description for widely varying systems, challenged by widely varying combinations of stress, is daunting.

Stressed ecosystems often recover once the stress has been removed. However, sometimes human assistance is required, and this process is called ecological restoration or rehabilitation. Restoration has as its goal the return of an ecosystem to a close approximation of its condition prior to the stress and the recreation of a functioning, self-regulating system that is integrated into the ecological landscape in which it occurs. The practice of ecological restoration often involves the reconstruction of physical conditions present prior to the stress, chemical cleanup, and biological manipulation, including revegetation and the reintroduction of native species.

Stress Assessments

Studies of environmental stress can have different purposes. In some studies, the purpose is accounting, i.e., what is the existing condition of this biological system? This question can be important for the purposes of disclosure, national environmental accounting, prioritization, and remediation. Studies of environmental stress can also be used for prediction, i.e., will this action cause a problem or which action is better? Predictive studies are used to register chemicals, rank risks, design processes, etc. Another distinct purpose for studies of environmental stress is to provide early warning of conditions that, if left unchecked, will result in significant damage to human quality of life. By detecting damage before it is of a magnitude that is unacceptable, crises can be averted.

Appraisal

Studies of environmental stress can assess the condition of biological systems that exist at a particular point in time. When repeated over time, trends in condition can be assessed. Appraising the condition of a biological system can confirm that environmental quality is adequate or can serve to define an existing environmental problem. Many countries are undertaking a national accounting of the health of their ecological systems. For example, the Canadian State of the Environment Reports and the Environmental Monitoring and Assessment Program in the US measure the condition of rivers, lakes, forests, wetlands, arid lands, and agroecosystems. Although some of these programs measure common sources of environmental stress, as well as biological response, others focus solely on response. Once a problem is found, additional studies to diagnose the problem would include measures of environmental stress.

Prediction

Often, the easiest way to maintain environmental quality is to prevent damage before it occurs. This strategy requires prediction of the future. Such predictions can come from observations of the effects of similar stress on similar systems or from extrapolations or models from the effects of dissimilar stresses or the effects on dissimilar systems.

The ability to extrapolate the measured effects of environmental stress at one level to consequences at a higher hierarchical level depends on the use of mechanistic models that describe the interaction of component parts. The model is then calibrated, i.e., compared to observed behavior of a system under environmental stress and adjusted to maximize the accuracy of its predictions. Currently, there are a few calibrated models for large-scale predictions. Also, it is unlikely that calibrated models will be available for some processes because their spatial and temporal scales make testing impractical or unethical. All hypotheses and theories are more readily accepted if they have withstood rigorous testing. Predictions of environmental stress are no exception. As a general rule, multiple lines of evidence published in peer-reviewed professional journals whose contents have been reviewed by respected professionals result in acceptance both by the person carrying out research on environmental stress and by mainstream science. Generally, validation occurs in two primary ways: (1) a designed test of a hypothesis derived from a theory, especially by those having nothing to do with its development, and (2) consilience with other well-accepted and tested

theories. Predictions of environmental stress are particularly likely to be challenged outside the scientific community because taking precautionary measures to avoid conditions estimated to cause stress often requires changing societal and industrial practices and sometimes engenders costs. Environmental stress associated with global warming is a good illustrative example. Limiting greenhouse gases in the atmosphere will affect the lives of almost every person on the planet, as would significant global warming. In this case, the entire planet is the experimental unit, making designed tests of impact at the same hierarchical level as that of the environmental problem impossible since there is no control planet available. As a consequence, much uncertainty about the probable effect of increased greenhouse gases will persist. However, management decisions about such gases must be made and must be based on the best available information; managers must act even though there is uncertainty accompanying any estimate of risk. Fortunately, most forms of environmental stress, such as exposure to potential toxicants, are much more easily validated.

Early Warning

Monitoring is a systematic and orderly gathering of data to ensure that previously established quality control conditions are being met. Biomonitoring applies this activity to the detection of environmental stress. In this case, the goal is to provide an early warning that unacceptable levels of environmental stress have occurred. As is the case in an intensive care room in a hospital when heart and respiration rates are monitored and unacceptable conditions are detected, an immediate action is mandatory. In any form of quality control, the more rapidly the information becomes available, the more quickly corrective action can be taken. Extensive use of information technology has made complete automation of some monitoring systems, including triggering the remedial action, possible. However, in the absence of carefully selected goals and objectives related to the decisions to be made, it can also generate huge amounts of unnecessary and inappropriate data.

Setting the corrective action threshold low to ensure early detection of deleterious change seems prudent, but it can also produce false-positive readings. A false positive is an indication that some deleterious effect has occurred when in fact none has occurred. Emergency team response to eliminate stress can be quite expensive and unpopular with management. False positives are usually most numerous when monitoring is in the early developmental stages and decrease substantially as experience is gained. Avoiding false positives by setting the action threshold well beyond the response threshold will probably result in false negatives. A false negative is information that no deleterious effects have occurred when in fact some have occurred. Thus, ecosystem damage occurs because no corrective action alert is produced. These false signals are clearly a matter of prime importance in the design of all monitoring systems. The dilemma is that one wishes to detect environmental stress at the earliest possible moment, but this may result in false signals since sensitive end points are often highly variable.

The Biospheric Life Support System

Earth is now in its sixth biosphere. This sixth biosphere provides unique conditions in which the genus *Homo*, including *Homo sapiens*, evolved and flourished. The biosphere also provides the resources on which the human economy is based. In terms of biospheric function, a species must not only be present but present in sufficient numbers to be effective ecologically. The present biosphere has already suffered from a substantial loss of species due to both extinction and biotic impoverishment. The tipping point for the present biosphere is unknown. Changes caused by passing a tipping point are irreversible; consequently, prudence dictates avoiding such a tipping point.

Future Trends

In view of unprecedented growth in human population and the desire to raise the standard of living above the subsistence level for most of the world's people, there are new challenges for those studying environmental stress. The level of environmental stress from food production and energy usage may increase. Simultaneously, there is a great pressure to increase food production to feed the increasing human population, and the need to protect intact biological systems and ensure their robust functioning so that they can continue to provide necessary ecosystem services will become more pressing. Some approaches to these problems are of great interest.

The World Commission on Environment and Development (1987) of the US published *Our Common Future*, which focuses attention on the future condition of the planet, arguably more so than any publication that preceded it. The commission defined sustainable development as

> development that meets the needs of the present without compromising the ability of future generations to meet their own needs.

This presents a curious combination of words because "sustain" means to continue and "development" is usually associated with growth. However, infinite growth on a finite planet is clearly not possible.

In December 1989, the General Assembly of the United Nations (UN) attempted to address the problems identified in *Our Common Future* (World Commission on Environment and Development, 1987) by organizing the UN Conference on Environment and Development (popularly known as the Earth Summit), which was held in Rio de Janeiro in June 1992. This conference resulted in a heightened awareness of the interrelatedness of environmental degradation and stress, population development, and depletion of natural resources. Previously, all had been viewed as separate problems, but now attempts are being made to address them as an interactive system. The resulting Rio Declaration on Environment and Development endorsed the following principles: (1) Nations should not cause damage to the environment of other states and areas beyond their borders; (2) eradicating poverty and reducing disparities in worldwide standards of living are indispensable requirements for sustainable development; (3)

the polluter, in principle, should pay the cost of pollution; (4) states should discourage or prevent trans-boundary movements of activities and the substances that endanger health or environment; and (5) scientific uncertainty should not be a reason for postponing urgent measures to prevent environmental degradation.

Sustainable use of the planet will not be possible if excessive environmental stress impairs the ecological life support system that provides necessary ecosystem services. As a consequence, keeping environmental stress at tolerable levels has a direct bearing on the quality of life for humans. Events beyond human control, such as a large extraterrestrial object striking earth, make it impossible to guarantee that reducing environmental stress will ensure sustainability.

The theory of "weak sustainability" asserts that human society is sustainable provided that the aggregate stock of manufactured and natural assets is not decreasing. Thus, the loss of the whaling industry would not impair sustainability if the proceeds of liquidation are invested in industries of comparable income-producing potential. Pearce and Atkinson (1993) dispute the assumption that natural and human-made capital are sustainable in this context. They assert that strong sustainability requires that natural capital stocks be held constant, independently of the human-made capital. A weak sustainability scenario would permit considerably more environmental stress than a strong sustainability scenario. Conventional monetary analyses are biased against strong sustainability. For example, at a discount rate of 5%, the current value of ecological services for an American life span (about 76 years) from the present on is approximately 2.5 ¢. Using this approach, the farther into the future one projects, the less valuable natural systems appear, thus diminishing the significance of environmental stress.

Reducing environmental stress requires the cooperation of the entire society. One of the most important components is the industrial system. Fortunately, the field of industrial ecology (IE) is developing worldwide (Hawken, 1993). The goal of IE is to reduce environmental stress at all stages: (1) extraction of raw materials, (2) processing, (3) disposal of manufacturing wastes, (4) packaging, and (5) reincorporation into the environment at the end of the product's life in a nonstressful way, ideally in a way that enhances ecological integrity. IE's primary goals are to (1) reuse materials as much as possible, (2) reduce energy consumption per unit produced, and (3) design both processes and products so that they can be reincorporated into the environment with minimal stress. Books such as *Engineering within Ecological Constraints* (Schulze, 1996), which was produced by the National Academy of Engineering, are directed toward achieving this goal.

See also: Adaptation. Biodiversity, Origin of. Carrying Capacity, Concept of. Ecosystem, Concept of. Energy Use, Human. Extinction in the Fossil Record. Functional Diversity. Global Species Richness. Population Dynamics

References

Barrett GW, Van Dyne GM, and Odum EP (1976) Stress ecology. *Bio Science* 26: 192–194.

Blaustein AR and Wake DB (1995) The puzzle of declining amphibian populations. *Scientific American* 272(4): 52–57.

Bormann RH and Likens GE (1979) *Pattern and Process in a Forested Ecosystem*. New York: Springer-Verlag.

Flannery TF (1999) Debating extinction. *Science* 283: 182–183.

Geckler JR, Horning WB, Neiheisel TM, Pickering QH, and Robinson EL (1976) *Validity of Laboratory Tests for Predicting Copper Toxicity in Streams, EPA600/3-76-113*. Springfield, VA: National Technical Information Service.

Hawken P (1993) *The Ecology of Commerce*. New York: Harper Collins.

Johnston CA, Detenbeck NE, Bonde JP, and Niemi GJ (1988) Geographic information systems for cumulative impact assessment. *Photogrammetric Engineering and Remote Sensing* 54: 1609–1615.

Kelly JR and Harwell MA (1989) Indicators of ecosystem response and recovery. In: Levin SA, Harwell MA, Kelly JR, and Kimball KD (eds.) *Ecotoxicology: Problems and Approaches*, pp. 9–35. New York: Springer-Verlag.

Odum EP (1985) Trends expected in stressed ecosystems. *Bio Science* 35: 419–422.

Pearce D and Atkinson G (1993) Capital theory and the measurement of sustainable development. *Ecological Economics* 8(2): 103–108.

Rapport DL, Regier HA, and Hutchinson TC (1985) Ecosystem behavior under stress. *American Naturalist* 125: 617–640.

Rees WE (1996) Revisiting carrying capacity: Area-based indicators of sustainability. *Population and Environment* 17: 195–214.

Schindler DW (1990) Experimental perturbations of whole lakes as tests of hypotheses concerning ecosystem structure and function. *Oikos* 57: 25–41.

Schulze PC (1996) *Engineering Within Ecological Constraints*. Washington, DC: National Academy Press.

Woodwell GM (1974) The threshold problem in ecosystems. In: Levin SA (ed.) *Ecosystem Analysis and Predictions*, pp. 9–21. Alta, VT: SIAM Institute for Mathematical Society.

World Commission on Environment and Development (1987) *Our Common Future*. Oxford: Oxford University Press.

Subspecies, Semispecies, Superspecies

James Mallet, University College London, London, UK

© 2013 Elsevier Inc. All rights reserved.

A Brief History of Subspecific Taxonomy

Variation Below the Level of Species

Since the invention of binominal nomenclature by Linnaeus, there has been a conflict between "splitters" who named more or less well-defined local populations as separate species, and "lumpers" who ignored geographic variation, and united local variants into a single species. The problem was compounded by early systematists' belief that species had an Aristotelian essence, each fundamentally different from similar essences underlying other species. To Linnaeus' followers, it seemed important to decide, which level of variation was fundamental. The terms "genus" and "species" both result from Aristotelian philosophy, and although Linnaeus is usually credited with establishing the species as the basal taxonomic unit, he confused matters, after recognizing that some plant species were of hybrid origin, by suggesting that genera were a more important taxonomic level (i.e., a separately created kind) than species.

Once evolution was accepted, it became clear that variation at all levels in the taxonomic hierarchy was due to more or less similar causes; the only difference between variation above the level of genus or species and below was one of degree. Darwin realized that species could evolve from intraspecific varieties, and he used the term "species" in a new and nonessentialist sense: "… the complete absence, in a well-investigated region, of varieties linking together two closely allied forms, is probably the most important of all the criterions of their specific distinctness … Geographical distribution is often brought into play unconsciously and sometimes consciously; so that forms living in two widely separated areas, in which most of the other inhabitants are specifically distinct, are themselves usually looked at as distinct; but in truth this affords no aid in distinguishing geographical races from the so-called good or true species" (Darwin, 1874). Using this criterion, Darwin classified all human races as members of the same species. Darwin showed convincingly that there was no essential difference between species and varieties; species were simply varieties, which had diverged more, and which could coexist without intermediates being common. However, with his term "varieties" Darwin did not clearly distinguish between polymorphic variants within populations and the identifiable geographic populations normally today considered as geographic races or subspecies. To Darwin the distinction was unimportant, because polymorphic variants, clinal variation, geographic races or subspecies, and good species formed a continuum. Darwin demonstrated that this continuum was excellent evidence for an evolutionary origin of the taxa we call species.

The Trinominal Revolution in Zoology

Many systematists wished to preserve the purity of the simple genus–species binominal nomenclature, but by the 1850s, there were enormous stresses. It began to be realized that identifiable geographic replacement forms were an important intermediate stage between insignificant local variants and good species. Some lumped these replacement forms as varieties within species, whereas others continued describing these replacement forms as separate species: practices varied widely, leading to considerable confusion. Although some Europeans had long advocated naming marked geographic forms as subspecies, the accumulation of major North American museum collections during the great push of colonization and railway construction westwards was probably the most important catalyst of a revolutionary new systematics. In this new approach, nomenclature consisted of a trinominal: genus–species–subspecies, which is still the dominant taxonomic practice today. The maxim was: "intergradation (at the boundary between two geographic replacement forms) is the touchstone of trinominalism." Examples from commonly observed birds, which intergrade are, in North America, the eastern rufous-sided towhee (*Pipilo erythrophthalmus erythrophthalmus*) replaced in the west by the spotted towhee (*Pipilo erythrophthalmus maculatus*), and in Europe the carrion crow (*Corvus corone corone*) found in the south and west, replaced by the hooded crow (*Corvus corone cornix*) in Italy, and the north and east of Europe. Among the ornithologists responsible for this revolution in North America was Elliott Coues, who published a catalog of American birds in 1872 incorporating an early version of this trinominal nomenclature, in which subspecies were prefixed by "var." and Robert Ridgway, who finally dropped the "var." in his own 1881 summary of North American bird nomenclature.

The American Ornithologists' Union soon adopted this policy, and the idea then spread to Europe, particularly England where Walter Rothschild began amassing his vast collection of birds and butterflies, and had hired excellent and productive staff, the ornithologist Ernst Hartert and the entomologist Karl Jordan, to curate and describe the new material. Jordan was particularly important in spreading the idea of trinominal nomenclature to entomologists, and in promoting the abandonment of other subspecific nomenclature. He was regarded by the Rothschilds as the clever member of the staff (Rothschild, 1983), and published enormous systematic treatises as well as papers on the theory of systematics, justifying trinominal nomenclature and the recognition of the subspecies as a valid, identifiable taxon in its own right. Both Jordan and Hartert were Germans who contributed to and read European as well as English journals, and in Germany a similar revolution was taking place. Thus, these systematic ideas were able to spread to the rest of Europe in the time when science was often highly parochial. The standard trinominal nomenclature for subspecies, and the abandonment of other named ranks, such as semispecies and superspecies, soon became established in the International Code of Zoological Nomenclature, and has remained there ever since.

Meanwhile, botany, Linnaeus' major expertise, remained much less prescriptive. Many infraspecific ranks are still considered valid taxa under the International Code of Botanical Nomenclature. For instance, *Saxifraga aizoon* var. *aizoon* subvar. *brevifolia* f. *multicaulis* subf. *surculosa* is a valid scientific name in botany, with names for variety, subvariety, form, and subform as well as genus and species ranks. Cultivated strains, hybrids, subspecies, are also considered valid nomenclatural ranks in botany, whereas subspecies do not necessarily refer specifically to geographic populations, as is the case in zoology. The remainder of this article is concerned largely with zoological terminology. It is important to realize, however, that both the botanical and the zoological codes avoid controversy by not specifying how to decide whether a particular group of specimens is a species or infraspecific taxon. The goal of these codes is to merely regulate nomenclature once a decision of rank has been reached.

Theories of Divergence: Superspecies, Semispecies, and Subspecies

It is hard to imagine the diversity of ideas by which the systematists of 100 years ago explained geographic variation. At that time, evolution by natural selection was far from generally accepted, in fact many believed that it had been disproved. One of the most influential ornithologists of the time was Otto Kleinschmidt, who believed that all species suddenly came into being long ago, and since then had remained completely separate. Replacement forms or subspecies had diverged from the main species but, in Kleinschmidt's view, subspecies could never evolve into new species as the Darwinians supposed. To distinguish his new species concept from the older one in which geographical replacements might be named as separate species, Kleinschmidt called his theory of variation the *Formenkreis* (ring of forms) theory. The *Formenkreis* theory fitted neatly with, and indeed promoted the new practices of naming subspecies and trinominal nomenclature.

In those times many somewhat peculiar explanations competed to explain geographic variation and speciation, including Kleinschmidt's "nonspeciation" theory, Lotsy's hybridization theory, saltational evolution via mutation, inheritance of acquired characters, as well as natural selection. In Britain, Jordan and Rothschild argued eloquently and influentially against any new terminology (including *Formenkreis*) that had theoretical implications, and proposed incorporating as little evolutionary theory into taxonomy as possible, in view of the lack of agreement among scientists at the time. Rothschild and Jordan (1895, 1903), supported by Hartert, agreed both with the nomenclatural practice of naming subspecies, and also that subspecies were valid real taxa that could evolve into full species. They argued that the Linnaean term "species" should be retained for the whole group of races, and that the geographic races were not true species, they were simply subspecies or incipient species.

Others felt that the term species was too emotive to be used in the new, multiple-subspecies sense. Some scientists continued following the *Formenkreis* doctrine, and had begun to name quite distinct taxa, which did not intergrade at their boundaries, as subspecies. This situation led in the 1920s and 1930s to the neo-Darwinian ornithologist Bernhard Rensch scrapping the term *Formenkreis* because of its theoretical limitations, and instead substituting two new terms, *Rassenkreis* (circle of races) and *Artenkreis* (circle of species). *Rassenkreise* were again considered to be equivalent to species, composed of races or subspecies. However, now there was an additional layer in the taxonomy, of groups of *Rassenkreise* that replaced one another geographically, the *Artenkreise*. Thus an *Artenkreis* could consist of multiple *Rassenkreise*. Rensch and many others believed that the subspecies was an incipient species, of which the geographic replacement species, *Artenkreise* were a further development, until finally divergence was sufficient to allow complete geographic overlap, whereupon new *Rassenkreise* could again be formed.

These terms did not catch on, and most scientists came to the conclusion that the *Rassenkreise* were equivalent to the species taxa used by Linnaeus and Darwin. Probably a major reason that we do not use these multiple taxonomic terms in zoology is due to the prolific work published in English by another German, Ernst Mayr. Mayr had worked for Walter Rothschild, and knew Hartert. After Walter Rothschild was blackmailed by a lover, his enormous bird collection of 280,000 skins was sold in 1932 to the American Museum of Natural History, where Mayr happened to have been hired as a curator. Mayr's experience of ornithology, contact with the European literature, and burgeoning friendship with the geneticist Theodosius Dobzhansky (who convinced him of the lack of evidence for inheritance of acquired characters), lent a unique opportunity to influence the course of systematics and evolutionary biology. Mayr did not waste this opportunity. Mayr used ideas underlying Rensch's terms, but renamed them in English. The *Rassenkreis* became simply the species or polytypic species, with its geographic races being subspecies, whereas the *Artenkreis* became the superspecies, and its component parts semispecies, that is, replacement species, not very divergent but which may hybridize occasionally where they overlap. Mayr successfully blended the local species concept of Poulton and Dobzhansky based on interbreeding with the geographic *Rassenkreis* idea of species, incorporating Jordan and Rothschild's ideas about subspecies. He renamed this combination of ideas "the biological species concept," a term, which has since remained strongly associated with Mayr's name. His many influential articles and books promoted a new program of species study, a science of the species that is still with us today. Central to Mayr's system was the belief that discrete taxa such as species or subspecies would normally diverge in "allopatry," that is, in complete geographic isolation.

The Subspecies Today

Modern Views of Subspecies and Semispecies

The views of Darwin, Wallace, Rensch, and Mayr that geographic replacement forms, subspecies and semispecies, are in fact incipient species, has few critics today. Most geographic replacement species (i.e., semispecies, which intergrade only rarely when they meet) must indeed have evolved from previously interbreeding subspecies. Modern genetic data have done nothing to cast doubt on this idea. Meanwhile,

superspecies became a specialized term reserved for groups of semispecies that intergraded little at their boundaries. Because neither of these are valid ranks in the zoological or botanical codes, taxonomists normally name such replacement semispecies at the species rank.

Under the trinominal approach, taxonomists were now required to describe subspecies, which has been never seen as a particularly noble activity in comparison to the description of species, especially recently. A strong attack on the zoological subspecies was mounted by Wilson and Brown (1953). Both were systematists working on ants, a group particularly riddled with poorly conceived trinominals and other named varieties at the time. Wilson and Brown argued that subspecies rarely, if ever, could be justified on the basis of multiple characters, and that therefore they were not real taxa. The only real taxa were species, which in a sense were self-defining because interbreeding prevented divergent genes from flowing from one species to another. Subspecies that interbred at their boundaries, however, were not so endowed, so that genes and morphological characters could flow between them. Good examples of subspecies were put forward, which undoubtedly would be hard to justify on multiple character grounds. This single paper was enormously influential on systematics in the US, and generations of insect systematists trained at Harvard and Cornell, where Wilson and Brown worked, together with their own many intellectual descendants, and their students' students in turn, have eschewed the practice of naming subspecies.

However, recent work shows that many subspecies separated by hybrid zones do in fact differ at multiple morphological, behavioral, and genetic characters (Barton and Hewitt, 1985). For instance, the toad *Bombina bombina* meets its relative *Bombina variegata* across a broad front in Europe. The two forms hybridize freely in the contact zone – although the hybrids can be shown to suffer some inviability – and so should really be classified as members of the same species under polytypic or biological species concepts. However, it has always seemed natural to place such well-defined forms in separate species in spite of the fact they have not truly speciated. The two taxa differ strongly in call, morphology, skin thickness, the sizes of water bodies used for breeding, and egg size, as well as in mitochondrial DNA and protein sequences. The levels of differentiation suggest that the *Bombina* taxa have evolved separately for many millions of years. This situation of multiple character changes has now been shown to be true across many examples of subspecies as well as (semi-) species separated by hybrid zones. Effective gene flow can be shown to be almost completely blocked by hybrid zones such as these, even if hybridization is frequent (Barton and Hewitt, 1985). Thus, although some named subspecies undoubtedly merited Wilson and Brown's scorn, genetic evidence shows that there are plenty of local replacement forms, which hybridize at their boundaries but which do form real identifiable taxa that are not reproductively isolated, and are therefore valid subspecies under the Wilson and Brown criteria.

Subspecies, Species, and Conservation

This opposition among modern taxonomists to subspecies taxa can be traced as one catalyst of the recent diagnostic version of the phylogenetic species concept. The adherents of this view of species, led by the ornithologist Cracraft (1989), proposed a radical species concept so that even a single fixed character difference may define a geographic form as a separate species. Multiple character justification is not considered necessary, even at the species level. The practical result of this new concept is that many local forms, having been downgraded to subspecies in polytypic species, are again being recognized as species, leading to taxonomic inflation (Isaac *et al.*, 2004). In birds and butterflies, which often have many morphologically or genetically distinct subspecies, this could easily result in a 2–10-fold increase in the number of species.

It is probable that the revision of geographic forms upward to the level of species is being driven not only by theoretical considerations, but also by existing conservation legislation, which proposes that endangered species are the valuable units to be conserved. If a reserve contains a taxon recognized as a species rather than just as a local subspecies, it may be seen as more valuable for conservation purposes. The potential consequences for biodiversity and conservation of the continued instability of the term "species" are detailed in the article Species, Concepts of. Today's conservationists are reducing emphasis on species conservation, and are becoming increasingly aware of biodiversity at all the levels of the hierarchy of life, including well-marked subspecies. Thus, the legislative need for differentiating local races as species may ultimately become less important provided that future legislation falls more into line with the prevailing biological thought.

Further Reading

Much of the historical overview in this article is covered by the excellent reviews of Stresemann (1936, 1975), Mayr (1982), and Rothschild (1983), as well as by other sources already cited. A recent biography of Karl Jordan (Johnson, 2003) and work on the connections between the species concepts of Darwin, Wallace, Poulton, Karl Jordan, and Ernst Mayr (Mallet, 2004a, b, 2008, 2009) may also be of interest.

See also: Speciation, Process of. Species, Concepts of

References

Barton NH and Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* 16: 113–148.

Cracraft J (1989) Speciation and its ontology: The empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: Otte D and Endler JA (eds.) *Speciation and its Consequences*, pp. 28–59. Sunderland, MA: Sinauer Associates.

Darwin C (1874) *The Descent of Man and Selection in Relation to Sex*, 2nd edn. London: John Murray.

Isaac NJB, Mallet J, and Mace GM (2004) Taxonomic inflation: Its influence on macroecology and conservation. *Trends in Ecology and Evolution* 19: 464–469.

Johnson KR (2003) *Karl Jordan: A Life in Systematics. PhD Dissertation*. Corvallis: Oregon State University.

Mallet J (2004a) Poulton, Wallace and Jordan: How discoveries in *Papilio* butterflies initiated a new species concept 100 years ago. *Systematics and Biodiversity* 1: 441–452.
Mallet J (2004b) Species problem solved 100 years ago. *Nature* 430: 503.
Mallet J (2008) Mayr's view of Darwin: Was Darwin wrong about speciation? *Biological Journal of the Linnean Society* 95: 3–16.
Mallet J (2009) Alfred Russel Wallace and the Darwinian species concept: His paper on the swallowtail butterflies (Papilionidae) of 1865. *Gayana* 73: 42–54.
Mayr E (1982) *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Belknap.
Rothschild M (1983) *Dear Lord Rothschild: Birds, Butterflies and History*. London: Hutchinson.
Rothschild W and Jordan K (1895) A revision of the Papilios of the eastern hemisphere, exclusive of Africa. *Novitates Zoologicae* 2: 167–463.
Rothschild W and Jordan K (1903) A revision of the lepidopterous family Sphingidae. *Novitates Zoologicae* 9(supplement): I–cxxv, 1–813.
Stresemann E (1936) The Formenkreis theory. *Auk* 53: 150–158.
Stresemann E (1975) *Ornithology From Aristotle to the Present*. Cambridge, MA: Harvard University Press. [Reprinted edition of Stresemann's 1951 "Entwicklung der Ornithologie." Translated by Epstein HJ and Epstein C; Cottrell GW (ed.); foreword and epilogue by Mayr E].
Wilson EO and Brown WL (1953) The subspecies concept and its taxonomic application. *Systematic Zoology* 2: 97–111.

Subterranean Ecosystems

David C Culver, American University, Washington, DC, USA
Tanja Pipan, Karst Research Institute at ZRC SAZU, Postojna, Slovenia

© 2013 Elsevier Inc. All rights reserved.

Glossary

Anchialine (or anchihaline) Haline water, usually with restricted exposure to open air, and always with subterranean connections to the sea.
Epikarst Interface zone between soil and rock characterized by fractures and small solution pockets; an ecotone between surface and subterranean water in karst.
Hypotelminorheic A persistent wet spot, a kind of perched aquifer, rich in organic matter, underlain by a clay layer typically 5–50 cm beneath the surface, with a drainage area typically of less than 10,000 m^2.
Interstitial The spaces between sediment particles, especially in alluvial deposits.
Milieu souterrain superficiel (MSS) The interstitial spaces deep in the soil and mantle/bedrock interface, as found in glacially fragmented zones. By origin, the milieu souterrain superficiel, or mesovoid shallow stratum, thus the acronym MSS.
Stygobiont Aquatic species that are obligate subterranean dwellers.
Stygophile Aquatic species that can live and reproduce in both subterranean and surface environments.
Troglobiont Terrestrial species that are obligate subterranean dwellers; sometimes used for aquatic species as well.
Troglomorphic Pertaining to the morphological, behavioral, and physiological characters that are convergent in subterranean species.
Troglophile Terrestrial species that can live and reproduce in both subterranean and surface environments; sometimes used for aquatic species as well.

The Subterranean Domain

The subterranean domain, literally the domain below the Earth, includes both air- and water-filled underground habitats. The most celebrated of these are caves. However, the subterranean domain is much more extensive and complex than caves. Subterranean cavities occur in a wide variety of situations. Among them are large spaces such as caves and small spaces such as the interstitial openings between sand grains (Botosaneanu, 1986). Many large spaces, especially karst caves, formed by the action of acidic waters, are in carbonates (particularly limestone) and evaporites (particularly gypsum).

Landscapes in which the primary agent molding the landscape is dissolution rather than erosion are called karst landscapes. Comprising ~15% of the Earth's surface, karst represents 75% of the land area of Cuba, 45% of Slovenia, 25% of France and Italy, and 40% of the US east of Tulsa, Oklahoma. Caves are present in Cambrian rocks, more than 500 million years ago, to Holocene limestone, less than 10,000 years old. Although caves have a reputation as both exotic and rare habitats, they are actually quite common. More than 100,000 caves are known from Europe, and nearly 50,000 are known from the US. Caves come in many sizes and varieties, ranging from a few meters long to Mammoth Cave in Kentucky with a surveyed length of more than 500 km.

All subterranean habitats, whether they have large or small spaces, share several key features (Culver and Pipan, 2009). First, there is a permanent absence of light. Second, productivity is extremely limited, and except for rare cases, primary productivity is absent. Third, almost all systems of water- or air-filled cavities have biological activity. Fourth, there is reduced environmental variability relative to surface conditions. Thus, subterranean organisms must contend with complete darkness, limited food, and at least a reduction in seasonal cues. Because they are quite different, it is convenient to consider aquatic and terrestrial habitats separately.

Aquatic Subterranean Habitats

Caves are an important part of aquatic ecosystems in karst terrains, but are not the only components (**Figure 1**). Water enters the subterranean karst system at the rock–soil interface, which typically has many small solution pockets and cavities with complex horizontal and vertical pathways. Eventually, water percolating through the epikarst reaches a cave stream, which in turn exits at a spring and flows into a surface river. Beneath the cave stream is a permanent saturated (phreatic) zone that itself often contains large cavities.

Epikarst is a perched aquifer, and a major ecotone between surface water and cave water. Water is transmitted vertically either through conduits or small fissures to the phreatic zone. Lateral transmission occurs through poorly integrated lateral openings. Its principal characteristic is its heterogeneity, with many semi-isolated solution pockets whose water chemistry is also quite variable. It forms more or less a permanently saturated zone, close under the surface.

The water in cave streams comes from infiltration from the epikarst and from sinking streams. Cave streams formed largely from percolating epikarst water tend to be very stable temporally, whereas cave streams formed largely from sinking surface streams show greater temporal variability. Flow rates of cave streams range from a few milliliters per minute to tens of cubic meters per minute.

A rather unusual habitat, intermediate between an aquatic and terrestrial habitat, is the cave hygropetric. It is where there is a thin layer of water moving over vertical surfaces, resulting from water percolating from above. The flow of the water is

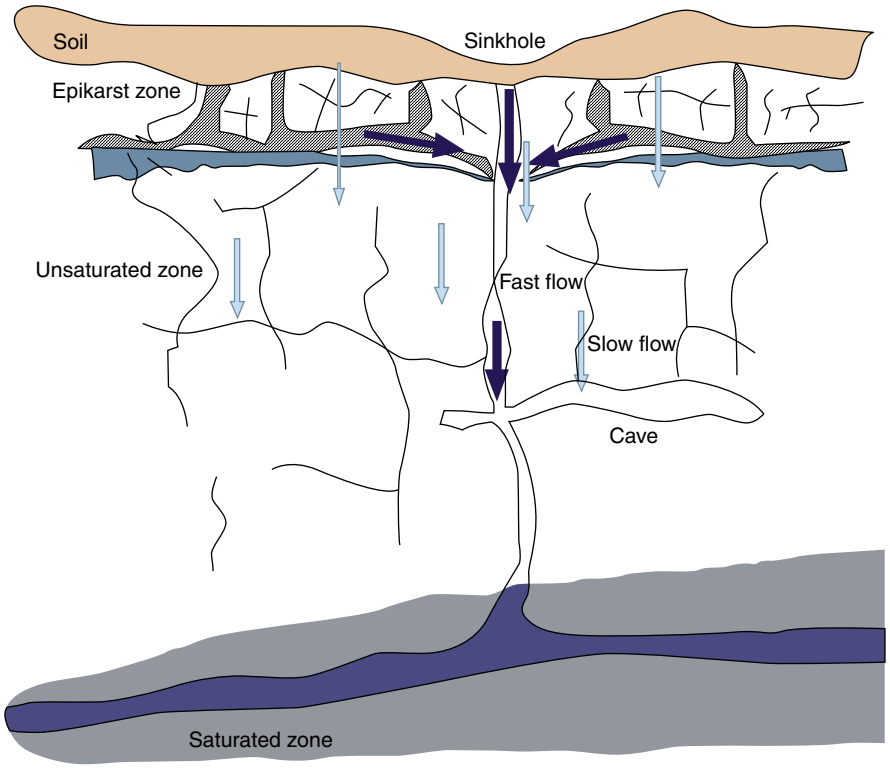


Figure 1 Idealized cross section of a karst area.

usually laminar rather than turbulent, and it is well oxygenated and often relatively rich in organic matter.

Phreatic water (permanent groundwater) in karst areas is in most ways similar to phreatic water in nonkarst areas except that karst phreatic water has high secondary porosity. Phreatic water has very slow flow rates and consequently long residence times – decades or even centuries. Wells and artesian springs in karst areas often connect with water-filled underground voids, tens to hundreds of meters deep.

As described below, epikarst, cave streams, and the phreatic zone contain distinct communities. Other important aquatic habitats in karst landscapes include intermittent lakes (a much understudied habitat), anchialine water (haline water, usually with restricted exposure to open air, and always with subterranean connections to the sea), hydrothermal waters (both in caves and in springs), and springs (a flow of water rising or issuing naturally out of the earth). It is worth noting that springs come in a wide variety of sizes and shapes, and in countries with extensive karst development such as Slovenia and Croatia, there are multiple words that subdivide the English word "spring." A specialized aquatic karst habitat has been found in extensive areas of western Australia. These are thin carbonate deposits in isolated, though sometimes extensive pockets formed by evaporation of near-surface waters in very arid climates. They probably have some functional analogies with epikarst, both being relatively shallow subterranean habitats with capacity for water storage, and relatively rich in organic matter.

The aquatic interstitial habitat is comprised of the water-filled spaces between grains of unconsolidated and saturated sediments. These habitats occur in littoral sea bottoms and beaches, freshwater lake benthos, alluvial plains, riverbeds, and the hyporheic zone (the porous aquifer beneath and lateral to streams). The hyporheic zone of rivers can also be viewed as an ecotone between surface and groundwater. Stability of environmental variables in interstitial habitats generally increases with depth in these habitats. Grain size determines interstitial space as well as many chemical parameters such as oxygen concentration. At a microscale, interstices are interconnected, and the porous medium is heterogeneous and constitutes a mosaic of microenvironments. Interstitial habitats also occur at depths greater than 1 km, and most water wells have microbial community, both Eubacteria and Archaea. In some cases, the microbes probably colonized the habitat as it was being formed. The interstitial habitat is a very ancient one, especially along the margin of seas.

A final aquatic subterranean habitat fits uneasily into any classification scheme. This habitat type is perched aquifers, typically of less than 10,000 m^2 catchment area, underlain by an impermeable layer, usually clay. Operationally, it is a permanent wet spot in the woods isolated from other surface and groundwater, usually with a black color derived from decaying leaves. This habitat was given the name hypotelminorheic by Mestrov, and the exit of the water to the surface is a seepage spring, according to the classification of Kresic. Also little studied; it may be quite extensive, particularly in regions where there are clay strata.

Terrestrial Subterranean Habitats

The range of terrestrial subterranean habitats is much more restricted than that of aquatic subterranean habitats. Terrestrial

subterranean habitats must of course be above the water table, and in general are less well developed in nonkarst areas. The epikarst habitat has largely been studied by hydrogeologists, but in fact this zone also contains many air-filled cavities, and terrestrial species are routinely collected from epikarst. Both solution caves and lava tubes are important terrestrial habitats. Lava tubes form by the surface-cooling and solidification of lava flows. Under some circumstances, abandoned mines can be an important habitat for hibernating and maternity colonies of bats.

Terrestrial subsurface habitats also occur outside of caves. These include cracks and small voids in volcanic rocks and a superficial zone of rock fissures and debris slopes in a variety of rocks, including limestone, schists, gneiss, etc. Generally occurring at a depth of few meters, this habitat was called *milieu souterrain superficiel* (MSS) in French, or mesovoid shallow substratum in English. The MSS is generally found in mountains in temperate zones but apparently not in the tropics, where voids are usually filled with laterites and clay.

Deep-soil (edaphic) habitats have not traditionally been part of subterranean biology. Deep-soil habitats share all the characteristics of other subterranean habitats except for one – deep-soil habitats tend to be rich in organic matter. However, this distinction is relatively arbitrary and there are other relatively energy-rich subterranean habitats such as the hypotelminorheic and cave guano communities (*see* Some Representative Cave Communities).

Sources of Energy in Subterranean Environments

All subsurface environments share an absence of light and hence an absence of photosynthesis. With the exception of a few caves and possibly most deep-groundwater habitats with significant chemoautotrophy, there is no significant primary production, that is, no photoautotrophy. With these exceptions, all subsurface food webs rely on the import of surface organic matter in one form or another. For the most part, subterranean habitats are energy poor. The famous Romanian speleobiologist Emil Racovitza stated that many subterranean organisms are ''carnivores by predilection but saprophages by necessity'' (translation by the authors).

For most aquatic subsurface habitats, organic matter is brought in by the action of water. Dissolved organic matter (DOM) and fine particulate organic matter (FPOM) are important energy sources in interstitial and hypotelminorheic habitats. The source of the organic matter is surface production, and as organic matter is utilized by subterranean organisms, both its quality and quantity decline. DOM and FPOM are also important in subterranean karst habitats. Vertically percolating water from the epikarst (and ultimately surface water) collects in pools and streams, and is often relatively rich in organic matter. In addition, considerable numbers of both terrestrial and aquatic organisms wash out of the epikarst into pools and streams as well. Although some of this ''rain'' of animals, especially copepods, are able to colonize the pools and streams, others, such as Collembola, survive short periods at most, and are food resources for the cave inhabitants. In many caves, sinking streams (swallets) bring in additional organic matter – not only DOM and FPOM but also coarse particulate organic matter (CPOM), including leaves and twigs. In a sense, many cave streams are not different from a low-productivity stream on the surface, where allochthonous input greatly exceeds autochthonous production.

For many subsurface terrestrial communities, a major source of energy is the organic matter that is stranded along stream banks as a result of the processes outlined in the previous paragraph. In addition, another source of energy in many caves is the result of the movement of animals in and out of the cave. In many North American caves, cave crickets of the subfamily Rhaphidophorinae periodically leave caves to forage and deposit their eggs in sandy substrates in the cave. These eggs are, in turn, the major part of the diet of some troglobiotic carabid beetles. The fecal material of the cave crickets and the occasional dead cave cricket bodies are an important food supply for other troglobionts. Fecal matter left by other cave visitors is also important in many caves. The most important source of fecal material is bats. In most temperate zone caves, only a few bats use the cave, and their guano is the basis for a group of species whose food base is transitory organic matter. In those temperate zone caves that serve as maternity sites or hibernacula, and in many tropical caves, large bat colonies produce enough guano for there to be a community specialized on the guano. Most of these species are guano specialists rather than cave specialists.

There are other allochthonous sources of food. In a few caves, such as the lava tubes on the Canary Islands and some MSS habitats, the major source of organic matter is wind-carried particles and debris, which penetrate into the subterranean habitats through cracks and fissures. In lava tubes and some shallow limestone caves, roots penetrating through the ceiling are an important food supply, and in some Hawaiian lava tubes there are species that specialize in this resource. Roots are also an important food source in the aquatic calcrete habitats of arid Western Australia.

The possibility of a subterranean chemoautotrophic ecosystem has long intrigued biologists. Movile Cave, a small cave near the Black Sea in Romania, is such an ecosystem. This cave, discovered during geochemical prospecting in 1986, has no natural entrance. Its H_2S-, CH_4-, and NH^{4+}- enriched waters are slightly thermal. A series of air bells occur throughout the cave in which atmospheric composition is enriched in CO_2 and CH_4 and depleted in O_2. The water surface is covered with a floating mat of bacteria and fungi. Chemoautotrophic production was demonstrated by incorporation of bicarbonate ion by the mat and by the presence of ribulose-1,5-biphosphate carboxylase-oxygenase, an enzyme characteristic of autotrophic organisms. A Gram-negative rod-shaped bacterium was isolated from the mat that was an obligate autotroph, using the energy of the oxidation of thiosulfate and sulfide. Furthermore, this autotrophic production is the basis not only for the aquatic food web but also for the terrestrial food web – the only documented case of a chemoautotrophic terrestrial ecosystem. Several other sulfide-rich caves, including Frassasi Caves in Italy, Villa Luz Cave in Mexico, and Cesspool Cave in Virginia, also exhibit many features of chemoautotrophic production.

Many phreatic deep systems are chemoautotrophic (actually chemolithoautotrophic). The presence of bacteria in the

deep subsurface was first suggested by Edwin S Bastin in the 1920s. He suggested that the curious presence of hydrogen sulfide and bicarbonate in water extracted from oil was the result of the activity of sulfate-reducing bacteria at depth. It was nearly 60 years before the technology was developed to sample bacteria at depth. To do so the solution of several formidable technical problems was required. The first was to obtain samples uncontaminated by surface microbes, and the second was to characterize the microbes. Characterization of the samples is difficult because of the unusual environmental conditions in which the microbes live, making standard culturing techniques of limited value. Much of our knowledge of deep-groundwater microbes is derived from sampling wells near former nuclear weapons production sites in Washington State and in South Carolina. Bacteria and Archaea have been found at depths of more than 2 km. Their presence in sedimentary rock may be the result of the slow infiltration of water from the surface or they may have been present at the time of the deposition of the rock. The possibility that they were present at the time of deposition is plausible, given the depth of the water and its age based on stable isotope ratios. Microbes present in igneous rock must have been deposited after it formed since it was too hot at the time of formation. In these habitats, chemolithotrophs, which are forms that use inorganic carbon as a carbon source and obtain energy from the oxidation of reduced inorganic chemicals, predominate. The possible great age of some of these microorganisms combined with their ability to survive in the absence of organic matter and sunlight makes it probable that they are more representative of some of the earliest forms of life on the planet than are surface-dwelling microorganisms. They may also be useful analogs for understanding what life on other planets, such as Mars, might be like. More than 9000 strains of organisms have been cataloged from subsurface habitats.

A variety of chemical reactions are used by Eubacteria and Archaea to provide the energy for biosynthesis. Among the reactions occurring in deep-groundwater is methanogenesis:

$$4H_2 + CO_2 \rightarrow CH_2 + 2H_2O, \quad \Delta G = -32.5 \text{ kcal}$$

The methanogenic reaction uses hydrogen gas as the energy source and inorganic CO_2 as the carbon source. Hydrogen derives from the reaction of the oxygen-poor water with iron-bearing minerals. Methanogens can persist indefinitely without any supply of carbon from the surface.

Taxonomic Survey of Subterranean Life

Many organisms spend some or all of their life cycle in caves, particularly cave entrances. The moist rock of cave entrances provides habitat for many mosses and ferns, such as the Hart's tongue fern (*Phyllitis scolpendrium*). The entrance and twilight zone of a cave are refuges from temperature extremes of the surface. Some species, such as the spider *Meta menardi*, are specialized for the surface–subsurface ecotone at cave entrances. The entrance and twilight zones of caves are relatively predator-free, at least for vertebrates. The neotropic oil bird *Steatornis caripensis* nests in caves; many other birds do so as well, albeit on a less regular basis. Many cave entrances in the eastern US have one or more phoebe (*Sayornis phoebe*) nests. The best-known visitors to caves are bats. Depending on the species, bats use caves as maternity colonies, as hibernacula, and as temporary roosts during the warmer months of the year. More than half of the species of bats found in the eastern US use caves throughout or for part of the year. About 95% of the more than one million gray bats, *Myotis grisescens*, hibernate in eight caves. Many bat species are at risk because of the vulnerability of their hibernating and maternity sites to disturbance. The recently discovered disease, White Nose Syndrome, is rapidly decimating North American cave-inhabiting bats.

Many species spend their entire life cycle in caves. In the case of terrestrial species, troglobionts have an obligate dependence on caves and must complete their entire life cycle in caves. Troglophiles can complete their life cycle in caves, but they can also complete their life cycle in surface habitats. The equivalent terms for aquatic species are stygobionts and stygophiles. Another useful term is troglomorphic, which denotes the morphological syndrome associated with life in caves – loss of eyes and pigment and the elaboration of extra-optic sensory structures. Most but not all troglobionts are troglomorphic. Those that are not, likely have not been isolated in caves for a long enough time for troglomorphic characteristics to evolve, or live in relatively food-rich subterranean habitats.

The remarkable cave amphibian, *Proteus anguinus anguinus*, was the first stygobiont to be mentioned in scientific writing. Writing in 1689, Valvasor reported that local residents thought that this amphibian, common in caves in Slovenia, was the larvae of dragons. Its scientific description was accomplished by Laurenti in 1768. The first troglobiont, the cave beetle *Leptodirus hochenwartii* (**Figure 2**), was described by Schmidt from the same region in 1832. Currently, more than 4000 troglobionts and 2000 stygobionts have been formally named, and at least several times that number probably exist.

In all subterranean environments, a wide variety of heterotrophic bacteria and Protozoa occur, even in deep-groundwater sites. Early interest in autotrophs in caves centered on the potential role of nitrifiers such as Nitrobacter in

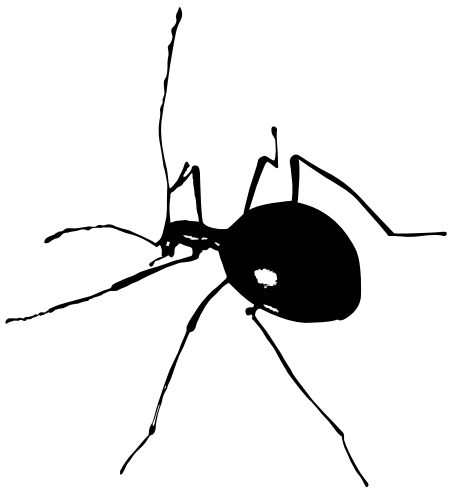


Figure 2 The first described troglobiont, the beetle *Leptodirus hochenwartii* from a cave in Slovenia. The length of the beetle is 8 mm. Reproduced with permission from Matija Gogala, Ljubljana, Slovenia.

the formation of saltpeter (KNO_3). There is a wide diversity of metabolic pathways among autotrophs, most important in the carbon-fixing pathways: the Calvin cycle in bacteria, the acetyl-coenzyme A pathway in Archaea, and the reductive citric acid cycle found in Archaea and a few bacteria. There is also considerable diversity in the electron donor and acceptor among subsurface autotrophs. Electron donors include H_2, NH^{4+}, S, and Fe^{2+}; acceptors include O_2, $SO_4{}^{2-}$, and NO^{3-}.

Except for a few protozoans, the Eukarya is usually not found in deep-groundwater. A wide taxonomic array of organisms is stygobiotic or troglobiotic, including the following phyla: Porifera, Hydrozoa, Platyhelminthes, Nemertina, Nematoda, Mollusca, Rotifera, Annelida, Onychophora, Arthropoda, and Chordata.

Among stygobionts, the Crustacea predominate: At least half of the described species are Crustacea. Among the Crustacea, the orders Remipedia, Spelaeogriphacea, and Thermosbaenacea are exclusively stygobiotic. Numerically, crustacean stygobionts are dominated by Copepoda (especially in the interstitial and the epikarst) and Amphipoda (especially in caves). Some Amphipoda genera are particularly speciose. For example, nearly 175 stygobiotic species of the amphipod genus *Stygobromus* are known from North America, and more than 200 stygobiotic species of the amphipod genus *Niphargus* in Europe. Other important groups include the planarians (Platyhelminthes: Tricladida), snails (Mollusca: Gastropoda), salamanders (Chordata: Caudata), and fish (Chordata: Pisces). The surface-dwelling ancestors of these species were largely benthic, and many species were adapted to environments with dim light and a detrital food base.

Among troglobionts, the Arachnida, Diplopoda, and Insecta predominate. At least 90% of the species are in these groups. Among the Arachnida, spiders (Araneae), false scorpions (Pseudoscorpionida), and harvestmen (Opiliones) are common in temperate caves and in the tropics. Other groups, such as Schizomida, Ricinulei, and Amblypygi, are also important. Among the Insecta, springtails (Collembola) and bristletails (Diplura) are common, but the beetles (Coleoptera) predominate. More than 2000 troglobiotic beetles are known. One genus in the US (*Pseudanophthalmus*) accounts for nearly 250 species, and the European genus *Duvalius* has approximately the same number. Among the cholevid beetles, 23 troglobiotic genera are known from the Pyrenees alone. The surface-dwelling ancestors of many of these species occurred in leaf litter in forests and are believed to have colonized caves during climatic vicissitudes of the Pleistocene.

Although bats are not troglobionts since they spend part of their life cycle outside of caves foraging for food, their dependence on caves is every bit as profound. Among the nearly 1100 described species of bats, at least one-third utilize caves as day roosts, hibernacula, or maternity sites. For many species, especially those in the Vespertilionidae and Molossidae, there is an obligate dependence on caves. Some species form giant hibernating colonies: more than 90% of *M. grisescens* found east of the Mississippi hibernate in three caves in which numbers may reach more than 1 million. The Mexican free-tailed bat, *Tadarida brasiliensis*, forms large summer maternity colonies, with up to 20 million individuals in a single cave. This species declined in numbers due to the widespread use of dichlorodiphenyltrichloroethane (DDT).

Adaptations to Subterranean Life

Compared with related surface-dwelling species, subterranean animals are characterized by the absence or reduction of a series of features, particularly eyes and pigment. Subterranean animals are also characterized by an elaboration of extra-optic sensory structures as well as other morphological, ecological, physiological, and behavioral adaptations (**Table 1**). This represents a remarkable convergence among subterranean animals, the morphological aspects of which are called troglomorphy. Of course, not all features are convergent. For example, interstitial species tend to be smaller than surface-dwelling sister taxa while cave species tend to be larger.

What makes subterranean animals particularly interesting in this regard are the obvious and unusual reductions, often called regressive evolution or "evolution in reverse." Unlike some aspects of evolution, such as the evolution of sex and the origin of evolutionary novelty, eyelessness and regressive evolution in general have always appeared easy to explain. To an orthogeneticist such as Albert Vandel writing in the mid-twentieth century, eye loss in subterranean animals was a by-product of a senescent phyletic line. Animals were not blind because they were in caves; rather, they were in caves because they were blind. To a neo-Lamarckian such as Packard writing in the late nineteenth century, eye loss was a classic example of the role of disuse. To a neo-Darwinian, eye loss in subterranean animals was the result of an evolutionary tradeoff resulting from natural selection for elaborated extra-optic sensory structures. To neutral mutationists, eye loss in cave animals was easily explained by the accumulation of selectively neutral, structurally reducing mutations made possible by the lack of stabilizing selection. In a modern context, theories that account for regressive evolution in subterranean cave animals are classified into two groups: those that involve natural selection directly or indirectly and those that involve neutral mutation and genetic drift (Wilkens *et al.*, 2000).

One of the best-studied organisms in this regard is the amphipod *Gammarus minus* which occurs in springs and caves in much of the central Appalachians in West Virginia and Virginia. Populations independently isolated in caves show a strongly convergent morphology of appendage elongation and eye and pigment reduction. There was a strong correlation between the fitness components of mating and egg number with relative appendage length and eye size. Furthermore, this correlation was opposite for cave and spring populations. However, natural selection is not the whole explanation for eye and pigment reduction in *G. minus*. The rate of change of reduced structures is consistently greater by an order of magnitude than the rate of change of elaborated structures (**Table 2**). Elaborated structures changed as a result of natural selection, whereas reduced structures changed as a result of both natural selection and the accumulation of structurally reducing, selectively neutral mutations. The increased variability of eyes in recently isolated cave populations is also most easily explained as a result of neutral mutation.

The genetics of eye loss are best understood in the Mexican cave fish *Astyanax mexicanus*. Like *G. minus*, it has apparently been isolated in caves for less than 1 My. Cave populations show, relative to surface-dwelling populations, greatly reduced eyes, reduced pigment, reduced optic tecta, and increased jaw

Table 1 Morphological, ecological, physiological, and behavioral adaptations to subterranean life

Morphological	*Ecological and physiological*	*Behavioral*
Reduced eyes, pigment, and wings	Reduced metabolism	Reduced aggregation
Specialization of extra-optic sensory organs	Resistance to starvation	Reduced response to alarm cues
Cuticular thinning	Reduction of circadian rhythm	Increased sensitivity to vibrations
Modification of feeding structures	Reduced egg number; increased egg volume	Reduced intraspecific aggressive behavior
Reduction of scales (fish)	Increased life span	
Elongation of appendages		

Table 2 Evolutionary rates of divergence for different morphological characters of a cave population of the amphipod *Gammarus minus* from its ancestral spring population[a]

Morphological Character		Δ	*Relative rate*
Size	HL	4.0×10^{-7}	0.02
Eyes	EN	2.4×10^{-5}	1.00
	EA	1.5×10^{-5}	0.62
Antennae	P1	8.2×10^{-7}	0.03
	L1	4.1×10^{-7}	0.02
	N1	1.1×10^{-6}	0.05
	P2	1.1×10^{-6}	0.05
	L2	8.9×10^{-7}	0.04
	N2	1.4×10^{-6}	0.06

[a]The measure of divergence, Δ, is independent of the size of the character and depends only on its variability. The two populations compared are Organ Cave and Organ Cave spring, about 2 km apart in southern West Virginia. Time since divergence was set to 5×105 years. Traits were head length (HL), eye ommatidia number (EN), surface area of compound eye (EA), peduncle length of first antenna (P1), flagellar length of first antenna (L1), number of flagellar segments of first antenna (N1), and second antenna analogs (P2, L2, and N2). In the last column, all rates are compared with that of ommatidia number, which was set to 1.

Source: Reprinted from Culver DC, Kane TC, and Fong DW (1995) *Adaptation and Natural Selection in Caves. The Evolution of Gammarus minus*. Cambridge, MA: Harvard University Press.

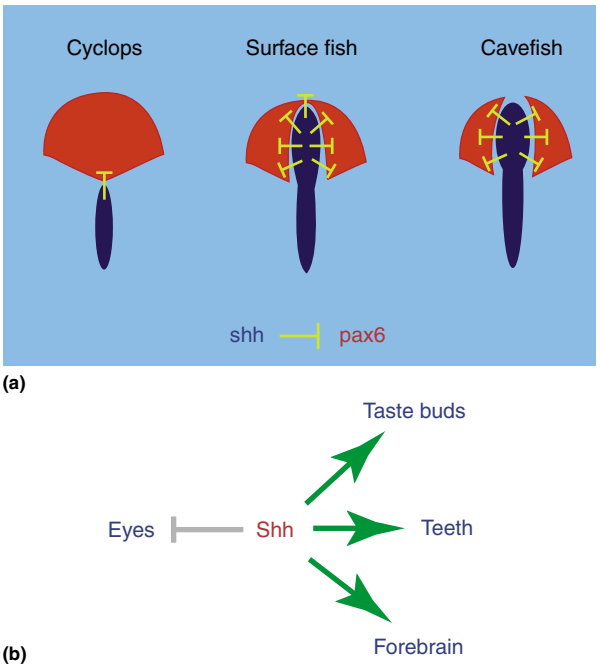


Figure 3 (a) Diagram showing the relationship between *shh* (dark blue) and *pax6* (red) expression at the neural plate stage in cyclopic (left), surface fish (middle), and cavefish (right) embryos. (b) Model showing positive (arrows) and negative (blocked line) pleiotropic effects of shh signaling on constructive traits and eyes, respectively. Reproduced from Jeffery WR (2005) Evolution of eye degeneration in cavefish: The return of pleiotropy. *Subterranean Biology* 3: 1–12, with permission from Société Internationale de Biospéologie.

size, taste bud number, and increased cranial neuromast size and number. The manner of eye development and degeneration, the ability to experimentally restore eyes by lens transplantation, and the analysis of gene expression patterns in *A. mexicanus* make it a particularly useful model for understanding the evolution of eye degeneration. Recent key discoveries are (1) the expansion of Hedgehog (*Hh*) signaling along the anterior midline, which suppresses expression of the *pax6* gene (encoding transcription factor that functions upstream in eye gene hierarchies) in the developing eye fields and (2) control of eye degeneration by Hh-mediated lens apoptosis (cell death) in cavefish embryos. William Jeffery concluded from these and other studies that the best explanation for eye degeneration in *Astyanax* cavefish is natural selection for beneficial traits that are negatively linked to eye development by pleiotropy involving the Hh signaling pathway. The roles of *Hh* and *pax6* genes are shown in Figure 3.

Among subterranean organisms, adaptations have been most thoroughly studied in cave animals. After eye and pigment reduction, the most characteristic troglomorphic traits are the elaboration or elongation of extra-optic sensory structures (Table 1). Among arthropods, the elongation of antennae is especially noticeable. A striking example is the cave beetle *L. hochenwartii* found in the Dinaric Mountains of Italy and Slovenia (Figure 2). Vertebrates such as the North American cave fish in the Amblyopsidae display a striking hypertrophy of the lateral line system. These changes are also reflected in the central nervous system. Areas associated with vision decrease in size and areas associated with the other senses increase (Figure 4). The potential advantage of elaboration and increase in extra-optic sensory structures is to compensate for the absence of light by increasing tactile and chemical sensory sensitivity. In the resource-poor environment of caves, metabolic efficiency is often increased, particularly among species at or near the top of the food chain. Activity is often reduced, as is the metabolic rate. The frequently observed decline in intraspecific aggressive behavior of troglobionts and stygobionts may also in part be an evolutionary response to food scarcity. Food scarcity is also apparently a

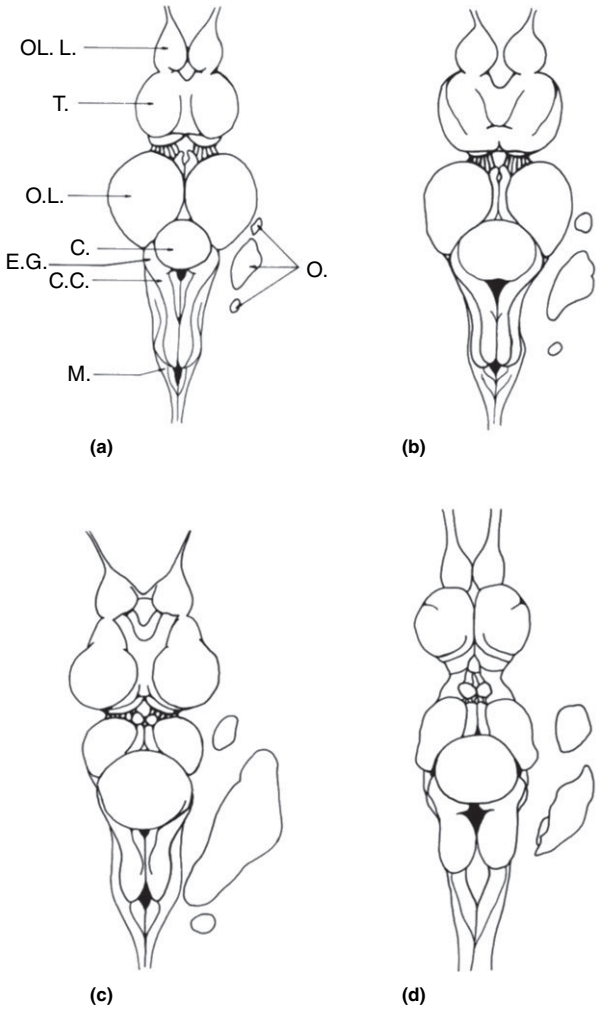


Figure 4 Brain morphologies of four fishes in the family Amblyopsidae: (a) *Chologaster cornuta*, a surface-dwelling species; (b) *Forbesichthys agassizi*, a stygophile found in both springs and caves; (c) *Amblyopsis rosae*, a stygobiont; and (d) *Speoplatyrhinus poulsoni*, a stygobiont. OLL, olfactory lobe; T, telencephalon; OL, optic lobe; C, cerebellum; EG, emenentia granularis; CC, cristae cerebelli; M, medulla oblongata; O, otoliths. Note the relative hypertrophy of extra-optic parts in stygobionts. Reproduced from Poulson TL (1963) Cave adaptation in amblyopsid fishes. *American Midland Naturalist* 70: 257–290, with permission from AMN.

strong selective force in life-history patterns of cave animals. Cave animals live longer, begin reproduction at a later age, produce fewer offsprings, and have larger eggs – all adaptations to a relatively constant, food-poor environment. Perhaps, the most striking example of life-history modifications is that of a population of the crayfish *Orconectes australis australis* in Shelta Cave, Alabama. In an extensive mark-recapture study lasting more than 6 years, John E Cooper found that the age at first reproduction was ~35 years, and that the crayfish lived to an age of more than 100 years (Culver, 1982).

Animals in other subterranean habitats also show many of the same adaptations since they share an aphotic, resource-poor environment, in common with cave animals. However, there are two other environmental factors that exert

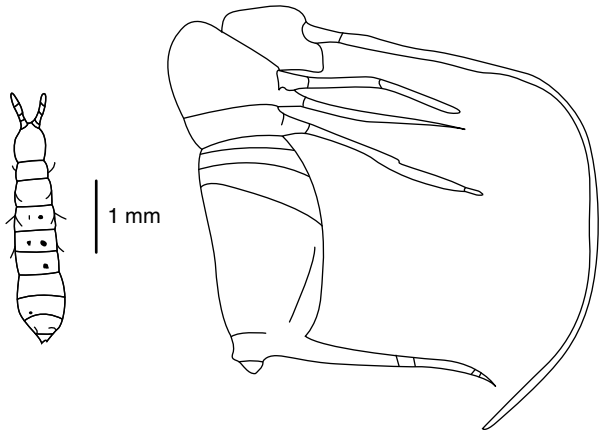


Figure 5 A typical soil collembolan (*Onychiurus armatus*, left) and a highly troglomorphic collembolan (*Pseudosinella christianseni*, right). Reproduced with permission from K. A. Christiansen, Grinnell, Iowa.

strong selective pressures not present in cave environments: intragranular space size in interstitial habitats and oxygen levels in aquatic interstitial habitats. Interstitial animals tend to be both thin and small since they are constrained by interstice diameter. The morphological contrast between cave and interstitial animals can be quite striking. A terrestrial example comparing a cave- and deep-soil-dwelling Collembola is shown in **Figure 5**. The small size of many aquatic interstitial species may occur as a result of progenesis, which is the precocious sexual maturation of an organism resulting in an adult descendent exhibiting the juvenile morphology of its ancestor. This process also results in reduced extremities, spines, and bristles. More than cave dwellers, interstitial stygobionts must be able to survive under conditions of low oxygen. Many interstitial crustaceans are able to regulate respiration rate by varying the ventilating activity of their pleopods according to oxygen availability.

Metabolic adaptations of deep-groundwater microorganisms are just beginning to be understood. Unlike the situation with eukaryotic organisms, the basic tenet of microbial ecology, Beijerinck's principle, holds in the deep subsurface: "Everything is everywhere, the environment selects." The environment in this case is an extremely food-poor one, with little or no organic matter. The major characteristic of phreatic microorganisms is the widespread occurrence of facultative or obligate chemoautotrophy. The Gibbs-free energy available in chemoautotrophic responses is much less than that of photosynthetic organisms, reinforcing the concept of resource scarcity from a physical chemical standpoint.

Colonization and Speciation in Subterranean Environments

The terms troglobiont, stygobiont, troglophile, and stygophile are an ecological classification, which attaches great importance to the isolation of species in subterranean habitats. Troglophiles and stygophiles, species which can complete their life cycle in both surface and subterranean habitats, are often important components of cave communities. In many

cases, especially in the tropics, they outnumber troglobionts and stygobionts. Troglobionts and stygobionts are typically but by no means always troglomorphic. The evolution of troglobiotic (stygobiotic) species that are troglomorphic (stygomorphic) has received the most attention.

There are two components to evolution in subterranean habitats: colonization and subsequent isolation of the subterranean populations usually as a result of the extinction of the surface-dwelling populations. Colonization may be active or it may be passive, that is, "colonization under constraint." An example of active colonization is the movement of surface-dwelling Homoptera in lava fields in Hawaii and elsewhere into lava tubes and cracks in the lava in which root sap is an abundant resource. In some cases, speciation has proceeded without the extinction of the surface populations. The mechanism of speciation is either parapatric or sympatric. Active colonization may be particularly important for interstitial organisms, such as copepods. The classic case of colonization under constraint is that of many troglobionts in north temperate caves. Many of the surface-dwelling ancestors of these troglobionts were leaf-litter-dwelling invertebrates. During Pleistocene interglacials, as warming occurred, many of these species were either forced into caves or up mountain slopes to cooler conditions. As warming continued, surface populations became extinct. For most cases of colonization under constraint, the factor that forced animals into caves, such as Pleistocene interglacials or the Messinian salinity crisis in the Mediterranean, was also the factor that resulted in the extinction of surface populations.

Recent molecular analysis allows a more detailed look at the biogeographic history of the surface- and cave-dwelling isopod *Asellus aquaticus*. The species itself is apparently very old, and is part of a Paratethys fauna (8–12 My) unrelated to more recent Mediterranean invaders. In the Dinaric karst in Slovenia and extreme northeast Italy, *A. aquaticus* has invaded cave systems at three separate times between 3.6 Ma and 800,000 years ago, including two colonizations of the Postojna-Planina Cave System (PPCS). The oldest lineage split is after the Messinian Salinity Crisis but before the Pleistocene. The present pattern is the result of fragmentation and migration (**Figure 6**).

By whatever mechanisms organisms colonize subterranean environments and by whatever geological or climatological event surface populations become isolated in subterranean environments, subsequent dispersal is generally extremely

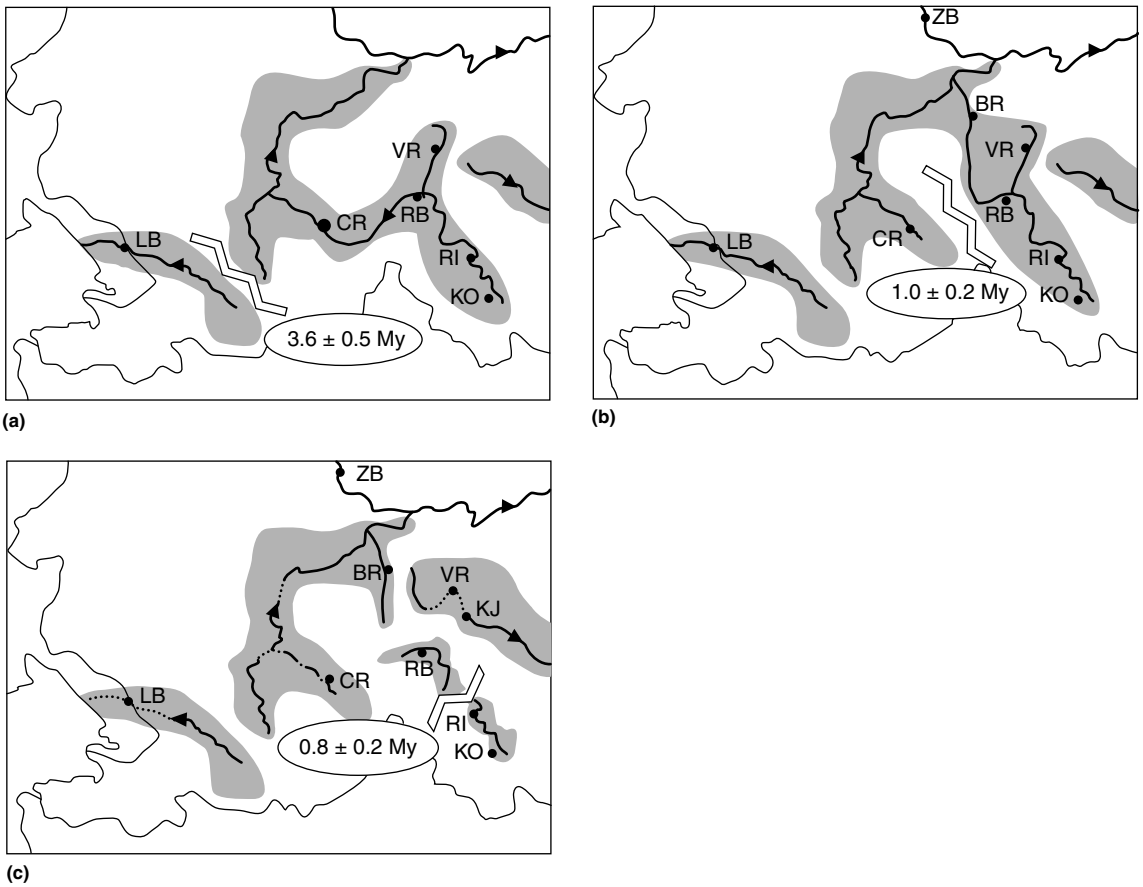


Figure 6 Graphical representations of a hypothetical chain of events that best explains the current distribution of haplotypes of *Asellus aquaticus* in the Slovenian Dinaric karst. A zigzag line indicates fragmentation events that separated drainages. Approximate dimensions of the map are 160 × 120 km. Localities are BR, Ljubljansko Barje (ditch); CR, Cerknica Polje; KJ, Krka Cave; KO, Kočevje Polje; LB, Trebiciano Cave (Italy); RB, Ribnica Polje (spring); RI, Kočevje Polje-Rinža River; VR, Vršnica Cave; and ZB, Zbilje Reservoir (Sava River). Reproduced from Verovnik R, Sket B, and Tronteli P (2004) Phylogeography of subterranean and surface populations of water lice *Asellus aquaticus* (Crustacea: Isopoda). *Molecular Ecology* 13: 1519–1532, with permission from Wiley.

limited. Most subterranean species have very restricted ranges. For example, more than 60% of the obligate subterranean-dwelling species described from the US are known from a single county, and about half of these are known from a single locality. Species with large ranges, at least for subterranean species, have on closer examination been shown to be complexes of morphologically identical but genetically distinct species and semispecies. For example, the European cave salamander *P. anguinus* ranges more than 500 km from Italy to Montenegro but actually represents several invasions of a common surface-dwelling ancestor that is extinct. For some species, there is little gene flow at distances of less than 1 km. This makes subterranean species excellent markers of biogeographic events. Once species are isolated in subsurface habitats, which are often a refuge from the vicissitudes of surface climatic change such as temperature and salinity, their distribution may leave a record of ancient events, especially continental drift. Examples abound in the Crustacea. The distributions of the order Syncarida and the amphipod family Bogidiellidae are Pangaean and predate the pre-Triassic breakup of Pangaea. The stygobiotic copepod genera *Rheocyclops* from North America and *Graeteriella* and *Speocyclops* from Europe are probably closely related taxa that differentiated after the breakup of Laurasia; their isolation presumably dates back to the late Mesozoic. The order Remipedia and the ostracod genus *Danielopolina* have a Tethyan distribution. The distribution of the amphipod genus *Longipodacrangonyx* in Morocco closely matches the embayments of the Eocene sea, with little movement since then (**Figure 7**). In most of these cases, extinction of surface populations was a result of marine regression. Of course, not all subterranean species are as old as the examples given here, but they do indicate the apparent great age of some subterranean groups combined with highly restricted dispersal.

A final example is the remarkable cave clam, *Congeria kusceri*, the only known stygobiotic clam. Known from a series of caves in Bosnia and Herzegovina, Croatia, and Slovenia, *C. kusceri* is the only living representative of an otherwise extinct genus, known from the late Miocene. Marine ancestors were apparently victims of the Messinian salinity crisis 6 Ma, and *C. kusceri* is likely a relic of Pliocene lakes and rivers in the region. As the lakes and rivers disappeared under ground, *C. kusceri* went with them.

Some Representative Subterranean Communities

Epikarst Community in PPCS, Slovenia

Epikarst communities are best sampled by collection and filtration of drip water over extended periods of time. Animals in dripping water are the ones that have been swept out of their primary habitat above the cave passage. Direct sampling of epikarst habitats is not possible due to the small size of the cavities and their inaccessibility. The best-sampled epikarst community is that of PPCS, where 20 drips were continuously sampled for more than 1 year. The drips exhibited considerable environmental heterogeneity. Drip rate itself ranged between 1.7 and 202 ml min^{-1} (mean = 35.3, median = 8.8). Conductivity varied between 238 and 548 μS cm^{-1}, probably

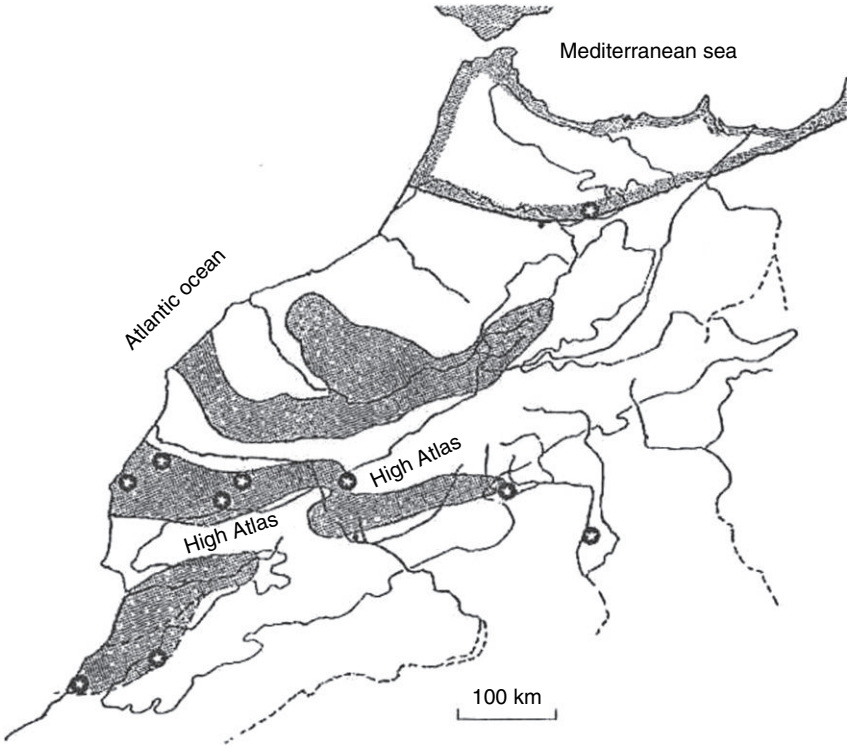


Figure 7 Localities of all known sites (stars) of *Longipodacrangonyx* amphipods plotted on a sketch map of Morocco showing the three marine embayments during the Eocene (shaded areas). Reproduced from Boutin C (1994) Phylogeny and biogeography of metacrangonyctid amphipods in North Africa. *Hydrobiologia* 287: 49–64, with permission from Springer.

reflecting different residence times of the water in the subsurface. A total of 23 copepod species (4 Cyclopoida and 19 Harpacticoida) were found in the drips, including 8 undescribed species. Six of these 23 species are endemic to PPCS. A total of five species were in a single genus – *Bryocamptus*. Species composition was extremely heterogeneous; differences in drip rate, ceiling height, and nitrate concentrations were correlated with differences in species composition. There was also a strong distance effect, with drips closer than 100 m being more similar than those farther away. The scale of variation of epikarst communities is remarkable – even drips less than 100 m apart often have different species. This probably reflects the semi-isolated nature of the small cavities that comprise the epikarst (**Figure 1**).

Cave Stream Community in PPCS, Slovenia

Consisting of 17 and 6 km of passages, respectively, connected by 2 km of flooded corridors, the PPCS has more known species of stygobionts than any other cave or other subterranean site. The sinking river in the main passages is inhabited by a rich assortment of stygobionts, stygophiles, and accidental surface species. Both streams drain a catchment area of more than 100 km^2. Above the cave passages is the rich epikarst community. Among the 48 stygobionts is the European cave salamander *P. anguinus anguinus*. Both the snail (8 species) and crustacean (16 species) faunas are rich. The marine origin of some of the stygobiotic species is evident in the hydrozoan *Velkovrhia enigmatica* and the cirolanid isopod *Monolistra racovitzae racovitzae*. Three species of the amphipod genus *Niphargus*, as well as populations of *A. aquaticus* were isolated in the cave at different times. The Postojna-Planina Cave System is one of the best-studied caves in the world, and parts of it have been heavily visited by tourists since 1818.

Phreatic Community of Edwards Aquifer, Texas, USA

San Marcos Spring is an artesian spring that serves as an exit for the Edwards Aquifer, a large cavernous limestone reservoir that lies along the Balcones escarpment. The aquifer occupies an area of 10,000 km^2. Two thirds of the 27 stygobionts are crustaceans, and endemism is high as well. The subterranean amphipod fauna in this system may be the most taxonomically diverse of its kind in the world and includes 10 species of gammaridean amphipods, representing eight genera in four families. The vertebrate stygobiotic fauna is rich with two catfishes (*Satan eurystomus* and *Trogloglanis pattersoni*) and the Texas blind salamander *Typhlomolge rathbuni*. Oil and peat deposits above the aquifer as well as long-term and gradual input of organic material into the aquifer through numerous sinkholes on the adjacent Edwards Plateau are the primary inputs into the system. The aquifer is threatened by excessive draw down, both by agricultural interests and by the city of San Antonio. Lowering of the aquifer also threatens species limited to the springs of the Edward Aquifer, including two fish (*Etheostoma fonticola* and *Gambusia georgei*) and two aquatic beetles (*Heterelmis comalensis* and *Stygoparnus comalensis*) on the US Endangered Species List.

Anchialine Community of Walsingham Caves, Bermuda

This complex of anchialine caves (ones with a direct connection to the sea) is ~1 km long, with most of the passages submerged. With a freshwater lens and a redox boundary between fresh and saltwater, Walsingham Cave likely has considerable productivity as the result of chemoautotrophic and heterotrophic production. The fauna is entirely aquatic, with 37 species, 29 of which are crustaceans. Among the crustaceans is a representative of the exclusively subterranean order Mictacea and rich ostracod and copepod faunas with six species each.

Spring Community of Organ Cave Resurgence, West Virginia, USA

Organ Cave is a large, complex cave aligned along the axis of a minor syncline and has more than 40 km of surveyed passages, most of which have permanent streams. The cave waters with a catchment area of ~15 km^2 drain at a single resurgence, which empties directly in Second Creek. Base-level discharge is about 125 l s^{-1}. As is typical of large karst springs, environmental conditions (temperature, chemistry, etc.) are very stable, and the water resurging at the spring has been underground at least as long as the cave water, in this case typically 3–6 days. A total of 19 macroscopic invertebrate and two fish species have been found in the cave. Only four of these species have been found in the spring. Of these, three are rare stygophiles (a snail, a stone fly, and an amphipod). The fourth, *G. minus*, is a genetically and morphologically distinct population from that of the cave. This separation is remarkable, given that the cave has been explored to within 1 km of the resurgence. The remaining 1 km is probably a water- and sediment-filled tube.

Hygropetric Community of Vjetrenica, Bosnia and Herzegovina

Vjetrenica is a moderately sized (8 km) cave on the edge of Popovo Polje about 20 km from the Adriatic Sea. It has very diverse aquatic habitats including a 180 m long lake (Veliko Jezero) and hygropetric habitats. Thin sheets of permanently flowing water over vertical rock make up the hygropetric habitat. Two species of Coleoptera in the subfamily Leptodirinae (*Hadesia vasiceki vasiceki* and *Nauticiella stygivaga*) are specialists in the hygropetric, apparently filter feeders. Other species, both aquatic (the amphipod *Typhlogammarus mrazeki*) and terrestrial (the isopod *Aspioniscus*) are often found there as well. In general, the cave has a very rich stygobiotic and troglobiotic fauna (see **Table 3**).

Chemoautotrophic Community of Movile Cave, Romania

Movile is a small (<300 m), mostly water-filled cave near the coast of the Black Sea with no natural connection to the surface. Extensive chemoautotrophic production in the form of sulfur oxidation occurs in the cave system. Eighteen stygobionts are known from the cave, and the chemoautotrophic production also supports a terrestrial community of 27 troglobionts. About two-thirds of the fauna is endemic to the

Table 3 Caves with more than 25 stygobionts (A) or 25 troglobionts (B). Data compiled from various sources (see Culver and Pipan 2009)

Site name	*Country*	*Number of species*	*Region/ecology*
A. Stygobionts			
Postojna Planina Cave System	Slovenia	48	Dinarides
Vjetrenica	Bosnia & Hercegovina	40	Dinarides
Walsingham Cave	Bermuda	37	Anchialine
Triadou Aquifer	France	34	Phreatic
Robe River	Australia	32	Phreatic
Križna jama	Slovenia	29	Dinarides
Logarček	Slovenia	28	Dinarides
Šica-Krka System	Slovenia	27	Dinarides
Edwards Aquifer	Texas, USA	27	Phreatic
B. Troglobionts			
Postojna Planina Cave System	Slovenia	36	Dinarides
Cueva de Felipe Reventón	Canary Islands, Spain	36	Lava tube
Vjetrenica	Bosnia & Hercegovina	30	Dinarides
Peştera Movile	Romania	29	Chemoautotrophic
Cueva del Viento	Canary Islands, Spain	28	Lava tube
Cueva del Sobrado	Canary Islands, Spain	28	Lava tube
Mammoth Cave	Kentucky, USA	26	Longest cave

cave or groundwater system that the cave intersects. Among the unusual components of the fauna are an endemic leech and an endemic water scorpion (*Nepa anophthalma*). Based on analysis of mitochondrial DNA sequences, *N. anophthalma* was isolated in the cave 2 Ma. The high diversity is made possible by the high productivity, and the high endemism is a result of the lack of contact with the surface.

Communities in Groundwater Calcretes, Western Australia

In arid Western Australia, a series of shallow aquifers of Oligocene age, formed by groundwater near the surface, have been colonized by a remarkable variety of stygobionts, including diving beetles, amphipods (Crangonyctoidea and Melitoidea), and isopods (Phreatocoidea). Aside from organic matter in percolating water, tree roots that penetrate the calcrete are an important food source. Although regional diversity is high, overlap among related species is low, and different isolated calcretes have different but related species. The fauna was probably isolated in these calcretes during the drying periods of the Tertiary.

Interstitial Communities in Lobau Wetlands, Austria

Formed by a meander arm, the Lobau wetlands are part of the floodplain of the Danube River near Vienna. Extensive studies of the superficial gravel sediments, which are up to 20 m thick, using pumps and minivideo cameras in shallow wells revealed a complex habitat with areas of differing oxygen and permeability and a rich fauna. In a 1 km^2 area, more than 100 species were found, one third of which were stygobionts. Stygobionts were more common at depths of greater than 5 and 100 m away from surface water. The microcrustaceans – Copepoda and Ostracoda – were particularly diverse.

Interstitial Communities of Rhone River, France

After the Danube, the Rhone River is the most thoroughly studied river with respect to its interstitial fauna. Thirty-eight stygobiotic species, dominated by 23 crustaceans and 10 oligochaetes, have been found in the interstitial waters of the Rhone. There are considerable faunal differences depending on the distance from the main channel of the river, depth, and upwelling and downwelling.

Hypotelminorheic Communities of Rock Creek, District of Columbia, USA

Rock Creek is a small urban stream in the metropolitan Washington, DC, area. Most of the stream and a substantial portion of the watershed is in parkland. Small seeps, with catchment areas usually less than 0.01 km^2 and underlain by clays at a depth of 25–50 cm, occur along the banks of the creek, usually 20–50 m above the stream. Five species of the amphipod genus Stygobromus as well as the isopod *Caecidotea kenki*, and the snail *Fontigens bottimeri* are found in these seeps. Compared with other subterranean habitats, organic material is abundant and all of the seeps are lined with dead and decaying leaves. Most, if not all of the seeps, dry up at least occasionally and animals likely survive by burrowing into the clay layer.

Terrestrial Community of Gua Salukkan Kallang-Towakkalak, Indonesia

This immense river cave system, with more than 20 km of passage and a large bat population, is unusual among lowland tropical caves in the richness of its troglobiotic fauna. Food resources are abundant in the cave. There are large amounts of flood debris along the riverbanks, and bat and swiftlet guano is scattered along the underground galleries. In addition to

troglobionts, there are many species specialized on guano as well. Arachnids and Collembola dominate the 21 troglobiotic species.

Terrestrial Community of Mammoth Cave, Kentucky, USA

The longest cave in the world with more than 500 km of passage, Mammoth Cave harbors 28 troglobionts. Most of the cave lies within the boundaries of Mammoth Cave National Park. The long length of the cave is due in part to a sandstone caprock that reduces erosion and loss of upper-level passages. In the large upper-level passages with sandy-bottomed floors (an unusual cave habitat), trogloxenic cave crickets (*Hadenoecus subterraneus*) deposit their eggs. Since *H. subterraneus* regularly leave the cave to feed, their eggs are an important energy source for beetles. Predaceous beetles, *Neaphaenops tellkampfi*, specialize in predation on cricket eggs. Such specialization is unusual because of the scarcity of resources in most caves. Also noteworthy is the co-occurrence of five species of trechine beetles in the closely related genera *Neaphaenops* and *Pseudanophthalmus*.

MSS Communities of Ariège, France

The MSS is a subterranean habitat that occurs typically in mountainous regions in deposits of scree not infilled by soil. In its structure, it bears some resemblance to dry epikarst. Communities are similar to nearby terrestrial cave communities, but with some species unique to the MSS. Among these are the cholevid beetles such as *Speonomus colluvii* and *Mehadiella paveli*. The majority of species, both troglophiles and troglobionts, are found in both caves and the MSS – arachnids, beetles, and Collembola.

Terrestrial Epikarst Communities of Organ Cave, West Virginia, USA

In a sampling of 13 drips in Organ Cave for a period of 30 days, 176 terrestrial arthropods were collected in dripping water, indicating an abundant arthropod fauna in the epikarst. Nearly all of these individuals, except for a few troglobiotic Collembola, were troglophiles. Epikarst communities in general have been little studied.

Terrestrial Communities of Vjetrenica, Bosnia and Herzegovina

Vjetrenica is a moderately sized (8 km) cave on the edge of Popovo Polje. It is one of the most faunistically rich caves in the world (see **Table 3**) largely due to its geographic position and ecological heterogeneity. The terrestrial fauna of 30 troglobionts depends on food brought in by moving water, and to a much lesser extent by animals moving in and out of the cave. The most common species in the main corridor is *Antroherpon apfelbecki*. Some species have very restricted distributions within the cave; for example, the centipede *Typhloglomeris caeca* is limited to large clay deposits.

Terrestrial Lava Tube Community of Kazumura Cave, Hawaii, USA

Located on the island of Hawaii, it is the longest lava tube in the world, more than 60 km long, covering a straight line distance of 32.2 km, and a vertical extent of more than 1000 m. Delicate tree roots penetrate the thin soils and lava, and provide a food source for the terrestrial fauna. Among the species feeding on the roots are a cixiid plant hopper, *Oliarus polyphemus*, a noctuid moth in the genus *Schrankia*, and a gryllid cricket, *Caecomobius varius*. It is noteworthy for both the main food source and the unusual species (at least for subterranean habitats).

Bat Colony of Bracken Cave, Texas, USA

More than 20 million Mexican free-tailed bats (*T. brasiliensis*) form the largest known bat maternity colony in the world, producing millions of young each year. It is one of the largest, if not the largest, concentrations of vertebrate biomass. About 30 km from downtown San Antonio, it is protected by Bat Conservation International. The concentration is large enough to substantially increase the temperature of the roost environment.

Geography of Cave Biodiversity

Although other subsurface habitats have not been sufficiently studied to analyze biogeographical patterns, caves have been. On a worldwide basis, more than 10,000 caves have been biologically investigated, and there is an extensive amount of information about the number of stygobionts and troglobionts at individual caves and karst wells. Because of the restricted possibilities for dispersal and the restricted abilities of stygobionts and troglobionts to disperse, species numbers in any one cave or karst well are low relative to the regional fauna. On a worldwide basis, there are not sufficient data to analyze regional faunas of similar areas, but there are enough data to analyze the geographic pattern of fauna-rich caves.

The nine caves and karst aquifers with more than 25 stygobionts are listed in **Table 3a**. In this case, a strong

Table 4 Species richness of epikarst copepods in well-sampled caves

Cave	*Country*	*Total number of Copepods*	*Number of Stygobiotic Copepods*
Velika Pasica	Slovenia	12	9
Postojna-Planina Cave System	Slovenia	23	16
Škocjanske Jame	Slovenia	17	10
Dimnice	Slovenia	12	11
Županova Jama	Slovenia	16	14
Organ Cave	US	10	7

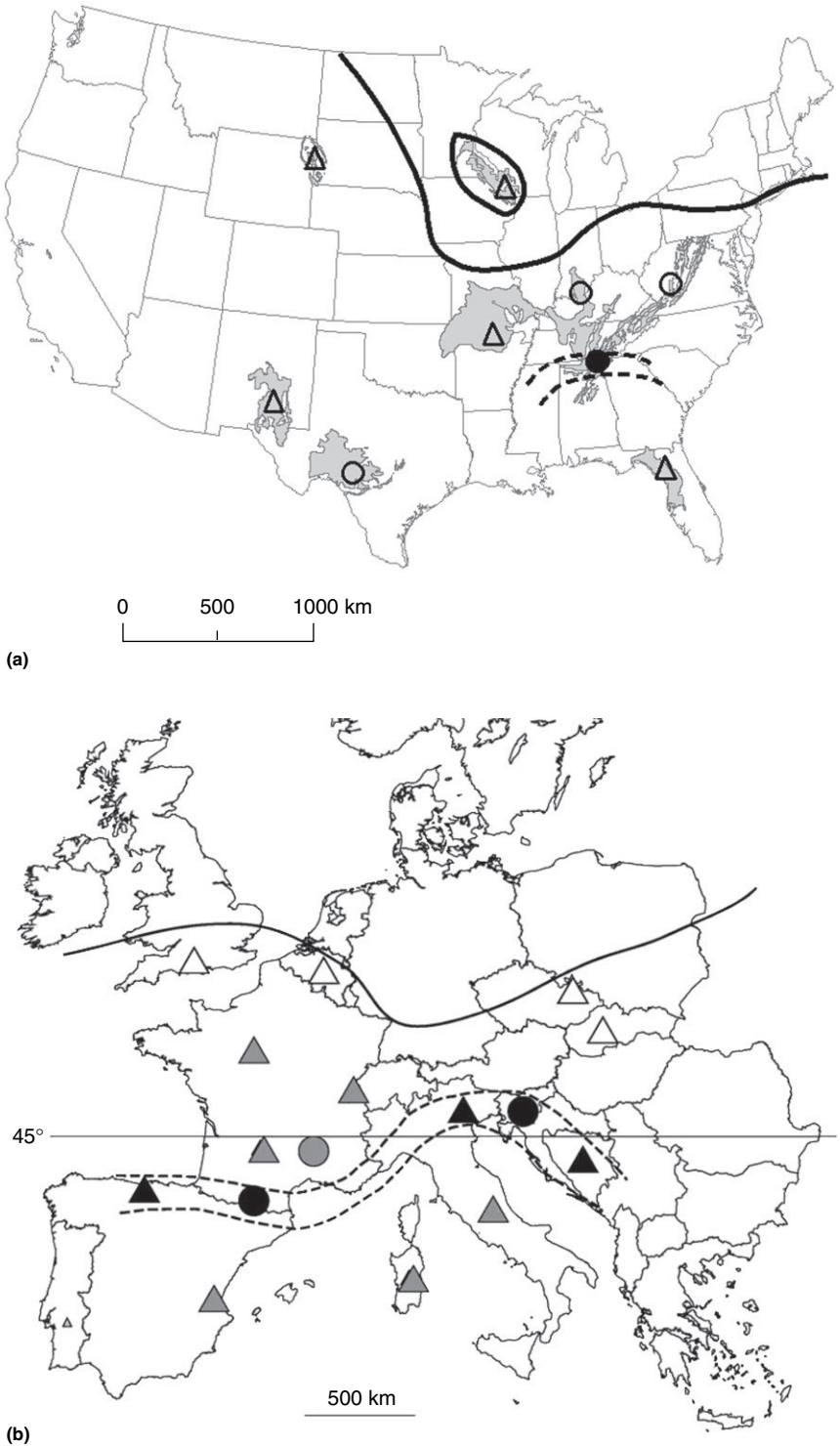


Figure 8 (a) Map of species richness patterns of North American troglobionts. Major karst areas of eastern and central US are shown in light gray. Although there are many caves in the western United States and Canada, there are no large karst areas, and no areas of rich fauna. The open triangles are areas with few if any troglobionts, the gray triangles are areas with less than 50 species, usually much less than 50. The three open circles are three areas analyzed in this study with less than 50 species in 5000 km^2 of area or less. The black circle is the diversity hotspot in northeast Alabama. The boundary of the Pleistocene ice sheet is shown as a solid line. A pair of dashed lines indicates the hypothesized position of the high-diversity ridge. (b) Map of species richness patterns of European troglobionts. The open triangles are areas with few if any troglobionts, the gray triangles are areas with less than 50 species, usually much less than 50, and the gray circle is Ardeche, with less than 50 species in 5000 km^2 of area or less. The black circles are the diversity hotspots in Slovenia and Ariege. Black triangles are other possible diversity hotspots. The boundary of the Pleistocene ice sheet is shown as a scored solid line. A pair of dashed lines indicates the hypothesized position of the high-diversity ridge. Adapted from Culver DC, Deharveng L, Bedos A, *et al.* (2006) The mid-Latitude biodiversity ridge in terrestrial cave fauna. *Ecography* 29: 1, 120–128.

geographical component is evident. More than half of the caves are in the Dinarides. The Dinaric karst has two unique features – it is adjacent to the Mediterranean Sea and karst development as measured by cave passage density is higher than elsewhere. Adjacency to the Mediterranean afforded opportunities for invasion of marine taxa during periods of drying, such as the Messinian salinity crisis. The remaining cases, with one exception, are entirely phreatic or with major phreatic components. The exception is Walsingham Cave in Bermuda, a high-energy anchialine cave. Thus, the mark of history is obvious with the predominance of examples from one geographic region – the Dinaric karst. The reason that phreatic habitats are so well represented is less clear but it may be that the contiguous area of these aquifers is much greater than the contiguous areas of other aquatic subterranean habitats.

The seven caves with more than 25 troglobionts are listed in **Table 3b**. Both ecological and geographic components are evident. The ecological component is that high-productivity caves, those in Canary Islands very close to the surface (three caves) and caves with high chemoautotrophy (one cave), are well represented. For the remaining caves, two are in the Dinarides. The exception is Mammoth Cave in the US, which is aberrant because of its great length (>500 km), much longer than any other limestone cave. The caves in the Dinarides are also in karst regions of Europe with high secondary productivity. Thus, both productivity and history play a role.

Not included in the above analysis is the fauna of epikarst. Epikarst in only a handful of caves has been sampled but it is clear from these results that the richness of epikarst stygobionts may rival that of the stygobiotic fauna in other subterranean aquatic habitats (**Table 4**). In the six caves in which the epikarst has been extensively sampled, the number of stygobiotic copepods ranged between 7 and 16! In several of these caves (e.g., Organ Cave, Županova Jama, and Škocjanske Jame), stygobiotic epikarst copepod species outnumber all other stygobionts known from the particular cave. The potential for discovery of new species in epikarst is enormous.

Although there are insufficient data for a global comparison of species richness by region, there are considerable data available for Europe and North America (**Figure 8**). When extensively sampled cave regions of between 2000 and 5600 km^2 are compared, a ridge ($\sim$42–46° in Europe and 34° in North America) of high biodiversity occurs in temperate areas of high productivity and cave density. As was the case with individual caves of high troglobiotic diversity, the pattern reflects a strong dependence of terrestrial cave communities on long-term surface productivity (as reflected in actual evapotranspiration). This dependence may explain the unique biodiversity pattern of terrestrial cave invertebrates.

Conservation and Protection of Subterranean Habitats

Subterranean sites present special conservation and protection problems. Literally out of sight, they are often neglected in decisions concerning conservation priorities. When they are considered, the high levels of endemism make it difficult to set priorities among endemic species. Once the decision has been made to protect cave and other subsurface sites, several principles apply. The protection of bat sites is most effectively done by gating the cave. However, there is considerable risk to the bats if the gate is not properly designed. Improperly designed gates can change environmental conditions in the cave, making it unsuitable for bats, and if the bars on the gate are not properly spaced bats can suffer considerable predation at the gate. Second, protection of the subterranean aquatic fauna requires protection of the surface riparian habitat. In the case of cave streams, their upstream sources, whether they are sinking streams or sinkholes, must be protected. In the case of epikarst species, it is the surface immediately above the cave that must be protected. Similar protection schemes are needed for the terrestrial riparian fauna. Third, protection of deep-groundwater species requires protection from both excessive draw down and contamination of the aquifer. Fourth, protection of troglobionts that rely on the flow of organic matter into the cave entrance, such as those *Neaphaenops* beetles that are predators of cricket eggs, must include protection of the entrance area and any foraging area of the species responsible for bringing organic matter into the cave. Finally, conservation and protection of subsurface sites requires greater public awareness and appreciation.

See also: Coastal Beach Ecosystems. Ecosystem Function Measurement, Aquatic and Marine Communities. Endangered Ecosystems. Estuarine Ecosystems. Lake and Pond Ecosystems

References

Botosaneanu L (ed.) (1986) *Stygofauna mundi*. Leiden, The Netherlands: Brill.
Boutin C (1994) Phylogeny and biogeography of metacrangonyctid amphipods in North Africa. *Hydrobiologia* 287: 49–64.
Culver DC (1982) *Cave life*. Cambridge: Harvard University Press.
Culver DC and Pipan T (2009) *The Biology of Caves and Other Subterranean Habitats*. Oxford: Oxford University Press.
Culver DC, Deharveng L, Bedos A, *et al*. (2006) The mid-latitude biodiversity ridge in terrestrial cave fauna. *Ecography* 29: 120–128.
Culver DC, Kane TC, and Fong DW (1995) *Adaptation and Natural Selection in Caves. The Evolution of Gammarus Minus*. Cambridge: Harvard University Press.
Jeffery WR (2005) Evolution of eye degeneration in cavefish: The return of pleiotropy. *Subterranean Biology* 3: 1–12.
Pipan T (2005) *Epikarst – A promising habitat*. Ljubljana, Slovenia: ZRC Publishing, Karst Research Institute at ZRC-SAZU.
Poulson TL (1963) Cave adaptation in amblyopsid fishes. *American Midland Naturalist* 70: 257–290.
Verovnik R, Sket B, and Trontelj P (2004) Phylogeography of subterranean and surface populations of water lice *Asellus aquaticus* (Crustacea: Isopoda). *Molecular Ecology* 13: 1519–1532.
Wilkens H, Culver DC, and Humphreys WF (eds.) (2000) *Subterranean Ecosystems*. Amsterdam: Elsevier.

Succession, Phenomenon of

HH Shugart, University of Virginia, Charlottesville, VA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Autotrophic successions These generate energy from internal processes (photosynthesis).
Climax community According to some theories of succession, the end result of succession in which successional change ends with a community that does not change and that is in equilibrium with the climate.
Disturbance Major alternations in vegetation due to events such as wildfires, hurricanes, landslides, and human clearing.
Heterotrophic successions Dependent on already fixed energy, such as the successional of communities associated with the decomposition of dead logs.
Individualistic view of succession The concept that succession is a consequence of species interacting with one another and their environment.
Primary succession Succession on newly exposed substrates such as a sandbar or rubble at the foot of a receding glacier.
Progressive succession Successions in which the dynamic changes are in the directions of increasing species diversity, structural complexity, greater biomass, and increased stability. Retrogressive successions are in the opposite directions.
Secondary succession Succession on existing substrate (soil) following a disturbance.
Succession The pattern of change expected in a community over time after a disturbance or after a new substrate has been exposed.

Introduction

Ecological succession is an ordered progression of structural and compositional changes in ecosystems toward an eventual stable condition. Primary successions are initiated on new substrates, such as a new volcanic island, a new sandbar in a river, or rubble fields at the foot of a receding glacier. Secondary successions involve the recovery of vegetation on established soils from land abandonment and from disturbances such as wildfires, hurricanes, or human alterations to vegetation. Other significant dichotomies used to categorize succession involve whether a given successional sequence is

1. Progressive (dynamically changing in the directions of increasing species diversity, structural complexity, greater biomass, and increased stability) or retrogressive (in the opposite directions).
2. Autotrophic (generating energy from internal processes) or heterotrophic (dependent on already fixed energy, such as the successional of communities associated with the decomposition of dead logs).
3. Autogenic (changing due to interactions from inside the system) or allogenic (changing in response to changes in external variables).

Ecological succession is an early concept in ecology and was essential in the early definitions of ecological communities. Although the basic concepts of succession are easily understood, debates about succession have led to considerable confusion and discussion. The mechanisms that drive ecological succession and its very existence as a natural phenomenon have been the subject of continual debate among ecologists. McIntosh (1985) identifies two contrasting views that typify traditional natural history and can also be found in current discussions of theoretical ecology. This dichotomy, which can be used to organize theories about ecological succession, contrasts mechanistic explanations and organic (holistic) explanations of the causes of succession. For mechanistic theories about succession, known laws explain the actions of the individual parts of a system and the whole system is the sum of these parts and their interactions. In the case of organic or holistic explanations, the whole system, its existence and design, explains the actions of the parts. Mechanistic explanations are often considered as the more modern interpretation of succession, but these modern views may have been the first that developed and certainly developed as early as the organic views. Organic explanations of succession were most popular, at least in the US, between the 1920s and the early 1950s and have had a considerable impact on land use and conservation policies.

Organic Explanations of Ecological Succession

Organic explanations seek to understand succession in terms of the principles that operate at the level of the whole system. In the 1920s and 1930s, Clements and particularly John Phillips were ascribing to ecological communities the attributes of a superorganism – a highly organized and coevolved assemblage of plants and animals interacting in a dynamic system. The ecosystem concept had its roots in debates regarding the organization and dynamics of natural systems. It was Tansley's negative view of Clements' and Phillips' interpretation of the community as a superorganism that, in 1935, inspired his development of the ecosystem concept as an alternative to the organic term "community."

Clementsian Concepts of Ecological Succession

Between 1905 and 1935, Clements promoted a dynamic plant ecology built around a supraorganismic view of the ecological

community. At the time of their conception, Clements' ideas represented a significant emphasis on the dynamics of vegetation and were an important early attempt to develop a formal theory predicting the pattern and expected change in ecological communities. The underlying conviction was that evolution and internal interactions would produce a homogeneous regional climax vegetation or community of regular species composition. The development of this climatic climax community was ecological succession, which was viewed as the community analog of the embryological interactions that produce an organism. In this view, succession was typified by a progressive sequence of seral stages or seres, communities that sequentially replaced one another over time until the climax stage was reached. Successional development was the result of a set of processes:

1. Nudation – the creation of a bare area (or partially bare area) to initiate succession.
2. Migration – the arrival of organisms at the location.
3. Ecesis – the establishment of organisms at the location.
4. Coaction – the interactions, particularly competition, among the organisms.
5. Reaction (or facilitation) – the modification of the site by the organisms and the subsequent change in the relative abilities of the organisms to establish and survive.
6. Stabilization – the development of a stable community called the climax community.

Stabilization is less a process and more a consequence of the iterative reapplication of the migration, ecesis, coaction, and reaction processes until a stable community is reached. Clements' concept of the climax community was that there was only one stable vegetation type that was in equilibrium with the regional climate. Succession was the orderly, predictable, progressive, and linear development toward this climax community.

Application of Clementsian Concepts

The Clementsian succession paradigm had a pronounced effect on ecology in the US in the first half of the twentieth century (much less so in Europe), and it shaped many of the laws and policies on the use of public lands. Earlier ecologists developed some of these ideas, and others evolved over Clements' and associates' scientific careers. Important aspects in the Clemetisian paradigm, many of which continue to be a part of natural land management today, are as follows.

Community-Level Management

The idea that one could use the state of ecological communities to evaluate their past and present conditions and to predict their future is a significant contribution of Clements and associates. Although it has been established by paleoecological studies that communities are not necessarily stable over century and millennial timescales, short-term changes and disturbances are often seen as changes in communities. Conservation groups often make considerable effort to preserve unique communities (as well as unusual or important species). Wildlife and endangered species management is often based on maintaining particular communities or habitat types appropriate to the survival of focal species of animal or plants.

Indicator Species Concept

Clements believed that the presence or absence of particular species could be used to assess the state of a community and its potential for agricultural conversion. For example, the presence of poisonous or distasteful plants on a range indicates overgrazing, the occurrence of certain species of Lupine (*Lupinus plattensis*) in a Nebraska prairie connotes a deep soil suitable for tilling and agriculture, and land with plants such as *Salicornia* might be so loaded with salt as to be unreclaimable. One could use indicator species to determine the past history of a landscape and to predict its future changes and potential uses.

The Climax Community Concept

The climax community is the community that is stable in a given climate condition and is the ultimate product of successional processes. Current vegetation maps, particularly for large regions, display the expected climax community or the potential natural vegetation expected in a region rather than the actual vegetation found there. Parks and wilderness areas are managed to maintain the natural climax community that is typical of the region. As an issue in conservation of diversity, often considerable effort is made to preserve unique communities (as well as species).

The Progressive Nature of Succession

Clements viewed succession as being progressive (moving in a positive direction) toward the climax community. The succession progressed toward more diverse, more stable, and more desirable communities. This is a concept associated with eighteenth-century intellectuals and tied to other concepts such as the divine design of natural systems and ideas about the balance of nature and of the antiquity of certain ecological communities. Whether from Clements or earlier sources, these ideas have a considerable influence on the esthetics of conservation in the valuation of wilderness and the assessment of the importance of preserving certain species rather than others.

Perhaps the most enduring legacy from Clements is his pioneering attempt to synthesize quantitative observations about succession into a unified theory. Even strong detractors of the details of Clements' concepts are still involved in understanding succession as a general phenomenon.

Alternative Organic Theories on Succession

Clements' theory is the most frequently presented organic succession theory, so much so that textbook writers often term Clementsian theory to be classic succession theory. However, all American and British ecologists did not accept these ideas. This rejection was certainly the case for ecologists from continental Europe. Some of these ecologists emphasized a more individual species-oriented mechanistic description of succession. Others shared an interest in the holistic causes of successional change but differed from Clements in significant details. Cowles, who, in 1899, was one of the first American ecologists to study ecological succession, characterized succession as "a variable approaching a variable:" Succession was considered the change of a system perturbed away from – but

moving toward – an equilibrium that is itself changing. Cowles believed that the climax community was never reached. In 1935, Tansley contrasted Clements' climatic climax or monoclimax theory with a polyclimax theory that had the climax vegetation of a region as a mosaic of local vegetation climaxes related to local conditions and disturbance history. In 1913, Cooper studied the forest succession on Isle Royale (Michigan) and found that the mature forest was a mosaic of patches of different ages and not the uniform climax community expected in Clementsian succession. He also believed that succession took multiple pathways and was not a linear progression of changes in seral stages. Similarly, in 1901, Cowles noted, "Succession is not a straightline process. Its stages may be slow or rapid, direct or tortuous and often they are retrogressive" (*Botanical Gazette* **31**, 73–108, 145–182). Recognizing that there is and has been considerable difference in opinion among ecologists who use organic explanations of ecological succession, there is an even stronger contrasting of concepts between Clementsian succession and mechanistic explanations of ecological succession emphasized by some ecologists.

Mechanistic Explanations of Ecological Succession

Some of the debate about the nature of succession concerns the degree to which the vegetation can be arranged into communities that are natural units of biological organization. Introductions of modern biology or ecology texts often have diagrams of biological organization that illustrate an organizational hierarchy from cells to tissues to individuals to populations to communities, etc. Such progressions convey the idea that the community is a unit of organization as demonstrable at a level such as the existence of the liver or the spleen as an organizational unit is at some other level.

The American Gleason (in 1926) and the Russian Ramensky (in 1924) emphasized that vegetation was mostly the consequence of the chance arrival of species at a location and the subsequent interactions among the available species to produce the observed pattern of relative species abundance. These interactions did not produce distinct unit communities. The vegetation was believed to vary continuously with the changes in the underlying environmental conditions. According to this individualistic view of ecological succession, the process of succession was a consequence of species interacting with one another in the context of the environment to produce vegetation dynamics or successional change. Today, most ecologists subscribe to this individualistic view of ecological succession.

Descriptive Mechanistic Models of Succession

When one compares modern descriptive models of succession with the Clementsian model, it is apparent that there is a substantial difference in the spatial scale considered by the two schools. Clements' climax community was considered by him to be a phenomenon that occurred over large areas. This is evident from the fact that Clementsian succession proceeded toward a regional-scale climax. Also, the union of the climax community (vegetation) and associated animals was a biome of which there were thought to be only 13 in nontropical North America. Clements believed that when the community in a given location was too small to have all of the species represented, succession might take different courses. However, he also believed that the succession processes he had described (nudation, ecesis, coaction, etc.) would still operate at these smaller scales.

In 1987, Pickett and colleagues produced a table of the mechanisms and causes of succession that can be compared directly with the earlier writing of Clements (**Table 1**). In this comparison, one sees that many of the processes deemed

Table 1 Clementsian succession processes compared with a mechanistic explanation of the causes of succession[a]

General causes of succession	*Contributing processes of succession*	*Modifying factors*	*Clementsian succession analog processes*
Site availability	Coarse-scale disturbance	Size, severity and timing of disturbance, dispersion of species	Nudation and migration
Differential species availability	Dispersal	Landscape configuration, dispersal agents	Migration
	Propagule pool	Time since last disturbance, land use conditions	Migration and ecesis
	Availability of resources	Soil condition, topography, microclimate, site history	Ecesis
Differential species performance	Ecophysiology	Germination requirements, photosynthesis rates, growth rates, population differentiation	Coaction
	Life history	Photosynthesis allocation pattern, timing of reproduction, mode of reproduction	
	Competition	Hierarchy of competitive interactions, the presence of competitors, identity of competitors, within-community disturbances, predators, herbivores, resource base	
	Herbivory, predation, disease	Climate cycles, predator cycles, plant vigor, plant defenses, community composition, patchiness	
	Environmental stress	Climate cycles, site history, prior occupants	Reaction (in part)
	Allelopathy	Soil chemistry, soil structure, microbes, neighboring species	Not included as process

[a]The first three columns are from a review by Pickett STA, Collins SL, and Armesto JJ (1987) Models, mechanisms and pathways of succession. *Botanical Review* 53: 335–371.

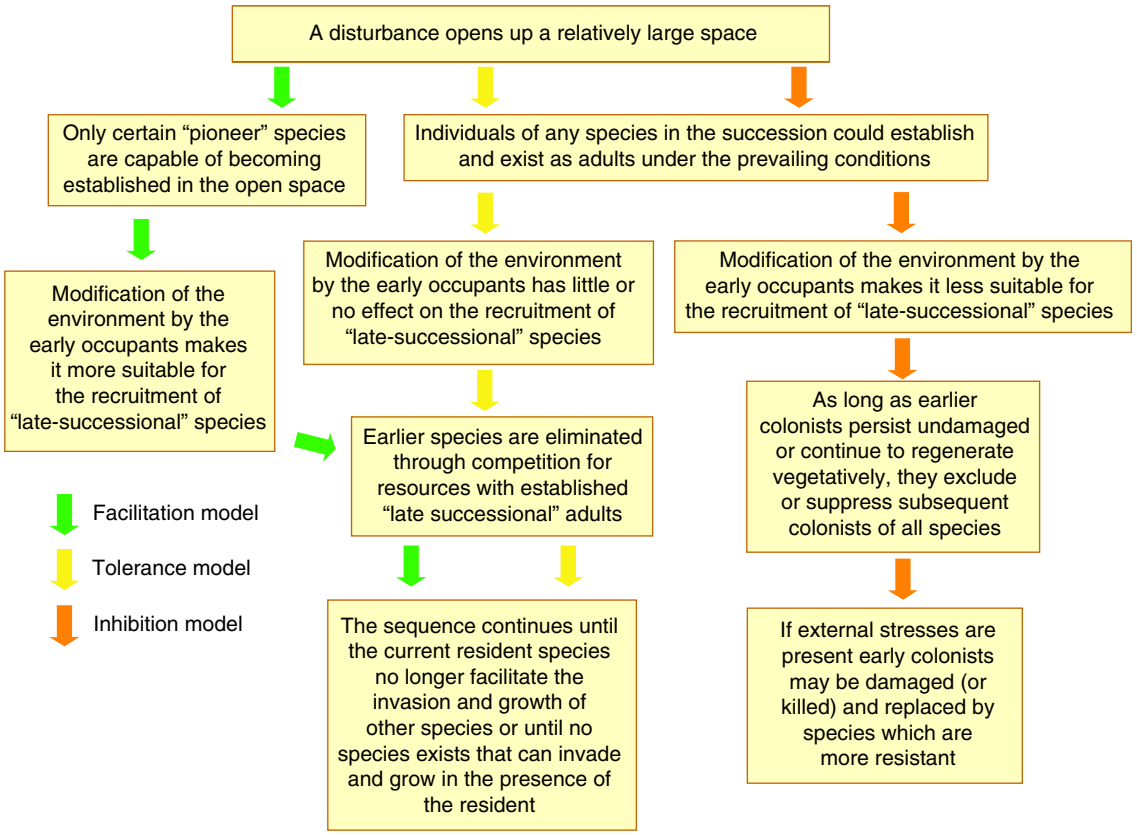


Figure 1 Mechanistic models of ecological succession. Reproduced from Connell JH and Slatyer RO (1977) Mechanism of succession in natural communities and their role in community stability and organization. *American Naturalist* 111: 1119–1144.

important by Clements (notably, nudation, migration, ecesis, coaction, and reaction) are also represented by a mechanistic explanation of succession, albeit with different names and differing emphases. The differences are most pronounced with regard to the importance of the reaction (facilitation). Certainly in many primary successions, the changes produced by one set of species appear necessary for the success of the next. For example, the rapidly growing willows (*Salix* sp.) that stabilize sandbars in rivers seem to be a necessary preamble to the success of subsequent species; the ability of alder (*Alnus* sp.) to symbiotically fix nitrogen increases the fertility of sites uncovered by receding glaciers; and the organic acids produced by blue-green algae, lichens, and mosses speed the breakdown of granite and the development of a thin soil to support grasses, herbs, and even trees in primary successions on granite outcrops.

However, some species appear to block the success of others and hold sites against species that might otherwise succeed them. In some secondary successions, all the species involved in the succession are present as seeds or other propagules from the initiation. In these cases, the familiar successional sequence of grasses and herbs yielding to shrubs and then to trees may reflect a difference in the rate of growth of individuals present from the start. Evolutionarily, it is difficult to explain why a species would evolve to help another take over a site it could otherwise occupy.

Connell and Slatyer (1977) developed a descriptive model of the succession processes based on a mechanistic understanding of succession (**Figure 1**). They used the reaction/facilitation issue to frame three models of succession based on mechanisms of interaction (the facilitation model, tolerance model, and inhibition model). In **Figure 1**, the facilitation model is most like Clementsian succession as typically interpreted. The three models are different and have different implications for land management and particularly land reclamation. If one had the objective of restoring degraded landscapes, then one might speed the restoration by eliminating established species in the case of the inhibition model, but this would be ill advised in the facilitation model (**Figure 1**).

The Landscape as a Mosaic

In addition to an expansion of the facilitation concept associated with Clementsian succession, there has also been an emphasis on understanding the mosaic nature of vegetation. Historically, this emphasis derives from earlier concepts of cyclical succession (exemplified by the cyclical replacement series involving a forest canopy gap) and the polyclimax (the mature forest as a mosaic trees, gaps, and recovering gaps). Theories about the dynamics of landscape mosaics have been developed in forests to a significant extent, probably because, regardless of other sources of spatial heterogeneity, a forest canopy is a mosaic of tree crowns (**Figure 2**).

Starting with a small plot of land in a mature forest dominated by a single large tree, the large tree shades the



Figure 2 The forest canopy as a mosaic as seen in the fall colors of a mixed-wood forest, southern Quebec.

ground and reduces the survival of smaller trees and seedlings below. There may be a few smaller shade-tolerating trees that survive under the large tree, but these are strongly suppressed in their rate of growth. The large tree dominates the resources (light, water, and nutrients) that are available at the site and blocks other trees from growing at the location. When this tree dies, the forest floor (where there had previously been little chance of a young tree's survival) becomes a nursery for small seedlings and saplings. There is adequate light and other resources, and hundreds of small trees survive and begin to grow toward the canopy. As these trees grow, they begin to compete with one another. Some of the trees lose to more vigorously growing competitors. Eventually, one tree manages to win the race to be the local canopy tree and begins to eliminate the others. This represents the closure of the cycle, with a large tree again dominating the site. Over the course of time, this cycle is reinitiated by the death of the new dominant replacement tree. The implications of the cyclical nature of small-scale forest dynamics were clearly elucidated by Cooper in 1913 and by Watt in a classic paper in 1947.

The expected changes in the amount of living material (biomass) over multiple iterations of a gap generation and filling should create a saw-toothed shape curve. This curve drops abruptly with the death of a canopy dominant and then builds biomass as the regenerating trees grow, compete, and occupy the site (**Figure 3**, top). The distances between the teeth in the saw-toothed, small-scale biomass curve are determined by how long a particular tree lives and how much time is required for a new tree to grow to dominate a canopy gap.

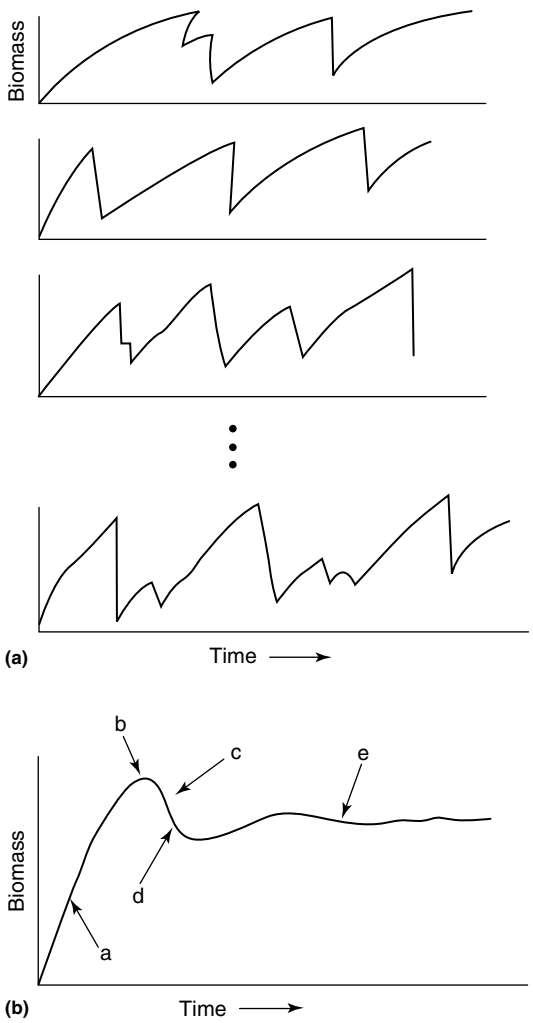


Figure 3 Biomass dynamics for an idealized landscape. The response is from a relatively large, homogeneous area composed of small patches with gap-phase biomass dynamics. (Top) The individual dynamics of the patches that are summed to produce the landscape biomass dynamics. (Bottom) The landscape biomass curve rises (a) as all of the patches are simultaneously covered with growing trees. Next, local decreases in biomass are balanced by the continued growth of large trees at other locations and the curve levels at a maximum value (b). If the trees have relatively similar longevities, there is a period in which several (perhaps the majority) of the patches that comprise the forest mosaic all have deaths of the canopy-dominant trees and the curve decreases (c). Eventually, the local biomass dynamics become desynchronized (d). and the landscape biomass curve varies about some level (e). Reproduced from Shugart HH (1998) *Terrestrial Ecosystem in Charging Environments*. Cambridge, UK: Cambridge University Press.

The larger scale biomass dynamics (**Figure 3**, bottom) is a simple statistical consequence of summing the dynamics of the parts of the landscape mosaic. If there has been a synchronizing event, such as a clear-cutting or other disturbance, one would expect the landscape mosaic biomass curve to increase as all of its parts are simultaneously covered with growing trees (**Figure 3a**). If the trees over the area have relatively similar longevities, there is also a subsequent period when the deaths (and biomass) on plots where a canopy tree

happens to die are balanced by those on plots where large trees are still growing. During this period, the loss balances gain in biomass and the curve levels (**Figure 3b**). If there is sufficient synchronization in the saw-tooth curves of the component plots, this is followed by a period during which many of the pieces that comprise the forest mosaic all have deaths of the canopy-dominant trees (**Figure 3c**) and landscape biomass decreases. Over time, the local biomass dynamics become desynchronized and the biomass curve varies about some level (**Figure 3d**). This mosaic of repairing gaps with all different stages of recovery represented on the landscape can be considered as the mature forest.

The mature forest should have patches with all stages of gap-phase dynamics and the proportions of each should reflect the proportional duration of the different gap replacement stages. Whitmore (1982) believed that this was the expected pattern and process for all forests and asserted that forests of the world are fundamentally similar, despite huge differences in structural complexity and floristic richness, because processes of forest succession and many of the autecological properties of tree species, worked out long ago in the north temperate region, are cosmopolitan. There is a basic similarity of patterns in space and time because the same processes are at work.

Whitmore is referring to mosaic forest canopies as a consequence of gap replacement processes. The occurrence of such patterns has been documented for several different kinds of forests. The presence of shade-intolerant trees occurring in patches in mature undisturbed forest is another observation consistent with the mosaic dynamics view of mature forests. The scale of the mosaics in many natural forests is somewhat larger than one would expect from gap filling of single tree gaps, indicating the importance of the phenomena that cause multiple tree replacements. Also, relatively long records (approximately 40 years in most cases) of forest structure and composition indicate a tendency for the forest composition to fluctuate with species showing periods of relatively weak recruitment of individuals to replace large trees and strong recruitment in other periods.

Quantitative Mechanistic Models of Succession

A consequence of the recognition of the mosaic nature of vegetation and an emphasis on more mechanistic representation of the succession process has been the development of quantitative models that can predict changes in vegetation structure (Shugart, 1998). Several of these models simulate the successional dynamics by accounting for the birth, growth, death, and interactions with the environment for the hundreds of individual plants living on a small plot of land. The predictions of hundreds of these plots are then combined to obtain a prediction of the change in ecological landscapes. Because they simulate the fates of each of the millions of plants involved in a landscape succession, these models are called individual-based models. These models require considerable computation efforts but they can be solved relatively easily as a consequence of the increased computational power of modern computers.

An advantage of individual-based models is that the following implicit simplifying assumptions associated with other modeling approaches (e.g., the Markov process or differential equation-based models) are not necessary: (1) The unique features of individuals are sufficiently unimportant to the degree that individuals are assumed to be identical and (2) the population is perfectly mixed so that there are no local spatial interactions of any important magnitude. Most ecologists are interested in variation in individuals (a basis for the theory of evolution and a frequently measured aspect of plants and animals) and appreciate spatial variation as being quite important. These assumptions seem particularly inappropriate for trees that are sessile and that vary considerably in size over their life span. This may be one of the reasons why tree-based forest models are among the earliest and most widely elaborated of this genre of models.

One group of individual-based models simulates the establishment, diameter growth, and mortality of each tree in an area the size of a gap left by the death of a canopy tree and are called gap models. Gap models can be used as an example of the more general individual-based modeling approach. In most gap models, calculations are on a weekly to annual time step. Early gap models were developed for a size unit (approximately 0.1 ha) approximately that of a forest canopy gap. Gap models feature relatively simple protocols for estimating the model parameters. For many of the more common temperate and boreal forest trees, there is a considerable amount of information on the performance of individual trees (growth rates, establishment requirements, and height/diameter relations) that can be used in estimating the parameters of such models. Gap models have simple, general rules for interactions among individuals (e.g., shading and competition for limiting resources) and equally simple rules for the birth, death, and growth of individual trees (based on the natural history of each species).

Gap models differ in their inclusion of processes that may be important in the dynamics of particular sites being simulated (e.g., hurricane disturbance, flooding, and formation of permafrost) but share a common set of characteristics. These latter characteristics involve an emphasis on the demography and natural history of plant species, relatively general rules for physiological tradeoffs among species, and an emphasis on the understanding of successional processes at the whole plant level. Each individual plant is simulated as an independent entity with respect to the processes of establishment, growth, and mortality. This feature is common to most individual tree-based forest models and provides sufficient information to allow the computation of species and size-specific demographic effects. Gap model structure emphasizes two features important to a dynamic description of vegetation pattern: (1) the responses of the individual plant to the prevailing environmental conditions and (2) how the individual modifies these environmental conditions. The models are hierarchical in that the higher level patterns observed (i.e., population, community, and ecosystem) are the integration of plant responses to the environmental constraints defined at the level of the individuals.

Succession and Biodiversity

Since the initial formulation of a progressive, holistic concept of ecological succession, there has been a tendency to

associate positive attributes to mature communities. Hence, there is an expectation for increasing successional age to be associated with increased biotic diversity. There are certainly ecological systems (notably many forest systems) that demonstrate this pattern, but there are other examples of ecological successions that show the highest levels of species diversity (measured as the number of species or by standard indices of diversity) at intermediate successional ages (e.g., shortgrass prairie in Colorado) or even at initial successional stages (e.g., boreal forests in areas of Canada). Succession on sand dunes on the coast of Queensland, Australia, has the highest species richness in shrubdominated communities at the beginning and end. The pattern of diversity in different successions seems to be an idiosyncratic consequence of the attributes of the participating species and the environmental conditions at a given location.

Biodiversity on Disturbed Landscapes

At the landscape level, the overall biodiversity can be related to the frequency and intensity of disturbance and there is evidence from both theoretical work and observations that an intermediate level of ecological disturbance can produce the most diverse landscapes (Huston and Smith, 1987). This occurs in part because disturbed landscapes have a mixture of species able to successfully occupy the differently aged patches created by the disturbance history of a given landscape. One would generally expect highly heterogeneous landscapes to be more diverse. Disturbances also prevent particularly well-adapted species from occupying the entire landscape.

Grime (1979) developed a triangle based on three primary plant response strategies that can be used to develop rules that can be used to predict the proportions of each strategy (and associated life-forms) expected under a particular environmental regime. Grime recognized two types of external factors limiting the biomass of plants. The first was stress, involving the conditions that restrict plant productivity (shortage of light, H_2O, mineral nutrients, etc.). The second was disturbance, involving partial or total destruction of plant biomass (activities of herbivores, diseases, fire, frosts, etc.). These two external factors can operate independently so that there are four possible combinations of high or low stress and high or low disturbance.

Grime reasoned that the combined action of high stress and high disturbance created a condition from which the vegetation could not regenerate. Low-stress and low-disturbance environments would ultimately favor species that were able to compete effectively against other species (competitor strategy), high-stress and low-disturbance environments should be dominated by plants of species able to tolerate the particular stress (stress-tolerator strategy), whereas low-stress and high-disturbance environments should favor short-lived, fast-growing species (ruderal strategy). The general problem exemplified by Grime's work in the development of these primary plant strategies is one of identifying plant functional types. An essential basis of

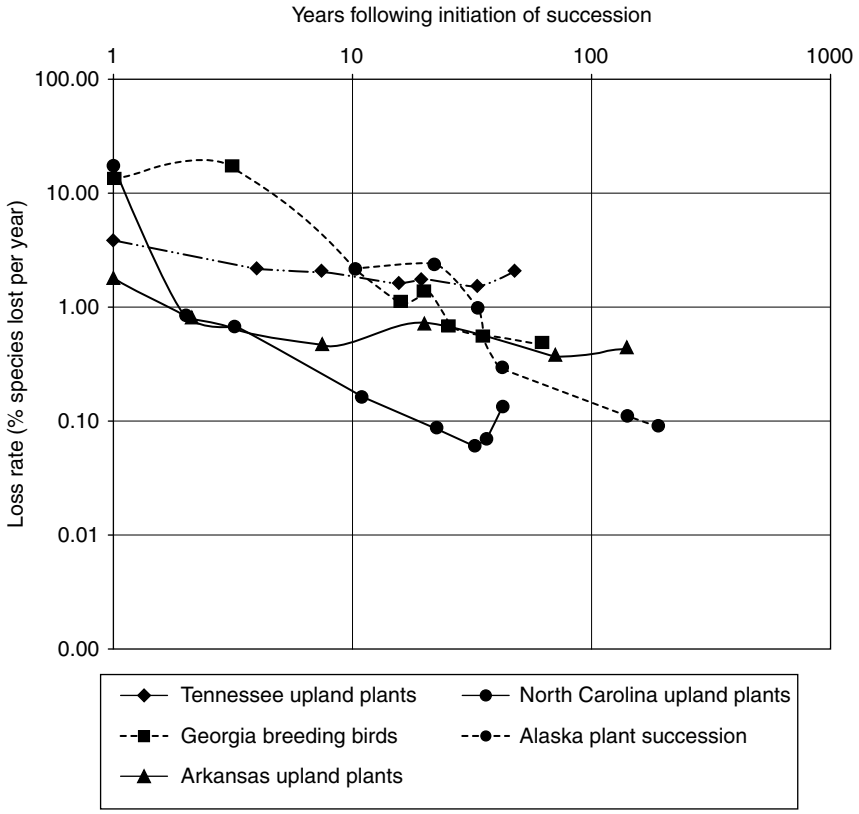


Figure 4 Loss rates of species from different successional sequences. Reproduced from Shugart HH and Hett JM (1973) Succession: similarities of species turnover rates. *Science* 180: 1379–1381.

this as well as other functional classifications of plants is that of tradeoffs – the idea that, due to the underlying rules that derive from the species physiology and natural history, a single species cannot simultaneously be the best as a stress tolerator, a competitor, and a ruderal species.

The Gain and Loss of Species with Succession

One can view the richness of species on a parcel of land with the same disturbance history as a consequence of the gain and loss of species. The gain of species involves factors such as the migration and establishment of species (or the species gaining sufficient abundance or size so that it can be sampled). There has been debate as to whether the gain of species at a site over succession is a consequence of the different species present at the site growing and developing at different rates or is actually due to species establishing themselves in an ordered sequence. The succession is a consequence of different growth rates of an initial inoculum of seeds of all the species involved germinating and growing at different rates and is sometimes called the initial composition model (Egler, 1954). The idea that one set of species is added to the community after a previous set has modified the site and lost the site has been termed the relay floristics model (Egler, 1954). To separate these two models, a particular succession study has to be sampled intensively enough to actually detect all the species at a location, which rarely occurs. Considering a wide range of successions of plant communities, animals associated with succession, and heterotrophic successions (such as the progression of changes in a decaying log), one finds that the appearance of species with abundances sufficient to be counted varies with the particular collection of species and with the changes in the physical environment associated with the succession.

The loss of species from successional communities tends to be somewhat more regular in its pattern of variation. Considering a wide range of communities (both autotrophic and heterotrophic successions), one finds that the rate of species local extinction through succession tends to decrease logarithmically over successional time. Successional sequences are often sampled and reported using a more or less logarithmic sampling regime (e.g., communities in a successional study might be sampled at 1, 2, 5, 10, 17, 35, 60, and 100 years). This reflects the pattern that one tends to sample successional sequences using designs where a proportion of the species in a site of a given age were found in the younger sites. **Figure 4** illustrates this pattern for five different successional sequences in different locations. The initial loss rates of species are on the order of approximately 10% of the species found on a plot disappearing each year. In the later successional stages, this decreases to less than 0.1% species lost per year (**Figure 4**). Because this pattern occurs across a wide range of successional sequences, this does not appear to be a consequence of the later successional species living longer (as is often the case in forest successions).

See also: Disturbance, Mechanisms of. Ecosystem, Concept of. Ecosystem Function, Principles of. Indicator Species. Landscape Diversity

References

Connell JH and Slatyer RO (1977) Mechanisms of succession in natural communities and their role in community stability and organization. *American Naturalist* 111: 1119–1144.

Egler FE (1954) Vegetation science concepts. I. Initial floristic composition – A factor in old-field vegetation development. *Vegetation* 4: 412–417.

Grime JP (1979) *Plant Strategies and Vegetation Processes*. Chichester, UK: Wiley.

Huston MA and Smith TM (1987) Plant succession: Life history and competition. *American Naturalist* 130: 168–198.

McIntosh RP (1985) *The Background of Ecology: Concept and Theory*. Cambridge: Cambridge University Press.

Pickett STA, Collins SL, and Armesto JJ (1987) Models, mechanisms and pathways of succession. *Botanical Review* 53: 335–371.

Shugart HH (1998) *Terrestrial Ecosystems in Changing Environments*. Cambridge, UK: Cambridge University Press.

Shugart HH and Hett JM (1973) Succession: Similarities of species turnover rates. *Science* 180: 1379–1381.

Whitmore TC (1982) On pattern and process in forests. In: Newman EI (ed.) *The Plant Community as a Working Mechanism*, pp. 45–59. Oxford: Blackwell British Ecological Society Special Publication No. 1.

Sustainability and Biodiversity

Jeannine Cavender-Bares, University of Minnesota, St Paul, MN, USA
James Heffernan, Florida International University, Miami, FL, USA, and Duke University, Durham, NC, USA
Elizabeth King, Princeton University, Princeton, NJ, USA, and University of Georgia, Athens, GA, USA
Stephen Polasky, University of Minnesota, St Paul, MN, USA
Patricia Balvanera, Universidad Nacional Autónoma de México, Morelia, México
William C Clark, Harvard University, Cambridge, MA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity The variability among living organisms from all sources, including diversity within species, between species, and of ecosystems.

Ecosystem services Ecosystem functions involving exchanges of nutrients, energy, waste, and materials through the interactions of organisms and the physical environment that contribute to human well-being.

Human well-being Meeting of physical, psychological, and spiritual needs through material security, health, social relations, personal security, freedom, and other factors.

Inclusive wealth The summed total of capital assets or human and Earth system properties that can be used to provide the flow of goods and services that contribute to well-being.

Intergenerational equity Equality and fairness in well-being of people in future different generations relative to current generations.

Natural capital The capacity of the natural world to provide services that contribute to human well-being. These may include providing services including food and fiber production; regulating services, such as water flow and quality, or temperature regulation of the atmosphere; cultural services, such as recreation or meeting spiritual needs; and provision services such as soil formation and nutrient cycling.

Poverty A multidimensional concept that describes the inability to meet human needs; it includes economic, human, political, socio-cultural, and protective aspects of human capabilities, including income, livelihoods, health, education, empowerment, rights, status, dignity, and vulnerability.

Shadow prices The added value (or inclusive economic value) of an additional unit of a capital asset, such as widgets, regulation of the Earth's climate system, or human knowledge.

Social equity The fair, just, and equitable distribution of the stocks or flows of capital assets that contribute to human well-being.

Sustainable development The process of meeting human needs of the present without compromising the ability of future generations to meet their own needs.

Introduction: Biodiversity and Sustainability

Biological systems exhibit extraordinary diversity, whether considering the genetic variation within species, the differences among the more than 8 million recognized species found on the Earth, or the range of environments inhabited and shaped by those organisms. Over the past two centuries, the expansion of human populations, resource demands, and influence on the Earth's landscapes is the driving force behind a dramatic, planet-wide reduction in biodiversity at all of these levels (genes, species, and ecosystems). Estimates of species extinction over that time period range from 100 to 1000 times background levels (Millennium Ecosystem Assessment, 2005), a magnitude of biodiversity loss that has been matched only five times in Earth's history, and the first mass extinction known to be caused by a living species. That this catastrophic loss of diversity is linked with a dramatic increase in both the human population and its overall well-being raises fundamental questions for efforts to provide for human needs and preserve the planet's biodiversity, including whether the human population and its well-being can be maintained in the face of declining biodiversity.

The idea of sustainability – that the fruits of nature, if harvested at moderate rates, may be reaped indefinitely – is ancient wisdom. However, translating that verity into workable policies is difficult and elusive. As economies and human population (**Figure 1(a)**) have grown rapidly during the past 200 years, exploitation of ecosystems for human gain has usually ignored sustainability and often depleted biodiversity. This may change in the next several decades as land transformation and human appropriation of ecosystem services surge toward natural limits and the growth rate of the human population declines toward zero. Still, more than 950 million people face hunger during at least a part of each year (14% of world population) emphasizing the tension between priorities for human welfare versus those for the conservation of species and ecosystems.

Should attempts to improve the material conditions of human life be constrained by attempts to ensure the long-term survival of habitats and species? The connection between sustainability and biodiversity is neither conceptually clear nor practically straightforward, but it is of fundamental significance. Whether slowing the loss of biodiversity is necessary or sufficient for the transition to a sustainable society depends not only on the material relationships between diversity and the provision of ecosystems services to human society, but also on the philosophical basis for the value of biodiversity. The three central rationales that have motivated biodiversity

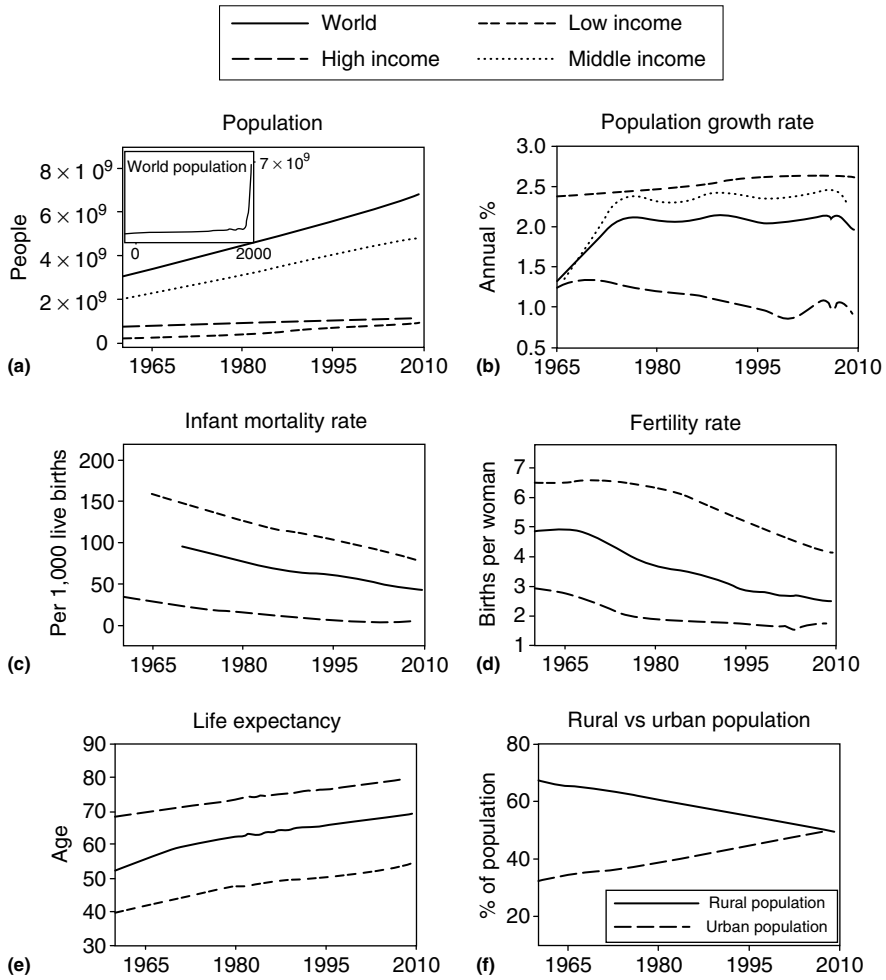


Figure 1 Human demography trends. In recent decades, world population has increased exponentially reaching 7 billion (a) although annual population growth has slowed (b) as human fertility rates have declined (d). Infant mortality rates have declined (c) whereas life expectancy has increased (e) providing an indication of increases in human well-being. Meanwhile, most humans now live in urban environments because the fraction of the population in rural regions has consistently declined whereas that in urban populations has consistently increased over the last half century (f). Data from the World Bank (www.worldbank.org) and the Population Reference Bureau (www.prb.org).

conservation can be summarized as: (1) the ethical perspective that other organisms that evolved over millennia have a right to exist and a claim to planetary resources irrespective of human needs, (2) spiritual and cultural practices that value the existence of other organisms, and (3) empirical linkages between biodiversity and the provision of ecosystem services. Although these rationales are philosophically compatible and even overlapping, they have divergent implications for the prioritization of biodiversity conservation and the meeting of human needs. The ethical perspective (1) suggests that preservation of biodiversity should be considered a priority equal to human well-being. The spiritual/cultural perspective (2) takes as given that human well-being is dependent on the preservation of other species. In either case, a trajectory of development that erodes biodiversity is, by definition, unsustainable. Although both views motivate biodiversity conservation, they do not consider the contributions of biodiversity to the material well-being of society. As such they are likely to fall short in conflicts over land-use that highlight immediate trade-offs between biodiversity conservation and production of goods and services for human consumption.

The ecosystem services justification is fundamentally anthropocentric, in that it makes the value of biodiversity contingent on its relationship to human well-being. Despite the weak philosophical protection it affords to biodiversity conservation, the potential for the ecosystem services perspective to align conservation and human development goals gives it the greatest prospects for widespread political adoption (MEA, 2005). In 1987, the World Commission on Environment and Development (WCED) chaired by Gro Harlem Brundtland declared: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs." Under this directive, the transition to sustainability will require that we understand the complex and reciprocal relationships between biodiversity and human well-being, both now and in the future.

Box 1 Sustainability and the theory of inclusive wealth

In theory, inclusive wealth considers all the contributions to human well-being from all planetary resources. These "capital assets" support, provision, and regulate goods and services now and in the future. The well-being or utility of individuals (U_i), who exist within a social state, x, where the social state characterizes everything that might influence well-being. Social well-being at time t $V(x(t))$ is an aggregate function of the well-being of individuals in society

$$V(x(t)) = V(U_1(x(t))\ \ U_2(x(t))\ \ \ldots\ \ U_n(x(t))$$

Social well-being from time t forward is measured by

$$V(t) = V(x(t)) + V(x(t+1))/(1+\delta) + \cdots$$

where δ is the discount rate by which future values are discounted when compared to present values. The discount rate is an area of much controversy. Perhaps the well-being of future generations should not be discounted at all, or perhaps discounting is reasonable. Leaving aside the controversy, sustainability is said to be achieved when the flow of well-being is nondeclining, $\mathrm{d}V/\mathrm{d}t \geq 0$. Since well-being is subjective and not directly observable sustainability cannot be measured in this way. However, we can attempt to measure assets and their values. The theory becomes more tangible if we consider human well-being in terms of all of the capital assets or system properties, $K(t)$, that can be used to provide the flow of goods and services that contribute to well-being

$$V(t) = V(K(t), M)$$

where $K(t) = (K_1(t), K_2(t), \ldots, K_H(t))$ and M is the evolving political economy or the resource allocation mechanism that involves institutions. Capital assets that contribute to well-being come in many forms, including manufactured capital, natural capital, human capital, social capital, and knowledge capital. Taking the derivative of $V(K(t), M)$ with respect to time generates the following expression:

$$\frac{\mathrm{d}V(K(t), M)}{\mathrm{d}t} = \sum_{h=1}^{n} \frac{\partial V(K(t), M)}{\partial K_h}\ \ \frac{\mathrm{d}K_h(t)}{\mathrm{d}t}$$

The added value of an additional unit of capital asset h is its 'shadow price' P_h and can be defined as

$$P_h(K(t), M) = \frac{\partial V(K(t), M)}{\partial K_h(t)}$$

Inclusive savings or investment in capital asset h is

$$I_h(t) = \frac{\mathrm{d}K_h(t)}{\mathrm{d}t}$$

Inclusive wealth, which is the value of all capital assets, is

$$W(t) = \sum_{h=1}^{H} P_h(K(t), M)K_h(t)$$

and the change in the value of inclusive wealth through time is given by

$$\frac{\mathrm{d}W(t)}{\mathrm{d}t} = \sum_{h=1}^{n} p_h(K(t), M)\frac{K_h(t)}{\mathrm{d}t} = \sum_{h=1}^{n} p_h(K(t), M)I_h$$

Nondeclining social well-being through time is equivalent to nondeclining inclusive wealth, so that we can summarize sustainability from the perspective of the theory of inclusive wealth as being a trajectory through time where the total value of all capital assets that contribute to human well-being is nondeclining:

$$\frac{\mathrm{d}V(K(t), M)}{\mathrm{d}t} = \frac{\mathrm{d}W(t)}{\mathrm{d}t} = \sum_{h=1}^{H} P_h(K(t), M)I_h(t) \geq 0$$

A major challenge for the integration of biodiversity conservation and providing for human well-being through the framework of sustainability is the appropriate valuation of biodiversity. This challenge begins with determining the relationships between biodiversity and ecosystem function, and between ecosystem functions and their value of in terms of well-being, relationships which are incompletely understood. But it also extends to and ultimately resides in questions of scale and equity, both within and among generations. That is, measures necessary to meet human needs locally may compromise biodiversity globally and vice versa, and measures necessary to meet human needs or maintain biodiversity today may compromise one or both of these aims in the future. How can we determine the course that provides greater benefit to human welfare now or for humans in generations to come? A framework that considers inclusive wealth (**Box 1**), or the contributions of all natural and human-derived systems to human well-being, may help guide this search.

At the heart of potential tradeoffs between biodiversity conservation and human well-being are decisions about alternative land-uses, which must balance conservation against

uses that contribute immediately and directly to human welfare, such as agricultural production (Foley *et al.*, 2011). These are entangled with issues of equity by the fact that the world's biodiversity hotspots tend to be found in countries with the highest human population density and the lowest per capita income (**Figure 2**). These same countries, mostly in the tropics, have seen much of their best agricultural land committed to the production of sugar, tea, coffee, chocolate, and other luxury goods for temperate-zone markets for centuries, and increasingly to meet rising global demand for soya and palm oil in biodiversity hotspots such as Indonesia and Amazonia. Ultimately, the possible paths toward sustainability are constrained by the realities of the history and geography of human society and Earth's biodiversity.

The complexities of relationships among biodiversity and human well-being present fundamental challenges to defining the objectives of a sustainable future; thus a transition toward sustainability must be a search rather than a march. Regardless of the specific path, a transition toward sustainability – in which biodiversity and the ecosystem services they provide are sustained and human needs met now and in the future – will require significant social, political, and technological changes during the next few generations. The good news is that this is a time period in which human population is expected to level off within the century due to declining fertility rates (**Figure 1(d)**); hence, it is possible to think of a sustainability transition on the timescale of the demographic transition drawing to a close during the twenty-first century. At the same time, even present levels of human population and consumption have proven devastating to the planet's biota. Understanding how past and future biodiversity loss influence the ecosystem and environmental services that contribute to human welfare is critical to the search for a sustainable future. Awareness of long-term trends and transitions, together with indicators to inform our searches, are important contributions that science can provide in addition to developing means for

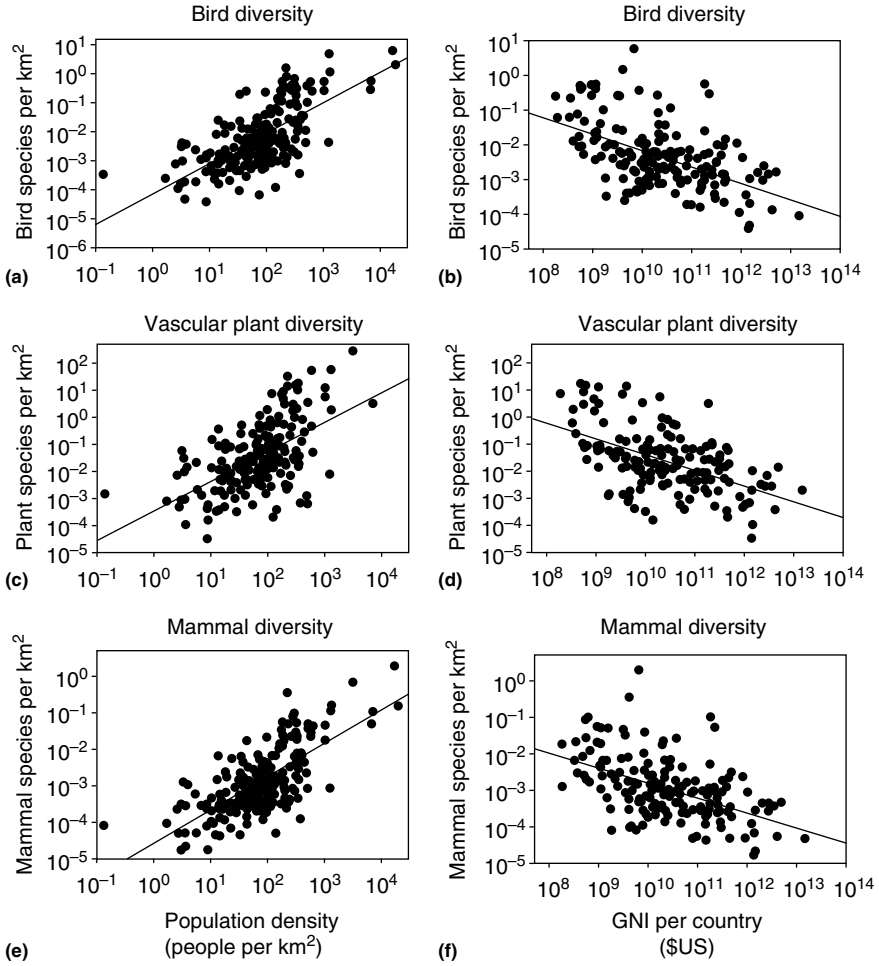


Figure 2 Bird diversity (a, b), vascular plant diversity (c, d), and mammal diversity (e, f) by country in relation to population density (people per km^2) and Gross National Income (GNI; $US), respectively. GNI is the value of all products and services generated within a country in one year (GDP) plus the income lost to or gained from other countries through debt and interest payments. The relationships show that species diversity, one measure of biodiversity, is highest in countries where population density is highest and income is lowest, posing challenges for biodiversity conservation. In a–f, each point represents a single country. Lines are least squares regressions fitted to the data. Species per km^2 are calculated from the number of species per country divided by the total area of the country. Data on bird, mammal and plant diversity are from the World Resources Institution (earthtrends.wri.org); population density and GNI data are from the World Bank.

reconnecting human prosperity to the diverse and essential riches of the natural world.

Sustainability, Sustainable Development, and Carrying Capacity

The definition of a sustainable material economy remains problematic: Which stocks must be preserved? Which flows must be conserved and at what levels? Is there a single numeraire that can indicate sustainability or its opposite? There are no definitive answers to these questions, but the definition of sustainable development put forward by the Brundtland Commission (meeting the needs of the present without compromising those of future generations) contains within it two key concepts: The concept of needs, in particular the essential needs of the world's poor with an emphasis on alleviating poverty; and the idea of limitations imposed by the state of technology and social organization on the environment's ability to meet present and future needs.

In a simple consumer–resource interaction, resource use, or consumption is sustainable when the rate of extraction is equal to or less than the rate at which the resource can replenish itself. In such systems, for a given extraction rate and replenishment rate, there exists a finite limit (or carrying capacity) of the size of the consumer population, which can be sustained. When consumers exceed the carrying capacity, this can result in a self-limiting feedback, because with a decline of the resource base, the consumers' abundance also declines. In this light, ecosystems are essentially suites of resources that support both human and other organisms. And as such, they also have a finite limit (or carrying capacity) to support other populations. However, for humans, self-limiting feedbacks rarely manifest themselves as they would in a simple consumer–resource relationship. People obtain highly diverse and complex portfolios of resources and services from ecosystems, which allows numerous forms of buffering from self-limitation. Some resources can substitute for others in maintaining livelihoods. Thus, with some notable exceptions, the birth and death rates of human populations are rarely in lockstep dependence with a single ecosystem service. Also, our consumer–resource interactions are frequently spatially buffered. Humans live and work in virtually all parts of the planet, exchanging goods in a global economy. Richer societies have often transferred the burdens of their unsustainable practices onto poorer ones by harnessing goods and services produced in poor regions, and avoiding some of the costs of those goods in the form of environmental degradation. Humans have also developed technological innovations that buffer us from the harsh consequences of exceeding carrying capacities by enhancing the rate of resource replenishment, as is the case with fertilizer use on crops. And there is of course temporal buffering, wherein the consequences of unsustainable resource use do occur but are unobserved because the feedbacks take decades to play out.

For nearly all human populations, trends during the past several decades in indicators of well-being, such as declining infant mortality rates (**Figure 1(c)**) and increasing life expectancy (**Figure 1(e)**, income and education, indicate that human welfare is increasing. However, given the many ways by which we buffer and escape direct suffering even when we exceed carrying capacities for ecosystem services, this outcome may not be entirely surprising. Buffering mechanisms, changes in social organizations, technological advances, and the values we place on different indicators of well-being can all obfuscate the finite nature of resource availability. However they do not change the fact that resources are still ultimately finite. They also do little or nothing to direct attention to the ways that overutilization of one resource can impact the availability of other resources and services. The destruction of biodiversity may be deemed a negligible cost to increasing ecosystem services like food production, but when we reduce a resource's abundance, we reduce the carrying capacity for other organisms that rely on the resource. This can lead to cascading losses of biodiversity, which comes full circle to limit resources on which humans depend. This cycle can be seen in the collapse of fisheries, the loss of species in logged tropical forests, and loss of agricultural soil productivity – all cases where exploitation of one resource has led to unexpected losses of other resources and services.

As human numbers and consumption of natural resources increased over the last century, so did concern that the impact of our species on the natural world would have irreversible, ultimately self-destructive effects. Using the idea of carrying capacity as a framework, ecologists have estimated that humans now appropriate slightly less than half the net primary production on land and approximately the same fraction of fresh water (Vitousek *et al.*, 1986). Such estimates indicate that further economic growth or improvements in human well-being – even if they occur – may not continue to increase the size of the material economy on which life depends. Indeed, it seems likely that achieving sustainable economies will require decreasing the burdens of wasted energy, discarded materials, and pollution that are now imposed on the environment. Technology plays a strategic role in this aspiration.

The Inclusive Wealth Framework for Sustainability

There are many perspectives from which to consider sustainability ranging from single indicators of human well-being that mimic Gross Domestic Product (GDP) to composite indicators to even more complex approaches. One type of approach, dubbed the 'dashboard' (Stiglitz *et al.*, 2009) involves gathering and analyzing a series of indicators that are relevant to evaluating environment, social, and economic progress simultaneously. At small spatial scales, particularly in agroecosystems found in many regions of Latin America, once such dashboard approach, the Framework for the Evaluation of Natural Resource Management Systems Incorporating Sustainability Indicators (MESMIS) has demonstrated promise (Speelman *et al.*, 2007). MESMIS considers indicators to evaluate system sustainability in the environmental, economic, and social arenas and provides criteria for sustainable management systems that emphasize (1) productivity, (2) stability, resilience and reliability, (3) adaptability, (4) equity, and (5) self-reliance. The effort has provided critical guidance in decision-making at the community level over the past decade. At larger spatial and sociopolitical scales, such as at national or global levels,

the inclusive wealth economic framework (Arrow *et al.*, 2004; Dasgupta, 2008), provides a theoretical framework for developing composite indicators and charting a path for sustainability (**Box 1**). The inclusive wealth framework follows an economic approach that defines sustainability as nondeclining human well-being over time. In doing so, it takes a strictly human-focused perspective thereby ignoring the value of planetary resources for other species unless they benefit humans. However, it does allow for valuing of biodiversity to humans spiritually, materially, and for their importance in providing ecosystem services. Acknowledging this tension, the inclusive wealth framework offers guidance for human decision-making for achieving sustainability considering resource use, ecosystem services, and human well-being.

Human well-being is difficult to define, given its subjective nature. It is a function of material security, health, social relations, personal security, freedom, and other factors. Essential to human well-being is that physical, psychological, and spiritual needs which are met. In keeping with the United Nations World Commission on Environment and Development (WCED) (1987) definition of sustainable development, a sustainable development path is one in which human well-being is nondeclining. A sustainable economy, therefore, is one which follows a development trajectory in which intergenerational well-being is nondeclining.

Inclusive wealth includes the contributions to wealth from all capital assets and the value of these assets is determined by their contribution to the current and future production or provision of goods and services (**Box 1**). In other words, sustainability is achieved if the change in the total value of all capital assets that contribute to human well-being and investment in those assets is nondeclining.

Theoretically, it is possible to calculate changes in well-being through time by tracking the change in inclusive wealth, which is a function of both the shadow prices of capital assets and investment or depreciation in the capital assets that contribute to well being (**Box 1**). The theory provides an elegant means to focus attention on what needs to be measured to evaluate sustainability. Actually measuring inclusive wealth, however, presents an extreme challenge given that it requires quantifying shadow prices, a process fraught with difficulties and unknowns. Evaluating shadow prices for natural resources, for example, requires evaluating losses in terms of all present and future well-being that would result reductions in these resources. This necessitates understanding exactly how assets contribute to provision of goods and services and how goods and services contribute to human well-being; achieving such estimates for spiritual, aesthetic, or nontangible goods may be particularly problematic. Provisioning of environmental services depends also on systems dynamics and how evolving conditions will impact the stocks and flows of capital assets. In particular, the risks associated with catastrophic changes (i.e., tipping points) are difficult to assess and therefore to manage appropriately. Present formulations of inclusive wealth do not generally include how capital and services are distributed within or among economies. Although equity considerations can be incorporated into this framework, the real challenge is achieving societal agreement on how this is best done. Despite the many challenges, which highlight the nature of the search rather than a march toward sustainability, the inclusive wealth framework provides one approach to considering sustainability that incorporates the stocks and flows through time of the major forms of capital assets that contribute to human well-being.

Table 1 Examples of ecosystem services

Supporting	*Regulating*
Nutrient cycling	Global Climate regulation
Soil formation	Flood regulation
Primary production	Disease regulation
Oxygen production	Water purification
Pollination	Shade/temperature regulation
Provisioning	*Cultural*
Food	Aesthetic
Freshwater	Spiritual
Wood and Fiber	Educational
Fuel	Recreational
Pharmaceuticals	Identity

Source: Adapted from the Millennium Ecosystem Assessment: www.MAweb.org

Capital assets that contribute to human well-being go far beyond resources that are typically valued in monetary terms. All of the planetary and human-created resources that influence human welfare must be considered. These assets can be broken down into major categories, including (1) manufactured capital, which includes all the goods and services produced by humans, (2) human resources and capabilities, such as human health, human knowledge, and governmental, social, and cultural institutions, and (3) natural capital, which includes all the goods and services provisioned, regulated, and supported by Earth systems (e.g., the atmosphere and climate; terrestrial, freshwater, and oceanic ecosystems; and so forth) (**Table 1**). Natural capital assets, for example, include system properties such as soil, hydrology, vegetation, habitat, interactions between species and trophic levels, etc. that provide ecosystem/environmental services that contribute to human well-being. Such services include plant pollination, pest regulation, pollution reduction, renewable resource conservation, soil fertility, food and fiber production, freshwater, regulation of water flow and quality, flood regulation, nutrient regulation, carbon sequestration, disease regulation, recreation, aesthetics, pharmaceuticals, etc. Biodiversity is a critical component of natural capital assets from which these services flow (section Biodiversity as a Form of Natural Capital). Ecosystems and Earth systems, more generally, provide services in often complex and interacting ways that are often poorly understood and monitored. The maintenance of natural capital assets and their ability to provide services are influenced by human use and management. The governance, institutions, cultural values, social relations, incentives, regulations, markets, and so forth that drive how humans use and manage these assets are thus critical to the maintenance of and investment in natural capital.

Biodiversity as a Component of Natural Capital

Biodiversity represents a distinct component of natural capital and is often considered essential to sustainability. Species

extinction is one of the only indications of irreversible environmental loss widely accepted by laypersons; what is its practical significance, both biologically and socially? Is human well-being dependent on the survival of biodiversity, and if so how?

First, it is worth remembering that whereas biodiversity is often used to refer to the richness of species in a given region, it is a much broader term that includes the sum total of the living resources on Earth. According to the 1993 Convention on Biodiversity, biodiversity is the variability among living organisms from all sources, including diversity within species, between species and of ecosystems. The variety of life this encompasses as well as the complex and dynamic relationships among components of biodiversity make it much more than simply the number of species per unit area.

Second, an increasing number of scientific studies demonstrate empirical linkages between biodiversity and ecosystem services that contribute to human well-being (see article from Quijas and Balvanera, *Links between biodiversity and ecosystem services*). Quantitative syntheses of the large amount of experimental studies that have manipulated species richness and measured multiple ecosystem functions have shown that higher biodiversity provisions more services such as increased stability in the face of perturbation, higher productivity and thus higher provision of food fodder or wood, and better regulation of erosion, pests, and pathogens.

Ehrlich and Ehrlich (1981) hypothesized that losses of biodiversity have consequences for ecosystem function much like losses of redundant rivets in airplane wings: Initial losses of species should be accompanied by minimal change in the functioning of ecosystems because some fraction of species are redundant in the processes they perform in nature. However, at some point, loss of species lead to rapid declines in ecological function, much like the loss of one too many rivets can lead to failure of an airplane wing. Of the hundreds of scientific studies that have examined the empirical links between biodiversity and ecosystem functions, such as nutrient cycling and productivity, the vast majority support the rivet-redundancy hypothesis for individual ecosystem functions (Cardinale *et al.*, 2011). In agricultural systems, for example, different crop varieties that can adapt to altered climates or are resistant to different diseases provide redundancy and reduce risks. In the face of climate change or invasive pathogens, if one crop variety fails, another may persist. If one pollinating insect species declines, a similar species may provide the same service. Meta-analyses that have examined multiple ecosystem services simultaneously (Hector and Bagchi, 2007; Isbell *et al.*, 2011) demonstrate, however, that many more species are required to provide multiple ecosystem services than a single service alone. Such multidimensional investigations demonstrate that although each ecosystem function or service may depend on only a fraction of the species in the ecosystem, the number of critical species continues to increase as more functions and services are considered. The end result is that that most species, including rare species, turn out to be critical to the integrity of ecosystems.

Despite the wealth of evidence linking biodiversity to human well-being in experimental and theoretical contexts, at landscape, regional, within country, and among countries spatial scales, biodiversity, and ecosystem services may be at conflict, given competing land-uses. Agriculture, for example, contributes to food provision at the cost of preserving habitat for other species. In contrast, a forest preserve allows for the maintenance of multiple species, which contributes to the maintenance of carbon stocks and the regulation of water-availability and related services, but does not allow for the extraction of timber or other products for human consumption.

It is possible to design landscape management schemes that allow for the simultaneous maintenance of biodiversity and agricultural areas (Porter-Bolland *et al.*, 2011). Diverse agroecosystems can be established within a matrix of conserved forest patches that allow for multiple species to find suitable habitat, while food is produced. Also, it is possible to design a network of sites for conservation that could maximize the conservation of biodiversity and that of ecosystem services. By identifying the key areas for conservation of biodiversity and those for the provision of ecosystem services, sites can be chosen to maximize both. Well-managed natural ecosystems that conserve biodiversity can provide income and jobs through trade, tourism, crafts, and sustainable food production. At the same time, productivity of existing agricultural systems in regions where they fall woefully short of possible production could prevent unnecessary expansion of land conversion to agriculture (Vandermeer and Perfecto, 2007; Foley *et al.*, 2011; Tilman *et al.*, 2011). Polasky and colleagues (2008) modeled an "efficiency frontier" of possible land-use scenarios in which biodiversity and economic output can both be increased far beyond current land management practices. Such options provide a path forward toward sustainability while recognizing that there are challenges, costs, and both winners and losers in the transition.

Biodiversity, Poverty, and Equity

Around the world, the Convention on Biological Diversity has unified nearly every country in a commitment to reduce the loss of biodiversity. Likewise, there is almost universal adoption of poverty reduction as a national goal, as evidenced by the widespread embracement of the Millennium Development Goals (Barrett *et al.*, 2011). These may seem on the surface to be independent endeavors, but they are important, complex interactions between the dynamics of poverty and biodiversity, occurring across spatial and temporal scales.

The well-known latitudinal gradient in biodiversity from the poles to the equator corresponds to a latitudinal gradient in poverty and human population, such that hotspots of biodiversity are found in some of the worlds' poorest and most densely populated nations (**Figure 2**). The interactions between biodiversity and poverty can lead to complex feedbacks that: synergistically benefit both biodiversity and poverty; create tradeoffs between aims to alleviate poverty or conserve biodiversity; or create detrimental feedbacks that eventually lead to worsening human and environmental conditions (Roe *et al.*, 2011). Therefore, the dual aims of biodiversity conservation and poverty reduction cannot be addressed independently of each other, nor can they be strictly pursued within nations without considering global dynamics. Instead, we must recognize and understand the relationships

that exist between biodiversity and poverty, and evaluate them integratively at local, regional, and global scales. Social and economic equity arises as a cross-cutting issue, shaping the trajectories of human–environment interactions and thus affecting sustainability.

At global, continental, and many national levels, trends indicate increases in average standards of human well-being despite declines in biodiversity (MEA, 2005). This would suggest that biodiversity conservation is not relevant to, or not necessary for, human development. However, such aggregated scales blur the fact that national average indicators of well-being can increase whereas the economic welfare of the poorest is stagnant or decreasing. Furthermore, when we focus on regions with extreme poverty, these areas are frequently regions of the greatest rates of biodiversity loss. Numerous factors contribute to this troubling co-occurrence.

Such regions are most commonly located in the tropics and in developing countries, where people often engage in traditional subsistence or small-scale production systems. These segments of society tend to rely heavily on local biodiversity to meet an array of their livelihood needs – food, fiber, fuel, construction materials, medicine, pollination services, cultural services, etc. Also, these societies also tend to have high levels of multiple indicators of poverty, such as little wage employment or cash income, poor access to education, health, and social services, and political marginalization (**Figure 3**). This leads to a spatial correlation of biodiverse areas and poverty.

A major driver of biodiversity loss globally is demand and consumption of natural resources from urban areas and from the developed or rapidly developing world (DeFries *et al.*, 2010). Urban populations have continued to increase in all parts of the globe over the last century (**Figure 1(c)**) with increases in consumption that put pressure on agriculture and the natural resource bases at the expense of biodiversity. The loss of biodiversity in poor rural areas of developing countries occurs as local residents and governments are driven by such demands to exploit their natural capital to increase national economic growth. Natural resource extraction and land conversion – such as timber extraction, mineral extraction, expansion of agriculture or aquaculture, and urban growth – are often undertaken in environmentally irresponsible ways, due to some combination of weak governance, perverse incentives, and economic desperation. Whether out of economic necessity,

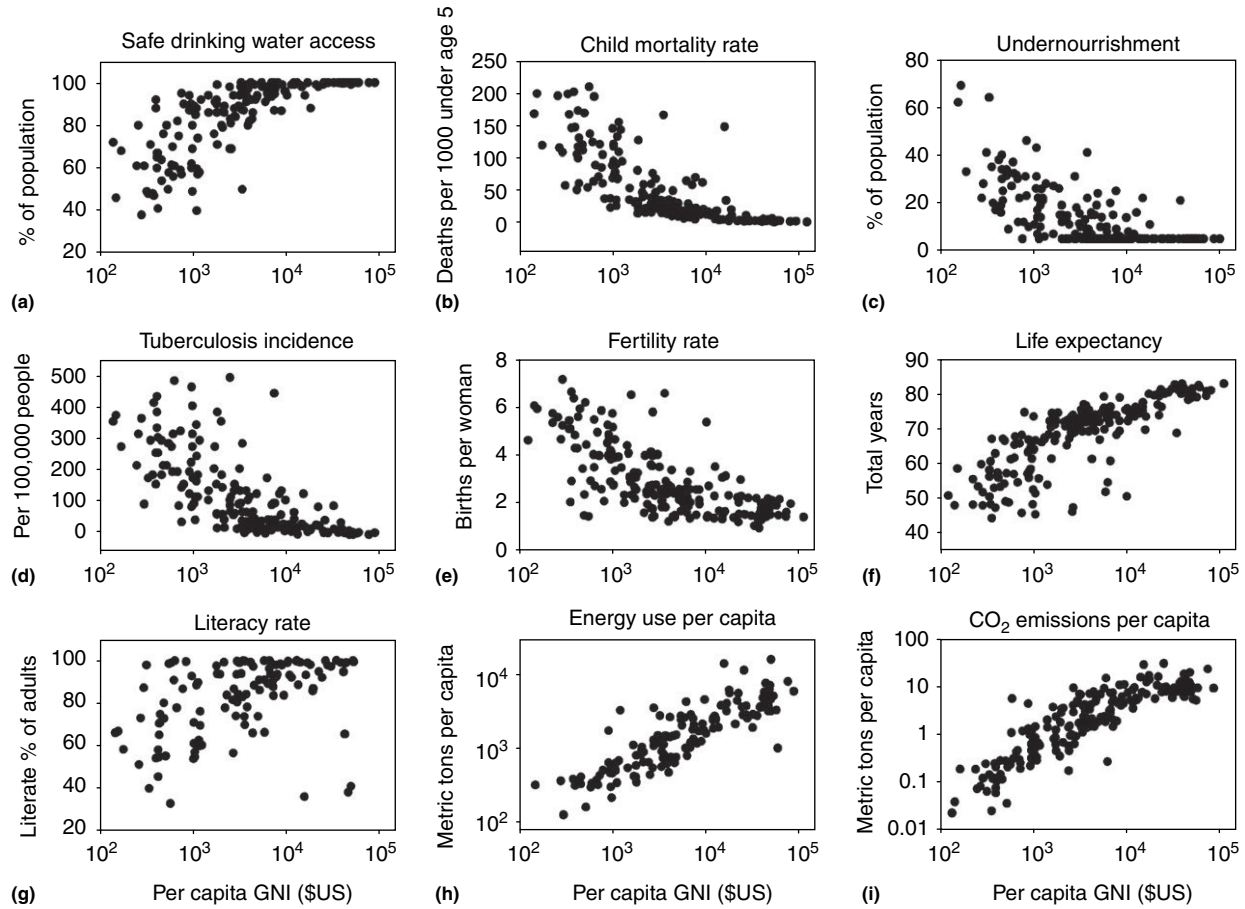


Figure 3 Indicators of human well-being and consumption rates are associated with monetary wealth. Wealthier nations have higher access to safe drinking water (a), lower child mortality rates (b), lower fractions of their populations undernourished (c), and lower rates of tuberculosis (d), an indication of disease risk. Wealthier nations also have lower fertility rates (e) and higher life expectancy (f), and higher literacy rates (g), an indication of demographic transitions. At the same time, consumption rates of energy (per capita) (h) and emissions of CO_2 (i) increase with wealth, indicating that gains in well-being come at a cost for the global environment. Data from the World Bank.

incentivization, greed, or negligence, governments and extractive industries often prioritize short-term economic gains above long-term sustainability. Even where environmental protection policies exist, higher levels of government may have little capacity to enforce them, and the potential economic gains from destructive exploitation negate incentives to attempt enforcement.

Such biodiversity loss can reinforce and increase poverty. When natural capital is converted into financial or manufactured capital, social, and political inequities often contribute to outcomes of increased poverty. Poor people living in proximity to biodiverse areas tend to accrue relatively fewer benefits, and often bear a greater burden of costs, from commercial resource exploitation. Poor people are disempowered politically, economically, and in some cases face cultural prejudice. As a result, they tend to receive the minimum compensation that larger polities can feasibly offer, ranging from forced land appropriation and resettlement, to conversion to an industrial wage labor force with reduced local authority and autonomy.

Local people also lose ecosystem services, such as wild and semiwild food sources, soil erosion control, water purification and recharge, pollinator services, and cultural values associated with biodiverse areas. Since poor rural populations rely more directly on those services, their loss exacerbates the challenges of poverty and decreases options for improving livelihoods. The poor become even poorer as they try to cope with land loss, meager wages, and the disappearance of the safety net of ecosystem services. Under such pressures, poor people must often resort to destructive resource use themselves in order to survive. Biodiversity loss and increased poverty are therefore both linked to social inequity in economic development. Certain development sectors, such as oil extraction in Africa and Latin America, forest conversion for soy production in Amazonia, and timber extraction in southeast Asia, have come to epitomize these dynamics.

International trade, global markets, and aid relations profoundly influence development trajectories in biodiverse regions. Many developed nations place higher political priority on environmental protection within their own territories, with the result that wealthier nations now import huge proportions of natural resources from developing countries. When donor aid is coupled to restructuring developing economies to prioritize exports, inequities at an international scale can disempower developing countries to conserve natural resources. This model of economic development is meant to increase wealth for developing countries. However, due to weak governance capacity, corruption, and pre-existing inequities within those countries, they can instead lead to greater wealth inequality and create poverty traps for the most vulnerable segments of society. The enticement of quick profits for countries in dire financial straits may generate perverse incentives for governments to tolerate or even prioritize practices that threaten biodiversity and perpetuate poverty. Coupled political inequities, between developed and developing countries, as well as within countries, can derail efforts for sustainable development and biodiversity conservation.

These outcomes of global trade are becoming more widely acknowledged. In response, developing countries are demanding more attention – and compensation – be paid for the environmental burdens relegated to the developing world (Roberts and Parks, 2009). This is spurring an important current trend in sustainable development – to dismantle perverse incentives and create local, equitable economic benefits for maintaining biodiversity. Formulating win–win situations is feasible in theory, yet is difficult to achieve in practice. There are now hundreds of thousands of coupled conservation-poverty alleviation projects worldwide, which vary in their relative emphasis on conservation or poverty reduction, and in their strategies for overcoming inequities and perverse incentives (Adams *et al.*, 2004).

There are successful and failed examples in every kind of project linking biodiversity conservation and poverty alleviation: community based natural resource management, integrated conservation and development programs, payments for ecosystem services, sustainable biodiversity utilization, and the recently burgeoning enterprise of cash payments for intact ecosystems for carbon dioxide sequestration (so-called REDD, Reducing Emissions for Deforestation and Forest Degradation, programs). Given the case-specificity in every project, there is consensus on only a few common attributes of projects that yield greater successes. In general, successful programs are carefully designed to account for the specific nature of linkages and feedbacks between poverty and biodiversity in any given case. And equity is recurringly important. Conservation-poverty initiatives require collective action and regulation. If the governance and benefits are not perceived as equitable and legitimate, there tends to be a breakdown of trust, cooperation, and compliance, and project failure ensues (Ostrom, 1990).

Conservation-poverty programs also vary in their approaches to evaluate changes in human well-being and poverty reduction. The inclusive wealth framework is proposed to integrate the values of biodiversity as well as other forms of wealth, into an aggregate assessment of well-being. It is important to note, however, that in the context of assessing poverty reduction, the framework has limitations. As noted, sociopolitical and economic inequality are major drivers of biodiversity loss and persistence of poverty. Inclusive wealth is based on aggregate summations for a given economic unit, and currently does not account for inequities that exist within that unit. Incorporating metrics of economic inequality, such as the Gini coefficient, would add consideration of the status of the poorest as an additional factor in gauging societal well-being. It would also address the concern that empirically, economic inequality is a strong predictor of biodiversity loss (Mikkelson *et al.*, 2007; Holland *et al.*, 2009).

Although there is much to debate and development in the study of biodiversity, poverty, and equity, there are likewise some key common trends that are critical to understanding the challenge of sustaining biodiversity. Biodiversity loss is largely driven by grossly disproportionate consumption in wealthy nations and in urban environments, and the resource exploitation and land conversion in developing countries necessary to support that consumption. There is geographic colocation of poor people and areas of high biodiversity, and those are also the regions where biodiversity loss is most rapid. The dynamics of biodiversity loss and poverty feedback on one another, which creates both situations of entrenched poverty as well as opportunities to potentially engineer win–win situations. Around the world, there are a panoply of efforts to achieve such

win–win outcomes, yet there is weak consensus as to the best ways to do so, or how to evaluate success. Environmental economics, political ecology, sustainability science, and conservation biology are among the host of academic disciplines striving to generate knowledge that can facilitate efforts to simultaneously sustain biodiversity and ecosystem services while alleviating poverty.

Diversity, Resilience, and Sustainability

Evaluating natural capital assets and their likely responses to decreases in resource use and consumption within the inclusive wealth framework are complicated by the fact that Earth system may not respond to change in easily predicted ways. A fundamental question that emerges is how stable are ecosystems and Earth systems in response to perturbation. Ecologists have long been interested in the temporal stability of structure and function of ecosystems, but have defined stability in two distinct ways. The first, termed engineering resilience by Gunderson (2000) defines resistance in terms of the sensitivity of ecosystems to disturbance (i.e., the magnitude of change in response to disturbance of some size) and resilience as the rapidity with which the system recovers to its predisturbance state. In many though by no means all cases (see article from Quijas and Balvanera, *Links between biodiversity and ecosystem services*), more diverse ecological communities are both less sensitive to change, and recover more rapidly to predisturbance values. Ecosystems that are more stable by these metrics might allow for ecosystem services to be obtained with greater regularity and predictability, and thus be of greater value. The shortcoming of studies based on resistance and engineering resilience (sensu Gunderson) is the presumption that ecosystems have a single equilibrium to which they ultimately return. The growing evidence that many ecosystems have more than one equilibrium (i.e., they possess alternative stable states) has more profound implications for relationships between diversity and sustainability.

The notion that ecosystems may shift abruptly between alternative stable states arose first from theoretical models but has since been demonstrated rigorously in natural systems. Some of the most convincing examples of such 'critical transitions' are found in shallow temperate lakes, which can switch rapidly from an oligotrophic clear water system to a eutrophic pea-green soup after nutrient additions pass a threshold (or 'tipping point') (Scheffer *et al.*, 2001). Importantly, reversing such a transition is extremely difficult, because positive feedbacks are reoriented to favor the new equilibrium. Another prominent example of such ecosystem behavior is the transition between woodlands and grasslands in arid and semiarid regions. Feedbacks among fire, herbivory, and the ability of plants to capture runoff all influence such landscapes. The same concepts of feedbacks, thresholds, and critical transitions have been applied to the apparently sudden collapse of human societies such as those in Central America, Easter Island, and elsewhere, generally due to collapse of environmental conditions. The existence of thresholds at the ecosystem scale also raises the notion that there may be planetary boundaries for critical factors, such as ocean acidification levels, atmospheric CO_2 concentrations, temperatures, species extinction levels, and so forth, beyond which Earth systems may switch into alternative states unfavorable to human well-being (Rockström *et al.*, 2010).

The capacity of an ecosystem to absorb external shocks without undergoing such transitions between stable states has been termed 'ecological resilience,' which is the second way ecologists have measured stability. A resilient system, like a ball in a cup, will return to its original equilibrium except in cases of severe disturbance. In a system with low resilience, only a small disturbance may be needed to move past the lip of the cup (the 'basin of attraction') and into another state. Human activities and other external drivers can reduce resilience and increase the probability of state shifts by weakening the interactions that reinforce the original state. Management of ecosystems with alternative stable states requires managing for resilience.

The concept of alternative stable states is important to the integration of biodiversity and sustainability in several respects: First, at the local and global scale, biodiversity may itself be subject to catastrophic declines as the human influences on ecosystems intensify. Rather than a linear decline with resource use (**Figure 4(a)**), biodiversity may exhibit threshold responses where minor stressors have little or no impact on biodiversity, but major stressors lead to severe losses (**Figure 4(b)**). In other words, there is no guarantee that diversity responds in a smooth fashion to environmental pressures – in fact, most evidence points to the contrary. The ongoing mass extinction associated with humans may represent the crossing of such a threshold. At the global scale, loss of species diversity is effectively permanent and irreversible, due to the contingent nature of evolutionary processes and the sheer length of time required for new species to evolve. But even at the local scale, loss of foundation species may effectively prevent reestablishment of original diversity levels after removal of stressors – meaning that biodiversity may exhibit alternative stable states within individual ecosystems (**Figure 4(c)**).

The response of ecosystem processes (and thus ecosystem service provision) to biodiversity can also be nonlinear. Rather than linear decreases (**Figure 4(d)**), relationships between species diversity and ecosystem functions such as productivity exhibit little change initially, followed by increasing large declines at low species number (**Figure 4(e)**; Cardinale *et al.*, 2011). In general, such relationships reflect redundancy among species' functional roles within ecosystems. How often feedbacks between ecosystem function and biodiversity are strong enough to generate alternative stable states (**Figure 4(f)**) is unknown.

Whether considering the response of biodiversity to environmental change, or the response of ecosystem services to biodiversity, nonlinear relationships have important implications for sustainability. Obtaining ecosystem services generally requires human institutions and infrastructure tailored for that purpose. Thus transitions to alternative states often incur significant costs due to the loss of services, even if the potential services from the new state are equal to those of the old. The risk of such losses is magnified by the fact that the location of thresholds and the potential loss of ecosystem services associated with a state transition are often difficult to determine ahead of time (Scheffer *et al.*, 2009). Global measures of sustainability such as inclusive wealth must

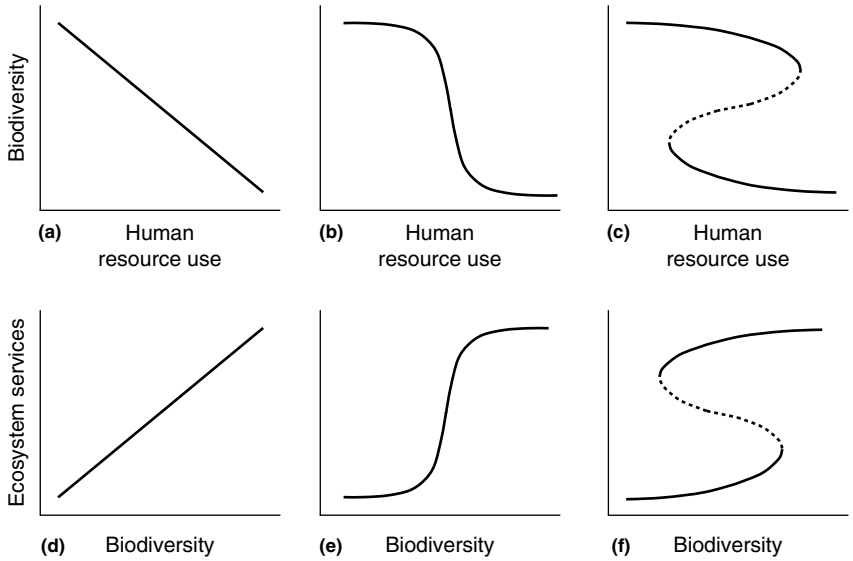


Figure 4 Potential relationships between human resource use and biodiversity (top row), and between biodiversity and ecosystem service provision (bottom row). Panels (a), (b), and (c) show linear, threshold, and hysteretic responses, respectively, of biodiversity to human resource use. Panels (d), (e), and (f) show linear, threshold, and hysteretic responses, respectively, of ecosystem service provision to biodiversity. Although the rate and costs of biodiversity loss are often assumed to be linear (a, d), some evidence suggests that these relationships exhibit sharp breaks at some level of human resource use (b) or biodiversity (e). A third possibility is that these relationships have alternative states (c, f), so that loss of biodiversity or the services that depend on them may not recover even when the underlying stressor (resource use or biodiversity loss) is ameliorated. Note that it would be possible for these relationships to exhibit one form when considered globally, and others when considered at the scale of individual ecosystems.

account for such risks if they are to guide or inform a sustainability transition.

One important implication of the inclusive wealth framework is that of intergenerational equity – the needs of the present balanced against those of future generations. Threshold relationships between biodiversity and ecosystem services (and among many other components of ecological systems) exacerbate the tension between maximizing current provision of ecosystem services and maintaining the capacity of ecosystem to provide services indefinitely, particularly when the potential for alternative states is considered. Available evidence suggests that maintaining diverse ecological communities is an important component of managing for resilient ecosystems and ensuring the future provision of ecosystem services.

Trends

Biodiversity Trends – The Millennium Ecosystem Assessment (2005) reports that extinction rates are 100 to almost 1000 times the background level of extinction based on the fossil record. Two thirds of the world's ecosystem services that depend on biodiversity are declining. These alarming trends led scientists to consider the current epoch of humanity the 'Sixth Extinction' (Leakey and Lewin, 1995). The largest driver of biodiversity loss is habitat change. On land, such changes occur through land use changes. Both globally and particularly in heavily indebted poor countries, conversion of land to agriculture (**Figure 5(e)**) has been on the rise in the last half century. Such trends occurred much earlier in many wealthier regions of the globe. Land conversions and land degradation also degrade or destroy ecosystems and the services they provide to humans. Increasing need for fuelwood and land for agriculture, together with industrial logging, have resulted in global losses of forest at a rate of 16 million hectares per year during the 1990s. World deforestation, mainly the conversion of tropical forests to agricultural land, has decreased over the past decade (2000–2010) to approximately 13 million hectares of forests per year but continues at an alarmingly high rate in many countries (Global Forest Resources Assessment, 2010).

In coastal and ocean ecosystems, habitat change is driven by loss of coral reefs through ocean acidification, pollution, climate change, species invasions, overexploitation, and damage to sea floors due to trawling. In coastal ecosystems, Worm *et al.* (2006) report dramatic declines in populations of ecologically and economically important species: In fishery taxa that have maintained viable population sizes over the last 1000 years, 40% have collapsed (their populations have dropped below 10% of their maximum population size) since 1800. More than half of the world's coral reefs face changes in species composition, obliteration, and other major ecosystem effects. In open ocean ecosystems, the situation is even more dire: 80% of fish and invertebrate taxa in our global fisheries have collapsed in the last 50 years. These collapses have resulted in major losses of ecosystem services (fisheries, nursery habitat, and filter function) and with concomitant increases in risks (such as beach closures, harmful algal blooms, fish kills, oxygen depletion, coastal flooding, and species invasions).

Covering less than 1% of the Earth's surface, freshwater ecosystems have lost the largest proportion of species and habitat when compared with other ecosystems on land or with the oceans. Physical modification or water withdrawal from

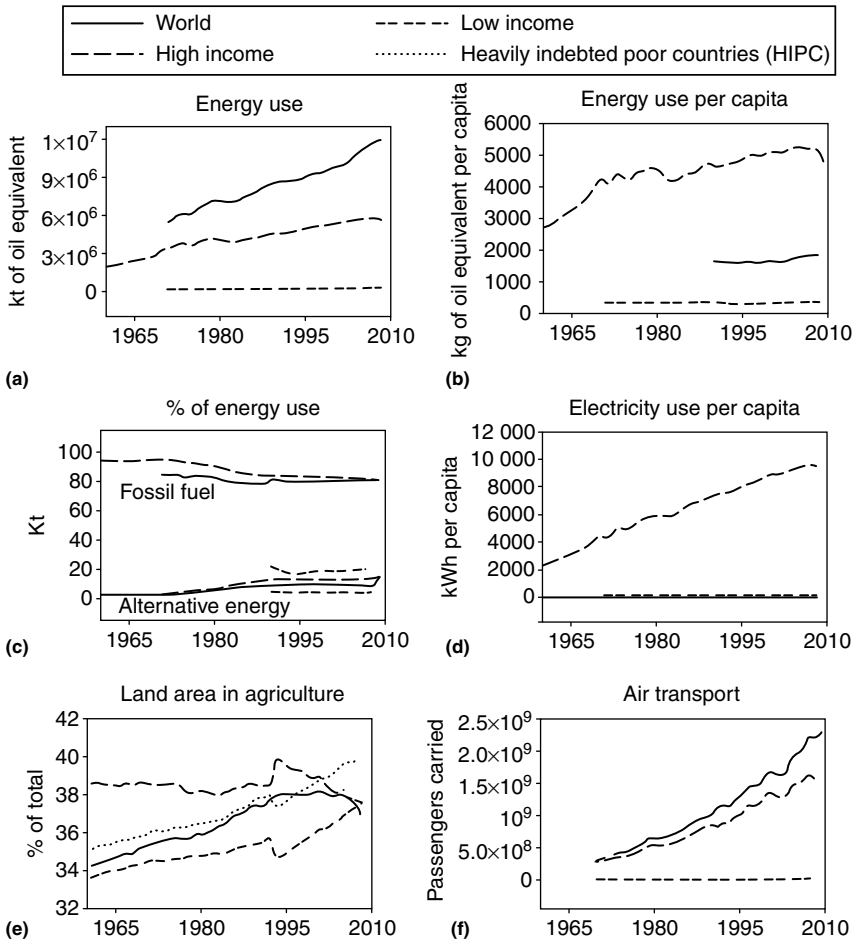


Figure 5 Global trends in total and per capita energy use and land area for agriculture. (a) Total energy use in kt of oil equivalents for the Word, high income nations and low income nations since 1960. High income and World totals continue to rise, whereas low income totals remain very low. (b) Per capita energy consumption has increased fairly steadily in high income nations since 1960 with a slight down-turn in recent years. (c) Fossil fuels still comprise the vast majority of energy consumption in the word, particularly in high income nations. In low income nations, fossil fuel energy use as a percentage of total energy use is much lower. Alternative energy sources, such as wind, solar, and nuclear remain very low percentages of total energy use throughout the world. (d) Electricity use per capita continues to increase steadily in high income nations but is stable in low income nations and for the World. (e) Land are for agriculture as a percentage of total land area has remained fairly stable in high income nations; it continues to increase in low income nations, and it is steadily increasing in heavily indebted poor countries, surpassing that of high income nations. (f) Air transport in terms of passengers carried an indicator of energy use and consumption that is afforded only by the wealthy, has grown steadily since 1970. This trend is driven largely by high income nations. Low income nations show very low and near constant levels. Data from the World Bank.

rivers, continuing overfishing, dam building and river development, and contamination have driven species losses and will continue to place greater threats on freshwater ecosystems. Many estuaries and bays have also deteriorated because of activities associated with land development and fishing pressure, undermining ecosystem services. All the habitat transformations have profound impacts on human well-being. Biodiversity losses affect the livelihood of local communities that depend on diverse ecosystems for food, tourism, and ecosystem stability.

Future projections of species extinction rates based on habitat change current population status of threatened or endangered species indicate that extinction rates will continue to rise to ten times their current rates (MEA, 2005), although the methods for generating such predictions are difficult (He and Hubbell, 2011). Measuring and monitoring for biodiversity loss presents a major challenge given that possibly >85% of species on land and >90% in the ocean have still not been identified or described (Mora *et al.*, 2011).

Trends in health, consumption, and communication – At the same time that biodiversity indicators show increasing cause for alarm in our ability to manage for sustainability, indicators of human well-being continue to show positive trends (**Figure 1(c)** and **1(e)**). But these benefits to human well-being have come at the cost of consumption of non-renewable resources (**Figure 5(a)–(d)**), homogenization of the biota, emission of climate altering greenhouse gases (**Figure 6(a)** and **6(b)**), and pollutants to human health (**Figure 6(c)** and **6(d)**). Inequities in the distributions of costs and benefits are apparent, as discussed in section Diversity,

Resilience, and Sustainability. Indicators such as disease incidence, literacy rate, safe drinking water access, and life expectancy indicate that wealthier nations have higher material well-being in terms of health, education, and access to basic resources (**Figure 3**). Wealthy nations have high per capita energy consumption rates and CO_2 emissions rates (**Figure 3(h)** and **(i)**). As such, wealthy nations appear to be reaping the benefits of consumption but the costs of this consumption is paid for globally. CO_2 emissions per capita have begun to level off in wealthy nations; globally, however, yearly emissions continue to increase as a result of population growth and consumption in fast developing countries. In the past 50 years, total annual nitrous oxide pollution has remained relative constant and particulate matter has markedly declined (**Figure 6(c)** and **(d)**). Electricity use per capita and total energy use of high income countries continues to rise. Air travel in wealthy nations and globally has risen sharply. At the same time, humans are increasingly connected and have the capacity to share ideas, cultural values, and transfer technology rapidly. The exponentially increasing trends in internet use and cell-phone use (**Figure 7(a)** and **(b)**) attest to the increasing global connectedness. The same global interconnectedness that creates pressures on biodiversity may also facilitate mechanisms such as REDD that provide incentives to address climate change while sustaining biodiversity.

Meeting human needs while preserving the Earth's life support systems for future generations will require a worldwide acceleration of today's halting progress in a transition toward sustainability. Whether humanity can take advantage of human capital, knowledge, and technological advances to develop processes and governing mechanisms to achieve sustainability will determine much about the fate of the Earth's biodiversity.

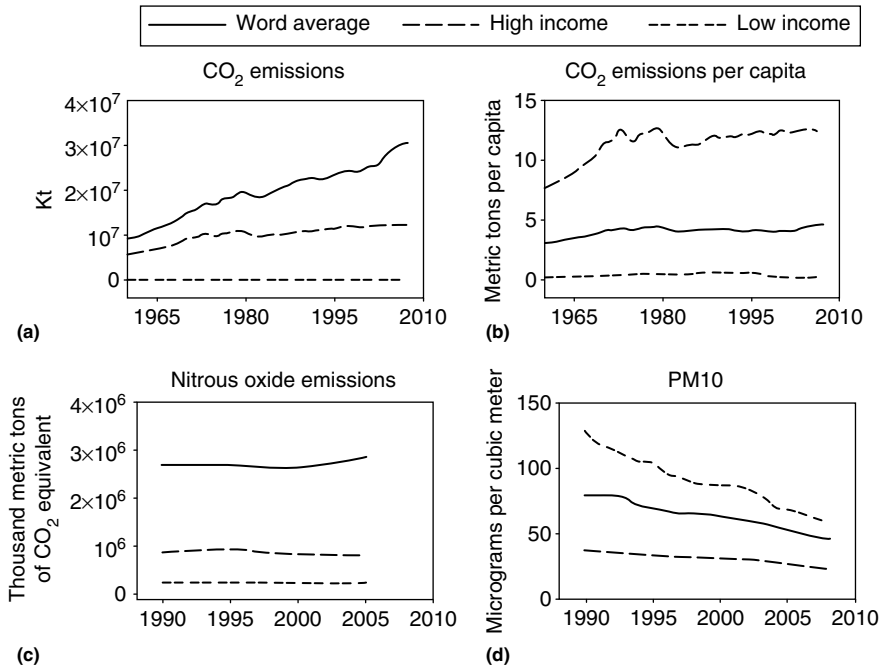


Figure 6 Pollution levels show contrasting trends. Although global CO_2 emissions have risen sharply (a), people in wealthy nations are largely responsible (a, c). Nitrous oxide emissions have stabilized (c) whereas particulate matter has decreased (d). Data from the World Bank.

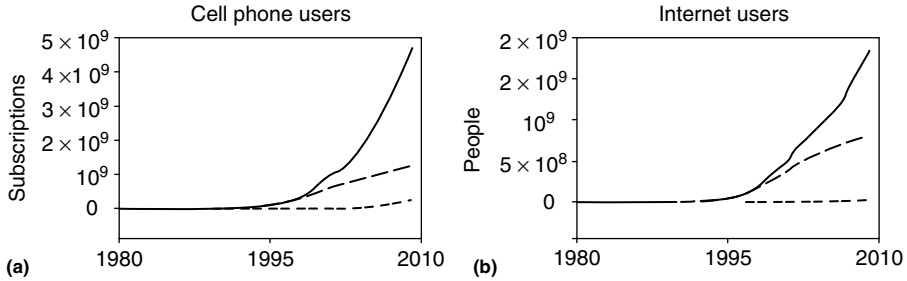


Figure 7 Exponential increase in global cell phone subscriptions (a) and internet users (b) in recent decades provides an indication of the change in communication and access to information that human societies have undergone. Data from the World Bank.

Acknowledgements

This work developed from a multi-institution Distributed Graduate Seminar on Sustainability Science sponsored by the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant #EF-0553768), the University of California, Santa Barbara, and the State of California and by the Long-Term Ecological Research network, the University of Minnesota Institute on Environment, and the Harvard Kennedy School of Government. We are grateful to participants of the distributed graduate seminar for valuable discussions that contributed to this chapter.

References

Adams WM, Aveling R, Brockington D, *et al.* (2004) Biodiversity conservation and the eradication of poverty. *Science* 306: 1146.

Arrow KJ, Dasgupta L, Goulder G, *et al.* (2004) Are we consuming too much? *Journal of Economic Perspectives* 18(3): 147–172.

Barrett CB, Travis AJ, and Dasgupta P (2011) On biodiversity conservation and poverty traps. *PNAS* 108(34): 13907–13912.

Cardinale BJ, Matulich KL, Hooper DU, *et al.* (2011) The functional role of producer diversity in ecosystems. *American Journal of Botany* 98(3): 572–592.

Dasgupta P (2008) Nature in economics. *Environmental and Resource Economics* 39: 1–7.

DeFries RS, Rudel T, Uriarte M, and Hansen M (2010) Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience* 3: 178–181.

Ehrlich PR and Ehrlich AH (1981) *Extinction: The Causes and Consequences of the Disappearance of Species*. New York: Random House.

Foley JAN, Ramankutty KA, Brauman ES, *et al.* (2011) Solutions for a cultivated planet. *Nature* 478: 337–342.

Food and Agricultural Organization of the United Nations (FAO) (2010) *Global Forest Resources Assessment, 2010 (FRA 2010).* http://www.fao.org/forestry/fra/fra2010/en/http://www.fao.org/forestry/fra/fra2010/en/

Gunderson LH (2000) Ecological Resilience – its theory and application. *Annual Review of Ecology and Systematics* 31: 425–439.

He F and Hubbell SP (2011) Species–area relationships always overestimate extinction rates from habitat loss. *Nature* 473: 368–371.

Hector A and Bagchi R (2007) Biodiversity and ecosystem multifunctionality. *Nature* 448: 188–190.

Holland TG, Peterson GD, and Gonzalez A (2009) A cross-national analysis of how economic inequality predicts biodiversity loss. *Conservation Biology* 23: 1304–1313.

Isbell F, Calcagno V, Hector A, *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature* 477: 199–202.

Leakey and Lewin (1995) *The Sixth Extinction: Biodiversity and Its Survival.* New York: Anchor Books.

Mikkelson GM, Gonzalez A, and Peterson GD (2007) Economic inequality predicts biodiversity loss. *PLoS ONE* 2: e444.

Millennium Ecosystem Assessment (2005) Ecosystems and Human Well-Being: Biodiversity Synthesis. Page available at http://www.MAweb.org. World Resources Institute, Washington, DC.

Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B (2011) How many species are there on earth and in the ocean? *PLoS Biology* 9(8): e1001127.

Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge, UK: Cambridge University Press.

Polasky S, Nelson E, Camm J, *et al.* (2008) Where to put things? Spatial land management to sustain biodiversity and economic returns. *Biological Conservation* 141: 1505–1524.

Porter-Bolland L, Ellis EA, Guariguata MR, Ruiz-Mallén I, Negrete-Yankelevich S, and Reyes-García V (2011) Community managed forests and forest protected areas: An assessment of their conservation effectiveness across the tropics. *Forest Ecology and Management*,http://dx.doi.org/10.1016/j.foreco.2011.05.034.

Roberts JT and Parks BC (2009) Ecologically unequal exchange, ecological debt, and climate justice. *International Journal of Comparative Sociology* 50: 385–409.

Rockström J, Steffen W, Noone K, *et al.* (2009) A safe operating space for humanity. *Nature* 461: 472–475.

Roe D, Thomas D, Smith J, Walpole M, and Elliott J (2011) Biodiversity and poverty: Ten frequently asked questions – Ten policy implications. *IIED Gatekeeper 150.* London: International Institute for Environment and Development.

Scheffer M, Bascompte J, Brock WA, *et al.* (2009) Early-warning signals for critical transitions. *Nature* 461(3): 53–59.

Scheffer M, Carpenter S, Foley JA, Folke C, and Walker B (2001) Catastrophic shifts in ecosystems. *Nature* 413(6): 591–596.

Speelman EN, López-Ridaura S, Aliana Colomer N, Astier M, and Masera OR (2007) Ten years of sustainability evaluation using the MESMIS framework: Lessons learned from its application in Latin American case studies. *International Journal of Sustainable Development & World Ecology* 14: 345–361.

Stiglitz JE, Sen A., and Fitoussi J-P (2009) *Report by the Commission on the Measurement of Economic Performance and Social Progress.* www.stiglitz-sen-fitoussi.fr

Tilman D, Balzer C, Hill J, and Befort BL (2011) Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*: www.pnas.org/cgi/doi/10.1073/pnas.1116437108.

United Nations World Commission on Environment and Development (WCED) (1987) *Report of the World Commission on Environment and Development, 1987. Our Common Future*, 43 pp. Oxford: Oxford University Press.

Vandermeer J and Perfecto I (2007) The agricultural matrix and a future paradigm for conservation. *Conservation Biology* 21: 274–277.

Vitousek PM, Ehrlich PR, Ehrlich AH, and Matson PA (1986) Human appropriation of the products of photosynthesis. *BioScience* 36: 368–373.

Worm B, Barbier EB, Beaumont N, *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.

Systematics, Overview

Quentin D Wheeler, Arizona State University, Tempe, AZ, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Character An attribute that is constantly distributed among all members of a species or a clade.
Homology Two structures hypothesized to share a common ancestry.
Homoplasy Similar attributes not due to common ancestry but rather to convergent evolution, reversals, or mistaken interpretation.
Macroevolution Evolution above the species level.
Microevolution Evolution within species.
Monophyletic group An ancestral species and all of its descendant species; evidenced by synapomorphy.
Synapomorphy Shared, derived similarity.
Three-taxon statement Two species (or taxa) are more closely related to one another than either are to a third; the most basic cladistic hypothesis.

Introduction

The terms taxonomy and systematics are frequently interchanged in common usage with little loss of meaning and they will be so treated here. Two accounts of the relationships between these terms exist, however. According to one, championed by neo-Darwinists Simpson (1961) and Mayr (1969), systematics encompasses broader studies of evolutionary relationships, whereas taxonomy narrowly concerns itself with classification and naming practices. Roy A. Crowson (1970), however, treated taxonomy as the broader field with the ultimate aim of phylogenetic classifications. Systematics, in this interpretation, is the narrower endeavor concerned with working out phylogenetic relationships in order to inform classifications (Wheeler, 2008, Introduction). Precisely defined, *systematics* (or systematic biology) is that subset of taxonomy concerned with delimitation of species and analysis of characters and cladistic relationships. *Taxonomy* is the broader field with the ultimate goal of a predictive phylogenetic classification and informative Linnaean names.

What is Systematics?

Systematic biology is a historical and comparative science providing a counterbalance to general or functional biology (Nelson and Platnick, 1981). As such, systematics is nonexperimental, yet is based on rigorously testable hypotheses about homology, species status, and monophyly. Systematics is at once the most fundamental and inclusive of the biodiversity sciences: most fundamental because it discovers, describes, and makes identifiable species and provides the names by which all biological information is communicated; most inclusive because it ultimately gathers together, interprets in an historical evolutionary framework, and synthesizes all relevant evidence whether molecular, anatomical, ontogenetic, or fossil. Taxonomy is essential to the environmental sciences, but is itself more fundamentally an evolutionary science. Its theories are deeply rooted in concepts related to speciation, character transformation, and macro-evolution. Taxonomy uses insights from population genetics, geography, and ecology to understand species and their spatial distributions, and it provides a general conceptual framework within which biological observations may be interpreted in an evolutionary framework.

Taxonomists focus on patterns of character distributions that distinguish species on the one hand and unite them into monophyletic groups on the other. Population biologists, in contradistinction, focus on the processes of speciation. These fields necessarily overlap at the species level but are fundamentally different in their theories, methods, and aims, the distinction aligning with that of pattern versus process (Eldredge and Cracraft, 1980). The boundary between population genetics and taxonomy was blurred in the early twentieth century by those who rejected studies of phylogeny as speculative and embraced exclusively experimental work below the species level, so-called ''population thinking'' (e.g., Mayr, 1942). This bias toward experimental research did much to increase the stature of genetics and unnecessarily and indefensibly decrease that of taxonomy. Interest in predictive classifications was renewed when taxonomists became early adopters of computer technology. Phenetics, at the time also called numerical taxonomy, posited that degree of similarity corresponds with community of descent. Pheneticists claimed to be more objective and endeavored to use all available data. It turned out that the results were repeatable, but only with the same algorithm and dataset. Further, the choice of algorithm was in itself subjective and it was ultimately concluded that such methods did not reliably retrieve evolutionary relationships. The next advance in theoretical systematics, arguably the most important advance since Darwin, was Hennig's (1966) phylogenetic systematics. At last, taxonomic theories were explicitly testable.

Hennig reasoned that an infinite number of special purpose, artificial classifications were possible, each meeting the need of a small subgroup of biologists well but not generalizable to all situations. He recognized that life was so diverse that the only general classification system that is reasonably useful to all biologists must be based on the one thing that all life shares in common: evolutionary history. Phylogenetic classifications are optimally informative (Farris, 1979), facilitating both the retrieval of what is known about species and predictions about what is not. Formal Linnaean names associated with phylogenetic classifications have proven highly efficient communication tools and are integral to modern databases from GenBank to the Encyclopedia of Life.

Taxonomy is commonly misunderstood, largely because of its historical basis. Evolutionary history, not unlike the sequence of events in human history, is not replicable. The myriad circumstances and preconditions of ancestral species cannot be enumerated, much less recreated. Experiments are designed to control variables and to make predictions about one possible outcome among a universe of possible outcomes. Because history is a single sequence of events, no such universe exists. The tasks of systematics include the interpretation of the transformation of characters, recognition and description of species, and reconstruction of cladistic relationships — that is, the relative recency of common ancestry among species. The most unique aspects of taxonomic work include the following.

Historical perspective: Taxonomy is not infrequently viewed with suspicion because of its nonconformity to standard practices of experimental biology. There is a good reason for this: Taxonomy is not experimental. Phylogenetic history is unrepeatable; even if the clock could be turned back, it is unlikely that the course of events would be the same given the many contingencies of evolution (Gould, 1989). Fortuitously, and unlike human history, evidence of phylogeny is preserved in the genome of organisms and the complex structures coded by genes. Taxonomy is no less scientific than experimental biology. However, it does require a different epistemology for its justification. Details of hypothesis testing in taxonomy will be discussed later in this section.

Comparative method: Taxonomy is comparative at virtually every level. Whether assessing variation within species, hypothesizing homology among characters, or choosing among several equally parsimonious cladograms, taxonomists are often comparing large numbers of specimens or species. No other subdiscipline of biology is comparative to the same degree.

Worldwide collections: This comparative aspect of taxonomy means that taxonomists must have access to specimens for every species of a group being studied, regardless of where on Earth these species might live. To make such work practical, taxonomists have assembled vast collections in major museums and herbaria. Although many biodiversity scientists need to study the same organisms or phenomena at the same study sites for extended periods of time, the same constraints do not exist for taxonomy. Just the opposite is true: Taxonomists require samples from the most disparate reaches of the geographic ranges of species and clades. To unambiguously sort out species, the taxonomist must see as much infra-specific variation as possible as well as study variation within closely related species. To complete as comprehensive a study of an entire group as possible, the taxonomist must examine every species possible within that group regardless of where or when they live or lived. This is a very different set of requirements than that facing the experimental biologist. Natural history collections are costly to assemble, laborious to maintain, and may only be accumulated with the combined efforts of many over very long periods of time. The great natural history collections of the world have been built up over a period of centuries and even the best of them contain but a fraction of all the world's species making international collaboration among scientists and curators essential.

Taxon inventories: Building taxonomic research collections demands that scientists with taxon-specific knowledge be encouraged and permitted to collect in as many locations within the range of their group as is possible. Long series of specimens are needed to address infra-specific variation. Two centuries of taxonomic practice has taught us that making species decisions in isolation, i.e., describing species based on material artificially restricted by geopolitical boundaries, leads to redundant species descriptions, synonymy, and taxonomic confusion.

Taxon scholars: Taxonomists are typically specialists and may spend their entire careers working on one or a few higher taxa. This is not surprising when it is recognized that a single genus or family may include hundreds or thousands of species. Also, good taxonomy requires a level of scholarship unmatched in experimental biology. A competent taxonomist must master all published literature in the taxon since 1758 (1753 for plants), learn all relevant anatomical terminology, understand the gathering, limitations and analysis of molecular, phylogenetic, and fossil data, know modern theories of species, characters, and phylogeny, learn to comply with a complex code of international rules governing botanical or zoological nomenclature, know their way equally well around field sites, museums, libraries, and laboratories, and master a range of technological tools. Taxonomist's scholarship is not bounded by geography, ecosystem, or even time: one must compare, interpret, and understand hundreds of characters across scores or hundreds of species at a time, constrained only by theories of monophyly.

Philosophy: In the philosophy of science, the demarcation principle draws the line between science and nonscience. To be scientific, claims of knowledge must be stated in such a way that further observation might show them to be false (Popper, 1968). Questions not open to critical empirical testing and possible rejection, such as "Is there a God?" are simply outside the limits of science. Stripped of the familiar trappings of experimental biology, how can the historical patterns constructed by taxonomists be scientific? Phylogenetic systematics finds epistemic justification in its hypothetico-deductive method (e.g., Gaffney, 1979). Knowledge does not exist with absolute certainty. Rather, "confidence" (in the common rather than statistical sense) increases in ideas in proportion to the number of times that they are subjected to, and survive, critical testing. The most rigorous scientific claims are the most vulnerable ones. For example, all-inclusive statements such as "all swans are white" are highly testable. Although observation of a million white swans cannot prove this claim about the world, seeing a single black swan is sufficient to show it to be false. Such all-or-nothing statements are in fact potentially falsified by a single observation.

An initial hypothesis can derive from anywhere – specific observations or merely conjecture, it does not matter. Hypotheses not founded on prior observations or based on few or poor observations, of course, are less likely to withstand critical testing. The point is that where a hypothesis comes from is secondary to whether it is formulated so as to be testable. From the initial hypothesis, generalizations and predictions are made about the natural world. Every member of a species, for example, is hypothesized to share a unique combination of characters. And a resolved cladogram predicts that

every character or species subsequently discovered will conform to the predicted pattern of relationships.

Cladistics or phylogenetic systematics seeks to reconstruct patterns of evolutionary history and to construct natural classifications. Darwin acknowledged the pioneering work on phylogenetic "trees" by Ernst Haeckel in Germany (**Figure 1**), but a century would pass before the theoretical advances were made that were necessary to make phylogenetics rigorous (Hennig, 1966). While frequently ignored, the technical distinction between a cladogram and a (phylogenetic) tree is a useful one (see Nelson and Platnick, 1981). A cladogram is merely a theory of the relative relationships among species: A and B are more closely related to one another than either is to C, for example. A tree, however, specifies common ancestors, whether named or hypothetical. Thus, a cladogram is more general and may be consistent with multiple trees.

Because it is easier to develop rigorous theories of cladograms than trees, taxonomists often emphasize the distinction between pattern and process. This is not to deny that evolution exists or is responsible for the cladistic patterns: the fact of evolution is a necessary and sufficient background assumption for doing cladistics. Nor does it imply that studies of processes in evolution cannot or should not be pursued. To the contrary, the cladogram assists in the interpretation of evolutionary processes by narrowing explanations to those consistent with an explicit chronology and process studies are a necessary complement to taxonomic studies.

In practice, it is very helpful to have a cladistic framework before interpreting evolutionary processes. And it is not coincidental that evidence of the relationships among species and their characters preceded and laid the groundwork for Darwin's theory of evolution. Without a pattern of relationships, there is nothing for evolution to explain. It is noteworthy that some patterns of relationship have been recognized for a thousand years or more in spite of radically changing world views from creationism to modern evolutionary theory. Early classifications divided species into groups that had or lacked some property. Thus there were vertebrates and invertebrates. These have been described as "A" and "not-A" groups. "A" groups are defined based on a shared attribute (in our example backbone), while "not-A" groups are defined based on the absence of something. Given that cucumbers and racing bicycles also lack backbones, it is not surprising that groups like invertebrates generally prove to be artificial. The contemporary period of taxonomy may fairly be described as the concerted push to eliminate "not-A" groups and replace them with "A" groups.

The Evidential Basis of Taxonomy

Taxonomy is comparative and integrative, summarizing all attributes of species and higher taxa that are heritable, observable, and comparable as characters. For the taxonomist, a *character* is an attribute of a species or clade that is constantly distributed among all its members (Nixon and Wheeler, 1992). In the context of phylogeny, a synapomorphy (character) is shared by an ancestral species and its descendant species. A character is a theoretical construct and may occur in its original or in some subsequently modified form (Platnick, 1979). The character forelimb is shared among all quadrupeds, for example, irrespective of whether it appears as the arm of a man, the wing of a pterodactyl, or the flipper of a penguin. Similarly, a particular character "wing" (not homologous but analogous with structures by the same name in insects, bats, or pterodactyls) is shared by all birds, even when expressed as a flipper in penguins.

Attributes shown to be inconstantly distributed within a species are not characters, but traits. Traits are also known as polymorphisms and are the stuff of microevolutionary studies. Changing frequencies of alternative traits in populations provide crucial evidence of selection at work. Important for population genetics, the variability of traits makes them inappropriate for taxonomic studies at and above the level of species.

Any attribute that is heritable, constantly distributed within a species or clade, and differing between species or clades is a potential character. Character sources are not limited to, but include anatomy, embryology, paleontology, molecular genetics, biosynthetic pathways, and ethology. Hennig (1966) used the term *holomorphology* to express the totality of characters and emphasized that the source of characters is less important than how they are analyzed. Simple observations with a hand lens can be rigorously analyzed, while the most expensive sequence data can be rendered unreliable if improperly analyzed. "Total evidence" analyses are those that combine all available sources of data. An alternative practice is to exclude data perceived to be weak or incomplete and to map it onto a reconstructed cladogram.

Mature sciences often develop theories that are scientific (testable) yet not directly derived from observations possible with the senses. Subatomic physics is based on the existence of particles never seen by anyone, but these ideas have been shown to map precisely onto predicted outcomes in power plants and atomic bombs. In taxonomy, characters represent an intellectual tradition traceable to the idea of homology articulated by Richard Owen in 1842 and made consistent with evolutionary theory by Lankester in 1870, who differentiated homology from homoplasy (Nelson and Platnick, 1981). Even choices of what specimens to compare are relevant when a character may only be present at particular periods of the life cycle. Specimens at these character-bearing periods are called *semaphoronts*; thus, it may be necessary to compare only third-instar larvae or plants that have flowered in order to determine presence or absence of a character.

The availability of large amounts of DNA sequence data has added exciting new evidence for phylogenetic analysis that has both challenged and corroborated long-standing hypotheses (e.g., Hervé *et al.*, 2009). And as complete genomes are sequenced, molecular phylogenetic analyses are entering a new era of phylogenomics (Hervé *et al.*, 2005). As important as molecular evidence has become, it is not a panacea and is, as Hennig said, only as good as how it is analyzed. It is a common sense observation that as characters become more complex they store and convey more phylogenetic signal. A substitution of one nucleic acid for another at a single site along a strand of DNA is more easily replicated than the neck of a giraffe or even a complexly folded protein. We have only begun to deeply address the question "what is a character?" in the context of molecular evidence and some analytical

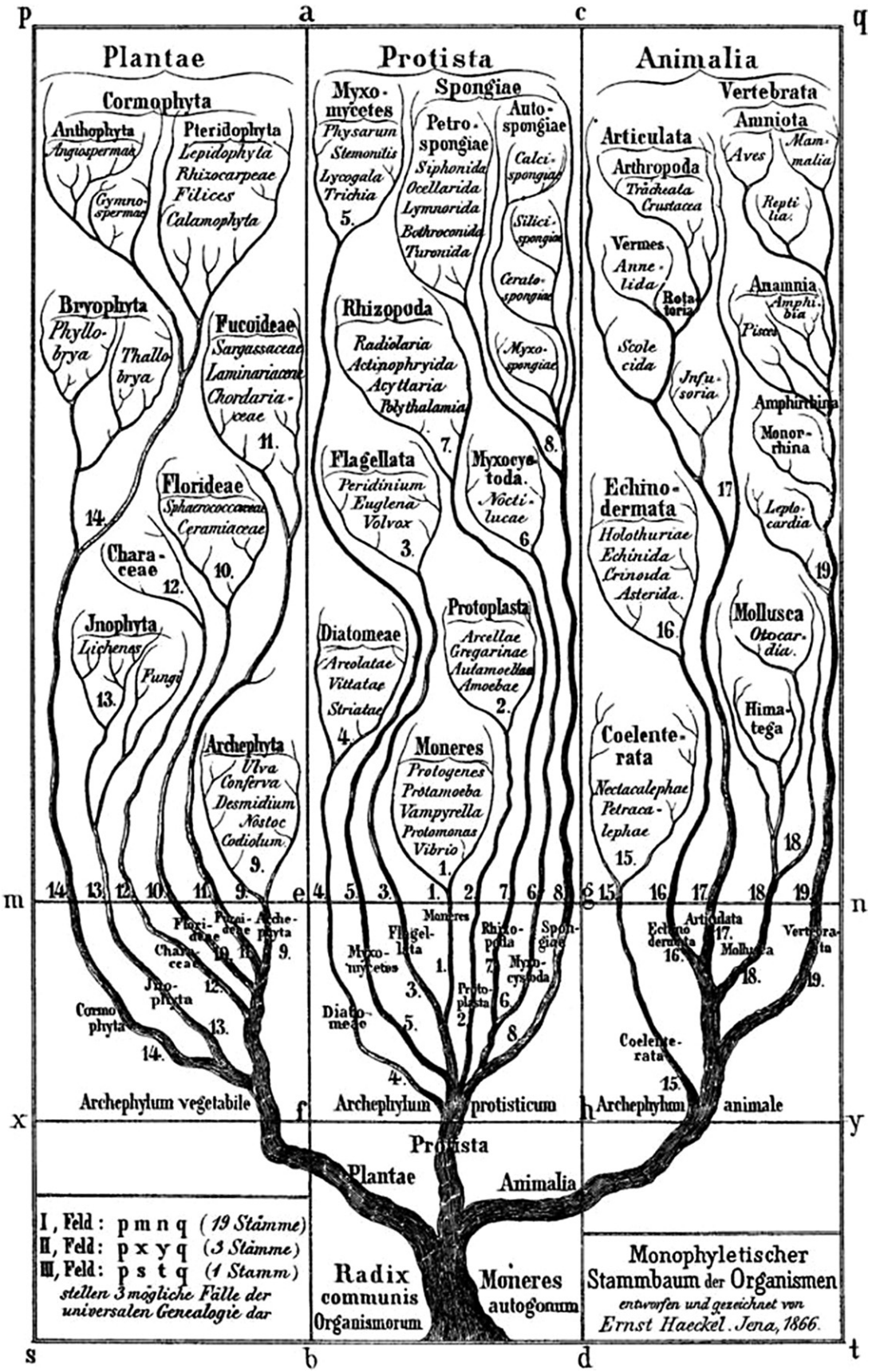


Figure 1 Treelike diagram depicting Ernst Haeckel's hypothesized phylogeny of life, cited for its genius by Charles Darwin.

methods of simple data appear to be throwbacks to phenetic rather than strictly phylogenetic analysis (Williams *et al.*, 2010).

Historical Context

The history of taxonomy is a long one, with roots lost in prehistory. Rudimentary taxonomy, it could be argued, is not unique to humans. Animals must distinguish food from foe in daily life, and mimicry complexes add further evidence of the observation powers of predators. The "Father of biological classification" is Aristotle who, in the third century BC, authored the first formal classifications. Aristotle's ideas would shape biological thought for the next 1500 years. As Europe emerged from the Dark Ages, an interest in the medicinal uses of plants drove the compilation of *herbals* and renewed interest in classification. Cesalpino's rudimentary species concept in *De Plantis Libri* was based on his observation that plant seeds gave rise to like kinds of plants was followed by John Ray's idea of species in the mid seventeenth century that included most of the elements of modern species concepts. It is estimated by ethno-biologists that perhaps 50,000 species are used by humans around the world presently and it is not unfair to say that the growth of taxonomic knowledge is directly correlated with the rise of human civilization and continued improvement of human welfare.

The Swedish naturalist Carolus Linnaeus (**Figure 2**) is considered the "Father of Taxonomy" and initiated the systems of scientific names and classification that we use today. Linnaeus gave to us the now standard set of hierarchical categories of classification: species, genus, family, order, class, phylum, and kingdom. The hierarchic nature of his categories, each more inclusive than the one below, proved preadapted for contemporary phylogenetic classifications, which are also hierarchic. He also devised Latin binominals as the names of species. The first word the genus name and the second the specific epithet. As the number of known species grew, this simple rule enabled us to use time and again descriptive and easily remembered epithets in multiple genera. Linnaeus knew and classified about 10,000 species in his lifetime. While this was a fitting achievement in his age of discovery, it also laid the foundation for the nearly two million species discovered and named since.

Comparative anatomical studies by Cuvier, Buffon, Lamarck, and Owen, among others, paved the way for the next major advance in taxonomy, evolution. It has often been claimed that Darwin's theory had little impact on taxonomy, but this is ingenuous. That the good descriptive work that had been a necessary precursor to Darwin could continue to be done in largely the same way after Darwin does not suggest that taxonomists somehow missed the evolutionary revolution. To the contrary, taxonomists had already been grappling with explanations for the patterns of similarity seen in anatomical and fossil evidence and were quick to see evolution as the explanation they had sought for the patterns in their classifications. John Henry Comstock, the founder of Cornell's entomology department, wrote in 1893 that "the description of a species, genus, family, or order, will be considered incomplete until its phylogeny has been determined so far as is possible with the data at hand."

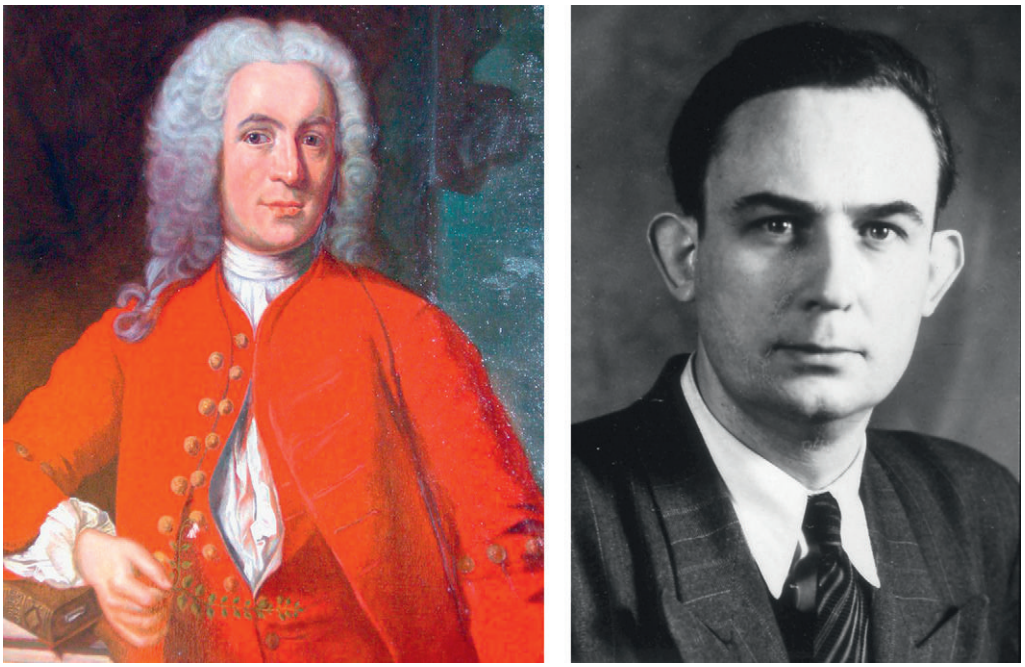


Figure 2 Prominent figures in history of taxonomy. (Left) Carolus Linnaeus (1707–1778), originator of modern classification and nomenclature, binominal nomenclature, and the hierarchic system of categorical ranks. He was also author of *Systema Naturae*, whose tenth edition in 1758 marks the beginning of recognized names for most animal taxa. (Right) Willi Hennig (1913–1976), originator of phylogenetic systematic or cladistic theory that realized Haeckel's and Darwin's vision a century earlier for classifications that reflect phylogeny through differentiation of synapomorphy and monophyly from related concepts. He was the author of *Grundzuge einer Theorie der Phylogenetischen Systematik* in 1950 and the highly influential *Phylogenetic Systematics* in 1966.

Although the goal of phylogenetic classification had become clear to post-Darwinian taxonomists, the method for achieving it remained enigmatic. The vacuum was soon filled instead with "evolutionary" or "eclectic" taxonomy that attempted to express both common ancestry and degree of divergence in classifications. Assumptions about evolutionary processes were speculative at best, and imposing degree of divergence on classifications led to another suite of problems. Many higher taxa corresponded to evolutionary grades rather than monophyletic clades.

In part, a reaction to the subjectivity of evolutionary taxonomy, phenetics, or numerical taxonomy developed putatively objective alternatives. The central claim of phenetics proved incorrect: Overall similarity did not reveal phylogenetic relationships. Rather than abandon phenetics, its practitioners admitted this failure and said instead that evolution was simply unknowable and that we should settle instead for objective, repeatable, artificial classifications. Many of the massive morphometric data sets compiled by pheneticists were subsequently shown to be suspect. Independence among their characters was questioned and in some studies the same ratios were clearly being measured over and over. More damning to phenetics, however, were two additional observations. First, their results were only repeatable given the same matrix (i.e., given no addition or subtraction of taxa or characters) and the same algorithm. By this time, there were scores of algorithms, any of which might give a different result from the next. Although each was "objective," the choice among them was purely subjective. Second, J. S. Farris demonstrated in 1979 that the phylogenetic system was superior in information content, the very criterion used to promote phenetics. Farris had delivered a deathblow from which phenetics would never recover. Adding insult to injury, he successfully hijacked the phrase "numerical taxonomy" to apply to his parsimony-based phylogenetic algorithms.

Hennig's 1950 *Grundzuge Einer Theorie der Phylogenetischen Systematik* had been read and cited by a handful of North American taxonomists, a few of which, such as Pedro Wygodzinski, had applied his principles. Almost immediately after the publication of Hennig's *Phylogenetic Systematics* in 1966 (a new, English-language book), a theoretical revolution was afoot. Phylogenetic systematics developed rapidly along two parallel lines of scholarship that would converge years later. James S. Farris was transforming "numerical taxonomy" from the trivially objective phenetic methods to sophisticated computer-aided phylogenetic analysis. In so doing, Farris also laid the groundwork necessary to cope with massive numbers of characters that would later emerge from molecular sequence techniques. Simultaneously, Gareth Nelson, Norman Platnick, Eugene Gaffney, Donn Rosen, Joel Cracraft, E.O. Wiley, and other scientists associated primarily with the American Museum of Natural History were exploring the connection between the philosophy of science of Sir Karl Popper, Willi Hennig's phylogenetic theory, and biogeographic work by Lars Brundin. The result was a powerful epistemic justification of Hennig and the transformation of cladistics to a rigorous science in its own right distanced from some of the evolutionary process assumptions common to taxonomic writings prior to Hennig and, to a lesser extent, inherent in Hennig's writings.

Linnaean Nomenclature

There are, from time to time, proposals to abandon or replace the Linnaean system of scientific names. While the Linnaean system is not perfect and continues to be improved on, its fundamental effectiveness for managing and communicating taxonomic information is proven. The fact that the Linnaean system has been used continuously by zoologists since 1758 (and botanists since 1753) is not, of course, an accident or a slavish adherence to tradition. Pre-Darwinians recognized patterns evident in character distributions that were used by Linnaeus and his students to group "related" animals and plants together. For Linnaeus, these relationships were the result of choices of Creation, but regardless of the causal process assumed, the evidence was expressible in hierarchic form. As the evolutionary implications of such patterns of resemblance became clear (as, for e.g., to Haeckel, Lankester, and other nineteenth century biologists), improved explanations for the same undeniable patterns became available. Given a rigorous method for the study of phylogeny by Hennig, a classification system that was hierarchic in nature was a necessity to convey relationships. The Linnaean system, due to its hierarchic logical structure and nested ranks, ideally met this need.

Elements of Biodiversity

When biodiversity is viewed from a functional perspective it can be described in terms of a series of increasingly inclusive units from organism to populations, ecosystems, and ultimately the biosphere as a whole. Viewed from a taxonomic perspective, biodiversity is seen in an evolutionary hierarchy progressing from individual to species to series of increasingly inclusive higher taxa or clades. For a great many practical purposes, the species may be seen as the element of organization, whether as the least inclusive taxon in a phylogenetic classification or the kinds of microbes, plants, and animals interacting in ecological networks.

Consider two islands of equal size with ten species each. On one island are ten closely related species of a single insect genus. On the other island are one species each of ten higher taxa (i.e., one insect, one tree, one nematode, etc.). In one trivial sense, the islands have equal diversity with ten species each. In a different sense, the second island has more diversity. In measuring biodiversity of ecosystems and places, it is important to consider phylogenetic diversity as well as species numbers. Were all 20 of these hypothetical species unique endemics and we could save only one of the islands, the author suspects that most of us would find it an easy decision to favor phylogenetic diversity over homogeneity.

In recent years, several numerical indices have been proposed to account for biodiversity at and above the species level. When biodiversity is seen through the lens of taxonomy, in terms of both species and phylogenetic diversity, needs and priorities for research and conservation can appear very different. Combined with research results from ecology, population biology, migration ethology, and assessments of bio-economics and so-called biodiversity hot spots, these

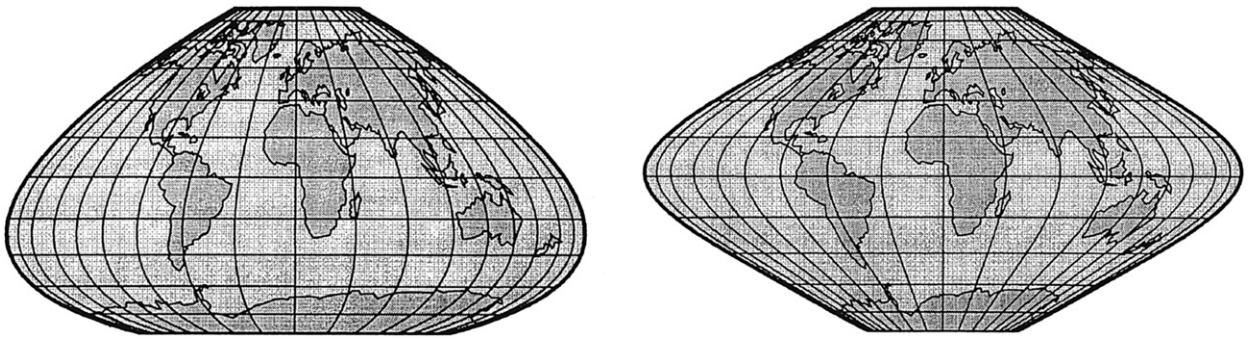


Figure 3 Visualization of two competing hypotheses of the global distribution of biodiversity. (Right) The tropical model, based on a simple count of numbers of species, argues that most biodiversity is equatorial with decreasing levels toward either pole. (Left) The Platnick model takes into account both phylogenetic diversity and species diversity and suggests that most biodiversity exists in both the tropical and the southern temperate regions, resulting in a pear-shaped globe.

indices create a fact base for decision making in natural resource management and conservation biology.

Using the imagery of the world to represent broad-scale geographic models of biodiversity distribution, we can see three phases in the twentieth century. First was what Nelson and Platnick (1981) called the "Sherwin–Williams" model, with biodiversity originating in Northern Hemisphere centers of origin and species spilling southward in decreasing abundance. This traditional view was biased by many factors, including a focus on certain vertebrate taxa, but especially the location of major universities, museums, and herbaria in Europe and North America.

As tropical forests were explored in greater detail, it soon became obvious that there were large numbers of species at and near the equator. From this species counting exercise, a tropical biology mythology posited that the tropics are most diverse, with biodiversity levels diminishing latitudinally as one moves northward or southward. There are many taxa with enormous numbers of tropical species, and based on existing species counts the tropics unequivocally house more species than any other region in the world. This tropical view morphed the world into one with a much exaggerated equatorial region.

For many taxa, the southern hemisphere is significantly more species-rich than the northern. Further, when phylogenetic uniqueness is taken into account, there exist a disproportionate number of relictual clades found only in the far south. This suggests a radically alternative view of global biodiversity distribution that was described by Norman Platnick as pear-shaped, with most species found in the tropics and southern hemisphere and with diminishing numbers northward. This is the emerging view of biodiversity as a whole (**Figure 3**).

Given the many advances in population genetics, molecular genetics, and phylogenetic systematics, one might suspect we would have converged on one modern concept of species by now. Unfortunately, the twentieth century was marked by an increased number of competing definitions. This has been fueled partly by the confusion of the evidence and goals of population biology with that of taxonomy, beginning about 1940 (Wheeler, 2008, Introduction). Hennig (1966) was careful to articulate the distinction between what he termed tokogeny and phylogeny. Tokogeny has to do with birth relationships among individuals within a population or single species. Phylogeny, however, refers to relationships among *species*. Above this tokogenetic–phylogenetic line are phylogeny, macroevolution, and much of the subject matter of taxonomy. Below are population genetics, various stages of incomplete speciation, and the subject matter of population genetics. Species are at once the upper limit of the one science and the lower limit of the other. It is the importance of species to questions above and below this line that has contributed to friction and confusion over their definition (**Figure 4**).

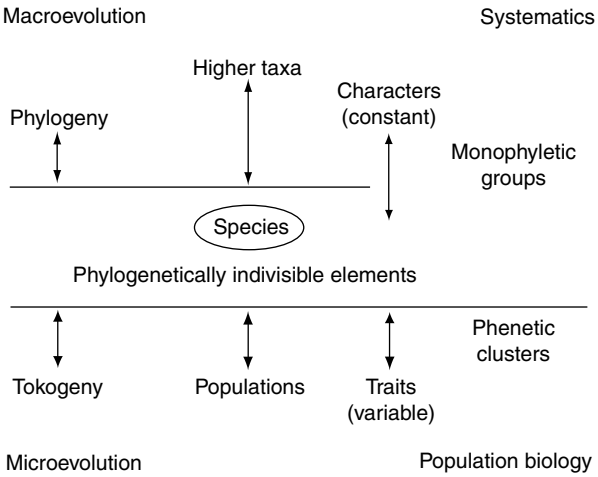


Figure 4 "Line of death" depicting the significance of species in Hennig's distinction between tokogeny or birth relationships (shared among individuals within and between populations of a single species) vs. phylogeny or patterns of relative recency of common ancestry shared among species. Hennig's line – species – also corresponds with divisions between taxonomy and population biology, characters and traits, microevolution and macroevolution, and upward between species and higher taxa.

Predictive Classifications

Cladistic hypotheses are expressed graphically as cladograms and in written form as phylogenetic classifications. Conventions exist by which the branching structure of the cladogram can be retrieved from a hierarchic list of taxon names, such as

phyletic sequencing (Schuh, 2000). Species are the least inclusive taxonomic category, although some authors recognize subspecies. In order that classifications and higher taxon names reflect the most recent ideas about phylogeny, they must be periodically revised.

In the interest of precision, Hennig (1966) introduced terminology that expressed the fundamental concepts in his theory. Monophyletic groups or clades include an ancestral species and all (and only) its descendant species. Paraphyletic groups are those that include an ancestral species and some, but not all, of its descendant species. Polyphyletic groups do not share a common ancestral species. Each kind of group can also be defined based on the evidence supporting it. Synapomorphies are shared derived characters (apomorphies or evolutionary novelties) and indicate monophyly. Symplesiomorphies or shared primitive features used to group species result in paraphyly, that is, groups including some but not all descendant species. Errors in interpreting characters sometimes assert that independent characters are the same, leading to unrelated species being classified together polyphyletically. Instances of convergent evolution, parallelism, or reversal to ancestral states are collectively known as homoplasy (in contrast to similarity due to common ancestry, homology).

Figure 5 shows examples of such groups referring to a primitively flightless silverfish, a damselfly, a beetle, and a fly. The best available insect classification suggests that fly + beetle constitutes a monophyletic group whose sister-group is the damselfly. The silverfish is then sister to the other three combined. Were the silverfish and damselfly grouped together based on their comparatively simple metamorphosis, they would be paraphyletic because simple metamorphosis is a shared, primitive similarity in comparison to holometabolism. Were the small, aristate antennae of the damselfly and fly taken as similarity, the resulting group would be polyphyletic since it is based on convergent similarity and not on common ancestry. This example was used for simplicity. In practice, paraphyletic and polyphyletic groups involve a much broader scale of grouping errors.

Although Hennig's terminology has been applied below the species level, this is a questionable practice as it violates assumptions in his theory that explain why historical patterns are to be expected in phylogenetic systems and not in tokogenetic systems. Hennig's groupings were groupings of species (not individuals or populations) marked by completely transformed characters (not partially transformed ones). Hennig's method thus required recognizing species prior to cladistics analysis and thus granted special status to species.

Another trend is to insist that species be "monophyletic." Not only does this violate Hennig's original meaning of the term, it also imposes a constraint on species that is, at least in certain theoretical situations, untenable. Consider a widespread ancestral species from which a tiny peripheral population is isolated. If that isolate attains full species status, it may well be "monophyletic" but its ancestral species clearly is not. This kind of situation is potentially common enough to suggest that insistence on unique common ancestry within species is ill advised. In this hypothetical example, the ancestral species clearly does not cease to exist and the only derived characters that is possesses are shared also with the daughter species. Such cases are not recognized as such in practice because the evidence would merely show a sister-group relationship between mother and daughter species. This is another example of why taxonomists work with general cladograms rather than specific phylogenetic trees: either the mother species in this instance has a unique apomorphy not yet discovered or it is the "paraphyletic" ancestor of the isolate species. Taxonomists leave it to other evolutionary biologists to speculate as to the particular processes behind this pattern.

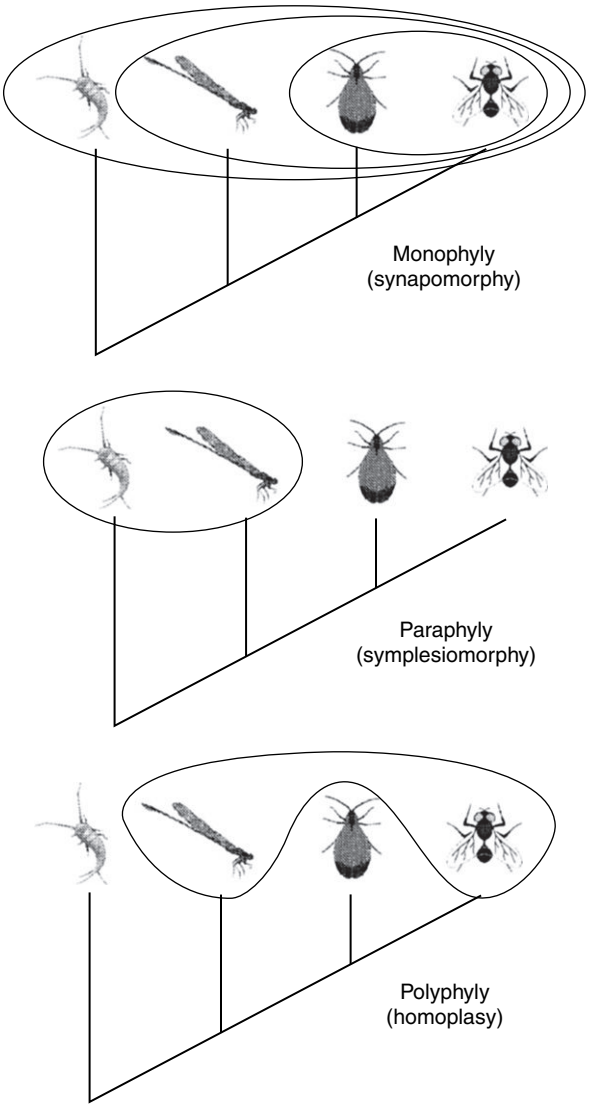


Figure 5 Example of Hennig's concepts of monophyly (based on synapomorphy), paraphyly (based on symplesiomorphy), and polyphyly (based on homoplasy) illustrated by four hexapods: silverfish, damselfly, beetle, and fly. See text for discussion.

We have already seen that many species concepts are presently in use. For many practical purposes, this is not in itself problematic. For others, however, it should be of great concern. In the absence of a uniformly applied species concept, certain fundamental comparisons become impossible. Are there more species in a comparable sized area of tundra or rain forest? Are there more species of moths or mammals?

These are critically important questions for conservationists and ecologists, but when dozens of species concepts are in use it is not at all clear what numbers to compare to find the answers. In the context of evolutionary biology it is interesting to seek explanations for why one lineage produces an order of magnitude more species over the same span of time as another. Again, unless the species counts are comparable, we cannot legitimately pose the question.

In zoology, the biological species concept gained nearly universal nominal acceptance during the middle of the twentieth century. The author says nominal because most textbooks and taxonomic publications claimed allegiance to the biological species concept, but very few ever gathered the data necessary to rigorously apply it. The biological concept was popularized by Mayr (1942) and others who advocated "population thinking" and is fundamentally tested by the ability to interbreed. Rosen (1979) tested the concept with freshwater fishes from Guatemala and demonstrated that the ability to interbreed is a shared-primitive condition and that the truly evolutionarily interesting phenomenon is the loss of the ability to interbreed. Botanists never found the biological species concept to be adequate for their needs, partly because of a greater frequency of hybridization among species. More recently, the biological concept has been attacked from several perspectives (cf. chapters in Wheeler and Meier, 2000) and there is growing support for the phylogenetic species concept.

Working simultaneously, two pairs of authors produced nearly identically worded versions of a phylogenetic species concept: Eldredge and Cracraft (1980) and Nelson and Plantick (1981). Their phylogenetic species concept is very similar to the morphological concepts in broad use by taxonomists before they were set aside in favor of the biological species concept, but it is formulated specifically in a phylogenetic framework. According to the phylogenetic species concept, a species is simply the smallest aggregation of (sexual) populations or (asexual) lineages diagnosable by a unique combination of character states. The argument has been made that the phylogenetic species concept is the unified species theory we have sought (Wheeler, 2010), but this is not yet a consensus view.

Debate continues regarding the best alternative. This battle over species concepts is far from a mere academic exercise (**Table 1**). Any effort to inventory or document biodiversity, to compare the biodiversity present in one habitat or geographic place versus another, and any wording in regulations aimed at natural resource management or biodiversity conservation will inevitably be described in terms of the numbers and kinds of species affected. Adopt the wrong concept, and we grossly under- or over-estimate biodiversity with potentially catastrophic effects. It is important that this debate proceed with all due haste, but it is no less important that we arrive at the correct answer and not simply an expeditious or democratically popular one. Every student of biodiversity has a responsibility to consider this question carefully and fully and has a stake in the outcome.

The Missions of Systematics

Through a major initiative by the International Willi Hennig Society, the Society of Systematic Biologists, the American Society of Plant Taxonomists, and the Association of Systematics Collections, with funding from the U.S. National Science Foundation, the systematics community laid out a vision for the future of the field in the mid 1990s. Its visionary 1994 report, *Systematics Agenda 2000: Charting the Biosphere* (and parallel technical report), returned the focus of the systematics community to three traditional core missions that are energized by recent theoretical and technological advances and made urgent by the biodiversity crisis. Collectively, these missions address the most fundamental scientific questions about biodiversity (Cracraft, 2002).

Mission 1: To Discover, Describe, and Inventory Earth's Species

This mission seeks to answer several of the most fundamental of all biodiversity questions, such as "What and how many species exist?" "How are they related?" "What properties do they share with other species?" and "Where do they occur?" In slightly more than two centuries, taxonomists have documented nearly 2 million species (Chapman, 2009; IISE, 2012). An estimated ten million species of plants and animals (eukaryotes) remain to be discovered and described. Because few microbes (bacteria, archaea, etc.) can be cultured in the laboratory, our knowledge of their diversity lags behind. However, recent advances in metagenomics have allowed scientists to sample the DNA diversity of microbes directly from their environments. At the completion of the decadal survey of marine life, microbiologists speculated that the oceans alone may home to as many as 20 million microbial species.

Species exploration is on the threshold of a Renaissance as cyber-enabled taxonomy, or cybertaxonomy, comes into its own (Wheeler, 2004, 2008). Constraints of the past that were

Table 1 Alternative species concepts[a]

Biological species concept (Mayr, 2000 p. 17): "groups of inter-breeding natural populations that are reproductively isolated from other such groups."
Hennigian species concept (Meier and Willmann, 2000 p. 31): "reproductively isolated natural populations or groups of natural populations. They originate via the dissolution of the stem species in a speciation event and cease to exist through either extinction or speciation."
Autapomorphic species concept (Mishler and Theriot, 2000 p. 44): "the least inclusive taxon recognized in a formal phylogenetic classification … grouped into species because of evidence of monophyly" (see original for more extensive wording).
Evolutionary species concept (Wiley and Mayden, 2000, p. 73): "an entity composed of organisms that maintains its identity from other such entities through time and over space and that has its own independent evolutionary fate and historical tendencies."
Phylogenetic species concept (Wheeler and Platnick, 2000, p. 58): "the smallest aggregation of (sexual) populations or (asexual) lineages diagnosable by a unique combination of character states."

[a]Citations from Wheeler and Meier (2000).

largely due to limited access to research resources are falling to the wayside one by one. The Global Biodiversity Information Facility has mobilized millions of museum records; the Biodiversity Heritage Library project is digitizing the natural history literature since 1758; the Encyclopedia of Life is building a Web page for every species; software to facilitate descriptive work is evolving rapidly toward electronic monography; institutions are busily populating digital archives with images of type and rare specimens; and remotely operable instruments are coming online that directly connect researcher and museum specimen. As cybertaxonomy flourishes in the decades ahead, the rate of species discovery and frequency with which species hypotheses are tested and improved will dramatically increase. And taxonomy will be increasingly democratized, closing a century's old gap between developed countries with great museums and libraries and developing nations with great species diversity, and removing obstacles to the quality of work achievable by motivated amateurs. Equally exciting is the facilitation of international cooperation among taxonomists and natural history museums. Recent Planetary Biodiversity Inventory projects (Knapp, 2008; Page, 2008) demonstrated the efficiency of teamwork in taxonomy and several strategies have been proposed to discover and describe 10 million species in no more than 50 years (Wheeler *et al.*, 2012).

Mission 2: To Analyze Phylogeny and Predictively Classify Species

Reconstruction of phylogeny is important to life science research in general and to predictive and maximally informative classifications in particular. All biological structures and functions can and should be seen in their historical (evolutionary) context. Phylogenetic classifications provide this conceptual framework that has proven instrumental to advances in fields as diverse as anatomy, behavior, and chemical ecology. Phylogenetic classifications allow us to efficiently store and retrieve what we know about biodiversity. Importantly, they also allow us to predict that we do not yet know. By virtue of knowledge of the biology of related species classified in the same genus, family, or kingdom, we can predict shared attributes of species not yet studied in detail. As we confront the exploration of tens of millions of additional species, such a predictive framework is essential.

Mission 3: To Disseminate Data, Information, and Knowledge

Cybertaxonomy is revolutionizing how taxonomic data, information, and knowledge are created, accessed, and used. The gold standard for taxonomic information remains the revision or monograph. Such studies comprehensively review and compare all available evidence, efficiently and simultaneously testing the validity of known species and describing new ones. In highly diverse taxa such as many flowering plants and insects, revisions were often completed only once or a few times each century. As research resources become available in digital form, it is possible to reconceive this process. In the future, it is likely that teams of specialists will collaborate to assemble and continually improve on comprehensive knowledge bases for higher taxa that will fulfill the traditional role of the monograph. In any case, as a taxonomy-specific cyberinfrastructure matures, we can expect vastly more and more reliable taxonomic data, information, and resources.

Importance of Taxonomy

The biodiversity crisis (e.g., Wilson, 1992) has both increased the importance of taxonomic information and made the growth of taxonomic knowledge an urgent and high priority. How much farther can we push our understanding of the function of ecosystems without knowledge of what species exist and how they interact in complex webs? Without baseline data on what species exist and where, how can we hope to detect invasive species, shifts in floras and faunas in response to climate change, or losses of species diversity? And unless we explore, discover, and describe species now – expanding natural history museums and herbaria to document their diversity – we severely limit the detail with which we can continue to reconstruct and study the origin and diversification of life on our planet. The reasons that taxonomy is important to science and society include the following.

Conceptual Framework

Understanding the origin, organization, and function of the biosphere requires the input of many geological and biological disciplines from ecology to molecular genetics, population genetics, and phylogenetic systematics. While existing distributions and combinations of species are shaped by climate, physiography, soils, and competition, they cannot be fully understood without also understanding history. Many interactions and adaptations are constrained by phylogeny: a species can only function and change within the band of genetic information inherited from its ancestral species. A phylogenetic conceptual framework is a necessary complement to virtually any comparative question in the life sciences.

Language of Biodiversity

Formal botanical, zoological, and microbiological nomenclatures provide the language for communication about biodiversity. Although names never convey the full story of phylogeny, they are nonetheless powerful communication tools. If one beetle family, Staphylinidae, is mentioned, nearly two million species that are not rove beetles are immediately excluded… that is saying a lot. If two beetle families classified in the Staphylinoidea, say Leiodidae and Staphylinidae, are mentioned, it is known that the two families are related and share a common ancestry and the evolutionary novelties (synapomorphies) associated with it. We know from the "-idae" endings that the two are both ranked at the family level and therefore one is not included within the other. At the species level, the generic part of the binomial tells us something of relationships. And the specific epithet frequently tells us something about the species itself, often being descriptive of the organism (such as *Quercus rubra*, the red oak) or its origin (such as *Trechus appalachiensis*, a ground beetle that lives in the Applachian Mountains).

Although common names are not governed by enforcible rules, convention within taxonomic communities sometimes tells us useful information, too. For example, when entomologists use two words in a common name it usually points to one taxon while combined it points to others. For example, house flies, crane flies, and deer flies all belong to the order Diptera or true flies, whereas dragonflies, butterflies, and fireflies represent unrelated groups (Odonata, Lepidoptera, and Coleoptera, respectively).

Information Storage and Retrieval

Linnaean binomial species names have served as unique identifiers for species since 1758 and were easily adapted to serve the same purpose equally well in electronic information systems. Nearly every modern database from GenBank to the Encyclopedia of life makes use of scientific names to organize, aggregate, and retrieve information. Innovative work on taxonomic indexing is expanding its role (Patterson *et al.*, 2005).

Evolutionary Record

The collections of specimens, tissue samples, DNA sequences, recordings, and other material evidence of biodiversity curated in natural history museums and botanical gardens around the world represents the most extensive and permanently useful record of the results of evolution. No action today in response to the biodiversity crisis is more important that exploring species diversity and preserving evidence of it in collections. We cannot anticipate where future curiosity and discoveries will lead us, but we can be assured that collections will remain a rich source of answers (NSTC, 2008). Our generation is the last that will have access to such a complete record of 3.8 billion years of evolution. Aiming to expand and develop collections as a comprehensive record of biological diversity at and above the species level should be a top priority for this century.

Environmental Baseline

Only by completing an inventory of earth's species can we create a baseline of environmental information sufficient to enable us to detect, monitor, measure, and adapt to biodiversity decline. Biodiversity sustainability is appropriately seen as a major goal for conservation biology and ethical land management, in large part to assure long-term ecological services. Unless basic taxonomic studies are completed, however, we will remain incapable of making fact-based environmental decisions and policies. Simply put, unless we know what kinds of plants, animals, and microbes exist in the first place, and where they are found in the biosphere, how can we hope to achieve biodiversity sustainability or even know if we have done so (Wheeler *et al.*, 2012)?

Biomimicry

The history of life on earth is written in innovative adaptations to selection pressures. Natural selection has, for nearly four billion years, undertaken rigorous trial-and-error experiments to find sustainable and effective solutions for survival and problem-solving. Time and again, humans turn to biodiversity for ideas to solve our most pressing problems, and we are rarely disappointed. From antibiotics to sharkskin Olympic swimming suits, Mother Nature yields creative answers. Yet most such discoveries have relied on serendipity. A deliberate effort to mine biodiversity for answers, an emerging area called biomimicry (Benyus, 1997), is relatively new. As cybertaxonomy matures, one goal should be to create the shortest paths to promising answers. Among the most promising tools are species descriptions annotated to ease the retrieval of evolutionary novelties and phylogenetic classifications that allow the prediction of where to search for species with a particular property.

Appendix

List of Courses

1. Introductory Taxonomy or Systematic Biology
2. Introductory Conservation or Sustainability
3. General Biology

See also: Biodiversity, Evolution and. Cladogenesis. Darwin, Charles (Darwinism). Nomenclature, Systems of. Phylogeny. Species, Concepts of. Species Diversity, Overview

References

Benyus JM (1997) *Biomimicry: Innovation Inspired by Nature*. New York: Harper-Collins.

Chapman AD (2009) *Numbers of Living Species in Australia and The World*. Canberra: Australian Biological Resources Study.

Cracraft J (2002) The seven great questions of systematic biology: An essential foundation for conservation and the sustainable use of biodiversity. *Annals of the Missouri Botanical Garden* 89: 127–144.

Crowson RA (1970) *Classification and Biology*. London: Heinemann.

Eldredge N and Cracraft J (1980) *Phylogenetic Patterns and The Evolutionary Process*. New York: Columbia University Press.

Farris JS (1979) The information content of the phylogenetic system. *Systematic Zoology* 28: 483–519.

Gaffney ES (1979) An introduction to the logic of phylogeny reconstruction. In: Cracraft J and Eldredge N (eds.) *Phylogenetic Analysis and Paleontology*, pp. 79–111. New York: Columbia University Press.

Gould SJ (1989) *Wonderful Life*. New York: Norton.

Hennig W (1966) *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Hervé P, Delsuc F, Brinkmann H, *et al.* (2005) Phylogenomics. *Annual Review of Ecology, Evolution, and Systematics* 36: 541–562.

Hervé P, Derelle R, Lopez P, *et al.* (2009) Phylogenomics revives traditional views on deep animal relationships. *Current Biology* 19: 706–712.

IISE, State of Observed Species 2011. International Institute for Species Exploration, Tempe, Arizona, retrieved from http://species.asu.edu/SOS (2012).

Knapp S (2008) Taxonomy as a team sport. In: Wheeler QD (ed.) *The New Taxonomy*, vol. 62, pp. 33–53. Boca Ratan: Taylor and Francis.

Mayr E (1942) *Systematics and the Origin of Species*. New York: Columbia University Press.

Mayr E (1969) *Principles of Systematic Zoology*. New York: McGraw-Hill.

National Science and Technology Council (2008) *Scientific Collections: Mission-Critical Infrastructure of Federal Science Agencies*. Washington, DC: Office of Science and Technology Policy.

Nelson G and Platnick NI (1981) *Systematics and Biogeography*. New York: Columbia University Press.

Nixon KC and Wheeler QD (1992) Extinction and the origin of species. In: Novacek MJ and Wheeler QD (eds.) *Extinction and Phylogeny*, pp. 119–143. New York: Columbia University Press.

Page LM (2008) Planetary biodiversity inventories as models for the new taxonomy. In: Wheeler QD (ed.) *The New Taxonomy*, pp. 55–62. Boca Ratan: Taylor and Francis.

Patterson DJ, Remsen D, Marino WA, and Norton C (2005) Taxonomic indexing – Extending the role of taxonomy. *Systematic Biology* 55: 367–373.

Platnick NI (1979) Philosophy and the transformation of cladistics. *Systematic Zoology* 28: 537–546.

Popper K (1968) *Logic of Scientific Discovery*, 2nd edn. New York: Harper.

Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bulletin of the American Museum of Natural History* 162: 267–376.

Schuh RT (2000) *Systematic Biology*. New York: Columbia University Press.

Simpson GG (1961) *Principles of Animal Taxonomy*. New York: Columbia University Press.

Wheeler QD (2004) Taxonomic triage and the poverty of phylogeny. *Philosophical Transactions of the Royal Society, Series B* 359: 571–583.

Wheeler QD (ed.) (2008) *The New Taxonomy*. Boca Ratan: Taylor and Francis.

Wheeler QD (2010) What can we learn from 20th century concepts of species? Lessons for a unified theory of species. In: Jahn I and Wessel A (eds.) *Für eine philosophie der biologie*, pp. 43–59. Munchen: Kleine Verlag.

Wheeler QD, Knapp S, Stevenson DW, *et al.* (2012) Mapping the biosphere: Exploring species to understand the origin, organization and sustainability of biodiversity. *Systematics and Biodiversity* 10: 1–20.

Wheeler QD and Meier R (eds.) (2000) *Species Concepts and Phylogenetic Theory: A Debate*. New York: Columbia University Press.

Williams DM, Ebach M, and Wheeler QD (2010) Beyond belief: The steady resurrection of phenetics. In: Williams DM and Knapp S (eds.) *Beyond Cladistics: The Branching of a Paradigm*, pp. 169–195. Berkeley: University of California Press.

Taxonomy, Methods of

Richard Irwin Vane-Wright, University of Kent, Cantenbury, UK; Natural History Museum, London, UK, and Canterbury Christ Church University, Cantenbury, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Autapomorphy An advanced or derived character state unique to an individual taxon (of any rank).
Clade A complete ancestor–descendant lineage (cf., monophyletic group).
Cladogram A dendrogram expressing estimated cladistic relationships among taxa; a cladogram has no direct connotation of ancestry and the long axis does not connote time.
Dendrogram A branching, nonreticulate diagram expressing nested hierarchical relationships or similarities (or both) between entities (e.g., taxa) such that the entities only appear at the tips of terminal branches.
Diagnosis A set of attributes sufficient to define, characterize, or identify a given taxonomic group or separate it from others.
Grade A group based on, for example, the functional level of organization or overall similarity rather than ancestor–descendant relationships.
Monophyletic group A group comprising a given (hypothetical) ancestor and all its descendants; in a restricted sense, within a given set of terminal taxa, all the members of a subset arising from a common ancestor that has given rise to no other member(s) of the entire set.
Paraphyletic group A group comprising a given (hypothetical) ancestor and only some of its descendants.
Phenogram A dendrogram expressing the overall similarity (long axis) between terminal taxa and sequentially linked groups of terminal taxa; the nodes do not connote specifiable characters and the long axis does not connote time.
Polythetic group A natural taxonomic group in which the terminal taxa are not known to share any universal unique character(s) but are nonetheless united by overall similarity (phenetics) or global parsimony (cladistics).
Sister groups Two terminal taxa or monophyletic groups that share a common ancestor that has not given rise to any other taxon under consideration.
Symplesiomorphy Character present in all subordinate taxa of a group under consideration.

> Taxonomy is a highly controversial subject, and the issues are inextricably bound up with philosophical disputes that have endured for centuries. The problems are so important that no biologist can totally avoid facing them – Ghiselin (1969).

> No one would think much of a chemist who confused water and benzene "because they look alike." – Cain (1954).

Taxonomy plays the central role in the scientific discipline of systematics (Stuessy, 2009, p. 8), although, according to some authors (e.g., Schuh and Brower, 2009), the two are largely or even entirely synonymous. Systematics makes an essential contribution to our understanding of biological diversity, including its origins, distribution, and maintenance. The primary task of taxonomy is systematization: to establish and provide an account of biological order among the diversity of organisms. This involves enumerating the kinds of living things that exist and have existed in the past and determining the patterns of difference and connection among them. By giving expression to these patterns through naming inclusive sets, subordinate sets, and the least included entities (e.g., higher classes, genera, species, subspecies, and varieties), taxonomists have produced the general reference system of biological classification: a categorical arrangement of named, diagnosable groups and components (taxa) that can be used to refer to all known living and extinct organisms.

Although effective taxonomy predates Darwinism, systematics is now closely linked to the theory of evolution. Descent with modification gives justification to the general form of classification adopted: hierarchical, non-overlapping sets rather than fuzzy sets or periodic tables. The fundamental methods of taxonomic systematization are few. Ultimately, there may prove to be only two – overall similarity methods (grouping by degree of genetic or phenetic similarity) and hierarchical methods (grouping by an inclusive phylogenetic or cladistic relationship) – but the theoretical bases of taxonomy are various. This has led to a proliferation of approaches, often

with variants. To understand alternative methods of taxonomy, it is necessary to appreciate the philosophical differences between these approaches, and much of this article is devoted to exploring these differences rather than describing techniques. The methods of cladistics, currently the dominant approach, are the subject of a separate review (*see* See Also).

The Tasks of Taxonomy

Taxonomists perform five main functions: discrimination (primary discovery or recognition of basic units of biological diversity in nature, notably species, also entailing a formal description and diagnosis of such taxa), systematization (by comparison, assessing the similarities, differences, and relationships among taxa), classification (production of summary schemes that encapsulate current knowledge of taxa and their main inter-relationships), symbolization (application of names to taxa and classes: technical nomenclature), and identification (secondary recognition of taxa: providing means to match unidentified material to the established system).

This article focuses on different approaches to primary discrimination and comparison of taxa (systematization) and the ways in which this knowledge can be expressed as summary schemes (classification). The procedures of taxonomy as an information retrieval system, which includes making checklists, catalogs, and databases, identification (naming specimens), and technical nomenclature (application and regulation of names: naming taxa), are not dealt with here. Nor, explicitly, are the means of gathering taxonomic data, for example, from anatomy, karyology, biochemistry, and nucleic acid chemistry, and their corresponding subdisciplines (morphotaxonomy, cytotaxonomy, chemotaxonomy, molecular systematics, etc.). Although such subdisciplines are often described as taxonomic methods, in the wider context of systematics and taxonomy as a whole, they are subsidiary techniques. Much the same methods and problems of interpretation apply, whatever the source of empirical data. This also applies to experimental taxonomy (the investigation of taxa based on predictions from alternative systems in an attempt to gather data bearing directly on particular taxonomic problems; see the discussion in Stuessy, 2009, p. 20) and morphometrics (quantitative comparison of related taxa based on continuously varying characters; see, e.g., MacLeod and Forey, 2002).

Disagreements regarding taxonomic methods can lead to major differences in classification. Such discrepancies range from disputes over the validity of species, subspecies, and infrasubspecifics at one end of the scale (e.g., Funk and Omland, 2003) to the extreme opposite, where the composition of even the most inclusive categories (domains and their component kingdoms) remains uncertain (e.g., Foster *et al.*, 2009). By classifying the classifiers, it may be possible to identify some of the fundamental reasons, but not all taxonomic disagreements are due to the method alone because historical precedent and subjectivity still intervene. Moreover, in practice, a great deal of constructive taxonomic work is carried out with little reference to philosophy or explicit method, being achieved by pragmatic (empirical) intervention, notably the extension or modification of existing parts of the system (e.g., by description of new species and establishment of synonymy).

Building Blocks: Individuals and Characters

Taxonomy is an empirical activity. On the basis of characters derived from sensory data, individual organisms or life cycles can be discriminated and then divided among or gathered into groups, and groups within groups, and so on to produce a taxonomic system. This can be done either top–down (divisive methods, as in Linnean divisions) or bottom–up (agglomerative, as in most phenetic clustering methods). Although individual organisms and characters are the basic elements in this process, their definitions are not straightforward.

Individual Organisms

Organisms multiply, typically from spores or fertilized eggs, to reproduce the corresponding parental stage through a cycle of growth and differentiation. At any point along such an ontogenetic pathway, a particular organism can be referred to as an individual: fertilized egg, embryo, larva, subadult, adult, or postreproductive adult. However, some organisms multiply by splitting or propagation, rendering the parent–offspring distinction uncertain. Thus, land plants can spread by runners so that what appears to be a field of separate individuals can be identical clones of one original plant. In the case of social organisms (e.g., ants and termites) in which very large numbers of workers, guards, or other castes may be produced that never have an opportunity to reproduce, the entire colony has some of the properties of an individual. These distinctions are important because the reliability of the taxonomic process, other factors being equal, is dependent on the number of individual organisms sampled as well as the number of characters recognized and recorded. Multiple sampling from essentially the same individual organism can distort our assessments of natural variation or the lack of it and lead to erroneous conclusions (Wiens, 1999).

Attributes and Characters

Taxonomists compare organisms by means of characters. Characters are abstractions derived from the detectable attributes of individual organisms or social groups (e.g., consider "large, two curved unforked horns on head, two-color vision, herd with tails erect" as descriptive of wildebeest among mammals). To be informative, it is obvious that characters observed must not be universal throughout the entire group of organisms under investigation. The distribution of characters is the primary interest, and the degree to which they differentiate, coincide, and conflict will largely determine their usefulness for taxonomy.

In general, characters can be divided into two sorts: continuous (e.g., size, from large to small) and discrete (e.g., paired horns vs. no horns). However, this is not absolute but more a matter of scale. Thus, different individuals of a particular kind of fly might bear every possible combination of 1–20 hairs on the thorax: is this continuous or discrete

variation? In addition, coding can be arbitrary (how large is "large?") and subdivision can be ambiguous (MacLeod and Forey, 2002). If we observe different individuals with straight, curved, and forked horns, do we simply regard these as three separate characters, three states of one character (form of horn), or the intersection of two binary alternatives – horns straight/curved and horns simple/forked (with the expectation of being able to distinguish, in theory at least, two sorts of forked horns)? Moreover, there are other possibilities for coding such variables. The decision that authors make can affect the analysis (e.g., by encoding spurious information derived from inapplicable characters). Thus, if an animal lacks horns, and this is coded as three separate pieces of data (not forked, not curved, and not straight), such a redundancy can have undesirable effects on the analysis (Strong and Lipscomb, 1999).

In theory, unit characters must not only be nonredundant but also homologous and independent (Pimentel and Riggins, 1987). If different parts of the body are functionally interdependent or developmentally linked, it will be misleading to count these attributes as separate characters. This can occur when a single gene has a pleiotropic effect, influencing, for example, the color of one organ and the form of another. At the extreme, it is clear that multiple characters should not be created by logical correlation (e.g., treating both the circumference and the diameter of a circular organ as two unit characters). Some organs (e.g., parts of the male genitalia of insects) may appear so complex that counting them as one character seems unreasonable, but any stopping rule for subdivision may be uncertain, whereas other characters that appear simple may prove to be complex (e.g., a single functional bone formed by fusion during development, representing several characters in comparison with related taxa in which the equivalent bones have not fused).

Special and General Classifications

On the basis of characters held in common, individual organisms can be grouped into a large number of classes, which are of two general kinds. On the one hand, individuals can be grouped in terms of a particular attribute (e.g., green, round, four-legged, marine, planktonic, nocturnal, pollinating) or by small combinations of attributes to give prescriptions such as plankton-feeding marine organisms or nonflowering, epiphytic semiparasitic plants. On the other hand, they can be placed into categories of species, genera, families, orders, and so on. The former are regarded as artificial or special classes and are generally defined by reference to given attributes, whereas the latter are viewed, ideally, as natural kinds or natural groups that are discovered but cannot be defined a priori (although they can be diagnosed a posteriori).

Special classes frequently overlap, such as the overlay of plankton feeders, not all of which are marine, and marine organisms, not all of which feed on plankton, to define the special class plankton-feeding marine organisms; also, the same organism may recur in many different special classifications (e.g., eagle as flying organism, predator, or nest builder, and in the conjunction of all three). In contrast, natural groups typically form nested, nonoverlapping sets in which each kind of organism or group appears only once. Thus, within the general classification of birds, all eagles are included as members of the family Accipitridae, which also includes vultures, buzzards, hawks, and kites.

Evolution, Genetic Relationships, and the Natural Hierarchy

Natural groups comprise individuals with many attributes in common, whereas individuals belonging to special classes have relatively few shared characters. In practice, the difference between, for example, large marine animals and Cetacea (whales and dolphins) is simply that individuals of the latter natural class have far more in common than those of the former. Even so, the word "natural" has had many connotations in the context of taxonomy, including the essence of things classified (Aristotle), rationality (e.g., as in God's design), similarity, and explanatory power (Gilmour, 1940). Modern biology, however, has a unique theory that underpins its totality: the theory of organic evolution.

Ideas about evolution can be divided into a general theory of descent with modification and special theories about the processes affecting that descent (e.g., orthogenesis; natural, group, kin, sexual, and species selection; adaptive radiation; molecular drive; molecular clocks). Most systematists agree that the general theory of evolution not only provides a compelling justification for seeking one natural, general classification of living organisms but also suggests the basis on which this classification is most securely founded. A hierarchical pattern of exclusive sets and inclusive subsets can reflect the primarily divergent sequence of ancestor–descendant relationships.

Hennig (1966) used the term "tokogenetic" relationships for the reticulate genealogical links that occur between parents and offspring within an evolving, sexually reproducing population. When an entire population splits to form two or more divergent subsystems within each of which tokogenetic relationships are maintained but between which processes of genetic recombination are largely or entirely discontinued, speciation (or phylogenesis) occurs. If we describe two extant taxa as phylogenetically more closely related to each other than to any third taxon (regarding them as sister taxa), we imply that the two do not share tokogenetic relations now but did so in the past within a common ancestor that they do not share with any other living taxon. Such ideas were basic to Hennig's concept of phylogenetic systematics. However, many organisms reproduce without genetic recombination (asexual) and thus lack tokogenetic relationships. Moreover, the huge variety of sexual processes suggests that tokogenetic relationships are not only nonuniversal but also may differ fundamentally in the many lineages where they do occur (Margulis and Sagan, 1984).

Species and other taxa grouped to reflect their phylogenetic relationships are expected, by virtue of their historical connections, to share far more characters in common than members of artificially formed groups. This expectation extends to unknown attributes, making phylogenetic classifications highly predictive. However, another type of genetic relationship is increasingly recognized as important in organic evolution: lateral gene transfer. The chimerical nature of

lichens as fungal/algal symbionts has long been recognized. Following the work of Margulis, it is accepted that at least two of the cellular organelles found in all eukaryotes originated through symbiosis of fundamentally separate organisms. Evidence is accumulating from molecular phylogenetics that the genomes of many major life-forms are chimeras formed from multiple original sources. The extent to which hybridization and lateral gene transfer undermines current approaches to natural classification is uncertain, but at least at the domain level, a nonreticulate hierarchy based exclusively on divergence appears to be unrealistic (Doolittle, 1999).

Differing Philosophies and Methods of Taxonomy

Taxonomic methods have developed over time. The massive edifice of taxonomic classification involving millions of terminal (e.g., species, varieties) and higher (e.g., phyla, families, genera) taxa, on which the scientific study, practical use, and conservation of biodiversity depends has been built up over centuries. This system has been produced by thousands of different minds using different methods, working with different knowledge, under different influences, and often seeking different goals. Some parts have been revised and reworked repeatedly and others hardly at all. To understand the strengths, weaknesses, and limitations of the current distributed taxonomic system, it is necessary to appreciate the ways in which systematists of contrasting persuasions have sought order in nature and tried to reflect that order in biological classification. In the sections that follow, the various approaches and methods are reviewed in a historical sequence.

Essentialism, Idealism, and Preevolutionary Taxonomy

In his early writings, it is clear that Linnaeus, the founding father of modern taxonomy, was trying to detect the pattern of Creation in the classes and species that he recognized and named. Linnaeus' system was, superficially at least, very simple. Having divided all life-forms into animals and plants, these kingdoms were then divided successively, on the basis of one or a few defining characters, into a series of smaller units down to the level of the genus (Linnaeus's genera were roughly equivalent to modern families, but the names he applied are nonetheless now deemed to be of genus rank). Within each of these genera, a large number of terminal species were usually recognized, and each was given a short, diagnostic description. For each species, Linnaeus also accepted that nameable variations could occur but considered that these did not represent the fundamental plan of God's work.

According to Hull (1965), Linnaeus started out as an essentialist, claiming that he belonged to an intellectual tradition founded on Aristotle's methods of logical division (Linnaeus later modified his views considerably). Aristotle was perhaps more concerned with the classification of our knowledge about living things, or even the generation of knowledge through the process of classification, than he was with the classification of organisms as such. The young Linnaeus was trying to discover or reveal deep knowledge: the natural order of God's Creation. Through his search for an external criterion of verity, Linnaeus' concepts of species and higher taxa were fundamentally removed from those of the nominalists (*see* Nominalism).

Following Hull and Mayr, it has been argued that essentialism and related but distinct ideas of the early Greek philosophers, including Plato's idealism, have negatively influenced both the practice and the understanding of taxonomy. According to the typological species concept, every species was thought to have its own idealized plan or design. The task of the taxonomist was then to recognize each of these theoretical designs, and describe, divine, or define the essential features of these types so that individual, real organisms could then be assigned to them. Thus, for a Platonist, "taxonomic names are the names, not of organisms, but of concepts" (Ghiselin, 1969). Even today, many taxonomists consider given species as concepts (e.g., fulfilling some ideal as gene pools, potential interbreeding units, or mate recognition systems) rather than empirical entities that are discovered.

Idealism has influenced concepts regarding major groups of organisms, especially animal phyla, which many morphologists believed could be formulated as a series of fundamentally different body plans (ground plans or Baupläne). The notion of the ground plan was strongly developed by the ideal morphologists, who evidently affected some of the developing ideas of Hennig (*see* Phylogenetic Systematics and Cladistics) in the 1930s and 1940s, even though he flatly rejected their conclusions about the relationship between morphology and phylogenetics (Hennig, 1966, p. 11). According to Hull (1965), by supposedly ridding taxonomy of its essentialist and idealist burdens, and moving on to an empirical and relativistic evolutionary framework, Mayr, Hennig, and other modern systematists have liberated the science of taxonomy from a philosophical bind that held back progress for more than 2000 years. Hull's evaluation, however, has recently been challenged as based on no more than myth (Winsor, 2006; Schuh and Brower, 2009, p. 45). Moreover, Hull's supposed good news appears to be premature (*see* Conclusions).

Empirical and Inductive Taxonomic Methods

According to Mayr (1969), empiricism should be included as an atheoretical approach embracing the plausible, commonsense view that once sufficient knowledge has been gathered about organisms, a natural system of classification will simply emerge or become self-evident. Although it seems difficult to accord such a process the status of a general method, particular empirical observations can and often do render theoretical disputes irrelevant.

When Vaughan Thompson discovered in 1829 that barnacles develop from a nauplius larva, they were readily transferred from the Mollusca (where they had often been placed) to the Crustacea. Although this change in classification could in retrospect be justified by appeal to arguments about evolution, overall similarity, bauplans, or synapomorphies, to the empiricist, this would all appear unnecessary. To raise a brood of insects and find that all the males belong to one nominal genus and all the females to another leads to an instant appreciation of generic synonymy, without any need for appeal to theory.

The most complete expression of empirical classification occurs outside science in the form of folk taxonomy. Because folk taxonomy is well developed even in illiterate cultures, instead of inquiring into method or philosophy, it can only be understood by description and comparison. Berlin (1992) identified seven major features of ethnobiological categorization: recognition is given to the most distinctive local species; classification is based on affinities that humans observe (not on cultural significance); systems have a limited hierarchic structure; recognized taxa are distributed among a few mutually exclusive universal ranks, approximately equivalent to kingdom, life-form (e.g., plant), family, genus, species, and variety; in different ethnic systems, taxa at each rank show striking similarities regarding their number of subordinate taxa; taxa at generic and specific levels have an internal structure in which some included members are viewed as prototypical and others as reminiscent of other taxa at the same rank, giving rise to a type of fuzzy set classification; and a large majority of the taxa differentiated, most notably at the generic level, correspond to groups recognized in formal taxonomy.

These regularities and correspondences are highly suggestive of the fact that there is a natural biological classification that is almost literally self-evident and independent of the observer, and that a significant number of natural elements and groupings can be recognized simply through extensive knowledge and contact with nature. Although the young Linnaeus thought he was discovering the handiwork of God, it seems likely he was also involved in formalizing preexisting European ethnobiological knowledge.

Thus, as empirical knowledge accumulates, a common sense or consensus view of certain issues arises, and in some cases, such a view may seem irrefutable. Even so, there are many problems in taxonomy that cannot be decided (e.g., questions of relative rank or the status of paraphyletic groups) without an explicit theoretical framework.

Nominalism

One view of taxonomy is that, even though we may desire a general system, there are no independent means for assessment. The only objective realities are individual organisms and all taxonomic groups are man-made abstractions (“categories of thought” according to Louis Aggasiz). Named groups are convenient for collective reference but have no independent basis separate from the human mind, simply being useful pigeonholes for dividing up or handling diversity. Many biologists have held similar views – apparently including Darwin, who once commented on species as a “term … arbitrarily given for the sake of convenience to a set of individuals closely resembling each other” (see Mayr, 1963, p. 14; also Ghiselin, 1969).

By embracing the idea that taxonomic groups are established for convenience, nominalism has links with special classifications. If special classifications such as trees or four-footed land vertebrates are convenient, then there is no obvious reason why the nominalist should reject them, except insofar as it can be demonstrated that some comparable grouping has more heuristic or predictive power (e.g., trees extended to Tracheophyta and literal Tetrapoda extended to include snakes, birds, whales, and bats); in other words, they are in some sense more convenient. Convenient for who then becomes the question.

Berlin’s (1992) review suggests that folk taxonomies are not nominalist, which perhaps might have been expected, but seminatural and based on extensive empirical knowledge. Folk systems often have terms for conspicuous natural groups such as Mammalia, but it is not clear to what extent paraphyletic or even polyphyletic taxa are also included. Because folk taxonomies have terms for life-forms, which are in effect grades, this suggests that humans may possess an innate mixed strategy for classification, sometimes grouping on palpable characters, whereas at other times forming groups on the basis of general resemblance, Gestalt, or jizz.

Although Panchen (1992) suggests that there are no extant practitioners of nominalism, the view that species are real whereas all higher taxonomic groups are artificial is still quite widely expressed (even though quite misguided: contrast the difficulty of recognizing many species of mice with the ease with which the class Mammalia can be recognized by many consistent features; or the many dubious species of *Aster* recognized within the abundantly distinct Asteraceae: Stuessy, 2009). There are major disagreements regarding the existence of fundamental differences between terminal taxonomic components (e.g., species) and higher groups (e.g., polytypic genera and families), as accepted by Simpson (1961, p. 114) and Mayr Winsor (2006, p. 157), but rejected by Gilmour (1940) and Nelson (1989).

Evolutionary Systematics: Grades and Clades

The Darwinian revolution provided a rationale external to the human mind for the basis of taxonomy: hierarchical relationships among living things are explicable as the result of organic evolution, and taxonomists should strive to reflect these patterns as the logical basis of a general classification. However, if taxonomists looked to Darwin for a more detailed lead, what did they find? Unfortunately, at various places in the *Origin*, Darwin appeared to shift between general evolutionary statements (“the natural system is founded on descent with modification”), phylogenetic assertions (“the arrangement of groups within each class … must be strictly genealogical in order to be natural”), and nominalism (“I look at the term species, as one arbitrarily given for the sake of convenience”). Moreover, the preevolutionary taxonomic hierarchy had served the Darwinians well, presaging a continuing debate to this day: is the theory of evolution relevant to the pursuit of taxonomy only as justification for seeking a single general system or should we build into the taxonomic method theories about the way in which evolution has occurred?

Whatever the precise reason, the immediate impact of Darwinism on method was very limited. Taxonomy remained highly individualistic, and idealist, nominalist, empiricist, and even creationist views all continued, although often cloaked in evolutionary language. The first move to formulate a general method post-Darwin did not occur until the 1920s, later epitomized by two books by Julian Huxley: *The New Systematics* (1940) and *Evolution: the Modern Synthesis* (1942). This

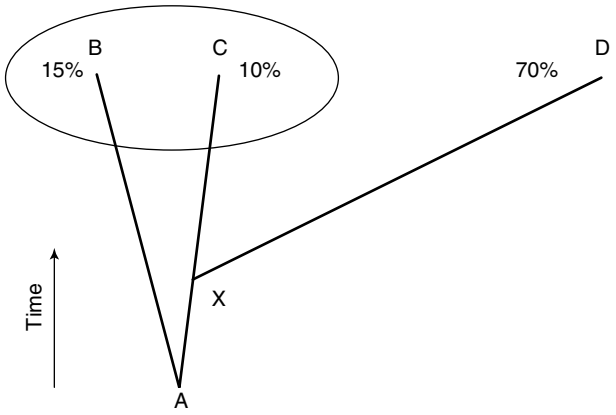


Figure 1 Evolutionary systematization. Horizontal axis, anagenesis; vertical axis, time. In this hypothetical case, the slope of each branch corresponds to the rate of anagenesis, with the percentage values representing the degrees of difference in the three lineages from the ancestral species (A). B is grouped with C because the two are more like their common ancestor than either is to D (even though D shares a common ancestor with C at X that it does not share with B). Note that this method requires a means of measuring similarity or distance (phenetics), estimating phylogenetic relationships (cladistics), and knowledge of ancestors (paleontology). Based on Mayr E (1974) Cladistic analysis or cladistic classification? *Zeitschrift für Zoologische Systematik und Evolutionsforschung* 12: 94–128, and Patterson C (1982) Cladistics and classification. *New Scientist* 29: 303–306.

emergent approach was developed and defended most notably by George Gaylord Simpson and Ernst Mayr, and it became evolutionary systematics or evolutionary taxonomy.

With respect to Darwin, the evolutionary systematists argued that the natural system (**Figure 1**) should reflect descent with modification, to include the processes by which evolutionary change occurs, the measurable degree of modification (anagenesis), and the temporal sequence of divergence (cladogenesis). According to Simpson (1961, p. 107), evolutionary systematics requires that a natural classification should be "consistent with all that can be learned of the phylogeny of the group classified," but he was emphatic that this does not mean that natural classification be based on phylogenetic relationships alone – a view also clearly stated by Gilmour in *The New Systematics* (1940). This led Simpson to adopt a very broad notion of monophyly as "the derivation of a taxon through one or more lineages … from one immediately ancestral taxon of the same or lower rank."

To be natural, according to Mayr (1969), a classification must have explanatory, predictive, and practical values but also be emendable in the light of new evidence or understanding. Mayr then proposed that the fundamental basis for naturalness is the proportion or the number of genes held in common by any two taxa: "If we knew the entire genotype of each organism, it would be possible to undertake a grouping of species that would accurately reflect their 'natural affinity'" (p. 81).

The empirical approach closest to this new ideal of evolutionary systematics is the use of pairwise distance data representing, in some way, the entire genome (e.g., immunological data and notably DNA–DNA hybridization data used by Sibley and co-workers for bird classification). In addition, complete nucleotide sequences for the entire genomes of various organisms are now becoming available for direct comparison (e.g., Chimpanzee Sequencing and Analysis Consortium, 2005). These approaches, however, founder on some of the fundamental problems of phenetics. Mayr's proposal that "genes-in-common" provides the ultimate arbiter of natural classification, however, is an important concept because it encapsulates the only well-articulated rival to the phylogenetic nexus idea first suggested by Darwin ("the arrangement of groups … must be strictly genealogical in order to be natural"). If lateral gene transfer undermines a strictly hierarchical approach, then the estimation of genes-in-common will certainly increase in importance as the basis of a natural system.

In practice, evolutionary systematics became a syncretistic, all-embracing method that included a regard for the absence of characters as informative and insisted on the primacy of paleontology for revealing phylogenetic sequences (**Figure 1**). The method's most striking characteristic involved conflation of the two methods that soon sought to replace it: a desire to give expression to genetic distances (grades or anagenesis, as reflected in phenetics) and, at the same time, to ancestor–descendant relationships (clades, or cladogenesis, as reflected in cladistics). Because these two methods stem from fundamentally different philosophies, this led to an inevitable arbitrariness (**Figure 1**). Thus, Simpson (1961, p. 107) was happy to write that although "taxonomy is a science … its application to classification involves a great deal of human contrivance and … there is a leeway for personal taste, even foibles." This lack of explicitness led to the demise of evolutionary systematics as the leading method in taxonomy, although it still has some active proponents (e.g., Stuessy, who terms this approach "phyletics"). During its development, however, Mayr in particular made an enormous contribution, especially to ideas on the taxonomy of species, with which he was preoccupied as the basic unit of classification (Mayr, 1963, p. 11).

Phenetics and Operationalism

Phenetics (as numerical taxonomy) emerged in the late 1950s, its origin associated with, among others, Charles Michener, Arthur Cain, and especially Robert Sokal and Peter Sneath. In Sokal and Sneath's (1963) *Principles of Numerical Taxonomy*, any evolutionary approach is avoided in favor of an operational method based on a direct comparison of phenotypes. As many characters as possible of the organisms to be compared, both continuous and discontinuous, are measured and counted from operational taxonomic units (OTUs), which can be individuals or samples of individuals from conventionally recognized taxa (typically species). On the basis of a matrix of variation in all features across all OTUs, the OTUs are then compared by overall similarity, affinity, or phenetic distance. Such measures can be obtained by transforming the raw matrix (**Table 1**) to yield the proportion of all character matches (affinity) or mismatches (distance: **Table 2**) for every pairwise combination of OTUs. The results are displayed by means of a network (**Figure 2(a)**) or OTUs are linked to each other by a clustering algorithm to produce a phenogram (**Figure 2(b)**).

Table 1 Amino acids found at eight particular positions in the myoglobin chain of four operational taxonomic units (OTUs): (A) Human, (B) Alligator, (C) Tuna fish, and (D) *Heterodontus* Shark[a]

OUT	*Characters*							
	0	*1*	*2*	*3*	*4*	*5*	*10*	*11*
A	–	G	L	S	D	G	V	L
B	M	E	L	S	D	Q	V	L
C	–	–	–	–	–	A	V	L
D		–	–	–	–	T	V	N

[a]Single letters stand for particular amino acids and dashes for positions not represented, based on alignment of the four complete myoglobin sequences.
Source: Reproduced from Patterson C (1980) Cladistics. *Biologist* 27: 234–240.

Table 2 Data of **Table 1** transformed into a distance matrix, based on the proportion of mismatches summed across all eight positions, for all six pairwise comparisons of the four OTUs[a]

	A	*B*	*C*	*D*
A	0	0.375	0.625	0.75
B		0	0.75	0.875
C			0	0.25
D				0

[a]These distances are replicated in the branch lengths of an unrooted network (**Figure 2(a)**). Taken as reciprocals to give measures of similarity, the data can also be used to produce a phenogram (**Figure 2(b)**).

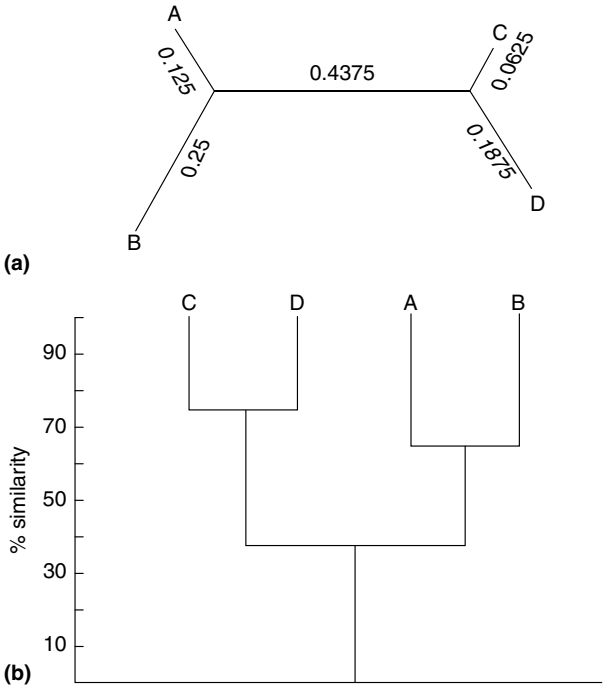


Figure 2 Phenetic systematization. (a) The values in **Table 2** between the four operational taxonomic units (OTUs) A–D represented by a distance network. The scaled lengths of the line segments are such that all the relative values in **Table 2** are satisfied, but the angles of the four branches leading to the terminals have no special meaning, (b) Horizontal axis, linkages; vertical axis, overall similarity. By taking reciprocals of the values in **Table 2**, the data yield a similarity matrix – the basis of a phenogram. With a value of 75% similarity, (C + D) are the first OTUs clustered. Comparing A with B, and (C + D) with A and B separately, (A + B) form the next most similar cluster, linked at 62.5% similarity. The two clusters can then be linked together at 37.5% similarity based on three of the eight attributes occurring in one or both members of the two groups. Other linking procedures could be adopted.

Although such a procedure appears objective at first, many different ways have been proposed to measure pairwise similarity or dissimilarity, and many different clustering methods have also been devised. Most clustering algorithms used in phenetics are sequential, agglomerative, hierarchic, and nonoverlapping (SAHN). Among this class of methods there are subclasses (e.g., single linkage, complete linkage, and average linkage), which in turn have variations (Sneath and Sokal, 1973). The phenograms produced by SAHN algorithms are dendrograms and thus strictly hierarchical, lacking reticulations, appearing similar to a Linnaean hierarchy or a cladogram.

Initially, it was expected that an agreement would be reached on best if not ideal procedures, but this proved to be an illusion. As early as 1966, Hennig (p. 85) commented that "even the most recent authors present and recommend their own methods (for measuring morphological differences) but never explain why their method deserves preference over those of their predecessors." Despite this difficulty, the pheneticists pursued their ideas with vigor based on four key hypotheses proposed by Sneath and Sokal that can be viewed as attempts to realize Mayr's goal of measuring genes-in-common (Panchen, 1992, p. 135). The nexus hypothesis stated that every character is likely to be affected by more than one gene, and every active gene is likely to affect more than one character. The nonspecificity hypothesis suggested that no major or distinct group of genes would exclusively affect one class of attributes (e.g., morphology and behavior); thus, classifications based on one character class or set should not differ from classifications based on another class or set, and there would be no a priori reason to prefer particular data sets to others. The hypothesis of the factor asymptote supposed that as data were added, the information content of the classification would reach an asymptote, whereas the hypothesis of the matches asymptote proposed that as the number of characters sampled increased, the similarity value for any pair of OTUs would tend toward a final, asymptotic value, thus conferring stability on the classificatory system.

Regrettably, for phenetics, as discussed by Panchen (1992) and even acknowledged by Sneath and Sokal as early as 1973, these hypotheses, apparently basic to the validity of the methods, are now largely discredited (in particular, the all-important nonspecificity hypothesis). More generally, phenetics offers no justification for choosing a hierarchical system unless it is accepted that the general theory of evolution dictates this as the most efficient representation. However, because phenetics conflates homologous and nonhomologous characters, and neither nodes nor branch lengths can be directly related to hypotheses about particular characters, the origin of characters, or even the degree of genetic change, the notion of overall similarity lacks analytical power.

Moreover, although it was originally expected that its procedures would lead to stability in classification, several

factors preclude this (Eldredge and Cracraft, 1980, p. 176). In addition to the technical ambiguities already noted, every new taxon requires complete reanalysis, with the unique attributes and combinations of attributes usually affecting every branch length and often many branching points in the phenogram. Of course, other taxonomic methods, including cladistics, are not invulnerable to change due to new discoveries, but in cladistic analyses, the effects are interesting (because they can be related to hypotheses about homology) and cladists never laid serious claim to the idea that stability in the face of new evidence was an important justification for the method.

Cladistics also has its own algorithmic problems, but it is always possible to work out what different procedures are doing with respect to the data and the inferences drawn. With phenetic methods, the nature of the algorithms and the form of the results are not separable. Thus, a basic problem of numerical phenetics is the lack of a clear criterion of choice (such as parsimony or even a model of evolution) by which one result can be judged against another ("the interpretation of evidence:" Schuh and Brower, 2009, p. 41). Even Mayr's concept of genes-in-common will not repair the difficulty because of innumerable problems related to the definition of a gene, duplication, alignment, position effects, and so on. The only other external arbiter is the indefinable goal that James Farris called Gilmour-naturalness ("a system of classification is the more natural the more propositions can be made regarding its constituent classes"), enthusiastically embraced by Sneath and Sokal (1973). This led Panchen (1992) to suggest that the original philosophy of numerical phenetics is closer to empiricism (not nominalism, as suggested by Mayr). Given enough observations and an appropriate technique, the numerical taxonomists expected that a single, stable, and predictive general system would simply emerge.

Although phenetics has been judged wanting, it has had a lasting and positive influence on current taxonomic methods and still flourishes in areas such as bacterial taxonomy, where the general absence of tokogenetic relationships negates many advantages of a phylogenetic system. The most important contributions of phenetics have been technical, notably the use of data matrices, precise comparison across all taxa under scrutiny, and the adoption of algorithmic analytical methods. Advances in other methods, notably cladistics, have emerged not only from vigorous debate but also from the general adoption, post-Hennig, of data matrices and their exploration through other types of numerical algorithms.

Phylogenetic Systematics and Cladistics

The fundamental methods of phylogenetic systematics were established by Hennig, who wanted to base systematization and classification directly on the historical branching patterns of the phylogenetic nexus. He realized that it would never be possible to know the course of history precisely, and that the system would always be provisional. His goal was, by means of appropriate methods, to produce a phylogenetic system that would "approximate more closely than any other *the ideal system* (italics added) that reflects the phylogenetic relationships absolutely correctly" (Hennig, 1966, p. 29).

Table 3 Data from **Table 1** transformed to show only coincidences of positive attributes among the four OTUs[a]

Taxa	*Character*			
	2	*3*	*4*	*11*
A	*	*	*	*
B	*	*	*	*
C				*
D				

[a]Characters 2–4 link (A + B), and character (11) links (A + B + C). *denotes characters present. These data are most efficiently represented by the nested-set pattern (D(C(A + B))); compare with the pattern ((D + C)(A + B)) implied by the phenetic analysis (**Figure 2**). In this analysis, four characters (0, 1, 5, 10) in **Table 1** are uninformative because there is no positive coincidence (5), only shared absence (0), both (1), or only shared coincidence (10).

Hennig's method was the search for the sister group, epitomized as the three-taxon problem. For any group of three natural taxa, the expectation is that two have a common ancestor not shared by the third. Thus, for the trio shark, tuna, and human, of the three possible combinations of two from three, the one with tuna and human grouped to the exclusion of shark accords with what we know of ancestry (see **Table 3**). According to Hennig's view, we should then recognize a taxon linking tuna and human as sister groups (Osteichthyes) but not one linking shark and tuna: from a phylogenetic perspective, Pisces (fish) are a paraphyletic or non-group. Thus, Hennig defined phylogenetic relationships solely in terms of common ancestry (propinquity of descent) and did not include similarity, distance, grade, adaptive zone, or any comparable concepts in his assessments.

This line of argument caused Mayr and other evolutionary systematists to object very strongly to Hennig's approach, accusing him of an unjustifiable restriction of the concept of relationship. For this reason, Mayr referred to Hennig and his followers as cladists (from clade or cladogenesis) to criticize them for their narrow approach – with the unexpected result that the term was happily adopted by the growing band he sought to reprove (even to founding a journal with the title *Cladistics*).

However, how could clades be recognized in the absence of prior knowledge or without acceptance of sequences simply read from the fossil record? Arguably, Hennig's greatest methodological advance was the introduction of relativism to the concept of characters. He recognized three classes with respect to a set of taxa under study: autapomorphies (characters unique to individual taxa), symplesiomorphies (characters present in all the taxa), and synapomorphies (characters shared by subsets of two or more taxa). According to Hennig, only characters of the third class provide evidence of grouping at the specified level of inclusiveness. Thus, a synapomorphy indicative of a natural group of species within the taxa under review (e.g., a genus) becomes a symplesiomorphy when addressing questions of relationship among species within that genus so delimited. Ideally, each and every character would be informative of relationships at just one level in the total hierarchy. Thus, virtually all the characters that might seem to unite shark and tuna to the exclusion of human (e.g., gills and fins) are actually characters that relate to more inclusive levels

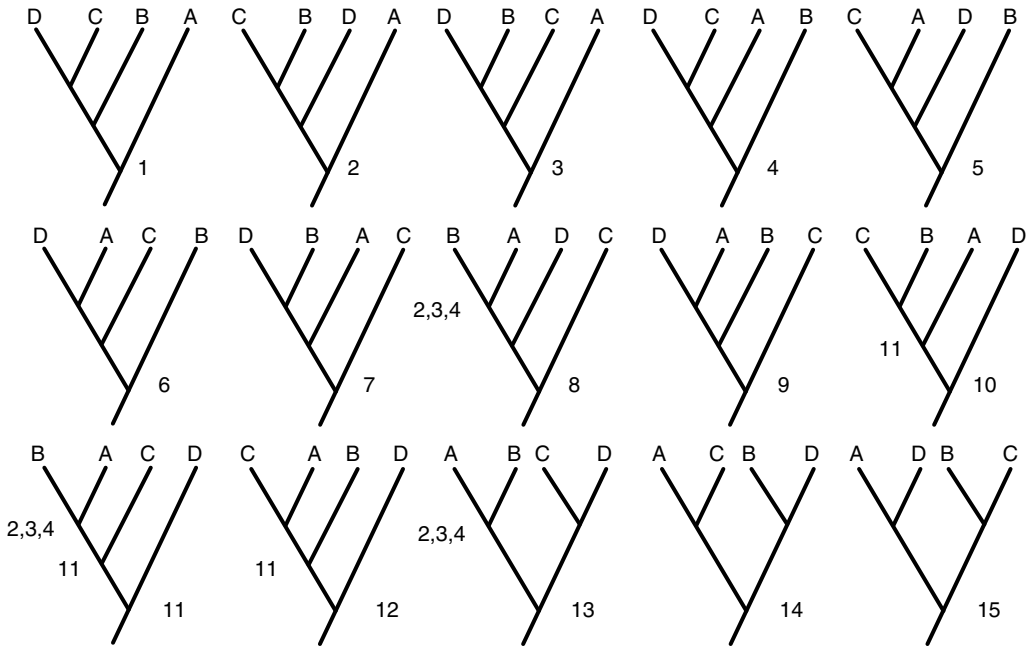


Figure 3 Cladistic systematization. Horizontal axis, convention; nodes, pattern of homologous characters. The putative synapomorphies in **Table 3** plotted on all 15 possible arrangements (as fully resolved cladograms) for four terminal taxa. Numbers at given nodes indicate the shared attributes from **Table 3**; arrangement 11 is the only one that reflects all four attributes, all of which can then be construed as homologous characters. Based on Patterson C (1980) Cladistics. *Biologist* 27: 234–240.

of the hierarchy (e.g., Chordata and Vertebrata). Observations such as absence of mammary glands or absence of hair are not characters linking fish but the counter conditions of autapomophies (of the Mammalia).

To see this in practice, the data in **Table 1** have been transformed (**Table 3**) to show only potential synapomorphies, i.e., the positive attributes shared by two or three of the four taxa, and thus able to provide evidence of grouping within (A–D). In this simple example, there are no conflicting distributions, and there is just one maximally efficient solution. Efficiency means the hierarchical arrangement that captures all (or, where there is conflict, the largest number) informative characters (putative synapomorphies). In **Figure 3**, all 15 possible fully resolved groupings of terminal taxa A–D are presented, with the characters that they can summarize marked at the relevant nodes. Arrangements 1–7, 9, 14, and 15 do not reflect any of the characters. Arrangements 10 and 12 capture one, and 8 and 13 capture three, but only arrangement 11 captures all four. The best cladistic arrangement, on the available data, is therefore (((A + B)C)D).

Patterson (1980) summarized the axioms of cladistics as homologous characters having a hierarchical pattern in nature in which this pattern is efficiently expressed by cladograms, with the nodes connoting homologies shared by the organisms so grouped. The search for the sister group therefore reduces to finding the cladogram that summarizes the potentially homologous characters as parsimoniously as possible. Once this best-fit cladogram has been found, the attributes at each node are hypothesized as homologous characters shared by the taxa subtended at that node.

One of the most challenging features of cladistics (or at least transformed cladistics – *see* Current Practice: Variations on a Cladistic Theme) is the link between attributes, grouping, parsimony, characters, and homology. An attribute is accepted as a character if the weight of evidence from all other attributes under scrutiny suggests that it is a homologous feature peculiar to a particular group. In **Table 4**, two potential synapomorphies (x and y) have been added to the distribution of characters in **Table 3**. Given these additional data, and presupposing the independence of all the attributes, the most efficient grouping of the four taxa based on positive shared features now changes and highlights a problem with the interpretation of attribute 11: either it is not relevant as a character within the group and has undergone reversal in taxon D or it is not a character at all (it is nonhomologous).

Table 4 Imaginary character matrix for four taxa and six characters[a]

Taxa	*Character*					
	2	*3*	*4*	*11*	*x*	*y*
A	*	*	*	*		
B	*	*	*	*		
C				*	*	*
D					*	*

[a]Characters 2–4 and 11 correspond to **Table 3**. *denotes characters present. Additional characters x and y are in conflict with character 11. The most parsimonious solution, assuming that all characters are independent, is to group the four taxa as ((A + B)(C + D)). This implies that attribute 11 is no longer considered homologous (not the same in C as in (A + B)) or that it is symplesiomorphous with respect to the entire group (A + B + C + D) and does not appear in taxon D (e.g., in evolutionary terms, this would imply that character 11 had been secondarily lost in taxon D).

Information about this could be sought, for example, by studying the ontogeny or some more fundamental quality of attribute 11 in A–C, revealing a particular scientific strength of the cladistic method: conflicts can be resolved by character analysis (Kitching *et al.*, 1998) involving recourse to additional data derived from previously unsampled or unused attributes or by investigation into the homology of the conflicting characters (they are not simply aggregated at face value, as in phenetics).

If characters x and y in **Table 4** were found to be due to the pleiotropic effects of a single gene, then rejection of attribute 11 as a character relevant to resolving the relationships within the group would be premature. If, however, character 11 was a certain color produced, for example, by a pigment shared by taxa A and B, but the same hue was produced by a totally different type of pigment in taxon C, this would be consistent with the idea that attribute 11 as originally formulated (color) was a noncharacter. The similarities and differences between taxa A and C in this regard should then be rescored as two new attributes (e.g., based on chemistry), lending further support to the (A + B) grouping and removing conflict with the (C + D) grouping. The convergence in color between (A + B) and C could then be seen to be in the eye of the human observer. Such convergences are of intense interest for the study of evolution, but once detected, they play no part in the process of natural classification.

Thus, the relativistic nature of cladistics is emphasized. Based on the entire (attribute × taxon) matrix, the cladograms chosen and attributes thereby specified as homologous characters are probability statements: "Common ancestry versus convergence is tested by topographical correspondence (on the cladogram). The resulting explanation is a statement of maximal likelihood rather than a denotion of lawful relations or processes" (Rieppel, 1988, p. 166; see Hennig, 1966, p. 29).

General Procedures

The main operations involved in taxonomic research are listed as follows in an idealized order. In practice, this order is rarely followed precisely, and certain steps may be omitted in part or even altogether:

1. Individual specimens, or individuals representing selected taxa (ideally at least all immediately subordinate taxa currently recognized within the group under study), including material representing suitable outgroups for comparison, are chosen.
2. A selection of attributes to be scored across all samples of taxa under study is made (in original work, this will usually be a mixture of known and novel features).
3. The manifestations of these attributes are systematically recorded in an individual or a taxon/attribute data matrix (in the process of doing this, it is normal for other potentially informative features to be recognized and these may be added to the list under stage 2; it is not unusual at this stage for additional study material also to be called for, when variation has become apparent or specimens prove to be incomplete or otherwise unsuitable).
4. Systematization of the data is carried out by cladistic analysis to produce a cladogram based on homologous characters defining each node (or, for phenetic analysis, by an agglomerative numerical technique to produce a phenogram); in cladistic studies, this may involve one or more rounds of Hennigian reciprocal illumination, notably with respect to testing the coherence of putative homologies that seem to be in conflict (incongruent).
5. Inclusive monophyletic groups inferred from the parsimonious distribution of homologous characters are thus linked together and then named as taxa, to produce a hierarchical classification.
6. The homologous characters are used to help develop diagnoses for the various taxa recognized (as a potential aid for identification, information on characters that are absent may also be included, and keys, tables, or computerized identification programs may also be constructed at this point).
7. Publication of the results will ideally include the provision of necessary descriptions, indications (diagnoses), specifications of type specimens, and formulation of names in accordance with the relevant code (e.g., zoological, botanical, and bacteriological codes) as well as a record of the complete character matrix and specification of the analytical procedures used (e.g., computer programs and options and models in the case of maximum likelihood analyses).

Cladistic and numerical phenetic methods are essentially the same up to stage 3, but a crucial difference occurs at stage 4, the process of systematization. Stages 5 and 6 also differ, but both methods involve some arbitrariness with respect to where the phylogenetic nexus is to be cut or the phenogram to be divided (*see* From System to Classification). At stage 6, there is also some difference in that the very act of cladistic analysis gives the primary characters for diagnosis, whereas the numericist is faced with an additional searching process to enumerate a sufficient set (usually a polythetic set) of characters, whereby given members of a specified portion of the phenogram can be recognized in isolation. Fundamental differences in approaches to systematization have already been discussed (*see* Differing Philosophies and Methods of Taxonomy). The problems inherent in stage 5 are discussed next.

From System to Classification

Classification involves translating a systematization scheme into words (or numbers). Based on the methods of logical division, Linnaeus (who established many of the conventions of formal classification) placed all known organisms within a descending series of fully nested hierarchical categories. The major ranks in the Linnean system were kingdom (e.g., plants and animals), class (e.g., Mammalia, Aves, Pisces, and Insecta), order (e.g., Hemiptera, Lepidoptera, Coleoptera, and Diptera), genus (e.g., *Culex*, *Tipula*, and *Musca*), and species (e.g., *Musca domestica* (housefly) and *Musca vomitoria* (bluebottle)). For Linnaeus, systematization and classification were the same. Modern classifications attempt to summarize far more complex schemes. In this brief review, only the major differences between the major twentieth-century methods are considered.

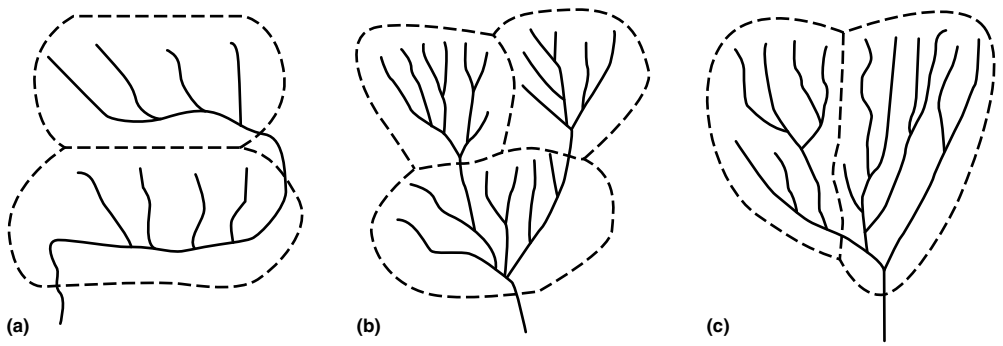


Figure 4 Evolutionary classification. The original caption reads: "Three hypothetical phylogenies … divided into taxa as shown by the broken lines. A is evidently less arbitrary … than B, and B somewhat less than C." Redrawn from Figure 7 in Simpson GG (1961) *Principles of Animal Taxonomy*. New York: Columbia University Press.

According to the evolutionary method, classification should be viewed as a useful art, meaning that in addition to scientific analysis, human ingenuity is also needed to produce a practical classification. Simpson (1961) recognized three principles: a classification should reflect the most biologically significant relationships among the organisms, it should be consistent with the relationships on which it is based, and it should be as stable as possible without contravening the first two principles. These ideas were elaborated to include, most notably, grades and clades (anagenesis and cladogenesis), both of which were regarded as significant for classification. This led to the idea that, even if literally everything relevant were known about the connections among the members of a group, "innumerable different classifications could be made consistent with those interrelationships … Selection among those alternatives is decidedly an art" (Simpson, 1961, p. 110).

To appreciate the difficulties resulting from such a view, consider **Figure 4**. In cladistic terms, three paraphyletic groups (basal groups in A–C) are delimited, together with just four of the 34 subclades depicted. Why were these particular groupings chosen? Such a mysterious art proved difficult to follow (especially when the prescription for creating such trees in the first place is also imprecise). As discussed previously (*see* Evolutionary Systematics: Grades and Clades), Mayr urged that classifications be based on genes-in-common. Because it is a distance method, this transformation in evolutionary systematics encounters the same problems of classification faced by phenetics. Recently, the demonstration that the human and chimpanzee genomes are extremely similar (Chimpanzee Sequencing and Analysis Consortium, 2005) has led many mammalogists and primatologists to abandon the paraphyletic or grade family Pongidae in favor of an expanded and strictly monophyletic Hominidae, thus abandoning the position adopted for decades by the evolutionary systematists, who argued repeatedly that the large gap between *Homo* and the rest of the great apes was reason enough to place them in separate families (Simpson, 1961; Mayr, 1963). Even so, in this particular case and despite the small genetic or genomic distance, the highly autapomorphic nature of the human condition continues to give rise in some to misgivings about classifying *Homo sapiens* as just another ape (Guldberg, 2010) whereas, more generally, a few taxonomists continue to call for the formal recognition of paraphyletic groupings (e.g., Hörandl and Stuessy, 2010).

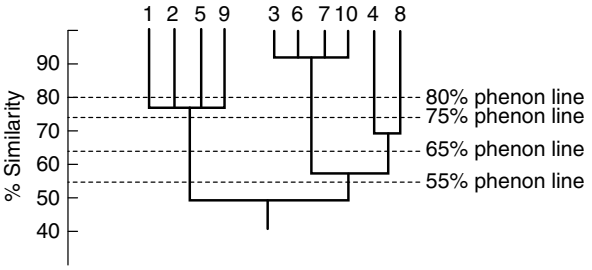


Figure 5 Phenetic classification. The formation of phenons from a phenogram. Redrawn from Figure 5.33 in Sneath PHA and Sokal RR (1973) *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. San Francisco: Freeman.

Sneath and Sokal (1973) proposed that good classifications have three desirable properties: naturalness, ease of manipulation, and practicality for information retrieval. Their concept of naturalness, as already discussed, was based on the ideal of incorporating all codified information about a very large numbers of characters. Manipulation related to the practical relationship between a classification and its degree of hierarchical structure (useful for memorizing) and its utility (e.g., for the construction of identification keys). Convenience for information retrieval was also viewed as important, but not if it conflicted with natural classification.

The typical product of systematization by phenetic methods is the phenogram (Schuh and Brower, 2009, p. 12). How is such a dendrogram turned into a summary classification of named groups? An early proposal was the use of phenon lines (**Figure 5**). A phenon line cuts across the phenogram at a particular similarity level. The lines must be straight and are not allowed to "bend up and down according to … whim" (Sneath and Sokal, 1973), not only delimiting groups but also determining their rank. Thus, in **Figure 5**, if OTUs 1–10 are species, the 80% phenon line could indicate seven subgenera, the 75% line four genera, the 65% line three subfamilies, and so on.

Such a procedure is not objective. The most fundamental problem is simply that the choices of percentage similarity levels and ranks are arbitrary, both within groups and between them. Other procedures, such as McNeill's method based on the assignation of rank levels to each node, also involve arbitrary limits and adjustments. The difficulty for phenetics is

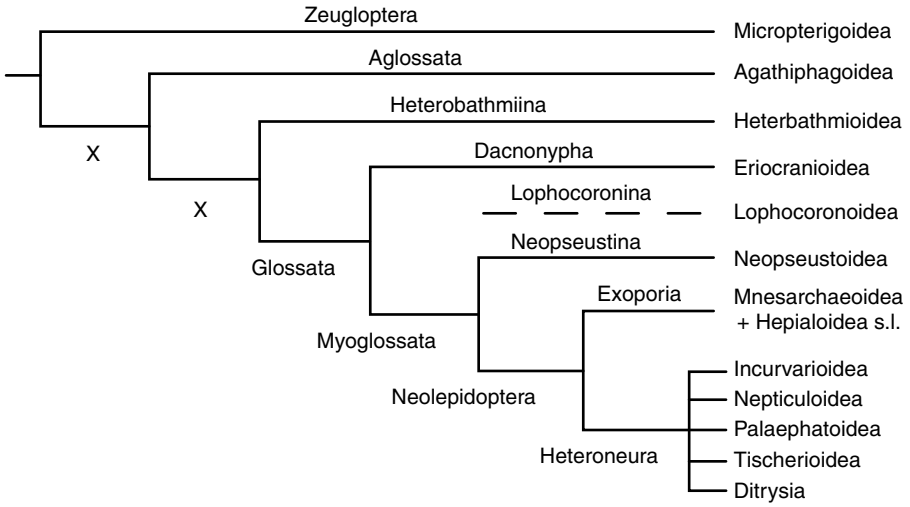


Figure 6 Cladistic classification. Labeled cladogram of the main lineages of the order Lepidoptera (slightly modified from Scoble (1992), Figure 185). In this classification, two of the major monophyletic groups recognized are not named (X) and the ranks of the terminals are noncoordinate. The dashed line to the Lophocoronoidea indicates that the relationships of this group are uncertain: they may form the sister group of the Exoporia, the Neolepidoptera, or the Myoglossata (Scoble, p. 203). Noncoordinate terminals and unnamed major lineages also appear in a recent phylogenetic scheme for the Ditrysia. Reproduced from Figure 3 in Regier JC, Zwick A, Cummings MP, *et al.* (2009) Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): An initial molecular study. *BMC Evolutionary Biology* 9: 280. doi:10.1186/1471-2148-9-280.

that "there is no necessary structural relationship between the pattern of clusters in hyperspace and the inclusive Linnean hierarchy into which they are converted" (Panchen, 1992, p. 151). If this is so, then there can be no nonarbitrary method for converting a phenogram (or any other distance-based method of systematization, including DNA–DNA hybridization trees) into a ranked summary classification.

Cladistic classification is based on the tenet that every putative monophyletic group can be named, but polyphyletic and paraphyletic groups should not. In a grand, top–down cladistic classification of all life, all coordinate monophyletic taxa (sister groups) could be placed at the same rank. Alternatively, taxa of the same geological age could be given the same rank (Hennig, 1966). Or classifications could be formed bottom upwards by linking all terminal sister species starting with the lowest-ranking pair or pairs and building up classes and ranks based on sister groups and ascending ranks held coordinate across the entire hierarchy. Such grand visions remain almost wholly impractical due to our very imperfect knowledge. Moreover, such systems would cause massive proliferation of named groups and ranks and be unstable in the face of most newly discovered taxa and every older geological find. In terms of producing a useful summary, cladistic classification thus faces potential problems of impracticality as well as arbitrariness.

In practice, when classifications are based on cladograms an attempt is usually made, within conventional ranks (e.g., order, family, and genus), to give expression to major monophyletic groups by naming inclusive taxa at intermediate levels. However, this usually means abandoning any equation of classificatory rank with clade level and any attempt to give all groups in the cladogram formal recognition. Thus, in a labeled cladogram of the main lineages of the order Lepidoptera (**Figure 6**), not all putative monophyletic groups are named and all terminals (despite being almost completely noncoordinate) are classified at the same superfamily rank, except one of the cladistically most inferior groups, the crown group Ditrysia, which retains its higher, subordinal status accorded in earlier, precladistic classifications!

One partial solution to these difficulties is the sequencing convention, first proposed by Nelson. This attempts to combine the practicalities of a limited Linnean hierarchy with a listing that enables the entire (cladistic) hierarchy to be recovered. A sequence of taxa named at the same rank indicates that the first is sister to all following at that rank, the next is sister to the remainder, and so on. Taxa of uncertain position can be annotated as *incertae sedis* and those comprising unresolved polytomies as *sedis mutabilis*. In presenting these sequences, comprehension is greatly facilitated by indentation. Thus, the cladogram for the major groups of Lepidoptera in **Figure 6** can be converted into a written classification (**Table 5**).

In conclusion, the idea that classification can simply be equated with systematization is a vestige of preevolutionary taxonomy and should be abandoned (Minelli, 1993, p. 14). In practice, reflecting a Simpsonian view, classifications are made consistent, as far as possible, with current systematization. Named para and polyphyletic assemblages are increasingly rejected whenever, *pace* Hörandl and Stuessy (2010), strong evidence of nonmonophyly becomes apparent. Such groups can be very misleading for comparative biology, biogeography, ecology, and extinction studies (although dicots, fish, reptiles, and invertebrates seem impossible to eliminate). Thus, dinosaurs are only extinct if we insist that birds are not part of the group – in other words, their extinction is a matter of semantics rather than biology if, by excluding the birds, we continue to treat a paraphyletic Dinosauria as if it were a natural entity.

Table 5 Sequenced, tabular classification of the Lepidoptera, reflecting Scoble's cladogram (**Figure 4**)

Order Lepidoptera
Suborder Zeugloptera
Suborder Aglossata
Suborder Heterobathmiina
Suborder Dacnonypha
Suborder Lophocoronina incertae sedis
Suborder Neopseustina
Suborder Exoporia
Infraorder Mnesarchaeoidea
Infraorder Hepialoidea
Suborder Heteroneura
Infraorder Incurvariioidea sedis mutabilis
Infraorder Nepticuloidea sedis mutabilis
Infraorder Palaephatoidea sedis mutabilis
Infraorder Tischerioidea sedis mutabilis
Infraorder Ditrysia sedis mutabilis

Restraint is also necessary to avoid unstable changes to the formal classification arising from partial or inconclusive evidence (careful use of consensus trees may be helpful here) and overelaborate nomenclature for the numerous hierarchical levels. With the real prospect of international consensus systems covering all major groups of organisms being developed on the Internet, a balance between stability and universally accessible periodic revision to reflect fundamental advances can be anticipated. Such IT-based bioinformatic systems should also solve most problems of information retrieval, including the ever-present difficulties of alternative classifications, synonymy, and misidentifications (the last being the most pernicious of all taxonomic problems affecting information retrieval).

Current Practice: Variations on a Cladistic Theme

Phenetics was afflicted by a proliferation of clustering algorithms, each tending to yield different results for which no logical criterion of choice was available. The vaunted objectivity of the approach dissolved, leaving the field of systematization open to the apparently more consistent and decisive methods of phylogenetic systematics. However, matters are rarely this simple.

In the 1970s, a partial schism opened between those phylogeneticists who followed directly in Hennig's footsteps, approaching character transformation from an evolutionary perspective, and the so-called transformed cladists (Patterson, 1982), who held that particular theories of evolution were unnecessary for cladistics, adopting instead a taxic approach to cladistic analysis (Kitching *et al*., 1998). The transformed cladists embraced computer algorithms, character matrixes, and global parsimony, leaving polarization of characters to the algorithms and outgroup choice. For Hennig, the nodes in a cladogram were hypothetical ancestors; for the transformed cladists, the nodes indicate the hierarchical pattern of homologous characters. These differences have emerged in debate over congruence versus total evidence methods, now treated formally as partitioned versus simultaneous data analysis. Because many cladists claim to have demonstrated the superiority of simultaneous analysis for making cladograms, this debate links in turn to a seemingly fundamental disagreement over whether or not to use Fisher's maximum likelihood statistics in phylogeny reconstruction, as first proposed by Edwards and Cavalli-Sforza and later implemented by Joseph Felsenstein as a general method (Huelsenbeck and Crandall, 1997). For a recent account and critique, see Schuh and Brower (2009, pp. 135–170).

However, the real source of these disagreements seems to be one of different scientific agendas: those in favor of model-based likelihood methods and partitioned data sets are mainly seeking insights into evolutionary processes, most notably those affecting molecular evolution within different lineages or under different constraints. Those rejecting likelihood models in favor of global parsimony and simultaneous analysis of heterogeneous data sets are primarily concerned with tests of homology and the construction of cladograms (but see Wiens, 1999). Both approaches have different strengths and weaknesses, and both make simplifying assumptions in pursuing these different goals (Kitching *et al*., 1998, p. 165). Because most work affecting taxonomy requires the use of morphological data, or a combination of morphological and molecular data, for which plausible process models cannot readily be elaborated for simultaneous analysis using likelihood, the significance of likelihood methods for taxonomy per se (rather than understanding phylogenetic processes; Huelsenbeck and Crandall, 1997) remains to be demonstrated and is clearly rejected by some (e.g., Schuh and Brower, 2009).

However, with respect to the reconstruction of phylogeny there are several other methods, global parsimony and maximum likelihood being only the main contenders. These alternatives include James Lake's rate-invariant method for molecular data and pairwise distances (e.g., DNA–DNA hybridization). Even within the parsimony techniques generally adopted for cladogram construction, at least one fundamentally different challenge to global parsimony has emerged: three-item statement analysis, first proposed by Nelson and Platnick. Three-item analysis discards any presupposition that characters undergo evolutionary transformation and, by decomposing a taxon $\times$ attributes matrix into an analysis of all possible three-taxon statements supportable by the data, focuses entirely on evidence for sister group relationships rather than character transformations. Although not widely acclaimed or used, this method is operational and often yields more or less different results from those of conventional parsimony methods, thus raising some fundamental issues. For brief but informative discussions about this debate, see Kitching *et al*. (1998, Chapter 9) and Schuh and Brower (2009, pp. 133–135).

Conclusions

Taxonomy involves the search for a general pattern of order among living and extinct organisms, from which a general reference system (or classification) is derived. As such, taxonomy is the primary discipline of biodiversity. Fundamental philosophical disagreements about the nature of human

knowledge and reality have given rise to several basic methods. All recent approaches agree that systematics is empirical and should be based largely (evolutionary systematics and maximum likelihood molecular systematics) or entirely (phenetics and cladistics) on the comparison of data in the form of characters observed and abstracted from specimens. The methods differ somewhat with respect to the type of data preferred, and fundamentally with regard to methods of analysis used to reveal pattern, and how the system, once obtained, is translated into a written classification.

Cladistics has become the dominant method of systematization (where a preferred method is made explicit). The research program of cladistics is based on the view that the ideal general system should reflect the phylogenetic nexus. However, the true nexus cannot be known. If the precise course of evolution, taking account of all the relevant details of every lineage, is real but unknowable, then the natural system cannot be discovered but only approximated by an indirect process of estimation. Because cladistic classification can only admit monophyletic groups (based on the selection of a particular cladogram), all admissible groups are necessarily defined because they must have at least one unique (presumptively) homologous character (or a unique loss within a more inclusive group that is specified by some other homologous feature). If so, as suggested by both Rieppel (1988, p. 25) and Panchen (1992), the cladistic method is essentialist because the search for the sister group depends on the discovery of locally unique and thus defining shared characteristics for each and every group. Thus, Schuh and Brower (2009, p. 258) define apomorphy as a group-defining feature. Moreover, it has long been understood that although evolution is not bound by parsimony, scientific method admits nothing better than parsimony as the criterion of choice among competing hypotheses. Insofar as the divergent model of evolutionary history that justifies searching for a hierarchical pattern is flawed (Doolittle, 1999), the methods of cladistics may need further modification, notably to deal more adequately with hybrid origins and lateral gene transfer.

In conclusion, we may have to accept the seeming paradox that although the principle to which cladistics aspires is natural (i.e., the groups we seek to recognize exist regardless of human perception, and are thus to be discovered rather than defined; Rieppel, 1988, p. 163), its empirical methods appear in some sense to be inescapably essentialist (Panchen, 1992, p. 344). The basic method of taxonomy may be equivalent to the construction of a universal, never to be experienced directly but derived from the sensory experience of particulars (samples of individuals and characters) by means of an unending iterative sequence of analysis, hypothesis formation, testing, and reanalysis. This is a manifestation of the two ways of seeing explored by Rieppel (1988): the world of being owes its form to the processes of becoming, but those evolutionary processes can only be inferred indirectly from the patterns we observe in nature.

Finally, it is notable that approaches to the measurement of biodiversity that seek to maximize the number of expressible genes held in networks of germplasm banks or functional ecosystems depend on models incorporating information about branching sequences (cladistic relationships) and branch lengths (anagenesis). In practice, the taxonomic hierarchy and the raw data matrix (before removal of homoplasy and autapomorphies) from which it has been derived will often be the best or the only available data for use in such models. For a review of this field, see Crozier (1997); see also Humphries (2006).

See also: Cladistics. Cladogenesis. Differentiation. Evolution, Theory of. Genes, Description of. Genetic Diversity. Nomenclature, Systems of. Nucleic Acid Biodiversity: Rewriting DNA and RNA in Diverse Organisms. Phenotype, a Historical Perspective. Phylogeny. Speciation, Theories of. Species, Concepts of

References

Berlin B (1992) *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton, NJ: Princeton University Press.

Cain AJ (1954) *Animal Species and Their Evolution*, p 11. London: Hutchinson.

Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437: 69–87.http://dx.doi.org/10.1038/nature04072.

Crozier RH (1997) Preserving the information content of species: Genetic diversity, phylogeny, and conservation worth. *Annual Review of Ecology and Systematics* 28: 243–268.

Doolittle WF (1999) Phylogenetic classification and the universal tree. *Science* 284: 2124–2128.

Eldredge N and Cracraft J (1980) *Phylogenetic Patterns and the Evolutionary Process*. New York: Columbia University Press.

Foster PG, Cox CJ, and Embley TM (2009) The primary divisions of life: A phylogenomic approach employing composition-heterogeneous methods. *Philosophical Transactions of the Royal Society* 364: 2197–2207.http://dx.doi.org/10.1098/rstb.2009.0034.

Funk DJ and Omland KE (2003) Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* 34: 397–423.

Ghiselin MT (1969) *The Triumph of the Darwinian Method*. Berkeley: University of California Press. (1984 reprint by University of Chicago Press, Chicago).

Gilmour JSL (1940) Taxonomy and philosophy. In: Huxley J (ed.) *The New Systematics*, pp. 461–474. Oxford: Oxford University Press.

Guldberg H (2010) *Just Another Ape?*. Exeter: Imprint Academic.

Hennig W (1966) *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Hörandl E and Stuessy TF (2010) Paraphyletic groups as natural units of biological classification. *Taxon* 59(6): 1641–1653.

Huelsenbeck JP and Crandall KA (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics* 28: 437–466.

Hull DL (1965) The effect of essentialism on taxonomy – two thousand years of stasis. *The British Journal for the Philosophy of Science* 15: 314–326, 16: 1–18.

Humphries CJ (2006) Measuring diversity. In: Leadlay E and Jury SL (eds.) *Taxonomy and Conservation*, pp. 141–161. Cambridge: Cambridge University Press.

Kitching IJ, Forey PL, Humphries CJ, and Williams DM (1998) *Cladistics*, 2nd edn. Oxford: Oxford University Press.

MacLeod N and Forey PL (2002) *Morphometrics, Shape and Phylogenetics*. London: Taylor & Francis.

Margulis L and Sagan D (1984) Evolutionary origins of sex. *Oxford Surveys in Evolutionary Biology* 1: 16–47.

Mayr E (1963) *Animal Species and Evolution*. Cambridge, MA: Harvard University Press.

Mayr E (1969) *Principles of Systematic Zoology*. New York: McGraw-Hill.

Mayr E (1974) Cladistic analysis or cladistic classification? *Zeitschrift für Zoologische Systematik und Evolutionsforschung* 12: 94–128.

Minelli A (1993) *Biological Systematics. The State of the Art*. London: Chapman & Hall.

Nelson G (1989) Species and taxa: Systematics and evolution. In: Otte D and Endler JA (eds.) *Speciation and its Consequences*, pp. 60–81. Sunderland, MA: Sinauer.

Panchen AL (1992) *Classification, Evolution, and the Nature of Biology*. Cambridge, UK: Cambridge University Press.

Patterson C (1980) Cladistics. *Biologist* 27: 234–240.

Patterson C (1982) Cladistics and classification. *New Scientist* 29: 303–306.

Pimentel RA and Riggins R (1987) The nature of cladistic data. *Cladistics* 3: 201–209.

Rieppel OC (1988) *Fundamentals of Comparative Biology*. Basel: Birkhäuser Verlag.

Regier JC, Zwick A, Cummings MP, *et al.* (2009) Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): An initial molecular study. *BMC Evolutionary Biology* 9: 280.http://dx.doi.org/10.1186/1471-2148-9-280.

Schuh RT and Brower AVZ (2009) *Biological Systematics: Principles and Applications*, 2nd edn. Ithaca: Cornell University Press.

Scoble M (1992) *The Lepidoptera*. Oxford: Oxford University Press.

Simpson GG (1961) *Principles of Animal Taxonomy*. New York: Columbia University Press.

Sneath PHA and Sokal RR (1973) *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. San Francisco: Freeman.

Sokal RR and Sneath PHA (1963) *Principles of Numerical Taxonomy*. San Francisco: Freeman.

Strong EE and Lipscomb D (1999) Character coding and inapplicable data. *Cladistics* 15: 327–362.

Stuessy TF (2009) *Plant Taxonomy: The Systematic Evaluation of Comparative Data*, 2nd edn. New York: Columbia University Press.

Wiens JJ (1999) Polymorphism in systematics and comparative biology. *Annual Review of Ecology and Systematics* 30: 327–362.

Winsor MP (2006) The creation of the essentialism story: An exercise in metahistory. *History and Philosophy of the Life Sciences* 28: 149–174.

Temperate Forests

John A Silander Jr., University of Connecticut, Storrs, CT, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 607–626, © 2001, Elsevier Inc.

Glossary

Evapotranspiration The process of transferring moisture from the earth to the atmosphere by evaporation of water and transpiration from plants: actual evapotranspiration as observed at a locality or potential evapotranspiration, given unlimited water availability.

Neogene The Miocene and Pliocene Epochs, accorded the status of a period when the Tertiary is considered an era.

Physiognomy The outward appearance or morphology of a community as determined by the growth forms of the dominant plants present.

Sclerophyll A plant with tough, leathery, evergreen leaves, usually associated with drought resistance.

Sere The series of stages in an ecological succession sequence.

Temperate region Any locality with at least 1 month of frost (for continental areas) or with one or more months with a mean temperature lower than 18 °C (for maritime-influenced areas), and with at least 4 months with a mean temperature higher than 10 °C.

Tertiary The earlier part of the Cenozoic Era, occurring from about 65 to 2 million years ago.

Overview

The temperate forests are globally important and unique. They host the largest and oldest organisms in the world. They serve as the world's major source of timber and wood products and are perhaps the only forests with some proven potential for sustainable management. The biomass of at least some temperate forests stands exceeds that of any tropical forest. The temperate forests of the world also provide critical ecosystem services locally and globally. Recent evidence indicates the global importance of carbon sinks in the temperate forest zone, especially in eastern North America. On a landscape level, temperate forests are critical to modulating hydrological, nitrogen, and carbon cycles. Although the biodiversity of temperate forests is typically much lower than that of tropical forests, some temperate forests approach the biodiversity observed at larger spatial scales in the tropics. Temperate forest biodiversity hot spots with high levels of endemism rival in importance those anywhere. They have a unique evolutionary history divergent from either the tropics or the boreal regions. Moreover, Northern and Southern Hemisphere forests are as different from each other as either is from tropical or boreal forests. This contrast reflects striking differences in climate, biogeography, evolutionary history, and the impact of humans.

The objectives here are to provide an overview of the distribution of temperate forests globally; their structure and composition; evolutionary history; diversity, endemism, and rarity; ecosystems services provided; current conservation status; and current threats.

Global Distribution Patterns of Temperate Forest Systems

A global view of the forests of the world reveals everything from broad continuous expanses of trees to mosaics of small forest patches in the landscape and from dense, closed-canopy stands to open wooded parklands. As one progresses from the equator through the humid zones at midlatitude to tree line in the polar regions, changes in forest structure and composition typically occur gradually. Demarcating where the temperate forests begin and end along this continuum is difficult. Indeed, defining "temperate" is difficult. Of the many published maps portraying the extent of the temperate forest biome, few agree on boundaries. The broadest, most arbitrary definition of temperate forests includes all forested areas north or south of the tropics of Cancer and Capricorn, respectively. More common, general macroclimatic factors have been used to define boundaries. In such cases, the results are maps of the potential distribution of forests. Because of the pervasive effects of human and natural disturbances, such maps are considerably more extensive than those depicting extant forests (compare **Figures 1** and **2**). Maps of existing forests are generally derived from remotely sensed images, but even these vary in what is portrayed as forest. In consequence, **Figure 2** does contain acknowledged, inherent errors in misclassification and omission.

Figure 1a shows where globally temperate forested landscapes may potentially be found, delimited by the humid temperate domain or ecoregion. The boundaries are based on macroclimate, which distinguishes this zone from the polar ecoregion with boreal forests, the humid tropical region with tropical forests, and the dry ecoregions dominated by arid grassland/savanna or desert vegetation. The latitudinal boundaries are set by thermal regime as modified from Köppen and Trewartha, who developed the most commonly used climate classification scheme. These zones are similar to those of Holdridge and Walter but differ in detail. The temperate region defined in **Figure 1** includes any locality with at least 1 month with frost (for continental areas) or with 1 or more months with a mean temperature lower than 18 °C (for maritime-influenced areas) and with at least 4 months with a

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00142-8

mean temperature higher than 10 °C. Moisture availability sets the remaining boundaries, with the humid temperate domain bounded by where precipitation equals or exceeds potential evapotranspiration. Most other maps of temperate forest biomes employ variations on this theme.

Biome or community boundaries are only approximate, reflecting the difficulty inherent in delineating features that are fuzzy spatially and temporally. Boundaries are often either broad ecotones or a mosaic of patches. Moreover, these patterns will change as the climate, biotic composition, or disturbance regimes inevitably change over time. Consequently, any map will be at best an abstract representation of reality.

Within any biome or ecoregion, subdivisions of convenience may be designated. In **Figure 1**, the humid temperate domain can be subdivided into subclimatic zones: marine (areas with temperature fluctuation moderated by oceanic influences and with elevated moisture availability, in some cases producing rainforests) and continental (areas with comparatively greater temperature fluctuation and greater probability of drought). Temperature zones are delineated as well: subtropical (defined as having 8 or more months with mean temperatures higher than 10 °C), continental hot temperate (4–7 months with temperatures higher than 10 °C, warmest month higher than 22 °C, and coldest month lower than 0 °C), and continental warm temperate (as for hot temperate, but with warmest month lower than 22 °C).

The excess of annual precipitation over evapotranspiration becomes less as one moves away from oceanic influences on continents at midlatitudes. Thus, temperate forests tend to be replaced by grasslands in central North America, central and eastern Europe, central eastern Asia, and eastern South America. This may be augmented by fire and grazers or browsers. The boundary between forest and grassland may be a broad transitional, open wooded parkland with scattered trees, or a broad mosaic of forest and grassland patches with forests restricted to favorable soils, sites with more moisture, or sites protected from fire.

Forests may also be replaced as aridity increases in Mediterranean climates with winter rainfall–summer drought regimes. As fire becomes a pervasive element in the landscape, forests tend to be replaced by sclerophyllous shrublands, thickets, or sometimes open woodlands. Again, the transition may be abrupt or a gradual mosaic, with forest patches restricted to favorable soils and/or moist sites protected from fires.

The boundaries between the temperate forested regions and the adjacent boreal or tropical regions are often imprecise. In both cases, there can be a very broad transition zone with considerable overlap in species composition. The subtropical–tropical boundary set by the 18 °C mean monthly isocline is quite arbitrary. Some have attempted to show that this boundary approximately corresponds to the natural poleward limits of the distribution of palms.

Boreal forests are distinguished only for the Northern Hemisphere as typically conifer-dominated forests within specified climate regimes (i.e., monthly mean temperatures all

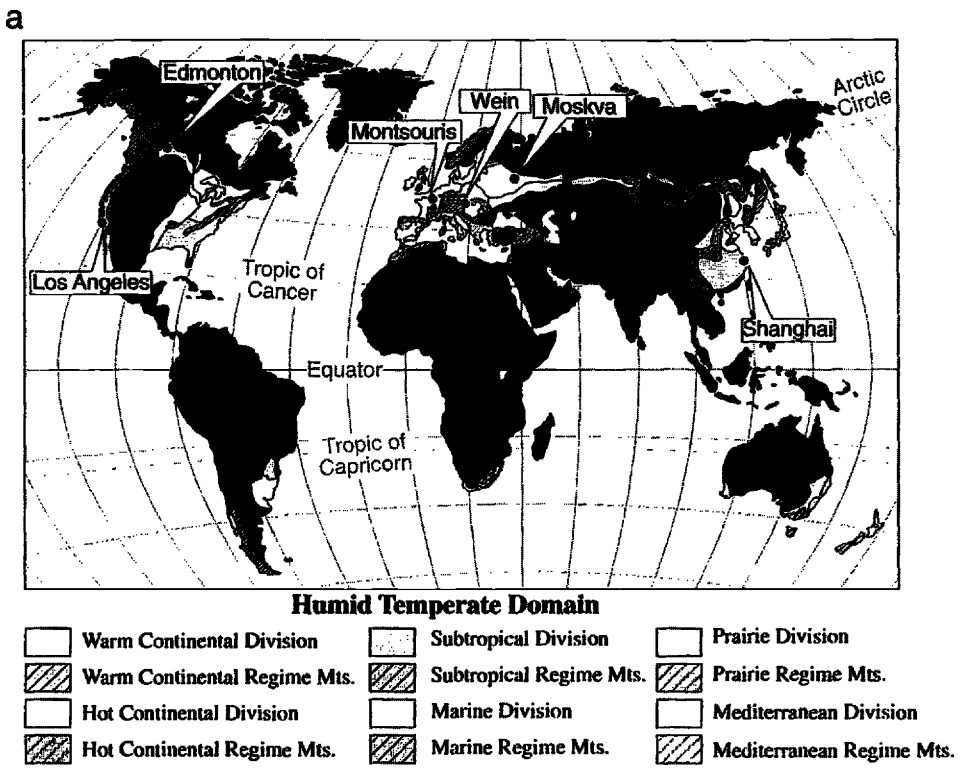


Figure 1 (a) Global distribution of the humid temperate ecoregions. These correspond to the potential distribution of temperate forest, grassland, and Mediterranean shrublands globally. (b) Representative climate diagrams for the various divisions of the humid temperate ecoregion domain (Reproduced with permission from Figure 6.2 and 6.3 in Bailey RG (1998) *Ecoregions: The Ecosystem Geography of the Oceans and Continents*. New York: Springer-Verlag.).

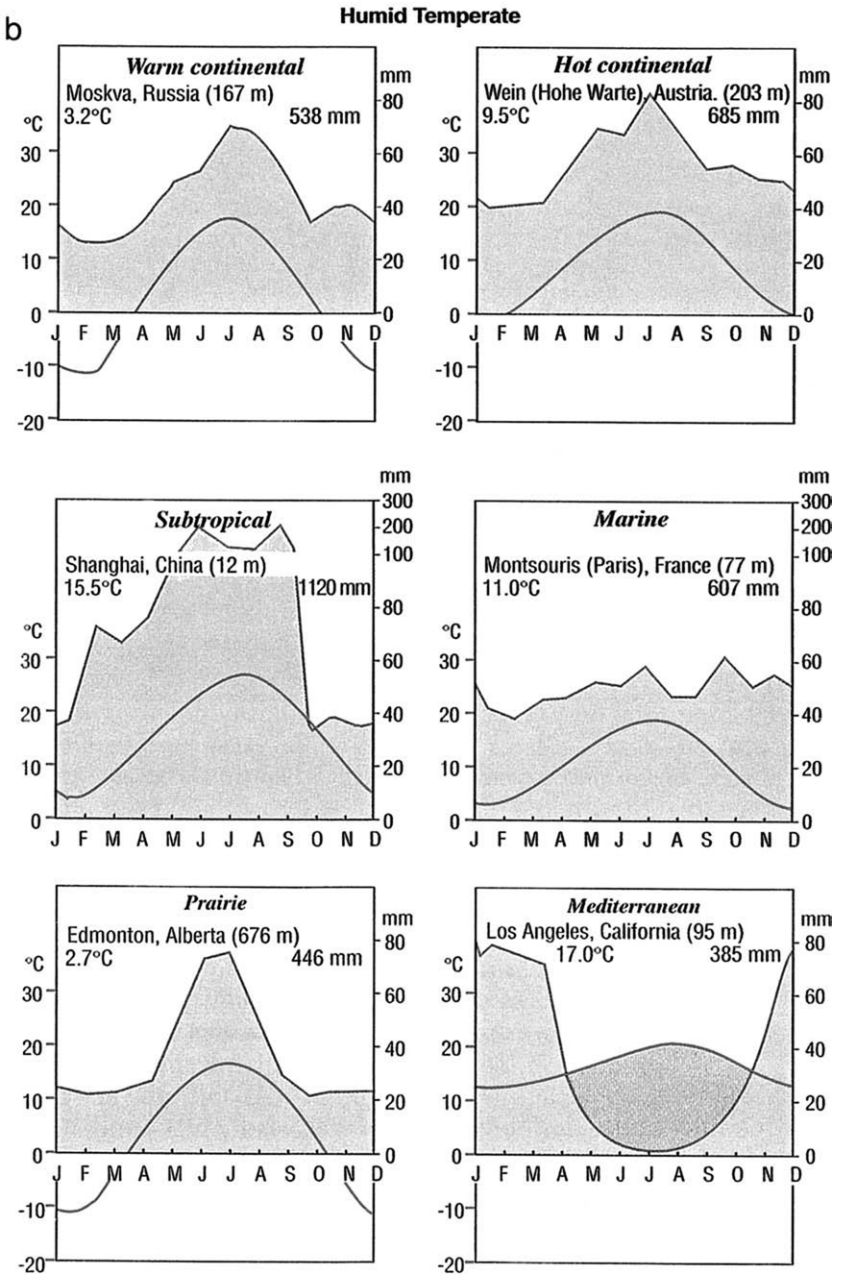


Figure 1 Continued

lower than 22 °C but 1–3 months with means higher than 10 °C). This boundary approximately corresponds with the mean position of the summer polar front in the Northern Hemisphere.

Confusion remains regarding the classification of conifer-dominated forests. The boreal forests of the Northern Hemisphere and the conifer forests of the North American Pacific Northwest, the Asian Pacific northeast, and the Southern Hemisphere are treated differently by different authors. Using the previously mentioned temperature criteria, all of the Southern Hemisphere forests, except perhaps the southern tip of South America, would be classified as temperate, as would most of the Pacific forests of Asia and North America. The climatic definition for the temperate zone forests will be used here.

There are parallel inconsistencies in classifying certain subtropical forest regions. For example, northern Indian forests, which are climatically subtropical, are usually classed as tropical. The same is true for forests with some temperate affinities in subtropical regions of southern Brazil.

Dealing with mountainous areas is problematic since they typically contain multiple ecoregions (tropical or temperate to boreal and alpine depending on elevation and latitude). In the Americas, Africa, eastern Asia, and Australasia, one can find discontinuous bands of forest with temperate affinities extending from sea level in the temperate zone well into the

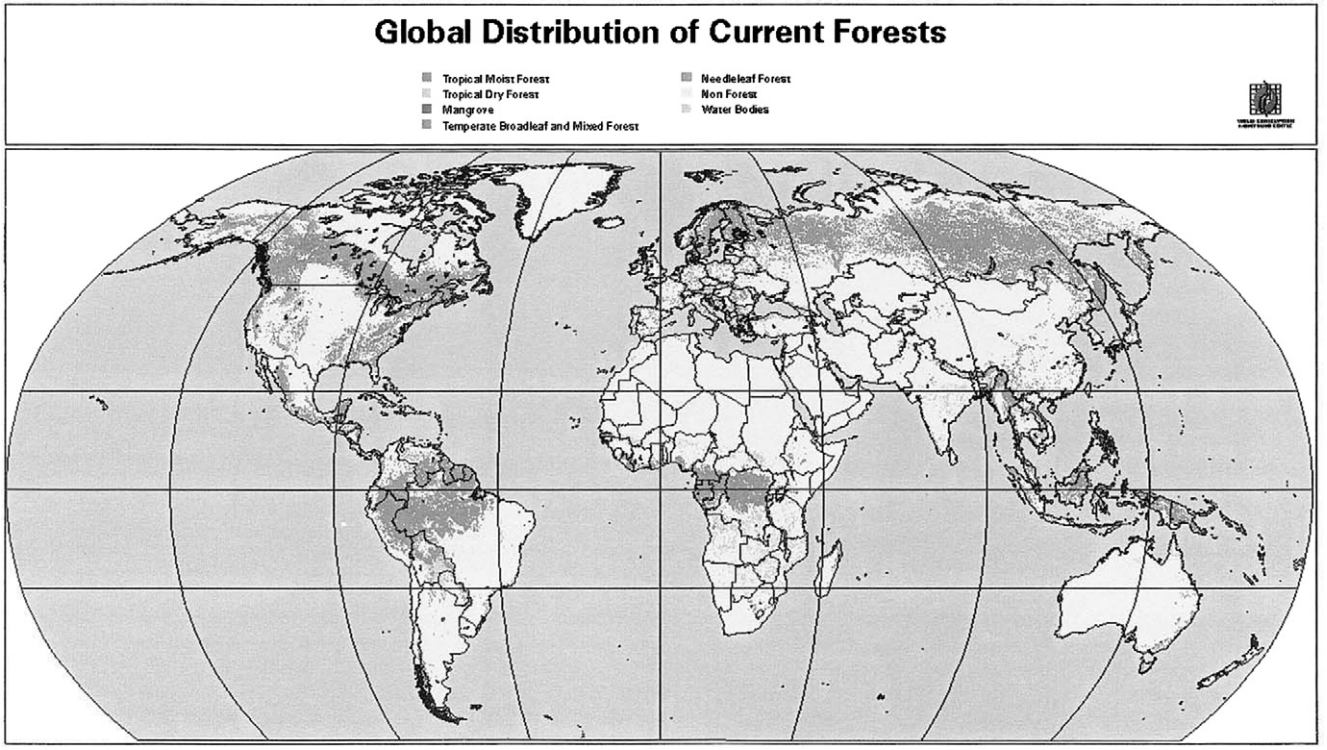


Figure 2 The current distribution of forests globally. The data on which this map is based were assembled by the World Conservation Monitoring Centre in collaboration with the World Wide Fund for Nature and published on the World Wide Web (http://www.wcmc.org.uk/forest/data/wfm.htm).

tropics at higher elevations. The most common resolution is to classify all forest elements as part of unclassified mountainous regions, as part of the surrounding domain, or as separately classified biome units within the domain. One consequence of this inconsistent classification is to obfuscate biotic patterns in some of the most important global hot spots of biodiversity.

The temperate forest biomes, thus defined, comprise about 14,600,000 ha (estimates vary). By far the largest actual or potentially forested landscapes occur in the Northern Hemisphere (80% or more) (cf. **Figure 1** and **Table 2**). The regional biomes include (i) eastern North America from the Atlantic coast west to about 95° latitude and from about 45° latitude south to about 28°; (ii) western North America from about 35° north to about 60° (and mainly from the Sierra–Cascade ranges, west); (iii) western and central Europe from the Atlantic coast north to about 60° and east through eastern Europe, but excluding the Mediterranean coastal zone and much of Spain, and then extending in a narrow strip around 55° east across Russia to west central Asia; (iv) a small, discontinuous temperate forest zone in the Middle East, especially along the south coast of the Black Sea, to the southern Caucasus and to the southern Caspian Sea; (v) eastern Asia from about 50° south to about 25° in southern China and from Japan and the Pacific coast northwest to about 120° and southwest to about 100°; and (vi) northern south Asia (India) and adjacent areas.

In the Southern Hemisphere the extent of temperate forests is much more restricted in extent: (i) eastern, coastal Australia from about 25° south to Tasmania, plus the southern tip of Western Australia; (ii) most of New Zealand; (iii) southern Chile and adjacent Argentina from about 40° south to about 55°; (iv) a small area of southern Brazil just below the tropic of Capricorn, plus adjacent Paraguay and Argentina; and (v) small patches of coastal and interior forest in south and southeastern South Africa.

Table 1 summarizes the potential and current extent of forested areas within each of the 11 temperate forest biomes, from World Conservation Monitoring Centre (WCMC, see **Figure 2**), World Wide Fund for Nature (WWF), and the Food and Agriculture Organization data based on approximately 25 forest types. The more conservative estimate of potential forest (**Table 1**, column 2) is for mesic forest coverage only. Maximum potential extent (column 3) includes all types of dry forest, woodlands, and thickets but not savannas or other sparsely treed landscapes (i.e., forest cover <30%). The corresponding ranges under current forest extent are shown in the fourth column. These data are only approximate. The base data were obtained from different sources, with different categories or forest type classes reported for different regions. In some cases, conflicting information on forest cover is reported from different sources.

Almost 80–90% of the temperate forest biome types are found in the Northern Hemisphere. For current forest cover, the data are only slightly less for northern dominance. The major northern forest biomes of Europe, eastern North America, and east Asia all cover about the same potential area. However, eastern North America has by far the largest cover remaining (40–45%). The Middle East has the smallest percentage of potential forest cover (13–18%), followed by east Asia (<20%) and Europe (25–30%); western North America has the highest percentage of remaining forests (60–75+%). Estimates for western North America are only approximate (estimated from WCMC forest type information and other data), with the eastern and southern boundaries for this system fuzzy and disjunct. Similarly, the estimates for forest cover in south Asia are only approximate. For some regions these data can be deceptive. Much of the forest covering Europe, especially in western Europe, is intensively managed plantation or seminatural forest. The same is true for Japan, even though it retains more than one-third of its potential forest

Table 1 Temperate forest cover by regions

Forested region	*Potential extent of temperate (mesic) forest cover ($km^2 \times 10^6$)*	*Maximum potential forest extent (including dry forests, woodlands, and thickets) ($km^2 \times 10^6$)*	*Current estimated extent of temperate forest cover ($km^2 \times 10^6$)*	*Conservation areas (IUCN classes I–VI) ($km^2 \times 10^3$)*
Europe (including Mediterranean)	3.30	3.91	0.85–1.1	44.0–54.0
Russia	1.12	1.13	0.26–0.36	11.0–36.0
East Asia	3.21	3.79	0.62–0.75	24.0–26.0
North America	4.26	4.72	2.13–2.17	117.0–132.0
Eastern north	3.56	3.72	1.63–1.57	—
Western (Pacific)	0.7	1.0	0.5–0.6	—
Middle East	0.36	0.61	0.05–0.11	1.0–4.0
South Asia	0.87	1.34	0.20–0.31	20.0–34.0
South America	0.7	1.8	0.4–0.52	45.0–69.0
Southern (Chile and Argentina)	0.60	0.8	0.25–0.30	51.0–54.0
Southeastern (Southern Brazil and adjacent countries)	0.14	1.0	0.12–0.22	12.0–15.0
Australia	0.45	1.64	0.03–0.66	4.0–44.0
New Zealand	0.23	0.25	0.04–0.08	17.0–18.0
Southern Africa	0.1	0.4	0.01–0.1	4.0–6.0
Total	14.6	19.6	4.7–6.2	287.0–423.0

Table 2 Worldwide forest cover

Forested biome	*Potential extent of mesic forest cover ($km^2 \times 10^6$)*	*Maximum potential extent (including dry forests, woodlands, and thickets) ($km^2 \times 10^6$)*	*Current estimated extent of forest cover ($km^2 \times 10^6$)*
Boreal forests	12.2	18.5	8.0–11.5
Tropical forests	29.0	40.0	14.2–15.6
Temperate forests	14.6	19.6	4.6–6.8
Total	55.8	78.1	26.8–33.9

cover. China retains almost 12% of its original forest, but most of this is in the boreal to northernmost, mixed conifer forests.

Southern temperate forests are fragmented into six or more regional biomes, all much smaller in extent than their northern counterparts. The extreme ranges in forest cover tabulated for some of the southern biomes reflect the large contribution of dry forest systems within the regions. The largest of the southern biomes comprises the forests of southern Chile and adjacent Argentina. The smallest of all temperate forest biomes is that found in South Africa, perhaps rivaled in size only by the forests of Western Australia. Most significantly, the South African forests are the richest of all temperate forests in tree families, genera, and possibly species (certainly if the complete tree flora is included).

Compared to other forest system globally, the temperate forest biomes covers a slightly greater area than boreal forests. Boreal and temperate forests together approximate the tropical forests of the world (**Table 2**).

General Characterization of Temperate Forests

Physiognomic Features

Each of the previously discussed 11 temperate forest regions may be characterized by general physiognomic features (i.e., the outward appearance and structure of the dominant vegetation):

1. Eastern North America: dominated by broadleaf deciduous forests, mixed with conifers to the north, locally dominated by conifers under drier, successional conditions or in fire-prone areas in the southeast and northwest, plus small patches of broadleaf evergreen forests in the south. Closed forest systems predominate.
2. Western North America: dominated by evergreen conifers with broadleaf trees contributing little to the forests; some northwestern coastal areas support rainforests (i.e., rainfall in excess of 2000–3000 mm per annum). The eastern edge of this biome is discontinuous and grades to open conifer woodlands and montane boreal forests. To the south it grades to a mosaic of conifer or broad-leafed evergreen forests and woodlands or shrublands.
3. Europe: mainly broadleaf deciduous forests, mixed with conifers to the north and in mountainous areas, and to the south grading to a mosaic of broadleaf evergreen forests (many sclerophyllous leafed), conifer-dominated forests, and shrublands under Mediterranean climate influence.
4. Middle East: mainly broadleaf deciduous forests with a mosaic of broadleaf evergreen, sclerophyllous, and conifer forests and woodlands. Rainforests occur very locally near the southeastern Black Sea coast.
5. Eastern Asia: mainly broadleaf deciduous forests mixed with conifers to the north and broadleaf evergreen forests to the south; mostly mesic closed forests. Evergreen broad-leafed rainforests occur locally in southeastern Japan, one of the rarest forest types found in the Northern Hemisphere.
6. South Asia: dominated by broad-leafed evergreen to semievergreen monsoonal forests; temperate mixed forests (locally rainforests) occur in the foothills of the Himalayas.
7. Australia: dominated by broadleaf evergreen forests (mostly sclerophyllous); small patches of closed forest (<70% cover) and extensive areas of open forests grading to woodlands. Very restricted patches of rainforest are found in southeastern Australia and western Tasmania.
8. New Zealand: dominated by conifers and mixed with broadleaf evergreen forest patches, especially on the north island and locally on the south island. The western forests in New Zealand are rainforests.
9. Southern South America: dominated by broadleaf evergreen forests with some conifers. Broadleaf deciduous forests prevail in the southernmost areas and at higher elevation in the Andes; rainforests are restricted to the Pacific slope south of about 40° latitude.
10. Southern Brazil: characterized by the presence of southern conifer forests; elsewhere in this zone the forests are dominated by evergreen or semideciduous angiosperms with patches of open forest and thicket to the west.
11. South Africa: mixed broadleaf evergreen (many sclerophyllous leafed) with some conifers forming a very patchy mosaic in the landscape with thickets, shrublands, and savannas.

Dominant Floristic Features

The community dominance and floristic affinities for these regions as they exist today can be characterized very broadly. Detailed descriptions and characterizations may be found elsewhere. There are broad floristic affinities among the forested biomes in the Northern Hemisphere with many shared

families and genera both now and in the fossil record. These include older lineages from the Tertiary flora of Asiamerica and recent lineages that evolved under cooler and/or drier climates. The Southern Hemisphere forests are very different floristically, with few important families or genera shared with the north. There are some Gondwanan lineages in common now or in the fossil record. However, the floras today reflect considerable divergence with many tropical affinities and many fewer common links than are seen in the north.

In eastern North America, in the northeast birch, maple, beech, and hemlock (*Betula, Acer, Fagus*, and *Tsuga*) dominate the landscape, with the latter two tending to form monodominant stands late in succession. Farther south and west the forests tend to be dominated by oaks (*Quercus* spp.) or hickories (*Carya* spp.). The central sections, especially in the central to southern Appalachians, tend to be the richest, with a diminished tendency toward dominance by one or a few canopy species, although the highest regional tree diversity is found in the southeast. In the extreme south, small patches of evergreen forest are found in protected areas dominated by evergreen oaks and *Magnolia*. Much of the southeast is now dominated by pine (*Pinus taeda*) plantations. Fire successional pines also dominate parts of the northwest. Drier sites throughout tend to be dominated by oaks or conifers, especially pines; wetter sites are dominated by conifers (e.g., *Tsuga* to the north and *Taxodium* to the south) or by locally adapted broad-leafed deciduous species (e.g., *Ulmus, Nyssa*, and *Acer*).

Western North American forests are dominated by a small number of large, long-lived conifer species; deciduous angiosperms are only minor components. The Pacific Coast rainforests are dominated by hemlocks (*Tsuga*), firs (*Abies*), spruce (*Picea*), and/or cedar (*Thuja*) from Alaska south to Washington State. Douglas fir (*Pseudotsuga*) becomes important in the central coast. Timber industries in these zones tend to actively manage the landscape for monospecific stands of native species (e.g., Douglas fir). Fragmented stands of redwoods (*Sequoia*), which globally are the tallest trees, occur more southerly extending to central California. On the drier eastern slopes of the coastal mountains (central sections), pines, Douglas fir, and sometimes poplars (*Populus*) dominate the landscape. Further south in the Sierras one finds local dominance by *Sequoiadendron*, the most massive tree globally, and *Pinus aristata*, the longest lived tree. Drier, lower elevation landscapes in the south (California) may be dominated by broad-leafed evergreen forests of tanoaks (*Lithocarpus*) and madrone (*Arbutus*) or open oak, pine, or mixed woodlands, which grade into Mediterranean shrublands or grasslands.

European forests are highly disturbed and fragmented following centuries of human habitation. Central European forests tend to be dominated by beech (*Fagus*) on many intermediate sites and by various oaks (*Quercus*) on drier and slightly wetter sites or very acidic sites. The wettest sites tend to be dominated by birches (*Betula*). Many of these forests have been converted to *Picea* or *Pinus* plantations. For example, German forests have gone from 90% deciduous broadleaf domination to 80% *Picea* plantations. To the south there is a transition to dominance by evergreen oaks and pines under Mediterranean influence. To the north and throughout much of the central uplands, conifers (*Pinus, Picea, Abies*, and *Larix*) can locally dominate the landscape or there are mixed forests with beech and birch. Many of the European forests not in plantation are still actively managed for timber products.

The Middle Eastern forests of northern Turkey to the Caucasus and the southern Caspian Sea are probably the poorest known and least studied of temperate forests. In the western Mediterranean-influenced zone, one finds sclerophyllous forests and woodlands dominated by pines and oaks. Along the coast of the Black and Caspian Seas and the southern Caucasus (Colchian and Hyrcanian regions) one finds highly diverse, mesic deciduous forests with oaks, maples, beech, chestnut, and many other species. At higher elevations the forests are dominated by beech with conifer-dominated or mixed stands above. Drier sites in the landscape are dominated by open oak forests or woodlands. Many of these forests have experienced a long history of human occupation and associated agriculture with overgrazing.

East Asia has the most diverse forests in the Northern Hemisphere. Along the Pacific coast in northern Japan, southern Russia, and China, one finds forests similar to those of the North American Pacific Northwest, with conifer dominance but also mixed with broad-leafed species of maples, birches, limes (*Tilia*), and elms. Central China has been extensively cultivated for centuries, and there is very little forest cover left. Remnant tracts here indicate a rich diversity of mixed deciduous trees mentioned previously plus many other genera, including oaks, elms, poplars, ash (*Fraxinus*), and rowan (*Sorbus*), with a rich understory. Locally, the understory may be dominated by bamboos. The same pattern is seen in Korea and Japan. In southern Japan, much of the forest is managed for native *Cryptomeria*, and in the north it is managed for *Abies* or *Picea*. South of the Yangtze River in China and in eastern and southern Japan the broad-leafed evergreen species increase in dominance as one approaches subtropics. These are the most diverse temperate forests in the Northern Hemisphere, with many of the same genera listed previously. However, these have become highly fragmented through human disturbance. In southern China there is a shift to forests with strong tropical affinities.

The temperate forests of south Asia are difficult to categorize. Most of the lowland and premontane remnant forests of northern India are climatically subtropical but the flora has strong tropical affinities. There is local strong dominance by Sal (*Shorea robusta*) and bamboos. In the northern hill forests there is a transition to strong temperate affinities with high species diversity. Oaks mixed with Lauraceae dominate the forest, but maples, *Castanopsis*, and *Magnolia* occur. In the montane zone diverse oak forests are mixed with conifer (*Abies, Picea*, and *Pinus*) forests and patches of *Rhododendron*. As is the case in much of Asia, these forests have been long affected by human disturbance. Throughout the lower elevations small stands of *Eucalyptus*, teak, pine, or *Populus* plantations are common.

The subtropical forests of the eastern mountains and coastal areas of Australia and the small areas of forest in Western Australia are dominated by the numerous sclerophyllous *Eucalyptus* species. Locally, stands tend to be dominated by one or only a few species. The eucalypt forests tend to form open-canopy stands grading to woodlands. The tallest angiosperm trees are found here. Only on Tasmania and in scattered pockets along the eastern mountain chain is there

Table 3 Tree species diversity patterns across the temperate forested region

Region	*Families*	*Genera*	*Species*	*Genus:family*	*Species:genus*	*Forest biome maximum extent ($km^2 \times 10^3$)*
Europe	21	43	124	2.0	2.9	3300–3910
East Asia	67	177	876	2.6	4.9	3210–3790
Eastern North America	46	90	253	2.0	2.8	3560–3720
Western (Pacific) North America	24	47	131	2.0	2.8	700–1000
Chile	29	40	83	1.4	2.1	330–370
Southern Brazil	25	45	77	1.8	1.7	<100
Southeast Australia	37	78	331	2.1	4.2	300–700
New Zealand	47	74	212	1.6	2.9	230–250
Southeast South Africa	88	280	598	3.2	2.1	20–50

local dominance by closed-canopy or temperate rainforest species, including the southern beech (*Nothofagus*) and various southern conifers (*Dacrydium, Phyllocladus, Arthrotaxus*, or *Araucaria*). Throughout much of southeastern Australia plantations of Monterey pine (*P. radiata*) have become a pervasive component of the landscape.

The forests of New Zealand tend to be either multi-storied mixed conifer–broadleaf with species composition varying across the landscape or low-diversity southern beech (*Nothofagus*) forests. The mixed forests dominating in the lowlands have a scattered overstory of *Agathis* in the north or various podocarps (e.g., *Podocarpus, Dacrycarpus*, and *Phyllocladus*), with a sub-canopy of Lauraceae (*Beilschmeidia*), Myrtaceae (*Metrosideros*), Cunoniaceae (*Weinmannia*), and many other families and genera. Evergreen *Nothofagus* forests may form pure dense canopies in subalpine areas and may be a component in the lowland forests along with other broad-leafed species.

The forests of southern South America are confined to Chile and adjacent areas of Argentina. They vary from small remnants of sclerophyllous forests and woodlands in the Mediterranean zone to the species-rich Valdivian rainforests, the species-poor but still extensive north Patagonian and Magellanic forests, and depauperate deciduous *Nothofagus* forests at higher elevations and the interior south. The sclerophyll forests were dominated by *Acacia caven* and other species, with deciduous *Nothofagus* forests at high elevation. The Valdivian forests may be either broadleaf dominate, with *Nothofagus, Eucryphia* (Eucryphiaceae), *Laurelia* (Monimaceae), *Weinmannia*, and other species, or mixed with conifers (*Podocarpus, Araucaria*, or *Fitzroya* (Cupressaceae)). The north Patagonian/Magellanic forests are dominated by evergreen *Nothofagus* mixed with *Podocarpus, Weinmannia*, and *Drimys* (Winteraceae). To the south, one of the deciduous *Nothofagus* species (*N. pumilo*) forms a pure stand or is mixed with *N. betuloides* at timberline. Many of the remaining Chilean forests are being clear-cut for chips and converted to *P. radiata* or *Eucalyptus* plantations.

The southeastern forests of Brazil have many south temperate affinities. These forests are characterized by the presence of *Araucaria*. However, other south temperate components include *Podocarpus, Weinmannia*, and *Drimys*, plus Sapindaceae, Proteaceae, and Myrtaceae. These forests have been largely cleared for agriculture. The subtropical forests to the west in adjacent Paraguay and Argentina have more tropical affinities.

The forests of South Africa are some of the richest in tree species of any in the temperate zone. However, this is also the smallest of all temperate forest biomes, and it is highly fragmented into many small forest patches. It is not clear how forest cover has changed during the Pleistocene. These forests do have south temperate affinities, with the presence of *Podocarpus*, Cunoniaceae, and Proteaceae, but the majority of the temperate forest flora have tropical affinities. Despite these tropical affinities, the level of tree endemism is high for a continental area contiguous with tropical forests. The "Afromontane" forest elements extend from southernmost South Africa at sea level to the mountains through northeastern Africa. The coastal, Indian Ocean (Maputaland/Pondoland) forests are quite different floristically, with high diversity and many local endemics. The vast majority of the forested landscape is now in *P. radiata* or *Eucalyptus* plantation.

Trends in Biodiversity

Comparing trends in biodiversity within and among regions is fraught with difficulties. The results can differ depending on the scale of the sample unit compared (i.e., 0.001 vs 1, 1000, or 100,000,000 ha). For many regions of the world, data are available only for a limited scale range. For example, in east Asia there are very few accessible data records for small plots (0.1–100 ha). In other regions (e.g., the Middle East), species diversity numbers are either estimates for large areas or entirely lacking. For many records the number of tree species may be accurately reported but the number of herbaceous, especially ephemeral, plants may be significantly undercounted or not reported at all. Even simply listing tree species numbers can be misleading because authors vary widely in delimiting the threshold size for what constitutes a tree. Likewise, authors vary in classifying vegetation type with which tree taxa may be associated (closed forest to sparsely treed parkland). Nevertheless, some trends appear robust.

Table 3 summarizes regional tree taxon richness from a variety of different sources, and **Figure 3** plots tree richness tallies against area for 75 forest sample sites throughout the temperate zone (spanning 10^{-2} to 10^{8}ha). Despite this large

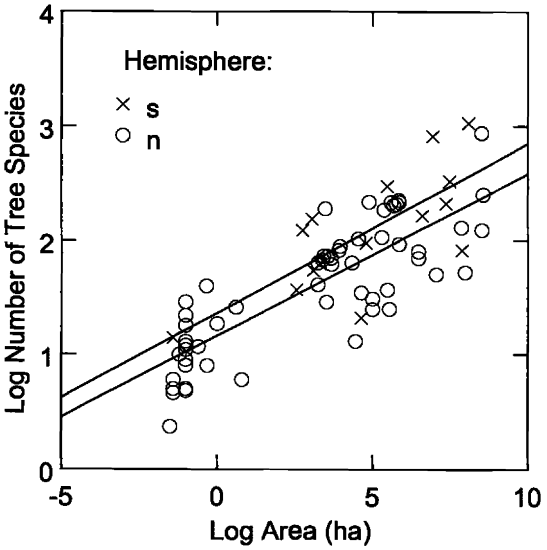


Figure 3 The relationship between tree species numbers and area sampled is shown, log transformed for both variables. The samples are from a wide range of different sources for most of the temperate zone forest biomes. These are simply classified here as Northern (n) and Southern (s) Hemisphere localities with best fit linear regression lines. There were significant differences between hemispheres ($p = 0.024$, $n = 75$, $r^2 = 0.695$) and among regions ($p < 0.001$, $r^2 = 0.916$).

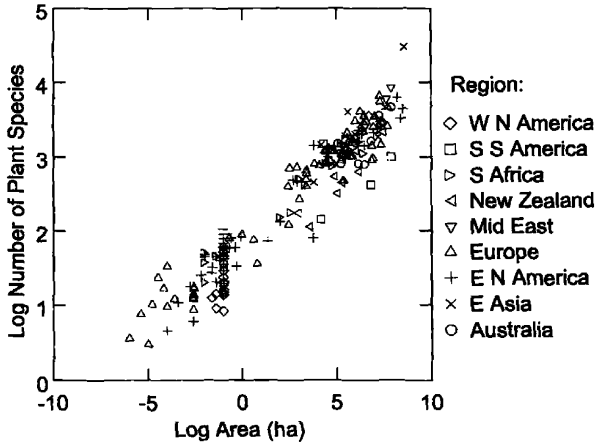


Figure 4 The relationship between total vascular plant diversity and area samples is shown, log transformed for both variables. The samples are from a variety of different sources covering most of the temperate zone forest biomes. The samples are classified by the biome region from which they were obtained. Significant differences are found between hemispheres ($p = 0.005$, $r^2 = 0.924$, $n = 198$) and among regions ($p < 0.001$, $r^2 = 0.938$).

range in areas and the inherent variation in estimates and counts, there are significant differences in tree richness among biomes and between hemispheres. For these data, greater tree species richness occurs in the Southern Hemisphere across the full range of areas surveyed (Figure 3). This is also reflected in the absolute number of families and genera. For species, the totals are similar, but the tabulated survey covers only a small part of the diverse tree flora for Australia and South Africa, respectively, and only part of southeastern Australia and Tasmania and Kwazulu-Natal province and adjacent Transkei. If all the tree species are included for South Africa alone (well over 1000 species in 370 genera and 97 families), the species numbers would be higher in the Southern Hemisphere despite the fact that southern forests only cover 10–20% of the area of northern forests. On an area basis, the taxon richness of the southern forests is at least an order of magnitude greater than that of the north.

This high southern diversity is contributed largely by the flora of South Africa, arguably one of the richest per unit area of any biome globally. Australia and New Zealand also contribute to this southern richness, each having a taxon richness per unit area of 4–10 or even 100 times that of northern forested regions. Just the rainforests of New South Wales and Victoria, covering less than 200,000 ha, have more than 250 tree species (not all included in Table 3). Even the Chilean forests, perhaps the most depauperate in tree species of any temperate forest biome, have more taxa per unit area than any northern biome (except for total species in east Asia).

Among northern temperate forest biomes, east Asia has by far the richest tree flora. Europe and western North America are the most depauperate, with Europe having the lowest tree taxon diversity per unit area globally.

Figure 4 shows analogous species area data for all vascular plants tallied for about 200 plots or regions (spanning 10^{-6} to 10^{8} ha) across all temperate forest biomes. Here, the hemisphere trends are reversed. The north has a slight but significantly higher total vascular flora than the south across the range of sample areas. East Asia and the Middle East (the latter, a smaller sample size) tend to have the highest vascular plant diversity, and New Zealand, South America, and western North American forests tend to have the lowest vascular plant diversities. Forested systems in eastern North America, Europe, Australia, and South Africa tend to be intermediate across areas sampled. There is a dramatic decrease in species diversity toward the pole, more so in the Southern Hemisphere. In pairwise comparisons at intermediate sample areas the trends tend to hold. The vascular plant flora for east Asia at 0.5–10 ha is significantly richer than that of Europe or eastern North America. However, few differences can be seen among regions at the smallest plot sizes (<0.01 ha).

Trends in alpha, beta, and gamma diversity vary considerably within and among regions. Some of the large- to intermediate-scale patterns in species richness are undoubtedly related to spatial landscape heterogeneity. Regionally, high levels of species richness are associated with mountains (e.g., the Smoky Mountains (east North America), the Pyrenees, Alps, or Balkan Mountains (Europe), and the Sichuan Mountains or Mount Halla (east Asia)). Complex mosaics of vegetation types (e.g., southeast Australia or Pondoland (South Africa)) or multiple successional states (e.g., Indiana Dunes National Seashore) all tend to support significantly higher diversity than adjoining areas. Even for the United Kingdom, with an exceptionally small flora overall, regional or county floras that encompass a diversity of habitats and seral stages may have as many species as are found in many similar-sized temperate regions elsewhere. Within Europe there many more tree species in the southern,

Mediterranean-influenced region with complex spatial heterogeneity (almost 100 versus 12 common and 25 uncommon species in northern and central Europe). Perhaps the high tree and overall species diversity in South Africa is related to the extreme spatial heterogeneous and natural fragmentation of the vegetation (forest and otherwise). Locally or regionally, lower species diversity tends to be found under closed conifer forests and late successional stands dominated by trees that cast deep shade. Often, the highest diversity will be found in intermediate successional stage forests or in a landscape with a mosaic of seral stages. Locally, higher diversity will tend to be found on richer, fine-textured soils with circum neutral pH with higher cation exchange capacity and more humus. This contrasts with diversity trends elsewhere. For example, higher diversity tends to be found on poorer soils in Mediterranean shrubland biomes.

Compared to tropical forests, the species richness (of trees or all vascular plants) in the temperate zone will usually be smaller than that of comparable areas in the tropics, especially at the 0.05- to 10-ha scale. Most temperate forest stands tend to be dominated by one or a few species, with the other tree species being uncommon or rare. In the majority of tropical forest stands local dominance by one or a few species is rare; greater evenness is common. Perhaps only on a regional basis may floristic diversity of temperate forests (e.g., southern China and South Africa) begin to approach those observed in the tropics.

An alternative measure of global richness patterns may be obtained by focusing on clades or higher taxonomic richness. The data from **Table 3** show more genera and families in the Southern Hemisphere forests sampled. Also, the ratios of species to genus and genus to family tend to be lower in the south. These trends indicate a potentially richer phylogenetic diversity (using higher taxa in lieu of cladistic information) in the south. A complementary, global perspective is provided by the British Natural History Museum (NHM) global mapping of plant family richness (395 in total) on 10° latitude/longitude grids. The trends reveal higher family diversity on average in the tropics, with the richest cell in Southeast Asia, and a dramatic decline poleward. However, family richness is as high in southern China as it is anywhere else in the world (other than Southeast Asia), and about as high in the central southern United States as in most of South America (including Amazonia) and southern Africa versus tropical Africa. Centers of plant family richness in the temperate zone at the 10° grid cell included southern east Asia, southern North America, and southern Africa, with lower diversity in Europe and temperate Australia and South America.

Endemism and Range Size Rarity

Endemism (species restricted to specific localities) and taxon rarity have received a similar amount of attention as has biodiversity. How do levels of plant endemism compare among temperate forested regions and with other biomes? Statistics on plant endemism tend to be available for only a limited range of spatial scales (often countywide, occasionally for states or provinces, and infrequently for localities) and are often incomplete or lacking for poorly known regions such as the tropics. Nevertheless, there are some trends in the data that are available. Boreal and cold temperate forested regions in the Northern Hemisphere tend to have very low levels of endemism based either on absolute numbers or the percentage expected on average per unit area (e.g., northern and western European countries (0–2%) and the eastern United States (0 to <1%)). Regions that include warm temperate to subtropical regions with topographically heterogeneous landscapes (and therefore heterogeneous climates and vegetation types), and especially with isolated mountain ranges, tend to have higher than expected endemism per unit area (e.g., Bulgaria (9%), Turkey (31%), western North America, and countrywide for China (56%)). Larger islands and peninsulas tend to have higher than expected species diversity (e.g., Korea (14%), Florida (12%), Japan (37%), and New Zealand (82%)). The Southern Hemisphere forested regions tend to have some of the highest levels of floristic endemism globally (50–80% for Australia, New Zealand, Chile, and South Africa). In part this reflects their geographic isolation. The previous data are for species endemism for the entire flora. Tree species endemism is more difficult to determine and likely to be much lower. For example, in South Africa about 25% of the more than 1000 tree species are confined approximately to the borders (and enclaves) of the country and about 40% to southern Africa (versus 70–80% for the entire flora). Regional tree endemism is much lower (e.g., about 8% for Pondoland and 15% for Maputaland forest regions). However, very high levels of tree species continental endemism are found in southern South America: 85% of woody species and 34% of genera are endemic.

Compared to tropical forests, the absolute numbers of endemics are generally lower in the temperate zone, except perhaps for some of the Southern Hemisphere temperate zone. This reflects the larger floras in the tropics, which are still poorly known for many localities. However, on the basis of expected percentage endemism per unit area the trends are less clear. For example, although Venezuela (38%), Panama (13%), and the Congo (29%) have higher than expected levels of plant endemism, Columbia (4%), Nicaragua (<1%), and Nigeria (4%) have lower than expected endemism. On a subregion or local basis the tropics may well have substantially higher levels of endemism than equivalent localities in the temperate zone, but this pattern remains to be shown conclusively.

Species endemism is only one way to assess rarity or the geographical restrictions of taxa for conservation purposes. Range size rarity of taxa or clades evaluated in the absence of political or other arbitrary boundaries may be a more consistent and robust means of judging rarity. A global survey by the British NHM of range size rarity was done on a 10° grid for selected plant taxa. This showed that there were no clear trends between warm temperate and tropical regions in range size rarity. Many, perhaps the majority of the prime hot spots, are in temperate grids that include forested regions (e.g., southern South Africa, central China, southeastern Australia, and central to south-central Chile, with secondary centers in southern Chile, northern New Zealand, elsewhere in China, and the southeastern United States). Clade rarity and endemism need evaluation as an alternative to enumerating species endemism or range size rarity.

Explanations for Patterns of Diversity and Endemism

A variety of hypotheses have been put forward to explain local and global patterns of species distributions and therefore richness. One author lists as many as 120 named hypotheses for variation in species richness. Geological and biogeographic history provides one set of important related explanations for observed trends in richness and endemism. This may well explain some of the patterns within and between hemispheres.

Historical and Geological Explanations

In the Triassic period global plate tectonic activity had united major landmasses of the world to form a single supercontinent (**Figure 5**) which created the potential for a common biota. However, by the beginning of the Cretaceous this landmass began to break up, forming northern and southern landmasses. This, together with the formation of a broad sea, effectively isolated the northern Asiamerica continent from the southern Gondwanaland. The angiosperm and gymnosperm floras on these separate landmasses thus began to evolve independently. This history is reflected in both the Tertiary floras of the fossil record and the floras we see today. Until approximately 40 million years into the Tertiary the climate was warm, moist, and relatively stable. In the Northern Hemisphere forests spanned North America and Eurasia into the present-day Arctic. In the mid-Cretaceous (100 million years ago (mya)) these forests were dominated by both conifers (Pinaceae, Ginkgoaceae, and Taxodiaceae) and angiosperm taxa. Early examples of widespread flowering plant genera include *Magnolia*, *Betula*, and *Platanus*. Many modern northern genera soon followed in the fossil record (e.g., *Quercus, Castanea*, *Carya, Ulmus, Juglans*, and *Acer*), also spanning the Northern Hemisphere.

By the beginning of the Neogene, approximately 25 mya, the climate began to change, becoming cooler and drier in certain regions as a consequence of mountain building and shifts in ocean currents. In the cold northern regions, conifers were favored, giving rise to the boreal forests. At midlatitude, summer drought and/or winter cold favored deciduous angiosperms. Midcontinent aridity also gave rise to the temperate steppe grasslands and deserts. This, together with Pleistocene glaciation, undoubtedly contributed to divergence in the north temperate forest floras and their biodiversity.

In Europe and central Asia, east–west trending mountain ranges and arid zones blocked the southward retreat of the forests as the glaciers advanced. In consequence, there was apparently greater species extinction here than elsewhere in the Northern Hemisphere, and geographic isolation prevented recolonization from temperate or tropical locations elsewhere. In contrast, in east Asia, continental glaciation was less extensive and continuous connections with tropical wet forests permitted temperate forest species to retreat south and advance north with climate oscillations; there were many fewer extinctions and thus greater retention of an older phylogenetic diversity. This tropical connection undoubtedly accounts for the greater tropical affinities seen in the flora of east Asia today. In eastern North America the mountains trend north–south and thus migrations with the glacial oscillations were possible. However, the tropical connections were largely blocked by seas or arid zones. Compared to Asia, there were more extinctions with fewer opportunities for recolonizations by taxa with tropical affinities. In eastern North America the phylogenetic lineages have a more recent and more northerly balance. Western North America suffered the greatest extinctions. Extensive mountain building, accompanying aridity, and the development of a Mediterranean climate to the south favored conifers or sclerophyllous trees with divergent evolutionary lineages from those of the mesic Tertiary forests. To the north, cool, wet climates favored conifer forests (many with tertiary affinities) at the expense of angiosperms.

The forests of the Southern Hemisphere began similarly in the Cretaceous. Warm, mesic, mixed conifer and angiosperm forests spanned Gondwanaland. These were dominated by southern conifers, including Auracariaceae and Podocarpaceae, and southern angiosperms, including Cunonaceae, Proteaceae, Myrtaceae, and Sapindaceae. In addition to these, Casurinaceae and the genus *Nothofagus* were present throughout, except on what became Africa. Gondwanaland began to break up early in the Cretaceous (**Figure 5**) well before Asiamerica. Southern Africa and India had become isolated by the early Cretaceous. This undoubtedly accounts for Africa's greater dissimilarity with the rest of the southern temperate flora. India eventually joined Asia and today it has primarily north temperate or Asian tropical floristic affinities. Not until early in the Tertiary did South America, Antarctica, and Australia finally separate. Climate changes in the Southern Hemisphere were less dramatic during the Neogene and Pleistocene than they were in the north. Aridity and glaciation were more localized; the climate remained less continental. In consequence, the forest flora retains more characteristics of the warm, mesic Tertiary flora than is the case in the north.

During the Neogene, temperate southern Africa experienced increased aridity. With climatic oscillations the forests expanded and contracted. Many of the Gondwanan elements became extinct, leaving only a couple of species of Podocarpaceae and a minor contribution of Cunonaceae and Proteaceae to the forest flora today. Connections along the east coast of southern Africa allowed connections with the tropical forest flora, and today the temperate South African forests have the strongest tropical affinities of any temperate region and the greatest tree diversity. Exceptionally high levels of spatial and temporal environmental heterogeneity may contribute to speciation here as well. With substantial tropical affinities also found in Australia and New Zealand, the southern forests as a group have a much stronger representation of tropical plant families than do the northern forests.

During the Pliocene the southern temperate forest region of Chile and adjacent Argentina became isolated from the rest of South America by the Andes and arid zones to the north and east. The isolation and relatively small extent of these forests undoubtedly account for the low diversity and high endemism found here. Only in the extreme south of Chile and at high elevations with lower mean and absolute temperatures are deciduous (*Nothofagus*) forests found in the Southern Hemisphere.

Australia and New Zealand retain much of their Tertiary flora. The maritime climate of New Zealand has changed relatively little during the Tertiary. Its long isolation from other

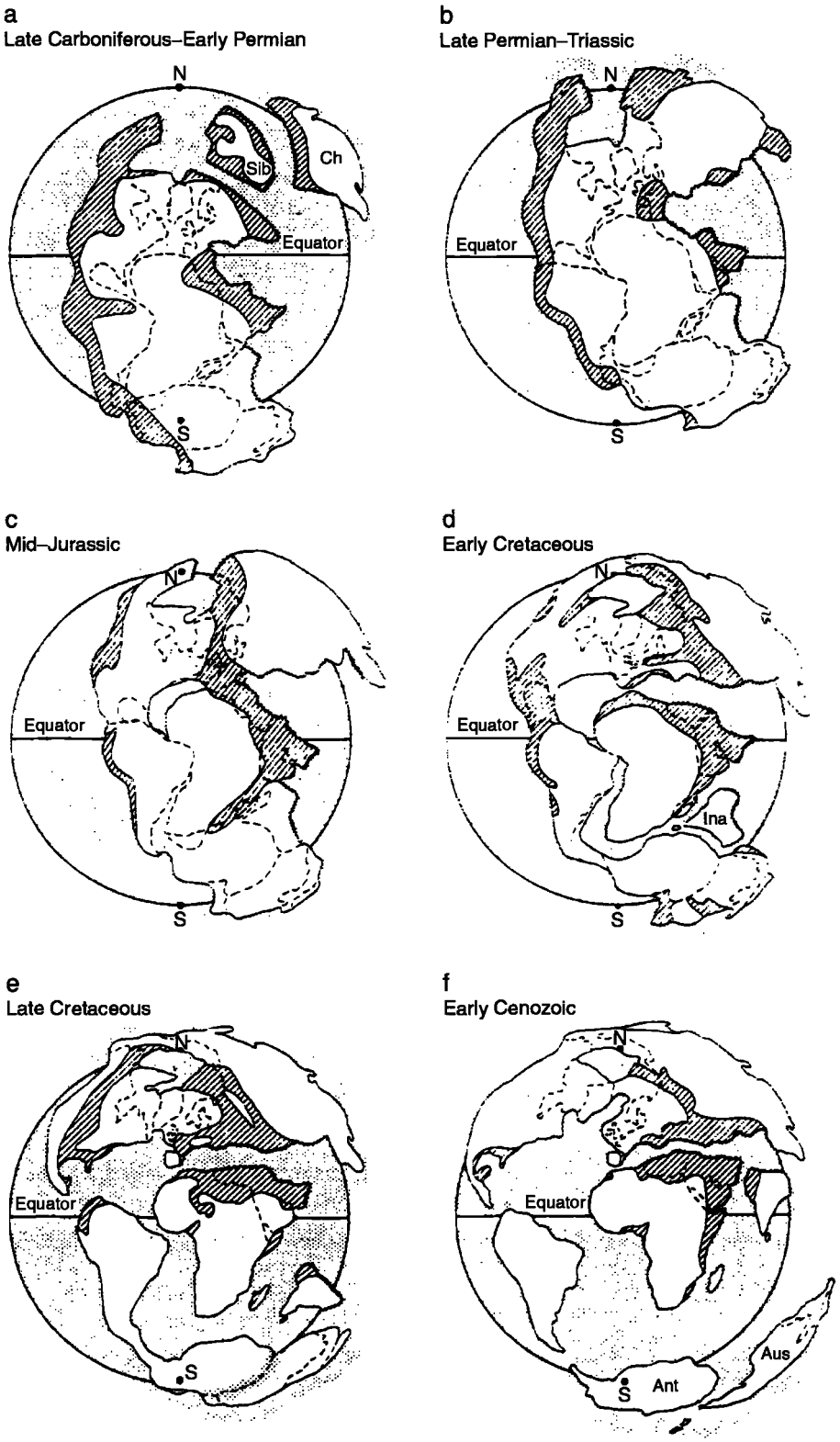


Figure 5 Changes in the locations of continental landmasses are shown from the Mesozoic to the early Tertiary in Lambert equal area projections. Areas on the back projection of the spheres are folded out for view. Dashed lines indicate shoreline of modern continents, hatched areas are epicontinental seas, and stippled areas are oceans. Reproduced from Cox CB and Moore PD (1985) *Biogeography*. Oxford: Blackwell, with permission from Wiley.

landmasses is responsible for the high level of endemism. Increased aridity in Australia in the Neogene substantially affected the temperate forest structure and composition. Explosive radiation occurred in the genus *Eucalyptus* (Myrtaceae) and *Acacia* (Mimosaceae). Unique among temperate forests, there are more than 500 species of *Eucalyptus* in Australia (many restricted to temperate forests). These are sclerophyllous-, drought-, and fire-tolerant tree species that dominate most of the temperate forest and woodlands in the east and west. Only in moist, cool patches in the southeast or at higher elevations are remnants of the tertiary flora found. *Nothofagus*, Araucariaceae, Podocarpaceae, and other groups are part of the rich tree flora.

The previous account indicates that it is not simply history or geology but a complex process that incorporates the role of temporal and spatial heterogeneity, geographic isolation, recent Pleistocene effects, and the role of climate stability that jointly affect the patterns of taxon distributions, diversity, and endemism we see today. What about other explanatory variables?

Productivity

The differential effects of solar radiation figure prominently as an explanatory hypothesis, measured either directly or indirectly via actual or potential evapotranspiration (i.e., AET or PET) or productivity. There is conflicting evidence on the relative importance of this as an explanatory variable versus historical factors, at least at the regional scale. Examined at the plot level ($\sim$1 ha) across regions, correlations are found between diversity and AET. There are also correlations at the regional and latitudinal scales (boreal vs temperate vs tropical forests). However, other evidence shows that differences at regional and local scales can be well explained by historical factors after taking into account AET. Global joint correlations between species richness and productivity were examined by the British NHM for plant family diversity on a 10° grid scale. This study showed higher family richness than expected from productivity in southern Africa, southern North America, and southern east Asia and lower or expected diversity elsewhere in the temperate zone. Undoubtedly both sets of factors may be involved, but these remain only correlates; cause–effect relationships have not been demonstrated.

Spatiotemporal Heterogeneity and Other Explanations

The roles of disturbance regimes and environmental or habitat heterogeneity at local and regional levels have figured prominently as explanatory hypotheses for richness patterns. Higher diversity is predicted under intermediate levels of disturbance. On a local or landscape level there is considerable evidence to support this idea. Midsuccessional forests and successionally diverse landscape support a richer flora. For example, in northeastern North America many of the extinct or threatened species are those that occurred in open or early successional landscapes that were common 100 years ago but have since disappeared with the homogenizing reforestation of the landscape.

The predominant natural disturbances in temperate forest zones include fire and windstorm. Fires are most prevalent in the drier forests of western North America, Australia, South Africa, and the Mediterranean Basin. In many of these systems, periodic fires contribute to higher landscape biodiversity. The role of fires in other forest systems is less clear. Large cyclonic storms affect mainly eastern coastal, midlatitude regions in the Northern Hemisphere (eastern North America and Asia). Tornadic events are most prevalent in midcontinental North America. Such storms contribute to heterogeneous successional stages in the landscape and thus higher diversity. Occasionally, human activity can contribute to higher landscape diversity. Historically, Europeans husbanding the landscape for diverse forest products undoubtedly increased the local diversity over what it would have been naturally. Humans creating a moderate spatially and temporally heterogeneous landscape ironically may contribute slightly to increased regional diversity. The higher regional species richness in mountainous areas associated with the manifold habitats was discussed previously. Another example is the extremely rich South African forests (and other biomes). The high local and regional richness has been largely attributed to the environmental heterogeneity found at this scale.

In addition, there are many other explanatory variables for patterns in species richness, some of which were intimated or included in the topics discussed previously: environmental stability or predictability, abiotic rarifaction, land area, seasonality, aridity, range limits and geometric constraints, and many more. For many of these factors it is easy to establish correlation and much more difficult to establish cause–effect relationships.

Ecosystem Services

Temperate forests provide important ecosystem services globally, regionally, and locally. Temperate forests contribute about 17% to global net primary productivity (versus about 49% for tropical forest systems and 8% for boreal forests). However, recent evidence from atmospheric and oceanic CO_2 data point to temperate forests, especially those in eastern North America, as globally important carbon sinks. The magnitude and causes of the net carbon uptake by temperate biomes are uncertain. However, it may be related to the reforestation that has occurred during the past century, especially in North America. CO_2 fertilization, anthropogenic N deposition, and global warming may contribute as well.

At local to landscape levels there are tight links among forest structure, composition, and species richness; soil attributes; mineral and hydrological cycles; and human disturbances. In the moist temperate zone species richness tends to be higher on soils that are better drained, warmer, and finer textured, with greater NO_3–N and P availability and higher cation saturation and lower Al^+ (toxic) levels, all associated with higher pH. These attributes are associated with many calcareous soils. Thinner calcareous soils, which apparently create greater spatial heterogeneity in forest canopies, are also associated with higher local richness. Hence, soil attributes are intimately related to pattern in species diversity. The extent to which these features are altered will directly affect biodiversity;

the extent to which the community structure and composition are changed will feed back on these variables.

The strongest correlates of productivity tend to be soil texture and N cycling, which in turn are correlated with moisture retention, cation exchange, and maximum levels of humus accumulation. These correlations are strongest in undisturbed forests. Also, ecosystem control by soil texture extends over long time frames. Human disturbances can profoundly affect these links. Soil compaction with intensive forest management will have a cascading and long-term effect on nutrient and hydrological cycles. As the studies at Hubbard Brook Experimental Forest have demonstrated, forest clearing will dramatically alter watershed hydrology (increased water loss and sediment loss), nutrient cycling (with elevated nutrient loss), microclimate, and species composition for years or decades. Changes in forest species composition such as with conifer plantations will have an equally large effect. Substituting homogeneous conifers stands for mixed broadleafed forests will increase C/N and lignin/C ratios, which in turn reduces decomposition rates, N mineralization rates, and pH, with cascading effects on cation exchange capacity and diversity in the soil flora and fauna. For example, in German *Picea* plantations, N cycling between canopy and forest soils is reduced by 75% from that observed in native beech forests. Moreover, with greater stem interception of precipitation, the conifer forest soils also become drier. In consequence, it may not be possible to successively reintroduce beech forests on these sites without large-scale soil amendments.

Even small changes in the broadleaf forest composition can alter decomposition rates and nutrient status. For example, sugar maple (*Acer saccharum*) and ash (*Fraxinus*) promote higher N mineralization, and sugar maple accumulates calcium. In contrast, oak-dominated forests have lower N mineralization. Substituting exotic species in plantations can dramatically alter ecosystem processes. *Eucalyptus* and *Melaleuca* tend to significantly lower water tables where they have been introduced. In the northeastern United States, Japanese barberry (*Berberis thunbergii*) has become a seriously invasive exotic, forming a continuous shrub layer. Under barberry canopies soil ammonium N levels are elevated, the soil flora and fauna are altered, and native species richness is depressed.

Conservation Status

Table 1 shows the approximate extent of protected forests in each of the regional biomes based on WCMC global data for International Union for the Conservation of Nature (IUCN) conservation protection categories I–VI. Globally for temperate forests, 6 or 7% of the remaining forests receive some level of protection. This represents about 1 or 2% of the total temperate forest biome extent. Compared to other forest systems, temperate forests are slightly better protected than boreal forests (6 or 7% versus 5 or 6%) but apparently less well protected than tropical forests as a whole (10–12%).

There is considerable variation both within and among regions in the level of forest protection as well as among forest types. For example, temperate freshwater swamp forests are afforded the least protection (2.7% globally), whereas Southern Hemisphere evergreen broadleaf forests receive the highest level of protection (22.6% globally). Overall, Southern Hemisphere forests are afforded significantly better protection as a group than are northern forests.

The level of protection varies considerably among regions and countries within the temperate zone: Less than 3% of east Asian temperate forests are protected, but this varies with forest type from about 5% for cool to cold temperate and subtropical forests to 1% or less for most warm to hot temperate forests types (except for warm temperate rainforests, which only cover a tiny area and are about 15% protected). There is considerable variation among the countries of east Asia in the amount of forest afforded protection: almost 10% for Japan to less than 5% of China and less than 1% of North Korea. For Europe (excluding Russia) about 8% for the various forest types are given some protection. This also varies considerably among countries: As much as 25% of the broadleafed deciduous forests in Germany are protected, whereas $<1\%$ in Bosnia and 2% in Russia are protected. For North America, the extent of forest protection varies from about 15% for conifer-dominated temperate rainforests to 7% for the northern cool or cold temperate forests and about 2% for the warm to hot temperate and subtropical forests versus more than 9% for North American boreal forests. The forests of the Middle East, which are the most reduced in extent, are afforded the least protection over all. Less than 3% of the forests are protected in this zone. This varies from $<1\%$ in Georgia and Azerbaijan to 1% in Turkey and perhaps as much as 12% in Iran. About 8% of India's temperate forests have some level of conservation protection. For temperate Australia about 9% of the forests are protected; for New Zealand the figure is about 43%. For South Africa, about 24% of its southernmost temperate forests are protected. For South America, about 23% of the temperate forests are afforded some protection, but the vast majority (more than 90%) of these are the wet evergreen forests of the south, mostly located in Chile. Less than 2% of the sclerophyllous and dry temperate forest are protected. An estimated 8% of Brazil's temperate mixed conifer forests are protected.

Overall, the most well protected forests, on a relative scale, are the Southern Hemisphere temperate rainforests, the Pacific wet forests of western North America and east Asia, the northernmost temperate forest (mixed with boreal elements), and mixed temperate forests in mountainous areas. The most poorly protected are dry and sclerophyll leafed forests of the Northern Hemisphere, wet and rainforest broadleaf deciduous and evergreen forests in the Northern Hemisphere (very limited in extent), and moist temperate deciduous and evergreen broad-leafed forests of the Southern Hemisphere. East Asia has the highest percentage of temperate forest types receiving no protection, followed by the Middle East and South America.

In some ways these data are misleading. Many of the forests classified as receiving protection are plantation or otherwise highly managed forests. These are largely native species forests in Europe, North America, and Asia, but they tend to be managed as homogeneous, even-aged, monocultures with consequent reduced diversity and ecosystem services. Natural, undisturbed, or old-growth forests comprise only a small fraction of the remaining forests: western Europe, 1%; eastern North America, 1%; eastern Asia, 1%; Australia, about 5% (but probably $<1\%$ is unlogged); South Africa, $<1\%$.

Somewhat better off are New Zealand with about 25%, northwest (Pacific) North America with about 13%, and southern Chile with about 45%.

Some temperate forests are unique in that significant reforestation has occurred during the past century (about 1 or 2% increase on average per annum), primarily in Europe and eastern North America. However, most of these net gains in forest cover are intensively managed (in Europe and the southeastern United States). In northeastern North America, during the past century the landscape has reverted naturally from 50–90% agricultural to mostly forested, following the extensive abandonment of agriculture. Most of this forest is highly fragmented and fairly homogeneous, being of similar age. Consequently, the ecological functionality of these forests is limited. Elsewhere, temperate forest cover continues to decline at rates that vary from 1 to 10% pa.

Conservation and Protection Strategies

How does one develop a strategy for conserving or protecting the remaining temperate forests of the world? Some authors focus on "hot spots" of diversity or endemism. Globally, most attention has been focused on tropical systems and recently Mediterranean systems. Temperate forest-dominated systems have received less attention. For example, of Conservation International's (CI's) 24 hot spots, only New Zealand is a temperate zone forest-dominated locality. However, in his most recent iteration of "megadiversity countries," Mittermeyer highlights 4 of 17 countries that contain important temperate forest biomes: China, the United States, Australia, and South Africa. Three of these (Australia, China, and South Africa) rank among the top 12 countries worldwide in species richness across phyla and species endemism. The four largely temperate countries also rank among the top globally in the number of IUCN Red Data Book (RDB) "threatened" plant species. Although these predominantly temperate floras are better known, as a group they have more RDB species than all of the tropical countries together. The WWF and the IUCN have also identified approximately 240 centers of plant diversity globally. Of these, about 15% represent temperate forest-dominated systems and another 10% represent other temperate biomes (e.g., shrublands, grasslands, and arid lands). Only 1 or 2% are boreal or polar, and the balance (~74%) are tropical.

There have been similar attempts to identify hot spots regionally. For example, WWF identified more than 100 forest hot spots across Europe, the Mediterranean Basin, and the Middle East. Regional centers of diversity can also be detected from taxon turnover. Data from the Atlas Florae Europaeae project, which compared the joint diversity of Pinaceae and Fagaceae across Europe at a 50-km grid scale, show highest joint species diversity in these two families in the Balkans, the southern Alps, the Carpathians, and the southern Pyrenees. This obviously reflects the high spatial heterogeneity in these localities.

Alternatively, conservation assessments may include such factors as taxon or clade irreplaceability, minimum area sets, minimum viable niche space, or ecosystem integrity. A study was done by the British NHM to select conservation priority areas for selected plant groups. This exercise was done on a 10° grid globally and regionally for all of Europe on a 50-km grid. The global analysis revealed top-priority sites in eastern China, southeastern Australia, and central Chile along with six tropical areas, plus secondary centers in southern Chile, the southeastern and western United States, central and southern China, northern Japan, southwestern Europe, eastern and southwestern Australia, and southern Africa along with nine secondary centers in the tropics. Within Europe, many priority sites were identified in the Balkan Peninsula and margins of the Mediterranean (which include forested lands). However, there are prioritized areas elsewhere throughout Europe at lower densities.

World Resources Institute (WRI) provides another perspective on priority sites for forest conservation. They recognize "frontier forests"—large tracts of intact forested ecosystems sufficient to maintain viable populations of all indigenous species. This perspective includes large, wide-ranging predators and migratory species and takes into account the prevailing natural disturbance patterns. These are thus considered intact, fully functional forested ecosystems or landscapes. Very few of these are in the temperate zone, and most are under medium or high threat. This contrasts with the tropical and boreal zones, in which there are many more large tracts identified as frontier forest, even if there are substantial areas at risk. No intact forested landscapes are found in temperate Europe or Africa. In North America only one small patch of transitional boreal forest in central Ontario and discontinuous tracts of conifer forests in coastal British Columbia and Alaska (all under threat) can be considered intact landscapes. In Asia, only a few small forest patches in the inaccessible mountains of central and south-central China, boreal transition forest patches along the border between China and Russia, and the Primorski Krai region of Pacific Russia are considered intact temperate forest landscapes. All but one small patch are under threat. In South America the only frontier temperate forest is in southern Chile and adjacent Argentina. Most of this region is considered to be at risk. In Australia, only a small rainforest patch in Tasmania is classified as intact. Small patches of frontier forest occur on the west coast of the south island of New Zealand and one patch in the central north island. All are at risk except for the Tasmania rainforest.

Major Threats to Conservation

Many major threats exist for conserving the biodiversity and ecosystems services provided by temperate forests locally, regionally, and globally. Certainly a major threat is homogenization of the landscape. This is a result of intensively managing forests as near monocultures of either native or exotic species, managing landscape for similar forest age and size classes, introducing invasive exotic species, and relying globally on only a few taxa for forest plantations (e.g., *P. radiata* or *Eucalyptus*). The consequences of these landscape management strategies are reductions in local or regional diversity and alternation of many ecosystem processes (e.g., nutrient and hydrological cycles) and soil attributes. Large-scale increases in timber harvesting may well have a

negative impact on carbon source–sink relationships and atmospheric CO_2 levels.

Increasing urbanization of the landscape and continued harvesting of ever more remote forests will lead to a more fragmented landscape. The likelihood of these fragments maintaining ecosystem functionality is ever diminished. Within the next few decades, it is likely that no fully intact, functional forested landscapes will remain in the temperate zone, given current trends.

Atmospheric pollution from acid precipitation, and the associated ecosystem N and S loadings, remain prevalent in the north temperate zone and will have long-term effects. There is evidence that calcium is being rapidly depleted in acidic soils as a consequence of long-term acid precipitation. This will have cascading effects on forest soils and on major tree species performance, particularly those most sensitive to calcium levels such as sugar maple.

Projected global warming will have a major effect on northern temperate forests. The main direct effect of global warming will likely be seen in the boreal forest zone, with decreased periods of snow cover and hence changes in surface albedo. This will likely favor the expansion of temperate forests well north into the boreal zone. The community response patterns will be determined by dispersal characteristics of the biota and the availability of propagules in source populations. This in turn will be dictated by the configuration of forest fragments in the landscape.

In the Southern Hemisphere, the effects of global warming will be much smaller due to the moderating maritime influences. Atmospheric pollution effects are very negligible or very localized. The major threats to temperate forests here are deforestation, especially in South America, and the pervasive effects of introduced exotic species in plantation or as escapes. *Eucalyptus* and northern pines (especially *P. radiata*) are planted in monocultures across the Southern Hemisphere, with negative effects on biodiversity, soil attributes, and hydrology.

Conservation Objectives and Research Needs

Where do we need to focus attention to improve our understanding of temperate forests and develop a more effective conservation plan? Clearly we need better and more accurate means of mapping temperate forests and assessing their conservation status. There are large discrepancies and errors in assessing the current and potential forest cover for all forest types. We need more information on patterns of species occurrences (simply the presence/absence is sufficient) at a variety of scales throughout the temperate zone. Because distribution (and hence richness) patterns vary across spatial scales, one cannot simply rely on particular, arbitrary sample sizes (e.g., 0.04-, 0.1-, 1-, or 50-ha plots). Effort should be placed on inventorying poorly known areas, such as the Middle East. We need to have a better, predictive understanding of the links between forest structure and composition, ecosystem functions, and the effects of human disturbances.

There needs to be a concerted, manifold effort at developing a variety of different conservation strategies for temperate forests. These include targeting more protection for forested biomes with high ratios of people to forest area and forest biomes or forest types that are poorly protected. Examples include east Asian forests in general, north temperate broad-leafed rainforests, and subtropical dry forests. Moreover, concerted effort should be placed on protecting the few remaining intact frontier forests, especially those at high risk, and old-aged forest stands particularly where these comprise miniscule components of the landscape. However, the strategy needs to be inclusive, focusing on conserving the complete spatial and temporal heterogeneity of the landscape, even if this necessitates some human husbandry of the landscape.

Strategies for identifying priority conservation sites should not simply target sites with high species diversity and/or endemism. Alternative means of evaluation need to be incorporated as well, such as phylogenetic richness and irreplaceability, range rarity, ecosystem or landscape integrity, landscape heterogeneity that incorporates migration in the face of natural and human-made disturbances, and minimum niche space for all biotic components of the landscape. Only with a manifold creative approach can we hope to come to grips with the conservation of this critical suite of temperate forest biomes.

See also: Boreal Forest Ecosystems. Deforestation and Land Clearing. Endemism. Forest Ecology. Grassland Ecosystems. Hotspots. Tropical Forest Ecosystems

References

Armesto JJ, Rozzi R, and Caspersen J (2000) Past, present, and future scenarios for biological diversity in South American temperate forests. In: Chapin T and Sala O (eds.) *Future Scenarios for Biological Diversity*. New York: Springer-Verlag.

Bailey RG (1998) *Ecoregions: The Ecosystem Geography of the Oceans and Continents*. New York: Springer-Verlag.

Cox CB and Moore PD (1985) *Biogeography*. Oxford: Blackwell.

Dudley N (1992) *Forests in Trouble: A Review of the Status of Temperate Forests Worldwide*. London: Earth Resources Research.

Ehrlich PR (1996) Conservation in temperate forests: What do we need to know and do? *Forest Ecol. Management* 85: 9–19.

Francis AP and Currie DJ (1998) Global patterns of tree species richness in moist forests: Another look. *Oikos* 31: 598–602.

Latham RE and Ricklefs RE (1993) Continental comparisons of temperate-zone tree species diversity. In: Ricklefs RE and Schluter D (eds.) *Species Diversity in Ecological Communities. Historical and Geographical Perspectives*. Chicago: University of Chicago Press.

Ovington JD (1983) *Ecosystems of the World 10: Temperate Broad-Leafed Evergreen Forests*. Amsterdam: Elsevier.

Qian H and Ricklefs RE (1999) A comparison of the taxonomic richness of vascular plants in China and the United States. *Am. Nat.* 154: 160–181.

Rhode K (1992) Latitudinal gradients in species diversity: The search for the primary cause. *Oikos* 65: 514–527.

Röhrig E and Ulrich B (eds.) (1991) *Ecosystems of the World 7: Temperate Deciduous Forests*. Amsterdam: Elsevier.

Schultze E-D, Bazzaz FA, Nedelhoffer KJ, Koike T, and Takatsuki S (1996) Biodiversity and ecosystem function of temperate deciduous broad-leafed forests. In: Mooney HA, Cushman JH, Medina E, Sala OE, and Schultze E-D (eds.) *Scope 55: Functional Roles of Biodiversity: A Global Perspective*. Chichester, UK: Wiley.

Terrestrial Ecosystems

Stephen Roxburgh and Ian Noble, Australian National University, Canberra, ACT, Australia

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 637–646, © 2001, Elsevier Inc.

Glossary

Biome A group of ecosystems which are subject to similar climatic conditions and which share a similar range of vegetation structures, animal diversity, and soil types. Biomes represent major regional groupings of ecological systems and are therefore discernible at the global scale.

Ecosystem All of the organisms living within or utilizing a particular area, together with their surrounding nonliving environment.

Ecosystem classification A procedure for grouping together similar ecosystems based on a predefined set of criteria, e.g., vegetation structure, climate, land-use history, and soil type.

Latitudinal biodiversity gradient The well-documented trend for there to be more species in tropical ecosystems compared with temperate or polar ecosystems.

The Earth's Terrestrial Ecosystems

Ecosystem Classifications: Overview

Terrestrial ecosystems cover approximately 148 million km^2, corresponding to 29% of the total surface area of the earth. They include such diverse habitats as the frigid regions around the poles, the searing heat of tropical deserts, and lush temperate and tropical rainforests. This enormous variety poses great difficulties for the study of ecosystems at regional to global scales, and it has led to the development of many classification systems which seek to reduce this complexity to a manageable level. Indeed, the development of such classification systems, and the identification of the spatial extent of the Earth's major ecosystem types, is a fundamental first step in managing the world's biological systems.

Despite the importance of this problem, the classification of the world's ecosystems into a consistent and useable framework is not without its problems. First, ecosystems usually grade into one another, hence making it difficult to mark ecosystem boundaries on a map. Second, there is no single agreed upon definition of an "ecosystem." To some researchers it represents an area within which the same broad mix of species can be found. To others it is merely an arbitrary area delimited solely for practical purposes. Third, different criteria can be used to generate different kinds of classifications. For example, human-induced land-use change is included in some classifications but not in others. Ecosystem classification systems are therefore approximations of reality and should be viewed as a set of tools which provide a simplified overview of the relative extents and spatial arrangements of the world's major ecosystem types.

Ecosystems are usually classified according to the dominant vegetation type for a region, which is in turn determined largely by geographical and environmental factors. Vegetation is considered the most appropriate criteria for classifying ecosystems because it constitutes the majority of biomass and also because of its role, through the process of photosynthesis, as the primary fixer of energy on which all other life depends.

The Biome Classification

The broadest units of ecosystem classification are called biomes. Biomes are complex mixtures of a large number of similar ecosystem types which share a similar physiognomy (vegetation structure) and climate but do not necessarily share the same species. Indeed, one of the strengths of the biome approach is that ecosystems with different species compositions, but with similar vegetation structures and ecosystem processes, can be compared. For example, the tropical rainforests of Central America and northern Australia are both recognized as belonging to the same biome "tropical rainforest," even though there would be little overlap in species composition between the two regions. The major climatic variables on which the biome classification is based are average annual temperature and average annual precipitation.

Many biome classifications have been suggested. These have classified the world into as little as 8 and into as many as 18 separate biomes. **Figure 1** shows the classification of Cox and Moore, which maps 10 major terrestrial biomes. One feature of the spatial distribution of the world's biomes is that many are fragmented over the surface of the earth. For example, the "Mediterranean" biome is found in the areas surrounding the Mediterranean sea and also in Australia, South Africa, South America, and North America.

A brief description of the major terrestrial biomes is given in the following sections. More complete accounts of each can be found in the appropriate sections of this encyclopedia.

Tundra

Tundra occurs around the Arctic Circle at latitudes above where trees can survive and also at high altitudes above the tree line in other regions. The vegetation is of a very low stature, typically comprising low shrubs, sedges, grasses, lichens, and mosses. The climate is characterized by long, cold winters, during which average monthly temperatures seldom exceed 0 °C and, at the highest latitudes, there is constant darkness for weeks at a time. Soils are permanently frozen (permafrost), making water unavailable to plant growth.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00143-X

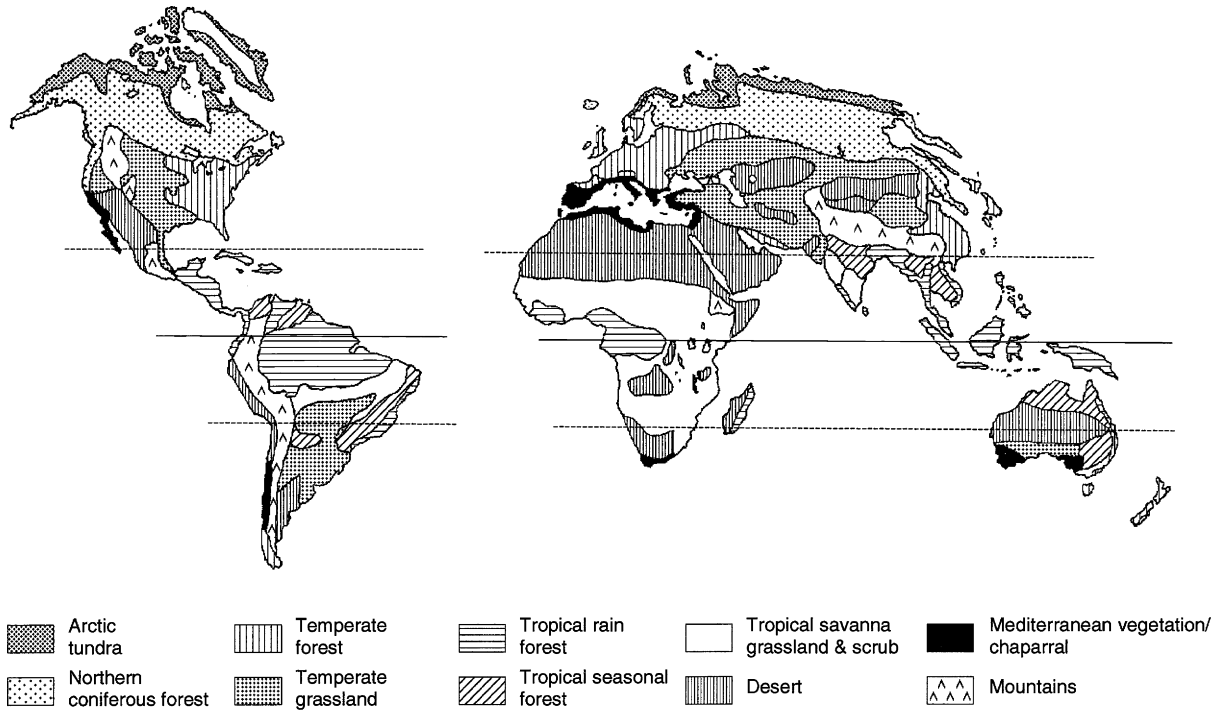


Figure 1 The world's major biomes, based on the classification of Cox CB and Moore PD (1985) *Biogeography. An Ecological and Evolutionary Approach.* Oxford: Blackwell, with permission from Wiley.

Also, liquid water is available for only short periods at a time, thus making these ecosystems effectively "polar deserts." Due to these harsh climatic conditions and brief growing season (usually less than 60 days), plant growth and decomposition is slow and ecosystems are very slow to recover from disturbance. Animal diversity is similarly low, with a small number of permanent bird and mammal species and a number of seasonal insects. Many bird species and some mammals migrate to the tundra during the brief summer.

Taiga (Boreal Forest and Northern Coniferous Forest)

The taiga forms a broad band adjacent to the tundra biome, extending across North America, northern Europe, and northern Asia. The dominant vegetation consists of evergreen conifers, with some representation of deciduous broadleaved species in some environments. Within this biome it is common to find large areas of land dominated by only one or a few coniferous tree species. Winters are typically long and cold, contrasting with summers that, although brief, can be both warm and wet. The soils are characterized as being nutrient poor and acidic. Characteristic animals include insect- and seed-eating birds, small mammal herbivores such as the snowshoe hare, larger browsers such as elk and moose, and carnivorous mammals such as the lynx, wolf, and grizzly bear. Many bird species migrate to the taiga during the summer to feed on swarms of seasonally breeding insects.

Temperate Forests

The temperate forest biome includes a wide range of forest types, including the temperate rainforests of the Pacific Coast of North America, southern Chile, and New Zealand; the temperate deciduous forests of eastern North America, eastern Asia, and Europe; and temperate evergreen forests such as the sclerophyllous eucalyptus forests of Australia and the southern beech forests of Chile and New Zealand. The forest structure is generally well developed, with various strata including emergent trees, canopy trees, small trees, shrubs, and a herbaceous layer. Along with an enhanced diversity of plant species and habitat structure, there is an associated general increase in the diversity of the fauna. The growing season is approximately 6 months. Because of the wide range of forest types and environments represented in this biome, soil nutrition, precipitation, and water availability are also highly variable.

Tropical and Temperate Grasslands

The grassland biome includes ecosystems from a wide range of geographical regions and environments. In the tropics these include the savanna, which can also be associated with an open cover of shrubs or trees. Examples of temperate grasslands can be found in North America (prairies), Europe and Asia (steppe), South America (pampas), and South Africa (veld). A common feature is the importance of herbivory, particularly in the savanna, which has a high diversity of ungulate browsers. In comparison with tropical grasslands, temperate grasslands support a much lower faunal diversity. Fires are also an important component in many grassland ecosystems, in which repeated fires can prevent the establishment of woody species, thus preventing the grasslands from developing into shrublands or forests. Climate and soils vary with grassland type and region, with the tropical savanna experiencing distinct wet and dry seasons and temperate grasslands experiencing cold winters and hot summers.

Mediterranean Vegetation (Chaparral)

Examples of the Mediterranean biome can be found between approximately 30° and 40° latitude in both hemispheres and are associated with wet winters and dry summers. Mediterranean vegetation occurs most extensively on the coast surrounding the Mediterranean Sea, but it is also found as the "chaparral" in California and northwest Mexico, the "matorral" in Chile, the "fynbos" around the Cape of South Africa, and also in the southern regions of Australia. Because precipitation in this biome is generally less than in grasslands, and there are regular periods of drought, the vegetation is characterized by shrubs and small tress which have small, leathery "sclerophyllous" leaves adapted to drought stress. Fire is also a common feature of Mediterranean vegetation, and many plant species share adaptations to persist under a regime of repeated fire, e.g., resprouting from belowground and seeds that require fire to germinate. In general, soils are nutrient poor. One striking feature of this biome is that high levels of endemism are common, but adaptations to the Mediterranean environment in geographically separate regions are very similar. This has been seen as evidence for convergent evolution, in which plant and animal species from different taxonomic groups respond independently, but similarly, in response to the same environmental pressures.

Deserts

Deserts occur in climates characterized by extreme aridity. Water limitation in these areas can arise due to many factors, for example, being located in hot tropical regions in which evaporation exceeds precipitation such as the Sahara desert; rainshadow regions in which the presence of a mountain range influences precipitation patterns, resulting in dry conditions on the leeward side of the range, such as in the Patagonian desert in Argentina; and coastal deserts in which prevailing winds prevent precipitation occurring, such as in the Namib desert in southern Africa. Continental-scale effects can also be important, where the interiors of major continents are too distant from moist airflow to receive sufficient rainfall. Vegetation in deserts is sparse, with shrubs being the most common life-form. Deserts are also characterized by succulents (e.g., cacti) and many ephemeral annual and perennial species which have dormancy as some part of their life cycle, enabling them to actively grow and reproduce when rain does occur. The animals of desert regions have also evolved in response to the harsh conditions. Examples include nocturnal activity to avoid the heat of day and dormancy during the periods between rainfall events.

Tropical Forests

This biome includes tropical seasonal forests, which occur in humid tropical climates with a pronounced dry season, as well as tropical rainforests. Tropical seasonal forests are characterized by a mix of evergreen and deciduous species, with some forest types dominated solely by deciduous species. In contrast, tropical rainforest thrives in regions in which rainfall is abundant and evenly distributed throughout the year and temperatures are uniformly warm. Tropical rainforests are noted for supporting the highest recorded diversities of both plants and animals. The soils of tropical forests are typically nutrient poor, with decomposition being rapid and with the nutrients stored mostly within the plant biomass.

Other Classification Systems

Although large-scale "biome" classification systems are useful in gaining an appreciation of the spatial extent of the world's major ecosystem types, for many applications a finer resolution is often required. Two approaches which aim to increase the detail but still retain a manageable number of categories are considered.

First, Olson's classification is an attempt to map the world's ecosystems as they were in 1980. Included in the classification are natural or seminatural ecosystems (e.g., "Tundra" and "Tropical dry forests") as well as those ecosystems that have undergone human modification (e.g., "Warm crops" and "Paddyland"). Olson's classification recognizes 44 different ecosystem complexes compared with the 10 major biomes in **Figure 1**.

A different approach was adopted by Holdridge in his "life zone" classification. This classification aims to predict the potential vegetation that an area could support rather than what ecosystems might actually be occurring there. The classification is based on two main variables, labeled "Biotemperature" and "Annual precipitation." Annual precipitation is the average annual amount of rainfall a region receives, without taking into account its seasonal pattern. Biotemperature is the average annual temperature, excluding temperatures below 0 °C and above 30 °C. Note that these two climatic variables are essentially the same as those used to define the biomes discussed previously. A third variable in the Holdridge classification, "Potential evapotranspiration ratio," combines precipitation and biotemperature into an index which summarizes the water availability for a particular region and quantifies the balance between the amount of water entering the system through precipitation and that lost through the combined processes of evaporation and transpiration. Evaporation is the physical process whereby water is returned directly to the atmosphere from plant and soil surfaces, and its rate is directly influenced by the amount of incoming solar radiation. Transpiration is the biological process whereby soil water is taken up by plant roots, transported through the plant's vascular systems, and evaporated from the leaf surfaces through pores called stomata. Potential evapotranspiration ratios less than 1.0 indicate an excess of water in the ecosystem, and ratios greater than 1.0 indicate a deficit.

The classification produces 38 major life zone categories, ranging from hot-wet tropical forests and hot-dry temperate and tropical deserts to cold-dry polar deserts (**Figure 4**).

The Holdridge classification has been used in global vegetation models to predict major changes in ecosystems types in response to global climate change, and Olson's classification has been similarly used to model the effects of global change, and in particular the dynamics of the global carbon cycle and carbon storage.

Global Patterns of Biodiversity Across Terrestrial Ecosystems

Description of Known Patterns

One of the best documented patterns of biodiversity is the tendency for tropical ecosystems to harbor more species than

temperate or polar ecosystems, i.e., there is a gradient in diversity stretching from lower latitudes (high diversity) to higher latitudes (low diversity). Some examples of different terrestrial organisms which show this pattern are shown in **Figure 2**. The trend is also seen in a wide range of marine organisms.

It is worth noting that these examples include only a subset of the total biodiversity for any given region, e.g., land-breeding birds or swallowtail butterflies. Furthermore, none of these examples are global in extent, with two of the four (landbirds and mammalian quadrupeds) being based on data from only a single continent.

The reason for the focus on small groups of species, at subglobal scales, is simply because the data required to perform global-scale, multispecies studies of biodiversity are still incomplete. Nevertheless, despite some minor exceptions, the majority of studies that have investigated latitudinal trends in diversity have shown that tropical regions are indeed characterized by high levels of species diversity.

Although distributional information for groups of individual species at the global level is in general poorly known, that for the higher taxonomic groupings (e.g., families) is considered more complete. One example of this approach, compiled by Williams and coworkers, is shown in **Figure 3**. It shows the global distribution of the combined number of families of flowering plants, amphibians, reptiles, and mammals. Together, these four groups of families are considered to

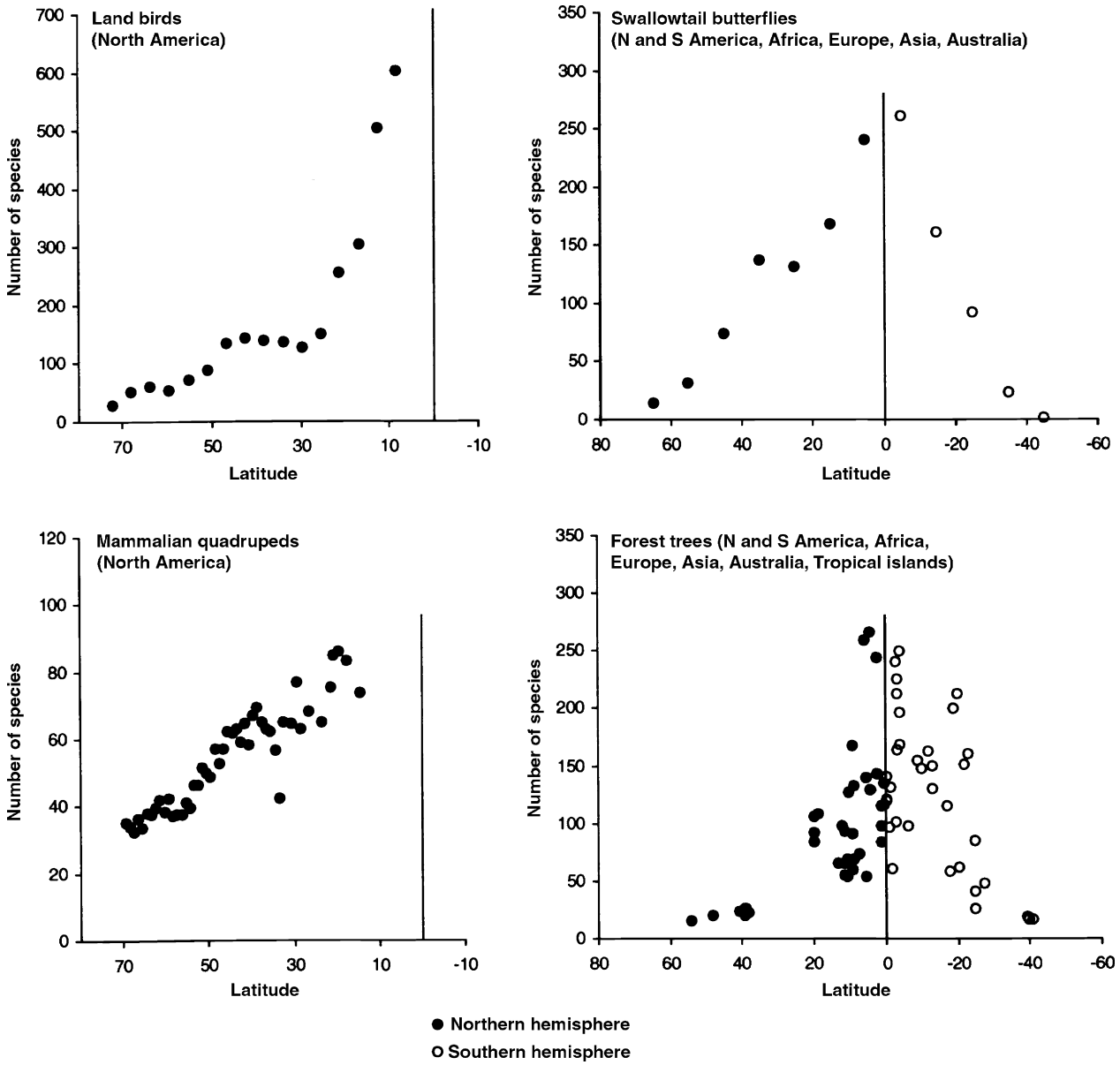


Figure 2 Examples of the latitudinal gradient in diversity, showing higher species richness near the equator (0° latitude) for all species groups. Swallowtail butterfly data redrawn from Sutton and Collins (1991). Land bird data-points represent diversity within areas approximately 230,000 km^2 (MacArthur RH and Wilson EO (1967)). Mammalian quadraped data-points represent average diversity within areas approximately 62,500 km^2 (Rosenzweig ML (1995)). Forest tree data-points represent diversity within 0.1 ha plots (Gentry AH (1988).

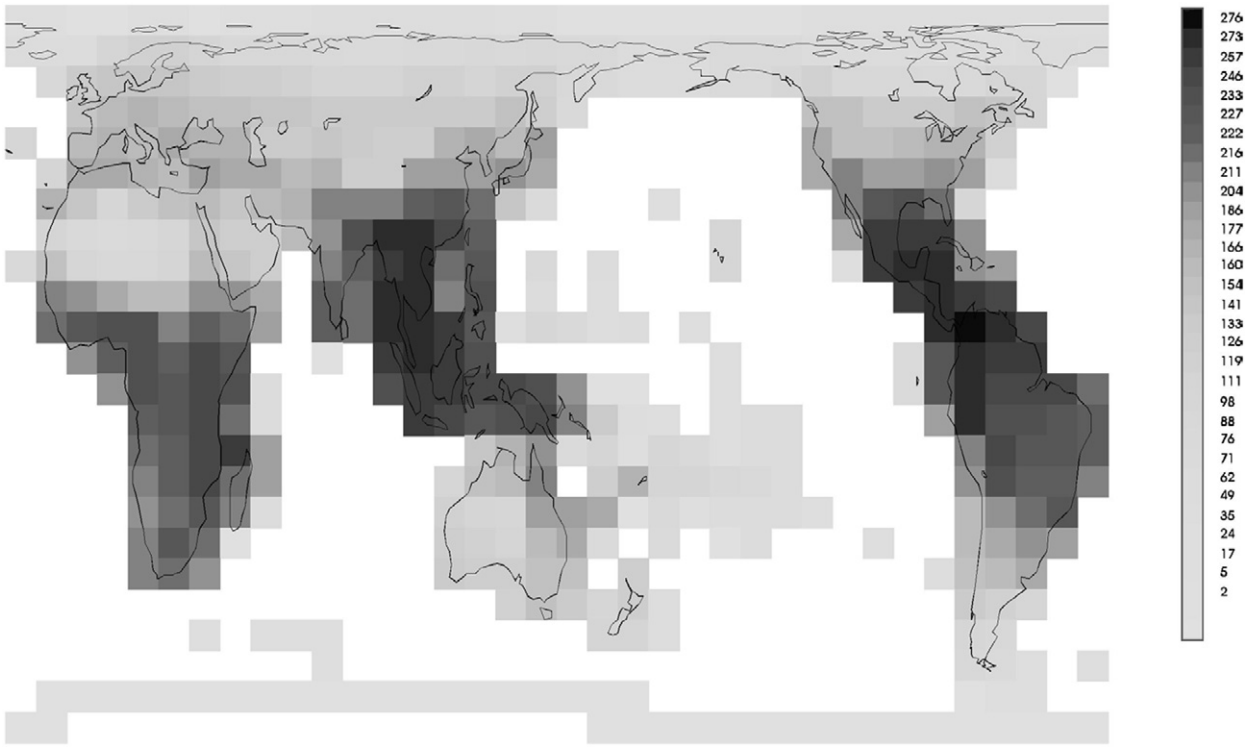


Figure 3 The global distribution of the combined family richness of flowering plants, amphibians, reptiles, and mammals (Reproduced from Williams PH, Gaston KJ, and Humphries CJ (1997) Mapping biodiversity value worldwide: Combining higher-taxon richness from different groups. *Proc. R. Soc. London B* 264: 141–148.). Grid cells are approximately 611,000 km^2.

be the most completely known. It has been estimated that for the chordates 90% of all living species have been identified and for plants 84.4%. These values are far in excess of any other taxonomic group. **Figure 3** shows that the trends at the family taxonomic level mirror those of individual species, with a concentration of families in the tropics, particularly around Central America and Southeast Asia, decreasing toward the higher latitudes in each hemisphere.

Figure 4 shows the relationship between the distribution of the family richnesses in **Figure 3** relative to the life zone classification of Holdridge. There is a consistent pattern, with warm/wet tropical ecosystems containing many more families than either the hot-dry (desert) or cold-wet (polar) ecosystems. Incorporating the spatial information of **Figure 3** into the Holdridge classification allows the trends in family richness across different ecosystem types to be displayed visually. The advantage of this approach, over traditional richness vs latitude relationships, is that different ecosystem types occurring at the same latitude but with different richnesses are able to be separated. For example, tropical rainforests and the Sahara desert both occur at tropical latitudes, but the latter is conspicuously lower in diversity.

A second major trend in terrestrial biodiversity is a decrease in diversity with increasing altitude. As altitude increases there is an associated change in the environment that is analogous to the changes that occur when moving from tropical to temperate latitudes. For example, temperature decreases with increasing altitude, and the total land surface area also tends to decrease with altitude (simply because mountains tend to be broader at the base than at the peaks). Because these changes appear to mirror those observed in the tropical to temperate gradient, many researchers have suggested that the factors controlling the pattern of lower diversity at higher altitudes are likely to be the same as those controlling the latitudinal gradient.

Possible Explanations

Although many plausible hypotheses have been suggested to explain the observed latitudinal trend in diversity, many of them have proved to be controversial and there is no consensus on which of the explanations, if any, are of primary importance in determining actual diversity gradients. A major source of the difficulty is that most of the explanations are based on correlative evidence, and it is well-known that correlation does not necessarily imply causation. Correlative evidence predominates because appropriate experimental tests are extremely difficult to devise at the spatial and temporal scales required.

For any given area on the earth, the number of species present (S) is due to many processes acting over a range of spatial and temporal scales. These processes can be summarized into three broad categories and combined into the following simple formula:

S = species which have immigrated from outside the area (dispersal)

+ species which have arisen through speciation

− species which have been lost due to extinction.

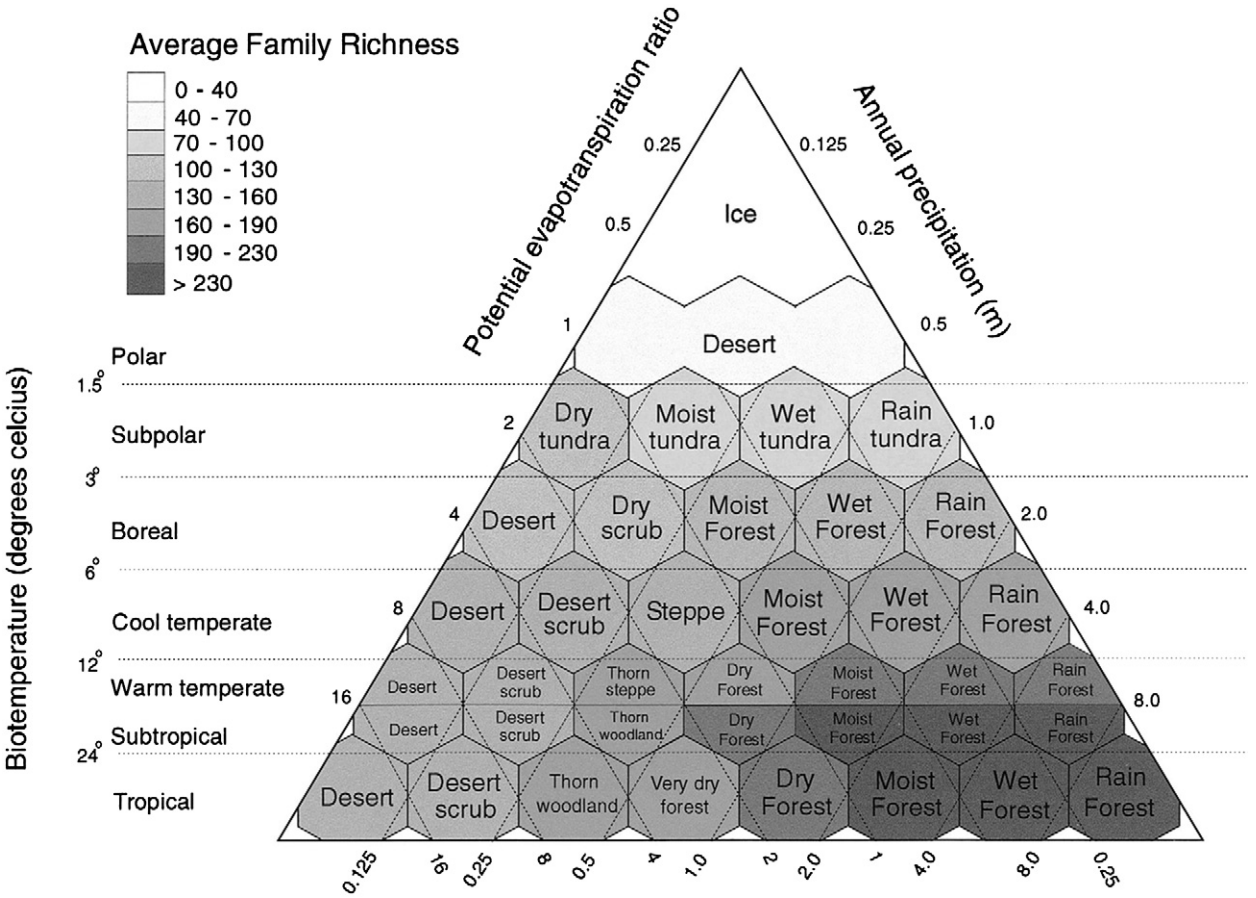


Figure 4 The global distributions of combined family richness incorporated into the Holdridge classification, which predicts 38 vegetation types based on the climatic variables 'biotemperature' and 'annual precipitation', as explained in the text. The scale is average combined family richness (flowering plants, amphibians, reptiles, and mammals) per 611,000 km^2 grid cell calculated for each of the major life zones. Reproduced with permission from Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge Univ. Press.

At large spatial scales (e.g., regional or between biome), dispersal can be considered of limited importance compared with speciation or extinction. Therefore, the hypotheses which aim to explain the latitudinal gradient in diversity must have, as a basis, processes which enhance speciation rates and/or reduce extinction rates in the tropics or, alternatively, which reduce speciation rates at higher latitudes and/or increase extinction at those latitudes. The hypotheses that have been suggested to explain the enhanced diversity of the tropics are described within this broad framework.

Time

One of the earliest explanations for the latitudinal diversity gradient is that tropical regions are more ancient, and therefore less disturbed, than their temperate counterparts. The greater age of the tropics implies that there has been more time for speciation to occur, hence a longer time for species to accumulate. Particular emphasis was placed on glaciations because it was assumed that temperate regions are impacted upon by glacial activity more so than the tropics. However, the fossil record does not support this view. For example, there is evidence that the latitudinal gradient for angiosperms has existed for at least 110 million years. There is also evidence that during the most recent glaciation both temperate and tropical forests decreased substantially, with tropical forests being replaced by grassland and savanna. This suggests that tropical regions may not be as ancient or stable as once thought.

However, if the tropical rainforest remnants were more numerous, and their destruction less complete than in the temperate zone, then this continual isolation and reforming of tropical communities might promote diversity through the mechanism of allopatric speciation. Allopatric speciation is considered by many to be the most important of the speciation-promoting mechanisms. It occurs when a geographical barrier subdivides a population, such that dispersal and genetic exchange between the subpopulations is prevented. These sub-populations continue to evolve in isolation, eventually diverging enough so that two new species are produced. In this scenario, higher diversity in the tropics might be promoted due to enhanced allopatric speciation due to historical geological events, possibly combined with an increased extinction rate in the temperate regions.

Area

Calculated on a land area basis, tropical and subtropical ecosystems are approximately five times more extensive than

temperate ecosystems and approximately twice as extensive as boreal and tundra ecosystems. The reason is due to the earth being approximately spherical so that a zone of a constant width running around the equator (the tropics) will contain much more area than a zone of equivalent width running around the earth's surface closer to the poles. Therefore, tropical ecosystems, which contain the highest diversities, are also those which cover the largest land area.

Rosenzweig suggested three reasons why a greater land area in the tropics might result in greater diversity. They are all based on the assumption that the geographical ranges of tropical species are, or were in the past, larger than the ranges of species in other latitudes. First, a larger geographical range leads to a larger total population size, and a larger population size would be expected to buffer the species against accidental extinction. Second, a larger geographical range contains a greater number of potential refuges where the species might survive following catastrophic disturbances or climatic events, again buffering the species against accidental extinction. Third, a larger range has more chance of being subdivided by a geographical barrier, hence increasing the probability of allopatric speciation.

Habitat Heterogeneity

All other things being equal, larger areas also contain a greater variety of habitats than smaller areas, and hence they are capable of supporting more species, i.e., there is opportunity for the evolution of species to fill these available habitats. Because tropical ecosystems tend to be larger in extent than those in other regions, it follows that they should contain a greater number of species.

It has also been suggested that the increase in the diversity and complexity of vegetation as one approaches the tropics can enhance diversity by providing animals with a greater variety of arboreal habitats. This increased faunal diversity can enhance diversity because many of these species can become a resource on which species higher in the food chain can be supported. Although the increased vegetation complexity in the tropics potentially explains the increased diversity in animal species, it cannot explain the high plant diversity because there is no evidence to suggest that plant habitats are more diverse in tropical regions.

Despite many studies which illustrate a positive relationship between habitat heterogeneity and diversity, care must be taken with the interpretation of these results because habitats are often defined by the species that occur within them, hence leading to circular reasoning. To be a valid hypothesis, habitat heterogeneity must be defined independently from the species occupying those habitats.

Solar Energy

Tropical regions receive, per unit area and per unit time, greater amounts of solar radiation than any other ecosystems. This is again due to a spherical Earth, whereby light energy at higher latitudes intercepts the earth's surface at a more oblique angle compared with the tropics. Although the amount of incoming solar radiation directly affects productivity, and relatedly, the water balance of a region, it also has the potential to accelerate the rate at which new species evolve.

An acceleration of the evolutionary process in tropical regions can arise due to two factors. First, the greater amount of available energy might explain the tendency for tropical species to have shorter generation times. A shorter generation time means more offspring per unit time, hence faster evolution. Second, increased solar radiation would be expected to increase mutation rates, thereby increasing the genetic variability on which natural selection acts, leading to an increase in the rate of speciation. These two factors combine so that over evolutionary time more species would be expected to have evolved in tropical regions rather than those at higher latitudes.

Species Interactions

Species interactions, and in particular competition, are central to many hypotheses addressing the latitudinal diversity gradient. A common ingredient to these hypotheses is that there are factors in the tropics which allow competitive exclusion to be avoided or which allow shifts in the "ecological niche" of a species so that competition is no longer intense, enabling the coexistence of many species. The ecological niche is a multidimensional summary of the environmental ranges and resource tolerances of a species. Species with broad niches are considered generalists, and species with narrow niches are considered specialists. Most important, the amount of niche overlap between species is a measure of the amount of competition experienced. For example, two species with overlapping niches require the same range of resources and environmental conditions; hence, they would be expected to compete strongly, and in the absence of any other factors the weaker competitor of the pair would be expected to be eliminated by the stronger. This is known as the "competitive exclusion principle."

It has been suggested that competition between species in the tropics is of greater importance, and is potentially more intense, than that in temperate regions. The reason given is that the tropical environment is more predictable and less prone to large disturbance events which cause indiscriminate mortality; therefore, population sizes are greater so that species interactions are more likely. High competitive pressure and a stable environment allow natural selection to narrow the niche requirements of the species, allowing the environment and resources for which they are competing to be partitioned to a finer degree than otherwise possible. This would result in more species coexisting per unit area of habitat in the tropics than in temperate areas. This hypothesis assumes that competition is not as important in temperate and higher latitudinal regions, and that in these regions the physical environment plays a greater role in forcing natural selection and evolution. The validity of this assumption remains an open question.

A second hypothesis assumes there are more predators (including herbivores) in the tropics. Predators remove prey biomass, thereby reducing competition among the prey species by reducing the demand on available resources. Under this hypothesis, competition is also less intense in the tropics, not because the species have evolved to avoid it through partitioning the environment more efficiently but because prey biomass is reduced to such an extent that competition is significantly reduced. The assumption that predation is

more important in tropical ecosystems also remains an open question.

Productivity

Productivity is defined as the rate at which biomass is accumulated per unit area. Many hypotheses have attempted to relate the productivity of different ecosystems to diversity; however, they are not always easy to disentangle from one other and from other hypotheses (e.g., the solar energy hypothesis).

One line of reasoning suggests that low-productivity environments (e.g., those at higher latitudes) can support only a limited resource base (e.g., a low-productivity environment might be able to support a herbivore species but at a population density too low to support a predator higher in the food chain). In contrast, in a high-productivity tropical environment a herbivore population might be large enough to support several predatory species. Also, in high-productivity environments food might be more abundant, allowing the opportunity for predators and herbivores to specialize more finely, resulting in a greater number of species (see the discussion in the previous section on the ecological niche).

Another suggestion is that although total productivity over the course of a year might be higher in the tropics than that in other regions, instantaneous rates at other latitudes are much higher (e.g., the flush of plant growth during spring at more temperate latitudes). This effect, combined with low nutrient availability in the tropics, would result in overall lower population growth rates in tropical species, which would tend to slow rates of competitive exclusion and hence extinction. Although slow rates of competitive exclusion cannot explain enhanced diversity on their own, when population sizes are continually impacted upon by external factors such as disturbances, under certain conditions the slow competitive exclusion can be converted into species coexistence, thus promoting higher diversity.

Summary

There are consistent trends in global biodiversity across the various ecosystem types which broadly correlate with latitude. Many hypotheses have been proposed to explain this gradient; however, the spatial and temporal scales over which the processes of speciation and extinction take place make rigorous testing of these hypotheses extremely difficult. As a result, there is still no consensus on which of the hypotheses might be the most important in determining actual gradients. A further difficulty in identifying the primary causes of the latitudinal diversity gradient is that many of the hypotheses are intimately related, with the potential for many of them to be acting in unison to determine the diversity of a particular ecosystem. It is also likely that the relative importance of different hypotheses varies both across different ecosystems and through time.

Because the latitudinal diversity gradient appears to be a universal pattern across a wide range of different kinds of organisms, it is likely that there must be some common explanation. The challenge for the future is to identify which factors, particularly which combinations of factors, are the primary causes of this pattern.

Appendix

List of Courses

1. MSc Conservation Biology, Foundations of Natural Science for Conservation (D1879), Durrell Institute of Conservation and Ecology (DICE), University of Kent, Canterbury CT2 7NR.
2. MSc / Postgraduate Diploma in the Biodiversity and Taxonomy of Plants, Royal Botanic Garden Edinburgh, Scotland
3. MSc Advanced Methods in Taxonomy and Biodiversity, Imperial College London, UK.
4. ENTO4004, Insect Taxonomy and Systematics, University of Sydney, Australia.
5. Systematics and Biotic Diversity, Cornell University, Ithaca, New York State, USA.

See also: Ecosystem Function Measurement, Terrestrial Communities. Endangered Terrestrial Invertebrates. Invertebrates, Terrestrial, Overview. Marine Ecosystems

References

Cox CB and Moore PD (1985) *Biogeography. An Ecological and Evolutionary Approach*. Oxford: Blackwell.

Gentry AH (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann. Missouri Bot. Gard.* 75: 1–34.

Holdridge LR (1967) *Life Zone Ecology*. San Jose, CA: Tropical Science Centre.

Leemans R (1990) *Global Data Sets Collected and Compiled by the Biosphere Project, Working Paper*. Austria: IIASA-Laxenburg.

MacArthur RH and Wilson EO (1967) *The Theory of Island Biogeography*. Princeton, NJ: Princeton Univ. Press.

Olson JS, Watts JA, and Allison LJ (1983) Carbon in Live Vegetation of Major World Ecosystems, Report No. ORNL-5862. Oak Ridge, TN: Oak Ridge National Laboratory.

Pianka E (1974) *Evolutionary Ecology*. New York: Harper & Row.

Rosenzweig ML (1995) *Species Diversity in Space and Time*. Cambridge, UK: Cambridge Univ. Press.

Sutton SL and Collins NM (1991) Insects and tropical forest conservation. In: Collins NM and Thomas JA (eds.) *The Conservation of Insects and Their Habitats*, pp. 405–424. London: Academic Press.

Williams PH, Gaston KJ, and Humphries CJ (1997) Mapping biodiversity value worldwide: Combining higher-taxon richness from different groups. *Proc. R. Soc. London B* 264: 141–148.

World Conservation Monitoring Centre (1992) *Global Biodiversity: Status of the Earth's Living Resources*. London: Chapman & Hall.

The Global Environment Facility: Financing the Stewardship of Global Biodiversity

Mark Zimsky, Gustavo Fonseca, Jaime Cavelier, Dirk Gaul, Jean-Marc Sinnassamy, Yoko Watanabe, and Ming Yang, Global Environment Facility, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity (BD) BD is defined by the Convention on Biological Diversity (CBD) as "the variability among living organisms from all sources, including, *inter alia*, terrestrial, marine, and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species, and of ecosystems."

Biosafety Biosafety is a term used to describe efforts to reduce and eliminate the potential risks resulting from biotechnology and its products. For the purposes of the Biosafety Protocol of the CBD, this is based on the precautionary approach, whereby the lack of full scientific certainty should not be used as an excuse to postpone action when there is a threat of serious or irreversible damage.

Mainstreaming The persistence of BD, including threatened species that are not solely dependent on site-based action, require the sustainable management of landscape and seascape mosaics that include protected areas and a variety of other land and resource uses outside of these protected areas. Measures that help reduce the negative impacts that productive sectors exert on BD, particularly outside of protected areas and highlight the contribution of all components of BD to ecosystem functioning, economic development, and human well-being is often referred to as "BD mainstreaming."

Payment for ecosystem services (PESs) The definition of "PESs" varies widely from narrow market-based definitions with emphasis on direct transactions between providers and beneficiaries (sometimes with specifications on the nature and conditions of the transaction between buyers and sellers) to broader schemes, where those who benefit from ecosystem services pay those who provide such services. Overall, PES for GEF has involved arrangements between buyers and sellers of environmental goods and services, where those who pay are fully aware of what they are paying for, and those who sell proactively and deliberately engage in resource use practices aimed at securing the provision of the services.

Protected area system "sustainability" The GEF defines a sustainable protected area system as one that: (1) has sufficient and predictable financial resources available, including external funding, to support protected area management costs; (2) effectively protects ecologically viable representative samples of the country's ecosystems and provides adequate coverage of threatened species at a sufficient scale to ensure their long-term persistence; and (3) retains adequate individual and institutional capacity to manage protected areas such that they achieve their conservation objectives.

The Global Environment Facility (GEF) The GEF unites 182 member governments – in partnership with international institutions, nongovernmental organizations (NGOs), and the private sector – to address global environmental issues. As an independent financial organization, the GEF provides grants to developing countries and countries with economies in transition for projects related to BD, climate change, international waters, land degradation, the ozone layer, and persistent organic pollutants (POPs). These projects benefit the global environment linking local, national, and global environmental challenges and promoting sustainable livelihoods. The GEF partnership includes 10 agencies: the United Nations (UN) Development Program; the UN Environment Program; the World Bank; the UN Food and Agriculture Organization; the UN Industrial Development Organization; the African Development Bank; the Asian Development Bank; the European Bank for Reconstruction and Development; the Inter-American Development Bank; and the International Fund for Agricultural Development. Recently, the GEF Council has approved a pilot program to accredit national agencies for direct access to GEF resources, becoming a pioneer in this modality among multilateral funding mechanisms. The Scientific and Technical Advisory Panel provides technical and scientific advice on the GEF's policies and projects. The GEF serves as financial mechanism for the following conventions: CBD, UN Framework Convention on Climate Change, Stockholm Convention on POPs, and UN Convention to Combat Desertification. The GEF, although not linked formally to the Montreal Protocol on Substances That Deplete the Ozone Layer, supports implementation of the Protocol in countries with economies in transition.

Introduction to the GEF and GEF Support to Biodiversity Conservation: Background and Context

What is the Global Environment Facility?

The Global Environment Facility (GEF) was established in October 1991 as a $1 billion pilot program of the World Bank to assist in the protection of the global environment and to promote environmentally sustainable development. The GEF would provide new and additional grants and concessional funding to cover the "incremental" or additional costs associated with transforming a project with national benefits into one with global environmental benefits. The United Nations

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00212-4

(UN) Development Program, the UN Environment Program, and the World Bank were the three initial partners implementing GEF projects.

In 1992, at the Rio Earth Summit, the GEF was restructured and moved out of the World Bank system to become a permanent, separate institution. The decision to make the GEF an independent organization enhanced the involvement of developing countries in the decision-making process and in implementations of the projects. However, since 1992, the World Bank has served as the Trustee of the GEF Trust Fund and provided administrative services.

As part of the restructuring, the GEF was entrusted the financial mechanism for both the UN Convention on Biological Diversity (CBD) and the UN Framework Convention on Climate Change (UNFCCC). In partnership with the Montreal Protocol (MP) of the Vienna Convention on Ozone Layer Depleting Substances, the GEF started funding projects that enable the Russian Federation and nations in Eastern Europe and Central Asia to phase out their use of ozone-destroying chemicals.

The GEF subsequently was also selected to serve as a financial mechanism for two more international conventions: The Stockholm Convention on Persistent Organic Pollutants (POPs, 2001) and the UN Convention to Combat Desertification (UNCCD, 2003).

The GEF unites 182 member governments in partnership with international institutions, civil society, and the private sector. As the world's largest public financial organization dedicated to global environmental issues, the GEF provides grants to developing countries and countries with economies in transition for projects related to biodiversity (BD), climate change (CC), international waters, land degradation (LD), the ozone layer, and POPs. These projects link local, national, and global environmental challenges while promoting sustainable livelihoods.

Over its 20 year history, the GEF has directly invested $9.5 billion, leveraging another $42 billion toward more than 2700 projects in 165 countries. Through its Small Grants Program (SGP), the GEF has also made more than 12,000 small grants directly to NGOs and community organizations.

The GEF partnership includes 10 agencies: the UN Development Program, the UN Environment Program, the World Bank, the UN Food and Agriculture Organization, the UN Industrial Development Organization, the African Development Bank, the Asian Development Bank, the European Bank for Reconstruction and Development, the Inter-American Development Bank, and the International Fund for Agricultural Development. Recently, the GEF Council has approved a pilot program to accredit national agencies for direct access to GEF resources, becoming a pioneer in this modality among multilateral funding mechanisms. The Scientific and Technical Advisory Panel provides technical and scientific advice on the GEF's policies and projects.

The GEF also serves as financial mechanism for the following conventions:

- CBD
- UNFCCC
- Stockholm Convention on POPs
- UNCCD
- The GEF, although not linked formally to the MP on Substances That Deplete the Ozone Layer, supports implementation of the protocol in countries with economies in transition.

The GEF and the CBD

The CBD provides the global policy framework to address BD issues. The CBD also provides the guidance under which the GEF operates to assist countries in meeting their obligations under the Convention. In other words, the GEF is the financial instrument of the Convention and is the only binding multilateral agreement in this area, with 190 parties. The objectives of the CBD are defined in Article 1 as "…the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding."

The relationship between the Conference of the Parties (COPs) of the CBD and the GEF is ruled by a memorandum of understanding (Convention on Biological Diversity, 1992). In accordance with Article 21 of the Convention, the COP determines policy, strategy, program priorities, and eligibility criteria for access to and utilization of financial resources available through the financial mechanism, including monitoring and evaluation. In translating the COP guidance to operational policy for implementation, the Secretariat, in consultation with the GEF Agencies, assesses how the guidance can best be implemented. The GEF defines new or strengthened strategic objectives and approaches, modalities, operational criteria, procedures, and any other process needed and presents this for Council approval. In applying COP guidance in project operations, the GEF and its implementing agencies support country-driven, national priority projects and programs endorsed by relevant GEF focal points (i.e., full- and medium-sized projects, SGP, and enabling activities).

GEF Funding for BD

The level of investment in national BD projects that would be warranted by the global benefits of those projects is frequently greater than the investment warranted by local development or environmental priorities. The GEF provides grants in support of national projects to cover the difference – the incremental cost of achieving the desired level of global benefits generated by BD conservation and sustainable use.

The GEF's BD portfolio has been the largest focal area portfolio in terms of grant money awarded since the facility began operations and has accounted for approximately 31% of the total GEF grants to developing countries and those with economies in transition. Since 1991, the GEF has provided about $3.1 billion in grants and leveraged about $8.3 billion in cofinancing to support approximately 990 projects that addressed the loss of globally significant BD in more than 155 countries (**Figure 1**).

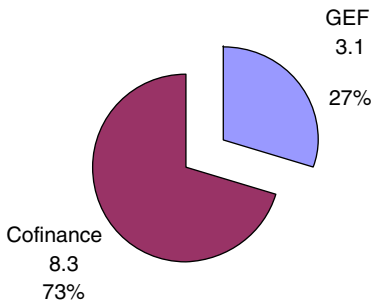


Figure 1 GEF Biodiversity Investments FY 1991–2010 (US$ billion). The GEF's biodiversity portfolio has been the largest focal area portfolio in terms of grant money awarded since the Facility began operations.

System for Transparent Allocation of Resources

The GEF allocates resources in an indicative way to GEF-eligible countries during a replenishment period through the System for Transparent Allocation of Resources (STAR). It was developed during 2009–2010 to upgrade the Resource Allocation Framework (RAF), which was the former GEF resource allocation system used in the fourth replenishment period of the GEF (GEF-4).

In the fifth replenishment period of the GEF (GEF-5), the STAR covers three focal areas: BD, CC, and LD. Other focal areas and programs may be covered by the STAR in the future GEF replenishment periods.

The main benefits of the STAR for eligible countries are predictability of funding and flexibility in programming. This will enhance planning and contribute to country ownership of GEF projects and programs.

The funding envelope that a given country can access during a replenishment period is called "indicative allocation." Indicative allocations are calculated based on a combination of the GEF benefits index (GBI), the GEF performance index (GPI), and a social and economic index based on the Gross Domestic Product (GDPI).

The GBI represents the global environmental benefits that can be generated for each focal area in a specific country. The GPI is specific to a country and is the same across all focal areas. The GDPI is a new design feature of the STAR. The rationale to have the new index in the STAR is that poorer countries need additional funding to build the capacity that is required to ensure the success of GEF project development and implementation. The GDPI is the same for all focal areas in a country.

A GBI in the STAR is an index to measure GEF's investment benefits in a specific focal area in a country. In the STAR, there are three GBIs: GBIBD, GBICC, and GBILD. They are directly related to environmental indicators for three focal areas: BD, CC, and LD. The GBIs in a focal area represent a country's relative share of GEF potential benefits that can be generated by a fixed amount of resource input in the focal area. The higher the GBI in a focal area, the higher are the investment benefits to be generated.

The GPI in the STAR is an index to provide a relative ranking of eligible countries' performance and capacity to deliver potential global environmental benefits. The GPI is the same across all focal areas in a country. The calculation of the GPI is based on a country's past and current performance in project development and implementation, and on the quality of each country's policies and institutions. The GPI therefore provides a proxy measure of each country's capacity to successfully implement GEF programs and projects to deliver the potential global environmental benefits identified by the GBI.

The GDPI is designed so as to increase the allocation of countries with low GDP per capita, decrease the allocation of countries with higher GDP per capita, and keep the allocations with middle GDP per capita unchanged. The higher the GDP per capita, the larger impact the GDPI will have.

The calculation of indicative country allocation scores for BD under the STAR is based on three indices: (1) the GBIBD, (2) the GPI, and (3) the GDPI.

The same GBIBD data were used in the RAF and in the STAR. In the RAF, the GBIBD score consisted of two components: (1) terrestrial score (80%) and (2) marine score (20%). In the STAR, the weights were slightly modified – terrestrial (75%) and marine score (25%).

For BD, the "benefits index" is built utilizing the updated GEF-4 benefits index, but with an increased weight for the marine component from 20% to 25%. The GBI makes maximum use of the available, scientifically reliable information for a cross-country assessment of terrestrial and marine BD. The GBI reflects the complex, highly uneven distribution of species and threats to them across the ecosystems of the world, both within and across countries. In terrestrial BD, the GBI accounts for ecoregion representation, species representation, threatened species, and threatened ecoregions. In marine BD, it accounts for species representation.

Evolution of the GEF BD Strategy: From GEF Inception to the Present

GEF Operational Strategy and GEF Operational Programs

Until the formulation of GEF's first targeted BD strategy for implementation during GEF-3 (2003–2006), the GEF BD portfolio was built on the GEF Operational Strategy and Operational Programs, as well as the guidance provided to GEF from the COP. The GEF Operational Strategy defines the 10 operational principles for development and implementation of the GEF's work program. The GEF Operational Programs in BD were built on the general Operational Strategy and defined, by ecosystem type, specific criteria by which the GEF projects were further characterized and evaluated. Earlier implementation of the GEF BD program emphasized eligibility based on fit with one or more of the five BD operational programs: arid and semiarid ecosystems, coastal–marine and freshwater ecosystems, forest ecosystems, mountain ecosystems, and agro-BD.

GEF's BD Strategy

The Formulation of the First BD Strategy for GEF-3

In response to the two external evaluations undertaken by the BD program, the Second Program Study of the GEF

Biodiversity Program and the Second Overall Performance Study, the GEF developed a strategy for GEF-3 to focus the GEF's investment strategy on four priorities:

1. catalyzing the sustainability of protected areas,
2. mainstreaming BD in the production of landscapes/seascapes and sectors,
3. capacity building for the implementation of the UN CBD Cartagena Protocol on Biosafety (CPB), and
4. generation and dissemination of best practices for addressing current and emerging BD issues.

The main purpose for sharpening the investment focus of GEF resources was to apply scarce GEF resources in a manner that most effectively catalyzes actions to maximize global environmental benefits. The priorities for GEF-3 internalized the most pertinent recommendations that had emerged from the evaluation exercises and provided a framework for the entire portfolio. The GEF developed its first targeted BD strategy in GEF-3 to complement and further focus its operational programs and to respond to evaluation findings (Biodiversity Program Study, 2004).

The GEF-3 strategy incorporated principles to achieve lasting BD conservation and sustainable use and thereby: (1) placed greater emphasis on sustainability of results and the potential for replication; (2) moved beyond a project-based emphasis to strategic approaches that strengthened country enabling environments (policy and regulatory frameworks, institutional capacity building, science and information, and awareness); (3) mainstreamed BD conservation and sustainable use in the wider economic development context; and (4) increased support for sustainable use and benefit sharing. The changes implemented in the GEF-3 strategy formed the foundation on which subsequent GEF strategies have been built. The strategy for each new phase has maintained continuity with these basic tenets of sustainability while incorporating new findings on good practice in BD conservation and sustainable use.

BD Strategy for GEF-4 and GEF-5

The GEF revised its strategy for GEF-4 (2007–2010) (GEF BD Strategy, GEF-4, and GEF-5) based on the implementation experience gained during GEF-3 and in response to evolving thinking in the conservation community about the drivers of BD loss. The GEF-funded Millennium Ecosystem Assessment identified the most important direct drivers of BD loss and degradation of ecosystem goods and services as being habitat change, CC, invasive alien species, overexploitation, and pollution (Millennium Ecosystem Assessment (2005)). These drivers are influenced by a series of indirect drivers of change, including demographics, global economic trends, governance, institutions and legal frameworks, science and technology, and cultural and religious values. The BD strategy in GEF-4 addressed a subset of the direct and indirect drivers of BD loss and focuses on the highest leverage opportunities for the GEF to contribute to sustainable BD conservation.

To achieve the goals of the GEF-4 BD strategy, four complementary and mutually reinforcing objectives were implemented: (1) improving the sustainability of protected area systems, the most predominant and dedicated land use globally for BD conservation; (2) mainstreaming BD conservation and sustainable use into production sectors that impact BD; (3) safeguarding BD by building country capacity to implement the CPB and prevention, control, and management of invasive alien species; and (4) capacity building to support the implementation of the Bonn Guidelines on access to genetic resources and benefit sharing. Underpinning these responses, GEF supported institutional capacity building and the development of the appropriate policy frameworks to ensure sustainable BD conservation.

The GEF-4 strategy formally replaced the previous structure of operational programs and balanced the need for continuity in the investment strategy, while focusing more explicitly on specific interventions for sustaining conservation over the long term.

The GEF BD strategy is consistent with the integrated approaches to BD conservation and sustainable use promoted by the ecosystem approach, the primary framework for action under the CBD. Together, these strategic objectives will make a substantial contribution to implementing most of the Millennium Development Goals, particularly environmental sustainability and poverty reduction, while meeting the priorities identified by the COP of the CBD.

The GEF-5 (2010–2014) strategy maintains coherence with the GEF-4 strategy, while proposing refinements to the strategy's objectives based on COP-9 guidance, advances in conservation practice, and advice from the GEF's Scientific and Technical Advisory Panel. The ninth meeting of the COPs of the CBD acknowledged that the GEF-4 strategy served as a useful starting point for the GEF-5 strategy and requested GEF to build on it for the fifth replenishment based on the 4-year framework of program priorities developed by COP-9.

The goal of the BD focal area is the conservation and sustainable use of BD and the maintenance of ecosystem goods and services. To achieve this goal, the strategy encompasses five objectives:

1. improve the sustainability of protected area systems,
2. mainstream BD conservation and sustainable use into production landscapes/seascapes and sectors,
3. build capacity to implement the CPB,
4. build capacity on access to genetic resources and benefit sharing, and
5. integrate CBD obligations into national planning processes through enabling activities.

GEF Financing of Sustainable Forest Management and its Contribution to BD Conservation

Forests cover almost 4 billion ha, or 30%, of the world's total land area. They are found virtually everywhere, from the boreal forest to the humid tropics. The broad variety of forest types – in tropical, subtropical, Mediterranean, temperate, and boreal regions – together account for two-thirds of all terrestrial ecoregions. They provide a wide range of environmental services, including BD conservation, water supply, carbon sequestration, flood control, and protection against soil erosion and desertification.

Equally important, forest ecosystems contain at least an estimated 80% of the Earth's BD. Tropical forests are

particularly rich in species. Although covering only approximately 10% of the total terrestrial surface, they correspond to approximately 50% of the area covered by world's forests and are home to considerably more than 60% of all terrestrial and freshwater BD.

The UNFCCC, the CBD, and the UNCCD all emphasize the importance of the conservation, sustainable use, and management of forests in achieving their respective Convention objectives. As operating entity of the financial mechanisms for all three Conventions, the GEF has since 1991 supported more than 300 projects and programs that focus on various actions dealing with forest conservation and management in developing countries. During the same period, the GEF has allocated approximately US$1.6 billion to forest initiatives, supplemented by more than US$5.0 billion in cofinancing. GEF has continuously increased its financial flows for forest-related activities throughout its successive replenishment periods.

Drawing on guidance from the related conventions, from its eligibility criteria, and from associated strategic programs, the GEF has funded projects that can be broadly classified into three categories:

1. forest conservation (primarily protected areas and buffer zones),
2. sustainable use of forests (forest production landscapes), and
3. sustainable forest management (SFM) (addressing forests and trees in the wider landscape).

Historically, most of the GEF's investments were dedicated to forest conservation. However, investments in the latter two categories have increased considerably with the introduction of a dedicated program in SFM.

Until 2006, GEF supported forest projects drawing from the focal areas of BD and LD. In response to the growing international recognition of the role of forests in delivering global benefits across a variety of themes, in 2007, the GEF Council approved the broadening of the GEF's efforts in the field of SFM. Adopting the framework of the UN Forum on Forests, the GEF considers SFM "a dynamic and evolving concept that aims to maintain and enhance the economic, social, and environmental value of all types of forests, for the benefit of present and future generations." The inclusion of Land Use, Land-Use Change and Forestry (LULUCF) into the climate change strategy enabled support to projects and programs that deliver global benefits across all three GEF focal areas that are directly relevant to forests: BD, CC, and LD.

Through its pilot SFM Program in GEF-4 (2007–2010), the GEF took early action in the Reducing Emissions from Deforestation and Forest Degradation (REDD+) and LULUCF arenas by providing resources for pilot projects focusing on cross-sectoral cooperation. In the current funding cycle GEF-5 (2010–2014) and in line with the Copenhagen Accord that calls for "…substantial finance to reduce emissions from deforestation and forest degradation…," the GEF is expanding its support to actions reducing deforestation and will provide up to $1 billion for the implementation of a dedicated SFM/REDD+ Program throughout the period 2010–2014, during which the GEF will manage a separate $250 million funding envelope for SFM/REDD+. This envelope operates as an incentive mechanism for developing countries to invest up to $750 million of their allocations from BD, CC, and LD for more comprehensive SFM/REDD+ projects and programs. Altogether, the GEF will thus make up to $1 billion for SFM/REDD+ funding available throughout the course of GEF-5. This investment is expected to leverage substantial additional funding from external sources.

Monitoring the GEF BD Portfolio

Monitoring Results at the Portfolio Level

The independent GEF Evaluation Office develops the policy, related guidelines, and administrative procedures for monitoring and evaluation in the GEF. The policy and guidelines help project managers and Agency and GEF Secretariat staff plan and conduct monitoring and evaluation. The GEF Monitoring and Evaluation Policy provides norms and standards for the GEF Secretariat and the GEF Evaluation Office. The Policy explains the concept, role, and use of monitoring and evaluation within the GEF; establishes minimum requirements for how projects should be monitored and evaluated in line with international standards; and assigns roles and responsibilities for these tasks. The GEF Agencies plan and implement their project monitoring and evaluation, in line with their own systems and procedures and based on these minimum requirements.

The GEF introduced BD "tracking tools" in GEF-3 to measure progress in achieving the outputs and outcomes established at the portfolio level for GEF-3 in the BD focal area. (GEF website, http://www.thegef.org). These tools have evolved with each successive phase of the GEF based on experience applying the tools and changes in GEF's portfolio results framework.

The tracking tools are applied three times during the life of the project: at project initiation to establish a baseline, then at project midterm, and at project completion. Project outcomes from each GEF phase cohort are aggregated for analysis of directional trends and patterns at a portfolio-wide level to inform the development of future GEF strategies and to report to the GEF Council on portfolio-level performance in the BD focal area as the projects are completed and evaluations conducted.

Learning at the Portfolio Level

The GEF Results-Based Management (RBM) framework and approach includes an emphasis on portfolio monitoring and learning by placing particular attention on using monitoring information for accountability, internal management, learning, and knowledge management.

In support of the GEF RBM and based on a review of evaluations of the BD focal area, a select number of learning objectives were identified for the BD focal area and these were included in the GEF-5 BD strategy to be implemented and lead by the GEF Secretariat in collaboration with the GEF Agencies.

The three learning objectives share in common a dual-fold purpose, in that the results will contribute to strengthening

GEF's capacity to deliver on its own mandate and the broader global public good of enhanced knowledge to catalyze change in BD conservation practice. Three learning objectives are proposed for implementation in GEF-5 and include:

1. Learning Objective One: "Enhancing Impacts and Outcomes through Improved Understanding of Protected Area Management Effectiveness." The learning objective is to better correlate protected area management effectiveness as recorded by a scorecard (Management Effectiveness Tracking Tool (METT)) to the successful conservation and sustainable use of BD within a protected area. This learning objective will be accomplished through a series of country case studies and field visits to select countries that have been applying the METT over an extended period of time in their protected area system and that are also collecting quantitative data on the status of BD and protection within the system.
2. Learning Objective Two: "Enhancing Social Impacts through Improved Understanding of the Causal Relationships between Protected Area Management and Local Community Welfare." This learning objective seeks to answer the following question, "What has been the impact of protected areas on human welfare?" This learning objective will be accomplished through a systematic review of the literature as well as complementary case studies when these are designed to focus on elucidating potential causal relationships.
3. Learning Objective Three: "Enhancing Impacts through Improved Understanding of the Causal Relationships between Popular Mainstreaming Approaches and Conservation Outcomes." As a leader in supporting innovative incentive-based and information-based mainstreaming approaches, the GEF has observed an increase in the number of funded projects using certification, Payment for ecosystem services (PESs), and ecosystem service valuation. Thus, the GEF has an opportunity to contribute the evidence base of these approaches by supporting work to answer the following question, "How do certification, PES, and transfers of information about the distribution and values of ecosystem services affect conservation and sustainable use outcomes, and in what circumstances are they likely to be most effective?" This learning objective will be accomplished primarily through the support of prospective experimental and *quasi*-experimental project designs. When feasible, quantitative retrospective studies in programs that have received GEF funding will also be supported.

The GEF network of agencies, partner government, and nongovernment executing agencies and country-based staff will be the main users of the findings derived from the portfolio monitoring and learning review process.

Given the extensive investment that the GEF has made in protected areas over the course of its existence ($1.89 billion of GEF resources, which supported 2302 protected areas spanning 634 million ha and 700 globally threatened species), priority has been placed on first implementing learning objective one, "Enhancing Impacts and Outcomes through Improved Understanding of Protected Area Management Effectiveness," through five country case studies as a priority for the first 2 years of GEF-5. The first learning mission was undertaken in Zambia in 2010 and the full report can be found at: http://www.thegef.org. The report represents a new era of portfolio monitoring and learning at the GEF, where the portfolio is used as a source for developing the evidence base for effective conservation investments.

Looking Ahead: What is Next for GEF

A New GEF: Seeking BD Impacts Commensurate with the Scale of the Threats

To continue to fulfill its function as the financial mechanism of the CBD, the GEF must evolve and be strengthened. The emphasis on BD as an individual focal area will remain at GEF, not only to highlight the specificities of dealing with an irreplaceable global good whose value to society remains to be fully assessed but also to address the ways to deal with existing and emerging threats, as well as identify rapidly expanding opportunities to act strategically. GEF also needs to support larger programs composed of many complementary projects, which themselves could include resources from different GEF focal areas, adding to the BD-specific investments.

The reformed GEF created a Natural Resources Management team, breaking the silos that existed in the focal areas of BD, LD, and international waters. This structure allowed for strategic cross-fertilization to occur and enabled new multifocal area programs to emerge, such as the SFM/REDD + program. The other pioneering programs introduced during GEF-4 (e.g., Pacific Alliance for Sustainability and the China Biodiversity Partnership and Framework for Action, among others) have opened the opportunity to pool resources from BD, CC mitigation and CC adaptation, LD, and international waters into more integrated programmatic action plans tackling multiple issues and fostering synergy among a diverse set of interventions. Such a scale of action also promotes a higher degree of efficiency in the overall conservation investment.

Nothing can better illustrate this change in the scope of planned interventions, and in the ambition of delivering outcomes that are fundamental to BD but also result in multiple benefits, than GEF's SFM Strategy. Tropical forests, in addition to conserving global BD, sustaining rural livelihoods, and providing spiritual and cultural havens for local and remote populations, are also among the largest and most important providers of ecosystem services on Earth, fundamental to maintaining our planet's long-term health and stability. Tropical forests have now been identified as a significant part of the overall CO_2 emission reduction scheme, given that they are responsible for 20% of the current greenhouse gas emissions problem.

As the GEF gained experience through the implementation of its SFM program, it built the foundation for a more ambitious global forest initiative in GEF-5, incorporating more explicit CC mitigation objectives. Threats to forests and opportunities for their conservation and sustainable management arise from a variety of sectors. These include agriculture expansion, shifts in global commodity markets, infrastructure development, and energy. But, more importantly, the role of forests in the global carbon equation is solidifying in policy

Box 1

Key accomplishments of the GEF in biodiversity

- The GEF biodiversity focal area program has provided approximately $2.9 billion in grants and leveraged an additional $8.2 billion in cofinancing to support 990 projects in more than 155 countries.
- The GEF is the largest funding mechanism for protected areas worldwide. GEF has invested in more than 2302 protected areas, covering more than 634 million ha, an area equivalent to Greenland and Mongolia put together. The GEF has provided more than $2.9 billion to fund protected areas, leveraging an additional $8.2 billion in cofinancing from project partners.
- The GEF leads the world in establishing financing mechanisms to sustainably finance and operate national protected areas systems in developing countries. It has supported more than 90 projects that involve conservation trust funds, PESs schemes, revolving funds, private sector and village funds, and other innovative financial mechanisms to provide steady, reliable funding for protected area management, and biodiversity conservation in developing countries.
- The GEF is recognized as a pioneer in supporting more than 40 conservation trust funds worldwide, investing more than $300 million in total.
- GEF's PESs portfolio includes more than 42 projects with explicit PES components. Investments have been made in the development of national systems of PES, regional, or local schemes with investments from the private sector, and private–public partnerships.
- The GEF is recognized as the first provider of capacity building in the area of biosafety, where the GEF has invested more than $115 million and leveraged more than $113 million.
- The GEF has supported the development of National Biosafety Frameworks in 123 countries, contributing to a rapid ratification by countries of the CPB and has built the capacity for the countries' effective participation in the Biosafety Clearing House mechanism.
- The GEF has supported 71 countries to effectively implement their National Biosafety Frameworks and the CPB.
- The GEF has developed strong partnerships with civil society organizations, including NGOs and indigenous and local communities, through its biodiversity program. The GEF SGP has provided small grants to more than 6945 biodiversity projects proposed by NGOs and community-based organizations in 121 countries, with a total GEF funding of $51 million. The Critical Ecosystem Partnership Fund is another GEF partnership mechanism, with a program budget of more than $125 million, which has reached out to more than 1000 civil society organizations in 33 countries to help conserve the world's most important biodiversity hotspots. More recently, the GEF has partnered with the World Bank and the International Union for Conservation of Nature, together with the French Fonds Français pour l'Environnement Mondial and private sector partners, in the Save our Species (SOS) Initiative. SOS aims to become a significant provider of funding in support of the conservation of globally threatened species, as well as raising global awareness to the plight of biodiversity worldwide.

circles, and the GEF is now programmatically prepared to act swiftly in this arena. The effective implementation of the SFM strategy will thus require a more holistic, wide-reaching approach.

The imperative of dealing with CC has introduced a new reality to GEF's programs spanning BD outcomes. Although it is necessary to act in the short term, before important windows of opportunities to safeguard global BD close, it is no longer sufficient to solely address the most immediate BD conservation challenges. For example, pressing action must start to be complemented by long-term sustainability interventions to cope with the anticipated effects of CC. In some regions, entire protected area systems will need to be the focus of adaptation strategies, as species and communities adjust to a changing environment.

At the same time, the opportunities to combine multiple goals in natural resources management are growing. We now know how interdependent the objectives of combating desertification, introducing sustainable land management, and improving productive systems all are for successful BD conservation strategies to emerge.

The GEF must also continue to expand its efforts in building the capacity of institutions to act as effective stewards of their environmental goods. Capacity building is likely the most enduring of investments, but one whose results are felt mostly in the medium and long term.

In the long term, BD conservation must be mainstreamed into all development sectors. There is an emerging consensus among environmental economists that changes in BD affect the provision of ecosystem goods and services, particularly those on which rural communities depend most. Although the magnitude of this dependence is yet to be determined, most studies agree that BD conservation is a very worthwhile investment, in particular for BD-rich countries. The GEF is open to embrace these new challenges and opportunities, experimenting with large-scale projects, and programs on mainstreaming BD into productive landscapes **Box 1**.

Acknowledgment

This article draws heavily on the GEF report 'Financing the Stewardship of Global Biodiversity' (GEF, 2008) of which Mark Zimsky was the primary author. The author acknowledges the permission of GEF to reproduce sections of that report in this article.

Appendix

List of Courses

1. Conservation Biology
2. International Development
3. Rural Development
4. Biodiversity Conservation
5. Environmental Economics.

References

GEF Biodiversity Strategy for GEF-4 and GEF-5 and GEF Biodiversity Tracking Tools can be found at http://www.thegef.org

Global Environment Facility, Office of Monitoring and Evaluation, Biodiversity Program Study, 2004. http://www.thegef.org

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-being: Synthesis*. Washington, DC: Island Press.

Secretariat of the Convention on Biological Diversity (2005) *Handbook of the CBD, including its Cartagena Protocol on Biosafety*, 3rd edn. Montreal, Canada.

Zimsky M, Fonseca G, Cavelier J, *et al*. (2008) *Financing the Stewardship of Global Biodiversity*. Washington DC: Global Environment Facility.

The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity

Jeffrey J Opperman, The Nature Conservancy, Chagrin Falls, OH, USA
Gerald E Galloway, University of Maryland, College Park, MD, USA
Stephanie Duvail, Institute of Research for Development, Nairobi, Kenya

© 2013 Elsevier Inc. All rights reserved.

Glossary

Conveyance The movement of water through an area. For example, by increasing the cross-sectional area available to convey floodwaters, floodplains can reduce flood stage.
Ecosystem-based adaptation Conservation and management of natural ecosystems such that they can provide services that help society adapt to perturbations from climate change.
Ecosystem service Products and processes generated by functioning ecosystems that economically benefit society.
Floodplain Low-lying land adjacent to a river and subject to periodic inundation from overflow from the river and other sources.
Flood-recession agriculture A pattern of agriculture in which crops are planted on floodplains as floodwaters recede. The growing plants benefit from several factors influenced by the flood, including moist soil, elevated water table, and freshly deposited soil and nutrients.
Hydrologic connectivity The ability for water to move between portions of a river system, such as between a river channel and floodplain surfaces.
Levee A structure built along rivers intended to contain floodwaters and prevent flooding of land behind the structure. Also referred to as "dikes." Note that natural levees can form along rivers through geomorphic processes.
Nonstructural flood management Flood-risk management approaches that do not depend on controlling water with engineered structures. Examples include land-use planning, building codes, insurance, and use of the natural and beneficial functions of floodplains, including storage and conveyance of floodwater.
Residence time Hydrological residence time is the length of time that a given unit of water remains in a given place and thus reflects the exchange rate of water at that place. Residence time can be calculated in many ways. One simple method is by dividing the volume of the area of interest (e.g., floodplain site) by the flow rate.
Riparian forest Forests that grow adjacent to rivers and within floodplains, generally composed of trees tolerant of periodic flooding and riparian trees often require flooding to promote their regeneration.

Introduction

Floodplains are dynamic and productive ecosystems whose immense values are often overshadowed by losses of life and property during large flood events. Despite the attention generated by disasters, floodplains support rich ecosystems and provide critically important benefits to people in diverse ways. In some regions, periodic inundation of the floodplain is essential to the provision of benefits, such as those from fisheries or flood-recession agriculture, so named because planting occurs as the floodwaters recede. In other settings, floodplain benefits depend on disconnecting the river from the floodplain, such as through levees that protect flood-intolerant crops growing on rich alluvial soil. In this article, the authors review how maintaining or restoring floodplain connectivity can promote the full spectrum of floodplain benefits. While the benefits that require flooding are clearly enhanced by river–floodplain connectivity, the benefits that require the absence of flooding can also be enhanced through strategically located connections between rivers and floodplains as part of comprehensive flood-risk management. In other words, some designated locations are allowed to flood in order to reduce flood risk for other areas. In this way, floodplains act as "green infrastructure" complementing the services provided by engineered infrastructure, such as dams and levees, for managing flood risks.

In this article, the authors review floodplain processes and how connectivity between the river and floodplain drives productivity, diversity and a range of outputs and services that directly benefit people and nature. The conversion of floodplains from largely natural habitats to a variety of other land uses and the management challenges associated with human use and occupation of floodplains is then described. Changes in land use, demographic patterns, and climate are all contributing to an increase in flood risk. The article summarizes the benefits of managing floodplains for connectivity and then concludes with several case studies that illustrate the multiple benefits for people and nature of restoring or maintaining this connectivity.

Key Floodplain Processes and Benefits

Floodplains are complex systems with physical, biological, economic, and social components (Naiman, Decamps *et al.*, 2005). They are created by flows of water and sediment, which influence, and are influenced by, biological processes. Floodplains are often converted to agricultural and urban development and these land uses are often protected by infrastructure

such as dams and levees. Owing to these complex and interacting components, professionals of numerous disciplines contribute to floodplain science, policy, and management: hydrologists, geomorphologists, ecologists, engineers, planners, sociologists, and policy makers. As a result, there are numerous definitions of floodplains (Nanson and Croke, 1992; Alexander and Marriot, 1999).

From a geomorphic perspective, a floodplain is a depositional feature composed of sediment derived from overbank and in-channel processes of an adjacent water body (Nanson and Croke, 1992; Knighton, 1998). A hydraulic floodplain is a surface that is inundated by a flood of a defined recurrence interval (e.g., the 100-year floodplain that has a 1% probability of being inundated in any given year) whether or not that surface is alluvial (Alexander and Marriot, 1999). This definition is often used by hydrologists and engineers (Nanson and Croke, 1992). For example, the National Flood Insurance Program, administered by the US Federal Emergency Management Agency, requires that communities participating in the Program regulate development in the 1% annual probability hydraulic floodplain (Mount, 1995). Ecologists focus on how hydrological and geomorphic processes shape ecosystem processes (Mahoney and Rood, 1998; Ahearn, Viers *et al.*, 2006) and emphasize the importance of a diversity of flow levels and other hydrologic characteristics (Poff *et al.*, 1997; Opperman *et al.*, 2010).

Hydrological, Geomorphic, and Ecological Processes

Floodplain ecosystems are shaped by the interaction of hydrologic, geomorphic, and ecological processes. Though often viewed as distinct elements – with the river as water and the floodplain as "land" – in reality, a river and its floodplain are one integrated system for conveying water and sediment (Knight and Shiono, 1996).

Tockner and Stanford (2002) characterize river hydrology as "by far the single most important driving variable in floodplains." Floodplains are highly variable and heterogeneous systems due to the natural variability of river flows (Poff *et al.*, 1997). Riverine hydrology exerts a strong influence on floodplain ecological and biogeochemical processes and, because river flows drive the processes of erosion and deposition that create floodplain topography, also shapes the physical template on which ecosystem processes occur (Poff *et al.*, 1997; Ward *et al.*, 2002). Hydrological processes influence floodplain ecosystems by controlling patterns of connectivity, residence time, and the flows that allow the exchange of organisms, carbon, and nutrients between portions of the landscape (Wiens, 2002).

These linkages between hydrological processes and floodplain geomorphic and ecological processes are driven by a variety of flow levels and other hydrologic parameters. For example, long-duration, high-frequency flood events can be particularly important for food-web productivity and fish reproduction (Junk *et al.*, 1989; Williams *et al.*, 2009), while less frequent, higher magnitude events exert stronger influences on floodplain morphology (Trush *et al.*, 2000). An extensive review of floodplain processes is beyond the scope of this article, but several comprehensive reviews are available (Welcomme, 1979; Junk *et al.*, 1989; Sparks *et al.*, 1990; Nanson and Croke, 1992; Bayley, 1995; Wohl, 2000; Ward *et al.*, 2002; Opperman *et al.*, 2010).

Multiple Benefits of River–Floodplain connectivity

Hydrologically connected floodplains support productive ecosystems and generate a broad range of benefits for people (Gren *et al.*, 1995). These benefits can be characterized as ecosystem services – products and processes generated by functioning ecosystems that economically benefit society (National Research Council (NRC), 2004; Brauman *et al.*, 2007). In their review on the value of the world's ecosystem services, Costanza *et al.* (1997) found that floodplains were the second ranked ecosystem type, behind only estuaries, in terms of their per-hectare value to society. Despite representing $<2\%$ of Earth's terrestrial land surface area, floodplains provided approximately 25% of all "terrestrial" (i.e., nonmarine) ecosystem service benefits (Costanza *et al.*, 1997). Other researchers have attempted to quantify the benefits provided by floodplains. For example, Sheffer *et al.* (2002) concluded the costs of replacing the goods and services provided by functioning floodplains and found that each hectare of floodplain has a replacement cost of approximately US $150,000.

Below, several of these benefits are briefly reviewed, with an emphasis on three: agricultural and pastoral productivity, fisheries, and flood attenuation.

Floodplains support high levels of biodiversity (Tockner and Stanford, 2002) for a variety of reasons. Because floodplains often contain a mosaic of habitats – river, wetland, grassland, and forests of varying successional stage – they frequently have high levels of beta diversity (Odum, 1978; Salo *et al.*, 1986), and this habitat mosaic is maintained by fluvial processes that require river–floodplain connectivity. Floodplains also contribute to regional species richness because they often provide unique habitat types not otherwise found in the region (Sabo *et al.*, 2005), such as when a large river flows through a very dry region, or when a region has been largely converted from natural land covers to agriculture or urban development. Within extensively converted landscapes, floodplains can support the last remains of natural habitat and so can continue to support populations of species that would otherwise be extirpated from the region (Odum, 1978). In addition to providing habitat for endangered species, floodplains can support game species (e.g., deer, waterfowl) and provide other recreational and open-space benefits. As discussed below (*see* Fisheries), floodplains connected to rivers can support productive fisheries and thus contribute to commercial and recreational fishing.

During overbank flooding, floodwaters spread out on floodplains and, due to slower water velocities on the floodplain, much of the sediment in transport is deposited there. Because nutrients such as phosphorous are largely adsorbed to sediment particles, this deposition can reduce the loads of sediment and some nutrients in rivers and thus improve water quality for downstream waterbodies, such as estuaries and near-shore marine habitats (Noe and Hupp, 2005). Biogeochemical processes within floodplain wetlands, such as denitrification, can also reduce nitrogen loads in river water (Burt and Pinay, 2005; Valett *et al.*, 2005). Floodplain

reconnection and restoration of floodplain wetlands is therefore recommended as a strategy to reduce nutrient pollution to important waterbodies such as the Chesapeake Bay (Noe and Hupp, 2005) and the Gulf of Mexico (Mitsch *et al.*, 2001). Owing to sediment deposition during recurrent overbank flooding, portions of floodplains can have deep, fertile soil, which can support productive forests (Brinson, 1990; Yarie *et al.*, 1998). Thus, in addition to nutrient sequestration, floodplains can also sequester carbon within rapidly growing trees.

During overbank flooding, a portion of floodwaters can percolate into the shallow groundwater. For example, Valett *et al.* (2005) reported that approximately half of the floodwater entering an experimental semiarid floodplain in New Mexico (USA) contributed to groundwater recharge. The Yolo Bypass, a managed floodplain in California's Sacramento Valley (*see* Floodplain Reconnection for Flood-Risk Reduction), is frequently inundated for long periods of time, contributing to recharge of the groundwater. This recharge supported a valuable "groundwater bank" during a drought (Jercich, 1997). In arid, semiarid, or Mediterranean climate regions, overbank flooding can contribute to ecosystem and agricultural productivity by recharging groundwater and thus increasing the water available during the growing season, particularly if flooding occurs during or relatively soon before a period of drought stress (Robertson *et al.*, 2001). Groundwater recharge is often one of the most important ecosystem services provided by floodplains (Schuyt, 2005), supporting much of the agriculture in parts of Africa (Acharya and Barbier, 2002; Barbier, 2003).

Floodplain Agriculture

This linkage between groundwater recharge and plant productivity is responsible for one of the primary socio-economic benefits of connected floodplains. In Africa, Latin America, and Asia, floodplains that remain hydrologically connected to rivers provide important sites for agriculture and grazing. The flood pulse promotes vital resources for populations in arid and semiarid climates, such as crop production and forage for grazing livestock, by replenishing soil nutrients through sediment deposition and/or recharging groundwater. Depending on various factors (the intensity of water control, the relative importance of rainfall and flood, the type of socio-political institution controlling the natural resources, and the various users' strategies) floodplain agriculture takes different forms along a gradient from rainfed through recession agriculture to irrigation. Below, examples are drawn from Africa to illustrate the mechanisms of floodplain cropping and pastoralism.

Flood-recession agriculture is the most ancient form of floodplain agriculture. Where floods are exogenous (i.e., the majority of the runoff is derived from upstream areas distant from the floodplain), such as in Saharan and Sahelian Africa prior to extensive river regulation, flood-recession agriculture has been the main source of food for vast numbers of people, exemplified by Nile agriculture before completion of the Aswan High Dam. Prior to flow regulation from dams, the Senegal River floodplain supported approximately 200,000 ha of sorghum during the flood-recession, which was the primary food source for more than 500,000 people (Lericollais and Schmitz, 1984). The crop was seeded in November when the annual flood was receding and harvested from February to April, during the dry season. Flow regulation from dams on the Senegal River negatively impacted floodplain agriculture and grazing, although managed flood releases have subsequently restored some floodplain productivity (*see* Floodplain Reconnection for Restoring Livelihoods) (Hamerlynck and Duvail, 2003) The Tana River floodplain in Kenya supported more than 115,000 Pokomos farmers who depended on recession agriculture prior to dam construction in the 1980s (Emerton, 2003).

In more humid climates, flood-recession agriculture is generally practiced in combination with rainfed agriculture and with various other activities such as fishing, forestry, and livestock-keeping. In these situations, flood-recession agriculture may not be the primary source of food, but it plays an important complementary role to rainfed agriculture. In Tanzania's Lower Rufiji floodplain, more than 100,000 farmers practice flood-recession agriculture within a diversified agricultural system. Farmers plant rainfed maize for the short rains in October–November, which they harvest in December and January. They subsequently plant a second crop of rice that is watered by a combination of an exogenous flood peak (derived from upstream runoff) and local moisture from the long rains occurring between March and May. This crop is harvested in June. During years in which the flood provides sufficient groundwater recharge, the farmers then implement the so-called "Mlao" cultivation of pumpkin, Maize, and cow peas. As a form of flood-recession agriculture that occurs during the dry season, "Mlao" cultivation provides an important "safety net" if either the rainfall failed or the flood was so high it drowned the rice (Duvail and Hamerlynck, 2007).

For vulnerable groups with limited or no access to capital, flood-recession agriculture has many advantages compared to irrigated agriculture. Flood-recession agriculture requires no infrastructure or pumping and no capital outlay except for the seeds. It is also well-adapted to the high hydrological variability of African river systems as crops are only planted in areas likely to succeed – the areas of recent inundation that have fertile and wet sediments. Floods generally remove the dry season vegetation, which greatly reduces the need for soil preparation and weeding. In summary, the workload and capital investment for flood-recession agriculture are low and the productivity is high (Park, 1992; Mollard and Walter, 2008).

People can also convert floodplain resources into food through mobile pastoralism. In Africa, livestock keepers bring their cattle from hundreds of kilometers away to graze post-flood floodplain pastures during the dry season after upland areas have been fully grazed. Floodplain grasses such as *Echinochloa stagnina* ("Bourgou") *or Echinochloa colona* are particularly valued by pastoralists because of their high nutrient content. Approximate estimates suggest that African floodplain pastures can sustain between 40 and 80 head of cattle per km^2 for up to 8 months, whereas the average surrounding dryland Sahelian pastures can only support 6–8 head per km^2 for 4 months (Marie, 2009). For example, in the inner Niger delta in Mali, approximately 20,000 km^2 of flooded surface area, encompassing 1600 km^2 of *E. stagnina*, supports approximately 1.2 million cattle (60 head per km^2; Marie, 2009).

Fisheries

River systems that continue to inundate extensive floodplains support the most productive freshwater fisheries in the world. Floodplains support river fisheries by boosting total productivity and by providing favorable habitat conditions for fish. Productive river–floodplain fisheries are characterized by flood pulses that are predictable and frequent (i.e., a regular annual flood pulse that occurs during the same season) and are of long duration. River systems with these hydrological characteristics are often characterized by a fish fauna with numerous species that have evolved to take advantage of floodplain resources and habitat (King *et al.*, 2003; Winemiller, 2004).

Floodplains support productive fisheries because floodplain habitats often provide beneficial conditions for fish. During floods, the main channels of rivers have high-velocity, turbid water. Floodplains can provide refuge from these harsh conditions with slower-velocity water due to lower gradients and increased hydraulic roughness from trees and other vegetation. During prolonged flooding, floodplain water can become less turbid than river water and warmer, which can benefit fish in cold or temperate regions (Sommer *et al.*, 2001). Flooded terrestrial vegetation and aquatic plants provide abundant cover from predation and, compared to channel habitats where both prey and piscivorous predators can become concentrated, extensive flooding allows fish to disperse over wide areas. These conditions – slow water with extensive cover and reduced concentrations of piscivorous fish – can be ideal for larval and juvenile fish (Welcomme, 1979).

Floodplain habitats also offer abundant food resources for fish, including terrestrial items such as insects, seeds, fruits, and leaves (Chick *et al.*, 2003; Winemiller, 2004), and aquatic food sources. The high primary productivity within the water column of floodplains (Ahearn *et al.*, 2006) supports concentrations of zooplankton and aquatic invertebrates that can be orders of magnitude greater than within the adjacent river (Sommer *et al.*, 2004; Grosholz and Gallo, 2006). The abundant resources on floodplains can provide important nutrition to fish prior to spawning (Moyle *et al.*, 2004) and for larval and juvenile fish after hatching (Ross and Baker, 1983; Sommer *et al.*, 2001; Agostinho *et al.*, 2004; Grosholz and Gallo, 2006).

Illustrating linkages between floodplain inundation and fish productivity, Welcomme (1979) reported that the productivity of fisheries in the Danube River was directly proportional to the extent and duration of inundation of floodplains, and Risotto and Turner (1985) found that the amount of bottomland hardwood forest, which they used as a proxy for floodplain, was a significant predictor of the amount of fish biomass produced by various sections of the Mississippi River and its tributaries. Sommer *et al.* (1997) found that Sacramento splittail year-class strength was correlated with the duration of flooding.

Because floodplains provide high-resource environments, many fish species time their reproduction to coincide with the annual flood pulse and/or spawn within inundated floodplains so that their larvae and juveniles can rear on floodplains (Welcomme, 1979). In tropical rivers, many of the most abundant and economically important fish species spawn only within floodplain habitats, with spawning periods coinciding with predictable, annual flooding (Welcomme, 1979; Hogan *et al.*, 2004).

Rivers that exhibit annual flood pulses onto extensive floodplains have significantly higher productivity of fish per unit area than waterbodies that lack a dynamic flood pulse, including regulated rivers or reservoirs, a phenomenon characterized as the "flood-pulse advantage" (Bayley, 1991, 1995). For African floodplains, Welcomme (1979) estimated an annual production of 50 kg of harvestable fish biomass per ha flooded. The Mekong River supports the largest freshwater fishery in the world and its productivity is derived from the extensive floodplains and lakes inundated by the annual flood pulse. Approximately 60 million people derive their primary source of protein from the fish harvest. Because much of this harvest consists of small fish consumed whole, fish also provide a primary source of calcium and other micronutrients in a region with little access to dairy products (Coates *et al.*, 2003). In the Rufiji River (Tanzania), fish diversity and lake productivity were higher in the well-connected floodplain lakes than in those rarely connecting to the river (Hamerlynck *et al.*, 2011).

In developing countries, small-scale freshwater fish harvests total 14 million tons, with 66% from Asia, 25% from Africa, and 4% from Latin America. These fisheries provide employment to 60 million people and are a primary source of cash income to rural families in Africa and southern Asia. Annual harvest from the Mekong River averages 2 million tons and West African river harvests exceed 1 million tons (UNEP, 2010).

Flood Attenuation

As described above (*see* Key Floodplain Processes and Benefits), floodplains are geomorphic surfaces that are created by flood processes. Rivers in alluvial valleys have the cross-sectional area to convey flows within their banks up to a certain threshold; flows above that level are conveyed by both the river and floodplain. The frequency with which floodplains carry flow varies between rivers, based on characteristics such as river and valley gradient, sediment load, watershed drainage area, and precipitation and runoff regime (Welcomme, 1979; Mitsch and Gosselink, 2000), and ranges from nearly every year to once every few years. In general, floodplains flanking high gradient rivers, and rivers with small drainage areas, have more frequent and "flashier" periods of inundation, whereas floodplains along very low gradient rivers with large drainage areas have fewer peak events with a smoother flood pulse (Mitsch and Gosselink, 2000). For example, portions of the Amazon floodplain are regularly flooded by a single pulse that lasts up to two-third of the year due to the river's enormous drainage basin with seasonal runoff patterns (Bayley, 1989).

Hydrologically connected floodplains influence flood dynamics and can reduce flood risk to other areas (**Figures 1** and **2**). Floodplains increase the cross-sectional area to move floodwaters (conveyance) and can also store water and thus reduce the height (stage) of floodwaters ("peak-shaving"). As an illustration of floodplains' capacity to act like a storage reservoir, the extensive flood basins that historically flanked the Sacramento River (California, USA) temporarily stored and then released approximately 5 bcm (billion cubic meters) annually, effectively delaying peak outflow from the river system by a month (The Bay Institute, 1998).

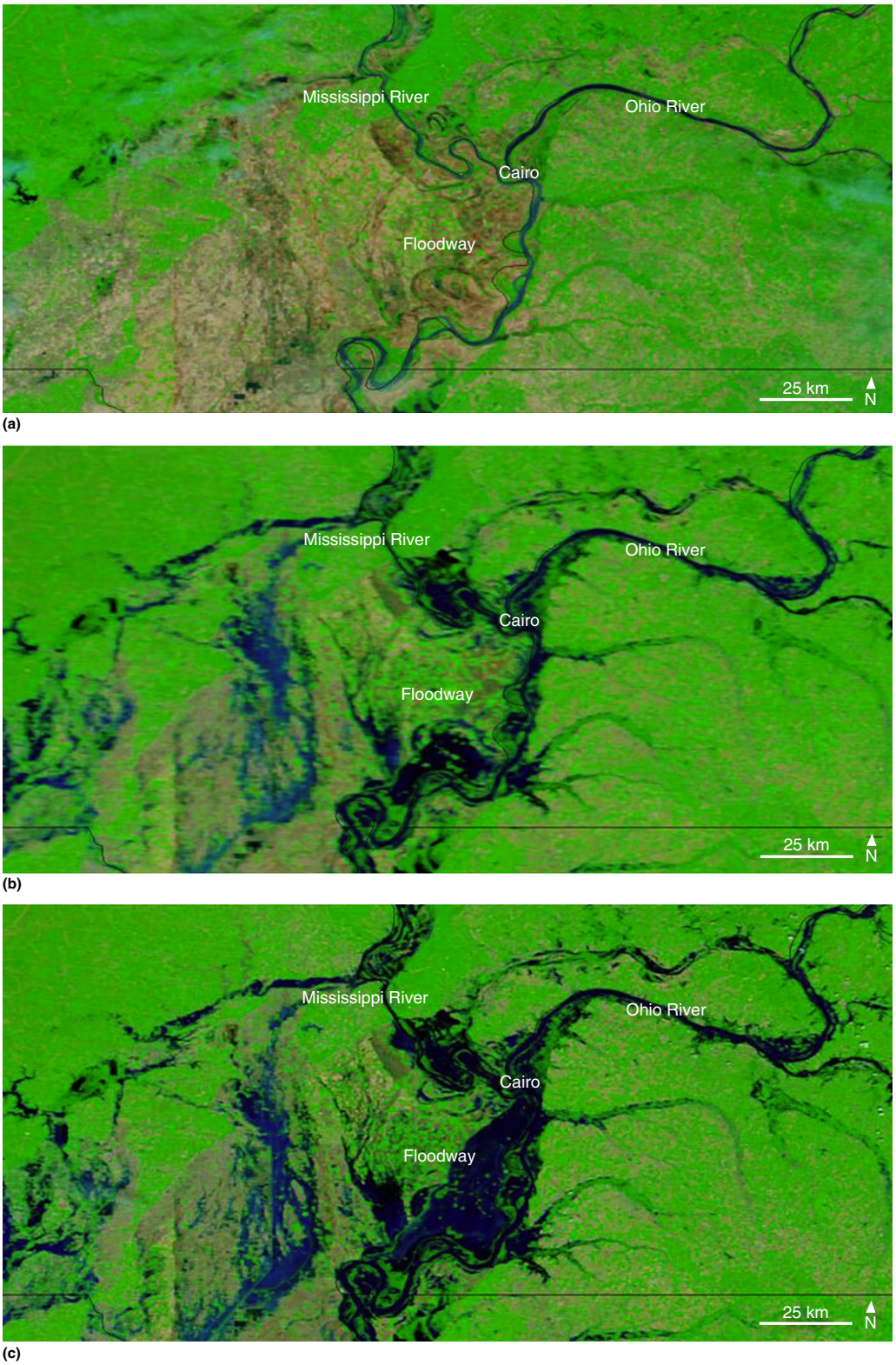


Figure 1 The Bird's Point-New Madrid Floodway in southeastern Missouri (USA). The floodway is southwest of the confluence of the Ohio and Mississippi rivers at Cairo, Illinois. (a) The Ohio-Mississippi confluence in May 2010 with river discharge within banks; area denoted as "floodway" is dry. (b) The Ohio-Mississippi confluence in early May 2011 during the 2011 flood. Note the floodplain inundation along both rivers and within the southern end of the Bird's Point-New Madrid Floodway. (c) The same location after the floodway levee was breached and floodwaters have entered the floodway. Figure 2 illustrates how using the floodway influenced river stage during the flood. Images courtesy of NASA/GSFC, Rapid Response, http://earthdata.nasa.gov/data/near-real-time-data/rapid-response#

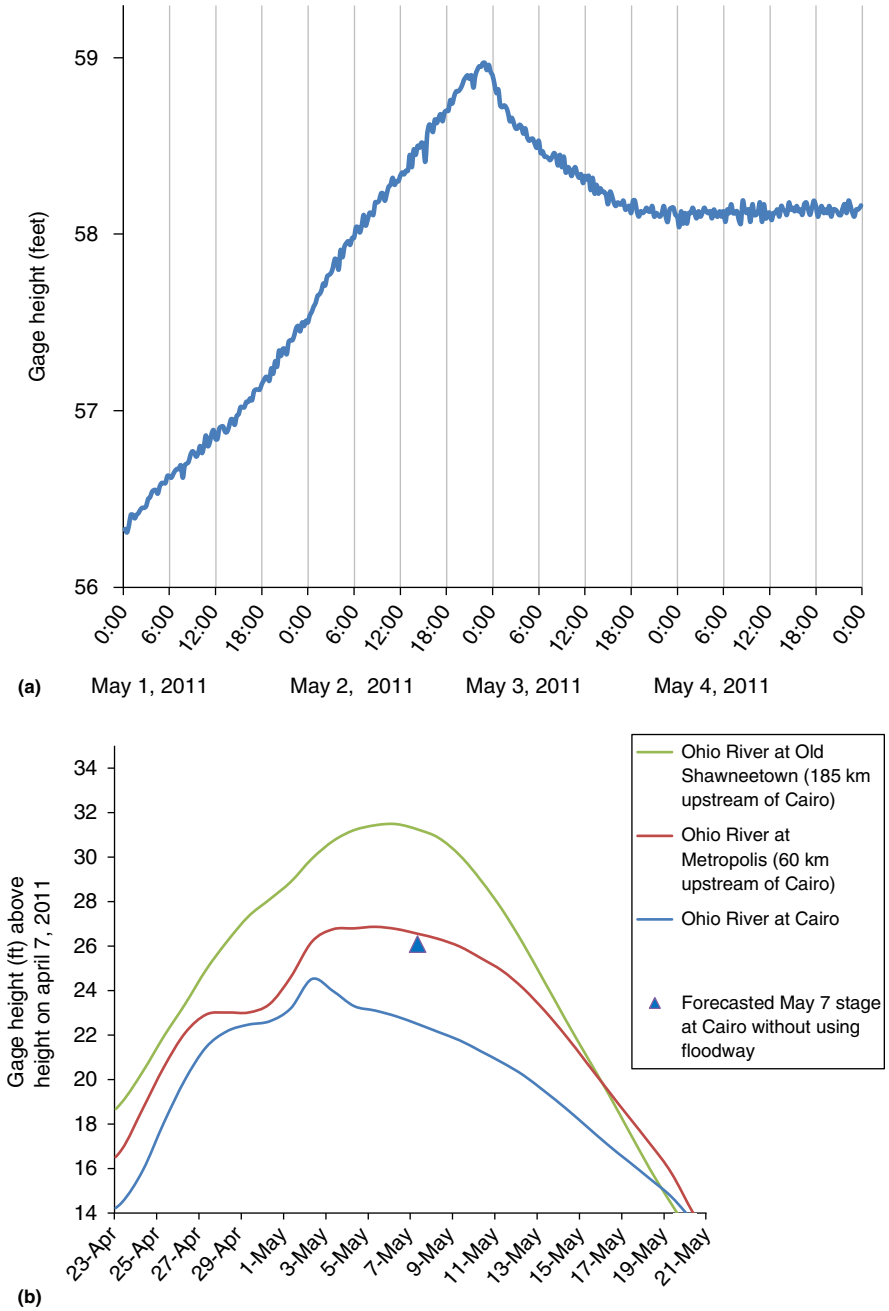


Figure 2 River gages that show the effects of opening the Bird's Point-New Madrid Floodway to increase conveyance in the Mississippi River and to reduce backwater flooding along both Ohio and Mississippi rivers. (a) Real-time gage readings for the USGS gage on the Ohio River at Dam 53 near Grand Chain, IL, approximately 30 km upstream of the floodway. The floodway was opened at 23:00 (11 pm) on 2 May, and the gage shows a rapid response with stage declining approximately one foot in 18 h. (b) River stage on the Ohio River at Cairo, IL, adjacent to the floodway (blue line), Metropolis, Kentucky (red line) approximately 60 km upstream of Cairo, and Old Shawneetown, Illinois (green line), approximately 185 km upstream of Cairo. For display, river stages at all three gages are plotted relative to their height on 7 April 2011. (On their respective gages, Cairo = 36.5 ft; Metropolis = 34.8 ft; Old Shawneetown = 24.8 ft on 7 April). At Cairo, river stage began to decline on 3 May after opening of floodway and declined for the rest of the flood event. The gage at Old Shawneetown shows that the Ohio River flood did not crest upstream until 6 May. Thus, the declining stage at Cairo reflected the increased floodplain conveyance due to the opening of the floodway and not a recession in the flood. The 7 May gage elevation at Cairo was 3.6 ft below the Corps' forecast for Cairo without operation of the floodway. Metropolis and Old Shawneetown gages operated by USGS and Cairo gage operated by the US Army Corps of Engineers. Reprinted with permission from Ted Sommer.

Flood-risk reduction of floodplains can be one of the most important benefits provided by floodplains. For example, a study that assigned economic value to floodplains in Cook County, Illinois (USA) attributed 86% of their value to the regulation of floodwaters (Sheaffer *et al.*, 2002) and Costanza *et al.*, (1997) concluded that "disturbance regulation" represented 37% of the value of floodplains. The presence of extensive hydrologically connected floodplains can reduce flood risks for downstream communities along a river (Klijn *et al.*, 2004) and, conversely, the cumulative loss of upstream floodplain storage, for example, through constriction of the floodplain by levees, can increase flood stages, and thus risk (Criss and Shock, 2001; Pinter *et al.*, 2006).

Along rivers flanked by levees, both peak shaving and increases in conveyance can occur through unintentional levee failures, which allow previously disconnected floodplains to convey and store floodwaters and reduce flood levels downstream. These unintentional flood-risk reduction benefits of floodplains occurred due to levee breaks in the 1927 Mississippi River flood, sparing the city of New Orleans (Barry, 1997), as well during floods on the Missouri and Mississippi Rivers in 1993 (Interagency Floodplain Management Review Committee (IFMRC), 1994) and the Sacramento and San Joaquin rivers in 1997 (McCarthy, 1997). For example, after a levee failed at Sny Island, allowing the flooding of 44,000 acres, the flood stage at Hannibal, Missouri lowered from 31.5 to 27 feet in 2 days (Tobin, 1995). The role of both intentional and unintentional floodplain storage and conveyance are discussed in greater detail below.

Deriving Benefits from Connected Floodplains

The various benefits described above require three primary characteristics: hydrological connectivity, a range of flow levels, including high flows that exceed bankfull, and sufficient spatial scale for the benefits to accrue to a meaningful level (Opperman *et al.*, 2010).

Although exchanges of groundwater between the river and floodplain can influence hydrological conditions for floodplain wetlands and vegetation, surface-water connectivity is required for the performance of geomorphic processes and the exchange of sediment, organic material and organisms between river and floodplain. As river stage rises, surface-water connectivity can occur through lateral overflow, breaks in natural levees (such as sloughs or side channels), unintentional levee breaches, and gates, weirs, or other control structures. Connectivity is maintained during a flood's recession although the dominant direction of flow is reversed as the floodplain drains back to the river. Following disconnection of surface waters, floodplains can continue to drain back into the river through subsurface pathways and contribute to river baseflow (Ward, 1989).

Floodplain processes are supported by a broad range of flow levels and flow events, ranging from base flows to infrequent, high-magnitude floods (Poff *et al.*, 1997). High-magnitude flood events perform geomorphic work, such as erosion and deposition, and maintain and create floodplain habitat features. Long-duration flooding can be particularly important to many of the benefits described above (Junk *et al.*, 1989; Williams *et al.*, 2009), such as groundwater recharge and fisheries productivity, as fish reproduction and juvenile rearing can require weeks to months of inundation.

A variety of spatial variables influence the production of benefits from floodplains, such as the floodplain's topographic features and position within the river basin. The extent of connected floodplain is critically important to the production of benefits as floodplain productivity, for fisheries or agriculture, and flood-risk reduction benefits increase directly with area. For floodplain benefits to accrue to a meaningful level, sufficient area of floodplain must be connected.

Changes to Floodplains and Challenges for Water and Floodplain Management

For millennia, people have occupied and used floodplains, drawn by their flat land with fertile soil, proximity to transportation, water supply and water pollution dilution, and the aesthetics of water itself. As a result, floodplains have undergone dramatic conversion globally and are considered one of the most converted and threatened ecosystem types on the planet (Tockner and Stanford, 2002). Particularly in the earlier developing world, most large rivers have become disconnected from a significant portion of their floodplains. For example, less than 10% of historic floodplains remains hydrologically connected to rivers in California's Central Valley (Bakker, 1972; Barbour *et al.*, 1991; The Bay Institute, 1998) and Mississippi River floodplain forests below the Ohio River have declined by 80% from their historic extent (Llewellyn *et al.*, 1995). In Africa, most of the large rivers have been dammed for irrigation projects causing considerable changes to floodplain extent and processes and the economies that depended on them (Adams, 1996; World Commission on Dams, 2000).

In this section, the authors review the processes that disconnect floodplains from their rivers and facilitate land-use conversions. These changes and conversions have produced many benefits, such as the globally important agricultural economies along major rivers such as the Yangtze and Mississippi. However, these changes have also produced impacts and unintended consequences, such as losses of traditional food-production systems (Adams, 1996; Richter *et al.*, 2010) or increases in flood risk caused by upstream efforts to contain flood flows between levees (Criss and Shock, 2001; Pinter *et al.*, 2006). After describing the processes contributing to floodplain disconnection, the authors review a set of water and floodplain management challenges that can be addressed, in part, by maintenance of restoration of floodplain connectivity.

Just as floodplain benefits require connectivity and flow variability, changes to connectivity and flow are the primary actions leading to floodplain disconnection and land-use conversion. Levees and floodwalls are built to contain floodwaters and physically disconnect floodplains from river surface water. Levees have been constructed along major rivers for centuries and were first built on the Yangtze and Yellow Rivers more than 4000 years ago, and on the Rhine, Mississippi, and Sacramento Rivers in the eighteenth and nineteenth centuries (Sayers *et al.*, 2011). Levees are generally built to protect existing development investments on floodplains, such as agriculture and residential and commercial buildings, and/or to facilitate future

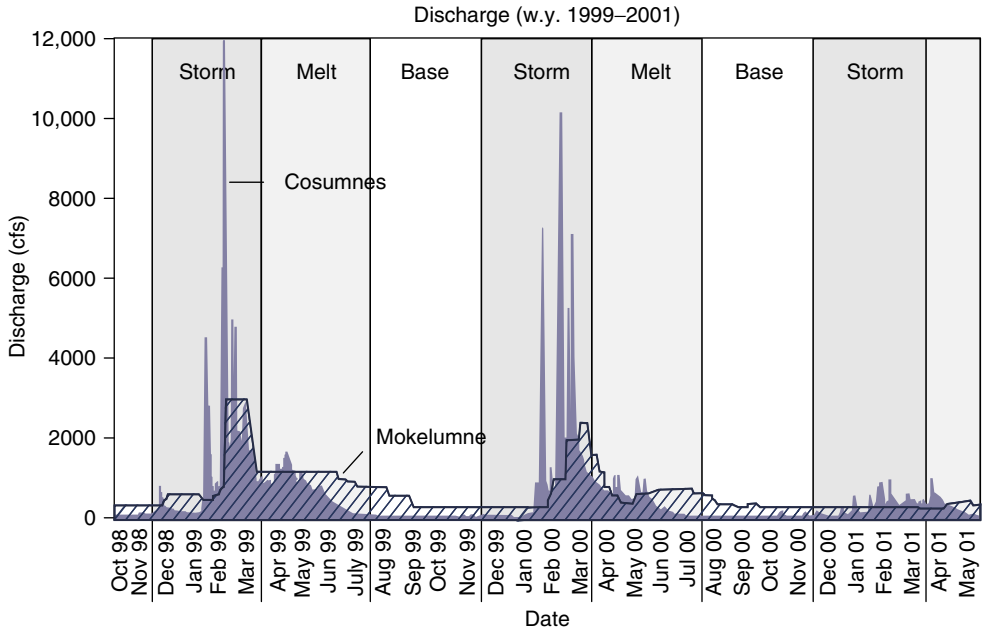


Figure 3 Discharge of two adjacent rivers in California's Central Valley. The Cosumnes is essentially unregulated and displays natural flood peaks (solid line). The Mokelumne (blue line with hatching) is regulated by two large dams that provide flood control, in addition to hydropower and water-supply storage. Reprinted with permission from Jeff Mount.

investment. The USA currently has approximately 23,000 km of levees managed by the US Army Corps of Engineers and potentially more than 160,000 km of nonfederal levees (National Levee Safety Committee, 2009). Dams regulate river flows, with flood-control dams operated specifically to store floodwaters and reduce downstream flood peaks (**Figure 3**). Dams operated for water supply, irrigation, and hydropower also often reduce downstream flood peaks because they capture water during periods of high flow to store and release later during periods of low flow (Ligon *et al.*, 1995).

Land and Water Management Challenges

Current water and floodplain management, including the extensive conversion of floodplains from largely natural to developed landscapes, has produced numerous benefits for society. However, this land and water management has also created a legacy of impacts, and current approaches may not be adequate for future challenges. The sections below describe several of these challenges, both legacy and future, which can be addressed in part through floodplain reconnection or conservation.

Declines in Freshwater Biological Diversity and Productivity

The extensive conversion of floodplains and alteration of rivers has contributed to dramatic declines in biodiversity and populations of commercially, recreationally, or culturally important species. Freshwater species are imperiled at far higher rates than terrestrial or marine species (Ricciardi and Rasmussen, 1999) and a primary cause is fragmentation, flow alteration, and river–floodplain disconnection caused by dams and levees (Poff *et al.*, 1997; Richter *et al.*, 1997; Tockner and Stanford, 2002). Freshwater fisheries have declined dramatically from rivers following widespread floodplain disconnection and flow alteration. For example, the Missouri River commercial fish harvest declined 80% from the 1880s to the 1990s due to flow regulation and levee construction (Galat *et al.*, 1998).

Impacts to River-Dependent Communities and Livelihoods

These declines in floodplain extent and function have had negative impacts on millions of river-dependent people worldwide (Richter *et al.*, 2010). Along the lower Tana River in Kenya numerous indigenous communities derived socio-economic values from the annual inundations of the floodplain and delta, including fisheries, flood-recession agriculture, and pastoralism (Terer *et al.*, 2004). The construction of the series of hydropower dams in the upper basin altered flow regimes including decreasing peak flows during the wet season (Maingi and Marsh, 2002). The resulting reduction of flooded surface area, flood peak duration, and meandering dynamics have negatively affected floodplain productivity and forest regeneration and therefore the local economies (Hamerlynck *et al.*, 2010).

In the 1980s, the Senegal River Basin Development Authority, representing Mali, Guinea, Senegal, and Mauritania developed a project on the Senegal River to produce hydropower, improve navigation, and develop irrigated agriculture on 375,000 ha of the valley's floodplains. The project consisted of two large dams: a multipurpose dam at Manantali in Mali (hydropower and irrigation) and a "salt wedge" dam at Diama close to the river mouth, to prevent saltwater intrusion. Expectations for the productivity of irrigated agriculture have not been met, in spite of huge investments (1500–6500 US$ per ha) (Hamerlynck *et al.*, 2005). The

negative impacts of the dams on fisheries (Bousso, 1997), the environment, and human health (Verhoef, 1996) have been considerable and the livelihoods of the valley's inhabitants have been seriously affected (Adams, 1999). For example, flood-recession sorghum farming diminished drastically in the 1970s and 1980s due to both drought and flow regulation from upstream dams, but rebounded somewhat after the failure of the irrigation schemes and the implementation of the managed flood releases from Manantali (*see* Floodplain Reconnection for Restoring Livelihoods) (LeRoy, 2008).

Rising Flood Risk and Aging Infrastructure

Although flood-management infrastructure in the USA has allowed greater economic utilization of floodplains, it has also compromised many of the natural benefits provided by river–floodplain systems and encouraged development in places vulnerable to periodic damaging floods (Association of State Floodplain Managers, 2008). Flood damages in the USA have increased over the past century, from less than US$1B annually in the 1930s to more than US$6B currently (in constant dollars) (Pielke *et al.*, 2002), even as billions have been spent to build and maintain dams and levees to manage floodwaters.

Several trends suggest that flood risk will continue to increase in the USA and in other developed countries. First, development continues in places vulnerable to flooding whereas upstream land-use changes, such as an increase in impervious surfaces, continue to alter runoff and downstream flood dynamics (Pinter, 2005). Many of the nation's dams require extensive rehabilitation, with an estimated 5-year price tag of US$12.5B including 1743 high hazard dams in need of repair. Levee maintenance is chronically underfunded (Leavenworth, 2004b), requiring an estimated 5-year investment of US$50B (American Society of Civil Engineers, 2009). Further, forecasts indicate that climate change may lead to greater flood magnitudes and frequencies. Evaporation rates increase with air temperature, and warmer air can hold more water vapor, increasing the intensity of heavy precipitation events (Kundzewicz *et al.*, 2008). Owing to these trends, flood risks are greater than commonly perceived and are growing.

Competition between Water-Management Sectors

Reservoirs that provide flood control often also provide multiple purposes, such as water supply, hydropower, and recreation, and these various objectives can compete with each other. Flood control is maximized by maintaining low reservoir levels during periods of flood risk whereas hydropower generation is maximized by high reservoir levels. Water supply risk can increase when reservoirs are lowered for flood control; for example, if the basin enters a period of drought after reservoir drawdown. These conflicts may be exacerbated by climate change, which is forecasted to increase the "flashiness" of precipitation and runoff and thus may increase both flood and drought risk in many parts of the world (Kundzewicz *et al.*, 2008).

Floodplain Reconnection and Conservation

This section reviews the methods to maintain or restore floodplain connectivity, encompassing both physical connectivity and the flow regime. Several examples of floodplain reconnection for a variety of primary objectives are then described.

Methods to Maintain or Restore Floodplain Connectivity

Hydrologically connected floodplains provide a range of benefits. Achieving these benefits requires either maintaining existing floodplain connectivity or reconnecting floodplains. Connected floodplains can be maintained through real-estate transactions such as acquisition, easements, or government funding programs, such as the Wetland Reserve Program. For example, to maintain the capacity for floodplains to be inundated along the Savannah River (Georgia/South Carolina, USA), The Nature Conservancy (TNC) is working with willing landowners to acquire or place under easements floodplain land to maintain it in low-intensity land uses (Opperman *et al.*, 2010). In the 1970s, the US Army Corps of Engineers studied alternatives for reducing flood risk for Boston from the Charles River and concluded that acquisition and protection of floodplain wetlands could provide similar protection as new levees and flood-control reservoirs for approximately one-tenth the cost. The Corps acquired 3440 ha in the Charles River floodplain for approximately US$10 million (Postel, 2005). In addition to acquisition or easements, emerging markets for ecosystem services, including carbon and nutrient sequestration, floodwater storage, and recreation may be able to provide revenue to landowners that maintain floodplains connected to rivers (Opperman *et al.*, 2009, 2010).

Reconnecting floodplains can be accomplished through physical reconnection (removing barriers between river and floodplains), hydrological reconnection (such as the release of environmental flows), or a combination of the two. Managing dam operations to release high flows to mimic floods is a specific example of environmental flows (Postel and Richter, 2003).

Floodplain connectivity on the Cosumnes River (California, USA), an unregulated river draining the Sierra Nevada (**Figure 3**), was restored by cutting breaches through an existing levee. Levee breaches were also used on Mollicy Farms to reconnect 6400 ha of floodplain to the Ouachita River (Louisiana, USA) (Opperman *et al.*, 2010). In many locations in the USA, levees have been set back further from the river to reconnect portions of the floodplains, such as on the Sacramento, Bear and Feather rivers in California (Williams *et al.*, 2009; Opperman *et al.*, 2010). Along both the Sacramento and Mississippi Rivers, flood-management strategies evolved from a "levees only" approach, which disconnected nearly the entire floodplain from the river, to the reconnection of large areas of floodplain (tens of thousands to hundreds of thousands of hectares) as flood bypasses or floodways (US Army Corps of Engineers, 2004). The Yangtze River (China) has similar features known as Flood Detention Areas, while the Meuse, Rhine, and Elbe rivers in northern Europe have a system of polders that can be intentionally flooded to reduce flood stages in the river (Klijn *et al.*, 2004; Forster *et al.*, 2005) . For these features, river–floodplain connectivity is regulated by weirs, fuseplug levees, or gates. The frequency of connection of these managed floodplains ranges from nearly every year (Yolo Bypass on the Sacramento River; Sommer *et al.* (2001)) to once or twice a

century (New Madrid Floodway on the Mississippi and the Jingjiang Flood Detention Area on the Yangtze) (US Army Corps of Engineers, 2004).

The operation of dams can be modified to restore characteristics of the flood pulse to increase river–floodplain connectivity and promote floodplain benefits. Flood releases from reservoirs can be designed to replicate various characteristics of historical flood patterns, such as magnitude, season, duration, or rate of change (Postel and Richter, 2003). For example, an experimental flood released from Alamo Dam on the Bill Williams River (Arizona, USA) favored the regeneration and growth of native riparian willows (*Salix spp.*) over invasive Tamarisk (Shafroth *et al.*, 2009). On the Truckee River (Nevada, USA), dam operators mimicked the gradual recession of the snowmelt hydrograph and promoted regeneration of the willow-cottonwood floodplain forest (Rood *et al.*, 2003). Several dams in Africa have released experimental floods intended to restore downstream floodplain ecosystems and livelihoods (Hamerlynck and Duvail, 2003; Loth, 2004).

Managed flood releases cannot address all impacts from dams, but can restore some ecosystem functions and thus continue to ensure the livelihoods of the downstream populations (Acreman, 2003). A variety of methods can be used for developing environmental flow recommendations, including managed floods, (Tharme, 2003) and integrating these environmental flows into dam operations (Harman and Stewardson, 2005; Beilfuss, 2010; Esselman and Opperman, 2010).

For rivers with natural or close-to-natural flow regimes, restoration of physical connectivity will allow restoration of floodplain processes and many associated benefits. For rivers below dams, restoring physical connectivity will likely not restore all floodplain processes, depending on the degree and type of flow alteration. For example, a levee setback project on the Sacramento River at Hamilton City will increase the connectivity between river and floodplain for a range of flow types. However, without alternative flow management from the upstream Shasta Dam, the reconnected floodplain will only rarely experience ecologically important long-duration spring flooding, which was frequently common (Williams *et al.*, 2009; Opperman *et al.*, 2010). Similarly, projects focused on restoring flood releases from dams may also need to consider the downstream physical connectivity. For example, restoration of floodplain processes on the Logone River (Cameroon) required both managed flood releases and removal of barriers to river–floodplain connectivity (Loth, 2004).

Floodplain Reconnection for Restoring Ecosystems and Services

The Cosumnes River is one of the only remaining major rivers draining the western slope of the Sierra Nevada that does not have any large dams on the mainstem and, therefore, high flows on the Cosumnes are mostly unchanged from their natural condition (Moyle *et al.*, 2007) (**Figure 3**). TNC, along with numerous partners, established the Consumnes River Preserve (CRP) along the lower river. In 1985, an agricultural levee adjacent to the CRP breached, allowing flooding and sediment deposition. Riparian trees, including cottonwoods and willows grew rapidly in the freshly deposited sediment. This area, now referred to as the "accidental forest," became part of the preserve in 1987 and currently has a tall canopy of cottonwoods with an understory of valley oaks (Tu, 2000; Swenson *et al.*, 2003). Owing to the successful riparian regeneration following the 1985 breach, TNC sought to increase connectivity between the river and floodplain through additional breaches in the remnant agricultural levee. After hydrological modeling indicated that a new breach would not increase flooding for neighboring landowners, TNC opened up two new breaches in the levee in 1995 and 1997. (Swenson *et al.*, 2003). Prior to the breaches, the levee contained floods up to about the Q_5 (0.20 exceedance probability in any year). After the breaches, hydraulic connectivity initiates at flows only half the size of the mean annual flood, with an annual exceedance probability of 0.95 (Mount *et al.*, 2003).

Florsheim and Mount (2002) demonstrated the significance of rare, large flow events in creating dynamic floodplain topography on the Cosumnes. Sand splay complexes formed at levee breaches increase topographic heterogeneity on the floodplain and promote recruitment of riparian trees (Tu, 2000; Florsheim and Mount, 2002). The Cosumnes floodplain is frequently flooded for long durations (weeks to months) in the spring, which is particularly important for promoting food-web productivity on floodplains. This type of flooding was once common in California's Central Valley but now, due to flow regulation and levees, is relatively rare (Williams *et al.*, 2009). During this long-duration flooding, the floodplain is the site of high algal productivity (Ahearn *et al.*, 2006), which leads to concentrations of zooplankton and invertebrates that are 10–100 times greater on the floodplain than in the river (Grosholz and Gallo, 2006). This productivity benefits native fish (Moyle *et al.*, 2007) and Jeffres *et al.* (2008) demonstrated that juvenile salmon rearing on the floodplain grow significantly faster than juvenile salmon rearing in adjacent river habitats.

At Mollicy Farms along the Ouachita River (Louisiana, USA), a tributary to the Mississippi River, TNC also opened levee breaches to restore connectivity to 6400 ha of marginal agricultural land on the floodplain. The restoration is projected to have local benefits for wildlife but a key feature of the restoration project is research and monitoring to determine the potential for large-scale reconnection to promote system-scale benefits. Researchers at Mollicy Farms are quantifying the ecosystem services of recreation, carbon sequestration, and nutrient sequestration. This research is intended to further understand the potential for markets for these ecosystem services and for these markets to provide revenue to landowners that would be comparable to crops (Opperman *et al.*, 2010). Nutrient sequestration could potentially be compensated through a market-based approach to reduce nutrient loading to the Gulf of Mexico, which results in a massive "dead zone" each year (Mitsch *et al.*, 2001).

Floodplain Reconnection for Restoring Livelihoods

Cameroon's Logone River was dammed in 1979, forming Lake Maga, and subsequent flow regulation diminished the extent of annual flooding downstream, negatively impacting floodplain-dependent communities. In the 1990s,

dam managers released experimental floods intended to reinundate the floodplain and restore economic activities such as grazing and fishing (Loth, 2004). In the experimentally flooded area, cattle numbers increased 260% in part due to an increase in cover of nutritious perennial grasses such as *Echinochloa pyramidalis* and *Oryza longistaminata*. Fish harvest within the flooded area was 90 kg ha^{-1} (Loth, 2004).

On the Senegal River, flow regulation from dams greatly reduced flooding and, by the early 1990s, most of the Mauritanian lower delta of the Senegal River had turned into a saline desert. In the dry season, hypersaline conditions downstream of Diama Dam were degrading natural vegetation communities (*Avicennia germinans* mangrove, *Sporobolus robustus* grasses, and *E. colona* pastures). Between 1994 and 1996, water managers constructed embankments and sluicegates to manage flows to simulate the predam hydrodynamics on the former floodplain. Through a participatory approach, the water requirements of various local user groups (fishers, livestock keepers, and gatherers) and of the National Park authority were analyzed, modeled, and reconciled with the imperatives of the basin's water-management authority. A number of flood scenarios to increase floodplain productivity were proposed, discussed, tested, and formalized in a management plan approved by all stakeholders and government institutions in 1997. The effect of the managed flood releases on various components of the ecosystem was quantified and the flood scenario improved through an iterative consultative process (Hamerlynck and Duvail, 2003).

The Sahelian ecosystems proved resilient and most of the floodplain functions were restored after a few years of managed flooding, although restoration of the floodplain-associated Acacia and mangrove forest took longer (at least 10 years) (Hamerlynck and Duvail, 2003). The restoration of floodplain productivity allowed local inhabitants to resume their traditional livelihood activities. An economic evaluation has shown that the restored area produces about 65 US$ of added value per ha per year through the extractive activities of the local communities (Moulaye Zeine, 2004). By stemming the rural-to-urban migration of young men this economic revival has also contributed to an improved social equilibrium. Interestingly, the stakeholders recommended, and managers were able to implement, a period of inundation during the May dry season when the floodplain would have been dry under the natural hydrological pattern. This counter-season inundation has increased the productivity of the floodplain compared to the natural flooding pattern and provides access to water for cattle during the dry season and can serve as emergency pasture (Duvail and Hamerlynck, 2003).

Floodplain Reconnection for Flood-Risk Reduction

Traditional approaches to managing floods have emphasized structural solutions, such as levees and dams. Increasing flood damages in spite of massive investments in structural approaches highlighted the need for nonstructural measures, such as land-use planning, building codes, insurance, and use of the natural and beneficial functions of floodplains, including storage and conveyance. Accompanying the increased emphasis on nonstructural measures, many water managers and floodplains managers now refer to "flood-damage reduction" or "flood-risk management" rather than "flood control" (Klijn *et al.*, 2004). The concept of flood-risk management acknowledges that no set of actions can guarantee the prevention of flooding; structural measures can fail or have their design level of protection exceeded and thus some level of residual risk will be inevitable. As a result, flood-risk management proposes that flooding challenges should be addressed through a portfolio of measures, including both structural and nonstructural, emphasizing those actions which provide for sustainability and long-term use of the flood attenuation characteristics of natural systems. The use of floodways and off-river storage areas are receiving significant attention as new plans are prepared to deal with the growing challenges of climate change and increasing population in flood-prone areas (Klijn *et al.*, 2004; Freitag *et al.*, 2009; Opperman *et al.*, 2009; Sayers *et al.*, 2011).

Although gaining renewed attention, the strategy of using floodplains to store and convey floodwaters is not new; the reconnection of floodplains to reduce flood risks is a feature of flood management on major rivers such as the Yangtze, Mississippi, and Sacramento. For both the Mississippi and the Sacramento, the floodplain reconnection occurred in the aftermath of major flood disasters that revealed the weakness of the existing paradigm for flood management, a "levees only" approach that attempted to confine the river between levees and essentially disconnect the entire floodplain (Kelley, 1989; Barry, 1997).

In 1927, a historic flood on the lower Mississippi overtopped, or caused the failure, of dozens of levees, inundating 70,000 km^2 and displacing 600,000 people (US Army Corps of Engineers, 2004). Hundreds, and more likely thousands, of people died (Barry, 1997). River and floodplain management prior to the flood had relied on levees to contain floodwaters and was composed of multiple, disjointed efforts. Following that flood, the US Army Corps of Engineers developed a comprehensive plan for navigation and flood protection in the lower Mississippi Valley. The plan proposed and later implemented by the Corps included use of storage behind dams, setback of levees to provide more room for conveyance of floodwaters, and the use of floodways and backwater storage areas to relieve pressure on levees protecting major urban areas (and thereby reduce the necessity to raise levees to meet growing increases in flood flows).

The system that was developed included four portions of the historic floodplain that were designated as floodways – the Birds Point-New Madrid (in southeast Missouri; **Figure 1**) and the West Atchafalaya, Morganza, and Bonnet Carré (all in Louisiana). During the highest flow events, floodwaters can leave the river and enter the floodways through fuseplug levees or gates. The floodways can carry a significant portion of water during floods, lowering flood stages elsewhere and reducing pressure on levees (**Figure 2**). For example, for the "project design flood" (the maximum flood with a reasonable probability of occurring), the New Madrid floodway is designed to convey nearly one-fourth of the flow below the confluence of the Ohio and Mississippi rivers and, near the Gulf of Mexico, the main channel of the Mississippi is designed to carry approximately 40% of total flow with the remainder moving through floodways, primarily the Old River Diversion and the

Morganza Floodway in the Atchafalaya Basin (US Army Corps of Engineers, 2008).

The effectiveness of floodplains for contributing to flood-risk management was demonstrated during major floods in the lower Mississippi Valley in 1973 and 2011 when use of floodways in Missouri and Louisiana relieved pressure on levees protecting Cairo, Illinois, and the major urban industrial areas from Baton Rouge to the Gulf of Mexico along the Mississippi. Diversion of Mississippi waters into the floodways increased conveyance in key locations and allowing water to accumulate in the backwater areas (where major tributaries enter the Mississippi) reduced peak stages. Although the flood of 2011 had a greater volume than the 1927 flood, no levee failures or deaths occurred in 2011.

The Sacramento River (California, USA) followed a similar trajectory as the Mississippi: disjointed river and floodplain management that relied almost strictly on levees, disastrous floods and levee failures exposing the flaws of a "levees only" approach, followed by a systematic approach to floodplain and river management that featured large-scale reconnection of floodplains. Flood management in the Sacramento Valley relies to a great extent on a system of bypasses, portions of the historic floodplain connected to the river by weirs. When river stage exceeds the weir elevation, floodwaters enter the bypasses. The Yolo Bypass, a portion of the low-lying Yolo Basin that flanked the Sacramento River, was reconnected to the Sacramento River in the 1930s (**Figure 4**). During major floods, the Bypass carries 80% of the volume of floodwaters while the river channel, adjacent to the city of Sacramento, carries only 20% (Sommer *et al.*, 2001). As described further below, the Yolo Bypass is a central feature for flood-management in the Sacramento Valley. Beyond flood management, the Bypass provides a variety of other benefits. Approximately two-thirds of the Bypass is in active agriculture with annual crops. The Bypass also provides open space, recreation (primarily duck hunting and wildlife viewing), and groundwater recharge that proved valuable as a groundwater bank during a drought (Jercich, 1997; Sommer *et al.*, 2001). The Bypass is also considered one of the most important remaining floodplain habitats in the Central Valley for ecologically important Spring flooding (Williams *et al.*, 2009) and for providing habitat for native birds and fish (Sommer *et al.*, 2001). Owing to high floodplain productivity, juvenile Chinook rearing in the Bypass grow faster than juveniles rearing in the adjacent Sacramento River (Sommer *et al.*, 2001) and the Bypass has population-level benefits for the obligate floodplain spawner, the Sacramento splittail (Sommer *et al.*, 1997).

California provides several other examples of projects that reconnected floodplains with a primary objective of contributing to flood-risk reduction. The city of Napa selected a flood management plan based on levee setbacks, bypass channels, and floodplain reconnection to lower flood risk, rejecting a typical approach of dredging and tall floodwalls that would have been considerably more expensive and would have visually and physically disconnected the river from the town (Postel, 2005). On the Bear and Feather rivers (tributary to the Sacramento), the Three Rivers Levee Improvement Authority selected levee setbacks as the most effective way to overcome erosion risk to certain levee segments and to increase conveyance to reduce backwater flooding (Leavenworth, 2004a). Though driven by flood management objectives, the projects will restore thousands of hectares of riparian forest.

Removal of buildings and relocation of residents from floodplains can be a component of flood-risk management (US Army Corps of Engineers, 2001). Generally, this approach is implemented to avoid repetitive losses in areas with high flood risk. In the USA, due to the sensitivity of federal land acquisition, federally supported relocations are voluntary. For example, the town of Valmeyer, Illinois, which was inundated by the 1993 Mississippi River flood, was relocated from the floodplain of the Mississippi to bluffs approximately 2 miles away using funds from the Hazard Mitigation and Grant Program of the Federal Emergency Management Agency and other sources (Bello and Eisler, 2008). Because of the very high cost of developed land, most proposals for, or analyses of, large-scale floodplain reconnection focus on agricultural or other land that is either uninhabited or sparsely settled (Klijn *et al.*, 2004; Sparks and Braden, 2007).

Floodplain Connectivity to Increase the Operational Flexibility of Multipurpose Reservoirs

As described above, within multipurpose reservoirs, the various water-management purposes can compete with each other for reservoir allocation. Forecasts for increased risk of both drought and flood with climate change (Kundzewicz *et al.*, 2008) may exacerbate the conflict between flood control, which benefits from lowering reservoir levels in preparation for the potential need to store and regulate flood inflows, and purposes that benefit from full reservoirs (hydropower, water supply, release of environmental flows, and recreation).

Large-scale floodplain reconnection, or maintenance of connectivity, can potentially shift some of the flood-management burden onto downstream floodplains and thus reduce flood-management obligations within the reservoir. Coupling floodplain and reservoir management can potentially increase the flexibility and resiliency of multipurpose reservoirs. With reduced flood-management responsibility, a multipurpose reservoir could generate more hydropower or have greater water-supply security heading into droughts. For example, Opperman *et al.* (2011) summarize an analysis of alternative operations for a cascade of dams on the upper Yangtze River, which suggests that reducing flood storage from the reservoirs and shifting flood-risk management to the downstream floodplain can provide equivalent or lower flood risk, greater environmental performance, and hundreds of millions of dollars more in annual hydropower revenue because the reservoirs can be maintained at higher levels.

The Yolo Bypass, described above, provides another illustration of this concept (**Figure 4**). Although the Bypass was built decades before a system of multipurpose reservoirs in the Sacramento Valley, it effectively illustrates how a geographically large hydrologically connected floodplain can increase the flexibility of multipurpose reservoirs (Opperman *et al.*, 2011). By conveying 80% of the volume of floodwaters during major floods, the Bypass floodplain relieves pressure on levees and substitutes for billions of cubic meters of reservoir flood storage. Opperman *et al.* (2011) present an analysis of the role

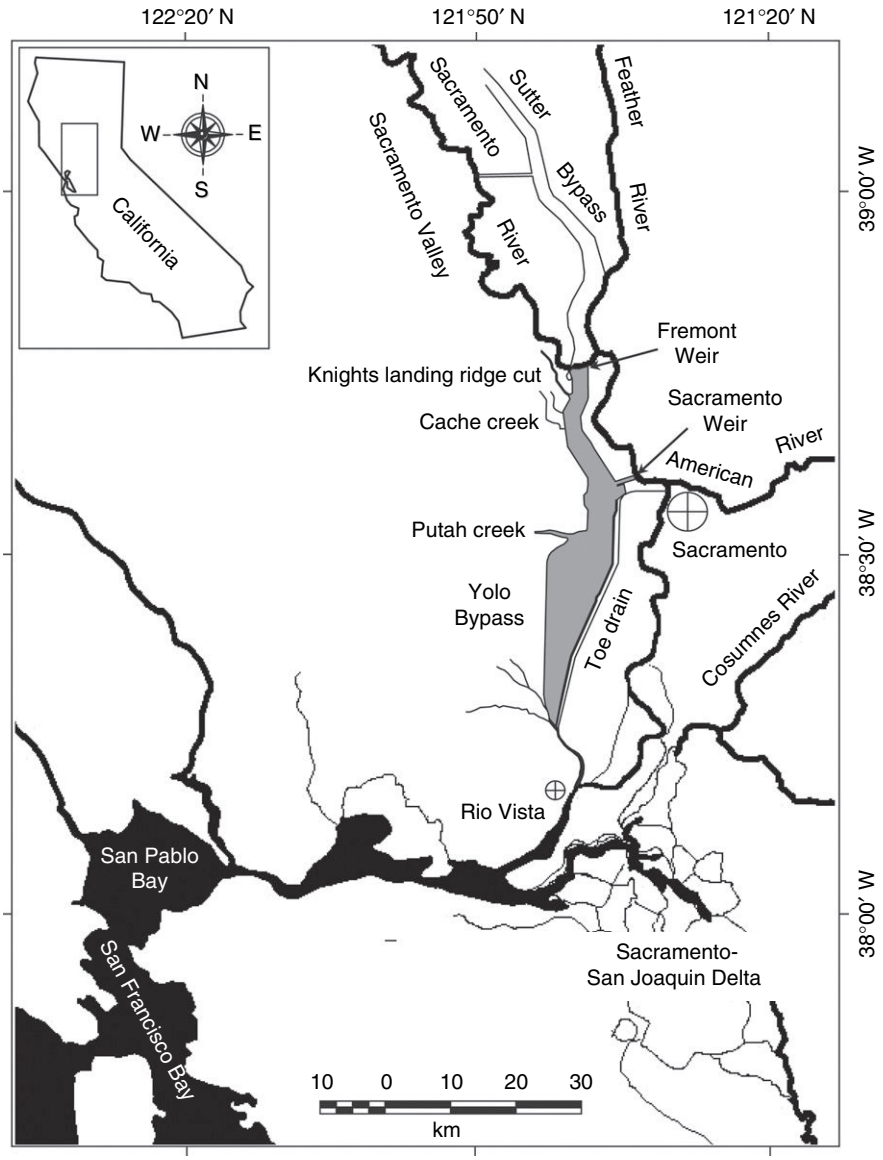


Figure 4 Location of the Yolo Bypass (shaded gray), and its major features, within the lower Sacramento Valley, California.

the Bypass played in the Sacramento Valley flood management during a flood in 1986 (a 50–80-year recurrence interval flood at the city of Sacramento). During the 3-day peak of the 1986 flood (19–21 February) 3.3 bcm flowed through the bypass, which is approximately the same volume as the combined flood-control storage of the six major reservoirs that manage floodwaters in the Sacramento River basin (**Figure 5**). During this peak, the multipurpose reservoirs' flood storage volume was nearly full and the Sacramento River was essentially at its channel capacity (i.e., greater discharge within the channel would have lead to levee overtopping). Thus, the 3.3 bcm that flowed through the Bypass during the peak 3 days could not have been stored or conveyed by any other element of the existing flood–control system. The "green infrastructure" of the Bypass floodplain was providing a service that could not have been provided by the existing engineered infrastructure.

Managing the volume of water associated with the 1986 flood without the Bypass floodplain, let alone a 100- or 200-year event, would require some combination of (1) additional flood-control infrastructure, including new dams and extensive rising of levee height to increase channel capacity; and (2) considerable reallocation of existing reservoir storage from water supply to flood control. The first would require significant funding – in a state where maintenance of the current levee system is chronically underfunded (Leavenworth, 2004b) – while the second would impinge on California's already strained water supply resources. Thus, the hydrologically connected floodplain of the Yolo Bypass is essential for the operational integrity and flexibility of the current system of water storage and allocation and provides a considerable economic value in the form of avoided costs for additional infrastructure.

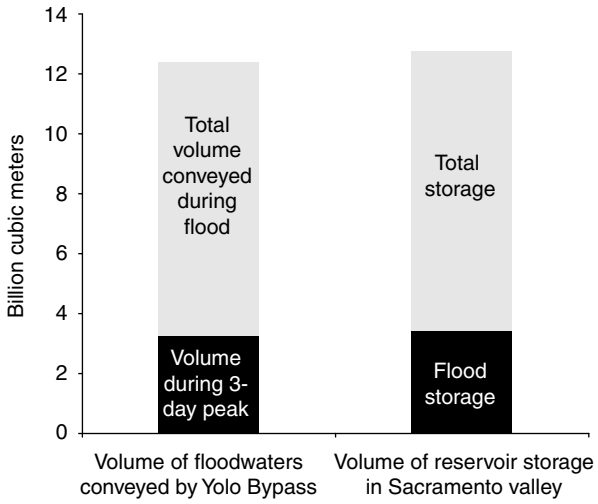


Figure 5 Volume of water conveyed by the Yolo Bypass during the 1986 flood event and volume of water in Sacramento River reservoirs.

Conclusions

Hydrologically connected floodplains provide multiple benefits to people and nature. From the perspective of environmental resources, floodplains are among the most diverse and productive ecosystems on the planet. Much of their value to people flows directly from that biological productivity and diversity and from processes that require floodplain inundation. These include fisheries, groundwater recharge, nutrient sequestration, and flood-dependent forage vegetation and flood-recession agriculture. Most of these floodplain benefits are associated with frequent and long-duration flooding, such as an annual, extended flood pulse.

However, floodplains are also globally important as sites of high agricultural productivity with crops that are not tolerant of flooding (e.g., corn, wheat, and soybeans). By storing and conveying floodwaters, hydrologically connected floodplains also can promote the production of these benefits by reducing flood risk for flood-intolerant agriculture. Flood-risk reduction services of hydrologically connected floodplains can occur across a gradient of floodplain management, encompassing extensive "natural" floodplains, strategically located levee setback projects, frequently inundated flood bypasses that encompass a mix of natural and agricultural land use (e.g., the Yolo Bypass), and floodways that are only used during relatively rare, high magnitude floods (e.g., the floodways on the Mississippi River). Through this gradient of floodplain management, floodplains that serve to reduce flood risk can continue to support productive agriculture and provide revenue to landowners, as exemplified by the Yolo Bypass (Sommer *et al.*, 2001). Flood-tolerant crops and current or potential markets for ecosystem services such as groundwater recharge, recreation, and nutrient and carbon sequestration can also support private landowners and agricultural economies (Opperman *et al.*, 2009; Opperman *et al.*, 2010).

Although the previous three sections described floodplain restoration for specific purposes, such as promoting flood-dependent livelihoods, a hallmark of floodplain connectivity is that it tends to produce multiple benefits. For example, while the case study for the Senegal Delta focused on livelihood benefits, the project also produced important benefits for biodiversity. The Yolo Bypass was conceived strictly as a flood-control project, but decades later scientists recorded its immense value for the native species of California's Central Valley, providing some of the most important remaining floodplain habitat in the lower Valley. Even the rarely used floodways of the Mississippi River had high food-web productivity from weeks to months of inundation in 2011, and will likely contribute a strong year class of fish, based on the fish response to the 1993 flood on the Mississippi (Jones and Noltie, 2007).

The feasibility of implementing floodplain reconnection projects, or maintaining existing connectivity, can be improved if these multiple benefits can be identified and captured. The levee setback project at Hamilton City was only financially feasible because it could identify benefits to both flood-risk reduction and ecosystem restoration, and it drew on funding sources for both services. The floodplain reconnection project at Mollicy Farms includes a study of the values provided by ecosystem services such as recreation and carbon and nutrient sequestration to determine if the "stacking" of these multiple services can provide comparable revenue to agriculture for landowners of flood-prone cropland (Opperman *et al.*, 2010).

The challenge of accounting for the full diversity of benefits provided by hydrologically connected floodplains has contributed to floodplains being undervalued within studies for development projects. In the developing world, people use floodplains for a wide variety of purposes depending on the season or type of year (e.g., wet or dry), including fishing, flood-recession agriculture, irrigated agriculture, pasture, timber and nontimber forest products, and wildlife. Often most or all of these uses exist outside of market economies and thus can be quite difficult to measure and enter into cost–benefit analyses to compare alternatives for water management within a development project. Focusing on a single purpose of floodplains, such as comparing flood-recession agriculture to irrigated agriculture, will not account for the full value that floodplains may provide to rural people. Failure to fully account for these values is undoubtedly one of the primary reasons that many water development projects have not produced net positive benefits for rural people (Richter *et al.*, 2010). Going forward, funders and planners of development projects should ensure that the full range of values of hydrologically connected floodplains is accurately measured to facilitate sound analyses and decisions.

These multiple values should also be considered by agencies and organizations tasked with helping society manage changing risks, including risks associated with climate change. Although the hydrological implications of climate change for total precipitation are less certain than forecasts of temperature rise, most forecasts suggest an increase in the frequency of extreme hydrological events, such as floods and droughts (Kundzewicz *et al.*, 2008). Floodplain reconnection or conservation can be considered a form of "ecosystem-based adaption" to climate change and can potentially address both types of extreme events. By storing and conveying floodwaters, floodplains can contribute to flood-risk reduction objectives. By contributing to groundwater recharge and reducing the need for flood-control storage volume within multipurpose

reservoirs, hydrologically connected floodplains can also contribute to increased water security to manage drought risk. Corresponding to ecosystem-based adaption projects generally, floodplain reconnection or conservation offers a broad range of benefits beyond managing risks associated with potential hydrological perturbations from climate change. Because of these multiple benefits, floodplain projects can be considered "no regrets" projects, meaning they still produce values even if the predicted climate-change perturbations do not materialize.

Appendix

List of Courses

1. Water Resources Management
2. Aquatic Ecology (river ecology and floodplain ecology)
3. Floodplain Management/flood-risk Management
4. Sustainable Development

See also: Agriculture, Sustainable. Conservation and the World's Poorest of the Poor. Ecosystem Services. Fish Conservation. Fishes, Biodiversity of. Freshwater Ecosystems. Freshwater Ecosystems, Human Impact on. Impact of Ecological Restoration on Ecosystem Services. Plankton, Status and Role of. Riparian Landscapes. River Ecosystems. Salmon. Valuing Ecosystem Services. Wetland Creation and Restoration. Wetlands Ecosystems

References

Acharya G and Barbier E (2002) Using domestic water analysis to value groundwater recharge in the Hadejia–Jama'are floodplain, Northern Nigeria. *American Journal of Agricultural Economics* 84: 415–426.

Acreman MC (2003) *Environmental Flows: Flood flows*. Washington, DC: The World Bank.

Adams A (1999) Social impacts of an African dam: Equity and distributional issues in the Senegal River Valley. *Report for the World Commission on Dams (WCD)*.

Adams WM (1996) Economics and hydrological management of African floodplains. In: Acreman MC and Hollis GE (eds.) *Water Management and Wetlands in Sub-Saharan Africa*, pp. 21–33. Gland, Switzerland: IUCN.

Agostinho AA, Gomes LC, Verissimo S, and Okada EK (2004) Flood regime, dam regulation and fish in the Upper Parana River: Effects on assemblage attributes, reproduction and recruitment. *Reviews in Fish Biology and Fisheries* 14: 11–19.

Ahearn DS, Viers JH, Mount JF, and Dahlgren RA (2006) Priming the productivity pump: Flood pulse driven trends in suspended algal biomass distribution across a restored floodplain. *Freshwater Biology* 51: 1417–1433.

Alexander J and Marriot SB (1999) Introduction. In: Marriot SB and Alexander J (eds.) *Floodplains: Interdisciplinary Approaches*, pp. 1–13. London: Geological Society of London. Special Publications.

American Society of Civil Engineers (2009) *Report Card for America's Infrastructure*, 168 pp. Washington, DC: American Society of Civil Engineers.

Association of State Floodplain Managers (2008) *Natural and Beneficial Floodplain Functions: Floodplain Management-More than Flood Loss Reduction*. Madison, WI: Association of State Floodplain Managers.

Bakker E (1972) *An Island Called California*. Berkeley, CA: University of California Press.

Barbier EB (2003) Upstream dams and downstream water allocation: The case of the Hadejia–Jama'are floodplain, northern Nigeria. *Water Resources Research* 39: 1311–1319.

Barbour M, Pavlik B, Drysdale F, and Lindstrom S (1991) California vegetation: Diversity and change. *Fremontia* 19: 3–12.

Barry JM (1997) *Rising tide: The Great Mississippi Flood of 1927 and How it Changed America*. New York: Touchstone.

Bayley PB (1989) Aquatic environments in the Amazon Basin, with an analysis of carbon sources, fish production, and yield. In: Dodge DP (ed.) *Proceedings of the International Large River Symposium*, vol. 106, pp. 399–408. Canadian Special Publication of Fisheries and Aquatic Sciences.

Bayley PB (1991) The flood pulse advantage and the restoration of river–floodplain systems. *Regulated Rivers: Research and Management* 6: 75–86.

Bayley PB (1995) Understanding large river–floodplain ecosystems. *BioScience* 45: 153–157.

Beilfuss R (2010) Modelling trade-offs between hydropower generation and environmental flow scenarios: A case study of the lower Zambezi River Basin, Mozambique. *International Journal of River Basin Management* 8: 331–347.

Bello M and Eisler P (2008) Illinois town finds life does go on after floods. Newspaper article June 19, 2008, *USA Today*.

Bousso T (1997) The estuary of the Senegal River: The impact of environmental changes and the Diama dam on resource status and fishery conditions. In: Remane K (ed.) *Fisheries, Aquaculture and the Environment*, pp. 45–65. Oxford, UK: Fishing News Books.

Brauman KA, Daily GC, Duarte TK, and Mooney HA (2007) The nature and value of ecosystem services: An overview highlighting hydrologic services. *Annual Review of Environment and Resources* 32: 67–98.

Brinson MM (1990) Riverine forests. In: Lugo AE, Brinson M, and Brown S (eds.) *Forested Wetlands: Ecosystems of the World*, pp. 87–142. Amsterdam: Elsevier.

Burt TP and Pinay G (2005) Linking hydrology and biogeochemistry in complex landscapes. *Progress in Physical Geography* 29: 297–316.

Chick JH, Cosgriff RJ, and Gittinger LS (2003) Fish as potential dispersal agents for floodplain plants: First evidence in North America. *Canadian Journal of Fisheries and Aquatic Sciences* 60: 1437–1439.

Coates D, Poeu O, Suntornratana U, Tung NT, and Viravong S (2003) *Biodiversity and Fisheries in the Lower Mekong Basin*, 30 pp. Phnom Penh: Mekong River Commission.

Costanza R, d'Arge R, deGroot R, *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature* 387: 253–260.

Criss RE and Shock EL (2001) Flood enhancement through flood control. *Geology* 29: 875–878.

Duvail S and Hamerlynck O (2003) Mitigation of negative ecological and socio-economic impacts of the Diama dam on the Senegal River Delta wetland (Mauritania), using a model based decision support system. *Hydrology and Earth System Sciences* 7: 133–146.

Duvail S and Hamerlynck O (2007) The Rufii River flood: Plague or blessing? *International Journal of Biometeorology* 52: 33–42.

Emerton L (2003) *Tana River, Kenya: Integrating Downstream Values into Hydropower Planning*. Gland, Switzerland: IUCN.

Esselman P and Opperman JJ (2010) Overcoming information limitations for the prescription of an environmental flow regime for a Central American river. *Ecology and Society* 15(1): 6. http://www.ecologyandsociety.org/vol15/iss1/art6/

Florsheim JL and Mount JF (2002) Restoration of floodplain topography by sand-splay complex formation in response to intentional levee breaches, Lower Cosumnes River, California. *Geomorphology* 44: 67–94.

Forster S, Kneis D, Gocht M, and Bronstert A (2005) Flood risk reduction by the use of retention areas at the Elbe River. *International Journal of River Basin Management* 3: 21–29.

Freitag B, Bolton S, Westerlund F, and Clark J (2009) *Floodplain Management: A New Approach for a New Era*. Washington, DC: Island Press.

Galat DL, Fredrickson LH, Humburg DD, *et al.* (1998) Flooding to restore connectivity of regulated, large-river wetlands – Natural and controlled flooding as complementary processes along the lower Missouri River. *Bioscience* 48: 721–733.

Gren IM, Groth KH, and Sylven M (1995) Economic values of Danube floodplains. *Journal of Environmental Management* 45: 333–345.

Grosholz E and Gallo E (2006) The influence of flood cycle and fish predation on invertebrate production on a restored California floodplain. *Hydrobiologia* 568: 91–109.

Hamerlynck O and Duvail S (2003) *The Rehabilitation of the Delta of the Senegal River in Mauritania*. Gland, Switzerland and Cambridge, UK: IUCN.

Hamerlynck O, Duvail S, Ould Messaoud B, and Benmergui S (2005) The restoration of the Lower Delta of the Senegal River, Mauritania (1993–2004). In: Symoens JJ (ed.) *Coastal Ecosystems of West Africa, Biological Diversity – Resources – Conservation. Proceedings at the Conference at the Belgian Academy of Sciences*, pp. 195–2010. Brussels, Belgium: Foundation for the Promotion of Scientific Research in Africa.

Hamerlynck O, Duvail S, Vandepitte L, *et al.* (2011) To connect or not to connect – floods, fisheries and livelihoods in the Lower Rufiji floodplain lakes. *Tanzania. Hydrological Sciences – Journal des Sciences Hydrologiques* 56: 1436–1451.

Hamerlynck O, Nyunja J, Luke Q, Nyingi D, Lebrun D, and Duvail S (2010) The communal forest wetland, rangeland and agricultural landscape mosaics of the Lower Tana, Kenya: socio-ecological entity in peril. In: Bélair C, Ichikawa K, Wong BYL, and Mulongoy KJ (eds.) *Sustainable Use of Biological Diversity in Socio-Ecological Production Landscapes, Background to the Satoyama Initiative for the Benefit of Biodiversity and Human Well-Being*. Convention on Biological Diversity Technical Series no. 52, pp. 54–62.

Harman C and Stewardson M (2005) Optimizing dam release rules to meet environmental flow targets. *River Research and Applications* 21: 113–129.

Hogan ZS, Moyle PB, May B, Vander Zanden MJ, and Baird IG (2004) The imperiled giants of the Mekong. *American Scientist* 92: 228–237.

Interagency Floodplain Management Review Committee (IFMRC) (1994) Sharing the challenge: Floodplain management into the 21st century. *Report of the Interagency Floodplain Management Review Committee to the Administration Floodplain Management Task Force*. Washington DC: Executive Office of the President.

Jeffres CA, Opperman JJ, and Moyle PB (2008) Ephemeral floodplain habitats provide best growth conditions for juvenile Chinook salmon in a California river. *Environmental Biology of Fishes* 83: 449–458.

Jercich SA (1997) California's 1995 Water Bank Program: Purchasing water supply options. *Journal of Water Resources Planning and Management – ASCE* 123: 59–65.

Jones BD and Noltie DB (2007) Flooded flatheads: Evidence of increased growth in Mississippi river *Pylodictis olivaris* (Pisces: Ictaluridae) following the great Midwest flood of 1993. *Hydrobiologia* 592: 183–209.

Junk WJ, Bayley PB, and Sparks RE (1989) The flood pulse concept in river–floodplain systems. In: Dodge DG (ed.) *Proceedings of the International Large River Symposium*, vol. 106, pp. 110–127. Canadian Journal of Fisheries and Aquatic Sciences Special Publication.

Kelley R (1989) *Battling the Inland Sea: Floods, Public Policy, and the Sacramento Valley*. Berkeley, CA: University of California Press.

King AJ, Humphries P, and Lake RS (2003) Fish recruitment on floodplains: The roles of patterns of flooding and life history characteristics. *Canadian Journal of Fisheries and Aquatic Sciences* 60: 773–786.

Klijn F, van Buuren M, and van Rooij SAM (2004) Flood-risk management strategies for an uncertain future: Living with Rhine river floods in the Netherlands? *AMBIO* 33: 141–147.

Knight DW and Shiono K (1996) River channel and floodplain hydraulics. In: Anderson MG, Walling DE, and Bates PD (eds.) *Floodplain Processes*, pp. 139–181. New York: John Wiley & Sons.

Knighton D (1998) *Fluvial Forms and Processes*. London: Arnold.

Kundzewicz ZW, Mata LJ, Arnell NW, *et al.* (2008) The implications of projected climate change for freshwater resources and their management. *Hydrological Sciences Journal – Journal Des Sciences Hydrologiques* 53: 3–10.

Leavenworth S (2004a) Farms in path of levee projects. Goal is better flood protection. *Sacramento Bee*. Sacramento, CA.

Leavenworth S (2004b) Rising risk. *Sacramento Bee*, A1 pp. Sacramento, CA.

Lericollais A and Schmitz J (1984) La calebasse et la houe. Techniques et outils des cultures de décrue dans la vallée du fleuve Sénégal. In: Mollard E and Walter A (eds.) *Le sorgho de décrue dans la vallée du fleuve Sénégal*, pp. 33–38. Paris: IRD.

LeRoy X (2008) Le sorgho de décrue dans la vallée du fleuve Sénégal. In: Mollard E and Walter A (eds.) *Agricultures Singulières*, pp. 33–38. Paris: IRD.

Ligon FK, Dietrich WE, and Trush WJ (1995) Downstream ecological effects of dams. *BioScience* 45: 183–191.

Llewellyn DW, Shaffer GP, Craig NJ, *et al.* (1995) A decision-support system for prioritizing restoration sites on the Mississippi River alluvial plain. *Conservation Biology* 10: 1446–1455.

Loth P (ed.) (2004) *The Return of the Water: Restoring the Waza Logone Floodplain in Cameroon*. Gland Switzerland and Cambridge, UK: IUCN.

Mahoney JM and Rood SB (1998) Streamflow requirements for cottonwood seedling recruitment – An integrative model. *Wetlands* 18: 634–645.

Maingi JK and Marsh SE (2002) Quantifying hydrologic impacts following dam construction along the Tana River, Kenya. *Journal of Arid Environments* 50: 53–79.

Marie J (2009) Le Niger va-t-il devenir les eaux de la discorde? In: Raison JP and Magrin G (eds.) *Des fleuves entre conflits et compromise*, pp. 77–124. Paris: Karthala.

McCarthy E (1997) Perspective on the New Year's floods. *Western Water* 4–13.

Mitsch WJ, Day JW, Gilliam JW, *et al.* (2001) Reducing nitrogen loading to the Gulf of Mexico from the Mississippi River Basin: Strategies to counter a persistent ecological problem. *Bioscience* 51: 373–388.

Mitsch WJ and Gosselink JG (2000) *Wetlands*. New York: John Wiley & Sons.

Mollard E and Walter A (eds.) (2008) *Agricultures singulières*. Paris: IRD.

Moulaye Zeine SA (2004) *Evaluation de l'impact économique du Parc National du Diawling*. Nouakchott, Mauritania: IUCN.

Mount JF (1995) California Rivers and Streams: The Conflict between Fluvial Processes and Land Use. *Berkeley, CA*. University of California Press.

Moyle PB, Baxter RD, Sommer TR, Foin TC, and Matern SA (2004) Biology and population dynamics of Sacramento splittail (*Pogonichthys macrolepidotus*) in the San Francisco Estuary: A review. *San Francisco Estuary and Watershed Science 2*: article 3.

Moyle PB, Crain PK and Whitener K (2007) Patterns in the use of a restored California floodplain by native and alien fishes. *San Francisco Estuary and Watershed Science* 5. http://escholarship.org/uc/item/6fq2f838

Naiman RJ, Decamps H, and McClain ME (2005) *Riparia: Ecology, Conservation, and Management of Streamside Communities*. Amsterdam: Elsevier Academic Press.

Nanson GC and Croke JC (1992) A genetic classification of floodplains. *Geomorphology* 4: 459–486.

National Levee Safety Committee (2009) *Report to Congress on Recommendations for a National Levee Safety Program (Draft)*. Washington, DC: National Levee Safety Committee. http://www.leveesafety.org/docs/NCLS-Recommendation-Report_012009_DRAFT.pdf

National Research Council (NRC) (2004) *Valuing Ecosystem Services: Toward Better Environmental Decision Making*. Washington, DC: The National Academies Press.

Noe GB and Hupp CR (2005) Carbon, nitrogen, and phosphorus accumulation in floodplains of Atlantic Coastal Plain rivers, USA. *Ecological Applications* 15: 1178–1190.

Odum EP (1978) Ecological importance of the riparian zone. In: Johnson RR and McCormick JF (eds.) *Strategies for Protection and Management of Floodplain Wetlands and Other Riparian Ecosystems*. Callaway Gardens, GA: US Department of Agriculture Forest Service GTR-WO-12.

Opperman JJ, Galloway GE, Fargione J, Mount JF, Richter BD, and Secchi S (2009) Sustainable floodplains through large-scale reconnection to rivers. *Science* 326: 1487–1488.

Opperman JJ, Luster RA, McKenney BA, Roberts MD, and Meadows AW (2010) Ecologically functional floodplains: Connectivity, flow regime, and scale. *Journal of the American Water Resources Association* 46: 211–226.

Opperman JJ, Warner A, Girvetz EH, Harrison D, and Fry T (2011) Integrated reservoir–floodplain management as an ecosystem-based adaptation strategy to climate change. *Proceedings to American Water Resources Association 2011 Spring Specialty Conference on Climate Change and Water Resources*. Baltimore, MD: American Water Resources Association.

Park TK (1992) Early trends toward class stratification: Chaos, common property, and flood recession agriculture. *American Anthropologist, New Series* 94: 90–117.

Pielke Jr. RA, Downton MW, and Barnard Miller JZ (2002) *Flood Damage in the United States, 1926–2000: A Reanalysis of National Weather Service Estimates*. Boulder, CO: University Corporation for Atmospheric Research.

Pinter N (2005) One step forward, two steps back on US floodplains. *Science* 308: 207–208.

Pinter N, Ickes BS, Wlosinski JH, and van der Ploeg RR (2006) Trends in flood stages: Contrasting results from the Mississippi and Rhine River systems. *Journal of Hydrology* 331: 554–566.

Poff NL, Allan JD, Bain MB, *et al.* (1997) The natural flow regime. *BioScience* 47: 769–784.

Postel S (2005) *Liquid Assets: The Critical Need to Safeguard Freshwater Ecosystems*. Washington, DC: Worldwatch Institute.

Postel S and Richter B (2003) *Rivers for Life: Managing Water for People and Nature*. Washington, DC: Island Press.

Ricciardi A and Rasmussen JB (1999) Extinction rates of North American freshwater fauna. *Conservation Biology* 13: 1220–1222.

Richter BD, Braun DP, Mendelson MA, and Master LL (1997) Threats to imperiled freshwater fauna. *Conservation Biology* 11: 1081–1093.

Richter BD, Postel S, Scudder T, Lehner B, Churchill A, and Chow M (2010) Lost in development's shadow: The downstream human consequences of dams. *Water Alternatives* 3: 14–42.

Risotto SP and Turner RE (1985) Annual fluctuations in abundance of the commercial fisheries in the Mississippi River and tributaries. *North American Journal of Fisheries Management* 5: 557–574.

Robertson AI, Bacon P, and Heagney G (2001) The response of floodplain primary production to flood frequency and timing. *Journal of Applied Ecology* 38: 126–136.
Rood SB, Gourley CR, Ammon EM, *et al.* (2003) Flows for floodplain forests: A successful riparian restoration. *Bioscience* 53: 647–656.
Ross ST and Baker JA (1983) The response of fishes to periodic spring floods in a southeastern stream. *American Midland Naturalist* 109: 1–14.
Sabo JL, Sponseller R, Dixon M, *et al.* (2005) Riparian zones increase regional species richness by harboring different, not more, species. *Ecology* 86: 56–62.
Salo J, Kalliola R, Hakkinen I, *et al.* (1986) River dynamics and the diversity of Amazon lowland forest. *Nature* 322: 254–258.
Sayers, PB, Galloway GE, Penning-Roswell E, Shen F, Chen Y, and Le Quesne T (2011) *Flood Risk Management: A Strategic Approach*. London: UNESCO on behalf of WWF-UK/China and the General Institute of Water Design and Planning, China.
Schuyt KD (2005) Economic consequences of wetland degradation for local populations in Africa. *Ecological Economics* 53: 177–190.
Shafroth PB, Wilcox AC, Lytle DA, *et al.* (2009) Ecosystem effects of environmental flows: Modelling and experimental floods in a dryland river. *Freshwater Biology* 55: 68–85.
Sheaffer JR, Mullan JD, and Hinch NB (2002) Encouraging the wise use of floodplains with market-based incentives. *Environment* 44: 33–43.
Sommer T, Baxter R, and Herbold B (1997) Resilience of splittail in the Sacramento-San Joaquin estuary. *Transactions of the American Fisheries Society* 126: 961–976.
Sommer T, Harrell B, Nobriga M, *et al.* (2001) California's Yolo Bypass: Evidence that flood control can be compatible with fisheries, wetlands, wildlife, and agriculture. *Fisheries* 26: 6–16.
Sommer TR, Harrell WC, Solger AM, Tom B, and Kimmerer W (2004) Effects of flow variation on channel and floodplain biota and habitats of the Sacramento River, California, USA. *Aquatic Conservation – Marine and Freshwater Ecosystems* 14: 247–261.
Sommer TR, Nobriga ML, Harrell WC, Batham W, and Kimmerer WJ (2001) Floodplain rearing of juvenile Chinook salmon: Evidence of enhanced growth and survival. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 325–333.
Sparks RE, Bayley PB, Kohler SL, and Osborne LL (1990) Disturbance and recovery of large floodplain rivers. *Environmental Management* 14: 699–709.
Sparks RE and Braden JB (2007) Naturalization of developed floodplains: An integrated analysis. *Journal of Contemporary Water Research and Education* 2007: 7–16.
Swenson RO, Whitener K, and Eaton M (2003) Restoring floods on floodplains: Riparian and floodplain restoration at the Cosumnes River Preserve. In: Faber PM (ed.) *California Riparian Systems: Processes and Floodplain Management, Ecology, and Restoration. 2001 Riparian Habitat and Floodplains Conference Proceedings*, pp. 224–229. Sacramento, CA: Riparian Habitat Joint Venture.
Terer T, Ndirtu GG, and Gichuki NN (2004) Socio-economic values and traditional strategies of managing wetland resources in Lower Tana River, Kenya. *Hydrobiologia* 527: 3–15.
Tharme RE (2003) A global perspective on environmental flow assessment: Emerging trends in the development and application of environmental flow methodologies for rivers. *River Research and Applications* 19: 397–441.
The Bay Institute (1998) *From the Sierra to the Sea: The Ecological History of the San Francisco Bay–Delta Watershed*. San Francisco: The Bay Institute.
Tobin GA (1995) The levee love affair: A stormy relationship. *Water Resources Bulletin* 31: 359–367.
Tockner K and Stanford JA (2002) Riverine floodplains: Present state and future trends. *Environmental Conservation* 29: 308–330.
Trush WJ, McBain SM, and Leopold LB (2000) Attributes of an alluvial river and their relation to water policy and management. *Proceedings of the National Academy of Sciences* 97: 11858–11863.
Tu M (2000) *Vegetation patterns and processes of natural regeneration in periodically flooded riparian forests in the Central Valley of California*. Dissertation, Department of Plant Biology, University of California.
US Army Corps of Engineers (2001) *Non-Structural Flood Damage Reduction within the Corps of Engineers*. U.S. Army Corps of Engineers. http://www.nwo.usace.army.mil/nfpc/non_struc_flood_damage_reduction.pdf
US Army Corps of Engineers (2004) *The Mississippi River and Tributaries Project*.
US Army Corps of Engineers (2008) *The Mississippi River & Tributaries Project: Designing the Project Flood*.
UNEP (2010) *Blue Harvest: Inland Fisheries as an Ecosystem Service*. Penang, Malaysia: WorldFish Center.
Valett HM, Baker MA, Morrice JA, *et al.* (2005) Biogeochemical and metabolic responses to the flood pulse in a semiarid floodplain. *Ecology* 86: 220–234.
Verhoef H (1996) Health aspects of Sahelian floodplain development. In: Acreman MC and Hollis GE (eds.) *Water Management and Wetlands in Sub-Saharan Africa*, pp. 35–50. Gland, Switzerland: IUCN.
Ward JV (1989) The four-dimensional nature of lotic ecosystems. *Journal of the North American Benthological Society* 8: 2–8.
Ward JV, Tockner K, Arscott DB, and Claret C (2002) Riverine landscape diversity. *Freshwater Biology* 47: 517–539.
Welcomme RL (1979) *Fisheries Ecology of Floodplain Rivers*. London: Longman Group Ltd.
Wiens JA (2002) Riverine landscapes: Taking landscape ecology into the water. *Freshwater Biology* 47: 501–515.
Williams PB, Andrews E, Opperman JJ, Bozkurt S, and Moyle PB (2009) Quantifying activated floodplains on a lowland regulated river: Its application to floodplain restoration in the Sacramento Valley. *San Francisco Estuary and Watershed Science* 7.
Winemiller KO (2004) Floodplain river food webs: Generalizations and implications for management. In: Welcomme RL and Petr T (eds.) *Proceedings of the Second International Symposium on the Management of Large Rivers for Fisheries*, pp. 285–309. Phnom Penh, Cambodia: Food and Agricultural Organization (FAO) of the United Nations.
Wohl EE (2000) Geomorphic effects of floods. In: Wohl EE (ed.) *Inland Flood Hazards: Human, Riparian, and Aquatic Communities*, pp. 167–193. Cambridge, UK: Cambridge University Press.
World Commission on Dams (2000) *Dams and Development: A New Framework for Decision-Making*. London: Earthscan Publications.
Yarie J, Viereck L, Van Cleve K, and Adams P (1998) Flooding and ecosystem dynamics along the Tanana river. *BioScience* 48: 690–695.

Thermophiles, Origin of

Anna-Louise Reysenbach and Margaret L Rising, Portland State University, Portland, OR, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 647–654, © 2001, Elsevier Inc.

Glossary

Archaea One of three domains of life; from the Greek *archaios* (ancient, primitive); prokaryotic cells; membrane lipids predominantly isoprenoid glycerol diethers or diglycerol tetraethers; formerly called archaebacteria.

Bacteria One of three domains of life; from the Greek *bacterion* (staff, rod); prokaryotic cells; membrane lipids predominantly diacyl glycerol diesters; formerly called eubacteria.

Biomarker A macromolecule unique to a particular organisms or group of organisms such that its detection alone would suggest the presence of the organism or group of organisms.

Chemolithotroph Organism deriving its energy from the oxidation of inorganic compounds.

Eukarya One of three domains of life; from the Greek *eu-* (good, true) and *karion* (nut; refers to the nucleus); eukaryotic cells; cell membrane lipids predominantly glycerol fatty acyl diesters.

Heterotroph Organism deriving its energy from the oxidation of organic compounds.

Hyperthermophiles From the Greek *hyper-* (over), *therme-* (heat), and *philos* (loving); includes organisms that grow best at temperatures warmer than 80 °C.

Photoautotroph Organism that uses light as its source of energy (photosynthesis) and inorganic carbon (CO_2) as its sole carbon source.

Planetesimals Small, solid bodies similar to meteors in composition but revolving in orbit around a central gaseous nucleus, as do planets around the sun.

Small subunit ribosomal RNA RNA (about 1500 bases in prokaryotes) that functions as part of the ribosome and the sequence permits the inference of evolutionary relationships among organisms.

Thermophiles From the Greek *therme-* (heat) and *philos* (loving); includes organisms that grow best at temperatures between 50 ° and 80 °C.

Introduction

At the time when liquid water–a prerequisite for life as we know it–appears in the geological record (3.8 billion years ago), Earth was a hot, anoxic environment and under constant bombardment by meteors, many of which could have virtually vaporized the oceans. Early Earth, therefore, would have been an attractive home to heat-loving thermophiles and their extreme cousins, the hyperthermophiles, where thermophily would have offered a great selective advantage. The geochemical and thermal characteristics of deep-sea hydrothermal vents and terrestrial hot springs are thought to approximate conditions on Earth at the earliest possible time that it could have supported life, providing modern analogs for testing the early evolution of life hypotheses. In addition, the relatively recent use of the evolutionarily conserved molecules such as the small subunit ribosomal RNA (ss rRNA) sequences in phylogenetic analyses of all life has placed the thermophiles closest to the root of the universal tree of life (**Figure 1**). On this tree, hyperthermophiles decidedly monopolize the lowest and shortest branches, suggesting that they may be related to the earliest microbes to inhabit Earth. However, the proposal that thermophiles may have originated early in Earth's history is much debated. Therefore, this article also considers alternative scenarios for the origins of thermophiles.

Thermophiles and Hyperthermophiles Defined

All life can be placed in one of the three domains of life: the Archaea, Bacteria, or Eukarya. Archaea and Bacteria are prokaryotes, and the Eukarya are eukaryotes. However, phylogenetically, the Archaea are as different from the Bacteria as the Bacteria are from the Eukarya. Thermophiles are found in all domains as multicellular and unicellular organisms, such as fungi, algae, cyanobacteria, and protozoa, and they grow best at temperatures higher than 45 °C. In contrast, the extreme thermophiles, or hyperthermophiles, grow best at temperatures higher than 80 °C and are almost exclusively restricted to the Archaea, with only two hyperthermophilic orders in the Bacteria, namely, the Thermotogales and Aquificales. Commonly, thermophiles and hyperthermophiles are found associated with deep-sea and terrestrial hydrothermal vents. Thermophiles have also been obtained from deep (up to 3500 m), hot, subterranean areas, Jurassic oil-bearing sandstone and limestone formations, and suitable manmade environments such as smoldering coal refuse piles, coal-containing uranium mines, and boiling wastewaters from geothermal power plants.

Ecological Niche and Metabolic Diversity of Hyperthermophiles

Deep-sea hydrothermal and terrestrial hydrothermal vents form primarily as the result of plate tectonics, either as a result

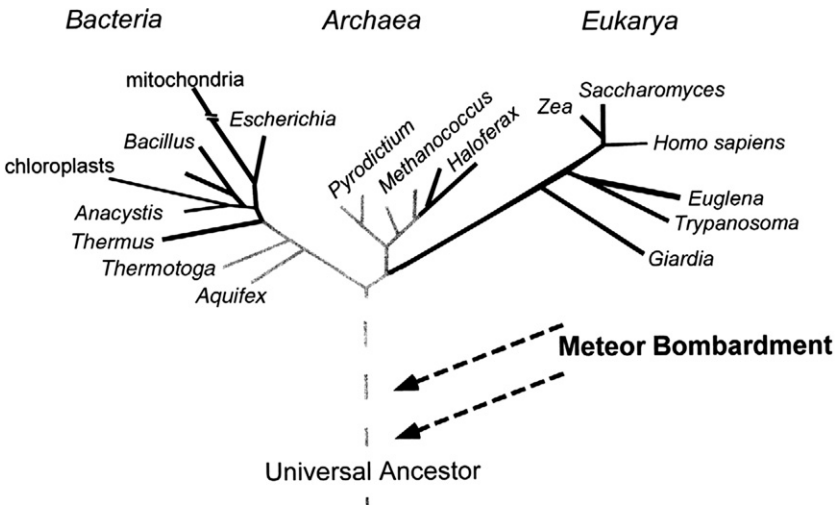


Figure 1 Rooted phylogenetic tree based on the small subunit rRNA molecule. The tree is not drawn to scale.

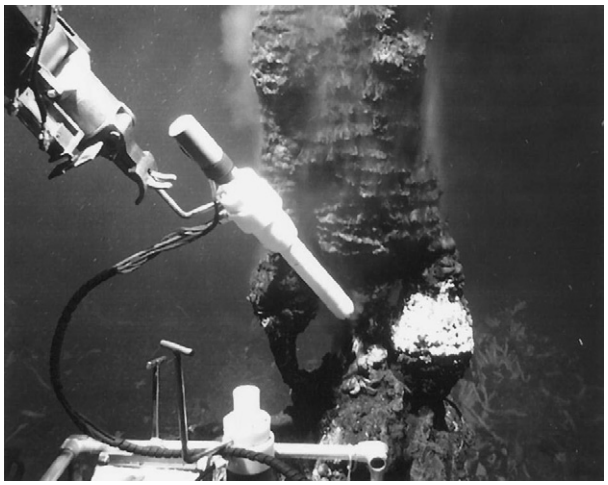


Figure 2 Sampling a deep-sea hydrothermal vent chimney from the eastern Pacific Ocean. The porous sulfide chimney structures can reach several meters high and are ideal habitats for hyperthermophiles.



Figure 3 Moose Pool, Yellowstone National Park, Wyoming. This acidic thermal spring contains elemental sulfur that accumulates on the gas bubbles (see inset). The archeon *Sulfolobus acidocaldarius* was isolated from this spring in the late 1960s by Thomas Brock.

of seafloor spreading or as a tectonic plate moves across a hot spot. In both scenarios, fissures and faulting occur in the earth's crust, permitting water (seawater, groundwater, or rainwater) to percolate downward. As the fluid moves through the earth's crust and approaches the magma chamber, the fluid heats and reacts with the surrounding rocks, adding some minerals and gases (e.g., iron, manganese, carbon dioxide, hydrogen, and hydrogen sulfide) and removing others (e.g., sulfate and magnesium in seawater). The fluid is finally forced back to the surface as a highly altered hot fluid, its chemistry representing a history of its travels through the earth's crust. At deep-sea hydrothermal vents, this fluid can reach temperatures of 400 °C, remaining in a liquid state due to the hydrostatic pressures at these depths (>2000 m). It rapidly mixes with cold, oxygenated water, and the minerals rapidly precipitate, giving the fluid a smoky appearance (**Figure 2**). These vents are aptly named "black smokers." As the minerals precipitate they create porous sulfide–mineral structures called "chimneys." At terrestrial hot springs, the fluid may be ejected forcefully as a geyser or simply bubble into a thermal spring or mudpot (**Figure 3**).

The high temperatures and unusual geochemical characteristics of deep-sea and terrestrial hydrothermal vents create ecological niches that are exploited by hyperthermophilic Bacteria and Archaea. The geochemical milieu provides different energy sources (electron donors) and carbon dioxide for a chemolithotrophic existence. The porous chimney structures provide temperature gradients along which thermophiles can situate themselves, and thermal springs create temperature gradients as the fluid moves away from the source. Comparable to photoautotrophs that harvest light energy to fix inorganic carbon, these chemolithotrophs harvest inorganic chemical energy to fix inorganic carbon. It has been calculated that the mixing of superheated thermal fluids at deep-sea vents with cold, oxygenated seawater causes a significant geochemical disequilibrium, significant enough to lower the

Gibbs free energy and help drive biologically mediated redox reactions (Shock, 1996). The food chain does not stop there, though. Many thermophiles are heterotrophs, consuming organic carbon produced from biological activity. Additionally, at terrestrial vents, due to the available light energy, the springs are often colonized by a rich diversity of thermophilic phototrophs, including the cyanobacteria, which are oxygenic phototrophs that use water as the reductant and evolving oxygen in the process. Anoxygenic phototrophs are also found at terrestrial vents, primarily using the readily available hydrogen sulfide as the reductant in photosynthesis, releasing elemental sulfur in the process. Both of these types of phototrophs are part of a complex community with heterotrophs, and together they form thick green, purple, and orange microbial mats typical of many terrestrial thermal areas throughout the world.

The majority of hyperthermophiles are anaerobes, with a few exceptions that are microaerophilic, requiring low oxygen concentrations for growth. In contrast, aerobiosis is more common among the thermophiles because oxygen dissolves more readily in the lower temperatures of the cooler thermal springs. The thermophiles that have been isolated from deep-sea vents generally grow at pH values near neutrality. However, thermophiles from terrestrial hot springs have been grown at pH values from 1 to 10.

Although the highest temperature for life has not been determined, it has been suggested that 150 °C may represent that threshold. In contrast, the lower temperature limit for most hyperthermophiles is 60 °C, a temperature fatally hot for most other organisms. Even 90 °C is too low for one of the more extreme hyperthermophiles, *Pyrolobus fumarii*, which has a maximum temperature for growth of 113 °C. However, for these thermophiles, the inability to grow does not result in death; they are capable of entering a dormant state when exposed to cold conditions and will thrive again when returned to favorable temperatures. Thermophiles are therefore very well suited to the deep-sea hydrothermal vent environment, in which the hot fluids are continually being mixed with cold, oxygenated seawater.

Molecular Mechanisms for Thermophily

On a molecular level, thermophily is accomplished in many ways, contributing together to overall stability and success of life at high temperatures. In general, the amino acid composition of hyperthermophilic enzymes is surprisingly similar to that of homologous mesophilic enzymes. Because of the amino acid sequence similarity, it has been proposed that heat resistance is explained, instead, by the manner in which the polypeptide chains are folded on themselves, forming their tertiary conformational stability. It is a polypeptide's three-dimensional or tertiary shape in critical locations on the polymer that permits an enzyme to function. Proteins do not spontaneously fold into their active form after synthesis. Instead, many require the assistance of molecular chaperones for proper folding. A specialized form of molecular chaperone, called a thermosome, is of particular importance to hyperthermophiles because it assists proteins in refolding correctly after denaturation due to heat exposure. Also called heat shock proteins, thermosomes effectively increase heat tolerance of organisms so they can function at higher temperatures. For instance, at 108 °C, approximately 80% of the protein of *Pyrodictium occultum* consists of a heat-induced thermosome, allowing it to survive 1 h of autoclaving at 121 °C (Stetter, 1998).

Additional thermal resistance of DNA is conferred by the presence of DNA topoisomerases and histones and, to a lesser extent, by the base composition of the genome. All hyperthermophiles known to date have been shown to possess reverse gyrase, a unique type I DNA topoisomerase that causes stabilizing, positive supertwists in the DNA helix. Additionally, archaeal (but not bacterial) hyperthermophiles possess histones which significantly increase the temperature at which DNA denatures.

Because an increased guanine + cytosine (G + C) nucleic acid content increases the melting temperature of DNA, one would predict that hyperthermophiles have a higher G + C content in their genomes. In many cases, the opposite is true. However, analysis of the genome reveals that certain genes essential to survival, such as the ss rRNA gene, have higher G + C contents.

The presence of ethers instead of esters in cellular membranes also contributes to the stability of hyperthermophiles at high temperatures. Cell membranes of Bacteria and Archaea are generally distinguished by the presence of glycerol-linked ether lipid moieties in Archaea and ester lipid moieties in Bacteria. An exception to this rule is the presence of a glycerol ether lipid in *Thermotoga maritima*, one of the few bacterial hyperthermophiles. The presence of the ether lipids probably increases stability of cellular membranes against hydrolysis at high temperatures. Membranes of the archaeal hyperthermophiles contain lipids derived from diethers or tetraethers and are highly resistant against hydrolysis at high temperatures.

Origin of Thermophiles: Fossil Record and Universal Phylogenetic Tree

Were thermophiles the original ancestors to all life? Much of the debate about origins of life and early evolution of life on Earth centers on this question. Is thermophily an acquired feature derived from adaptation? As more information accumulates, the resolution of these questions becomes less clear. However, many different scenarios are emerging based on fossil record analyses, RNA and protein phylogenies, and our understanding of the conditions that prevailed on early Earth.

Early Earth and Its Hyperthermophilic Niche

Perhaps the most compelling evidence supporting the proposal that hyperthermophiles are living fossils of the first life that arose on Earth is that modern ecological niches of hyperthermophiles fit our current view of the conditions on primitive Earth, which was anoxic, hot, and volcanically active. During its first half billion years of existence (4.6–4.0 billion years ago), the temperature of the surface of Earth probably exceeded 100 °C. From recent studies of ancient sedimentary rocks of the Isua formation in Greenland, traces of biologically

produced carbon were detected, suggesting that the invention of life had already taken place 3800 billion years ago. Thus, sometime prior to 3800 billion years ago, it is likely that the first living organisms appeared at a time when Earth, although cooling, was much hotter than it is now, and all the essential geochemical energy sources were present for a chemolithotrophic and thermophilic existence. Heat-tolerant organisms such as hyperthermophiles would have been at a distinct advantage over heat-sensitive organisms.

In addition to the generally hot and anoxic conditions on early Earth, the bombardment of early Earth by meteors may have caused a bottleneck event that favored survival of hyperthermophilic organisms. An examination of the impact craters on Earth's geologically inactive moon strongly suggests that the earth was heavily bombarded by meteors from its formation 4.6 billion years ago until about 3800 billion years ago. A large meteor impact, of which there were many, would have caused intense local heat and sent sufficient debris into the atmosphere to cause global cooling and a reduction in light intensity, detrimentally impacting photo-synthesizing organisms. There is also evidence of extremely large meteor impacts that would have generated enough energy to heat the earth's surface to 2000 K and virtually vaporize the world's oceans. In either scenario, hyperthermophiles would have been uniquely adapted to survive. They could have tolerated lower light levels because many are chemolithotrophs, making a living independent of photosynthesis. Any life adapted to live near ocean-floor volcanic centers would have had the greatest chance of surviving one of these ocean-evaporating events.

The Fossil Record

Additional evidence for early ancestors of microbial life can be found in the rock record. Rocks older than 3500 billion years are highly metamorphosed and deformed, precluding the preservation of morphological fossils. However, the 3800-billion-year-old Isua rocks from Greenland, mentioned previously, offer indirect evidence that life existed earlier than 3500 billion years ago. These sedimentary rock formations are associated with liquid water, a prerequisite for life as we know it. The rocks contain geochemical evidence of past biotic activity preserved within the minerals. Grains of apatite (calcium phosphate) contain a significant portion of carbonaceous inclusions that are isotopically "light," suggesting biological activity. Microorganisms preferentially incorporate the lighter carbon-12 isotope over the carbon-13 isotope when inorganic carbon is fixed into organic carbon using the enzyme ribulose bisphosphate carboxylase/oxygenase (Rubisco). This isotopic evidence suggests these putative life-forms were carbon-fixing chemolithotrophs or phototrophs and not heterotrophs.

Three hundred million years later, the first microfossils appear in rocks (about 3500 billion years old) from Western Australia and South Africa. The fossils are simple rod-shaped and filamentous bacteria, indistinguishable morphologically from any similar-shaped bacteria today. Many are reminiscent of photosynthetic bacteria, such as modern cyanobacteria. However, there is some doubt that much oxygen evolved from photosynthesis into early Earth's atmosphere. Isotopic signatures associated with these fossils do suggest a lighter carbon preference, similar to that observed in the 3800-billion-year-old rocks. Were these first fossils perhaps thermophiles? Unless we are able to identify specific thermophilic biomarkers in these fossils, we may never know the answer.

The Molecular Fossils

Although they are not a substitute for the existence of microfossils, evolutionarily conserved macromolecules within living organisms function as "molecular fossils," permitting an inference of relatedness to other organisms and of evolutionary distance from a hypothetical common ancestor. Carl Woese's pioneering work with the ss rRNA molecule (16S rRNA in prokaryotes) led to the generation of the universal phylogenetic tree (**Figure 1**), in which life falls within three domains: the Archaea, Bacteria, and Eukarya. The suitability of ss rRNA as the macromolecule for evolutionary comparisons is based on several critical reasons. The ss rRNA exists in every living organism, its function in every living organism is constant, it is an ancient molecule (assuming protein synthesis was necessary in the earliest of life), it has undergone only moderate changes in nucleic acid sequence when compared between diverse biological domains, and the size of the molecule is large enough to contain considerable information but small enough to be manageable.

Molecular Phylogenies

The relatedness of organisms to each other as suggested by the universal phylogenetic tree is markedly different than the previously proposed phylogenies. Most notably, the three domains of biological life (the Bacteria, Archaea, and Eukarya) replace the five kingdoms (animals, plants, fungi, protista, and monera). The Eukarya domain includes the multicellular animals, plants, fungi, and protista. Distributed throughout the Bacteria and Archaea are the unicellular prokaryotes. Using alternative markers, such as 23S rRNA, RNA polymerase, elongation factor Tu, F_1F_0 ATPase β-subunit, RecA protein, and HSP60 heat shock protein, reveals general agreement with the major lineages in the ss rRNA tree. However, some discrepancies are evident.

One can root the ss rRNA tree by comparing it with a paralogous marker that arose from gene duplications (such as EF-Tu or ATPase) prior to the diversification of the three primary domains. This rooted tree reveals that all hyperthermophiles, whether bacterial or archaeal, comprise the deepest and earliest lineages within the tree (**Figure 1**). In other words, they may be the closest living relatives to a common universal ancestor. However, mutational rates of ss rRNA and their associated base sequence changes are not consistent enough through time nor within lineages to assign a specific date to branching events, and so one cannot assign a clock to this tree. Nonetheless, the shorter branches leading to the hyperthermophiles indicate a slower evolutionary rate, whereas the longer branches leading to their mesophilic relatives suggest a faster evolutionary rate. It is possible that the hyperthermophiles have short lineages because they are so highly adapted to their ecological niche that they have no need for further adaptation. It is also possible that certain evolutionary

constraints are imposed on thermophiles due to the extreme selectivity of the high-temperature environment. Nevertheless, the placement of thermophiles at the base of the universal phylogenetic tree strongly suggests that they are most closely related to the common ancestor of all life. In this sense, they are not unlike other "living fossils," or organisms whose morphology has changed very little based on comparisons of modern and ancient fossilized specimens, such as horseshoe crabs, club mosses, *Welwitschia*, and *Gingko bilboa*.

Although hyperthermophiles may be most closely related to the universal ancestor of all life, their metabolic machinery is anything but primitive. Viewed independently of the rooted ss rRNA universal phylogenetic tree, their heat-tolerant adaptations appear sophisticated or even highly evolved. If an analysis of another universal and highly conserved macromolecule placed hyperthermophiles away from the root of its phylogenetic tree, it would severely shake the topology of the phylogenetic tree discussed previously. A phylogenetic analysis of the conserved DNA-directed RNA polymerase has done just that (Klenk *et al.*, 1999). Although overall the RNA polymerase-based phylogeny corresponds very well with the ss rRNA-based phylogeny, the positions of the hyperthermophilic bacteria *A. pyrophilus* and *Thermotoga maritima* are in the middle of the RNA polymerase tree instead of in the most deeply rooted branch in the Bacteria as suggested by ss rRNA phylogeny. Other examples also exist in which the ss rRNA tree does not quite hold true. Additional analyses of universal and highly conserved macromolecules such as DNA-directed RNA polymerases or whole genomes may one day provide a comprehensive universal phylogenetic tree that may or may not place hyperthermophiles near its root.

The debate continues, and evidence against an early origin of thermophily accumulates. If mesophily preceded thermophily, then versions of heat-tolerant mechanisms that characterize hyperthermophiles would have evolved first in mesophiles and been exploited by hyperthermophiles to open up new ecological niches. Structural analysis of reverse gyrase and *Taq* polymerase support such a proposal, suggesting they evolved from mesophilic ancestors (Forterre, 1996). Additionally, the lipid moieties in hyperthermophilic Bacteria are not homologs of the lipid moieties in hyperthermophilic Archaea. Instead, they are analogs with opposite stereochemistry, suggesting the ether-based lipid moiety feature evolved independently in the Bacteria and Archaea instead of from a common ancestor (Forterre, 1996).

Lateral Gene Transfer and the Genetic Annealing Model

Lateral (horizontal) gene transfer and variable rates of evolution and mutation may have played a significant role in the evolution of life and may offer an explanation for the confusing results of the analyses of non-ss rRNA molecules (Woese, 1998; Doolittle, 1999). Lateral gene transfer refers to the exchange of genetic material from one organism to another. It is contrasted against vertical gene transfer, in which genetic information is passed vertically from parent to offspring. Bacteria and Archaea, which as a general rule have simpler cell designs, exhibit considerable horizontal gene transfer. In contrast, the Eukarya, which contain highly evolved cell designs, generally do not engage in horizontal gene transfer. Woese proposed in his "genetic annealing model" that the universal ancestor was not a discrete entity but rather a diverse community of primitive cells that evolved as a unit, engaging in horizontal gene transfer on a scale even greater than that occurring in Bacteria and Archaea today and developing into three different communities, which in turn gave rise to the three primary lines of descent as defined by the ss rRNA tree (Woese, 1998).

The ss rRNA universal tree, then, is not a conventional organismal phylogenetic tree but rather a history of the evolution of central components of the ribosome, with the deeply rooted branches of the universal tree representing a "gene tree" and not an "organismal tree." In other words, by the time the three primary lines of descent emerged, and the tree started to take form, self-replicating organisms had not yet taken form. Instead, "life," with its associated exchange of genetic information, existed in communal entities by way of lateral gene transfer. Additional cell complexity and function needed to evolve before there was life as we envision it today. Nonetheless, these communal entities could have evolved in the conditions on early Earth.

Theories of the Origin of Life

Theories of the origin of life are highly debated. Did life originate here on Earth, or did it develop on another planet, such as Mars, which shared similar planetary conditions 4.0 billion years ago? Did life originate in environments analogous to deep-sea hydrothermal vents, or did the first biological molecules form as a result of the reaction of electrical discharges within a prebiotic soup of chemicals? Closely tied to much of this debate are models pertaining to early Earth's atmosphere and the possible energy carbon sources (organic and inorganic). Given that this article deals with the origins of thermophiles, we focus our discussion on a possible high-temperature origin of life, highlighting some arguments that do not support this hypothesis.

In light of Earth's hot, reduced, and anoxic origins, it has been proposed that life may have arisen in environments very similar to present-day deep-sea hydrothermal vents. Here, life could take refuge and survive the planetesimal bombardment of early Earth in an environment rich in redox energy and inorganic carbon. The first biological entities could evolve rapidly as anaerobic, hyperthermophilic chemolithotrophs. The rapid mixing of superheated hydrothermal fluid, rich in reduced minerals, with cold oxygenated sulfate-rich seawater at deep-sea hydrothermal vents creates thermodynamic disequilibrium conditions that favor the production of organic molecules (Shock, 1996). The abundance of charged minerals associated with deep-sea vents led to the proposal by Gunter Wächtershäuser that the original source of reducing power for carbon fixation (and therefore a chemolithotrophic origin of life) may have come from exergonic "pyrite-pulled reactions"—the oxidative formation of pyrite (FeS_2) from ferrous sulfide (FeS) and hydrogen sulfide (H_2S or SH^-). Charged surfaces such as pyrite would attract and bind any negatively charged molecule in solution, such as carbonate, phosphate, and sulfide. These molecules would be maintained in sufficient proximity for subsequent metabolic interactions by "surface bonding," or anionic bonding to the positively

charged pyrite surface, resulting in the formation of the first biomolecules on a charged surface. Many are skeptical of these high-temperature scenarios for the origin of life, offering evidence that many of the biomolecules of life, such as RNA, are not stable at high temperature and therefore would not support the proposal that RNA arose prior to DNA in an early RNA world (Miller and Bada, 1988). Furthermore, others have generated models to show that the ancestral rRNA would have a moderate G + C content, contrary to all thermophilic rRNAs which are generally rich in G + C contents (Galtier *et al.*, 1999).

The "panspermia" hypothesis offers another scenario for the origin of life. It holds that life originated elsewhere in the galaxy and that microorganisms were propelled through space to Earth or, alternatively, that exogenous organic carbon arriving on planetesimals fueled a heterotrophic origin of life on Earth. Chyba and Sagan (1992) estimate that approximately 4 billion years ago about 100,000,000 kg/year of organic carbon was delivered by interplanetary dust particles, and about 10,000,000,000 kg/year of organic carbon was produced by postimpact plumes caused by meteor, asteroid, and comet bombardment. From what is known about the hostile environment on early Earth and the timing of the appearance of the first microfossils, life became considerably complex within a relatively short period of time. Therefore, life either evolved very rapidly after its inception or, consistent with the panspermia hypothesis, it was raining down on Earth from elsewhere in the galaxy. Because the panspermia hypothesis involves high temperatures as particles enter the earth's galaxy, it embraces, at the very least, a thermotolerant origin of life.

Conclusion

Morphological features of early microfossils generally offer little information about their relatedness to extant microbes. Consequently, microbiologists have come to rely on the molecular record to define phylogenetic relationships and infer the history of microorganisms. The absence of a good microbial fossil record prevents the assignment of a time line on the molecular record. The ss rRNA universal phylogenetic tree places hyperthermophiles in the deepest and shortest branches of the tree, implying that they are the closest living relative to the common universal ancestor of life. However, as we sequence more microbial genomes and create detailed phylogenies from other molecules, the situation becomes increasingly confusing, with lateral gene transfer perhaps "muddying" the phylogenetic record. However, high-temperature conditions that prevailed on early Earth and its subsequent cooling favor the likelihood that at some time during Earth's early history thermophiles took advantage of the geochemical energy supplied by the hydrothermal fluid and evolved into the highly adapted chemolithotrophic and heterotrophic life that is found at high temperatures.

See also: Archaea, Origin of. Eukaryotes, Origin of. High-Temperature Ecosystems. Origin of Life, Theories of. Psychrophiles. Vents

References

Chyba C and Sagan C (1992) Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: An inventory for the origins of life. *Nature* 355: 125.

Corliss J (1990) Hot springs and the origin of life. *Nature* 347: 624.

Doolittle W (1999) Phylogenetic classification and the universal tree. *Science* 284: 2124.

Forterre P (1996) A hot topic: The origin of hyperthermophiles. *Cell* 85: 789.

Galtier N, Tourasse N, and Gouy M (1999) A nonhyperthermophilic common ancestor to extant life forms. *Science* 283: 220.

Jakosky B (1998) *The Search for Life on Other Planets*. New York: Cambridge Univ. Press.

Klenk H, Meier T, Durovic P, *et al.* (1999) RNA polymerase of *Aquifex pyrophilus:* Implications for the evolution of the bacterial *rpoBC* operon and extremely thermophilic bacteria. *J. Mol. Evol.* 48: 528.

Madigan M, Martinko J, and Parker J (1997) *Brock Biology of Microorganisms*, 8th edn. New York: Prentice Hall.

Miller S and Bada J (1988) Submarine hot springs and the origin of life. *Nature* 334: 609.

Mojzsis S, Arrhenius G, McKeegan K, Harrison T, Nutman A, and Friend C (1996) Evidence for life on earth before 3800 million years ago. *Nature* 384: 55.

Shock E (1996) Hydrothermal systems as environments for the emergence of life. In: Bock G and Goode J (eds.) *Evolution of Hydrothermal Ecosystems on Earth (and Mars?)*. Chichester, UK: Wiley.

Stetter K (1998) Hyperthermophiles: Isolation, classification, and properties. In: Horikoshi K and Grant W (eds.) *Extremophiles: Microbial Life in Extreme Environments*. New York: Wiley-Liss.

Wächtershäuser G (1988a) Pyrite formation, the first energy source for life: A hypothesis. *Syst. Appl. Microbiol.* 10: 207.

Wächtershäuser G (1988b) Before enzymes and templates: Theory of surface metabolism. *Microbiol. Rev.* 52: 452.

Woese C (1998) The universal ancestor. *Proc. Natl. Acad. Sci. USA* 95: 6854.

The Value of Biodiversity

Partha Dasgupta, University of Cambridge, Cambridge and University of Manchester, Manchester, UK
Ann P Kinzig and Charles Perrings, Arizona State University, Tempe, AZ, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Partha Dasgupta, volume 2, pp. 291–304, © 2001, Elsevier Inc.

Glossary

Accounting prices Accounting or shadow prices are measures of the social opportunity cost of resource use. The accounting price of a resource is the increase in social well-being that would be enjoyed if a unit more of the resource was made available at no cost.

Ecological functions The combinations of ecosystem component and ecosystem processes that enable ecosystems to support primary production and to maintain populations of species.

Economic valuation The methods used to identify society's willingness to pay for the nonmarket benefits obtained from ecosystems.

Ecosystem processes Biogeochemical cycling, including primary production (photosynthesis), nutrient and water cycling, and materials decomposition, and the flow, storage, and transformation of materials and energy through an ecosystem, including through food web processes such as pollination, predation, and parasitism.

Ecosystem services The benefits (i.e., goods and services) people obtain from ecosystems.

Production functions The combinations of capital stocks, material inputs, and technology to produce valued outputs.

Regulating services The benefits obtained from the regulation of ecosystem processes, including flood protection, climate regulation, human disease regulation, water purification, air quality maintenance, pollination, and pest control.

Resilience The capacity of an ecosystem to maintain functionality when subject to stress or shock.

Social discount rates The rate at which society is willing to substitute future consumption for present consumption.

Biological Resources and Biodiversity

Traditionally economists have approached the use people make of biodiversity through the individual species those people exploit: the fish that are caught in freshwater and marine capture fisheries; the trees that are harvested from either managed or natural forests; the livestock raised for meat, hides, and other animal products in pastoral systems; the food and cash crops produced in arable systems; or the medicinal plants extracted from forests, grasslands, and wetlands. To do this, they have relied on models of the growth of individual populations in which the functioning of the rest of the ecosystem is taken to be exogenous. Well known illustrations of this viewpoint include the use of the logistic function to chart the time path of the population of a single species of fish, the study of predator–prey interactions by means of variants of the Lotka–Volterra equations, the estimation of growth in biomass of a species of trees at a given site, and so on. A prominent concern has been to determine the rates at which a single resource would be harvested in different institutional settings. Thus, not only have socially optimum harvest rates been analyzed (Clark, 1976), but economists have also determined harvest rates when harvesters have free or open access to the resource (e.g., Gordon, 1954; Dasgupta, 1982).

This approach has given an understanding of the effect on harvest rates of harvesting costs, the rates at which harvesters discount future costs and benefits, the productivity of the resource *in situ*, the "worth" of the harvest to harvesters, and the property-rights regime in which the harvesting is done. It has also given an understanding of the implications of changes in either the abundance or the intrinsic growth rates of exploited species. The approach has offered much less insight, however, into the economic effects of changes in the various measures of species richness (alpha, beta, or gamma diversity) or in the diversity within and between functional groups of species.

An alternative approach has focused on the ecosystem that is exploited when a particular species is harvested. In this approach, the resource switches from the individual species to the biotic and abiotic components of the system that regulate its growth. Instead of treating the role of the ecosystem as constant, as it is in the carrying capacity of a Lotka–Volterra population growth model, the approach allows investigation of the interactions between ecosystem components induced by human exploitation. There the focus is on system feedbacks through, for example, the energy flows involved in trophic interactions and the distribution and flows of biochemical substances in soils and water bodies, and of gases and particulates in the atmosphere. Where the state variables in the traditional approach are population "quantities" (e.g., herd sizes, expressed in numbers, or biomass units), in the ecosystem approach they may also include quality indices for air, soil, or water. The motivation is to understand the way in which the exploitation of ecosystems alters their usefulness to humankind by changing the biotic and abiotic processes that underlie various ecosystem functions, and hence the services they yield.

Economic studies of global warming, eutrophication of lakes, the management of rangelands, purification of water in watersheds, and the pollution of estuaries are examples of such endeavor (Mäler, 1974; Ludwig *et al.*, 2003; Perrings and Walker, 2005; Chichilnisky and Heal, 1998). Such studies have

provided valuable insights into the effects on ecosystems of the character of economic activities, as driven by technology, costs and revenues, discount rates, and the property-rights regime that governs access to the ecosystem. In this approach the focus of study is the ecosystem as a renewable natural resource system: a mix of functional groups of species that support a range of ecosystem processes, and hence a range of ecosystem services. Many ecosystems are deliberately "simplified" (through the removal of pests, pathogens, predators, or competitors) in order to increase their value for particular purposes. The best examples of this are to be found in agriculture, aquaculture, forestry, and urban systems. The biodiversity in such systems is managed to enhance production of particular services. But even the most simplified ecosystems – monocultures in agriculture and forestry and cities – are still ecosystems. The services they provide depend on the composition of species they contain, the energy and nutrient flows within the system, and the interactions between species.

This article asks how the second of these approaches informs the understanding of the value that biodiversity has to individual farmers, foresters, and fishers; how that differs from the value it has to wider society; and why the difference matters. All ecosystems generate an array of services that affect peoples' well-being in different ways. All ecosystems are also subject to property rights or access rules that affect the rights and responsibilities of users, and that determine which services they take into account and which they do not. Biodiversity change is a problem for society if it has either significant social impacts that are ignored by users or unacceptable distributional consequences. This article considers both the inefficiency that results when biodiversity use involves external costs and the inequity that results when biodiversity change bears disproportionately on the poor.

Biodiversity and Ecosystem Services

The Millennium Ecosystem Assessment (MA) defined ecosystem services to include the full array of benefits people obtain from ecosystems distinguishing four broad benefit streams: the provisioning, cultural, regulating, and supporting services (Millennium Ecosystem Assessment, 2005). The MA provisioning services describe the processes that yield foods, fibers, fuels, water, biochemicals, medicines, pharmaceuticals, and genetic resources. Many of these products are subject to well-defined property rights and are priced in the market. Others are not. The MA cultural services comprise a set of largely non-consumptive uses of the environment including both the well-established market benefits of recreation and tourism and the spiritual, religious, esthetic, and inspirational well-being that people derive from the "natural" world, along with the value to science of the opportunity to study that world. The MA regulating services limit the effect of stresses and shocks to the system; include erosion control or soil stabilization, water purification, and waste treatment; and the regulation of air quality, climate, hydrological flows, pests, disease, and natural hazards. Finally, the MA supporting services comprise the ecosystem processes that underpin production of all these services, including soil formation, photosynthesis, primary production, nutrient, carbon, and water cycling.

The value of biodiversity derives from the value of the final goods and services it produces. To estimate this value, one needs to understand the "production functions" that link biodiversity, ecosystem functions, ecosystem services, and the goods and services that enter into final demand. Recent ecological research on the linkages between biodiversity and ecological functioning has improved the understanding of ecological processes involved in the production of a number of ecosystem services (Loreau *et al.*, 2002; Naeem *et al.*, 2009). It has also improved the understanding of the role played by the diversity of species in the production of these services. Species are related through functional traits that make them more or less perfect substitutes in executing particular ecological functions. Individual species near perfect functional substitutes for other species if they share a full set of traits with those other species. Conversely, they are "singular" if they possess a unique set of traits (Naeem, 1998). Species are also related through ecological interactions – trophic relationships, competition, parasitism, facilitation and so on – that make them more or less complementary in executing ecological functions (Thebault and Loreau, 2006).

The value of individual species depends both on their substitutability/complementarity in the functioning of ecological systems and on the way in that functioning has been affected by land use. The simplification of agroecosystems to privilege particular crops or livestock strains necessarily affects the array of services that system delivers, partly because the number of functions performed reduces with the number of species (Hector and Bagchi, 2007) and partly because each species in a system generally performs multiple functions. Ecosystems are systems of "joint production." Individual systems generate multiple services. It follows that part of the cost of simplification is the ecosystem services foregone as a result. So, for example, modern agriculture has focused on increased yields per hectare, but has compromised a range of other ecosystem services including water supply, water quality, habitat provision, pollination, and soil erosion control (Millennium Ecosystem Assessment, 2005). It has also affected the capacity of the modified system to function over a range of environmental conditions – that is, its stability or resilience.

There is a vast and confusing literature on the relationship between diversity (or complexity) and stability or resilience (see, e.g., May, 1972; Pimm, 1984; Schulze and Mooney, 1993; Tilman *et al.*, 1996; Kinzig *et al.*, 2001), with many studies appearing to contradict each other. This is in part due to definitional ambiguities, in part due to analysis at different levels of organization, and in part due to differing results in theoretical versus experimental versus observational approaches. Most studies suggest that diversity (in the form of a greater number of species) enhances stability for community-level properties, though these results do not necessarily extend to the population level, where diversity can enhance, erode, or have little impact on stability (Tilman *et al.*, 1996; Mccann, 2000; Muradian, 2001). Most of these studies, though, address stability and resilience in a particular sense – how diversity affects the ability of a system to return to a previous (possibly equilibrium) state (Gunderson, 2000). They say less about resilience in the alternative sense (sensu Holling) – the contributions that diversity might play in diminishing the probability that a system might access entirely new system

configurations, or "flipping" into a new basin of attraction. (A system could be highly resilient in the second sense but not in the first – highly dynamic within a given basin of attraction, never returning to any given state for very long, but never leaving that basin of attraction, either.)

The relationship between diversity and resilience in this second sense is less clear. The vulnerability of low-diversity systems, such as agricultural monocultures, and the persistence of high-diversity systems, such as tropical forests, have often been cited as evidence of a relationship between diversity and resilience. Goodman (1975) takes these examples to task. In addition, there are equally compelling examples of low diversity systems that appear to be highly resilient – rice paddies have persisted in the same places for hundreds of years (Geertz, 1963; Lansing, 1991), and desertified and biologically depauperate regions are highly resistant to restoration efforts. (This latter example illustrates that resilience is not always a good thing.) These contradictions indicate the hazards of associating system-level properties with system-level outcomes. Degradation of soils, for instance, can cause both low diversity and high resilience, without there being a direct causal link between the two. Similar observations have been made about the role of productivity in high-diversity systems – are high-diversity systems productive, or are highly productive systems diverse?

Ecological theory suggests that increasing species number or complexity within a community should increase the number of equilibrium points or stable states (May, 1977; Deangelis and Goldstein, 1978). All else being equal, that would seem to imply more thresholds and lower resilience. Of course, all else would not be equal – it is entirely possible that increased diversity and complexity increase both the number of alternative states that can be accessed and the tenacity with which the system remains in any particular state. But both the theoretical and the empirical evidence for a strong relationship between diversity and resilience (sensu Holling), when correctly sorted, are still ambiguous.

Diversity can, of course, play a role in enhancing the functioning of ecosystems. More functional guilds will mean provisioning of a greater number of ecosystem services. Those services may or may not be resilient. What will enhance resilience (continued delivery) of a particular service is the diversity of species within the functional guild responsible for that service (e.g., within the guild of nitrogen fixers). If each species within a guild has a differential response to the environment, this would allow sustained function under environmental variability. This is known as functional redundancy.

Why should one necessarily expect species within a functional guild to have differential responses to the environment? Ecological theory suggests, in fact, that this is exactly what one should expect. In spite of perceptions of "nature red in tooth and claw," evolution by natural selection, in fact, serves to diversify the niches of similar species and reduce competition among them (hence the diversification of beak sizes and shapes among the finches on the Galapagos Islands, each suited to different seed-gathering strategies). In most ecological communities it was also observed that just a few species make up the bulk of the biomass, with many more occurring in low abundance. The hypothesis is that both the breadth of services provided and their resilience are facilitated by these patterns. In particular, one might expect the dominant species to be functionally dissimilar (providing different services, but reducing direct competition through functional dissimilarity), with the lower-abundance species being functional redundants of the dominant species (but with different responses to environmental variability or stressors). Walker *et al.*, 1999 demonstrated just these outcomes in a grassland ecosystem subjected to grazing pressure . In the absence of grazing, dominant species were functionally dissimilar, whereas minor species were functionally similar to a dominant species. When dominant species declined or were eliminated by grazing, their functionally similar minor counterparts increased in abundance, thus maintaining certain functions within the disturbed grassland.

Although much remains to be done to understand the joint production of ecosystem services in particular ecosystems, the interactions between services, and the impact of changes in the relative abundance of species within and between functional groups, this article can now identify at least some of the consequences that biodiversity change for ancillary ecosystem services. The canonical bioeconomic models developed by Clark to understand the exploitation of marine mammals and fisheries (Clark and Munro, 1979) clarified the conditions required for the optimal extraction of particular populations, establishing the capital theoretical basis for exploiting biological stocks. The extension of such work to cover the exploitation of multiple species has made it possible to understand the effect of species interactions on co-produced ecosystem services (Tilman *et al.*, 2005; Eichner and Pethig, 2005; Brock and Xepapadeas, 2002; Perrings and Walker, 1997, 2005). This, in turn, has made it possible to track the physical effect of biodiversity change that lies outside the market, and to begin to estimate its social cost. This article first considers the generic issues involved in the diagnosis and cure of market failures, and then identifies their implications for the economics of biodiversity change.

Market Failure

A central concern among economists concerned with the environment has been to devise ways by which it would be possible to ascertain the "value" of natural resources and the services they provide. This is because for many natural resources and the services they support, markets simply do not exist. The interactions between biodiversity, ecological functioning, ecosystem services, and the production of the things that people care about most directly are not signaled by market prices. In some cases, markets do not exist because the costs of negotiation and monitoring are too high. One class of examples is provided by economic activities affected by ecological interactions involving long geographical distances (e.g., the effects of upland deforestation on downstream activities hundreds of miles away); another, by large temporal distances (e.g., the effect of carbon emission on climate in the distant future, in a world where forward markets are nonexistent because future generations are not present today to negotiate with the present generation). Then there are cases (e.g., the atmosphere and the open seas) where the nature of the resource makes private property rights impractical and so

keeps markets from existing; whereas in others, ill-specified or unprotected property rights prevent their existence, or make markets function wrongly even when they do exist. In short, market failures mean that people are not confronted by the real cost of their behavior (but *see* Market Failure). This gives rise to the phenomenon of externalities (i.e., exchanges among people that take place without their consent). (The early literature on ecological economics identified market failure as the underlying cause of environmental problems. (Pigou, 1920; Lindahl, 1958; Arrow, 1971; Meade, 1973; Mäler, 1974; Baumol and Oates, 1975; Dasgupta and Heal, 1979)

Problems arising from an absence of forward markets for "transactions" between the present generation and those to appear in the distant future are no doubt ameliorated by the fact that humans care about their children's well-being and know that they, in turn, will care for theirs, in an intergenerational sequence. This means, by recursion, that even if people do not care directly about the well-being of their distant descendants, they do care about them indirectly. However, there is a distinct possibility that one's implicit concern for the distant future via such recursion is inadequate, due, say, to institutional failure in other spheres of economic activity. This is why economists have argued that market rates of interest do not reflect socially desirable discount rates (Lind, 1982; Arrow *et al.*, 1996; Portney and Weyant, 1999). In short, market failure involves misallocation of resources not only in the present but also across time.

In each of these cases, the market prices of goods and services fail to reflect their social worth; typically, they are less than their social worth. In economics, the social worth of goods and services is called accounting prices (sometimes, shadow prices). The accounting price of a resource is the increase in social well-being that would be enjoyed if a unit more of the resource were made available at no cost. So a resource's accounting price is the difference between its market price and the tax (or subsidy) that ought to be imposed on it. Needless to say, accounting prices reflect social objectives, ecological and technological constraints, and the extent to which resources are available.

It should be noted that externalities do not create market distortions; they are a form of market distortion. The presence of externalities leads to a wedge between market prices and accounting prices. Generally speaking, laissez-faire economies are not much good at producing publicly observable signals of environmental scarcities. To illustrate, if there were free access to a resource base, the market price of the resource, *in situ*, would be zero. However, if the resource were in limited supply, its accounting price would be positive. So, there is a directional bias in environmental externalities: market failure typically results in an excessive use of the natural resource base, not an insufficient one.

One way to improve matters would be to impose regulations on resource users, for example, restrictions on effluent discharges, quota on fish harvests, and bans on logging. Another would be to introduce a system of taxes, often called Pigovian taxes (in honor of Pigou, who first discussed the difference between private and social costs in the context of environmental pollution (Pigou, 1920)). Pollution charges, charges on the amount of fish harvested, and stumpage fees are examples. The idea underlying Pigovian taxes is to bring market prices (inclusive of taxes) in line with accounting prices. Each of the two schemes, quotas and taxes, has its advantages and disadvantages, some of the differences between the two becoming salient once it is recognized not only that ecological processes are stochastic, but also resource users and government agencies do not have the same information about local ecology, say, for example, the cost of waste disposal (Weitzman, 1974; Dasgupta, 1982). This article cannot discuss this in details, but three points are worth noting. First, the two schemes are distributionally not equivalent: under a quota system resource rents are captured by harvesters and polluters, whereas under a tax system they are collected by the tax authority. Second, the imposition of Pigovian taxes provides greater incentives to resource users to explore resource-saving technological improvements. This is because if the users are taxed, they pay more for the resource than they would have if they had been issued quotas instead. Third, environmental taxes, when properly designed, remove market distortions. In addition, there is a presumption that tax revenues, thus collected, would enable the government to reduce distortionary taxes (e.g., taxes on earned income). There is, thus, a presumption that Pigovian taxes yield a "double dividend" (Bovenberg and Goulder, 1996; Goulder, 1995) (but see Bohm, 1996), a rhetorical phrase that has been much used to persuade governments to impose "green" taxes. Matters of public finance have been a recurrent theme in ecological economics since the 1970s (see, especially, Carraro and Siniscalco, 1996; Baumol and Oates, 1975; Cropper and Oates, 1992). A hybrid policy instrument, which involves the government issuing a fixed number of transferable licenses, combines some of the features of quotas and Pigovian taxes. For example, the scheme resembles quotas, in that, resource rents are not captured by government, and it resembles Pigovian taxes, in that at the margin license, holders pay the accounting price of the resource for its use. (See Tietenberg (1990) for a good discussion of transferable licenses, both in theory and in practice.)

Institutional Failures and Poverty: Global Versus Local Environmental Problems

Aside from market failure, another long-recognized source of difficulty is what may be called government failure. For example, Binswanger argued that, in Brazil, the exemption from taxation of virtually all agricultural income (allied to the fact that logging was regarded as proof of land occupancy) provided strong incentives to the rich to acquire forest land and to then deforest it (Binswanger, 1991). He argued that the subsidy the government thereby provided to the private sector was so large, that a reduction in deforestation (via a removal of subsidies) was in Brazil's interests, not merely in the interests of the rest of the world. That is, he concluded that Brazil had much to gain from reducing then current rates of deforestation. (In a wider discussion of the conversion of forests into ranches in the Amazon basin Schneider (1995) has demonstrated that the construction of roads through the forests has also been a potent force.) Since that time the government of Brazil has reached a similar conclusion. It has changed the incentives it offers to landowners and has entered into

negotiations with other countries that are potential beneficiaries of reduced deforestation in Brazil. At the heart of this decision is the recognition that there are interactions between ecosystem services generated by tropical forests. Land clearance for agriculture has consequences for biodiversity, water supplies, microclimatic regulation, and macroclimatic regulation (via carbon sequestration). Indeed, Brazil is now one of the main players in the development of the Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD) scheme. The scheme aims to create a financial value for the carbon stored in forests, offering incentives for developing countries to reduce emissions from forested lands (Malhi *et al.*, 2008; Phelps *et al.*, 2010; Arriagada and Perrings, 2011).

This said, it is important to note that the causes of environmental problems are not limited to market and government failure; problems also arise because such microinstitutions as the household can function badly. In poor communities, for example, men typically have the bulk of the political voice. One should then expect public investment in, say, resource regeneration to be guided by male preferences, not female needs. On matters of afforestation in the drylands, for instance, one should expect women to favor planting for fuelwood and men for fruit trees, because it is the women and children who collect fuelwood, whereas men control cash income (and fruit can be sold in the market). This explains why, even as the sources of fuelwood continue to recede, fruit trees are often planted (Dasgupta, 1993b).

That political instability (at the extreme, civil war) is a direct cause of resource degradation is obvious. What is not obvious is that it is a hidden cause as well. Political instability creates uncertainty in property rights. In its presence, people are reluctant to make the investments that are necessary for environmental protection and improvement: the expected returns on such forms of investment are low. In a study comprising 120 countries, Deacon has offered statistical evidence of a positive link between political instability and forest depletion (Deacon, 1994).

Taken together, these examples reflect the environmental consequences of institutional failure. They have a wide reach, and in recent years, they have often been discussed within the context of the thesis that environmental degradation, such as eroding soil, receding forests, and vanishing water supplies, is a cause of accentuated poverty among the rural poor in poor countries. There is truth in this. But there is also accumulated evidence that poverty itself can be a cause of environmental degradation (Dasgupta, 1993a, 2001). This reverse causality occurs because some natural resources (e.g., ponds and rivers) are essential for survival in normal times, whereas others (e.g., forest products) are also a source of supplementary income in times of acute economic stress. Under changing circumstances (e.g., economic development in urban centers), social norms that previously had maintained long-term economic relationships among members of a community tend to break down. Some (e.g., the able bodied and mobile) gain, whereas others (e.g., women, the old, and the very young) lose and become poorer. In extreme cases the breakdown of social norms also means that local resources that earlier were subject to communitarian regulations become "open access," with all the attendant consequences.

These links between rural poverty and the state of the local natural resource base in poor countries offer a possible pathway along which poverty, resource degradation, and even high fertility feed on one another in a synergistic manner over time (Dasgupta, 1993a, 1995, 2000). Experience in sub-Saharan Africa and Pakistan is not inconsistent with this (Cleaver and Schreiber, 1994; Filmer and Pritchett, 1996). Indeed, an erosion of the local natural-resource base can make certain categories of people destitute even while the economy's gross national product (GNP) increases. The thought that entire populations can always be relied upon to make the shift from resource-based, subsistence existence to a high-income, industrial one is belied both by evidence and by theory.

These two causes of resource degradation, namely, institutional failure and poverty, pull in different directions and are together not unrelated to an intellectual tension between the concerns people share about global warming and acid rain, which sweep across regions, nations and continents, and those matters (such as, for e.g., the decline in firewood or water sources) which are specific to the needs and concerns of the poor in as small a group as a village community. Environmental problems present themselves differently to different people. In part, it is a reflection of the tension this article has just noted and is a source of misunderstanding of people's attitudes. Some people, for example, identify environmental problems with poverty and unprecedented population growth in the South, whereas others identify them with wealth and unprecedented expenditure patterns in the North. Even though debates between the two groups often become shrill, each vision is in part correct. There is no single environmental problem and, so, no single valuation problem; rather, there is a large collection of them. Thus, growth in industrial waste and resource use has been allied to increased economic activity; in the former Socialist bloc neither preventive nor curative measures have kept pace with the production of waste. Moreover, the scale of the human enterprise, both by virtue of unprecedented increases in the size of the world's population and by the extent of economic activity, has so stretched the capabilities of ecosystems that humankind can today rightly be characterized as Earth's dominant species (Vitousek *et al.*, 1997). These observations loom large not only in ecological economics, but also in the more general writings of environmentalists and in the professional writings of ecologists in the West.

However, economic growth itself has brought with it improvements in the quality of a number of natural resources. The large-scale availability of potable water, and the increased protection of human populations against both water- and airborne diseases in industrial countries, have in great measure come in the wake of growth in national income these countries have enjoyed over the past 200 years or so. Moreover, the physical environment inside the home has improved beyond measure with economic growth. For example, cooking in South Asia continues to be a central route to respiratory illnesses among women. Such positive links between economic growth and environmental quality often go unnoted by environmentalists in the North. This lacuna may be yet another reflection of the fact that it is all too easy to overlook the enormous heterogeneity of Earth's natural resource base, ranging as it does from the atmosphere, oceans, and

landscapes to water holes, grazing fields, and sources of fuelwood. Both this heterogeneity and the diversity of the human condition across the globe constantly need to be kept in mind in discussions of the value of biodiversity in different locations.

Valuing Resources and Evaluating Projects

Since institutional failures abound in human dealings with Earth, the commercial profitability of economic activities, say, of investment projects (projects for short), is frequently not an adequate measure of their social worth. So recourse should be taken to social cost–benefit analysis, the purpose of which is to estimate the impact of projects on human well-being, now and in the future. Notice that, if undertaken, a project would be a perturbation to the economy. So, for example, a project consisting of the construction of a dam would be a perturbation to an economy without the dam. The economic forecast *sans* the project can be thought of as the *status quo*.

Analyzing the consequences of a project would involve estimating the need for labor, intermediate products, raw materials, and output, as well as predicting the ecological effects of the project. These consequences need to be specified for each future period (*see* Discontinuous Value Functions for a formalization). Since there is never sufficient knowledge to make precise estimates of the consequences, project evaluators should quantify estimates of the uncertainties, preferably in terms of probabilities. This means that, in general, project designers ought to model the integrated ecological and economic system. Unhappily, in practice this is infrequently done.

To arrive at a good estimate of a project's social benefits and costs, one should in principle value each and every commodity involved in it. The procedure devised by economists is to select some readily measurable bundle of goods ordinarily consumed and to define the "value" of any other commodity as the amount of the bundle society would be willing to give up for it. This is a workable way for estimating the commodity's accounting price. The net social benefit of a project in any given period of its life is obtained by multiplying the project's inputs and outputs in that period by their corresponding accounting prices and adding them (outputs of "goods" are taken to be positive; output of "bads" and inputs are taken to be negative). Using a suitable discount rate (often called the social discount rate), the net social benefits yielded by a project in each period are added. Projects that yield a positive present discounted value of net social benefits are recommended for acceptance; those that yield a negative present discounted value of net social benefits are rejected. The theory of social cost-benefit analysis has been developed by economists over the past 50 years, and is now, to all intents and purposes, complete (Dasgupta *et al.*, 1972; Little and Mirrlees, 1974; Dasgupta and Mäler, 2000). Daily *et al.* (1999) provides a nontechnical account of the role of social cost–benefit analysis in environmental management.

A prior exercise (i.e., prior to conducting social cost–benefit analysis) is to estimate accounting prices. A great deal of work in ecological economics has been directed at discovering methods for estimating accounting prices of natural resources. It is as well to remember that the kinds of resources this article mentions are on occasion of direct use in consumption (as with fisheries), on occasion indirectly, as inputs in production (as with plankton, which serves as food for fish), and sometimes in both (as with drinking and irrigation water). The value may be utilitarian (e.g., as a source of food, or as a keystone species), it may be esthetic (e.g., a shoreline), or it may be intrinsic; indeed, it may be all these things.

Economists have devised various methods for estimating accounting prices. As would be expected, the prices of some natural resources are easier to estimate than those of others. There are now standard techniques for determining the accounting prices of irrigation water, fisheries, timber, and agricultural soil (Anderson, 1987; Repetto *et al.*, 1989; Solorzano *et al.*, 1991; Vincent and Ali, 1997). They involve estimating the resource's use-value. For example, the value of a piece of agricultural land would be the present discounted value of the flow of net profits it is expected to generate from cultivation, minus the environmental damage caused by the pesticides and herbicides to be used. Such an approach can be also used for estimating losses associated with water logging and overgrazing. Reductions in air- or water-borne pollution can be valued in terms of improvements in health (e.g., reductions in the number of days people would be expected to be ill; see, e.g., (World Bank, 1992)). Other techniques have been devised for valuing "amenities," such as places of scenic beauty. One popular method involves asking people hypothetical questions concerning their willingness to pay for preserving the amenity (this is called the "contingent-valuation method," or CVM, for short); another involves estimating from sample surveys the distribution of costs visitors from different locations have incurred to view the site (this is called the "travel-cost method"). A third involves inferring how much people are willing to pay for enjoying the amenity (e.g., clean air) from the commercial value of land at sites that offer the amenity (this is called the "hedonic price" of land). (See Mitchell and Carson, (1989) and Freeman (2003).) Good descriptions of these methods and their application to valuing ecosystem services can be found in Freeman (2003).

Two things that complicate the valuation of biodiversity change are uncertainty in the future use-values of these resources and irreversibility when they are lost. The twin presence of uncertainty and irreversibility implies that preservation of the stock has a value in addition to its current use-value, namely, the value of extending society's set of future options. (For a review of the challenges posed by these two issues, see Perrings and Brock (2009).) Future options have an additional worth because with the passage of time, more information is expected to be forthcoming about the resource's use-value. This additional worth is often called an option value (Arrow and Fisher, 1974; Henry, 1974; Fisher and Hanemann, 1986). The accounting price of a resource is, at the very least, the sum of its direct use-value and its option value.

These techniques enable one to estimate the use-value of a given resource. As it happens, the resource's accounting price may well exceed this. Why? The reason is that there may be additional values "embodied" in a species, an assemblage of species, or a landscape. This is sometimes captured in the idea that living resources have an "intrinsic" worth as living resources. It would be absurd to suppose that the value of a blue

whale is embodied entirely in its flesh and oil or that the value of game in Kenyan safari parks is simply the present-discounted value of the flow of tourists' willingness-to-pay to view them. The idea of intrinsic worth of living things is inherent not only in traditional religious systems of ethics but also in modern ethical theories. Economists tend to think of this as the nonuse value of resources. The question is not so much whether living creatures have nonuse value (intrinsic worth), but rather of ways of assessing this worth. Since it is very difficult to get a quantitative handle on intrinsic worth, the correct thing to do may often be to take note of it, keep an eye on it, and call attention to it in public debate if the stock is threatened with destruction.

It can be concluded that the social worth of natural resources can be decomposed into three parts: their use value, their option value, and their nonuse value. The components appear in different proportions, depending on the resource. For example, oil and natural gas would not be thought to have very low nonuse value relative to their use value. In contrast, endangered primates would be thought to have very limited use value relative to their nonuse value. And so on.

For both use and nonuse values, wherever species are identified in the functions that describe ecosystem processes and functions, one can also identify their marginal impact on the output of things people care about. The marginal value of an incremental change in the abundance of any species derives from the value of the services it yields. Derivation of that value requires specification of the functions that connect species to directly valued goods or services (Barbier, 2007) or that connect ecosystems and the services they produce (Barbier, 2008). Whether one uses market prices or revealed or stated preference methods to obtain an estimate of willingness to pay for the directly valued goods or services is more or less irrelevant. What is important is that some form of "production function" approach is needed to estimate the value of a marginal change in the biodiversity that supports the directly valued good or service. A good example of the combination of methods is Allen and Loomis, (2006). They combine willingness-to-pay estimates obtained using stated preference methods for the conservation of directly valued higher-trophic-level species with ecological data on trophic relationships to derive estimates of implicit willingness to pay for the conservation of species lower down the food chain.

It is as well to emphasize that the purpose of estimating environmental accounting prices is not to value the entire environment; rather, it is to evaluate the benefits and costs associated with changes made to the environment due to human activities. Prices, whether actual or accounting, have significance only when there are potential exchanges from which choices have to be made (e.g., when one has to choose among alternative investment projects). Thus, the statement that a particular act of investment can be expected to degrade the environment by, say, 1 million dollars annually has meaning, because it says, among other things, that if the investment were not to be undertaken, humanity would enjoy an additional 1 million dollars of annual benefits in the form of environmental services. The statement also has operational significance: the estimate could (and should) be used for calculating the rate of return attributable to the investment in question.

Biodiversity: Necessity or Luxury?

The starting point is that the value of biodiversity derives from its role in the production of things that people care about. This extends from clean water, through the production of foods, fuels, fibers, and pharmaceuticals, all the way to the spiritual, cultural, or scientific value of intact ecosystems or totemic species. It follows that the value people implicitly place on biodiversity change depends on the value of the ecosystem services they want. For those who are most directly threatened by pests or diseases, the most important consideration may be pest or pathogen elimination. For those engaged in the production of foods, fuels, or fibers, the most important consideration may be the elimination of crop predators (e.g., insect pests), competitors (e.g., weed species), and diseases (e.g., blights and rusts). For those engaged in conservation or ecotourism the most important consideration may be the preservation of native species and the elimination of exotic species. For those engaged in managing the impacts of trade or travel, the most important consideration might be preventing movement of potentially invasive plants, insects, or diseases. Since the services that people care about are likely to differ depending on geographical location, culture, income, and so on, the value they attach to changes in biodiversity is also likely to differ. And since many services are increased by focusing on a relatively small number of species, people may be expected to put a positive value on reducing species richness as often as they put a positive value on enhancing it.

This helps putting into perspective the common view that the rich value biodiversity more than the poor: that biodiversity conservation is a "luxury" good. The natural resource base is far from being a luxury. As the source of a multitude of ecosystem services necessary to the lives and livelihoods of the majority of people on the planet, the natural-resource base is a necessity. The way that the poorest of people depend on the resource base may differ from the way that the rich do so, but both do depend on it. (The way that the poor depend on the resource base is explored in greater detail in Dasgupta (2001, 1993a) and Dasgupta and Mäler (1997).

This perspective, of viewing natural resources as luxuries, first found expression in World Bank (1992), where it was suggested that there is an empirical relationship between gross domestic product (GDP) per head and concentrations of industrial pollutants. Based on the historical experience of OECD countries, it was argued in the document that when GDP per head is low, concentrations of atmospheric pollutants (e.g., sulfur dioxide (SO_2)) increase as GDP per head increases, but when GDP per head is high, concentrations decrease as GDP per head increases further. In short, it was found that the functional relationship between GDP per head and concentrations of industrial pollutants has an inverted-U shape (**Figure 1**). Among economists this relationship has been christened the "environmental Kuznets curve." (It is a misnomer. The original Kuznets curve, which was an inverted U, related income inequality to real national income per head on the basis of historical cross-country evidence.)

Panayotou had reported the inverted-U shape in cross-country data on GDP per head and deforestation and emissions of SO_2, nitrogen oxides (NO_x), and particulate matters (Panayotou, 1992). The logic underlying the environmental

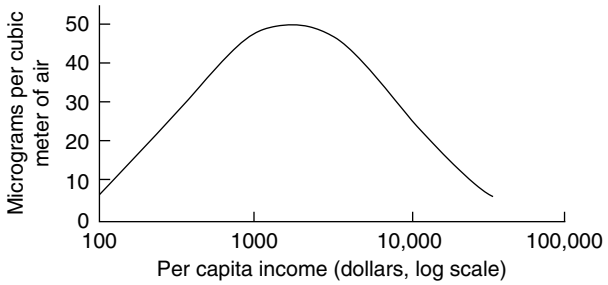


Figure 1 Urban concentrations of sulfur dioxide. Reprinted from figure 4 in *World Development Report* (1992), with permission from World Bank.

Kuznets curve is that resource degradation is reversible: degrade all you want now, you can always recover the stock later, because Earth can be relied upon to rejuvenate it. It is known, however, that this presumption does not apply to biodiversity loss. The presence of ecological thresholds implies that damage to ecosystems can be irreversible. As an overarching metaphor for "trade-offs" between manufactured wealth and resource degradation, the environmental Kuznets curve has to be rejected. (For more extensive discussions of the environmental Kuznets curve, see Arrow *et al.* (1995), and the responses it elicited in symposia built round the article in *Ecological Economics*, 1995, 15(1); *Ecological Applications*, 1996, 6(1); and *Environment and Development Economics*, 1996, 1(1); and see the special issue of *Environment and Development*, 1997, 2(4).)

The relation between income and biodiversity loss has been approached in three ways in this literature. In one approach, deforestation has been used as a proxy for biodiversity loss (using the species area relationship to explain the link between changes in forest area and biodiversity loss). The evidence for any well-defined relation between income and biodiversity loss using this metric is extremely weak (Dietz and Adger, 2003; Mills and Waite, 2009). A second approach uses the National Biodiversity Risk Assessment Index (NABRAI) developed by Reyers *et al.* (1998). This too has failed to find evidence for a statistically significant relation between biodiversity loss and income (Mozumder *et al.*, 2006). A third approach has focused on the direct measures of threat contained in the IUCN's Red List, and finds a statistically significant relation between the natural log of *per capita* income and the number of threatened species that is linear in the case of plants; "U" shaped in the case of amphibians, reptiles, fishes and invertebrates, and inverted "U" shaped in the case of birds (Naidoo and Adamowicz, 2001). More recently, Perrings and Halkos have reconsidered this latter approach using the most recent IUCN data. Controlling for climate, population density and protected areas, they find a strongly quadratic relation between *per capita* Gross National Income and the number of threatened species. They conclude that the interest people have in the abundance of other species is indeed sensitive to income. In poor countries where agriculture is expanding most rapidly at the extensive margin (by bringing new land into cultivation), people have a direct interest in habitat conversion, and in the elimination of weeds, pests and predators. They may also have an interest in conserving their biological heritage, and many poor countries have committed a significant share of their total land area as protected areas, but their willingness to pay for conservation is strictly limited by their income – their ability to pay (Perrings and Halkos, 2010).

Substitution Possibilities

This section and the following one are based on Dasgupta *et al.* (2000a, b) and Dasgupta and Maler (2004). In fact the belief that constraints arising from resource depletion can be overcome as countries become wealthier in terms of their manufactured- and human-capital assets is frequently based on a subtler thought than the one that underlies the environmental Kuznets curve. The belief is based on possibilities of substitution.

Resource constraints facing an economy can be eased by four types of substitution. First, there can be substitution of one thing for another in consumption (nylon and rayon cloth substituting for cotton and wool, pulses substituting for meat, and so forth). Second, manufactured capital can substitute for labor and natural resources in production (the wheel and double-glazing are two extreme examples). Third, novel production techniques can substitute for old ones. For example, the discovery of effective ways to replace the piston by the steam turbine (i.e., converting from reciprocating to rotary motion) was introduced into power plants and ships a little more than a hundred years ago. The innovation was an enormous energy saver in engines. Fourth, and for this article most importantly, natural resources themselves can substitute for one another. This involves the thought that, as each resource (e.g., each species) is depleted, there are close substitutes lying in wait, either at the same site or elsewhere. If this were true, then even as constraints increasingly bite on any one resource base, humanity would be able move to other resource bases, either at the same site or elsewhere. The enormous additions to the sources of industrial energy that have been realized (successively human and animal power, wind, timber, coal, oil and natural gas, and most recently, nuclear) are a prime historical illustration of this possibility. (But these shifts have not been without unanticipated collective costs. Global warming, associated with the burning of fossil fuels (an "externality"), did not feature in economic computations in earlier decades. (See Dasgupta (1993b) for a less coarse partition of substitution possibilities than the four-way classification.)

Humans have been "substituting" one thing for another since time immemorial. Even the conversion of forests into agricultural land in England in the middle ages was a form of substitution: large ecosystems were transformed to produce more food. However, the pace and scale of substitution in recent centuries have been unprecedented. Landes has argued that substitution made the Industrial Revolution in England in the eighteenth century (Landes, 1998). The extraordinary economic progress experienced in Western Europe and North America since then (during the past two centuries GDP per head in Western Europe has increased some twelve-fold), and in East Asia more recently, has also been a consequence of substitution. Spatial dispersion of ecosystems has enabled this to happen. The ecological transformation of rural England in

the middle ages presumably reduced the nation's biodiversity, but it increased income without any direct effect on global productivity.

But that was then and this article discusses the here and now. A question currently much debated is whether it is possible for the scale of human activity to be increased substantially beyond what it is today, without placing undue stress on the major ecosystems that remain. In any event, the cost of substituting manufactured capital for natural resources can be high. Low-cost substitutes could turn out to be not so low cost if accounting prices were used in the costing, not market prices. Even when accounting prices are not used, degrading natural capital and substituting it with manufactured capital can be uneconomic. Studies of the well-known Catskill Watershed ecosystem in New York State demonstrated the overwhelming economic advantages of conservation of the watershed over construction of water-purification plants. Chichilnisky and Heal, for example, showed that independent of the other services the Catskill Watershed provides, and ignoring the annual running costs of 300 million US dollars for a filtration plant, the capital costs alone showed a more than six-fold advantage for investing in the natural-capital base (Chichilnisky and Heal, 1998).

Degradation of a natural-resource base (e.g., destruction of native populations of flora and fauna) not only affects the volume and quality of ecosystem services the base provides, but it also challenges the system's resilience, which is its capacity to absorb disturbances, or perturbations, without undergoing fundamental changes in its functional characteristics. The way to interpret an ecosystem's loss of resilience is to view it as having moved to a new stability domain, which is another way of saying that the system, having crossed a "threshold," has been captured by a different attractor (Levin *et al.*, 1998; Levin, 1999; Brock *et al.*, 1999). Sudden changes in the character of shallow lakes (e.g., from clear to eutrophied water), owing to increases in the input of nutrients (Scheffer and Carpenter, 2003) and the transformation of grasslands into shrublands, consequent upon nonadaptive cattle-management practices (Perrings and Walker, 2004) provide two examples. Human societies have on occasions been unable to avoid suffering from unexpected flips in their local ecosystems because of this.

In particular ecosystems, biodiversity has been shown to be key to ecosystem resilience. However, even today it is a popular belief that the utilitarian value of biodiversity is located mainly in the potential uses of genetic material (e.g., for pharmaceutical purposes), or in other words, that its social worth is almost wholly an option value. Preservation of biodiversity is seen as a way to hold a diverse portfolio of assets with uncertain payoff. But as other contributions to this Encyclopedia make clear, biodiversity, appropriately conceived, is essential for the maintenance of a wide variety of services on which humans depend for survival. This has the important corollary that to invoke the idea of substitutability among natural resources in order to play down the use value of biodiversity, as has often been done (Simon, 1981), is a wrong intellectual move. The point is this: if an ecosystem's biodiversity is necessary for it to be able to continue providing humans with its services, the importance of that same biodiversity cannot be downplayed by the mere hope that for every species, there are substitute species lying in wait within that same ecosystem. In short, there is an inconsistency in this line of reasoning. Recall Ehrlich and Ehrlich's famous analogy relating species in an ecosystem to rivets in an airplane (Ehrlich and Ehrlich, 1981). One by one, perhaps, species may disappear and may not be missed. Eventually, however, the cumulative effect of loss of biodiversity will lead to the crash of ecosystem functioning, just as the cumulative loss of redundant rivets will lead to the crash of an airplane.

Discontinuous Value Functions

How do discontinuities in the social worth of ecosystems affect valuation exercises and social cost-benefit analysis? To answer this, it helps to formalize.

Consider an ecosystem describable by N state variables, indexed by i and j ($i, j = 1, 2, \ldots, N$). For concreteness, one may think of each state variable as reflecting the population size of a particular species. (As noted in the section Biological Resources and Biodiversity, problems of environmental pollution can be formulated in a similar manner.) Denote time by t (≥ 0) and let S_{it} be the population size of i at t. Time is taken to be a continuous variable. Imagine, therefore, that the dynamics of the ecosystem can be described by a system of (nonlinear) differential equations. For expositional ease, one can assume for the moment that the system is deterministic.

Let the net reproduction rate of i at t be F_{it}. Since the ecosystem is coupled, F_{it} is a function of the stocks at t. This can be written as $F_{it}(S_{1t}, S_{2t}, \ldots, S_{Nt})$, for $i = 1, 2, \ldots, N$. It is assumed that ecologists have estimated these functions. Assume next that the ecosystem dynamics are autonomous. This means that F_{it} is not an explicit function of t. So drop the subscript t from F_{it} and write the function as $F_i(S_{1t}, S_{2t}, \ldots, S_{Nt})$. In all the familiar applications of this framework with which we are familiar, F_i is taken to be a differentiable function. Assume this.

The analysis begins at $t = 0$ (the "present"). Denote by X_{it} the rate at which species i is harvested at time t. Now imagine that economists have studied the human–ecosystem interactions in question. They have enquired into the structure of property rights, demand conditions, government policies, and so forth. On the basis of this they have concluded that harvests are based on an implicit policy, in that they are time autonomous and are functions solely of stocks. So one may write $X_{it} = X_i(S_{1t}, S_{2t}, \ldots, S_{Nt})$. Assume that X_i is piece-wise continuous and possesses right- and left-partial derivatives everywhere. This is a technical assumption and a good one. For example, optimal policy functions for those ecosystem management problems that have been studied have been found to possess this property (Skiba, 1978; Brock *et al.*, 1999). Moreover, actual harvest rates have frequently been known to be approximately constant over time. So both sets of example satisfy the assumption. (Actual harvest rates frequently display time trends, say, because population and income grow. Time trends in X_{it} would render the system of eqn [1] nonautonomous. This article restricts the discussion to autonomous systems because it is understood that the mathematics of autonomous system is better than that of nonautonomus ones. But experience with simple nonautonomous systems

suggests that the arguments offered next cover them as well.) No doubt some of the X_{it}s would be zero. For example, it could be that only one species in the ecosystem is ever harvested (because, say, it is the only one that has economic worth). One should think of X_{it} as a forecast. It should be stressed that $X_i(S_{1t}, S_{2t}, \ldots, S_{Nt})$ is not necessarily a socially optimal harvest-policy function. It can be an actual policy function within an imperfect institution (e.g., the ecosystem could be one to which there is free access).

The rate of increase of S_{it} is the difference between F_i and X_{it}. Therefore, given the economists' forecast for X_{it}, mathematicians would be able to forecast S_{it} by solving the "coupled" system of differential equations:

$$dS_{it}/dt = F_i(S_{1t}, S_{2t}, \ldots, S_{Nt}) - X_i(S_{1t}, S_{2t}, \ldots, S_{Nt}), \quad \text{for all } i \quad [1]$$

For simplicity of exposition, assume that the social worth of the ecosystem is autonomous in time. Then express that worth by a scalar V. Since V would be a function of the stocks, and may be written as $V(S_{1t}, S_{2t}, \ldots, S_{Nt})$. V is the value at t of the entire ecosystem. It is the maximum amount society should be "willing to pay" at t for the ecosystem's survival if the stocks of the N resources were $S_{1t}, S_{2t}, \ldots$, and S_{Nt}, respectively. Any alternative use of the site (e.g., conversion into an urban center) would have to be worth at least V if the alternative were to be acceptable. The form of V would depend on the availability of substitutes for those species that are harvested. Again, to keep the mathematical notation from getting out of hand, assume that there are no substitutes available at low cost from outside the ecosystem (e.g., because the community doing the harvesting is not near other sources of livelihood). From eqn [1] it is possible to use the forecast on harvest rates to determine forecasts on stocks. This in turn makes it possible to forecast the time path of V.

At one level the valuation problem is now "solved": $V(S_{1t}, S_{2t}, \ldots, S_{Nt})$ would be the value of the ecosystem at t. It would be the social worth of the ecosystem at t. The problem is that V is typically a nonlinear function, which means that it is hugely difficult to estimate. The task of valuing ecosystems would be made much easier if recourse were taken to estimating accounting prices. The advantage would be this: since accounting prices reflect the social worth of marginal units of the various populations, one could use such prices to construct a linear index of the ecosystem's value. This article turns to this discussion.

Assume for the moment that V is differentiable everywhere. Let P_{it} be the accounting price of i at t. From the discussion in the section Institutional Failures and Poverty: Global versus Local Environmental Problems and from the assumption that no substitute resources are near in hand for the human community in question, it is known that (if substitutes were available, P_{it} would be the minimum of $\partial V/\partial S_{it}$ and the accounting price of the substitute. It is better to avoid such complications here.)

$$P_{it} \alpha \partial V/\partial S_{it}, \quad \text{for all } i \text{ and all } t \quad [2]$$

At time t, the value of species i would be $P_{it}S_{it}$. It follows that the value of the ecosystem itself would be $C_iP_{it}S_{it}$. Notice that this is a linear function of stocks, the weights being accounting prices.

In the section Institutional Failures and Poverty: Global versus Local Environmental Problems it is noted that a "project" can be thought of as a perturbation of the forecast X_{it}. So a project can be denoted as $(\otimes X_{1t} \otimes X_{1t}, \otimes X_{2t}, \ldots, \otimes X_{Nt})$, for $t \geq 0$. ($\otimes$ denotes an operator signifying "small difference".) (Note that for all i and all t, $\Delta X_{it} = \Sigma_j[\partial X_i/\partial S_{jt}][dS_{jt}/dt]\Delta t$.) Some of the $\otimes X_{it}$s would be zero. Nevertheless, the project would be expected to perturb future stocks of all the resources, since this is what a strongly coupled ecosystem would be expected to display.

Let r be the social rate of discount and let C_{it} be the unit cost of harvesting i at t. (C_{it} could depend on stock sizes at t. For example, the unit cost of fishing depends not only on the technology available for fishing and the price of fishing equipment, but also on the stock in the fishery: the larger the stock, the smaller the unit cost.) It follows that the present discounted value of the flow of net social benefits from the project is:

$$\int_0^\infty e^{-rt}[\mathbb{C}_i(P_{it} - C_{it}) \otimes X_{it}]dt \quad [3]$$

If expression [3] is positive, the project should be accepted; if it is negative, the project should be rejected.

It can be argued that projects, as it has been defined here, are merely "small" perturbations, whereas redirecting economic activity so as to avoid damaging an ecosystem irrevocably could involve drastic change. But it should be noted that one way to conceptualize a "large" perturbation is to regard it as the sum of a large number of small perturbations. A large perturbation (i.e., a large project) could then be evaluated by repeated use of expression [3].

However, if this route is not adopted, social cost–benefit analysis of large projects requires the project evaluator to estimate the large changes in V consequent upon the adoption of large changes in economic policy. Accounting prices, reflecting as they do the social worth of marginal units of the various resources (expression [2]), would then not suffice: the evaluator would need to integrate over the marginal units so as to estimate "consumer surpluses," to use a term familiar in economics.

So far, so good. But there is a problem with the account: it is unreasonable to assume that V is differentiable, even continuous, everywhere. Ecosystems are nonlinear systems (eqn [1]). So even if it were reasonable to suppose that V is differentiable everywhere else, it would be wrong to suppose that it is even continuous at loci of points separating different basins of attraction (i.e., at separatrices). (An important early contribution, showed via an example that if harvest functions are optimal, V is continuous even at points where harvests are discontinuous. This means that accounting prices are not uniquely specified at such points (Skiba, 1978). However, V can be shown to possess right- and left-partial derivatives there. So, accounting prices can be used for evaluation purposes, even though they are not uniquely given at every point on the space of resource stocks.)

But if the X_is are not optimal, V can be discontinuous at points on a separatrix. This causes problems, because accounting prices cannot even be defined at such points. Let the

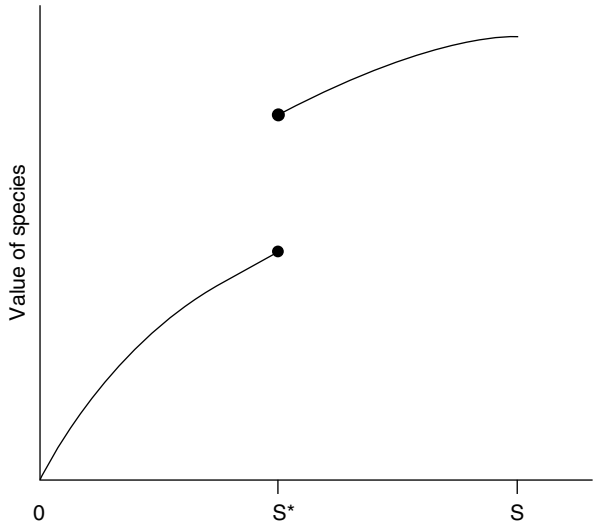


Figure 2 Nonlinear model depicting value of species in relation to function of the stock, where the ecosystem comprises a single species.

implications of this for biodiversity valuation and social cost-benefit analysis be studied.

Experience with nonlinear models of ecosystems tells that, under the assumptions made here, there could be at most a countable number of separatrices. This is fortunate, because it means that points on the stock space that are "troublesome" are nongeneric. So assume this. In **Figure 2** the matter is illustrated in the context of an ecosystem comprising a single species. The figure depicts the case where the separatrix is a single point, S^*, reflecting a threshold. For example, it could be that, under the harvesting policy $X(S)$, the species would become extinct if its population were below S^*, but would be harvested in a sustainable manner if the population more than S^*. So stocks to the right and left of S^* represent different basins of attraction. Reasonably enough, **Figure 2** depicts a case where the value of the species, $V(S)$, is an increasing function of the stock. It is assumed to be continuous (indeed, differentiable) everywhere except at S^*, where it jumps (an irrevocably dying population being a lot less valuable than a sustainable one). Of course, the location of S^* depends on $X(S)$: change the harvesting policy slightly, and S^* will shift slightly. The influence of $X(S)$ on S^* is something that has to be estimated if ecologists and economists are to offer policy advice.

Now, except by fluke the stock at $t=0$ would be different from S^*. So then assume it is different. If the project is sufficiently small, the account of social cost–benefit analysis discussed earlier remains valid: the system would not cross into a different basin of attraction. But a good theory should be extendable to fluke cases. Moreover, actual projects are frequently not "small," so acceptance of a project or its rejection could mean that the ecosystem is eventually in one basin of attraction rather than in another. How can this theory be extended the theory to handle the possibility that the ecosystem crosses into a different basin of attraction? In particular, is the repeated use of expression [3] a feasible means of evaluating projects?

It is as well to be clear where the problem lies if one were to try using expression [3]. The problem lies in that an accounting price cannot be defined at S^*. This means that a project that involves the stock passing through S^* cannot be evaluated by means of a linear index of social profitability. The height of the jump would have to be estimated and put to use in social cost–benefit analysis. Ecologists and economists would have to combine their expertise to locate S^* and identify the functional form of $V(S)$, both on the right and on the left of S^*. Estimate the height of the jump involves measuring "consumer surpluses", a point noted earlier. In short, at least one small project in the series of small projects that add up to the large project in question would not be assessable by means of expression [3]. This makes for difficulties.

Having noted this, there is a way to avoid the problem. The authors have been studying deterministic systems. Introducing uncertainty about the location of S^* can help matters by smoothing the value function. To see how, imagine that $V(S)$ represents the expected value of the resource's social worth at S. If the location of S^* were a smooth probability distribution, $V(S)$ would be a continuous, even a differentiable function. In this case an accounting price of the resource would be definable at all S (S^* being a smooth random variable). A linear index of the social profitability of projects could then be constructed. The methods of social cost-benefit analysis outlined earlier would remain valid.

It is not often that introducing realism simplifies analysis. Valuing biodiversity would seem to be an exception.

Conclusions

The US National Academy of Science described the fundamental challenge in valuing ecosystem services as "providing an explicit description and adequate assessment of the links between the structure and functions of natural systems, the benefits (i.e., goods and services) derived by humanity, and their subsequent values" (Heal *et al.*, 2005, p. 2). Despite the strides already taken in meeting this challenge, however, there are few economic studies of the value avoiding thresholds. Nor are there many studies of the value of the capacity of systems to keep functioning when subject to shocks or stresses. This is partly because there is still disagreement among ecologists of the role of biodiversity in maintaining either the stability or the resilience of ecosystems. At small scales an increase in species richness and the diversity of overlapping functional groups of species enhances the level of functional diversity, and hence both ecological stability (Tilman *et al.*, 2005) and resilience. At larger scales, the evidence is less clear. However, it is likely that maintaining a portfolio of options in ecosystems does strengthen their capacity to respond to shocks and stresses in constructive and creative ways. There is certainly evidence for this is coupled social–ecological system. In agricultural systems, for example, greater crop diversity has been shown to reduce the variance of both agricultural yields and farm incomes (Di Falco, 2007). Whereas the homogenization of agroecosystems has sometimes increased yields in the short run, it has often been at the cost of increasing variance in yields due to the increased vulnerability to pests and pathogens it brings. This in turn increases reliance on

pesticides that are potentially damaging to both humans and other species. In this case, the value of biodiversity – at least the value of the portfolio of cultivated species in the system – is the value to society of keeping crop production risks within acceptable bounds.

The wider economic literature has tended to identify biodiversity with conservation, and sometimes even more narrowly with the preservation of endangered species. It has also tended identify biodiversity change with biodiversity loss. But the extirpation or extinction of endangered species, important though it is, is only one dimension of biodiversity change. Biodiversity change affects the functioning of ecosystems in ways that potentially affect a large array of interdependent ecosystem services. Actions designed to enhance one service frequently compromise other services. In very many cases the effect of biodiversity change is to alter the capacity of the system to function over a range of environmental conditions. So the effect may not be immediate. In fact it may not be felt for many years. The valuation of biodiversity requires that one understands how a change in the mix of species within some functional group affects human wellbeing at various spatial and temporal scales. It also requires that one understands which effects of biodiversity change are captured in market prices and which are not.

It is not always the case that biodiversity loss is bad. Many very important ecosystem services depend on the simplification of ecosystems, and that frequently means the extirpation of particular populations. Protecting people against the effects of disease means controlling pathogens, even to the point of driving them to the extinction in the case of smallpox and rinderpest. The economic problem of biodiversity change is not that ecosystems have been simplified; it is that many of the less direct consequences of simplification are not taken into account by those responsible. The indirect effects of simplification on other ecosystem services, on other people, or on other places are frequently ignored. Such externalities of biodiversity change frequently impose costs on those who either have no voice (future generations) or are least able to cope with the effects (the poorest among the present generation). That is the economic problem of biodiversity change.

References

Allen BP and Loomis JB (2006) Deriving values for the ecological support function of wildlife: An indirect valuation approach. *Ecological Economics* 56: 49–57.

Anderson D (1987) *The Economics of Afforestation*. Baltimore, MD: Johns Hopkins University Press.

Arriagada R and Perrings C (2011) Paying for international environmental public goods. *Ambio* 40(7): 798–806.

Arrow K, Bolin B, Costanza R, *et al.* (1995) Economic growth, carrying capacity, and the environment. *Science* 268: 520–521.

Arrow KJ (1971) Political and Economic Evaluation of Social Effects of Externalities. In: Intriligator M (ed.) *Frontiers of Quantitative Economics*, vol. I., North Holland, Amsterdam.

Arrow KJ, Cline WR, Mäler K-G, Munasinghe M, Squitieri R, and Stiglitz JE (1996) Intertemporal equity, discounting, and economic efficiency. In: Bruce JP, Lee H, and Haites EF (eds.) *Climate Change 1995: Economic and Social Dimensions of Climate Change, Contribution of Working Group III to the Second Assessment Report of IPCC, the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

Arrow KJ and Fisher A (1974) Environmental preservation, uncertainty and irreversibility. *Quarterly Journal of Economics* 88: 312–319.

Barbier EB (2007) Valuing ecosystem services as productive inputs. *Economic Policy* 178–229.

Barbier EB (2008) Ecosystems as natural assets. *Foundations and Trends in Microeconomics* 4: 611–681.

Baumol WM and Oates W (1975) *The Theory of Environmental Policy: Externalities, Public Outlays and the Quality of Life*. Englewood Cliffs, NJ: Prentice-Hall.

Binswanger H (1991) Brazilian policies that encourage deforestation in the Amazon. *World Development* 19: 821–829.

Bohm P (1996) *Environmental Taxation and the Double-Dividend: Fact or Fallacy?*. London UK: University College. GEC Working Paper Series No. 96-01.

Bovenberg AL and Goulder L (1996) Optimal environmental taxation in the presence of other taxes: general equilibrium analyses. *American Economic Review* 86: 766–788.

Brock WA, Mäler K-G, and Perrings C (1999) *The Economic Analysis of Non-linear Dynamic Systems*. Stockholm, Sweden: Beijer International Institute of Ecological Economics.

Brock WA and Xepapadeas A (2002) Biodiversity management under uncertainty: species selection and harvesting rules. In: Kristrom B, Dasgupta P, and Lofgren K (eds.) *Economic Theory for the Environment: Essays in Honour of Karl-Goran Maler*, pp. 62–97. Cheltenham: Edward Elgar.

Carraro C and Siniscalco D (1996) *Environmental Fiscal Reform and Unemployment*. Dordrecht, Netherlands: Kluwer.

Chichilnisky G and Heal G (1998) Economic returns from the biosphere. *Nature* 391: 629–630.

Clark CW (1976) *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*. New York, NY: John Wiley.

Clark CWCFH and Munro GR (1979) The optimal exploitation of renewable resource stocks: problems of irreversible investment. *Econometrica* 47: 25–47.

Cleaver KM and Schreiber AG (1994) *Reversing the Spiral: The Population, Agriculture, and Environment Nexus in Sub-Saharan Africa*. Washington, DC: World Bank.

Cropper M and Oates W (1992) Environmental economics: a survey. *Journal of Economic Literature* 30: 675–740.

Daily G, Söderqvist T, Aniyar S, *et al.* (1999) *Valuing Resources*. Stockholm, Sweden: Beijer International Institute of Ecological Economics.

Dasgupta P (1982) *The Control of Resources*. Oxford: Blackwell.

Dasgupta P (1993a) *An Inquiry into Wellbeing and Destitution*. Oxford: Clarendon Press.

Dasgupta P (1993b) Natural resources in the age of substitutability. *Handbook of Natural Resource and Energy Economics* (ed. by A.V. Kneese & J.L. Sweeney). North Holland, Amsterdam, Netherlands.

Dasgupta P (1995) Population, poverty, and the local environment. *Scientific American* 272: 40–45.

Dasgupta P (2000) Reproductive externalities and fertility behaviour. *European Economic Review* 44: 619–644.

Dasgupta P (2001) *Human Wellbeing and the Natural Environment*. Oxford, New York: Oxford University Press.

Dasgupta P and Heal GM (1979) *Economic Theory and Exhaustible Resources*. Cambridge, UK: Cambridge University Press.

Dasgupta P, Levin S, and Lubchenco J (2000a) Economic pathways to ecological sustainability: challenges for the new millennium. *Bioscience* 50: 339–345.

Dasgupta P, Levin S, Lubchenco J, and Mäler K-G (2000b) *Ecosystem services and economic substitutability*. Stockhom, Sweden: Beijer International Institute of Ecological Economics.

Dasgupta P and Mäler K-G (1997) *The Environment and Emerging Development Issues*. Oxford: Clarendon Press.

Dasgupta P and Mäler K-G (2000) Net National Product, Wealth, and Social Wellbeing. *Environment and Development Economics* 5: 69–93.

Dasgupta P and Maler K-G (2004) *The Economics of Non-Convex Ecosystems*. Dordrecht, Boston, London: Kluwer Academic Publishers.

Dasgupta P, Marglin S, and Sen A (1972) *Guidelines for Project Evaluation*. New York, NY: United Nations.

Deacon RT (1994) Deforestation and the rule of law in a cross section of countries. *Land Economics* 70: 414–430.

Deangelis DL and Goldstein RA (1978) Criteria that forbid a large, non-linear food-web model from having more than one equilibrium point. *Math Biosci* 41: 81–90.

Di Falco SACJP (2007) On the role of crop biodiversity in the management of environmental risk. In: Kontoleon A, Pascual U, and Swanson T (eds.) *Biodiversity Economics: Principles, Methods, and Applications*, pp. 581–593. Cambridge: Cambridge University Press.

Dietz S and Adger N (2003) Economic growth, biodiversity loss and conservation effort. *Journal of Environmental Management* 68: 23–35.

Ehrlich PR and Ehrlich AH (1981) *Extinction: The Causes and Consequences of the Disappearance of Species*. New York, NY: Random House.
Eichner T and Pethig R (2005) Ecosystem and economy: An integrated dynamic general equilibrium approach. *Journal of Economics* 85: 213–249.
Filmer D and Pritchett L (1996) *Environmental degradation and the demand for children*. Washington, DC: World Bank.
Fisher A and Hanemann M (1986) Option value and the extinction of species. *Advances in Applied Microeconomics* 4: 169–190.
Freeman AMI (2003) *The Measurement of Environmental and Resource Values: Theory and Methods*, 2nd edn. Washington, DC: Resources for the Future.
Geertz CL (1963) *Agricultural involution*. Berkeley, CA: University of California Press.
Goodman D (1975) The theory of diversity-stability relationships in ecology. *Q Rev Biol* 3: 237–266.
Gordon HS (1954) The economic theory of common-property resources. *Journal of Political Economy* 62: 124–142.
Goulder L (1995) Environmental taxation and the double dividend: a reader's guide. *International Tax and Public Finance* 2: 157–183.
Gunderson LH (2000) Ecological resilience-in theory and practice. *Annu Rev Ecol Syst* 31: 425–429.
Hector A and Bagchi R (2007) Biodiversity and ecosystem multifunctionality. *Nature* 448.
Henry C (1974) Investment decisions under uncertainty: the "Irreversibility Effect". *American Economic Review* 64: 1006–1012.
Kinzig AP, Pacala SW, and Tilman D (2001) The functional consequences of biodiversity: Empirical progress and theoretical extensions. *Monographs in Population Biology* 33. Princeton, NJ: Princeton University Press.
Landes D (1998) *The Wealth and Poverty of Nations*. New York, NY: W.W. Norton.
Lansing JS (1991) Priests and programmers: Technologies of power in the engineered landscapes of Bali. Princeton, NJ: Princeton University Press.
Levin S (1999) *Fragile Dominion: Complexity and the Commons*. Reading, MA: Perseus Books.
Levin S, Barrett S, Aniyar S, *et al.* (1998) Resilience in natural and socioeconomic systems. *Environment and Development Economics* 3: 222–262.
Lind RC (ed.) (1982) *Discounting for Time and Risk in Energy Planning*. Baltimore, MD: Johns Hopkins University Press.
Lindahl ER (1958) Some controversial questions in the theory of taxation. In: Musgrave RA and Peacock AT (eds.) *Classics in the Theory of Public Finance*. London, UK: MacMillan.
Little IMD and Mirrlees JA (1974) *Project Appraisal and Planning for Developing Countries*. London, UK: Heinemann.
Loreau M, Naeem S, and Inchausti P (2002) *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*. Oxford: Oxford University Press.
Ludwig D, Carpenter S, and Brock W (2003) Optimal phosphorous loading for a potentially eutrophic lake. *Ecological Applications* 13: 1135–1152.
Mäler K-G (1974) *Environmental Economics: A Theoretical Inquiry*. Baltimore: Johns Hopkins Press.
Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, and Nobre CA (2008) Climate Change, Deforestation, and the Fate of the Amazon. *Science* 319: 169–172.
May RM (1972) *Stability and Complexity in Model Ecosystems*, 2nd ed Princeton, N.J.: Princeton University Press.
May RM (1977) Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* 269: 471–477.
Mccann KS (2000) The diversity-stability debate. *Nature* 405: 228–233.
Meade JE (1973) *The Theory of Externalities*. Geneva, Switzerland: Institute Universitaire de Hautes Etudes Internationales.
Millennium Ecosystem Assessment (2005) *Ecosystems and Human Wellbeing: General Synthesis*. Washington, DC: Island Press.
Mills JH and Waite TA (2009) Economic prosperity, biodiversity conservation, and the environmental Kuznets curve. *Ecological Economics* 68: 2087–2095.
Mitchell RC and Carson RT (1989) *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.
Mozumder P, Berrens RP, and Bohara AK (2006) Is there an Environmental Kuznets Curve for the risk of biodiversity loss? *The Journal of Developing Areas* 39: 175–190.
Muradian R (2001) Ecological thresholds: A survey. *Ecological Economics* 38: 7–24.
Naeem S (1998) Species redundancy and ecosystem reliability. *Conservation Biology* 12: 39–45.
Naeem S, Bunker D, Hector A, Loreau M, and Perrings C (2009) *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*. Oxford: Oxford University Press.
Naidoo R and Adamowicz WL (2001) Effects of Economic Prosperity on Numbers of Threatened Species. *Conservation Biology* 15: 1021–1029.
Panayotou T (1992) *Environmental Kuznets curves: empirical tests and policy implications. Harvard Institute for International Development*. MA: Harvard University.
Perrings C and Brock W (2009) Irreversibility in Economics. *Annual Review of Resource Economics* 1: 219–238.
Perrings, C and Halkos G (2010) Biodiversity loss and income growth in poor countries: the evidence. *ecoSERVICES Working Paper*. ASU, Tempe.
Perrings C and Walker B (1997) Biodiversity, resilience and the control of ecological-economic systems: The case of fire-driven rangelands. *Ecological Economics* 22: 73–83.
Perrings C and Walker BH (2004) Conservation in the optimal use of rangelands. *Ecological Economics* 49: 119–128.
Perrings C and Walker BH (2005) Conservation in the optimal use of rangelands. *Ecological Economics* 49: 119–128.
Phelps J, Guerrero M, Dalabajan D, Young B, and Webb EL (2010) What makes a REDD country? *Global Environmental Change* 322–332.
Pigou AC (1920) *The Economics of Welfare*. London, UK: Macmillan.
Pimm SL (1984) The complexity and stability of ecosystems. *Nature* 307: 321–326.
Portney PR and Weyant JP (eds.) (1999) *Discounting and Intergenerational Equity*. Washington, DC: Resources for the Future.
Repetto R, *et al.* (1989) *Wasting Assets: Natural Resources and the National Income Accounts*. Washington, DC: World Resources Institute.
Reyers B, Van Jaarsveld AS, Mcgeoch MA, and James AN (1998) National Biodiversity Risk Assessment: A Composite Multivariate and Index Approach. *Biodiversity and Conservation* 7: 945–965.
Scheffer M and Carpenter SR (2003) Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends in Ecology & Evolution* 18: 648–656.
Schulze ED and Mooney HA (1993) *Biodiversity and Ecosystem Function*. New York: Springer-Verlag.
Simon JL (1981) *The Ultimate Resource*. Princeton, NJ: Princeton University Press.
Skiba AK (1978) Optimal growth with a convex–concave production function. *Econometrica* 46: 527–539.
Solorzano R, *et al.* (1991) *Accounts Overdue: Natural Resource Depreciation in Costa Rica*. Washington, DC: World Resources Institute.
Thebault E and Loreau M (2006) The relationship between biodiversity and ecosystem functioning in food webs. *Ecological Research* 21: 17–25.
Tietenberg T (1990) Economic instruments for environmental regulation. *Oxford Review of Economic Policy* 6: 17–33.
Tilman D, May RM, Polasky S, and Lehman CL (2005) Diversity, productivity and temporal stability in the economies of humans and nature. *Journal of Environmental Economics and Management* 49: 405–426.
Tilman D, Wedin D, and Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379: 718–720.
Vincent J and Ali RM, and Associates (1997) *Environment and Development in a Resource-Rich Economy: Malaysia under the New Economic*. Cambridge, MA: Harvard Institute for International Development.
Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of Earth's ecosystem. *Science* 277: 494–499.
Walker B, Kinzig AP, and Langridge J (1999) Plant attribute diversity, resilience, and ecosystem function: The nature and significance of dominant and minor species. *Ecosystems* 2: 95–113.
Weitzman ML (1974) Prices vs. quantities. *Review of Economic Studies* 41: 477–491.
World Bank (1992) *World Development Report*. New York, NY: Oxford University Press.

Threatened Birds

S Jacob Socolar and David S Wilcove, Princeton University, Princeton, NJ, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Nigel J. Collar, volume 2, pp. 395–406, © 2001, Elsevier Inc.

Glossary

Critically endangered Facing extremely high risk of extinction in the wild, based on specific criteria (see IUCN, 2001).

Data deficient Appropriate data are lacking to apply the International Union for Conservation of Nature (IUCN) criteria.

Endangered Facing very high risk of extinction in the wild, based on IUCN criteria.

Extinct There is no reasonable doubt that the last individual has died, based on exhaustive surveys at appropriate times in proper habitat throughout historical range.

Extinct in the wild Although extinct, individuals survive in captivity.

Least concern Not qualifying or close to qualifying for any threat category.

Near-threatened Not currently meeting any of the criteria for the above-mentioned threat categories, but close to qualifying or likely to qualify soon.

Vulnerable Facing a high risk of extinction in the wild, based on IUCN criteria.

Introduction

The concept of a threatened bird is easy to grasp but surprisingly difficult to define. A clear idea of what constitutes a threatened bird requires firm notions of taxonomy, geographical scale, and the degree of peril that qualifies the bird as threatened. This article focuses primarily on the endangerment of species at the global scale, but nationally and locally threatened forms, including subspecies, are also considered. The authors' emphasis on global endangerment reflects the premise that total extinction is a worse outcome than local extirpation, although this premise is not necessarily self-evident (Kareiva, 2011).

The discussion on threatened birds is enlightening simply because of the amount that is known about birds compared with other major taxonomic groups. As such, threatened birds offer a window into the broader topic of the loss of biodiversity. Scientists and conservationists have accumulated a wealth of information about threatened birds. Scientists now know in great detail the status of nearly every bird species and the specific threats they face. Based on their knowledge of threats and experience with management techniques, they have increasingly refined their notions of what conservation actions need to be prioritized (Wilson *et al.*, 2009). Few (if any) other taxonomic groupings permit this level of analysis. Partly for this reason, avian conservation biology has provided conservation biology at large with a wealth of case studies, insights, and technical knowledge.

This article aims to provide an overview of threatened birds around the world. Information and statistics about global patterns of bird endangerment frame the discussion, enabling the identification of the most important threats faced by threatened birds today. A survey of these threats is followed by a review of management techniques and conservation opportunities. Detailed discussion of threats and conservation measures is beyond the scope of this article. Instead, brief summaries of the issues are complemented with specific examples.

Numbers and Trends in the Endangerment of Birds

Taxonomists currently recognize more than 10,000 species of birds in the world, many of which are threatened. Rather than making sense of the global situation on a species-by-species basis, it is helpful to recognize broad trends in bird endangerment. This section discusses some of the most basic questions about global patterns: How many threatened birds are there, and how severe is their plight? Is their status improving or deteriorating? Where do they occur and in which habitats? How does the picture differ when viewed through global versus regional and national lenses? This sort of information allows us to recognize globally important threats to birds and provides the background against which to appreciate specific examples.

Such a discussion requires precise criteria to identify species that are in peril of extinction as well as the empirical data to apply those criteria correctly. In the last 45 years, global species endangerment has been registered principally through the Red Data Book program of the International Union for Conservation of Nature (IUCN), with BirdLife International in charge of the Red List for birds (BirdLife International, 2000, 2011). The IUCN has developed objective criteria to categorize the degree of a species' endangerment based on the severity of its plight, applying thresholds for range and population size and decline designed to reflect the probability of extinction (IUCN, 2001). Hereafter, this article capitalizes terms such as "Endangered" or "Extinct" to refer to particular IUCN Red List categories (IUCN, 2001); the authors use the term "threatened" to refer to imperiled species in general, comprising the ensemble of Critically Endangered (CR), Endangered (EN), and Vulnerable (VU).

Notwithstanding all we know about birds, incomplete data still cause obstacles to their conservation. The species-level taxonomy of birds is unresolved and continues to change. Furthermore, the information required to apply the IUCN's threat categories is sometimes not known. As a result, a few

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00421-4

bird species cannot be assigned a threat level currently, whereas in nearly all others, the threat level is based on estimation, as opposed to precise empirical data. Nevertheless, compared with most taxonomic groups, bird conservation benefits from an extraordinary wealth of information, and threatened birds benefit from an extraordinary degree of conservation effort.

Globally Threatened Birds in 2011: Numbers and Outlook

The IUCN Red List (2011) and the companion publication Threatened Birds of the World (BirdLife International, 2000) are the definitive references on global patterns of bird endangerment. The 2011 update to the IUCN Red List recognizes 10,052 species extant after AD 1500. Of these, 132 (1%) species are now Extinct (EX). Of the remaining species, 189 (2% of extant species) are Critically Endangered (CR), 381 (4%) Endangered (EN), and 683 (7%) are Vulnerable (VU), representing 13% of the world's currently extant avifauna. A further 844 species (8%) are considered Near-Threatened (NT), indicating that they may become threatened in the future, whereas almost all remaining species are classified as Least Concern (LC). Insufficient information exists to assess the status of only 62 species (1%), classified as Data Deficient (DD). Thus, 2159 bird species, or 22% of the world's avifauna, are identified as being of global conservation concern.

Across the board, these totals are higher than the equivalent totals from previous assessments (Collar *et al.*, 1994; BirdLife International, 2000). However, using such figures to infer temporal trends in bird endangerment is not straightforward. Most changes in the threat status of bird species reflect taxonomic revisions or newly available population data rather than genuine changes in the status of bird populations. The "splitting" of species-level taxa can cause threatened bird species to appear out of thin air, especially because the populations of the newly split taxa necessarily will be lower than their total population. Similarly, novel information about a species' population may warrant a change in the threat assessment without reflecting any change in the actual status of the species.

One illustration of the temporal trends in bird endangerment comes from the fact that the vast majority of red-listed birds have decreasing populations. This trend arises because the outlook for many bird species continues to deteriorate, but it is partly circular because severe population declines trigger inclusion on the Red List. Only species with irremediably small ranges or populations are listed in absence of an active threat or population decline. Birds like the Lava Gull (*Larus fuliginosus*) VU, with 300–400 pairs maximum, and the Cocos Flycatcher (*Nesotriccus ridgwayi*) VU, found only on an island of less than 25 km^2, destined to remain forever threatened, and the price of their survival is eternal vigilance.

Another general picture of the deteriorating status of threatened birds derives from the Red List Index, a metric developed by the IUCN to track changes in Red List status due to genuine changes in population status (Butchart *et al.*, 2004). For birds, the Red List Index has consistently shown negative trends since 1988, and the pattern holds across all major geographic regions of the planet when their avifaunas are disaggregated. Unmistakably, birds across the globe are becoming more threatened.

Globally Threatened Birds: Regions, Countries, and Habitats

Most threatened bird species occur in Asia and South America, with comparatively few in Africa. The top 10 countries for numbers of threatened birds are Brazil (122), Indonesia (119), Peru (98), Colombia (94), China (84), India (77), USA (76), Philippines (73), Ecuador (72), and New Zealand (69). USA ranks high on the list largely because of Hawaii and its Pacific territories; otherwise, Asia and South America predominate. Together, the top 10 countries are home to 729 (58%) of the 1253 species of threatened birds in the world.

Large numbers of Endangered and Critically Endangered birds occur on islands – 99 species on the Pacific islands of Oceania (including the Hawaiian archipelago but not the Australian mainland) alone, 50 in Indonesia, 32 on the Caribbean Islands, and 27 in the Philippines. The contemporary situation is but a narrow snapshot of a centuries-old pattern of bird endangerment and extinction on islands (*see* Extinct Birds). In a historical sense, the noteworthy feature of present bird endangerment is the large number of threatened mainland birds, including 165 Endangered and Critically Endangered species just in South America. Ricketts *et al.* (2005) note that the growing roster of Endangered and Critically Endangered species of birds and other taxa on continents portends the movement of the anthropogenic extinction wave from islands to the more species-rich continents (*see* Causes of Endangerment in Birds).

Because most of the world's bird species (76%) live in forests, it should come as no surprise that most threatened birds are forest dwellers – in fact, the figure is precisely 76%. This figure represents a balance between rampant forest destruction across the globe and the fact that the largest tracts of forest, in Amazonia, the Congo basin, and New Guinea, are – despite the destruction visited on them – still too extensive, and the species they contain are too widespread, to have resulted in more than a small number of listings from these areas. Shrublands harbor 27% of threatened bird species, inland wetlands harbor 16%, and grasslands harbor 16%. Marine habitats collectively, including coastal and intertidal areas, are home to 13% of the threatened bird species. The percentages sum to more than 100 because many bird species utilize multiple habitat types in different parts of their range or at different times of year. In the former case, preservation of either habitat may be sufficient to permit the survival of the species; the latter case often requires the conservation of both habitats.

Globally Threatened Birds: Taxonomic Groups

Particular taxonomic groups with high proportions of threatened birds often reflect a group's collective sensitivity to a particular threat, and a survey of threatened bird orders will help to frame the discussion of threats later in this article. Although nearly every extant bird order contains at least one threatened species (exceptions being the loons (Gaviiformes) and mousebirds (Coliiformes)), some orders and families have been particularly hard-hit. Of the 129 extant tubenoses (Procellariiformes), 59 (46%) are threatened. Within the Procellariiformes, the 22 extant species in the albatross family (Diomedeidae) include 17 threatened species (77%),

many of them Critically Endangered. Eleven of 18 (61%) penguins (Sphenisciformes) are threatened, as are 97 of 355 (27%) parrots (Psittaciformes), 74 of 291 (25%) game birds (Galliformes), and 58 of 205 (28%) rails and cranes (Gruiformes). The rails also feature heavily among the recently extinct birds, largely due to their evolutionary diversification on islands. To place these percentages in perspective, recall that the overall percentage of threatened birds is "only" 13%.

Regionally Threatened Birds in USA and UK

Patterns of bird endangerment and conservation priorities at regional and global scales are not necessarily identical. Because bird conservation organizations have availed themselves of high levels of interest and funding to thoroughly assess the status of birds in the USA and the UK, these countries provide a lens to understand patterns at the regional scale and to appreciate the plight of declining birds that miss global IUCN criteria for endangerment. In the USA, the American Bird Conservancy and the National Audubon Society have developed the WatchList, classifying the entire avifauna into the threat categories red, yellow, and green (Butcher *et al.*, 2007). Similarly, in the UK, the Royal Society for the Protection of Birds (RSPB) has developed a list of Birds of Conservation Concern, categorizing the avifauna in the categories red, amber, and green (Eaton *et al.*, 2009).

The 2007 WatchList applies criteria based on population size, range size, threats, and population trend to determine the status of US birds (Panjabi *et al.*, 2005). The WatchList gives "red" status to 59 birds of the continental USA (including Alaska) and to an additional 33 from Hawaii. "Yellow" status is accorded to 117 birds of the continental USA. Together, these birds represent about a quarter of the avifauna of the continental USA. Many of these birds are not threatened under global criteria, as they may be relatively common in other parts of the world. For example, the Yellow Rail (*Coturnicops noveboracensis*) is listed as "red" on the WatchList but is placed in the Least Concern category on the Red List. Only 37 birds on the continental USA WatchList (21% of continental species on the WatchList), mostly in the red category, are globally threatened (VU or higher) based on IUCN criteria; many of the remaining birds are NT. Of the 52 birds on the RSPB's Red List, only 2, the Balearic Shearwater (*Puffinus mauretanicus*) CR and the Aquatic Warbler (*Acrocephalus paludicola*) VU, are globally threatened under IUCN criteria. The remainder is listed based on severe historical declines in the UK without recent recovery or based on more recent declines in UK breeding population, wintering population, or breeding range, since either 1969 or 1984. Twenty-seven of the UK red-listed birds are species of European conservation concern (BirdLife International, 2004). Because IUCN status was a trigger for inclusion on the UK red list, no globally threatened birds were left off, but the Balearic Shearwater's listing was not triggered by any other criterion. The amber list includes an additional 126 species with declining UK populations (but not severely enough to merit inclusion on the red list), low UK populations, or localized UK distributions. Also included are all remaining species of European conservation concern.

The USA and the UK share critical cultural and economic similarities, and trends in the local endangerment of their birds might not hold everywhere in the world. However, the detail of the WatchList and RSPB's Birds of conservation concern provides the best available picture of regional trends. Both regional lists designate more birds as threatened than that listed by BirdLife International. This scenario applies particularly to the UK, where birds that are secure at the continental scale may still be threatened regionally. In other cases, different evaluation schemes produce different results. Perhaps regional-scale assessments can afford to list relatively less imperiled species without the risk of muddling global conservation priorities. Most importantly, these lists make it clear that large numbers of globally common birds that miss the IUCN criteria are nevertheless in steep decline.

Extinct Birds

Since AD 1500, at least 132 bird species have gone extinct (EX), and a further four are extinct in the wild (EW). Strikingly, 121 (92%) of these species were confined to islands, mostly in the Pacific, and an additional three were confined to single lakes (BirdLife International, 2011). Moreover, losses since AD 1500 represent only the tip of the iceberg in terms of human-caused extinction of birds. Bird extinctions in the more distant past may be inferred from subfossil data. The spread of Polynesian seafarers across the islands of the Pacific caused between 820 and 1960 bird extinctions, making it one of the largest vertebrate extinction events ever recorded (Steadman, 2006). The number of individual extirpated populations is up to 10 times higher.

Because extinction is only demonstrated by documenting the absence of a species over its entire range, mainland extinctions may be harder to detect than island extinctions. Still, the overwhelming historical preponderance of island extinctions is unambiguous.

Because of the increasing contemporary plight of mainland birds, the few recent mainland extinctions are of heightened interest. Most of these have occurred in eastern North America, where the Labrador Duck (*Camptorhynchus labradorius*), Great Auk (*Pinguinus impennis*), Carolina Parakeet (*Conuropsis carolinensis*), Passenger Pigeon (*Ectopistes migratorius*), and likely the Bachman's Warbler (*Vermivora bachmanii*) and Eskimo Curlew (*Numenius borealis*) have all vanished. The Ivory-billed Woodpecker (*Campephilus principalis*) is almost certainly gone as well, although controversy lingers about the possibility of its survival (Fitzpatrick *et al.*, 2005; Sibley *et al.*, 2007). Compared with extinct island birds, many of the extinct mainland species ranged over vast areas or occurred in vast numbers, indicating that factors beyond range and population size can put birds at risk for endangerment and eventual extinction. In the cases of the Great Auk, Carolina Parakeet, Passenger Pigeon, Eskimo Curlew, and Ivory-billed Woodpecker, hunting/direct persecution has been implicated as an important factor in their demise. The Bachman's Warbler may have been extirpated by destruction of its breeding and wintering habitats. Too little is known about the Labrador Duck to make conclusive statements about the causes of its decline.

Causes of Endangerment in Birds

Although some birds are at risk by virtue of their naturally small ranges or populations, the vast majority of threatened birds are imperiled due to the actions of humans. Some 82% are affected by habitat loss and degradation, most often related to agriculture (BirdLife International, 2011). Forty percent are adversely affected by other species in their environment, including invasive exotics or natural competitors that exacerbate additional threats. Hunting and trapping threaten 38% of threatened birds, whereas climate change and severe weather events potentially threaten 33%. In general, small ranges or low populations make birds particularly sensitive to these threats (Note that the percentages sum to more than 100 because most threatened birds are imperiled by multiple factors).

Habitat Loss and Degradation

When habitats change, the birds that lived in them frequently cannot persist in the new landscape. Habitat loss is a major factor in bird endangerment because by the turn of the millennium, at least one-third of the Earth's terrestrial surface had been reshaped by human activities, such as agriculture, ranching, logging, and industrial development (Vitousek *et al.*, 1997). In the humid tropics, where bird diversity reaches its peak, forests were lost or degraded at a rate of approximately 8 million hectares per year during the 1990s (Achard *et al.*, 2002). Deforested land was primarily logged and/or converted to agriculture with major threats, including logging, cattle ranching, soybean farming, and oil palm plantations. The trend is expected to continue. By the year 2050, the Earth's burgeoning population will require up to 1 billion hectares of additional agricultural lands to feed itself (Tilman *et al.*, 2001).

Habitat Destruction

In the simplest cases, birds become imperiled by habitat destruction when their habitat is obliterated over large portions of their range. For example, the Cebu Flowerpecker (*Dicaeum quadricolor*) CR is endemic to the Philippine island of Cebu. Long feared extinct due to the near complete obliteration of forest on the 5088 km^2 island, it was rediscovered in 1992 in a patch of degraded forest totaling less than 2 km^2 (Dutson *et al.*, 1993). Because some bird species require much larger territories than others, their range requirements may span several orders of magnitude. On the nearby island of Mindanao, pairs of Philippine Eagles (*Pithecophaga jefferyi*) CR maintain territories of 133 km^2 on average, including 68 km^2 of forest (BirdLife International, 2011). Thus, despite the presence of about 9000 km^2 of forest remaining within its range, this species too is severely threatened due to habitat loss.

A layer of complexity arises when birds use multiple habitats at different times of the year or different regions of a habitat in different years. This is the situation that is faced by all migratory or nomadic bird species. Migratory species can depend on the conservation of multiple disjunct sites comprising multiple habitats. For example, despite attempts to manage the Golden-cheeked Warbler (*Dendroica chrysoparia*) EN on the shrinking and fragmented oak–juniper woodlands of its small US breeding range, the species also is threatened by the destruction of high-elevation oak woodlands in its Mesoamerican wintering range (Rappole *et al.*, 2003). Similarly, nomadic species may depend on the conservation of vast blocks of contiguous habitat. The Purple-winged Ground-Dove (*Claravis godefrida*) CR is a nomadic species intimately linked to the mast seedings of woody bamboos (Areta *et al.*, 2009). The bamboos occur patchily in space and have intermast intervals of approximately 30 years, so the survival of the dove depends on sufficiently extensive areas of forest to ensure that a suitable patch of masting bamboo is generally available, and that the birds can find their way between such patches.

Other examples of this space/time vulnerability include the Thick-billed Parrot (*Rhynchopsitta pachyrhyncha*) EN, Andean Flamingo (*Phoenicopterus andinus*) VU, Resplendent Quetzal (*Pharomachrus mocinno*) NT, and perhaps even the Passenger Pigeon (*E. migratorius*) EX. The parrot shows the same trait as the Purple-winged Ground-Dove because it is dependent on pine seed, a notoriously unpredictable resource. Most, if not all, Thick-billed Parrots are believed to be nomadic, but as their native pine forests in Mexico's Sierra Madre are further fragmented, there is a danger that a cone-crop failure will leave the last populations "stranded" far from food. The flamingo moves between lakes in search of appropriate foraging resources (which shift over time) and is therefore exposed to the possibility that human damage to even a small number of sites may one day leave the species with nowhere to go (BirdLife International, 2011). Postbreeding quetzals make complex short-distance movements to several different areas, so that multiple tracts of forest may be needed to ensure the long-term survival of viable populations of this species (Powell and Bjork, 1994). Bucher (1992) shows how even the Passenger Pigeon was a specialist on seeds produced in masting events whose scale and geographic location varied from year to year.

Habitat Degradation and Fragmentation

Habitat degradation need not entail the wholesale elimination of suitable habitat to pose a threat to species. Pollution, changing land management practices, and even indirect effects of other species on habitat can cause bird populations to decline. The plight of the Junin Grebe (*Podiceps taczanowskii*) CR results from pollution and the management of lake levels. The grebe is entirely confined to Lake Junin, Peru, and declined following the contamination of the lake with mining wastes (Fjeldsa 1984) and the regulation of the lake level by mining companies (BirdLife International, 2011). The Paradise Parrot (*Psephotus pulcherrimus*) EX of Australia disappeared probably due to new burning patterns, livestock grazing, and drought, all of which diminished the availability of the plants that the parrot fed on (Joseph, 1988). Changes in Australian fire regimes also afflict the food supply of the Golden-shouldered Parrot (*Psephotus chrysopterygius*) EN and the habitat of the Noisy Scrub-bird (*Atrichornis clamosus*) VU and Mallee Emu-wren (*Stipiturus mallee*) EN, both of which require areas that have not burned too recently. In the Longleaf Pine savannahs of the southeastern USA, fire suppression is the problem. A lack of fires has permitted the growth of a broadleaf understory, rendering the habitat unsuitable for the Red-cockaded Woodpecker (*Picoides borealis*) VU (Ligon *et al.*, 1986) and

Bachman's Sparrow (*Peucaea aestivalis*) NT. A great variety of subtle changes on the land can cause bird populations to decline.

Species may decline if their habitat becomes fragmented, even if large amounts remain. Local populations in a fragment may die out due to either stochastic processes or Allee effects, and the fragment may not be recolonized if dispersal is poor. This situation has been documented for understory-dwelling insectivorous birds in Amazonian forest fragments (Stouffer and Bierregaard,1995; Ferraz *et al.*, 2007). Alternatively, fragmentation may open bird populations to pressures arriving from the fragmenting matrix, including arriving nest predators and brood parasites (*see* Natives Threatened by Other Natives). In the midwestern USA, these factors appear to turn forest fragments into population sinks for the Wood Thrush (*Hylocichla mustelina*) LC, a yellow species on the WatchList (Robinson and Wilcove, 1994).

Biotic Threats

Sometimes, despite the existence of appropriate habitat, birds may become threatened as a result of the assemblage of other species present there. The introduction of alien (nonnative) predators, competitors, or diseases can result in precipitous declines of native bird species, especially on islands. In other cases, competition or predation from native species can exacerbate the plight of species that are already declining in the face of other threats. Finally, some species depend on particular native species and run into trouble if these native species decline.

Alien Species

On oceanic islands where mammalian predators of birds and nests were historically absent, their introduction has resulted in massive bird declines and numerous extinctions. Across the world's oceanic islands, greater numbers of mammalian introductions (rats, cats, stoats, etc) correlate with higher rates of extinction of native birds (Blackburn *et al.*, 2004). In a few cases, it has been possible to link these declines to specific predators. Because oceanic avifaunas evolved in the absence of mammalian predators, they are behaviorally and physically adapted in ways that leave them highly vulnerable when continental predators establish within their ranges, through either the direct or indirect agency of humans.

These behavioral adaptations are not degenerative. The Dodo (*Raphus cucullatus*) EX was given its name from the Portuguese slang for "stupid," *doido*. But island animals that are especially tame, or that nest in what to human eyes are ludicrously undefended places, or that have lost the function of their wings are not evolutionary failures. On the contrary, these seemingly disadvantageous attributes are the result of continuing evolutionary pressures. Wings cost energy to carry and maintain; and in the absence of predators, they offer no return on the investment to such energy. Shyness costs its possessors dearly if less shy creatures have more time to exploit whatever resource is at stake. Nesting in inaccessible places is needlessly expensive if there is no risk in nesting on the ground in the open. Moreover, the relatively stable conditions on tropical islands tend to promote marked K-selected traits (e.g., slow reproductive rates). The factors that render island birds vulnerable to aggressive, fast-breeding, continental animals may have been favored by eons of narrower, often intraspecific competition.

Island birds have been, and continue to be, massively impacted by rats (Moors *et al.*, 1992): Magenta Petrel (*Pterodroma magentae*) CR, Tuamotu Sandpiper (*Prosobonia cancellata*) EN, Polynesian Ground-dove (*Gallicolumba erythroptera*) CR, Seychelles Paradise-flycatcher (*Terpsiphone corvina*) CR, and Rarotonga Monarch (*Pomarea dimidiata*) EN are a few of the birds whose declines are directly linked to the invasion of their islands by rats. Species on larger islands face introduced predator assemblages that include cats, and the large island size makes predator control or eradication particularly difficult. The New Caledonian Rail (*Gallirallus lafresnayus*) CR (perhaps EX), Cuba's Zapata Rail (*Cyanolimnas cerverai*) CR, the Samoan Moorhen (*Gallinula pacifica*) CR (perhaps EX), and virtually the entire surviving endemic avifauna of Hawaii (notably the honeycreepers, Drepanididae) are good examples of this difficulty. Mongooses, which were often released to devour rats or control snakes, are major threats to species, such as Hawaiian Duck (*Anas wyvilliana*) EN and St. Lucia's Semper's Warbler (*Leucopeza semperi*) CR (perhaps EX). Introduced diseases may also cause disaster among local bird populations. Accidentally introduced mosquitoes carrying avian malaria and pox devastated Hawaii's land birds and continue to pose a serious problem, which may worsen with climate change (van Riper III *et al.*, 1986; BirdLife Internaional, 2011). Numerous other diseases, including avian botulism, West Nile virus, and avian cholera are just a few diseases that have caused worrisome declines in threatened birds (Butchart *et al.*, 2010). Mammals are not the only alien predators to cause harm. In Guam, most of the island's native birds were extirpated by the accidentally introduced brown tree snake (*Boiga irregularis*) (Savidge, 1987).

In other cases, introduced competitors (rather than predators) cause problems. New Zealand's Kaka (*Nestor meridionalis*) EN and Yellowhead (*Mohoua ochrocephala*) EN face enormous difficulties now as the introduced wasps compete with them for honeydew, an important foodstuff (breeding success in the Kaka correlates with honeydew intake in the previous autumn). The White-headed Duck (*Oxyura leucocephala*) EN faces long-term extinction through hybridization with its New World counterpart, the Ruddy Duck (*Oxyura jamaicensis*) LC, which is spreading through the White-headed Duck's range from an original feral population in Britain (BirdLife Internaional, 2011).

Natives Threatened by Other Natives

Sometimes threats come not from introduced exotics but from natives expanding their ranges, often owing to human modifications of habitat. The spread of cowbirds (*Molothrus* spp.) in the USA and the Caribbean islands is particularly worrying. Birds, such as Kirtland's Warbler (*Dendroica kirtlandii*) NT and Black-capped Vireo (*Vireo atricapillus*) VU, need constant effort programs to reduce cowbird brood parasitism to tolerable levels. Similarly the spread of the Pearly-eyed Thrasher (*Margarops fuscatus*) LC, a nest-hole competitor, through the Caribbean islands has been viewed with alarm by the Puerto Rican Amazon (*Amazona vittata*) CR recovery teams.

Researchers classify such species as Conservation Dependent in recognition of the possibility that they will need perpetual care (Scott *et al.*, 2005).

In the face of declines due to other factors, even normal pressures from native species may become problematic. On New Zealand's Chatham Islands, nest predation by the native Weka (*Gallirallus australis*) LC has contributed to the decline in numbers of Chatham Oystercatcher (*Haematopus chathamensis*) EN, an endemic species already threatened by the introduced mammalian predators. On Bermuda, White-tailed Tropicbirds (*Phaethon lepturus*) LC outcompete Bermuda Petrels (*Pterodroma cahow*) EN for nests and must be controlled in the vicinity of petrel colonies.

A separate class of cases involves birds that depend on particular native species. When these native species decline, it spells trouble for the birds that depend on them. For example, the Cuban Kite (*Chondrohierax wilsonii*) CR is threatened by habitat loss and persecution, but the harvesting of tree snails, its major foodstuff, has exacerbated its problems (BirdLife Internaional, 2011). Shorebirds that utilize Delaware Bay, USA as a spring stopover site have been hard-hit by the harvesting of Horseshoe Crabs there for fish bait. Most of the world's Semipalmated Sandpipers (*Calidris semipalmatus*) LC and most of the eastern North American population of Red Knots (*Calidris canutus*) LC feed on the eggs of spawning Horseshoe Crabs, putting on fat reserves that enable them to continue northward to their Arctic breeding grounds. Both species have shown serious declines as a result of the Horseshoe Crab harvest, and both are yellow species on the WatchList.

Complicated pathways involving multiple species and indirect effects can also degrade bird habitat. The decline of Northern Wheatears (*Oenanthe oenanthe*) LC in the Netherlands apparently results from grass and shrub encroachment following the decline of rabbits in their dune habitat (van Oosten, 2010). The extirpation of apex predators from huge portions of the planet's terrestrial surface is likely responsible for a suite of underappreciated habitat changes and threats, including mesopredator release (Estes *et al.*, 2011), which has been tied to declines of North American songbirds (Wilcove, 1985; Soulé *et al.*, 1988).

Hunting, Trapping, and Poisoning

Some birds are threatened due to capture and/or mortality directly caused by humans. In many cases, the killing of birds is deliberate (whether rational or not), but in other cases it is incidental to other activities. The former cases include hunting for sustenance and trapping for the pet trade; both phenomena have produced major declines in some birds, and overharvesting contributed to approximately one-third of extinctions since 1500 (Butchart *et al.*, 2010). Avian families with especially high rates of endangerment often suffer from particular threats in combination, with habitat loss or alien species compounded by unsustainable harvesting (Collar *et al.*, 1997). For example, the parrots (Psittacidae) and curassows and guans (Cracidae) have declined due to habitat loss, but the parrots also experience intensive trapping for the pet trade, and the Cracids are intensively exploited for food. Trapping has contributed to the endangerment or extinction of approximately one-quarter of the world's parrots. Among the Cracidae, the curassows (genera *Crax*, *Mitu*, *Nothocrax*, and *Pauxi*) have been especially hard-hit by hunting. Hunting for food has contributed to the extinction in the wild of the Alagoas Curassow (*Mitu mitu*) and to the endangerment of 6 of the 13 remaining species, with a seventh species Near-Threatened (BirdLife Internaional, 2011). Historically, extinctions have resulted from the overzealous persecution of threatened species by specimen collectors. The curse of economic value combined with restriction to a beleaguered habitat type confers a strong likelihood of endangerment.

In addition to hunting and trapping for sustenance and sale, some birds are persecuted accidentally. Accidental bycatch in longline fisheries pressures all 22 of the world's albatross species (Diomedeidae), 17 of which are globally threatened. Birds attempt to take floating frozen bait, become hooked, and eventually drown. Spoon-billed Sandpipers (*Eurynorhynchus pygmeus*) CR wintering in Southern Asia have been trapped by hunters targeting larger species (Bird *et al.*, 2010). The Marbled Teal (*Marmaronetta angustirostris*) VU is primarily threatened by habitat loss but is also vulnerable to accidental capture in fishing nets (BirdLife Internaional, 2011). Vulture (genus *Gyps*) populations in South Asia have plummeted due to inadvertent poisoning by the veterinary drug diclofenac, with three of five of India's vulture species Critically Endangered for this reason (Shultz *et al.*, 2004). In Africa, vultures like the Lappet-faced Vulture (*Torgos tracheliotos*) VU are inadvertently poisoned by strychnine bait set for predator control (Ogada and Keesing, 2010). Other large raptor species, including the California Condor (*Gymnogyps californianus*) CR and the Spanish Imperial Eagle (*Aquila adalberti*) VU, are susceptible to inadvertent poisoning by the ingestion of lead shot. Many raptors are shot due to the belief, generally incorrect, that they pose a threat to poultry or livestock. For example, the Cuban Kite (*C. wilsonii*) CR feeds on tree snails, but it is still shot as a perceived threat to poultry (BirdLife Internaional, 2011).

Small Ranges and Large Aggregations

When birds occur in small ranges, localized disturbances can have catastrophic effects on the entire population. The entire population of Montserrat Oriole (*Icterus oberi*) CR lives essentially in the shadow of an active volcano on islands where hurricanes are likely. As noted previously, a strikingly high proportion of threatened and extinct birds occur on islands, where a small amount of habitat destruction might completely obliterate their habitat.

Similarly, species that concentrate in a small area for even part of their life cycle may be abnormally exposed to danger. For example, although they forage at sea, the entire world population of Ascension Frigatebird (*Fregata aquila*) VU breeds on a single stack smaller than a municipal parking lot. The Semipalmated Sandpiper and Red Knot depend on specific staging grounds (i.e., the Delaware Bay) on their northbound migration (*see* Biotic Threats). Even birds that are briefly drawn to individual fruiting trees are liable to suffer: In October 1997, as many as 40 Visayan Hornbills (*Aceros waldeni*) EN on a single tree were shot over the course of a single day.

Climate Change

Global climate change may already be affecting bird populations across the globe. In the Netherlands, declines in the breeding population of pied flycatchers have been tied to the earlier emergence of leaves in the spring. The flycatchers depend on abundant caterpillars to feed their young. In apparent response to warmer temperatures, the caterpillars have been emerging earlier. Although the flycatchers have been laying their eggs earlier in the spring, they have not been able to do so early enough to synchronize feeding their offspring with the peak numbers of caterpillars (Both *et al.*, 2006). In places as disparate as California and Peru, montane species are ascending to higher elevations (Tingley *et al.*, 2009; Forero-Medina *et al.*, 2011).

It is difficult, however, to definitively implicate anthropogenic climate change in the contemporary imperilment of threatened birds. For the most part, global climate change is a specter looming on the horizon. Particularly vulnerable bird groups include those that depend on pack ice (e.g., Ivory Gull (*Pagophila eburnea*) NT) or occur mostly at the highest elevations of mountain ranges (e.g., Bicknell's Thrush (*Catharus bicknelli*) VU), are vulnerable to emerging or introduced infectious diseases (e.g., the native land birds of Hawaii), or any bird restricted to a specific habitat type in a narrow geographic area. There are many examples of this latter type. For example, the Hooded Grebe (*Podiceps gallardoi*) EN of southern Argentina and Chile occurs only in a few alkaline lakes and was adversely affected by a single year of drought in 2006 (BirdLife International, 2011). The consequences of long-term climatic shifts for a bird such as this could be catastrophic.

Approaches to the Conservation of Threatened Birds

When it is possible to identify specific threats to threatened birds, it is usually possible to identify potential solutions. The implementation of these solutions, however, is often technically difficult, politically difficult, or costly. To present a portfolio of solutions, a brief discussion of ground rules for establishing priorities is warranted.

Goals and Priorities

For many conservationists, the notion of endangered species suggests that preventing extinctions should be a high priority – even the top priority – but valid, alternative goals exist. The preservation of ecosystem services, the preservation of intact ecosystems, the preservation of migration as a phenomenon of abundance, and the preservation of keystone or umbrella species – none of these goals overlaps perfectly with the prevention of extinctions. Even if the preservation of extinctions is taken as the goal, priorities may be required to determine which extinctions to focus on preventing. Should global extinctions always take precedence over local extirpations? Species over subspecies? Monotypic genera and families over species with close relatives? Need some species be triaged? Which species are sufficiently common that they do not require attention?

It is beyond the scope of this article to answer these questions, but it is worth noting that different governments and organizations have answered them in different ways. For example, the US Endangered Species Act provides regulatory protection and management resources to subspecies and vertebrate populations, in addition to full species. Under the US Endangered Species Act, the decision to designate a species, subspecies, or population as endangered is supposed to be made solely on scientific grounds, a laudable principle. In reality, however, political, economic, and other nonbiological considerations sometimes muddle the designation process (Wilcove, 2010). Political and other considerations certainly affect the subsequent efforts to save the species, subspecies, or population in question.

Problems of Definition: Scale

Because policy and legislation affecting threatened species is often formulated at the national level, the conservation of globally threatened species does not always take precedence over the conservation of regionally threatened but globally widespread species. Species may be declared nationally endangered without being globally threatened. For example, the Spotted Owl (*Strix occidentalis*) NT is an endangered species under US law, but it still misses the IUCN criteria due to its populations in Mexico. In other cases, no trend toward global endangerment is present at all. The tiny US population of Snail Kite (*Rostrhamus sociabilis*) LC is endangered in that country despite robust populations in the tropical Americas. Although the Snail Kite is not globally threatened, US endangered species legislation legitimately devotes resources to its conservation.

Managing Priorities

The bewildering array of decisions confronting conservationists has resulted in attempts to systematically identify strategies that satisfy multiple priorities simultaneously. This is especially the case for habitat conservation (*see* International Cooperation and Legislation). It stands to reason that in some cases the preservation of a threatened habitat can benefit numerous threatened bird species simultaneously. For example, it is possible to delimit regions that total less than 1.4% of the Earth's terrestrial surface, but contain more than 28% of the world's bird species (Myers *et al.*, 2000). Presumably, targeted conservation of such "biodiversity hotspots" could provide the maximum benefit for the limited resources available.

Similarly, BirdLife International identified Endemic Bird Areas (EBAs), regions with significant numbers of endemic birds (Stattersfield *et al.*, 1998). The importance of EBAs lies in the fact that if an EBA is entirely destroyed, its endemic avifauna will face global extinction. At the same time, BirdLife has set out to identify Important Bird Areas (IBAs), places that hold at least one globally threatened species, several rare species, or major concentrations of common species. IBAs identify the priority localities for conservation both within and outside hotspots and EBAs.

Site and Habitat Conservation and Management

In response to habitat loss and degradation, the commonest conservation approach is to protect and manage landscapes

for the benefit of the conservation targets. The key tool for conservation is typically the protected area. Management may occur on protected areas (where even the decision to leave the landscape as undisturbed as possible constitutes a management decision of sorts) or outside of protected areas, for example in agricultural or urban landscapes. Typically, preservation is the primary response in cases of habitat loss (e.g., conversion of tropical forests to oil palm plantations), whereas management is important in cases of habitat degradation (e.g., grasslands overgrazed by livestock). These strategies often compliment and overlap one another, as in the case of prescribed burns within state parks to manage habitat for the Florida Scrub-Jay (*Aphelocoma coerulescens*) VU. But if a region is to meet a specified agricultural production target, there may be a difficult tradeoff between employing low-impact farming over a large area and high-impact farming concentrated into a smaller area but accompanied by habitat preservation (Phalan *et al.*, 2011).

International Cooperation and Legislation

Bird conservation extends across international boundaries in three principal ways. First, the birds themselves physically cross the boundaries. The conservation of migratory species frequently depends on the participation of multiple nations and multiple polities within nations along the migratory route. A Swainson's Thrush (*Catharus ustulatus*), for example, flying from its wintering grounds in western Brazil to its breeding grounds in southern Manitoba, Canada, may pass through (or over) 10 different countries and more than 40 states, provinces, departments, and other major subnational jurisdictions (Wilcove, 2008). Second, certain threats such as international trade in threatened birds or global climate change are international in scope and must be dealt with at the international level. Third, some nations may require international assistance to implement conservation policies domestically.

Although trade is not a strong factor in the endangerment of birds in general, it is important for a few groups of species, most notably the parrots. The major international instrument for the control of trade is the Washington Convention, universally known as the Convention on International Trade in Endangered Species (CITES), which has been in operation since the 1970s. Animals and plants may be registered in three ways: on Appendix I, which essentially prohibits all movement of the species in question; on Appendix II, which prohibits all commercial trade except under license; and on Appendix III, which allows a particular nation to prohibit trade across its borders irrespective of a species' status elsewhere.

Because of the immense volume of traffic in parrots, and the problem of identification in so diverse a family (approximately 350 species), all but three species were placed on Appendix II in 1981. This move had the intention, if not the effect, of giving protection to the more threatened parrots – naturally their rarity increased their desirability among bird fanciers – which, if listed alone on the Appendix, could easily have been traded indiscriminately under other names without customs officials necessarily being able to identify them. Appendix II species have quotas set by exporting countries, in theory based on data that show the exploitation to be sustainable. Where the evidence suggests that it is not, movement to Appendix I is supposed to bestow immunity.

Curiously – but as a measure of the power that some trade interests can exert – listing on Appendix I can actually stimulate trade not only while the species is still at the proposal stage for upgrading but also following listing, on the basis of its enhanced rarity value. This happened to the Hyacinth Macaw (*Anodorhynchus hyacinthinus*) EN, whose rarity in the wild today is primarily a consequence of a trapping blitz in the 1980s. In general, CITES trade controls have some effect and allow for very useful monitoring of changes in trading fashions over time. However, for truly prized species such as Lear's Macaw (*Anodorhynchus leari*) EN and Spix's Macaw (*Cyanopsitta spixii*) CR, both from Brazil, there is little that any control system can do to eliminate smuggling: The financial inducements are simply too strong. The same tends to be true in Indonesia for exquisite songsters, such as the Straw-headed Bulbul (*Pycnonotus zeylanicus*) VU and good lookers, such as the Bali Starling (*Leucopsar rothschildi*) CR.

Other important international instruments for threatened birds are the Ramsar Convention (for important wetlands), Bonn Convention (for migratory species), and Berne Convention (for European species). The Convention on Biological Diversity ought to be a key aspect of threatened species conservation across the planet, but considerable effort by nongovernmental organizations is needed to ensure the inclusion of global priorities in the national conservation strategies that each party to the convention is obliged to produce and implement.

The United Nations' Reducing Emissions from Deforestation and Forest Degradation program (REDD and REDD+, the latter encompassing managed forests) has the potential to provide an international mechanism for the conservation and sustainable management of large amounts of forests and other carbon-rich ecosystems, while also potentially slowing the onset of global climate change, although there remain many questions about how such a program will be implemented and monitored.

Management Techniques for "Critically Endangered" Birds

Habitat Restoration

Major conservation projects almost invariably involve some habitat restoration work. However, for species at the brink of extinction through habitat loss, the emphasis falls more directly on the rapid replanting of food plants and land areas. Lear's Macaw appears to be constrained by the availability of licuri palms (*Syagrus coronata*), many stands of which show no sign of regeneration owing to cattle grazing: Programs have long been planned to establish many new groves within the species' range. The Cebu Flowerpecker survives in a few areas of heavily degraded forest that total less than 10 km^2: Efforts are under way to reforest adjacent areas within the next 30 years. There is now an emerging discipline of restoration ecology, and it is likely that a great deal of conservation energy in the twenty-first century will be

channeled into reconfiguring habitats that were ruined during the twentieth century.

Control and/or Restriction of Aliens and Natives

Eradication of alien predators and pests from islands has been pioneered in New Zealand, where several small offshore islands have been rid of various mammals both to preserve resident breeding species and to translocate stricken native species from the main islands (Clout and Craig, 1995). This painstakingly systematic restoration of islands – mainly involving the elimination of cats, rats, and mustelids – has been essential to the survival of birds, such as the Black Petrel (*Procellaria parkinsoni*) VU and Kakapo (*Strigops habroptilus*) CR, and the expertise generated is now being exported to other islands in the Pacific and Indian Oceans.

Researchers on Mauritius recently discovered the reason why the native Pink Pigeon (*Nesoenas mayeri*) EN and Mauritius Fody (*Foudia rubra*) EN have managed to persist in the face of rat predation. These birds now nest almost exclusively in a grove of exotic *Cryptomeria japonica*, whose bark produces a sticky gum and whose leaves consist of spiny needles, both of which discourage rats from getting at nests. Curiously, then, in rare cases, it appears that exotic vegetation can help rather than hinder native species in their struggle against exotic predators; in this instance, the planting of more *Cryptomeria* is clearly called for.

Captive Breeding

Captive breeding has several major drawbacks because even when the breeding program is successful, problems of imprinting, migratory patterns, and techniques for survival in the wild plague reintroduction attempts. An additional serious problem is the possibility of disease transmission to wild birds from captive stock (Snyder *et al.*, 1996). Nevertheless, captive breeding has achieved several outstanding successes. The Northern Bald Ibis (*Geronticus eremita*) **CR** and California Condor both flourish in captivity while efforts to improve environmental conditions in the wild continue. The Lord Howe Woodhen (*Gallirallus sylvestris*) EN and Guam Rail (*Gallirallus owstoni*) EW, members of a notoriously vulnerable family (because so many rails reached islands in the past and proceeded to lose the power of flight), only persist thanks to *ex situ* regimes. The Socorro Dove (*Zenaida graysoni*) EW survives only because a few Californian bird fanciers took and bred specimens at a time before its native island was overrun by exotic predators. In 1998 the Bali Starling, was within eight individuals of becoming extinct in the wild owing to relentless poaching for trade, but zoos throughout the world have bred the species such that many hundreds and perhaps thousands survive for possible reintroduction when poaching has finally been brought under control.

There is also a role for veterinarians in intensive *in situ* management programs, through their expertise in providing appropriate advice on toxicity of nest box materials, control of nest parasites, composition and secure provision of supplementary foods, disease screening, and minimization in cross-fostering and translocation exercises.

Reintroduction and Translocation

The extirpation of a species at a discrete site may have been caused by a short-term or remediable factor, so that restocking with individuals from elsewhere can be undertaken. Such an endeavor requires careful planning (Black, 1991). A series of feasibility assessments should determine the site's continuing ecological suitability (the original constraining factors must no longer operate); the threats it faces; the availability of appropriate stock; the socioeconomic implications; and local, national, and international awareness needs. The birds must be in optimal condition; and if captive-bred, they should have been reared in disease-free conditions and in such a way as to be behaviorally and genetically fit for independence in nature. Monitoring of the project is crucial, and it should be documented for circulation to other biologists contemplating such efforts.

Bird species that have benefited from well-designed reintroduction programs include the Peregrine Falcon (*Falco peregrinus*) LC and Chatham Islands Snipe (*Coenocorypha pusilla*) VU. Those that were not included are Cheer Pheasant (*Catreus wallichii*) VU in Margalla Hills, Pakistan (fox predation) and Thick-billed Parrots in the USA (the wild-caught birds flew back to Mexico; the captive-bred ones did not move very far and were unusually vulnerable to predation by raptors).

Translocation to previously unoccupied sites has become a greatly valued technique, if only as a temporary measure while efforts are concentrated on habitat restoration or predator eradication at the native site. Beneficiaries of this approach include Tongan Megapode (*Megapodius pritchardii*) EN, *Vini* lorikeets, Kakapo, Seychelles Magpie-Robin (*Copsychus sechellarum*) EN, and Seychelles Warbler (*Acrocephalus sechellensis*) VU. The work done in preparation for the Seychelles Warbler translocations showed that target islands had many times the insect abundance of the host island, Cousin. This was clearly related to the absence of predation pressure on the insect fauna, and – unsurprisingly but still very strikingly – when birds were released onto the target islands, they began breeding almost immediately and very rapidly expanded their numbers. On Cousin Island, the birds had been so packed that they bred only very slowly, with one offspring tending to stay on territory and help at the nest, waiting for a parent to die (Komdeur *et al.*, 1997). This observation underscores the value of the food resource base in managing threatened birds and leads to the next point.

Supplementary Feeding

The value of increasing food availability for species whose populations need rapid growth seems to have been perceived only relatively recently, perhaps beginning with work on Peregrine Falcons (see Temple, 1978). The technique has been adopted for threatened birds on Mauritius, principally with the Mauritius Kestrel (*Falco punctatus*) VU, for which extra food clearly enhanced reproductive output (Jones *et al.*, 1995). New Zealand workers have found the same with the Kakapo: Normally the species breeds only in response to major masting events, which may occur once every 5 years, but with dietary supplements, it appears to be capable of breeding every year. On Hokkaido, Red-crowned Cranes (*Grus japonensis*) EN have recovered from near extirpation through a combination

of better protection and extensive food provision. Clearly, as a relatively short-term measure, supplementary feeding can be crucial in producing an unnaturally rapid reproductive output, a very desirable effect in critically low populations where the genetic value of every individual needs to be maximized.

Nest-Site Provision or Enhancement

Where nest sites are limiting (often the case with hole-nesting species), the provision of nest boxes or the enhancement of natural cavities may be appropriate. Puerto Rican Amazons ignored the former but benefited from the latter. Mauritius Kestrels greatly benefited from the creation of nesting ledges. The provision of grilles at burrow entrances has helped prevent White-tailed Tropicbirds from appropriating Bermuda Petrel burrows (*see* Natives Threatened by Other Natives).

Cross-Fostering and Cross-Breeding

Many large raptors and most cranes lay two eggs but normally rear only one young, so biologists attempting to increase productivity of threatened forms of such birds have long sought to make use of the expendable second eggs by taking them for hatching and captive breeding. The most famous instance of cross-fostering involved the Black Robin (*Petroica traversi*) EN, whose population fell to five individuals in 1980 but recovered by placing some eggs in the nests of Tomtits (*Petroica macrocephala*). In the case of the Whooping Crane (*Grus americana*) EN, an attempt was made to establish a second population by placing such eggs in the nests of Sandhill Cranes (*Grus canadensis*) LC, but this did not result in a breeding population, possibly as a result of imprinting.

Cross-breeding with another subspecies was regarded as the last hope of the Dusky Seaside Sparrow (*Ammospiza maritimus nigrescens*), but the last stock died out before the endeavor could succeed. However, it has been used on the endemic Norfolk Island race *undulata* of the Southern Boobook (*Ninox novaeseelandiae*), which by 1986 had been reduced to a single female. After nest boxes were erected to overcome an immediate shortage in nesting sites, two male nominate male boobooks were introduced to the island; one of them paired with the female and in due course produced a string of hybrid offspring. The taxonomic (and legal) implications of this technique pose complexities, but it clearly represents a means of preserving genetic diversity.

Caveats

Although the ability of conservation professionals to rescue threatened birds via hands-on techniques, such as translocation, food supplementation, and cross-fostering is truly impressive, it is emphasized that the use of such techniques is feasible for only a small number of the world's threatened birds, principally those species residing in wealthier countries. The future for most of the world's threatened birds will hinge on whether or not their remaining habitats can be protected and properly managed.

See also: Birds, Biodiversity of. Conservation Efforts, Contemporary. Endangered Amphibians. Endangered Mammals. Endangered Reptiles. Modern Examples of Extinctions

Appendix

List of Courses

1. Conservation Biology
2. Conservation Policy
3. Field Ornithology

References

Achard F, Eva HD, Stibig HJ, *et al.* (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297: 999–1002.

Areta JI, Bodrati A, and Cockle K (2009) Specialization on Guadua bamboo seeds by three bird species in the Atlantic Forest of Argentina. *Biotropica* 41: 66–73.

Bird JP, Lees AC, Chowdhury AU, Martin R, and Ul Haque E (2010) A survey of the Critically Endangered Spoon-billed Sandpiper *Eurynorhynchus pygmeus* in Bangladesh and key future research and conservation recommendations. *Forktail* 26: 1–8.

BirdLife Internaional (2000) *Threatened Birds of the World.* Barcelona, Spain and Cambridge, UK: Lynx Edicions and BirdLife International.

BirdLife International (2004) *Birds in Europe: Population Estimates, Trends and Conservation Status.* Cambridge, UK: BirdLife International.

BirdLife International (2011) *IUCN Red List for Birds.* Cambridge, UK: BirdLife International.

Black JM (1991) Reintroduction and restocking: Guidelines for bird recovery programmes. *Bird Conservation International* 1: 329–334.

Blackburn T, Cassey P, Duncan R, Evans K, and Gaston K (2004) Avian extinctions and mammalian introductions on oceanic islands. *Science* 305: 1955–1958.

Both C, Bouwhuis S, Lessells CM, and Visser ME (2006) Climate change and population declines in a long-distance migratory bird. *Nature* 441: 81–83.

Bucher EH (1992) The causes of extinction of the passenger pigeon. In: Power DM (ed.) *Current Ornithology,* vol. 9, pp. 1–36. New York, NY: Plenum Press.

Bucher G, Niven D, Panjabi A, Pashley D, and Rosenberg K (2007) The 2007 WatchList for United States birds. *American Birds* 107: 18–25.

Butchart S, Collar N, Stattersfield A, and Bennun L (2010) Conservation of the world's birds. In: del Hoyo J, Elliott A, and Christie D (eds.) *Handbook to the Birds of The World,* vol. 15. Barcelona, Spain: Lynx Ediciones.

Butchart S, Stattersfield A, Bennun, *et al.* (2004) Measuring global trends in the status of biodiversity: Red List Indicies for birds. *Public Library of Science Biology* 2: e383.

Clout MN and Craig JL (1995) The conservation of the critically endangered flightless birds in New Zealand. *Ibis* 137: S181–S190.

Collar NJ, Crosby MJ, and Stattersfield AJ (1994) *Birds to Watch 2. The World List of Threatened Birds.* Norwich, UK: Page Bros/BirdLife International.

Collar NJ, Wege DC, and Long AJ (1997) Patterns and causes of endangerment in the New World avifauna. *Ornithological Monographs* 48: 237–260.

Dutson G, Magsalay P, and Timmins R (1993) The rediscovery of the Cebu Flowerpecker *Dicaeum quadricolor*, with notes on other forest birds on Cebu, Philippines. *Bird Conservation International* 3: 235–243.

Eaton MA, Brown AF, Noble DG, *et al.* (2009) Birds of conservation concern 3: The population status of birds in the United Kingdom, Channel Islands and the Isle of Man. *British Birds* 102: 296–341.

Estes J, Terborgh J, Brashares J, *et al.* (2011) Trophic downgrading of planet earth. *Science* 333: 301–306.

Ferraz G, Nichols J, Hines J, *et al.* (2007) A large-scale deforestation experiment: Effects of patch area and isolation on Amazon birds. *Science* 315: 238–241.

Fitzpatrick J, Lammertink M, Luneau Jr M, *et al.* (2005) Ivory-billed Woodpecker (*Campephilus principalis*) persists in continental North America. *Science* 308: 1460–1462.

Fjeldsa J (1984) Three endangered South American grebes (*Podiceps*): Case histories and ethics of saving species by human intervention. *Annals Zoologici Fennici* 21: 411–416.

Forero-Medina G, Terborgh J, Socolar SJ, and Pimm SL (2011) Elevational ranges of birds on a tropical montane gradient lag behind warming temperatures. *Public Library of Science ONE* 6: e28535.

IUCN (2001) *IUCN Red List Categories and Criteria: Version 3.1.* IUCN Species Survival Commission (eds.) http://www.iucnredlist.org

IUCN (2011) *IUCN Red List of Threatened Species: Version 2011.2*. http://www.iucnredlist.org (accessed 10 November 2011)
Jones CG, Heck W, Lewis R, *et al.* (1995) The restoration of the Mauritius Kestrel Falco punctatus population. *Ibis* 137: S173–S180.
Joseph L (1988) A review of the conservation status of Australian parrots in 1987. *Biological Conservation* 46: 261–280.
Kareiva P (2011) Apex consumers: Holy grail for conservation? *Science Chronicles Best of 2011* 3–5.
Komdeur J, Daan S, Tinbergen J, and Mateman C (1997) Extreme adaptive modification in sex ratio of the Seychelles Warbler's eggs. *Nature* 385: 522–525.
Ligon JD, Stacey PB, Conner RN, Bock CE, and Adkisson CS (1986) Report of the American Ornithologists' Union Committee for the Conservation of the Red-Cockaded Woodpecker. *The Auk* 103: 848–855.
Myers N, Mittermeier R, Mittermeier C, Fonesca G, and Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
Moors PJ, Atkinson IAE, and Sherley GH (1992) Reducing the rat threat to island birds. *Bird Conservation International* 2: 93–114.
Ogada DL and Keesing F (2010) Decline of raptors over a three-year period in Laikipia, Central Kenya. *Journal of Raptor Research* 44: 129–135.
Panjabi A, Dunn EH, Blancher PJ, *et al.* (2005) *The Partners in Flight handbook on species assessment*. Version 2005. Partners in Flight Technical Series No. 3. Rocky Mountain Bird Observatory website: http://www.rmbo.org/pubs/downloads/Handbook2005.pdf
Phalan B, Onial M, Balmford A, and Green R (2011) Reconciling food production and biodiversity conservation: Land sharing and land sparing compared. *Science* 333: 1289–1291.
Powell GVN and Bjork RD (1994) Implications of altitudinal migration for conservation strategies to protect tropical biodiversity: A case study of the Resplendent Quetzal *Pharomacrus mocinno* at Monteverde, Costa Rica. *Bird Conservation International* 4: 161–174.
Rappole JH, King DI, and Diez J (2003) Winter- vs. breeding-habitat limitation for an endangered avian migrant. *Ecological Applications* 13: 735–742.
Ricketts T, Dinerstein E, Boucher T, *et al.* (2005) Pinpointing and preventing imminent extinctions. *Proceedings of the National Academy of Sciences of the United States of America* 102: 18497–18501.
Robinson SK and Wilcove DS (1994) Forest fragmentation in the temperate zone and its effects on migratory songbirds. *Bird Conservation International* 4: 233–249.
Savidge J (1987) Extinction of an island forest avifauna by an introduced snake. *Ecology* 68: 660–668.
Scott J, Goble DD, Wiens JA, *et al.* (2005) Recovery of imperiled species under the Endangered Species Act: The need for a new approach. *Frontiers in Ecology and the Environment* 3: 383–389.
Shultz S, Baral HS, Charman S, *et al.* (2004) Diclofenac poisoning is widespread in declining vulture populations across the Indian subcontinent. *Proceedings of the Royal Society of London Series B* 271: S458–S460.
Sibley D, Bevier L, Patten M, and Elphick C (2007) Ivory-billed or Pileated Woodpecker? *Science* 315: 1495–1496.
Snyder NFR, Derrickson SR, Beissinger SR, *et al.* (1996) Limitations of captive breeding in endangered species recovery. *Conservation Biology* 10: 338–348.
Soulé ME, Bolger DT, Alberts AC, *et al.* (1988) Reconstructed dynamics of rapid extinctions of chaparral-requiring birds in urban habitat islands. *Conservation Biology* 2: 75–92.
Stattersfield A, Crosby M, Long A, and Wege D (1998) *Endemic Bird Areas of the World: Priorities for Biodiversity Conservation*. Cambridge, UK: BirdLife International.
Steadman D (2006) *Extinction and Biogeography of Tropical Pacific Birds*. Chicago, IL: University of Chicago Press.
Stouffer PC and Bierregaard RO (1995) Use of Amazonian forest fragments by understory insectivorous birds. *Ecology* 76: 2429–2445.
Temple SA (ed.) (1978) *Endangered Birds: Management Techniques for Preserving Threatened Species*. Madison, WI: University of Wisconsin Press.
Tilman D, Fargione J, Wolff B, *et al.* (2001) Forecasting agriculturally driven global environmental change. *Science* 292: 281–284.
Tingley MW, Monahan WB, Beissinger SR, and Moritz C (2009) Birds track their Grinnellian niche through a century of climate change. *Proceedings of the National Academy of Sciences of the United States of America* 106: 19637–19643.
van Oosten H (2010) *25th International Ornithological Congress*. Why Northern Wheatears (Oenanthe oenanthe) do not Return After Restoration of Coastal Dune Grasslands in the Netherlands. http://www.acquaviva.com.br/ioc2010/
van Ripper III C, van Riper S, Goff M, and Laird M (1986) The epizootiology and ecological significance of malaria in Hawaiian land birds. *Ecological Monographs* 56: 327–344.
Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of Earth's ecosystems. *Science* 277: 494–499.
Wilcove DS (1985) Nest predation in forest tracts and the decline of migratory songbirds. *Ecology* 66: 1211–1214.
Wilcove DS (2008) *No Way Home: The Decline of the World's Great Animal Migrations*. Washington, DC: Island Press.
Wilcove DS (2010) Endangered species management: The US experience. In: Sohdi NS and Erlich PR (eds.) *Conservation Biology for All*, pp. 220–235. Oxford, UK: Oxford University Press.
Wilson KA, Carwardine J, and Possingham HP (2009) Setting conservation priorities. *Annals of the New York Academy of Sciences* 1162: 237–264.

Threatened Species: Classification Systems and Their Applications

Rebecca M Miller, IUCN Red List Unit, Cambridge, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biodiversity (from biological diversity) The variety of all life on Earth. Includes genetic diversity (diversity within species), organismal diversity (diversity between species), and the diversity of ecosystems and ecological processes.

Ecosystem services Benefits that humans derive from the resources and processes supplied by natural ecosystems (e.g., water quality, climate regulation; pollination; food, fuel, and raw materials; nutrient cycling).

Endangered Often interpreted generically to mean having a relatively high risk of extinction (*see* Threatened). In many systems that classify extinction risk, a specific category that indicates a particular degree of extinction risk (e.g., in the IUCN Red List Categories and Criteria, Endangered falls between Vulnerable and Critically Endangered; on the US Endangered Species List, Endangered is the highest category of risk, followed by Threatened).

Endemic species Species found only within a specific geographic area on Earth (e.g., a habitat type or region) and nowhere else.

Extinction risk The likelihood that a taxon will disappear or die out within a given area (e.g., one country or the entire world) and over a definable time period.

Taxon (plural taxa) Taxonomic group or entity at any level such as a genus, species, subspecies, or population.

Threatened Facing a relatively high risk of extinction in a definable time period. May apply to any taxonomic (e.g., genus, species, subspecies, variety) or geographic (e.g., population, subpopulation) rank. Often used interchangeably with endangered (though see the definition of endangered). In some systems that classify extinction risk, "Threatened" is a category indicating a specific level of extinction risk.

Threatened species list (also often called a red list or red data book) Publication defining the extinction risk of assessed taxa within a given region. In its most minimal form, just a list of names and categories of threat, but frequently also includes additional supporting information about the biology, ecology, threats, and conservation of the taxa assessed.

Introduction

There is no doubt that the world is facing a biodiversity crisis of epic proportions. The rate of species loss is greater now than at any time in human history, with extinctions occurring at rates hundreds of times higher than historical extinction rates. The question at the forefront of the conservation debate today is, What can we do about it?

To answer this fundamental question requires information about the current status and trends of biodiversity. We need to know what is being lost, where it is disappearing most quickly, what is causing the declines, and, perhaps most important, what measures must be taken to protect and restore species and their habitats. As evidence accumulates that the loss of biodiversity will have major consequences for ecosystem functions and services, and therefore for human health and well-being (Millennium Ecosystem Assessment, 2005), the arguments for finding solutions to these questions become even more compelling.

Inventories or lists that aim to quantify species extinction risk can provide insight into many of these pressing issues. The principle aim of a threatened species assessment is to estimate a species' risk of extinction in a comparable, repeatable, transparent, and objective manner. Threatened species lists, in turn, have multiple conservation applications. They permit us to quantify the magnitude of the contemporary extinction crisis so as to monitor the status of biodiversity and measure current trends; they draw attention to the plight of threatened species and help mobilize political and public support for conservation measures; and they help guide conservation efforts, allowing us to optimize investment of the limited conservation resources available by taking action where it is most needed.

These lists have strong scientific, political, and public relevance. While conservationists debate how many species are threatened, where they are found, how quickly they are moving toward extinction, and what to do to protect them, decision makers are looking for guidance on how best to set priorities to protect species, and the public seeks information about why other living creatures are disappearing, what our role is in the process, and what they as individuals can do to help.

In truth, extinction encompasses more than just species loss: global species extinction results from the progressive loss of individuals and populations, causing a decline in genetic diversity and ultimately leading to reduced ecosystem function as key species drop out of the system. Biodiversity loss could be measured on any number of alternative scales such as biomes, habitats, populations, or genes. But many consider species the fundamental unit of biodiversity (Wilson, 1992), and species are the most common and easily measured unit for assessing the state of biodiversity; moreover, threatened species are among the most visible and easily understood symbols of the rising tide of extinctions (Wilcove, 2010), making them an emotive and politically powerful measurement of biodiversity loss.

This chapter explores the history and current status of threatened species assessment, including a brief outline of extinction vulnerability, a summary of the various methods employed to assess the threat status of species, and a more

detailed look at what these assessments can tell us about the status of biodiversity today and how they are used in conservation planning. It does not attempt to be comprehensive, as many of these concepts are explored in more detail in other chapters, but rather places the concept of threatened species assessment in context within the wider world of conservation planning, prioritization, and action.

History of Threatened Species Lists

The first calls of alarm at rapid species loss were raised more than a century ago. By the early 1900s, several works had been published highlighting recent extinctions in British animals (Harting, 1880; Hudson, 1894) and declining and extinct wildlife around the world (e.g., Anon, 1928; Dollman, 1937; Johnson, 1937). In 1913, William Temple Hornaday warned of the imminent extinction of American game species and drew support for legislation and legal reforms to protect wildlife (Hornaday, 1913). The path toward using threatened species lists to advocate legislative protection measures had been opened.

In the 1940s and 1950s, threatened species publications expanded to include information on the factors contributing to the depletion of the world's mammalian and avian faunas (e.g., Allen, 1942; Harper, 1945; Greenway, 1958). Following this lead, the International Union for Conservation of Nature (IUCN) spearheaded the concept of Red Data Books (RDB) as a "register of threatened wildlife that includes definitions of degrees of threat" (Fitter and Fitter, 1987). These publications were used widely in scientific, political, and popular contexts to highlight the world's most threatened species (Mace *et al.*, 2008), cementing IUCN's role as a leader in global threatened species assessment.

Early Red List assessments, lacking any standardized criteria, relied on the experience and opinion of experts. Until the late 1960s, a variety of qualitative terms and categories were used to define different levels of extinction risk. Rather than focusing solely on probability of extinction, these initial threat categories were a mixture of several issues, including likelihood of extinction, causes of threat, severity of threat, the nature of population vulnerability, and priority for conservation action. In addition, they depended on subjective interpretations rather than objective criteria and were thus open to biases of individual preference or political influence (Mace *et al.*, 2008).

Seeking a more rigorous, objective, and repeatable system, and in view of the increasing incorporation of threatened species lists and categories into international, national, and local legislation that mandated or prohibited certain actions, IUCN initiated a review of the Red List categories in 1984 (Fitter and Fitter, 1987; Mace *et al.*, 2008). An extensive period of development, consultation and testing across a wide range of taxonomic groups resulted in a revised risk-ranking system of data-driven quantitative criteria (Mace and Lande, 1991; IUCN, 1994), with which several threatened species lists were produced (Baillie and Groombridge, 1996; Oldfield *et al.*, 1998). The categories and criteria were subsequently finalized in 2001 (IUCN, 2001).

Many other systems for assessing extinction risk were developed from or in parallel to the IUCN system. At least 25 categorization systems are used across the North American continent alone (de Grammont and Cuarón, 2006), giving an indication of the total number of different categorization systems that must exist worldwide. Many of these systems suffered the same dilemma as the early IUCN system in mixing the concept of extinction risk with priority for conservation (see **Box 1**); although this has been remedied in some systems, many others still conflate the two.

Inspired by the IUCN RDB and the expanding global awareness of species declines, national RDB began to appear in the late 1960s. The US federal government produced a publication on rare and endangered wildlife in 1966; the first RDB for the former USSR was published in 1978 (Scott *et al.*, 1987). In 1968, the Brazilian government published its first Official List of Threatened Plant and Animal Species, using the IUCN categories in place at the time. There has since been a virtual explosion of national lists of threatened and endangered species; 3562 current and historical threatened species lists have been reported for the European nations alone (Köppel *et al.*, 2003). These lists have become important tools for conservation and today play a major role in informing and influencing national legislation, international conventions, conservation planning and resource allocation, and scientific research.

Vulnerability to Extinction

The assemblage of species that exist at any one time in history is a function of two processes: extinction of existing species and the evolution of new species. Indeed, of all the species that have ever lived, only 2% to 4% are alive today (May *et al.*, 1995). However, conservative estimates indicate that the current rate of extinction is 100–1000 times higher than historical rates (Baillie *et al.*, 2004). Other estimates suggest that current extinction rates may be 1000–10,000 times higher than historical rates of extinction (Pimm *et al.*, 1995; Baillie *et al.*, 2004). We are thus now approaching what many scientists believe to be the sixth great extinction event (Wilson, 1992; Myers, 1993; Barnosky *et al.*, 2011), which could ultimately result in the loss of a large proportion of the biodiversity on the planet today. Although the ultimate fate of all species is extinction, by definition threatened species face a relatively higher risk of extinction in the near future than nonthreatened species. What then makes a species vulnerable to extinction?

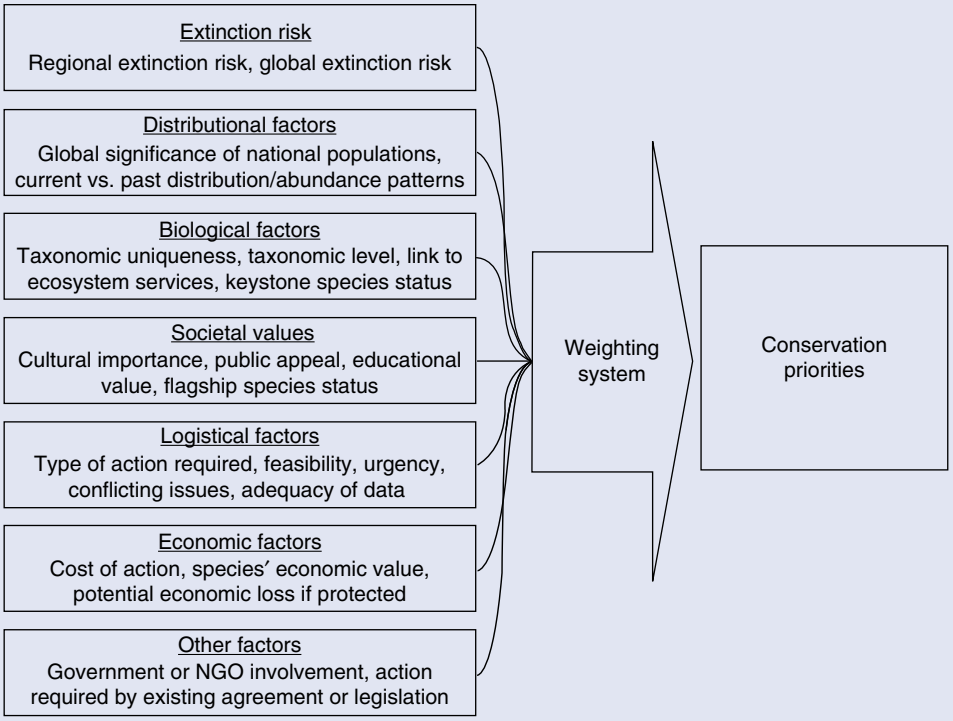
Extinction risk is a function of intrinsic characteristics and extrinsic factors and the complex manner in which these various biological traits and threatening processes interact (Mace *et al.*, 2008). Intrinsic traits demonstrated to be associated with a high risk of extinction include high trophic level, low population density, slow life history or low fecundity, high habitat specificity, and small geographical range size (Bennett and Owens, 1997; Purvis *et al.*, 2000). Species with these traits are more sensitive to environmental instability and are less able to adapt to or recover from environmental or ecological changes.

Population size and the ability of a population to recover from natural or anthropogenically induced declines are

Box 1

Assessing extinction risk versus setting conservation priorities

Extinction risk is an important factor to consider when deciding whether or not to devote scarce attention and resources to the conservation of a particular species, but it is certainly not the only one (**Box 1 Figure**). Indeed, a high risk of extinction does not imply a high priority for action. The ultimate success of conservation efforts often depends more on societal and economic factors than biological ones, and when it comes to deciding how best to allocate the scarce resources available, difficult decisions must be made. Fundamental questions that must be answered include but are not limited to: What needs to be done to conserve the species? Do we have the required knowledge and expertise? Are sufficient funds available? What are the prospects for successful intervention? What economic and cultural importance does the species have? Is it taxonomically, genetically, or ecologically unique? What are the conditions within the region? What is its global status? What proportion of the global population is found within the region? Although extinction risk assessment should be as objective as possible, ultimately the setting of conservation priorities involves value judgments and is therefore part of a broader sociocultural and political debate. Even when incorporating subjective values into a flexible and adaptable weighting system, however, effective priority-setting mechanisms should always be explicit and transparent.



Box 1 Figure Factors to consider when establishing conservation priorities (the factors shown are illustrative, not exhaustive). The variables considered will depend on the purpose of the priority-setting exercise and conditions within the region. Different factors may be summed, multiplied, or otherwise weighted to achieve an ultimate ranking of conservation priorities. Reprinted from Miller RM, Rodríguez JP, Aniskowicz-Fowler T, *et al.* (2007) National threatened species listing based on IUCN criteria and regional guidelines: Current status and future perspectives. *Conservation Biology* 21: 684–696.

Despite this reality, national threatened species lists often move beyond assessing extinction risk to incorporate conservation priority setting or degree of protection afforded to the species into the listing criteria. Even when the assessments reflect extinction risk only, many countries interpret these lists as a direct priority for conservation action by incorporating the assessments directly into national legislation. Some countries have resolved this issue by producing at minimum two separate lists: (1) a scientific list of nationally threatened species that is produced using objective and explicit criteria designed purely to assess risk of extinction and (2) a list of species that are priorities for receiving conservation funding and attention, which is often produced using a weighting system and incorporating a number factors in addition to threat status. Numerous other methods that consider extinction risk alongside other factors have been proposed for prioritizing species for conservation action (e.g., Avery *et al.*, 1995; Dunn *et al.*, 1999; Bollmann *et al.*, 2002; Rodríguez *et al.*, 2004, de Grammont and Cuarón, 2006; Miller *et al.*, 2006). It is important to make this clear distinction to prevent blurring the line between extinction risk assessment, which is a scientific endeavor, and conservation priority setting, which is very much a value-based judgment.

References: Mace and Lande, 1991; Avery *et al.*, 1995; Collar, 1996; Dunn *et al.*, 1999; Gärdenfors, 2001; Gärdenfors *et al.*, 2001; IUCN, 2001; Bollmann *et al.*, 2002; Rodríguez *et al.*, 2004; de Grammont and Cuarón, 2006; Miller *et al.*, 2006; Mace *et al.*, 2007; Miller *et al.*, 2007; Hoffmann *et al.*, 2008; Brito *et al.*, 2010.

particularly important in determining extinction risk (Purvis *et al.*, 2000). "Rare" species, those with small populations or geographic ranges (or both), are far more likely to go extinct than widespread or higher-density species, as are species with fluctuating populations and those suffering rapid declines (Pimm and Jenkins, 2010). However, rarity does not in itself cause extinction. It is when those small populations are confronted with novel or altered environmental conditions or

threatening processes and find it difficult to adapt that they begin to slide toward extinction.

The manner in which different threat processes interact with different biological traits has a significant impact on extinction risk. For example, persecution and introduced predators have been correlated with an increase in extinction risk in birds with large body size and long generation times, whereas habitat loss has been correlated with an increase in extinction risk in small-bodied birds that have specialized habitat requirements (Owens and Bennett, 2000). Carnivores within reserves are more likely to go extinct if they have large home ranges (Woodroffe and Ginsberg, 1998).

Extinction risk can have as much or more to do with the driving threat factors as with intrinsic biological traits. Regardless of the size of a species' population, how quickly it breeds, or the extent of its current range, if the pressures put on it are too great and remedial measures are not put in place quickly enough, the species may not be able to recover. This has been witnessed in species such as the passenger pigeon *Ectopistes migratorius*, which went from being one of the most numerous birds on the planet to being globally extinct in less than 100 years as a result of overhunting and habitat loss.

Classification Systems for Assessing Extinction Risk

The finality of extinction makes it a very powerful motivating force for developing conservation actions to prevent its occurrence, but before we can determine what should be done to protect at-risk species, those species most at risk of extinction must be identified.

Many different systems have been developed to assess extinction risk, each with differing aims and objectives, using a range of demographic and ecological parameters (e.g., Millsap *et al.*, 1990; Master, 1991; Lunney *et al.*, 1996; Cofré and Marquet, 1999; IUCN, 2001; Regan *et al.*, 2004). Some systems are qualitative and strongly based on expert judgment, where threat level is generally inferred from information such as population size, geographic range, threat to the species, or whether the species already benefits from conservation actions. Other systems, however, use quantitative methods that rely on specific rules or point-scoring systems to assign categories of threat. Many systems use a mixture of both quantitative and qualitative attributes. Likewise, some systems are designed purely to evaluate risk of extinction, whereas others focus on ranking species to receive priority conservation attention, and many purport to do both.

This variation is particularly notable at the regional and national levels. Many countries have designed their own classification systems to meet their specific needs by incorporating the parameters they deem most relevant; these needs are sometimes defined as much by legislative and political constraints as by ecological considerations. The resulting lists may reflect extinction risk, rarity, cultural importance, conservation value, population decline, conservation priorities, international responsibility for protection, or a combination of several of these (Miller *et al.*, 2007), yet they are often generally called threatened or endangered species lists. Furthermore, the definitions of categories on these various lists are rarely consistent, so commonly used terms such as "threatened" and "endangered" sometimes mean generally "at risk of extinction" and sometimes imply a specific level of extinction risk; the precise level of threat for any given term will differ from list to list (de Grammont and Cuarón, 2006). Adding even more confusion, the criteria used are often neither explicit nor transparent and can be significantly subjective (de Grammont and Cuarón, 2006; Feria Arroyo *et al.*, 2009).

Unavoidable though it may be, having multiple classification schemes is inconvenient and even problematic, as assessments made with different methodologies are inherently difficult to compare and are often incompatible (de Grammont and Cuarón, 2006; Miller *et al.*, 2007). This makes understanding and interpreting lists from different countries impossible, which in turn can hamper efforts to consolidate data across regions and impedes large-scale species protection (de Iongh *et al.*, 2003; Miller *et al.*, 2007; Zamin *et al.*, 2010).

Although different classification systems often reflect different objectives, some systems have been found to yield similar threat rankings when applied to the same taxa (O'Grady *et al.*, 2004; Feria Arroyo *et al.*, 2009). However, much variability in the category assigned to a given species still exists, both between systems and within them. Achieving consistency may be highly dependent on who is doing the assessments, as different assessors often come to different conclusions (Regan *et al.*, 2005).

Disagreements between systems can lead to a lack of confidence in species assessments, but many of these differences may be justified as each system has been designed to emphasize different attributes, depending on the purpose for which it was created (O'Grady *et al.*, 2004). It is therefore important to remember the intention of the list when interpreting its results. Nevertheless, species assessments are more likely to be reliable if they are carried out by groups of experts working together, using standardized and transparent protocols, as discussion and peer review are critical to reducing inconsistencies (Regan *et al.*, 2005). Given the immense uncertainty in the population and distribution data available for most species, as well as the important effects that varying attitudes toward uncertainty can have on assessments (Akçakaya *et al.*, 2000), some variability is inevitable. However, using a standard system may facilitate understanding of results and data comparisons across regions (Zamin *et al.*, 2010).

Assessment of Global Species Extinction Risk

The most influential and widely used methodology for evaluating global extinction risk – the IUCN Red List Categories and Criteria (IUCN, 2001) – has been in use and undergone continuous refinement for five decades, evolving from a system based primarily on subjective expert opinion to a more rigorous rule-based system (Mace *et al.*, 2008). The system is designed to reflect the probability of a taxon (most often species, but occasionally subspecies or populations) becoming extinct over a given time period. It aims to provide a standardized, consistent, and transparent method for assessing extinction risk, thereby increasing the objectivity and scientific credibility of the assessments.

Over the past several decades, the IUCN Red List of Threatened Species has developed from a listing of species

assessed a priori because of their perceived threatened status to comprehensive or representative assessments of entire taxonomic groups, which include a range of detailed assessment data (Rodrigues *et al.*, 2006). This strong history combined with the thousands of scientists and experts around the world that comprise the IUCN Species Survival Commission network has given the system a unique credibility and recognition worldwide. Today the IUCN Red List is considered one of the most authoritative sources of information on the global conservation status of plants and animals (Lameroux *et al.*, 2003; de Grammont and Cuarón, 2006; Rodrigues *et al.*, 2006).

The IUCN Red List Categories and Criteria classify species by extinction risk, highlighting those species that are most likely to become extinct in the near future given current knowledge of population trends, range and habitat availability, population size and structure, and recent, current, or projected threats acting on the species (see **Box 2**). The system consists of one set of criteria that are applicable to all species and measure the symptoms of endangerment, not the causes. For example, a species is listed under criterion A based on the rate of population decline over a specific time period, regardless of the causes of the decline. Species can therefore be listed as soon as they display signs of imperilment, even if the cause of threat is unknown. The system is designed to be precautionary but realistic, rather than evidentiary (IUCN SPSC, 2010).

To make listings objective and transparent, extensive documentation is provided to justify the assessment. Each assessed species is cataloged with available information on population size and trends, distribution, ecology and habitat preferences, use and trade information, ecosystem services, threats, and recommended conservation actions. This documentation can be used to inform conservation and policy decisions that involve human activities, development, and business; guide species conservation efforts; direct scientific research; and generate public awareness.

Of course, no assessment system is perfect. In developing a comprehensive methodology applicable to all taxonomic groups across all geographic scales, compromises had to be made (Mace *et al.*, 2008). The system has been criticized for several reasons, including requiring information that is nonexistent or insufficient for many species, incorporating subjective terms that can be misinterpreted and misused, measuring declines over unobtainable or unrealistic time frames (especially for long-lived species), grouping species into broad categories that mask widely differing extinction risk, and having unrealistic threshold values for certain taxa (e.g., Pimenta *et al.*, 2005; Robbirt *et al.*, 2006; Freeman, 2008; Godfrey and Godley, 2008; but see de Grammont and Cuarón, 2006; Miller *et al.*, 2007; Mace *et al.*, 2008). Some analyses have shown that listings tend to be overly precautionary (Heard *et al.*, 2011), and some subjectivity in assessments remains due to data uncertainty and semantic interpretations (Akçakaya *et al.*, 2000). Assessments also come under scrutiny when species of political or economic importance are listed as threatened and when funding opportunities are linked to threatened status (Mace *et al.*, 2008; Gillespie *et al.*, 2011), though this is true of all classification schemes. These complexities notwithstanding, the IUCN system remains the most widely accepted and cited system in use today and has had a significant influence on biodiversity conservation worldwide.

Assessment of Regional and National Extinction Risk

When it comes to species conservation, the impact of global extinction is clearly more significant than regional extinction, for as long as the species is still present elsewhere, the possibility for recovery always exists. However, loss of local biodiversity can be of major importance to the country or region from which the species disappeared. Moreover, national and regional risk assessments can serve as early warning signs of local impacts, triggering conservation actions that can possibly stave off more widespread declines (Miller *et al.*, 2007), and they may also contribute to informing global conservation efforts and improving our understanding of patterns of global biodiversity loss, especially when the information they contain is incorporated into global conservation efforts like the IUCN Red List (Rodríguez *et al.*, 2000, 2006; Gärdenfors *et al.*, 2001; Brito *et al.*, 2010; Zamin *et al.*, 2010). Incorporating species' conservation status into national and regional decision making is especially important for protection and recovery efforts, because most conservation action and legislation take place at the national and local scale (Miller *et al.*, 2007; Mace *et al.*, 2008; Rodríguez, 2008; Brito *et al.*, 2010).

Significant conceptual differences exist between global and regional extinction risk assessment, primarily because species do not conveniently segregate their populations based on political borders. Biological populations are often distributed across borders, and populations outside the region being assessed can have substantial influence on the long-term survival of populations inside the region, thereby making evaluation of the part of the population resident within a target region very complicated (Gärdenfors, 2001). Deciding which taxa should be assessed adds another layer of complexity, because most countries have nonnative introduced species and nonbreeding migrants and vagrants present within their borders, along with their native historical fauna and flora (Gärdenfors, 2001).

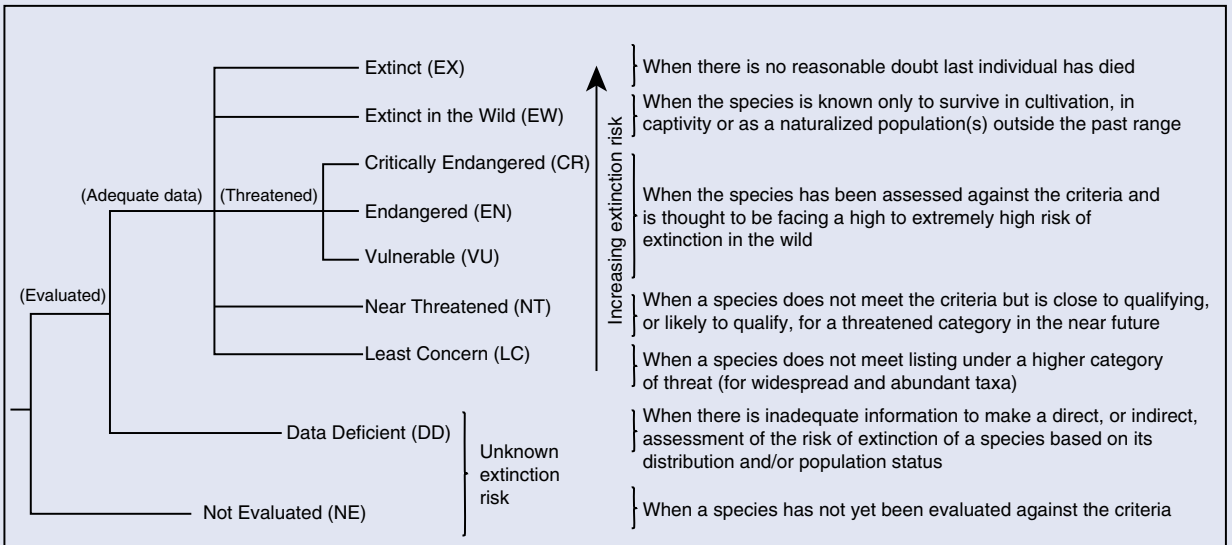
Global conservation status will not always coincide with regional or national conservation status (Gärdenfors, 2001; Gärdenfors *et al.*, 2001; Mace *et al.*, 2008). Certain species may not be at risk globally but can be highly threatened within a particular region (e.g., species at the edge of their geographic range where their population and range size are small). Other species may be threatened globally but be relatively secure within a region; the dugong (*Dugong dugon*), for example, is listed as Vulnerable on the IUCN Red List due to declines throughout at least a third of its global range, yet it appears to be stable throughout much of its Australian range and is not listed as threatened there.

These differences in status at the global and regional levels often cause concern among national assessors, especially when governments have difficulty reconciling these differences in their conservation legislation. The issue may be mitigated by increasing communication and data sharing between global and national assessors (Rodríguez *et al.*, 2000; Miller *et al.*, 2007; Zamin *et al.*, 2010). However, when using different

Box 2

IUCN Red List Categories and Criteria: How does the system work?

The IUCN Red List Categories and Criteria are used to assess wild populations in their natural range. Species are assigned to one of nine categories based on criteria linked to population trend, population size and structure, geographic range size, amount of occupied habitat, and trends in the amount of habitat (**Box 2 Figure**). Species listed as Critically Endangered, Endangered, or Vulnerable are collectively described as 'threatened'. These three hierarchical categories are defined qualitatively by decreasing probabilities of extinction over increasing time scales and quantitatively by the thresholds in the criteria (Mace *et al.*, 2008).



Criterion	Critically Endangered	Endangered	Vulnerable	Qualifiers and notes
A1: reduction in population size	≥90%	≥70%	≥50%	Over 10 years/three generations* in the past, where causes of the reduction are clearly reversible AND understood AND ceased
A2–4: reduction in population size	≥80%	≥50%	≥30%	Over 10 years/three generations * in past, future or combination
B1: small range (extent of occurrence)	<100 km²	<5000km²	<20,000 km²	Plus two of (a) severe fragmentation and/or few locations (1, ≤5, ≤10); (b) continuing decline; (c) extreme fluctuation
B2: small range (area of occupancy)	<10 km²	<500km²	<2000 km²	Plus two of (a) severe fragmentation/few locations (1, ≤5, ≤10), (b) continuing decline, (c) extreme fluctuation
C: small and declining population	<250	<2500	<10,000	Mature individuals. Continuing decline either: (1) over specified rates and time periods; or (2) with (a) specified population structure or (b) extreme fluctuation
D1: very small population	<50	<250	<1000	Mature individuals
D2: very restricted population	N/A	N/A	<20 km² area of occupancy or ≤ five locations	Capable of becoming Critically Endangered or even Extinct within a very short time frame
E: quantitative analysis	≥50% in 10 years/three generations*	≥20% in 20 yeas/five generations*	≥10% in 100 years	Estimated extinction-risk using quantitative models, (e.g., population viability analyses)

Box 2 Figure The IUCN Red List Categories and Criteria. Reprinted from Rodrigues ASL, Pilgrim JD, Lamoreux JF, Hoffmann M, and Brooks TM (2006) The value of the IUCN Red List for conservation. *Trends in Ecology and Evolution* 21: 71–76.

*Whichever is longer. Different species with vastly different life histories may be at a similarly high risk of extinction but for different reasons. The five criteria (A–E) in the Red List methodology thus represent different issues that contribute to increased extinction risk such as population decline or a very restricted population or distribution. Not all the criteria will be appropriate for all species, but at least one should capture the symptoms of threat for every threatened species, be it a marine fish with a high rate of population decline due to overexploitation or a terrestrial mollusk with a restricted and declining range as a result of intense habitat loss and degradation. To qualify for listing in any of the threat categories, a species must meet the threshold values for that category for any one of the five criteria, including all relevant subcriteria.

The Red List criteria operate independently of one another, so meeting the thresholds for one criterion does not mean the rules for another criterion also apply. A species listed as Endangered under criterion B may have far more than the 250 mature individuals required to list it as Endangered under criterion D, for example. All taxa are assessed against every criterion, but only one criterion must be met to list the species; in the final listing, all criteria met at the highest level of threat are given.

To prevent erroneous declarations of extinction, the Red List assessment of 'Extinct' is extremely precautionary; the tag 'possibly extinct' has therefore been created to identify those Critically Endangered species that are likely to be extinct, but for which confirmation is required (IUCN SPSC, 2010).

For species listed as Not Evaluated and Data Deficient, no estimation of extinction risk has been made; these species may or may not be threatened and should not necessarily be treated as secure. Similarly, although Least Concern species do not meet any of the criteria for a threatened category, they may be undergoing slow declines that are below the threshold rates or have suffered severe declines too far in the past to qualify for a category of threat. Such species may still require conservation attention.

Particular biological characteristics are incorporated into the criteria to standardize the thresholds across the broad range of species that may be assessed, considering their variable life histories. For example, 'generation length' is used to scale time-based measures – such as a population decline over a given time period – for the different rates at which taxa survive and reproduce. These terms have very specific definitions that must be interpreted correctly in order for the listing to be correct (see IUCN SPSC, 2010).

For more details, see IUCN (2001) *IUCN Red List Categories and Criteria. Version 3.1*. Gland, Switzerland and Cambridge, United Kingdom: IUCN Species Survival Commission, IUCN; Mace GM, Collar NJ, Gaston KJ, *et al.* (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology* 22: 1424–1442, and IUCN SPSC (Standards and Petitions Subcommittee) (2010) *Guidelines for using the IUCN Red List Categories and Criteria. Version 8.1*. Available at http://www.iucnredlist.org/documents/RedListGuidelines.pdf. Mace *et al.* (2008) give a particularly thorough explanation of the theory and justification of the system, including a detailed description of the logic behind each criterion.

classification systems at different geographic scales, there will inevitably be variability among countries in assessment methods and therefore in species assessments (Rodríguez *et al.*, 2004; Eaton *et al.*, 2005; Miller *et al.*, 2007). An exception is for endemic species assessed using the same classification system at both global and regional scales; these assessments should always concur.

As of 2010, at least 109 countries (~56% of the countries of the world) from virtually every region had produced a national red data book, national red list, or other national list of threatened species, with the fewest coming from Oceania, Africa, and the Caribbean (**Figure 1**; Zamin *et al.*, 2010). These assessments are carried out by government agencies, nongovernmental organizations, or individuals and involve a wide diversity of assessment systems but with few exceptions (e.g., the European Red List initiative) are rarely coordinated across larger regions. This makes large-scale conservation action challenging, as outlined previously.

Regional Assessments Using the IUCN Red List Criteria and Regional Guidelines

The IUCN Red List Categories and Criteria are designed for use at the global level and can lead to incorrect assessments when used at a subglobal scale. This is because small isolated populations often face a higher risk of extinction than large widespread populations (Lande, 1998; O'Grady *et al.*, 2004), so artificially dividing a larger biological population into smaller, more restricted subpopulations located within a political border can cause the individual subpopulations to be assessed as having a higher extinction risk than they actually face (Miller *et al.*, 2007). To resolve this issue, IUCN developed specific guidelines that adapt the criteria for use at the regional level by taking into account the effect that subpopulations outside a region have on the likelihood of that species' extinction within the region (Gärdenfors *et al.*, 1999; Gärdenfors, 2001; Gärdenfors *et al.*, 2001; IUCN, 2003; Miller *et al.*, 2007). These guidelines use data from populations in neighboring regions to determine whether immigration from outside the region may increase or decrease extinction risk within the region; if necessary, the Red List category is then adjusted by uplisting or downlisting the species accordingly, to more appropriately reflect the long-term extinction risk of the subpopulation being assessed (IUCN, 2003).

At last count, 76 countries were using the IUCN methodology for their national red lists (~39% of all countries; Mace *et al.*, 2008), though not all of these countries employ the regional guidelines as they should (Miller *et al.*, 2007). The system has also been adopted for the European Red List initiative for assessments at the pan-European and EU-27 levels. Despite the quantitative and data-reliant nature of the IUCN methodology – a reason many national assessors claim it cannot be used without modification at the national level (Miller *et al.*, 2007) – it has been used successfully in many countries where data are scarce such as Ecuador (e.g., Tirira, 2011) and Mongolia (e.g., Ocock *et al.*, 2006), as well as for poorly studied taxonomic groups such as Venezuelan invertebrates (Rodríguez and Rojas-Suárez, 2008), Croatian cave-dwelling fauna (Ozimec *et al.*, 2009), and Swedish fungi (Gärdenfors, 2010).

Regional Assessments Based on the IUCN Red List Criteria

Many countries have based their national assessment criteria on the IUCN system, modifying and adding to the IUCN Red List Categories and Criteria as they deem necessary to meet their particular needs, or they have incorporated elements of the IUCN methodology into their national criteria (e.g., Australia (www.environment.gov.au), Austria (Zulka *et al.*, 2003), Canada (www.cosewic.gc.ca, www.sararegistry.gc.ca), Japan (www.biodic.go.jp)).

For example, to list a species under the Australian Environment Protection and Biodiversity Conservation Act (1999), the Threatened Species Scientific Committee may consider a range of biological information to inform its judgments about a species' threat status, including indicative thresholds and qualitative criteria that have been adapted from the IUCN criteria. In Canada, the Committee on the

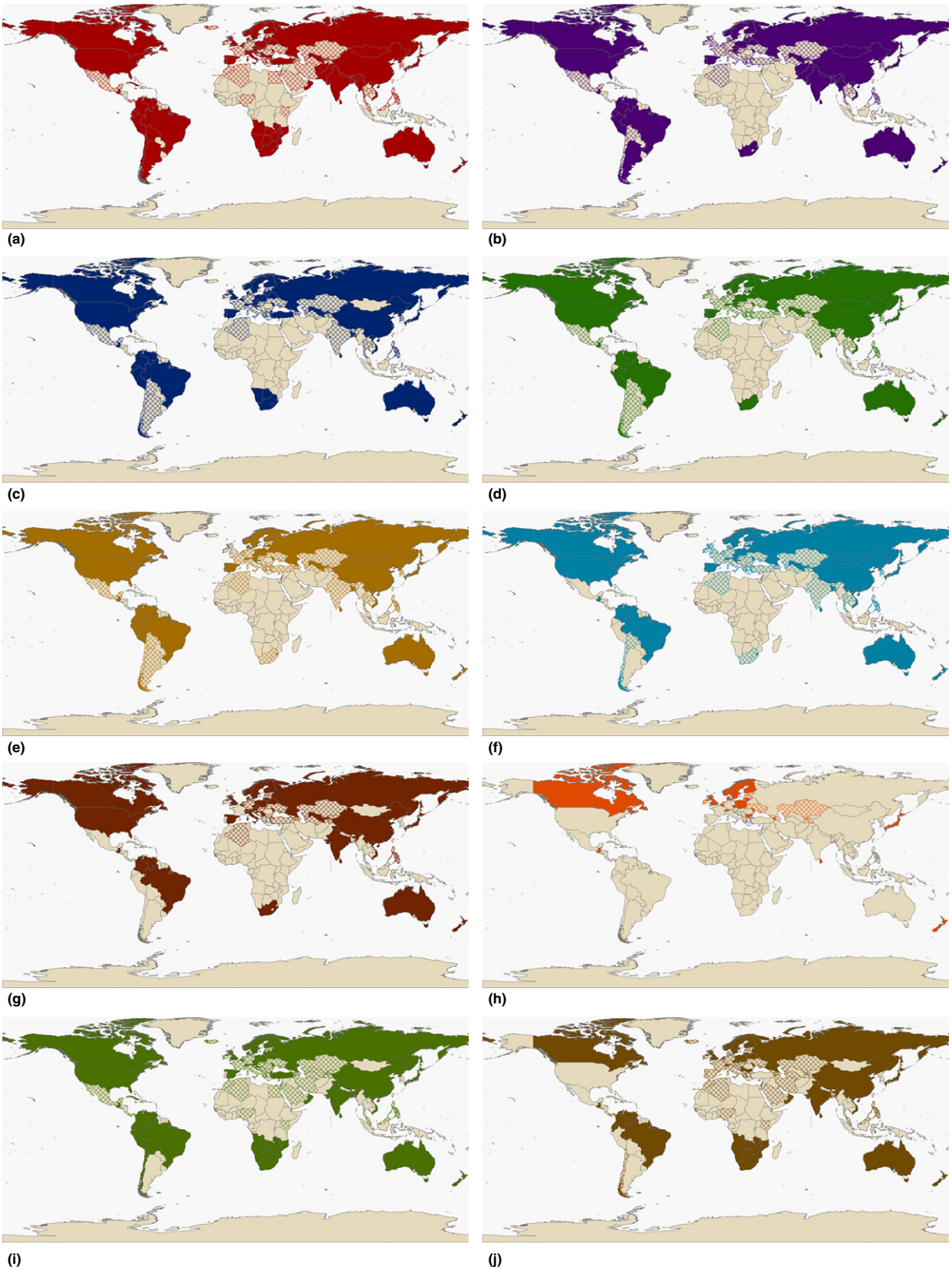


Figure 1 Distribution of (a) countries with at least one national red list and countries with national red lists that cover (b) mammals, (c) birds, (d) amphibians, (e) reptiles, (f) fishes, (g) invertebrates, (h) fungi and lichen, (i) vascular plants, and (j) nonvascular plants. Solid shading indicates lists $<$10 years old; patterned shading indicates lists $\geq$10 years old. Reprinted from Zamin TJ, Baillie JEM, Miller RM, Rodríguez JP, Ardid A, and Collen B (2010) National red listing beyond the 2010 target. *Conservation Biology* 24: 1012–1020.

Status of Endangered Wildlife in Canada (COSEWIC) first assesses the extinction risk of candidate species using a revised version of the IUCN system (having eliminated the Critically Endangered category and changed the name of the Vulnerable category to "Threatened"); however, the government is not obligated to protect threatened species and considers other factors along with threat status when determining which species to legally protect under the Species at Risk Act (SARA).

These systems are effective within the individual countries and often result in strong protective measures for imperiled species. It becomes problematic, however, when a country claims to use the IUCN system but in fact has modified it substantially. It is often difficult to determine which elements of the IUCN methodology have been adopted and how, and this ambiguity can cause substantial confusion. The IUCN regional guidelines were developed to adapt the IUCN system to a subglobal scale and should be used unaltered if the integrity of the system is to be maintained (IUCN, 2003). Of course, many national systems based on the IUCN criteria were developed before the IUCN regional guidelines were published in an effort to fit the IUCN Red List Categories and Criteria to the national scale, and they are now cemented so firmly in governmental legislation that amendment is difficult.

Other Approaches

A few countries and organizations have created their own unique systems for evaluating and listing threatened species. Perhaps the most well known is the United States' Endangered Species Act (ESA); passed in 1973, the ESA is one of the oldest and strongest national laws created to protect imperiled species. The ESA lists species in one of two categories of risk: "endangered" species are in danger of extinction throughout all or a significant portion of their range, whereas "threatened" species are likely to become endangered within the foreseeable future. However, the listing criteria are notably vague, and placement of species in a given category depends entirely on expert judgment. Species are listed based on any of the following factors (ESA, Section 4(a)(1)):

1. the present or future destruction, modification, or curtailment of its habitat or range;
2. overutilization for commercial, recreational, scientific, or educational purposes;
3. disease or predation;
4. the inadequacy of existing regulatory mechanisms; and
5. other natural or manmade factors affecting its continued existence.

Although the ESA lists species "solely on the basis of the best scientific and commercial data available" (ESA, Section 4(b)(1)(A)), in reality political and economic concerns often come into play when deciding whether to list a species (Shelden *et al.*, 2001), as is true in many countries. Listing a species has substantial impact, as the ESA affords direct legal protection to all listed species and their habitats on both publicly and (except for plants) privately owned lands. Federal agencies, private landowners, and citizens are prohibited from harming these species in any way, including damaging important habitats, though permits allowing such activities can be obtained and mitigation plans can be developed to compensate for damage done to listed species or their habitats.

Many species no doubt require such strict measures for their survival, and the ESA is commendable for providing such rigid protection, but the unfortunate result has been that some commercially important or politically sensitive species have not been listed, specifically to avoid the issue of enforcing this protection. The status of hundreds of species that are likely at risk of extinction thus remains undefined, preventing them from benefiting from nonlegally binding protection measures – a regrettable consequence of combining legal protection with threatened status assessment (see **Box 1**).

Another influential classification scheme, the NatureServe Conservation Status methodology, has been used extensively over the past two decades by a network of natural heritage programs and conservation data centers throughout the Western hemisphere. NatureServe assessments provide information on global, national, and subnational (state or province) conservation status for North American animals, plants, and fungi; ecological communities; and ecological systems (collectively called *elements* of biodiversity) (Faber-Langendoen *et al.*, 2009; Master *et al.*, 2009). Although the system includes global assessments, it has been designed primarily for use in the US and Canada.

In the NatureServe system, "elements" are assessed against 10 biological or external factors, grouped into three general categories, that may affect their ability to persist in the wild (Faber-Langendoen *et al.*, 2009):

- *Rarity*: Population Size, Range Extent, Area of Occupancy, Number of Occurrences, Number of Occurrences or Percent Area with Good Viability/Ecological Integrity, and Environmental Specificity.
- *Trends*: Long-term and Short-term Trend in population size or area.
- *Threats*: Threat Impact (generated by considering the scope and severity of the major threats), and Intrinsic Vulnerability (used only if Threat Impact is unknown).

A numerical score is computed based on weightings assigned to each factor and some conditional rules, which is translated to a final rank. This rank can be adjusted if the assessor deems it necessary, though the reasoning must be documented. The final assessment is recorded using a scale of 1–5, G1 being critically imperiled on a global scale and G5 being globally secure; a prefix of N or S indicates national (N1–N5) and subnational assessments (S1–S5). An element can be listed across multiple categories and at several geographic scales; for example, the knobby brain coral (*Diploria clivosa*) is listed as G4G5S2, indicating that there is uncertainty whether it is secure or apparently secure at the global level, whereas it is imperiled at the subnational level (in this case, Florida). Several other symbols can be added to the final rank, such as Q for questionable taxonomy or U for unrankable due to insufficient or conflicting data. There are also categories for extinct (GX/NX/SX) and possibly extinct (GH/NH/SH).

Like the IUCN system, the purpose of the NatureServe conservation status ranks is to classify species by extinction risk in a scientific, objective, and consistent manner. Efforts have

recently been made to standardize the definitions for shared terms in the IUCN and NatureServe methodologies, allowing information to be communicated between organizations and countries and the same data to be used for either system (Master *et al.*, 2009).

Applications of Threatened Species Assessments

From its origins as a general interest in rare and declining wildlife, the science of threatened species assessment has blossomed into a massive conservation theme with far-reaching influence on conservation on the ground (and in the water). Although threat status alone should not determine conservation priorities (Possingham *et al.*, 2002; Rodrigues *et al.*, 2006; Miller *et al.*, 2006, 2007; see Box 1), threatened species lists have become one of the most important tools in the conservationist's toolbox, with threatened species assessment data being widely incorporated into international conservation policies and priority-setting strategies.

Determining the Status of Biodiversity

Threatened species assessment data can provide important insights into not only the status of the individual species assessed but also the overall status of biodiversity and geographical patterns of threat, which is essential information for monitoring and directing effective conservation action.

For example, IUCN has evaluated the global threat status of 59,508 species of animals, plants, and fungi (IUCN, 2011). Since 1500, nearly 800 species have been documented as having gone Extinct, whereas 64 are Extinct in the Wild (IUCN, 2011). The best estimates of the proportion of threatened species among completely assessed groups ranges from 0% to 63% (**Table 1**). Alarmingly, nearly one-fifth (19%) of extant vertebrate species are threatened with extinction (Hoffmann *et al.*, 2010); the true value of this estimate falls between 16% (assuming no Data Deficient species are threatened) and 33% (assuming all Data Deficient species are threatened). Although only 1% of described invertebrate species have been assessed, those few groups that have been completely assessed generally show high levels of threat (**Table 1**). Plants fare no better, especially the ancient conifer and cycad lineages (**Table 1**). Though trends in completely assessed plant groups to date do not likely represent the overall status of plants (Hilton-Taylor *et al.*, 2009), preliminary results from a representative sample of plant species suggest that 22% may be threatened (http://threatenedplants.myspecies.info/), a figure similar to the proportion of threatened vertebrates.

Threatened species lists can also give insight into the status of different ecosystems. Freshwater ecosystems appear to be particularly at risk, as demonstrated by the threat levels of amphibians, freshwater crabs, and crayfish (**Table 1**); although dragonflies are also freshwater dependent, they are less likely to become highly threatened for a variety of reasons (see Clausnitzer *et al.*, 2009). Likewise, some marine groups exhibit high levels of threat – for example, 33% of cartilaginous fish and corals (IUCN, 2011), 36% of marine mammals (Schipper *et al.*, 2008), 27.5% of seabirds, and 86% of marine turtles (Polidoro *et al.*, 2009). This is unsurprising considering the large number of overlapping threats marine species face: high and often unregulated fishing pressure, rapid coastal development, pollution, and changing ocean conditions under climate change (Polidoro *et al.*, 2009). Nonetheless, much is still unknown about threat status in marine species, and many groups exhibit notably high levels of data deficiency (**Table 1**).

Threatened species are concentrated in tropical regions, and these concentrations are disproportionately high even when accounting for their high overall species richness, as exemplified by marine and terrestrial vertebrates (**Figure 2**). This is not surprising given that these areas combine high species richness, high endemism, and high human pressure (Schipper *et al.*, 2008; Hoffmann *et al.*, 2010). Data Deficient species are also highly concentrated in the tropics (Hoffmann *et al.*, 2010); in regions where land and water resources are being rapidly degraded, many DD species may be dangerously close to extinction (Schipper *et al.*, 2008). Plant and invertebrate geographic threat patterns appear to be broadly similar to those of vertebrates (e.g., Carpenter *et al.*, 2008; Clausnitzer *et al.*, 2009; Collen *et al.*, 2009; Polidoro *et al.*, 2010).

Threatening processes themselves are distributed non-randomly worldwide. For example, Li and Wilcove (2005) found that the most pervasive threat to vertebrates in the US is habitat destruction, which affects more than 92% of imperiled species. In China, however, overexploitation is the major threat to vertebrates, affecting 78% of imperiled species, followed by habitat destruction (70%). Maps correlating the presence of threatened species with the primary threats affecting them can be especially informative for conservation decision making (**Figure 3**).

Informing Conservation Action for Individual Species and Guiding Scientific Research

Threatened species lists should be supported by additional data that justify the species assessments (though this is not always the case); these data can be used to help develop conservation programs and guide scientific research and monitoring to fill knowledge gaps. For example, the most important anthropogenic drivers of threat for biodiversity today are habitat destruction, introduction of alien species, overexploitation, disease, and climate change (Mace *et al.*, 2005); the significance of different threats, however, varies both within and between taxonomic groups (**Figure 4**). The IUCN Red List includes a description of the threats affecting listed species and a series of proposed conservation actions to mitigate those threats, thus providing direct recommendations for species-specific measures to be taken (Baillie *et al.*, 2004; Rodrigues *et al.*, 2006). On oceanic islands, for instance, eradication of introduced predators is often a crucial step in restoring native species that have been decimated by these hungry invaders (Ortiz-Catedral *et al.*, 2009).

It is encouraging that several studies have shown that active conservation can be effective in improving the status of biodiversity (Sodhi *et al.*, in press). Hoffmann *et al.* (2010), for example, found that despite an overall decline in vertebrate

Table 1 Number of species in each IUCN Red List category and proportion of threatened species for all vertebrates and all major completely or randomly assessed invertebrate and plant groups. IUCN Red List categories: EX, Extinct; EW, Extinct in the Wild; CR, Critically Endangered; EN, Endangered; VU, Vulnerable; NT, Near Threatened; LC, Least Concern; DD, Data Deficient. Threatened includes Vulnerable, Endangered, and Critically Endangered Species. Asterisks indicate groups in which estimates are derived from a randomized sampling approach. Lower estimates of percent threatened assume no DD species are threatened, upper estimates assume all DD species are threatened, and best estimates assume DD species are threatened in the same proportion as data-sufficient species

Taxonomic level		*EX (%)*	*EW (%)*	*Subtotal*	*CR (%)*	*EN (%)*	*VU (%)*	*Number of threatened species*	*NT (%)*	*LC (%)*	*DD (%)*	*Number of species assessed*	*Lower estimate of percent species threatened: Threatened/Extant assessed*	*Best estimate of percent species threatened: Threatened/(Extant assessed-DD)*	*Upper estimate of percent species threatened: (Threatened + DD)/Extant assessed*
Vertebrates															
Class	Mammals	76 (1)	2 (0.04)	78	191 (3)	447 (8)	496 (9)	1134	324 (6)	3122 (57)	836 (15)	5494	21	25	36
Class	Amphibians	37 (1)	2 (0.03)	39	495 (8)	761 (12)	654 (10)	1910	386 (6)	2377 (38)	1600 (25)	6312	30	41	56
Class	Birds	132 (1)	4 (0.04)	136	190 (2)	372 (4)	678 (7)	1240	838 (8)	7751 (77)	62 (1)	10,027	13	13	13
Class	Reptiles*	0 (0)	0 (0)	0	28 (2)	108 (7)	133 (9)	269	70 (5)	873 (58)	288 (19)	1500	18	22	37
Class	Cartilagenous fish	0 (0)	0 (0)	0	25 (2)	43 (4)	113 (11)	181	132 (13)	241 (23)	487 (47)	1041	17	33	64
Class	Bony fish*	0 (0)	0 (0)	0	34 (2)	35 (2)	102 (7)	171	34 (2)	934 (65)	297 (21)	1436	12	15	33
Class	Hagfish	0 (0)	0 (0)	0	1 (1)	2 (3)	6 (8)	9	2 (3)	35 (46)	30 (39)	76	12	20	51
Plants															
Class	Conifers	0 (0)	0 (0)	0	22 (4)	58 (9)	97 (16)	177	81 (13)	332 (54)	24 (4)	614	29	30	33
Class	Cycads	0 (0)	4 (1)	4	53 (17)	65 (21)	74 (24)	192	63 (21)	45 (15)	3 (1)	307	63	63	64
Functional group[a]	Sea grasses	0 (0)	0 (0)	0	0 (0)	3 (4)	7 (10)	10	5 (7)	48 (67)	9 (13)	72	14	16	26
Functional group[a]	Mangroves	0 (0)	0 (0)	0	2 (3)	3 (4)	6 (9)	11	7 (10)	46 (69)	3 (4)	67	16	17	21
Invertebrates[b]															
Functional group[a]	Reef-building corals	0 (0)	0 (0)	0	5 (1)	25 (3)	201 (24)	231	175 (22)	297 (35)	142 (17)	845	27	33	44
Family[c]	Lobsters	0 (0)	0 (0)	0	0 (0)	0 (0)	0 (0)	0	1 (0.4)	160 (65)	86 (35)	247	0	0	35
Family[d]	Freshwater crabs	0 (0)	0 (0)	0	34 (3)	53 (4)	116 (9)	203	18 (1)	428 (33)	631 (49)	1280	16	31	65
Family[e]	Freshwater crayfish	4 (1)	0 (0)	4	49 (8)	62 (11)	31 (5)	142	46 (8)	262 (44)	136 (23)	590	24	32	47
Order	Dragonflies and damselflies*	0 (0)	0 (0)	0	24 (2)	41 (3)	57 (4)	122	64 (4)	785 (52)	527 (35)	1498	8	13	43
	Total	**249**	**12**	**261**	**1153**	**2078**	**2771**	**6002**	**2246**	**17,736**	**5161**	**31,406**	**19**	**23**	**36**

[a]These species do not represent a unique phylogenetic taxon (e.g., class) but rather a functional group of similar species representing a distinctive habitat or ecological function.

[b]No invertebrate group has been comprehensively or representatively assessed at the class level. All comprehensively or representatively assessed invertebrate groups with current data on the IUCN Red List are presented here, regardless of taxonomic level and including functional groups.

[c]Within the order Decapoda, the lobsters represent the families Glypheidae, Polychelidae, Nephropidae, Enoplometopidae, Palinuridae, and Scyllaridae.

[d]Within the order Decapoda, the freshwater crabs represent the families Gecarcinucidae, Potamidae, Potamonautidae, Pseudothelphusidae, and Trichodactylidae.

[e]Within the order Decapoda, the freshwater crayfish represent the families Astacidae, Cambaridae, and Parastacidae.

Source: Data for reptiles, bony fish, freshwater crayfish, and dragonflies and damselflies from Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science* 330: 1503–1509. All other data from IUCN (2011) *IUCN Red List of Threatened Species. Version 2011.1.* Available at: http://www.iucnredlist.org

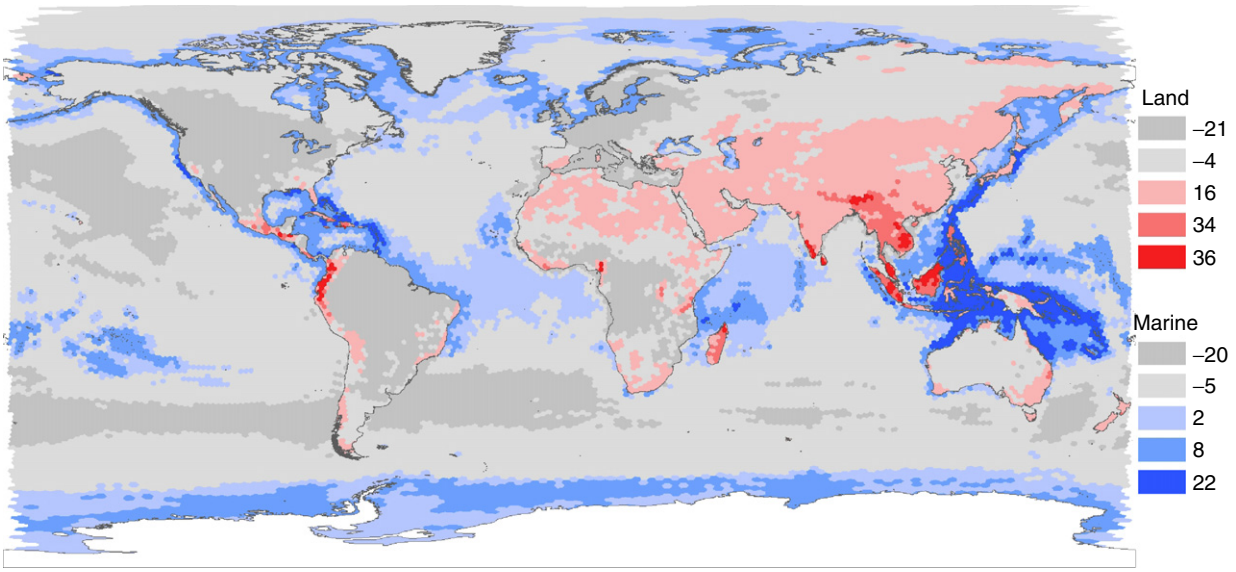


Figure 2 Residuals of the relationship between total number of vertebrate species and total number of threatened vertebrate species per cell, where positive values (red) represent cells with higher threat than expected for their richness alone, and negative values (gray) represent cells with equal or lower threat than expected for richness alone. Reproduced from Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science* 330: 1503–1509, with permission from AAAS.

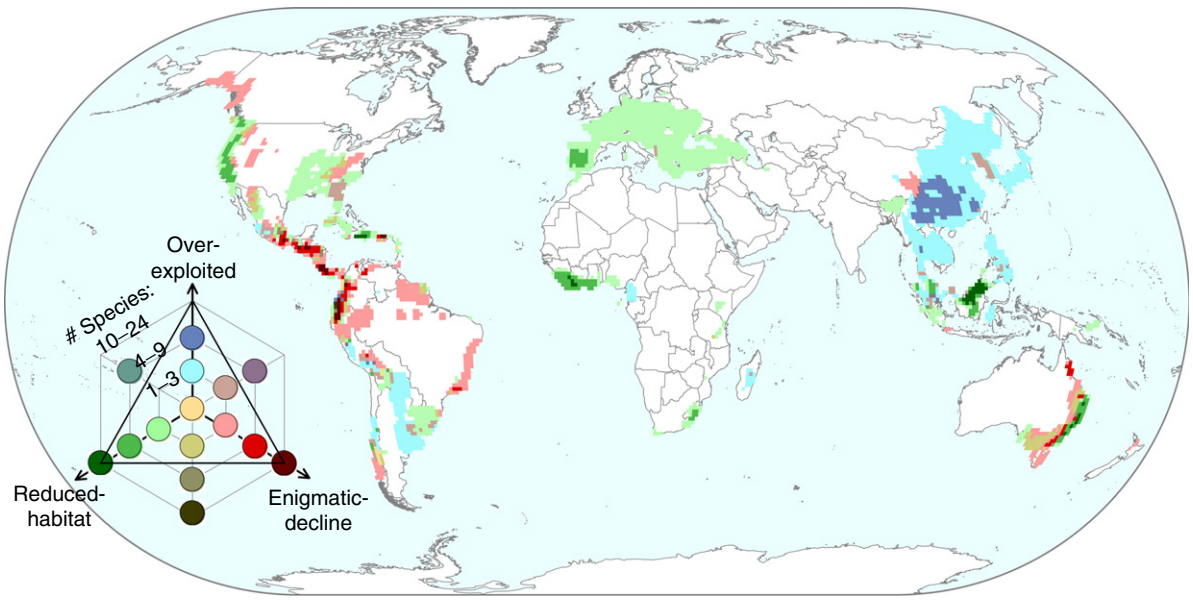


Figure 3 Geographical pattern of the dominant causes of rapid decline in amphibian species: overexploited (shades of blue); reduced habitat (shades of green); and enigmatic decline (shades of red). Where two threat types overlap in the same 1° cell, the color referring to the threat type with the larger number of rapidly declining species in that cell is indicated on the map. Intermediate colors are shown in cases of equal numbers of species experiencing two types of decline in the same cell, as shown in the key. Darker colors correspond to larger numbers of rapidly declining species of any type (not just of the dominant type in the cell in question). Reproduced from Stuart SN, Chanson JS, Cox NA, *et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786, with permission from AAAS.

threat status, 64 vertebrate species listed on the IUCN Red List improved in status as a result of conservation actions; without these conservation successes, vertebrate threat status would have declined by at least an additional 20%. Monitoring changes in the status of targeted species as a result of specific conservation measures can thus help measure the effectiveness of those actions.

Threat status is one of the key criteria for identifying Evolutionarily Distinct and Globally Endangered (EDGE) species, which are phylogenetically unique species that are frequently

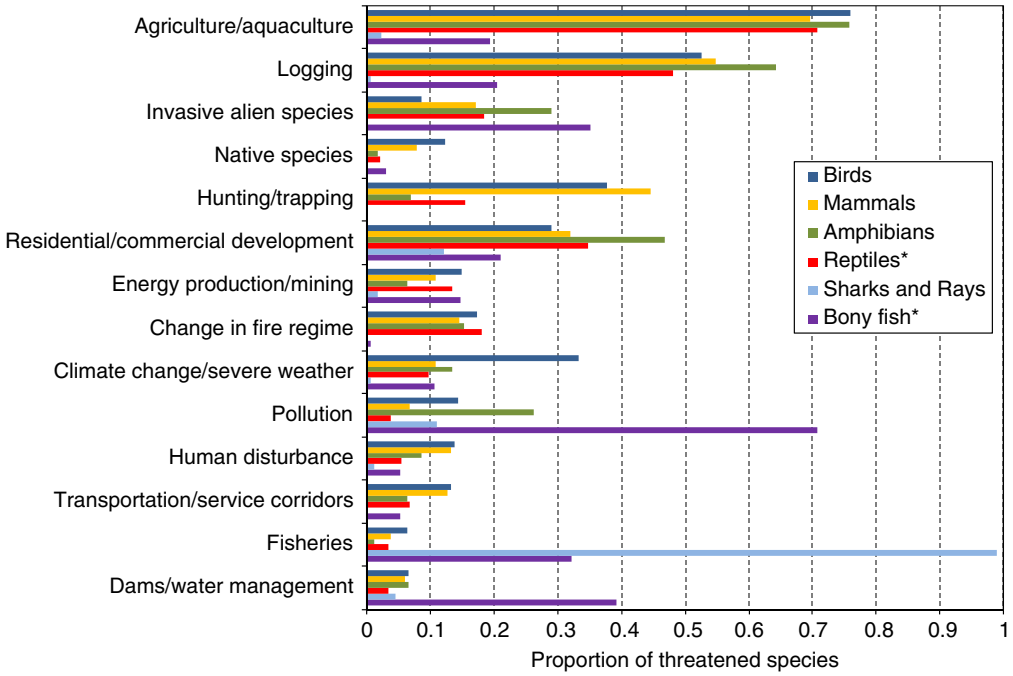


Figure 4 Proportion of threatened vertebrates affected by different major threats. Asterisks indicate groups in which estimates are derived from a randomized sampling approach. Data from Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science* 330: 1503–1509.

overlooked in existing conservation frameworks (Isaac *et al.*, 2007). The EDGE program aims to maintain genetic diversity by focusing conservation attention, resources, and research on threatened species that represent a significant amount of unique evolutionary history.

Good knowledge of a species' population, distribution, biology, threats, and conservation needs and regular monitoring of its threat status over time are essential requirements for any effective conservation strategy. Not enough data exist, however, to complete threat status assessments for many species, some of which may be highly threatened; considerable information gaps remain among even the most well-studied species (Hoffmann *et al.*, 2008). These data deficiencies highlight areas where further study is needed to build baseline information on population size, range, habitat and ecology, threats to species, and the trends in all these measures (Baillie *et al.*, 2004; Hoffmann *et al.*, 2008). Data Deficient species therefore represent high research priorities (Butchart and Bird, 2009). Compiling threatened species lists can reveal information gaps and stimulate data collection in the field (Miller *et al.*, 2007), especially in highly diverse regions that have been poorly studied, where a significant proportion of Data Deficient species are found (Baillie *et al.*, 2004; Hoffmann *et al.*, 2010).

Informing Identification of Important Sites for Biodiversity

Given that habitat destruction is the primary threat for most terrestrial and freshwater species (Baillie *et al.*, 2004), one of the most imperative conservation responses is establishing protected areas to safeguard the habitats in which these species live (Bruner *et al.*, 2001). This would be ideally accomplished by identifying comprehensive networks of protected sites that collectively help to minimize biodiversity loss (Pressey *et al.*, 1993; Margules and Pressey, 2000). Strong political agreements are helping to drive this work; for example, Target 11 of the Convention on Biological Diversity's 2011–2020 Strategic Plan specifically calls for conservation of terrestrial, inland water, coastal, and marine areas of importance for biodiversity, through comprehensive, representative, and well-connected systems of effectively managed protected areas.

Biodiversity assessment data like that found on threatened species lists can be used to help determine gaps in the existing protected area network. Rodrigues *et al.* (2004) used species distribution maps to assess how effectively terrestrial vertebrates are represented in the existing protected area network and identify gaps in coverage; they found that 20% of globally threatened species were not covered in any part of their range, compared to 12% of all species analyzed. Such analyses are also relevant at regional (e.g., Rondinini *et al.*, 2005) and national (e.g., Jaffre *et al.*, 1998) scales.

Various initiatives exist for identifying site-scale conservation targets essential to maintaining biodiversity, which can help to strategically fill these gaps. For instance, European Natura 2000 sites are selected specifically to ensure long-term survival of Europe's most valuable and threatened species and habitats. Presence of threatened species is one of the key criteria for designating Ramsar wetland sites, Ecologically and Biologically Significant Areas (specifically for marine species; CBD COP 9 Decision IX/20), Important Bird Areas (IBAs; Osieck and Mörzer Bruyns, 1981), and important plant

areas (IPAs; Anderson, 2002), among others. The latter two initiatives fall under the umbrella of the Key Biodiversity Areas (KBA) approach (Eken *et al.*, 2004; Langhammer *et al.*, 2007), which aims to identify globally important sites to ensure the persistence of biodiversity based on the presence of threatened, congregatory, or range- or biome-restricted species.

A particularly important subset of KBAs are Alliance for Zero Extinction (AZE) sites (Ricketts *et al.*, 2005), which hold the last remaining population of Critically Endangered or Endangered species as listed on the IUCN Red List. AZE sites reflect the topmost tier of KBA sites where action is urgently needed (**Figure 5**). Several dozen countries have identified KBAs as part of their commitment under the Convention on Biological Diversity's Program of Work on Protected Areas (Brooks, 2010), and the process is underway in many more (Langhammer *et al.*, 2007). A regional KBA approach has been piloted in Turkey, where the IUCN Red List and nationally listed threatened species were used to identify KBAs at both the global and regional scale (Eken *et al.*, 2004).

Policy-Based Actions

Concerns over the mounting biodiversity crisis and the associated impacts on human well-being have led to the creation of international conservation policies with far-reaching aims. The most influential of these is arguably the Convention on Biological Diversity (CBD): in 2002, the 188 CBD government parties agreed an explicit mandate for improving the status of threatened species, by "significantly reducing the rate of biodiversity loss by 2010." Known as the 2010 Target, this ambitious goal inspired a flurry of conservation activity as signatory countries attempted to both meet the target and devise methods for evaluating progress toward this goal.

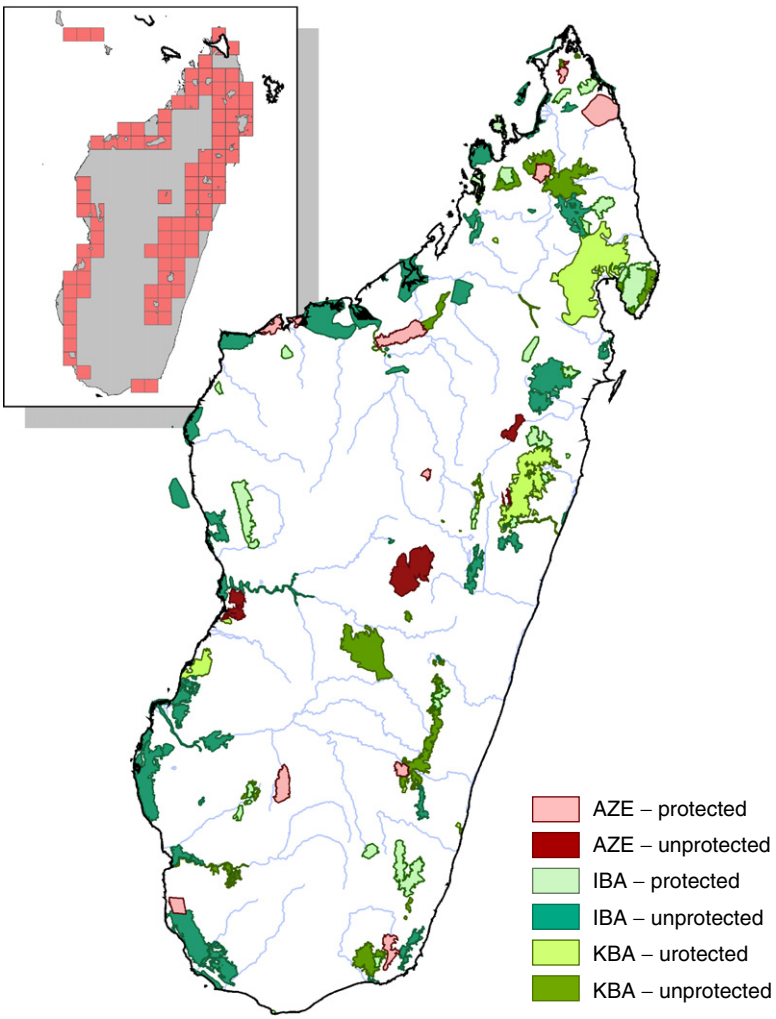


Figure 5 Site-scale conservation priorities in Madagascar, reflecting the nested and complementary nature of sites identified at different scales. Key Biodiversity Areas (KBA; $n=117$) identified from the distributions of threatened species in eight taxonomic groups (mammals, birds, amphibians, freshwater fish, reptiles, arthropods, gastropods and plants); Important Bird Areas (IBA; $n=78$), the avian subset of KBAs; Alliance for Zero Extinction (AZE) sites ($n=16$) contain the entire population of Critically Endangered or Endangered species. Inset: urgent priorities (pink grid squares) identified at a larger scale for expanding the network of protected areas in Madagascar, ranked according to an extinction risk index (data for mammals, birds, amphibians, and turtles). Reprinted from Hoffmann M, Brooks TM, da Fonseca GAB, *et al.* (2008) Conservation planning and the IUCN Red List. *Endangered Species Research* 6: 113–125, with permission from Inter Research.

Although the 2010 Target was not met (Mace *et al.*, 2010), the 20 "Aichi Targets" adopted as part of the 2011–2020 CBD Strategic Plan (COP 10 Decision X/2) built on the original goal. Information from threatened species lists can be used to measure progress in at least 13 of the 20 Targets (*see* Importance for Monitoring State of Biodiversity: Developing Indicators of Global Biodiversity Status). Achieving Target 12 (extinction and decline of threatened species is prevented and their status improved) in particular will depend on understanding threatening processes and addressing their impacts successfully. National CBD commitments will result in increasing efforts to protect and recover threatened species as the new strategic plan is translated into CBD-mandated National Biodiversity Strategies and Action Plans.

Assessment data from threatened species lists and change in the status of threatened species can inform several other international environmental agreements. For example:

- The CBD 2010 Target was adopted into Target 7b of the United Nations Millennium Development Goal 7 (MDG 7): Ensure environmental sustainability.
- Appendix I of the Convention on Migratory Species (CMS) lists migratory species threatened with extinction; parties to the convention strive to strictly protect these species and their habitats and remove any obstacles to their migration.
- Targets 2, 7, and 8 of the CBD Global Strategy for Plant Conservation call for the status assessment of all known plant species and *in situ* or *ex situ* conservation of 75% of threatened species.
- One criterion for designating a wetland for special protection under the Ramsar Convention is if the wetland "supports vulnerable, endangered, or critically endangered species" as defined by national or international threatened species programs (e.g., national threatened species lists, the IUCN Red List) or frameworks (e.g., CMS appendices).
- The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) aims to ensure that international trade in specimens of wild animals and plants does not threaten their survival; strict controls on the trade of species listed on the three appendices are established, with species on Appendix I (species threatened with extinction that are or may be affected by trade) afforded the most protection (trade is permitted only in exceptional circumstances).

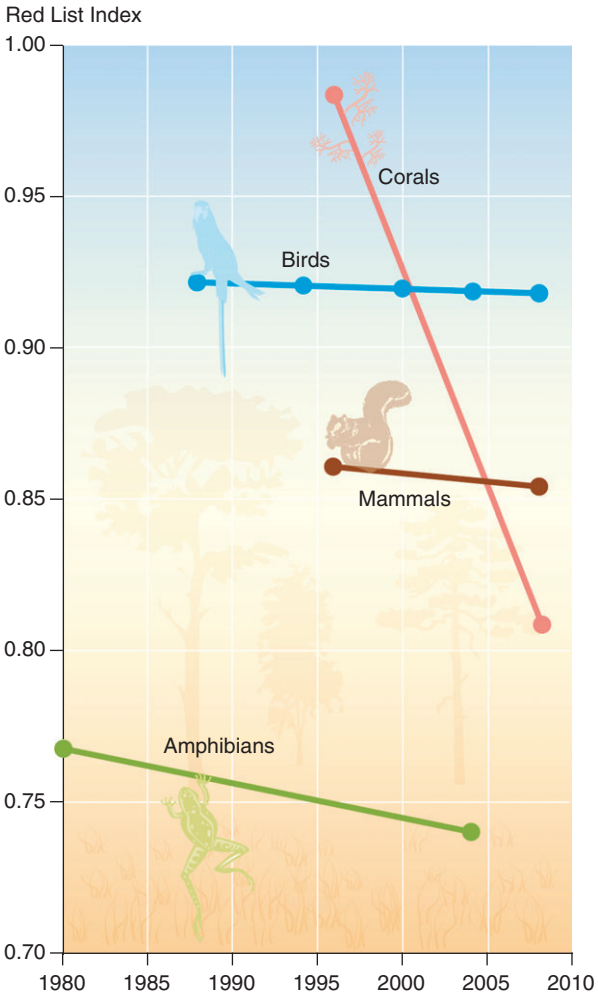
It is worth discussing CITES in more detail, as many misinterpret it to be a method for measuring extinction risk. CITES lists species in three appendices according to their conservation status and the risk they face from international trade. The CITES appendices are not traditional lists of threatened species, as they do not aim purely to report vulnerability to extinction; rather, they highlight those species for which trade may jeopardize their survival. Only those species that are demonstrably threatened by international trade should be listed on CITES, though these species are often absent from CITES appendices, whereas others not threatened by trade are erroneously listed on CITES (Stuart *et al.*, 2008). Furthermore, CITES criteria simply provide guidance; there is no obligation to list a species that meets the criteria. Consequently, listing a species under CITES often sparks hot debate, especially if the decision is controversial or involves a species with a lucrative international trade market. Many argue that CITES is more political than conservation-oriented, but ultimately the fact that thousands of species are afforded international protection due to their placement on the CITES appendices makes the system immensely influential in the fate of threatened species that are internationally traded.

The European Union has pioneered the implementation of conservation policies at the regional level, helped no doubt by their cohesive political stature. Since 1982, European signatory countries to the Convention on the Conservation of European Wildlife (the Bern Convention) have committed to conserving wild flora and fauna, with particular emphasis on "endangered and vulnerable" species and habitats. The Birds and Habitats Directives followed from the Bern Convention and required European Union member states to maintain and restore favorable conservation status of the endangered and vulnerable species listed in their annexes. This is often achieved through the designation of Natura 2000 sites; outputs of the European Red List initiative can be used to inform placement of species on the directives as well as Natura 2000 site identification. Donald *et al.* (2007) found that the Birds Directive has brought demonstrable benefits to European bird populations. Inspired by the CBD 2010 Target, in 2003 European ministers of environment adopted the even-stricter target of halting biodiversity loss by 2010. Although these conservation aims have not always been met, the European example shows that international policies can be effective in addressing conservation issues over wider regions (Donald *et al.*, 2007).

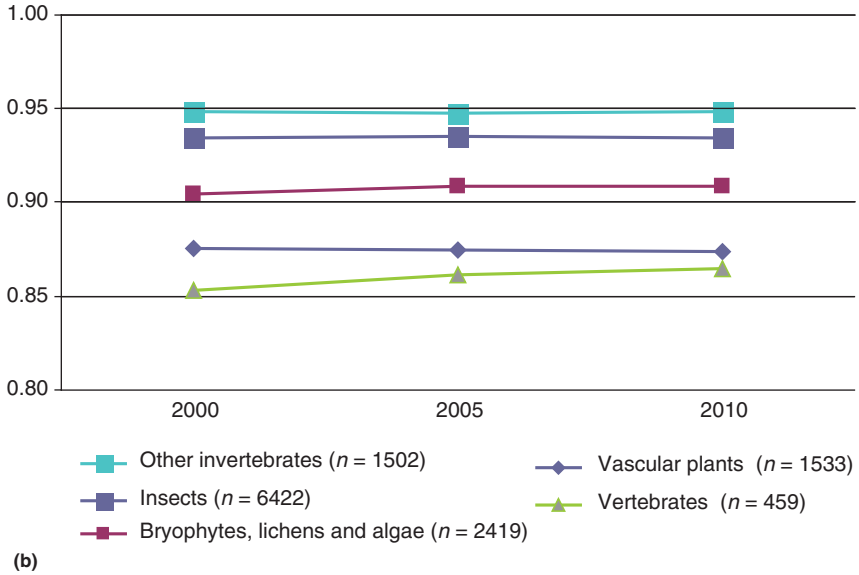
Biodiversity assessment data can also be used to inspire and inform national development policies and legislation to protect species and the sites they inhabit (e.g., in the areas of land-use planning, transport, energy, poverty reduction strategies, and national biodiversity strategies and action plans) (Rodrigues *et al.*, 2006). Moreover, in many countries there is a direct connection between threatened species lists and conservation legislation, as exemplified by the US Endangered Species Act (*see* Classification Systems for Assessing Extinction Risk). As important and influential as national conservation policies are, they must be crafted carefully; basing protective legislation directly on threat categories can have unanticipated consequences, such as when governments prohibit access to threatened species and thus impede their scientific study (Rodrigues *et al.*, 2006), and may not lead to optimal expenditure of resources (Possingham *et al.*, 2002).

Importance for Monitoring State of Biodiversity: Developing Indicators of Global Biodiversity Status

Since the adoption of the CBD 2010 target, much effort has been dedicated to developing indicators capable of measuring changes in the state of biodiversity over time, assessing the impact of policy and management responses, and identifying priorities for action (Walpole *et al.*, 2009). One such indicator is the IUCN Red List Index (RLI; Butchart *et al.*, 2004, 2005, 2007), which monitors trends in species' threat status and is thus a key measure of whether the rate of biodiversity loss is escalating or being reduced. The RLI was adopted as one of the 22 headline indicators to measure progress toward the CBD


Red List Index
1.00
0.95
0.90
0.85
0.80
0.75
0.70
1980
1985
1990
1995
2000
2005
2010
Corals
Birds
Mammals
Amphibians

(a)


1.00
0.95
0.90
0.85
0.80
2000
2005
2010
Other invertebrates (n = 1502)
Insects (n = 6422)
Bryophytes, lichens and algae (n = 2419)
Vascular plants (n = 1533)
Vertebrates (n = 459)

(b)

2010 Target and subsequently incorporated into other global policies (e.g., MDG 7, CMS) and regional initiatives (e.g., Streamlining European Biodiversity Indicators, SEBI-2010).

The RLI shows trends in the overall extinction risk of species based on the number of species that move between IUCN Red List categories over time as a result of genuine improvements or genuine deterioration (i.e., category changes due to improved knowledge or revised taxonomy are excluded) (Butchart *et al.*, 2004, 2007). It can be calculated for any set of species that has been assessed at least twice.

At the global level, a RLI of comprehensively assessed groups on the IUCN Red List shows that the rate of biodiversity loss is increasing (Butchart *et al.*, 2010; Secretariat of the Convention on Biological Diversity, 2010); coral species in particular are moving rapidly toward extinction, while amphibians are, on average, more threatened than birds, mammals, or corals (**Figure 6(a)**). At a national level, the situation in Sweden appears less grim (**Figure 6(b)**): the rate of biodiversity loss is relatively constant for most species and even improving for vertebrates, though ultimately 1942 species are still at risk of extinction (Gärdenfors, 2010).

The RLI can and has been used to monitor trends in biodiversity status at the regional (e.g., IUCN, 2010), national (e.g., Gärdenfors, 2010), and subnational (e.g., Quayle *et al.*, 2007) levels (Bubb *et al.*, 2009). It can also be disaggregated by taxonomy, body size, or ecology or by habitat, ecosystem, region, or biogeographic realm (e.g., Butchart *et al.*, 2005; Stuart *et al.*, 2008; Hilton-Taylor *et al.*, 2009; Hoffmann *et al.*, 2011); threat (Butchart, 2008; McGeoch *et al.*, 2010); or benefit to humans (e.g., internationally traded food and medicine species; Butchart *et al.*, 2010). A sampled RLI approach has been developed, based on the rate at which a stratified sample of species moves through the IUCN Red List categories (Baillie *et al.*, 2008), which will enable us to track conservation status across a broader, more representative range of taxonomic groups.

Resource Allocation

Demand for human and financial resources to support conservation efforts far outweighs the supply. Given this harsh reality, conservation agencies and funding bodies often use species' threat status to help focus priorities for global conservation resources and to guide conservation investment. For example, IUCN Red List data have been incorporated into the system for resource allocation (which determines how much money each focal country can receive) of the Global Environment Facility (GEF), the largest funder of projects to improve the global environment (GEF has allocated $9.5 billion (US) since 1991). Numerous funding organizations draw on threatened species information to decide where to allocate resources, such as the $130 million (US) Critical Ecosystem Partnership Fund, the $6 million (US) Mohammad bin Zayed Species Conservation Fund, and the $4 million (US) Conservation Leadership Program.

One downside to using threat category to inform resource allocation is that species in lower threat categories and Data Deficient species may be disregarded when allocating resources for conservation (Gillespie *et al.*, 2011). Allocating the most funds to the most-threatened species could in theory also fail to maximize conservation returns, as less-threatened taxa can often be recovered with a smaller investment, leaving more money to go around (Possingham *et al.*, 2002), although this concern may be offset by the fact that most threatened species live in those areas of the world where conservation interventions are cheapest (*see* Determining the Status of Biodiversity. And despite the sums spent on conservation annually, the amount required to even begin safeguarding biodiversity remains a distant goal (though reachable through modification of government expenditure) and a disproportionate amount is spent in wealthy countries where the number of threatened species is lower (James *et al.*, 1999, 2001).

Informing Environmental Safeguards

Economic development and poverty alleviation have historically been considered incompatible with biodiversity conservation. Today, however, protecting and conserving biodiversity, including the services and products that natural habitats provide, and sustainably managing natural resources are increasingly being recognized as fundamental to lasting sustainable development. Appropriate environmental safeguards – tools and policies specifically designed to minimize the level of environmental risk (including offsetting any inevitable environmental impacts) and to maximize the environmental benefits of development projects – are essential to guaranteeing that biodiversity conservation and other environmental objectives are integrated into development plans.

The loss of biodiversity undermines not only the species and habitats directly affected but also the human communities that rely on natural systems for their livelihoods (Millennium Ecosystem Assessment, 2005). Leading development banks and financial institutions such as the World Bank, European Bank for Reconstruction and Development, and GEF have adopted policies that ensure the projects they fund support and promote the protection and restoration of natural environments; data from threatened species lists can feed directly into these policies. For example, Performance Standard 6 (Biodiversity Conservation and Sustainable Management of

Figure 6 IUCN Red List Indices (RLI) of species survival, showing proportion of species expected to survive into the near future without additional conservation action. An RLI value of 1 indicates that all species are Least Concern, whereas a value of 0 indicates that all species have gone Extinct. A downward trend indicates that the expected rate of species extinctions is increasing (i.e., the rate of biodiversity loss is increasing); a horizontal line implies that extinction risk is constant; an upward trend means that the rate of biodiversity loss is being reduced. (a) The global extinction risk of birds, mammals, amphibians, and reef-forming corals is increasing, with corals showing the most dramatic declines. (*Source*: IUCN. Reprinted with permission from Secretariat of the Convention on Biological Diversity (2010) *Global Biodiversity Outlook 3*. Montréal: Secretariat of the Convention on Biological Diversity.) (b) The expected rate of extinctions in Sweden is improving somewhat for vertebrates but declining for vascular plants; the rate of biodiversity loss for other taxonomic groups remains relatively unchanged. (Reprinted with permission from Gärdenfors U (ed.) (2010) *Rödlistade arter i Sverige – The 2010 Red List of Swedish Species 2010*. SLU, Uppsala: ArtDatabanken.)

Living Natural Resources) in the International Finance Corporation's Policy and Performance Standards on Social and Environmental Sustainability explicitly makes reference to threatened species under the IUCN Red List criteria, prohibiting any project activity that leads to measurable adverse impacts or a direct reduction in any globally or regionally listed Critically Endangered or Endangered species (IFC Performance Standard 6(17)).

Protocols for ensuring environmental stewardship in project development have also been widely incorporated into international, national, and state laws and treaties. Article 14 of the CBD, for example, specifically calls for parties to implement environmental impact assessment (EIA) procedures. Conducting an EIA during the planning stage of a project life cycle has become standard practice for many companies and organizations, including national, state, and local governments. Though it is often difficult to predict the impacts of development on biodiversity (Meynell, 2005), whether a project is likely to affect threatened species or their habitats is generally one key consideration in deciding whether the project may proceed or how to mitigate the environmental impacts the project may have.

Challenges and Future Directions

We have come a long way toward understanding and measuring extinction risk in the past century, and we are well on our way to integrating this knowledge into proactive conservation policies and programs. However, some tough challenges remain to be met if we are to fully understand how and where the multiple drivers of threat impact different species and how best to respond to minimize biodiversity losses.

Addressing both taxonomic and geographic knowledge gaps is a huge priority. Data remain scarce for many taxonomic groups and those data that do exist are biased toward terrestrial species and temperate zones (Mace *et al.*, 2005), hampering efforts to measure the state of biodiversity worldwide. Most threatened species lists (including the IUCN Red List) lack important information on plants, invertebrates, fungi, and freshwater and marine species (Stuart *et al.*, 2010), whereas national lists are missing entirely from many countries, especially in Oceania, Africa, and the Caribbean (Zamin *et al.*, 2010). The Sampled Red List approach (Baillie *et al.*, 2008) is a big step toward addressing taxonomic knowledge gaps, but expanding global biodiversity datasets to meet the geographic gaps will require enlisting many new experts worldwide (Stuart *et al.*, 2010).

Long-term investment in iterative biodiversity assessment is needed at all levels. As knowledge of biodiversity increases and intensifying threats change the ecological landscape, our measures of species' extinction risk can and will change dramatically. If we hope to develop effective mechanisms for protecting and recovering species and to monitor the conservation actions we put in place, then we must stay abreast of changes in threat status. Establishing a representative "barometer of life" (Stuart *et al.*, 2010) and maintaining and improving assessment data will require a sincere commitment not only by the scientific community but also by the governments and funding bodies that support these efforts (e.g., approximately $400,000 (US) year^{-1} is needed to maintain the IUCN Red List mammal data set alone; Hoffmann *et al.*, 2011).

Our long-term knowledge of trends must catch up with our knowledge of status. Despite various data gaps, we have a substantial body of knowledge that can provide insight into the state of biodiversity today. However, we are only beginning to gauge how this status is changing over time, additional indicators to measure these trends are needed. Monitoring trends in status is increasingly important for determining whether we are effectively reducing the rate of biodiversity loss, but to do so we need a time series of comparable measures (Mace *et al.*, 2005). This requires that assessment methods as well as the groups being assessed be consistent over time, and links to the need for repeated and consistent assessments.

Understanding the multiple and interacting impacts that climate change may have on species and how these impacts affect extinction risk is another critical task (Akçakaya *et al.*, 2006). Despite the burgeoning literature on potential and realized climate change impacts, on-the-ground evaluations of vulnerability assessments and predictions are rare. IUCN has begun to address this issue, adding codes describing various climate change threat types to its assessment protocols and updating guidance on how to interpret the Red List Categories and Criteria in a climate change context (IUCN SPSC, 2010). Methods for assessing species' climate change susceptibility, designed to be used during or complementary to extinction risk assessment, are also emerging (Foden *et al.*, 2009; Young *et al.*, 2011). Evaluating the accuracy of extinction risk predictions based on climate change impacts is an important new frontier.

Threatened species lists at all spatial scales would benefit from increased communication and cooperation between countries, as well as better linkages between national and global assessment efforts (Rodríguez *et al.*, 2000; Miller *et al.*, 2007). The national "red listing" network is rapidly expanding, and with it interest in using the IUCN methodology to assess species extinction risk. Integrating the diffuse network of national assessors into the global IUCN Red Listing process could bring about the dual benefits of expanding information on the world's threatened species and building local scientific capacity for generating and using these data to effect conservation action on the ground (Rodríguez, 2008). Existing experts such as IUCN Species Survival Commission members can play an important role in helping to build and strengthen such a network by acting as advisers to national assessment efforts. Regular training of both global and national assessors is imperative to maintain and improve assessment rigor and consistency, as well as increase awareness of the multiple issues that affect extinction risk assessment. In addition, standardizing classification systems between countries and with the IUCN Red List would help facilitate knowledge transfer and allow for the development of coordinated strategies for data collection and biodiversity conservation across wider regions.

Finally, our biggest challenge over the next few decades will be to respond to the massive ecological shifts already underway and to stem the rising tide of extinction. The information contained in threatened species assessments will be fundamental to informing our progress and directing conservation strategies. Tackling the many drivers of threat to reduce rates of biodiversity loss will be key to success but is undoubtedly the

most difficult challenge that lies ahead. Yet considering the rapid expansion of conservation science and the explosive growth of threatened species lists in the past several decades, we should be up to the challenge. For our own future as much as for the future of the biodiversity on our planet, it is imperative that we get it right.

Appendix

List of Courses

1. Conservation Biology
2. Wildlife Ecology and Conservation
3. Biodiversity, Conservation and Management
4. Environmental Studies, Policy and Natural Resource Management
5. Conservation Policy

See also: Biodiversity State and Trends in Southeast Asia. Climate Change and Extinctions. Climate Change: Anticipating and Adapting to the Impacts on Terrestrial Species. Comparing Extinction Rates: Past, Present, and Future. Conservation and People. Conservation Efforts, Contemporary. Conservation Genetics. Conserving Biodiversity Outside Protected Areas. Convention on Biological Diversity. Corals and Coral Reefs. Defining, Measuring, and Partitioning Species Diversity. Deforestation and Land Clearing. Diseases, Conservation and. Endangered Amphibians. Endangered Freshwater Invertebrates. Endangered Mammals. Endangered Marine Invertebrates. Endangered Plants. Endangered Reptiles. Endangered Terrestrial Invertebrates. Evolution in Response to Climate Change. Extinction, Causes of. Extinction in the Fossil Record. Fish Conservation. Framework for Assessment and Monitoring of Biodiversity. Global Declines of Amphibians. Habitat Loss and Fragmentation. Human Impact on Biodiversity, Overview. Implications of Urbanization for Conservation and Biodiversity Protection. *In Situ*, *Ex Situ* Conservation. Introduced Species, Impacts and Distribution of. Latent Extinction—The Living Dead. Latitudinal and Elevational Range Shifts under Contemporary Climate Change. Loss of Biodiversity, Overview. Mammals, Conservation Efforts for. Mammals (Pre-Quaternary), Extinctions of. Marine Conservation in a Changing Climate. Marine Mammals, Extinctions of. Marine Protected Areas: Static Boundaries in a Changing World. Mass Extinctions, Concept of. Mass Extinctions, Notable Examples of. Modern Examples of Extinctions. Natural Extinctions (not Human Influenced). Natural Reserves and Preserves. Plant Conservation. Plant Invasions. Rainforest Loss and Change. Role and Trends of Protected Areas in Conservation. Species, Concepts of. Species Diversity, Overview. Species–Area Relationships. Subspecies, Semispecies, Superspecies. Threatened Birds

References

Akçakaya HR, Butchart SHM, Mace GM, Stuart SN, and Hilton-Taylor C (2006) Use and misuse of the IUCN Red List Criteria in projecting climate change impacts on biodiversity. *Global Change Biology* 12: 2037–2043.

Akçakaya HR, Ferson S, Burgman MA, Keith DA, Mace GM, and Todd CR (2000) Making consistent IUCN classifications under uncertainty. *Conservation Biology* 14: 1001–1013.

Allen GM (1942) *Extinct and Vanishing Mammals of the Western Hemisphere*. New York: ACIWP.

Anderson S (2002) *Identifying Important Plant Areas*. London, UK: Plantlife International.

Anon (1928) Notes on animals which have become totally extinct through human agency during the nineteenth century. *Journal of the Society for the Preservation of the Fauna of the Empire, New Series* 7: 108–109.

Avery M, Gibbons DW, Porter R, Tew T, Tucker G, and Williams G (1995) Revising the British Red List for birds: the biological basis of U.K. conservation priorities. *Ibis* 137: 232–239.

Baillie J and Groombridge B (eds.) (1996) *1996 IUCN Red List of Threatened Animals*. Gland, Switzerland and Cambridge, UK: IUCN.

Baillie JEM, Collen B, Amin R, *et al.* (2008) Toward monitoring global biodiversity. *Conservation Letters* 1: 18–26.

Baillie JEM, Hilton-Taylor C, and Stuart SN (eds.) (2004) *2004 IUCN Red List of Threatened Species. A Global Species Assessment*. Gland, Switzerland and Cambridge, UK: IUCN.

Barnosky AD, Matzke N, Tomiya S, *et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57.

Bennett PM and Owens IPF (1997) Variation in extinction risk among birds: Chance or evolutionary predisposition? *Proceedings of the Royal Society of London Series B – Biological Sciences* 264: 401–408.

Bollmann K, Keller V, Müller W, and Zbinden N (2002) Prioritäre Vogelarten für Artenförderungsprogramme in der Schweiz. *Der Ornithologische Beobachter* 99: 301–320.

Brito D, Ambal RG, Brooks T, *et al.* (2010) How similar are national red lists and the IUCN Red List? *Biological Conservation* 143: 1154–1158.

Brooks TM (2010) Conservation planning and priorities. In: Sodhi NS and Ehrlich PR (eds.) *Conservation Biology for All*, pp. 199–219. New York: Oxford University Press, Inc.

Bruner AG, Gullison RE, Rice RE, and da Fonseca GAB (2001) Effectiveness of parks in protecting tropical diversity. *Science* 291: 125–128.

Bubb PJ, Butchart SHM, Collen B, *et al.* (2009) *IUCN Red List Index – Guidance for National and Regional Use*. Gland, Switzerland: IUCN.

Butchart SHM (2008) Red List Indices to measure the sustainability of species use and impacts of invasive alien species. *Bird Conservation International* 18: 245–262.

Butchart SHM, Akçakaya HR, Chanson J, *et al.* (2007) Improvements to the Red List Index. *PLoS ONE* 2: e140.

Butchart SHM and Bird J (2009) Data Deficient birds on the IUCN Red List: what don't we know and why does it matter? *Biological Conservation* 143: 239–247.

Butchart SHM, Stattersfield AJ, Baillie J, *et al.* (2005) Using Red List Indices to measure progress towards the 2010 target and beyond. *Philosophical Transactions of the Royal Society B* 360: 255–268.

Butchart SHM, Stattersfield AJ, Bennun LA, *et al.* (2004) Measuring global trends in the status of biodiversity: Red List Indices for birds. *PLoS Biol* 2: e383.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: indicators of recent declines. *Science* 328: 1164–1168.

Carpenter KE, Abrar M, Aeby G, *et al.* (2008) One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* 321: 560–563.

Clausnitzer V, Kalkman VJ, Ramc M, *et al.* (2009) Odonata enter the biodiversity crisis debate: The first global assessment of an insect group. *Biological Conservation* 142: 1864–1869.

Cofré H and Marquet PA (1999) Conservation status, rarity, and geographic priorities for conservation of Chilean mammals: an assessment. *Biological Conservation* 88: 53–68.

Collar NJ (1996) The reasons for Red Data Books. *Oryx* 30: 121–130.

Collen B, Ram M, Dewhurst N, *et al.* (2009) Broadening the coverage of biodiversity assessments. In: Vié JC, Hilton-Taylor C, and Stuart SN (eds.) *Wildlife in a Changing World – An analysis of the 2008 IUCN Red List of Threatened Species*, pp. 67–75. Gland, Switzerland: IUCN.

de Grammont PC and Cuarón AD (2006) An evaluation of threatened species categorization systems used on the American continent. *Conservation Biology* 20: 14–27.

de Iongh HH, Bánki OS, Bergmans W, and van der Werff ten Bosch MJ (eds.) (2003) *The Harmonization of Red Lists for Threatened Species in Europe*. Leiden: The Netherlands Commission for International Nature Protection.

Dollman G (1937) Mammals which have recently become extinct and those on the verge of extinction. *Journal of the Society for the Preservation of the Fauna of the Empire* New Series 30: 67–74.

Donald PF, Sanderson FJ, Burfield IJ, Bierman SM, Gregory RD, and Waliczky Z (2007) International conservation policy delivers benefits for birds in Europe. *Science* 317: 810–813.

Dunn EH, Hussell DJT, and Welsh DA (1999) Priority setting tool applied to Canada's landbirds based on concern and responsibility for species. *Conservation Biology* 13: 1404–1415.

Eaton M, Gregory RD, Noble DG, *et al.* (2005) Regional IUCN red-listing: The process as applied to birds in the United Kingdom. *Conservation Biology* 19: 1557–1570.

Eken G, Bennun L, Brooks TM, *et al.* (2004) Key Biodiversity Areas as site conservation targets. *BioScience* 54: 1110–1118.

Faber-Langendoen D, Master L, Nichols J, *et al.* (2009) *NatureServe Conservation Status Assessments: Methodology for Assigning Ranks*. Arlington, VA: NatureServe.

Feria Arroyo TP, Olson ME, García-Mendoza A, and Solano E (2009) A GIS-based comparison of the Mexican national and IUCN methods for determining extinction risk. *Conservation Biology* 23: 1156–1166.

Fitter R and Fitter M (1987) *The Road to Extinction*. Cambridge, UK and Gland, Switzerland: IUCN/UNEP.

Foden W, Mace G, Vié JC, *et al.* (2009) Species susceptibility to climate change impacts. In: Vié JC, Hilton-Taylor C, and Stuart SN (eds.) *Wildlife in a Changing World – An Analysis of the 2008 IUCN Red List of Threatened Species*, pp. 77–87. Gland, Switzerland: IUCN.

Freeman MMR (2008) Challenges of assessing cetacean population recovery and conservation status. *Endangered Species Research* 6: 173–184.

Gärdenfors U (2001) Classifying threatened species at national versus global levels. *Trends in Ecology and Evolution* 16: 511–516.

Gärdenfors U (ed.) (2010) *Rödlistade arter i Sverige – The 2010 Red List of Swedish Species 2010*. SLU, Uppsala: ArtDatabanken.

Gärdenfors U, Hilton-Taylor C, Mace GM, and Rodríguez JP (2001) The application of IUCN Red List criteria at regional levels. *Conservation Biology* 15: 1206–1212.

Gärdenfors U, Rodríguez JP, Hilton-Taylor C, *et al.* (1999) Draft guidelines for the application of IUCN Red List criteria at national and regional levels. *Species* 31–32: 58–70.

Gillespie GR, Scroggie MP, Roberts JD, Cogger HG, Mahoney MJ, and McDonald KR (2011) The influence of uncertainty on conservation assessments: Australian frogs as a case study. *Biological Conservation* 144: 1516–1525.

Godfrey MH and Godley BJ (2008) Seeing past the red: flawed IUCN global listings for sea turtles. *Endangered Species Research* 6: 155–159.

Greenway JC (1958) *Extinct and Vanishing Birds of the World*. New York: ACIWP.

Harper F (1945) *Extinct and Vanishing Mammals of the Old World*. New York: ACIWP.

Harting JE (1880) *British Animals Extinct Within Historic Times. With Some Account of British Wild Cattle*. London: Trübner.

Heard M, Smith KF, and Ripp K (2011) Examining the evidence for chytridiomycosis in threatened amphibian species. *PLoS ONE* 6: e23150.

Hilton-Taylor C, Pollock CM, Chanson JS, Butchart SHM, Oldfield TEE, and Katariya V (2009) State of the world's species. In: Vié JC, Hilton-Taylor C, and Stuart SN (eds.) *Wildlife in a Changing World – An Analysis of the 2008 IUCN Red List of Threatened Species*, pp. 15–41. Gland, Switzerland: IUCN.

Hoffmann M, Brooks TM, da Fonseca GAB, *et al.* (2008) Conservation planning and the IUCN Red List. *Endangered Species Research* 6: 113–125.

Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science* 330: 1503–1509.

Hoffmann M, Belant JL, Chanson JS, *et al.* (2011) The changing fates of the world's mammals. *Philosophical Transactions of the Royal Society of London B –Biological Sciences* 366: 2598–2610.

Hornaday WT (1913) *Our Vanishing Wildlife*. New York: New York Zoological Society.

Hudson WH (1894) *Lost British Birds*. London: Society for the Protection of Birds.

Isaac NJB, Turvey ST, Collen B, Waterman C, and Baillie JEM (2007) Mammals on the EDGE: Conservation priorities based on threat and phylogeny. *PLoS ONE* 2: e296.

IUCN (1994) *IUCN Red List categories*. Gland, Switzerland: IUCN Species Survival Commission, IUCN.

IUCN (2001) *IUCN Red List Categories and Criteria. Version 3.1*. Gland, Switzerland and Cambridge, United Kingdom: IUCN Species Survival Commission, IUCN.

IUCN (2003) *Guidelines for Application of IUCN Red List Criteria at Regional Levels. Version 3.0*. Gland, Switzerland, and Cambridge, United Kingdom: IUCN Species Survival Commission, IUCN.

IUCN (2010) *IUCN Red List Indices (RLI) for European Terrestrial Vertebrates: Mammals, Amphibians and Reptiles*. Gland, Switzerland: Technical Report.

IUCN (2011) *IUCN Red List of Threatened Species. Version 2011.1*. Available at: http://www.iucnredlist.org.

IUCN SPSC (Standards and Petitions Subcommittee) (2010) *Guidelines for using the IUCN Red List Categories and Criteria. Version 8.1*. Available at: http://www.iucnredlist.org/documents/RedListGuidelines.pdf.

Jaffre T, Bouchet P, and Veillon J-M (1998) Threatened plants of New Caledonia: Is the system of protected areas adequate? *Biodiversity and Conservation* 7: 109–135.

James A, Gaston KJ, and Balmford A (1999) Balancing the Earth's accounts. *Nature* 401: 323–324.

James A, Gaston KJ, and Balmford A (2001) Can we afford to conserve biodiversity? *BioScience* 51: 43–52.

Johnson E (1937) List of vanishing Gambian mammals. *Journal of the Society for the Preservation of the Fauna of the Empire* New Series 31: 62–66.

Köppel C, Jansen F, Burton J, Schnittler M, and Hirneisen N (2003) A statistical survey on European red lists. In: De longh HH, Bánki OS, Bergmans W, and van der Werff ten Bosch MJ (eds.) *The Harmonization of Red Lists for Threatened Species in Europe*, pp. 59–75. Leiden: The Netherlands Commission for International Nature Protection.

Lameroux J, Akçakaya HR, Bennun L, *et al.* (2003) Value of the IUCN Red List. *Trends in Ecology and Evolution* 18: 214–215.

Lande R (1998) Genetics and demography in biological conservation. *Science* 241: 1455–1460.

Langhammer PF, Bakarr MI, Bennun LA, *et al.* (2007) *Identification and GAP Analysis of Key Biodiversity Areas: Targets for Comprehensive Protected Area Systems*. Gland, Switzerland: IUCN.

Li Y and Wilcove DS (2005) Threats to vertebrate species in China and the United States. *BioScience* 55: 147–153.

Lunney D, Curtin A, Ayers D, Cogger HG, and Dickman CR (1996) An ecological approach to identifying the endangered fauna of New South Wales. *Pacific Conservation Biology* 2: 212–231.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Washington DC: World Resources Institute.

Mace G, Masundire H, Baillie J, *et al.* (2005) Biodiversity. In: Hassan R, Scholes R, and Ash N, (eds.) *Millennium Ecosystem Assessment: Current State and Trends: Findings of the Condition and Trends Working Group. Ecosystems and Human Well-being* vol. 1. pp. 77–122. Washington DC: Island Press.

Mace GM, Collar NJ, Gaston KJ, *et al.* (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology* 22: 1424–1442.

Mace GM, Cramer W, Díaz S, *et al.* (2010) Biodiversity targets after 2010. *Current Opinion in Environmental Sustainability* 2: 3–8.

Mace GM and Lande R (1991) Assessing extinction threats: Toward a reevaluation of IUCN threatened species categories. *Conservation Biology* 5: 148–157.

Mace GM, Possingham HP, and Leader-Williams N (2007) Prioritizing choices in conservation. In: Macdonald DW and Service K (eds.) *Key Topics in Conservation Biology*, pp. 17–34. Oxford, United Kingdom: Blackwell Publishing.

Margules CR and Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.

Master LD (1991) Assessing threats and setting priorities for conservation. *Conservation Biology* 5: 559–563.

Master L, Faber-Langendoen D, Bittman R, *et al.* (2009) *NatureServe Conservation Status Assessments: Factors for Assessing Extinction Risk*. Arlington, VA: NatureServe.

May RM, Lawton RH, and Stork NE (1995) Assessing extinction rates. In: Lawton JH and May RM (eds.) *Extinction Rates*, pp. 1–24. Oxford, UK: Oxford University Press.

McGeoch MA, Butchart SHM, Spear D, *et al.* (2010) Global indicators of biological invasion: Species numbers, biodiversity impact and policy responses. *Diversity and Distributions* 16: 95–108.

Meynell P-J (2005) Use of IUCN Red Listing process as a basis for assessing biodiversity threats and impacts in environmental impact assessment. *Impact Assessment and Project Appraisal* 23: 65–72.

Miller RM, Rodríguez JP, Aniskowicz-Fowler T, *et al.* (2006) Extinction risk and conservation priorities. *Science* 313: 441.

Miller RM, Rodríguez JP, Aniskowicz-Fowler T, *et al.* (2007) National threatened species listing based on IUCN criteria and regional guidelines: Current status and future perspectives. *Conservation Biology* 21: 684–696.

Millsap BA, Gore JA, Runde DE, and Cerulean SI (1990) Setting priorities for the conservation of fish and wildlife species in Florida. *Wildlife Monographs* 111: 1–57.

Myers N (1993) Questions of mass extinction. *Biodiversity and Conservation* 2: 2–17.

O'Grady JJ, Burgman MA, Keith DA, *et al.* (2004) Correlations among extinction risks assessed by different systems of threatened species categorization. *Conservation Biology* 18: 1624–1653.

Ocock J, Baasanjav G, and Baillie JEM (eds.) (2006) *Mongolian Red List of Fishes. Regional Red List Series, vol. 3.* London: Zoological Society of London.

Oldfield S, Lusty C, and MacKinven A (1998) *The World List of Threatened Trees.* Cambridge: World Conservation Press.

Ortiz-Catedral L, Ismar SMH, Baird K, Brunton DH, and Hauber ME (2009) Recolonization of Raoul Island by Kermadec red-crowned parakeets *Cyanoramphus novaezelandiae cyanurus* after eradication of invasive predators, Kermadec Islandsarchipelago, New Zealand. *Conservation Evidence* 6: 26–30.

Osieck ER and Mörzer Bruyns MF (1981) *Important Bird Areas in the European Community.* Cambridge: International Council for Bird Preservation.

Owens IPF and Bennett PM (2000) Ecological basis of extinction risk in birds: Habitat loss versus human persecution and introduced predators. *Proceedings of the National Academy of Sciences of the United States of America* 97: 12144–12148.

Ozimec R, Bedek J, Gottstein S, *et al.* (2009) *Red Book of Croatian cave-dwelling fauna.* Zagreb: Ministry of Culture, State Institute for Nature Protection, Republic of Croatia.

Pimenta BVR, Haddad CFB, Nascimento LB, Gonçalves Cruz CA, and Pombal Jr. JP (2005) Comment on "Status and trends of amphibian declines and extinctions worldwide". *Science* 309: 1999b.

Pimm SL, Russell GJ, Gittleman JL, and Brooks TM (1995) The future of biodiversity. *Science* 269: 347–350.

Pimm SL and Jenkins CL (2010) Extinctions and the practice of preventing them. In: Sodhi NS and Ehrlich PR (eds.) *Conservation Biology For All*, pp. 181–199. New York: Oxford University Press, Inc.

Polidoro BA, Carpenter KE, Collins L, *et al.* (2010) The loss of species: Mangrove extinction risk and geographic areas of global concern. *PLoS ONE* 5: e10095.

Polidoro BA, Livingstone SR, Carpenter KE, *et al.* (2009) Status of the world's marine species. In: Vié JC, Hilton-Taylor C, and Stuart SN (eds.) *Wildlife in a Changing World – An Analysis of the 2008 IUCN Red List of Threatened Species*, pp. 55–65. Gland, Switzerland: IUCN.

Possingham HP, Andelman SJ, Burgman MA, Medellín RA, Master LL, and Keith DA (2002) Limits to the use of threatened species lists. *Trends in Ecology and Evolution* 17: 503–507.

Pressey RL, Humphries CJ, Margules CR, Vane-Wright RI, and Williams PH (1993) Beyond opportunism – key principles for systematic reserve selection. *Trends in Ecology and Evolution* 8: 124–128.

Purvis A, Gittleman JL, Cowlishaw G, and Mace GM (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London B – Biological Sciences* 267: 1947–1952.

Quayle JF, Ramsay LR, and Fraser DF (2007) Trend in the status of breeding bird fauna in British Columbia, Canada, based on the IUCN Red List Index method. *Conservation Biology* 21: 1241–1247.

Regan TJ, Burgman MA, McCarthy MA, *et al.* (2005) The consistency of extinction risk classification protocols. *Conservation Biology* 19: 1969–1977.

Regan TJ, Master LL, and Hammerson GA (2004) Capturing expert knowledge for threatened species assessments: A case study using NatureServe conservation status ranks. *Acta Oecologica* 26: 95–107.

Ricketts TH, Dinerstein E, Bouchers T, *et al.* (2005) Pinpointing and preventing imminent extinctions. *Proceedings of the National Academy of Sciences of the United States of America* 102: 18497–18501.

Robbirt KM, Roberts DL, and Hawkins JA (2006) Comparing IUCN and probabilistic assessments of threat: Do IUCN Red List criteria conflate rarity and threat? *Biodiversity and Conservation* 15: 1903–1912.

Rodrigues ASL, Andelman SJ, Bakarr MI, *et al.* (2004) Effectiveness of the global protected area network in representing species diversity. *Nature* 428: 640–643.

Rodrigues ASL, Pilgrim JD, Lamoreux JF, Hoffmann M, and Brooks TM (2006) The value of the IUCN Red List for conservation. *Trends in Ecology and Evolution* 21: 71–76.

Rodríguez JP (2008) National Red Lists: The largest global market for IUCN Red List Categories and Criteria. *Endangered Species Research* 6: 193–198.

Rodríguez JP, Ashenfelter G, Rojas-Suárez F, García Fernández JJ, Suárez L, and Dobson AP (2000) Local data are vital to worldwide conservation. *Nature* 403: 241.

Rodríguez JP, Rojas-Suárez F, and Sharpe CJ (2004) Setting priorities for the conservation of Venezuela's threatened birds. *Oryx* 38: 373–382.

Rodríguez JP and Rojas-Suárez F (eds.) (2008) Libro Rojo de la fauna Venezolana. 3*a* ed. Caracas, Venezuela: Provita y Shell Venezuela, S. A.

Rondinini C, Stuart S, and Boitani L (2005) Habitat suitability models and the shortfall in conservation planning for African vertebrates. *Conservation Biology* 19: 1488–1497.

Schipper J, Chanson JS, Federica Chiozza, *et al.* (2008) The status of the world's land and marine mammals: Diversity, threat, and knowledge. *Science* 322: 225–230.

Scott P, Burton JA, and Fitter R (1987) Red Data Books: The historical background. In: Fitter R and Fitter M (eds.) *The Road to Extinction*, pp. 1–5. Cambridge, UK and Gland, Switzerland: IUCN/UNEP.

Secretariat of the Convention on Biological Diversity (2010) *Global Biodiversity Outlook* 3. Montréal: Secretariat of the Convention on Biological Diversity.

Shelden KEW, DeMaster DP, Rugh DJ, and Olson AM (2001) Developing classification criteria under the U.S. Endangered Species Act: Bowhead Whales as a case study. *Conservation Biology* 15: 1300–1307.

Sodhi NS, Butler R, Laurance WF and Gibson L (in press) Conservation successes at the micro-, meso- and macroscales. *Trends in Ecology and Evolution.* doi: 10.1016/j.tree.2011.07.002.

Stuart SN, Chanson JS, Cox NA, *et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306: 1783–1786.

Stuart SN, Hoffmann M, and Chanson JS (eds.) (2008) *Threatened Amphibians of the World.* Barcelona, Spain; Gland, Switzerland and Arlington, Virginia, USA: Lynx Editions, IUCN and Conservation International.

Stuart SN, Wilson EO, McNeely JA, Mittermeier RA, and Rodríguez JP (2010) The barometer of life. *Science* 328: 177.

Tirira DG (ed.) (2011) *Libro Rojo de los mamíferos del Ecuador. 2ª ed. Publicación Especial sobre los mamíferos del Ecuador 8.* Quito: Fundación Mamíferos y Conservación, Pontificia Universidad Católica del Ecuador y Ministerio del Ambiente del Ecuador.

Walpole M, Almond RE, Besançon C, *et al.* (2009) Tracking progress toward the 2010 biodiversity target and beyond. *Science* 325: 1503–1504.

Wilcove DS (2010) Endangered species management: The US experience. In: Sodhi NS and Ehrlich PR (eds.) *Conservation Biology For All*, pp. 220–235. New York, USA: Oxford University Press.

Wilson EO (1992) *The Diversity of Life.* Cambridge, Massachusetts: Belknap Press.

Woodroffe R and Ginsberg JR (1998) Edge effects and the extinction of populations inside protected areas. *Science* 280: 2126–2128.

Young B, Byers E, Gravuer K, Hall K, Hammerson G, and Redder A (2011) *Guidelines for Using the NatureServe Climate Change Vulnerability Index 2.1.* Arlington, VA: NatureServe.

Zamin TJ, Baillie JEM, Miller RM, Rodríguez JP, Ardid A, and Collen B (2010) National red listing beyond the 2010 target. *Conservation Biology* 24: 1012–1020.

Zulka KP, Eder E, Höttinger H, and Weigane E (2003) Threat descriptors and extinction risk – the Austrian Red List concept. In: De Iongh HH, Bánki OS, Bergmans W, and van der Werff ten Bosch MJ (eds.) *The Harmonization of Red Lists for Threatened Species in Europe*, pp. 103–110. Leiden: The Netherlands Commission for International Nature Protection.

Timber Industry

S Kellomäki, J Kouki, P Niemelä, and H Peltola, University of Eastern Finland, Finland

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 655–666, © 2001, Elsevier Inc.

Glossary

Biodiversity of forests Variability in genetic structures, species composition, and habitat properties in forest ecosystems.

Boreal forests A forest zone dominated by coniferous species, covering the northern latitudes of North America, the Nordic countries, and Russia.

Forest management Manipulation of the properties of tree populations or modification of the properties of the habitats occupied by these populations.

Sustainable forestry Management and utilization of forest resources in a way that balances human needs with undisturbed functioning of the forest ecosystem, and a long-term maintenance and sustainability of the resources and biodiversity of forest ecosystems.

Timber industry Any kind of forest-based production using timber as a raw material.

Main Features of Boreal Forests, with Special Reference to Finland

Forests currently cover about 4030 million hectare (Mha), or 31% of the world's land surface, and represent the largest land-based ecosystem types in the world. Forest species, along with related species communities and ecosystems, represent a large proportion of the global biological diversity and are of great importance in maintaining the functioning and structure of the biosphere. The boreal forests alone represent an area of 1700 Mha and account for about 20% of the world's industrial timber supplies. Commercially, the most important coniferous species are pine (*Pinus*), spruce (*Picea*), fir (*Abies*), larch (*Larix*), juniper (*Juniperus*), thuja (*Thuja*), and hemlock (*Tsuga*), whereas the most common deciduous species in these forests are aspen (*Populus*), birch (*Betula*), willow (*Salix*), and alder (*Alnus*).

The boreal forests in northern Europe or Fennoscandia (including Norway, Sweden, Finland, and northwestern Russia) are probably the most intensively utilized of all; that is, they form 80% of the total forest area in Europe (1005 Mha) and provide about 40% of the timber used in Europe (**Table 1**). The stocking ($68\ m^3\ ha^{-1}$) and growth rate ($1.1\ m^3\ ha^{-1}\ year^{-1}$) in these forests are substantially lower than in more southerly regions, however, and consequently the timescale in forest production is long, with rotations of 40–160 years according to species and region. Growth and felling are currently well balanced, and the growth may even exceed the needs of the timber industry.

Forestry in northern Europe is mainly based on native tree species which invaded this region postglacially. In Finland, for example, Scots pine (*Pinus sylvestris*) is the dominant species throughout the country (about 45% of total stem volume) wherever nutrient and water supplies are limited at the site. Scots pine can also dominate at more fertile sites, but both Norway spruce (*Picea abies*) (about 37% of total stem volume) and birch (*Betula pendula* and *Betula pubescens*) can win out over it under such conditions. Birch species (about 18% of total stem volume) are most common on the most fertile soils.

In contrast with many developing countries, forest coverage in Fennoscandia has remained at the current level for decades or even increased slightly. Forest utilization in Finland has been occurring for centuries. Toward the late 1800s Finnish forests were intensively used for slash-and-burn cultivation and tar production without any actions to promote new tree generation. As a result, large areas, especially in southeastern Finland, became dominated by relatively young forests 100 years ago. Effective prevention of slash-and-burn cultivation and forest fires together with active timber-oriented

Table 1 Main features of the Boreal Forest Zone in Europe

Countries representing boreal forests in Europe	*Examples of tree species of importance to forestry*	*Main parameters for forests*	
Finland, Sweden, Norway excluding southwest part, northwestern Russia	*Pinus sylvestris*, *Picea abies*, *Betula pendula*, *Betula pubescens*, *Populus tremula*, *Pinus contorta*, *Larix sibirica*	Area (Mha)	808
		Total volume(Mm^3)	55,031
		Total growth(Mm^3)	878
		Total cut (Mm^3)	643
		Balance (Mm^3)	+235

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00368-3

management have led to an increase in the areas of closed-canopy forest during the past century.

In general, the number of species per unit area is low at northern latitudes. This also holds for Finland, where the estimated total number of species is about 50,000, of which about 43,000 are known. The reason for the relatively low total number of species is the short time that has elapsed since the last glaciation (10,000 years), with the consequence that immigration is still occurring. These species and the subsequent biodiversity involve a large contribution from natives of the eastern taiga (e.g., flying squirrel, Ural owl, and Siberian jay). Most of these taiga species are connected with spruce forests. Furthermore, the amount of dead or decaying wood is generally high in boreal forests, especially in the early phase of the succession after a fire or windstorm. Consequently, a high proportion of the forest species (20–25%) are dependent on dead wood (800 coleopterans, 1000 dipterans, 1000 fungi, 200 lichens, etc.). Many of these species are also specialized in living on recently burned tree material, whereas a high proportion lives in peatland forests or on mires. Most of these species need a moist and shady habitat.

Dynamics of Boreal Forest Ecosystems and Management Implications

Management: Control of the Structure and Functioning of the Forest Ecosystem

Forests are ecosystems in which trees and other green plants occupy the site and intercept solar energy under the control of climatic and edaphic factors (**Figure 1**). The interaction of populations with their environments forms complex food webs, in which the solar energy flows from producers to consumers. The links between the organisms are the key to the management of a forest ecosystem. Proper manipulation of forest dynamics will allow the production of timber or other items, with the management needs being related to the management objectives and to the changes in the structural properties of the forest ecosystem occurring with time. The management history and past structure of a forest ecosystem may also have a crucial impact on future management. Production in a forest can be maintained and increased by manipulation of the genetic properties of the tree populations or modification of the properties of the habitats occupied by these populations. This takes place by controlling the long-term functional and structural development of the forest ecosystems (the forest succession) in order to induce them to produce the items defined in the management goals. In other words, any kind of forest-based production is a result of environment, genotype, and the interaction between these.

Successional Dynamics of Forest Ecosystems and Management Implications

The long-term growth and development of tree populations are subject to disturbances due to wind, fire, and attacks by pests and pathogens. Under boreal conditions in particular, such disturbances cause cycling in the dynamics of the forest ecosystem by returning the succession to an earlier phase. For example, a fire that kills old trees provides space for regeneration. For larger forest areas, this implies that forest landscapes contain a mosaic of tree stands of varying species composition and age, from open areas and saplings to mature trees.

During the succession, each stand or part of the mosaic undergoes a gradual change involving alterations in species composition and the accumulation of organic matter in trees and soils (**Figure 2**). Many broad-leaved tree species can make better use of the abundant supply of resources at fertile sites than can coniferous trees. Broad-leaved species typically dominate the initial phase of the succession after a fire, and

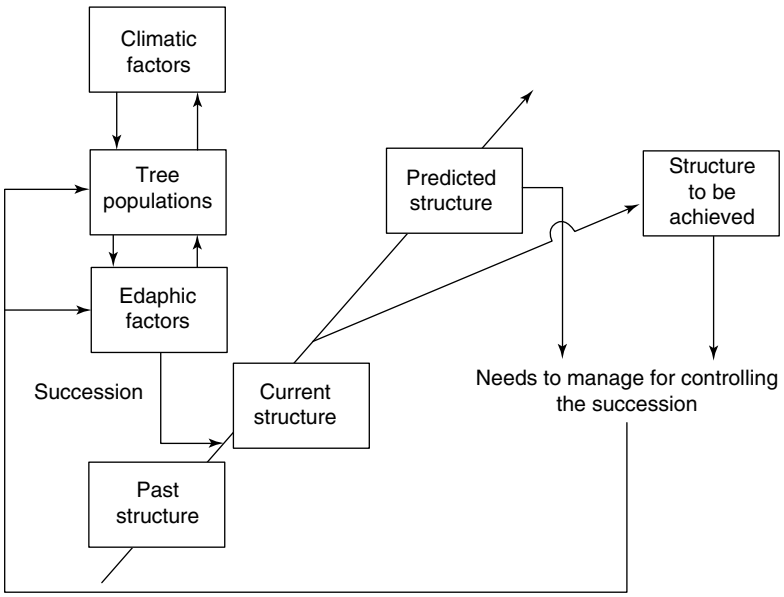


Figure 1 Schematic presentation of the structure and functioning of a forested ecosystem as an interaction between climatic and edaphic factors and the populations of organisms, with implications for management.

wherever site fertility is sufficiently high the changing conditions will later start to support shade-tolerant conifers more than the broad-leaved trees. This brings an invasion of conifers such as Norway spruce to the site, which finally replaces the broad-leaved species unless a disturbance causes the succession to return to an earlier phase. These disturbances and the dynamics of the tree species result in substantial variability in the mass of organic matter in forest ecosystems. This cycling determines the management of the forest ecosystem and the timber yield obtained.

In terms of the energetics of the ecosystem, the energy intercepted by the trees and other green plants and the energy used at different trophic levels in the ecosystem's food webs is balanced by means of enhanced formation of dead and decaying wood and other organic matter. In this context, the food webs provide complex feedbacks through which the populations of the various species interact, with consequent control of the ecosystem dynamics. The food webs tend to increase in complexity during the succession.

Disturbance Dynamics of Forest Ecosystems and Management

The disturbance dynamics of forest ecosystems provide the framework in which management occurs. Two categories of disturbance may be distinguished which drive the succession: autogenic and allogenic (hence, we may speak of an autogenic or allogenic succession). Allogenic and autogenic successions are typical processes operating simultaneously in boreal forests.

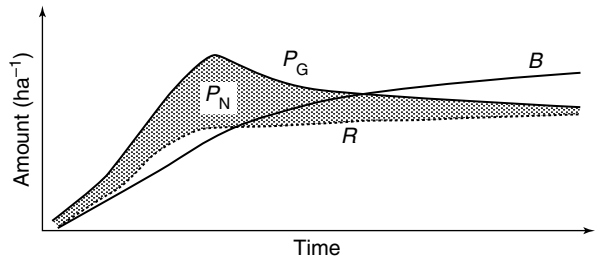


Figure 2 Schematic presentation of the course of total growth (P_G), net growth (P_N), respiration loss (R), and the accumulation of mass (B) in trees during the succession.

In the autogenic succession, the structural dynamics of the forest ecosystem are driven by the regeneration, growth, and death of single trees, that is, minor disturbances related to their life cycle (gap-phase dynamics). The deaths of single trees and the consequent gaps in the canopy release resources and enhance regeneration and growth. The concept of autogenic succession is applied in the form of selection forestry or uneven-aged management, in which single trees are the basic object of management. This system is widespread in the temperate zone, in which autogenic succession predominates in the dynamics of forest ecosystems.

Allogenic succession refers to dynamics driven mainly by major disturbances induced by fires, gales, and excessive snowfall. Under these conditions many trees may die simultaneously, releasing space for new trees to regenerate and grow to form a population with individuals of more or less the same age (cohort-phase dynamics). In allogenic succession the stand is the basic unit driving the structural dynamics of the forest ecosystem. The concept is applied to management in the form of standwise forestry, or even-aged management, in which single stands form the basic object of management. This system is commonly employed in the boreal zone, in which allogenic succession predominates in the dynamics of forest ecosystems.

Management involves the use of a variety of measures to control the supply of resources and the properties of tree populations. The measures used to control the supply of resources may be classified into two categories in relation to the disturbances driving the succession methods liberating resources to a large extent (allogenic measures related to the allogenic succession) and methods liberating resources to a lesser extent (autogenic measures related to the autogenic succession) (**Table 2**).

Allogenic measures mainly control the properties of the soil system, and subsequently the nutrient cycle and soil moisture, with the greatest enhancement of growth and regeneration of pioneer species being achieved through disturbance of the soil surface. Site preparation refers to

Table 2 Schematic presentation of the effects of given management measures on site properties and tree populations[a]

Measure	*Effects on site properties and populations*				
	Nutrients	*Moisture*	*Temperature*	*Soil physics*	*Species*
Allogenic measures					
Site preparation	+++	++	+++	+++	+++
Ditching	++	+++	++	++	+++
Prescribed burning	+++	+	++	++	+++
Fertilization	+++				+++
Regenerative cutting	++	++	++	+	+++
Autogenic measures					
Selection for regeneration	+	+	+	+	+
Precommercial thinning	+	+	+	+	+++
Commercial thinning	++	++	+	+	++
Pruning	+	+	+		+

[a] +, Slight effect; ++, moderate effect; +++, substantial effect.

mechanical measures used to reduce the effects of the ground vegetation on saplings and to adapt the physical and chemical properties of the soil. In this context, ditching is aimed only at lowering the groundwater level and reducing excess soil moisture. Prescribed burning affects the chemical properties of the soil and eliminates ground vegetation interference with sapling growth. Fertilizing implies the addition of nutrients at the site, leading to a substantial increase in the availability of nutrients for tree growth. Regenerative felling refers to terminal felling aimed at promoting reforestation of the site through natural regeneration by means of natural seeding or artificial regeneration by means of the sowing of seed or planting.

Autogenic measures are mainly represented by spacing, leading to enhancement of growth and regeneration among both pioneer and climax species. This is achieved through the thinning of sapling stands (precommercial thinning) and more mature stands (commercial thinning). The tending of sapling stands is aimed at proper spacing, as is also true of thinnings, and it also implies the elimination of unwanted tree species. Pruning involves the removal of dead or living branches from the lower crown in order to increase the amount of knot-free wood obtainable from the stem. Selection for regeneration refers to the removal of single mature trees or small groups of mature trees in order to create canopy gaps in which seedlings can become established and grow.

Management of Forests for Optimal Productivity and Timber Production

Among the key issues to be addressed in management is the optimal distribution of tree species and stand age (or tree size) over the forest area in order to maximize long-term growth. Assume a forest area of 100 ha to be divided into six compartments (stands) allocated to initial age classes (i.e., 2, 10, 22, 38, 82, and +82 years) in different ways. Calculations by Kellomäki (1998, pp. 225–226) show growth during the next 10 years to be highest in the stand with an initial age of 38 years, with this growth being five times that in the stand with an initial age of 85 years. Consequently, a distribution skewed toward dominance by old stands represents the highest level of stocking but about 15% less growth than that for a normal distribution (**Figure 3**). Where young stands are dominant, the total growth is only one-third of that for a normal distribution, whereas an even distribution gives 70–80% of the total growth achieved with a normal distribution.

In the short term, the maximization of timber production would thus require an age structure representing a normal distribution, that is, a forest area that is fragmented into single stands of variable age and stocking with the inevitable impact that this will have on species with a habitat corresponding to mature or old-growth forests. In the long term, this age distribution will gradually lead to dominance by old stands, with increasing stocking and decreasing growth – a situation that is less than optimal for timber production. Terminal felling of stands previously of intermediate age but now mature will lead to a shift in the age distribution toward dominance by young stands with low stocking and absolute growth but high relative growth. The maximization of timber production over several rotations in a sustainable manner thus requires that the forest area should be divided into sections representing a mix of sapling, pole, and mature stands in such a way that the overall age distribution implies dominance by pole stands with high stocking and high growth. The forest area will still be fragmented in this case, which will mean major impacts on species with a preference for mature or old-growth forests.

Impacts of Timber-Oriented Management on Biodiversity: Comparison between Managed and Natural Forest Areas

Impact Mechanisms

Among the key issues in managing biodiversity in commercial forests is the maintenance of the functional and structural properties of the forest ecosystem, which will create and maintain heterogeneity among the habitats available for species to occupy. Two levels of heterogeneity may be distinguished: allogenic heterogeneity between stands and autogenic heterogeneity within stands. These concepts are related to allogenic and autogenic succession. Allogenic heterogeneity creates a landscape mosaic, whereas autogenic succession creates further heterogeneity within each element of the mosaic in terms of an even-aged stand structure or multilayered tree canopies (**Table 3**). The contribution of allogenic and autogenic heterogeneity to the total heterogeneity of the landscape is probably multiplicative rather than additive.

The total impact of forestry on the availability of suitable habitats is difficult to predict. The optimization of age distribution for timber production may increase allogenic heterogeneity and fragmentation of the landscape. In particular, the remaining patches of mature or old-growth forests may be too small to sustain viable populations or too isolated for recolonization of areas where a population has temporally become extinct. Simultaneously, management may reduce autogenic heterogeneity in two ways. First, it may avoid or even curb successional processes that give rise to within-stand structural components and spatial patterns that are important for forest biodiversity (**Table 2**). Second, regeneration practices and precommercial and commercial thinnings can reduce or even remove autogenic heterogeneity. This tendency is quite obvious in standwise management, especially if pure single tree populations in even-aged stands are preferred for reasons connected with timber production.

The critical question in the previous context is to what extent it is possible to maintain in commercial forests the processes, structural components, and patterns typical of mature or natural forests as listed in **Table 3**. Many of these elements are interrelated, and practically all are somehow related to different types of disturbances and the subsequent succession. The amount of coarse woody debris is highest during the early successional phases in natural forests, whereas in managed forests no dead or dying trees are left at this stage. The scarcity of such debris (dead and decaying wood) is generally accepted as the main reason for the disappearance of many species from the boreal forests that are being managed

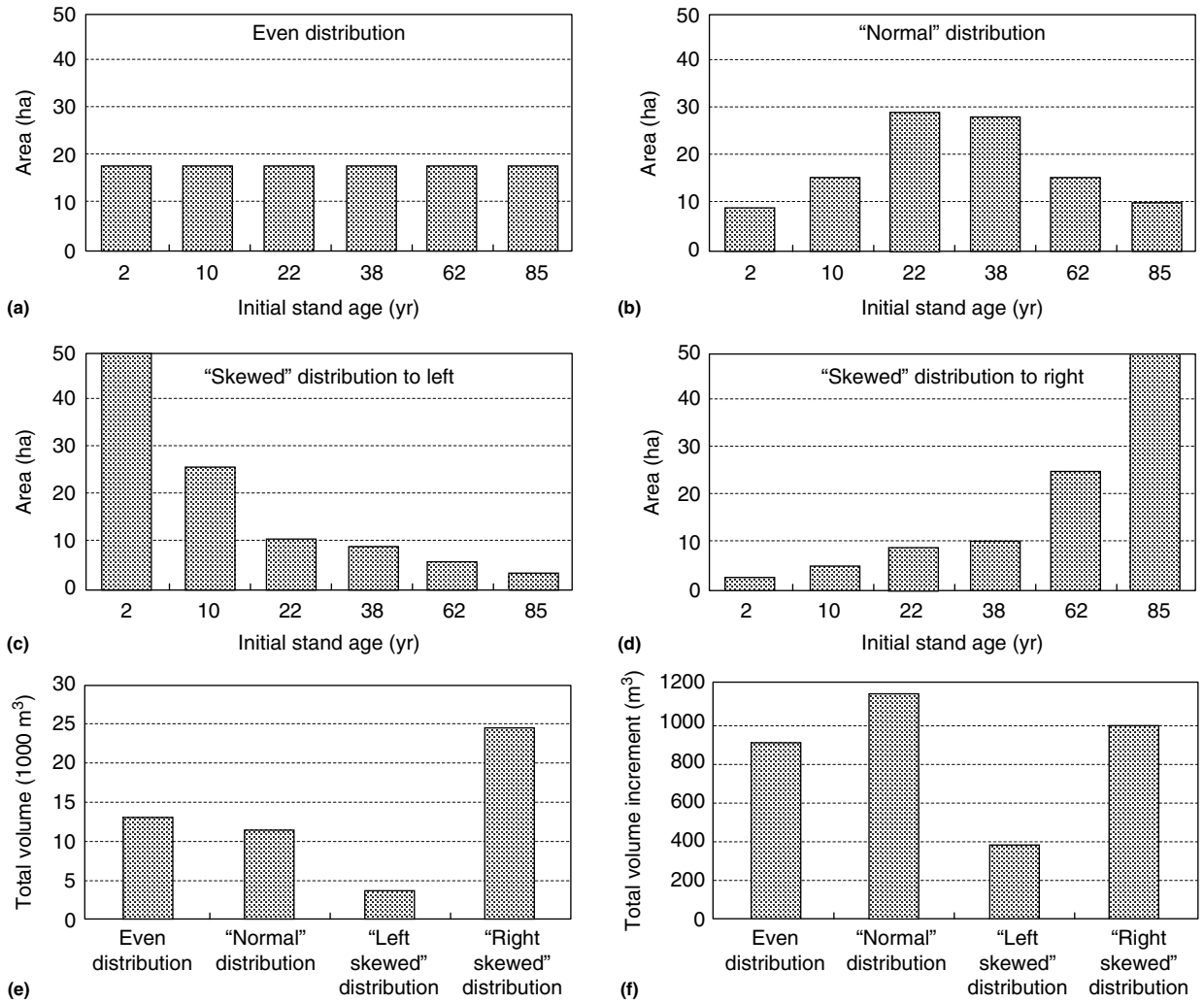


Figure 3 Computational example of how the age class distribution of tree stands can influence the productivity of a forest area. Distribution of tree stands with an overall area of 100 ha: (a) no single age class dominant (even distribution), (b) dominated by intermediate age classes (normal distribution), (c) dominated by young age classes (distribution skewed to the left), and (d) dominated by old age classes (distribution skewed to the right). (e and f) Stocking and growth, respectively, over the forest area for the various distributions.

for timber production. In Finland, for example, more than 20% of the species with a preference for forest habitats are dependent on dead wood (e.g., there are about 1000 fungi and about 800 beetles whose occurrence is related to the presence of such material). This implies that 36% of the total number of species threatened by forestry practices are ones whose occurrence is related to the presence of dead wood.

It was not known for certain until recently whether species depending on dead wood require closed-canopy old-growth habitats or whether they could make use of younger successional phases if a sufficient amount of dead wood was available. Recent evidence suggests (**Figure 4**) that the latter may be possible and that many threatened species may survive in managed forests if certain properties typical of natural forests are maintained or even increased. However, many threatened species have highly specialized habitats, for example, sand ridges, riverbanks, or large, overmature aspen trees (*Populus tremula*). The majority of these key habitats are of marginal importance for timber-oriented forestry since they represent sites of low productivity or are quite small in extent. Given careful planning and management, these key habitats could be set aside, which could have a major impact on the conservation of threatened species which prefer forest habitats. Along with measures to increase dead wood in commercial forests, this would provide new possibilities for reducing the impact of timber-oriented management on threatened species.

Properties of Forests under Intensive and Extensive Management: An Example

The impacts of timber-oriented management are exemplified in the following by referring to a comparison of selected beetle populations in an intensively managed forest area and an extensively managed one with a quite different management history (Siitonen and Martikainen, 1994; Siitonen *et al.*, 1995). This material also provides a unique opportunity to contrast our current understanding with experimental data representing the long-term effects of forestry on the properties

Table 3 Structural components, spatial patterns, and processes important for biodiversity in natural forests but lacking or altered in managed stands

Processes
Postfire succession
Succession with tree species replacement
Self-thinning
Gap formation
Snag and log formation
Decomposition of coarse woody debris
Structural components
Very old pine and spruce trees
Old broad-leaved trees, particularly aspen and sallow
Trees with abundant growth of epiphytic lichens
Broken, stag-headed, and leaning trees
Trees with holes and cavities
Dead standing trees (snags)
Fire-scarred trees, snags, and stumps
Large fallen logs in various stages of decomposition
Spatial patterns
A developed understory of tree saplings and shrubs
Mixed stands with both conifers and broad-leaved trees
Uneven-aged stand structure
Multilayered tree canopies
Patchy distribution of trees

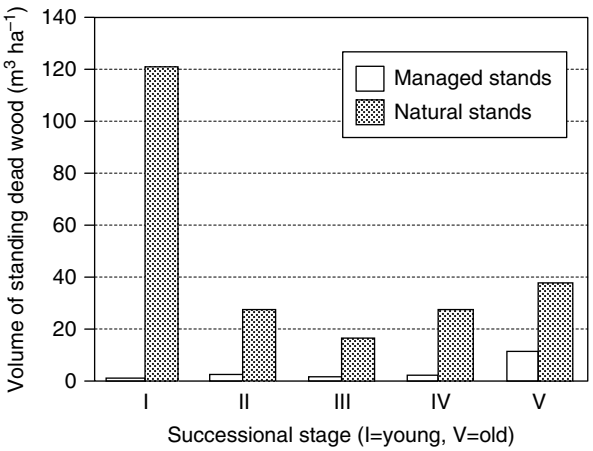


Figure 4 Volume of standing dead wood in different successional stages of managed and natural Norway spruce stands (data provided by Anneli Uotila, University of Eastern Finland).

and biodiversity of boreal forests. The studies of Siitonen and Martikainen (1994) and Siitonen *et al.* (1995) are widely used in the following discussion.

The two areas are located on either side of the border between Finland and Russia (62 or 63° N, 30–32° E) (**Figure 5**). Before World War II, both areas belonged to Finland and were utilized and managed in similar ways. The border, created in 1944, split them in a biogeographically random manner.

The forests on the Finnish and Russian sides of the border have been used and managed quite differently for the past 50 years. Those in Finland have been subject to intensive use and a form of management that has included balancing growth

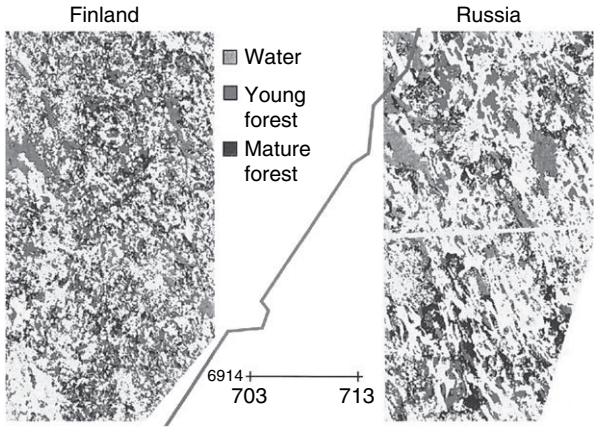


Figure 5 Landsat image (27 May 1992) of part of the forest area on both sides of the border between Finland and Russia. Water is indicated in blue, young forest in red, and mature forest in green (Siitonen *et al.*, 1995).

with felling, systematic regeneration in the form of plantations established in clear-felled areas, or natural regeneration by means of seed trees or shelter wood fellings in favor of coniferous species. Precommercial thinnings and intermediate thinnings have been used to reduce the role of deciduous trees, and the latter have also been used to harvest the growth and yield that would otherwise have been lost in the natural self-thinning associated with the succession. Furthermore, the length of the rotation has been determined with the aim of maximizing the timber yield (80–100 years), which has further reduced the formation of dead and decaying wood. Additional reductions in dead wood have been achieved by effective fire fighting and the logging of trees blown down by storms or broken by heavy snowfall. The small-scale ownership of forest plots has further enhanced the fragmentation of the forests as each owner has aimed to maximize her or his timber yield and income.

On the Russian side of the border, clear-felled areas were not systematically regenerated, and thus fertile sites were occupied by mixtures of coniferous and deciduous trees. Since no precommercial thinnings or intermediate thinnings took place, the dominance of deciduous trees throughout the rotation was substantially higher than on the Finnish side. Later, dying trees were not logged in connection with thinnings, with the consequence that there is now a large amount of dead and decaying wood present. Furthermore, a low utilization rate of stems and the leaving of deciduous trees of low commercial value in the clear-felled areas were characteristic of logging practices. Simultaneously, uncontrolled forest fires were common and timber destroyed in other natural disturbances (e.g., storms and heavy snowfalls) remained unharvested. In other words, human intervention in the regeneration and growth of the forests has been substantially less than that on the Finnish side of the border.

Landscape and Local Patterns in Relation to Management

Systematic field samples from both sides of the border and analyses of satellite images indicated that the mean area of

individual tree stands was quite similar but old-growth stands were substantially larger on the Russian side. Furthermore, the largest areas covered by homogeneous tree stands were larger on the Russian side. Most of the large stands on the Russian side were the result of extensive forest fires. However, the number of separate stands relative to the total area was larger on the Finnish side, that is, the Finnish forest landscape is more fine grained and fragmented. This difference is quite evident even on the satellite image, which indicates the dominance of a coarse-grained landscape on the Russian side (**Figure 5**).

Closer analysis of the properties of the forests shows that deciduous tree species are more common in the Russian forests than in the Finnish ones throughout the area (**Figure 6**). Mature aspen in particular has been effectively eradicated from the Finnish forests because of its tendency to host a serious fungus disease that affects sapling stands dominated by pine. The most striking differences, however, concern sapling stands on fertile sites, which are dominated by Scots pine and Norway spruce in Finland but both Pendula and Pubescent birch in Russia. Coniferous tree species naturally dominated the sapling stands on poor sites in Finland, but they also did so on fertile sites, whereas in Russia deciduous species exceeded 20% of the total number of stems per hectare even on poor sites in many cases. In other words, the forest cover more frequently represented a mixture of tree species on the Russian side than on the Finnish side.

In terms of dead and decaying wood, the difference between the two areas was clear; that is, the mean amount of dead wood in the Finnish forests was 4 $m^3\,ha^{-1}$ (range 0.1–26 $m^3\,ha^{-1}$) and that in the Russian forests was 29 $m^3\,ha^{-1}$ (range 0.1–213 $m^3\,ha^{-1}$) (**Figure 7**). It is worth noting that the amount of coarse woody debris is high even in sapling stands in the Russian forests (up to 79 $m^3\,ha^{-1}$), mainly as a result of forest fires. Here, the dead wood represents mainly dead trees (standing and lying), whereas in the Finnish forests it mainly comprises logging residues (stumps, branches, and tops of stems). These patterns held good over the whole range of variability in fertility and the age of sites and stands.

Beetle Fauna in Forests under Intensive and Extensive Management

Siitonen *et al.* (1995) assessed the impact of forest management on biodiversity in terms of the number of beetles species caught with window traps operating over one growing season. The total yield was about 29,000 individual beetles representing 623 species. There were no major differences in the total number of species or in the main groups of species, except for the species preferring dead wood, with the consequence that the total number of saproxylic species was 179 in the Finnish forests and 213 in the Russian ones (**Table 4**).

It was found that 17 of the 20 most frequent species with a preference for dead wood were more frequent on the Russian side of the border. These represented mainly species with a preference for decaying birches (e.g., *Trichius fasciatus*, *Hylocoetus dermestoides*, *Leptura quadrifasciata*, and *Anaspis arctica*) or for the final phases of decay in coniferous trunks

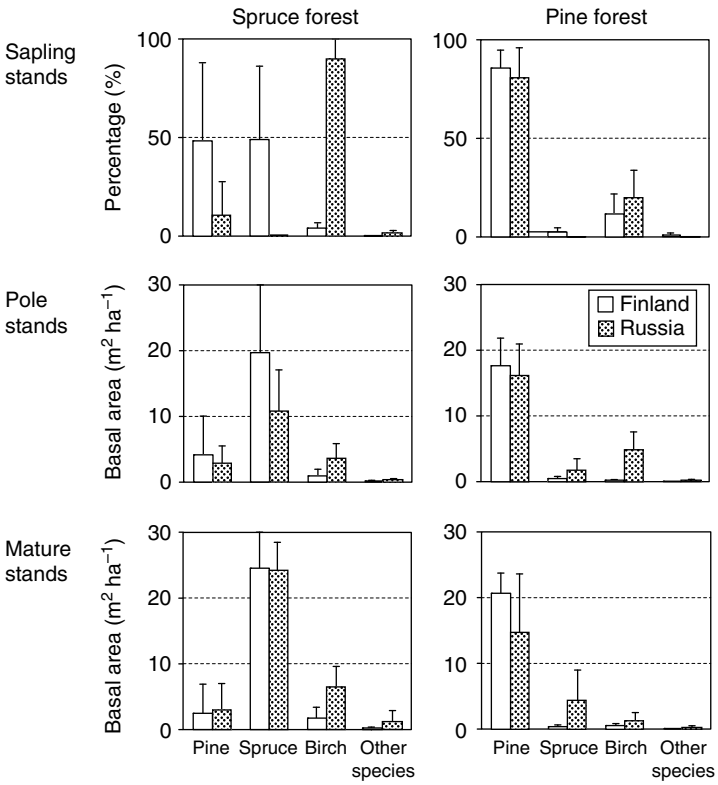


Figure 6 Tree species composition of sapling stands, pole stands, and mature stands in areas on the Finnish and Russian sides of the border (Siitonen *et al.*, 1995).

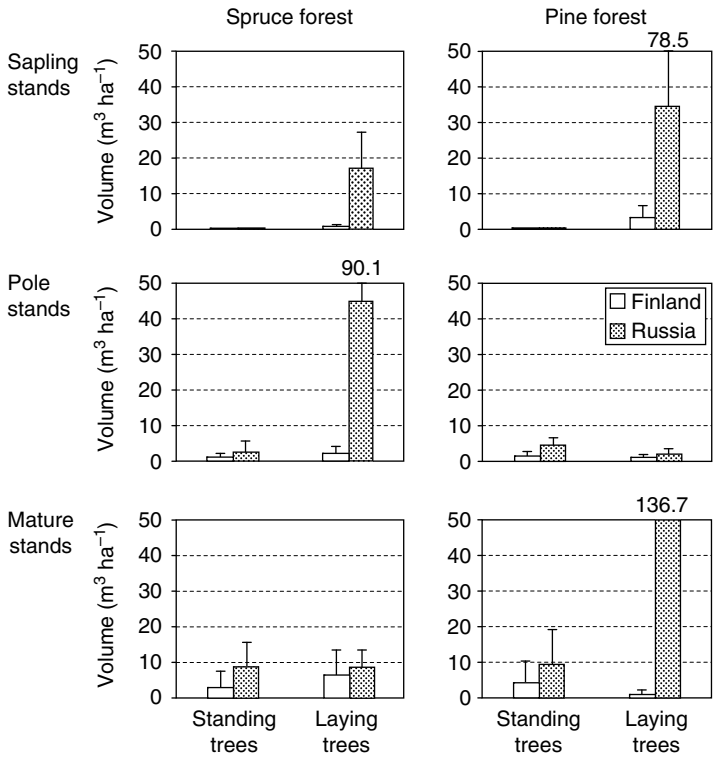


Figure 7 Volume of dead wood as a function of tree species and maturity of tree stands on the Finnish and Russian sides of the border. Spruce forest represents higher fertility than pine forest (Siitonen *et al.*, 1995).

Table 4 Total numbers of Beetle individuals and species as a function of their dependence on the occurrence of dead wood (Siitonen *et al.*, 1995)

Parameter	*Finnish forests*	*Russian forests*	*Total*
No. of individuals			
Obligatory saproxylic	3645	4964	8609
Facultative saproxylic	1893	2202	4095
Other species	8657	7663	16,320
Total	14,195	14,829	29,024
No. of species			
Obligatory saproxylic	179	217	253
Facultative saproxylic	55	58	64
Other species	245	232	307
Total	479	507	623

(e.g., *Anoplodera virens*, *Ampedus balteatus*, *Ampedus tristis*, and *Hadrobregmus pertinax*). The numbers of common and fairly common species in both countries (found consistently less than 25 times in Finland) were very similar, but the number of the rarest species (found consistently less than 12 times in Finland) was substantially greater on the Russian side (**Table 5**).

The occurrence of species with a preference for dead wood was related to the amount of dead wood present in the tree stands, as shown in **Figure 8**. However, there was no correlation between the occurrence of other species and the amount of dead wood. The numbers of specimens of other species were inversely correlated with the basal area of trees in the stand, however, which implies that there are more beetle species in sapling stands and pole stands than in mature or old-growth stands, regardless of tree species and site fertility.

Conclusions

Forests are a natural resource of prime importance, and they provide human communities with tangible items such as timber and wildlife and intangible benefits such as scenic beauty, wind reduction, and urban noise abatement. A unique feature of forest production is that its control requires only a small external input. Natural processes exert the greatest control, in the form of the biological diversity that controls the functioning and structure of forested ecosystems. The management and utilization of forest resources in a way that balances human needs with the undisturbed functioning of the forest ecosystem is the key to a form of sustainable forestry which provides ample space for maintaining biodiversity in the world's forests.

The current biodiversity of the boreal forests is a result of long-term evolution and changes in the environment and is thus related to the successional dynamics of the forest ecosystem. The long-term growth and development of tree populations is subject to disturbances due to wind, fire, and attacks by insects and pests, and under boreal conditions in particular, these disturbances cause a cyclic effect in the dynamics of the forest ecosystem by returning it to earlier phases

Table 5 Rarest beetle species with a preference for dead wood presented in order of abundance based on the total number of individuals (Siitonen *et al.*, 1995)

Species	*Finnish forests*	*Russian forests*
Found ≤ 12 times in Finland		
Pseudeuglenes pentatomus	–	1
Phymatura brevicollis	–	1
Hylis foveicollis	–	1
Enicmus planipennis	–	1
Enicmus apicalis	–	1
Epuraea deubeli	–	1
Epuraea longula	–	1
Cryptophagus confusus	–	1
Corticaria obsoleta	–	1
Found ≤ 25 times in Finland		
Stagetes borealis	1	30
Hylis procerulus	–	19
Euplectus fauveli	–	8
Lacon fasciatus	–	5
Triplax rufipes	–	4
Biploporus minutus	–	4
Mycetophagus quadripustulatus	–	3
Atrecus longiceps	–	3
Episernus angulicollis	–	2
Cryptophagus subdepressus	–	2
Corticaria crenicollis	–	2
Cis dentatus	–	1
Callidium aeneum	–	1
Epuraea muehlii	–	1
Orthocis linearis	1	–
Total	2	94

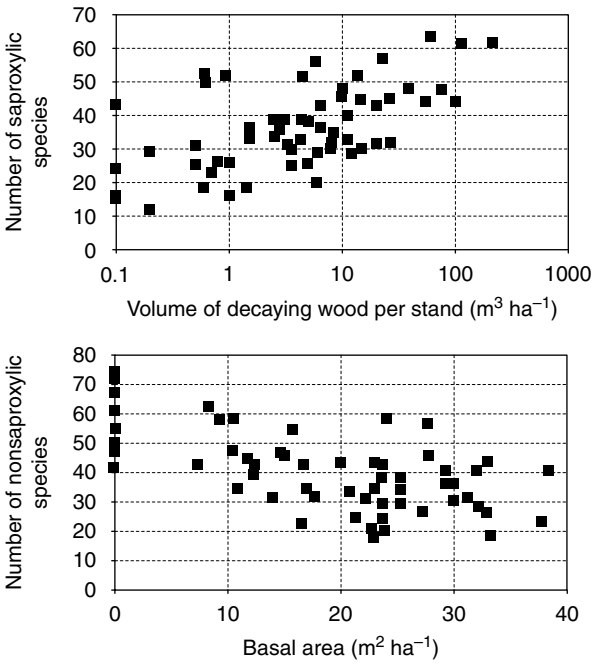


Figure 8 Correlation between the number of saproxylic species and the volume of decaying wood ($r=0.548$, $d.f.=58$, linear scale after logarithmic transformation of the x-axis) and correlation between the number of nonsaproxylic species and the basal area of the tree population ($r=-0596$, $d.f.=58$, linear scale) (Siitonen *et al.*, 1995).

in the succession. For example, a fire that kills old trees provides space for regeneration, enhancement of growth, and the accumulation of mass. Considered over large areas, this means that under natural conditions the boreal forests represent a mosaic of tree stands of varying age from saplings to mature trees and of varying tree species composition. The maximization of growth and timber yield implies that the area of mature or old-growth forests may be smaller than might be desirable in order to conserve the most specialized species.

The forest resources of northern Europe have been utilized intensively during the past 100 years, and this has inevitably had an impact on the current diversity of the forests. The management goal of maximizing the timber yield will inevitably lead to fragmentation of the forests and structural properties which deviate substantially from those arising under the influence of major disturbances. An increasing harvesting rate in terms of regular thinnings, together with a preference for coniferous species and a shorter rotation, will reduce the proportion of deciduous trees and the occurrence of dead and decaying wood. Furthermore, effective fire fighting and measures aimed at increasing the capacity of forests to resist the force of winds and heavy snowfalls will reduce the formation of dead wood and the occurrence of burnt wood, for which many species are specialized.

The previously mentioned tendencies will detract from the conditions which many rare and endangered species need in order to survive, as indicated by the positive correlation between the occurrence of these species and the amount of dead wood in the boreal forests. The increasing amount of dead wood probably indicates that there are more specialized microhabitats available under these conditions than in forests with a small amount of dead wood. These highly specialized species seem to be those affected most by the utilization and management of forest resources. However, the occurrence of species other than those confined specifically to dead wood seems to be quite similar regardless of the management history of the forests.

The protection of threatened biota may be based solely on a network of strictly protected areas. Currently, less than 10% of the productive forestland in Finland is strictly protected, and opportunities to increase this within a short time span seem to be limited. Most of the valuable old-growth areas are already protected. Perhaps a better potential for protecting forest biodiversity and threatened species can be found in managed forests, in which the application of silvicultural practices that aim at restoring the natural properties of younger forests can be developed. Natural characteristics can be promoted quite rapidly in young managed forests since such forests are created continuously along with timber harvesting. However, most of the threatened species have highly specialized habitat requirements. By preserving these habitats (key habitats) in forest management, viable populations of these species can be maintained. Consequently, strictly protected areas and biodiversity-oriented management should be regarded as important complementary elements rather than alternatives when planning and developing new forestry practices (Kouki *et al.*, 2000).

See also: Boreal Forest Ecosystems. Forest Ecology

References

Esseen PA, Ehnström B, Ericson L, and Sjöberg K (1997) Boreal forests. *Ecological Bulletins* 46: 16–47.

Kellomäki S (1998) Forest resources and sustainable management. *Papermaking Science and Technology Book 2*. Jyväskylä, Finland: Gummerus Oy.

Kouki J, Löfman S, Martikainen P, Rouvinen S, and Uotila A (2000) Spatio-temporal patterns in the Fennoscandian forest fragmentation and habitat requirements of wood-associated threatened species. *Scandanavian Journal of Forest Research* 15.

Siitonen J and Martikainen P (1994) Occurrence of rare and threatened insects living on decaying *Populus tremula*: A comparison between Finnish and Russian Karelia. *Scandanavian Journal of Forest Research* 9(2): 185–191.

Siitonen J, Martikainen P, Kaila L, Nikula A, and Punttila P (1995) Kovakuoriaslajiston monimuotoisuus eri tavoin käsitellyillä metsäalueilla suomessa ja Karjalan Tasavallassa. *Metsäntutki-muslaitofesen Tiedonantoja* 564: 43–63. (in Finnish).

UN-ECE/FAO (1992) *The Forest Resources of the Temperate Zones – The Forest Resource Assessments 1990*. New York: United Nations.

Tourism, Role of

Ralf Buckley, Griffith University, Gold Coast, QLD, Australia

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Richard W. Braithwaite, volume 5, pp 667–679, © 2001, Elsevier Inc.

Glossary

Conservation tourism Tourism which yields a net positive benefit for conservation of biodiversity, i.e. where contributions to conservation outweigh environmental impacts.

Ecotourism Commercial tourism which is based on relatively undisturbed natural environments, and which takes steps to minimize environmental impacts, provide environmental education to clients, and contribute to environmental conservation.

Nature-based tourism Tourism where nature provides the primary setting or attraction.

Tourism Travel for leisure, recreation, or vacation; for economic statistics, generally only overnight travel is included.

Visitor Term used by protected area (parks) management agencies to include everyone entering the park, either as an individual or a client of a commercial tour operator.

Introduction

Biodiversity is a tourist attraction. Tourism creates impacts on biodiversity, but can also help conserve it. There is a mass-market wildlife tourism industry which relies on large charismatic icon species, but also a smaller, more sophisticated and specialized sector focusing on rare birds, mammals, reptiles, and plants. Tourism and recreation are the main uses permitted in most of the world's protected areas. Many people visit parks only for recreation; they create impacts and impose management costs, but may also provide a political constituency for protected area agencies. In developing nations, tourism can also be a significant source of conservation funding, in private as well as public reserves. There are also private reserves in developed nations, but these rely more on philanthropy than tourism.

Across both developed and developing nations, the degree of interaction between the tourism industry and the conservation sector is small relative to the overall scale of either. Protected areas are established and funded by governments principally for conservation; the major threats to biodiversity are from large-scale primary and extractive industries, both legal and illegal, and from continued expansion of human populations and residential settlement; and most tourism is urban. Worldwide, it has been estimated that about one-fifth of the global tourism industry operates in outdoor natural environments; and only a tiny fraction of 1% contributes directly to biodiversity conservation. Some of these contributions, however, seem to be critical for particular species and ecosystems (Buckley, 2009, 2010a).

Geography

Tourism has a well-defined though continually changing global geography. Most tourists originate in rich Northern cities; and they travel both to other such cities, and to rural, protected and wilderness areas in both the developed and developing world. The patterns in the numbers of tourists, the activities they do, and the money they spend is determined by a combination of human and physical geographies. In particular, the use of tourism as a conservation tool is highly vulnerable to wars, crime, terrorism, and political instability in destination areas; and also to marginal changes in the economic conditions, currency exchange rates, and relative airfares on different routes.

Both biodiversity and conservation also have their own geographies, so that species are under much greater threats in some parts of the world than in others. Either as a tool or a threat in conserving biodiversity, tourism is most significant in areas where other threats are most severe and least effective. To improve resilience of protected areas to climate change, for example, other sources of impact need to be reduced, and these include impacts from tourism and recreation. Political and financial support from tourism, however, may be more critical where conservation reserves or endangered species are poorly protected because of factors such as poaching, invasive species, encroachment, or extractive industries, than in areas where parks agencies are operating effectively.

The significance of any contribution from tourism to conservation depends on the economic scale of the tourism sector in the area concerned, and on the effectiveness of local laws and institutions to minimize the impacts of tourism and harness it for conservation. The first aspect is straightforward. Some, at least, of the world's most biodiverse regions do indeed receive large numbers of international tourists. The second aspect is highly problematic. There is an extensive literature on: whether or not public protected areas should charge fees; how these should be structured; whether or not they outweigh visitor management and fee collection costs; who should use any net funds raised; and for what purposes. There is also an extensive literature on private organizations which run tourism operations specifically to support conservation.

There does not seem to be any nation which has successfully harnessed tourism to support conservation at a large

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00308-7

scale. Indeed, it is by no means clear that this is a valid public policy goal, except in countries with strong overall economic dependence on nature-based tourism. For countries with other large industry sectors, or a large urban and resort tourism industry, it may well be more logical to fund conservation from resource royalties and corporate taxation across all sectors. Logging and mining, farming, and fisheries are continually reducing the conservation value of public lands outside the protected estate. Tourism is indeed a major user of protected areas, but still has lower impacts on conservation than those of the extractive industries.

Attractions

A significant subsector of the tourism industry worldwide relies on particular plant and animal species as the primary attractions for commercial tourism products. There is a broad distinction between products and packages based on the opportunities to see a single charismatic icon species, and those which provide opportunities to see a diverse assemblage of different species (Buckley, 2010b). Examples of the former include: the Big Five of sub-Saharan Africa; the big cats such as tiger and jaguar; bears and pandas; whales and dolphins; and the larger primates such as gorilla, orangutan, and chimpanzees. There are substantial wildlife tourism subsectors which rely on each of these. There are also individual icon species which are perceived by tourists as small and cute rather than large and dangerous. Examples include meerkat, lemurs, bushbabies, and sugar gliders. Lemurs as a group, for example, are a major attraction for tourists to visit Madagascar.

There is a substantial international birdwatching tourism subsector which focuses on the diversity of bird species encountered, in the shape of lifetime lists. There are particular icon species which birdwatchers want to see in individual destination areas, but they choose destinations which will give them a chance to see many species in the same area. This also applies for tourism products based on flowering plants, known as wildflower tours, where people want to see a diversity of spectacular species flowering simultaneously. There are specialized subsidiaries, such as those which focus on orchids or proteas. There are also tourism products which focus on single individual trees, typically those which are very large and old or those which have some historical cultural significance.

There are some wildlife species which are perceived as iconic to particular landscapes, and used in tourism marketing even if few tourists actually see them. Examples include: wolves in wilderness areas of North America and parts of Europe; the Mikado pheasant in the mountains of Taiwan; and narwhal in boat-based tourism in the Arctic. While mammals and birds are by far the most important wildlife for tourism, there are also a number of reptiles and amphibia which are valued by tourists. Examples include the Komodo dragon in Indonesia, frilled lizard in Australia, nesting turtles on tropical beaches worldwide, chameleons in Madagascar, and nocturnal frog hunts at various rainforest lodges. There are small-scale tourism operations based around particular plant species such as the rafflesia flower of tropical southeast Asia, the welwitschia plant of the Namib Desert, and the spiny octopus tree of Madagascar. There are also a number of landscapes famed for tourism which owe their form and attractiveness to particular types of vegetation. Examples include the redwoods of California, the saguaro and ocotillo of the American southwest, the heather moors of Scotland, the peat bogs of Ireland, the savannahs of Africa, the grasslands of central Asia, the eucalypt forests of Australia, and many more.

Impacts

At a global scale, tourism currently contributes around 5% of total anthropogenic carbon dioxide emissions and around 14% of global greenhouse gas impacts when other gases are also considered (Scott *et al.*, 2010). About four-fifths of this is from air travel. Climate change is already affecting the range and survival of species and ecosystems worldwide, and tourism is thus contributing to these impacts. Airlines and car rental companies offer carbon offsets, but few passengers purchase them. The effectiveness of current offset programs is highly questionable, since they are based on marginal adjustments between land uses and energy generation technologies which have high environmental impacts and a history of public subsidies. At current offset prices, the additional cost is far below the levels which would reduce air travel through economic effects (Tol, 2007; Buckley, 2011).

The majority of tourist accommodation is urban, and hotels and holiday apartments are subject to the same environmental regulations and planning procedures as other urban buildings of comparable size. They create impacts on biodiversity through *pro rata* contributions, firstly, to the impacts of providing construction materials, operational consumables, and water and energy; and secondly, to the generation and disposal of solid and liquid wastes and atmospheric emissions. The tourism industry in many countries also relies heavily on large stand-alone resorts, especially in coastal and mountain areas. Ski resorts and associated residential developments can have considerable impacts on mountain ecosystems, including clearance, terrain modification, fresh water extraction, noise, light, and roadkill. These impacts are exacerbated as the resort operators respond to climate change by extending retail precincts and residential property developments, and expanding both summer activities and winter snowmaking (Pickering and Buckley, 2010). In Australia, for example, ski resorts are one of the principal factors threatening the endangered mountain pygmy possum.

In pristine or relatively undisturbed natural environments of high value for biodiversity conservation, activities associated with tourism and recreation can generate a variety of direct and local impacts. The degree of impact depends strongly on the type of activity and on when, where, and how it is carried out (Buckley, 2004). There are many instances where helicopters and light aircraft flown at low altitude, for example, have created significant local impacts on caribou, polar bear, dall sheep, nesting pelicans, and other species. An off-road vehicle (ORV) driven slowly and carefully can often approach wildlife within meters or even closer, without causing disturbance; whereas ORVs driven fast and heedlessly can cause significant damage and disturbance to soil, vegetation, and wildlife. Such impacts are particularly severe for

competitive events, whether formal or informal, such as desert races, dune driving, and mudbogging. There are a number of reported instances where off-road vehicles have caused major losses in local populations of small mammals, ground-nesting birds, reptiles and amphibians, and beach-dwelling crabs. Both wheeled vehicles and livestock can also act as vectors of weeds and pathogens. Hikers also disperse weed seeds and fungal propagules on footwear and clothing (Pickering and Mount, 2010; Whinam *et al*., 2005). Hikers and other non-motorized outdoor recreationists, especially with dogs, can also cause disturbance to birds and mammals, which may affect foraging energetics if repeated frequently.

Conservation

The potential for tourism to contribute to conservation is promoted extensively by tourism industry associations, but this is largely a political lobbying tool. There are indeed a small number of commercial tourism enterprises which do appear to make a positive net contribution to conservation and biodiversity, but these are exceptions: The majority of the mainstream tourism sector makes a large net negative impact on biodiversity through the mechanisms outlined in the preceding section. Tourism industry lobbyists routinely suggest, for example, that merely giving people the opportunity to visit areas of high conservation value will somehow convert those people to active conservation lobbyists, but there is no evidence that this actually happens. Tourism is often touted as one justification for the declaration of new protected areas, but these lobbying efforts are largely by conservation groups rather than the tourism industry itself.

Where public parks agencies charge visitor entrance or activity fees, commercial tour clients generally have to pay these in the same way as independent recreational visitors, though sometimes at slightly different rates. Commercial tour operators in public parks may also have to pay a basic license or permit fee, though this is usually at a nominal level (Buckley, 2003a). In developed nations, funds raised through visitor fees typically do not cover the costs of parks visitor infrastructure and management, so commercial tour operators are effectively subsidized by general tax revenue. In developing nations, especially where international visitors pay significantly higher fees than local residents, these revenues may indeed make a significant contribution to park's agency budgets. The broader economic activity associated with inbound park-based international tourism may also provide significant political support for conservation at a national level.

The case of Madagascar is often cited; though this case also exemplifies the extreme vulnerability of conservation tourism to political instability. At least Madagascar's lemurs are well known, so that international concern was quickly created following the 2009 coup. Other countries, for example, in West Africa, may suffer similar depredations of endangered species through poaching, logging, and hunting, without anyone knowing. It is not clear, however, whether international public knowledge of Madagascar is linked to wildlife tourism, or solely to wildlife documentaries and the cartoon movie *Madagascar*. While there are indeed some individual tour operators who advertise the conservation significance of their wildlife tours to Madagascar and similar areas, the main conservation mechanism seems to be from local economic contributions during the tours, rather than any larger-scale political action by the clients.

In a number of developing nations, private reserves and community lands are similar in scale, or indeed larger in total area, than the public protected area estate. Such reserves are not necessarily managed in the same way as national parks, but they do at least show a broad-scale change in land use, for example, from cattle grazing to conservation. The principal financial incentive for land owners to establish private reserves, in countries such as the USA, Brazil, and South Africa where they are common, is through land tax and income tax systems. Only a small proportion relies on tourism as a significant income stream; and in many cases this is based on hunting safaris rather than wildlife watching. Either approach can potentially yield benefits for biodiversity more broadly, but this depends on the details of land management practices.

A proportion of these private and community reserves, especially in sub-Saharan Africa and South America, are indeed managed for conservation and are funded specifically by commercial wildlife tourism (Buckley, 2010a, b). The best known examples operate at the upper end of the market, with luxury lodges offering good opportunities to watch icon wildlife species at close range. Less luxurious models are also widespread and numerous, but somewhat less visible.

There have been a number of attempts to catalogue enterprises marketing themselves as ecotourism (Buckley, 2003b; Zeppel, 2006), but only a small proportion of these make a net positive contribution to the conservation of biodiversity (Buckley, 2008, 2010a). In one or two cases, it has been possible to quantify the contribution as a definable percentage of the remaining global population of particular endangered species which owe their existence specifically to conservation tourism. There is as yet, however, no aggregate estimate of the total contribution to conservation of biodiversity which may be made by the commercial tourism sector, even without allowing for conservation costs through the large-scale impacts of mainstream tourism.

Conclusions

Links between tourism and biodiversity may be summarized as follows: (1) About one-fifth of the global tourism industry relies directly on biodiversity. (2) Most of the mainstream tourism industry generates diffused but large negative impacts on biodiversity, plus smaller localized direct impacts from some resorts and activities. (3) Outdoor and nature tourism generates broad political, and to some degree financial, support for protected areas, especially in developing nations, but this is uncertain and unquantified. (4) A small number of specialized conservation tourism enterprises do make demonstrable net positive contributions to conservation of particular endangered species, principally through the establishment, management, and funding of private and community reserves. (5) This model, however, depends heavily on political stability, currency exchange rates, and low airfares, and is itself a contributor to global climate change. (6) There is

continued contention, even at the most expert level, as to whether conservation is better served by coupling it or decoupling it from commercial tourism, and by what precise mechanism. (7) Many different approaches have been tried or suggested, with various degrees of success depending on the detailed circumstances. While the number of case studies is continuing to increase, general patterns remain elusive. (8) Tourism has definite negative and potential positive effects on biodiversity, and has not yet been harnessed routinely so as to live up to its potential as a conservation tool. Of course, the commercial tourism industry does not see itself as a conservation tool, but as a set of profit-making enterprises, so this situation is likely to continue.

See also: Conservation Efforts, Contemporary. Conserving Biodiversity Outside Protected Areas. Corals and Coral Reefs. Economic Value of Biodiversity, Measurements of. Ecosystem Services. Endangered Mammals. Endangered Plants. Hotspots. Identifying Conservation Priorities using a Return on Investment Analysis. Indigenous Peoples and Biodiversity. Mammals, Conservation Efforts for. Market Economy and Biodiversity. Natural Reserves and Preserves. Ocean Ecosystems. Property Rights and Biodiversity. The Value of Biodiversity. Trends in Nature Recreation: Causes and Consequences. Wildlife Management

References

Buckley RC (2003a) Pay to play in parks: An Australian policy perspective on visitor fees in public protected areas. *Journal of Sustainable Tourism* 11: 56–73.
Buckley RC (2003b) *Case Studies in Ecotourism*. Wallingford: CAB International.
Buckley RC (ed.) (2004) *Environmental Impacts of Ecotourism*. Wallingford: CAB International.
Buckley RC (2008) World wild web: Funding connectivity conservation under climate change. *Biodiversity* 9: 71–78.
Buckley RC (2009) *Ecotourism: Principles and Practices*. Wallingford: CAB International.
Buckley RC (2010a) *Conservation Tourism*. Wallingford: CAB International.
Buckley RC (2010b) Wildlife tourism. In: Buckley RC (ed.) *Adventure Tourism Management*, pp. 139–156. Wallingford: CAB International.
Buckley RC (2011) Tourism under climate change: Will slow travel supersede short breaks? *Ambio* 40: 328–331.
Pickering CM and Buckley RC (2010) Climate response by the ski industry: The shortcomings of snowmaking for Australian resorts. *Ambio* 39: 430–438.
Pickering CM and Mount A (2010) Do tourists disperse weed seed? A global review of unintentional human-mediated terrestrial seed dispersal on clothing, vehicles and horses. *Journal of Sustainable Tourism* 18: 239–256.
Scott D, Peeters P, and Gössling S (2010) Can tourism deliver its "aspirational" greenhouse gas emission reduction targets? *Journal of Sustainable Tourism* 18: 393–408.
Tol RSJ (2007) The impact of a carbon tax on international tourism. *Transportation Research Part D* 12: 129–142.
Whinam J, Chilcott N, and Bergstrom DM (2005) Subantarctic hitchhikers: Expeditioners as vectors for the introduction of alien organisms. *Biological Conservation* 121: 207–219.
Zeppel H (2006) *Indigenous Ecotourism: Sustainable Development and Management*. Wallingford: CAB International.

Traditional Conservation Practices

Carl Folke and Johan Colding, Stockholm University and Beijer International Institute of Ecological Economics, Stockholm, Sweden

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 681–694, © 2001, Elsevier Inc.

Glossary

Institutions Humanly devised constraints that shape human interaction and the way societies evolve through time; made up of formal constraints (rules, laws, constitutions), informal constraints (norms of behavior, conventions, self-imposed codes of conduct), and their enforcement characteristics.

Natural disturbance Any relatively discrete event in time that disrupts ecosystem community or population structure and changes resources, substrate availability, or the physical environment; key for structuring biological communities and for maintaining resilience in ecological systems.

Resilience The system's capacity to absorb disturbance and conserve opportunity for self-organization and evolution. Resilience has to do with how resistant the system is to fundamental reorganization such as a phase shift into another stability domain.

Social taboo A prohibition imposed by social custom or as a protective measure.

Traditional ecological knowledge A cumulative body of knowledge, practice, and belief, evolving by adaptive processes and handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and with their environment.

Traditional peoples Variously referred to as indigenous peoples, native peoples, or tribal peoples; peoples with cultures that differ from those of the mainstream of a national population, having distinct ethnic languages with locally evolved resource management systems; currently some 5000 distinct indigenous peoples exist in the world.

Introduction

Even traditional peoples with a relatively high human population number were able to utilize rain forest areas without destroying these environments and surrounding biological communities (Primack, 1993). However, the specific objective of these local practices is not necessarily directed toward the conservation of species, their habitats, and ecosystems. Rather, such practices are often geared for sustainable use of local resources and ecosystems, with biological conservation resulting as an indirect outcome. They are often tied to cultural belief systems, which makes it difficult to separate the belief component from actual management practices and the ecological knowledge system on which they build. Knowledge, practices, and beliefs tend to intermingle in most traditional management systems (Gadgil *et al.*, 1993). This constitutes the basis behind traditional ecological knowledge (TEK), denoting that resource management patterns are the products not only of a people's physical environment and its resources but also of their cultural perceptions of the environment and its resources (Ruddle, 1994).

Cultural belief systems of various peoples have protected species, their habitats, and even smaller ecosystems. For example, several verses in the Vedas and Upanishads mention conservation and protection of plants and animals, indicating that traditional conservation practices of many rural and indigenous groups of India go as far back as the Vedic period (5000 B.C.). In fact, a great deal of social mechanisms, such as social taboos, may be highly adaptive from an ecological perspective and contribute to biodiversity conservation (Colding and Folke, 1997). The term *conservation* is here used in the sense of sustainable use of natural resources for human benefit, without compromising the interests of future generations (WCED, 1987).

This article will provide examples of a diverse set of traditional management practices and institutions that exist among local resource users. We will illustrate that such practices play an important role for *in situ* conservation of biological resources. We deal with traditional management practices that lead to the sustainable use of biological resources and local institutions and belief systems that impose regulations on the use of species and access to ecosystems.

The word *traditional* refers here to the historical and cultural continuity of resource management, recognizing that societies are constantly redefining what is considered "traditional." The term "local resource users" refers to both traditional peoples and small-scale societies of Western countries with locally evolved management systems. Local resource users generally depend on a rather limited resource procurement base to provide them with a wide diversity of resources (Gadgil *et al.*, 1993). Many do not have access to fossil fuel dependent technology or capital markets. Their day-to-day survival depends on proper interpretations and knowledge of the dynamics of their local resources and ecosystems. Thus, they often have a stake in managing their resources for long-term endurance.

Many of the examples of management practices and local institutions presented in this article derive from the case study anthology edited by Berkes and Folke (1998). This anthology predominantly deals with local resource users that display success of long-term environmental management by using practices that contribute to building resilience in local

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00144-1

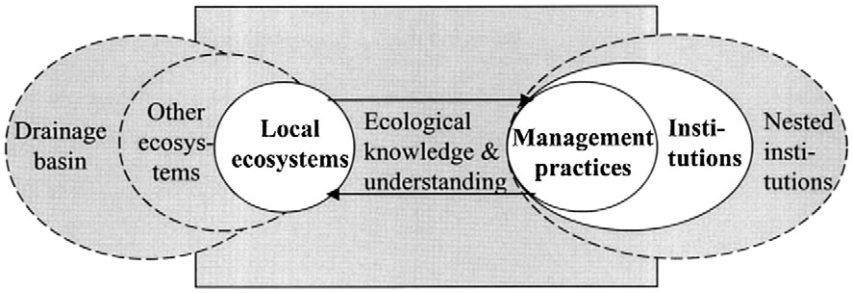


Figure 1 Focus of analysis of linked social–ecological systems. Ecological knowledge and understanding is a critical link between complex and dynamic ecosystems, adaptive management practices, and institutions.

Table 1 Management practices of local resource users

1. Ecological monitoring, and practices, framing access to and use of species and habitats
   - Monitoring the status of the resource
   - Total protection of certain species
   - Protection of vulnerable life histories of species
   - Protection of specific habitats
   - Temporal restrictions of harvest
2. Practices of multiple-species management, resource rotation, and succession
   - Multiple-species and integrated management
   - Resource rotation
   - Management of succession
3. Practices related to the dynamics of complex ecosystems
   - Management of landscape patchiness
   - Watershed management
   - Managing ecological processes at multiple scales
   - Responding to and managing pulses and surprises
   - Nurturing sources of ecosystem renewal

Source: Reproduced from Folke C, Berkes F, and Colding J (1998) Ecological practices and social mechanisms for building resilience and sustainability. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

ecosystems. A diverse set of management practices and social mechanisms from temperate and tropical regions and traditional and contemporary society were identified and analyzed in recognition of their importance for building resilience in combined social–ecological systems (Levin *et al.*, 1998). The focus of the analysis is presented in **Figure 1**.

The examples stem from a wide range of local resource users, including hunters and gatherers, herders, fishers, agriculturists, and small-scale communities of industrial nations. This article describes these practices in relation to their nature conservation functions. Most of the practices described are interrelated and have multiple ecological functions. Nonetheless, we have categorized them in the manner outlined in **Table 1**.

As indicated in **Table 1**, the first category of practices includes ecological monitoring and regulation of the use of biological resources and ecosystems. (*see* Ecological Monitoring and Practices Framing Access to and Use of Species and Habitats) Section II of this article. The second category mainly concerns traditional agroforestry practices, with an emphasis on their effect for biological conservation. (*see* Practices of Multiple-Species Management, Resource Rotation, and Succession). The third category describes management practices that are related to the dynamics of complex ecosystems, in particular the processes that structure ecosystems at different temporal and spatial scales. (*see* Practices Related to the Dynamics of Complex Ecosystems). Such conservation practices, often at the watershed and landscape levels, are important in securing a flow of natural resources and ecosystem services on which the local community depends.

In the concluding section, we stress lessons that can be learned from traditional conservation practices in building resilience through improved management of biodiversity. Resilience conserves options and opportunity for ecosystem renewal and evolutionary change, and its dynamic maintenance is a prerequisite for all forms of biological conservation.

Ecological Monitoring and Practices Framing Access to and Use of Species and Habitats

The practices described below are often determined by local social institutions. These are often informal, based on traditional norms and conventions. Some appear to be more directly associated with resource management activities, while others appear to be more closely associated with cultural belief systems.

Monitoring the Status of the Resource

Monitoring the status of the resource is very common among local resource users (Folke *et al.*, 1998). Monitoring leads to the acquisition of local ecological knowledge and helps local users respond to resource and ecosystem dynamics. Both qualitative and quantitative indicators develop as the result of experience and ecological knowledge based on monitoring (Johannes, 1998). Monitoring may provide information about location and timing of target resources (i.e., resource procurement) and also the condition and dynamics of individual populations and particular ecosystems for generating target resources. Ecological monitoring also provides the basis for the spatial and temporal regulations of resource use in local communities. Thus, various social response mechanisms evolve as a result of ecological monitoring. These are often in the form of local institutions, such as social taboos, area rotation regulations, and seasonal closures. The list of the importance of ecological monitoring for local resource users can be made very long. We will provide a few examples.

For resource procurement reasons, the Cree of James Bay, Canada, monitor geographic distribution, migration patterns, and individual behavior of caribou, as well as the sex and age composition of the caribou herd, including the presence or absence of predators. They also predict herd size of caribou by examining the fat content of caribou (Berkes, 1998). The Inuit of Quebec and the Inuit and Innu of Labrador also use fat content as an indicator of health of both individual animals and the herd. Such monitoring may provide useful information for resource procurement and in the enforcement of hunting regulations when prey is low in abundance.

Similarly, the Cree of James Bay monitor the interaction between beaver and vegetation (Berkes, 1998). Such monitoring leads to knowledge about beaver populations and their ecological interaction with other biological resources. For example, when aspen is low in a particular area, trappers know that the beaver population is low in abundance.

Knowledge about the effect of natural disturbance regimes on target resources and other species is also based on ecological monitoring. Such knowledge requires monitoring over extensively long time periods and may be passed on from one generation to the next. For example, Cree beaver trappers know that 3–4 years after a fire occurs in an area, beavers will start to inhabit the area again (Berkes, 1998).

Also, monitoring combined with ecological knowledge may lead to resource procurement methods that do not deteriorate the habitats of target resources and secure a sustainable use of target resources. For example, in Tokelau in the South Pacific, islanders use, based on detailed knowledge of octopus behavior, a special octopus stick to extract the animal, which obviates the need for the destructive crushing of the coral or the use of poison (Ruddle, 1994). Another example is the use of the clam rake in the Maine soft-shell clam fishery, which is the only tool allowed for clam harvesting (Hanna, 1998). Such methods reduce the risk that a resource is overexploited, as compared to more advanced forms of technological methods.

Local institutions for management, based on ecological monitoring, may either be permanent in time such as in the two examples above, or "kick in" occasionally in times of resource scarcity. For example, ecological monitoring underlies the Cree's decision when to rotate beaver trapping grounds and fishing areas. Cree fishers rotate fishing grounds based on a declining catch per unit of effort and rest these sites when needed (Berkes, 1998). In this way, overfishing is limited. Coastal communities in Maine monitor clam populations to help determine the areas needing enhancement (Hanna, 1998). Such response mechanisms represent fine-tuned ways of responding to ecological feedback and are employed in order to avoid overexploitation of subsistence resources.

The widespread uses of closed seasons in Oceania are based on observations founded on local ecological knowledge about the spawning periods of key fish species, and fishing is prohibited during such periods. Pacific island groups often know when and where fish aggregate to spawn. At such aggregation sites, fishers monitor yearly changes of fish stock size and composition and reduce their fishing effort when stocks seem to be low (Johannes, 1978; Hviding, 1988).

Similarly, the use of fishing taboos among Pacific islanders at particular sea areas is based on local knowledge about the fluctuations of reef fish populations. On Satawal Island in the Central Caroline Island of Micronesia, fishing off one reef section is prohibited by taboo in order to conserve a breeding ground to supply the rest of the reef with resources (Ruddle, 1994). On Tokelau, the *lafu* (fishing taboo) is the most explicit conservation measure. *Lafu* is invoked by the Council of Elders. It bans all fishing in specific areas on the main reef and is announced to permit stock recovery.

Such examples allowing for ecosystem renewal is evident among a number of local resource users. For example, pastoral groups in arid and semiarid Africa monitor ecosystem change by way of tracking to determine daily movements of herds. This includes tracking ecological processes, such as climate, soil content, groundwater availability, forage availability, temporal environmental variability, and environmental degradation. Various forms of indicators are used for this, such as the behavior of fauna, specific plant indicators, soil color and texture, forage quantity and quality, and vegetation/plant composition (Niamir-Fuller, 1998).

Due to the proximity of local resource users to their resource base, ecological monitoring is often feasible on a daily basis. This facilitates detection and response to ecological change. Quite often, monitoring activities may be ascribed to particular individuals in local communities, such as resource stewards, elders, or shamans (Berkes and Folke, 1998).

Total Protection of Certain Species

Total protection of certain species is another practice that is particularly common among traditional peoples. For example, flora and fauna in India have been protected by indigenous belief systems for millennia and have been revered as the vehicles of Gods and Godesses.

Of particular interest are the different kinds of social taboos imposed on species by traditional societies throughout the World. Colding and Folke (1997) found that specific-species taboos protect threatened species as well as species considered keystone and/or endemic by ecologists. It was estimated that about 30% of the identified taboos protect species listed as threatened by IUCN. **Table 2** displays a number of ecologically important and threatened species protected through taboos. This indicates the role that traditional local institutions, such as taboos, may have in biological conservation.

Protection of Vulnerable Life Histories of Species

This practice reduces the danger of overharvesting and the depletion of a population of target resources. For example, in the Maine fisheries, it is prohibited to gather lobsters with eggs (Acheson *et al.*, 1998). Traditional fishing castes in the Bhandara district of Maharastra, India, never disturb the spawning aggregations of freshwater fish in hill streams, and the Phasepardhis of Maharastra traditionally let loose pregnant does caught in their snares (Gadgil, 1987). The Cree of James Bay never kill or disturb nesting geese (Berkes *et al.*, 1995). The Tukano Indians of Colombia impose taboos on the collection of bird eggs and avoid the collection of reptiles during their breeding season as well as protect fish spawning

Table 2 Threatened species that are avoided as a result of taboos

Species	*Popular name*	*Local resource users/locality*	*IUCN status*[a]
Kinosternon oaxacae	Oaxaca mud turtle	Pima Bajo, Papago, Yuman, U.S.A./Mexico	I
Chelonia mydas	Green sea turtle	Inhabitants of Buzios Island, Brazil	E
Naja oxiana	Oxus cobra	Local protection in vicinity of temples in India	K
Melanosuchus niger	Black caiman	Piro of Peru	V
Heloderma suspectum	Gila monster	Riverine Pima, Papago, U.S.A./Mexico	V
Pavo muticus	Green peafowl	Tamil Nadu, Rajasthan, Gujarath, India	V
Gorillas gorilla	Gorilla	Edo state, Nigeria	V
Colobus polykomos	Black and white colobus	Villages of Boabeng and Fiema, Ghana	V
Pan troglodytes	Chimpanzee	Edo state, Nigeria	V
Thomomys umbrinus emotus	Southern pocket gopher	Riverine Pima, Papago, Maricopa, U.S.A./Mexico	V
Perognathus alticola	White-eared pocket mouse	Riverine Pima, Papago, Maricopa, U.S.A./Mexico	V
Dipodomys gravipes	San Quintin kangaroo rat	Riverine Pima, Papago, U.S.A./Mexico	E
Dipodomys microps leucotis	Houserock chiseltoothed kangaroo rat	Riverine Pima, Papago, U.S.A./Mexico	K
Canis lupus	Grey Wolf	Bishnois of Thar desert, Rajasthan, India	V
Tremarctos ornatus	Spectacled bear	Achuar of Ecuador/Peru	V
Panthera tigris	Tiger	Local protection in vicinity of temples in India	E
Felis concolor	Puma	Maricopa, Yuman speakers, U.S.A./Mexico	E
Tapirus bairidi	Central American tapir	Coshiro-wa-teri of Brazil/Venezuela; Achuar of Ecuador	V
Myrmecophaga tridactyla	Giant anteater	Coshiro-wa-teri of Brazil/Venezuela; Achuar of Ecuador/Peru	V
Pridontes maximus	Giant armadillo	Achuar of Ecuador/Peru	V
Antilope cervicapra	Blackbuck	Bishnois of Thar desert, Rajasthan, India	V

[a]IUCN status: E, endangered; I, indeterminate; V, vulnerable; K, insufficiently known; R, rare.

Source: Reproduced from Colding J and Folke C (1997) The relations among threatened species, their protection, and taboos. *Conserv. Ecol.* (online) 1(1), 6. Available from the Internet: http://www.consecol.org/vol1/issl/art6

aggregation sites in rivers (Reichel-Dolmatoff, 1976). Size restrictions are sometimes employed among South Pacific islanders on slow-moving or sessile marine species that are particularly susceptible to over-harvesting (Johannes, 1978).

Protection of Specific Habitats

This practice is commonly found among local resource users. For example, pastoralists of arid and semiarid Africa use buffer zone areas of Sahelian rangelands which are protected from grazing except in the case of emergencies (Niamir-Fuller, 1998).

Also, whole forests, forest patches, coast stretches, rivers, or ponds may to various degrees be protected for human resource use. Usually, such areas are set aside by religious taboos and considered sacred to community members. In India, sacred groves were once extremely common. A sacred grove is a small part of a forest set aside for spiritual or religious purposes. The sizes of such protected areas vary from a clump of 5–10 trees to as much as 50 ha or more (Gadgil and Vartak, 1976). Sacred groves still exist in many parts of contemporary India—for example, in the Khasi Hills in Assam, in the Arvalli ranges of Rajasthan, all along the Western Ghats in the southern peninsula, in the districts of Bastar and Sarguja in Madya Pradesh, and in the Chanda district in Maharastra. Gadgil and Chandran (1992) indicate that the Indian traditional shifting cultivation system, *jhum* (described *in* Practices of Multiple-Species Management, Resource Rotation, and Succession: Management of Succession), is associated with sacred groves.

Kenya has sacred groves all along its coast, known as *kayas* (homesteads), used for ceremonies and burials (Wilson, 1993). So do the Yoruba of Ara in southwestern Nigeria (Warren and Pinkston, 1998). In South America, the Kuna Indians of Panama have spirit sanctuaries, places where spiritual animals, plants, or demons are believed to reside (Chapin, 1991). According to cosmological beliefs, the Kuna must respect these areas often located on choice agricultural land. The Cocnucos and Yanaconas of Colombia have similar sanctuaries (Redford and Maclean Stearman, 1993). The Tukano of the Uaupés basin on the Brazil–Colombia border reserve the forested river margin for fish and fishing. Fishing may be restricted to as little as 38% of the total river margin available (Chernella, 1987). The result is a management system that allows for, yet distinguishes, human use areas and animal refuge areas. Any deforestation of the river edge is prohibited.

The importance of such cultural beliefs for preserving patches of ecosystems should not be underestimated. New species of plants, and species that have disappeared from other areas, are still being discovered within sacred groves (Mohannan and Nair, 1981). A botanical survey in a Nigerian sacred grove yielded 330 plant species as compared to only 23 in surrounding nonprotected areas (Warren and Pinkston, 1998). Sacred groves are also important for maintaining ecological services, such as preserving local hydrological cycles, preventing soil erosion, serving as firebreaks, and serving as areas of recruitment of species, allowing for ecosystem renewal in face of various disturbances.

Temporal Restrictions of Harvest

This practice is adopted among some local resource users. The idea is that certain biological resources are protected from

exploitation for certain periods of time. Among local resource users, the imposition of temporal taboos regulates access to resource(s) on either a sporadic, daily, weekly, or monthly basis (Colding and Folke, 2000).

For example, clans of Tikopia in the Solomon Islands impose sporadic taboos on particular foodstuffs they are associated with (Chapman, 1985). Several different durations of closed seasons exist among Vanuatu fishing villages, ranging from 1 month to 5 years (Johannes, 1998).

At the Sakumo and Djange Lagoons in Ghana, taboos are imposed on fishing during a particular day every week (Ntiamoa-Baidu, 1991). In India, taboos imposed on a monthly basis appear quite widespread. For example, many castes abstain totally from consumption of fish, poultry, and meat, and suspend all hunting as well, in the Hindu month of Sravana (roughly August), which is the peak of the main rainy season over most of India (Gadgil, 1987). Similarly, taboos on hunting certain animals from July to October exist in many Indian villages. The Oraons, the fourth largest tribal group in India, observe a taboo of hunting any wild animal or bird during the months of June and July (Xaxa, 1992).

In Papua New Guinea, the Maopa people in the Marshall Lagoon impose a seasonal ban on hunting (Kwapena, 1984). Every 3–4 years at the end of their cultivation work, a bush area is set on fire to chase and hunt any animals found. This traditional hunting ritual lasts for a week. Once a particular animal-rich area is selected and hunted on, it will be left alone for another 3–4 years. This ritualistic way of hunting allows for hunting areas to regenerate. It may also mimic natural disturbance regimes. Groups of Canadian Amerindian hunters use a similar system for rotating hunting, fishing, and trapping areas by their periodic resting (Berkes, 1998).

In the Maluku Islands of eastern Indonesia, temporal prohibitions on gathering are imposed on terrestrial plant and animal resources, such as coconut, nutmeg, areca, and cuscus, as well as on marine species, including fish, trochus, shell, sea cucumber, and seaweed. This traditional management institution, known as *sasi*, is continually changing, depending on changing social and ecological conditions (Soselisa, 1998).

Locally decided and implemented closed periods is one of the measures used for managing kombu kelp (*Laminaria augusta*) at Hokkaido Island, Japan. Regulations initiated for kelp harvesting as early as 1881 stated that villagers could not harvest kombu in the evening or on a rainy day, when good quality kombu cannot be produced (Iida, 1998).

Practices of Multiple-Species Management, Resource Rotation, and Succession

This category constitutes management practices that largely can be described as agroforestry practices, although hunter and gatherers may employ methods that draw on the same principles. Agroforestry is the generic name used to describe "an old and widely practiced land use system in which trees are combined spatially and/or temporally with agricultural crops and/or animals" (Farrell, 1987). They are common in traditional agricultural systems throughout the world, especially in the tropics, but were once common in temperate regions as well (Beets, 1990).

The three types of practices described in the following are often related to the efficient use of local ecosystems in terms of space, nutrient availability, climatic conditions, and soil conditions. These practices serve to actively manage species diversity and intensity of use of subsistence resources for ecosystem renewal and species recruitment. Such practices often create habitat heterogeneity on the landscape scale, affecting forest structure and species composition by creating a mosaic of forest patches of different ages (Primack, 1993). Thus, the outcome of these practices often enhances diversity of biological resources at the local level (Orejuela, 1992).

Multiple-Species and Integrated Management

The cultivation of several species of crops on the same piece of land is often referred to as polycultures or intercrops. There is an enormous variety of this type of practice, and they may involve the mix of annual crops with other annuals, annuals with perennials, or perennials with perennials. Crops may be sown in different spatial arrangements and may range from the growing of two crops in alternate rows to complex assemblies of a dozen or more species (Altieri, 1987). Integration of crop cultivation with animal species is also common in traditional agroforestry systems.

Polyculture cropping systems constitute at least 80% of the cultivated area of West Africa and predominate in other parts of Africa as well as in the Latin American tropics and many Asian countries (Liebman, 1987). Frequently, more yield can be harvested from an area sown in polyculture than from an equivalent area of monoculture without the application of synthetic fertilizers, pesticides, and field machinery. Polyculture cropping systems may provide several social–ecological benefits as compared to monocultures (Liebman, 1987)—for example, by reducing the risk of total crop failure and by contributing to an increase of nutritional returns to local resource users.

Polycultures may also promote soil water conservation and nutrient recycling as well as reduce soil erosion and insect pests. Some polyculture systems may also reduce crop diseases and increase weed control. Polycultures may thus provide for many ecological services, including the maintenance and enhancement of biological diversity. In fact, diversity is used to enhance productivity and resilience (Berkes *et al.*, 1995).

The two most common polyculture systems in Java, Indonesia, are *Talun-kebun* (rotation between mixed gardens and tree plantation) and *pekarangan* (home garden intercropping system including animals) (Altieri, 1987). A *Talun-kebun* is an indigenous Sudanese agricultural system, consisting of three stages that each serves a different function. In the first stage, *kebun*, a mixture of annual crops is usually planted, mainly for cash income. After 2 years, tree seedlings begin to grow in the field, and the *kebun* gradually evolves into a *kebun-campuran*, where annual crops are mixed with half-grown perennials. This stage promotes soil and water conservation. After the annuals are harvested, the field is usually abandoned for 2 or 3 years and becomes dominated by perennials, the *talun* stage (perennial crop garden). The *talun* can be turned back into a *kebun* after the forest is cleared or be planted to rice paddy, depending on whether irrigation water is available.

A *Talun-kebun* may be composed of up to 112 species of plants, of which about 40% provide for building materials and fuel wood, 20% are fruit trees, 15% are vegetables, and the remainder is used for ornamentals, medicinal plants, spices, and cash crops (Altieri, 1987).

The integration of animals in polyculture systems is a common practice in diverse regions of the world. In Southeast Asia, rice ecosystems include diverse animal species. For example, domestic ducks may be integrated in paddy rice cultivation to allow for insect and weed control, and fish species, such as common carp, may be integrated and harvested at the end of the rice-growing season. The rice/fish culture differs considerably from country to country and from region to region (Altieri, 1987).

In China, as in Southeast Asia, crops, chicken, duck, and fish are often integrated to provide for a very high overall production through waste recycling and use of residues (Yan and Yao, 1989). Also, natural control of insects and weeds is provided for in these systems (Beets, 1990). In Indonesia, traditional systems combined rice and fish culture, and wastes from this system often flowed downstream into brackish water aquaculture systems (*tambak*). The *tambaks* were polyculture ponds, often combining fish, vegetation, and tree crops (Gadgil *et al*., 1993).

In Rajasthan, India, local farmers grow crops and rear livestock under and among trees. The Bishnoi, inhabiting one of the driest climatic zones on Earth, integrate the cultivation of food crops and animal rearing among islands of khejri trees, *Prosopis cineraria*, which are strongly protected by religious taboo (Sankhala, 1993). The Bishnoi keep as many khejri trees as possible on their farms, providing them with material for fencing, firewood, and fodder and pods for cattle as well as providing excellent microclimatic conditions for crop cultivation. Even threatened species, such as the blackbuck and gray wolf, benefit from these farming systems (Colding and Folke, 1997).

In many parts of Africa, "farm trees" can also be found within and adjacent to farm fields. The trees are actively protected, harvested, and managed by farmers to yield construction poles, fuel wood, fodder, edible fruits, nuts and leaves, medicines, and other products without unduly competing with associated annual crops (Beets, 1990).

Resource Rotation

Resource rotation exists among several communities of local resource users (Folke *et al*., 1998). Chisasibi Cree hunters rotate trapping areas (ideally) of beavers on a 4-year cycle to allow for the recovery of beavers to these areas. They use a similar rotational practice for resting fishing areas, using a traditional pattern of fishing in remote lakes on a 4- or 5-year cycle. Also, Cree goose hunters rotate hunting areas on a 7-day cycle in order to reduce disturbance to feeding and resting geese and to harvest for subsistence needs with a minimum disruption of the large population that passes through an area (Gadgil *et al*., 1993).

The Awa Indians of southwestern Colombia and neighboring Ecuador hunt echimid rodents on a cyclical basis, which allows sufficient time for the recovery of wild populations (Orejuela, 1992).

Arid and semiarid African pastoralists migrate seasonally in accordance with plant availability, determined by precipitation and lunar cycles (Niamir-Fuller, 1998). The yearly cycle of nomads and their cattle provides a rotational management system that enables the recovery of heavily grazed rangelands. For example, the Wodaabe Fulani follow the lunar cycle when moving to new pastures, which means that the camp is moved every 2–3 days. By contrast, the Rufa'a al Hoi of Sudan move to a new pasture every 204 days. The Fulani of northern Sierra Leone once practiced "shifting pasturage," whereby they heavily stocked an area for 2–3 years and then moved elsewhere and rested the first area for 15–20 years. These are all examples of small-scale movement, or micromobility, practiced by many Sahel pastoralsist tribes (Niamir-Fuller, 1998). Local resource users enforce grazing rotation in parts of the Hindu–Kush Himalayas in a similar fashion (Jodha, 1998).

Management of Succession

Polycultures and crop rotation often involve management of succession, as exemplified by the different systems of shifting cultivation that exist in the world. Shifting cultivation is defined by FAO (1982) as "a farming system in which relatively short periods of cultivation are followed by relatively long periods of fallow." Although part of polycultures, it can be distinguished by fallow periods that are ideally very long.

Shifting cultivation, or "slash-and-burn" cultivation or "swiddens," involves the clearing of a plot of land, usually a forest area, its use for a few years, and, as soil fertility declines, its abandonment in favor of another plot of land to be cleared in the same fashion. It is one of the oldest forms of agriculture and most present-day agricultural systems have evolved from it. In shifting cultivation, agriculture becomes a sequential cropping of crops and noncrops. Presently, about 3–500 million people, or about 40% of the total agricultural population of developing nations, depend on shifting cultivation for their daily livelihood. In total, shifting cultivation covers about 30% of the world's exploitable soils (Beets, 1990).

Shifting cultivation is common in all tropical areas and was once common in temperate regions as well. It has received much attention as one of the major degrading processes in tropical forest areas due to human population increase that greatly has led to shortages in fallow periods with subsequent loss of soil fertility. Generally, if fallow is less than 5–7 years, land degradation occurs and species diversity may be greatly reduced (Berkes *et al*., 1995). However, if adequate long fallow is allowed for, shifting cultivation may be highly sustainable (e.g., Posey, 1985).

Extremely high crop diversity is a characteristic of many traditional shifting cultivation systems. For example, Philippine swidden cultivators can distinguish over 600 plant species (Beets, 1990). In general, swiddens do not compare in complexity to the surrounding forest. However, when shifting cultivation is analyzed as an agroforestry system, i.e., the use of trees is also taken into account, then the overall result of managing forest patches can lead to an enhancement of biodiversity (Berkes *et al*., 1995). For example, the Runa Indian managed swiddens in the Ecuador Amazon increase species diversity in 5-year-old fallows as compared to unmanaged

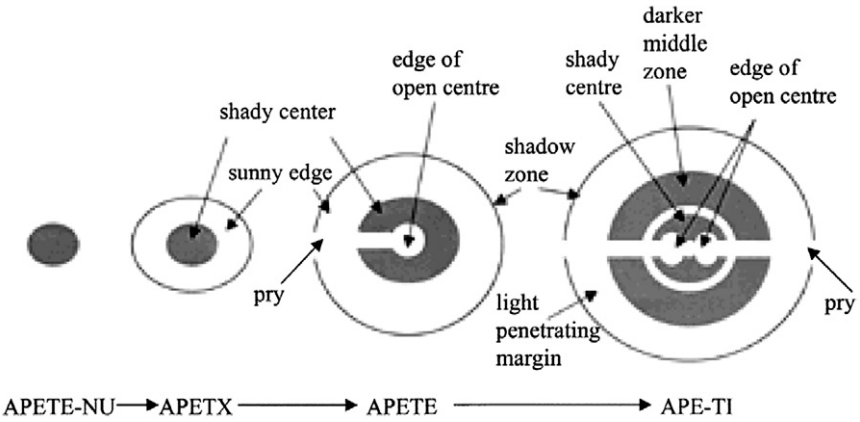


Figure 2 Among the Kayapo Indians of Brazil, enhancement of biodiversity is facilitated for by the creation of forest islands, *apete*. Through a number of management methods, this behavior promotes patchiness and heterogeneity in the landscape in time and space. Reproduced from Gadgil M, Berkes F, and Folke C (1993) Indigenous knowledge for biodiversity conservation. *Ambio* 22: 2–3.

fallows. Between 14 and 35% of this enhanced species diversity was attributed to direct planting and protection of secondary species.

Different forms of swiddens and related fallow systems exist in the world. In tropical Mexico, as well as in other Mesoamerican countries, *milpa* (maize fields) is widely practiced (Alcorn and Toledo, 1998). *Milpa* involves the clearing of new fields in high forests, or secondary regrowth forests, for maize cultivation. If more than one successive crop of maize is taken in a short fallow *milpa*, weedy species come to dominate the plot and forest regeneration may not occur. *Milpa* is governed by strong informal institutions, which reinforces reciprocity and community-based control of natural resources. Farmers who mismanage their *milpa* are labeled as witches and punished by social pressure. Each stage of the *milpa* cycle is also named and marked by ritual activities, rendering *milpa* a strong sociocultural practice.

The Awa Indians of southwestern Colombia and neighboring Ecuador practice a shifting agriculture known as slash–mulch (Orejuela, 1992). It involves the cutting of natural vegetation and the mulching of this material for a temporary agricultural field. Maize and/or short-cycle varieties of red and white beans are planted by the slash–mulch method. Once yields decline after several cycles of harvest, the field is left fallow and permitted to regrow as an enriched secondary forest. A similar system involving mulching is used to clear lots to raise cattle.

In northeastern India, detailed studies on shifting agriculture (*jhum*) have described multispecies systems involving 4 to over 35 crop types based on locally adapted native strains (Berkes *et al.*, 1995). This practice requires sophisticated local ecological knowledge, including the use of soil nutrients by adequate changes in the crop mixture depending on the length of *jhum* cycles and the consequent availability of soil nutrients. On hill slopes, farmers combine r-strategist species (cereals and legumes) with K-strategists, with emphasis on vegetative growth, such as leafy vegetables (Berkes *et al.*, 1995).

An interesting form of shifting cultivation by the Kayapó Indians of Brazil constitutes the creation of forest islands known as *apete* (Posey, 1985). This begins as small mounds of vegetation about 2 m in diameter (*apete-nu*). As planted crop and tree seedlings grow and the planted area expands, the taller vegetation in the center of the mounds is cut to allow light to enter. A full-grown *apete* has an architecture that creates zones that vary in shade and moisture (**Figure 2**). The species mix includes medicinal plants, palms, and vines that produce drinking water. Of a total of 120 species found in 10 *apete*, Posey (1985) estimated that 75% may have been planted. *Apete* constitutes the manipulation of semi-domesticated plants on which the Kayapó could survive during times of warfare. These old forest islands have been scattered for millennia in known spots throughout the forest and savanna.

Practices of multiple species, resource rotation, and succession management contribute to biodiversity conservation relative to more modern and technology-dependent ways of exploiting resources and ecosystems. By actively using biodiversity in production, they contribute to biodiversity conservation in areas outside protected areas and reserves.

Practices Related to the Dynamics of Complex Ecosystems

Apete is also a good example of traditional practices that actively promote patchiness and heterogeneity at the landscape level. Many agroforestry systems make use of patchiness for multifunctional benefits, although patchiness may also be used in smaller cultivation systems to provide for crop protection and natural pest management (Altieri, 1987).

Management of Landscape Patchiness

Such management is employed by Sahelian pastoralists in order to mimic the variability and unpredictability of the landscape. For example, an appropriate mix of herding animals is used to utilize different vegetation types and patches in a dynamic fashion (Niamir-Fuller, 1998). The progressive widening of grazing radius around wells as the wet season

advances is also an example of active management at the landscape level. This practice is employed by the Maasai of Kenya to leave enough forage around wells for the dry season (Niamir-Fuller, 1998).

Fire management among Australian aborigines and northern Canadian Amerindians was practiced widely to open up clearings (meadows and swales), corridors (trails, traplines, ridges, grass fringes of streams and lakes), and windfall forests. These clearings provided improved habitats for ungulates and waterfowl, thus increasing hunting success, and the corridors and windfall areas improved accessibility (Gadgil *et al.*, 1993).

The use of different elevation zones for cultivators in the eastern Himalayas and other parts of the world represents practices that make use of habitat heterogeneity and create patchiness at the landscape level. The ancient *ahupua'a* system of Hawaii utilizes different elevation zones, determined by precipitation, that are used for various integrated farming practices (Costa-Pierce, 1987). In the ancient *ahupua'a*, both freshwater and seawater fish ponds were integrated with agriculture, and river valleys were managed as integrated systems, from the upland forest (left uncut by taboo) all the way down to the reef (Costa-Pierce, 1987).

Watershed Management

The *ahupua'a* also constitutes an example of traditional watershed management systems. Southeast Asia and Oceania had, and to lesser extent still have, a number of these prescientific ecosystem management practices. Examples include the Yap *tabinau*, the Fijian *vanua*, and the Solomon Islands *puava* (Ruddle *et al.*, 1992). These all refer to generically similar watershed-based management systems. For example, the *vanua* concept is an integrated human–nature component that regards the land, water, and human environment as one unit, one and indivisible (Ruddle *et al.*, 1992). Similarly, the *puava* includes all resources and land in a watershed, from the top of the mainland mountains to the open sea outside the barrier reef (Hviding, 1990). In all these cases, the social group inhabiting the ecosystem unit is considered to be part of the system.

Managing Ecological Processes at Multiple Scales

There is some evidence that traditional management systems may be useful in managing ecological processes at multiple scales. In *milpa*, (described earlier *in* Practices of Multiple-Species Management, Resource Rotation, and Succession: Management of Succession), food crops are managed on a 1- to 3-year scale, and some tree crops and products on a 30-year scale (Alcorn and Toledo, 1998).

James Bay Cree hunters seem to be simultaneously managing beaver populations on a 4- to 6-year scale, lake fish on a 5- to 10-year scale, and caribou on an 80-to 100-year scale (Berkes, 1998). The holistic forestry of the Gitskan people of northern British Columbia simultaneously manages the production of fiber over several square kilometers with ecological processes involving soil bacteria at the spatial scale of a few square meters (Pinkerton, 1998).

Responding to and Managing Pulses and Surprises

Natural disturbances, such as fire, hurricanes, pest outbreaks, and heavy grazing, are inherent to the internal dynamics of ecosystems and often set the timing of ecosystem renewal processes (Holling *et al.*, 1995).

An example of a practice that responds to disturbance and manages pulses and surprises is the establishment of range reserves within the annual grazing areas of African herders. These reserves provide an emergency supply of forage that serves as buffer when disturbance, such as drought, challenges the process and function of the dryland ecosystem (Niamir-Fuller, 1998). At the landscape level, this functions to maintain the resilience of both the ecosystem and the social system of the herders. Such practices may be considered ecological adaptations to unpredictable, low-rainfall environments and provide "ecological insurance" to the local communities.

Sacred groves in India absorb disturbance by serving as firebreaks for cultivated areas and villages (Gadgil *et al.*, 1998). At the same time, these sacred habitats function as recruitment centers for species regeneration and ecosystem renewal. In times of natural disturbances and other emergencies, such areas of refugia are essential for providing social–ecological resilience.

Many of the traditional forms of polyculture previously described may provide for pest and insect protection that contributes to building resilience (Altieri, 1987). The Warlis of India control pests by placing certain kinds of tree branches in their paddy fields. This practice serves to attract birds for insect control and buffers against outbreaks of various pest populations (Pereira, 1992).

Nurturing Sources of Ecosystem Renewal

Many traditional societies seem to actively nurture sources of ecosystem renewal by creating small-scale disturbances. Traditional agroforestry practices such as shifting cultivation create forest gaps and enable people to produce crops or enhance wild foods without disrupting natural renewal processes (Berkes *et al.*, 1995). Such practices may even enhance genetic and ecological diversity in the benign mosaic landscape of forest, fallow, and gardens (Orejuela, 1992). Also, traditional rotational cycles employed by hunters and gatherers, (described *in* Practices of Multiple-Species Management, Resource Rotation, and Succession), allow for species recovery of wild populations.

Aboriginal use of fire in as geographically diverse areas as Canada, Australia, and California had many elements and principles in common (Berkes *et al.*, 1995). Until the late 1940s, Amerindians of northern Alberta, Canada, regularly used fire to create clearings (meadows and swales), corridors (trails, traplines, ridges, grass fringes of streams and lakes), and windfall forests (Lewis and Ferguson, 1988). These clearings provided improved habitat for species such as ungulates and waterfowl. Australian aborigines possessed detailed technical knowledge of fire and used it to improve feeding habitat for game and to assist in the hunt itself (Lewis, 1989).

African herders behave like a disturbance by following the migratory cycles of the herbivores from one area to another

(Niamir-Fuller, 1998). Pulses of herbivore grazing contribute to the capacity of the semiarid grasslands of Africa to function under a wide range of climatic conditions. If this capacity of the ecosystem to deal with pulses is reduced, an event that previously could be absorbed can flip the grassland ecosystem into a relatively unproductive state, dominated and controlled by woody plants for several decades (Walker, 1993).

Concluding Remarks

One of the 10 principles of the *Global Biodiversity Strategy* concerns the linkage of biodiversity with cultural diversity and the conservation of the two together. At a broader level, traditional conservation practices as a part of cultural diversity may offer benefits in terms of local biological and ecological understanding, sustainable resource management systems, implementation of protected areas, development planning, and environmental impact assessment. Each of these potential benefits is related to biodiversity conservation, either directly or indirectly. Biodiversity is recognized as an important component of sustainable use. It is in this sense it is being conserved through traditional conservation practices as illustrated in this article through practices framing access to and use of species and habitats, practices of multiple-species management, and practices related to biodiversity conservation in complex dynamic ecosystems. Such practices seem to contribute to maintaining and building ecological resilience *sensu* Holling (1986).

Traditional conservation practices rely on the accumulation of ecological knowledge and understanding over many generations, and knowledge is embedded in local institutions and transmitted culturally. Practices are generally site and context specific although the ecological knowledge embedded in them may be spatially and temporally transmitted.

The locally generated knowledge and associated conservation practices are not necessarily complete in the sense of understanding of all aspects of an ecosystem and its dynamics. However, many traditional conservation practices are in line with the shifting scientific view on the nature of ecosystems as nonlinear, multiequilibrium, and full of surprises, threshold effects, and system flips. Predictability and controllability are not limited by the scientific data available but by the very nature of ecological systems.

Traditional ecological knowledge systems, based on detailed observations of the dynamics of the natural environment, feedback learning, social system–ecological system linkages, and resilience-enhancing mechanisms, seem akin to adaptive management (Berkes and Folke, 1998). Traditional conservation practices parallel adaptive management in their reliance on learning-by-doing and the use of feedback from the environment to provide corrections for management practice.

The parallels between adaptive management and indigenous management systems are probably not accidental. Flexible social systems that proceed along by learning-by-doing are better adapted for long-term survival than are rigid social systems that have set prescriptions for resource use. In light of this, adaptive management in modern society could be seen as a replication of traditional ecological knowledge systems in the framework of contemporary science. It is a sort of rediscovery of principles applied in traditional social–ecological systems. It is a search for sustainable use of ecosystems and biological diversity as a means of survival, a social and institutional response to resource scarcity and management failure. Even though there are no doubt major differences between the two, adaptive management may be viewed as the scientific analogue of traditional ecological knowledge and practices because of its integration of uncertainty into management strategies and its emphasis on practices that confer ecological resilience.

See also: Agriculture, Traditional. Conservation Movement, Historical. Historical Awareness of Biodiversity. Indigenous Peoples and Biodiversity. Religious Traditions and Biodiversity. Social and Cultural Factors. Sustainability and Biodiversity

References

Acheson JM, Wilson JA, and Steneck RS (1998) Managing chaotic fisheries. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Alcorn JB and Toledo VM (1998) Resilient resource management in Mexico's forest ecosystems: The contribution of property rights. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Altieri MA (1987) *Agroecology: The Scientific Basis of Alternative Agriculture*. Boulder, CO: Westview.

Balée W (1992) People of the fallow: A historical ecology of foraging in lowland South America. In: Redford KH and Padoch C (eds.) *Conservation of Neotropical Forests. Working from Traditional Resource Use*. New York: Columbia Univ. Press.

Beets WC (1990) *Raising and Sustaining Productivity of Smallholder Farming Systems in the Tropics*. Alkmaar, Holland: AgBé Publishing.

Berkes F (1998) Indigenous knowledge and resource management systems in the Canadian subarctic. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Berkes F and Folke C (1998) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Berkes F, Folke C, and Gadgil M (1995) Traditional ecological knowledge, biodiversity, resilience and sustainability. In: Perrings CA, Mäler K-G, Folke C, Holling CS, and Jansson B-O (eds.) *Biodiversity Conservation*. Dordrecht/Norwell, MA: Kluwer Academic.

Chapin M (1991) Losing the way of the great father. *New Scientist*, Aug. 10.

Chapman M (1985) Environmental influences on the development of traditional conservation in the South Pacific region. *Environ. Conserv.* 12(3).

Chernella J (1987) Endangered ideologies: Tukano fishing taboos. *Cult. Surv.* 11(2).

Colding J and Folke C (1997) The relations among threatened species, their protection, and taboos. *Conserv. Ecol.* (online) 1(1), 6. Available from the Internet: http://www.consecol.org/vol1/issl/art6

Colding J and Folke C (2000) The taboo system: Lessons about informal institutions for nature management. *The Georgetown Int'l Envtl. Law Review* 12(2).

Costa-Pierce BA (1987) Aquaculture in ancient Hawaii. Integrated farming systems included massive freshwater and seawater fish ponds. *BioScience* 37(5).

FAO (1982) *Improved Production Systems as an Alternative to Shifting Cultivation*, FAO Soils Bulletin No. 52. Rome: FAO.

Farrell JG (1987) Agroforestry Systems. In: Alatieri MA (ed.) *Agroecology: The Scientific Basis of Alternative Agriculture*. Boulder, CO: Westview.

Folke C and Berkes F (1998) *Understanding Dynamics of Ecosystem–Institution Linkages for Building Resilience* (Beijer Discussion Paper Series No. 112). Beijer International Institute of Ecological Economics, Stockholm, Sweden: The Royal Swedish Academy of Sciences.

Folke C, Berkes F, and Colding J (1998) Ecological practices and social mechanisms for building resilience and sustainability. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Gadgil M (1987) Social restraints on exploiting nature: The Indian experience. *Development: Seeds of Change* 1987: 1.

Gadgil M and Chandran MDS (1992) Sacred groves. In: Sen G (ed.) *Indigenous Vision: Peoples of India Attitudes to the Environment*. New Delhi: Sage Publications.

Gadgil M and Vartak VD (1976) The sacred groves of Western Ghats in India. *Econ. Bot.* 30.

Gadgil M, Berkes F, and Folke C (1993) Indigenous knowledge for biodiversity conservation. *Ambio* 22: 2–3.

Gadgil M, Hemam NS, and Reddy BM (1998) People, refugia and resilience. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Hanna SS (1998) Managing for human and ecological context in the Maine soft shell clam fishery. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Holling CS (1986) Resilience of ecosystems; local surprise and global change. In: Clark WC and Munn RE (eds.) *Sustainable Development of the Biosphere*. Cambridge, UK: Cambridge Univ. Press.

Holling CS (1996) Cross-scale morphology, geometry, and dynamics of ecosystems. In: Samson FB and Knopf FL (eds.) *Ecosystem Management. Selected Readings*. New York: Springer-Verlag.

Holling CS, Schindler DW, Walker BW, and Roughgarden J (1995) Biodiversity in the functioning of ecosystems: An ecological synthesis. In: Perrings C, Mäler K-G, Folke C, Holling CS, and Jansson B-O (eds.) *Biodiversity Loss. Economic and Ecological Issues*. Cambridge, UK: Cambridge Univ. Press.

Hviding E (1988) *Marine Tenure and Resource Development in Marovo Lagoon, Solomon Islands: Traditional Knowledge, Use, and Management of Marine Resources, with Implications for Contemporary Development* (FFA Report No. 88/35). Honiaro, Solomon Islands: South Pacific Forum Fisheries Agency.

Hviding E (1990) Keeping the sea: Aspects of marine tenure in Marovo Lagoon, Solomon Islands. In: Ruddle K and Johannes RE (eds.) *Traditional Marine Resource Management in the Pacific Basin: An Anthology*. Jakarta, Indonesia: UNESCO/ROSTSEA. Jln. M. H. Thamrin No. 14.

Iida T (1998) Competition and communal regulations in the Kombu Kelp (*Laminaria angustata*) harvest. *Hum. Ecol.* 26(3).

Jodha N (1998) Reviving the social system–ecosystem links in the Himalayas. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Johannes RE (1978) Traditional marine conservation methods in Oceania and their demise. *Annu. Rev. Ecol. Syst.* 9.

Johannes RE (1998) The case for data-less marine resource management: Examples from tropical nearshore finfisheries. *Trends Ecol. Evol.* 13.

Kwapena N (1984) Traditional conservation and utilization of wildlife in Papua New Guinea. *Environmentalist* 4(Suppl. 7).

Levin SA, Barrett S, Aniyar S, *et al.* (1998) Resilience in natural and socioeconomic systems. *Environ. Dev. Econ.* 3.

Lewis HT (1989) Ecological and technical knowledge of fire: Aborigines versus park managers in Northern Australia. *Am. Anthropol.* 91.

Lewis HT and Ferguson TA (1988) Yards, corridors and mosaics: How to burn a boreal forest. *Hum. Ecol.* 16.

Liebman M (1987) Polyculture cropping systems. In: Altieri MA (ed.) *Agroecology: The Scientific Basis of Alternative Agriculture*. Boulder, CO: Westview.

Monahan CN and Nair NC (1981) Kunstlaria Prain: A new genus record for India and a new species in the genus. *Proc. Indian Acad. Sci.* 90.

Nelson JG and Serafin R (1992) Assessing biodiversity: A human ecological approach. *Ambio* 21(3).

Niamir-Fuller M (1998) The resilience of pastoral herding in Sahelian Africa. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Ntiamoa-Baidu Y (1991) Conservation of coastal lagoons in Ghana: The traditional approach. *Landscape Urban Plan.* 20.

Orejuela JE (1992) Traditional productive systems of the Awa (Cuaiquer) Indians of Southwestern Colombia and neighboring Ecuador. In: Redford KH and Padoch C (eds.) *Conservation of Neotropical Forests. Working from Traditional Resource Use*. New York: Columbia Univ. Press.

Pereira W (1992) The sustainable lifestyle of the Warlis. In: Sen G (ed.) *Indigenous Vision: Peoples of India Attitudes to the Environment*. New Delhi: Sage Publications.

Pinkerton E (1998) Integrated management of a temperate montane forest ecosystem through wholistic forestry: A British Columbia example. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Posey DA (1985) Indigenous management of tropical forest ecosystems: The case of the Kayapo Indians of the Brazilian Amazon. *Agrofor. Syst.* 3(2).

Posey DA (1992) Interpreting and applying the "reality" of indigenous concepts: What is necessary to learn from the natives? In: Redford KH and Padoch C (eds.) *Conservation of Neotropical Forests. Working from Traditional Resource Use*. New York: Columbia Univ. Press.

Primack RB (1993) *Essentials of Conservation Biology*. Sunderland, MA: Sinauer.

Redford KH and MacLean Stearman A (1993) Forest-dwelling native Amazonians and the conservation of biodiversity: Interests in common or in collision? *Conserv. Biol.* 7(2).

Reichel-Dolmatoff G (1976) Cosmology as ecological analysis: A view from the rainforest. *Man* 11.

Ruddle K (1994) Local knowledge in the folk management of fisheries and coastal marine environments. In: Dyer CL and McGoodwin R (eds.) *Folk Management in the World's Fisheries: Lessons for Modern Fisheries Management*. Niwot: Univ. Press of Colorado.

Ruddle K, Hviding E, and Johannes RE (1992) Marine resources management in the context of customary tenure. *Mar. Res. Econ. 7.*

Soselisa HL (1998) Marine Sasi in Maluku, Indonesia. *Out of the Shell. Coastal Resources Res. Netw. Newsl.* 6(3).

Turner BL, Clark WC, and Kates WC (1990) *The Earth As Transformed by Human Action: Global and Regional Changes in the Biosphere over the Past 300 Years*. Cambridge, UK: Cambridge Univ. Press.

Walker BH (1993) Rangeland ecology: Understanding and managing change. *Ambio* 22: 2–3.

Warren DM and Pinkston J (1998) Indigenous African resource management of a tropical rainforest ecosystem: A case study of the Yoruba of Ara, Nigeria. In: Berkes F and Folke C (eds.) *Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience*. Cambridge, UK: Cambridge Univ. Press.

Wilson A (1993) Sacred forests and the elders. In: Kemf E (ed.) *Indigenous Peoples and Protected Areas*. London: Earthscan Publications Ltd.

Wilson E (1992) *The Diversity of Life*. Cambridge, MA: The Belknap Press of Harvard Univ. Press.

World Commission on Environment and Development (WCED). (1987) *Our Common Future*. Oxford: Oxford Univ. Press.

Xaxa V (1992) Oraons: Religion, customs and environment. In: Sen G (ed.) *Indigenous Vision: Peoples of India Attitudes to the Environment*. New Delhi: Sage Publications.

Yan J and Yao H (1989) Integrated fish culture management in China. In: Mitsch WJ and Jörgensen SE (eds.) *Ecological Engineering: An Introduction to Ecotechnology*. New York: Wiley.

Translocation as a Conservation Strategy

Jessica J Hellmann, University of Notre Dame, Notre Dame, IN, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Captive breeding Collection of organisms from nature followed by human husbandry to maximize the reproduction and survival of an organism of interest, usually in a zoo, botanic garden, or other controlled environment.

Conservation Steps taken by people to preserve the diversity of life on Earth and the health of natural ecosystems, usually to counteract a force that erodes or reduces natural processes.

Introduction Placement of individual organisms by people into a habitat for purposes of establishing a population in a place where it does not currently occur.

Managed relocation Intentional movement by humans of organisms from a place where that organism lives currently to a new region where it is likely to persist in the future due to human-caused climate change.

Reintroduction The placement of individual organisms by people into a habitat for purposes of establishing a population where a previous population has gone locally extinct.

Supplementation The purposeful introduction of additional organisms to an existing population to increase mean fitness or population size.

Translocation The intentional movement of organisms by people from one or more locations to other locations.

Overview of Translocations for Conservation

Humans have exploited nature and its creatures for thousands of years. But as threats to nature from pollution, overexploitation, invasive species, and other side effects of human civilization emerged, people began to worry about the protection of nature. And they began to take steps to conserve organisms and ecosystems for future use. This intervention to prevent the decline of nature involves many tools and strategies, including the movement of seeds, shoots, and animals from one location to another. This *translocation* of organisms has a long history because economic, social, and even accidental reasons compel us to pick things up and deposit them somewhere else. This article discusses why and how people move organisms, and sometimes entire populations, to advance the cause of conservation. It is informed by the successes and failures of past translocation efforts, and it raises new reasons why translocations may play an increasing role in conservation planning in the future. Translocations for conservation are rarely a panacea because they usually address conservation problems for an individual species in a specific location, but they can be a powerful tool that may work when other conservation strategies fail. In recent years, translocation has reemerged as an active topic of debate, and still more research can shed light on the appropriate and inappropriate uses of translocation for emerging conservation objectives.

Translocations in Traditional Conservation

Conservationists are people who work to reduce the social, economic, and environmental factors that diminish natural ecosystems. To *conserve* is to retain something as it was before, to prevent it from disappearing or declining. Conservationists include park and wildlife managers, nongovernmental organizations, landowners, and private citizens, and they use a number of tools or techniques to achieve their goals. Conservationists can, for example, remove a threat or stressor to an ecosystem through pollution control or invasive species removal; they can restore an ecosystem through cleanup and species planting; and they can set aside land or water to be left untouched by humans. Sometimes they might even move species from one area to another to restore a native habitat, to save a species from extinction, or to enable genetic mixing and reduce negative effects of inbreeding in small populations.

Conservation biologists are scientists who study the process of conservation and the strategies that humans can use to sustain nature and help it flourish. Conservation biologists evaluate the effectiveness of conservation techniques, suggest ways to achieve conservation goals, and point out where conservation intervention is needed because natural resources are declining. The consultation of conservation biologists is critical to designing conservation interventions that help, rather than undermine, the cause of preservation. Just like conservationists, conservation biologists have long been interested in understanding the costs and benefits of moving species to achieve conservation goals.

Conservation biologists divide the various ways of moving species into categories based on whether the movement (1) is reconstructing some past condition, (2) is to an already-occupied locality, or (3) is to a new location. All of these movements can be called *translocations* because they require picking an organism up from one area and putting it in another. Some translocations involve seed storage or captive breeding that separates the stages of collection and deposition, but all organisms come from an original source population because we cannot create nature from anything other than nature itself.

Reintroductions involve moving organisms from an area of current occupancy to areas where historic populations became locally extinct. Reintroductions are critical to restoration — the reconstruction of a habitat to some historic state. For example, as nearly all of the native tallgrass prairie in Wisconsin, Iowa, and Illinois was converted to agriculture, scientists at the University of Wisconsin with help from the Civilian

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00345-2

Conservation Corp initiated the first prairie restoration in 1936 (Allison, 2002). Since then, prairie planting and restoration has become one of the only ways that diverse prairie habitats exist in the Midwest, providing habitat for animals, cycling nutrients, cleaning storm water, and taking up and storing carbon. Most restorations use locally sourced seed so that historic genetic structure can be maintained or replicated, and attempts are made to create communities that are historically accurate. Reintroductions are also an essential tool of endangered species management. One of the most famous examples of single-species reintroduction is the California condor. When two captive-bred birds were released in 1987, they represented a culmination of 14 years of research and breeding worth $20 million (Cohn, 1993). Today, there are more than 100 condors living in the wild, though the species remains on the list of critically endangered species. Though primarily consumers of carrion, condors are similar to other restoration projects that restore diminished ecosystem functioning through reintroduction of top predators. These predators exert control on large herbivores and thereby encourage greater ecosystem productivity (e.g., see Ripple and Beschta, 2004).

Supplementation is the addition of organisms to an existing population. Fisheries and timber managers often supplement existing populations to increase the number of sport fish or the quantity of board feet. Supplementation can also be used to introduce genetic diversity into an existing population where that diversity will help the population expand or persist for a longer period of time. For example, eight panthers from Texas were introduced into south Florida in the mid-1990s to reduce inbreeding depression in the declining Florida panther. Subsequent studies found that kittens born to Texan mothers and Floridan fathers had a higher survival to adulthood than purebred cats (Pimm *et al.*, 2006). This benefit may explain recent increases in the abundance and habitat use of Florida panthers that appears to have staved off imminent extinction.

Introductions involve the placement of organisms into a new area, generally where a species has not occurred historically. Introduction has been a rare occurrence in conservation and natural resource management because most scientists feel that placing species where they have not occurred previously can be damaging to other species. In some cases, introductions have been pursued near a species' historic range where habitat is sufficiently similar and introduced populations can flourish. This is the case for Tule elk introduced to Owens Valley of California, for example, where a rare subspecies was placed in an area east of its historic range. Some species also have been introduced to escape a threat that exists in their historic range. This was the case for koala that were introduced to islands in Westernport Bay, Victoria, in the late 1800s and early 1900s to avoid threats of hunting, disease, and habitat loss in their native mainland range (Short and Smith, 1994). These introductions were later used as a source for reintroductions back to mainland Australia, where local populations had gone extinct.

Emerging Motivations for Translocation

This broad role for translocation in traditional conservation, to restore and preserve biodiversity, serves as a backdrop for new ideas about translocation. This time it could be used to combat worldwide shifts in several environmental conditions. For one, climate strongly affects where species live, and climate change causes populations to expand into new areas and disappear from others. In addition, human activities, such as the addition of fertilizers are pushing communities into altered states with differing species composition and differing system functioning. Human populations also continue to grow, increasing stress on rural areas to provide food and other services such as timber, minerals, and freshwater; and massive tropical deforestation continues at an alarming rate, threatening the heartland of global biodiversity.

Each of these factors – global climate change, nutrient application, land-use change, and human population growth – are changing the "baseline" of life on Earth. They make it increasingly difficult for conservationists to achieve the objective of preserving (or reconstructing) nature as it used to be. These factors of global environmental change are, in fact, undermining the very definition and objectives of conservation and conservation biology. We will need to substitute a new, evolving definition of conservation that is forward-looking and works to maintain biodiversity as it shifts due to human forces. Scientists, governmental agencies, stakeholders, and the public are not finished grappling with a new interpretation of conservation biology, but possible objectives include maintaining as much global biodiversity as possible (e.g., genetic, species, and habitat diversity) and maximizing ecosystem services such as flood control, nutrient cycling, and aesthetic values. Perhaps an ethic of sustainability will emerge in conservation biology, where natural resources are available for future generations to enjoy as much as they have been enjoyed by previous generations, but the identity of those resources changes steadily over time. In this way, ecosystems are preserved not as they have been but as they can be – in a dynamic state that is often managed by people to maintain diversity and function.

In this new paradigm of conservation where future ecosystems are "preserved" in a dynamic, active way, an expanded role for translocations may appear. For example, species might be moved from areas where they historically lived to new areas so that they do not go extinct due to climate change. Species might also be moved to new areas because changing conditions have enabled a service-delivering species to live in a new area where it can provide a valuable ecosystem function such as purifying water or controlling floods. Organisms might also be moved within a range to introduce new genotypes to existing populations, genotypes that are resistant to warmer temperatures or climatic extremes. Other translocations could be imagined to combat other forms of environmental change – such as habitat loss due to large-scale hydropower development or habitat incursion from human refugees that are escaping sea-level rise or extreme drought. Translocation for these purposes have not yet been articulated or named, but we can look forward and imagine their possible use.

Consider the example of the island scrub Jay that lives only on Santa Cruz Island off the California coast. The species is not listed as threatened or endangered, but we can anticipate future threats from emerging disease such as West Nile virus, rising sea levels, changing habitat from climate change, and small population size. Morrison *et al.* (2011) suggest that an

integrated strategy include captive breeding, vaccination against West Nile, and introduction onto a second island to increase population size and spread the risk of local extinction. Such integrated strategies, including translocations for new population establishment, may become standard practice for precautionary species management.

Managed Relocation

Though translocation may play several roles in a world dominated by global change, at least one role has been recently named and explored in the literature. This idea was first termed *assisted migration* to describe how humans might help species "migrate" from one location to another as the climate changes. Avoiding the potential for policy makers and the public to misinterpret migration, this idea also has been called *assisted colonization*. This term has been criticized, however, because *colonization* only allows for population establishment and not persistent tending, monitoring, and even removal in the event of unanticipated, negative outcomes. Most recently, the term *managed relocation* was coined to imply a suite of activities that one might do to place a species in a new location *and* manage it there once it has established. Managed relocation was defined by Richardson *et al.* (2009), arising from an international working group formed to study the pros and cons of the idea. It is an "intervention technique aimed at reducing negative effects of climate change on defined biological units such as species, populations, or ecosystems. It involves the intentional movement of biological units from current areas of occupancy to locations where the probability of future persistence is predicted to be higher." The lack of personality in this terminology, however, is a significant downside when explaining the idea to the public. It remains to be seen which term will stick; likely, multiple terms will be used in differing contexts.

There have been very few examples of managed relocation to date, but new conversations about it are arising all of the time. A few studies have examined the process of population establishment in areas outside a species' range through experiments that test the viability of managed relocation (e.g., Marsico and Hellmann, 2009; Willis *et al.*, 2009). Some citizen groups concerned with the preservation of particular species also have moved organisms to new areas (e.g., *Torreya taxifolia* via the Torreya Guardians, a not-for-profit group). As of 2011, no managed relocation activities were being officially pursued by regional or national governments, though the issue is being deliberated as agencies grapple with the management implications of climatic change.

Factors Determining Translocation Success

Whether the motivation for translocation is classic conservation or in response to modern global change, a number of factors have been identified that increase the success of population establishment (Griffith *et al.*, 1989; Wolf *et al.*, 1996). In the case of animals, for example, Fisher and Lindenmayer (2000) identified the following as increasing the chances of successful translocation: (1) removal of the original reason for population decline in the case of reintroduction; (2) using wild stock, as opposed to stock that was bred in captivity, whenever possible; (3) including as many individuals in the translocation as possible; and (4) considering nonecological factors such as public relations, participation by a reputable agency or actors, and the long-term commitment of stakeholders and other parties. These factors should prevent subsequent declines (1 and 2), reduce the need to perform the translocation multiple times (2), minimize genetic bottlenecks (3), and maximize social acceptability (4). We need additional research on the factors determining the success of translocations for a wider range of taxonomic groups to inform best-practice guidelines, however (Milton *et al.*, 1999). In addition, social scientific research is critical to examine when people value translocations and when they do not, particularly if translocations expand as a conservation tool as in the case of managed relocation.

Risks and Costs of Translocation

Although translocation may sometimes be a valuable conservation approach, it also comes with significant downsides. The risks of translocations are magnified by the complexity of nature and our relative lack of understanding about the processes that control nature. The implementation of translocation can also be very costly, particularly if it is employed in a thoughtful way using risk assessment and experimental research to guide the implementation process.

The first problem with translocation is the risk of overexploitation to source populations. Even if a captive breeding program is used to maximize the number of introduced individuals and minimize taking from natural populations, nature is the ultimate source of all translocation material. If a species or population is highly threatened, then the removal of even a small number of individuals could increase the risk of local (or global) extinction or increase the risk of inbreeding or other genetic problems of small populations. If a translocation fails to establish, then multiple extractions from a source population might be necessary. This risk of overexploitation has caused some to suggest that highly endangered species should not to be targets for managed relocation, for example.

The second problem is that translocations run the risk of creating pest or nuisance populations where they are introduced. Even where a species has been reintroduced where it occurred historically, the ecological and social context of that ecosystem may have changed so that the introduced species is now considered harmful, as in the case of some ranchers perceiving reintroduced wolves as nuisance predators. In the case of managed relocation, introduced species could reduce the abundance of other species in the introduction area, risking local extinction of native species and altering ecosystem function and community composition. To explore this risk, a recent paper built a database of current invasive species in North America and assumed that invasive species with a North American origin are useful analogs for cases of managed relocation (Mueller and Hellmann, 2008). The authors found that the overall risk of invasion from species moved within North America was relatively low; approximately 14% of all

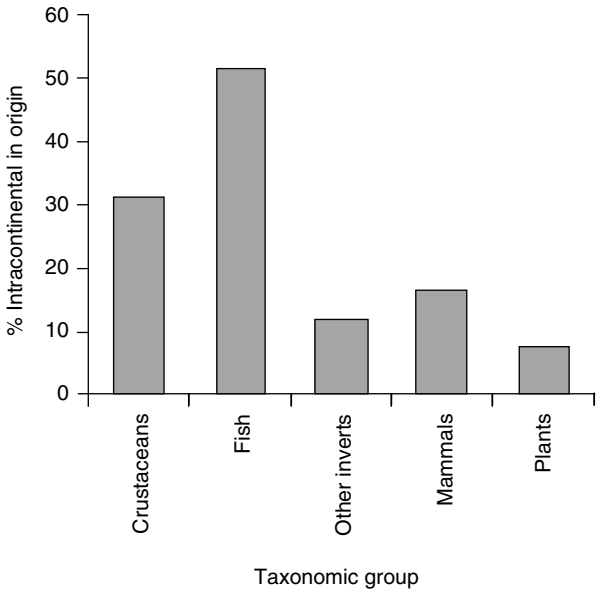


Figure 1 Percentage of species invasive in North America that originate from elsewhere in North America, according to a database of invasives assembled by Mueller and Hellmann (2008). These species of "intracontinental origin" are analogous to species that might be translocated poleward to minimize biodiversity losses due to climate change (i.e., managed relocation). In some taxonomic groups (e.g., plants), a small percentage of currently invasive species come from within North America, but in other groups the percentage of invasive forms of nearby origin is quite high (e.g., fish). In these later cases, translocations under climate change could have a high likelihood of creating nuisance species.

invasive species in the database came from within North America. However, these 14% of species were just as likely to be rated as very harmful in terms of ecological impact as they were likely to have low or minimal impact. Further, the risk of intracontinental invasion was considerably higher in some groups than others, with as many as 50% of invasive freshwater fish species having a North American origin, for example (**Figure 1**). These findings suggest that the overall risk of creating nuisance species via managed relocation could be low or infrequent, but the activity could be ill-advised for some species and will undoubtedly be damaging in some instances.

The counterargument to the risks raised and explored by Mueller and Hellmann (2008) is that invasive species are not representative of organisms that might be moved via managed relocation or other kinds of intracontinental translocations. For example, rare or endangered species that might be helped via managed relocation may be less likely to become nuisance species than those species already identified as invasive and thus included in the Mueller and Hellmann (2008) database. Still other researchers point out, however, that our understanding of the factors that cause a species to become a nuisance are not well known. More sophisticated risk analyses for managed relocation in particular species and taxonomic groups is needed. Social science and legal research also is needed to shed light on who might bare responsibility for damages due to translocations when they occur. These damages might not always be ecological or economic in nature; they also could include a degradation of cultural or social values for people with established ties to the region or ecosystem where a translocation takes place.

A third risk of translocation is hitchhikers or incidental introductions that can inadvertently happen during a translocation. Consider, for example, a pest insect or disease that might accidentally be introduced on an endangered plant moved via managed relocation. In these instances, quarantine measures followed by zoos in captive-breeding programs should be followed by land mangers considering translocation, especially if translocations become more widespread due to increasing global change.

Finally, economic costs of translocations vary widely (Lindburg, 1992), depending on the extent of habitat preparation, captive breeding, ease of obtaining source material, and other factors. If the number of translocations increases in the future, then the total costs of the activity could be considerable, possibly outweighing the costs of dealing with the causes of global change directly. For example, if translocation becomes necessary to conserve a large number of species and ecosystems in the face of climate change, then it could be more costly than reducing greenhouse gas emissions. Litigation that might be associated with translocations could push costs higher.

The Ethics of Translocation

When humans move organisms from one location to another – whether in an attempt to recreate nature or to enhance its future function – we are tinkering with nature. This comes with a number of possible benefits and risks as described above, but it also comes with important ethical and moral considerations. These moral and ethical dilemmas strike at the heart of humanity's relationship with nature. Some people believe that nature exists to service humankind and that proper management and wise use are the core values of conservation. Others think that humans must be responsible stewards of nature based on more than just self-interest — that stewardship comes out of respect for a creator, the power of nature, or the astonishing process of evolution that has produced life as we know it. Still others think that the proper place for human society is outside of nature altogether, such that wilderness and wild spaces are best left untouched by humans. Each perspective can lead to a different evaluation of translocation.

As profound changes to the global environment occur, we must confront difficult ethical questions. For example, do we bear moral or ethical responsibility for stress or death to translocated organisms in terms of animal welfare (e.g., see Singer, 1975)? Or is the preservation of a species or a genotype a sufficient objective? Is it appropriate to try to undo the effects of climate change by further interfering with environmental processes via managed relocation, or is it arrogant to think that humans could correctly engineer systems under climate change (Minteer and Collins, 2010)? Are we morally and ethically obligated to minimize the effects of climate change on biodiversity because we caused the problem in the first place? Exploring the ethical issues of translocation can help society envision the future that we want and the kind of nature that we hope to keep for future generations.

Conclusions

The idea that humans should work to preserve the functioning of ecosystems and the genetic diversity that makes up nature is well established socially and scientifically. How best to achieve these goals, however, is not always clear. A diverse set of tools are available to conservationists, and one of these tools is translocation – the purposeful movement of organisms from one location to another. Many translocations for conservation have been successful in that they preserve biodiversity with minimal negative side effects. Nonetheless, the fear of unintended consequences form a constant backdrop for translocation, as do nontrivial economic costs. It is unlikely that we can anticipate all of the potential risks and costs of translocation at any time in the future given the complexities of natural systems and human values.

We must increasingly recognize that few ecosystems and habitats remain untouched by humans, and the power of humanity to transform nature is profound. Translocations may be a tool for reducing some of those ecological and social damages. As we reevaluate the objectives of conservation and shift from a reconstructive and historic-looking perspective to one of dynamic shifts and change caused by humanity, our views on translocation will undoubtedly grow and change. New research is needed that evaluates translocation for conservation purposes and produces the best practices for multiple taxonomic groups.

Appendix

List of Courses

1. Conservation Biology
2. Environmental Ethics
3. Global Environmental Change
4. Wildlife Management

See also: Captive Breeding and Reintroduction. *In Situ*, *Ex Situ* Conservation. Introduced Species, Impacts and Distribution of. Restoration of Biodiversity, Overview. Threatened Species: Classification Systems and Their Applications. Traditional Conservation Practices. Wildlife Management

References

Allison SK (2002) When is a restoration successful? Results from a 45-year-old tallgrass prairie restoration. *Ecological Restoration* 20: 10–17.

Cohn JP (1993) The flight of the California condor. *BioScience* 43: 206–209.

Fisher J and Lindenmayer DB (2000) An assessment of the published results of animal relocation. *Biological Conservation* 96: 1–11.

Griffith B, Scott JM, Carpenter JW, and Reed C (1989) Translocation as a species conservation tool: Status and strategy. *Science* 245: 477–480.

Lindburg DG (1992) Are wildlife introductions worth the cost? *Zoo Biology* 11: 1–2.

Marsico TD and Hellmann JJ (2009) Dispersal limitation inferred from an experimental translocation of *Lomatium* (Apiaceae) species outside their geographic ranges. *Oikos* 118: 1783–1792.

Milton SJ, Bond WJ, Du Pleissis MA, *et al.* (1999) A protocol for plant conservation by translocation in threatened fynbos. *Conservation Biology* 13: 735–743.

Minteer BA and Collins JP (2010) Move it or lose it? The ecological ethics of relocating species under climate change. *Ecological Applications* 20: 1801–1804.

Morrison SA, Sillett TS, Ghalambor CK, *et al.* (2011) Proactive conservation management of North America's lone insular bird species, the island scrub-jay. *BioScience* 61: 1013–1021.

Mueller JM and Hellmann JJ (2008) An assessment of invasion risk from assisted migration. *Conservation Biology* 22: 562–567.

Pimm SL, Dollar L, and Bass Jr OL (2006) The genetic rescue of the Florida panther. *Animal Conservation* 9: 115–122.

Richardson DM, Hellmann JJ, McLachlan JS, *et al.* (2009) Multidimensional evaluation of managed relocation. *Proceeding of the National Academy of Sciences* 106: 9721–9724.

Ripple WJ and Beschta RL (2004) Wolves and the ecology of fear: Can predation risk structure ecosystems? *BioScience* 54: 755–766.

Short J and Smith A (1994) Mammal decline and recovery in Australia. *Journal of Mammalogy* 75: 288–297.

Singer P (1975) *Animal liberation: A new ethics for our treatment of animals*. New York, NY: Random House.

Willis SG, Hill JK, Thomas CD, *et al.* (2009) Assisted colonization in a changing climate: A test-study using two UK butterflies. *Conservation Letters* 2: 45–51.

Wolf CM, Griffith B, Reed C, and Temple SA (1996) Avian and mammal translocations: Update and reanalysis of the 1987 survey data. *Conservation Biology* 10: 1142–1154.

Trends in Nature Recreation: Causes and Consequences

Patricia Zaradic, Red Rock Institute, Bryn Mawr, PA, USA
Oliver RW Pergams, Olive-Harvey College, Chicago, IL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Biophilia The innate tendency to focus on life and lifelike processes, or the "innately emotional affiliation of human beings to other living organisms" (Wilson, 1984). According to Wilson, we humans have an affinity for the natural world that has evolved over millennia and is part of our genes, just like our tendency to be territorial or to protect our young.

BLMV Bureau of Land Management Lands Visitation = (Total number of annual visits to all Bureau of Land Management lands in the USA)(Total population of the US).

Duck Stamps Federal Migratory Bird Hunting and Conservation Stamps, pictorial stamps produced by the US Postal Service for the US Fish & Wildlife Services as the federal licenses required for hunting migratory waterfowl.

NFV National Forests Visitation = (Total number of annual visits to all national forests in a nation)(Total population of a nation).

NPV National Parks Visitation = (Total number of annual visits to all national parks in a nation)(Total population of a nation).

SPV State Parks Visitation = (Total number of annual visits to all state parks in a nation)(Total population of a nation).

Videophilia The new human tendency to focus on sedentary activities involving electronic media (Pergams and Zaradic, 2006).

Virtual nature Nature experienced vicariously through electronic means.

Why We Care

The greatest threat to conservation and the environment may be more subtle than bulldozers and chainsaws. Our review of studies on attitudes toward the environment suggests that direct contact with nature, especially as children, is the most critical influence on later attitude toward the environment (Zaradic and Pergams, 2007). Family vacations and spending time with family and other mentors outdoors is cited as a major influence on later environmental attitude and activism (Chawla, 1998). Yet for the period 1987–2003, *per capita* visits to national parks (NPV, America's iconic family nature vacation) have steadily declined (Pergams *et al.*, 2004; Pergams and Zaradic, 2006), with this trend continuing today (Pergams and Zaradic, unpublished data). Moreover, most long-term trends point to a general and fundamental shift away from people's participation in nature-based recreation with reliable long-term *per capita* visitation measures of nature recreation peaking between 1981 and 1991, declining approximately −1.2% per year, and totaling to date 18–25% (Pergams and Zaradic, 2008).

This declining trend in nature-based recreation is of enormous importance. Kellert (1996) describes human cultural learning and experience as exerting a fundamental shaping influence on the content, direction, and strength of people's nature-related values. Similarly, it has been found that environmentally responsible behavior results from direct contact with the environment (Hungerford and Volk, 1990) and that people must be exposed to natural areas as children if they are to care about them as adults (Duda *et al.*, 1998). Extended periods spent in natural areas, especially with a mentor, seem to create the most environmentally responsible behavior (Matthews and Riley, 1995) and increased involvement in biodiversity conservation (Chawla, 1998). Moreover, as today's adult role models spend less time in nature, this generation of children is also likely to follow suit (Louv, 2006). Accordingly, the authors think it probable that any major decline in the value placed on natural areas and experiences will greatly reduce the value people place on biodiversity conservation.

Successful stewardship of the environment depends on recognizing our impact and interactions with nature. The less time we spend in nature, the less likely we are to recognize our human impact and respond ecologically and effectively. As global population grows, so does the need for and extraction of natural resources and the impact of humans on the planet. Humans are getting to a point in resource needs and global impact at which efficient use of resources is increasingly important. Rising CO_2 and global warming, overfishing of the oceans, air pollution and declining air quality concerns, pesticides detected in the most remote reaches of the wild, together form a plethora of indicators of our ability to impact the planet on a grand scale. Paying attention to these indicators (the canaries in our global coal mine) is the key to successful conservation stewardship and to provisioning the human species with clean water, air, and food in the face of global change.

Spending substantial time in nature is a selfish act in the best sense of the word; it satisfies our human–animal longing for multisensory interactions with "other," it feeds our need to belong to a greater context beyond self, and it serves the individual and collective species purpose of increasing our survival. We humans, as part and parcel of the machinery of nature, are exquisitely designed through the process of evolution to sense, perceive, and respond to natural inputs. We, as collections of specialized receptors and integrative neural machinery replete with large cerebrums and opposable thumbs, are uniquely suited to attend and respond to changes

in the natural environment – as long as we are paying attention.

Conservation in the way the term has historically been applied may be outdated and inadequate. Lands set aside, sequestered separate from human impact, the outdoor equivalent of dusty museum dioramas, can ultimately only capture a small portion of the landscape and of the planet's biodiversity. There are few if any places on earth that are untouched by human use or impact. Whether it is pesticides traveling on the wind or acid rain, wildlife vectors transporting human-manufactured compounds, or transoceanic travelling trash, we have left an impression on the globe. Even the most wild and protected landscapes are at the mercy of wind and rain carrying compounds from the more civilized and heavily impacted corners of the globe.

Moreover, even if we managed to conserve, in this instant in time, all the land required to fulfill our conservation goals, without the political and social will to maintain that status these lands would not stay conserved for long. As human populations continue to grow, practice consumerism, and act unsustainably, so will pressure to exploit wild lands for their resource value. Conservation and nature stewardship must compete with all other human priorities. Fiscal and development decisions are made in the broader context of human priorities such as food, water, and housing security. Where nature stewardship falls on the scale of human priorities depends on providing for basic human needs while maintaining the interconnection with nature, the ultimate source.

Just as we track conservation goals and environmental indicators, the key to conservation's future may be tracking trends in human behaviors and attitudes regarding nature. The goal of this type of tracking is to examine trends and behaviors relevant to creating value in and connection to nature and the environment. The more interactive and prolonged the nature exposure, particularly when facilitated by a mentor, the more likely a higher degree of connection to nature will result. Further research is required to determine the duration, frequency, intensity, or wildness of the nature "dose" needed to instill an environmental ethic. Is viewing nature online half or one-tenth as effective as seeing that image live? How does watching a bird at a birdfeeder compare to spotting a bird on a woodland walk? There is likely to be a spectrum of outdoor activities, from those that occur outdoors but have very little to do with nature (such as a football game) to those that represent almost complete immersion in wild nature (such as solo wilderness backpacking). If the goal of tracking or encouraging nature exposure is to produce conservation ethics, then it is likely that a camping trip or a walk in the woods is a more reliable indicator of nature exposure than field hockey or a softball game. In the absence of more detailed research, determining the activities that are most relevant to instilling an environmental ethic requires keeping in mind the goal of these nature experiences: creating connection to the type of biodiversity and landscapes we would like to conserve.

Finally, in tracking trends in human behavior, it is important to remember that intentions do not necessarily translate to action, as is clear after every New Year's Day in the USA, when resolutions often fall by the wayside of habits. Therefore, surveying intentions of spending time in nature or purchases made related to nature recreation is not the same as actual time spent in nature. Buying a fishing pole that sits in the hall closet is fundamentally a different experience than going fishing. Wearing a fleece jacket and carrying a water bottle to school is not equivalent to roughing it in the outdoors. In choosing activities as indicators of environmental ethics, it is important to keep in mind that ultimately, actual experiences with nature are required to develop a later conservation ethic.

Behind all threats to biodiversity and the environment is our perceived value of and connection to nature. The basis of all political and social will for conservation is the value we place on nature. Whether motivated by aesthetics, spiritual ideals, or self-preservation, how connected we feel to nature and how much we value nature will determine the decisions we make and priorities we set regarding nature stewardship.

Biophilia Replaced With Videophilia

Historical Trends

The downtrend in US NPV was first identified in 2004 and then compared with changing cultural trends in 2006 (Pergams *et al.*, 2004; Pergams and Zaradic, 2006). After many decades of status as the iconic American family vacation, it was not until the late '80s that the popularity of NPV began to decline. So from 1939 (the start of available data) until 1987, for 50 years NPV was an increasingly popular American pastime. Popularity is measured in *per capita* attendance; in other words, attendance at national parks was increasing above that which might be expected due to population growth alone. This is important because conservation and nature stewardship decisions are made in a competing market of interests. Maintaining the same number of visitors as the US population grows is in effect losing market share, in the same way that maintaining the same number of vinyl record sales as the population grows is an example of losing market share to other forms of music media. The steadily increasing popularity, measured as *per capita* NPV between 1939 and 1987 occurred in spite of many social and economic upheavals during that time frame, including recession, depression, World War II, changing demographics, the rise of interstate travel, innovations, and inventions. The period after 1987 is equally remarkable for its steady decline in *per capita* visits (**Figure 1(a)**). Somehow, something shifted after 1987 that impacted the popularity of NPV like no cultural, social, or economic shift in the 50 years prior.

Examining the Decline

National parks represent a national-scale measure of nature recreation. The decline in this very popular activity suggests a national, social change in nature recreation. To discover the cause of the decline, a suite of potential explanatory variables with logical links to park declines were compared with park visitation data to examine overall and percent year-to-year change in trends. Comparisons required annual, national data, extending back to at least 1987, the start of the decline.

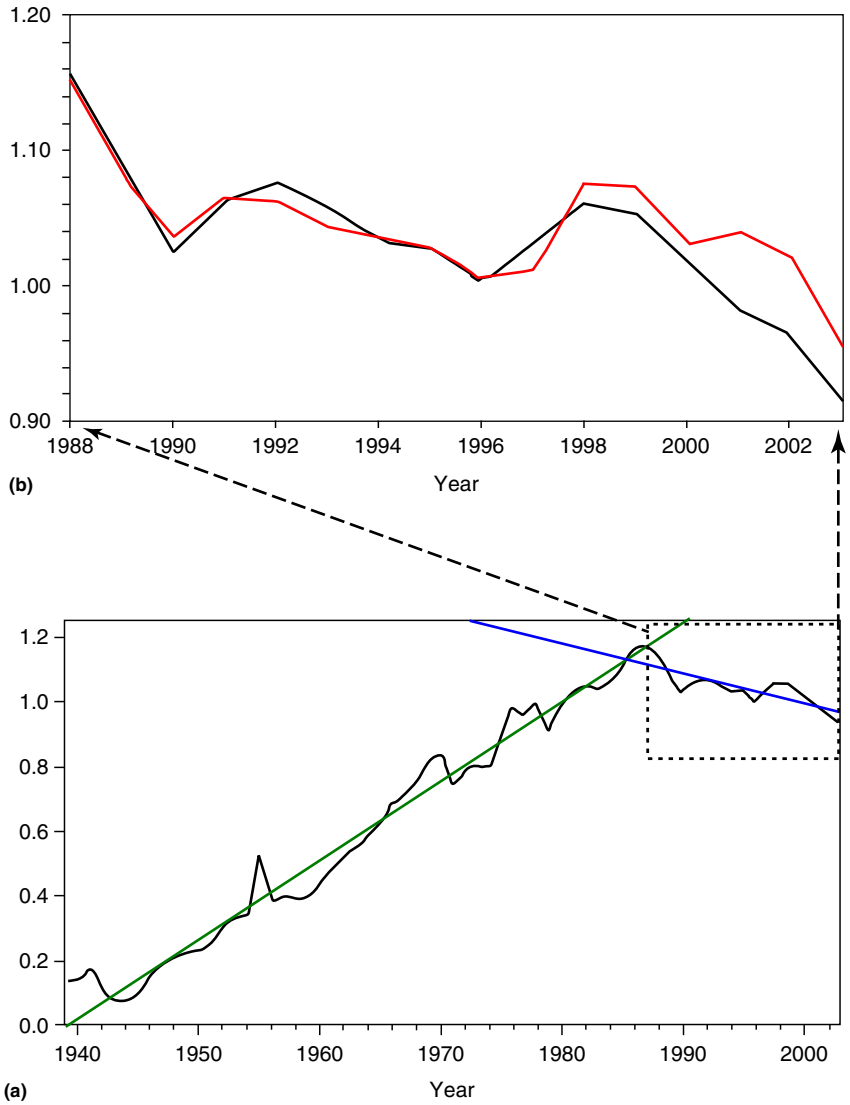


Figure 1 (b) *Per capita* US national park attendance from 1939 until 2003 is graphed in black. In green is a linear regression calculated using park attendance between 1939 and 1987; and in blue, from 1988 until 2003. (a) The attendance portion of the graph from 1988 until 2003 in the dashed box is expanded to show both the actual attendance data (black) and the multiple linear regression model (red) of four entertainment media variables and oil prices. Reproduced from Oliver RW, *et al.* (2006).

Variables were considered significantly correlated if both overall and Spearman correlations of percent year-to-year changes were significant; significantly correlated variables were further combined in a stepwise multiple linear regression model to attempt to explain and predict the decline. Comparing the year-to-year percent difference removed the possible effects of heteroskedacticity (similarity of trending time series) alone causing the significant result.

In Pergams and Zaradic (2006), 13 variables that might impact park attendance and for which the authors have data were compared with NPV data. Those variables included factors that might impede people's ability to get to the parks, potentially negative experiences at the parks, some alternative nature recreation trends, and finally the increase in electronic media use. The variables representing factors which could impact people's ability to get to the park were: oil prices (proxy for the price of gas), number of vacation days, and median family income. A potentially negative experience at the park due to overcrowding or park upkeep was measured by evaluating funding for national parks and park capacity at the 10 most popular parks. Modern alternative nature venues were evaluated by using foreign travel trends as a proxy for ecotourism and thru-hiking of the Appalachian Trail as a proxy for extreme outdoor recreation. Finally, five electronic media variables were used to measure various types of modern media recreation: hours spent watching TV, home movies, movies in theatres, video games, and the Internet.

Of the variables considered, increased use of video games, home movies, theatre attendance, and Internet combined with inflation-adjusted oil prices explains most of the 16-year decline in *per capita* US NPV (linear regression model, $r=0.975$, adjusted multiple $r^2=0.925$, SE$=0.015$, $F=37.800$, $p<.0001$,

Figure 1(b)). Other factors were either not significant or did not add to the explanatory value of the model. For example, federal funding in inflation-adjusted dollars rose during that period, the most popular parks continued to increase in visitor capacity through the mid-1990s, and the aging baby boomers were still of prime family vacation age at the time (Pergams and Zaradic, 2006).

Videophilia

Of course, correlation alone does not demonstrate causation. However, a raw r value of 97.5% in a multiple linear regression denotes a huge amount of explanatory power. Moreover, given the logical link and the extremely tight relationship between the rise of these potentially causal factors and the decline in park visitation, either the relationship is causal or both the rise in electronic media and the decline in park visitation are closely tied to some greater social shift in values and recreation choices. Finally, some of these electronic media choices essentially came into existence around the time the park visits began declining. The average person in the USA went from spending 0 h year^{-1} on the Internet in 1987 to spending 174 h year^{-1} on the Internet in 2003, and from spending 0 h year^{-1} playing video games in 1987 to spending 90 h year^{-1} in 2003 (Pergams and Zaradic, 2006). Again, home video games and Internet use essentially came into existence around the time park visitation started declining, and this increases the likelihood of causality to some extent. Altogether the average person in the USA spent 327 more hours/year on these entertainment media in 2003 than he or she did in 1987, an incredible increase in time (Pergams and Zaradic, 2006). Interestingly, television did not appear to have the same dramatic impact in relation to park visits. It appears that there is something particularly compelling about the personalized, immediate gratification of current electronic media entertainment. Most recently, baby boomers were found to average the most daily screen time at just more than 9.5 h/day, but all other age groups were strikingly similar, at roughly 8.5 h/day (Hess and Brill, 2009).

This declining trend in nature-based recreation is of enormous importance. We may be seeing evidence of a fundamental shift away from people's appreciation of nature ("biophilia," Wilson, 1984) to "videophilia," defined as "the new human tendency to focus on sedentary activities involving electronic media."

Electronic media, in addition to taking up our time in absolute terms, is essentially a sedentary activity, whereas NPV is much less so. The increase in sedentary lifestyles has been suggested as an explanatory factor for the rise in obesity in the USA (Flegal *et al.*, 2002; Skidmore and Yarnell, 2004). Time spent involved in sedentary media activities has also been implicated in reducing the amount of time spent in physical activity (Skidmore and Yarnell, 2004; Fotheringham *et al.*, 2000; Utter *et al.*, 2003). It seems likely that in addition to taking up our time, the increase in sedentary electronic media more broadly impacts recreation choices. National parks are losing their market share, whether directly to time spent on electronic media or indirectly to recreation choices favored by sedentary, electronic media consumers. Either way, this does not bode well for our society's long-term connection to nature and support of biodiversity conservation.

National and International Trends in Nature Recreation

How Widespread Is the Decline?

After publication of the national parks and videophilia research, there were some suggestions that perhaps the national park trends were due to factors limited to the US park system, such as reduced interpretive staff, changes in admission fees, or park use structure. Perhaps other nature areas that allowed alternate recreation (such as Forest Service lands that allowed all-terrain vehicle (ATV) use) were successfully outcompeting national parks for market share. Sceptics also pointed to the rise in outdoor adventure goods sales as indicators of increased nature recreation.

To determine how indicative the decline and pattern in NPV is of broader trends in nature-based recreation, Pergams and Zaradic (2008) identified 16 long-term time series assessing various types of nature-based recreation. As before, data series had to represent nationwide long-term annual measures of actions rather than intentions and had to stretch as far back as 1987, the start of the decline.

People, Places, and Activities

The nature recreation data series fell into three categories, places, activities, and people: annual visitation measured at various nature places, frequency of participation in nature-based recreation activities, and market surveys of the American people regarding natural lands visits and nature recreation to independently verify the other lines of data. By using a variety of approaches to evaluate nature-based participation, Pergams and Zaradic (2008) could determine the robustness of their initial evaluation. US visitation measures included data from: NPV, SPV, Forest Service, and Bureau of Land Management (BLMV). Measures of nature recreation activities included national rates of hiking, backpacking, and camping, as well as hunting and fishing license data. National survey data included measures of annual rates of camping, hiking/backpacking, and visits to state and federal parks. In addition to trends in the USA, Pergams and Zaradic, 2008 were able to gather NPV data from Spain and Japan.

Rather than an anomaly restricted to national parks, the results of the analysis suggest a persistent and pervasive decline in *per capita* nature recreation (Pergams and Zaradic, 2008). National survey data corroborate declines detected in natural lands visitor data at approximately the same rate over the same time frame (**Figure 2**). Long-term nature visitation data was highly significantly correlated among the various park visitation variables. Long-term nature use data sets (greater than 50 years) suggest the typical decline in *per capita* nature recreation is -18% to -25% so far, starting between 1981 and 1991, and is declining at -1.0% to -1.3% per year (**Figure 2** and Pergams and Zaradic, 2008). These similarities and the highly significant correlation among various public land visitation variables corroborate a general longitudinal

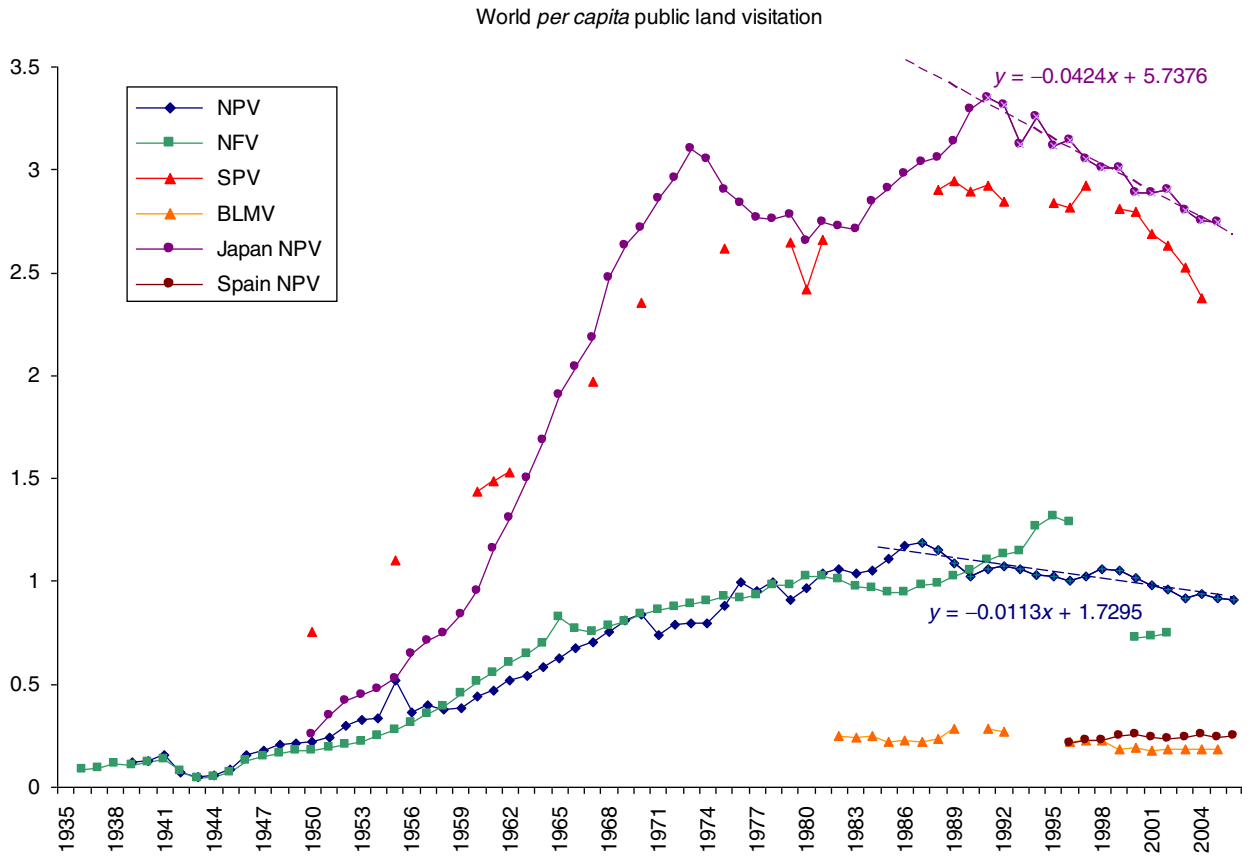


Figure 2 Annual *per capita* visitation to various US and international public lands. Included are US national parks (variable NPV, range of time series 1939–2006, N = 68), US state parks (SPV, 1950–2003, 24), US national forests (NFV, 1939–2002, 61), US Bureau of Land Management sites (BLMV, 1982–2005, 20), Japanese national parks (Japan NPV, 1950–2005, 56), and Spanish national parks (Spain NPV, 1996–2006, 11). Linear regressions for declines from identifiable peaks in NPV (1987) and Japan NPV (1991) are represented by lines and equations. Reproduced with permission from Pergams ORW and Zaradic PA (2008) Evidence for a fundamental and pervasiv shift away from nature-based recreation. *Proceedings of the National Academy of Sciences, United States of America* 105: 2295–2300.

decline in visitation to natural areas rather than an isolated decline in US NPV (Pergams and Zaradic, 2008).

International Trends

Moreover, the trend in declining nature use could well extend beyond US political and cultural boundaries. Japan's 56 years of *per capita* NPV data was among the most highly correlated with all of the long-term US public land data, in both untransformed and difference model comparisons (Pergams and Zaradic, 2008). Spain's national park data was limited to post-1995, well after declines detected in most of the other longer-term data sets. As compared with both the USA and Spain, the Japanese visit their national parks much more frequently (more than three times a year on average at peak, as compared with just more than once a year for Americans at US national parks and approximately once every 4 years for Spanish citizens in Spain, **Figure 3**). Japanese NPV trends are extraordinarily similar to those for Americans in US state parks (**Figure 2**, Pergams and Zaradic, 2008, $\rho = 0.928$, $p < .0005$); perhaps due to Japan's smaller size, Japan's national parks are more readily accessible.

The similarity between visitor patterns at Japan's national parks and US parks suggests that the cause of the decline is also similar. Of the various suggested causes for the US decline in nature recreation, very few would also hold true for simultaneous declines in Japanese nature recreation, with the exception of our common fascination with videophilia.

Nature Places Declining in Popularity

United States National Forest data and US national park data stand out among the most highly correlated time series for both the high overall correlation and the length of these two data sets (Pergams and Zaradic, 2008, $\rho = 0.931$, $p < .0005$, N = 61). Discounting the probably inflated National Forest visitor (NFV) data in the mid-1990s, both US national park and NFV show steady increases for 50–55 years before a considerable decline. Even given the differences in counting methods and missing years of visitor data in the late 1990s, it is remarkable that the last time the national forests saw *per capita* visitors as low as 2002 was almost 40 years earlier (**Figure 2**).

Most US nature exposure as detected in our data sets is through state park visitation (SPV) (**Figure 3**, on average three

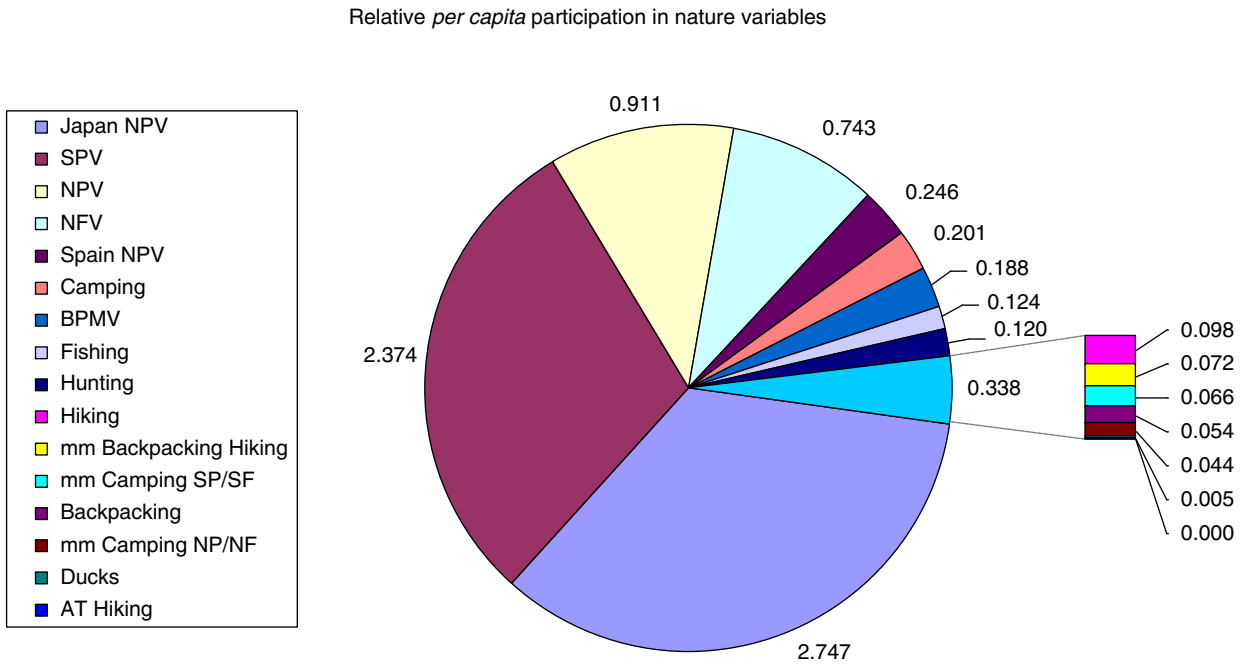


Figure 3 Annual *per capita* participation for all 16 of the nature recreation variables included in any part of the authors' analysis. The figure is meant merely to compare relative *per capita* participation in the recreation choices the authors address.

times a year per American at peak, as compared with approximately one visit a year at peak for national parks or forests). Although nationally reported numbers for SPV are sporadic, the last 15 years of data on SPV suggest a decline similar to that seen in Japan's national parks (approximately −19% total and −1.3% annual). The decline in BLMV properties, although a much smaller component of US nature exposure (**Figure 3**), is highly correlated with both the overall trend and year-to-year declines in SPV (Pergams and Zaradic, 2008).

Survey the People

US public surveys from two market survey sources independently corroborate the decline reported from park visitor counts (Pergams and Zaradic, 2008; Figure S2). The range of US data included in our comparisons covers all public nature spaces for which national visitor use is available. The fact that all of these US public land time series (as well as Japan's) are among the most highly correlated (Pergams and Zaradic, 2008) suggests that public nature spaces in the USA and Japan are generally similarly responding to changes in nature participation (**Figure 2**). Moreover, the many significant short-term correlations in declining public land use in the USA and Japan (Pergams and Zaradic, 2008) suggest that there has been a general national and potentially international fundamental shift in people's participation in nature recreation over the last 20 years.

Camping data from two market survey sources independently corroborate the decline reported from park visitor counts (Pergams and Zaradic, 2008). The Mediamark camping survey questions refer specifically to camping within national parks and forests (mmCampingNP/NF) or state parks and forests (mmCampingSP/SF). Furthermore, the decline in annual camping as detected by both of the Mediamark surveys was significantly correlated with the decline in NPV. Mediamark survey results for declining camping in national parks and forests were also significantly correlated with SPV (Pergams and Zaradic, 2008). The consistency between survey results and the trends in public lands visitors suggest that declines detected in the visitor data are not due to changing counting methods by the parks, but rather represent actual visitor declines.

Nature Activities Declining in Popularity

Camping is the largest recreation component of the *per capita* pie chart, a choice for approximately one in five Americans, more popular in *per capita* participation than hunting or fishing (**Figure 3**). As such, a trend of fewer and fewer Americans going camping is especially notable. Along with the Mediamark annual surveys, data from Statistical Abstracts surveying the frequency of camping in any venue also suggests a decline since 1987 (Pergams and Zaradic, 2008). This survey and visitor data together suggest that, rather than a change in recreation venue, we are detecting a real shift away from nature as a recreation choice.

The range of *per capita* participation in all variables is very large: each Japanese individual visits their national parks on average 2.747 times per year (351 million visits total), whereas each USA individual finishes the Appalachian Trail on average 0.000002 times per year (<700 visits total), or six orders of magnitude less (**Figure 3**). It is therefore important to realize that although all of these trends are of interest, some of them involve many more people than others, and are much more important when discussing national or global trends. The only

counter trends to the decline in nature recreation come from a small minority of hikers and backpackers. Survey data suggest that hiking in all venues increased from 0.08 *per capita* participation in 1987 to 0.098 in 2002. Most of the hiking and backpacking participation survey results were significantly negatively correlated with the general decline in nature recreation (Pergams and Zaradic, 2008). Rather than attracting a new sector to nature recreation, the small but steady growth in the hiking and backpacking market may reflect individuals who were previously campers choosing day hikes instead of overnight camping trips.

Fishing and hunting were next in popularity after camping (**Figure 3**). They are closely correlated (Pergams and Zaradic, 2008) and both increased in popularity until the early 1980s. Hunting has managed to hang on to most of its market share since its 1983 high; however, fishing has experienced a considerable *per capita* decline (-25% from its 1981 peak, an average of -1.0% a year). This may be related to various overfishing and pollution issues decreasing access to fish populations, contrasted with exploding deer populations (which are largely anthropogenic effects). The decline in fishing is highly significantly correlated with the decline in visitors to US public lands since 1987 (Pergams and Zaradic, 2008). The US duck stamps time series is a subset of the much larger recreation category of US hunting licenses (*per capita* participation 0.005 vs 0.120, **Figure 3**). Duck hunting regulations are sometimes more complex, often have higher equipment costs (decoys, boats, and dogs), and often require access to relatively rare habitat (wetlands) than many other forms of hunting (Stephens, 2007). This may in turn relatively deter recruitment of young duck hunters (Babcock, 2007). We would add as conjecture that although the number of ducks in the USA has only recently increased (e.g., $+14\%$ in 2007; Ducks Unlimited, 2007), the number of deer has been exploding for a long time (Stromayer and Warren, 1997).

Pervasive Decline in Nature Recreation

In conclusion, all major lines of evidence point to a general and fundamental shift away from people's participation in nature-based recreation with most reliable long-term *per capita* visitation measures of nature recreation peaking between 1981 and 1991, and declining at about -1.2% per year, and total to date 18–25% (Pergams and Zaradic, 2008). In addition, the cultural shift away from recreation in nature appears to extend outside of the USA to at least Japan. Our research suggests that the root cause is likely to be videophilia. Other factors may be responsible, but they would have to be large enough in scale and potentially international in impact, as well as timely enough in instigation, to generate this type of shift across the USA and potentially around the globe. Regardless of the root cause, the evidence for a pervasive shift away from nature-based recreation is clear.

Implications for Health and Childhood Development

Direct, Indirect, and Vicarious Nature

Declining nature experiences for people across the USA and potentially internationally is likely to have broad implications for conservation. On the surface, declining *per capita* visits could decrease impact on natural areas and in that way seem beneficial. However, decreasing affinity for nature areas corresponding to increasing videophilia seems likely to present much greater long-term costs in loss of the constituency of support for conservation.

Teachers, environmental education, and exposure to nature through other venues (such as reading about nature) are credited with influencing environmental sensitivity, but to a much lesser degree than direct actual experience of natural areas (Chawla, 1998, 1999; Kahn, 2002; Kellert, 2002; Duda *et al.*, 2003; Wells and Lekies, 2006). Similarly, direct sustained contact with nature best cultivates children's environmental knowledge and concern (Vaske and Kobrin, 2001; Fisman, 2005). Direct experience with nature appears to uniquely effect childhood development as compared with other types of nature encounters (Zaradic and Pergams, 2007).

Kellert (2002) describes three modes of experiencing nature: direct, indirect, and vicarious. Direct experience involves actual physical contact with natural settings and nonhuman species. This is the spontaneous play in a forest, creek, neighborhood park, backyard, or even vacant lot. Although these settings are impacted by human manipulation to some degree, they function largely independently of human intervention. Indirect experience of nature involves physical contact but in a much more controlled and restricted setting. Examples include zoos, nature centers, aquariums, natural history and science museums, and domesticated animals such as cats and dogs. Vicarious experience of nature occurs in the absence of actual physical contact with the natural world. Nature in this form ranges from stylized and symbolic art to photographs, videos, and virtual webcam tours of natural areas.

A comparative review by Kellert (2002) suggests that of the three modes of contact, direct experience with nature plays the most significant role in cognitive and evaluative development. Direct experience of nature offers a multiplicity of sights, sounds, smells, and tactile stimuli shifting continuously in space and time. The spontaneity and complexity of these sensory experiences engage a wide range of adaptive and problem-solving responses, requiring alertness and attention (Sebba, 1991). In contrast, the more structured, indirect experiences of nature do not require the same level of spontaneous engagement and do not exert the same types of long-term developmental impacts on children (Kellert, 2002; Pyle, 2002). Similarly, Wells and Lekies (2006) found that direct contact with "wild" as compared with "domesticated" nature before age 11 is a particularly potent pathway to shaping environmental attitudes and behaviors in adulthood. Indirect experiences may exert the greatest positive effect in conjunction with direct encounters in familiar natural settings (Kellert, 2002). The least sensory engaging and spontaneous type of nature contact, vicarious experience through an electronic outlet, is becoming more prevalent. More families live in or near urban settings with more time committed to wage earning (Duda *et al.*, 2002). Internet, video game, and home movie use continues to increase dramatically as direct contact with nature via outlets such as national parks decreases (Pergams and Zaradic, 2006, unpublished data). What are the long-term impacts of this trend for childhood development and conservation?

A Sedentary Society

Increasing time spent in sedentary indoor media activities has been implicated in the decline in time available for outdoor physical activities (Skidmore and Yarnell, 2004; Fotheringham *et al.*, 2000; Utter *et al.*, 2003). An increasingly sedentary lifestyle in the USA is one of the explanatory factors in the rise in obesity (Flegal *et al.*, 2002; Skidmore and Yarnell, 2004). Frequency of being overweight and obese in US children older than 6 years and adolescents has doubled in the last two to three decades (Deckelbaum and Williams, 2001). Moreover, the spread of this behavior is resulting in similar doubling rates of obesity in developing countries (Deckelbaum and Williams, 2001). Even younger children are at risk because of this increasingly sedentary lifestyle, with more than 22 million children younger than 5 years of age overweight across the world (Deckelbaum and Williams, 2001).

Adult activity choices are significant for childhood development because children often depend on adults both for access to nature and as role models for recreational choices. Dong *et al.* (2004) provide an index of how much energy we currently spend on various activities during a 24-h day in our sedentary society. In 1992–1994, 7515 adults (weighted to be representative of the contiguous 48 states) were surveyed to create a detailed report of each activity performed in the previous 24 h. An energy expenditure index was created by multiplying duration and intensity of each activity by each individual and summing across all individuals. The top five ranked activities (representing 42.9% of all energy expended) are: driving, office work, watching TV/movie, childcare, and activities performed while sitting quietly. "Watching TV/movie, home or theater" and "activities performed while sitting quietly" (presumably including time on the Internet and playing video games) were together 14.4% of the total. The study was not designed to compare indoor versus outdoor activities, but it is still interesting that the two clearly outdoor recreational activities ("fishing and hunting" and "gardening") were together only 1.5% of the total, and ranked 22nd and 27th, respectively. Given the activity choices adults are making, it should be no surprise that children are also spending less time in outdoor recreation as compared with time in front of a video screen.

Children and Videophilia

American Academy of Pediatrics guidelines recommend 0 h/day of screen time for children less than 2 years old and $\leq$2 h/day for older children. The guidelines also recommend that there be no electronic media at all in young children's rooms. In contrast, in a 2003 survey of 1000 individuals, parents of young children aged 0.5–6 years, Rideout *et al.* (2003) had a number of startling findings. In a typical day, 43% of children $<$2 years old watch television every day, and 26% have a television in their bedroom; 68% of children $<$2 years old spend $>$2 h a day using screen media. The larger sample (children aged 0.5–6 years) used screen media a little under 2 h/day. Almost all homes (95%) with children 0.5–6 years old had at least one VCR or DVD player, and 27% of children had one in their bedroom. On average, these children spend 40 min/day watching home movies, with 25% watching videos every day. Video games still seem to start a little later, but by age 4–6, 50% of children play them, and 25% play several times a week or more, with an average time played of a little more than 1 h/day. These 2003 values are probably doubled or tripled now.

Children and the Natural Environment

Research has recently begun to explore children's relationships to the natural environment. Children younger than 13 years living in the USA reportedly spend on average only about half an hour of unstructured time outdoors each week, consisting of gardening, boating, camping, picnicking, pleasure drives, walking, and hiking (Hofferth and Sandberg, 2001). Bartlett (1997) describes behavioral and emotional problems as a result of lack of opportunity to play outside. Hüttenmoser (1995) found that 5-year-old children limited in playing outdoors by automobile traffic exhibited poorer social, behavior, and motor skills and had fewer playmates than children not so limited. Grahn *et al.* (1997) found that children attending a day care facility surrounded by orchards, pastures, and woodlands (and where the children went outdoors every day regardless of weather) had better motor coordination and greater attention capacity than did children who attended an urban day care center surrounded by tall buildings. Wells (2000), in one of the few studies to employ a longitudinal design, found that children who moved to housing with more nature nearby tended to have higher levels of cognitive functioning than children who moved to housing with less nature. Taylor *et al.* (2001) found that activities in green settings tend to lower the symptoms of children with attention deficit disorder (ADD). Taylor *et al.* (2002) examined the effect of nearby nature on girls aged 7–12 years living in Chicago public housing. The relative naturalness of window views was predictive of the girls' self-discipline, as defined by ability to concentrate, inhibit impulses, and delay gratification.

Wells and Evans (2003) interviewed 337 rural third to fifth graders to determine whether vegetation near their homes provided a buffer to stressful situations such as family relocation, being picked on or punished at school, or being subject to peer pressure. The children's emotional well-being was assessed by both the children themselves and their parents. Stressful life events had less impact on psychological distress under "high nature" conditions than under "low nature" conditions. This buffering appears to be greatest for those at most risk: those experiencing the highest levels of life stress.

Early TV and ADD

Christakis *et al.* (2004) applied logistic regression to the National Longitudinal Survey of Youth, containing data for some 2600 children at ages 1 and 3. The hypothesis tested was that early television viewing (at ages 1 and 3) is associated with attentional problems at age 7. Children watched an average of 2.2 h (standard deviation (SD): 2.91) of television per day at age 1 and 3.6 h (SD: 2.94) of television per day at age 3. They corroborated the hypothesis even after controlling for a number of potentially confounding factors, including prenatal

substance abuse, gestational age, maternal psychopathology, and socioeconomic status. A 1-SD increase in the number of hours of television watched at either age 1 or age 3 was associated with a 28% increase in the probability of having attentional problems at age 7.

The developmental physiological mechanisms speculated to be responsible relate to the rapid development of the brain in young children. Environmental variables, including visual and auditory experiences, are thought to greatly influence the number and density of neuronal responses (Turner and Greenough, 1985; Greenough *et al.*, 1987; Wallace *et al.*, 1992). However, in contrast to the pace with which real life unfolds, television can portray rapidly changing scenes and events. Also, this speed of portrayal is not at all chronologically constant, but rather in itself varies greatly from moment to moment, scene to scene, show to show, movie to movie. Thus this unevenness and rapidity in depicting the passage of time, during the critical period of synaptic development in early childhood, is speculated to affect the number and density of neuronal responses (Christakis *et al.*, 2004). Although no similar research exists specifically for video games or movies, given their similarity one would be surprised if the same factors did not apply.

Internet and Socialization

One of the major trends in electronic media is the increase in use by children of younger ages. The National Center for Education Statistics (USDE, 2005) used a weighted sample representing approximately 58.3 million children aged 3 years and older in nursery school through the 12th grade in October 2003. A survey of this sample describes use of computer and Internet technologies by age group: 67% of children in nursery school were already computer users, as were 80% of those in kindergarten. About one-quarter (23%) of nursery school children were already Internet users, rising to about 32% in kindergarten. By high school, + nearly all students (97%) use computers and a large majority (80%) use the Internet (USDE, 2005).

The exponential rise in Internet use over a mere two decades from its initial availability offers little time to study its long-term developmental effects. However, these trends in adult Internet use at home clearly must have direct implications on time spent interacting with children, and indirect implications for parents as role models and gatekeepers of children's recreational choices. Nie and Hillygus (2002a) explore the effects of the Internet on interpersonal communication and sociability, collecting data from time diaries of their subjects. Internet use at home (but not at work) is shown to have a strong negative impact on time spent with friends and family as well as time spent on social activities. Similarly, Internet use during weekend days is more strongly related to decreased time spent with friends and family and on social activities than Internet use during weekdays. The relationship is highly significant: for every hour spent online at home or on the weekend, there is a corresponding 41 min less spent with family members ($\beta = -0.69$, t-stat $= -6.65$, $p < .001$). Nie and Hillygus (2002b) refined the parameters of this reallocation: in other words, what activities are affected, rather than in whose company they are performed. Not surprisingly, Internet time seems to be reallocated most from discretionary activities, especially social activities, hobbies, reading for pleasure, and television viewing. However, Internet time is also reallocated from nondiscretionary activities such as housework and childcare, especially by heavy (≥ 61 min/day) Internet users. These findings seem to strongly corroborate the hypothesis that the Internet has created a shift in people's time allocation away from family and friends, and that this loss of time is suffered by a wide spectrum of specific activities.

Kraut *et al.* (1998) examine the social and psychological impacts of the introduction of the Internet to a household. The behaviors of 169 people in 73 households were tracked over the first 1–2 years of Internet use. Again, greater use of the Internet was associated with declines in communication with family members and in reduced social contact outside the home. In addition, increased Internet was associated with increases in loneliness and depression.

Media and Fear

Television and other electronic media probably provide most incidental information about nature (Bixler and Carlisle, 1994). As well as transmitting neutral information about nature, media depiction of nature's hazards can teach fear. Fears of evolutionarily relevant natural dangers such as snakes and spiders seem easier to acquire and harder to extinguish than fear of post-technological dangers such as cars and guns (Heerwagen and Orians, 2002). Children exposed to realistic media depictions of life-threatening events such as fire and drowning report feeling more at risk from those events and less likely to engage in activities related to the depicted tragedies (Bixler and Carlisle, 1994). Hence media depictions of nature tend to focus on and sensationalize nature's hazards, promoting the perception that natural areas are dangerous.

Similarly, for parents, sensationalism in the media has greatly reinforced paranoia over children being abducted while playing outdoors. Finkelhor *et al.* (1992) report 200–300 stereotypical criminal child kidnappings/year (approximately 0.00008% of the population) as opposed to the tens of thousands per year suspected in the public imagination. The best inoculation against inaccurate and sensationalized depictions of nature is an accurate frame of reference for nature built from early firsthand experiences (Bixler and Carlisle, 1994).

Children's firsthand experience in natural settings plays an important role in developing environmental attitudes and preferences for nature (Heerwagen and Orians, 2002). Without early and regular exposure to nature, urban dwellers find the unfamiliarity of wilderness settings uncomfortable and overwhelming (Bixler and Carlisle, 1994), preferring built settings to natural environments (Heerwagen and Orians, 2002).

Biophilia must be nurtured to truly take hold (Verbeek and de Waal, 2002). Active behaviors that require direct involvement with live animals, such has fishing, hunting, and birdwatching, lead to the most consistent environmental attitude and knowledge (Duda *et al.*, 2003). Values such as reverence and respect for nature appear to be derived in part from direct involvement (Matthews and Riley, 1995; Duda *et al.*, 2003).

Conservation and Virtual Nature

One of the most potentially controversial issues pertaining to videophilia is the issue of "virtual nature," here defined as "nature experienced vicariously through electronic means." What are the costs, what are the benefits? For there are indeed (at least in theory) benefits. Certainly, the Internet has provided unprecedented access to natural areas for many people. From a viewpoint solely focused on *Homo sapiens*, some of the relaxing and educational aspects of nature can in fact be delivered through vicarious experience (Levi and Kocher, 1999) on a video screen, or as technology advances, through a virtual reality space. An example is the National Park Service webcam showing Old Faithful in Yellowstone National Park (National Park Service, 2006). In particular, children, not able to travel to the Amazon rainforest or Old Faithful on their own, can do so online without the need for car, plane, or parent. Bugs, snakes, spiders, predators, and pests can all be left behind in virtual nature. There is even a little something for conservationists: endangered species and sensitive sites can be visited with little or no impact on habitat via webcam, and there is greatly reduced maintenance of natural areas when they do not have any real visitors. Lastly, whether or not we believe in the potential benefits of videophilia to conservation, we must unfortunately at least admit the possibility that the very strong current increase in videophilia (Pergams and Zaradic, 2006) is a juggernaut in whose way we cannot stand. This might leave virtual nature as the best related conservation option.

The potential costs of virtual nature to conservation are, of course, huge, and may be compounded in future generations. Environmental awareness, primarily nurtured through parents and adult mentors while outdoors (Chawla, 1998), may be the cost of replacing real nature with virtual nature, and this possibility should not be underestimated. Moreover, today's children are tomorrow's parents, but potentially with greatly decreased connection to nature. If children experience Old Faithful primarily through a webcam, where is the nurturing connection that in the past was provided by sharing the experience with family or an adult mentor? If parents and children are sharing the experience of virtual nature, will the nurturing connection be to the subject matter (nature) or to the media (the virtual experience)? From a conservation context, will they still go see it in real life? Will they still pay their tax in dollars to maintain Yellowstone?

Lacking direct experience of our natural environment, we lack the most immediate feedback of our impact on nature. Experiencing nature vicariously, we may not be aware of the sensory experiences we are missing through the virtual medium, but the real nature we leave outside will continue to receive the footprint of our presence on the planet. If we spend our recreational time on virtual tours, will we still pick up litter at our local parks? Or will we opt for solutions consistent with our simulated experience of nature? In Los Angeles, it was proposed to replace highway median plantings with plastic trees, rather than deal with the larger issue of air pollution (Levi and Kocher, 1999). In the long run, spending less time in real nature seems likely to speed the onset of environmental generational amnesia. This psychological phenomenon, described by Kahn (2002), is the tendency for each generation in its youth to take the existing environmental conditions as normal, although with each ensuing generation the amount of environmental degradation increases.

If we generationally redefine normal conditions, we might well increasingly use experiences of vicarious nature as our benchmark rather than the local environment. A study of the effects of simulated experience through commercially available nature images suggests this type of vicarious contact with spectacular natural environments increases support for national parks and forests but decreases support for preservation and acquisition of local natural areas (Levi and Kocher, 1999). Thus vicarious nature experiences tend to sensationalize nature's spectacular habitats, generating the perception that local natural areas are relatively lackluster.

Given that the frequency of direct contact with nature depends on local availability (particularly for children, due to their dependence on caregivers for access; Louv, 2006), the devaluing of local natural areas as compared with sensationalized virtual views of nature seems likely to further reduce direct experiences, and perforce increase our reliance on vicarious nature experiences. Moreover, the recent decline in *per capita* NPV (Pergams *et al.*, 2004; Pergams and Zaradic, 2006) suggests that although vicarious exposure to spectacular habitats may increase the perceived value of national parks, this increased value does not translate to increased direct experiences. In fact, the results of the authors' research suggest just the opposite, that increased use of virtual electronic media is significantly correlated with decreased direct experiences with national parks (Pergams and Zaradic, 2006). Concurrent with the almost 25% decline in NPV, Internet use increased from none in 1987 to a per person annual average of 174 h by 2003, and average overall annual electronic media use (home movies, theatre movies, video games, and Internet use) rose by an average of 327 h per person (Pergams and Zaradic, 2006). Thus virtual access to nature is, in our data, already very strongly associated with loss of direct contact with actual nature. In other words, visiting Old Faithful on the web more seems to result in visiting Old Faithful in person less.

Videophilia Implications for the Future

Videophilia has both direct and indirect implications for childhood development and the future of conservation. We may well be seeing evidence of a fundamental shift away from biophilia, "the innate tendency to focus on life and lifelike processes" (Wilson, 1984, elaborated with Kellert and Wilson, 1993, and by Kahn, 1997) to videophilia. Moreover, adult attitudes toward the environment are often nurtured through family vacations and spending time outdoors with family and mentors (Chawla, 1998). If parents, mentors, and children are indeed spending less recreational time in nature (Louv, 2006), the implications for conservation and impacts on childhood development may be compounded in future generations.

Most directly, increasing videophilia has been implicated in negative psychological and physical effects. Some of the negative effects are linked to the sedentary nature of videophilia and reduced time available for outdoor physical activities and nature experiences. Conversely, outdoor play and nature experience have proven beneficial for cognitive functioning, reduction in symptoms of ADD, increase in self-discipline, and emotional well-being at all developmental

stages. Other detrimental effects of videophilia seem to be more related to its potential for isolation and much faster pace than real time. High levels of children's electronic media consumption is correlated with attentional problems and increases in loneliness and depression.

Potentially more complex are the implications of virtual nature. Virtual nature, here defined as "nature experienced vicariously through electronic means," has some potential benefits, particularly for children who are dependent on adults for access to many natural areas. Although a sedentary videophilic outlet, the Internet provides unprecedented access for children to natural areas, delivering some of the relaxing and educational aspects of nature vicariously. Yet, given increasing Internet usage by younger children, virtual nature appears to directly compete with time previously allocated to the more beneficial direct contact with the outdoors. In addition, virtual nature experiences may further discourage direct contact with local nature by sensationalizing nature's hazards and habitats, generating the perception that local natural areas are simultaneously both dangerous and lackluster.

Ultimately, to more conclusively resolve the long-term impact of videophilia, further research is needed to answer the questions of exactly what videophilia means to (1) children's development, (2) their success and happiness as adults, and (3) specifically, their environmental consciousness. The authors propose the initiation of a large-scale longitudinal study to determine the effects of videophilia on children's development and future outcomes. Such research would follow controlled groups of children from birth to adulthood, and quantify differences in as many areas as possible, correlated to exposure to nature and/or videophilia. Physical and mental health, educational achievement, career choices, and economic success as adults would all be among areas considered in addition to environmental awareness. The proposed research would require the cooperation of a number of experts from a wide variety of fields. As daunting as the proposal might seem, children are already exposed, at earlier and earlier developmental stages, to an uncontrolled videophilia treatment. More daunting is the prospect of today's children (tomorrow's parents) in a culture devoid of contact with the evolutionary driver and life-support system that is our natural world.

Conservation's Current Base of Support

A Narrow Base

Coincident with the decline in popularity of nature-based recreation, a review of recent trends suggests that current public support for conservation has a narrow base, and that existing support is not strong enough to make conservation or the environment a high priority (Earthjustice, 2005; Nordhaus and Shellenberger, 2007). A 2004 survey of the electorate determined that Americans are increasingly indifferent to environmental issues, with the environment ranking lower than most other voter priorities (Shellenberger and Nordhaus, 2004; Earthjustice, 2005; Nordhaus and Shellenberger, 2007). Finally, populations globally and within the USA are increasingly urban, meaning there simply is less contact with nature in any form (UNPD, 2005; Kareiva, 2008). All of these trends combine to suggest that one of the greatest threats to conservation may be declining public support due to a progressively smaller population engaged in outdoor recreation and as a result less willing to support conservation activities.

Not All Nature Recreation Is Equal

Not all types of nature recreation are equal when it comes to generating conservation support. Some types of nature recreation are more likely to generate a strong constituency of conservation supporters. In addition to the impact of broader economic forces on support for conservation (Pergams *et al.*, 2004; Butler, 2009), participation in specific types of nature recreation may be more indicative of conservation support, potentially with some time delay between the participation and the support. Zaradic *et al.*, (2009) used annual contributions to four large conservation nongovernmental organizations (NGOs: The Nature Conservancy, World Wildlife Fund, Sierra Club, and Environmental Defense Fund) as a longitudinal measure of conservation support. The nature exposure time series were updated from those utilized in Pergams and Zaradic (2008). Conservation contributions reflect the relative priorities of the donors, as well as the state of the US national economy and changes in both corporate and personal income and wealth (Pergams *et al.*, 2004). As appropriate, values were weighted *per capita* and/or inflation-adjusted to 2001 dollars using the appropriate time series of gross domestic product (GDP) deflators from the US Federal Reserve.

Contributions to the major conservation, NGOs were positively correlated with *per capita* backpacking and hiking 11–12 years earlier, but negatively correlated with less elite outdoor activities, such as public lands visitation or fishing. A chart comparing *per capita* annual conservation NGO contributions with *per capita* backpacking and hiking time series is given in **Figure 1**. The backpacking and hiking series are shown lagged in time to match the significantly correlated year of investment. Backpacking/hiking time series show longer-term increases, but with peaks in 1998–2000 and declines since then.

The declines in backpacking/hiking since 2000 could imply a significant problem for conservation support. Backpacking is defined in the survey data as 1+ overnight trips, including wilderness camping, so backpackers require time and gear, both of which are relatively expensive. The two sets of survey data for backpacking, and the most extreme version of backpacking (complete thru-hiking of the 2000+ mile Appalachian Trail) are measures of similar, exclusive, high-cost and time-intensive nature recreation. Backpacking tends to be more attractive to younger age groups. Most backpackers are in the 25–44 age range and overwhelmingly European American (NPS, 2001; Cordell *et al.*, 2005). Using the most parsimonious interpretation of the time lags, these primarily European American former backpackers become a significant source of NGO dollars 11–12 years after backpacking.

The Two Americas of Conservation Support

The type and timing of nature exposure makes a great deal of difference when it comes to creating later conservation support.

The interpretation of Zaradic *et al.*, 2009 is that there are effectively two Americas when considering the pathway from nature exposure to conservation support: an elite backpacking/hiking group and a broader public lands visitation group. If this is true, then it has profound consequences for future generations and prospects for conservation support. Conservation organizations seem to be receiving donations from a very narrow group of relatively elite outdoor enthusiasts. The three variables that were highly positively correlated with conservation NGO contributions are all variations of backpacking. Backpacking represents the least popular of the four classes of recreation variables studied (Pergams and Zaradic, 2008). The current *per capita* rate of backpacking is 0.054; in other words, the average American goes backpacking once every 18.5 years (Pergams and Zaradic, 2008). Each hiker or backpacker translates to $200–$300 annually in future NGO contributions. Given that most backpackers are 25–44 and adding the 11–12 year time lag, this would most likely be in middle age, presumably near their income prime. The demographics of this group are consistent with the description of the small fraction of the electorate that considers the environment a top priority: overwhelmingly European American, mostly college educated, higher income, and older than 35 (Earthjustice, 2005). Furthermore, based on the lagged impact of hiking/backpacking on investment, conservation NGOs have been benefiting from the tail end of a decade-old rise in the popularity of backpacking and hiking (**Figure 1**). The most recent data show a decline in hiking/backpacking popularity since 1998–2000 (Pergams and Zaradic, 2008). The authors project the negative effect of reduced hiking/backpacking frequency on NGO revenues to begin in approximately 2010–2011, and to continue through at least 2018 (**Figure 1**).

In contrast to backpacking, nature recreation at public lands appears much less likely to translate to conservation NGO investment. Moreover, there is apparently relatively little backpacking and hiking occurring during public lands visitation. For example, a recent survey of 849 Yellowstone NPV groups asked what primary activities were their reasons for visiting the park (Manni *et al.*, 2007, Figure 43). Seven activities accounted for 92% of reasons: sightseeing/taking a scenic drive (59%), viewing wildlife/birdwatching (16%), boardwalk/geyser basin (9%), camping in developed campgrounds (3%), day hiking (3%), viewing roadside/trailside exhibits (1%), and overnight backpacking (backcountry) (1%). All other activities were $<1\%$ of reasons. With backpacking and hiking combined being only 4% of the reasons, people gave for visiting Yellowstone, there is probably little ambiguity between public lands visitation data and data acquired specifically about backpackers and hikers.

In addition, there appears to be a large income gap of (very roughly) 30–70% between backpacker/hikers and public lands visitors. Based on a 30-year review of outdoor recreation literature by the US Department of Agriculture (Loomis, 2005), average income (in 1996 inflation-adjusted dollars) of backpacker/hikers is $70,222. In contrast, the median household income for SPV and NPV (in 1996 inflation-adjusted dollars) is $42,100–$44,438 (South Carolina DPRT, 2001). The average income of sport fishermen of $43,410 is comparable with that of park visitors. Similar income gaps can be detected using market surveys of specialty magazines: *Field and Stream* for fishing/hunting (2008 annual circulation 16,500,000) has an average reader income of $54,838 (Echo Media, 2009a), whereas *Backpacker* magazine for hiking/backpacking (circulation 2,790,000) has an average reader income of $70,950 (Echo Media, 2009b).

Increased exposure to nature through public land visits (SPV, NPV, NFV, and BLMV) is significantly negatively correlated with conservation NGO contributions. Fishing is strongly associated with public lands; typically 80–85% of all fishing takes place on public water bodies (State of Virginia, 2007). People who get their primary nature exposure from visiting public lands or fishing are much more diverse in age and ethnic composition (NPS, 2001; Cordell *et al.*, 2005). For example, in the African American community, fishing is more than seven times as popular (in *per capita* participation) as backpacking and more than twice as popular among the Hispanic community (Cordell *et al.*, 2005). Similarly, NPV is approximately three times more popular among Hispanics and four times more popular among the African American community as compared with backpacking (NPS, 2001). The negative correlation between public land visitation and NGO contributions appears to hold both for short-term conservation investment (4–7 years after the nature experience) and for the longest-term measures we could obtain (approx. 16 years later). The authors hypothesize that people are more likely to invest in what they know from firsthand experience. Indeed, it may be that high levels of public land recreation might create a sense that access to the outdoors is what is important rather than preserving less accessible landscapes through conservation NGOs. Whatever the reason, it isimportant to acknowledge that all nature participation is not equal in terms of generating support for conservation, and that to the extent that conservation needs to broaden its support base, there is a need to understand what type of nature recreation creates the strongest commitment to conservation.

Need to Diversify

Given the trends of increasing US population diversity, urbanization, and economic and cultural changes, the authors fear that the currently narrow base of conservation NGO supporters will become even narrower. To avoid becoming marginalized, the conservation movement will need to diversify its outreach strategy, engaging novel and diverse constituencies. Strategies for doing so may require either more of the "right types" of nature exposure or entirely different approaches to ethnic or socioeconomic groups who are not likely to engage in hiking and backpacking. Ultimately, the fate of biodiversity and intact ecosystems may depend less on rates of habitat loss or invasive species than on public perception of whether conservation should be supported at all.

Demographics and Conservation's Future Base

Current conservation support depends primarily on a very narrow group of relatively elite outdoor enthusiasts: overwhelmingly European American, mostly college educated, higher income, older than 35, former backpackers and hikers 11–12 years after their peak of outdoor participation (Zaradic *et al.*, 2009). Given the trends of increasing US population diversity, urbanization,

and economic and cultural changes, this narrow base of conservation supporters is likely to become even narrower (Zaradic *et al.*, 2009). By 2042, the non-white population in the USA is predicted to become the majority of the population. Today's young people are leading this demographic shift, that is, 40% of US youth aged 25 and younger are non-white, 44% of youth aged 18 and younger are non-white, and 50% of children aged 5 and younger are non-white (US Census, 2010). Nationally and globally, the field of conservation will depend on the support of this coming generation.

Speaking to and attracting a younger and more diverse urban constituency requires targeted outreach. Programs such as the The Nature Conservancy's Leaders in Environmental Action for the Future (LEAF) program target young people characteristic of the growing demographic shift in diversity, and at a young adult age during which outdoor participation tends to decline sharply (Outdoor Foundation, 2010). The primarily (85%) non-white LEAF participants and alumni represent the types of diverse future leaders and environmentalists that will be needed to support global conservation goals. Spending sustained time in nature, particularly with a mentor, is thought to be one of the foundations for nurturing the types of environmental attitudes and values leading to long-term conservation support (Bögeholz, 2006; Lang, 2006; Wells and Lekies, 2006; Zaradic and Pergams, 2007). The LEAF partnership model is based on providing a sustained nature experience, facilitated by mentors. By combining what students are learning in environmental high schools with real-world nature stewardship experience in the field with a mentor, The Nature Conservancy's LEAF seems to create an ideal environmental foundation (Bailey *et al.*, 2004; Rickinson *et al.*, 2004) for its participants. LEAF alumni are outdoor enthusiasts, environmental advocates, and volunteers, and they influence their peers and community (Zaradic and Pergams, 2011).

Conservation Based on Purely Selfish Motives

Ecosystem Services

Against a background of increased videophilia and concurrent decreased nature participation, we have seen a new field of applied research come into being. The new field of ecosystem services attempts to economically value the biodiversity that directly serves human needs. To value nature's systems completely and coherently in our chaotic world is, of course, impossible. We try anyway, for the sake of the many people who place little or no intrinsic value on nature. Some of these people are from developed countries, and may have already become distanced from nature though videophilia. Others are from less-developed countries and may be forced to prioritize immediate human needs (like food and shelter) over conservation.

In any case, ecosystem services valuations exist so that conservationists are allowed at the table where ministers of commerce, transportation, defense, etc. sit, in an attempt to alter their behavior and policies. Until conservation is translated into monetary value, it is viewed by most people as a curiosity. As a result, conservationists need to remind everyone that looking only at immediate commercial payoffs (say, from fisheries) will consistently undervalue biodiversity, given the enormously broader range of tangible and intangible services fisheries provide. The challenge is to deduce the value that people do or should place on species (should, because species are often providing indirect benefits not immediately appreciated by observers), including the intangible aspects of that value. The value of biodiversity to human health consists of such undervalued and little-discussed benefits.

Health Benefits

In developed nations, local ecosystems are degraded with the assumption that health care and alternate access to provisions ameliorates any health impact. Many developed nations enjoy a fair amount of *per capita* investment in health care (United Nations, 2007, 2009), and the average lifespan, prenatal care, and few disability days in those countries reflects a high level of health. As a result, in developed nations, mortality is increasingly related to environmental issues and a sedentary lifestyle (WHO, 1997, p. 132, 2000), including diseases such as coronary disease, cancer, and diabetes (WHO, 2008).

There may be some nonsubstitutable health benefits of active exposure to intact ecosystems. Parks and other local green spaces, Internet/webcam access to wild nature, and outdoor recreation campaigns may be attempts at creating a substitute for the benefits of access to intact ecosystems. The marginal health value and the degree of substitutability of these nature access alternatives remain unquantified unknowns. However, the increasing mortality related to sedentary lifestyle and stress-related disorders suggest that the quantity and quality of current substitutes is inadequate. If regular access to intact ecosystems can be shown to address the medical ills of modern developed life, then this supports a market for conserving local intact ecosystems and also, through ecotourism, for intact ecosystems throughout the globe.

There appear to be links between videophilia and decreasing participation in nature-based recreation. There is a further body of work that suggests active exposure to nature reduces stress and provides other aspects of physical and psychological well-being. It is possible that access to intact ecosystems may provide the greatest added health benefit in those very developed areas where intact ecosystems are most rare. Although the added health benefit in years of longevity may be relatively small, it is of the greatest marginal benefit to the wealthiest sector of the global population: those on the most developed, highest-GDP parts of the world. It may be that those areas that have access to the best modern medical care in fact have the most marginal health benefits to gain from access to intact nature. If videophilia and its accompanying ailments are the side effects of living in an urban, developed environment, and regular access to nature is the cure, then the same market forces that drive exploitation can be harnessed to drive conservation.

Proposed Research

In closing, the authors would like to suggest research uniting most of the themes in this article. They propose construction

of a complex, multivariate model with the capacity to predict conservation effort in the USA with a high degree of accuracy. Such a model would elucidate the underlying economic factors and nature experiences that sustain conservation, and show how these factors relate to specific aspects of natural systems. This model would be of enormous use to conservation entities of all sorts, from governmental agencies to large NGOs to grassroots organizations. It would also further our understanding of the social and economic motivators driving the US conservation movement.

In addition to evaluating national-scale longitudinal trends, the authors suggest identifying smaller populations with comparable nature exposure, as well as corresponding control groups. Here their goal is twofold: (1) they wish to determine the extent to which their models can be utilized on smaller, regional scales and (2) they would like to characterize as precisely as possible the age, setting, and type of nature exposure that leads to the highest value being placed on natural areas and conservation. These latter results would be extraordinarily useful for environmental educators, and are particularly cogent in light of the No Child Left Inside Act (H.R. 3036 and S. 1981; see also Pergams, 2008).

Over the last 7 years, the authors have developed a line of research on the relationships among natural systems, biodiversity conservation, and economic forces. Pergams *et al.* (2004) predicted revenues to major conservation NGOs and acreage of new federal protected lands based on corporate and individual incomes with a high degree of accuracy ($p<.0005$) and predicted the drop in conservation investment during the current economic crisis. Of further interest, one of the variables gathered (*per capita* NPV in the USA) was shown to be in decline since 1987, after a previous 50 years of steady increase.

Pergams and Zaradic (2006) tested possible proximate causes of this drop in nature recreation and found that a model consisting of four entertainment media usage variables (hours of video games, home movies, theatre attendance, and Internet use) as well as oil prices explained the decline (adjusted multiple $r^2=0.925$, $p<.0001$). The authors coined the word "videophilia" to describe this phenomenon.

The breadth of decline in nature exposure was revealed in Pergams and Zaradic (2008). Beyond US national parks, the authors examined 16 time series in nature visitation across state and federal public lands (both in the USA and abroad)

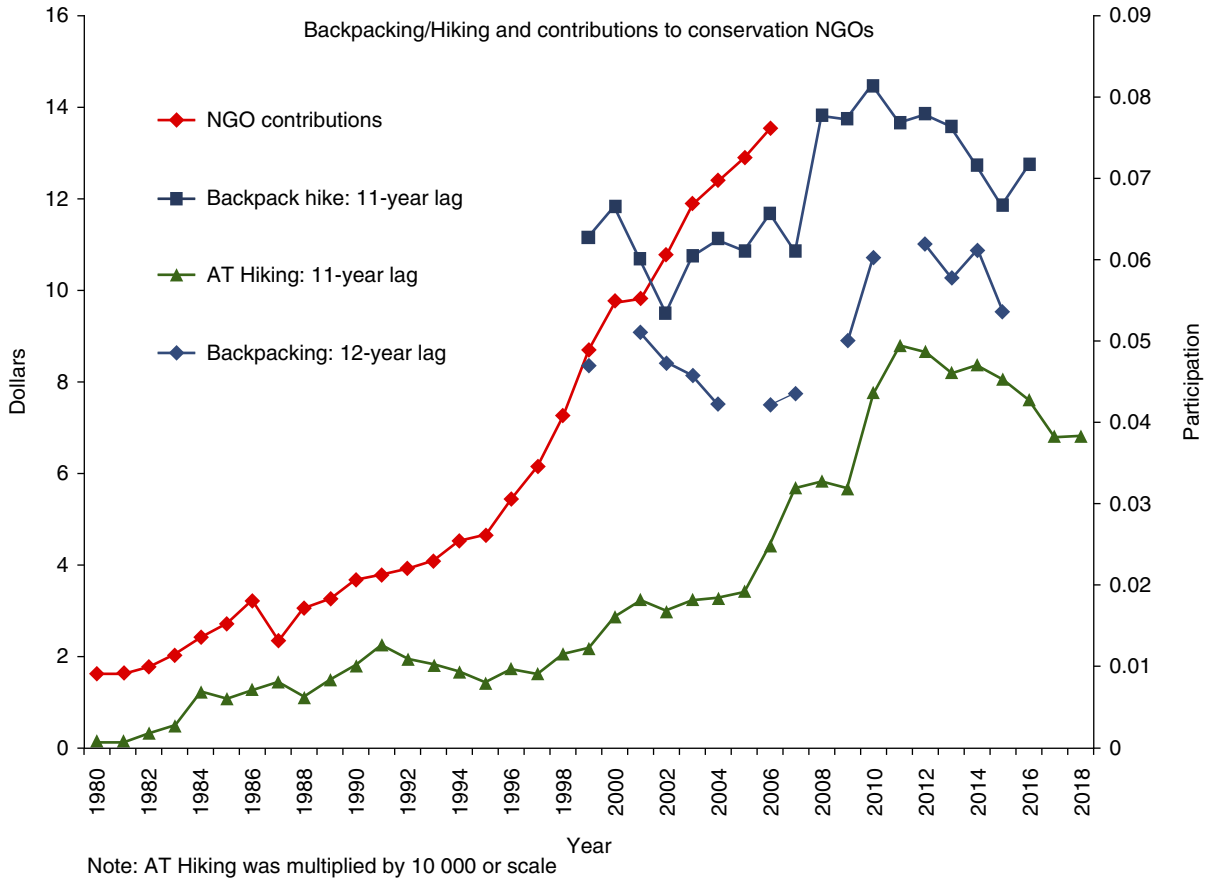


Figure 4 Comparing annual *per capita* contributions to conservation NGOs and participation in backpacking and hiking. Gaps in the backpacking time series are due to lack of data for those years. "NGO contributions" is aligned with actual year of contribution on the *x*-axis. The backpacking and hiking series are shown lagged by +11 or +12 years to match the significantly correlated year of NGO revenue. In other words, data for "NGO Contributions" for 2006 are depicted on the graph at *x*-axis year 2006 along with the correlated data for "Backpack Hike" from 1995, "AT hiking" from 1995, and "Backpacking" from 1994. Based on the correlation between these variables, a decline in NGO contributions would be projected until at least 2018.

as well as other nature-related recreation activities. Typical declines in nature recreation began between 1981 and 1991 and proceeded at rates of −1.0% to −1.3% per year, with a total to date of −18% to −25%. The drop in nature recreation is indeed generalized in the USA (and potentially in other developed nations such as Japan) and has been characterized as possibly the world's greatest environmental threat (Kareiva, 2008).

Zaradic *et al.*, 2009 suggests that the type and timing of nature experience may determine future conservation investment. Time spent hiking or backpacking is correlated with increased conservation contributions 11–12 years later. However, contributions are negatively correlated with past time spent on activities such as public lands visitation or fishing (**Figure 4**).

The authors' published findings are represented as solid arrows in the conceptual map below (**Figure 5**). The dotted arrows represent proposed research. The map itself depicts the relationships the authors perceive between independent variables (ovals) of (1) nature exposure, (2) electronic media usage, and (3) economic forces; and dependent variables (rectangles) of (1) biodiversity conservation effort and (2) quality of natural systems. The authors propose research to evaluate the following questions:

1. Nature Exposure → Biodiversity Conservation. Does nature exposure influence future environmental consciousness? Research in the social sciences (and common sense) suggests that nature values are highly influenced by direct contact with nature (Hungerford and Volk, 1990; Wells and Lekies, 2006), particularly in childhood (Duda *et al.*, 1998; Zaradic and Pergams, 2007) and with a role model (Matthews and Riley, 1995; Chawla, 1998). Zaradic *et al.*, 2009 evaluated whether various forms of nature recreation and exposure to nature are significantly associated with later conservation effort and quantify the degree of these relationships and the time lags between them. Experimental work could further validate and explore the relationship between timing and duration of specific types of nature recreation and later conservation ethics.
2. Natural Systems → Nature Exposure. There is a feedback loop between the quality and accessibility of natural systems and the amount of time people are willing to spend in these systems. The authors propose quantifying these relationships and the time lags involved.
3. Economic Forces → Natural Systems. A growing economy has positive effects on conservation effort (Pergams *et al.*, 2004). However, a growing economy is associated with increased consumption of natural resources; habitat degradation, loss, and fragmentation; pollution; climate change; etc. The authors propose the characterization of how such individual aspects of our natural systems are affected by individual economic forces.
4. Biodiversity Conservation → Natural Systems. Increased conservation effort no doubt results in improved aspects of natural systems. But which aspects of biodiversity

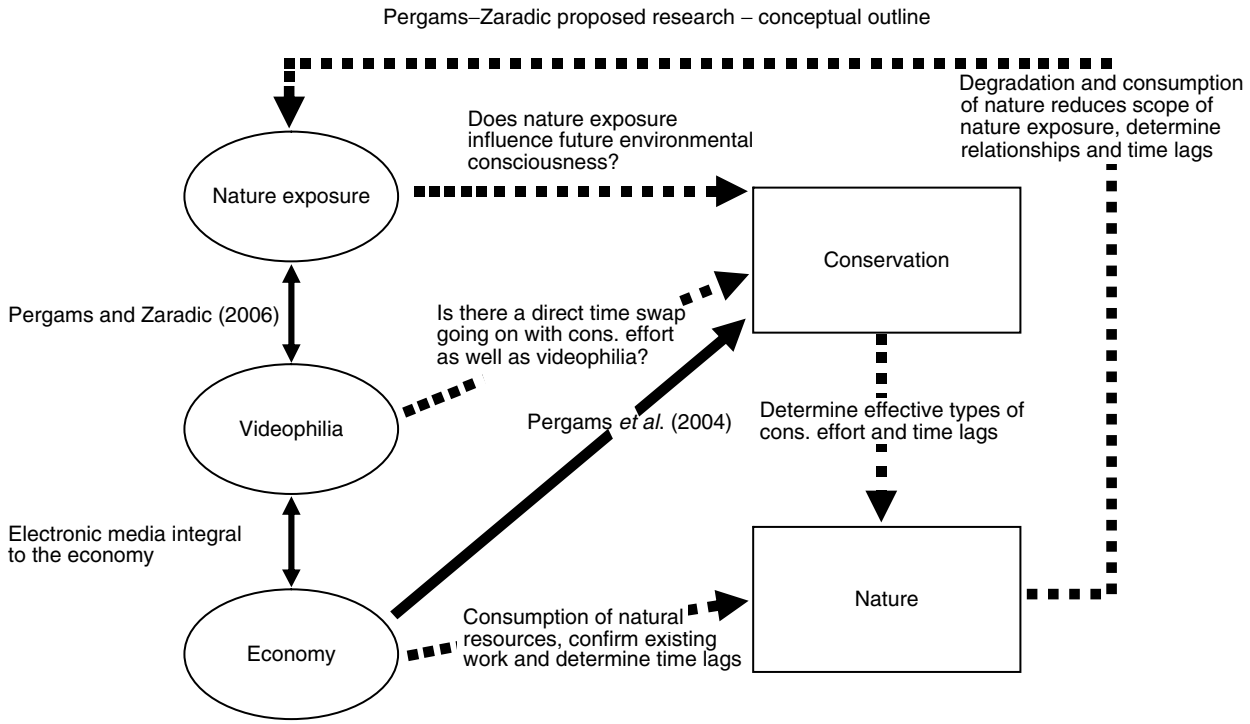


Figure 5 A conceptual map of current and suggested research on the connections between nature exposure, videophilia, and economics and their impact on conservation and the quality of natural systems. The map depicts the relationship between independent variables (ovals) of (1) nature exposure, (2) electronic media usage, and (3) economic forces; and dependent variables (rectangles) of (1) biodiversity conservation effort and (2) quality of natural systems. The authors' published findings are represented as solid arrows in the conceptual map. The dotted arrows represent proposed research.

conservation result are most efficient? How long does it take, on average, for these improvements to have effect?

5. Videophilia→Biodiversity Conservation. The authors have shown the likelihood of a time swap between use of electronic media and nature exposure (Pergams and Zaradic, 2006). They would like to evaluate whether there is also a direct time swap between use of electronic media and effort put into conservation, what types of conservation effort are affected, and what the time lags are.

Appendix

List of Courses

1. Conservation Biology
2. Environmental Biology
3. Ecology
4. Sustainability
5. Introductory Biology

See also: Ecosystem Services. Sustainability and Biodiversity

References

Babcock M (2007) Fewer hunters: Ducking trend with clinics. *Great Falls Tribune Online*. Published 6 September 2007 and accessed 21 September 2007 at http://www.greatfallstribune.com/apps/pbcs.dll/article?AID=/20070906/LIFESTYLE05/709060321.

Bailey TR, Hughes KL, and Moore DT (2004) *Working Knowledge: Work-Based Learning and Education Reform*. New York: Routledge Falmer.

Bartlett SN (1997) No place to play: Implications for the interaction of parents and children. *Journal of Children and Poverty* 3: 37–48.

Bixler RD and Carlisle CL (1994) Observed fears and discomforts among urban students on field trips to wildland areas. *Journal of Environmental Education* 26: 24–33.

Bögeholz S (2006) Nature experience and its importance for environmental knowledge, values and action: Recent German empirical contributions. *Environmental Education Research* 12: 65–84.

Butler RA (2009) *Economic Crisis Hits Conservation but may Offer Opportunities, Says TNC President*. March 03, Available at: http://news.mongabay.com/2009/0303-tercek_interview.html. Accessed 21 April 2009.

Chawla L (1998) Significant life experiences revisited: A review of research on sources of environmental sensitivity. *Journal of Environmental Education* 29: 11–21.

Chawla L (1999) Life paths into effective environmental action. *Journal of Environmental Education* 31: 15–26.

Christakis DA, Zimmerman FJ, DiGiuseppe DL, and McCarty CA (2004) Early television exposure and subsequent attentional problems in children. *Pediatrics* 113: 708–713.

Cordell HK, Green GT, Leeworthy VR, Stephens R, Fly MJ, *et al.* (2005) United States of America: Outdoor recreation. In: Cushman G, Veal AJ, and Zuzanek J (eds.) *Free Time and Leisure Partcipation: International Perspectives*. Wallingford, UK: CAB International.

Deckelbaum RJ and Williams CL (2001) Childhood obesity: The health issue. *Obesity Research* 9: S239–S243.

Dong L, Block G, and Mandel S (2004) Activities contributing to total energy expenditure in the United States: Results from the NHAPS Study. *International Journal of Behavioral Nutrition and Physical Activity* 1: 1–11.

Ducks Unlimited (2007) *Duck Numbers up Slightly Overall*. Webpage published on 11 July 2007 and accessed on 14 June 2011 at http://www.ducks.org/news/1305/Ducknumbersupslightl.html.

Duda MD, Bissell SJ, and Young KC (1998) *Wildlife and the American Mind: Public Opinionon and Attitudes Toward Fish and Wildlife Management*. Harrisonburg, VA: Responsive Management.

Duda MD, De Michele PE, Zurawski C, *et al.* (2002) *Factors Relating to Hunting and Fishing Participation Among the Nation's Youth*. Harrisonburg, VA: Responsive Management.

Earthjustice (2005) Roadmap Toward a New Ecological Majority. *American Environics*. Available at: http://www.americanenvironics.com/PDF/Road_Map_for_Ecological_Majority_AE.pdf. Accessed 15 December 2008.

Echo Media (2009) *Market Reports for a) Field and Stream, b) Backpacker Magazine*. Available at: (a) http://www.echo-media.com/MediaDetail.asp?IDNumber=7174. (b) http://www.echo-media.com/MediaDetail.aspIDNumber=4376. Accessed 24 July 2009.

Finkelhor D, Hootaling G, and Sedlak A (1992) The abduction of children by strangers and nonfamily members: Estimating the incidence using multiple methods. *Journal of Interpersonal Violence* 7: 226–243.

Fisman L (2005) The effects of local learning on environmental awareness in children: An empirical investigation. *The Journal of Environmental Education* 36: 39–50.

Flegal KM, Carroll MD, Ogden CL, and Johnson CL (2002) Prevalence and Trends in Obesity Among US Adults, 1999–2000. *Journal of the American Medical Association* 288: 1723–1727.

Fotheringham MJ, Wonnacott RL, and Owen N (2000) Computer use and physical inactivity in young adults: Public health perils and potentials of new information technologies. *Annals of Behavioral Medicine* 22: 269–275.

Grahn P, Mårtensson F, Lindblad B, Nilsson P, and Ekman A (1997) *Ute på dagis. Stad and Land, Nr. 145 [Outdoor Daycare. City and country]*. Hässleholm, Sverige: Norra Skåne Offset.

Greenough WT, Black JE, and Wallace CS (1987) Experience and brain development. *Child Development* 58: 539–559.

Heerwagen JH and Orians GH (2002) The ecological world of children. In: Kahn Jr. PH and Keller SR (eds.) *Children and Nature: Psychological, Sociocultural, and Evolutionary Investigations*, pp. 29–64. Massachusetts Institute of Technology.

Hess M and SA Brill (2009) Video Consumer Mapping Study: Key Finding Report. *Report prepared for the Center for Research Excellence*. Accessed 15 June 2011 at http://researchexcellence.com/committees/vcm_finalreport.pdf.

Hofferth S and Sandberg J (2001) Changes in American Children's Time, 1981–1997. In: Hofferth SL and Owens TJ (eds.) *Children at the Millemmium: Where Have We Come From, Where Are We Going?* pp. 193–232. Oxford, England: Elsevier Science.

Hungerford H and Volk T (1990) Changing learner behaviour through environmental education. *Journal of Environmental Education* 21: 8–21.

Hüttenmoser M (1995) Children and their living surroundings: Empirical investigations into the significance of living surroundings for the everyday life and development of children. *Children's Environments* 12: 403–413.

Kahn Jr. PH (1997) Developmental psychology and the biophilia hypothesis: Children's affiliation with nature. *Developmental Review* 17: 1–61.

Kahn Jr. PH (2002) Children's affiliations with nature: Structure, development, and the problem and environmental generational amnesia. In: Kahn Jr. PH and Keller SR (eds.) *Children and Nature: Psychological, Sociocultural, and Evolutionary Investigations*, pp. 93–116. Cambridge, MA: Massachusetts Institute of Technology.

Kareiva P (2008) Ominous trends in nature recreation. *Proceedings of the National Academy of Sciences United States of America* 105: 2757–2758.

Kellert S (1996) *The Value of Life: Biological Diversity and Human Society*. Washington, District of Columbia: Island Press.

Kellert SR (2002) Experiencing nature: Affective, cognitive, and evaluative development in children. In: Kahn Jr. PH and Kellert SR (eds.) *Children and Nature: Psychological, Sociocultural, and Evolutionary Investigations*, pp. 117–152. Cambridge, MA: Massachusetts Institute of Technology.

Kellert SR and Wilson EO (1993) *The Biophilia Hypothesis*. Washington, DC: Island Press.

Kraut R, Lundmark V, Patterson M, Kiesler S, Mukopadhyay T, and Scherlis W (1998) Internet paradox: A social technology that reduces social involvement and well-being? *American Psychologist* 53: 1017–1031.

Lang S (2006) Camping, hiking and fishing in the wild as a child breeds respect for environment in adults, study finds. *ChronicleOnLine*. 13 March 2006.

Levi D and Kocher S (1999) Virtual nature: The future effects of information technology on our relationship to nature. *Environment and Behavior* 31: 203–226.

Loomis J (2005) *Updated Outdoor Recreation Use Values on National Forests and Other Public Lands*, 26 pp. General Technical Report PNW-GTR-658. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research

Station. Available at: ftp://ftp-fc.sc.egov.usda.gov/Economics/recreate/pnw_gtr658.pdf Accessed 24 July 2009.
Louv R (2006) *Last Child in the Woods: Saving Our Children From Nature-Deficit Disorder*. New York: Alogonquin.
Manni MF, Littlejohn M, Evans J, Gramann J, and Hollenhorst SJ (2007) Yellowstone National Park Visitor Study, Summer 2006. *Visitor Services Project Report 178, Park Studies Unit*, pp. 102. Moscow, ID: University of Idaho. Available at: http://psu.uidaho.edu/files/vsp/reports/178_YELL_rept.pdf?PHPSESSID=46fdea30ff23d9b103260058713213cb. Accessed 24 July 2009.
Matthews BE and Riley CK (1995) *Teaching and Evaluating Outdoor Ethics Education Programs*. Vienna, Virginia: National Wildlife Federation (ERIC Document Reproduction Service No ED 401 097).
National Park Service (NPS) (2001) *The National Park Service comprehensive survey of the American Public*. Available at: http://www.nature.nps.gov/socialscience/pdf/NatSurvTechRep.pdf Accessed 21 April, 2009.
National Park Service (2006) *The Official Website of Yellowstone National Park - The Old Faithful WebCam*. Available at: http://www.nps.gov/archive/yell/oldfaithfulcam.htm. Accessed 23 October 2006.
Nie NH and Hillygus DS (2002a) The impact of Internet use on sociability: Time-diary findings. *IT and Society* 1: 1–20. Available at: http://www.itandsociety.org.
Nie NH and Hillygus DS (2002b) Where does Internet time come from? A reconnaissance. *IT and Society* 2: 1–20. Available at: http://www.itandsociety.org.
Nordhaus T and Shellenberger M (2007) *Break Through: From the Death of Environmentalism to the Politics of Possibility*. New York: Houghton Mifflin.
Outdoor Foundation (2010) *Outdoor Recreation Participation Report 2010. The Outdoor Foundation*. Available at: http://www.outdoorfoundation.org/research.participation.2010.html
Pergams ORW (2008) Written testimony of Dr. Oliver R. W. Pergams before the Subcommittee on Early Childhood, Elementary, and Secondary Education, US House of Representatives Committee on Education and Labor, on April 22, 2008. Available at http://www.videophilia.org/uploads/PergamsTestimonyAbd.pdf. Accessed 15 June 2011.
Pergams ORW, Czech B, Haney JC, and Nyberg D (2004) Linkage of conservation activity to trends in the U.S. economy. *Conservation Biology* 18: 1617–1623.
Pergams ORW and Zaradic PA (2006) Is love of nature in the US becoming love of electronic media? 16-year downtrend in National Park visits explained by watching movies, playing video games, Internet use, and oil prices. *Journal of Environmental Management* 80: 387–393.
Pergams ORW and Zaradic PA (2008) Evidence for a fundamental and pervasive shift away from nature-based recreation. *Proceedings of the National Academy of Sciences, United States of America* 105: 2295–2300.
Pyle RM (2002) Eden in a vacant lot: Special places, species and kids in the neighborhood of life. In: Kahn Jr. PH and Keller SR (eds.) *Children and Nature: Psychological, Sociocultural, and Evolutionary Investigations*, pp. 305–328. Cambridge, MA: Massachusetts Institute of Technology.
Rickinson M, Dillon J, Teamey K, *et al*. (2004) A Review of Research on Outdoor Learning, *(Executive Summary)*. Shrewsbury, UK: National Foundation for Educational Research.
Rideout VJ, Vandewater EA, and Wartella EA (2003) *Zero to Six: Electronic Media in the Lives of Infants, Toddlers, and Preschoolers*, Publication no. 3378. Menlo Park, CA: Kaiser Family Foundation.
Sebba R (1991) The landscapes of childhood: The reflection of childhood's environment in adult memories and in children's attitudes. *Environment and Behavior* 232: 395–422.
Shellenberger M and Nordhaus T (2004) *The Death of Environmentalism: Global Warming Politics in a Post Environmental World*. San Francisco: Shellenberger and Nordhaus.
Skidmore PML and Yarnell JWG (2004) The obesity epidemic: Prospects for prevention. *QJM: An International Journal of Medicine* 97: 817–825.
South Carolina Department of Parks, Recreation and Tourism (DPRT) (2001) *National and State Parks: South Carolina 2001 Tourism Report Series*. Available at: http://www.scprt.com/files/Research/National_and_State_Parks.htm. Accessed 24 July 2009.
State of Virginia (2007) *Outdoor Recreation Issues, Trends and Survey Findings. Technical Report*. Available at: http://www.dcr.virginia.gov/recreational_planning/documents/vopchapt02.pdf. Accessed 10 June 2008.
Stephens S (2007) *Director of Conservation Planning, Great Plains Regional Office*. Bismarck, ND: Ducks Unlimited.
Stromayer KAK and Warren RJ (1997) Are overabundant deer herds in the Eastern United States creating alternate stable states in forest plant communities? *Wildlife Society Bulletin* 25: 227–234.
Taylor AF, Kuo FE, and Sullivan WC (2001) Coping with ADD: The surprising connection to green play settings. *Environment and Behavior* 33: 54–77.
Taylor AF, Kuo FE, and Sullivan WC (2002) Views of nature and self-discipline: Evidence from inner city children. *Journal of Environmental Psychology* 22: 49–63.
Turner AM and Greenough WT (1985) Differential rearing effects on rat visual cortex synapses. I. Synaptic and neuronal density and synapses per neuron. *Brain Research* 329: 195–203.
US Census Bureau (2010) *Statistical Abstract of the United States: 2010*, 129th edn. Washington, DC. http://www.census.gov/statab/www/. Tables: 603, 701, 1094, 1205.
US Department of Education (USDE) (2005) Rates of computer and Internet use by children in nursery school and students in kindergarten through twelfth grade: 2003. *Issue Brief Series from the National Center for Educational Statistics*. No. 2005-111.
United Nations (2007) *Fighting Climate Change: Human Solidarity in a Divided World. Human Development Report*. UNDP. Web: hdr.undp.org
United Nations (2009) *Overcoming Barriers, Human Mobility and Development. Human Development Report*. UNDP. Web: hdr.undp.org
UNPD (2005) *World Urbanization Prospects: The 2005 Revision*. New York: United Nations Population Division.
Utter J, Neumark-Sztainer D, Jeffery R, and Story M (2003) Couch potatoes or French fries: Are sedentary behaviors associated with body mass index, physical activity, and dietary behaviors among adolescents? *Journal of the American Dietetic Association* 103: 298–1305.
Vaske JJ and Kobrin KC (2001) Place attachment and environmentally responsible behavior. *The Journal of Environmental Education* 32: 16–21.
Verbeek JH and de Waal FBM (2002) The primate relationship with nature: Biophilia as a general pattern. In: Kahn Jr. PH and Keller SR (eds.) *Children and Nature: Psychological, Sociocultural, and Evolutionary Investigations*, pp. 1–28. Cambridge, MA: Massachusetts Institute of Technology.
Wallace CS, Kilman VL, Withers GS, and Greenough WT (1992) Increases in dendritic length in occipital cortex after 4 days of differential housing in weanling rats. *Behavioral Neural Biology* 58: 64–68.
Wells NM (2000) At home with nature: The effects of nearby nature on children's cognitive functioning. *Environment and Behavior* 32: 775–795.
Wells NM and Evans GW (2003) Nearby nature: A buffer of life stress among rural children. *Environment and Behavior* 3: 311–330.
Wells NM and Lekies KS (2006) Nature and the life course: Pathways from childhood nature experiences to adult environmentalism. *Children, Youth and Environments* 16: 1–24.
WHO (1997) Health and environment in sustainable development. *Five years after the summit*. Geneva, Switzerland: World Health Organization (WHO/EHG/97.8).
WHO (2000) Considerations in evaluating the cost-effectiveness of environmental health interventions. *Protection of the Human Environment*. Geneva: World Health Organization. WHO/SDE/WSH/00.10.
WHO (2008) *The top ten causes of death by broad income group, 2004*. World Health Organization. Fact Sheet #310. http://www.who.int/mediacentre/factsheets/fs310/en/index.html.
Wilson EO (1984) *Biophilia*. Cambridge, MA: Harvard University Press.
Zaradic PA and Pergams ORW (2007) Videophilia: Implications for childhood development and conservation. *Journal of Developmental Processes* 2: 130–144.
Zaradic PA and Pergams ORW (2011) *The Nature Conservancy's Long-Term LEAF Program Evaluation*. Red Rock Institute. Accessed 2011 June 14 http://www.nature.org/aboutus/diversity/leaf/prd_033183.
Zaradic PA, Pergams ORW, and Kareiva P (2009) The impact of nature experience on willingness to support conservation. *PLoS One* 4: 1–5. http://dx.plos.org/10.1371/journal.pone.0007367.

Trophic Cascades

Michael L Pace, University of Virginia, Charlottesville, VA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Apex predators Predators that have no predators of their own and are at the top of the food web.
Ecosystem services Values to humans provided by ecosystems.
Indirect effects Effects on a third species or trophic group owing to a direct interaction between two species or trophic groups. Trophic cascades are the result of the indirect effect of predator–prey interactions on a third group.
Mesopredators Predators in the middle of the food web, usually small- to medium-sized animals that are preyed on by other predators and that often benefit when apex predators decline.
Meta-analysis Method of summarizing outcomes across numerous studies by compiling measures of the effect size of treatments as, for example, measures of the differences in biomass of primary producers in the absence versus presence of predators.
Natural experiments When a change occurs in a system due to natural or anthropogenic forces, and comparisons can be made with conditions before the change and with conditions in an environment where the change did not occur. For example, the recovery of a predator may provide a natural experiment in which conditions before and after recovery and conditions inside and outside the recovery area are compared.
Omnivory A consumer that feeds on many types of prey including, for example, a diet of plant and animal sources.
Regime shifts Dramatic changes in an ecosystem from one distinct state to another distinct state. Regime shifts are abrupt and often difficult to reverse.
Resource facilitation Activity of one species promoting the availability of resources for another species.
Trophic cascades The effects of predatory feeding on prey that propagate down more than one link in a food web. Trophic cascades cause inverse patterns in the abundance or biomass of trophic groups at successive levels of a food web.
Trophic downgrading Ecological and environmental consequences associated with losses of apex predators from ecosystems.

Introduction: The Far-Reaching Effects of Predation

When predators reduce prey populations, the effect can propagate, causing an increase in prey lower in the food web. This dynamic interaction is a trophic cascade. Trophic cascades cause inverse abundance or biomass patterns across more than one trophic link in a food web (Pace *et al.*, 1999). Consider a two-level food chain of primary producers and herbivores. A trophic cascade would occur if a predator were introduced that reduced the biomass of herbivores, resulting in an increased biomass of primary producers due to eased grazing (**Figure 1**). However, not all predator–prey interactions propagate down food webs via trophic cascades. Extensive omnivory and multiple predators and prey embedded in complex food webs blunt trophic cascades (Polis and Strong, 1996). In addition, there are many behavioral, morphological, and chemical defenses that reduce the impacts of predators. Despite this complexity, trophic cascades are observed in terrestrial, marine, and freshwater systems and may cause transitory or permanent shifts in biomass of trophic groups within ecosystems.

The earliest studies of ecosystems provided a "bottom-up" perspective on the control of productivity. Physical–chemical forcing factors such as sunlight and nutrient availability controlled photosynthesis, and the resulting primary production determined production of secondary consumers via transfers of energy up the food web. Hairston, Smith, and Slobodkin (1960), in a famous paper, challenged this perspective by arguing that predators limited herbivores, reducing herbivory, with the consequence of a "green" world of lightly grazed plants. Although the Hairston, Smith, and Slobodkin argument was controversial, their paper was among the first descriptions of predatory "top-down" regulation. Paine (1980) first described the propagating effects of predators in the marine intertidal zone as "cascading alterations in structure."

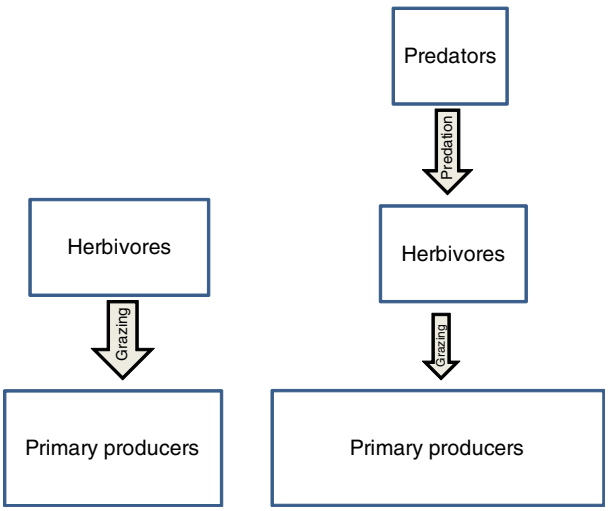


Figure 1 Two simple food chains depicting the impact of a trophic cascade. In the two-level food chain, herbivore grazing limits primary producers. Adding a predator to create a three-level food chain causes a decline in herbivores, which leads to reduced grazing and an increase in primary producers.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00396-8

Terrestrial researchers noted inverse patterns of biomass in some food chains of plants, rodents, and predators, especially those in high northern latitudes (e.g., Oksanen *et al.*, 1981). In freshwater systems, researchers observed that shifts in top fish predators caused strong cascading changes in prey fishes as well as in invertebrates, with consequences for primary producers (Hrbáček *et al.*, 1961; Power, 1990). Carpenter *et al.* (1985), drawing on examples from lakes, popularized the concept of trophic cascades and the impacts of apex predators that could transfer all the way down the food web.

Subsequent research has focused on the occurrence and relative importance of trophic cascades in a variety of ecosystems, with a particular interest in whether cascades influence primary producers. The term "trophic cascade" applies broadly across population, community, and ecosystem levels of organization. Hence, a species might be influenced by a trophic cascade with consequent effects on its abundance. Because of elk grazing, aspen trees (*Populus tremuloides*) were rare when wolves (*Canis lupus*) were absent from Yellowstone National Park (USA) (Ripple *et al.*, 2010). When wolves were reintroduced, aspen increased to abundances similar to levels observed before the extirpation of wolves. The distribution and abundance of the aspen population are significantly driven by a trophic cascade. The term "trophic cascade" is also used in the context of communities. Trophic cascades may affect feeding groups or entire trophic levels. Some researchers have advocated labeling this pattern of top-down regulation as "community cascades," restricting the use of the term "trophic cascades" to specific cases in which the impact manifests as changes in primary production (Polis *et al.*, 2000). In this view, cascades are important only if they affect how "green" (i.e., productive) or "brown" (i.e., unproductive), in the case of terrestrial vegetation, an ecosystem appears (Polis *et al.*, 2000). In practice, the term carries a broader if more diffuse application.

Trophic cascades are by definition dynamic. The outcomes of trophic cascades may result in a relatively permanent change in the structure of food webs and lead to distinct differences in the biomass of trophic levels, as indicated in **Figure 1**. Ecosystems are complex, however, with temporal and spatial heterogeneity in population processes and trophic interactions. Trophic cascades may induce compensatory responses and result in variability in system components rather than permanent change. Consider a lake food web with four trophic levels: piscivorous fish, planktivorous fish and invertebrates, zooplankton herbivores, and phytoplankton. Variable recruitment of the top predators may lead to enhanced periods of predation pressure by planktivores on zooplankton. This occurs when dominant, older age classes of the piscivores die off, and planktivores (including young, planktivorous-feeding life stages of the piscivore) recruit in a pulse. These changes create a period of strong planktivory with cascading consequences that ameliorate as the piscivore population re-establishes (Post *et al.*, 1997). Thus, trophic cascades are one of many forces leading to variability in food webs. This dynamical aspect of trophic cascades is probably widespread and counters the simple view that cascades are significant regulatory forces only when effects cause community-level alterations in biomass or productivity that propagate to primary producers.

Trophic Cascade Examples: Wet and Dry

Trophic cascades occur in terrestrial and aquatic ecosystems. The earliest examples of trophic cascades were found in coastal water and freshwater systems. Strong (1992) argued that cascades were "all wet" – meaning confined largely to aquatic systems – because of greater potential in aquatic ecosystems for "runaway" consumption of primary producers by herbivores. In contrast, terrestrial plants resist excess consumption through a variety of defenses. However, trophic cascades occur in a spectrum of ecosystems including many terrestrial environments (Pace *et al.*, 1999; Estes *et al.*, 2011). Further, many aquatic primary producers are also defended and resist herbivory. At this stage (*see* Trophic Cascades: When, Where, How Much?) it is not clear whether trophic cascades are more common in aquatic systems than in terrestrial ones, but they are certainly not "all wet."

Some of the best evidence for the importance of trophic cascades comes from the recovery of apex predators. These recoveries have provided researchers with the opportunity to study and understand the impacts of cascades, as in the following two examples.

Sea otters (*Enhydra lutris*) drive a well-understood trophic cascade in the coastal waters of northwestern North America. Urchins (echinoderms) eat kelp, but otters eat urchins, and through predatory control of urchin populations, kelp forests can flourish. This cascade was first discovered when otters were protected from hunting, allowing their populations to recover from near-extinction (Estes and Palmisano, 1974). Subsequent site-specific reductions of otters due to disease or predation have led to a reversal of otter control on urchins. When otters decline to low levels, urchin populations increase unchecked and drive the loss of kelp beds, creating barren bottom areas. As a consequence, these coastal ecosystems exhibit alternate states of kelp forest and urchin barrens due to effects that emanate from otter predation. Regime shifts between alternate states as evidenced in the otter–urchin–kelp interaction are a significant outcome of changes in trophic cascades.

Like otters, wolves were driven to near-extinction throughout much of their range due to overhunting, but under protective laws, wolves are recovering in many areas. The return of wolves to the Banff region of the Canadian Rocky Mountains provides evidence of trophic cascades and other indirect effects of wolf predation (Hebblewhite *et al.*, 2005). Where wolves re-established, the population density, female survival, and recruitment of elk were all negatively impacted. In the absence of wolves, aspen and willow browsing by elk was greater and recruitment of these trees was near zero (**Figure 2**). This effect on tree recruitment is the same as described for Yellowstone Park. Interestingly, the higher elk populations in areas of Banff not used by wolves were also associated with lower populations of beavers and song birds due to reductions in riparian vegetation caused by elk grazing (**Figure 2**). This is a case of an indirect effect of predators on other animal populations through a trophic cascade that changes habitat suitability and biodiversity.

The wolf and sea otter recoveries also illustrate the difficulty of studying trophic cascades that originate from top predators. Natural experiments as opposed to controlled manipulations provide the evidence supporting conclusions

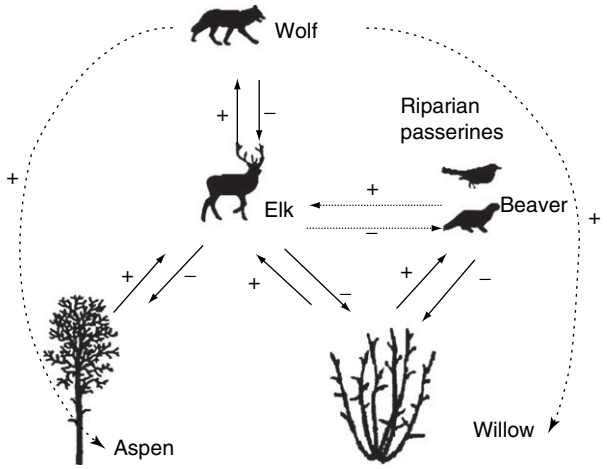


Figure 2 Trophic interactions for the Banff National Park area of Canada depicting the trophic cascade of wolves–elk–aspen/willow. Solid lines are direct predator–prey interactions, with +/− indicating the direction of the effect. Dashed lines represent indirect effects; for example, wolves have a positive indirect effect on aspen and willows, whereas elk have negative indirect effects on birds and beavers. Reproduced from Hebblewhite M, White CA, Nielvelt CG, *et al.* (2005) Human activity mediates a trophic cascade caused by wolves. *Ecology* 86: 2135–2144, with permission from ESA.

about trophic cascades in these cases. Research to establish these impacts was carried out over large areas, took many years, and required comparison among systems. To encompass and evaluate large predators with extensive ranges, researchers must often consider large spatial areas and long time scales. Many predators are also difficult to manipulate and do not fit within typical experimental units (e.g., cage enclosures). The study and understanding of predator impacts has lagged partly because of these difficulties and partly because top predators are absent and, unlike otters and wolves, have not recovered in their native ecosystems.

Shifts caused by predator introductions do not always result in a return to a previous state as in the wolf and otter cases. Novel predators may cause trophic cascades that move an ecosystem far from the "native" condition. On Christmas Island south of Java, red land crab (*Gecarcoidea natalis*) was an extremely abundant herbivore that suppressed forest understory plants. Explosive population growth of an invasive ant species (*Anoplolepis gracilipes*) wiped out the crabs over much of the island, allowing plant recruitment and the development of a lush forest understory (O'Dowd *et al.*, 2003). In locales where crabs remain abundant the forest floor is open with few saplings and little leaf litter, but where the island is dominated by ants, there is much higher plant biomass and richer plant diversity. Ants also altered the canopy by developing associations with scale insects collectively promoting the growth of molds, leading to canopy dieback. The ecosystem under ant dominance is now completely different (O'Dowd *et al.*, 2003).

Trophic Cascades: When, Where, How Much?

Are trophic cascades an ecological curiosity or a phenomenon of widespread importance, common in many systems? When and where do trophic cascades occur, and what factors are related to the strength of trophic cascades? These questions address the generality and significance of trophic cascades. Ecologists have posed many ideas about characteristics of food webs, diversity of producers and consumers, metabolic attributes, key traits of consumers, and other features that could influence the propensity and strength of cascades. One way to address these questions is by synthesizing the results of many studies through the techniques of meta-analysis. A common metric for assessing trophic cascades is the biomass of autotrophs (primary producers) with and without top predators. The log of the ratio of autotrophic biomass in the presence to absence of predators was compared for over 100 studies that included a wide diversity of methodologies and ecological characteristics (Shurin *et al.*, 2002; Borer *et al.*, 2005). Trophic cascades as measured by the log ratio are stronger in some types of systems (e.g., the marine benthos) and the strength of cascades is related to the metabolic efficiency of herbivores and predators. For example, trophic cascades are stronger when herbivores have a high metabolic efficiency (e.g., ectothermic organisms), allowing enhanced conversion of the autotrophic food to growth. Many other factors such as diversity of consumers or producers within a trophic group or rates of primary production had little influence when considered across a large suite of studies (Borer *et al.*, 2005). These types of synthetic studies also indicate limitations in the data available for meta-analysis. For example, there are relatively few studies from terrestrial systems and even fewer studies that have manipulated terrestrial endothermic predators (Borer *et al.*, 2005).

The generalities drawn from meta-analyses need cautious consideration because, in part, results are analyzed for studies that use a variety of methodologies and experimental scales (Whittaker, 2010). In lakes, for example, the patterns of trophic cascades across gradients of nutrient enrichment and primary production differ from the results of meta-analysis based on many different types of systems. Whole lake manipulations done with consistent methods over the same temporal scale illustrate that trophic cascades are stronger when nutrient loading (additions of nitrogen and phosphorus) and hence potential primary production is greater (Carpenter *et al.*, 2001). Lakes responded to increased nutrient loading with large increases in phytoplankton biomass when plantivores and small grazers were the dominant food web constituents (**Figure 3**). In contrast, in lakes dominated by piscivores and large zooplankton grazers, there was a limited increase in phytoplankton primary production and biomass across a range of nutrient loads (**Figure 3**). This result contrasts with the lack of any effect of primary production on the strength of trophic cascades found in meta-analyses (Borer *et al.*, 2005). Stronger trophic cascades in more productive ecosystems may be particular to lakes, or alternatively, data across many types of ecosystems may be too limited to have strong confidence in the results of meta-analyses.

The ecological features that promote or inhibit cascades remain an open question because: (1) methods in ecological research are variable among studies, making comparisons difficult; (2) trophic cascade studies typically require large spatial and temporal scales and studies at these scales are less common; and (3) the number of case studies is limited,

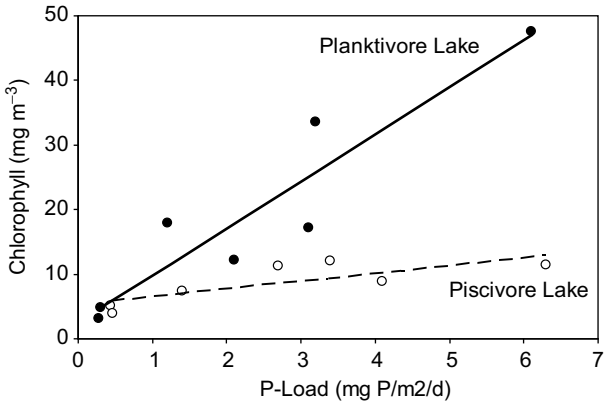


Figure 3 Trophic cascades in lakes affect the response of phytoplankton biomass measured as chlorophyll to increased nutrient loading measured as phosphorus input. Each point represents a different year for the two lakes – one with planktivorous fish as top predators and one with piscivorous fish as top predators. In the planktivore lake, herbivores were smaller, less effective grazers and phytoplankton biomass was greater at the given nutrient load. In the piscivore lake, herbivores were larger, more effective grazers and phytoplankton biomass was lower. Reproduced with permission from Terborgh J and Estes JA (eds.) (2010) *Trophic Cascades: Predators, Prey and the Changing Dynamics of Nature*. Washington, DC: Island Press.

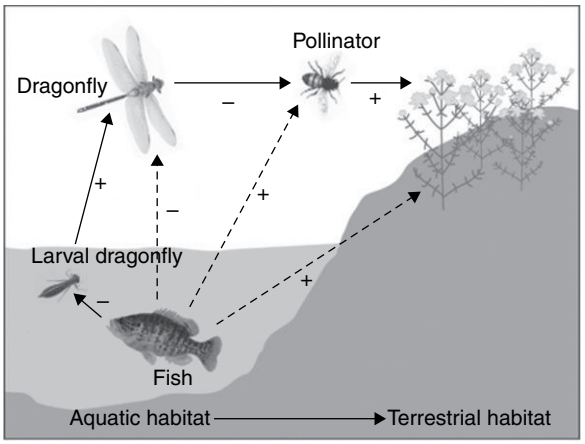


Figure 4 A crosshabitat trophic cascade of fish–dragonflies–insects that affects the pollination of riparian plants. Solid lines indicate the direct effects of trophic interactions. Dashed lines indicate the indirect effects of fish. Modified from Knight TM, McCoy MW, Chase JM, McCoy KA, and Holt RD (2005) Trophic cascades across ecosystems. *Nature* 437, 880–884, with permission from nature Publishing Group.

especially those that consider trophic cascades under varying ecological conditions, such as differences in nutrient loading. Ecologists can still profitably argue about whether trophic cascades are more likely in terrestrial versus aquatic systems, whether predator-induced switches to efficient herbivores are critical for driving cascades to primary producers, and how the characteristics and diversity of plant communities influence cascades. Simple answers may not emerge but instead may depend on the organisms and the context of particular ecosystems. Nevertheless, a recent volume that includes many case studies points to the importance of predators and cascading trophic interactions in ecosystems throughout the biosphere (Terborgh and Estes, 2010).

Consequences of Trophic Cascades

The cascading impacts of predators are not just about whether ecosystems have more or less primary production but include many other effects. Here a distinction is needed between direct predatory effects and the indirect effects inherent in trophic cascades. Recall that, by definition, beyond the direct effect of predator on prey, trophic cascades require an indirect effect at lower levels of the food web. Other indirect consequences result from the shifts in biomass and organism abundances induced by trophic cascades. These include consequences for processes related to biogeochemical cycling, crosshabitat linkages, and ecosystem services. Indeed, given the impacts of predators on phenomena such as fire frequency and disease (Estes *et al.*, 2011), it is likely that there are many other collateral consequences of trophic cascades yet to be discovered.

Biogeochemical processes are largely driven by the unique metabolic capabilities of microorganisms in concert with primary producers that utilize nutrients and generate the organic matter fueling elemental cycles. Trophic cascades driven by predatory animals can determine patterns of nutrient recycling and even the net direction of biogeochemical fluxes. For example, the presence or absence of piscivorous, largemouth bass (*Micropterus salmoides*) in experimental lakes fertilized with nutrients was the key to whether lakes were net sources or sinks of CO_2 (Schindler *et al.*, 1997). How did bass influence a physical–chemical process such as the flux of a gas between the atmosphere and a lake? In the absence of bass, planktivorous fish dominated, suppressing large-bodied zooplankton grazers and promoting greater phytoplankton production and hence uptake of CO_2. Under these conditions, lakes became undersaturated in CO_2 with a resultant movement of CO_2 from the atmosphere to the lake. In other experimental lakes with bass, zooplankton exerted greater control of phytoplankton production. Phytoplankton took up less CO_2 and in combination with the microbial respiration of terrestrial organic matter more CO_2 was produced than consumed. The bass-dominated lakes were net sources of CO_2 to the atmosphere.

Trophic cascades can spill across habitat boundaries via the movement, life histories, and predatory impacts of organisms. Dragonflies deposit their eggs in water and the larvae grow in the aquatic system, where they are a favored prey item of fish. Knight *et al.* (2005) compared ponds with and without fish and found many more dragonfly larvae in fishless ponds. This predator–prey interaction caused a trophic cascade that affected the pollination of riparian plants (**Figure 4**). Around the fishless ponds, dragonfly adults were more abundant, they preyed on pollinator insects, and as a consequence, there were fewer pollinator visits to plants such as St John's wort (*Hypericum fasciculatum*). The opposite conditions prevailed in and around ponds with fish – dragonfly larvae and adults were fewer, flies and bees were more abundant, and pollinator visits to *Hypericum* plants were more frequent. The trophic cascade

in this case was crosshabitat with pollination influenced by dragonflies as both prey and predators.

Humans harvest natural resources and receive many benefits from ecosystems. Resources and benefits that result from the operation of ecosystems are known as ecosystem services. Alterations of trophic cascades can change the services provided by ecosystems with important impacts on human welfare. For several centuries, the continental shelf off Nova Scotia, Canada, supported a highly productive cod (*Gadus morhua*) fishery. The cod population collapsed due to overfishing in the early 1990s. The loss of cod altered the system via a shift in cascading trophic interactions (Frank *et al.*, 2005). When cod collapsed, small pelagic fishes, crabs, and shrimp increased, zooplankton declined, and phytoplankton increased. Interestingly, this shift was observed despite the physical dynamics and the large food web complexity of the marine environment, which might be expected to obscure strong trophic cascades.

The collapse of cod had great economic impact, but the rise of the pelagic fishes and crustaceans led to a new fishery that currently produces a higher economic value than the prior cod fishery (Frank *et al.*, 2005). In this case, cascading trophic interactions changed with the resulting impact that the former ecosystem service was lost but a new service of greater value emerged. Although this may seem a positive outcome, the decline of cod had devastating social effects that extended beyond fishing. In most cases where ecosystem services are lost, they are expensive or in some cases impossible to replace. For example, in an agricultural region of south central Texas, USA, Brazilian free-tailed bats (*Tadarida brasiliensis*) prey on agricultural pests of cotton, reducing crop damage – a trophic cascade (Cleveland *et al.*, 2006). Brazilian free-tailed bats have a positive effect on both conventional and genetically modified cotton, reducing the number of sprayings of pesticides and hence costs to growers (Federico *et al.*, 2008). Other species of bats with similar impacts on agricultural systems are currently in decline in North America due to emerging diseases (e.g., a fungal infection of hibernating bats) and other factors (e.g., mortality from the growing use of wind turbines). The estimated value of bats in the US as agents of pest control may be over a billion of dollars (Boyles *et al.*, 2011). Loss of trophic cascades involving bats, insect pests, and crop plants would invoke significant costs due to lost production and increased expenses of treatment with pesticides.

Conserving Not Only the Big and Fierce but Also the Small and Timid

Colinvaux (1978) wrote a book explaining the general concepts of ecology including a problem posed in the book's title, "Why big fierce animals are rare (...)." The answer to this problem relates to the efficiencies of ecological food chains. You are more likely to encounter a squirrel in your backyard than a lion because squirrels feed lower on the food chain, mainly on a variety of plentiful plant items. Lions feed on vertebrate prey, which they must hunt, attack, and kill. Lions must pay a large cost in acquiring food, and there is a greater limitation of prey (fewer wildebeest than nuts). Energy limitation explains the rarity of large predators.

But predators are also rare because they have been persecuted by humans. In every ecosystem of the earth, apex predators have been killed because they either compete with or are a danger to humans. Apex predators are also targeted for food, for fur, and as trophies. Consequently, there is a second, more insidious answer to Colinvaux's problem. Large, fierce predators are rare because humans have diminished or totally eliminated their populations. The nexus of effects of this loss of top or apex predators has been recently dubbed "trophic downgrading" (Estes *et al.*, 2011). Generally, trophic downgrading of ecological systems has led to a loss of top-down regulation with changes in the abundance of consumers susceptible to predation as well as many consequences related to a lack of predators. For example, in the absence of wolves, abundant deer populations impact riparian vegetation, causing erosion and loss of shoreline integrity on stream banks (Ripple *et al.*, 2010). The global diminishment of the "big and fierce" has led to impoverished ecological interactions, some of which are now lost forever due to permanent predator extinctions.

Another consequence of the loss of large, apex predators is that smaller predators lower in the food web have become more abundant and have assumed in many systems the role of apex predators, a phenomenon known as mesopredator release. These mesopredators include a variety of species. In terrestrial environments they are often small- to medium-sized vertebrates. In aquatic systems they are either the younger and smaller age classes of the apex predator fishes or medium-sized fish species that were previously in the middle of the food web. Mesopredator release appears responsible for many ongoing changes in ecological systems. For example, the loss of large predators, including lions and panthers, has released olive baboons (*Papio anubis*) in Ghana, West Africa (Brashares *et al.*, 2010). In many locales, baboons are now hyperabundant. Baboons may shift to a more carnivorous diet when released, and there are correlations between increases in baboons and declines in numbers of small primates, in numbers of small ungulates, and in bird nesting success (Brashares *et al.*, 2010). In addition, marauding troops of baboons increasingly invade homes and destroy crops. This is one of many examples where mesopredator release has resulted in some species becoming abundant with attendant losses of biodiversity and negative impacts on humans (Prugh *et al.*, 2009).

Sustaining predators and restoring trophic cascades are important goals for conservation and restoration. Apex predators are a key target as conserving these species may also contribute to sustaining many attributes of the ecosystems that support them. For example, apex predators may promote species richness through trophic cascades and a variety of other mechanisms such as resource facilitation as when predators make carrion available for scavengers (Sergio *et al.*, 2008). Apex predators are also indicators. Where present, they are often associated with high biodiversity and greater ecosystem complexity. The absence of predators can indicate anthropogenic disturbance associated with lower diversity. Finally, because apex predators require large areas, their effective conservation can lead to the creation of large reserves and promote the establishment of corridors between conserved areas for predator movement (Sergio *et al.*, 2008).

With the perspective of trophic cascades, it is also clear that some intermediate species in the food web are necessary targets for conservation and restoration. An example is the European wild rabbit (*Oryctolagus cuniculus*). Within its native range, European rabbits are an important grazer and a key prey for many vertebrate predators, including two endangered species: the Iberian lynx (*Lynx pardinus*) and the Spanish imperial eagle (*Aquila adalberti*; Lees and Bell, 2008). Although the European rabbit is quite fecund and capable of rapid growth, populations are diminished in Iberia because of habitat loss, human hunting pressure, diseases, and predation. Hence, conservation of apex predators such as lynx depends, at least in part, on rabbit conservation, and rabbits are being restocked in areas of Spain and Portugal to promote predators and reestablish native trophic cascades. Maintaining the big and fierce also requires the small and timid.

Conclusion and Prospects

Trophic cascades are a feature of many ecological systems. Although cascades are now documented for many locales, the ecological rules that promote cascades are incompletely understood. This means, as in many areas of ecology, there is a need for theory that can better explain trophic cascades. The addition or removal of predators particularly as the result of human activity often makes trophic cascades apparent. A frontier research area is to gain a greater appreciation for the many consequences of trophic cascades, particularly those that impact ecosystem services. Another frontier is the restoration of lost trophic cascades. Trophic cascades should be partially restorable with the return of large apex predators to systems where they have been extirpated. Restoring and conserving these predators in a world of changing climate, growing human populations, and increased pressure on natural systems will require new knowledge about ecological systems, better surveillance of predator populations, and a willingness to manage ecosystems to promote predator survival. The alternative is a world of trophic downgrading and cascades lost.

See also: Biogeochemical Cycles. Carnivores. Ecosystem Services. Food Webs. Grazing, Effects of. Predators, Ecological Role of. Trophic Levels

References

Borer ET, Seabloom EW, Shurin JB, *et al.* (2005) What determines the strength of a trophic cascade? *Ecology* 86: 528–537.

Boyles JG, Cryan PM, McCracken GF, and Kunz TH (2011) Economic importance of bats in agriculture. *Science* 332: 41–42.

Brashares JS, Prugh LR, Stoner CJ, and Epps CW (2010) Empirical and conservation implications of mesopredator release. In: Terborgh J and Estes JA (eds.) *Trophic Cascades: Predators, Prey and the Changing Dynamics of Nature*, pp. 221–240. Washington, D.C.: Island Press.

Carpenter SR, Cole JJ, Hodgson JR, *et al.* (2001) Trophic cascades, nutrients, and lake productivity: Experimental enrichment of lakes with contrasting food webs. *Ecological Monographs* 71: 163–186.

Carpenter SR, Kitchell JF, and Hodgson JR (1985) Cascading trophic interactions and lake productivity. *BioSciences* 35: 634–649.

Cleveland CJ, Betke M, Federico P, *et al.* (2006) Economic value of the pest control service provided by Brazilian free-tailed bats in south-central Texas. *Frontiers in Ecology and Environment* 4: 238–243.

Colinvaux PA (1978) *Why Big Fierce Animals are Rare: An Ecologist's Perspective*. Princeton, NJ: Princeton University Press.

Estes JA and Palmisano JF (1974) Sea otters: Their role is structuring nearshore communities. *Science* 185: 1058–1060.

Estes JA, Terborgh J, Brasheres JS, *et al.* (2011) Trophic downgrading of planet earth. *Science* 333: 301–306.

Federico P, Hallam TG, McCracken GF, *et al.* (2008) Brazilian free-tailed bats as insect pest regulators in transgenic and conventional cotton crops. *Ecological Applications* 18: 826–837.

Frank KT, Petrie B, Choi JS, and Leggett WC (2005) Trophic cascades in a formerly cod dominated ecosystem. *Science* 308: 1621–1623.

Hairston NG, Smith FE, and Slobodkin LB (1960) Community structure, population control, and competition. *American Naturalist* 94: 421–425.

Hebblewhite M, White CA, Nielvelt CG, *et al.* (2005) Human activity mediates a trophic cascade caused by wolves. *Ecology* 86: 2135–2144.

Hrbáček J, Dvorakova M, Koninek V, and Prochazkova L (1961) Demonstration of the effects of the fish stock on the species composition of zooplankton and the intensity of metabolism of the whole plankton assemblage. *Verhanlungen Internationale Vereinigung fur Theoretische und Angewandte Limnologie* 14: 192–195.

Knight TM, McCoy MW, Chase JM, McCoy KA, and Holt RD (2005) Trophic cascades across ecosystems. *Nature* 437: 880–884.

Lees AC and Bell DJ (2008) A conservation paradox for the 21st century: The European wild rabbit *Oryctolagus cuniculus*, and invasive alien and an endangered species. *Mammal Review* 38: 304–320.

O'Dowd DJ, Green PT, and Lake PS (2003) Invasional 'meltdown' on an oceanic island. *Ecology Letters* 6: 812–817.

Oksanen L, Fretwell SD, Arruda J, and Niemelä P (1981) Exploitation ecosystems in gradients of primary productivity. *American Naturalist* 118: 240–261.

Pace ML, Cole JJ, Carpenter SR, and Kitchell JF (1999) Trophic cascades revealed in diverse ecosystems. *Trends in Ecology and Evolution* 14: 483–488.

Paine RT (1980) Food webs: Linkage, interaction strength, community infrastructure. *Journal of Animal Ecology* 49: 667–685.

Polis GA, Sears ALW, Huxel GR, Strong DR, and Maron J (2000) When is a trophic cascade a trophic cascade? *Trends in Ecology and Evolution* 15: 473–475.

Polis GA and Strong DR (1996) Food web complexity and community dynamics. *American Naturalist* 147: 813–846.

Post DM, Carpenter SR, Christensen DL, *et al.* (1997) Seasonal effects of variable recruitment of a dominant piscivore on pelagic food web structure. *Limnology and Oceanography* 42: 722–729.

Power ME (1990) Effect of fish on river food webs. *Science* 250: 811–815.

Prugh LR, Stoner CJ, Epps CW, *et al.* (2009) The rise of the mesopredator. *BioScience* 59: 779–791.

Schindler DE, Carpenter SR, Cole JJ, Kitchell JF, and Pace ML (1997) Influence of food web structure on carbon exchange between lakes and the atmosphere. *Science* 277: 248–251.

Ripple WJ, Rooney TP, and Beschta RL (2010) Large predators, deer, and trophic cascades in boreal and temperate ecosystems. In: Terborgh J and Estes JA (eds.) *Trophic Cascades: Predators, Prey and the Changing Dynamics of Nature*, pp. 141–161. Washington, D.C.: Island Press.

Sergio F, Caro T, Brown D, *et al.* (2008) Top predators as conservation tools: Ecological rationale, assumptions, and efficacy. *Annual Reviews of Ecology, Evolution, and Systematics* 39: 1–19.

Shurin JB, Borer ET, Seabloom EW, *et al.* (2002) A cross-ecosystem comparison of the strength of trophic cascades. *Ecology Letters* 5: 785–791.

Strong DR (1992) Are trophic cascades all wet? Differentiation and donor-control in speciose ecosystems. *Ecology* 73: 747–754.

Terborgh J and Estes JA (eds.) (2010) *Trophic Cascades: Predators, Prey and the Changing Dynamics of Nature*. Washington, D.C.: Island Press.

Whittaker RJ (2010) Meta-analyses and mega-mistakes: Calling time on meta-analysis of the species richness–productivity relationship. *Ecology* 91: 2522–2533.

Trophic Levels

Peter Yodzis, University of Guelph, Ontario, Canada

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 695–700, © 2001, Elsevier Inc.

Glossary

Basal species A species that eats no other species.
Ecological transfer efficiency The ratio of energy ingested by a population's predators to energy ingested by the population.
Food chain A sequential relationship of the form x_1 is eaten by x_2 which is eaten by x_3 which is eaten by … which is eaten by x_n.
Food web A specification of which species eat which in an ecosystem.
Trophic level of a species $1+$ a weighted average of the lengths of all food chains linking that species to basal species. Different weightings may be appropriate for addressing different questions.

Food Chains and Trophic Levels

Trophic ecology has to do with feeding relations—for instance, weasels eat mice, mice eat herbs—and is among the most basic organizing principles underlying biodiversity in natural ecosystems. One of the earliest attempts to identify ecosystem structure was based on trophic ecology, as follows. Some species in nature (mostly plants) do not eat anything; instead they utilize solar energy through photosynthesis and are called *primary producers* or *autotrophs*. Other species eat autotrophs and are called *herbivores* (Figure 1). *Carnivores* eat herbivores, *secondary carnivores* eat carnivores, *tertiary carnivores* eat secondary carnivores, and so on (Figure 1). These are, of course, highly aggregated entities: each of the categories just mentioned contains many species in any given ecosystem.

A *food chain* is a sequential relationship of the form x_1 is eaten by x_2 which is eaten by x_3 which is eaten by … which is eaten by x_n, where the entities x_i might be individual species, or they might be aggregations such as those just defined.

Early trophic ecology (approximately the 1930s through 1970s) viewed an ecosystem as a single food chain involving the aggregated entities just defined, as depicted in Figure 1. In this view of an ecosystem, autotrophs constitute "trophic level" 1, carnivores constitute "trophic level" 2, secondary carnivores are "trophic level" 3, and so on. (I am using "scare quotes" for the term "trophic level" because there are difficulties with this particular definition, *see* Food Webs and Trophic Levels.) The International Biological Program (IBP) of 1964–1974 was very much organized around this concept of ecosystem structure.

The concepts of food chain and trophic level are very closely related. Thus, in the formulation sketched in Figure 1, species in the nth trophic level are linked to primary production (autotrophs) through a food chain of length $n-1$.

There is another food chain that is very important in a great many ecosystems, namely, the one that is based on detritus. Detritus is certainly eaten by a wide variety of organisms, but it is not itself a living organism, so it cannot exactly be said to "eat" anything. However, biomass does move from living organisms into detritus as they decay after dying. If one is concerned to follow the recycling of specific nutrients, one needs to keep track of this entire dynamic. For purely trophic studies it is generally adequate (and far simpler) to treat detritus as though it were another organism that, like an autotroph, does not eat anything. Then autotrophs and detritus together are called *basal species*, and the detritivores are lumped together with the herbivores in trophic level number 1.

The Utility of Trophic Levels

The trophic level concept has been exceptionally durable: it has been one of the basic concepts of ecology for six decades and is one of the few ecological concepts contained in the vocabulary of most educated people. The reason for this distinguished place in the scheme of things is that the concept is both simple and useful. Furthermore, it is universal: it applies to all ecosystems.

Because of this universality, trophic levels enable us to compare the role of vastly different species in vastly different systems. For instance, we can discuss and understand a lake and the surrounding forest with a common language: the forest has its vegetation and its leaf litter; the lake has its phytoplankton and its dissolved organic matter (basal species). The forest has herbivorous insects, birds, and mammals; the lake has zoo-plankton (herbivores). And so on. We can use the same language to compare these two systems with any other ecosystem anywhere in the world.

This categorical and conceptual role can be made more quantitative and detailed, revealing important similarities and important differences among systems, by adopting a *bioenergetic* viewpoint, as follows.

Biological organisms contain caloric energy, which is transferred to organisms in the next step up a food chain: herbivores gain energy from consuming basal species, carnivores gain energy by consuming herbivores, and so forth. Each organism, or set of organisms such as a trophic level, *produces* energy at a certain rate. This is the maximum rate at which the next trophic level up the food chain could in principle ingest energy.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00145-3

The rate of energy production by a trophic level must necessarily be less than the rate of energy ingestion by that trophic level. First, not all energy ingested by an organism is available to be metabolized by that organism. Some of it will be lost to excretion. The ratio of metabolizable energy to ingested energy is called *assimilation efficiency* and is typically about 0.45 for herbivores and 0.85 for carnivores. Of the metabolizable energy, some is lost to respiration, being used up by the organism to carry out its various activities and also simply to live, and the remainder is available for the production of new tissue, which can in principle be consumed by the next trophic level. The ratio of energy production to metabolizable energy is called *production efficiency* and ranges from about 0.1 to 0.4 for invertebrate ectotherms to about 0.01–0.03 for endotherms. Furthermore, not all of the energy produced at a trophic level will actually be ingested by the next higher trophic level: much of it will be missed and end up as detritus. There is no generally agreed term for this form of energy loss, nor is there a great deal of quantitative data for it.

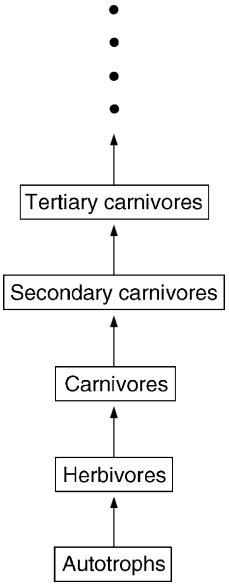


Figure 1 The IBP conception of an ecosystem as a linear sequence of trophic levels.

The *ecological transfer efficiency* of a trophic level is the ratio of (energy ingested from that trophic level by the next highest trophic level) to (energy ingested by that trophic level). It is the product of the three efficiencies explicated in the preceding paragraph. Ecological transfer efficiency ranges from 0.001 or smaller (depending upon losses to detritus) up to a maximum of about 0.5.

It quickly becomes apparent that energy is dissipated quite rapidly as we ascend a food chain. Suppose, for example, that each ecological transfer efficiency in a food chain is 0.1, which is rather high. Then of the energy produced by basal species (primary production), one-tenth is produced by herbivores. One-tenth of that, or one-hundredth of primary production, is produced by carnivores. One-tenth of that, or one-thousandth of primary production, is produced by secondary carnivores, and so on up the food chain. Less and less energy is available to higher trophic levels as we move up a food chain. One can visualize this phenomenon as a "pyramid of production" for a food chain (**Figure 2**).

This geometrically decreasing production as we move up a food chain, hence the rapidly decreasing available energy flow for the next higher trophic level, has been offered as one possible explanation for the apparently limited number of trophic levels in natural systems. Eventually, as we move up a food chain, the small fraction of primary production available to a putative next highest trophic level simply will not be enough to support a viable biological population.

The pyramid of production is an inescapable consequence of the dissipative processes, sketched above, that lead to ecological transfer efficiencies less than 1. There are a couple of similar "pyramids" that, while not universal in this way, are fairly typical of trophic levels. They follow from the circumstance that, for the most part, predators tend to be larger than their prey. For a predator to be larger (hence also faster and stronger) than its prey greatly facilitates the capture and consumption of prey.

Thus, as we move to higher trophic levels, we will, generally speaking, see larger animals. And yet, moving to higher trophic levels, these larger animals need to live on smaller energy production from the next trophic level down. As a result, there will usually be fewer animals at higher trophic levels. This "pyramid of numbers" is frequently, though not necessarily always, observed.

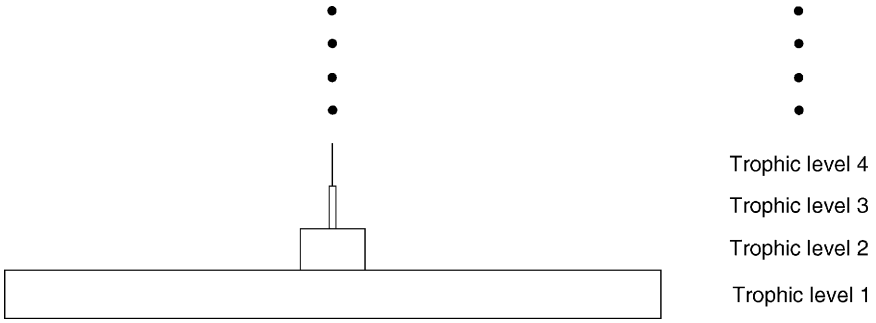


Figure 2 The pyramid of production in an ecosystem. The width of each layer is proportional to the rate of production of biomass in the corresponding trophic level. Because energy is dissipated in each transfer from one trophic level to the next, production decreases at an approximately geometric rate as trophic level increases.

An obvious exception to the pyramid of numbers emerges if we treat parasites and parasitoids as "predators"; they are almost always smaller than their "prey." Even though parasitism is tremendously widespread in nature, these are not really trophic relationships, and so most trophic studies do not include parasites or parasitoids.

What about total biomass [=(number of animals) × (weight of each animal)] at each trophic level? The number of animals tends to decrease as trophic level increases, while the weight of each animal tends to increase. The result is equivocal. Particularly in aquatic systems, where very small organisms at low trophic levels have very rapid rates of biomass turnover and can be grazed to quite low levels, one frequently (but not always) sees "inverted pyramids" of biomass, with more biomass at higher trophic levels. But terrestrial systems typically (though by no means always) display pyramids of biomass, with less biomass at higher trophic levels.

There are exceptions to this scheme, but they prove the rule; that is, they make sense in terms of the ideas underlying the scheme. For instance, some of the very largest animals, such as elephants and big ungulates, are herbivores. These animals are so large that they could not possibly range far enough to live by eating, say, lions. The only way to get a high enough energy density to support such large animals is by feeding directly on plants.

Just as energy propagates upward through food chains, so may chemical substances contained in organisms. This becomes particularly interesting when toxic contaminants are present. If those toxic substances are absorbed and/or ingested by animals at some trophic level, then, depending upon the rate at which they are excreted, there may be residues in the tissue consumed by higher trophic levels. Under some circumstances, the concentration of toxins may increase as trophic level increases, which is called *biomagnification*.

Food Webs and Trophic Levels

The conception of an ecosystem as a linear chain of "trophic levels" (**Figure 1**) is a useful starting point, but if we examine trophic relations with a higher degree of taxonomic resolution—that is, not lumping so many biological species together as we did in motivating **Figure 1**—we find quite a different trophic structure. A *food* web is a specification of which species eat which in an ecosystem. For instance, **Figure 3** is a food web for Wytham Wood, a forest near Oxford in England. An arrow from one kind of organism to another indicates that the organisms at the head of the arrow eat the organisms at the other end.

One can detect something like "trophic levels" here (partly because of the way I chose to draw the picture), but there is certainly not a simple flow of energy through a linear sequence of levels as in **Figure 1**; this picture is more "webby." For instance, we might put weasels at "trophic level" 5, because they

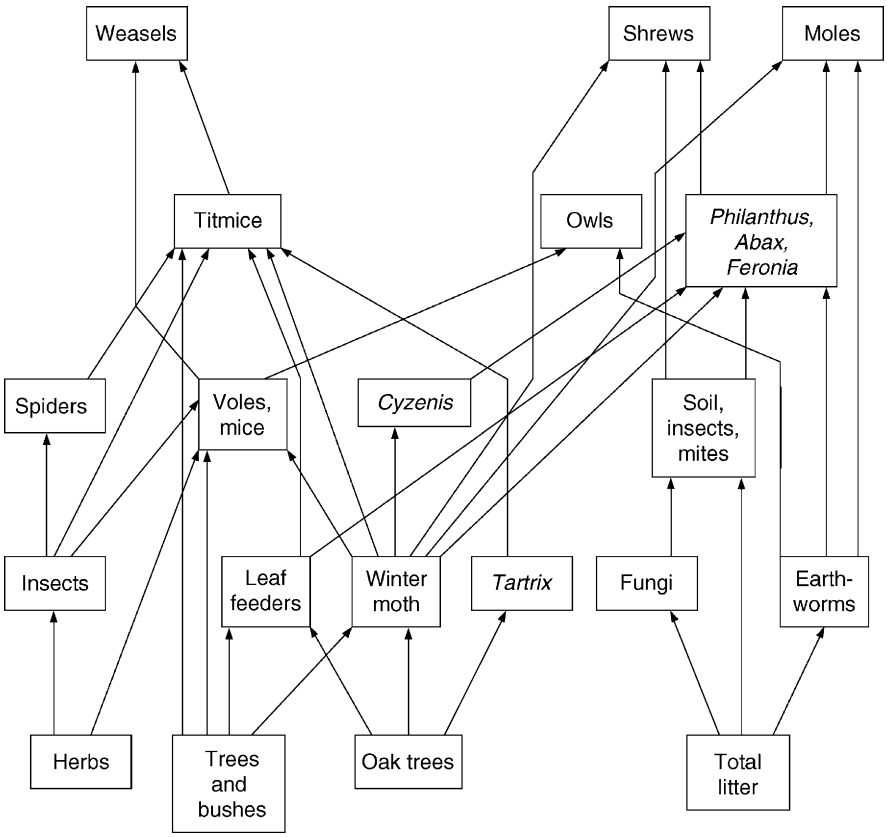


Figure 3 A food web for Wytham Wood, England (based on data in Varley GC (1970) The concept of energy applied to a woodland community. In: Watson A (ed.) *Animal Populations in Relation to Their Food Resources*. Oxford: Blackwell Sci).

Table 1 Food chains and trophic levels in the Wytham Wood food web of **Figure 3**

Species	*Number of chains to basal species of length*				*Trophic level using*		
	1	*2*	*3*	*4*	*Shortest chain*	*Longest chain*	*Mean length*
Insects	1	0	0	0	2	2	2.0
Winter moth	2	0	0	0	2	2	2.0
Tartrix	1	0	0	0	2	2	2.0
Leaf feeders	2	0	0	0	2	2	2.0
Earthworms	1	0	0	0	2	2	2.0
Fungi	1	0	0	0	2	2	2.0
Voles, mice	2	3	0	0	2	3	2.6
Spiders	0	1	0	0	3	3	3.0
Titmice	1	6	1	0	2	4	2.9
Cyzenis	0	2	0	0	3	3	3.0
Philanthus, Abax, Feronia	0	6	3	0	3	4	3.3
Soil insects, mites	1	1	0	0	2	3	2.5
Owls	0	3	3	0	3	4	3.5
Weasels	0	3	9	1	3	5	3.8
Shrews	0	3	7	3	3	5	4.0
Moles	0	3	6	3	3	5	4.0

eat titmice, which eat spiders, which eat insects, which eat herbs. But weasels also eat voles and mice, which eat herbs: this would put weasels at "trophic level" 3.

We do not want to throw away the trophic level concept altogether—it is too useful for that—but in the light of more refined data such as **Figure 3**, we need to refine our concept of trophic level. In fact, there are a number of different ways that we may define the term "trophic level," and it seems imprudent to insist that any one definition is "The Right" one. Rather, different trophic level concepts may be appropriate for different purposes.

The constant theme linking all trophic level concepts together is the idea that trophic level has to do with the lengths of food chains linking a species to basal species. We just have to bear in mind that a species will generally be linked to basals through several food chains, which might be of different lengths. For instance, there are 13 food chains that link weasels to basal species in the Wytham Wood food web of **Figure 3**. Three of these have length 2, 9 of them have length 3, and 1 of them has length 4. The relative importance assigned to these 13 food chains distinguishes several different definitions of "trophic level." Five commonly used definitions are listed here; following discussion of these, a sixth, in a considerably different spirit, will be addressed:

1. 1 + the length of the shortest food chain linking a species to some basal species.
2. 1 + the length of the longest food chain linking a species to some basal species.
3. 1 + the mean length of food chains linking a species to some basal species.
4. 1 + the weighted mean length of food chains linking a species to some basal species, where the weighting reflects energy flow through each food chain.
5. $\lambda + (\delta X_{\text{organism}} - \delta X_{\text{reference level}})/E$, where E is the average enrichment of a heavy isotope and $\delta X = [(R_{\text{sample}}/R_{\text{standard}}) - 1] \times 10^3$. Here X denotes the heavy isotope (for instance, ^{13}C, ^{15}N, or ^{34}S) and R denotes the heavy/light ratio (for instance, $^{13}C/^{12}C$, $^{15}N/^{14}N$, or $^{34}S/^{32}S$). The trophic level of the reference level organisms is λ: this might be basal species ($\lambda = 1$), or perhaps herbivores ($\lambda = 2$).

Definition 5 requires explanation. Certain heavy isotopes are enriched relative to the light isotope in biochemical reactions. As a result, the ratio of heavy to light isotope in an organism's tissue may bear a systematic relation to the ratio in that organism's diet. For instance, the ratio $^{15}N/^{14}N$ appears to be enriched by $E = 3.4\lambda‰$ ($\pm 1\lambda‰$) in a wide variety of organisms. Therefore, if we know the isotope ratio for organisms at some reference trophic level, we can use Definition 5 to calculate a trophic level from measurements of the isotope ratio in other organisms. The resulting numbers, which are a weighted average over all food chains from the reference level to the organisms in question, are probably fairly close to what we would get from Definition 4. This method requires far less effort than a direct calculation of Definition 4 would; for instance, there is no need to measure dietary proportions. It does require careful calibration of isotope ratios at the reference level, including allowance for possible differences among species at that level. In practice, use of the $^{13}C/^{12}C$ ratio has not been fruitful, in part because so little ^{13}C is enriched at each trophic transfer, but as of this writing, the $^{15}N/^{14}N$ ratio appears to hold promise as a tool for trophic studies.

Table 1 shows, for each nonbasal species in the Wytham Wood food web, the frequency distribution of food chains linking that species to basals and the consequent trophic level according to Definitions 1–3. (Definitions 4 and 5 require more data than are available for the Wytham Wood food web.) For instance, depending upon the relative importance attached to the 13 food chains linking weasels to basal species, we may put weasels at trophic level 3 (shortest chain; Definition 1), 5 (longest chain; Definition 2), or 3.8 (mean length; Definition 3).

Generally speaking, one would like to use something like Definition 4, even though it means replacing the notion of

discrete trophic levels with a trophic continuum. However, calculating trophic level in this way requires a tremendous amount of data. Definition 5 may be a good surrogate, but it still requires data beyond the food web itself. One is tempted to regard Definition 3 as a reasonably good substitute, but if we are thinking energetically, then because of the dissipation of energy as we move up a food chain, the shorter chains may well be more important energetically than the longer ones, so the equal weighting of Definition 3 may be deceptive. Definition 2 is what one has in mind implicitly when one draws tidy pictures such as **Figure 3**, but the existence of very long food chains can be deceptive. Animals that have a food chain to basals with 8, 9, or even 10 links exist, but they invariably also have chains no longer than 3 links, and these shorter chains are likely more important energetically. However, very long food chains may be particularly significant if one is concerned with biomagnification of toxin concentrations.

Another viewpoint, which ought to produce a sixth trophic level definition if it could be articulated precisely enough, emphasizes the *top-down* aspect of trophic relationships. This viewpoint, which has been put forward by S. Fretwell and L. Oksanen, counts a predator as one trophic level higher than its prey only if it significantly controls the biomass or dynamics of the prey species. Oksanen suggests that on this basis the distinction among carnivores, secondary carnivores, tertiary carnivores, and so on largely evaporates, leaving only three true trophic levels: basal species, herbivores, and carnivores—except in pelagic systems, where, due to the very small size of the primary producers and the consequent small size of zooplankton, planktivory and piscivory emerge as truly distinct trophic roles, permitting four trophic levels. This notion that there are actually only a few trophic levels despite the existence of some very long food chains is consonant with the implications of energetics noted in the preceding paragraph.

Thus, the term "trophic level" needs to be used with caution, and if we are to speak quantitatively of trophic levels, we need to specify exactly which definition we are using and to choose a definition that sheds the most light on the particular issues of concern.

See also: Ecotoxicology. Energy Flow and Ecosystems. Food Webs. Predators, Ecological Role of. Species Interactions

References

Polis GA and Winemiller KO (1996) *Food Webs: Integration of Patterns and Dynamics*. New York: Chapman & Hall.

Rundel PW, Ehleringer JR, and Nagy KA (eds.) (1989) *Stable Isotopes in Ecological Research*. New York: Springer-Verlag.

Varley GC (1970) The concept of energy applied to a woodland community. In: Watson A (ed.) *Animal Populations in Relation to Their Food Resources*. Oxford: Blackwell Sci.

Yodzis P (1989) *Introduction to Theoretical Ecology*. New York: Harper-Collins.

Tropical Forest Ecosystems

Gary S Hartshorn, World Forestry Center, Portland, OR, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Andepts Soils derived from andesitic volcanic ash.
Australasia Geographic region that encompasses Australia, eastern Indonesia, and nearby islands of the tropical southwest Pacific.
Great faunal interchange Development of the Central American isthmus facilitated the interchange of fauna from north to south and vice versa.
Indo-Malaya Geographic region from India to Malaysia and western Indonesia.
Intertropical convergence zone (ITCZ) Low-latitude region characterized by warm air masses and convectional rainfall; also known as the heat or thermal equator.
Neotropics Tropical portion of Western Hemisphere ranging from southern Mexico to northeastern Argentina, including the Caribbean and the tip of Florida; New World Tropics is synonymous.
Paleotropics Tropical regions of the Eastern Hemisphere extending from Africa through south and southeast Asia into the southwest Pacific; also known as Old World Tropics.
Rain shadow Less rainy sector on leeward side of mountains and protected valleys caused by descending warm air masses.
Terra firme Nonflooded landscape in tropical lowlands, usually characterized by highly weathered nutrient-poor soils.

Introduction

Though covering only 9% of the Earth's land, tropical forest ecosystems are famous for their exuberant greenness, impressive structure, and striking diversity of fauna and flora – and how poorly studied they are. Despite their attraction to nineteenth-century explorers such as Charles Darwin, Alexander von Humboldt, and Alfred Russell Wallace and the development of permanent field stations in the twentieth century (McDade *et al.*, 1994), we still know very little about the biodiversity in tropical forests and how species function and interact in these ecosystems. For example, the huge uncertainty about the total number of species on our planet is largely due to how little is known about the invertebrates living in tropical forest ecosystems. Even though the phrase "tropical rain forest" still conjures up vivid images and myths about this signature ecosystem, there is a great amount of variety and heterogeneity in the types and extent of tropical forest ecosystems (Leith and Werger, 1989). This essay first provides an overview of the principal physical factors influencing the distribution and variation in tropical forest ecosystems and then describes the major types and features of tropical forest ecosystems.

Physical Factors

Geography

Tropical forest ecosystems are usually differentiated into neotropical (New World Tropics) and paleotropical (Old World Tropics). In contrast to the Amazon's dominance of the Neotropics, the Paleotropics include two major regions: Africa and Asia (Richards and Phillips, 2005). African tropical forest ecosystems occur primarily in the Congo basin and the higher lands that border the basin. Outliers occur in a narrow coastal band of West Africa and in the Eastern Arc Mountains of East Africa (White, 1983). A narrow strip of tropical forests occurs along the eastern coast of Madagascar and on the western Ghats of India, as well as in patches on islands of the Indian Ocean. Because of striking biogeographic differences (e.g., Wallace's Line), tropical Asia is further differentiated into the Indo-Malayan and Australasian regions (Whitmore, 1984).

Tropical biogeography is profoundly influenced by the degree of insularity of major landmasses. For example, the lengthy isolation and connection of Australia, New Guinea, and other islands of Australasia facilitated the striking evolution and diversification of marsupials (e.g., tree kangaroos) in the absence of Indo-Malayan carnivores. The tree family Dipterocarpaceae dominates tropical forest ecosystems in the Indo-Malayan region, but this family has far less importance in Australasian tropical forests. Similarly, the long geological isolation of South America led to a distinctive fauna (e.g., anteaters, sloths, and tapirs) but many of these species were wiped out during the Great Faunal Interchange via the Central American isthmus (Bermingham *et al.*, 2005). In contrast, the continued isolation of oceanic islands (e.g., Galápagos and Madagascar) has enabled some of the unique fauna and flora to survive.

The latitudinal extent of tropical forest ecosystems is not defined by the astronomical limits of the tropic of Capricorn (23°30′ S) and the tropic of Cancer (23°30′ N). The ecological limits are determined by winter cold (which may not mean freezing temperatures) and inadequate rainfall to support forest vegetation. In general, tropical forest ecosystems extend farther poleward where moisture is adequate, such as in the moist lowlands of Australia, Madagascar, and Mexico. Where aridity precludes the presence of lowland forest (e.g., Argentina and India), tropical forest ecosystems may extend to higher latitudes on the flanks of cooler and moister mountains.

Though approximately 85 countries contain tropical forest ecosystems covering 18 million km^2, equivalent to about double the area of the USA, the top 10 countries comprise

Table 1 Top 10 countries with tropical forests

Rank	*Country*	*Forests (km^2)*	*Cover (%)*	*Annual change (%)*
1	Brazil	5,195,220	62	−0.42
2	Dem. Rep. Congo	1,541,350	68	−0.20
3	Indonesia	944,320	52	−0.71
4	Sudan	699,490	29	−0.08
5	India	684,340	23	+0.21
6	Perú	679,920	53	−0.22
7	Colombia	604,990	55	−0.17
8	Angola	584,480	47	−0.21
9	Bolivia	571,960	53	−0.53
10	Venezuela	462,750	52	−0.61
Subtotal top 10		11,968,820	49	−0.34

Source: FAO Global Forest Resources 2010.

66% of this total (see **Table 1**). The Amazon basin, shared by eight countries, has by far the largest area of tropical forest ecosystems, with five of the top 10 countries by extent of their forests. Vast tropical forest ecosystems also occur in the Congo basin and on the mega-islands of the tropical Far East (Borneo, New Guinea, Sumatra, and Sulawesi). Smaller but significant tropical forest ecosystems occur in the Caribbean, Mesoamerica, trans-Andean South America, coastal Brazil, West Africa, East Africa, Madagascar, South Asia, Southeast Asia, southern China, the Philippines, and many tropical Pacific islands.

Physiography

Regional features such as mountains and water bodies play an important role influencing climate, soils, and vegetation in addition to ecosystem functions. Because major mountain chains (above 2500 m elevation) block the normal flow of air masses, they are the most prominent physiographic feature influencing tropical forest ecosystems. The best example is the Andes, running along the western coast of South America. On the eastern (windward) side, copious rains fall usually throughout the year, whereas on the leeward (i.e., rain shadow) side, coastal deserts dominate – especially in Perú and Chile with the oceanic upwelling of the cold Humboldt Current. Many other mountain ranges (e.g., New Guinea, western India, East Africa, and Mesoamerica) have similar windward–leeward effects on climate and vegetation. Although lesser ranges may not create such striking differences, they provide impressive altitudinal gradients in physical features and they may also be biological barriers. Volcanic massifs (e.g., Mount Kilimanjaro), isolated mesas (e.g., tepuís of southern Venezuela), ancient geologic shields (e.g., Guianan and Brazilian), and uplifted limestone platforms (e.g., Yucatán, northern Borneo, Palawan, and Papua New Guinea) all contribute to the physiographic differentiation of tropical forest ecosystems.

Mountainous regions and high rainfall generate impressive river systems such as the Amazon, Brahmaputra, Congo, Ganges, Mekong, Niger, and Orinoco. These complex hydrographic systems both facilitate and impede ecological connections. The Amazon basin has three major types of streams: (1) white water (originating in the Andes and carrying relatively fertile sediments), (2) clear water (originating from terra firme in the lower basin), and (3) black water (originating from white sand in the basin). Tributaries cut deeply into mountain flanks and carry huge sediment loads. Their natural floodplains are vast, creating unique ecosystems such as the famous flooded forests in the Amazon basin. Major rivers also meander over surprisingly large areas (~25%), reworking the sediments into new landforms and creating oxbow lakes.

Continental islands (Australia), marine islands (e.g., Antilles, Galapagos, Hawaii, Madagascar, and south Pacific), islands on the continental shelf (e.g., Borneo, Hainan, Java, New Guinea, Philippines, Sri Lanka, Sulawesi, and Sumatra), and innumerable smaller islands are important physiographic features. Owing to degree of isolation, proximity to the "mainland," and colonization/extinction events, islands may develop distinctive fauna and flora that make significant contributions to tropical biodiversity.

Climate

Tropical climates are characterized by predictable temperature patterns with modest seasonal variation and precipitation regimes that are much more unpredictable and highly variable from year to year (Kellman and Tackaberry, 1997). Daily temperature range (on a sunny day) is far greater than the seasonal variation between the warmest and coolest months. In mountainous terrain, the daily temperature variation may be as large as 20 °C. Seasonal differences are usually determined by rainfall patterns, such that the dry season is called "summer" and the wet season is "winter." The rainy season(s) tracks the movement of the thermal equator, also known as the intertropical convergence zone (ITCZ), which is approximately 1 month behind the passage of the sun. The ITCZ brings a daily pattern of afternoon convectional rains, often via thunderstorms or the monsoon. When the ITCZ is seasonally away from the region, it is usually the dry season. When the sun is in the Southern Hemisphere, the northeast tradewinds dominate the Northern Hemisphere subtropics, for example, the Caribbean and Africa. In contrast to the persistent tradewinds, the Asian monsoon involves a reversal of winds, with a summer inflow of warm moist air and a winter outflow of cool, dry air.

In the Caribbean, the tradewinds pass over warm seawater, bringing considerable moisture and clouds to mountainous

land near the sea. The cloud cover and moisture create the renowned cloud forests of the windward Caribbean region, such as Monteverde (Costa Rica), Blue Mountains (Jamaica), El Verde (Puerto Rico), and Rancho Grande (Venezuela). On the opposite side of the mountains, descending tradewinds produce much more arid forest ecosystems on the leeward slopes. The same windward–leeward phenomenon occurs on the slopes and valleys of major mountain ranges such as the Andes and Himalayas (Hamilton *et al.*, 1995). The front ranges are exposed to moisture-laden air masses and orographic rainfall, while interior valleys are much more arid.

In tropical regions, considerable between-year variation in total rainfall and its seasonality is the normal pattern, regardless of the average annual rainfall. Long-term rainfall records indicate that the annual maximum is usually more than twice the amount of the driest year. Near the equator, there are usually two alternating rainy and dry seasons, each approximately 3 months in duration. As one moves away from the equator, one dry season (when the sun is in the other hemisphere) lengthens and the other shortens or disappears. Rainfall regimes tend not to be simple or predictable, because short rainless periods can occur during the rainy season or isolated rains can occur during the dry season. The El Niño Southern Oscillation (ENSO) phenomenon occasionally eliminates or greatly delays one of the seasons; that is, the rains may be very late in arriving or they may override a normal dry season.

Soils

Tropical soils have been influenced by major geologic phenomena such as plate tectonics, weathering of ancient landscapes, uplifting of mountain ranges, volcanic outputs, and oceanic submergence or emergence. Erosion of the ancient Guianan shield contributed the extensive white sand soils of the Río Negro basin, leaving the impressive flat-topped mesas ("tepuís") that dominate southern Venezuela. Many of the infertile, terra firme soils of the Amazon basin were eroded from the geologically ancient Brazilian shield. The active floodplains of several major western Amazonian rivers receive sediments eroded from the geologically young Andes, thus juxtaposing annually renewed, fertile floodplain soils with the infertile terra firme soils that are so typical of much of the Amazon basin.

Much volcanic activity has characterized the Quaternary, especially in mountainous regions of the Tropics. Volcanic ash deposits and mudflows tend to be more prevalent than lava flows (Hawaii is an exception). There are two major types of volcanic ash: andesitic and rhyolitic; the former is rich in certain nutrients, whereas the latter is high in silica. Soils developed from andesitic ash are termed "andepts" and are famous as the best soils for coffee or tea (e.g., Assam, Colombia, Central America, Hawaii, Jamaica, Java, and Kenya). The natural fertility of andepts is the foundation of the highly productive paddy rice cultures of Southeast Asia. In contrast, extensive soils derived from rhyolite in Honduras and Nicaragua have low agricultural potential and are dominated by open pine forests.

Tropical soils vary greatly in their principal chemical and physical properties, such as parent material, soil structure, drainage regime, and age. Variations in soil properties play an important role in the different types of tropical forest ecosystems; however, we do not know much about their controlling influence on the distribution of most tropical tree species. The more restrictive a site is, the more distinctive the vegetation may be. For example, the combination of nutrient-poor, white sand soils and a perched water table in southern Venezuela gives rise to a dwarf shrubby vegetation ("bana" or "caatinga") in a tropical rain forest region.

Major Forest Ecosystems

Tropical forest ecosystems are extremely rich in species of both fauna and flora. The highest tree species richness numbers – approximately 300 species per hectare – have been documented from the western Amazon and northern Borneo. Lowland forests of the tropical Far East and Amazonia are far richer in species than is the Congo basin. Even more striking is the fact that if one excludes tree species from the total floristic richness, tropical forest ecosystems are still the richest ecosystems in the world, due to the species diversity of nontrees such as epiphytes and climbers (Gentry, 1990). Given similar rainfall regimes, central Amazonian forests tend to have fewer epiphytes than Mesoamerican forests, whereas tree species richness may be three times greater in the former.

There are many classification systems for differentiating tropical forest ecosystems that use bioclimatic, structural, or floristic criteria (Holdridge *et al.*, 1971). Each has its purposes and merits, but they are too detailed for general overview purposes such as this article. Rather, a simple artificial system (Hartshorn, 2000) is used that first recognizes three elevational zones – lowlands (<500 m), low mountains (500–2000 m), and high mountains (>2000 m) – and then differentiates three major moisture regimes – perhumid (wet), humid (intermediate), and subhumid (seasonally dry). Brief overviews are also included for some distinctive tropical forest ecosystems such as mangroves, flooded forests, savannas, and páramo. The interested reader must keep in mind that these are arbitrary groupings along continuous gradients, analogous to differentiating the colors of the rainbow, and that it is impossible to assign regional altitudinal or rainfall limits for major types of tropical forest ecosystems.

Lowlands (< 500 m)

Classic lowland tropical forest ecosystems extend onto the foothills and lower slopes, but it is difficult to define an upper ecological limit for the lowlands. Some authors prefer an elevation (e.g., 1500 m), whereas others use the upper limit of typical lowland crops (such as the "coffee line") or more conventional bioclimatic, structural, or floristic criteria. To include the very representative lowlands of the Amazon and Congo basins (with base elevation of ~300 m) as well as extensive foothills, the arbitrary upper limit of 500 m for lowlands is used here.

Mangroves

The mangrove forest ecosystems, probably the best known tropical forest ecosystem, are characteristic of most tropical

coastal zones due to the ability of several unrelated tree species to grow in saline conditions. Mangroves are typical of estuaries, deltas, and low-energy coasts lacking strong wave action. Owing to greater tidal fluctuation and lower base flow of rivers associated with longer and more severe dry season, mangrove forests extend substantially farther inland on the western side of landmasses than on eastern coasts. For example, mangroves extend 190 km inland along the Gambia river in West Africa. Neotropical and West African mangrove forests have far fewer tree species than the mangroves of the Indian Ocean and the tropical Far East. Most mangrove species have evolved effective ways of dealing not only with their salty environment (e.g., leaves that exude salt through specialized glands), but also with the anaerobic sediments in which they grow. Mangrove ecosystems are important nurseries for shellfish as well as finfish and offer frontline buffering against tropical storms.

Flooded Forests

In the wet lowlands, flooded forests occur behind the mangroves (where salinity does not influence the floristic composition of the forests), along many major rivers such as the lower Amazon and in low-lying areas with poor drainage. The frequency, depth, and length of flooding appear to be important physical factors determining the composition and dominance of flooded forests. Many flooded forests are pure stands of a single tree species, such as *Campnosperma brevipetiolatum* (New Guinea), *Campnosperma panamensis* (Panama and Costa Rica), *Mauritia flexuosa* (Amazon basin), *Manicaria saccifera* (Neotropics), *Mora oleifera* (circum-Caribbean), *Nypa fruticans* (New Guinea), *Prioria copaifera* (Central America), *Pterocarpus officinalis* (Costa Rica), *Raphia taedigera* (Neotropics and West Africa), and *Shorea albida* (Borneo). We do not have a good understanding of why any one of these species is able to dominate a particular ecosystem.

In contrast to flooded forests dominated by a single species, more heterogeneous flooded forests occur farther from the sea or adjoining the pure stands. The dominant tree species of pure stands also occur in the more mixed forests, usually with far less dominance. Swamp forests may have very large trees, but tree density tends to be lower in swamp forests than in adjoining terra firme forests. Most tree species are not restricted to swamp or flooded forests, but may occur on terra firme, where rainfall is adequate and especially where there is no effective dry season. Even with a mixture of tree species, these flooded forests are not nearly as species rich as nearby terra firme forests.

Many tree species that occur in flooded forests have large seeds that float, as can be seen in the flotsam on tropical beaches. One of the most characteristic trees of tropical beaches, the coconut (*Cocos nucifera*, Palmae), has such a wide pantropical distribution that it is difficult to determine its geographic origin. Several tree genera with large, buoyant seeds have the same or a sibling species across the Atlantic Ocean, such as *Elaeis* (Palmae), *Pentaclethra* (Mimosaceae), *R. taedigera* (Palmae), and *Sacoglottis* (Humiriaceae).

Perhumid Forests

Lowland perhumid forests are well known as classic tropical rain forest (Richards, 1996). Curiously, these forest ecosystems are more typical of the foothills and lower slopes of front ranges than of the extensive Amazon or Congo lowlands. The absence of major mountain ranges in tropical Africa may by a factor in the comparative scarcity of perhumid forest ecosystems there, especially in the Congo basin. Coastal West Africa from Guinea to Liberia is one of the few regions on that continent with average annual rainfall exceeding 3000 mm. Other lowland perhumid forests (e.g., southeastern Mesoamerica, Colombian Chocó, Gulf Province of Papua New Guinea, north Borneo, and Chittagong in Bangladesh) may have annual rainfall of 6000–10,000 mm.

Lowland perhumid forests on adequately draining soils have high tree diversity, consistently exceeding 100 species per hectare for trees 10 cm or more in diameter (diameter at breast height (dbh)). Some centers of diversity usually exceed 200 tree species per hectare and a few exceptional areas (e.g., eastern Ecuador, northeast Peru, central Guyana, and northern Borneo) report more than 300 tree species per hectare. Though forest stature may be lower and tree density higher on poor soils, these forests typically have 500–600 trees per hectare (>10 cm dbh). Thus, the great majority of tree species are represented by only one individual per hectare. Lowland perhumid forests on terra firme soils with adequate drainage show considerable variation in the distribution of tree species over a mosaic of forest types.

Palms are very rich in species in lowland perhumid forest ecosystems. Because of their distinctive leaf form, palms are conspicuous in these forests, ranging from dwarf species (<1.5 m tall) to understory colonial species, to elegant single stems that reach the subcanopy, and to the famous climbing rattans. A few palm genera have radiated into several hundred species of climbing palms in the Indo-Malayan region; commercial rattan comes from robust climbing palms. In neotropical forests, the climbing palm, *Desmoncus*, has just a few dozen species and they are not nearly as abundant as the rattans of the tropical Far East. In some neotropical forests (e.g., La Selva in Costa Rica), palms comprise up to 25% of the stems (>10 cm dbh) in heterogeneous old-growth forests. Palms tend to be less abundant and less diverse in secondary forests.

Plants dependent on other plants for support, such as epiphytes, climbers, and stranglers, are abundant in lowland perhumid forests, especially on tree branches in the forest canopy. Many plant families have epiphytic species, but most typical of lowland perhumid forests are orchids, ferns, gesneriads, and bromeliads – the latter are restricted to the Neotropics. There are even epiphytic cacti (*Epiphyllum* and *Rhipsalis*) in lowland perhumid forests. The latter genus (with sticky, bird-dispersed seeds) is the only member of this neotropical plant family present in Africa. Lianas (woody climbers) are one of the most striking components of tropical forests because of their growth form, stem structure, and impressive size; they may be 35 cm in diameter and greater than 100 m long as they drape over the crowns of canopy trees. Many plant families are represented by lianas, but they are especially common in the Apocynaceae, Bignoniaceae, Dilleniaceae, Fabaceae, and Marcgraviaceae. In contrast to robust lianas, slender herbaceous climbers are typical of the Araceae, with hundreds of species in several genera.

Although some stranglers belong to the genus *Clusia* (Clusiaceae), most are figs (*Ficus*, Moraceae). A strangler

begins life as an epiphyte, sending down aerial roots, eventually forming an interconnected network of roots that embrace the host tree. A strangler may survive as an independent tree for many years after the host tree dies and decomposes, leaving a "hollow cylinder" as its testimony. Interestingly, the strangling part of this relationship has not been confirmed; death may be associated with the inability of the host tree to tolerate shading by the strangler. Each fig species is pollinated by a unique species of tiny fig wasps, representing one of the best examples of coevolution. Figs are important keystone food resources for vertebrate fruit eaters in many tropical forests.

Lowland perhumid forests are extremely dynamic because tree falls are frequent events. The fall of a large canopy tree creates a sizeable opening in the forest canopy and prime regeneration sites on the forest floor. Up to 50% of the tree species in heterogeneous old-growth neotropical forests may be shade intolerant and dependent on canopy gaps for successful regeneration. In addition to the well-known "fugitive" pioneer species (e.g., *Cecropia*, *Macaranga*, *Musanga*, *Ochroma*, and *Trema*) that characterize the rapid recolonization of disturbed areas, many valuable timber species – such as mahoganies, Spanish cedar, and dipterocarps – require gaps to reach the canopy. Numerous studies of many tropical forests over the past few decades indicate that turnover rates in the order of 100 years are the norm in tropical forests. Gap-phase dynamics appears to be an important ecological contributor to the maintenance of species-rich tropical forests.

Humid Forests

Lowland humid forests are the most widespread of tropical ecosystems, covering much of the Amazon and Congo basins, Mesoamerica, and the tropical Far East. Mean annual rainfall typically exceeds 1500 mm, but the upper limit is more related to the seasonality of rainfall. Although lowland perhumid forests have only a short or ineffective dry season, lowland humid forest ecosystems are characterized by a significant dry season. These latter forests are almost as equally impressive as the former (e.g., the tallest trees in Ghana reach 60 m); however, the number of deciduous trees is much greater in the seasonal forests. Epiphytic orchids and gesneriads are less abundant in the canopy of humid forests than in perhumid forests. Smaller, thin, and supple woody vines are more common in humid forests, perhaps related to the greater proportion of deciduous species in the forest canopy. Large palms are common subcanopy components in humid forest ecosystems. Because of fire tolerance and the adaptive strategy of building the trunk belowground, some of these palms (e.g., *Attalea speciosa*, *Copernicia* spp., and *Scheelea rostrata*) are aggressive colonizers of pastures created from humid forests and maintained by fire.

Seasonal rains and drought not only determine the ecological functions and processes, but also have strong influence on the populations of animals and plants (Leigh *et al.*, 1982). The synchronization of flowering with the dry season is truly impressive. Flowering during the dry season when tree crowns and lianas are deciduous facilitates visitation by pollinators such as large bees. The alternation of rainy and dry seasons and the use of seasonal changes to trigger physiological responses (i.e., flowering, insect emergence, nesting, and rearing of young) make these ecosystems unusually susceptible to climatic fluctuations such as ENSO events. The best example is the ENSO-caused rains during the dry season that occasionally limit flowering and cause famine among fruit-eating animals on Panama's Barro Colorado Island (BCI).

The pioneering studies of the spatial distribution of trees by Hubbell and associates on BCI indicate that very few tree species are hyperdispersed as suggested in the classic literature on tropical forests. Many species have the adults randomly distributed, while some species show clumping. For example, several caesalpinioid legumes (*Brachystegia laurentii*, *Cynometra alexandri*, *Gilbertiodendron dewevrei*, and *Julbernardia seretii*) form single-species stands in the Congo basin. A long-term study on the 50-ha plot at BCI has elucidated evidence for strong density dependence within species, suggesting an important role for interactions between plants and their predators and pathogens. The spatial distribution of plants has important consequences on pollination, breeding systems, herbivory, and successful regeneration.

One of the most famous tropical trees is the Brazil nut (*Bertholletia excelsa*, Lecythidaceae), which has a patchy distribution in Amazonian humid lowland forests. Because adult trees occur in clumps, there has been considerable conjecture about indigenous influence on the abundance of this valuable species. Detailed studies document that the clumping is natural, probably due to the scatter-hoarding behavior of forest agoutis (*Dasyprocta leporina*, Heteromyidae) that cache seeds for later use. Another neotropical tree species with an unusual distribution is the true mahogany (*Swietenia macrophylla*, Meliaceae), which is reasonably abundant in the subtropical humid forests of northern Mesoamerica and the southern Amazon basin but quite scarce in the intervening tropical forests. Because of its high-value timber, there have been numerous attempts to promote or increase its abundance in natural forests, but they have mostly failed. Recently, it has been suggested that mahogany requires large-scale disturbance such as by fire or river meanders, creating the high light regime needed for establishment and growth of young trees.

One of the more fascinating historical aspects of this ecosystem is the development of truly impressive human civilizations such as the Maya in northern Mesoamerica and the Khmer in Kampuchea (Cambodia). Both cultures developed intensive agriculture based on sophisticated engineering and management of water during the dry season through the use of raised fields, irrigation canals, and paddies. The Maya are believed to have also favored particularly useful plants such as mahogany, *Manilkara zapota* (Sapotaceae), and *Brosimum alicastrum* (Moraceae). Approximately 900 CE, the classic Maya civilization collapsed, perhaps due to overpopulation and/or a prolonged, severe drought. Given our current understanding of tropical forest dynamics, it is quite unlikely that the humid lowland forests now occupying the classic Maya region are successional forests.

Subhumid Forests

Lowland subhumid tropical forest ecosystems are much more extensive in the Paleotropics (e.g., Madagascar, India, and Indochina) than in the Neotropics. The high-value teak (*Tectona grandis*, Verbenaceae) forests of Burma, Thailand, and Java are representatives of this ecosystem, as are the vast sal forests

(*Shorea robusta*, Dipterocarpaceae) of northern South Asia. These ecosystems also occupy significant areas in the Caribbean, along the Pacific coast of Mesoamerica (Bullock *et al.*, 1995) and northeast Brazil. Mean annual rainfall seldom exceeds 1500 mm, but the determining ecological factor is the long and severe dry season. Because of the historic pastoralist tradition in this ecosystem and the prevalence of nearly annual fires, most lowland subhumid forests have been destroyed or severely degraded by centuries of use (Laurance and Peres, 2006). One of the most notable exceptions is the well-studied dry forest of Guánica on the south coast of Puerto Rico (Lugo and Lowe, 1995). This island is unusual in that the economic pressures to use land were not nearly as intense in the twentieth century compared to most developing countries.

Lowland subhumid forest ecosystems have far lower tree species richness than humid and perhumid forests. They are also much shorter in height and simpler in structure, but virtually all woody species are deciduous. In low-stature forests, the shrub layer is often dense and very spiny; thin lianas are abundant (including the *Vanilla* orchid), but epiphytes are less rich in species and abundance. Terrestrial cacti and bromeliads are more common, especially on extreme sites such as rock outcrops. At the latitudinal limits of the Tropics, lowland subhumid forest ecosystems include chaparral, open pine forests, and mixed oak–pine forests – depending primarily on moisture availability. Evergreen forests occur in riparian situations and in other moist sites such as in karst (limestone) terrain, where underground moisture is available during the long dry season.

In contrast to the pure stands in some freshwater swamps mentioned earlier, pure stands in lowland subhumid forest ecosystems usually can be attributed to edaphic factors, such as rhyolitic ash (tuff or pumice) or montmorillonite clay (black cotton soils). One of the more distinctive forests is the pure stands of the only lowland tropical oak species, *Quercus oleoides*, that has a patchy distribution on rhyolitic ash deposits from southern Mexico to northwestern Costa Rica. Appreciable areas of lowland subhumid forest ecosystems in Mesoamerica exhibit single-species dominance, such as *Celaenodendron mexicanum* (Euphorbiaceae), *Crescentia alata* (Bignoniaceae), *Erythrina fusca* (Fabaceae), and *Parkinsonia aculeata* (Caesalpiniaceae). Despite the attribution of these single-dominant stands to restrictive soil factors, there is precious little clarity as to why a species is able to attain such dominance.

Savannas

Savannas are defined as having a continuous cover of grass, but trees may be conspicuous components of savanna landscapes. Though vast areas in the major tropical regions meet the criteria for savannas, it is much more difficult to ascertain if it is natural or derived through human activities (e.g., burning). Tropical Africa epitomizes the image and grandeur of savannas with vast expanses of grass, sparse trees, migratory herds of herbivores, and their predators. Recently, the role of elephant herds in transforming woodland to savanna has been well documented. Extensive savannas also occur in the tropical Far East and the Neotropics (the llanos of Colombia and Venezuela, Rupununi in Guyana, and Beni in Bolivia).

The extensive pastures of Mesoamerica are considered to be derived savannas created by almost annual burning to "freshen" pastures and suppress woody invaders. Neotropical savannas have some very characteristic and fire-resistant tree species, such as *Acrocomia vinifera* (Palmae), *Byrsonima crassifolia* (Malpighiaceae), and *Curatella americana* (Dilleniaceae). Lowland pine (*Pinus caribaea* var. *hondurensis*, Pinaceae) savannas occur in northeastern Nicaragua, much of lowland Honduras, part of the Guatemalan Petén, and east central Belize. In contrast to the derived savannas of Mesoamerica, the pine savannas appear to be natural ecosystems maintained by frequent fire. The flat pine savannas of Belize occur on sandy soils that flood during the rainy season and undergo severe drought in the dry season.

Low Mountains

For purposes of this essay, low-mountain ecosystems occur between 500- and 2000-m elevation. The rather arbitrary upper limit represents the maximum extent of tropical lowland flora near the equator. In many regions, the transition to montane flora more typically occurs by 1500 m, while nearer the latitudinal limits of the Tropics, this transition can be as low as 500 m. This elevational limit for tropical plants is more noticeable with crops such as highland coffee (*Coffea arabica*, Rubiaceae) or tea (*Camellia sinensis*, Theaceae). Low-mountain ecosystems typically have a mixture of tropical and temperate fauna and flora; for example, pines and oaks are conspicuous trees in northern Mesoamerican forests. Some tree genera have pairs of closely related species, with one occurring at lower elevations and the other in higher elevations in Mesoamerica, for example, *Billia* (Hippocastanaceae), *Brunellia* (Brunelliaceae), *Hedyosmum* (Chloranthaceae), and *Symplocos* (Symplocaceae). In some ways, low-mountain ecosystems between 500 and 2000 m are transitional between lowland tropical and tropical montane floras. These ecosystems are typical of the major mountain ranges as well as the many smaller chains and high plateaus (e.g., Brazilian Shield, Serengeti Plain, and Great African Plateau).

Subhumid Forests

Low-mountain subhumid forest ecosystems are common on leeward slopes and in valleys and plateaus protected from orographic rain. Because of favorable climate, this ecosystem has had a long history of human activities, leaving little natural vegetation. Subhumid forests are relatively open and low in stature, with an abundance of spiny trees and shrubs, especially Mimosaceae. In northern Mesoamerica and continental Asia, conifers (*Pinus* spp.) are common in subhumid forest ecosystems. *Pinus kesiya* and *Pinus merkusii* occur in the Malay archipelago, while *Pinus wallichiana* and *Pinus roxburghii* are common on the lower flanks of the Himalayas.

This ecosystem is extensive on the plateaus that border both the Amazon and Congo basins. The Great African Plateau is dominated by miombo woodland (dominated by *Brachystegia*, *Julbernardia*, and *Isoberlinia*, Caesalpiniaceae), while the Brazilian shield is characterized by similar open forest ("cerrado"). The Brazilian cerrado are low open forests or woodlands with characteristically crooked trees that occur on

soils high in aluminum and low in pH. Cerrado vegetation exhibits considerable variation, ranging from grass-dominated savanna to fairly dense forest; however, most of the plant species are of tropical origin. Where soils are less restrictive, more typical tropical forest exists that is surprisingly tall (to 40 m), fairly open canopy, with a high proportion of deciduous species.

Humid Forests

Humid forest ecosystems often ring lower, more arid regions on low mountains, because of more moisture and cooler temperatures, and/or less severe drought. Low-mountain humid forests show enormous variation of forest type, structure, and floristic composition due to their widespread distribution and great heterogeneity of physiography, soils, and rainfall patterns. An excellent example of this variation occurs in northern Mesoamerica, where pine forests dominate the upper slopes and broad-leaved forests occur in the valleys. Pines are characteristic of low-fertility soils and frequently burnt sites, with *Pinus oocarpa* often dominant. The broad-leaved forests usually have a mixture of tropical (e.g., *Alchornea*, Euphorbiaceae) and temperate tree genera. *Parinari excelsa* (Chrysobalanaceae) has an unusually broad distribution in Africa, dominating humid forests in both the Congo lowlands and the Guinea highlands.

Impressive tropical forests occur on fertile soils (i.e., andepts) in Mesoamerica, with a high density of canopy trees (typically in the Lauraceae and Sapotaceae) plus occasional tall (>50 m), emergent trees of *Ulmus mexicana* (Ulmaceae). Also occurring in these mid-elevation humid forests are canopy trees of Podocarpaceae (Southern Hemisphere conifers), Juglandaceae, and Proteaceae. The latter tree family (that includes *Macadamia*) is especially diverse and abundant in the Queensland Wet Tropics of northeastern Australia.

Perhumid Forests

These high-rainfall ecosystems are common on the flanks of low mountains and ranges that receive moisture-laden clouds from the east. When the mountains are oriented more or less perpendicular to the prevailing air currents, they nurture the famous cloud forests of the Caribbean, Mesoamerica, northern and eastern Andes ("ceja de selva" and "yungas"), and even on isolated volcanic massifs. Perhumid forests on low mountains may have tree species richness comparable to lowland perhumid forests. They also exhibit complex vertical structure and fairly tall trees; the latter may be related to the better drainage and anchorage typical of dissected topography.

Canopy trees are usually evergreen (with abundant epiphylls) and with heavy epiphyte loads, including bryophytes, shrubs, and even small trees. The abundant mosses in the canopy give these perhumid forests a distinctive brownish hue visible from small planes. Epiphytic Ericaceae shrubs are usually very abundant in cloud forests, providing year-round nutrition to nectar-feeding birds. The sizeable and heavy (when wet) loads of organic material in tree crowns cause frequent branch falls as well as stimulating the host tree to produce feeding roots on the branches to take up nutrients. In perhumid forests on nutrient-poor soils, numerous species of insect-trapping, epiphytic pitcher plants (*Utricularia*, Lentibulariaceae) are quite abundant.

High Mountains (>2000 m)

Because several tropical lowland species and crops may grow on protected sites up to 2000-m elevation, we use this as the lower limit of high-mountain forest ecosystems. High mountains in the Tropics have a temperate climate characterized by warm diurnal temperatures (on a sunny day) and cold nocturnal temperatures, including occasional freezes. In Africa, the fairly isolated mountains are defined as the Afromontane zone. Treeline varies considerably between approximately 3300 and 3800 m, depending on moisture, disturbance by fire, and proximity to the equator. Despite the temperate climate, most forest ecosystems have distinctive tropical features such as epiphytes and lianas. Nevertheless, the high-mountain tropical flora has much stronger affinities with temperate zone flora (Churchill *et al.*, 1995).

Subhumid Forests

In the Afromontane region of East Africa, *Juniperus procera* (Cupressaceae) forms pure stands; however, it is unclear if human disturbance is a contributing factor to this species' success. It is possible that high-mountain subhumid forests once existed on the Andean Altiplano, but centuries of human activities have destroyed most natural vegetation. The only remnants may be the small stands of *Polylepis* (Rosaceae) that occur up to 4000-m elevation in the central Andes – the highest forests in the world. The forest patches persist on rocky slopes that may provide natural protection from frequent fires, and soil temperatures may be slightly higher. *Polylepis* trees are small, twisted, and gnarled, with considerable trunk and branch dieback, especially on stand edges.

Humid and Perhumid Forests

The high elevations and cool temperatures tend to produce a preponderance of humid and perhumid forest ecosystems because of frequent cloud cover and low evapotranspiration. These ecosystems usually have the canopy dominated by holarctic genera (e.g., *Pinus*, *Quercus*, and Lauraceae), whereas the understory is more of a mixture of holarctic and tropical genera. Northern Mesoamerican forests are the center of diversity for oaks (*Quercus*, Fagaceae) and *conifers* (*Abies*, *Cupressus*, *Juniperus*, *Pinus*, *Podocarpus*, *Taxodium*, and *Taxus*). Except for *Podocarpus*, these are all holarctic genera. On poor soils, the coniferous forests tend to be open and frequently degraded by grazing and fire. More fertile and/or wetter sites also have conifers, but the forests are more of a mixture with oaks. Northern Hemisphere conifers do not occur naturally south of the Nicaraguan lowlands; thus oaks dominate the high-mountain forests of southern Mesoamerica and Colombia. The bamboos *Chusquea* in the Neotropics and *Arundinaria alpina* (both Poaceae) in East Africa are typically very abundant in the understory, where they may be impeding natural regeneration of trees. As with lower elevation perhumid forests, the large canopy trees carry huge epiphyte loads of mosses, tank bromeliads, and ericaceous shrubs.

Páramo

Although not strictly a forest ecosystem, high-mountain páramo merits brief mention as a unique high-mountain tropical ecosystem. Subalpine páramo occurs above treeline on high

mountains in the Neotropics (Ecuador to Costa Rica) and East African volcanic massifs. Neotropical páramo is characterized by a high density of shrubs, including *Chusquea*, *Hypericum* (Hypericaceae) and *Vaccinium* (Ericaceae), and occasional emergent trees of *Clusia* (Clusiaceae). Where drainage is poor, bogs and a distinctive flora are frequent. But the most striking feature in Andean páramo is the abundant postlike *Espeletia* (Compositae) that have a single thick stem topped by a roseate of furry leaves. Afroalpine páramo has giant *Lobelia* (Lobeliaceae) and giant *Senecio* (Compositae) that are very similar in growth form to *Espeletia*. The thick trunk is fire-resistant and the roseate of leaves may offer some protection from nightly freezing temperatures.

See also: Amazon Ecosystems. Biodiversity-Rich Countries. Hotspots. Mangrove Ecosystems. Rainforest Ecosystems, Animal Diversity. Rainforest Ecosystems, Plant Diversity. Rainforest Loss and Change

References

Bermingham E, Dick C, and Moritz C (eds.) (2005) *Tropical Rainforests: Past, Present and Future*. Chicago: University of Chicago Press.

Bullock S, Mooney H, and Medina E (eds.) (1995) *Seasonally Dry Tropical Forests*. New York: Cambridge University Press.

Churchill S, Balslev H, Forero E, and Luteyn J (eds.) (1995) *Biodiversity and Conservation of Neotropical Montane Forests*. Bronx, NY: New York Botanical Garden.

Gentry A (ed.) (1990) *Four Neotropical Forests*. New Haven, CT: Yale University Press.

Hamilton L, Juvik J, and Scatena F (eds.) (1995) *Tropical Montane Cloud Forests (Ecological Studies)*, vol. 110. New York: Springer-Verlag.

Hartshorn G (2000) Tropical and subtropical vegetation of Meso-America. In: Barbour M and Billings D (eds.) *North American Terrestrial Vegetation*, 2nd edn. New York: Columbia University Press.

Holdridge L, Grenke W, Hatheway W, Liang T, and Tosi Jr. J (1971) *Forest Environments in Tropical Life Zones: A Pilot Study*. San Francisco: Pergamon.

Kellman M and Tackaberry R (1997) *Tropical Environments: The Functioning and Management of Tropical Ecosystems*. New York: Routledge.

Laurance W and Peres C (eds.) (2006) *Emerging Threats to Tropical Forests*. Chicago: University of Chicago Press.

Leigh Jr. E, Rand S, and Windsor D (eds.) (1982) *The Ecology of a Tropical Forest: Seasonal Rhythms and Long-Term Changes*. Washington, DC: Smithsonian Institution Press.

Lieth H and Werger M (eds.) (1989) *Tropical Rain Forest Ecosystems (Ecosystems of the World)*, vol. 14B. Amsterdam: Elsevier.

Lugo A and Lowe C (eds.) (1995) *Tropical Forests: Management and Ecology*. New York: Springer-Verlag.

McDade L, Bawa K, Hespenheide H, and Hartshorn G (eds.) (1994) *La Selva: Ecology and Natural History of a Neotropical Rainforest*. Chicago: University of Chicago Press.

Richards P (1996) *The Tropical Rainforest: An Ecological Study*, 2nd edn. New York: Cambridge University Press.

Richards Y and Phillips O (2005) *Tropical Forests and Global Atmospheric Change*. Oxford, New York.

White F (1983) *The Vegetation of Africa*. Paris: UNESCO.

Whitmore T (1984) *Tropical Rain Forests of the Far East*, 2nd edn. New York: Oxford University Press.

Tropical Forest Regeneration

Robin L Chazdon, University of Connecticut, Storrs, CT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Chronosequence A series of sites that differ in age, but otherwise occur on similar soil types and environmental conditions within the same climatic zone.

Forest regeneration The process of forest regrowth on land that was formerly forested.

Old-growth forest A forest undergoing a late stage of succession that is relatively stable with regard to forest structure, ecosystem function, and species composition.

Primary succession The development of an ecosystem on a substrate previously devoid of vegetation.

Secondary succession The re-establishment of an ecosystem following a disturbance that removes existing vegetation.

Shifting cultivation A cyclic agricultural system widely used throughout the tropics involving cutting and burning of small patches of forest, followed by several years of crop cultivation, followed by a fallow period, during which forest regenerates and can be cleared again.

Introduction: Deforestation and Forest Regeneration in the Tropics

The twentieth century was a period of unprecedented deforestation in the world's tropical regions. There are hopeful signs that these rates are slowly declining due to increasing rates of reforestation (replanting) and natural forest regeneration. The rates of tropical deforestation declined from 16 million ha yr^{-1} during the 1990s to 13 million ha yr^{-1} from 2000 to 2010 (FAO, 2010). In Africa and Asia, less than 20% of the forests in 2010 were classified as primary (old-growth) forests, with no visible indications of human disturbance. Some tropical countries, such as Vietnam, Costa Rica, El Salvador, Puerto Rico, and India, have shown a net gain of forest cover over the past 10–20 years (Chazdon, 2008a). Although forest regeneration in the tropics cannot replace lost old-growth forests, regrowth forests offer suitable habitats for many forest-requiring species (Chazdon *et al.*, 2009b; Dent and Wright, 2009) and represent major sinks for atmospheric carbon, partially offsetting global carbon emissions (Pan *et al.*, 2011). Moreover, regenerating tropical forests provide essential sources of timber and nontimber products that support millions of forest-dependent people.

The small spatial scale of land-use change and the difficulty of distinguishing early stages of secondary growth from plantations or short-term fallows create significant challenges for assessing the area of tropical forests undergoing natural regeneration. Asner *et al.* (2009) classified at least 1.2% of the humid tropical forests around the world as forests undergoing long-term secondary regeneration at the end of the twentieth century. This highly conservative estimate does not include forests committed to regeneration before the 1990s and is based on selected areas with time series of satellite-based data. The increased prominence of tropical regrowth forests around the world signals an urgent need to understand the underlying biophysical and social factors that influence their regeneration following agricultural abandonment and natural disturbance.

Tropical forests undergo periods of disturbance and recovery at different spatial and temporal scales (Chazdon, 2003). Natural disturbances, such as hurricanes, floods, and fires, completely or partially remove forest cover and alter soils, with dramatic consequences for biodiversity and ecosystem functions (Whitmore and Burslem, 1988). Forest clearing for cultivation, establishment of pastures, or harvesting timber creates intensive and extensive disturbances. In the aftermath of large-scale disturbances, successional processes bring about changes in the species that form forest communities, species population sizes and structure, species interactions, and ecosystem properties. Successional pathways and rates of change vary widely according to the nature of prior land use, the proximity of old-growth forest, and the availability of fauna (Chazdon *et al.*, 2007).

Forest regeneration is a community- and ecosystem-level process of secondary succession on cleared land that was formerly forested. The successional process follows a progression of stages during which forest assemblages gradually become enriched with species and increase in structural and functional complexity. Old fields that replaced abandoned clearings, pastures, or cultivated fields, turn into young, regenerating forests dominated by fast-growing, highly dispersed pioneer tree species. Over the time, plant and animal species characteristic of nearby old-growth forests gradually replace successional pioneer species. Depending on the longevity of early-colonizing tree species, old-growth forests can reestablish within 100–200 years (Wirth *et al.*, 2009). But succession can become arrested or deflected following intensive and large-scale land uses that deplete soil fertility or reduce the availability of local biota. In these cases, restoration interventions are needed to assist in the establishment of native tree species.

Tropical Forest Succession Following Natural and Human Disturbances

Historical Legacies of Tropical Forest Disturbance

Today's tropical forests represent the legacy of successive waves of colonization, exploitation, cultivation, abandonment, and regrowth, shaped by human occupations, cultural change, natural disasters, and climate change. Hurricanes and fires also occurred repeatedly during prehistory, influencing the

composition and regeneration of tropical forests. Following a prehistoric hurricane at Laguna Negra in southern Nicaragua between 3830 and 2820 years ago, a series of fires occurred over approximately 200 years, followed by repeated cycles of forest regeneration and burning by fires over several thousand years (Urquhart, 2009). Volcanic eruptions destroyed former forests and gave rise to new ones in New Guinea, Sumatra, and Krakatoa. In Amazonian Ecuador, a catastrophic flood wiped out an entire forest area and its human inhabitants more than 500 years ago. The forest has been regrowing since then, but species composition remains distinct from nearby forests that were spared in the historic flood (Pitman *et al.*, 2005).

The legacies of ancient occupations by hunter-gatherers and later periods of extensive and intensive landscape transformation by agricultural societies are still evident in today's tropical forests. Early hunter-gatherers significantly altered forest cover, led to the geographic spread of key food plants, and contributed to the extinction of some of the Pleistocene megafauna. Forest dwellers practiced early cultivation of forest tubers and arboriculture of fruit trees, increasing the concentration, abundance, and geographic ranges of plant species used for food and shelter in their camps, settlements, and rock shelters. Late Pleistocene colonists in the New Guinea Highlands fashioned stone tools to clear trees, enhancing the supply of *Pandanus* seeds, an important staple food (Summerhayes *et al.*, 2010). These early occupations are associated with increasing frequency of fire and changes in hydrology. In the Neotropics, Paleoindians hunted, fished, and gathered fruits along rich alluvial floodplains and bluffs of major rivers and along coastlines. Human populations extended into upland areas, where they engaged in broad-spectrum hunting and collecting in areas away from streams and practiced small-scale gardening.

Indigenous peoples of Mesoamerica cultivated trees and managed forest patches for more than 3000 years. The Maya planted homegardens and managed forest fallows and older forests. Forest gardens were so widespread during the Mayan Preclassic period that contemporary forests of the southeastern Peten, eastern Guatemala, and western Belize are considered to be anthropogenic in origin (Campbell *et al.*, 2006). Aggregations of useful tree species in forests surrounding archeological sites today provide strong evidence of silvicultural management by ancient peoples (Ross, 2011). The persistence and management of tree cover in Mayan landscapes likely favored the rapid regeneration of forest following the Maya collapse several hundred years later after the Spanish Conquest. In the Mirador Basin of the northern Peten region of Guatemala, sediments of Lago Puerto Arturo reveal a history of forest clearing beginning more than 4000 years ago, followed by 2000 years of agricultural activity, and sudden abandonment around AD 1000, at the end of the Late Classic Period (Wahl *et al.*, 2006). Legacies of ancient arboriculture systems are still evident in the current vegetation of the Solomon Islands and throughout Melanesia and Oceania.

The development and spread of agriculture profoundly impacted the structure and composition of tropical forests. Human activity over more than hundreds to several thousand years also produced changes in soil properties and landscape hydrology across the Brazilian Amazon basin and in regions of Columbia, Ecuador, Peru, Venezuela, Bolivia, the Guianas, and West Africa (Woods, 2009). In the depths of the Darien forests of Panama, pollen cores document nearly 4000 years of land clearing and maize cultivation (Bush and Colinvaux, 1994). These areas were abandoned abruptly about 320 years ago, shortly after the Spanish conquest, and have not been resettled. The forests surrounding these areas are now in a late stage of secondary succession. As canopy trees often live greater than 300 years, these forests – and many other so-called "primary" forests in Mesoamerica – are still undergoing gradual successional changes in composition and stand structure, as current canopy trees may represent only the second generation of trees since this forest began to regenerate (Bush and Colinvaux, 1994).

Forest Dynamics and the Successional Continuum

Nonequilibrium theories have displaced equilibrium views, shedding new light on the nature of successional change and responses to disturbance. There is no particular moment when a forest reaches a stable or a "climax" state, as disturbances frequently occur, even during late stages of succession (Chazdon, 2008b). Characterization of species as "early-successional" versus "late-successional" often involves arbitrary distinctions based on the nature of the forests where different species colonize, recruit, and reproduce. Replicated landscape-level surveys of species abundance in second-growth and old-growth forests can be used to classify species according to their specialization in different forest types using rigorous statistical approaches.

Forest successional stages can be defined based on three major criteria: total aboveground biomass, age, or size structure of tree populations, and species composition. These characteristics change at different rates and vary differentially with the spatial scale of measurement. Multiple successional pathways can be observed within a particular tropical region, often reflecting differences in prior land use (Mesquita *et al.*, 2001). Most methods used to detect forest disturbance or change are based on measures of forest structure, such as basal area, which tend to approach stability more quickly than measures of species composition (Letcher and Chazdon, 2009). The only way to distinguish old secondary forests from forests that have not been disturbed by humans or catastrophic forces of nature is through detailed knowledge of species growth and regeneration patterns and careful studies of past stand history.

Successional processes occur in all forests at different spatial and temporal scales. The term "old-growth forest" describes forests at a late stage of succession that are relatively stable with regard to these criteria. Forest dynamics do not stop when a forest has reached a late stage of succession, but they shift toward localized (endogenous) disturbances such as tree-falls or local flooding that do not uniformly characterize the entire forest assemblage.

Successional Trajectories and Stages

Successional trajectories are influenced by the scale, frequency, and intensity of previous disturbance or land use, by soil

texture and nutrient availability, by the nature of remnant vegetation, and by conditions following disturbance, including types of management, colonization of invasive species, and dispersal of seeds from surrounding forest areas (Chazdon, 2003, 2008b). Mesquita *et al.* (2001) describe two distinct successional pathways on abandoned agricultural land in Central Amazonia. In sites where the period of intensive land use was more than 4 years and fire was used during the initial clearance of forest, species of *Vismia* dominate early stages of succession. In contrast, *Cecropia* species dominate in areas where pastures were used for less than 2 years in both burned and unburned sites. Floristic turnover in *Cecropia* stands showed strong and consistent negative frequency dependence, whereas *Vismia* stands exhibited little or no frequency dependence (Norden *et al.*, 2011). Patterns of early species colonization and dominance following land abandonment strongly affect successional changes in vegetation structure and composition (Chazdon, 2008b).

Dividing successional pathways into distinct stages or phases is a useful approach that enables comparative studies and examination of ecological processes that affect transitions in the forest structure, composition, and ecosystem properties (Chazdon, 2008b, **Table 1**). The boundaries between successive successional stages are fuzzy, but the temporal sequence of these stages generally follows consistent patterns. **Table 1** illustrates the vegetation dynamics processes that occur during successive stages of tropical forest succession. During the stand initiation stage, pioneer trees colonize and establish, herbaceous vegetation declines, and shade-tolerant tree species begin to establish as seedlings. Canopy closure ushers in the stem exclusion stage, when shade-intolerant species of trees, lianas, and shrubs below the canopy become suppressed and die, whereas shade-tolerant species recruit in the subcanopy and canopy and continue to establish as seedlings. During the long understory-reinitiating stage, long-lived pioneer trees in the canopy begin to die, forming canopy gaps, which increase heterogeneity in light availability. This stage is associated with turnover of species in the canopy, whereas the earlier stem exclusion stage is associated with turnover of species in the understory. The old-growth stage begins with the mortality of the last remaining pioneer trees that established during the stem initiation stage (Wirth *et al.*, 2009). This stage is the most diverse for trees and epiphytes and is characterized by high spatial heterogeneity and high functional diversity (**Table 1**).

Chronosequences versus Dynamics

Our current understanding of secondary forest succession in the tropics emerges primarily from chronosequence studies (Brown and Lugo, 1990). A chronosequence is a series of sites that differ in age since abandonment or disturbance, but otherwise occur on similar soil types and environmental conditions within the same climatic zone under the same historic land use. Sites are replicated in space as a substitute for replication in time. Chronosequence studies rely on the

Table 1 Vegetation dynamics processes associated with stages of secondary succession and time since disturbance in tropical forests

Time since disturbance/successional stage			
0–15 years *Stand initiation*	*15–50 years* *Stem exclusion*	*30–200 years* *Understory reinitiation*	*> 200 years* *Old-growth*
Germination of seeds in seed-bank	Canopy closure	Mortality of canopy trees	Mortality of pioneer cohort in canopy
Resprouting of remnant trees	High mortality of lianas and shrubs	Formation of small canopy gaps	Range of gap sizes
Colonization of short- and long-lived pioneer trees	Recruitment of shade-tolerant seedlings, saplings, and trees	Canopy recruitment and reproductive maturity of early-colonizing species	Recruitment of shade-tolerant and gap-requiring canopy species and emergents
Rapid height and diameter growth of woody species	Growth suppression of shade-intolerant trees in understory and subcanopy	Increased heterogeneity in understory light availability	Spatial heterogeneity in biomass and microtopography
High mortality of herbaceous species	High mortality of short-lived, pioneer trees	Seedling and sapling establishment of shade-tolerant tree species	Large woody debris
High rates of seed predation	Dominance of long-lived pioneer trees	Tree recruitment of early-establishing shade-tolerant species	Maximum diversification of trees and epiphytes
Seedling establishment of shade-tolerant tree species	Development of canopy and understory tree strata		
	Seedling establishment of shade-tolerant tree species		

Successional stages are based on the framework Oliver and Larson (1990). Actual time ranges vary widely among tropical forest regions and land uses.

critical assumption that fields abandoned at different times in the past experienced the same changes, processes, and conditions during their formation. This space-for-time substitution allows the study of a wide range of successional stages and makes it possible to sample considerably more sites than could be accommodated through site-intensive long-term studies. Chronosequence approaches are used to study primary succession in the tropics following landslides, volcanic eruptions, and river channel flooding.

Chronosequence studies require detailed knowledge of land-use history and time since disturbance or since abandonment of each stand. In seasonal tropical forests, trees with annual growth rings can often be used to date forest age or past fire events. Few tropical forest chronosequences include sites more than 80–100 years old. Well-replicated samples from different forest patches within each forest age class are needed for robust chronosequence studies. Ideally, study sites should be selected without systematic bias in forest attributes or land-use history, but this is not always feasible. Increasing intensification of land-use and landscape change in the tropics over the past 150 years creates challenges for chronosequence studies.

Long-term studies of vegetation dynamics in successional forests often reveal idiosyncratic patterns in individual sites driven by initial species composition, site factors, land-use history, and landscape composition. With few exceptions, these patterns are not consistent with chronosequence predictions. Chronosequence data for basal area, stem density, and species density of trees >10 cm diameter were compared with time-series data over 9 years on vegetation dynamics in eight successional stands in northeastern Costa Rica (Chazdon *et al.*, 2007). Only basal area showed a consistent trend between the chronosequence and the time series. A similar analysis compared chronosequence data with 3-year time series data for 10 successional sites in Chiapas, Mexico, regenerating on shifting cultivation fallows (van Breugel *et al.*, 2006). In this case, fallow age was not a significant predictor of either plot basal area or stem density across the chronosequence (Chazdon *et al.*, 2007). Feldpausch *et al.* (2007) tested chronosequence predictions of secondary forest growth and biomass accumulation with 4 years of time-series data in 10 secondary forests growing on abandoned pastures in Central Amazonia. The youngest forests consistently accumulated less biomass per year than the predicted trend due to a low initial density of trees following a longer duration in pasture (Feldpausch *et al.*, 2007).

When sites within a chronosequence share similar land-use history, environments, and species pools, patterns of successional change are more consistent within and across plots of different age since abandonment. In tropical dry forest of Oaxaca, Mexico Lebrija-Trejos *et al.* (2010) observed successional changes over 4 years within 16 plots ranging from 1 to 60 years old that generally matched chronosequence trends, particularly when pioneer and nonpioneer species were considered separately. Chronosequences and long-term, plot-based studies should be viewed as complementary approaches to the study of successional rates and processes. Besides the difficulty in matching land-use history and environmental conditions across sites, chronosequence studies do not provide direct information on successional processes or vegetation dynamics. The only way to quantify these processes is to monitor changes in forest structure, composition, and function overtime within sites of known disturbance history and landscape surroundings.

Sources of Regeneration

Following abandonment of cultivated fields or pastures or after a forest-clearing disturbance, new vegetation recruits arise from five major sources: the soil seed bank, newly dispersed seed, resprouting damaged stems, and advance regeneration – the local pool of existing seedlings and saplings that survived the disturbance. Many mature forest species are capable of resprouting from roots or stems after forest clearing and burning (Kammesheidt, 1999). The vigorous growth of sprouts confers a competitive advantage during the early stages of forest succession. In abandoned shifting cultivation fallows in Eastern Amazonia, 81–86% of the individuals and 68–81% of trees originated from sprouts (Vieira and Proctor, 2007). The density of resprouts in shifting cultivation fallows in a tropical moist forest region of Paraguay declined with fallow age (Kammesheidt, 1999). Resprouting is also an important mechanism for tree regeneration following logging and cyclones. As land-use intensity increases, as in frequent burning, weeding, and bulldozing of pastures, sprouting is no longer a source of new recruits following abandonment. Resprouting is a particularly important mechanism of recruitment in tropical dry forests, where the rates of seedling establishment may be reduced by desiccation and plant species are adapted to aboveground dieback during droughts.

In large abandoned pastures and swidden fallows, the main seed sources for initial colonization during secondary succession come from either the soil seed bank or newly dispersed seed. Many pioneer trees with small seeds in the seed bank require direct exposures of red light for germination (Vazquez-Yanes and Orozco-Segovia, 1993). In abandoned pastures of Chiapas, Mexico, the seed bank was more important than the seed rain for the recruitment of early-successional tree species, but the opposite was true for late-successional species (Benitez-Malvido *et al.*, 2001). In most cases, seeds of late-successional species are rarely found within the soil seed bank (Vazquez-Yanes and Orozco-Segovia, 1993).

The soil seed bank is severely depleted in tree and shrub species following many years of grazing and burning to establish and maintain pasture or plant crops, placing an even greater reliance on the arrival of seeds transported by wind or delivered by frugivorous vertebrates. The rate of new seed dispersal decreases with clearing size and distance from forest edge, retarding the initiation of secondary succession in large clearings (Cubina and Aide, 2001). However, if birds and bats are attracted to perching sites on fruiting shrubs or trees within pastures or fallows, the rate of succession accelerates through increased quantity and diversity of seed dispersal and enhanced recruitment of woody species, including species characteristic of mature forests. Seedling establishment of vertebrate-dispersed species during initial stages of succession is enhanced beneath remnant trees compared with areas away from remnants within the same site (Guevara *et al.*, 1986).

Besides their role in nucleating succession, remnant trees can facilitate forest regeneration in at least three other ways. First, isolated trees serve as sources of seed for tree regeneration beneath other isolated remnant trees and as pollen and seed sources for regenerating trees in forest fragments and secondary forests. Second, isolated pasture trees store seed of forest tree species in canopy soil that accumulates within their crowns (Nadkarni and Haber, 2009). Third, isolated remnant trees in pastures continue to influence tree recruitment beneath their crowns beyond the initial stages of succession, even after canopy closure (Schlawin and Zahawi, 2008).

Effects of Prior Land Use and Soil Fertility

The intensity, extent, severity, and duration of land use have direct consequences on sources of regeneration, soil nutrient availability, and conditions for early seedling establishment, including competition with crop species and weeds. When initial stand disturbances are small, land use is of a short duration, and clearings are embedded in a forest matrix, tree species and biomass accumulate rapidly. Intensive land use, short fallow times, and frequent burning reduce soil nutrient availability, increase soil compaction, reduce soil water-holding capacity, and restrict seed availability and seedling establishment. Exotic pasture grasses can also negatively affect the establishment of woody plants. During repeated cycles of shifting cultivation in Kalimantan, Indonesia, the species diversity of both early- and late-successional species declined with increasing numbers of cycles, due to changes in the surrounding landscape that affected local species availability (Lawrence, 2004).

Secondary succession generally proceeds more rapidly on young, fertile, volcanic soils, compared with weathered, nutrient-poor soils (Guariguata and Ostertag 2001). In shifting cultivation fallows of white-sand vegetation in the Venezuelan Amazon, saplings of mature forest tree species took 60 years to become more abundant than early-successional species (Saldarriaga *et al.*, 1988). In contrast, sapling assemblages in 15–20-year-old secondary forests on abandoned pastures on volcanic soils in northeastern Costa Rica were already becoming dominated by mature forest species (Norden *et al.*, 2009). Across five regions of the Amazon Basin, the rates of forest regrowth, as measured by stand height, were the highest in the Altamira region, which had the highest soil fertility (Moran *et al.*, 2000). In the Bragantina region of Brazil, 150 years of land clearance for agriculture has led to slow and impoverished forest regrowth due to low seed availability and pool soil quality (Tucker *et al.*, 1998).

Successional forests across five regions of the Amazon basin showed faster rates of height growth in shifting cultivation fallows than in either abandoned pastures or in fields used for mechanized agriculture, and shifting cultivation fallows generally had higher rates of aboveground biomass accumulation than more intensive land uses (Moran *et al.*, 2000). Expanding this study to nine locations, Zarin *et al.* (2005) found that the rates of biomass accumulation in secondary forests stands across the Amazon basin were strongly affected by the number of previous burns, which often serves as a proxy for the number of shifting cultivation cycles. On average, stands with five or more previous burns showed a reduction in biomass accumulation of more than 50% compared with stands that burned only once or twice (Zarin *et al.*, 2005). Repeated cycles of shifting cultivation can also reduce the levels of available soil nutrients during fallow periods. After three cultivation–fallow cycles in tropical dry forests of the Yucatan Peninsula in Mexico, the levels of available phosphorus declined by 44%, creating a negative feedback cycle of increasing phosphorus limitation in regenerating vegetation (Lawrence *et al.*, 2007).

Recovery of Ecosystem Processes

When a tropical forest is cleared, burned, cultivated, or grazed, a substantial fraction of the living biomass is lost from the ecosystem. The fate of carbon and nutrients in the soil following forest clearing varies widely depending on the type and the intensity of land use, as well as soil texture and clay content. In a humid tropical region of Mexico, conversion of forests to pastures or cornfields led to the loss of 95% of the aboveground pool of carbon, 91% of the aboveground pool of nitrogen, and 83% of the aboveground pool of phosphorous (Hughes *et al.*, 2000). As pastures are generally not tilled or cultivated, little or no carbon is lost in the top 100 cm of soil. Conversion of forest to active pasture in northeastern Costa Rica did not lead to significant loss of soil carbon (Powers, 2004). Conversion of tropical old-growth forests to cropland resulted in a 25% loss of soil organic carbon, twice as much as when old-growth forests are converted to pasture (Don *et al.*, 2011).

The losses of carbon and nutrients following forest clearing and burning are gradually reversed during forest regeneration. Lost soil organic matter is restored due to the decomposition of litterfall and root biomass production. Regenerating forests rapidly accumulate nutrients in vegetation, litter, and soil, particularly during the first 20 years (Brown and Lugo, 1990). The rapid development of leaf biomass contributes to litterfall and to the rapid accumulation of total biomass during the stand initiation stage (**Table 1**). Woody biomass (stems and roots) accumulates over a longer time frame, as trees continue to grow in height and diameter over several decades (Brown and Lugo, 1990).

Approximately 45–50% of this biomass is composed of carbon. During the first 20 years of succession, the rates of aboveground carbon accumulation are generally higher in moist and humid forests than in dry forests, and on former agricultural land than former pastures (Marin-Spiotta *et al.*, 2008). Regenerating forests in the tropics are major sinks for carbon; the global carbon sink formed by tropical regrowth forests was estimated to reach 1.7 ± 0.5 PgC years from 2000 to 2007, greater than the carbon sink from intact tropical forests (Pan *et al.*, 2011).

Chronosequence studies have been widely used to estimate the rate of accumulation of carbon and nutrients in above- and belowground stocks during secondary forest regeneration. Based on a chronosequence study in the Venezuelan Amazon, Saldarriaga *et al.* (1988) estimated that forests regrowing following shifting cultivation require 190 years to reach levels of aboveground biomass observed in old-growth forests. In tropical dry forests of the Yucatan Peninsula in Mexico, aboveground biomass was estimated to reach old-growth forest levels within 65–120 years of succession (Read and

Lawrence, 2003). Chronosequence studies in humid Neotropical forests have found higher levels of aboveground biomass in older secondary forests than in old-growth forests (Letcher and Chazdon, 2009). In these cases, the higher aboveground biomass in older secondary forests could be due to the lower biomass of subcanopy and canopy palms, which are more abundant in old-growth forests. In contrast to aboveground carbon stocks, soil carbon stocks do not change appreciably over time during forest regeneration on former pastures (Marin-Spiotta *et al.*, 2009). Aboveground pools of nitrogen, phosphorus, and sulfur increased with age during forest regeneration in humid forests of Los Tuxtlas in Mexico (Hughes *et al.*, 1999) and in Central Amazonia (Feldpausch *et al.*, 2004). Fixation of nitrogen by legumes during early succession allows the rapid recovery of nitrogen levels in shifting cultivation fallows in Amazonia (Gehring *et al.*, 2005).

Biodiversity and Tropical Forest Regeneration

Functional Diversity of Vegetation

Tropical forest regeneration follows a general theme of replacement of shade-intolerant, fast-growing species with shade-tolerant, slower-growing species. These changes bring about transitions in the dominance of different plant growth forms and functional types. Grasses, herbs, vines, and shrubs dominate recently abandoned fields, but decline in abundance as the forest canopy closes and reduces light availability (Chazdon, 2008b). Much of this replacement takes place during the stand initiation and stem exclusion stages beneath the developing canopy of long-lived pioneer trees, which themselves become replaced by both shade-tolerant and gap-requiring tree species only during later stages of succession. Pioneer and shade-tolerant species in tropical forests exhibit contrasting values of many leaf, wood, and seed traits that are important determinants of plant growth and survival (Bazzaz and Pickett, 1980). Poorter *et al.* (2004) ranked 15 tree species from lowland moist forest in Bolivia according to their successional position, based on abundance patterns during succession following shifting cultivation. Six leaf traits showed significant linear relationships with species successional position. Specific leaf area, leaf water content, leaf nitrogen content, and leaf phosphorus content all declined from the early- to the late-successional position, whereas the leaf carbon : nitrogen ratio and lignin content increased. Leaf lifespan also increased in a linear manner from early- to late-successional position, whereas herbivory rates declined (Poorter *et al.*, 2004). Wood density of plant stems and branches also shows successional trends; fast-growing species that dominate early in succession generally have low wood density, whereas slow-growing species that dominate later in succession generally have high wood density (Poorter, 2008).

Changes in Species Composition

Tropical old-growth forests are among the most species-rich terrestrial assemblages on earth. How species composition changes over long time scales during succession remains poorly understood, as few studies have followed such changes for more than a few years. Compositional changes are driven by complex interactions among local site factors, past site history, and by dynamic landscape and regional factors. The structural, functional, and taxonomic diversity of vegetation increases during succession. The analysis of successional trends in species composition is a challenging prospect for at least five reasons. First, almost all of our knowledge on changes in species composition during succession is based on short-term studies (less than 50 years in duration) or on chronosequence studies. Comparisons of species richness or diversity across a chronosequence are insufficient to assess similarity of assemblages, due to the high rates of species turnover during succession and site-specific trajectories of species composition (Chazdon *et al.*, 2007). Second, assessments of the long-term recovery of a particular species during succession assume that no changes in species composition or population abundance have occurred over this same time interval in old-growth areas that serve as reference points. Third, the classification of old-growth specialists in a particular region or landscape is rarely based on robust statistical procedures and is highly subjective. Fourth, in many regions, old-growth forests used as reference points and for classifying specialists do not occur within the same elevational zone or soil type as second-growth forests, owing to the landscape variation in deforestation and land abandonment. Thus, spatial differences in composition may be confounded with successional changes. Fifth, secondary forests are often cleared before there is an opportunity to assess long-term successional trajectories and species gain in these forests (Chazdon *et al.*, 2009b).

In humid forest regions, woody species composition recovers slowly but steadily in abandoned pastures and shifting cultivation fallows (Guariguata and Ostertag, 2001; Chazdon, 2008b). In the Southern Atlantic Forest region of Brazil, only 3.1% of the original forest cover remains. Piotto *et al.* (2009) inventoried trees >5 cm DBH (diameter at breast height) in 12 secondary forests in three age classes (10, 25, and 40 years) following shifting cultivation of manioc. The similarity of species composition between second-growth forests and old-growth forests in the region increased with stand age, due to the gradual establishment of old-growth tree species (Piotto *et al.*, 2009). The proportion of trees classified as pioneer species declined steadily with forest age class, and more than half of the species found in 40-year-old stands were shared with neighboring old-growth forests. The percent of known local endemics and Atlantic forest endemics also increased with age across these stands, whereas widespread species decreased.

Across a chronosequence in tropical rain forests of lowland NE Costa Rica, species composition and species richness of stems >2.5 cm DBH did not differ significantly between older regenerating forests (30–42 years old) and old-growth forests (Letcher and Chazdon, 2009). The proportion of old-growth species present in second-growth forests increased significantly with age since abandonment (**Figure 1**), suggesting that the establishment of old-growth species increases steadily with time during at least the first 50 years of succession in this region. In seedling assemblages of 15–25-year-old regenerating forests, the proportion of old-growth species increased to 59–75%, reflecting conditions favorable for seedling recruitment (Norden *et al.*, 2009).

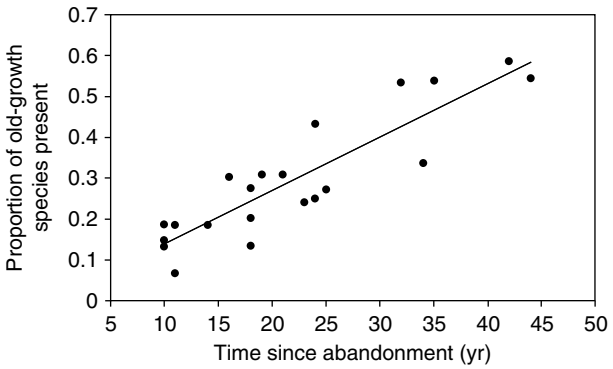


Figure 1 The proportion of old-growth woody species occurring in secondary forest assemblages increases significantly with time since pasture abandonment in regenerating lowland rain forests of northeastern Costa Rica (Letcher and Chazdon (2009); $R^2 = 0.79$; $P < .001$). Data are based on 0.10 ha vegetation samples of stems > 2.5 cm DBH in 21 secondary forest stands.

Few studies have assessed the long-term recovery of tree species composition and endemism of trees during secondary succession. Regenerating forests on former coffee plantations in the Blue and Port Royal Mountains of Jamaica did not differ significantly from adjacent old-growth forests in basal area, the overall number of species, or the number of endemic species (Chai and Tanner, 2011). Endemic species were less abundant in the second-growth forests, however, despite 150 years of succession. More than half (54%) of the species present in nearby old-growth forest were present in the regrowing plantations, suggesting that old-growth tree species are gradually establishing in these forests (Chai and Tanner, 2011).

Recovery of animal populations and species composition during succession depends heavily on recovery of vegetation (Chazdon *et al.*, 2009b). Regenerating forests provide different types and amounts of resources (food and shelter) for animals at different stages. Based on a review of 10 studies, Karthik *et al.* (2010) reported that 70% of old-growth forest birds had returned within 25 years of abandonment of shifting cultivation fields. Bird species diversity and composition in 14–19-year-old second-growth forests of Jari, Brazil, are distinct from surrounding old-growth forests however, suggesting that many years are needed for establishment of old-growth specialist bird species (Barlow *et al.*, 2007b). Second-growth forests supported a higher abundance of ungulate browsers and small-bodied primates than old-growth forests but lower abundance of large-bodied birds and large primates (Parry *et al.*, 2007). Studies of 15 taxonomic groups in this region showed idiosyncratic responses to land-use change; 95% of old-growth orchid bee species occur in second-growth, but less than 40% of old-growth tree and liana genera were found in second growth (Barlow *et al.*, 2007a). Volant invertebrates and vertebrates (butterflies, birds, and bats) generally show faster rates of influx into regenerating forests than do nonvolant animals (Chazdon *et al.*, 2009b). Of the 30 species of nonvolant mammals detected in old-growth forests in the humid forest region of Los Tuxtlas, Mexico, only 11 were detected in 25–35-year-old forests and eight in 5–15-year-old forests (Estrada *et al.*, 1994).

Dent and Wright (2009) examined 65 studies that compared faunal diversity (birds, reptiles, amphibians, invertebrates, and mammals) in tropical old-growth and forests regenerating following complete clearance. Of the 114 regenerating sites, only 11 were more than 40 years old. On an average, the mean proportion of old-growth animal species present in regenerating forests was 58%. More than 50% of species present in regenerating forests less than 10 years old were also present in old-growth. Moreover, the composition similarity (presence–absence-based Sørensen Index) between old-growth and young regenerating forests was similar to levels of similarity evaluated exclusively among old-growth sites (Dent and Wright, 2009). Among the 22 studies that included regenerating forests of different ages, faunal similarity between regenerating and old-growth forest increased significantly with increasing time since abandonment. This increase in similarity is due to an increasing percent of old-growth animal species that are present in regenerating forests over time, reaching 80% in forests older than 50 years (Dent and Wright, 2009). Species composition of animal communities in forests regenerating following intensive agriculture or pasture was less similar to animal communities in old-growth forests than those recovering from shifting cultivation or that were cleared without cultivation. Regenerating forest adjacent to old-growth forests had higher levels of faunal similarity to old-growth forest than forests that were not adjacent (Dent and Wright, 2009).

It is important to recognize that regenerating tropical forests cannot support all species that lived in local forests before they were cleared (Dent and Wright, 2009). Species with highly specialized dietary or habitat requirements are often missing from regenerating forests. Large-bodied canopy frugivores and understory insectivores are adversely affected by habitat disturbance and are often restricted to old-growth forests (Karthik *et al.*, 2010). Furthermore, patches of regenerating forests are often small, isolated, and close to roads, further reducing their occupancy by rare species and species that cannot cross unforested habitat (Lees and Peres, 2009). Hunting pressures are also very high in regenerating forests, because they are often close to agricultural fields and human communities (Karthik *et al.*, 2010).

Seed Dispersal and Predation

The lack of seed dispersal to abandoned areas can be a major obstacle to forest regrowth. For 50–90% of the tree and shrub species in tropical forests, dispersal is assisted by animals (Howe and Smallwood, 1982). But most frugivorous animals will not venture into abandoned areas unless they encounter suitable resources for feeding or perching. Open conditions in clearings expose frugivores to increased predation risk and offer a lower availability of fleshy fruits. These factors lead to a steep decline in seed rain with increasing distance from the forest edge (Duncan and Chapman, 2002). Isolated remnant trees, shrubs, and pioneer trees in abandoned pastures and cultivated fields attract frugivores, leading to higher densities of seed deposition and recruited seedlings (Laborde *et al.*, 2008).

Bats and birds are the major biotic dispersal agents for early-successional trees and shrubs in the tropics. In humid

forest areas of Chiapas and Mexico, bats dispersed more seeds than birds did in early-successional habitats (Medellin and Gaona, 1999). Birds tend to deposit seeds beneath fruiting trees or carry fruit to perches or nests, where they drop or defecate seeds (Corlett, 1998). Although bats also deposit seeds below fruiting trees, they also defecate small seeds during flight. Bat-dispersed seeds are therefore more likely to reach abandoned areas that lack perching or nesting sites than bird-dispersed seeds. Phyllostomid bats play a particularly important role in seed dispersal during both primary and secondary succession (Muscarella and Fleming, 2007). Species that often dominate early stages of succession in the Neotropics are concentrated in four of the five top plant families that frugivorous bats consume (Solanaceae, Moraceae, Piperaceae, and Clusiaceae; (Muscarella and Fleming, 2007). With the exception of Moraceae (*Ficus* and *Musanga* species), bat-dispersed plants rarely dominate early stages of succession in the Old World. Early-successional fleshy fruits in the tropical Africa and Asia are dispersed primarily by birds (Gonzales *et al.*, 2009).

Animal-mediated seed dispersal is poorly studied beyond the stand initiation stage. Early-successional forests are dominated by small-seeded species, whereas large-seeded species increase in importance during later successional stages (Westoby *et al.*, 2002). Frugivore body size is the major determinant of the ability to feed on fruit of a given size. Thus, large seeds (and the large fruits that contain them) are generally dispersed by large-bodied frugivores. One exception to this trend is the dispersal of large seeds (>8 mm diameter) by small-bodied tent-roosting bats, such as *Artibeus watsonii* in humid Neotropical forests (Melo *et al.*, 2009). Seed deposition beneath leaf tent roosts of this species significantly increased the species richness and abundance of large-seeded species in regenerating forests of northeastern Costa Rica (Melo *et al.*, 2009). Small-bodied primates can also be important dispersers of large seeds in secondary forests. Mixed-species groups of tamarins (*Saguinus mystax* and *Saguinus fuscicollis*) dispersed 63 tree species with seeds >1 cm diameter in a 9-year-old regenerating forest adjacent to old- growth forest in northeastern Peru (Culot *et al.*, 2010).

The spatial distribution of vertebrate-dispersed seedlings, saplings, and trees in regenerating forests may reflect a long-term legacy of remnant trees and their interactions with frugivorous vertebrates. In montane forests of Costa Rica, remnant trees influenced the composition of regeneration 3 and 14 years after abandonment; late-successional species were more abundant than early-successional species only beneath remnant trees. In the Caribbean lowlands of Costa Rica, tree and sapling density decreased with increasing distance from remnant trees 23 years after pasture abandonment (Schlawin and Zahawi, 2008). Remnant trees also influenced the species composition of recruited trees, favoring the establishment of species with seeds >1 cm diameter.

Dispersal agents vary in their importance during the early, middle, and late stages of forest regeneration (**Table 1**). Initially, if no remnant vegetation is present, wind and bats play particularly important roles in the dispersal of small-seeded species. Once perches become available, frugivorous birds become frequent visitors and increase the taxonomic diversity, size, and functional diversity of seeds arriving to regenerating forests. As regenerating vegetation becomes more developed and diverse, increasing availability of resources for roosting and feeding encourages a wide diversity of small- and large-bodied vertebrates to become regular visitors and residents in regenerating forests, carrying with them the fruits and seeds of old-growth species from the surrounding areas. The progression of vegetation during regeneration facilitates colonization of new species. During mid- and late-successional stages, the only barriers to seed dispersal are distance from seed sources of rare or infrequently fruiting trees, availability of dispersal agents, and time. Even after 100 years of regeneration, forests in the central catchment reserve of Singapore were depleted of many tree species above 30 cm DBH found in contiguous old-growth forests in the region (Turner *et al.*, 1997). Large-seeded old-growth tree species likely require large-bodied frugivores to disperse their seeds, but these species are now extinct or extremely rare in Singapore (Corlett, 1992).

Following dispersal, seeds are subject to attack by vertebrates, invertebrates, and fungi. The rates of seed predation are high in pastures, often exceeding 50% (Scariot *et al.*, 2008). Small seeds (<0.2 g) dispersed into pastures are consumed by ants and other insects, whereas medium-sized seeds (0.2–4.0 g) are often consumed by rodents (Myster, 2008). In regenerating forests, mammals appear to be more important seed predators than insects for large-seeded species such as palms. The rates of seed predation were higher in shifting agriculture fallows than in mature forest gaps (Uhl, 1987), and the removal rate of seeds declined with age since abandonment (Pena-Claros and de Boo, 2002).

Declines in vertebrate abundance due to hunting can decrease the levels of seed predation, leading to changes in the species composition of regenerating vegetation. Although these effects are poorly studied in successional forests, declines in populations of mammalian seed predators may lead to increased recruitment of large-seeded species or species that are less susceptible to insect predation (Stoner *et al.*, 2007).

Conclusions: The Broader Context of Tropical Forest Regeneration

Most regenerating forests in the tropics are found in relatively small patches within a matrix of old-growth forests, logged forests, and agricultural fields or pastures. Regeneration pathways and rates of change strongly depend on the surrounding landscape conditions, which are often highly dynamic. Moreover, landscape conditions are often related to the duration and intensity of prior land use that influence regeneration at a local level. Ultimately, forest regeneration is part of a broader socio-ecological system that governs patterns of land use and abandonment, the spatial configuration of remnant forests, conservation of wildlife, hunting, harvesting, and management of forest products (Grau and Aide, 2008).

The distribution of forest patches of different ages within a landscape is a key indicator of the potential for forest regeneration on recently abandoned land (Chazdon *et al.*, 2009b). Young second-growth forests have higher rates of clearance than older second-growth forests or old-growth forests (Etter *et al.*, 2005). Many factors affect the persistence

of regenerating forests within tropical landscapes, such as soils, topography, road access, and proximity to urban areas (Crk *et al.*, 2009). In Puerto Rico, older second-growth forests occur on limestone substrates, at higher elevations, on steep slopes, farther from roads, and in areas with less pasture (Helmer *et al.*, 2008).

Although there is no simple predictor of the likelihood or pathway of forest regeneration within a tropical region, conserving intact old-growth forests or large forest fragments is the best recipe for future forest regrowth. Natural regeneration will proceed more quickly and will have a higher composition of old-growth species in areas bordering old-growth forests and in regions where flora and fauna are protected from harvesting and hunting (Chazdon *et al.*, 2009b). Conservation measures aimed at protecting a diversity of forest types are needed to ensure a long-term future for regenerating forests in tropical regions (Chazdon *et al.*, 2009a).

See also: Rainforest Loss and Change. Succession, Phenomenon of. Tropical Forest Ecosystems

References

Asner GP, Rudel TK, Aide TM, Defries R, and Emerson R (2009) A contemporary assessment of change in humid tropical forests. *Conservation Biology* 23: 1386–1395.

Barlow J, Gardner TA, Araujo IS, *et al.* (2007a) Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. *Proceedings of the National Academy of Sciences of the United States of America* 104: 18555–18560.

Barlow J, Mestre LAM, Gardner TA, and Peres CA (2007b) The value of primary, secondary and plantation forests for Amazonian birds. *Biological Conservation* 136: 212–231.

Bazzaz FA and Pickett STA (1980) Physiological ecology of tropical succession: A comparative review. *Annual Review of Ecology and Systematics* 11: 287–310.

Benitez-Malvido J, Martinez-Ramos M, and Ceccon E (2001) Seed rain vs. seed bank, and the effect of vegetation cover on the recruitment of tree seedlings in tropical successional vegetation. In: Gottsberger G and Liede S (eds.) *Life Forms and Dynamics in Tropical Forests*, pp. 1–18. Stuttgart: J. Cramer.

Brown S and Lugo AE (1990) Tropical secondary forests. *Journal of Tropical Ecology* 6: 1–32.

Bush MB and Colinvaux PA (1994) Tropical forest disturbance: Paleoecological records from Darien Panama. *Ecology* 75: 1761–1768.

Campbell DG, Ford A, Lowell KS, *et al.* (2006) The feral forests of the Eastern Peten. In: Balee W and Erickson CL (eds.) *Time and Complexity in Historical Ecology,* pp. 21–55. New York: Columbia University Press.

Chai S and Tanner E (2011) 150 year legacy of land use on tree species composition in old secondary forests of Jamaica. *Journal of Ecology* 99: 113–121.

Chazdon RL (2003) Tropical forest recovery: Legacies of human impact and natural disturbances. *Perspectives in Plant Ecology Evolution and Systematics* 6: 51–71.

Chazdon RL (2008a) Beyond deforestation: Restoring forests and ecosystem services on degraded lands. *Science* 320: 1458–1460.

Chazdon RL (2008b) Chance and determinism in tropical forest succession. In: Carson W and Schnitzer SA (eds.) *Tropical Forest Community Ecology,* pp. 384–408. West Sussex, UK: Wiley-Blackwell Publishing.

Chazdon RL, Harvey CA, Komar O, *et al.* (2009a) Beyond reserves: A research agenda for conserving biodiversity in human-modified tropical landscapes. *Biotropica* 41: 142–153.

Chazdon RL, Letcher SG, van Breugel M, *et al.* (2007) Rates of change in tree communities of secondary Neotropical forests following major disturbances. *Philosophical Transactions of the Royal Society B-Biological Sciences* 362: 273–289.

Chazdon RL, Peres CA, Dent D, *et al.* (2009b) The potential for species conservation in tropical secondary forests. *Conservation Biology* 23: 1406–1417.

Corlett RT (1992) The ecological transformation of Singapore, 1819–1990. *Journal of Biogeography* 19: 411–420.

Corlett RT (1998) Frugivory and seed dispersal by vertebrates in the Oriental (Indomalayan) Region. *Biological Reviews* 73: 413–448.

Crk T, Uriarte M, Corsi F, and Flynn D (2009) Forest recovery in a tropical landscape: What is the relative importance of biophysical, socioeconomic, and landscape variables? *Landscape Ecology* 24: 629–642.

Cubina A and Aide TM (2001) The effect of distance from forest edge on seed rain and soil seed bank in a tropical pasture. *Biotropica* 33: 260–267.

Culot L, Lazo F, Huynen MC, *et al.* (2010) Seasonal variation in seed dispersal by tamarins alters seed rain in a secondary rain forest. *International Journal of Primatology* 31: 553–569.

Dent DH and Wright SJ (2009) The future of tropical species in secondary forests: A quantitative review. *Biological Conservation* 142: 2833–2843.

Don A, Schumacher J, and Freibauer A (2011) Impact of tropical land-use change on soil organic carbon stocks – A meta-analysis. *Global Change Biology* 17: 1658–1670.

Duncan RS and Chapman CA (2002) Limitations of animal seed dispersal for enhancing forest succession on degraded lands. In: Levey DJ, Silva WR, and Galetti M (eds.) *Seed Dispersal and Frugivory: Ecology, Evolution, and Conservation*, pp. 437–450. New York: CAB International.

Estrada A, Coates Estrada R, and Meritt Jr. D (1994) Non flying mammals and landscape changes in the tropical rain forest region of Los Tuxtlas, Mexico. *Ecography* 17: 229–241.

Etter A, McAlpine CA, Pullar D, and Possingham H (2005) Modeling the age of tropical moist forest fragments in heavily-cleared lowland landscapes of Colombia. *Forest Ecology and Management* 208: 249–260.

FAO (2010) *Global Forest Resources Assessment 2010.* Rome, Italy: Food and Agricultural Organization of the United Nations.

Feldpausch TR, Prates-Clark CD, Fernandes ECM, and Riha SJ (2007) Secondary forest growth deviation from chronosequence predictions in central Amazonia. *Global Change Biology* 13: 967–979.

Feldpausch TR, Rondon MA, Fernandes ECM, *et al.* (2004) Carbon and nutrient accumulation in secondary forests regenerating on pastures in central Amazonia. *Ecological Applications* 14: S164–S176.

Gehring C, Vlek PLG, de Souza LAG, and Denich M (2005) Biological nitrogen fixation in secondary regrowth and mature rainforest of central Amazonia. *Agriculture Ecosystems & Environment* 111: 237–252.

Gonzales RS, Ingle NR, Lagunzad DA, and Nakashizuka T (2009) Seed dispersal by birds and bats in lowland Philippine forest successional area. *Biotropica* 41: 452–458.

Grau HR and Aide M (2008) Globalization and Land-Use Transitions in Latin America. *Ecology and Society* 13(2): 16. http://www.ecologyandsociety.org/vol13/iss2/art16/

Guariguata M and Ostertag R (2001) Neotropical secondary forest succession: Changes in structural and functional characteristics. *Forest Ecology and Management* 148: 185–206.

Guevara S, Purata SE, and van der Maarl E (1986) The role of remnant forest trees in tropical secondary succession. *Vegetatio* 66: 77–84.

Helmer EH, Brandeis TJ, Lugo AE, and Kennaway T (2008) Factors influencing spatial pattern in tropical forest clearance and stand age: Implications for carbon storage and species diversity. *Journal of Geophysical Research* 113: G02S04, doi:10.1029/2007JG000568.

Howe HF and Smallwood J (1982) Ecology of seed dispersal. *Annual Review of Ecology and Systematics* 13: 201–228.

Hughes RF, Kauffman JB, and Jaramillo VJ (1999) Biomass, carbon, and nutrient dynamics of secondary forests in a humid tropical region of Mexico. *Ecology* 80: 1892–1907.

Hughes RF, Kauffman JB, and Jaramillo VJ (2000) Ecosystem-scale impacts of deforestation and land use in a humid tropical region of Mexico. *Ecological Applications* 10: 515–527.

Kammesheidt L (1999) Forest recovery by root suckers and above-ground sprouts after slash-and-burn agriculture, fire and logging in Paraguay and Venezuela. *Journal of Tropical Ecology* 15: 143–157.

Karthik T, Govindhan Veeraswami G, and Kumar Samal P (2010) Forest recovery following shifting cultivation: An overview of existing research. *Tropical Conservation Science* 2: 374–387.

Laborde J, Guevara S, and Sanchez-Rios G (2008) Tree and shrub seed dispersal in pastures: The importance of rainforest trees outside forest fragments. *Ecoscience* 15: 6–16.

Lawrence D (2004) Erosion of tree diversity during 200 years of shifting cultivation in Bornean Rain Forest. *Ecological Applications* 14: 1855–1869.
Lawrence D, D'Odorico P, Diekmann L, *et al.* (2007) Ecological feedbacks following deforestation create the potential for a catastrophic ecosystem shift in tropical dry forest. *Proceedings of the National Academy of Sciences of the United States of America* 104: 20696–20701.
Lebrija-Trejos E, Meave JA, Poorter L, *et al.* (2010) Pathways, mechanisms and predictability of vegetation change during tropical dry forest succession. *Perspectives in Plant Ecology Evolution and Systematics* 12: 267–275.
Lees AC and Peres CA (2009) Gap crossing movements predict species occupancy in Amazonian forest fragments. *Oikos* 118: 280–290.
Letcher SG and Chazdon RL (2009) Rapid recovery of biomass, species richness, and species composition in a forest chronosequence in Northeastern Costa Rica. *Biotropica* 41: 608–617.
Marin-Spiotta E, Cusack DF, Ostertag R, and Silver WL (2008) Trends in above and belowground carbon with forest regrowth after agricultural abandonment in the Neotropics. In: Myster RW (ed.) *Post-Agricultural Succession in the Neotropics*, pp. 22–72. New York: Springer.
Marin-Spiotta E, Silver WL, Swanston CW, and Ostertag R (2009) Soil organic matter dynamics during 80 years of reforestation of tropical pastures. *Global Change Biology* 15: 1584–1597.
Medellin RA and Gaona O (1999) Seed dispersal by bats and birds in forest and disturbed habitats of Chiapas, Mexico. *Biotropica* 31: 478–485.
Melo F, Rodriguez-Herrera B, Chazdon RL, *et al.* (2009) Small tent-roosting bats promote dispersal of large-seeded plants in a Neotropical forest. *Biotropica* 41: 737–743.
Mesquita RCG, Ickes K, Ganade G, and Williamson GB (2001) Alternative successional pathways in the Amazon Basin. *Journal of Ecology* 89: 528–537.
Moran EF, Brondizio E, Tucker JM, *et al.* (2000) Effects of soil fertility and land-use on forest succession in Amazonia. *Forest Ecology and Management* 139: 93–108.
Muscarella R and Fleming TH (2007) The role of frugivorous bats in tropical forest succession. *Biological Reviews* 82: 573–590.
Myster RW (2008) *Neotropic Post-dispersal Seed Predation Post-Agricultural Succession in the Neotropics*. New York: Springer.
Nadkarni NM and Haber WA (2009) Canopy seed banks as time capsules of biodiversity in pasture-remnant tree crowns. *Conservation Biology* 23: 1117–1126.
Norden N, Chazdon RL, Chao A, *et al.* (2009) Resilience of tropical rain forests: Tree community reassembly in secondary forests. *Ecology Letters* 12: 385–394.
Norden N, Mesquita RCG, Bentos TV, *et al.* (2011) Contrasting community compensatory trends in alternative successional pathways in central Amazonia. *Oikos* 120: 143–151.
Oliver CD and Larson BC (1990) *Forest Stand Dynamics*. New York: McGraw-Hill Inc.
Pan Y, Birdsey RA, Fang J, *et al.* (2011) A large and persistent carbon sink in the world's forests. *Science* 333: 988–993.
Parry L, Barlow J, and Peres CA (2007) Large-vertebrate assemblages of primary and secondary forests in the Brazilian Amazon. *Journal of Tropical Ecology* 23(23): 653–662.
Pena-Claros M and de Boo H (2002) The effect of forest successional stage on seed removal of tropical rain forest tree species. *Journal of Tropical Ecology* 18: 261–274.
Piotto D, Montagnini F, Thomas W, *et al.* (2009) Forest recovery after swidden cultivation across a 40-year chronosequence in the Atlantic forest of southern Bahia, Brazil. *Plant Ecology* 205: 261–272.
Pitman NCA, Ceron CE, Reyes CI, *et al.* (2005) Catastrophic natural origin of a species-poor tree community in the world's richest forest. *Journal of Tropical Ecology* 21: 559–568.
Poorter L (2008) The relationships of wood-, gas- and water fractions of tree stems to performance and life history variation in tropical trees. *Annals of Botany* 102: 367–375.
Poorter L, de Plassche MV, Willems S, and Boot RGA (2004) Leaf traits and herbivory rates of tropical tree species differing in successional status. *Plant Biology* 6: 746–754.
Powers JS (2004) Changes in soil carbon and nitrogen after contrasting land-use transitions in northeastern Costa Rica. *Ecosystems* 7: 134–146.
Read L and Lawrence D (2003) Recovery of biomass following shifting cultivation in dry tropical forests of the Yucatan. *Ecological Applications* 13: 85–97.
Ross NJ (2011) Modern tree species composition reflects ancient Maya "forest gardens" in northwest Belize. *Ecological Applications* 21: 75–84.
Saldarriaga JG, West DC, Tharp ML, and Uhl C (1988) Long-term chronosequence of forest succession in the upper Rio Negro of Colombia and Venezuela. *Journal of Ecology* 76: 938–958.
Scariot A, Vieira DLM, Sampaio AB, *et al.* (2008) Recruitment of dry forest tree species in Central Brazil pastures. In: Myster RW (ed.) *Post-Agricultural Succession in the Neotropics*, pp. 231–244. New York: Springer.
Schlawin J and Zahawi R (2008) 'Nucleating' succession in recovering neotropical wet forests: The legacy of remnant trees. *Journal of Vegetation Science* 19: 485–492.
Stoner K, Riba Hernandez P, Vulinec K, and Lambert J (2007) The role of mammals in creating and modifying seedshadows in tropical forests and some possible consequences of their elimination. *Biotropica* 39: 316–327.
Summerhayes G, Leavesley M, Fairbairn A, *et al.* (2010) Human adaptation and plant use in highland New Guinea 49,000 to 44,000 years ago. *Science* 330: 78–81.
Tucker J, Brondizio E, and Moran E (1998) Rates of forest regrowth in Eastern Amazonia: A comparison of Altamira and Bragantina regions, Para State, Brazil. *Interciencia* 23: 64–73.
Turner IM, Wong YK, Chew PT, and Binlbrahim A (1997) Tree species richness in primary and old secondary tropical forest in Singapore. *Biodiversity and Conservation* 6: 537–543.
Uhl C (1987) Factors controlling succession following slash-and-burn agriculture in Amazonia. *Journal of Ecology* 75: 377–407.
Urquhart G (2009) Paleoecological record of hurricane disturbance and forest regeneration in Nicaragua. *Quaternary International* 195: 88–97.
van Breugel M, Martinez-Ramos M, and Bongers F (2006) Community dynamics during early secondary succession in Mexican tropical rain forests. *Journal of Tropical Ecology* 22: 663–674.
Vazquez-Yanes C and Orozco-Segovia A (1993) Patterns of seed longevity and germination in the tropical rainforest. *Annual Review of Ecology and Systematics* 24: 69–87.
Vieira I and Proctor J (2007) Mechanisms of plant regeneration during succession after shifting cultivation in eastern Amazonia. *Plant Ecology* 192: 303–315.
Wahl D, Byrne R, Schreiner T, and Hansen R (2006) Holocene vegetation change in the northern Peten and its implications for Maya prehistory. *Quaternary Research* 65: 380–389.
Westoby M, Falster DS, Moles AT, *et al.* (2002) Plant ecological strategies: Some leading dimensions of variation between species. *Annual Review of Ecology and Systematics* 33: 125–159.
Whitmore TC and Burslem DFRP (1988) Major disturbances in tropical rainforests. In: Newbery DM, Prins HHT, and Brown ND (eds.) *Dynamics of Tropical Communities*, pp. 549–565. Oxford: Blackwell Science Ltd.
Wirth C, Messier C, Bergeron Y, *et al.* (2009) Old-growth forest definitions: A pragmatic view. In: Wirth C, Gleixner G, and Heimann M (eds.) *Old-Growth Forests: Function, Fate and Value* 207. pp. 11–33. New York: Springer.
Woods WI (ed.) (2009) *Amazonian Dark Earths: Wim Sombroek's Vision,* vol. 43. New York: Springer.
Zarin DJ, Davidson EA, Brondizio E, *et al.* (2005) Legacy of fire slows carbon accumulation in Amazonian forest regrowth. *Frontiers in Ecology and the Environment* 3: 365–369.

True Bugs and Their Relatives, Diversity of

Carl W Schaefer, University of Connecticut, Storrs, CT, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Beak (or rostrum) Thin elongate mouthparts, adapted for piercing the organisms on which the bugs feed (the mouthparts of bugs are similar in function to, but quite different in structure from, those of mosquitoes).

Food pump Sclerotized pump used by hemipterans to suck up the juices (plant or animal) on which they feed.

Hemelytron Heteropteran's forewing; so called because the anterior part of the wing is opaque and the posterior part is clear (membranous).

Salivary pump Sclerotized pump used by hemipterans to inject saliva into the organisms on which they feed.

Scent glands Glands through which heteropterans secrete various compounds for various purposes, including defense, attraction of mates, and attraction of both sexes. Adult heteropterans have scent glands on their metathoraces, and nymphal heteropterans as well as some adults have scent glands on the dorsa of their abdomens.

Scutellum The dorsal portion of the mesothorax, in most heteropterans triangular or shield-shaped (hence "little shield").

Suborders of Hemiptera

The number of actual suborders has recently come into question, because of the uncertain validity of Homoptera. However, three suborders are frequently accepted in most entomology texts: Homoptera, Heteroptera, and Coleorrhyncha. Nevertheless, many homopterists recognize two orders, Homoptera and Hemiptera (= Heteroptera). The reasons for this confusion are themselves confusing and have to do with the importance that heteropterists and homopterists give to the groups they study; this importance in turn is based on the extent to which the students of one group understand the work of the students in -the other groups. Those who have undertaken the most thorough study of all hemipteran groups conclude that they share too many characteristics to be treated as separate orders.

"Homoptera"

This group includes the cicadas, leafhoppers, aphids, scale insects, whiteflies, and their relatives. Nearly all feed on plants (a few on fungi), and are not predacious; and many are important pests of crops and ornamentals. Recent work on the origins and (especially) the evolutionary relationships of groups within "Homoptera" has led to the conclusion that the group did not have a unique ancestor common to all homopterans and to no other hemipterans. As a result, the higher classification of "Homoptera," and the systematic status of the group itself, are deeply in question. That is why the author places the term "Homoptera" in quotation marks: The term is useful in grouping together a number of insects, but it is not useful in suggesting they compose a single evolutionary unit.

Coleorrhyncha

Members of the single family (Peloridiidae) have a classic Gondwana distribution: Australia, New Zealand, southern South America. They are associated with moss in *Nothofagus* (southern beech) forests and probably feed on the moss. Small, somewhat flattened, often with lovely semitransparent expansions of the head and thorax, these insects are poorly known and deserve much more study. Some believe Coleorrhyncha (formerly included in Homoptera) to be the sister group of Heteroptera – that is, that these two groups had a common ancestor. The evidence for this conclusion is stronger than that for the earlier inclusion of coleorrhynchans in Homoptera, where they were once thought to be a sort of "nonmissing link" between Heteroptera and Homoptera. Recently a fine monograph on coleorrhynchans, with 13 genera and 37 species, was written by Burckhardt (2009); this also includes the biogeography of these interesting insects.

Suborder Heteroptera

There are about 42,000 described species of Heteroptera (Henry, 2009) and (a much rougher estimate) perhaps 25,000 species remaining to be described: a total of about 67,000 species of Heteroptera altogether. (Note: In what follows, the estimates of groups of Heteroptera come from Henry (2009).) This estimate may understate the number of undescribed species. Like all hemipterans, they feed by piercing other organisms and sucking their juices. Unlike Homoptera and Coleorrhyncha, however, Heteroptera includes many predacious insects, and many aquatic ones. Heteroptera are more closely allied (evolutionarily) with Coleorrhyncha than with Homoptera, although some have recently argued that the closest relatives of Heteroptera are certain homopteran groups.

Heteroptera: General

Why "True" Bugs?

Heteropterans are the only group of insects that entomologists agree may legitimately be called "bugs"; hence the name entomologists themselves use, "true bugs." Nonentomologists call all insects (and other arthropods) "bugs," and, although entomologists disapprove, they are resigned to this use.

Why are heteropterans called "bugs"? The reason almost certainly lies in the original Middle English meaning of the word: "bug" (as *bugge*) meant originally a *wraith* or *specter*, something unseen that emerges from the shadows and does harm. This is exactly what the best known heteropteran does: the bedbug (*Cimex lectularius* L.) sucks blood from the sleeping victim and disappears; in the morning only the irritating welt from the bite remains, its cause and origin unknown. Clearly, an unseen nocturnal shadow has caused this damage: a bug. (This original meaning of *bugge* is retained in Modern English in such words as *bugbear*, *bugaboo*, and *bogeyman*; one may also wonder if it is retained in the verb "to bug," meaning (1) to irritate and (2) to use a small and secret device for spying – a form of damage.)

Characteristics

Heteropterans share many structural features with other insects, of course; other features with other insects with incomplete metamorphosis; and of course some characteristics with other hemipterans. Still more aspects of their morphology are unique. A superb early account of structure and function is given in Weber (1930).

Like all insects, heteropterans have three basic body regions: head, thorax, abdomen; and, like those of nearly all insects, these regions are specialized: the head for grasping and preparing food and for exploring the environment, the thorax (with legs and wings) for locomotion, and the abdomen for processing food and for reproduction. Like most insects, heteropterans have three pairs of five-jointed legs and two pairs of wings, a pair of compound eyes (usually large, because true bugs are mostly diurnal and therefore sight is an important sense), a pair of antennae, and an aedeagus (or penis) for the introduction of sperm into the female, who in turn has an ovipositor for laying eggs.

Like other insects with incomplete metamorphosis, immature bugs look like their adults, except for the lack of functional wings and genitalia and for different body proportions. As the immatures develop, wings appear externally as small pads, which enlarge at each molt; later, the genitalia too may be discerned, and often the last immature stage can be sexed.

Heteropterans share more specific features with other hemipterans: The mouthparts are modified into a long sharp tube, with two internal openings. Through one of these, salivary fluid is pumped into the plant or animal (plant only, in the case of nonheteropteran hemipterans), and through the other opening the fluids are sucked up. A set of pumps within the head accomplishes these activities. This mechanism and method of feeding are perhaps the best evidence that all hemipterans are closely related, and this in turn is why most entomologists consider all hemipterans to compose a single order.

Heteropterans, like other hemipterans, have certain characteristic features of the digestive system and the excretory system (although heteropterans lack the "filter chamber" independently evolved in several homopteran groups). The primitive 12–13 ganglia of the ventral nervous system are fused to a single one (or to four in a few homopteran groups). The wing venation is reduced (less so in many homopterans); the chromosomes are holocentric (diffuse centromere). Many hemipterans can produce sounds, usually by scraping a "pick" over a corrugated surface. Often these structures appear the same and occur on the same parts of the body. But in most cases, these sound-producing structures have evolved independently; the appearance is constrained by their function, and their occurrence on the same body parts is constrained by the fact one body part must move and be in contact with an unmoving part.

Heteropterans themselves are more diverse in habit and habitat than the other hemipteran groups. Heteroptera contains the only hemipterans that live in, and sometimes on, the water; the only hemipterans that are partly marine; the only carnivorous hemipterans; and the only hemipterans that feed on vertebrate (including human) blood. It is not surprising, then, that within the suborder Heteroptera is a wide variety of sizes and shapes, nor is it surprising that structures of heteropterans (legs, etc.) are highly modified in some groups.

The beak, or rostrum, characteristic of Hemiptera in general, in some plant-feeding heteropterans may be longer than the body, and in some predacious heteropterans be so short as not to reach the thorax. It may be slender or stout, and the number of segments may be reduced from the original five. The beak arises closer to the back of the head in heteropterans than in homopterans, and thus a region of the head lies just in front of the rostrum's base; the presence of this region in coleorrhynchans is one reason why Coleorrhyncha and Heteroptera are thought to have had a common ancestor.

Highly characteristic of Heteroptera alone is the division of the forewing into two regions, the posterior one, clear and membranous (like the hindwing), and the anterior portion, harder, often leathery, and opaque: hence the name "heteroptera" ("different wings"). The two wings on each side are coupled together with a device on the forewing; this is found only in Heteroptera and Coleorrhyncha. Heteropteran sperm, too, has some unique features; coleorrhynchan sperm has not been examined.

Characteristic of nearly all heteropterans is a triangular (pointing backward) part of the top of the thorax, the *scutellum* ("little shield"); in a few heteropterans, especially in the group Pentatomoidea, the scutellum covers all or most of the dorsal surface of the insect, except the head (but including the wings).

Unique to Heteroptera are scent glands. These are found on the abdomen of immature heteropterans (from one to four pairs on either side of the midline). These glands are usually nonfunctional in adults, although in many species they do produce scent; the known number of such species is increasing with increasing study, and the function of the scent from these dorsal abdominal glands is becoming more clear. Adult heteropterans have a pair of glands (regardless of their possession of functional dorsal glands) on the last (third) thoracic segment. These lateral metathoracic scent glands are thought to repel predators (and have been shown to repel ants), but they certainly have other functions as well, including the attraction of conspecifics of the opposite (and sometimes of both) sex(es).

Heteroptera do not differ biologically from other insects and other hemipterans. They do the three basic things that all organisms do – feed, reproduce, and disperse – and, because these things are basic to survival as a species, bugs do them as

much as other winged insects. However, bugs are remarkably more diverse than other hemipterans (and indeed than other insects with incomplete metamorphosis) in their choices of food and of habitats and in their interactions with their environments (for the latter, see below).

Heteropterans feed like other hemipterans. Their sharp elongated mouthparts are adapted for piercing; one internal pump pushes saliva into the food source, and another sucks back the food. In most cases, this food has been all or partly digested by salivary enzymes. Few bugs feed on the watery contents of the xylem in plants, and so bugs lack the filter chamber that in many homopterans removes and excretes excess water (and sugars). (Diuresis does occur in blood-sucking bugs, however: see below.)

The salivary enzymes and substances of plant-feeding bugs serve other functions than digestion alone. Many phytophagous bugs form a small feeding cone on the surface of the plant where the mouthparts are inserted. This cone, made from salivary secretions, helps stabilize the mouthparts as they penetrate to the source of food. Other salivary products help break down cell walls and release the cell contents; other products mimic plant auxins and "fool" the plant into mobilizing nitrogen-rich materials at the point of feeding.

As a general rule, with many exceptions, plant-feeding heteropterans feed on the highly nutritional reproductive parts of plants – flowers, developing fruits and seeds, ripened fruits and seeds, and fallen fruits and seeds. This may be a matter of size: Many plant-feeding bugs are relatively large (i.e., larger than the majority of homopterans); larger hemipterans are better able to penetrate plant parts to the seeds within and to disperse to find flowers, fruits, and seeds – which are transitory resources. Smaller heteropterans, like many homopterans, feed on leaves, a resource more permanent, easier to penetrate, but of less nutritional value.

The flat bugs (Aradomorpha – see below) have very long, very slender mouthparts with which they feed on the linear array of fungal mycelia below the bark of dead or dying trees. A few of these bugs live in termite nests, where they probably feed also on fungi. With very few exceptions, these are the only bugs that feed on fungi.

Indeed, the great majority of phytophagous heteropterans feeds on flowering plants, angiosperms. Several small groups feed on gymnosperms, and a very few bugs feed on mosses or ferns. It seems likely that the great radiation of heteropterans – like the radiations of so many other insect groups – occurred at the same time as, and perhaps in response to, the radiation of flowering plants.

Within the angiosperms, some groups are less preferred than others: grasses, for example, and legumes. Each of these groups has developed defenses against being eaten, and in each case one or more groups of bugs has overcome the defense and radiated on the plant group. It is significant that these two groups (grasses and legumes) have few *individual* species feeding on them but, rather, a few groups of species (subfamilies, families). This suggests that an early ancestor broached the defense (of grass or legume) and its descendants radiated on this newly available food source.

Many bugs are predacious. Indeed, it has been argued that this was the original way of life of Heteroptera (an argument that some dispute). The prey is other insects and often their eggs or other immature stages. Some of the smallest heteropterans are predacious, on arthropod eggs, mites, and scale insects and aphids. It was long thought that predacious bugs inserted their mouthparts and sucked up fluids inside the prey, much as plant feeders suck sap. Recently, it has been demonstrated how much more complex, and interesting, the situation really is (Cohen, 1998). A predacious bug actively pumps digestive enzymes and fluid into the prey. The enzymes digest the soft internal organs of the prey, and the fluid and digested food are sucked up. This "solid-to-liquid feeding" process is repeated over and over, the mouthparts within the prey moving throughout its body, probing even into the legs and antennae to digest the muscles there. There remains but a husk, a shell, the exoskeleton of the prey, all internal organs having been digested, liquefied, and sucked into the predator. As a consequence, a predacious bug need feed on only a few prey, because all of the prey except its exoskeleton is eaten. The implications of this for biological control using heteropterans are considerable (see below).

The predacious habit has arisen often in the Heteroptera, secondarily, probably more often than phytophagy. This suggests that the search for more concentrated and usable nitrogen is as important in this group as it is in others.

A few bugs feed on vertebrate blood. Heteropterans' mouthparts, like those of mosquitoes, are long and sharp, and so adapted for sticking into liquid-filled organisms. It seems inevitable that some would take vertebrate blood, although few feed on birds and none on cold-blooded vertebrates (with one exception). With few exceptions, the vertebrates fed on are small mammals and bats – and man. Members of two closely related families (Polyctenidae and Cimicidae) feed on bats, and a few members of the latter family feed on humans: the bedbugs, some of whose relatives also feed on poultry. Members of the assassin bug subfamily Triatominae feed on small mammals, and some of them on humans; these spread the protozoan trypanosome which causes Chagas disease.

The problem of feeding on vertebrate blood includes the problem of ingesting red blood cells through a narrow tube. The physics of this type of feeding have been studied, and maximal feeding is related to the inside diameter of the feeding tube, the concentration of red blood cells, and the duration of feeding. Such feeding also results in the accumulation of blood fluids. These are excreted (diuresis) in a manner roughly analogous to the excretion of plant juices by homopterans ("honeydew").

Two other heteropterans feed on vertebrates: a birdnest-living lygaeid (the "seed bug" family) apparently feeds only on the blood of the host bird; another group of lygaeids apparently feeds on small mammals in Africa (and may there occupy the niche occupied in Latin America by Triatominae). Giant water bugs (Belostomatidae) have been reported to attack baby ducklings (this may be a legend), and some species specialize in taking the fluids of small fish.

No heteropterans feed exclusively on dead or decaying organic matter (i.e., none is a saprophage). However, there are many anecdotal accounts of various bugs found feeding on bird droppings, carrion, and dead insects. It is likely these bugs sought water or a source of concentrated soluble nitrogen.

The great majority of heteropterans reproduce like other insects: Sexes find one another, there may be a brief period of

courtship, mating occurs (fertilization is internal), and the fertilized eggs are laid near or on the appropriate food for the young.

The sexes find one another probably *via* sex attractants given off by one or the other sex, or perhaps by both (depending on the species; however, very few species have been studied). Courtship, if it occurs at all, is brief. In a few species, males fight for females; the enlarged legs of certain males are used to knock one another off a leaf to which sex attractants are drawing females. However, there seems to be very little intrasexual rivalry in Heteroptera. Mating is often end-to-end, or in many species the male lies at an angle across the female; in all cases both sexes are ventralside down. In some end-to-end species, the female, larger than the male, may be spotted moving around dragging her mate with her; or she may feed casually on a leaf, with the smaller male dangling helplessly in the air, attached only by the genitalia.

Soon after mating occurs, eggs are laid singly, or in characteristic batches, sometimes in several layers, depending on the species and, often, on the family. A female lays several batches of eggs in her lifetime, each one usually requiring a separate mating. The eggs of many species have tiny openings (micropyles) through which the sperm enter; other small openings (aeropyles) provide gas exchange; and some species' eggs have a round or oval line of weakening, through which the insect will hatch.

Eggs may be round, oval(!), cylindrical, or barrel-shaped. Many are brightly colored. The egg of Rhopalidae ("scentless plantbugs") is mounted on an elongate pedicel (like eggs of some neuropterans), presumably to protect it from predators.

Parental care of the eggs occurs in a few species. As is well known, the female of some giant water bugs (Belostomatidae) glues her eggs onto the back of the male, who carries them until they hatch. By sweeping water across them, he keeps them aerated and prevents the attack of fungi. The females of some stinkbugs (Acanthosomatidae) and some lace bugs (Tingidae) protect their eggs and early immature stages, guarding them from attack from parasites and such predators as ants. Experiments have shown that in all these instances, the protection is effective.

The author stresses, however, that with respect to reproduction, as with respect to feeding and all other aspects of heteropterans' lives, we know very little about the lives of some bugs, and nothing about the lives of most.

One strange exception, unique to a few heteropterans, is "traumatic insemination." One of the male's claspers is sharp and sometimes scimitar-shaped. With it he slashes the female's abdomen and sperm is deposited *via* a duct running along the clasper. The sperm finds its way to the storage receptacle of the female. In the evolutionarily more advanced members of this group, a special pad of tissue on one side of the female's abdomen receives the sperm, which then moves to the storage site, passing on its way *through* the cells. This is the situation in the common bedbug. In evolutionarily less advanced species (certain members of the bedbug's family and some Anthocoridae), there is no such tissue, and the sperm are simply deposited in the abdomen. Counting the resulting mating scars allows one to learn how often a female has been mated.

Like those of most insects, the eggs of heteropterans hatch within the year; often they wait through cold seasons (in temperate regions) or dry ones (in tropical regions). Hatching often occurs through a weakened area of the egg and is aided by an "egg burster" on the hatchling, which is soon lost.

After hatching, the immature bug molts several times, the last time to adult; five immature stages occur in nearly all bugs, and four in the rest. The first immature of several plant-feeding bugs does not feed. Blood-feeding bugs require a bloodmeal to molt to the next stage; molting may require a full meal in other bugs (especially predacious ones), but in fact the stimulus to molt in nonbloodfeeders is unknown.

Each successive stage is larger than its predecessor, of course, and in each the wings are more fully developed. Development of the wings is so consistent in Heteroptera, that the immature stage can often be determined by the relative development of the wings. The sex can often be detected in fourth-stage, and always in fifth-stage, immatures, as the outlines of the external genitalia become discernible.

In general, the immatures feed on the same organisms as the adults. Predacious immatures may take smaller prey, or they may collaborate in taking larger prey. The young of some plant-feeding bugs feed on different plants than the adults, perhaps in response to differing seasonal availabilities of different plants; perhaps also some plants provide better food for growth and others better food for reproduction.

In temperate regions, most bugs go through a single life cycle, overwintering as egg or adult (there being no pupal stage). Eggs overwinter embedded in twigs or in leaf litter; diapausing adults overwinter in litter or just below the soil surface, near their host plants. In spring, overwintered adults mate, lay eggs, and the hatchlings actively seek food (those of some species do not feed, however). Overwintered eggs hatch usually in midspring and become adult in midsummer or late summer, mating and laying eggs that will overwinter. In warmer regions (and, for some species, in temperate regions), there may be more than one generation in a year, and these may overlap. In regions with dry seasons, adults or eggs may estivate. Aquatic heteropterans may overwinter as eggs in or near the water; several species are active below the ice in temperate-region winter; and adults of water striders and their relatives often overwinter on land, rather far from water.

Heteropterans vary widely in their ecological adaptations, both to the biotic and the abiotic aspects of their habitats. Their relationship with their sources of food is of particular significance, both ecologically and because it is this relationship that determines their impact on humans.

Because heteropterans have sucking mouthparts, they must take in liquid or semiliquid food and can accommodate only the smallest of particulate matter. This constraint lies at the base of their relationships to their food.

Plant-feeding heteropterans vary in their degree of host specificity. Some species feed on only a single species of plant, others on members of the same genus, others on many members of the same plant family; and many bugs feed very widely indeed. One coreid is listed from plants in 90 genera representing 35 plant families! It is of interest that a single genus may contain both very highly host-specific species and species feeding on a great variety of plants; this phenomenon is not restricted to Heteroptera.

Not surprisingly, the more host specific a bug is, the more closely its life cycle is tied to the host's. Eggs hatch when the

preferred parts of the preferred host plant are available: Hatch occurs later in the plant's life cycle if the bug prefers reproductive parts, earlier if the bug prefers somatic parts. The triggers for hatch are not known in nearly all cases. The eggs of some mirids are laid in woody tissue of the host and prepare to hatch when the spring uptake of moisture causes the wood to swell and exert pressure. Several other species of Miridae specialize on grasses, and as these different grasses succeed one another in a field over several weeks, so do the mirids. Again, the dynamics and phenology of this succession remain to be studied.

Mirids in particular, many of which feed on shortlived annuals, are closely tied to their hosts' life cycles. The adults of several species of heteropterans feed on different plants than do the immatures. This may reflect differing needs of the stages: immatures need nutrients for growth, adults need nutrients for reproduction; or, the switch from one host to another may merely reflect the fact that these bugs live longer than the plants preferred by the immatures.

The newly hatched immatures of many phytophagous bugs do not feed at all, or take only water; some species take up symbionts from the egg-shells. This period of time when feeding does not occur may be a "hedge" against a delayed appearance of the host plant. It would be interesting to seek a correlation between species with nonfeeding hatchlings and the uncertainty of their host plants' appearance.

Many heteropterans prey on other insects; a few feed on vertebrates; and only a very few feed, occasionally, on noninsect arthropods. In general, predacious bugs feed on prey of appropriate size: very small predators feed on small prey, such as insect (and mite) eggs, mites, scales, and aphids. Large predators feed on larger prey. However, many larger prey organisms (such as caterpillars) are slow-moving and have defenses against predators on their body surfaces; such organisms are available to predators that avoid the body's surface (by penetrating it with their beaks). And so many small predacious bugs will attack large prey, often collectively; such bugs include small adults and the small early immatures of predators whose adults are large. There is some evidence that the collective effect of the predators' injected digestive enzymes benefits all the predators feeding on the single prey. These collective associations are all of the same species.

Aquatic and semiaquatic bugs are nearly all predacious. Those bugs that live beneath the water's surface prey on other small aquatic invertebrates (although one group of water bugs feeds on snails, and another on small fish). Bugs that live on the water's surface (semiaquatic) prey on land invertebrates that fall into the water and become trapped on its surface.

Predacious bugs – whether land, aquatic, or semiaquatic – are either sit-and-wait predators or stalk-and-pounce predators. The great majority attacks its prey on foot, not in flight.

Bugs that feed on vertebrate blood live associated with their hosts, not on them. The hosts are vertebrates that live in enclosed habitats: burrows, nests, caves – and houses. The bugs spend the day in cracks and crevices in these habitats and, at night, venture forth to feed. When feeding, the bug injects an anti-irritant, so that only later is the fact of feeding known. The host list is rather limited: a few birds that live communally (swallows, domestic poultry, and a few others), humans (which also live communally), and especially bats – two families of heteropterans (Cimicidae and Polyctenidae) appear to have arisen as specialists on bats. One may speculate that the bedbug itself (family Cimicidae) became acquainted with humans when the latter took shelter from the cold in caves, during early Ice Ages; indeed, the only wild bedbug collected came from a bat cave in the northern Middle East.

Surprisingly little is known about the influence of abiotic factors on the individual lives of heteropterans or on bugs' population or community ecologies.

The effects of, and the reactions of bugs to, light and humidity are very poorly known. A few species disperse (i.e., large numbers move from one place to another, although not at regular intervals), and this movement in a few species is influenced by amount of moonlight. In other cases, dispersal is a response to deteriorating environment, whether physical (drying, for example) or biological (dearth of food). Several such bugs (e.g., species of cotton stainers, *Dysdercus*) lose their flight muscles after such dispersal, presumably to divert resources to reproduction.

The existence of circadian rhythms has been documented as influencing a very few physiological phenomena of a very few bugs.

The greatest enemy of heteropterans is other arthropods – this is of course true of all groups of insects. Small heteropterans are fed on by small predacious arthropods, and large heteropterans by large ones. Large heteropterans are also eaten by insectivorous vertebrates – fish, reptiles, birds, and small mammals. All bugs are attacked by ants.

It has been thought that the scent glands, whose possession is a unique characteristic of heteropterans, are defensive against predators. This is probably so, but it is certainly not their only function. Nevertheless, the few actual experiments and observations that have published support the idea that predators – especially ants and vertebrates – are repelled by the glands' secretions. More work on this is needed.

Many true bugs are brightly colored, arrayed in contrasting patterns of black with red, orange, yellow, or white. These bugs, said to be warningly colored, are assumed to be distasteful to predators (at least, to visually orienting predators, like vertebrates) and to warn such predators away. Experiments with the large milkweed bug (family Lygaeidae) confirm that this is the case. The bug feeds on the intensely bitter sap of milkweed and sequesters the cardenolides (which are also strong ATPase inhibitors); these have no effect on the bug itself.

There are many other bugs with conspicuous patterns like the milkweed bug's. Many of these bugs are also bad-tasting, but certainly many are not. Batesian mimicry (where a palatable organism looks like an unpalatable one, and thereby gains protection) is common in Heteroptera. So is Müllerian mimicry, where two unpalatable species resemble one another and presumably each gains from the protection afforded by the other. The difference between Batesian and Müllerian mimicry is graded, not absolute. Also, we need experimental evidence to show what type of mimicry two phylogenetically different but similar bugs represent.

In making these assessments, one should also realize that the types of patterns, and colors, "available" (in the ecological-evolutionary sense) to these bugs are limited. The dark parts of the pattern must be black or near black, and the colored parts red to yellow, because of pigment limitations. The types of

patterns are also limited by the smallness of the canvas (the insect's body) on which they are to be painted. There may also be some genetic constraint, for many of the patterns are much the same: white spots or a line across the wings, for example (although this may not be for contrast but for pattern disruption; see below). For this reason, similar warning patterns may have arisen by chance, and not from a selective pressure for mimicry.

A few predator–prey mimicry complexes are known. In these, a brightly colored predacious species closely resembles its usual prey. Some have dubbed this "aggressive mimicry," thinking the predator "fools" the prey into allowing approach. This supposes that the prey sees as do human beings, something that is most unlikely. More probably, these are Batesian (or Müllerian) mimicry complexes, and what the prey loses in associating with the predator, it more than gains by being protectively – warningly – colored like the predator. This suggestion needs experimental testing.

Such complexes often involve large populations. Other (nonheteropteran) warningly colored species also live in large populations. Perhaps the mass of bright color warns motile predators (like birds) away from a distance.

Other bugs protect themselves by mimicking inedible bits of their environment. At least seven distantly related families and superfamilies contain species that mimic ants; and within several of these families such mimicry has evolved several times independently. Ants are unattractive as food, because of their high content of formic acid, and many other insects mimic ants. In the Heteroptera there are species that look so much like ants that even experts may be fooled. When mimics look like particular species of ants (rather than like ants in general), they look like *local* species. The immatures of some Alydidae look like small species of ants when young and like larger species of ants as they grow.

Three facts constrain such mimicry: First, large bugs cannot look like ants, because ants are small. Second, the bug must be at least somewhat ant-shaped: slender and elongate. This is probably why ant mimicry has not developed in the very large superfamily Pentatomoidea, whose members are large and squat. Third, ants are characterized by a constriction between thorax and abdomen (the "wasp waist"). Large bugs can do nothing about the first constraint. Small bugs over evolutionary time may overcome the third constraint, by "developing" a similar "waist." However, many true bugs "fake" the constriction: Where on an ant the "waist" would be, the bug (otherwise dark, like an ant) has white markings. Viewed from a slight distance, against the dark background of a branch or the ground, these markings give the appearance of an *absence* of body – hence, a "waist." A sequence of such antlike appearances may be seen in the Arhaphinae, a neotropical group of Largidae, and in the Micrelytrini, a tropical group of Alydidae only distantly related to Largidae. Many unrelated heteropteran ant mimics have one or more dorsal spines, presumably to harm the tender mouth tissues of a vertebrate not deterred by the antlike appearance.

Less common is the mimicry of wasps, which occurs in some adult Alydidae, and doubtless in a few other bugs.

Other heteropterans look like twigs, or leaves, or other things of no interest to a predator. Many run or fly rapidly and erratically.

Many bugs can defend themselves more actively, as anyone who has been "bitten" by a backswimmer or giant water bug knows. Predacious bugs especially, with their stout beaks designed to be driven through the hard exoskeleton of another insect, can defend themselves effectively against vertebrate predators – or against insect collectors.

Aquatic bugs would seem to be a delightful prey for fish. Yet they are not so frequently eaten; indeed, even though trout live in the same streams as water striders, stomach contents indicate the striders are rarely eaten. Many water bugs "bite" fiercely (as most aquatic biologists know), and it seems likely that their scent glands help protect them, although the evidence for this is very slim.

Like many other animals, some true bugs break up their body outline by using spots or lines of white against a dark background (the "zebra effect"); indeed, the antlike appearance of some bugs may have arisen this way. Many Coreidae have a white line, straight or zigzagged, across the dark wings; many Pyrrhocoroidea have white spots in the same place, as do some members of other groups. Like the zebra's white stripes, these bits of white break up the bug's outline and make it more difficult to be seen – especially when moving in an erratic way.

Heteropterans seem not to be heavily parasitized. Some groups of parasites are specialized for parasitizing certain bugs (the subfamily Phasiinae of the dipteran family Tachinidae, for example); the eggs of heteropterans may at times be parasitized by hymenopterans; and certainly some species of heteropterans are quite heavily parasitized. But overall, the group seems less afflicted with parasites than (for example) the Lepidoptera.

The females of some lace bugs (Tingidae) ward off predators and parasites by protecting their eggs and young, and this behavior occurs in a few other heteropteran groups. Some heteropterans bury their eggs partly or completely in the substrate, probably to prevent their being parasitized: Many lace bugs, for example, embed their eggs on the lower surface (less conspicuous) of leaves and cover them with frass, thus concealing them from sight and (perhaps) from scent.

Finally, like all arthropods, heteropterans wear their skeletons on the outside. This makes their bodies hard and often smooth. As a result, several kinds of potential predators (especially ones with jaws) simply cannot grasp or crush heteropterans (or many other arthropods). For example, there are few records of predacious beetles feeding on bugs.

Economic Importance

Many heteropterans are serious pests of crops, and one group transmits the trypanosome protozoan that causes Chagas disease. This group is the Triatominae, a subfamily of assassin bugs (Reduviidae) that lives in the New World Tropics (with a few exceptions) and in which *Trypanosoma cruzi* develops. These bugs feed on vertebrate blood. Several live in association with humans and can transmit *Trypanosoma cruzi* to them. The result is Chagas disease, which afflicts millions of people a year in Latin America, often killing its victims and costing several billions of dollars in lost lives, lost productivity, and disease prevention and treatment (see Schofield, 1994).

A genus of Triatominae lives in India, but so far Chagas disease does not occur there. Very rarely, the disease occurs in the southern United States, where the bugs, the trypanosomes, and the wild rodent hosts all may be found.

More widespread is the damage caused to plant crops by heteropterans. Several species of stinkbugs, called collectively Sunn pests, are the chief destroyers of wheat and barley in a broad swathe from eastern Europe through the Middle East into Pakistan. In parts of this range, these bugs may destroy up to 70% of the crop (see Javahery, 1995). The southern green stinkbug is a major pest of many crops throughout the world. In the United States, the chinch bug *Blissus leucopterus* (Lygaeoidea: Blissidae) is sometimes a serious pest of wheat. Cotton stainers (Pyrrhocoridae) damage cotton and cotton seed in many semitropical parts of the world (but not the United States). Lygus bugs (Miridae) are very serious pests of a wide variety of crops, especially in the temperate parts of North America. In the last few years, there has been a resurgence of bedbugs, which one expert has called a "pandemic," in all parts of the world. This "pandemic" has perhaps been caused by the increased movement of people throughout the world. Many other heteropterans are general or specific pests, locally or worldwide, on many crops: A detailed account of these pest insects may be found in Schaefer and Panizzi (2000).

A few heteropterans are important in biological control. A very few feed on weeds and provide some control. Many are predacious and feed on other insects. However, like most predators, most heteropterans are not prey-specific and cannot usually be depended on to control completely pests on a crop; moreover, they do not feed on many prey items. These insects would serve better as components of an Integrated Control Program. Among the families that have been studied are Reduviidae, Berytidae, Nabidae, Anthocoridae, (some) Miridae, Geocoridae, and Pentatomidae: Asopinae; members of several aquatic families have been tried in controlling mosquitoes, without much success (see Schaefer and Panizzi, 2000).

As noted above, predacious heteropterans feed on only a few prey. When it was believed that these bugs removed from their prey only the latter's liquids, it was also thought that each predator must feed on many prey in order to get enough nourishment. Now we know this is not so. As a result, those using these bugs in biocontrol programs must rethink the bugs' efficiency, because each bug actually feeds on far fewer prey items than had been thought. Many prey–predator models, and considerable biocontrol planning, have been rendered invalid by the recent discovery of this mode of feeding (Cohen, 1998).

Heteroptera: Specific

The suborder Heteroptera contains eight infraorders: Enicocephalomorpha, Dipsocoromorpha, Gerromoropha, Leptopodomorpha, Nepomorpha, Cimicomorpha, Pentatomomorpha, and Aradomorpha; "-morpha" means "in the form of," and each infraorder is derived from the generic name of a member: Hence the often pen- and tongue-shattering concatenation of syllables. These groups, and the families that compose them, are discussed very well in Schuh and Slater (1995), and, in less detail, by Henry (2009) and Schaefer (2009).

Of these, the first two infraorders contain relatively few species (400 and 320, respectively), of poorly known insects; these are all predacious and many live in or on ground debris; it is certain that many more species will be discovered when this type of habitat is explored in the Tropics. They are of interest because they are phylogenetically primitive.

The Gerromorpha contain about 2000 species, many of them in the Gerridae, the water striders, or water skaters, or "Jesus bugs," which are common everywhere skimming the surface of still waters or the edges of moving waters. Here they capture insects that fall into the water; they rarely if ever capture aquatic organisms. One group of water striders is among the very few insects that live on or in saltwater: the marine water striders live on the ocean's surface, often far from land. Because they live on the water, not in it, gerromorphans are called "semiaquatic" bugs, or, in an earlier terminology, Amphibicorisae.

The Leptopodomorpha too are semiaquatic, most of the 380 or so species living near water but neither on it nor in it; a few species live in the splash zone of the sea. All are predacious. Most belong to the family Saldidae, the shore bugs, which are fairly common on rocks in or near water, where their ability to jump and conceal themselves in crevices frustrates the best of collectors.

The 2300 species of Nepomorpha, aquatic bugs (Hydrocorisae), live below the water's surface. Here are the familiar backswimmers (Notonectidae), water boatmen (Corixidae), giant water bugs (Belostomatidae; some Indian species may reach 4 inches in length), and water scorpions (Nepidae); there are several other nepomorphan families that have more tropical species than temperate. All of these are predacious, although water boatmen feed also on freshwater algae. Some giant water bugs specialize on freshwater snails and may be useful in controlling those harboring the schistosomiasis platyhelminth. Others sometimes cause some damage in fisheries, where they may attack young fish. From time to time, someone suggests using some of these aquatic bugs to control mosquitoes, but the attempts have yet to succeed.

The Cimicomorpha, with more than 26,400 species, is the largest of the heteropteran infraorders. This is because it contains the Miridae, or plant bugs, which, with almost 10,400 species, is the largest heteropteran family. Cimicomorpha also hold the Reduviidae, whose 6900 species make it the second-largest heteropteran family. Cimicomorphans are basically (primitively) predacious, but several groups have secondarily become phytophagous: the small family Thaumastocoridae (19 species) feeds on palms; the Tingidae (2100 species), or lace bugs, feeds on leaves of many plants and may occasionally become serious pests; and most members of the Miridae feed on plants, although many mirids are predacious. It is possible that the evolutionary success of Miridae (as reflected in the family's many species) derives from mirids' association with annual weedy plants, plants that grow rapidly and yield to other species; mirids, too, develop rapidly, and many are quite host specific. Many of the predacious Miridae are site-specific, living on a single species of plant. If the plant should be of economic significance, these mirids can be useful in biological control.

Of the wholly predacious groups, the assassin bugs, Reduviidae, are the most important. Some are specialists (unusual for predators), one group feeding on millipedes, another living in spider webs and stealing food from the spider; yet another feeds on termites and sometimes lures them from the mound. The group's biocontrol potential has been little studied, except recently in southern India (see Schaefer and Panizzi, 2000; Ambrose, 1999). As mentioned above, members of the subfamily Triatominae feed on vertebrate blood, and the neotropical members transmit the Chagas disease trypanosome to humans.

All 110 species of another cimicomorphan group, Cimicidae, feed on vertebrate blood, especially on that of cave-dwelling vertebrates such as bats, communally nesting birds – and (earlier) man. The bedbugs (*Cimex lectularius* L. and *C. hemipterus* (Fabricius)) are parasites of man. *C. lectularius* feeds only on man and is one of a very few insects that occur *only* with man (the housefly and the head and body louse are among the others). *Cimex hemipterus* is at times a pest of poultry. *C. lectularius* is primarily a temperate-zone species, although it extends into the Tropics; *C. hemipterus* is tropical.

Other important cimicomorphan groups include the damsel bugs (Nabidae), a worldwide family of nearly 400 species, common in crop fields in North America and often studied for their biocontrol potential; and Anthocoridae, minute flower bugs, a group (440 species), also worldwide, of very small and often strikingly patterned insects much studied for their control of pests, especially in greenhouses.

The seventh infraorder, Pentatomomorpha, is also large, with at least 16,200 species. One of its families is third largest in the Heteroptera: Pentatomidae, with about 4700 described species (Note: In what was once called the "Lygaeidae," there were nearly 3600 species, which would make this the fourth largest family in Heteroptera. However, many of these subfamilies have been raised to family rank, and are now part of the superfamily Lygaeoidea (Henry, 1997).) Pentatomomorpha is one of only two infraorders that appear to have arisen as plant feeders, although one important subgroup, the subfamily Asopinae (family Pentatomidae), is secondarily predacious and important in biocontrol programs. Most pentatomomorphs feed on the nitrogen-rich reproductive parts of plants, especially on their ripe and ripening seeds. Pentatomomorpha is also the most difficult infraorder to spell correctly.

A major pentatomomorph group is the Lygaeoidea (see above), or seed bugs and milkweed bugs; most of these feed on seeds, although a few groups feed on grasses and members of one of these, Blissidae (435 species) or chinch bugs, sometimes become pests on wheat and grass (including turf). Another major group is the Pentatomidae (stinkbugs) (4700 species), which also feed on reproductive parts, although many others feed on somatic tissue. Included here is the southern green stinkbug, *Nezara viridula* (L.), a major pest of many crops across the world. Here also are the Sunn pests, a group of pentatomids and some Scutelleridae (a family closely related to Pentatomidae; 450 species) that ravage wheat and barley throughout the Middle East and surrounding areas. Several small families are related to Pentatomidae, including the Cydnidae (765 species), many of which live in the soil and suck from the roots of plants. Members of Scutelleridae are often brilliantly and iridescently colored (although Sunn pests are not). Females of many Acanthosomatidae (nearly 200 species) guard their eggs from parasites and their young from predators.

The Coreidae (leaf-footed bugs) (nearly 1900 species) and Pyrrhocoridae (cotton stainers) (340 species) are related to Lygaeoidea. The former includes the squash bug (*Anasa tristis* (De Geer)) and a group of bugs with expansions on their legs ("leaf-footed" bugs); some of the latter (in the Neotropics) are brilliantly colored, and others (in North America) come to houses for warmth when winter sets in. The largest genus in Pyrrhocoridae is the tropical *Dysdercus*, many of whose species are serious pests of cotton, which they damage in part by direct feeding on the seeds. Greater damage is caused by their feeding punctures' providing entry into the cotton boll of boll rot disease organisms; these organisms, and the bugs' excreta, destroy and stain the cotton fibers (hence "cotton stainers"). One other species, the European *Pyrrhocoris apterus* L., is famous for its role in the discovery of "paper factor," a discovery that led to the development of natural and artificial juvenile hormone analogs, useful in insect control.

The last (eighth) infraorder is Aradomorpha (about 1940 species), also a plant-feeder, and until recently included as a member (superfamily Aradoidea) of the Pentatomomorpha, from whose other members it differs in several significant respects. Among these are the long mouthparts and the lack of long sensory hairs arranged in characteristic patterns on the underside of the abdomen; pentatomomorphans have shorter mouthparts (like those of other heteropterans) and have these abdominal sensory hairs (unlike most other heteropterans).

The Aradomorpha are a small group of brownish to grayish flattened bugs, nearly all of which live under the bark of dead or dying trees, where they feed on the long mycelia of fungi; adapted for this are the very long very slender mouthparts, which at rest are kept coiled in a special pouch within the head. These bugs range from about 3 to 10 mm long and are often wingless; the upper surface of the body is frequently "bumpy" or pebbly and the legs and antennae are short. One species, *Aradus cinnamomeus* Panzer, differs from the others and feeds on the sap of several pine species; from time to time it becomes a serious pest of commercial pines in northern Europe. It has a 2-year life cycle, and the discovery that some populations reproduce in oddnumbered years, and others in even-numbered years, has led to interesting biogeographical and population studies. Nine other species (family Termitaphididae), distantly related to all the others (family Aradidae), live in termite nests, where they lay their eggs among the termites. Only a few millimeters long, these bugs lack an ovipositor, eyes, and wings and probably feed on fungi within the nests. Very little indeed is known about them.

Conclusions

Heteropterans share fundamental properties of structure and biology, especially of feeding: their elongated mouthparts so well adapted for the sucking up of the fluids of other organisms. Yet the diversity of heteropterans is remarkable, in size, form, habitat, and structure. Groups now considered rare are probably in fact common, as collecting in habitats hitherto disregarded has shown: Enicocephalomorpha and

Dipsocoromorpha are probably far more common than we believe, but live in soil and debris habitats in the Tropics, habitats no one studies. Recent work on tree-canopy habitats in the Tropics has revealed assemblages of unknown or poorly known heteropterans found (apparently) nowhere else. Too many entomologists have for too long scorned insects that cannot be collected with nets. Once such other collecting methods as searching the ground's surface became common, so did groups (like Lygaeoidea) once thought relatively scarce.

It follows that the diversity of Heteroptera, like the diversity of all other groups, is constrained not by evolution but by man's knowledge and by man's willingness to seek diversity in different habitats: For the diversity is there.

What the author has sketched above is based in truth on a remarkably small number of observations and experiments on a remarkably small number of heteropterans. One hopes those observations are accurate and those heteropterans are representative. Much more work, on many more heteropterans, as well of course as much more collecting, will make even more clear just how diverse and fascinating true bugs are.

See also: Insects, Overview

References

Ambrose DP (1999) *Assassin Bugs*. Enfield, New Hampshire: Science Publisher.

Burckhardt D (2009) Taxonomy and phylogeny of the Gondwanan moss bugs or Peloridiidae (Hemiptera, Coleorrhyncha). *Deutsche Entomologische Zeitung* 56: 173–235.

Cohen AC (1998) Solid-to-liquid feeding the inside(s) story on extra-oral digestion in predaceous Arthropods. *American Entomologist* 44: 103–116.

Henry TJ (1997) Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea. *Annals of the Entomological Society of America* 90: 275–301.

Henry TJ (2009) Biodiversity of Heteroptera. Chapter 10. In: Foottit RG and Adler PH (eds.) *Insect Biodiversity, Science and Society*, 1st edn. Hoboken, New Jersey: Wiley-Blackwell. (Note: An account of the insects themselves).

Javahery M (1995) *A Technical Review of Sunn Pests*. United Nations, Regional Office of the Near East, Cairo, Egypt: FAO.

Schaefer CW (2009) Prosorrhyncha (Heteroptera and Coleorrhyncha). In: Resh VH and Cardé RT (eds.) *Encyclopedia of Insects*, 2nd edn. San Diego, CA: Elsevier. (Note: An account of the insects themselves).

Schaefer CW and Panizzi AR (2000) *Heteroptera of Economic Importance*. Boca Raton, FL: CRC Press.

Schofield CJ (1994) *Triatominae: Biology and Control*. West Sussex, England: Eurocommunica Publications.

Schuh RT and Slater JA (1995) *True Bugs of the World (Hemiptera Heteroptera)*. Ithaca, NY: Cornell University Press.

Weber H (1930) *Biologie der Hemipteren: Fine Naturgeschichte der Schnabelkerfe*. Berlin: Verlag-Springer.

U

Ultraviolet Radiation

Andrew R Blaustein, Oregon State University, Corvallis, OR, USA
Catherine Searle, Georgia Institute of Technology, Atlanta, GA, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Andrew R. Blaustein and Nareny Sengsavanh, volume 5, pp 723–732, © 2001, Elsevier Inc.

Glossary

Chlorofluorocarbon (CFC) A type of hydrocarbons that are nonreactive, nonflammable organic molecules in which both chlorine and fluorine atoms replace some of the hydrogen atoms. Use of CFCs is one of the primary causes of stratospheric ozone depletion.
Montreal Protocol A treaty signed by a number of nations to curtail CFC production.
Ozone A form of oxygen in which three atoms of oxygen occur together. Ozone forms a natural layer in the stratosphere that blocks living organisms from harmful ultraviolet radiation from the sun.
Photolyases DNA repair enzymes that repair damage caused by exposure to ultraviolet light.
Ultraviolet radiation Radiation with wavelengths shorter than violet light and with more energy.

Properties of Light

Light is a natural phenomenon that allows us to see objects, shapes, and colors. It radiates from the sun, the stars, a flame, or a light bulb. Light encompasses a broad spectrum of radiation that includes gamma rays, ultraviolet (UV) rays, infrared rays, microwaves, and radiowaves (Figure 1). These types of radiation are all classified as nonvisible light and consist of the longest and shortest wavelengths and frequencies. A small section of the spectrum is the visible light spectrum. The wavelengths of this region range from 380 to 750 nanometers (nm) and are the only part of the electromagnetic spectrum (Figure 1) that can be detected by the human eye. When light shines on an object, some of the light is absorbed and converted to heat; another percentage is scattered or dispersed in various directions; some are transmitted; and the rest may be reflected depending on the material on which the light strikes.

UV Radiation

The electromagnetic spectrum is continuous, but the types of electromagnetic radiation do not begin or end at precise points along the spectrum. For example, red light shades into invisible infrared (below red) radiation and violet light shades into invisible UV (beyond violet) radiation. The sun is the major source of UV radiation for the Earth. UV is responsible for producing suntans and vitamin D in the human body. In humans, overexposure to UV can lead to serious and sometimes irreparable harm. It can cause mutations in cellular DNA, which can ultimately lead to significant alterations in cells, the main cause of cancer. Other damaging effects of UV include premature aging, blindness, and sterilization. UV radiation also has the ability to kill microorganisms, plants, and animals.

UV can be divided into four wave bands. These are vacuum UV (<200 nm), UV-C (200–280 nm), UV-B (280–315 nm), and UV-A (315–400 nm). At the earth's surface, vacuum UV and UV-C are not present because of their absorption by various gases such as oxygen and ozone. The formation of atmospheric oxygen and a stratospheric ozone layer was essential for the evolution of life on earth. The ozone layer shields the terrestrial surface from harmful UV radiation. Unfortunately, through anthropogenic emissions of chlorofluorocarbons (CFCs) and other gases, the ozone layer has been adversely affected. It has thinned and has developed "holes" in polar regions. Thus, there is potential for increased UV radiation hitting the earth's surface.

The ozone hole over Antarctica has dramatically increased since its discovery in the 1970s. At midlatitudes, ozone levels have also continued to decrease. Stratospheric ozone levels are at their lowest point since measurements began, so the current UV-B radiation levels are thought to be close to the maximum. Global ozone measurements from satellites from 1979 to 1993 show increases in UV-B radiation at high and midlatitudes of both hemispheres, but only small changes in the tropics. These estimates assume that cloud cover and pollution

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00147-7

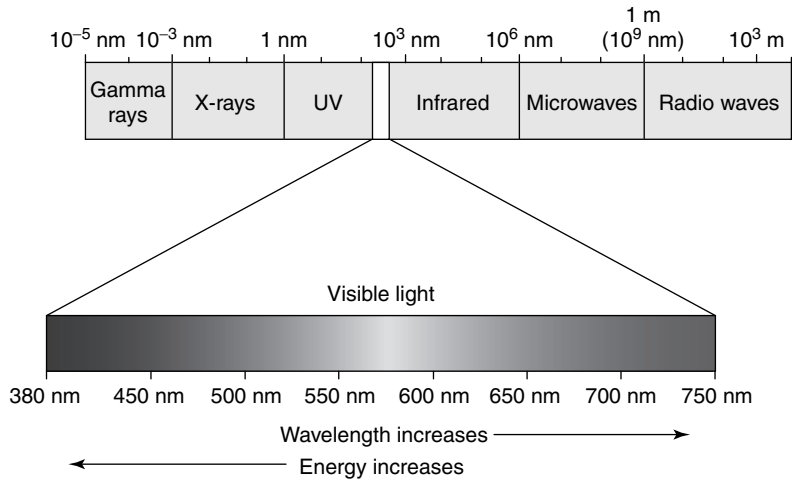


Figure 1 The electromagnetic spectrum. Visible light and other forms of electromagnetic energy radiate through space as waves of various lengths. Reprinted from figure 10.5 in Campbell NA (1996) Biology, 4th edn, with permission from Benjamin Cummings.

have remained constant over the years. Increases in surface erythemal (sunburning) UV radiation relative to values obtained in the 1970s are estimated to be about : 7% at Northern Hemisphere midlatitudes in winter and spring and 4% in summer and fall; 6% at Southern Hemisphere midlatitudes throughout the year; 130% in the Antarctic in spring; and 22% in the Arctic in spring. The resulting increases in UV radiation threaten humans, animals, plants, and microorganisms in both terrestrial and aquatic ecosystems.

The pioneering work of Rowland and Molina showing that CFCs were responsible for the decrease of stratospheric ozone caused considerable concern among atmospheric and environmental scientists. They reasoned that CFCs in the stratosphere would be subjected to intense UV radiation, which would break them apart, releasing free chlorine atoms.

All the chlorine of a CFC molecule would eventually be released due to further photochemical breakdown. The free chlorine atoms could then damage stratospheric ozone, forming chlorine monoxide (ClO) and molecular oxygen. Molecules of ClO can react to release more chlorine and an oxygen molecule.

These reactions are part of the chlorine cycle because of the continuous regeneration of chlorine as it reacts with ozone. Chlorine is the catalyst in the reaction because it promotes the chemical reaction without being used up. Because chlorine can last for 40–100 years, every chlorine atom in the atmosphere can potentially break down 100,000 molecules of ozone. CFCs are especially damaging because they are transport agents that constantly move chlorine atoms into the stratosphere and because chlorine atoms are removed from the stratosphere very slowly.

By 1978, most Scandinavian countries and the US banned the use of CFCs in spray cans. In 1983, most European countries proposed a voluntary reduction in the use of CFCs. As a result of increasing concern over the effects of ozone depletion, an international agreement was signed to reduce and eventually eliminate certain anthropogenic ozone-destroying substances. A treaty, known as "The Montreal Protocol," was signed by a number of nations on 16 September 1987, in Montreal, Quebec. The protocol which consisted of a plan that would dramatically cut CFC production was recognized as the first worldwide effort to solve a massive environmental problem. Since 1987, more than 150 countries signed the agreement to phase out all the uses of CFCs by 2000. Despite this agreement, the worsening environmental news about CFCs and ozone depletion in the 1990s led certain nations to adopt stricter measures, limiting CFC production.

Several industrial companies have developed CFC substitutes. CFCs and several other chemicals that contribute to ozone depletion have been phased out in the US and several other countries. Nevertheless, existing stockpiles of these chemicals can be used until the deadline. Moreover, developing countries are on a different time course and will phase out CFCs by 2006. Unfortunately, CFCs are very stable and estimates suggest that those in use today will continue to deplete stratospheric ozone for 50–150 years. Atmospheric scientists believe that the Antarctic ozone hole will reappear each year until about 2050. However, because of the Montreal Protocol and other measures to limit ozone-depleting substances, there is some optimism about ozone depletion being curtailed.

Bromine (Br) is another important source of ozone depletion. Br is present in smaller quantities than chlorine but is more destructive on an atom-to-atom basis. The largest source of Br is methyl bromide, which comes naturally from the oceans and wildfires. However, a large portion of methyl bromide is human-made and is used as a fumigant. Another important source of bromine is fire extinguishers. Like CFCs, methyl bromide has a long atmospheric lifetime (approximately 72 years). Regulations on the use of methyl bromide are being debated.

Measurement of Environmental UV Radiation

Although significant advances have been made, measuring UV radiation is difficult. There are significant differences in how

various devices measure UV radiation. Moreover, there are a number of factors that affect UV radiation. For example, there are geographical and seasonal variations in UV. Thus, recent studies have confirmed that there are generally higher UV-B levels at lower latitudes in the US. Other measurements confirm latitudinal differences in Europe, Asia, and New Zealand.

Spectral measurements show higher summer values of both UV-A and UV-B radiation in New Zealand and Australia compared with Germany due to the yearly cycle of the sun–Earth distance and due to the lower stratospheric ozone levels in the southern hemisphere and higher air pollution levels in Germany. Additional UV data are being accumulated for midlatitudes, and eventually, a more complete picture of the UV situation at these latitudes will emerge.

Cloud cover also affects UV-B radiation levels on the earth's surface. Data from several locations in the US suggest that the monthly average UV levels are reduced by 10–50% by cloud cover, depending on the location and season. Aerosols (small particles suspended in the air) may also reduce UV levels in polluted areas. The magnitude of this effect is highly variable and may depend on the number of particles and their chemical and physical composition.

UV levels are expected to increase with increasing surface elevation above sea level due to a thinner atmosphere overhead. Measurements in a remote area of Chile, for example, showed increases of 4–10% per km in elevation. Other locations showed larger vertical gradients of up to 40% per km near Santiago, Chile, and 9–23% per km in the Swiss Alps.

Because of the high spatial and temporal variability of surface UV radiation and the difficulty of maintaining calibration within instruments, it is difficult to provide global UV climatology and represent long-term UV trends based on ground monitoring stations alone. However, satellite-based measurements can provide global coverage and continuous long-term monitoring. Yet, it is difficult to use remote satellite data in an attempt to estimate UV levels from specific microhabitats on earth. The derivation of surface UV irradiance from satellites is indirect because satellites detect radiation reflected by the atmosphere and the surface of the earth. The use of radiative transfer models is necessary to relate transmission, reflection, and atmospheric absorption. These models have been useful in showing general changes in UV radiation reaching the surface, and are computed for clear skies using satellite ozone measurements.

Effects of UV Radiation on Biological Systems

DNA Repair

At the terrestrial surface, most UV radiation of biological concern is in the 280–315 nm (UV-B) band. UV light induces a variety of photoproducts in DNA. UV photoproducts impede gene expression by blocking transcription, and they can cause cell death or mutagenesis. There are a number of responses to DNA damage by living cells. The two fundamental processes for repairing UV-induced DNA damage are direct reversal of DNA damage or its excision. These processes can be explained by analogy. If we consider a rope with a knot, we can either undo the knot (direct reversal of damage) or cut the knot out and replace the rope with a new piece (excision of the knot). This descriptive analogy roughly represents how photoreactivating enzyme, photolyase, and excision repair, respectively, work to undo DNA damage.

Thus, harmful photoproducts can be removed by photoreactivating enzyme, photolyase. In some species, this is the most important mechanism for DNA repair. In excision repair, when DNA molecules are damaged, a segment of the strand of DNA containing the damage is cut out by one repair enzyme, and the resulting gap is filled in with nucleotides (building blocks of DNA) properly paired with nucleotides in the undamaged strand.

Effects of Increased Solar UV Radiation on Terrestrial Plants

Obviously, it is extremely important to understand the effects of increasing solar UV radiation on terrestrial plants. Plants comprise most of the living matter in terrestrial ecosystems. Damage to plants within natural ecosystems and in agricultural systems can have far-reaching consequences on other organisms, including humans.

Both the physiological and developmental processes of plants are affected by UV-B radiation. Several studies have shown that some plant species in greenhouses, growth chambers, and in the field have reduced growth and leaf area under close to ambient UV-B conditions compared with plants grown under reduced UV-B levels. Increased UV-B radiation may greatly affect photosynthesis. In certain species, such as soybean, sunflower, and corn seedlings, solar UV-B radiation reduced photosynthesis by about 15% when a 12% ozone depletion was simulated. UV-B radiation can also alter the time of flowering and the number of flowers in certain species. Alteration of flowering time can have a severe impact on plants because pollinator availability may be subsequently affected. Anther walls can absorb more than 98% of incident UV-B radiation, and pollen walls contain UV-B-absorbing compounds, providing some protection from the negative effects of UV-B. However, after the transfer to the stigma, pollen may be directly exposed to UV-B radiation. Several experimental studies using Mylar or glass filters that shielded plants from UV-B radiation showed that flowering was enhanced under these regimes. Different studies have shown the different effects of UV-B on flowering in various species. For example, flowering was inhibited in plants such as *Melilotus* and *Trifolium*, but *Zea mays* and *Sorghum* were not affected. In sexually reproducing populations of a desert plant, the effects of UV-B radiation on growth and biomass seemed to accumulate in subsequent generations that were exposed to UV-B radiation. Thus, the effects of exposure to UV-B radiation may be amplified. Even roots of plants can be affected by elevated UV-B radiation. For example, the microorganisms associated with the roots of sugar maple trees were altered when the shoots of the trees were exposed to UV-B radiation.

The yields of certain crop plants may be significantly affected by increases in UV-B radiation. The available data on the effects of UV-B radiation on yield illustrate significant interspecific variability and, also, variability among cultivars complicated by differences in how the experiments were

performed. Nevertheless, some species and certain cultivars seem to be more tolerant to UV-B than others. For example, results from greenhouse and field tests suggest that the soybean cultivar "Forrest" was more tolerant to UV-B radiation than the cultivar "Shore." A study of 10 crop species in Florida showed that under UV-B radiation, yields were reduced by 5–90% in half of them, including wheat (5% reduction), potato (21% reduction), and squash (90% reduction). In this study, rice, peanut, and corn were not affected.

Several studies have shown that plant susceptibility to herbivores and pathogens is altered by UV-B radiation. For example, a number of studies have shown that ambient UV-B radiation can reduce insect herbivory of agricultural pests and native plants. Supplementation of solar UV-B radiation in field studies can reduce the population of herbivorous insects in certain systems. It is unclear why changes in herbivory occur. It is possible that the secondary compound plant defenses may be altered, or that UV-B directly affects the herbivores. However, most of these studies suggest that the changes in insect herbivory are due to changes in host plant tissues.

Experiments conducted in greenhouses and in the laboratory suggest that viral and fungal pathogens react differently to UV-B radiation. Some studies indicate that UV-B radiation promotes the severity of disease, whereas other studies indicate a reduction in infection severity. For example, cucumber plants first exposed to UV-B radiation were more susceptible to subsequent infection by fungal pathogens. However, if exposed to UV-B radiation after infection, there was no effect on the severity of the disease. Other studies have shown that when UV-B radiation is removed, there is increased incidence of fungal infection.

Because there are differences in tolerance to UV-B radiation in different species, it is suggested that a reduction of primary productivity in one plant species may lead to an increase in primary productivity in another more UV-B-tolerant species. Thus, it is possible that the overall productivity within an ecosystem may change, even if species composition is unaffected. Even if the plant species composition does not change, individual plant form may change, which could affect how these plants compete for sunlight, moisture, and nutrients. This could lead to significant changes in the overall characteristics of ecosystems.

Effects of Radiation on Aquatic Systems

The attenuation of solar radiation on the water column is dependent on a variety of factors. The transparency of the water to UV radiation is dependent on many characteristics of the water. For example, there can be highly turbid coastal waters that do not allow much UV-B to penetrate very deeply. In comparison, some ocean waters are clear enough so that the UV-B penetration can be for dozens of meters. In the Antarctic, 1% of the solar UV-B hitting the surface has been measured at a depth of 65 m. Solar UV-B has been shown to degrade dissolved organic carbon (DOC). Increased breakdown of DOC and subsequent consumption by bacteria increases the UV-B penetration in the water column.

Across life histories, trophic groups, habitats, and experimental venues, UV-B negatively affects the growth and survival of aquatic organisms. A meta-analysis investigating the effects of UV-B across life stages found that UV-B had a greater negative effect on embryonic growth compared with growth at the larval stages. However, in amphibians, the larval stage experiences greater mortality than the embryonic stage. Additionally, across taxonomic groups, protozoans appear to undergo the greatest reduction in growth rates compared with animals or plants.

Plankton

Globally, phytoplankton is the most important producer in aquatic ecosystems. Thus, damage to phytoplankton populations will affect higher trophic levels. A number of studies in a variety of aquatic ecosystems have shown that UV-B radiation affects the growth, survival, and distribution of phytoplankton. Phytoplankton exist on the top layers of the oceans and freshwater aquatic systems. The Antarctic is especially productive in phytoplankton, and the region is significantly impacted by UV-B because of the Antarctic ozone hole. Therefore, a number of studies on the effects of UV-B on phytoplankton have been conducted in that region. In certain experiments, productivity was two to four times higher in tanks where UV-A and UV-B were excluded. Pigmentation was also affected. *In situ* incubations of natural phytoplankton assemblages in the Antarctic waters indicated that photosynthesis was impaired by about 5% under the ozone hole. A similar result was found in the tropics. Screening of most UV <378 nm resulted in a 10–20% increase in photosynthesis. However, no significant decrease in stratospheric ozone have been observed in the tropics.

Many phytoplankton can actively move to different positions within a particular habitat. They may do this via flagella, cilia, or by utilizing buoyancy to adjust their position in the water column. Various chemical, magnetic, light, and gravity cues influence movement so that plankton can maintain specific positions within the water column. To cope with constantly changing environmental conditions, these organisms must constantly adjust their positions. If UV radiation affects motility, or the ability of phytoplankton to respond to external cues, this may negatively affect their growth and survival. There is growing evidence that many phytoplankton species are under stress from ambient levels of UV radiation.

Macroalgae and Seagrasses

In contrast to the motile phytoplankton, macroalgae and seagrasses are attached to their growing sites. Thus, they are restricted to certain depth zones. It is thought that this zonation is caused, at least in part, due to limits of visible light penetration at various depths. Some species may be more tolerant to solar radiation than others. Thus, increased levels of UV-B radiation may expose algae and seagrasses to levels that they have not encountered, perhaps affecting growth and photosynthesis. Several studies have shown that UV-B radiation inhibits photosynthesis in many of the red, brown, and green algae. Deep water algae were the most affected, whereas intertidal algae were the least sensitive.

The DNA of algae appears to be poorly shielded when compared to higher plants. For example, doses of UV-C radiation (at 254 nm) necessary to kill leaves of higher plants appear to be about four orders of magnitude greater than that

necessary to kill highly resistant algae. Flavenoids, highly effective UV-screening compounds found in higher plants, have not been found in algae. In higher plants, flavenoids exist in high enough concentrations in the epidermis of leaves so that in combination with cuticle waxes and other cell-wall components, the incidence of UV radiation is reduced by one or two orders of magnitude. However, algae do produce other UV-absorbing substances, including a yellow protein–carotenoid complex in some species and other substances that afford some protection from UV radiation.

An experimental study conducted in Canada illustrates how ecosystems may have complex responses to increased solar UV-B radiation. Solar UV radiation can reduce photosynthesis and growth in bottom-dwelling algal communities in shallow freshwater. However, in this study, it was observed that greater amounts of algae were accumulated in UV-exposed habitats than in UV-protected environments. UV-A and UV-B radiation inhibited insects that feed on algae. Because larval algal consumers are more sensitive to UV radiation than algae, algal abundance increased.

Invertebrates

Sunlight can be lethal to a wide variety of marine and freshwater plankton exposed for just a few hours or a few days. Mortality rates are usually lower in zooplankton that contain photoprotective compounds including pigments derived from dietary plant carotenoids, melanin, and substances known as mycosporine-like amino acids that gets absorbed in the UV-A and UV-B range. The presence of these photoprotective compounds suggests that there is significant selection pressure associated with the harmful effects of UV radiation.

A number of experimental studies have shown that natural levels of UV radiation are lethal to zooplankton. For example, one study in Pennsylvania showed that zooplankton communities exposed to ambient levels of UV-B for 3 days experienced significant mortality when they were not shielded from UV-B radiation. However, mortality was significant only in an oligotrophic lake, not in an eutrophic lake where light penetration is not as high.

Marine invertebrates differ considerably in their sensitivity to UV-B radiation. For example, while one species of crustacean may suffer about 50% mortality at current levels of ambient UV-B radiation at the sea surface, some shrimp can tolerate irradiances higher than those predicted for a 16% ozone depletion. Bottom-dwelling invertebrates of the ocean may also be affected by UV-B radiation. For example, cleavage in sea urchins is impaired by UV radiation. Marine organisms associated with coral reefs, such as sponges, bryozoans, and tunicates, are also adversely affected by UV-B radiation.

Corals are affected by UV radiation in a number of ways. Depending on the species and the particular ecosystem, a number of studies illustrate that UV radiation can cause death, inhibit growth, and contribute to coral bleaching. Bleaching, a well-recognized phenomenon, occurs principally via the loss or expulsion of symbiotic algae from the coral host tissue. When the algae are lost, the coral loses its characteristic color and the remaining white polyp is most noticeable. Although a number of events, including elevated seawater temperature and heavy rains, can contribute to bleaching, there is some evidence that UV radiation may also play a role in certain bleaching events.

UV radiation can inhibit photosynthesis in symbiotic algae. UV radiation affects respiration among corals and their symbiotic algae, but not in a consistent way. In some species, respiration increases under UV radiation, whereas in other species it may decrease and in others respiration may be unaffected.

UV radiation inhibits the growth of symbiotic algae in culture. Reproduction in corals may be affected by UV-B radiation. Broadcast spawning at night, a nearly universal phenomenon among many reef invertebrates and corals, might be related to avoiding UV-induced DNA damage as well as to reducing predation. In at least one study, it was observed that the larvae of reef corals from shallow water were more resistant to UV-B radiation than those from deeper water.

In the Antarctic, a number of marine invertebrates sustain UV-induced DNA damage. These include worms and crustaceans. Krill, copepods, and gelatinous zooplankton generally have transparent eggs, larvae, or adult stages that are pelagic and planktonic, and are often found in surface water for several months. Thus, these species are especially vulnerable to DNA damage from elevated UV-B exposure. These species are important components of the ecosystem, and damage to them could have significant consequences on the ecosystem.

Fishes and Amphibians

A study of Antarctic zooplankton, including larval fish, showed that they sustain DNA damage during periods of increased UV-B flux. Fish larvae in Antarctic marine ecosystems sustained UV-B-induced DNA damage greater than the lethal limit determined for Antarctic diatoms, and comparable to the lethal limit of damage for cultured goldfish cells. DNA damage has been shown to be especially correlated with daily UV-B flux in icefish eggs. Icefish larvae, however, showed patterns of DNA damage that correlated less significantly with daily UV-B flux. Antarctic fish appear to primarily use photolyase to remove harmful UV-B-induced photoproducts.

Other fish species may also be affected by UV-B radiation. For example, in the Arctic ecosystem, many economically important fish species, including cod, pollock, herring, and salmon, spawn in open shallow water and are subjected to increased solar UV-B radiation. Because many of their eggs are found near the surface, it is possible that marine fish productivity could decline in this region due to increased UV-B radiation. At this point in time, however, it is difficult to assess the impact of UV-B radiation on Arctic fish productivity. However, a laboratory experiment demonstrated that salmon exposed to UV-B radiation are more prone to fungal infections and skin lesions.

Several laboratory studies have shown that UV radiation affects the growth, development, and hatching success of certain amphibian species. These studies have shown that under simulated UV light of various intensities, amphibians may develop skin lesions, edema, eye damage, curvature of the body, and behavioral abnormalities. Under relatively low-level but prolonged doses of UV radiation, the mortality of embryos increases compared with controls that are shielded from UV radiation.

Amphibian Declines and UV Radiation

Amphibian populations are undergoing a serious decline in various regions of the world. Unfortunately, the causes for amphibian population declines have been difficult to assess. Much of the information on amphibian declines comes from observational or anecdotal accounts. Hypothesized causes for the declines include habitat destruction (the most obvious cause), climate change, pathogens, introduced exotic species, pollution, and increased UV radiation. These agents may act alone or in combination to contribute to the decline of amphibian populations.

The diversity of locations where amphibian populations have declined prompted consideration of atmospheric factors such as increased UV irradiance associated with depletion of stratospheric ozone. Several investigators have used field experiments to examine the potential role of UV-B radiation in amphibian population declines by measuring the mortality of embryos that were shielded from UV radiation compared with the embryos that were unshielded. Continuous high mortality in early life stages may ultimately contribute to a decline in the population level.

Field experiments from North America, Europe, and Australia show that ambient UV-B damages the embryos of certain amphibian species but not others. Results of these experiments by several different investigators strongly indicate that the hatching success of at least nine species of amphibians, from widely separated locales, is reduced under ambient UV-B radiation. This includes a taxonomically diverse group of two frog species, one toad species, two salamander species, and a newt from North America; two frog species from Australia; and a toad species from Europe. Some of these species are found in montane areas, and others are found at sea level. A key characteristic shared by these species is that they often lay their eggs in shallow water, where they are exposed to solar radiation.

Hatching success of several other amphibian species in North America, Australia, and Europe was not affected by UV-B radiation. This is not surprising because several studies have demonstrated differential sensitivity of amphibians to various abiotic factors. There may be variation in response to UV-B radiation, perhaps even within a species, at different locations. For example, embryos of western toads in Oregon are sensitive to ambient levels of UV-B, while those of a different subspecies in Colorado are unaffected.

Based on a limited sample, there is a correlation between resistance to UV-B and the activity of photoreactivating enzyme, photolyase. Species with the highest photolyase activities seem to be more resistant to UV-B radiation. Furthermore, of the species examined, frogs and toads generally show more photolyase activity than salamanders. It is also possible that nuclear excision repair is also being used to counter UV-induced DNA damage, but this has not been measured in amphibian species taken from the wild.

In nature, more than one environmental agent may affect an animal as it develops. This also seems true for amphibians developing at their natural field sites. Field experiments have been conducted to examine at least three factors that seem to interact synergistically with UV-B: a pathogenic fungus, low pH, and fluoranthene, a polycyclic aromatic hydrocarbon that may pollute aquatic environments impacted by petroleum contamination. A meta-analysis investigating the combined effects of UV-B with pH, contaminants, and disease found that UV-B radiation interacted synergistically with other stressors, resulting in greater-than-additive effects on survival. In certain combinations with UV radiation, these three agents increase mortality to levels greater than that contributed by UV radiation alone.

Obviously, UV-B cannot be regarded as contributing to all amphibian population declines. For example, UV-B radiation is unlikely to contribute to mortality in amphibians that are primarily nocturnal, live under dense forest canopies and lay eggs hidden from solar radiation, or have high photolyase activities.

Amphibian Deformities and UV Radiation

Reports of deformed amphibians have been given wide media attention, with the most common reports of deformities being frogs and toads with extra or missing limbs. Three major agents are being examined: pesticides, UV radiation, and parasitic trematode infection. However, only limited data are available on this subject based on experimental work. One laboratory study of northern leopard frogs showed that on exposure to 24 h of simulated ambient UV light, frogs developed hindlimb malformations. It remains unclear as to whether UV radiation can cause extra limbs in wild amphibian species. However, other deformities in wild species have been observed. These include edema, curvature of the spine, and lesions in larvae and newly metamorphosed amphibians. One report showed severe retinal abnormalities consistent with UV damage in a basking frog species.

Effects of UV Radiation on Biogeochemical Cycles

The effects of increased UV-B radiation on biogeochemical processes may be complex. For example, increased UV-B radiation may affect the magnitude and direction of trace gas on emissions and mineral nutrient cycling in terrestrial systems. Moreover, the effects on these processes may be species specific and may vary in different ecosystems. Increased UV-B radiation may alter the chemical composition of plant tissue, the photodegredation (breakdown) of dead plant matter, the release of carbon monoxide, and the activity of microbial decomposers, and nitrogen-fixing organisms.

In aquatic ecosystems, increased UV-B radiation may affect the processes that produce organic matter and the processes that degrade organic matter. There may be photodegradation of dissolved organic matter, which may lead to the production of organic acids and ammonium. Photoinhibition of surface aquatic organisms may also affect biogeochemical processes.

The potential effects of increased UV-B radiation on terrestrial and aquatic carbon, nitrogen, sulfur, oxygen, and metal cycles have been explored in a number of studies. These effects and similar effects on biogeochemical cycles in the atmosphere may aid or help impede the buildup of greenhouse gases and aerosols in the atmosphere.

Effects of UV Radiation on Humans

UV-B radiation does not penetrate deep into the body because most of it is absorbed in the superficial tissue layers. Therefore,

much of the UV radiation affects the skin and eyes. However, there are also systemic effects. UV-B is the main cause of sunburn and tanning and the formation of vitamin D_3 in the skin. UV-B also affects the immune system. UV-B can cause snowblindness and is a significant factor causing cataracts. It also contributes significantly to the aging of the skin and the eyes and is effective in causing skin cancer.

Eyes

Photokeratitis is the effect most attributable to exposure to UV radiation. It is similar in effect to a sunburn and often occurs after short-term exposure to UV radiation. The eyeball becomes inflamed and reddens, often accompanied by pain and photophobia (fear of light). This is frequently diagnosed in skiers as snowblindness.

Exposure of the eye to UV radiation can also have effects on the cornea, contributing to degeneration of its fibrous layer. Under certain conditions, exposure to UV radiation can cause an outgrowth of the outermost mucous layer over the cornea, resulting in the loss of transparency. Squamous cell carcinoma (SCC) of the cornea, a malignant neoplasm, can also be found after exposure to solar radiation. These diseases are associated with outdoor living or working near areas of high reflectance (e.g., near water, concrete, and sand).

Cataracts are the leading cause of blindness in the world. They are characterized by a gradual loss in the transparency of the lens of the eye due to oxidized lens proteins. This can lead to blindness unless the affected lens are removed. There is a correlation between certain types of cataracts and exposure to UV-B radiation. Several studies have suggested that the relative risk associated with increased exposure to sunlight and cortical cataracts (those that develop in the outer layer of lens protein) is between about one- and threefold. One model suggests that a sustained 10% loss of ozone worldwide would lead to an additional 30,000 blind people per year.

Sunburn

Sunburn is the most common effect of exposure to UV radiation. This results in the reddening of the skin and possibly blistering. Sensitivity to sunburn and tanning varies with pigmentation. Heavily pigmented individuals are less sensitive to sunburning than less heavily pigmented individuals. There are various categories with regard to sunburn and tanning sensitivity. The most sensitive individuals (skin type I) develop a moderate to severe burn and usually do not tan within an hour of exposure to the summer sun. These persons usually have very fair freckled skin, red or blond hair, and blue eyes. The most resistant individuals (skin type VI) are darkly pigmented and become more pigmented after exposure to the sun.

Photoaging

Exposure to sunlight ages the skin. This is known as photoaging and is characterized by wrinkles, altered skin pigmentation, and an overgrowth of abnormal elastic fibers in the dermis.

Skin Cancer

There are various types of skin cancers. These are basal cell carcinoma (BCC), SCC, and cutaneous melanoma (CM). The carcinomas of the skin are often referred to as the "nonmelanoma skin cancers (NMSC)." The NMSCs are clearly correlated with sunlight. They occur primarily in light-skinned individuals, and usually on the areas of the body most exposed to sunlight.

BCC is the predominant form of NMSC in light-skinned individuals. Individuals most susceptible to BCC are those with the lightest skin and a poor tanning ability. The incidence of BCC in light-skinned populations has been increasing in certain regions. Although it was originally thought that cumulative lifetime exposure to sunlight was directly related to developing BCC, recent information suggests that this may not be the case.

SCC is much less common than BCC but is much more common than CM in the US. Epidemiological data suggest that cumulative lifetime exposure to UV is a critical risk factor for developing SCC compared to other types of skin cancers. Several studies found an increase in the risk of developing SCC with incidence of childhood sunburns. However, this may be related to a high level of childhood exposure to sunlight rather than the number of sunburns *per se*.

CM is relatively rare compared with BCC or SCC. It accounts for only about 2–3% of the skin cancers associated with solar radiation, but it also accounts for most of the mortality. Like BCC, there does not seem to be a clear relationship between developing CM and cumulative lifetime exposure to UV radiation. Furthermore, CM often appears on areas of the body that are not the most heavily sun exposed. Several epidemiological studies have shown that exposure to sun during childhood increases the risk of developing CM. An additional risk factor is the appearance of freckles or moles.

Effects on the Immune System

In humans, the skin is the first line of defense against foreign bodies that may threaten an individual's health. Thus, the skin in combination with the immune system helps maintain health against infectious diseases, cancers, and parasites. The skin incorporates a number of cells from the immune system that can mount or influence immune responses to foreign substances. Substances entering the body, such as viruses, have to be "recognized" by the immune system as either "self" or "nonself" (foreign) entities. UV radiation can induce photochemical changes in the skin and potentially alter cell surface proteins that are used to determine "self" from "non self" entities. Thus, UV radiation can be immunosuppressive. The immunosuppressive effects of UV-B radiation can influence the outcome of melanoma and NMSCs, certain infectious diseases, some forms of autoimmunity, and allergy. For example, implants of UV-induced tumors in genetically identical mice are rejected in unexposed hosts but fail to be rejected by UV-exposed mice. UV-B radiation can inhibit local inflammatory responses within UV-irradiated skin. Thus, the response elicited by injection of an antigen into the skin of sensitized individuals may be diminished in UV-irradiated skin.

Some infectious agents can directly be harmed by exposure to UV radiation, whereas others are unaffected. Immunosuppression may also decrease an individual's resistance to certain infectious diseases. In animal models, human infectious diseases have been shown to be influenced by exposure to UV-B radiation. These diseases include herpes, tuberculosis, trichinella, candidiasis, leshmaniasis, listeriosis, and Lyme disease. Effects include suppression of immune responses to the organisms or their antigens, reactivation of latent infections, increased body loads of infectious organisms, decreased resistance to reinfection, and reduced survival. The impact of UV-B exposure on antigen-presenting cells in the skin suggests the possibility that UV-B radiation may exacerbate or ameliorate autoimmune diseases such as Lupus or human immunodeficiency virus (HIV). At the very least, UV-induced immune suppression could affect the course of certain diseases within human populations.

Appendix

List of Courses

1. General Biology
2. General Chemistry
3. Environmental Science
4. Photobiology

See also: Biogeochemical Cycles. Endangered Amphibians. Endangered Reptiles. Greenhouse Effect. Plankton, Status and Role of. Seagrasses

References

AMBIO (1995) Environmental effects of ozone depletion: 1998 assessment, 24: 137–196.

Bancroft BA, Baker NJ, and Blaustein AR (2007) Effects of UVB radiation on marine and freshwater organisms: A synthesis through meta-analysis. *Ecology Letters* 10: 332–345.

Bancroft BA, Baker NJ, and Blaustein AR (2008) A meta-analysis of the effects of ultraviolet B radiation and its synergistic interactions with pH, contaminants, and disease on amphibian survival. *Conservation Biology* 22: 987–996.

Blaustein AR, Kiesecker JM, Chivers DP, *et al.* (1998) Effects of ultraviolet radiation on Amphibians: Field experiments. *American Zoologist* 38: 799–812.

Häder D-P (ed.) (1997) *The Effects of Ozone Depletion on Aquatic Ecosystems*. Georgetown, TX, USA: RG Landes Co.

Tevini M (ed.) (1993) *UV-B Radiation and Ozone Depletion*. Boca Raton, FL: Lewis Publishers.

UNEP (1998) *Environmental Effects of Ozone Depletion: 1998 Assessment*. Nairobi, Kenya: United Nations Environment Programme.

Urban–Suburban Biodiversity

Elizabeth M Cook, Rebecca L Hale, and Ann P Kinzig, Arizona State University, Tempe, AZ, USA
J Morgan Grove, USDA Forest Service, Baltimore, MD, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Bottom up Those impacts or controls on urban biodiversity that originate from human actions at smaller scales, such as household or neighborhood scales. Tends to increase heterogeneity (diversity) of outcomes.

Ecosystem services The benefits people obtain from ecosystems, including provisioning services, regulating services, cultural services, and supporting services.

Ecotone A transition zone between two or more different ecosystem types.

Hinterland A region lying beyond an urban area, but often influenced by it.

Human ecology A type of ecology focused on a specific species, *Homo sapiens*. Human ecology may be thought of as a subset of social ecology, which is a life science focusing on the ecological study of various social species such as ants, bees, wolves, dolphins, or orangutans.

Social differentiation A term taken from biology to describe the specialization of functions in a society and to characterize societies over time. Theories of social change propose that increased social differentiation emerges as societies increase in size and complexity. Differentiation is frequently accompanied by a need for increased coordination and increased interdependence in larger and more complex societies.

Synanthropic species Species that live near humans, and benefit from the types of habitats humans create.

Top down Those impacts or controls on urban biodiversity that originate from human actions at larger scales, such as the level of a city or above. Tends to homogenize outcomes.

Urban ecology The study of urban systems from an ecological perspective; an emerging field within ecology that strives to understand human interactions in ecological systems in and around urban areas and to develop theories and analyses that include human communities as fundamental components of ecological systems.

Introduction

In 2010, for the first time ever, more humans lived in urban environments than rural ones. From the advent of agriculture 10,000 years ago, when humans first began gathering in ever-larger settlements, we have transformed ourselves from *Homo Rusticus* to *Homo Urbanus*.

Over this time, we have transformed our cities as well. The largest city in 3100 BCE was Memphis, Egypt, with 30,000 people. At the advent of the Common Era, Rome had nearly half a million inhabitants. The first city to reach a million was Baghdad, Iraq, toward the end of the first century (Chandler, 1987). Today, there are 21 megacities – metropolitan areas with upward of ten million inhabitants.

The ecosystems people create in cities – either intentionally or unintentionally – are unique in Earth's history. Plants from many major global biogeographic regions – having no coevolutionary history – can be found within a few yards of each other. Cities contain previously wild animal species that have become domesticated, and domesticated animals that have become "naturalized" and free-ranging members of the urban fauna. Cities also exclude certain organisms that cannot tolerate the urban environment, such as the high fragmentation or low area of suitable habitat found there. Species moving through cities – such as birds, insects, reptiles, or small rodents – encounter new biotic communities, new patterns of fragmentation, and fundamentally different resource availabilities.

There are two major reasons scientists study these unique assemblages. The first is precisely because this is the environment most people now experience. Their interactions with "nature" – and their experience of the tangible and intangible goods and services this nature can offer – will be profoundly influenced by the biodiversity found in cities. Second, these unique assemblages give scientists an opportunity to test, accept, or modify ecological theories developed in more "pristine" environments (Collins *et al.*, 2000). Since the vast majority of the world's ecosystems are now human-impacted, if not human-dominated, it is worth understanding how human interactions with the environment structures biological communities and ecological outcomes.

The Founding of Cities and Affected Biomes

Humans, through their creation of urban and near-urban environments, can profoundly affect the distribution and abundance of populations, species, communities, and ecosystems. They do so in many ways, through fragmentation, the mobilization of formerly recalcitrant or scarce resources, preferential cultivation, or destruction. These on-site or near-site disturbances are discussed more systematically later in the text. But before turning to that discussion, it is worth noting that not all biomes are equally susceptible to the disruptions caused by urban settlements. There are certain characteristics that make sites more or less desirable for dense human

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00317-8

settlements. These characteristics have changed over time as our technologies and societies have advanced, allowing acquisition of resources from afar or the defense of even physically vulnerable locations. Nonetheless, we are more apt to find cities nestled in valleys than perched on mountaintops; more likely to find ancient cities on fertile rather than barren soils; more likely to find cities situated on ecotones or "transition zones" – for instance, between the piedmont and coastal plain – because those locations allow access to more diverse resources.

Consider the appearance of the very first urban centers. They would have begun, in many cases, as villages – and later fortresses – around which agricultural fields were established. There would have to be a favorable climate, soil, and a ready supply of water. The labor required for clearing would have to be minimal. There would also have to be locally available productive grains and legumes that were a part of the first "packages" of cultivatable materials. Frequently (but not always), the emergence of agriculture as a dominant means of food procurement over hunting and gathering was aided by the availability of domesticable animals, to be used both as draft animals and as protein supplements. It is only in sites sharing these characteristics that enough food could have been produced to allow population densities to become high enough to form settlements somewhat larger than villages, on their way to being urban centers. We know, in fact, many of the places where agriculture first arose – in Southwest Asia by 8000 BC; Greece, Cyprus, and the Indian subcontinent by 6500 BC; Egypt by 6000 BC; Spain by 5200 BC; and Britain by 3500 BC. The thick forests of the wet tropical regions and the unfavorable climate of the boreal regions made these areas less favorable for the early emergence of intensive agriculture, and thus these biomes were not initially impacted by urban centers.

Later, as civilizations advanced, the need for local agricultural self-sufficiency may have been offset to some extent by the emergence of trade routes. Some urban centers would have been established in areas of particularly favorable mineral or gemstone deposits. Similarly, circumstances of chance or talent would have allowed the emergence of an artisan class within an urban center; both types of sites would have served to anchor the locations of trade routes. Other urban centers may have emerged merely because they were particularly favorable stopping-off places – just a day's journey up or down the road from a final destination – or because they were at the crossroads of two routes. Nonetheless, there would have been certain characteristics of these sites as well that meant some biomes were more susceptible to disruption from urban centers than were other biomes – those along waterways, coastways, or valley bottoms, for instance.

In current times, technology has allowed the establishment of urban centers in new locations and biomes. Near-global transportation of food products, improvements in our ability to clear land, and our ability to control indoor climate have extended urban centers into deserts, wet tropical forests, and boreal regions. Nonetheless, many of the large urban centers remain in areas of the highest terrestrial productivity. A high percentage of cities continue to be located along waterways or coasts. Much of the growth in urban centers in the coming century will come in those countries with the fastest growing populations; these countries are largely located in the subtropical or tropical regions of both the Northern and the Southern Hemispheres. Thus, when one thinks about the types of biomes – and thus the types of flora and fauna – likely to be most significantly impacted by urban centers, one should concentrate on temperate forests and grasslands, coastal and riparian zones, savanna/woodland systems, and wet and dry tropical forests. Boreal forests and montane or alpine ecosystems are less likely to be impacted. Desert and arid systems will become increasingly vulnerable as populations grow in the northern regions of sub-Saharan Africa and the North American Southwest, among other locations.

In addition to significant spatial differentiation in the distribution of cities and consequent impacted biomes, there is temporal differentiation as well. Technology not only allows the founding of urban centers in previously unsettled areas but also changes the interactions of a city with its surroundings and the geographic region over which resources are acquired or wastes dispersed. Advances in transportation allow housing expansion in areas farther removed from industrial or commercial centers; advances in architecture allow vertical rather than horizontal expansion of the city center.

Finally, the situation with respect to regional and global levels of development and trade can also profoundly influence the growth and structure of cities. Older cities can display the residual characteristics of their structures of a century ago – residential areas close to the city center, the robust industrial base, or skeleton of that industrial base – that fueled its urban growth. Newer cities, or newer sections of cities, can reflect the dramatic increase in commercial activity fueled by increasingly global and service-oriented industries and the desire for larger homes and more land that comes with economic development.

A Framework for Evaluating Biodiversity in Cities

In the past two decades, scientists have begun augmenting their study of biodiversity with analyses centered on ecosystem services – the tangible and intangible benefits people receive from ecosystems. Human interactions with the environment are often motivated by a need to secure or increase ecosystem services – for example, the food ecosystems provide, or the protection ecosystems might offer from floods or fire. Thus, understanding the ecological patterns people create means understanding the services and benefits they seek to secure.

The recently conducted Millennium Ecosystem Assessment (MA) divided ecosystem services into four categories (MA, 2005). These are:

- *Provisioning services*, which are the tangible products that humans consume, including food, fiber, fuel, and water.
- *Regulating services*, which includes the capacity of ecosystems to prevent or minimize damage from, for example, floods, fires, disease outbreaks, or climate change.
- *Cultural services*, which includes both tangible and intangible benefits provided by ecosystems, such as recreational opportunities, sense of place, or spiritual or cultural well-being.
- *Supporting services*, which are the basic ecosystem processes – such as nutrient cycling or soil formation – needed to deliver all other services.

Many ecosystem services – such as climate regulation through large-scale carbon sequestration – will primarily take place away from cities, in the world's forests and grasslands capable of storing significant carbon. But other services – such as disease regulation, natural-hazard regulation, freshwater provisioning, air-quality regulation, and recreation – need to be produced in or near cities in order to benefit urban residents. If we are to understand how changes in biodiversity alter ecosystem services, we need to understand these relationships in the ecosystems people are creating and experiencing.

Because this is an encyclopedia of biodiversity, the focus here is primarily on the ways in which people are manipulating biodiversity to achieve a set of desired services. People manipulate the biodiversity in cities to achieve certain services, and they do so both "directly" and "indirectly." That is, some species or taxa are the subjects of deliberate (direct) attempts to secure an environmental good (e.g., a pleasing yard) or avoid an environmental bad (e.g., crown fires near a home). These deliberate manipulations always have "spillover" effects, however, impacting species other than those targeted. Simple examples would include the influences of watering, weeding, and fertilizing practices on soil fauna; or the impact of domesticated animals on bird and small rodent populations (see **Tables 1–3**).

People also make decisions that impact urban biodiversity and ecosystem services at various scales – from homeowner to neighborhood association to city manager to multinational organizations governing trade. At each scale, people are pursuing different goals, have different tools available to them to achieve those goals, and operate under different constraints. In the four sections that follow, we examine these goals, tools, and constraints at four different scales – household (individual), neighborhood, cities and their hinterlands, and networks of cities. We refer to those decisions and impacts occurring at smaller scales (household and neighborhood) to be "bottom-up" processes – outcomes at larger scales are the integrated (and often uncontrolled) impact of decisions and actions taken at smaller scales. Because of variations in preferences, values, and economic capacities, bottom-up processes tend to create heterogeneity in the urban landscape, as individual residents or neighborhoods pursue different goals and outcomes. Larger-scale decisions often exert "top-down" pressures on outcomes, either again through deliberate manipulations (e.g., exclusion of certain plants in cities for safety or health reasons) or unintended outcomes (e.g., invasive species that might accompany trade goods). Top-down processes tend to have a homogenizing influence, with many policies or actions having a similar impact across a city or collection of cities. After examining key issues at each scale, we return to the question of how (or indeed whether) our understanding of urban ecology challenges traditional ecological theory.

Individual and Household Scales

Residential properties compose at least a quarter of land in many cities (Gaston *et al.*, 2005; Mathieu *et al.*, 2007), which means that household-level decisions can have significant impacts on urban biodiversity patterns. Within the outdoor space surrounding homes, residents directly and indirectly affect biodiversity through yard management and "bottom-up" control of species composition, habitat, resource flows, or disturbances. Household management decisions often reflect residents' personal preferences, such as for easy to maintain or esthetically appealing landscapes (Larson *et al.*, 2010). However, landscaping decisions are also constrained by residents' capacities (e.g., financial capital, time constraints, and social status) to choose their ideal preferences (Templeton *et al.*, 1999); social pressures to maintain certain yard types; neighborhood or municipal regulations (e.g., planting restrictions); or the decisions of past land owners (Boone *et al.*, 2010; Cook *et al.*, 2011). Residents most consciously manage their yard for cultural services (e.g., esthetics and sense of place), provisioning services (e.g., vegetable or flower gardens); and regulating services (e.g., pest, pathogen, climate, or fire control). Other services, such as nutrient cycling, are less often direct goals of management decisions of urban residents, but instead are indirectly impacted by other activities.

At the scale of households, individuals' management practices alter species composition, such that urban vegetation and fauna often differ from the surrounding native ecosystem (Hope *et al.*, 2006). In particular, people purposefully introduce, care for, and manage certain "desirable" plants in their yards, directly affecting vegetation biodiversity (Martin *et al.*, 2003; Smith *et al.*, 2006; Loram *et al.*, 2008). In other cases, people remove "undesirable" plants through weeding or the application of herbicides. In addition, fertilizer and watering affect nutrient cycling, increase plant primary production, and support vegetation outside of its typical growing season or range. With available financial and social capital, residents can choose specific landscaping or gardens, for example, to meet their esthetic preferences or attract or deter particular species, such as hummingbirds and scorpions, respectively. As such, urban vegetation diversity and cover are positively related to residents' socioeconomic status (Hope *et al.*, 2003; Grove *et al.*, 2006).

Residents' intentional manipulation of yard vegetation leads to many direct and indirect impacts on faunal biodiversity ranging from soil organisms to mammals (see **Tables 1–3**). The abundance and diversity of fauna that can be maintained within a managed area or yard is directly related to the available habitat, yard type, vegetation density, and distribution. Faunal composition, particularly ground arthropods, can also be directly managed with application of pesticides. Residents choose and maintain groundcover (e.g., grass) and other plant species that in turn influence faunal habitat characteristics, such as microclimate and associated resources. Yard size and pervious area represent the available living space for many organisms and impact the distribution of many urban faunal species. For example, yard area is positively related to bird, butterfly, and small mammal abundance and diversity (Baker and Harris, 2007; Evans *et al.*, 2009).

Finally, humans directly impact the availability and distribution of resources, such as food and water, which in turn regulate urban plant and faunal populations and their interactions. In residential yards, seasonal variation of resources is often decreased relative to the preurban ecosystems, whereas resource abundance is often also increased (Shochat *et al.*, 2006). For example, bird feeders and nesting boxes create resources for bird species that would not otherwise remain

Table 1 Human activities that directly target biodiversity in cities, and some of the unintended impacts of such activities. Check marks indicate the *main* taxa impacted by particular activities, but impacts are not limited to these taxa. (Symbols broadly represent common groups of vegetation and fauna: e.g., **: vegetation; **: birds; **: mammals (domesticated and wild); **: insects and arthropods; **: reptiles; **: microorganisms**

Human action	*Targeted taxa*	*Intentional and unintentional impacts*	*Other impacted taxa*					
Household level								
Pets		Pets alter the diversity of birds and mammals in the city. Escaped pets can naturalize and become part of the urban fauna. Cats can predate on birds, reptiles, insects, and rodents.		✓	✓	✓	✓	
Bird feeding		People actively feed birds, or plant vegetation that will attract birds. This has implications for the insects, small mammals, and rodents birds may prey on.			✓	✓	✓	
Household, neighborhood, and city level								
Landscaping		Humans create preferred landscapes. This can create broader habitats in neighborhoods with shared sociocultural preferences for certain landscapes. The homogeneity or heterogeneity of these neighborhood habitats will influence birds, insects, reptiles, and small mammals moving through the urban setting. Soil properties and microbial communities are influenced by the chosen plant communities.		✓	✓	✓	✓	✓
Watering & fertilizing		Watering to promote plant growth can attract birds, insects, and small mammals to the city from the hinterlands, particularly in places where water is scarce. Watering and fertilizing affect soil nutrient properties and soil microbial communities.		✓	✓	✓	✓	✓
Pest control		People control the prevalence of harmful or annoying insects and rodents, which can have impacts on beneficial insects and rodents, with consequences for plants. Pesticides can alter soil properties and soil microbial communities.	✓					✓
Neighborhood or city level								
Planting restrictions		Neighborhood associations may restrict landscaping choices, and city governments can ban invasive or allergenic plants. This influences the habitat experienced by all other organisms.		✓	✓	✓	✓	✓

Table 2 Some of the decisions and activities of city governments directed toward urban design and infrastructure (e.g., biodiversity is not the target) with significant consequences for biodiversity (symbols same as **Table 1**)

Human action	*Intentional and unintentional impacts*	*Primary affected taxa*
Zoning	Zoning restricts land use, and the vegetation associated with that land use (e.g., manufacturing vs. residential).	
Construction	Compacts or removes soils, altering soil microbial communities.	
Roads	Can create conduits for movement of plants, mammals, and reptiles. Roadside deposits from car exhaust can alter soil microbial communities.	
Flood control & groundwater recharge	Cities create retention basins, canals, and levees to redistribute and control water. This affects birds, insects, small mammals, and reptiles that will converge on water sources. The altered soil properties of retention basins will affect soil microbial communities.	

Table 3 A classification of human impacts that affect biodiversity in urban settings

Alteration of habitat
Destruction of habitat
Creation of habitat
Fragmentation of habitat
Alteration of resource flows
Reduction in net primary production
Increase in regional temperature
Mobilization and concentration of nutrients
Dispersion of toxins
Diversions of water and changes in timing of availability
Degradation of water quality
Alteration of disturbance regimes
Increase in small-scale disturbance
Attempts to decrease large-scale disturbance
Unintentional shifts in disturbance
Alteration of species composition
Preferential cultivation or destruction of native species
Introduction of nonnative species

year round (Gaston *et al.*, 2005; Fuller *et al.*, 2008; Davies *et al.*, 2009). Similarly, watering practices directly provide a needed resource for many urban fauna, and water and fertilizer together, through their impact on plant communities, can alter habitat availability. Those species that can thrive on the new, urban patterns of resource availability are called synanthropic species (literally "with humans"). Predator populations are also impacted through both exclusion (e.g., wild cats and some raptors that are ill-suited to urban environments) and introductions (e.g., domesticated pets). When both faunal resources and faunal predators are changed, there can be significant impacts on trophic dynamics (Faeth *et al.*, 2005; Shochat *et al.*, 2006, 2010). Other disturbances to biotic populations include pruning and trimming of vegetation, which can alter patterns and timing of pollination and seed dispersal, as well as natural pest regulation within cities.

Neighborhood Scale

Biodiversity at the neighborhood scale is often an "emergent property" dependent on the personal characteristics and yard management decisions of many individual urban residents. Ordinances enacted by neighborhood associations or historic preservation societies, however, constrain the landscaping practices of urban residents, and thus have resulting implications for neighborhood biodiversity and ecosystem services. Neighborhood regulations vary by social and biophysical context, but generally focus on collectively maintaining property values and neighborhood cohesion. Still, neighborhood "bottom-up" controls affect urban biodiversity through direct and indirect management for ecosystem services, as well as altered species composition, habitat, resources, and disturbances. Neighborhood ordinances focus on managing for services similar to those at the individual-household scale, including esthetics and a place for recreation (cultural services); and pest, microclimate, flood, and fire control (regulating services). In addition, neighborhood associations may implement measures – such as restrictions on water use (e.g., for flood control) or on certain plant species – to minimize environmental bads; yet, these restrictions can also affect biodiversity.

Neighborhood ordinances frequently regulate plant composition and thus impact habitats and resources for fauna. Specifically, regulations often focus on landscape plantings (e.g., the amount and type of grass and woody plants) in order to visually maintain a cohesive community. For example, in Phoenix, Arizona, neighborhoods with homeowner associations have planting regulations that result in yards with fewer trees, more shrubs, and less turfgrass than neighborhoods without homeowner associations (Martin *et al.*, 2003). Planting regulations in turn impact microclimates and species habitats, and indirectly influence the amount and timing of fertilizer, pesticides, and water applied to a given landscape. The existence of these neighborhood-level regulations, along with the propensity of people to move to neighborhoods where other residents share their landscaping preferences, may mean that the neighborhood scale is of particular relevance in understanding the urban matrix of habitats and corridors that affect urban species (Goddard *et al.*, 2010).

At the neighborhood scale, biotic diversity is also impacted by structural characteristics of the neighborhoods and the households comprising that neighborhood. For example, vegetation diversity and cover are inversely related to housing density and neighborhood age (Marco *et al.*, 2008;

Hope *et al.*, 2006). Similarly, faunal distribution is positively related to proximity to natural habitats and corridors, which provide a link between fragmented habitat areas within the city (Germaine *et al.*, 1998; Daniels and Kirkpatrick, 2006; Loss *et al.*, 2009). However, fragmentation between houses, neighborhoods, or land use types introduces the urban equivalent of "ecotones" – transition zones between pavement and yard, or a lawn and a garden, or a park and a commercial mall. The preponderance of edges in a highly fragmented landscape increases the abundance of edge-dwelling species and decreases the abundance of those species particularly sensitive to edge effects.

Cities and Their Hinterlands

City-scale decisions about land use can act as "top-down" controls on biodiversity by constraining the spatial patterns of land use or through ordinances on land use and land management practices (e.g., limitations on lawn watering). City-scale decisions also affect smaller, localized areas of land owned and managed by the city, such as parks and public right of ways, as well as habitats that are less frequently managed by individuals, such as rivers and canals. Finally, cities are surrounded by nonurban land, their "hinterlands," on which they can exert influence and from which urban residents may obtain certain ecosystem services. Variation of management practices and regulations among cities is driven by some of the same considerations as those of individuals (such as wealth and culture). However, the influence of city-scale decisions on biodiversity is also constrained by the environment itself, for example, topography, climate, and availability of water resources.

As noted earlier, the ecosystem services that are needed by city residents include provisioning services (e.g., food, fuel, water, and building materials); cultural services (including city "identity" and recreational opportunities for urban residents); and regulating services (such as flood, water quality, disease, and fire control). Some of these services are provided within a city, such as recreation in urban parks, food provisioning by urban agriculture, climate regulation by urban green space, and waste (nutrient) removal by wastewater treatment plant microorganisms. Other services are obtained via connections with urban hinterlands: water supply may be obtained through protecting a city's watershed, recreational opportunities also exist outside of a city, and waste is often disposed of outside of a city's borders (e.g., wastewater delivered downstream, landfills, and deposition of air pollutants downwind). Although cities often do not have jurisdictional control over their hinterlands, they do exert indirect control as a strong center of resource consumption and demand for ecosystem services. Finally, the way in which cities control services can impact plant and faunal distribution patterns. Many cities employ technological or hard engineering approaches to provide certain services, whereas other cities rely more heavily on green infrastructure (Gill *et al.*, 2007; Tzoulas *et al.*, 2007). For example, stormwater management ranges from hard engineering, such as pipes or concrete structures, to green infrastructure such as multiuse floodways, retention basins, and rain gardens. This example illustrates that how a city provides services has a potential impact on habitats and biodiversity within and around the city.

Decisions and actions taken at the city level can also have direct impacts on urban species composition especially through management of city-owned property. Urban protected areas, such as urban forests, often house and protect native vegetation. Parks and gardens have mixes of native and nonnative vegetation. These areas are often managed to provide certain services to residents, such as playing fields, space for vegetable gardens, or decorative flower gardens and arboretums. The selection of species and resulting species composition is dependent on these goals. Although plant diversity in parks is directly controlled by city managers, diversity of other taxonomic groups may be strongly influenced by the surrounding land use and socioeconomic factors such as wealth (Kinzig *et al.*, 2005). Cities can also have "top-down" control on residential decisions about their own properties, for example, by banning certain plants such as cottonwood and olive trees, for health concerns.

Cities impact species composition indirectly both within and beyond city boundaries through connections with agricultural areas that supply food to the city and recreational areas used by urban residents. Urban demand for certain food can drive crop selection by farmers. Recreational use of public lands by urban residents may alter species composition through the transport of invasive (and native) species by boats and cars (Hansen and Clevenger, 2005; Rothlisberger *et al.*, 2010). Finally, as cities expand, biodiversity is directly impacted by new development and changing land uses (McDonald *et al.*, 2008).

The spatial pattern of parks, public right of ways, and stream corridors is often very important in terms of how much habitat is available for higher trophic levels such as birds, reptiles, amphibians, and mammals. Park size can matter greatly for large animals with greater territory needs. Furthermore, although avian diversity is often related to socioeconomic patterns at the neighborhood scale (Kinzig *et al.*, 2005), other studies have found that avian diversity is also affected by the plant diversity of streetways and parks (White *et al.*, 2005; Sandstrom *et al.*, 2006). Some species thrive in the urban landscape. Plants whose seeds are carried along roads by wind or in the tread of vehicles do not face a fragmented landscape, but a highly connected one (Hansen and Clevenger, 2005). In contrast, fragmentation of suitable habitat is often dangerous for species that migrate seasonally between habitats, such as amphibians during breeding season.

The availability of resources, such as food, nutrients, and water, are often greatly varied by decisions made at the city scale. At the city scale, many of these altered resource flows are intentional, though only indirectly related to biodiversity. For example, imported food eventually ends up in wastewater, dramatically increasing nutrient availability in downstream ecosystems (Lauver and Baker, 2000; Baker *et al.*, 2001). However, storage of water for human use in reservoirs reduces the availability of water to downstream systems and creates new freshwater habitats upstream of cities.

As dense areas of human settlement, cities have large populations and property to protect from disturbances such as floods and fires (Radeloff *et al.*, 2005). With active and effective management, these disturbances can often be reduced within cities. However, cities also provide new human

sources of fire, and the built environment is especially prone to flooding due to altered hydrology (Paul and Meyer, 2001). Some management of disturbances in cities may only displace disturbances elsewhere. For example, stormwater drainage reduces flood damage within cities, but drastically changes the hydrology and disturbance regimes of downstream ecosystems. This has led to a reduction in species richness of fish and invertebrates in many areas (Paul and Meyer, 2001). However, prevention of disturbances may negatively impact disturbance-adapted species.

Networks of Cities and Their Hinterlands (Regional and Global Scales)

Although city officials, neighborhood associations, and individuals determine many features of individual cities, cities are located within a complex web of interconnected metropolitan areas and nonurban ecosystems. Although fewer decisions are directly made at this regional or global network scale, it is important to note some of the emergent functions of these networks and the effects they might have on biodiversity at the city scale and globally. Unlike cities and individuals, networks of urban areas rarely have goals or ecosystem services that they are intentionally and collectively pursuing (although note the emergence of networks of city mayors, such as the US Conference of Mayors Climate Protection Agreement, planning together for climate change adaptation). Instead, one can think of networks of cities as a scale at which the effects on biodiversity are primarily unintentional and indirect. A fundamental function of city networks is to redistribute goods, resources, species (including diseases), and people between cities separated by great distances. Thus, cities are connected specifically through the exchange of goods, people, and ideas. Finally, cities can act as regional centers of consumption, and therefore can indirectly affect land management and biodiversity in regions that produce goods and services for the city (Luck *et al.*, 2001).

Resources such as nutrients are redistributed globally and regionally through trade networks. A key example is the global fertilizer trade that provides nutrients and resources needed to grow and support a variety of crops at regional, city, and local scales (Cordell *et al.*, 2009). Global networks of cities also profoundly affect biodiversity via the intentional and unintentional movement of species (Levine and D'Antonio, 2003). Landscaping and ornamental plants are perhaps the most common intentionally traded species within the urban context, but animals are also intentionally introduced to urban areas, such as starlings and ornamental fish (Whittington and Chong, 2007). Trade in species clearly affects the biodiversity of the cities importing new species, but it may also negatively impact biodiversity in the exporting region if that species is threatened or endangered. Although there are various national and international agreements and legislation regarding the sale of endangered species, illegal markets still exist.

The effects of intentional movement of species might be insignificant compared to the effects of unintentional species transport around the globe. These effects are felt primarily through the movement of invasive species and diseases. Aquatic invasives are transported largely through shipping (Hulme, 2009), and therefore primarily affect port cities. Terrestrial invasives, however, easily travel along roadways, transported by wind or in the treads of vehicles (Hulme, 2009). Animal and plant diseases can easily spread from city to city inadvertently through trade, much of which is inadequately monitored to prevent disease spread (e.g., Whittington and Chong, 2007). Some connections between disease and biodiversity cover long distances and complex market relationships (e.g., relationships between Bovine Spongiform Encephalopathy (BSE) in the UK and the abundance of North American grassland birds; Nocera and Koslowsky, 2011).

Finally, the movement of people around the globe and subsequent shifts in local demographics and culture is another significant way in which the connections between cities can constrain or alter local patterns of biodiversity. Immigration to cities can have substantial effects on urban biodiversity due to cultural changes especially with regard to landscaping preferences. In fact, increasingly, the flora and fauna of cities are becoming more homogeneous within the US and globally (McKinney, 2006).

Applying and Testing Ecological and Social Theories in Urban Systems

It has been argued that studying ecology and biodiversity in cities is essential because this is where most people live – and thus we need to understand the relationship between human activities and ecosystem services in the urban biosphere and its impact on human well-being. But there is a second scientifically driven reason to study biodiversity: it offers us a chance to test ecological theories, and to ask whether our theories, developed in more pristine (less human-dominated ecosystems), apply when humans dominate the system (Collins *et al.*, 2000). Do our theories – about how species assemblages change, interact, scale with area, transform resource flows, and withstand disturbance regimes – hold when we have biotic communities almost "haphazardly" assembled (from an ecological perspective), and with no coevolutionary history? To what extent can humans – themselves the product of evolution – be treated as "just another species" in the ecosystem, and when do we need new theories – borrowed from the social sciences or synthetically created across disciplines – about human behaviors and their influence on ecological systems?

Arguing from first principles, one could assert that ecological theory should continue to apply. Humans are, after all, "just another species," the product of evolutionary processes. Analogs of many human activities – species consumption and dispersal, "fertilizing" of patches through waste disposal, and disturbance through fire or "grazing" (mowing) – can be found among the nonhuman species with whom we share the planet.

Consider, for instance, the use of a "gradient approach" for assessing biodiversity patterns in urban settings. Many studies examine the relationships between biodiversity and certain indicators of human pressure, activity, or disturbance along an urban to rural gradient – indicators such as population density, percent impervious surface, or distance from urban center (Blair, 1999). One assumption implicit in this approach is that the characteristics of the individuals occupying the space – their preferences and the capacity to realize those preferences – matter very little to outcomes. We need to only

know how closely they are packed together (population density), or their preference for urban or suburban landscapes (distance from city center), or their propensity for spreading to know what impact they will have on biodiversity. This approach is consistent with current ecosystem- and community-ecological theory, which primarily tends to treat humans as an exogenous perturbing force on the landscape (or ignore them completely). In other words, relying only on current theories we need primarily to know something about the magnitude of disturbance inflicted, or the gross flow of mobilized resources, to understand the impact humans will have on biodiversity outcomes. This suggests, for example, that a thousand "snowbirds" (transplants from the Midwest) living in a square kilometer on the fringe of Phoenix should have the same impact on biodiversity as 1000 Guatemalan refugees living in the same area. This seems unlikely, given what we have discovered over the past decades about human interactions with urban ecosystems – and indeed the environment more generally (see, e.g., Warren *et al.*, 2010; Hegmon *et al.*, 2008). Groups of different ages or cultural backgrounds prefer different landscapes and value different ecosystem services. In addition, socioeconomic status will enable or constrain the capacity to realize those landscapes. To build on this example and the idea of testing and adapting existing theories – both social and ecological – to address urban systems, the authors provide a brief history of urban ecology.

The study of cities from an ecological perspective in the US began with the Park *et al*. landmark 1925 publication, *The City* (Park *et al.*, 1925), which formally introduced human ecology as a new research agenda for sociology and the study of cities in America. Their research focused on many social changes that were a result of rapid expansion of America's urban areas due to the mass immigration of people from Europe and rural America. The explosive growth of the city, the confluence of people from diverse backgrounds, the breakdown of old ways, and the changes that were necessary for a viable new urban life caught the imagination of the authors.

Although Park *et al*. drew on the work of European social and biological scientists, such as Malthus, Darwin, and Spencer, the initial development of human ecology in America was influenced significantly by and contemporaneous with the emerging fields of plant and animal ecology in America. The Chicago School – as it came to be known – conceived of human ecology as an extension of the developing fields of plant and animal ecology.

The Chicago School articulated and developed an approach to human ecology that drew on and paralleled earlier ecological work. First, Park and his colleagues applied a community ecology approach to the complexities of urban society in order to uncover a set of regular social patterns and processes in the apparent confusion of the urban melting pot. For instance, Park *et al*. employed ecological concepts such as succession, competition, and metabolism to describe stages of human community structure (organization) and function (processes): specifically, indicators of social disorganization such as disease, crime, vice, insanity, and suicide. Second, the Chicago School conceived of the city as a closed and functional system (community) that could be treated as an organism or "superorganism." Park and his colleagues also focused on the spatial and temporal dimensions of the city.

A significant product of this work was Burgess' ideal model of the city which the Chicago School used to describe and measure the city's spatial differentiation and development into zones and areas-within-zones through processes of concentration, centralization, segregation, invasion, and succession.

The Chicago School's conception of human ecology was criticized strongly by social scientists for several reasons, including opposition to the use of biological factors to explain individual human behavior and social structures; a singular reliance on competition as the mechanism for explaining the organization of economic functions and the spatial distribution of human populations and services; and the use of macroscale processes and functional approaches to explain individual behavior from both a conceptual and a statistical point of view. Many of these criticisms were addressed in later studies in human ecology.

As in the example of the "snowbirds" and Guatemalans, it is not enough to treat humans or human ecological systems as monolithic blocks. Just as performance tradeoffs and the emergence of different strategies for resource acquisition and reproduction can influence the patterns of biodiversity on the landscape, differences in the structure and organization of human social systems need to be considered in order to understand how humans impact and manipulate the biodiversity surrounding them.

All social species are characterized to varying degrees by patterns and processes of social differentiation. In the case of *Homo sapiens*, social differentiation or social morphology has been a central focus of sociology since its inception. In particular, social scientists have used concepts of social identity (age, gender, class, caste, and clan) and social hierarchies (wealth, power, status, knowledge, and territory) to study how and why human societies become differentiated.

Social hierarchies – or social differentiation – is an important concept for understanding human ecological systems because it affects the allocation of critical resources (natural, socioeconomic, and cultural) and hence the patterns of and impacts on biodiversity. In essence, social differentiation determines "who gets what, when, how, and why." This allocation of critical resources is rarely equitable. Unequal access to and control over critical resources is a consistent fact within and between households, communities, regions, nations, and societies.

Wealth is access to and control over material resources in the form of natural resources, capital (money), or credit. The unequal distribution of wealth is a central feature of human ecological systems, and one that must be accounted for when examining the driving forces behind past changes in the environment and predicting the impacts of future policies and conditions. Power is the ability to alter others' behavior through explicit or implicit coercion. The powerful (often elites with political or economic power) typically have access to resources that are denied to the powerless. Status is access to honor and prestige and the relative position of an individual (or group) in an informal hierarchy of social worth. Status is distributed unequally, even within small communities, and high-status individuals may or may not have access to either wealth or power. For instance, a minister or an imam may be respected and influential in a community even though he or she is neither wealthy nor has the ability to alter coercively other people's behavior. Knowledge is access to or control over

specialized types of information (technical, scientific, religious, and so forth). Not everyone within a social system has equal access to different types of information. Knowledge often provides advantages in terms of access to and control over the critical resources and services of social institutions. Finally, territory is access to and control over critical resources through formal and informal property rights.

These types of hierarchies – or differentiation – are important constraints that affect the allocation or supply of critical resources. It is also important to examine and understand variations in demand. For instance, additional social characteristics that may affect the demand for critical resources are often related to demographics or ethnic or religious backgrounds, since different age groups and various cultures may have different needs from, or attitudes toward, their environment and each other.

Social differentiation within human ecological systems also has a spatial dimension that is usually characterized by patterns of territoriality that lead to spatial heterogeneity on many scales. A spatial understanding of social differentiation in an ecological context enables researchers to ask "who gets what, when, how, why, and where?" and subsequently, to ask about the reciprocal relationships between spatial patterns of sociocultural and biophysical patterns and processes of a given area (Grove and Burch, 1997).

Today, it is increasingly difficult to determine where traditional biogeophysical ecology ends and human ecology begins (Golley, 1993). The articulation of a biosocial approach to human ecological systems has occurred over the past 20 years. Importantly, this approach is continuously advanced by the development of and discourses between theories related to plant, nonhuman animal, and human ecologies and social sciences. Indeed this integration among ecological and social sciences is absolutely essential if we are to understand the functioning and dynamics of cities in an ecological landscape, and if we are to be able to both understand and predict the impacts of humans in urban centers on biodiversity.

See also: Air Pollution. Biofuels and Biodiversity: The Implications of Energy Sprawl. Comparing Extinction Rates: Past, Present, and Future. Ecological Footprint, Concept of. Ecosystem Services. Evaluation of Ecosystem Service Policies from Biophysical and Social Perspectives: The Case of China. Evolution, Theory of. Extinction, Causes of. Globalization Effects on Common Plant Species. Government Legislation and Regulations in the United States. Human Impact on Biodiversity, Overview. Human Impacts on Ecosystems: An Overview. Implications of Urbanization for Conservation and Biodiversity Protection. *In Situ*, *Ex Situ* Conservation. Introduced Species, Impacts and Distribution of. Land-Use Issues. Landscape Corridors. Loss of Biodiversity, Overview. Market Economy and Biodiversity. Modern Examples of Extinctions. The Multiple Benefits of River–Floodplain Connectivity for People and Biodiversity

References

Baker L, Hope D, Xu Y, Edmonds J, and Lauver L (2001) Nitrogen balance for the Central Arizona–Phoenix (CAP) ecosystem. *Ecosystems* 4: 582–602.

Baker PJ and Harris S (2007) Urban mammals: What does the future hold? An analysis of the factors affecting patterns of use of residential gardens in Great Britain. *Mammal Review* 37: 297–315.

Blair RB (1999) Birds and butterflies along an urban gradient: Surrogate taxa for assessing biodiversity? *Ecological Applications* 9: 164–170.

Boone CG, Cadenasso ML, Grove JM, Schwarz K, and Buckley GL (2010) Landscape, vegetation characteristics, and group identity in an urban and suburban watershed: Why the 60s matter. *Urban Ecosystems* 13: 255–271.

Chandler T (1987) Four Thousand Years of Urban Growth: An Historical Census. Lewiston, New York: St. David's University Press.

Collins JP, Kinzig A, Grimm NB, *et al.* (2000) A new urban ecology. *American Scientist* 88: 416–425.

Cook EM, Hall SJ, and Larson KL (2011) Residential landscapes as social–ecological systems: A synthesis of multi-scalar interactions between people and their home environment. *Urban Ecosystems*

Cordell D, Drangert J-O, and White S (2009) The story of phosphorus: Global food security and food for thought. *Global Environmental Change* 19: 292–305.

Daniels GD and Kirkpatrick JB (2006) Does variation in garden characteristics influence the conservation of birds in suburbia? *Biological Conservation* 133: 326–335.

Davies ZG, Fuller RA, Loram A, Irvine KN, Sims V, and Gaston KJ (2009) A national scale inventory of resource provision for biodiversity within domestic gardens. *Biological Conservation* 142: 761–771.

Evans KL, Newson SE, and Gaston KJ (2009) Habitat influences on urban avian assemblages. *Ibis* 151: 19–39.

Faeth SH, Warren PS, Shochat E, and Marussich WA (2005) Trophic dynamics in urban communities. *Bioscience* 55: 399–407.

Fuller RA, Warren PH, Armsworth PR, Barbosa O, and Gaston KJ (2008) Garden bird feeding predicts the structure of urban avian assemblages. *Diversity and Distributions* 14: 131–137.

Gaston KJ, Warren PH, Thompson K, and Smith RM (2005) Urban domestic gardens (IV): The extent of the resource and its associated features. *Biodiversity and Conservation* 14: 3327–3349.

Germaine SS, Rosenstock SS, Schweinsburg RE, and Richardson WS (1998) Relationships among breeding birds, habitat, and residential development in greater Tucson, Arizona. *Journal of Applied Ecology* 8: 680–691.

Gill SE, Handley JF, Ennos AR, and Pauleit S (2007) Adapting cities for climate change: The role of the green infrastructure. *Built Environment* 33: 115–133.

Goddard MA, Benton TG, and Dougill AJ (2010) Beyond the garden fence: Landscape ecology of cities. *Trends in Ecology & Evolution* 25: 202–203.

Golley FB (1993) *A History of the Ecosystem Concept in Ecology: More Than the Sum of the Parts*. New Haven, Connecticut: Yale University Press.

Grove JM and Burch Jr. WR (1997) A social ecology approach to urban ecosystem and landscape analyses. *Urban Ecosystems* 1: 259–275.

Grove JM, Troy AR, O'Neil-Dunne JPM, Burch WR, Cadenasso ML, and Pickett STA (2006) Characterization of households and its implications for the vegetation of urban ecosystems. *Ecosystem* 9: 578–597.

Hansen M and Clevenger A (2005) The influence of disturbance and habitat on the presence of non-native plant species along transport corridors. *Biological Conservation* 125: 249–259.

Hegmon M, Peeples M, Kinzig A, Kulow S, Meegan CM, and Nelson MC (2008) Social transformation and its human costs in the prehispanic US southwest. *American Anthropologist* 110: 313–324.

Hope D, Gries C, Casagrande D, Redman CL, Grimm NB, and Martin C (2006) Drivers of spatial variation in plant diversity across the Central Arizona–Phoenix ecosystem. *Society and Natural Resources* 19: 101–116.

Hope D, Gries C, Zhu WX, *et al.* (2003) Socioeconomics drive urban plant diversity. *Proceedings of the National Academy of Sciences* 100: 8788–8792.

Hulme PE (2009) Trade, transport and trouble: Managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* 46: 10–18.

Kinzig A, Warren P, Martin C, Hope D, and Katti M (2005) The effects of human socioeconomic status and cultural characteristics on urban patterns of biodiversity. *Ecology and Society* 10: 23.

Larson KL, Cook EM, Strawhacker CS, and Hall SJ (2010) The influence of diverse values, ecological structure and geographic context on residents' multifaceted landscaping decisions. *Human Ecology* 38: 747–761.

Lauver L and Baker L (2000) Mass balance for wastewater nitrogen in the Central Arizona–Phoenix ecosystem. *Water Research* 34: 2754–2760.

Levine JM and D'Antonio CM (2003) Forecasting biological invasions with increasing international trade. *Conservation Biology* 17: 322–326.

Loram A, Thompson K, Warren PH, and Gaston KJ (2008) Urban domestic gardens (XII): The richness and composition of the flora in five UK cities. *Journal of Vegetation Science* 19: 321–330.

Loss SR, Ruiz MO, and Brawn JD (2009) Relationships between avian diversity, neighborhood age, income, and environmental characteristics of an urban landscape. *Biological Conservation* 142: 2578–2585.

Luck M, Jenerette G, Wu J, and Grimm NB (2001) The urban funnel model and the spatially heterogeneous ecological footprint. *Ecosystems* 4: 782–796.

Marco A, Dutoit T, Deschamps-Cottin M, Mauffrey JF, Vennetier M, and Bertaudiere-Montes V (2008) Gardens in urbanizing rural areas reveal an unexpected floral diversity related to housing density. *Comptes Rendus Biologies* 331: 452–465.

Martin CA, Peterson KA, and Stabler LB (2003) Residential landscaping in Phoenix, Arizona, USA: Practices and preferences relative to covenants, codes and restrictions. *Journal of Arboriculture* 29: 9–17.

Mathieu R, Freeman C, and Aryal J (2007) Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. *Landscape and Urban Planning* 81: 179–192.

McDonald R, Kareiva P, and Formana R (2008) The implications of current and future urbanization for global protected areas and biodiversity conservation. *Biological Conservation* 141: 1695–1703.

McKinney ML (2006) Urbanization as a major cause of biotic homogenization. *Biological Conservation* 127: 247–260.

Millennium Ecosystem Assessment (MA) (2005) *Ecosystems and Human Well-being: Synthesis*. Washington, DC: Island Press.

Nocera JJ and Koslowsky HM (2011) Population trends of grassland birds in North America are linked to the prevalence of an agricultural epizootic in Europe. *Proceedings of the National Academy of Sciences* 108: 5122–5126.

Park RE, Burgess EW, and McKenzie RD (eds.) (1925) *The City*. Chicago, Illinois: University of Chicago Press.

Paul MJ and Meyer JL (2001) Streams in the urban landscape. *Annual Review of Ecology and Systematics* 32: 333–365.

Radeloff V, Hammer R, Stewart S, Fried J, Holcomb S, and McKeefry J (2005) The wildland–urban interface in the United States. *Ecological Applications* 15: 799–805.

Rothlisberger J, Chadderton W, McNulty J, and Lodge D (2010) Aquatic invasive species transport via trailered boats: What is being moved, who is moving it, and what can be done. *Fisheries* 35: 121–132.

Sandstrom U, Angelstam P, and Mikusinski G (2006) Ecological diversity of birds in relation to the structure of urban green space. *Landscape and Urban Planning* 77: 39–53.

Shochat E, Warren PS, Faeth SH, McIntyre NE, and Hope D (2006) From patterns to emerging processes in mechanistic urban ecology. *Trends in Ecology & Evolution* 21: 186–191.

Shochat E, Lerman SB, Anderies JM, Warren PS, Faeth SH, and Nilon CH (2010) Invasion, competition, and biodiversity loss in urban ecosystems. *Bioscience* 60: 199–208.

Smith RM, Thompson K, Hodgson JG, Warren PH, and Gaston KJ (2006) Urban domestic gardens (IX): Composition and richness of the vascular plant flora, and implications for native biodiversity. *Biological Conservation* 129: 312–322.

Templeton SR, Yoo SJ, and Zilberman D (1999) An economic analysis of yard care and synthetic chemical use: The case of San Francisco. *Environmental and Resource Economics* 14: 385–397.

Tzoulas K, Korpela K, Venn S, *et al.* (2007) Promoting ecosystem and human health in urban areas using green infrastructure: A literature review. *Landscape and Urban Planning* 81: 167–178.

Warren P, Harlan S, Boone C, Lerman SB, Shochat E, and Kinzig AP (2010) Urban ecology and human social organisation. In: Gaston K (ed.) *Urban Ecology: Ecological Reviews*, pp. 172–201. Cambridge: Cambridge University Press.

White J, Antos M, Fitzsimons J, and Palmer G (2005) Non-uniform bird assemblages in urban environments: The influence of streetscape vegetation. *Landscape and Urban Planning* 71: 123–135.

Whittington RJ and Chong R (2007) Global trade in ornamental fish from an Australian perspective: The case for revised import risk analysis and management strategies. *Preventive Veterinary Medicine* 81: 92–116.

V

Valuing Ecosystem Services

Marc N Conte, Stanford University, Stanford, CA, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Compensated demand curve Demand curves represent how changes in price impact the quantity demanded of a given good. The compensated demand curve is a demand curve that ignores the income effect of a price change, only taking into account the substitution effect. To do this, utility is held constant and is unaffected by the change in the price of the good.

Compensating variation Compensating variation is the amount of money a consumer would need to be paid, or the amount of money that the consumer would need to pay, to be just as well off after a change in the price of a product which the consumer might buy.

Consumer surplus Consumer surplus is commonly defined as the area beneath the uncompensated demand curve and above the market-clearing price that consumers must pay for the good.

Contingent valuation A method of eliciting stated estimates of willingness to pay or willingness to accept for a given non-market good through hypothetical situations in which the provision of the good changes.

Ecosystem services Benefits that accrue to society as a result of functioning ecosystems.

Environmental production function Analog to the more traditional production functions from economics, in which inputs are combined using a given technology to create an output for consumption. Environmental production functions are used to describe how functioning ecosystems produce goods and services enjoyed by society.

Equivalent variation Equivalent variation is the amount of money that must be given to (or taken from) an individual in lieu of a change in the price of a consumed good to make them just as well off as they would have been with the price change.

Hedonic pricing model A means of attributing portions of the price of a good with many characteristics (e.g., a house or a car) to each of those characteristics.

Revealed preference A method of exploiting the relationship that exists between individual behavior and environmental attributes to assign value to non-market goods.

Stated preference A method of eliciting the value of a non-market good through an individual's stated behavior in a hypothetical setting.

Travel cost method The travel cost method estimates the value of non-market environmental goods, using the cost of arriving at locations with different environmental amenities to estimate the value of changes in the quality of the sites.

Uncompensated demand curve Demand curves represent how changes in price impact the quantity demanded of a given good. The uncompensated demand curve illustrates the consumer's consumption bundle at each wealth and price combination, assuming that she is a utility-maximizer.

Introduction

When choosing which patches of old-growth Douglas firs to harvest in the mountains of Oregon, a timber company might take note of the tree density or the slope, as these factors directly impact the profit of the company's operations. It is less likely that the company would voluntarily avoid harvesting trees from a steep hillside because the resulting increased release of sediment from the denuded hillside could increase costs for the reservoir operator downstream. Less likely still is the possibility that the timber company would factor in the impacts of extracting old-growth trees on the population viability of different forest-dwelling species.

In practice, the logging industry in the Pacific Northwest did not voluntarily adjust its operations in response to the negative environmental impacts associated with the clear-cut harvest of old-growth trees from the Pacific temperate rain forest. The clear-cut harvest of timber in Oregon, which began in the late 1820s, continued as the dominant management approach until the National Forest Management Act passed in 1976. The act did not prohibit clear-cutting, though it did restrict application of this harvest technique. In 1989, the US Fish and Wildlife Service designated the northern spotted owl as a threatened species in Washington, Oregon, and northern California, restricting land-use activities, including logging, on its habitat under the Endangered Species Act (ESA) of 1973.

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00343-9

These pieces of legislation set the status quo operation of the timber industry at odds with efforts to ensure the viability and continued function of these forest ecosystems.

Many loggers in the region voiced their displeasure with enforcement of the ESA, as they felt that their way of life was being sacrificed for a cause that did not clearly benefit social well-being. Following the listing of the northern spotted owl as a species covered by the ESA, bumper stickers appeared in the Pacific Northwest in support of the loggers. Examples include "Kill a Spotted Owl – Save a Logger" and "I Like Spotted Owls – Fried" (Satchell, 1990). Though framed in the context of species protection, the ESA has been implemented as a means of conserving habitat, presumably out of the belief that habitat conservation provides value to society. That said, this value may not accrue to the individuals whose activities are regulated, and the salience of this value to beneficiaries may not be obvious when discussed in terms of species viability. These facts contribute to the tension between forces of conservation and forces of conversion. This tension is what may have motivated the reframing of the habitat conservation issue in terms of ecosystem services, which deal, by definition, with the value of nature to humans.

The potential benefits to society of conserving these temperate rain forests in Oregon include but are not limited to maintenance of water quality in coastal tributaries that are home to commercially important salmon species and are used for human recreation, retention of sediment that might impact downstream reservoirs regarding both hydropower production and flood-prevention functions, and the sequestration and storage of carbon whose release into the atmosphere is associated with expected damages. In addition to this sample of benefits from intact forests in the Pacific Northwest, these forests are also essential habitat to multiple species of animals and plants whose survival depends on their continued existence. None of this suite of benefits is as directly tangible as the consumption of timber from the harvest of Douglas firs or the income associated with this harvest; however, this lack of familiarity and tangibility does not imply a lack of social value.

The choice between the conservation and conversion of extant habitat that continues to play out in the Pacific Northwest and elsewhere around the globe illustrates the trade-off that exists between the private benefits of conversion and the social benefits of conservation. Ensuring that land- and resource-use decisions provide the greatest amount of benefit to society requires the identification and quantification of the trade-offs inherent in such decisions. To achieve this goal, we must be able to value the benefits of habitat conservation, which must generally be forgone to enjoy the benefits of intensive resource management.

Under certain conditions, self-interested actors operating within a market system can achieve an efficient use and allocation of resources. However, several characteristics of land and other natural resources typically prevent markets for these goods from achieving their efficient use. There are external costs associated with habitat conversion, and several benefits provided by extant habitat are public goods: it is difficult to exclude downstream users from receiving the benefits of improved water quality (i.e., this good is nonexcludable), and carbon sequestration and storage by forests reduces the probability of climate change for every individual on the planet (i.e., this good is nonrival). These two features of land and resource use have contributed to inefficient habitat conversion across the globe.

It would seem that regulatory and public policy mechanisms could increase social welfare through the increased protection of habitat and the provision of associated ecosystem services relative to the outcomes of unregulated markets. There are trade-offs associated with habitat conservation just as there are with habitat conversion, and there will be both winners and losers from such decisions. Those asked to forgo private returns in order to ensure increased social benefits will tend to challenge the implementation of such regulation. Quantifying the welfare implications of public policies aimed at ensuring the continued provision of ecosystem services requires an understanding of how ecosystem function is affected by human activity, how these effects alter the provision of services from these ecosystems, and how this change in provision affects social welfare.

Overview of Ecosystem Services

The Millennium Ecosystem Assessment (MEA, 2005) introduces four different categories of ecosystem service. The different characteristics across service categories impact efforts to value the individual services within each category. The first category, provisioning services, includes the services that are most familiar to society, such as agricultural products, seafood, and timber, among others. Provisioning services are ecosystem services that are directly consumed by humans and are often examples of renewable resources, a class of resource that has been well-studied by resource economists for nearly 70 years. These services are generally traded in traditional markets, and their enjoyment by society requires concerted effort by some members of society (e.g., the time, capital, and labor of farmers to harvest each year's crop). Even with markets for these provisioning ecosystem services, there are still externalities associated with their harvest, and the management of these provisioning services can significantly impact the quantity of service provision as well as its value to society. The value of these services comes from their consumptive use.

Regulating services impact human well-being through their regulation of natural processes. Examples of regulating services include the sequestration and storage of carbon by vegetation that helps to regulate climate, the mitigation of surface flows associated with storm events by vegetation that helps to regulate flooding extent, and the retention of nutrients and other pollutants by vegetative cover on the landscape that helps to regulate the quality of downstream water bodies. Regulatory services are typically not associated with markets, so government intervention is essential to ensure that there is an adequate provision of these services on the landscape. These services are typically provided without direct costly activity by humans. The value of these services comes from their use, though these services are quasi-public goods in that they are simultaneously enjoyed by groups of beneficiaries (i.e., they are nonrival and nonexcludable to segments of the population).

Cultural services cover a broad array of services related to both the direct and indirect social benefits of functioning natural systems. This umbrella of services runs from the enjoyment of outdoor recreation to the spiritual calming and

restoration of either access to or knowledge of areas of scenic beauty and cultural importance. Esthetic benefits such as the enjoyment of scenic coastal views, as well as the educational value of functioning natural systems, are included in the cultural ecosystem services category as well. Recreational opportunities and esthetic benefits are unique among the cultural services in that occasionally markets can be used to value these services, whereas the cultural, educational, and spiritual services are typically not associated with markets. The value of these services exists due to interaction with nature, directly (e.g., the satisfaction from catching or eating salmon) or indirectly (e.g., the benefits of having water available to irrigate crops that are eventually grown to feed humans) or through use, as well as the knowledge of or satisfaction from the existence of nature, which environmental economists refer to as "nonuse" value. Furthermore, there are both consumptive use values, as in the case of eating freshly caught fish, and nonconsumptive use values, as in the case of time spent canoeing the tributaries of Chesapeake Bay.

The final category of ecosystem services defined by the Millenium Ecosystem Service (MA) is the supporting services. These services represent natural processes that are directly responsible for the composition of each ecosystem though perhaps only indirectly related to the ecosystem services provided by that system. Examples of these processes include primary production, nutrient cycling, and soil formation. There are no markets for these services, and they are not typically valued directly; instead, they are acknowledged as contributing to the other categories of services that provide more easily quantifiable value to society. In economic terms, these services can be thought of as inputs to the production of the other categories of ecosystem services that are directly valued by society.

Ecosystem services are the outputs that result from a combination of different inputs. Cell phones, and many other products that we consume, are produced using a combination of raw materials, labor, and capital through a specific technology. Similarly, ecosystem services are produced by a combination of ecosystem components, both biotic and abiotic, through the function of that ecosystem. An ecological production function specifies the output of ecosystem services that are provided by an ecosystem given its condition and processes. These functions will vary across space due to site- and ecosystem-specific relationships.

Once an ecological production function is specified, researchers can quantify the impact of landscape change on the level of ecosystem-service outputs. In the past two centuries, human alteration of ecosystems on a large scale, such as the conversion of extant habitat to single-crop agriculture, or monoculture, led to an increase in some provisioning services (e.g., food production) at the expense of many regulating, cultural, and supporting services (MEA, 2005; Vitousek *et al.*, 1997). In the example of timber harvest in the Pacific Northwest, the easily definable and tangible value of Douglas firs for their timber led to massive consumption of this ecosystem service and a simultaneous loss of most other services associated with temperate coastal forests.

The production function through which biophysical processes in an ecosystem are converted into ecosystem services is impacted by several ecosystem characteristics. The plant and animal species that live within an ecosystem are important features of the system. Although the dominant vegetation type associated with an ecosystem has been shown to impact the quantity of ecosystem services produced by the system, the relationship between biodiversity and ecosystem services is less clear.

Ecosystem Services and Biodiversity

The traditional motivation for conservation has been the protection of wildlife, especially the charismatic species, that live in the areas of interest for conservation. Generally, this approach sets tangible benefits of intensive land management against the fate of different species chosen to represent the health of the ecosystem of interest. In many instances, as in the case of the northern spotted owl, the value of the representative species to society is not easily defined, making it difficult to compare to the forgone profits associated with habitat conservation.

Although there is a consensus among natural scientists that human action has altered the species composition of many ecosystems through species invasions and extinctions, there are still uncertainties regarding the impacts of varying species composition and diversity on ecosystem function (Hooper *et al.*, 2005). Experimental work conducted in the grassland prairies of Minnesota suggest that species diversity increases the productivity of a system as measured by biomass (Tilman *et al.*, 1997). However, it is not well understood if such results will extend to impacts of altered diversity for non–primary producer species.

The introduction of the term "ecosystem services" and the focus on the benefits to society of functioning ecosystems was partially motivated by the challenge of convincingly conveying the importance to humanity of ensuring the continued existence of various animal and plant species. Although the value of the northern spotted owl might not be clear to the average policy maker or landowner, that same policy maker or landowner might be familiar with the importance of clean water to society. By connecting the forests that are home to the owl to the idea of water filtration by the trees in the forest, the societal trade-offs between logging and forest conservation may become more salient to people affected by the choice between forest conservation and conversion.

Maintaining biodiversity is not typically considered an ecosystem service. Although there is some expectation that a focus on ecosystem services will ensure the continued existence of different species, the shift in focus to ecosystem services from species conservation clearly aims to highlight the value of functioning ecosystems to humanity. Understanding the different approaches that can be applied to value nature is useful in determining which approach might be most appropriate for a given policy context or land-use decision.

Definition of Value

In discussing the value of nature, there are two alternative philosophical approaches that can be used to frame the issue. The anthropocentric approach claims that natural things and functioning ecosystems are of value only to the extent that

they provide satisfaction to humans. The biocentric approach asserts that the satisfaction of humans as well as that of other species should be considered in determining the value of extant habitat and nature. Goulder and Kennedy (2011) provide a more-comprehensive discussion of the concepts that are briefly introduced here.

Although the anthropocentric approach assigns value based solely on benefits that accrue to humans, this does not imply that it offers a defense for the total exploitation of nature. This approach asserts that we should assign value and therefore protect nature and other species only because humans gain satisfaction or well-being from doing so. As such, if people feel that the protection of nature or other nonhuman species is truly beneficial to human satisfaction or well-being, then the anthropocentric approach will place significant value on nature and these other life forms.

The total value that is calculated under the anthropocentric approach represents the sum of values from a suite of different benefits that accrue to humans from nature. Broadly, the total value is the sum of both use and nonuse values. As previously mentioned, use values can be categorized as direct or indirect use values. Finally, nonuse values, which involve no direct or indirect interaction between humans and the benefits from nature, can provide significant satisfaction to members of society.

Unlike the anthropocentric approach, the biocentric approach gives weight directly to the well-being of all species. As a result, the biocentric approach allows for the possibility that other species and the habitats in which they exist can be of value even if the value does not directly or indirectly accrue to humans. This component of value, which is unrelated to humans, is often referred to as "intrinsic value."

The impetus for calculating the value of ecosystem services is to help inform resource users and policy makers in evaluating potential resource uses and policy mechanisms. These decisions are frequently evaluated through benefit–cost analysis, which ranks alternatives based on their net returns to society, identifying the net gain to society of choosing one policy alternative over another.

The benefit–cost technique involves summing all of the gains associated with a given policy option and then subtracting all of the losses of that same option to establish the net gain of each option before identifying the policy option with the greatest net benefit. Part of the controversy surrounding the ESA in the USA is that the act specifically forbids the use of economic considerations to determine which species are eligible for protection. As a result, all species – from the red wolf and the North American cougar to the American burying beetle – are equally eligible to have their habitat protected by the act. This biocentric approach to land-use regulation, which unanimously passed the Senate in 1973, has become quite contentious as the value of conservation is measured in terms of the benefits afforded to the listed species, whereas the benefits of habitat conversion are measured using the anthropocentric approach and are therefore more easily relatable to landowners. By preventing application of benefit–cost analysis, the ESA has often been criticized as placing an overly onerous burden on private landowners (Brown and Shogren, 1998). Part of the reason for this criticism may be due to the fact that the connection between the survival of the less-charismatic species protected under the ESA and the benefits of their survival to humans is tenuously understood at best.

The concept of ecosystem services is quite useful in highlighting the anthropocentric value of nature. By valuing each of the services provided, it is possible to quantify the anthropocentric value of a piece of habitat. The remaining sections of this chapter will focus on the anthropocentric approach to the valuation of ecosystem services. Having defined the approach to valuation that we will discuss in this chapter, we next consider the importance of assigning value to nature.

Assigning Value to Nature

When an objective is attained in the least-cost means, efficiency is a key concept and sought-after outcome in the discipline of economics. Under certain conditions, independent actors in competitive markets can achieve efficient social outcomes in the absence of government intervention. However, when these conditions do not hold, as is generally the case with regard to natural resources, government action is necessary to achieve efficiency. Generally, when the cost effectiveness or efficiency of a policy choice is being evaluated, the benefit–cost methodology is used to calculate what resources had to be used to achieve the given choice and what the benefit, or value, of that choice is to society. **Figure 1** illustrates the inefficiency that arises in the absence of government regulation when the production of a good, here converted habitat, has higher social costs than private costs. This negative production externality is a classic example of conditions under which government regulation can increase the efficiency of market outcomes. It is possible to measure

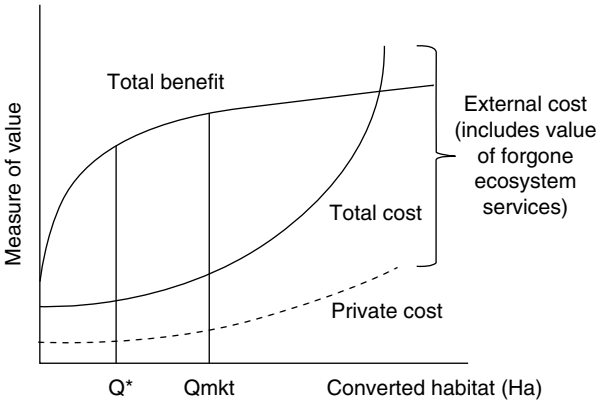


Figure 1 The external costs of habitat conversion. In the absence of regulation, Qmkt units of habitat will be converted because this amount of conversion maximizes net private benefits (total benefits minus private costs). Qmkt is more than the socially optimal amount of conversion because the costs of habitat conversion that accrue to other members of society, including the loss of the benefits associated with the various ecosystem services provided by the habitat, are not included in the maximization of private net benefits. By valuing these ecosystem services, we are able to identify the optimal amount of habitat conversion (total benefits minus total costs) and design policies to achieve this level of conversion.

the benefits and costs of a choice in different units. However, identifying the net benefits of an action is more straightforward if the benefits and costs of the action are measured in the same units.

Since before the passage of the ESA in 1973, there has been discomfort associated with the idea of placing an economic value on nature. Some of this concern has been rooted in the belief that by discussing nature in dollar terms, the value of conserving key habitat or protecting certain species would actually prove to be less than the value to society of converting this habitat into intensive agriculture production or another intensively managed land-use system. One other potential cause for the unwillingness to use financial terms to measure the importance of nature is related to the distinction between anthropocentric and biocentric values. By thinking of the importance of nature in economic terms, we are implicitly ignoring its biocentric value, thereby placing the welfare of humans ahead of that of all other living organisms. At the same time, efforts to define habitat value in the context of species diversity and viability have failed to resonate with broad elements of society.

The key goal of valuing nature is to ensure that resource-use decisions are being made in ways that result in the greatest returns to society. The shift to a focus on ecosystem services reflects an effort to emphasize the anthropocentric value of nature.

Ecosystem services are unlike most other market goods in that multiple services can be simultaneously produced by the same piece of habitat (e.g., a forest parcel provides carbon sequestration and storage, nutrient and sediment retention, and storm-peak mitigation). Because we are using the value of different ecosystem services to quantify the value of conserving extant habitat, we must take care to accurately value all services provided by a given patch of habitat. When policy makers and resource users choose to focus on the value of a single service, unintended consequences that induce additional costs on society can occur (e.g., Jackson *et al.*, 2005).

It is possible to estimate the value of ecosystem services in ways that more faithfully represent true social preferences than the approach of simply summing biophysical units, which implicitly assumes that each service is of equal per-unit value to society, a condition that may not hold in reality. The outcomes of these techniques are also consistent with the benefit–cost approaches that are used to determine the land- and resource-use decisions that provide the greatest benefits to society.

Quantifying the Value of Ecosystem Services

This section presents some of the economic theories and methods used to measure the welfare changes associated with changes in the provision of various ecosystem services. This effort has been undertaken in several different contexts previously (Goulder and Kennedy, 1997, 2011; Salzman, 1997; Heal, 2000; de Groot *et al.*, 2002; Boyd and Banzhaf, 2007). Bockstael and Freeman (2005) provide a very thoughtful and comprehensive discussion of assigning value to nonmarket goods in the broader context of environmental externalities, and the forthcoming discussion closely follows Sections 3 and 4 of their chapter in Volume 2 of *The Handbook of Environmental Economics*.

We present a general framework to measure the welfare implications associated with changes in the provision of various ecosystem services. The discussion here focuses on the direct welfare impacts of changes in ecosystem-service provision or prices. Understanding the full welfare impact of these changes requires an understanding of their impact on production of firms in different industries, which can be estimated using dynamic general equilibrium models.

The ecosystem services provided by a chunk of habitat are a function of the geologic, biophysical, and spatial characteristics of that piece of habitat. The function describing the production of ecosystem services from a piece of habitat will differ for each service of interest. Let $ES_{jx} = f_j(A_x)$ represent the quantity of ecosystem service j (e.g., nutrient retention by vegetation or storm-peak mitigation) produced by a piece of habitat x, with attributes given by A_x. We introduce the index x in recognition of the fact that the production of a given ecosystem service on a piece of landscape or seascape will vary spatially. There is a significant challenge in the practice of linking biophysical outputs with the inputs required by valuation efforts (e.g., hydrological models might estimate the tons of phosphorous retained by forests, but additional effort is required to relate that metric to the metric of lake clarity, which impacts lake swimming and the value of nutrient retention in the context of recreation).

Let H_x represent the measure of human activity that impacts the provision of ecosystem services on the xth unit of habitat. H_x might include increased water quality due to more-stringent regulation of effluent, human disruption of salmon nurseries, or changes in the area of forests due to increased zoning for residential development. So, we now have $ES_{jx} = f_j(A_x,H)$ as the expression that describes how a piece of habitat's characteristics and human activity on that piece of habitat combine to impact the production of the different ecosystem services provided by habitat on a given piece of landscape or seascape. The total production of ecosystem service j in the study area of interest is given by ES_j, where $ES_j = \Sigma_X\, ES_{jx}$. Certain ecosystem services, such as the production of agriculture, fish, or timber for harvest or storm-peak mitigation by forests, will directly impact human welfare, whereas others, such as the availability of water for crop irrigation or the presence of insect pollinators near agricultural land, will matter to humans because they enter as inputs to a production process, which we will describe as $Z(ES_j)$.

In the case of logging operations in the US Pacific Northwest, H might describe the conversion of coniferous forest to bare soil due to clear-cut logging of old-growth Douglas firs. This activity will have different impacts on the various ecosystem services provided by the previously forested land. With regard to the service of timber production, the harvest of the firs represents a realization of this service and a direct increase in welfare. Regarding the service of storm-peak mitigation, converting the forest to bare soil represents a reduction in the provision of this service, which will accordingly correspond to a direct decrease in welfare. Removing the firs from a portion of the landscape may increase the water yield and therefore the water available for downstream irrigation, which will indirectly increase welfare.

Space is especially important in the valuation of ecosystem services. First, the location of service-providing habitat in the landscape, not just the habitat type, can impact provision. Generally, services are provided by ecosystems at a location that is some distance away from the segment of the human population that benefits from this service (e.g., forests in upstream portions of watersheds help to reduce the volume and timing of surface flows associated with storm events that reaches downstream populated areas). Next, different services provide value at different spatial scales (e.g., carbon sequestration provides benefits to people around the globe, whereas filtration of surface water used for drinking downstream is only valuable to the population within the same watershed) making it challenging to estimate the full value of a given piece of functioning ecosystem with a single analysis at a given scale. Hein *et al.* (2006) provide further commentary on this challenge.

The links that exist between ecosystem function, the provision of ecosystem services that are of use to humans, the effects of human activity on the provision of these services, and the welfare impacts of service provision can be represented with a simple model, which does not imply that identifying any one of these links is a simple matter. Let W represent the level of economic welfare. Then the model that describes the direct and indirect utility provided by a given ecosystem service, indexed by ES_j, can be expressed as:

$$ES_j = \Sigma_X ES_{jx} = \Sigma_X f_j(A_x, H_x) \quad [1]$$

$$Z(ES_j) \quad [2]$$

$$W = W(ES_j, Z(ES_j)) \quad [3]$$

We are able to track the welfare impacts of proposed regulation or changes in land use by predicting how these regulatory and behavioral changes will impact the production of the focal ecosystem service and then using eqns [1]–[3] to calculate the resulting welfare change in monetary terms. The response of firms and individuals to regulation is by no means a certain outcome, and the prediction of these responses is a difficult goal that lies outside of the scope of this chapter.

Before proceeding with the exposition of estimating the value of a single service, we briefly return to the fact that habitat provides multiple ecosystem services simultaneously. The value of a given patch of habitat is given by the sum of the value of the suite of services provided by that patch, or $V_x = \Sigma_J W(ES_j, Z(ES_j))$. To accurately estimate the value of nature, we must accurately estimate the value of each of the provided services.

The model presented in eqns [1]–[3] represents a simplified view of the problem that de-emphasizes several challenges. Although we already briefly mentioned the issue of space, tracking the differential impacts of location on the provision of ecosystem services as well as appropriately incorporating the differential impacts of human activity on important characteristics of parcels across the landscape is a nontrivial task. That said, key spatial characteristics should be incorporated in the model. The temporal dimension also provides motivation to approach estimation of welfare impacts from ecosystem service provision with care. The issue of time also warrants careful consideration, particularly because certain ecosystem services represent stocks (e.g., the salmon available for harvest), whereas others represent flows (e.g., the sediment retained by a forest). Furthermore, if considering welfare effects of long-term changes in service provision, it is important to remember that the ecosystem production function may vary across time due to climate change, and the production function that takes ecosystem services as inputs may also vary through time as technology changes, so that calculations made ignoring such changes will provide inaccurate estimates of welfare impacts.

In the following simple model of individual utility, we ignore uncertainty, allowing us to pinpoint how changes in inputs to the utility function alter utility. Let us first consider the case in which an individual's utility is a function only of private goods purchased, where the individual attempts to maximize her utility by choosing quantities of these goods to consume given an exogenous set of product prices and a fixed amount of personal income, denoted by M. The constrained utility optimization faced by the consumer is given by

$$\max U = U(X) \text{ subject to } P'X \leq M \quad [4]$$

where X is a vector of n goods and services and P is a corresponding price vector. The solution to this problem results in a set of n Marshallian demand functions given by

$$x_i = x_i(P, M), \quad i = 1, \ldots, n \quad [5]$$

where eqn [5] represents the consumption of a given good that allows the individual to achieve her highest level of utility given her income and the prices of all goods. Substituting eqn [5] into eqn [4] results in an indirect utility function, representing utility as a function of prices and income based on the consumption decisions made by the individual when facing these prices with income level M

$$U = V(P, M) \quad [6]$$

Ecosystem services can be introduced in three different ways to a consumer's utility function. First, for services such as production of salmon or air quality in Yosemite Valley, ES_j could enter the utility function in eqn [4] directly. Other services such as vegetative pollutant retention that maintains downstream water quality for consumption might enter the utility function as an input to the production of a household produced good – in this case, the health status of the individual as given by $Z(ES_j)$. Regardless of whether the given ecosystem service directly impacts utility or only has an effect on utility through the production of Z, ES_j would appear as an argument of the indirect utility function described in eqn [6].

The third mechanism through which ecosystem services might impact an individual's utility is by affecting the production process of one of the private goods that the individual consumes. In this case, the welfare effects of a change in service provision would manifest themselves through changes in the price of the privately produced goods or inputs to their production, which could alter one or more elements of P or could alter M or both. As an example, it might be the case that the harvest of fir trees for their timber could result in a reduction in the capacity of the landscape to prevent sediment

from being carried into reservoirs during rainfall events. This result means that the sedimentation of the reservoir could occur at a faster rate than the operator of the hydropower station might have planned. As such, there could be an increase in the cost of power production, a resulting increase in the price of electricity, or possibly a reduction in rents to the hydropower owner and the income of plant employees. Under different conditions, the change in income alone could be used to measure the welfare impact, whereas other conditions would call for the inclusion of opportunity costs in the calculation of the welfare effects (See Freeman, 2003 for more details of this issue).

We have isolated the mechanisms through which ecosystem services impact individual utility. The change in utility resulting from a change in ecosystem-service provision could theoretically then be used to quantify the value of these services to the individual. Unfortunately, utility change is neither measurable for individuals nor additive across individuals, so we cannot use utility itself as a measure of welfare change. Instead, we can use the amount of money that must be offered or taken from an individual in conjunction with a change in the provision of ecosystem services to ensure that she maintains a constant utility level as our measure of the welfare impact of changes in ecosystem-service provision.

If we are interested in a small or marginal change in ES_j, we can identify the resulting marginal welfare change by differentiating the indirect utility function with respect to ES_j and M to identify the change in M necessary to compensate for the change in ES_j:

$$-\frac{dM}{dES_j} = \frac{\partial V/\partial ES_j}{\partial V/\partial M} \quad [7]$$

The expression in eqn [7] will be positive if ecosystem-service provision increases and negative if provision decreases. If the changes in ecosystem-service provision and the resulting welfare impacts are relatively large or nonmarginal, then the concepts of compensating and equivalent variation are used to measure the welfare effects of these changes. Compensating variation represents the amount of money that must be taken away from (or given to) an individual to make her just as well off after the change in ecosystem-service provision as she was before the change. We can implicitly define compensating variation using the indirect utility function as

$$V(P^1, ES_j^1, M^1 - CV) = V(P^0, ES_j^0, M^0) \quad [8]$$

where superscript 0 represents initial levels and superscript 1 denotes subsequent levels of the parameters, respectively. Presented in this way, the compensating variation has the same sign as the welfare change. For example, if the only change is an increase in ecosystem-service provision, meaning that $ES_j^1 > ES_j^\circ$, while $M^1 = M^\circ$ and $P^1 = P^\circ$, then the individual will clearly have to give up some income to be indifferent between the initial and subsequent parameter values, so that $CV > 0$.

Although compensating variation uses the utility level of the initial condition as the base case, equivalent variation uses the utility level of the subsequent condition as the base case. The equivalent variation is defined as the amount of money that must be given to (or taken from) an individual in lieu of a change in the quantity of ecosystem-service provision to make her just as well off as she would have been with the change in provision. We can implicitly define equivalent variation using the indirect utility function as

$$V(P^1, ES_j^1, M^1) = V(P^0, ES_j^0, M^0 + EV) \quad [9]$$

Once again, we see that equivalent variation, defined in this way, has the same sign as the welfare change.

Having identified the three pathways through which changes in the provision of ecosystem services might impact welfare and introduced two measures to quantify the welfare impacts of such changes, we can now explore how compensating and equivalent variation can be used to quantify these welfare impacts. Thus far we have implicitly expressed *CV* and *EV* in terms of the indirect utility function, which expresses an individual's utility as a function of prices and income based on her consumption decisions made in an attempt to maximize utility with her fixed amount of income. We can now explicitly define *CV* and *EV* in terms of an individual's expenditure function.

The expenditure function is obtained by identifying the consumption decisions that would be made by an individual who would like to achieve a threshold level of utility while minimizing the cost of her consumption choices. The expenditure function is derived by solving the following constrained optimization problem:

$$\min P'X \quad \text{s.t. } U(X, ES_j) \geq U^0 \quad [10]$$

where P represents the vector of product prices, X represents the vector of consumption units, and U° represents the threshold level of utility that the individual is attempting to achieve. It should be noted that the objective in expression [10] represents the dual to the utility maximization problem presented in expression [4].

By substituting the cost-minimizing demand for X into the objective function in eqn [10], we can identify the minimum amount of money necessary to achieve the targeted level of utility, given market prices and ecosystem-service provision,

$$e = e(P, ES_j, U^0) \quad [11]$$

We are now able to explicitly describe compensating and equivalent variation of a change in prices or ecosystem-service provision. The compensating and equivalent variations for a price change are given by

$$CV = e(P^0, ES_j^1, U^0) - e(P^1, ES_j^0, U^0) \quad [12]$$

and

$$EV = e(P^0, ES_j^0, U^1) - e(P^1, ES_j^0, U^1) \quad [13]$$

whereas those for a change in the provision of the focal ecosystem service are given by

$$CV = e(P^0, ES_j^1, U^0) - e(P^0, ES_j^0, U^0) \quad [14]$$

and

$$EV = e(P^0, ES_j^0, U^1) - e(P^0, ES_j^1, U^1) \quad [15]$$

where

$$U^i = V(P^i, ES^i_j, M) \quad i = 0, 1 \qquad [16]$$

The majority of the benefits that accrue to humans as a result of functioning ecosystems are goods that are not traded on traditional markets. As a result, most of these services will impact welfare through changes in the quantity of provision as opposed to impacts on the production of market goods by firms that will have price effects. That said, agriculture, fisheries, and timber are important components of the global economy, and changes in the provision of these ecosystem services will impact welfare not only through changes in the quantity of the service itself but also through changes in the market prices for these goods.

Thus far, we have focused on estimating the impacts of changes in the price or quantity of a particular ecosystem on an individual's welfare. In the context of benefit–cost analysis, we would like to understand which resource-use options provide the greatest benefits to society. Doing so requires a method of aggregating the estimates of individual welfare impacts over the relevant portion of society. Moving from individual to social welfare outcomes could be accomplished by summing up the welfare measures across relevant individuals who might be differentially affected due to variation in preferences or because the welfare changes themselves vary. Still, if the change results in net welfare gains to some and net welfare losses to others, when should such a change be deemed beneficial to society? One approach involves summing the *CV*s across all affected individuals, with the policy change deemed beneficial if this sum is positive (Kaldor, 1939) (remember that *CV* represents the compensation necessary for those made worse off by the policy to accept it as well as the willingness to pay of those who benefit from the policy to ensure it occurs). This approach recommends such policy changes that make it possible for those who benefit from the change to compensate those who suffer from the change. In practice, those transfers do not occur, so this method does not truly address the potential equity issues of policy evaluation, particularly when it is possible for the benefits to accrue unevenly across income levels in society. The inability to clearly resolve distributional concerns regarding policy changes is a key reason that benefit–cost analysis can be insufficient to decide policy matters.

In addition to the equity issues associated with aggregating estimates of individual welfare changes to a measure of social welfare impacts, ecosystem services face aggregation challenges due to spatial variation in welfare impacts. As previously mentioned, the welfare impacts of different ecosystem services will accrue over different spatial scales. For this reason, care must be taken in identifying the appropriate group of individuals affected by changes in the price or quantity of ecosystem-service provision. Bateman *et al.* (2006) provide a thoughtful discussion of how estimates of individual welfare impact might be obtained to minimize potential biases in the resulting estimates of social welfare impacts. We now turn to a discussion of how observable information can be used to estimate *CV* and *EV*, the theoretically accurate measures of welfare in the context of ecosystem services.

Welfare Measures with Marshallian Demand Functions

As a point of origin for this portion of the chapter, we must review the distinction between two popular methods for estimating demand. The ordinary, uncompensated, or Marshallian demand function represents the bundles of goods that a consumer would choose to purchase under each price and income combination, assuming she is able to perfectly solve her utility-maximization problem, as specified in eqn [4]. The Hicksian or compensated measure of demand is derived in a way to isolate the substitution effects of changes in prices from the associated income effects of such changes.

Briefly, a change in the price of a good that is in an individual's optimal consumption bundle (i.e., the good is a component of vector X in expression [10]) has two effects on an individual's consumption choices. First, the change in the price of the good will impact the individual's consumption of that good because the price change will make the good relatively cheaper (for a price reduction) or relatively more expensive (for a price increase) than the other goods in the individual's consumption bundle. This effect of the price change is known as the "substitution effect," because the individual will substitute toward the good whose price has changed from the other goods (for a price reduction) or away from the good whose price has changed toward the other goods (for a price increase). The change in the price of the good will also impact the individual's overall consumption bundle because the price change will either make the consumer more wealthy (for a price reduction) or less wealthy (for a price increase). This effect of the price change is known as the income effect, because the change in the price of the good impacts the income available to be spent on the consumption of goods in the consumption bundle, which will alter total consumption. Hicksian demand is derived in such a way to uniquely identify the substitution effect of a price change by preventing or compensating for any simultaneous income effect.

If we could observe this compensated demand curve, then the calculation of *CV* or *EV* would be quite straightforward as previously described. In reality, when a change in the price of a good does not take place, be it an ecosystem service or not, it is not possible to isolate the substitution effects of this change from the income effects. As a result, we cannot generally derive Hicksian demand curves from observed behavior.

When dealing with uncompensated demand functions, which are derived from observed behavior, we are unable to isolate the substitution effects of a change in price from its income effects (**Figure 2**). As such, we are unable to directly derive accurate estimates of *CV* and *EV* from Marshallian demand functions.

The measure of welfare associated with Marshallian demand is known as consumer surplus, *CS*. Consumer surplus is defined as the area beneath the demand curve and above the market-clearing price from the origin to the market-clearing quantity of consumption for the good in question. Consumer surplus can be affected by both changes in the price of a good as well as changes in the quantity or quality of the good

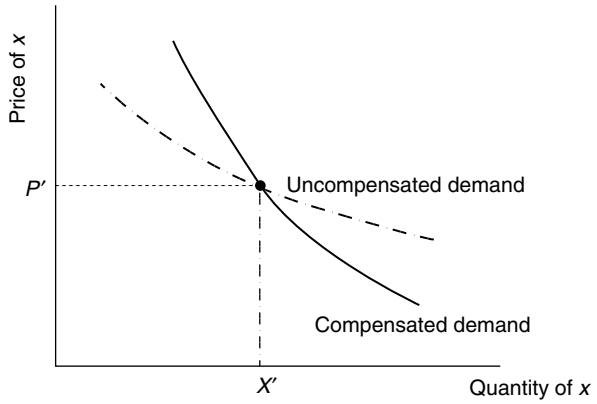


Figure 2 Compensated and uncompensated demand curves. The compensated demand curve illustrates the relationship between the price of a good and the quantity purchased, holding constant other prices and the individual's utility. The uncompensated demand curve illustrates the relationship between the price of a good and the quantity purchased, assuming that preferences, other prices, and income are held constant. The two curves intersect at point (x',p') because the individual's income at that point is just sufficient to achieve the utility level that is held constant on the compensated demand curve. The compensated demand curve is less responsive to price changes than the uncompensated demand curve, because the income effect, which is controlled for in the compensated demand curve, for normal goods is in the same direction as the substitution effect (i.e., a lower price leads to an increase in quantity purchased both because the good is relatively cheaper than other goods (substitution effect) and because the individual has more wealth to spend at the lower price (income effect)).

provided. Ecosystem services for which markets exist, such as provisioning services, will be able to impact welfare through price changes, whereas services for which there are no markets will have effects on welfare through changes in quality and quantity of the service provided.

The consumer surplus associated with a price change is given by the area beneath the demand curve between the initial and subsequent prices. In other words,

$$CS = -\int_{p_1^0}^{p_1^1} x_1(p_1,\hat{P},ES,M)dp_1 \qquad [17]$$

where x_1 represents the good whose price is changing, $\hat{P}$ represents the vector of prices for all other goods, p_1^0 represents the initial price of the good of interest, p_1^1 represents the price of this good after the price change, and the other variables are as previously defined.

The definition presented in eqn [15] represents the change in consumer surplus associated with a price change in the good of interest. This calculation is common for market goods. Therefore, the calculation in eqn [17] will be of use for provisioning services that undergo a change in market price. However, as we have explained above, the remaining categories of ecosystem services – regulating, cultural, and supporting services – are typically nonmarket goods. For these types of goods, the change of interest is related to the quantity or quality of provision (or both). In the absence of a market, the consumer surplus is defined as the area beneath the demand curve from the origin up to the observed level of consumption. In other words,

$$CS = -\int_{0}^{ES_1^1} ES_1(P,\widehat{ES},M)dES_1 \qquad [18]$$

where P represents the vector of prices for all market goods consumed, including provisioning ecosystem services, $\widehat{ES}$ represents the levels of all other ecosystem services consumed, and M represents income. The change in consumer surplus from a change in the quantity of ecosystem-service provision is found by differencing the consumer surplus associated with the original and shifted Marshallian demand curves.

It is clear that the observable (CS) and accurate (CV, EV) measures will differ under many conditions due to the presence of an income effect in observed conditions. Willig (1976) determined the precise differences between the measures, which are driven by the income elasticity of demand for the good in question and the magnitude of consumer surplus as a fraction of total income or expenditures. He concludes that the differences between the welfare measures are relatively small and basically trivial for most realistic conditions. Randall and Stoll (1980) conduct similar analysis for changes in the quantity of the nonmarket good of interest and also determine that consumer surplus is a suitable approximation for CV and EV, though these findings may not be as robust as those of Willig (Hanemann, 1991). Still, these findings are supportive of the use of consumer surplus, which can be empirically derived, as a welfare measure. It should be noted that subsequent efforts have demonstrated practical means of estimating CV and EV directly from observable data (Hausman, 1981; Irvine and Sims, 1998).

Estimating Social Value

As previously described, it is relatively straightforward to derive estimates of the social value of changes in the price or quantity of a good for which a market exists. For many goods, including several ecosystem services, markets do not exist. As a result, the methods previously described, which depend on knowledge of the consumer's demand function for the good of interest, can only be applied using novel methods to identify the demand functions for these goods of interest. The next two sections briefly introduce the methodologies of several different approaches to this problem, which can be broadly grouped into two categories. The first category of approaches is referred to as "revealed preference" because these methods utilize observed behavior to estimate demand for nonmarket goods. The next category is referred to as "stated preference" because these approaches utilize information gleaned from individuals through their responses to surveys, questionnaires, and interviews. Each broad category of approaches to assigning value to nonmarket goods has its benefits and costs, which we will discuss in the next two sections, beginning with revealed-preference methods.

Revealed-Preference Methods

These methods to value nonmarket goods use observed behavior as well as information from other markets to estimate

the demand for a particular nonmarket good. These approaches are referred to as "revealed-preference methods" because an individual's behavior in existing markets is used to reveal her preference for the nonmarket good being studied. There are two principal revealed-preference methods that are relevant to the topic of ecosystem services: the travel cost method for visits to a recreation site, and hedonic valuation, which uses the property value of a home to estimate the value of changes in ecosystem-service provision. Before introducing these two well-known approaches, we first consider an approach to value ecosystem services that are substitutes for market goods.

The specific revealed-preference method that is most appropriate for use in the valuation of a nonmarket ecosystem service depends on the relationship that exists between the ecosystem service of interest and other goods. In 1974, Karl-Göran Mäler laid the foundation for subsequent efforts to understand how preference information about a public good could be derived from the relationship between the public good and a substitute, market good. The correspondence between the estimated and actual welfare impacts of changes in the amount of ecosystem service provision depend on the substitutability of the ecosystem service and the market good.

Several ecosystem services – such as nutrient and sediment retention, storm-peak mitigation, and possibly air filtration – are perfect substitutes for goods that appear in an individual's utility function. In this case, where the ecosystem service and the market good are equally well suited to provide a valuable amenity to the individual, it is possible to identify the compensating variation due to a change in the amount of ecosystem-service provision using only information about the price of the substitute market good and the relationship between the market good and the ecosystem service in the production of the valuable amenity.

For example, consider an individual who maximizes utility, $U(X,Z,ES)$, subject to a budget constraint, $M - P'X - p_zZ = 0$, where ES is the ecosystem service of interest, Z is the market-good substitute for ES, and X is a composite of all other goods. It is possible to recast this problem as one in which the desired good or service provided by Z or ES, S, directly enters the utility function, so that the previous expression becomes $U(X,S(Z,ES))$. As an example, we can think of ES as the nutrient retention provided by forests and Z as the water filtration provided by a water-treatment plant, where S is the quality of water available for consumption downstream of the forest.

In the case of perfect substitutes, under certain conditions it is possible to identify the compensating variation of a change in the provision of the focal ecosystem service, ES. Let us imagine that the individual currently produces S by purchasing some amount of Z and by enjoying the benefits of the current amount of ecosystem-service provision, ES°. If the amount of ecosystem-service provision were to change (from ES° to ES^{1}, with $ES^{1} < ES^{\circ}$), and the reduction in the production of S due to the reduction in service provision could be perfectly offset by an increased purchase of Z, then the compensating variation associated with this change in ES is given by the cost of the increased purchase of Z necessary to return the individual to her original consumption of S. Similarly, if $ES^{1} > ES^{\circ}$, and some amount of Z is still purchased, then the compensating variation associated with this increase in ES is given by the reduction in cost associated with the production of the initial amount of S.

It is important to note that these results only hold if the individual is using a combination of both ES and Z to produce S before and after the change in ES. If this is true, and the individual is at an interior solution of her constrained utility-maximization problem, then we do not need information about her preferences to estimate the welfare impacts of a change in the focal ecosystem service. So, in the case of perfect substitutes between an ecosystem service and a market good, when the individual is at an interior solution before and after the change in ES, then the amount of money necessary to return the individual to her initial level of S is the same amount of money necessary to return her to her initial level of utility. If the ecosystem service of interest is an imperfect substitute for a market good, then the above result does not hold, though it does serve as a bound on the resulting welfare impact. Bartik (1988) explores the results for imperfect substitutes, and the key points are thoughtfully explained in Section 7.2.2 of Bockstael and Freeman (2005).

The finding that the amount of money required to return an individual to her initial consumption of a good that is equally well provided by an ecosystem service at no cost or a market good does not mean that we can use actual expenditures on the market-good substitute as a direct measure of the value of the ecosystem service. The reason for this point is that the previous finding holds when we are able to isolate the substitution effect of a change in the quantity of ecosystem service provided from the income effect of this change in service provision (i.e., in the context of Hicksian demand). In real-world situations, individuals are unable to isolate the income and substitution effects in this way, meaning that the actual expenditure on the perfect substitute for the ecosystem service following a change in the quantity of ecosystem service available may change even more than expected due to the opportunity to adjust consumption of other goods in the utility function. Bockstael and McConnell (2007) and Bartik (1988) provide further explanations about the relationship between actual expenditures and the actual welfare change as a result of a change in an environmental amenity.

Several notable nonmarket ecosystem services (e.g., nutrient and sediment retention related to water quality, sediment retention above reservoirs, carbon sequestration and storage, storm-peak mitigation) seem to have perfect substitutes among market goods that might allow for valuation estimates based on changes in expenditures on these market goods. Other nonmarket ecosystem services (e.g., recreation, esthetic or cultural benefits, and filtration of air pollutants by vegetation) can be thought of as complements and not substitutes to specific market goods. It is possible to estimate the value of changes in the quantities of these ecosystem services by envisioning them as attributes of heterogeneous goods in order to isolate how changes in the provision of these services impacts the willingness to pay for the complementary good. Two of the best-known revealed-preference valuation approaches, the travel cost method and hedonic price studies of the housing market, are examples of this approach.

The hedonic travel cost approach (Brown and Mendelsohn, 1983) differs from the travel cost procedure (Hotelling, 1949; Clawson, 1959) in that it assigns value to characteristics of destinations as opposed to the different destinations themselves. In this approach, ecosystem services can be thought of as a measure of the quality of a location that could be a destination for recreation. In this context, individuals gain utility from the consumption of a composite market good, the amount of recreation opportunities enjoyed, and the quality of that recreation experience. In this case, there is not a market price for recreation, but the cost of producing the recreation experience can be used as a measure for the price of recreation. This price can be quantified using the costs of transportation and the individual's time, which is estimated using a function of the wage rate. If the transportation and time costs are constant, then the marginal cost of producing the recreation experience can be thought of as the price of the recreation experience. By observing the visitation to sites with different characteristics, as well as the cost of arriving at those sites based on the individual's point of origin, it is possible to estimate the value of changes in the quality of the sites, as measured by changes in provision of ecosystem services. This approach has been used to consider the impact of water quality on recreation (e.g., Phaneuf *et al.*, 2008).

A similar approach can be used to estimate the value of ecosystem services using observed variation in housing prices. In this approach, we assume that the individual's utility is a function of a composite good, as well as the various attributes associated with the individual's choice of housing. The price of any house will be determined by the quantities of the different attributes associated with the given house (e.g., square footage, number of bedrooms, air quality at the site, water quality at a nearby lake, pond, or coast). The hedonic price function maps housing attributes, including the relevant ecosystem service, into market prices for houses. The derivative of this function with respect to any attribute can be interpreted as the marginal implicit price of that attribute. There are some complications with using this approach to determine the welfare impacts of changes in ecosystem service provision at different sites due to the fact that the level of service provision is a choice variable in this approach (see Rosen, 1974; Bartik, 1987; Epple, 1987 for detailed discussion of the causes of this outcome and Bockstael and Freeman (2005) for a thoughtful summary of this issue).

The hedonic property value models have been successfully applied to estimate the value of several different ecosystem services. Leggett and Bockstael (2000) use the hedonic pricing method to determine the impact of water clarity, which is related to nutrient and sediment retention provided by extant habitat around the property, on house prices around Chesapeake Bay. Bin *et al.* (2009), consider the value to society of riparian-zone buffers, through their impacts on water quality, using the hedonic pricing method.

With revealed-preference methods, analysis of actual behavior is used to identify the value of the nonmarket good through its relationship with other goods or services that have observed prices. The hedonic pricing method involves deconstructing the price of a market good and attributing portions of the total value to certain good characteristics (i.e., square footage and number of bedrooms for a house) to identify the value of a particular nonmarket amenity. The travel cost method can be used to identify the social value of clear water through changes in visitation rates to a river or lake based on cost of travel and the clarity of the water.

For several ecosystem services, observed behavior within markets for related goods can be used to identify the value of these services by understanding how changes in the provision of these services is related to the demand for the related market goods. These revealed-preference approaches require the availability of detailed socioeconomic data on a fine scale (i.e., visitor or parcel-level resolution). One shortcoming of the revealed-preference techniques is that they fail to account for the nonuse value of different nonmarket goods, meaning that they may only provide lower bound estimates of the value of nature, capturing only the values stemming from use of the services provided. Stated-preference methods of nonmarket valuation can be used to identify the full value of such ecosystem services, including nonuse values.

Stated-Preference Methods

Stated-preference methodologies rely on the stated behavior of individuals in a hypothetical setting to identify the value placed on the nonmarket good of interest. Although there are several techniques based on stated preferences, such as conjoint analysis, choice experiments, and contingent valuation, the contingent valuation method is the most commonly applied in the context of ecosystem services (see Farber and Griner (2000) and Powe *et al.* (2005) for examples of the application of conjoint analysis and choice experiments, respectively). In the contingent valuation approach (Mitchell and Carson, 1989; Bateman and Willis, 1999), the researcher uses a survey to elicit information from the respondent about the value that she attaches to a given nonmarket good using her stated willingness to pay (WTP) to experience an increase in the focal good or her stated willingness to accept (WTA) to tolerate a reduction in the focal good.

We can return to the previous definitions of compensating and equivalent variation to illustrate how WTP and WTA are related to these established measures of value. As noted, let the individual's indirect utility, $V()$, be a function of the prices of market goods consumed, P, the amount of ecosystem services consumed, ES, and the individual's income level, M. Consider the situation in which the amount of ecosystem services consumed changes, while the prices of consumed market goods and income are held constant; then the compensating variation, CV, is given by

$$V(P,ES^1,M-CV)=V(P,ES^0,M) \qquad [19]$$

and the equivalent variation, EV, is given by

$$V(P,ES^1,M)=V(P,ES^0,M+EV) \qquad [20]$$

If $ES^1>ES^\circ$ – namely, the change in ecosystem services is viewed as a desirable outcome and an improvement in the amount of service provision – then $CV>0$ and $EV>0$. In this case, CV measures the individual's maximum WTP to ensure that the change occurs, and EV represents the individual's minimum WTA to forgo the change. If the change in ecosystem-service provision is viewed as unfavorable, then $CV<0$

and $EV<0$, with CV measuring the individual's WTA to tolerate the welfare-reducing change, whereas EV measures the individual's WTP to avoid the change.

The contingent valuation approach involves providing individuals with hypothetical situations in which there is a change in the level of ecosystem services provided (e.g., there is a potential project to restore a wetland that will increase nutrient pollutant retention so that the quality of downstream water bodies will be improved) and then eliciting from those individuals estimates of their WTP or WTA to estimate their perceived value of the focal ecosystem service. This example illustrates the importance of providing very specific information about the change of interest in the survey, particularly in the context of ecosystem services, whose joint provision may potentially result in multiple ecosystem services being impacted within a given hypothetical scenario. In the parenthetical example of restoring a wetland, such action would contribute additional storm-peak mitigation, increased opportunity for bird-watching and other recreational activities, and sediment retention in addition to the retention of nutrient pollutants focused on by the survey, making it difficult to identify the value of the nutrient retention if the survey were not thoughtfully designed. Diamond and Hausman (1994) present some challenges that must be overcome in interpreting responses to contingent valuation surveys as credible estimates of WTP and WTA, whereas Hanemann (1994) provides an overview of how survey design and study methodology can address these challenges.

Carson and Hanemann (2005) provide a clear and thorough discussion of the steps that must be taken to move from the responses elicited through a contingent valuation study to estimates of WTP or WTA for the hypothetical change in the nonmarket good of interest. Briefly, the process entails conceiving of the survey responses as a random variable, which might be the outcome of individuals trying to maximize utility functions that contain a random component, from the perspective of the researcher. Next, a link must be made between the distribution of WTP or WTA that will result from this introduction of randomness and the distribution of responses to the survey. As Carson and Hanemann (2005) note, this linkage depends on the survey format, particularly regarding its use of either open-ended or closed-ended, single-bound discrete-choice formats. In the former case, for a survey of WTP for an increase in the provision of a specific ecosystem service, the respondent would be asked, "How much are you willing to pay for the change from ES° to ES^{1} (where $ES^{1}>ES^{\circ}$)?" In the latter case, for the same survey the respondent would be asked, "Would you vote to support the change from ES° to ES^{1} if it would cost you $\$A$?" With the open-ended format, the response directly reveals the respondent's value of CV, whereas the closed-ended format response provides an interval in which the respondent's value of CV must lie. To appropriately estimate the relevant measure of value, the researcher must acknowledge the impact of survey design and other assumptions on the relationship between survey response and the distribution of the relevant measure of value. Again, Carson and Hanemann (2005) provide a thoughtful description of these issues.

If thoughtfully designed, contingent valuation surveys can be used to estimate the full value of changes in ecosystem-service provision. However, the challenges of providing adequate information to ensure that the respondents are assigning value to the specific service of interest are not trivial, and significant thought must go into the methodological development of such studies. Finally, the hypothetical nature of such estimates – meaning that the respondents are not actually being asked to pay the amount of money that they indicate as their WTP in the survey – can be used to question the validity of valuation estimates derived using this approach.

As with the revealed-preference methods of estimating the value of ecosystem-service provision, there are obstacles that must be avoided in applying the contingent valuation methodology. That said, this method has been applied by frequently cited studies to estimate the welfare effects of changes in ecosystem-service provision. Desvousges *et al.* (1987) use this approach to determine the impact of enhanced water quality on recreational experience on a span of river in Pennsylvania. Loomis *et al.* (2000) use contingent valuation to estimate the overall economic impacts of restoration along a stretch of the South Platte River in Colorado, which would affect the provision of several different ecosystem services simultaneously (wastewater dilution, water purification, erosion control, recreation opportunities) as well as habitat for fish and wildlife. Given the shortcomings of both the revealed- and stated-preference approaches, some authors have combined the methodologies with the hope of improving the accuracy of amenity value estimates. These dual approaches have been applied to the value of coastal water quality (Hanley *et al.*, 2003) and to the willingness to pay for recreational trout fishing in Nevada (Englin and Cameron, 1996). While such efforts require additional effort in data collection, such exertion may be rewarded with more precise estimates of amenity value.

Conclusion

As we are increasingly faced with the choice between habitat conversion and conservation, the importance of achieving efficient resource use increases as well. To understand the full benefits and costs of habitat conversion and conservation decisions, we must be able to quantify the value of conserved habitat. This article has reviewed several of the existing methods available for such a task. That said, identifying the value of different resource-use outcomes does not necessarily guarantee that the most desirable outcome will be achieved, or even that there will be agreement about the most desirable outcome. Even as our understanding of ecosystem function and its value to society increases, the issues of equity and political feasibility offer challenges to the successful implementation of policies aimed at achieving efficient resource use.

Appendix

List of Courses

1. Graduate Course in Environmental Science and Management
2. Advanced Undergraduate or Introductory Graduate Course in Environmental and Resource Economics
3. Graduate Course in Ecosystem Services

See also: Biodiversity, Human Well-Being, and Markets. Economic Value of Biodiversity, Measurements of. Economics of the Regulating Services. Modeling Terrestrial Ecosystem Services. Priority Setting for Biodiversity and Ecosystem Services. The Value of Biodiversity

References

Bartik TJ (1987) The estimation of demand parameters in hedonic price models. *Journal of Political Economy* 95: 81–88.

Bartik TJ (1988) Evaluating the benefits of nonmarginal reductions in pollution using information on defensive expenditures. *Journal of Environmental Economics and Management* 15: 111–127.

Bateman I, Day BH, Georgiou S, and Lake I (2006) The Aggregation of Environmental Benefit Values: Welfare Measures, Distance Decay, and Total WTP. *Lincoln University Discussion Paper 114*.

Bateman IJ and Willis KG (1999) *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. New York: Oxford University Press.

Bin O, Landry CE, and Meyer GF (2009) Riparian buffers and hedonic prices: A quasi-experimental analysis of residential property values in the Neuse River Basin. *American Journal of Agricultural Economics* 91(4): 1067–1079.

Bockstael NE and Freeman AM (2005) Welfare theory and valuation. In: Maler K-G and Vincent JR (eds.) *Handbook on Environmental Economics* 2: 517–570.

Bockstael NE and McConnell K (2007) *Environmental Valuation with Revealed Preference: A Theoretical Guide to Empirical Models*. Berlin: Springer.

Boyd J and Banzhaf S (2007) What are ecosystem services? The need for standardized environmental accounting units. *Ecological Economics* 63: 616–626.

Brown G and Mendelsohn R (1983) The Hedonic travel cost method. *Review of Economics and Statistics* 66(3): 427–433.

Brown G and Shogren J (1998) Economics of the endangered species act. *Journal of Economic Perspectives* 12(3): 3–20.

Carson R and Hanemann WM (2005) Contingent valuation. In: Maler K-G and Vincent JR (eds.) *Handbook on Environmental Economics*, vol. 2, pp. 821–936. Amsterdam: Elsevier.

Clawson M (1959) *Methods of Measuring Demand for the Value of Outdoor Recreation*. Resources for the Future Reprint No. 10.

De Groot RS, Wilson MA, and Boumans RMJ (2002) A typology for the classification, description, and valuation of ecosystem functions, goods, and services. *Ecological Economics* 41: 393–408.

Desvousges WH, Smith VK, and Fischer A (1987) Option price estimates for water quality improvements: A contingent valuation study for the Monongahela River. *Journal of Environmental Economics and Management* 14(3): 248–267.

Diamond PA and Hausman JA (1994) Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives* 8(4): 45–64.

Englin J and Cameron TA (1996) Augmenting travel cost models with contingent behavior data. *Environmental and Resource Economics* 7(2): 133–147.

Epple D (1987) Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. *Journal of Political Economy* 87: 59–80.

Farber S and Griner B (2000) Using conjoint analysis to value ecosystem change. *Environmental Science and Technology* 34(8): 1407–1412.

Freeman AM (2003) *The Measurement of Environmental and Resource Values: Theory and Methods*, 2nd edn. Washington DC: Resources for the Future, Inc.

Goulder LH and Kennedy D (1997) Valuing ecosystem services: philosophical bases and empirical methods. In: Daily G (ed.) *Nature's Services*, pp. 23–48. Washington, DC: Island Press.

Goulder LH and Kennedy D (2011) Interpreting and estimating the value of ecosystem services. In: Kareiva P, Tallis H, Ricketts TH, Daily GC, and Polasky S (eds.) *Natural Capital: Theory and Practice of Mapping Ecosystem Services*, pp. 15–34. New York: Oxford University Press.

Hanemann WM (1991) Willingness to pay and willingness to accept: How much can they differ? *American Economic Review* 81: 635–647.

Hanemann WM (1994) Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8(4): 19–43.

Hanley N, Bell D, and Alvarez-Farizo B (2003) Valuing the benefits of coastal water quality improvements using contingent and real behavior. *Environmental and Resource Economics* 24(3): 273–285.

Hausman JA (1981) Exact consumers' surplus and deadweight loss. *American Economic Review* 71: 662–676.

Heal GM (2000) *Nature and the Marketplace: Capturing the Value of Ecosystem Services*. Washington, DC: Island Press.

Hein L, van Koppen K, de Groot RS, and van Ierland EC (2006) Spatial scales, stakeholders and the valuation of ecosystem services. *Ecological Economics* 57: 209–228.

Hotelling H (1949) *Letter quoted by RE Prewitt in Economic Study of the Monetary Evaluation of Recreation in National Parks*. Washington, DC: United States Department of Interior.

Hooper DU, Chapin III FS, Ewel JJ, *et al.* (2005) Effects of Biodiversity on Ecosystem Functioning: A Consensus of Current Knowledge. *Ecological Monographs* 75(1): 3–35.

Irvine IJ and Sims WA (1998) Measuring consumer surplus with unknown Hicksian demands. *American Economic Review* 88: 314–322.

Jackson RB, Jobbagy EG, Avissar R, *et al.* (2005) Trading water for carbon with biological carbon sequestration. *Science* 310(5756): 1944–1947.

Kaldor N (1939) Welfare propositions of economics and interpersonal comparisons of utility. *Economic Journal* 49: 549–552.

Leggett CG and Bockstael NE (2000) Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management* 39: 121–144.

Loomis J, Kent P, Strange L, Fausch K, and Covich A (2000) Measuring the total economic value of restoring ecosystem services in an impaired river basin: results from a contingent valuation survey. *Ecological Economics* 22: 103–117.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Wellbeing: Synthesis*. Washington, DC: Island Press.

Mitchell RC and Carson RT (1989) *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.

Phaneuf DJ, Smith VK, Palmquist RB, and Pope JC (2008) Integrating property value and local recreation models to value ecosystem services in urban watersheds. *Land Economics* 84(3): 361–381.

Powe NA, Garrod GD, and McMahon PL (2005) Mixing methods within stated-preference environmental valuation: choice experiments and post-questionnaire qualitative analysis. *Ecological Economics* 52(4): 513–526.

Randall A and Stoll JR (1980) Consumer's surplus in commodity space. *American Economic Review* 71(3): 449–457.

Rosen S (1974) Hedonic prices and implicit markets: product differentiation in perfect competition. *Journal of Political Economy* 82: 34–55.

Salzman J (1997) Valuing ecosystem services. *Ecology Law Quarterly* 24: 887–904.

Satchell M (June 25, 1990) The endangered logger: Big business and little bird threaten a northwest way of life. *US News & World Report* 108(25): 27–29.

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, and Siemann E (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277: 1300–1302.

Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of earth's ecosystems. *Science* 25: 494–499.

Willig RD (1976) Consumer's surplus without apology. *American Economic Review* 66: 589–597.

Vents

Cindy Lee Van Dover, College of William & Mary, Williamsburg, VA, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 747–753, © 2001, Elsevier Inc.

Glossary

Autogenic species A species that creates habitat for other organisms; this habitat would not exist if the autogenic species was absent.
Chemoautotrophic Primary production of organic matter using energy derived from chemical reactions rather than from photons.
Dyke intrusion Upwelling magma; may reach the surface of the crust to flow out as lava and solidify as basalt.
End-member fluid In black smokers, the hot, chemically modified fluid that develops at depth in the crust and exits undiluted by seawater.
Euphotic zone Sunlit upper layer of the ocean that can support photosynthesis.
Heterotrophy Organism that relies on organic material (other organisms, pieces of organisms, dissolved organic material) for its nutrition.
Hydrothermal plumes Created by rising black smoker effluents that mix turbulently with the surrounding seawater, becoming dilute and eventually reaching neutral buoyancy (typically ~200 m above the black smoker source); once neutral, they spread laterally and are detectable by chemical and physical tracers for kilometers.
Mid-ocean ridge Crustal accretion boundaries between major tectonic plates that make up the ocean crust.
Phylotypes Molecularly differentiated taxa.
Plume prospecting Use of chemical and physical anomalies in hydrothermal plumes to locate hydrothermal vents.
Spreading rate The rate at which tectonic plates move apart from each other.
Symbiont One of a pair of coexisting, mutually supportive organisms; in vent invertebrate-bacteria symbioses, the "symbiont" colloquially refers to the bacterium associated with the host invertebrate tissue; endosymbiont—within the host invertebrate (intra- or extracellularly); episymbiont—on the outer surface of the host invertebrate.
Tectonism Forces that disturb and dislocate the Earth's crust (earthquakes).
Vicariant event A historical event that isolates species' populations and leads to speciation (e.g., the closing of the Isthmus of Panama); usually considered in a geological context, but human activities also generate vicariant events.

Midocean Ridges and the Hydrothermal Setting

Although deep-sea hydrothermal vents occur in a variety of geological settings, including isolated back-arc spreading centers (such as the one that lies behind the Mariana Trench) and seamounts (e.g., Loihi, the incipient Hawaiian island), most hydrothermal systems lie along the network of volcanic midocean ridges that girdle the globe. Midocean ridges are the boundaries between the tectonic plates that make up the ocean crust; as such, they are the locus of episodic volcanism and earthquake activity associated with magmatic dyke intrusions and seafloor spreading. Most of the midocean ridge (and thus its hydrothermal systems) lies well below the euphotic zone, at depths greater than 2000 m. Where ridges intersect continents (e.g., at the head of the Gulf of California) or shoal to volcanic islands (e.g., Iceland and the Azores), hydrothermal systems may be found at shallower depths and even intertidally.

The geological pavement of deep-sea vents is typically hard basalt. Local relief of the basalt varies from flat, featureless sheets to smooth lobate surfaces a few to tens of centimeters high to rounded pillows of a meter or more relief. There are a small number of locales where sedimentation rate is in excess of volcanism, resulting in soft-sedimented bottoms with intruding basalt sills. The two best known sediment-hosted hydrothermal systems are in Guaymas Basin (Gulf of Mexico), where terrestrial runoff and pelagic production rates are high, and portions of the Gorda Ridge (off Oregon) and the northern Juan de Fuca Ridge (off Vancouver, Canada), where terrigenous inputs from the continental margin are significant.

Hot springs develop on midocean ridges where sea-water can penetrate ~2 km into the porous basalt crust to reach hot rock. In this "reaction zone" at depth, seawater reacts with the basalt, losing all of its magnesium and oxygen and becoming enriched in metals and reduced compounds, including hydrogen sulfide. The resulting acidic, hot water is thermally buoyant and rises to the seafloor to exit as high-temperature (typically 350 °C) fluids. As the vent fluid exits, it mixes with the cold, nearly neutral seawater and precipitates mineral sulfides of copper, iron, and zinc, forming substantial mineral chimneys known as black smokers. Where venting is persistent over time, chimneys coalesce to form larger, hydrodynamically and topographically complex sulfide edifices tens to hundreds of meters in diameter. Plumes emitted from black smokers rise ~200 m, until they become neutrally buoyant and spread out laterally. Hundredths of a degree deviations from ambient temperatures, elevated particulate concentrations, and modified chemistry allow water column geochemists to undertake "plume prospecting" by towing sensor packages at the depth of the neutrally buoyant plume. This prospecting method

permits extensive mapping of ridge segments to yield integrated measures of hydrothermal output. Plume prospecting is also critical in locating new hydrothermal fields in unexplored ocean basins.

Not all of the vent fluid escapes as high-temperature fluid. Large volumes of warm, sulfide-laden water are emitted as "diffuse flow" vents where ambient seawater has mixed with the end-member fluid. It is this warm-water flow that sustains productive populations of free-living chemoautotrophic, thermophilic microorganisms (up to 115 °C) and dense invertebrate populations (typically <40 °C). These diffuse flows may issue from porous surfaces of sulfide-mineral structures or directly from cracks and fissures in basalt lavas and can contribute an order of magnitude more flow than the focused output of black smokers at some sites. Because there are large gradients in chemistry and temperature within vent ecosystems, there is a large potential for diverse microbial and invertebrate types.

Hydrothermal vent fields, typically consisting of multiple discrete zones of focused and diffuse flows, are the basic unit of hydrothermal activity on a ridge axis. Vent fields vary in size from hundreds to thousands of square meters and represent islands of productive habitat within an otherwise relatively barren, hard-substrate environment.

The tectonic plates that make up the crust do not all move apart at the same rate. Spreading rate is a function of magma supply, with fastest spreading centers (100–170 mm $year^{-1}$) associated with well-developed magma chambers and supporting a higher spatial frequency of hydrothermal systems than slow-spreading systems (moving apart at 10–50 mm $year^{-1}$). As will be seen below, spreading rate may have a profound influence on diversity and biogeography of vent invertebrate communities because of its control on the spacing between vents. Other features vary with spreading rate, including the relative longevity of venting sites and bathymetric isolation of near-bottom waters. The two end members of the spreading rate continuum include the slow-spreading Mid-Atlantic Ridge, characterized by long-lived (thousands of years) vent systems and deep, basinlike valleys within which vents are located, and the fast-spreading East Pacific Rise, where vents are relatively short-lived (months to years to decades) and occur on topographic highs.

Microbial Diversity

Free-Living Microorganisms

Deep-sea hydrothermal vent ecosystems are celebrated as sites where chemosynthesis by microorganisms sustains primary production and a huge biomass of primary and secondary invertebrate consumers. The generation of biomass in the system is driven primarily by the microbial oxidation of sulfide, which yields biochemical energy for the fixation of CO_2. Inasmuch as the oxygen used in sulfide oxidation comes from seawater and ultimately from photosynthesis, the larger part of the vent ecosystem cannot be said to be totally independent of sunlight.

The potential metabolic diversity of vent microorganisms is large, with both aerobic and anaerobic poises. Reduced sulfur substrates, iron, manganese, ammonia, methane, and hydrogen can all serve as electron donors, leading to sulfide, sulfur, and thiosulfate oxidation, iron and manganese oxidation, nitrification, methane and hydrogen oxidation, denitrification, sulfur and sulfate reduction, and methanogenesis. While microbial isolates may be induced to be strictly autotrophic in the laboratory, mixotrophy may be the rule in the natural environment, with assimilation of simple organic compounds (heterotrophy) whenever they are available. Physiological diversity of microorganisms may also be great, with microbial types optimized for life in extreme temperatures (perhaps to >115 °C; superthermophiles), high temperatures (50–115 °C; hyperthermophiles and thermophiles), intermediate temperatures (10–50 °C; mesophiles), and cold temperatures (<10 °C; psychrophiles). Microorganisms can be pressure tolerant or barophilic (unable to survive low pressure). Microbial habitats include low-temperature diffuse-flow environments, buoyant plumes, interstices of sulfide structures, and surfaces of rocks and organisms. There is also growing evidence for an expansive, deep subsurface biosphere within the porous upper layer of the ocean crust, wherever temperatures are compatible with life.

Diversity of microbial communities at the level of species or phylotype is one of the outstanding issues in microbial ecology. Isolates cultured from deep-sea samples may only deliver a few percent of the actual diversity within the system. Current research combines traditional culture methods with molecular characterization of natural communities. Within vent systems, microbial communities may be dominated by a single phylotype, as in the case of bacteria associated with mineral sulfides at black smoker chimneys on the Mid-Atlantic Ridge. Dominance by a small number of phylotypes is also observed at the Loihi Seamount hydrothermal vents. The dominance of microbial communities by a small number of taxa is reminiscent of species-abundance characteristics observed in studies of diversity in marine invertebrate communities from "extreme" environments. It is, however, too soon to conclude that there is any single pattern of microbial diversity in the free-living microbial populations at hydrothermal vents. An even greater unknown is the degree to which vent microorganisms express biogeographic patterns. The current dogma is that free-living bacteria have unlimited dispersal capabilities in the open ocean because they can drift indefinitely in an inactive state. Hydrothermal vents may prove to be an ideal system in which to test this hypothesis with rigor at a global scale.

Biotechnology industries have embraced the vent ecosystem as a source of novel microbial organisms and enzymes that operate under extreme conditions. The prospect of thermostable or barophilic enzymes in particular has led to systematic assays of deep-sea microbial communities for industrial applications. Where microorganisms live at meso- and thermophilic temperatures in organic-rich sediments, they are implicated in transformation and decomposition of freshly cracked organic material. Microbial hydrocarbon processors are obvious targets for extraction and design of enzyme systems by petroleum and environmental waste management industries. Microorganisms from vents are also screened for exceptional therapeutic properties. Biotechnology interests thus place an added emphasis on discovery and isolation of as many microorganisms as possible from vent ecosystems.

Bacterial–Invertebrate Symbioses

Deep-sea hydrothermal vents are distinguished from their shallow-water and terrestrial hot-spring counterparts by the hugely successful associations between chemoautotrophic, endosymbiotic microorganisms and their macroinvertebrate hosts. Endosymbiotic bacteria are now known from a number of deep-sea invertebrate groups, including the bivalve and gastropod mollusks, vestimentiferan and pogonophoran annelids, and sponges. It is the vestimentiferan tubeworms that exhibit the greatest anatomical and physiological accommodation of the symbionts. As adults, tubeworms lack a mouth and gut, having instead a special organ—the trophosome—derived from gut tissues and made up of host cells filled with bacteria, all bathed in hemoglobin-rich fluids. Symbiont acquisition in tubeworms is by ingestion from the environment when the worm is in its larval and/or postlarval stage and still has a digestive system. Is there more than one tubeworm symbiont species in the environment? Based on DNA–DNA hybridization techniques, the sulfide-oxidizing chemoautotrophic endosymbionts of two co-occurring species of vestimentiferans belonging to different genera (*Riftia pachyptila* and *Tevnia jerichonana*) were found to be identical. Geographically distant populations of *R. pachyptila* have been shown to host this same symbiont, although DNA fingerprinting can discriminate symbionts among populations. A geographically disjunct population of the vestimentiferan worm, *Ridgeia piscesae*, from within the same ocean basin also hosted this same symbiont. A second symbiont was discovered in the vestimentiferan *Lamellibrachia* sp., but, because this species occurs in the Gulf of Mexico, this may be a basin-scale difference in the distribution of infective symbiont species.

Vesicomyid clams found at hydrothermal vents also host chemoautotrophic bacteria and have nonfunctional guts. In contrast to the vestimentiferan tubeworms, symbionts are transmitted maternally in the clams and there is, accordingly, strict fidelity between a symbiont species and its host species. A map of phylogenetic relationships among species of vesicomyid clams is remarkably congruent with the phylogenetic map of their respective symbiont species, and is evidence of cospeciation of host and symbiont.

There is physiological diversity in microbial endosymbionts of vent taxa. Sulfide oxidizers predominate in vestimentiferan tubeworms, but there is at least one instance of a tubeworm that hosts a methanotrophic symbiont. Within bivalves, some symbionts use sulfide; others use the more oxidized and less toxic forms of sulfur, including thiosulfate or elemental sulfur. Several vent mussel species exhibit a dual symbiosis, harboring both methanotrophs and sulfur oxidizers. This diversity of symbiont capabilities within individuals increases metabolic flexibility and thus the range of environmental conditions under which the mussels may thrive.

Episymbiotic bacteria also occur in well-known associations with certain invertebrates. One of the best examples is the bacterial "fur coat" of the pompeii worm, *Alvinella pompejana*. This large polychaete worm lives in galleries on the sides of black smoker chimneys and experiences one of the steepest gradients of temperature along its body length of any invertebrate (40–50 °C). The dorsal surface of *A. pompejana* is colonized by a morphologically diverse consortium of microorganisms, including large filamentous bacteria visible to the naked eye. The anatomical relationship between bacteria and worm is exquisite in its complexity, but the functional role of the members of the consortium remains unclear. Both trophic and detoxifying roles have been suggested. Episymbiotic, chemoautotrophic bacteria are also known from the chitinous cuticle of vent shrimp, *Rimicaris exoculata*. In this instance, the bacteria are thought to contribute to the diet of the shrimp; they are also likely to play some role in sulfide detoxification. Molecular characterization of the shrimp episymbionts shows them to be a single phylotype that matches the phylotype of the dominant free-living microbial species associated with the sulfide minerals on which the shrimp lives.

While there is much known and far more to be learned about microbial diversity at the interface between vent fluid and seawater, recent speculation about a vast, deep subsurface biosphere has fueled efforts to tap into the fluid-filled voids of the upper ocean crust. Efforts to understand microbial abundance, diversity, and function within such a biosphere may yield ancient lineages or otherwise novel species adapted to unusual conditions.

Invertebrate Diversity

It is the lush, gardenlike accumulations of invertebrate biota that make hydrothermal vent ecosystems a prized subject for television documentaries. Tubeworms often star in the title role, with clams and mussels as supporting actors, but this view of a simple ecosystem is little more than a Hollywood gimmick for audiences with short attention spans. The idea that vent systems have low invertebrate diversity is both pervasive and potentially misleading and is in large part due to the paucity of quantitative studies of diversity at vents. There is also a widespread failure to scale diversity measures at vents to the truly minute global acreage of vent habitat compared to extensive areas of soft-sediment deep sea or terrestrial rain forest, where diversity is so celebrated. These cautionary notes aside, there is an *a priori* reason to expect that diversity at vents might be low: Sulfide is a potent toxin, poisoning a major enzyme system of cellular respiration (cytochrome c oxidase) and bringing to a halt the aerobic production of ATP. Multicellular organisms that live at vents all must have some means of avoiding sulfide toxicity. This requirement may be a fundamental determinant of higher order invertebrate diversity at vents—only those groups that have efficient sulfide detoxification systems can occur where sulfide is present.

If we consider diversity at the phyletic level, vents are depauperate in several marine phyla normally found in the deep-sea benthos, including Porifera, Echinodermata, and the lophophorates (Phoronida, Bryozoa, Brachiopoda). The scarcity of sponges (Porifera) and echinoderms is attributed to the lack of excretory and blood vascular systems in these groups and a consequent inability to tolerate elevated levels of sulfide or other toxic compounds associated with hydrothermal vents; where these phyla are represented at vents, the animals are associated with the periphery of a vent field or other circumstances where sulfide levels are likely to be low (as in a waning vent site). The colonial motif and asexual reproduction are

also apparently unsuccessful at vents. Colonial coelenterates such as gorgonians, antipatharians, and hydroids are common animals of the deep sea but so far are virtually unknown at vents. An exception is a colonial siphonophore (distantly related to the Portuguese Man o' War) that is often found in large numbers at dying vents.

A number of hypotheses have been put forth to explain the absence of colonial organisms at vents, including intense competition and the ephemerality of vents, which may preclude any advantage of vegetative reproduction. Colonial organisms would also be at a disadvantage by virtue of the connectivity between individuals, since exposure of one individual within the colony to sulfide jeopardizes the aerobic metabolism of the entire colony. Encrusting colonial species are difficult to posit, because the vent effluent is buoyant and would tend to lift off recruiting polyps. Further, if established, an encrusting species could cap the flow, resulting in either diversion of the flow to less resistant paths or pooling of the flow with concomitant, life-threatening elevation of sulfide concentrations and temperature. While strategies to cope with these kinds of physical and chemical challenges to colonial organisms are not inconceivable, they have so far not been observed.

Origins of Vent Invertebrate Taxa

The majority of species that occur at vents ($\sim$90%) are known only from vents. This high degree of endemicity is difficult to prove, given the lack of effective means of (and interest in) collecting organisms from nonvent, deep-sea basalts. Nevertheless, it is a distinctive fauna. Where does it come from? There is no single source. Some species belong to nonvent deep-sea genera (e.g., the squat lobster, *Munidopsis lentigo*); others have closest relatives among the shallow-water invertebrates (e.g., the polychaete *Ophryotrocha*). Several species, notably symbiont-containing invertebrates (tubeworms, clams, and mussels), have closest alliances to species known from other chemosynthetic ecosystems, such as brine and hydrocarbon seeps and whale skeletons. Some vent families (e.g., scale worms in the polychaete family Polynoidae) are broadly represented in the marine environment. Still others appear to be specialized taxa, known only from hydrothermal systems—the polychaete family Alvinellidae is a good example. Finally, some species appear to be most closely related to ancient Paleozoic or Mesozoic taxa, relict lineages that have found refuge in hydrothermal systems. These include several stalked barnacle species and an archaeogastropod limpet. Some taxa routinely found at vents have been hugely successful, undergoing substantial radiation. These include the siphonostome copepods, the archaeogastropod limpets, and the polynoid polychaetes; within each of these groups there are now dozens of described species.

Biogeography

When vents were first discovered in the eastern Pacific, there was much speculation about how many vents there might be, where they were, and whether the fauna would be cosmopolitan. We know now that there are likely to be hundreds, if not thousands, of hydrothermal vent fields along the midocean ridge system and that they can be found in every ocean basin. Furthermore, we know that the fauna is not cosmopolitan. Distinct biogeographic provinces occur, defined in part by the tectonic fabric and history of the ocean plates. Along a continuous ridge system such as the East Pacific Rise, which stretches from the head of the Gulf of California down to high latitudes in the Southern Hemisphere, the fauna of the hydrothermal vents undergoes subtle changes. The same species of tubeworms, mussels, and clams generally dominate the biomass (mussels are absent from the northernmost locales), but the details of their relative abundance change from site to site. These dominant taxa (especially the tubeworms and mussels) are autogenic bioengineers, creating complex three-dimensional habitat that becomes occupied by countless individuals of crustacean, mollusk, and polychaete species. We know little about how the ranges of the hundreds of smaller invertebrate species that live in association with these autogenic taxa vary along the length of the East Pacific Rise nor of how these ranges relate to the life history characteristics of each species.

At its northern limit in the Gulf of California, the East Pacific Rise goes terrestrial as a strike-slip fault (the San Andreas), reemerging in the submarine environment off northern California as the Mendocino fracture zone. From there, the spreading ridge axis between the Pacific and North American plates heads northward as the Gorda, Juan de Fuca, and Explorer Ridge systems. The vent fauna of these northern ridges is distinct from that of the East Pacific Rise at the species level, but there are alliances at the generic and familial levels. Verena Tunnicliffe (University of Victoria) notes that at one point in the geological history of these ocean plates, the East Pacific Rise and northern ridges were one continuous system. Overriding by the North American Plate was a vicariant event, bisecting the ridge system and resulting in subsequent development of the vent faunas in isolation and of two distinct, formerly related, biogeographic provinces.

Other biogeographic provinces are known: Deep-sea hydrothermal vents at back-arc spreading centers associated with western Pacific microplates support a distinctive invertebrate fauna at the species level, although at higher taxonomic levels, affiliation to the two eastern Pacific biogeographic provinces is patent. This affiliation may be related to the direct link that existed between the eastern and western Pacific via the now extinct Kula Ridge more than 40 million years ago. Deep-sea vents along the Mid-Atlantic Ridge support an even more disparate group of species and genera. Based on biogeographic characterization of these four major geographic regions, we expect that ridge systems in unexplored ocean basins (the Arctic and the Indian Oceans in particular) will yield hundreds of new species. Exploration and discovery remain hallmarks of the study of diversity and biogeography at vents.

Gene Flow and Genetic Diversity

Deep-sea hydrothermal vents on midocean ridges are distributed as linear, insular, and ephemeral arrays of endemic species. To persist in the face of certain local extinction, survival of

a species in any region is only possible where propagules can disperse to other vents. An understanding of gene flow along these arrays is essential if we are to understand species distributions and speciation processes at vents. Two basic models of gene flow, borrowed from island biogeography, have been applied to genetic data from vent populations. In the "stepping-stone" model, gene exchange occurs locally, between neighboring populations, and genetic exchange decreases with increasing distance. In the "island" model, there is thorough mixing of gene pools over a regional scale. Vent taxa conform to no single model of gene flow. Tubeworms on the East Pacific Rise, for instance, appear to follow a stepping-stone model, while mussels on the same ridge follow the island model. Still other species display a ridge-based isolation model, where topography of the ridge system (e.g., deep east–west fracture zones that offset the shallower north-south-trending ridge) creates a genetic filter. Studies of genetic diversity in vent systems by Robert Vrijenhoek (Monterey Bay Aquarium Research Institute) and his students highlight the significance of the number of populations in sustaining genetic diversity: vent species with abundant populations have higher genetic diversity than species with fewer populations. Genetic diversity in vent organisms (measured as percentage of polymorphic loci within a population) is also correlated with the order of establishment of species at nascent vents, with early colonists having double the genetic diversity of later colonizing species.

Biodiversity

For two decades, basic community descriptors of species richness and diversity eluded vent ecologists. Sampling was qualitative, with uneven efforts, and typically resulted in lists of species for vent sites without regard to distributions within specific microhabitats. An exception is the work of Fred Grassle (Rutgers University), who compared diversity of soft-sediment vent infauna with infaunal diversity in the surrounding soft-sediment, nonvent environment. His quantitative data demonstrated that diversity in vent muds is low compared to the relatively high species diversity of the "normal" deep sea. Further, at vents, one or two species account for 70 – 90% of the infaunal individuals, whereas in the nonvent deep sea, abundances are more evenly distributed among species, with the most common species making up less than 20% of the total. Similar results were found in studies by others of soft-sediment meiofauna. But soft sediments are the exception rather than the rule for deep-sea hydrothermal vents.

What about hard-substrate biodiversity at vents? Species richness measured at a site on the Mid-Atlantic Ridge based on replicate samples within a single vent habitat (mussel beds) and species–effort curves has been compared to values obtained using a comparable sampling effort at an intertidal mussel bed in south-central Alaska. The intertidal, photosynthetic site supported more than twice as many species than the deep-sea chemosynthetic site, consistent with the idea that diversity at vents is low. But diversity in intertidal mussel beds is well-known to be geographically variable; some intertidal mussel beds (e.g., along the coast of Japan) are reported to have species richness values comparable to what is found at the Mid-Atlantic Ridge vent site.

Vents allow us to study patterns of regional diversity over different scales of spatial frequency of habitat and in the absence of confounding climatic and anthropogenic effects. The Mid-Atlantic Ridge is a slow-spreading ridge system, with vents spaced relatively far apart (hundreds of kilometers). Diversity is predicted to be higher on faster spreading centers where vents are more closely spaced. With the same sampling methods as used at the Mid-Atlantic Ridge vent site, diversity has been measured at two vent fields on the ultrafast spreading southern East Pacific Rise. Diversity at these vents is comparable to that of the Alaskan intertidal mussel bed. Thus, whether one considers diversity at vents to be high or low depends only on which end members one chooses to compare.

Of theoretical importance is the fact that patterns in diversity at midocean ridge hydrothermal vents may be controlled by fundamental properties associated with plate tectonics and volcanism that determine the spatial frequency of vent habitat. Where vents are closely spaced, the likelihood of extinction is relatively low, and species tend to accumulate. Where vents are far apart, a premium is placed on effective dispersal, extinction becomes much more likely, and allopatric speciation is enhanced. This hypothesis leads to the prediction that slow-spreading systems, with distant vents, support a greater number of biogeographic provinces, each with lower species diversity than fast-spreading systems with closely spaced vents.

Area – spreading rate relationships have also been proposed as controls on species diversity at vents. In general, the well-known area – species relationship is driven by increasing numbers of habitats with increasing area. While presumably relatively unimportant when comparing diversity within habitat types (e.g., mussel beds), the increasing vent area (and, by inference, habitat diversity) with spreading rate should contribute to increasing species diversity within vent fields. Quantitative assessment of species diversity within all representative microhabitats of vent fields and measures of areal extent of fields along ridges of different spreading rates is not a tractable approach at this time.

See also: Archaea, Origin of. High-Temperature Ecosystems. Marine Ecosystems. Microbial Biodiversity. Thermophiles, Origin of

References

Gebruk AV, Galkin SV, Vereschaka AL, Moskalev LI, and Southward AJ (1997) Ecology and biogeography of the hydrothermal vent fauna of the Mid-Atlantic Ridge. *Adv. Mar. Biol.* 32: 93–144.

Humphris SE, Zierenberg RA, Mullineaux LS, and Thomson RE (eds.) (1995) *Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions*. Heavily oriented toward the physical, chemical, and geological contexts, but with some outstanding reviews of vent biology/ecology. Washington, DC: American Geophysical Union.

Jollivet D (1996) Specific and genetic diversity at deep-sea hydrothermal vents: An overview. *Biodivers. Conserv.* 5: 1619–1653.

Juniper SK and Tunnicliffe V (1997) Crustal accretion and the hot vent ecosystem. *Philos. Trans. R. Soc. London, A* 355: 459–474.

Karl DM (ed.) (1995) *The Microbiology of Deep-Sea Hydrothermal Vents*. Includes review articles on the geological and geochemical setting, free-living microbial

communities, symbioses, thermophiles, microbe–metal interactions, plume microbiology, and biogeochemical cycles. Boca Raton, FL: CRC Press.

Peek AS, Feldman RA, Lutz RA, and Vrijenhoek RC (1998) Cospeciation of chemoautotrophic bacteria and deep-sea clams. *Proc. Natl. Acad. Sci. USA* 92: 7232–7236.

Tunnicliffe V, McArthur AG, and McHugh D (1998) A biogeographical perspective of the deep-sea hydrothermal vent fauna. *Adv. Mar. Biol.* 34: 353–442.

Van Dover CL (1995) Ecology of Mid-Atlantic Ridge hydrothermal vents. In: Parson LM, Walker CL, and Dixon DR (eds.) *Hydrothermal Vents and Processes*, pp. 257–294. Special Publication 87. London: Geological Society.

Van Dover CL (2000) *The Ecology of Deep Sea Hydrothermal Vents*. A comprehensive overview of geology, chemistry, microbiology, trophic and reproductive ecologies, evolution and biogeography, and community dynamics of vent communities, with chapters on cognate communities and the relevance of vent ecosystems to the origin of life and astrobiology. Princeton, NJ: Princeton Univ. Press.

Vrijenhoek RC (1997) Gene flow and genetic diversity in naturally fragmented metapopulations of deep-sea hydrothermal vent animals. *J. Hered.* 88: 285–293.

Vertebrates, Overview

Carl Gans, University of Texas at Austin, Austin, TX, USA
Christopher J Bell, University of Texas at Austin, Austin, TX, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 755–766, © 2001, Elsevier Inc.

Glossary

Chordate A member of the group Chordata. The Chordata includes the most recent common ancestor of tunicates and cephalochordates and all of that ancestor's descendants. Tunicates, lancelets, hag-fishes, and vertebrates are all chordates.

Ectoderm An embryonic tissue that provides the future outside layer of the animal.

Ectothermy A method of body temperature control in which the animal utilizes external sources for gaining and giving up heat, thus achieving temperature control without affecting metabolic rate.

Endothermy A method of body temperature control in which the animal modifies its metabolic rate to achieve the desired body temperature.

Neural crest An embryonic tissue intermediate between neurectoderm and ectoderm, with cells migrating widely to their final destination. This tissue gives rise to anterior skeletal elements, many portions of the future head and pharynx, and all pigment cells. Sometimes also referred to as mesectoderm.

Neurectoderm An embryonic tissue that gives rise to the central tube of the nervous system.

Notochord A stiff, flexible, longitudinal rod running along the middorsal portion of the chordate body. It is situated dorsal to the coelom and ventral to the central tube of the nervous system.

Pharynx The anterior portion of the alimentary canal, characterized by lateral buds that provide skeletal support for the gill region.

Tuberculum interglenoideum An anterior projection of the first (cervical) vertebra in salamanders. The tuberculum interglenoideum bears articular facets that insert into the foramen magnum of the skull and provide additional articulation points between the skull and the vertebral column.

Introduction

The vertebrates include most of the major groups of animals that humans encounter and interact with on Earth, and in common parlance, the term "vertebrate" is often equivalent to "animal." Vertebrates include all the fishes, amphibians, reptiles, birds, and mammals; they are found on every major landmass and in all of the world's oceans and major seas. The fossil record of the group is excellent and extends back to the late Cambrian period of the early Paleozoic, more than 500 million years ago. The vertebrates are members of a nested series of more inclusive groups, extending all the way down the tree of life to a common ancestor they share with bacteria, the most primitive life forms known. In order to properly understand the vertebrates, it is necessary not only to briefly discuss the characteristic features unique to the group but also to evaluate the features they share with their closest living relatives, the hagfishes, lancelets, and tunicates. These organisms plus vertebrates constitute the group Chordata. In the pages that follow, we summarize some generalized characteristics of chordates and vertebrates, examine the early history of these groups as it is revealed from the study of the fossil record, and provide explicit definitions and diagnoses for all of the major chordate groups.

General Vertebrate Characteristics

The diversity among vertebrate groups makes it possible to utilize functional analyses and paleontological data to reconstruct a phylogenetic sequence of vertebrates and to place the currently surviving (extant) vertebrate groups into a historical sequence. This sequence permits the development of explanatory schemes for some conditions observed in humans. Consequently, the study of the comparative morphology and physiology of vertebrates has proved important for the understanding of the human condition. These studies in turn led to the development of curricular variants for students training themselves for careers in human medicine. Examples are approaches to applied physiology and morphology of the precursors' conditions. Such instructional variants may be presented as the sequential development of the history of a topic such as gas exchange, digestion, or embryology. Alternatively, the study may involve a series of steps, each explaining the condition in one of the subgroups of animals representing putative precursors for the vertebrate condition (i.e., basal chordates or echinoderms), or historical stages in the development of groups of vertebrates (i.e., fishes, reptiles, and mammals).

Although a few arthropods and mollusks achieve large, indeed giant, sizes, most of the largest species, whether terrestrial, aquatic, or aerial, are vertebrates, and this latter group includes the largest animals known. However, some vertebrates are small, almost tiny, and this poses interesting questions of how they perform the basic functions of reproduction, locomotion, respiration, food processing, and waste discharge. Vertebrate size reduction involves two strategies: (1) maintenance of the adult pattern at a reduced body size and (2) the acquisition of metamorphosis, meaning that the life processes, and the structures that facilitate them, change markedly at one or more stages and sizes of development.

Vertebrates may be visualized as elongate, tubular creatures possessed of a central tubule, this being a gut or alimentary

canal. The anterior end of the gut forms a mouth and buccal structures, whereas the posterior end is modified for the storage and ultimate elimination of waste. The central tubule is placed in a lined cavity, the coelom, that also houses various diverticulae of the gut that support aspects of such visceral functions as gas exchange (i.e., gills and lungs, heart, and vessels), digestion (pancreas and liver), fluid balance (kidneys), and reproduction (gonads). The entire animal is coated by a multilayered epidermis that protects the internal tissues from mechanical stresses and can serve as a physiological barrier. The space between coelom and epidermal layer is occupied by the somatic musculoskeletal system that facilitates locomotion and supports and maintains the shape of the coelom and hence of the entire animal.

The musculoskeletal system involves the cartilages and bones of the internal skeleton and the striated somatic musculature. Externally, the system underlies the epidermis, this portion forming a dermis having skeletogenic properties. Two additional median and tubular structures complete this architectural scheme. The first tube, the notochord, lies just dorsal to the gut and is filled with vacuolated cells, their turgor pressure turning the notochord into a stiffened skeletal rod. The second, still more dorsal, tube consists of neuronal cells that involve sensory function and transmission of signals to the musculature.

The entire animal may have various appendages such as fins or limbs, each having characteristic internal skeletal supports and motile capacity developed by the contraction of the attaching striated musculature, which is subdivided into myotomes, paired blocks that parallel the notochord. The notochord provides a capacity for resisting longitudinal stresses, while the appendages involve support of the animal above the substrate and the transmission of laterally directed force patterns.

Whereas this general pattern seems to characterize vertebrate structure, the group shows major curiosities. The various groups of nonvertebrate creatures can generally be diagnosed by a relatively few characteristics. In contrast, there are no fewer than 25 new characteristics shared by vertebrates (**Table 1**). These involve multiple and major physiological systems, for instance the multiple senses, the head skeleton, the vertebral column, and the digestive system. This clear diversity of vertebrate specializations has long raised questions about the adaptive reasons to explain why these comparable characteristics are modified in vertebrates and why and how a multiplicity of such characteristics might be associated and consequently subject to possibly simultaneous selection. Actually, the answer to this conundrum has been available for some 30 years. Rather than looking at the functions of the many "vertebrate characteristics," namely, aspects seen in the adult condition, it is more informative to base comparisons on the stages of vertebrate embryology or development. The traditional rules that posited firm conditions, such as that the ectoderm formed all the epidermis and the mesoderm all of the striated musculature, soon broke down. In particular, the neural crest, an embryonic tissue intermediate between the ectoderm and the neurectoderm, proved to be the source of many diverse tissues. Similarly, the neurogenic epidermal placodes formed a set of embryonic tissues that provided a source for many shared vertebrate characteristics (**Table 1**).

Table 1 Derived characters of vertebrata and their embryonic origin[a]

Vertebrate characters	*Embryonic origin*
Nervous system	
Sensory nerves with ganglia, cranial	NC, P
Sensory nerves with ganglia, trunk	NC
Peripheral motor ganglia	NC
Second- and higher-order motor neurons	NC
Forebrain	NC?
Chromatophores	NC
Paired special sense organs	
Nose	P
Eyes (accessory organs)	NC?, (P)
Ears	P
Lateral-line mechanoreceptors	P
Lateral-line electroreceptors	P
Gustatory organs	NC?, P
Pharyngeal and alimentary modifications	
Cartilaginous bars	NC
Branchiomeric muscle	ME
Smooth muscle of gut	MH
Calcitonin cells	NC
Chromaffin cells, adrenal cortex	NC
Circulatory system	
Gill capillaries, endothelium	ME
Major vessels, trunk	M
Wall of aortic arches	NC
Muscular heart	MH
Skeletal system	
Anterior neurocranium and sensory capsules	NC
Cephalic armor and derivatives	NC
Armor of trunk	NC??, ME?

[a]M, mesoderm; ME, mesoderm, epimere; MH, mesoderm, hypomere; NC, neural crest; NC?, intermediate between neurectoderm and neural crest; P, placodes; (P), placodes provide lens and peripheral component, but not sensory tissues. Origin of the various trunk armors is uncertain.

Source: Modified from Gans C (1993) Evolutionary origin of the vertebrate skull. In: Hanken J and Hall BK (eds.) *The Skull: Patterns of Structural and Systematic Diversity*, pp. 1–35. Chicago: Univ. of Chicago Press.

The most interesting aspect was the discovery that many otherwise indistinguishable adult tissues could be derived from different embryonic sources, depending on their position in the embryo. Thus cephalic striated muscle has a histological pattern equivalent to that of the striated muscle of the trunk but is of neural crest rather than mesodermal origin. Similarly, the bones of the skull are derived from neural crest and the more posterior and serially equivalent vertebral bones are of (sclerodermal) mesodermal origin.

Consequently, the vertebrate characteristics either are derived from neural crest of neurogenic epidermal placodes or are induced by these tissues. In contrast to searching for 25 functional issues, the developmental approach makes it possible to recognize two or three associated aspects and note that these provided the cause of the vertebrate condition.

Early Chordate and Vertebrate History

Vertebrates are members of a more inclusive group of animals, the chordates. The history of both chordates and vertebrates

extends back into the Cambrian period in the early Paleozoic. Although the fossil record of early chordates is generally quite poor, several fossil localities have produced fossils that are likely to represent the remains of early chordate animals. Controversy over the interpretation of anatomical features of these animals is in part due to the fact that the important characteristics by which we recognize chordates are soft anatomical features that are not easily fossilized. Exquisite preservation of soft-bodied organisms in the Burgess Shale in Canada (middle Cambrian, approximately 530 million years ago) resulted in the discovery of several specimens of a chordate-like animal named *Pikaia gracilans*. These important fossils appear to display major chordate structural features, including myotomes (segmented muscle blocks) and a stiff, longitudinal rod running along the middorsal portion of the body, the notochord. Some controversy regarding the precise phylogenetic position of *Pikaia* persists due to a lack of adequate description of these fossils in the primary literature. Recent discoveries in early Cambrian deposits in China stimulated renewed interest in the origin and diversification of chordates. *Yunnanozoon lividum* was described as an early representative of the cephalochordates (Chen et al., 1995; Chen and Li, 1997), and *Haikouela lanceolata* was recently identified as a potential craniate from the early Cambrian (Chen et al., 1999) but these interpretations are controversial (Shu et al., 1996). The earliest fossils attributed to the Vertebrata are enigmatic bits of bone from the late Cambrian described under the name *Anatolepis heintzi* The vertebrate affinity of these fossils was challenged on various grounds, but recent analyses showed they contain dentine, a form of the calcium phosphate mineral apatite (the major mineral component of bone) that is found only in vertebrate skeletal tissues (Smith et al., 1996).

Homoiosteles and stylophorans, Paleozoic animals with a highly asymmetrical body form and a calcite skeleton resembling that of echinoderms, are another problematic extinct group. Fossils range in age from middle Cambrian to middle Devonian. The body can be divided into two distinct regions: a compact, asymmetrical region and a long segmented appendage. Interpretation of these fossils as "calcichordates," a distinct assemblage of stem-group chordates (see Jefferies, 1986, 1997) is controversial. Most authors treat them as aberrant echinoderms, and the controversy over their affinity is based in part on interpretations of orientation and locomotion (see Parsley, 1997).

Vertebrate Classification

The classification of chordates has a long and complicated history. Traditionally, the vertebrates were divided into fishes, which occupied aquatic environments, and the tetrapods, which mainly occupied terrestrial habitats. At present, there are two major classification systems in use. The Linnaean classification system was first published by Linnaeus about 1758 and organizes major divisions of life into kingdoms (Plantae, Animalia, Fungi, etc.), with major body plans organized into phyla (singular: phylum). The basic ranks descend from phylum to class, order, family, genus, and species. This classification system is familiar to most teachers and students and remains the dominant classification taught to elementary and high school students in the United States (**Figure 1**). In the past several decades, a nonranked

Kingdom Animalia
- Phylum Chordata (chordates)
  - Subphylum Urochordata (tunicates)
  - Subphylum Cephalochordata (lancelets)
  - Subphylum Vertebrata (hagfish + vertebrates)
    - Class Myxini (hagfishes)
    - Class Cephalaspidomorphi (lampreys)
    - Class Chondrichthyes (chimaeras, sharks, skates, rays)
    - Class Osteichthyes (bony fishes)
      - Subclass Actinopterygii (teleosts, etc.)
        - (many orders)
      - Subclass Sarcopterygii (lobe-finned fishes)
        - Order Crossopterygii (includes coelacanths)
        - Order Dipnoi (lungfishes)
    - Class Amphibia (amphibians)
      - Order Gymnophiona (caecilians)
      - Order Anura (frogs and toads)
      - Order Caudata (salamanders)
    - Class Reptilia (reptiles)
      - Order Testudines (turtles)
      - Order Rhynchocephalia (tuatara)
      - Order Lacertilia (lizards)
      - Order Serpentes (snakes)
      - Order Amphisbaenia (amphisbaenians)
      - Order Crocodylia (alligators and crocodiles)
    - Class Aves (birds)
      - (many orders)
    - Class Mammalia (mammals)
      - Order Monotremata (echidnas and platypus)
      - Order Marsupialia (marsupials)
      - (18 placental orders)

Figure 1 A traditional, Linnean classification of the major chordate groups. The major vertebrate groups are ranked into eight categories, all given the equivalent rank of class.

classification system was described by Hennig, a European entomologist. This classification system, termed "phylogenetic systematics" or "cladistics," organizes taxa into nested sets of monophyletic groups based on common ancestry (a monophyletic group contains an ancestor and all of that ancestor's descendants). A classification system based on natural groupings of taxa (based on ancestor–descendant relationships) is thus achieved. Phylogenetic systematics is rapidly becoming an important alternative classification system and is increasingly being utilized in biological sciences (including paleontology). Although we provide ancestry-based definitions for all the major chordate groups discussed below, we include three figures to illustrate the differences in classification systems now in use. **Figure 1** illustrates a version of a Linnaean classification system, in which taxa are placed in equivalent ranks. **Figures 2** and **3** are graphic representations of cladistic classifications, showing the hierarchical arrangement of nested taxa (**Figure 2** shows the relationships of the major chordate groups, and **Figure 3** shows the relationships of the major groups nested within the taxon Tetrapoda).

Definitions and Diagnoses of Major Chordate Groups

The major chordate groups discussed below are *defined* in terms of their ancestry; in other words, membership in a given taxon is dependent upon shared ancestry. Taxa are *diagnosed* by

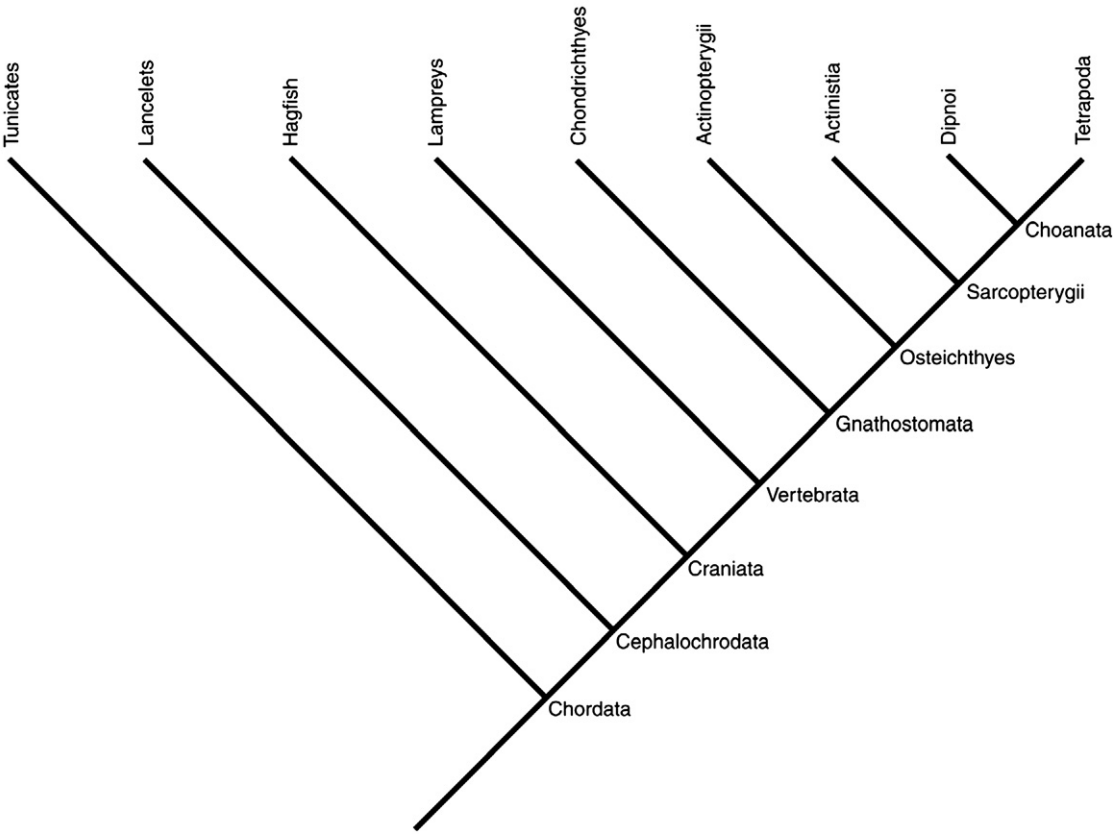


Figure 2 A graphic representation of a cladisitic classification of chordates. The major groups are arranged in a nested hierarchy, and taxon names are defined based on shared ancestry. The groups nested within Tetrapoda are depicted in **Figure 3**. See text for definitions and diagnoses.

a list of characteristics hypothesized to have first appeared in the common ancestor of the members of the group; this assemblage of characters can be used to identify a taxon (see Rowe, 1987, for further discussion).

Chordata

Chordata is defined as the most recent common ancestor of tunicates and cephalochordates, and all of that ancestor's descendants. Chordates share a common, generalized body plan and can be diagnosed by four features shared by members of all the major groups: a notochord, pharyngeal slits, a hollow dorsal nerve tube, and a postanal tail (an extension of the body posterior to the anus). The notochord is a stiff, flexible rod running dorsal to the coelom along the length of the body beneath the central nervous system. In many vertebrates, the notochord is almost completely replaced by blocks of bone forming vertebrae. Pharyngeal slits are lateral openings from the pharynx (an organ situated behind the mouth that constitutes part of the digestive tract). These features are found in all chordates at some point during their lifetime, but they are not retained as fully functional units throughout life in all chordates.

The tunicates (urochordates) are a group of basal chordates represented by approximately 3000 species. They are entirely marine. Some tunicates spend their entire life as pelagic organisms, floating in the water column. One group of tunicates, the sea squirts, undergo a dramatic metamorphosis from a planktonic larval form to a sessile adult (attached to a substrate); a notochord is present in the tail of the larval stage but is resorbed in the adult during metamorphosis. The pharynx enlarges in the adult and expands to form a barrel-shaped branchial basket through which a constant flow of water passes as the animal filter-feeds.

Cephalochordata

Cephalochordata is defined as the most recent common ancestor of lancelets and Craniata, and all of that ancestor's descendants. Cephalochordates are diagnosed by the presence of segmented muscle blocks. The lancelets are marine organisms that live in shallow, tropical and temperate seas surrounding all the continental land masses except Antarctica. Two major groups are recognized and are classified under the names *Branchiostoma* (23 species) and *Epigonichthyes* (7 species; Poss and Boschung, 1996). Lancelets lack a well-differentiated head region. A rudimentary brain is present, but significant regions of the brain found in vertebrates appear to be absent. A recent volume addressing many aspects of lancelet biology (Gans *et al.*, 1996) provides an important foundation for future research on this group.



Figure 3 A graphic representation of a cladisitic classification of tetrapods. The major groups are arranged in a nested hierarchy, and taxon names are defined based on shared ancestry. See text for definitions and diagnoses.

Craniata

Craniata is defined as the most recent common ancestor of hagfishes and Vertebrata, and all of that ancestor's descendants. Craniates are diagnosed by the presence of a differentiated head, with well-developed paired sense organs designed to act as distance receptors. There are 43 described species of hagfishes; all are marine organisms, and the group has a worldwide distribution. Hagfishes are scavengers, feeding on dead and moribund fishes, worms, mollusks, and crustaceans. They lack recognizable vertebral structures but have weakly developed, irregularly shaped plates that are positioned almost randomly along the nerve chord in the tail region. Their eyesight is poor, but they have well-developed senses of smell and touch. Reproductive biology is poorly understood; only a few embryos are known. Hagfishes were not known as fossils until 1991. The first reported fossil was found in Pennsylvanian age rocks in Illinois (Bardack, 1991). It differs from living forms in at least a few features (better developed eyes, different position of the gills, and some aspects of the feeding apparatus).

Vertebrata

Vertebrata is defined as the most recent common ancestor of lampreys and Gnathostomata, and all of that ancestor's descendants. As noted above, Vertebrata is diagnosed by a suite of characters (see **Table 1**). In all vertebrates there is at least some development of vertebral elements. The basal members of Vertebrata, the lampreys, have only rudimentary development of these structures. Lamprey vertebral elements consist of tiny distinct cartilages, situated above the dorsal nerve cord. The homologies of these elements with vertebral units in other vertebrates are not known. Many adult lampreys are parasites on other fish, and sometimes act as scavengers. They have a circular mouth with a rasping tongue; the mouth is used like a suction cup to attach the lamprey to a fish, and the tongue is used to open a hole in the body cavity of the fish, providing access to the body fluids on which the lampreys feed. Some lampreys spend significant portions of their adult lives in marine habitats, but all lampreys return to freshwater to breed. There are approximately 32 living species of lampreys.

Many earlier classifications placed lampreys, hagfishes, and a diverse suite of extinct jawless fishes into the class Agnatha ("without jaws"). A great diversity of jawless marine vertebrate groups populated Ordovician, Silurian, and Devonian seas. One enigmatic group, the conodonts, were first discovered in 1856 and subsequently were found to be remarkably abundant in marine rocks dating from the late Cambrian through the Triassic. For over a century, conodonts were known only from abundant, tiny, phosphatic, toothlike fossils that showed

a remarkable diversity in form. These fossils were, and remain, important tools for establishing correlation and relative chronology of Paleozoic and early Mesozoic rocks. The systematic position of conodonts was debated for decades, and various authorities suggested affinities with many different animal groups. In 1983, a conodont fossil with soft-body preservation was reported for the first time (Briggs et al., 1983). Subsequent finds that preserve greater anatomical detail revealed several features that can be interpreted as a notochord, a dorsal nerve cord, myotomes, and a distinct head region with very large eyes. These features suggest that conodonts may well have been chordates, or possibly vertebrates, but the relationships of conodonts and other jawless fishes are still quite controversial (for reviews, see Benton, 1997; Zimmer, 2000).

Gnathostomata

Gnathostomata is defined as the most recent common ancestor of Chondrichthyes and Osteichthyes, and all of that ancestor's descendants. Gnathostomata is diagnosed by the presence of jaws, an enlarged forebrain, and a skeleton in the paired pectoral and pelvic appendages. The evolution of jaws provided new opportunities for vertebrates by opening the door to truly predatory habits, with new food-handling capabilities and a broader diversity in potential diet. Among living vertebrates the chondrichthyans constitute the basal members of Gnathostomata. There are approximately 850 species of living chondrichthyans, including the sharks, skates, rays, and chimaeras. The majority are marine, but a few species inhabit freshwater environments. There is no bony skeleton; the skeleton is predominantly cartilaginous, with only minor calcification occurring throughout. Males have special structures called claspers that help facilitate internal fertilization. The sharks first appear in the fossil record in the late Devonian, chimaeras appear in the Carboniferous, and rays appear in the Jurassic. The whale shark (*Rhincodon typus*) is the largest living fish, achieving lengths of up to 15 meters.

Osteichthyes

Osteichthyes is defined as the most recent common ancestor of Actinopterygii and Sarcopterygii, and all of that ancestor's descendants. Osteichthyes is diagnosed by the presence of lungs (see below) and by most of the bony skeleton being formed from cartilaginous precursors. Many bony fishes have an elongated sac, the gas bladder, situated dorsal to the digestive tract. Gas bladders can serve either a buoyancy control function (swim bladder) or a gas exchange function (respiratory gas bladder). In the latter condition, they are highly vascularized and serve as a supplementary reserve for oxygen to be used in respiration. Although gas bladders and lungs differ in their position relative to the digestive tract (lungs are ventral to the digestive tract) and lungs tend to be paired while gas bladders are single, their developmental origins and gross anatomical features are quite similar. These similarities provide some evidence that the two structures are homologous. Actinopterygii includes a remarkably diverse assemblage of fishes, including more than half of the living species of known vertebrates; over 24,000 valid species are known (see Eschmeyer, 1998). The skeleton of actinopterygians is extensively ossified and the internal pectoral skeleton has broad bony plates that facilitate support of fin rays.

Sarcopterygii

Sarcopterygii is defined as the most recent common ancestor of Actinistia and Choanata, and all of that ancestor's descendants. Sarcopterygii is diagnosed by the presence of fleshy, lobed appendages, with differentiated proximal limb bones in the pectoral and pelvic fins, and by the presence of a ring of four or more bony plates in the eye (the sclerotic, or scleral, ring). Actinistia includes at least one, and possibly two, species of living coelacanths. These fishes fist appear in the fossil record in the Devonian and make their last appearance in the fossil record near the end of the Cretaceous, approximately 70 million years ago. The discovery of a extant species of coelacanth off the coast of South Africa in 1938 was therefore quite surprising. Waters near the Comoro Islands (situated between the northern tip of Madagascar and the African mainland) yielded a second specimen in 1952, and subsequently approximately 200 additional specimens were recovered. An additional population of coelacanths was recently discovered in Indonesia (Erdmann et al., 1998; Forey, 1998). Analysis of the DNA of the Indonesian coelacanth suggests that it might belong to a new species, but this interpretation is controversial (Holden, 1999).

Choanata

Choanata is defined as the most recent common ancestor of Dipnoi and Tetrapoda, and all of that ancestor's descendants. Choanata is diagnosed by the presence of folded tooth enamel and choanae (internal nostrils that open directly into the palatal portion of the roof of the mouth). The lungfishes (Dipnoi) first appear in the fossil record in the Devonian and are represented by six living species. Their modern distribution is restricted to the southern continents of Africa, Australia, and South America. The Australian lungfish is the least derived of the three; it cannot live out of water but can survive in stagnant pools by coming to the surface to gulp air. The South American and African lungfishes can survive out of water for relatively long periods of time.

Tetrapoda

Tetrapoda is defined as the most recent common ancestor of Amphibia and Amniota, and all of that ancestor's descendants. Tetrapoda is diagnosed by the presence of differentiated fingers and toes and by a bony joint formed between the occipital condyles of the skull and the anterior vertebral elements.

Extant amphibians share three morphological characteristics: (1) loss of several skull bones found in extinct amphibian groups; (2) short or absent ribs; (3) pedicellate dentition (an upper tooth crown resting on a lower pedicel). The relationships among the three major extant amphibian groups are not well supported, but combined analyses of morphological and molecular data favor a sister taxon

relationship between Anura (frogs and toads) and Caudata (salamanders). Batrachia is defined as the most recent common ancestor of Anura and Caudata, and all of its descendants. In this arrangement, Batrachia and Gymnophiona (the caecilians) are sister taxa (**Figure 1**; see also Cannatella and Hillis, 1993).

There are approximately 165 species of living caecilians. All extant caecilians are limbless, and their distribution encompasses tropical and subtropical regions of Central and South America, Africa, India, and Southeast Asia. Most caecilians are fossorial, spending the majority of their life underground. One South American group (*Typhlonectes*) is secondarily aquatic. The otic capsules, exoccipital, basioccipital, and parasphenoid ossifications of the braincase are fused into a solid unit (the os basale). Caecilians possess a pair of sensory organs called tentacles that are unique among vertebrates. The tentacles are protrusible organs, usually situated between the external naris and the orbit, and probably function in chemoreception. Although the orbits of caecilians are often covered with skin (and sometimes bone), the eyes are photoreceptive in most species that have been examined. In some caecilians, the eye is closely associated with the tentacle and in at least one species, they eye is protruded from the skull during protrusion of the tentacle (Nussbaum, 1992). The derived groups of caecilians fuse many elements of the skull and lack a tail. The fossil record of caecilians is notably poor but extends back at least as far as the Cretaceous. The Jurassic *Eocaecilia* shares many characteristics with extant caecilians (including the reported presence of a tentacular foramen) but retains limbs and other anatomical features not seen in extant forms (Jenkins and Walsh, 1993).

There are approximately 415 described species of extant salamanders, the majority of which are found in the Northern Hemisphere. A significant radiation of tropical South American salamanders extends as far south as southern Brazil and Bolivia. Caudates share a number of morphological features of the skeleton, including the loss of the quadratojugal ossification and the presence of intravertebral spinal foramina in the atlas vertebra. Most salamanders also have a reduced number of bones in the pectoral girdle relative to extinct amphibian groups and frogs (caecilians lack a pectoral girdle), and they have an anteriorly projecting process on the atlas (the first vertebra), the tuberculum interglenoideum, which forms accessory articulations between the skull and the vertebral column. Members of one large assemblage of caudates (the Plethodontidae) lack lungs, and this condition is also found in one gymnophionan. The early history of salamanders is poorly known, but fossils belonging to Caudata (as defined above) are known from the Mesozoic. Stem-group salamanders (outside of Caudata as defined above, but on the evolutionary stem leading to modern salamanders) are reported from the Jurassic. A recent review of Mesozoic taxa was provided by Evans and Milner (1996).

There are more than 4100 described species of extant anurans, distributed on all the major continental land-masses except Antarctica. Skeletal features shared by frogs include the fusion of forearm bones (radius and ulna), fusion of the lower leg bones (tibia and fibula), elongation of the ankle bones, reduced number of trunk vertebrae (10 or fewer among extant frogs), and presence of a urostyle (a rodlike, posterior bony extension of the vertebral column). Many of these skeletal features are associated with jumping ability and evolved as early as the early Jurassic (Shubin and Jenkins, 1995). The fossil record of the group is extensive (see recent review by Sanchiz, 1998) and at least two stem-group frogs are now know from the early Triassic (Evans and Borsuk-Bialynicka, 1998).

Many amphibians have a larval stage followed by metamorphosis into the adult form. Metamorphosis was defined by Duellman and Trueb (1986, p. 173) as "a series of abrupt postembryonic changes involving structural, physiological, biochemical, and behavioral transformations." The transformations that take place in frogs are the most dramatic of any within the Tetrapoda. The majority of frog species have a larval stage in the life cycle (following hatching from an egg and preceding development of the mature adult body form) known as a tadpole. The tadpole has a more or less oval-shaped body with a tail; there is no distinct head region, but the mouth includes a beak and a number of rows of chitinous toothlike structures. During metamorphosis many morphological and physiological changes take place that completely transform the animal's appearance and the way it interacts with its environment (e.g., the tail is resorbed, limbs grow, the axial and appendicular skeleton is dramatically transformed, the digestive tract shortens, gills disappear, and respiration function is transferred to lungs). Many frogs show variants of direct development.

Amniota

Amniota is defined as the most recent common ancestor of Reptilia and Mammalia, and all of that ancestor's descendants. Amniota is diagnosed by a suite of characters, including presence of an amniote egg (see below), caniniform teeth, two or more vertebrae in contact with the pelvic girdle, internal fertilization, and keratin (a protein that acts as the building block for scales, nails, hooves, hair, and feathers). The amniote egg includes a series of extraembryonic membranes that surround the developing embryo and provide all the nutritional, waste disposal, and gas exchange requirements during development. In reptiles (including birds) and the Monotremata (basal mammals), the embryo and extraembryonic membranes are encapsulated in an egg with either a leathery exterior or a hard shell. Because this is a self-contained fluid-filled system, eggs can be laid away from water. Although fishes, amphibians, and most reptiles are often referred to as "cold-blooded," many species in these groups manage to maintain elevated body temperature at a constant level. This is achieved by shuttling between heat sources and heat sinks (ectothermy). However, birds and mammals generally maintain a constant temperature by changing the basal metabolic rate (endothermy). There is substantial evidence suggesting that some of the fossil reptiles did maintain body temperatures above the level of the environment. There has been much argument about the question of whether the mechanism by which this was achieved was ectothermy or endothermy. The earliest known members of the Amniota appear in the fossil record during the Pennsylvanian period.

Reptilia

Reptilia is defined as the most recent common ancestor of Chelonia and Sauria, and all of that ancestor's descendants. Reptilia is diagnosed by several features of the skull, including the presence of suborbital fenestrae (paired openings in the skull that are situated ventral to the orbital region).

There are approximately 260 species of extant Chelonia (turtles). They are found on all continental land-masses except Antarctica and also occupy all of the world's tropical and temperate oceans and seas. All turtles lack marginal teeth, but the earliest known stem-group turtles (*Proganochelys*) from the Triassic have a few small teeth on the palatal portion of the skull; all extant chelonians lack palatal teeth. The palate is firmly attached to the braincase. The presence of a hard, bony shell is one of the most distinguishing features of turtles. The upper portion of the shell (the carapace) is formed in part from the vertebral column and ribs. During development these elements fuse with dermal ossifications to form a generally continuous, hard bony shell. The generally flat, ventral plate (the plastron) of the shell is formed anteriorly by dermal bones and skeletal elements that are probably homologous with elements of the pectoral girdle in other vertebrates. The bony shell of most turtles is covered by a series of keratinous scutes; in some forms, these scutes are lacking and a leathery skin covers the carapace and plastron. Turtles are unique among vertebrates in having the pectoral girdle enclosed by the rib cage. Turtles first appear in the fossil record during the Triassic.

Sauria

Sauria is the most recent common ancestor of Lepidosauria and Archosauria, and all of that ancestor's descendants. Sauria is diagnosed by the presence of a specialized ankle joint, long and generally slender hind limbs, and relatively larger leg muscles.

Lepidosauria can be defined as the most recent common ancestor of Squamata and the extant rhynchocephalian (*Sphenodon*; see below), and all of that ancestor's descendants. Lepidosauria is diagnosed by various features of the skull and the presence of a transverse cloacal slit, a distally notched tongue, and a modified middorsal scale row. Rhynchocephalians were a very diverse and widespread group of lepidosaurs during the Mesozoic era, but by approximately 60 million years ago, they were extinct everywhere except in New Zealand. The last surviving rhynchocephalians are restricted to a few islands in New Zealand. These relict populations are classified in the genus *Sphenodon*, and two species are currently recognized.

Squamata includes all the lizards and snakes as well as the fossorial amphisbaenians. A number of anatomical features are shared by these three groups, including aspects of the skull, postcranial skeleton, musculature, and tongue. All squamates have paired, evertible hemipenes. Many squamate species lack, or have reduced, limbs and girdles.

There are approximately 4150 species of extant lizards, widely distributed on all the continents except Antarctica. A number of extinct groups were fully marine, but today only one species utilizes a marine habitat (the marine iguana of the Galapagos Islands feeds on marine algae). Many groups of lizards have lost or reduced limbs. The two North American species of *Heloderma* are venomous. Lizards first appear in the fossil record in the Jurassic.

Both morphological and molecular data indicate that snakes evolved from within lizards, but their precise phylogenetic position among other squamates is unresolved. There are approximately 2700 species of extant snakes, with a distribution encompassing tropical oceans and seas and all continents except Antarctica. Several species spend most or all of their lives in open ocean habitats. Most snakes are limbless, but some basal members retain vestigial pelvic girdle and limb elements. A few poorly known Mesozoic snakes had well-developed limbs. All snakes are carnivorous, and several lineages of advanced snakes evolved potent venom that lets them subdue prey. Most snake species have an extremely modified and flexible skull structure that permits them to eat animals much more bulky than themselves and to ingest prey that is of a greater diameter than their own head. Snakes first appear in the fossil record during the Cretaceous, but their fossil record as a whole is relatively poor and consists predominantly of isolated vertebrae.

Amphisbaenia is a small group of squamates with approximately 160 extant species. They are found in Africa, the Iberian Peninsula, the Middle East, Mexico, from Panama to Argentina in South America, and Georgia and Florida in the southeastern United States. As a result of their subterranean habits and the fact that almost one-third of the described species are known from only one specimen, many aspects of their biology and systematics are not well understood. Three species of *Bipes* retain anterior limbs, but all other species are limbless. The compact and well-ossified skull is the primary burrowing tool and is covered with relatively large scales. The scales of the skin are usually rectangular and are arranged in rings that encircle the body and allow rectilinear locomotion. Amphisbaenians first appear in the fossil record in the late Paleocene of North America.

Archosauria is defined as the most recent common ancestor of Crocodylia and Aves, and all of that ancestor's descendants. Archosaurs are diagnosed by a number of anatomical features and also by behavioral characteristics such as nest building and parental care of young. The two extant groups of archosaurs represent only a tiny fraction of the past diversity of this group, which includes the Mesozoic flying reptiles (pterosaurs), all the dinosaurs, and many extinct crocodylians.

There are 22 extant species of Crocodylia (alligators, crocodiles, and caimans), most of which are found in the Tropics. Two species of alligators extend into temperate regions in North America and China. The earliest known crocodylians were terrestrial, but all living species are found in aquatic habitats. Some extant species are found in nearshore marine habitats, but none are fully marine. Crocodylians have a complete secondary palate that forms the roof of the mouth (similar to the configuration seen in mammals). Stem-group crocodylians are known from the late Triassic, but the earliest true crocodylians are found in early Jurassic rocks.

Aves is defined as the most recent common ancestor of all extant birds, and all of that ancestor's descendants. There are approximately 9700 species of extant birds. In traditional

classifications, Aves was considered to be a separate class of vertebrates, equivalent in rank to the Amphibia, Mammalia, and Reptilia. Cladistic classifications recognize birds as a distinctive group of vertebrates, but one that is nested within Reptilia, specifically within the Dinosauria. Although not accepted by all ornithologists, there is strong morphological support for this hypothesis. Birds and dinosaurs share such morphological features as an in-turned femoral head, opposable thumb, and a perforate acetabulum. An additional suite of characters indicate that birds are nested within theropod dinosaurs. Aves (as defined above) is diagnosed by the loss of teeth, the presence of a pneumatic foramen in the humerus (upper arm bone), and fusion of uncinate processes with the ribs. Most extant bird lineages appear in the fossil record at least by the early Cenozoic, and several are reported from the Cretaceous.

Mammalia

Mammalia is defined as the most recent common ancestor of Monotremata and Theria (marsupials plus placentals), and all of that ancestor's descendants. Mammalia is diagnosed by a large number of anatomical and physiological features, including the presence of mammary glands and a unique jaw joint. Mammary glands in females produce milk to nourish the rapidly growing young during postnatal development. In most vertebrates the jaw joint is formed between the quadrate bone of the skull and the articular bone of the lower jaw, but in mammals, the jaw joint is formed between the dentary bone (the only bone in the lower jaw of mammals) and the squamosal bone of the skull. Bones homologous with the quadrate and articular are still present in mammals, but their positional relationship and function has been transformed. The incus and malleus in the middle ear of mammals are homologous with the quadrate and articular (respectively) in other vertebrates. The 4810 species of extant mammals can be found on all continents and major bodies of water on Earth.

Monotremata is represented by three living species (duck-billed platypus and two species of echidna) that are found only in Australia and New Guinea, but the fossil record of this group extends back to the Cretaceous and reveals that monotremes were once found in South America. The monotremes are the only mammals that lay eggs. The marsupials are found today only in Australia and surrounding islands, North America, and South America. Their earliest appearance in the fossil record is in the Cretaceous of North and South America, and fossils are known from all continental landmasses. Marsupials differ from other mammals in a number of physiological and anatomical features, especially those related to reproduction. The gestation period in marsupials is very short compared with that of other mammals and the young are born in a nearly embryonic state. After birth, development takes place in a pouch (or sometimes under a fold of skin) on the mother's abdomen. The Placentalia includes a diverse assemblage of mammals that include many of the most familiar mammals and the only mammals to adopt a completely or predominantly marine lifestyle (whales, dolphins, porpoises, seals, sea lions, walruses) and to achieve powered flight (bats). Human beings are members of a placental mammal group, the Primates.

See also: Amphibians, Biodiversity of. Birds, Biodiversity of. Endangered Amphibians. Endangered Mammals. Endangered Reptiles. Fishes, Biodiversity of. Mammals, Biodiversity of. Reptiles, Biodiversity of

References

Bardack D (1991) First fossil hagfish (Myxinoidea): A record from the Pennsylvanian of Illinois. *Science* 254: 701–703.
Benton MJ (1997) *Vertebrate Paleontology*, 2nd ed. Princeton, London: Chapman & Hall.
Briggs DEG, Clarkson ENK, and Aldridge RJ (1983) The conodont animal. *Lethaia* 16: 1–14.
Cannatella DC and Hillis DM (1993) Amphibian relationships: Phylogenetic analysis of morphology and molecules. *Herpetol. Monogr* 7: 1–7.
Chen J-Y, Dzik J, Edgecombe GD, Ramsköld L, and Zhou G-Q (1995) A possible Early Cambrian chordate. *Nature* 377: 720–722.
Chen J-Y, Huang D-Y, and Li C-W (1999) An early Cambrian craniate-like chordate. *Nature* 402: 518–522.
Chen J-Y and Li C-W (1997) Early Cambrian chordate from Chengjiang, China. *Bull. Mus. Nat. Sci.* 10: 257–273.
Duellman WE and Trueb L (1986) *Biology of Amphibians*. New York: McGraw-Hill.
Erdmann MV, Caldwell RL, and Moosa MK (1998) Indonesian 'king of the sea' discovered. *Nature* 395: 335.
Eschmeyer WN (1998) Catalog of Fishes California Academy of Sciences Center for Biodiversity Research and Information Special Publication 1.
Evans SE and Borsuk-Bialynicka M (1998) A stem-group frog from the Early Triassic of Poland. *Acta Palaeontol. Pol.* 43: 573–580.
Evans SE and Milner AR (1996) A metamorphosed salamander from the Early Cretaceous of Las Hoyas, Spain. *Philos. Trans. R. Soc. London, B* 351: 627–646.
Forey P (1998) A home from home for coelacanths. *Nature* 395: 319–320.
Gans C (1993) Evolutionary origin of the vertebrate skull. In: Hanken J and Hall BK (eds.) *The Skull: Patterns of Structural and Systematic Diversity*, pp. 1–35. Chicago: Univ. of Chicago Press.
Gans C, Kemp N, and Poss S (eds) (1996) The lancelets (Cephalochordata): A new look at some old beasts *Isr. J. Zool.* 42(Suppl.).
Holden C (1999) Dispute over a legendary fish. *Science* 284: 22–23.
Jefferies RPS (1986) *The Ancestry of the Vertebrates*. Cambridge, UK: Cambridge Univ. Press.
Jefferies RPS (1997) How chordates and echinoderms separated from each other and the problem of dorso-ventral inversion. In: Waters JA and Maples CG (eds.) *Geobiology of Echinoderms*, pp. 249–266 *Paleontological Society \Papers 3.*
Jenkins Jr. FA and Walsh DM (1993) An Early Jurassic caecilian with limbs. *Nature* 365: 246–250.
Nussbaum RA (1992) Caecilians. In: Cogger HG and Zweifel RG (eds.) *Reptiles and Amphibians*, pp. 52–59. New York: Smithmark Publishers.
Parsley RL (1997) The echinoderm classes Stylophora and Homoiostelea: Non Calcichordata. In: Waters JA and Maples CG (eds.) *Geobiology of Echinoderms* pp. 225–248. *Paleontological Society Papers 3.*
Poss SG and Boschung HT (1996) Lancelets (Cephalochordata: Branchiostomatidae): How many species are valid? *Isr. J. Zool.* 42: S13–S66.
Rowe T (1987) Definition and diagnosis in the phylogenetic system. *Syst. Zool.* 36: 208–211.
Sanchiz B (1998) *Encyclopedia of Paleoherpetology: Salientia, Part 4*. München: Verlag Dr. Friedrich Pfeil.
Shu D, Zhang X, and Chen L (1996) Reinterpretation of *Yunnanozoon* as the earliest known hemichordate. *Nature* 380: 428–430.
Shubin NH and Jenkins Jr. FA (1995) An Early Jurassic jumping frog. *Nature* 377: 49–52.
Smith MP, Sansom IJ, and Repetski JE (1996) Histology of the first fish. *Nature* 380: 702–704.

Vicariance Biogeography

Christopher John Humphries, The Natural History Museum, London, UK

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 767–779, © 2001, Elsevier Inc.

Glossary

Area of endemism An area recognized by congruent distributions of two or more groups of organisms.
Biogeography The study of the distributions of organisms on Earth and why the patterns exist.
Cladistic biogeography The combination of cladistics (phylogenetic systematics) with vicariance biogeography. A method that searches for patterns of relationship among areas of endemism.
Cladistics (phylogenetic systematics) A method of phylogeny reconstruction concerned with branching patterns. Sister group relationships are hypothesized on the basis of shared derived, or synapomorphic, characters.
Components Elements of groups of areas, or taxa, as determined by the branching pattern of a cladogram. For example, in a group comprising three taxa A, B, and C, when B is more closely related to C, there are two components, an ABC component and a BC component.
Dispersal The movement, or spread, of an organism from one area to another independent from other organisms and of historical events that change the natural distribution of the organism (see **Figure 1**).
Ecological biogeography Spatial patterns of organisms considered in terms of their interactions with other organisms and the environment.
Endemic A taxon, or group of taxa, restricted to a proscribed area and found nowhere else on Earth.
Historical biogeography Spatial patterns of organisms interpreted in terms of their concordance to patterns of Earth history.
Panbiogeography The examination of distributions and relationships on a worldwide scale; term introduced by Croizat (1964).
Plate tectonics (including continental drift) The theory that the Earth's crust is composed of plates that move relative to each other through seafloor spreading and subduction. As continents are found on plates, continental drift is one result of this movement.
Vicariance The existence of closely related taxa or biota in separate, disjunct areas, which have been separated by the formation of a natural barrier (vicariance event).
Vicariance event The splitting of a taxon or biota into two or more geographical subdivisions by the formation of natural barriers, e.g., mountain building, glaciation, or stream capture (see **Figure 1**).

Historical Biogeography

Biogeography means many different things to biologists and geographers of different persuasions depending on their outlooks and purposes of enquiry. It has been clearly noticed and well documented through accumulation of systematic distribution data for most groups of organisms over the past two centuries of global exploration that there are many organisms that exhibit particular distribution patterns on Earth. In efforts to interpret these patterns, biogeographers pose the question: "What lives where and why?" The usual answers invoke either ecological or historical explanations, or a combination of both (Nelson and Platnick, 1981; Craw *et al.*, 1999; Humphries and Parenti, 1999).

For the most part ecological biogeographers classify distribution patterns as communities in a general hierarchy of organization. Populations and species live in habitats that belong to different land classes in ecosystems within biomes. Few would dispute that many similarities of climate, topography, habitat, and landscape exist between, for example, the rain forests of South America, Africa, and Southeast Asia, but, when examined closely, the resident organisms in each rain forest are specifically and generically quite different. Similar species in similar ecological niches belong to phylogenetically disparate taxa that have particular phylogenetic and distribution histories that have occurred through geological time. Historical biogeographers focus on the older developments of global biotas in an effort to study the history of the Earth rather than relatively more recent events such as short-term species interactions and recent perturbations in distribution due to climatic changes.

Data for historical biogeography come from comparative biologists and systematists who include distribution information in their monographs and revisions. A major synthesis is found in the works of Croizat, especially in *Space, Time, Form: The Biological Synthesis* (1964; see Craw *et al.*, 1999). Croizat considered geological events and changes in Earth history and the "form-making" of species as two parts of a continuous historical process. To him, evolution of the Earth and the evolution of organisms are part and parcel of one series of historical events. Recently, greater attention has been given to these ideas, kindled by the general acceptance of plate tectonics and the notion that the continents are continually reshaped and changing in their spatial geometry (see Craw *et al.*, 1999). At the same time, there has been a widening interest in the use of cladistics for systematic investigation of organisms (Hennig, 1966). Today, cladistics is considered as the primary means of reconstructing the relationships of organisms. Comparing different groups of organisms sharing similar distribution patterns and occupying similar areas of

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00149-0

endemism with historical patterns in geology and Earth history allows biogeographers to provide retrodictive hypotheses of divergence and explanations of past and present diversity of organisms (Humphries and Parenti, 1999).

Vicariance, Dispersal, and Extinction

Particular species are endemic to certain areas today either because their ancestors originated there and their descendents have survived to the present day or because the ancestors occurred elsewhere and later some of the descendents dispersed into new areas. Such a difference invokes two, very different, historical explanations which in a shorthand way are described as vicariance or dispersal modes of diversification. In a vicariance explanation, an ancestral species (**Figure 1a**, 1) divides into two separate populations when a physical barrier is formed across the population such that neither of the two new populations can cross over the barrier (**Figure 1a**, 2). With time, the two separated populations evolve into two closely related species, shown here as black-and-white ovals (**Figure 1a**, 3). In a dispersal explanation, the range of one species is already bounded at least in one direction by a barrier (**Figure 1b**, 1), which is later crossed by some members of the population (**Figure 1b**, 2). If the organisms of the new population survive, successfully interbreed, and yet remain isolated from the "ancestral" population, they eventually evolve into a different species (**Figure 1b**, 3). Thus, in the vicariance explanation the appearance of the barrier is seen as the cause of the disjunction of the two species, whereas in the dispersal explanation the barrier is considered to be older than the disjunction.

Examples of Vicariance and Dispersal

By Victorian times and the emergence of evolutionary theory, it was clear that the distinction of dispersal and vicariance explanations was emerging. As described by Nelson and Platnick (1984), the English botanist Joseph Hooker (1817–1911) provided perhaps the earliest coherent vicariance explanations of plant distributions. Hooker studied the plants of the Southern Hemisphere in South Africa, Australia, southern South America, and New Zealand and concluded that their pattern of distribution suggested a former, more extensive flora, which had broken up by geological and climatic causes. Hooker compared the southern floras with those of the Northern Hemisphere and also concluded that the various southern floras were more closely related to each other than any were to the floras of the Northern Hemisphere. As a consequence, he also boldly concluded that divergence of the northern and southern floras must have occurred before the fragmentation of the southern floras. His vicariance hypothesis considered that initially a worldwide flora was first subdivided by a barrier into the northern and southern floras and then the latter were further subdivided by expanses of ocean—barriers severing once continuous floras that no longer exist today.

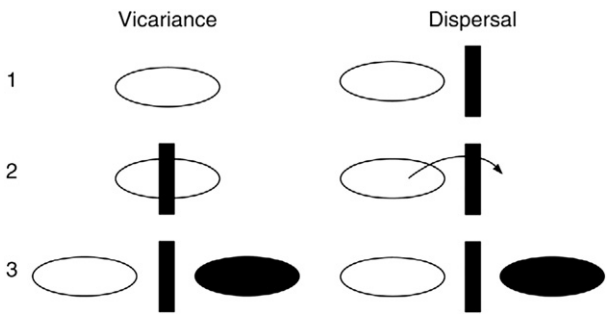


Figure 1 Vicariance and dispersal explanations: (a) vicariance; (b) dispersal. See text for explanation.

Nelson and Platnick (1984) compared Hooker's views with fellow Englishman Alfred Russell Wallace, another great biologist of the late nineteenth century. Wallace disagreed entirely with Hooker's hypothesis. In 1876, he said that the north and south division of biotas was not a vicariance pattern but represented the fact that

> "… the northern continents are the seat and birthplace of all higher forms of life, while the southern continents have derived for the greater part, if not the whole, from the north; … it implies the erroneous conclusion that chief southern lands—Australia and South America—are more closely related to each other than to the northern continent."

At that time the prevailing geological theory was one of fixed continents and slow, gradual changes in both Earth history and biogeographical developments. Hooker saw a close connection between cause and effect and said that the basic pattern of relationships divided the northern from the southern taxa, and the pattern was explained by vicariance events. Wallace, heavily influenced by Darwinian selection theory on the other hand, did not see a necessary connection between distribution and Earth history. For him, the basic pattern was east–west and the pattern was explained by dispersal from northern centers of origin.

Plate Tectonics

It has become clear that the nineteenth century was characterized by the dominance of British Empire and European colonialism throughout the world. It was clear that the "higher forms of [human] life" inhabiting the north were invading and dominating elsewhere, particularly the countries of the Southern Hemisphere (Nelson and Platnick, 1984). It was no accident that the "favored races" were of northern origin as indicated in the title of Darwin's (1859) famous book: *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggles for Life*. Although throughout his life his ideas changed, Wallace's earlier views on biogeography were part of Darwinian orthodoxy (Nelson and Platnick, 1984). Many groups of animals and plants were seen created in "centers of origin" in the Northern Hemisphere to then disperse and colonize other parts of the globe. Following the central dogma of Darwinism, the colonizing organisms were seen as the surviving fittest and superior competitors to all others that lay before them. The less fit southern forms

became thrust aside and eventually extinct. This rather fanciful scenario has been repeatedly described in textbooks, even within the 1990s, and is sometimes known as the "mono-boreal relict hypothesis" (Nelson and Platnick, 1984).

The origin of taxa from "centers of origin" and by dispersal to other areas is heavily entrenched in biogeographic literature, an idea that goes back to the Age of Reason in the eighteenth century. For example, Linnaeus believed that if the Earth was changing, with continents growing in size, and that the process had operated in the past

> "… the continent in the first ages of the world lay immersed under the sea, except a single island in the midst of this immense ocean; where all animals lived commodiously, and all vegetables were produced in the greatest luxuriance …" (translated from Linnaeus, 1781, p. 77; see Nelson and Platnick, 1980).

Linnaeus knew that different organisms had particular lifestyles and he suggested that the "center of origin" must have been a primordial island in the Tropics and "bore a very lofty mountain" inhabited at different altitudes by species with different ecological requirements. As new land emerged beyond this island and as the seas receded, the resident organisms migrated by various means to colonize the new parts with suitably similar ecological conditions (Nelson and Platnick, 1980). Embedded deeply within these comments are two important ideas: that species originate and disperse from a "center of origin" and that regularities in distribution are controlled by ecological conditions.

During the nineteenth century adherence to dispersal could be attributed to contemporary geological and geophysical theory. Such a belief allowed only stable continents separated by permanently wide oceans. Biogeographers, such as Hooker, who considered former land connections and favored vicariance explanations, were scorned as being out of line with the prevailing view (Nelson and Platnick, 1984). Even though the late nineteenth century and the first half of the twentieth century produced occasional theories that suggested continents were mobile and forever changing in shape and position through continental drift, mountain building, and erosion, dispersal theories still dominated the literature. It was not until the 1960s that geophysical studies eventually showed that continents and ocean bottoms were not permanent geographical features and that a general acceptance, one of the great paradigm shifts in science, came about. Early theories of continental drift were based on a model of the world that had a fixed diameter and that all of the present-day continents were joined up into one supercontinent, Pangea. During the Jurassic, some 180 million years ago, the supercontinent broke up into two major landmasses, the southern supercontinent, Gondwanaland, and the northern supercontinent, Laurasia. Laurasia contains those areas that have subsequently crystallized out into Europe, North America, and Asia, but exclude the Indian subcontinent. From the Cretaceous onward, the southern end of the world, Gondwanaland, broke up into those areas we now know as Antarctica, Australia, India, Africa, and South America.

The breakup of Gondwanaland lends credence to vicariance biogeography and hence Hooker's theory. The similarities among floras of the Southern Hemisphere suggest former land connections. It also revived the theory that Hooker's hypotheses of north–south connections exist for plants and contradicted Wallace's theory of east–west connections based on animals. As Nelson and Platnick (1984) asked, how can we determine which hypothesis is correct? Are both theories correct to some degree, or is there another hypothesis that should be erected which subsumes both into one explanation? The interesting point is that if the explanation is a general one, what process caused it and what methods are required to uncover the historical pattern? Paradoxically, through attempts to resolve the relationship between space, form, and time in terms of vicariance and dispersal, including application of track analysis in panbiogeography (Croizat, 1964; Craw *et al.*, 1999) and cladistics in vicariance biogeography (Nelson and Platnick, 1981), it was clear that a more general issue was at stake. This was the question of how to interpret patterns of plant and animal distributions in a systematic way to provide independent theories of biogeography, rather than rely on geophysical and geological theories as prior explanations for present-day distribution of biotas (see Humphries and Parenti, 1999). To give some flavor of the rather fitful debate this idea created, the stages of application of cladistic methods that grew out of vicariance biogeography are briefly described.

Cladistic Methods In Biogeography

The analysis of patterns in cladistic biogeography starts with finding the species or higher taxa that exist in the different biogeographic areas and to discover their relationships. Cladograms, or branching diagrams, are used to express hypotheses of relationship between organisms. An example is shown in **Figure 2a**, where among the four species, C and D are more closely related to each other than either are to B, and also species B, C, and D are more closely related to each other relative to A. The cladogram specifies two inclusive

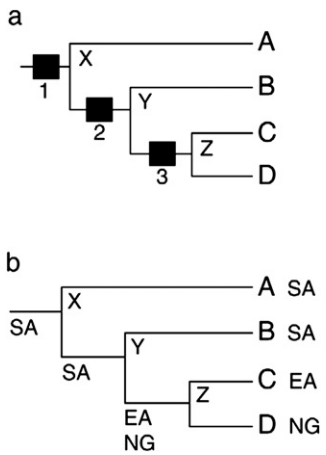


Figure 2 (a) Cladogram depicting the relationships between four species. A–D are four recent species, X–Z are hypothetical ancestors, and 1–3 are group-defining characters. (b) Area cladogram showing application of the progression rule to species A and B in South America (SA), C in eastern Australia (EA), and D in New Guinea (NG).

informative groups of species, BCD and CD. Cladograms are produced by the distribution of characters or features shared by the different species or taxa. Thus, in **Figure 2a**, character 3 specifies the group CD, character 2 specifies group BCD, and character 1 specifies ABCD as the overall group. Characters 2 and 3 then specify unique groups that have characters not shared by any other organism. If such features are taken to be present in the ancestors of each group, then the cladogram can be taken as specifying the pattern of evolution or common descent of that group. Seen in this way, the cladogram of **Figure 2a** may be taken to specify that taxa C and D have a unique ancestor Z that is not the ancestor of A or B and also that taxa B, C, and D have an ancestor Y that is not the ancestor of A. CD and BCD are monophyletic groups, compared, for example, with A and D, which is a polyphyletic group, and A and B, which comprise a paraphyletic group.

Cladistics was applied to biogeography in the 1960s at about the same time when theories of plate tectonics and shifting continents were becoming respectable. However, dispersal hypotheses to explain distribution patterns were still in vogue and thus the earliest applications in biogeography used cladograms to determine "centers of origin" of monophyletic groups. The German dipterist Willi Hennig (1966) reasoned that there is a recognizable, close relationship between a species and the space each occupies. He assumed that dispersal patterns are unique for each taxonomic group and that each had an independent history.

The Progression Rule

Central to Hennig's (1966) method to find a center of origin for a group of taxa from a particular cladogram was the idea that phylogenetically primitive members of a monophyletic group will be found near the center. That is, they are occupied by the species nearest to the basal nodes of cladograms. Within a continuous range of species of a monophyletic group, it was plausible that a series of characters transform and run parallel with distribution in space, such that the youngest members would be on the geographic periphery of the group.

To apply the "progression rule," one works backward considering the areas in the terminal positions of the cladogram and then gradually working down to the base of the cladogram, assigning distributions to the internal nodes or the common ancestors. If the areas occupied by the descendents do not overlap, they are combined together in assigning the probable distribution of the common ancestor. If they do overlap, the unshared element is eliminated. Consider the four species in **Figure 2b**. Species A and B occur in South America, C occurs in eastern Australia, and species D occurs in New Guinea. Working from the most derived species, C and D, the distributions clearly do not overlap; hence they are added together to give the area for ancestor Z as eastern Australia plus New Guinea. For ancestor Y the descendent taxa do not overlap, with B occurring in South America and ancestor Z in eastern Australia and New Guinea. The two distributions do not overlap so the area for ancestor Y is South America, eastern Australia, and New Guinea. Moving down the cladogram to ancestor X, descendent species A occurs in South America and ancestor Y in South America, eastern Australia, and New Guinea. The two distributions thus overlap in South America, and eastern Australia and New Guinea are eliminated as the unshared area. Working back up the cladogram, the simplest hypothesis is to consider that South America is common to all three ancestral species, X, Y, and Z, and that species C and D dispersed to eastern Australia and New Guinea from South America. The conclusion from analysis of this biogeographic pattern is that the center of origin for the group is South America.

To demonstrate this early use of cladograms in biogeography, a simple example can be seen in the caddis fly distribution of *Wormaldia* (Ross, 1974; see Humphries and Parenti, 1999). *Wormaldia* and other caddis flies have been a lifelong study of Ross because they occur in almost all freshwater systems throughout the globe, show a high degree of endemicity in all areas, and are very speciose. Ross considered the geographical distribution of the nine species of the *Wormaldia kisoensis* complex of caddis flies. Eight species are present in the western Pacific from northwestern Borneo in the south and Japan in the north and one species occurs in the Smoky Mountains of eastern North America (**Figure 3**). The cladogram of phylogenetic relationships is shown diagrammatically in **Figure 4**. Like Hennig, Ross (1974) assumed that the species nearest to the base of the stem of the cladogram denotes the ancestor of the group. He applied also the progression rule to the analysis to determine the center of origin of the group. In *Wormaldia* the most derived species pair occurs in Japan and eastern North America, which led Ross to propose the dispersal hypothesis of origin of the group in the western Pacific and a single dispersal of one species across the Bering straits to North America (**Figures 3** and **4**).

Brundin's classic studies of chironomid midges (Brundin, 1966) showed that the southern temperate areas of South America, Southern Africa, Tasmania, southeast Australia, and New Zealand are inhabited by 600–700 species. Trans-Antarctic relationships are a recurring phenomenon throughout the group, so by way of an example consider the midges of subfamily Diamesinae, which show a double distribution: two major groups present in both Northern and Southern Hemisphere temperate areas but absent from the Tropics. The largest and most widespread is the relatively generalized tribe Heptagymi, represented by eleven species in Andean South America, two species in southeastern Australia, and five species in New Zealand (**Figure 5**). Its sister group is the relatively more apomorphic and monotypic tribe Lobodiamesini of New Zealand.

The cladogram in **Figure 5** shows that there are a total of 25 terminal taxa in the Southern Hemisphere areas of South America, New Zealand, Australia, and South Africa and three groups in Laurasia. The monotypic genus *Heptagyia* occurs in South America and *Paraheptgyia* has five South American species. According to Brundin, the southeastern Australian subgroup of two species is a younger evolved offshoot of the older South American group including *Heptagyia*. Brundin considered the Australian taxa to have dispersed from Patagonia or east Antarctica at stem 6 by the end of the Paleocene because they have derived characters and because the stem species (indicated by 1, 2, and 4) never occurred in Australia. The other stem (2a) includes the genera *Reissia*, with three

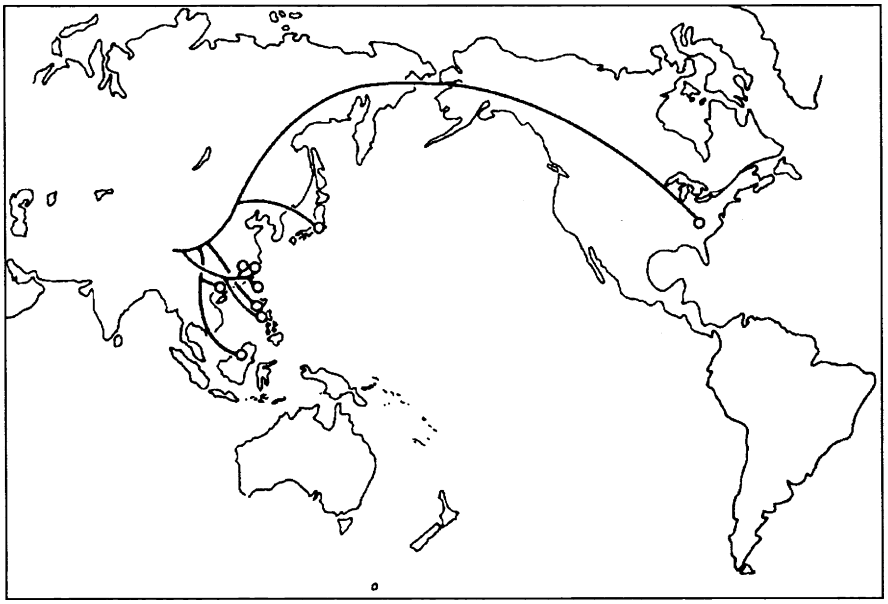


Figure 3 Distribution of the *Wormaldia kisoensis* group of caddis flies. The circle in Japan is *Wormaldia kisoensis*, and that in North America is *Wormaldia mohri*. (Redrawn from Ross HH (1974) *Biological Systematics*. Reading, PA: Addison-Wesley, with permission from CUP.)

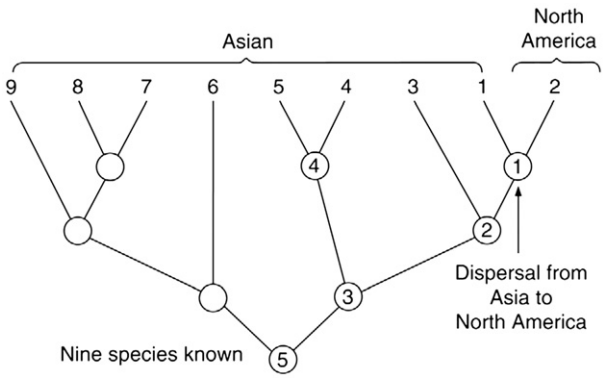


Figure 4 Cladogram for the *Wormaldia kisoensis* group of caddis flies. (Adapted from Ross HH (1974) *Biological Systematics*. Reading, PA: Addison-Wesley, with permission from CUP.)

species in South America, *Limaya*, with two species in South America, and *Maoridiamesa*, with five species in New Zealand. Brundin (1981) considered that *Maoridiamesa*, on a different stem from the Australian *Paraheptagyia* group, agreed well with plate tectonic theory for an early separation of New Zealand from western Antarctica in the Upper Cretaceous. The fact that *Maoridiamesa* is a comparatively younger, derived offshoot of an older group in South America was, according to Brundin, evidence of long-distance dispersal from South America via west Antarctica to New Zealand of stem species 4a rather than a vicariance event. In other words, Brundin thinks that the *Maoridiamesa* group is younger than the areas in which it occurs.

The progression rule can be interpreted as one particular optimization procedure applied to cladograms. It has problems in the sense that it can specify a different center of origin when applied to different data. Changes in data occur for a host of different reasons, but largely through the discovery of known taxa in new places, the discovery of new taxa, and reinterpretation of relationships. Patterson (1981) showed that the progression rule was very sensitive to the discovery of new taxa, and particularly fossils. The discovery of fossils simply added more areas to the problem and thus increased the number of possibilities for finding new centers of origin. Thus, for example, if we added two new fossils, V and W, to the example described earlier (see **Figures 2** and **6**), then the possibility of a center of origin in North America and in turn Europe would alter the interpretation significantly. Patterson considered that for groups of organisms with a significant fossil record, the progression rule is heavily influenced by fossils and this would be especially true in the areas where the best fossils occurred. The progression rule is dependent on the notion that organisms occur in a center of origin. The problem is that when the progression rule is applied to a cladogram, it can always specify a center whether it may, or may not, have occurred. Suppose, for the sake of argument, the species of **Figure 6b** in Europe, North America, South America, eastern Australia, and New Guinea were caused by the breakup of continents, a vicariance explanation; then the progression rule would provide a spurious hypothesis impossible to test.

Vicariance Biogeography

The problem of interpreting cladograms as dispersal scenarios is the same as treating them as phylogenetic trees rather than as hierarchical classifications to show groups with relatively more inclusive groups of set membership. Such a procedure requires making ad hoc assumptions not based on the information on which they are based. Furthermore, interpreting individual cladograms as having individual

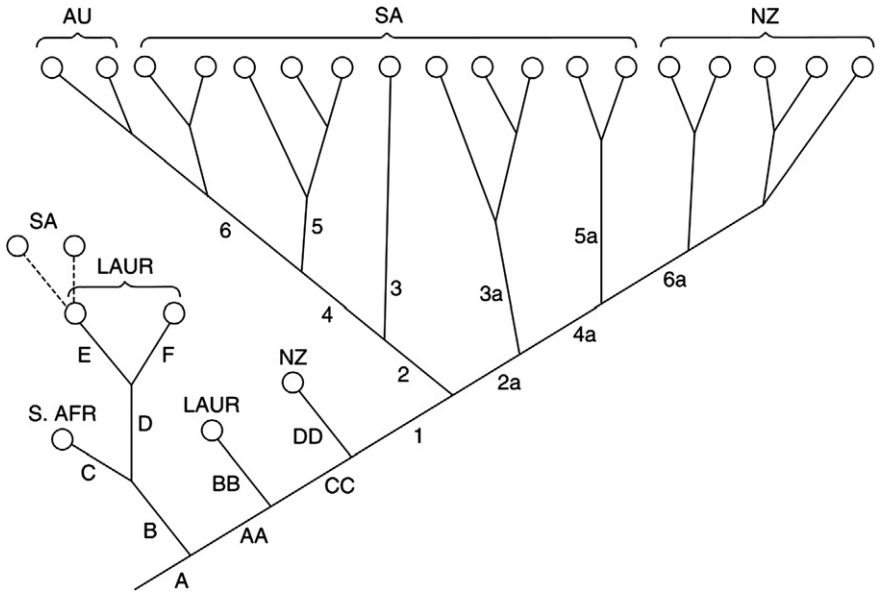


Figure 5 Partial area cladogram of the subfamily Diamesinae (Diptera; Chironomidae). Some named lineages are as follows: 1, Heptagyini; 3, *Heptagyia;* 4, *Paraheptagyia;* 3a, *Reissia;* 5a, *Limaya;* 6a, *Maoridiamesa;* A, Diamesinae; C, Harrisonini; E, Diamesini; F, Protanypodini; BB, Boreoheptagyini; DD, Lobodiamesini. (Reproduced with permission from Figure 3.7 in Brundin L (1981) Croizat's panbiogeography versus phylogenetic biogeography. In: Nelson G and Rosen DE (eds.) *Vicariance Biogeography: A Critique*, pp. 94–158. New York: Columbia Univ. Press.)

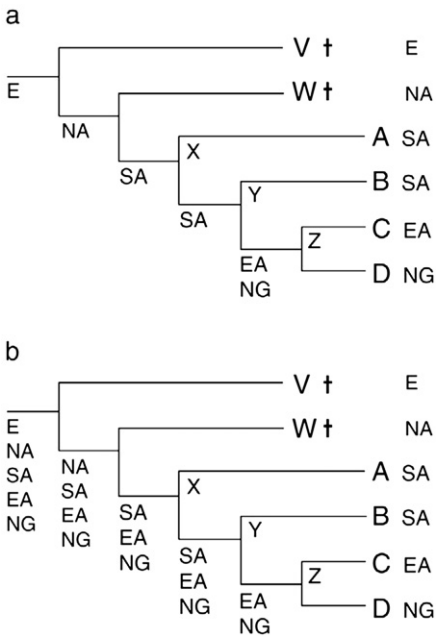


Figure 6 (a) Fossils and centers of origin. A–D represent four Recent species, A and B in South America (SA), and C and D in eastern Australia (EA) and New Guinea (NG), respectively. V and W are fossils, as indicated by the crosses distributed in North America (NA) and Europe (E), respectively. Areas on the internal nodes are ancestral areas assigned using the progression rule. (b). The same cladogram but applying a vicariance hypothesis. See text for explanation.

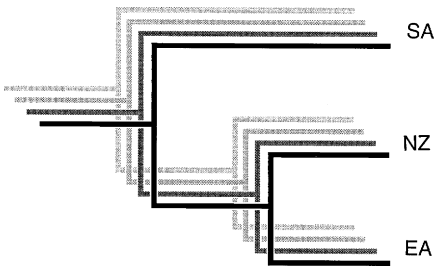


Figure 7 Four hypothetical groups of organisms showing the same area relationships.

histories leads to certain conceptual difficulties. A crucial one is how do we interpret repetitious distribution patterns? If we have many unrelated taxonomic groups repeating a pattern of distribution between major continents, such as South America, eastern Australia, and New Zealand (**Figure 7**), it is improbable that the pattern was caused by separate unique dispersal events. This implies that species in each different group separately makes its own way from one continent to the other. A logical and more robust conclusion would be to suggest that at one time, the continents were in contact and that the present-day pattern was due to the breakup of a formerly continuous biota.

The breakthrough in application of cladistic reasoning to biogeography came with the realization that disjunct distributions could occur because of such vicariance events. This is based on the idea that present-day endemic taxa occur in the same areas as their ancestors and the Recent taxa evolved *in situ*. Where dispersal models explain disjunctions by dispersal across pre-existing barriers, vicariance models explain them by the appearance of barriers fragmenting ancestral species ranges (**Figure 1**). To historical biogeographers what became particularly clear was the important idea that distribution data are insufficient to resolve decisively either dispersal or vicariance as the cause of disjunct distribution patterns.

Platnick and Nelson (1978) argued that one should not worry about the cause of a disjunct distribution pattern between different related areas of endemism but whether or not it conforms to a general pattern of relationships shown by other groups of taxa endemic to the areas occupied. Thus, there is an analogy to the three-taxon statements in systematics as the most basic units for expressing relationships. In biogeography three-area statements are the most basic units for expressing relationships between areas of endemism. The generality of the area cladograms can be examined by comparison with other unrelated taxonomic groups endemic to the relevant areas and corroboration of a particular pattern is equivalent to a general statement for the relative recent ancestry of the biotas under scrutiny. However, there has been a long, drawn out saga of trying to find ways for applying the principles of cladistic biogeography because initial applications of the method (e.g., Rosen, 1976) encountered problems of incongruence between different groups of organisms. Theoretically, it should be possible to connect every area of endemism of the world into one huge general statement of interrelationships. However, our perception of the world is less than perfect for a variety of reasons—extinction, dispersal of widespread taxa, and restricted distributions of taxonomic groups. The significance of these problems can be demonstrated with poeciliid fish in Middle America.

Poeciliid Fish In Middle America

To discover relationships of areas, at least two groups of taxa must be available for the same or similar set of areas. Rosen (1978, 1979) examined two groups of fishes, *Heterandria* and *Xiphophorus* from the Mesoamerican region. Both genera have close relatives elsewhere but each has a monophyletic subgroup inhabiting 11 areas in southern tropical Mexico, south to eastern Honduras in *Xiphophorus*, and further south to eastern Nicaragua in *Heterandria* (**Figure 8**). Area cladograms for the two genera, which indicate areas occupied by particular species, are shown in **Figures 9** and **10**.

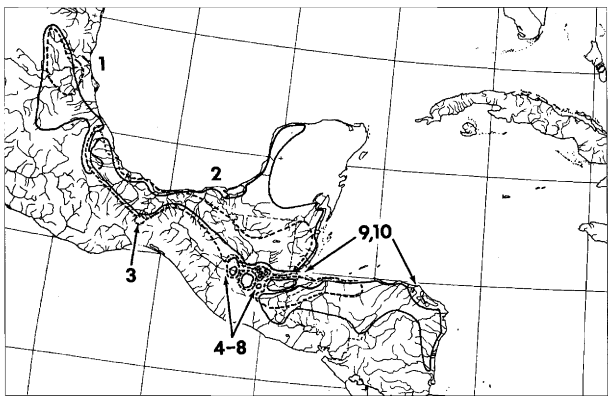


Figure 8 Map to compare the distributions of the species of *Heterandria* (full line) and *Xiphophorus* (broken line). Numbers refer to areas defined by the occurrence of taxa. (Reproduced from Figure 45 in Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* 162: 267–376, with permission from Society of Systematic Biology.)

Component Analysis

Cladistic biogeography would be uncomplicated if all groups of organisms were each represented by one taxon in each of the smallest identifiable areas of endemism. This is not the case. Unique patterns may be meaningful and cannot at the same time be incongruent with patterns determined from

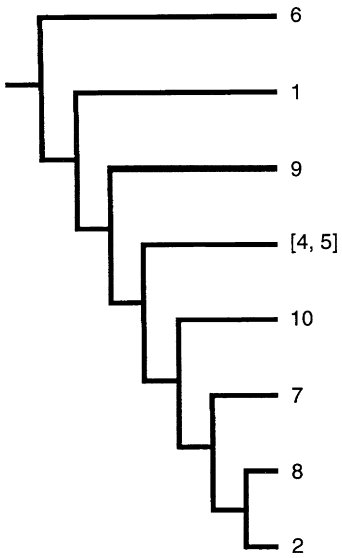


Figure 9 Area cladogram for *Heterandria*. The numbers refer to areas denned by the species as shown on the map in **Figure 8**. (Redrawn from Figure 48 in Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* 162: 267–376, with permission from Society of Systematic Biology.)

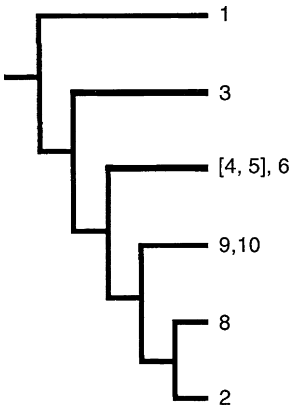


Figure 10 Area cladogram for *Xiphophorus*. The numbers refer to areas defined by the species as shown on the map in **Figure 8**. (Redrawn from Figure 49 in Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* 162: 267–376, with permission from Society of Systematic Biology.)

other taxa. Nevertheless, incongruence between two or more cladograms can occur for a variety of historical reasons. Particular groups of organisms can exhibit older or younger patterns than the groups to which they are being compared. Also, some taxa have dispersed and become widespread and by so doing have obscured the historical signal in the area patterns. Comparing several different groups of organisms is redundant in the sense that the same pattern is repeated over and over or areas are represented by two or more taxa. With extinction the observed pattern creates a spurious historical signal. All of these problems effectively cause sampling errors and lead to wrong predictions of the general patterns of area interrelationship. Nelson and Platnick (1981) used cladistic logic and component analysis to make comparisons between area cladograms to yield the maximum resolution in the general-area cladograms (see **Box 1**).

Heterandria and *Xiphophorus*

To illustrate just how the methods to extract the best possible patterns of relationship work, consider the two groups of poeciliid fish genera *Heterandria* and *Xiphophorus* (**Figures 8–10**) based on the work of the American ichthyologist Donn Rosen (1978, 1979). The 11 identifiable areas of endemism, each determined by substituting species of *Heterandria* or *Xiphophorus* for the areas in which each occurs, include areas 1, 2, 4, 5, 8, 9, and 10 common to both groups. Areas 4 and 5 are occupied by one species from each group and thus treated as one area, (4–5). An area 11 was described by Rosen as a putative hybrid area between areas 4–5 and 2, but to simplify matters, it is excluded here. A comparison of the two cladograms shows that *Xiphophorus* is less informative than *Heterandria* because it has two widespread species in areas (4–5), 6, 9, and 10 and is absent from area 7. In *Heterandria* areas (4–5), 6, 9, and 10 are all occupied by recognizable endemic species. However, the pattern is not quite complete in that there is no species present in area 3.

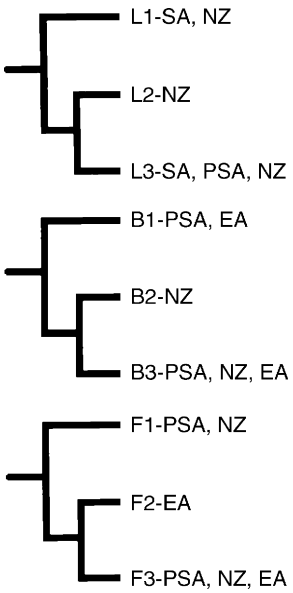


Figure 11 Three cladograms for three groups of lizards (L1–L3), birds (B1–B3), and fishes (F1–F3). The areas in which species occur include South Africa (SA), New Zealand (NZ), Pacific South America (PSA), and eastern Australia (EA). L2, B2, and F2 are endemics to NZ and EA, respectively. L1, B1, and F1 are widespread in two areas and L3, B3, and F3 are each widespread in three of the four areas. Lizards are not present in EA, and birds and fishes are not present in SA.

Box 1 Component analysis

To determine the degree of congruence between different groups of organisms occupying similar areas, the problems of "missing areas," widespread taxa, and redundancy are taken into consideration using component analysis and consensus.

Components are the elements of a group of areas, or group of taxa, as determined by the branching pattern of a cladogram. For example, in a group comprising three taxa (or areas) A, B, and C, when B is more closely related to C, there are two components, a general uninformative ABC component and an informative BC component.

Components of cladograms in vicariance biogeography can be manipulated to extract historical signal from even the most recalcitrant cases (**Figure 11**). For example, consider the four areas of endemism, South Africa (SA), Pacific South America (PSA), New Zealand (NZ), and eastern Australia (EA). Now consider three groups of organisms distributed in these four areas; three species of lizards (L1–L3) in SA, PSA, and NZ, and three species each of birds (B1–B3) and fish (F1–F3) in PSA, NZ, and EA. The three numbered informative nodes (1–3) on the cladograms express the relationships for each species in the lizards, birds, and fishes. Notice that each group of organisms occurs in three areas, and in each case there is one area not present (EA for the lizards and SA for the birds and fish). Also, for each group there is one species endemic to an area, i.e., L2 in PSA, B2 in NZ, and F2 in EA. The remaining species are more widespread, occurring in either two or three areas (respectively L1, B1, and F1 in two areas and L3, B3, and F3 in three areas).

Because of widespread species and the lack of representation of one area in each of the area cladograms, at first sight relationships between the four areas seem difficult to resolve (**Figure 11**). However, it can be assumed that each species originated in one area with extinction, dispersal, and failures to vicariate, causing the more complex patterns. Taking this view, we can examine the informative three-area relationships in the area cladograms to see if they tell us anything about the general-area relationships. Thus, in the lizard cladogram, L1 in SA, L2 in NZ, and L3 in PSA is the only informative statement that can be represented as SA(PSA, NZ) and scored as 0(1, 1) in a matrix (**Table 1**, column 1). The "missing" area (EA) is scored as a dash. Informative area statements for birds are PSA(NZ, EA) and EA(NZ, PSA), and for fishes PSA(EA, NZ) and NZ(EA, PSA) similarly scored in the matrix (**Table 1**, columns 2–5). Consequently, for the three cladograms there are five informative three-area statements (**Table 1**). By using a standard parsimony analysis, the best fitting tree for the five statements is shown in **Figure 12**. Thus, despite the complexity of the patterns in the original three groups (**Figure 11**), if there is one history that best explains the data, then it might be a vicariance hypothesis for the four areas as offered in **Figure 12**—(SA, PSA, NZ, EA) (SA(PSA, NZ, EA)) (SA(PSA(NZ, EA)))—and shown as an area cladogram in **Figure 13**, indicating the relationships of the four areas of endemism.

By using the assumption that one sequence of historical events can explain the patterns observed in both *Heterandria* and *Xiphophorus* and then using the logic applied for the hypothetical example used for the breakup of Gondwana (Box 1), a single result can be obtained. Platnick (1981) noted that if widespread taxa are uninformative, they cannot at the same time be incongruent. We can assume that the information on areas 6 and 9 in *Heterandria* is correct. Thus the incongruent information in the same areas for *Xiphophorus* (areas 6 or 4–5, but not both, and 9 or 10, but not both) is due either to dispersal or a failure to speciate in response to vicariance events that gave the pattern in *Heterandria*. Taken on their own, widespread taxa are uninformative. However, when considered with other cladograms, informative results are possible. Similarly, "absence" data (e.g., area 3 in *Heterandria* and area 7 in *Xiphophorus*) can never be incongruent with information at hand so unique areas should be placed in terms of whatever pattern exists (i.e., area 7 by *Heterandria* and area 3 by *Xiphophorus*). By using these assumptions, the *Xiphophorus* cladogram allows the populations in areas 9 and 6 to occur in any of 12 positions and the *Heterandria* cladogram allows area

Table 1 Informative components (three-item statements) of the three cladograms in **Figure 12** (scored as 0(1,1))

Area	*Components expressed as three-item statements*				
	1	*2*	*3*	*4*	*5*
Root	0	0	0	0	0
South Africa (SA)	0	–	–	–	–
New Zealand (NZ)	1	1	1	1	0
Pacific South America (PSA)	1	0	1	0	1
Eastern Australia (EA)	–	1	0	1	1

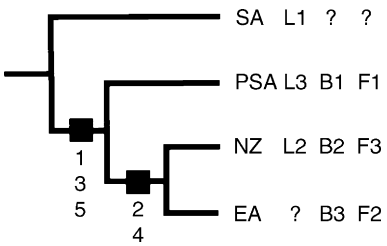


Figure 12 General-area cladogram for lizards (species of lizards (L1–L3), birds (B1–B3), and fishes (F1–F3). The acronyms for areas are as in **Figure 12**. The five informative components (1–5) identify area relationships.

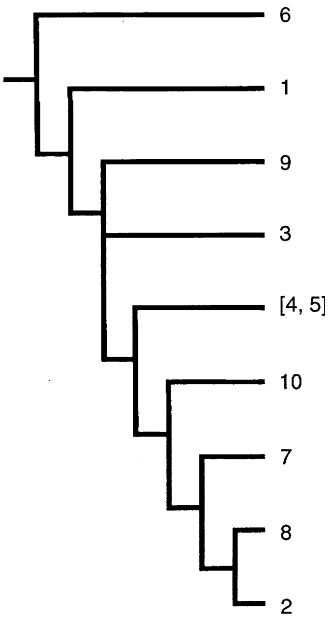


Figure 14 Consensus tree of the three possible intersecting cladograms of *Heterandria* and *Xiphophorus*. See text for explanation.

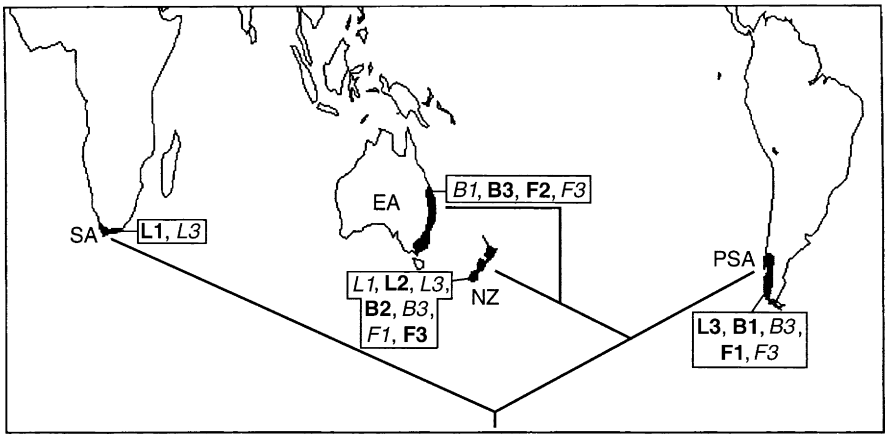


Figure 13 Sequence of vicariance events (1–4) for some areas of Gondwanaland hypothesized from the general-area cladogram in **Figure 12**. The area acronyms are as for **Figure 11**. This is just one scenario out of several where an ancestral area gradually divides into four separate areas by successive vicariance events. The taxa are given the same abbreviations as in **Figure 11** and those providing the relational signals are written in bold and roman type.

3 to float onto any of the branches of the cladogram. Thus, of the 425 allowable trees for *Xiphophorus* and the 27 for *Heterandria*, the analysis yields three possible intersections. These are summarized as a consensus tree (**Figure 14**).

By using this method, we have an area hypothesis that accounts for all 10 areas of endemism (i.e., excluding the hybrid area 11) that can be recognized from the two genera of fishes *Heterandria* and *Xiphophorus*. If the pattern has been created due to changes in Earth history, the question that we could ask now is: "What might have been the historical factors in Mesoamerica to cause this pattern and how might these be compared with the given biological distribution?" Ideally, we would require that geological information be assembled into cladograms, in the same way as biological cladograms, so that biotic and historical patterns can be compared. Until such time as geological data can be ordered for a more informative comparison, one thing we can say is that the observed patterns in species relationships of *Heterandria* and *Xiphophorus* in Mesoamerica have been formed over a period of at least the past 80 million years and are most likely to have been brought about by vicariance events where life and Earth evolved together (Humphries and Parenti, 1999; Rosen, 1978).

Conclusions

At this time, methods of vicariance or cladistic biogeography have proved useful to analyze and compare biotic patterns at the highest resolution and organized in such a way as to compare them to independent sources of data such as geological patterns. Cladistics is a general method of determining class and subclass relations whatever the source of data without recourse to evolutionary narrative. A cladistic approach to Earth history makes it possible to express biogeographic interrelationships as hierarchical relations from biotic information. The development of cladistic methods that cater for the possibilities to consider complicating events but without ever using such information prior to analysis provides a system that allows the possibility of generating general biogeographic hypotheses even from seemingly ambiguous patterns. In other words, vicariance biogeography is a general empirical procedure without any prior assumptions in the analysis about dispersal, vicariance, redundancy, or extinction events but which at the same time never denies that they occur.

See also: Biogeography, Overview. Cladistics. Darwin, Charles (Darwinism). Diversity, Community/Regional Level. Endemism. Island Biogeography

References

Brundin L (1966) Transantarctic relationships and their significance as evidenced by midges. *K. Sven. Vetenskapsakad. Handl. (Ser. 4)* 11: 1–472.

Brundin L (1981) Croizat's panbiogeography versus phylogenetic biogeography. In: Nelson G and Rosen DE (eds.) *Vicariance Biogeography: A Critique,* pp. 94–158. New York: Columbia Univ. Press.

Craw RC, Grehan JR, and Heads MJ (1999) *Panbiogeography: Tracking the History of Life*. Oxford: Oxford Univ. Press (Oxford Biogeography Series No. 11).

Croizat L (1964) *Space, Time, Form, the Biological Synthesis*. Caracas: Published by the author.

Hennig W (1966) *Phylogenetic Systematics*. Urbana, IL: Univ. of Illinois Press.

Humphries CJ and Parenti LR (1999) *Cladistic BiogeographyAnalysing the Patterns of Animal and Plant Distributions*, 2nd ed. (Oxford Biogeography Series No. 12). Oxford: Clarendon.

Nelson G and Platnick NI (1980) A vicariance approach to historical biogeography. *Bioscience* 30: 339–343.

Nelson G and Platnick NI (1981) *Systematics and Biogeography: Cladistics and Vicariance*. New York: Columbia Univ. Press.

Nelson G and Platnick NI (1984) *Biogeography*. Burlington, NC: Carolina Biological Supply Co. (Carolina Biology Readers No.119).

Patterson C (1981) Methods of palaeobiogeography. In: Nelson G and Rosen DE (eds.) *Vicariance Biogeography: A Critique*, pp. 446–489. New York: Columbia Univ. Press.

Platnick NI (1981) Widespread taxa and biogeographic congruence. In: Funk VA and Brooks DR (eds.) *Advances in Cladistics: Proceedings of the First Meeting of the Willi Hennig Society*, pp. 223–227. Bronx, NY: The New York Botanical Garden.

Rosen DE (1976) A vicariance model of Caribbean biogeography. *Syst. Zool.* 24: 431–464.

Rosen DE (1978) Vicariant patterns and historical explanation in biogeography. *Syst. Zool.* 27: 159–188.

Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* 162: 267–376.

Ross HH (1974) *Biological Systematics*. Reading, PA: Addison-Wesley.

Water Funds: A New Ecosystem Service and Biodiversity Conservation Strategy

Rebecca L Goldman-Benner, The Nature Conservancy, Durham, NC, USA
Silvia Benitez, The Nature Conservancy, Quito, Ecuador
Alejandro Calvache and Aurelio Ramos, The Nature Conservancy, Cartagena de Indias, Colombia
Fernando Veiga, The Nature Conservancy, Curitiba, Brazil

© 2013 Elsevier Inc. All rights reserved.

Glossary

$ All $ are US dollars unless otherwise stated.
Biodiversity The diversity and variety of life on earth.
Direct payments An exchange of money where a consumer of a service pays the provider for provision of that service.
Ecosystem services The benefits that ecosystems can provide to people.
Forest enrichment Planting trees in degraded or damaged forests to help recover forest density and/or to improve biodiversity by planting species otherwise unable to colonize and regenerate or that are threatened or vulnerable.
Global environment facility (GEF) An independent financial organization providing grants to developing countries and countries with economies in transition for projects related to biodiversity, climate change, international waters, land degradation, the ozone layer, and persistent organic pollutants.
Indirect payments An in-kind payment in the form of materials, labor, or time rather than payments of money.
Livelihood investments A nonmonetary payment to compensate for opportunity costs of a change in land use can include alternative livelihoods, alternative food sources, and education investments, among others.
Payments for ecosystem services (PES) Defined in purely economic terms as a voluntary transaction where a well-defined ecosystem service(or a land-use likely to secure that service) is 'bought' by a (minimum one) ecosystem service buyer from a (minimum one) ecosystem service provider if and only if the ecosystem service provider secures ecosystem services provision (conditionality).
Payments for watershed services A subset of PES that involves a similar financial transaction but the focus is watershed-based hydrologic services.
Revegetation Replanting grasses, bushes, and other types of vegetation on cleared and bare soils.

Introduction

Water funds are an innovative ecosystem services-based approach to conservation that use a sustainable finance mechanism created through investments by water users and other stakeholders to ensure the long-term provision of critical hydrologic services from a watershed. Investments are essentially paid into an endowment fund. Fund contributions are used to improve management of small-scale farms and ranches in the watershed and to strengthen the conservation of protected areas. Through such investments, water funds ensure the long-term protection of critical terrestrial and freshwater systems while providing a clean, reliable source of water – the key hydrologic services for the downstream water users. Tailored to local conditions, these funds are rapidly replicating throughout Latin America, most prominently in Ecuador, Colombia, and Brazil (**Figure 1**).

Biodiversity in Peril

Latin America and the Caribbean (LAC) are among the richest regions in biodiversity worldwide (FAO, 2011). South America alone accounts for half of the terrestrial biodiversity in the world and contains the world's most biodiversity-rich region: the eastern slope of the Andes. LAC includes some of the world's most biodiverse countries, for example, Brazil, Colombia, Ecuador, Mexico, Peru, and Venezuela (Mittermeier *et al.*, 1997; Bovarnick *et al.*, 2010), and contains one quarter of the 33 global biodiversity hotspots – regions of high species endemism and high rates of habitat loss – in only 15% of the Earth's land area. Additionally, the region contains close to 800 million hectares of forested areas (more than 35% of the world's forests), 570 million hectares of wild savannas, 700 million hectares of agricultural lands, and 27% of the planet's available drinking water. From timber and nontimber forest

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00330-0

products to water and nutrient regulatory services, biodiversity is of strategic importance for conservation and development in LAC.

Globally, human impacts are intensifying across Earth's lands and waters, leading to unprecedented biodiversity loss. In all, 33–50% of the Earth's land surface has been altered by the human action (Vitousek *et al.*, 1997) causing huge biodiversity losses, and declines in biodiversity are greater in fresh waters than in the most affected terrestrial ecosystems (Sala *et al.*, 2000). As a result some 10–30% of all mammal, bird, and amphibian species are threatened with extinction (Levin and Levin, 2004; Kiesecker *et al.*, 2004). Despite occupying only a tiny percentage of the planet's surface (0.8%), on a hectare-to-hectare basis, freshwater ecosystems are richer in species than the more extensive terrestrial and marine ecosystems (Revenga and Mock, 1999), and yet, freshwater habitats and species are proportionally more severely degraded and threatened than ecosystems on land or in the ocean.

Of protected areas worldwide, 20% are located in LAC (ECLAC, 2010) and are under threat due to poor governance, small budgets, and weak enforcement of legislation. One major threat is land conversion for agriculture and livestock (Barbier, 2004; CEPAL, 2005) as LAC, especially Central America and the Andes-Amazon region, has long been one of the cradles of global agriculture (Diamond, 2002; Purugganan and Fuller, 2009). Overexploitation also causes huge losses in marine and aquatic systems due to overfishing, more fleets, and permissive government incentives. Another growing cause of biodiversity loss that has not been thoroughly researched or noted is the spread of invasive species largely as a result of trade but also because of human travels via airplane or road. The roads themselves are highways for invasive species. Other threats to biodiversity include development and tourism. Direct threats to freshwater systems include dams, water abstraction, flow modification, overexploitation, overharvesting, pollution, deforestation, and invasive species (MA, 2005; Dudgeon *et al.*, 2006). Additionally, water quality and water flows are impacted by surrounding land uses.

Population growth and climate change further threaten LAC biodiversity. Population growth will require greater investments in resource development (IEA, 2007), but there is increased pressure to find ways to balance development needs with those of biodiversity conservation (Kiesecker *et al.*, 2009). Climate change could cause major ecosystems to shift and species to attempt to migrate, and indirect impacts such as storm intensity and frequency can provide further opportunities for previously nonthreatening exotic species to become invasive species. Additionally, climate change threatens to alter seasonality and annual precipitation patterns, further exacerbating threats to water resources (Vörösmarty *et al.*, 2000). For example, about 50 million people in the Andes could lack dry-season water for drinking, irrigation, sanitation, and hydropower. These threats require attention.

Freshwater systems are intimately tied to surrounding terrestrial systems, so conserving one requires conserving the other. Despite numerous efforts to protect watersheds or establish drinking water projects, few programs address the link with protected areas that were often created to protect water sources (McAlpine and Wottom, 2009; Nel *et al.*, 2009). While evidence suggests that it is more cost effective to protect than mitigate, the costs of watershed management have been almost universally neglected in water pricing. Worse still, these costs have not been evaluated against operational costs for water treatment or investment costs for new infrastructure. Recent evidence of shrinking clean water supplies and perceived water insecurity have made businesses and water utilities look at freshwater as they never have before: a valuable good that is produced, sold, and consumed and deserves investment. Increasingly, this investment is in the form of payment for ecosystem services (PES) projects to conserve freshwater systems.

Ecosystem Services and Payment for Ecosystem Services: Rapidly Expanding Conservation Approaches

In the latter half of the 1990s, researchers began using ecosystem service values to demonstrate the importance of natural capital (e.g., Costanza *et al.*, 1997), and with the publication of the Millennium Ecosystem Assessment (MA, 2005), the importance of biodiversity to secure livelihoods became increasingly evident, particularly through the delivery of ecosystem services – the benefits that nature can provide to people (Daily, 1997). In the past decade, more studies have looked at ecosystem service values and at the biodiversity values either directly associated with ecosystem services such as the economic benefits to pollination from conservation (e.g., Ricketts *et al.*, 2004) or the indirect, all-inclusive values such as the economic benefits and costs avoided by not creating protected areas (e.g., Bezaury-Creel, 2009). These have led to an increasing understanding of the relationship between biodiversity conservation, ecosystem services, and human well-being.

Inevitably, ecosystem service approaches to biodiversity conservation will require trade-offs that might imply that the conservation of one service is done at the expense (degradation) of another (e.g., Heal *et al.*, 2001; Pereira *et al.*, 2005; Rodríguez *et al.*, 2005, 2006). Of particular interest is the possibility that the recent emphasis on ecosystem services could actually detract from biodiversity conservation. For example, if projects are designed and managed with special attention to how people benefit, will that take attention away from biodiversity (Kareiva *et al.*, 2008)? Naidoo *et al.* (2008) found that locations selected for conserving ecosystem services would conserve only 22–35% as many species as locations selected for preserving biodiversity. What's more, only 16% of World Bank biodiversity-focused development projects resulted in a win–win for biodiversity and human well-being (Tallis *et al.*, 2008). That said, optimizing an ecosystem for one or a few services can prove beneficial provided the ecosystem can function as a whole (e.g., Tallis and Kareiva, 2006). As such, ecosystem services are one of many biodiversity conservation approaches.

Despite such trade-offs, ecosystem service approaches to conservation confer advantages, not the least of which is providing opportunities for conservation in new areas. Thinking in an ecosystem services framework can broaden the scope for conservation by expanding where we view conservation opportunities and what those opportunities are (Tallis *et al.*, 2009;


Mature and strengthened water funds
Water funds created and operating
Water funds in design
Evaluate potential to establish a water fund
United States
Mexico
Cuba
Jamaica
Haiti
Dominican Republic
Puerto Rico
Guadeloupe
Martinique
Grenada
Trinidad and Tobago
Belize
Guatemala
Honduras
El Salvador
Nicaragua
Costa Rica
Panama
Venezuela
Colombia
Guyana
Suriname
French Guiana
Ecuador
Peru
Bolivia
Brazil
Paraguay
Uruguay
Argentina
Chile
AMAZON BASIN
GUIANA HIGHLANDS
BRAZILIAN HIGHLANDS
ANDES MOUNTAINS
PAMPAS
PATAGONIA
Gulf of Mexico
Caribbean Sea
Atlantic Ocean

Goldman *et al.*, 2008). Additionally, that approach can diversify stakeholder interest in and funding streams for conservation (Goldman *et al.*, 2008), so it is one of many useful approaches for conservation. While there is still uncertainty about what factors are likely to contribute to successful ecosystem service projects (Perrot-Maître, 2006; Asquith and Wunder, 2008; Engel *et al.*, 2008; Jack *et al.*, 2008; Daily and Matson, 2008) and debates about their use continue (e.g., Goldman and Tallis, 2009), these approaches continue to proliferate.

Ecosystem service projects can take a wide variety of forms and serve a variety of functions (Tallis *et al.*, 2009), but in applied conservation, PES projects are widely touted as an innovative finance mechanism (Salzman, 2005; Wunder, 2007; Jack *et al.*, 2008). In the last decade, PES schemes have received considerable attention (e.g., Landell-Mills and Porras, 2002; Pagiola *et al.*, 2007; Jack *et al.*, 2008; Sommerville *et al.*, 2009) with their focus on providing financial incentives to land owners or managers whose conservation actions benefit others when those benefits would otherwise not be compensated (Sommerville *et al.*, 2009).

Payment for watershed services (PWS) projects, a subset of PES approaches, are a significant portion of ecosystem services schemes (many others relate to carbon) and often involve water users paying "suppliers" for the delivery of clean, consistent water supplies (Brauman *et al.*, 2007; Krchnak, 2007; Asquith and Wunder, 2008; Porras *et al.*, 2008). Numerous case studies of these approaches exist in the developed (e.g., Perrot-Maître, 2006) and developing (e.g., Stanton *et al.*, 2010) world. In 2002, there were only 41 proposed and ongoing PWS schemes in developing countries (Landell-Mills and Porras, 2002), but the number more than doubled in 6 years as a recent review documented 50 ongoing, eight advanced proposals, and 37 preliminary proposals (Porras *et al.*, 2008). Both freshwater ecosystems (wetlands, aquifers, rivers, and lakes) and terrestrial ecosystems provide hydrologic services (MA, 2005) and help to regulate the quantity and quality of water accessible to people, but the value associated with this service is difficult to capture; there is no easily obtainable market value. Water funds are a mechanism that uses the value of these hydrological services to enable biodiversity conservation. **Figure 1** is the map of ongoing and proposed water funds.

Water Funds: Protecting Biodiversity Through Ecosystem Service Payments

The Nature Conservancy (TNC), in collaboration with numerous partners, has successfully replicated a series of PES projects called water funds across Latin America. In these funds, payments from water users help protect the ecosystems that provide important hydrologic services and help improve management of working landscapes to ensure regular supplies of clean water. TNC has partnered with the Inter-American Development Bank (IDB), the Global Environment Facility (GEF) and FEMSA Foundation to improve the water funds model, to further replicate it, and to explore new sources of private and public funding for watershed conservation and management. The US Agency for International Development (USAID) has been another critical partner in replicating water funds, particularly in Ecuador.

Through water funds, TNC has more than matched the rapid advancement of watershed-based ecosystem service payment schemes in the global developing world in just one region (the Northern Andes) using just one project approach. In addition to funds not supported by TNC, today there are about 45 TNC-supported water funds in some stage of development in LAC (**Figure 1**; **Table 1**) ten of which are operational. What makes the water fund model so successful? Water funds are driven by water users, and the approach has the potential to be successful in a variety of political and institutional settings throughout the world. They have many of the elements for success identified in similar schemes (Porras *et al.*, 2008; Landell-Mills and Porras, 2002) but they go beyond the basics to become truly participatory, long-term, adaptable, multi-institutional, sustainable watershed conservation projects that with strong leadership and advocacy set them apart from the others.

History of the Water Fund Concept

About 80% of the water for the city of Quito, Ecuador (nearly two million people) comes from three protected areas and their buffer zones: Cayambe-Coca Reserve, Antisana Reserve, and Cotopaxi National Park. These areas contain about 5% of Ecuador's land area, include headwaters of more than 20 rivers and six larger watersheds, and provide critical habitat for many of Ecuador's bird and mammal species, but a variety of activities threaten this biodiversity and the availability of the clean, regular water supply. These threats include insufficient budgets to adequately control and conserve protected areas and buffer zones to prevent ecosystem conversion, illegal logging, and deforestation, among others. Much of the pressure on the land is from the people living in the watershed, as they depend on natural resources for their livelihoods. Available productive land is diminishing as soils lose nutrients, forcing families to move up the watershed into the natural ecosystems – a mixture of forest and *páramo* (high altitude grasslands) – that are the key hydrologic regulators of the system. Conversion means diminishing water services to people downstream, but keeping watershed communities out is unjust and unsustainable.

In the late 1990s, with the results of the USAID- and Fundación Antisana-funded SUBIR I and II projects in hand, TNC approached the mayor of Quito to demonstrate that protecting Quito's watersheds was crucial if citizens were to continue to enjoy the same regular supply of clean water in the future. TNC wanted to create a finance mechanism for

Figure 1 TNC water funds in Latin America and the Caribbean – mature, created, and in design. An illustration of mature, created, and in-design water funds throughout Latin American and the Caribbean where blue dots indicate mature, operational water funds; green dots indicate water funds that have been formally created but are only just starting operations; and red dots indicate water funds that are currently in the process of being designed. Data for this figure comes from Esri, DeLorme, NAVTEQ, TomTom, Intermap, iPC, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), and the GIS User Community.

Table 1 Current status, location and number of beneficiaries of water funds projected portfolio in Latin America and the Caribbean

No	*Water fund name*	*Country*	*Status of WF*				*Population*
			Idea	*Feasibility*	*Created*	*Consolidated*	*Beneficiaries*
1	Agua por la Voda(East Cauca Valley)	Colombia					920,000
2	*Cartagena*	Colombia					892,545
3	*Medellin*	Colombia					2,700,000
4	*Sierra Nevada de Santa Marta*	Colombia					TBD
5	*Bogota Agua Somos*	Colombia					6,840,116
6	*Cali*	Colombia					2,100,00
7	*Manizales*	Colombia					TBD
8	*Barranquilla*	Colombia					TBD
9	FONAG, *Quito*	Ecuador					2,300,000
10	FONOPA, *Cuenca Paute*	Ecuador					800,000
11	Tunguahau/*Ambato*	Ecuador					350,000
12	Guayas, Guayaquil	Ecuador					TBD
13	Procuencas, Zamora	Ecuador					25,000
14	Ayampe, *Puerto López*	Ecuador					16,000
15	Merida	Venezuela					630,000
16	Sixaola River	Panama					10,000
17	Terraba	Costa Rica					TBD
18	Sao Paulo/PCJ Watershed	Brazil					9,000,000
19	Sao Paulo/Upper Tiete Watershed	Brazil					TBD
20	Paraiba do Sul Watershed	Brazil					100,000
21	Guandu/Rio de Janeiro	Brazil					8,000,000
22	Camboriu Watershed	Brazil					1,000,000
23	Pipirupau/Brazilia	Brazil					350,000
24	Taquarussu/Palmas	Brazil					100,000
25	Minas Gerais PES	Brazil					TBD
26	Espirito PES state Program	Brazil					300,000
27	Sao Paulo State PES	Brazil					TBD
28	Parana PES State PES	Brazil					TBD
29	Rio de Janeiro PES State Program	Brazil					TBD
30	La Tigra, Tegucigalpa	Honduras					2,000,000
31	*Guatemala City*	Guatemala					2,900,000
32	Rivera Maya-Cancún	Mexico					2,000,000
33	*Monterrery*	Mexico					4,000,000
34	Sierra Madre de Chiapas	Mexico					800,000
35	Aquafondo, *Lima*	Perú					9,000,000
36	Arequipa	Perú					1,000,000
37	Trujillo	Perú					751,000
38	Piura	Perú					500,000
39	Santiago/Valparaiso	Chile					TBD
40	La Paz	Bolivia					2,300,000
Total							58,098,166.00

watershed conservation, and then Ecuador Program Director, Roberto Troya, obtained the support of the mayor, the municipality and the Quito water company (Empresa Municipal de Alcantarillado y Agua Potable de Quito – EMAAP-Q – the key water user) to do so.

The financial approach was carefully evaluated and selected because it was ecologically sustainable, legal, politically viable, efficient, and participatory (Krchnak, 2007). In January 2000, it was decided that voluntary donations from stakeholders would go into a trust fund that would form the financial basis of the water fund. The Quito Water Conservation Fund – Fondo para la Conservación del Agua or FONAG – was thus created as an endowment fund now receiving money from government, public utilities, electric companies, private companies, and nongovernment organizations. The donations were invested and managed by an independent organization through a trust, and initially, only the interest was used to fund activities. Over time, the financial model has shifted; now some contributions go into the endowment and some immediately finance conservation activities.

FONAG initially had two main members: TNC and EMAAP-Q. Other water users have since joined such as the Quito Electric Company (La Empresa Eléctrica de Quito – EEQ) in 2001 and private organizations including a beer company (Cervecería Andina) in 2003, the Swiss Agency for Development and Cooperation (COSUDE) in 2005, and a water bottling company (Tesalia Springs Co.) in 2005. Incentives for participation vary but are complimentary. The main incentive for EMAAP-Q and the other major water users was avoiding or reducing future costs for water treatment and

supply, functions provided by the conserved ecosystems. For TNC, the incentive was long-term financing for conserving protected areas (Goldman *et al.*, 2010a).

FONAG is governed by a board of directors comprised of water users who have contributed to the fund. The board approves the annual operational plan and reports, conducts audits, and reforms the bylaws. It also has a technical secretariat that acts as executive director.

In 2004, TNC invited the mayor of Quito to visit the PES mechanism in New York City, the example upon which FONAG was created. As a result of this visit, the mayor helped support and pass a municipal ordinance dedicating 2% of water fees to FONAG. TNC and the FONAG secretariat have been instrumental in helping explain and explore the benefits nature can provide to people, in this case, hydrologic services. USAID was another critical component of FONAG's success and has been a key player in strengthening FONAG and in financing a variety of the fund's activities.

The water fund concept has been improved dramatically since FONAG. Several have used cutting-edge science to meet challenges like the funds in Bogotá and the Cauca Valley of Colombia and in the Paute watershed near Cuenca, Ecuador. New users like major hydropower companies and big agricultural producers are involved and are beginning to understand and learn about the value of nature to their business operations. TNC has proposed a model and a business plan to replicate water funds in new places around the world.

The Institutional Workings of Water Funds

Clean water is a basic human necessity: not only do people require it for drinking, but many industries require it for production including everything from electricity to beer. In many watersheds, this life-supporting service comes from surface flows and aquifers. Nowhere is this truer than in the Northern Andes region of South America. Here, surface flows from high altitude head waters supply water to diverse users most notably several large cities (e.g., Bogota and Quito), hydropower companies, and in some cases, agricultural producers. These users have a vested interest in maintaining clean, regular water flows at the lowest cost and in managing supply risks. Water funds emerged from these concerns.

Water Funds are an innovative way to pay for nature's services – clean freshwater and biodiversity – by investing in conservation projects that protect the healthy habitat from which the services derive. Each water fund has its own set of objectives and goals, but in general, they invest to: (1) improve or maintain water quality and water quantity for downstream users; (2) maintain regular flows of water throughout the year; (3) maintain or enhance natural ecosystem biodiversity, both freshwater and terrestrial; and (4) improve or maintain human well-being and quality of life for upstream human communities (Goldman *et al.*, 2010a).

Historically, water flows and water cleanliness have been naturally regulated by predictable precipitation patterns and natural vegetation, but deforestation and the degradation of natural ecosystems, increased human demand for water, and a changing climate are all threatening and disrupting the ability of ecosystems to provide these services. Thus, water purification and storage services are now increasingly provided by water filtration plants or regular dredging of reservoirs to remove unwanted sediments. Water funds aim to restore or maintain nature's ability to provide these critical hydrologic services. In the long run, it could be cheaper and more sustainable than investing in business as usual, that is, gray infrastructure.

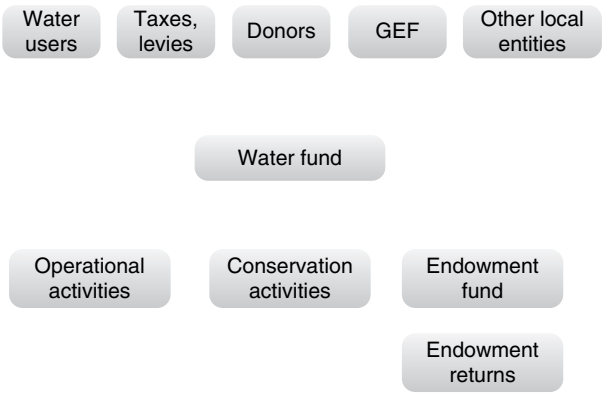


Figure 2 Water fund general model including financial flow and operational process. A wide variety of water users and other donors have invested money in water funds. Depending on the fund, portions of investments are directly invested in operational activities and conservation activities (protection and restoration) while the rest of the fund money is invested in an endowment fund. Capital from the endowment fund similarly funds operational and conservation activities.

Figure 2 summarizes the basic water fund model. The funds attract contributions from water users such as water utilities and local industries. These contributions build the funds' capital, including an endowment fund. In turn, endowments are invested in a wide range of assets (e.g., money market, bonds, and stocks) and the revenue generated from those investments provides long-term, secure funding for conservation projects like creating and strengthening protected areas, helping neighboring landowners switch to conservation-friendly management practices, paying for conservation easements, and financing environmental initiatives for local communities. Species benefit from having larger and better protected territories, communities benefit from a healthy watershed and improved land management, and large water users benefit from the resulting reduced water treatment costs. In addition, investments in watershed management reduce the risk of future clean water shortages that promotes long-term economic growth. With watershed conservation as a common objective, water funds create a governing body often bringing together public and private partners to manage it.

Investing in Watersheds: Providing the Hydrologic Services

Water funds invest in nature and in people by resolving potential conflicts between the resource needs of watershed communities and the conservation needs for water-service supply. Thus, investments focus on:

1. Managing public protected areas; and
2. Implementing best management practices on productive systems.

Ensuring good conservation management and providing incentives to conserve and restore natural ecosystems can provide long-term protection for the watershed but only if the incentives are sufficient to compensate people for any negative impacts.

Public Protected Area Management

The Northern Andes Region contains 20% of the World's biodiversity in an area that covers only 0.2% of the world's surface; 17% is protected and a relatively intact mixture of native forests and *páramo*. However, growing demand for land and access to resources to support livelihoods is increasingly threatening protected areas, and the level of government investment is not sufficient. For example, in 2005 according to the Ministry of Environment in Ecuador, there were 31 legally declared protected areas and there was a budget shortfall of ~$9.5 million (all $ as USD unless otherwise stated). For even the most basic management the gap was still ~$3.6 million. Despite numerous efforts, few programs address the link between managing protected areas and drinking water (Echavarria, 2002; Benitez *et al.*, 2010). *Páramo*, in particular, is a crucial water regulator; disturbance can damage or destroy its hydrologic functioning (Buytaert *et al.*, 2006, 2007). Water funds do address the link between drinking water and protected areas and can supplement budgets by cofinancing park guards, creating community-based ecotourism programs, and improving infrastructure management (**Table 2**). Park guards paid by the water fund are from the local communities ensuring community participation and acceptance, and providing a stable source of income for participating families.

Implementing Best Management Practices

In the watersheds, families rely on crop and ranchlands for income in and around protected areas. Managing these lands can have major impacts on water quality, the timing of flows (particularly floods), fires, and freshwater biodiversity. Water funds seek to minimize these impacts by providing direct and indirect payments to families for land management that includes setting aside conservation areas (e.g., fencing off riparian areas and headwaters), restoring and revegetating priority areas, silvopastoral management (e.g., live fences, forage plants, and rotating pastures), creating agroforestry systems, and/or creating tourism facilities (**Table 2**).

Specific conservation management practices supported by the water funds will vary by location, but investing in land management is essential for providing ecosystem services.

Table 2 Conservation activities of water funds

Activities	*Description*	*Type of strategy*	*Relation with water and biodiversity*		
			Area maintenance (conservation)	*Best practices*	*Restorations*
1. Public protected area management plan implementation					
1.1. Co-finance park guards	Improve control of high-risk conservation areas	Threat abatement	■		
1.2. Community-based eco-tourism programs	Reduce threats to buffer zones through income substitution	Threat abatement	■	■	■
1.3. Improve infrastructure management	Support best practices on management of current and new infrastructure in the park	Threat abatement	■	■	■
2. Best practices at the farm or productive unit					
2.1. Set aside conservation areas	Environmental payment for areas set aside for conservation on farmland: along streams, headwaters or forest connectivity	Conservation	■		
2.2. Set aside areas and restoration	Restoration payments and future environmental payments for areas set aside for conservation and restoration	Conservation			■
2.3. Silvopastoral systems	Improve productivity of farm with more environmentally-friendly cattle ranching practices such as live fences, forage plants, rotation of pastures	Best practices		■	
2.4. Agroforestry systems	Introduce environmental practices in the farm	Best practices		■	
2.5. Tourism facilities	Income substitute for land use practices	Conservation	■		■

These services can benefit the landholder (by enhancing soil stability and nutrient cycling), other people in the watershed (by providing clean and consistent water flows), as well as the broader ecosystem (by protecting habitat and biodiversity). Changing management practices is not, however, without cost. Restricting access to fertile soils in natural ecosystems and encouraging conservation management on productive lands involves trade-offs. Families are left with less land, so adequate, appropriate compensation is necessary.

The Payments: Compensating Impacts

Water funds must be flexible as they are established in geographically diverse places and in different legal and institutional settings. This includes the type of compensation provided to watershed communities. In the Northern Andes, indirect payments and livelihood investments have worked best. Indirect payments are in the form of materials and training to improve land management such as supplies for fencing and seeds for revegetation. Ideally, conservation management will enhance farm/ranch productivity by producing on-farm ecosystem services such as soil stabilization and enhanced soil fertility, but these benefits will not be immediate and are not guaranteed. In the shorter term, conservation management agreements include livelihood investments such as environmental education programs, alternative income sources like guinea pig farms, alternative food sources like organic vegetable gardens, and expanded capacity for production like building a milk bottling plant in the community to reduce shipping costs and payments for outside bottling fees, among others. As described later (*see* The Case of Brazil), in addition to indirect payments, Brazil uses mostly direct payments to landowners to compensate them for the opportunity costs of their traditional land uses.

Financial Sources and Management

Flexibility is also required in financial management, as are transparency and longevity, to ensure success under different legal and institutional frameworks. The core financial mechanism is a trust fund or endowment fund (**Figure 2**). This unique fiduciary structure involves an independent financial institution. A trustee (nonpartisan) manages and distributes payments to recipients (e.g., watershed communities and/or park guards) based on decisions made by the trustors, that is, the water fund board that generally comprises the main water users and other stakeholders. National and regional rules and regulations dictate the preferable investment scheme for the fund (private, nongovernment, etc.). All the trust funds have long-term contractual arrangements explicitly defining their use.

Water user contributions to the trust create a principal that accrues interest revenue for conservation projects. Ideally, the principal is untouched and only the interest is used. TNC and partners have learned, however, that in order to secure more partners and more contributions, at times it is important to use part of the principal in the first few years to demonstrate tangible progress as interest takes time to accumulate (**Figure 2**). Thus, the initial financial arrangement must include what percentage of the fund, if any, can be used. Some of the water funds, Fondo del agua para la conservación de la cuenca del río Paute (FONAPA – Water Fund for the Conservation of the Paute river watershed), for example, have designated a portion of the trust to go directly to conservation activities. This makes sense in smaller watersheds that are less likely to have large capital investments earning significant interest in a reasonable time.

Water funds can be financed from a variety of private and public sources including the following:

- *Water users* (*e.g., water utilities, bottling companies*): potentially the largest beneficiaries, the users are also often the largest financial contributors.
- *Citizens:* a fundraising proposal for the Bogota water fund was to get donations from citizens through their water bills. Contributions from the general public (low contribution, high volume) can be a significant source of funding.
- *Taxes, levies, and public programs*: working with existing local regulations, taxes, fees, or special-purpose contributions can be a strong source of funding. In FONAG, a municipal ordinance requires the water company to direct 2% of tariffs to the fund. In countries such as Colombia and Brazil, water laws obligate water users, municipalities, and environmental regional authorities to invest resources in the watershed, and water funds have managed to capture their interest since they leverage other money, are participatory, and can function as implementation arms for public funding. In some cases, this might be an ideal revenue source for water funds.
- *Grants, international organizations, and private foundations*: funding from bilaterals, multilaterals, or independent foundations can play a strategic role in the first 2 years of a water fund by helping to establish it including funding feasibility studies. TNC has provided this support in the past. Grants are also used to fund specific activities in the conservation plan. USAID, for example, has been very important in supporting water funds in the Andean region.
- *GEF*: funding will be directed to critical start-up costs, specifically to the endowment fund, and will co-finance outreach and feasibility studies. In Brazil, money will be spent only on technical support and conservation activities.
- Financial returns generated from the endowment fund.

Revenue sources vary from one water fund to the next depending on a variety of factors such as the legal framework for water policy, private sector opportunities, environmental service provisions, and governance strategies. Based on current experience, the majority of funding will come from the public sector and from water users. Funding through grants, international organizations, and private foundations, while proportionately small, is essential in ensuring that the water funds have a solid foundation.

Governance

Water funds are governed not by a separate nonprofit or nongovernment entity; rather, they are governed by contractual partnerships which create a collaborative, stakeholder-based, decision-making board that elects and approves

a technical secretariat that calls meetings and implements decisions. In addition, the board can establish technical committees to help guide its investment decisions. The main contributors to the fund necessarily have a voice on the board, but noncontributors and indirect contributors can also participate. For example, in the Procuencas water fund in Zamora, Ecuador, the Ministry of the Environment, a noncontributor, has a seat on the board as their input and support is essential. In Tungurahua, Ecuador, numerous, important indigenous communities in the watershed will be affected by the water fund and have invaluable knowledge about the watershed, so the German Technical Cooperation Agency GIZ donated money in their names so they have representatives on the board. TNC contributes both directly and indirectly, and in many cases has a representative on water fund boards. In the Cauca Valley in Colombia, each of the nine watersheds that feed the East Cauca Valley has a grassroots, nonprofit organization that has been working with the communities for many years. The directors of these organizations have seats on the technical committee to ensure community input.

The board sets priorities and makes decisions on investments and approves activities, but a lesson learned from FONAG was to ensure a scientific or technical basis for those decisions; thus, most are now based on feasibility or analytical studies and often on advice from a technical committee composed of scientists and engineers. The overall obligation of the board, with input from the committee(s), is to create and implement a strategic and an operational plan that include specific objectives and the means to achieve them.

How is a Water Fund Created?

While there cannot be a single formula for creating a water fund given the diversity of their features, there is a set of general steps that should be followed to establish a successful one. These steps are the result of a decade of lessons learned by TNC in replicating, improving, and advancing funds across Latin America (see Calvache *et al.*, 2012 for more detail).

Step 1: Can We Even Think About a Water Fund Here?

The first step is to determine if the right service users and service providers exist in the same area. This requires an ecosystem services assessment to determine if specific areas for generating the hydrologic services upon which a specific set of water users depends can be defined. There must be an explicit, measureable relationship between a specific area, the services the ecosystems in the area can provide, and the stakeholders in the area. Three important questions guide this step: (1) What is or are the service(s) that the water fund will prioritize for protection, conservation, restoration, and where is the long-term opportunity? (2) What is the area of influence for generating these services? and (3) Who are the key water users interested in these services, what is their interest, and how might the value of these services be internalized in their cost-benefit analyses? This type of work is best done by establishing a core working group that includes a main water user when possible.

Step 2: Will a Water Fund be Successful Here and How?

A water fund will most likely be effective if technical, financial, and legal studies are conducted beforehand to demonstrate the importance of the water fund to water users and to provide information for setting investment priorities if and when the fund is created. Determining what studies are needed is best done by creating a working group that includes the key stakeholders identified in Step 1. Work can then be divided among the participants. The detail and depth of the studies will depend to a great extent on available data and resources, but they should attempt to answer the following questions: (1) What is the demand and supply status of current hydrologic services? (2) What are the potential benefits or impacts – environmental and socioeconomic – of the water fund? (3) Where are priority areas for water fund investments? (4) What is the cost of maintaining or improving hydrologic services? (5) How will the provision of hydrologic services change as land use and climate change? and (6) What services other than hydrologic services can the water fund maintain? There are a growing number of tools (e.g., InVEST, SWAT, and ECOSAUT) to analyze the feasibility of a water fund and the potential impacts of its investments.

A final critical study is a legal and institutional assessment/analysis. Water funds ideally need to be transparent, independent, and permanent. An institutional analysis can help to identify legal hurdles and opportunities and to define the most appropriate financial and governance structure for the fund given the particularities of the country. National, regional, and local laws are all important to consider and follow in water fund design.

Step 3: Designing and Negotiating the Fund

With the stakeholder analysis and the feasibility studies in hand, the structure of the fund can be designed, contracts can be negotiated, the governance mechanism can be determined, administrative principles can be defined, and a contact can be signed. As water funds involve stakeholders from a wide variety of institutions, it is essential to have a clear definition of roles and responsibilities. The structure should be specified in a contract that formalizes the partnership. From experience, TNC has found that a useful structure is the previously described board of directors, technical secretariat, and technical committee(s). The board is a formal public–private partnership working under a mutually agreed, legal contract to ensure that each stakeholder has a role and the obligation and incentive to carry it out.

Finally, a process for handling administrative details must be determined. These include managing fund activities, the money in the fund, and other administrative matters that arise during formation and implementation.

Step 4: Details, Details, Details

Taking into account the studies in Step 2, the basic components that will allow the fund to achieve its social, environmental, economic, and institutional goals are then decided. The goal is to establish the first board of directors that will then appoint the technical secretariat. In addition, a

technical committee should be created to help guide decision making.

After these entities are functioning, the board should develop a strategic plan and a financing plan. The strategic plan lays out the basic components of the fund's operational plan that should reflect the priority areas and activities in the feasibility assessments. The strategic plan should highlight the objective of the investments, the goals and how they will be achieved, a timeline, and costs. Fundraising is a critical part of water funds particularly to finance current activities, and bringing more water users into the fund can increase the amount of money in the trust thereby enhancing sustainability.

Step 5: Activities Begin

Operationalizing the fund means executing the plans created in Step 4, that is, implementing the proposed investments and all the activities needed to achieve the goals. Execution is mainly the responsibility of the technical secretariat, but this should be supported by the technical committee that can provide advice and suggestions to improve implementation.

Actual implementation means further prioritizing activities for investment with a view to maximizing efficiency – achieving the greatest impact for least cost. Scenario-based, land-use change analyses can help determine the best activities for achieving the goals. The major activities implemented thus far by water funds have been described previously.

The financial management plan created in Step 4 should include goals and a timeline for acquiring future funding and for capitalizing the trust. Implementation includes launching communication and fundraising plans. Interest in participating in a fund can increase or decrease, so it is important to constantly attract new donors and participants.

Finally, a plan should be developed to disseminate results beyond just the stakeholders in the water fund. This can often be done by contracting a marketing and communications firm. Fundraising can also be done by contracting an appropriate organization.

Step 6: Evaluating and Monitoring

A final, but very important, step is to monitor and evaluate outcomes to ensure the goals are being achieved and to promote adaptive management. A monitoring plan is a tool to help maintain and expand the technical and financial support for the water fund effort. It should cover socioeconomic, biophysical, institutional, and economic impacts. Monitoring outcomes can help to do the following: (1) ensure activities implemented are achieving desired goals; (2) make adaptive management possible allowing for changing and adjusting activities; (3) improve communications about the fund's activities and benefits; and (4) improve water fund transparency.

Examples of Water Funds: Ecuador and Colombia

As described previously, water funds are flexible ecosystem-services-based payment mechanisms adapted to different localities, regions, and countries. The following are two examples of water funds that have succeeded in different contexts. FONAG in Quito, Ecuador was the first water fund and is now demonstrating its impact. The Water Fund for Life and Sustainability – Fondo de Agua por la Vida y la Sostenibilidad – FAVS in Colombia is a water fund in the Cauca Valley whose water users are an association of sugarcane farmers.

FONAG – Quito, Ecuador

Previously we described FONAG; here we focus on outcomes. In 2000, FONAG had $21,000 invested in the trust fund – money from EMAAP-Q ($20,000) and from TNC ($1.000). By 2008, the trust fund had grown more than 250 fold to about $5.4 million and is now (2011) nearly $9 million. In 2008, this endowment yielded about $800,000 in interest to spend on conservation projects (FONAG 2008). In 2007, FONAG, with the support of its board members, helped to pass a municipal bylaw requiring the Quito water company (EMAAP-Q) to provide 2% of its revenue to the water fund (an increase from the initial 1% commitment).

FONAG was initially structured so that only endowment interest was used on conservation activities along with other money leveraged by the fund but not invested in the endowment. These investments were often substantial. From its initiation through 2008, $7.1 million was donated to FONAG as matching funds from a variety of other donors. USAID, in particular, has been one of the most important supporters of FONAG throughout; other donors include InWEnt – Capacity Building International, Germany; IDB, and EcoFund (FONAG 2008). This past year, the board decided to change the financial model and in the future, 30% of funds will be invested in implementing current and future activities and 70% will remain in the endowment.

FONAG uses the revenue from the water fund to finance various programs and projects. The programs currently underway are controlling and monitoring conservation areas, restoring natural vegetation, environmental education and outreach, training in watershed management, communications, integrated watershed management, and hydrological monitoring. The main beneficiaries of the activities are the communities close to the water sources. They receive permanent support from FONAG through various programs. From 2000 to 2010 FONAG has achieved the following:

- helped to conserve the watersheds that provide 80% of Quito's water (population of about 2 million);
- had a positive impact on 500,000 ha/1.2 million acres of land;
- enrolled 30,500 children in environmental education programs;
- enriched standing forest to improve forest density and quality on ~600 ha of land/year from 2006 to 2010;
- reforested 2033 ha/5023 acres of land with over two million trees;
- hired, trained, and salaried 11 park guards from local communities to help conserve protected areas; and
- engaged over 200 families in community development projects in rural basins.

Recent monitoring and evaluation projects have helped to demonstrate FONAG's impact. By analyzing the biodiversity of conservation areas supported by water fund investments and comparing them to areas with no such investments and a "pristine" reference site, it was found that the species composition of water fund-supported conservation areas are similar to the reference site but differ considerably from areas without water funds. There have also been reductions in grazing impacts and fire damage in several vast areas where the water fund supports protection. Finally, the analysis demonstrated that the conservation activities of the water fund helped maintain the ecology of the sites intact in an economically efficient manner (Boucher, 2011).

An analysis of water-related impacts revealed that waterways with water fund investments have greater ecological integrity, improved riparian and aquatic habitat quality, reduced erosion, and a more balanced temperature than waterways without fund investments. Additionally, the richness and diversity of macroinvertebrates in water fund waterways were greater than waterways with no water fund investments indicating better water quality and increased ecosystem integrity. These results were not true for each individual site analyzed (eight sites were compared with eight similar nonwater fund sites) which helped to demonstrate where adaptive management might be needed to ensure that more goal-appropriate activities are implemented in those areas (Encalada *et al.*, 2011).

In terms of impacts to the people living in the watershed, communities not affected by FONAG were compared socioeconomically with communities affected by the water fund across a number of variables. The results revealed that communities with FONAG interventions had the following: (1) a great diversity of plant and animal species in the watershed area; (2) greater investments in subsistence farming particularly in horticulture and meats that had lead to a reduction in expenditures by households in the market and increased food security; and (3) improved agricultural management. Communities impacted by FONAG also had a better understanding of their legal rights and had fewer community conflicts (Delgado and Mosquera, 2011).

As the oldest of the TNC water funds, FONAG has demonstrable results both in terms of actions implemented and outcomes. FONAG is only now developing studies to help prioritize investments; thus, the positive outcomes of the fund could potentially increase.

Fondo de Agua Por La Vida y La Sostenibilidad – Cauca Valley, Colombia

The FAVS water fund – Water for Life and Sustainability – in the Cauca Valley of Colombia lies near the city of Cali and serves a watershed that contains a huge number of sugarcane producers, an important export and domestic crop for the country. The water users are these sugarcane growers and about 900,000 people residing in cities in the watershed. As in FONAG, providing hydrologic services is dependent on conserving upstream natural ecosystems and managing the lands of rural watershed communities.

Led by TNC in June 2009, FAVS was launched by bringing together a variety of partners: ASOCAÑA (Colombia's sugarcane producers association), CVC (Corporación Autónoma regional del Valle del Cauca, the local environmental authority), Vallenpaz (a peace and justice organization), and nine grassroots organizations representing nine subwatersheds that feed the valley. Other stakeholders have since joined including the sugarcane growers association (Procaña), the technical branch of the sugarcane association (Cenicaña), and two other grassroots organizations representing two additional subwatersheds. In 2009, the fund contained $1.8 million in capital from the sugarcanes producers association (ASOCAÑA). In the first 2 years of its operations, FAVS raised more than $5 million from both the private and public sectors (UNICEF, PAVCO, Colombian Oil Company ECOPETROL, Regional Environmental Authority CVC, others). Current negotiations with the sugarcane growers association (Procaña) are likely to bring in an additional $2.3 million per year to the fund.

The principle objectives of FAVS are to increase the natural vegetation in the watersheds in order to maintain their hydrologic services, to conserve biodiversity, to provide water for consumption and for industry, to irrigate lands, to provide recreation, to generate energy, and to protect fishing resources. The Cauca Valley has the most productive sugarcane land in the world; damage to the region's forests and water resources threatens to reduce production significantly, which could cost the industry $33 million each year, according to the sugarcane producers association (ASOCAÑA) (based on TNC estimates of water yield and climate change that basically indicate that the length of irrigation cycles will decrease).

Activities in the watersheds that the fund is helping to support include isolating riparian areas to keep cattle and crops from riversides, protecting headwaters, reforesting and restoring landscapes, revegetating pasturelands, and conserving natural ecosystems. Ultimately the aim is to conserve and restore about 30,000 ha of land. In addition, through livelihood investments the water fund is promoting food security for local communities, environmental education, and capacity building to ensure sustainable production. FAVS aims to benefit 1500 families directly and to help provide water for 1.25 million people. To date, FAVS has spent more than $1.4 million on conservation projects: $0.1 million on the first water fund project; $ 0.6 million in the first call for proposals cycle and $ 0.7 million in the second proposal cycle. With these investments, FAVS has built more than 80 km of protected river fences; protected more than 87 headwaters; conserved over 250 ha of land through restoration and natural regeneration; converted more than 80 ha of critical land along streams to sustainable cattle ranching, supported the Las Hermosas Natural Park, supported nine schools by promoting environmental education, and supported at least 24 families.

FAVS has also helped to demonstrate the feasibility and importance of doing analytical studies to set priorities for water fund investments. In addition, it was the first water fund to test a methodology for ensuring that priorities for activities were based on cost and return on investment. The goal was to identify priority areas for FAVS investment, establish quantitative ecosystem service goals, and develop a portfolio of the most efficient activities. To achieve it TNC and partners used a watershed scoring process and a modeling tool called InVEST (Integrated Valuation of Ecosystem Services and Trade-offs),

developed by the Natural Capital Project (Tallis *et al.*, 2010). This evaluation involved several steps. First, a conservation activity (forest enrichment, reforestation, fencing, or silvopastoral practices) was assigned to each part of the landscape based on the behavior of landowners in the region and successful investments made in the watershed over the last 20 years. These assignments implicitly considered factors such as opportunity costs and land owners' willingness to change their activities (this willingness was evaluated in collaboration with the watershed grassroots community organizations). The landscape was ranked to highlight the places where possible conservation investments were likely to yield the greatest improvement in water yield and erosion control. Factors included in the ranking were those known to affect the hydrological response of the services such as slope, soil depth, distance to stream or water body, aspect, elevation and precipitation. Data from historic conservation investments in each watershed were used to estimate how much the proposed conservation activity in each location would cost (Goldman *et al.*, 2010b).

Then the landscape ranking and cost information were combined to select the highest ranked locations for each activity, tallying costs until the target budget level was met. The activities selected across the landscape formed the water fund's investment portfolio. This process was repeated for five budgets ranging from the level of investment currently committed by the fund ($10 million) to double that amount. With this set of investment portfolios as scenarios for future management, InVEST was used to estimate the ecosystem service returns from each (Goldman *et al.*, 2010b), specifically for erosion and annual water yield providing preliminary estimates of return on investment. The model provides a relative change (%) expected as spending progresses. One of the model's results demonstrated that for one watershed, erosion control benefits would increase from 1% in year 1 to 14% in year 5 based on water fund investments (The Natural Capital Project, 2011). Basically, InVEST helps to determine the most efficient investment portfolio for providing the two main services of interest: annual water yield and avoided sedimentation.

Work is now being done on this water fund to incorporate the impacts of climate change into the scenarios and the activity assessment. This impact assessment helps to determine if current activities promoted by the water fund are adapted for climate change and will allow for the design of activities to promote resilient ecosystems (Goldman *et al.*, 2010a). In particular, new ecosystem services maps incorporating suggested changes based on climate vulnerabilities were prepared and discussed with key stakeholders. These maps will complement the water fund portfolio in order to make sure that FAVS activities will be aligned with climate change analysis. This information was complemented with a local expert's suggestions in terms of the main problems related to land management and best practices. A complete conceptual model was developed for this workshop with local communities to help them to understand the links between land use practices and problems derived from climate change, and priority strategies for climate change adaptation were identified.

Not only did FAVS use studies to guide the investments of the water fund, but the fund also includes community members on the governing board. The watersheds that feed the Cauca Valley have suffered extreme levels of violence and disruption in the last decade; thus, providing the communities with a voice in how the fund operates ensures sustainability and equity. FAVS is also unique in that the main water users are agricultural producers, which demonstrates the potential for including a variety of stakeholders in water funds. A final characteristic that contributed to the fund's success is that investments are building on a legacy of watershed conservation projects already underway in the region funded by ASOCAÑA in collaboration with the grassroots community organizations. The water fund is helping to formalize and expand these prior investments by bringing the water users and providers together in a partnership to ensure joint decision making and scientifically sound investment choices.

The Case of Brazil

In the Atlantic Forest and other biomes of Brazil, TNC and partners have had great success in replicating a similar watershed conservation model called the water producer concept that has now been added to the water fund portfolio of TNC (*see* Water Fund Platform). The water producer concept was first developed by the National Water Agency (ANA) and first implemented by ANA, TNC, and several state and local partners. It recognizes the positive externalities generated by landowners living in the headwaters of watersheds when they implement forest and soil conservation and restoration activities and that people downstream benefit from these positive externalities. The concept is based on the premise that those who benefit should pay directly for that which benefits them, in this case watershed conservation. Thus landowners are compensated for their opportunity costs for generating the hydrologic services downstream users require.

Brazil's Atlantic Forest is the most densely populated region in Latin America as it is home to 11 major cities including Sao Paulo and Rio de Janeiro and 70% of Brazil's population. Since 2006, TNC and partners have been working in the watersheds of the Atlantic Forest that provide 50% of Sao Paulo's drinking water and 80% of Rio de Janeiro's. Brazil holds more water than any nation, but only 1/12 of that water can be found along the southeastern coast where nearly half of the population lives (Bradley, 2010). The forest is also one of the most biodiverse areas on earth, with 5% of the earth's vertebrates and 8% of the planet's plants. Unfortunately, much of the forest is becoming rapidly degraded and only 12% remains. Clearing and fragmentation is the main threat to its viability in the long term.

Deforestation threatens the provision of a clean, reliable water supply and signals an urgent need to protect and restore Brazil's Atlantic Forest. Water producer projects involve direct payments from water users that go toward maintaining and enhancing watershed forest cover. The incentive for water users is sustainable clean, regular water supplies at a lower cost. In Sao Paulo, for example, pollution and sediment in the water have doubled treatment costs in 8 years (Bradley, 2010). Investments in reforestation provide a natural pollution and sediment filtration system that can help curb these costs in the future.

The implementation of water producer projects in Brazil has followed three main routes (Veiga and Gavaldão, 2011). The first one is a result of the National Water Policy that established a water-user fee and required the proceeds to go toward maintaining and increasing watershed health. In addition, it created watershed committees that comprise water users and representatives of governments and civil society and have the legal power to decide the best way to spend the proceeds. The implementation of the policy and the creation of the watershed committees is ongoing gradually throughout Brazil and has started with more urbanized states such as São Paulo, Minas Gerais and Rio de Janeiro where the water-use conflicts are more prominent. The fundraising potential of these watershed schemes is significant. One example is the Piracicaba-Capivari-Jundiaí Watershed (PCJ), one of Brazil's key watersheds, which raises approximately $25 million per year in water user fees.

Creating a water producer project from this policy arrangement involves several steps. The concept is presented to the watershed committees as one of the best investments that the committee could make to guarantee water in quality and quantity from the watersheds in their jurisdiction by explaining that the concept is based on direct payments to compensate watershed landowners for the positive externalities they generate when they restore and protect their lands. The first water producer project created this way began in 2006 in the PCJ Watershed where the watershed committee allocated $250,000 for a first pilot in the watershed matching the funds allocated by a coalition of partners including ANA, the Environmental and Agricultural State Agencies of São Paulo State (SMA-SP and SAA-SP), and TNC (Veiga, 2009). These same partners created a project management unit that manages the pilot and creates plans for its replication within the watershed. Since this first pilot, more initiatives have been replicated in PCJ and other watersheds in the country, including the Guandu Watershed which provides water for the Rio de Janeiro metropolitan area.

A second means by which water producer projects have been created requires municipal and state legal approval explicitly for developing PES schemes. These laws are important for two reasons. They can provide a legal framework for PES implementation, and they can enable and facilitate the use of public funds for PES schemes. For example, one water producer project was based on a 2005 law in Extrema, Minas Gerais, that gave the municipality the ability to pay landowners who achieved specific targets related to soil conservation measures, rural sanitation, and forest conservation and restoration. In addition, this law allowed the first direct payment to a water producer to be made in Brazil in February, 2007. This initiative was a partnership between the municipality, ANA, the State Forestry Agency, and TNC, and support was also received from the PCJ watershed committee. In addition to direct payments for environmental services, the landowners also receive financial, technical, and in-kind support to reach conservation targets. Another example at the state level occurred in 2008 in Espírito Santo where a law was passed that created a water fund called FUNDÁGUA that was primarily funded by a percentage of the oil royalties collected by the state. FUNDÁGUA stipulates that most of the fund's revenues should go toward paying landowners in priority watersheds for the water services they provide. The first payments started in March 2009; by July 2011 approximately 300 landowners received direct payments from FUNDÁGUA.

With these examples, several other municipalities and states have been discussing and implementing laws and PES programs. TNC has been a partner in nearly all of the initiatives, especially those at the state level realizing the great potential these mechanisms have to conserve watersheds throughout the region. At the federal level in Brazil, there is a PES bill proposal under discussion at the National Congress that was informed by these first municipal and state initiatives.

A third means through which a water producer project has been created is more similar to the water funds in other parts of Latin America and has emerged more recently in Brazil. Here a key water user such as a water utility company takes the lead to implement a PES scheme. One example is a project headed by the Balneário Camboriú Municipality Water utility (EMASA), located in one of the most important touristic cities in Southern Brazil. This utility understood the benefits of investing in watershed conservation to reduce its water treatment costs and so has allocated $1.5 million to the PES scheme. Using other water fund projects as examples, the utility and their partners (the municipalities of Camboriú, Balneário Camboriú, ANA, TNC, and the local watershed committee, among others) intend to create a municipal PES water fund to guarantee the initiative's sustainability in the long term. These schemes are also starting to be replicated. For example, SANEATINS, the water utility company of Palmas, the capital of Tocantins state, is now initiating a water fund.

A final, very nascent method for creating a PES scheme similar to a water fund in Brazil involves companies that are trying to offset their water footprints. One mechanism by which they can offset water consumption could be to invest in a water producer project. This possibility is still under development but can be considered another important source of funding for the long term, especially if the project is set up like other water funds in Latin America.

In short, the financial mechanisms in Brazil are slightly different from those of the water funds of the Northern Andes region. In general, however, the funds use an annual distribution model where fees or other sources of funding are collected and distributed each year rather than going into a trust. The payments are direct and are used to both conserve standing forest and to reforest critical areas, and they are generally based on opportunity costs to the farmer from having to reforest and/or protect and/or improve management of their land. Payments tend to range from $29–$100 per acre per year depending on the project site and on institutional arrangements. There are additional criteria that are specific to the sort of scheme proposed such as a landowner's willingness to participate, slope intensity, and forest quality among others that can be combined with opportunity costs. The water fund supplies materials and capacity (hiring local people) to enable reforestation. Restoration costs also vary but are round $1600 to $2800 per acre. The landowner's responsibility is to protect and maintain reforested or forested areas in exchange for a payment based on the location and size of restored and/or conserved parcels.

Water Fund Platform

With all the valuable experience it has gained in developing diverse approaches to watershed conservation through PES schemes, TNC in collaboration with three main partners – IDB, GEF, and FEMSA Foundation, the largest Coca-Cola bottling company in the world – is launching a water fund platform. Using this platform, the characteristics of the water funds in the Northern Andes and the PES approaches in Brazil will be extended to water funds more broadly throughout Latin America, thereby conserving and enhancing valuable freshwater and terrestrial biodiversity while providing important hydrologic services to people.

In the next 5 years TNC and its partners under the Latin American Water Funds Partnership aim to create, implement and capitalize at least 32 water funds (**Figure 1**; **Table 1**). This will comprise investments of over $27 million to protect seven million acres of water sheds in Ecuador, Colombia, Peru, Brazil, Mexico and other countries in the LAC region (Marx, 2011).

The goals of the platform include the following:

- a total of 32 self-sustaining water funds of which 20 are new;
- conservation of more than seven million acres of important ecosystems;
- effective mechanisms for adapting natural ecosystems that supply water to climate change;
- secure access to water supply for more than 50 million people, industries, and rural areas;
- savings in water treatment costs and reduced productivity loss because of water shortages and reductions in energy supply in rural areas and for industries;
- water funds capitalized with at least $27 million (total across all funds) from public and private local funds after 5 years;
- at least $143 million in conservation projects, total across all water funds, on watersheds leveraged from other institutions.

As of December 2011, the platform encompassed 45 water funds in various stages of implementation (**Figure 1**; **Table 1**), most of which were just starting operations or were in the initial stages of assessing feasibility and viability.

See also: Biodiversity and Ecosystem Services. Conservation and People. Economic Value of Biodiversity, Measurements of. Economics of the Regulating Services. Ecosystem Services. Human Impacts on Ecosystems: An Overview. Impact of Ecological Restoration on Ecosystem Services. The Value of Biodiversity. Valuing Ecosystem Services

References

Asquith N and Wunder S (eds.) (2008) *Payments for Watershed Services: The Bellagio Conversations*. Santa Cruz de la Sierra: Fundación Natura Bolivia.

Barbier E (2004) Agricultural expansion, resource booms and growth in Latin America: Implications for long-run economic development. *World Development* 32: 137–157.

Benitez S, Blanco A, Cole J, Ibáñez M, Rodríguez JJ, and Halloy S (2010) Using water funds to finance watershed conservation in the Andes and Costa Rica. *Mountain Forum Bulletin*, January 2010.

Bezaury-Creel JE (2009) *El Valor De Los Bienes y Servicios Que Las Áreas Naturales Protegidas Proveen a Los Mexicanos*. Mexico: The Nature Conservancy Programa México – Comisión Nacional de Áreas Naturales Protegidas.

Boucher T (2011) *A Terrestrial Ecological Impact Assesement of the Quito Water Fund*. Arlington, VA: The Nature Conservancy.

Bovarnick A, Alpizar F, and Schnell C (eds.) (2010) *The Importance of Biodiversity and Ecosystems in Economic Growth and Equity in Latin America and the Caribbean: An Economic Valuation of Ecosystems*. United Nations Development Programme

Bradley T (2010) Brazilian Water Protection a $100 Million Market? *National Geographic News*, June 4, 2010.

Brauman KA, Daily GC, Duarte TK, and Mooney HA (2007) The nature and value of ecosystem services: An overview highlighting hydrologic services. *Annual Review of Environment and Resources* 32: 67–98.

Buytaert W, Celleri R, De Bièvre B, Cisneros GW, Deckers J, and Hofstede R (2006) Human impact on the hydrology of the Andean páramos. *Earth-Science Reviews* 79: 53–72.

Buytaert W, Iñiguez V, and De Bièvre B (2007) The effects of afforestation and cultivation on water yield in the Andean páramo. *Forest Ecology and Management* 251: 22–30.

Calvache A, Benítez S, and Ramos A (2012) Fondos de Agua: Conservando la Infraestructura Verde. Guía de Diseño. *Creación y Operación. Bogotá*. Colombia: The Nature Conservancy, Fundación Femsa y Banco Interamericano de Desarrollo.

CEPAL (2005) El Nuevo patron de desarrollo de la agricultura en América Latina y el Caribe. *Panorama 2005*. Santiago del Chile, Chile: United Nations.

Costanza R, d'Arge R, De Groot R, *et al.* (1997) The value of the world's ecosystems and natural capital. *Nature* 387: 253–260.

Daily GC (1997) Introduction: What are ecosystem services? In: Daily G (ed.) *Nature's Services: Societal Dependence on Natural Ecosystems*, pp. 1–10. Washington, DC: Island Press.

Daily GC and Matson P (2008) Ecosystem services: From theory to implementation. *Proceedings of the National Academy of Sciences* 105: 9455–9456.

Delgado A and Mosquera H (2011) *Análisis de Impacto Socio-Económico De Las Inversiones De FONAG*. Arlington, Virginia: The Nature Conservancy.

Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature* 418: 700–707.

Dudgeon D, Arthington AH, Gessner MO, *et al.* (2006) Freshwater biodiversity: Importance, threats, status, and conservation challenges. *Biological Review* 81: 163–182.

Echavarria M (2002) Financing watershed conservation: The FONAG water fund in Quito, Ecuador. In: Pagiola S, Bishop J, and Landell-Mills N (eds.) *Selling Forest Environmental Services: Market-Based Mechanisms for Conservation and Development*, pp. 91–102. London and Sterling: Earthscan.

Economic Commission for Latin America and the Caribbean (ECLAC) (2010) *Economics of Climate Change in Latin America and the Caribbean: Summary 2010*. Santiago, Chile: United Nations. Santiago, Chile: United Nations.

Encalada A, Ibarra LC, and de la Paz MC (2011) Diagnóstico de la integrdiad ecológica y la calidad del agua de los ríos en las zonas de manejo del FONAG. *Informe Final, Laboratorioa de Ecología Acuática de la Universidad San Francisco de Quito*. The Nature Conservancy 2011.

Engel S, Pagiola S, and Wunder S (2008) Designing payments for environmental services in theory and practice: An overview of the issues. *Ecological Economics* 65: 663–674.

Food and Agriculture Organization (FAO) (2011) *State of the World's Forest*. Rome, Italy: FAO.

Goldman RL, Tallis H, Kareiva P, and Daily GC (2008) Field evidence that ecosystem service projects support biodiversity and diversify options. *Proceedings of the National Academy of Sciences* 105: 9445–9448.

Goldman R, Benitez S, Calvache A, *et al.* (2010b) *TEEBcase: Water Funds for Conservation of Ecosystem Services in Watersheds*. Colombia: TEEBweb.org.

Goldman RL, Benitez S, Calvache A, and Ramos A (2010a) *Water Funds: Protecting Watersheds for Nature and People*. Arlington, Virginia: The Nature Conservancy.

Goldman RL and Tallis H (2009) A critical analysis of ecosystem services as a tool in conservation projects: The possible perils, the promises, and the partnerships. In: Ostfeld R and Schlesinger W (eds.) *Annals of the New York Academy of Sciences: The Year in Ecology and Conservation Biology*, pp. 63–78. USA: Wiley-Blackwell Publishing.

Heal G, Daily GC, Ehrlich PR, *et al.* (2001) Protecting natural capital through ecosystem service districts. *Stanford Environmental Law Journal* 20: 333–364.
International Energy Agency (IEA) (2007) *World Energy Outlook 2007*. Paris: IEA. 11 November 2008. www.worldenergyoutlook.org
Jack BK, Kousky C, and Sims KRE (2008) Designing payments for ecosystem services: Lessons from previous experience with incentive-based mechanisms. *Proceedings of the National Academy of Sciences* 105: 9465–9470.
Kareiva P, Chang A, and Marvier M (2008) Development and conservation goals in World Bank projects. *Science* 321: 1638–1639.
Kiesecker JM, Belden LK, Shea K, and Rubbo MJ (2004) Amphibian declines and emerging disease. *American Scientist* 92: 138–147.
Kiesecker JM, Copeland H, Pocewicz A, *et al.* (2009) A framework for implementing biodiversity offsets: Selecting sites and determining scale. *Bioscience* 59: 77–84.
Krchnak KM (2007) *Watershed Valuation as a Tool for Biodiversity Conservation*. Arlington, VA: The Nature Conservancy.
Landell-Mills N and Porras I (2002) Silver bullet or fools' gold? *A Global Review of Markets for Forest Environmental Services and Their Impact on the Poor*. London, U.K: International Institute.
Levin PS and Levin DA (2004) The real biodiversity crisis. *American Scientist* 90: 6–9.
Marx Eric (2011) Briefing: Water resources. *The Water Funds Solution: Water, Water Everwhere*…. Ethical Corporation October 2011.
McAlpine KG and Wottom DM (2009) Conservation and the delivery of ecosystem services. *Science for Conservation* 295: 1–81.
Millennium Ecosystem Assessment (MA) (2005) Millennium ecosystem assessment. *Ecosystems and Human Well-Being: Current State and Trends*. Washington, DC: Island Press.
Mittermeier RA, Gil PR, and Mittermeier GC (1997) Megadiversity: Earth's Biologically Wealthiest Nations. *Cemex*. Mexico: Conservation International.
Naidoo R, Balmford A, Costanza R, *et al.* (2008) Global mapping of ecosystem services and conservation priorities. *Proceedings of the National Academy of Sciences USA* 105: 9495–9500.
Nel JL, Roux DJ, Abell R, *et al.* (2009) Progress and challenges in freshwater conservation planning. *Aquatic Conservation: Marine and Freshwater Ecosystems* 19: 474–485.
Pagiola S, Ramirez E, Gobbi J, *et al.* (2007) Paying for the environmental services of silvopastoral practices in Nicaragua. *Ecological Economics* 64: 374–385.
Pereira HM, Reyers B, Watanabe M, *et al.* (2005) Condition and trends of ecosystem services and biodiversity. In: Capistrano D, Samper C, Lee MJ, and Raudsepp-Hearne C (eds.) *Ecosystems and Human Well-Being: Multi Scale Assessments, Vol. 4. Findings of the Sub-global Assessments Working Group of the Millennium Ecosystem Assessment*, pp. 171–203. Washington, DC: Island Press.
Perrot-Maître D (2006) *The Vittel Payments for Ecosystem Services: A "Perfect" PES Case?* London, U.K: International Institute for Environment and Development.
Porras I, Greig-Gran M, and Neves N (2008) All that glitters: A review of payments for watershed services in developing countries. *Natural Resource Issues No. 11*. London, U.K: International Institute for Environment and Development.
Purugganan MD and Fuller DQ (2009) The nature of selection during plant domestication. *Nature* 457: 843–848.
Revenga C and Mock G (1999) *Pilot Analysis of Global Ecosystems: Freshwater Systems and World Resources 1998–1999*. Washington, DC: World Resources Institute.
Ricketts TH, Daily GC, Ehrlich PR, and Michener CD (2004) Economic value of tropical forest to coffee production. *Proceedings of the National Academy of Sciences USA* 11: 12579–12582.
Rodríguez JP, Beard Jr. TD, Agard J, *et al.* (2005) Interactions among ecosystem services. In: Carpenter SR, Pingali PL, Bennett EM, and Zurek MB (eds.) *Ecosystems and Human Well-Being: Scenarios. Vol. 2. Findings of the Scenarios Working Group, Millennium Ecosystem Assessment*, pp. 431–448. Washington, DC: Island Press.
Rodríguez JP, Beard TD, Bennett EM, *et al.* (2006) Trade-offs across space, time, and ecosystem services. *Ecology and Society* 11: 28.
Sala OE, Chapil III FS, Armesto JJ, *et al.* (2000) Global biodiversity scenarios for the Year 2100. *Science* 287: 1770–1774.
Salzman J (2005) Creating markets for ecosystem services: Notes from the field. *New York University Law Review* 80: 870–961.
Sommerville MM, Jones JPG, and Milner-Gulland EJ (2009) A revised conceptual framework for payments for environmental services. *Ecology and Society* 14: 34–47.
Stanton T, Echavarria M, Hamilton K, Ott C (2010) State of watershed payments: An emerging marketplace. Ecosystem Marketplace. http://forest-trends.org/publication_details.php?publicationID=2438
Tallis H, Goldman R, Uhl M, and Brosi B (2009) Integrating conservation and development in the field: Implementing ecosystem service projects. *Frontiers in Ecology and the Environment* 7: 12–20.
Tallis H, Kareiva P, Marvier M, and Chang A (2008) An ecosystem services framework to support both practical conservation and economic development. *Proceedings of the National Academy of Sciences USA* 105: 9457–9464.
Tallis HM and Kareiva P (2006) Shaping global environmental decisions using socio-ecological models. *Trends in Ecology and Evolution* 21: 562–568.
Tallis HT, Ricketts T, Nelson E, *et al.* (eds.) (2010) *InVEST 1.004 beta User's Guide*. Stanford University: The Natural Capital Project.
The Natural Capital Project and The Nature Conservancy (2011) Improving Conservation Investment Returns for People and Nature in the East Cauca Valley. Colombia: The Natural Capital Project. 2011.
Veiga F (2009) Hydrological Services Payment in Brazil. In: *Beyond Carbon: Biodiversity and Water Markets*, pp. 44–48. Ecosystem Marketplace. http://ecosystemmarketplace.com/documents/acrobat/ECM%20Beyond%20Carbon_eng.pdf
Veiga F and Gavaldão M (2011) Iniciativas de PSA de Conservação de Recursos Hídricos na Mata Atlântica. In: Guedes F and Seehusen S (eds.) *PSA na Mata Atlântica: Lições Aprendidas e Desafios*. Brasília: MMA. orgs.
Vitousek PM, Mooney HA, Lubchenco J, and Melillo JM (1997) Human domination of Earth's ecosystems. *Science* 277: 494–499.
Vörösmarty CJ, Green P, Salisbury J, and Lammers RB (2000) Global water resources: Vulnerability form climate change and population growth. *Science* 289: 284–288.
Wunder S (2007) The efficiency of payments for environmental services in tropical conservation. *Conservation Biology* 21: 48–58.

Wetland Creation and Restoration

William J Mitsch, Florida Gulf Coast University, Naples, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Mitigation wetland Wetland created, restored, or enhanced to replace lost wetlands.

Wetland Land that is saturated with water long enough to promote wetland or aquatic processes as indicated by poorly drained soils, hydrophytic vegetation, and various kinds of biological activity, which are adapted to a wet environment.

Wetland bank A site where wetlands and/or other aquatic resources are restored, created, enhanced, or in exceptional circumstances, preserved expressly for the purpose of providing compensatory mitigation in advance of authorized impacts to similar resources.

Wetland creation The conversion of a persistent upland or shallow water area into a wetland by human activity.

Wetland restoration The return of a wetland from a disturbed or altered condition caused by human activity to a previously existing condition.

Introduction

Wetlands have been described as "ecological supermarkets" because of the extensive food chain and rich biodiversity that they support (Mitsch and Gosselink, 2007). They play major roles in the landscape by providing unique habitats for a wide variety of flora and fauna. They support a wide variety of plant, and microbial species where the "critters" come from both terrestrial and deepwater aquatic systems to reproduce, find refuge, eat, or be eaten. More recently, wetlands are being described as important water-quality enhancement ecosystems and flood mitigation systems ("natures' kidneys") and carbon sinks and climate stabilizers on a global scale. These ecosystem services of wetlands are now recognized worldwide and have led to wetland conservation, protection laws, regulations, and management plans. Wetlands have become the *cause célèbre* for conservation-minded people and organizations throughout the world, in part, because they have become symptoms of our systematic dismantling of the world's water resources and in part because their disappearance represents an easily recognizable loss of natural areas to economic "progress." But there is optimism by some that wetlands can be created and restored and that the ecosystem services such as biodiversity protection can be thus restored. Loss rates of wetlands around the world and the subsequent recognition of wetland values have stimulated restoration and creation of these systems (Mitsch and Jørgensen, 2004).

Wetland restoration involves returning a wetland to its original or previous wetland state, whereas wetland creation involves conversion of uplands or shallow open-water systems to vegetated wetlands. Wetland restoration and creation can occur for replacement of habitat, for coastal restoration, and for restoration of mined peatlands. Generally, wetland restoration and creation first involve establishment or re-establishment of appropriate natural hydrologic conditions, followed by establishment of appropriate vegetation communities. Although many of these created and restored wetlands have become functional, there have been some cases of "failure" of created or restored wetlands generally caused by a lack of proper hydrology. Creating and restoring wetlands needs to be based on the concept of self-design whereby any number of native propagules can be introduced, but the ecosystem adapts and changes itself according to its physical constraints, and success should not solely be determined by specific plant and animal presence.

Definitions

Wetlands have many distinguishing features, the most notable of which are the presence of standing water for some period during the growing season, unique soil conditions, and organisms, especially vegetation, adapted to or tolerant of saturated soils. Wetlands are unique because of their hydrologic conditions and their role as ecotones between terrestrial and aquatic systems. Terms such as swamp, marsh, fen, and bog have been used in common speech for centuries to define wetlands and are frequently used and misused at present. Formal definitions have been developed by several federal agencies in the US, by scientists in Canada and the US, and through an international treaty known as the Ramsar Convention. These definitions include considerable detail and are used for both scientific and management purposes.

The most common definition of wetlands used in the world is the one developed as part of the Convention on Wetlands of International Importance Especially as Waterfowl Habitat, better known as the Ramsar Convention:

> areas of marsh, fen, peatland or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish, or salt including areas of marine water, the depth of which at low tide does not exceed 6 m.

This definition, which was adopted at the first meeting of the convention in Ramsar, Iran, in 1971, states that wetlands may incorporate riparian and coastal zones adjacent to the wetlands and islands or bodies of marine water deeper than 6 m at low tide lying within the wetlands. This definition does not include vegetation or soil and extends wetlands to water depths of 6 m or more, well beyond the depth usually

considered wetlands in the US and Canada. The rationale for such a broad definition of wetlands is based on its original intention to include habitats for water birds.

In North America, there are at least five separate word definitions of wetlands (see Mitsch and Gosselink, 2007 for a summary of all). The one that is used to legally protect and define wetlands in the USA is the so-called US Army Corps of Engineering definition:

> The terms "wetlands" means those areas that are inundated or saturated by surface or groundwater at a frequency and duration sufficient to support and that under normal circumstances do support a prevalence of vegetation typically adapted for life in saturated soil conditions. Wetlands generally include swamps, marshes, bogs, and similar areas. (33 CFR 328.3(b); 1984)

Canada has developed a specific national definition of wetlands that may be the best of all:

> land that is saturated with water long enough to promote wetland or aquatic processes as indicated by poorly drained soils, hydrophytic vegetation, and various kinds of biological activity, which are adapted to a wet environment. (Zoltai and Vitt, 1995; Warner and Rubec, 1997)

It is illustrative that the first two definitions above retreat to common terminology used to define wetlands such as "marsh," "fen," "peatland," "swamp," and "bog." **Table 1** lists many of the common terms that are used around the world to define wetlands. These common terms have specific meanings to scientists yet are widely used by the public, often well beyond the generally accepted scientific definitions. In some cases, words such as "swamp" have different meanings in North America and Europe. "Wetland" is a relatively new term in the scientific literature (one of the first uses of it was in a US Fish and Wildlife publication in 1956), but wetland definitions now abound, but there is not one that will prove satisfactory to all users (Mitsch and Gosselink, 2007). Different definitions have been formulated by geologists, soil scientists, hydrologists, botanists, ecologists, economists, political scientists, public health scientists, lawyers, and the public at large. Wetlands are not easily defined, especially for legal purposes, because they have a considerable range of hydrologic conditions, because they are found along a gradient at the margins of well-defined uplands and deepwater systems, and because of their great variation in size, location, and human influence (**Figure 1**). No absolute answer to "What is a wetland?" should be expected, but a combination of the three definitions above is more than sufficient to provide adequate description of wetlands from a structural and functional views.

Global Extent of Wetlands

The extent of the world's wetlands is generally thought to be from 5 to 9 million km^2 or about 4–6% of the land surface of the Earth (**Table 2**). Estimated areas of wetlands in parts of North America are fairly robust with 43.6 million ha in the lower 48 states, 71 million ha in Alaska, and 127 million ha in Canada; these collectively represent about 30% or more of the world's wetlands. The loss of wetlands in the world caused by humans is difficult to determine, but it is probably similar to the 50% loss rate estimated for the lower 48 states of the US, with high rates of loss in Europe and parts of Australia, Canada, and Asia.

Threats to Wetlands

Wetland impacts have included both wetland alteration and wetland destruction. In earlier times, wetland drainage was considered the only policy for managing wetlands in the western world. With over 70% of the world's population living on or near coastlines, coastal wetlands have long been destroyed through a combination of excessive harvesting, hydrologic modification and seawall construction, coastal development, pollution, and other human activities. Likewise, inland wetlands have been continually affected, particularly through hydrologic modification and agricultural and urban development. By 1985, 56–65% of wetlands in North America and Europe, 27% in Asia, 6% in South America, and 2% in Africa had been drained for intensive agriculture. Other human activities such as forestry, stream channelization, aquaculture, dam, dike, and seawall construction, mining, water pollution, and groundwater withdrawal all had impacts, some severe, on wetlands (**Table 3**). Wetlands are degraded and destroyed indirectly as well through alternation of sediment patterns in rivers, hydrologic alternation, highway construction, and land subsidence. Wetlands are also managed close to their natural state for certain objectives such as fish and wildlife enhancement, agricultural and aquaculture production, water-quality improvement, and flood control. Management of wetlands for coastal protection has now taken on more significance.

Wetland Creation and Restoration

The literature on wetland creation and restoration has exploded like no other wetland topic. Mitsch (1994) and Streever (1999) provided early regional overviews and case studies of wetland restoration from around the world. Some general principles and details of techniques are included in the book *Ecological Engineering and Ecosystem Restoration* (Mitsch and Jørgensen, 2004). A critique of the policies and techniques of wetland creation and restoration in the US was published as NRC (2001).

Wetland restoration refers to the return of a wetland from a disturbed or altered condition caused by human activity to a previously existing condition. The wetland may have been degraded or hydrologically altered and restoration then may involve reestablishing hydrologic conditions to re-establish previous vegetation communities. *Wetland creation* refers to the conversion of a persistent upland or shallow water area into a wetland by human activity. *Wetland enhancement* refers to a human activity that increases one or more functions of an existing wetland. One type of created wetland, a *constructed wetland*, refers to a wetland that has been developed for the primary purpose of contaminant or pollution removal from wastewater or runoff. This last type of wetland is also referred to as a *treatment wetland*.

Table 1 Common terms used for various wetland types in the world

Billabong – Australian term for a riparian wetland that is periodically flooded by the adjacent stream or river.
Bog – A peat-accumulating wetland that has no significant inflows or outflows and supports acidophilic mosses, particularly *Sphagnum*.
Bottomland – Lowland along streams and rivers, usually on alluvial floodplains, that is periodically flooded. When forested, it is called a bottomland hardwood forest in the southeastern and eastern US.
Carr – Term used in Europe for forested wetlands characterized by alders (*Alnus*) and willows (*Salix*).
Cumbungi swamp – Cattail (*Typha*) marsh in Australia.
Dambo – A shallow treeless wetland in which water collects in the rainy season; term is used in southern Africa.
Delta – A wetland–river–upland complex located where a river forms distributaries as it merges with the sea; there are also examples of inland deltas such as the Peace–Athabasca Delta in Canada and the Okavango Delta in Botswana.
Fen – A peat-accumulating wetland that receives some drainage from surrounding mineral soil and usually supports marshlike vegetation.
Lagoon – Term frequently used in Europe to denote a deepwater enclosed or partially opened aquatic system, especially in coastal delta regions.
Mangal – Same as mangrove.
Mangrove – Subtropical and tropical coastal ecosystem dominated by halophytic trees, shrubs, and other plants growing in brackish to saline tidal waters. The word ''mangrove'' also refers to the dozens of tree and shrub species that dominate mangrove wetlands.
Marsh – A frequently or continually inundated wetland characterized by emergent herbaceous vegetation adapted to saturated soil conditions. In European terminology, a marsh has a mineral soil substrate and does not accumulate peat. See also Tidal freshwater marsh and salt marsh.
Mire – Synonymous with any peat-accumulating wetland (European definition); from the Norse word ''myrr.'' The Danish and Swedish word for peatland is now ''mose.''
Moor – Synonymous with peatland (European definition). A highmoor is a raised bog; a lowmoor is a peatland in a basin or depression that is not elevated above its perimeter. The primitive sense of the Old Norse root is ''dead'' or barren land.
Muskeg – Large expanse of peatlands or bogs; particularly used in Canada and Alaska.
Oxbow – Abandoned river channel, often developing into a swamp or marsh.
Pakihi – Peatland in southwestern New Zealand dominated by sedges, rushes, ferns, and scattered shrubs. Most pakihi form on terraces or plains of glacial or fluvial outwash origin and are acid and exceedingly infertile.
Peatland – A generic term of any wetland that accumulates partially decayed plant matter (peat).
Playa – An arid- to semiarid-region wetland that has distinct wet and dry seasons. Term used in the southwest US for marshlike ponds similar to potholes, but with a different geologic origin.
Pocosin – Peat-accumulating, nonriparian freshwater wetland, generally dominated by evergreen shrubs and trees and found on the southeastern Coastal Plain of the US. The term comes from the Algonquin for ''swamp on a hill.''
Pothole – Shallow marshlike pond, particularly as found in the Dakotas and central Canadian provinces, the so-called prairie pothole region.
Raupo swamp – Cattail (*Typha*) marsh in New Zealand.
Reedmace swamp – Cattail (*Typha*) marsh in the UK.
Reedswamp – Marsh dominated by *Phragmites* (common reed); term used particularly in Europe.
Riparian ecosystem – Ecosystem with a high water table because of proximity to an aquatic ecosystem, usually a stream or river. Also called bottomland hardwood forest, floodplain forest, bosque, riparian buffer, and streamside vegetation strip.
Salt marsh – A halophytic grassland on alluvial sediments bordering saline water bodies where water level fluctuates either tidally or nontidally.
Sedge meadow – Very shallow wetland dominated by several species of sedges (e.g., *Carex, Scirpus*, and *Cyperus*).
Slough – An elongated swamp or shallow lake system, often adjacent to a river or stream. A slowly flowing shallow swamp or marsh in the southeastern US (e.g., cypress slough). From the Old English word ''sloh'' meaning a watercourse running in a hollow.
Swamp – Wetland dominated by trees or shrubs (US definition). In Europe, forested fens and wetlands dominated by reed grass (*Phragmites*) are also called swamps (see Reedswamp).
Tidal freshwater marsh – Marsh along rivers and estuaries close enough to the coastline to experience significant tides by nonsaline water. Vegetation is often similar to nontidal freshwater marshes.
Turlough – Areas seasonally flooded by karst groundwater with sufficient frequency and duration to produce wetland characteristics. They generally flood in winter and are dry in summer and fill and empty through underground passages. Term is specific for these types of wetlands found mostly in western Ireland.
Vernal pool – Shallow, intermittently flooded wet meadow, generally typical of Mediterranean climate with dry season for most of the summer and fall. Term is now used to indicate wetlands temporarily flooded in the spring throughout the US.
Vleis – Seasonal wetland; term used in southern Africa.
Wad (pl. Wadden) – Unvegetated tidal flat originally referring to the northern Netherlands and northwestern German coastline. Now used throughout the world for coastal areas.
Wet meadow – Grassland with waterlogged soil near the surface but without standing water for most of the year.
Wet prairie – Similar to a marsh but with water levels usually intermediate between a marsh and a wet meadow.

Source: Reproduced from Mitsch WJ and Gosselink JG (2007) *Wetlands*, 4th edn., 582 pp. New York: John Wiley & Sons, Inc.

Significant efforts now focus on the voluntary restoration and creation of wetlands. Part of the interest in wetland creation and restoration stems from the fact that we are losing or have lost much of this valuable habitat. Often interest is less voluntary and more in response to government policies such as ''no net loss'' in the US that require the replacement of wetlands for those unavoidably lost. New Zealand, which has lost 90% of its wetlands, has made major efforts to restore marshes and other wetlands in the Waikato River Basin on North Island and in the vicinity of Christchurch on South Island. In southeastern Australia, restoration of the Murray–Darling watersheds, particularly the riverine

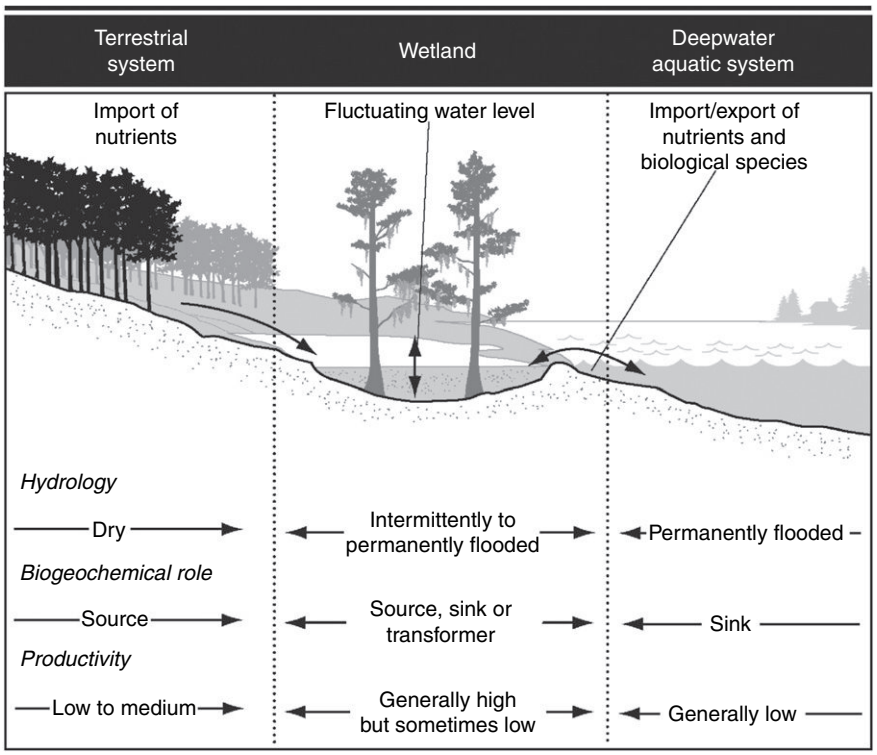


Figure 1 Position and function of wetlands in the landscape between terrestrial and deepwater aquatic ecosystems. Reprinted from Mitsch WJ and Gosselink JG (2007) *Wetlands*, 4th edn., 582 pp. New York: John Wiley & Sons, Inc.

Table 2 Estimated area of wetland ecosystem types in the world

Type of wetland	*Area, $\times 10^6$ ha*
Coastal wetlands	
Tidal salt marshes	10*
Tidal freshwater marshes	2*
Mangrove wetlands	24
Inland wetlands	
Freshwater marshes	95
Freshwater swamps and riparian forests	109
Peatlands	350
Total	580

*Number estimated.

Source: Reproduced from Mitsch WJ, Gosselink JG, Anderson CJ, and Zhang L (2009) *Wetland Ecosystems*, 295 pp. New York: John Wiley & Sons, Inc.

billabongs, has become a major undertaking, while coastal plain wetland restoration and creation are occurring in southwestern Australia. There are concerted efforts to restore mangrove forests in the Mekong Delta of Vietnam, along South American coastlines where shrimp farming has destroyed thousands of hectares of mangroves and around the Indian Ocean to provide tsunami and typhoon protection for coastal areas. Tidal marshes have been created along much of China's eastern coastline, and restoration is now occurring in the Yangtze Delta in Shanghai. Wetland restoration and creation are being proposed or implemented on very large scales to prevent more deterioration of existing wetlands (Everglades in Florida), to mitigate the loss of fisheries (Delaware Bay in Eastern USA), to reduce land loss and provide protection from hurricanes (Mississippi Delta in Louisiana), to stabilize a watershed and provide water-quality improvement (Skjern River, Denmark), and to solve serious cases of overenrichment of coastal waters (Baltic Sea in Scandinavia; Gulf of Mexico in USA).

Mitigating Wetland Habitat Loss

Wetland protection regulations in the US and now elsewhere have led to the practice of requiring that wetlands be created, restored, or enhanced to replace wetlands lost in developments such as highway construction, coastal drainage and filling, or commercial development. This is referred to as the process of "mitigating" the original loss, and these "new" wetlands are often called *mitigation wetlands* or *replacement wetlands*. Replacement wetlands are designed to be at least of the same size as the lost wetlands, but more often a *mitigation ratio* is applied so that more wetlands are created and/or restored than are lost. For example, a mitigation ratio of 2:1 means that 2 ha of wetlands will be restored or created for every hectare of wetland lost to development. Considerable controversy exists, for example, in the US, on the question as to whether wetland loss can be mitigated successfully or if it is essentially impossible (NRC, 2001). Robb (2002) reviewed several years' efforts of mitigating wetland loss in Indiana and suggested, based on failure rates of various wetland types, that there should be the following mitigation ratios: 3.5:1 for forested wetlands; 7.6:1 for wet meadows, 1.2:1 for freshwater marshes, and 1:1 for open-water systems.

Table 3 Human actions that cause direct and indirect wetland losses and degradation in the world

Cause	*Estuaries*	*Floodplains*	*Freshwater marshes*	*Lakes/Littoral zone*	*Peatlands*	*Swamp Forest*
Cause direct wetland losses and degradation						
Agriculture, forestry, mosquito control drainage	xx	xx	xx	x	xx	xx
Stream channelization and dredging; flood control	x		x			
Filling – solid-waste disposal; roads; development	xx	xx	xx	x		
Conversion to aquaculture/mariculture	xx					
Dikes, dams, seawall, levee construction	xx	x	x	x		
Water pollution – urban and agricultural	xx	xx	xx	xx		
Mining of wetlands of peat and other materials	x	x		xx	xx	xx
Groundwater withdrawal		x	xx			
Cause indirect wetland losses and degradation						
Sediment retention by dams and other structures	xx	xx	xx			
Hydrologic alteration by roads, canals, etc.	xx	xx	xx	xx		
Land subsidence due to groundwater, resource extraction, and river alternations	xx	xx	xx			

xx, common and important cause of wetland loss and degradation; x, present but not a major cause of wetland loss and degradation. Blank indicates that effect has not been validated.

Source: Reproduced from Dugan P (1993) *Wetlands in Danger*, 192 pp. London: Reed International Books.

One of the most interesting strategies that the private sector and government agencies have developed to deal with the piecemeal approach to mitigation of wetland loss is the concept of a *mitigation bank*. A mitigation bank is defined as "a site where wetlands and/or other aquatic resources are restored, created, enhanced, or, in exceptional circumstances, preserved expressly for the purpose of providing compensatory mitigation in advance of authorized impacts to similar resources" (*Federal Register*, 28 November 1995, "Federal Guidance for the Establishment, Use, and Operation of Mitigation Banks"). In this approach, wetlands are built in advance of development activities that cause wetland loss, and credits of wetland area can be sold to those who are in need of mitigation for wetland loss. Banks are seen as a way of streamlining the process of mitigating wetland loss and, in many cases, providing a large, fully functional wetland rather than small, questionable wetlands near the site of wetland loss.

Wetland Conservation Programs in Agricultural Lands

Conservation programs are now in place to encourage individual farmers in the US to restore wetlands on their land. Both the Conservation Reserve Program (CRP) and the Wetlands Reserve Program (WRP) under the US Department of Agriculture have led to significant areas of wetlands being restored or protected. CRP guidelines, announced in 1997, give increased emphasis to the enrollment and restoration of *cropped wetlands*, that is, wetlands that produce crops but serve wetland functions when crops are not being grown. The CRP also encourages wetland restoration, particularly through hydrologic restoration. In the CRP, participants voluntarily enter into contracts with the US Department of Agriculture to enroll erosion-prone and other environmentally sensitive land in long-term contracts for 10–15 years. In exchange, participants receive annual rental payments and a payment of up to 50% of the cost of establishing conservation practices. The WRP is another voluntary program established in 1990 and is specific for wetland restoration; it offers landowners the opportunity to protect, restore, and enhance wetlands on their property and provides funds for the farmer to do so.

Forested Wetland Restoration

There is less experience at forested wetland restoration and creation compared to herbaceous marshes, despite the fact that these wetlands have been lost at alarming rates. Forested wetland creation and restoration are different from marsh creation and restoration because forest regeneration takes decades rather than years to complete, and there is more uncertainty about the results. Much riparian forest restoration in the US has centered on the lower Mississippi River alluvial valley, where more than 78,000 ha were reforested by federal agencies over the 10-year period 1988–1997 (King and Keeland, 1999), primarily with bottomland hardwood species and, to a lesser extent, deepwater swamp species. This is a small contribution to the restoration of this alluvial floodplain where 7.2 million ha of bottomland hardwood forest were estimated to have been lost (Hefner and Brown, 1985).

Hydrologic and Water Quality Restoration of Watersheds

Lines often blur between wetlands created and restored for habitat restoration and those restored for water quality and hydrology improvement. In fact most wetlands that are restored or created are done so for both habitat recreation and for other ecosystem services. Four large freshwater wetland restorations around the world are presented as case studies here. One of the largest wetland restoration projects in the world is the Florida Everglades where an attempt is made to restore, at least to some degree, the natural hydrologic conditions in the Everglades that are left (see Case Study 1). Another example of a proposed watershed restoration is a large-scale wetland and riparian forest restoration and

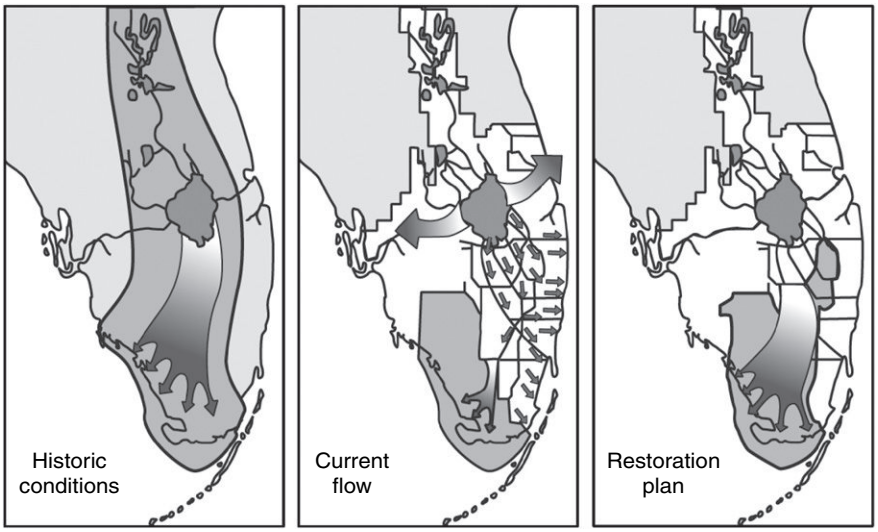


Figure 2 Florida Everglades in historic conditions, current conditions, and planned restoration conditions. Reprinted from Mitsch WJ and Jørgensen SE (2004) *Ecological Engineering and Ecosystem Restoration*, 411 pp. New York: John Wiley & Sons, Inc.

creation, on the order of millions of hectares, being proposed to help solve a major coastal pollution problem in the Gulf of Mexico (see Case Study 2). Restoration done for water quality improvement also has major advantages of also providing habitat restoration and flood mitigation in addition to water quality improvement. The Skjern River Restoration Project in Europe provides a major river/riparian wetland restoration project in a previously drained landscape in Denmark's largest watershed (Case Study 3), while a fourth example of a hydrologic restoration at a culturally significant location in the world that is well underway is the restoration of the Mesopotamian Marshlands of Iraq (Case Study 4).

Case Study 1 – Restoring the Florida Everglades

The restoration of the Florida Everglades, the largest wetland area in the US, actually involves several separate initiatives being carried out in the 4.6 million ha Kissimmee–Okeechobee–Everglades (KOE) region in the southern third of Florida (**Figure 2**). Overall, the Everglades restoration, as now planned by the US Army Corps of Engineers and the South Florida Water Management District, will cost over $12 billion and will be carried out over the next 20 years or more. The basic plan involves restoring something closer to the original hydrology of the KOE region, by sending less of the water from the upper watershed to the Calosahatchee River to the west and the St Lucie Canal to the east (see **Figure 2**, middle diagram) and directing more of the water to the Everglades itself south of Lake Okeechobee (see **Figure 2**, right). Specific problems in the Everglades have developed because of (1) excessive nutrient loading to Lake Okeechobee and to the Everglades itself, primarily from agricultural runoff, (2) loss and fragmentation of habitat caused by urban and agricultural development, (3) spread of *Typha* and other invasives and exotics to the Everglades, replacing native vegetation, and (4) hydrologic alteration due to an extensive canal and straightened rivers system built by the US Army Corps of Engineers and others for flood protection and maintained by several water management districts.

One major restoration project in the KOE region that has received a lot of attention is the restoration of the Kissimmee River. As a result of the channelization of the river in the 1960s, a 166-km-long river was transformed into a 90-km-long, 100-m-wide canal, and the extent of wetlands along the river decreased by 65%. The restoration of the Kissimmee River is a major undertaking to reintroduce the sinuosity to the artificially straightened river. The river restoration work, expected to be completed in stages over the next several decades, will return some portion of lost wetland habitat to the riparian zone and will also provide sinks for nutrients that are otherwise causing increased eutrophication in downstream Lake Okeechobee.

Everglades restoration also involves halting the spread of high-nutrient cattail (*Typha domingensis*) through the low-nutrient sawgrass (*Cladium jamaicense*) communities that presently dominate the Everglades. Since the main causes of the spread of cattails are nutrients and especially phosphorus emanating from agricultural areas in the basin, 16,000 ha of created wetlands, called Stormwater Treatment Areas (STAs) have been created for phosphorus control from the agricultural area (**Figure 3**). A prototype of the STAs, a 1500 ha site, called the Everglades Nutrient Removal (ENR) project has operated since mid-1994 (Reddy *et al.*, 2006).

Case Study 2 – Creating and Restoring Wetlands to Solve the Gulf of Mexico Hypoxia

A hypoxic zone has developed off the shore of Louisiana in the Gulf of Mexico where hypolimnetic waters with dissolved oxygen less than 2 mg l^{-1} O_2 now extend over an area of 1.6–2.0 million ha (Rabalais *et al.*, 2001). Nitrogen, particularly nitrate-nitrogen, is the most probable cause; 80% of the nitrogen input is from the 3 million km^2 Mississippi River basin. The basin represents 41% of lower 48 states of the USA.

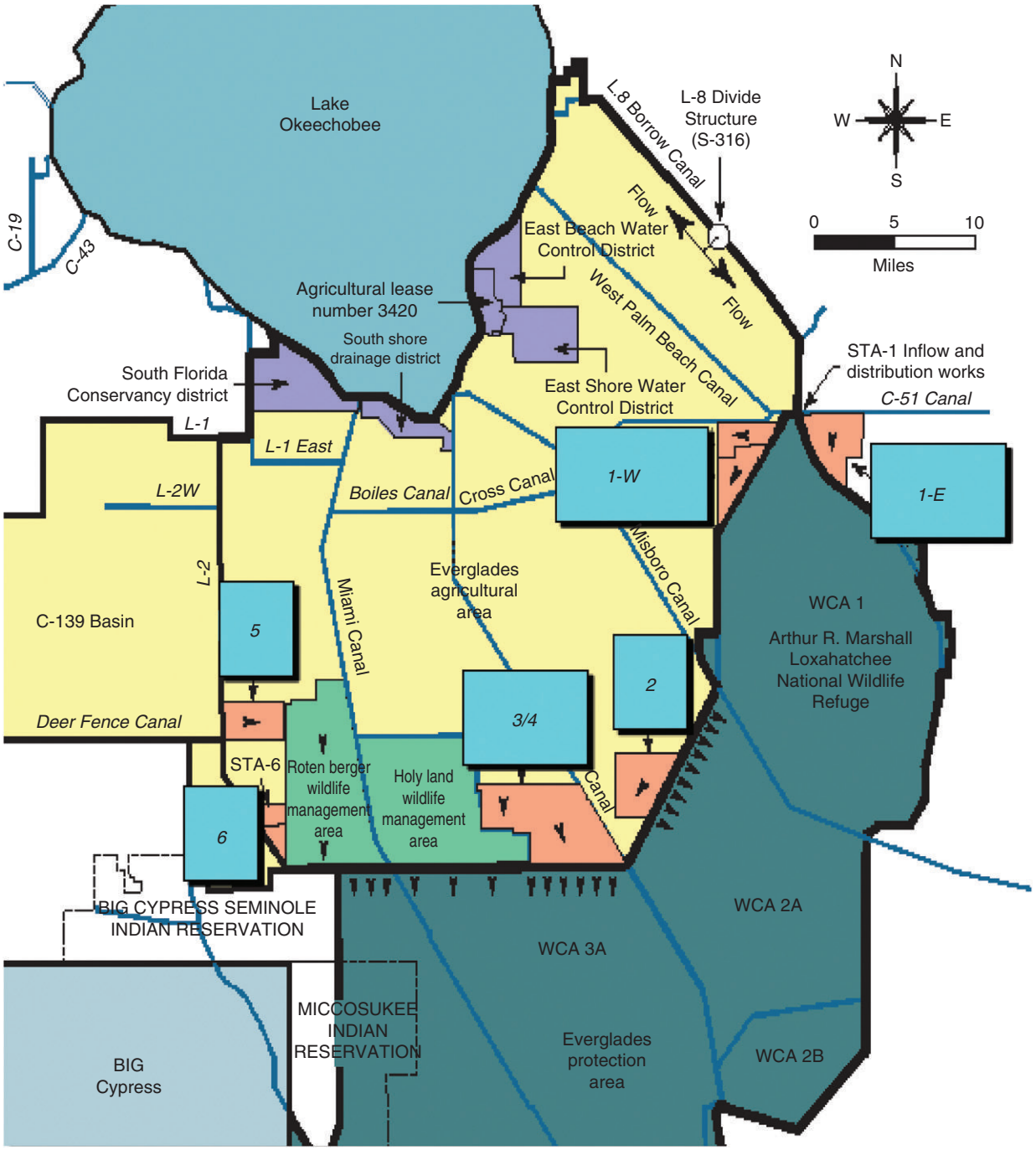


Figure 3 Map of the location of 16,000 ha of Stormwater Treatement Areas (STAs) created north of the Florida Everglades to reduce phosphorus pollution from the Everglades Agricultural Area. Source: South Florida Water Management District, West Palm Beach, Florida.

The control of this hypoxia is important in the Gulf of Mexico because the continental shelf fishery in the Gulf is approximately 25% of the US total. A number of options were investigated for controlling nutrient flow into the Gulf (Mitsch *et al.*, 2001). In the end, there are three general approaches that involve either revision of agronomic approaches or wetland creation and riparian restoration that make the most sense (**Figure 4**). Two million ha of restored and created wetlands and restored riparian buffers have been recommended as necessary to provide enough nitrogen retention to substantially reduce the nitrogen entering the Gulf of Mexico (Mitsch *et al.*, 2001; Mitsch and Day, 2006). That area is less than 1% of the Mississippi River Basin. Interestingly, Hey and Phillipi (1995) found that a similar scale of wetland restoration would be required in the Upper Mississippi River Basin to mitigate the effects of very large and costly floods such as the one that occurred in the summer of 1993 in the Upper Mississippi River Basin.

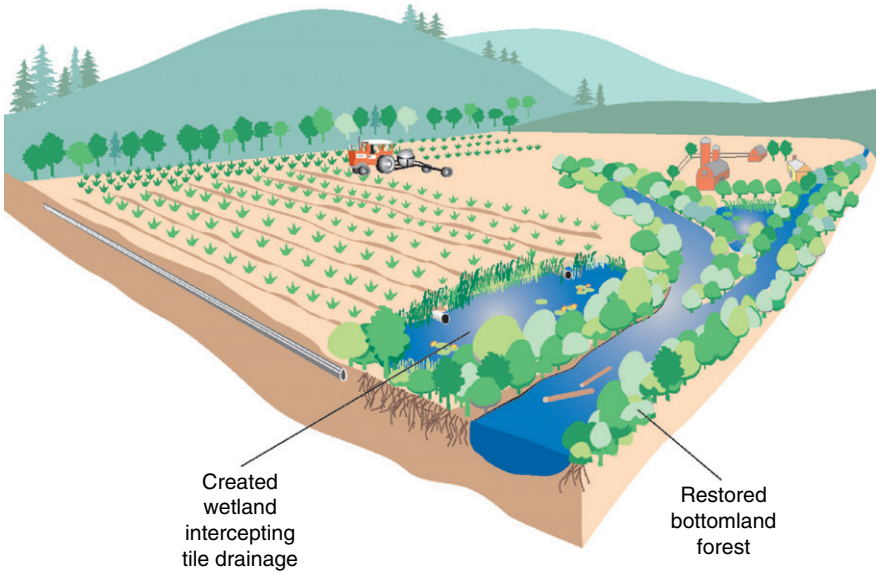


Figure 4 Schematic of agricultural landscape in Mississippi River Basin with wetlands and riparian zones for controlling nitrogen pollution from agricultural fields. Reprinted from Mitsch WJ and Gosselink JG (2007) *Wetlands*, 4th edn., 582 pp. New York: John Wiley & Sons, Inc.

Case Study 3 – The Skjern River Restoration Project, Denmark

River restoration has flourished since the early 1990s in Denmark. The Skjern River in west central Jutland drains more than 11% of Jutland, and the flow of water in the Skjern is the largest in Denmark. In Denmark's largest drainage project ever, 4000 ha of wet meadow was converted into arable land, and the lower Skjern was straightened to a fraction of its former meandering self. By the late 1980s, the river was essentially a straight line to the Ringkobing Fjord on the North Sea, eliminating thousands of hectares of marshland, meadows, and river habitat. The channelized river was diked, canals were built, and pumps were installed to hasten the downstream movement of water from the land. This public works project cost DDK 30 million (about US$3.6 million) and was considered a success by the agricultural community at first as grains could now be grown in the formerly wet region. But the environment was paying a heavy price with this artificial river. The self-cleansing ability of the river was lessened, the downstream fjord was becoming polluted with nutrients and sediments, and the land that was draining began to subside due to peat oxidation and loss of water – up to 1 m or more in some locations. A few years after the drainage, it appeared that another drainage project might be necessary, but the Danish Parliament (Folketing) passed a Public Works Act in 1998 with a huge majority that called for the restoration of the lower Skjern River and earmarked about DDK 254 million for this project. The project is being implemented in three phases for three reaches of the river. Nineteen kilometers of the river and 2200 ha of the river valley wetlands were restored over 1998–2002 (**Figure 5**) by putting back the meanders of the river wherever possible, removing dikes along the river to allow adjacent meadows to once again be flooded, and moving the dikes far away from the river to prevent flooding of farmland outside of the project area. The project was successful in substantially increasing the biodiversity of the region with aquatic macrophytes, invertebrates, amphibians, and mammals such as otters (Pedersen *et al.*, 2007).

Case Study 4 – Restoration of the Mesopotamian Marshlands

The Mesopotamian Marshlands of southern Iraq and Iran, found at the confluence of the historic Tigris and Euphrates Rivers, were 15,000–20,000 km^2 in area as recently as the early 1970s but were drained and diked, especially in the 1990s, to less than 10% of that extent by 2000. Among the main causes are upstream dams and drainage systems constructed in the 1980s and 1990s that altered the river flows and eliminated the flood pulses that sustained the wetlands. Since the overthrow of Saddam Hussein's dictatorship in 2003 in Iraq, there has been a concerted effort by the Iraqis and the international community at restoring the marshlands (Richardson *et al.*, 2005). The restoration has often occurred with local residents breaking dikes or removing impediments to flooding. Remote sensing images showed that at least 37% of the wetlands were restored by 2005, and *Frontiers in Ecology and the Environment* (Vol. 3, No. 8, October 2005) reported that year "at least 74 species of migratory waterfowl and many endemic birds have been sighted in a survey of Iraq's marshland… ." It was also reported that as many as 90,000 Marsh Arabs have returned to the wetlands already (Azzam Alwash, personal communication, 2006). Alwash, the Director of the Eden Again effort, has suggested that perhaps as much as 75% of the marshlands will be restored. There are still several questions that remain unanswered about whether full restoration can occur, including whether adequate water supplies exist given the competition from Turkey, Syria, and Iran, and within Iraq itself, and whether landscape connectivity of the marshes can be re-established (Richardson and Hussain, 2006).

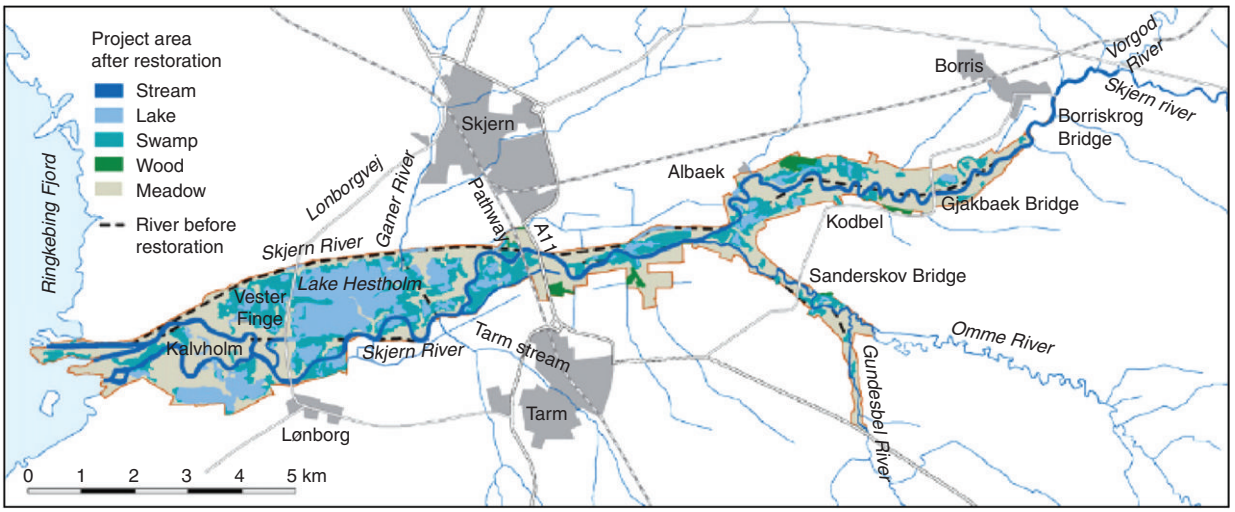


Figure 5 Skjern River restoration area in western Denmark illustrating restoration and floodplain (meadow, lake, and swamp) restoration and river before restoration. Reproduced from Pedersen ML, Andersen JM, Nielsen K, and Linnemann M (2007) Restoration of the Skjern River and its valley: Project description and general ecological changes in the project area. *Ecological Engineering* 30: 131–144.

Peatland Restoration

Peatland restoration is a relatively new type of wetland restoration compared to other types and potentially could be the most difficult (Gorham and Rochefort, 2003). Early attempts with peatlands occurred in Europe, specifically in Finland, Germany, UK, and The Netherlands. Increased peat mining in Canada and elsewhere has led to increased interest in understanding if and how mined peatlands can be restored. When peat surface mines are abandoned without restoration, the area rarely returns through secondary succession to the original moss-dominated system (Quinty and Rochefort, 1997). There is promise that restoration can be successful but because (1) surface mining causes major changes in local hydrology and (2) peat accumulates at an exceedingly slow rate, restoration progress will be measured in decades rather than years. In the 1960s and 1970s, block harvesting of peat was replaced by vacuum harvesting in southern Quebec and in New Brunswick, necessitating the development of different restoration techniques. While traditional block-cutting of peat left a variable landscape of high ground and trenches, vacuum harvesting left relatively flat surfaces bordered by drainage ditches. Abandoned block-cut sites appear to revegetate with peatland species more easily than do vacuum-harvested sites, and the latter can remain bare for a decade or more after mining (Rochefort and Campeau, 1997).

Case Study 5 – Bois-des-Bel Peatland, Québec, Canada

Despite the vast expanses of peatlands in the world, whole-ecosystem experiments on this type of wetland are rare. Bois-des-Bel peatland, located about 200 km northeast of Québec City, on the southern shore of the St Lawrence River in Québec, Canada, is a whole-ecosystem research site where scientists are evaluating the pace of peatland restoration after peat mining (Rochefort *et al.*, 2003). The entire peatland is about 210 ha; the research area is about 11.5 ha of peatland that was drained in 1972 and mined by a vacuum extraction technique from 1973 to 1980. When mining stopped, a 2-m peat deposit remained. Restoration began in 1999 on 8.4 ha of the site, with the remaining as an unrestored control. The restored area was divided into four zones, each of which has two shallow pools (13 m × 5 m × 1.5 m max depth) for aquatic and amphibian habitat. Line Rochefort and her students and colleagues at Université Laval, Quebec City, and at other Canadian universities have established the site as a long-term ecosystem research site here to investigate the revegetation of mined peatlands (Price *et al.*, 1998; Rochefort *et al.*, 2003; Waddington *et al.*, 2003, 2008; Isselin-Nondedeu *et al.*, 2007). The restoration involved terracing to produce better water distribution, reintroduction of *Sphagnum* diaspores harvested from a nearby natural wetland, and reflooding by blocking drainage ditches. The moss carpet increased by about 12 cm by 2007 and was three times the thickness that it was in 2003. *Sphagnum* cover by 2005 was 60% of the area in the restored sites compared to only 0.25% in the nonrestored sites (Isselin-Nondedeu *et al.*, 2007). The restored sites also exported less than half the dissolved organic carbon than did the cutover peatland sites (Waddington *et al.*, 2008).

Coastal Wetland Restoration

There is a great deal of interest in coastal wetland restoration. Early pioneering work on salt marsh restoration was done in Europe, China, North Carolina, Chesapeake Bay and Delaware Bay in the USA, and along the coastlines of Florida, Puerto Rico, and California. Some of this coastal wetland restoration has been undertaken for habitat development as mitigation for coastal development projects. For coastal salt marshes in Eastern US, the cordgrass *Spartina alterniflora* is the primary choice for coastal marsh restoration; but the same species is considered an invasive and unwanted plant on the west coast of North America. Both *Spartina townsendii* and *S. anglica* have

been used to restore salt marshes in Europe and in China although both species are considered to be invasive by some in China and New Zealand. Salt marsh grasses tend to distribute easily through seed dispersal, and the spread of these grasses can be quite rapid once the reintroduction has begun, as long as the area being revegetated is intertidal, that is the elevation is between ordinary high tide and low tide.

Restoring mangrove swamps in tropical regions of the world has some similar characteristics to restoring salt marshes in that the establishment of vegetation in its proper intertidal zone is the key to success. But that is generally where the similarities end. Mangrove restoration is more cosmopolitan in that it has been attempted throughout the tropical and subtropical world (Lewis, 2005); salt marsh restoration has been attempted primarily on the Eastern North American and Chinese coastlines and to some extent in Europe and the west coast of North America. Salt marsh restoration can often rely on water-borne seeds distributing through an intertidal zone; mangrove restoration often involves the physical planting of trees. In countries such as Vietnam, mangrove declines have been attributed to the spraying of herbicides during the Vietnam war, and immigration of people to the coastal regions, leading to cutting of lumber for timber, fuel, wood, and charcoal.

Mangroves are being cleared for construction of aquaculture ponds at unprecedented rates in Vietnam and many other tropical coastlines of the world (Benthem *et al.*, 1999). Most of the edible shrimp sold in the US and Japan are produced in artificial ponds constructed in mangrove wetlands in Thailand, Indonesia, and Vietnam. These products are the result of massive destruction of mangrove forests and are sold in the US and Japan at very low prices. More than 100,000 ha of abandoned ponds located in former mangrove swamps currently exist in these countries (R. Lewis, personal communication). In Vietnam, mangroves are being restored and protected to provide coastal protection and coastal fisheries support. In the Philippines, despite a presidential proclamation in 1981 prohibiting the cutting of mangroves, it is estimated that the country was still losing 3000 ha yr^{-1} in the late 1990s (2.4% yr^{-1}; de Leon and White, 1999). Shrimp ponds last only about 5–6 years before they develop toxic levels of sulfur and are then abandoned, and more mangroves are destroyed. These abandoned ponds present a challenge for mangrove restoration.

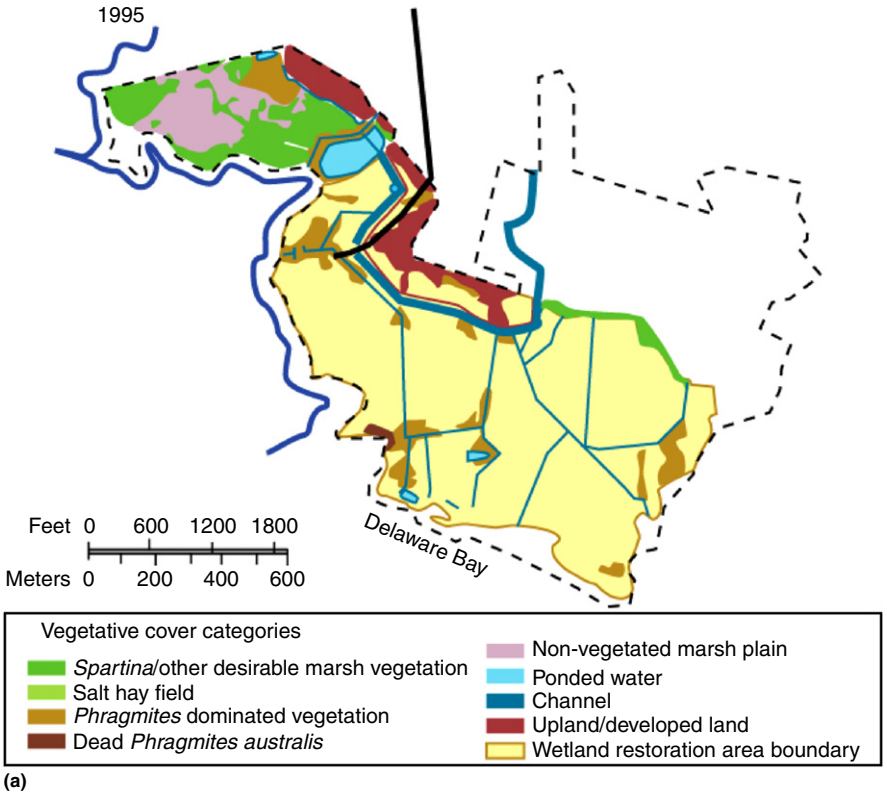
Several studies emphasized the importance of restoring tidal conditions, including salinity, to marsh or mangove areas that had become more "freshwater" because of isolation from the sea. In cases such as this, the restoration is simple – remove whatever impediment is blocking tidal exchange. Case Study 6 describes a salt marsh restoration in Delware Bay where that is exactly the case – restore the natural tidal hydrology, and the vegetation and aquatic species will follow.

Case Study 6 – Delaware Bay Salt Marsh Restoration

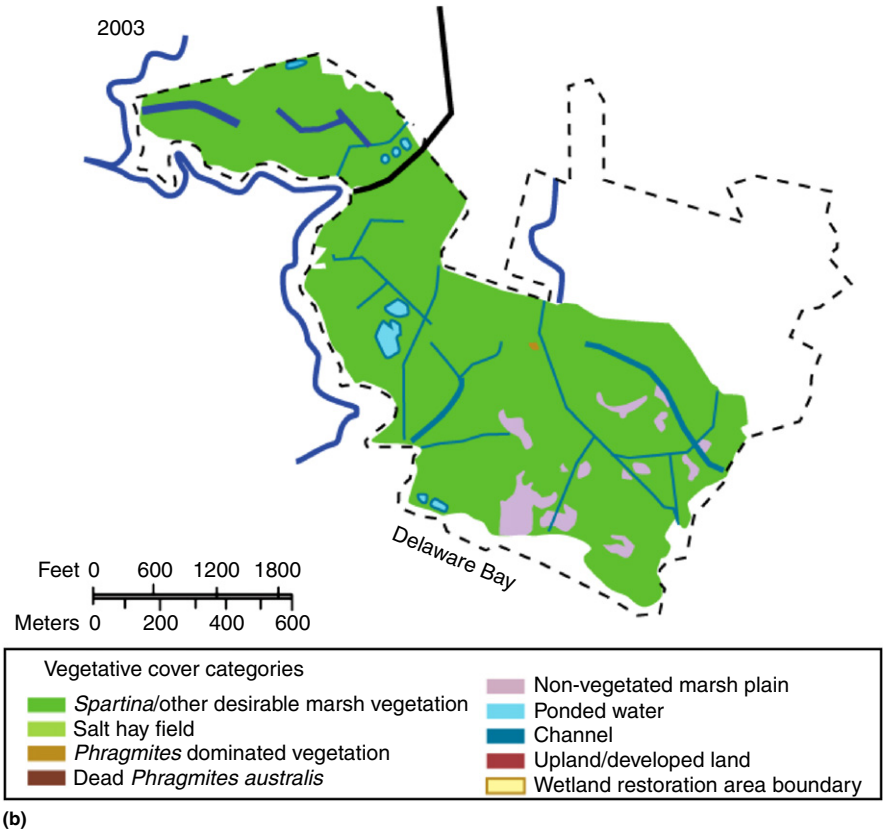
A large coastal wetland restoration project in Eastern USA involves the restoration, enhancement, and preservation of 5000 ha of coastal salt marshes on Delaware Bay in New Jersey and Delaware in northeastern USA (Teal and Peterson, 2005). This estuary enhancement, being carried out by New Jersey's electric utility with advice from a team of scientists and consultants, was undertaken as mitigation for the potential impacts of once-through cooling from a nuclear power plant operated on the Bay. The reasoning was that the impact of once-through cooling on fin fish, through entrainment and impingement, could be offset by increased fisheries production from restored salt marshes. Because of uncertainties involved in this kind of ecological trading, the area of restoration was estimated as the salt marshes that would be necessary to compensate for the impacts of the power plant on fin fish times a safety factor of four. There are two distinct approaches being utilized in this project to restore the Delaware Bay salt marshes:

1. Reintroduce flooding: The most important type of restoration involves the reintroduction of tidal inundation to about 1800 ha of former diked salt hay farms. Many marshes along Delaware Bay have been isolated by dikes from the bay, sometimes for centuries, and put into the commercial production of "salt-hay" (*Spartina patens*). Hydrologic restoration was accomplished by excavating breaches in the dikes and, in most cases, connecting these new inlets to a system of recreated tidal creeks and existing canal systems. From a hydrodynamic perspective, in those marshes where tidal exchange was restored, the development of an intricate tidal creek density from the originally constructed tidal creeks has been impressive. The "order" of the stream channels increased from 5 or less to well over 20 from 1996 through 2004. The number of small tributaries increased from "dozens" to "hundreds" at all three salt hay farm sites that were reopened to tidal flushing. Hydrologic design did occur in "self-design" fashion after only initial cuts by construction of the first-order channels.
2. Reduce *Phragmites* domination: In another set of restoration sites in Delaware and New Jersey, restoration involves the reduction in cover of the aggressive and invasive *Phragmites australis* in 2100 ha of nonimpounded coastal wetlands. Alternatives that were investigated include hydrological modifications such as channel excavation, breaching remnant dikes, microtopographic changes, mowing, planting, and herbicide application.

Results of this study were reported in several early journal articles (e.g., Weinstein *et al.*, 1997, 2001; Teal and Weinstein, 2002) and more recently in a journal special issue of 10 papers related to this salt marsh restoration (Peterson *et al.*, 2005). For the salt hay farms that were flooded, typical goals include a high percent cover of desirable vegetation such as *Spartina alterniflora*, a relatively low percent of open water, and the absence of the invasive reed grass *Phragmites australis*. Reestablishment of *Spartina alterniflora* and other favorable vegetation has been rapid and extensive. Approximately 64% of one site was dominated by *Spartina alterniflora* after only two growing seasons and 87% by the fifth year after construction (**Figure 6**). Tidal restoration was completed at a larger site and major revegetation by *Spartina alterniflora* and some *Salicornia* already occurred with 71% of the site showing desirable vegetation after only four growing seasons. The study has



(a)



(b)

Figure 6 Restoration of Dennis Township salt marsh on Delaware Bay, New Jersey, USA, from initial conditions in 1995 to restored salt marsh in 2003. Reproduced from Hinkle R and Mitsch WJ (2005) Salt marsh vegetation recovery at salt hay farm wetland restoration sites on Delaware Bay. *Ecological Engineering* 25: 240–251.

shown that the speed with which restoration takes place is dependent on three main factors: the degree to which the tidal "circulatory system" works its way through the marsh; the size of the site being restored; and the initial presence of *Spartina* and other desirable species. No planting was necessary on these sites as *Spartina* seeds arrive by tidal fluxes; but the design of the sites to allow that tidal connectivity (and hence the importance of appropriate site elevations relative to tides) was critical. Self-design works when the proper conditions for propagule disbursement are provided.

River Delta Restoration

As large rivers connect to the sea, multitributary deltas tend to develop, allowing the river to discharge to the sea in many channels. Many of these rich-soil deltas are among the most important ecological and economic regions of the world, from the ancient Nile delta in Egypt to the modern-day Mississippi River delta in Louisiana. There should be two major ecological resource goals of delta areas: (1) protecting and restoring the functioning of the deltaic ecosystems in the context of a geologically dynamic framework and (2) controlling pollution from entering the downstream lakes, oceans, gulfs, and bays. Delta restoration should have this dual emphasis where possible – ecosystem enhancement of the delta itself and improvement of coastal water-quality downstream. The best strategy for delta restoration when "land building" is a prerequisite (see Louisiana case study below) is to restore the ability of the river to "spread out its sediments" in deltaic form as wide an area as possible, particularly during flood events and by not discouraging (or encouraging and even creating) river distributaries. When river distributaries are not possible on a large scale due to navigation requirements or population locations, then restoring and creating riverine wetlands and constructing river diversions to divert river water to adjacent lands may be the best alternatives to maximize nutrient retention and sediment retention. In some cases, this involves the conversion of agricultural lands back to wetlands; in other cases, the dikes that "protect" wildlife protection ponds or retain rivers in their channels only need to be carefully breached to allow lateral flow of rivers during flood season.

Case Study 7 – Louisiana Delta Restoration

Louisiana is one of the most wetland-rich regions of the world with 36,000 km^2 of marshes, swamps, and shallow lakes. Yet Louisiana is suffering a rate of coastal wetland loss of 6600–10,000 ha yr^{-1} as it converts to open water areas on the coastline, due to natural (land subsidence) and human causes such as river levee construction, oil and gas exploration, urban development, sediment diversion, and possibly climate change. There has been, since the early 1990s, a major interest in reversing this rate of loss and even gaining coastal areas, particularly freshwater marshes and salt marshes, the loss of which are the major symptom of this "land loss." Clearly since the disaster in Louisiana and New Orleans caused by Hurricane Katrina in 2005, there is intense interest in restoring the Louisiana delta. But the enormous cost of reestablishing human settlements and putting back levees that were breeched during the hurricane has led some to doubt if there are enough resources to carry out the needed wetland restoration as well (Costanza *et al.*, 2006). Yet the wetlands are vital to the long-term survival of New Orleans. River diversions are becoming a large part of the delta restoration in Louisiana. The deltaic wetlands in Louisiana cover more than 20,000 km^2 and are critical to building land and wetlands from open water in the delta and may be critical nutrient removal sites for abatement of Gulf hypoxia (see Case Study 2 above). The Caernarvon freshwater diversion is one of the largest of several diversions currently in operation on the Mississippi River in Louisiana. The diversion structure is on the east bank of the river below New Orleans 131 km upstream of the Gulf of Mexico. The structure is a five-box culvert with vertical lift gates with a maximum flow of 226 $m^3 s^{-1}$. Freshwater discharge began in August 1991 and discharge from then until December 1993 averaged 21 $m^3 s^{-1}$; current minimum and maximum flows are 14 and 114 $m^3 s^{-1}$, respectively, with summer flow rates generally near the minimum and winter flow rates 50–80% of the maximum (Lane *et al.*, 1999). The Caernarvon diversion delivers water to the Breton Sound estuary, a 1100 km^2 area of fresh, brackish, and saline wetlands.

Creating and Maintaining the Proper Hydrology

The key to restoring and creating wetlands is to develop appropriate hydrologic conditions. Groundwater inflow is often desired because this offers a more predictable and less seasonal water source. Surface flooding by rivers gives wetlands a seasonal pattern of flooding, but such wetlands can be dry for extended periods in flood-absent periods. Depending on surface, runoff and flow from low-ordered streams can be the least predictable. Often wetlands developed under these conditions are isolated pools and potential mosquito havens for a good part of the growing season; their design should be carefully considered. It is generally considered to be optimum to build wetlands where they used to be and where hydrology is still in place for the wetland to survive. But tile drainage, ditches, and river downcutting have often changed local hydrology from prior conditions. Most biologists have difficulty in estimating hydrologic conditions whereas engineers often overengineer control structures that need substantial maintenance and are not sustainable.

Soils

Often choice of the site of wetland creation and restoration is limited by property ownership. If a choice exists, a wetland that is restored on former wetland (hydric) soils is much preferred over one constructed on upland soils. Hydric soils develop certain color and chemical patterns because they have spent long periods flooded and thus under anaerobic conditions. The soil color is mostly black in mineral hydric soils because iron and manganese minerals have been converted to reduced soluble forms and have leached out of the soil. In most cases, developing wetlands on hydric soils has three advantages:

1. hydric soils indicate that site may still have or can be restored to appropriate hydrology;

2. hydric soils may be a *seedbank* of wetland plants still established in the soil; and
3. hydric soils may have the appropriate soil chemistry for enhancing certain wetland processes. For example, mineral hydric soils generally have higher soil carbon than do mineral nonhydric soils. This soil carbon, in turn, stimulates wetland processes such as denitrification and methane production.

Otherwise, it is possible to create wetlands on upland soils, and in the long run, these soils will develop characteristics typical of hydric soils such as higher carbon content and seed banks.

Natural Succession versus Horticulture

To develop a wetland that will ultimately be a low-maintenance one, natural successional processes need to be allowed to proceed. The best strategy is usually to introduce, by seeding and planting, as many native choices as possible to allow natural processes to sort out the species and communities in a timely fashion. Wetlands created or restored by this approach are called *self-design wetlands*. Providing some help to this selection process, for example, selective weeding, may be necessary in the beginning, but ultimately the system needs to survive with its own successional patterns unless significant labor-intensive management is possible. A somewhat different approach, called *designer wetlands*, occurs when specified plant species are introduced, and the success or failure of those plants are used as indicators of success or failure of that wetland. This is akin to horticulture.

An important general consideration of wetland design is whether plant material is going to be allowed to develop naturally from some initial seeding and planting or whether continuous horticultural selection for desired plants will be imposed. Odum (1987) suggested "in many freshwater wetland sites, it may be an expensive waste of time to plant species which are of high value to wildlife.... It may be wiser to simply accept the establishment of disturbance species as a cheaper although somewhat less attractive solution." Reinartz and Warne (1993) found that the way vegetation is established can affect the diversity and value of the mitigation wetland system. Their study showed that early introduction of a diversity of wetland plants may enhance the long-term diversity of vegetation in created wetlands. The study examined the natural colonization of plants in 11 created wetlands in southeastern Wisconsin. The wetlands under study were small, isolated, depressional wetlands. A 2-year sampling program was conducted for the created wetlands, aged 1–3 years. Colonization was compared to five seeded wetlands where 22 species were introduced. The diversity and richness of plants in the colonized wetlands increased with age, size, and proximity to the nearest wetland source. In the colonized sites, *Typha* spp. comprised 15% of the vegetation for 1-year wetlands, and 55% for 3-year wetlands, with the possibility of monocultures of *Typha* spp. developing over time in colonized wetlands. The seeded wetlands had a high species diversity and richness after two years. *Typha* cover in these sites was lower than in the colonized sites after 2 years.

Case Study 8 – To Plant or Not Plant Created Wetlands

A study where the effects of planting versus not planting have been observed for over 15 years is at two experimental wetlands at the Olentangy River Wetland Research Park in central Ohio (**Figure 7**; Mitsch *et al.*, 1998, 2005, 2012). In essence, both wetlands were different degrees of self-design since there were no expectations as to what the ultimate cover would be, and there was no "gardening" to get to any endpoint. After 3 years, both wetlands were principally dominated by soft-stem bulrush *Schoenoplectus tabernaemontani* (=*Scirpus validus*) and were thought to be similar. Researchers found that both planted and unplanted wetlands converged in most of the 16 ecological indicators (eight biological measures and eight biochemical measures) in those 3 years (Mitsch *et al.*, 1998). After 6 years, however, several communities of vegetation continued to exist in the planted basin, but a highly productive monoculture of *Typha* dominated the unplanted basin (Mitsch *et al.*, 2005).

This study has now completed over 15 years of observations where one full-scale wetland was planted, and an identical wetland remained unplanted. The wetlands showed that there are a few differences in wetland function that persist a decade planting in planted and unplanted (naturally colonizing) wetland basins that could be traced to effects of the initial planting (Mitsch *et al.*, 2012). Some values that we appreciate in wetlands, for example, carbon sequestration and amphibian production, were higher in the naturally colonizing wetland while other values such as macrophyte community diversity were higher in the planted wetland. Planting did have an effect, but whether to plant depends on the original objective of the wetland. If plant diversity is desired, then planting makes sense. If productivity and carbon sequestration are desired, it may be a waste of effort to plant unless there are no sources of plant propagules (seed banks or inflowing rivers, for example). In any regard, there appears to be a lingering long-term effect on ecosystem function caused by planting.

Estimating Success

There are few satisfactory methods available to determine the "success" of a created or restored wetland or even a mitigation wetland created to replace the functions lost with the original wetland (Mitsch and Wilson, 1996). It is clear from the many studies of created and restored wetlands that some cases are successes, while there are still far too many examples of failure of created and restored wetlands to meet expectations. In some cases, expectations were unreasonable, as when endangered species habitat was to be established in a heavily urbanized environment (Malakoff, 1998). In such cases, the original wetland should not have been lost to begin with. Where expectations are ecologically reasonable, there is optimism that wetlands can be created and restored and that wetland function can be replaced. The spotty record to date is due to three factors:

1. little understanding of wetland function by those constructing the wetlands;
2. insufficient time for the wetlands to develop; and
3. a lack or recognition or underestimation of the self-design capacity of nature.



Figure 7 Experimental and other created and restored wetlands at the Olentangy River Wetland Research Park, Columbus, Ohio.

Understanding wetlands enough to be able to create and restore them requires substantial training in plants, soils, wildlife, hydrology, water quality, and engineering. Replacement projects and other restorations involving freshwater marshes need enough time, closer to 15 or 20 years than to 5 years, before success is apparent. Restoration and creation of forested wetlands, coastal wetlands, or peatlands may require even more time. Peatland restoration could take decades or more. And, of course, forested wetland restoration generally takes a lifetime.

Principles for Proper Ecological Engineering of Wetlands

Some general principles of ecological engineering that apply to the creation and restoration of wetlands are outlined below (Mitsch and Jørgensen, 2004):

1. Design the system for minimum maintenance and a general reliance on self-design.
2. Design a system that utilizes natural energies, such as the potential energy of streams as natural subsidies to the system.
3. Design the system with the hydrologic and ecological landscape and climate.
4. Design the system to fulfill multiple goals but identify at least one major objective and several secondary objectives.
5. Give the system time.
6. Design the system for function, not form.
7. Do not overengineer wetland design with rectangular basins, rigid structures and channels, and regular morphology.

Zedler (2000a) had suggested the following additional ecological principles that should be applied to wetland restoration:

1. Landscape context and position are crucial to wetland restoration. See #3 design principle above. Wetlands are always a function of the watershed and ecological setting in which they are placed.
2. Natural habitat types are the appropriate reference systems. This suggests that while we may know how to build ponds, for example, are those the natural habitats of the area, even if they do increase waterfowl?
3. The specific hydrologic regime is crucial to restoring biodiversity and function. See # 2 and 3 design principles. In many cases, such as the Florida Everglades, the restoration is being done in the face of a massive change in the hydrologic character of the landscape.
4. Ecosystem attributes develop at different paces. Give the system time; see # 5 above. Hydrology develops quickly, vegetation over several years, and soils over decades. Yet we are quick to review and criticize created and restored wetlands after a couple of years.
5. Nutrient supply rates affect biodiversity recovery. There are low-nutrient and high-nutrient wetland systems. Low-nutrient wetlands are often more difficult to create or restore. High-nutrient inflows cause wetlands to go for power instead of biodiversity. In other cases, insufficient nutrients were the cause of wetland failure as was the case in southern California salt marshes (Zedler, 2000b).
6. Specific disturbance regimes can increase species richness. This can clearly be the case if we allow the word "disturbance" to include flood pulses, fire, and even tropical storms.
7. Seed banks and dispersal can limit recovery of plant species richness. This is why restoring wetlands with seed banks can be so important. Another solution is to have a hydrologically or biologically "open" system with a multitude of inputs of propagules (plants, animals, and microbes) more likely.
8. Environmental conditions and life history traits must be considered when restoring biodiversity. See our discussion on self-design versus designer wetlands below.
9. Predicting wetland restoration begins with succession theory. Again, design principle # 5 says we need to give the system time. Ecological succession cannot be accelerated without other consequences. This also supports our contention that one must understand wetland science first before attempting to create and restore wetlands.
10. Genotypes influence ecosystem structure and function. This is an important but often overlooked principle of wetland restoration. Species are not the same everywhere. This has been shown in common garden experiments on *Spartina alterniflora* (Seliskar, 1995) and freshwater rush *Juncus effusus* (Weihe and Mitsch, 2000). A brackish/freshwater wetland plant with this several genotypes that has invaded many natural and restored wetlands in the USA is *Phragmites australis*. This creates a difficult problem in wetland creation and restoration as wetland managers must be able to distinguish between the invasive "bad" *Phragmites* and the native "good" *Phragmites*.

Robin Lewis and Kevin Erwin, two wetland consultants from Florida, have spent about five decades restoring and creating wetlands around the world. Their experience has led to the following 15 recommendations (Lewis *et al.*, 1995):

1. Wetland restoration and creation proposals must be viewed with great care, particularly when promises are made to restore or re-create a natural system in exchange for a permit.
2. Multidisciplinary expertise in planning and careful project supervision at all project levels is needed.
3. Clear, site-specific measurable goals should be established.
4. A relatively detailed plan concerning all phases of the project should be prepared in advance to help evaluate the probability of success.
5. Site-specific studies should be carried out in the original system prior to wetland alteration if wetlands are being lost in the project.
6. Careful attention to wetland hydrology is needed in design.
7. Wetlands should, in general, be designed to be self-sustaining systems and persistent features of the landscape.
8. Wetland design should consider relationships of the wetland to the watershed, water sources, other wetlands in the watershed, and adjacent upland and deepwater habitat.

9. Buffers, barriers, and other protective measures are often needed.
10. Restoration should be favored over creation.
11. The capability for monitoring and mid-course corrections is needed.
12. The capability for long-term management is needed for some types of systems.
13. Risks inherent in restoration and creation and the probability of success for restoring or creating particular wetland types and functions should be reflected in standards and criteria for projects and project design.
14. Restoration for artificial or already altered systems requires special treatment.
15. Emphasis on ecological restoration of watersheds and landscape ecosystem management requires advanced planning.

The above principles and practices of wetland creation and restoration are based on wetland science (hydrology, biogeochemistry, adaptations, and succession). If one is interested in creating and restoring wetlands, one should first become an expert in wetland science. When we attempt to create or recreate natural ecosystems, Mother Nature is in control. In all situations of wetland creation and restoration, human contribution to the design of wetlands should be kept simple and should strive to stay within the bounds established by the natural landscape. As stated by Boulé (1988) "simple systems tend to be self-regulating and self-maintaining." We should recognize that Nature remains the chief agent of self-design, ecosystem development, and ecosystem maintenance; humans are not the only participants in these processes. We have referred to these self-design and time requirements for successful ecosystem restoration and creation as invoking "*Mother Nature and Father Time*" (Mitsch and Wilson, 1996; Mitsch *et al.*, 1998, 2012).

Appendix

List of Courses

1. Wetland Ecology
2. Ecological Restoration
3. Ecological Engineering
4. River Restoration

See also: Ecosystem, Concept of. Energy Flow and Ecosystems. Impact of Ecological Restoration on Ecosystem Services. Mangrove Ecosystems. Riparian Landscapes. River Ecosystems. Wetlands Ecosystems

References

Benthem W, van Lavieren LP, and Verheugt WJM (1999) Mangrove rehabilitation in the coastal Mekong Delta, Vietnam. In: Streever W (ed.) *An International Perspective on Wetland Rehabilitation*, pp. 29–36. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Boulé ME (1988) Wetland creation and enhancement in the Pacific Northwest. In: Zelazny J and Feierabend JS (eds.) *Proceedings of the Conference on Wetlands: Wetlands: Increasing Our Wetland Resources*, pp. 130–136. Washington, DC: Corporate Conservation Council, National Wildlife Federation.

Costanza R, Mitsch WJ, and Day JW (2006) A new vision for New Orleans and the Mississippi delta: Applying ecological economics and ecological engineering. *Frontiers in Ecology and the Environment* 4: 465–472.

Dugan P (1993) *Wetlands in Danger*, 192 pp. London: Reed International Books.

Gorham E and Rochefort L (2003) Peatland restoration: A brief assessment with special reference to *Sphagnum* bogs. *Wetlands Ecology and Management* 11: 109–119.

Hefner JM and Brown JD (1985) Wetland trends in the southeastern United States. *Wetlands* 4: 1–11.

Hey DL and Philippi NS (1995) Flood reduction through wetland restoration: The Upper Mississippi River Basin as a case study. *Restoration Ecology* 3: 4–17.

Hinkle R and Mitsch WJ (2005) Salt marsh vegetation recovery at salt hay farm wetland restoration sites on Delaware Bay. *Ecological Engineering* 25: 240–251.

Isselin-Nondedeu F, Rochefort L, and Poulin M (2007) Long-term vegetation monitoring to assess the restoration success of a vacuum-mined peatland (Québec, Canada). In: *International Conference Peat and Peatlands 2007*, pp. 153–166.

King SL and Keeland BD (1999) Evaluation of reforestation in the Lower Mississippi River alluvial valley. *Restoration Ecology* 7: 348–359.

Lane RR, Day Jr. JW, and Thibodeaux B (1999) Water quality analysis of a freshwater diversion at Caernarvon, Louisiana. *Estuaries* 22: 327–336.

de Leon ROD and White AT (1999) Mangrove rehabilitation in the Philippines. In: Streever W (ed.) *An International Perspective on Wetland Rehabilitation*, pp. 37–42. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lewis III RR (2005) Ecological engineering for successful management and restoration of mangrove forests. *Ecological Engineering* 24: 403–418.

Lewis RR, Kusler JA, and Erwin KL (1995) Lessons learned from five decades of wetland restoration and creation in North America. In: Montes C, Oliver G, Molina F, and Cobos J (eds.) *Bases Ecologicas para la Restauracion de Humedales en la Cuenca Mediterránea*, pp. 107–122. Spain: Consejeria de Medio Ambiente, Junta de Andalucía, Andalucía.

Malakoff D (1998) Restored wetlands flunk real-world test. *Science* 280: 371–372.

Mitsch WJ (ed.) (1994) *Global Wetlands: Old World and New*, 967 + xxiv pp. Amsterdam: Elsevier Scientific Publishing Co.

Mitsch WJ and Day Jr. JW (2006) Restoration of wetlands in the Mississippi–Ohio–Missouri (MOM) River Basin: Experience and needed research. *Ecological Engineering* 26: 55–69.

Mitsch WJ, Day Jr. JW, Gilliam JW, *et al.* (2001) Reducing nitrogen loading to the Gulf of Mexico from the Mississippi River Basin: Strategies to counter a persistent ecological problem. *BioScience* 51: 373–388.

Mitsch WJ and Gosselink JG (2007) *Wetlands*, 4th edn., 582 pp. New York: John Wiley & Sons, Inc.

Mitsch WJ, Gosselink JG, Anderson CJ, and Zhang L (2009) *Wetland Ecosystems*, 295 pp. New York: John Wiley & Sons, Inc.

Mitsch WJ and Jørgensen SE (2004) *Ecological Engineering and Ecosystem Restoration*, 411 pp. New York: John Wiley & Sons, Inc.

Mitsch WJ, Wang N, Zhang L, Deal R, Wu X, and Zuwerink A (2005) Using ecological indicators in a whole-ecosystem wetland experiment. In: Jørgensen SE, Xu F-L, and Costanza R (eds.) *Handbook of Ecological Indicators for Assessment of Ecosystem Health*, pp. 211–235. Boca Raton, FL: CRC Press.

Mitsch WJ and Wilson RF (1996) Improving the success of wetland creation and restoration with know-how, time, and self-design. *Ecological Applications* 6: 77–83.

Mitsch WJ, Wu X, Nairn RW, *et al.* (1998) Creating and restoring wetlands: A whole-ecosystem experiment in self-design. *BioScience* 48: 1019–1030.

Mitsch WJ, Zhang L, Stefanik KC, *et al.* (2012) Creating wetlands: Primary succession, water quality changes, and self-design over 15 years. *BioScience* 62: 237–250.

National Research Council (2001) *Compensating for Wetland Losses under the Clean Water Act*, 158 pp. Washington, DC: National Academy Press.

Odum WE (1987) Predicting ecosystem development following creation and restoration of wetlands. In: Zelazny J and Feierabend JS (eds.) *Wetlands: Increasing Our Wetland Resources. Proceedings of the Conference Wetlands: Increasing Our Wetland Resources. Corporate Conservation Council*, pp. 67–70. Washington, DC: National Wildlife Federation.

Pedersen ML, Andersen JM, Nielsen K, and Linnemann M (2007) Restoration of the Skjern River and its valley: Project description and general ecological changes in the project area. *Ecological Engineering* 30: 131–144.

Peterson SB, Teal JM, and Mitsch WJ (eds.) (2005) Delaware Bay Salt salt Marsh marsh Restorationrestoration. *Special Issue of* Ecological Engineering 25: 199–314.

Price J, Rochefort L, and Quinty F (1998) Energy and moisture considerations on cutover peatlands: Surface microtopography, mulch cover and *Sphagnum* regeneration. *Ecological Engineering* 10: 293–312.

Quinty F and Rochefort L (1997) Plant reintroduction on a harvested peat bog. In: Trettin CC, Jurgensen MF, Grigal DF, Gale MR, and Jeglum JK (eds.) *Northern Forested Wetlands: Ecology and Management*, pp. 133–145. Boca Raton, FL: CRC Press/Lewis Publishers.

Rabalais NN, Turner RE, and Wiseman Jr. WJ (2001) Hypoxia in the Gulf of Mexico. *Journal of Environmental Quality* 30: 320–329.

Reddy KR, Kadlec RH, Chimney MJ, and Mitsch WJ (eds.) (2006) The Everglades everglades Nutrient nutrient Removal removal Projectproject. *Special Issue of Ecological Engineering* 27: 265–379.

Reinartz JA and Warne EL (1993) Development of vegetation in small created wetlands in southeast Wisconsin. *Wetlands* 13: 153–164.

Richardson CJ and Hussain NA (2006) Restoring the Garden of Eden: An ecological assessment of the marshes of Iraq. *BioScience* 56: 447–489.

Richardson CJ, Reiss P, Hussain NA, Alwash AJ, and Pool DJ (2005) The restoration potential of the Mesopotamian marshes of Iraq. *Science* 307: 1307–1311.

Robb JT (2002) Assessing wetland compensatory mitigation sites to aid in establishing mitigation ratios. *Wetlands* 22: 435–440.

Rochefort L and Campeau S (1997) Rehabilitation work on post-harvested bogs in south eastern Canada. In: Parkyn L, Stoneman RE, and Ingram HAP (eds.) *Conserving Peatlands*, pp. 287–294. Walingford, UK: CAB International.

Rochefort L, Quinty G, Campeau S, Johnson K, and Malterer T (2003) North American approach to the restoration of *Sphagnum* dominated peatlands. *Wetlands Ecology and Management* 11: 3–20.

Seliskar D (1995) Exploiting plant genotypic diversity for coastal salt marsh creation and restoration. In: Khan MA and Ungar IA (eds.) *Biology of Salt-Tolerant Plants*, pp. 407–416. Pakistan: Department of Botany, University of Karachi.

Streever W (ed.) (1999) *An International Perspective on Wetland Rehabilitation*, 338 pp. Dordrecht: Kluwer Academic Publishers.

Teal JM and Peterson SB (2005) Introduction to the Delaware Bay salt marsh restoration. *Ecological Engineering* 25: 199–203.

Teal JM and Weinstein MP (2002) Ecological engineering, design, and construction considerations for marsh restorations in Delaware Bay, USA. *Ecological Engineering* 18: 607–618.

Waddington JM, Greenwood MJ, Petrone RM, and Price JS (2003) Mulch decomposition impedes recovery of net carbon sink function in a restored peatland. *Ecological Engineering* 20: 199–210.

Waddington JM, Tóth K, and Bourbonniere R (2008) Dissolved organic carbon export from a cutover and restored peatland. *Hydrologic Processes* 22: 2215–2224.

Warner BG and Rubec CDA (eds.) (1997) *The Canadian Wetland Classification System*. Waterloo, Ontario: National Wetlands Working Group, Wetlands Research Centre, University of Waterloo.

Weihe PE and Mitsch WJ (2000) Garden wetland experiment demonstrates genetic differences in soft rush obtained from different regions (Ohio). *Ecological Restoration* 18: 258–259.

Weinstein MP, Balletto JH, Teal JM, and Ludwig DF (1997) Success criteria and adaptive management for a large-scale wetland restoration project. *Wetlands Ecology and Management* 4: 111–127.

Weinstein MP, Teal JM, Balletto JH, and Strait KA (2001) Restoration principles emerging from one of the world's largest tidal marsh restoration projects. *Wetlands Ecology and Management* 9: 387–407.

Zedler JB (2000a) Progress in wetland restoration ecology. *TREE* 15: 402–407.

Zedler JB (2000b) *Handbook for Restoring Tidal Wetlands*, 464 pp. Boca Raton, FL: CRC Press.

Zoltai SC and Vitt DH (1995) Canadian wetlands: Environmental gradients and classification. *Vegetatio* 118: 131–137.

Wetlands Ecosystems

Barbara L Bedford, Cornell University, Ithaca, NY, USA
Donald J Leopold and James P Gibbs, State University of New York, Syracuse, NY, USA

© 2001 Elsevier Inc. All rights reserved.

This article is reproduced from the previous edition, volume 5, pp 781–804, © 2001, Elsevier Inc.

Glossary

Benthic Located on the bottom of a water body, or pertaining to bottom-dwelling organisms.
Biological diversity The variety of different forms of life on earth, usually considered to include variety at all levels of biological organization from genetic variation to variety seen in species, populations, communities, and ecosystems.
Bog An acidic peatland that receives all water inputs from the atmosphere and hence has low pH (usually 3.4–4.6), low alkalinity, and negligible concentrations of base cations in surface waters; typically dominated by mosses of the genus *Sphagnum* and other plants able to tolerate low pH and low nutrient availability – that is, members of the heath family (Ericaceae).
Calcareous fen A term variously used for wetlands with high calcium carbonate in water and soils. In the strict sense, the term refers to peat-accumulating wetlands with surficial deposits of calcium carbonate and a distinctive flora of calciphilic species. These wetlands typically develop where substantial amounts of cold, calcium-carbonate rich groundwater discharge to the surface.
Detritivores Organisms that consume dead plant or animal material.
Emergent vegetation Rooted plants that tolerate saturated or flooded soil but not extended periods of submersion.
Floodplain Broad, flat areas adjacent to rivers and streams that become inundated during floods.
Hydrologic regime The aggregate set of variables describing the behavior of water in wetlands, including water depth, magnitude and frequency of water level fluctuations, seasonal pattern of water fluctuations, direction and velocity of water flows, and water residence time.
Hydrophytes Plants capable of growing and persisting in saturated or flooded soils.
Littoral The shoreline zone where sunlight penetrates to a wetland's substrate and supports rooted plant growth.
Macroinvertebrates Organisms lacking backbones and generally visible to the human eye.
Marsh A wetland dominated by herbaceous species of plants (e.g., cattails, sedges, grasses, floating ferns) and typically occurring on mineral soils. Floating-leaved and submerged species of plants dominate the parts of marshes that grade into open water. Although the surface layer of some marsh soils may consist of shallow accumulations of organic matter over mineral soils, the depth of this layer is usually less than the minimum required for the site to be classified as peatland.
Peat A type of soil formed from accumulation of the remains of dead plants, with high organic matter content (minimally 20% and as high as 90% by dry weight) and depth greater than 30 to 40 cm. Peat forms in some wetlands because anoxic conditions, low pH, low temperatures and short growing seasons, or plant tissues highly resistant to decay prevent rates of decomposition from keeping up with rates of organic matter production.
Peatlands Wetlands in which dead organic matter accumulates to form peat more than 30 to 40 cm deep; typically includes bogs, poor fens, and rich fens but some swamps also occur on peat.
Poor fen An acidic peatland similar to a bog but having a somewhat higher surface water pH (4.3–5.8) because it receives some groundwater or surface water in addition to precipitation; typically dominated by *Sphagnum* mosses or sedges. But for the presence of distinctive poor-fen indicators, vegetation may superficially resemble bog vegetation or sedge-dominated rich fens that lack *Sphagnum*.
Rich fen An alkaline peatland that receives significant inputs from groundwater or surface water; typically having surface water pHs >6.0, higher alkalinity, and distinctly higher concentrations of base cations than bogs and poor fens; dominated by brown mosses (true mosses, largely of the Amblystegiaceae) and small sedges.
Species richness The number of different species in a specified area.
Submerged vegetation Plants that grow and reproduce while completely submerged.
Swamp Wetlands dominated by woody species. Although the term has been variously used, it is used here in the restricted sense of forested wetland, either on mineral soils or peat. However, while many peatland ecologists classify forested fens on deep peat as swamps, they would not call a bog forest a swamp.

Diversity of Wetland Ecosystems

Wetlands, and the organisms adapted to these unique environments, are as diverse as the myriad combinations of climate, topography, and geology that cause different types of wetlands to form. By definition, wetlands are ecosystems whose physical, chemical, and biological characteristics are determined by the constant or recurrent presence of water at

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00314-2

or near the soil surface. Beyond this commonality, however, lies a seemingly infinite number of ways in which climate and landscape features combine to create wetlands with different: magnitudes and seasonal patterns of water level fluctuation, paths and rates of water movement, soils, water chemistry, nutrient status, and species composition. Out of this landscape diversity rises the large number of different types of wetland ecosystems seen worldwide.

Wetlands occur on every continent except Antarctica and in every climatic regime. The best available estimates indicate that they cover only between 4 and 6% of the Earth's land surface. Within this relatively small area, however, an astonishingly high number of different wetland ecosystem types occur. A recent global summary (Mitsch, 1994) hints at this diversity when it describes the world's wetlands in terms of coarse-scaled differences in climatic setting, geomorphologic setting, and major wetland type. According to that summary, wetlands "are found (a) in humid, cool regions as bogs, fens, and tundra; (b) along rivers and streams as riparian wetlands, seasonally flooded forests, and back swamps; (c) in the delta's of the world's great rivers; (d) along temperate, subtropical, and tropical coastlines as salt marshes, mudflats, and mangrove swamps; and (e) in arid regions as inland salt flats, seasonal playas, and vernal pools". They also are found as (f) marshes, wet meadows, and swamps in depressions that occur in a variety of landscapes and climatic settings; (g) marshes, wet meadows, and swamps on extensive flat areas where the land surface slope is minimal and climate is humid; and (h) marshes, bogs, and fens along the shores of lakes in many climatic regions. Wetlands even occur on (i) slopes as blanket bogs, seepage fens, and vegetated springs.

Finer-scaled differences in climate, geomorphologic setting, and biogeographic region yield an even greater diversity of wetland ecosystem types. Mangrove swamps, for instance, do not exhibit uniform species composition worldwide. The mangrove ecosystems of the Indo-western Pacific have a far greater number of plant species than do the mangrove swamps of Central America, and they certainly should be considered different ecosystems for conservation purposes. The nutrient-poor sawgrass (*Cladium jamaicense*) marshes of the Everglades differ profoundly from the nutrient-rich giant reed grass (*Phragmites australis*) marshes of the Camargue in France's Rhone delta in land form, water chemistry, vegetation, and animal life. Neither of these marsh ecosystems bears much similarity to the thousands of small marshes that dot the prairie pothole region of Canada and the north central United States. The montane cushion-plant herbaceous bogs of Chile contain flora and fauna quite distinct from that of the *Sphagnum* moss-dominated bogs of the Northern Hemisphere, and from the domed bogs of Sarawak. The shore pine (*Pinus contorta*) bogs of Vancouver Island's Pacific coast contain many plant species not seen in Canada's continental bogs. Seasonally flooded forests along rivers include ecosystems as dissimilar as the sometimes monotypic palm swamps of South America, the moderately species-rich hardwood swamps of the southeastern United States (**Figure 1**), and the highly species-rich *varzea* forests of the Amazon basin. Forested wetlands (**Figure 2**) also include those occurring on steep windward slopes of Puerto Rico's Luquillo Mountains where annual precipitation as high as 5000 mm keeps the clayey soils continuously saturated. In Finland, where about one-third of the land is peatland, at least 30 different mire types are recognized. Some of the bogs and fens of Ireland and Scotland, Scandinavia, Finland, Alaska and Canada, the former Soviet Union, and Maine form intricately patterned peatlands, some of the most unique land forms on earth.



Figure 1 Southern deepwater swamp, Congaree National Monument, South Carolina, USA.



Figure 2 Forested wetland in upstate New York, USA.

Classification of the Major Wetland Ecosystems of the World

Because wetlands take so many forms, occur in many different physiographic and climatic settings, and form part of the history and culture of many regions of the world, their diversity is not easily cataloged. Different regional and colloquial names for them abound, as do formal classification systems. Different common names may refer to the same type of wetland with similar species, while the same name may be applied in different parts of the world to wetlands with quite different species. For example, the terms mire, moor, muskeg, and peatland all refer to similar types of wetlands with similar plant growth forms and many common species. On the other hand, the term swamp is used for a variety of forested wetlands in the United States but for areas dominated by giant

reed grass (*Phragmites*) and other herbaceous species in parts of Europe and Africa. Fens dominated by woody species may be called swamps, forested fens, or forested wetlands. In the strictest sense, bogs are peat-accumulating wetlands that receive water only from precipitation. However, the term also is applied to poor fens, which superficially resemble true bogs, and more loosely to any wet vegetated environment that offers less than firm footing.

Numerous formal classifications of the different types of wetlands have been developed. They give some indication of the vast diversity of wetland ecosystems but cannot be used to enumerate that diversity. Because they have been developed for different purposes and from different scientific perspectives, they vary in the criteria used for classification and in the level of resolution to which they classify wetlands. Nonetheless, the existence of literally dozens of wetland classifications worldwide implies the existence of such a high diversity of types that they must be grouped into similar classes for practical reasons of inventory, mapping, management, and scientific communication.

The high diversity of wetland ecosystems also is revealed by considering the many different bases for wetland classification: geomorphologic setting in the landscape, genesis (how they formed), shape or form (e.g., surface morphology, basin morphometry, morphology of the underlying mineral terrain), hydrology (e.g., flooding frequency and duration), sources of water supply, physiognomy of the vegetation, species composition of the vegetation, soil type, water chemistry, and various combinations of these factors. Given that each of these variables represents not a single factor but rather a continuum, the extremely large number of possible combinations becomes apparent. Thus, most classifications in wide use today are broadly based, interdisciplinary, and hierarchical (i.e., consisting of several nested levels with increasing detail only at lower levels in the hierarchy). Species composition, the basis for diversity within wetlands and the focus of many older classifications, is usually not a defining characteristic except at the lowest levels in current classifications. Rather, similarities in hydrology and geomorphology are more likely to form the highest levels of classification, reflecting the importance of the interaction of water with the landscape in determining the biotic characteristics of wetlands.

Nonetheless, the terms bog, fen, marsh, and swamp persist in both scientific and popular literature because of their long history of use and the powerful images they evoke. These terms are not distinguished primarily on the basis of geomorphology or hydrology; nor are they distinguished on any single criterion, nor used consistently. Perhaps the key distinction among them in terms of species diversity is their source or sources of water, which determine hydrologic regime and water chemistry. Greater surface water inputs generally lead to greater water level fluctuations and higher concentrations of base cations and nutrients in marshes and swamps on mineral soils than in bogs and fens. In this chapter, these terms follow current American usage (*see* Glossary).

The current classification systems used in the United States (Cowardin *et al.*, 1979) and Canada (National Wetlands Working Group, 1997) reveal much about the diversity of wetlands in these two countries. For example, the U.S. classification is hierarchical, recognizing 5 major wetland *systems* and 10 *subsystems*, each of which is divided further into 4 to 8 *classes*. A *system* is defined as "a complex of wetlands and deepwater habitats that share the influence of similar hydrologic, geomorphologic, chemical, or biological factors" (Cowardin *et al.*, 1979, p. 4). The five systems are the marine, estuarine, riverine, lacustrine, and palustrine. The term *class* "describes the general appearance of the habitat in terms of either the dominant life forms of the vegetation or the physiography and composition of the substrate—features that can be recognized without the aid of detailed environmental measurements." Classes can be further divided into *dominance* types based on dominant plant species for vegetated sites or on the dominant sedentary or sessile animal species where vegetation is not the dominant cover. Special *modifiers* are used for water regime (type and duration of flooding), water chemistry, soils, and human alterations (e.g., excavated, impounded, diked). Fifty-five classes emerge before one ever gets to the level of dominant species, water regime, or water chemistry.

Most of Canada's wetlands fall within just one system (palustrine) of the U.S. classification. Canada's classification differentiates many types of palustrine wetland, especially peatlands. This classification also is hierarchical but with just three levels: *class, form,* and *type.* Five *classes* are recognized: bog, fen, marsh, swamp, and shallow water. *Forms* are defined on the basis of surface morphology of the wetland, position in the landscape, morphology of the underlying terrain, tidal effects, and proximity to surface water bodies. The Canadian system recognizes 16 bog forms and 4 subforms, 12 fen forms and 7 subforms, 8 marsh forms with 18 subforms, 8 swamp forms with 19 subforms, and 5 shallow water forms with 24 subforms. The 18 *types* based on the general physiognomy of the vegetation (e.g., hardwood treed, coniferous treed, low shrub) modify these 121 forms and subforms. Differences in species composition are left to individual users.

Only the wetlands of western Europe are as well documented as those of the United States and Canada. Specific sites on other continents have received considerable attention but, with few exceptions, comparative regional treatments began to emerge in the English language only 15 years ago, with most appearing since 1990 (*see* References). The exceptions include Chapman's books—for example, *Salt Marshes and Salt Deserts of the World* (1960), *Mangrove Vegetation* (1976), *and Wet Coastal Ecosystems* (1977). However, despite the general lack of scientific studies in many areas of the world, as of 1991, 62 countries from all continents had signed the Convention on Wetlands of International Importance Especially as Waterfowl Habitat (Ramsar Convention) and had designated 527 sites covering more than 30 million ha as being of international importance. Five international wetland symposia have been organized since 1980 by INTECOL (The International Society of Ecology), with a sixth planned for the year 2000 in Quebec City.

Major Regional Wetland Ecosystems

Efforts to classify wetlands into unique entities often obscure, either spatially or temporally, critical connections among parts of what might be viewed as integrated ecological systems.

For example, various classification systems segment the great delta of the Mississippi and Atchafalaya rivers into many types of floodplain forest, freshwater marsh, salt marsh, streams, and lakes. They assign each prairie pothole in the prairie region of the United States and Canada to one of about 20 classes of prairie wetlands. The mangrove forests of Central America are made to seem distinct from the adjacent beds of sea grasses. Functionally, however, these diverse wetlands are linked and might be better viewed as major regional wetland ecosystems. Water, sediments, plant propagules, fish, birds, other animals, genetic material, and pollutants move among them. One type changes through time to become another. Diversity emerges in both space and time from the components as well as their interconnectedness.

In fact, prior to human modification of them, most wetlands existed as complexes of several different types functioning as integrated systems linked by movements of water, sediment, nutrients, plants, and animals. The vast rain-fed and riverine floodplains that form the Sudd in southern Sudan, for example, consist of a complex and dynamic mix of different types of grassland (e.g., *Echinochloa pyramidalis* or *Hyparrhenia rufa* dominated), wild rice marshes (*Oryza longistaminata*), woodlands (*Acacia seyal* or *Balanites aegyptiaca* dominated), and open water that cover more than 45,000 km^2. The great floodplain of the Orinoco River and its tributaries form a nearly 10-million ha wetland complex of streams, ponds, shallow oxbow lakes, riverine forest, marshes, and seasonally-flooded palm (*Copernicia tectorum*) savannas and grassland in Venezuela and eastern Columbia. The eastern shore of Lake Ontario, New York, is bordered by a 17-mile long interconnected system of streams, shallow ponds, marshes, fens, and shrub swamps. Most temperate zone estuaries include freshwater, brackish, and saltwater marshes along with tidal creeks and mudflats. Mangrove swamps, saline lagoons, sea grass beds, and coral reefs form interconnected systems along shallow coasts in subtropical and tropical regions.

These complexes, or assemblages of wetlands within regions, represent another level of wetland diversity that only partially remains today. Current classification systems do not capture such regional ecosystems, though clearly they exist. Many people recognize wetland complexes such as the Everglades and the Prairie Pothole Region as regional ecosystems. Classification systems, however, only recognize hierarchy in types, not in assemblages of wetlands within regions. Current efforts by conservation organizations and government agencies to move to ecoregional mapping attempt to overcome this limitation.

Continental and Regional Diversity

Climate and geomorphology, along with biogeographic factors, control the diversity of wetlands on different continents and in different regions. Wetland development is promoted where precipitation annually exceeds evapotranspiration on a long-term basis. In such climates, wetlands develop in a number of landscape settings, including depressions in the land surface, areas where minimal land surface slope slows surface water runoff, along breaks in the slope of the land surface, and even on slopes if precipitation greatly exceeds evapotranspiration. In semiarid, arid, and desert climates, wetlands occur far less frequently, but they are found (a) in areas with depressions or lowlying land along rivers or coasts where runoff converges from extensive uplands or mountains, (b) along coasts regularly inundated by tides, (c) in areas where permafrost or other impermeable layers below ground prevent or slow infiltration of surface water, and (d) in spring-fed oases.

In addition to topographic variation, variation in the composition and stratigraphy of surficial geological deposits promotes diversity in the types of wetland that occur within a region. The composition of subsurface materials plays a major role in determining the chemistry of water entering wetlands, and thereby the plant species composition of wetlands. For example, wetlands occurring in regions dominated by limestone deposits tend to have high-pH waters rich in base cations that promote the growth of distinctive rich fen vegetation. In contrast, waters draining areas with granitic bedrock tend to be low in pH and base cations; such areas more frequently contain bog-like vegetation. Areas with thick deposits of glacial till can contain a mix of rock types that, in turn, may support a variety of wetland types. A diversity of types in a given region also is promoted where surface topography and subsurface stratigraphy of geological deposits produce groundwater flow paths of varying lengths. Wetland water chemistry is strongly influenced not only by water source but also by the length of groundwater flow paths that enter the wetland and the number of different groundwater flow systems that feed the wetland. Different flow systems may contribute water with different chemistry because the water originates from different geological deposits. If the water originates from the same geological materials, longer flow paths may yield higher concentrations of solutes because longer flow paths generally correlate with longer groundwater residence time (i.e., longer contact time with geological materials).

Biogeographic factors also influence the number of wetland types occurring in different regions and continents. Aquatic plants generally are known for their remarkably wide geographic range, but many species have smaller ranges determined in part by biogeographic factors (i.e., the climatological, geological, glacial, and floristic history of a region). The distributions of these more restricted species help define distinctive biogeographic regions that contain wetlands with species compositions unlike those from other biogeographic regions. For example, shallow ponds that support many species of the Atlantic Coastal Plain flora might be classed broadly as emergent marshes, but this would miss their unique contribution to the diversity of North American wetlands. The flora of these coastal plain ponds, which has an unusual abundance of rare and uncommon species, bears little resemblance to the emergent marshes of other biogeographic regions.

While no systematic comparison has been made of wetland ecosystem diversity in different regions, the factors controlling regional wetland diversity are understood. All other things being equal, regions with the greatest climatic and geomorphologic variation are likely also to contain the highest diversity of wetland types. If wetland types are defined to the level of plant species composition, then biogeography enters into the equation. Thus, continental diversity will vary as a

function of the number and combination of different climatic, geomorphologic, and biogeographic regions occurring on that continent.

Diversity Within Wetland Ecosystems

Plant Species Richness within Wetlands

At the regional and global scale, wetlands clearly exhibit high diversity of types and an associated high diversity of plant species. Within-habitat (or *alpha*) diversity, and between-habitat (or *beta*) diversity, however, is highly variable among wetlands. Wetlands include some of the most species-rich plant communities in the world, as well as numerous examples of extensive stands dominated by a single species. Familiar examples of single-species wetlands include mangrove swamps, tidal marshes of salt marsh cordgrass (*Spartina alterniflora or Spartina patens*), cattail (*Typha* spp.) marshes, and papyrus (*Cyperus papyrus*) swamps. Species-rich herbaceous communities include calcareous fens with as many as 20 to 30 species of vascular plants and bryophytes per square meter. Intermediate examples include abandoned beaver ponds with 3 to 12 species per square meter and a range of coastal wetlands (bottomland hardwoods, wet prairies, fresh marsh, and salt marsh) from the southern United States with 1 to 17 species per square meter.

Change in plant species composition across gradients within wetland sites depends on the amount of change that occurs within a site in the physical and chemical variables controlling species distributions. A single small depression may encompass a wide range of water depths and the associated changes in species composition. Although adjacent zones may share some species, the deep-water end of the gradient, which may be dominated by floating or submerged species, shares no species with the sedge- and forb-dominated wet meadow zone at the other end of the gradient. Depending on the slope and width of the floodplain, wetlands along major rivers may include areas that experience very different frequencies of flooding and consist of very different plant communities, from the open water and marshes of backwater sloughs to hardwood-dominated forests along natural levees. Erosion and deposition of sediments within the floodplain create surfaces of varying ages that support plant communities of different successional status and composition. Much of the species diversity in the Peruvian Amazon has been attributed to the heterogeneity created by long-term patterns of erosion, deposition, and plant succession.

In other wetlands, environment changes little over relatively large distances, for example, where the accumulation of peat has filled in basins and minimized topographic changes. Though the presence of hummocks and hollows introduces small-scale heterogeneity, mean water depth, pH, and nutrient availability change little as one moves across the bog surface. Plant species composition in plots spaced at great distances, but at equivalent microtopographical elevation, may be quite similar.

Although many studies have examined plant species richness in small plots, systematic comparisons are difficult to make because plot size and sample number vary among studies. Given that species richness is known to increase with increasing sample area, conclusions cannot be drawn about relative richness among samples if sample areas are not equal. Nonetheless, some general statements can be made. First, the number of plant species present generally decreases as water depth and frequency of saturation or flooding increase. For example, in his study of Wisconsin plant communities, Curtis (1959) reported that the average number of species per stand ranged as follows: 7 in stands of submerged aquatics, 11 for emergent aquatics, 28 to 29 in sedge meadows, 44 in wet prairie, and 62 in wet-mesic prairie.

Second, the effect of moving water on diversity depends on flood velocity, magnitude, and frequency. Some degree of disturbance from flooding increases diversity by removing live biomass and litter, thereby creating openings for germination of new species. Extreme or very frequent flooding, however, lowers diversity. Few species can withstand both the high stress of extended periods of anoxia and the frequent disturbance of biomass removal.

Third, the vegetation of saline wetlands tends to be less species-rich than freshwater wetlands. For example, coastal salt marshes of the eastern United States rarely contain more than two species in 0.25-m^2 plots (**Figure 3**). Plots of comparable size in freshwater herbaceous communities frequently contain 10 or more species. Increasing the plot size in salt marshes adds few species but is likely to significantly increase species number in many types of freshwater wetlands. For example, the number of vascular and bryophyte species in rich fens in New York increases from about 20 in 1-m^2 plots to 40 in 25-m^2 plots to 65 in 100-m^2 plots (**Figure 12**).

Fourth, in the temperate and boreal zones, *alpha* diversity generally increases across the range of peat-land types from bog to rich fen for vascular species. Continental raised bogs in eastern North America may have only 20 to 25 vascular species per site, one of the most impoverished vascular floras in North America. Rich fens may contain as many as 140 vascular species at a single site. The bryophyte diversity also tends to increase from bogs to rich fens in eastern North America. Continental raised bogs usually have fewer than 20 species of bryophytes in 100 m^2 while rich fens usually have between 20 and 60 bryophytes in a similar-sized plot. However, Vitt and others (1995) reported that bryophyte diversity in peatlands of continental western Canada could be similar for bogs and



Figure 3 Salt marsh in southern New Jersey, USA.

fens. Species richness in larger plots (2500 m^2) ranged from about 5 to 24 in bogs and poor fens and from about 5 to 28 in rich fens. Thus, even though the highest species richness was found in extreme rich fens, any individual bog could contain as many species of bryophytes as a rich fen. At the landscape scale, rich fens are distinctly more species rich than bogs and poor fens. Because their species composition varies more among sites, rich fens have greater diversity when an entire region is considered. In both Britain and continental western Canada, bogs have only about half the number of bryophyte species as fens.

Fifth, relative to the landscapes in which they occur, many types of wetlands appear to be islands of diversity. Fens in northeastern Iowa support 28% of the regional and 18% of the total state flora even though they currently cover only 0.01% of the land surface. Boreal old-growth swamp forests constitute only 5% of Sweden's forest land area but harbor more than half of all vascular forest plant species found in Sweden and about a third of the total Swedish boreal flora of bryophytes. Almost a third of the federally listed threatened and endangered plant species in the United States are wetland dependent. Furthermore, for some types of wetlands, rare and uncommon plant species tend to be associated with the most species-rich sites.

Finally, latitudinal patterns in plant species diversity of wetlands do not appear to parallel those for terrestrial environments. In sharp contrast to terrestrial habitats, wetlands in the tropics are not necessarily more species rich than those in the temperate and boreal zones. According to Ellison 2004, the mangrove swamps, palustrine forested wetlands, palm swamps, and hardwood swamp forests of Central America all exhibit low species diversity. On the basis of a comparison of representative aquatic families and aquatic habitats, Crow (1993) also concluded that aquatic plant diversity was higher in both warm and cool temperate latitudes than in the tropics. On the other hand, some tropical wetlands can be highly diverse. Ellison 2004 reported high plant species richness (292 species of 0.5 m in height or greater) in riparian forest fragments in Belize. Additional systematic and standardized studies are needed to address this issue.

Major Groups of Wetland Animals

In terms of abundance, biomass, and species richness, macroinvertebrates are the most important wetland-dependent animals in freshwater wetlands. Four groups—the insects, mollusks, crustaceans, and annelids—make up the majority. The aquatic insects are the most functionally and taxonomically diverse, and include forms that occur throughout all habitats (the bottom, deep waters, and shallow, vegetated areas) in most wetlands. The midges (Chironomidae, Diptera) constitute perhaps the most widespread group of aquatic insects, and larval forms of midges typically represent the most abundant macroinvertebrates in wetlands. Other ecologically important, widespread groups include the mosquitoes, dragonflies and damselflies (Odonata), mayflies (Ephemeroptera), caddisflies (Trichoptera), stoneflies (Plecoptera), craneflies (Tipulidae), and water beetles (e.g., Corixidae, Belostomatidae, and Notonectidae). Among these groups, only the water beetles are entirely aquatic; members of other groups typically have aquatic larval stages that undergo a synchronized transformation ("hatch") into adult forms, which may live for periods ranging from 1 to 2 days (some mayflies) to weeks or months (many dragon-flies).

The mollusks include the many species of filter-feeding clams and herbivorous snails. The most widespread and common are fingernail clams. Larger clams and mussels are equally widespread but less common, although they are sometimes abundant along large rivers and lake fringes. Most large clams dwell on the bottom of wetlands, whereas the fingernail clams often occur in submerged vegetation. Snails also occur in many wetlands and live upon stems, stalks, and leaves of aquatic vegetation, and graze on epiphytes.

The crustaceans, like the aquatic insects, are commonly encountered throughout most freshwater wetlands. Especially common in open-water and upon aquatic vegetation are isopods and amphipods. In protected open water areas, small zooplankters, including cladocerans and copepods are abundant. Crayfish are common bottom-dwelling crustaceans in many wetland ecosystems. The last major group of aquatic invertebrates, the annelids (mainly oligochaetes), are not particularly diverse taxonomically. However, they can be extremely abundant in some wetlands, where they burrow in the wetland bottom or attach to aquatic vegetation.

Vertebrate animals represent another important component of the wetland fauna and include the amphibians, reptiles, fishes, birds, and mammals. The amphibians are composed of two main groups, the frogs and toads, and the salamanders. Most species of frogs and toads have complex life cycles that include an aquatic, usually herbivorous, larval stage adapted for rapid growth (the tadpole stage), and a terrestrial, carnivorous adult stage adapted for dispersal among wetlands (**Figure 4**). Their relative use of wetlands versus uplands varies among groups. For example, the major group of frogs in North America, the Ranidae, includes (a) members that may be mostly terrestrial but return to wetlands to breed and sometimes to overwinter (e.g., members of the leopard frog complex, e.g. *Rana pipiens*) and (b) other species (e.g., bullfrogs, *Rana catesbeiana*) that are primarily aquatic and typically leave wetlands only to disperse to new areas.



Figure 4 Frog and duckweed in freshwater marsh in upstate New York, USA.

Among North American salamanders, all are wetland dependent except for members of the Plethodontidae, although even this family includes the Desmognathinae, a large group of stream-associated species. In salamanders, wetland dependency ranges from strictly aquatic (e.g., mudpuppies (*Necturus*), hellbenders (*Cryptobranchus*), Conger eels (*Amphiuma*), and sirens (*Siren*)), to semiaquatic in groups that spend much of the growing season in wetlands but overwinter on land (e.g., newts (Salamandridae)), to primarily terrestrial species that return to wetlands only to breed (e.g., the vernal pool-breeding mole salamanders (Ambystomidae)).

In terms of numbers, species richness, and biomass, the majority of wetland fishes are small forage fishes. In North America, the primary groups are the killifishes (*Fundulus*), shiners (*Notropis*), sunfishes (*Lepomis*), and mosquito fishes (*Gambusia*). In deeper-water wetlands or wetlands connected to permanent water bodies, larger, bottom-feeding species occur, such as Ictalurid catfish (e.g., bullheads) and carp (*Cyprinus carpio*), as well as carnivorous species such as pickerel (*Esox*), perch (*Perca*), and bass (*Micropterus, Morone*).

Turtles are the most diverse group of wetland reptiles. Most turtles are highly aquatic and leave water only to lay their eggs and to disperse to new habitats. Commonly encountered, highly aquatic turtles in North America include the snapping turtle (*Chelydra serpentina*), mud and musk turtles (Kinosternidae), and softshells (Trionychidae). The Family Emydidae includes many other highly aquatic turtles, such as cooters, sliders, painted turtles (*Pseudemys and Chrysemys*), map turtles and sawbacks (*Graptemys*), and terrapins (*Malaclemys*), as well as other partially aquatic species that travel regularly among disjunct wetlands (e.g., spotted turtle (*Clemmys guttata*)) or that use wetlands only for hibernation (e.g., wood turtles (*Clemmys insculpta*)). Other wetland-dependent reptiles include the water snakes (e.g., *Nerodia*) that use wetlands primarily for feeding on fish, frogs, and crayfish, and rest above water on hydrophytes or on land at other times. Most wetland-dependent snakes are viviparous, that is, bear live young, and therefore do not undertake migrations to terrestrial habitats to lay eggs, as do turtles. Many other snakes, however, and some lizards, live primarily in uplands but feed opportunistically in wetlands (e.g., king snakes (*Lampropeltis*), rat snakes (*Elaphe*), and garter snakes (*Thamnophis*)).

Several large groups of birds occur nearly exclusively in wetlands. These include many carnivorous (mainly fish-eating) species, such as the loons (Gaviidae), herons and bitterns (Ardeidae), several waterfowl species (Anatidae), the kingfishers (Alcedinidae), terns (Sternidae), and several raptors (Accipitridae, e.g., ospreys (*Pandion haliaetus*)). Omnivorous groups include the grebes (Podicipedidae), many species of diving and dabbling ducks (Anatidae), cranes (Gruidae), rails (Rallidae), and gulls (Laridae). Primarily insectivorous species include the shorebirds (Charadriidae) and many songbirds. Unlike other wetland inhabitants, birds are highly mobile and often use disjunct wetlands seasonally or even daily (**Figure 5**).

The last group of wetland-associated vertebrate animals, the mammals, include few, wholly wetland-dependent species. Most prominent are rodents such as the muskrat (*Ondatra zibethicus*) and beaver (*Castor canadensis*). Other obligate wetland mammals include several primitive, carnivorous shrews



Figure 5 Foraging assemblage of scarlet ibis and barefaced ibis in the seasonally flooded grasslands of central Venezuela (Apure state). Photo by Mark Gregory.

("water shrews" in the genus *Sorex*), some lagamorphs such as the swamp rabbit (*Sylvilagus aquaticus*) and marsh rabbit (*S. palustris*), and mustelids such as the river otter (*Lutra canadensis*). Wetlands are a critical resource, however, for many otherwise terrestrial species, which use wetlands for feeding and cover, such as moose (*Alces alces*), raccoon (*Procyon lotor*), and mink (*Mustela vison*).

Animal communities of saltwater wetlands can differ substantially from those in freshwater systems because of the effects of salinity on animal metabolism. Generally speaking, analogous taxa occur in freshwater and saltwater wetlands, such that similar niches are occupied, but species richness of animals generally declines along a freshwater-brackish-saltwater gradient. Reptiles from the Gulf of Mexico region provide an example: about 24 species occur in fresh marsh, 16 in brackish marshes, but just 4 in salt marshes. Of the salt marsh species, few are salt marsh specialists. Some terrapins are largely restricted to salt marsh habitats, but other regularly encountered salt marsh inhabitants, such as cooters and alligators, actually prefer freshwater situations. Amphibians show a particularly sharp trend along the wetland salinity gradient. Owing to their highly permeable skins and inability to counteract the drying conditions produced by high osmotic pressures of salt water, very few species can survive in saline wetlands and even then occur only infrequently (e.g., some toads).

Relatively few breeding birds specialize on salt marsh habitats, although sparrows and rails are well represented. However, salt marshes are noted for their importance as stopover sites for the vast populations of migrating species that breed in freshwater wetlands that freeze in winter. Waterfowl, shorebirds, songbirds, and owls can occur in brackish and salt marsh habitats in large numbers during the colder months resulting in very diverse, if unstable, assemblages during the winter period. Thus, though occupied for relatively short periods, and not by a large number of breeding species, coastal marshes nevertheless are a critical habitat in the annual cycle of many bird species.

Mammal communities in saltwater wetlands are generally quite species poor. However, the occasional effects of high densities of aquatic rodents (e.g., muskrats and nutria) can

be of ecological significance. As in freshwater systems, these species can radically alter the wetland environment through so-called "eat-outs" of the vegetation. Otherwise, saline wetlands are infrequently visited by species more dependent on freshwater wetlands, such as otter, mink, or raccoon.

Overlap of freshwater and marine species, in conjunction with resident species that can tolerate the substantial fluctuations in salinity, produces highly complex, species-rich assemblages of fish and crustacean species in saltwater wetlands. Though characteristically species rich, these assemblages are temporally quite unstable because of daily tidal cycles, with freshwater-associated species descending into salt marsh systems on outgoing tides and marine species arriving on incoming tides. In particular, many marine fishes and crustaceans, especially shrimp, use salt marshes as nursery areas, particularly for post-larval and juvenile forms. Here they take advantage of the lower numbers of predators in salt marshes relative to the open ocean and have access to the high productivity of salt marshes.

Controls of Species Diversity

Within particular wetland ecosystems, plant and animal diversity reflects the interacting influence of a number of abiotic and biotic variables that operate across a range of spatial and temporal scales. Diversity in wetlands cannot be understood without attention to these gradients, the scales at which they operate, and their interactions. Of the physical and chemical variables, hydrologic regime, mineral ion concentrations, pH, nutrient availability, salinity, and disturbance have been shown to play particularly strong roles. They largely determine the number of species that are physiologically capable of living in a site. Important biotic variables include primary productivity and interspecific competition. Along with abiotic factors, they influence which plant and animal species actually occur within the site. Because the composition, structure, and density of vegetation form key elements of habitat for animal species, these gradients also strongly influence animal diversity.

Depending on the type of wetland, different abiotic and biotic variables assume greater or lesser importance in determining species distributions. Hydrologic regime is important in all wetlands for both plant and animal species. Saturation with water severely restricts diffusion of oxygen into the soil while microbes rapidly deplete available soil oxygen. Hydrologic regime, therefore, largely determines the distribution and frequency of the most universal stress to which wetland plants and animals must adapt—anoxia in the water column, shallow soils, and rooting zone of plants. Salinity generally assumes importance only in coastal wetlands and in the saline wetlands of semiarid and arid regions. However, inland salt marshes occur in the Great Lakes region and some of the wetlands of the prairie pothole region of North America are saline. Mineral ion concentrations and pH affect species composition in all types of wetlands but have particular significance in peatlands that exhibit both extremes of the gradient in these factors. The combination of low nutrient availability, low pH, and low concentrations of calcium and magnesium makes bogs inhospitable environments for most plant and animal species. The extremely high pH and calcium concentrations of some rich fens are associated with high species richness and a large number of rare and uncommon species of plants and snails. Nutrient availability, which is lower in all undisturbed peatlands than in mineral soil wetlands, has become an issue for many types of wetlands as humans have increased nutrient inputs to wetlands.

Flowing water, especially at faster rates and larger volumes, exerts an additional mechanical stress on wetland species in some environments. Plants growing in river floodplains, tidal marshes, and other wetlands subject to flooding with rapidly flowing water must be able to withstand the force of moving water or must have life histories that allow them to germinate and reproduce in periods between floods. Well-developed root systems and flexible stems are essential for perennial plant species. Animals in these environments must be able to move out of the reach of flood waters, either into sediments or the tops of trees or out of the flooded area.

Many types of disturbance—factors that remove plant biomass—affect wetland diversity. Extended periods of inundation with deep water kill most emergent plants, which allows new species to germinate when water levels drop. Muskrats and nutria eat marsh vegetation. Moose consume wetland shrubs and aquatic vegetation. Beavers fell trees, alter hydrologic regimes with their dams, and destroy existing vegetation in digging for mud to strengthen the dam. The flooding created by beaver ponds destroys flood-intolerant vegetation. Tides and storm surges deposit patches of debris that bury vegetation in coastal wetlands. Hurricanes rip trees out of swamps and deposit sediment and other debris in wetlands. Humans harvest timber from swamps in many parts of the world. In Africa, southeastern Asia, and parts of South America, humans harvest other wetland plants for various uses, including firewood, food, and shelter. Reeds (*Phragmites communis*) are still harvested in The Netherlands and England as thatching for roofs.

Change over time as well as space is the rule in wetlands. Because climate drives hydrology, seasonal, annual, and longer-term changes occur in water depth and all the chemical and physical factors it regulates. Extended periods of wet or dry years dramatically change the distribution of both plant and animal species. Some wetlands systems, such as the prairie potholes of North America, have 10- to 20-year cycles of wet to dry climate. In other regional wetlands, such as the deep-water cypress swamps of the southeastern United States, wet to dry cycles occur over much longer time periods. The effect of this temporal variation at many scales is to increase overall species diversity by promoting a dynamic community composition and spatial distribution among years. In wetlands where organic matter accumulates as peat, even more dramatic changes occur with time; the landscape itself is transformed. Over hundreds to thousands of years, a mineral-rich, deep-water pond with high pH and relatively high nutrient availability can change to a raised bog with low pH, low nutrient availability, and low mineral ion concentrations (**Figure 6**). The plants and animals of a pond are replaced with those of the bog. Long-term deposition of mineral sediments in estuaries can create similarly dramatic changes in topography, landform, and the hydrologic regime for wetland organisms.



Figure 6 Raised bog, coastal Maine, USA.

Controls of Plant Species Richness

The question of what controls patterns in plant species richness is an old one that has received renewed attention with observed declines in species diversity over much of the globe. In wetlands, the question has been resolved only in part. The effects of many single factors are relatively well understood because of the over-riding importance of some physical and chemical variables in determining the pool of species physiologically capable of living in a site. Spatial and temporal heterogeneity, long recognized as critical determinants of species richness in other systems, are also positively correlated with species richness in wetlands. However, the relative importance of interacting sets of variables, operating in different types of wetlands and over different spatial and temporal scales, is only beginning to be understood. Because humans have increased nutrient inputs to many wetlands, the question of the relationship of plant species richness to nutrient availability, community productivity or biomass, and competition continues to motivate many wetland studies. Controls operating at regional and global scales have been little studied.

Gradients in several physical and chemical factors show strong correlation with species richness in wetlands, especially at the ends of the gradients where stress on plant growth is high. While more than 6700 species, approximately one-third of the total United States vascular flora, grow in wetlands, few species have evolved to persist in wetlands with low pH, high salinity, continuous inundation with shallow water, or high flooding frequency. Hence, bogs (low pH), the low marsh zone of salt marshes (high salinity), deep-water marshes (continuous inundation), and wetlands in the active channels of rivers (high flooding frequency) are all species-poor. In general, species richness increases with decreasing stress due to abiotic factors.

Numerous studies have shown that water level fluctuations are critical to maintaining species diversity in a wide variety of wetlands. Alternating periods of inundation and exposure of the soil create temporal and spatial gradients of water depth that provide conditions for germination of the largest number of species. Species diversity always declines when water levels are artificially stabilized. Several studies have shown that coastal plain ponds, Great Lakes shoreline vegetation, prairie wetlands, and Carolina bay wetlands all require fluctuations in water depth to maintain overall species richness.

Spatial heterogeneity increases the number of species co-occurrences in wetlands through a number of mechanisms. The presence of hummocks and hollows in peatlands creates small-scale gradients in pH and moisture that allow species with a wider range of tolerances to coexist in relatively small areas. The presence of mounds and tree bases in floodplain swamps provides sites for germination of a number of species that do not germinate or persist in standing water. Old muskrat mounds serve the same function in cattail marshes. Microtopographic variation in riparian wetlands causes spatial variation in flood frequencies that results in higher species richness. Deposition of wrack (dead plant debris) by tides and storms increases species richness in salt marshes by producing openings in dense stands of *Spartina*; these openings have higher salinity than the surrounding vegetation and are colonized by more salt-tolerant species (e.g., *Distichlis spicata*). High variation in substrate pH, type, texture, and age creates microhabitat variation along arctic, boreal, and tropical rivers and increases the number of plant species occurring in riparian vegetation.

Some of the spatial heterogeneity associated with higher species richness in wetlands results from various types of disturbance. In general, richness first increases and then decreases along gradients of increasing disturbance so that richness is highest at intermediate points along the gradient. In many environments, lack of disturbance allows a few highly competitive species to displace other species through the process of competitive exclusion. Disturbance disrupts this process and allows other species to persist. However, richness decreases if disturbance is too frequent because few species are adapted to very high disturbance frequencies.

Fertility, or nutrient availability, appears to be another strong control on plant species richness. However, few studies have actually measured nutrient availability because of the difficulty of doing so. Rather, nutrient availability has been inferred from surrogates such as community biomass, percent soil organic matter, or substrate texture. Use of these surrogates generally suggests that species richness is highest at moderate levels of community biomass and presumably, therefore, at moderate levels of nutrient availability. Close inspection of the data shows that richness can be low or high at low to intermediate levels of community biomass but almost always decreases above some threshold of high biomass. The amount of biomass at which that threshold occurs, however, varies greatly among wetland types. Furthermore, the relationship between richness and biomass appears to be scale-dependent. When data from wetlands with different vegetation types are examined, richness follows a unimodal relationship with community biomass. Within small plots within types, the relationship breaks down.

The complex interplay of abiotic factors with biological variables such as competition and productivity has been examined in only a few freshwater wetlands and seldom through experimental manipulation of variables. Salt marshes have received more attention. Results from these studies emphasize that the relative importance and direction of the effect of biotic variables on species richness change along gradients of stress. Abiotic variables play the strongest role where stress is high but the presence of some species can ameliorate the stress for other species (e.g., by lowering salinity through shading,

which decreases concentration of salts due to evapotranspiration). Species richness is thereby increased above what abiotic conditions might otherwise allow through positive interactions among species. Nonetheless, the number of species tends to decline monotonically with increasing stress from abiotic factors. Where stress in minimal, biotic control of species richness is greater. Here, variables related to density effects (e.g., productivity, biomass, light) are more likely to control variation in species richness, probably through competitive effects of dominant species on subordinate ones. In such environmentally benign wetlands, productivity is likely to be high and a few species are likely to dominate unless the process of competitive displacement is disrupted by disturbances that remove plant biomass and allow other species to establish.

The interplay of these local biotic and abiotic factors is played out against the background of regional and historical controls on the species pool (i.e., the number of species available to colonize a site). While local and current conditions determine which plant species do colonize a site, the regional species pool sets the upper limit to the number of species that could colonize the site. The geological, evolutionary, and ecological history of a region, along with its area, controls this pool. Geological age, processes of speciation and extinction, geographic dispersal, and chance historical events all leave their imprint on the species pools of particular regions. Few attempts have been made to relate the species pool to the number of species actually co-occurring in a site because of the difficulty of obtaining adequate data sets.

Controls of Animal Species Diversity

Patterns of diversity and productivity in wetland animals follow gradients strongly related to hydrologic regime, primarily because of its effects on the composition, structure, and productivity of wetland vegetation and on oxygen availability within the water column and wetland sediments. Oxygen levels most influence the distributions of wholly aquatic, nonmigratory wetland animals (i.e., the macroinvertebrates and fishes). The productivity and diversity of wetland invertebrates and fish, and, consequently, many water birds and mammals, are highest in semipermanent wetlands (e.g., those flooded during most of the growing season but dry otherwise). Most amphibians, however, occur in transient pools. Trends in reptile diversity and productivity along this inundation gradient are not clear.

Semipermanent and permanent water bodies are the primary habitats of fish, which generally lack the ability of many other wetland inhabitants to disperse overland or enter dormancy when wetlands dry. Thus, areas of sufficiently oxygenated water must occur in wetlands throughout the year to support fish populations, especially those of large-bodied species. Large, seasonally flooded wetlands with pools or channels that remain permanently flooded, and lakes with well-developed littoral vegetation tend to support the most diverse fish communities. Similarly, fish-eating birds and mammals frequent the habitats occupied by their fish prey. In contrast, temporary, unvegetated pools are the primary breeding habitats of amphibians, offering two main advantages. First, amphibian larvae are able to exploit a pulse of primary productivity that occurs shortly after pools fill with water. Second, temporary pools will not support fish, major predators on amphibian eggs and larvae.

Many other wetland animals exploit intermittently flooded wetlands. Waterfowl, particularly egg-laying females, frequently travel among shallow pools, ditches, and puddles to feed on the pulse of invertebrate production that occurs following flooding. These feeding pools often are remote from nesting areas at more permanent wetlands, where deep waters and dense vegetation protect nesting birds and their eggs from predators. Similarly, snakes and turtles often can be found in temporarily flooded wetlands, and seek refuge from subsequent low-water periods by migrating or entering dormancy in a wetland's sediments.

Much of the variation in cover and food for wetland fauna, and hence in animal diversity, is associated with variation in the species diversity, density, and structural characteristics of wetland plants. Many wetland animals show affinities for particular life-forms of hydrophytes (e.g., submerged plants, floating-leaved plants, or emergents), and some are associated with specific plant taxa. For example, most aquatic invertebrates seek protection from predation and safe oviposition sites in macrophytes. In general, plants with dissected leaves support more invertebrates than plants with broad leaves. Macroinvertebrate diversity and density generally are greater in vegetated than open water areas, and, within vegetated areas, are greatest in areas of mixed emergent and submerged vegetation. Shading by erect hydrophytes also influences macroinvertebrate abundance; unshaded areas of submerged vegetation often support dense populations owing, in part, to elevated water temperatures in unshaded areas. Similarly, shallow, vegetated wetland habitats fringing deep-water areas are critical nursery habitats for fishes, many species of which use stands of particular wetland plant species as spawning habitat. These areas provide protective cover for eggs and larval fishes and food resources for young fishes. Like aquatic insects, many fishes prefer unshaded, shallow areas where warm water temperatures hasten the development of eggs and larvae.

Two key water chemistry parameters that influence distributions and diversity of wetland animals are acidity and salinity. The relationship of acidity to the distribution of wetland animals can be complex. For example, many aquatic, fish-eating birds frequent wetlands with moderately acidic waters (pH 5.5–6.0). Although fish populations often are reduced in wetlands with waters of pH <6.0, suspended material may precipitate from the water column at low pH. The resulting increase in water transparency may improve encounter rates with prey and capture success for birds that pursue their prey underwater, and thereby compensate for decreased fish density. The reduction or loss of fish from moderately and strongly acidified waters also reduces competition between fish and insectivorous birds for macroinvertebrate prey. This may explain why dabbling ducks frequent waters of relatively high acidity that fish-eating ducks avoid. Highly acidic waters (i.e., pH <5.0) generally support few fish or amphibians and have low diversity of aquatic invertebrates. Water bird and mammal use of such wetlands is limited.

Salinity strongly limits animal diversity in wetlands, primarily through effects on the composition of plant and invertebrate communities and through its direct physiological

effects on wetland animals. Difficulty in maintaining water balance leads many dabbling ducks, particularly those brooding young, to avoid wetlands with highly saline waters. Owing to their highly permeable skin, amphibians are intolerant of even moderately saline conditions. For similar reasons, highly saline conditions often limit the survival and hatching success of fish eggs and survival of larval fish. Many mammals, such as muskrats, are tolerant of highly saline conditions, but populations often are limited by the low quality of forage plants that grow in saline wetlands.

Three spatial characteristics of wetlands and their distribution in the landscape—size, connectivity, and isolation—also influence the diversity of wetland animals within and among sites. For example, small wetlands often support fewer species of birds than large wetlands, and isolated wetlands support fewer species than wetlands in complexes. Aquatic, fish-eating birds are particularly sensitive to wetland area and often shun small wetlands (<10 ha). Small wetlands may not provide a sufficient quantity of habitat or prey for many large-bodied species, and short inter-wetland distances may provide certain species with alternate foraging sites while minimizing their time in flight. Larger wetlands generally support a wider diversity of vegetative life-forms and, therefore, may provide a greater variety of habitats for wetland species than smaller wetlands. This may account for the frequently observed pattern of increased fish species diversity in larger than smaller wetlands.

The proximity of diverse habitat types within a wetland, and connectivity among wetland environments and with upland habitats, is critical to maintaining high diversity of wetland animals, particularly amphibians, some birds, and fish. Aquatic breeding and terrestrial nonbreeding habitats must be contiguous for amphibian populations to persist. Local breeding populations undergo frequent extinction and recolonization events, and individuals often are exchanged among populations. The importance of migration in amphibian population ecology is indicated by the presence in many species of a juvenile stage adapted strictly for dispersal. Destruction of larval or adult habitats, or the connections between them (e.g., riparian corridors) often drives local amphibian populations to extinction.

Proximity of diverse habitat types within a wetland is critical to many breeding water birds and fish. Some water birds, such as rails, dwell within dense stands of emergent vegetation, while others, such as bitterns, herons, and shorebirds, forage mainly along the interface of wetland vegetation and open water. Still others, such as grebes and many dabbling ducks, are associated with floating and submerged vegetation, while loons and cormorants forage in deep, open waters. Most of these birds, however, build their nests within dense stands of emergents, on small islands, or in trees or shrubs.

Fish assemblages provide another example of the importance to wetland animals of spatial heterogeneity and connectivity in wetland habitats. Most large species of bottom-feeding and predatory species undergo small-scale migrations between feeding and spawning areas in shallow wetlands and resting areas in deeper, open-water habitats. These migrations often occur nightly, with fishes returning to deep waters to rest during the day. Critical to these migratory movements are the links established between shallow, vegetated areas and deep-water zones by seasonal regimes of wetland flooding. Finally, many fishes undergo seasonal migrations over hundreds of kilometers into river floodplains and return to the river channels as waters recede.

Some wetland animals create some of the spatial and temporal heterogeneity that promotes animal diversity in wetlands. Beavers alter stream channels and transform riparian forests into freshwater marshes and ponds with stands of dead timber. Beaver activities also permanently modify soils beneath dam sites and even change local topographies; beaver dams reduce current velocities and cause massive silt depositions, which remain long after the dams disintegrate. High levels of herbivory by muskrats and geese can result in "eat outs," whereby essentially all emergent plant material in a marsh is consumed over a few years. Such drastic, animal-caused modification of the wetland environment not only affects habitat availability for other wetland animals but can alter patterns of energy flow and biotic productivity within wetland ecosystems. Other, less drastic examples of habitat modification include the activities of alligators, which construct small basins that serve as critical refuges for many fishes, and important feeding areas for fish-eating birds, during low water periods (**Figure 8**). Crayfish burrows serve in many areas as moist refuges for water snakes, salamanders, and frogs.

Functional Diversity

Wetland plants exhibit a diverse array of adaptations that allow many species to tolerate stresses associated with soil saturation, alternating periods of flooding and drought, high salinity, low pH, and low nutrient availability. Other adaptations assist wetland plants in dispersal and establishment. Animals also exhibit an array of adaptations to wetland environments, although unlike plants, many simply can move to escape stresses. Collectively, these adaptations explain much of the high species richness and diversity of many wetlands, especially those with strong hydrological and chemical gradients.

Adaptations to Anoxic Environments

Many wetland plants have one or more morphological and anatomical adaptations that allow them to tolerate soil saturation and anoxia for short to long time periods, primarily by allowing more oxygen to reach the plant root system. Key morphological adaptations include (a) aerenchyma, air spaces in roots and stems that allow oxygen diffusion from stems above water to roots; (b) hypertrophied lenticels, enlarged openings in stems and roots that allow gas exchange between internal plant tissue and the atmosphere; (c) adventitious or stem roots developed above the water line; and (d) the ability to grow new roots under anoxic conditions. In black mangrove (Avicennia), pneumatophores, vertically growing air roots, absorb oxygen that is transported to the connected, submerged, lateral growing roots. Prop roots of red mangrove (*Rhizophora*) function in much the same way. In both mangrove species, oxygen enters the plant through lenticels exposed above water and diffuses to roots in anoxic sediments.

Many woody species of alluvial floodplains have extensive, shallow root systems placed where sediments are least likely to experience oxygen deficits.

Plant physiological adaptations generally involve tolerance to low soil oxygen and specialized chemical reactions. Specialized reactions include an accumulation of malate instead of ethanol, the production of high levels of nitrate reductase, and a reduction in ethanol production by reducing alcohol dehydrogenase activity. Additionally, rhizosphere oxygenation is an important mechanism for many species and can lessen toxic concentrations of some anaerobic soil compounds.

Wetland animals, with their characteristically high metabolic rates, have developed a variety of adaptations to low levels of oxygen and carbon dioxide. Perhaps most obvious is development of specialized regions of the body for gas exchange. Examples include gills in fish and crustacea, parapodia in polychaetes, and highly vascularized tissues on the lower lips of some tropical fishes or in the cloacas (urogenital openings) of turtles. Other physical adaptations include modification of respiratory pigments to improve oxygen-carrying capacity in invertebrates. For example, midge larvae (*Chironomus*) are often colored brightly red, indicating the unusually high concentrations of hemoglobin present in these organisms, permitting them to survive long periods of hypoxia. Some vertebrates, particularly fishes, also increase densities of circulating red blood cells and thereby their oxygen-holding capacity.

Physiological adaptations of animals primarily involve shifts in metabolic pathways. For example, bivalves use alternative biochemical pathways, primarily a switch to glycolytic fermentation, to increase energy production under anoxic conditions. Some invertebrates also diversify the by-products of glycolysis to avoid toxic accumulation of any single compound, particularly ethanol. Turtles are remarkable for the ability of these lung-breathers to remain under water submerged in sediments for months during the winter season. Biochemical adaptation for natural anoxia tolerance in turtles includes well-developed antioxidant defenses that minimize or prevent damage by reactive oxygen species during the reoxygenation of organs after anoxic submergence. Last, many invertebrates store large quantities of respirable carbohydrate, usually glycogen, for breakdown and oxygen-liberation during periods of anoxia.

Behavioral adaptations also are critical and widespread, including dormancy or low locomotor activity during periods of oxygen stress, and migration from hypoxic to oxygen-rich environments. Many animals in low-oxygen situations have developed means of moving water more rapidly across respiratory surfaces. For example, benthic animals often use a variety of behavioral means (fanning, retreating into and out of burrows) to ventilate their burrows and increase the water flow across membranes during times of hypoxia.

These adaptations combine in interesting ways in particular taxa. Some fishes highly adapted to mud and shallow-water swamps cope by aestivating in mucous cocoons or by migrating overland while air breathing through modified swim bladders (e.g., lungfish). Some aquatic microinvertebrates, such as Collembola (spring-tails) and mites (Acari), have developed a "physical gill" that traps air gathered at the water surface in body surface hairs. Breathing of the trapped air, while underwater, occurs via a tracheal system, which opens to the body surface. Some aquatic insects, including mosquito larvae and chrysomelid beetles, tap the air within the aerenchyma of plant roots using a highly specialized, spinelike siphon attached to their abdomens. Last, some fly larvae use snorkel-like devices that extend above the surface of liquid mud or anoxic water and that permit the animal to air-breathe while remaining submerged in the anoxic substrate.

Adaptations to Salinity and Drought

Although halophytes (i.e., plants capable of persisting in saline environments) are not restricted to wetlands, these species dominate saline wetlands, such as inland and coastal salt marshes, and coastal fringe forests. Some species, such as various mangroves (*Avicennia*, *Laguncularia*, and *Rhizophora* spp.), are facultative halophytes and have mechanisms to inhibit salt absorption by their roots and to excrete salt from their leaves. Because nonhalophytes ("glycophytes") exhibit a range of tolerance to a salinity gradient, salinity and flooding regime determine plant species composition and richness along a gradient from salt marsh to tidal freshwater marsh. Plant species that use the C_4 pathway of photosynthesis are well adapted to habitats subjected to drought stress. Saline wetlands, therefore, often are dominated by C_4 plant species because the salt content makes water uptake more difficult. Plant species with C_4 photosynthesis are much more efficient in fixing carbon and in water use and have higher rates of net photosynthesis, growth, and dry matter production than C_3 plants.

Little is known about adaptations of wetland animals to salt stress. The fundamental problem of salt stress is its effect on water movement at the cellular level, which interferes with metabolic processes. Most simple microinvertebrates are osmoconformers whose internal salt concentration tracks that of the external environment. More complex wetland animals, particularly vertebrates, typically osmoregulate—that is, retain internal salt concentrations independent of those external to the body. Osmoregulation primarily involves moving ions across body membranes against the concentration gradient and is thus an energy intensive process. This process is usually associated with specialized renal organs (kidneys and antennal glands) and salt-secreting nasal or rectal glands. Behavioral adaptations, particularly short-term migrations, are common in mobile species. In such a manner, both freshwater and marine species can occur in salt marshes, retreating during unfavorable periods either upstream (freshwater species) or downstream (marine species) during the course of the daily tidal cycle.

Adaptations to Low Nutrient Availability and High Acidity

Nitrogen fixation is relatively uncommon in wetland vascular plants but the few examples are noteworthy. Nitrogen fixation by sweet gale (*Myrica gale*) can contribute 3 to 4 g N m^{-2} yr^{-1} to *Sphagnum* bogs. Nitrogen fixation rates of speckled alder (*Alnus rugosa* = *A. incana ssp. rugosa*) have been measured at 85 to 167 kg N2 ha^{-1} yr^{-1}. Some nonsymbiotic, aerobic and anaerobic bacteria, and blue-green algae are also significant nitrogen fixers in wetlands (**Figure 7**).



Figure 7 Acidic peatland, coastal Maine, USA.



Figure 8 Alligator in pool at Everglades National Park, southern Florida, USA.



Figure 9 Yellow pitcher plants in pocosin, South Carolina, USA.



Figure 10 Acidic peatland in the Adirondacks, New York, USA.

Most of the carnivorous species in the world occur in wetlands and other oligotrophic habitats. The relationships between insect capture, mineral nutrition (particularly nitrogen and phosphorus), and plant growth and reproduction have been investigated for numerous species. The relatively large diversity of organisms that can survive in the "pitchers" of some carnivorous plants also has been studied (**Figure 9**). Although benefits to carnivorous plants have been demonstrated, there is limited evidence that carnivory is necessary for most species to survive.

Evergreen, sclerophyllous shrubs that generally belong to the Ericaceae (heath family) are characteristic of nutrient-poor peatlands. The role of this foliage type in wetlands is uncertain, but it has been suggested that plant species with evergreen leaves may be more efficient in using nutrients than species with deciduous leaves. These leaves, which persist for up to four years, extend the period that the plant can photosynthesize, which conserves energy in the plant because all of the leaves do not have to be renewed each year. Furthermore, evergreen species tend to resorb more nutrients from senescing leaves than other species, thus increasing nutrient use efficiency. High nutrient resorption and long retention of nutrients in evergreen sclerophyllous leaves may be especially important in phosphorus-poor habitats such as bogs.

One of the most important groups of plants in colder, generally acidic wetlands of the world are *Sphagnum* mosses, a large genus consisting of more than 135 species (**Figure 10**). These mosses can thrive in conditions highly stressful to most plants (e.g., low pH, low nutrient availability, and saturated soils). They are the primary peat producers worldwide and have the capacity to modify the environment in ways that favor their own growth. Once established, *Sphagnum* mosses can further acidify their environment through cation exchange processes and by excretion of polygalacturonic acids. They

also paludify their environment through their high water-holding capacity and their role in peat accumulation which can block water drainages. *Sphagnum* species also can tolerate great desiccation. Three major gradients exert the greatest influence on the bryophyte flora of peatlands: acidity–alkalinity, dry–wet, low light-high light. Numerous *Sphagnum* species can occur even within a relatively small peatland because of the affinities of particular species to very specific pHs, other water chemistry variables, height above water table, and light.

Adaptations to Reproducing in Wetlands

Seeds of many wetland species have features that enable their dispersal by water (hydrochory). Red mangrove (*Rhizophora mangle*) produces seeds that germinate while on the parent tree, resulting in viviparous seedlings. The germinated seed can root if it lands on sediment. Otherwise, the seedling floats until a suitable substrate is reached.

Seed banks, the store of seeds that accumulate in soils, play a critical role in the long-term vegetation dynamics of many types of wetlands worldwide. Seed banks allow wetland vegetation to adjust to rising and falling water levels and help maintain compositional diversity across water depth gradients and from year to year and over longer climatic cycles. Most wetland species do not germinate under water and only a few survive more than a year or two of continuously deep water. A series of wet years can eliminate most vegetation, but when water levels fall, many species that existed only as seeds in the seed bank germinate in the wet exposed mud. Thus, fluctuating water levels and seed banks drive wetland succession. Seed banks are particularly important in maintaining the species composition and zonation of wetlands that have a large number of annual species (e.g., tidal freshwater marshes). An understanding of the seed bank is essential in wetland management, especially in restoration efforts.

Seed density in wetland soils, usually determined by germination tests, typically ranges from 1000 to 14,000 seeds $m-2$ with density and viability decreasing with increasing depth in the substrate. Graminoids (grasses and sedges) are often the most abundant group in wetland seed banks, with seeds of broad-leaved herbaceous and woody species far less abundant. Although the composition of the seed bank and plant species present in the wetland can be highly correlated, in many cases the seed bank harbors a much greater species richness than found in the extant vegetation. Seeds of some species are transient, while others can persist for decades.

The greatly enlarged, rounded bases of water tupelo (*Nyssa aquatica*) (**Figure 11**) and buttressed bases of bald- and pond cypress (*Taxodium spp.*) develop especially in deeply flooded conditions. These swollen bases likely aid in anchoring the tree. The "knees" of *Taxodium* also may function in this manner. *Taxodium* buttresses appear to provide a refugium above water for numerous species of herbaceous and woody plants that become established in the crevices where organic materials have accumulated over time. In some deep-water swamps, alternative colonization sites are limited for many plant species.

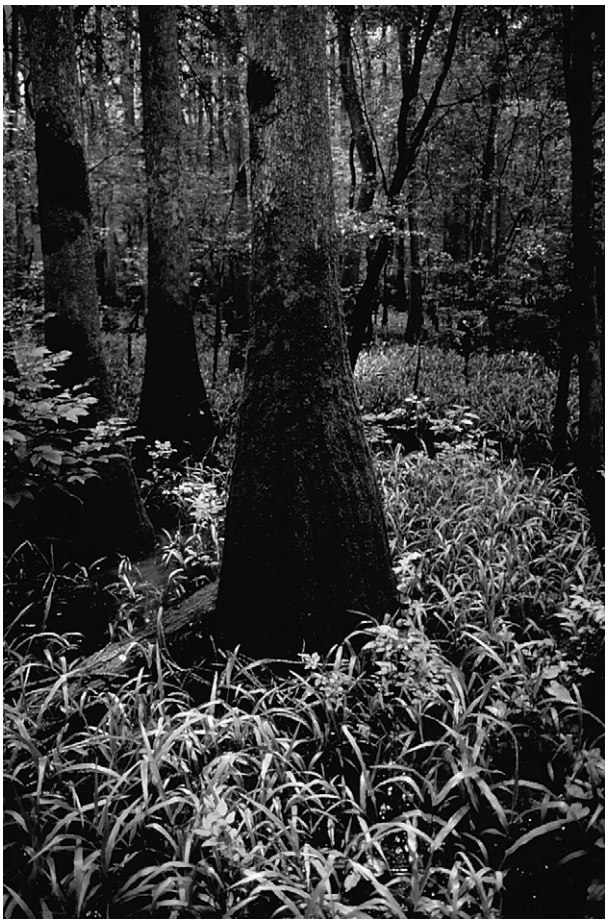


Figure 11 Water tupelo in southern deepwater swamp, South Carolina, USA.

Threats to Wetland Diversity

Worldwide, the loss of wetlands is estimated at about 50%. Functionally, the losses are much greater because extensive areas of boreal wetlands have been little altered while many different types of regional wetlands have been reduced by 80 to 90% elsewhere (e.g., California, southern Ontario, southwestern France, New Zealand). Wetland diversity and numerous plant and animal species have been impacted by a host of anthropogenic activities and natural disturbances for centuries. While single and cumulative effects of these activities have been assessed for many wetland species (primarily wildlife) and functions, few studies have examined the effects of these activities on species diversity. Because human activities often modify wetland hydrology and water chemistry, one should expect significant changes in diversity given the well-established relationships between wetland hydrology, water chemistry, and species distributions.

Despite the relatively recent appreciation of wetland functions and values, wetlands continue to be destroyed. A notable example is the Pantanal of Brazil, the world's largest wetland and home to many endangered or threatened plant and animal species. Various development plans, including

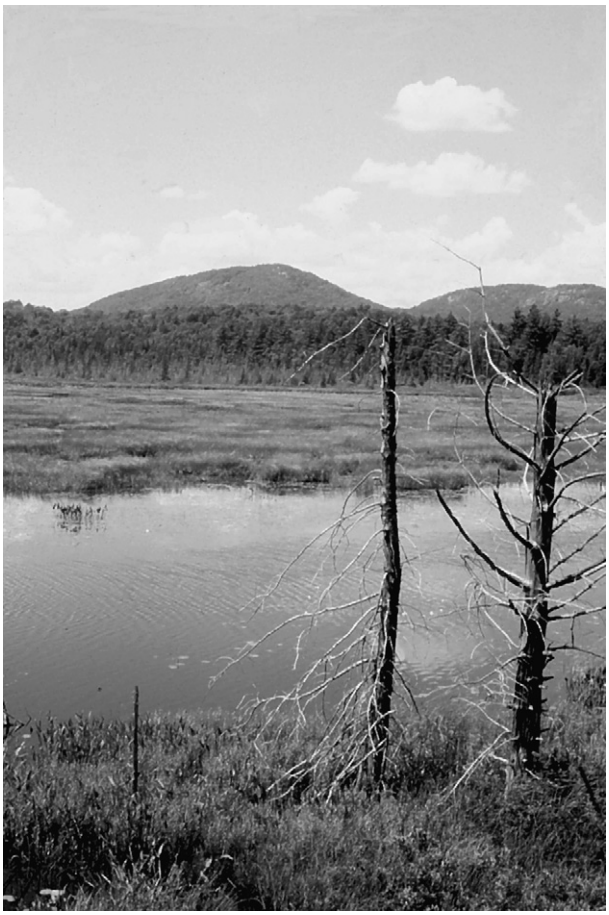


Figure 12 Freshwater marsh in the Adirondacks, New York, USA.

dam and highway construction, threaten the diversity of this vast wetland. Although some human activities may be able to be carried out in a sustainable manner in wetland habitats, habitat preservation is the key to preserving wetland diversity around the world.

Direct Habitat Destruction

By far, the majority of palustrine wetland loss in North America has been due to agricultural activities, especially direct conversion of wetlands to agricultural land. Agricultural uses have converted over 80% of the Mississippi River bottomland hardwood forest. About 90% of the wetlands in California, Ohio, and Iowa have been destroyed, primarily for agriculture. Wetlands adjacent to agricultural lands also can be adversely affected because drainage effects can extend a significant distance into the wetland. Coastal wetlands, most first altered by agricultural activities, including grazing and salt production, are now impacted by draining and filling for urban and industrial development. Drainage of wetlands on organic substrates leads to accelerated decomposition, erosion of organic matter, and release of carbon dioxide (a "greenhouse" gas) into the atmosphere.

Riparian wetlands in the western United States can be seriously degraded by cattle grazing, which adversely affects the adjacent aquatic ecosystem. Wildlife species composition and diversity are modified because heavy grazing reduces the structural complexity of riparian wetlands.

Timber harvest, when done without extensive drainage and total land clearing, generally converts forested wetland to earlier successional stages, or changes species composition and structure. Logging, specifically the extraction of cut trees through mechanical skidding and transport of heavy log loads, adversely modifies soil physical properties. Early logging of southern forested wetlands was typically done with no regard for sustainability. Consequently, forest composition and structure were dramatically altered. Logging roads can modify wetland hydroperiods such that productivity is raised or lowered. Occasionally, especially in northern forested wetlands, a heavy timber harvest can raise the water table, flooding already established species and decreasing the regeneration sites for new plants. Whole-tree timber harvest of oligotrophic peatlands can greatly reduce a site's nutrient pool.

Tropical intertidal zones are being increasingly altered for mariculture, especially mangrove wetlands for shrimp farming. Many salt marshes have been ditched and diked, and often converted to impoundments, to favor waterfowl.

Peat extraction, primarily for fuel or horticultural uses, has destroyed millions of hectares of peatlands in North America, Great Britain, and northern Eurasia. Where the mined peatlands have been abandoned or actively managed for conservation values, the vegetation can slowly recover. Which species come to occupy the site depends on the hydrology of the site, availability of propagules, and time since abandonment.

Gravel and sand mining can adversely affect wetlands directly by habitat destruction, and indirectly by changing the hydrology and water chemistry of an adjacent wetland. These mining activities can be especially damaging to ombrotrophic and minerotrophic peatlands (e.g., bogs and fens) because many of the characteristic species occur at very specific water levels and chemistry parameters.

Direct Alteration of Wetland Hydrology

Humans alter wetland hydrology primarily by building dams at wetland outlets to regulate the depth and duration of inundation, and by constructing channels and ditches to drain wetlands and prevent prolonged flooding. Artificially regulated water levels often result in impoverished animal communities because few species of wetland animals have life histories adapted to the stable water regimes that result. This decrease in diversity occurs even where water control structures were originally built with the intention of benefiting wetland animals (e.g., at wildlife refuges). Management strategies for many water impoundments now seek to emulate the "natural" flooding regimes that were present before impoundments were built. Wetland drainage has obvious and drastic effects on wetland animal communities and has resulted in a loss of about half of the original habitat available to the wetland fauna of the United States. Among the most imperiled wetland habitats are the easily drained, intermittently flooded basins and pools that are critical to a large proportion of the wetland fauna.

Dam construction and water diversion projects also have significant effects on downstream riparian vegetation by altering peak and minimum flows, decreasing erosion and deposition of sediment, and lowering floodplain water tables. Altering the natural variability of river flow can adversely affect the species composition and structure of riparian wetland vegetation. Dredging, channelization, and levee construction have greatly impacted extensive wetland areas in the Mississippi River Delta.

Groundwater extraction is a primary cause of wetland modification because it lowers the regional water table and alters recharge and discharge patterns within the wetland and surrounding landscape. Irrigation with water supplied by aquifers that are recharged in part by wetlands has led to declines in the area of major regional wetlands in North America (e.g., the prairie pothole and Nebraska sandhills regions).

Although plant species of coastal wetlands may be well adapted to flooding, increases in the level or duration of flooding above that at which species have long persisted can cause serious deterioration of these wetlands. Sea-level rise and saltwater intrusion are important factors in the degradation and loss of wetland forests and coastal marshes in the southeastern United States. Increased flooding and salinity have been particularly damaging to forests dominated by baldcypress.

Landscape Fragmentation

Other human activities impoverish the wetland fauna in more subtle ways, even where wetlands receive nominal legal protection. Destruction of the uplands adjacent to wetlands destroys the habitat interface critical for amphibians and reptiles that migrate between wetland and upland habitats. Reductions in the connectivity among wetlands (e.g., by construction of roads and levees) blocks the migration routes of amphibians, fish, and birds, and can lead to local population extinctions. In contrast, dredging of canals that link major river systems causes sudden mixing of fish and reptile faunas that may have evolved in isolation for thousands of years.

Other Threats

Anthropogenic inputs of nutrients and toxins drastically alter the chemical environment of wetlands and render many sites unsuitable for wetland plants and animals. Atmospheric deposition of nitrogen and sulfur is particularly high in parts of western Europe but is also high in the northeastern United States. Atmospheric deposition may be especially damaging to peatlands dominated by bryophytes (mosses and lichens), because these species are very sensitive to changes in nutrients and acidity-alkalinity. Wetlands in agricultural landscapes receive excess nitrogen via ground water draining from agricultural fields and excess phosphorus primarily via surface water runoff. Residential development also contributes excess nitrogen and phosphorus to wetlands. Coastal oil spills can adversely affect the vegetation of salt marshes and tidal freshwater wetlands.

Invasive, exotic wetland species typically cause a substantial loss of native wetland species and greatly alter wildlife habitat. In otherwise arid regions, invasive species use substantial amounts of water and can desiccate watercourses. Invasive species generally respond favorably to altered hydrologic regimes, substrate disturbance, or changes in water quality, especially eutrophication. Introduction of exotic wetland animals, for example, carp, nutria (*Myocaster coypu*), and bullfrogs, outside their native range have greatly altered wetland environments throughout much of North America, mostly to the detriment of the native wetland fauna.

Purple loosestrife (*Lythrum salicaria*), an emergent, herbaceous species of Eurasian origin, has become a serious invasive species of open wetlands in eastern North America. A single individual of this species can produce an average of 2,700,000 seeds. Another very serious invasive species in wetlands of the eastern United States is the common reed (*Phragmites communis*), which can produce 200 to 300 culms m^{-2} in fresh and brackish wetlands through an extensive network of rhizomes. Recent biological control methods (i.e., leaf- and root-feeding beetles) appear promising for *L. salicaria*.

Salt cedars (*Tamarix* spp.) are rapid, arborescent invaders of riparian wetlands in the southwestern United States, especially downstream from dams where flooding is minimal. Another invasive tree species of riparian wetlands in the western United States is Russian-olive (*Elaeagnus angustifolia*). In southern Florida, the invasive tree, melaleuca (*Melaleuca quinquenervia*), has been aggressively colonizing shallow wetlands. Two other tree species that seriously threaten southern Florida wetlands are Australian pine (*Casuarina*) and Brazilian pepper (*Schinus terebinthifolius*). Glossy buckthorn (*Rhamnus frangula*) is becoming a serious problem in wetlands around the eastern Great Lakes basin.

Some native species become invasive following some hydrologic changes or an increase in nutrients and dominate wetlands previously occupied by other, more desirable species. Cattail is regarded as a threat to southern Florida marshes that have been dominated by sawgrass (*Cladium jamaicense*). The spread of cattail is believed due mostly to increased phosphorus input from agricultural lands and anthropogenic changes to the regional hydrology. Cattail is also considered a weed in many wetlands and aquatic habitats that are used for agriculture or to provide water resources (e.g., rice fields, irrigation canals, recreational lakes, and reservoirs). Mechanical (e.g., cutting, water level modification, fire, shading), chemical (selective versus nonselective herbicides), and biological control methods vary in their effectiveness, which varies widely among species. Maintaining water levels over 0.5 m is effective in controlling spread of cattail.

Natural Disturbances

Severe flooding, hurricanes, and fire are the major natural disturbances that greatly affect wetlands. Although wetland plant and animal communities are adapted to each of these disturbances, events of exceptional magnitude and duration, or that occur during unusual times of the year, can have profound effects. A serious concern following any of these severe events is the spread of invasive species.

During the 1993 flooding of the Mississippi River, submerged aquatic, emergent wetland, and floodplain species

were adversely affected by extreme sediment deposition, uprooting due to wave action, and probably increased inputs of agricultural chemicals. Recolonization of open spaces left by plant mortality can happen quickly by germination of local or upstream plant propagules.

Hurricanes especially affect the mangrove forests of the Gulf of Mexico and Caribbean, and the deep-water and alluvial swamps of the southern United States. These coastal forests, where many species have shallow root systems, are often subjected to the highest winds and succumb to windthrow. Trees that are especially resistant to windthrow (e.g., *Taxodium distichum*) may lose most of their crown but are able to reestablish a new canopy quickly by extensive sprouting along the bole. Seedling regeneration is generally higher after such disturbances because of increased light and microsite availability. Storm surges of saltwater occasionally cause greater damage than high winds to species intolerant of high salinity.

Forested wetlands on substantial peat deposits, such as the black spruce (*Picea mariana*) and tamarack (*Larix laricina*) swamps of the boreal forest and baldcypress swamps of the southeastern United States, have naturally burned during periods of extreme drought. In northern climates, fires can convert the forest to a more open peatland. In the southeastern swamps (e.g., Okeefenokee Swamp), fires can convert forested wetlands to prairie-like, tree-less wetlands. Fire can favor the regeneration of some tree species, like *Chamaecyparis thyoides* (Atlantic white-cedar), which regenerate best under full sunlight conditions on saturated peat. In the Everglades of southern Florida, open sawgrass marshes naturally burned every 1 to 5 years, ignited by lightning strikes from May to October when water levels are highest and substrates are saturated. However, human-set fires are most common from November to May when soils can be dry enough to ignite. These fires, to which the vegetation is not well adapted, can cause substantial ecological damage.

Restoring and Maintaining Wetland Diversity

Because of extensive loss and degradation of wetlands throughout the world, and continuing loss and degradation in much of the world, maintaining wetland diversity will require a strategy that includes restoration. The ultimate goal in any specific wetland restoration project is to have a self-sustaining, functioning ecosystem. The following discussion focuses on techniques to achieve this goal. However, in terms of restoring wetland diversity, the goal must be a broader one—best addressed at the landscape, regional, and continental scale—to restore the structure, function, and self-sustaining properties of a wide diversity of wetland types. Without attention to restoration of the full complement of wetland types, wetland diversity cannot be restored at any but the smallest scales.

Restoring self-sustaining wetland ecosystems initially, and at later times, may require various engineering, horticultural, and wildlife management activities (National Research Council, 1992). The intensity of these activities depends on whether the site was initially an upland (thus requiring wetland "creation"), a wetland that has been drastically disturbed (e.g., through conversion to agricultural crop production), or disturbed to lesser degrees. In some cases, a wetland may have reached a successional stage that does not support a species or group of species of concern, and manipulation is required to maintain these species. Nontidal emergent wetlands are generally easiest to restore. Long-term monitoring of site hydrology and vegetation is required to determine if and when project objectives are not being met.

Because wetlands serve many functions, and seldom can all functions be restored in a single site, restoration objectives must be clearly defined for every project. Restoring wetlands as habitat for rare and uncommon species, or for overall biological diversity, especially requires clear statements of objectives, as well as thorough knowledge of the ecology of the species of concern. Wetlands that could be correctly described as functioning wetlands have failed to attract the species for which they were restored. Consider the restoration of a bottomland hardwood forest. The restoration could be viewed as successful if a certain percentage of planted oak (*Quercus*) species and individuals were to become established after a given period of time. However, such a planting would be a failure in terms of restoring the diversity of neo-tropical migratory birds, which would fail to use the site because they require a more complex forest structure. In general, a higher overall plant and animal diversity is achieved by planting an array of fast-and slow-growing plant species, and where appropriate a mix of woody and herbaceous species tolerant of the range of hydrological and chemical conditions at a site.

The first step to restoring wetland plant composition, structure, and function is to ensure a range of hydrological conditions by implementing various techniques. Depending on the existing condition of the site and restoration goals, these techniques may include reestablishing stream flow or tidal fluctuations, restoring flood regimes, halting drainage, and reestablishing topography and microtopography. These conditions can be attained by excavating basins or channels, constructing dikes, grading an undulating topography, installing water control structures, plugging ditches, and removing drainage tile lines. If the site is to be used for the reintroduction of rare species, there are additional considerations, especially genetic, political, and legal issues.

Once a range of hydrological conditions is established, plant species best adapted to each specific hydrological condition can be selected. Many wetland restorations fail because the relatively few plant species selected grow poorly or die because conditions are too wet or dry for too long. If the goal of a particular restoration is to restore functions such as flood control or nutrient retention, then relatively few plant species need to be planted. If the goal of a project is to restore the natural heritage or educational values of a wetland, or to reestablish the compositional and structural complexity of highly diverse wetlands, then many plant species should be established.

Sources of native trees, shrubs, sedges, grasses, and herbaceous species include plant nurseries and other wetlands, especially those that are planned for future, permitted alteration. Seeds of many wetland species can be collected and grown for future transplanting or direct seeding. Some woody species can be established from stem cuttings. Many herbaceous species can be propagated from divisions off the parent. Spreading soil collected from a donor wetland, especially

those planned for future alteration, will provide the greatest diversity of plant and animal organisms. Numerous woody and herbaceous plant species can arise from transplanted soil, via whole plants, seeds, roots, rhizomes, and stolons. Including leaf litter, detritus, and, for forested wetlands, large amounts of coarse woody debris are beneficial.

Herbivores cause many wetland restoration projects to fail, directly (by feeding) and indirectly (by burrowing, by dam building). Invasive species, such as *Lythrum salicaria*, *Phragmites australis*, and *Tamarix* spp., also must be controlled if plant and animal diversity are desired.

Understanding the ecological factors that control the diversity of animals in wetlands is key to altering these factors to restore, enhance, or create wetlands for the benefit of wetland fauna. Because hydrology underpins so much of wetland ecology, manipulation of wetland hydrology through water-level regulation is the most common and cost-effective way to manage wetland habitats for wetland fauna. More specifically, water-control structures can be used to flood or drain wetlands and thereby alter plant communities and the habitats available for wildlife. By drawing down water levels, a germination phase can be produced, which can be followed by gradual reflooding once plant communities become established. Timing, duration, and degree of drawdown (for example, shallow, growing season drawdowns versus complete overwinter drawdowns) influence subsequent plant species composition and the types of wildlife later attracted. Several years of stable water levels of moderate depth following drawdown are often required to establish the submerged plant communities that are important to many wetland animals (invertebrates, ducks, fish). Manipulation of wetlands to produce early stages of plant succession can create a diversity of habitat niches for wetland animals, particularly if practiced within a wetland complex in which management of different units is staggered for different successional stages. This practice permits the wetland fauna, elements of which can be quite mobile, to track habitat conditions over time and persist over the long-term within the same local area. Manipulation of wetland hydrology also has been used to manipulate salinity in coastal areas by blocking natural drainages, thereby permitting the fresh water that accumulates to leach out salt. The resultant shift in plant communities to more salt-intolerant species also changes wildlife communities.

In addition to water-level manipulation, a number of artificial procedures can be used to modify wetlands to enhance wildlife habitats. These include creation of openings in dense emergent vegetation through cutting, application of herbicides or, more permanently, by blasting. Planting of select forage species is sometimes attempted but is usually expensive and rarely successful. Provision of artificial nest sites for select species, especially ducks, has been one of the most common and visible forms of wetland management to benefit wildlife. Artificial nesting and loafing sites can be constructed for birds. Large, whole tree boles, including partially hollowed-out trees, can be placed within or adjacent to the restoration site. While perhaps aesthetically pleasing, the overall contribution of such efforts to boost local populations is at best modest given the small fraction of regional populations that can be supported by such structures. These artificial procedures, in general, are expensive and time-consuming and should be considered a minor complement to a larger strategy of natural management that uses hydrological manipulation to produce habitat changes over large areas in a cost-effective manner.

The tradeoffs between community types must be articulated and considered in management and planning. For example, flooding bogs and meadows to produce marshes, a not uncommon occurrence in North America, entails losses of certain species, many rare or unusual, and gains of others, many common and sought after by recreational users. In general, maintaining wetlands in a productive and natural state that meets the needs of the entire wetland fauna is more likely to meet diverse public needs than is species-specific management (e.g., for waterfowl only).

Conclusions

Despite lack of systematic comparisons, the existing scientific literature on wetland ecosystems is unambiguous in conveying the high level of diversity supported by wetlands. The source of this high diversity lies in the exceptionally large number of different types of wetlands that occur worldwide. Because wetlands occur on every continent except Antarctica, in every climatic region, in most biogeographic regions of the globe, and in a wide array of geological and topographic settings, the number of different types of wetlands created by the myriad combinations of these factors is very large and has yet to be catalogued. The extraordinary richness of plant and animal species that depend on wetlands arises from this diversity of types. Relative to the area they occupy, wetlands support a disproportionately large fraction of the world's rare, endangered, and threatened plant and animal species. Maintaining the high diversity of wetland species means maintaining a high diversity of wetland types.

See also: Estuarine Ecosystems. Lake and Pond Ecosystems. River Ecosystems. Wetland Creation and Restoration

References

Chapman VJ (1960) *Salt Marshes and Salt Deserts of the World*. New York: Interscience.
Chapman VJ (1976) *Coastal Vegetation*, 2nd edn. Oxford: Pergamon Press.
Chapman VJ (1977) *Wet Coastal Ecosystems*. Amsterdam: Elsevier.
Cowardin LM, Carter V, Golet FC, and LaRoe ET (1979) *Classification of Wetland and Deepwater Habitats of the United States*. U.S. Fish and Wildlife Service Publication FWS/OBS-79/31. Washington, DC: U.S. Fish and Wildlife Service.
Crow GE (1993) Species diversity in aquatic angiosperms: Latitudinal patterns. *Aquatic Botany* 44: 229–258.
Crum H (1988) *A Focus on Peatlands and Peat Mosses*. Ann Arbor: The University of Michigan Press.
Curtis JT (1959) *The Vegetation of Wisconsin*. Madison, WI: University of Wisconsin Press.
Ellison AM (2004) Wetlands of Central America. In: Whigham DF, Dykyjova D, and Hejny S (eds.) *Wetlands of the World*, vol. 2. Dordrecht, The Netherlands: Kluwer.
Finlayson M and Moser M (eds.) (1991) *Wetlands*. New York: Facts on File.
Good RE, Whigham DF, and Simpson RL (eds.) (1978) *Freshwater Wetlands: Ecological Processes and Management Potential*. New York: Academic Press.

Gore AJP (ed.) (1983) *Mires: Swamp, Bog, Fen and Moor.* New York: Elsevier.
Lugo AE, Brinson M, and Brown S (eds.) (1990) *Forested Wetlands.* New York: Elsevier.
Mitsch WJ (ed.) (1994) *Global Wetlands: Old World and New.* New York: Elsevier.
Mitsch WJ and Gosselink JG (2000) *Wetlands*, Third Edition. New York: John Wiley and Sons, Inc.
National Research Council (1992) *Restoration of Aquatic Ecosystems: Science, Technology, and Public Policy.* Washington, DC: National Academy Press.
National Wetlands Working Group (1997) *The Canadian Wetland Classification System.* Ontario, Canada: Wetlands Research Centre, University of Waterloo.
van der Valk A (ed.) (1989) *Northern Prairie Wetlands.* Ames, IA: Iowa State University Press.
Vitt DH, Li Y, and Belland RJ (1995) Patterns of bryophyte diversity in peatlands of continental western Canada. *The Bryologist* 98(2): 218–227.

Wildlife Management

David Saltz, Ben Gurion University of the Negev, Israel
Gary C White, Colorado State University, Fort Collins, CO, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by David Saltz, volume 1, pp 823–829, © 2001, Elsevier Inc.

Glossary

Density dependence Change in the birth and death rate of a population as a result of changes in population density; usually in the form of increased mortality rate and decreased fecundity as density increases.
Focal species A species that can be monitored and managed to maintain the integrity of the ecological system of which they are a part.
Information-theoretic approach A method of comparing between multiple competing models explaining a phenomenon based on the information lost (estimated distance from the truth) by each model.
Intermediate disturbance hypothesis A hypothesis claiming that some level of disturbance is necessary to maintain biodiversity.
Minimum viable population (MVP) A method for assessing the minimum size a population must be, to bring its extinction over a given time span to below a predetermined level.
Multimodel inference A method used to assess the uncertainty in model selection by calculating likelihoods among multiple competing hypotheses (models), most commonly based on Akaike's Information Criteria (AIC).
Population viability analysis (PVA) A method of assessing the probability of extinction for a given population within a specified time span.
Stochastic processes Processes that affect the dynamics of a population through random chance alone.
Sustained yield The removal of a constant number of animals from a wild population that can be maintained through time while keeping the population at a predetermined size.

Introduction

Although the term wildlife can be expanded to include all living organisms, the science of wildlife management, as an academic program offered and taught in universities and as written in textbooks, generally refers only to vertebrates.

Five major goals of wildlife management can be recognized.

1. Maximizing harvest/yield over time.
2. Maintaining population levels to support harvest and sport hunting.
3. Preventing extinction and increasing population survival probability.
4. Maintaining and managing the integrity of ecosystems and landscapes of which wildlife populations are a part.
5. Controlling wildlife to minimize damage to human crops and assets caused by wild populations or to return the ecosystem to some predetermined state.

As a science, wildlife management relies heavily on the understanding of ecological theory and processes, but differs from the science of ecology in its research objectives. Ecological research is targeted at formulating new hypotheses and providing the data to support or reject them. Thus, the species or setting selected to perform the research is that which is best suited to test a specific ecological theory. In contrast, research in wildlife is often dictated by real needs of specific problem-species or specific environments. The management of wild populations requires four processes: assessing the problem and defining the management objectives for the target population, monitoring and analyzing the ecology of the target population, formulating and implementing management actions based on the identified problems and ecology of the target population, and evaluating the implemented methodology. These processes can be carried out on a local or global scale (e.g., migratory waterfowl). In addition, because wildlife management is an applied science, involving what is often a national asset, public opinion strongly influences its goals. Thus, education, law enforcement, and public relations are also legitimate and appropriate aspects of wildlife management.

Harvest and the Concept of Sustained Yield

The science and profession of wildlife management have evolved over time based on perceived needs and ecological understanding. Originally, the field of wildlife management addressed only game species and harvest issues, recognizing wildlife as a crop that is a product of the land. The realizations that populations may go extinct due to overharvesting and past harvest rates affect the future performance of wild populations were the main forces driving the development of the science (Leopold, 1933). The recognition that populations can be overharvested is not new (a review of past examples of game management can be found in Aldo Leopold's book *Game Management*). As far back as biblical times, harvesting wildlife intelligently was considered important (Deuteronomy 22:6). According to Marco Polo's writings, Kublai Khan imposed hunting regulations throughout his Mongol Empire. The first refuge in the modern western world was established

by the British Parliament in 1869. However, the formulation of wildlife management as a quantitative science occurred only during the early part of twentieth century, driven mostly by the concept of "sustained yield" and its maximization (e.g., maximum sustained yield (MSY)). Leopold's *Game Management* was a keystone to this process, as was President Roosevelt's doctrine of conservation.

Sustained yield refers to a constant level of harvest that when applied, will allow the harvested population to stabilize at a given size. This can be achieved by removing an amount equal to the population's rate of increase. In the absence of limiting factors, the rate of increase of a population should remain unchanged, and, therefore, absolute growth will increase with population size (more animals have more offspring). However, populations do not grow indefinitely, which suggests that at some point, the population rate of increase begins to decline as the population increases. This density-dependent response is due to a decline in the amount of resources available per individual, which causes a subsequent reduction in reproductive success and increased mortality. The classic logistic growth equation, introduced by Verhulst in 1838, describes the growth curve of such a population, depicting an S-shaped curve of population size against time (**Figure 1a**). The rate of increase of the population at time t is the first derivative (dN/dt) and is given by

$$\frac{dN}{dt} = r_{\max} N\left(1 - \frac{N}{K}\right) \qquad (1)$$

where N is the population size at time t, $r_{\max}$ is the maximum growth rate per individual when resources are not limiting, and K is the maximum size of the population that can be supported over time in a particular environment (K is often termed "carrying capacity"). As can be clearly seen, there are two cases when population growth is 0; when $N = 0$ and when $K = N$. Thus a sustained yield requires that the population be kept below K. Between $N = 0$ and $N = K$, growth rate initially increases with population size and then decreases (**Figure 1b**). Each point on this curve represents the sustained yield for a given population size, and the highest point is the MSY.

We calculate MSY by finding the maximum value of dN/dt. This can be done analytically by taking the second derivative of the equation depicting the growth curve of the population and equating it to 0. In the case of the logistic equation,

$$\frac{d(dN/dt)}{dN} = r_{\max}\left(1 - 2\frac{N}{K}\right) = 0 \qquad (2)$$

This equation equals 0 when $r_{\max} = 0$ (which is meaningless) or when $2\frac{N}{K} = 1$. Rearranging the latter, we find that $N = \frac{K}{2}$; that is, the population grows the fastest when at $\frac{1}{2}K$. To find the rate of increase at that point, we simply substitute $\frac{1}{2}K$ for N in equation (1) and find that the maximum growth rate and, therefore, MSY, is $r_{\max}\frac{K}{4}$.

Note that except for the point of MSY, a given yield (growth rate) can be achieved at two different population sizes, one on either side of MSY$\left(\frac{K}{2}\right)$ – the upper and lower sustained yields. Thus it appears that one may achieve the same level of sustained yield with a small population as with a large one. However, the two points differ considerably in terms of stability. The points to the right of MSY are stable because moderate overharvesting will increase the growth rate of the population, thus compensating for the excessive harvest (and vice versa). In contrast, the points to the left of MSY are unstable because overharvesting will cause a decrease in the rate of increase. As a result, the next harvest will further decrease the population and rate of increase and so on. Initially the impact may appear minor, but this "snowballing" effect can very quickly lead to the demise of the population.

Clearly, real populations do not behave deterministically nor does their density-dependent response follow that of the logistic model. Density-dependent responses may vary depending on the species and environment, skewing the dN/dt curve; the age structure of the population must be considered; the environment (carrying capacity) constantly changes; other species (predators, competitors, etc.) interact with the population; and hunters may select for specific animal traits (sex, age group, size, etc.). Thus, much of the research on harvested populations has been directed at developing more sophisticated models with better predictive capabilities and estimating the parameters needed for them (Getz and Haight, 1989; Saether, 2001).

Although many advances have been made in this field, to date, harvest strategies in many places around the world are still implemented on a trial-and-error basis trying to just maintain a level that will support hunting and produce revenue. However, even under such protocol, harvest theory has made an important contribution in demonstrating that uncontrolled harvest can very quickly lead to the demise of populations. Furthermore, while the logic of a sustained yield set somewhat above MSY is obvious, nonecological factors may have a strong influence on harvest policy. In many

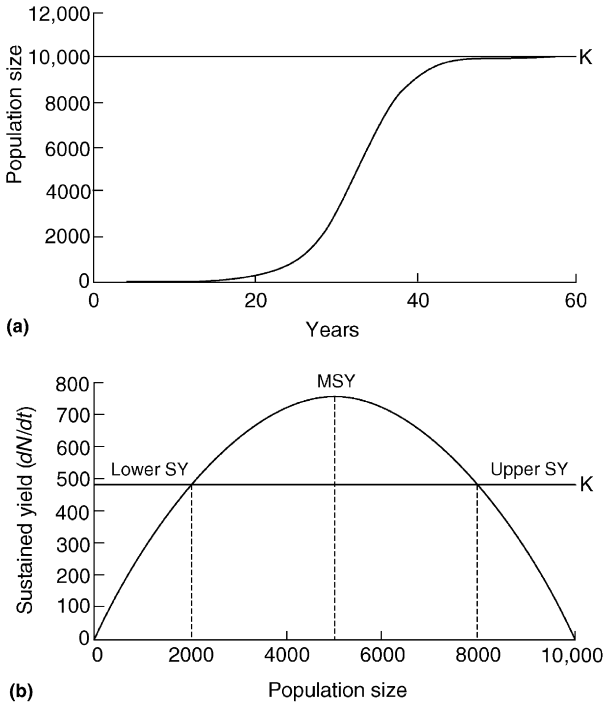


Figure 1 The logistic population growth curve as a function of time (a) and sustained yield as a function of population size (b).

species, the long-term monetary income from a sustained yield may be less than the yield offered by the economic market on regular investments (Clark, 1976). In other words, economically, a greater profit can be realized by selling licenses to hunt the entire population and investing the profits than by limiting the number of licenses to a sustained yield that will ensure the continued existence of the population. Thus, in recent years considerable effort has been directed at assessing the economic value of wildlife for nonconsumptive uses such as tourism and conservation (Fennell, 2008).

Endangered Species and Reducing the Risk of Extinction

With the biodiversity crisis coming into focus during the second half of the twentieth century, the science of conservation began to evolve. Initially, conservation to preserve an endangered species and wildlife management were treated as separate sciences. Eventually, the sciences merged in the area where they overlap (wildlife), and managing wildlife populations for conservation became an integral part of wildlife management. Topics such as management for recovery and sustainability, the dynamics of small population (especially as impacted by demographic, genetic, and stochastic environmental processes), fragmentation, and the importance of populations to the integrity of the ecosystem and biodiversity became a central focus of wildlife management. In terms of management for conservation, Caughley and Gunn (1996) recognized two central paradigms: the small population paradigm and the declining population paradigm.

The sensitivity of small wildlife populations to extinction was first recognized and expressed in 1938 in a book written by the famous American ecologist Allee. In his book (Allee, 1938), Allee writes: "The general conclusion seems to be that different species have different minimum populations below which the species cannot go with safety, and that in some instances this is considerably above the theoretical minimum of one pair." Four types of stochastic processes (Shaffer, 1981) play an important role in the dynamics of small populations and contribute to their increased risk of extinction: demographic stochasticity, genetic stochasticity (including loss of genetic variance and inbreeding depression), environmental stochasticity, and catastrophes.

The accelerated loss and fragmentation of wildlife habitat during the second half of the twentieth century restricted many wildlife populations and required that management steps be taken to minimize these stochastic effects. The size of refuges and other areas set aside for conserving wild populations must be large enough to support a population equal to or larger than the minimum viable population (MVP) (Belovsky, 1987). Corridors connecting small populations are currently considered important (Bennett, 1999), but their effectiveness has not yet been fully demonstrated. Where the establishment of corridors is not feasible, translocation has been implemented. For instance, in South Africa, mountain zebras are translocated by airlift between isolated populations. Habitat improvement and resource supply during limiting periods of harsh environmental conditions are also being used to minimize population declines and reduce the impact of stochastic processes and risk of extinction.

The management of declining species requires that the cause of decline be identified and managed directly or indirectly. Human-related decline of wildlife populations in the twentieth century can be attributed to three main factors: overharvesting, loss and fragmentation of habitat, and the introduction of alien species. With the improved control of harvest by science and legislation, loss of habitat and fragmentation became the major causes of wildlife population declines. This process has highlighted the importance of considering space in wildlife management. Management efforts have been concentrated on land allocation for conserving wildlife and securing refuges and linkages (corridors) between patches (fragments) of adequate habitat. Roads and fencing are some of the major causes of fragmentation, and special overpasses, underpasses, and gates have been developed to ensure wildlife movement.

Managing Wildlife for Conserving the Integrity of Ecosystems and Landscapes

Wildlife populations are important for maintaining biodiversity in the system of which they are a part. Many wild populations are critical for the functioning of the entire community. Beavers (*Castor canadensis*) are a classic example of how a single species engineers change in the ecosystem beneficial to many other species. Prairie dogs (Cynomys spp.) are another example. Thus, in many instances, populations are managed in order to sustain species richness of the ecosystem of which they are a part. For example, research has shown that both overgrazing and undergrazing will reduce plant species richness. This is known as the intermediate disturbance hypothesis (Connell, 1978). In certain cases, herbivore populations are managed to maintain the diversity of the plant community. Elk (*Cervus elaphus*) populations were regularly controlled to maintain the existing landscape in Yellowstone National Park. Wolves (*Canis lupus*) were reintroduced (Fritts and Carbyn, 1995) into Yellowstone with the hope that they would control the elk population thus restoring cottonwood stands along the river-banks. The control, however, was not manifested by direct predation, but rather by changes in the behavior of elk that avoided the riverbanks due to the increased risk imposed by the newly present wolves (Ripple and Beschta, 2004).

The use of focal species as a tool to manage ecosystems gained popularity in the 1990s but has waned in the past decade. The underlying assumption is that by managing the focal species, the entire system, or a significant part of it, can be secured. Indicator, keystone, and umbrella species are the main types that have been suggested as focal species through which systems can be monitored, conserved, or managed.

Indicator species are species that testify to the well-being (health) of the entire ecosystem. These species are the first to respond to deterioration of the ecosystem, since they are more sensitive. By monitoring and managing the indicator species population only we can estimate and protect the welfare of the entire ecosystem. We assume that as long as the status of the

indicator species is satisfactory, the entire system is operating adequately.

Keystone species are species that play an important role in the ecosystem. These may be animals that hold most of the biomass in the ecosystem or that influence many of the other species or functions of the ecosystem. In this way, if the population of a keystone species is removed from the ecosystem, the ecosystem changes dramatically. Therefore, by managing only the population of keystone species, we are managing most of the ecosystem. In recent years, species that modify the environment (termed "ecosystem engineers") have become a point of interest (VanNimwegen *et al.*, 2008).

Umbrella species are species that require a large area containing many types of habitats to sustain a viable population. Often these are large-bodied homeotherms with large home ranges. By securing a tract of land that is large enough to sustain a viable population of these species, many others will come under the same protection. The focus here is on leaving enough area for the umbrella species and other members of the ecosystem, but there are no direct management implications for this approach.

Control of Wildlife

The everexpanding human population ensures that human–wildlife conflicts will continue. Most of these conflicts are within the realm of damage to crops, equipment, and other assets. Thus, a major component of wildlife management is wildlife control. A good example is managing habitats around airports to reduce gull activity near runways that endangers aircraft during takeoff. Often, damage caused by wildlife is a result of human manipulation of the environment, such as outbreaks of herbivores following the eradication of predators and high concentration of canids around human waste sites leading to outbreak of rabies. Wildlife management strives to implement control measures in a manner enabling the other goals of wildlife management (sustained yield or maintaining biodiversity) to be met.

The four processes necessary for wildlife management generally apply to wildlife damage control, with some differences. These processes are the following: (1) Defining the problem. The species causing the damage must be identified, the type of damage determined, and extent of the damage evaluated. (2) Knowledge of the general ecology, behavior, and dynamics of the species causing the damage must be obtained and evaluated with special regard to its response to various control techniques. (3) The combination of (1) and (2) defines the methods that are feasible and economical to control the damage. (4) The methodology used must be evaluated. Unless the cost of control is less than the losses due to the damage, management should not be implemented.

Wildlife Management Techniques

Because wildlife management is an applied science, methods and techniques for both research and implementation are important. As a result, a considerable amount of research has been directed at developing, assessing, and improving techniques. Techniques for managing wildlife are aimed at studying, reducing, increasing, or maintaining the population at its current level while securing its integrity (i.e., preventing loss of genetic diversity, enhancing long-term survival, preventing epizootics, etc.) and the integrity of its ecosystem. The techniques can be classified generally into two categories.

1. Applied techniques for manipulating populations by impacting survival, reproductive success, or distribution.
2. Techniques for studying, analyzing, and assessing population.

Many of the more common techniques are published in a series of *Wildlife Techniques Manuals* published by The Wildlife Society. The first of this series, edited by Mosby, was published in 1960. The most recent issue (the 6th edn.) was published in 2005 and edited by Braun. The manual covers five main topics: teaching and communication skills, study design and analytical techniques, investigational techniques, managing populations, and habitat management.

Applied Techniques for Managing Wildlife

Applied techniques may be resource (habitat) manipulation, mechanical, behavioral, immunization and immunocontraception, biological, and direct population control.

Habitat and community manipulation: This refers to managing target species by impacting important resources and other species that interact with them (i.e., habitat manipulation, or impacting predator–prey relationships and competing species, and species that are ecosystem engineers). By controlling key elements that affect the species abundance (water, shelter, predators, etc.), it is possible to manipulate its density and distribution. The increase in computer power enabled the use of spatially explicit models and Geographic Information Systems techniques to project the impacts of various habitat manipulations (Verner *et al.*, 1986).

Mechanical: The simplest and most successful method of controlling wildlife damage is exclusion by fencing. The advantages of fencing are that it is relatively fail-proof and nonlethal. However, it is expensive and labor intensive, demanding constant upkeep. Furthermore, fencing large tracts of land is a major cause of fragmentation.

Behavioral: These are modern forms of scarecrows that operate on the animal's senses and are, therefore, acoustic, olfactory, or visual. A good example of visual repellents, other than the classic scarecrow, is the silhouettes of birds of prey that are placed on large glass windows to prevent other birds from slamming into them. Examples of acoustic deterrents are gas cannons that produce explosions at irregular time lapses or recordings of alarm calls of other animals. Ultrasonic deterrents work by using sounds at a high decibel level (above 120 dB), outside the range of human hearing, to cause a painful stimulus to animals with sensitive hearing, such as canids. Olfactory repellents are chemical compounds. Two types of chemical compounds can repel animals: (1) compounds that link a food source with an induced sickness through a behavioral process called conditioned taste aversion; and (2) compounds that repel through the unpleasant stimuli of the nervous system. The first type are properly

referred to as aversive agents, while the second type are true sensory repellents.

Immunization and immunocontraception: These techniques require the administration of a vaccine either directly, by capturing and injecting, or indirectly, by oral administration with bait. A major advancement in wildlife management in the past decade is the control of rabies by oral immunization. Immunocontraception is a very attractive method of control because it is potentially highly species specific, but it so far has had limited success.

Biological: Biological control has been mostly effective against insects. There are only a few cases in which biological control was effectively applied to wildlife. Of these, the introduction of *Myxoma* virus to control rabbits introduced into Australia is the most famous. Although complete eradication was not achieved, the virus now holds the rabbit population at 20% of its original size.

Direct population control: Direct population control refers to the removal (harvesting) or addition (reintroduction and translocation) of animals. Removal is usually used when the animal is a game species and can also be controlled by hunting license quotas. Its main advantages are that it is species specific, the number of animals removed can be closely controlled, and it has limited impact on the environment. Reintroductions and translocations are important methods for enhancing the viability of existing populations and reducing the species risk of extinction.

Studying, Analyzing, and Population-Assessing Techniques

A major prerequisite of sound management is the accurate estimates of population size and dynamics, and determination of how these are influenced by other factors. Consequently, a significant part of wildlife science today is directed at the development and improvement of estimation techniques for assessing population density, survival, reproduction success, occupancy, well-being, age structure, density-dependent responses, and interspecific interactions. Such data are the basic requirements for estimating sustainable harvest quotas in different and variable environments and assessing extinction probability and other risks.

Rigorous statistical methods for assessing density and survival based on maximum likelihood estimation techniques, are continuously being refined. These methods evolved from a landmark paper (Seber, 1970) on estimating survival from the recovery of dead animals, and the insight of David R. Anderson of the US Fish and Wildlife Service who recognized the potential of Seber's work. Computer software such as DISTANCE, SURVIV, ESTIMATE, JOLLY, BROWNIE, and more recently MARK, were developed specifically for wildlife research and management (Dinsmore and Johnson, 2005).

Research and management of wildlife are carried out over large geographic areas, often with limited access and visibility. Considerable effort has been directed toward developing remote data collection techniques using indirect methods of assessing population condition and nutrition by using individual animal physiological indicators (Servello *et al.*, 2005).

Methods for assessing the viability of small populations (population and habitat viability analysis, (PHVA)) and estimating the minimum population size necessary to ensure its long-term existence (MVP) have been developed based on stochastic theory (Burgman *et al.*, 1993). Commercial software such as RAMAS and Vortex are available for PHVA and MVP analysis; however, in most cases and due to the detailed knowledge required, the data are insufficient for a reliable analysis.

Clearly the greatest change in the analysis of data from wild populations in the past decade is the use of the Kullback-Leibler information-theoretic approach and multimodel inference (as opposed to the standard Popperian approach). This approach was integrated into the field of wildlife management and evolved within it, driven by the need to deal with multiple potential predictor variables (Burnham and Anderson, 2002). The approach has consequently filtered into the basic-ecology literature and reflects a paradigm shift in this and related fields.

See also: Endangered Ecosystems. Endangered Mammals. Indicator Species. Keystone Species. Sustainability and Biodiversity

References

Allee WC (1938) *The Social Life of Animals*. New York: Norton & Company.
Belovsky G (1987) Extinction models and mammalian persistence. In: Soule ME (ed.) *Viable Populations for Conservation*, pp. 35–58. Cambridge, UK: Cambridge University Press.
Bennett AF (1999) *Linkages in Landscapes*. Gland, Switzerland: IUCN.
Burgman MA, Ferson S, and Akcakaya HR (1993) *Risk Assessment in Conservation Biology*. London, UK: Chapman & Hall.
Burnham KP and Anderson DR (2002) *Model Selection and Multi-Model Inference: A Practical Information – Theoretic Approach*, 2nd edn. NY: Springer.
Caughley G and Gunn A (1996) *Conservation Biology in Theory and Practice*. Cambridge, MA: Blackwell Science.
Clark CW (1976) *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*. New York: John Wiley & Sons.
Dinsmore SJ and Johnson DH (2005) Population analysis in wildlife biology. In: Braun CE (ed.) *Techniques for Wildlife Investigations and Management*, 6th edn., pp. 154–184. Bethesda, MD: The Wildlife Society.
Fennell D (2008) *Ecotourism*, 3rd edn. New York: Taylor and Francis.
Fritts SH and Carbyn LN (1995) Population viability, nature reserves, and the outlook for gray wolf conservation in North America. *Restoration Ecology* 3: 26–38.
Getz WM and Haight RG (1989) *Population Harvesting*. Princeton, NJ: Princeton University Press.
Leopold A (1993) *Game Management*. New York, NY: Charles Scribner's Sons.
Ripple WJ and Beschta RL (2004) Wolves and the ecology of fear: Can predation risk structure ecosystems? *BioScience* 54: 755–766.
Saether B (2001) Special issue on Harvesting. *Wildlife Biology* 7(3): 129–244.
Seber GAF (1970) Estimating time-specific survival and reporting rates of adult birds from band returns. *Biometrika* 57: 313–318.
Servello FA, Hellgren EC, and McWilliams SR (2005) Techniques for Wildlife Nutritional Ecology. In: Braun CE (ed.) *Techniques for Wildlife Investigations and Management*, 6th edn., pp. 554–590. Bethesda, MD: The Wildlife Society.
Shaffer ML (1981) Minimum population sizes for species conservation. *BioScience* 31: 131–134.
VanNimwegen RE, Kretzer J, and Cully JF (2008) Ecosystem engineering by a colonial mammal: How prairie dogs structure rodent communities. *Ecology* 89: 3298–3305.
Verner J, Morrison ML, and Ralph CJ (1986) *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates*. Madison, WI: The University of Wisconsin Press.

Worms, Annelida

Kristian Fauchald, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Chaetae Bristles composed of chitin cylinders glued together by scleroprotein emerging from parapodia or the body wall in many annelids, including most polychaetes and in the oligochaet families of the clitellates, but present in most of all major annelid taxa.

Nuchal organ Sensory organs with nerves coming from the posterior part of the brain or from the nerve ring, usually present as ciliated pits or short ridges at the posterior end of the prostomium, present in most polychaetes species, also present in the sipunculans.

Parapodium Segmentally arranged outpockets from the body wall, usually divided into two branches (rami), the dorsal notopodium and the ventral neuropodium; present in most polychaete species, but extremely variably developed and sometimes entirely absent.

Peristomium Region around the mouth of the larvae; if distinct in adults may carry peristomial cirri.

Prostomium Morphologically anteriormost part of the body; may carry antennae and eyes and carry at least part of the brain.

Segment In annelids a section of body set off from the rest of the body by septae at both ends; containing a separate part of the secondary body cavity (coelom) with paired ventral ganglia, and usually nephridia, gonads, transverse muscles and segmental blood vessels carrying externally parapodia with chaetae and gills in most polychaetes, chaetae in oligochaetes. Distinct segments present in posterior end of siboglinids (pogonophorans and vestimentiferans), indistinct in most other annelids.

Trochophore Larva characteristic especially of annelids and mollusks, consisting of an episphere, a ciliated girdle around the middle called the prototroch, and the hyposphere. In feeding trochophores the mouth is located behind the prototroch and the anus is terminal, usually surrounded by a second ciliated band, called the telotroch. Segmentation takes place between the prototroch and the telotroch and is initiated immediately in front of the telotroch.

Annelida

Traditionally the phylum Annelida consisted of three classes: Polychaeta, Oligochaeta and Hirudinea. The Hirudinea were shown to be derived from one of the oligochaeta groups, so the two groups are now fused to Clitellata. The clitellates have been shown to be closely linked to one of the polychaete groups, so Polychaeta is now considered a paraphyletic group in relation to the clitellates.

Two small, mainly deep-water phyla, Vestimentifera and Pogonophora have been shown to be a single group, and to belong to a single family, Siboglinidae, a sister taxon to the sabellid polychaetes; thus these two phyla are now considered annelids belonging to one of the polychaete groups.

Recently the spoon worms, phylum Echiura, have been linked directly to the annelids, probably most closely related to the clitellates. Finally, the peanut worms, phylum Sipuncula, have been shown to be related to the annelids, but the position is still debated.

One group, Myzostomida, parasitic on echinoderms, was considered a member of the Polychaeta; they were then removed from the Annelida and associated with the Platyhelminthes, but they are now back among the annelid-related groups, in a somewhat uncertain position.

The evidence for this reorganization of many of the worm-like phyla comes only in part from morphology or embryology, the most convincing evidence has been supplied by molecular genetics.

Polychaeta is a commonly encountered group of annelids (segmented worms) best represented in the marine environment, although several species are present in freshwater and a few are known from moist soil on land. The Polychaeta, the many-bristled worms, include about 15,000 described species. Estimates of the total number of species varies, but judging from the rates of description of new species it is likely that more than 20,000 species will eventually be recognized, based on morphological characters in addition to the study of all genetically fixed properties of the worms. No single feature can be used to diagnose the taxon uniquely; until recently the presence of the nuchal organs were considered such a feature, but even this structure is missing in several polychaetes and present in Sipuncula.

Clitellata contains the earth worms and their aquatic relatives and the leeches. The clitellates evolved from the polychaetes, but the exact relationship has not yet been diagnosed.

Morphology

This section is based mainly on the polychaetes, since these are perhaps the most complex of the annelids.

The annelids are composed of three body regions, derived from regions present in the larvae. In the polychaetes the presegmental prostomium and peristomium are followed by a segmented body and a postsegmental pygidium, which carries the anus. Most polychaetes are less than 100 mm in length

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00170-2

with 100 or fewer segments; many are very small, with as few as 10 segments, and are adapted to life between sand grains in sandy beaches. Some species may become very large with 2–3000 segments and lengths between 1 and 3 m; the longest recorded specimen was more than 6 m long when collected in Port Jackson, Australia. These very large worms may be as old as 50–100 years, but the documentation for this is rather poor. In general, small species are short lived compared to larger relatives. Perhaps most polychaetes live 12–18 months, while others go through three or four generations in a year. The general body shapes may be earthworm-like with cylindrical bodies and very little in the way of external appendages, but more usually the protruding parapodia and chaetae make the worms appear ragged. The ventral side may be flattened with two longitudinal ridges on either side of the mid ventral nerve chord; the dorsal side is usually convex. This description fits best for mid-sized free-living worms; the tubicolous worms are often cylindrical with inconspicuous parapodia, the very small species are often slender and the parasitic species may be flattened and disc-shaped.

Of the clitellates the oligochaetes are slender; most aquatic species are small, the terrestrial ones may be very large, such as the giant earthworms present in Australia. The leeches are usually flattened and have suckers developed at one or both ends of the body.

The echiurans are often large, rounded or sausage shaped, without any visible segmentation and with a pair of, or sometimes more, hooks just behind the mouth; the prostomium is spoon shaped or sometimes extremely elongated T-shaped.

The sipunculans are often more or less peanut shaped (hence the common name) when the anterior end is retracted; when this end is extended they become elongated drop shaped with a crown of feeding appendages anteriorly.

Head

The head consists of the prostomium and peristomium in the adults, but one or more of the first segments may be fused to this structure (see **Figures 1–3**). The prostomial part may carry antennae and eyes, and the fused anterior segments may have a pair of tentacular cirri per fused segment; these cirri are the dorsal and ventral cirri of each segment participating in the head structure. In Canalipalpata the head end can be complex and the number of segments included in the head may be difficult to count. Some of these worms may have a tentacular crown (e.g., the sabellids and serpulids) or a very large number of antennae (e.g., the terebellids). Certain polychaetes lack anterior appendages entirely (e.g., the capitellids and maldanids).

In clitellates the head consists of the prostomium and peristomium and lacks appendages. In the echiurans the prostomium is most often short and spoon shaped, but may be extremely elongated and distally T-shaped. In the sipunculans the prostomium is a rounded lobe when distinct and the peristomium consists of branching groups of feeding appendages.

Segments

Segments are always added just in front of the pygidium, that is, from the posterior end of the segmented body. They are often first seen as low ridges just in front of the pygidium; parapodia, chaetae and other externally obvious structures are differentiated as more segments are added and, as a consequence, each segment no longer has the position of the last segment. The total number of segments may be fixed and limited (e.g., 20, 25, etc., in certain polynoids, maldanids and opheliids). In many species the number of segments varies with relatively narrow ranges; for example, many neridid species have 80–120 segments. Some species have an unlimited number of segments and the worm keeps on adding segments throughout life. Normally each segment has a pair of parapodia (see **Figure 4**). In many species the parapodia are poorly developed and marked only by the position of the chaetae; in other taxa the parapodia are complex structures with a variety of lobes and lamellae. In well-developed parapodia the two rami often differ. The notopodial are smaller (or shorter) than the neuropodia in many species; when they are completely missing the parapodia are called uniramous. In the identificatory literature a number of terms (e.g., sequiramous) are used to characterize the level of reduction of the notopodia. These terms are used inconsistently and it is better to specify the structure observed than to apply a poorly understood term. When the parapodial are best developed each ramus has a chaetal lobe supporting one or more bundles (or fascicles) of chaetae and pre and postchaetal lamellae covering the bases of the emerging chaetae. The chaetal lobe is supported by one or more acicula in many species. A dorsal cirrus is found on the dorsal edge of the notopodium or on the adjacent body wall and a ventral cirrus is along the ventral edge of the neuropodium; both cirri are mainly sensory structures. In addition to the pre and postchaetal lamellae various other lamellar or finger-shaped structures may be associated with each ramus. A complex terminology is associated with shape and position of each structure. Clitellates have distinct segments, often subdivided into two or more rings, but lack parapodia. In echiurans, larvae may have more or less distinct internal segmentation distinct mainly as groups of muscles; the sipunculans lack distinct segmentation, note however, that the sipunculans have the body cavity separated into an anterior and posterior cavity organized so that the eversion of the anterior end is supported from musculature in the anterior cavity.

The Pygidium

The pygidium may be a rounded or tapering simple structure with the anus opening either dorsally or terminally, but frequently it carries anal (or pygidial) cirri. Commonly one or two pairs of cirri are present, but in some taxa it may be ornamented with numerous cirri or flattened lamelliae (e.g., maldanids), and in the opheliids it may have an open hood with many papillae, both inside the hood and along the margin.

In the sipunculans the digestive tract is U-shaped and as a consequence the anus opens at the base of the anterior introvert.

Chaetae

The chaetae, the bristles, may be in bundles, however, in most cases the ''bundles'' consist of two or more short, oblique rows

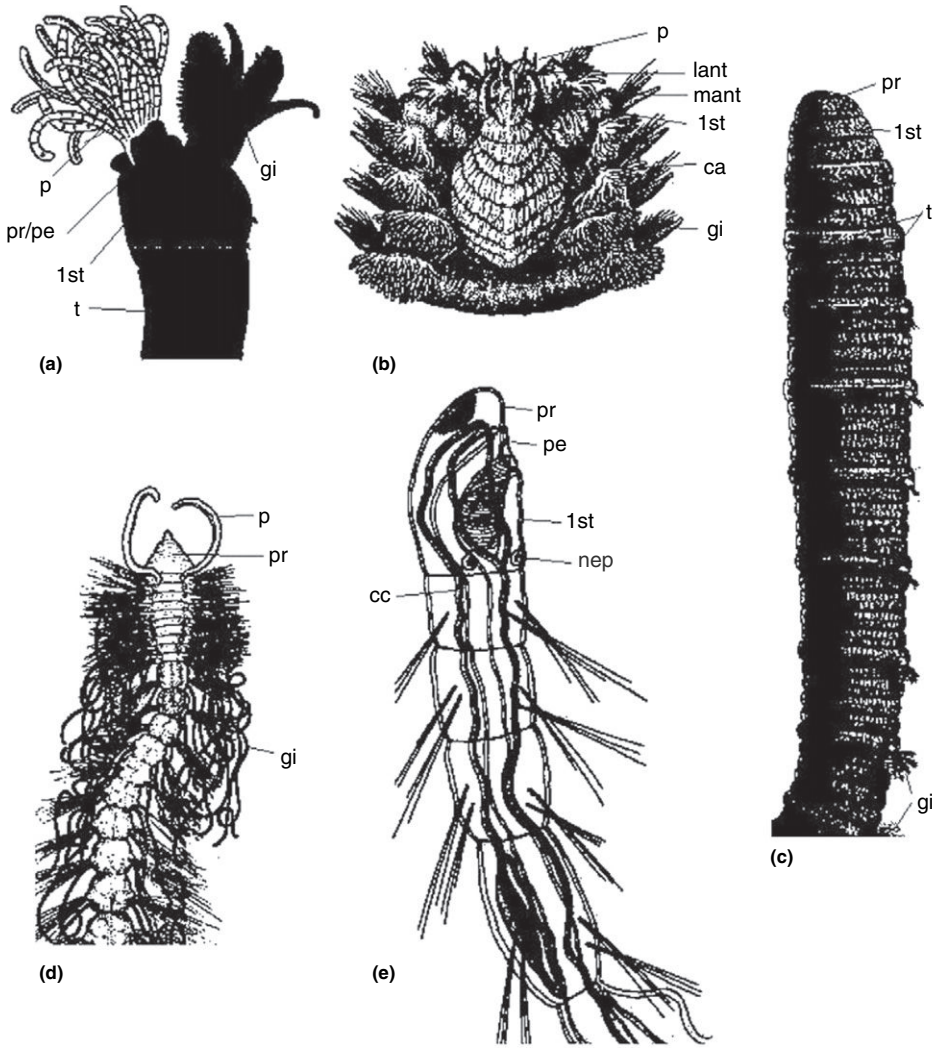


Figure 1 Polychaete anterior ends. (a) Ampharetid, lateral view. (b) Amphinomid, dorsal view. (c) Arenicolid, lateral view. (d) Cirratulid, dorsal view. (e) Ctenodrilid, lateral view. Modified from Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

of a few chaetae each. Other polychaetes have chaetae arranged in flattened fascicles usually arranged dorsoventrally. Each chaeta is produced from an invaginated epidermal cell (chaetoblast). The outer cell membrane of each chaetoblast is covered with microvilli, each of which produces a slender cylinder of chitin. These cylinders are glued together by scleroprotein produced by cells lining the invagination. All chaetae are formed the same way, through complex interactions between production of cylinders and glue secreted to keep the chaeta together and give the external shape. The chaeta may be tapering capillaries, but most chaetae are much more complex (**Figure 5**). A species may have a single kind of chaetae or two or more kinds; in some parapodia each chaeta may be different, but more usually groups of chaetae with similar structure are present. New chaetae are formed throughout the life of the worms, and distribution of different kinds of chaetae is carefully controlled and changes as the worm grows. In rows of identical chaetae new chaetae are produced at one end of the fascicle and drop out at the other end (e.g., hooks in terebellids). In syllids that often have few chaetae, and each chaeta differs from all other chaetae in a parapodium, replacement takes place in a one-to-one fashion. Special kinds of chaetae such as acicula, which are rods supporting long parapodia and various large spines often show signs of wear, an indication that they are not replaced or at least not replaced as often as other chaetae. The base of the chaetae (chaetal sac) is anchored with muscles running across the body cavity, making it possible to protrude and retract the whole fascicle or bundle since in most cases only a single set of muscles are present. Acicula and large spines often have their own investment of muscles and are independently movable.

The oligochaete taxa of the clitellates each have a limited number of kinds of chaetae, most have capillaries or spines. The echiurans have minimally a pair of large hooks, usually located ventrally behind the mouth. The sipunculans lack chaetae, but have epidermal hooks formed directly externally on the body.

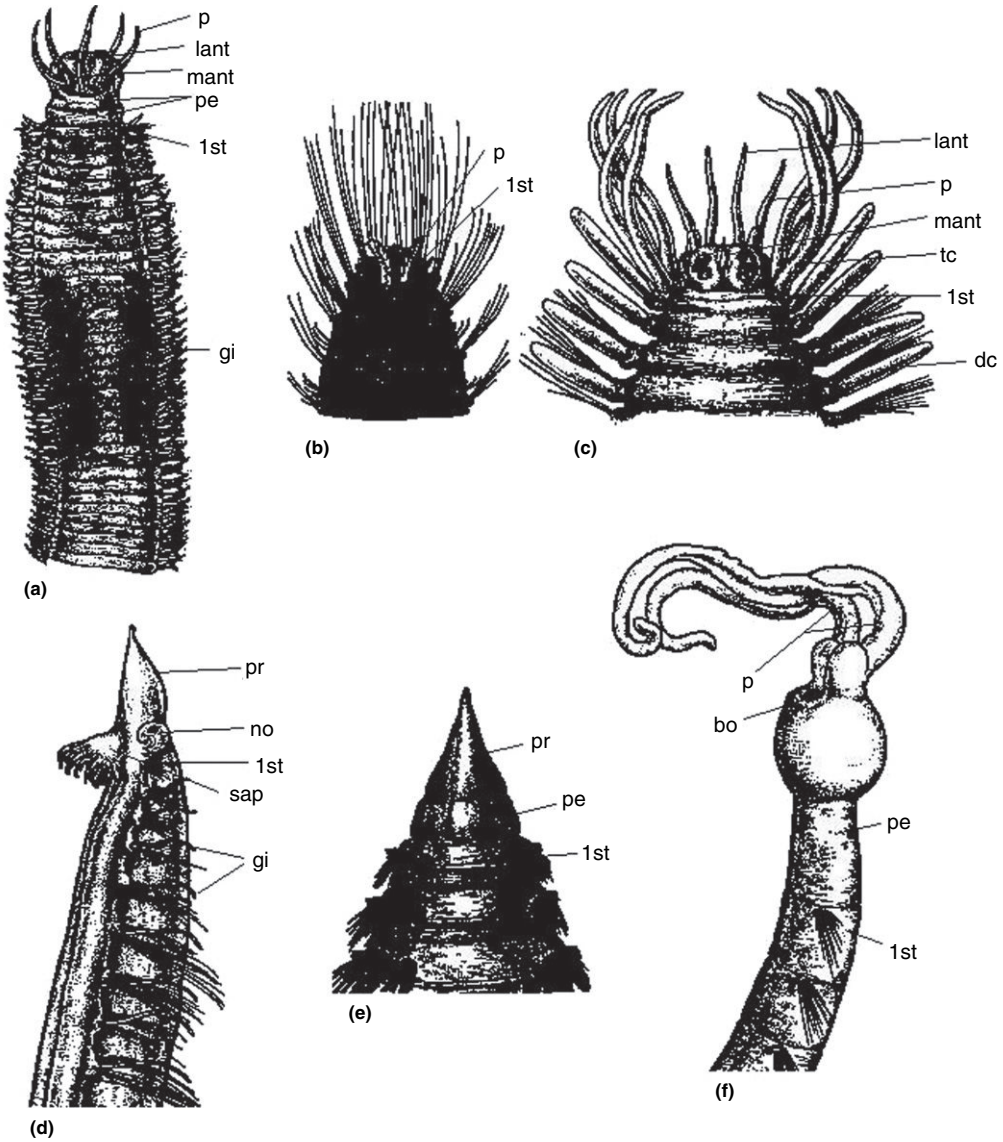


Figure 2 Polychaete anterior ends. (a) Eunicid, dorsal view. (b) Flabelligerid, dorsal view. (c) Hesionid, dorsal view. (d) Opheliid, lateral view. (e) Orbiniid, dorsal view. (f) Oweniid, lateral view. Modified from Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

The Nervous System

The nervous system consists of a dorsal brain connected by a double nerve ring running around the esophagus to double ventral nerve chords. The anterior part of the brain is located in the prostomium, but penetrates well into the peristomium or even into the next several segments in many species. In most groups the two parts of the ventral nerve chord are wholly or partially fused into a single entity; each segment always contains a pair of ganglia from which the segmental nervous system runs to the parapodial muscles, nephridia, gonads, and circulatory system. The gut is invested from a separate set of nerves from the lower part of the brain (stomatogastric nerves). Antennal nerves run from the anterior part of the brain, and a pair of nerves associated with the posterior part of the brain runs to the palps. Indeed, Orrhage has found a pattern in the emergence of nerves and internal connections in polychaete brains, showing that the limited number of nerves can be identified from one group to the next based on position and function.

The ventral nerve chord may contain two or more giant fibers associated with rapid retraction into burrows or tubes consisting of very thick nerves running from the anterior end innervating the longitudinal muscles directly with only a minimum of synapses. Other motor fibers are segment-to-segment loops passing along undulatory locomotory waves.

The clitellates have a nervous system resembling the one present in the polychaetes, In other annelids the main nervous system resembles the one present in the polychaetes, but segmented nature of other nerves is absent.

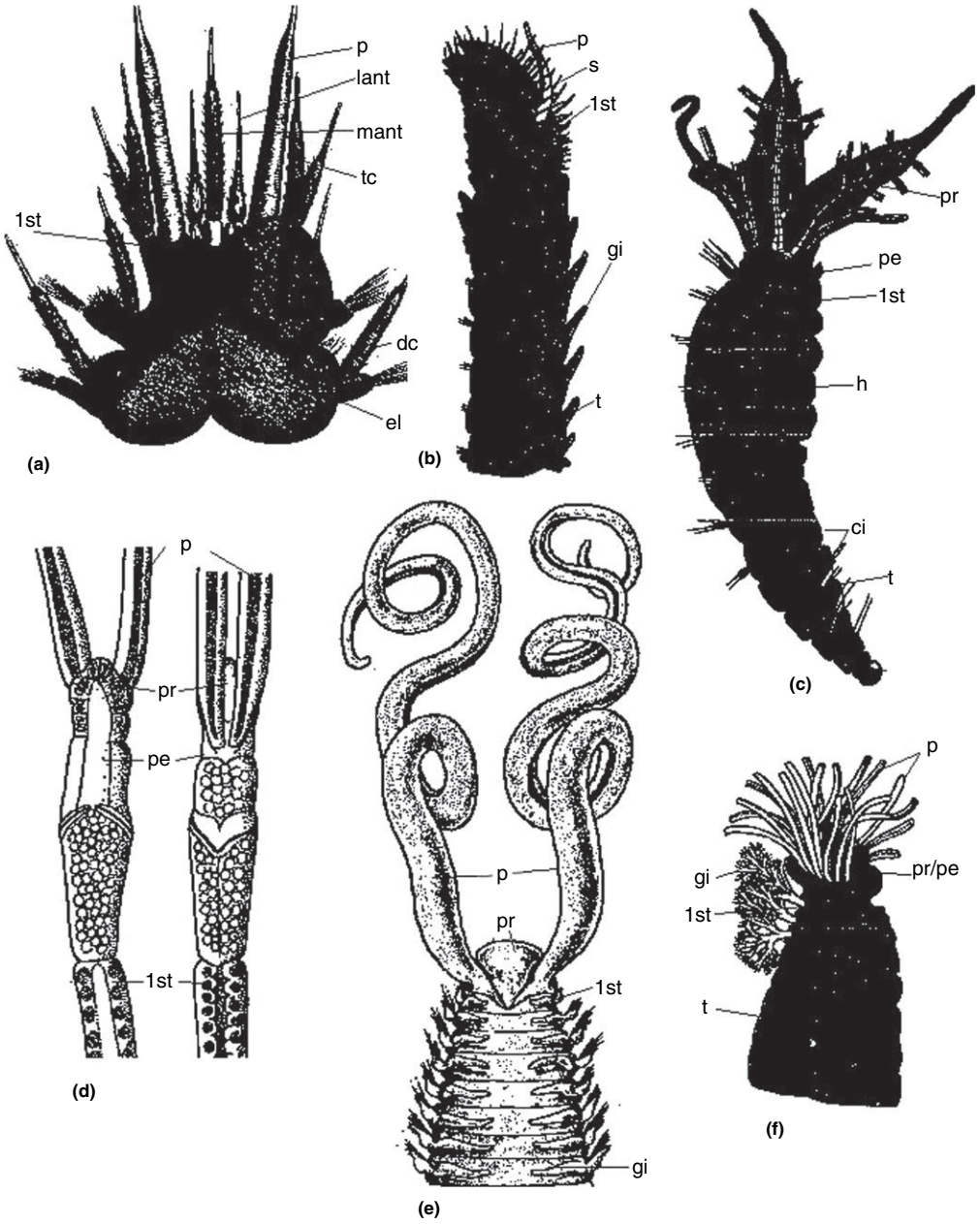


Figure 3 Polychaete anterior ends. (a) Polynoid, dorsal view. (b) Sabellariid, lateral view. (c) Sabellid, lateral view, whole body. (d) Siboglinid, dorsal and ventral view. (e) Spionid, dorsal view. (f) Terebellid, lateral view. Modified from Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

Sense Organs

Eyes are mostly eye-spots located on the prostomium or even directed on the upper surface of the brain. In some polychaetes the eyes become much larger when the species are sexually mature (e.g., nereidids), and the epidermis over the eye becomes a translucent lens. In a group of pelagic polychaetes (e.g., alciopin phyllodocids) the eyes are a very large camera construction with a distinct lens and with focal capabilities of considerable complexity. In the polychaete eyes, the sensory cells penetrate through a pigment layer (retina) into the space in which the lens is supported and the optical nerve is collected behind the pigment layer, thus the retina is reversed compared to vertebrate eyes.

The nuchal organs are usually located in pits, in short grooves or on short ridges near the junction between the prostomium and the first segment. The nuchal organs may be very large, complexly folded structures (caruncles) in the euphrosinids and amphinomids. And the epaulets present in some syllids are also nuchal organs. In some spionids the nuchal organs stretch along most of the anterior region of the worm as narrow, ciliated attached ridges. The presence of nuchal organs has been considered diagnostic for the polychaetes in relation to the clitellates, which lack these

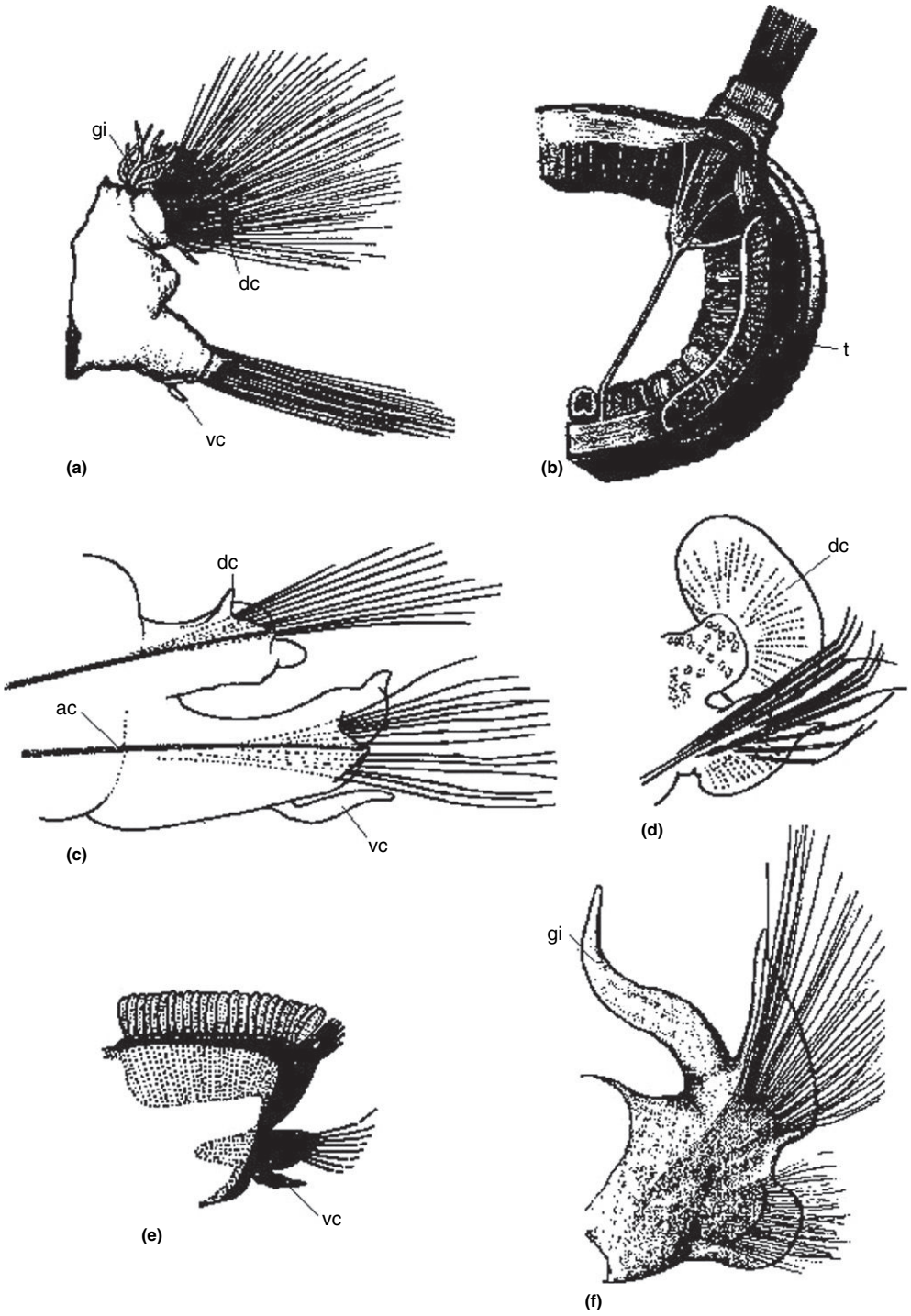


Figure 4 Polychaete parapodia. (a) Amphinomid; note similar rami. (b) Arenicold; note neuropodia modified to a torus and the musculature linking the base of the chaetae to the ventral body wall. (c) Paralacydoniid; note unequal rami and the notable presence of aciculae. (d) Phyllodocid; note large, foliose dorsal cirrus. (e) Chrysopetalid; note that some of the notopodial chaetae are large, flattened paleae covering the dorsum. (f) Spionid; note the dorsal flattened gill. Modified from Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

structures. However, some polychaetes lack nuchal organs, and sensory structures closely similar to the nuchal organs are present in sipunculans, so the presence of these sensory organs may be a shared feature of the annelids and the loss among certain annelids may be a secondary phenomenon.

Ciliated lateral organs are located between the parapodial rami in many polychaetes, they are usually tufts of cilia in a shallow pit, but may form eversible papillae in some taxa. Similarly constructed patches of cilia have been reported from a variety of annelids and may form patterns of sensory organs

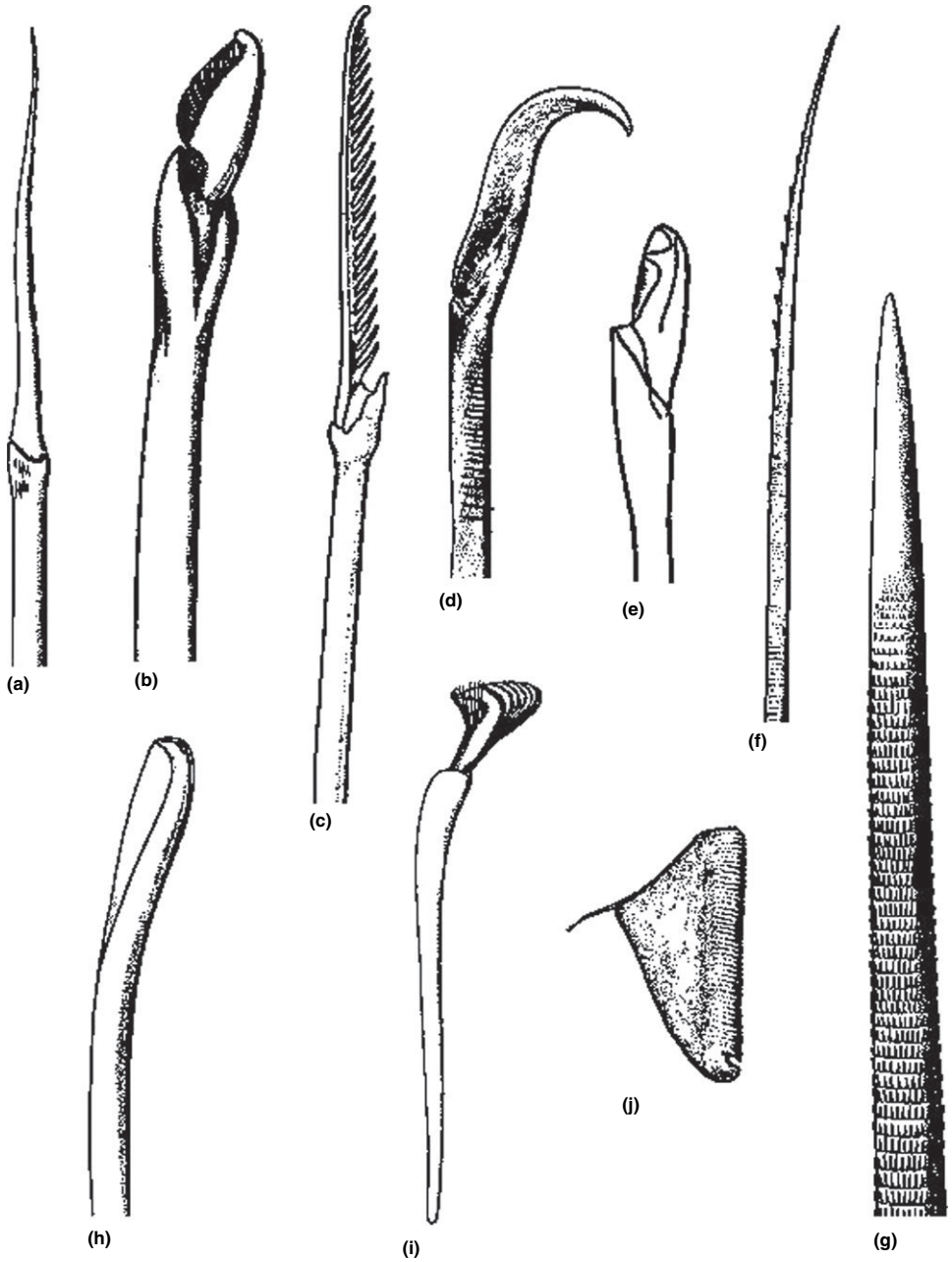


Figure 5 Polychaetae chaetae. (a–e) Compound chaetae showing a variety both of hinges and of the distal ends of the chaetae (a, sigalionid; b, nereidid; c, syllid; d, flabelligerid; e, eunicid). (f) Capillary chaeta of hesionid. (g) Spine of a polynoid. (h) Hook of a spionid; note the two teeth at the end. (i) Hook of a maldanid; note the many teeth at the end and the spray of fibers under the main tooth. (j) Uncinus of a chaetopterid; note the straight edge with the many small teeth along it. Modified from Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

characteristic of each of the major groups. All these ciliated structures are assumed to be at least partially chemosensory, but the experimental evidence is usually missing.

Musculature

Musculature includes two pairs of longitudinal muscle bundles; these bundles are usually ovate in cross-section, but especially in species with massive longitudinal muscles, they are folded over, presumably to avoid buckling during strong contraction. The ventral pair is usually heavier than the dorsal pair, which in some worms (e.g., terebellids) form a thin layer between the upper edge of the notopodial and the dorsal midline. Circular muscles are usually considered to be present; however, at best they are present near the septal boundaries. Most of the other segmental muscles are associated with the parapodia, the chaetae, or are dorsoventral transverse muscles.

Whether these antagonists to the longitudinal muscles represent the remnants of a former complete layer of circular muscles has yet to be determined.

However, it appears likely since most other annelids have the longitudinal muscles represented by a relatively even layer and a distinct circulatory muscle layer is present. The clitellates have longitudinal bundles of muscles, and the small aquatic species lack circular muscles; however, the terrestrial earthworms have a well-developed layer of circular muscles.

Septa

Septa separating the segments are rarely complete, but in many cases they are sufficiently muscularized, with radiating and circular muscles to allow each septum to clamp down around the digestive tract and at least temporarily isolate each segment. In certain families one or more of the anterior septa may be complete (gular membranes) and may isolate an anterior part of the body cavity from the rest of the body.

Other than the polychaetes the septa in other annelids are reduced or missing except the separation isolating an anterior chamber of the body cavity from the rest of the body in the sipunculans.

Digestive Tract

The digestive tract is usually a more or less straight tube running from mouth to anus. The larval stomodaeal region (infolded ectodermal region in the mouth) may be elaborated into an eversible pharynx (proboscis). This pharynx may be axial or may include only the ventral side; it may be sac like or muscular. Jaws are present in some taxa; they are thickened regions of the cuticle and as such consist primarily of sclerotinized protein in which a variety of metal ions may be imbedded, or they may be calcified. Paired salivary glands are often present. The middle of the gut may be invested with a large blood sinus or is at least well supplied with segmental blood vessels. In postlarval siboglinids (Pogonophora and Vestimentifera) the gut is blind ending at both ends and the lumen of the gut is largely obliterated, but a very inflated specialized gut wall is present. The cells of this gut wall contain numerous commensal bacteriae. The scale worms have numerous segmentally arranged blind sacs attached to the midgut; at least some part of the intermediary metabolism takes place in these sacs, which also in part supplements the circulatory system in bringing nutrients to the tissues in which they are needed. Some polychaetes have a looped gut; in these taxa the septation is missing in the region that contains the loop. Looping allows for an increased length of gut in relation to the size of the worms.

The gut in other annelids is most frequently similar to the one in the polychaetes in that it runs from mouth to anus, except that it frequently is rather long and folded. The exception is Siouncula in which the gut is U-shaped, with the ascending and descending parts spiraled around each other. The anus is at the anterior end of the main body at the base of the retractable anterior part of the body.

Gills

Gills are present in a variety of taxa. They are usually associated with the notopodial but may be present anywhere along the body; they are recognizable as thin-walled extensions from the body wall containing a vascular loop in which the two limbs of the loop are linked through capillaries. This definition excludes the structures called gills in capitellids and glycerids, both of which lack a circulatory system; these structures cannot at this time be considered homologous with the gills present in taxa with closed circulatory systems. Functionally, any thin-walled extension from the body wall will have a gas-exchange function whether it structurally belongs to a given category of structure or not. This issue is of considerable phylogenetic interest in that gills must have evolved independently several times, while in other clades, morphologically different structures, including the notopodial cirri, have taken on respiratory functions.

Gills are absent in most nonpolychaete annelids; some leeches have extensions from the body wall with respiratory functions and the anterior feeding apparatus in the sipunculans have a similar function. The echiurans have paired anal sacs that they periodically flush, presumably bringing water into the sacs, which would bring oxygenated water into close proximity to the digestive tract.

Circulatory System

Circulatory systems may be complete, consisting of a dorsal, often contractile vessel, in which the blood flows anteriorly, one or two major ventral vessels, and capillaries connecting these vessels to the gut and to the rest of the organs. Some polychaetes lack the capillary beds and thus have an open circulatory system. And small species often lack a circulatory system entirely, presumably relying on diffusion for transportation of oxygen, carbon dioxide, and so forth. Shared absence of a circulatory system is not considered a feature of phylogenetic importance; it seems to be linked to a reduced body size.

Nephridia

Nephridia may be present already in the larvae as protonephridia; postlarval nephridia may be either protonephridia or metanephridia. Protonephridia open into the body cavity with intracellular slits; metanephridia open with a funnel. In all postlarval polychaetes, the inner end of the nephridia is located in the segment preceding the one in which the duct opens to the exterior. About 100 years ago Goodrich identified a series of possible fusions between the nephridia and the exit ducts for the gonads (e.g., protonephromixia, metanephromixia, mixonephridia). New information suggests that this classification may not represent the situation accurately, and that it is more complex than previously understood. Nephridia are often present in many segments, especially when the segments are similar in structure; however, in many polychaetes the number of pairs of nephridia may be limited to four or five. Excretory nephridia are always located anterior to nephridia, which functionally become gonoducts when the worms are sexually mature.

Gonads

Gonads may be present in many segments, or they may be limited to several or even just a few segments in the middle of the body. Sexes are most frequently separate, but hermaphroditic taxa are scattered throughout the group. The gonads are usually very simple, part of the peritoneal lining of each segment, sometimes attached to the transverse or oblique muscles. In many polychaetes only the first few cell cleavages leading to eggs and sperm take place in the gonads proper, the rest of the development takes place in the coelomic fluid in which the eggs and sperm float. The developing spermatocytes may have incomplete early mitotic cleavages, leading to flattened sheets or rounded balls of cells in which the head end of the sperm is embedded in the shared structure; these structures are called sperm morulae. The yolk of the developing eggs may be furnished by nurse cells, either in a layer or as two strings. However, in many species no nurse cells are present, but the coelomocytes function at least in part to furnish the building blocks of yolk for the eggs and the general nutrients to the developing sperm.

Reproductive Biology

Most species spawn in the water, the site of fertilization as well as development up to settlement; indeed, several species swarm on the surface, such as the palolo worms of tropical waters, sometimes while emitting luminescence (e.g., *Odontosyllis*). A few species have modified some of the parapodia for copulation, but this is rare and limited to certain small species. However, we are increasingly finding that females may store sperm in sperm receptacula, which may be segmentally arranged at or near the anterior end. Sperm morphology has been related to the modes of fertilization. The round-headed sperm with a short acrosome and a long tail has been considered primitive since it is present in taxa that spawn into the water column. When fertilization takes place elsewhere, the sperm usually has another structure, often with greatly elongated heads and short tails.

In taxa with both feeding and nonfeeding (or brooded) larvae, eggs are much larger in the taxa in which the larvae do not feed. Nonfeeding larvae are present in all major groups of polychaetes and in some families they are the only kind of larvae present. The shift between feeding and nonfeeding larvae must have evolved more than once; the direction in which this shift has taken place cannot be considered settled in many cases.

Early development may take place in open water, in deposited egg masses, in brood chambers of various sorts, or internally, in the body of the females. After the early cleavages, which follow a pattern called spiral cleavage, as in many related phyla, and gastrulation, a trochophora larva develops. The trochophore consists of any upper episphere with a sensory tuft of cilia on the top and often a pair of small eyes. The episphere is separated from the hyposphere by an encircling prototroch, a ciliary band that in the polychaetes may consist of anything from a single row of ciliated cells to a broad densely ciliated girdle. The mouth is located on the ventral side, immediately posterior to the prototroch and is often associated with a longitudinal midventral ciliary band, the gastrotroch. At the posterior end of the larva, a third ciliary band, the telotroch, is often also present. In feeding larvae the gut is fully developed and a pair of larval nephridia is present. When metamorphosis starts, segmentation starts posteriorly, just in front of the telotroch, internally the segments are first visible as paired mesoderm blocks on either side of the gut. These blocks eventually become hollow and the mesoderm forms the peritoneal lining and muscles and most other structures in the developing segment. The space inside each block becomes the secondary body cavity, the celom, which eventually contacts the exterior through nephridial ducts. Metamorphosis in most polychaetes is a gradual process of adding more segments with a progressive reduction in the prototroch and other larval structures; the body proportions of a newly settled juvenile may reflect what the adult worm will eventually look like, but in many cases considerable change in relative proportions of various body regions takes place during the later development. The body of the polychaetes may appear to consist of a series of repeated units, but as we become more familiar with postmetamorphic development it is becoming increasingly obvious that the polychaetes maintain integrated bodies with distinct proportionality between different regions.

The clitellates are hermaphroditic and, especially the terrestrial ones, have developed complex mechanisms to ensure fertilization and protection of the fertilized eggs. The clitellum, the excretory girdle will during the copulation secrete a cocoon in which the larvae develop. Clitellate development is always without a trochophore as described above for the polychaetes. Both the echiurans and the sipunculans have a trochophore. In the echiurans, metamorphosis follows a relatively brief trochophore stage. One unusual feature of the bonellid echiurans is that the larvae have indeterminate sex. If the metamorphosing larva settles directly on the substrate, it will become female; however, if the larva settles on the elongated prostomium of a female, it will become a dwarf male and live permanently connected to the female, essentially becoming a ready source of sperm when the female spawns.

The sipunculans follow the first trochophore with a prolonged larval stage in which they may stay in the plankton for several months. If an early trochophore happens to hit the substrate prematurely, it will not metamorphose, but will most likely die.

Systematics, Diversity, and Phylogeny

The phylum Annelida has recently undergone a massive restructuring as indicated above. The following outline of these changes follows some of the major changes, but all issues are yet to be resolved.

Traditionally the polychaetes were separated into errants and sedentaries (two equivalent groups in terms of numbers of species), since the 1950s this system has been replaced by one in which the group is separated into many orders, but without a clearly stated phylogenetic scheme to support the groups (see Fauchald 1977 for an example). The classification presented here was proposed by Rouse and Fauchald (1997)

(Figure 6); it will certainly be improved when more information becomes available. Especially important will be a clearly recognized linkage between the clitellates and the polychaetes. In particular, the group called Scolecida in Rouse and Fauchald turned out to be unsatisfactory when genetic information became available.

In addition to the approximately 80 families traditionally included among the polychaetes, a major change was

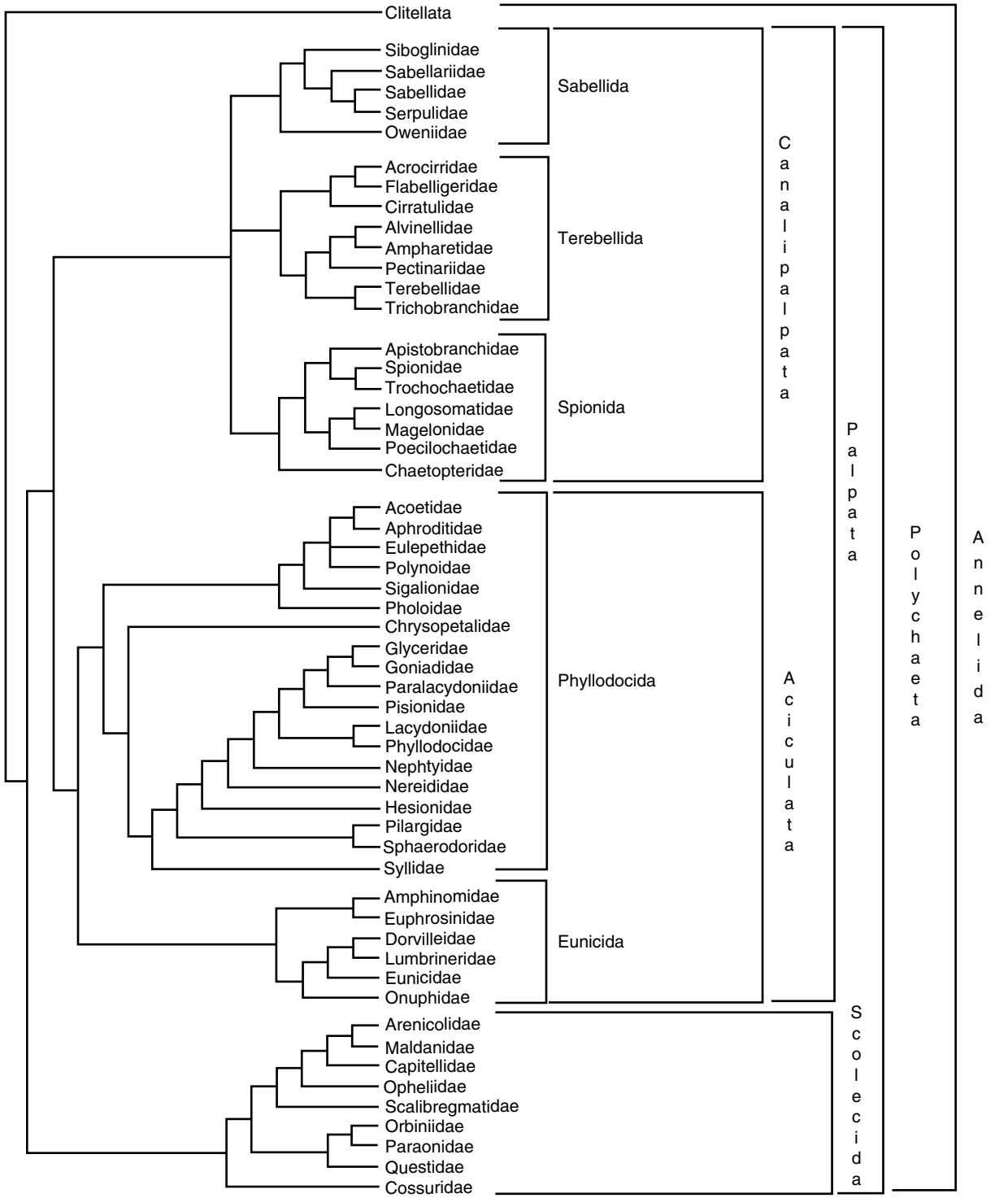


Figure 6 Diagram showing the current systematic subdivision of the Polychaeta. Only 54 of the 80 families have been included; most of the remaining families can be readily associated with the taxa where they fit but are sufficiently poorly known or lack so many features that including them would have distorted the current analysis. Eventually, we expect to be able to incorporate all families. Modified from Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

proposed by Rouse and Fauchald. The pogonophorans (including the vestimentiferans) have been considered a separate pylum (phyla). Their major features match the polychaetes well, and they are included here as single family polychaetes, Siboglinidae. Some of the unusual features, such as the "absence" of a gut are incorrect; the siboglinids have most tissues associated with the gut, but in the adults the gut is closed both anteriorly and posteriorly and the gut lumen is obliterated. Evidence from molecular systematics furnishes good criteria for including the siboglinids among the polychaetes, even if the exact placement among the canalipalpates is still uncertain.

Rouse and Fauchald included the myzostomids among the polychaetes; this group has been considered distinct for more than 100 years, and recently it has been shown that the apparent similarity to the polychaetes is incidental (see especially papers by Eeckhaut and collaborators). The myzostomids are associated with echinoderms and have a fossil record stretching back to the paleozoic, so they certainly have a long separate evolutionary history. Recent molecular studies link the myzostomes to the annelids again.

The fossil record of the polychaetes is, not surprisingly, uneven, considering that they are soft bodied with very few, readily fossilizable structures. They were present in the middle Cambrian (in the Burgess Shale) in taxa resembling in certain respects, both recent canalipalpates and aciculates. In the same material is also found *Wiwaxia*, which may or may not have been a polychaete, but which appears to share certain details with members of a recent aciculate polychaete family, Chrysopetalidae. Jaws of various polychaetes have been reported present from the Ordovician on, these jaws were originally considered a separate group, Scolecodonta, but are now considered disarticulated jaws of aciculate polychaetes, especially the euniceans, which are still common in all kinds of environments. Whole body fossils have also been reported from other Paleozoic strata, and the presence of recent families has been suggested. Many of the whole-body fossils lack the diagnostic characters of recent taxa so the presence is often difficult to document. Calcified tubes similar to those present among the serpulids are present from early paleozoic strata on, some of these tubes may have belonged to nonannelid taxa, but certainly some of them were serpulids. The upshot of the scattered records shows that most major taxa of polychaetes must have been present by the Ordovician and some clades may be even older than that. This suggests that each group has a long separate evolutionary history. Most of the groups were thus present before Pangaea was formed, which agrees with the very wide dispersal of all major polychaete taxa: they have been rafting on the continents. However, there is much evidence to show that individual species are capable of very rapid dispersal as invasive species, probably mostly through transportation of adults.

Polychaetes are important components of the bottom fauna at all depths of the oceans from beach sands to the soft mud of the deep-sea trenches. Recently a rich fauna of very unusual polychaetes has been found in both hot vents and cold seeps. Various polychaetes are also present in the pelagic environment throughout their life, more commonly the polychaetes have benthic adults but are present in the plankton for larvae for a varying period of time. In certain areas the polychaetes may make up as much as 90% of the standing crop of benthic macrofauna, so any account of life in and on the sea bottom must take into account the activities of these worms.

Polychaetes often appear to be very widely dispersed geographically, and some species are considered cosmopolitan. Careful studies have in many cases shown that the records on which these distribution maps were based are erroneous. More than one species have been called by the same name: later investigators have shown that what was originally thought to be a single taxon often represented several distinct taxa. This correction of geographical records is important for our understanding of the biodiversity of many areas, and is in part the result of increased quality of analysis based on the development of statistical techniques and the emergence of molecular systematics or has been demonstrated in experimental studies. Many polychaetes are genuinely widely dispersed however, possibly because they can be transported on ship bottoms and in ballast water. Other forms of accidental introductions also have been demonstrated. Importation of a species of abalone from South Africa to California for use in aquaculture brought along small sabellid polychaetes that burrow into the shell of the South African abalone; the sabellid escaped and is now doing damage to native abalones not only in aquaculture but also in the wild. Other polychaetes, especially members of the family Spionidae, have been transported with importation of oysters, they burrow into the shells of oysters, forming mud blisters, which may have serious consequences for the oyster industry in certain areas.

Polychaetes may move around at the surface of the substrate on which they live, but even the most active species are usually cryptic, nestling in cracks and crevices, under the cover of sea weed or debris or along the edge of rocks where these meet sand or mud. Many polychaetes from a variety of higher taxa form tubes or burrows in which they live more or less permanently. Some, such as the onuphid *Hyalinoecia*, are capable of moving their quill-like tubes along, but others, such as the related *Diopatra* are sessile; in this genus the tube consists of a gray matrix which, where it reaches above the substrate, is decorated with shell fragments, wood or sea grass debris, and sea weed. The sabellids, especially the larger species, form tough, but pliable tubes, but the closely related serpulids have their tubes impregnated with calcium carbonate. Some species, such as certain glycerids, live in complex galleries with several openings on the surface; the worms move around in the galleries so that they can intercept small prey organisms moving around on the surface of the mud or sand.

The presence of polychaete tubes creates a microenvironment in many areas so that the sheer presence of these worms modifies sedimentation patterns and influences the relationship between the surface of the sediment and the overlying water.

In general, polychaetes feed on anything organic. Some taxa are obligate carnivores, but most species will feed basically on any kind of plant and animal fragments present, alive or dead. Several species take food particles out of the water column, mostly by creating small back eddies in flowing water so that particles in transport in the water will drop out of the flow. Only the chaetopterids truly filter the water for contained food and are specialists of that mode of feeding.

The nereidid polychaetes have been shown to be capable of filtering as well, but they are not morphologically structured for this activity. This may be true for several other polychaetes that are capable of building temporary tubes and driving water through the tube with undulatory motions. Incomplete mucus plugs will fetch particles in the stream. Many polychaetes feed by capturing small, light-weight particles in mucus strands and transporting them to the mouth along the palps, feeding mainly on the light fraction of particles in bottom load transport at the water–sediment interface, often creating or taking advantage of irregularities in the bottom. Some polychaetes feed at depth at the bottom of their tubes (maldanids) and are thus capable of turning over sediments. The arenicolds are especially well known for their feeding in the intertidal sands and mud flats, these worms are large and as a consequence have considerable effect on the life conditions for all other inhabitants in the bottom. They are buried sufficiently deeply to avoid being the prey of plovers and other birds.

Certain polychaetes, such as the capitellids, include some of the most pollution-resistant species in the sea, capable of surviving in inner harbor waters where the organic pollution load is very heavy. Other polychaetes, such as some glycerids are extremely sensitive to change in sediment composition and oxygen levels in the water. This range of reactions to disturbance has made polychaetes much used in environmental research. Most of the taxa used have short generation spans and at most a very short larval life, making them ideal subjects for laboratory studies of heavy metal ions toxicity. Most of the species used in this kind of research can be maintained under very simple conditions and have contributed to giving polychaetes a reputation for being generally pollution resistant, but it is likely that this is a sampling error. The species used in research easily go into culture. The more sensitive taxa are much more difficult to keep in culture, but may be more generally characteristic of the group as a whole in terms of physiological requirements.

The clitellates include three ecologically very different groups: the small, fresh-water and marine microdriles; the large, terrestrial earthworms, and the leeches. The microdriles are at least partially very widely dispersed geographically, but thanks to increased study, characteristic faunal elements are becoming better documented. The earthworms include several species that have been transported by humans in soil. Many of them are invasive and move into new environment easily.

The leeches are interesting in that they are essentially carnivores, or parasitic on animals, but phylogenetical they are derived from groups that appear to have been feeding basically on detrital particles. The leeches are present in all three major environments, marine, freshwater, and terrestrial. The terrestrial leeches are linked to high-humidity environments. The best known leeches include the bloodsucking leeches, *Hirudo medicinalis*, which were used in bloodletting for curative purposes in Western medicine. Some leeches are still used in surgery in helping with removing blood clots and other material at or near an incision. The leeches that have been sold as medicinal leeches turn out to belong to more than one taxon; but apparently the curative role is independent of species.

Echiurans are present in soft sandy and muddy marine environments where they can build up very dense populations. The best known echiuran is perhaps *Urechis caupo* (the fat innkeeper) from the US west coast. This is a very large, sac-like echiuran living in burrows in which it hosts a whole fauna of associated animals, most of whom live by stealing the food-nets produced by their host.

The sipunculans are present in most marine environments, they usually live in burrows from which they can evert their anterior end with the feeding tentacles and essentially brush up food particles from the surrounding of their burrows. They are often found in empty snail shells, in cracks and crevices in rocks and corals and even in the mangrove root mats in tropical environments. Because of their extended larval life, many species have been reported to be widely dispersed, but exact documentation of the relationships among the populations is often missing.

Appendix

List of Courses

1. Invertebrate Zoology Classes at Upper-Division Levels
2. Systematic Zoology and Phylogeny Classes
3. Morphology
4. Reproductive Biology
5. Systematics, Ecology, and Phylogeny

See also: Invertebrates, Marine, Overview. Worms, Nematoda. Worms, Platyhelminthes

References

Brusca RC and Brusca GJ (1990) *Invertebrates*, xviii, 922 pp. Sunderland, MA: Sinauer Associates Inc.

Fauchald K (1977) The polychaete worms. Definitions and keys to the orders, families, and genera. *Natural History Museum of Los Angeles County, Science Series* 28: 1–188.

Rouse GW and Fauchald K (1997) Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.

Rouse GW and Pleijel F (2001) *Polychaetes*, xi, 354 pp. New York: Oxford University Press.

Struck TH, Schult N, Kusen T, *et al.* (2007) Annelid phylogeny and the status of Sipuncula and Echiura. *BMC Evolutionary Biology 2007* 7: 57.

Worms, Nematoda

Scott Lyell Gardner, University of Nebraska-Lincoln, Lincoln, NE, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Anhydrobiosis State of dormancy in various invertebrates due to low humidity or desiccation.
Cuticle Noncellular external layer of the body wall of various invertebrates.
Gubernaculum Sclerotized trough-shaped structure of the dorsal wall of the spicular pouch, near the distal portion of the spicules; functions for reinforcement of the dorsal wall.
Hypodermis Cellular, subcuticular layer that secretes the cuticle of annelids, nematodes, arthropods (see epidermis), and various other invertebrates.
Pseudocoelom Body cavity not lined with a mesodermal epithelium.
Spicule Bladelike, sclerotized male copulatory organs, usually paired, located immediately dorsal to the cloaca.
Stichosome Longitudinal series of cells (stichocytes) that form the posterior esophageal glands in *Trichuris*.
Stoma Mouth or buccal cavity, from the oral opening and usually includes the anterior end of the esophagus (pharynx).
Synlophe In numerous Trichostrongylidae, an enlarged longitudinal or oblique cuticular ridge on the body surface that serves to hold the nematodes in place on the gut wall.
Vermiform Worm shaped with tapering form both posteriorly and anteriorly.

What is a Nematode? Diversity in Morphology

General Characteristics and Synapomorphies

Despite numerous assertions that appear in the literature stating that nematodes are morphologically conservative, the contrary is actually true: Species of the phylum Nemata are extremely variable in their morphological characteristics. Because of this diversity, almost any broad statement regarding their anatomy probably should be tempered or qualified. Nevertheless, in general, nematodes are nonsegmented worms that generally lack external appendages. Most are vermiform, with tapering anterior (**Figure 1(a)**) and posterior ends (**Figure 1(b)**), cylindrical in cross section, and covered with a usually translucent (**Figure 1(c)**), flexible acellular cuticle secreted by an underlying cellular hypodermis. The cuticle may be smooth (**Figure 2**) or ornamented with rings (**Figure 3(a–c)**), longitudinal striations, spines (**Figure 3(b)**), or spikes or it may have well-developed wing-like structures called lateral alae very commonly on the externo-lateral surfaces (**Figure 4**). Some marine species have long setae (genus *Draconema* Figures 5(a), (b), and 6), which are modified sensory papillae that are probably used in movement and in detecting their environment. In contrast to all other nematodes some marine species have eye-spots that enable them to detect light in their environment (e.g., genus *Thoracostoma* **Figure 7**).

The fine-structure of the external body-cuticle is complex, acting to protect the animal from the external environment allowing it to remain homeostatic inside. The cuticle can be extremely resistant, and depending on the nematode species and its life-history attributes, for example, the cuticle of the nematode may be able to resist digestion in the most inhospitable stomachs in the vertebrate world. Whereas, nematodes may be extremely delicate, able to exist intact only within the osmotically balanced tissues of other animals (e.g., members of the Order (O) Filaroidea) and if moved from the isosmotic solution to one with fewer salts, they may explode and die in a most amazing display.

Nematodes are called "pseudocelomates" because in most forms, their celom is not completely lined with mesodermally derived cells and they are triploblastic. Possessing no circular body muscles, movement is accomplished by contraction and relaxation of longitudinal muscles in apposition to a hydrostatic skeleton. The body wall, is flexible and very strong. All nematodes maintain their form because their body fluids (in the hydrocoel) are under a positive pressure, relative to their environment, analagous to the way a water balloon maintains its shape. The cross-section in **Figure 8** shows *Vexillata* (a nematode that occurs in rodents of the superfamily Geomyoidea, and recently discovered in other rodents) the round shape of the body under the cuticle supports spines or aretes (the whole structure being called the synlophe). Nematodes have a complete digestive system with an anterior stoma just behind the mouth and usually a tri-radiate or triple muscle-pumping type of esophagus (**Figure 1(a)** and **(c)**) that can be muscular or glandular in structure (or both as occurs in certain species of the Filaroidea). The intestine, tubular in form, (**Figure 1(c)**) is usually a single cell in thickness and is lined on the peritoneal side with a thin collagenlike material and, internally, the cells are lined with microvilli to increase the absorptive surface area of the tubular intestine (Grassé, 1965; Maggenti, 1981; 1991a). The tubular intestine extends from the posterior of the esophagus straight to the anus or cloaca, sometimes showing outpouched cecae (blind sacs) or diverticulae near the esophageal–intestinal junction **Figure 9**.

Most nematodes are sexually dimorphic with separate sexes (diecious or amphigonus) and oviparous; however, some are ovoviparous and some are viviparous. Males usually have a cuticularized spicule or pair of spicules that are used to assist in the transfer of sperm to the females. Many species also have a gubernaculum (seen extended from the cloaca of a species of *Ansiruptodera* in **Figure 23**) that guides the spicule during copulation and also eversion of the spicules. Some species are hermaphrodytic and in these cases the nematode produces both sperm and ova from the ovotestis of the same

 http://dx.doi.org/10.1016/B978-0-12-384719-5.00169-6

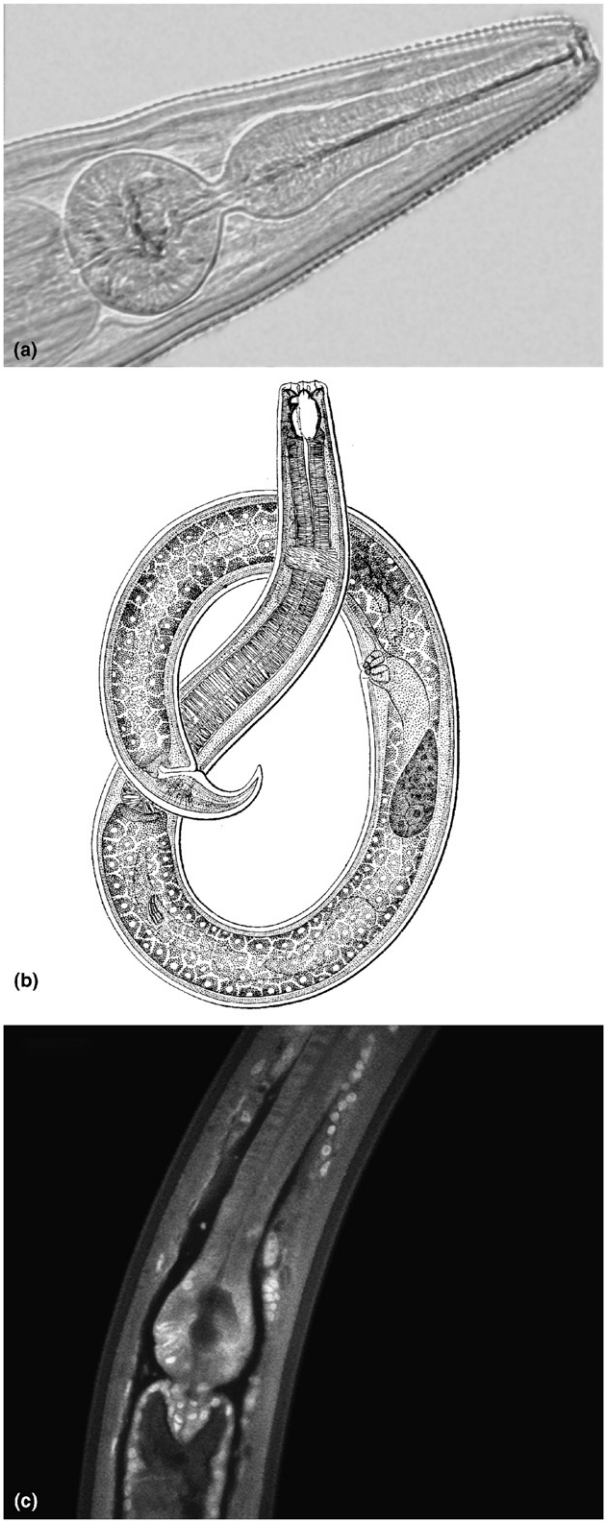


Figure 1 (a) Anterior end of *Didelphoxyuris*, the pinworm of South American Marsupials, showing the small anterior stoma at the narrow end, followed by the well-developed muscular esophagus with a large posterior bulb. The esophagus is tripartite and serves as a muscular pumping organ, pumping food into the intestine. (b) Whole drawing of a predatory nematode of the genus *Mononchus* showing the general structures of a nematode including the tapering anterior and posterior ends and the ventral vulval opening just posterior to midbody. The stoma is heavily cuticularized and bears a spike or spine that is used to puncture the cuticle of the prey. (c) Confocal microscopic image of the exophageal bulb and posterior part of the esophagus of a species of *Pratylenchus*. The cuticle can be seen as the outer light outline of the nematode with an underlying brighter layer showing the cellular hypodermis. The bulb shaped posterior part of the esophagus attaches via a valve to the single cell thick-tubular-intestine, shown here as an optical section just posterior to the esophageal bulb.

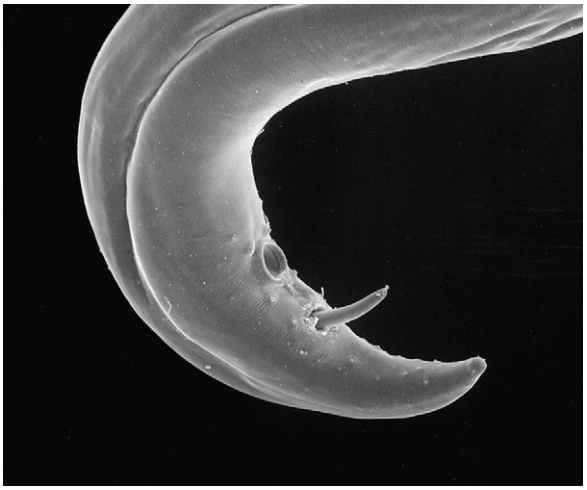


Figure 2 Posterior end (tail) of a male *Paraspidodera* showing the relatively smooth cuticle, a sucker just anterior to the opening of the cloaca, and one spicule protruding from the cloaca.

individual at different times during ontongeny (Maggenti, 1991a; Malakhov, 1994).

A synapomorphy of the Nemata is the presence of non-contractile axonlike myoneural processes or extensions that run from the contractile or body portion of muscle cells to the neural junctions of the nerve cords.

Their width is usually less than 2 mm, even when extremely long. However even here there are exceptions, with the giant kidney worm, *Dioctophyma renale* occurring in mustelids and canids sometimes attaining a size of greater than 15 mm by 1000 mm. The eggs of nematodes are also very similar in size with most ranging from 50–100 µm long by 20–50 µm wide with the exception being those of the antarctic marine nematode *Deontostoma timmerchioi* (40 mm in length of body) that has eggs that range from 870–1100 µm long by 240–350 µm wide (huge by nematode standards).

As mentioned above, nematodes are extremely diverse in size, shape, and structure. The following will only briefly touch on the expansive subject of nematode morphological diversity; for additional reading and exploration see the essential works and references therein of Maggenti (1981); and Grassé (1965 a and b).

External Covering – The Cuticle

An example of diversity in shape is that of the complex cuticular aretes found in species of the O. Strongylida: Trichostrongyloidea (parasitic in vertebrates) in which the exocuticle is modified into a series of cuticular aretes called the synlophe

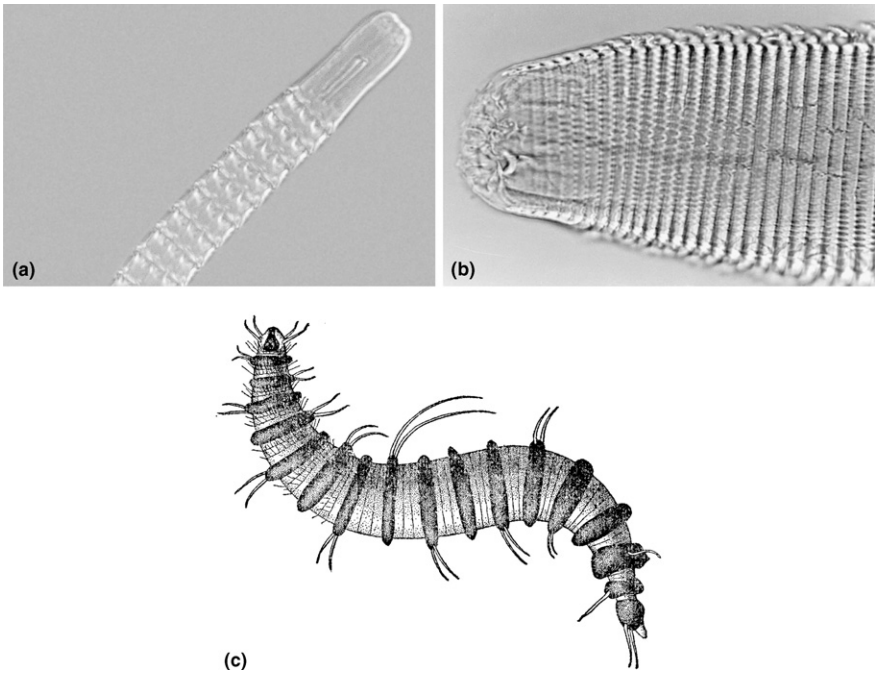


Figure 3 (a, b, c) Anterior ends of marine nematodes showing rings and spines on the cuticle. (c) Whole drawing of a marine nematode of the genus *Desmoscolex* showing heavily cuticularized rings and long setae.

Figure 12. In these forms, it is thought that the ridges running down the length of the body of the nematode are used in maintaining their position in the intestine of their hosts (Figure 8). Other forms have an exocuticle composed of serrated ridges (Figures 3(a), (b), 10, and 11), bumps, or may be very smooth (Figure 2). Some groups (Heterakoidea – parasitic commonly in birds and mammals) possess large cuticularized suckers just anteriad to the cloaca that are surrounded by sensory papillae that evidently allow the male to find the female in the intestinal tract of the host as in *Paraspidodera* of rodents (Figures 2 and 13) and *Ansiruptodera* from rodents (Figure 23). Marine nematodes of the genus *Draconema* (Figures 5(a) and 6) and *Desmoscolex* (3c) have a wildly modified cuticle relative to most nematodes and very large hair-like and segmented cuticular setae that the nematodes use for both movement and detecting its environment (Figures 5(a) and 6). One curious structure that occurs in all Nemata is the amphid, a highly variable sensory organ that can be very obvious as in *Desmodera* (Figure 10) or very inconspicuous as in *Chambersiella* (Figure 18).

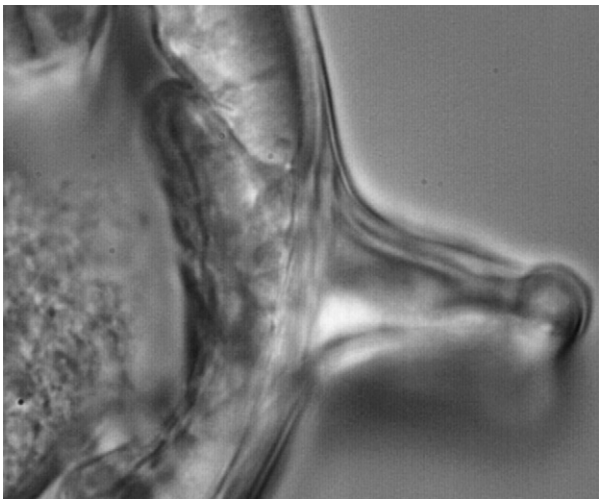


Figure 4 Lateral alae of *Aspidodera*, a parasite of South American xenarthrans, marsupials, and rodents. The ala or wing runs the length of the body possibly providing support for the nematode in the gut of its host. The cellular hypodermis of the nematode can be seen just inside the translucent cuticular layer.

The Alimentary Canal

The mouth of the nematode (Figure 14(a)) provides the anterior opening to the external environment, which connects to the stoma leading posteriad to the muscular esophagus (Figures 15–18). The stoma may be quite reduced or absent as in the case of members of the Trichostrongyloidea (O. Strongylida) or they may be well developed such as that found in *Clarkus* (Figure 16) and *Mononchus* (Figure 1(b)) and capable of inflicting damage on prey items or it may be modified into horny toothlike structures that are used to attach to the intestinal villi of the host animal (e.g., *Ancylostoma* Figure 19). In mammalian gut parasites of the genus *Trichuris* (Adenophorea: Trichuridae) the stoma is lacking and the esophagus is formed of stichocytes comprising a stichosome, which is glandular in function, but many other plant-parasites such as the Adenophorean ectoparasitic, below ground root-feeders (capable of transmitting viruses between and among plants *Xiphinema*, *Longidorus*, or *Dorylaimus:* O. Dorylamida: Longidoridae) have a tubular stoma with the posterior parts of the stoma modified into a spear that is used to penetrate plant

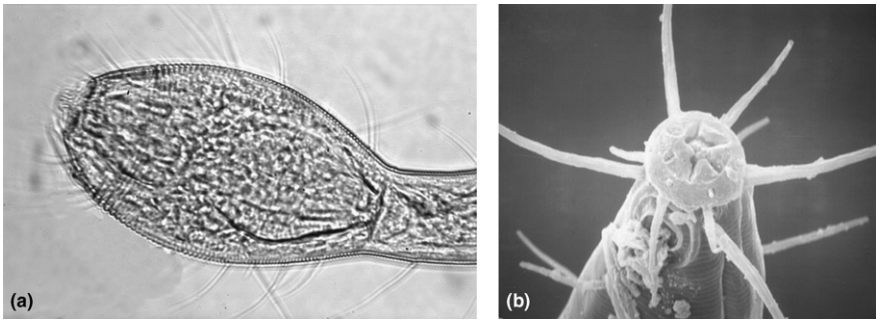


Figure 5 (a, b) The expanded head with visible cuticularized stoma of a species of *Draconema*, a marine nematode. Long setae that are probably used to move and to sense the environment are visible in this photograph. (b) Scanning electron micrograph of the anterior end of a marine nematode showing the long setae and a well-developed amphid just ventral and posterior to the lip rings.

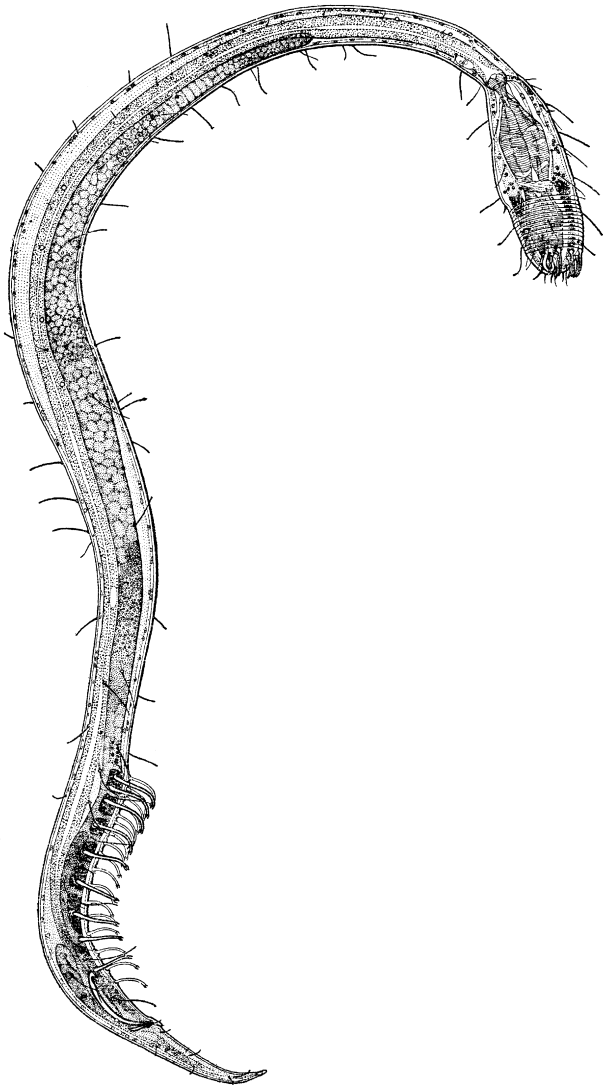


Figure 6 A whole drawing of a male of *Draconema* sp. illustrating the large setae, spicules, testis, and large expanded stoma and head.



Figure 7 Cuticularized "eye-spots" in the marine nematode *Thoracostoma.* Located at about the mid-part of the esophagus, this nematode can detect light in its environment.

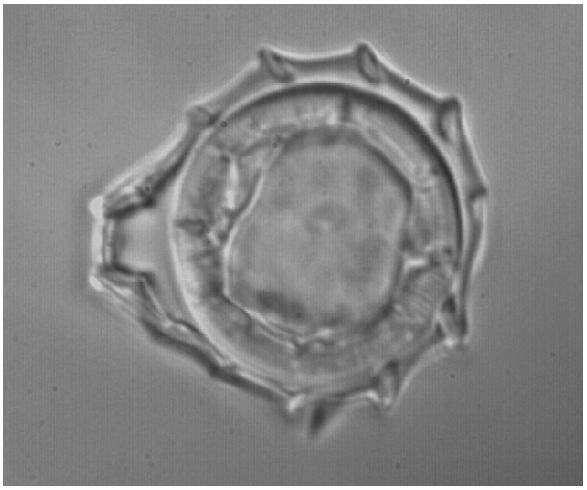


Figure 8 Cross section at midbody of *Vexillata armandae*, a species of nematode parasitic in rodents of the genus *Perognathus* in New Mexico, USA. This photograph shows the spines (aretes) in the cuticle, called the synlophe in Trichostrongyloidea. Inside is the hypodermis lining the body of the nematode. The reproductive tract and the intestine are not visible in this photograph.

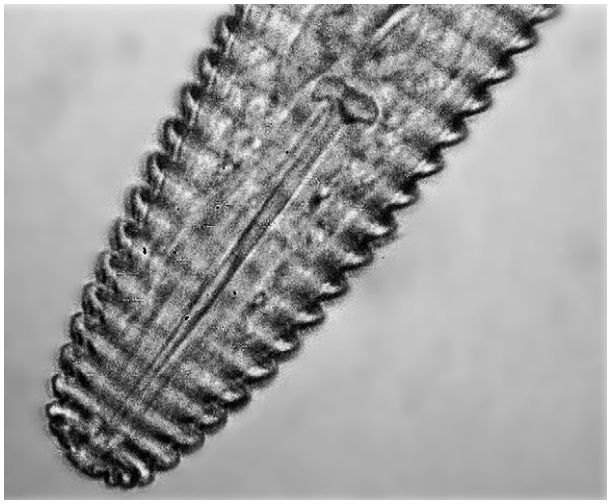


Figure 9 Anterior end of a species of *Criconemoides*, an external root feeding plant parasite. The rings of the cuticle can be seen from the head end posteriad, and the strong cuticularized stomatal spear is clearly visible in this photograph.

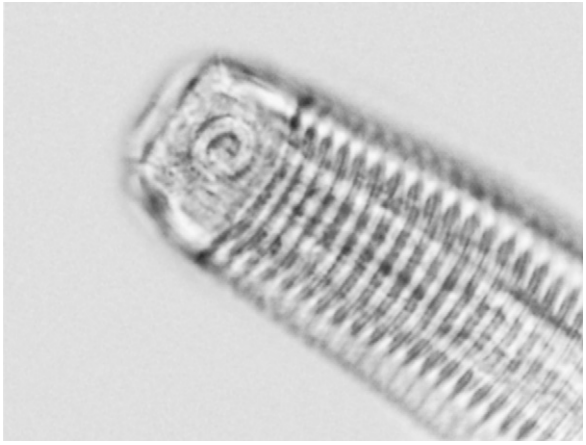


Figure 10 Anterior end of a marine nematode of the genus *Desmodera* showing the rings of the cuticle and a well-developed circular amphid.

cell-walls as in *Dorylaimus stagnalis* (**Figure 20**). The Secernentean plant endoparasitic nematode of the genus *Pratylenchus* (**Figure 21(a)**) also has a modified stomatal spear (much more delicate than those found in *Dorylaimus* sp.) with which it penetrates plant cells as the nematode moves through the tissues and cells of the plant (usually below ground).

The Reproductive System

The reproductive system of most animal parasitic nematodes is adapted to produce extremely large numbers of eggs (e.g., *Ascaris*). Some characteristics of larger animal parasitic nematodes include very large body, two large ovaries, and an equally large uterus. Free-living nematodes generally have much smaller bodies, and are therefore individually less prolific. However, the reproductive systems of most nematodes have the following basic structural similarities: Most

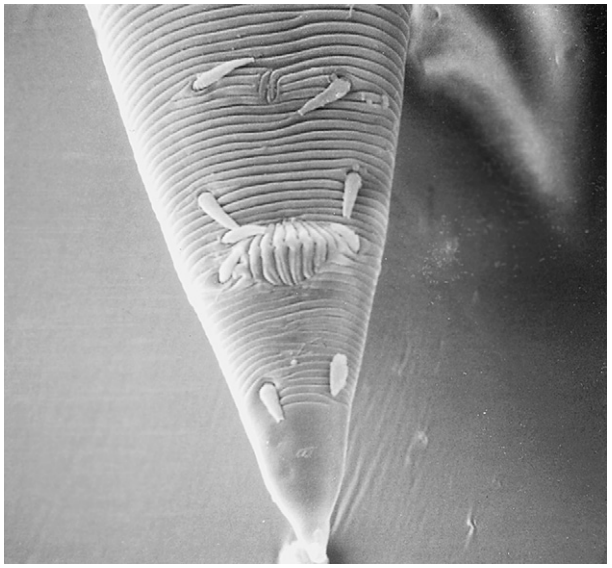


Figure 11 Scanning electron micrograph of the posterior end of a male marine nematode showing rings of the cuticle and small papillae lateral to the cloaca.

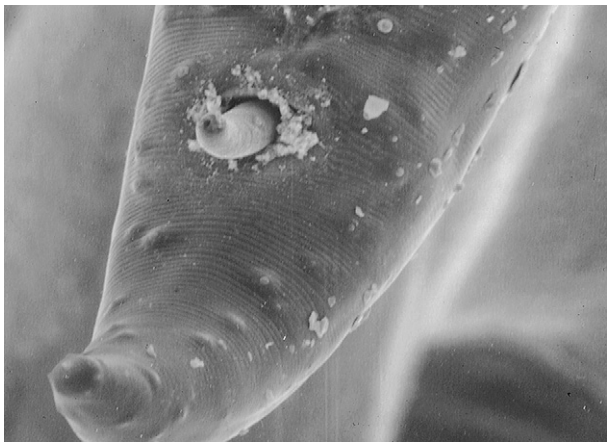


Figure 12 Scanning electron micrograph showing the posterior end of a species of *Paraspidodera* (a parasite of the cecum of rodents of the genus *Ctenomys* in Bolivia) with a spicule everted.

females have two ovaries, one anterior (as in *Pratylenchus* **Figure 21(b)**) and one posterior, each connected to an oviduct and uterus with the uteri connecting to the vagina, terminating in the vulva. Males usually have one testes as in the Secernentea or two as in the Adenophorea, a seminal vesicle and a vas deferens (Maggenti, 1991a) connecting to the outside via the cloaca (**Figure 13**), which is a joining of the reproductive system and the rectum. Secondary sexual organs in male nematodes are usually much more pronounced and variable than that of the female (**Figures 2**, **6**, **11**, **12**, **13**, **22**, **23**, **24**, and **25**) and most males have one or two spicules (**Figure 22**), a spicular pouch, sometimes with a spicular sheath (**Figure 24**), and a gubernaculum. (**Figure 23**). Some males of the O. Strongylida have a well-developed copulatory bursa (**Figure 22**).



Figure 13 Close up scanning electron micrograph of the sucker and cloacal opening of a species of *Paraspidodera* (a parasite of the cecum of rodents of the genus *Ctenomys* in Bolivia). Numerous sensory papillae can be seen on the cuticle around the cloaca.

Some plant-parasitic nemas produce eggs directly into the plant where the nematode lives (in the case of species of the genus *Pratylenchus,* the female produces eggs that hatch within the tissues of the plant and the juveniles begin feeding), whereas in others such as *Heterodera* (which are ectoparasites) the body of the female fills with eggs, forming a sac that eventually dries and transforms into a resistant cyst with the juveniles within capable of resisting environmental extremes.

The Ubiquitous Nature of Nematodes

There are few habitats on earth unoccupied by nematodes. Over a century of both biological surveys and informal collecting has led many biologists to believe that the Phylum Nemata is probably the most ubiquitous of all animal groups. Early in the history of scientifically based biological investigations, pioneers of microscopy opened a new window into the previously unseen microscopic world of the soils.

Descriptions of the "invisible" life seen by pioneers of microscopy such as Antony van Leeuwenhoek evoked images of a wonderfully diverse and dynamic community of worms and other organisms. Subsequent investigations by other early researchers began to open up the unseen world of the nematodes.

Early Views of Nematode Diversity – The Human Perspective

Humans have been parasitized by nematodes from the earliest times. Eggs of the pinworm *Enterobius vermicularis* and the whipworm *Trichuris trichuria* occur in coprolites dated to about 7000 years old (y.o.) from dry areas of Peru and eggs of the hookworm *Ancylostoma duodenale* have been reported from coprolites dated to around 7230 y.o. in eastern Brazil (Arujo *et al.*, 2008). *Ascaris lumbricoides* has been positively identified from human coprolites dated to about 28,000 y.o. from caves in France, but this is the only occurrence of a record this old. The seeming dearth of other nematodes from human remains older than about 7000 years appears to be due to the fact that organic material comprising the coprolites themselves do not preserve well enough to last that long (Karl J. Reinhard, personal communication).

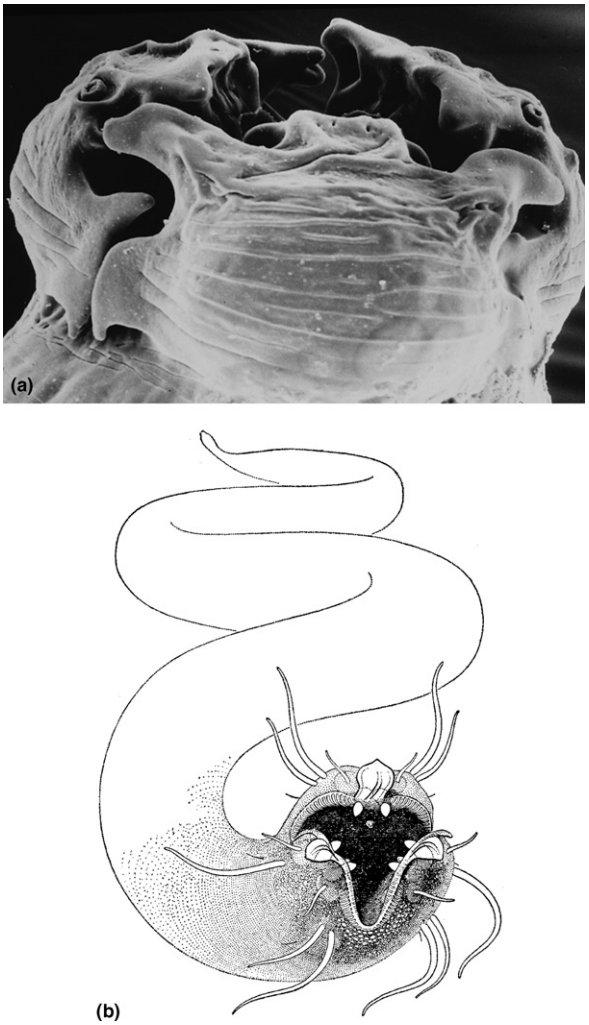


Figure 14 (a) Anterior end of a species of *Paraspidodera* (a parasite of the cecum of rodents of the genus *Ctenomys* in Bolivia) showing three huge lips, the stomatal opening in the middle of the lips and large anteriorly directed sensory papillae on lips 2 and 3 near the outer part of the photograph. (b) Head-on-view of a predaceous nematode as might be seen from the perspective of a mild mannered bacterial feeding form such as *Caenorhabditis* just before it is devoured.

In what is thought to be the oldest surviving written account of *Ascaris* in humans (dated to approximately 4700 y.o. (in China)), foods to avoid and a description of the symptoms of humans infected with these worms were accurately given (Maggenti, 1981). In the area of the Nile River Valley, early Egyptian physicians recorded the presence of both *Ascaris* and *Dracunculus* (the "Guinea Worm") in an ancient papyrus manuscript (written by Egyptian physicians around 3552–3550 y.o.) which was obtained and translated by the Egyptologist "Ebers" in 1872 (Chitwood and Chitwood 1977; Maggenti, 1981). In the extant literature, the first mention of a nematode from a nonhuman animal was by Hippocrates



Figure 15 Anterior end of a species of *Ansiruptodera* from a rodent from Bolivia showing the anterior end of the muscular esophagus attaching to the partially muscular and cuticularized stoma.

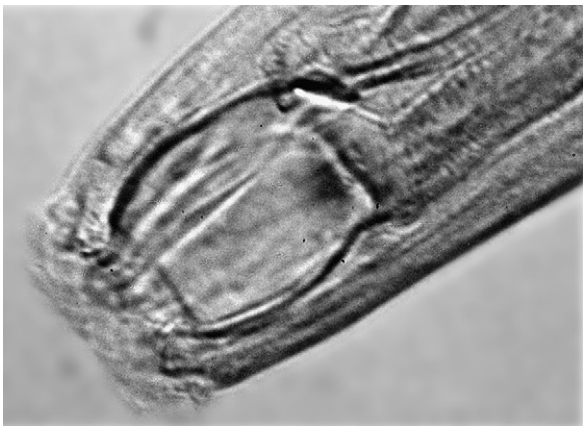


Figure 16 Heavily cuticularized stoma of a predaceous nematode of the genus *Clarkus*. This species is a soil predator living in rhizosoil of grapes in Northern California.

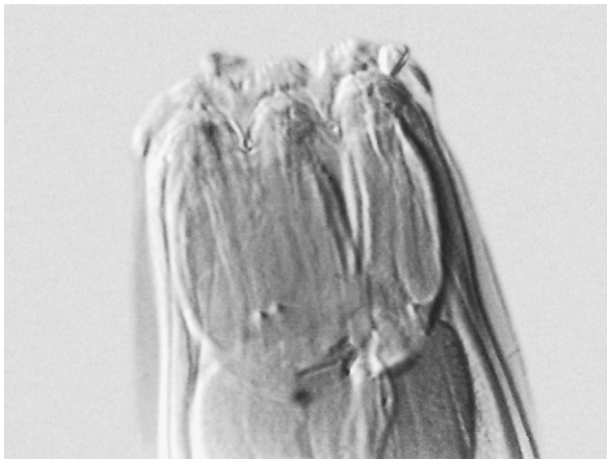


Figure 17 Anterior end of *Miconchus* sp. showing well-developed lips with sensory neurons running posteriad from the anteriorly directed sensory papillae on each lip.

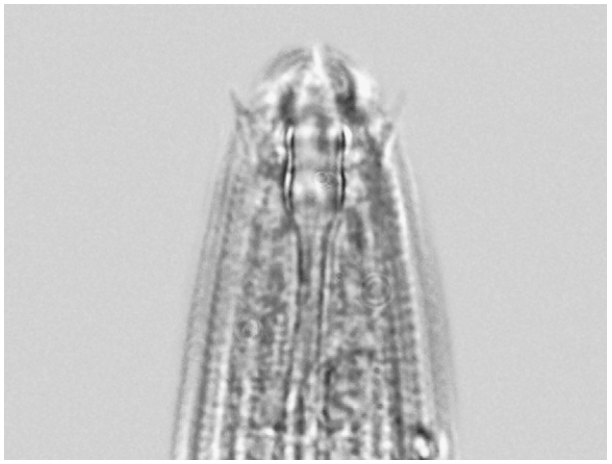


Figure 18 Delicate but well cuticularized stoma of *Chambersiella*.

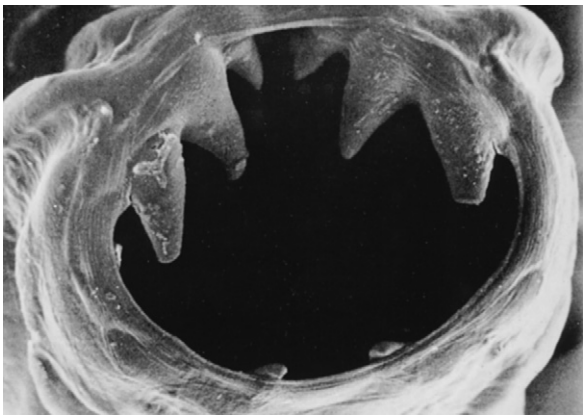


Figure 19 Scanning Electron Micrograph looking into the mouth of a species of mammalian parasite of the genus *Ancylostoma*. The teeth are used to attach to the villi of the host mammal.

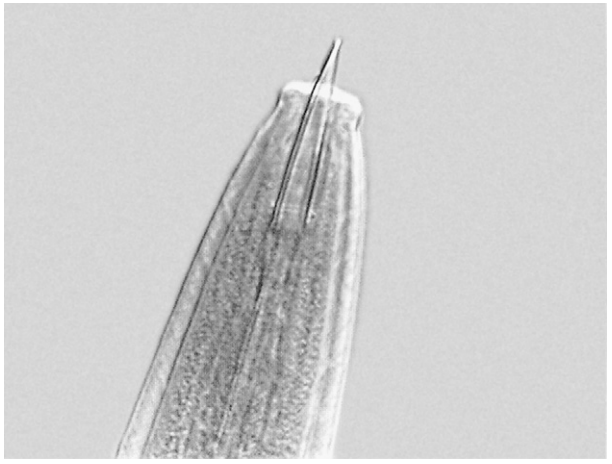


Figure 20 Anterior end of *Dorylaimus stagnalis*, a plant root ectoparasite showing the well-developed stomatal spear that is used to penetrate plant root cells.

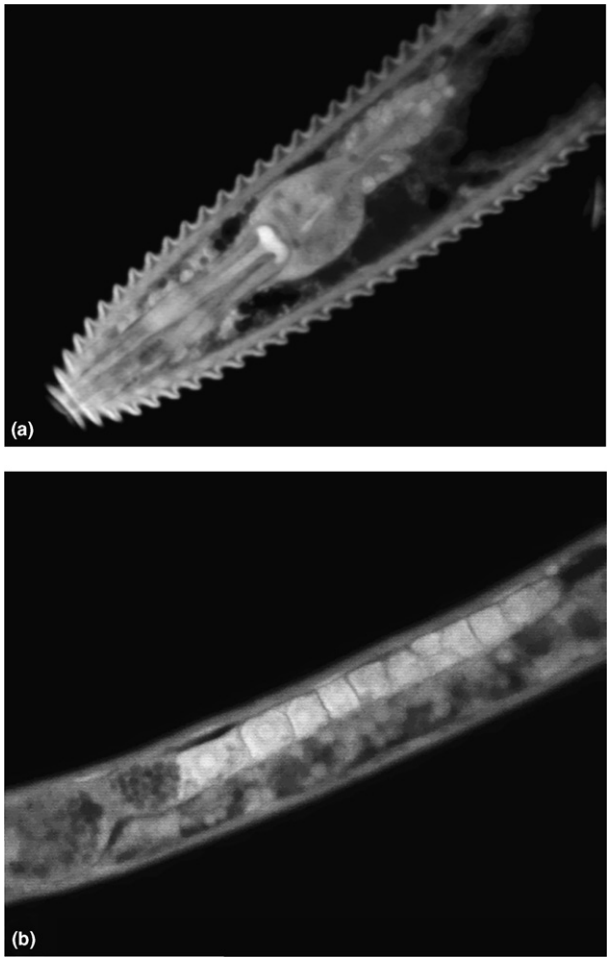


Figure 21 (a) Confocal image of *Pratylenchus* sp. showing the well-developed spear which the nematode uses to penetrate plant cells while moving through the roots of the plant. The muscles and glandular part of the esophagus are also visible. (b) Confocal image of *Pratylenchus* showing the anteriorly directed ovary with individual ovocytes and their nuclei visible.

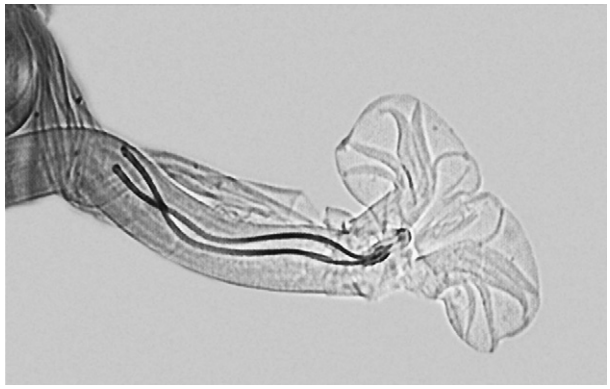


Figure 22 Posterior end of a male trichostrongyloidea showing the well-developed copulatory bursa, bursal rays, and the long thin paired spicules.

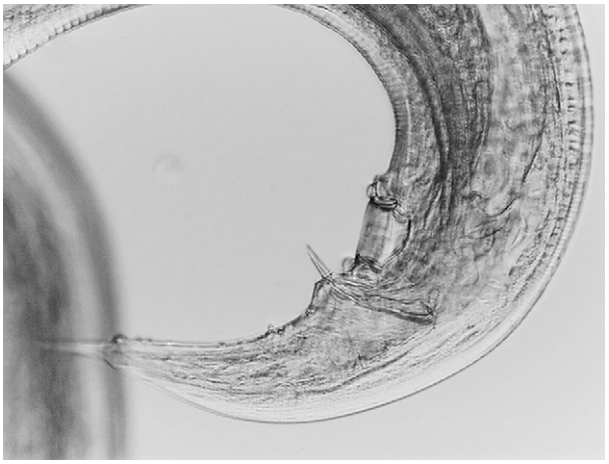


Figure 23 Posterior end of *Ansiruptodera* from a rodent of the genus *Oxymycterus* from Bolivia. The gubernaculum can be seen protruding from the cloaca of this specimen. The cuticularized sucker is also easily visible surrounded by sensory papillae.

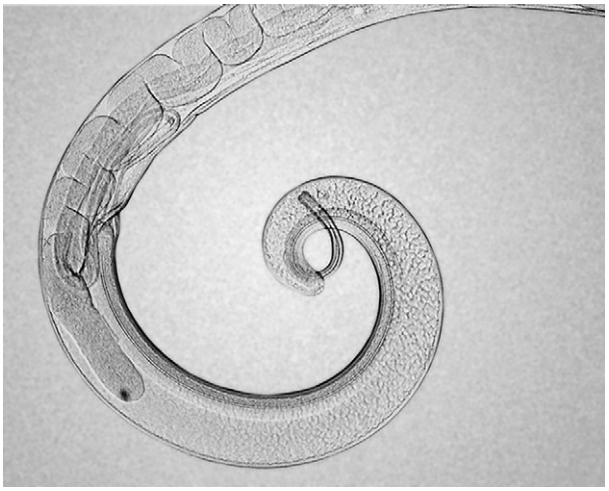


Figure 24 Posterior end of *Trichuris* from a Bolivian species of *Ctenomys*. The long spicule and everted spicular sheath can be seen.

about 2430 years ago where he described the occurrences of pinworm nematodes of both horses and human females. From that time, little further was discovered until Albertus Magnus and Demetrios Pepagomenos (in the thirteenth century, Rausch, 1983) recorded nematodes from falcons (also see Chitwood and Chitwood, 1977).

With the development of the microscope and the emergence of Europe from the dark-ages, knowledge of nematodes as parasites of plants and animals and of free-living forms expanded rapidly. It was found that nematodes occurred everywhere people looked, in fact, Anton van Leeuwenhoek first recording the presence of vinegar eels "*Anguillula aceti*" in his vinegar stored for personal use in a letter dated 21 April 1676 although he was not aware that others had reported finding nematodes some time earlier (Dobel, 1932).

Recent estimates of infections of humans with common human parasitic nematodes give the following numbers

(Crompton, 1999): Of a total human population size of about 6 billion individuals (in the year 2000) the strongylid hookworms, Ancylostoma duodenale and Necator americanus infect about 1,298,000,000 (22%) and the large intestinal nematode, Ascaris lumbricoides occurs in about 1,472,000,000 (25%) people at any one time in the world. Obviously, many people harbor more than one species of nematode at a time, and it is common for people to sport Ascaris, Necator, Trichuris, and Enterobius simultaneously. The author provides the following estimate to indicate just how important these organisms are in the web of life on earth. To put the number of infections of just humans in perspective, the author made the following extrapolations: An adult female Ascaris produces 200,000 eggs per day at an average rate of about 5grams of eggs per year. Actual data for Ascaris in humans that are infected show an average of 18 worms per infected person. Given that ½ of these are females, the author calculated that nine worms per person will produce about 45 grams of eggs in the feces of the host per year. In one year the total population of Ascaris in humans worldwide is conservatively estimated to produce 66,240,000 kg or 66,240 metric tons (72,864 tons (English)) of eggs; this is equal in weight to about 348 large adult blue whales, 8,832 adult male elephants, or 364 fully loaded railroad coal cars.

Estimates of the number of human infections in the year 2000 by other species of parasitic nematodes are shown in **Table 1**. At the present time, approximately 138 species of nematodes have been reported from humans (Crompton, 1999) with 32–36 being host specific.

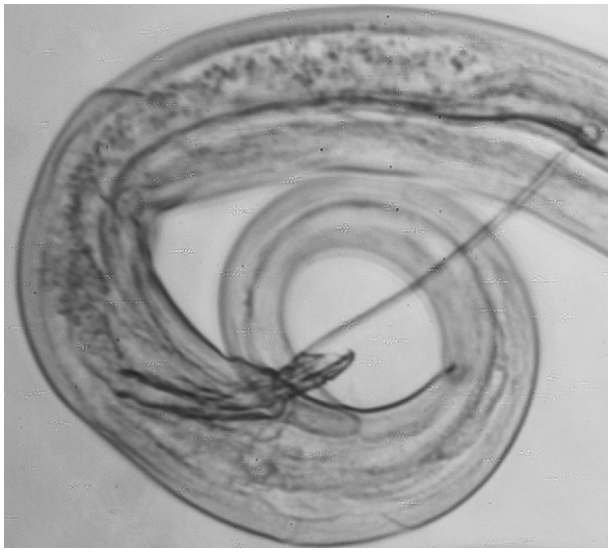


Figure 25 Posterior end of the filaroid nematode *Litomosoides* from *Ctenomys opimus* from high-altitude western Bolivia. Note the dimorphic nature of the spucules in this species one being long and filamentous the other being short and stubby.

The Science of Nematode Diversity

Nathan A. Cobb, one of the first scientists in North America to advance nematology, was a student of the renowned German zoologist, Ernst Haeckel. After just a few years of research spanning the globe from Europe, Australia, and North America, Cobb amassed a huge amount of knowledge and came to have a deep appreciation of the immense number of species that existed. With scientific knowledge based on keen observational skills, he understood the nature of both the great numerical density and species diversity of nematodes in all habitats of the globe that he examined. Thus armed, he wrote the following:

> *"In short, if all the matter in the universe except the nematodes were swept away, our world would still be dimly recognizable, and if, as disembodied spirits, we could then investigate it, we should find its mountains, hills, vales, rivers, lakes, and oceans represented by a film of nematodes. The location of towns would be decipherable, because for every massing of human beings there would be a corresponding massing of certain nematodes. Trees would still stand in ghostly rows representing our streets and highways. The location of the various plants and animals would still be decipherable, and, had we sufficient knowledge, in many cases even their species could be determined by an examination of their erstwhile nematode parasites.*

> *We must therefore conceive of nematodes and their eggs as almost omnipresent, as being carried by the wind and by flying birds and running animals; as floating from place to place in nearly all the waters of the earth; and as shipped from point to point throughout the civilized world in vehicles of traffic."* (From Cobb (1914).

As if challenged by this assertion, scientists have tested Cobb's hypotheses by examining the extremes of the biosphere on earth to evaluate the limits of nematode life. Through these investigations, biologists have now shown that nematodes are living and reproducing everywhere on Earth

Table 1 Numbers of common nematode infections in humans worldwide

Species of Nematode	*Numbers Infected*	*Distribution*
Ancylostoma duodenale and *Necator americanus*	1,298,000,000	World-wide
Ascaris lumbricoides	1,472,000,000	World-wide
Brugia maylayi and *B. timori*	13,000,000	South Pacific, SE. Asia, India
Dracunculus medinensis	80,000	Sub-Saharan Africa and Yemen
Loa loa	13,000,000	West and Central Sub-Saharan Africa and Yemen
Onchocerca volvulus	17,660,000	Central and South America and Sub-Saharan Africa
Strongyloides stercoralis	70,000,000	Temperate regions
Trichuris trichiura	1,049,000,000	World-wide
Enterobius vermicularis	400,000,000	Temperate regions

Source: Data from Crompton DWT (1999) How much human helminthiasis is there in the world? *Journal of Parasitology* 85: 397–403.

that water exists in a liquid state even for short periods of time annually.

Diversity of Habitats and Distribution

General

The most obvious ecological character that defines habitats for members of the phylum Nemata is that they all are aquatic animals – to move, live, eat, and reproduce, nematodes must exist in an aqueous environment. This environment includes soils, muds, sands, plants, and animals. They can be found living in soils with moisture contents as low as 5–10%, but in the majority of these cases, the nemas are associated with the roots of plants. It is evident that the environment in which nematodes live constrains their ultimate size. Soil dwelling species live in the water film of the interstices of soil particles. Nematodes, must live and carry out all functions of life in these spaces, thus free-living, or plant-parasitic soil-dwelling forms are usually extremely small. Morphological diversification of nematodes can operate only within the constraints of their life-history parameters, and thus, evolutionary pathways for nematodes living in an interstitial – soil or sediment/sand envionment are limited. Thus these forms have limited abilities to diversify into the nonaquatic regions of the earth, and their size is constrained by this fact.

In contrast to the small sizes of plant-parasitic, free-living marine, freshwater, or soil nematodes (some adults can be as small as (*Criconemoides*) with an adult length of around 250 μm), species that occur as parasites of animals are free from the physical constraints on their size and the bodies of species in mammals are, therefore, relatively large. In fact, the largest nematode thus far recorded is *Placentonema gigantissima* from the blue whale with a length of more than 8 meters (Maggenti, 1981). The ratio between the smallest adult nematode to the largest can be calculated as (0.250 mm (*Criconemoides*)/8,000 mm (*Placentonema*) = 0.000031), compared to the ratio between a shrew and a blue whale (50 mm/24,4000 mm = 0.12), indicating that the size differences in nematodes are three orders of magnitude greater than the mammals.

In the marine-benthic environment Lambshead (1993) has estimated (based on transect data from deep sea benthic samples) that there may be as many as 100 million species of marine infaunal nematodes. As exorbitant as that estimate seems, deep-sea nematodes have been show to be extremely rich in species diversity. One marine sediment sample from the east-Pacific benthos was reported to contain 148 species from a total of only 216 individual specimens examined (Lambshead 1993). At the present time the true extent of species diversity in the marine benthos can only be imagined because fewer than 20 studies of nematode community structure from marine-benthic habitats have been reported (Boucher and Lambshead, 1995). Relatively recent data from Boucher and Lambshead, 1995 shows that in marine environments the highest diversity in nematodes occurs in abyssal benthic sediments.

We now know that great numbers of individuals and species of nematodes live in sediments on the ocean floor distributed from the intertidal continental margins to the benthos of the abyssal zones, even though these marine habitats remain unexplored. Nematodes occur in tissues and organs of all species of vertebrates that have been studied and some, such as *Physaloptera spp*. live, feed, and reproduce in the strongest stomach acids of mammals, birds, and reptiles. Dessicated specimens from both the Arctic and Antarctic have been rehydrated to form viable colonies and living nematodes have been found in the limited meltwater in some of the dry-valleys of Antarctica, an area that is probably one of the most extreme biotopes on the earth.

Extreme Biotopes

Some of the most extreme soil habitats on the Earth exist in the Dry-Valleys of the Antarctic where the annual mean air temperature is −20 °C, and soil temperatures at a 5 cm depth for the two "summer months" range from −2.7° to 15.9 °C. In this area, no vascular plants grow and mosses and lichens are rare; this is the only terrestrial soil-system known where nematodes are the final consumers and are at the apex of the food chain (in this case it seems more of a chain than a web). Three nematode species exist in these dry soils, *Scottnema lindsayae*, a microbivore (feeding on bacteria and yeast), *Plectus antarcticus* (a bacterial feeder), and *Eudorylaimus antarcticus*, an omnivorous predator that presumably feeds on individuals of the other two species (Powers *et al*., 1998). In another biological extreme, Death Valley in California has recorded some of the highest temperatures in North America and the soils of the valley are teeming with nematodes, many of which have been discovered to possess similar adaptive traits to those found in the cold deserts of Antarctica.

Individuals of some species of Nemata are capable of resisting extended periods of dessication, for example, 3rd stage juveniles of *Anguina tritici* have been dried for more than 20 years in a state of anhydrobiosis in which all metabolic activities are shut down (Maggenti, 1981). These nematodes have been shown to have specialized proteins that fold into stable/preserved structures as the organism dries. In this state of crypto or anhydrobiosis, individual nematodes can remain viable through incredible extremes of temperature, desiccation, hypoxia, and even synthetic nematicides designed to kill nematodes (of course the regular biochemical processes that nematicides interfere with are not operational, so the nematode does not notice this particular assault). When water again becomes available, the animal in fact comes back to life when the molecular structures rehydrate and the proteins and enzymes spring back into normal operation.

Habitat Diversity

Nematodes occupy every conceivable life-history niche. There are benthic deep-water marine forms that appear to consume mostly diatoms and others such as species of the genus *Draconema* (**Figures 5(a)**, **6**) that are mostly associated with marine algae, but it is still unclear what they actually eat. Nematodes of the genus *Dirofilaria* live in the aorta and left ventricles of canids and are transmitted by mosquitoes from dog to dog. Species of nematodes live in the hearts of sharks,

and other species occur by the thousands weighing more than one kilogram in the stomach of a pilot whale (SLG pers. obs.), and, as the human consumption of raw marine fish (ceviche and sashimi) has increased, transfer of juvenile *Anasakis, Terranova,* and other anasakines from the fish intermediate host to humans is occurring more commonly and these nematodes are turning up as parasites in the stomach of humans.

Some species feed on fungi in the soil whereas others are trapped and are themselves consumed by different species of fungi. Still other nematodes like *Mononchus* (**Figure 1(b)**) and *Clarkus* (**Figure 16**) are predatory and hunt and eat other nematodes in the soil environment. Free-living bacterial feeding nematodes have been shown to be integral parts in the carbon and nitrogen cycles in healthy soils (Ferris *et al.*, 1998) and recent survey work has shown that undisturbed or natural soils in noncultivated habitats can have as many as 20 times more species than soil from similar areas that have been under cultivation (Al Banna and Gardner, 1996).

Abundance – Estimates and Facts

In addition to the large numbers of species that may occur in any given habitat, nematodes also occur in very great densities, for instance, in sheer numerical density of individuals in any given habitat, nematodes exceed even the mites and beetles combined. More than 90,000 nematodes were recorded from a single decomposing apple, and one report showed 1 cc. of marine mud contained 45 nematodes representing 19 species. Nematodes in marine estuaries occur at high numerical densities with reports of 4,420,000 m^{-2} in surface mud and 527,000,000/acre in the top 3 inches of sand in the Massachusetts coast. Counts and extrapolations for relatively moist soils (10–70% moisture content) worldwide show that in the uppermost levels, nematodes occur in mind boggling abundance: 7–9 billion/acre in undisturbed sod in North China; from 800,000,000 to >1 billion per acre (representing just 35 species), in Utah and Idaho and around 3 billion per acre in low-lying alluvial soils of Europe and other areas of North America.

How do Nematodes Affect the Biosphere?

Because of the huge numbers of nematodes that have been shown to occur in plants, animals, soils, and the benthos, there has been much speculation about the role of nematodes in basic biological processes occurring in the soils of the earth. Cobb (1914) speculated that some nematodes "are beneficial" however he also noted that this area of study of this was still in its infancy. In fact, in 2000, this area of study is still just developing and recent work has shown that nematodes can be good indicators of biodiversity (Bongers and Ferris, 1999). Gardner and Campbell (1989) showed that mammalian parasites with complex life cycles may serve as excellent indicators of areas of high biological diversity. Because of the multifarious nature of parasitic nematodes in mammals, it is expected that these kinds of species may provide biologists with additional tools for identification of areas of high biological diversity.

The fact that parasitic nematodes occur in such high prevalence and numerical densities in mammals should give us pause, there is obviously a huge energy drain on any population of mammal that we should care to analyze, and this energy drain probably causes significant decreases in the number of offspring in any given population over long periods of time.

Soils and Plants

Recent studies indicate that nematodes play a substantial role in the cycling of carbon and nitrogen in the soil environment (Bongers and Ferris, 1999) and it has been shown that the numbers of bacterial feeding nematodes increases as the bacteria increase with annual warming of the soils. In the rainforest of Cameroon, average nematode abundances of $2.04 \times 106/m^2$ of rainforest soil were found indicating that these nematodes play a significant role in carbon flux (CO_2 and CH_4) in this rain forest site (Lawton *et al.*, 1996).

Nematodes occur in, on, and around the roots, bulbs, rhizomes, stems, and leaves of plants. They can cause galls in both the somatic and germinal tissues of the plants, where the nematodes can encyst, and dry, waiting for the next stage in their life cycle. There are species that are almost fully endoparasitic, exiting the plant root or stem only as juveniles. Others are fully ectoparasitic, such as species of *Xiphinema* (the vector of grape fan-leaf virus). Some species spend part of their life-cycle in a plant and part of it in the soil. As mentioned earlier, seed parasites of the genus *Anguina* may spend most of their existence as juveniles in a dried state, the nematodes rehydrating, molting and then as adults crawling up the outside of the plant during periods of high relative humidity and then laying eggs in the seed head. The seeds are consumed by the developing juveniles, and when they dry, the seed coats protect the also dried nematodes within, and are distributed through the environment as are normal seeds (Maggenti, 1981). Nematodes of the genus *Pratylenchus* use their spear to move through the roots of the plants that they infect, penetrating the plant cells with repeated jabs of their stomatal armature.

Predators, Entomopathogenic Forms, Fungal, and Bacterial Feeders

Many nematodes that are found in the soil are either predaceous forms eating other nematodes, or forms that prey on mites, or other soil macro-organisms. The more spectacular predators such as *Mononchus* and *Clarkus* have specialized buccal structures with which to puncture the cuticles of other nematodes (**Figures 1(b)**, **16**). Microbivorous fungal and bacterial feeding nematodes are also extremely abundant with species specializing in their feeding habits on bacteria, fungi, diatoms, and other microscopic organisms. Some, such as juveniles of species of the genera *Heterorhabditis* and *Steinernema* carry special bacterial for the genus *Photorhabdus* in their digestive system. The 3rd stage juveniles wait in the soil until an unwary insect passes nearby, the nematode then homes in on and penetrates the hapless insect, making its way into the hemocoel where it releases the bacteria, which proliferate, killing the host. The nematode then feeds on the

bacterial colony and reproduces in the insect, eventually again producing 3rd stage juveniles, which leave the carcass of the insect and disperse into the soil, waiting there for another insect to invade (for more specific details of the lifecycle of these entomopathogenic nematodes see Gaugler and Kaya, 1990).

Bacterial feeding forms occur in the soil in extremely high numbers. In soils that have not been disturbed, and have a good layer of organic matter, large numbers of all kinds of nematodes occur. One of the most well-known groups is the Rhabditida or the rhabditid nematodes. These forms feed on bacteria and yeasts growing in the soil, and are typically found in high numbers in moist soils with high organic content. The most famous of these forms are members of the genus *Caenorhabditis* of which the complete genome of *C. elegans* has recently been sequenced.

Aquatic and Marine Nematodes

This is an area that is wide open for future biologists. How do the trillions upon trillions of individuals and the millions of species that occur in the oceans really affect the biosphere? Nothing is known on the subject at the present time.

How many Species of Nemata?

Estimates of Number Described

Estimates of the numbers of species of nematodes that are known (i.e., described species) vary widely. However, in 1819, Rudolphi summarized what was known of the nematodes and he recorded 11 genera and about 350 species. Just 115 years later in 1934, Filipjev reported that 4601 species of nematodes had been described with about half free-living and the other half parasitic. By 1950, Libbie Hyman estimated that approximately 9000 species were described (based on her analysis of the Zoological record with descriptions being recorded at a rate of about 200 new species being described per year). In 1981, Maggenti's summary showed around 15,000 described species. The author's analysis from counting additions to the Zoological Record shows that in a 5 year period from 1992 to 1996 numbers of descriptions were relatively stable with approximately 776 new species described (average of approximately 155 descriptions per year). From 1996 to 1998 the numbers of descriptions decreased to 118 per year, most likely due to the continued retirement and expiration of taxonomists.

How Many Species of Vertebrate Parasitic Nematodes Exist?

All species of vertebrates examined thus far serve as hosts for at least one species of parasitic nematode. Some mammalian hosts harbor many species of nematodes that are distributed through several orders and families. Some of these nematodes are highly host-specific, surviving and reproducing successfully only in host individuals comprising a single species or perhaps a closely related group of species. Other nematodes show less specificity, being much more likely to jump from one suitable vertebrate host to another during opportune times during their life-history (Brant and Gardner, 2000).

Within a host, many different types of habitats may be occupied by nematodes. As a species, *Homo sapiens* harbors around 35 species of host-specific parasitic nematodes (Chitwood and Chitwood, 1977). To illustrate the diversity of habitats in a single animal host, humans will be used as an example. In a human, nematodes can occur as juveniles in muscle tissues, usually smooth muscle like the diaphragm or the tongue (*Trichinella*), in the mucosa of the intestine (*Strongyloides*), as migrating forms in blood, and lungs (*Ascaris*), in blood or lymph as microfilariae (Filarioids of humans), as adults in the small and large intestines (*Ancylostoma*, *Necator*, *Ascaris*), as adults in mesentaries and subcutaneous tissues (*Onchocerca*, *Loa*, *Wuchereria*) and *Enterobius* in the large intestine and cecum.

Cobb was well acquainted with animal parasitic nematodes, and his familiarity with host-specificity in the Nemata led him to estimate that as early as 1914, well over 80,000 species of nematodes would eventually be found parasitizing vertebrates alone. Similar estimates could be derived for parasites of invertebrates and plants resulting in totals of free-living and parasitic species numbers far in excess of the approximately 27,000 species of nematodes known today.

> Cobb (1914) stated *"There must be hundreds of thousands of species of nematodes. Of this vast number only a very few thousand have been investigated, and of these, comparatively few with any degree of thoroughness."*

Present estimates are that only approximately 14,000–16,000 species of parasitic nematodes have been described from all taxa of slightly more than 48,000 currently recognized species of vertebrates. In addition, many species of wild mammals each harbor more than two host-specific species of nematode so the above estimate of 8900 species of nematodes only from mammals can be considered very low.

As of this writing, the natural history, development, and transmission parameters of only around 561 speceis are known (Anderson, 1992). As parasites of only vertebrates, Anderson (1992) estimated that there were about 2300 described genera distributed among 256 families comprising about 33% of all nematode genera known. This percentage is approximately equal to the percentage of genera of Nemata presently known in marine and freshwater habitats. It is agreed by most workers that this bias is due to the larger number of parasitologists working on selected groups of Nemata relative to the number of specialists working on the free-living marine and freshwater forms.

If each of the approximately 4450 known species of mammals were infected with only two species of host-specific nematodes, we would expect to find a minimum of 8900 species of parasitic nematodes only in the Class Mammalia.

Pinworms, nematodes of the Order Rhabditida Superfamily Oxyuroidea (**Figure 1**) show high levels of host specificity and it is well-known that almost all species of rodents and primates have one species-specific pinworm, (for instance, *Homo sapiens* harbors *Enterobius vermicularis*). Both recent and historical studies have shown that pinworm nematodes exhibit high levels of coevolution, i.e., concomitant host-parasite

speciation (and host-specificity) with their primate and rodent hosts (Hugot, 1999). At the present time, slightly more than 716 species of Oxyuroidea have been described with the vertebrates hosting about 496 species and invertebrates about 217 species. The greatest diversity in the Oxyuroidea to be found in the future is expected to come from examination of the Arthropoda especially beetles, cockroaches (4000 species described and more than 20–30,000 expected to be found) and millipedes (17,000 species described and more than 60,000 species expected to be found). This gives huge numbers of pinworms occurring just in the cockroaches and the millipedes if each species harbors its own species of pinworm. At least two species of pinworms (genus *Thylastoma*) occur in laboratory colonies of *Periplaneta americana* and more are expected to be found in free-living populations (J. P. Hugot, personal communication).

When adequate surveys are completed, large numbers of species of Oxyuroidea are also expected to be described from Neotropical rodents of the family Muridae. Up to the present time, only around seven species of pinworms have been described from Neotropical murids and these rodents potentially host from 400 to 800 undescribed species of Oxyuroidea, (given that only one to two new species of Oxyuroid nematode is found in each species of rodent examined). J. P. Hugot, personal communication.

Recent comparative studies in the Oxyuroidea of rodents shows that the larger the body-size of the rodent, the larger the body size of oxyurid nematodes that it harbors (Morand, *et al.*, 1996).

Comparative Nematode Diversity of New World Subterranean Rodents Geomyidae

Papers on nematode parasites from rodents of the Nearctic family Geomyidae covering the dates from 1857 to the present were reviewed. Combined with field-collected specimens from the early 1970s up to the present time, the author discovered that six of the approximately 11 nematode parasites reported from pocket gophers in North America, are host specific to only the Geomyidae (**Table 2**).

Of members of the vertebrate Class Mammalia, one of the most complete sets of nematode parasite data exists for rodents of the family Geomyidae. Of the approximately 35 known species of pocket gophers (Wilson and Reeder, 1993), at the present time only 15 species have been surveyed for parasitic nematodes. From those 15 species, six species of nematodes are known to be host-specific only to geomyids. Some nematodes such as the strongylid *Ransomus rodentorum* and the heligmosomid *Heligmosomoides thomomyos* and filarioids of the genus *Litomosoides* have been reported from more than one species of gopher, other nematode species such as *Vexillata vexillata* occur only in gophers of the tribe Thomomyini (genus *Thomomys*) but do not appear to be host-species specific.

These mammals occur in an extremely wide and ecologically variable geographic area (from southern Manitoba and British Columbia south to extreme northern Colombia) therefore, there may be many more undescribed or undetected species of nematodes in these hosts than this analysis provides. In addition, no studies on genetic diversity of nematodes (or any endoparasites) in these rodents have been published, therefore, levels of genetic variation in these nematodes are unknown and the true genetic diversity that exists will probably cause an increase in the number of nematode species that are recognized.

There is little if any evidence of phylogenetic coevolution of the nematode parasites and their pocket gopher hosts. However, all species listed are specific to species of the family Geomyidae and both *Litomosoides* and *Vexillata* appear to exhibit some level of phylogenetic host specificity with two closely related North American species of *Litomosoides* being found only in geomyids (Brant and Gardner, 2000) and species of *Vexillata* occurring more generally in members of the Geomyoidea.

Table 2

Nematode species	*Classification and Location in Host*
	Sub-Phylum Adenophorea
Trichuris fossor	Cecum and large intestine
	Sub-Phylum Secernentea
Ransomus rodentorum	Cecum/small intestine
Vexillata vexillata	Duodenum
Vexillata convoluta	Duodenum
Heligmosomoides thomomyos	Duodenum
Litomosoides thomomydis	Mesentaries
Litomosoides westi	Mesentaries

Nematodes of Tuco Tuco's (Ctenomyidae)

A review of the nematode parasites occurring in Neotropical rodents of the genus *Ctenomys* indicates a considerably more depauperate fauna of nematodes as compared with the nearctic Geomyidae (**Table 3**). Data collected from 1984 up to the present time indicate that nematodes of the genera *Trichuris* and *Paraspidodera* have cospeciated with their hosts and exhibit different levels of phylogenetic congruence relative to their hosts. In addition, nematodes of the trichostrongyloid (O. Strongylida) genus *Pudica* were encountered only two times from the same species of *Ctenomys* in one locality (from a sample of more than 500 individuals and more than six species of hosts examined). The occurrence of *A. caninum* in *Ctenomys* appears to be a capture, as it only occurred in areas

Table 3

Nematode species	*Classification*
	Sub-Phylum Adenophorea
Trichuris (>3 spp.)	Cecum and large intestine
	Sub-Phylum Secernentea
Ancylostoma caninum (host capture in synanthropic species of rodents?)	Duodenum
Pudica sp. Host capture from Muroid rodents.	Duodenum
Litomosoides (2 spp.)	Peritoneal cavity and mesenteries
Paraspidodera (>6 spp.)	Cecum and large intestine

where dogs, humans, and ctenomyids lived in relatively close proximity (banana fields in lowland Santa Cruz, Bolivia).

Although most of the pocket gophers examined carefully generally harbor from one to several species of nematodes in the small intestine, very few nematodes are found in samples of tuco-tucos. Even though the genus *Ctenomys* contains almost 40 species, comparatively few species of nematodes have been described or reported from them. This lack of parasites in a wide-open group of mammals might be a result of rapid speciation in the mammal group with parasites failing to keep up with the speciation rate of the mammals themselves and actually loosing parasites through time; there is some evidence that the ctenomyids have speciated rapidly in their history (Lessa and Cook, 1998). The lack of a diverse fauna of nematodes in these mammals could also be due to historical accident whereby the ancestor of the ctenomyids had a low diversity of nematode parasites (for whatever reason) thus giving rise to a phylogenetic lineage of mammals lacking a diverse fauna of parasites. But why have they not picked up more parasitic nematodes from other syntopic species of mammals?

There is some evidence of host-switching in nematodes of the genus *Litomosoides* in that two species occur in *Ctenomys opimus* in high-altitude Western Bolivia, but these nematodes have not been reported from any other species of *Ctenomys* from throughout the Neotropics. Superficially this indicates a host-capture event from some other lineage of mammals (Brant and Gardner, 2000). Another example is the fact that nematodes of the genus *Pudica* found in *Ctenomys* have diverse relatives in other species of muroid rodents in South America but none in ctenomyids, leading to the conclusion that most are now found in the tucos because of host-switching events, and not phylogenetic coevolution.

Nematodes of the genus *Paraspidodera* are found only in Hystricognath rodents in the Neotropical Region and these nematodes appear to have had a long historical, coevolutionary association with ctenomyids, showing varying levels of both cospeciation and host-switching. The adenophorean whip worm genus *Trichuris* occurs in many diverse groups of rodents in the neotropics and at the present time, no analyses have been done to examine the levels of coevolution with ctenomyids.

The multifarious nature of nematode diversity in subterranean mammals in the Nearctic and Neotropical Regions requires at the minimum that phylogenetic hypotheses for each group of mammals and their nematodes be developed and then each host-group can be compared with each parasite-group. Much more detailed work needs to be paid to collecting parasites from some of the unknown species of *Ctenomys* throughout the Neotropics. The same can be said about the level of knowledge of parasitic nematodes in the Geomyidae in the northern Neotropics and southern Nearctic regions.

Molecular Diversity in the Nemata

Several molecules have been used to begin to assess phylogenetic and genetic diversity within the Nemata and the number of investigations using molecular methods to try to quantify the diversity of the nematodes is rapidly increasing (Dorris *et al.*, 1999; Blaxter *et al.*, 1998; Al-Banna *et al.*, 1997; Nielsen, 1996). However, because of the extremely large number of species that may exist, examination of levels of molecular or genetic diversity in representatives of the group as a whole can be said as just the beginning even though massive amounts of molecular data on nematodes are now literally pouring into the literature stream. The summary papers by Blaxter *et al.* (1998); Dorris *et al.* (1999); Meldal, *et al.*, 2007; Holterman *et al.* (2006); Van Megan *et al.* (2009); and Bik *et al.* (2010) indicate the utility and power of estimating the molecular–phylogenetic relationships among the Nemata using ribosomal DNA and other molecular sequence data. Initial studies of the molecular diversity within and among several lineages of the Strongylida show that the genetic diversity among these taxa is relatively low (Chilton *et al.*, 1997). From these works, it is clear that molecular phylogenies will provide robust tests of the hypotheses of morphological relationships among the Nemata. As more regions of DNA are used to examine the relationships among the nematodes, we expect additional clarification of both the deep phylogenetic branches (Bik *et al.*, 2010) that are at present relatively obscure in the molecular phylogeny of the Nemata and the more rapidly evolving branch tips that represent extant species with valuable genetic information

Relationships to other Animal Groups

A recent analysis grouped the nematodes, gastrotrichs, priapulids, kinorhynchs, and the loriciferans into a group (superphylum) called the Cycloneuralia (**Figure 26**) based on the circular shape of the brains in these groups (Nielsen *et al.*, 1996). The aforementioned study and at least one other (Zrzavy, 1998) using 18s rDNA sequences showed that the Nemata share a common ancestor with members of the phylum Nematomorpha but there was a shuffling of the other groups out of the "Cycloneuralia" (**Figure 27**).

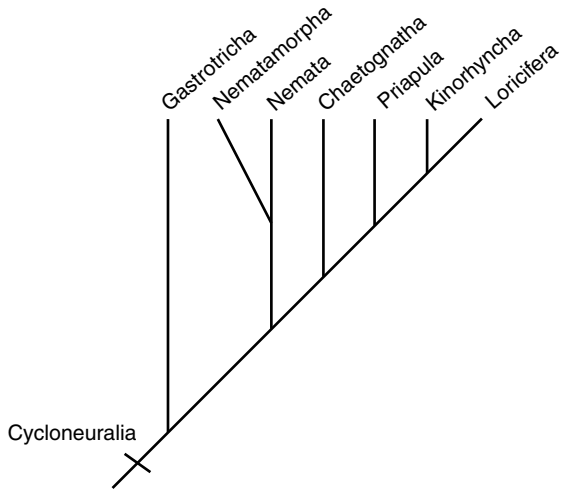


Figure 26 Phylogenetic tree showing the relationships of the Nemata to the rest of the Cycloneuralia.

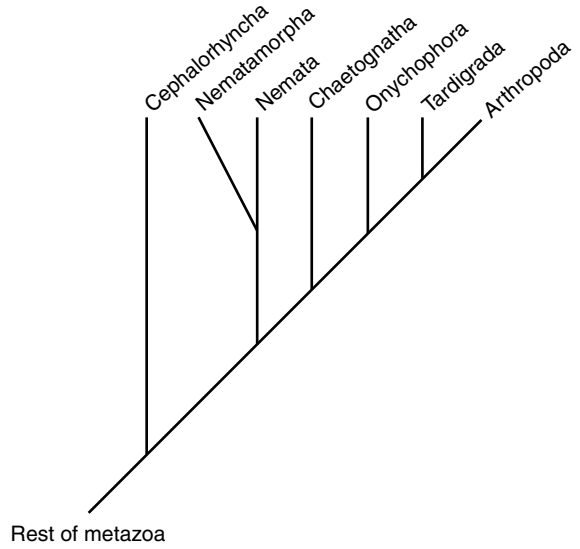


Figure 27 Phylogenetic tree showing the relationship of the Nemata to the rest of the Animalia.

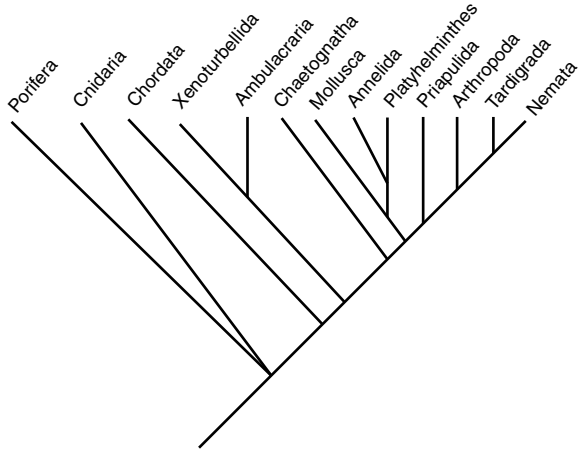


Figure 28 Phylogenetic tree of the Bilateria showing the sister group of the Nemata to be members of the phylum Tardigrada or the moss bears.

More recently, Lartillot and Philippe (2008) presented a molecular phylogeny of the Bliateria where they show the Nemata as sharing a most recent common ancestor with members of the phylum Tardigrada and are close to the phylum Arthropoda in their final tree (**Figure 28**).

As mentioned above, nematodes are soft-bodied and mostly very small organisms, any larger forms that existed in prehistory were probably parasites of vertebrates, however, these left no fossil traces. The only fossil nematodes that are known are insect parasitic or plant parasitic forms that occur very rarely in amber inclusions, but these are also not very old. Because there are no fossil records of nematodes of Cambrian or Precambrian ages, estimates of the age of the Nemata have been only speculation up to the present time and without fossils; it is difficult to calibrate molecular clocks for the nematodes. However, through application of various models of molecular evolution and molecular clock theory, estimates of the time of divergence of the nematodes from the rest of the animal groups appear to be approximately 1177–1179 million years (Wang *et al.*, 1999) (this study showed a basal origin of the nematodes on a phylogenetic tree, in contrast to the relatively derived placement in the analyses shown in **Figures 26–28**).

Most authors consider that nematodes arose from a Marine ancestor. The tri-radiate esophagus and the tubular body indicate a possible primarily sedentary existence with the posterior end attached to the substrate and the anterior end freely encountering the marine environment from all sides, thus the secondarily derived somewhat radial symmetry (Maggenti, personal communication).

Classification and Groups of the Nemata

Above the level of the order, confusion reigns relative to the classification and systematic arrangement of the nematodes. Maggenti (1991) is usually followed in this regard and his analyses followed corroborated historical analyses in recognizing the two main sub-phyla: the Secernentea and the Adenophorea. Recent work shows that these groups are substantiated both in morphological and molecular analyses although competing phylogenetic hypotheses and associated classifications have also relatively recently been proposed (Dorris *et al.*, 1999; Blaxter *et al.*, 1998; Adamson, 1989). Even more recent work by Meldal *et al.* (2007); Holterman *et al.* (2006); Van Megen *et al.* (2009) and Bik *et al.* (2010) have provided some larger scale analyses of the Nemata using sequencing of ribosomal DNA and other conserved genes.

A systematic treatment of the nematodes was recently undertaken by Hodda (2007) where he reviews the systematics of the group and proposes a classification that incorporates molecular, developmental, and morphological advances. He reviews the classification of the Nemata as a phylum (although he uses Nematoda as the name). Here the author would like to point out that Maggenti (1981, 1991a, and 1991b) used Nemata as the name for the phylum as he often pointed out that "Nematoda" is a class level name left over from the old classification when nematodes were considered a class in the phylum Aschelminthes. If the nematodes were considered a class under a different phylum name, then Class Nematoda would be appropriate, as it is, Phylum Nemata is more correct (A.R. Maggenti, personal communication).

Since 1949, at least eight authors have provided classifications for members of the phylum Nemata (Malakhov, 1994; Brooks and McLennan, 1993; Maggenti, 1991). Of these, the classifications of Maggenti (1981; 1991) have proven to be the most useful summary of all nematodes (free-living, and parasitic); however, phylogenetic hypotheses have recently been proposed based on both molecular and morphological characteristics. This does not necessarily mean that because a phylogeny has been proposed that a classification from that phylogenetic tree will result (Maggenti, 1991a and Brooks and McLennan, 1993).

At this point in time, the synthetic classification proposed by Hodda (2007), which takes the arrangement down to the level of the family, appears to be the most easily useable and informative new classification. He presents the groups in a

table format, and as noted above, uses information from many different fields of study of the Nemata to generate the classification. One of his goals was to minimize the changes in taxonomic rank of well established groups to achieve stability and he generally follows the use of superorders (e.g., Maggenti, 1991).

Future Knowledge of Nematodes

The author hopes that this summary treatment provides readers with sufficient knowledge to allow more in-depth research on nematodes. The group is so large, and so ecologically, morphologically, and phylogenetically diverse that to attempt to discuss the diversity of the group in such an abbreviated way is practically futile at best. N. A. Cobb (1914) stated this clearly on the last page of his famous "Nematodes and Their Relationships" (pp. 490).

> *"The foregoing fragmentary sketch may indicate to the student, as well as to the general reader, the vast number of nematodes that exist, the enormous variety of their forms, and the intricate and important relationships they bear to mankind and the rest of creation."*

As more data on distribution and function of nematodes throughout the biosphere are obtained, the importance of this group of worms will surely be realized. We are just beginning to explore the oceans, and we are probably losing more species of nematodes from rainforest clearing than will ever be ultimately found, described, and classified.

The author hopes that mankind will generate more interest in the microscopic world of the Nemata, and that this section will provoke the reader into action and disprove what Van Leeuwenhoek stated about his fellow man on 28 September, 1715 in which he said: "...And over and above all, most men are not curious to know: nay, some even make no bones about saying, What does it matter whether we know this or not?" (Dobel, 1932; pp. 324–325).

Acknowledgments

Special thanks to the staff and students of the Manter Laboratory. Also thanks to Sue, Grant, and Clark Gardner for their support during this work.

See also: Census of Marine Life. Diseases, Conservation and. Diversity and Conservation of Neotropical Mammals. Diversity, Organism Level. Endangered Ecosystems. Endangered Plants. Fishes, Biodiversity of. Global Species Richness. Loss of Biodiversity, Overview. Mammals, Biodiversity of. Mammals, Conservation Efforts for. Neotropical Seasonally Dry Forests. Poverty and Biodiversity. Soil Biota, Soil Systems, and Processes

References

Adamson M (1989) Constraints in the evolution of life histories in zooparasitic nematoda. In: Ko RC (ed.) *Current Concepts in Parasitology*, pp. 221–253. Hong Kong: Hong Kong University Press.

Al-Banna L, Williamson VM, and Gardner SL (1997) Phylogenetic analysis of nematodes of the genus *Pratylenchus* using nuclear 26s rDNA. *Molecular Phylogenetics and Evolution* 7: 94–102.

Al Banna L and Gardner SL (1996) Nematode Diversity of Native Species of *Vitis* in California. *Canadian Journal of Zoology* 74: 971–982.

Anderson RC (1992) *Nematode Parasites of Vertebrates. Their Development and Transmission*, pp. 578. Wallingford, UK: CAB International.

Arujo A, Reinhard KJ, Ferreira LF, *et al.* (2008) Parasites as probes for prehistoric migrations? *Trends in Parasitology* 24: 112–115.

Bik HM, Lambshead JD, Thomas WK, *et al.* (2010) Moving towards a complete molecuar framework of the Nematoda: A focus on the Enoplida and early-branching clades. *BMC Evolutionary Biology* 10: 353. 14 pages.

Blaxter ML, De Ley P, Garey JR, *et al.* (1998) A molecular evolutionary framework for the phylum Nematoda. *Nature* 392: 71–75.

Bongers T and Ferris H (1999) Nematode community structure as a bioindicator in environmental monitoring. *Trends in Ecology and Evolution* 14: 224–228.

Boucher G and Lambshead PJD (1995) Ecological biodiversity of marine nematodes in samples from temperate, tropical, and deep-sea regions. *Conservation Biology* 9: 1594–1604.

Brant SV and Gardner SL (June 2000) Phylogeny of species of the genus *Litomosoides* (Nemata: Onchocercidae) evidence of rampant host-switching. *Journal of Parasitology* 83: 545–554.

Brooks DR and McLennan DA (1993) Parascript. *Parasites and the Language of Evolution*. Washington, DC: Smithsonian Institution Press.

Chan MS (1997) The global burden of intestinal nematode infections – fifty years on. *Parasitology Today* 13: 438–443.

Chilton NB, Gasser RB, and Beveridge I (1997) Phylogenetic relationships of Australian strongyloid nematodes inferred from ribosomal DNA sequence data. *International Journal for Parasitology* 27: 1481–1494.

Chitwood BG and Chitwood MB (1977) *Introduction to Nematology*, pp. 334. Baltimore: University Park Press.

Cobb NA (1914) Nematodes and their relationships. In: *Yearbook of Department of Agriculture for 1914*, pp. 457–490. Washington, DC: Government Printing Office.

Crompton DWT (1999) How much human helminthiasis is there in the world? *Journal of Parasitology* 85: 397–403.

Dobel C (1932) *Antony Van Leeuwenhoek and His "Little Animals,"* pp. 435. NY: Harcort, Brace and Co.

Dorris M, De Ley P, and Blaxter ML (1999) Molecular Analysis of Nematode Diversity and the Evolution of Parasitism. *Parasitology Today* 15: 188–193.

Ferris H, Venette RC, van der Meulen HR, *et al.* (1998) Nitrogen mineralization by bacterial feeding nematodes: Verification and measurement. *Plant and Soil* 203: 159–171.

Gaugler R and Kaya HK (1990) *Entomopathogenic Nematodes in Biological Control*, pp. 365. Boca Raton, FL: CRC Press.

Grassé P-P (1965) *Traité de Zoologie. Anatomie, Systématique, Biologie.* Némathelminthes (Nématodes) Nématodes - Gordiacés - Rotifères - Gastrotriches - Kinorhynques. Tome IV. Fascicule II, III. pp. 1–1497.

Holterman M, van den Wurff A, van den Elsen S, *et al.* (2006) Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution* 23: 1792–1800.

Huber JCh (1906) Demetrios Pepagomenos uber die Wurmer in den Augen der Jagdfalken. *Zool. Annal.* 2: 71–73.

Hugot JP (1999) Primates and their pinworm parasites: The Cameron hypothesis revisited. *Systematic Biology* 48: 523–546.

Lambshead PJD (1993) Recent developments in marine benthic biodiversity research. *Océanis* 19: 5–24.

Lartillot N and Philippe H (2008) Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. *Philosophical Transactions of the Royal Society B* 363: 1463–1472.

Lawton JH, Bigtnell DE, Bloemers GF, *et al.* (1996) Carbon flux and diversity of nematodes and termites in Cameroon forest soils. *Biodiversity and Conservation* 5: 261–273.

Lessa E and Cook JA (1998) The molecular phylogenetics of Tuco-tucos (genus *Ctenomys*, Rodentia: Octodontidae) suggests an early burst of speciation. *Molecular Phylogenetics and Evolution* 9: 88–99.

Maggenti AR (1981) *General Nematology*, pp. 372. New York: Springer-Verlag.

Maggenti AR (1991a) Nemata: Higher Classification. In: *Manual of Agricultural Nematology*, pp. 147–187. New York: Marcel Dekker, Inc.

Maggenti AR (1991b) General Nematode Morphology. In: *Manual of Agricultural Nematology*, pp. 3–46. New York: Marcel Dekker, Inc.

Malakhov VV (1994) Nematodes. *Structure Classification and Phylogeny*, pp. 286. Washington, DC: Smithsonian Intitution Press.

Meldal BHM, Debenham NJ, De Ley P, *et al.* (2007) An improved molecular phylogeny of the Nematoda with special emphasis on marine taxa. *Molecular Phylogenetics and Evolution* 42: 622–636.

Morand S, Legendre P, Gardner SL, *et al.* (1996) Body size evolution of Oxyurid (Nematoda) parasites – the role of hosts. *Oecologia* 107: 274–282.

Musser GG and Carleton MD (1993) Family Muridae. In: Wilson DE and Reeder DM (eds.) *Mammal Species of the World: A Taxonomic and Geographic Reference*, pp. 501–755. Washington, DC: Smithsonian Institution Press.

Nielsen C, Scharff N, and Eibye JD (1996) Cladistic analyses of the animal kingdom. *Biological Journal of the Linnean Society* 57: 385–410.

Powers L, Mengchi H, Freckman-Wall D, *et al.* (1998) Distribution, community structure, and microhabitats of soil invertebrates along an elevational gradient in Taylor Valley, Antarctica. *Arctic and Alpine Research* 30: 133–141.

Rausch RL (1983) Chapter 5. The biology of avian parasites: Helminths. In: Farner DS, King JR, and Parkes KC (eds.) *Avian Biology,* vol. VII. NY: Academic Press.

Wang D, Kumar S, and Hedges B (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proceedings of the Royal Society of London, Series B* 266: 163–171.

Van Megen H, Van den Elsen S, Holtermann M, *et al.* (2009) A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* 11: 927–950.

Zrzavy J, Mihulka S, Kepka P, *et al.* (1998) Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* 14: 249–285.

Worms, Platyhelminthes

Janine N Caira, University of Connecticut, Storrs, CT, USA
D Timothy J Littlewood, The Natural History Museum, London, UK

© 2013 Elsevier Inc. All rights reserved.

Glossary

Acetabulum Ventral sucker used for attachment and locomotion in digeneans.
Acoelomate Lacking a body cavity; see parenchyma.
Cercomer Posterior parenchymous or muscular extension of body; homology among groups is controversial; some view both the hook-bearing regions of monogeneans and some larval cestodes to be modifications of the cercomer.
Cirrus Male intromittent copulatory organ consisting of distal, invaginable portion of ejaculatory duct that is evaginated (by turning inside out) through the genital pore during copulation; may or may not be armed with spines, microtriches, etc.; see penis.
Cortex The outermost region of parenchyma in taxa in which two distinct regions are present; inner boundary often marked by conspicuous circular muscle fibers; see medulla.
Duogland adhesive system Combination of a cement, a viscid gland producing adhesive material, and a releasing gland; usually associated with cells of ventral epidermis; provides temporary adhesion to the substrate for many interstitial flatworm species.
Gonochoristic Individual organisms exhibit only one of two sexes.
Haptor The posterior attachment organ of monogeneans.
Hermaphroditic Individual organisms exhibit both sexes.
Medulla The innermost region of parenchyma in taxa in which two distinct regions are present; outer boundary often marked by conspicuous circular muscle fibers; see cortex.
Microthrix A specialized surface extension of the neodermis of cestodes, characterized by an electron-dense cap composed of numerous cap tubules separated from a core of distal cytoplasm by a baseplate (plural microtriches); aid in attachment and possibly absorption of nutrients.
Monozoic Possessing only a single set of reproductive organs; see polyzoic.
Neodermis A syncytial, nonciliated epidermis with cell bodies and nuclei (often referred to as cytons or pericaryon) laying below other body wall elements; unique to the Neodermata; completely replaces epidermis, usually at metamorphosis when larva encounters the first host.
Parenchyma The connective tissue between cells of ectoderm and endoderm layers that fills spaces among organ systems; as platyhelminths are acoelomates, the space among internal organs is completely filled with parenchyma.
Penis Male intromittent copulatory organ consisting of a muscular, often papilliform, retractable structure that is protruded (without turning inside out) through the genital pore during copulation; may or may not be armed with spines, etc.; see cirrus.
Polyzoic Possessing multiple serial sets of reproductive organs; condition found in most cestode orders.
Proglottid Compartmentalized regions of cestode body; each proglottid contains one or more sets of reproductive organs; not homologous with segments of coelomates.
Rhabdite A rod-shaped, proteinaceous secretory product found in the epidermis of most free-living platyhelminths; known functions include mucus production for ciliary gliding, cocoon formation, prey capture, and predator repulsion.
Scolex Attachment organ in adult eucestodes; often bearing suckers, loculi, or hooks.
Statocyst Static sense organ in which the movement of a statolith is detected, thereby indicating an animal's orientation.
Strobila The portion of the body of cestodes exhibiting serial repetition of proglottids.

The Phylum Platyelminthes is presently considered to contain approximately 30,000 species, although that number is likely to be a gross underestimate. These animals are usually bilaterally symmetrical, many are dorsoventrally flattened; they generally lack a body cavity and most are hermaphroditic. None of these features, however, is unique to this phylum, and thus, none can be counted on for the definitive recognition of membership in the group. That is not to say the phylum does not represent a natural assemblage of taxa. Molecular analyses support its monophyly and the platyhelminths appear to be unique in their possession of protonephridial cilia that lack accessory centrioles. Unfortunately, the latter feature can be observed only with transmission electron microscopy and thus it is not especially convenient for the casual identification of platyhelminths. In many cases, it is more useful to turn to the rather distinctive features that characterize the individual platyhelminth subgroups.

Because they lack a protective outer body covering, water figures prominently in the lives of the platyhelminths, and indeed, the majority of species are found in moist or aquatic environments. Marine, freshwater, and terrestrial environments have all been colonized, the latter to a lesser extent than the former two environments. The phylum includes species spanning a wide range of adult lifestyles, from those that are

free-living interstitial dwellers to those that are symbiotic with invertebrates or obligate (internal or external) parasites of vertebrates. Consequently, the group includes species ranging enormously in size (200 μm–80 m), shape, form, and occasionally color. Species that are obligate parasites of vertebrates belong to four distinctive, potentially monophyletic groups: Monogenea, Aspidogastrea, Digenea, and Cestoda. A number of these are of significant medical or veterinary importance. There exists little evidence to suggest that the remaining platyhelminths, many of which are free-living members of the meiofauna and some of which exist as either commensals or parasites of invertebrates, form a single cohesive group, although collectively they are often referred to as "turbellarians." Present knowledge suggests that the parasitic groups are more speciose than the nonparasitic groups. However, there is no question that numerous free-living and parasitic taxa await discovery and formal description.

General Features

A glance at the defining features of each of the major platyhelminth groups reveals that the characteristics that vary most among, and even within, these groups are primarily those associated with the digestive and reproductive systems. However, in the parasitic taxa, the features associated with attachment also vary markedly. When present, the digestive system generally lacks an anus. Thus, there is no unidirectional flow of food materials through this system; in many species food enters and solid wastes exit through the oral opening. Most platyhelminths are hermaphroditic, with each individual possessing both male and female reproductive organs. Several major features of the female reproductive system are of particular importance. The female gonads of platyhelminths are either homocellular, consisting of only a single type of cell that produces both ova and vitelline material (i.e., yolk and shell precursors), or heterocellular, consisting of two cell types, one that produces ova and one that produces vitelline materials. The former configuration results in the production of endolecithal eggs thus vitelline material is incorporated directly into the cytoplasm of the ova; the latter configuration results in the production of ectolecithal eggs, as vitelline cells are packaged along with, but independent of, the zygote in the egg. Heterocellular platyhelminths usually, but not always, conduct these functions in separate organs, often possessing one, or occasionally more, ovaries that produce ova, and a vitellarium consisting of vitelline cells that produce nutritive yolk and shell precursors.

In addition, most platyhelminths possess a formalized excretory system consisting of two (rarely one or three) sets of protonephridia, which may be connected at several points throughout the body or may join at a common excretory vesicle. The nervous system and sense organs are present to a greater or lesser extent among different taxa. Specialized organs functioning in respiration and circulation are entirely lacking in many groups of platyhelminths. Thus, respiration generally occurs by means of diffusion of gases directly through the surfaces of the body, a process that is greatly facilitated by the relatively flattened nature of the platyhelminth body. The "circulation" of materials throughout the various parts of the platyhelminth body is also accomplished by diffusion. However, a number of digeneans are known to possess a "lymphatic system" with contractile vessels and cellular inclusions that may have a circulatory function. The details of most organ systems are difficult to observe without the use of specialized staining and microscopic techniques.

One final, rather unusual feature of the platyhelminths is that the differentiated nonreproductive (somatic) cells of the body are unable to divide. Thus, mitosis does not occur in the differentiated somatic cells of platyhelminths; rather, these cells are replaced as necessary by undifferentiated stem cells located below the outer layers of the body, which undergo mitosis. This phenomenon helps to explain the remarkable regenerative abilities of certain platyhelminth groups.

Major Groups

Diversity in the platyhelminths is most effectively represented at the taxonomic rank of order (or superfamily in the digeneans). Within the phylum, this rank is relatively informative and is much more numerically manageable than lower ranks such as family or genus (i.e., there are 71 orders (or superfamilies in the case of the digeneans) versus 391 families and 3847 genera). However, it is important to recognize that traditionally, not only have taxonomists working on each of the three major groups exhibiting obligate parasitism (Digenea, Monogenea, and Cestoda) generally worked independently from one another, but they have also worked independently from the taxonomists studying the platyhelminths exhibiting more free-living life history strategies. Thus, the concepts of higher ranks, such as order, are not necessarily comparable among the major groups. For example, the 716 genera of monogeneans are currently distributed among ten orders, whereas the 1592 genera of digeneans are distributed among only three orders. Nonetheless, diagnoses are presented for orders in all groups except the digeneans, which are treated at the level of superfamily. The numbers of families and genera for each of the orders (or superfamilies in the case of the digeneans) are given in **Table 1**. This treatment reflects recent taxonomic revisions and discoveries, and thus, in addition to revised estimates of families and genera, includes two turbellarian orders, four cestode orders and eight digenean superfamilies not included in the first edition of this article. Given the lack of a recent comprehensive treatment, the numbers presented below for the monogeneans were developed in consultation with I. Whittington and D. Kritsky. The Acoela and Nemertodermatida have been removed from consideration because they are no longer considered to belong to the Platyhelminthes, rather they appear to be basal bilaterians, or possibly even early divergent deuterostomes.

For more detailed information on specific groups, readers are referred to the works of Cannon (1986), Rieger (1998), and Tyler *et al.* (2006–2010) for the non-neodermatan (i.e., free-living and invertebrate associated) groups; Yamaguti (1971); Schell (1985), and the CABI keys (Gibson *et al.*, 2002; Jones *et al.*, 2005; Bray *et al.*, 2008) for the digeneans;

Table 1 Platyhelminth diversity

Figure number	*Order or superfamily, etc.*	*Number of families*	*Number of genera*
	Platyhelminthes	**391**	**3847**
	"**Turbellaria**"	**107**	**776**
2	Catenulida	5	11
3	Macrostomida	3	27
4	Haplopharyngida	1	1
5	Polycladida	25	108
6	"Lecithoepitheliata"	2	4
7	Bothrioplanida	1	1
8	Proseriata	12	62
9	Kalyptorhynchia	16	151
10	"Dalyellioida"	7	67
11	Temnocephalida	4	23
12	"Typhloplanoida"	7	129
13	Prolecithophora	8	31
14	Tricladida	12	154
15	Fecampiida	4	7
16	**Subphylum Neodermata**	**284**	**3071**
	Class Monogenea	**63**	**716**
	Subclass Monopisthocotylea	**25**	**390**
17	Capsalidea	1	46
18	Dactylogyridea	11	244
19	Gyrodactylidea	8	51
20	Lagarocotylidea	1	1
21	Monocotylidea	3	47
22	Montchadskyellidea	1	1
	Subclass Polyopisthocotylea	**38**	**326**
23	Chimaericolidea	1	2
24	Diclybothriidea	2	15
25	Mazocraeidea	33	279
26	Polystomatidea	2	30
	Class Trematoda	**152**	**1604**
27	**Subclass Aspidogastrea**	**4**	**12**
	Subclass Digenea	**148**	**1592**
28	Allocreadioidea	4	107
29	Azygioidea	1	4
30	Bivesiculoidea	1	4
31	Brachylaimoidea	8	22
32	Bucephaloidea	1	24
33	Clinostomoidea	2	9
34	Cyclocoeloidea	3	24
35	Diplostomoidea	6	87
36	Echinostomatoidea	8	82
37	Gorgoderoidea	12	95
38	Gymnophaloidea	5	40
39	Haploporoidea	2	31
40	Haplosplanchoidea	1	9
41	Hemiuroidea	12	123
42	Heronimoidea	1	1
43	Lepocreadioidea	10	171
44	Microphalloidea	19	244
45	Microscaphidioidea	2	24
46	Monorchioidea	2	47
47	Opisthorchioidea	3	128
48	Paramphistomoidea	12	107
49	Plagiorchioidea	23	105
50	Pronocephaloidea	6	47
51	Schistosomatoidea	3	53
52	Transversotrematoidea	1	4

(*Continued*)

Table 1 Continued

Figure number	*Order or superfamily, etc.*	*Number of families*	*Number of genera*
	Class Cestoda	**69**	**751**
53	Amphilinidea	2	6
54	Gyrocotylidea	1	1
	Subclass Eucestoda	**66**	**744**
55	Bothriocephalidea	4	48
56	Caryophyllidea	4	41
57	Cathetocephalidea	1	2
58	Cyclophyllidea	18	400
59	Diphyllidea	2	2
60	Diphyllobothriidea	3	16
61	Haplobothriidea	1	1
62	Lecanicephalidea	4	17
63	Litobothriidea	1	1
64	Nippotaeniidea	1	2
65	Proteocephalidea	2	49
66	Rhinebothridea	2	13
67	Spathebothriidea	2	5
68	Tetrabothriidea	1	6
69	"Tetraphyllidea"	4	65
70	Trypanorhyncha	16	76

Yamaguti (1963); Schell (1985) and Boeger and Kritsky (2001) for the monogeneans; and Schmidt (1986) and Khalil *et al*. (1994) for the cestodes. However, the treatments of these groups by Hyman (1951) and the five sections covering the platyhelminths in the Traité de Zoologie (Grassé, 1961) are classics and remain among the more comprehensive sources of detailed information on the platyhelminths. The many works of Klaus Rohde (e.g., 1997), Ulrich Ehlers (e.g., 1985), Peter Ax (e.g., 1996), and Tim Littlewood (e.g., 2006), among others, have done much to further our understanding of the morphology and phylogenetic relationships among these groups.

Discussions of diversity are most effectively organized around hypotheses of the phylogenetic relationships among the groups involved. Although many aspects of the phylogenetic relationships among the members of this phylum remain unresolved, the tree in **Figure 1** illustrates some of the aspects of these relationships for which supporting morphological and molecular data are starting to accumulate. The presentation of diversity that follows is generally organized around the topology of this tree. The non-neodermatan orders are treated in the sequence in which they appear from left to right on this tree; orders or superfamilies of each of the three major neodermatan groups are treated in alphabetical order. A more detailed treatment of the phylogenetic relationships among the various groups is provided in the section on Phylogenetic Relationships.

The authors have presented an illustration of a representative of each platyhelminth order (or superfamily for the digeneans). A consistent stippling pattern has been used throughout these figures for each different organ system so that the conditions of the major organs are readily visible and comparable among figures and among major groups. The organs represented by each stippling pattern are labeled in Figure 9 (**Plate 1**) and Figure 28 (**Plate 3**).

"Turbellaria"

Catenulida

See Figure 2 (**Plate 1**). Some with tail; lamellated rhabdites absent, unlamellated rhabdoids present; duogland adhesive system absent; epidermal ciliation sparse, with cilia able to reverse beat; some with V-shaped ciliated oral furrow; brain lobed or not; statocyst often present, usually with one but occasionally with up to six statoliths and several parietal cells; mouth ventral, opening into simple pharynx; intestine in form of simple elongate sac; some lacking mouth and obtaining nutrients from endosymbionts in trophosome; protonephridial system unpaired; with single hermaphroditic gonad or separate male and female gonads; sperm aflagellated; penis present or absent, with common genital pore; vitellarium absent; eggs endolecithal; most undergo asexual multiplication *via* paratomy (subdivision of the body into multiple parts, each of which develops a head) or parthenogenesis (development of an embryo without fertilization).

Habitat: All species are free-living, primarily in freshwater, but some in marine and terrestrial environments.

Macrostomida

See Figure 3 (**Plate 1**). Frontal glands sometimes present; some with spatulate posterior region bearing adhesive papillae; lamellated rhabdites present; duogland adhesive system present; brain not encapsulated; statocyst absent; simple pharynx present, sometimes tubular; intestine in form of simple elongate sac, sometimes with pre-oral blind sacs; with protonephridial system paired; testes compact, single or paired; sperm aflagellated; seminal vesicle present or absent; prostatic vesicle present or absent; penis present or absent; penis stylet present or absent, female pore anterior to male

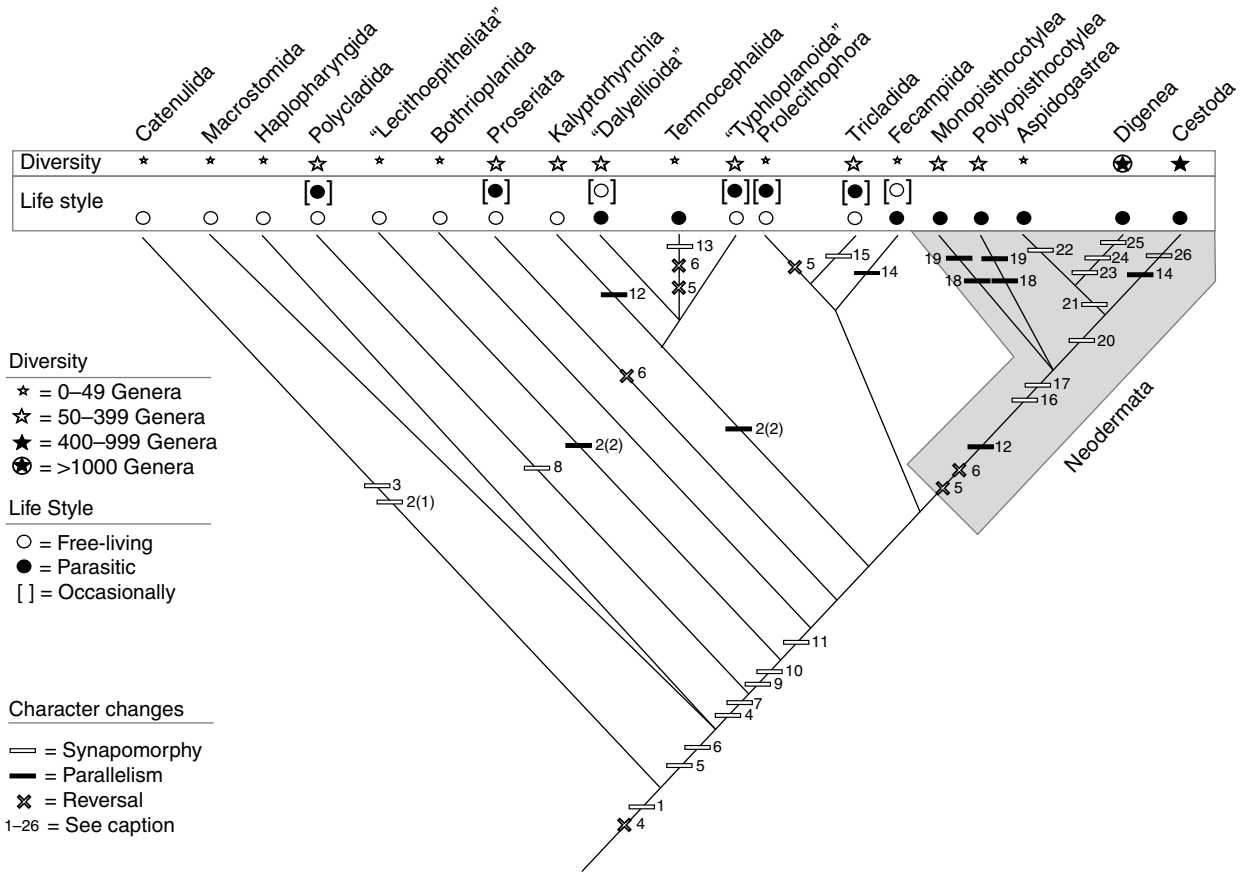
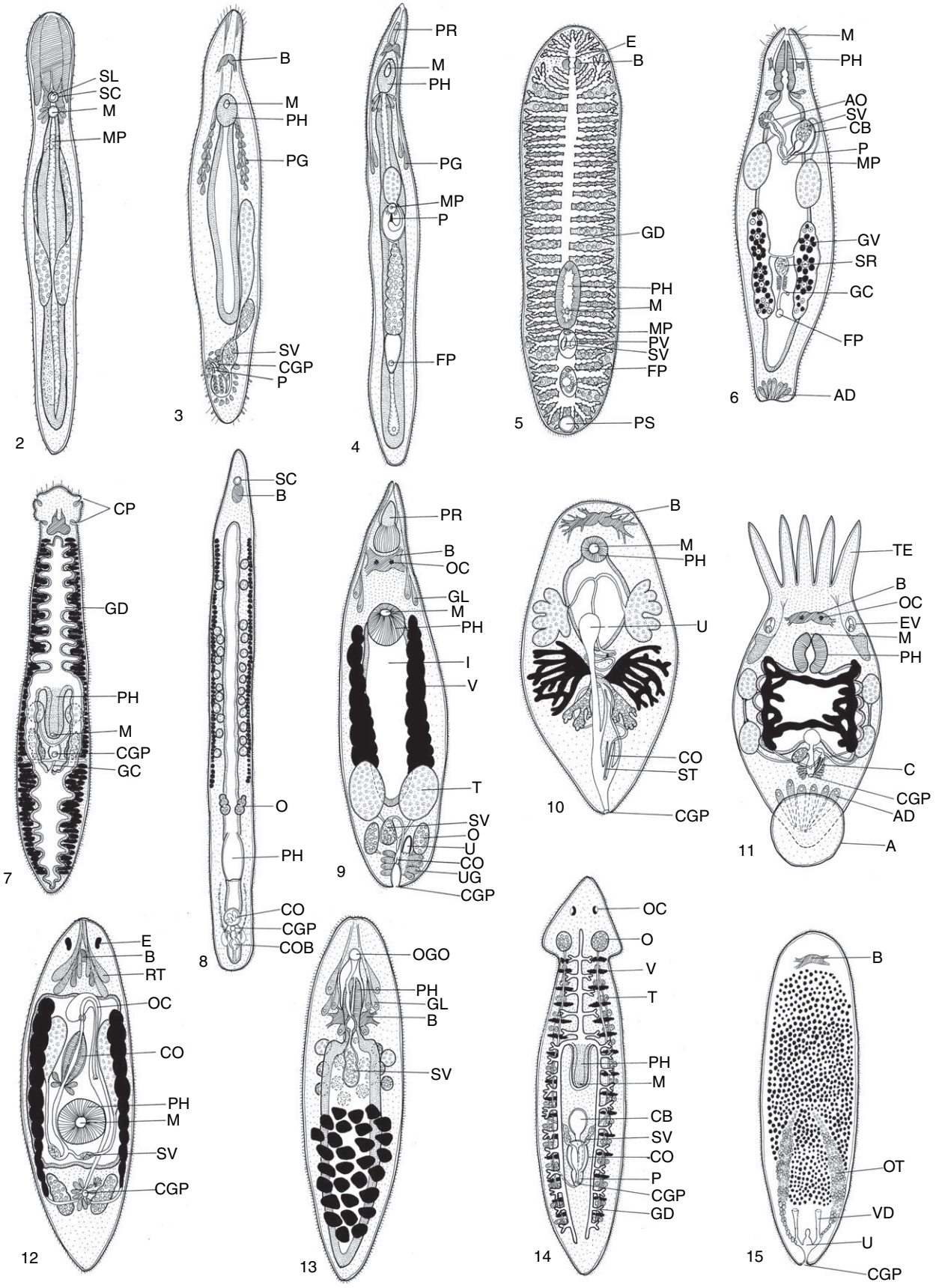


Figure 1 Hypothesized phylogenetic relationships among major platyhelminth groups. State "0" should be assumed prior to point at which character is marked on tree; state "1" should be assumed if character is marked on tree but no state is given. Solid bars indicate homoplasious characters; x indicates character reversals. Topology is synthesized from various recent molecular systematic estimates of phylogeny; where there exist no morphological apomorphies there is often consistent support from molecular data.

1. Accessory centriole in protonephridial flame bulb cilia: 0 = present; 1 = absent
2. Protonephridial type (sensu Rhode, 1991): 0 = Type 2; 1 = Type 1; 2 = Type 3
3. Protonephridial system: 0 = paired; 1 = unpaired
4. Sperm flagellation: 0 = no flagellum; 1 = biflagellated
5. Rhabdites: 0 = absent; 1 = present (CHECK distribution!)
6. Duogland adhesive system: 0 = absent; 1 = present
7. Arrangement of microtubules in sperm flagella when present: 0 = 9 + 2; 1 = 9 + '1'
8. Intestinal branching: 0 = absent; 1 = extensive
9. Configuration of female gonad: 0 = homocellular; 1 = heterocellular
10. Eggs: 0 = endolecithal; 1 = ectolecithal
11. Relationship between ovary and vitelline tissues: 0 = in same organ; 1 = in separate organs
12. Flagellum of sperm cell: 0 = not incorporated into cytoplasm of sperm cell; 1 = incorporated into cytoplasm of sperm cell
13. 2–12 anterior tentacles: 0 = absent; 1 = present
14. Digestive system: 0 = present; 1 = absent
15. Extra, embryonic intestine: 0 = absent; 1 = present
16. Form of outer body layer: 0 = cellular epidermis; 1 = syncytial neodermis
17. Epidermal cells: 0 = not shed by larva; 1 = shed by larva
18. Haptor: 0 = absent; 1 = present
19. 4 rhabdomeric ocelli at some time in their lives: 0 = absent; 1 = present
20. Complex life cycle: 0 = absent; 1 = present
21. Epidermal cells of larva: 0 = not separated from one another by neodermis; 1 = separated from one another by neodermis
22. Oviduct construction: 0 = not divided into chambers; 1 = divided into chambers
23. Miracidium: 0 = absent; 1 = present
24. Sporocyst/redia: 0 = absent; 1 = present
25. Cercaria: 0 = absent; 1 = present
26. Microtriches: 0 = absent; 1 = present


2
SL
SC
M
MP
3
B
M
PH
PG
SV
CGP
P
4
PR
M
PH
PG
MP
P
FP
5
E
B
GD
PH
M
MP
PV
SV
FP
PS
6
M
PH
AO
SV
CB
P
MP
GV
SR
GC
FP
AD
7
CP
GD
PH
M
CGP
GC
8
SC
B
O
PH
CO
CGP
COB
9
PR
B
OC
GL
M
PH
I
V
T
SV
O
U
CO
UG
CGP
10
B
M
PH
U
CO
ST
CGP
11
TE
B
OC
EV
M
PH
C
CGP
AD
A
12
E
B
RT
OC
CO
PH
M
SV
CGP
13
OGO
PH
GL
B
SV
14
OC
O
V
T
PH
M
CB
SV
CO
P
CGP
GD
15
B
OT
VD
U
CGP

pore; ovary single or paired; vitellarium absent; eggs endolecithal; genital pores separate or combined; ability to asexually produce zooids present or absent.

Habitat: All species are free-living in marine, brackish, or freshwater environments.

Haplopharyngida

See Figure 4 (**Plate 1**). With protrusible anterior proboscis not connected to digestive system; lamellated rhabdites present; duogland adhesive system present; brain enclosed in connective tissue capsule (encapsulated), bilobed; statocyst absent; mouth opens into simple muscular pharynx; intestine in form of simple elongate sac; protonephridial system paired; single testis present; sperm aflagellated; penis present, with stylet and circle of hard, straight needles; single ovary; vitellarium absent; eggs endolecithal; with single posterior/ventral pore.

Habitat: Both species free-living in marine environments.

Polycladida

See Figure 5 (**Plate 1**). Body oval or elongate, or thick and broad with folded margins; with (Cotylea) or without (Acotylea) pseudosucker posterior to female genital pore; anterior or nuchal (neck) tentacles present or absent; ventral prostatoid-like organs present or absent; dorsal surface smooth or papillate; body margins of some with hard spines; may be multicolored with complex patterns; lamellated rhabdites present; duogland adhesive system present; statocyst absent; numerous ocelli usually present, marginal or scattered anteriorly, some arranged in one or two pairs of clusters (cerebral and tentacular), some ventral; pharynx variable in position, ruffled and vertically oriented or sometimes consisting of anteriorly directed tube, rarely consisting of multiple tubes arranged ventrally; intestine usually with numerous radiating branches, branches sometimes anastomosing, sometimes with anal pores; protonephridial system paired; testes numerous, scattered between branches of gut; sperm biflagellated, flagella not incorporated into body of sperm, $9+'1'$ microtubule arrangement; true seminal vesicle present or absent; intromittent organ absent or in form of penis or cirrus; penis stylet present or absent; cirrus armed or not; prostatic vesicle present or absent; bursa copulatrix present or absent; some with multiple copulatory complexes; ovaries numerous, dorsal to branches of gut; Lang's vesicle (possibly for digestion of excess sperm) opening into vagina present or absent; uterus present or absent (in form of expanded oviducts in some); vitellarium absent: eggs endolecithal; male and female gonopores usually separate (male anterior to female) but sometimes combined; some with hypodermic impregnation.

Habitat: Predominantly marine, a few in freshwater or brackish water; mostly free-living, some interstitial, some symbiotic with other invertebrates, for example, wrapped around the abdomen of hermit crabs, associated with gorgonian corals in mantle cavities of gastropod or lamellibranch molluscs, in genital bursa of ophiuroid or echinoid echinoderms; some live in association with their food, for example, in association with tunicate colonies feeding on zooids, or in association with oyster colonies feeding on oyster tissue; some seek shelter in empty shells.

''Lecithoepitheliata''

See Figure 6 (**Plate 1**). Some with weak posterior adhesive zone; some with pigmentation; lamellated rhabdites possibly present; duogland adhesive system present; statocyst absent; mouth anterior; pharynx variably bulbous, anterior; intestine in form of simple elongate sac; protonephridial system paired; testes compact or follicular, one or many; sperm biflagellated, flagella not incorporated into body of sperm, $9+'1'$ microtubule arrangement; seminal vesicle present or absent; penis present, usually with stylet; muscular copulatory bulb present or absent; accessory male organ present or absent, if present opening into male atrium, prostatic vesicle present or absent; male pore opening to outside or into pharyngeal cavity; female organ single or tetrapartite, in form of heterocellular germovitellarium (vitelline-producing cells surrounding ovum-producing cells); seminal receptacle present or absent; genitointestinal canal sometimes present; female pore ventral; eggs ectolecithal (yolk not produced in oocytes). The group is likely not monophyletic.

Habitat: All species are free-living in freshwater, marine, or terrestrial environments.

Bothrioplanida

See Figure 7 (**Plate 1**). With posterior papillae (i.e., *Haftpapillen*); body ciliated, with two pairs of ciliated pits; statocyst absent; pharynx plicate, posteriorly directed; gut splitting laterally around pharynx and fusing again posterior to pharynx, with lateral diverticula; with one pair of protonephridium;

Plate 1 ''Turbellaria'' (Figures 2–15) 2. Catenulida: adult of *Retronects* sp. 3. Macrostomida: adult of *Bradynectes* sp.; 4. Haplopharyngida: adult of *Haplopharynx* sp. 5. Polycladida: adult of *Cestoplana* sp. (Beauchamp in Grassé, 1961); 6. ''Lecithoepitheliata'': adult of *Gnosenesima* sp. 7. Bothrioplanida: adult of *Bothrioplanus* sp. (Beauchamp in Grassé, 1961); 8. Proseriata: adult of *Parainvenusta englarorum* (Curini-Galletti, *et al*., 2010); 9. Kalyptorhynchia: adult of *Cystiplana* sp. 10. ''Dalyellioida'': adult of *Syndesmis* sp. 11. Temnocephalida: adult of *Temnocephala* sp. (Baer in Grassé, 1961); 12. ''Typhloplanida'': adult of *Promesostoma* sp. (Ehlers, 1974); 13. Prolecithophora: adult of *Prolecithoplana* sp. (from several literature sources); 14. Tricladida: adult of *Dugesia* sp. 15. Fecampiida: adult of *Fecampia balanicola*. Abbreviations: A, adhesive disc; AD, adhesive glands; AO, accessory male organ; B, brain; C, cirrus; CB, copulatory bulb; CGP, common genital pore; CO, copulatory organ; COB, copulatory bursa; CP, ciliated pits; E, eyespot; EV, excretory vesicle; FP, female genital pore; GC, genitointestinal canal; GD, gut diverticula; GL, gland cells; GV, germinovitellarium; I, intestine; M, mouth; MP, male genital pore; O, ovary; OC, ocellus; OGO, orogenital opening; OT, ovotestis; P, penis; PG, proboscis glands; PH, pharynx; PR, proboscis; PS, pseudosucker; PV, prostatic vesicle; RT, rhabdoid tracts; SC, statocyst; SL, statolith; SR, seminal receptacle; ST, stylet; SV, seminal vesicle; T, testis; TE, tentacle; U, uterus; UG, uterine gland; V, vitellarium; VD, vitelline duct.

testes paired, compact, possibly producing only degenerate sperm; ovaries small, paired, at distal end of vitelline duct; vitellarium follicular; with single gonopore; genitointestinal canal extending from genital atrium to posterior arm of gut.

Habitat: Single known species free-living usually in freshwater environments, occasionally terrestrial.

Proseriata

See Figure 8 (**Plate 1**). Epithelial nuclei insunk or not; body of some with feeble or prominent tactile bristles; body uniformly ciliated (sometimes with exception of caudal tip) or cilia restricted to ventral surfaces and head; head with or without ciliated pits; lamellated rhabdites absent; duogland adhesive system present; brain encapsulated or not; statocyst present or absent; pigmented ocelli present or absent; pharynx plicate, tubular; intestine linear or tripartite, dorsal to pharynx; protonephridial system paired or tripled, occasionally single; testes single, paired or multiple, usually follicular, rarely compact; sperm biflagellated, flagella not incorporated into body of sperm, $9+'1'$ microtubule arrangement; usually with single male copulatory; occasionally with accessory (prostatoid) organ; bursa present or absent; with one pair of compact ovaries, ovary subdivided into smaller follicles consisting of oocytes surrounded by accessory cells in some; vitellarium follicular, follicles in paired lateral rows; vagina externa present or absent; genitointestinal canal usually present; eggs ectolecithal; genital pores common or separate.

Habitat: Free-living, predominantly in marine environments; a few occur in freshwater; occasionally ectocommensal on marine crustaceans.

Kalyptorhynchia

See Figure 9 (**Plate 1**). Body usually ciliated throughout; rarely with subepidermal spicules of crystalline calcium carbonate; rarely with anterior girdle of highly vacuolated epidermal cells; anterior sheathed proboscis usually present, divided or not, with or without hooks, gland cells, or specialized apex; lamellated rhabdites present in some; duogland adhesive system present; statocyst absent; ocelli present or absent; pharynx variable in position (anterior or posterior), bulbous, usually oriented ventrally, with or without sphincter or hard knobs, not associated with proboscis; gut usually in form of simple elongate sac; protonephridial system paired; testes single or paired, or in form of median row of follicles; sperm biflagellated, flagella not incorporated into body of sperm, $9+'1'$ microtubule arrangement; internal seminal vesicle in copulatory bulb or not; copulatory organ with stylet or not; or three separately opening prostatic organs with their own spiny cirri, or one cirrus, or copulatory organ with no penial or cirrus structures; prostate vesicle incorporated in testis or not; prostate with spiny structures or not; ovary and vitellarium separated or not (ovo-vitellarium), paired or single; uterus generally present, with or without accessory piece; seminal receptacle present or absent, paired or not; bursa present or absent; eggs ectolecithal; usually one but sometimes two gonopores.

Habitat: All species free-living, primarily in marine or brackish water, several in freshwater, environments.

''Dalyellioida''

See Figure 10 (**Plate 1**). Lamellated rhabdites present in some; duogland adhesive system absent; cilia completely covering body, or just ventral; statocyst usually lacking; paired ocelli present or absent; mouth, pharynx, and intestine present or occasionally absent; pharynx bulbous, oriented anteriorly if present; intestine in form of simple elongate sac if present; paired protonephridial system present or occasionally absent; testes paired, compact or lobed; sperm biflagellated, flagella not incorporated into body of sperm, $9+'1'$ microtubule arrangement; with or without male copulatory organ; if present male organ with or without stylet; ovary and vitellarium separate or occasionally combined; ovary single or paired, lobed; vitellarium single or paired, branched or unbranched; seminal receptacle present or absent; uterus present; copulatory bursa present or absent; single genital pore present or absent; at least one taxon is gonochoristic and sexually dimorphic (*Kronborgia*). The group is likely not monophyletic.

Habitat: Some species free-living in marine, freshwater, or terrestrial environments; many species commensal or parasitic: for example, in connective tissue of tube feet of asteroid echinoderms, in hemocoel of crabs, shrimp, amphipods, and isopods, in the mesenchyme of myzostomid annelids, in kidneys and gonoducts or mantle cavity and gut of gastropod molluscs, in stomach of lamellibranch molluscs, in mesenchyme of the turbellarian *Plagiostomum*, in either the coelom or the digestive tract of holothurioid, echinoid, or crinoid echinoderms; some also in sipunculans.

Temnocephalida

See Figure 11 (**Plate 1**). Usually with two to 12 anterior tentacles, tentacles usually adhesive; with posterior adhesive region, usually circular but occasionally crescentic or divided into two or more smaller regions, often pedunculate; some with lamellated rhabdites, if present restricted to anterior region of body; duogland adhesive system absent; locomotory cilia usually absent; outer body layer divided into a series of syncytial epidermal plates; statocyst absent; paired ocelli usually present; mouth terminal or ventral; pharynx tubular or bulbous, usually oriented anteriorly, sometimes reduced; intestine in form of simple sac or with weak lateral diverticula; paired protonephridial system present, usually with two lateral excretory vesicles; with one to 10 pairs of testes; sperm biflagellated, flagella not incorporated into body of sperm, $9+'1'$ microtubule arrangement; male organ usually armed, usually with eversible cirrus mounted on tubular or conical shaft; ovary single; vitellarium lateral or scattered in form of large follicles over intestine; eggs ectolecithal.

Habitat: Entirely freshwater; ectocommensal, primarily on external body surfaces, gills, or lining of branchial chambers of decapod crustaceans, occasionally found on amphipods, molluscs, turtles, and hemipterans.

''Typhloplanoida''

See Figure 12 (**Plate 1**). Very small; anterior usually modified with prominent rhabdoid tracts or rarely with an impermanent sheathed proboscis, or with a permanent terminal invagination with proboscis glands; dermal and adenal rhabdites usually present; duogland adhesive system present; cilia uniform throughout all surfaces of body or restricted to ventral surface; brain undivided; statocyst absent; paired ocelli present or absent; tubular spinous, buccal region occasionally present; mouth usually midventral or posterior, occasionally anterior; pharynx usually ventrally oriented, plicate or bulbous; intestine in form of simple elongate sac; paired protonephridial system present; testes single or paired; sperm biflagellated, flagella not incorporated into body of sperm, 9 + '1' microtubule arrangement; single or paired seminal vesicles sometimes present; male system often with funnel-shaped stylet; sheathed, unarmed cirrus present; penis papilla sometimes present; ovary and vitellarium usually separate; ovary single, paired, or unpaired; vitellarium paired or a column of serially arranged follicles; usually with bursa copulatrix and seminal receptacle; vagina usually absent; eggs ectolecithal; male and female pores separate or combined, male anterior to female; common orogenital pore present or absent. The group is likely not monophyletic.

Habitat: Species known from freshwater, terrestrial, and marine environments; mostly free-living; a few known from the body surfaces of polychaete annelids.

Prolecithophora

See Figure 13 (**Plate 1**). With or without ciliated groove around anterior of body; anterior adhesive disc usually absent (present in Hypotrichinidae); extensive epidermal spicules present or absent; cilia uniform throughout all surfaces of body or restricted to ventral surface; brain encapsulated or not; statocyst absent; one to three pairs of ocelli present or absent; mouth anterior, midventral or posterior; with simple, plicate or weakly bulbous pharynx; intestine in form of simple elongate sac, well defined (surrounded by tunic of connective tissue) in most; paired protonephridial system present; testes, one or two, compact or diffuse; sperm biflagellated, flagella not incorporated into body of sperm, mitochondria with elaborate outer cell membrane foldings; seminal vesicle in some; male copulatory stylet rarely present; ovary and vitellarium not fully separated in some, but separated in others; ovary compact or diffuse, paired or not; vitellarium usually diffuse; accessory female openings present or not; eggs ectolecithal; male pore opening anterior to female opening but into common atrium; combined mouth and gonopore or gonopore opening separate from mouth.

Habitat: Predominantly marine, some inhabiting freshwater; free-living, or symbiotic (commensal or parasitic) on gills, on body surfaces, in mantle cavity of gastropod and lamellibranch molluscs, or on surface of small crustaceans; rarely parasitic on anal and branchial regions of fish; some of these associations may be phoretic, with the worm using the host merely for transport.

Tricladida

See Figure 14 (**Plate 1**). Usually large, flattened, elongate; body sometimes pigmented, terrestrial forms often colorful; some with triangular, semilunate, or pointed anterior region of body; caudal adhesive disc present or absent; lamellated rhabdites present in some; duogland adhesive system present; cilia uniform throughout all surfaces of body or restricted to ventral surface; statocyst absent; with many ocelli or one pair of ocelli (cave dwelling species lack ocelli), ocelli sometimes with distinct lens; mouth in middle or posterior half of body; pharynx plicate or tubular, posteriorly directed, usually one, occasionally more pharynges; intestine tripartite with one anterior and two posterior branches, posterior branches fused or not, diverticula sometimes present; anterior branch sometimes anterior to brain; protonephridial system paired; testes, two or many; sperm biflagellated, flagella not incorporated into body of sperm, 9 + '1' microtubule arrangement; seminal vesicle present or absent; male copulatory structures complex; penis usually pyriform, usually unarmed, but sometimes armed with spines or stylet; prostate organ usually absent; one pair of small ovaries usually at anterior of body, often anterior to brain; female copulatory structures variable, with or without copulatory bursa; accessory genital organs may be present in form of extra bursae each with separate pore; vitellarium follicular, extensive; genitointestinal canal sometimes present; eggs ectolecithal; gonopore usually single, rarely two or three; unique temporary embryonic pharynx (i.e., "blastomerenanarchie") and intestine replaced by definitive pharynx during development to adult stage.

Habitat: Predominantly free-living in marine, brackish water, freshwater, or moist terrestrial environments (suborders associated with habitats); one family known only from freshwater caves; some symbiotic with other invertebrates, for example, in mantle cavity of gastropods, or gill lamellae of horseshoe crabs, or on dorsal surface of skates.

Fecampiida

See Figure 15 (**Plate 1**). Evenly ciliated (dorsally and ventrally), with rhabdites (without ciliated groove); pharynx variabilis, or lacking; oral–genital pore terminal (or mouth, pharynx, and intestine lacking, or with pad of epidermal cells on dorsal side of body); ovaries and vitellaria separated; ovary and testes separated or combined in ovotestis; gonopore posterior or absent; thickening of cutaneous muscles.

Habitat: Symbiotic with invertebrates such as molluscs, crustaceans, polychaetes, and occasionally teleosts.

Subphylum Neodermata

Ehlers championed the recognition of the remaining platyhelminth groups as the Neodermata because they share a unique outer body covering in the form of a neodermis, which is sometimes also referred to as a tegument (**Figure 16**). Neodermatans generally begin life with a single-layered, cellular, ciliated epidermis, like most of the non-neodermatan platyhelminth groups. But, as they mature this outer epidermal layer is shed and replaced with a neodermis. This

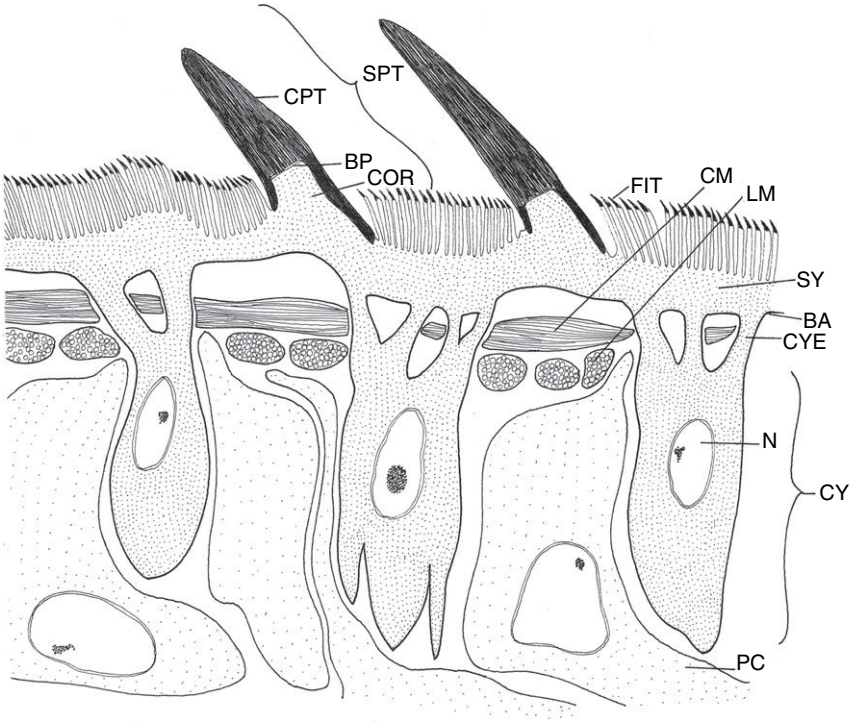


Figure 16 Neodermis of a cestode. Abbreviations: BA, basal lamina; BP, base plate; CM, circular muscle; COR, core; CPT, cap tubules; CYE, cytoplasmic extension; FIT, filithrix; LM, longitudinal muscle; SPT, spinithrix; SY, syncytium.

transformation usually occurs when the first larval stage encounters its host. The neodermis is characterized by the fact that the cytoplasm of neodermal cells is continuous throughout the surface of the animal as there are no cell membranes separating adjacent cells in this layer (i.e., it is a syncytial layer). Furthermore, the nuclei of these cells are generally located below the basal lamina and outer muscle layers of the body, in structures called cytons or pericarya. These are the nuclei of the stem cells, the cytoplasm of which grew outward to form the syncytial layer of the neodermis. The cytons are connected to the outer syncytial layer by thin cytoplasmic extensions. Furthermore, the configuration of the sperm flagella is unusual in neodermatans in that rather than remaining separate from the main sperm body, they are incorporated into the cytoplasm of the sperm body. Unlike other platyhelminths, the neodermatans are generally associated with one or more vertebrate hosts at some point in their lives.

There are four major neodermatan groups: the Monogenea (i.e., the Monopisthocotylea and Polyopisthocotylea), the Aspidogastrea and the Digenea (collectively referred to as the Trematoda), and the Cestoda. The orders of these major groups, or superfamilies in the case of the digeneans, are treated separately below.

Class Monogenea

Adult Monogenea (or Monogenoidea) are primarily external parasites of cold-blooded aquatic, or occasionally terrestrial, vertebrates. However, a few occur internally in the body cavity or very rarely the digestive system of their hosts and one parasitizes the eyes of hippopotamuses. The majority of monogeneans possess a ciliated, free-swimming larval stage known as an oncomiracidium, although the larvae of several groups are unciliated.

Adults generally possess both anterior and posterior attachment structures. The form of the anterior attachment structure varies among groups, consisting of one or two sucker-like organs, pseudosuckers, or grooves. In some cases this region is equipped with gland cells. In the past this anterior structure has been referred to as a prohaptor. The posterior attachment structure also varies widely in morphology among taxa and its form is the primary distinction between the two subclasses of monogeneans, the monopisthocotyleans and the polyopisthocotyleans. This structure is generally now referred to as a haptor, although in the past the term opisthaptor has also been used. This structure is also present in the oncomiracidial stage. The haptor may be symmetrical or asymmetrical, may bear glands, suckers, and may be subdivided into numerous loculi. In addition, the haptor of most monogeneans is armed with a series of hardened structures greatly facilitating attachment to the host. The terminology of these structures is based to some extent on their morphology, but also to a large extent on when they appear in the course of development. For example, the small hooked structures that are usually found around the periphery of the haptor of the ciliated larval stage, or oncomiracidium, and which persist into the adult stage of some monopisthocotyleans, are termed hooklets or uncinuli. The larger, more centrally located hooked structures are generally termed hamuli or anchors, depending on whether or not they are present in the

oncomiracidium. The more complex structures consisting of multiple sclerites, which are found only in polyopisthocotyleans, are termed clamps. Four rhabdomeric ocelli are present in at least one stage of development. The mouth is almost always located in the anterior region of the body. Most monogeneans possess a conspicuous muscular pharynx and an intestine that is usually divided into two, or rarely one, blind cecum; the caeca are occasionally branched or anastomose posteriorly. All monogeneans are hermaphroditic. The male and female genital pores are ventral and usually combined. The number of testes ranges from one to many among different groups. Most groups possess a single ovary that varies somewhat in form from compact to double inverted U-shaped. The male intromittent organ can be a penis or a cirrus, but the exact form of this feature is not well known in some groups. The terminal genitalia are often complex. The excretory system consists of two sets of protonephridia that open separately to the outside, usually in the anterior of the body, but that may be connected at some point in the body.

Molecular work has called the monophyly of the class Monogenea into question. This is reflected in the unresolved nature of this node in the tree presented in **Figure 1**. However, the mutual monophyly of the subclasses Monopisthocotylea and Polyopisthocotylea appears to be well supported. Members of these subclasses are relatively easy to recognize. Most conspicuously, the Polyopisthocotylea possess a haptor that is subdivided into multiple suckers or bears distinctive attachment structures known as clamps, and the Monopisthocotylea, possess a haptor that, although it may bear loculi, is undivided and, although it may bear anchors or hamuli, never bears clamps. Furthermore, unlike the monopisthocotyleans, the polyopisthocotyleans possess a genitointestinal canal such that their genital ducts open into their intestine. Boeger and Kritsky (2001) provide a number of additional features unique to each subclass. The interrelationships among the orders in each subclass are controversial. The orders in each Subclass are treated below in alphabetical.

To make the following diagnoses more concise, unless mentioned, it should be assumed that an order possesses the following: a posterior, discoidal, undivided, symmetrical haptor, a conspicuous muscular pharynx, a gut consisting of two blind intestinal ceca, paired testes, biflagellated sperm, flagella incorporated into body of sperm, 9+'1' microtubule arrangement, a single ovary, a vitellarium that is follicular with follicles arranged in two extensive lateral fields, operculate ectolecithal eggs, and a single, ventral genital pore (**Plate 2**).

Subclass Monopisthocotylea

Capsalidea

See Figure 17 (**Plate 2**). Anterior end with paired petaloid head organs or paired glandular pseudosuckers; haptor sometimes subdivided by septa into central area and multiple peripheral loculi, with 14 marginal hooklets and two or four central anchors and two anterior sclerites; mouth ventral; intestinal ceca fused posteriorly, with lateral and axial diverticula; testes, two to many; sperm microtubules absent; vas deferens forming bipartite external seminal vesicle and simple internal vesicle, or seminal vesicle absent; penis sac present or absent; penis present or absent; penis elongate, muscular, spines absent; some with cirrus; prostatic complex present; vagina present or absent; seminal receptacle present or absent; genital apertures usually combined, marginal; oviparous.

Habitat: Adults parasitic on skin or gills of marine teleosts (including remoras), or on skin or gills of elasmobranchs.

Dactylogyridea

See Figure 18 (**Plate 2**). With two or more pairs of head organs, with or without cephalic glands; posterior half of body covered with anteriorly directed cuticular spines or not; haptor with one or two pairs of anchors supported by one to four transverse bars, with or without accessory plaques, occasionally with two suckers and median terminal anchor complex, usually 14 or 16 marginal hooklets; mouth subterminal, ceca fused posteriorly or not, with lateral diverticula or not; testes usually one; sperm with one axoneme; vas deferens encircling left cecum; seminal vesicle usually present; prostatic complex present; male copulatory organ sclerotized, elongated, muscular, ovary compact, lobed or tubular; vitellarium follicular or occasionally divided into tubular lobules in form of frond-like sprays spreading in two or three lateral groups (Protogyrodactylidae), coextensive with intestinal ceca; seminal receptacle and vagina present or absent; oviparous; eggs oval or tetrahedral; genital aperture median or marginal.

Habitat: Adults parasitic on gills and skin of marine and freshwater, occasionally brackish water, teleosts; rarely on gills or in heart of elasmobranchs.

Gyrodactylidea

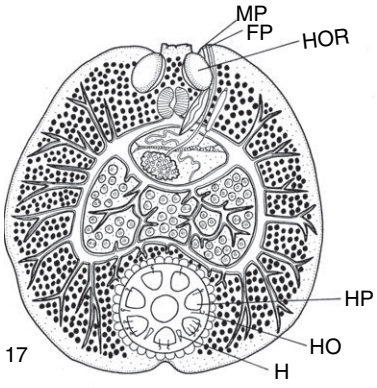
See Figure 19 (**Plate 2**). With two head organs; haptor usually with one pair of anchors often supported by dorsal and ventral transverse bars with 16 marginal hooklets; mouth ventral, intestinal ceca, one or two, usually not fused posteriorly; testis single; sperm microtubules absent; vas deferens encircling left caecum; male copulatory organ sclerotized or muscular, spines present or absent; accessory piece present; vitellarium symmetrical, near posterior part of intestinal cecae; vagina absent or present, opening ventrally near left margin of body; genital pore marginal; most species viviparous; eggs oval or tetrahedral.

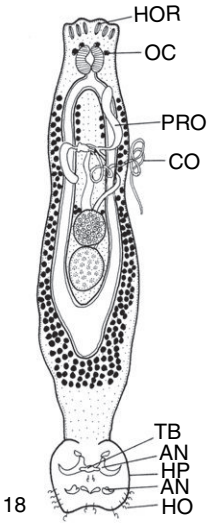
Habitat: Adults parasitic on gills or skin of freshwater and marine teleosts, in stomach and intestine of amphibians, and gills and tentacles of squid; udonellids on marine crustaceans.

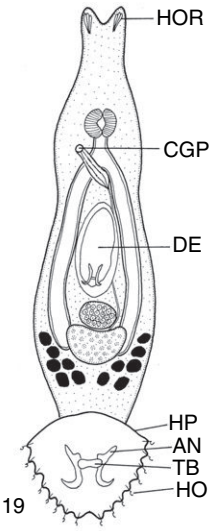
Lagarocotylidea

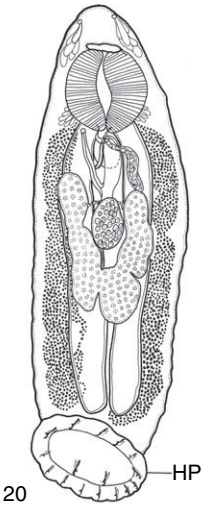
See Figure 20 (**Plate 2**). Oral sucker-like organ absent; with cephalic glands; haptor with 14 submarginal and 2 subcentral hooklets; mouth midventral; ceca not fused posteriorly, without lateral diverticula; with single U- or V-shaped testis; vas deferens encircling left caecum; seminal vesicle present; prostatic reservoir dorsal to copulatory complex; male copulatory organ sclerotized, with accessory piece; ovary compact; vitellarium consisting of two dense bilateral bands of follicles; seminal receptacle absent; vaginal pore dextro-ventral at level of genital pore; eggs without filaments; genital aperture median.

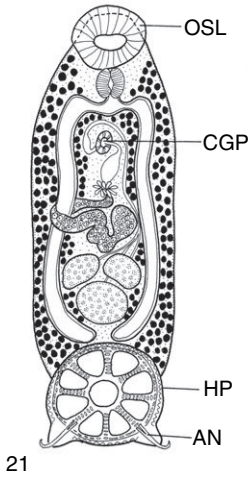
Habitat: Adults parasitic in the cloaca of salamanders.

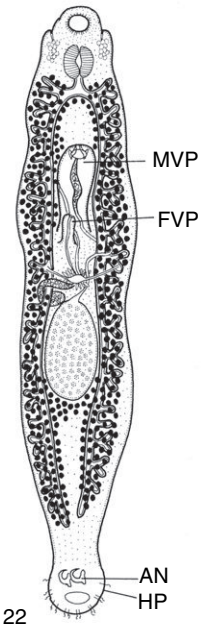

MP
FP
HOR
HP
HO
H
17

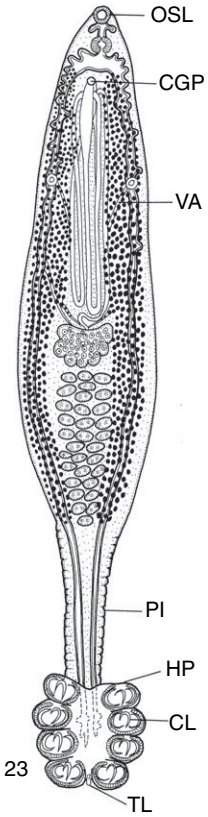

HOR
OC
PRO
CO
TB
AN
HP
AN
HO
18

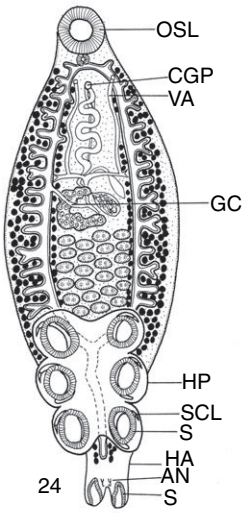

HOR
CGP
DE
HP
AN
TB
HO
19

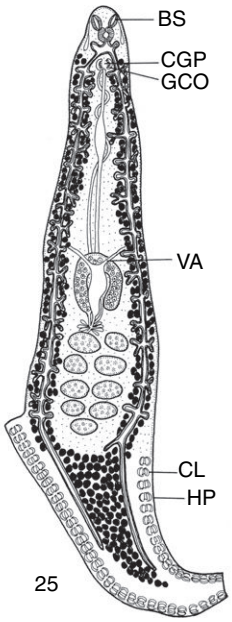

HP
20

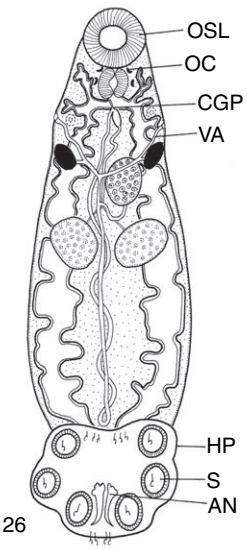

OSL
CGP
HP
AN
21


MVP
FVP
AN
HP
22


OSL
CGP
VA
PI
HP
CL
TL
23


OSL
CGP
VA
GC
HP
SCL
S
HA
AN
S
24


BS
CGP
GCO
VA
CL
HP
25


OSL
OC
CGP
VA
HP
S
AN
26

Monocotylidea

See Figure 21 (Plate 2). With single anterior sucker or several sucker-like depressions or paired head organs (lobes) with cephalic glands; haptor aseptate or divided by septa into loculi, often with one pair of anchors, with or without marginal papillae, usually with 14 marginal hooklets; ocelli present or absent; mouth ventral, intestinal ceca occasionally fused posteriorly, occasionally with lateral diverticula; with one to several testes; sperm microtubules lying along one quarter of cell periphery, with two axonemes one of which is reduced; male copulatory organ usually sclerotized; ovary tubular, usually encircling right intestinal caecum; ventral vaginal pore sometimes present; genital pore usually median.

Habitat: Adults parasitic on skin, gills, and sometimes nasal bulbs of elasmobranchs; occasionally in coelom, oviducts, and cloaca of elasmobranchs; occasionally in cloaca of holocephalans.

Montchadskyellidea

See Figure 22 (Plate 2). Haptor with one pair of anchors lacking supporting bars, with 14 marginal hooklets; mouth subterminal, gut diverticula present; testis single; male copulatory organ sclerotized; vas deferens encircling left caecum; accessory piece present; ovary encircling right caecum; one midventral vagina present; genital pores separate, opening on separate ventral papillae.

Habitat: Adults parasitic in gut of marine teleosts.

Subclass Polyopisthocotylea

Chimaericolidea

See Figure 23 (Plate 2). Single, weak oral sucker-like organ usually present; body with highly contractile, sometimes with very long, posterior peduncle; haptor with four pairs or four alternating rows of two short-stalked symmetrical clamps, and terminal lappet-bearing hooks; ceca with lateral diverticula, extending posteriorly into haptor; with numerous testes; seminal vesicle present or absent; cirrus muscular, armed or not; prostatic complex present or absent; ovary lobed or branched; with two lateral vaginae opening into vitelline duct; seminal receptacle present or absent; uterus with numerous longitudinal loops, with or without lateral branches.

Habitat: Adults parasitic on gills or in gill chamber of holocephalans.

Diclybothriidea

See Figure 24 (Plate 2). With anterior muscular, paired ventral "bothria" or oral sucker-like organ; haptor with three pairs of sessile suckers, with medial, haptoral appendix bearing three pairs of hook-like sclerites or one pair of terminal suckers and two small anchors; each haptoral sucker with single, large, curved, hooked sclerite; mouth terminal or nearly so; ceca united posteriorly, with lateral diverticula; testes numerous; male intromittent organ usually armed; ovary tubular, convoluted; with two lateral vaginae; uterus preovarian; vitellarium sometimes extending posteriorly into haptor; eggs with or without polar filaments, occasionally forming chains.

Habitat: Adults parasitic on gills of sturgeon, paddlefish (chondrosteans), and elasmobranchs.

Mazocraeidea

See Figure 25 (Plate 2). With two anterior buccal suckers, occasionally enlarged in form of discoid adhesive organ with conical papillae; haptor symmetrical or asymmetrical, set off from body by distinct peduncle or not, distinctly bilobed or not, usually with four to numerous pairs of clamps, without marginal hooklets, terminal lappet usually present; terminal lappet with one to three pairs of similar or dissimilar hook-like sclerites, bifid or not; testes numerous; copulatory apparatus consisting of sheaf of long spines or longitudinally striated sheath; male intromittent organ present or absent, armed or unarmed; ovary tubular, winding, convoluted, rarely compact or inverted U-shaped; vitellarium sometimes extending posteriorly into haptor; seminal receptacle present or absent; vagina present or absent, paired or not; vaginal aperture spined or not, conical sclerite in front of vaginal pore or not; eggs with two filaments; genitointestinal canal present; genital pores common or separate, genital pore sometimes sucker-like; in rare cases (Diplozoidae) two adults permanently fused in form of an X.

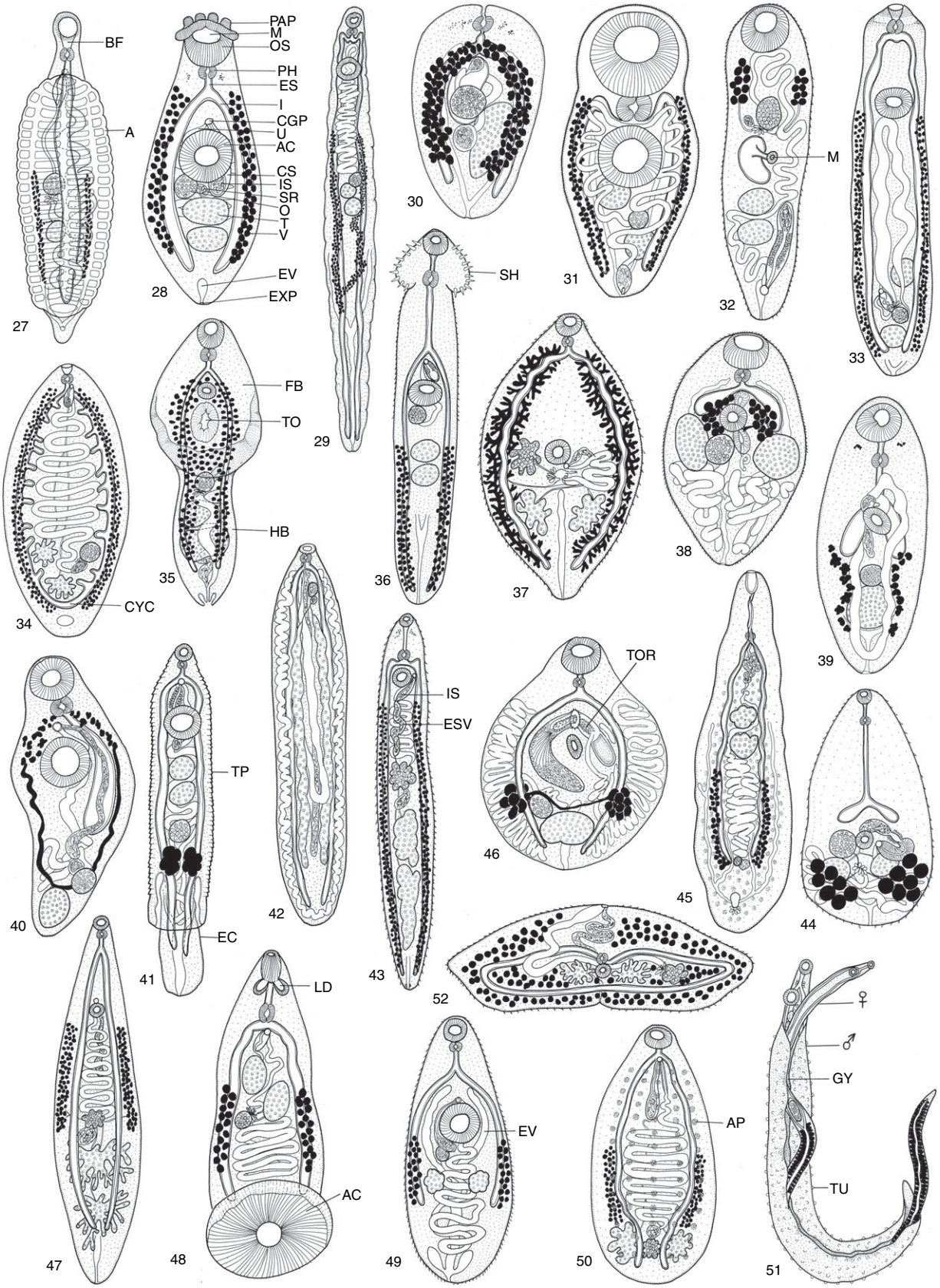
Habitat: Adults parasitic on gills or occasionally mouth cavity of marine, or sometimes freshwater teleosts; sometimes on crustaceans in mouth cavity of teleosts.

Polystomatidea

See Figure 26 (Plate 2). With anterior muscular oral sucker-like organ; haptor with three but occasionally one pair of suckers, usually with one to two pairs of hamuli, sometimes with accessory spines or spurs, often with marginal hooklets, sometimes with medial appendix-like prolongation; ocelli usually absent; mouth terminal or subterminal; ceca occasionally fused posteriorly, often with lateral diverticula; testes, one to many; seminal vesicle present; male intromittent organ usually armed with hooks or spines; genitointestinal canal present; ovary compact; vitellarium rarely compact; uterus anterior or opposite ovary or reaching posterior of body; with two lateral vaginae.

Habitat: Adults parasitic in mouth, pharynx, esophagus, urinary bladder of frogs and toads and gills of tadpoles; mouth, esophagus, nasal passages, or urinary bladder of

Plate 2 Monogenea (Figures 17–26) 17. Capsalidea: adult of *Tristoma* sp. 18. Dactylogyridea: adult of *Ancylodiscoides* sp. 19. Gyrodactylidea: adult of *Gyrodactylus* sp. 20. Lagarocotylidea: adult of *Lagarocotyle salamandrae* (Kritsky, *et al.* 1993); 21. Monocotylidea: adult of *Monocotyle* sp. 22. Montchadskyellidea: adult of *Montchadskyella* sp. (Bychowsky, *et al.*, 1970); 23. Chimaericolidea: adult of *Chimaericola* sp. 24. Diclybothriidea: adult of *Erpocotyle* sp. 25. Mazocraeidea: adult of *Heteraxinoides* sp. 26. Polystomatidea: adult of *Neodiplorchis* sp. Abbreviations: AN, anchor; BS, buccal sucker; CGP, common genital pore; CL, clamp; CO, copulatory organ; DE, developing embryo; FP, female genital pore; FVP, female ventral papilla; GC, genitointestinal canal; GCO, genital corona; H, hook; HA, haptoral appendix; HO, hooklet; HOR, head organ; HP, haptor; MP, male genital pore; MVP, male ventral papilla; OC, ocellus; OSL, oral sucker-like organ; PI, posterior isthmus; PRO, prostatic reservoir; S, sucker; SCL, sclerite; TB, transverse bar; TL, terminal lappet; VA, vagina.


BF
A
27
PAP
M
OS
PH
ES
I
CGP
U
AC
CS
IS
SR
O
T
V
EV
EXP
28
29
30
SH
31
M
32
33
FB
TO
HB
35
34
CYC
36
37
38
39
IS
ESV
TOR
46
45
44
40
TP
EC
41
42
43
52
LD
AC
48
47
EV
49
AP
50
♀
♂
GY
TU
51

turtles; mouth and gills of Australian lungfish; urinary bladder of urodeles; eyes of hippopotamuses.

Class Trematoda

Subclass Aspidogastrea

See Figure 27 (**Plate 3**). With ventral holdfast organ in form of linear series of 20–100 suckers, numerous transverse rugae, or extensive adhesive disc subdivided by septa into numerous loculi; disc with or without papillae; neodermis with microvilli in form of hemispherical microtubercles; buccal funnel at anterior end of body; oral sucker usually absent; oral lobes occasionally present; with one or two blind intestinal ceca; excretory vesicle present, V-shaped, excretory pore near posterior of body; with one, two, or many testes; sperm biflagellated, flagella incorporated into body of sperm, 9 + '1' microtubule arrangement; cirrus present; cirrus-sac present or absent; ovary single; vitellarium follicular or tubular, if follicular usually with multiple lateral follicles, rarely single median tube flanked by follicles, if tubular, two tubes; uterus generally long; Laurer's canal present or absent; Mehlis' gland present; eggs ectolecithal; single common genital pore.

Habitat: Adults endoparasites in kidneys, pericardial chamber, gill lamellae, or intestine of freshwater and occasionally marine lamellibranch molluscs, kidneys (renal cavities) and pericardial chamber of marine and freshwater gastropod molluscs, occasionally guts of turtles or freshwater and marine teleosts, gall bladder and bile ducts of elasmobranchs, or in lumen of rectal glands of Holocephali.

Subclass Digenea

Almost all digeneans are endoparasites as adults in various organ systems of vertebrates. They are monophyletic and are most easily characterized by their possession of a unique series of developmental stages that include miracidium, sporocyst, redia, cercaria, and metacercaria. Variation in life cycles seen among digenean species suggests that perhaps only a subset of these ontogenetic stages should be considered to distinguish the subclass. Most adult digeneans possess an anterior sucker (or oral sucker), which is usually (Figure 28, **Plate 3**), but not always (Figure 32, **Plate 3**), associated with the mouth. Most also possess a second sucker (or acetabulum) that is usually midventral (Figure 49, **Plate 3**), but occasionally posterior in position (Figure 48, **Plate 3**). The body of some species is divided into two parts (Figure 35, **Plate 3**). Other major variations in adult body form include a posterior telescoping ecsoma (Figure 41, **Plate 3**), a sucker-like ventral adhesive tribocytic organ (Figure 35, **Plate 3**), one or more anterior collars of large spines (Figure 36, **Plate 3**), and paired anterior retractile proboscides. Most species possess an anterior mouth that opens into a bulbous pharynx and an intestine consisting of two or occasionally one blind cecum, with or without diverticula. However, the ceca in some species are connected to one another posteriorly (Figure 34, **Plate 3**) or open into the excretory bladder. The excretory system generally consists of two sets of protonephridia, both of which empty into a common excretory vesicle. This vesicle is variable in shape among taxa and usually opens at the posterior end of the body. With the exception of the schistosomes (Figure 51, **Plate 3**) and some didymozoids, all digeneans are hermaphroditic. The male intromittent organ is usually a cirrus, which may or may not be armed with spines, and may or may not be surrounded by a cirrus-sac. Testes are usually oval, but occasionally lobed or tubular; testes can be single, paired, or numerous. Sperm may be stored inside or outside the cirrus-sac in an internal or external seminal vesicle respectively; some taxa possess both vesicles. Most digeneans possess a single ovary that may be round, lobed, or tubular in form. Digeneans lack a vagina; instead, sperm enters the female system through the uterus, which opens, generally in combination with the male system, into a common genital pore. Some species possess a seminal receptacle in the form of a basal expansion of the uterus or as diverticula of the Laurer's canal. A Laurer's canal, which may or may not open to the outside, may be present or absent. The function of the Laurer's canal is unknown, but it is considered by some to represent the remnant of a vagina. The vitellarium consists of columns of numerous follicles along both sides of the body in most taxa, but is occasionally present as one or a pair of compact follicles or vitelline chords.

With at least 1592 nominal genera, the digeneans are by far the most diverse of the platyhelminth groups. Unlike all of the other major platyhelminth groups, superfamily, rather than order, is the highest taxonomic category of relative stability within the subclass. Unfortunately, although the phylogenetic relationships among some of the major digenean taxa have been explored (e.g., Cribb *et al.*, 2001; Olson *et al.*, 2003), the

Plate 3 Trematoda (Figures 27–52) 27. Aspidogastrea: adult of *Aspidogaster* sp.; 28. Allocreadioidea: adult of *Crepidostomum* sp.; 29. Azygioidea: adult of *Otodistomum veliporum* 30. Bivesiculoidea: adult of *Treptodemoides fukenensis* 31. Brachylaimoidea: adult of *Leucochloridium vogtianum* 32. Bucephaloidea: adult of *Bucephalopsis* sp. 33. Clinostomoidea: adult of *Odhneriotrema* sp. 34. Cyclocoeloidea: adult of *Typhlocoelum* sp. 35. Diplostomoidea: adult of *Neodiplostomum* sp. 36. Echinostomatoidea: adult of *Echinochasmus* sp. 37. Gorgoderoidea: adult of *Paragonimus* sp. 38. Gymnophalloidea: adult of *Pseudogymnophallum alcae* 39. Haploporoidea: adult of *Saccocoelioides octavus* 40. Haplosplanchoidea: adult of *Provitellotrema crenimugilis* 41. Hemiuroidea: adult of *Parahemiurus* sp. 42. Heronimoidea: adult of *Heronimus mollis* 43. Lepocreadioidea: adult of *Echeneidocoelium* sp. 44. Microphalloidea: adult of *Microphallus* sp. 45. Microscaphioidea: adult of *Paradeuterobaris novaguineae* 46. Monorchioidea: adult of *Postmonorchis orthopristis* 47. Opisthorchioidea: adult of *Clonorchis* sp. 48. Paramphistomotoidea: adult of *Megalodiscus* sp. 49. Plagiorchioidea: adult of *Styphlotrema* sp. 50. Pronocephaloidea: adult of *Quinqueserialis* sp. 51. Schistosomatoidea: male and female adults of *Schistosoma* sp.; 52. Transversotrematoidea: adult of *Transversotrema* sp. Abbreviations: A, adhesive disc; AC, acetabulum; AP, adhesive papillae; BF, buccal funnel; CGP, common genital pore; CS, cirrus-sac; CYC, cyclocoel; EC, ecsoma; ES, esophagus; ESV, external seminal vesicle; EV, excretory vesicle; EXP, excretory pore; FB, forebody; GY, gynecophoral canal; HB, hindbody; I, intestine; IS, internal seminal vesicle; LD, lateral diverticula of oral sucker; M, mouth; O, ovary; OS, oral sucker; PAP, papilla; PH, pharynx; SH, spiny head collar; SR, seminal receptacle; T, testis; TO, tribocytic organ; TOR, terminal organ; TP, transverse plications; TU, tubercles; U, uterus; V, vitellarium.

relationships among the 25 superfamilies and 148 families remain controversial and poorly understood. The digeneans are treated here following the classification scheme embraced in the recent CABI keys (see Gibson *et al*., 2002; Jones *et al*., 2005; Bray *et al*., 2008); superfamilies are considered below in alphabetical order. It is important to note, however, that the monophyly of many of these superfamilies also remains uncertain.

To make the following diagnoses more concise, unless mentioned, it should be assumed that a superfamily possesses the following features: no specialized body regions or structures, both suckers (oral sucker and acetabulum) are present, the mouth opens into the oral sucker, a muscular pharynx is present, the gut consists of two blind intestinal ceca, they are hermaphroditic, the testes are compact and paired, the sperm are biflagellated and the flagella are incorporated into the body of the sperm, the flagella of the sperm possess a 9 + '1' arrangement of microtubules, a cirrus and cirrus-sac are present, the ovary is single, the vitellarium is follicular and the follicles are arranged in two lateral fields, there is a common genital pore, the eggs are ectolecithal and the epidermal cells of the miracidium are covered with cilia.

Allocreadioidea

See Figure 28 (**Plate 3**). Oral sucker with or without one or three pairs of muscular papillae; body surfaces usually unspined; genital pore opens near acetabulum; excretory vesicle I-shaped. Developmental stages: Cotylocercous xiphidiocercaria, ophthalmoxiphidiocercaria, trichocercous, or homalometrine cercaria; metacercaria usually encysts in invertebrate hosts, rarely on substrate.

Habitat: Adults parasitic in digestive tract and occasionally gall bladder of marine and freshwater teleosts, or occasionally gall bladder of freshwater turtles, or intestine of chameleons; some species progenetic in insects.

Azygioidea

See Figure 29 (**Plate 3**). Body surface usually unspined; genital pore opens anterior to acetabulum; sinus organ present; uterus usually entirely preovarian; excretory vesicle Y-shaped. Developmental stages: Cercaria nonoculate furcocercous, furcocystocercous, or cystocercous; metacercaria unknown.

Habitat: Adults parasitic in stomach or body cavity of elasmobranchs (sharks and rays) and in stomach of freshwater teleosts and holosteans (gar, bowfin, etc.).

Bivesiculoidea

See Figure 30 (**Plate 3**). Oral sucker and acetabulum absent; body surface with minute spines; testis single; gential pore midventral; uterine seminal recepticle present; excretory vesicle V-shaped. Developmental stages: Cercaria furcocercous, cystophorous, oculate furcocystocercous.

Habitat: Adults parasitic in intestine of marine and occasionally freshwater teleosts.

Brachylaimoidea

See Figure 31 (**Plate 3**). Body surfaces spined or not; ovary usually between testes; ovary and testes often close to posterior extremity of body; genital pore opens in hindbody, often near posterior extremity; excretory vesicle Y-shaped. Developmental stages: Cercaria furcocercous or cercariaeum with rudeimentary or no tail that may or may not leave the sporocyst; metacercaria encysted or unencysted in second intermediate host or within sporocyst in second intermediate host.

Habitat: Adults parasitic in intestine of aquatic birds and terrestrial mammals, birds and occasionally reptiles and amphibians.

Bucephaloidea

See Figure 32 (**Plate 3**). Body surface usually spined; oral sucker and acetabulum absent; anterior holdfast organ a rhynchus, with or without sucker; mouth opening on ventral surface; intestinal caecum simple, not bifurcating; genital pore usually terminal; excretory vesicle I-shaped. Developmental stages: Cercaria of bucephalid type; metacercaria encyst in teleosts.

***Habitat*:** Adults parasitic in digestive system or sometimes in swim bladder or body cavity of marine and freshwater teleosts and occasionally intestine of salamanders.

Clinostomoidea

See Figure 33 (**Plate 3**). Oral sucker weak, enveloped in collar-like anterior end of body, pharynx rudimentary or small; genital pore opens in posterior third of body. Developmental stages: Brevifurcate cercaria; metacercaria encysts on vertebrate hosts.

Habitat: Adults parasitic in mouth and esophagus of alligators and piscivorous birds.

Cyclocoeloidea

See Figure 34 (**Plate 3**). Oral sucker usually absent, occasionally present; acetabulum absent or, if present, weakly developed; intestinal ceca usually fused posteriorly to form cyclocoel; genital pore opens near pharynx. Developmental stages: Cercaria usually tailless or with rudimentary tail; metacercaria encysts within sporocyst or redia in molluscan host.

Habitat: Adults parasitic in nasal sinuses, nasolacrimal sinus, infraorbital sinuses, respiratory tract, and occasionally coelom or intestine of birds.

Diplostomoidea

See Figure 35 (**Plate 3**). Body generally divided into anterior forebody and posterior hindbody; tribocytic organ present posterior to acetabulum; reproductive organs generally concentrated in hindbody; genital pore usually opens in posterior of body. Developmental stages: Cercaria variable; metacercaria generally encysts on vertebrate hosts, very rarely invertebrate hosts.

Habitat: Adults parasitic in digestive system of birds and reptiles, especially crocodilians and snakes, sometimes amphibians, sometimes mammals including dogs, cats, dolphins, and some monotremes.

Echinostomatoidea

See Figure 36 (**Plate 3**). Spiny head collar often present; two anterior, retractile, spiny proboscides rarely present; body usually spined; ceca sometimes with lateral diverticula; hermaphroditic-sac sometimes present, often with cirrus-sac, cirrus occasionally absent; genital pore variable in position. Developmental stages: Cercaria echinostome-like,

megalurous, gymnocephalous megaperid or haplosplanchnid; metacercaria usually encysts on aquatic vegetation, sometimes on vertebrates, rarely in invertebrates.

Habitat: Adults parasitic in the intestinal tract, gall bladder, bile duct, liver, lungs, nasal cavity, conjunctival sac, bursa fabricius, and orbit of vertebrates including a variety of marine and freshwater teleosts, birds, crocodiles, turtles, snakes, lizards, marine mammals including pinnepeds and cetaceans, and a variety of terrestrial herbivorous mammals.

Gorgoderoidea

See Figure 37 (**Plate 3**). Body surfaces spined or not; acetabulum well developed, sessile or protuberant; intestinal caeca usually two, occasionally one, may unite posteriorly; cirrus-sac present or absent; genital pore usually anterior to acetabulum; vitellarium follicular or dendritic, usually limited in extent; uterus usually in hindbody; excretory vesicle variable. Larvae: Microcercous xiphidiocercaria; metacercaria encysts on invertebrates or vertebrates.

Habitat: Adults parasitic in intestine, gall bladder, bileducts, pancreatic ducts, body and pericardial cavities, urinary bladder, ureters, swim bladder, nasal and paranasal sinues, lungs, brain, or encysted in skin and in wall of digestive tract of a diversity of vertebrates (elasmobranchs, teleosts, herptiles, birds, and mammals).

Gymnophalloidea

See Figure 38 (**Plate 3**). Mouth opening into oral sucker or acetabulum (bucephalids); cirrus-sac present or absent; cirrus sometimes absent; genital pore variable in position. Developmental stages: Cercaria often produced in branched sporocysts; cercaria often with rudimentary tail or tailless, often gasterostomous; metacercaria encysts on vertebrate hosts or within sporocyst in molluscan host; very rarely encysts in invertebrates.

Habitat: Adults parasitic in liver, bile duct, gall bladder, digestive tract of freshwater and marine teleosts, or large intestine, cloaca, bursa fabricius of birds, occasionally mammals (such as, for example, hedgehogs, rabbits, ruminants).

Haploporoidea

See Figure 39 (**Plate 3**). Body surfaces usually spined; eye-spot pigment usually present, dispersed in forebody; oral sucker with or without lobes or hood; testes usually one, occasionally two; cirrus-sac absent; genital pore anterior to acetabulum; excretory vesicle I-, Y-, or V-shaped.

Habitat: Adults parasitic in digestive system of marine, freshwater, and estuarine teleosts.

Haplosplanchoidea

See Figure 40 (**Plate 3**). Body surfaces unspined; oral sucker normal, occasionally replaced by disc with or without lobes; intestinal caecum single; testis usually one, rarely two; cirrus-sac usually absent, rarely present; genital pore anterior to acetabulum; excretory vesicle Y-shaped.

Habitat: Adults parasitic in digestive system of marine, freshwater, and estuarine teleosts.

Hemiuroidea

See Figure 41 (**Plate 3**). Body generally elongate; sometimes with posterior telescoping ecsoma; body surface unspined but often with transverse plications; ceca sometimes fused to one another posteriorly, or fused posteriorly with excretory vesicle; usually hermaphroditic but one group (didymozoids) occasionally gonochoristic (if gonochoristic adult sexually dimorphic, male and female usually encysted in pairs); hermaphroditic duct almost always present; cirrus-sac often absent; vitellarium often in form of rosette or tubular, or in two compact masses. Developmental stages: Miracidium spinous, lacking cilia; cercaria cystophorous; metacercaria usually unencysted in invertebrate or vertebrate hosts, occasionally within redia in molluscan host.

Habitat: Adults parasitic in digestive tract or swim bladder of freshwater and marine teleosts and rarely elasmobranchs; some associated with branchial cavities of marine teleosts and elasmobranchs; occasionally in mouth of frogs or stomach of sea snakes.

Heronimoidea

See Figure 42 (**Plate 3**). Body surfaces unspined; oral sucker often poorly developed; acetabulum absent; genital pore at level of pharynx; testes elongate, tubular, extend for length of body; vitellarium in form of slender elongate bands, extend for much of length of body; excretory vesicle tubular. Developmental stages: cercaria amphistomate; metacercaria apparently lacking, infection of definitive host is through ingestion of cercaria in the molluscan host.

Habitat: Adults parastic in lungs and trachea of freshwater turtles.

Lepocreadioidea

See Figure 43 (**Plate 3**). Body surfaces spined or not; with or without circumoral spines; oral sucker occasionally with papilliform extensions; acetabulum ventral or rarely at posterior end of body; ceca open into excretory vesicle or not; one or usually two intestinal ceca present; with two to many testes; cirrus-sac usually present, sometimes absent; hermaphroditic duct in some; external seminal vesicle in some; internal seminal vesicle bipartite in some. Developmental stages: Gymnophalous cercaria, produced in redia; metacercaria encysts on invertebrates, or on substrate, rarely in sporocyst or redia in molluscan host.

Habitat: Adults parasitic almost exclusively in digestive tract, rarely gall bladder or urinary bladder of marine teleosts; one record from lungs of reptile.

Microphalloidea

See Figure 44 (**Plate 3**). Body surfaces spined or not; oral sucker may be surrounded by semicircular collar of spines; acetabulum poorly developed, rarely absent; gut usually present, occasionally lacking or poorly developed (with short ceca); cirrus-sac present or absent; genital atrium sometimes with complicated accessory organs; vitellarium usually in restricted symmetrical clusters of large follicles; excretory system V-, Y-, or I-shaped. Developmental stages: Xiphiodiocercaria of ubiquita, microphallous, or virgulate type; metacercaria encysts on invertebrates or vertebrates, or occasionally in sporocyst in the molluscan host.

Habitat: Adults parasitic in many organ systems of most vertebrate classes.

Microscaphidioidea

See Figure 45 (Plate 3). Body surfaces spined or not; oral sucker and acetabulum absent; part of all of ventral surface may function as accessory attachment organ; some with ventral papillae; pharynx with or without paired intra or extramural sacs; cirrus-sac absent; uterine seminal receptacle often present; excretory system entirely or partly reticular. Developmental stages: cercaria apharyngeal monostomous; metacercaria encysting on vegetation.

Habitat: Adults parasitic in digestive system of marine and freshwater teleosts, and turtles.

Monorchioidea

See Figure 46 (Plate 3). Body surfaces usually spined; oral sucker rarely with ring of spines; testes one, two, or rarely 8; uterus usually with numerous coils; terminal organ associated with terminal part of uterus in some; excretory vesicle I-, V-, or Y-shaped. Developmental stages: cercaria often microcercous or completely lacking tail, some with stylet; metacercaria often progenetic (i.e., producing eggs), encysting in molluscs, planarians, aquatic annelids and larval insects.

Habitat: Adults parasitic in digestive system of marine and freshwater teleosts.

Opisthorchiodea

See Figure 47 (Plate 3). Body sometimes spined; oral sucker sometimes with ring of spines; cirrus-sac usually lacking; hermaphroditic duct sometimes present opening into genital atrium; acetabulum sometimes enclosed in ventrogenital sac, with or without gonotyl; seminal vesicle when present often bipartite. Developmental stages: Cercaria generally with ocelli, pleurolophocercous, parapleurolophocercous or gymnocephalus cercaria; metacercaria encysts on vertebrates, rarely invertebrates.

Habitat: Adults parasitic in stomach, intestine, occasionally bile ducts of marine teleost fishes, snakes, lizards, birds, and mammals; occasionally ovary of freshwater teleosts; occasionally progenetic encysted in musculature of frogs.

Paramphistomoidea

See Figure 48 (Plate 3). Usually unspined; oral sucker often with paired lateral diverticula, occasionally one unpaired diverticulum; acetabulum posterior or subterminal, sometimes with muscular lip, occasionally lacking; body occasionally with caudal appendages; usually with lymphatic system (usually with one to three longitudinal trunks); genital sucker sometimes present; ventral pouch sometimes present; cirrus-sac present or absent; hermaphroditic-sac in some. Developmental stages: Cercaria lacking acetabulum or with posterior acetabulum, develops from redia or branched sporocyst; metacercaria usually encysts on vegetation, but occasionally on frogs, metacercarial stage occasionally lacking.

Habitat: Adults parasitic in digestive tract of marine and freshwater teleost fishes, snakes, marine turtles, amphibians, birds and a wide diversity of marine and terrestrial mammals (including humans), or in lungs and trachea of turtles and tortoises.

Plagiorchioidea

See Figure 49 (Plate 3). Body usually spined; oral sucker occasionally with pair of muscular lateral papillae; acetabulum rarely vestigial; fields of vitellarium usually restricted anteriorly and posteriorly; excretory vesicle Y-shaped. Developmental stages: Cercaria of leptocercous, microcercous, cystocercus, armatao, ornatae, virgulate, ubiquita, or corcacriaeum type; often developing in daughter sporocysts in gastropod or lamellibranch molluscs; metacercaria encysts in terrestrial arthropods or gastropods.

Habitat: Adults parasitic in digestive system of mammals (especially bats), turtles, amphibians, reptiles, and occasionally fishes; liver and bile ducts of mammals, birds, and reptiles; cloaca and oviducts of birds and mammals; bursa fabricius of birds; urinary bladder of fishes and amphibians; coelom or pericardial chamber of elasmobranchs; occasionally progenetic in leeches and crustaceans.

Pronocephaloidea

See Figure 50 (Plate 3). Body surfaces spined or unspined; acetabulum absent; often with ventral longitudinal ridge or series of ventral papillae; head collar present or absent; pharynx absent; male and female pores occasionally separate, pore(s) variable in position; uterus usually in form of transverse coils anterior to ovary; vitellarium follicular, rarely in form of single lobed mass; excretory vesicle V- or Y-shaped, sometimes extremely modified in structure; posterior invaginations of body forming secondary excretory vesicle or protrusible tentacles associated with excretory vesicle. Developmental stages: Cercaria monostomate, bi or trioculate, with posterior adhesive glands at posterior of body; metacercaria encysts on substrate or on invertebrate hosts, rarely in sporocyst or redia within molluscan host.

Habitat: Adults parasitic in oviduct, urinary and gallbladders, pancreas, liver, respiratory tract, tissues or digestive system of teleosts, marine turtles, marine iguanids, birds, mammals (including dugongs, cetaceans, bats, rodents, and insectivores), occasionally in bursa fabricius of birds.

Schistosomatoidea

See Figure 51 (Plate 3). Gonochoristic or hermaphroditic, if gonochoristic, male with ventral gynecophoral canal in which female resides for much of her life; surface of male body often with tubercles, with or without spines; pharynx usually lacking; both oral sucker and acetabulum present, or one or both lacking; ceca of gut often anastomosing at several points along length; one, two or many testes; genital pores separate or combined. Developmental stages: Cercaria apharynegeate, brevifurcate, sometimes also lophophorate; metacercarial stage lacking; cercaria develops into schistosomula; first host usually a mollusc but rarely an annelid.

Habitat: Adults parasitic in blood vessels or heart of freshwater and marine teleosts, elasmobranchs, holocephalans, usually freshwater (but sometimes marine) birds and mammals (including humans), crocodiles, and marine and freshwater turtles.

Transversotrematoidea

See Figure 52 (Plate 3). Body transversely elongated; spined; ceca united posteriorly to form transverse cyclocoel; excretory

vesicle curved to right, receiving collecting vessels from both sides of body; cirrus absent; genital pore anterior. Developmental stages: Cercaria brevifurcate with pair of anterior lateral appendages; metacercarial stage lacking; cercaria infects definitive host cutaneously and develops directly into adult stage.

Habitat: Adults parasitic under scales of freshwater, brackish water, and marine teleosts.

Class Cestoda

The cestodes, or tapeworms, conspicuously differ from almost all other platyhelminth groups in their lack of all elements of a digestive system, the only exceptions perhaps being the amphilinideans and gyrocotylideans, the anterior attachment organ of which is considered by some to represent a vestige of a digestive system. Thus, in general, the cestodes have no mouth, pharynx, esophagus, or intestine at any stage in their development. Rather, they absorb nutrients, in the form of very small molecules, directly through the outer layer of their body. There is evidence to suggest that their unique neodermal surface extensions, known as microtriches, assist in this process. Microtriches typically consist of an outer electron-dense cap composed of cap tubules, a baseplate, and a basal core consisting of an extension of the cytoplasm of the tegument (**Figure 16**). The entire structure is surrounded by a plasma membrane. These structures may also assist with attachment. Microtriches appear to be of two basic types. The larger form, termed a spinithrix (pl. spinitriches) (SPT in **Figure 16**), is greater than 200 nm in width; the smaller form, termed a filithrix (pl. filitriches) (FIT in **Figure 16**), is less than or equal to 200 nm in width.

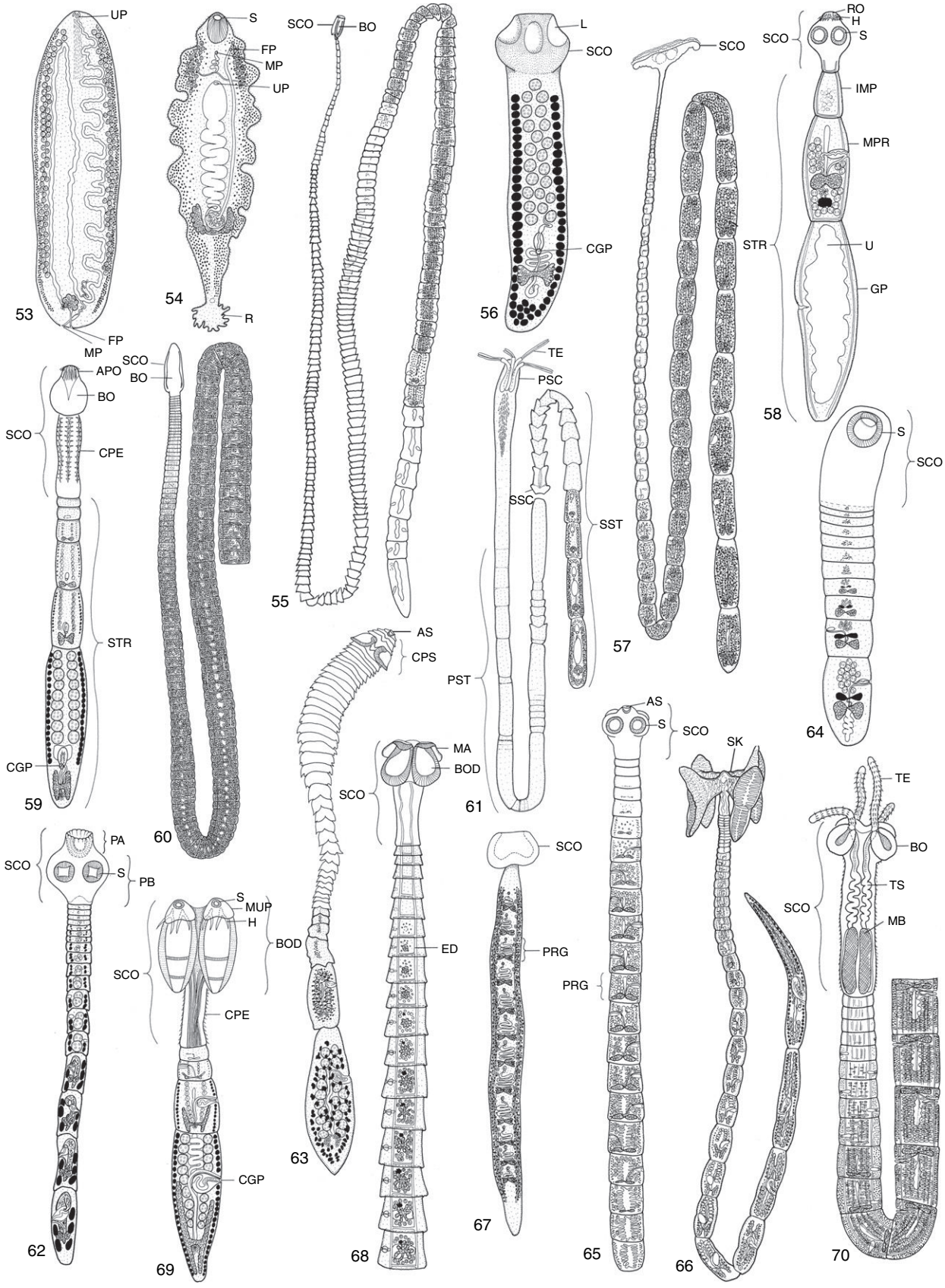
The majority of tapeworms have a body consisting of an attachment region called the scolex, followed by a series of compartments, termed proglottids, that are collectively referred to as the strobila. Scolex morphology varies significantly among the various tapeworm orders. The scolex is simple in some groups, but it is more commonly equipped with attachment structures such as hooks, suckers, loculi, or hooked tentacles. In some groups, the scolex is subdivided into two or four flexible bothria or four muscular bothridia. Bothridia can be distinguished from bothria in their possession of discrete inner and outer membranes bounding their musculature, visible most readily in histological section.

The proglottids of most tapeworms are hermaphroditic, each containing one, or less commonly two or more, sets of male and female reproductive organs. A few orders of tapeworms (specifically the amphilinideans, gyrocotylideans, and caryophyllideans) lack proglottids and also possess only a single set of male and female reproductive organs rather than multiple sets (i.e., they are monozoic rather than polyzoic). The members of one order of tapeworms, the spathebothriideans, lack distinct proglottids but possess multiple sets of reproductive organs arranged in a linear series (i.e., they are nonproglottized but polyzoic). The majority of tapeworms, however, are proglottized, and each proglottid bears one or more sets of reproductive organs (i.e., they are proglottized and polyzoic). In proglottized tapeworms, a germinitive region located directly behind the scolex consecutively produces the proglottids of the strobila. Thus, the proglottid nearest the scolex is the youngest, and that furthest from the scolex is the oldest. In some tapeworms, the posterior-most proglottids, on reaching a certain level of maturity, drop from the strobila (apolytic) and may continue development independent of the main body of the worm. In other tapeworms, the proglottids are retained on the strobila essentially throughout the life of the animal (anapolytic). In the former instance the proglottids may complete the production of eggs independent of the strobila; in the latter case proglottids on the strobila produce eggs that are released through uterine pores or tears. Degrees of apolysis are often recognized. Thus, cestodes that drop mature proglottids are considered to be euapolytic; those that drop immature proglottids are considered to be hyperapolytic. Proglottids filled with eggs are termed gravid. Consecutive proglottids either overlap (craspedote) or not (acraspedote). A very few species exhibit proglottids that, rather than being hermaphroditic, are either male or female. Because these proglottids are found on the same strobila, individuals of these species are considered to be hermaphroditic, but the proglottids themselves are not. There is, however, evidence of gonorchorism (separate sexes) in individuals of a few tapeworm species in which all of the proglottids of the strobila are either entirely male or entirely female.

With at least 751 recognized genera, the cestodes are the second most diverse group of platyhelminths. The most recent comprehensive systematic treatment of the group is that of Khalil *et al.* (1994), in which 14 orders were recognized. However, since then, the class has been revised to consist of 18 orders. The bothriate tapeworm orders are considered to be earlier divergent lineages than the acetabulate tapeworm orders (i.e., those with suckers or bothridia), but the relationships among the acetabulate orders remain controversial, as does membership in some of the elasmobranch-parasitizing orders.

The terms Cestoidea, Cestodaria, and Eucestoda are also sometimes used in reference to tapeworms. The term Cestoidea is synonymous with the term Cestoda; the authors have followed Hyman (1951) and Grassé (1961) in using the term Cestoda here. In the past, the amphilinideans and gyrocotylideans have been collectively referred to as the Cestodaria. However, evidence against the monophyly of this group is mounting, and, thus, the Cestodaria is not recognized here. The term Eucestoda is used to collectively refer to the cestodes other than the amphilinideans and gyrocotylideans. The monophyly of the Eucestoda is fairly well accepted; members of this subclass can be distinguished from the amphilinideans and gyrocotylideans, for example, in their possession of a scolex and also in that their sperm lack mitochondria.

To make the following diagnoses more concise, unless otherwise indicated, it should be assumed that an order possesses an anterior scolex of some form, lacks all vestiges of a digestive system, and possesses a strobila consisting of numerous proglottids that each contain one full set of both male and female reproductive organs (polyzoic). These proglottids contain numerous, compact testes filling much of the region anterior to the ovary, sperm lacking a mitochondrion, flagella that are incorporated into the main body of sperm, $9+'1'$ microtubule arrangement in the flagella of the sperm, a cirrus


53
UP
FP
MP
54
S
FP
MP
UP
R
55
SCO
BO
56
L
SCO
CGP
57
SCO
STR
58
RO
H
S
SCO
IMP
MPR
STR
U
GP
59
SCO
APO
BO
CPE
STR
CGP
60
SCO
BO
61
TE
PSC
SSC
SST
PST
62
PA
SCO
S
PB
63
AS
CPS
SCO
64
S
SCO
65
AS
S
SCO
66
SK
67
SCO
PRG
PRG
68
MA
BOD
SCO
ED
69
S
MUP
H
BOD
SCO
CPE
CGP
70
TE
BO
TS
SCO
MB

armed with microtriches, a cirrus-sac, a single posterior ovary, a vitellarium consisting of numerous follicles that extend in two lateral bands, one down each side of the proglottid, a Mehlis' gland located posterior to the ovarian bridge, a uterus that is sacciform, and male and female systems that open through a combined lateral genital pore, with pores of consecutive proglottids alternating irregularly down the length of the strobila.

Amphilinidea

See Figure 53 (**Plate 4**). Distinct scolex absent but with anterior coniform "sucker;" one set of male and female reproductive organs (monozoic) present; testes in two lateral bands; sperm with mitochondrion; cirrus-sac absent, ovary varying in shape; preformed uterine pore anterior; uterus often N-shaped or sinuous; seminal receptacle present; genital pores usually separate, posterior.

Habitat: Adults parasitic in body cavity of chondrosteans, some freshwater and marine teleosts, and Australian freshwater turtles.

Gyrocotylidea

See Figure 54 (**Plate 4**). Distinct scolex absent but with anterior "sucker," rosette usually present at posterior end of body; body usually "spined" (spines are likely to be microtriches), "spines" limited in distribution or not; excretory system reticulate; one set of male and female reproductive organs (monozoic) present; sperm with mitochondrion; cirrus-sac absent; ovary follicular, in V- or U-shaped band around seminal receptacle; seminal receptacle present; uterus sinuous, in median region of anterior part of body, terminal portion expanded into sac, opening into preformed ventral uterine pore; genital pores separate, both anterior.

Habitat: Adults parasitic in spiral intestine of holocephalans.

Subclass Eucestoda

Bothriocephalidea

See Figure 55 (**Plate 4**). Small to large tapeworms; scolex with two lateral bothria, occasionally with posterior appendages, occasionally divided into two or several vertical loculi, with or without muscular apical disc, occasionally replaced by scolex deformatus or peudoscolex; each bothrium usually unarmed, occasionally armed with one pair of hooks or apical series of hooks; proglottization usually complete; proglottids usually craspedote, usually wider than long; anapolytic, usually with one set of male and female reproductive organs per proglottid, rarely with two sets per proglottid; cirrus unarmed or armed; combined genital pore dorsal, median, sublateral or lateral, irregularly alternating; ovary posterior, medullary, mostly bilobed; vitelline follicles usually cortical, medullary or both, in two extensive lateral bands or circumsegmental; uterus variable, consisting of winding uterine duct and compact or diverticulate uterine sac; preformed uterine pore present or absent, if present opening posterior to genital pore.

Habitat: Adults parasitic in intestine of teleosts, rarely salamanders.

Caryophyllidea

See Figure 56 (**Plate 4**). Scolex usually with one to three pairs of weakly developed shallow depressions or grooves (considered to be loculi if temporary, bothria if more permanent), occasionally with terminal introvert or disc; with one set of male and female reproductive organs (monozoic); ovary H- or dumbbell-shaped; vitellarium usually preovarian, occasionally postovarian follicles also present; seminal receptacle present or absent; uterus often sinuous; female pore a combined utero-vaginal pore; genital pores separate or combined, median, ventral.

Habitat: Adults parasitic in intestine of siluriform and cypriniform freshwater fishes.

Cathetocephalidea

See Figure 57 (**Plate 4**). Large worms; scolex in form of single transversely expanded fleshy organ, lacking all signs of acetabula, bothria, and armature; usually with one strobila per scolex, rarely with multiple strobilae per scolex; proglottids acraspedote, euapolytic; with one set of male and female reproductive organs per proglottid; testes in single field; ovary H-shaped in dorso-ventral view; vitelline follicles circumsegmental; preformed uterine pore absent; vagina extending anterior to cirrus-sac; combined genital pore lateral; genital pores alternating irregularly along length strobila.

Habitat: Adults parasitic in spiral intestine of whaler sharks.

Cyclophyllidea

See Figure 58 (**Plate 4**). Scolex usually with four round suckers; suckers armed with spines or not; often with apical organ in form of single sucker (rarely four apical suckers), some with apical, often protrusible rostellum; rostellum armed with one or two rings of hooks or not, with or without rostellar pouch; scolex divided into scolex and metascolex, or with pseudoscolex in some; proglottids craspedote or acraspedote;

Plate 4 Cestoda (Figures 53–70) 53. Amphilinidea: adult of *Amphilinia* sp. (Gibson in Khalil, *et al.*, 1994); 54. Gyrocotylidea: adult of *Gyrocotyle* sp. (Gibson in Khalil, *et al.*, 1994); 55. Bothriocephalidea: adult of *Bathycestus brayi.* (Kuchta and Scholz, 2004); 56. Caryophyllidea: adult of *Paraglaridacris* sp. (Mackiewicz in Khalil, *et al.*, 1994); 57. Cathetocephalidea: adult of *Sanguilevator yearsleyi*; 58. Cyclophyllidea: adult of *Echinococcus* sp.; 59. Diphyllidea: adult of *Echinobothrium* sp.; 60. Diphyllobothriidea: adult of *Diphyllobothrium* sp.; 61. Haplobothriidea: adult of *Haplobothrium* sp.; 62. Lecanicephalidea: adult of *Corrugatocephalum* sp.; 63. Litobothriidea: adult of *Litobothrium daileyi*; 64. Nippotaeniidea: adult of *Nippotaenia* sp. (Bray in Khalil, *et al.*, 1994); 65. Proteocephalidea: adult of *Proteocephalus* sp.; 66. Rhinebothriidea: adult of *Rhinebothrium megacanthophallus*; 67. Spathebothriidea: adult of *Spathebothrium* sp.; 68. Tetrabothriidea: adult of *Tetrabothrius* sp.; 69. "Tetraphyllidea": adult of *Acanthobothrium* sp.; 70. Trypanorhyncha: adult of *Pterobothrium* sp. Abbreviations: APO, apical organ; AS, apical sucker; BO, bothrium; BOD, bothridium; CGP, common genital pore; CPE, cephalic peduncle; CPS, cruciform pseudosegments; FP, female genital pore; GP, gravid proglottid; H, hook; IMP, immature proglottid; L, loculus; MA, muscular appendage; MB, muscular bulb; MP, male genital pore; MPR, mature proglottid; MUP, muscular pad; PA, pars apicalis; PB, pars basalis; PRG, proglottid; PSC, primary scolex; PST, primary strobila; R, rosette; RO, rostellum; S, sucker; SCO, scolex; SK, stalk; SSC, secondary scolex; SST, secondary strobila; STR, strobila; TE, tentacle; TS, tentacle sheath; U, uterus; UP, uterine pore.

apolytic or anapolytic; usually with one but rarely two strobilae per scolex; usually with one but occasionally two sets of male and female reproductive organs per proglottid; strobila rarely with proglottids that are entirely either male or female, or male and female proglottids that alternate regularly along strobila; testes, one to many, usually anterior to ovary, occasionally posterior to ovary; cirrus-sac usually present; cirrus usually armed with spiniform microtriches, occasionally with stylet, occasionally with accessory sac; external and internal seminal vesicles present or absent; ovary dumbbell-shaped, or asymmetrical; vitellarium usually single compact mass, occasionally bilobed mass, usually posterior to ovary; vagina present or absent, replaced functionally by cirrus-sac and modified male duct, or replaced by supplementary ducts running between seminal receptacles in consecutive proglottids in some taxa; uterus ring-shaped, or saccate, with or without lateral branches (diverticula), or reticular, transverse in some, replaced by parauterine organ(s) in some, replaced by uterine capsules or parenchymatous egg capsules in some, uterus single or double, usually with preformed uterine pore; common genital pore lateral, rarely midventral; genital pores occasionally regularly alternating or unilateral along strobila; genital atrium conspicuous or not, with or without prominent radial musculature; vagina usually posterior to cirrus-sac, occasionally dorsal to ventral to cirrus-sac.

Habitat: Adults parasitic in intestine of amphibians, snakes, lizards, birds, or mammals.

Diphyllidea

See Figure 59 (**Plate 4**). Scolex with two bothria, proximal surfaces of bothria occasionally with spines; scolex with apical organ; apical organ usually armed with one row of large hooks on dorsal and ventral surfaces, often with continuous or interrupted row of smaller hooklets on lateral margins, occasionally with small spines; bothria supported by cephalic peduncle; cephalic peduncle often armed with eight columns of straight hooks with trifid bases; proglottids acraspedote; generally apolytic; testes usually filling proglottid anterior to ovary; ovary H-shaped in dorso-ventral view; lateral bands of follicles of vitellarium occasionally extending medially; preformed uterine pore absent; vagina posterior to cirrus-sac; common genital pore median, ventral.

Habitat: Adults parasitic in spiral intestine of rays, skates, and sharks of families Hemiscyllidae, Scyliorhinidae, and Triakidae (most marine).

Diphyllobothriidea

See Figure 60 (**Plate 4**). Medium to large tapeworms; scolex with two bothria (one dorsal and one ventral), occasionally with posterior appendages, occasionally divided into two or several vertical loculi, with or without muscular apical disc, unarmed; proglottization usually complete; proglottids craspedote or acraspedote, usually wider than long; anapolytic, usually with one set of male and female reproductive organs per proglottid, but occasionally multiple sets per proglottid; cirrus-sac usually thin-walled; cirrus unarmed; genital pore ventral, median, or submedian; ovary, posterior, medullary, mostly bilobed; vitelline follicles usually cortical and circumsegmental; uterus in form of spiral tube; preformed uterine pore opening posterior to genital pore.

Habitat: Adults parasitic in intestine of tetrapods, mostly mammals.

Haplobothriidea

See Figure 61 (**Plate 4**). Body bearing primary and secondary scoleces and strobila; primary scolex with four tentacles with spiniform microtriches at bases, each tentacle with muscular sac into which it can be withdrawn, bothria, bothridia and suckers absent; primarily strobila proglottized, but primary proglottids never develop reproductive organs on primary strobila; proglottized regions of primary strobila separate from remainder of primary strobila and develop into secondary strobilae that produce proglottids, each proglottid bearing one set of reproductive organs; anterior-most proglottid of each secondary strobila modified as secondary scolex; secondary scolex with four shallow anterior indentations surrounding central, apical dome; cirrus-sac median, anterior, cirrus armed with microtriches; testes numerous, arranged in two lateral fields; ovary inverted U-shaped; follicles of vitellarium arranged in two lateral fields, confluent in midline at anterior and posterior regions of proglottid; uterus with dilated uterine sac; preformed uterine pore present or absent; vagina posterior to cirruss-sac; seminal receptacle present; genital pores separate, median, anterior; eggs operculate.

Habitat: Adults parasitic in intestine of the bowfin, *Amia calva*.

Lecanicephalidea

See Figure 62 (**Plate 4**). Scolex usually divided into two parts: if present an anterior region (pars apicalis) variable in form, elongate (myzorhynchus), pad-like (metaporhynchus), sucker-like, or consisting of a crown of nonhooked tentacles; posterior region (pars basalis) usually cushion-like, often bearing four circular suckers or bothridia; proglottids, craspedote or acraspedote, usually apolytic; testes, three to many, arranged in a single or multiple columns; external seminal vesicle usually present, often conspicuous, extending to near posterior margin of proglottid; ovary bilobed, multilobed or irregular in dorso-ventral view; follicles of vitellarium sometimes encircling proglottid; preformed uterine pore absent; vagina present or absent, almost always posterior to cirrus-sac; cirrus armed or not; combined genital pore lateral; genital pores alternating irregularly along length strobila.

Habitat: Adults parasitic in spiral intestine of marine elasmobranchs, primarily rays.

Litobothriidea

See Figure 63 (**Plate 4**). Worms small; scolex with cup-shaped apical sucker and three to five pseudosegments, some pseudosegments cruciform, with or without posterior spine-like structures; proglottids slightly craspedote, euapolytic; testes arranged in two to multiple columns; cirrus armed; ovary inverted U-shaped in dorso-ventral view; vitelline follicles encircling proglottid; preformed uterine pore absent; vagina opening at same level as cirrus-sac; combined genital pore lateral; genital pores alternating irregularly along length strobila.

Habitat: Adults parasitic in spiral intestine of marine lamniform sharks.

Nippotaeniidea

See Figure 64 (Plate 4). Scolex in form of single apical sucker; proglottids acraspedote, anapolytic or apolytic; testes restricted to previtelline medulla of proglottid; cirrus-sac thin-walled; convoluted ejaculatory duct present; seminal vesicle present or absent; ovary dumbbell-shaped in dorso-ventral view, vitellarium consisting of two symmetrical compact masses, anterior to ovary; uterus in form of transverse coils in medulla, lacking preformed uterine pore; vagina opening posterior to cirrus-sac.

Habitat: Adults parasitic in digestive system of freshwater teleosts.

Proteocephalidea

See Figure 65 (Plate 4). Scolex usually with four round suckers or bothridia; suckers or bothridia uni-, bi-, tri-, or tetra-loculate; apex of scolex sometimes with armed rostellum, or large sucker, or occasionally with apical organ in form of two lappets; metascolex sometimes present as wrinkled region posterior to suckers that may or may not cover suckers of scolex; scolex with or without gland cells; proglottids craspedote or acraspedote, usually anapolytic; longitudinal muscle bundles usually conspicuous, occasionally inconspicuous; ovary usually dumbbell-shaped in dorso-ventral view; uterus usually with lateral branches (diverticula) when gravid, opening through one or more longitudinal apertures; vagina usually anterior to, but occasionally posterior to, cirrus-sac; genital pores rarely separate, lateral, alternating irregularly along length strobila; distinction between cortex and medulla usually marked.

Habitat: Adults parasitic in intestine of freshwater teleosts, occasionally in amphibians, terrestrial lizards (such as, e.g., *Varanus*, monitor lizards) and snakes.

Rhinebothriidea

See Figure 66 (Plate 4). Small to medium worms; scolex with four stalked, unarmed bothridia with facial septa; with or without myzorhynchus and marginal septa; proglottids craspedote or acraspedote, generally euapolytic; post-poral testes usually lacking; cirrus armed; ovary H-shaped in dorso-ventral view; vitelline follicles usually in two lateral fields, occasionally encroaching on median line; preformed uterine pore absent; vagina opening anterior to cirrus-sac; combined genital pore lateral; genital pores alternating irregularly along length strobila.

Habitat: Adults parasitic in spiral intestine of marine and freshwater batoid elasmobranchs.

Spathebothriidea

See Figure 67 (Plate 4). Scolex usually bearing one or two sucker-like attachment organs, but occasionally absent; osmoregulatory system in form of canals on each side of strobila, with irregular commissures and lateral pores; body proglottization absent; with multiple sets of male and female reproductive organs serially repeated down body (polyzoic); testes in two irregular lateral bands that extend throughout the length of the body; ovary rosette-shaped or bilobed, median, posterior to genital pores; small seminal receptacle occasionally present; uterus often sinuous, with median pore; male and female pores separate, opening close to one another along median line, usually irregularly alternating in adjacent proglottids between opening dorsally and ventrally.

Habitat: Adults parasitic in intestine of freshwater, brackish water and marine chondrosteans and teleosts.

Tetrabothriidea

See Figure 68 (Plate 4). Scolex usually with four muscular bothridia; bothridia round, rectangular or occasionally triangular in form, often with auriculate muscular appendages; auricles sometimes fused to form apical "organ;" apical rostellum present or absent; proglottids craspedote; anapolytic; testes, few to many; ovary bilobed or highly lobed in dorso-ventral view; vitellarium in form of single compact mass, usually anterior to ovarian bridge; uterus transverse, with one or more dorsal pores; seminal receptacle present or absent; vagina occasionally armed; genital pores combined or not; genital atrium often muscular, sometimes with distinct male atrial canal, or genital papilla, occasionally armed; genital pores usually unilateral along length of strobila; vagina usually posterior to cirrus-sac.

Habitat: Adults parasitic in intestine of seabirds, pinnipeds, and cetaceans.

"Tetraphyllidea"

See Figure 69 (Plate 4). Scolex divided into four muscular sessile or pedunculated bothridia of a variety of forms; bothridia with or without apical muscular pad, with or without one or two pairs of uni, bi, or tripronged hooks, may or may not be further subdivided into loculi; muscular pad with or without one or three accessory suckers; metascolex occasionally present; proglottids craspedote or acraspedote, generally euapolytic; rarely with proglottids that are either male or female (or female proglottids retain male copulatory apparatus); ovary usually H-shaped, occasionally A-shaped in dorso-ventral view; vitelline follicles sometimes completely encircling proglottid, rarely condensed around ovary; uterine pores sometimes present, but not preformed, or uterus opening through median slits in uterine wall; vagina opening anterior to cirrus-sac; genital pores rarely separate, alternating irregularly or rarely unilateral along length of strobila. The order is almost certainly not monophyletic as currently configured.

Habitat: Adults parasitic in spiral intestine of marine and freshwater elasmobranchs or very rarely holocephalans.

Trypanorhyncha

See Figure 70 (Plate 4). Scolex with four retractable, hollow tentacles each armed with numerous spiral rows of hooks (hook patterns of enormous taxonomic importance); each tentacle everted and retracted by rhyncheal apparatus consisting of tentacle sheath, and retractor muscle originating within posterior muscular bulb (tentacles and rhyncheal apparatus rarely absent); scolex with two or four bothria and elongated peduncle housing rhyncheal apparatus; proglottids craspedote or acraspedote, apolytic or anapolytic; usually with one set of male and female reproductive organs per segment, but occasionally with two sets; cirrus-sac or hermaphroditic-sac usually present; external, internal and accessory seminal vesicles present or absent; ovary usually H-shaped in dorso-ventral view; follicles of vitellarium lateral or completely

encircling proglottid; uterus tubular or branched, with or without preformed uterine pore; vagina usually opening posterior to cirrus-sac, occasionally dorsal or ventral to cirrus-sac; common genital pore often associated with muscular lips.

Habitat: Adults parasitic in spiral intestine, stomach, rarely gall bladder, of elasmobranchs.

Nervous System

A bewildering variety of neural organization and complexity characterizes the platyhelminths (e.g., see Halton and Gustafsson, 1996). The basic plan involves an apical subepidermal brain consisting of multiple neurons, which can be uni, bi (especially in Neodermata), or multilobed. This is a true brain because it controls reflexes in the peripheral nerve net. The number of nerve cells (neurons) in the brain varies between 50 and 550 in free-living species, but this number is less well known for the neodermatans. The brain is generally connected to two, or sometimes more, main longitudinal nerve cords, each consisting of axons with cell bodies distributed at irregular intervals along the length of the cords. These main nerve cords are usually connected to one another by numerous transverse commissures in a ladder-like arrangement. In addition, a peripheral array of nerve plexuses (or nerve-nets) is found throughout the body. The attachment organs and the components of the female reproductive system involved in egg formation are especially well supplied with nerve plexuses, but plexuses are also found associated with the pharynx and intestine and below the surface and below the muscle layers of the body. There are separate plexuses for sensory, integrative and motor activities. The central nervous system (brain plus main cords) is more emphasized in neodermatan groups than in non-neodermatan groups. Cestodes may exhibit regional differences in the complexity of the subtegumental plexuses along the length of the body. Transmission electron microscopy suggests that platyhelminths exhibit an extensive diversity of types of sensory receptors.

Several different neuronal cell types are present, including unipolar, bipolar, and multipolar cells. Although unipolar neurons are generally confined to the ganglia of the brain and longitudinal nerve cords, cells of plexuses are dominated by bipolar cells. Individual neurons are highly secretory and contain a variety of types of vesicles in their cytoplasm. Platyhelminths utilize a diversity of signaling molecules and at least two different groups of neuropeptides. The nervous systems of platyhelminths appear to possess compounds similar to all of the major neurotransmitters known in taxa as complex as vertebrates. In addition, there are two groups of neuropeptides that appear to be unique to flatworms: neuropeptide F and FMRPamide-related peptides (FaRPs) both of which are fairly extensively distributed among the various platyhelminth groups. The existence of glial or glial-like cells in platyhelminths remains somewhat controversial.

Feeding and Digestion

Other than a few fecampiids, dalyelloids, and the cestodes, flatworms generally have a digestive system divisible into a foregut and a caecum. The foregut leads from the mouth to the blind-ended sac-like caecum that is lined with a thin gastrodermis. Free-living species tend to be microphagous or predatory and feed on bacteria, unicellular alga, protists, and almost all types of invertebrates. Characteristically, the free-living species have a highly muscular suctorial pharynx. Depending on the species, the pharynx can be protruded, inserted, applied to, extended over, or used to envelop and swallow whole the various foodstuffs they actively seek or encounter. Food passes from the mouth *via* the pharynx to the esophagus and then to the caecum. The gastrodermis has glandular and phagocytic cells that digest and incorporate the nutrients. A number of free-living platyhelminths (e.g., prolecithophorans, rhabdocoels, polyclads) harbor endosymbiotic algae that serve as an additional source of food. The solid waste products of free-living species are often released *via* a genitointestinal canal, rather than a separate anus; in some groups (e.g., polyclads), anal pores assist with this function.

Symbiotic and parasitic platyhelminths vary their feeding habits according to the hosts on, or in, which they live. Groups that possess suckers may have numerous secretory cells and sensory structures surrounding the mouth and sucker. Although similar in basic design to the foregut of free-living groups, the parasitic groups tend to exhibit distinct cell type configurations in the gastrodermis. In these groups, the junction between the foregut and caecum is invariably abrupt. Aspidogastreans have a single cell type in their gastrodermis that alternates cyclically between a secretory–absorptive phase and an autophagic–exocytotic phase. Digeneans have a syncytial or cellular gastrodermis with only one cell type, and digestion takes place extracellularly in the lumen of the caecum. There is evidence, however, that some digeneans are capable of acquiring nutrients directly through microvillar extensions of their body surfaces. In contrast, digestion in the caecum of monogeneans is intracellular. The polyopisthocotylean monogeneans are blood-feeders. The nonpigmented cells in these taxa are thought to support and protect the pigmented digestive cells in the gastrodermis. In addition, at least some of these taxa possess a buccoesophageal canal between the buccal cavity and the esophagus that permits regurgitation of intestinal contents. The monopisthocotyleans feed on the epidermal cells and mucus of their host and have only a single gastrodermal cell type involved with endocytosis and digestion.

In the Cestoda and some symbiotic fecampiids all vestiges of a digestive system have been lost. These species rely solely on acquiring nutrients through their outer body wall. Among the Cestoda, and indeed all other neodermatans, the neodermis plays an active role in the biochemical transport of nutrients. In cestodes the surface area of the neodermis is markedly increased by the presence of microtriches, which enhance the availability of readily available nutrients to be actively absorbed into the parasite's body.

Excretion and Osmoregulation

The excretory system of platyhelminths is thought to function primarily in osmoregulation, with the removal of metabolic

wastes generally occuring by way of diffusion through the outer body layer. However, removal of excess water can also be achieved through the gut, and in free-living species it is not uncommon to find excretory products stored in various tissues. In platyhelminths, this system is made up of a series of protonephridia and is thus commonly referred to as a protonephridial system. It consists of a series of protonephridia connected to collecting ducts that empty into one or more major canals termed nephridioducts; these either open to the outside of the body through one or more nephridiopores or empty into a common excretory vesicle which opens to the outside through a common nephridiopore. All protonephridia consist of at least one flame bulb bearing two to many cilia, and a canal cell and sometimes also a cap cell. Rohde (1991) recognized three distinct types of protonephridia. In Type 1, found only in catenulids, each flame bulb has a cap cell, lacks a canal cell, and bears only two cilia, the rootlets of which extend along the collecting weir. In Type 2, found in most platyhelminths and also many other phyla including nemertines and gnathosomulids, each flame bulb has a cap cell and multiple cilia, the rootlets of which extend into the cap cell; each flame bulb is also associated with a canal cell that it may share with other flame bulbs. In Type 3, found only in kalyptorhynchs, "dalyellioids," temnocephalids, and potentially also "typhlophanoids" and the "lecithoepitheliatans," each flame bulb bears multiple cilia, but lacks a cap cell; each flame bulb is also associated with a canal cell that it usually shares with other flame bulbs. In all cases, the flame bulb is associated with a filtering apparatus, called a weir, which consists of processes of the cap and canal cell that are surrounded by a membrane. The beating of the cilia in the flame bulb draws body fluids across the membrane of the weir into the collecting ducts. One or two sets of protonephridia are present in all platyhelminths except a few catenulids in which excretion is thought to take place solely through the digestive syncytium. The protonephridial system is most developed in freshwater species. Indeed, this system is considered to have played a major role in allowing platyhelminths to invade freshwater habitats.

The protonephridial system of most platyhelminth taxa is paired (but the pairs are sometimes connected medially), consisting of two nephridioducts, one running along the lateral margin of each side of the body; these ducts connect to many protonephridia throughout their length. Some of the catenulids are exceptional in their possession of a single dorsal nephridioduct that opens through a single posterior nephridiopore; some proseriatans also possess a single set of protonephridia, but some possess three sets of protonephridia. In most of the non-neodermatan groups, the two nephridioducts open to the outside of the body through separate nephridiopores that are found in the anterior half of the body, generally on the ventral surface. In most temnocephalideans, the two nephridioducts open into separate, saccate excretory vesicles, located in the anterior of the body on either side of the pharynx, and the left and right sets of protonephridia are generally connected by two transverse ducts, one at the anterior and one at the posterior aspect of the system. The protonephridial system of monogeneans is similar to that of temnocephalidans, but the anterior transverse duct is absent in the polyopisthocotyleans.

The aspidogastreans possess a protonephridial system that is also much like that of the temnocephalidans but without the transverse ducts. The protonephridial system of digeneans differs somewhat from that of the aspidogastreans in that the two nephridioducts generally open into a common posterior excretory vesicle. This vesicle takes on a number of different forms in different taxa; for example, it can be V-, Y-, or I-shaped. Unlike other platyhelminth taxa, the arrangement of the protonephridia relative to the nephridioducts is regular and consistent in the cercarial stage of digeneans among individuals of the same species, but differs among individuals of different taxa, and thus this "flame cell formula" is often of significant taxonomic value. The larger number of flame bulbs and the difficulty of seeing most of the elements of the protonephridial system in the more robust bodies of the adult digeneans reduce the taxonomic utility of the flame bulb arrangement in adults.

Among tapeworms, the amphilinideans possess a protonephridial system consisting of a network of collecting ducts, rather than just two lateral nephridioducts. This network empties into an elongated excretory vesicle that opens to the outside through a posterior nephridiopore. The flame bulbs are usually found in pairs, grouped in clusters throughout the length of the body. The gyrocotylidean system also includes two longitudinal nephridioducts that are connected to one another by a number of secondary canals. The main nephridioducts open into a nephridiopore on the dorsal surface in the anterior half of the body. In most of the other cestode groups the protonephridial system consists of two dorsal and two ventral longitudinal nephridioducts, the ventral ducts generally being larger in diameter than the dorsal ducts; the nephridioducts are connected to multiple flame bulbs throughout the length of the body. Some cestodes do, however, have up to 12 longitudinal ducts. The pairs of dorsal and ventral nephridioducts are connected to one another in the scolex, where they may undergo extensive winding. The left and right nephridioducts are often connected to one another by way of transverse excretory ducts located in the posterior end of each proglottid; this condition is most commonly found in anapolytic eucestode species. In some groups, the longitudinal vessels are connected to one another with numerous lateral nephridioducts, resulting in a network-like arrangement of vessels. In the first proglottid formed (located at the posterior end of the strobila) the ventral pair of nephridioducts often empty into a small vesicle at the posterior end of this proglottid. However, once this proglottid drops from the strobila, the ventral vessels open to the outside separately in the posterior end of the terminal proglottid of the strobila. Rohde (1991) provides a useful overview of the evolution of the protonephridial system in platyhelminths.

Reproduction

Most platyhelminth species undergo sexual reproduction. Given the predominance of hermaphroditism in this phylum, there is some question about the relative frequency and importance of self- versus cross-fertilization in most species. In the few taxa in which this issue has been investigated in detail, cross- rather than self-fertilization appears to be the dominant

form of sexual reproduction. Gonochorism does, however, occur in some triclads, rarely in prolecithophorans, in a few fecampiids (e.g., *Kronborgia*), within the digenean family Schistosomatidae, and in a few tetraphyllidean and cyclophyllidean cestodes. To date, no instances of gonochorism in aspidogastreans or monogeneans are known. Most platyhelminths are simultaneous hermaphrodites, each individual possessing a full complement of male and female reproductive systems. There is a general tendency for the male system to mature before that of the female (protandry). Sperm transfer usually occurs during copulation. But, in a few species of macrostomids, catenulids, triclads, polyclads, rhabdocoels, polyopisthocotylean monogeneans, and cestodes, sperm is injected directly through the body wall and then migrates through the body to the female reproductive organs. Fertilization is always internal.

Asexual multiplication in one form or another is also found in a number of platyhelminths, either in the adult or, most commonly in neodermatans, in one or more of the larval stages. Adults of many of the catenulids and macrostomids undergo forms of asexual multiplication either as paratomy (regeneration before fission) or as architomy (regeneration after fission). These phenomena result in a linear chain of individuals that eventually separate and each develop into a mature worm.

Some adult triclads possess the ability to regenerate complete individuals from, in some cases, a very small portion of the body of the original individual. For example, it is possible to divide an adult freshwater triclad posterior to the pharynx and produce two complete adult individuals. More dramatic is the ability of adults of species belonging to the triclad genus *Phagocata* to divide into a number of separate fragments, each of which subsequently encysts and develops into a complete adult worm. Their ability to regenerate has made freshwater triclads favored animals for studies in wound healing and developmental genetics. Asexual reproduction in the form of polyembryony occurs in a diversity of free-living groups such as the triclads, rhabdocoels, "lecithoepitheliatans," and rhabdocoels.

With the exception of the aspidogastreans, asexual multiplication is found in at least some members of all of the neodermatan groups. Although it is uncommon among the monogeneans, it is considered to be present in the form of sequential polyembryony (hyperviviparity) in some members of the order Gyrodactylidea. Rather than releasing eggs containing oncomiracidia, an adult can produce a second generation juvenile *in utero* (without sexual reproduction) that in turn can produce a third generation juvenile *in utero*, until one single individual may yield in excess of 20 individuals of the subsequent generation from within its own body, apparently without the intervention of a second individual or of gametes. As a result, the body of such asexually multiplying individuals in effect resembles a series of Russian dolls. There is evidence, however, of sperm in the seminal receptacle of some of these embryos, which has led some to question this as a truly asexual form of multiplication.

Digeneans typically undergo asexual multiplication in the larval stages that parasitize the first, usually molluscan, intermediate host. All digeneans possess one or more generations of sporocysts or one or more generations of rediae, or one or more generations of both sporocysts and rediae. Asexual multiplication occurs in the generation of all larval stages developing within any generation (mother or daughter) of both sporocysts and redia. This form of asexual multiplication, termed polyembryony, results in multiple genetically identical individuals of the subsequent larval stage that develop within the saccate body of the parent generation. Thus, first generation sporocysts (mother sporocyts) produce either multiple second generation sporocysts (daughter sporocyts) or multiple first generation redia (mother redia), which in turn produce multiple second generation redia (mother redia) or multiple second generation redia (daughter redia), respectively. Second generation sporocysts produce either multiple cercariae or multiple redia that then subsequently produce multiple cercariae.

In tapeworms, asexual multiplication occurs in the terrestrial members of the order Cyclophyllidea that develop through larval stages known as coeneri and hydatid cysts. The fluid-filled cavity of these larvae is lined with a layer of germinative tissue that produces multiple protoscoleces, in the case of coeneri, and daughter hydatids that in turn are lined with germinative tissue that produces multiple protoscoleces, in the case of hydatid cysts. Each protoscolex has the potential to become an adult worm. The reproductive potential in these groups is enormous, with the hydatid cysts of some species producing hundreds of thousands, perhaps even millions, of protoscoleces.

Ontogeny

There are a remarkable variety of larval forms found among the platyhelminths, especially within and among the classes of neodermatans. In general, the life cycle of each class of neodermatan (see Figures 71–73) is characterized by one or more unique larval stages (**Plate 5**). This diversity reflects the great

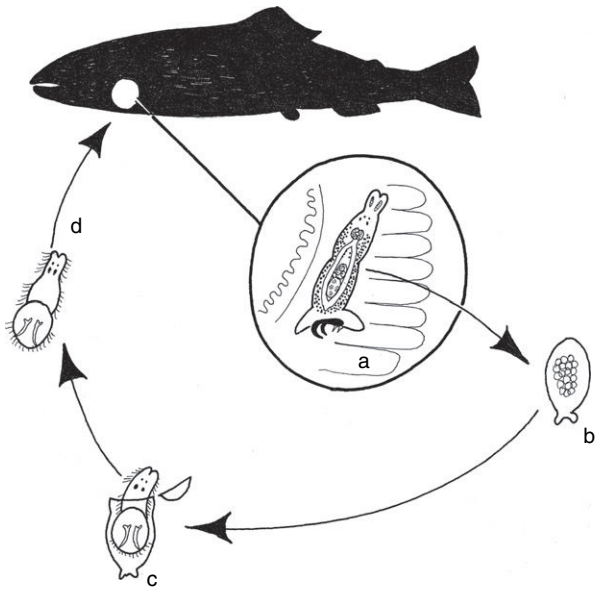


Figure 71 Generalized monogenean life cycle. (a) Adult on gills of definitive host (usually an aquatic vertebrate), undergoes sexual reproduction, and produces eggs; (b) eggs released into water; (c) oncomiracidium hatches from egg; (d) oncomiracidium finds definitive host.

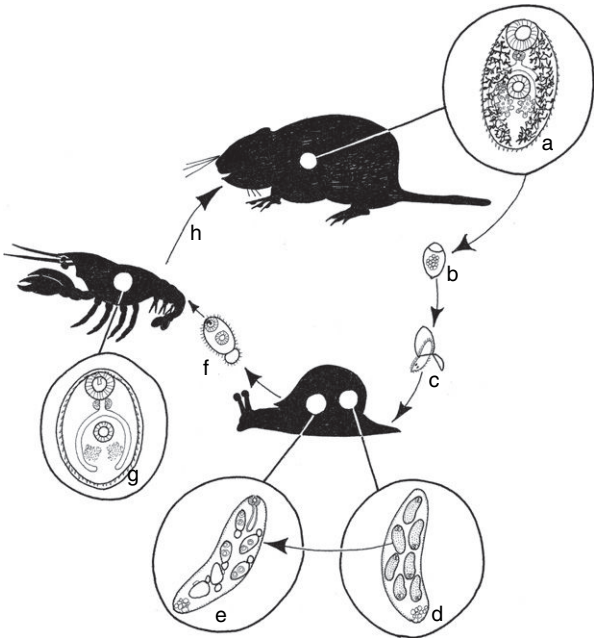


Figure 72 Generalized digenean life cycle. (a) Adult in definitive host, undergoes sexual reproduction and produces eggs; (b) eggs released with feces of host; (c) miracidium hatches from egg and penetrates first intermediate host (usually a mollusc); (d) miracidium develops into sporocyst containing numerous redia; (e) redia leave sporocyst and produce numerous cercariae; (f) cercaria escapes from first host and finds second intermediate host; (g) cercaria sheds tail, encysts, and develops into metacercaria; (h) second intermediate host containing metacercaria is eaten by definitive host.

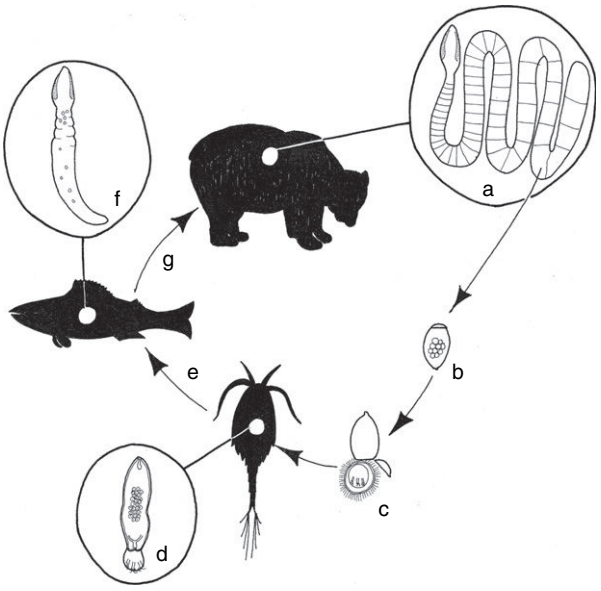


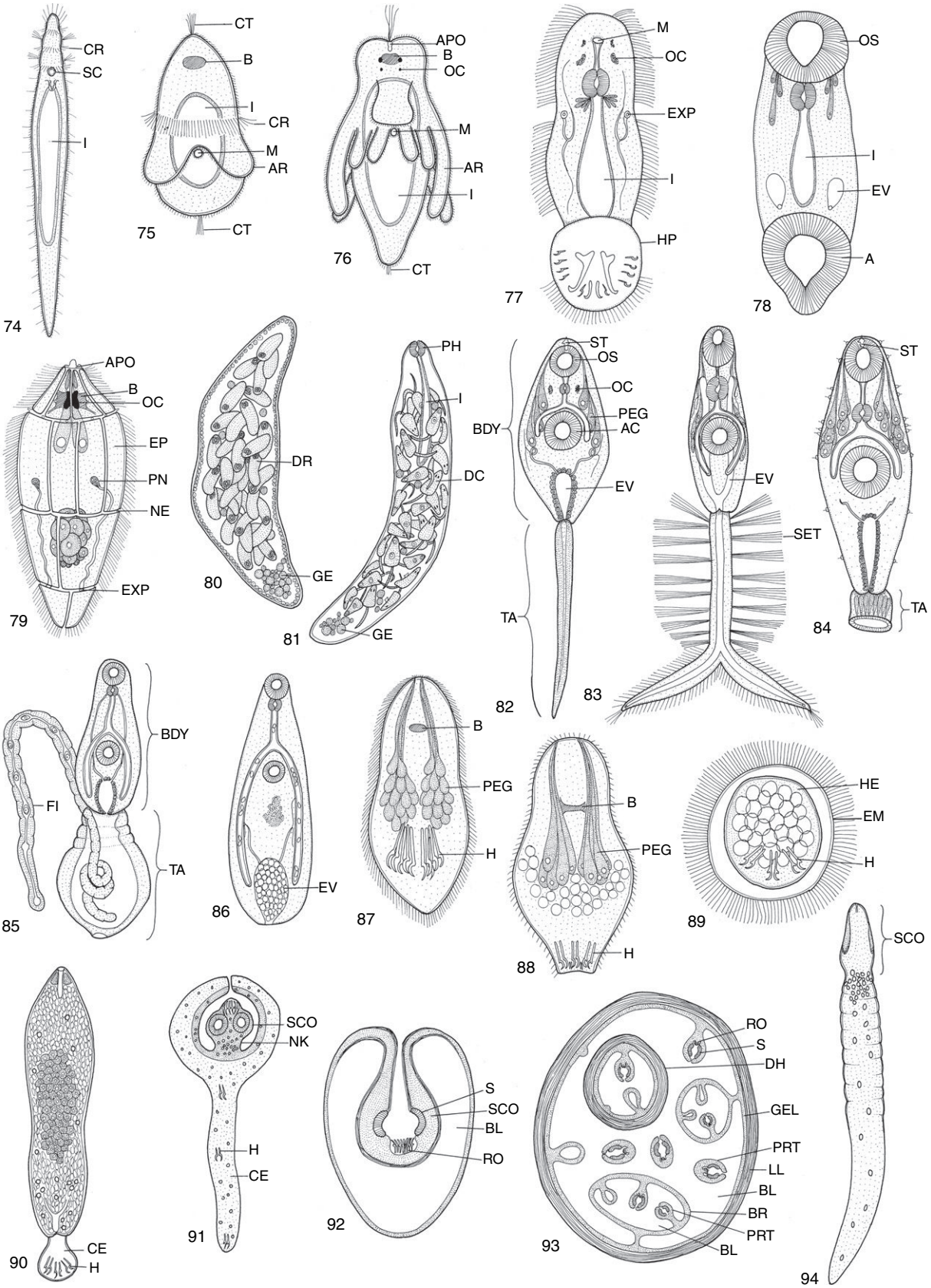
Figure 73 Generalized life cycle of aquatic cestode. (a) Adult in definitive host, undergoes sexual reproduction, and produces eggs; (b) eggs released with faeces of host; (c) coracidium larva hatches from egg and penetrates first intermediate host (usually a copepod); (d) coracidium develops into procercoid; (e) first host eaten by second intermediate host; (f) procercoid develops into plerocercoid; (g) second host eaten by definitive host (usually a piscivorous mammal).

range of habitats invaded and life history strategies employed. Most non-neodermatans undergo direct development such that the zygote develops directly into a form resembling a young adult. As discussed by Ruppert (1978), however, there are several notable exceptions to this pattern. One species of catenulid develops from a zygote into a pelagic larval stage known as a Luther's larva (Figure 74, **Plate 5**); the larva is vermiform in shape, ciliated throughout, but in addition possesses several anterior rings of elongate cilia that aid in swimming. The zygotes of some polyclads develop into free-swimming, ciliated larval stages bearing arms. These larvae are generally referred to as Götte's larvae (Figure 75, **Plate 5**) or Müller's larvae (Figure 76, **Plate 5**) depending on whether they possess four or seven to ten arms, respectively. This variation in developmental pattern is seen among individuals of some polyclad species. Götte's larvae are found only in members of the polyclad suborder Acotylea; Müller's larvae are known from members of both suborders of polyclads. There does not seem to be a phylogenetic pattern to the distribution of these larval stages among the polyclads. For example, the polyclad genus *Stylochus* includes species that develop into Götte's larvae, species that develop into Müller's larvae, and species that undergo direct development.

In general, the monogenean life cycle consists of two stages, the adult and a free-swimming larva, which is usually ciliated, known as the oncomiracidium. This larva usually possesses one or two pairs of ocelli, and is conspicuous in its possession of a posterior attachment structure known as a haptor, which is usually armed with hooklets (uncinuli) and sometimes also with hamuli (Figure 77, **Plate 5**). The oncomiracidium differs from the ciliated larvae of other neodermatans in that its cilia are often restricted to three distinct regions. When the oncomiracidium hatches from an egg, which is sometimes anchored to the substrate, it swims or creeps to find its vertebrate host. An oncomiracidial stage is mostly lacking from monogeneans such as the gyrodactylideans, which are unusually in that they are generally viviparous and undergo sequential polyembryony. A generalized oviparous monogenean life cycle is shown in **Figure 71**.

Prior to developing into the adult form, aspidogastreans pass through a larval stage called a cotylocidium (Figure 78, **Plate 5**), which is often but not always ciliated. This larval stage is free-swimming in at least some species. However, the life cycles of many species are not yet fully understood. In some cases it appears that the cotylocidium hatches from the egg and parasitizes the same host as the parent individual from which it was released. In other cases it appears that the cotylocidium hatches from the egg and moves into the water column to find a new host. In yet other cases the cotylocidium never hatches from the egg; rather, the egg is ingested by the molluscan host. In some species the life cycle is direct, involving only a single host. In other species, the life cycle may involve a second host which may be infected by ingesting infected molluscs.

Digeneans undergo complex development in which certain life cycle stages are able to produce multiple copies of new generations before becoming sexually reproductive adults. Consequently the term "larva" is used loosely when discussing life cycle stages in the Digenea. The digeneans pass through a series of larval stages that are either parasitic or free-living,


74
CR
SC
I
75
CT
B
I
CR
M
AR
CT
76
APO
B
OC
M
AR
I
CT
77
M
OC
EXP
I
HP
78
OS
I
EV
A
79
APO
B
OC
EP
PN
NE
EXP
80
DR
GE
81
PH
I
DC
GE
82
ST
OS
OC
PEG
AC
BDY
EV
TA
83
EV
SET
84
ST
TA
85
BDY
FI
TA
86
EV
87
B
PEG
H
88
B
PEG
H
89
HE
EM
H
90
CE
H
91
SCO
NK
H
CE
92
S
SCO
BL
RO
93
RO
S
DH
GEL
PRT
LL
BL
BR
PRT
BL
94
SCO

depending on their association with this sequence of hosts. A generalized digenean life cycle is illustrated in **Figure 72**. As can be seen, the basic sequence of larvae is as follows: miracidium, sporocyst, redia, cercaria, and metacercaria. The miracidium (Figure 79, **Plate 5**) possesses cilia on epidermal plates, or occasionally on tufts or bars. The neodermis is formed when the miracidium sheds its epidermal cell plates, thus all stages following the miracidium have an outer layer that is a neodermis. In most taxa the miracidium develops into a sporocyst, but it occasionally may develop into a redia. Sporocysts and redia are both sacciform. The main difference between them is that the sporocyst has none of the components of a digestive system (Figure 80, **Plate 5**), whereas the redia usually possesses at least a mouth, a pharynx, and often a single intestinal caecum (Figure 81, **Plate 5**). These stages vary morphologically among species in their possession of lateral lobes, and in some cases birth pores; the sporocysts of some species are branched. It is quite common for digeneans to possess more than a single sporocyt or redial generation in their life cycle. A first generation sporocyst (mother sporocyst) is essentially a miracidium that has lost its epidermal plates, whereas the second generation sporocysts (daughter sporocysts) are produced through asexual multiplication within the primary sporocyst. A first generation redia (mother redia) is most commonly produced through asexual multiplication within a primary or secondary sporocyst, but exceptionally develops directly from a miracidium that has shed its epidermal plates. Second generation rediae (daughter rediae) are produced through asexual multiplication within the primary redia). In most taxa, the final result of several generations of sporocysts or redia is the development of a number of cercariae. Each cercaria generally possesses an anterior body and a posterior tail. The cercaria is by far the most morphologically diverse of the digenean larval forms. Indeed, classification of the digenean families is based to a large extent on cercarial morphology. An elaborate nomenclature based on morphology exists to describe these cercarial types (e.g., see Cable, 1956). Examples of some of the more distinctive types include ophthalmoxiphidiocercaria (Figure 82, **Plate 5**), furcocercous (Figure 83, **Plate 5**), microcercous (Figure 84, **Plate 5**), and cystocercous cercariae (Figure 85, **Plate 5**). The metacercarial stage is usually achieved when the cercaria drops its tail and secretes a protective cyst around itself. The metacercaria is essentially a juvenile version of the adult worm (Figure 86, **Plate 5**). However, variations on this general developmental pattern occur: the metacercarial stage may be lacking or it may be preceded by an unencysted mesocercarial stage.

Associations of these larval forms with their hosts are quite predictable. In general, the miracidium and cercaria are free-living forms. The miracidium hatches from the egg and seeks the first (molluscan) host; the cercaria escapes from the mollusc and seeks either the second intermediate or the definitive host. The sporocyst and redial stages generally parasitize the first (molluscan) host. The metacercarial stage encysts either on aquatic vegetation or in or on the tissue of a second intermediate host, which can be either an invertebrate or vertebrate, depending on the digenean species. Exceptions to this general pattern do occur. In some species, the miracidium is not free-swimming; rather, it remains within the egg, which is subsequently eaten by the first host. In some species, the cercaria does not leave the mollusc to swim freely; rather, it continues its development into a metacercaria, encysting within the body of the sporocyst or redia within the molluscan host. In species that lack metacercaria, such as the schistosomes, the cercaria penetrates the tissues of the definitive host directly. This stage is referred to as a schistosomula once it drops its tail and enters the host.

Both amphilinideans and gyrocotylideans possess a ciliated larval stage known as a lycophore larva (Figure 87 and Figure 88, respectively; **Plate 5**). This larva possesses 10 hooks at the posterior end of its body; these hooks are of similar size and shape in gyrocotylideans but consist of six large and four small hooks, which also differ in shape, in the amphilinideans. This larval stage is also sometimes referred to as a decacanth. The life histories of many species in these two cestode groups are not well known. It appears that, in at least some amphilinideans, the lycophore does not escape from the egg until it is eaten by a small crustacean. In this intermediate host, the lycophore burrows through the gut and comes to rest in the hemocoel, where it develops into a procercoid by shedding its ciliated epidermal cells, elongating, and developing a posterior tail. It subsequently develops into a plerocercoid by shedding its tail and undergoing further elongation. In another case, the lycophore larva hatches and actively penetrates through the cuticle of the intermediate crayfish host. Development into the adult stage occurs only when the intermediate host is eaten by an appropriate vertebrate, at which point the plerocercoid sheds its tail and becomes sexually mature. It is possible that gyrocotylideans do not require an intermediate host to complete their development.

Most eucestodes require at least two hosts to complete their life cycles, but it is also common for additional paratenic hosts to occur in the life cycles of these platyhelminths. Paratenic hosts are not required, but serve to bridge food chain gaps,

Plate 5 Larval stages (Figures 74–94) 74. Catenulida: Luther's larva (Ruppert, 1977); 75. Polycladida: Götte's larva (Ruppert, 1977); 76. Polycladida: Müller's larva (Ruppert, 1977); Figure 77. Monogenea: oncomiracidium; 78. Aspidogastrea: cotylocidum; 79. Digenea: miracidium; 80. Digenea: sporocyst; 81. Digenea: redia; 82. Digenea: ophthalmoxiphidiocercaria; 83. Digenea: furcocercous cercaria; 84. Digenea: microcercous cercaria; 85. Digenea: cystocercous cercaria; 86. Digenea: unencysted metacercaria; 87. Cestoda: gyrocotylidean lycophore larva (Beauchamp in Grassé, 1961); 88. Cestoda: amphilinidean lycophore larva (deBeauchamp in Grassé, 1961); 89. Cestoda: coracidium; 90. Cestoda: procercoid (Joyeux and Baer in Grassé, 1961); 91. Cestoda: cysticercoid (scolex retracted); 92. Cestoda: cysticercus (scolex retracted and invaginated); 93. Cestoda: hydatid cyst 94. Cestoda: plerocercoid. Abbreviations: A, adhesive disc; AC, acetabulum; APO, apical organ; AR, arm; B, brain; BDY, body; BL, bladder; BR, brood sac; CE, cercomer; CR, ciliary ring; CT, ciliary tuft; DC, developing cercaria; DH, daughter hydatid cyst; DR, developing redia; EM, embryophore; EP, epidermal cell; EV, excretory vesicle; EXP, excretory pore; FI, filament; GE, germ cells; GEL, germinal layer; H, hook; HE, hexacanth embryo; HP, haptor; I, intestine; LL, laminated layer; M, mouth; NE, neodermis; NK, neck; OC, ocellus; OS, oral sucker; PEG, penetration glands; PH, pharynx; PN, protonephridium; PRT, protoscolex; RO, rostellum; S, sucker; SC, statocyst; SCO, scolex; SET, setae; ST, stylet; TA, tail.

thereby facilitating infection of appropriate vertebrate hosts. A generalized aquatic eucestode life cycle is illustrated in **Figure 73**. In general, eucestode zygotes initially develop into a larval stage known as a hexacanth embryo (Figure 89, **Plate 5**). This embryo characteristically possesses three pairs of hooks. The hexacanth may or may not be surrounded by a layer termed an embryophore. A hexacanth surrounded by an embryophore is referred to as an oncosphere. An oncosphere in which the outer region of the embryophore is ciliated is referred to as a coracidium and is usually free-swimming. Some cestode hexacanths or oncospheres are surrounded by an egg shell; this shell possesses a cap-like operculum in some groups. A hexacanth enters the first intermediate host either when it is eaten or by actively swimming to and then penetrating this first host. The first intermediate host is an invertebrate in most eucestode groups, but some terrestrial cyclophyllideans utilize vertebrate as first intermediate hosts. The developmental fate of the hexacanth remains to be discovered in many cestode species; this is especially true of the trypanorhynchs and tetraphyllideans. However, existing data suggest that, in general, the larval stages that follow the hexacanth vary somewhat in form among the different cestode orders. These forms are generally given different names depending on their morphology. The key features for distinguishing among larval types are the presence or absence of a scolex, the number of scoleces, whether the scolex (or scoleces) is invaginated (turned inside out) or retracted (withdrawn) into the main body of the larva, and whether a cavity known as a primary lacuna is present. Although current terminology does not accommodate all known combinations of these features, some common larval forms are given (see following). For example, procercoids are solid-bodied, lack a scolex, and usually possess a posterior extention of the body considered a cercomer by some (Figure 90, **Plate 5**). Plerocercoids possess a scolex that is neither invaginated nor retracted (Figure 94, **Plate 5**). Cysticercoids possess a conspicuous scolex that is retracted into the tissue of the neck alone, or, in some cases, the neck is then retracted into the main body of the larva (Figure 91, **Plate 5**). Cysticerci possess one (Figure 92, **Plate 5**), or more invaginated scoleces. Cysticerci with multiple scoleces are referred to as either hydatid cysts (Figure 93, **Plate 5**) or coeneri, depending on whether they have the ability to generate daughter cysts or not, respectively. The reproductive potential of hydatid cysts is enormous as each scolex (termed a protoscolex in hydatids) can go on to become either an adult worm or another protoscolex, if released from the original cyst within the first host. The numerous marine eucestode larvae referred to as "*Scolex pleronectes*" or "*Scolex polymorphus*" are likely to be the plerocercoid larvae of tetraphyllideans.

Phylogenetic Relationships

The tree shown in **Figure 1** provides an overview of the current understanding of the phylogenetic relationships among the major platyhelminth groups. The topology of this tree was derived from a combination of molecular analyses (e.g., Lockyer *et al*., 2003; Willems *et al*., 2006). The morphological characters that have been mapped onto the tree include distinguishing characters for some of the major groups as well as characters that potentially support the phylogenetic relationships shown. It is of note that, the Acoelomorpha (Acoela + Nemertodermatida), are no longer considered members of the Platyhelminthes; instead they are considered to be basal bilaterians or even early divergent deuterostomes, perhaps allied with the enigmatic *Xenoturbella*. The phylogenetic relationships of the platyhelminths in general are treated below. Readers are directed to Littlewood and Bray (2001) and Littlewood *et al*. (2004) for more detailed treatments of the phylogenetic relationships of individual groups.

In truth, few unique features (synapomorphies) have been identified for the phylum Platyhelminthes. None of the characters normally used to describe the group (i.e., dorsoventral flattening, acoelomate, bilateral symmetry) is unique to this group, nor are some of these characteristics (e.g., dorsoventral flattening) characteristic of all members of the group. A candidate synapomorphy for the phylum is the lack of accessory centrioles from the base of the cilia of the protonephridial flame bulbs, but unfortunately, this feature is visible only with transmission electron microscopy. The lack of sperm flagella is a condition seen in the three most basal platyhelminth orders (i.e., Catenulida, Macrostomida, and Haplopharyngida), but the remaining platyhelminth groups exhibit the much more widespread condition of biflagellated sperm.

Some morphological and recent molecular data appear to confirm the monophyly of many of the major platyhelminth groups but have highlighted the need for further systematic research. Although the nonmonophyly of the non-neodermatan platyhelminths, previously referred to collectively as the "turbellarians," is now well accepted, the relationships among the non-neodermatan platyhelminth orders remain poorly understood. There is general agreement that the catenulids represent the earliest divergent extant lineage of platyhelminths. The members of this order are exceptional in that they have lost their sperm flagellation and possess a single, rather than a paired, protonephridial system; they also possess protonephridia of Type 1 (i.e., with a cap cell, only two cilia, and no canal cell). In addition, the catenulids lack several distinctive features that unite most or all of the remaining playthelminths as a group referred to collectively as the Rhabditophora. For example, rhabditophorans and catenulids differ in the genetic codes they use to translate mitochondrial protein-coding genes; catenulids share the same mitochondrial code as other protostome invertebrates but rhabditophorans translate AAA as asparagine (usually lysine), and AUA as isoleucine (usually methionine). In addition, whereas most rhabditophorans possess rhabdites and the duogland adhesive system, the catenulids do not. All platyhelminths except the catenulids, macrostomids, and haplopharyngids share a unique arrangement of flagellar microtubules (9 + '1' rather than 9 + 2) in their sperm, which are biflagellated. All of the platyhelminths except most members of these three orders and the polyclads share the possession of an ovary that is heterocellular (rather than homocellular) and thus produce eggs that are ectolecithal (rather than endolecithal). With the exception of the "lecithoepitheliatans," this suite of heterocellular platyhelminths also possesses vitelline and ovarian cells that are arranged in separate organs, rather than combined into a single germinovitellarium. Apart from the fact that the polyclads

possess extensive branching of their intestinal caeca, temnocephalidans generally exhibit a series of anterior tentacles, and triclads exhibit an extra embryonic intestine, robust morphological synapomorphies for the non-neodermatan platyhelminth orders have not been forthcoming. The proseriatans lack the duogland adhesive system, but so do the temnocephalans and neodermatans. Prolecithophorans have lost their rhabdites, but so too have the temnocephalans and neodermatans. The fecampiids have lost their digestive system, but so too have the cestodes. Molecular data support recognition of the Kalyptorhynchia + "Dalyellioida" + Temnocephalida + "Typhyloplanida" as a monophyletic group, and, in fact, these four orders have collectively been referred to as the Rhabdocoela owing to their possession of a distinctive pharynx and barrel-like intestine. These four orders also share protonephridia of Type 3, but this feature is also found in the "lecithoepitheliatans." Also comprising a group supported by molecular data are the Prolecithophora, Tricladida, and Fecampiida; however, a morphological synapomorphy for this group remains to be identified.

The monophyly of the Neodermata, consisting of the monopisthocotyleans and polyopisthocotyleans (i.e., monogeneans) as well as aspidogastreans, digeneans, and the cestodes, is now among the most well-established aspects of platyhelminth relationships. The neodermatans are characterized by their possession of a unique syncytial outer body layer known as the neodermis, which generally replaces the cellular epidermis, and is shed when the first larval stage encounters the first host. In addition, the sperm of neodermatans is unusual among platyhelminths (except the kalyptorhynchs) in that the two flagella are actually incorporated into the cytoplasm of the sperm, rather than remaining separate. All neodermatans also lack rhabdites and the duogland adhesive system. The neodermatans are obligate parasites, most of vertebrates, at some time during their lives.

All Monogenea exhibit a posterior haptor. They are further unique among the platyhelminths in their possession of a ciliated larval stage known as the oncomiracidium (with the exception of the few viviparous forms) and also of four rhabdomeric ocelli at some point in their lives. However, the union of the two subclasses (Monopisthocotylea and Polyopisthocotylea) in the Class Monogenea is contradicted by molecular phylogenetic analysis of nuclear ribosomal RNA genes, and mitochondrial genomes, depending on the analysis and gene set; these molecular data suggest paraphyly of the monogeneans, with either Monopisthocotylea or Polyopisthocotylea as the earliest divergent neodermatan lineage. Nonetheless, the monophyly of each of these two subclasses appears to be robust. In each case it is supported by morphological features as articulated above and in the most extensive morphological phylogenetic treatment of the group (see Boeger and Kritsky, 2001). The nonmonophyly of the Class as suggested by molecular work may reflect a rapid divergence of the two subclasses rather than lack of support for monogenean monophyly. The interrelationships among the monogenean orders within these two subclasses are also not universally agreed on.

The Aspidogastrea, Digenea, and Cestoda are united by their complex life cycles. The close affinities between the Aspidogastrea and Digenea, collectively referred to by some as the Trematoda, are supported by their possession of epidermal cells in their first larval stage (cotylocidia and miracidia, respectively) that are separated from one another by portions of the neodermis. In addition, the first or primary host in the life cycle of the trematodes is a mollusc. The Aspidogastrea are unique in their possession of oviducts separated into chambers by septa and and also potentially in their possession of neodermal microvilli in the form of hemispherical microtubercles. The Digenea are distinctive in their possession of several unique larval stages, the miracidium, sporocyst, redia, and cercaria. The phylogenetic relationships among the digeneans have been explored relatively recently in both morphological (e.g., Cribb *et al.*, 2001) and molecular (e.g., Olson *et al.*, 2003) contexts, but no higher-level classification scheme is generally accepted. More detailed analysis is required before these relationships can be resolved with any confidence.

There is little doubt that the cestodes are monophyletic. All cestodes lack essentially all elements of the digestive system and exhibit unique extensions of the neodermis called microtriches (**Figure 16**). The Eucestodes are particularly unusual in that their sperm lack mitochondria. The relationships among the cestode orders, and also the circumscription of some orders, remain controversial despite the fact that this issue has received much recent attention (e.g., Waeschenbach *et al.*, 2007). The recent erection and resurrection of five cestode orders (Bothriocephalidea, Cathetocephalidea, Diphyllobothriidea, Litobothriidea and Rhinebothridea) attests to the active state of cestode phylogenetics. The monophyly of the Tetraphyllidea is particularly problematic; the order is almost certainly not monophyletic even with the recent extraction of the Rhinebothriidea.

Host Associations

It is useful to consider host associations in general terms. The tendencies of platyhelminths to live in association with other organisms are summarized on the tree in **Figure 1**. One of the most striking aspects of this mapping is that it becomes apparent that all but six of the major platyhelminth groups (catenulids, macrostomids, haplopharyngids, "lecithoepitheliatans," bothrioplanidans, and kalyptorhynchs) include at least some species that live in association with other organisms. Thus, symbiosis, in its broadest sense, is a widely distributed feature among the platyhelminths. Within the major platyhelminth groups, however, symbiosis varies in its occurrence. Only the neodermatan groups and the temnocephalids consist entirely of obligate symbionts. The dalyelloids and fecampiids are predominantly symbiotic, although a few free-living species are known. The remaining groups (polyclads, proseriatans, triclads, prolecithophorans, and "typhloplanoids") are predominantly free-living, but all of these groups also include at least some species that are symbiotic.

The hosts of almost all of the symbiotic non-neodermatans are invertebrates; molluscs, arthropods, and echinoderms figure prominently in this role. The neodermatans are unlike almost all of the other symbiotic platyhelminths in that they are generally associated with at least one vertebrate host group at some point in their lives. The aspidogastreans are somewhat exceptional in that some species are known only from freshwater clams and snails. However, records of aspidogastreans from the digestive system and its associated organs, in turtles,

elasmobranchs, and teleosts are not uncommon. Although uncharacteristic of the group as a whole, some monogeneans (e.g., gyrodactylids) are found associated with cephalopods, and others with crustaceans that parasitize teleosts or elasmobranchs.

With the exception of the monogeneans, the neodermatan groups are also unique in that their life cycles generally involve more than a single host species. All digeneans possess at least two different hosts in their life cycles. In general, the first (intermediate) host is a mollusc, while the final (definitive) host is a vertebrate. In species possessing life cycles involving more than two hosts, there is one (or sometimes more) additional intermediate (or paratenic) host that is either an invertebrate or a vertebrate depending on the digenean group. In such taxa the host that precedes the vertebrate is generally an invertebrate, often some sort of mollusc or arthropod. The monogeneans are exceptional in this respect as their life cycles involve only a single host.

Symbiotic platyhelminths are found either in or on their hosts. The non-neodermatan groups are often found on, rather than in, their hosts. The exceptions are the fecampiids, "dalyelloids," which are most commonly found inhabiting the coelom or digestive tract of their hosts. Among the neodermatans, the monogeneans are unique in that they are generally found on (i.e., ectoparasitic), rather than in, their hosts whereas the remaining neodermatan groups are most commonly found as endoparasites of their respective hosts.

Medical Importance

The medical and economical importance of the various platyhelminth groups is closely tied to the type and intimacy of the host associations of the groups. Species that use vertebrates as hosts are obviously of much greater concern to humans than those that do not. Thus, the neodermatans are by far the most medically and economically important platyhelminth groups. Space limitations do not permit discussion of the economic consequences of infections with neodermatans, but some of the medically important species are briefly discussed below.

Within the neodermatans only the digeneans and cestodes include species that infect humans. Among the 25 superfamilies of digeneans, human parasites are essentially restricted to four. In the majority of these species it is the adult stage that parasitizes humans. Many of these species are zoonotic; that is, their normal hosts are usually vertebrates other than humans. Humans acquire infections with such species when they unwittingly ingest the metacercarial stage. These zoonotic groups infect a variety of different organs in their human hosts. They include a total of eight species of lung digeneans in the superfamily Gorgoderoidea, five species of liver digeneans in the superfamily Opisthorchioidea, and seven species in the superfamilies Echinostomatoidea, Opisthorchioidea, and Gorgoderoidea that inhabit the digestive tract of their hosts. Finally, human infection with cercariae of species of the superfamily Schistosomatoidea that normally infect birds and aquatic mammals is responsible for an irritating skin condition in humans known as cercarial dermatitis or swimmer's itch.

Only six species of digeneans parasitize humans as their normal final (definitive) hosts. These include one species of liver fluke in the superfamily Opisthorchioidea and five species of blood flukes in the superfamily Schistosomatoidea. The unusual life cycle of the schistosomes requires that humans come into contact with the free-swimming cercariae of these species to acquire an infection. Schistosomes, which cause a disease known as schistosomiasis in humans, are without question the most deadly of the digeneans. Hundreds of millions of people globally are infected with these species. Schistosomes are responsible for the deaths of tens of thousands of people each year.

Species that parasitize humans belong to only two of the 18 orders of tapeworms. The adults of at least four species of diphyllobothriideans (all of which are members of the genus *Diphyllobothrium*) have been reported from the digestive system of humans; evidence suggests that all of these species are zoonotic in humans, as they normally parasitize other fish-eating vertebrates as adults. Plerocercoids of some species of diphyllobothriideans of the genus *Spirometra* are known to parasitize the musculature and subcutaneous tissues of humans, causing a condition known as sparganosis. By far the most pathogenic of the human tapeworms, however, are the cyclophyllideans. Nine species in five genera parasitize humans. These species are known from humans as either larval or adult stages. Larval stages are generally much more pathogenic than adult stages. The hydatid cysts, coeneri, and cysticerci of *Echinococcus*, *Multiceps*, and *Taenia* species, respectively, are arguably the most life-threatening of the human-infecting cestodes. These larval stages often develop in difficult-to-treat extraintestinal sites, such as the brain, or grow extremely large and, in some cases, invasively into one or more organs of the body, making surgical removal of the parasite difficult. Thus, infection with any of these larval stages can be fatal. Infection with an adult tapeworm is generally debilitating rather than life-threatening, because these stages remain in the digestive system and thus are relatively easy to treat. With the exception of *Taenia solium*, it appears that tapeworm infections in humans are also all zoonotic. Infection with adult tapeworms can be prevented by properly cooking food. Infection with larval stages usually involves contact with fecal matter from humans or other animals.

See also: Parasitism. Worms, Annelida. Worms, Nematoda

References

Ax P (1996) *Multicellular Animals: A New Approach to the Phylogenetic Order in Nature*, vol I. New York: Springer-Verlag.

Boeger WA and Kritsky DC (2001) Phylogenetic relationships of the Monogenoidea. In: Littlewood DTJ and Bray RA (eds.) *Interrelationships of the Platyhelminthes*, pp. 92–102. London: Taylor and Francis.

Bray RA, Gibson DI, and Jones A (eds.) (2008) *Keys to the Trematoda*, vol. 3. Wallingford: CABI.

Cable RM (1956) Scientific survey of Porto Rico and the Virgin Islands. *Marine Cercariae of Puerto Rico, XVI, Part 4*, pp. 491–577. NY: NewYork Academy of Sciences.

Cannon LRG (1986) *Turbellaria of the World*. Brisbane, Australia: Queensland Museum.
Ehlers U (1985) *Das Phylogenetische System der Plathelminthes*. Stuttgart: Gustav Fischer Verlag.
Cribb TH, Bray RA, Littlewood DTJ, Pichelin SP, and Herniou EA (2001) The Digenea. In: Littlewood DTJ and Bray RA (eds.) *Interrelationships of the Platyhelminthes*, pp. 168–199. London: Taylor and Francis.
Grassé PP (1961) *Traité de Zoologie: Plathelminthes, Mésozaires, Acanthocéphales, Némertiens* Volume IV. Paris: Masson et Cie.
Halton DW and Gustafsson MKS (1996) Functional morphology of the platyhelminth nervous system. *Parasitology* 113: 47–72.
Hyman LH (1951) *The Invertebrates II. Platyhelminthes and Rhynchocoela. The Acoelomate Bilateria*. New York: McGraw-Hill.
Gibson DI, Jones A, and Bray RA (eds.) (2002) *Keys to the Trematoda*, vol. 1. Wallingford: CABI.
Jones A, Bray RA, and Gibson DI (eds.) (2005) *Keys to the Trematoda*, vol. 2. Wallingford: CABI.
Khalil LF, Jones A, and Bray RA (1994) *Keys to the Cestode Parasites of Vertebrates*. Wallingford: CAB International.
Littlewood DTJ (2006) The evolution of parasitism in flatworms. In: Maule AG and Marks NJ (eds.) *Parasitic Flatworms: Molecular Biology, Biochemistry, Immunology and Physiology*, pp. 1–36. Wallingford: CAB International.
Littlewood DTJ and Bray RA (eds.) (2001) *Interrelationships of the Platyhelminthes*. London: Taylor & Francis.
Littlewood DTJ, Telford MJ, and Bray RA (2004) Protostomes and Platyhelminthes – the worm's turn. In: Cracraft J and Donoghue MJ (eds.) *Assembling the Tree of Life*, pp. 209–234. New York: Oxford University Press.
Lockyer AE, Olson PD, and Littlewood DTJ (2003) Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): Implications and a review of the cercomer theory. *Biological Journal of the Linnean Society* 78: 155–173.
Olson PD, Cribb TH, Tkach VV, Bray RA, and Littlewood DTJ (2003) Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). *International Journal of Parasitology* 33: 733–755.
Rieger RM (1998) 100 Years of research on 'Turbellaria'. *Hydrobiologia* 383: 1–127.
Rohde K (1997) The origins of parasitism in the Platyhelminthes: A summary interpreted on the basis of recent literature. *International Journal of Parasitology* 27: 739–746.
Rohde K (1991) The evolution of protonephridia of the Platyhelminthes. *Hydrobiologia* 227: 315–321.
Ruppert EE (1978) A review of metamorphosis of turbellarian larvae. *Settlement and Metamorphosis of Marine Invertebrate Larvae*, pp. 65–81. New York: Elsevier.
Schell SC (1985) *Handbook of Trematodes of North America North of Mexico*. Moscow, ID: University Press of Idaho.
Schmidt GD (1986) *CRC Handbook of Tapeworm Identification*. Boca Raton, FL: CRC Press Inc.
Tyler S, Schilling S, Hooge M and Bush LF (compilers) Turbellarian taxonomic database. Version 1.6. http://turbellaria.umaine.edu (2006–2010).
Waeschenbach A, Webster BL, Bray RA, and Littlewood DTJ (2007) Added resolution among ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with complete small and large subunit nuclear ribosomal RNA genes. *Molecular Phylogenetics and Evolution* 45: 311–325.
Willems WR, Wallberg A, Jondelius U, *et al.* (2006) Filling a gap in the phylogeny of flatworms: Relationships within the Rhabdocoela (Platyhelminthes), inferred from 18S ribosomal DNA sequences. *Zoologica Scripta* 35: 1–17.
Yamaguti S (1963) *Systema Helminthum. Monogenea and Aspidocotylea*, vol. IV. New York, NY: Interscience Publishing (John Wiley & Sons).
Yamaguti S (1971) *Synopsis of Digenetic Trematodes of Vertebrates*, vols. I and II. Tokyo: Keigaku Publishing Co., Ltd.

Z

Zoos and Zoological Parks

Anna Marie Lyles, Bio-Gist Ventures, Princeton, NJ, USA
Dan Wharton, The Chicago Zoological Society, Brookfield, IL, USA

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by Anna Marie Lyles, volume 5, pp 901–912, © 2001, Elsevier Inc.

Glossary

AZA Association of Zoos and Aquariums, a North American professional organization.
Ex situ Outside the natural range. Hence, a captive breeding program in Europe for a South American species would be an *ex situ* program. (Conversely, a Chinese program to breed native panda bears in captivity is an *in situ* breeding program because it is within the natural range.)
Exotic animals Animals that have not been domesticated for use by humans in agriculture or as pets; also connotes that the wild animals are from another region of the world.
Founders Animals that were collected from their natural range and that have produced progeny in captivity (and are assumed to be unrelated). Thus, these animals are the founding stocks of the captive population. Also generalized to include animals with wild parents that are collected in order to supplement the genetic diversity of an established captive population.
International Species Information System (ISIS) A central repository for zoo animal records founded in 1973; has also developed the computer programs that most zoos use for keeping their animal records, as well as software used by zoos for studbooks.
Studbooks Also called genealogies, registries, or pedigrees, a database of historic and living animals intended to go beyond passively recording bloodlines and records; typically kept in computer databases that are also useful for actively managing the genetics, demography, and husbandry of captive populations.

Introduction

Zoo is a handy three-letter word, and thus in this article "zoos" will be used in a very general way to include any public or private institutions that house live exotic animals, largely for the purposes of exhibition to the public. Zoos are permanent establishments that are open to the public to provide education, recreation, and cultural enjoyment. Zoos, for purposes of this discussion, are not transient shows such as circuses or petting zoos. Nor are zoos private collections of exotic animals. "Zoos," defined thus, include the facilities that are traditionally considered zoos, such as zoological parks and aquariums, and also many wildlife centers and "bioparks," as well as some insect houses and *ex situ* breeding centers. There are currently over 1200 well-established zoos worldwide.

Zoos reach a vast audience; for example, North American accredited zoos attract more visitors, currently estimated at over 175 million annually, than the major professional sports combined. Worldwide, zoo attendance is estimated at over 600 million visitors (Gusset and Dick, 2011). The best modern zoo exhibits are sublime, invoking in visitors a sense of awe and respect for nature. Zoos are increasingly operated by specialized, professional staff in partnership with government and private interests, with the express goal of conserving wildlife. The main focus of this article is on how zoos are evolving into institutions for wildlife conservation, and also on the art and science of managing biodiversity in zoos.

The Changing Role and Facade of Zoos

Zoos from Ancient to Recent Times

Zoos probably arose as places of spectacle and entertainment shortly after civilization became urbanized. The first recorded zoological collection dates to 2500 BC, when several thousand wild animals are documented in Saqqara, Egypt. Ancient Egyptians kept many wild species for use in religious ceremonies, or perhaps because they hoped to tame them. These ancient animal collections were acquired as tribute gifts or via expeditions that were mounted specifically to collect wildlife.

Zoo collections have sprung up independently on many continents and in many civilizations. For example, Chinese

emperors as far back as 1150 BC had walled animal collections they called "parks of knowledge." When Cortés reached Mexico City in 1519, he witnessed, and then destroyed, the Aztec ruler Montezuma's immense zoo and gardens. Montezuma's zoo was so large that 300 keepers were employed for the aviaries alone, and it even housed malformed humans as part of the collection. Most of the zoological collections of antiquity were imperial and exclusive – symbols of wealth and power. They were menageries in the sense that they were collections of living trophies to be managed.

Collections during Roman times were also for the amusement of citizens. Large numbers of wild animals were used for blood sport and spectacles; holding areas that also allowed public viewing, called *viveria*, were associated with arenas. After the fall of Rome, collections of wildlife were not so common in Europe. During Europe's Middle Ages, imperial menageries once again became more significant. Wild animals were used as diplomatic gestures – as tributes and ceremonial gifts.

The roots of public zoos are somewhat contemporaneous with the emergence of Europe's great cities. London gained a zoo, of sorts, in 1252 when Henry III transferred his menagerie to the Tower of London; London's citizens were pressed to support the menagerie, and in turn the public could view some of the animals, such as a polar bear and an elephant. Urban menageries, under the sponsorship of merchants and other citizens, began to emerge, first in Frankfurt in the late 1300s, then in the 1500s at The Hague and in Augsburg.

The rising profile of zoos can be traced to the Renaissance, when trade and exploration brought more animals into Europe. Colonial officers in far-flung empires served as conduits for animal requests. Public interest was most piqued by the wildlife of Africa and the Far East. The first zoological garden, which combined displays of plants and animals, was established in Versailles by Louis XIV in 1665; in 1794 this collection became a division of Paris' natural history museum, creating a formal link to zoological research that was later emulated by many other zoos and museums. The first "modern" zoo that survives to this day is probably the Schönbrunn Zoo in Vienna, originally established as a private collection in the 1770s, but to which the public were intermittently admitted.

The nineteenth century has been hailed as the century of science and the golden age of museums. This century witnessed an explosive increase in the numbers of zoos and museums. Scientific societies sprang up, including zoological societies. Zoological research was nominally based in nineteenth-century zoos, but the inquiry was mostly limited to collecting and classifying. Local governments tended to operate zoos of this era with the aim of providing recreation and edification for the increasingly better-educated citizenry that arose with the Industrial Revolution. Zoos were sources of civic pride, and they competed to exhibit as many kinds of animals as possible. The rise of print media made it possible for civic leaders to compare collections more readily with rival zoos. Natural social groupings were largely ignored, and frequently solitary animals were displayed. Generally species were grouped to correspond with scientific classification, for instance, into bird, reptile, or carnivore houses. These animal collections were viewed as analogous to collections of art, and even the zoo architecture of the period came to be considered art.

An innovation at the beginning of the twentieth century was habitat exhibits, as pioneered by Carl Hagenbeck at Tierpark, outside Hamburg, Germany. In these zoos of living panoramas, it became increasingly common to display animals along geographic lines rather than strictly taxonomic ones. Habitat exhibits were well suited to zoo designs that intended to teach ecological principles, and thus led to the display of animals in more natural social groupings. Eventually mixed-species exhibits became more common, and these conglomerations might even be grouped by ecological biomes such as "polar zone" or "tropical zone," rather than along strictly geographical or taxonomic lines.

Later in the twentieth century, world culture became increasingly electronic and graphic. Zoos found themselves in competition with wildlife film documentaries. No longer would the animal displayed in a sterile cage suffice for an increasingly sophisticated audience. "Immersion" exhibits became the ideal for most new zoo development. Immersion exhibits are ones that seek to make visitors feel that they are in the habitat with the animals. This may be accomplished by incorporating the same details of rocks, logs, or plants into both the animal and the public space, using these same details to hide any cage elements, and also by special effects such as mist machines and animal call background soundtracks that help create an illusion. Barriers to separate animals and people may be relatively invisible (e.g., glass) or carefully placed so they are unobtrusive. Special feeders and techniques of animal management may be used to encourage animals to be active and visible – for example, hidden heating or cooling pads for animals to lounge on, depending on season, or feeders that intermittently spray out small amounts of food near a viewing station to encourage visible foraging behavior. Zoos often try to craft the visitor space so that visitors do not see other visitors while observing the animals. In many instances, the designers' intent is to evoke an almost religious experience of awe and respect for nature, much as the grand cathedrals did in early eras.

Rise of the Zoo Conservation Ethic

Zoos, as we have seen, arose as places of spectacle, not as conservation organizations. But many modern zoos have chosen wildlife conservation as their primary mission. The conservation ethic has developed in zoos over the past century in parallel with the nonzoo world, and in many cases one can point to zoos as the places that incubated the people and ideas that have led to important conservation developments (Norton *et al.*, 1995). The National Zoo in the United States, founded in 1887, is considered to be the first zoo created to preserve endangered species.

Conservation of wildlife and wildlife habitat has become the major goal of most of the prominent zoos and zoo professional organizations (IUDZG/CBSG (IUCN/SSC), 1993). Although altruism is part of the reason for the shift, compelling utilitarian reasons also drive these lofty goals. As humanity expands and wild places contract, exotic animals are getting scarcer, thereby reinforcing the zoo conservation mission while presenting the reality that zoos must become net producers of wildlife if their animal exhibits are to remain vibrant, inspirational, and of a multifaceted significance to

conservation. As regulators seek to simultaneously protect their natural resources and their agricultural sectors, governments make it ever more difficult to import and export wildlife. Wild animals may carry diseases that will harm livestock and perhaps even spread to the human populace; increasing understanding of disease has led to import restrictions or quarantines. For example, the bird-importing era, at least in the United States, slowed markedly in the early 1970s, when quarantine laws were enacted to guard against Newcastle's disease. In the 1990s the Wild Bird Trade Act was passed in the United States to restrict importation of any birds protected by CITES, the Convention on International Trade in Endangered Species of Wild Fauna and Flora. Most professional animal dealers have gone out of business due to the increasing restrictions, and zoos collectively have supported these regulations as part of their conservation mission.

Simultaneously, zoos have needed to import fewer animals because they became more proficient at keeping and breeding wildlife. Zoo keeping is a complicated art and science, and zoo keepers and managers have found it useful to pool their knowledge. Furthermore, single zoos are rather like islands from a population biology point of view, in that they are not big enough to sustain viable populations. So cooperative breeding programs for animals have become key, while cheap, fast transport, usually by air, makes it feasible for distant zoos to exchange breeding stocks. In many regions of the world, the leading zoos and their staff now belong to professional organizations. There are also world zoo organizations that help link the regions. These organizations have codes of ethics and accreditation (certification) programs. For example, North America's American Zoo and Aquarium Association (now renamed the Association of Zoos and Aquariums or AZA) was founded in 1924 and currently includes 223 accredited institutional members (this is about one-tenth of the number of animal exhibitors that are licensed with the US government).

Public support is yet another motive for many zoos to become conservation organizations. Zoos experience challenges in the competition for public funding. Modern zoos are costly, with expensive professional staffs, sophisticated veterinary medicine facilities, and a public that demands complex, educational, and artistic exhibits (Conway, 1986). Municipal governments are increasingly burdened by social programs and other competing demands for funds and now only fund a fraction of zoo operations compared to earlier times (on average 21% of the zoo budget in North America). But while the funding situation is tough in the richer nations, it is often quite desperate in poorer regions of the world. Most zoo-based conservation is being done by the "haves" whereas much of the need for conservation and access to endangered species is in the "have-not," largely tropically located zoos of the world.

There is enormous public support for environmental causes and grants available for conservation programs. Despite a vocal minority of critics who question the ethics of keeping animals in zoos, zoos are maintaining enormous popularity as they further refine their role as advocates of, and players in, wildlife conservation (Falk *et al.*, 2007). Zoos have often contributed significantly to the economies of their regions. AZA figures indicate $15 billion in annual economic activity among the AZA-accredited zoos and the support of 142,000 jobs. For instance, the Chicago Zoological Society estimates that their Brookfield Zoo generates $150 million to the Chicago area whereas the four zoos and aquariums run by the Wildlife Conservation Society in New York City are worth approximately $350 million to New York.

Animal Welfare in Zoos

Advocates for animals are becoming increasingly strident, even to the extent of using terrorist tactics. Their ire is more typically directed against laboratories that use animals, but zoos also have come under fire. Some animal rights extremists have expressed a goal of putting zoos out of existence. Extreme though these views may be, they signal a shift in public attitudes toward animal welfare. Many people now consider animals as fellow creatures with needs and feelings that should be respected. Zoos that have not become attuned to perceptions of animal welfare have often experienced an erosion of public support, often regained by a renaissance of rebuilding on existing or new sites and hiring some of the best professionals in the business. Some of the poorer, old-fashioned zoos have closed.

These animal advocates often confuse animal rights with animal welfare, and their anthropomorphic ideas about what animals need are often at odds with what is best for animals. For example, well-meaning protestors tried to prevent the transfer of a gorilla named Timmy, who had never procreated, to a zoo in New York that had considerable success in breeding gorillas as well as many potential fertile mates for Timmy. The protestors did not want Timmy to be separated from his current, infertile female partner. However, in nature gorillas do not form monogamous bonds and they do transfer from one social group to another. Eventually, after a court found the transferring and recipient zoos to be in total compliance with both the Endangered Species Act and the Animal Welfare Act, Timmy was moved, formed bonds with new females, and has begotten numerous offspring.

A related problem that we shall return to in the section Keeping Biodiversity is the problem of limited space in zoos. Thus, breeding unwanted, "surplus" animals is a problem for zoos because surplus animals inhabit valuable space. Some zoos take the apparently logical stance, from a utilitarian animal management vantage, of culling their surplus stocks or selling them to the private collectors. In private hands, some animals have ended up on game ranches where they are shot for sport. This has caused major public relations imbroglios, as a result of which zoos have had to examine their ethical policies and restrict their animal trade to accredited zoos. Zoos have developed contraceptive and husbandry techniques to prevent unplanned breeding, and they have developed scientific breeding plans to keep population demographics aligned with available space. Furthermore, there has been a trend of exchanging animals without fees as they become increasingly priceless.

Improved contraception for animals is just one example of how health care for exotic animals is becoming increasingly sophisticated. Improved veterinary care deserves much of the credit for allowing zoos to shift from relatively sterile concrete and steel cages to more natural housing, which is often,

literally, dirty (as the substrate is frequently dirt or other organic matter). Most zoos now have their own veterinarians and veterinary hospitals. Some even have pathologists, endocrinologists, and nutritionists providing preventive care, better drugs, and diagnostics. The understanding of wild animals' nutritional needs is advancing rapidly, and there are now numerous commercial brands of food for exotic animals, such as "Flamingo Fare." Nutritionists are even improving on fish diets (which are expensive and become depleted of certain vitamins during shipping and storage) for animals such as penguins and sea lions by devising a gelled, artificial fish in a fish shape. Most zoos have "browse" programs to grow and harvest fresh plant foods for their denizens.

There are numerous technical periodicals, books, and other publications related to zoo husbandry, including *Zoo Biology*, the zoo profession's peer-reviewed science journal, *International Zoo News*, the *International Zoo Yearbook*, *The Shape of Enrichment*, and *The Dodo*. As a testament to the improving welfare of zoo animals, zoos are now recording many longevity records, and zoo animals frequently senesce and die of old age – an almost unheard of occurrence in nature. As zoos have become more successful at keeping exotic animals alive and meeting their physical needs, the attention of zoo biologists has begun to shift toward behavior.

Zoos have also added applied psychology to their toolkits. They now must reassure their public that animals are not distressed but are potentially "happy." New techniques are being developed to stimulate zoo animals, reduce psychological stress, and elicit cooperation rather than coerce desired behaviors (Shepperdson *et al.*, 1998; Watters and Wielebnowski, 2009). Some zoos train animals to perform in shows, which is not a new phenomenon. But modern shows demonstrate natural behaviors to the public and keep the performing animals occupied, rather than emphasizing cute tricks or an animal's fierce nature. These new techniques range from behavior management and operant conditioning to environmental enrichment. Behavioral techniques aim to banish boredom, eliminate stereotypic tics, and otherwise encourage behavioral patterns that approximate those of animals in nature. For example, the Chicago Zoological Society has made a major commitment to behavioral science as it applies to zoo animal management through their Center for the Science of Animal Welfare (see http://www.brookfieldzoo.org/czs/Animal-Care/Center-for-Science-and-Animal-Welfare).

Emerging Trends

Zoos, as we have seen, have evolved from imperial menageries to the modern habitat and immersion exhibits that feature animals in more natural settings and social groupings. In the twenty-first century, we expect this trend toward more ecological themes to continue, with more zoos featuring animals in diverse, mixed-species exhibits. Zoos that predominantly have these diversity-designed exhibits have been given the moniker "biopark," or "biodome." This sort of zoo may feature botanical collections housed with terrestrial and aquatic vertebrates and invertebrates. There is a considerable art in developing these mixed-species exhibits, because zoos must ensure that cohoused animals do not decimate the foliage or each other. Animals that do wreak havoc with plants or other animals are either excluded from the biopark or managed separately, although their exhibit boundaries may be crafted carefully so that this is not obvious. Some zoos are extending this idea to focus on the ecology of a certain region using both zoo exhibits and more museum-like exhibits; typically the focus may be the region in which they are located – for example, the "living desert museums" in the American southwest near Tucson, Arizona and Palm Springs, California.

Also emerging in the twenty-first century seems to be a new trend, which we shall call "conservation edu-tainment." These thematic exhibits have a storyline to which visitors may be introduced, sometimes through videos or costumed actors, as they wait to enter the exhibit. This visit may even include an amusement-style ride; for example, Sea World visitors to the polar region exhibit experienced a "tundra buggy" simulated ride before viewing polar bears and other creatures. Some call this the "Disneyfication" of zoos, after a zoo, "Disney's Animal Kingdom," that the Walt Disney Company opened in Florida in the late 1990s; one feature is a drive-through, safari-style bus ride, where the storyline is that visitors are joining a patrol to catch poachers.

The actual educational value of these new-style zoo exhibits is debatable, and a still greater unknown is whether the lessons learned by visitors will translate into conservation actions. Conservation education, zoos have argued, is one of their major contributions, one that is becoming increasingly important as citizens become further distanced from nature. Zoos are trying to justify these claims, and also to do a better job of conservation education. Thus, zoos are beginning to scrutinize their exhibits and programs using evaluation techniques. Evaluation methods, including surveys, interviews, or observation of visitors, are sometimes employed during the design phase to improve the message before expensive construction begins. They may also be used at other times to help fine-tune exhibits or otherwise inform about how the zoo is working or what the audience is like. The science of evaluation is rapidly developing in zoos and museums, and it will probably become an increasingly important part of zoo planning for quantifiable conservation outcomes and mission accomplishment (Falk *et al.*, 2007).

As the world becomes more wired to the Internet, zoos are following suit. Many zoos now have websites where visitors can find information about the zoo, and many also use social networking media. Some zoo databases, including the vast International Species Information Service (ISIS) registry of zoo animals, are now available at Internet websites. The newly launched Zoo Information Management System (ZIMS), aims to increase information sharing about collections for better animal management and population sustainability. Some zoos are incorporating electronic media into their educational efforts, either integrated with exhibit graphics or separately. Cooperative breeding plans, and the data needed to manage them, will increasingly be shared via electronic means. And some website programs (for example, one for trumpeter swans) allow school children and others to dynamically follow the movements of reintroduced animals via satellite tracking. Virtual zoos will not keep real animals, which is the primary business of zoos, so zoos will probably continue to find ways to use the Internet that enhance rather than radically change what they do.

As the supply of wild-born founders dwindles, and more amenities are expected for the animals in hand, zoos are abandoning the one-of-each-kind collecting approach. Within-zoo species diversity is decreasing in order to make room for zoos to manage larger, more natural social groups and to house a breeding nucleus of at least several pairs and their young offspring of nonsocial species (Gibbons *et al.*, 1995). Many zoos are trying to allocate a proportion of their animal space for off-exhibit breeding efforts, and some are operating special facilities devoted to breeding of endangered species – mostly *ex situ* but also some *in situ*. Many zoos are developing specializations, concentrating their efforts on the specific groups of animals that they can keep best and exhibit most dramatically. Zoo workers are also becoming more specialized, both along taxonomic lines and also professionally into fields such as educators, veterinarians, nutritionists, population biologists, behaviorists, life-support system operators, pathologists, and horticulturists. With the increasing need for cooperation, more formal committees and other structures are developing through the professional organizations.

Choosing Animals to Promote Biodiversity Conservation

Space, the Final Limit

Modern zoos are expensive to build and upgrade. In North America alone, zoos spend millions of dollars per year on capital or physical improvements, with much larger expenditures for major new exhibits or zoo facilities. A major exhibit can cost in the tens of millions of dollars and typically takes at least 3 years to plan and build. Maintaining a zoo and its animals is also very expensive. Therefore, there are practical limits to how many zoos can be supported and thus to how many animals can be maintained by zoos (Conway, 1986). For example, the number of animals that are presently kept in the 223 accredited zoos of North America is approximately 870,000 from among 6000 species.

Thus, conservation-oriented zoos choose their inhabitants with care in order to maximize the biodiversity that they preserve in their finite space. A biodiversity trade-off must be made between the number of species that are maintained by zoos versus the genetic diversity or size of the captive populations that are maintained. Zoo planners have to think long-term because they are trying to maintain viable populations for decades or centuries. In many cases, zoo populations are, essentially, closed gene pools because it is difficult or impossible to acquire new lineages or wild-caught "founders." Extinction in nature is the most obvious cause for a closed captive population, and zoos work with several "extinct" species such as the Prezwalski's horse. But species that are extant in nature can also be effectively isolated in zoos by government import or export regulations, concerns about disease transmission, rarity in nature, political issues, or the extreme difficulty and expense of collecting some species. In most cases zoo planners also try to "save all the pieces" by managing their populations for genetic variability. Genetic variability is an achievable objective since anything approaching natural selection is by definition impossible while the animals are under human care. Arguably, this is a prudent and responsible strategy should the animals be needed someday for restocking wild habitats. Habitats are also changing, so the theoretically successful genetic profile of the past might not be the same in the future. By breeding for variability, i.e., retaining as much of variability as was present in the original wild-caught founders as possible, the potential for readaptation to old or new habitats is maximized.

Hence, zoo population plans usually strive to breed animals better than would be done randomly without a scientific framework, and they often try to do this by maintaining small, well-managed populations that are projected to be genetically and demographically viable well into the future. This usually requires a captive population of several hundred animals. In this effort, attention must often be paid to picking the right unit for conservation, or the evolutionarily significant unit. Zoos have funded a considerable amount of applied molecular genetic research in order to determine subspecies and species characteristics, thereby avoiding the possibility of breeding animals that are similar-looking but are in fact different species. The converse is just as problematic when members of the same species are perceived as distinct and maintained separately, taking up valuable space and/or compromising the viability of both populations with suboptimal numbers.

Metamanagement of Zoo Biodiversity

Each zoo is an island too small to do much conservation breeding on its own. To facilitate cooperation, organized conservation programs involving captive animals are beginning to develop around the world through regional zoo professional associations. Most of the labor that goes into these programs is pro bono or volunteer. Much of the work is accomplished through workshop meetings and via committee, at considerable personal and institutional expense.

The basic unit of these programs is the studbook, which is the database of historic and living animals in a particular captive population. Studbooks are usually maintained through the efforts of one "studbook keeper." Populations that are monitored with studbooks can be managed through regional or even global captive breeding plans, usually overseen by a committee, which specifies where each animal should be and whether it should be bred.

To prioritize how much captive space various populations should be allocated, zoo specialist groups for each kind of animal have formed, and are known as "taxon (for taxonomic) advisory groups." Taxon advisory groups typically work closely with the appropriate taxonomic and scientific specialist groups of the World Conservation Union (IUCN), as well as with scientific advisors. Regional taxon advisory groups tend to develop priorities that are complementary to those in other areas; for example, Europe may choose to allocate space to breed a different species of penguin than North America. Taxon advisory group priorities are formalized into "regional collection plans," which ideally include a justification for including or excluding each species or subspecies in the taxonomic group, as well as setting a target population size for each managed population. These plans may also specify areas where research is needed, or where field conservation should

be a priority. Tools are being developed to facilitate regional planning, including metacollection software that has been developed by Australian zoos.

The representation of biodiversity in zoos is biased toward animals that are furry, large, brightly colored, or otherwise captivating for people. In general, the more closely related to humans, the greater the representation. So most kinds of primates are kept in zoos and managed through cooperative programs, whereas only a fraction of rodents or bats are represented. Similarly, most kinds of flamingos can be found in captivity, whereas only a fraction of the small perching birds are kept. A smaller proportion of birds are maintained in organized programs than mammals, but only a fraction of the scaly creatures (fish, reptile, and amphibian) are kept. Very few invertebrates are represented in captivity, but this seems to be changing as butterfly exhibits and insect displays are gaining in popularity. Invertebrate husbandry is relatively primitive for both aquatic and terrestrial species; the public generally considers the collection of wild invertebrates acceptable, so there has been less pressure to develop husbandry techniques. The situation for fish, and most other nonmammalian aquarium denizens, is similar to that of invertebrates.

Institutional Planning

Zoos as individual institutions are also becoming very mission-oriented. Many zoos have developed mission statements and master plans for growth and development. These plans provide the overall organization of the zoo exhibits and the visitor traffic patterns as well as locations of facilities for merchandise, refreshments, and bathrooms. (Bathroom location is amazingly important for a positive visitor experience!) Master plans usually designate the kind of educational messages the zoo will try to convey and the themes of the exhibits. They may be informed by practical considerations of what animals and plants will do best in the zoo, and these may be based on local climate or the experience and expertise of the zoo's staff.

Within the context of the master plan, zoo planners, typically curators, will develop an animal collection plan. Ideally, curators will develop these plans after studying the regional collection plans that have been developed for metamanagement of zoo collections. This collection plan will usually list each species that the zoo keeps or wishes to obtain, along with a justification for what the species will contribute to the zoo and to conservation of the species. The plan should say how many individuals there are of each sex, and how many are desired. The plan will specify where these animals will be housed. If the species is managed by a cooperative breeding plan, then the institutional collection plan should include the recommendations for which animals are to be bred and which are not. If not part of a cooperative plan, the curators should have a breeding and possibly a research plan of their own. Institutional collection plans usually need to be updated annually.

Keeping Biodiversity

Zoo Animal Records Keeping

Basic to managing animals and improving husbandry is keeping track of the inventory and gathering data. The memory of zoo staff is no longer sufficient to keep these records, as animals live longer and breed better, the government expects more information, and both animals and keepers are more mobile. Zoos are tracking an ever-increasing amount of information about their animals. Each animal in the inventory, current and historic, is accessioned, which means it is assigned a unique identification number. The animal also must be physically identifiable, so most are tagged or marked; this can be done with traditional bands and tags or with more high-tech injectable microchips. Records are kept for diverse variables, including characteristics that allow an animal to be individually identified, dates of birth and death and other life-stage transitions, sex, parents, offspring, current and previous locations, health notes, cage mates, previous owners, physical parameters such as measurements, species and subspecies, and behavioral observations. Zoos are standardizing their data collection and records so that all this information can be pooled.

Keeping these records has turned into a specialty of its own, and the keepers of the data are usually called "registrars." Registrars may also be responsible for some of the government permits and paperwork associated with modern zoo keeping. These data are generally organized into a computer database. The most widely used program is Animal Records Keeping System (ARKS) that was developed by ISIS in 1985. ARKS generates a variety of standard reports. In addition, ARKS is useful for looking up simple questions about individual animals (for example, potential exposure to disease) as well as for mining the data to seek trends or clues to larger questions.

Most zoos send their ARKS data to ISIS at regular intervals so that it is backed up and also so that parts of the central database can be combined to create an international inventory. This international data is useful for various other purposes such as cooperative breeding programs, scientific research, or for governmental authorities. A weakness of ARKS is its limited ability to manage data for group-living or colonial species. Thus ARKS is of little use for managing animals such as schooling fish, many invertebrates, colony-nesting birds, or other animals that are kept in groups with multiple adult males and females. Similarly, animals that breed in relatively large enclosures (such as ducks in large ponds or fish in large tanks) are difficult to keep track of with individual records. Because of this weakness, animal records for aquaria and invertebrate collections are not standardized. As genetic identification becomes less expensive and more available, the possibility for tracking the genetics and demographics of large colony collections becomes much more feasible.

Zoo Population Management Strategies

Zoos have historically made much of their "babies," as births were and still are exciting events of much interest to the public. Prior to the development of management plans, rare breeding successes for a particular species tended to be replicated and replicated again. Then zoo curators would notice that there was a surplus of the species and they could no longer find good homes for the babies. Thus, all zoos would stop breeding the now-abundant species all at the same time. Some years later, as the stocks of the animal were aging and

starting to die off, there would again be room and demand for babies, and curators would try to resume breeding. But all too often, it would be too late for the aging stocks and the captive population would be lost.

To prevent these boom-and-bust breeding cycles, beginning with the launch of the Species Survival Plan (SSP) concept in 1980, zoo biologists have developed population management strategies and tools (Gibbons *et al.*, 1995). Key to zoo population management is gathering information about all the animals in the managed, captive population (the managed population often is not identical to the captive population because nonaccredited zoos and private collections are often not part of the cooperative program). The population data are kept in studbooks, which contain genealogical information as well as demographic information. As in zoo records keeping, ISIS distributes the most popular database for studbooks, a program called SPARKS, which stands for Single Population Analysis and Records Keeping System. The studbook programs integrate with computer software such as PM 2000, lately updated as PMx (PM = Population Management), that performs much of the demographic and genetic analyses that are essential to zoo population management.

Two principal kinds of conservation breeding programs are the current standards. The first is a program that participants essentially sign a contract to join, and agree to abide by the decisions of an elected committee. This type of program was first introduced in the United States, and was named the SSP; nearly identical programs are known by other names in other regions of the world. The second kind of program is a less intensive one, in which recommendations are issued but individual zoos elect whether or not to cooperate. In North America this second kind of program is known as a Population Management Plan (PMP). These two kinds of programs follow similar population management strategies but differ more in the intensity and politics of the actual management. The SSP also has conservation, research, and education components, whereas the PMP is strictly for population management. Recently, it became mandatory for AZA-accredited zoos holding SSP species to participate fully in the SSP-managed programs.

Planning for conservation breeding programs generally begins by determining the purpose of the program. Will animals from the population eventually be returned to the wild, or will the primary purpose of the population be for exhibition in zoos? The managers should decide what the appropriate population is for management purposes – for example, whether to limit the management to a subset of the animals that are known to derive from founders of the same subspecies, while excluding those that are of more uncertain ancestry. Once these decisions are made, demographic and genetic goals are set. Genetic and demographic analyses are done using the studbook data and specially designed software. The analyses are used to estimate parameters, which are needed for the long-range population forecasts that are used to guide the goals. Goal setting is done with the help of analytical tools that consider the number of founders (amount of genetic diversity that was likely to be captured), the effective population size N_e, the current population size, the rate of population growth, the number of years since the program was founded, and the generation time. Typically, conservation breeding programs aim to preserve greater than 90% of the original diversity for at least 100 years. Sometimes it becomes apparent that demographic intervention is needed, perhaps to increase the population growth rate, lengthen the generation time, or increase the number of spaces available so that the population can grow. At other times managers realize that more founders must be imported, if possible from zoos in another region of the world, but sometimes from nature.

As an aside, zoos have sustained much criticism from conservationists for taking animals from the wild for exhibition, including charges that this has led to the depletion of wild stocks. Zoos rebut that the number of animals they take are trivial compared with animals that die through habitat loss, hunting, or other human activities. Good captive management reduces the number of founders, or wild-born breeding stock, needed for long-term population viability. Theoretically, the acceptable number of founders is 12–25 individuals picked at random from a wild population (25 being ideal since each additional founder beyond 25 is unlikely to possess genes not already present in the 25). However, many captive populations start with only a fraction of this number. Even with good management, new founders may need to be added, since not all founders will produce the ideal number of offspring. Founders should be collected with minimal impact on the wild population. The way that this is best accomplished depends on the natural history of the species and the specifics of the program. For example, one low-impact approach to collecting founders is to salvage orphaned or injured nonreleasable wild animals; this opportunistic approach has the disadvantage that it may be slow and that disabled wildlife can constitute poor breeding stock. In extreme cases, such as those of the California condor and black-footed ferret, all of the remaining wild individuals were brought into captivity at the point it was calculated that the wild population was on a trajectory to zero. In both cases, this proved to be the salavation of the entire species. Descendants of these last remaining wild specimens became the ancestors of captive-born animals used to reestablish several new wild populations in the species' historic range.

Once the goals of the population are determined, and plans are made to redress problems such as too few founders, a detailed conservation breeding plan is developed. These plans give a recommendation for each animal in the managed population, which will include whether the animal should be moved, if it should be bred, which animal it should be paired with, and whether any special actions are needed, such as fertility testing.

Scientific principles are given a heavy weight in the planning. Zoo biologists tend to start with demography, using a Leslie matrix approach that incorporates both current and historic data. The demographic software tools were developed at the National Zoological Park by Jon Ballou and Laurie Bingaman-Lackey. Demographic data sets for zoos are often small, but contain highly detailed information compared to wild populations. Therefore zoo demographic estimates are often very crude and may require considerable tweaking. Early in the analysis, zoo biologists look at the population's age structure and sex ratio. If these are skewed, efforts may be needed to create a better demographic balance. The next

consideration is rates of fertility and infant and adult mortality; these estimates are used to project the number of breeding pairs that will be required in order to end up with the desired number of births that will balance deaths and also generate the recommended population growth or shrinkage.

"Genetic representation" of individual animals is the primary variable considered in deciding the animal-by-animal breeding priorities, except that aging, genetically underrepresented animals with only a few breeding years left may be given special breeding priority. Genetic algorithms, largely developed by the Chicago Zoological Society's Robert Lacy, use pedigree-based genetic simulations to assess the genetics of each animal and potential pairing in the population. Managers attempt to simultaneously maximize three parameters when recommending breeding matches: (1) good matches will improve the genetic diversity of the population (i.e., by increasing underrepresented lineages), (2) good matches will attempt to balance the genetic value of male and female, and (3) good matches will avoid inbreeding with close relatives. Balancing the genetic value, which is done by looking at a parameter called mean kinship, or the effective number of close kin, is important because it breeds big families with other big ones and little families with other little ones. New ways of visualizing mean kinship are now being used in PMx so that mismatching of animals with different mean kinship values becomes less important. This is especially useful in managing polygynous species where other difficulties can emerge if genetic management trumps social management.

Although the science of conservation breeding has advanced considerably, practical considerations also must enter into planning. Wild animal keeping in practice is a complex art, and it would be foolish to ignore details such as the distance of a proposed transfer and behavioral and physiological observations. For example, there is no point in recommending a pairing of two gorillas that are incompatible. Similarly, if an animal has not bred after many attempts at several zoos, there may be a physiological problem, or perhaps managers have not yet tried the right trick.

Group-breeding species are a special challenge to zoo population biologists, as they are for zoo animal records keepers. The methods outlined above work less optimally for populations in which a significant proportion of the pedigree is unknown, and also when managers cannot control which individuals mate with which, as is generally the case with group-breeding animals. An approach that is sometimes used is to completely ignore the pedigree and manage the species as a subdivided population, with occasional transfers of animals between demes or subpopulations. But the pedigree-free approach often throws out large amounts of information about a population and results in overly conservative management that will require larger populations than might otherwise be needed. Similarly, species whose life history is more size-dependent than age-dependent are not well served by current methods, and may be better managed using stage-based population management. As zoos move beyond management of animals with relatively simple breeding biology, such as tigers, and tackle species with more complex life histories, new population approaches will be needed. Similarly, as the "brave new world" of reproductive technology develops, frozen zoos will need to become integrated into population planning.

Frozen Zoos and Other High-Tech Solutions

An alluring prospect is that someday we may be able to preserve biodiversity in bottles on a warehouse shelf. Another enticing prospect is the use of reproductive technology to propagate animals that have not reproduced naturally. New technologies are making these prospects possible in some cases.

"Assisted reproduction" is a catchall phrase for a variety of veterinary or endocrine interventions used to promote breeding (Gibbons *et al.*, 1995). Artificial insemination, using fresh or frozen sperm, has been in use for many years in the domestic livestock industry and also for some zoo animals. The techniques can be readily adapted to zoo animals if there is a closely related domestic model in which the technique has been developed; for example, rare species of cows can benefit from research with domestic cows. Zoos have also employed other techniques, such as *in vitro* fertilization with harvested eggs, or embryo transfer both within and between species. The idea of growing a rare species in the womb of a common one is appealing, and has been achieved in some instances.

"Frozen zoos" are being developed by several zoos. In these superchilled repositories, sperm, eggs, and embryos are committed to long-term storage. Now that cloning is also a real possibility, frozen zoos of living cell lines are also being developed (Ryder and Bernirschke, 1997). Saving biodiversity in freezers is tantalizing, but there are many limitations to this approach and hurdles to be overcome.

Assisted reproductive techniques are difficult and costly to adapt to animals that do not have a closely related model, and also for those with complex reproductive physiology, for example, animals with induced ovulation. By and large, the techniques for animals with shelled eggs have been limited to artificial insemination. We cannot, as of yet, grow embryos in a test tube, so those derived from a frozen zoo must be grown in a real womb (or egg). This means a real population of surrogates will be needed to mother the frozen zoo, although in some cases these surrogate mothers can be of a different species. Surrogate mothers may produce animals with species-inappropriate behaviors and possibly even a hybrid phenotype.

Frozen zoos will not preserve learned animal behaviors (culture), and they will not eliminate the need to keep real populations. Frozen zoos, cloning techniques, and the like, as San Diego's Oliver Ryder predicts, will probably become powerful adjuncts to the management of populations. So far, genetic and demographic models for managing zoo populations have not factored in frozen resources, but some day they probably will. This will allow managers to keep smaller numbers of living animals. In addition, relative to the transfer of whole animals, the transfer of sperm, eggs, and embryos between wild and captive populations can be done with far less risk of cotransferring disease-causing pathogens. Thus, these techniques may be highly useful in importing new gene lines into small populations, both wild and captive.

Supporting Diversity in Nature

The Ark Paradigm

Enamored, perhaps, with their newfound ability to breed and manage exotic animals, zoos and their commentators were for

a time fixated on the "ark" paradigm of conservation. That is, like the biblical Noah, they talked about saving species in captivity until sometime in the future when the flood of human population shall recede and there will again be the space, and also the will, to restore Nature. Zoos have actually had some success saving species this way. The list of species that have been rescued from extinction by taking refuge for some time in zoos is ever-growing and includes black-footed ferret, California condor, European bison, Mongolian wild horse, Arabian oryx, Partula snails, and Guam rails. However, some species have become extinct during their captive sojourn and no opportunity has arisen to restore some others – for example, several bird species of Hawaii. We are seeing that restoration of a fauna may be only a remote possibility and the utility of reintroduction overstated in some ways. Where species have been rescued by breeding and reintroduction techniques, it is nothing short of miraculous and remains a model worth pursuing and perfecting.

The Post-Ark Paradigm

Maintenance of captive stocks has several functions, but the singular purpose is not, as it is popularly misperceived, for future reintroduction efforts. Zoos may indeed rescue individual animals and sequester them in the safety of captivity. As zoo professionals look beyond the ark, they are developing many other ways to contribute to the conservation of exotic animals.

Exciting, educational exhibits are potentially zoos' preeminent contribution to conserving animals and their habitats. Worldwide, zoos are estimated to attract over 600 million visitors annually. The majority of these visitors live in urban areas and otherwise have little contact with wildlife. Zoos are beginning to study and develop approaches for mobilizing visitors for conservation through innovative exhibit and educational programs. North America's accredited zoos report that each year 12 million students visit and enjoy onsite education programs, and in addition 85,000 teachers profit from teacher training. Some zoos are even developing "museum schools" in which a zoo or museum actually serves as an alternative school for a small number of students who typically combine rigorous college preparatory curricula with hands-on or applied learning in the zoo setting (more information about museum schools can be found at http://www.astc.org). Zoo-based conservation education programs need not end at the exit gates. Zoo educators and graphic artists can also develop conservation curricula that can be used in other parts of the world, including the native range of some endangered species. There is a remarkable dearth of educational materials about most endangered species within their native ranges, and any effort that zoos can make to redress this problem will be most helpful. Zoos may also need to develop curricula that more directly address the biodiversity crisis, and they may need to reach out more effectively to decision makers (politicians) and leaders of commerce.

Zoos can serve as refuges for displaced, injured, or confiscated wildlife. Thus, zoos help wildlife authorities enforce regulations by providing homes for seized wildlife. In addition, many zoos maintain patches of artificial or natural habitats for native wildlife and can provide significant breeding refuges for some birds and butterflies, and for fish and small animals that need relatively little space for a viable population. The Amphibian Ark, an international program (see http://www.amphibianark.org/), now addresses the amphibian crisis by engaging zoos in field research and conservation as well as inexpensive captive propagation programs.

Zoos sponsor significant basic research, often studies that would be difficult or impossible to accomplish with animals in the wild. Furthermore, zoos invest a considerable amount of money in research; for example, accredited zoos in North America spent $51 million on scientific research in 1998. While developing technology and methodology for captive animals under relatively controlled conditions, zoos create expertise that is useful *in situ* – for example, guidelines for medication dosages.

Direct financial support of field conservation is another way that zoos are helping preserve biodiversity (Gusset and Dick, 2011). Since 2006, accredited North American zoos have been involved in more than 3700 conservation projects in 100 countries, spending approximately $90,000,000. Another form of *in situ* conservation funding is for zoos in the wealthier nations to help improve the conservation impact of zoos in less-developed nations. Many zoos also sponsor member tours to see wildlife in nature, and they try to make arrangements that will have a positive impact on conservation.

The Ark Revisited

Human populations are still increasing and wildlife will continue to disappear. More direct conservation measures will be needed from all zoos. Zoos will be called on to provide refuge for more species as the wild populations dwindle. Shrinking wild populations in postage-stamp parks may need to be managed more like zoo populations, with infusions of unrelated lineages, potentially from zoo stocks. Intensive, joint, metapopulation management of *in situ* and *ex situ* wild animals will require both scientific and political innovation. Zoos may indeed help to save some wildlife from the flood of humanity, but they probably will not be acting as closed vessels. Rather, expect to see zoo-based conservation continue to become increasingly dynamic and interactive.

Appendix

List of Courses

1. Zoology
2. Conservation Biology
3. Population Biology

See also: Breeding of Animals. Captive Breeding and Reintroduction. Conservation Efforts, Contemporary. Education and Biodiversity. Ethical Issues in Biodiversity Protection. *In Situ*, *Ex Situ* Conservation. Natural Reserves and Preserves

References

Conway WG (1986) The practical difficulties and financial implications of endangered species breeding programmes. *International Zoo Yearbook* 24/25: 210–219.

Falk JH, Reinhard EM, Vernon CL, Bronnenkant K, Deans NL, and Heimlich JE (2007) *Why Zoos & Aquariums Matter: Assessing the Impact of a Visit to a Zoo or Aquarium*. Silver Spring, MD: Association of Zoos & Aquariums.

Gibbons, Jr. EF, Durrant BS, and Demarest J (eds.) (1995) *Conservation of Endangered Species in Captivity*. Albany, NY: State University of New York Press.

Gusset M and Dick G (2011) The global reach of zoos and aquariums in visitor numbers and conservation expenditures. *Zoo Biology* 30(5): 566–569.

Hoage RJ, Deiss, and William A (eds.) (1996) *New Worlds New Animals: From Menagerie to Zoological Park in the Nineteenth Century*. Baltimore, MD: Johns Hopkins University Press.

IUDZG/CBSG (IUCN/SSC) (1993) *The World Zoo Conservation Strategy; the Role of the Zoos and Aquaria of the World in Global Conservation*. Brookfield, IL: Chicago Zoological Society.

Norton BG, Hutchings M, Stevens EF, and Maple TL (eds.) (1995) *Ethics on the Ark: Zoos Animal Welfare and Wild-life Conservation*. Washington, DC: Smithsonian Institution Press.

Ryder OA and Benirschke K (1997) The potential use of cloning in the conservation effort. *Zoo Biol* 16: 295–300.

Shepperdson D, Mellen J, and Hutchins M (eds.) (1998) *Second Nature: Environmental Enrichment for Captive Animals*. Washington, DC: Smithsonian Institution Press.

Watters JV and Wielebnowski N (eds.) (2009) Special issue on zoo animal welfare. *Zoo Biology* 28(6): 501–651.

GLOSSARY

Abiotic Referring to the non-living components of an ecosystem; i.e., the chemical and physical environment.
Ablation A direct transfer of water molecules from ice crystals to the vapor phase, without transition through the liquid phase.
Abundance 1. The number of individuals of a given species in a region, area, or site. 2. Specifically, the average parasite load of an entire host population (including uninfected individuals).
Abyssal Referring to the deepest part of the ocean.
Abyssal plain A large, relatively flat area of the ocean bottom, typically 2000 to 6000 m in depth.
Acceptance A behavior pattern in which hosts treat brood parasitic eggs as if they were their own eggs.
Accidental epiphyte A plant that normally grows terrestrially but that occasionally grows to maturity on or in a tree.
Acclimation 1. Another term for ACCLIMATIZATION (see next). 2. Specifically, a process of acclimatization that occurs in a controlled, experimental setting, such as a laboratory.
Acclimatization Reversible physiological or morphological changes that an organism experiences in response to changing environmental conditions. Such physiological changes enable the organism to tolerate (acclimatize to) the new environmental conditions.
Accounting price See SHADOW PRICE.
Accretionary Accumulating; describing a gradual increase or growth, as of sand on a beach.
Acetylcholinesterase (AChE) The enzyme responsible for breaking down the neurotransmitter acetylcholine (ACh) at nerve synapses, thereby preventing hyperexcitation of cholinergic pathways in the nervous system. The target of organophosphate and carbamate insecticides.
Acid deposition Anthropogenic acidification of terrestrial and freshwater ecosystems by (primarily) sulfuric acid, derived from sulfur dioxide produced by burning oil and coal and deposited in rain and snow (**acid rain**), directly as particles (**dry deposition**) and as cloud droplets.
Acid neutralizing capacity (ANC) The amount of strong acid that is required to decrease the pH of water to a defined chemical end point. Acid neutralizing capacity is measured in units of $\mu eq\ l^{-1}$. It is a measure of the ability of water to neutralize inputs of strong acids.
Acidification The process of soil or water becoming more acidic, with a loss of base cations (calcium, magnesium, potassium) and an increase in hydrogen and aluminum ion concentrations, resulting in losses of plant biodiversity and dominance by acid-tolerant species.
Acidophilic Acid-loving; describing organisms that prefer or require a low pH environment. Such an organism is an **acidophile**.
Acidophytic, Acidophyte Alternate form of ACIDOPHILIC, ACIDOPHILE (see previous).
Acid-sensitive Describing an area with low supply rates of acid-neutralizing basic cations. These areas generally occur in high-elevation forest regions and are characterized by shallow soil.
Acquisition costs The costs of acquiring property rights to a parcel of land. This can be total (if the land and title are sold to a conservation agent) or partial. Partial transfers of property rights include short-term rental, conservation easements, and contracts between conservation agents and land owners who exchange money for land management that enhances conservation value.
Active restoration A process of actively assisting a system in recovery beyond simply removing the disturbance.
Acute Describing an exposure to an environmental stress that is brief in relation to the temporal scale of the biological system exposed.
Acute toxicity A damaging effect caused by a single short period of exposure to high concentrations of a pollutant.
Adaptation 1. In general, characteristics or behavior patterns of an organism that suit it to its conditions of existence. 2. Specifically (under the theory of descent with modification), a derived feature of an organism that increases the survival and reproductive success of the organism by virtue of the possession of that feature.
Adaptationism The doctrine that all important evolutionary processes are dominated by natural selection, and that all significant biological characters increase an organism's fitness.
Adaptive radiation The evolution of ecological and phenotypic diversity within a rapidly multiplying lineage.
Adaptive shift A change in the nature of a trait (morphology, ecology, or behavior) that enhances survival and/or reproduction in an ecological environment different from that originally occupied by the organism.
Adaptive syndrome A coordinated set of adaptations.
Additionality The achievement of environmental gains beyond what would have occurred under usual conditions.
Adipocytes 1. **Brown adipocytes**. Specialized thermogenic cells unique to mammals and probably derived from skeletal muscle, containing numerous mitochondria and small droplets of triacylglycerols. Under neural and endocrine control, brown adipocytes can switch rapidly from inactivity to the highest known metabolic rate, generating heat from oxidation of fatty acids and/or glucose in mitochondria uncoupled from ATP synthesis by uncoupling proteins. 2. **White adipocytes**. Large cells unique to vertebrates that store triacylglycerols, take up and release fatty acids, and produce various signal molecules and receptors that enable them to respond to circulating and locally acting agonists.
Adipose fin A small, fleshy fin without supporting spines or rays, set far back on the dorsal surface of many catfish, salmon, characins, and other groups.
Adipose tissue Storage tissue unique to vertebrates and best developed in tetrapods, consisting almost entirely of expandable adipocytes (see above) that contain a large droplet of triacylglycerols, along with vascular and neural tissue, and often contiguous or interchelated lymphoid tissues.

Advective precipitation A process in which moist air passes over cool ground by advection (wind) and is cooled.
Aeolian Wind-blown; relating to or caused by the action of the wind.
Aeolianites Wind-deposited sediments.
Aerenchyma Porous root tissue, especially well developed in wetland plants, that allows diffusive flux of oxygen from above-ground tissues to root tips. This tissue supports both the respiratory demand of the root tissues and allows oxygen to leak into the surrounding soil.
Aerial root In mangrove species such as Rhizophora, a root branching out from the stem some distance above the soil surface. Lenticels (pores) in the aerial portion of these roots enable gas exchange to take place, through aerenchyma tissue (see previous), with the respiring underground portions of the root.
Aerobe An organism that utilizes or requires the presence of oxygen for growth.
Aerobic 1. Occurring in the presence of oxygen. 2. Utilizing or requiring oxygen for growth.
Aerosols Microscopic airborne particles.
Affiliative behavior Behavior that is supportive and brings individuals together.
Afforestation The process of planting trees in areas currently devoid of them; e.g., in desert regions or a previously wooded area that had been converted to cropland or pasture.
Afrotropical One of the major zoogeographical regions of the world. See ZOOGEOGRAPHICAL REGION.
Agonistic behavior Behavior that is associated with fighting and competition in some way, as by individuals competing for the same resource; it may be aggressive, threatening, or submissive.
Agricultural intensification See INTENSIFICATION.
Agricultural land See PRIME AGRICULTURAL LAND; DEGRADED AGRICULTURAL LAND.
Agrobiodiversity 1. The variety and variability of living organisms that contribute to food and agriculture in the broadest sense, and that are associated with cultivating crops and rearing animals within ecological complexes. 2. More generally, all the organisms present in an agricultural landscape.
Agroecosystem Any ecosystem under agricultural management practices, as for purposes of food, fiber or biofuel production.
Agroenergy The production of biofuels as plant crops in agriculture, as distinct from biofuels produced synthetically or from algal culture.
Agroforestry A collective name for land-use systems and practices in which woody perennials are deliberately integrated with crops and animals on the same land management unit. This integration can be either in a spatial mixture or in a temporal sequence.
Agroscape A collective term for the agricultural, ranching, and plantation countryside, with its roads, irrigation ditches, buildings, and so on; contrasted with the wildland countryside that is not directly managed by humanity (though strongly impacted by it). The agroscape intergrades with wildlands in the form of woodlots, abandoned fields, poor soil sites, hedgerows, and edges of wildlands.
Air pollution Lowering of air quality due to the release of toxic materials by factories, automobiles, fires, and other human activities.
Airshed By analogy with a watershed, a definable area of the atmosphere that has commonality with respect to the dispersion of emissions; all parts of an airshed (e.g. the area around a power plant) are exposed to similar conditions of air pollution.
Alate An imago (adult insect) still possessing its wings.
Albedo The fraction of light hitting a surface that is then reflected from the surface.
Alevin A life cycle stage in salmon in which they are newly hatched with an unabsorbed yolk sac from which they obtain nourishment.
Algae (singular **alga**) Primitive plant-like organisms that generally obtain energy through photosynthesis. These organisms can be unicellular, filamentous, or colonial microscopic forms (**microalgae**), or they can be macroscopic (**macroalgae**) consisting of a primitive plant body (thallus) that lacks vascular tissue in most species.
Algal Having to do with or belonging to the algae (see previous).
Algal bloom A rapid increase in algae population in an aquatic system, freshwater or marine; so called in analogy with the blooming of flowers. Can result from higher concentrations of nutrients (particularly phosphorus and nitrogen) in waters, leading to increased growth of algae and green plants. A **harmful algal bloom (HAB)** is such an event involving toxic or otherwise harmful phytoplankton, which presents an environmental or public health threat.
Alien 1. A species or organism in a given area in which it does not naturally occur, present there because it has been intentionally or accidentally introduced from another region or ecosystem as a result of human activities. May also be described as *exotic*, *non-indigenous*, or *non-native*. 2. Having to do with or being such a species or organism.
Allee effect A positive association between average individual fitness and population size over some finite interval. For example, the growth rate of a population may be reduced at low densities, as from the increased difficulty of finding a mate. [Described by U.S. ecologist Warder Clyde Allee, 1885–1955.]
Allele One of two or more different forms of a gene.
Allelochemical A bioactive secondary metabolite with no known major function in primary metabolism but with the capacity to affect the growth, survival, or behavior of target organisms. Many allelochemicals act as chemical defense agents, although they may also mediate beneficial interactions with the targets.
Alley cropping See CONTOUR HEDGEROW INTERCROPPING.
Allochthonous Originating in a place other than that where it is now sited; e.g., outside the ecosystem where it is found.
Allometric Having to do with or exhibiting allometry (see next).
Allometry The relationship between a given biological feature and body mass. For example, basal metabolic rate varies with body mass raised to the power three-quarters, rather than in direct proportion with body mass, so that an animal twice as large does not need twice the amount of food to supply its metabolic needs.

Allopatric Describing populations, species, or taxa occupying separate, non-overlapping geographical areas.
Allopatric speciation The process of genetic divergence between geographically separated populations, leading to distinct species.
Allopatry The fact of populations, species, or taxa occupying different geographical areas.
Allopolyploid 1. Having two or more dissimilar chromosome sets derived from interspecific hybridization and chromosome doubling. 2. An organism of this type.
Alluvial fan An aggregation of fine sediments forming a conical fan when deposited by streams or rivers.
Alpha (α) diversity The diversity of species, often estimated as species richness, within a local community or site. Compare BETA DIVERSITY; GAMMA DIVERSITY.
Alpine Referring to the life zone above the climatic high-elevation tree line, irrespective of latitude. Though originally describing the Alps mountain range of Europe (from a pre-Roman word for high mountains), the term is now applied globally.
Altruism Behavior that reduces the agent's lifetime reproductive fitness while enhancing that of another or others in the same population. Such behavior is **altruistic**.
AMO Atlantic Multidecadal Oscillation.
Ammonites An extinct group of shelled invertebrates related to modern octopus, squid, and nautilus; their fossils are often referenced to identify geologic time periods.
Amniota The taxon of vertebrates that includes mammals, birds, and reptiles (including extinct relatives). So called because they possess a set of protective membranes for the embryo: the *amnion*, chorion, and allantois.
Amniote One of the Amniota (see previous).
Amniotic Relating to the amnion, or an organism possessing this (see AMNIOTA above).
Amphibian A member of a class of vertebrates (the **Amphibia**), comprising frogs and toads (order Anura), newts and salamanders (order Caudata), and caecilians (order Gymnophiona). Amphibians typically return to water to breed and pass through an aquatic larval stage with gills. They have a moist skin without scales, which is permeable to water and gases.
Amphibite A species requiring both surface waters and subterranean waters to complete its life cycle.
Amplification host A species that increases the rate of transmission of pathogens.
AMS date Accelerator Mass Spectrometry date; a radiocarbon date obtained by directly measuring the amount of carbon 14 in a sample with an accelerator mass spectrometer.
Anadromous Describing aquatic organisms that spawn and undergo early development in fresh water, migrate to sea to grow and mature, and then return to freshwater to reproduce; e.g., salmon, lampreys, sturgeons. Contrasted with CATADROMOUS.
Anadromy The fact of being anadromous (see previous).
Anaerobe An organism able to grow in the absence of oxygen.
Anaerobic 1. Occurring without the presence of oxygen. 2. Not requiring oxygen for growth.
Anagenesis Evolutionary changes occurring within an ancestor–descendant lineage. Contrasted with CLADOGENESIS.
Anamniotic Describing organisms (e.g., frogs) whose eggs are not surrounded by the complex protective membranes that distinguish the amniotic eggs of reptiles, birds, and mammals (the Amniota).
ANC Acid neutralizing capacity.
Ancestral 1. Describing a trait or condition that appears in an ancestor and also in a descendant species. 2. Describing taxa with such traits or conditions.
Anchialine Describing haline (salty) water, usually with restricted exposure to open air, and always with subterranean connections to the sea. Also, **anchihaline**.
Ancient lake Classification for a lake with a persistence of more than 100,000 years.
Andepts Soils derived from andesitic (high mountain) volcanic ash; so called because associated with the Andes mountains.
Angiosperms Flowering vascular plants in which the seeds, which are produced from the ovules of the flowers, are enclosed in fruits developed from the matured ovaries; the largest class of vascular plants, as opposed to cryptogams (mosses, etc., which do not produce seeds) and conifers (in which the seeds are not enclosed in fruits).
Anhydrobiosis An ability of certain organisms to lose virtually all water from their bodies under certain adverse conditions, yet remain viable when subsequently rehydrated.
Animal unit month (AUM) A measurement expressed as the quantity of forage needed to sustain one 1000-lb cow or five sheep for one month; used to manage grazing livestock.
Anion A negatively charged atom or molecule.
Annelid One of the Annelida, a large, relatively species-poor phylum of segmented worms.
Annual A plant that survives year to year as a seed rather than as a mature plant.
Anoxia A condition in which little or no oxygen is present.
Anoxic 1. Lacking oxygen. 2. Specifically, describing ocean or fresh waters that are depleted of dissolved oxygen.
Antennae (singular **antenna**) Paired sensory appendages on the head of arthropods. In crustaceans the first pair are also known as **antennules**.
Anthropocene The current era beginning with the Industrial Revolution, characterized by the increasingly significant impact of human activities on the Earth's biodiversity, land cover, biogeochemical cycling, water quality and availability, and other major features.
Anthropocentric Human-centered in perspective, especially with regard to the inherent value of the natural world.
Anthropocentrism The position that only human beings have moral worth or intrinsic value.
Anthropochoric Geographic distribution due to human agency.
Anthropogenic 1. Caused by humans; resulting from human activities. 2. Specifically, referring to environmental effects or changes caused or influenced by human activity; e.g., global warming.
Apex predator A predators that has no predator of its own and thus occupies a food web's highest trophic level.
Aphyllous Having no leaves; leafless; e.g., most cactus plants.
Apomictic Having to do with or involving apomixis (see next).

Apomixis (plural **apomixes**) A mode of clonal propagation by seeds, the embryos of which are totally genetic copies of the source plant. Seeds are produced without fertilization, but often pollination is required to induce apomictic seed production. Apomictic plants can also reproduce sexually, but this is very rare.
Apomorphic Describing derived characteristics, relative to a sister group.
Apomorphy A derived character or character state.
Aposematic Describing an organism that is rendered less susceptible to predation by somehow making known its obvious unpalatability, as by conspicuous odors, sounds, or coloration. The fact of doing this is **aposematism**.
Apparent competition An indirect interaction between two prey species in which one experiences more intense predation, because the presence of the other brings about a change in predator abundance or behavior.
Aquaculture The commercial farming of aquatic organisms, including fish, mollusks, crustaceans, and aquatic plants.
Aquifer A geological formation that contains groundwater and allows its movement.
Arboreal 1. Having to do with or involving trees. 2. Living in or on trees.
Arboricolous Growing on or living in trees, or off the ground in large woody shrubs or on fallen tree trunks.
Archaea Microorganisms that are similar to bacteria in size and simplicity of structure but very different in molecular organization, often living in extreme environments. They comprise (in addition to eukaryotes and bacteria) one of the three domains of life.
Archaic ungulate One of an extinct group or grade of mammals that possibly gave rise to hoofed mammals. Also known as a CONDYLARTH.
Area cladogram A branching diagram that uses phylogenetic relationships to trace geographical changes (either dispersal events or barriers to dispersal).
Area of endemism An area characterized by the congruent distributions of two or more groups of organisms.
Arid Describing a climate lacking in moisture, especially one having insufficient rainfall to support many trees or woody plants. Such a condition is **aridity**.
Arrhenotokous Describing a reproductive mode in which unfertilized eggs develop into haploid males and fertilized eggs develop into diploid females. This mode is **Arrhenotoky**.
Arthropod One of the Arthropoda, the phylum of joint-legged animals including insects, spiders, crustaceans, and others; the most species-rich phylum existing today.
Artificial ecosystem An ecosystem dominated by human artifacts and human-induced materials and energy flows.
Assemblage The collection of animals and plants found together in a patch of habitat.
Assisted colonization The translocation of species outside of their native range to allow them to track changes in climate.
Associational refuge A situation in which a potential prey escapes or deters a consumer by associating with another organism, thus interfering with the ability of the consumer to locate or attack the prey.
Assortative mating The tendency of like to mate with like, whether it is based on similarities in genotypes or similarities in phenotypes.
ATBI All taxa biodiversity inventory; the thesis, first suggested by D. H. Janzen, that it might be feasible to produce a complete species list for all the organisms living in one place; e.g., a hectare of tropical forest. This goal has thus far proved elusive.
Atlantic Multidecadal Oscillation (AMO) A mode of variability occurring in the North Atlantic Ocean, principal expressed in the sea surface temperature (SST).
Atmospheric deposition The process by which materials are deposited from the atmosphere to the Earth's surface.
Atmospheric lifetime A time scale describing the rate of removal of a gas from the atmosphere, generally defined as the time for a 63.2% decrease (the e-folding time) in the global atmospheric burden of the gas; a period of many years for major greenhouse gases.
ATP synthetase A membrane-associated protein complex that couples exergonic fluxes of $H+$ across the membrane to ADP phosphorylation, or vice versa.
AUM Animal unit month.
Australasian One of the major zoogeographical regions of the world. See ZOOGEOGRAPHICAL REGION.
Autapomorphy An advanced or derived character state unique to an individual taxon (of any rank).
Autocatalysis The self-acceleration of certain chemical reactions that yield catalytically active products. The reaction rate of autocatalytic processes increases exponentially.
Autochthonous Referring to the production of organic matter within an ecosystem, for example, primary production by aquatic plants.
Autogenic Describing a species that creates habitat for other organisms, in a habitat that would not exist for them if the autogenic species were not present.
Autotroph Literally, a self-feeder; an organism that is able to utilize inorganic carbon (carbon dioxide) as the sole carbon source for growth; for example, green plants and certain bacteria. Such an organism is **autotrophic**.
Autotrophy The ability of organisms to grow and reproduce independent of external sources of organic carbon compounds. Compare HETEROTROPHY.
Avirulent Not virulent; describing a parasite that has little impact on its host's fitness.
AZA Association of Zoos and Aquariums, a North American professional organization.
Background extinction (rate) The natural rate of extinction over geological time, measured from the fossil record and independent of anthropogenic influence; characteristic of the major part of the evolution of life.
Background selection Selection that takes place at loci other than that of interest. Such events can strongly increase the intensity of genetic drift as well as hinder the action of selection at a locus of interest. These effects are normally modeled by reducing the effective population size applicable to the locus of interest.
Bacteria (singular **bacterium**) An entirely microbial domain of single-celled organisms, existing in various shapes and generally found throughout the Earth as well as in organic matter and the living bodies of humans and other organisms. They are evolutionarily distinct from Archaea, and one of the three domains of life along with Archaea and Eukaryota.
Bacteriophage A virus that infects bacteria.

Baiji A freshwater marine mammal, *Lipotes vexillifer*, living in China's Yangtze River; last verifiably observed in 2002 and now regarded as extinct after an extensive 2007 search failed to find any specimens. This would be the first known extinction of a cetacean species in recorded human history. Also known as a **Yangtze (River) dolphin** or **Chinese river dolphin**.

Baleen Modified, hardened epithelial structures suspended from the upper jaws of mysticete whales, having the form of multiple elongate plates with a fibrous fringe on the inside edge. They allow whales to filter zooplanktonic or fish prey from sea water.

Band The basic economic, social, and political unit of a hunter-gatherer society.

Barophile An organism able to survive in conditions of extreme atmospheric pressures, at great depths of the ocean or solid earth. **Barotolerant** organisms can survive at up to 400 atm below the water or earth, but grow best at 1 atm. **Barophilic** organisms grow best at higher pressures, up to 500–600 atm. **Extreme barophiles** grow best at depths of 700 atm or more. (Each gradient of one atmosphere in the ocean is equal to 10 feet below the surface.)

Barophilic Literally, pressure-loving; describing an organism able to endure high atmospheric pressures; e.g., certain microorganisms living at great ocean depths. See BAROPHILE (previous).

Barotolerant See BAROPHILE.

Basal 1. Designating a species that eats no other species and thus represents the first part of a food web. 2. Having to do with or describing such a species: *basal* plants.

Basal metabolism The minimum amount of energy expended by an adult animal that has not eaten recently, at normal body temperature during rest (usually sleep).

Basement complex The layer of highly metamorphosed sedimentary rocks, mainly of late Pre-Cambrian age, that underlies much of tropical Africa.

Batesian mimicry A form of mimicry in which the target organism is rendered less susceptible to predation by its resemblance in morphology or coloration to a different species that is dangerous or unpalatable to the predator. [Described by English naturalist Henry Walter Bates, 1825–1892.]

Batha A traditional term for semishrub communities, mainly of seral vegetation of old field succession, in the Mediterranean area of the Near East.

Beneficial management practice (BMP) Any of various methods to prevent or reduce the movement of sediment, nutrients, pesticides, and other chemical contaminants from the land to surface or groundwater, or otherwise to protect water quality from the adverse effects of agricultural activities.

Bennettitales (Cycadeoidales) An order of extinct gymnosperms composed of two families having both unisexual and bisexual reproductive structures.

Benthic 1. Relating to or located on the bottom of the sea or another aquatic environment such as a freshwater lake, river, or stream. 2. Living at the bottom of a body of water: a *benthic* organism.

Benthos 1. Those organisms living on, in, or near the seabed or at the bottom of some other aquatic environment, including those that reside on hard and soft bottom surfaces and others that reside between sediment grains. 2. The area at the bottom of the sea or another body of water.

Beringia The so-called Bering Land Bridge, a region at the northwestern corner of North America separated from Asia only by the shallow waters of the Bering Strait. When ocean levels declined during glaciation, Asia and North America were connected by land in this region; thus it is theorized to be the means by which humans migrated from Asia to the Americas.

Beta (β) diversity The degree of turnover in species (and changes in their abundances) among communities or sites along a gradient or within a larger area. Compare ALPHA DIVERSITY; GAMMA DIVERSITY.

Bet-hedging A strategy of an organism, living in a variable environment, that leads to low variation in fitness. In general, such a trait provides the organism greater net fitness over a range of environmental conditions than would a trait specialized for a single environment. A **bet-hedging trait** is expected to evolve when the environment in which a species lives fluctuates over a fixed range of conditions that is sufficiently broad for fitness to vary significantly, and when it is unpredictable which state the environment will take in the immediate future.

Between-habitat diversity The degree of dissimilarity in species composition between two habitats; the greater the number of species shared by two habitats, the lower the between-habitat diversity.

Big bang reproduction See SEMELPARITY.

Binomen A two-part name of a species, consisting of the name of the genus in which it is placed, followed by a word peculiar to that species.

Binomial nomenclature The scientific practice of naming species by a two-part combination of a specific epithet preceded by the genus to which the species belongs; e.g., *Ectopistes migratorius*, the passenger pigeon.

Bioaccumulation The accumulation of toxicants or contaminants, such as mercury, by successive organisms in a food chain.

Bioassay The use of living cells or organisms to make measurements of the amounts or activity of toxins or other substances.

Biocentric Valuing the existence and diversity of all biological species.

Biocentrism The belief that all living beings have moral worth or intrinsic value.

Biochemical oxygen demand (BOD) A measure of the amount of dissolved oxygen required by the microbial community to decompose organic matter by aerobic metabolism.

Biocontrol See BIOLOGICAL CONTROL.

Biocultural axiom The concept that biological and cultural diversity are mutually dependent and geographically coterminous.

Biocultural diversity The diversity of life in all its manifestations, in both the biological and cultural realms, which are interrelated (and likely have coevolved) within a complex social–ecological adaptive system.

Biodiesel A liquid biofuel used as a substitute for, or an additive to, standard diesel fuel.

Biodiversity (abridged from **biological diversity**) The full diversity of life on Earth, encompassing all levels of biological

organization, from molecules to genes, to populations, to species, to ecosystems and the entire biosphere. It is a multidimensional concept difficult to measure, and in practice most focus is on diversity at the species level.

Biodiversity assessment The identification and classification of species, habitats, and communities within a given area or region; can be done to evaluate whether management is necessary to conserve biological diversity.

Biodiversity hotspot See HOTSPOT.

Biodiversity offset A measures intended to compensate for adverse biodiversity impacts, with the goal of no net loss and preferably a net gain of biodiversity of species composition, habitat structure, and ecosystem function.

Biodiversity restoration See RESTORATION.

Bioelectricity 1. An electric current generated by living tissue. 2. Electricity produced by burning biomass.

Bioenergetics The flow and transformation of energy within organisms, or between organisms and their environment.

Bioenergy A form of energy that can be extracted from biomass, including biodiesel from fats and oils, bioethanol from sugars, starches, and lignocellulosic material, and bioelectricity and bioheat from the combustion of lignocellulosic material.

Bioethanol Alcohol made from sugar (e.g., sugarcane), starch (e.g., corn grain), or cellulosic materials (e.g., cornstalks, wheat straw, prairie grasses, and trees) for use as a biofuel; sugar and starches are converted to ethanol through fermentation, while cellulosic material is converted via enzymatic hydrolysis or synthesis gas fermentation.

Biofuel Any of a variety of fuels in some way derived from biomass, including firewood, methane, or liquid transportation fuels obtained from plants or other biological materials.

Biofuel land-use change The conversion of forest land and pasture to cropland for the purpose of growing plants for biofuel. The anticipated reduction in GHG emissions resulting from the substitution of biofuel for traditional fossil fuels is thus offset by release of carbon from biomass and soils, and there are a variety of other impacts such as loss of biodiversity and overexploitation of water resources, as well as market effects such as higher food prices (and possible food scarcity) due to reduction in the amount of land devoted to food crops.

Biogeochemical processes The various chemical, physical, geological, and biological processes that collectively govern the composition of the biosphere.

Biogeochemistry The scientific study of biotic controls on the chemistry of the environment and geochemical controls on the structure and function of ecosystems.

Biogeographic realm One of the largest divisions of the historic and evolutionary patterns of terrestrial plants and animals. In such an area, plants and animals evolved in the isolation created by barriers such as oceans, mountain ranges, or deserts. It can be a single continent, or it can span continents.

Biogeographic region 1. A large region with characteristic assemblages of natural communities and species, distinct from other regions. 2. See BIOGEOGRAPHIC REALM.

Biogeography The study of the geographic variation of nature, including variation in any biological characteristics (e.g., body size, population density, or species richness) on a global scale. **Historical biogeography** studies this variation over time.

Bioheat 1. A fuel or heating system based on or employing a biological source. 2. Specifically, a fuel composed of conventional heating oil blended with a percentage of biodiesel (typically 2–20%).

Biological assay See BIOASSAY.

Biological carbon pump See CARBON PUMP.

Biological control The use of naturally occurring or introduced natural enemies (e.g., predators and parasitoids) to control agricultural pests, as opposed to non-natural controls such as insecticide sprays.

Biological diversity See BIODIVERSITY.

Biological energetics See BIOENERGETICS.

Biological integration The fact of life functions being organized in distinct, self-regulating units (cells, tissues, organs, organisms, populations, communities, and ecosystems).

Biological soil crust The combination of non-vascular and non-seed-forming plants that commonly occur on the soil surface of many arid and semiarid ecosystems. May consist of mosses, lichens, liverworts, green algae, and cyanobacteria.

Bioluminescence The production of light by living organisms; e.g., fireflies, anglerfish.

Biomass 1. The combined mass or weight of all the members of a given population or stock at a given time, or, on the average, over a certain time period. 2. The mass of biological material remaining after the removal of water from it by oven-drying at 70–100 °C; often expressed as mass per area of ground surface. **3**. A collective term for all organic substances that can be used as energy sources, including wood, straw, agricultural and industrial plant residues, municipal organic wastes, and crops grown specifically for fuel. Applies to relatively recent (non-geologic) fuel sources, excluding fossil fuels.

Biomass pyramid A graphic representation of the biomass relationships of a community, in which the total amount of biomass at each successive trophic level is proportional to the width of the pyramid at the appropriate height.

Biome A major terrestrial ecological community and landscape type, characterized by more or less uniform physiognomy of its potential natural vegetation and with characteristic fauna and flora, such as the tropical rainforest, the warm deserts, or the temperate grasslands.

Biomimetic 1. An artificial compound, device, or structure that mimics or was inspired by nature. 2. Imitating or modeled after nature or a natural process. **3**. Having to do with or employing biomimetics (see next).

Biomimetics The study of natural structures, processes, and systems as a model for human technology. Also known as **biomimicry**.

Bionomic equilibrium In a biological common property resource, an equilibrium between resource production rate and exploitation rate, characterized by the dissipation of economic rents.

Bionomics The mode of life of a species.

Biophilia The innate tendency of humanity to focus on life and lifelike processes and to have an affinity for the natural world, said to have evolved over millennia and to be part of human inheritance.

Bioquads An occurrence record of organisms used in biodiversity research, consisting of four elements: species names, location, time, and source.
Bioregion Another term for a BIOGEOGRAPHIC REALM or REGION.
Biosafety Efforts to reduce and eliminate the potential risks of biotechnology and its products.
Biosphere 1. All the plants and animals on Earth, together with the habitat they require for long-term survival. 2. The portions of the Earth inhabited by living beings.
Biosphere reserve A specific area of land or water identified as part of an international conservation program to protect representative ecosystems throughout the world in ways compatible with sustainable development. The program is administered by the United Nations World Network of Biosphere Reserves.
Biosynthesis The biochemical process by which a plant produces molecules of utility or adaptation.
Biota Fauna and flora; the totality of organisms in a given place.
Biotic 1. Having to do with the living components of an ecological community (organisms, species). 2. Having to do with life; living.
Biotic homogenization the process by which the genetic, taxonomic, or functional similarities of regional biotas increase over time.
Biotic impoverishment The decline of a species to the point of having little or no ecological significance.
Biotic interaction A situation in which the growth or behavior of one species affects another species; such interactions can be antagonistic (e.g., competition for limiting resources, predation, herbivory, or parasitism) or facilitory (e.g., pollination or other forms of mutualism).
Biotic similarity A measure of the degree to which two or more samples or assemblages are similar in species composition. This can be expressed by any of various **biotic similarity indices**.
Biotope A region that is distinguished by particular environmental conditions (climate, soil, altitude, etc.) and thus contains a characteristic assemblage of organisms.
Biotopic Having to do with or constituting a biotope (see previous).
Biramous Relating to or having an extension of the body (usually an appendage) with two branches.
Bivoltine 1. A species with two generations per year. 2. describing such a species.
Black box An informal term for an entity that can be examined at the system level without specifying its internal contents.
Black-water A term for Amazonian rivers that are dark because of dissolved tannic matter and are acidic (pH about 4).
Bleaching The loss of symbiotic zooxanthellae (photosynthetic algae) from corals. Bleaching is usually caused by elevated sea surface temperatures, but it can also be caused by low temperature, sedimentation, extreme salinities or light levels, and bacterial infection.
Bloom See ALGAL BLOOM.
Blubber Specialized superficial adipose tissue found only in pinnipeds (seals), cetaceans (whales and dolphins), and sirenians (sea-cows), serving as thermal insulation as well as for energy storage.
blue–green algae Former name for the CYANOBACTERIA.
BMP Beneficial management practice.
BOD Biochemical oxygen demand.
Bog An acidic peat land that receives all water inputs from the atmosphere and hence has low pH (usually 3.4–4.6), low alkalinity, and negligible concentrations of base cations in surface waters; typically dominated by mosses of the genus Sphagnum and other plants that are able to tolerate low pH and low nutrient availability; i.e., the heath family (Ericaceae).
Booklungs In arachnids, one to four pairs of abdominal respiratory organs consisting of a thin, multi-folded membrane (the book's "pages") over which blood circulates and that is open to an air-filled cavity on the outside, itself open to the exterior via a spiracle. Gases passively diffuse back and forth across the membrane.
Boreal Relating to the northern high latitudes (but not polar), regions that contain large continental landmasses and thus continental climates with moderate summers and long, severely cold winters. [From Greek *boreas*, the north wind.]
Boreal forest A forest zone dominated by coniferous species, covering the northern latitudes of North America, the Nordic countries, and Russia.
Bottom-up 1. Describing a process in which physical parameters (such as nutrients) allow increased primary productivity and the effects of this cascade up through higher trophic levels; e.g., nutrient additions to a lake resulting in increased phytoplankton growth, this increasing resource of food allowing increased number of herbivorous zooplankton, and the abundance of zooplankton leading to increased densities of zooplankton-eating fishes. 2. Describing impacts or controls on urban biodiversity that originate from human actions at smaller scales, such as household or neighborhood scales. Thus, **bottom-up effect** or **force**. Compare TOP-DOWN.
Brachyodont A low-crowned cheek tooth commonly found in omnivorous or burrowing animals.
Brachypterous Having rudimentary or unusually short wings.
Brackish Describing water that is saline but less salty than seawater, as when fresh and salt water mix in an estuary.
Breed 1. A subgroup within a species distinguished from similar subgroups by identifiable external characteristics, typically developed by managed reproduction; e.g., the Brahman or Angus cattle breeds. 2. To cause an animal or plant to produce offspring in a controlled manner.
Brood parasite An organism (e.g., certain birds) that makes use of individuals of the same or another species for the rearing of its young; a strategy known as **brood parasitism**.
Broodstock Fish or shellfish from which a first or subsequent generation may be produced in captivity, whether for growing as aquaculture or for release to the wild for stock enhancement.
Brown adipocytes See ADIPOCYTES.
Browse 1. Leaves, twigs, or reproductive parts of shrubs, woody vines, and trees available for animal consumption. 2. The act of consuming such plant parts. 3. To consume plants in this manner.
Bryophyte A nonvascular plant of the division Bryophyta (a moss, liverwort, or hornwort).

BS Bundle sheath.
Bulk density The mass of undisturbed dry soil per unit volume; a measure of soil compaction.
Bundle sheath (BS) tissue A cell layer at the periphery of the vascular bundles in leaves. In C4 plants, the reactions of the photosynthetic carbon reduction cycle are localized in this layer, in contrast with C3 plants where they occur in the mesophyll tissue. In comparison to C4 species, the BS cells of C3 species are small, with few chloroplasts.
Bush-fallow cropping An agronomic system in which soil fertility is maintained by allowing native vegetation to regenerate following several years of cropping.
Bushmeat Primates or other wild animals hunted illegally for food or income.
Bycatch Unwanted fish or other living creatures caught unintentionally in fishing gear during an effort to catch something else, and then discarded; also called **bykill** if discarded dead.
C3 plant A plant in which carbon dioxide is first fixed into a compound containing three carbon atoms before entering the Calvin cycle of photosynthesis. Most trees and temperate plants are C3 plants.
C4 plant A plant in which carbon dioxide is first fixed into a compound containing four carbon atoms before entering the Calvin cycle of photosynthesis. Most tropical grasses are C4 plants.
Caatinga A native South American word for an open place; applied to semi-arid ecosystems of eastern South America.
CAFO Concentrated animal feeding operations; a feeding facility that confines animals for more than 45 days in an area that does not produce vegetation during the growing season.
Calanoid copepods Zooplankton belonging to the Crustacea, very abundant in ocean waters. They play a major ecological role in the food web of the pelagic ecosystem as grazers of phytoplankton and as a food source (e.g., for larval and adult fish).
Calcareous 1. Mostly composed of calcium carbonate ($CaCO_3$). 2. Specifically, describing soils with high calcium carbonate content, which buffers them against changes in pH.
Calcareous fen 1. A peat-accumulating wetland with surface deposits of calcium carbonate and a distinctive flora of calciphilic species. Such wetlands typically develop where substantial amounts of cold, calcium-carbonate rich groundwater discharge to the surface. 2. More generally, any wetland with a high calcium carbonate content in its water and soils.
Calciphilic Describing a plant or other organism that thrives in a calcium-rich environment. Such an organism is a **calciphile**.
Cambrian explosion A term for the period of time from about 570 million to 530 million years ago during the Cambrian geological period, when the diversity of multicellular life greatly increased in a relatively short time (by the standards of geologic time). Also known as the **Cambrian radiation**.
CAM metabolism Crassulacean acid metabolism; a carbon uptake system utilized by many succulent plants. These plants fix carbon dioxide during the night when evapotranspiration is relatively low, storing it as organic acids. The stored CO_2 is then released during the day and refixed in the typical light reactions of photosynthesis, thereby increasing water use efficiency.
Canopy The highest tree layer in the vertical stratification of forests. The herbaceous foods for canopy-dwelling animals are found among the branches and leaves of the canopy.
Cap and trade A market-based incentive for businesses or other entities, setting an environmental target such as a maximum amount of pollution or minimum quantity of habitat preservation (the "cap"), and allowing this compliance obligation to be exchanged across firms to meet the objective at lowest cost (the "trade").
Capoeira A Brazilian term for secondary forest on cleared ground.
Captive breeding Human-mediated breeding and rearing of wildlife under controlled conditions, as in a zoo.
Carbon capture See CARBON SEQUESTRATION.
Carbon cycle The process by which carbon moves among its major reservoirs (atmosphere, terrestrial biosphere including freshwater systems, soil, oceans, sediments, and so on) by various chemical, physical, biological, and geological processes. These movements may be natural flows (e.g., plant respiration) or the result of human activity (e.g., burning of fossil fuels). The carbon cycle plays a key role in global climate because carbon dioxide is a major contributor to the greenhouse effect (see).
Carbon footprint 1. The total amount of carbon dioxide and other greenhouse gases produced to directly and indirectly support human activities. 2. The amount of this associated with one specific entity, such as a nation, city, building, or residence, or an individual person. See also ECOLOGICAL FOOTPRINT.
Carbon pump The biologically mediated transfer of carbon in the Earth's oceans from the euphotic (uppermost) zone to the ocean's interior, maintaining a vertical gradient in inorganic carbon concentrations. Major processes involved in the pump include photosynthetic consumption of inorganic carbon and production of organic matter in the euphotic zone, gravitational settling of particulate organic matter, and subsequent remineralization of organic matter back to inorganic carbon at greater depths.
Carbon sequestration The process of removing carbon dioxide from the atmosphere and rendering it inaccessible to carbon-cycling processes (and thus mitigating climate change), as by reforestation.
Carbon sink See SINK.
Carboxylase An enzyme that catalyzes the formation of a C–C bond between inorganic C (CO_2 or HCO_3^-) and an organic molecule.
Carnivorous Describing an organism that obtains its food by consuming other animals, especially one that does so exclusively or mainly in this manner. Such an organism is a **carnivore**.
Carnivory The fact or process of feeding exclusively or mainly on the flesh of other animals.
Carrying capacity The maximum population that can be sustained indefinitely in a given area without changing the ecosystem in ways that will eventually reduce the sustainable population. This balance between population and resources is a dynamic one that involves changes in technology and other factors.

Caste A group of individuals in a colony that are both morphologically distinct and specialized in behavior.
Catadromous Describing aquatic organisms that spawn and undergo early development in the ocean, migrate to fresh water to grow and mature, and then return to the ocean to reproduce; e.g., freshwater eels, some mullets. Contrasted with ANADROMOUS.
Catadromy The fact of being catadromous (see previous).
Catch The total amount of fish or other aquatic organisms landed during a certain period by the operation of fishing gear. This includes fish discarded at sea (*bycatch*), or caught by lost or abandoned gear (*ghost fishing*).
Catena A regularly recurring sequence of soil and vegetation types developed on an undulating landscape.
Cation A positively charged atom or molecule.
Cation exchange capacity (CEC) The capacity of a soil to adsorb cations such as calcium, magnesium, and potassium on the surfaces of soil particles and organic matter.
Caudal rami (singular **ramus**) In some crustaceans, a thin branching extension from the posterior segment of the body, usually paired.
CBD Convention on Biodiversity.
Cellulosic ethanol Ethanol produced by means of the conversion of cellulose, the primary structural molecule of most plants, into glucose that can then be fermented into ethanol.
Centinelan extinction The disappearance of a species before its formal recognition by science.
Cerci In insects and certain other arthropods, paired, usually unsegmented structures at the end of the abdomen, sometimes used by males during mating to grasp the female's abdomen.
Cerotegument A hydrophobic secretion layer covering the body of certain arthropods adapted to temporary submersion.
Cerrado A Brazilian term for the large areas of savanna and savanna forests that are the dominant vegetation of the Planalto of Central Brazil.
Cetacean One of the Cetacea, the order of mammals that includes the whales, dolphins, porpoises, and allies.
CFCs Chlorofluorocarbons.
Chaco A region located between the Paraná basin, the central Brazilian shield, and the Andes. Ecosystems present in the Chaco region include temperate grasslands, savannas, and arid and semiarid environments.
Chaetae (singular **chaeta**) Bristles composed of chitin cylinders held together by scleroprotein emerging from parapodia or the body wall in many annelids.
Chain-termination sequencing A method of DNA sequencing that employs four reactions, one for each nucleotide, and terminates the copy at each occurrence of a given base.
Change matrix A method of classifying land use change over a given time period, by following the exact path of change; i.e., how much land was transferred from a given use to other uses, how much came into that use, and what remained unchanged during the period.
Chaparral Evergreen sclerophyllous-leaved shrublands that cover large areas of California.
Character An attribute that is constantly distributed among all members of a species or a clade.
Character displacement Divergence in a morphological character between two species when their distributions coincide in the same ecological environment, compared to overlapping of the character when the two species are geographically separated.
Charismatic species See FLAGSHIP SPECIES.
Chela (plural **chelae**) A pincer-like claw, as in crustaceans and arachnids.
Chelate Having a pincer-like claw.
Chelicerae (singular **chelicera**) Mouth parts of the Chelicerata, a subphylum of arthropods including arachnids, horseshoe crabs, and sea spiders. They are at most three segmented appendages, usually two, and typically the distal segment acts against the penultimate to grab or hold prey or objects.
Chemical defenses Compounds used by plants to deter or poison herbivores and pathogens.
Chemoautotroph Literally, a chemical feeder; an organism able to synthesize organic compounds by the oxidation of energy-rich inorganic sources, without the use of light energy. Such an organism is **chemoautotrophic**.
Chemoautotrophy The use of energy-yielding chemical reactions as an energy source for the synthesis of organic matter from inorganic precursors.
Chemolithotroph, Chemolithotrophy See CHEMOAUTOTROPH, CHEMOAUTOTROPHY.
Chemoorganoheterotroph An organism that requires organic substrates from its environment for its for growth and development. Such an organism (e.g., herbivores, carnivores) is **chemoorganoheterotrophic**.
Chemoorganoheterotrophy A form of metabolism that relies on oxidation of organic compounds for energy, reducing equivalents, and carbon.
Chemostat An apparatus for growing microorganisms in a continually replenished medium.
Chemotroph, chemotrophy See CHEMOAUTOTROPH, CHEMOAUTOTROPHY.
Chlorofluorocarbons (CFCs) A category of hydrocarbons that are nonreactive, non-flammable organic molecules in which both chlorine and fluorine atoms replace some of the hydrogen atoms. Also known by the brand name **Freon**. Once widely used for various industrial purposes such as aerosol propellants, refrigerants, and cleaning solvents, CFCs have now been phased out because of their tendency to deplete the Earth's ozone layer.
Chondrichthyeans Sharks and their relatives.
Chordate A member of the group Chordata. This includes the most recent common ancestor of tunicates and cephalochordates and all of that ancestor's descendants. Tunicates, lancelets, hag-fishes, and vertebrates are all chordates.
Chromatin A complex of DNA and associated proteins that when assembled make up the chromosomes of eukaryotes.
Chromophore An organic molecule that absorbs photosynthetically active radiation and either carries out photochemistry or transfers excitation energy to another chromophore that carries out photochemistry. It is always bound to a polypeptide in the photosynthetic apparatus. Such a molecule is **chromophoric**.

Chromosome A long, threadlike structure in the nucleus of a cell, composed of DNA and protein.
Chronic toxicity A damaging effect caused by a long period of exposure to low or moderate concentrations of a toxin.
Chronosequence A series of sites that differ in age, but otherwise occur on similar soil types and environmental conditions within the same climatic zone.
Circannual Describing a phenomenon having a recurring cycle of approximately one year. **Circannual rhythms** recur with such a pattern.
Clade A group of organisms, such as a species, that all share a common ancestor.
Cladistic Having to do with or based on cladistics (see next).
Cladistics A method of classifying organisms based entirely on monophyletic taxonomic groupings within a phylogeny; i.e., the grouping includes all descendants of a common ancestor and no others, and taxonomic units that are paraphyletic or polyphyletic are excluded. A **cladist** is someone who practices cladistics.
Cladistic biogeography The combination of cladistics (see previous) with vicariance biogeography; a method that searches for patterns of relationship among areas of endemism.
Cladogenesis Branching evolution; the origin of diverging new forms from an ancestral lineage. Contrasted with ANAGENESIS.
Cladogram A general graphic representation of estimated cladistic relationships among taxa; it has no direct connotation of ancestry and the long axis does not connote time.
Clear-water A term for Amazonian rivers that have neither mud nor humic matter and are usually slightly acidic.
Climate change 1. Changes in global temperature, rainfall, and other aspects of climate (occurring or predicted). 2. Specifically, such a change resulting from or exacerbated by increased human production of carbon dioxide and other greenhouse gases. See also GLOBAL WARMING.
Climate sensitivity Long-term change in global mean surface temperature following an increase in carbon dioxide (CO_2) concentrations in the atmosphere, especially the significant increase since the onset of the industrial era (post-1750).
Climate teleconnections A term for the fact of climate anomalies being related to each other at large distances.
Climatic model A mathematical description of the flows of energy, momentum, and materials around the atmosphere and oceans, typically run on large computers to stimulate the Earth's climate, both naturally and when disturbed by greenhouse gases or another forcing or disturbance.
Climax (community) According to some theories of succession, the end result of succession in which successional change ends with a community that does not change and that is in equilibrium with the climate.
Clique A group of organisms in a food web in which any pair of species in the group shares at least one prey species.
Cloaca The common chamber in which the reproductive, excretory, and digestive tracts of amphibians unite before exiting the body.
Clonal 1. having to do with or being a clone (see next). 2. Specifically, describing a vegetative mode of propagation and expansion of plants by runners, tillers, or plant fragment dispersal, very important in the alpine life zone. Clonal plants also produce sexual offspring by seed, but their clonal propagules often show higher survival. All clonal offspring of one source plant have the same genome.
Clone 1. A process of producing similar populations of genetically identical individuals through asexual reproduction, as in bacteria, insects, and plants. 2. A group of individual plants formed by vegetative reproduction from a parent. The term applies particularly to those trees and plants that reproduce by creeping stems or by suckers from roots, forming circular patches of genetically identical individuals. **3**. A population of cells all descended from a single cell, or a number of copies of a single DNA fragment. 4. A process of artificially producing a multicellular organism that is genetically identical to another; e.g., a sheep or cow.
Cloning Any process in which a clone is produced, either naturally or intentionally (see previous).
Clupeid Any of various pelagic fish of the family Clupeidae, including the anchovy, sardine, and herring, which feed on plankton.
Cnidaria The marine invertebrate phylum containing the reef-building corals.
Coadapted gene complex Chromosomes, loci, or genes that are adapted to function well together.
Coalescent Looking back in time through a genealogy to a point where two genetic lineages come together in a common ancestor.
Coastal Relating to areas including estuaries, semi-enclosed seas, and shallower regions of the ocean, including areas influenced by rivers and runoff from land.
Coastal sage scrub Semiwoody shrublands dominated by drought-deciduous species, occurring in semiarid areas along the coast of southern California and the interior transition from chaparral to desert.
Coastal zone A zone whose terrestrial boundary is either the inland extent of tidal influence or the inland penetration of marine aerosols within the atmospheric boundary layer, whichever is greater. Its seaward limit is either the outer extent of the continental shelf or the limits of territorial waters, whichever is greater.
Code In nomenclature, one of the sets of internationally agreed rules governing the scientific names of organisms.
Coenosis Another term for a COMMUNITY.
Coevolution Evolved mutual adaptations between species, in which each species influences the evolution of the other.
Coevolutionary Having to do with or caused by coevolution (see previous).
Coevolutionary arms race A term for the continuing bouts of coevolving defenses and counterdefenses that occur between hosts and parasites.
Coevolve To undergo a process of coevolution (see above).
Coexistence 1. The fact of two or more species being found in the same place at the same time. 2. The ability of species to persist together in an ecological community indefinitely in the absence of any major environmental change.
Coextinction A process in which the extinction of one species leads to the loss of another species.
Cohesion The sum total of forces or systems that hold a species together; cohesion mechanisms include isolating

mechanisms in sexual species as well as stabilizing ecological selection.

Cold spot A region that is unusually lacking in a particular group of species (e.g., birds), compared with average sites in the same geographic region. Contrasted with a HOT SPOT.

Collector curve A plot of the number of species found versus the number of individuals examined. The number of species generally rises very quickly at first and then levels off toward an asymptote as fewer new species are found per unit of individuals collected.

Colonial organism A colony (see below) of single-celled organisms; postulated to be the first evolutionary step from unicellular to multicellular organisms.

Colonization The process by which a species moves into a new area where previously it was not present.

Colony A grouping of individual species of the same species living closely together; e.g., a collection of interconnected coral polyps. Often the polyps are connected with living tissue, but sometimes tissue death means that only skeletal material connects them.

Command and control A policy instrument in which regulated firms must meet individual environmental targets, often requiring the use of specific technologies. Such policies contrast with MARKET-BASED incentives, which allow firms more flexibility in meeting aggregate environmental targets.

Commensal Characterized by commensalism (see next); living in association with (e.g., on or in) another organism.

Commensalism An association of organisms in which one obtains benefit with no effect on the other. [From a term meaning "sharing a table;" i.e., sharing food.]

Commodification The act of making something a commodity (see next).

Commodity A traded item in the marketplace that is not differentiated from other similar types. An example of an agricultural commodity is Number 2 Yellow Corn, which has certain market characteristics and can be produced by many different cultivars of corn.

Common pool resource (CPR) A resource (such as an ecosystem) large enough that excluding potential beneficiaries from its use is difficult, and each individual's consumptive use (e.g., harvesting a boatload of fish or a truckload of forest products) reduces what is available to others. Also known as a **common property resource**.

Commons Resources owned in common; i.e., that are available to the public in general. Can include both public property and private property to which the public has certain rights. See also TRAGEDY OF THE COMMONS.

Community 1. An assemblage of species coexisting in a given area and interacting with each other through trophic and spatial relationships. 2. A designated subset of such an assemblage, such as the avian community or the vascular plant community.

Community structure The web of potential biological interactions among members of a community that may be characterized in terms of diversity, complexity, hierarchy, and stability.

Compact city theory The theory that more compact cities, with greater population density, are more efficient, use less energy per capita, and have less environmental impact per capita than less compact cities.

Comparative analysis Pattern-seeking across natural ecosystems, based on the premise that existing gradients and variations generate regularities that can be discerned by statistical analysis of the composite aggregate, of component variables, or of mechanistic connections.

Compensation depth The ocean depth where photosynthesis and respiration are in balance.

Competition 1. The use or defense of a resource by an individual that reduces the availability of the resource for another individual; e.g., a desirable growing site for a plant, or a food source or shelter for an animal. 2. A reduction in one species' growth rate because of the effects of another species, as by exploitation of a common resource.

Competitionism The view that competition regulates populations.

Competitive exclusion Extinction or exclusion of one species by another in a given area, through competition for resources.

Competitive release The expansion of a species-realized niche that is associated with the absence or removal of competition from other species.

Complementarity Resource partitioning among consumer species that results in greater utilization of available resources.

Complementary conservation The application of a range of conservation techniques (including *ex situ* and *in situ*) to conserve a target taxon, with one technique acting as a backup to another.

Complex adaptive system A system of multiple interconnected diverse elements whose localized interactions are subject to autonomous selection, leading to continual change and hierarchical structure.

Compositional unit A virtual subunit that contains as many types (such as species) as the actual subunits do on average, but does not share types with the other virtual subunits.

Condylarth Another term for an ARCHAIC UNGULATE.

Congener Another species that belongs to the same genus as the one being referenced.

Connectivity The linkage of habitat, land cover, or ecological processes from one location to another, or throughout an entire landscape.

Connectivity web A food web that illustrates only feeding links without reference to strength of interaction or energy flow.

Consensus cladogram or tree A branching diagram that summarizes the common branching patterns from two or more cladograms.

Conservation Active efforts by people to preserve the diversity of life on Earth and the health of natural ecosystems, usually to counteract a force that erodes or reduces natural processes. Includes the planned management of natural resources to prevent their exploitation, destruction, or neglect.

Conservation agriculture Farming practices that maintain a permanent or semi-permanent organic soil cover, such as a growing crop or a dead mulch to protect the soil physically and to feed soil biota, which improves soil quality.

Conservation biology An integrative approach to the protection and management of biological diversity that uses appropriate principles and experiences from the natural sciences, the social sciences, and various resource management fields.

Conservation concession A type of payment for ecosystem services in which a buyer purchases the right to manage a large tract of natural habitat for conservation purposes instead of commercial exploitation.

Conservation planning The process of identifying spatially explicit priorities and actions for the conservation of a region's biodiversity or natural capital.

Conservation practice or **effort** Any action taken to maintain environmental quality or to produce environmental improvements.

Conservation refugees A term for people who, as a result of the establishment of a protected area, have been displaced without fair compensation from the place where they lived, hunted, foraged, or grazed livestock.

Conservation tillage Tillage that reduces soil disruption in order to conserve organic matter and water and reduce erosion.

Conservation tourism Tourism that yields a net positive benefit for conservation of biodiversity, i.e. where contributions to conservation outweigh environmental impacts. See also ECOTOURISM.

Conserved Of a scientific name, one that the appropriate international body has decided should continue to be used for an organism in cases where a strict application of the Code would mean it had to be replaced.

Conspecific 1. Belonging to the same species. 2. An organism of the same species as another.

Constitutive defense A defense manufactured and maintained by a plant, regardless of whether it has been attacked by an herbivore or pathogen.

Consumer surplus A measure of the welfare that people gain from the consumption of goods and services, typically measured as the difference between the price that they are willing and able to pay for a good or service and the actual market price.

Contemporary (environmental) factors The current abiotic and biotic characteristics of an area, including chemical and physical aspects, as well as the composition of the biotic community.

Continental drift The movement of continents as a result of plate tectonic processes. This was especially important for Southern Hemisphere continents.

Continental rise An area at the base of the continental slope (see below) between 3000 and 4000 m, where the bottom slope is slight and sediments often accumulate.

Continental shelf A region of ocean bottom extending from the low water mark at the edge of continents to a depth (greater than 200 m) at which the incline increases markedly and the continental slope (see next) begins.

Continental slope The ocean bottom extending from the edge of the continental shelf at an average incline of greater than 4° (3000–10,000 m) at which the slope decreases and the continental rise (see above) begins.

Contour hedgerow intercropping An agricultural system in which crops are regrown in the interspaces between rows of planted woody shrub or tree species that usually are legumes, and in which the woody species are periodically pruned during the cropping season. Also called ALLEY CROPPING.

Contracted vegetation A term for vegetation restricted to wadis that receive additional water supply.

Control A process or technique of managing the population or limiting the spread of invasive species to reduce the damage they cause after becoming established in a new habitat. See also BIOLOGICAL CONTROL.

Convective mixing Vertical atmospheric mixing produced by the increasing density of a fluid in the upper layer, especially during winter in temperate and polar regions.

Convention on Biological Diversity (Biodiversity) (abbreviated **CBD**) An international convention that originated at the Earth Summit of Rio de Janeiro in 1992. Its objectives are stated as "the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding." To date 193 nations have either ratified, acceded to, approved, or accepted the Convention, with the U.S. being the most prominent exception.

Convergence The evolution of similar characters in genetically unrelated or distantly related species, usually as the result of selection in response to similar environmental pressures in different places. Also, **convergent evolution**.

Conveyance The movement of water through an area. For example, by increasing the cross-sectional area available to convey flood waters, floodplains can reduce flood stage.

Cooperation See SOCIAL COOPERATION.

Cooperative breeding A social system in which more than two individuals care for young.

Coppice The regrowth of a felled tree from the stump.

Coral reef A benthic environment characterized by reef-building corals with symbiotic dinoflagellates.

Coral Triangle A name for the region, located in the waters of southeast Asia (e.g., Indonesia and the Philippines), that contains the highest diversity of corals and other marine organisms.

Core and satellite model A network of habitat patches in which centrally located core patches harbor stable populations that provide emigrants (or colonists) to ephemeral populations in satellite patches.

Core range The central part of the home range of an animal, as opposed to its outlying areas.

Correlation The process of determining the relative age of geological strata.

Corridor A linear strip of habitat connecting two otherwise isolated patches of suitable habitat, especially in a landscape altered by human activity. Corridors can be established through the conservation of existing habitat or by ecological restoration. Also, HABITAT CORRIDOR.

Cosmopolitan Being present everywhere; globally distributed.

Countryside (habitat) A landscape in which native and human-influenced habitats are combined.

Countryside biogeography A field aimed at understanding the diversity, distribution, and conservation of species in agricultural and other human-modified landscapes.

Cover (type) A term for a type of mapped land feature (e.g., forest, freshwater, or grassland).The dominant natural cover

type in a landscape is often used to approximate habitat. See also LAND COVER.

Cowbird An American blackbird in the genus *Molothrus*, which contains important brood parasites.

CPR common-pool resource.

Craniomandibular The anatomical site where the head, or cranium, and the jaw, or mandible, meet.

Crenarchaeota One of two kingdoms of organisms of the domain Archaea, including sulfur-metabolizing, extreme thermophiles. [From Greek *crene* (spring, fountain), for the resemblance of these organisms to the ancestor of the Archaea, and *archaios* (ancient).]

Crepuscular Describing organisms that are characteristically active in dim light conditions, such as at dusk or dawn.

Critical habitat A habitat required to maintain a threatened or endangered species, often itself also threatened by destruction, disturbance, or modification.

Critical load A quantitative estimate of an exposure to a pollutant below which significant harmful effects on specified elements of the environment do not occur (according to present knowledge).

Critically endangered See ENDANGERED.

Cropland Any land used for growing crops, including both sites with long-growing crops such as fruit and nut trees and those with temporary crops or temporarily fallow.

Crop mixture (mixing) An agricultural system in which two or more crop (tree) species are grown simultaneously on the same field in alternative rows or mixture, with no distinct row arrangement.

Crop rotation The practice of growing a sequence of crops in a single field over different growing seasons.

Cropping system An integrated system of agricultural production, including crop selection, management, harvest, distribution, processing, and marketing.

Crossbreeding The mating of animals of different breeds to produce commercial market animals or breeding animals. The resulting **crossbred** animals often exhibit improved fitness and performance compared to their purebred parents.

Cross-resistance The ability of a single gene or mechanism to confer resistance to more than one toxin.

Crown group A clade including a group of modern species, the common ancestor, and all its descendants, including extinct lineages.

Cryopreservation The long-term storage of gametes, embryos, or somatic cells in liquid nitrogen.

Cryoturbation A process by which soil particles and stones are moved and mixed by the frequent formation and subsequent melting of ice crystals in the soil column.

Cryptic 1. Describing an organism that is concealed or obscured by the similarity of its appearance to the surrounding environment. 2. Belonging to a cryptic species (see next).

Cryptic species Species that are virtually indistinguishable by normal morphological analysis and are, instead, defined by a combination of genetic, behavioral, and other characters.

Cryptoendolithic Living within the surface of rocks.

Cryptogam A plant that reproduces by spores or gametes rather than seeds; this includes bryophytes and lichens.

Ctenidia (singular **ctenidium**) The molluskan gill, comprising a characteristic muscular axis from which multiple filaments arise. The relative placement of ciliary bands, nerves, muscles, and blood spaces is similar in all extant taxa.

Cuckoo One of a family of birds, the Cuculidae, of which approximately half the known species (61 of 125) are obligate interspecific brood parasites.

Cultivar A cultivated variety or genetic strain of a domesticated food plant.

Cultivated plants A collective term for plants that are grown for their produce.

Cultural diversity The variation of human groups distinguished through such features as beliefs, lifeways, dress, food, languages, sexual behavior, forms of productive organization, art, and conceptions of nature.

Cultural ecology The study of how culture influences the interactions between a human population and the ecosystems in which they reside. Also, ECOLOGICAL ANTHROPOLOGY.

Cultural ecosystem An ecosystem produced by the long-term interaction of wild plants and animals with human activities.

Cultural eutrophication The acceleration of nutrient over-enrichment caused by humans; can result from direct, point-source pollution (sewers) or by diffuse, nonpoint-source pollution (such as fertilizer runoff from farm fields).

Cultural services Ecosystem services (see) that satisfy cultural needs, such as recreational, religious, spiritual, or artistic experiences.

Culture The system of socially learned, shared, and patterned ideas, institutions, behaviors, and material products that distinguishes a particular society.

Cursorial Adapted for running.

Cuticle A noncellular external layer of the body wall of various invertebrates.

Cyanobacteria A division of microorganisms that are related to the bacteria but are capable of photosynthesis. They are prokaryotic and represent the earliest known form of life on Earth. Formerly known as BLUE–GREEN ALGAE.

Cyanophage A virus that infects cyanobacteria.

Cyanotoxin A group of toxins produced by various freshwater, marine, and terrestrial cyanobacteria.

Cycadeoidales Another name for the BENNETTITALES.

Cyperaceae The formal name of the sedge family of plants.

Cytochrome P-450 monooxygenases A ubiquitous group of enzymes involved in the NAPDH-mediated oxidation and metabolism of a broad range of endogenous and exogenous substrates.

Cytotoxicity The ability of a chemical compound to kill cells in an experimental system, such as a petri dish. Such a compound is **cytotoxic**.

Damage costs Costs associated with damages to economic activities arising from conservation programs; e.g., damage to crops and livestock from wild animals living in protected areas adjacent to human settlements or direct wildlife attacks on humans.

DDP Dryland Development Paradigm.

Decomposer An organism that acts to break down dead or decaying plants and animals; e.g., bacteria, fungi, worms.

Decomposition The breakdown of organic matter such as dead plants and algae to release carbon and nutrients.

Deep ecology An activist philosophy that advocates significant personal and political change to protect the natural environment.

Deep sea 1. Volumes of water or areas of ocean bottom at depths greater than 200 m. 2. **deep-sea**. Having to do with this area.

Definitive host For parasites with complex life cycles, the host in which the parasite reproduces sexually. Also, FINAL or PRIMARY HOST.

Deforestation The most intense form of timber extraction, in which forests are converted for another land use such as agriculture or ranching, either permanently or for a long term. Can be specifically defined as depletion of forest crown cover to less than 10 percent density.

Degradation A temporary or permanent loss of ecosystem structure, productivity, and native species diversity. Directly related to declines in biodiversity, degradation encompasses soil impoverishment and hydrological alterations that diminish water availability. See also DESERTIFICATION.

Degraded 1. Having undergone a process of degradation (see previous). 2. Specifically, describing agricultural land in which the productivity has been reduced as a result of human actions.

Delphinidae The family of odontocete cetaceans that includes the dolphins. Delphinids are the most diverse family of marine mammals, with 36 species currently recognized, more than one quarter of the known modern species of marine mammals.

Deme A localized breeding group or subpopulation, often confined to a single habitat patch.

Dendrogram A branching, nonreticulate diagram expressing nested hierarchical relationships or similarities (or both) between entities (e.g., taxa), such that the entities only appear at the tips of terminal branches.

Denitrification The reduction by bacteria of nitrate or nitrite to gaseous nitrogen products, mainly N_2 and N_2O.

Density dependence The condition that environmental factors influence population growth rate in relation to population size. Density dependence usually is seen as a linear, inverse relationship between population growth rate and population density (i.e., population growth decreases as density increases) and may occur if individuals compete or predators are more effective as a prey population increases.

Density independence The absence of environmental factors that influence population growth as a function of density. This may occur if mortality removes a fixed percentage of a population, independent of population size.

Denticles Small tooth-like projections such as those that occur on shark skin.

Depauperate Of low diversity; lacking in species richness. Contrasted with SPECIOSE.

Depletion 1. The near or total exhaustion of a resource through human overexploitation. 2. Specifically, a reduction in population of a fishery to a level low enough to reduce or threaten future productivity. **3**. See OZONE DEPLETION.

Deposit feeder An organism that feeds primarily by ingesting organic material occurring on or between sediment grains.

Deposition 1. The delivery of material inputs to the Earth's surface from the atmosphere. 2. The rate of influx of a substance in this manner, usually expressed as mass per unit area of ground.

Derived 1. Describing a trait or condition that appears in a descendant species and is changed from the ancestral condition. 2. Describing taxa with such traits or conditions.

Desertification Reduction of the biological potential of land that ultimately leads to the creation of desert-like conditions, through factors such as a decline in primary productivity, elimination of biological cover, shifts in plant diversity, or degradation of soils. Can result from natural climatic variation but usually occurs primarily through overexploitation of marginal lands by humans for agriculture; e.g., overgrazing by livestock.

Desertification myth The argument that the popular conception of African deserts expanding inexorably across fertile lands exaggerates the extent to which this process is happening, and overemphasizes the role of human activity in the process.

Desmostylia The only known fully extinct order of marine mammals, closely related to the sirenians. Desmostylians were large herbivorous quadripedal animals living in shallow waters during Oligocene and Miocene times, possibly resembling the modern hippopotamus in ecology, behavior, and appearance. They are known in the fossil record only from coastal regions of Oregon and Washington State.

Detoxification 1. A lessening of the toxic properties of some substance. 2. Specifically, a reduction in the toxic effect of a pollutant by its chemical or biological transformation to another less toxic chemical.

Detrital shunt A process in which energy and nutrients from the saprotrophic (detritus-eating) web re-enter the plant herbivore predator food web when detritivores are eaten by predators that also eat plants, herbivores, or other predators.

Detritivore An organism that consumes dead plant or animal material. Such an organism is **detritivorous**.

Detritus Dead organic matter, in either soluble or particulate form.

Diapause A state of dormancy in some animals that is induced by a so-called token environmental cue, such as day length. The token cue serves as a reliable indicator of a coming onset of harsh environmental conditions, but is not by itself harsh.

Diaspore A plant part distributed by dispersal, regardless of its developmental and morphological origins. It may be a naked seed, a seed enclosed in a fruit, or many seeds enclosed in a fruit, also a bulb or length of rhizome.

Diastema A large gap between incisors and premolars that is thought to allow an animal enough space to manipulate food with the tongue.

Diazotrophs Free-living or symbiotic bacteria or archaea that are capable of reducing N_2 to a biologically utilizable form of nitrogen (N) or ammonia (NH_3).

Dichotomous tree A cladogram in which all nodes are bifurcate (split into two parts).

Dichotomy Two taxa arising from one node.

Diffused vegetation A term for vegetation occupying all slopes and most habitats.

Digitigrade 1. Literally, walking on the digits; a posture where the majority of the weight is borne by the metacarpal and metatarsal bones and the heel/palm is raised off the

ground. 2. An organism that walks in this manner; e.g., dogs, cats, and various other mammals, or walking birds. Compare PLANTIGRADE, UNGULIGRADE.

Dilution effect The phenomenon by which increases in biodiversity are associated with decreases in the transmission, risk, or incidence of infectious disease.

Dilution host A species that reduces rates of transmission of pathogens.

Dinitrogen N_2; the most abundant form of nitrogen (N) in the atmosphere.

Dioecious Describing plants that have the male and female reproductive organs borne on separate individuals of the same species.

Diploid 1. Describing an individual with two sets of chromosomes, usually received from different parents, in contrast with a HAPLOID organism such as most microorganisms, the gametophyte stage of higher plants and some parthenogenetic forms (e.g., male ants and honeybees). 2. An organism of this type. It may have the same form or allele of a gene on both members of a chromosome pair, in which case it is *homozygous*, or in different forms (alleles), in which case it is *heterozygous*.

Diptera The group of insects to which the flies belong, including gnats and mosquitoes.

Dipterocarp A member of the Dipterocarpaceae family (comprising 22 genera) of South Asian and African timber trees.

Direct land-use change (DLUC) Conversion of land from a non-crop use to the production of bioenergy crops.

Direct use values Benefits gained from environmental goods and services that involve some form of direct physical interaction, such as recreational activities, resource harvesting, drinking clean water, breathing unpolluted air, and so on. Compare INDIRECT USE VALUES.

Disease Infection by parasitic organisms that can cause weakness, decreased reproduction, and death.

Disease dynamics Changes in incidence or prevalence of disease through time.

Disease incidence The number of new cases of a disease per population per unit time, often expressed as a rate.

Disease prevalence The number of disease cases per population.

Disharmonic Describing a fauna in which a strongly reduced number of major lineages is represented, as in oceanic islands, due to their improbable colonization from distant sources.

Disharmony The non-random representation of biota among natural colonists of oceanic islands as compared with mainland systems. Species that are more dispersive will have a higher representation in the biota, the effect being accentuated by the tendency of these few successful colonists to diversify in to multiple species through adaptive radiation.

Disparity The range of diverse morphological architectures present in higher taxa such as classes, phyla, and kingdoms; used to explain the origins and maintenance of this diversity of life-forms and body plans.

Dispersal 1. Any movement or spread of an organism from one area to another independent of other organisms and of historical events that change the natural distribution of the organism. 2. The movement of an individual organism away from its place of origin to the place where it breeds, or the movement by an adult from one breeding location to another.

Dispersal limitation The degree to which the location of an organism is constrained by the location of its parent; the inability to move to and colonize another location.

Disruptive selection Selection acting to preserve extreme phenotypes in a population. Speciation usually involves disruptive selection, because intermediates (hybrids between incipient species) are disfavored. Compare STABILIZING SELECTION.

Disruptor See ENDOCRINE DISRUPTOR.

Dissipative beach A wide, flat beach.

Distance–decay A decrease in the similarity of community composition with increasing spatial distance.

Distinct (population) segment A subspecific unit that can be afforded legal protection under the U.S. Endangered Species Act.

Disturbance Any relatively discrete event that disrupts ecosystem, community, or population structure and changes resources, substrate availability, or the physical environment, particularly one that results in increased mortality of plants or animals. Disturbances may be natural events such as wildfires, hurricanes, tornados, or landslides, or human effects such as logging or land clearing.

Disturbance regime The collective spatial, temporal, physical, and ecological characteristics of a disturbance process operating in an area.

Divergence 1. A fixed genetic difference between two species or populations. 2. The process of evolving such a difference.

Divergence-with-gene-flow A model explaining the process of species formation (speciation) in which new species diverge in the face of gene flow; i.e., the movement of genes within a group that results from mating with immigrant individuals.

Diversity A quantitative measure of the taxonomic or physiognomic composition of coexisting organisms at one of three scales: *alpha* (local diversity), *beta* (diversity change between adjacent local sites), and *gamma* (regional diversity). See also BIODIVERSITY.

Diversity index A numerical measure combining information about the diversity of species present in a sample or habitat and information about their relative abundances.

Diversity principle The principle that there is a general geographical coincidence between high concentrations of biological diversity and cultural diversity.

Diversity-productivity hypothesis The proposal that greater diversity would lead, on average, to greater total biomass or productivity.

Diversity-stability hypothesis The proposal that ecosystems containing more species would be more stable.

DLUC Direct land-use change.

DNA Deoxyribonucleic acid, the crucial molecule of inheritance that stores genetic information passed from one cell to another and from one generation to succeeding generations, composed of four nucleotides: adenine (A), cytosine (C), guanine (G), and thymine (T).

DNA barcoding A method of delimiting species by means of the clustering of short stretches of DNA sequence data called "barcodes," usually from mitochondrial DNA.

DNA fingerprinting A technique used to amplify small amounts of DNA to identify individuals, often for forensic purposes.

DNA probe array A collection of short, single-stranded, uniquely addressable segments of DNA attached to a solid surface; used to detect the presence or expression of specific genes via the principles of nucleic acid hybridization. A **protein probe array** is similar but with proteins serving as sensing elements.
DNA slippage A polymerase error leading to a change in the length of a simple repetitive sequence.
Dodder Viney parasitic plants of the genus *Cuscuta* that form haustoria (specialized nutrient absorbing structures) on the stems of their hosts. The unrelated parasitic vine, *Cassytha*, is often referred to as Laurel Dodder.
Domain The highest level of taxonomic rank used to classify organisms; the three domains are Archaea, Bacteria, and Eukarya, which differ from each other in fundamental genomic and phenetic properties.
Domesticate 1. To purposely adapt an animal or plant from a wild to a controlled state for some human use. The process of doing this is **domestication**. 2. An animal or plant that has been adapted in this way.
Dominant (species) 1. Species that have the greatest influence on ecosystem structure and function, by virtue of their high abundance, biomass, and/or coverage. Examples include trees in forests, grasses in grasslands, mussels in rocky intertidal habitats, and kelp or coral in near-shore subtidal habitats. 2. Having to do with or being such a species.
Donor control A condition in a food web in which consumer population growth is affected by their resources but consumers do not affect the renewal rate of these resources and hence cannot depress them.
Dormancy Any state of reduced metabolic activity of an organism. Typically, dormant organisms have associated characteristics such as cessation of development, absence of reproduction, and enhanced resistance to harsh environmental conditions.
Dormant In a state of dormancy (see previous).
Dorsal The part of an organism that is the "back" or upper part of the body. Contrasted with VENTRAL which is the "front" or lower part.
Drift The influence of stochastic demographic variability (such as birth, death, and migration rates) on biotic composition in an area.
Drought-deciduous Describing plant species that lose their leaves during the dry season as soil moisture becomes limited.
Dryland Development Paradigm (DDP) A conceptual framework to aid in understanding the complexity of desertification and dryland development, by identifying and synthesizing the factors important to the research, management, and policy communities.
Drylands A collective term for arid and semiarid croplands, pastures, rangelands, and subhumid woodlands where the index of aridity is less than 0.65; these cover more than two-fifths of the land surface of the globe, and are home to more than one-third of the human population.
Dyke intrusion Magma that upwells from the mantle; may reach the surface of the Earth's crust to flow out as lava and solidify as basalt.
Dynamical model A mathematical description of a system with components that vary in time.
Dynamical system A set of rules defining how certain variables change with time. Ecological models are dynamical systems representing what are believed to be vastly more complex dynamical systems in nature.
Dynamical variable A quantity in a dynamical system that changes with time according to the rules of the system.
EBM Ecosystem-based management.
Ecoagriculture A land-use approach that achieves agricultural production and sustains rural livelihoods, in ways that also protect wild plant and animal species and the natural ecosystem services upon which both humans and wildlife depend.
Ecocriticism The interdisciplinary study of literature, history, religion, and philosophy with an emphasis on habitats, evolutionary biology, and environmental problems.
Ecogeography The analysis of a species' ecological, geographical, and taxonomic characteristics to assist in the formulation of collection and conservation priorities.
Ecohumanism See HUMANISM.
Ecological 1. Having to do with ecology, the study of the relationship of organisms to each other and their environment. 2. Having to do with the natural environment, especially its protection and sustainable use.
Ecological anthropology See CULTURAL ECOLOGY.
Ecological biodiversity See BIODIVERSITY.
Ecological biogeography See BIOGEOGRAPHY.
Ecological community See COMMUNITY.
Ecological footprint The "mark" that people leave on the natural environment; a quantitative calculation to estimate the cumulative environmental impact of a human individual or population, especially in the context of daily activities. See also CARBON FOOTPRINT.
Ecological function See ECOSYSTEM FUNCTION.
Ecological integrity The balanced state of an ecosystem under normal environmental conditions, including its capacity to absorb disruption and also to recover from it.
Ecological niche See NICHE.
Ecological process 1. Any of the direct or indirect interactions between species, such as grazing, predation, competition, responses to disturbance, and so on, that cause spatial or temporal patterns in the distribution and abundance of the species. 2. See ECOSYSTEM PROCESS.
Ecological refugee A term for people who have lost access to their traditional base of natural resources yet have very limited access to resources through markets.
Ecological release The expansion of habitat, or ecological environment, often resulting from release of species from competition.
Ecological restoration See RESTORATION.
Ecological sorting The process by which organisms colonize and persist in novel environments, use novel resources, or form novel associations with other species, as a result of the suite of traits that they carry at the time they encounter the new condition. This provides an important mechanism by which new species can fit in to an existing community without adaptation.
Ecological stoichiometry The quantitative relationship between chemical elements (especially carbon, nitrogen, and phosphorus) arising from ecological processes and interactions such as nutrient acquisition, biomass formation,

remineralization, trophic transfer, sedimentation, and dilution.

Ecological transfer efficiency The ratio of energy ingested by a population's predators to the energy ingested by the population itself.

Ecological transition The concept that human societies tend to be in growing disequilibrium with their biophysical environment as they increasingly deplete natural resources and degrade their habitat, thus reaching new levels of environmental impact.

Ecologist A scientist who studies the relationship of organisms to each other and their environment.

Ecology The scientific study of the relationships of organisms with each other and with their environment, at levels ranging from the interaction of two individual organisms to the entire biosphere; it includes issues such as the nature, distribution, abundance, and diversity of species, the flow of energy and materials, and the effects of human activity on natural systems.

Ecomorph A local population or group associated with a particular ecological niche.

Economic botany The scientific effort to identify and characterize economically useful plants and the products derived from them.

Economic efficiency Another term for PARETO EFFICIENCY; i.e., an equilibrium condition in which all goods in an economy are allocated in such a way that no one can be made better off without making someone else worse off.

Economic rent The flow of net economic benefits (revenue minus costs) derived from the exploitation of a resource asset.

Economic value (valuation) In this context, the various methods used by economists to determine the willingness of individuals to pay for the benefits they derived from ecosystems, many of which are not marketed.

Ecoregion A relatively large unit of land or water containing a distinct assemblage of natural communities that share a large majority of their species and ecological dynamics, and also share similar environmental conditions, and interact ecologically in ways that are critical for their long-term persistence.

Ecospecies A species considered in the context of its ecological characteristics and usually including several ecotypes (see) that are capable of interbreeding.

Ecosystem A community of organisms (plants, animals, and microbes) and their physical environment (air, soil, water, and so on) interacting as an ecological unit in some defined geographical space, such as a lake, a wetland, a forest, or an agricultural landscape.

Ecosystem approach A strategy for the integrated management of land, water, and living resources that promotes conservation and sustainable use in an equitable manner.

Ecosystem-based adaptation Conservation and management of natural ecosystems so that they can provide services to help society adapt to perturbations from climate change.

Ecosystem-based management (EBM) An environmental management approach that recognizes the full array of interactions within an ecosystem, including humans, rather than considering single issues, species, or ecosystem services in isolation.

Ecosystem disservices The converse of ecosystem services (see); functions or processes provided by ecosystems that have a negative impact on people; e.g., disease agents, agricultural pests.

Ecosystem diversity See BIODIVERSITY.

Ecosystem engineer A species that directly or indirectly modulates the availability of resources (other than itself to other species) by causing physical state changes in biotic or abiotic material and in so doing modifying, maintaining, or creating habitats. Examples are animals that affect an ecosystem by physically changing it (such as beavers building dams) or by being present in such great numbers that they have a disproportionate impact (such as large grazing animals). Also applied to organisms that move parts of the soil volume for their own use (such as earthworms making macropores that permit the flow of large amounts of water through the soil). Also, **ecosystem engineer species**.

Ecosystem function(ing) Any of the physical, chemical, and biological processes or attributes that contribute to the self-maintenance of an ecosystem (e.g., nutrient cycling, primary production, food web systems).

Ecosystem goods Commodities, such as timber and seafood, supplied free to humanity by natural ecosystems.

Ecosystem health The condition of an ecosystem, whether natural, managed, or human-dominated, that is free from influences that would damage or destroy its characteristic structures and functions.

Ecosystem integrity See ECOLOGICAL INTEGRITY.

Ecosystem management The management of an ecologically defined geographical area as an integrated whole for both sustainable yields of natural resources and maintenance of ecological processes and species populations.

Ecosystem people A description for people who directly meet the bulk of their resource requirements from a limited area near their habitation, as through gathering or low input agriculture and animal husbandry.

Ecosystem process Another term for ECOSYSTEM FUNCTION.

Ecosystem resilience See RESILIENCE.

Ecosystem services A broad term for the many and varied benefits that people obtain from natural ecosystems. These include *provisioning* services such as food and water; *regulating* services such as flood and disease control; *cultural* services such as aesthetic and recreational benefits and spiritual values; and *supporting* services, such as nutrient cycling, which maintain the conditions for life on Earth.

Ecosystem structure The sum of the biotic (living) and abiotic (nonliving) components of an ecosystem, and the interactions among them.

Ecotonal Having to do with or occurring in an ecotone (see next).

Ecotone A zone or transition between adjacent ecological systems, having a set of characteristics uniquely defined by space and time scales, and by the strength of the interactions between adjacent ecological systems.

Ecotone effect The pattern of increased species richness (number of species) and abundance in ecotones and the occurrence of unique ecotonal species.

Ecotourism Commercial tourism that focuses on relatively undisturbed natural environments, and that takes steps to minimize environmental impacts, provide environmental education to clients, and contribute to environmental conservation.

Ecotoxicology The use of test organisms (e.g., the water flea, *Daphnia*) to study the toxicity, pathways of accumulation, and breakdown of chemicals, particularly those manufactured by humans (e.g., pesticides).

Ecotype A genetically different type within a species that reflects an obvious advantage in a given environment. Ecotypic traits are retained when individuals are transplanted into a different environment where these traits have no advantage.

Ecotypic Having to do with or being an ecotype (see previous).

Ecozone A major division of the Earth's land surface, based on the general groupings of the organisms found there.

Ectoderm An embryonic tissue that provides the future outside layer of the animal.

Ectoparasite A parasite living on the surface of a host.

Ectotherm An animal whose body temperature has minimal internal regulation and instead varies with external conditions; e.g., snakes.

Ectothermic Describing animals whose body temperature is regulated mainly by the exchange of heat with the surrounding environment; in popular use, "cold-blooded."

Ectothermy A method of body temperature control in which the animal utilizes the Sun and other external sources for gaining and giving up heat, thus achieving temperature control without affecting metabolic rate. Contrasted with ENDOTHERMY.

Edaphic Related to or caused by particular soil conditions, as of texture or drainage, rather than by physiographic or climatic factors.

Edge effect The effect of the juxtaposition of contrasting environments on an ecosystem.

Edge-of-field loss A term for the nitrogen and phosphorus that is lost or exported from fields in agricultural production.

Effective population size (Ne) The size of an ideal population that would have the same rate of increase in inbreeding or decrease in genetic diversity as the actual population under study.

Effective species A virtual species that has the same proportional abundance as the actual species do on average. The number of effective species, or effective number of species, is species diversity.

Efficiency 1. The degree of success in reaching a conservation goal (e.g., maximizing the number of species in a given area) relative to the cost (or to some surrogate for cost, such as number of areas). 2. A condition of economic equilibrium. See PARETO EFFICIENCY; KALDOR-HICKS EFFICIENCY.

Efficiency frontier A condition where it is not possible to improve on any particular objective without decreasing some other objective; can be used to determine the maximum feasible outcome that can be achieved for a landscape.

Effluent line The level on a beach face where the water table intersects the surface.

Egg bank An accumulation of long-lived diapausing eggs (i.e., those eggs that persist in diapauses for longer than a single growing season) of aquatic invertebrates in the sediments of marine or freshwater habitats.

Egg rejection Host responses to parasitism that include the ejection of parasitic eggs, abandonment of parasitized nests (usually with renesting), or a new cup built over a parasitized clutch. See also MAFIA EFFECT.

EIA Environmental impact assessment.

EID Emerging infectious disease.

El Niño Southern Oscillation (ENSO) A periodic climate pattern that takes place across the tropical Pacific Ocean at intervals of approximately five years, characterized by warmer than normal ocean surface temperature and greater air surface pressure in the eastern and western tropical Pacific Oceans, respectively. The other phase of the cycle, the **La Niña Southern Oscillation**, is characterized by colder than normal surface waters and lower air surface pressure. ENSO events are associated with changes in rainfall patterns in the tropics and with an influence on the weather at higher latitudes, and also with changes in fisheries and ecosystem processes. [From Spanish *El Niño*, the male child (i.e., the Christ child); based on a traditional name for warmer ocean currents occurring around Christmastime.]

Electrophoresis The separation of charged molecules (such as polypeptides or polynucleotides) between the two poles of an electric field.

Elemental stoichiometry See ECOLOGICAL STOICHIOMETRY.

Elfin woodland Small, stunted tree growth characteristic of forests at higher elevations in warm, moist regions.

Elite capture The argument that benefits from a conservation or a development initiative may be "captured" by local elites who have the economic or political power to determine who will benefit from such an initiative.

Eltonian niche The role, or occupation, of a species in a community, likened to the way a person's profession establishes their role in the community. [Conceived by the British ecologist Charles Elton.]

Eluviation The process in which dissolved or suspended materials in soil move from a higher to a lower level.

Emergent Describing vegetation rising above a certain level or surface; e.g., a very tall tree that rises above the general level of the forest canopy, or a plant that grows in water but rises partially above the surface; e.g., cattails.

Emerging infectious disease (EID) A disease of infectious origin whose incidence in humans has recently increased and may continue to increase in the near future; it may be a newly recognized disease or a **re-emerging** disease that was previously recognized. EIDs are associated with various causal factors such as changes in land use or human-wildlife interactions, or climate change (e.g., spread of West Nile virus).

Enchytraeid One of the Enchytraeidae, segmented worms that are related to but smaller than earthworms.

Endangered Facing a very high risk of extinction in the wild, based on specific criteria. In some systems that classify extinction risk, *endangered* is the highest category of risk, followed by *threatened*. In others it is second highest after *critically endangered*.

Endangered species A species that is at substantial risk of extinction throughout all or a significant portion of its range.

Endemic 1. Describing a species or other taxonomic group that is restricted to or occurs only in a specified geographic location or region. The actual size of the area is arbitrary (an **endemic species** might be found only in North America or only on a small island). 2. A member of such a species or group.

Endemic language One of the languages that are restricted to a single country and, like their species counterparts, hold a high percentage of the unique traits in human language as a whole.

Endemism The fact of a species or other taxonomic group being found only in a specific location of limited size, rather than being widely distributed. Also, **endemicity**.

Endocrine disruptor A chemical that interferes with the endocrine system, the means of chemical communication (mainly performed by hormones) within an organism.

Endogenous 1. Inside; occuring within a system. 2. Specifically, occurring within an organism or a population of organisms. Contrasted with EXOGENOUS.

Endogenous forces Forces within a population that affect the dynamics of the population.

Endogenous process A process that emerges from biotic and abiotic interactions internal to the system of interest. Examples include vegetation growth and succession, dispersal of organisms within a landscape, and topographic influences on microclimate. Some human activities may be considered endogenous (e.g., crop rotation systems) if they are well established and responsive to local needs.

Endogenous risk The concept that private citizens and public servants have some control over the set of probabilities and outcomes that define the relevant states of the world.

Endoparasite A parasite living inside a host.

Endophyte 1. A fungus or other parasitic organism residing or growing within plant tissue. Such an organism is **endophytic**. 2. The portion of such an organism that is embedded inside the host tissue.

Endosperm Nutritive storage tissue in the seeds of most angiosperms.

Endosymbiont An organism living in a long-term association inside a host organism for their mutual benefit; e.g., methanogenic bacteria that live inside an aerobic protozoa and utilize waste H_2 produced by the host). Such an association is **endosymbiotic**.

Endotherm An animal whose body temperature is mainly regulated by its own metabolism and varies only slightly with external conditions; e.g., mammals.

Endothermic Describing animals who regulate their body temperature mainly through internal metabolic means; in popular use, "warm-blooded."

Endothermy A method of body temperature control in which the animal modifies its metabolic rate to achieve the desired temperature, largely independent of the ambient temperature. Contrasted with ECTOTHERMY.

Enemy release hypothesis The idea that alien species have a better chance of establishing and becoming dominant when released from the negative effects of natural enemies that, in their native range, lead to high mortality rates and reduced productivity.

Energetic web A food web depiction that quantifies the amount of energy (or material) that flows across the links joining species.

Energy 1. A fundamental physical concept, defined classically as the capacity to perform work. **Potential energy** is this capacity stored as a consequence of its position, structure, shape, or state. **Kinetic energy** is this capacity as manifested in the motion of matter. The joule (J) is the common SI unit of energy, where 1 J equals the amount of energy required to increase the temperature of 1 g of water by 1° K. 2. The use of this capacity by humans to perform useful functions, such as heating or cooling, lighting, transportation, powering devices and tools, and so on. See also RENEWABLE ENERGY; NON-RENEWABLE ENERGY; INDUSTRIAL ENERGY; TRADITIONAL ENERGY.

Energy balance 1. The difference between incoming solar energy and outgoing terrestrial radiation for the Earth as a whole. 2. The relative energy flows into and out of a living system, such as an ecosystem or an individual organism. Also, **energy budget**.

Energy efficiency A measure of the performance of an energy system. **First Law efficiency** (based on the first law of thermodynamics) equals the ratio of desired energy output to the energy input. **Second Law efficiency** (based on the second law) equals the ratio of the heat or work usefully transferred by a system to the maximum possible heat or work usefully transferable by any system using the same energy input.

Energy sprawl An increase in the agricultural land area required for the production of dedicated biofuel crops in place of food crops, natural and semi-natural habitats, and other land uses, occurring directly from land-use change or indirectly due to biofuel-related market conditions.

Engineer See ECOSYSTEM ENGINEER.

ENSO El Niño Southern Oscillation.

Entropy A measure of the degree of disorder, or randomness, in a given system at a given time; greater entropy means less energy available to do work. The entropy of an isolated system cannot decrease when the system undergoes a change.

Environment 1. The entirety of the biological world and its physical surroundings. 2. The living and non-living surroundings of a particular organism.

Environmental 1. Relating to the natural environment and the impact of human activity on it. 2. Describing efforts to protect the viability and diversity of the natural world.

Environmental ethics The philosophical discipline that describes and specifies proper human relationships with the natural world.

Environmental filter(ing) Constraints on the presence and abundance of species in local communities or regional species pools, due to abiotic conditions.

Environmental impact assessment (EIA) A formalized process, including public consultation, in which all relevant environmental consequences of a project are identified and assessed before authorization is given to proceed.

Environmental Kuznets curve See KUZNETS CURVE.

Environmental stress An action, agent, or condition that impairs the structure or function of a biological system.

Environmental sustainability 1. The fact of maintaining ecosystem conditions that will sustain the biosphere now and into the future. 2. The complex of environmental, economic, social, and cultural systems necessary to accomplish this.

Environmental variables The abiotic and biotic elements of an area including vegetation, soil, topography, species competition, geology, prevalence of diseases and hydrologic characteristics, among others. Habitat selection can be positively or negatively related to these environmental variables.
Enzyme A protein involved in catalytic reactions.
Eolian Another spelling of AEOLIAN.
Epigenetic Having to do with molecules that are the product of gene expression.
Epikarst 1. The interface zone between soil and rock, characterized by fractures and small solution pockets. 2. An ecotone between surface and subterranean water in karst.
Epiphyl A plant that grows on the leaf surface of another plant.
Epiphyte A non-parasitic plant that uses another plant, usually a tree, as mechanical support but does not derive nutrients or water from it. Such a plant is **epiphytic**.
Epiphytic material Live and dead canopy vascular and nonvascular plants, associated detritus, microbes, invertebrates, fungi, and crown humus.
Epistasis An interaction between different alleles at two or more loci, such that the phenotype differs from what would be expected if each locus acted independently.
Equalizing mechanisms Biotic or abiotic forces giving rise to equal net per-capita population growth rates of individuals of different species in a community.
Equity See SOCIAL EQUITY.
Eremic Relating to or occurring in desert regions.
ESS Evolutionarily stable strategy.
Estuary A semi-enclosed coastal body of water that has a free connection with the open sea and within which the water is brackish; i.e., it has a moderate level of salinity due to the seawater being diluted by fresh water from land drainage.
Ethical holism The principle that complex aggregates have intrinsic value, such as species, ecosystems, or human groups.
Ethics Codes that exert a palpable influence on human behavior. Embedded in world views, ethics provide models to emulate, goals to strive for, and norms by which to evaluate actual behavior.
Ethnobiology 1. The study of the reciprocal interactions between people and the biological organisms in their local environment. 2. The study of the ways in which different cultures interact with and make use of the organisms in their environment, both in the present and the past. Ethnobiology can be divided into **ethnobotany** with respect to human interaction with plants and **ethnozoology** with respect to animals.
Ethnobotany See ETHNOBIOLOGY (previous).
Ethnoecology An interdisciplinary study that explores how nature is perceived by human groups and how humans use natural resources. May focus specifically on the traditional environmental knowledge and practices of local groups, as exemplified by their resource management.
Ethnoscience The study of local peoples' knowledge of biological organisms and the physical environment, especially their systems of classification and naming. See also ETHNOBIOLOGY.
Ethnozoology See ETHNOBIOLOGY.
Ethology The study of the behavior and social relations of animals.
Eucalyptus A diverse genus of trees and shrubs including species that grow quickly to great height; the dominant tree genus of Australia today and also widely cultivated in tropical and subtropical areas elsewhere.
Eukarya or Eukaryota One of the three domains of life, along with the Archaea and Bacteria, consisting of the organisms that have a membrane-enclosed nucleus and other organelles. Also spelled **Eucarya**.
Eukaryote A cellular organism in which the genome of the cell is stored in chromosomes enclosed in a membrane-bound nucleus; all protists (algae and protozoa), fungi, plants, and animals are eukaryotes. Also spelled **eucaryote**. Compare PROKARYOTE.
Eukaryotic Relating to or being a eukaryote (see previous).
Euphotic Describing the top layer of a water body through which sufficient light penetrates to support net photosynthesis.
Euphotic zone The uppermost sunlit layer of the ocean, lake, or another body of water, in which there is sufficient light energy to support net photosynthetic growth. Can vary from up to about 200 m in the open sea to as little as 1 m in turbid lake waters. Also, PHOTIC ZONE.
Euryarchaeota One of two kingdoms within the domain Archaea, including halophiles, methanogens, and some anaerobic, sulfur-metabolizing, extreme thermophiles.
Eurythermal Able to tolerate a wide range of temperatures. An organism with this ability is a **eurytherm**.
Eusociality Cooperative behavior among individuals of the same species, characterized by reproductive division of labor, overlap of generations, and cooperative nesting. Such behavior is **eusocial**.
Eutherian One of the Eutheria, the mammals in which the embryo is attached to the mother by a placenta.
Eutrophic 1. Describing an ecosystem that is high in production, or a species typical of high-production environments. 2. Specifically, describing an aquatic ecosystem characterized by high phytoplankton production and high nutrient concentrations and loadings. Compare MESOTROPHIC; OLIGOTROPHIC.
Eutrophication 1. The fact of an ecosystem being or becoming high in production. 2. Specifically, the process by which a lake or other aquatic environment becomes enriched with nutrients such as nitrogen; this stimulates plant growth and often leads to losses of plant biodiversity with dominance by algal species, and the depletion of oxygen from the water. Also, **eutrophy**.
Evaporative aerodynamics The study of the influence of aerodynamics on the development of both leaf structure and foliage distribution throughout a plant community.
Evapotranspiration The process of transferring moisture from the earth to the atmosphere by the evaporation of water and transpiration from plants.
Evenness An equitable condition in the abundance of the species present in a local community or region.
Evolution Descent with modification; morphological or genetic change in species over time. Small changes that do not lead to reproductive isolation among members of a group are referred to as *microevolution*. Speciation, or the generation of new species, is referred to as *macroevolution*.

Evolutionarily stable strategy (ESS) An optimal strategy that, if adopted by all the members of a population, can resist invasion by any new (mutant) strategy; i.e., it is stable with respect to evolutionary change.
Evolutionary equilibrium The hypothesis that frequencies of acceptance of brood parasitism reflect an equilibrium between costs and benefits of host defenses against parasitism.
Evolutionary lag The hypothesis that hosts lack defenses against brood parasites because the defenses have not yet had time to evolve.
Evolutionary legacy Genetic resources that are the product of past evolutionary events and also represent the raw material for future evolution.
Exaptation A process in which a trait first evolves to serve one particular function, but subsequently comes to serve another.
Existence value The benefit derived simply from knowing that something exists, without making use of it; e.g., a certain natural ecosystem.
Exoenzymes Digestive enzymes excreted by fungi into the environment to digest materials externally.
Exogamy Reproduction in which the individuals are not closely related to each other. Such a reproductive pattern is **exogamous**.
Exogenous 1. Outside; occuring outside a system. 2. Specifically, occurring outside an organism or a population of organisms. Contrasted with ENDOGENOUS.
Exogenous forces Those forces outside a population that affect the dynamics of the population.
Exogenous process A process external to the system of interest. For landscapes, exogenous processes would include climate change or species migration into the landscape. Human activities may be considered exogenous if they are novel, permanent (e.g. a housing development), or driven by external forces (e.g., the world market for timber).
Exons Parts of a gene sequence transcribed and translated into a protein; usually interrupted by *introns* which are spliced out of messenger RNA before translation.
Exotic Occurring in an area outside of its historically known range; not native or indigenous.
Exotic animal 1. A wild animal; i.e., an animal that has not been domesticated for use by humans in agriculture or as a pet. 2. A member of an exotic species (see next).
Exotic species A species occurring outside its natural range or potential natural dispersal range, especially one whose presence there is due directly or indirectly to human activity.
Exploitation Any use of a natural resource or other ecosystem service by humans. See also OVEREXPLOITATION.
Exploitation (exploitative) competition Interaction among two or more species that use a common resource that is limited, in which one species benefits more than the other.
Ex situ Literally, "outside its position;" not in its natural setting (*in situ*), but rather in a laboratory, zoo, or other such artificial setting.
Ex situ conservation 1. "Off-site" conservation; conservation efforts in which the species or population is removed from its natural habitat and managed in a protected environment such as a zoo, botanic garden, or nursery, or stored as dormant seed or gametes for future propagation and use. 2. A captive breeding program that take place in an artificial setting but one within the natural range of the animal. Compare IN SITU CONSERVATION.
Extant Living; having some living representatives at a particular time; i.e., not extinct.
Extent In landscape ecology, the overall size of a study area.
Externality An external cost or benefit imposed on others as the result of some economic activity.
Extinct Not living; describing a species or other group that has no living members.
Extinct in the wild A status for a species that has no living members in its natural setting but has some surviving members in a zoo or other such artificial environment.
Extinction The fact of being extinct, or the process of becoming extinct. Can be *global* (worldwide) or *local* (in a specific area). In the fossil record, it is the last fossil occurrence of individuals of a given species.
Extinction cascade A chain of extinctions triggered by the extinction of a particular species on which many others depend.
Extinction rate The number or proportion of species or other taxa becoming extinct per unit time or after an important geological temporal boundary.
Extinction risk The likelihood that a species or other taxon will disappear or die out within a given area (e.g., in one country or the entire world) and over a definable time period.
Extinction spasmorpulse A catastrophic burst of extinctions, peaking in less than a millennium.
Extinction vortex A complex of interacting mechanisms that lead to the extinction of populations; these mechanisms include loss of genetic diversity as well as environmental and demographic stochasticity.
Extirpation Local extinction; the fact of a group of organisms ceasing to exist in a particular location but still existing elsewhere.
Extremophilic Describing an organism that grows better at, or requires for growth, extreme environmental conditions such as very high (or low) temperature, pressure, and so on.
Extremophile An "extreme lover;" an organism that lives and thrives in a habitat characterized by extremes of temperature, acidity/alkalinity, salinity, light, pressure, or other environmental factors that would prove detrimental or lethal to most other organisms.
Eynbos Evergreen sclerophyll vegetation dominating the Cape Region of South Africa.
Facilitation A process in which one species benefits another, as by mediating stresses.
Facultative Having options; able to function in more than one way; e.g., a parasite that can also live without a host; bacteria that can live in both aerobic and anaerobic environments. Compare OBLIGATE.
Facultative carnivore A carnivorous animal that eats mainly meat but that also consumes non-animal food.
Facultative epiphyte A plant or lichen that commonly grows epiphytically and terrestrially, usually exhibiting preference for one or the other habit in a particular habitat.
Fallow 1. Describing a phase in agriculture when no crop is grown in a cultivated field and volunteer vegetation can grow or regrow; typically done to restore soil fertility or to avoid overproduction. 2. **Lie fallow**. Of a field, to be in this state.

Family In the classification of evolutionary lineages, a category of related organisms that ranks above a genus and below an order. A family usually contains many genera.
Farming intensity or intensification See INTENSIVE AGRICULTURE.
Fat body Storage tissue of arthropods that stores lipids and has endocrine and immunological functions.
Fauna The animal species living in a given area, along with the flora (plant species).
Feedback A change in a system component that triggers effects which will eventually change the original component again. Feedbacks can be positive (*self-reinforcing*) or negative (*self-dampening*).
Feedback loop An ecological response to a human action having environmental consequences, followed by a human behavioral response to ecological changes having economic consequences.
Feed conversion The efficiency of farmed animals to incorporate given feed into biomass; usually expressed as the ratio of weight of feed provided to flesh biomass harvested.
Feedstock The plants or plant-derived products (e.g., pellets) that can be used to produce biofuels. *First-generation* feedstocks are currently under wide cultivation for the production of ethanol and biodiesel, such as corn, soy, sugarcane, and palm oil. *Second-generation* feedstocks (e.g., switchgrass) are currently not widely cultivated for biofuels but have high potential yields.
Fellfield Dry, windswept habitat of cold regions comprising mineral soil, gravels, stones, and rock, dominated by cryptogams (especially mosses and lichens) and, outside the Antarctic continent, by compact cushion-forming phanerogams and short grasses.
Felsic rock Continental crustal rocks relatively rich in silicon and aluminum.
Femtoplankton A classification of plankton according to size; those plankton of the smallest size (0.02–0.2 mm).
Field-deployable Describing a sensor or other such instrument capable of operating outside of a conventional laboratory setting, exposed to the natural elements such as being submerged below the sea surface.
Final host See DEFINITIVE HOST.
Fingerprint A pattern produced by DNA fragments or lipids that represents a community or species.
Fire ecology The study of the occurrence of natural fires in a given ecosystem and the effect of this on the system; e.g., the fact of certain plants germinating via fire-activated seeds.
Fire regime The characteristic pattern of wildfire and bushfire in a given ecosystem, including type, frequency, intensity, severity, season, and extent.
Fire suppression An intentional effort to prevent wildfires in a forest ecosystem; can significantly disturb the natural functioning of the system.
First contact 1. The initial human arrival on a landmass, followed by human colonization. 2. The initial interaction of two previously unconnected cultures; e.g., aboriginals and European settlers.
First-contact extinction A wave of extinction of species native to a continent or island, following the first arrival of humans in that area.
First (fundamental) theorem of welfare economics The mathematical proof that if competitive markets exist for all goods and there are no externalities or other market imperfections, then the market system leads to an outcome of PARETO EFFICIENCY (see).
Fishery 1. A collective effort to gather, collect, or catch wild aquatic wildlife or plants for recreational or commercial purposes, especially fish and shellfish for use as food. 2. A site where such an effort takes place.
Fishing effort The level of fishing activity quantified in terms of the number and power of vessels and duration of fishing.
Fissipedia The suborder of the Carnivora with divided toes, as distinguished from those with finned feet (Pinnipedia).
Fitness The relative contribution of offspring to the next generation. An individual, genotype, or phenotype whose progeny constitutes a large proportion of the succeeding generation has high fitness.
Fixation The process by which certain plants and microorganisms produce organic molecules from inorganic gases. See NITROGEN FIXATION.
Fixation index (F_{ST}) An estimate of the extent of genetic variation among subpopulations. It ranges from 0 to 1, with higher measures indicating more genetic subdivision and lower values indicating less subdivision.
Flagship (species) A charismatic or well-known species that is associated with a given habitat or ecosystem and that may serve as a symbol and rallying point to increase public awareness of the need for conservation action or climate change remediation there; e.g., polar bears in the Arctic.
Flexibility The degree to which alternatives exist for the purpose of reaching a particular conservation goal.
Floodplain Broad, flat areas adjacent to rivers and streams that periodically become inundated during recurring floods. Also spelled **flood plain**.
Flood-recession agriculture A pattern of agriculture in which crops are planted on floodplains as flood waters recede. The growing plants benefit from several factors influenced by the flood, including moist soil, elevated water table, and freshly deposited soil and nutrients.
Flood resistance Submersion ability over a time span of weeks or months.
Flood tolerance Submersion ability over a time span of a few hours up to several days.
Floristic 1. Having to do with flowers or a flora. 2. Describing a plant community that contains a homogeneous suite of species sorted by some classificatory program.
Floristics The study of the type, number, and distribution of plants in a given geographical setting.
Fluctuating resources theory A theory formulated for plants, predicting that pronounced fluctuations in resource availability enhance community invasibility if they coincide with the availability of sufficient propagules to initiate an invasion. In full, **fluctuating resources theory of invasibility**.
Flux the flow of any material measured as quantity per area per unit of time.
Focal landscape An area being evaluated in an HRV (historic range of variability) analysis.
Focal species A species that can be monitored and managed to maintain the integrity of the ecological system of which it is a part.

Foliage projective cover (FPC) The horizontal cover of leaves (in crowns of varying foliage density) in the overstory and understory strata of a plant community.
Food chain A linear representation of the feeding links between consumers and their resources, based on the pattern of a given organism consuming something earlier in the chain and then itself being consumed by something later in the chain. A simple version might show plants (producers) being consumed by plant-eating animals which in turn are consumed by meat-eating animals. A food chain can be a set of species within a food web.
Food pump A sclerotized pump used by hemipterans to suck up the juices (plant or animal) on which they feed.
Food pyramid A graphic representation of the energy relationships of a community, in which the total amount of energy available at each successive trophic level is proportional to the width of the pyramid at the appropriate height.
Food web A set of trophic (feeding) relationships among species in a community; schematically, the pattern often resembles a web of species each connected by trophic interactions with other species.
Footprint The area of land or water covered or affected by a project. This can include direct physical coverage (i.e., the area on which the project actually stands) and the area indirectly affected by the project (through disturbances such as noise emanating from the project). See also CARBON FOOTPRINT; ECOLOGICAL FOOTPRINT.
Foraminifera An order of animal-like protists, many of which secrete calcareous skeletons (tests).
Forb or Phorb A term for nongraminoid herbaceous plants (non-grasses); a shortened form of Euphorbiaceae, the family to which they belong.
Forcing An initial agent of a climate change, such as an increase in solar irradiance (positive forcing) or an accumulation of atmospheric aerosols, as from a volcano (negative forcing).
Forest change process Any activity on forest land that can affect the stand or site and, in particular, the biological diversity. Can include forest degradation or fragmentation, or the establishment of new plantations.
Forest degradation Damage to forest ecosystems through human activities that does not result in the total elimination of forest cover.
Forest enrichment The process of planting trees in degraded or damaged forests to help recover forest density and/or to improve biodiversity by planting species otherwise unable to colonize and regenerate, or that are threatened or vulnerable.
Forest fragmentation See FRAGMENTATION.
Forest land Descriptive term for an area with a minimum tree height of 5 m, a tree canopy cover exceeding 10 percent, and a minimum area of 0.5 ha.
Forest line Another term for the TREE LINE.
Forest management A pattern of human activity intended to derive economic benefit or other utility through multiple and sustainable use of a forest; a complex concept integrating ecological, socioeconomic, and institutional perspectives. Forests can be managed for timber or nontimber forest products, forage for animals, watershed protection, or recreational use.
Forest regeneration The process of forest regrowth on land that was formerly forested; e.g., following a fire.
Forest stand A relatively homogeneous forest landscape unit that can be distinguished from neighboring units by age or composition.
Forest tundra See SUBARCTIC.
Formation A large-scale (subcontinental or continental) vegetation complex defined primarily on the basis of a combination of dominating life forms.
Fossil Anything found in a strata of rocks or sediments that is recognized as the remains of an organism from a former geological time.
Fossil fuels Forms of stored energy produced by the action of pressure and temperature on organic matter buried over geologic time. The major types of fossil fuels are oil, natural gas, and coal.
Fossorial 1. Burrowing; adapted for living or digging underground. 2. A burrowing animal.
Foundation species The group of critical species whose effects and interactions define much of the structure of a community.
Founder effect A random genetic sampling in which only a few organisms derived from a large population initiate a new population. Since these founders carry only a small fraction of the parental population's genetic variability, radically different gene frequencies can become established in the new colony.
Founders 1. See FOUNDER EFFECT (previous). 2. Animals that were collected from their natural range and that have produced progeny in captivity (and are assumed to be unrelated). Thus, these animals are the founding stocks of the captive population. 3. More generally, animals with wild parents that are collected in order to supplement the genetic diversity of an established captive population.
Fragmentation The disruption of an extensive habitat into smaller, isolated patches.
Freon A brand name for CHLOROFLUOROCARBONS.
Front edge A term for the stable margin of a species' range, where population size is growing.
Frugivorous Describing an animal that eats fleshy fruits; e.g., certain birds and primates. Such an animal is a **frugivore**.
Frugivory The fact or process of feeding on fleshy fruits.
Fruiting body The sexual reproductive structure of a fungus.
Frustule The hard, silica-containing skeleton of diatoms (green algae).
Fry Juvenile salmon that have absorbed their yolk sac and emerged from the water bottom.
F_{ST} Fixation index.
Function Any process carried out by an organism or group of organisms. See also ECOSYSTEM FUNCTION.
Functional biodiversity A characterization of a biota that relates to variation in the importance (e.g., relative abundance, biomass, and relative frequency of occurrence) and functional attributes of taxa, and estimated by components such as guild richness, evenness, dominance, and diversity.
Functional connectivity The processes of dispersal and gene flow of organisms among habitat patches.
Functional constraint A limit to the kinds of mutations that can appear in a gene or protein without compromising function. Also called a SELECTIVE CONSTRAINT.

Functional dependency The dependency of one species on another to complete a particular ecosystem process or function; e.g., the dependency of plants on decomposers, or decomposers on plants, to recycle nutrients.
Functional diversity The range and value of those species and organismal traits that influence ecosystem functioning.
Functional extinction Extinction in effect though not necessarily literally in the death of all individuals; can describe a species that is extinct in the wild, that exists but is on the threshold of extinction; that is no longer capable of reproducing, and so on.
Functional gene A recognizable, heritable segment of DNA that encodes for proteins involved with cell metabolism.
Functional groups Polyphyletic suites of species that share ecological characteristics and play equivalent roles in natural communities and ecosystems. Commonly, organisms with convergent anatomical, morphological, physiological, behavioral, biochemical, or trophic characteristics are grouped together. Also called **functional effect groups**.
Functional redundancy An ecosystem property by which more than one species performs a given role so that when one species disappears, the loss of the ecosystem's efficiency as a whole is relatively small. However, if several such species are lost, the system is likely to essentially collapse. Also called **functional similarity**.
Functional trait space A space where species are located according to their functional traits. This may be multidimensional, with each functional trait representing a separate dimension.
Functional traits Those characteristics of species that define their ecological roles.
Functional type A group of species that share morphological and physiological characteristics that result in a common ecological role.
Functional web See INTERACTION WEB.
Fungus A member of the kingdom Fungi, such as mushrooms, molds, or the fungal part of lichens. Members of this kingdom are heterotrophic (i.e., they require a preformed organic source, rather than being able to make their own food), eukaryotic, have rigid cell walls containing chitin, and reproduce by means of spores.
Furca (plural **furcae**) **1**. A forked structure. 2. Specifically, the telson (posterior segment) of crustaceans.
Fynbos A habitat type in southern Africa that is characterized by thickets and low shrubs, in which fire plays a dominant role in ecosystem maintenance. Plant endemism is particularly high in these areas.
GABA receptor One of a class of receptors for GABA (gamma-aminobutyric acid) found in both vertebrates and invertebrates nervous systems, where they mediate rapid, mostly inhibitory synaptic transmission. Insect GABA receptors are the target site for certain insecticides and pesticides.
GABI Great American Biotic Interchange; see GREAT AMERICAN FAUNAL INTERCHANGE.
Gadoid 1. Describing the fish family Gadidae, which includes the true cod and other important food fish. 2. A fish belonging to this family.
GAFI Great American Faunal Interchange.
Gamete A mature sexual reproductive cell, such as a sperm or egg, that can unite with another such cell to form a new organism.
Gametogenesis The developmental process for the formation of functional male and female haploid reproductive cells that combine at fertilization to produce the zygote.
Gametophyte In an organism with an alternation of generations, the haploid, gamete-producing phase.
Gamma (γ) diversity The total number of species within a large area or landscape, encompassing a number of samples from different communities. Compare ALPHA DIVERSITY; BETA DIVERSITY.
Gamma landscape diversity The total number of patch types contained within a geographic region.
Gap 1. An opening in forest vegetation created by disturbances such as clearing, logging, fires, diseases, storms, or the natural death of a tree. The process by which this occurs is **gap dynamics**. 2. A "hole" in the ecological safety net; i.e., a species or class not well represented in areas designed to protect biodiversity. See GAP ANALYSIS (next).
Gap analysis A biodiversity policy tool that compares the actual distribution of a large region's species and vegetation classes to what is found in areas managed for long-term protection of biodiversity. Species and classes that are poorly represented in protected areas are gaps in the protection of biodiversity, and thus are candidates for proactive protection.
Garrigue A name for the low-growing secondary evergreen shrublands that dominate extensive areas of the Mediterranean Basin.
GCM General circulation model.
GBA Global Biodiversity Assessment.
GEF Global Environment Facility.
Gene The functional unit of heredity; the place (locus) on the chromosomes of an organism that encodes for a particular function.
Gene bank A facility where genetic material is stored in the form of seeds, pollen, embryos, or semen, or, in some cases, as live plants or animals living in a field or a greenhouse or other installation.
Gene conversion A process of nonreciprocal (unidirectional) transfer of genetic information.
Gene flow The movement of genes between populations, usually via immigration and mating of whole genotypes, but sometimes single genes may undergo horizontal gene transfer via transfection by microorganisms.
Gene frequency The frequency of an allele within a population.
Gene pool The total genetic diversity found within an individual population or species at a given time.
General circulation model (GCM) A computer model developed to simulate global climate; GCMs are widely used for global climate change predictions.
Generalist A brood parasite species that may parasitize many (up to more than 200) host species.
Genetic Having to do with genes or the study of genes (genetics).
Genetically modified organism (GMO) An organism produced through the introduction of a gene or genes that were transferred from another organism, usually by means of

biotechnology. There is current debate over the role of GMOs in agriculture.

Genetic architecture A characterization of the number and types of genes that underlie a particular trait, and their interactions.

Genetic code The language that specifies how DNA will be translated into amino acid sequences by means of three nucleotide "words" (codons) that specify the 20 amino acids and regulators of transcription (*start* and *stop* codons).

Genetic diversity Genetic variety found within or among species; this allows the population or species to adapt and evolve in response to changing environments and natural selection pressures.

Genetic drift The process of random changes in gene frequency, leading to changes in the genetic composition of a population. The effect of this process is particularly important in small population sizes.

Genetic erosion A loss of genetic diversity in a species, often caused by anthropogenic factors such as land clearing or urbanization.

Genetic load The extent to which an individual is inferior to the best possible individual in the population in terms of fitness, as a consequence of the genes it contains.

Genetic recombination The physical exchange of genetic material between a pair of chromosomes during meiosis.

Genetic variation Differentiation in traits that is attributable to genetic differences among the individuals in a population.

Genetics The scientific study of genes, in the context of an individual cell or organism, the passing of genes from parent to offspring, and the distribution, variation, and alteration of genes in populations.

Genome The complete set of genes of an individual organism.

Genomic Having to do with the genome (see previous).

Genomic cluster See GENOTYPIC CLUSTER.

Genomics The discipline of genetics that focuses on the study of the complete genetic content of species, often by determining the entire DNA sequence of organisms and the expression and function of their genes in relation to various stimuli or metabolic perturbations.

Genotype The genetic makeup of an organism. Compare PHENOTYPE.

Genotypic Relating to the genetic composition of an organism.

Genotypic cluster In a local area, a single group of individuals recognizable on the basis of multiple, unlinked inherited characters or genetic markers.

Gentes (singular **gens)** 1. Brood parasite species with a host-specific lineage. 2. Specifically, lineages of cuckoos in which individual females specialize on a single host and lay mimetic eggs.

Geographic information system (GIS) Any of various computational systems for storing, managing, manipulating, and displaying spatial data. Because landscape models are inherently spatial, strong links to GIS are either built into such models or are required to interface with the model's data.

Geomorphological Having to do with or used in geomorphology (see next).

Geomorphology The study of the surface configuration of the Earth, especially the nature and evolution of current land forms, their relationships to underlying structures, and the history of geological activity as represented by such surface features.

Geophyte A plant with succulent storage tissues such as bulbs, corms, or rhizomes.

Geospatial Having to do with locations on the Earth.

Geothermal Having to do with or produced by the internal heat that exists below the surface of the Earth. Thus, **geothermal energy, geothermal heat(ing)**.

Germplasm 1. Sources of hereditary material. 2. In animal breeding, the breeding animals and fresh or frozen sperm cells, ova, and embryos.

GHG greenhouse gas.

ghost fishing A term for the fact that lost, discarded, or abandoned fishing gear (such as **ghost nets**) can still continue to catch fish.

Gini–Simpson index An index that quantifies the probability that two entities (such as individuals) drawn at random from the data set represent different types (such as species). Often used as a diversity index, but not an expression of diversity itself. [Named for the statisticians C. W. Gini and E. H. Simpson.]

GIS Geographic information system.

Global Biodiversity Assessment (GBA) An independent, peer-reviewed analysis of the biological and social aspects of biodiversity, commissioned by the United Nations Environment Programme.

Global change disturbance An alteration or disturbance significant enough to have global impact on the natural conditions of the atmosphere (such as GHGs produced by human activities), or on various regions throughout the world (such as large-scale deforestation or the widespread introduction of exotic, invasive species).

Global circulation The regular circulation pattern of the Earth's atmosphere, which generates the world's basic climate types; the system involves an Intertropical Convergence zone of low pressure near the equator, subtropical high-pressure belts near the Tropics of Cancer and Capricorn, trade winds flowing from these high-pressure belts toward the equatorial low, and westerly winds in the middle latitudes.

Global climate change See CLIMATE CHANGE.

Global Environment Facility (GEF) An independent financial organization providing grants to developing countries and countries with economies in transition, for projects related to biodiversity, climate change, international waters, land degradation, the ozone layer, and persistent organic pollutants.

Global environmental sustainability See ENVIRONMENTAL SUSTAINABILITY.

Global extinction The fact of a species becoming extinct worldwide, as opposed to *local extinction* which applies to a particular area.

Globalization The process in which businesses and financial markets operate internationally, or increase the extent to which they do so, largely as a result of increased trade, transportation, and communications.

Global warming The warming trend of the Earth's surface temperatures from the late 19th century onward and

particularly in the last several decades, caused by the incremental injection of carbon dioxide and other greenhouse gases (GHGs) into the atmosphere from human activities such as deforestation and fossil fuel consumption. Though literally the term could also apply to earlier natural warming periods in Earth's history, in practice it refers specifically to the current period. See also CLIMATE CHANGE.

Glycogen An insoluble polymer of glucose; the main storage carbohydrate in animals.

GMO genetically modified organism.

Gnetales A small relict order of chiefly tropical or xerophytic gymnosperms that includes the genera Gnetum, Welwitschia, and Ephredra.

Gondwana The southern supercontinent, named after an ancient kingdom in India. Gondwana drifted apart over many millions of years, giving rise to the current Southern Hemisphere land masses of South America, Africa, Australia, and India. Also, **Gondwanaland**.

Gondwanic having to do with or situated in the former supercontinent of Godwana (see previous).

Gonochoristic Describing an individual that is either a male or a female throughout life. This condition is **gonochorism**.

GOS Global Ocean Sampling; a program to study the genetic diversity of microorganisms in the world's oceans.

Grade Another term for a PARAPHYLETIC GROUP.

Gradient A transitional geographic area in which environmental conditions vary.

Gradualism The evolutionary accumulation of small individual quantitative changes in a population, ultimately leading to qualitative change. Contrasted with SALTATION.

Grain A component of spatial scale used in landscape ecology; the size of a single observation.

Graminoids The grasses (family Poaceae) and and also grass-like plants, mostly sedges (family Cyperaceae) and rushes (family Juncaceae).

Granivorous Describing animals that eat seeds or achenes (grass fruits). Such an animal is a **granivore**.

Graze To consume herbaceous plants, including grass, grass-like plants, and forbs.

Grazer A consumer that removes only a part of each prey it attacks and thus rarely kills a prey in the short term. In the traditional sense, this applies to animals that eat plants (herbivores), but it can also apply to nonlethal microbial pathogens and to animals that cause tissue loss, but not death.

Grazing intensity The frequency and closeness of grazing.

Grazing pressure Stress on plant populations due to a high proportion of grazing animals compared to forage available in the given area.

Great American Faunal Interchange The major exchange of fauna (e.g., land mammals, birds) between South and North America after the formation of the Isthmus of Panama, about 2.7 million years ago. Also known as the **Great American Biotic Interchange** or simply **Great Faunal Interchange**.

Great Chain of Being See SCALA NATURAE.

Greenhouse effect The process that traps solar radiation in the lower atmosphere and makes the Earth warmer, in analogy with the way a greenhouse warms incoming air to facilitate the growing of plants.

Greenhouse gas (GHG) A gas in the atmosphere, such as water vapor or carbon dioxide, that can affect climate through the absorption and re-emission of terrestrially emitted infrared radiation. Grennhouse gases are so called because they contribute to the greenhouse effect (see previous).

Grilse An Atlantic salmon that spends only one year at sea before returning to spawn.

Grinnellian niche The sum of the requirements and behaviors expressed by a species in the habitat where it normally occurs. [Described by U.S. naturalist Joseph Grinnell.]

Groundwater A "reservoir" of water occurring below ground in saturated soils and beneath geologic formations.

Group selection Selection between different populations or sub-populations based on attributes of the entire group, where these attributes usually are either selected against or not favored at the level of individual selection.

Growth-form See LIFE-FORM.

Guild A group of species in a community that use similar environmental resources in similar ways.

Ha Hectare.

HAB Harmful algal bloom. See ALGAL BLOOM.

Habitat A living space; the place where an organism or species normally lives, including both its biotic and abiotic elements.

Habitat corridor or connection See CORRIDOR.

Habitat fragmentation The process by which a once continuous area of habitat is divided into two or more disconnected parts, as by roads, farms, fences, logging, and other human activities. It occurs when habitat is lost from the interior, rather than the edge, of a large block of cover.

Habitat island One of the resulting patches that occurs as a result of habitat fragmentation.

Habitat loss A change to an area that prevents a species from living there (e.g. conversion of forest to crop field).

Habitat patch A spatially contiguous area in which environmental variables are similar.

Habitat remnant A small area that is relatively undisturbed and that represents a much larger historical ecosystem.

Habitat restoration Active modification of the current state of a degraded habitat in order to return it to a former, preferred state.

Habitat selection The act or fact of an organism's choosing a combination of available abiotic and biotic elements for the purpose of fulfilling its life history events.

Habitat structure The physical aspects of a habitat; may be characterized by such measures as complexity, heterogeneity, regularity, stratification, and fractal dimensionality.

Hadley cell A pattern of air circulation that causes deserts at about 30° north and 30° south of the equator. The Sun heats the equator, air rises, pressure decreases, moist air moves north and south, cooling as it goes, and then falls onto the spinning earth, drying as it falls. The falling dry air increases pressure preventing the incursion of moist air. [Named for the English meteorologist George Hadley.]

Halophile A salt lover; an organism that is able to tolerate and/or grows optimally at high salt concentrations. Such an organism is **halophilic**.

Haploid 1. Describing a cell with only one complete set of chromosomes, in contrast with a DIPLOID cell with two sets. 2. An organism having such cells.

Haplorrhini One of the two recognized suborders of primates, along with the Strepsirrhini (see); includes New World and Old World monkeys, apes, and tarsiers.
Harmful algal bloom (HAB) See ALGAL BLOOM.
Hatchery A facility in which aquatic species are spawned and resulting juvenile progeny are reared for at least part of their life cycle.
Haustorium (plural **haustoria**) An organ found in parasitic plants that penetrates vascular tissue of the host plant and forms a functional bridge for uptake of water and nutrients from the host.
Heat equator See THERMAL EQUATOR.
Heath Open vegetation dominated by dwarf undershrubs, usually of the family Ericaceae.
Hectare A metric measure of surface equal to 10,000 m^2 or 2.471 acres (U.S.). Abbreviated **ha**.
Helminth One of several classes of parasitic worms, including cestodes (tapeworms), trematodes (flukes), nematodes (roundworms), and monogeneans (flatworms).
Hemelytron The forewing of a heteropteran insect; so called because the anterior part of the wing is opaque and the posterior part is clear (membranous).
Hemiepiphyte A plant that lives part of its life cycle as an epiphyte (i.e., growing nonparasitically on another plant).
Hemimetabolous Having incomplete metamorphosis, that is, showing gradual change from molt to molt, with externally developing wing pads, and lacking any larval and pupal stages.
Hemiparasite A parasite that is also able to carry out photosynthesis; e.g., mistletoe. Also called a SEMIPARASITE; compare HOLOPARASITE.
Herbaceous Having the characteristics of an herb; i.e., having little or no woody tissue and typically persisting for only one growing season in its aboveground growth.
Herbal 1. Having to do with herbs, or medicines derived from herbs. 2. A book or other compilation recording accumulated knowledge of medicinal plants.
Herbicide A chemical used to suppress or kill plants, or to severely interrupt their normal growth processes.
Herbicide resistance The inherited ability of a plant to survive and reproduce following exposure to a dose of herbicide normally lethal to the wild type. Resistance may be naturally occurring, or induced by techniques such as genetic engineering or selection of variants produced by tissue culture or mutagenesis.
Herbicide tolerance The inherent ability of a species to survive and reproduce after herbicide treatment. This indicates that there was no selection or genetic manipulation to make the plant tolerant; i.e., it is naturally tolerant. Thus contrasted with herbicide resistance; see previous.
Herbivorous Describing animals that feed exclusively or mainly on plants. Such animals are **herbivores**.
Herbivory The act or fact of feeding on plants. A distinction can be made between feeding patterns of grazing or browsing, which do not destroy the plant, and one where a single animal eats an entire plant, which is characterized as predation rather than herbivory.
Heritability The proportion of the variance in a trait that is due to additive genetic effects.
Heritable Able to be inherited.
Hermatypic Reef-building; referring to coral species.
Heterochromatin Regions of chromosomes that in general do not include coding DNA, always remaining condensed during a cell's life cycle.
Heterocyst The site for N_2 fixation in some species of filamentous cyanobacteria. A heterocyst is a differentiated cell that may be present terminally or intercalary. It lacks photosystem II, but photosystem I is present, providing a means to generate additional energy (ATP) required to fuel N_2 fixation.
Heterogamete Either of a pair of gametes that differ in form, size, or behavior. The production of such gametes is **heterogamety**.
Heterogeneity 1. A variation in environmental conditions through space (spatial) or time (temporal). 2. The fact of being heterogeneous (see next).
Heterogeneous Made up of differing components; having different parts or elements. Compare HOMOGENEOUS.
Heterokont One of a monophyletic group of eukaryotes, mostly algae, having cells with heterodynamic flagella. Also, STRAMENOPILE.
Heteropteran An insect of the order Heteroptera, the so-called true bugs.
Heterostyly The fact of having two or more kinds of flowers with stamens and pistils of different lengths. Such a plant is **heterostylous**.
Heterotroph Literally, a feeder on others; an organism that is dependent on organic material from an external source to provide carbon for growth; e.g., vertebrates. Such an organism is **heterotrophic**.
Heterotrophic respiration The respiration of those life forms that feed on photosynthetic organisms, including bacteria, fungi, and animals.
Heterotrophy The mode of nutrition in which organisms grow and reproduce on the basis of external sources of organic carbon sources. Compare AUTOTROPHY.
Heterozygosity The presence of alternate alleles at a given locus in a diploid organism (e.g., A_1A_2). Compare HOMOZYGOSITY.
Heterozygous Having two different alleles at a different locus.
Hexapod A member of the Hexapoda, the group including insects and their primitive relatives.
Hibernation A dormant state in certain animals, characterized by inactivity and lower metabolic rates; typically a strategy to conserve energy when food supplies are reduced (as in winter in cold climates).
Hierarchical (habitat) selection The fact that an organism's selection of habitat occurs at varying scales of space and time. First-order habitat selections define species range; second-order define the site of an individual's home range within that species range; third-order specify within home range selections; and fourth-order are fine-scale daily habitat selections.
Hierarchical partitioning The process of partitioning regional diversity between alpha (local) and beta (larger) components.
High seas Those regions of the ocean outside national boundaries, usually beyond 200 miles of any nation's coast.

Hill numbers A continuum of diversity measures that quantify diversity in units of equivalent numbers of equally abundant species. [Developed by British ecologist Mark Hill.]
Hinterland A region lying beyond an urban area, but often influenced by it.
Historic rainforest All land areas with a potential to support the tropical rainforest as determined by the climatic and physiographic conditions, whether they are currently forested or not. This concept provides a baseline to calculate the rainforest loss in a country or region over a historic period.
Historic range of variability (HRV) A description of the full range of landscape characteristics that occurred in the past. Also, **historic range of variation**.
Historical biogeography See BIOGEOGRAPHY.
Historical ecology The study of how human societies and ecosystems change and in turn transform one another through time in local and regional landscapes. Data are drawn from geology, archeology, history, and various other sources.
Historical processes The evolutionary and ecological processes that occur in the recent or distant past including immigration, emigration, speciation, extinction, drift, and/or past selection by the environment.
Holarctic A biotic realm in the temperate, boreal, and arctic regions of the Northern Hemisphere with strong affinities among the regional floras and faunas.
Holistic management A philosophy and practice of livestock grazing management, used by a minority of ranchers in Africa, North America, and Australia, attempting to integrate economic, social, and ecological goals.
Holobiont A collective term for all the organisms engaged in a symbiotic relationship; e.g., the coral animal and the diverse assemblage of microbes that live associated with the coral.
Holocene The epoch in time extending from about 11,700 years ago, approximately when the last ice age ended, to the present day.
Holometabola Insects characterized by complete metamorphosis, a wingless larval stage, and an intermediate pupal stage.
Holomorphology The totality of characters that are indicative of phylogeny (molecular, behavioral, and so on).
Holoparasite A parasitic plant that lacks chlorophyll and cannot photosynthesize. Compare HEMIPARASITE.
Holoplankton See ZOOPLANKTON.
Homeostasis The ability of an organism or other biotic system to maintain a state of relative internal stability in the presence of external change.
Home range The area where an animal typically spends its time; the expanse traversed in meeting life requirements over a given period.
Homogeneous Made up of components that are of the same kind; having uniform parts or elements. Compare HETEROGENEOUS.
Homogenization A blending of species throughout the globe, largely as a result of trade and transportation.
Homoiohydry The ability of an organism to maintain a constant internal water balance independent of fluctuating environmental conditions. Contrasted with POIKILOHYDRY.
Homologous Denoting common ancestry; describing structures, processes, sequences, behaviors, and so on that are derived from a common ancestral structure.
Homology The existence of a similarity that is explained by common ancestry.
Homonymy The fact of the same scientific name being applied to two or more species.
Homoplastic Relating to or exhibiting homoplasy (see next).
Homoplasy A character or character state acquired by parallel or convergent evolution that bears resemblance to a character in a different group.
Homozygosity In a diploid organism, the presence of the same alleles at a given locus (e.g., A_1A_1). Compare HETEROZYGOSITY.
Homozygous Possessing the same gene form (allele) on both chromosomes.
Horizon See SOIL HORIZON.
Horizontal transmission The transmission of pathogens within generations of hosts.
Host The individual that is parasitized by a parasite.
Host range The range of different organisms in which a virus (or a plasmid) is able to replicate successfully.
Hotspot 1. An area with exceptionally high biodiversity, especially one threatened by habitat loss. Called "hot' because such areas are often selected as priority targets for conservation or management efforts. 2. A region that is unusually rich in a particular group of species (e.g., birds), compared with average sites in the same geographic region; contrasted with a COLD SPOT. Also spelled **hot spot**.
HRV historic range of variability.
Human-altered Describing a habitat or ecosystem that has been significantly changed by humans; e.g., the conversion of natural forest into pasture or the reclamation of a wetland for agriculture.
Human-dominated Describing a habitat or ecosystem where human presence and activities strongly influence living and nonliving ecosystem components.
Human ecology A type of ecology focused on the species *Homo sapiens* (humans), including their relationship with their natural, manufactured, social, and cultural, and environments. Can be considered a subset of SOCIAL ECOLOGY.
Humanism A philosophy focusing on the needs, interests, and values of humans. **Ecohumanism** is the belief that this must include concern for the natural environment because of its value to humans.
Human well-being People's condition and capacities allowing secure access to food, water, energy, and shelter of adequate quality to meet their needs, and provide good health as well as social connections.
Husbandry The practice of raising livestock and growing crops.
Hutchinsonian niche The set of environmental conditions, or opportunities, that will permit a species to exist indefinitely. [Named for the British-American ecologist G. Evelyn Hutchinson.]
Hutchinsonian ratio The body size ratio of larger species over smaller species in a pair of species; according to this coexisting species are expected to have a greater body size ratio than would occur by chance, because competition prevents coexistence of species that are morphologically too similar.
Hybrid An individual that results from crossing between individuals from two or more populations, races, subspecies,

species, and so on, including individuals that are F1s, F2s, and the set of all backcrosses.

Hybrid inviability A phenomenon whereby one or both sexes of progeny produced in crosses between two different species is unable to survive.

Hybridization Mating between individuals of two different populations, usually classified as separate species; mixing of their genomes is not required.

Hybrid sterility A phenomenon in which one or both sexes of progeny produced in crosses between two different species are able to survive, yet are unable to reproduce.

Hybrid swarm A mixture of hybrid genotypes, including F1s, F2s, and backcrosses, due to hybridization between two or more species that co-occur in a locality.

Hybrid zone An area of overlap or point of contact between two populations that are distinguishable based on one or more heritable characters where viable or partially fertile hybrids are formed.

Hydrologic(al) Having to do with water or the study of water.

Hydrologic connectivity The ability for water to move between portions of a river system, such as between a river channel and floodplain surfaces.

Hydrologic regime The aggregate set of variables describing the behavior of water in wetlands, including water depth, magnitude and frequency of water level fluctuations, seasonal pattern of water fluctuations, direction and velocity of waterflows, and water residence time.

Hydrology The scientific study of the properties, distribution, movement, and effects of water on the Earth's surface.

Hydrophilous 1. Water-loving; growing or thriving in water. 2. Pollinated by or in water.

Hydrophily Underwater pollination developed by most seagrasses and many freshwater angiosperms.

Hydrophyte A plant capable of growing and persisting in saturated or flooded soils. Such a plant is **hydrophytic**.

Hydrothermal plume A plume created by the heat of volcanic processes in the ocean, rising as "black smoker" effluents that mix turbulently with the surrounding seawater, then become dilute and eventually reach neutral buoyancy (typically at about 200 m above the black smoker source. Once neutral, these plumes spread laterally and are detectable by chemical and physical tracers.

Hypercycle A cyclic sequence of self-supporting reactions between primitive prebiotic biomolecules, nucleic acids, and proteins. The **hypercyclic** organization of reacting systems is postulated to explain the spontaneous emergence of replicating systems. As a result of the **hypercyclic reaction principle**, functional characteristics of replicating molecules (i.e., phenotypes) are connected to their hereditary characteristics or their genotypes.

Hyperparasite A parasite living on or in another parasite. Such an organism is **hyperparasitic**.

Hyperparasitism The fact of living parasitically on or in another parasite.

Hyperparasitoid A parasitoid of a parasitoid; e.g., a parasitoid wasp larva living in a parasitoid fly larva that feeds on a beetle.

Hypersaline More salty than sea water.

Hyperthermophile An extreme heat-lover; an organism that grow optimally at temperatures higher than 80° C. Such an organism is **hyperthermophilic**.

Hyperthermophily The fact of growing optimally at temperatures higher than 80^0 C.

Hyphae (singular **hypha**) The feeding (vegetative) threads of most fungi. The hyphae of one organism are sometimes referred to as *mycelia*.

Hypodermis The cellular, subcuticular layer that secretes the cuticle of annelids, nematodes, arthropods, and various other invertebrates.

Hypognathous Describing the position of the head when the mouth is directed toward the ventral side of the body.

Hyporheic zone The area of saturated soils beneath a stream or a river channel; it interacts with the stream through the processes of hydraulic upwelling and downwelling.

Hypotelminorheic A persistent wet spot, a kind of perched aquifer, rich in organic matter, underlain by a clay layer typically 5–50 cm beneath the surface, with a drainage area typically less than 10,000 m^2.

Hypoxia 1. A lack of oxygen; a decreased oxygen supply. 2. Specifically, a condition in which water has depressed dissolved oxygen levels that generally are too low to sustain healthy populations of most fish species for extended periods.

Hypoxic 1. Lacking oxygen; characterized by decreased levels of oxygen. 2. Specifically, describing water with depressed oxygen levels; many fish will actively avoid hypoxic areas.

Hypsodont A high-crowned cheek tooth commonly found in grazing mammals.

Hysteretic response A phenomenon in which the response of a physical system to an external influence depends not only on the present magnitude of that influence, but also on the previous history of the system.

Idiobiont A parasitoid that develops on a paralyzed, incapacitated host.

Igapó 1. Vegetation periodically flooded by acidic black-water or clear-water rivers, as in the Amazon basin. 2. A floodplain with such vegetation.

ILUC Indirect land-use change.

Imago An adult insect; the final developmental stage when insects possess wings.

Immobilization The result of microbial decomposition of organic matter in which nutrient elements (particularly nitrogen) are taken up into the microbial biomass, rendering them unavailable for plant uptake. See also MINERALIZATION.

Impoverishment On an island, a paucity of species relative to the adjacent mainland.

Inbreeding Mating between close relatives; this results in progeny that carry two copies of the same allele at many loci throughout the genome.

Inbreeding depression Reduction in fitness due to inbreeding.

Incidence The number of samples containing at least one of a given species in a census.

Inclusive fitness An individual's lifetime reproductive output if not assisted by others, plus the added reproduction it confers on each relative it helps, times the proportion r of genes identical by recent descent it shares with that relative. It thus represents the common interest of the autosomal genome, which is what natural selection maximizes.

Inclusive wealth The sum total of capital assets or human and Earth system properties that can be used to provide the flow of goods and services that contribute to well-being.
Index of aridity The ratio of mean annual precipitation to mean annual potential evapotranspiration.
Indicator A species or community that enables observers to evaluate environmental conditions and detect changes.
Indicator species are normally surrogates of other species in the area of interest and are usually sensitive to environmental change. Environmental variables may also be used as indicators.
Indigenous 1. Originating in and living naturally in a certain place; native. 2. Relating to indigenous peoples (see below).
Indigenous knowledge Traditional understandings of indigenous peoples, often orally transmitted from generation to generation.
Indigenous peoples Groups of human beings who have resided in the same geographical area for many generations, usually many centuries, and who possess legends, proverbs, genealogies, languages, and other unique cultural features linking them to this site.
Indirect effect An effect on a third species or trophic group owing to a direct interaction between two species or trophic groups. Trophic cascades are the result of the indirect effect of predator–prey interactions on a third group.
Indirect land-use change (ILUC) Conversion of land from one land-use category to another, induced by the expansion of biofuel production elsewhere.
Indirect use values Benefits gained from environmental goods and services that are derived from supporting and protecting activities as opposed to direct use of the resource, such as coastal habitats serving as breeding and nursery grounds for off-shore fisheries, coral reefs providing protection against storm surge, or upper watershed forests controlling soil erosion and runoff. Compare DIRECT USE VALUES.
Individual-based (abundance) data A common form of data in biodiversity surveys. The data set consists of a vector of the abundances of different species; used when an investigator randomly samples individual organisms.
Individualism (of species) The observation that distributions and abilities of most species do not exactly coincide because each species is the product of a unique evolutionary history. Coevolution and convergence may occur, but rarely are these processes so complete that groups of species are distributed as a unit or superorganism.
Individual selection Selection driven by differences in the net reproduction of individual organisms.
Individual transferable quota (ITQ) A type of market-based incentive in which regulators set an aggregate limit on harvests of an open access resource (typically fish), assign rights to the allowable harvest to individual participants, and permit participants to trade those rights.
Induced defense A plant defense, including both chemical and physical defenses, that is produced, at least in final form, only after the plant has been damaged by herbivores or pathogens.
Industrial agriculture A modern form of agriculture that differs from TRADITIONAL AGRICULTURE (see) in the use of elaborate and expensive machinery, the control of pests with toxic chemicals rather than through biological controls, fertilization by synthetic rather than organic products, excessive consumption of water, and farm ownership and management by corporations rather than individuals. Industrial agriculture has significantly greater environmental impact than traditional agriculture. Also known as **industrialized agriculture**.
Infauna Animals living within sediments.
Infection The presence of pathogens within a host's body.
Infectious disease A state of compromised health in a host caused by a pathogen infection.
Input environment A collective term for all energy and materials moving into a given system.
Insecticide A toxic chemical used to suppress or kill insects, or to severely interrupt their normal growth processes.
Insecticide resistance The inherited ability of an insect to survive and reproduce following exposure to a dose of insecticide normally expected to be lethal to individuals of this species.
Inselberg An isolated hill, rock, dome, small mountain, or the like that stands out from the surrounding lower elevations. [German for "Island Mountain."]
In situ Literally, "in position;" in its natural setting, as opposed to a laboratory, zoo, or other artificial setting (*ex situ*).
In situ conservation 1. Conservation of natural ecosystems and habitats, and the maintenance and recovery of viable populations of species, in their natural surroundings. 2. Conservation of domesticated or cultivated species in the surroundings where they have developed their distinctive properties. Compare EX SITU CONSERVATION.
Instar A stage between molts of immature insects.
Instrumental value The value of something in terms of human interests or desires.
Insurance hypothesis The principle that biodiversity "insures" ecosystem function, because more species guarantee that some will continue to function even if others fail, thereby enhancing the capacity to recover from and mitigate the risks caused by disturbance. According to this, biodiversity can be said to have **insurance value**, because it reduces risk for those who make use of a given ecosystem service.
Integrated conservation and development A project that links biodiversity conservation, usually of a protected area, with local socioeconomic development activities in which alternative incomes or livelihoods are introduced to reduce pressure on local natural resources.
Integrated weed management An approach for suppressing weeds that combines information on the biology and ecology of the weed with all available control technologies, so that no one method is used exclusively.
Intensification See INTENSIVE AGRICULTURE.
Intensity 1. The number of parasites per infected host. 2. See INTENSIVE AGRICULTURE.
Intensive agriculture The management of agricultural land to increase production, mainly through high-yielding crop varieties, chemical fertilizers and pesticides, increased tillage and irrigation, shorter rotations, larger fields, and greater mechanization. May also include converting natural habitat to cropland, fragmenting natural habitat, avoiding fallows, and removal of hedgerows. See also INDUSTRIAL AGRICULTURE.

Interaction web A subset of species that through their interactions and their responses to abiotic factors make up the dynamic core of food webs or communities. These webs include keystone species, dominants, and other strong interactors. Also known as a **functional web**.

Intercropping The agricultural system of growing two or more crops simultaneously on the same field. According to row arrangement and relative growth time, intercropping can be divided in to three types. In **mixed intercropping** the crops are grown with no distinct row arrangement. In **relay intercropping** the crops are grown during part of the life cycle of each. (A second crop is planted after the first crop has reached its reproductive stage of growth but before it is ready for harvest.) In **strip intercropping** the crops are grown in different strips that are wide enough to permit independent cultivation but narrow enough for the crops to interact agronomically.

Interference competition Interaction among two or more species that use a common resource that is not limiting, in which one species is harmed by having its access to the resource restricted.

Intergenerational equity Equality and fairness in the well-being of people in future different generations relative to current generations.

Intermediate disturbance (hypothesis) The proposal that some level of disturbance is necessary to maintain biodiversity.

Intermediate host A host that harbors sexually immature, developing stages of a parasite. Also known as a SECONDARY HOST.

International Species Information System (ISIS) A central repository for zoo animal records, founded in 1973; it has also developed the computer programs that most zoos use for keeping their animal records.

Interspecific Between species; involving organisms of different species. Contrasted with INTRASPECIFIC.

Interspecific brood parasite A brood parasites that lays its eggs in the nests of other species.

Interstices The spaces between sediment particles, such as sand or silt grains.

Interstitial Having to do with or living in interstices (see previous).

Intertropical convergence zone (ITCZ) A low-latitude region characterized by warm air masses and convectional rainfall.

Intraguild predation Species interaction in which predators feed on herbivores as well as other predators, and trophic levels are not well defined.

Intraspecific Within species; involving organisms of the same species. Contrasted with INTERSPECIFIC.

Intraspecific parasitism A parasitic association of members of the same species.

Intrinsic rate of growth A parameter, often designated as the variable *r*, reflecting the rate of change in population size due to births and deaths when resources are not limiting.

Intrinsic value The value of something independent of its value to people.

Introduced species Species transported by humans to regions outside their natural geographic range.

Introduction The placement of individual organisms by people into a habitat for purposes of establishing a population in a place where it does not currently occur.

Introgression The permanent transfer of genes from one or more species, subspecies, or race into another by means of hybridization and backcrossing.

Introns Generally noncoding segments of genes, situated between exons and removed before translation of messenger RNA to proteins.

Invasibility The properties of a community, habitat, or ecosystem that determine its inherent vulnerability to invasion.

Invasion The process in which an organism enters an ecosystem from some other geographic region, often as the direct or indirect result of human activity, and disrupts the ecosystem's composition, structure, or function.

Invasional meltdown The compounded effect on an ecological community of successful and successive invasions of alien species benefiting from the previous introduction of the other.

Invasive Describing an organism that has entered (invaded) an ecosystem from elsewhere, and that has the potential to spread over the area and displace indigenous species, causing environmental and/or economic harm. Thus, **invasive species, invasive plant**, and so on.

Invasiveness The fact of being invasive; the features of an alien organism, such as their life-history traits and modes of reproduction, that define their capacity to invade.

Invasive weeds Introduced plant species that establish self-maintaining populations and spread into new areas where they frustrate human intentions in production and natural landscapes.

Invertebrate An animal without a backbone (vertebrae); i.e., most of the animal kingdom, ranging from unicellular protozoa to multicellular, complex organisms such as insects.

Invisible hand The term used by the Scottish economist Adam Smith to describe the capacity of a decentralized market system to attain efficiency through self-governing mechanisms.

IPAT A mathematical formula for calculating human impact on the environment; expressed as Impact = Population x Affluence x Technology.

Irreplaceability The degree to which a region is necessary as part of a comprehensive conservation strategy. A region with high levels of endemism has high irreplaceability, because it cannot be replaced by others if the endemic species are to be conserved.

Irreplaceable Having the property of irreplaceability (see previous).

ISIS International Species Information System.

Island biogeography theory The theory that the number of species in the biota of an island (or island-like habitat) is in dynamic equilibrium, with species frequently going extinct on the island and new species frequently arriving.

Isocline A line along which some property remains constant.

Isolating mechanism Any intrinsic or extrinsic mechanism or barrier that inhibits the free exchange of genes between populations.

Iteroparity The fact of being capable of repeat breeding, or the process of doing this. Compare SEMELPARITY.
Iteroparous Capable of breeding more than once.
ITCZ Intertropical convergence zone.
ITQ Individual transferable quota.
IUCN International Union for Conservation of Nature; a global environmental organization dedicated to conserving biodiversity, founded in 1948 and based in Switzerland. Noted especially for the IUCN Red List (see next).
IUCN Red List A widely cited, comprehensive compilation of the extinction risk of species on a local or global basis. In full, IUCN Red List of Threatened Species.
Jack Term for a sexually precocious male salmon that spends one winter or less at sea before returning to its natal stream to spawn.
Joint determination The fact that ecological and economic systems mutually affect each other.
Jumping gene Another term for a TRANSPOSON.
Kaldor–Hicks efficiency An economic condition in which those participants who are made worse off by a change to the system could in theory be compensated by those who are made better off, and would be willing to accept this compensation. An alternative to the economic equilibrium described by PARETO EFFICIENCY. [Named for the British economists Nicholas Kaldor and John Hicks.] Also known as **potential Pareto efficiency**.
Karst A landscape in which the primary geomorphic agent is solution rather than erosion; typically formed in carbonates, landforms including caves, sinkholes, blind valleys, and large springs.
Kelt A spawned-out Atlantic salmon or steelhead that is returning to the ocean to recuperate and spawn again.
Key innovation A property of a species that promotes the process of speciation.
Key-industry species Prey of intermediate trophic status that support a large group of consumers.
Keystone species A species whose effects on the structure, dynamics, and functioning of the community is disproportionately large relative to its abundance. Also, **keystone**.
Kinetic energy See ENERGY.
Kleptoplastic Having the ability to harvest chloroplasts from algae; these are then used by the secondary host to make photosynthetic products. This phenomenon is **kletoplasty**.
Koinobiont A parasitoid that develops on a mobile, active host.
Korarchaeota A proposed third kingdom within the domain Archaea, including a small group of ribosomal RNA sequences retrieved from geothermally heated sediments. [From Greek *koros*, young man and *archaios*, ancient; for the early divergence of this group during the ancient evolution of the Archaea.]
Kranz anatomy Specialized anatomy of the C_4 plant leaf in which the bundle sheath (BS) tissue is enlarged and enriched with chloroplasts, whereas the mesophyll is reduced in size and often forms a lighter green halo around the BS. This produces a wreath-like ('Kranz' in the original German) appearance.
K-T An abbreviation for Cretaceous-Tertiary, usually in reference to the K-T event or boundary (see next).
K-T event A notable large-scale extinction event occurring about 65 million years ago, at the boundary of the Cretaceous and the Tertiary (**K-T boundary**), involving the extinction of about 70 percent of extant species, including the dinosaurs and many other land-dwelling vertebrates.
Kuznets curve A hypothesized trend in which pollution and other environmental degradation initially rise, then level off, and eventually decrease as incomes increase and the national economy develops. [Proposed by the Russian-American economist Simon Kuznets.]
Kwongan Evergreen heathlands that cover extensive areas of southwestern Australia.
Land cover The physical materials covering the surface of the earth, natural or man-made; e.g., water, forests, grassland, bare ground, buildings, roads, and so on.
Land degradation Reduction or loss of the biological and economic productivity and complexity of terrestrial ecosystems; see DEGRADATION.
Land equivalent ratio (LER) The area needed for sole cropping to produce the same amount of crops as would be produced in one hectare of intercropping or mixed cropping.
Land ethic A philosophy of proper treatment of lands and the natural communities on them, as opposed to valuing land merely as property.
Landrace A traditional local variety of a domesticated animal or plant species that has developed largely by natural processes, through adaptation to the physical and cultural environment in a given geographic region. The term TRADITIONAL VARIETY is sometimes applied to plant landraces. [From the sense that it is a race (breed) peculiar to the land where it lives.]
Landscape 1. The visible features of an area of the natural terrestrial world, including physical aspects such as mountains or rivers, and also native vegetation. 2. A specific area having consistent features of this type; e.g., a desert or forest landscape.
Landscape complementation The degree to which a landscape provides all necessary cover types for a species that requires multiple cover types to complete its life cycle.
Landscape connectivity See CONNECTIVITY.
Landscape fragmentation See FRAGMENTATION.
Landscape heterogeneity The variability in land cover types within a given landscape.
Landscape legacy The long-term impact of historical events on ecosystem structure, composition, or function.
Landscape matrix See MATRIX.
Landscape patch See PATCH.
Land sparing The spatial separation of biodiversity conservation from agriculture.
Land use The set of human activities within a landscape and their impacts on natural processes, such as physical, chemical, and biological cycles. Land use is often classified in major categories such as cropland, pasture, plantation forest, and natural forest.
Land-use change The changes in the way land is used at a given location, for example, when a forested land is converted to agricultural fields or when a forestry plantation is changed to other types of use.
La Niña Southern Oscillation See EL NIÑO SOUTHERN OSCILLATION.

Larvae (singular **larva**) Early stages in the development of an organism; for invertebrates, often morphologically quite distinct from the adult.
Larval Having to do with or being larvae.
Larval parasite An organism that is parasitic only at a larval stage.
Late Quaternary The most recent period of the geologic time scale.
Late Quaternary extinction (event) Selective prehistoric extinctions, typically catastrophic, eliminating within the past 40,000 years two-thirds or more of the large land mammals of America, Australia, and Madagascar and at least half the species of land birds on remote islands of the Pacific. Humans were present or suspected to be present in virtually all cases of such extinctions.
Latent parasitism Parasitism occurring without obvious symptoms.
Latitudinal biodiversity gradient The well-documented trend for a greater number of species in tropical ecosystems than in temperate or polar ecosystems.
LCA Life cycle assessment.
Leading edge The dynamic margin of a species' range, where it is expanding into neighboring unoccupied areas.
Leaf area index The single-sided total leaf area of vegetation per unit area of land.
Leaf beetle One of the Chrysomelidae, a large, widely distributed family of beetles that includes several significant agricultural pests.
Leakage A term for the displacement of environmentally harmful activities outside the spatial or temporal boundaries of a policy or project.
Least concern A classification for organisms not qualifying or close to qualifying for any extinction threat category.
Legacy ecosystem An ecosystems that carries an intact "memory" of a change in some physical, chemical, or structural attribute or a change in species composition that modifies the resistance and resilience characteristics of the ecosystem.
LER land equivalent ratio.
Lethal equivalent A group of mutant alleles that would cause an average of one death if homozygous; for example, one lethal equivalent might represent two mutant alleles, each with a 50 percent probability of causing death, or any other combination of mutant alleles that would produce an average of one death.
Levee 1. A structure built along rivers that is intended to contain floodwaters and prevent flooding of land behind the structure. Also referred to as a **dike**. 2. A similar natural feature having the same effect, formed through geomorphic processes.
Liana Any of various woody climbing plants with relatively long, thick stems that use trees as support to gain access to the canopy; they generally grow in mature habitats.
Lichen A composite organism consisting of a fungus (the *mycobiont*) and an alga or a cyanobacteria (the *phycobiont*) living in a symbiotic relationship.
Life cycle assessment (LCA) An analytic framework to assess the full environmental impacts of a given product or service, from raw material extraction, through production, use, and disposal. Also known as **life cycle analysis**.
Life-form The size and stature of a plant under natural life conditions. Environmental constraints can cause the life-form to differ substantially from that of a plant that develops under more favorable conditions, where genotype morphology, as the **growth-form**, finds full expression. For instance, the growth form *tree* may be modified to the life-form *shrub* in the alpine zone.
Light use efficiency (LUE) For crops, the slope of the linear relationship between accumulated biomass and cumulative intercepted light.
Limiting nutrient See NUTRIENT-LIMITED.
Limiting resource 1. The resource that most limits primary production in ecosystems (often nitrogen). 2. Any ecosystem resource that is the one most able to restrict the ability of a certain organism or species to grow, survive, and reproduce in that area; it may be abiotic (e.g., temperature) or biotic (e.g., predators). Also, **limiting factor**.
Limnology The study of the ecological and physical characteristics of lakes and other inland bodies of water. One who conducts such studies is a **limnologist**.
Linnean system or hierarchy The hierarchical arrangement of inclusive categorical ranks in which, apart from the lowest rank, subordinates the members of all lower ranks; it contrasts with exclusive classifications (such as SCALA NATURAE). [From the original version of such systems devised by the Swedish botanist Carl Linnaeus.]
Litter Dead organic matter, such as leaves and branches, that is deposited as the top soil layer.
Littoral 1. having to do with or situated near the shore of a body of water. 2. having to do with or living in the littoral zone (see next).
Littoral zone The shoreline zone of a lake or river, between land and open water, where sunlight penetrates to a wetland's substrate and supports rooted plant growth.
Livelihood investment A non-monetary payment to compensate for opportunity costs of a change in land use; can include alternative livelihoods, alternative food sources, and education investments, among others.
Living dead An individual devoid of the ecological circumstances that allow it to be a reproductive member of its population, but able to live out its own physiological life. Living dead species are most easily observed as large trees remaining on the agroscape, but they are also present in natural ecosystems.
Llanos A region of tropical South America, on the east side of the Andean range of Colombia and Venezuela. Llanos landscapes occupy lower elevations and the landforms vary from flat to rolling terrain. The dominant vegetation is savannah with areas of dry forest, gallery forest, and morichales.
Local adaptation The process by which populations of organisms become adapted to their local environment or ecological niche. Populations that are **locally adapted** generally exhibit the highest fitness in their native environment and reduced fitness in other, contrasting locations or conditions.
Local community A set of individual organisms in a region of space that is typically small, in which local speciation is assumed to have a negligible impact on community dynamics,

and which may be a part of or connected to a much larger metacommunity.

Local extinction The fact of a species ceasing to exist in an area that it had historically occupied.

Local species richness The number of species found at a local site; distinguished from regional or global species richness that aggregates the number of species across a number of individual sites.

Locus (plural **loci**) The physical location of a gene on a chromosome.

Logging The operation of harvesting trees, sawing them into appropriate lengths, and transporting them to a sawmill.

Logistic growth Regulated population growth that follows the logistic equation $dN/dt = rN(1 - N/K)$. Populations growing according to this equation increase rapidly at low densities and the growth rate decreases as they approach carrying capacity (K).

Lognormal distribution A statistical distribution that is normal or Gaussian (a bell curve) when the logarithm of the original data is used. Species abundance data from many communities fit this distribution well.

Lotka–Volterra competition model A model describing the per-capita growth rates of two competing species as linear, declining functions of the abundances of each species. [Developed independently by the mathematicians Albert J. Lotka and Vito Volterra.]

LSU rDNA The gene in DNA that codes for the ribosomal RNA in the large subunit (LSU) of the ribosome. This is 23 S rRNA in prokaryotes and 28 S rRNA in eukaryotes.

LUE light use efficiency.

Lysogenic A bacterium that has a temperate virus integrated into its genome. So called because it may **lyse** (be destroyed or decomposed) at some later time as a result of the virus excising from the genome and initiating replication (induction). This occurs spontaneously with low probability, or with high probability in response to a stress to the host cell.

Lytic virus A virus that does not integrate into the host genome, but initiates replication and ultimately lyses the cell.

Ma An abbreviation of the Latin "mega-annum;" used by paleontologists to indicate dates millions of years before the present time.

Macroalgae See ALGAE.

Macrobiota Those organisms large enough to be seen with the naked eye, usually greater than 1 mm.

Macroclimate A generalized climate; see MICROCLIMATE.

Macrocosm A context or system that reflects one of its components on a large scale.

Macroevolution Evolution leading to the existence of new species.

Macrofauna Invertebrate organisms living in or on sediments and large enough to be retained on a 300- or 500-μm sieve.

Macroinvertebrate An organism lacking a backbone (vertebrae) and generally visible to the unaided human eye.

Macroparasite A parasite large enough to be seen with the naked eye, such as a leech or worm.

Macroplankton A classification of plankton according to size; those plankton of relatively large size (2–20 cm).

Macropod A marsupial in the family Macropodidae, the kangaroos and wallabies.

Macropterous Having long or large wings or fins.

Macroscopic Large enough to be visible to the unaided human eye.

Macrotidal Describing a spring tide range of more than 4 m.

Mafia effect or behavior A term for the fact that hosts have been observed to tolerate the presence of the eggs or nestlings of brood parasites rather than ejecting them, because the parasite can retaliate by destroying the nest entirely. [Likened to the way citizens will endure the existence of the Mafia in their community, because any opposition would result in extreme retaliation.]

Mafic Describing rocks relatively rich in magnesium and iron.

Major ecosystem type Groups of ecoregions that share minimum area requirements for conservation, response characteristics to major disturbance, and similar levels of biodiversity (i.e., the rate of species turnover with distance).

Malacology The scientific study of mollusks. One who conducts such studies is a **malacologist**.

Malesia The Malay Peninsula, Malay archipelago, and New Guinea. Biogeographically, this combines Sundaland, Wallacea (including the Philippines), and New Guinea.

Mammal A member of the order Mammalia, including species that provide milk for their young and have fur or hair. They typically (but not exclusively) bear live young.

Managed relocation Intentional movement by humans of organisms from a place where they currently live to a new region where they are likely to persist in the future, especially as a strategy to deal with human-caused climate change.

Management The efforts of humans to select, plan, organize, and implement programs designed to achieve specified goals. In the context of the natural environment, such activities can range from protective measures to ensure that human influences on natural resources are minimized to greater interventions required to maintain diversity, install facilities, control populations, and eliminate unwanted elements.

Mangal A mangrove tree habitat, or the trees themselves.

Mangrove 1. Any of various trees and large shrubs, especially of the genus Rhizophora, that have intertwining aerial roots and grow densely in salty coastal swamps of tropical and subtropical regions. 2. An ecosystem dominated by mangroves and associated vegetation; a unique type of system that is often the focus of conservation or restoration efforts.

Mantel test A statistical test of the correlation between two matrices that is often used in ecology; e.g., a geographic and genetic distance matrix in landscape genetics. [Named for U.S. statistician Nathan Mantel.]

Mantle In mollusks, an ectodermal tissue layer that secretes calcium carbonate in the form of spicules or shell.

Maquis Vegetation cover of the Mediterranean Basin dominated by evergreen sclerophyllous (leathery-leaved) shrubs and trees.

Marginal value The economic value (**marginal benefit** or **cost**) of the next incremental unit of something. In an ecological context, marginal values are those associated with managing the next small unit of an ecosystem in a particular way (as by preserving, rather than clearing, the next unit of forest). They can also be the partial contribution of natural capital to a final good that is produced with other inputs; e.g., the marginal value of irrigation water for crop production is

the value of the incremental crop yield that can be attributed to irrigation, rather than to labor, fertilizer inputs, and the like.
Marimba Elevated flooring made of wood, used in flooded areas for placing cattle or plants during river flood periods.
Mariculture The commercial farming of aquatic organisms specifically in marine (saltwater) environments.
Marine Having to do with or situated in the ocean, including coastal, intertidal, and near shore settings as well as the open ocean.
Marine protected area (MPA) A clearly defined marine space recognized, dedicated, and managed to achieve the long-term conservation of nature with associated ecosystem services and cultural values.
Marine reserve A subset of a marine protected area (see previous); an area of ocean completely protected from all destructive and extractive activities, such as fishing, mining, or other consumptive use.
Market-based Describing a policy, program, incentive, and so on that is implemented by the encouragement of certain financial incentives rather than by government mandates.
Market-based command and control A policy instrument in which regulated firms must meet individual environmental targets, often requiring the use of specific technologies. Such policies contrast with market-based incentives (see next), which allow firms more flexibility in how to meet aggregate environmental targets.
Market-based incentive A policy designed to change behavior by altering the price paid, or cost borne, by a person, business, or other entity engaged in an activity that affects biodiversity.
Market-based instrument A policy mechanism in which a change in technology, behavior, or the mix of inputs and outputs is encouraged through financial incentives, such as subsidies, taxes, or market creation.
Market economy An organization of economic activity in which private individuals buy and sell goods at prices that (ideally) balance supply and demand.
Market failure In a market economy, a situation in which the allocation of goods and services to consumers does not take place in an efficient way. This can occur when the free market confers benefits on certain participants for acting in a way that is not beneficial to the system as a whole.
Market-mediated effect Changes in the supply and demand of other products due to the production of a given product.
Market value chain The sequence of activities involved in producing raw materials and transforming them into products that can be purchased by final consumers. It can be used as a strategy to create incentives for farmers to conserve landraces with traits that are valued by consumers.
Marsh A wetland dominated by herbaceous species of plants (e.g., cattails, sedges, grasses, floating ferns) and typically occurring on mineral soils. Floating-leaved and submerged species of plants dominate the parts of marshes that grade in to open water.
Marsupial One of the mammals (e.g., kangaroos) without a placenta and with a pouch to carry the young.
Marsupium The folds of skin that envelope the mammary glands and provide protection to infants: found in both metatherian (marsupial) and protherian (monotreme) mammals.
Mass extinction One of the rare events when extinction rates escalate and extinction reaches devastating magnitudes in a relatively short period of geological time. Mass extinction events affect species from a wide variety of taxa and have a global impact. There have been five mass extinction events in Earth's history with magnitudes of 75 percent or more of species loss.
Mass ratio hypothesis A hypothesis stating that ecosystem processes are largely determined by the dominant contributors of the overall plant biomass; i.e., dominant species will exert greater influence on processes than will subordinate species.
Mast fruiting (seeding) The production of many fruits (seeds) by a plant every two or more years, in regional synchrony with other plants of the same species. *Community-wide* mast fruiting occurs with other plants of different species.
Matorral Evergreen sclerophyllous (leathery-leaved) shrublands that dominate large areas of central Chile.
Matrix The most extensive and interconnected mosaic of habitats in a particular landscape; the dominant component of the landscape that exerts a major influence on ecosystem processes.
Maximum sustainable yield The maximum average level of natural resources that can be harvested over time without reducing future potential; e.g., the maximum fish catch that can be removed from a population over a long time period.
MDNA Mitochondrial DNA.
Mechanistic model A model of population dynamics that is explicit about resource consumption, the relationship of consumption with demography, and mortality factors.
Median or medial Along or toward the midline of the body.
Mediterranean-type ecosystem (MTE) A habitat characterized by mild wet winters and warm dry summers. MTEs occur in the Mediterranean region itself, California, central Chile, the Cape Region of South Africa, and southwestern and South Australia, all at northern or southern latitudes of 30–35°.
Megacity A modern city with a large, expanding population, characterized by high consumption levels of energy, water, and food from sources outside the city.
Megadiversity The fact of an area or a country possessing a much larger proportion of species than would be expected by its extent, latitudinal position, and other factors. Such a region is **megadiverse**.
Megafauna Large terrestrial animals that are wolf-sized, deer-sized, and larger. Common standards for inclusion are a body weight of 44 kg (100 lbs) or 100 kg (220 lbs). The term is often used in reference to the many species of large New World mammals that became extinct during the Pleistocene epoch.
Megaplankton A classification of plankton according to size; those plankton of the largest size (2–20 cm).
Meiofauna Animals small enough to pass through a 500-μm sieve but large enough to be retained on a 44- or 63-μm sieve.
Meroplankton See ZOOPLANKTON.
Mesectoderm See NEURAL CREST.
Mesh size The size of the holes in a fishing net.
Mesic Describing a climate with medium moisture supply, neither too dry nor too wet, with sufficient rainfall to support trees and woody plants.
Mesocarp The middle layer of the pericarp; e.g., the flesh of fruits such as plums and cherries.

Mesophile A "middle lover;" an organism (especially a microorganism) that grows optimally at temperatures between 20 and 50 °C. Such an organism is **mesophilic**.
Mesophily The fact of living and growing optimally at moderate temperatures.
Mesophyll (M) tissue Photosynthetic cells that are located between the arrays of vascular bundles and BS (bundle sheath) cells of a leaf.
Mesoplankton A classification of plankton according to size; those plankton of intermediate size (0.2–2 cm).
Mesopredator A predator in the middle of a food web, usually small-to medium-sized animals that are preyed on by other predators and that often benefit when apex predators decline.
Mesotrophic 1. Intermediate in production or species. 2. Specifically, describing an aquatic ecosystem characterized by moderate phytoplankton production and moderate nutrient concentrations and loadings. Compare EUTROPHIC; OLIGOTROPHIC.
Mesotrophy The fact of being intermediate in production, or the existence of species typical of intermediate-production environments.
Meta-analysis A method of summarizing outcomes across numerous studies by compiling measures of the effect of treatments; e.g., measures of the differences in biomass of primary producers in the absence or presence of predators.
Metabolic Having to do with or involved in metabolism (see next).
Metabolism The array of chemical reactions that occur in the cells of living organisms to import, transform, and transport the energy and materials needed to sustain life.
Metacommunity A set of communities; an assemblage of species in a region of space that is typically large and composed of many local communities.
Metagenomes The collective array of genomes of multiple organisms in a natural environment.
Metagenomic Involving or based on metagenomics (see next).
Metagenomics Literally, "beyond genomics;" the process of sampling the entire genetic complement of a natural environment by means of DNA isolation, sequencing, annotation, and informatics analysis. An approach to determining the genetic information of natural microbial populations, and thus of the great majority of microorganisms that cannot easily be cultured.
Metapopulation A group of local populations belonging to the same species that reconnect via regular migration to each other's habitat patches. The state of existence of most species and thought to be the original, natural state of species in general.
Metatherian A member of the infraclass of mammals that includes the marsupials.
Metazoa Those organisms whose individuals have more than one cell (multicelled animals).
Metazoan 1. having to do with or being one of the Metozoa (see previous). 2. Such an organism.
Methane clathrate Methane (CH_4) gas frozen in ice within deep submarine sedimentary pore space or permafrost.
Methanogens Anaerobic archaebacteria that convert $CO_2 + H_2$ or acetate into CH_4. Such organisms are **methanogenic**.
Methyl mercury (MHg) The form of mercury largely produced by sulfate-reducing bacteria that strongly bioaccumulate along food chains.
Microalgae See ALGAE.
Microarray A collection of microscopic DNA spots attached to a solid surface; used in the analysis of gene expression or genetic variation.
Microbe 1. A microorganism; a (usually) single-celled organism that cannot be seen with the naked eye. 2. In popular use, such an organism that causes disease.
Microbial Being, pertaining to, or caused by microbes.
Microbiology The scientific study of microorganisms.
Microclimate 1. The small-scale climate that plants, small animals, and soil microbes experience, and that differs substantially from the macroclimate reported by weather stations. This difference is related to surface warming by the sun or cooling at night, as well as wind shelter effects, and is largely driven by relief, exposure, ground cover, and plant stature. 2. In general use, any localized climate peculiar to a particular area.
Microcosm A small, representative system that has analogies to a larger system. Oceanic islands are often considered microcosms of continental areas because the composition of species on islands, and the processes of formation and decline, are analogous but occur on a much smaller scale.
Microevolution A process of evolution within species, not leading to speciation.
Microfauna Very small, typically microscopic organisms that are described as having animal-like qualities. Many inhabit water films or pore spaces in leaf litter and in the soil.
Microhabitat A small, self-contained environmental unit occupied by a specific subset of interacting species.
Microorganism An organism that can only be observed with a microscope; bacteria, archaea, protists, fungi, and small metazoans, typically smaller than 0.5 mm in equivalent diameter and with mass < 5–10 g. The category may also include viruses, though some classifications exclude these as not independently living.
Microparasite A parasite too small to be seen with the naked eye.
Micropterous Having small or shortened wings or fins.
Microsatellite See REPETITIVE DNA.
Microtidal Describing a spring tide range of less than 2 m.
Mid-domain effect A hypothesis based on random distribution of species ranges within a bounded domain, resulting in maximum species richness in the middle of the domain.
Mid-ocean ridge One of the crustal accretion boundaries between the major tectonic plates that make up the ocean crust.
Milieu souterrain superficiel (MSS) The interstitial spaces deep in the soil and mantle/bedrock interface, as found in glacially fragmented zones. Also, **mesovoid shallow stratum**.
Mimicry 1. A defensive mechanism in which organisms of two species resemble each other, ordinarily because at least one of them is distasteful or harmful and thus unattractive to predators. 2. Evolution by natural selection in which a

character is favored because it closely resembles one found in another species.

Mineralization The availability of inorganic nutrients in the microbial decomposition of organic detritus, occurring after the IMMOBILIZATION phase in which microbial demand for nutrient elements is satisfied.

Minimum viable population (MVP) The minimum size that a population must be to persist over a given time span.

Miombo A woodland habitat type widespread in south-central Africa, characterized by numerous species of the tree genera *Brachystegia* and *Isoberlinia*, which form nearly closed canopies. Fire is an annual event in this habitat, which supports relatively low populations of large mammals.

Mistletoe Woody parasitic plants of the closely related families Loranthaceae or Viscaceae with haustoria that originate as root tissue and penetrate stems or roots (in Loranthaceae) of the host plants. Also capable of photosynthesis.

Mitigation 1. Measures taken to reduce environmental impacts to the point where they will have no adverse effects. 2. Specifically, actions designed to reduce the amount that the climate will change. These actions generally involve reducing greenhouse gas emissions and sequestering carbon dioxide.

Mitigation wetland A wetland created, restored, or enhanced to replace lost wetlands.

Mitochondria (singular **mitochondrion**) Eukaryotic cellular organelles used for cellular respiration and energy production.

Mitochondrial DNA (**mtDNA** or **mDNA**) A small, maternally inherited molecule of DNA found only in the mitochondria (see previous).

Mixed intercropping See INTERCROPPING.

Mixing ratio The ratio of the concentration of a given gas to the concentration of air. It can be defined in terms of volume or mass.

Mixotrophic Feeding in both an autotrophic and a heterotrophic manner. Such an organism is a **mixotroph**.

Mixotrophy A form of nutrition involving both autotrophic (photosynthetic) and heterotrophic carbon acquisition.

Mobile element See TRANSPOSON.

Model A construct, either physical, mathematical, or computational, that is a simplification of reality developed to meet certain objectives.

Moisture deficit The difference between the existing water content of soil and the water content of the same soil at field capacity, i.e., after saturation followed by drainage.

Molecular clock A hypothesis that mutation rates and substitution rates do not vary among lineages in a tree, so that the timing of divergences can be dated from known times of divergence in the fossil record. Because rates of evolution along a lineage can vary, two different kinds of clocks (a *fixed* or a *relaxed* clock) can be used, depending on whether the rate of evolution is allowed to vary.

Molecular phylogenetics The study of the history of evolutionary relationships according to similarities and differences among (usually extant) organisms in molecular characters (such as DNA).

Molecular phylogeny A phylogeny constructed from comparisons of homologous genetic material.

Mollusk One of the Mollusca, a large phylum of invertebrates including among others the cephalopods (e.g., squid, octopus), bivalves (e.g., clams, oysters, mussels, scallops) and gastropods (snails and slugs). Also spelled **mollusc**.

Monitoring The repeated collection and analysis of observations and measurements to evaluate changes in populations of species and environmental conditions.

Monoculture 1. The agricultural practice of growing one single crop on a given field over a sequence of growing seasons. 2. The presence of a single species, or single genotype, of crop plants or other organisms.

Monoecious Having male and female flowers on the same plant.

Monogamy The reproductive pattern in which one male mates with one female exclusively. Such a pattern is **monogamous**.

Monograph effect The argument that the apparent existence of intensive evolutionary activity in a given interval actually reflects the fact that researchers have thoroughly studied a particular section of the fossil record which reveals this activity, while other such investigations elsewhere might just as well have produced the same number of previously unknown species. [From the idea that such a study produces various monographs on the subject.]

Monophyletic Describing a group in which all members are descended from a single common ancestor and all descendants of the ancestor are classified in the group. Compare PARAPHYLETIC; POLYPHYLETIC.

Monophyletic group A true, historical, evolutionary lineage consisting of an ancestor and all of its descendants; defined by shared, derived characters.

Monophyly The fact of being an evolutionary unit that includes a most recent common ancestor and all of its descendants.

Monotreme A mammal in the family Monotremata, which includes spiny anteaters and the platypus.

Monsoon (system) A wind system covering the eastern half of Asia in which winter cooling of the large landmass produces strong, stable high pressure, with clear skies and outward flow of cold, dry air, and summer warming produces low pressure, drawing wet air masses, with clouds and rain, inward from the adjacent oceans.

Montreal Protocol A treaty signed by a number of nations to curtail the production of CFCs (Chlorofluorocarbons), which are known to deplete the Earth's ozone layer.

Moor An expanse of open vegetation, dominated especially by dwarf Ericaceae and *Sphagnum* mosses, on peaty soil.

Morichal A plant community characteristic of tropical savannas; it is seasonally flooded and the presence of the palm *Mauritia flexuosa* is conspicuous.

Morphodynamics Interactions in a beach and surf zone environment between the physical structure and the water and sediment movement.

Morphological Involving or based on the structure and form of an organism.

Morphology The scientific study of the structure and form of organisms, including both outward appearance and internal features.

Morphometrics A statistical procedure by which morphological traits are quantified and analyzed.

Morphospecies A species defined on the basis of morphology alone. This criterion has been traditionally used with many eukaryotic organisms, when information about interbreeding or molecular composition was not available. In microorganisms this approach definitely underestimates real diversity.
Mosaic A spatial patch distribution of vegetation communities.
MPA Marine protected area.
MPA network A collection of individual MPAs (marine protected areas) operating cooperatively and synergistically to fulfill ecological aims more effectively and comprehensively than individual sites could alone.
MSS Milieu souterrain superficiel.
MtDNA Mitochondrial DNA.
MTE Mediterranean-type ecosystem.
Mulch A layer or covering of organic materials such as leaves, shredded bark, dried grass, compost or the like, spread around plants to enrich the soil, help retain moisture, protect against temperature extremes, and so on.
Müllerian mimicry A form of mimicry in which two or more unpalatable species resemble each other, with the effect that predators are more likely to avoid any species with this appearance. [Described by the German biologist Fritz Müller.]
Multidimensional niche A tempero-spatial region defined by meeting a set of different requirements for the viability of a particular kind of organism.
Multifunctional cropping An agricultural system that incorporates different combinations of herbaceous crops, trees, and animals within a single farm unit.
Multimodel A single integrated model that links together models utilizing different mathematical or computational approaches.
Multiple resistance The occurrence of more than one resistance mechanism in the same individual or pest population.
Multivoltine See PLURIVOLTINE.
Mushroom A large, fleshy fruiting body consisting of a cap with gills or pores on the undersurface (or sometimes flat or with teeth or folds), usually on a stalk, and producing sexual spores; used especially for members of the Basidiomycota. A poisonous mushroom is popularly called a **toadstool**.
Musth The annual reproductive period in adult male elephants, characterized by extreme aggressiveness and secretions from temporal glands.
Mutant An organism, individual, or genetic character resulting from an instance of mutation (see next).
Mutation A heritable change in the nucleotide sequence of an organism.
Mutation rate The frequency with which new mutations arise per generation.
Mutualistic A mutually beneficial association or interaction, temporary or ongoing, between or among organisms of different species.
Mutualism Describing an interaction between organisms of different species that provides survival and/or reproductive benefits to both of them.
MVP Minimum viable population.
Mycelium (plural **mycelia**) The vegetative part of a fungus, made up of a mass of branching threadlike filaments (hyphae).
Mycobiont See LICHEN.
Mycoparasite A fungus whose host is another parasite.
Mycorrhiza (plural **Mycorrhizae**) A root tip that is infected with fungi in a mutually beneficial partnership. The fungal hyphae explore large volumes of bulk soil, absorbing nutrients and transferring them to the plant; the plant supplies the fungus with the organic carbon necessary for growth and energy production. The two basic types are arbuscular mycorrhizae (endomycorrhizae), in which fungal hyphae penetrate into the plant root cells; and ectomycorrhizae, in which fungal hyphae grow between plant root cells and form a thick sheath over the root tip, but do not penetrate.
Mycotroph A plant that obtains energy through mycorrhizal associations with saprophytic or other biotrophic fungi. Such an association is **mycotrophic**.
Myrmecophily The ability to form symbiotic associations with ants. Such an organism is **myrmecophilous**.
Mysticete One of the Mysticeti, the suborder of cetaceans that includes the four families of baleen whales. Most mysticetes are quite large, including the blue whale which reaches 33 m in length and is recognized as the largest animal in the Earth's history. About 10 percent of the known modern species of marine mammals are mysticetes.
Myth A story, usually of unspecified origin and at least partially traditional, that ostensibly relates actual events to explain some practice, belief, institution, or natural phenomenon; myths provide models for human behavior and often emphasize behaviors that support conservation of biodiversity.
N_2 fixation See nitrogen fixation.
Nanoplankton A classification of plankton according to size; those plankton of very small size (2–20 mm).
Native Occurring naturally in a given area; indigenous.
Native habitat 1. The community of species and the physical environment that occurred in a region before human intervention, especially the naturally occurring vegetation. 2. The particular natural habitat of a given species.
Native species A species that evolved and is naturally distributed in a given region.
Natural capital The capacity of the natural world to provide services (ecosystem services) that contribute to human well-being. These may include provisioning services, such as food and fiber production; regulating services, such as water flow and quality, or temperature regulation of the atmosphere; cultural services, such as recreation or spiritual needs; and supporting services, such as soil formation and nutrient cycling. Fossil fuels and minerals constitute non-living natural capital, which is generally not renewable on time scales of value to society (and is relatively unaffected by changes in ecosystem status).
Natural disturbance Any relatively discrete event in time that is not directly caused by humans and that disrupts ecosystem, community, or population structure and changes resources, substrate availability, or the physical environment.
Natural enemy Any species that consumes or parasitizes another species, including predators, herbivores, parasites, pathogens, and parasitoids.

Natural experiment A change in a system, due to either natural or anthropogenic forces, that provides the basis for comparison with conditions before the change and with conditions in another setting where the change did not occur. For example, the recovery of a predator may provide a natural experiment in which conditions before and after recovery and conditions inside and outside the recovery area are compared.
Natural forest An area of land that supports a minimum of 10 percent tree cover that has arisen as a result of natural processes of establishment and succession.
Natural insurance value See INSURANCE HYPOTHESIS.
Naturalized Describing alien plants that reproduce consistently and sustain populations over many life cycles without the input of resources or direct intervention by humans; they often recruit offspring freely, but mainly near adult plants, and do not necessarily invade natural ecosystems.
Natural product A compound, organic in nature (e.g., composed of carbon, hydrogen, oxygen and nitrogen), that is produced by a living organism for a purpose other than basic life functions, and that provides some adaptive value to the producing organism.
Natural regeneration The process of replacement by natural seedlings in gaps created by the selective cutting of marketable timber.
Natural resources Components of the natural environment that have utility or value to humans. *Renewable* resources either can be replenished within a relatively short time (assuming sustainable use) such as timber or fish stocks, or they have an ongoing existence (e.g., wind energy). *Nonrenewable* resources exist only in a finite quantity (e.g., gold) or could only be replenished over geologic time spans (e.g., oil).
Natural resource(s) management The management of natural systems to maintain or provide yields of wood, water, and other resources for human use.
Natural selection The variation among individuals in their ability to survive and contribute offspring to the next generation. To the extent that the variation in offspring number is associated with genetic differences among individuals in the population, natural selection results in change in the population's genetic composition and tends to increase its average fitness.
Nature 1. The totality of the external physical world, including the plants and animals that inhabit it and the processes that occur in it, as opposed to human existence and activities. 2. The sense of this external world as an agent governing physical processes, as in "*Nature* abhors a vacuum."
Nearctic One of the major zoogeographical regions of the world. See ZOOGEOGRAPHICAL REGION.
Nearly neutral Describing a mutation whose population genetic dynamics are influenced by both genetic drift and selection.
Near-threatened Not currently meeting any of the criteria for an extinction threat category, but close to qualifying or likely to qualify soon.
Nekton Actively swimming pelagic organisms.
Neogene The Miocene and Pliocene Epochs, accorded the status of a period when the Tertiary is considered an era.
Neotenic A secondary reproductive form with juvenile morphological characters. Neotenics derive from larvae, nymphs, pseudergates, or workers through at least one special molt.
Neotropical One of the major zoogeographical regions of the world, ranging from southern Mexico through Central and South America, including the Caribbean and the tip of Florida. Also called the NEOTROPIC or NEW WORLD TROPICS. See ZOOGEOGRAPHICAL REGION.
Neritic Having to do with or living in an area of the ocean with relatively shallow waters, from the low tide mark to a depth of 200 m. This area is the **neritic zone**.
Net immobilization See IMMOBILIZATION.
Net mineralization See MINERALIZATION.
Net primary production The rate of production of biomass by plants and other primary producers such as algae; this forms the basis for all food webs. Also called **net primary productivity**.
Neural crest An embryonic tissue in vertebrates intermediate between the neurectoderm and ectoderm, with cells migrating widely to their final destination. This tissue gives rise to anterior skeletal elements, many portions of the future head and pharynx, and all pigment cells. Also, MESECTODERM.
Neurectoderm An embryonic tissue that gives rise to the central tube of the nervous system.
Neutral mutation A mutation whose effect on fitness is either non-existent or so small that its fate is controlled by genetic drift rather than by selection.
Neutral theory The principle that the majority of nucleotide substitutions result from the random fixation of neutral or nearly neutral mutations.
New production Primary production in the ocean maintained by assimilation of new sources of nitrogen (N) to the euphotic zone, such as nitrate (NO_3^-) or N_2. Compare REGENERATED PRODUCTION.
New World Tropics see NEOTROPICAL.
NGO Nongovernmental organization.
Niche The unique position of a particular species in an ecosystem in relation to other species and with respect to its resource use, habitat requirement, and other functions within the larger community.
Niche complementarity The differential acquisition and use of resources by different species; niches of different species differ which results in more complete utilization of resources in spatially or temporally complex environments.
Niche differentiation Differences in the morphology, physiology, or behavior of species that can influence their abundances, dynamics, and interactions with other species, including the ability of various competing species to coexist.
Niche dimensions Ranges of values of environmental measurements in an ecological niche. A range of temperatures, salinities, or oxygen concentrations may be niche dimensions for a population of fish.
Niche overlap The proportion of available resources that are shared by two species, especially in the context of a single resource that limits population growth.
Nitrification The step in the nitrogen cycle in which ammonia is oxidized and converted to nitrite and then oxidized again to form nitrate; a biogeochemical process mediated by soil microbes.
Nitrogenase A metalloenzyme complex that catalyzes the process of nitrogen fixation (see below).

Nitrogen deposition The process by which reactive forms of nitrogen are deposited to the Earth's surface through either wet or dry deposition.
Nitrogen fixation The reduction of atmospheric dinitrogen (N_2) to ammonia (NH_3) or other organic or inorganic compounds by bacteria, lightning, or photochemistry.
Nitrogen saturation Nitrogen supply in excess of the capacity of the ecosystem to retain it.
Nitrophytic Describing species of plants that either tolerate or require high levels of nutrients, especially nitrogen. Such an organism is a **nitrophyte**.
Node A branching point on a cladogram corresponding to a hypothetical common ancestor.
Non-avian dinosaurs The Dinosauria excluding modern birds (Aves); i.e., the extinct forms popularly considered to be the only dinosaurs.
Nonexcludable See PUBLIC GOOD.
Nongovernmental organization (NGO) An organization that is not an agency or part of a government, typically one that is not-for-profit and not owned by individuals.
Nonindigenous 1. Not native; not originating in the site that it now occupies. 2. Specifically, describing crop plants that did not originate in the range where they are now cultivated. Also, **nonnative**.
Nonpoint source (NPS) A diffuse source of pollution not attributable to any single specific discharge (e.g., agricultural runoff, urban runoff, or atmospheric deposition). Contrasted with POINT SOURCE.
Non-renewable Describing a resource that cannot be grown or produced at a level which over time will sustain its present consumption rate; for example, because it exists only in a finite quantity (e.g., gold) or it could only be replenished over geologic time spans (e.g., oil). Contrasted with RENEWABLE.
Non-renewable energy Those forms of energy whose transformation consumes the energy source. The major ones are the fossil fuels (coal, oil, and natural gas) and nuclear power (because it requires the mineral uranium). Compare RENEWABLE ENERGY.
Nonrival See PUBLIC GOOD.
Non-structural flood management Flood-risk management approaches that do not depend on controlling water with engineered structures. Examples include land-use planning, building codes, insurance, and use of the natural and beneficial functions of floodplains, including storage and conveyance of flood waters.
Nonsynonymous mutation A nucleotide change in a protein coding region that changes the encoded amino acid.
Non-use ecosystem services Benefits that are derived from ecosystems without actually using them, such as the benefit derived simply from knowing a certain ecosystem exists (existence value).
Nonuse values Values ascribed to the natural environment that are independent of any human interaction with it; for example, an individual valuing the pure existence of a natural habitat or ecosystem or wanting to "bequeath" it to future generations. Compare DIRECT USE VALUES; INDIRECT USE VALUES.
North Atlantic Oscillation (NAO) The difference in atmospheric pressure between Ponta Delgada, Azores, and Stykkisholmur, Iceland. The variations in the NAO are usually associated with a positive or negative phase related to changes in the direction and strength of the westerly wind as well as in sea surface temperature. During a positive phase, winters over Scandinavia are warmer and the opposite for a negative phase. Recently, the fluctuations observed in the NAO have been related to changes in ecosystem processes.
No-take area A classification for a marine area that is permanently or temporarily completely closed to any form of extraction, and where no disturbance of any kind is allowed.
Nothofagus A genus of Southern Hemisphere trees with a classic Gondwanic (see) distribution in southern South America, southeastern Australia, New Zealand, New Caledonia, and New Guinea.
Notochord A stiff, flexible, longitudinal rod running along the mid-dorsal portion of the chordate body. It is situated dorsal to the coelom and ventral to the central tube of the nervous system.
NPS Nonpoint source.
Nuchal organ Any of various sensory organs with nerves coming from the posterior part of the brain or from the nerve ring, usually present as ciliated pits or short ridges at the posterior end of the prostomium. Found in most polychaete worms and also present in the sipunculid worms.
Nucleotide A subunit of DNA and RNA, composed of a ringed five-carbon sugar, a ringed nitrogen-rich base, and phosphates. Nucleotides are often referred to as base pairs because individual types form complementary hydrogen bonds (G-C and A-T in DNA, G-C, and A-U in RNA) to make double-stranded DNA and RNA molecules.
Nucleus 1. The compartment of a cell in which DNA is stored on chromosomes; it functions as the central control unit of the cell. 2. The dense region at the center of an atom, consisting of protons and neutrons. **3**. Any crucial central part around which other parts are grouped.
NUE Nitrogen-use efficiency; a measure of photosynthesis or growth relative to the nitrogen content of the leaves or plant.
Nunatak An isolated ice-free summit of a mountain whose bulk is permanently below the surface level of an ice sheet. [From the Inuit term for this feature.]
Nurse plant A plant that shelters and facilitates the growth of others.
Nutrient Any of the various inorganic chemicals (particularly nitrogen, phosphorus, and silicon) required for the survival and growth of animals and plants, including crops and phytoplankton.
Nutrient cycle The repeated pathway of mineral elements, such as carbon, nitrogen, phosphorus, and water, from the environment through organisms and back into the environment.
Nutrient-enriched Having elevated concentrations of nitrogen and phosphorus nutrients, especially in an aquatic ecosystem. This condition is **nutrient enrichment**.
Nutrient-limited Having growth limited by the relative amount of a certain nutrient, or by the proportion of two nutrients to each other. This condition is **nutrient limitation**.
Nutrient-saturated Having a nutrient or nutrients available well in excess of the amount required for growth. This condition is **nutrient saturation**.
Nutrient stoichiometry 1. The study of changes in the relative proportions of critical nutrients available in the environment, relative to differences in the allocation of these

elements in organisms. 2. The study of the relative availability of critical elements in ecosystems, as both an influence on and an overall result of interactions among biota and system biogeochemistry.

Nutrient uptake capacity The instantaneous rate of nutrient acquisition, usually measured in brief (1–2 h) incubations. Uptake capacity reflects the abundance of transport sites on the root cell membranes and their affinity for nutrient ions.

Nutritionalism The view that bottom-up forces (i.e., nutrients) regulate populations.

Nymph An immature individual that is on the developmental pathway to the imago and that possesses external wing buds.

Obligate Obliged or required; able to function in only one way; e.g., a parasite that cannot live without a host; bacteria that can only live in an aerobic environment. Also, **obligatory**. Compare FACULTATIVE.

Obligate carnivore A carnivorous animal that eats only meat and not non-animal food.

Obligate epiphyte A plant that always grows on another plant for structural support, but derives no nutrients from the host.

Ocean acidification Ongoing changes in carbonate chemistry of the Earth's oceans (e.g., decrease in pH and decrease in carbonate ions), caused by their uptake of anthropogenic carbon dioxide from the atmosphere.

Ocean observatory A collection of airborne and in-water sensors for characterizing oceanic chemical, physical, and biological parameters that may vary over time.

Odontoceti The suborder of the cetaceans that includes the ten families of toothed whales, including sperm whales, beaked whales, dolphins, porpoises, belugas, narwhals, and river dolphins. Approximately half of the known modern species of marine mammals are odontocetes.

ODP Ozone depletion potential.

Offset activities The set of activities identified to achieve no net loss or a net gain of biodiversity in the specific context of the development project concerned.

Old-growth (forest) A forest undergoing a late stage of succession that is relatively stable with regard to forest structure, ecosystem function, and species composition.

Old World Tropics See PALEOTROPICS.

Oligonucleotide A short, single-stranded nucleic acid molecule either obtained from an organism or synthesized chemically.

Oligotrophic Describing an ecosystem or water mass that is low in nutrients and generally characterized by low-standing stocks of living biomass. Compare EUTROPHIC; MESOTROPHIC.

Oligotrophy The fact of being low in production, or the existence of species typical of low-production environments.

Ombrotrophic Describing ecosystems or plant species only receiving nutrients in precipitation. This condition is **ombrotrophy**.

Omnivorous Describing a consumer that feeds on various food sources; e.g., a predator with a diet of plants and other animals as a primary food source. Such an animal is an **omnivore**.

Omnivory The fact of feeding at more than one trophic level, as when a predator consumes not only prey species but also herbivores and other predators.

Ontogenic Having to do with or being part of ontogeny (see next).

Ontogeny The sequence of embryonic and postembryonic changes in the characters of an organism in the course of its life. Also, **ontogenesis**.

Open access resource 1. A resource that is nonexcludable; i.e., it can be exploited by any person, regardless of the consequences of that exploitation for other would-be users. 2. A resource for which property rights are not recognized or clearly defined.

Open community A community that receives immigrants from external sources and may export emigrants as well.

Operational taxonomic unit (OTU) A group of organisms regarded as being distinct from other groups, based on any clearly defined set of variables.

Opportunistic species A species typical of transient, unstable, unpredictable, frequently disturbed, or periodically extreme environments, usually having strong dispersal abilities and faster growth, smaller size, and shorter life spans than potential competitors.

Opportunity costs Costs of foregone opportunities; i.e., a measure of what could have been gained by the next-best use of a resource had it not been put to the current use. In terrestrial protected areas where extractive uses are forbidden, the opportunity cost represents the highest-value extractive use for that land. It is important to include opportunity costs to track the full set of consequences of conservation planning.

Option value The potential benefits of a resource not derived from actual use, expressing a preference for conservation of the resource against some probability that it can be used at a later date.

Organelles Membrane-bound structures within eukaryotic cells that carry out specific functions and often contain their own DNA; typical organelles are mitochondria and chloroplasts.

Organic 1. Relating to or coming from living beings as opposed to inanimate matter. 2. Describing farming practices (**organic agriculture** or **farming**) that rely on biological inputs and diversity for soil fertility and pest and disease suppression; **organic farms** that are certified by government policies or private organizations are not allowed to use synthetic pesticides or fertilizers.

Organismal Having to do with living or formerly living beings (organisms).

Organismal diversity The number and relative abundance of all species of organisms living in a given area.

Ornithischian One of the Ornithischia, the so-called bird-hipped dinosaurs such as horned and duck-billed dinosaurs.

Orobiome A mountain range characterized by a particular climatic pattern and a characteristic sequence of vegetation zones.

Orogenesis The process of formation of mountains.

Orographic clouds Clouds that are formed as moist air rises over mountains.

Orthophyllous Pollinated by birds.

OTU Operational taxonomic unit.

Outbreeding A mating system in which individuals breed with other individuals that are not closely related to them.

Outgroup A species or other taxonomic group chosen to establish the root of any phylogeny.

Outpost An individual, isolated tree or small group of trees appearing above the tree line.
Output environment A collective term for all energy and materials moving out of a given system.
Overcompensatory growth In grazed plants, the reaction to tissue removal by enhanced primary production compared to that in undefoliated controls.
Overexploitation The harvesting of a natural resource, such as fish or timber, at a rate more rapidly than it can be naturally replenished; the opposite of sustainable use.
Overfishing The process of extracting marine organisms faster than they can reproduce.
Overgraze 1. To carry on a process of overgrazing; consume forage at a level that is not sustainable. 2. The fact or process of doing this.
Overgrazing The consumption of forage to the extent that declines in productivity or desirable species composition are probable (i.e., degradation). It represents excessive use in terms of the amount, duration, and frequency of plant utilization.
Oviparous 1. Describing reproduction in which the eggs hatch outside the mother's body. 2. Describing an organism with this mode of reproduction; e.g., birds. The fact of reproducing in this way is OVIVIPARITY. Compare OVOVIVIPAROUS; VIVIPAROUS.
Oviposition The fact or process of laying eggs.
Ovipositor A modified appendage in some animals used for egg-laying (e.g., wasps) and also for other purposes in some cases.
Ovoviviparous 1. Describing reproduction in which the embryos develop within eggs retained by the mother within her body until they hatch. 2. Describing an organism with this mode of reproduction; e.g., some fish and reptiles. The fact of reproducing in this way is OVOVIVIPARITY. Compare OVIPAROUS; VIVIPAROUS.
Oxic 1. Having oxygen. 2. Specifically, describing an environment or habitat in which oxygen is present.
Oxidizing capacity A self-cleansing ability of the atmosphere through oxidative reactions.
Oxygen isotopic composition The ratio of common isotopes of oxygen (16O and 18O) analyzed in a mass spectrometer and reported in delta notation (d18O) relative to a standard; reported in parts per thousand and used as a proxy for paleo temperature and global ice volume.
Ozone A form of oxygen in which three atoms of oxygen occur together, extremely important to life on Earth because of its absorption of solar ultraviolet radiation in the stratosphere's **ozone layer**. Ozone is also a greenhouse gas.
Ozone depletion A decline in the total volume of the Earth's ozone layer, observed since the 1970 s and attributed to the presence there of certain industrial chemicals (e.g., CFCs) that are now generally banned.
Ozone depletion potential (ODP) The relative amount of decline in the total volume of the ozone layer that a given substance can potentially cause, based on a scale in which trichlorofluoromethane (a well-known depleting agent) has a value of 1.0.
Pacific Decadal Oscillation (PDO) A long-term El Niño-like pattern of Pacific climate variability. The climate oscillations ENSO (El Niño Southern Oscillation) and PDO have similar spatial climate fingerprints, but very different behavior in time; e.g., 20th century PDO events persisted for 20 to 30 years, while typical ENSO events persisted for 6 to 18 months.
Pacific salmon Any of several species in the genus *Oncorhynchus* that always die after spawning (and hence are semelparous). Five Pacific salmon species occur in North America: Chinook, coho, sockeye, chum, and pink. Western trout in the genus *Oncrorhynchus* are closely related to Pacific salmon; however, like the Atlantic salmon (*Salmo salar*), these species can spawn in more than one year.
Paedomorphosis Reproduction while retaining at least some larval characteristics.
Palearctic One of the major zoogeographical regions of the world. See ZOOGEOGRAPHICAL REGION.
Paleo- A prefix meaning very old; ancient; used to refer to ancient time periods.
Paleoclimatology The study of ancient climates, on the basis of evidence from fossils and other traces left in the geological record.
Paleotropics Tropical regions of the Eastern Hemisphere extending from Africa through south and southeast Asia into the southwest Pacific. Also known as the Old World Tropics.
Palynology The study of pollen and spores.
Palynomorphs Pollen and spores produced by plants.
Pampa(s) A region of temperate South America, west of the Andes in Argentina. It contains several ecosystem types, including temperate grasslands, dry forest, and xeric shrublands.
Panbiogeography The examination of biogeographical distributions and relationships on a worldwide scale.
Pantanal The largest wetland of South America, drained by the Paraguay River and its tributaries.
Páramo 1. An ecosystem of the high Neotropical mountains, found in South America above the tree line in the northern Andean mountains of Colombia, Venezuela, Ecuador, Peru, and Bolivia. The dominant vegetation is similar to an alpine meadow or grassland; cacti and giant-rosette plants are conspicuous. 2. Any such cold, humid landscape located above the tree line.
Parapatric Describing taxonomic entities whose geographic ranges are distinct but adjacent to each other.
Parapatry The relationship of organisms whose ranges do not significantly overlap but are adjacent to each other.
Paraphyletic Describing a group that includes a common ancestor and some, but not all, of its descendants. Compare MONOPHYLETIC; POLYPHYLETIC.
Paraphyletic group A group of taxa that includes some, but not all, of the descendants of the most recent common ancestor for the group.
Paraphyly The fact of being an evolutionary unit that includes a most recent common ancestor and some, but not all, all of its descendants.
Parapodia (singular **parapodium**) Segmentally arranged outpockets from the body wall of certain invertebrates, usually divided into two branches (rami), which are the dorsal notopodium and the ventral neuropodium. Present in most polychaete species, but extremely variably developed and sometimes entirely absent.
Parasitic Being, having to do with, or caused by a parasite.

Parasite An organism that obtains its nutrition from a living host (and usually lives inside or on it), and in so doing harms the host.
Parasitism A close association of two organisms in which one, the parasite, depends on the other (the host), deriving some benefit from it and damaging it.
Parasitoid An organism that spends a significant portion of its life history attached to or within a single host organism; unlike a true parasite, it ultimately sterilizes or kills the host.
Parataxonomist A lay person trained to collect and carry out preliminary identification of botanical and zoological specimens that will be used for taxonomic purposes.
Pareto efficiency or optimality An economic equilibrium in which any change to the system will decrease the economic well-being of at least one participant; such a system is **Pareto efficient**. If a change can be made to an economic system that will benefit at least one individual without harming anyone else, that is a **Pareto improvement**. [A concept introduced by the Italian economist Vilfredo Pareto.] See also KALDOR–HICKS EFFICIENCY.
Pareto principle The so-called 80–20 rule; the idea that for many events and conditions, 20 percent of the agents produce 80 per cent of the effects or results. Examples might be 20 percent of a given area containing 80 percent of its total biodiversity, or 20 percent of the residences in a community occupying 80 percent of the total land cover. [Based on Vilfredo Pareto's observation that 80 percent of the land in Italy was owned by 20 percent of the people.]
Parr Older juvenile salmon that are distinguished by **parr marks**; i.e., prominent oval spots on their sides.
Parsimony The general scientific principle that given alternative explanations or hypotheses for a set of observations or data, the most corroborated is the **parsimonious** one; i.e., the one that requires the fewest ancillary or additional hypotheses.
Parthenogenesis Reproduction in which eggs are not fertilized by males.
Passive crown fire A term for a wide range of crown fire behavior, from torching of individual crowns to burning of dense patches of vegetation.
Passive recovery The process of allowing a system to recover on its own without intervention beyond removing the disturbance. Similarly, **passive restoration**.
Pastoral(ist) Having to do with or practicing pastoralism.
Pastoralism 1. The agricultural practice of raising livestock. 2. A traditional, often nomadic economic and cultural system based on the raising of livestock. **3**. Traditional agricultural life and work in general, or a literary or artistic description of this.
Pasture Lands in which the native grasses and other vegetation have been removed in favor of cultivated grasses or other forage plants to be used by grazing domestic animals.
Patagium The skin that stretches across the arms and legs of flying and gliding mammals.
Patch A relatively homogenous area within a larger landscape that differs markedly from its surroundings.
Patchiness Variation from place to place (or time to time) in the abundances of animals or plants, caused by the interaction of numerous processes.
Pathogen The causative agent of an infectious disease, typically a virus, bacterium, protozoan, fungus, or multicellular parasite. Such an agent is **pathogenic**.
Payments for ecosystem services (PES) Financial rewards for individuals who protect habitats so that the provision of natural goods and services can continue; payments may be made by private parties that benefit from the delivery of those ecosystem goods and services, or by a governmental or other agency. **Payments for watershed services** is a subset of PES that involves a similar financial transaction, but with the focus on watershed-based hydrologic services.
PCR Polymerase chain reaction.
Peat 1. A type of soil formed from the accumulation of the remains of dead plants, with high organic matter content (minimally 20% and as high as 90% by dry weight) and depth greater than 30 to 40 cm. Peat forms in some wetlands because anoxic conditions, low pH, low temperatures, and short growing seasons, or plant tissues highly resistant to decay, prevent rates of decomposition from keeping up with rates of organic matter production. 2. Peat fuel (see next).
Peat fuel Dried peat biomass used as a fuel; a traditional fuel source in various northern regions; e.g., Ireland.
Peatland A wetland in which dead organic matter accumulates to form peat more than 30 to 40 cm deep; this typically includes bogs, poor fens, and rich fens but some swamps also occur on peat.
Pedipalps The second pair of pre-oral appendages in spiders and other arachnids. They are multi-segmented and primitively leg-like. They may be raptorial (grasping) or sensory (like antennae) or used as walking legs.
Pelagic 1. Located in or relating to the water column in aquatic environments. 2. The open-water part of the aquatic environment that is far from the shore and the bottom bed.
Pelagic organism A free-swimming (nekton) or floating (plankton) organism that lives exclusively in the water column.
Pelagic zone The open-water portion of the ocean, a lake, or another body of water, beyond the littoral zone.
Penetrometer A device used to test the compressive strength of soil, based on the energy required to push a probe of defined size and shape into the soil.
PEP carboxylase Phosphoenolpyruvate carboxylase; the initial carboxylation enzyme in C_4 photosynthesis.
Peptide A compound composed of several amino acids that are linked together by amide bonds.
Perennating structures The tissue that carries a plant into a new growing season; e.g., buds of trees, shrubs, and so on; bulbs of some plants such as tulips; or seeds of annuals.
Perfect competition The ideal form of a market economy, in which all goods are owned by perfectly rational, perfectly informed individuals, all goods can be bought and sold in markets, and no single participant can determine the price at which transactions take place. Such an economy is **perfectly competitive** and is said to have PARETO EFFICIENCY (see).
Periodic parasite A parasite that visits a host at intervals.
Periphyton Bacteria, fungi, algae, and sessile microfauna growing attached to substrata (sediments, rock, plants, animals, sand).
Peristomium In annelid worms, a region around the mouth of the larvae; if distinct in adults it may carry peristomial cirri.

Permafrost The phenomenon of ground materials (e.g., soil and bedrock) that are at or below 0–1 °C for two or more years. Usually, but not always, any water present is in the solid state. The Arctic is characterized by the presence of large, continuous areas of permafrost that have the form of lower soil layers permanently frozen with a shallow active layer that freezes and thaws each year and accommodates below-ground biological activity.
Permanence A property of communities that ensures the long-term coexistence of species because community trajectories (variation in numbers through time) have no species approaching very low numbers.
Permanent parasite A parasite associated with a host for long periods.
Persistence The sustained existence of species or other elements of biodiversity.
Perturbation A temporary change in one or more dynamical variables or parameters of an ecosystem or other dynamical system, due to external factors.
PES Payments for ecosystem services.
Pesticide A chemical substance used for killing, preventing, controlling, or mitigating a pest organism.
Pesticide resistance The inherited ability of a pest to survive and reproduce following exposure to a dose of pesticide normally expected to be lethal to individuals of this species.
Petaloid Describing flowers with petals present, and typically showy.
pH The free hydrogen ion concentration of an aqueous solution, measured on a logarithmic scale and used to express the acidity or alkalinity of the solution. At 25° C a neutral solution, such as pure water, has a pH value of 7. A solution with a pH below 7 is considered to be acidic and one above 7 is considered to be alkaline. Because pH is a logarithmic scale, the amount of strong acid required to decrease the pH by one unit value increases greatly. Water with a pH of 5 has 100 times the hydrogen ion concentration of a pH 7; a pH of 4 has 10 times the hydrogen ion concentration of pH 5, or 1000 times the hydrogen ion concentration of a pH 7 solution. [Said to be a shortened form of Potential (or Power) of hydrogen.]
Phagotroph An organism that feeds by ingesting solid food particles (e.g., bacteria). Such an organism is **phagotrophic**.
Phagotrophy The fact or process of feeding on particulate organic matter.
Phanerogam A plant with conspicuous reproductive organs (flowers); reproduction takes place by seeds.
Phanerozoic The geological interval of abundant animal fossils, beginning approximately 545 million years ago.
Pharynx An organ in animals, serving as part of the respiratory system. In fish it is the anterior portion of the alimentary canal, characterized by lateral buds that provide skeletal support for the gill region. In humans it is the part of the throat below the mouth and nasal cavity, and above the esophagus and larynx.
Phenetic Describing a taxonomic classification or grouping based purely on overall similarity.
Phenetic biodiversity A characterization of a biota that relates to variation in phenotypic characteristics (traits) of taxa, and estimated by components that reflect the size of the phenotypic hypervolume defined by constituent taxa, or the dispersion of phenotypes within that hypervolume.
Pheneticist A scientist who practices or studies the discipline of phenetics (see next).
Phenetics The study of phenotypes, especially the grouping of organisms into taxa on the basis of estimates of similarity. Pheneticists use matrices of overall similarity rather than parsimony to construct a phenogram (see next) as an estimate of phylogeny.
Phenogram A branching diagram showing overall levels of phenetic resemblance rather than cladogenesis.
Phenological Having to do with or based on phenology (see below).
Phenological shift A change over years in the mean of the phenological value of a trait within a population.
Phenology The study of the timing of recurring natural events, such as leafing or flowering, the initiation of reproduction, or the return date of migrant birds.
Phenomenological model A model that relates observable variables without postulating the underlying mechanisms (physical, biological, and so on). For example, the species–area relationship is a phenomenological model because it does not explicitly state the mechanisms leading to the relationship between the number of species and the area.
Phenophase A given phenological event that is being recorded, such as first leafing.
Phenotype The observable characteristics or traits of an organism, including its appearance and behavior. Compare GENOTYPE.
Phenotypic Relating to the observable characteristics of an organism.
Phenotypic plasticity The ability of an organism to change its phenotype in response to changes in the environment.
pH number or value See pH.
Phorb Another spelling of FORB.
Phoresis A method of long-range dispersal in which the dispersing animal attaches itself to another animal (e.g., beetle, wasp, or bird) that carries the disperser along with it until the disperser drops off or disembarks. Also, **phoresy**.
Phoretic having to do with or practicing phoresis (see previous).
Photic zone Another term for the EUPHOTIC ZONE.
Photoautotroph An organism able to synthesize organic compounds from water and inorganic nutrients, using the energy in photons of light.
Photoautotrophic Using light as an energy source for synthesis of organic matter from inorganic precursors.
Photoautotrophy A type of autotrophy (self-feeding) in which organisms gather light energy in order to reduce carbon dioxide to organic carbon; characteristic of green plants, most algae, and some prokaryotes.
Photochemical source (or **sink**) The production (or loss) of an atmospheric gas involving photodissociation or chemical reactions.
Photodissociation The destruction of a molecule through absorption of solar radiation and subsequent breaking of one or more chemical bonds.
Photolithoautotrophy A form of metabolism that depends on light as an energy source, inorganic compounds (e.g., H_2O, H_2S) for reducing equivalents, and CO_2 as a source of carbon.

Photoorganoheterotrophy A form of metabolism that utilizes light as a source of energy and organic compounds as sources of both reducing power and carbon.
Photoperiod The length of the daylight period. Photoperiod is sensed by green tissues, and is involved in phenological events such as leaf abscission in autumn.
Photoperiodism The fact of reacting physiologically to changes in the duration of the daylight period.
Photorespiration A biochemical process in which O_2 is assimilated during the oxygenation of Rubisco (Ribulose-1,5-bisphosphate carboxylase oxygenase), and CO_2 is given off in the metabolism of the products of Rubisco oxygenation.
Photosynthesis A fundamental chemical process in which green plants (and blue-green algae) utilize the energy of sunlight and other light to convert carbon dioxide and water into carbohydrates, with chlorophyll acting as the energy converter. This releases oxygen and is the major source of atmospheric oxygen. Photosynthesis provides green plants with their energy requirements, and allows other organisms to obtain nutrients from these plants, either directly or indirectly (by consuming organisms that consume green plants).
Photosynthetic Relating to or involved in a process of photosynthesis (see previous).
Phototroph An organism utilizing the energy of sunlight or other light for growth.
Phototrophic Describing energy metabolism based on light energy.
Phototrophy The ability to obtain nutrition or energy directly from sunlight and other light.
Phrygana Dwarf shrubland of the eastern Mediterranean Basin, characterized by a dominance of species with seasonal leaf dimorphism.
Phycobiont See LICHEN.
Phylesis A line or lines of descent between ancestors and descendants.
Phyllopod A leaf-like appendage.
Phylogenetic Having to do with or based on phylogeny (see below).
Phylogenetic diversity Adjusted diversity measures that take into account the degree of relatedness among a set of species in an assemblage. Other things being equal, an assemblage of closely related species is less phylogenetically diverse than a set of distantly related species.
Phylogenetics The study of evolutionary relatedness among various groups of organisms (e.g., species or populations).
Phylogenetic tree See PHYLOGENY (next).
Phylogeny The evolutionary relationships among taxa (groups of related organisms), often portrayed with some kind of tree-like branching diagram (phylogenetic tree), with the branches representing speciation events and the lengths of branches representing time since divergence.
Phylogeography Studies concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially within and among closely related species.
Phylogram A cladogram (see), especially one interpreted to show terminal taxa and ancestral nodes (intermediate specificity between a generalized cladogram and specific phylogenetic tree).
Phylotypes Molecularly differentiated taxa.
Phylum The highest level of taxonomic division in the animal kingdom, followed in descending order by class, order, family, genus, and species.
Physiognomy The general appearance of a given area, especially a natural area. For example, any desert has a certain recognizable physiognomy, regardless of species composition.
Physiography The landforms that lend shape and character to the continental landscape, for example, the configuration of mountains and drainage basins.
Phytochorion (plural **phytochoria**) A region within which a substantial proportion of the flora is endemic.
Phytophagous Feeding on plants; herbivorous.
Phytophagy Plant feeding; herbivory.
Phytoplankton Plant plankton; i.e., usually single-celled photosynthetic plankton, largely algae and cyanobacteria.
Phytotoxic Harmful to plants.
Picophytoplankton The smallest ($<$ 2 mm) size class of photoautotrophic plankton i.e., the picoplankton (see next).
Picoplankton 1. A classification of plankton according to size; those plankton of extremely small size (0.2–2 mm). 2. The fraction of aquatic biomass in the size fraction that passes through 3.0 and 0.8 μm filters, and is retained on a 0.1 μm filter.
Pinniped A member of the Pinnipedia, marine mammals with fin-like limbs including the seals, sea lions, fur seals, and walrus. Approximately one-quarter of the known modern species of marine mammals are pinnipeds. Pinnipeds are in the mammalian order Carnivora.
Pioneer One of the first groups of species to colonize a newly formed or denuded habitat.
Piscivore A predator that consumes fish in aquatic habitats.
Placenta In most mammals, the organ providing nourishment to the developing fetus via the umbilical cord.
Placental Having a placenta; see EUTHERIAN.
Planktivore A predator that consumes zooplankton in aquatic habitats.
Plankton Small organisms with no or limited powers of locomotion that float freely in the water column of lakes or the ocean and are largely dispersed by turbulence and other water movements. Phytoplankton (plant plankton) are plankton with photosynthetic pigments and zooplankton are the animals of the plankton.
Planktonic Having to do with or being one of the many types of small organisms that float freely in the water column.
Planktotrophic Feeding on plankton to survive and grow. This feeding pattern is **plankotrophy**.
Planned biodiversity Crops and livestock purposefully introduced and maintained in an agroecosystem by humans, as opposed to the associated natural biodiversity that already exists there (unplanned biodiversity).
Plantation An area of land that supports planted forest or food or fiber crops, usually for commercial exploitation.
Plantation monoculture A plantation forest involving a single species. The original biodiversity at the site is not recovered, though the protective function and many of the original ecological services may be re-established.
Plant breeding The systematic process of genetic improvement of plants for food and agricultural uses.

Plantigrade 1. Walking on the soles of the feet. 2. An organism that walks in this manner; e.g., humans, bears, mice. Compare DIGITIGRADE, UNGULIGRADE.
Planula (plural **planulae**) The ciliated planktonic larval stage of a coral.
Plasmid A piece of DNA (usually circular) which can replicate independently of the chromosome of a cell and which codes for supplementary (noncore) functions.
Plasticity The ability of a plant to respond to temporal changes or spatial variation in environmental conditions by altering the size or the distribution of its parts. These are phenotypic rather than genetic changes.
Plastron 1. In aquatic insects, a set of body structures that trap a thin layer of air against the body; oxygen from the surrounding water is added by means of diffusion to the same extent that oxygen is withdrawn by breathing. 2. The flat lower portion of the shell of a turtle.
Plate tectonics 1. A modern geological theory of tectonic activity, according to which the earth's crust is divided into a small number of large, rigid plates whose independent movements relative to one another cause deformation, volcanism, and seismic activity along their margins. 2. The study of the movement of these plates and of how this process has been involved in the evolution of Earth's crust.
Pleiotropic Describing a gene that has more than one independent phenotypic effect.
Pleiotropy The fact of one gene affecting more than one trait.
Pleistocene The geologic period from 2 million to 10,000 years before the present, which was characterized by a series of glacial advancements (ice ages) separated by warmer interglacial periods.
Pleopod An abdominal appendage in crustaceans.
Plesiomorphic Describing an ancestral or primitive character, relative to sister group.
Plesiomorphy An apomorphic character or character state that specifies a more inclusive group than that under consideration.
Pliocene An epoch within the Cenozoic, the most recent prior to the Quaternary, from about 5.3 to 2.6 million years ago. The Pliocene is characterized by largely modern flora and fauna, and is the most recent epoch to have experienced extended periods globally warmer than the present.
Plow horizon The homogenous soil layer of an agricultural field, typically at 0–20 cm depth, that results from the field being plowed; designated as the "Ap" horizon in soil morphology. See also SOIL HORIZON.
Plume prospecting The process of detecting chemical and physical anomalies in hydrothermal plumes to locate hydrothermal vents.
Plurivoltine Describing species with several generations of offspring per year. Also, MULTIVOLTINE.
Pneumatophores Roots growing vertically from underground roots that spread laterally from the main stem, occurring in some genera of mangrove, such as *Avicennia* and *Sonneratia*. They typically stand 10–20 cm above the soil surface, enabling gas exchange to take place with the underground roots.
Podzol A soil type highly characteristic of boreal forests and developed as a consequence of podzolization (see next).
Podzolization The process of acid leaching in soil, in which clay, organic particles, and mineral ions (primarily iron and aluminum) are carried downwards and deposited in the B soil horizon, leaving an impoverished and leached A horizon above. This occurs as a consequence of low temperatures and precipitation in excess of the needs of evapotranspiration. See also SOIL HORIZON.
Poikilohydry A condition in certain organisms, such as some aquatic algae, in which their internal water balance varies with changes in ambient humidity; i.e., they are not able to compensate for fluctuating humidity. Contrasted with HOMOIOHYDRY.
Point source A specific, readily identifiable input of pollution, such as when treated wastes are discharged from a municipal, industrial, or agricultural facility to receiving waters through a pipe or a drain. Contrasted with NONPOINT SOURCE.
Poisson distribution 1. A discrete probability distribution used to model the number of events of a given type occurring in a specific time and/or space. 2. In this context, the probability of distribution of numbers of organisms per area if the organisms are distributed completely at random in the space; e.g., the number of prairie dogs in a square mile of prairie. [Introduced by the French mathematician Siméon Poisson.]
Polarity The relative occurrence of derived versus ancestral character states.
Pollard A tree repeatedly cut at about 6–10 ft above ground, to produce successive crops of wood.
Pollution The introduction into air, water, or soil of contaminant substances in quantities that are threatening to living resources, biological processes, and/or human health and activities.
Polyandrous Describing a female that has more than one male mate.
Polyandry A mating system in which a female mates with more than one male.
Polyculture 1. The agricultural practice of growing two or more crops either simultaneously or in sequence on the same piece of land. 2. The presence of a diversity of species, or genotypes, of crop plants or other organisms.
Polygamous Reproducing with several different mates.
Polygamy A reproductive pattern in which males and females both have several different mates.
Polygynous Describing a male that has more than one female mate.
Polygyny A mating system in which a male mates with more than one female.
Polymerase chain reaction (PCR) An enzymatic technique for selectively copying (amplifying) a specific segment of DNA across several orders of magnitude. A common variation of this is **quantitative PCR (qPCR)**, wherein the progression of the amplification reaction is recorded optically in real time, providing a means for rapidly estimating the abundance of a particular sequence of DNA in a given sample.
Polymorphic Having to do with or showing a genetic difference between two or more individuals of the same species.
Polymorphism The existence at intermediate frequencies of two or more variants at a given gene locus within a population.
Polyp The basic unit of the coral animal, consisting of a simple tubular structure with a mouth and ring of tentacles at one end. The polyp sits in a skeletal cup called the **corallite**.
Polypeptide A single chain of amino acids.

Polyphyletic Describing a group of animals that, although grouped together in a phylogeny or evolutionary tree, do not share a common ancestor. Compare MONOPHYLETIC; PARAPHYLETIC.
Polyphyletic group A taxon for which all included species (or taxa) do not share a most recent common ancestors.
Polyphyly The fact of being a group that does not include the most recent common ancestor of all of its members.
Polyploid Having more than one set of homologous chromosomes.
Polyramous Relating to or having an extension of the body (usually an appendage) with multiple branches.
Polythetic group A natural taxonomic group in which the terminal taxa are not known to share any universal unique character(s) but are nonetheless united by overall similarity (phenetics) or global parsimony (cladistics).
Polytomy The fact of many branches arising from a single node or an unresolved (partial or whole) cladogram.
Polyunsaturated fatty acids Fatty acids with two or more double bonds in the chain of carbon atoms. They occur in distinct families defined by the positions of the double bonds, of which the commonest are n-3 and n-6.
Poor fen A term for an acidic peat land similar to a bog but having a somewhat higher surface water pH (4.3–5.8), because it receives some ground water or surface water in addition to precipitation; typically dominated by sphagnum mosses or sedges.
Population A group of individuals of a particular species that live in a given region and that are more likely to mate with one another than with individuals from another such group elsewhere. A population is usually a subset of the entire species.
Population dynamics Changes in population size through time, and the processes that bring about these changes.
Population regulation The constraint of positive population growth. The study of population regulation deals with the factors that cause this constraint, such as predation or competition for food.
Population stability The tendency for populations to return to a previous size after a disturbance, such as increases due to immigration, or reductions due to predation, hunting, or disease. Stable populations may be locally stable (return after small disturbances) or globally stable (return after severe or catastrophic disturbances).
Population viability analysis (PVA) A method of assessing the probability of extinction for a given population within a specified time span.
Poricidal anthers In certain plants, anthers that function similarly to salt shakers; they must be vibrated at the right frequency (by buzz pollinators) to dispense pollen grains through the pores at the end of the anthers.
Postemergence herbicides Herbicides applied after the emergence of the specified weed or crop.
Posttranscriptional modification The process of cutting and splicing messenger RNA, in some cases to produce alternative transcripts and proteins with different structural or regulatory properties (e.g., sex determination in *Drosophila*).
Potential energy See ENERGY.
Potential (natural) vegetation The vegetation cover that would develop naturally in an area and become stable (not replaced by a subsequent stage) if all outside disturbances were eliminated.
Poverty A pronounced deprivation in well-being in which a person lacks opportunities, empowerment, and security. Typically measured by income level or by the ability to purchase food and other staples.
Precautionary principle The principle that when there is a threat of serious or irreversible environmental damage, lack of full scientific certainty about outcomes should not be used as a reason to postpone measures to avoid that damage.
Precession The drift of the timing of the point in the year when Earth is closest to the Sun (**perihelion**), which influences the distribution of insolation with season and latitude.
Predation An interaction between two species in which one species consumes the other.
Predator An animal that naturally preys on others to obtain its food.
Predator-prey relationship The ongoing ratio in a given habitat of the population size of a predator and that of its target prey.
Predatory Having to do with or exhibiting the behavior of predation.
Pre-emergence herbicides Herbicides applied prior to the emergence of the specified weed or crop.
Preservation 1. The school of conservation philosophy that emphasizes the protection of natural features and landscapes from human exploitation. 2. The fact of taking active measures to maintain ecosystem viability.
Preserve Another word for a RESERVE.
Presoldier In termites, a transitional morph that precedes the soldiers; an unsclerotized individual whose head shows signs of soldier differentiation.
Prevalence The proportion of hosts in a population that are infected or diseased.
Prevention Efforts to reduce the probability of an invasive species being introduced and subsequently becoming established in a new habitat.
Primary compounds See SECONDARY COMPOUNDS.
Primary forest A forest that appears to be undisturbed by human influence.
Primary hemiepiphyte A plant that begins its life cycle anchored in a tree crown and ultimately becomes rooted in the ground.
Primary host See DEFINITIVE HOST.
Primary mutualism The initial association between a symbiont and a host.
Primary producer An organism that is capable of producing biomass from inorganic compounds.
Primary production 1. The process in which organic matter is synthesized from carbon dioxide by photosynthesis. 2. The amount of energy produced by this process. Also, **primary productivity**.
Primary succession The development of an ecosystem on a substrate previously devoid of vegetation, such as a sandbar.
Prime Describing agricultural land that has the best combination of physical (e.g., moisture status, soil depth, or slope condition) and chemical (e.g., nutrient status and salinity) characteristics for producing food, forage, fiber, and oil seed crops.

Primitive Another term for ANCESTRAL.
Priority The principle in scientific nomenclature that the name with the first date of valid publication is generally the one to be used.
Priority areas Areas that need the most urgent management; e.g. action to avert a threat or to meet a conservation goal (such as to increase the likelihood of persistence of valued species).
Priority effect The effect of the existing presence of an individual or taxon on newly arriving individuals; can be negative or positive.
Prisoner's dilemma Uncertainty about others' behavior leading to choices that are individually rational but collectively irrational; based on a classic example in game theory in which two prisoners will collectively benefit if they act in a cooperative way, but they act selfishly in the interest of individual benefit. Can be used as a context for describing or analyzing animal behavior.
Private good A good whose consumption typically occurs as the result of a market transaction between buyer and seller; called *private* because one person's consuming it reduces the amount available to others, and the provider of the good can exclude some people; e.g., those who are not willing to pay for it. Examples of private goods are a pound of apples or a gallon of gasoline. Compare PUBLIC GOOD.
Procaryote Another spelling of PROKARYOTE.
Process-based model(ing) A model explicitly based on mathematical representations of processes or mechanisms, such as animal or seed dispersal. For example, population dynamics models describe the evolution over time of the population abundance by accounting for processes such as birth, death, and dispersal.
Producer See primary producer.
Producer surplus A measure of the welfare that producers gain from selling goods and services, usually measured as the difference between the actual market price and the minimum amount that they would be willing to accept for the good or service.
Production The amount of newly formed biomass of an individual, population, or trophic level, including the organic matter lost during the period of observation through processes such as respiration, excretion, injury, or predation.
Production-function An approach that models ecosystem services as the relationship between ecological and human inputs (e.g., the structure and functions of an ecological system, and human labor and capital) and outputs valued by humans.
Productive benefits Values that are bought and sold in markets, such as honey and timber, or that enhance goods and services sold in markets, such as real estate.
Prognathous Describing the position of the head when the mouth opening is directed forward.
Progressive Describing successions in which the dynamic changes are in the directions of increasing species diversity, structural complexity, greater biomass, and increased stability. *Retrogressive* successions are in the opposite directions.
Progressive Era The period in American political history, especially coinciding with the presidency of Theodore Roosevelt (1901–1909), in which the environmental conservation movement gained definition and popular support.
Prokaryote A single-celled organism that has its DNA arranged as a circular molecule not contained within a nucleus, and which reproduces by simple fission. Prokaryotes are members of the domains Archaea and Bacteria, as opposed to EUKARYOTES, all other organisms that do have a cell nucleus.
Propagule Any part of an organism capable of developing into a new individual, such as a seed, spore, egg, bud, and so on.
Propagule pressure A concept that encompasses variation in the quantity, quality, composition, and rate of supply of alien organisms resulting from the transport conditions and pathways between source and recipient regions.
Property right The legal and social construct under which owners of property cannot be deprived of it without the payment of compensation that is acceptable to them.
Prosimians A category of primates that are neither monkeys or apes, such as galagos, lorises, lemurs, and tarsiers.
Prostomium The most forward part of the body in annelid worms; may carry antennae and eyes and at least part of the brain.
Protandrous Existing first as a functional male, then as a female.
Protandry A breeding system in which individuals change sex from male to female.
Protein A three-dimensional macromolecule constructed of amino acids that is formed on the basis of an RNA sequence.
Protein probe array See DNA PROBE ARRAY.
Protherian One of the Protheria, egg-laying mammals that are mostly extinct but include the platypus and echidna.
Protist One of a diverse collection of mostly unicellular eukaryotes that are defined by their exclusion from other groups; i.e., they are not animals, plants, or fungi. Includes protozoa, simple algae and fungi, and slime molds and water molds.
Protogynous Existing first as a functional female, then as a male.
Protogyny A breeding system in which individuals change sex from female to male.
Protozoa A group of mostly unicellular eukaryotic organisms, usually classified as animals because they ingest food and are capable of movement. However, some forms regarded as protozoa, such as slime molds and euglenoid forms, have the properties of plants.
Protozoan (plural, **protozoans** or **protozoa**) One of the protozoa (see previous). Also, **protozoon**.
Provenance test An experiment, particularly in forestry, in which plants from different source locations are grown together in common conditions, often in multiple locations, as a basis for assessing genetically based differences in trait values.
Provisioning services Ecosystem services (see) that supply humans with needed provisions, such as water, food or fibers, biochemicals for use in medicine or industry, and energy sources (e.g., biomass fuels, hydropower). Also, **provision services**, **providing services**.
Proximate Directly related; immediate; e.g., the agent that causes a species to become extinct. Also, **proximal**.
Psammophilic Sand-loving; living or thriving in sandy soil.

Pseudergate A false worker; among termites, a temporarily nonreproductive individual serving the colony in nutrition, construction, or brood care, which results from a late, reversible deviation from the pathways to the imago and is characterized by reduced wing buds compared to nymphs of the same stage.
Pseudocoelom A body cavity (coelom) not lined with a mesodermal epithelium.
Pseudoextinction The disappearance of a species from the fossil record due not to the death of all its members but rather to an evolutionary change that results in it being reclassified as a new species.
Psychrophile An organism that grows better at low temperature or that requires low temperature for growth. Such a condition is **psychrophily**.
Public good A good whose consumption is both *nonrival* and *nonexcludable*. That is, one person's consuming it does not preclude another (a rival) from doing so, and the provider of the good cannot exclude those who have not paid for it. Examples of public goods are clean air, a fireworks display, or a lighthouse. Compare PRIVATE GOOD.
Pull of the recent A term reflecting the fact that recently formed rock is more accessible and offers better preserved fossils; this may produce a bias toward the conclusion that diversity has increased toward the present time.
Puna An ecosystem of the high Andean mountains, found above the tree line from Bolivia to southern Peru. It has four distinct ecological regions: wet, dry, thorny, and desert. The dominant vegetation types vary from grassland/shrublands to desert.
Punctuated equilibrium The morphological stasis of fossil species over long periods of time, punctuated by seemingly abrupt speciation and morphological change.
Purebred An animal that is a member of a recognized livestock breed, as determined by pedigree records, geographical location, or knowledge of the breeding structure of the herd or flock.
Pure culture An organism growing in the absence of all other organisms.
PVA Population viability analysis.
Pycnocline The water layer in which density changes most rapidly with depth and separates the surface mixed layer from the deeper waters.
Quasispecies Hierarchically ordered populations of mutants resulting from erroneous copying of a genotypic ancestor. Whole mutant distributions behave like single species and are subjected to natural selection.
Quaternary The past 2 million years (approximately) of Earth's history, including the Pleistocene and Holocene (or Recent) epochs, particularly characterized by the extreme fluctuations in global temperature patterns that produced the ice ages. See also LATE QUATERNARY.
Queens and kings A term for females and males that actively reproduce in a colony.
Quiescence A state of dormancy in which inhibition of development depends directly on environmental factors.
R0 The basic reproductive ratio of a parasite or pathogen. For a microparasite, this describes the number of new infections generated by a single infected host entering an entirely susceptible population. For a macroparasite, R0 is the number of adult offspring produced by a single adult parasite over its lifespan.
Radiation All the species descended from a single common ancestor; tends to be a more diverse group in comparison with others of a similar age.
Radiative forcing A measure used to express and compare the potential of contemporary climate system changes (vs. preindustrial ones) to perturb the Earth's energy balance; reported in watts per square meter (W/m^2). A positive radiative forcing tends to warm the Earth's surface, and a negative one tends to cool the surface.
Radiocarbon dating The dating of organic matter in geological samples using as the standard the radioactive decay of Carbon-14, which has a half-life of about 5, 750 years. When a plant or animal dies it has approximately the same level of Carbon-14 as the atmosphere, but this decreases from that point onward. By dating tree rings of known age, and other methods, radiocarbon determinations can be calibrated.
Radula In mollusks, a ribbon-like structure bearing transverse rows of small "teeth," typically used to collect food and move it into the mouth.
Rainshadow A less rainy sector on the leeward side of mountains and protected valleys, caused by descending warm air masses.
Rami See CAUDAL RAMI.
Range 1. The extent of the area in which a given organism or species is found. 2. See **rangeland**.
Range core See CORE RANGE.
Range expansion The process in which a species spreads into new areas (usually new regions, rather than local-scale movements) due to natural dispersal, though the expansion may be assisted by human-mediated changes to the environment. The concept can be applied to both native and alien species.
Rangeland Land on which the potential native vegetation is predominately grasses, grass-like plants, forbs, or shrubs. Rangelands include prairies, marshes, tundra, wet meadows, savannas, shrubland steppe, chaparral, desert grasslands, and woodlands.
Range periphery The marginal parts of the distribution of a species.
Range state A country in which an animal is found.
Rank-abundance curve A plot of the number of individuals in each species ordered by their rank, starting with the most abundant species and ending with the most rare.
Rapoport effect A pattern in which the size of a species' geographic range increases toward higher latitudes. [Named for the Argentinean ecologist Eduardo Rapoport, who observed this pattern.] Also known as **Rapaport's rule**.
Rare A classification for species with small populations globally that are not presently listed as endangered or vulnerable, but are at risk because of their limited population size.
Rarefaction A statistical interpolation method of rarefying or thinning a reference sample by drawing random subsets of individuals (or samples), in order to standardize the comparison of biological diversity on the basis of a common number of individuals or samples.
rDNA See RECOMBINANT DNA.

Reaction center A bacteriochlorophyll–protein complex that performs the primary photochemical reactions of photosynthesis.
Reaction norm The relationship between the phenotypic expression of a single genotype and the environmental conditions under which the phenotype is shaped.
Reactive nitrogen All forms of nitrogen except atmospheric dinitrogen gas, including all radiatively, photochemically, and biologically active inorganic forms (e.g., NH_3, NH_4^+, NO_x, HNO_3, N_2O) and organic molecules (e.g., proteins, urea, etc.).
Realm See BIOGEOGRAPHIC REALM.
Rear edge A stable margin of a species' range where population size is declining.
Recalcitrant Describing seeds that have high moisture content and lose viability if moisture drops below a critical amount limiting their storage for long periods.
Recipient control A situation in which consumers substantially depress populations of their resources.
Reciprocal selection Evolutionary change as a result of species acting as selective agents on each other through interspecific interactions.
Reclamation 1. A revegetation or land management goal that includes a lower diversity of species and may include substitutions by introduced species. 2. The conversion of an area of the ocean, a lake, a riverbed, or a wetland to dry land. 3. The restoration of depleted or desert land to productive agricultural land. 4. The cleansing of wastewater for return to the natural environment.
Recolonization The reappearance of a species in an area where it has earlier been present, then disappeared.
Recombinant DNA (rDNA) A DNA sequence obtained by artificially combining two or more different DNA strands, thus creating a sequence that would not be found in nature. Recombinant DNA is widely used in biotechnology, biomedicine, and research; e.g., comparative phylogenetic studies.
Reconstructed evolutionary tree A phylogeny linking contemporary species, and thus containing no information about lineages that have gone extinct.
Recruitment The arrival of new juvenile animals or plants into a habitat or an older stage of the population; e.g., the entry of juvenile fish into the adult stock.
Recruits A term for young fish that are first caught by a fishery.
Recycling 1. The process of reusing industrial materials (e.g., aluminum, glass, plastic, paper) after the initial use, to preserve useful materials and reduce the need for manufacturing new materials (as well as for other benefits). 2. The transpiration and evaporation of water from the land surface, followed by reprecipitation elsewhere within the continent.
REDD Reduced emissions from deforestation and forest degradation; a form of payment for ecosystem services in which the carbon storage value of forests threatened by degradation or clearance is financially recognized through payments to forest owners (usually, but not exclusively, nation-states) to conserve the forest.
Redfield ratio The ratio of carbon, nitrogen, phosphorus, and oxygen for normal ocean plankton and deep-ocean nitrate and phosphate pools, on the basis of time-averaging for a large-area. Also, **Redfield stoichiometry**. [Described by the U.S. oceanographer Alfred C. Redfield.]
Red list See IUCN RED LIST.
Red tide The popular name for a harmful algal bloom (HAB) having a reddish or brownish coloration. See ALGAL BLOOM.
Reference area An undisturbed or natural area chosen to compare with a restored site to determine the success of restoration.
Reflective beach A narrow, steep beach; so called because it reflects a major part of the incoming waves.
Reforestation The re-establishment of trees and understory plants at a site previously occupied by forest cover. This can occur by natural means or by some form of human intervention, including planting seedlings or sowing seeds.
Refugia Arctic land areas that were not covered by ice sheets or glaciers during the last glaciation. Consequently, some biota could survive there and recolonize adjacent areas when the ice retreated. Therefore, these areas are associated with high biodiversity and endemism. Refugia usually occurred in coastal areas (now continental shelves) and ice-free mountain tops or "nunataks."
Regenerated production Primary production in the ocean derived from recycling of nutrients within the euphotic zone such as by uptake of ammonia (NH_4^+) or dissolved organic nitrogen. Compare NEW PRODUCTION.
Regeneration 1. The ability of organisms to regrow lost or destroyed tissue or body parts. 2. See FOREST REGENERATION.
Regime The specific, relatively stable structural and functional state of a given ecosystem.
Regime shift A large, persistent, often abrupt change in the structure and function of an ecosystem, associated with a change in dominant feedback processes. Regime shifts are abrupt and often difficult to reverse.
Region A large geographic area that contains more than one landscape.
Regional pool Species occurring within a defined geographical area or in a collection of samples from within such an area.
Regulating services The ecosystem services (see) obtained from the regulation of environmental processes, including flood protection, climate regulation, human disease regulation, water purification, air quality maintenance, pollination, and pest control.
Rehabilitation The process of re-establishing the productivity and some, but not necessarily all, of the plant and animal species thought to be originally present at a site. For ecological or economic reasons, the new ecosystem might also include species not originally present at the site. The protective function and many of the ecological services of the original system may be re-established.
Reintroduction The placement of individual organisms by humans into a habitat for the purpose of establishing a population where a previous population has gone locally extinct.
Relative abundance distribution A frequency distribution depicting the number of species in a community as a function of the number of individuals comprising each species.
Relay intercropping See INTERCROPPING.

Relict 1. Surviving; being a survivor. 2. A surviving species of an otherwise extinct group of organisms. **3**. A formerly numerous and widespread species that now persists only in small numbers in a limited area.

Relictualization The survival of species on islands or other relatively small areas, after those on larger mainland areas have suffered extinction.

Remediation The removal of toxicants from a contaminated environment, using chemical, physical or biological means.

Remote detection The assessment of a particular variable or parameter from afar, without a local human presence; e.g., a substance that could be dangerous to investigate at close range.

Remote sensing The technique of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation (e.g., by means of a satellite image).

Renewable Describing a resource that can be grown or produced at a level which over time will sustain its consumption rate, assuming sustainable use in the present; for example, wood, agricultural plants, fish stocks. Contrasted with NON-RENEWABLE.

Renewable energy Those forms of energy (**renewables**) whose transformation does not consume the ultimate source of the energy, harnessing instead sources such as solar radiation, wind, the motion of water, and heat from subsurface water and geologic formations. The major forms of renewable energy are solar, wind, geothermal, hydropower, and biomass. Compare NON-RENEWABLE ENERGY.

Renewable fuel A fuel obtained from a renewable source, such as ethanol from corn.

Renewable Fuel Standard (RFS) A U.S. federal government mandate requiring certain quantities of four categories of biofuels to be used (e.g., by the blending of biofuels into gasoline), with the total reaching 136 billion liters (36 billion gallons) of ethanol equivalent by 2022.

Renosterfeld Evergreen, needle-leaved shrubland on richer soils of the Cape Floristic Region of South Africa; dominated by the resinous shrub *Elytropappus rhinocerotis* (Asteraceae), whose name literally means "rhinoceros bush."

Rent See ECONOMIC RENT.

Repetitive DNA Regions of DNA that include the same DNA sequence repeated several hundred or thousand times; regions with repeated segments that involve only 2–5 base pairs of DNA are called **microsatellites**.

Representation The occurrence of species (or other attributes) within a set of selected areas.

Reproductive isolation The existence of intrinsic barriers (e.g., large rivers, mountain ranges) to the production of offspring.

Reproductive success The number of an organism's offspring surviving to sexual maturity, relative to the number produced by others in the same population.

Reptile A member of a class of vertebrates (the Reptilia), including turtles and tortoises (order Testudinata), lizards, snakes, and worm-lizards (order Squamata), the tuatara (order Rhynchocephalia), and crocodiles and alligators (order Crocodylia). Reptiles typically lay eggs with a leathery, impermeable shell and have a dry, horny skin with scales, plates, or scutes.

Rescue effect A reduction in the likelihood of local extinction in a habitat patch by the presence of emigrants from other habitat patches.

Reserve An area in which the primary management objective is to fully protect native species and ecosystems from direct human modification.

Reserve selection algorithm A mathematical technique used to maximize efficiency in the selection of protected areas for conservation. The efficiency criteria vary with circumstances but may, for example, be the minimum number of reserves with every species represented, or minimum cost.

Reservoir competence The degree to which a host is capable of supporting pathogen proliferation and transmission to other hosts or to vectors.

Reservoir host A host organism that can carry and transmit a parasite or pathogen to other host species, while suffering little or no negative effects itself.

Residence time The length of time that a given unit of water remains in a given place and thus reflects the exchange rate of water at that place. It can be calculated in various ways; e.g., by dividing the volume of the area of interest (e.g., a floodplain site) by the flow rate.

Residual impact The remaining adverse impact on biodiversity after appropriate avoidance, minimization, and restoration actions have been taken according to the mitigation hierarchy.

Resilience 1. The ability of an ecosystem to recover to conditions similar to those that existed prior to the imposition of a disturbance. 2. The rate at which such a recovery occurs.

Resistance 1. The degree to which an ecosystem is able to remain unchanged when subjected to a disturbance; e.g., climate change. 2. The ability of an individual to minimize the effects of a disease or other harmful agent. **3**. The fact of a harmful pest having reduced susceptibility to a given pesticide.

Resolution 1. The magnitude of the biological response to a toxic chemical. 2. The unit size of the information presented in an ecological model.

Resource 1. Anything that is used or consumed by an organism for growth, maintenance, reproduction, or survival, including food, space, and shelter. 2. Any substance, material, process, or feature that has value for human use, either natural or man-made. See also NATURAL RESOURCE.

Resource catchment The locality from which the resources consumed by a group of people are derived.

Resource conservation See CONSERVATION.

Resource facilitation The fact of the activity of one species promoting the availability of resources for another species; e.g., a plant that provides shelter (or greater access to sunlight) for another.

Resource partitioning An ecological arrangement in which two or more species use different, non-overlapping resources in a given habitat, such as birds foraging for insects in different locations within a tree or canopy.

Response curve The relationship between a species probability of occurrence or its density (number of individuals per unit area) and the values of a given environmental variable.

Resprouting In woody plants that are damaged but not killed by fires, the growing of new leaves, either from the base or from epicormic shoots (new shoots originating beneath the bark).
Restoration A process of assisting in the recovery of damaged, degraded, or destroyed ecosystems, by the manipulation of organisms and ecological processes to create self-organizing systems that resemble predisturbance structure and functioning and promote biodiversity. The study of such a process is **restoration ecology**.
Reticulate evolution A pattern of speciation in which new species arise from interspecific hybridization between two species, coupled with establishment of reproductive isolation.
Retinal The rhodopsin chromophore, responsible for the color of rhodopsin pigments. Also called **retinaldehyde** or **vitamin A aldehyde**.
Retrogressive See PROGRESSIVE.
Revegetation The process of replanting grasses, bushes, and other types of vegetation on cleared and bare soils.
RFS Renewable Fuel Standard.
Rhizodeposition The mixture of sloughed cells, mucilages, and low molecular-weight sugars, amino acids, and other compounds leaked from root cells, which are deposited in the soil adjacent to the surface of fine roots.
Rhizoplane The soil directly in contact with the plant root surface.
Rhizosphere The volume of soil adjacent to, and strongly influenced by a plant root; usually considered to extend about 2 mm from the root surface.
Ribosomal RNA (rRNA) A segment of ribonucleic acid (RNA) that serves as a structural element of a ribosome. Ribosomes are macromolecules that translate messenger RNA into protein. They are universally distributed molecule among cellular life forms and are widely used to infer evolutionary relationships.
Ribotyping The fingerprinting of genomic DNA restriction fragments that contain all or part of the ribosomal RNA gene, in order to identify and distinguish among strains or species (which are thus defined as **ribotypes**).
Rich fen An alkaline peat land that receives significant inputs from ground water or surface water, typically having surface water pHs 4–6.0, higher alkalinity, and distinctly higher concentrations of base cations than bogs and poor fens. Such terrain is dominated by brown mosses (true mosses, largely of the Amblystegiaceae) and small sedges.
Richness See SPECIES RICHNESS.
Riparian having to do with rivers; located in or on a river.
Riparian forest A forest that grows adjacent to rivers and within floodplains, generally composed of trees tolerant of periodic flooding and riparian trees that require flooding to promote their regeneration.
Riparian zone The interface between rivers or streams and terrestrial uplands. These zones are three-dimensional areas of direct interaction between the terrestrial and aquatic ecosystems, and they often contain unique species and high levels of biological diversity.
Risk A situation in which outcomes are stochastic (not deterministic) but each alternative has a known probability of occurrence. Risk may be idiosyncratic (specific to an individual) or systemic (shared).
River continuum A series of conceptual, unifying models to explain and predict the structure and function of river ecosystems.
RNA Ribonucleic acid, composed of nucleotides as is DNA, but in RNA the base uracil (U) replaces the base thymine (T); RNA molecules form important structural and regulatory parts of cells.
Root length density (RLD) Root length per unit of soil volume (m/m^3 or cm/cm^3).
Root parasite A parasite that forms haustoria (penetrating parts) on the roots of the host plant.
Root weight density (RWD) Root fresh or dry weight in unit soil volume (g/cm^3).
Rostrum 1. A beak; any of various beaklike structures in different organisms. 2. Specifically, elongated mouth parts in certain insects, adapted for piercing the organisms on which they feed.
Rotation See CROP ROTATION.
rRNA See RIBOSOMAL RNA.
Rubisco (RuBisCO; RuBP) Ribulose-1,5-bisphosphate carboxylase oxygenase, the central enzyme of photosynthesis.
Rubisco capacity The maximum rate of carboxylation (incorporation of carbon dioxide into carbohydrates), determined principally by the amount of Rubisco (see previous) and other proteins needed for photosynthesis.
Rumen The first, and usually largest, chamber of the four-chambered stomach found in ruminant ungulates; the place where cellulose digestion occurs.
Ruminant 1. A grazing mammal with a complex four-chambered stomach (including a rumen). Common ruminants include cattle, sheep, goats, deer, antelopes, and giraffes. 2. Having to do with or being such an organism.
Runoff Water that is left over after evapotranspiration; the source of fresh water in streams and rivers.
Russell cycle A major change in the biota of the western English Channel from 1925 to 1972, in which first there was a change from one fairly stable configuration of organisms to another very different one, followed by a return to the first. [Named for British marine biologist Frederick Stratten Russell.]
SAD Species abundance distribution.
Sahul A collective term for New Guinea, Australia, and their islands.
Saltation 1. An abrupt, large-scale evolutionary change. Contrasted with GRADUALISM. 2. The movement of sand or other particles "jumping" across the surface due to wind or water action.
Salt wedge Water from the ocean, with higher salt content, that moves into an estuary along the bottom of the water column, beneath less dense fresh water that has moved into the same area from a river. The salt content makes the ocean water heavier than the fresh (riverine) water, so that under calm conditions, the estuarine water becomes density-stratified.
Sampling effect The hypothesis that diversity might influence an ecosystem process because of the greater chance that a given species trait would be present at higher diversity, and the effect of its presence on ecosystem functioning.
Sandwich hybridization A method for detecting and quantifying specific sequences of nucleic acid. A **sandwich**

hybridization assay employs a capture probe to immobilize a target molecule onto a solid support and a signal probe to reveal its presence.
Saprotroph A detritus eater; an organism that feeds on dead organic matter; e.g., a fungus or bacterium. Also, **saprovore**. Such an organism is **saprotrophic**.
SAR Species-area relationship.
Satellite virus A virus that does not encode all of the structural or enzymatic components required to replicate itself in a given host cell, but which can replicate through coinfection with a helper virus which provides the missing functions.
Saturated fatty acid A fatty acid with the maximum complement of hydrogen atoms and no double bonds between carbon atoms.
Saturation The upper limit to species richness within a community, set by species interactions, and independent of the pool of colonists for which that community is accessible.
Saurischian One of the Saurischia, the so-called reptile-hipped dinosaurs, including sauropods (see next) and theropods.
Sauropod One of the Sauropoda, quadripedal herbivorous dinosaurs that are the largest animals ever to have lived on land, including such well-known genera as *Diplodocus*, *Brachiosaurus*, and *Apatosaurus* (*Brontosaurus*).
Savanna A mixed tree and grass system, with trees ranging from sparse to medium dense, with an understory of grasses. Savannas are open, well lit systems. Rainfall usually occurs in the warmer, summer months with a dry period of 2–8 months. Fires are typical across savannas during drier months, and occur at intervals from 1 to 50 years. Also spelled **savannah**.
SC Synaptonemal complex.
Scala naturae "Nature's ladder;" the medieval concept of life as being arranged in a hierarchical chain or ladder, with God and the angels at the top and living things on Earth below, placed according to their perceived complexity. Thus a typical progression would be inanimate materials (minerals, soil), plants, insects, fish, amphibians, reptiles, birds, mammals, humans. An influential belief in the early days of modern biology but eventually abandoned, though reminders of it still exist, as in the idea that mammals, birds, and reptiles constitute the "higher" vertebrates. Also known as the **Great Chain of Being**.
Scale The distance or time over which a significant change in some quantity of interest can be measured.
Scale dependence A condition in which either the form or the parameters of a relationship between two variables (e.g., species richness and latitude) is contingent on spatial or temporal attributes.
Scale of effect The spatial extent over which the habitat and the matrix characteristics in the surrounding landscape influence individual, population, or community processes in a focal area.
Scansorial 1. Adapted for or capable of climbing. 2. Using both the forest floor and canopy for movement and seeking resources.
Scent glands Glands through which many types of organisms secrete various compounds for purposes such as defense, marking of territory, attraction of mates, and attraction of both sexes.
Scleractinia The taxonomic order of cnidarians, marine invertebrates including the reef-building corals.
Sclerophyll A plant with tough, leathery, evergreen leaves, usually associated with drought resistance.
Sclerophyllous Describing leaves with a leathery texture due to the presence of **sclerenchyma** cells with large amounts of lignin and cellulose in their tissues.
Scutellum "Little shield;" the dorsal portion of the insect mesothorax, in most heteropterans it is triangular or shield-shaped.
Seagrasses A collective term for grasslike flowering plants that are able to grow submerged in marine and estaurine waters, to which they are restricted.
Seasonal dimorphism A trait in which plant species change the morphology of leaves throughout the year to adapt themselves better to prevailing temperature and water stress.
Secchi depth Measurement of the transparency of water by determining the depth at which a white **Secchi disk** (diameter 20 cm) lowered from the surface disappears out of the sight of the observer. [Developed by the Italian scientist Father Angelo Secchi.]
Secondary compounds Chemical compounds that serve as defenses in plants; contrasted with **primary compounds** that are involved with processes of primary metabolism, such as photosynthesis and cellular respiration.
Secondary extinction The extinction of a species resulting from the extinction of another species on which it relies.
Secondary forest An area of previously forested land that was subsequently degraded or deforested through human or natural action but which now supports regenerating or mature natural forest.
Secondary hemiepiphyte A plant that begins its life cycle as a terrestrial seedling, ascends a tree, and can later lose root connections with the ground, including both lianas (see) and vines.
Secondary host See INTERMEDIATE HOST.
Secondary metabolite An unusual compound that does not play a role in primary metabolism (e.g., production and storage of energy). Such metabolites were initially thought to be waste products but are now known to have adaptive or defensive value.
Secondary mutualism The incorporation of a primary mutualism into a secondary host; the primary host either becomes the symbiont or is lost.
Secondary production The production of new living material through tissue produced by heterotrophic organisms, which consume living or dead material from other organisms. Compare PRIMARY PRODUCTION.
Secondary succession The re-establishment of an ecosystem on existing substrate (soil), following a disturbance that removes existing vegetation.
Seed bank An accumulation of dormant plant seeds that persist in dormancy for longer than a single growing season.
Seeding 1. The process of planting or sowing seeds. 2. Describing plants whose seeds germinate abundantly after fires, thus quickly re-establishing a population from seeds.

Segregation The allocation of genetic variants (alleles) to different gametes during sexual reproduction.
Selection The differential survival or reproductive success of individuals, associated with differences in phenotype or genotype.
Selection regime A situation in which a particular set of evolved traits is favored; thus a newly established population will evolve in a different direction from the ancestral population.
Selective constraint See FUNCTIONAL CONSTRAINT.
Selective logging A less intense form of timber extraction involving removal of select trees, often based on species, size, or age class. The forest canopy is typically maintained under this process.
Selectivity A phenomenon in which some plants are killed with doses of herbicides that have little or no effect on other plants.
Self-incompatibility A genetically based system that prevents fertilization of an egg by pollen that shares one or more alleles at a self-incompatibility locus; a mechanism to prevent self-fertilization in plants.
Self-organization 1. A process in which pattern at the coarse scale of a system emerges solely from numerous interactions among the lower-level components of the system. 2. Specifically, the spontaneous formation of complex structural aggregates from biological macromolecules such as proteins, nucleic acids, or lipids. The assembly pathways are usually determined by physicochemical properties of the constituent molecules.
Self-pollination The pollination of a flower by pollen from the same flower, or from another flower on the same plant.
Semaphoront An assemblage of individual organisms at a particular point in time in their developmental (ontogenetic) history.
Semiparasite Another term for a HEMIPARASITE.
Semelparity The fact of breeding only once and then dying, or the process of doing this; sometimes called "big bang" reproduction because it is both conclusive and large-scale. Compare ITEROPARITY.
Semelparous Describing an organism that produces offspring (flowers and fruits in plants) only once in its lifetime before dying.
Seminatural Describing vegetation that owes its character to human activity, but in which the plants are wild, not sown nor planted; e.g., coppice woods; most heathland; some savannas. Seminatural vegetation can provide suitable habitat for a large number of native species.
Semishrub A plant whose new growth is herbaceous but whose older base parts are woody. Also called a SUBSHRUB.
Semi-steppe batha Semishrub communities developing on most soil types at the boundary of the Mediterranean zone of the Near East.
Sensu Latin for "in the sense of;" used to describe how precisely or generally a term is defined. A *sensu stricto* usage employs the narrower sense of the term; a *sensu lato* usage is more inclusive. An example would be the term *symbiosis* which can be defined strictly as a mutually beneficial relationship between organisms. or more broadly as any long-term association, regardless of effect.
Sere 1. One of a series of stages in an ecological succession sequence, or the entire set of stages. 2. Describing withered, very dry vegetation, or an area with such vegetation.
Serotiny The fact of seeds being stored on a plant with dispersal triggered by fire or another such event.
Setae (singular **seta**) Bristle- or hair-like structures, such as the extensions of the cuticle that are characteristic of arthropods and related phyla.
Settlement Land used for human habitation, including cities, towns, villages, rural communities, roads, and so on.
Sexual selection Selection that acts directly on mating success through direct competition between members of one sex for mates, through choices made between the two sexes, or through a combination of both.
Shadow price A measure of the social opportunity cost of resource use; the net increase in societal well-being that would be enjoyed if a unit more of the resource were made available, other factors being equal. Since shadow prices reflect the social scarcities of capital assets, it is only in exceptional circumstances that they equate to market prices.
Shannon entropy An index that quantifies the uncertainty regarding which type (such as species) is picked when one entity (such as an individual) is drawn at random from the data set of interest. Often used as a diversity index. [Described by U.S. mathematician Claude Shannon.]
Shelf See CONTINENTAL SHELF.
Shifting cultivation A cyclic agricultural system widely used throughout the tropics; it involves cutting and burning of small patches of forest, followed by several years of crop cultivation, then a fallow period during which the forest regenerates and can be cleared again. Also, **shifting agriculture**.
Shoot growth Biomass production (per hectare) of foliage shoots produced annually in the overstory of a plant community.
Shotgun DNA A technique that isolates DNA (deoxyribonucleic acid) fragments from soil or water samples and compares them to known DNA libraries to determine the presence of novel species.
Sibling species A pair of closely related, morphologically similar species (usually sister species).
Siderophores Chemicals secreted by roots (primarily non-protein-forming amino acids), which complex with insoluble metal ions, bringing them into solution and permitting their transport to and uptake into the root.
Silvics The scientific study of forest trees, including their life history, growth, reproduction, and response to environmental change.
Silviculture The branch of forestry concerned with the theory and practice of controlling forest establishment, composition, and growth.
Simpson (diversity) index An index of diversity based on species abundance data. If p_i is the frequency of species i in a community of n different species, then Simpson diversity $= (1 - \Sigma_i^n = p_i^2)$ and is equal to the probability that two randomly chosen individuals from a given community are of the same species.
Simulation buffer The area surrounding a context landscape that allows for the spread of spatial processes into this landscape.

Simulation landscape The area being simulated to produce an HRV (historic range of variability; see) time series that usually includes both the context landscape and the simulation buffer.

Sink 1. A physical reservoir of a compound that can take up additional quantities of this compound from the air and thus reduce its atmospheric concentration. 2. Specifically, a component of the carbon cycle that is a net accumulator of carbon annually, such as the ocean or the forests.

Sink population A deme (localized breeding group) in which death rates exceed birth rates. Sink populations will go extinct without an input of immigrants.

Sirenian One of the Sirenia, the order of marine mammals that includes the manatees, dugongs, and seacows.

Sister groups Two terminal taxa or monophyletic groups that share a common ancestor that has not given rise to any other taxon under consideration.

Sister species (taxa) A pair of species (taxa) that have arisen from a single speciation event; each is the other's closest relative.

Site history The collection of past events involving natural disturbance or land use conversion at a given site.

Smolt A downstream migrating salmon that has begun the physiological transition to seawater.

SMRS Specific mate recognition system.

SNP Single nucleotide polymorphism; population variation at one nucleotide position.

Social cooperation A mutually beneficial association, temporary or permanent, among members of the same species.

Social costs The total costs to society of an action. These may exceed (or in some cases be less than) the private costs, which are the costs of that action to the individual or institution executing it.

Social differentiation The specialization of functions in a human society, as societies increase in size and complexity over time.

Social discount rate The rate at which society is willing to substitute future consumption for present consumption.

Social ecology The ecological study of various social species such as ants, bees, wolves, dolphins, and orangutans (or humans).

Social equity The fair, just, and equitable distribution of the stocks or flows of capital assets that contribute to human well-being.

Soil aggregate A crumb-sized unit of soil, composed of aggregated soil minerals, microbes, and soil microfauna, which are cemented together by a combination of biological materials such as polysaccharide secretions, fungal hyphae, and precipitated carbonates or silicates and other chemical substances. Aggregates are classified by their size and stability in water (disintegrating versus retaining their structure and integrity).

Soil horizon One of the distinct horizontal layers making up the soil. They range from the rich, organic upper layers of humus and topsoil known as the **O Horizon** and **A Horizon**, through the **E Horizon** (an intermediate leaching layer), to the underlying rocky layers of subsoil (**B Horizon**), regolith (**C Horizon**), and bedrock (**R Horizon**).

Soil organic matter (SOM) Organic substances, including a wide variety of carbohydrates, proteins, lipids, waxes, phenolic, and humic compounds, that accumulate in soil as a result of both plant and microbial growth. These include materials of low molecular weight that are rapidly decomposed to carbon dioxide; larger compounds that may be slowly decomposed over years to decades; and large, complex, aromatic substances, which may be stable within the soil for millennia.

Soil salinization index The molar ratio of soda and potash over alumina, determined by whole-rock chemical analysis, and used as a proxy for paleotemperature.

Soil structure A spatial arrangement of soil particles.

Soil texture The relative abundance of sand, silt, and clay particles in the soil. Particle size distribution determines the distribution of pore sizes, which in turn strongly affects the behavior of water in the soil.

Solar-heated Describing water whose temperature significantly exceeds the regional temperatures of lakes and streams. Many of the thermally tolerant and thermal-opportunist groups of organisms are thought to have evolved in shallow solar-heated ponds and water margins.

Sole cropping The agricultural practice of having one crop variety grown alone in pure stand at normal density.

SOM Soil organic matter.

Source population A deme (localized breeding group) in which birth rates exceed death rates. Sources produce an excess of individuals who will emigrate from that deme.

Source-sink dynamics Dynamics between populations that are net exporters of individuals (source populations) and populations that are net importers (sink populations).

Southern Ocean The circumpolar ocean in the Southern Hemisphere between the Subtropical Front and the continent of Antarctica.

Spatial Having to do with physical space; measured by or in the context of space.

Spatial autocorrelation The fact that observations are not independent identically distributed, but are correlated over some distance in space. **Positive autocorrelation** indicates that nearby values are similar; **negative autocorrelation** indicates that nearby values are dissimilar.

Spatial heterogeneity In community sampling, variation in species capture probabilities across space; this may be due to habitat variation, intraspecific clumping, interspecific association, or other factors.

Spatial subsidy Input from other habitats of organic carbon, nutrients, and prey or the movement of consumers. These resources can significantly influence the energy, carbon, and nutrient budget of recipient habitats.

Spawner A mature individual of a stock responsible for reproduction.

Specialist brood parasite A species that parasitizes only one or a few host species.

Specialization An evolutionary adaptation to a particular mode of life or habitat.

Speciation Development of species; the evolutionary process by which new species arise in a biological system.

Species A fundamental category for describing organisms, defined in several ways but typically on the basis of reproduction; i.e., members of a species can interbreed freely

with one another to produce fertile offspring, but cannot do this with members of other species. In taxonomic classification, a species is a subdivision of a genus.
Species abundance distribution (SAD) The number of species in a community as a function of abundance.
Species accumulation curve A representation of the total number of species observed in an assemblage, plotted against some measure of cumulative sampling units. Similarly, **species-area curve** or **model**.
Species-area relationship (SAR) The relationship between the size of a given habitat or area sampled across habitats, and the number of species (typically of a specific taxonomic group) found in that area.
Species Distribution Model A model used to describe and understand the geographic distribution of species, usually through interpolation and/or extrapolation of empirically derived correlations between environmental factors and the presence or absence of the species.
Species diversity A measure of diversity that incorporates both the number of species in an assemblage (species richness) and the way their abundances (i.e., number, biomass, and cover) are distributed (species evenness).
Species evenness The degree to which all species have the same proportional abundance.
Species flock(s) An aggregate of closely related species that share a common ancestor and are endemic to a geographically circumscribed area.
Species lifespan The time between the first appearance of a species in the fossil record and its final disappearance. This time span is typically in the range of millions of years.
Species niche See NICHE.
Species optimum The value of a given environmental variable at which a species reaches its maximum probability of occurrence or abundance.
Species packing The study of how species on the same trophic level coexist in a limited region or context.
Species pool The set of species that are able to colonize a habitat of interest.
Species range See RANGE.
Species response curve See RESPONSE CURVE.
Species richness The total number of species present in a habitat or ecosystem, with each species given an equal weighting; this is the simplest measure of biodiversity for that area.
Species turnover See TURNOVER.
Specific mate recognition system (SMRS) One of the fertilization and mate recognition systems in the recognition concept of species; i.e., the factors leading to premating compatibility within a species.
Specificity The ability to relate a chemical effect to a distinct biological function or organism.
Speciose Of high diversity; having many species. Contrasted with DEPAUPERATE.
Spermatophore In various species (e.g., insects), a chitinous container produced by the male to hold sperm. It may be attached to the substrate for the female to locate, or passed to the female from the male during mating.
Spicule In certain nematodes, a bladelike, sclerotized male copulatory organ, usually paired, located immediately dorsal to the cloaca.
Spillover The transmission of a pathogen from reservoir hosts to other species that generally do not sustain the pathogen.
Spinneret One of usually three, rarely four pairs of modified terminal abdominal appendages in spiders, bearing one to hundreds of hollow spigots from which silk is drawn (spun).
Spore A walled, typically unicellular reproductive structure found in many organisms, adapted for dormant survival in unfavorable environmental conditions, sometimes over vast time periods. Capable of developing into a new individual without fusion with another sexual body if conditions are favorable.
Sporophyte In an organism with an alternation of generations, the diploid, spore-producing phase.
Sprawl inequality The concept that the majority of habitat loss from urban expansion is due to a minority of new housing unit creation, typically lower density sections at the edges of the urban area. Similarly, a minority of households live on land parcels that cover the majority of the urban area.
Spreading rate The rate at which tectonic plates move apart from each other.
SSU rDNA The gene in DNA that codes for the ribosomal RNA in the small subunit (SSU) of the ribosome. This is 16 S rRNA in prokaryotes and 18 S rRNA in eukaryotes.
Stabilizing mechanism A biotic or abiotic force that introduces a tendency for species to increase when rare, giving rise to niches and possible stable coexistence of multiple species in a community.
Stabilizing selection Selection that favors intermediate phenotypes (hybrids between incipient species). Compare DISRUPTIVE SELECTION.
Stage The lowest-ranking unit in the hierarchy of terms for stratigraphic intervals that can be recognized on a global scale.
Stasis 1. A state of stability or equilibrium; lack of change. 2. Specifically, the persistence of organisms that greatly resemble their fossil forebears, through long geological periods of time. Also, **stasigenesis**.
Stem A term for the branch on a cladogram between node and term.
Stem parasite A parasitic plant that forms haustoria on the stems of a host.
Stenotopic Occupying a highly specialized habitat, or tolerating a narrow range of environmental conditions.
Stewardship 1. The management of a thing for someone or something else. 2. Specifically, the ethical concept that humans have the responsibility to serve as a **steward**, or watchful caretaker, with respect to the natural environment, acting to protect natural resources and preserve biological diversity.
Stichosome A longitudinal series of cells (**stichocytes**) that form the posterior esophageal glands in the worm genus *Trichuris*, which includes several species that are intestinal parasites in mammals.
Stochastic Having to do with or being the result of chance and random events or processes.
Stochastic event (process) An event (process) that occurs by chance; a random phenomenon such as a lightning strike.
Stochasticity The fact of occurring by chance; the condition of a system whose future cannot be precisely predicted from its

current status, even with a full knowledge of all the factors involved (due to random forces inherent in the system).
Stochastic system A dynamical system whose detailed future behavior cannot be accurately predicted, even in principle.
Stoichiometry The ratios of different elements in a material. For example, a plant tissue requires not only carbon, but also nitrogen, phosphorus, and other elements in certain **stoichiometric** proportions.
Stoichiotrophic imbalance A unique, forced trophic state in an aquatic ecosystem that develops when one nutrient, generally nitrogen (N) or phosphorus (P), but not both, is altered, either due to enrichment or management-related nutrient control. It can occur throughout the eutrophication process, depending on the enrichment sources or timing of N or P controls.
Stoma 1. The singular form of STOMATA (see below). 2. A natural opening in the body, such as the mouth. **3**. An opening in the body that is artificially created for medical purposes.
Stomata (singular **stoma**) **1**. An array of minute pores in the surface of a leaf, surrounded by specialized guard cells that regulate the degree to which the pores are open or closed. They are the medium for gas exchange across the leaf surface, as carbon dioxide and oxygen enter and water vapor and oxygen exit. 2. The entire complex including both these pores and the accompanying guard cells. Also, **stomates**.
Stomatal or **stomal** Having to do with or being part of the stomata (see previous).
Stomatal index The percentage of stomatal openings compared with epidermal cells; used with fossil plant cuticles as a proxy for past atmospheric CO_2 levels.
Storage effect A general mechanism for the maintenance of biodiversity within a single habitat, based on differences between competing species in their responses to environmental conditions. The storage effect can allow co-existence by different species, if each species reproduces successfully under the conditions favorable for it and can survive through unfavorable periods in the resistant stage (e.g., when the competing species dominates).
Strain A population of cells all of which are descended from a single cell.
Stramenopile Another term for a HETEROKONT.
Strata The plural form of STRATUM (see below).
Stratification The formation of distinct layers with different densities, which inhibits mixing. Can apply to layers of rock, levels of the ocean, or divisions of the atmosphere.
Stratigraphic Of or having to do with stratigraphy (see below).
Stratigraphic interval All (and only those) rocks formed during a specific interval of geologic time.
Stratigraphic range The strata through which a taxon is known to have been extant.
Stratigraphic section 1. An outcrop of rock (with fossils) or a drilled core. 2. A composite for a region in which fossil ranges and stratigraphic events have been summarized into a synthesized rock column.
Stratigraphy The scientific study of rock strata, especially the distribution, deposition, and age of sedimentary rocks.
Stratum (plural **strata**) **1**. A unit of sedimentary rock or soil that is distinguishable from other units on the basis of properties and/or relative position. 2. Any particular layer or level in a horizontal system, such as a hierarchical society.
Strepsirrhini One of the two recognized primate suborders, along with the Haplorrhini (see); it includes the galagos, lorises, and lemurs.
Stridulatory (stridulation) apparatus An organ of sound production based on the mechanism of rubbing one part of the body against another.
Strip intercropping See INTERCROPPING.
Strong interactor A species that has a major effect on the species (one or a few) with which it interacts. Communities and ecosystems may have various strong interactors, and such species may occur at all trophic levels. Some, but not all strong interactors are keystone species.
Structural connectivity The distance between habitat patches and the nature of the intervening landscape matrix.
Structural intricacy The complex of patterns of links or connections among species in a community that are created by biotic interactions, shared pathways of nutrient and energy flow, or phylogenetic relationships.
Stygobiont A member of one of the aquatic species that are obligate subterranean dwellers.
Stygophile A member of a species that can live and reproduce in both subterranean and surface environments.
Subarctic The ecotone (ecological boundary zone) of the Northern Hemisphere connecting the treeless tundras to the north with the taiga or coniferous boreal forests to the south. The area is characterized by the presence of scattered, deciduous or coniferous trees of low stature and is sometimes termed the **forest tundra**.
Subchelate Describing organisms in which the chela (claw-like appendage) terminal segment only folds back on the next segment as opposed to extending.
Sublittoral Another term for NERITIC.
Sub-nivean Describing an environment beneath a temporary, seasonal, or permanent covering of snow.
Subordinate Describing species that have a minor influence on ecosystem structure and function, presumably because of their lesser abundance and biomass compared to dominant species.
Subpolar Relating to or located in the regions between the polar and temperate zones.
Subshrub Another term for a SEMISHRUB.
Subsurface depletion A condition in which metabolic breakdown of organic matter from surface waters can cause depletion of dissolved oxygen concentration, leading to very low oxygen (hypoxia) or no oxygen (anoxia).
Subtropical 1. Located in the regions lying either above the Tropic of Cancer (23.5° N) or below the Tropic of Capricorn (23.5° S). 2. More generally, involving the climates and landscapes characteristic of this region. The **Subtropics** experience mild winters with little frost or snow and have a wide variety of rainfall patterns.
Succession The pattern of change in species that occupy an area over time, caused by biotic or abiotic factors that benefit some species at the expense of others.
Succulent Having fleshy tissues that store moisture.
Sulfur bacteria Bacteria that depend on the chemotrophic or phototrophic oxidation of reduced sulfur compounds for energy generation.

Sundaland The Malay Peninsula, along with Sumatra, Java, Borneo, and associated islands.
Supercontinent 1. The amalgamation of various continental masses into a single mass through continental drift. 2. Specifically, such a land mass of the geologic past; e.g., the supercontinent of Pangea (from 300 to approximately 190 million years ago) included most continental areas other than those that comprise East Asia.
Supplementation The purposeful introduction of additional organisms to an existing population to increase mean fitness or population size.
Supply chain Those stages in the production and use of a given product through which materials and energy flow can be traced.
Support(ing) services Processes that support the functioning of an ecosystem so that it can supply benefit to humans, such as soil formation or nutrient cycling. See ECOSYSTEM SERVICES.
Supralittoral 1. The portion of beaches, rocky shores, and other marine habitats that typically lie above the high tide. 2. Having to do with or occurring in this area.
Susceptible Describing a potential host for a pathogen that is neither currently infected nor immune to infection.
Suspension feeder An organism that feeds primarily on particles of organic material suspended in the water above the bottom.
Sustainability The fact of being sustainable; the ability of an individual or collective human activity to continue indefinitely without depleting the resources consumed in it, degrading the ecosystem providing these resources, or otherwise disturbing the balance between human needs and the preservation of the natural environment.
Sustainable 1. Able to endure; continuing to function over time. 2. Describing a natural system that is able to remain diverse and productive over time, based on current conditions. 3. Describing practices and policies in dealing with natural systems that are intended to ensure their viability over time; e.g., the sustainable use of a forest or fish stock.
Sustainable agriculture Agricultural practices that are environmentally sound, culturally sensitive, socially acceptable, and economically viable.
Sustainable development The process of meeting human needs of the present without damaging the ecosystems that produce these resources or compromising the ability of future generations to meet their own needs.
Sustainable forestry The management and utilization of forest resources in a way that balances human needs with undisturbed functioning of the forest ecosystem, and a long-term maintenance and sustainability of the resources and biodiversity of forests.
Sustainable use The exploitation of wildlife in a manner that avoids depletion of the resource (e.g., limited hunting, ecotourism).
Sustained yield The management and harvesting of renewable resources in a manner that does not exceed their rate of replacement or reproduction; e.g., the removal of a constant number of animals from a wild population that can be maintained through time while keeping the population at a predetermined size.
Swamp 1. A wetland dominated by woody species. 2. Any low-lying area without cultivated vegetation where water collects.
Swiddening A term for diverse types of horticulture in which a small section of forest is cut and burned to plant crops in a temporary garden.
Symbiont One of a pair of coexisting, mutually supportive organisms.
Symbiosis 1. In the strict sense, a relationship between two organisms of different species that benefits both partners; i.e., another term for MUTUALISM. 2. In the broader sense, any intimate, protracted association between organisms of different species (parasitism, commensalism, mutualism, or phoresis).
Symbiotic Having to do with or taking part in symbiosis.
Symmetry The correspondence in size, form, and arrangement of parts on opposite sides of a two- or three-dimensional object. **Bilateral symmetry** is a body plan in which similar anatomical parts (e.g., the eyes) are arranged on opposite sides of a median axis, so that the individual in effect has two matching halves.
Sympatric Describing populations, species, or taxa that occur together in the same geographical area within the dispersal range of one another.
Sympatric speciation The process in which genetic divergence between populations occupying the same geographic range leads to distinct species.
Sympatry The fact of populations, species, or taxa occurring together in the same geographical area.
Symplesiomorphy Shared, primitive similarity; a character present in all subordinate taxa of a group under consideration.
Synamorphic Having shared, derived similarity.
Synanthrope A species that lives near humans, and benefits from the types of habitats that humans create; the term does not apply to pets or farm animals. Many synanthropes are regarded as human pests, such as rats or crows. An organism that lives in this manner is **synanthropic**.
Synanthropy The fact of an organism being ecologically associated with human habitats on an ongoing basis.
Synapomorphy A character state that is derived relative to the ancestral state (plesiomorphy). Synapomorphies are special cases of homology that are evidence of monophyly among the taxa that bear them relative to taxa that bear the plesiomorphic state.
Synaptonemal complex (SC) A three-part proteinaceous structure that affects the number and distribution of chromosomal crossovers and converts these crossovers into functional chiasmata (the point of contact between paired chromatids).
Synchrony The temporal matching of periodic biological events; e.g., coordination of adult activity in an insect predator with the flowering of the plant on which it feeds. Such matching is **synchronous**.
Synergism An interaction among many factors which collectively may have a much larger (or smaller) effect than the sum of the effects of each factor acting independently. Such an interaction is **synergistic**.
Synergist A chemical used at sub-lethal concentrations to inhibit particular groups of detoxifying enzymes and therefore to implicate these enzymes in resistance.

Synlophe In numerous Trichostrongylidae (nematode worms), an enlarged longitudinal or oblique cuticular ridge on the body surface that serves to hold the nematodes in place on the gut wall.

Synonymous mutation A nucleotide change in a protein coding region that does not change the encoded amino acid.

Syntaxon (plural **syntaxa**) A category of vegetation typology based on the floristic–sociological approach (also known as the **Braun-Blanquet approach** or **method**); the basic syntaxon rank is association, which further groups into alliances, which group into orders, and which group into classes.

Systematics The area of biology that focuses on the classification of living things based on their relationships and evolutionary history.

System functions Processes that occur in an ecosystem that are typically measured as changes in the carbon or nutrient cycles regulated by decomposers, consumers, or primary producers.

Systems theory A framework for understanding the dynamics of interconnected components; can be applied to the study of complex biological systems.

Taiga 1. A Russian term applied to Eurasian conifer forests described as damp and almost impenetrable, or to a coniferous forest with no non-coniferous tree species except *Betula* and *Populus*. 2. More generally, another term for a BOREAL FOREST.

Targeted conservation Strategic land allocation and application of management practices, such as buffers, corridors, constructed wetlands, stepping stones, and reserves, to maintain species populations and ecosystem services in concert with human land use.

Target taxon The species or species group that a conservation action is focused upon.

Taungya A Burmese term for an agricultural system in which crops are interplanted with plantations of trees. As the trees grow and shade the areas, cultivation of crops is abandoned.

Taxon (plural **taxa**) A general term used to indicate groups of related organisms at any level in a taxonomic hierarchy. In the classification of evolutionary lineages, the term can refer to species, genera, families, orders, and so on. Higher taxa are those above the species level.

Taxon cycle A hypothesis proposed by E. O. Wilson to describe the common niche expansion among colonizers, and subsequent long term trend toward differentiation, coupled with isolated species narrowing their niches and restricting their ranges, resulting in eventual extinction.

Taxonomic Having to do with or based on taxonomy (see below).

Taxonomic (bio)diversity A characterization of a biota that relates to variation in the importance and taxonomic identities of taxa, estimated by components such as species richness, evenness, dominance, diversity, and rarity.

Taxonomic impediment A term for the idea that currently there is not a sufficient amount of taxonomic expertise and tools to identify and describe the world's biodiversity.

Taxonomic inflation The process by which the numbers of species in the checklist of a group increases due to a change in species concept rather than due to new discoveries of previously unknown species.

Taxonomic rank The level in the Linnaean hierarchy to which a particular taxon or group of species is assigned. For example, mammals are at the level of the class (Mammalia) and hippopotamuses are at the level of family (Hippopotamidae).

Taxonomic resolution The level of genetic variation of the taxa being considered.

Taxonomic richness The total number of species, genera, and families from a given time period or fossil excavation.

Taxonomy The classification of species and higher taxa, ultimately based on phylogeny.

Taxon sorting The argument that in certain taxa, species undergo relatively rapid speciation and extinction, while in others the species speciate and die out relatively slowly; over time this difference will mean that the highly volatile families become extinct more quickly than the less volatile ones.

Techno-ecosystem Technology-based ecosystem; a system in the contemporary world that is fundamentally distinct from natural ecosystems in features such as the use of energy sources other than sunlight (e.g., fossil fuels, nuclear power), an urbanized concentration of human population, and the generation of substantial amounts of air and water pollutants and waste materials.

Tectonic Having to do with the earth's crust, or with forces or conditions that cause the movement of this crust. See PLATE TECTONICS.

Tectonism The action of forces that disturb and dislocate the Earth's crust; earthquakes are part of this process.

Tegmina (singular **tegmen**) Thickened forewings.

TEK Traditional ecological knowledge.

Teleconnection A causal link between patterns of weather in two locations, or between two atmospheric occurrences, which are very far apart, such the El Niño Southern Oscillation, where unusually high sea surface temperatures off the west coast of South America can be connected with droughts in Indonesia.

Telson The most posterior segment of the body in crustaceans, usually flattened and often armored.

Temperate Pertaining to the climates and landscapes of the mid-latitude regions, which are seasonally warmer and cooler, with winter frost even in most coastal areas and at least partial dormancy or collapse of the vegetation.

Temperate region Any locality with at least one month of frost (for continental areas) or with one or more months with a mean temperature lower than 18° C (for maritime-influenced areas), and with at least four months with a mean temperature higher than 10° C.

Temperate virus A virus that has a high probability of integrating into the host genome rather than initiating a lytic infection.

Temporal Having to do with time; measured by or in the context of time.

Temporal dispersal The emergence of individuals from dormancy over a range of years (or another time interval), when those individuals entered dormancy in a single year. Often the years (or other time intervals) have different environmental qualities for growth and reproduction.

Temporal heterogeneity In community sampling, a variation in species capture probabilities over time; this may

arise from temporal environmental variation, migration, speciation, or other factors.

Temporal migration The avoidance of harsh environmental conditions by an individual organism that enters dormancy before conditions become harsh and emerges from dormancy when favorable conditions return.

Temporary parasite A parasite found in or on a host only for short periods.

Tepui A Venezuelan term for the sandstone table mountains of the Guayana Highland or "Lost World" region.

Terra firma Solid ground; the dry land mass of the Earth's surface.

Terra firme Non-flooded landscape in tropical lowlands, usually characterized by highly weathered nutrient-poor soils.

Terra firme forest Continuous hardwood forest of the nonflooded or upland parts of the Amazon rainforest.

Terrestrial 1. A species that lives on land (as opposed to marine or freshwater settings) for the majority of its life cycle. 2. Describing or being such a species.

Terricolous Living in or on the ground; being a soil inhabitant.

Tertiary The earlier part of the Cenozoic Era, occurring from about 65 million to 2 million years ago.

Tetrapod One of the Tetrapoda, the terrestrial vertebrate classes of amphibians, reptiles, birds, and mammals; so named because they all (at least primitively) possess four legs.

Thelytoky A reproductive mode in which unfertilized eggs develop into diploid females. An organism reproducing in this manner is **thelytokous**.

Therapod One of the Therapoda, the so-called beast-footed dinosaurs, a diverse group of bipedal animals including huge carnivores such as *Tyrannosaurus rex* and also small nonflying animals that are the ancestors of modern birds.

Therapsid One of the Therapsida, an order of reptiles that existed during the Late Paleozoic and Early Mesozoic, from which mammals are believed to have evolved.

Thermally tolerant Describing organisms able to live at temperatures significantly higher than the regional norm. They can also exist at cooler temperatures, but may not compete well in the latter situation.

Thermal opportunist An organism that has developed life history characteristics that permit the exploitation of temporary cooler spots in thermal systems.

Thermal system An outflow of geothermally heated water, especially one at the surface of the earth where light is present.

Thermal vent Geothermally heated water issuing from cracks in the ocean floor. Very high temperatures and pressures and the complete absence of light are characteristic of these vents.

Thermophile Heat lover; an organism that grows optimally at temperatures between 50° and 80^0 C. Such an organism is **thermophilic**. Compare HYPERTHERMOPHILE.

Thermophily The fact of being adapted to live at high temperatures and (usually) unable to survive at lower temperatures.

Thoracopod An appendage on the thorax (paired in crustaceans).

Threatened Facing a relatively high risk of extinction in a definable time period. In ranking systems for extinction risk, it is typically a category less critical than ENDANGERED.

Threatened species A species that faces a significant risk of extinction in the near future over all or a significant portion of its range.

Three-taxon statement The idea that two species (or taxa) are more closely related to one another than either are to a third; the most basic hypothesis in cladistics.

Threshold theorem In epidemiology, the principle that if the density of susceptible individuals in a population falls below some critical threshold, then a disease may not be able to maintain itself in the population.

Timber line The boundary of closed, tall forest with timber-size stems; it commonly occurs at an elevation that is 50–150 m lower than the tree line.

Tipping point The point at which a complex system goes into disequilibrium and is replaced by a new system; e.g., the transitional point separating two alternate regimes or configurations in the structure and function of an ecosystem.

Token cue An environmental cue that induces a state of dormancy in certain animals, such as reduced day length. The cue serves as a reliable indicator of a coming onset of harsh environmental conditions, but is not in itself harsh.

Tokogeny Birth relationships responsible for reticulate patterns of relationships within and among populations.

Tolerance An individual's capacity to sustain damage by an enemy with limited decrease in fitness.

Top-down 1. Describing a process in which feeding by higher trophic-level organisms has cascading effects on lower trophic levels; e.g., when fish (predators) consume zooplankton (herbivores) and, by lowering the numbers of these herbivores, allow increased densities of photosynthetic phytoplankton (primary producers). 2. Describing impacts or controls on urban biodiversity that originate from human actions at larger scales, as at the level of a city or above. Thus, **top-down effect** or **force**. Compare BOTTOM-UP.

Topkill Loss of the aerial biomass in trees or other plants as a result of fire, although the plant as a whole still survives and can resprout from root-stocks. Seedlings and saplings are particularly vulnerable to topkill.

Top predator See APEX PREDATOR.

Torpor A condition, usually in birds and mammals, when the body temperature falls a few degrees below the ambient temperature for a short period of time. It differs from HIBERNATION in the extent to which the body temperature falls and the length of the resting state.

Total allowable catch The level of permitted catch established by fishery regulations.

Tourism Travel for leisure, recreation, or vacation; for economic statistics, generally only overnight travel is included. See also ECOTOURISM.

Trace gas A gas present in the atmosphere in low mixing ratios (typically less than a few ppm). Some trace gases, including methane, nitrous oxide, and ozone, are greenhouse gases. Others are reactive in the atmosphere and influence the greenhouse effect indirectly. For example, carbon monoxide, produced by internal combustion engines and certain industrial processes, competes with methane for the available oxidizing capacity of the atmosphere, and so an increase in carbon monoxide emission leads indirectly to a higher methane concentration.

Tracheae 1. In terrestrial vertebrates, a tube connecting the pharynx or larynx to the lungs, allowing the passage of air. 2. In invertebrates, a system of hollow, branched or unbranched air-conducting tubes used for respiration, opening via abdominal spiracles. They may or may not extend into the cephalothorax or legs.

Tracheophyte A vascular plant; i.e., one having tissues that conduct water, minerals, and photosynthetic products through the plant.

Trade-off A negative correlation between traits, such that a benefit due to changes in the value of one trait is associated with a cost produced by changes in the value in another trait.

Traditional Following a pattern that has continued over time; according to beliefs, customs, practices, and so on that have persisted from generation to generation.

Traditional agriculture An indigenous form of ecologically based agriculture resulting from the coevolution of local cultural and environmental systems; generally characterized by energy sources such as human and animal power, hydropower, or wind as opposed to elaborate machinery, the use of biological fertilizers and pest controls as opposed to chemical ones, and crop variety as opposed to monoculture. Traditional agriculture has significantly less environmental impact than INDUSTRIAL AGRICULTURE (see).

Traditional ecological knowledge (TEK) A cumulative body of knowledge, practice, and belief, evolving by adaptive processes and handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and with their environment. TEK is applied by indigenous and other peoples in their interactions with their habitats and in their system of natural resource use and management. Also known as **traditional environmental knowledge.**

Traditional peoples Peoples with cultures that differ from those of the mainstream of a national population, typically having distinct languages and locally evolved resource management systems.

Traditional society A group in which knowledge, practice, and belief are handed down through generations largely by cultural transmission. The tradition itself evolves by adaptive processes, but not all tradition is necessarily adaptive.

Traditional variety See LANDRACE.

Tragedy of the commons A process in which an open access resource is overexploited because individual users cannot expect to realize the potential benefits of resource conservation for the community in general.

Trailing edge The dynamic margin of a species' range where it is becoming locally extinct, causing the range to retract.

Trait A quality or characteristic that an organism may possess through inheritance.

Transaction costs 1. The costs associated with negotiating an economic exchange. 2. Specifically for a terrestrial conservation effort, the costs in addition to the price of a transfer of property rights for a given parcel of land, such as the costs of property search, negotiation with individual landholders, and approval for title transfer. For example, carbon sequestration projects involving afforestation or reforestation can be beneficial for conservation, but high transaction costs often limit their viability.

Transgenic Having a gene or genes that were transferred from another organism (**transgenes**), most often by means of biotechnology.

Transient species Describing plant species that are present as scattered seedlings or small immature individuals; many of these species occur as dominant or subordinate species in neighboring vegetation associated with different environmental conditions or management regimes.

Translocation The intentional movement of organisms by people from one or more locations to other locations.

Transmission dynamics Changes in the rates of pathogen transmission through time.

Transovarial transmission The vertical transmission of a pathogen through its passage from mother to offspring.

Transpiration The process by which vascular plants take up water through the roots, transport it through the xylem, and evaporate it from the surface of the leaves.

Transport host A host that harbors sexually immature stages of a parasite that do not develop.

Transposon A segment within a chromosome that replicates itself by transposition to a new location within a host genome, or from a genome to other extrachromosomal DNA in the cell such as a plasmidorviral genome. Also known as a **transposable element**.

Tree A cladogram (diagram of organismal relatedness) or a cladogram interpreted in more evolutionary detail.

Tree line The high-elevation and/or high-latitude limit of forested landscape, usually fragmented at this point. Most often there is no literal "line" visible in the landscape, so tree line position represents a convention. Also spelled **treeline**.

Tree line ecotone The whole transitional area from the lower timber line (see) to the higher tree species line (see next).

Tree species line The highest point above the tree line at which individual trees or small groups of trees are found.

Triacylglycerol An ester of glycerol and three fatty acids. The commonest lipid storage molecule in animals and green plants.

Triage A system of prioritization for making decisions when resources are scarce and the need for responses is widespread. The *tri-* prefix denotes the concept that priorities can be separated into three groups: those situations where conditions do not require intervention; those where intervention is likely to be futile, and the intermediate group that could benefit from intervention (this last being the focus of prioritization). [Originally applied to a French system of battlefield medicine in World War I that divided casualties into those likely to live without immediate treatment, those likely to die even with treatment, and those who could survive if given immediate treatment.]

Trichobothria (singular **trichobothrium**) In arachnids and various insects, long, delicate, slender setae set in broad, shallow innervated sockets in the cuticle. Trichobothria are sensitive to vibration or near-field air movement and are a major sense organ of arachnids.

Triramous Relating to or having an extension of the body (usually an appendage) with three branches; e.g., triramous antennae have three main branches or flagellae.

Tritomy A condition of three undifferentiated branches or clades. Also, **trichotomy**.

Trochophore A type of larva characteristic especially of annelids and mollusks, swimming freely by means of a ciliated girdle around the middle.
Troglobiont Terrestrial species that are obligate subterranean dwellers; the term is sometimes used for aquatic species as well.
Troglobitic Describing cave-dwelling organisms that are adapted to life in complete darkness.
Troglobite An organism whose entire life cycle takes place in a cave or other subterranean environment. Contrasted with a TROGLOPHILE (see below).
Troglomorphic Pertaining to the morphological, behavioral, and physiological characters that are convergent in subterranean species.
Troglophile A cave lover; a terrestrial species that lives in a cave or other subterranean environment but that also can live and reproduce in surface environments; sometimes used for aquatic species as well. Such an organism is **troglophilic**.
Trophic Relating to feeding; describing types of food consumed and patterns of feeding behavior.
Trophic cascade The effect of predatory feeding on prey that propagates down more than one link in a food web. Trophic cascades cause inverse patterns in the abundance or biomass of trophic groups at successive levels of a food web.
Trophic complexity The number and types of organisms that feed at different trophic levels within a community.
Trophic downgrading Ecological and environmental consequences associated with losses of apex predators from ecosystems.
Trophic input Energy entering a community via pathways other than local primary production, and often at higher levels in the food chain. For example, nutrients from salmon carcasses are a periodic trophic input to the forest community adjacent to salmon streams.
Trophic interaction A feeding relationship between two species; these include predation, herbivory, parasitism, bacterivory, frugivory, or any other interaction that involves individuals of one species consuming individuals or parts of individuals from another species.
Trophic level An abstract classification to describe subsets of species that acquire energetic resources in a similar way on a subset of species (e.g., top carnivores feed on primary carnivores which feed on herbivores which feed on primary producers). In natural systems, most species do not feed strictly on the level below them, making the concept of trophic level a difficult term to assign operationally to species.
Trophic status A ranking system for aquatic ecosystems, based on the amount of organic production and nutrient (N, P) levels.
Trophic web Another term for a FOOD WEB.
Trophogenic Describing a region where net production of organic matter occurs by photoautotrophy or chemoautotrophy.
Tropholytic Describing a region where respiration and decomposition of organic matter occur in the absence of primary production.
Tropical 1. Located in the region lying between the Tropic of Cancer (23.5° N) and Tropic of Capricorn (23.5° S). 2. More generally, involving the climates and landscapes characteristic of this region. The **Tropics** are essentially frost free in the lowlands and permit biological activity as long as water is available.
Tropical dry forest A type of forest found in tropical to subtropical regions that has distinct rainy and dry seasons. Many tropical dry forest plants are adapted to withstand high temperatures and seasonal droughts.
Tropical rainforest A forest that occurs below 1000 m in elevation, experiences high, relatively constant temperatures, and receives at least 200 mm of rainfall per year.
TRY A global database of plant functional traits, maintained by the Max Planck Institute for Biogeochemistry in Jena, Germany and governed by an international steering committee.
Tuberculum interglenoideum An anterior projection of the first (cervical) vertebra in salamanders. It bears articular facets that insert into the foramen magnum of the skull and provide additional articulation points between the skull and the vertebral column.
Tundra The vegetation characteristically occupying the Arctic.
Turnover A measure of the amount by which the species composition changes among subunits in a data set.
Type The specimen, culture, or other element on which the scientific name of an organism is based and that fixes the name's application.
Type I community An assemblage of species in which the richness is a constant proportion of the number of species occurring in a larger geographic unit where the community is embedded.
Type II community An assemblage of species in which the richness is independent of the number of species in a larger geographic unit where the community is embedded.
Typification The practice of designating a single specimen (or, formerly, a series) as the name-bearer for a species; or a species for a higher taxon, especially genera.
Ultraviolet radiation Radiation with wavelengths shorter than violet light and with more energy.
Umbrella species A species with large habitat needs or other requirements whose conservation results in many other species also being conserved at the ecosystem or landscape level, thus it acts as an "umbrella" protecting these species.
Uncultured Describing prokaryotic organisms whose presence has been detected by molecular methods in the environment but that have not been cultured under artificial laboratory conditions.
Understory The broad spectrum of plants at the ground level of forests that often provide forage for grazing animals. These transitory plants are often relatively sparse in dense forests with closed canopies.
Underutilized Describing a species whose potential for contributing to food security, health (nutritional/medicinal), income generation, and environmental services has not been exploited.
Unguligrade 1. Walking only on a hoof or tip of the digit. 2. An organism that walks in this manner; i.e., the hoofed mammals such as deer, horses, and cattle. Compare DIGITIGRADE, PLANTIGRADE.
Uniramous Relating to or having an extension of the body (usually an appendage) with one branch.

Unity of type Similarities among organisms leading diverse ways of life under diverse conditions of existence, going beyond any functional need for similarity. Under the theory of descent with modification, these similarities result from common ancestry.

Unplanned biodiversity Soil biota, herbivores, carnivores, decomposers, and any other species that naturally exist in or colonize an agroecosystem, as opposed to crops and livestock purposefully introduced and maintained there by humans (planned biodiversity).

Uptake capacity See NUTRIENT UPTAKE CAPACITY.

Upwelling An ocean phenomenon in which cooler, more nutrient-rich water rises up from the bottom of the ocean to the surface. This occurs where currents of contrasting temperatures meet. The extra nutrient influx allows for specific and rich **upwelling ecosystems** in these areas (e.g., off the coasts of Chile and Peru, and southwest Africa).

Urban ecology The study of urban systems from an ecological perspective; an emerging field within ecology that strives to understand human interactions in ecological systems in and around urban areas and to develop theories and analyses that include human communities as fundamental components of ecological systems.

Urbanization 1. A process in which an area becomes more urban; i.e., increases in population, building density, and so on. 2. The relocation of people to large cities from rural areas and smaller towns and villages.

Urban sprawl Urban development that occurs at a relatively low population density and/or occupies a large amount of land.

Use values Values that are obtained by using a natural resource, such as timber, fuelwood, water, or landscapes.

Utilitarian Describing the school of conservation philosophy that emphasizes the value of the natural world in contributing to human well-being, especially in terms of economic standard of living.

Vapor pressure deficit The difference between the water vapor content of the air and the saturated water vapor content of the air at its current temperature; this is the driving force of transpiration.

Variety See CULTIVAR.

Varzea Brazilian term for the floodplain of the whitewater rivers of the Amazon basin.

Vector An animal, typically an arthropod, that transmits pathogens from one host to another; e.g., mosquitoes are vectors of malaria.

Vegetation The total plant cover of an area.

Vegetation megazone A large-scale vegetation complex characteristic of a zonobiome.

Vegetative Describing a form of asexual reproduction in plants; a process by which new individuals arise without production of seeds or spores. Thus, **vegetative reproduction, propagation, multiplication**, and so on.

Ventral The part of an organism that is the "front" or lower part of the body. Contrasted with DORSAL which is the "back" or upper part.

Vermiform Worm-shaped, with a tapering form at both the posterior and anterior.

Vertical transmission The transmission of pathogens across generations of hosts.

Vestigial Describing an organ or structure that is present only in rudimentary form (probably lost from a previous condition).

Viable 1. Having life; able to live and grow. 2. Specifically, describing a population deemed to have a reasonable chance of long-term persistence; the sense of what is a reasonable chance of survival and what constitutes a long term can vary according to the defining entity. The fact of being viable is **viability**.

Vibratory papillae In certain insects, mobile, grooved, rod-like appendages arising from the distal edge of the first thoracic segment, used for communicating.

Vicariance The existence of closely related taxa or biota in separate, disjunct areas, which have been separated by the formation of a natural barrier (vicariance event; see next).

Vicariance event A historical event that isolates species' populations and leads to speciation (e.g., the closing of the Isthmus of Panama); usually considered in a geological context, but human activities also generate such events. Also, **vicariant event**.

Virad A suffix denoting viruses that belong to a certain order (e.g., viruses in the order Picornavirales are picornavirads).

Viral Having to do with or caused by a virus.

Viremia A concentration of viruses in the blood of a host.

Virid A suffix denoting viruses that belong to a certain family; e.g., viruses belonging to the family Phycodnaviridae are picornavirids.

Virion An individual viral particle.

Virulence The degree to which a parasite reduces the probability of survival or the reproductive capacity of its host.

Virulent Capable, or highly capable, of causing disease.

Virus A genetic element containing either DNA or ribonucleic acid (RNA) that can replicate only in cells but is characterized by having an extracellular state; an infectious agent causing disease in all forms of organisms. 2. A type of disease or infection cause by such an agent; e.g., Ebola virus, West Nile virus.

Viviparous 1. Describing reproduction in which the young are born alive. 2. Describing an organism with this mode of reproduction; e.g., most mammals. The fact of reproducing in this way is VIVIPARITY. Compare OVIPAROUS; OVOVIVIPAROUS.

VOC Volatile organic compound; many VOCs are emitted by plant leaves, of which the most abundant is the unsaturated hydrocarbon isoprene. VOCs have various functions in plants, including protection against heat stress damage and damage by insect herbivores.

Volunteer (vegetation) A plant or plants growing on its own instead of being deliberately planted by a farmer; usually applied to a desirable plant rather than a weed.

Voucher A representative specimen of a plant or an animal that is properly collected, prepared, and preserved in a herbarium or museum to facilitate identification of the species.

Vulnerable Facing a high risk of extinction in the wild.

Wadi A gully, stream bed, or other water course that is dry except during a rainy season; used especially of such features in northern Africa and southwest Asia.

Wallacea Lands east and north of Borneo, west of New Guinea and east of Wallace's Line (see next), including the Philippines, Sulawesi, Lombok, Sumbawa, Sumba, Timor, the

Moluccas (e.g., Halmahera and Seram), Flores, and the other islands of Nusa Tenggara.

Wallace's (Wallace) Line A hypothetical line marking the division between the Oriental and Australasian zoogeographic regions. [First depicted by the English naturalist Alfred Russel Wallace.]

Water-holding The ability of soil to retain water; e.g., sandy soils have very low water-holding capacity, while clay soils have higher capacity.

Water pollution Lowering of water quality due to input of sewage, pesticides, agricultural runoff, and industrial wastes that can harm aquatic plants and animals.

Water quality standards Risk-based requirements that set site-specific allowable pollutant levels for individual water bodies, such as rivers, lakes, streams, and wetlands. These standards are set by designating uses for the water body (e.g., recreation, water supply, aquatic life, agriculture) and applying certain criteria to protect the designated uses.

Watershed The drainage basin contributing water, organic matter, dissolved nutrients, and sediments to a stream or a lake.

Water-use designation The appropriate water uses specified for a given body of water, such as public water supply, protection of fish, shellfish, and wildlife, or recreational, agricultural, industrial, or navigational purposes.

Water-use efficiency (WUE) a measure of the amount of photosynthesis or growth per unit of water lost in transpiration.

Weak interactor A species having little effect on other species, at least under average conditions. Under some circumstances, weak interactors may occupy important roles in ecological communities as a result of changes that lead to temporary increases in their abundance, size, or biomass.

Weed An unwanted plant that interferes with the growth of desirable plants, especially one that is unusually persistent and harmful.

Weed control The process of reducing or suppressing weeds in a defined area to an economically acceptable level, without necessarily eliminating them.

Weevil A member of the beetle family Curculionidae.

Wetland A terrestrial ecosystem with a water table that is permanently close to the surface.

Wetland bank A site where wetlands and other aquatic resources are restored, created, enhanced, or preserved.

Wetland creation The conversion of a persistent upland or shallow water area into a wetland by human activity.

Wetland restoration See RESTORATION.

White-water A term for Amazonian rivers that are muddy with much suspended sediment and are neutral or only slightly acidic.

Wild Not domesticated; describing those species that have not been brought under human management.

Wilderness A large area that remains essentially unmanaged and unmodified by human beings. Also called a **wildland**.

Wildland fuel Any live or dead biomass available for burning.

Wildlife All nonhuman and nondomesticated animals.

Wildlife-friendly Describing farm management practices as benign to wildlife as possible; can include reduction in pesticides and chemical fertilizers and retention of patches of non-farmed habitats and less-intensively farmed semi-natural habitats.

Wildwood Wholly natural forest not affected by human activities such as agriculture and pasturage, widespread in prehistoric times (including earlier interglacial periods) and still surviving in remote places.

Willingness to pay The highest payment a person would make in return for a good or service.

Woodland An ecosystem in which the mature overstory tree or tall shrub cover makes up a high percentage of the area. Typically, a well-developed herbaceous understory is present. The mature tree cover is usually greater than 60 percent.

Wood-pasture A human-managed ecosystem combining trees for fuelwood and timber with grazing land for domestic animals.

Worker A female in a social insect colony that is not able to reproduce; so called because they perform "helping" tasks in the colony, such as foraging for food or caring for larvae.

WUE Water-use efficiency.

Xenobiotic 1. A toxic substance not naturally produced within organisms. 2. Having to do with or being such a substance.

Xeric Describing a dry environment.

Xeromorphic Describing a form that has developed in response to highly arid conditions.

Xylem Water-conducting tissue in the stem or transport roots of vascular plants.

Zero-growth isocline In a graph with axes describing factors important to population growth (e.g., resource abundance), a line along which a population's growth rate is zero.

Zonation The tendency of climate, soil, and natural landscape types to occur in distinct latitudinal zones (e.g., tropical, subtropical, and temperate), as influenced by the global atmospheric circulation system.

Zonobiome A broad ecological topographical unit characterized by a certain climatic pattern.

Zoogeographical region or province One of the major geographical divisions of the world, identified for the study of the distribution of land mammals. They are the Palearctic (Europe, North Africa, Asia north of the Himalayas), the Nearctic (North America), the Neotropical (South America), the Afrotropical (Africa south of the Sahara), the Oriental (Asia south of the Himalayas), and the Australasian (Australia, New Zealand and nearby islands). The latter two feature a zone of overlap known as Wallacea. Antarctica constitutes an additional region.

Zoogeography The study of the geographical distribution of animal life on Earth.

Zoonosis A disease of humans caused by pathogens that proliferate within and are transmitted from nonhuman vertebrate animals; e.g., plague, Lyme disease, West Nile virus. Such a disease is said to be **zoonotic**.

Zooplankton Those animals that are found passively drifting or weakly swimming in the water column. They can be divided into two major categories: **holoplankton**, which spend their entire lives as plankton, and **meroplankton**, which spend part of their life cycle as plankton and part on the

sea floor as benthic invertebrate larvae or asnekton (e.g., fish larvae).

Zooprophylaxis The use of wild or domestic animals to divert potential pathogen vectors from feeding on humans.

Zooxanthellae A diverse group of symbiotic, photosynthetic dinoflagellate algae associated with corals and other tropical marine invertebrates. Such organisms are **zooxanthellates**.

Index

Notes
Cross-references terms in *italics* are general cross-references, or refer to subentry terms within the main entry (the main entry is not repeated to save space). Readers are also advised to refer to the end of each article for additional cross-references – not all of these cross-references have been included in the index cross-references.

The index is arranged in set-put style with a maximum of four levels of heading. Major discussion of a subject is indicated by bold page numbers. Page numbers suffixed by *T* and *F* refer to Tables and Figures respectively. *vs.* indicates a comparison.

Where index subentries and sub-subentries pertaining to a subject have the same page number, they have been listed to indicate the comprehensives of the text.

A

aardvark **1**:608, **4**:690*T*, **4**:695
abalone **1**:190*T*, **4**:345–346
abamectin **5**:694
abandoned lands **1**:543
Abantiades hyalinatus **5**:390
Abbey, Edward **4**:653
Abbott's gray gibbon **6**:256–258*T*
Abel's classification system **2**:51*F*
Abies
  Asia **1**:272–273
  boreal forests **1**:628–630
  geographic boundaries **1**:629*F*, **1**:630*T*
  North America **7**:118
Abies alba **2**:552*F*
Abies amabilis **3**:639–640
Abies balsamea **1**:628–629, **1**:630*T*, **1**:631*T*, **1**:632*T*, **3**:440
Abies bifolia **1**:628–629, **1**:631*T*, **1**:632*T*
Abies delavayi **1**:272–273
Abies fabri **1**:272–273
Abies faxoniana **1**:272–273
Abies firma **1**:271
Abies fraseri **4**:361
Abies georgei **1**:272–273
Abies guatemalensis **3**:208
Abies homolepsis **1**:273
Abies lasiocarpa **1**:612, **1**:631*T*, **1**:632*T*, **2**:571
Abies mariesii **1**:273
Abies nephrolepis **1**:273–274, **1**:630*T*
Abies procera **3**:639–640, **3**:639*F*
Abies sachalinensis **1**:273–274, **1**:273
Abies sibirica **1**:273–274, **1**:630*T*
Abies squamata **1**:272–273
Abies veitchii **1**:273
Abies webbia **1**:272–273
abiu **4**:284*F*, **4**:285, **4**:287–290
above-ground–below-ground species interactions **2**:386
above-ground species interactions
  canopy heat capture **2**:387, **2**:387*F*
  characteristics **2**:386–387
  light interception/light use efficiency **2**:386–387
  shading effects **2**:387, **2**:387*F*
  spatiotemporal complementarity **2**:387–388
Abramis brama **4**:463
Abraxas grossulariata **6**:3
abyssal plains **2**:17, **5**:108*T*, **5**:117*F*
Abyssobrotula galatheae **3**:466
Acacia **1**:69, **1**:267, **1**:291–292, **2**:146, **3**:130*T*, **5**:221, **5**:489
Acacia aneura **1**:292, **2**:468
Acacia caven **5**:216
Acacia cornigera **2**:428
Acacia cyclops **6**:96–97
Acacia drepanolobium **1**:708, **6**:46–47
Acacia erubescens **1**:67
Acacia hebclada **1**:67
Acacia karroo **1**:63
Acacia kirkii **1**:67
Acacia koa **2**:655, **6**:171
Acacia laeta **1**:67
Acacia macracanta **3**:115–116
Acacia mangium **5**:605
Acacia newbrownii **1**:67
Acacia nilotica **1**:67
Acacia pachyceras **5**:484–485
Acacia raddiana **5**:484–485, **5**:486
Acacia saligna **6**:94*F*
Acacia seyal **7**:387
Acacia tortilis **1**:41, **1**:52, **1**:67, **2**:469–470, **5**:235, **5**:486
Academy of Natural Science (Philadelphia) **5**:405–406
Acaena antarctica **4**:107
Acaena magellanica **4**:107
Acaena splendens **4**:106
açaí **4**:286–287, **4**:290
Acanitholimon **1**:270
Acanthaceae **2**:374*T*, **3**:129*T*
Acanthamoeba **5**:573–574*T*, **6**:290–291
Acanthamoeba polyphaga mimivirus **5**:130*F*, **5**:135
acantharians **2**:21*F*, **6**:292
Acanthiza katherina **2**:75*F*
Acanthobothrium **7**:456–457*F*
Acanthobrama telavivensis **5**:473
Acanthocephala **2**:524*T*, **2**:645, **3**:684*F*, **5**:669–670*T*
Acanthodactylus **5**:467–469, **5**:468*F*
Acanthodactylus scutellatus **5**:472
Acanthogobius flavimanus **6**:457
Acanthomysis **5**:61
Acanthopagrus australis **5**:667*F*
Acanthophrynus coronatus **1**:211
Acanthopteroctetidae **5**:385–386
Acanthopteroctetoidea **5**:386*F*, **5**:387–388*T*
Acanthopterygii
  Atherinomorpha **3**:467
  Mugilomorpha **3**:467
  Percomorpha **3**:467–471
Acanthosomatidae **7**:294
Acanthotermes **4**:409*T*
Acanthotermes acanthothorax **4**:429*F*
Acanthuroidei **3**:470
Acanthus **1**:277, **5**:11*T*
Acari
  Amazon ecosystems **1**:246*T*
  characteristics **1**:213–215, **1**:214*F*, **4**:394*F*, **6**:587
  ecological characteristics **1**:214–215
  freshwater ecosystems **4**:374
  genera and species **1**:202*T*
  lake and pond ecosystems **4**:461*T*
  latitudinal range shifts **4**:604*T*
  major lineages
    Acariformes **1**:216–217
    Astigmata **1**:217, **6**:587–588
    characteristics **1**:216
    Holothyrida **1**:216
    Ixodida **1**:216
    Mesostigmata **1**:216, **6**:587–588
    Opilioacariformes **1**:216
    Oribatida **1**:217, **6**:587–588
    Parasitiformes **1**:216
    Prostigmata **1**:216–217, **6**:587–588
  number of species **4**:371*T*
  phylogenetic relationships **1**:205*F*
  phylogeny and taxonomy **1**:215–216
  reproduction and dispersal **1**:215
  soil systems **6**:587
Acariformes **1**:216–217
Acarina **1**:184*T*, **4**:298–299
Acartophthalmidae **3**:489–491*T*
Access to Biological Collection Data (ABCD) standards **1**:440–441, **1**:443
Accipiter **2**:141
Accipiter nisus **5**:474
Accipitridae **6**:308*T*
Account of the Foxglove and Some of Its Medical Uses: With Practical Remarks on Dropsy and Other Diseases **5**:706
acephate **5**:693–694
Acer **7**:118

Ceratherium **5**:224*T*
Acer japonicum **1**:271
Acer mono **1**:271
Aceros waldeni **7**:185
Acer palmatum **1**:271
Acer pensylvanicum **6**:433*F*
Acer rubrum **6**:115*T*, **6**:433*F*
Acer saccharum **6**:433*F*, **6**:459, **7**:125
Acer tetramerum **1**:272–273
acetic acid **5**:696
acetylcholinesterase (AChE) **4**:300–301, **4**:300*F*
Achanthes exigua **4**:114–115
Achatina fulica **1**:82, **5**:382, **6**:455
Achelousaurus **2**:519*T*
Achillea fragrantissima **5**:484–485
Acidianus **1**:219–220, **1**:221*T*
acidic deposition **1:1–14**
  biodiversity impacts **1**:144–145, **5**:522, **5**:527*F*
  emission occurrences and sources **1**:1–5, **1**:137*T*
  forest ecosystems
    environmental impacts **1**:5, **1**:5*T*
    soils
      alteration consequences **1**:5
      aluminum mobilization **1**:5–6
      calcium loss **1**:5, **1**:6*F*
      geographic distribution **1**:7*F*
      nutrient cation loss **1**:5, **1**:6*F*
      sulfur and nitrogen accumulations **1**:6
    surface water **1**:7*F*
    trees
      red spruce **1**:6–8, **1**:8*F*
      species declines **1**:6–8
      species distribution shifts **1**:8
      sugar maples **1**:8, **1**:8*F*
  freshwater ecosystems
    aluminum concentrations **1**:9–10
    anthropogenic impacts **3**:572–574, **3**:577–578
    aquatic biota response **1**:9–10, **1**:10*T*, **1**:11*F*
    pH changes **1**:9–10, **1**:10*T*, **3**:584–585, **3**:584*F*
    surface water acidification **1**:8–9, **1**:10*T*, **1**:144–145
  geographic distribution **1**:1–5, **1**:3*F*
  lake and pond ecosystems **4**:465–466
  research background **1**:1
  schematic diagram **1**:2*F*
acid phosphatase **5**:269*T*
acid rain
  ecosystem alterations **6**:151–152
  environmental threats **3**:166
  habitat degradation **3**:407, **3**:407*F*
Acinonyx jubatus **1**:52, **2**:357–358, **3**:190–191, **4**:693, **4**:715–716, **5**:472
Acipenser nudiventis **2**:528*T*
Ackama **3**:522*T*
Acnonyx jubatus **5**:369
Acoetidae **7**:417*F*
Acomys cahirinus **3**:674*F*, **5**:466
Acomys russatus **5**:466
Aconitum **1**:270
acorns **1**:709
acorn worms **4**:381*T*
acoustic communication **1**:656
acquired immune deficiency syndrome (AIDS) **1**:594
Acraeinae **1**:653
Acremonium **3**:638
Acrididae **1**:710
Acrididae (true grasshoppers and locusts) **3**:735
Acridoidea
  Acrididae (true grasshoppers and locusts) **3**:735
  characteristics **3**:734
  Near East ecosystems **5**:460*T*
  Pamphagidae (earth hoppers of pamphagids) **3**:734–735
  Romaleidae (lubber grasshoppers) **3**:734*F*, **3**:735
Acridotheres tristis **1**:77–78, **4**:353, **4**:364
Acridoxeninae **3**:727
Acrobates pygmaeus **4**:688–689
Acrobatidae **4**:688–689
Acrocephalus paludicola **7**:182
Acrocephalus sechellensis **1**:601–602, **7**:188
Acroceridae **3**:489–491*T*
Acrocirridae **7**:417*F*
Acrocomia aculeata **4**:286–287
Acrocomia vinifera **7**:274
Acropora cervicornis **2**:338–340, **3**:201*T*, **3**:202*T*
Acropora palmata **2**:338–340, **2**:528*T*, **2**:529*F*, **3**:201*T*, **3**:202*T*
Acrostichum **5**:11*T*
Acrostichum aureum **1**:277, **5**:18–19
Acrothoracica **2**:416–417*T*
acrsids **6**:289–293
Actinidia deliciosa **3**:128
Actinobacteria **2**:613–614
Actinomonas **6**:293
Actinophrys **6**:290*F*, **6**:293
Actinopodidae **1**:206, **6**:289*T*, **6**:292
Actinopterygii **2**:7*F*, **3**:463, **3**:463*F*, **6**:522–523, **7**:338
Actinosphaerium **6**:293
actiomycetes **3**:652*T*
Aculeata **5**:674
acylalanines **5**:695
Acyrthosiphon pisum **2**:139
Adaliabi punctata **3**:222
Adams consensus analysis **2**:42–44, **2**:43*F*
Adansonia digitata **1**:67, **5**:235
adaptation **1:15–20**
  adaptationism
    basic concepts **1**:18–19
    criticisms **1**:19
    post-adaptationism research **1**:19–20
  adaptive landscape theory **2**:510, **2**:511*F*, **6**:632
  current research areas **1**:19–20
  definitions
    evolution-based adaptations **1**:16–17
    general discussion **1**:16
    process-based adaptations **1**:16
  evidential perspectives
    general discussion **1**:17–18
    genomics **1**:18
    increased fitness **1**:18
    natural selection **1**:17–18
  historical perspectives
    biblical theories **1**:15
    classical era **1**:15
    Darwinian theory **1**:15–16, **2**:263–265, **2**:420–421
    pre-Darwinian theories **1**:15
adaptive radiation **1:21–36**
  action mechanisms
    ecological change phase **1**:27
    genetic change phase
      founder events **1**:26
      rapid proliferation and hybridization **1**:26–27
  biodiversity generation **1**:430
  case studies
    Caribbean lizards **1**:31*F*, **1**:32
    cichlids **1**:25*F*, **1**:29–30
    cultured bacteria **1**:33
    Galapagos finches **1**:27–28, **1**:28*F*
    Hawaiian *Drosophila* **1**:28–29
    Hawaiian honeycreepers **1**:29, **1**:29*F*
    Hawaiian silverswords **1**:30–32, **1**:30*F*
    Hawaiian swordtail crickets **1**:32–33
    Hawaiian *Tetragnatha* spiders **1**:32, **1**:33*F*
    land snails **1**:33
    research background **1**:27–28
    research summary **1**:33–34
    three-spined sticklebacks **1**:32, **1**:32*F*
  definition **1**:21
  future research areas **1**:34
  historical perspective **1**:21
  nonadaptive radiations **1**:21
  oceanic islands **5**:593–594
  predisposition theories **1**:25–26
  sister taxon analysis **1**:432
  underlying factors
    intrinsic versus extrinsic factors **1**:21–24, **1**:22*F*, **1**:23*F*
    key innovations **1**:22–24
    novel environment theory
      environmental change responses **1**:24–25, **1**:25*F*
      isolated landmasses **1**:25
      speciation rates **1**:25*F*
adaptive syndromes **4**:37
Addax nasomaculatus **1**:53, **3**:192
addiction treatment drugs **1**:594
Adelges piceae **4**:361
Adelges tsugae **4**:361
adelgids **4**:361
Adeloidea **5**:386, **5**:386*F*, **5**:387–388*T*
Adenanthos **5**:221
Adenium socotranum **1**:51–52
adenosine triphosphate (ATP) **2**:362*F*, **3**:479, **6**:12–13
Adenota kob thomasii **1**:52
Adenoviridae **5**:134*T*, **5**:135
Adephaga **1**:326, **1**:327*T*
Adesmia **5**:217
Adesmia caespitosa **4**:105–106
Adesmia salicornia **4**:106–107
Adesmia subterranea **4**:105–106

Adesmia villosa **4**:106–107
Adiantum capillus-veneris **5**:486
adipose tissue
  adipocyte size and number **7**:27
  anatomical arrangement **7**:27–30, **7**:28*F*, **7**:29*F*
  characteristics **7**:26–27, **7**:27*F*
  fatness measures **7**:30–31, **7**:30*F*
  functional role
    adipocyte metabolism **7**:31–32
    fatty acid composition **7**:32–33
    thermal insulation **7**:33–35, **7**:34*F*
ad-libitum sampling **4**:60–61
ADMADE project **7**:19–20
Adonis vernalis **1**:269
adult parasites **5**:657
Advanced Land Observation Satellite (ALOS) **6**:394
Advanced Very High Resolution Radiometer (AVHRR) **4**:567
Aëdes pembaensis **5**:13
Aegialitis **5**:11*T*
Aegiceras **1**:277, **5**:11*T*
Aegllops peregrina **3**:670
Aepyceros melampus **1**:51, **6**:572–573
Aepycerotinae **4**:696
Aepyornis **1**:55–56, **4**:657, **4**:730–731
Aepyornis maximus **1**:601, **5**:365
Aepyornithidae **1**:606
Aequorea victoria **1**:591–592
aequorin **1**:591–592
aerial photography **4**:567, **6**:392
aerobic respiration, microbial **5**:301*T*
Aeromonas hydrophila **3**:696
aerosols **2**:680, **3**:442
Aerva **2**:371*T*, **2**:376*F*
Aeschynomene rudis **1**:164
Aeschynomene sensitiva **1**:165
Aesculus hippocastanum **6**:105*F*
Aesculus turbinata **1**:271
aesthetic factors **1:37–44**
  biodiversity **1**:43, **3**:57
  cognition **1**:43
  cultural ecosystem services **5**:353–354
  definition **1**:37
  emotional responses
    negative emotional responses **1**:39–40
    positive emotional responses
      foods/food selection **1**:42–43, **1**:42*T*
      habitat selection **1**:41–42
      research background **1**:40–41
      restorative responses **1**:43
      symmetry **1**:40–41
    purpose **1**:37
  environmental information classifications **1**:38–39, **1**:38*T*, **1**:39*T*
  evolutionary considerations **1**:37–38
  human well-being **1**:437
  overexploitation **4**:156
  problem-solving factors **1**:39
  valuation measures **6**:170
Aextoxicaceae **5**:216–217
Aextoxicon **5**:622*F*
Afghanistan
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  deforestation **2**:449*T*
  ecosystems
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  urbanization **4**:560*F*
  wheat production **1**:87–88
aflatoxins **3**:634–635
Africa
  agricultural intensification effects **1**:130*F*
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
  butterflies **1**:658, **1**:659*T*
  C_4 plants **2**:371*T*
  climate **1**:46
  climate change effects **1**:130*F*
  crop distribution **1**:88*T*
  crop plant domestication **2**:659
  deforestation
    deforestation rates **1**:679*T*
    deforesting countries **2**:448*T*, **2**:449*T*
    economic factors **1**:677*F*
    forest product exports **2**:451*T*
    historical perspective **2**:447–448
    land use change **2**:31*F*
    logging operations **2**:454*T*
    total forest change **2**:452*T*
  desert ecosystems **2**:467–468
  Diptera **3**:488*T*
  domesticated ruminants **1**:702–703
  ecosystems **1:45–57**
    Afro-montane forests **5**:219
    alpine life zone **1**:149*T*
    climate **1**:46
    conservation status **3**:171–173*T*
    fire/fire regimes **1**:46–47
    future outlook **1**:56–57
    geographic characteristics **1**:45
    geologic history **1**:45–46
    Madagascar **1**:55–56
    past climatic fluctuations **1**:47
    phytogeographic/ecoclimatic zones
      Cape Region (South Africa) **1**:47–48
      Mediterranean North Africa **1**:47
      tropical Africa **1**:48
    plant diversity **6**:314–315
    river water composition **3**:562*T*
    soils **1**:46
    termites **4**:430*T*
    tropical Africa
      afroalpine ecosystems **1**:53
      bushlands **1**:51–52
      coastal ecosystems **1**:55
      deserts **1**:53
      forests **1**:48–49
      grasslands **1**:52
      lake and pond ecosystems **1**:54–55
      montane ecosystems **1**:53
      phytogeographic/ecoclimatic zones **1**:48
      seasonal tropical vegetation **1**:49–51
      shrublands **1**:52–53
      thickets **1**:51–52
      wetlands **1**:53–54
      woodlands **1**:49–51
    tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*
    vegetation classifications **1**:45
ecotones **3**:143*F*
ecotourism **1**:395–396
energy use **3**:252*T*
family planning programs **6**:207
fire/fire regimes **1**:46–47, **3**:435–436
fish diversity **3**:472–473, **3**:474
floodplain agriculture **7**:146
forest product exports **2**:451*T*
freshwater ecosystems **4**:370*T*
geologic history **1**:45–46
habitat–bird species diversity relationship **1**:613–614
habitat destruction **3**:403
indigenous peoples **4**:269*T*, **4**:271*T*
intercropping **2**:382–383
invasive species
  aquatic ferns (*Salvinia molesta*) **4**:346–347
  Lantana camara **4**:352–353
  Mimosa pigra **4**:353–354
lake and pond ecosystems **4**:461*T*
land-use trends **4**:559*F*
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*, **4**:729–730
logging operations **2**:454*T*
maize production **1**:86, **1**:88–89
mammalian evolution **4**:701*T*
mammalian extinctions **4**:705*T*
megaherbivores **5**:224*T*, **5**:225*T*, **5**:232*F*
multiple cropping systems **1**:119–120
myriapods **5**:429*T*
native crops **2**:663
natural history museums and institutions **5**:406
nitrogen deposition **5**:526*T*
parasitic infections **5**:671*T*
plant diversity **6**:82*T*
polyculture cropping systems **7**:230
postglacial extinctions **5**:364–365
primates
  primate diversity **6**:250–252, **6**:251*T*
  threatened species **6**:258*T*
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
resource management **6**:386–387, **6**:387*T*
rice production **1**:88
sacred sites **6**:384, **6**:384–385
savannas **1:58–74**
  biodiversity **1**:70, **1**:71*F*
  characteristics **1**:58–59
  classifications
    arid savannas **1**:65–67, **1**:66*F*
    dystrophic savannas **1**:65–67, **1**:66*F*
    eutrophic versus mesic savannas **1**:65–67, **1**:66*F*
    forest-savanna mosaics **1**:66*F*, **1**:68
    general discussion **1**:65–67
    grass and shrub savannas **1**:66*F*, **1**:67
    life-form associations **1**:66*F*, **1**:67
    shrublands **1**:66*F*, **1**:67
    tree and shrub savannas **1**:66*F*, **1**:67

Africa (*continued*)
woodland savannas **1**:66*F*, **1**:67–68
climate change effects **1**:71–72
emotional responses **1**:41
future outlook **1**:71–72
geographic distribution **1**:59–60, **1**:59*F*
global carbon cycle **1**:70–71
precipitation–temperature relationship **1**:59*F*
species richness **1**:71*F*
structure–function drivers
bottom-up controls **1**:60–61
characteristics **1**:60–61
fire/fire regimes **1**:61, **1**:63
herbivory **1**:61, **1**:65
nutrient intake regulation mechanisms **1**:63–65
regulating factors **1**:63
top-down controls **1**:61
water and nutrient intake **1**:60–61
woody vegetation dynamics **1**:61–63, **1**:62*F*, **1**:64*F*
termites **4**:430*T*, **4**:432*T*
tree cover variations **1**:60*F*, **1**:61–63, **1**:62*F*, **1**:64*F*
tropical biomes **1**:49–51, **1**:58–59
soils **1**:46
Sub-Saharan Africa **2**:663
Succulent Karoo **1:485–490**
biodiversity
biodiversity **2**:300
invertebrates **1**:488
plants **1**:485, **1**:486*T*, **1**:487, **1**:487*T*
vertebrates **1**:487*T*
bioregions **1**:485, **1**:486*F*
characteristics **1**:485, **1**:486*F*
climate change threats **1**:489
land degradation threats **1**:489
nitrogen deposition **5**:526*T*
plant growth forms **1**:488–489
threat management **1**:489
vegetation structure **1**:488–489, **2**:467–468
termites **4**:409*T*, **4**:430*T*, **4**:432*T*
transboundary protected areas **6**:489*T*
tropical forests **7**:269–270, **7**:270*T*
tropical rain forests
distribution **6**:324*T*
loss and change **6**:324*T*
model-based assessments **6**:324*T*
species richness loss **6**:328*T*
termites **4**:430*T*, **4**:432*T*
African acacia **4**:515
African black oystercatcher **6**:232
African bridal creeper **4**:362
African buffalo **1**:487–488, **4**:696
African clawed frogs **3**:695–696
African crystalline ice plant **4**:360
African elephants **4**:694–695
African finch **1**:604–605
African forest elephant **5**:225*T*
African groung squirrel **2**:473–474, **2**:474*F*
Africanized honey bees **2**:551–552*T*, **4**:183
African locusts **2**:549–550
African oil palms **1**:529*T*
African palm weevil **5**:603–604
African rhinoceros beetle **5**:603–604
African rice **2**:663
African savanna elephant **5**:225*T*, **5**:226
African scorpion **1**:204
African soapberry tree **1**:592
African wild ass **3**:192
African wild cat **5**:475
African wild dogs
conservation efforts
background and characteristics **4**:715–716, **4**:717*F*
conservation practices **4**:716–717
major threats **4**:716
range areas **4**:718*F*
disease outbreaks **3**:191
interspecific competition **2**:190, **2**:190*F*
suitable habitats **3**:190
African Wildlife Fund **2**:231–232, **4**:325–326
afroalpine ecosystems **1**:53
Afrolarcus aequalis **3**:736*F*
Afro-montane forests **5**:219
Afropavo congensis **1**:49
Afrosoricida **4**:690
Afrostyrax **5**:623*F*
Afrotheria clade **4**:685
Afrotropics **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
Afzelia africana **1**:68
Agalmaceros **4**:727*T*
Agama savignii **5**:472
Agaonidae **4**:178
Agaricus bisporus **3**:625*T*, **3**:637–638, **3**:637*T*
Agasicles hygrophila **4**:361
Agassiz, Louis **1**:578
Agastoschizomus lucifer **1**:212
Agathiphaga **5**:393*F*
Agathiphagoidea **5**:386*F*, **5**:387–388*T*
Agathis **7**:119
Agathosma **5**:219
Age of the Dinosaurs **5**:5–6
Ageratum conyzoides **1**:376
age–sex pyramid **6**:202, **6**:204*F*
aggregation model **2**:181–182, **2**:182*F*
Agkistrodon rhodostoma **5**:711
Aglaia **1**:263
Aglossata **5**:385, **5**:386*F*, **5**:387–388*T*
Agnatha
characteristics **3**:457, **7**:337–338
hagfish
characteristics **3**:456–457, **3**:457, **3**:458*F*, **7**:336–337
classification systems **7**:335, **7**:335*F*, **7**:336*F*
lampreys **3**:456–457, **3**:457–458, **3**:458*F*, **7**:337–338
species diversity **3**:684–685
Agoutidae **6**:309*T*
Agouti paca **4**:287
agoutis **2**:590*T*, **5**:14
Agra grace **3**:512–513, **3**:513*F*
agribusiness **4**:144
Agricultural Research Service Culture Collection **3**:652*T*
agriculture
agribusiness **4**:144
agrobiodiversity **1:126–135**
agricultural intensification
biodiversity conservation **2**:299
environmental resilience **1**:130–131, **1**:131*F*
rapid land-use change/world food demand **1**:128–129, **1**:129*F*, **1**:130*F*
sustainable agriculture **1**:129–130
wildlife protection **1**:131–132
agrobiodiversity loss **3**:34–35, **3**:35*F*
characteristics and functional role **1**:127–128, **1**:127*F*, **1**:128*F*
conservation factors **1**:133–134, **1**:134*F*
conservation practices **3**:33
ecological functions
characteristics **1**:132
community–ecosystem relationship **1**:132–133
genetic and population diversity **1**:132, **1**:132*T*
landscape diversity **1**:133
economic factors **3:31–44**
agrobiodiversity loss **3**:34–35, **3**:35*F*
economic/market instruments **3**:39–40
landrace displacement **3**:34, **3**:35*F*
market and institutional failures **3**:34–35, **3**:35*F*
planned versus unplanned biodiversity **3**:31–34
public versus private valuation **3**:36, **3**:36*F*
regional cultivars **3**:32*F*
total economic value **3**:36–37
underutilized plant species **3**:36
valuation measures **3**:35–36
economic/market instruments
International Treaty on Plant Genetic Resources for Food and Agriculture **3**:39–40
market value chain development **3**:39–40
payment for agrobiodiversity conservation services (PACS) **3**:40–41, **3**:41*F*
payment for ecosystem services (PES) **3**:40–41
voluntary direct payments **3**:40–41, **3**:41*F*
planned versus unplanned biodiversity **1**:127*F*, **3**:9, **3**:31–34
private and social values **1**:127–128, **1**:128*F*
regional cultivars **3**:32*F*
socioeconomic factors **1**:133–134, **1**:134*F*
total economic value
components **3**:36–37
direct use value **3**:37
global option value **3**:38–39
indirect use value **3**:37–38
national insurance value **3**:37–38
agroecosystems **3:9–15**
background information **3**:9
biodiversity–human health relationship **1**:359–360

future outlook **3**:15
landscape diversity
ecological consequences **3**:13
general discussion **3**:13
organism spillover **3**:13–14
spatial structure **3**:14–15, **3**:14*F*
nutrient cycling **1**:95–97, **1**:96*F*
pest control **1:373–385**
agricultural intensification effects **1**:374–375
agricultural landscapes **1**:380–381
background information **1**:373–374
conservation biological control **1**:382
crop diversity **1**:375–376
habitat heterogeneity **1**:381–382, **1**:381*F*
influencing factors **1**:374*F*
multiple trophic levels **1**:380
natural enemies hypothesis **1**:375–376
natural habitat proximity **1**:381
noncrop habitats **1**:380–381
noncrop plants **1**:376, **1**:376*F*
predator diversity **1**:376–377
research background **1**:373–374
summary discussion **1**:382–383
vegetation diversity **1**:375–376
planned diversity
characteristics **3**:9–11, **3**:10*F*
host–parasite interactions **3**:11–12
pest control **3**:11–12
productivity **3**:9–11, **3**:10*F*
soil processes **3**:12
stability **3**:9–11
planned versus unplanned biodiversity **1**:127*F*, **3**:9, **3**:31–34
termites **4**:432*T*
unplanned diversity
ecological services **3**:12–13
pest control **3**:12–13
soil processes **3**:13
weed management
biodiversity effects **4**:91–92
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
agroforestry **2**:382–383
biodiversity conservation
biodiversity-friendly farming **2**:299–300, **2**:300
biodiversity stewardship **2**:300
conservation approaches **2**:298–299
management practices **2**:299
biodiversity–ecosystem services
relationship **1**:346, **2**:298–299, **3**:345
biodiversity-friendly farming **1:418–429**
background information **1**:418–419
biodiversity conservation **2**:299–300, **2**:300
biodiversity stewardship **2**:300
disadvantages **2**:544
influencing factors **2**:544, **2**:554*F*
landscape-scale effects
general discussion **2**:544
landscape connectivity **2**:544–545
landscape heterogeneity **2**:544
local management strategies
chemical use minimization **2**:542
crop production **2**:542–544
disturbance events **2**:542–544
general discussion **1**:419–420
livestock grazing **2**:542–544
native vegetation benefits **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
pest control **2**:544
pollination **2**:544
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
structural complex agricultural areas **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
management practices **2**:299–300, **2**:300
research summary **2**:545–546
biodiversity impacts **7**:203*F*
biofuel crop production
agroenergy sprawl **1**:534–535
biodiversity-friendly energy-crop production strategies **1**:535–536
biodiversity management strategies **1**:533
biodiversity responses **1**:529–530, **1**:530*F*
first-generation biofuel crops
economic importance **1**:529*T*, **1**:530
feedstock management **1**:530–531
feedstock selection **1**:530
second-generation biofuel crops
economic importance **1**:530*F*, **1**:531–532
native polycultures **1**:532–533
perennial monocultures **1**:532
short-rotation woody biomass **1**:531–532
Chinese Grain-to-Green Program (GTGP)
background information **3**:379–380
biodiversity/biophysical effects **3**:379*F*, **3**:380
future outlook and challenges **3**:381–382
program goals **3**:379–380
socioeconomic effects **3**:380–381, **3**:381*F*
spatial distribution **3**:376*F*, **3**:379*F*, **3**:380
summary discussion **3**:382
commercial farming industry **4**:144
crop genetic diversity **1:85–94**
agricultural intensification **1**:86, **1**:122–123, **1**:374–375, **4**:558, **4**:565–566
agrobiodiversity conversation **3**:33
basic concepts **1**:87–89
commercialization effects
general discussion **1**:91–92
intensified environments **1**:92
non-intensified environments **1**:92–93
crop genetic resources **3**:39–40
genetic narrowing **1**:89–90, **1**:404–405
genetic trends **1**:90
landrace–modern varieties comparisons **1**:89
research summary **1**:93
seed selection evolution **1**:86–87
spatial distribution **1**:87–89, **1**:88*T*
cropland conversion **3**:427, **3**:428*T*, **4**:154, **4**:665*T*
cropping systems
agricultural production–ecosystem maintenance relationships **1**:116*F*
animal integration **1**:116
bush-fallow rotation systems **1**:114–115
ecosystem services **1**:114
intercropping **2:382–395**
above-ground–below-ground species interactions **2**:386
above-ground species interactions **2**:386–387
Africa **2**:382–383
Asia **2**:383, **2**:384*F*
below-ground species interactions **2**:388–389
crop pest and disease control **2**:391
definition **2**:382–383
importance **2**:382–383
intercropping patterns **2**:383
land equivalent ratio (LER) **2**:386
Latin America **2**:382–383
summary discussion **2**:392
yield classifications **2**:386
multifunctional agriculture **1**:115–117, **1**:115*F*, **1**:116*F*, **1**:116*T*, **1**:117*F*
resource conservation **1**:113–114
crop plant domestication **2:657–664**, **4:279–292**
basic concepts **4**:279
conscious versus unconscious selection **2**:660
crop maintenance **2**:660–661
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
evolutionary development
African Savanna belt **2**:659
East Asia region **2**:658–659
fertile crescent region **2**:658
founder crops **2**:658
pre-Columbian America **2**:659
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
growing purposes **2**:661
high care-level plants **4**:289*T*
historic land-use patterns **4**:570, **4**:570*F*
human–crop coevolution **2**:657, **4**:570, **4**:570*F*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
native crops **2**:662
reproductive isolation **4**:281*F*, **4**:285
reproductive systems

agriculture (*continued*)
characteristics **2**:659–660
self-pollination **2**:659–660
vegetative propagation **2**:660
research summary **4**:290–291
sowing and reaping impacts **2**:661–662
traditional agriculture **2**:657–658
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
crop production
biodiversity-based agricultural strategies **1**:123–124
biodiversity-friendly farming **2**:542–544
biodiversity–human health relationship **1**:359–360
biofuels
agroenergy sprawl **1**:534–535
biodiversity-friendly energy-crop production strategies **1**:535–536
biodiversity management strategies **1**:533
biodiversity responses **1**:529–530, **1**:530*F*
crop production impacts **4**:293
first-generation biofuel crops **1**:529*T*, **1**:530
second-generation biofuel crops **1**:530*F*, **1**:531–532
climate change effects **1**:130*F*
commercialization
basic concepts **1**:90–91
seed industries **1**:91
crop homogeneity **1**:460–461, **1**:461*F*
crop plant pathogens
Australia **1**:76*T*, **1**:81
Brazil **1**:76*T*, **1**:83
fungi **3**:633, **3**:633*T*
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81–82
United Kingdom **1**:76*T*, **1**:80
deforestation **2**:453
dryland syndrome **2**:483*F*, **2**:484*F*
ecosystem services modeling **5**:349–350
energy crops **1**:542–543, **1**:542*F*
genetic diversity **1**:360
global agriculture impacts **3**:427, **3**:428*T*, **4**:154
global cropland area **4**:558*F*
hemiparasites **4**:75–77
introduced crop plants **1**:75–76
land-use change impacts **1**:542–543, **1**:542*F*
land-use trends **4**:559*F*
monoculture crops **1**:359–360
parasitic disease vectors **5**:673
plant–human interactions **6**:57–58
pollination **1**:360, **5**:354–355
polycultures **1**:119–120
ultraviolet radiation effects **7**:298–299
unplanned diversity **3**:12–13
weed management
biodiversity effects **4**:91–92
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
wetland conservation programs **7**:371
deforestation **2**:452–453, **2**:452*T*, **4**:562–564, **4**:563*F*
floodplain agriculture **7**:146
gene banks **3:641–653**
conservation role **3**:641–642
ex situ collections **3**:641–642
nonseed gene banks
botanical gardens **3**:653
cryopreservation **3**:651–652
DNA banks **3**:652, **4**:319–320, **4**:321*T*
field gene banks **3**:33, **3**:652–653, **4**:319, **4**:321*T*
microorganisms **3**:652, **3**:652*T*
pollen **3**:650, **4**:319–320, **4**:321*T*
somatic and zygotic plant embryos **3**:650–651
spores **3**:650
vegetative tissues **3**:651
seed banks
collection value **3**:644–645
community seed banks **3**:644
current status **3**:643–644
effectiveness **3**:642
genetic considerations **3**:646
historical perspective **3**:642–643
longevity predictions **3**:645–646
management strategies **3**:646–647
scientific principles **3**:645–646
seed storage conditions **3**:645–646
wild species **3**:644
global agriculture **3:426–434**
biodiversity regime shifts and loss thresholds **3**:429–430, **3**:430*T*
future outlook
agricultural policies **3**:432–433
agriculture-related conservation strategies **3**:431
market-based approaches **3**:432–433
Millennium Ecosystem Assessment (MEA) **3**:430–431, **3**:431*T*
payment for ecosystem services (PES) **3**:432–433
sustainable agriculture **3**:430–431
impacts
climate change **3**:427–429, **3**:428*F*
general discussion **3**:427
habitat loss **3**:427, **3**:428*T*, **4**:154
invasive species **3**:429
water quality **3**:429
water quantity **3**:429
prevalence and distribution **3**:426–427, **3**:427*F*
research summary **3**:433
habitat destruction **3**:164–165, **6**:62
habitat loss and fragmentation **4**:50, **4**:51*F*, **4**:56
human impacts **4**:141, **4**:154
industrialized agriculture **1:85–94**
agricultural intensification **1**:86, **1**:122–123, **1**:374–375, **4**:558, **4**:565–566
crop genetic diversity
agrobiodiversity conversation **3**:33
basic concepts **1**:87–89
commercialization effects **1**:91–92
crop genetic resources **3**:39–40
genetic narrowing **1**:89–90
genetic trends **1**:90
landrace–modern varieties comparisons **1**:89
research summary **1**:93
spatial distribution **1**:87–89, **1**:88*T*
crop production
background information **1**:85–86
commercialization **1**:90–91
seed industries **1**:91
seed selection evolution **1**:86–87
Southeast Asia **1**:517, **1**:519*F*, **1**:520*F*
insecticide resistance **4:298–307**
cross-resistance **4**:302
diagnostic assays **4**:302–303, **4**:302*F*
gene homology **4**:301
management strategies **4**:305–306
multiple resistance **4**:302
mutation recurrence **4**:301–302
nonpest species **4**:306–307
origins **4**:299–300
prevalence **4**:298–299, **4**:298*F*, **4**:299*F*
research background **4**:298
research summary **4**:307
resistance gene selection
determinant factors **4**:303
ecological influences **4**:303–304
fitness costs **4**:305
genetic influences **4**:303
operational influences **4**:304–305
resistance mechanisms
characteristics **4**:300
increased detoxification **4**:300, **4**:300*F*
target site alterations **4**:300–301, **4**:300*F*
transgenic plants **4**:307
intercropping **2:382–395**
above-ground species interactions
canopy heat capture **2**:387, **2**:387*F*
characteristics **2**:386–387
light interception/light use efficiency **2**:386–387
shading effects **2**:387, **2**:387*F*
spatiotemporal complementarity **2**:387–388
Africa **2**:382–383
Asia **2**:383, **2**:384*F*
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391

crop pest and disease control
biological pest control **2**:391
genetic diversity effects **2**:391–392
push–pull strategies **2**:391
species diversity effects **2**:392
definition **2**:382–383
importance **2**:382–383
intercropping patterns **2**:383
Latin America **2**:382–383
overyielding
above-ground–below-ground species interactions **2**:386
above-ground species interactions **2**:386–387
below-ground species interactions **2**:388–389
classifications **2**:386
crop pest and disease control **2**:391
land equivalent ratio (LER) **2**:386
summary discussion **2**:392
international organizations **4**:328
introduced crop plants **1**:75–76
introduced species **1:75–84**
Australia
nonindigenous invertebrates **1**:76*T*, **1**:80–81
nonindigenous plants **1**:76*T*, **1**:80
nonindigenous vertebrates **1**:76*T*, **1**:80
Brazil
nonindigenous invertebrates **1**:76*T*, **1**:83
nonindigenous plants **1**:76*T*, **1**:82–83
nonindigenous vertebrates **1**:76*T*, **1**:83
economic impact **1**:75
India
nonindigenous invertebrates **1**:76*T*, **1**:82
nonindigenous plants **1**:76*T*, **1**:82
nonindigenous vertebrates **1**:76*T*, **1**:82
research summary **1**:83
South Africa
nonindigenous invertebrates **1**:76*T*, **1**:81
nonindigenous plants **1**:76*T*, **1**:81
nonindigenous vertebrates **1**:76*T*, **1**:81
United Kingdom
nonindigenous invertebrates **1**:76*T*, **1**:79–80
nonindigenous plants **1**:76*T*, **1**:79
nonindigenous vertebrates **1**:76*T*, **1**:79
United States
background information **1**:75–76
nonindigenous invertebrates **1**:76*T*, **1**:78
nonindigenous plants **1**:75–76
nonindigenous vertebrates **1**:76–77, **1**:76*T*
land clearance **2**:447–448
landscape legacies **4**:524–525
land use issues **4**:556–558
marginal, degraded, and abandoned lands **1**:543, **1**:544*F*
Mlao cultivation **7**:146
native crops **2**:662
Near East ecosystems **5**:475–477
Neolithic agriculture **2**:658
nutrient management–water quality relationship **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
freshwater invertebrates **4**:377
land-use change impacts **1**:107–108
leaching vulnerability index **1**:102*F*
management criteria **1**:97, **1**:97*T*
nitrogen (N) **1**:96*F*, **1**:98
nutrient input and sources **1**:96*F*, **3**:575*F*, **3**:576–577, **3**:576*F*
phosphorus (P) **1**:96*F*, **1**:98–99
regional nutrient imbalances **1**:98*F*, **1**:99–101, **1**:100*F*, **1**:102*F*
remedial benefical management practices (BMPs)
characteristics **1**:103*T*
cover crops **1**:103*T*, **1**:106
critical source area management **1**:103*T*, **1**:107
erosion potential **1**:103*T*, **1**:105–106, **1**:105*T*, **1**:106*F*
farm input decisions **1**:101, **1**:103*T*
fertilizer application methods **1**:103*T*, **1**:104
fertilizer application rate and frequency **1**:103*T*, **1**:104–105
implementation strategies **1**:101, **1**:103*F*
nutrient-use efficiency improvements **1**:101–104, **1**:103*T*
riparian/buffer zones **1**:103*T*, **1**:106–107
runoff potential **1**:103*T*, **1**:105
soil and plant testing **1**:103*T*, **1**:104
source management **1**:103*T*, **1**:104
transport management **1**:103*T*, **1**:105
research background **1**:95–97
research summary **1**:108
on-farm conservation **3**:33, **4**:316*F*, **4**:320, **4**:321*T*
overyielding
above-ground–below-ground species interactions **2**:386
above-ground species interactions
canopy heat capture **2**:387, **2**:387*F*
characteristics **2**:386–387
light interception/light use efficiency **2**:386–387
shading effects **2**:387, **2**:387*F*
spatiotemporal complementarity **2**:387–388
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391
classifications **2**:386
crop pest and disease control
biological pest control **2**:391
genetic diversity effects **2**:391–392
push–pull strategies **2**:391
species diversity effects **2**:392
land equivalent ratio (LER) **2**:386
summary discussion **2**:392
pest control **1:373–385**
agricultural intensification effects **1**:374–375
agricultural landscapes
characteristics **1**:380–381
habitat heterogeneity **1**:381–382, **1**:381*F*
natural habitat proximity **1**:381
noncrop habitats **1**:380–381
background information **1**:373–374
conservation biological control **1**:382
influencing factors **1**:374*F*
multiple trophic levels **1**:380
predator diversity
complementarity **1**:377–378
enhancement factors **1**:379–380
facilitation **1**:377
functional diversity (FD) **1**:378
functional redundancy **1**:378
insurance hypothesis **1**:378
intraguild interference **1**:378–379, **1**:378*F*
risk reduction/risk enhancement **1**:376–377
sampling/selection effects **1**:377
research background **1**:373–374
summary discussion **1**:382–383
vegetation diversity
crop diversity **1**:375–376
natural enemies hypothesis **1**:375–376
noncrop plants **1**:376, **1**:376*F*
plant biodiversity threats **6**:62
plant breeding **1:459–469**
commodity agriculture **1**:467
controversies **1**:459
crop homogeneity **1**:461*F*
exploration and trade **1**:465–466
gene pool concepts **1**:463–465, **1**:465*F*
genetic bottlenecks **1**:463
green revolution genetics **1**:467–468
legal frameworks **1**:466–467
modern scientific plant breeding **1**:460–461, **1**:461*F*
phenotypic variation **1**:461–463, **1**:462*F*
quality guidelines **1**:461*F*
reduced genetic diversity **1**:461–463
research summary **1**:468

agriculture (*continued*)
plant–human interactions **6**:57–58
pollinators
economic importance **6**:144–145
genetic engineering effects **6**:145
introduced species **6**:144–145
plant–pollinator mutualism threats **6**:145
pollen gene banks **3**:650, **4**:316*F*, **4**:319–320, **4**:321*T*
supporting services **1**:360, **5**:354–355
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
prime agricultural lands **1**:543
slash-and-burn systems **6:551–562**
alternative systems
improved short fallow systems **6**:560–561
pastures **6**:560
plantation crops **6**:560
basic concepts **6**:551
carbon sequestration **6**:561–562
clearing effects **6**:554
consequences
bacteria and fungi reductions **6**:556
heat evolution **6**:554–555
invertebrate loss **6**:556
plant species loss **6**:556
soil properties and nutrients **6**:555–556
vertebrate loss **6**:556
weeds **6**:556
cultivation phase effects
characteristics **6**:556–557
nutrient export **6**:558
pest and disease suppression **6**:557–558
prolonged cropping phase effects **6**:558
tillage **6**:556–557
weed and crop residue management **6**:558
weeding operations **6**:557
fallow/crop sequences **6**:553–554
fallow length **6**:552, **6**:552*F*
fallow phase effects **6**:558–559
geographic distribution **6**:553
historical perspective **6**:553
hunter–gatherer societies **6**:567
labor requirements **6**:553
purpose
biomass removal **6**:551
declining crop yields **6**:551–552
economic factors **6**:551
nutrient release and addition **6**:551
pest and disease avoidance **6**:551
weed avoidance and suppression **6**:551
scale dependence
spatiotemporal scales **6**:559–560, **6**:560*F*
topographic preferences and fragmentation **6**:559
traditional conservation practices **7**:231–232
tropical forests **6**:339
soil conservation
biotic community manipulation **6**:594–595
diversified cropping **6**:594
general characteristics **6**:593–594
organic amendments **6**:594
pollutant bioremediation **6**:594
reduced tillage practices **6**:593–594
twentieth century conservation **2**:283–284
South America **6**:608
Southeast Asia **4**:562–564, **4**:563*F*
sustainable agriculture **1:111–118**
agrobiodiversity **1**:129–130
basic concepts **1**:111
cropping systems
agricultural production–ecosystem maintenance relationships **1**:116*F*
animal integration **1**:116
bush-fallow rotation systems **1**:114–115
ecosystem services **1**:114
multifunctional agriculture **1**:115–117, **1**:115*F*, **1**:116*F*, **1**:116*T*, **1**:117*F*, **2**:300–301
resource conservation **1**:113–114
developmental background **1**:111–113
global agriculture **3**:430–431
global marketplace **1**:117
indicators
criteria **1**:113–114
ecosystem services **1**:114
resource conservation **1**:113–114
traditional agriculture **1:119–125**
background information **1**:119
benefits **1**:124–125
biodiversity-based agricultural strategies **1**:123–124
biodiversity conservation **1**:122–123
biodiversity features **1**:119–120
definition **1**:119
ecological interaction benefits **1**:122
global patterns **2**:657–658
polyculture cropping systems **1**:375–376, **7**:230–231
research summary **1**:124–125
traditional knowledge systems
environmental knowledge **1**:121
experimental approaches **1**:122
farming practices **1**:121
folk taxonomies **1**:121
research background **1**:120–121
water quality
global agriculture impacts **3**:429
nutrient management **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
freshwater invertebrates **4**:377
land-use change impacts **1**:107–108
leaching vulnerability index **1**:102*F*
management criteria **1**:97, **1**:97*T*
nitrogen (N) **1**:96*F*, **1**:98
nutrient input and sources **1**:96*F*, **3**:575*F*, **3**:576–577, **3**:576*F*
phosphorus (P) **1**:96*F*, **1**:98–99
regional nutrient imbalances **1**:98*F*, **1**:99–101, **1**:100*F*, **1**:102*F*
remedial benefical management practices (BMPs) **1**:101, **1**:103*F*, **1**:103*T*
research background **1**:95–97
research summary **1**:108
weed management
biodiversity effects **4**:91–92
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
wetland conservation programs **7**:371
Agrilus **1**:328, **1**:329
Agrilus planipennis **4**:361
Agrobacterium tumefacience **6**:354
agrobiodiversity **1:126–135**
agricultural intensification
biodiversity conservation **2**:299
environmental resilience **1**:130–131, **1**:131*F*
rapid land-use change/world food demand **1**:128–129, **1**:129*F*, **1**:130*F*
sustainable agriculture **1**:129–130
wildlife protection **1**:131–132
characteristics and functional role **1**:127–128, **1**:127*F*, **1**:128*F*
conservation factors **1**:133–134, **1**:134*F*
conservation practices **3**:33
ecological functions
characteristics **1**:132
community–ecosystem relationship **1**:132–133
genetic and population diversity **1**:132, **1**:132*T*
landscape diversity **1**:133
economic factors **3:31–44**
agrobiodiversity loss **3**:34–35, **3**:35*F*
economic/market instruments
International Treaty on Plant Genetic Resources for Food and Agriculture **3**:39–40
market value chain development **3**:39–40
payment for agrobiodiversity conservation services (PACS) **3**:40–41, **3**:41*F*
payment for ecosystem services (PES) **3**:40–41
voluntary direct payments **3**:40–41, **3**:41*F*
landrace displacement **3**:34, **3**:35*F*
market and institutional failures **3**:34–35, **3**:35*F*
planned versus unplanned biodiversity **3**:31–34
regional cultivars **3**:32*F*
total economic value components **3**:36–37

direct use value **3**:37
global option value **3**:38–39
indirect use value **3**:37–38
national insurance value **3**:37–38
valuation measures
public versus private valuation **3**:36, **3**:36*F*
research background **3**:35–36
total economic value **3**:36–37
underutilized plant species **3**:36
planned versus unplanned biodiversity **1**:127*F*, **3**:9, **3**:31–34
private and social values **1**:127–128, **1**:128*F*
regional cultivars **3**:32*F*
socioeconomic factors **1**:133–134, **1**:134*F*
agro-botanical gardens **6**:505
Agro-Ecological Zones (AEZs) **4**:540–544
agroecosystems **3:9–15**
background information **3**:9
biodiversity–human health relationship **1**:359–360
future outlook **3**:15
landscape diversity
ecological consequences **3**:13
general discussion **3**:13
organism spillover **3**:13–14
spatial structure **3**:14–15, **3**:14*F*
nutrient cycling **1**:95–97, **1**:96*F*
pest control **1:373–385**
agricultural intensification effects **1**:374–375
agricultural landscapes
characteristics **1**:380–381
habitat heterogeneity **1**:381–382, **1**:381*F*
natural habitat proximity **1**:381
noncrop habitats **1**:380–381
background information **1**:373–374
conservation biological control **1**:382
influencing factors **1**:374*F*
multiple trophic levels **1**:380
predator diversity
complementarity **1**:377–378
enhancement factors **1**:379–380
facilitation **1**:377
functional diversity (FD) **1**:378
functional redundancy **1**:378
insurance hypothesis **1**:378
intraguild interference **1**:378–379, **1**:378*F*
risk reduction/risk enhancement **1**:376–377
sampling/selection effects **1**:377
research background **1**:373–374
summary discussion **1**:382–383
vegetation diversity
crop diversity **1**:375–376
natural enemies hypothesis **1**:375–376
noncrop plants **1**:376, **1**:376*F*
planned diversity
characteristics **3**:9–11, **3**:10*F*
host–parasite interactions **3**:11–12
pest control **3**:11–12
productivity **3**:9–11, **3**:10*F*
soil processes **3**:12
stability **3**:9–11
planned versus unplanned biodiversity **1**:127*F*, **3**:9, **3**:31–34
termites **4**:432*T*
unplanned diversity
ecological services **3**:12–13
pest control **3**:12–13
soil processes **3**:13
weed management
biodiversity effects **4**:91–92
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
agroforestry **2**:382–383
Agromyzidae **3**:489–491*T*, **3**:492, **3**:493–494
Agromyzoinea **3**:489–491*T*
Agropyron **1**:269
Agropyron cristatum **6**:456–457
Agropyron repens **1**:270
agroscapes **4**:590–591, **4**:591*F*
Agrostis stolonifera **3**:703
Agrostis tenuis **2**:512
Agusan marsh **1**:276*T*
Ahamitermes **4**:409*T*
ahimsa **6**:565
ahipa **3**:128
Ahnenreihe **2**:51, **2**:51*F*
ahupua'a system **7**:233
Aichi Biodiversity Targets **2**:306–307, **2**:312, **2**:317–318, **7**:204–205
Ailanthus altissima **6**:96
Ailuridae **1**:687*F*
Ailuroedus crassirostris **2**:243–244
Ailuropoda melanoleuca **3**:190, **3**:373
Aiptasia pallida **5**:119*F*
airborne dust **2**:680
Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) **6**:394
air pollution **1:136–147**
biodiversity impacts
air quality guidelines **1**:145–146
birds **1**:139
evidentiary research
general discussion **1**:139–141
lichen biodiversity **1**:141–142, **1**:142*F*
nitrogen deposition effects **1**:142–144, **1**:143*F*, **1**:146*F*
ozone effects **1**:144
smelting industry impacts **1**:140–141, **1**:141*F*, **1**:142*F*
surface water acidification **1**:144–145
evolutionary responses **1**:138–139
long-term and short-term effects **1**:137–138
mammals **1**:139
species interaction effects **1**:139, **1**:140*F*
synergistic interactions **1**:138
depositional and environmental pathways **1**:137
environmental impact **1**:136–137, **1**:137*T*
global impacts **1**:145–146, **1**:146*F*
historical perspective **1**:136
lower atmosphere **1**:289–290
major sources **1**:136–137, **1**:137*T*
summary discussion **1**:146
air pollution monitors **1**:591
Aizoaceae **2**:374*T*
ajmaline **5**:712*T*
Akaike's Information Criterion (AIC) **6**:701
Akashiwo sanguinea **1**:477*T*
Akeley Gallery of African Mammals (AMNH) **5**:406–407, **5**:407*F*
'Akepa **1**:29*F*
'Akialoa **1**:29*F*
'Akiapola'au **1**:29*F*
akiapolaau **1**:618–619
Akodon molinae **2**:587*F*
'Akohekohe **1**:29*F*
Akosombo Dam **3**:260
Aktautitan **5**:224*T*
Alabama canebrake pitcher-plant **3**:210–211, **3**:212*F*
alachlor **5**:696
Alagoas Curassow **7**:185
Alamosaurus **2**:519*T*
alanine **6**:113*T*
Alashan Plateau **2**:469–470
Alaska pollack **6**:422*T*
Alauda arvensis **4**:259
Alaudidae **1**:710
Albania
botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
albatross **2**:543, **3**:446, **7**:181–182
albedo, surface **2**:680–681
Albertosaurus **2**:519*T*, **2**:520*T*
albinism **6**:3–5
Albizia antunesiana **1**:68
Alca impennis **1**:606, **1**:617, **1**:621
Alcaligenes **6**:32
Alcanivorax borkumensis **1**:592
Alcedinidae **6**:308*T*
Alcelaphus bucelaphus caama **3**:191
Alcelaphus buselaphus **1**:51
Alcelphinae **4**:696
Alces **4**:724–725*T*
Alces alces **6**:466, **7**:390
Alchornea bogotensis **1**:166–167
Alchornea castaniifolia **1**:163
Alcidae **1**:606
alcohol dehydrogenase **3**:665
alder
boreal forests **1**:628–630, **1**:630*T*
Holocene postglacial spread **2**:551–552*T*
aldicarb **5**:699*T*
Aldina heterophylla **1**:160
Aldrovanda **6**:44*T*
Alectoris chukar **3**:145–146
Alectra **4**:71
Alectra fluminensis **4**:76
Alectra orobanchoides **4**:76
Alectra picta **4**:76
Alectra vogelii **4**:76
Aleppo pine **1**:269, **2**:541*F*
alerce tree **3**:210, **3**:211*F*
Aleutian Island study **4**:443–446, **4**:444*F*, **4**:445*F*, **5**:58–60, **6**:246–247, **6**:247*F*
Alexandrine parakeet **1**:82

Alexandrium **1**:472–473
Alexandrium catanella **1**:476–477, **1**:477*T*
Alexandrium fundyense **1**:476–477, **1**:477*T*
Alexandrium minutum **1**:477, **1**:477*T*
Alexandrium tamarense **1**:470, **1**:476–477, **1**:477*T*, **6**:30
alfalfa **1**:529*T*, **2**:139
alfalfa weevil **1**:78
Alfisols **6**:123*T*
algae
  algae-based biodiesel fuel **1**:592–593
  Arctic ecosystems **1**:231*T*
  Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346, **4**:346*F*
  biomonitoring processes **3**:122*T*
  Census of Marine Life project **2**:18–19
  convergent evolution **3**:613–617, **3**:615*F*, **3**:616*F*
  coral reef ecosystems **2**:338
  energy flow **3**:246–247
  evolutionary development **2**:641–642
  forest canopies **3**:517–518
  freshwater ecosystems
    acidification effects **3**:584–585, **3**:584*F*
    freshwaters **3**:571–572, **3**:571*T*
    freshwater sediments **3**:572*T*
  harmful algal blooms **1:470–484**
    biogeographical range **1**:470
    bloom dynamics **1**:475–478
    chemical ecology **1**:478–479
    coastal ecosystems **5**:543
    ecogenomic sensors **2**:697*F*
    future outlook **1**:481–482
    global change effects **1**:479–481
    global distribution **1**:470, **1**:471*F*
    isozyme analyses **1**:475–478
    lake and pond ecosystems **3**:356–358, **3**:449, **5**:25
    life histories **1**:478*F*
    microsatellite studies **1**:475–478, **1**:477*T*
    nutrient enrichment impacts **1**:95–97, **1**:96*F*
    phylogenetic and taxonomic relationships
      dinoflagellates **1**:472, **1**:473*F*
      general discussion **1**:470–472
      haptophytes **1**:473–474, **1**:474*F*
      heterokonts **1**:475, **1**:476*F*
    population genetics **1**:475–478
    species interactions **1**:478–479
    toxin distribution and characteristics **1**:478–479
  indicator species **7**:206–207*F*
  invasive species **4**:345–346, **4**:346*F*, **4**:357, **4**:359–360
  known and estimated species **4**:384*T*
  lake and pond ecosystems **4**:461*T*
  latitudinal range shifts **4**:604*T*
  lipid storage **7**:32
  macroalgae
    estuarine ecosystems **3**:358–359
    freshwater ecosystems **3**:358
    seagrass ecosystems **6**:547
    ultraviolet radiation effects **7**:299–300
  marine ecosystems **5**:48
  microalgae
    eutrophication
      harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*
      long-term human influences **3**:354–356, **3**:356*F*, **3**:357*F*
      species shifts **3**:352–354, **3**:353*F*
  modern extinctions **5**:368*T*
  nomenclature code **5**:550–551, **5**:553*T*, **5**:555
  nonseed gene banks **3**:652*T*
  number of species **6**:586*T*
  nutrient enrichment impacts **1**:95–97, **1**:96*F*, **3**:351, **3**:575*F*, **3**:576–577, **3**:576*F*
  oceanic food web **3**:244–245, **5**:49*F*
  oil spills **3**:256
  photosynthetic symbioses **5**:118–120
  plant-derived drugs and chemicals
    domoic acid **6**:132, **6**:134*F*
    green algal metabolites **6**:132–135, **6**:134*F*
    kainic acid **6**:132, **6**:134*F*
  red tides **3**:357, **3**:449
  river ecosystems **6**:477–478
  surface water acidification **1**:10*T*
  symbiotic algae **7**:300
  symbiotic relationships **5**:306
  temperature tolerance ranges **4**:113–115
  temporal habitat diversity fluctuations **6**:33–34, **6**:33*F*
  threatened species **7**:206–207*F*
Algeria **2**:465*F*, **6**:519
alien species invasion
  amphibian species decline **3**:695–696
  aquaculture **1**:194–196
  biodiversity impacts **7**:203*F*
  contemporary conservation theories and practices **2**:254–255
  dispersal biogeography **2**:554–555, **2**:554*F*
  endangered reptiles **3**:217–218
  fire/fire regimes **3**:440
  forest ecosystems **3**:534
  globalization effects **3**:701–703, **3**:702*F*
  grazer-based ecosystem restoration and management **4**:16–17
  impacts and distribution
    biodiversity threats **2**:80
    continental ecosystems **4**:155
    coral reef ecosystems **2**:341
    distribution patterns
      aquatic habitats **3**:411
      biodiversity loss **4**:663
      characteristics **3**:410–411
      island biogeography **3**:217–218, **3**:410–411, **3**:410*F*, **6**:83
    endangered mammals **3**:193
    energy flow **3**:246–247
    environmental impacts **3**:284
    fish species **3**:448
    freshwater invertebrates **4**:378
    human impacts **4**:143–144
    island biogeography **4**:155
    modern extinctions **5**:370
    neotropical mammals **2**:590
    onshore ocean waters **4**:155–156
    pathogens **4**:156
    plant–soil interactions **6**:125
    Southeast Asia **1**:518–520, **1**:521, **1**:522*F*, **4**:346–347
    terrestrial invertebrates **3**:222
    wetland ecosystems **7**:399
  invasive species–pathogen relationship **2**:529–530
  managed relocations **7**:238–239, **7**:239*F*
  marine sedimentary habitats **5**:113*F*, **5**:114
  river ecosystems **6**:481*T*, **6**:482
  taxonomic versus functional diversity **2**:655
  threatened birds **7**:184
aliphatic acids **5**:696, **6**:113*T*
Alismatales **2**:374*T*
alkaline phosphatase **5**:269*T*
alkaloids **2**:426*T*, **2**:427
Allee effect **2**:264–265, **4**:659, **6**:676–677, **7**:405
allele and genotype frequencies **4**:510, **6**:183
Allen, J. A. **1**:579
Allenrolfea vaginata **2**:469
allergic fungal sinusitis **3**:635
allethrin **5**:694, **5**:699*T*
Alliance for Zero Extinction (AZE) **6**:488, **7**:204, **7**:204*F*
Alliaria petiolata **6**:93–95, **6**:106*F*, **6**:107
Alligator mississippiensis **3**:217, **6**:401, **6**:411
alligators
  characteristics **6**:401
  river ecosystems **6**:479–480
  sustainable farming **3**:217
Alligator sinensis **6**:401
alligator weed **3**:703–704, **3**:704*F*, **4**:378
alligatorweed flea beetles **4**:361
Allionia **2**:376*F*
Allium cepa **2**:662–663
Allium chinense **2**:663
Allium porrum **2**:383, **2**:662–663
Allium sativum **2**:383, **2**:662–663
Allium tuberosum **2**:663
Allocasuarina decaisneana **2**:468
Allocasuarina pusilla **1**:304*F*
allocated quotas **2**:152
Allochrocebus preussi insularis **6**:256–258*T*
Allochrocebus preussi preussi **6**:256–258*T*
allochthonous inputs **3**:505
Allocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Allodontermes **4**:409*T*
Allogromia **6**:291–292
Alloherpesviridae **5**:134*T*, **5**:135
Allomyces **3**:627
allopatric speciation
  basic concepts **3**:144–145, **6**:621–624, **6**:621*F*, **6**:634
  dichopatric speciation **6**:621*F*, **6**:622–624, **6**:622*T*, **6**:623*F*
  general characteristics **6**:621–624, **6**:622*T*
  peripatric speciation **6**:621*F*, **6**:622*T*, **6**:624
  salamanders **6**:622–624, **6**:623*F*
allopolyploid/amphiploid speciation **4**:168, **6**:627, **6**:627*F*
Alloterposis **2**:376*F*

alloxydim sodium **5**:697
alluvial fans **2**:461, **2**:464*F*
almonds
  native crops **2**:662–663
  Near East ecosystems **5**:481–482
  seed protection **7**:26
Alnus **1**:628–630, **1**:630*T*
Alnus crispa **1**:630*T*
Alnus fruticosa **1**:630*T*
Alnus glutinosa **2**:551–552*T*, **2**:552*F*, **4**:173–174
Alnus hirsuta **1**:271
Alnus incana **1**:630*T*, **4**:173–174
Alnus rubra **3**:535
Alnus rugosa **1**:630*T*, **1**:632*T*, **7**:395–397
Alnus tenuifolia **1**:630*T*
Aloe dichotoma **1**:67, **1**:488–489, **2**:468, **2**:471*F*
Alopecurus pratensis **1**:269
Alopex lagopus **6**:240, **7**:27
Alosa sapidissima **3**:260, **5**:314–315
Alouatta fusca **3**:127*T*
Alouatta guariba guariba **6**:256–258*T*
Alouatta palliata mexicana **6**:256–258*T*
Alouatta palliata trabeata **6**:256–258*T*
Alouatta palliate **2**:587*F*
Alouatta pigra **6**:256–258*T*
Alouatta ululata **6**:256–258*T*
alpacas **1**:644*F*, **2**:475–476, **2**:476*T*
alpha diversity **1**:403, **2**:439–440, **2**:650, **4**:339–340, **6**:534, **6**:534
Alphaflexiviridae **5**:130*F*
alpine botanical gardens **6**:505
alpine butterflies **4**:516–517
alpine ecosystems **1:148–157**
  alpine life zone **1**:148–149, **1**:149*F*, **1**:149*T*
  animals and microorganisms
    animal diversity **1**:154–155, **1**:155*F*
    microbial diversity **1**:155, **1**:155*F*
  bioclimatic mountain belts **1**:149*T*
  biodiversity functions **1**:155–156
  global area **1**:149*T*
  global change impacts **1**:156–157, **1**:156*F*
  habitat diversity
    climatic microhabitats **1**:149–150, **1**:150*F*
    substrate diversity **1**:150–151
  high-elevation Andean ecosystems **4:96–110**
    background information **4**:96
    geographic distribution **4**:96–97, **4**:97*F*
    habitat characteristics **4**:96–97
    number of species **4**:102*T*
    páramo ecosystem
      characteristics **4**:101–103
      geographic distribution **4**:96–97, **4**:97*F*
      northern Andean alpine cushion formations **4**:103
      northern Andean dwarf *Polylepis* forest **4**:103
      number of species **4**:102*T*
      physical characteristics and patterns **4**:97–98
      plant diversity **4**:98–99, **4**:99*T*
      representative species **4**:102*F*
      subpáramo zone **4**:102–103
      superpáramo zone **4**:103
      true páramo zone **4**:103
    physical characteristics and patterns **4**:97–98
    plant diversity
      endemism **4**:99–100, **4**:99*T*
      origins and diversification **4**:100–101
      species richness **4**:98–99
    puna
      central Andean dwarf *Polylepis* forest **4**:105
      cushion bogs **4**:105
      desert puna **4**:104–105
      dry puna **4**:104
      geographic distribution **4**:96–97, **4**:97*F*
      humid puna **4**:104
      number of species **4**:102*T*
      physical characteristics and patterns **4**:97–98, **4**:103–104
      plant diversity **4**:98–99
      representative species **4**:104*F*
      salar vegetation **4**:105
    research summary **4**:107
    southern Andean steppes
      altitudinal zonation **4**:106–107, **4**:106*F*
      bogs and seeps **4**:107
      cushion plants **4**:106–107, **4**:106*F*
      geographic distribution **4**:96–97, **4**:97*F*
      latitudinal gradients **4**:106–107
      number of species **4**:102*T*
      physical characteristics and patterns **4**:97–98, **4**:105–107
      plant diversity **4**:98–99
      representative species **4**:105*F*
  plant diversity
    leaf mass ratio (LMR) **1**:152*F*
    morpho-type diversity **1**:151, **1**:151*F*
    physio-type diversity **1**:151–153
    reproductive diversity **1**:153
    specific leaf area (SLA) **1**:152, **1**:152*F*
    taxonomic diversity **1**:153–154, **1**:154*F*
  see also Antarctica
alpine life zone **1**:148–149, **1**:149*F*, **1**:149*T*
alpine mats **1**:274–275
alpine newt **4**:516–517
alpine snowbed herb **4**:516
Alpinia **1**:263
Alps
  altitudinal trends **1**:154*F*
  animal diversity **1**:155*F*
  leaf mass ratio (LMR) **1**:152*F*
  specific leaf area (SLA) **1**:152*F*
  vegetation surveys **3**:341–344*T*
  see also alpine ecosystems
Alsea Valley Alliance v. Evans (2005) **1**:671
Altai highlands **1**:262–263*T*
Alternanthera **2**:376*F*
Alternanthera philoxeroides **1**:165, **3**:703–704, **3**:704*F*, **4**:361
alternative agriculture **1**:112
Alticinae **1**:328
Altingiaceae **5**:622*F*
Altiplano **4**:97–98
Altman, Sidney **5**:631–632
Alucitoidea **5**:387–388*T*
aluminum (Al)
  aluminum mobilization **1**:5–6
  elemental partitioning **1**:554*T*
  forest ecosystems **1**:5–6
  freshwater ecosystems **1**:9–10
  relative abundance **1**:553*F*
Alvarez theory **3**:418–421
Alvernaviridae **5**:134*T*, **5**:138
Alvinella pompejana **7**:329
Alvinellidae **7**:417*F*
Alytes obstetricans **4**:664
Alzheimer's disease **1**:592
amago salmon **6**:522
'Amak-ihi **1**:29*F*
Amanita muscaria **3**:635*T*
Amanita ocreata **3**:635*T*
Amanita pantherina **3**:635*T*
Amanita phalloides **3**:635*T*
Amanita verna **3**:635*T*
Amanita virosa **3**:634, **3**:635*T*
amapá doce **4**:287–290
amaranth **2**:663–664
Amaranthaceae **2**:371*T*, **2**:374*T*
Amaranthus **1**:120, **2**:376*F*
Amaranthus caudatus **2**:663–664
Amaranthus cruentus **2**:663–664, **4**:288*T*
Amaranthus retroflexus **2**:368*F*
Amathusiinae **1**:653
amatoxins **3**:635*T*
Amazona vittata **6**:212, **7**:184–185
Amazon Basin **3**:105–107, **3**:106*F*
Amazon ecosystems **1:158–168**
  Amazonian savannas
    cerradão **1**:164–165
    flooded savannas **1**:165
    nonflooded savannas **1**:164–165
  background information **1**:158
  Biological Dynamics of Forest Fragments Project (BDFFP) **4**:663
  deforestation **4**:663, **6**:338–339
  ecosystem function measurements **3**:81*T*
  ecotones **3**:143*F*
  El Niño Southern Oscillation (ENSO) **3**:156–158
  floodplain ecosystems
    buritizal **1**:164
    igapó forest **1**:163–164
    lakes **1**:164
    mangrove forests **1**:164
    permanent swamps **1**:164
    pirizal **1**:164
    seasonal flooding **1**:162–163
    várzea forest **1**:163
  indigenous peoples **4**:273*F*, **4**:273*T*
  mammals **2**:591*F*
  montane fringing ecosystems
    Andes foothills **1**:166–167
    campos rupestres **1**:167
    characteristics **1**:165–166
    granite inselbergs **1**:166
    Guayana Highland **1**:166
    low hills **1**:166

Amazon ecosystems (*continued*)
sandstone tepuis **1**:166
myriapods **1**:246*T*, **1**:247, **5**:431
nonflooded terra firme ecosystems
cerradão **1**:161–162
dry semideciduous southern fringe forests **1**:161–162
general discussion **1**:158–160
rain forests **1**:158–160, **1**:159*F*, **1**:159*T*
white sand oligotrophic formations **1**:160–161
plant domestication **4:279–292**
basic concepts **4**:279
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
research summary **1**:167–168
riverine biota **6**:479
secondary forest **1**:167
terrestrial arthropods **1:245–253**
biotopes **1**:249–250
future research areas **1**:252
genetic characteristics **1**:251–252
habitat specificity **1**:249–250, **1**:250*T*, **1**:251*F*, **1**:252*F*
morphological characteristics **1**:250–251
number of species **1**:245, **1**:246*T*
origins and geographical distribution
Andes **1**:247
Carabidae **1**:247, **1**:250*T*
Geophilomorpha **1**:247
Meinertellidae **1**:247
neotropics **1**:246–247
nonflooded and/or floodplain forests **1**:247
Polydesmida **1**:247
Scolopendromorpha **1**:246–247
Scorpiones **1**:247
treeless floodplains **1**:247
species richness **1**:247–249, **1**:248*F*
taxonomy **1**:245, **1**:246*T*
threatened plant species **6**:84
transition ecosystems
babassu palm forests **1**:162
bamboo forests **1**:162
characteristics **1**:162
liana forests **1**:162
vascular epiphytes **3**:520*T*
Amazonian dark earth **4**:285–287, **4**:286*F*
Amazonian manatee **5**:81*T*
Amazon river dolphins **3**:192–193, **5**:65*T*
ambiregnal organisms **5**:554–555
Amblyopsis rosae **7**:55*F*
Amblypygi **1**:202*T*, **1**:205*F*, **1**:211, **1**:211*F*, **1**:246*T*
Amborella **5**:621*F*, **6**:19–20, **6**:20*F*
Amboseli National Park **5**:235
Ambrosia dumosa **2**:468–469
Ambulacraria **7**:434*F*
Ambystoma macrodactylum **4**:516
Ambystoma tigrinum **3**:166, **3**:166
Ambystoma tigrinum melanostictum **4**:515
Amebelodon **5**:224*T*
amebiasis **5**:671*T*
amebic dysentery **5**:660*T*
amebic meningoencephalitis **5**:660*T*
ameboid protozoa **6**:289–293, **6**:289*T*, **6**:290*F*
Ameiva festiva **6**:408
Ameiva leptophrys **6**:408
Ameiva quadrilineata **6**:408
Amelanchier **5**:218
Amentotaxus **1**:272–273
American Association for the Advancement of Science (AAAS) **3**:133–134
American bison **3**:188
American black duck **6**:689–690
American black oystercatcher **5**:57
American bullfrog **2**:529–530, **2**:530*F*, **3**:167
American burying beetle **3**:220–221, **3**:226
American chaffseed **3**:206
American chestnut **2**:528*T*, **2**:530–531
American elm **2**:528*T*
American golden plover **1**:605
American gray squirrel **2**:190
American grey squirrel **3**:191
American Livestock Breeds Conservancy **1**:645
American lobster **5**:464–465, **5**:464*F*
American mink **3**:192
American Museum of Natural History (AMNH)
Akeley Gallery of African Mammals **5**:406–407, **5**:407*F*
collection size and scope **5**:408*F*, **5**:409*F*
educational programs **3**:138
graduate and postdoctoral training **5**:416
historical perspective **5**:405–406
American opossums **4**:686–687, **4**:687*T*
American oyster **2**:264, **5**:61
American oyster drill **5**:61
American pika **2**:85
American robin **4**:174, **5**:723
American Samoa **4**:560*F*, **6**:507*T*
American shad **3**:260
American Southeast rivers **3**:184, **3**:185*F*
American toad **6**:153*F*
American Type Culture Collection (ATCC) **3**:652*T*
American water celery **6**:477–478
Amherstia nobilis **6**:77
Amia calva **3**:462–463
amino acids **5**:630–631, **6**:113*T*
2-aminopyrimidines **5**:695
Amitermes **4**:409*T*
Amitermes hastatus **4**:412–413, **4**:412*F*
Amitermitinae **4**:417*F*
amitraz **5**:694
amitrole **4**:90*T*, **5**:696
Ammi majus **5**:712*T*
Ammodorcas clarkei **1**:51
Ammodramus henslowii **4**:489–490, **4**:490*F*
Ammodramus nigrescens **6**:8
ammonia (NH_3)
ammonia-oxidizing bacteria (AOB) **5**:257–258
atmospheric concentrations **1**:282*T*
biogeochemical cycles **1**:560–561, **1**:562*F*
emission occurrences and sources **1**:1–5, **1**:137*T*
eutrophication **3**:366–367
geographic distribution **1**:3*F*
nitrogen cycle **5**:539*F*
nutrient cycling **1**:706
ammonia (NH_4)
eutrophication **3**:366–367
hydrothermal vent ecosystems **7**:328
ammonoids
adaptive radiation **1**:22*F*
ammonoid differentiation **1**:25
end-Cretaceous mass extinction **5**:173–175, **5**:174*F*
Late Devonian mass extinction **5**:170
Late Triassic mass extinction **5**:172–173
Ammophila arenaria **6**:96–97, **6**:121
Ammospiza maritimus nigrescens **7**:189
Ammotragus lervia **3**:193
Amniota **7**:339
amoebae **6**:586, **6**:587*F*
Amoebophrya ceratii **6**:295
Ampedus balteatus **7**:218–219
Ampedus tristis **7**:218–219
Ampharetidae **7**:417*F*
Amphibian Ark program **7**:478
amphibians **3:164–168**
alpine ecosystems **1**:154–155
anthropogenic activities and impacts **3**:401*F*
Asia
endemic species **1**:261*T*
threatened species **1**:278*T*
tropical rain forests **1**:263
background and characteristics **3**:164
biodiversity **1:169–178**
biogeographical models **1**:565–566, **1**:566*F*
conservation
as ecosystem components **1**:176
human uses **1**:176
major threats **1**:176–177
population declines **1**:177
current diversity **1**:169–170
evolutionary history **1**:169
functional anatomy **1**:170–171
future outlook **1**:177–178
geographic distribution **3**:691–693, **3**:692*F*
historical biogeography **1**:169–170
larval morphology and ecology **1**:173
life histories **1**:171
morphology **1**:170–171
postlarval behavior and ecology **1**:175–176
reproductive biology **1**:171, **1**:172

Central America **2**:25*T*
characteristics **7**:338–339
climate change studies **2**:66, **2**:92–93
conservation status **3**:164, **3**:164*T*
critical pH levels **6**:153*F*
desert ecosystems **2**:472–474
elevational range shifts **4**:605*T*
El Niño Southern Oscillation (ENSO) events **3**:156–158
endemism **3**:232*T*
environmental threats
  chemical pollution **3**:165–166
  climate change **3**:165
  commercial exploitation **3**:166–167
  disease **3**:166
  general discussion **3**:164–165
  habitat destruction **3**:164–165
  indicator species **4**:254–255
  introduced species **3**:167
  ultraviolet radiaion **3**:165
eutrophication impacts **3**:365–366
forest canopies **3**:513–514
fossil record **3**:688
freshwater ecosystems **3**:571–572, **3**:578–579, **3**:579*T*
genetic diversity **3**:665*F*
geographical variation **1**:614
global decline **3:691–699**
  causal factors
    exotic species invasions **3**:695–696
    general discussion **3**:695
    global climate change **3**:696–697
    habitat degradation/land conversion **3**:695
    human exploitation **3**:696
    infectious diseases **2**:527–529, **2**:528*T*, **2**:529–530, **2**:530*F*, **3**:349, **3**:696
    synergistic effects **3**:697
  ecological impact **3**:693
  future challenges
    captive breeding programs **3**:698
    conservation approaches **3**:697–698
    exotic species removal **3**:698
    habitat change mitigation **3**:697–698
    infectious agent attenuation **3**:698
  geographic extent **3**:693, **3**:694*F*
  global climate change
    elevated ultraviolet (UV–B) radiation **3**:697
    general discussion **3**:696–697
    temperature variability–disease immunity relationship **3**:697
    weather pattern shift effects **3**:697
  implications **3**:698
  phylogenetic investigations **3**:693–695
  systematic factors **3**:693–695
grassland ecosystems **4**:3
indicator species **4**:254–255, **7**:206–207*F*
infectious disease regulation **3**:349
introduced parasites–species decline relationships **2**:528*T*, **2**:529–530, **2**:530*F*
introduced species
  Australia **1**:80
  United States **1**:78
  major threats **7**:203*F*
  mangrove ecosystems **5**:13
  Mediterranean-type climate ecosystems
    Australia **5**:213*T*, **5**:221
    California **5**:213*T*
    Cape Region (South Africa) **5**:213*T*, **5**:220
    Chile **5**:213*T*, **5**:217
    Mediterranean Basin **5**:213*T*, **5**:218–219
  migratory patterns **5**:315–316
  modern extinctions **2**:168*F*, **2**:170*F*, **5**:368*T*
  Near East ecosystems **5**:460*T*
  North America **5**:565
  past, recent, and future extinction rates **2**:168*F*
  Permian–Triassic mass extinction **5**:4
  phenological shifts **5**:723
  population declines
    developmental deformities **3**:167–168
    protected habitats **3**:167–168
    ultraviolet radiation effects **7**:300–301
  river ecosystems **6**:479–480
  Southeast Asia **1**:515
  species richness gradients **4**:613*F*
  species richness maps **6**:709*F*
  Succulent Karoo **1**:487*T*, **1**:488
  surface water acidification **1**:10*T*
  threatened species **1**:176–177, **1**:515, **7**:201*T*, **7**:206–207*F*
  tourism industry **7**:223
  ultraviolet radiation effects
    deformities **7**:301
    growth and development **7**:300
    population declines **7**:300–301
  United States **1**:78
  urbanization threats **4**:564
  wetland ecosystems **7**:389–391
  wildfire effects **3**:156–158
Amphibolis **6**:542*T*
Amphibolis antarctica **6**:541*T*
Amphibolis griffithii **6**:541*T*
Amphidinium **1**:472, **1**:473*F*, **6**:294
Amphigyra alabamensis **3**:185*F*
Amphilinia **7**:456–457*F*
Amphilinidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Amphinomidae **7**:417*F*
Amphionidacea **2**:416–417*T*
amphioxus **3**:684–685
Amphipoda/amphipods
  classifications **4**:381*T*
  endangered species **3**:202*T*
  extinction vulnerability **2**:416–417*T*
  lake and pond ecosystems **4**:371–373, **4**:374*F*
  Lake Baikal **3**:183*F*
  mangrove ecosystems **4**:337
  marine sediments **5**:106*F*
  number of species **3**:182*T*, **4**:371*T*
  sandy beaches **2**:131–133
  seagrass ecosystems **6**:547–548
  subterranean ecosystems **7**:56–57, **7**:57*F*
Amphisbaenia **7**:340
Amphora coffaeformis **1**:475
Amphydasis cognataria **6**:3
amplified fragment length polymorphisms (AFLPs) **1**:477–478, **1**:479, **2**:274–275, **3**:663–664, **4**:510, **4**:515, **4**:519
Ampullaviridae **5**:130*F*
Amsinckia grandiflora **6**:441, **6**:455–456
Amsterdam albatross **1**:703
Amur leopards **3**:373
Amydrium **3**:522*T*
Amygdalus communis **2**:662–663, **5**:481–482
Amygdalus korschinskii **5**:485
Amygdalus ramonensis **5**:485
amyotrophic lateral sclerosis (ALS) **1**:592
Anabaena flos-aquae **6**:33*F*
Anabaena variabilis **2**:611*F*
Anabasis **1**:270, **2**:371*T*
Anabasis articulata **1**:268, **5**:485–486
Anabasis salsa **1**:270
Anabasis syriaca **5**:484–485
Anabropsis marmoratus **3**:731*F*
Anacanthotermes **4**:409*T*
Anacardium occidentale **4**:286–287, **4**:289*T*
Anadenobolus leucostigma martinicensis **5**:426*F*
anadromy **6**:526
anaerobic respiration, microbial **5**:301*T*
anagenesis
  basic concepts **2**:47–48, **2**:48*F*, **6**:620, **6**:621*F*
  definition **2**:47
  importance **2**:56–57
Anagyris foetida **1**:268
Anamirta cocculus **5**:712*T*
Ananas comosus **2**:663–664
Anancus **5**:224*T*
Anannas comosus **4**:287–290, **4**:289*T*
Anarthrophyllum cummingii **4**:106
Anasakis **7**:429–430
Anasa tristis **7**:294
Anas crecca **3**:365–366
Anaspidacea
  extinction vulnerability **2**:416–417*T*
  number of species **3**:182*T*
Anaspides tasmaniae **3**:181
Anaspis arctica **7**:218–219
Anas platyrhynchos **4**:362, **6**:689–690
Anas platyrhynchos diazi **6**:689–690
Anas rubripes **6**:689–690
Anas strepera **3**:365–366
Anas superciliosa superciliosa **4**:362
Anas wyvilliana **4**:362, **7**:184
Anatolepis heintzi **7**:334–335
Anatolian–Iranian desert **1**:262–263*T*
Anatrichus pygmaeus **5**:463*T*
anchialine communities **7**:58, **7**:59*T*
anchialine water **7**:50
Anchiceratops **2**:519*T*, **2**:520*T*
anchoveta **5**:688–690, **5**:690*F*, **6**:422*T*
ancient lakes **4**:459, **4**:463–464
ancient murrelets **5**:447
ancient polyploidy
  angiosperms **4**:169
  prevalence **4**:168–169
  tracheophytes **4**:169
Ancistrocladus korupensis **5**:713
Ancistrotermes **4**:409*T*

Ancylodiscoides **7**:448–449*F*
Ancylostoma **7**:422–424, **7**:426*F*
Ancylostoma caninum **1**:359, **7**:432–433, **7**:432*T*
Ancylostoma duodenale **7**:425–428, **7**:428*T*
Andaman Islands **1**:262–263*T*, **1**:275*T*
Andean Belt **3**:105–107, **3**:106*F*
Andean Biodiversity Strategy **2**:314
Andean Cordillera **5**:215–216
Andean flamingo **7**:183
Andean huemel **3**:193
Andean oak **4**:516–517
Andean titi **6**:256–258*T*
Andes
  Amazon ecosystems **1**:166–167
  edible plants **3**:130*T*
  high-elevation ecosystems **4:96–110**
    background information **4**:96
    geographic distribution **4**:96–97, **4**:97*F*
    habitat characteristics **4**:96–97
    number of species **4**:102*T*
    páramo ecosystem
      characteristics **4**:101–103, **7**:275–276
      geographic distribution **4**:96–97, **4**:97*F*
      northern Andean alpine cushion formations **4**:103
      northern Andean dwarf *Polylepis* forest **4**:103
      number of species **4**:102*T*
      physical characteristics and patterns **4**:97–98
      plant diversity **4**:98–99, **4**:99*T*
      representative species **4**:102*F*
      subpáramo zone **4**:102–103
      superpáramo zone **4**:103
      true páramo zone **4**:103
    physical characteristics and patterns **4**:97–98
    plant diversity
      endemism **4**:99–100, **4**:99*T*
      origins and diversification **4**:100–101
      species richness **4**:98–99
    puna
      central Andean dwarf *Polylepis* forest **4**:105
      cushion bogs **4**:105
      desert puna **4**:104–105
      dry puna **4**:104
      geographic distribution **4**:96–97, **4**:97*F*
      humid puna **4**:104
      number of species **4**:102*T*
      physical characteristics and patterns **4**:97–98, **4**:103–104
      plant diversity **4**:98–99
      representative species **4**:104*F*
      salar vegetation **4**:105
    research summary **4**:107
    southern Andean steppes
      altitudinal zonation **4**:106–107, **4**:106*F*
      bogs and seeps **4**:107
      cushion plants **4**:106–107, **4**:106*F*
      geographic distribution **4**:96–97, **4**:97*F*
      latitudinal gradients **4**:106–107
      number of species **4**:102*T*
      physical characteristics and patterns **4**:97–98, **4**:105–107
      plant diversity **4**:98–99
      representative species **4**:105*F*
  leaf mass ratio (LMR) **1**:152*F*
  nitrogen deposition **5**:526*T*
  plant species **6**:82*T*
  specific leaf area (SLA) **1**:152*F*
  terrestrial arthropods **1**:247
  vascular plant diversity **6**:82*T*
  see also alpine ecosystems
Andesianoidea **5**:386, **5**:386*F*, **5**:387–388*T*
Andorra
  botanical gardens **6**:507*T*
  urbanization **4**:560*F*
  vegetation maps **3**:346*T*
  vegetation surveys **3**:341–344*T*
Andrias davidianus **3**:373–374
Andricus californicus **4**:173–174
Androctonus spp. **1**:208
Andromeda **1**:277–278
Andromeda polifolia **1**:136
Andropadus virens **3**:144–145
Andropogon **2**:376*F*
Andropogon gayanus **6**:93
Andropogon gerardi **3**:437–438
Androstachys **5**:623*F*
andthiocyclam **5**:695
Anelloviridae **5**:134*T*, **5**:136
Anelosimus studiosus **5**:394
Anemone **4**:101
anemonefish **6**:574
Anemone patens **1**:269
anemones
  photosynthetic symbioses **5**:118–120, **5**:119*F*
  sandy beaches **2**:131–133
Aneurolepidium chinense **1**:269
Aneurophyton **5**:618*F*
Angarozonium amurense **5**:428
angelfish **3**:467–471, **3**:471*F*
angelim pedra **1**:160
angiosperms **5:613–627**
  abundance patterns **6**:73, **6**:74*F*
  adaptive radiation **1**:430
  ancient polyploidy **4**:169
  Arctic ecosystems **1**:231*T*
  Asia **1**:258, **1**:259*T*
  background information **5**:613
  beetles **1**:23, **1**:328–329
  biome distributions **1**:64*F*
  characteristics **6**:60
  diversification
    analytical methodologies **5**:619–620
    historical patterns **5**:625–626
    major events
      ASTERID I and ASTERID II Clades diversification **5**:622*F*, **5**:625
      ASTERID III and ASTERID IV Clades diversification **5**:622*F*, **5**:625
      early diversification events **5**:620–621, **5**:621*F*, **5**:622*F*
      general discussion **5**:620–621
      Rosid 1 Clade diversification **5**:621–625, **5**:623*F*, **5**:624*F*
    molecular clock-based models **5**:619–620
    research background **5**:614
  evolutionary development **2**:643, **6**:65–67, **6**:66*F*
  extinction rates **6**:73*F*
  insect-pollination angiosperm-radiation hypothesis **6**:48
  North America **5**:564–565
  number of species **6**:58*T*
  origination rates **6**:71–72, **6**:72*F*
  origins
    fossil record
      challenges **5**:615
      early angiosperms **5**:617–619
      general discussion **5**:615–617
      nineteenth century research **5**:613–614
      Phanerozoic **3**:686–687, **3**:687*F*
      potential angiosperm sister groups **5**:615–617, **5**:616*F*, **5**:617*F*, **5**:618*F*
    phylogenetic and morphological studies **5**:614–615, **5**:618*F*
    research background **5**:613–614
    research summary **5**:619
  Southern Hemisphere ecosystems **6**:617–618
  taxonomic distributions **3**:519*T*
anglerfish **3**:466–467, **3**:475*F*, **5**:122*F*, **5**:123
Angola
  botanical gardens **6**:519
  education level–fertility relationship **6**:204, **6**:205*F*
  tropical forests **7**:269–270, **7**:270*T*
  urbanization **4**:560*F*
  woodland savannas **1**:67–68
Angraecum cadetii **2**:328–329
Angraecum sesquipedale **2**:138, **2**:328–329, **6**:47–48
Anguilla **6**:519
Anguilla anguilla **5**:473
Anguina **7**:430
Anguina tritici **5**:672–673, **7**:429
Angulitermes **4**:409*T*
angwantibos **6**:251*T*
'Anianiau **1**:29*F*
Aniara sepulcralis **1**:250*T*, **1**:251*F*
anilazine **5**:695
anilides **5**:696
animal damage control **6**:342
animal fables **4**:653
animal parks **3**:137–138
Animal Quantitative Trait Locus database **1**:647
Animal Records Keeping System (ARKS) **7**:475
animal rights **3**:313–314
animal species
  alpine ecosystems **1**:154–155, **1**:155*F*
  animal breeding **3:100–107**
    challenges **1**:648
    conservation planning **1**:639–640
    conservation strategies

ex situ strategies **1**:645–646
general discussion **1**:642–645
in situ strategies **1**:642–645
crossbred animals **1**:639
current trends **1**:638–639
diversity challenges **1**:636
future outlook **1**:648
genetic structure **1**:637–638
global breeds **1**:638–639
livestock breeds **1**:637
management strategies
conservation strategies **1**:642–645
Food and Agriculture Organization (FAO) programs **1**:640–642, **1**:641
inventory, description, and characterization assessments **1**:641–642, **1**:642*F*
risk status classification **1**:641–642, **1**:642*F*, **1**:642*T*, **1**:643*F*, **1**:644*F*
new technologies
cloning **1**:647–648
cryopreservation **1**:646
gene mapping **1**:646–647
molecular characterization **1**:646–647
wild relative use **1**:640
animal pollination **5**:354–355
Arctic ecosystems **1**:231*T*
Asia **1**:258, **1**:261*T*
boreal forests **1**:631–632, **1**:633*F*
C_4 plants **2**:377–378
climate change studies **2**:66, **2**:67*F*, **2**:92–94, **2**:93*F*
conservation practices **2**:281–282, **4**:321*T*
density–biomass thinning laws **6**:161, **6**:165, **6**:166*F*
desert ecosystems **2**:471*F*, **2**:472–474, **2**:472*F*, **2**:473*F*, **2**:474*F*
disturbance ecology **2**:562–563, **2**:565–566
diversity–productivity relationship **2**:597–599, **2**:598*F*
dry Asian regions **1**:268
El Niño Southern Oscillation (ENSO) events **3**:156–158
fire/fire regimes **3**:438
food habits **3**:127–128, **3**:127*T*
forest canopies **3:511–515**
canopy architecture **3**:511–512
invertebrates **3**:512–513, **3**:513*F*, **3**:514*F*
research background **3**:511
research summary **3**:514–515
research techniques **3**:512, **3**:512*F*
vertebrates **3**:513–514
fossil record **3**:687–688, **3**:687*F*
gene banks **1**:645–646, **3**:651–652
global warming effects
helical animal burrows **4**:229
tropical animal geographic expansion **4**:229
global wildlife trade **3**:409*T*
grasslands **4**:3–4, **6**:331–333
guilds **4:33–38**
characteristics and structure **4**:33–35, **4**:34*F*
ecological boundaries **4**:37
predator guilds **1**:690–692
research benefits
functional community organization **4**:35–36, **4**:36*T*
manageable unit size **4**:35
trophic structure framework **4**:36
species individualism **4**:36–37
vaguely bounded categories **4**:37–38
Holocene postglacial spread **2**:550, **2**:551–552*T*, **2**:552*F*
hot deserts **6**:334
inbreeding depression **4**:248
introduced plants **1**:76
introduced species **3**:410–411, **3**:410*F*, **6**:444–445
living dead **4**:594–595
Madagascar **1**:55–56
mangrove ecosystems
intertidal gradient **4**:337, **4**:339*F*
marine fauna
bivalves **5**:17
characteristics **5**:14–15
crustaceans **5**:15–16, **5**:20*T*
ecosystem engineers **5**:16–17
fish **5**:15, **5**:15*F*, **5**:20*T*
grapsid crabs **5**:15–16, **5**:15*F*
hermit crabs **5**:16
meiofauna **5**:18
molluscs **5**:16–17, **5**:20*T*
mud crabs **5**:16
ocypodid crabs **5**:16, **5**:16*F*
root communities **5**:14–15
shrimp **5**:16
snails **5**:17–18, **5**:17*F*
terrestrial fauna
amphibians **5**:13
birds **5**:13–14, **5**:20*T*
characteristics **5**:12–13
insects **5**:13, **5**:13*F*
mammals **5**:14, **5**:20*T*
reptiles **5**:13, **5**:20*T*
modern extinctions **5**:368*T*
neoproterozoic animals **3**:682–683
neotropical seasonally dry forests **5**:490
nitrogen fixation **1**:560–561, **1**:562*F*, **6**:581–582, **6**:582*F*
oil-palm plantations **5**:604–605
past, recent, and future extinction rates **2**:168*F*
Phanerozoic diversity **3**:687–688, **3**:687*F*
plant–animal interactions **6:39–55**
antagonistic interactions
carnivorous plants **6**:44–45
coevolution **6**:42–43
community and ecosystem effects **6**:43–44
evolutionary responses **6**:41–42
plant consumers **6**:39–41
carnivorous plants
benefits **6**:45
fitness costs **6**:45
general discussion **6**:44–45
nonprey guests **6**:45
prey capture mechanisms **6**:44–45, **6**:44*T*
chemical defenses
digestibility reducers **6**:41
plant secondary chemicals **6**:42
repellents **6**:42
toxins **6**:41
classifications **6**:39, **6**:39*T*, **6**:40*F*
evolutionary responses
chemical defenses **6**:41–42
compensatory growth **6**:42
escape defenses **6**:42
mechanical defenses **6**:42
tolerance mechanisms **6**:42
fruit and seed dispersal
ants **6**:50
beetles **6**:50
birds **6**:50
dispersal syndromes **6**:50, **6**:51*T*
disperser types **6**:50
evolutionary dead ends **6**:52
fish **6**:50
frugivores **6**:50–51
importance **6**:49–50
insects **6**:50
mammals **6**:50
plant adaptations **6**:50
reptiles **6**:50
seed predators **6**:51–52
species persistence **6**:61–62
vertebrates **6**:50
mutualism
cheating studies **6**:45–46
fruit and seed dispersal **6**:49–50, **6**:61–62
plant–pollinator interactions **2**:474–475, **2**:475*F*, **6**:47
plant-protecting ants/plant-feeding ants **6**:46
specialization **6**:45–46
research summary **6**:52
trophic levels **6**:40*F*
plant defense system uses **2**:432
restoration efforts
basic concepts **6**:443–444
habitat restoration **6**:445–446
population sources **6**:446
site assessments **6**:445–446
species reductions or absences **6**:444–445
savannas **6**:331–333
seed dispersal **7**:283–284
speciation–extinction rate relationship **5**:439–440, **5**:440*F*
symbiotic relationships **5**:306
temperate/cold-desert shrublands **6**:333–334
temperate grasslands **1**:269
tropical Africa **1**:48–49
tropical rain forests **6:304–312**
Asia **1**:263
beetles **6**:306
birds **6**:307–309, **6**:308*T*
butterflies **6**:306–307, **6**:307*T*
frogs **6**:307
future outlook **6**:311–312
general characteristics **6**:306
geographical patterns **6**:309–310

animal species (*continued*)
geographic distribution **6**:304–305
high species richness **6**:310–311
historical perspective **6**:304
mammals **6**:309–310, **6**:309*T*
tundra ecosystems **6**:330–331
wetland ecosystems **7**:389–391, **7**:389*F*, **7**:390*F*, **7**:393–394, **7**:396*F*
wildfire effects **3**:156–158
see also livestock
animal welfare **7**:472–473
animism **6**:563–564
Aningeria altissima **1**:68
anisakis **5**:660*T*
Anisopodidae **3**:489–491*T*
Anisopteromalus calandrae **4**:306–307
Ankarana sportive lemur **6**:256–258*T*
Ankistrodesmus falcatus **6**:33*F*
Anklyosauria **2**:520*T*
Ankylosaurus **2**:519*T*, **2**:520*T*
Annamite region **4**:703–704
Annelida **7**:**408–419**
adaptability **4**:393
Antarctica **1**:184*T*
Census of Marine Life project **2**:7*F*
characteristics **4**:392, **4**:394*F*
classifications **7**:408
fossil record **3**:684*F*
geographic distribution **7**:418
global diversity **4**:385*T*
known and estimated species **4**:384*T*
lake and pond ecosystems **4**:461*T*
marine ecosystems **5**:50*T*, **5**:106*F*
morphology
anterior ends **7**:410*F*, **7**:412*F*
chaetae **7**:409–410, **7**:414*F*
circulatory system **7**:415
digestive tract **7**:415
general characteristics **7**:408–409
gills **7**:415
gonads **7**:415–416
head **7**:409
musculature **7**:414–415
nephridia **7**:415
nervous system **7**:410–411
posterior ends **7**:410*F*
pygidium **7**:409
segments **7**:409, **7**:413*F*
sense organs **7**:411–414
septa **7**:415
number of species **4**:371*T*, **5**:669–670*T*
phylogenetic and taxonomic relationships **7**:416–419, **7**:417*F*
phylogenetic relationships **7**:434*F*
Proterozoic–Cambrian transition **3**:396*F*
reproductive biology **7**:416
river ecosystems **6**:478*T*
wetland ecosystems **7**:389–391
Annonaceae **3**:129*T*, **6**:316*T*
Annona muricata **4**:289*T*
Annona palustris **1**:164
Anobiidae **1**:327*T*
Anodorhynchus hyacinthinus **7**:187
Anodorhynchus leari **7**:187
Anogeissus latifolia **1**:69–70, **1**:267
Anogeissus leiocarpus **1**:67
Anogramma ascensionis **6**:78
Anolis gingivinus **2**:187–188, **2**:187*T*
Anolis lizards
adaptive radiation **1**:31*F*, **1**:32
differentiation **2**:513
interspecific competition **2**:187–188, **2**:187*T*, **2**:533
Anolis wattsi **2**:187–188, **2**:187*T*
Anomalospiza imberbis **5**:502
Anomosetidae **5**:386
Anomospermum occidentale **1**:166–167
Anomura **2**:416–417*T*
Anona cherimola **2**:663–664
Anoplodera virens **7**:218–219
Anoplolepis gracilipes **3**:222, **5**:596–597, **7**:260
Anoplolepis longipes **6**:240–241
Anoplotermes **4**:409*T*
Anoplura **4**:392*T*
Anostostomatidae (wetas) **3**:731, **3**:731*F*
Anostraca **2**:416–417*T*, **3**:182*T*
anoxia **3**:351, **5**:543, **7**:394–395
Anser albifrons **1**:710
Ansiruptodera **7**:426*F*, **7**:427*F*
Antarctica
biological history **1**:179–180
Census of Marine Life project **2**:20–21
Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
Diptera **3**:488*T*
ecosystems **1**:**179–188**
anthropogenic impacts
alien species invasion **1**:185–186, **1**:186*T*
biological consequences **1**:186–187
historical perspective **1**:185–186
regional and global environmental change effects **1**:186
biodiversity
current studies **1**:183–184
origins and antiquity **1**:184
plants **1**:184*T*
terrestrial invertebrates **1**:184*T*
climate change effects **5**:89–91, **6**:617
freshwater ecosystems **1**:183
future trends
anthropogenic impacts **1**:185–186
climate change effects **1**:186–187
natural colonization processes
colonization mechanisms **1**:185
contemporary colonization **1**:184–185
propagule dispersion **1**:185
terrestrial biogeographic zones
continental Antarctic zone **1**:181, **1**:181*T*, **1**:182*T*, **1**:184*T*, **1**:186*T*
general characteristics **1**:180–181, **1**:180*F*
maritime Antarctic zone **1**:181, **1**:181*T*, **1**:182*T*, **1**:184*T*, **1**:186*T*
sub-Antarctic zone **1**:180–181, **1**:181*T*, **1**:182*T*, **1**:184*T*, **1**:186*T*
temperature ranges **1**:181*T*
terrestrial biota
fauna **1**:182, **1**:182*T*
flora **1**:182–183, **1**:182*T*
habitats **1**:181–182
microbiota **1**:182*T*, **1**:183
fish diversity **3**:474
geographic characteristics **1**:179, **1**:180*F*
glacial history **1**:179–180
mass extinctions **5**:6–7
protected area systems **6**:491*T*
psychrophiles **6**:299
tectonic history **1**:179–180
ultraviolet radiation effects **7**:300
Antarctic fur seals **5**:87*T*
Antarctic ozone hole **1**:288–289
antbirds **2**:180–181, **2**:181*F*
anteaters **2**:584*F*, **2**:590*T*, **4**:689, **4**:690*T*, **4**:701*T*, **7**:229*T*
antechinus **4**:687*T*
Antechinus **3**:188
Antechinus stuartii **4**:688
antelope
mangrove ecosystems **5**:14
number of species **4**:690*T*
species diversity **4**:696
tropical Africa **1**:49
Antelopinae **4**:696
Antennaria chilensis **4**:107
Anthaenantia **2**:376*F*
Antheraea polyphemus **5**:399*F*
Anthicidae **1**:327*T*
Anthocharis cardamines **6**:107
Anthomyiidae **3**:489–491*T*, **3**:493–494
Anthomyzidae **3**:489–491*T*
Anthozoa **3**:681*F*, **3**:685*F*, **5**:669–670*T*
anthracnose blight **2**:528*T*
Anthracotherriidae **5**:224*T*
Anthribidae **1**:327*T*
Anthropocene epoch **4**:146
anthropocentrism **3**:269
anthropogenic soils **4**:285–287, **4**:286*F*
Anthurium **6**:316*T*
Anthus spp. **2**:556, **2**:556*F*
Anthyllis cytisoides **1**:707–708
antibiotics **3**:638, **5**:694
antidiarrheal drugs **1**:594
Antidorcas australis **5**:364–365
Antidorcas marsupialis **1**:487–488
Antigua **6**:519
Antilocapra **4**:724–725*T*
Antilocapra americana **2**:503–505, **3**:256–257, **4**:696, **6**:332
antilocarpids **4**:696
Antilope cervicapra **7**:229*T*
antimalarial drugs **1**:593–594
antler moth **6**:159*F*
antlions **4**:392*T*
ant lions **3**:489–491*T*
ant-loving crickets **3**:730
Antonia ovata **1**:165
Antroherpon apfelbecki **7**:60
Antrolana lira **3**:185*F*
ants
alien species invasion **5**:596–597, **5**:598*F*
Amazon ecosystems **1**:246*T*
ant-associated epiphytes **3**:523–524
ant–plant interactions
benefits to ants/cost to plants **6**:46
benefits to plants **6**:46–47

desert ecosystems **2**:470*T*, **2**:474–475
intraguild interference **1**:378*F*
myrmecophytes **6**:46
myrmecotrophic epiphytes **6**:47
butterfly–ant symbiotic relationship
acoustic communication **1**:656
association categories **1**:656
food secretions **1**:655
myrmecophily **1**:655
semiochemical production **1**:655–656
characteristics **4**:392
domesticated ruminants **1**:709–710
fruit and seed dispersal **6**:50, **6**:51*T*
habitat–bird species diversity relationship **1**:612, **1**:613*F*
interspecific competition **2**:186–187, **2**:187*T*
invasive species **3**:222, **7**:260
kin selection theory **2**:322
leaf-cutting ants **4**:181
mangrove ecosystems **5**:13, **5**:13*F*
mutualism **2**:146, **4**:365
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*
number of species **4**:392*T*
oil-palm plantations **5**:604–605
plant tolerance–grazing relationship **1**:708, **2**:428–429, **6**:46–47
pollination **6**:142
social systems **2**:646
soil fauna **6**:590*T*
special associations **3**:495
anurans
characteristics **7**:339
evolutionary history **1**:169
forest canopies **3**:513–514
functional anatomy **1**:170–171
larval morphology and ecology **1**:174–175
life histories **1**:172
migratory patterns **5**:315–316
morphology **1**:170–171
Near East ecosystems **5**:464*T*
postlarval behavior and ecology **1**:175–176
reproductive biology **1**:172, **1**:173*F*
special features **1**:171
species richness **1**:170*T*
Anurogryllus muticus **3**:729
Aotidae **4**:692, **6**:250–252, **6**:251*T*
aoudad **3**:193
Apapanes **1**:29, **1**:29*F*
Apatornis **2**:519*T*
Apaturinae **1**:653
apes **4**:690*T*
apete **7**:232, **7**:232*F*
apex predators **4**:670–671, **7**:262–263
Apfelbeckia insculpta **5**:425
Aphanizomenon **6**:30
Aphanizomenon flos-aquae **6**:33*F*
Aphanomyces astaci **4**:378
Aphelinus semiflavus **4**:362
Aphelocoma coerulescens **7**:186–187
Aphelops **5**:224*T*
aphid flies **3**:489–491*T*
aphids **4**:298–299, **4**:361, **4**:392*T*, **5**:316
Aphroditidae **7**:417*F*
Apiaceae **3**:129*T*
Apicomplexa **5**:669–670*T*
apicomplexans **6**:296–297
Apicotermes **4**:409*T*
Apicotermitinae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:430*T*
Apioceridae **3**:489–491*T*
Apion **3**:512–513
Apionidae **1**:327*T*
Apis melifera scutellata **2**:551–552*T*
Apis mellifera
biodiversity–ecosystem services relationship **1**:348, **3**:347
introduced species **6**:144–145
recombination characteristics **6**:361*T*
reproductive behavior **4**:178–179
see also honeybees
Apistobranchidae **7**:417*F*
Apium graveolens **2**:383, **2**:662–663
Aplacophora **5**:376, **5**:376*F*
Aplodinotus grunniens **3**:469
Apluda **1**:70
Aplysina fistularis **5**:124*F*
APOBEC-1 **5**:577
Apocrita **4**:177–178
Apocynaceae **1**:24*F*, **3**:129*T*, **5**:489, **6**:316*T*
Apodanthaceae **4**:72, **4**:72*F*
Apodanthes **4**:72
Apodemus mystacinus **3**:674*F*, **5**:466
Apodemus sylvaticus **1**:367, **5**:466
Apodidae **6**:308*T*
Apodiformes **6**:308*T*
Apoditrysia **5**:386*F*, **5**:387–388*T*
apoendemics **3**:228
Apoidea **4**:178
apolipoprotein B **5**:577
apomorhpy **2**:33*F*, **2**:34–35
Apoplanesia paniculata **2**:26
Apothecaries Garden **6**:508
Appalachian Mountains **5**:562–563
Appalachian Orogeny **5**:562–563
apparent competition **6**:245–246
apple maggot fly **2**:140, **2**:725
apples **2**:662–663, **3**:32*F*
Approved List of Bacterial Names **5**:551–552
apricots **2**:383, **2**:384*F*, **2**:663, **3**:32*F*
Apteribis **4**:730–731
Apteryx **1**:606
Apteryx oweni **5**:366–367
Aptornis otidiformis **5**:366
Apuleia leiocarpa **1**:161–162
Apusomonas **6**:294
aquaculture **1:189–201**
background and characteristics **1**:190–191
biodiversity impacts **1**:190, **1**:191*F*, **3**:448–449, **3**:452, **4**:144, **7**:203*F*
carrying capacity **1**:196–198
chemical impacts **1**:195
disease impacts **1**:195–196
ecological footprint analysis (EFA) **1**:196–198, **1**:196*F*
ecological relationships **1**:192*F*
energy requirements **1**:194
farming systems and production methods **1**:190*T*
feed resources **1**:191–193, **1**:192*F*
feral animals **1**:194–196
functional role **1**:189–190
harvest, production, and consumption **1**:191–193, **1**:192*F*
integrated aquaculture **1**:198
land requirements **1**:193–194
life cycle analysis (LCA) **1**:196–198, **1**:197*F*
parasitic disease vectors **5**:672, **5**:672*F*
research summary **1**:198–199
resource use **1**:196–198
species groups **1**:190*T*
waste impacts **1**:194–196
water requirements **1**:194
wild fry capture **1**:192*F*, **1**:194
world fish catch **6**:420–421, **6**:420*F*
see also fish
AquaMaps initiative **3**:484–485, **3**:485*F*, **3**:486*F*
aquarium fish **2**:343–344, **4**:156
aquariums **2**:245
aquatic Diptera
freshwater habitats **3**:494–495
phytotelmata **3**:494–495
aquatic ecosystems
acidic and mercury deposition **1:1–14**
acidic deposition–mercury concentration linkage **1**:13
aluminum concentrations **1**:9–10
aquatic biota response **1**:9–10, **1**:10*T*, **1**:11*F*
emission occurrences and sources **1**:1–5
geographic distribution **1**:1–5, **1**:3*F*
mercury contamination **1**:10–13, **1**:11*F*, **1**:12*F*
mercury emissions and deposition **1**:2–5, **1**:4*F*
pH changes **1**:9–10, **1**:10*T*
research background **1**:1
schematic diagram **1**:2*F*
surface water acidification **1**:8–9, **1**:10*T*, **1**:144–145
biological impoverishment **6**:335–337, **6**:336*F*, **6**:341–342
international organizations **4**:329–330
nematodes **7**:429, **7**:431
nitrogen cycle **6**:152, **6**:153*F*
pH levels **1**:9–10, **1**:10*T*, **6**:152, **6**:153*F*
protected area systems **6**:494–498, **6**:496*F*, **6**:497*F*
species distribution modeling (SDM) **6**:700
trophic cascades **7**:259–260
ultraviolet radiation effects
amphibians **7**:300–301
fish **7**:300–301
macroalgae **7**:299–300
marine invertebrates **7**:300
plankton **7**:299
research background **7**:299
seagrass **7**:299–300
aquatic ferns (*Salvinia molesta*) **4**:346–347, **4**:347*F*
aquatic mites **1**:215
aquatic plants **3**:359–360, **4**:378

aquatic subterranean habitats **7**:49–50, **7**:50*F*
aquatic vascular plants **3**:359–360
aquatic vegetation **1**:292–297
aquatic warbler **7**:182
aquifers **1**:224*T*
Aquifoliaceae **1**:270–271
Aquila **1**:709
Aquila adalberti **7**:185, **7**:262–263
Aquila chrysaetos **1**:709
Aquila chrysaetus **5**:474
Aquila clanga **5**:474
Aquila verreauxi **5**:474
Aquilegia formosa **6**:47–48
Aquilegia pubescens **6**:47–48
Aquilegia sibirica **1**:269
Arabian desert **1**:262–263*T*, **2**:469–470
Arabian oryx **2**:243–244, **2**:245, **3**:196
Arabian peninsula **1**:259*T*
Arabidopsis lyrata **2**:636*F*
Arabidopsis thaliana
  molecular level diversity **2**:636*F*
  mutation rates **2**:628*T*
  parasite–host coevolution **2**:145
  phenotypic plasticity **2**:721–722
  recombination characteristics **6**:361*T*
  RNA editing **5**:575–576
arabinose **6**:113*T*
Arabis alpina **4**:519
Arabis hirsuta **6**:458
arable land **6**:592*F*
araçá **4**:286–287
Araceae **3**:129*T*, **3**:522*T*, **6**:313–318, **6**:316*T*
Arachis hypogaea **2**:389–390, **2**:663–664, **4**:287–290, **6**:553
Arachis hypogeae **4**:289*T*
Arachnida/arachnids **1:202–217**
  Acari
    characteristics **1**:213–215, **1**:214*F*, **4**:394*F*
    ecological characteristics **1**:214–215
    genera and species **1**:202*T*
    major lineages
      Acariformes **1**:216–217
      Astigmata **1**:217, **6**:587–588
      characteristics **1**:216
      Holothyrida **1**:216
      Ixodida **1**:216
      Mesostigmata **1**:216, **6**:587–588
      Opilioacariformes **1**:216
      Oribatida **1**:217, **6**:587–588
      Parasitiformes **1**:216
      Prostigmata **1**:216–217, **6**:587–588
    phylogenetic relationships **1**:205*F*
    phylogeny and taxonomy **1**:215–216
    reproduction and dispersal **1**:215
  Antarctica **1**:184*T*
  Araneae
    characteristics **1**:205–206, **1**:205*F*
    ecological characteristics **1**:205–206
    genera and species **1**:202*T*
    major lineages
      araneomorphs **1**:206–207
      Entelegynae **1**:207
      Haplogynae **1**:207
      Liphistiidae **1**:206
      mygalomorphs **1**:206
    phylogenetic relationships **1**:205*F*
    phylogeny and taxonomy **1**:206
    reproduction and growth **1**:206
  characteristics **1**:202–204
  diversity **1**:202–204, **1**:202*T*
  endangered species **3**:225–226
  feeding behaviors **1**:204
  habitats **1**:204
  insecticide resistance **4**:298–299
  known and estimated species **4**:384*T*
  Near East ecosystems **5**:460*T*
  number of species **4**:371*T*, **5**:669–670*T*
  Opiliones
    characteristics **1**:209–210, **1**:209*F*
    ecological characteristics **1**:210
    genera and species **1**:202*T*
    major lineages
      Cyphophthalmi **1**:210
      Dyspnoi **1**:210
      Eupnoi **1**:210
      Laniatores **1**:210–211
    phylogeny and taxonomy **1**:205*F*, **1**:210
    reproduction and growth **1**:210
  overharvesting/overcollecting **3**:223
  paleontology **1**:204–205
  phylogenetic relationships **1**:203*F*, **1**:205*F*
  phylogeny and taxonomy **1**:204
  reproduction and growth **1**:203–204
  riparian zones **6**:462–463
  Scorpiones
    characteristics **1**:207–208, **1**:207*F*
    ecological characteristics **1**:208
    genera and species **1**:202*T*
    major lineages **1**:209
    phylogenetic relationships **1**:205*F*
    phylogeny and taxonomy **1**:208–209
    reproduction and growth **1**:208
  smaller orders
    Amblypygi **1**:202*T*, **1**:205*F*, **1**:211, **1**:211*F*
    characteristics **1**:211
    Palpigradi **1**:202*T*, **1**:205*F*, **1**:212
    Pseudoscorpiones **1**:202*T*, **1**:205*F*, **1**:212–213, **1**:213*F*
    Ricinulei **1**:202*T*, **1**:205*F*, **1**:212, **1**:212*F*
    Schizomida **1**:202*T*, **1**:205*F*, **1**:212, **1**:212*F*
    Solifugae **1**:202*T*, **1**:205*F*, **1**:213, **1**:213*F*
    Uropygi **1**:202*T*, **1**:205*F*, **1**:211–212, **1**:211*F*
  subterranean ecosystems **7**:53
Arachnoidea **4**:392, **4**:394*F*
Aradomorpha **7**:294
Aradus cinnamomeus **7**:294
Araliaceae **3**:522*T*
Aral Sea **1**:262–263*T*, **3**:171–173*T*, **4**:377, **4**:461*T*
Araneae
  Amazon ecosystems **1**:246*T*
  characteristics **1**:205–206, **1**:205*F*
  ecological characteristics **1**:205–206
  genera and species **1**:202*T*
  major lineages
    araneomorphs **1**:206–207
    Entelegynae **1**:207
    Haplogynae **1**:207
    Liphistiidae **1**:206
    mygalomorphs **1**:206
  phylogenetic relationships **1**:205*F*
  phylogeny and taxonomy **1**:206
  reproduction and growth **1**:206
Araneidae **1**:184*T*, **1**:207
araneomorphs **1**:206–207, **1**:206
Araripesuchus rattoides **6**:401
Aratus pisonii **5**:15*F*, **5**:16
Araucaria **7**:118–119
Araucaria araucana **3**:117, **4**:105–107
Araucaria cunninghamii **6**:458–459
arbuscular mycorrhizal fungi (AMF) **3**:627–628
Arbutus **5**:218, **7**:118
Arbutus andrachne **5**:481–482
Arcella **6**:290–291
Arceuthobium **4**:71
Arceuthobium douglasii **4**:77
Arceuthobium tsugense **4**:77
Archaea **1:218–226**
  biodiversity **5**:254–255
  common ancestry theory **3**:397*F*
  domain Archaea **5**:286–288, **5**:287*F*, **5**:288*F*
  evolutionary development **2**:640–641, **5**:703–705
  historical research **1**:218
  known and estimated species **4**:384*T*
  marine ecosystems **5**:47–49, **5**:49*F*
  microbial biodiversity **3**:682
  nitrogen fixation **2**:611*F*, **2**:613–614
  number of species **5**:254*T*
  origin and evolution
    early respiratory processes **1**:224–225
    evolutionary development **1**:225–226
    high-temperature adaptations **1**:225
    phylogenetic relationships **1**:223–225, **1**:224*F*, **5**:285*F*, **5**:287*F*, **5**:288*F*
  phylogenetic relationships **1**:218–219, **1**:220*F*, **1**:310*F*, **5**:253*F*, **6**:584*F*, **7**:162*F*
  physiology
    characteristics **1**:219–220
    halophilic Archaea **1**:222–223, **1**:223*T*
    methanogenic Archaea **1**:220–222, **1**:222*T*
    nonextremeophilic Archaea **1**:223, **1**:224*T*
    thermophilic Archaea **1**:219–220, **1**:221*F*, **1**:221*T*
  taxonomy **1**:218–219
Archaefructus **5**:618–619
Archaeoglobus **1**:219–220, **1**:221*T*
Archaeognatha **1**:247, **4**:308–312, **4**:309*T*
Archaeopteris **5**:555, **5**:618*F*
Archeolepis mane **5**:389
archerfish **3**:467–471
Archey's frog **4**:194*T*
Archezoa **3**:329–330
Archosauria **7**:340
Archostemata **1**:326
Archotermopsis **4**:409*T*

Arctic ecosystems
  Census of Marine Life project **2**:20–21, **2**:21*F*
  climate change effects **5**:89–91
  Europe **3**:335*T*
  fish diversity **3**:474
  psychrophiles **6**:299
  terrestrial ecosystems **1:227–244**
    biodiversity filters **1**:238
    biodiversity patterns **1**:231–235, **1**:231*T*
    circumpolar diversity **1**:235–236
    current biodiversity status **1**:242
    environmental characteristics **1**:227–231, **1**:229*F*
    environmental consequences **1**:238–240, **1**:239*F*, **1**:240*F*
    general discussion **1**:227
    latitudinal biodiversity patterns **1**:232–235, **1**:233–234*T*, **1**:235*F*, **1**:236*F*
    natural resources/environmental exploitation **1**:240–241
    research organizations **1**:242–243
    research summary **1**:243
    threats and challenges **1**:241–242
    topographic impacts **1**:236–238, **1**:237*F*
    vegetation zones **1**:229*F*, **1**:230*T*
arctic fox **1**:236, **6**:240, **7**:27, **7**:34*F*
arctic ground squirrels **1**:631–632
Arctic Ocean Diversity (ArcOD) project **2**:20–21, **2**:21*F*
Arctic snowshoe hare **4**:697
arctic terns **1**:605, **5**:311–312
Arctiinae **5**:387–388*T*, **5**:388*F*, **5**:396, 5:397*F*, 5:400*F*
Arctocephalus galapagoensis **5**:87*T*
Arctocephalus gazella **5**:87*T*
Arctocephalus townsendi **5**:81*T*
Arctocephalus tropicalis **5**:87*T*
Arctodus **4**:724–725*T*, **4**:727*T*
Arctogeophilus glacialis **5**:428
Arctophila fulva **1**:232, **1**:236
Arctostaphylos **5**:215
Ardeidae **6**:308*T*
Ardeola ralloides **5**:476–477
Ardeotis kori **1**:52
Ardipithecus kadabba **2**:550–554
Ardipithecus ramidus **2**:550–554
area per se hypothesis **6**:640–641
Areca catechu **7**:26
Arecaceae **3**:129*T*
Arenaviridae **5**:130*F*
Arenicolidae **7**:417*F*
Arenosols **2**:461–464
Argentina
  botanical gardens **6**:519
  C_4 plants **2**:371*T*
  desert ecosystems **2**:469
  habitat–bird species diversity relationship **1**:613–614
  inland deserts **2**:461*F*
  invasive species **4**:345–346
  land cover classifications **6**:607*T*
  maize production **1**:86
  natural history museums and institutions **5**:406
  plant species **4**:102*T*
  seasonal rainfall **2**:465*F*
  Sistema Nacional de Datos Biológicos (SNDB) **1**:506–507
  species richness **1**:499*T*
  temperate forests **7**:116*T*, **7**:117
  tropical forests **6**:600–604*T*
  wheat production **1**:86
Argentine ants **1**:81, **3**:208–209, **3**:222, **6**:50, **6**:93–95
Argentiniforms **3**:464–465
argon (Ar)
  atmospheric concentrations **1**:282*T*
  elemental partitioning **1**:554*T*
  relative abundance **1**:553*F*
Argonne National Laboratory (ANL) **4**:539–540
Argulus **6**:26
Argyrolobium crotalarioides **5**:485
Argyroxiphium **1**:30–32
Argyroxiphium sandwicense ssp. *macrocephalum* **3**:208–209
Argyroxphium **1**:30*F*
Aridisols **6**:123*T*
Ariège (France) **7**:60
Aristida **2**:376*F*
Aristida **1**:268
Aristida adscensionis **1**:67, **2**:373*T*, **4**:446*F*
Aristida pennata **2**:468
Aristida stipoides **1**:67
Aristolochhiodeae **5**:618*F*
Aristoloqia **4**:281–285
Aristotle
  adaptive studies **1**:15
  common good theory **2**:326
  environmental studies **3**:320–321
  pharmacologically active compounds **5**:705
  taxonomic methods **5**:404–405, **7**:89–90, **7**:100
ark paradigm **6**:78, **7**:477–478
armadillos **2**:56*F*, **2**:584*F*, **2**:590*T*, **4**:689–690, **4**:690*T*
Armenia **6**:519
Armillaria **3**:633, **3**:633*T*, **6**:447
Armillaria mellea **3**:633*T*
Armillaria nabsnona **3**:629*F*
Armillaria solidipes **3**:633*T*
Armitermes **4**:409*T*
Army Corps of Engineers **7**:152–153
arnadillo **7**:229*T*
Arnold Arboretum (Harvard) **5**:405–406
aromatic acids **6**:113*T*
aromatic-L-amino-acid decarboxylase **5**:269*T*
Arrabidaea **5**:489
Arrhenius equation **2**:674–675
Arrhenius, Svante **4**:18, **5**:630
Arrhinoceratops **2**:519*T*
arrow poisons **2**:431–432
arrowroot **2**:663, **3**:32*F*, **4**:287–290
arrow worms **6**:25
arroyos **2**:461
arsenic **5**:692, **5**:695–696, **5**:699*T*
Arses kaupi **2**:75*F*
Arsinotheres **4**:701*T*
Artemesia annua **1**:593–594
Artemisia **1**:269
Artemisia artemisiifolia **1**:165
Artemisia frigida **1**:269
Artemisia herba-alba **1**:268
Artemisia monosperma **1**:268
Artemisia pauciflora **1**:270
Artemisia sibirica **1**:269
Artemisia sieberi **5**:483
Artemisia terrae-albae **1**:270
Artemisia tridentata **2**:468–469, **6**:443
artemisinin-based combination therapy (ACT) **1**:593–594
Artenkreis theory **7**:46
Arthrobacter **6**:448
Arthrobotrys **2**:644
Arthropoda/arthropods
  anthropogenic activities and impacts **7**:307*T*
  Census of Marine Life project **2**:7*F*
  characteristics **4**:392
  classifications and sampling methods **4**:381*T*
  domesticated ruminants **1**:709–710
  elevational range shifts **4**:605*T*
  forest canopies **3**:512–513, **3**:513*F*
  fossil record **3**:684*F*
  global diversity **4**:385*T*
  global species richness **3**:711
  insecticide resistance **4**:298–299, **4**:298*F*, **4**:299*F*
  introduced species
    Australia **1**:76*T*, **1**:81
    Brazil **1**:76*T*, **1**:83
    India **1**:76*T*, **1**:82
    South Africa **1**:76*T*, **1**:81
    United Kingdom **1**:76*T*, **1**:79–80
  known and estimated species **3**:710*T*, **4**:384*T*
  latitudinal range shifts **4**:604*T*
  marine ecosystems **5**:50*T*
  number of species **4**:371*T*, **5**:669–670*T*
  phylogenetic relationships **7**:434*F*
  Proterozoic–Cambrian transition **3**:396*F*
  radiation ecology **2**:399–400
  terrestrial arthropods
    Amazon ecosystems **1:245–253**
      biotopes **1**:249–250
      future research areas **1**:252
      genetic characteristics **1**:251–252
      habitat specificity **1**:249–250, **1**:250*T*, **1**:251*F*, **1**:252*F*
      morphological characteristics **1**:250–251
      number of species **1**:245, **1**:246*T*
      origins and geographical distribution **1**:245–247
      species richness **1**:247–249, **1**:248*F*
      taxonomy **1**:245, **1**:246*T*
Arthrotaxus **7**:118–119
Arthurdendyus vegrandis **3**:222
Artibeus watsonii **7**:284
artichokes **3**:32*F*
articulata **3**:685*F*
articulated halkieriids **5**:374*F*
artificial ecosystems **1**:279

Artificial Intelligence for Ecosystem Services (ARIES) software **5**:357–358, **5**:460–462*T*
artificial reefs **5**:61
Artiodactyla
  characteristics **4**:695–697
  convergent evolution **4**:701*T*
  desert ecosystems **2**:476*T*
  dietary tolerance **5**:225*T*
  late Quaternary extinction event (LQEE) **4**:727*T*
  number of species **4**:690*T*
  Paleocene–Eocene thermal maximum **5**:7–8
  representative species **2**:587*F*
  species diversity **2**:587, **4**:690*T*
  tropical rain forests **6**:309*T*
Artioposthia triangulata **3**:225, **6**:240
Artium tomentosum **1**:590
Artocarpus **1**:263
Artocarpus integrifolia **2**:663
Aruba **4**:560*F*, **6**:507*T*
arum lily **5**:220
arum lily frog **5**:220
Arunachal macaque **6**:256–258*T*
Arundinaria alpina **7**:275
Arundinella **2**:376*F*
Arundinella **1**:70
Arundinera **1**:70
Arundo donax **5**:486
Arundo mediterranea **5**:486
Arvicanthis ectos **5**:462
Arvicola terrestris **2**:552*F*
arylaminopropionic acids **4**:92*T*
arylesterase **5**:269*T*
aryloxyphenoxy propionate **4**:90*T*
arylsulfatase **5**:269*T*
Asan Bay **1**:276*T*
Ascaphus montanus **4**:517
Ascaphus truei **1**:172–173, **4**:517
Ascaris lumbricoides **5**:660–661, **5**:660*T*, **5**:671*T*, **6**:132, **7**:425–428, **7**:428*T*
Ascension frigatebird **7**:185
Ascension Island **6**:78
Ascetospora **5**:669–670*T*
Aschelminthes
  characteristics **4**:392
  lake and pond ecosystems **4**:461*T*
Aschiza **3**:489–491*T*, **3**:491–492
Ascidiacea **4**:381*T*
ascidians
  mangrove ecosystems **5**:20*T*
  marine environments **4**:382*F*
Asclepiadaceae **3**:129*T*, **6**:316*T*
Asclepias **1**:24*F*
Ascomycota **2**:644, **3**:628–629, **3**:628*F*
Ascothoracida **2**:416–417*T*
Asellus aquaticus **7**:56, **7**:56*F*
Asenophonus nasoniae **5**:679
asexual reproduction **5**:659–660
Asfarviridae **5**:136
ash borers **4**:361
Ashbya gossypii **3**:638–639
Ashgillean event **3**:415–418, **3**:416*T*, **5**:161*F*
ash trees **4**:361, **6**:124*F*, **7**:125
ashy red colobus **6**:256–258*T*
Asia
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
  C_4 plants **2**:371*T*
  crop distribution **1**:88*T*
  deforestation **1**:677*F*, **2**:448*T*, **2**:452*T*, **2**:454*T*
  Diptera **3**:488*T*
  domesticated ruminants **1**:702–703
  ecosystems **1:254–280**
    alpine life zone **1**:149*T*
    animal species **1**:258, **1**:261*T*
    artificial ecosystems **1**:279
    biogeographic provinces **1**:262–263*T*
    biomes **1**:258–260, **1**:262–263*T*
    boreal and polar Asia
      boreal forest **1**:262–263*T*, **1**:273–274
      characteristics **1**:273–274
      polar and mountain tundra **1**:262–263*T*, **1**:274–275
    climatic zonation **1**:257*T*, **1**:258, **1**:262–263*T*
    conservation status **1**:279, **3**:171–173*T*
    deserts **2**:469–470
    dry regions
      characteristics **1**:267–268
      Dzungaria **1**:270
      Gobi Desert **1**:270
      Mediterranean ecosystems **1**:262–263*T*, **1**:268–269
      Middle Asian deserts **1**:270
      subtropical deserts **1**:262–263*T*, **1**:267–268
      Tarim Basin **1**:270
      temperate grasslands **1**:262–263*T*, **1**:269
      temperate semi-deserts and deserts **1**:262–263*T*, **1**:269–270
      Tibetan Plateau **1**:270
    evolutionary history **1**:256*T*
    forest ecosystems **3**:532*T*
    geographic characteristics **1**:255
    geologic history **1**:256, **1**:256*T*
    humid monsoon region
      characteristics **1**:262–263*T*, **1**:270–271
      laurel forests **1**:270–271
      mountain regions **1**:272–273
      pine forests **1**:272
      temperate deciduous forests **1**:271
    physiographic characteristics **1**:255–258, **1**:255*F*
    plant species **1**:258, **1**:259*T*, **1**:260*T*
    regionalization **1**:258–260
    regional species richness controls **2**:602
    river water composition **3**:562*T*
    tropical Asia
      characteristics **1**:260–263
      humid tropical forest **1**:260–263, **1**:262–263*T*, **1**:266*T*
      monsoon forests (moist deciduous forests) **1**:262–263*T*, **1**:263–266
      tropical dry forest, scrub, and savanna **1**:262–263*T*, **1**:266–267
      tropical moist forest **1**:277*T*
      tropical rain forests **6**:314*T*
      tropical/subtropical humid mountain areas **1**:262–263*T*, **1**:266*T*, **1**:267
      vegetation belts **1**:266*T*
    tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*
    vegetation types and coverage percentage **1**:264–265*T*
    wetlands and coastal ecosystems **1**:275–279, **1**:275*T*, **1**:276*T*
endangered mammals **3**:194*T*
energy use **3**:252*T*
fish diversity **3**:473
floodplain agriculture **7**:146
freshwater ecosystems **4**:370*T*
grasslands and savannas **6**:333
habitat destruction **3**:403
indigenous peoples **4**:269*T*, **4**:271*T*
intercropping systems **2**:383, **2**:384*F*
invasive species **4**:353–354
land-use trends **4**:559*F*
Lepidoptera **5**:389*T*
logging operations **2**:454*T*
maize production **1**:86
megaherbivores **5**:224*T*, **5**:225*T*, **5**:232*F*
native crops
  East Asia **2**:663
  Southeast Asia **2**:663
  southwest Asia **2**:662–663
natural history museums and institutions **5**:406
plant species **6**:82*T*
primates
  primate diversity **6**:250–252, **6**:251*T*
  threatened species **6**:258*T*
protected area systems **6**:491*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
resource management **6**:386–387, **6**:387*T*
rice production **1**:86, **1**:88, **1**:388–389
savannas
  biodiversity **1**:70
  biome distributions **1**:64*F*
  characteristics **6**:333
  classifications
    dipterocarp savanna forests **1**:69*F*, **1**:70
    dry deciduous forests **1**:69–70, **1**:69*F*
    dry thorn forests **1**:69, **1**:69*F*
    general characteristics **1**:68–69
    moist sal savannas **1**:70
  climate change effects **1**:71–72
  future outlook **1**:71–72
  geographic distribution **1**:59–60, **1**:59*F*
  global carbon cycle **1**:70–71
  tree cover variations **1**:60*F*
temperate Asia **2**:448*T*
termites **4**:409*T*
threatened birds **7**:181
transboundary protected areas **6**:489*T*
tropical Asia **2**:448*T*
tropical rain forests
  distribution **6**:324*T*
  loss and change **6**:324*T*
  model-based assessments **6**:324*T*
  species richness loss **6**:328*T*
wheat production **1**:87–88

Asian balsam woolly adelgid **4**:361
Asian chestnut blight fungus **4**:361–362
Asian clam **5**:54, **5**:61
Asian elephants **3**:188, **4**:694–695, **5**:225*T*, 5:226, **5**:603, **5**:606
Asian emerald ash borer **4**:361
Asian hemlock woolly adelgid **4**:361
Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346, **4**:346*F*
Asian wild ass **1**:644*F*, **2**:469–470, **3**:191
Asia-Pacific Network for Global Change Research (APN) **4**:327
Asiatic black bear **3**:192
Asiatic knotweed **4**:351–352, **4**:352*F*
Asiatic wild ass **6**:573, **6**:576*F*, **6**:578–579
Asilidae **3**:489–491*T*, **3**:492, **5**:460*T*, **5**:464*T*
Asiloidea **3**:489–491*T*
Asobara **5**:678–679
Asota **5**:393*F*
Aspalathus **5**:219
asparagus **2**:662–663
Asparagus aphyllus **5**:481–482
Asparagus asparagoides **4**:362
Asparagus officinalis **2**:662–663
aspartic acid **6**:113*T*
aspen
  energy use impacts **3**:261
  fire regime alterations **6**:342–343
  habitat–bird species diversity relationship **1**:612
  Holocene postglacial spread **2**:551–552*T*
  plant tissue chemistry **6**:120*T*
  trophic cascades **7**:259
aspergillosis **3**:634*T*
Aspergillus **2**:644, **3**:634*T*
Aspergillus conidia **3**:629*F*
Aspergillus flavus **3**:634–635
Aspergillus nidulans **6**:355
Aspergillus niger **3**:638–639
Aspergillus oryzae **3**:638–639
Aspergillus parasiticus **3**:634–635
Aspergillus sydowii **2**:339–340
Asphodelaceae **1**:486*T*
Aspidisca **6**:295
Aspidodera **7**:422*F*
Aspidogaster **7**:450–451*F*
Aspidogastrea **5**:662–663, **7**:439–440*T*, **7**:441*F*, **7**:450–451*F*, **7**:451
Aspidontus taeniatus **2**:141–142
aspirin **5**:712*T*
Asplanchna **6**:26
Asplanchna priodonta **1**:10*T*
Asplenietea trichomanis **3**:340
Asplenium **6**:316*T*
Asplundia **3**:522*T*
assassin bugs **7**:292–293, **7**:293–294
assimilation efficiency **7**:264–265
assisted colonization **2**:109, **7**:238
Association of Zoos and Aquariums (AZA) **1**:662–663
Assulina **6**:290*F*
Astacida **2**:416–417*T*
Astasia **6**:293
Astatotilapia flaviijosephi **5**:470, **5**:471*F*
Asteiidae **3**:489–491*T*
Asteioinea **3**:489–491*T*
Asteraceae **1**:486*T*, **2**:374*T*, **2**:427, **3**:129*T*, 4:348–349, **6**:83
Asterales **2**:374*T*
Asteridae **4**:72*F*
ASTERID I and ASTERID II Clades diversification **5**:622*F*, **5**:625
ASTERID III and ASTERID IV Clades diversification **5**:622*F*, **5**:625
Asterionella formosa **3**:356*F*, **6**:30, **6**:33*F*
Asteroidea **5**:669–670*T*
asteroid impacts **3**:418–421
Asterozoa **3**:681*F*
Astigmata **1**:217, **6**:587–588
Astomaea seselifolium **5**:485
Astragalus **1**:270, **5**:215
Astragalus bethlehemiticus **5**:483–484
Astragalus echinus **5**:483–484
Astralagus geminiflorus **4**:103
Astrapotheria **2**:583*F*
Astrocaryum aculeatum **4**:288*T*
Astrocaryum jauari **1**:163
Astrocaryum vulgare **4**:286–287, **4**:288*T*
Astronium gracile **1**:162
Astronium lecointei **1**:162
Astroviridae **5**:134*T*, **5**:138
Astyanax mexicanus **3**:464, **7**:53–54
asulam **4**:90*T*, **5**:696
Aswan High Dam **3**:260, **7**:146
Atacama Desert **2**:469, **5**:215–216
Ateleopodomorpha **3**:465
Ateles belzebuth **6**:256–258*T*
Ateles chamek **6**:256–258*T*
Ateles fusciceps fusciceps **6**:256–258*T*
Ateles fusciceps rufiventris **6**:256–258*T*
Ateles geoffroyi azuerensis **6**:256–258*T*
Ateles geoffroyi geoffroyi **6**:256–258*T*
Ateles geoffroyi ornatus **6**:256–258*T*
Ateles geoffroyi vellerosus **6**:256–258*T*
Ateles geoffroyi yucatanensis **6**:256–258*T*
Ateles hybridus brunneus **6**:256–258*T*
Ateles hybridus hybridus **6**:256–258*T*
Ateles marginatus **6**:256–258*T*
Ateles spp. **3**:127*T*
Atelidae **4**:692, **6**:250–252, **6**:251*T*
Atelopus sp. **2**:86–87, **4**:664
Athene noctua **1**:79
Athericidae **3**:489–491*T*
Atherinomorpha **3**:467
Atherton scrub wren **2**:75*F*
athlete's foot **3**:634*T*
Atitlán grebe **1**:621
Atlantic bluefin tuna **2**:7–8
Atlantic Coastal Plain **5**:562–563
Atlantic cod **3**:123*T*, **4**:516, **5**:35, **6**:233
Atlantic croaker **5**:38
Atlantic gazelle **5**:364–365
Atlantic herring **5**:315, **5**:464–465, **5**:464*F*, **6**:422*T*
Atlantic humpback dolphin **5**:81*T*
Atlantic Multidecadal Oscillation (AMO) **4**:213
Atlantic Ocean **4**:386*T*
Atlantic salmon **1**:10*T*, **3**:260, **3**:448–449, **3**:579, **4**:516, **6**:522, **6**:524
Atlantic spotted dolphin **3**:192–193
Atlantic walrus **5**:81*T*
Atlantic white cedar **7**:400
Atlantis model **5**:460–462*T*
atmosphere
  environmental impacts **3**:279
  hydrological cycle **1**:553–554, **1**:555*F*
  water vapor **3**:561*T*
atmospheric epiphytes **3**:524
atmospheric gases **1:281–290**
  changing composition
    carbon dioxide (CO_2)
      biodiversity impacts **6**:126, **6**:127*F*
      causal factors **1**:282–284
      fossil fuel emissions **1**:283–284
      land use changes **1**:284, **4:539–554**
      monthly average measurements **1**:283*F*
    carbon monoxide (CO) **1**:286
    chlorofluorocarbons (CFCs) **1**:284–286, **1**:286*F*
    future outlook **1**:290
    historical perspective **1**:282–284
    methane (CH_4)
      annual average measurements **1**:285*F*
      causal factors **1**:284
      methane budget **1**:284
    nitrogen oxides (NO_x) **1**:282*T*, **1**:286–287
    nitrous oxide (N_2O) **1**:284
    nonmethane hydrocarbons **1**:282*T*, **1**:286
    sulfur gases **1**:287
  concentrations **1**:281–282
  dynamic global vegetation models (DGVMs) **2**:680
  future outlook **1**:290
  importance **1**:281–282
  lower atmosphere **1**:289–290
  savannas **1**:70–71
  stratospheric ozone
    concentration deviations **1**:289*F*
    current trends **1**:287–289, **1**:289*F*
    depletion causes and consequences **1**:289
    future outlook **1**:290
    occurrences and characteristics **1**:287–289
    projected trends **1**:289
    total ozone column **1**:287–289, **1**:288*F*
Atopogale cubana **4**:713–714
Atractosteus spatula **3**:462*F*
atrazine
  ecotoxicology **5**:700*T*
  environmental threats **3**:166
  herbicides **5**:696
Atrecus longiceps **7**:220*T*
Atrichornis clamosus **7**:183–184
Atriplex **2**:371*T*, **2**:376*F*
Atriplex confertifolia **2**:468–469
Atriplex glauca **5**:484–485
Atriplex halimus **5**:484–485
Atriplex rosea **2**:364*F*
Atropa belladonna **2**:427, **5**:712*T*, **6**:136
atropine **2**:427, **2**:432, **5**:712*T*, **6**:136, **6**:136*F*
Attalea maripa **4**:286–287
Attalea speciosa **1**:161

Attelabidae **1**:327*T*
attributional life cycle analysis (LCA) **4**:627–628
Atya lanipes **3**:247
aubergine **3**:32*F*
Aublysodon **2**:519*T*
Aucubaceae **1**:260*T*
Audubon Core Multimedia Resource Metadata Schema **1**:445
Audubon, John James **4**:651–652
Audubon Society **3**:139
Auerodendron pauciflorum **3**:206
auks **1**:606, **1**:617
Aulacaspis yasumatsui **6**:84
Aulacidea subterminalis **4**:349
Aulacigastridae **3**:489–491*T*
Aulacoeira subarctica **6**:33*F*
Aulacomnium turgidum **1**:232
Aulacomya ater **5**:57–58, **5**:60*F*
Aurelia aurita **7**:24
Aureophycus aleuticus **2**:6*F*
Auricularia auricula **3**:637*T*
Australia
  biodiversity richness **1**:501*F*
  bird diversity **1**:604
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:507*T*, **6**:510*T*
  butterflies **1**:658, **1**:659*T*
  C_4 plants **2**:371*T*
  countryside biogeography studies **2**:353–354, **2**:353*F*
  Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
  deforestation **1**:502*T*
  desert ecosystems **2**:468
  Diptera **3**:488*T*
  domesticated ruminants
    historical perspective **1**:702–703
    plant tolerance–grazing relationship **1**:707–708
  ecosystems
    alpine life zone **1**:149*T*
    biogeographic regions **1**:297–298, **1**:298*F*, **1**:299*F*
    characteristics **1**:291–292
    conservation status **3**:171–173*T*
    coral reefs **1**:500*T*
    deserts **2**:468, **2**:470*T*, **2**:471*F*
    floristic groups
      biogeographic regions **1**:298–300
      diagnostic species **1**:292–297
      Mediterranean-type climate ecosystems **5**:221
      TWINSPAN floristic groups **1**:292, **1**:294*T*, **1**:300*F*
    plant communities **1:291–306**
      biogeographic regions **1**:297–298, **1**:298*F*, **1**:299*F*
      characteristics **1**:291–292
      community diversity **1**:300, **1**:300*F*
      evaporative aerodynamics **1**:292, **1**:293*F*
      floristic groups **1**:292, **1**:294*T*
      Mediterranean-type climate ecosystems **5**:221
      research summary **1**:304
      species richness **1**:300, **1**:301*T*
      structural characteristics **1**:301*T*
    river water composition **3**:562*T*
    Southern Hemisphere **6:612–619**
      angiosperms **6**:617–618
      background information **6**:612
      biotic dispersal **6**:616
      climate change effects **6**:617
      climatic patterns **6**:614
      continental drift **6**:616
      dispersal ability **6**:613–614
      fossil record **6**:615–616
      general area cladograms **6**:614
      geographic distribution patterns **6**:613–614
      high southern latitude plant growth **6**:618
      historical research **6**:612–613
      long-distance dispersal **6**:616–617
      relaxed clock molecular dating analyses **6**:614–615
      research summary **6**:618–619
      vicariance biogeography **6**:616
    species richness
      dryland vegetation continua **1**:302, **1**:303*F*
      invasive species **1**:303
      invertebrates **1**:300–302, **1**:303*T*
      isolated rainforest stands **1**:303–304
      overstory and understory strata **1**:300, **1**:301*F*, **1**:302*F*
      postfire succession **1**:302–303, **1**:304*F*
      vertebrates **1**:300–302, **1**:302*F*
      wetland vegetation continua **1**:302
    tropical rain forests **6**:314*T*, **6**:317*T*
  endangered mammals **3**:193–194, **3**:194*T*
  endangered plant species **3**:206–208, **3**:207*F*
  endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
  energy use **3**:252*T*
  Environmental Resources Information Network (ERIN) **1**:506
  fire/fire regimes **3**:435–436
  fish diversity **3**:473, **3**:474
  forests **1**:292–297, **1**:294*T*
  freshwater ecosystems **4**:370*T*
  grasslands and savannas **1**:292–297, **1**:294*T*, **6**:333
  Great Artesian Basin **3**:186, **3**:187*F*
  Great Barrier Reef Marine Park **5**:446
  habitat–bird species diversity relationship **1**:613–614
  habitat destruction **3**:216–217
  heathlands **1**:292–297, **1**:294*T*
  indigenous peoples **4**:269*T*, **4**:270*T*, **4**:271*T*, **4**:272*F*
  institutional collections and capacity **1**:505*T*
  introduced species
    nonindigenous invertebrates
      arthropods **1**:76*T*, **1**:81
      crop plant pathogens **1**:76*T*, **1**:81
      livestock pathogens **1**:76*T*, **1**:81
    nonindigenous plants **1**:76*T*, **1**:80
    nonindigenous vertebrates
      amphibians and reptiles **1**:80
      birds **1**:80
      mammals **1**:76*T*, **1**:80
  invasive species
    aquatic ferns (*Salvinia molesta*) **4**:346–347, **4**:347*F*
    Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346
    Lantana camara **4**:352–353
    Mimosa pigra **4**:353–354, **4**:354*F*
    prickly pear (*Opuntia*) **4**:350–351
  Jenolan Karst Conservation Reserve **5**:444–445
  late Quaternary extinction event (LQEE) **4**:658*F*, **4**:727–728
  Lepidoptera **5**:389*T*
  mammals **4**:701*T*, **4**:705*T*
  marsupial range–abundance relationships **5**:436–437, **5**:437*F*
  marsupials **4**:683
  mass extinctions **5**:6–7
  Mediterranean-type climate ecosystems
    characteristics **5**:212–213
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:220–221
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:213*T*, **5**:221
    vertebrate diversity **5**:213*T*, **5**:221
  megaherbivores **5**:225*T*, **5**:232*F*
  myriapods **5**:429, **5**:429*T*
  natural history museums and institutions **5**:406, **5**:409*F*
  plant invasions **6**:93
  plant species **6**:82*T*
  population diversity **6**:172–173
  postglacial extinctions **5**:363–364, **5**:364*F*
  protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*, **6**:497
  reforestation programs
    mining impacts **6**:378
    North Queensland **6**:377
  scrub **1**:292–297, **1**:294*T*
  shrublands **1**:292–297, **1**:294*T*
  species reductions or absences **6**:445
  species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
  subterranean ecosystems **7**:59
  temperate forests
    distribution patterns **7**:116*T*, **7**:117
    diversity patterns **7**:119–121, **7**:119*T*, **7**:120*F*
    dominant floristic features **7**:118–119
    endemism **7**:121
    physiognomic characterization **7**:117
  termites **4**:409*T*, **4**:432*T*
  threatened birds **7**:181
  wheat production **1**:86
  woodlands **1**:292–297, **1**:294*T*
Australian desert tree frog **1**:173*F*
Australian horny devils **2**:471*F*
Australian Museum (Sydney) **5**:406, **5**:408*F*, **5**:409*F*
Australian National Herbarium **6**:506–508
Australian Network for Plant Conservation **6**:86

Australian pine **1**:279, **7**:399
Australian sea lions **5**:81*T*
Australian torrent frog **1**:172
Australian vedalia beetle **4**:361
Australian whipsnakes **6**:408–409
Australiasia **4**:271*T*
Australimyzidae **3**:489–491*T*
Australopithecus afarensis **2**:550–554
Austramphilina elongata **5**:664
Austria
  botanical gardens **6**:519
  vegetation surveys **3**:341–344*T*
Austrian wetlands **7**:59
Austrobaileya **5**:618*F*, **5**:621*F*
Austrocedrus chilensis **4**:105–107
Austrovenus stutchburyi **6**:160*F*
autapomorphy **2**:34–35, **2**:34*F*, **7**:104–106
Autecological Hypothesis **2**:133–134
autocatalysis **5**:630
autoimmune diseases **7**:302–303
Automeris postalbida **5**:391*F*
autonomous people **6**:221–222, **6**:221*T*
autonomous underwater vehicles (AUVs) **2**:698*F*, **2**:699
autopolyploid speciation **4**:168, **6**:626
autotrophs
  marine ecosystems **5**:48*F*
  trophic interactions **3**:6–7, **7**:264, **7**:265*F*
  trophic pyramid **4**:670–671
auxins **4**:92*T*
Avaceratops **2**:519*T*
Avahi cleesei **6**:256–258*T*
Avahi occidentalis **6**:256–258*T*
avalanches **2**:562–563
Avatar (2009) **4**:654
Avena sativa **1**:462–463, **2**:662–663, **3**:131–132
avermectins **1**:710–711
Averrhoa carambula **2**:663
aversive responses **1**:39–40
Aves *see* birds
Avicennia **1**:277, **5**:11*T*
Avicenniaceae **5**:11*T*
Avicennia germinans **7**:153–154
Avicennia marina **5**:487
Avimimus **2**:519*T*, **2**:520*T*
Avisaurus **2**:518, **2**:519*T*
avocados **2**:663–664
avocets **1**:710, **5**:476–477
Axel Heiberg Island **5**:562
Axelrodia **5**:617–619
axis deer **1**:82
Axymyiidae **3**:489–491*T*
ayahausca **1**:595
Ayala, Francisco **2**:185–186
aye-ayes **6**:251*T*
Aythya ferina **3**:365–366
ayurveda **3**:322
Azadindium **1**:472–473
Azadirachta indica **1**:589–590
Azadirachtin **1**:589–590
Azadiracta indica **5**:693
Azerbaijan **2**:469–470, **6**:519
azinphos-methyl **5**:693–694, **5**:699*T*
Azorella compacta **4**:104
Azorella filamentosa **4**:106–107
Azorella fuegiana **4**:106–107
Azorella madreporica **4**:101
Azorella monantha **4**:105–106
Azorella pedunculata **4**:103
Azorella selago **4**:106–107
Azorella trifoliata **4**:107
Azotobacter **2**:614–615
Azteca instabilis **1**:378*F*
Aztlanolagus **4**:722
Azuero howler **6**:256–258*T*
Azygioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452

B

Baas-Becking hypothesis **5**:271
babaçu palms **4**:287–290
babassu palm forests **1**:162, **1**:529*T*
baboons **4**:692, **6**:251*T*
bacaba **4**:286–287
Bachman's sparrow **7**:183–184
Bachman's warbler **5**:317–318, **7**:182
Bacillariophyceae **1**:476*F*, **3**:571*T*
Bacillariophyta **1**:475
Bacillarnavirus **5**:134*T*, **5**:138
Bacillus brevis **5**:710–711
Bacillus cereus **3**:122*T*
Bacillus licheniformis **5**:710–711
Bacillus thuringiensis **1**:78–79, **1**:393, **3**:222, **4**:366, **5**:395–396, **5**:402–403, **5**:694
back mangroves **1**:277
backpacking
  conservation support **7**:251–252
  habitat degradation **4**:566–567
  per capita contributions **7**:254*F*
  popularity **7**:246–247, **7**:246*F*, **7**:251
backswimmers **4**:371–373, **7**:293
backward smearing effect **3**:414–415
bacteria
  adaptive radiation studies **1**:33
  alpine ecosystems **1**:155
  ammonia-oxidizing bacteria (AOB) **5**:257–258
  bacterial–invertebrate symbioses **7**:328–329
  bacterial symbiosis **5:637–644**
    bacterial–eukaryotic symbiosis **5**:642
    evolutionary development **5**:638, **5**:639*F*
    evolutionary maintenance
      byproduct cooperation **5**:640–641
      partner choice **5**:641
      partner fidelity **5**:641
    mutualism
      evolutionary maintenance **5**:640–641
      origins **5**:638–640
    research summary **5**:642–643
    symbiont capture **5**:641–642
  biodiversity **1:307–316**
    diversity quantification **5**:254–255
    phenotypic trait distribution **1**:309–310, **1**:310*F*
    phylogenetic investigations **1**:308–309, **1**:310*F*, **1**:311*F*
    Proteobacteria symbionts
      α Proteobacteria **1**:312–313
      β Proteobacteria **1**:313
      γ Proteobacteria **1**:313
      general discussion **1**:312–313
    research background **1**:307–308
    research summary **1**:315–316
    16S ribosomal DNA (rDNA) sequences **1**:308–309, **1**:310*F*, **1**:311*F*, **1**:312*F*, **5**:252–254, **5**:292, **5**:296
    strain-level studies **1**:309
    symbiotic prokaryotes
      evolutionary development **1**:311–312, **1**:312*F*
      general discussion **1**:310–311
      phylogenetic affiliations **1**:312–313
      Proteobacteria **1**:312–313
      symbiont identification **1**:310–311
    uncultured free-living organisms
      hot spring environments **1**:315
      marine environments **1**:314
      16S ribosomal DNA (rDNA) sequences **1**:313–314, **1**:314*F*
      soil environments **1**:315
      strain-level studies **1**:313–314
      uncultured Archaea **1**:314
      uncultured bacteria **1**:314–315
bioluminescence **5**:121–123
biomonitoring processes **3**:122*T*
characteristics and properties **2**:524*T*
common ancestry theory **3**:397*F*
domain bacteria **5**:285–286, **5**:285*F*, **5**:287*F*
domesticated ruminants **1**:710
endosymbiotic bacteria **7**:328–329
evolutionary development **2**:640–641, **5**:637–638, **5**:703–705
filamentous bacteria **2**:12*F*
freshwater sediments **3**:572*T*
functional role
  element cycling
    aquatic sediments **5**:302*F*, **5**:303
    extreme environments **5**:304
    mineralization processes **5**:302–303
    terrestrial soils **5**:303–304
    water column **5**:302–303, **5**:303*F*
  metabolic diversity **5**:300–302, **5**:301*T*, **5**:302*F*
  principle properties **5**:300
  symbiotic relationships **5**:304–305
genetic diversity **1:317–325**
  circular DNA molecules **1**:325
  genetic analysis
    circular DNA molecules **1**:325
    complementation analysis **1**:323*F*, **1**:324
    general discussion **1**:321
    homologous recombination **1**:322–324, **1**:324*F*
    portable regions of homology **1**:324
    reversion/suppressor analysis **1**:321–322
    transposable elements **1**:324–325
  genetic exchange
    conjugation **1**:320, **1**:320*T*, **1**:321*F*
    importance **1**:319
    transduction **1**:320–321, **1**:320*T*, **1**:322*F*

bacteria (*continued*)
transformation **1**:319–320, **1**:320*T*
mutations
circular DNA molecules **1**:325
conditional mutations **1**:317–318, **1**:318*T*
genetic analysis **1**:321
mutant isolation and characterization **1**:318–319, **1**:318*T*, **1**:319*F*, **1**:320*F*
mutation effects **1**:317–318
occurrences **1**:317
transposable elements **1**:324–325
summary discussion **1**:325
transposable elements **1**:324–325
genetic exchange
conjugation **1**:320, **1**:320*T*, **1**:321*F*
importance **1**:319
transduction **1**:320–321, **1**:320*T*, **1**:322*F*
transformation **1**:319–320, **1**:320*T*
global species richness **3**:710–711
halobacteria **6**:9–11
Holocene postglacial spread **2**:551–552*T*
hydrothermal vents **4**:391
lake and pond ecosystems **4**:460, **4**:461*T*
marine ecosystems **5**:47–49
molecular level diversity **2**:635*T*
mutations
circular DNA molecules **1**:325
conditional mutations **1**:317–318, **1**:318*T*
genetic analysis
circular DNA molecules **1**:325
complementation analysis **1**:323*F*, **1**:324
general discussion **1**:321
homologous recombination **1**:322–324, **1**:324*F*
portable regions of homology **1**:324
reversion/suppressor analysis **1**:321–322
transposable elements **1**:324–325
mutant isolation and characterization **1**:318–319, **1**:318*T*, **1**:319*F*, **1**:320*F*
mutation effects **1**:317–318
occurrences **1**:317
transposable elements **1**:324–325
nematodes **7**:430–431
nitrogen fixation **6**:581–582, **6**:582*F*
nomenclature code **5**:551–552, **5**:553*T*
nonseed gene banks **3**:652, **3**:652*T*
number of species **5**:254*T*, **6**:583–584, **6**:586*T*
oil-eating bacteria **1**:592
parasitoid interactions **5**:678–679
phylogenetic relationships **1**:224*F*, **1**:310*F*, **5**:253*F*, **5**:285*F*, **5**:287*F*, **6**:584*F*, **7**:162*F*
rank-abundance curves **6**:346–347*F*
rare biosphere **6:345–352**
abundance considerations **6**:351–352
basic concepts **6**:345, **6**:346–347*F*
clone libraries **6**:349–350, **6**:349*T*
collector curves **6**:349–350, **6**:350*F*
global distribution **6**:351
high-throughput sequencing **6**:350–351
loss mechanisms **6**:351
microbial species numbers **6**:345–347
pure cultures versus sequences **6**:348–349
research summary **6**:352
species and phylogenetic concepts **6**:347–348, **6**:348*F*
slash-and-burn systems **6**:556
symbiotic relationships **5**:304–305
temperature tolerance ranges **4**:113–115
trophic pyramid **4**:670–671
upper atmosphere bacteria **6**:298–299
bacteriochlorophylls
light-harvesting pigment–protein complexes **6**:9–11, **6**:10*T*
primary photochemical processes **6**:11–12, **6**:11*T*
Bacteriological Code **5**:551–552
bacteriophages **6**:584–585
bacterioplankton
characteristics **6**:24–25
functionality **6**:31–32
trophic pyramid **4**:670–671
Bactrian camels **1**:644*F*, **2**:469–470, **2**:476*T*, **3**:191, **4**:697, **6**:333, **7**:29*F*
Bactris gasipaes **4**:284–285, **4**:284*F*, **4**:287–290, **4**:289*T*
bacuri **4**:286–287, **4**:287–290
badgers
adipose tissue **7**:34*F*
skull size comparisons **1**:692*F*
species diversity **1**:685, **4**:693
Badhamia utricularia **3**:631*F*
bagasse **3**:261
Bagehot, Walter **4**:652
Bahamas **6**:519
Bahamian hutas **3**:197
bahia grass **4**:170
Bahrain
animal species **1**:261*T*
botanical gardens **6**:507*T*
ecosystems
mangroves **1**:275*T*
threatened species **1**:278*T*
energy efficiency measures **3**:254*F*
plant species **1**:259*T*
baiji (Yangtze River dolphin/*Lipotes vexillifer*) **5**:75
Baikal (Lake)
ancient lakes species flocks
characteristics **4**:463*T*
hydrothermal vents **4**:459
invertebrates **4**:464, **4**:464*T*
ecosystem characteristics **1**:262–263*T*, **1**:279
endemic species **3**:179, **3**:183*F*, **4**:374–375
fish species **3**:444
photosynthesis **3**:560
Baikal oilfish **3**:467–471
Baileya multiradiata **2**:373*T*
baiting **1**:330
Baja California Desert **2**:468–469
bajadas **2**:461, **2**:464*F*
Balaena glacialis **5**:76–78, **5**:77*T*
Balaena mysticetus **5**:77*T*, **5**:86–87, **5**:87*T*
Balaeniceps rex **1**:53–54
Balaenoptera acutorostrata **3**:192–193
Balaenoptera acutorostrata scammoni **5**:88–89
Balaenoptera bonaerensis **5**:88–89
Balaenoptera borealis **5**:81*T*
Balaenoptera borealis borealis **5**:88–89
Balaenoptera brydei **5**:88–89
Balaenoptera musculus **1**:601, **3**:190, **5**:81*T*, **5**:82
Balaenoptera physalus **5**:81*T*, **5**:313–314
Balaenoptera physalus quoyi **5**:88–89
Balanites aegyptiaca **5**:483, **7**:387
Balanites aegytpica **1**:67
Balanites angolensis **1**:67
Balanites wilsoniana **1**:49
Balanophoraceae **4**:72*F*, **4**:73, **6**:315*T*
Balanops **5**:623*F*
Balantidium coli **6**:295
Balanus balanoides **2**:186, **2**:187*F*, **4**:334, **6**:232
Balanus cariosus **6**:723*T*
Balanus glandula **6**:723*T*
Balanus improvisus **3**:364
Bald Eagle Protection Act (1940) **3**:713*T*, **3**:716
bald eagles **3**:256, **3**:406, **4**:567
Baldwin principle **6**:7–8
balearic shearwater **7**:182
baleen whales
characteristics **4**:693–694, **5**:66–67
migratory patterns **5**:313–314
taxonomy **5**:65*T*
Bali starling **7**:187
Ball-Berry model **2**:675
Ballotetea undulatae **5**:483
Balsam fir **3**:440
balsam poplar **6**:120*T*, **6**:123–124
Baltic Sea **3**:171–173*T*
Baltic Sea ringed seal **5**:81*T*
bambara groundnuts **2**:659, **2**:663
Bambiraptor **2**:519*T*
bamboo **1**:162, **7**:118, **7**:275
bamboo lemur **6**:251*T*, **6**:256–258*T*
Bambusaceae **3**:129*T*
bananas **2**:384*F*, **2**:663, **3**:32*F*
Banau Sentarum **1**:276*T*
banded anteater **4**:687
banded pigs **5**:603
bandicoots **4**:687*T*, **4**:688
Banff National Park **7**:260*F*
Bangladesh
animal species **1**:261*T*
botanical gardens **6**:507*T*
deforestation **1**:677*F*
ecosystems
mangroves **1**:275*T*, **1**:277
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical moist forest **1**:277*T*
vegetation types and coverage percentage **1**:264–265*T*
education level–fertility relationship **6**:204, **6**:205*F*
plant species **1**:259*T*
rice production **1**:88
Banisteriopsis caapii **1**:595
Banksia **5**:221

Banksia marginata **1**:304*F*
Banksia ornata **1**:302–303, **1**:304*F*
Banks, Joseph **1**:577–578
bank vole **1**:365–368
banteng **1**:640
Banyuasin Musi River Delta **1**:276*T*
baobab trees **2**:468, **2**:471, **5**:235
Baptisia bracteata **1**:707
Baptornis **2**:519*T*, **2**:520*T*
Barbados **4**:560*F*, **6**:519
Barbarea vulgaris **1**:75
Barbary fig **4**:350–351
Barbary macaque **4**:692–693, **6**:251–252, **6**:256–258*T*
Barbeya **5**:623*F*
Barbourula kalimantanensis **1**:515
Barbus canis **5**:463*T*
barefaced ibis **7**:390*F*
bark beetles **1**:144, **5**:675–676, **6**:41, **6**:150*T*
bark epiphytes **3**:524
barks **3**:130*T*, **3**:131*T*
barley
  biofuel production **1**:529*T*
  Europe **2**:662–663
  genetic diversity **3**:670, **3**:674*F*
  intercropping patterns **2**:383
  malting traits **1**:461*F*, **1**:462
  Mediterranean Basin **2**:662–663
  Neolithic agriculture **2**:658
  quality guidelines **1**:460–461, **1**:461*F*
  regional cultivars **3**:32*F*
  root–soil interface **6**:113*T*
  southwest Asia **2**:662–663
barley and cereal yellow dwarf virus **2**:529–530, **2**:529*F*
Barlow's lark **1**:487–488
barnacles
  classifications **4**:381*T*
  interspecific competition **2**:186, **2**:187*F*
  intertidal ecosystems **4**:334
  predation effects **4**:334–335, **6**:232
Barnadesia **4**:101
Barneoudia **4**:101
Barneoudia major **4**:105–106
barn owls **5**:606–607, **6**:689–690
barn swallows **5**:312
Baronia brevicornis **1**:650
Baroniinae **1**:650
barracudas **3**:260, **3**:467–471
Barringtonia asiatica **1**:589
Barro Colorado Island (BCI) **4**:267, **6**:239–240, **6**:307, **7**:273
bar-tailed godwits **7**:27–28
Bartel, David **5**:633–636
Bartramia longicauda **5**:312
Bartram, William **4**:651–652
Bary, Heinrich Anton de **5**:116–117
Barytheriidae **5**:224*T*
Barytherium **5**:224*T*
basal cell carcinoma (BCC) **7**:302
basal species **7**:264
basalt **7**:327–328
Basidentitermes **4**:409*T*
Basidiobolus ranarum **3**:696
Basidiomycota **2**:644, **3**:628*F*, **3**:629–630, **3**:629*F*
Basin and Range Orogeny **5**:563
basket stars **4**:381*T*
Basque Center for Climate Change **5**:460–462*T*
bass **1**:10*T*, **1**:12*F*, **3**:123*T*, **3**:467–471, **6**:153*F*, **7**:261, **7**:390
Bassia (Chenolea) arabica **5**:484–485
Bates, Henry Walter **2**:424
Batesian mimicry **1**:656, **2**:424
Bateson, Gregory **6**:383
Bateson, William **6**:2
batfish **3**:466–467
bat flies **3**:489–491*T*
Bathybates **1**:25*F*
Bathycestus brayi **7**:456–457*F*
Bathyergus janetta **1**:487–488
Bathymodiolus thermophilus **4**:391
Bathynellacea **2**:416–417*T*, **3**:182*T*
Bathyplectes curculionis **4**:366
Bathysiphon **6**:292
Batrachochytrium dendrobatidis
  amphibians
    environmental threats **2**:86–87, **3**:166, **4**:664
    epidemiology **1**:519–520
    introduced parasites–species decline relationships **2**:527–529, **2**:530*F*
    population declines **1**:177, **3**:696
  characteristics **3**:627
  as disease vector **4**:505–506
  environmental threats **3**:166
Batrachospermum **3**:358
bats
  biodiversity-friendly farming **1**:420–421, **1**:423*F*
  boreal forests **1**:631–632
  characteristics **4**:691–692
  Chile **5**:217
  forest canopies **3**:513–514
  fruit and seed dispersal **6**:50, **6**:51*T*
  Great American Biotic Interchange (GABI) **2**:584*F*
  introduced parasites–species decline relationships **2**:527–529, **2**:529*F*
  mangrove ecosystems **5**:14
  migratory patterns **5**:313–314
  Near East ecosystems **5**:474–475
  new species **4**:704
  number of species **4**:690*T*
  pollination **6**:49*T*, **6**:141*T*, **6**:142–143
  pollution effects **3**:192–193
  seed dispersal **7**:283–284
  species diversity **2**:587
  subterranean ecosystems **7**:51–52, **7**:60
  Succulent Karoo **1**:487–488
  white-nose syndrome **4**:156
Bauera **5**:623*F*
Bauhinia **5**:489
bauxites **4**:228
bay-breasted warbler **2**:188–189, **2**:188*F*
Bayesian clustering analysis **4**:511, **4**:512*F*
Bayesian species delimitation **6**:684
Bdallophyton **4**:72
beaches
  coastal beach ecosystems **2**:**128–136**
    sandy beaches
      alongshore distribution **2**:131–133
      background information **2**:128
      beach index (BI) **2**:129
      future research areas **2**:135
      macrofauna surveys **2**:129–131, **2**:130*F*
      macroscale patterns **2**:131
      mesoscale patterns **2**:131–133
      microscale patterns **2**:133
      microtidal regimes **2**:128–129, **2**:129*F*
      morphodynamic states **2**:128–129, **2**:129*F*
      research summary **2**:135
      species richness hypotheses **2**:133–134
      threats and challenges **2**:134–135
      tidal regimes **2**:129
      two-dimensional model **2**:129, **2**:130*F*
      zonation patterns **2**:131–133
  intertidal beaches **5**:51
  sandy beaches
    alongshore distribution **2**:131–133
    background information **2**:128
    beach index (BI) **2**:129
    future research areas **2**:135
    intertidal ecosystems **4**:337–339
    macrofauna surveys **2**:129–131, **2**:130*F*
    macroscale patterns
      latitudinal trends **2**:131, **2**:131*F*
      morphodynamic effects **2**:131, **2**:132*F*, **2**:133*F*
    mesoscale patterns **2**:131–133
    microscale patterns **2**:133
    microtidal regimes **2**:128–129, **2**:129*F*
    morphodynamic states **2**:128–129, **2**:129*F*
    nature-based coastal protection strategies **5**:452–453, **5**:453*T*
    research summary **2**:135
    species richness hypotheses **2**:133–134
    threats and challenges **2**:134–135
    tidal regimes **2**:129
    two-dimensional model **2**:129, **2**:130*F*
    zonation patterns **2**:131–133
beach flies **3**:489–491*T*
Beagle expedition **2**:419–420, **4**:651–652
beaked whales **3**:192–193, **4**:693–694, **5**:65*T*, **5**:314
beans **1**:132*T*, **1**:462*F*, **2**:383, **2**:663–664, **6**:553
bearded pig **1**:517
bearded sakis **6**:251*T*
bearded seal **5**:90
bearded vultures **5**:474
beardfish **3**:466
Bear River **7**:152–153
bears
  adipose tissue **7**:34*F*
  biomedical research **1**:358–359
  boreal forests **1**:631–632
  convergent evolution **4**:701*T*
  Great American Biotic Interchange (GABI) **2**:584*F*

bears (*continued*)
late Quaternary extinction event (LQEE) **4**:724–725*T*
Near East ecosystems **5**:472
postglacial range expansion **2**:550, **2**:552*F*
prey–predator interactions **1**:709
species diversity **1**:686, **4**:693
urbanization threats **4**:564
beating **1**:330
Beauvaria bassiana **5**:694
beavers
boreal forests **1**:631–632
ecosystem/landscape conservation **7**:405–406
Great American Biotic Interchange (GABI) **2**:584*F*
late Quaternary extinction event (LQEE) **4**:724–725*T*
riparian zones **6**:466
river ecosystems **6**:479–480
wetland ecosystems **7**:390
bedbugs **7**:287–288, **7**:294
beech trees
greenhouse effect impacts **4**:28*F*
habitat–bird species diversity relationship **1**:611–612
North America **7**:118
bee-eaters **5**:474
bee flies **3**:489–491*T*, **3**:492, **6**:49*T*
bee hummingbirds **1**:601
bee lice **3**:489–491*T*
beer making **3**:638
bees
characteristics **4**:181, **4**:392
honey production **4**:183
introduced parasites–species decline relationships **2**:532
introduced species **6**:144–145
killer bees **4**:183
kin selection theory **2**:322
native vegetation **1**:419, **1**:419–420, **1**:420*F*
number of species **4**:392*T*
oil-palm plantations **5**:605
pollination **4**:183, **6**:49*T*, **6**:141*T*, **6**:142
social systems **2**:646
see also Hymenoptera
beeswax **4**:183
beetle flies **3**:489–491*T*
beetles **1**:326–331
acidification effects **3**:584–585
adaptive radiation **1**:430
Amazon ecosystems **1**:250, **1**:250*T*, **1**:251*F*, **1**:252*F*
angiosperms **1**:23, **1**:328–329
Arctic ecosystems **1**:231*T*
beetle–host plant interactions **1**:23, **1**:24*F*
characteristics **4**:392
diet **1**:327*T*
endangered species **3**:226
feeding behaviors **1**:328–329
forest canopies **3**:512–513, **3**:513*F*, **3**:514*F*
fossil record **1**:326
freshwater ecosystems **3**:579*T*
fruit and seed dispersal **6**:50
genetic diversity **3**:674*F*
geographic diversity **1**:329, **1**:329*F*
global diversity estimates **1**:330
habitat loss and fragmentation **3**:220–221
insecticide resistance **4**:298–299
interspecific competition **6**:670
major families
Buprestidae **1**:327*T*, **1**:328
Carabidae **1**:326, **1**:326–328, **1**:327*T*
Cerambycidae **1**:327*T*, **1**:328
Chrysomelidae **1**:327*T*, **1**:328
Curculionidae **1**:327*T*, **1**:328
Scarabaeidae **1**:327*T*, **1**:328
Staphylinidae **1**:327*T*, **1**:328, **5**:674
Tenebrionidae **1**:327*T*, **1**:328, **3**:513
mangrove ecosystems **5**:13
migratory patterns **5**:316
number of species **3**:182*T*, **4**:392*T*
occurrences **1**:326
pollination **6**:49*T*, **6**:140–141, **6**:141*T*
quantity estimates **1**:326–328
rare species **1**:330
research challenges **1**:329–330
research summary **1**:330–331
riparian zones **6**:462–463
sampling methods **1**:329–330
species and genera **1**:326, **1**:327*T*
species coexistence **6**:670
species richness **1**:328–329
subterranean ecosystems **7**:52–53, **7**:52*F*, **7**:60
systematic classifications **7**:91–93, **7**:92*F*
taxonomic patterns **1**:329
timber-managed boreal forests **7**:218–219, **7**:219*T*, **7**:220*F*, **7**:220*T*
tropical rain forests **6**:306
beets
genetic diversity **1**:463
native crops **2**:662–663
regional cultivars **3**:32*F*
Beggiatoa **6**:32
Begnas Tal **1**:276*T*
Begonia **6**:316*T*
Begonia eiromischa **1**:514
Beijing Botanic Garden **6**:510–511
Beijing Natural History Museum (BMNH) **5**:406
Beilschmiedia **1**:270–271
Belarus **6**:519
Belgium **4**:266–267, **4**:560*F*, **5**:529*F*, **6**:519
Belitung tarsier **6**:256–258*T*
Belize
botanical gardens **6**:507*T*
coastal zone management models **5**:465, **5**:465*F*
deforestation **2**:30–31, **2**:31*F*
species richness **2**:25*T*
urbanization **4**:560*F*
Bellanwilla-Attidiya marshes **1**:276*T*
Beloneuria georgiana **3**:185*F*
Belostomatidae **7**:389–391
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391
beluga whales **3**:192–193, **4**:693–694, **5**:65*T*, **5**:66–67, **5**:77*T*, **5**:79–80, **5**:314
Bemaraha woolly lemur **6**:256–258*T*
Bemisia tabaci **4**:298–299, **4**:299*F*
benalaxyl **5**:695
Benchmarks for Science Literacy **3**:133–134
benfluralin **5**:696
Bengalian rain forest **1**:262–263*T*
Bengal tigers **5**:14
Beni, Bolivia **1**:159*T*
Benin
botanical gardens **6**:519
deforestation **2**:449*T*
Beni Savannas **5**:496, **5**:497*F*
Beni titi **6**:256–258*T*
Bennettitales **5**:615–617, **5**:617*F*, **5**:618*F*
benodanil **5**:695
benomyl **5**:695, **5**:700*T*
bensulide **4**:90*T*
bentazon **4**:90*T*
Benthamiella azorella **4**:106–107
Benthamiella nordenskjoldii **4**:106–107
benthic biomass distribution **4**:380–381, **4**:380*F*
benthic ecosystems
ecosystem diversity **5**:53
energy flow **3**:243*F*
eutrophication impacts
long-term human influences **3**:354–356
species shifts **3**:352–354, **3**:353*F*
macrofauna **5**:106*F*
marine sediments **5**:105–107, **5**:106*F*
meiofauna **5**:106*F*
surface water acidification **1**:10*T*
water-column nitrogen:phosphorus (N:P) ratios **3**:350*F*
benthic foraminifera **6**:290*F*, **6**:291–292
see also Foraminifera
benzene **5**:693, **5**:695
benzimidazoles **5**:695, **5**:700*T*
benzoic acid **4**:90*T*
benzoin **5**:712*T*
benzonitriles **5**:697
Berberidopsis **5**:622*F*
Berberis empetrifolia **4**:106
Berberis thunbergii **6**:125, **7**:125
Berger–Parker dominance **4**:616*F*
Bergmann, C. **1**:579
Bergman's rule **2**:86
Bericostenus humicola **5**:426
Berlese sampling **1**:330
Berlin Dahlem Botanic Garden **6**:510–511
Bermuda **6**:519
Bermuda Atlantic Time-series Study (BATS) **2**:616–617
Bernardino Rivadavia Natural Sciences Museum **5**:406

Berne Convention **7**:187
berseem **2**:391
Bertholletia excelsa **1**:160, **4**:287–290, **4**:288*T*, **7**:273
Beryciforms **3**:467–468
Bessey's cactus classification system **2**:51, **2**:51*F*
β-N-methylamino-L-alanine (BMAA) **1**:592, **5**:704–705
beta diversity
  basic concepts **6**:659–660
  biogeographic scales **6**:660*F*, **6**:661, **6**:662*F*
  characterizations **6**:659–660, **6**:660*F*
  habitat and dispersal limitation scales **6**:660*F*, **6**:661
  partitioning methods **2**:439–440
  sampling scale **6**:660–661, **6**:660*F*
  species diversity measures **1**:403, **2**:650, **4**:339–340, **6**:534, **6**:534
  temporal beta diversity **6**:662, **6**:662*F*
Betaflexiviridae **5**:130*F*
Beta vulgaris **1**:463, **2**:662–663, **7**:24–25
betel nuts **7**:26
Bettongia lesueur **4**:363–364
Betula **3**:76*F*
Betula **1**:628–630, **1**:630*T*, **7**:118
Betula alagheniensis **1**:610
Betula albosinensis **1**:272–273
Betula alleghaniensis **1**:632*T*
Betulaceae **5**:621–625, **5**:623*F*
Betula costata **1**:271–272
Betula dahurica **1**:271–272
Betula ermani **1**:630*T*
Betula kenaica **1**:630*T*
Betula kusmisscheffi **1**:630*T*
Betula nana sensu lato **1**:232
Betula papyrifera **1**:628–629, **1**:630*T*, **1**:632*T*
Betula pendula **7**:212–213
Betula platyphylla **1**:269
Betula pubescens **1**:273–274, **1**:630*T*, **7**:212–213
Betula vericosa **1**:269, **1**:630*T*
between-habitat diversity **6**:534
Bhutan
  animal species **1**:261*T*
  botanical gardens **6**:519
  ecosystems
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  urbanization **4**:560*F*
Bhutan capped langur **6**:256–258*T*
Bibionidae **3**:489–491*T*, **3**:493–494
Bibionomorpha **3**:489–491*T*
Biblical stories **4**:650–651
bichirs **3**:461–462
Bicknell's thrush **7**:186
Bicosoeca **6**:293
Bicyclus anyana **6**:7
Bidens cernua **6**:457
Bidens pilosa **3**:704–705, **3**:704*F*
Biebricher, Christof **5**:632
Bienertia **2**:376*F*
Bienertia cycloptera **2**:365*F*
bigeye tuna **5**:35
big-headed ants **3**:222, **4**:365
big-headed flies **3**:489–491*T*
bighorn sheep **3**:193, **3**:197, **5**:313, **5**:318–319
Bignoniaceae **3**:522*T*, **5**:489, **6**:314, **6**:316*T*
bignose galaxia **4**:194*T*
Big Trout (Lake) **4**:461*T*
Bilateria **3**:395, **3**:396*F*
bilbies **4**:688
bilharzia **5**:660*T*, **5**:671*T*
billabong **7**:369*T*
billfish **3**:447, **3**:467–471
Billia **7**:274–275
Bill Williams River **7**:153
Bimastos syriacus **3**:674*F*
binapacryl **5**:695
binapicryl **5**:694
bioassays **3**:121, **3**:122*F*
Biocon experiment **4**:585–586, **4**:585*F*
BIODEPTH project **4**:586–587
biodiesel fuel **1**:592–593
biodiversity **1:399–410**
  air pollution impacts
    air quality guidelines **1**:145–146
    birds **1**:139
    evidentiary research
      general discussion **1**:139–141
      lichen biodiversity **1**:141–142, **1**:142*F*
      nitrogen deposition effects **1**:142–144, **1**:143*F*, **1**:146*F*
      ozone effects **1**:144
      smelting industry impacts **1**:140–141, **1**:141*F*, **1**:142*F*
      surface water acidification **1**:144–145
    evolutionary responses **1**:138–139
    long-term and short-term effects **1**:137–138
    mammals **1**:139
    species interaction effects **1**:139, **1**:140*F*
    synergistic interactions **1**:138
  amphibians **1:169–178**
    conservation
      as ecosystem components **1**:176
      human uses **1**:176
      major threats **1**:176–177
      population declines **1**:177
    current diversity **1**:169–170
    evolutionary history **1**:169
    functional anatomy **1**:170–171
    future outlook **1**:177–178
    historical biogeography **1**:169–170
    larval morphology and ecology **1**:173
    life histories **1**:171
    morphology **1**:170–171
    postlarval behavior and ecology **1**:175–176
    reproductive biology **1**:171, **1**:173*F*
  Antarctica
    current studies **1**:183–184
    origins and antiquity **1**:184
    plants **1**:184*T*
    terrestrial invertebrates **1**:184*T*
  aquaculture impacts **1**:190, **1**:191*F*, **7**:203*F*
  Archaea **5**:254–255
  bacteria **1:307–316**
    diversity quantification **5**:254–255
    phenotypic trait distribution **1**:309–310, **1**:310*F*
    phylogenetic investigations **1**:308–309, **1**:310*F*, **1**:311*F*
    Proteobacteria symbionts
      α Proteobacteria **1**:312–313
      β Proteobacteria **1**:313
      γ Proteobacteria **1**:313
      general discussion **1**:312–313
    research background **1**:307–308
    research summary **1**:315–316
    16S ribosomal DNA (rDNA) sequences **1**:308–309, **1**:310*F*, **1**:311*F*, **1**:312*F*, **5**:252–254, **5**:292, **5**:296
    strain-level studies **1**:309
    symbiotic prokaryotes
      evolutionary development **1**:311–312, **1**:312*F*
      general discussion **1**:310–311
      phylogenetic affiliations **1**:312–313
      Proteobacteria **1**:312–313
      symbiont identification **1**:310–311
    uncultured free-living organisms
      hot spring environments **1**:315
      marine environments **1**:314
      16S ribosomal DNA (rDNA) sequences **1**:313–314, **1**:314*F*
      soil environments **1**:315
      strain-level studies **1**:313–314
      uncultured Archaea **1**:314
      uncultured bacteria **1**:314–315
  basic concepts **1**:341–342, **1**:399–401, **1**:414, **2**:307–308, **4**:122–123
  biodiversity dynamics modeling **5:321–328**
    background information **5**:321–322
    phenomenological models
      dose–response models **5**:324, **5**:324*F*
      global species distributions **5**:324
      species–area relationships **5**:322–323
      statistical estimation models **5**:324
    process-based models
      integro-difference equations **5**:324–325
      multispecies metapopulation analysis **5**:325–326, **5**:325*F*, **5**:326*F*
      neutral theory **5**:326–327
      reaction–diffusion models **5**:324–325, **5**:325*F*
    research summary **5**:327
    species–area relationships
      countryside habitats **5**:323–324, **5**:323*F*
      native habitats **5**:322–323, **5**:322*F*
  biodiversity-friendly farming **1:418–429**
    background information **1**:418–419
    biodiversity stewardship **2**:300
    disadvantages **2**:544
    influencing factors **2**:544, **2**:554*F*
    landscape-scale effects
      general discussion **2**:544
      landscape connectivity **2**:544–545
      landscape heterogeneity **2**:544
    local management strategies
      chemical use minimization **2**:542

biodiversity (*continued*)
crop production **2**:542–544
disturbance events **2**:542–544
general discussion **1**:419–420
livestock grazing **2**:542–544
native vegetation benefits **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
pest control **2**:544
pollination **2**:544
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
structural complex agricultural areas **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
management practices **2**:299–300, **2**:300
research summary **2**:545–546
biodiversity loss **4:656–666**
causal factors
anthropogenic impacts **4**:661–662
biotic environment changes **4**:662–664, **4**:663
climate change **4**:664
deforestation **4**:663
disease **1**:519–520, **4**:664
ecological perspectives **4**:664–666
exploitation **4**:662
general discussion **4**:661–662
habitat conversion **4**:664–666, **4**:665*T*
habitat destruction **4**:662
habitat fragmentation **4**:662, **4**:663
persecution **4**:662
climate change effects
communities and ecosystems **2**:76
geographic and landscape scale variation **2**:74–75
individuals and populations **2**:75–76, **2**:75*F*
species interaction risks **2**:76–77
temporal match–mismatch (TMM) hypothesis **2**:76
deforestation **2**:454
extinction mechanisms
general discussion **4**:659
isolated populations **4**:659, **4**:661*F*
metapopulation theory **4**:660–661
extinction rates
average lifespan **4**:656*T*
current extinction rates **4**:658, **4**:659*T*
fossil record **4**:656
future extinction estimations **4**:658–659, **4**:659*F*, **4**:660
genetic narrowing **1**:404–405
mass extinction events **4**:657*F*
Pleistocene vertebrates **4**:658*F*
recent extinction rates **4**:658, **4**:659*T*
freshwater invertebrates **4**:376–377
general discussion **4**:656
human well-being **1**:435–436
hunting/overhunting **1**:405
isolated populations
Allee effect **4**:659
demographic stochasticity **4**:659–660, **4**:661*F*
environmental stochasticity **4**:660, **4**:661*F*
genetic bottlenecks **4**:660, **4**:661*F*
low population densities **4**:659
population size changes **4**:660, **4**:661*F*
Latin America **7**:352–353
marine sediments **5**:113–114
marine species **4**:664
plant communities **6**:63, **6**:85–87
biodiversity monitoring and assessment programs (BMAP) framework **3:545–559**
basic concepts
appropriate indicator selection **3**:548–549
functional role **3**:545–546
goals and objectives **3**:546, **3**:547*F*, **3**:548*T*
guiding principles **3**:546–548
legal and institutional framework **3**:549
process flow diagram **3**:547*F*
spatial and temporal scales **3**:546
conceptual ecosystem and landscape models **3**:549, **3**:550*T*
general discussion **3**:545
monitoring protocols design
assessment and monitoring components **3**:549–550, **3**:550*T*
data analysis **3**:553–554
data collection **3**:551
decision-making strategies **3**:555
evaluation assessments **3**:557
lessons learned **3**:557
sampling designs **3**:551
selection criteria **3**:550–551
stakeholder involvement **3**:556
biodiversity-rich countries **1:497–508**
background information **1**:497
case studies
Biota-Fapesp **1**:506
collection accessibility **1**:504–506
Comisión Nacional Para el Conocimiento y Uso de la Biodiversidad (CONABIO) **1**:505–506
Environmental Resources Information Network (ERIN) **1**:506
Instituto de Investigación de Recursos Biológicos Alexander von Humboldt **1**:507
Instituto Nacional de Biodiversidad (INBio) **1**:506
National Biodiversity Authority (NBA) **1**:507
Sistema Nacional de Datos Biológicos (SNDB) **1**:506–507
South African National Biodiversity Institute (SANBI) **1**:506
defining criteria
cultural diversity **1**:498
ecological heterogeneity **1**:497
endemic concentrations **1**:497
habitat diversity **1**:497
marine ecosystems **1**:498
species richness **1**:497
tropical forest ecosystems **1**:497–498
future outlook **1**:507
major threats
deforestation **1**:501–502, **1**:502*T*
general characteristics **1**:501–502
habitat loss and fragmentation **1**:501–502
overexploitation **1**:502
socio-democratic characteristics
biocultural diversity **1**:502–503
geographic distribution **1**:503*F*
indigenous peoples **1**:503
institutional collections and capacity **1**:504, **1**:505*T*
land ownership factors **1**:503
rural population distribution **1**:502–503
scientific infrastructure **1**:503–504
top ranked countries **1**:501, **1**:501*F*
underlying causes
coastal/marine contour **1**:500, **1**:500*T*
ecological heterogeneity **1**:500
general characteristics **1**:498
historical factors **1**:499–500
latitudinal location **1**:498, **1**:498*F*, **1**:499*T*
primary production factors **1**:498–499
tropical rain forests **1**:500–501
biodiversity stewardship **2**:297, **2**:300
components **1**:430
conservation **4:313–323**
basic principles **4**:315
biodiversity-friendly farming **1:418–429**
background information **1**:418–419
biodiversity stewardship **2**:300
chemical use minimization **2**:542
crop production **2**:542–544
disadvantages **2**:544
disturbance events **2**:542–544
influencing factors **2**:544, **2**:554*F*
landscape-scale effects **2**:544
livestock grazing **2**:542–544
local management strategies **1**:419–420
management practices **2**:299–300, **2**:300
native vegetation benefits **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
pest control **2**:544
pollination **2**:544
research summary **2**:545–546
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
structural complex agricultural areas **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
biogeographic research **1**:585–586
botanical gardens **6**:517
conservation products
biodiversity utilization **4**:316*F*, **4**:321–322
characteristics **4**:321, **4**:321*T*
product dissemination **4**:316*F*, **4**:321

energy sprawl **1**:528–529, **1**:529*F*
ex situ conservation
agrobiodiversity conversation **3**:33
botanical/zoological garden conservation **4**:316*F*, **4**:319, **4**:321*T*
field gene banks **3**:33, **4**:316*F*, **4**:319, **4**:321*T*
in vitro conservation **4**:316*F*, **4**:319
livestock park conservation **4**:316*F*, **4**:319, **4**:321*T*
maintenance strategies **1**:406
plant biodiversity **6**:63
pollen/semen/ovule/DNA conservation **4**:316*F*, **4**:319–320, **4**:321*T*
research programs and methodologies **4**:318–319
seed/embryo collection and storage **4**:316*F*, **4**:319, **4**:319*F*, **4**:321*T*
genetic diversity loss **4**:313–314
Global Environment Facility (GEF) **7:136–143**
habitat–niche relationship **4**:47–48
importance **4**:313
in situ conservation
agrobiodiversity conversation **3**:33
genetic reserve conservation **4**:316*F*, **4**:320, **4**:320*F*, **4**:321*T*
maintenance strategies **1**:405–406
on-farm conservation **3**:33, **4**:316*F*, **4**:320, **4**:321*T*
plant biodiversity **6**:63
research programs and methodologies **4**:316*F*, **4**:320
international organizations **4**:325–326
justice and equity
best management practices **4**:435
ecosystem services **4**:435
natural resources **4**:435
program support **4**:435
protected areas and heritage sites **4**:435–436
sustainability **4**:436
major threats **4**:314–315
marine protected areas **5:94–104**
benefits **5**:96–97
characteristics and functional role **5**:94–95
climate change threats **5**:100–101
ecosystem-based management **5**:95
effectiveness **5**:100
fish stocks **3**:486
geographic extent **5**:445–446, **5**:446*T*
global commitment **5**:99–100
socioeconomic benefits **5**:98–99, **5**:98*T*
threats and challenges **5**:100–101, **5**:447–448
molecular level diversity **2**:637
outside protected areas
biosphere reserves **2**:301
challenges **2**:292–293
conservancy/megaconservancy networks **2**:301, **2**:302*F*
conservation approaches **2**:293–294, **2**:294*T*
land and water protection and management **2**:294, **2**:294*F*, **2**:294*T*
living lands and waters **2**:294*F*, **2**:295–297, **2**:301–303
living landscape framework **2**:294–295, **2**:294*F*
systematic conservation planning **2**:293
working lands and waters **2**:294*F*, **2**:298–299, **2**:301–303
priority management **6:261–272**
carbon sequestration **5**:351–352, **6**:264*F*
carbon storage **5**:351–352, **6**:264*F*
challenges **6**:270–271
cost–benefit analyses **6**:265–266, **6**:266*F*
economic factors **6**:266–269, **6**:267*F*, **6**:268*F*, **6**:269*F*
ecosystem service targets **6**:263–265, **6**:264*F*
land-use patterns **6**:265–266, **6**:266*F*, **6**:268*F*, **6**:269*F*
livestock **6**:264*F*
mapping and modeling approaches **6**:263–265
research summary **6**:269–271
systematic conservation planning **2**:293, **6**:261–263
water provisions **6**:264*F*
protected area systems
characteristics **2**:289–292
classifications **2**:290
geographic distribution **2**:291*F*
global biome coverage **2**:292*F*
global growth estimates **2**:290*F*
research programs and methodologies
community-based conservation **4**:316*F*, **4**:320–321
conservation objectives **4**:316*F*, **4**:318
conservation products **4**:316*F*, **4**:321, **4**:321*T*
conservation strategies **4**:316*F*, **4**:318
ecogeographic surveys/preliminary survey missions **4**:316*F*, **4**:317–318
ex situ techniques **4**:316*F*, **4**:318–319
field exploration **4**:316*F*, **4**:318
goals **4**:315–317
in situ techniques **4**:316*F*, **4**:320
modeling approaches **4**:315–317, **4**:316*F*
project commission **4**:316*F*, **4**:317
target taxa selection **4**:315–317, **4**:316*F*
Southeast Asia **1**:520–521
species–area relationships (SARs)
background information **6**:646–648
incidence functions **6**:649–650
loss management strategies **6**:648–649
nestedness measures **6**:649–650
relaxation concept **6**:647
species loss **6**:643*F*, **6**:646–648, **6**:647*F*
summary discussion **6**:650
species concepts
ecological importance **6**:689
genetic differentiation **6**:689
spatiotemporal considerations **6**:689–690
species as real entities **6**:689
sustainable and integrated biodiversity conservation **4**:322, **4**:436
systematic conservation planning **2**:293, **6**:261–263
zoos/zoological parks
ark paradigm **7**:477–478
breeding population research **7**:474
frozen zoos **7**:477
institutional planning **7**:475
management strategies **7**:474–475
population management strategies **7**:475–477
post-ark paradigm **7**:478
record keeping plans **7**:475
research summary **7**:478
coral reef ecosystems
diversity patterns **2**:333–334
interspecific interactions **2**:334
major groups **2**:331
species diversity **2**:331–333, **2**:332*T*, **2**:333*T*
crustaceans **2**:396–397
cultural ecosystem services **1:332–340**
basic concepts **1**:333–335, **1**:333*F*, **2**:490*T*, **7**:315–316
cultural diversity effects **1**:335–336, **1**:335*F*
cultural services **1**:334*T*
desertification **2**:490*T*
general discussion **1**:333
habitat services **1**:334*T*
marine ecosystems **5**:451–452, **5**:452*T*
provisioning services **1**:334*T*
regulating services **1**:334*T*
sustainable development **1**:337–339
valuation language/value articulating institutions **1**:336–337, **1**:337*T*
definition
background information **1**:399
basic concepts **1**:399–401, **2**:307–308
contextual variations
attributes **1**:406–407
biological resource assets **1**:407–408
classifications **1**:406
derivations **1**:406
management objectives **1**:407–408
precautionary principle **1**:408
variation implications **1**:408–409
definitions **4**:122–123
ecological complexity versus biodiversity **2:201–212**
background information **2**:201
basic concepts **2**:204
biotic interactions and relationships **2**:202–204, **2**:203*F*, **2**:204*F*
community complexity **2**:203*F*, **2**:204–205, **2**:205*T*

biodiversity (*continued*)
entangled bank metaphor **2**:202–204
functional complexity **2**:203*F*, **2**:204–205, **2**:205*T*
historical perspective **2**:201–204
hypothetical relationships **2**:210–211, **2**:211*F*
measures and consequences
community complexity **2**:205–207
complexity measures **2**:205
functional complexity **2**:207–208
intricacy measures **2**:205*T*
phylogenetic complexity **2**:208, **2**:209*F*
usefulness **2**:208–209
phylogenetic complexity **2**:203*F*, **2**:204–205, **2**:205*T*
research summary **2**:210–211
schematic diagram **2**:203*F*, **2**:204*F*
significance
community complexity **2**:209–210
complexity in nature **2**:209–210
ecosystem functioning **2**:210
phylogenetic complexity **2**:210
theoretical perspectives **2**:210–211
ecological perspectives **3**:4–5
ecological succession
disturbed landscapes **7**:69–70
mature communities **7**:68–70
species gains and losses **7**:69*F*, **7**:70
economic importance **1:386–398**
ecosystem productivity **1**:386–388, **1**:392
ecosystem services **1**:389–391, **1**:395–396, **1**:396*F*
general discussion **1**:386
genetic knowledge **1**:389
insurance role **1**:388–389, **1**:392–393
market forces
basic concepts **1**:391–392
bioprospecting **1**:393–395
crop development **1**:395
drug development **1**:393–395
ecosystem services **1**:395–396, **1**:396*F*
insurance role **1**:392–393
productivity **1**:392
public versus private good **1**:391–392
research summary **1**:396–397
Tilman biomass study **1**:386–388, **1**:387*F*
ecosystem boundaries
analytical studies **3**:241–242
basic concepts **3**:238–239, **3**:238*F*
control mechanisms **3**:241, **3**:241*F*
food web–trophic level relationship **3**:240–241, **3**:241*F*
historical perspective **3**:238–239
trophic dynamic concept **3**:239–240
ecosystem diversity **1**:402, **3**:6–7
ecosystem function measurements **3**:84–86, **3**:84*F*
ecosystem services **1:341–356**
basic concepts **1**:344, **7**:315–316
biodiversity conservation **4**:435
biodiversity–ecosystem services relationship **1**:341, **3**:342, **3**:342
biodiversity impacts **7**:316
carrying capacity **7**:75
China **3:372–384**
background information **3**:372
Ecosystem Function Zones (EFZs) **3**:382
Grain-to-Green Program (GTGP) **3**:379–380
land-use planning and human development **3**:101
National Forest Conservation Program (NFCP) **3**:375–376
Nature Reserve System (NRS) **3**:372–373, **3**:373*F*
payment for ecosystem services (PES) **2**:227
summary discussion **3**:382
countryside biogeography **2**:352–353
cultural ecosystem services **1:332–340**
basic concepts **1**:333–335, **1**:333*F*, **2**:490*T*, **7**:315–316
cultural diversity effects **1**:335–336, **1**:335*F*
cultural services **1**:334*T*
desertification **2**:490*T*
esthetic services **5**:353–354
general discussion **1**:333
habitat services **1**:334*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:353
nature-based recreation models **5**:353
provisioning services **1**:334*T*
regulating services **1**:334*T*
spiritual services **5**:354
sustainable development **1**:337–339
valuation language/value articulating institutions **1**:336–337, **1**:337*T*
cultural services
basic concepts **7**:315–316
landscape scale **1**:344–345
regulating services **1**:334*T*
sustainability **7**:76*T*
desertification **2**:490*T*, **2**:491
economic importance **1**:396*F*
ecosystem functioning **1**:342–344, **3**:342
ecosystem processes **1**:342–344, **3**:342
energy flow **3**:247
global status **3**:46*T*
global urbanization impact
dependence interactions **4**:237–238, **4**:238*F*
environmental impact **4**:235
management strategies **4**:239*F*, **4**:241–242
habitat services **1**:334*T*
human well-being **1**:345, **1**:436, **7**:77–80, **7**:78*F*
marine ecosystems **5:451–468**
accounting assessment models **5**:455–457
American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries **5**:464–465, **5**:464*F*
bioeconomic models **5**:457–459
challenges **5**:455
coastal zone management (Belize) **5**:465, **5**:465*F*
conceptual framework **5**:452–453
coral reef protection and restoration **5**:465–466
cost–benefit analyses **5**:459–463
coupled social–ecological systems model **5**:453–455, **5**:454*F*
cross-system services **3**:96–98
cultural ecosystem services **5**:451–452, **5**:452*T*
decision-support tools **5**:459–463, **5**:460–462*T*, **5**:463*F*
future outlook **5**:466
integrated ecosystem assessments (IEAs) **5**:454–455
modeling approaches and uses **5**:453
multiple ecosystem services **5**:459–463
production function models **5**:457–459
provisioning services **5**:451–452, **5**:452*T*
regulating services **5**:451–452, **5**:452*T*
revealed references methods **5**:456–457
single ecosystem services **5**:457–459
stated preference methods **5**:457
supporting services **5**:451–452, **5**:452*T*
tradeoffs **5**:459–463
valuation measures **5**:457
wind farms **5**:466
marine sedimentary habitats
filtration processes **5**:112
general discussion **5**:109
nutrient cycling **5**:111
pollutant cycling **5**:111–112
schematic diagram **5**:111*F*
secondary production **5**:112
shoreline stability **5**:112
species diversity connection **5**:112–113
nature-based recreation **7**:253
payment for agrobiodiversity conservation services (PACS) **3**:40–41, **3**:41*F*
payment for ecosystem services (PES) **2**:226–227, **2**:298, **3**:40–41, **3**:432–433, **5**:151, **7**:353–355
priority management **6:261–272**
carbon sequestration **5**:351–352, **6**:264*F*
carbon storage **5**:351–352, **6**:264*F*
challenges **6**:270–271
cost–benefit analyses **6**:265–266, **6**:266*F*
economic factors **6**:266–269, **6**:267*F*, **6**:268*F*, **6**:269*F*
ecosystem service targets **6**:263–265, **6**:264*F*
land-use patterns **6**:265–266, **6**:266*F*, **6**:268*F*, **6**:269*F*
livestock **6**:264*F*

mapping and modeling approaches **6**:263–265
research summary **6**:269–271
systematic conservation planning **2**:293, **6**:261–263
water provisions **6**:264*F*
provisioning services
basic concepts **2**:490*T*, **7**:315–316
cultural ecosystem services **1**:334*T*
desertification **2**:490*T*
landscape scale **1**:344
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:350–351
sustainability **7**:76*T*
regulating services
basic concepts **2**:490*T*, **7**:315–316
cultural ecosystem services **1**:334*T*
desertification **2**:490*T*
global status **3**:46–47, **3**:46*T*
landscape scale **1**:344
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:351–352
structure–function relationship **3**:46–47, **3**:47*T*
sustainability **7**:76*T*
relationship characteristics and shape **1**:345, **3**:351
relationship examples
agricultural food production **1**:346, **3**:345
general discussion **1**:346
human infectious disease regulation **1**:348–350, **3**:349
pollination **1**:348, **3**:347
soil productive potential regulation **1**:346–348, **3**:347
research summary **1**:355
service providers **1**:345
social value **7:314–326**
compensated versus uncompensated demand **7**:321–322, **7**:322*F*
contingent valuation method **7**:324–325
estimation methods **7**:322
hedonic valuation estimation methods **7**:322–324
Hicksian demand function **7**:321–322
historical perspective **7**:314–315
Marshallian demand function **7**:321–322
revealed-preference estimation methods **7**:322–324
stated-preference estimation methods **7**:324–325
summary discussion **7**:325
travel cost estimation methods **7**:322–324
valuation measures **7**:317–318, **7**:317*F*
value definition **7**:316–317
value quantification **7**:318–321
spatially variable ecosystem services **3**:53
spatial scale analyses
experimental manipulation effects **3**:349
landscape scale **1**:351–353, **3**:353, **3**:354
local scale **1**:350–351, **3**:349
plant functional diversity **3**:354
regional to global scale **1**:353–355
spatiotemporal scales **1**:345–346
structure–function relationship **3**:46–47, **3**:47*F*
supporting services
animal pollination **5**:354–355
basic concepts **2**:490*T*, **7**:315–316
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:354–355
nutrient retention **5**:355–356
sustainability **7**:76*T*
valuation measures **7:167–179**
accounting price **7**:172
biological resource exploitation **7**:167–168
discontinuous value functions **7**:175–177, **7**:177*F*
ecological functions **7**:168–169
gross domestic product (GDP) **7**:173–174, **7**:174*F*
institutional failure **7**:170–172
local versus global stability **7**:170–172
market failure **7**:169–170
Millennium Ecosystem Assessment (MEA) **7**:168–169
necessity versus luxury perspective **7**:173–174
option value **7**:172
project evaluation **7**:172–173
resource evaluation **7**:172–173
social cost–benefit analysis **7**:172–173
substitution possibilities **7**:174–175
summary discussion **7**:177–178
water funds **7:352–366**
basic concepts **7**:352–353
Brazil **7**:363–364
Colombia **7**:362–363
design and implementation guidelines **7**:360
Ecuador **7**:361–362
financial sources and management **7**:357*F*, **7**:359
governance **7**:359–360
historical perspective **7**:355–357
Latin America and the Caribbean region **7**:354–355*F*, **7**:356*T*
operational and financial model **7**:357–358, **7**:357*F*
public protected area management **7**:358, **7**:358*T*
water fund platform **7**:364–365
watershed management **7**:357–358
energy flow
experimental research **3**:247
number of species effects **3**:245–246
research background **3**:245–247
top-down versus bottom-up regulation **3**:246–247
environmental ethics **3:309–318**
animal rights **3**:313–314
basic principles **3**:309
conservation biologists
personal choices **3**:317–318
professional considerations **3**:317
role models **3**:316–317, **3**:317*F*
environmental philosophies
deep ecology **3**:274–276, **3**:275*T*
dominant worldview **3**:275*T*
humanism **3**:276–277
goals
ecological integrity **3**:310–311, **3**:312*F*
economic factors **3**:309–310
long-term healthy ecosystems **3**:309–310, **3**:310*T*, **3**:311*F*
sustainability **3**:311–312
human flourishing
artistic expression **3**:273, **3**:274*F*
general discussion **3**:273–274
historical understanding **3**:274
life-support systems **3**:274
recreational activities **3**:273, **3**:273*F*
religious inspiration **3**:274
scientific knowledge **3**:273–274
human obligations **3**:271–272
human rights
challenges **3**:314–315
less-developed countries **3**:315–316, **3**:316*F*
logging industry–owl population controversy **3**:314–315, **3**:315*F*
property rights **3**:315
intrinsic value
anthropocentric denials **3**:269
economic arguments **3**:267
ethical holism **3**:269–270, **3**:270*F*
extensionist arguments **3**:267–268
nonextensionist arguments **3**:268–269, **3**:269*F*
management versus nonmanagement **3**:312–313, **3**:313*F*
religious traditions
Asian traditions **6**:383–384
Buddhist declaration on nature **3**:272, **6**:383
Christian declaration on nature **3**:272–273, **6**:381, **6**:381
Hindu declaration on nature **3**:273, **6**:383
Jewish declaration on nature **3**:273, **6**:381, **6**:382
Muslim declaration on nature **3**:273, **6**:381, **6**:382
pantheistic traditions **6**:384–385
stewardship **3**:272, **6**:381–383
virtue ethics **3**:275
estuarine ecosystems
characteristics **3**:299–300
ecological functions
biogeochemical cycles **3**:301
biological properties **3**:301
chemical modifications **3**:301

biodiversity (*continued*)
climate variations **3**:302
ecosystem modifications **3**:301
exchange of materials **3**:301
extreme events **3**:302
freshwater delivery **3**:301
general discussion **3**:300–301
physico–chemical properties **3**:301
sea level changes **3**:302
suspended matter **3**:301
temperature changes **3**:302
transport changes **3**:301
wind **3**:302
fish **3**:299–300
evolutionary perspective **1:411–417**
background information **1**:411
community stability
adaptive system complexity **1**:415
disturbance effects **1**:415–416
dynamic processes **1**:415
evolutionary dynamics **1**:416–417
evolutionary models **1**:416
evolutionary versus ecological time **1**:415–416
current patterns
general discussion **1**:414
saturation levels **1**:414–415
temporal changes **1**:414
fossil record
community interpretation and analysis **1**:413–414
mass extinctions **1**:412–413, **1**:413*F*
taxa diversity **1**:411–412, **1**:411*F*, **1**:412, **1**:412*F*
usefulness **1**:413
future research areas
current mass extinction events **1**:416
evolutionary dynamics **1**:416–417
research summary **1**:417
extinctions **1**:404–405
fish stocks
AquaMaps initiative **3**:484–485, **3**:485*F*, **3**:486*F*
bioquads **3**:484–485
challenges **3**:484
global catch data **3**:483–484, **3**:484*F*
marine protected areas **3**:486
species databases **3**:485
stock management **3**:485–486
forest ecosystems
community organization **3**:533–534
disturbance regimes **3**:530–533
ecosystem functions **3**:535
food webs **3**:533–534
forest structure and patterns **3**:530–533, **3**:532*F*
general patterns **3**:530, **3**:531*F*
human activity **3**:534
keystone species **3**:535
latitudinal gradients **1**:498*F*, **3**:530, **3**:530*F*
tree taxa **3**:532*T*
fossil record **3:678–690**
Archean era **3**:682
biodiversity drivers
climate change **3**:689
extinction events **3**:689–690
global heterogeneity **3**:688–689
plate tectonics **3**:688–689
evolutionary development
morphological disparities **3**:679
species richness **3**:678–679
fossil biodiversity
completeness measures **3**:680, **3**:681*F*
rock record completeness **3**:679–680
sampling challenges **3**:679
taxonomic reconstruction measures **3**:680–682
microbial biodiversity **3**:682
Phanerozoic marine biodiversity
Cambrian explosion **3**:683, **3**:684*F*
chordate diversity **3**:684–685
invertebrate diversity **3**:683–684, **3**:685*F*, **3**:686*F*
neoproterozoic animals **3**:682–683
taxonomic turnover **3**:685–686
Phanerozoic terrestrial biodiversity
animal diversity **3**:687–688, **3**:687*F*
chordates **3**:688
invertebrate diversity **3**:687–688, **3**:687*F*
plant diversity **3**:686–687, **3**:687*F*
Proterozoic era **3**:682, **3**:683*F*
research approaches **3**:678
freshwater ecosystems **3**:566–569, **3**:568*F*
functional diversity (FD) measures **3:597–608**
basic concepts **3**:597–598, **3**:598*F*
existing measures
combined functional richness–function divergence **3**:602–603, **3**:603*F*
functional divergence **3**:601–602, **3**:602*F*
functional evenness **3**:601, **3**:601*F*
functional richness **3**:598–601, **3**:599*F*, **3**:599*T*, **3**:600*F*, **3**:601*F*
general discussion **3**:598–601
Rao's quadratic entropy **3**:603, **3**:603*F*, **3**:606–608, **3**:607*F*
recommended measures **3**:603–604
key criteria **3**:598, **3**:599*T*
purpose **3**:597
research applications
ecosystem multifunctionality–functional divergence measures **3**:604–605, **3**:605*F*
general discussion **3**:604–605
grassland ecosystems **3**:604–605, **3**:605*F*
lacustrine fish communities **3**:605–606, **3**:606*F*, **3**:607*F*
phosphorus gradient–hierarchical partitioning measures **3**:606–608, **3**:607*F*
species richness–functional diversity measures **3**:605–606, **3**:606*F*, **3**:607*F*
fungi **3**:625–626
future research outlook **4**:125–126
generation mechanisms **1:430–434**
comparative analyses
independent evolutionary contrast analyses **1**:433–434, **1**:433*F*
trait–environmental character relationship **1**:433
fossil record **1**:430
key innovations **1**:432
reconstructed versus real evolutionary trees
extinction rate–speciation rate relationship **1**:431*F*
lineages-through-time plot **1**:432*F*
maximum likelihood estimation **1**:431*F*
modeling approaches **1**:430–432
phylogenetic reconstruction **1**:431*F*
sister taxon analysis **1**:432
genetic diversity **1**:401
geographic models **7**:91, **7**:91*F*
geologic timescale **3:678–690**
Archean era **3**:682
biodiversity drivers
climate change **3**:689
extinction events **3**:689–690
global heterogeneity **3**:688–689
plate tectonics **3**:688–689
evolutionary development
morphological disparities **3**:679
species richness **3**:678–679
fossil biodiversity
completeness measures **3**:680, **3**:681*F*
rock record completeness **3**:679–680
sampling challenges **3**:679
taxonomic reconstruction measures **3**:680–682
microbial biodiversity **3**:682
Phanerozoic marine biodiversity
Cambrian explosion **3**:683, **3**:684*F*
chordate diversity **3**:684–685
invertebrate diversity **3**:683–684, **3**:685*F*, **3**:686*F*
neoproterozoic animals **3**:682–683
taxonomic turnover **3**:685–686
Phanerozoic terrestrial biodiversity
animal diversity **3**:687–688, **3**:687*F*
chordates **3**:688
invertebrate diversity **3**:687–688, **3**:687*F*
plant diversity **3**:686–687, **3**:687*F*
Proterozoic era **3**:682, **3**:683*F*
research approaches **3**:678
global agriculture **3:426–434**
biodiversity regime shifts and loss thresholds **3**:429–430, **3**:430*T*
future outlook
agricultural policies **3**:432–433
agriculture-related conservation strategies **3**:431
market-based approaches **3**:432–433
Millennium Ecosystem Assessment (MEA) **3**:430–431, **3**:431*T*
payment for ecosystem services (PES) **3**:432–433
sustainable agriculture **3**:430–431
impacts
climate change **3**:427–429, **3**:428*F*

general discussion **3**:427
habitat loss **3**:427, **3**:428*T*, **4**:154
invasive species **3**:429
water quality **3**:429
water quantity **3**:429
prevalence and distribution **3**:426–427, **3**:427*F*
research summary **3**:433
Global Environment Facility (GEF) **7:136–143**
evolutionary development
biodiversity strategy **7**:138–139
operational strategy and programs **7**:138
functional role **7**:136–137
future outlook **7**:141–142
GEF benefits index (GBI) **7**:137–138
GEF performance index (GPI) **7**:137–138
global policy framework and guidance **7**:137
gross domestic product index (GDPI) **7**:137–138
key accomplishments **7**:142
monitoring and evaluation strategies
learning objectives **7**:140–141
portfolio level **7**:140
results-based management framework **7**:140–141
origins **7**:136–137
project investment and funding **7**:137, **7**:138*F*
System forTransparent Allocation of Resources (STAR) **7**:137–138
global urbanization **4:231–244**
demographic transition **4**:231–232
dependence interactions
biodiversity **4**:237
ecosystem services **4**:237–238, **4**:238*F*
general discussion **4**:237
environmental impact
ecosystem services **4**:235
general discussion **4**:234–235
habitat loss and fragmentation **4**:234–235
positive impacts **4**:236–237
protected area systems **4**:235–236, **4**:235*F*
geographic distribution **4**:235*F*
management strategies
ecosystem services **4**:239*F*, **4**:241–242
general discussion **4**:238–241
rural-to-urban migration **4**:240–241*F*, **4**:242
sustainable development **4**:239–241, **4**:239*F*
population trends **4**:231, **4**:231*F*
rural-to-urban migration **4**:232–234, **4**:233*F*, **4**:240–241*F*, **4**:242
governance systems **2:155–166**
advantages/disadvantages **2**:163–164
centralized regimes **2**:156–157
challenges **2**:155–156
complex adaptive systems
building blocks **2**:161–162
internal models **2**:161
parallel adaptive systems **2**:163–164
polycentric systems **2**:164–165
properties and mechanisms **2**:161
rule configurations and situations **2**:162–163, **2**:162*T*
self-organized resource governance systems **2**:160–161
tags **2**:161
local governance systems **2**:163–164
polycentric systems **2**:164–165
rules as tools
aggregation rules **2**:159–160, **2**:162*T*
boundary rules **2**:157–158, **2**:162*T*
choice rules **2**:158–159, **2**:162*T*
information rules **2**:159–160, **2**:162*T*
payoff rules **2**:159, **2**:162*T*
position rules **2**:159, **2**:162*T*
rule change experimentation **2**:160
rule characteristics **2**:157–158, **2**:162*T*
rule improvements **2**:160
scope rules **2**:159–160, **2**:162*T*
summary discussion **2**:165
grassland ecosystems
animal species diversity **4**:3–4
biodiversity change scenarios **4**:5–6, **4**:6*F*
characteristics **4**:3
ecosystem functionality **4**:4–5
future outlook **4**:5–6, **4**:6*F*
plant species diversity **4**:3
Relative Yield Total (RYT) **4**:4–5, **4**:5*F*
soil microbial diversity **4**:4
greenhouse effect impacts
anthropogenic effects **4**:24
biota types
at-risk organisms **4**:29
organisms likely to benefit **4**:29–30
community impacts **4**:30
current climate change evidence **4**:30–31, **4**:31
research methodologies
integrated research **4**:26–27
manipulation experiments **4**:25–26, **4**:26*F*
natural climate variability **4**:25
paleobiological data **4**:25
species distribution models **4**:24–25, **4**:24*T*
research summary **4**:31
response types
adjustment **4**:27
evolution **4**:27
extinction **4**:29
migration **4**:27–29, **4**:28*F*
historical perspective **4:121–126**
concept formation **4**:121–122
conservation biologists **4**:123–124
conservation challenges **4**:122
definitions **4**:122–123
Endangered Species Act (ESA) **4**:121–122
historical research **4**:121
international awareness efforts **4**:125
research outlook **4**:125–126
valuation measures **4**:124–125
human health **1:357–372**
biomedical research
bears as models **1**:358–359
cone snails as models **1**:359
nonhuman models **1**:358–359
components
general discussion **1**:357–358
human health impacts **1**:358
species diversity measures **1**:357–358
infectious disease incidence
disease management **1**:360–361, **1**:361*F*
emerging infectious diseases (EIDs) **1**:368–370, **1**:369*F*
global patterns **1**:368–370, **1**:369*F*
hantavirus pulmonary syndrome **1**:365–368, **1**:366*F*
Lyme disease **1**:361–362, **1**:363*F*
schistosomiasis **1**:368, **1**:368*F*
West Nile virus **1**:362–365, **1**:364*F*, **1**:365*F*
zooprophylaxis **1**:360–361
nutritional factors **1**:359–360
preventative/therapeutic agents **1**:359
summary discussion **1**:370
human impact **4:137–152**
basic principles
ambivalence **4**:139
environmental diversity **4**:137–138
human diversity **4**:137–138
impact types **4**:138
sociocultural factors **4**:138*T*
species and population levels **4**:138–139
biodiversity commodification
cities **4**:142
colonialism **4**:142–143
commercial farming industry **4**:144
human body **4**:142
international commerce **4**:143–144
modern transportation **4**:143–144
biodiversity conversion
agriculture **4**:141, **4**:154
population growth effects **4**:141–142
raised bed farming **4**:141
rice paddies **4**:141
Vavilov centers **4**:141
biodiversity modifications
extinction hypothesis **4**:139–140
fire/fire regimes **4**:140
horticultural practices **4**:140–141, **4**:154
biodiversity toxification
Anthropocene epoch **4**:146
industrialization **4**:145
militarization **4**:145–146
oil wells and pipelines **4**:145
climate change **4**:24
conservation efforts **4**:148–149
future outlook
conservation versus extermination **4**:147–148
diversity principle **4**:148
governmental limitations **4**:147
indigenous potential **4**:147
protected area systems **4**:147

biodiversity (*continued*)
uncertainty studies **4**:146–147
general discussion **4**:137
human well-being **1:435–439**
biodiversity loss **1**:435–436
commodity production **1**:436
direct and indirect valuation methods **1**:435–436
ecosystem services **1**:436
esthetic, spiritual, cultural, and recreational benefits **1**:437
insurance benefits **1**:436
knowledge benefits **1**:437
market factors
internalized benefits **1**:438–439
internalized values **1**:438
market efficiency **1**:437
public goods **1**:437–438
option value **1**:436–437
risk factors **1**:436
uncertainty factors **1**:436–437
hunter–gatherer societies **4**:164
international awareness efforts **4**:125
intertidal ecosystems **4**:339–340
justice and equity **4:434–441**
affected populations
colonial legacies and restitution **4**:439–440
consequences of conservation **4**:437–438
distribution of action, responsibility, and power **4**:438–439
distribution of rights, title, and access **4**:439
distribution of risks **4**:437–438
distribution of values **4**:439
basic concepts **4**:434–435
biodiversity conservation
best management practices **4**:435
ecosystem services **4**:435
natural resources **4**:435
program support **4**:435
protected areas and heritage sites **4**:435–436
sustainability **4**:436
importance **4**:434
nonhuman organisms **4**:436–437
research summary **4**:440
Key Biodiversity Areas (KBAs) **7**:203–204, **7**:204*F*
keystone species **3**:92–93, **3**:92*F*, **3**:246–247
large-scale biodiversity experiments **4:583–589**
background and characteristics **4**:583–584
community assembly
Biocon experiment **4**:585–586, **4**:585*F*
Ecotron large-scale controlled environment facility **4**:585
highly replicated laboratory experiments **4**:587, **4**:587*F*
historical perspective **4**:584–585, **4**:585*F*
large-scale field experiments **4**:585–586, **4**:585*F*, **4**:586*F*
multisite biodiversity experiments **4**:586–587
research applications
general discussion **4**:587–588
grassland restoration **4**:588, **4**:588*F*
tree diversity experiments **4**:588–589
species removal **4**:587, **4**:588*F*
underlying mechanisms **4**:584
literary perspectives **4:649–655**
basic concepts **4**:649
colonial expansion **4**:651
Darwinism **4**:651–652
ecocriticism **4**:653–654
ecology–industrialism relationship **4**:652
environmental and land ethics **4**:652–653
etiological tales of nature **4**:650
language and words **4**:649–650
literature–science relationship **4**:649
mass media **4**:654
natural disasters **4**:654
natural rights and animal fables **4**:653
pastoralism **4**:650–651
major threats
categories **4**:314–315, **4**:314*F*
exotic species invasions **2**:80
habitat change **2**:80
nutrient loading **2**:80
overharvesting/overexploitation **2**:80
pollution **2**:80
marine ecosystems
genetic diversity **5**:50
higher taxa **5**:47–49, **5**:50*T*
species **5**:49–50, **5**:50*T*
marine viruses **5**:127–129
market economy **5:145–155**
bioprospecting **5**:148
challenges **5**:152–154
command-and-control policies **5**:150, **5**:152–154
definitions **5**:145
ecotourism **5**:147–150
green products **5**:149
private ownership market economy **5**:145–147
public-based market approaches **5**:150–152
public goods **5**:146
summary discussion **5**:154
sustainable use strategies **5**:147–150
measurement and analysis **5:178–194**
categories **5**:178
challenges **5**:178–179
community diversity **1**:403
estimation methods
confidence judgments **5**:192–193
evaluation criteria **5**:193
real data versus simulated data **5**:192–193
selection criteria **5**:193–194
extrapolation estimators
basic concepts **5**:191–192
birth chains **5**:192
Michaelis–Menten equation **5**:192
species–area curves **5**:191–192
importance **5**:178
intertidal ecosystems **4**:339–340
modeling approaches **1**:403–404
richness estimators
bootstrap estimator **5**:184–185
Chao's abundance-based coverage estimator **5**:188–190
Chao's abundance-based estimator **5**:187–188, **6**:652–653
Chao's incidence-based coverage estimator **5**:190–191
Chao's incidence-based estimator **5**:188
definitions and terminology **5**:179
Fisher's α **5**:185–187
general discussion **5**:181
heterogeneous sample estimators **5**:191
jackknife estimator **5**:183–184
modeling approaches **5**:181
moment estimators **5**:185–187
observed species **5**:181–183, **5**:182*F*, **5**:183*F*
resampling methods **5**:183–184
sample coverage estimators **5**:188–190
sampling considerations
abundance data **5**:180–181
accumulation curve **5**:180*F*
biases **5**:179–180, **5**:180*F*
quadrats **5**:181
sampling scope **5**:180
surrogate data **5**:181
species diversity **1**:402–403, **2**:650, **6**:534, **6**:534
symbols **5**:179, **5**:179*T*
taxonomic diversity **1**:403, **3**:84, **6**:534, **6**:534
measurement and estimation methods **5:195–211**
biases **5**:197
biotic similarity
abundance-based similarity indices **5**:208–209, **5**:209*F*
basic concepts **5**:197
incidence-based similarity indices **5**:207–208
phylogenetic similarity indices **5**:210
total abundance–shared species-based similarity indices **5**:209–210
characteristics **5**:195–196
functional diversity **5**:196–197
phylogenetic diversity **5**:196–197, **5**:197*F*
sampling data
characteristics **5**:197–198
individual-based (abundance) data **5**:198
sample-based (incidence) data **5**:198
species diversity estimation methods
cladistic diversity **5**:205
Gini–Simpson index **5**:202–203
Hill numbers **5**:203, **5**:204*F*
metrics **5**:202–203
phylogenetic entropy **5**:205

Rao's quadratic entropy **3**:603, **3**:603*F*, **5**:205
Shannon entropy **5**:202–203
taxonomic/phylogenetic diversity **5**:204–207, **5**:205*F*, **5**:206*F*
species richness estimation methods
abundance-based coverage estimator (ACE) **5**:201
Chao estimator **5**:200–201
extrapolation estimators **5**:201–202, **5**:202*F*
incidence-based coverage estimator (ICE) **5**:201
nonparametric asymptotic richness estimators **5**:200–201
rarefaction measurement approach **5**:199–200, **5**:199*F*
species accumulation curve **5**:198–200, **5**:198*F*
taxonomic diversity **5**:196–197
microbial biodiversity **5:252–258**, **5:280–291**
background information **5**:252, **5**:280
basic concepts
Archaea **5**:254–255
bacteria **5**:254–255
eukaryotic diversity **5**:255
general discussion **5**:254–255
number of species **5**:254*T*
prokaryotic diversity **5**:255*F*
viruses **5**:255
biological significance **5**:282–283
cell counting measurement
characteristics **5**:261
microscopy **5**:261
most probable number (MPN) assays **5**:262
plate counts **5**:261–262
cellular constituent measurements **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
characterization studies
culture techniques **5**:253–254
genetic methodologies **5**:254
genomic research **5**:252–254
deep-sea hydrothermal vents
bacterial–invertebrate symbioses **7**:328–329
free-living microorganisms **7**:328
ecosystem functions **5**:257
evolutionary development **5**:280–282
future research areas **5**:291
global change effects
microbial diseases **5**:258
terrestrial and marine ecosystems **5**:257–258
maintenance mechanisms
basic principles **5**:255–256
dispersal rates **5**:256–257
disturbances **5**:256
environmental heterogeneity **5**:256
productivity **5**:256
major groups
domain Archaea **5**:286–288, **5**:287*F*, **5**:288*F*
domain bacteria **5**:285–286, **5**:285*F*, **5**:287*F*
domain Eukarya **5**:288–289, **5**:289*F*
fungi **5**:289–291, **5**:289*F*, **5**:290*F*
protists **5**:288–289, **5**:289*F*
measurement and analysis **5:259–270**
computational challenges **5**:268–269
culture-based measurement versus direct measurement **5**:260
environmental heterogeneity **5**:261
future research outlook **5**:269
importance **5**:259–260
measurement challenges **5**:260
measurement methods **5**:261
taxonomic resolution **5**:260–261
units of measurement **5**:259–260
measurement methods
cell counting **5**:261
cellular constituents **5**:262–264
characteristics **5**:261
computational challenges **5**:268–269
denaturing/thermal gradient gel electrophoresis (D/TGGE) **5**:264, **5**:264*T*, **5**:265*F*
DNA cloning and sequencing **5**:264*T*, **5**:265–266, **5**:266*F*
enzyme assays **5**:268, **5**:269*T*
fatty acid analysis **5**:267, **5**:268*F*
fluorescent *in situ* hybridization (FISH) **5**:264*T*, **5**:267
intergenic transcribed spacer (ITS) analysis **5**:264–265, **5**:264*T*, **5**:265*F*, **5**:266*F*
membrane hybridization **5**:264*T*, **5**:267
oligonucleotide arrays **5**:264*T*, **5**:267
polymerase chain reaction (PCR) technique **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
process/function measurements **5**:267–268
restriction fragment length polymorphisms (RFLPs) **5**:264*T*, **5**:265
solution hybridization **5**:264*T*, **5**:266–267
substrate utilization patterns **5**:267–268
nucleic acids **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
phylogenetic relationships **5**:253*F*
prevalence **5**:280–282, **5**:281*F*
research areas
Delft School tradition **5**:283–284, **5**:283*F*
research background **5**:283
16S ribosomal DNA (rDNA) sequences **5**:284–285
Woesean Reformation tradition **5**:284–285
research background **5**:252
restoration efforts
diversity evaluations **6**:446–447
establishment criteria **6**:450–451, **6**:450*F*
free-living saprobes **6**:448
importance **6**:448
soil food webs **6**:448–449, **6**:449*F*
spatiotemporal relationships **6**:447–448, **6**:447*F*
symbionts **6**:449–450, **6**:450*F*
moral considerations **3**:267–268
nature-based coastal protection strategies **5**:453–456, **5**:455*F*
neutral theory **5:510–518**
characteristics
dispersal **5**:511
general discussion **5**:511
speciation **5**:511
stochastic drift **5**:511
definition **5**:510, **5**:511
future outlook
general discussion **5**:516
testing programs **5**:516
theoretical developments **5**:516–517
theory applications **5**:517
general discussion **5**:510
goals **5**:510–511
historical background **5**:511–512
modeling approaches
Hubbell's spatially implicit neutral model **5**:512–514
spatially explicity neutral models **5**:514–515
speciation modes **5**:515
species abundance distributions (SADs) **5**:513–514, **5**:513*F*
species–area relationships (SARs) **5**:514
nonneutral extensions **5**:515–516
research summary **5**:517
schematic diagram **5**:512*F*
species abundance distributions (SADs) **5**:510–511, **5**:513*F*
species–area relationships (SARs) **5**:510–511, **5**:514
nitrogen cycle impacts
eutrophication **1**:95–97, **1**:96*F*, **5**:542–544
future outlook **5**:546
genetic diversity loss **5**:540–544
grasslands **5**:542, **5**:542*F*
nitrate deposition **6**:125–126
nitrogen dynamics **5**:544–546, **5**:545*F*
nitrogen oxide (NO_x) acidification **5**:544
North America **5:560–567**
anthropogenic impacts **5**:566
continental characteristics and diversity **5**:560–561
evolutionary changes **5**:565–566
invasive species **5**:566
organism groups
amphibians **5**:565
birds **5**:565
butterflies **5**:565
freshwater fish **5**:565
general discussion **5**:564–565
land mammals **5**:565
plants **5**:564–565
reptiles **5**:565
research summary **5**:566–567
terrestrial bioregions

biodiversity (*continued*)
boundaries and characteristics **5**:561–562, **5**:561*F*
Canadian Shield **5**:561*F*, **5**:562
Caribbean region **5**:561*F*, **5**:564
Eastern North America **5**:561*F*, **5**:562–563
Southernmost North America **5**:561*F*, **5**:563–564
Southwestern North America **5**:561*F*, **5**:563
Western North America **5**:561*F*, **5**:563
origins **1:491–496**
fossil record **1**:491
Phanerozoic Eon
fossil record **1**:411*F*, **1**:412*F*, **1**:491
marine diversity timeline **1**:492*F*, **3**:419*F*
mass extinctions **1**:412–413, **1**:413*F*, **1**:493, **1**:493*F*, **1**:494*F*, **1**:494*T*, **3**:419*F*
originations **1**:495, **1**:495*F*, **3**:425*F*
taxa diversity **1**:411*F*, **1**:412, **1**:412*F*
outside protected areas
challenges **2**:292–293
conservation approaches
classifications **2**:293–294, **2**:294*T*
land and water protection and management **2**:294, **2**:294*F*, **2**:294*T*
living landscape framework **2**:294–295, **2**:294*F*
living lands and waters
biodiversity stewardship **2**:297
biosphere reserves **2**:301
Community-Based Natural Resources Management (CBNRM) **2**:297–298
conservancy/megaconservancy networks **2**:301, **2**:302*F*
conservation approaches **2**:294*F*, **2**:295–297
conservation contributions **2**:301–303
ecotourism **2**:295–296, **2**:296*F*
educational programs **2**:295
financial sources **2**:295–296
payment for ecosystem services (PES) **2**:298
private conservation areas **2**:296–297, **2**:297*T*
property rights **2**:295
systematic conservation planning **2**:293
working lands and waters
biodiversity-friendly farming **2**:299–300, **2**:300
biodiversity stewardship **2**:300
biosphere reserves **2**:301
conservancy/megaconservancy networks **2**:301, **2**:302*F*
conservation approaches **2**:298–299
conservation contributions **2**:301–303
management practices **2**:299
multifunctional landscapes **2**:294*F*, **2**:300–301
pest control **1:373–385**
agricultural intensification effects **1**:374–375
agricultural landscapes
characteristics **1**:380–381
habitat heterogeneity **1**:381–382, **1**:381*F*
natural habitat proximity **1**:381
noncrop habitats **1**:380–381
background information **1**:373–374
conservation biological control **1**:382
influencing factors **1**:374*F*
multiple trophic levels **1**:380
predator diversity
complementarity **1**:377–378
enhancement factors **1**:379–380
facilitation **1**:377
functional diversity (FD) **1**:378
functional redundancy **1**:378
insurance hypothesis **1**:378
intraguild interference **1**:378–379, **1**:378*F*
risk reduction/risk enhancement **1**:376–377
sampling/selection effects **1**:377
research background **1**:373–374
summary discussion **1**:382–383
vegetation diversity
crop diversity **1**:375–376
natural enemies hypothesis **1**:375–376
noncrop plants **1**:376, **1**:376*F*
pharmaceuticals **5:703–715**
characteristics
marine organisms **5**:711
microbial products **5**:710–711
terrestrial animals **5**:711
terrestrial plants **5**:711–713, **5**:712*T*
ethical perspectives
bilateral research agreements **5**:714
citizens' rights **5**:715
Convention on Biological Diversity (1930) **5**:713–714
indigenous intellectual property rights **5**:714–715
sick and afflicted populations **5**:715
evolutionary perspective **5**:703–705
general discussion **5**:703
historical perspective
ancient and early research **5**:705
herbals and medicinal plants **5**:705–706
modern drug discovery approaches
bioassays **5**:707–708
isolation techniques **5**:708
research background **5**:706–707
structural determination methods **5**:708
vouchering **5**:707
rational screens
characteristics **5**:708–709
ecological screens **5**:709–710
ethnobotanical screens **5**:710
phylogenetic screens **5**:709
random screens **5**:709
phylogenetic diversity **6**:22
plant breeding **1:459–469**
commodity agriculture **1**:467
controversies **1**:459
crop homogeneity **1**:461*F*
exploration and trade **1**:465–466
gene pool concepts **1**:463–465, **1**:465*F*
genetic bottlenecks **1**:463
green revolution genetics **1**:467–468
legal frameworks **1**:466–467
modern scientific plant breeding **1**:460–461, **1**:461*F*
phenotypic variation **1**:461–463, **1**:462*F*
quality guidelines **1**:461*F*
reduced genetic diversity **1**:461–463
research summary **1**:468
plants **6:56–64**
angiosperms **6**:60
biodiversity distribution and loss **6**:79
biodiversity patterns
assessment measures **6**:60–61
endemism **6**:61
biodiversity restoration **6**:456–457
characteristics and classifications **6**:56
climate effects **2:115–127**
annual mean temperature **2**:119–120, **3**:119
area–productivity relationship **2**:117–118, **3**:118
biogeographic history **2**:125–126
competition **2**:123
disturbances **2**:115–117, **2**:123, **3**:118
energy availability **2**:123–124
experimental research **2**:126
extreme events **2**:124–125, **3**:124, **3**:125
general discussion **2**:115
genomic research **2**:126
global-scale distribution **2**:118–119
latitudinal gradients **2**:118–119, **3**:119
net primary production (NPP) **2**:117–118, **2**:120, **3**:118, **3**:121
observational research **2**:121–123, **3**:122
precipitation **2**:120, **3**:120
species–area relationships (SARs) **2**:115, **3**:116, **3**:117
summary discussion **2**:126
theoretical perspectives **2**:120–123
functional roles **6**:56–57
gymnosperms **6**:59–60
major threats
anthropogenic impacts **6**:62
conservation programs **6**:63
habitat destruction **6**:62
intrinsic factors **6**:61–62
natural disturbances **6**:62–63
species interactions **6**:61–62
nonvascular plants **5**:528, **6**:58–59, **6**:58*T*
plant–human interactions **6**:57–58
research outlook **6**:63–64
seedless vascular plants **6**:59
precautionary principle **1**:408
provisioning services
basic concepts **2**:490*T*
desertification **2**:490*T*

marine ecosystems **5**:451–452, **5**:452*T*
psychrophiles **6**:300
regulating services
basic concepts **2**:490*T*
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
religious traditions **6:380–388**
Asian traditions **6**:383–384
background information **6**:380
basic concepts **6**:380–381
Buddhist declaration on nature **6**:383
Hindu declaration on nature **6**:383
monotheistic traditions
Christian declaration on nature **6**:381, **6**:381
Jewish declaration on nature **6**:381, **6**:382
Judeo–Christian tradition **6**:381
Muslim declaration on nature **6**:381, **6**:382
stewardship **6**:381–383
pantheistic traditions
ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
indigenous peoples **6**:384, **6**:384–385
traditional ecological knowledge **6**:385
traditional social restraints **6**:385–386
research summary **6**:387–388
traditional social protection
background information **6**:385–386
biological communities/habitat patches **6**:386
critical life history stages **6**:386
long-lived (K-selected) species **6**:386
stewardship **6**:386
research approaches
community approach **2**:648, **2**:649*T*
ecosystem approach **2**:648–649, **2**:649*T*
general discussion **2**:648
individual approach **2**:648, **2**:649*T*
population approach **2**:648, **2**:649*T*
resource partitioning **6:429–437**
Clements–Gleason debate **6**:429
community dominance–diversity relationships **6**:431–432, **6**:432*F*
diversity assessment framework **6**:429–430, **6**:430*F*
diversity mechanisms
chance **6**:435
frequency-dependent effects **6**:435
general discussion **6**:435
spatial structure **6**:435–436, **6**:436*F*
niche differentiation
empirical evidence **6**:432–433, **6**:433*F*
relevance **6**:434–435
spatial heterogeneity **6**:432–433, **6**:433*F*
temporal heterogeneity **6**:433–434, **6**:434*F*
theoretical perspectives **6**:430–432, **6**:431*F*, **6**:432*F*
research background **6**:429
research summary **6**:436–437
restoration **6:453–460**
biologically diverse landscapes
landscape objectives **6**:458
novel/experimental opportunities **6**:458–459
regional efforts **6**:458
community restoration
abiotic constraints **6**:456–457
basic concepts **6**:456–457
biotic constraints **6**:457
management solutions **6**:457–458
genetic considerations
captive breeding programs **6**:454
challenges **6**:454
gene pool manipulation **6**:454
greenhouse propagation **6**:454
guidance **6**:454
research background **6**:453–454
goals and priorities **6**:453
single species reintroduction
abiotic constraints **6**:455
biotic constraints **6**:455
management solutions **6**:455–456
purposes **6**:454–455
riparian zones **6**:461–463, **6**:463*F*
scaling laws
alpha diversity **1**:402–403, **6**:534, **6**:534
beta diversity **1**:402–403, **6**:534, **6**:534
evolutionary biology **6**:535
gamma diversity **1**:402–403, **6**:534, **6**:534
genetic variability **6**:533–534
marine ecology **6**:535
scalable quantities **6**:532
scaling relations
general discussion **6**:538
number scales completely with area ($N \cong A^1$) **6**:538
species diversity distribution ($S(n_i \leq n_k) \cong S(A_i \leq A_k) \cong H(A_i \leq A_k)$) **6**:538
species number scales incompletely with area ($S \cong A^{\beta}$) **6**:538
species number scales incompletely with numbers ($S \cong N^{1/4}$) **6**:538
species number scales with total numbers ($S \cong N^{1/4} \cong A^{1/4}$) **6**:538
steeper scaling exponent ($S \cong A^{\beta}$, where $\beta_{island} > \beta_{block}$) **6**:538
steeper scaling exponent ($S \cong A^{\beta}$, where $\beta_{province} > \beta_{block}$) **6**:538
taxonomic diversity distributions **6**:532–533, **6**:533*F*
terrestrial ecology **6**:535
sociocultural factors **6:563–570**
language **4**:270–271, **4**:270*T*, **6**:566
religious traditions **6**:563–566
research summary **6**:569
resource management
ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
indigenous peoples **4**:272, **6**:384–385, **6**:384, **6**:566–568
traditional belief systems **6**:386, **6**:567–568
urban versus rural perceptions **6**:568–569, **6**:569*T*
soils
biotic diversity **6**:582*F*, **6**:583
disturbance effects **6**:588–589
domains
bacteria **6**:583–584
characteristics **6**:583
eukaryotes **6**:585–586, **6**:586*T*
phylogenetic relationships **6**:583
prokaryotes **6**:583
Protozoa **6**:586, **6**:587*F*
tree of life **6**:584*F*
evolutionary history **6**:582–583
research summary **6**:589
spheres of influence **6**:582*F*, **6**:583
viruses **6**:584–585
spatial changes **1**:404
species diversity
basic concepts **1**:401–402
measurement and analysis **1**:402–403, **2**:650, **6**:534, **6**:534
species packing **4:642–648**
basic concepts
ecological niche model **4**:643
general discussion **4**:644–645
laboratory experiments **4**:643–644
Lotka–Volterra model **4**:643
MacArthur–Levins model **4**:645
multiple species model **4**:645
niche separation–anatomical difference relationship **4**:645–646
biological metaphor **4**:642
research summary **4**:646–647
species diversity **4**:646
supporting services
basic concepts **2**:490*T*
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
sustainability **7:71–84**
biodiversity–population density relationship **7**:74*F*
carrying capacity **7**:75
demographic trends **7**:72*F*
ecological resilience **7**:80–81, **7**:81*F*
ecosystem services **7**:76*T*
global trends **7**:81–83, **7**:82*F*, **7**:83*F*
human well-being indicators **7**:77–80, **7**:78*F*
as natural capital component **7**:76–77
poverty reduction **7**:77–80
sustainability–biodiversity relationship **7**:71–75
sustainable development **7**:75
theory of inclusive wealth **7**:73, **7**:75–76
systems theory
basic concepts **2**:651–652
stability dynamics
basic concepts **2**:652–653
local asymptotic stability **2**:652–653
transient dynamic behavior **2**:653, **2**:653*F*
steady-state dynamics
basic concepts **2**:651–652, **2**:651*F*
colonization rates **2**:651–652, **2**:652*F*

biodiversity (*continued*)
community structure **2**:652
ecosystem processes **2**:652, **2**:652*F*
extinction rates **2**:651–652, **2**:652*F*
number of species **2**:651–652, **2**:652*F*
population densities **2**:652
species composition **2**:652
taxonomic diversity **7**:90–91, **7**:91*F*
temporal changes **1**:404
terrestrial ecosystems
nitrogen deposition **5:519–536**
acidification **5**:522, **5**:527*F*
anthropogenic impacts **5**:520–522, **5**:522*F*
biodiversity impacts **5**:522
biological nitrogen fixation **5**:520*F*
direct foliar toxicity **5**:523–525, **5**:527*F*
ecosystem-specific effects **5**:526–528
environmental impact **5**:519–520
eutrophication **5**:522
global distribution **5**:524*F*
global nitrogen budget **5**:523*F*
global nitrogen cycle **5**:520–522, **5**:521*F*
hotspots **5**:526*T*
influencing factors **5**:526–528
nitrogen as nutrient **5**:520, **5**:520*F*
primary drivers **5**:522, **5**:525*F*, **5**:527*F*
research methodologies **5**:525–526, **5**:528*T*, **5**:529*F*
secondary stressors **5**:522–523, **5**:527*F*
taxa-specific responses **5**:528
thermal ecosystems **4**:118–119
traditional conservation practices **7**:234
urban–suburban biodiversity **7:304–313**
anthropogenic activities and impacts **7**:307*T*, **7**:308*T*
city-scale decisions **7**:309–310
dryland syndrome **2**:483*F*
evaluation criteria **7**:305–306
global scale **7**:310
governmental policies **7**:308*T*
hinterlands **7**:309–310
historical perspective **7**:304
individual and household scales **7**:306–308
Millennium Ecosystem Assessment (MEA) **7**:305–306
neighborhood scale **7**:308–309
regional scale **7**:310
settlement–biome interactions **7**:304–305
socioecological theories **7**:310–312
valuation measures **3:55–58**, **7:167–179**
accounting price **7**:172
biological resource exploitation **7**:167–168
discontinuous value functions **7**:175–177, **7**:177*F*
ecological functions **7**:168–169
economic value **3**:55
gross domestic product (GDP) **7**:173–174, **7**:174*F*
institutional failure **7**:170–172
local versus global stability **7**:170–172
market failure **7**:169–170
market versus nonmarket values **3**:56–57
methodologies **3**:55–56
Millennium Ecosystem Assessment (MEA) **7**:168–169
necessity versus luxury perspective **7**:173–174
net primary production (NPP) **3**:58
nonuse benefits **3**:57
option value **7**:172
productive benefits **3**:56–57
project evaluation **7**:172–173
recreational and esthetic benefits **3**:57
resource evaluation **7**:172–173
site-specific valuations **3**:57–58
social cost–benefit analysis **7**:172–173
social value **3**:55
substitute services costs **3**:57
substitution possibilities **7**:174–175
summary discussion **7**:177–178
biodiversity monitoring and assessment programs (BMAP) framework **3:545–559**
basic concepts
appropriate indicator selection **3**:548–549
functional role **3**:545–546
goals and objectives **3**:546, **3**:547*F*, **3**:548*T*
guiding principles **3**:546–548
legal and institutional framework **3**:549
process flow diagram **3**:547*F*
spatial and temporal scales **3**:546
conceptual ecosystem and landscape models **3**:549, **3**:550*T*
general discussion **3**:545
monitoring protocols design
adaptive management
basic concepts **3**:556
decision-making strategies **3**:556–557
follow-up monitoring **3**:557
management action selection **3**:556
management approaches **3**:556
assessment and monitoring components **3**:549–550, **3**:550*T*
data analysis
initial data exploration **3**:554
methodologies **3**:553–554
nonparametric tests **3**:554–555
parameter estimation **3**:554
significance tests **3**:554
data collection **3**:551
decision-making strategies
adaptive management **3**:556
basic principles **3**:555
communication approaches **3**:555–556
data and information management **3**:555
information management goals **3**:555
evaluation assessments **3**:557
lessons learned **3**:557
sampling designs
fundamental principles **3**:551
parameter selection **3**:552–553
sampling objectives **3**:551–552
sampling units **3**:552–553
spatiotemporal scales **3**:552
selection criteria **3**:550–551
stakeholder involvement **3**:556
bioenergetics **7**:264–266
biofuels **1:528–539**
algae-based biodiesel fuel **1**:592–593
biodiversity-friendly energy-crop production strategies **1**:535–536
biodiversity restoration **6**:459
biofuel production **1**:528–529, **1**:529*F*, **1**:529*T*
crop production impacts **4**:293
crop selection and management
biodiversity management strategies **1**:533
biodiversity responses **1**:529–530, **1**:530*F*
first-generation biofuel crops
economic importance **1**:529*T*, **1**:530
feedstock management **1**:530–531
feedstock selection **1**:530
second-generation biofuel crops
economic importance **1**:530*F*, **1**:531–532
native polycultures **1**:532–533
perennial monocultures **1**:532
short-rotation woody biomass **1**:531–532
direct land-use change (DLUC) **4**:293–294
energy sprawl concerns **1**:528–529, **1**:529*F*
energy use impacts **3**:261–262
Global Trade Analysis Project (GTAP) model
consumption-side structure **4**:540*F*
emission factors **4**:548–549
ethanol production
estimated emissions **4**:549–550, **4**:549*T*
final analysis **4**:551–552, **4**:551*T*
marginal and average estimated emissions (first simulated group) **4**:550, **4**:550*T*
marginal and average estimated emissions (second simulated group) **4**:550–551, **4**:550*T*
marginal and average estimated emissions (third simulated group) **4**:551, **4**:551*T*
research limitations **4**:553
research summary **4**:552–553
well-to-wheel emissions **4**:551–552, **4**:552*T*, **4**:553*T*
land cover and land-use activities **4**:542*F*
livestock production structure **4**:542*F*, **4**:543*F*, **4**:543*T*
modeling process **4**:539–540
production-side structure **4**:541*F*
simulation models
basic concepts **4**:544–545
2001 simulation model **4**:544–545, **4**:544*T*, **4**:545*T*

2001–2006 simulation **4**:545–547, **4**:546*T*, **4**:547*T*
2006–2015 simulation model **4**:547–548, **4**:548*T*, **4**:549*T*
indirect land-use change (ILUC) **4:293–297**
avoidance measures **4**:296–297
carbon dioxide (CO_2) emissions
biomass emission estimation **4**:295–296
economic models **4**:294–295, **4**:294*F*, **4**:295*F*
food price increases and rationing **4**:296
global market impacts **4**:294–295, **4**:295*F*
land cover change emissions **4**:295–296
land cover maps **4**:295
total emission–fuel quantity relationship **4**:296
estimation uncertainties **4**:296
greenhouse gas emissions **4**:293–294
marginal land usage **4**:296–297
landscape–scale biodiversity considerations
agroenergy sprawl **1**:534–535
general discussion **1**:533–534
spatial considerations **1**:533–534
life cycle analysis (LCA) **1**:535–536, **4**:628–629
oil palm **1**:543–547, **5**:602
policy development impacts **1**:536
research summary **1**:536–537
wildlife habitat restoration **1:540–551**
cropping system comparisons **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
government policies and regulations **1**:540–541
land allocations
global-scale assessments **1**:547–549
landscape-scale assessments **1**:547–549, **1**:548*F*
local-scale assessments **1**:549
regional-scale assessments **1**:547–549
land-use change impacts
croplands and pasturelands **1**:542–543, **1**:542*F*
marginal, degraded, and abandoned lands **1**:543, **1**:544*F*
native ecosystems **1**:541–542, **1**:545*F*
prime agricultural lands **1**:543
oil-palm plantations **1**:540–541
summary discussion **1**:549
biogeochemical cycles **1:552–564**, **6:147–156**
carbon cycle **1:674–684**
atmospheric carbon dioxide (CO_2) **6**:147–149, **6**:148*F*
background information **1**:674
basic concepts
carbon dioxide (CO_2) concentrations **1**:674–676, **1**:675*F*, **6**:147–149, **6**:148*F*
carbon dioxide (CO_2) lifetime estimation **1**:680
deforestation flux **1**:676–678, **1**:677*F*, **1**:679*T*
fossil fuel emissions **1**:283–284, **1**:675*F*, **1**:678, **1**:679*F*
land sinks **1**:674–676
ocean sinks **1**:676
regional carbon stocks **1**:678*F*
schematic diagram **1**:675*F*, **6**:148*F*
transport systems **1**:675*F*, **1**:676
volatile organic compounds (VOCs) **1**:679–680, **1**:680*T*
weathering processes **1**:678–679
biological carbon (C) pump **2**:615–618, **2**:616*F*
climate change effects **6**:147–149
dynamic global vegetation models (DGVMs) **2**:679
ecosystem alterations
climate change effects **6**:149–150
coastal/marine ecosystems **6**:150–151
forest ecosystems **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
future outlook **1**:683–684, **1**:683*F*
management strategies
carbon capture and storage **1**:681
deforestation reduction **1**:682–683
government regulation **1**:680–681
iron fertilization **1**:681
Reducing Emissions from Deforestation and Forest Degradation (REDD) **1**:682–683
soil respiration **1**:681–682
tree plantations **1**:682
crustal elements
biological assimilation **1**:554–555
biological transport **1**:554
carbon pools **1**:554, **1**:555*T*
elemental partitioning **1**:553, **1**:554*T*
hydrological cycle **1**:553–554, **1**:555*F*
origins **1**:552–553
orogenesis **1**:554
phase state transitions **1**:553
relative abundances **1**:553*F*
vulcanism **1**:554
weathering processes **1**:553
ecosystem function measurements **3**:68–70
environmental impacts **3**:282
estuarine ecosystems **3**:301
evolutionary development **1**:563–564
global livestock **6**:342
historical perspective **6**:147
key reactions and cycles
biological pumps **1**:559–560, **1**:560*F*
biomineralization **1**:556–557, **1**:557*F*
carbon-based processes **1**:556, **4**:527
energy flow **3**:242–243
hydrolysis **1**:555–556
net primary production (NPP) **1**:559, **1**:560*T*
nitrogen cycle **1**:560–561, **1**:562*F*, **4**:527, **6**:581–582, **6**:582*F*
oceanic elemental distribution **1**:561
oceanic solubility **1**:559–560, **1**:560*F*
oxidation-reduction (redox) reactions **1**:555–556, **1**:556*T*
oxygen cycle **1**:557–558, **1**:558*F*
phosphorus cycle **1**:563
sequestration and burial **1**:558–559
sulfur cycle **1**:561–563
trace metals **1**:563
metabolic pathways **1**:552
nitrogen cycle
acid rain **6**:151–152
aquatic ecosystems **6**:152, **6**:153*F*
dynamic global vegetation models (DGVMs) **2**:680
forest ecosystems **6**:152
key reactions **1**:560–561, **1**:562*F*, **6**:581–582
landscape legacies **4**:527–528
nitrogen oxides (NO_x) **1**:560–561, **1**:562*F*, **6**:151–152
nitrogen saturation **6**:152
schematic diagram **6**:151*F*
tropospheric ozone precursors **6**:152–154
sulfur cycle **1**:558–559, **1**:561–563, **6**:154–155, **6**:154*F*
summary discussion **6**:155
trace pollutants **6**:155
ultraviolet radiation effects **7**:301
biogeography **1:576–587**
Australia **1**:297–298, **1**:298*F*, **1**:299*F*
background information **5**:271
basic concepts **1**:576
biodiversity conservation **1**:585–586
biogeographical models **1:565–575**
characteristics and functional role **1**:565–566, **1**:566*F*
correlative models **1**:566–569, **1**:566*F*
future research areas **1**:572–574
stochastic simulation models **1**:569–572, **1**:572*F*
birds **1**:616–617
Central America **2**:23–24, **2**:24*F*
cladistic biogeography
basic concepts **7**:344–345, **7**:344*F*
component analysis **7**:348–349, **7**:349, **7**:349*F*, **7**:350*F*, **7**:350*T*
dispersal biogeography **2**:555
Heterandria
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
progression rule **7**:344*F*, **7**:345–346, **7**:346*F*, **7**:347*F*
research summary **7**:351
Xiphophorus
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
definition **1**:576
dispersal biogeography **7**:343, **7**:343*F*, **7**:347*F*
harmful algal blooms **1**:470
historical research
background information **1**:576–578
eighteenth century **1**:577–578
nineteenth century **1**:578–579, **1**:579*F*

biogeography (*continued*)
persistent themes **1**:577*T*
island biogeography **4:395–407**
basic concepts **4**:395
data analysis
dendograph approach **4**:403*F*, **4**:404*F*, **4**:405
Hawaii Volcanoes National Park **4**:402*F*, **4**:403–405, **4**:404*F*, **4**:405*F*
improved display methods **4**:404*F*, **4**:405, **4**:405*F*
two-way table approach **4**:402*T*, **4**:404*F*, **4**:405
experimental sampling design
horizontal approach **4**:406
transect approach **4**:405–406, **4**:405*F*
vertical approach **4**:406
geobotanical research **4**:395, **4**:396*T*
island–mainland differences model **6**:645–646
living dead **6**:81
MacArthur–Wilson equilibrium theory **1**:583–585, **2**:239, **2**:239*F*, **3**:5, **4**:395–396, **4**:396*F*
mammalian evolution **4**:699
PABITRA (Pacific-Asia Biodiversity Transect) initiative **4**:406–407, **4**:406*F*
protected area systems **6**:493–494
quantitative sampling methods
challenges **4**:400
count-plot method **4**:400, **4**:401–403
distance methods **4**:400–401, **4**:401–403
point-centered quarter method **4**:401
sample plot size **4**:401*F*, **4**:403
wandering quarter method **4**:401
relevé sampling method
basic concepts **4**:399
cluster analysis **4**:403*F*, **4**:404*F*
predefinied strata **4**:399
size determinations **4**:399, **4**:400*F*
species determinations **4**:399
species quantity estimations **4**:399–400
scaling laws
spatial scales **4**:395
timescales **4**:395
species–area relationships (SARs) **6**:645–646
taxonomic versus functional diversity **2**:652*F*, **2**:655, **2**:655*T*
theoretical perspectives
biome theory **4**:396–397, **4**:397*F*
MacArthur–Wilson equilibrium theory **1**:583–585, **4**:395–396, **4**:396*F*
succession/regression theory **4**:397–398, **4**:398*F*
vegetation analysis
data analysis **4**:403–405
floristic checklist methods **4**:399
quantitative methods **4**:400
relevé method **4**:399
sampling methods **4**:399
Lepidoptera **5**:389–390, **5**:389*T*
mass extinction impacts **3**:421–422
microbial biogeography **5:271–279**
background information
Baas-Becking hypothesis **5**:271
general discussion **5**:271
molecular genetics **5**:271–272
distribution patterns
depth gradients **5**:273–274
distance–decay relationships **5**:274, **5**:274*F*
elevational gradients **5**:273–274
endemism **5**:272, **5**:273*F*
latitudinal gradients **5**:272–273
phylogenetic relationships **5**:273*F*
research background **5**:272
seasonal dynamics **5**:275
taxa–area relationships **5**:274–275, **5**:274*F*
taxa–time relationships **5**:275
future research considerations
habitat type **5**:278
pattern interpretation **5**:277–278
spatial scale **5**:278
taxonomic focus and resolution **5**:277–278
pattern interpretation
challenges **5**:275–276, **5**:276–277
contemporary environmental factors **5**:276, **5**:276*F*
future research considerations **5**:277–278
historical factors **5**:276, **5**:276*F*
hypothetical relationships **5**:275–276, **5**:276*F*
summary discussion **5**:278
regional species richness controls **2**:599–601, **2**:600*F*
South America
early history **3**:109
late Tertiary **3**:109
Quaternary period **3**:109–110
twentieth century research
geographic template
climatic regions **1**:581*F*, **1**:582*F*
continental drift **1**:583
environmental characteristics **1**:579–583
mean monthly water temperatures **1**:582*F*
plate tectonics **1**:580–583, **1**:584*F*
Pleistocene glacial cycles **1**:583, **1**:585*F*, **1**:586*F*
seafloor spreading **1**:583, **1**:584*F*
soil types **1**:581*F*
MacArthur–Wilson equilibrium theory **1**:583–585, **4**:395–396, **4**:396*F*
species–area/species–latitude relationships **1**:583–585
twenty-first century research **1**:583–585
vicariance biogeography **7:342–351**
basic concepts **7**:343, **7**:343*F*, **7**:346–348
component analysis **7**:348–349, **7**:349, **7**:349*F*, **7**:350*F*, **7**:350*T*
dispersal–vicariance analysis (DIVA) **2**:556, **2**:556*F*
Heterandria
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
historical biogeography **7**:342–343
Hooker's hypothesis **7**:343
Mesoamerican poeciliid fish **7**:348, **7**:348*F*
phylogenetic analyses **6**:22
plate tectonics **7**:343–344
research summary **7**:351
theoretical perspectives **2**:539–540, **7**:343–344
Xiphophorus
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
Biogeography of Deep-Water Chemosynthetic Ecosystems (ChEss) **2**:19–20, **2**:20*F*
bioindicators **3**:121, **3**:124–125
Bioko black colobus **6**:256–258*T*
Bioko drill **6**:256–258*T*
Bioko needle-clawed galago **6**:256–258*T*
Bioko Preuss's monkey **6**:256–258*T*
Bioko squirrel galago **6**:256–258*T*
biologging **2**:18–19, **2**:19*F*
Biological Dynamics of Forest Fragments Project (BDFFP) **4**:663
biologically prepared learning **1**:37–38
biological pumps **1**:559–560, **1**:560*F*
biological species concept **6**:620
bioluminescence **5**:121–123, **5**:122*F*
biomarkers **3**:121, **3**:123*T*
biomass
biomass turnover **3**:242–243
energy use impacts **3**:261
megaherbivores **5**:227–228, **5**:227*F*
Tilman biomass study **1**:386–388, **1**:387*F*
world energy use **3**:252*T*, **3**:253*T*
biomedical research **6**:254–255
biomimetics
gecko tape **1**:591
self-cleaning lotus paint **1**:590
shark swimsuits **1**:591
termite-mound architecture **1**:590–591
Velcro **1**:590
biomineralization **1**:556–557, **1**:557*F*
Biomphalaria glabrata **1**:368, **1**:368*F*
Biomyxa **6**:291
BIO NET-INTERNATIONAL **3**:140
bionomic equilibrium **2**:150–151, **2**:150*F*, **2**:151*F*
bioparasitic nongovernmental organizations (NGOs) **1**:597
bioparks **7**:473–474
biophilia **6**:563–564
biophilia versus videophilia
comparison studies **7**:242–244
future outlook **7**:250–251
future research outlook **7**:253–256, **7**:255*F*
historical trends **7**:242, **7**:243*F*
videophilia **7**:244
bioprobes **3**:121, **3**:122*F*

bioprospecting **1**:588–599
  commercial development **1**:588
  definition **1**:588
  economic importance **1**:393–395
  ethical perspectives
    benefits sharing progams **1**:598
    conservation practices **1**:598–599
    fair trade practices **1**:598–599
    intellectual credit **1**:599
    international agreements
      bioparasitic nongovernmental organizations (NGOs) **1**:597
      Convention on Biological Diversity (1930) **1**:596
      data return programs **1**:597–598
      indigenous knowledge **1**:596
      indigenous negotiators **1**:597
      Prostratin case study **1**:596–597
      public health care practices **1**:598
      2010 conference of the parties (COP(10)) meeting **1**:597
    sacred symbols protection **1**:599
    Shaman Pharmaceuticals **1**:598
    sustainable development **1**:598–599
  historical perspective **1**:588–589
  market economy **5**:148
  patents and intellectual property
    natural products
      exclusionary rights **1**:594–595
      Myriad Genetics legal case **1**:595
      patentability **1**:595
    traditional knowledge systems
      ayahausca **1**:595
      patent limits and costs **1**:595–596
      patent pools **1**:596
  targets, screens, and diagnostic analyses
    antifungal crop treatments **1**:590
    basic concepts **1**:589
    biomimetics
      gecko tape **1**:591
      self-cleaning lotus paint **1**:590
      shark swimsuits **1**:591
      termite-mound architecture **1**:590–591
      Velcro **1**:590
    diagnostic procedures and techniques
      air pollution monitors **1**:591
      green fluorescent protein (GFP) **1**:591–592
      litmus paper **1**:591
    disease prevention
      anti-schistosomiasis soap **1**:592
      motor neuron disease **1**:592
      seaweeds **1**:592
    environmental remediation
      algae-based biodiesel fuel **1**:592–593
      oil-eating bacteria **1**:592
    genetic analysis
      DNA fingerprinting **1**:593
      Pfu DNA polymerase **1**:593
      Taq polymerase **1**:593
    insecticides
      fish poisons **1**:589
      neem tree **1**:589–590
    pharmaceuticals
      addiction treatment drugs **1**:594
      antidiarrheal drugs **1**:594
      antimalarial drugs **1**:593–594
      cancer treatment drugs **1**:594
      human immunodeficiency virus (HIV) **1**:594
      pain treatment drugs **1**:594
    seagrass insulation **1**:593
bioremediation **6**:594
bioreporters **3**:121
bioresmethrin **5**:694, **5**:699*T*
BiosciEdNet (BEN) **3**:139
biosensors **3**:121
biosphere people **6**:221*T*
biosphere reserves **2**:301, **3**:279, **6**:489
biospheric life support system **7**:43
Biota-Fapesp **1**:506
biotic homogenization **3**:284, **3**:701, **6**:100–101*F*, **6**:101
biotic similarity **5:195–211**
  abundance-based similarity indices **5**:208–209, **5**:209*F*
  basic concepts **5**:197
  incidence-based similarity indices **5**:207–208
  phylogenetic similarity indices **5**:210
  total abundance–shared species-based similarity indices **5**:209–210
biotopes
  Amazon ecosystems **1**:249–250
  endemism **3**:233
Biovularia **6**:44*T*
bipirimate **5**:695
Biploporus minutus **7**:220*T*
bipyridilium herbicides **4**:90*T*, **4**:92*T*
bipyridiliums **5**:696, **5**:699*T*
birch
  boreal forests **1**:628–630, **1**:630*T*, **7**:212–213, **7**:218*F*, **7**:219*F*
  North America **7**:118
birch leaf miner **4**:182–183
bird flu **3**:411
BirdLife International **2**:231–232, **4**:325–326
birds **1:600–625**
  agricultural intensification effects **1**:129*F*
  air pollution impacts **1**:139
  alpine ecosystems **1**:154–155
  anthropogenic activities and impacts **3**:401*F*, **7**:307*T*
  Arctic ecosystems **1**:231*T*
  Asia
    dry Asian regions **1**:268
    endemic species **1**:261*T*
    threatened species **1**:278*T*
  beak shape **1**:604*F*
  biodiversity generation **1**:432, **1**:432*F*
  biodiversity–population density relationship **7**:74*F*
  biogeographical models **1**:565–566, **1**:566*F*
  biogeographic research **1**:578–579
  boreal forests **1**:631–632
  brood parasitism **5:501–509**
    adaptive behaviors **5**:501
    background information **5**:501
    conspecific brood parasitism (CBP)
      definition **5**:501
      prevalence **5**:508–509
      species survey **5**:502–503
    cowbird management **5**:507
    host–parasite coevolution
      cowbird–host systems **5**:504
      cuckoo–host systems **5**:503
      modeling approaches **5**:506
    host population dynamics **5**:506–507
    major groups **5**:502*T*
    number of species **5**:502*T*
    research outlook **5**:509
    social behavior
      foraging ecology **5**:507
      mating systems **5**:507–508
      spacing behaviors **5**:508
      vocal behaviors **5**:508
    species survey
      black-headed ducks **5**:502, **5**:502*T*
      conspecific brood parasitism (CBP) **5**:502–503
      cuckoo-finches **5**:502, **5**:502*T*
      Cuculidae **5**:501–502, **5**:502*T*
      Icteridae (cowbirds) **5**:502, **5**:502*T*
      Indicatoridae **5**:502, **5**:502*T*
      viduine finches **5**:502, **5**:502*T*
Census of Marine Life project **2**:12*F*
Central America **2**:25*T*
characteristics **7**:340–341
climate change studies **2**:66, **2**:67*F*, **2**:92–94, **2**:93*F*
countryside biogeography studies **2**:354–356, **2**:356*F*
deoxyribonucleic acid (DNA) evolution **2**:56*F*, **2**:57
desert ecosystems **2**:472–474, **2**:472*F*, **2**:475*F*
diets and feeding **1**:604, **1**:604*F*, **1**:604*T*
  as disease vectors **1**:362–365, **1**:364*F*, **1**:365*F*
diversity loss
  extinctions
    conservation approaches **1**:623
    current threats **1**:622–623, **1**:622*F*
    historical causes **1**:620–621, **1**:620*T*
    human colonization **1**:621–622
    natural causes **1**:620
  geographical variation **1**:624
  introduced species **1**:624
  species reductions **1**:623–624, **1**:623*F*
domesticated ruminants **1**:710
ecological diversity **1**:602–605
elevational range shifts **4**:605*T*
El Niño Southern Oscillation (ENSO) events
  precipitation increases **3**:158–159
  wildfire effects **3**:156–158
end-Cretaceous mass extinction **5**:173–175
endemism **3**:231*F*, **3**:232*T*, **3**:233*F*, **5**:592*T*
energy requirements **6**:404*F*
energy storage
  adipose tissue **7**:33–35
  migration **7**:35–36
  reproduction and lactation **7**:37–38
environmental change indicators

birds (*continued*)
farmland birds **4**:259
general discussion **4**:258–259
migratory songbirds **4**:259
peregrine falcons **4**:258–259
eutrophication impacts **3**:365–366
extinctions
conservation approaches **1**:623
current threats **1**:622–623, **1**:622*F*
estimation methods **4**:659*F*
global and regional trends **7**:182
historical causes **1**:620–621, **1**:620*T*
human colonization **1**:621–622
natural causes **1**:620
risk factors **2**:75*F*
fire/fire regimes **3**:438
flight constraints
body shape **1**:601
feathers **1**:600–601
metabolic rate **1**:601
size **1**:601–602, **1**:602*F*
wing shape **1**:601
flightless birds
Australia **4**:727–728
island biogeography **1**:605–606
Madagascar **4**:728–729, **5**:365
Oceania **4**:723*F*, **4**:730–731
food habits **3**:127–128
food webs **7**:266*F*, **7**:267*T*
forest canopies **3**:513–514
forest management impact **1**:453
freshwater ecosystems **3**:571–572, **3**:578–579, **3**:579*T*
fruit and seed dispersal **6**:50, **6**:51*T*
genetic diversity **3**:665*F*
geographical variation
altitude **1**:613
direct animal comparisons **1**:614
diversity loss **1**:624
habitat–bird species diversity relationship **1**:613
latitude **1**:613–614, **1**:614*F*
tropical–polar latitudinal gradient causes **1**:614–615
global wildlife trade **3**:409*T*
government legislation and regulations
Bald Eagle Protection Act (1940) **3**:713*T*, **3**:716
Migratory Bird Treaty Act (1918) **3**:713*T*, **3**:715–716
Wild Bird Conservation Act (1992) **3**:713*T*, **3**:716
grasslands **4**:3–4
habitat–bird species diversity relationship
feeding behaviors **1**:611*T*
foliage height diversity **1**:609–610, **1**:611*F*, **1**:611*T*
foraging behaviors **1**:611–612
geographical variation **1**:613
habitat comparisons **1**:608, **1**:608*F*
habitat complexity **1**:609–611, **1**:609*F*, **1**:610*F*
habitat structure **1**:611–612, **1**:611*F*, **1**:611*T*
productivity **1**:613
similar ecosystems–different continents **1**:608–609, **1**:609*T*
succession **1**:612–613, **1**:612*F*, **1**:612*T*, **1**:613*F*
Holocene postglacial spread **2**:550, **2**:551–552*T*
Important Bird Areas (IBAs) **7**:203–204, **7**:204*F*
indicator species **7**:206–207*F*
introduced parasites–species decline relationships **2**:528*T*
introduced species
Australia **1**:80
Brazil **1**:83
India **1**:82
South Africa **1**:81
United Kingdom **1**:79
United States **1**:76*T*, **1**:77–78
island biogeography
colonization **1**:617–618, **1**:618*F*
evolutionary development **1**:618–619
flightless birds **1**:605–606
fragmentation–extinction rate relationship **1**:619–620
human colonization impacts **1**:621–622
introduced species **1**:620
species–area relationships **1**:618*F*, **1**:619
late Quaternary extinction event (LQEE) **4**:721
latitudinal gradients **1**:498*F*
latitudinal range shifts **4**:604*T*
lineages-through-time plot **1**:432, **1**:432*F*
major threats **7**:203*F*
mangrove ecosystems **5**:13–14, **5**:20*T*
Mediterranean-type climate ecosystems
Australia **5**:213*T*, **5**:221
California **5**:213*T*
Cape Region (South Africa) **5**:213*T*, **5**:220
Chile **5**:213*T*, **5**:217
Mediterranean Basin **5**:213*T*, **5**:218–219
mercury contamination **1**:12*F*
metabolism **6**:404*F*
migration **1**:605, **1**:605*T*, **1**:614, **5**:311–313, **7**:35–36
mites **1**:214
modern extinctions **2**:168*F*, **2**:170*F*, **5**:368*T*
monogamy **6**:573
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*, **5**:472, **5**:473–475
New Zealand **5**:366
North America **5**:565
oil-palm plantations **5**:604–605, **5**:607
oil spills **3**:192–193, **3**:255
past, recent, and future extinction rates **2**:168*F*
pesticide toxicity **5**:701
phenological shifts
between-population variations **5**:720
between-species variations **5**:718–720
migration timing **5**:721
reproduction timing **5**:721
winter rest onset and termination timing **5**:721–722
pollination **6**:49*T*, **6**:141*T*, **6**:143
population dynamics models **4**:490*F*, **4**:491*F*
riparian zones **6**:463
risk status classification **1**:641–642, **1**:643*F*, **1**:644*F*
river ecosystems **6**:479–480
seagrass ecosystems **6**:547–548
seed dispersal **7**:283–284
Southeast Asia **1**:516
species abundance **1**:601–602, **1**:602*F*, **1**:603*F*, **6**:652*T*, **6**:653*T*
species–area relationships (SARs) **6**:659–660, **6**:660*F*, **6**:662*F*
species assemblages **6**:656*F*
species replacements
biogeographic regions **1**:616–617
causal factors **1**:615, **1**:616*T*
endemic bird areas (EBAs) **1**:615–616
evolutionary convergence **1**:617
ranges **1**:615, **1**:616*F*
species richness maps **6**:709*F*
subterranean ecosystems **7**:52–53
Succulent Karoo **1**:487–488, **1**:487*T*
taxonomic diversity **1**:606–608, **1**:607*T*
threatened species **7:180–190**
biotic threats
alien species **7**:184
climate change **7**:185–186
general discussion **7**:184
hunting, trapping, poisoning **7**:185
native–native threats **7**:184–185
small ranges **7**:185
causal factors
biotic threats **7**:184
general characteristics **7**:182–183
habitat loss and degradation **7**:183
conservation
critically endangered birds **7**:187–188
general discussion **7**:186
goals and priorities **7**:186
international cooperation and legislation **7**:187
priority management **7**:186
scale considerations **7**:186
site and habitat conservation **7**:186–187
critically endangered birds
alien and native restriction and control **7**:188
captive breeding programs **7**:188
challenges **7**:189
cross-fostering/cross-breeding **7**:189
habitat restoration **7**:187–188
nest site provision/enhancement **7**:189
reintroduction/translocation programs **7**:188
supplementary feeding programs **7**:188–189
current threats **1**:622–623, **1**:622*F*
global and regional trends
criteria **7**:180–181
extinct birds **7**:182
future outlook **7**:181
number of species **7**:181
regions and countries **7**:181

taxonomic groups **7**:181–182
United States/United Kingdom **7**:182
habitat loss and degradation
general discussion **7**:183
habitat degradation **7**:183–184
habitat destruction **7**:183
habitat fragmentation **7**:183–184
indicator species **7**:206–207*F*
Red List Index (RLI) **7**:201*T*
tourism industry **7**:223
tropical Africa **1**:49
tropical rain forests **6**:307–309, **6**:308*T*
United States **1**:76*T*, **1**:77–78
urbanization threats **4**:564
West Nile virus **1**:362–365, **1**:364*F*, **1**:365*F*
wetland ecosystems **7**:389–391, **7**:390*F*
wildfire effects **3**:156–158
zoonotic disease **1**:519–520
Bird's Point–New Madrid Floodway (Missouri) **7**:148*F*, **7**:149*F*, **7**:154–155
birdwing butterflies **3**:223
Birge, E. A. **3**:238–239
Birgus latro **3**:220, **3**:222–223, **3**:223*F*, **3**:226
biribá **4**:284*F*, **4**:285, **4**:287–290
Birnaviridae **5**:134*T*, **5**:137
birth and death rate analysis **3**:67–68
birth chains **5**:192
Bischoffia javanica **1**:277
Biserrula pelecinus **1**:708
Bismarck Islands **2**:189–190, **2**:189*F*
bison
animal breeding programs **1**:640
biodiversity effects **3**:188
grassland ecosystems **6**:332
Great American Biotic Interchange (GABI) **2**:584*F*
historical conservation movement **2**:281–282
late Quaternary extinction event (LQEE) **4**:722, **4**:724–725*T*
Bison bison **1**:640, **2**:281–282, **3**:188, **4**:3–4, **6**:332
Bison bonasus **3**:196
Biston betularia **2**:717, **6**:3
Bithynia tentaculata **4**:362
biting midges **3**:489–491*T*, **3**:492
Bitis arietans **6**:409
Bitis gabonica **6**:409
bitterns **7**:390
bitter vetch **2**:658
Bivalvia/bivalves
aquaculture **1**:190*T*
average lifespan **4**:656*T*
chemosynthetic symbioses **5**:120–121, **5**:120*F*
classifications **4**:381*T*
critical pH levels **6**:153*F*
fossil record **3**:681*F*, **3**:685*F*
hydrothermal vent ecosystems **4**:391, **7**:328–329
lake and pond ecosystems **4**:371–373
Late Ordovician mass extinction **5**:167
Late Triassic mass extinction **5**:172–173
mangrove ecosystems **4**:337, **5**:17
marine sediments **5**:106*F*
number of species **3**:182*T*, **4**:371*T*, **5**:669–670*T*
photosynthetic symbioses **5**:118–120, **5**:119*F*
physical characteristics **5**:376–377, **5**:377*F*
sandy beaches **2**:131–133
spatiotemporal distributions **5**:373–374, **5**:374*F*
species diversity **4**:386–387, **4**:387*F*
Bivesiculoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Biwa (Lake) **3**:171–173*T*, **4**:463*T*, **4**:464, **4**:464*T*
Bixa orellana **4**:289*T*
black abalone **3**:201*T*, **4**:391
black-and-white colobus monkeys **7**:229*T*
black-and-white ruffed lemurs **6**:256–258*T*
black-and-white warblers **4**:33–35, **4**:34*F*
black bass **4**:465
black-bearded saki **6**:256–258*T*
black bears **1**:631–632, **4**:518, **6**:50
Black Beauty (Sewell) **4**:653
black-bellied salamander **1**:173*F*
black-bellied seedcracker **3**:144–145
blackberry plants **1**:80
blackbilled magpies **4**:174
blackbirds **5**:474
blackbuck **7**:229*T*
blackburnian warbler **2**:188–189, **2**:188*F*
Blackburnian warblers **5**:309–310
black caiman **7**:229*T*
black-capped vireo **7**:184–185
black carbon **4**:295–296
black cherry **6**:121
blackchin tilapia **3**:481
black corals **3**:199
black-crowned Central American squirrel monkey **6**:256–258*T*
black eagles **5**:474
black-faced black spider monkey **6**:256–258*T*
black-faced lion tamarins **6**:256–258*T*
black flies **3**:489–491*T*, **4**:371*T*
black-footed ferrets **1**:664*T*, **2**:527–529, **3**:196, **3**:411, **6**:342
black francolin **1**:82
black gram **2**:663
black grama **4**:484–485, **4**:485*F*
black grouse **1**:631–632
black harrier **1**:487–488
black-headed ducks **5**:502, **5**:502*T*
black kites **5**:474
blacklegged ticks **1**:361–362, **1**:363*F*, **4**:505
blacklighting **1**:330
black lion tamarins **1**:666*T*, **6**:256–258*T*
black mold **3**:635
black mushrooms **3**:637*T*
black oak **6**:152–154
black pepper **2**:663, **3**:32*F*
black petrel **7**:188
Blackpoll warblers **5**:311–312
black rats **1**:365–368, **6**:240–241
black rhinos **1**:487–488, **4**:695, **5**:225*T*, **5**:226, **5**:449, **6**:578
see also megaherbivores
black robin **4**:194*T*, **5**:369, **7**:189
black scavenger flies **3**:489–491*T*
Black Sea bottlenose dolphin **5**:81*T*
black-shanked douc **6**:256–258*T*
black smokers
Endeavor Hydrothermal Vents Marine Protected Area **5**:447
marine invertebrates **2**:6*F*, **4**:391
occurrences and characteristics **7**:327–328
blacksnakes **6**:408–409
black-spined sea urchins **2**:337, **2**:338–340
black spruce ecosystems **1**:632*T*, **1**:633*F*, **7**:400
black stem rust **3**:633
black swans **3**:365–366, **5**:221
black-tailed godwits **5**:723
black-throated green warbler **2**:188–189, **2**:188*F*
black vulture **5**:603
blackwater streams **1**:512–513
bladder grasshoppers **3**:733, **3**:733*F*
Blakea **3**:522*T*
Blancocamelus **4**:724–725*T*
Blaniulus guttulatus **5**:424–425
Blastodinium **6**:295
Blastomyces dermatitidis **3**:634*T*
blastomycosis **3**:634*T*
Blastophaga psenes **6**:144–145
Blattodea **4**:308–312, **4**:309*T*
blennies **3**:467–471
Blephariceridae **3**:489–491*T*
Blephariceromorpha **3**:489–491*T*
blights **3**:633*T*
blindness **7**:302
Blissidae **7**:294
Blissus leucopterus **7**:293
blond capuchin **6**:256–258*T*
blond-masked titi **6**:256–258*T*
blood clotting **6**:3–5
blood pythons **5**:604
bloodroot **6**:50
blotched tiger salamanders **4**:515
blowflies **3**:489–491*T*, **3**:493
bluebells **6**:106*F*, **6**:107
blue carbon **6**:77
blue-cheeked bee-eaters **5**:473
blue cheese **3**:638–639
blue corals **2**:335, **3**:199
blue-eyed black lemur **6**:256–258*T*
bluefin tuna **2**:7–8, **2**:9*F*, **2**:10*F*, **3**:470
bluefish **3**:467–471
bluegill **4**:465
blue-green algae
characteristics **6**:25–26
energy flow **3**:246–247
freshwater ecosystems **3**:571–572, **3**:571*T*
marine ecosystems **5**:48
river ecosystems **6**:477–478
bluehead wrasse **6**:574
blue mussels **3**:355–356, **5**:35–36
blue orchard bee **1**:420*F*
Blue Ridge Mountains **5**:562–563
blue tits **1**:611–612, **1**:620, **5**:720
blue-tongued skinks **7**:32
blue whales **1**:601, **3**:190, **5**:65*T*, **5**:66–67, **5**:81*T*, **5**:82
blue whiting **4**:516, **6**:422*T*

blurred vision **3**:635*T*
boarfish **3**:467–471
Boas, Franz **3**:323
boat collisions **3**:192
bobcats **2**:584*F*
bobolinks **2**:67*F*, **2**:93*F*
bobtail squid **2**:323–324
Bodo **6**:293
bodonids **6**:293–295, **6**:294*F*
Boechera stricta **2**:636*F*
Boeckella **6**:26
Boeckella poppei **1**:183
Boerhavia **2**:376*F*
Boerhavia spicata **2**:373*T*
bofedales **4**:105
Bogidiella hebraea **5**:470–471, **5**:470*T*
Bogor Botanical Gardens **5**:448–449, **6**:508, **6**:510–511
bogs
  acidic environments **7**:395–397, **7**:396*F*
  Asia **1**:277–278
  definitions and terminology **7**:369*T*
  habitat–bird species diversity relationship **1**:613–614
  high-elevation Andean ecosystems **4**:105
  nutrient availability **7**:391, **7**:392*F*
  plant species richness **7**:388–389
  southern Andean steppes **4**:107
bogue **2**:356
Bohtrops alternata **3**:115
Boiga irregularis
  dispersal biogeography **2**:543, **2**:558–559
  diversity loss **1**:621, **4**:663
  invasive species **3**:218, **3**:410–411, **3**:410*F*, **4**:143, **4**:357, **7**:184
Bois-des-Bel Peatland, Quebec, Canada (case study) **7**:375
Bolax gummifera **4**:106–107
boletes **3**:629–630
Bolidophyceae **1**:476*F*
Bolivia
  botanical gardens **6**:519
  ecosystem services **7**:356*T*
  indigenous peoples **4**:270
  land cover classifications **6**:607*T*
  neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
  rain forests **1**:159*T*
  tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*, **7**:269–270, **7**:270*T*
  vascular epiphytes **3**:520*T*
  water funds **7**:356*T*
  wheat production **1**:87–88
Bolivina **6**:291–292
Bombacaceae **1**:263, **5**:11*T*, **6**:316*T*
Bombax munguba **1**:163
Bombina bombina **7**:47
Bombina variegata **7**:47
Bombix mori **6**:357–358, **6**:361*T*
Bombus affinis **2**:532
Bombus occidentalis **6**:144–145
Bombus terrestris **6**:144–145
Bombus vonesenskii **1**:420*F*
Bombycillidae **6**:308*T*
Bombycoidea **5**:386*F*, **5**:387–388*T*
Bombyliidae **3**:489–491*T*, **3**:492
Bombyx mori **5**:392, **5**:398–400
Bonasa umbellus **4**:62
Bonn Convention **2**:286, **7**:187
Bonnelli's eagle **5**:474
bonobos **1**:49, **4**:693, **6**:256–258*T*
bony fish
  bichirs **3**:461–462
  bowfin **3**:462–463
  characteristics **3**:456–457, **3**:457*F*, **6**:522–523
  classifications **3**:460–461
  coelacanths
    characteristics **3**:461, **7**:338
    hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
    speciation **2**:47
  gars **3**:462, **3**:462*F*
  lungfish
    characteristics **3**:461, **3**:461*F*, **7**:338
    hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
  major threats **7**:203*F*
  paddlefish **3**:462
  reedfish **3**:461–462
  species diversity **3**:684–685
  sturgeon **3**:462, **3**:462*F*
  teleosts
    Acanthopterygii
      Atherinomorpha **3**:467
      Mugilomorpha **3**:467
      Percomorpha **3**:467–471
    characteristics **3**:463, **6**:522–523
    clupeomorphs **3**:463–464
    elopomorphs **3**:463
    Euteleostei
      Ateleopodomorpha **3**:465
      Cyclosquamata **3**:465–466
      Lampriomorpha **3**:466, **3**:466*F*
      Paracanthopterygii **3**:466–467
      Polymixiomorpha **3**:466
      Protacanthopterygii **3**:464–465, **3**:465*F*
      Scopelomorpha **3**:466
      Stenopterygii **3**:465
    ostariophysans **3**:463–464
    osteoglossimorphs **3**:463, **3**:463*F*
  tetrapods
    characteristics **3**:461, **7**:338–339
    classification systems **7**:337*F*
  threatened species **7**:201*T*
bonytail chub **3**:260
boodie rats **4**:363–364
boojum trees **2**:471
Boomthang Valley **1**:276*T*
Boone and Crockett Club **2**:281
Boops boops **2**:356
Boops salpa **6**:547–548
Bootherium **4**:724–725*T*
bootstrap estimator **5**:184–185
bootstrap support **2**:44, **6**:19
Boraginaceae **1**:277, **2**:374*T*
Borassus aethiopium **1**:68
boreal Asia
  boreal forest **1**:262–263*T*, **1**:273–274
  characteristics **1**:273–274
  tundra **1**:262–263*T*, **1**:274–275
boreal forests **1:626–635**
  animal species **1**:631–632, **1**:633*F*
  Asia **1**:273–274
  biodiversity change scenarios **4**:6*F*
  Canadian Shield **5**:562
  climate
    characteristics **1**:626
    precipitation **1**:626–627, **1**:628*F*
    seasonal boundaries **1**:626
    temperatures **1**:626, **1**:628*F*
  climate change studies **2**:61*F*
  conservation
    clear-cutting threats **1**:634
    natural reserves and preserves **1**:634
    sustainable management **1**:634
  disturbance dynamics **7**:214–215, **7**:214*T*
  ecosystem function measurements **3**:77*F*
  Europe **3**:335*T*, **3**:336–340*T*
  Finland **7**:212–213, **7**:212*T*
  geography **1**:626, **1**:627*F*
  global agriculture impacts **3**:428*T*
  global biodiversity trends **4**:563*F*
  global distribution patterns **7**:117*T*
  litter deposition **6**:120*T*
  management strategies **7**:213, **7**:213*F*, **7**:214*T*
  precipitation–temperature relationship **1**:59*F*
  predation effects **6**:236–237, **6**:238*F*
  protected area systems **6**:492*T*, **6**:499*T*
  soils **1**:627, **6**:123*T*
  structure and function **7**:213, **7**:213*F*
  succession **1**:632–634
  successional dynamics **7**:213–214, **7**:214*F*
  timber-oriented management
    beetle species **7**:218–219, **7**:219*T*, **7**:220*F*, **7**:220*T*
    forest characteristics **7**:216–217, **7**:217*F*
    impact mechanisms **7**:215–216, **7**:217*F*, **7**:217*T*
    landscape patterns **7**:217–218, **7**:217*F*, **7**:218*F*, **7**:219*F*
    production optimization **7**:215, **7**:216*F*
    research summary **7**:219–220
  vegetation
    characteristics **1**:628–630
    geographic boundaries **1**:629*F*, **1**:630*T*
    latitudinal classifications
      closed forests **1**:629–630
      forest–tundra ecotone **1**:630
      general discussion **1**:629–630
      lichen woodlands **1**:630
    longitudinal classifications **1**:630–631, **1**:631*T*, **1**:632*T*, **1**:633*F*
  wildfires **1**:627–628
boreal toads **4**:515
Boreellus atriceps **1**:236
Borhyaenidae **4**:683, **4**:701*T*
boric acid **5**:695–696
Bornean agile gibbon **6**:256–258*T*
Borneo **1**:262–263*T*, **1**:514*T*, **3**:156–158, **6**:82*T*
Borophagus **4**:724–725*T*
Borrelia burgdorferi **1**:361–362, **1**:363*F*, **5**:415
Borschov **2**:376*F*
Boserup, Ester **6**:206

Bos gaurus **1**:640, **5**:225–226
Bos grunniens **3**:191
Bos javanicus **1**:640
Bosmina **6**:26
Bos mutus **1**:640, **3**:373
Bosnia and Herzegovina **6**:519
Bos taurus **3**:127*T*
Bostrychoidea **1**:327*T*
Boswellia **1**:69–70
Boswellia papyrifera **1**:67
botanical gardens **6**:504–521
  biodiversity conservation **4**:319, **4**:321*T*
  biodiversity-rich countries **1**:504, **1**:505*T*
  Brazil **6**:507
  classifications **6**:505
  climate change studies
    botanical and ecological research **6**:515–516
    challenges **6**:515
    ex situ conservation programs **6**:516
    germplasm utilization **6**:516
    in situ conservation programs **6**:516
    public awareness and educational programs **6**:515
  conservation biology **2**:245
  defining characteristics **6**:504–505
  educational programs **6**:511
  geographic distribution **6**:505–506, **6**:506*T*, **6**:507*F*
  global renaissance **6**:504
  herberia **6**:510–511
  historical perspective **6**:506–508
  horticultural contributions **6**:508
  human health issues **6**:518–519
  informal education experiences **3**:138
  medicinal plants **6**:518–519
  networks
    Botanic Gardens Conservation International (BGCI) **6**:512
    International Association of Botanic Gardens (IABG) **6**:511–512
    national and regional networks **6**:512–513
  nonseed gene banks **3**:653
  plant collections
    ex situ collections **6**:510
    geographical collections **6**:510
    global distribution **6**:510*T*
    taxonomic collections **6**:510
    thematic collections **6**:509
  plant conservation **4**:328, **6**:63, **6**:86
    ex situ conservation programs **6**:513, **6**:513–514, **6**:514
    Global Strategy for Plant Conservation (GSPC) **6**:514–515, **6**:515
    International Agenda for Botanic Gardens in Conservation **6**:514
  research collections and libraries **6**:508–509, **6**:509
  seed banks **3**:642–643
  summary discussion **6**:519
  sustainability movement
    biodiversity conservation support **6**:517
    challenges **6**:516–517
    educational programs **6**:517–518
    environmentally friendly practices **6**:517
    sustainable development support **6**:517
    waste reduction and recycling programs **6**:518
  see also natural history museums and institutions
Botanic Gardens Conservation International (BGCI) **4**:325–326, **6**:78, **6**:512
Botanischer Garten der Universitat Wien **6**:506–508
Botanischer Garten und Botanisches Museum (Berlin) **5**:411*F*
bot flies **3**:489–491*T*, **3**:494
Bothriocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Bothriocephalus acheilognathi **4**:364, **5**:668
Bothrioplanida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443–444
Bothrioplanus **7**:442–443*F*
Bothriuridae **1**:209
Bothrops jararaca **1**:359
Botrychium **6**:59
Botswana
  botanical gardens **6**:519
  edible plants **3**:130*T*
  megaherbivores **5**:234–235
botterbooms **2**:471
bottlenose dolphins **3**:256
bottle-trees **2**:468, **2**:471
bottomland **7**:369*T*
boulder fields **4**:336–337, **4**:336*F*
Bouteloua aristidoides **2**:373*T*, **4**:446*F*
Bouteloua barbata **4**:446*F*
Bouteloua gracilis **6**:441
Bouvier's red colobus **6**:256–258*T*
Bovidae **4**:690*T*, **4**:696, **6**:309*T*
Bowdichia virgiloides **1**:161
Bower's shrike-thrush **2**:75*F*
bowfin **3**:462–463
bowhead whales **5**:77*T*, **5**:86–87, **5**:87*T*
Bowie Seamount Marine Protected Area (Sglaan Kinghlas) **5**:446–447
boxfish **3**:467–471
Brachiaria deflexa **1**:67
Brachionus **6**:26, **6**:31
Brachionus calyciflorus **2**:726–727
Brachiopoda
  Census of Marine Life project **2**:7*F*
  classifications and sampling methods **4**:381*T*
  fossil record **3**:681*F*, **3**:684*F*
  global diversity **4**:385*T*
  hydrothermal vent ecosystems **7**:329–330
  known and estimated species **4**:384*T*
  Late Devonian mass extinction **5**:169–170
  Late Ordovician mass extinction **5**:167
  Late Triassic mass extinction **5**:172–173
  marine ecosystems **5**:50*T*
  number of species **5**:669–670*T*
  Proterozoic–Cambrian transition **3**:396*F*
Brachycera **3**:489–491*T*, **3**:493
Brachyceratops **2**:519*T*
Brachycybe **5**:423*F*
Brachydiastematherium **5**:224*T*
Brachydiceratherium **5**:224*T*
Brachylaimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Brachylophosaurus **2**:519*T*, **2**:520*T*
Brachypelma smithi **3**:225–226
Brachypodium distachyon **2**:364*F*
Brachypotherium **5**:224*T*
Brachypotoma **4**:722
Brachyramphus marmoratus **3**:256, **3**:314–315, **4**:564
Brachys **1**:329
Brachystegia **1**:67–68, **7**:274–275
Brachystegia laurentii **7**:273
Brachyteles arachnoides **4**:692, **6**:256–258*T*
Brachyteles hypoxanthus **6**:256–258*T*
Brachytrupinae **3**:729–730
Brachyura **2**:416–417*T*
Bracken Cave (Texas, USA) **7**:60
brackish wetlands **1**:292–297
Braconidae **4**:178, **4**:181, **5**:395–396, **5**:674
Bradynectes **7**:442–443*F*
Bradypodidae **6**:309*T*
Bradyporinae **3**:726
Bradypus torquatus **2**:587*F*
Bradyrhizobium **6**:450, **6**:581–582
brain size evolution **4**:698
Branchinecta gaini **1**:183
Branchiopoda **2**:400–401, **2**:401*T*, **4**:371*T*, **6**:26
Branchiostoma **7**:336
Branchiura **2**:401*T*, **2**:416–417*T*
Branta canadensis **1**:79, **3**:365–366, **5**:312
Brasiella argentata **1**:250*T*, **1**:251*F*
Brasiella pretiosa **1**:250*T*, **1**:251*F*
Brassicaceae **2**:138, **3**:129*T*
Brassica chinensis **2**:663
Brassicales **2**:374*T*
Brassica napus **2**:389
Brassica oleracea **2**:383, **2**:662–663
Brassica pekinensis **2**:383, **2**:663
Brassica rapa **1**:543–547, **2**:662–663, **3**:131–132, **6**:361*T*
Brassica spp. **1**:120, **3**:32*F*
Brassolinae **1**:653
Braulidae **3**:489–491*T*
Bray–Curtis index **6**:659–660
Brazil
  biodiversity richness **1**:501*F*
  Biota-Fapesp **1**:506
  botanical gardens **6**:507, **6**:519
  butterflies **1**:658*T*
  Convention on Biological Diversity (1930) **2**:313
  coral reef ecosystems **1**:500*T*
  cropland conversion **4**:665*T*
  deforestation **1**:502*T*, **1**:677*F*, **2**:449*T*, **2**:451*T*, **7**:170–172
  ecosystem services **7**:356*T*
  ecotones **3**:143*F*
  endangered mammals **3**:194*T*
  endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
  forest product exports **2**:451*T*
  grasslands and savannas **6**:332*F*
  indigenous peoples **4**:270*T*, **4**:273*F*, **4**:273*T*
  institutional collections and capacity **1**:505*T*
  introduced species
    nonindigenous invertebrates

Brazil (*continued*)
arthropods **1**:76*T*, **1**:83
crop plant pathogens **1**:76*T*, **1**:83
nonindigenous plants **1**:76*T*, **1**:82–83
nonindigenous vertebrates
birds **1**:83
mammals **1**:76*T*, **1**:83
invasive species **4**:353–354
land cover classifications **6**:607*T*
maize production **1**:86
natural history museums and institutions **5**:406
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
oil-palm plantations **1**:540–541, **5**:602
plant hotspots **6**:61
primate diversity **6**:252*T*
primate endemism **6**:252*T*
rain forests **1**:159*T*
reforestation programs **6**:377–378
shifting cultivation systems **7**:232, **7**:232*F*
species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
temperate forests
distribution patterns **7**:117
diversity patterns **7**:119–121, **7**:119*T*
dominant floristic features **7**:119
endemism **7**:121
threatened birds **7**:181
tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*, **7**:269–270, **7**:270*T*
tropical rain forests **6**:314–315, **6**:317*T*
vascular plant diversity **6**:82*T*
water funds **7**:356*T*, **7**:363–364
see also Amazon ecosystems
Brazilian cottontail **2**:590*T*
Brazilian free-tailed bats **7**:53, **7**:262
Brazilian pepper **4**:357, **7**:399
Brazil nut tree **1**:160, **4**:287–290, **7**:273
bread baking **3**:638
breadfruit **3**:32*F*
bread wheat **2**:662–663
bream **4**:463
breeding of animals **3:100–107**
challenges **1**:648
conservation planning **1**:639–640
crossbred animals **1**:639
current trends **1**:638–639
diversity challenges **1**:636
future outlook **1**:648
genetic structure **1**:637–638
global breeds **1**:638–639
livestock breeds **1**:637
management strategies
conservation strategies
ex situ strategies **1**:645–646
general discussion **1**:642–645
in situ strategies **1**:642–645
Food and Agriculture Organization (FAO) programs **1**:640–642, **1**:641
inventory, description, and characterization assessments **1**:641–642, **1**:642*F*
risk status classification **1**:641–642, **1**:642*F*, **1**:642*T*, **1**:643*F*, **1**:644*F*
new technologies
cloning **1**:647–648
cryopreservation **1**:646
gene mapping **1**:646–647
molecular characterization **1**:646–647
wild relative use **1**:640
Bremer support **2**:44
Brenthidae **1**:327*T*
Bretschneideraceae **1**:260*T*
Bretzia **4**:724–725*T*
brevetoxin A **6**:131–132, **6**:134*F*
Breviceps macrops **1**:487–488
Brevoortia tyrannus **3**:363–364
Brewer, William H. **2**:281
brewing **3**:638
Brexia **5**:623*F*
Bridgeoporus nobilissimus **3**:639–640, **3**:639*F*
bridled honeyeater **2**:75*F*
brie cheese **3**:638–639
brine shrimp **6**:26
bristlemouths **3**:465
bristletails **4**:392*T*
bristleworms **4**:371*T*
British Indian Ocean Territory
animal species **1**:261*T*
botanical gardens **6**:507*T*
plant species **1**:259*T*
threatened species **1**:278*T*
British Museum **5**:405
British soldier lichen **3**:637*F*
brittle greenhood **3**:211, **3**:213*F*
brittle stars **4**:381*T*
broad beans **1**:590
broad fish tapeworm **5**:660*T*
broadleaf forests
current global land cover **4**:557*F*
global biodiversity trends **4**:563*F*
net primary production (NPP) **1**:560*T*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
protected area systems **6**:492*T*, **6**:499*T*
tropical broadleaf forests **3**:428*T*, **6**:600–604*T*
Brocchinia **6**:44*T*
brocket deer **2**:590*T*
Brockman Family at Beachborough, The (Haytley) **2**:347, **2**:347*F*
broken-stick model **1**:402
bromacil **5**:696
Bromeliaceae **6**:44*T*, **6**:314, **6**:316*T*
Bromfield, Louis **1**:112
bromine (Br)
methyl bromide (CH_3Br) **1**:282*T*, **1**:284–286
ozone depletion **1**:289
ultraviolet radiation **7**:297
bromoxynil **4**:90*T*
Bromus inermis **1**:269, **6**:456–457, **6**:457–458
Bromus tectorum **1**:703, **2**:566, **3**:703, **4**:347–348, **4**:348*F*, **4**:360, **6**:93, **6**:125
Bronchocela smaragdina **1**:515–516
Brongniart, Adolphe **1**:578–579
Brontotheriidae **5**:224*T*
Brontotherium **5**:224*T*
Bronx Zoo **5**:405–406
brood parasitism **5:501–509**
adaptive behaviors **5**:501
background information **5**:501
conspecific brood parasitism (CBP)
definition **5**:501
prevalence **5**:508–509
species survey **5**:502–503
cowbird management **5**:507
fish **5**:503
host–parasite coevolution
cowbird–host systems
behavioral adaptations **5**:504–505
egg adaptations **5**:504
host defenses **5**:505
nestling and fledgling adaptations **5**:505
cuckoo–host systems
adaptive behaviors **5**:503
egg adaptations **5**:503
evolutionary development **5**:503
fecundity **5**:503
host defenses **5**:504
nestling and fledgling adaptations **5**:503–504
modeling approaches **5**:506
host population dynamics **5**:506–507
insects **5**:503
major groups **5**:502*T*
number of species **5**:502*T*
research outlook **5**:509
social behavior
foraging ecology **5**:507
mating systems **5**:507–508
spacing behaviors **5**:508
vocal behaviors **5**:508
species survey
black-headed ducks **5**:502, **5**:502*T*
conspecific brood parasitism (CBP) **5**:502–503
cuckoo-finches **5**:502, **5**:502*T*
Cuculidae **5**:501–502, **5**:502*T*
Icteridae (cowbirds) **5**:502, **5**:502*T*
Indicatoridae **5**:502, **5**:502*T*
viduine finches **5**:502, **5**:502*T*
brook trout **1**:10*T*, **1**:12*F*, **3**:192–193, **6**:153*F*
broomcorn millet **2**:663
broomrapes **2**:469, **4**:76
Brosimum alicastrum **7**:273
Brosimum potabile **4**:287–290
brotulas **3**:466–467
brown adipose tissue **3**:32, **7**:26–27
brown algae
evolutionary development **2**:642
freshwater ecosystems **3**:571–572, **3**:571*T*
grazing effects **4**:12
marine ecosystems **5**:48
modern extinctions **5**:368*T*
brown bats **2**:527–529, **2**:529*F*
brown bears **3**:190, **7**:34*F*
brown bullhead **1**:12*F*
brown lemur **6**:251
brown pelicans **3**:256
brown seaweed **2**:6*F*, **4**:12
brown spider monkey **6**:256–258*T*
brown tree snakes

dispersal biogeography **2**:543, **2**:558–559
diversity loss **1**:621, **4**:663
invasive species **3**:218, **3**:410–411, **3**:410*F*, **4**:143, **4**:357, **7**:184
brown trout **1**:12*F*, **3**:181
Brucella abortus **5**:449
brucellosis **3**:191
Bruchidae **1**:327*T*
Bruegel, Pieter, the elder **2**:347, **2**:347*F*
Brugia malayi **7**:428*T*
Brugia timori **7**:428*T*
Bruguiera **1**:277, **5**:11*T*
Bruguiera gymnorrhiza **5**:21
Bruguiera parviflora **5**:21
Brundtland Report **4**:326, **7**:13–14
Brunei Darussalam
animal species **1**:261*T*
botanical gardens **6**:507*T*
ecosystems
mangroves **1**:275*T*, **1**:513*T*
threatened species **1**:278*T*
tropical moist forest **1**:277*T*
vegetation types and coverage percentage **1**:264–265*T*
energy efficiency measures **3**:254*F*
plant species **1**:259*T*
see also Southeast Asia
Brunellia **7**:274–275
Brunellia comocladifolia **1**:166–167
Brunellia occidentalis **1**:167
brushtail possum **5**:370
Bryant, William Cullen **2**:280
Bryde's whales **5**:88–89
bryophytes
anthropogenic impacts **7**:399
Arctic ecosystems **1**:231*T*
boreal forests **1**:628–630
characteristics **6**:58–59, **6**:58*T*
conservation practices **6**:83
elevational range shifts **4**:605*T*
forest canopies
characteristics **3**:517–518
taxonomic distributions **3**:520, **3**:521*T*
fossil record **3**:686–687, **3**:687*F*
geographic distribution **6**:58–59
historical perspective **4**:168–169
indicator species **7**:206–207*F*
nonseed gene banks **3**:651
number of species **6**:58*T*, **6**:586*T*
threatened species **7**:206–207*F*
wetland ecosystems **7**:388–389
Bryozoa
Census of Marine Life project **2**:7*F*
classifications and sampling methods **4**:381*T*
convergent evolution **3**:615*F*
coral reef ecosystems **2**:332*T*
fossil record **3**:681*F*, **3**:684*F*
hydrothermal vent ecosystems **7**:329–330
lake and pond ecosystems **4**:374*F*
mangrove ecosystems **5**:20*T*
marine ecosystems **4**:382*F*, **5**:50*T*
number of species **4**:371*T*, **5**:669–670*T*
Bubalis bubalis arnee **1**:80
Bubalus bubalis **4**:354, **4**:364–365, **6**:333
bubonic plague **2**:551–552*T*
Bubulcus ibis **1**:710, **2**:551–552*T*, **4**:357, **5**:13–14
Bucconidae **6**:308*T*
Bucephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Bucephalopsis **7**:450–451*F*
Buchenavia macrophylla **1**:163
Buchloe dactyloides **1**:708
Buchnera aphidicola **1**:312
buckwheat **2**:663
Buddhism
biodiversity conservation **6**:564
Buddhist declaration on nature **3**:272, **6**:383
environmental influences **6**:384
buffalo **1**:644*F*, **4**:362, **5**:14
buffalo grass **1**:708
buffel grass **6**:93
buff ermine moth **6**:159*F*
Buffon, Georges Louis Leclerc, comte de **1**:577, **4**:651–652
Buffon's law **1**:577
buffy-headed marmoset **6**:256–258*T*
Bufo baxteri **3**:696
Bufo boreas **3**:165, **4**:515
Bufo bufo **3**:165
Bufo calamita **3**:166
Bufo marinus **1**:78, **1**:80, **1**:175, **2**:74, **3**:167, **3**:695–696
Bufonia ramonensis **5**:485
Bufo periglenes **2**:76, **2**:86–87, **3**:165, **3**:167–168, **3**:697, **4**:664
Bufo refus **3**:115
bufotenin **5**:711
Bufo viridis **5**:473
Bugenasaura **2**:519*T*, **2**:520*T*
bugs *see* true bugs
Bugula neritina **5**:711
buildings/built-up areas **4**:579, **4**:580*F*
Bulbitermes **4**:409*T*
Bulbocodium **1**:269
Bulbophyllum **6**:316*T*
Bulbostylis **2**:376*F*
bulbuls **5**:474
Bulgaria
botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
Buliminus labrosus **3**:674*F*, **5**:469–470
bulldozer millipedes **5**:429–430
bullfrogs **1**:78, **2**:529–530, **2**:530*F*, **3**:695–696, **6**:457, **7**:399
bullhead fish **3**:365*F*
bumble bees **1**:420*F*, **2**:532, **4**:183, **6**:144–145
bunchgrass **2**:469
Bunodosoma caissarum **5**:425
Bunopithecus **4**:692–693
Bunyaviridae **5**:130*F*, **5**:134*T*, **5**:139–140
Buphagus erythrorhynchus **5**:227
bupirimate **5**:695
Buprestidae **1**:327*T*, **1**:328
Buprestoidea **1**:327*T*
burdock plants **1**:590
buritizal **1**:164
Burkea africana **1**:68
Burkholderia **5**:293–294
Burkholderia pseudimallei **2**:636*F*
Burkina Faso
botanical gardens **6**:519
energy efficiency measures **3**:254*F*
Burma *see* Myanmar
Burman monsoon forest **1**:262–263*T*
Burmanniaceae **6**:315*T*
Burman rain forest **1**:262–263*T*
Burnupena **5**:57, **5**:60*F*
Burramyidae **4**:688–689
burrfish **3**:467–471
burros **1**:76–77
burrowing myriapods **5**:430
Bursera **3**:522*T*, **5**:489
Burseraceae **3**:522*T*
Burundi
botanical gardens **6**:519
deforestation **2**:449*T*
energy efficiency measures **3**:254*F*
burying beetles **3**:220–221, **3**:226
bushbucks **1**:50
bushcrickets **5**:460*T*, **5**:464*T*
bushes **7**:266*F*
bush-fallow rotation systems **1**:114–115
bush hyrax **4**:695
bush meat **6**:499
bush pigs **1**:55–56
bush wrens **1**:621
Buteo **1**:709
Buteo ferox **5**:474
Buteo jamaicensis **4**:516
Buthacus arenicola **5**:468–469
Buthacus yotvatensis **5**:468–469
Buthidae **1**:209
butterfish **3**:467–471
butterflies **1:650–661**
biodiversity indicators **4**:261
biodiversity restoration **6**:458–459
butterfly–ant symbiotic relationship
acoustic communication **1**:656
association categories **1**:656
food secretions **1**:655
myrmecophily **1**:655
semiochemical production **1**:655–656
butterfly caterpillars **1**:652*F*
characteristics **4**:392, **4**:393*F*
coevolution theory **2**:138–139
desert ecosystems **2**:473
domesticated ruminants **1**:709–710
elevational range shifts **4**:605*T*
El Niño Southern Oscillation (ENSO) events **3**:156–158
endangered species **3**:226
environmental change indicators **4**:257–258, **4**:259*F*
freshwater ecosystems **3**:579–580, **3**:579*T*
geographic diversity
latitudinal gradients **1**:657, **1**:657*T*, **1**:658*F*
neotropics **1**:657–658, **1**:658*T*, **1**:659*T*
regional patterns **1**:658, **1**:659*T*
spatiotemporal distributions **1**:658–659, **1**:659*F*, **1**:660*F*
species numbers **1**:657, **1**:657*T*
habitat–bird species diversity relationship **1**:612, **1**:613*F*

butterflies (*continued*)
habitat destruction **1**:659–660, **4**:516–517
hot spots **4**:261
introduced species **4**:363
latitudinal range shifts **4**:604*T*
life cycle
adult stage **1**:651*F*, **1**:654–655
caterpillars **1**:652*F*, **1**:654
egg stage **1**:654
general discussion **1**:654
host relationships **1**:655
pupa stage **1**:654
migratory patterns **5**:316
mimicry
basic concepts **1**:656
Batesian mimicry **1**:656
diversity considerations **1**:656–657
Müllerian mimicry **1**:656
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*
North America **5**:565
number of species **4**:371*T*, **4**:392*T*
overharvesting/overcollecting **3**:223
pesticides **3**:222
phylogenetic relationships **5**:384–385
pollination **6**:49*T*, **6**:141–142, **6**:141*T*
Southeast Asia **1**:514–515
taxonomic diversity **1**:650–654, **1**:651*F*
threatened species **3**:579–580, **3**:579*T*
tropical rain forests **6**:306–307, **6**:307*T*
urbanization threats **4**:564
wildfire effects **3**:156–158
butterfly fish **3**:467–471
Butyrospermum pyrodoxum **1**:68
Buxaceae **5**:621*F*
buzzards **1**:709, **5**:474
Byblidaceae **6**:44*T*
Byblis **6**:44*T*
bycatch **3**:446, **3**:446, **3**:447, **4**:156
Byrsonima crassifolia **4**:286–287, **4**:288*T*, **7**:274
Byrsonima verbascifolia **1**:165
Bythotrephes **6**:26

C

C_3 plants
anatomical modifications **2**:364*F*
energy requirements **2**:369*T*
grasses **1**:58–59, **1**:64*F*, **2**:654–655
historical atmospheric conditions **2**:367*T*
photosynthetic carbon reduction cycle **2**:361–362, **2**:362*F*
quantum yield **2**:369*F*
temperature effects **2**:369*F*
C_4 plants **2:361–381**
animal diversity impacts **2**:377–378
atmospheric carbon dioxide (CO_2) concentrations **2**:361
biogeography
C_4 annuals **2**:373*T*
general characteristics **2**:370–372
global biomes **2**:371*T*
latitudinal gradients **2**:370–372, **2**:370*F*
light intensity **2**:370
moisture availability **2**:372
principle life forms **2**:371*T*
salinity/salt tolerance **2**:371*T*, **2**:372–373
temperature parameters **2**:370–372
climate change effects **2**:378
ecological dominance **2**:377
economically significant C_4 plants **2**:362*T*
energy requirements **2**:369*T*
evolutionary development
carbon starvation hypothesis **2**:374–377
molecular clock analysis **2**:373–374, **2**:376*F*
polyphyletic origins
evolutionary centers **2**:375*F*
evolutionary lineages **2**:369*F*, **2**:373, **2**:374*T*, **2**:375*F*, **2**:376*F*
molecular clock analysis **2**:373–374, **2**:376*F*
taxonomic diversity **2**:373
exotic C_4 plants **2**:379–381, **2**:380*F*
grasses **1**:59*F*, **1**:64*F*, **2**:378–379, **2**:379*F*, **2**:654–655, **2**:683
historical atmospheric conditions **2**:367*T*
land-use change impacts **2**:378–379, **2**:379*F*
photosynthetic mechanisms
anatomical modifications **2**:364, **2**:364*F*
basic concepts **2**:362–364, **2**:363*F*
biochemistry **2**:363–364
C_3 photosynthetic carbon reduction cycle **2**:361–362, **2**:362*F*
diffusion barriers **2**:365–366
photorespiration **2**:361–362, **2**:362*F*
requirements **2**:363*T*
schematic diagram **2**:363*F*
single-cell C_4 photosynthesis **2**:364–365, **2**:365*F*
physiological significance
atmospheric carbon dioxide (CO_2) effects **2**:366–368, **2**:366*F*, **2**:368*F*, **2**:369*F*
energy requirements **2**:369*T*
historical atmospheric conditions **2**:367*T*
nitrogen-use efficiency **2**:367*T*, **2**:368–370
photorespiration **2**:366–368, **2**:366*F*
quantum yield **2**:369*F*
temperature effects **2**:367, **2**:369*F*
water-use efficiency **2**:367*T*, **2**:368–370
quantum yield **2**:369*F*
research summary **2**:381
temperature effects **2**:369*F*
tropical deforestation effects **2**:379–381, **2**:380*F*
tropical savannas **1**:58–59, **1**:59*F*, **1**:64*F*
vegetation–climate interactions **2**:683
Caatinga ecosystems **5**:496, **5**:497*F*
caatinga forests **3**:79–80, **3**:81*T*
cabbage
intercropping patterns **2**:383
native crops **2**:662–663
cabbage white butterfly **2**:551–552*T*
CABI Bioscience UK Centre **3**:652*T*
Cabomba **5**:618–619
cacao **2**:663–664, **3**:12, **4**:287–290, **7**:26
cacauí **4**:287–290
Cactaceae **3**:129*T*, **3**:130*T*, **4**:350–351, **5**:489, **6**:316*T*
cacti **3**:409*T*
Cactoblastis cactorum **4**:351, **4**:361, **5**:401–402, **6**:171
Cactospiza heliobates **1**:27–28
Cactospiza pallida **1**:27–28
cactus **2**:468–469
cactus finches **1**:27–28, **1**:28*F*
cactus flies **3**:489–491*T*
cactus moth **4**:361
caddisflies
cladistic biogeography **7**:345, **7**:346*F*
freshwater ecosystems
acidification effects **3**:584–585
organic pollution **3**:580–581, **3**:581*F*
threatened species **3**:579–580, **3**:579*T*
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
phylogenetic relationships **5**:385
wetland ecosystems **7**:389–391
Caddoidea **1**:210
Caecidotea kenki **7**:59
caecilians
biodiversity
geographic distribution **3**:691–693
historical biogeography **1**:169–170
characteristics **3**:691–693, **7**:339
functional anatomy **1**:170–171
larval morphology and ecology **1**:173
life histories **1**:171
morphology **1**:170–171
postlarval behavior and ecology **1**:175–176
reproductive biology **1**:171
special features **1**:170–171
species richness **1**:170*T*
Caecomobius varius **7**:60
Caenagnathus **2**:519*T*
Caenis macrura **5**:471
Caenolestidae **4**:687
Caenomorpha **6**:31
Caenorhabditis elegans **2**:628*T*, **6**:355
Caesalpiniaceae **3**:129*T*
Caeselapinia esplinata **6**:560
Cafeteria **6**:293
caffeine **2**:426*T*, **2**:427, **5**:712*T*
cahow **2**:279
caiaué **4**:285
Caiman latirostris **3**:115
caimans **3**:115
Cain, Arthur **7**:102–104
Cain, Stanley **2**:539–540
Cajanus cajan **2**:383
caju **4**:286–287
calabash **4**:281–285, **4**:283*F*
Caladiopsis **3**:522*T*
Calamagrostis purpurea **1**:634
Calamus **1**:263
Calandrinia **5**:217
Calandrinia acaulis **4**:107
Calandrinia affinis **4**:107
Calandrinia caespitos **4**:107

calanoids **6**:25
calanolides **1**:359
Calanus **4**:674
Calanus finmarchicus **5**:686–688, **5**:687*F*, **5**:688*F*
Calanus glacialis **4**:386*T*, **5**:690
Calanus helgolandicus **5**:686–688, **5**:687*F*, **5**:688*F*
Calathea allouia **4**:289*T*
Calcaritermes **4**:409*T*
Calceolaria **5**:217
Calceolaria uniflora **4**:106–107
Calcisols **2**:461–464
calcium (Ca)
  calcium carbonate ($CaCO_3$) **1**:554, **1**:555*T*, **1**:556, **6**:147–149
  elemental partitioning **1**:554*T*
  forest ecosystems **1**:5, **1**:6*F*
  relative abundance **1**:553*F*
  river water composition **3**:562*T*
  root–soil interface **6**:114*F*
calden **1**:708–709
Caldenia **5**:221
Caliciviridae **5**:134*T*, **5**:138
Calidris alba **5**:312
Calidris canutus **7**:185
Calidris fuscicollis **5**:312
Calidris pusilla **7**:32
Calidris semipalmatus **7**:185
California
  C_4 plants **2**:371*T*
  California Floristic Province **5**:215, **6**:82–83, **6**:82*T*
  flora **5**:481*T*
  Mediterranean-type climate ecosystems
    characteristics **5**:212–213
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:214–215
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:215
    vertebrate diversity **5**:213*T*, **5**:215
  plant diversity **3**:700–701
  population diversity **6**:172–173
  river–floodplain connectivity **7**:151*F*, **7**:155, **7**:156*F*, **7**:157*F*
California Academy of Sciences **5**:405–406, **5**:408*F*, **5**:409*F*
California black oak **6**:152–154
California condor **1**:621–622, **1**:664*T*, **2**:273, **5**:215, **6**:445, **7**:185, **7**:236–237
California mussels **4**:45–46, **6**:232
California sea lions **3**:192–193
California sea otter **5**:81*T*, **5**:83–84
California spiny lobster **4**:516
California thrashers **4**:40–42, **4**:41*T*, **4**:64
Calinaginae **1**:653
Callaeas cinerea **5**:370
Callicarpa japonica **1**:271
Callicebus barbarabrownae **6**:256–258*T*
Callicebus coimbrai **6**:256–258*T*
Callicebus modestus **6**:256–258*T*
Callicebus oenanthe **6**:256–258*T*
Callicebus olallae **6**:256–258*T*
Callidium aeneum **7**:220*T*
Calliduloidea **5**:387–388*T*
Calligonum **1**:268, **1**:270, **2**:371*T*, **2**:376*F*
Callimorpha jacobea **6**:159*F*
Callipepla gambelii **2**:475*F*
Calliphoridae **3**:489–491*T*, **3**:493
Callistophyton **5**:618*F*
Callithrix flaviceps **6**:256–258*T*
Callithrix kuhlii **6**:560
Callitrichidae **6**:250–252, **6**:251*T*
Callitrichinae **4**:692
Callixylon **5**:555
Callosciurus **5**:603–604
Callosobruchus maculatus **4**:365
Calluna **5**:218
Calluna vulgaris **1**:142, **4**:577–579, **5**:522–523, **5**:527*F*
Callyspongia sp. **5**:124*F*
Calochortus **3**:206
Calomys laucha **1**:365–368
Calophyllum **1**:359
Calophyllum brasiliense **1**:163
Calophyllum lanigerum **5**:713
Calophyllum teysmannii **5**:713
Caloplaca aurantia **3**:674*F*
Caloppapus **4**:101
Calosoma helenae **3**:222
Calothamnus **5**:221
Calothrix **2**:611*F*, **2**:612*T*, **2**:621, **2**:622*F*, **4**:119
Calothrix rhizosolenia **2**:612–613, **2**:612*T*
Calotropis procera **5**:486
Caltha appendiculata **4**:107
Caltha dionaefolia **4**:107
Calvaria major **6**:52
Calvin–Benson cycle **2**:361–362, **2**:362*F*, **2**:654–655
Calycanthus **5**:621*F*
Calycophyllum benthamii **1**:163
Calycophyllum spruceanum **1**:163
Calypte helenae **1**:601
Calyptogena magnifica **4**:391
Calyptranthes luquillensis **3**:206
Calyptratae **3**:489–491*T*
Camargo evenness index **4**:616*F*
Camarhynchus parvulus **1**:27–28
Camarhynchus pauper **1**:27–28
Camarhynchus psittacula **1**:27–28
Cambodia
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  deforestation **2**:449*T*
  ecosystems
    genetic richness **1**:514*T*
    mangroves **1**:275*T*, **1**:513*T*
    threatened species **1**:278*T*
    tropical moist forest **1**:277*T*
    vegetation types and coverage percentage **1**:264–265*T*
  edible plants **3**:130*T*
  education level–infant mortality relationship **6**:204, **6**:205*F*
  endangered mammals **3**:194*T*
  plant species **1**:259*T*
  rice production **1**:88
  see also Southeast Asia
Cambrian period
  arthropod radiation ecology **2**:399–400
  assemblage evolution **3**:611–612, **3**:612*F*
  crustaceans **2**:403–407
  differentiation **2**:514–515
  fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
  marine diversity **5**:168*F*, **5**:169*F*
  marine invertebrates **4**:390, **4**:390*F*
  mass extinctions
    extinction intensities **5**:162*F*
    lesser mass extinctions **5**:175–176
    marine animals **1**:413*F*, **1**:493*F*
    notable mass extinctions **5**:161*F*
    number of families **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
  metazoa **3**:683, **3**:684*F*
  molluscs **5**:373–374, **5**:374*F*
  originations **1**:495, **1**:495*F*, **3**:425*F*
  periodicity **3**:424*F*
  vertebrate history **7**:334–335
camel crickets **3**:732, **3**:732*F*
Camelidae bactrianus **4**:697
Camelidae/camelids
  evolutionary history **2**:583*F*
  grassland ecosystems **6**:332–333
  Great American Biotic Interchange (GABI) **2**:584*F*
  late Quaternary extinction event (LQEE) **4**:724–725*T*, **4**:727*T*
  species diversity **4**:697
Camelidae dromedarius **4**:697
camelina **1**:529*T*
Camellia sinensis **2**:663, **5**:712*T*, **7**:274–275
Camelops **4**:723–725, **4**:724–725*T*
camels
  adipose tissue **7**:28, **7**:29*F*
  Asia **1**:269
  desert ecosystems **2**:472–474, **2**:476*T*
  grasslands and savannas **1**:269, **6**:333
  introduced species **1**:80, **1**:81
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  mangrove ecosystems **5**:14
  risk status classification **1**:644*F*
  species diversity **4**:697
camel spiders **1**:213, **5**:460*T*
Camelus bactrianus **2**:476*T*, **3**:191, **6**:333, **7**:29*F*
Camelus dromedarius **2**:476*T*
Camelus ferus **1**:269, **2**:469–470, **2**:476*T*
Camelus thomasi **5**:364–365
camembert cheese **3**:638–639
Cameroon
  botanical gardens **6**:519
  endangered mammals **3**:194*T*
  indigenous peoples **4**:270*T*
  primate diversity **6**:252*T*
  river–floodplain connectivity **7**:153–154
  slash-and-burn systems **6**:555–556
  termites **4**:430*T*, **4**:432*T*
  tropical rain forests **6**:315–317, **6**:317*T*
Cameroon Preuss's monkey **6**:256–258*T*
Camillidae **3**:489–491*T*
camouflage **5**:678
Campanian stage **2**:517–518, **2**:519*T*, **2**:520*T*
Campephilus principalis **7**:182

CAMPFIRE project **7**:19–20
camphechlor **5**:693
camphor **5**:712*T*
Camphorosmae **2**:376*F*
campina forests **1**:160–161
camping **4**:566–567, **7**:246–247, **7**:246*F*
Campnosperma brevipetiolatum **7**:272
Campnosperma panamensis **7**:272
campos rupestres **1**:167
Campos Ruprestres savannas **5**:496, **5**:497*F*
Camptorhynchus labradorius **7**:182
Camptostemon **5**:11*T*
Camptotheca acuminata **6**:135–136, **6**:135*F*
camptothecin **6**:135–136, **6**:135*F*
Campylobacter jejuni **2**:636*F*
Canacidae **3**:489–491*T*
Canacomyrica monticola **3**:211–212, **3**:214*F*
Canada
  amphibians **5**:565
  birds **5**:565
  Bois-des-Bel Peatland, Quebec, Canada (case study) **7**:375
  boreal forests **1**:626, **1**:631*T*
  botanical gardens **6**:519
  C_4 plants **2**:371*T*
  deforestation **2**:448–449
  Endeavor Hydrothermal Vents Marine Protected Area **5**:447
  energy efficiency measures **3**:254*F*
  energy use **3**:252*T*
  freshwater fish **5**:565
  indigenous peoples **4**:270*T*
  land mammals **5**:565
  natural history museums and institutions **5**:405–406
  protected area systems **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*, **6**:496
  regional carbon stocks **1**:678*F*
  reptiles **5**:565
  species richness **1**:499*T*
  wetland ecosystems **7**:386
  wheat production **1**:86
Canada geese **1**:79, **3**:365–366, **5**:312
Canada thistle **1**:75
Canadian Botanical Conservation Network **6**:86
Canadian Foundation for the Conservation of Farm Animal Genetic Resources **1**:645
Canadian Heritage Rivers System (CHRS) **6**:496
Canadian Museum of Nature **5**:406
Canadian Shield **5**:561*F*, **5**:562
Canadian three-spined sticklebacks **1**:32, **1**:32*F*
Canaima National Park **5**:444
canals **1**:123–124
canary grass **4**:366
Canary Islands
  adaptive radiation **5**:593–594
  disharmony **5**:591–592
  endemism **3**:233, **5**:592*T*
  species diversification **5**:594
Canavalia ensiformis **2**:663–664
Canavalia gladiata **2**:663
Canavalia rosea **1**:55
canavanine **2**:426*T*, **2**:427
Cancer borealis **3**:610*F*
cancer treatment
  recombination applications
    chromosome translocations **6**:367
    DNA repair and recombination **6**:367
    mitotic recombination **6**:367
cancer treatment drugs **1**:594
Cancun Agreements **2**:318, **4**:21–22
Candida **3**:634*T*
Candida albicans **3**:634
Candida krusei **3**:638–639
candidiasis **3**:634*T*, **7**:302–303
Candolle, Augustin P. de **1**:577–578, **2**:539–540
cane grass **2**:468
cane mouse **1**:367–368
cane toad **1**:78, **1**:80, **2**:74, **3**:167, **3**:695–696
Cane Toads (1987) **4**:654
Canidae
  body mass **1**:687*F*
  ecological characteristics **1**:688*F*
  evolutionary history **2**:583*F*
  food categories **1**:690*F*, **1**:691*F*
  home range size **1**:687*F*
  late Quaternary extinction event (LQEE) **4**:724–725*T*, **4**:727*T*
  number of species **4**:690*T*
  species diversity **1**:685, **4**:693
  tropical rain forests **6**:309*T*
Caniformia **4**:693
canine hookworm **1**:359
Canis **1**:709, **4**:724–725*T*
Canis aureus **5**:475
Canis familiaris **1**:709, **3**:191, **6**:361*T*
Canis latrans **1**:621–622, **1**:709, **4**:517
Canis lupus
  conservation practices **7**:229*T*
  dispersal biogeography **2**:546–547
  ecosystem/landscape conservation **7**:405–406
  habitat selection **4**:64
  Near East ecosystems **5**:475
  prey–predator interactions **1**:709
  reintroduction programs **3**:196
  social behavior **6**:573
  trophic cascades **7**:259
  urbanization threats **4**:564
Canis lupus pallipes **5**:472
Canis mesomelas **6**:573
Canis simensis **3**:191
cankers **3**:633*T*
Cannabis sativa **2**:663, **5**:712*T*
Canna edulis **4**:289*T*
cannibalism **5**:394
canola oil **1**:529*T*, **1**:530, **7**:26
canonical correspondance analysis (CCA) **4**:514
canopy fogging **1**:330
canopy gap dynamics **6**:398
Cansaga Bay **1**:276*T*
Canterbury mudfish **4**:194*T*
Cantharellus formosus **3**:639–640
Cantharidae **1**:327*T*
Cantharoidea **1**:327*T*
Cao-Vit crested gibbon **6**:256–258*T*
cap-and-trade programs **5**:150–152
Caparaceae **2**:138
Cape buffalo **6**:572–573
Cape Floristic Region **5**:219–220, **5**:526*T*, **6**:61, **6**:82*T*
Capellipedium **1**:70
Cape May warbler **2**:188–189, **2**:188*F*
Cape Region (South Africa)
  characteristics **5**:212–213
  natural disturbances **5**:213–214
  physical characteristics **5**:213*T*, **5**:219
  speciation patterns **5**:214
  species richness **5**:213*T*
  vascular plant diversity **5**:213*T*, **5**:219–220
  vertebrate diversity **5**:213*T*, **5**:220
Cape rock lobster **5**:57–58, **5**:60*F*
Caperonia castaneifolia **1**:164
capers **3**:130*T*, **3**:131*T*
Cape Verde **6**:519
Capironia fistula **1**:165
Capitella capitata **4**:384
Capitellidae **7**:417*F*
Capitonidae **6**:308*T*
Capparidaceae **5**:489
Capparimyia savastani **5**:469–470, **5**:469*F*
Capparis **1**:69, **5**:489
Capparis aegyptia **5**:484–485
Capparis angulata **3**:115–116
Capparis decidua **5**:486
Capparis tomentosa **1**:52
capped gibbon **6**:256–258*T*
capped langur **6**:256–258*T*
Capra aegagrus **2**:476*T*
Capra falconeri **3**:197
Capra, Fritjof **6**:384
Capra hircus **2**:476*T*, **3**:211
Capra ibex **1**:640, **3**:197
Capreolus capreolus **4**:515–516, **5**:472
Caprimulgidae **6**:308*T*
Caprimulgiformes **6**:308*T*
Caprimulgus aegyptius **5**:474
Caprinae **4**:696
Capritermes **4**:409*T*
Capromyidae **4**:713
Capromys pilorides **2**:587*F*, **4**:713–714, **4**:714*F*
Capsalidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Capsella bursapastoris **3**:131–132
Capsella grandiflora **2**:636*F*
capsicum **2**:384*F*, **4**:284–285, **4**:284*F*
Capsicum annuum **2**:663–664
Capsicum anumm **4**:289*T*
Capsicum baccatum **4**:289*T*
Capsicum brasilianum **4**:289*T*
Capsicum chinense **4**:289*T*
Capsicum frutescens **2**:663–664
captafol **5**:695
captan **5**:695
captive breeding
  amphibians **3**:698
  biodiversity restoration **6**:454
  California condor **6**:445
  characteristics and functional role **1**:669
  critically endangered birds **7**:188
  disease management programs **2**:536–537

evolutionary significant unit (ESU) **1:668–673**
basic concepts **1**:668–669
ecological exchangeability **1**:668–669
evolutionary lineage versus ecological context approach **1**:670–672, **1**:670*F*
hatchery fish **1**:670–672, **1**:670*F*
Pacific salmon **1**:669–670, **1**:670*F*
research summary **1**:672
livestock park conservation **4**:319
mammal conservation **3**:196–197, **4**:710–711
Pacific salmon **1**:669–670, **1**:669*F*
reintroduction programs **1:662–667**
breeding strategies **1**:663–666
challenges **1**:663
elements of success **1**:665*T*
genetic and demographic management **1**:663–666, **1**:664*F*, **1**:664*T*
goals **1**:662–663, **1**:664*T*
mammal conservation **3**:196–197
reintroduction to the wild **1**:666–667, **1**:666*T*
translocations **7**:236–237
capuchin monkeys **2**:590*T*, **6**:251*T*, **6**:256–258*T*
capybaras **2**:590*T*, **4**:697, **4**:724–725*T*, **6**:332–333
Carabidae **1**:247, **1**:250*T*, **1**:326–328, **1**:326, **1**:327*T*, **1**:709–710
carabid beetles **1**:374–375, **4**:194*T*, **4**:266–267
Carabodes **4**:394*F*
Caraboidea **1**:327*T*
Carabus hemprichi **3**:674*F*
Caracal caracal **5**:475
caracaras **1**:624
Caracasia **3**:522*T*
Caragana **1**:269, **1**:270
carambula **2**:663
Carapa guianensis **1**:163–164
Carapus acus **5**:668
cará yams **4**:287–290
carbamates
fungicides **5**:700*T*
herbicides **5**:696
insecticides **5**:694, **5**:699*T*, **5**:700*T*
carbaryl **5**:699*T*, **5**:700*T*
carbendazim **5**:695
carbofuran **5**:700*T*
carbohydrate storage **7:23–38**
adipose tissue
adipocyte size and number **7**:27
anatomical arrangement **7**:27–30, **7**:28*F*, **7**:29*F*
characteristics **7**:26–27, **7**:27*F*
fatness measures **7**:30–31, **7**:30*F*
functional role
adipocyte metabolism **7**:31–32
fatty acid composition **7**:32–33
thermal insulation **7**:33–35, **7**:34*F*
invertebrates **7**:24
life histories
energy storage optimization **7**:35
hibernation **7**:36–37
migration **7**:35–36
reproduction and lactation **7**:37–38
material characteristics **7**:23–24
plants
growth cycles **7**:24–25
lipid composition **7**:25–26
predator protection **7**:26
seed storage **7**:24–25
Carbonaceous period
fossil record **3**:684*F*, **3**:685*F*
mass extinctions **1**:493*F*, **5**:157*F*, **5**:161*F*, **5**:162*F*
originations **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
carbon (C)
biogeochemical cycles
atmospheric carbon dioxide (CO_2) **6**:147–149, **6**:148*F*
climate change effects **6**:147–149
ecosystem alterations
climate change effects **6**:149–150
coastal/marine ecosystems **6**:150–151
forest ecosystems **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
energy flow **3**:242–243
key reactions **1**:556
landscape legacies **4**:527
net primary production (NPP) **1**:559, **1**:560*T*
partial pressure data **1**:559*F*
sequestration and burial **1**:558–559
black carbon **4**:295–296
blue carbon **6**:77
calcium carbonate ($CaCO_3$) **1**:554, **1**:555*T*, **1**:556, **6**:147–149
carbon cycle **1:674–684**
atmospheric carbon dioxide (CO_2) **6**:147–149, **6**:148*F*
background information **1**:674
basic concepts
carbon dioxide (CO_2) concentrations **1**:674–676, **1**:675*F*, **6**:147–149, **6**:148*F*
carbon dioxide (CO_2) lifetime estimation **1**:680
deforestation flux **1**:676–678, **1**:677*F*, **1**:679*T*
fossil fuel emissions **1**:283–284, **1**:675*F*, **1**:678, **1**:679*F*
land sinks **1**:674–676
ocean sinks **1**:676
regional carbon stocks **1**:678*F*
schematic diagram **1**:675*F*, **6**:148*F*
transport systems **1**:675*F*, **1**:676
volatile organic compounds (VOCs) **1**:679–680, **1**:680*T*
weathering processes **1**:678–679
biological carbon (C) pump **2**:615–618, **2**:616*F*
carbon storage and sequestration **5**:351–352, **6**:264*F*
climate change effects **6**:147–149
dynamic global vegetation models (DGVMs) **2**:679
ecosystem alterations
climate change effects **6**:149–150
coastal/marine ecosystems **6**:150–151
forest ecosystems **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
future outlook **1**:683–684, **1**:683*F*
management strategies
carbon capture and storage **1**:681
deforestation reduction **1**:682–683
government regulation **1**:680–681
iron fertilization **1**:681
Reducing Emissions from Deforestation and Forest Degradation (REDD) **1**:682–683
soil respiration **1**:681–682
tree plantations **1**:682
marine sedimentary habitats **3**:282, **5**:111
savannas **1**:70–71
carbon dioxide (CO_2) **4:539–554**
atmospheric concentrations
biodiversity impacts **6**:126, **6**:127*F*
biogeochemical cycles **6**:147–149, **6**:148*F*
C_4 plants **2**:361, **2**:376*F*
climate change effects **2**:378
dynamic global vegetation models (DGVMs) **2**:675
plant–soil interactions **6**:126, **6**:127*F*
schematic diagram **6**:148*F*
sources and concentrations **1**:282*T*
biodiversity–global warming correlation **4**:227–228, **4**:227*F*, **4**:664
biofuel production
biomass emission estimation **4**:295–296
economic models **4**:294–295, **4**:294*F*, **4**:295*F*
food price increases and rationing **4**:296
global market impacts **4**:294–295, **4**:295*F*
indirect land-use change (ILUC) **4**:293–294
land cover change emissions **4**:295–296
land cover maps **4**:295
total emission–fuel quantity relationship **4**:296
biogeochemical cycles **1**:556, **1**:559*F*, **6**:147–149, **6**:148*F*
biological transport processes **1**:554
biomass emissions **4**:295–296
C_4 plants **2**:361, **2**:366–368, **2**:366*F*, **2**:368*F*, **2**:369*F*, **2**:374–377
carbon storage and sequestration **5**:351–352
changing concentrations
biodiversity impacts **6**:126, **6**:127*F*
causal factors **1**:282–284
fossil fuel emissions **1**:283–284
land use changes **1**:284, **4:539–554**
monthly average measurements **1**:283*F*
climate change effects **1**:281–282, **2**:58–59, **3**:407–408
deforestation **4**:294–295
dynamic global vegetation models (DGVMs) **2**:679

carbon (C) (*continued*)
ecotoxicological studies **3**:118–119
enhanced greenhouse effect **4**:18–19, **4**:19*F*, **4**:19*T*
fire/fire regimes **3**:441–442
Global Trade Analysis Project (GTAP) model **4**:539–540
consumption-side structure **4**:540*F*
economic and environmental impacts **4**:540–544
emission factors **4**:548–549
estimated emissions **4**:549–550, **4**:549*T*
final analysis **4**:551–552, **4**:551*T*
land cover and land-use activities **4**:542*F*
livestock production structure **4**:542*F*, **4**:543*F*, **4**:543*T*
marginal and average estimated emissions (first simulated group) **4**:550, **4**:550*T*
marginal and average estimated emissions (second simulated group) **4**:550–551, **4**:550*T*
marginal and average estimated emissions (third simulated group) **4**:551, **4**:551*T*
modeling process **4**:539–540
production-side structure **4**:541*F*
productivity measures **4**:540–544
research limitations **4**:553
research summary **4**:552–553
simulation models **4**:544–545
2001 simulation model **4**:544–545, **4**:544*T*, **4**:545*T*
2001–2006 simulation model **4**:545–547, **4**:546*T*, **4**:547*T*
2006–2015 simulation model **4**:547–548, **4**:548*T*, **4**:549*T*
well-to-wheel emissions **4**:551–552, **4**:552*T*, **4**:553*T*
global warming
biogeochemical feedbacks **4**:23
future climate change **4**:21–22, **4**:22*F*
global change indicators **4**:224–226, **4**:225*F*
global warming forecasts **2**:62
marine feedbacks **4**:23
past climate change **2**:82–83, **4**:19, **4**:20*F*
terrestrial feedbacks **4**:23–24
grazing effects **1**:706, **6**:340–341
indicator species **4**:255–256, **4**:257*F*
land-use change carbon dioxide (CO_2) emissions **4:539–554**
Late Triassic mass extinction **5**:173
lifetime estimation **1**:680
oceanic solubility **1**:559–560, **1**:560*F*
photosynthesis mechanisms **6**:13–15, **6**:14*T*
proxy assessments **4**:225*F*, **4**:226, **4**:227*F*
radiative forcing effects **4**:20*F*
savannas **1**:70–71
carbon disulfide (CS_2) **1**:282*T*
carbon monoxide (CO)
atmospheric concentrations **1**:282*T*
biomass emissions **4**:295–296
changing concentrations **1**:286
climate impact **1**:281–282
dynamic global vegetation models (DGVMs) **2**:680
fire/fire regimes **3**:441–442
grazing effects **6**:340–341
sources **1**:286
carbon:nitrogen (C:N) stoichiometry **6**:121–123
carbon:nitrogen:phosphorus (C:N:P) stoichiometry **3**:506–507, **4**:675–677, **4**:676*F*
carbon pools **1**:554, **1**:555*T*
carbon sequestration **5**:351–352, **6**:264*F*, **6**:561–562
carbon starvation hypothesis **2**:374–377
carbon storage **5**:351–352, **6**:264*F*
carbonyl sulfide (COS) **1**:282*T*, **1**:287
chlorofluorocarbons (CFCs)
atmospheric concentrations **1**:282*T*
changing concentrations **1**:284–286, **1**:286*F*
climate impact **1**:281–282, **3**:450
ecotoxicology **3**:118–119
enhanced greenhouse effect **4**:19*T*
environmental ethics **3**:271
ozone depletion **7**:296–297
dissolved inorganic carbon (DIC) **1**:559–560, **1**:560*F*
elemental partitioning **1**:554*T*
green carbon **6**:77
hydrofluorocarbons (HFCs) **1**:284–286, **4**:19*T*
perfluorocarbons (PFCs) **1**:284–286
relative abundance **1**:553*F*
river water composition **3**:562*T*
total organic carbon (TOC) **3**:351, **3**:352*F*
Carboniferous period
amphibians **3**:688
arachnids **1**:204–205
arthropod radiation ecology **2**:399
boundary age **5**:1*T*, **5**:2
crustaceans **2**:404
fossil record **1**:411*F*, **1**:412*F*
marine diversity **5**:168*F*, **5**:169*F*
mass extinctions **1**:413*F*, **3**:419*F*, **3**:425*F*, **4**:657*F*
synapsids **5**:1–2, **5**:1*F*, **5**:4
temperate forests **7**:123*F*
terrestrial plant communities
abundance patterns **6**:74*F*
evolutionary development **6**:65–67, **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
carboxin **5**:695
carboxylesterase **5**:269*T*
Carcinus maenas **4**:366, **5**:61, **5**:658*F*
Cardamine **4**:194*T*
Cardamine glacialis **4**:107
Cardamine pratensis **6**:106*F*, **6**:107
cardamom **2**:663
cardenolide toxicity **1**:24*F*, **5**:397
cardinalfish **3**:467–471
Cardinium **5**:678–679
Cards familiaris **1**:77
Carduus nutans **4**:361, **4**:363
CARE International **2**:231–232
Caretta caretta **5**:472
Carex **1**:269, **1**:270, **1**:271, **1**:278–279, **5**:215
Carex aquatilis **1**:232
Carex caryophyllea **4**:80*F*
Carex stans **1**:232
Carex vallis pulchrae **4**:107
Careya arborea **1**:70
Caribbean lizards **1**:31*F*, **1**:32
Caribbean monk seal (*Monachus tropicalis*) **3**:191–192, **5**:75
Caribbean region
biodiversity loss **7**:352–353
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
deforestation **2**:448*T*, **2**:454*T*
ecosystem conservation **3**:171–173*T*
ecosystem services
payment for ecosystem services (PES) **7**:355
water funds **7**:354–355*F*, **7**:356*T*
watershed management
best management practices **7**:358–359, **7**:358*T*
compensating impacts **7**:359
public protected area management **7**:358, **7**:358*T*
endemic plant species **6**:82*T*, **6**:83
indigenous peoples **4**:269*T*
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*
logging operations **2**:454*T*
mammalian extinctions **4**:705*T*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
nitrogen deposition **5**:526*T*
plant conservation **6**:83
plant hotspots **6**:61
protected area systems **6**:491*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
terrestrial bioregions **5**:561*F*, **5**:564
caribou
late Quaternary extinction event (LQEE) **4**:724–725*T*
migratory patterns **5**:313
oil pipelines **3**:255
phenological shifts **5**:723
tundra ecosystems **6**:330–331
Carica papaya **2**:663–664, **4**:289*T*, **5**:712*T*
Caridea **2**:416–417*T*
Carlemanniaceae **1**:260*T*
Carlia rubigularis **3**:144–145
Carludovica **3**:522*T*
Carmichaelia hollowayi **4**:194*T*
Carnarvon xeric shrublands **2**:468
Carnegiea gigantea **2**:475*F*
Carnidae **3**:489–491*T*
Carnivora/carnivores **1:685–694**
adipose tissue **7**:33–35, **7**:34*F*
anthropogenic impacts **1**:693–694
body mass **1**:686–687, **1**:687*F*
Cape Region (South Africa) **5**:220

characteristics and occurrences **4**:693
Chile **5**:217
convergent evolution **4**:701*T*
desert ecosystems **2**:470*T*
disturbance ecology **2**:565–566
diversity evolution **1**:693–694
ecological characteristics **1**:686–687, **1**:688*F*
food categories **1**:689, **1**:690*F*, **1**:691*F*
foraging behaviors **1**:687–690
home range size **1**:687*F*
late Quaternary extinction event (LQEE) **4**:727*T*
number of species **4**:690*T*
phylogenic patterns **1**:688
predator guilds **1**:690–692
representative species **2**:587*F*
skull size comparisons **1**:692*F*
social systems **1**:692–693
species diversity **1**:685–686, **2**:587, **4**:690*T*
territoriality **1**:692–693
trophic interactions **7**:264, **7**:265*F*
trophic pyramid **4**:670–671
tropical rain forests **6**:309*T*
carnivorous plants
benefits **6**:45
fitness costs **6**:45
general discussion **6**:44–45
prey capture mechanisms **6**:44–45, **6**:44*T*
wetland ecosystems **7**:395–397, **7**:396*F*
Carnoidea **3**:489–491*T*
Carolina parakeet **2**:281–282, **7**:182
Caroxyloneae **2**:376*F*
carp **1**:190*T*, **4**:378, **7**:390, **7**:399
Carpenter's white-handed gibbon **6**:256–258*T*
Carpodacus mexicanus **2**:528*T*, **2**:551–552*T*
Carpophilus mutulatis **5**:673
Carposinoidea **5**:387–388*T*
carr **7**:369*T*
carrion flies **6**:49*T*
carrots
native crops **2**:662–663
regional cultivars **3**:32*F*
carrying capacity **1:695–701**
aquaculture **1**:196–198
background information **1**:695
current research areas
logistic growth equation **1**:699–700
metapopulation theory **1**:699–700
regulation mechanisms **1**:699
definitions **1**:696–697
determination methods **1**:698–699
ecological footprint analysis (EFA) **2**:704–705
historical perspective **1**:695–696
human carrying capacity **1**:700, **2**:704–705
logistic growth **1**:696, **1**:696*F*, **1**:697, **1**:697*F*, **1**:699–700
population dynamics **1**:700, **3**:3
population models **1**:697–698, **1**:698*F*
resource management **6**:563
stability theory **7**:3*F*, **7**:4
summary discussion **1**:700–701
sustainability **7**:75
U.S. population growth **1**:696, **1**:696*F*, **1**:697*F*
Carson, H. L. **6**:624
Carson, Rachel **1**:112, **2**:285–286, **3**:281–282, **4**:121–122, **4**:652–653
Cartagena Protocol on Biosafety **2**:306–307, **2**:307
cartilaginous fishes
characteristics **3**:456–457, **3**:458–459
chimaeras **3**:460, **3**:460*F*
rays **3**:459–460
sharks
characteristics **3**:456–457, **3**:458–459, **3**:459*F*
extinction theories **2**:517
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
skates **3**:459–460, **3**:460*F*
threatened species **7**:201*T*
Carya **7**:118
Caryocar glabrum **1**:159–160
Caryocar villosum **4**:287–290, **4**:288*T*
Caryodendron amazonicum **4**:288*T*
Caryophyllaceae **2**:374*T*
Caryophyllales **2**:374*T*, **2**:375*F*, **5**:622*F*
Caryophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cascades frog **3**:165
cascara **5**:712*T*
Casearia **5**:489
cashmere **1**:640
Caspian lowland desert **2**:469–470
Caspian seal **5**:81*T*
cassava **1**:132*T*, **2**:382–383, **2**:383, **2**:663–664, **3**:32*F*
cassava mealybugs **4**:183, **4**:361
Cassia **1**:267
Cassia angustifolia **5**:712*T*
Cassia occidentalis **1**:164
Cassia reticulata **1**:165
Cassidinae **1**:328
Cassiope tetragona **1**:232, **1**:236*F*
cassowaries **1**:606
Cassytha
economic impact **4**:77
hemiparasitism **4**:71
tropical rain forests **6**:315*T*
Cassytha filiformis **4**:77
Castanea dentata **2**:528*T*, **3**:411, **4**:361–362
Castanea henryi **2**:663
Castanopsis **1**:267, **1**:270–271, **7**:118
Castilleja **4**:71
Castor canadensis **6**:466, **7**:390, **7**:405–406
Castoridae **6**:309*T*
Castorimorpha **2**:586
Castoroides **4**:724–725*T*
castor oil beans **1**:529*T*, **3**:703, **3**:704*F*, **7**:26
Casuarina **1**:279, **5**:623*F*, **7**:399
Casuarinaceae **5**:621–625
Casuarina oligodon **6**:558–559
Casuarina pauper **1**:292
Casuarius **1**:606
Catagonus spp. **6**:332–333
catalases **5**:269*T*
Catalina Island mahogany **4**:167
catalytic hypercycle theory **5**:633–636, **5**:634*F*, **5**:635*F*
cataracts **7**:302
Cat Ba langur **6**:256–258*T*
Catenulida **7**:439–440*T*, **7**:440, **7**:441*F*, **7**:442–443*F*
caterpillars
butterfly life cycle **1**:652*F*, **1**:654
number of species **4**:371*T*
catfish **1**:190*T*, **3**:464, **6**:50, **6**:479, **7**:58, **7**:390
Catharanthus roseus **1**:359, **1**:594, **5**:706–707, **6**:57
Cathartidae **6**:308*T*
Catharus bicknelli **7**:186
Catharus ustulatus **7**:187
Cathaya **1**:272–273
Cather, Willa **4**:652
Cathetocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
cation exchange capacity (CEC) **6**:554–555
Catopsis **6**:44*T*
Catreus wallichii **7**:188
cats
adipose tissue **7**:34*F*
deoxyribonucleic acid (DNA) evolution **2**:56*F*
fungi-related diseases **3**:635–636
introduced species
Australia **1**:80
Brazil **1**:83
India **1**:82
predation effects **6**:240–241
South Africa **1**:81
United Kingdom **1**:79
United States **1**:76–77
late Quaternary extinction event (LQEE) **4**:724–725*T*
nematodes **5**:665*T*
predation effects **6**:237, **6**:240–241
species diversity **1**:686
cattails **1**:278–279, **7**:372, **7**:388–389
cattle **1:702–711**
animal breeding programs **1**:640
biotic relationships
birds **1**:710
invertebrate fauna **1**:709–710
natural predators **1**:709
parasites and pathogens **1**:710–711
plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
seed dispersal and germination **1**:708–709
small mammals **1**:710
domestication **1**:702–703
ecological role
desertification **1**:704
erosion **1**:704–705
grazing effects **1**:703–704
land degradation **1**:704–705
nutrient cycling **1**:705–706, **1**:705*F*
over-enrichment factors **1**:705–706
plant diversity–grazing intensity relationship **1**:704*T*
vegetation impacts **1**:703–704
food habits **3**:127*T*

cattle (*continued*)
introduced species
Australia **1**:80
Brazil **1**:83
South Africa **1**:81
United Kingdom **1**:79
United States **1**:76–77
risk status classification **1**:644*F*
cattle egrets **1**:710, **2**:551–552*T*, **4**:357, **5**:13–14
Caucasian–Iranian highlands **1**:262–263*T*
Caucasus **5**:526*T*, **6**:82*T*
caudates
characteristics **7**:339
evolutionary history **1**:169
functional anatomy **1**:170–171
larval morphology and ecology **1**:173–174
life histories **1**:171–172
morphology **1**:170–171
postlarval behavior and ecology **1**:175–176
reproductive biology **1**:171–172, **1**:173*F*
special features **1**:171
species richness **1**:170*T*
Caudofoveata **5**:376
Caudovirales **5**:130*F*, **5**:134–135, **5**:134*T*
caulerpa **1**:190*T*
Caulerpa racemosa **4**:358–359, **4**:359–360
Caulerpa taxifolia **3**:284, **3**:359, **3**:411, **4**:357, **4**:359–360
cauliflower **2**:383
cave clams **7**:57
cave crickets **3**:221, **3**:732, **7**:51–52, **7**:60
cave fish **3**:466–467
Cavendishia **3**:522*T*
caves **3**:221, **5**:431, **6**:299, **7**:49–50, **7**:50*F*
see also subterranean ecosystems
Caviomorpha **2**:583*F*, **4**:697
Cavitermes **4**:409*T*
Cayman Islands **5**:493*T*, **5**:495*F*, **5**:497*F*, **6**:519
Caytonanthus **5**:616*F*
Caytonia **5**:616*F*
Caytonia **5**:618*F*
Caytonia **5**:615–617
Caytoniales **5**:618*F*
Ceanothus **5**:215
Cebidae **4**:692, **6**:250–252, **6**:251*T*, **6**:309*T*
Cebuella pygmaea **2**:586
Cebu flowerpecker **7**:183
Cebus albifrons aequatorialis **6**:256–258*T*
Cebus albifrons malitiosus **6**:256–258*T*
Cebus albifrons trinitatis **6**:256–258*T*
Cebus albifrons versicolor **6**:256–258*T*
Cebus capucinus **2**:324, **5**:603
Cebus kaapori **6**:256–258*T*
Cebus macrocephalus **6**:254
Cebus olivaceus **6**:254
Cebus paella **3**:127*T*
Cech, Thomas **5**:631–632
Cecidomyiidae **3**:489–491*T*, **3**:492, **3**:493–494
Cecropia schrevenriana **3**:247
Cecropia spp. **2**:26, **2**:27–28
Cecropsis **5**:618*F*
Cedar Creek (Minnesota) prairie grassland field experiment **4**:585–586, **4**:585*F*
cedars
Asia **1**:268–269
boreal forests **1**:628–630
cedars of Lebanon **6**:76–79
North America **7**:118
Cedrela odorata **4**:595, **6**:560
Cedrela spp. **2**:25–26
cedro **6**:560
Cedrus **1**:268–269
Cedrus atlantica **1**:47, **1**:268–269
Cedrus brevifolia **1**:268–269
Cedrus deodara **1**:268–269, **1**:272–273
Cedrus libani **1**:268–269, **6**:76–79
Ceiba pentandra **1**:163
Celaenodendron mexicanum **7**:274
Celastraceae **3**:522*T*, **5**:623*F*
Celebes **1**:262–263*T*
celery
intercropping patterns **2**:383
native crops **2**:662–663
Celidophylla albimacula **3**:725*F*
cell phone use **7**:83*F*
cellulase **5**:269*T*
cellulose *see* termites
Celosia argentea **4**:288*T*
Celtis **5**:489
Celtis australis **1**:47
Celyphidae **3**:489–491*T*
Cenarchaeum symbiosum **6**:300
Cenchrus **1**:69
Cenchrus biflorus **1**:67
Cenchrus ciliaris **1**:67
Cenomanian/Turonian (C/T) mass extinction **5**:176
Cenothyla varians **1**:250*T*, **1**:251*F*
Cenozoic era
Antarctic ecosystems **6**:617
boundary age **5**:1*T*
crustaceans **2**:406
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*
global change indicators **4**:224–226, **4**:225*F*
mammal ages **2**:583*F*
mammal diversification **3**:688
mammal extinctions **4**:658*F*, **4**:721–725, **4**:727*F*
marine diversity **5**:168*F*, **5**:169*F*
marine invertebrates **4**:390–391, **4**:390*F*
mass extinctions **1**:413*F*, **3**:419*F*, **5**:157*F*, **5**:176
molluscs **5**:374*F*
originations **1**:495, **3**:425*F*
periodicity **3**:424*F*
temperate forests **7**:123*F*
Census of Antarctic Marine Life (CAML) **2**:21
Census of Coral Reefs (CReefs) project **2**:16
Census of Diversity of Abyssal Marine Life (CeDAMar) **2**:17, **2**:18*F*
Census of Marine Life **2**:**1–22**
abundance measures **2**:8–13, **2**:12*F*
census expeditions **2**:4*F*
census field sampling projects **2**:4*F*
discovered species **2**:3–8, **2**:5*F*, **2**:6*F*, **2**:8*F*
distribution and migration patterns **2**:7–8, **2**:9*F*, **2**:10*F*, **2**:11*F*
DNA barcoding programs **2**:13–14
goals and scope **2**:1–3
Klee diagrams **2**:6, **2**:7*F*
legacies **2**:14–15, **2**:14*F*
participating countries **2**:2*F*
population declines **2**:8–13, **2**:12*F*
public relations **2**:14–15
recreational fishing **2**:3*F*
research approaches
active geological regions **2**:15*F*, **2**:19–20, **2**:20*F*
central waters
biologging **2**:18–19, **2**:19*F*
characteristics **2**:17–19
dark zone **2**:15*F*, **2**:19, **2**:20*F*
light zone **2**:15*F*, **2**:18–19, **2**:19*F*
mid-water/bottom-water zones **2**:15*F*, **2**:19, **2**:20*F*
cold seeps **2**:15*F*, **2**:19–20
continental shelves
coastal zone **2**:15*F*, **2**:16, **2**:17*F*
human edge concept **2**:15–16, **2**:15*F*
nearshore zone **2**:15*F*, **2**:16, **2**:16*F*
realms of study **2**:15–16, **2**:15*F*
general discussion **2**:15–16
hidden zones
abyssal plain **2**:15*F*, **2**:17, **2**:18*F*
characteristics **2**:16–17
slope zone **2**:15*F*, **2**:17, **2**:18*F*
hydrothermal vents **2**:15*F*, **2**:19–20, **2**:20*F*
marine microorganisms **2**:15*F*, **2**:21, **2**:21*F*
polar oceans **2**:20–21, **2**:21*F*
seamounts **2**:15*F*, **2**:19–20, **2**:20*F*
research results **2**:3–8
research strategies **2**:1–3, **2**:3*F*, **2**:4*F*, **2**:5*F*
still-to-be-explored ocean **2**:5*F*
technological innovations **2**:13–14, **2**:13*F*
unknown/unexplored ocean **2**:8, **2**:11*F*
Census of Marine Zooplankton (CMarZ) **2**:18–19
Census of Seamounts (CenSeam) **2**:20, **2**:20*F*
Centaurea maculosa **4**:94, **4**:363, **6**:93–95
Center for Global Development **4**:330
Center for Plant Conservation (CPC) **6**:86
Centers for Disease Control (CDC) **5**:670–672
centipedes
characteristics **4**:392
developmental modes **5**:426
endangered species **3**:226
geographic distribution
endemism **5**:428
environmental characteristics **5**:427–428
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429, **5**:429*T*
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430

open-country myriapods **5**:430
special habitat myriapods **5**:430–431
morphology **5**:421–423, **5**:422*F*
number of species **5**:427*T*
soil fauna **6**:590*T*
subterranean ecosystems **7**:60
Centraalbureau voor Schimmelcultures (CBS) **3**:652*T*
Central Africa
deforestation **2**:448*T*
protected area systems **6**:491*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
Central African Republic
botanical gardens **6**:507*T*
energy efficiency measures **3**:254*F*
urbanization **4**:560*F*
Central America
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
C_4 plants **2**:371*T*
Convention on Biological Diversity (1930) **2**:314
deforestation **2**:448*T*, **2**:449*T*, **2**:452*T*, **2**:454*T*
ecosystems **2:23–32**
biodiversity
climatic regions/latitudinal range **2**:24–26
endemism **2**:26
floristic elements **2**:25–26
species richness **2**:25, **2**:25*T*
biogeography **2**:23–24, **2**:24*F*
global environmental change
climatic change **2**:30
general discussion **2**:30
land use change **2**:30–31, **2**:31*F*
high-elevation ecosystems **2**:30
life zones **2**:26–27
mangroves **2**:29
seasonally dry tropical forests **2**:28
temperate forests **2**:29–30
tropical cloud forests **2**:28–29
tropical rain forests **2**:27–28
endangered mammals **3**:194*T*
fire/fire regimes **3**:435–436
fish diversity **3**:472
intercropping systems **2**:382–383
lake and pond ecosystems **4**:461*T*
logging operations **2**:454*T*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
nitrogen deposition **5**:526*T*
oil-palm plantations **5**:601
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
terrestrial bioregions **5**:563–564
transboundary protected areas **6**:489*T*
Central American Biodiversity Strategy **2**:314
Central American tapir **7**:229*T*
central Andean dwarf *Polylepis* forest **4**:105
Central Asia
deforestation **2**:448*T*
desert ecosystems **2**:469–470
plant species **6**:82*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
Central Brazilian Shield **3**:105–107, **3**:106*F*
central chimpanzee **6**:256–258*T*
Central Europe **3**:341–344*T*
centralized regimes **2**:156–157
Central Ranges xeric scrub **2**:468
Centrarchidae **3**:468–469
centrioles **2**:327
Centrocercus urophasianus **2**:503–505, **3**:256–257
Centro Internacional de AgriculturaTropical (CIAT) **3**:643
Centro International de la Papa (CIP) **3**:652–653
Centro International de Mejoramiento de Mais y Trigo (CIMMYT) **3**:643
Centro Nacional de Pesquisa de Recursos Geneticos e Biotecno-logia (CENARGEN) **3**:652–653
Centropelma micropterum **1**:606
Centropodia **2**:376*F*
Centrosaurus **2**:519*T*, **2**:520*T*
Centrostephanus rodgersii **4**:516
Centrostigma macrophyllum **1**:161
Centruroides exilicauda **1**:208
Centruroides spp. **1**:208
CENTURY Soil Organic Matter model **5**:352, **5**:357–358
Cepaea hortensis **2**:141–142
Cepaea nemoralis **2**:141, **6**:5
Cephaelis ipecacuanha **5**:712*T*
Cephalocarida **2**:400–401, **2**:401*T*, **2**:416–417*T*
Cephalochordata **3**:457, **7**:336
Cephalohus **1**:49
Cephalopachus bancanus bancanus **6**:256–258*T*
Cephalopachus bancanus saltator **6**:256–258*T*
Cephalophinae **4**:696
Cephalopoda/cephalopods
fossil record **3**:681*F*, **3**:685*F*
number of species **5**:669–670*T*
physical characteristics **5**:379–380, **5**:380*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
Cephalorhyncha **7**:434*F*
Cephalorhynchus hectori **3**:192, **5**:81*T*
Cephalorhynchus hectori maui **5**:77*T*
Cephalotaceae **5**:221, **6**:44*T*
Cephalotaxaceae **1**:260*T*
Cephalotaxus **1**:272–273
Cephalotermes rectangularis **4**:429*F*
Cephalotus **6**:44*T*
Cephidae **4**:182–183
Cephoidea **4**:178
Cephus cinctus **4**:182–183
Cerambycidae **1**:327*T*, **1**:328
Ceraphronoidea **4**:178
Cerapterix graminis **6**:159*F*
Cerastes **5**:467–469
Cerastes vipera **5**:472
Cerastium beeringianum **1**:236*F*
Cerastium floccosum **4**:103
Ceratium **6**:29–30, **6**:294
Ceratium furca **5**:684*F*
Ceratium hirundinella **6**:33*F*
Ceratium tripos **6**:30
Ceratocystis fagacearum **3**:633, **3**:633*T*, **4**:363
Ceratocystis ulmi **2**:528*T*
Ceratonia **5**:218
Ceratonia siliqua **1**:269, **5**:480*F*, **5**:482–483, **5**:485
Ceratonotus steiningeri **2**:18*F*
Ceratophyllum **5**:621*F*
Ceratopogonidae **3**:489–491*T*, **3**:492
Ceratopsia **2**:520*T*
Ceratopteris **6**:59
Ceratopteris pteridioides **1**:164, **1**:165
Ceratotherium simum **4**:695, **5**:225*T*
Ceratotherium simum cottoni **3**:198
Cercartetus lepidus **4**:688–689
Cercidiphyllaceae **1**:260*T*, **5**:622*F*
Cercidiphyllum japonicum **1**:271
Cercocarpus betuloides **4**:167
Cercocarpus traskiae **4**:167
Cercocebus galeritus **6**:256–258*T*
Cercocebus lunulatus **6**:256–258*T*
Cercocebus sanjei **6**:256–258*T*
Cercomela schlegelii **1**:487–488
Cercomonas **6**:294
Cercopithecidae **4**:692, **6**:250–252, **6**:251*T*
Cercopithecus albogularis zammaranoi **6**:256–258*T*
Cercopithecus dryas **6**:256–258*T*
Cercopithecus erythrogaster erythrogaster **6**:256–258*T*
Cercopithecus kandti **6**:256–258*T*
Cercopithecus mitis schoutedeni **6**:256–258*T*
Cercopithecus roloway **6**:256–258*T*
Cercotrichas galactotes **5**:474
cereal leaf beetle **2**:551–552*T*
cereals
East Asia **2**:663
Europe **2**:662–663
Mediterranean Basin **2**:662–663
North and South America **2**:663–664
Southeast Asia **2**:663
southwest Asia **2**:662–663
Sub-Saharan Africa **2**:663
Ceriodaphnia **6**:26
Ceriops **1**:277, **5**:11*T*
Cerithidea **5**:17*F*
Cerithidea fuscata **3**:202*T*
Cerithium moniliferum **5**:662*F*
Ceroxylon echinulatum **4**:516
cerradão **1**:161–162, **1**:164–165
Cerrado **6**:82*T*
Cerrado ecosystems **5**:496, **5**:497*F*
Certhia brachydactyla **1**:618–619
Certhia familiaris **1**:618–619
Certhidea olivacea **1**:27–28
Certhiidae **4**:33–35, **4**:34*F*, **6**:308*T*
Certhilauda barlowi **1**:487–488
Certhilauda erythrochlamys **1**:487–488
Cervalces **4**:723–725, **4**:724–725*T*
Cervidae **2**:583*F*, **4**:696, **4**:727*T*, **6**:309*T*
Cervus **4**:724–725*T*
Cervus elaphus **2**:551–552*T*, **2**:719–720, **3**:127*T*, **4**:65, **4**:516, **5**:313, **5**:723, **6**:332, **7**:405–406

Cervus unicolor **5**:604
Ceryle rudis **5**:476–477
Cesalpino **7**:89–90
Cesspool Cave (Virginia, USA) **7**:51
cestodes
  Amphilinidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
  characteristics and properties **2**:524*T*
  domesticated ruminants **1**:710–711
  Gyrocotylidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
  life cycle **7**:463*F*
  number of species **5**:669–670*T*
  phylogenetic relationships **7**:441*F*
  species diversity **7**:439–440*T*, **7**:455–457
Cestoplana **7**:442–443*F*
cetaceans
  adipose tissue **7**:34
  Census of Marine Life project **2**:12*F*
  characteristics **4**:693–694, **5**:65*T*, **5**:66–67
  climate change effects **5**:89–91
  convergent evolution **4**:701*T*
  evolutionary history **5**:68–69
  number of species **4**:690*T*
  representative species **2**:587*F*
  species diversity **2**:587, **4**:690*T*, **4**:693–694
  taxonomy **5**:65*T*
Cetorhinus maximus **3**:478*F*
Ceuthophilus maculatus **3**:732*F*
Ceylonese monsoon forest **1**:262–263*T*
Ceylonese rain forest **1**:262–263*T*
Chaco dry forests **5**:496, **5**:497*F*
Chaco region **3**:105–107, **3**:106*F*
Chactidae **1**:209
Chad
  botanical gardens **6**:507*T*
  domesticated ruminants **1**:702–703
  energy efficiency measures **3**:254*F*
  indigenous peoples **4**:270*T*
  urbanization **4**:560*F*
Chad (Lake) **3**:171–173*T*, **4**:461*T*
Chaenophryne longiceps **5**:122*F*
Chaerilidae **1**:209
Chaeropodidae **4**:688
Chaetanthera apiculata **4**:106
Chaetanthera lycopodioides **4**:106
Chaetanthera pusilla **4**:106
Chaetoceros **1**:475, **2**:612*T*, **6**:30
Chaetodermomorpha **5**:376, **5**:376*F*
Chaetognatha
  Census of Marine Life project **2**:7*F*
  global diversity **4**:385*T*
  marine ecosystems **5**:50*T*
  number of species **5**:669–670*T*
  phylogenetic relationships **7**:433*F*, **7**:434*F*
  species diversity **4**:386*T*
Chaetomorpha linum **3**:359
Chaetopteridae **7**:417*F*
chaffinches **1**:618
Chagas' disease **5**:660*T*, **5**:668, **5**:671*T*, **7**:292–293
Chagos Islands **1**:262–263*T*
Chalai Lagoon **1**:276*T*
Chalan Beel **1**:276*T*
Chalcidoidea **4**:178, **4**:181, **5**:395–396, **5**:674
chalcidoid wasps **5**:674–675
Chamaecrista fasciculata **4**:248*F*
Chamae-cyparis **1**:272–273
Chamaecyparis lawsoniana **6**:63
Chamaecyparis obtusifolia **1**:273
Chamaecyparis thyoides **7**:400
Chamaedorea elatior **2**:27
Chamaeleo chamaeleon musae **5**:472
Chamaemyiidae **3**:489–491*T*
Chamaerops **5**:218
Chamaerops humilis **1**:47
Chamaesipho tasmanica **4**:335
Chamaportketes **4**:724–725*T*
Chambersiella **7**:426*F*
chameleons **5**:472
chamois **5**:313
Chamorro culture **1**:592
Chandler, Raymond **4**:652
Changhsingian Stage **5**:169*F*
Channa maculata **4**:465
Chanos chanos **3**:464
chanterelles **3**:629–630
Chao-2 estimator **3**:541–542
Chaoboridae **3**:489–491*T*
chaoborines **6**:26
Chao–Jaccard index **5**:209–210
Chao's abundance-based coverage estimator **5**:188–190
Chao's abundance-based estimator **5**:187–188, **6**:652–653
Chao's incidence-based coverage estimator **5**:190–191
Chao's incidence-based estimator **5**:188
Chao–Sørensen index **5**:209–210
Chapala (Lake) **4**:461*T*
char **3**:464–465
Chara **3**:365–366
characiforms/characins **3**:464, **3**:464*F*, **5**:315, **6**:50, **6**:479
character displacement **6**:720
character state biodiversity **1**:430
Charadriidae **1**:710
Charadrius alexandrinus **5**:476–477
Charadrius dubius **5**:476–477
Charaxinae **1**:653
charcoal manufacture
  deforestation **2**:453
  South America **6**:609
Charitoceros **4**:727*T*
Charles River **7**:152–153
Charlotte's Web(White) **4**:653
Charophyceae **3**:571*T*
Charybdis smithii **4**:386*T*
Chasmosaurus **2**:519*T*, **2**:520*T*
Chatham Island oyster catcher **4**:194*T*, **7**:185
Chatham Island snipe **7**:188
Chattonella **1**:475
Chaucer, Geoffrey **4**:650–651
cheaters **5**:640–641
cheatgrass **1**:703, **4**:347–348, **4**:348*F*, **4**:360, **6**:93, **6**:125
cheer pheasant **7**:188
cheese **3**:638–639
cheetahs
  conservation efforts
    background and characteristics **4**:715–716, **4**:717*F*
    conservation practices **4**:716–717
    major threats **4**:716
    range areas **4**:718*F*
  convergent evolution **4**:701*T*
  evolutionary bottlenecks **2**:357–358
  genetic diversity **3**:190–191
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  Near East ecosystems **5**:472
  research challenges **5**:369
  species diversity **1**:686, **4**:693
Cheilodactylus spectabilis **5**:33
cheilostome bryozoa **2**:332*T*
Cheirogaleidae **6**:250–252, **6**:251*T*
Chelicerata
  characteristics **1**:202–204
  phylogenetic relationships **1**:203*F*
Chelifer cancroides **1**:212–213
Chelojulus sculpturatus **5**:424
Chelonia **6**:400–401, **7**:339–340
Chelonia mydas **2**:279, **3**:217, **3**:218, **5**:472, **6**:78, **6**:410–411, **6**:547–548, **7**:229*T*
Chelydra serpentina **6**:405
chemical communication **4**:182
chemical contamination **3**:281–282
chemoautotrophs
  energy flow **3**:242
  evolutionary development **1**:563–564, **2**:640–641
  hydrothermal vent ecosystems **7**:328–329
  subterranean ecosystems **7**:58–59, **7**:59*T*
chemoheterotrophs
  evolutionary development **2**:640–641
  nitrogen fixation **2**:614–615
chemosynthetic ecosystems **7:327–332**
chemosynthetic symbioses **5**:120–121, **5**:120*F*
Chenopodiaceae **2**:371*T*, **2**:374*T*
Chenopodium **1**:120
Chenopodium album **2**:368*F*, **3**:703, **6**:433
Chenopodium berlandieri **2**:663–664
Chenopodium fremontii **2**:373*T*
Chenopodium quinoa **2**:484*F*, **2**:663–664, **3**:36
cherimoya **2**:663–664
cherries **2**:662–663
Chesapeake Bay
  case studies **3**:303–306, **3**:304*F*, **3**:305*F*, **3**:306*F*
  eutrophication impacts **3**:354–356, **5**:28*T*, **5**:29–30
chestnut blight fungus **2**:528*T*, **3**:628–629, **3**:633*T*, **4**:361–362
chestnut trees **2**:528*T*, **2**:530–531, **3**:411, **4**:361–362
Chetverikov, Sergei **3**:394
chevron skink **4**:194*T*
chewing lice **4**:392*T*
chiasmata **6**:353
Chicago School **7**:311
chickens
  deoxyribonucleic acid (DNA) evolution **2**:56*F*
  risk status classification **1**:644*F*

chickpeas
  intercropping systems **2**:390
  native crops **2**:662–663
  Neolithic agriculture **2**:658
chicory **3**:131–132
Chicxulub impact crater **5**:174–175
chiggers **1**:215
Chilades pandava **6**:84
Chile
  biodiversity-based agricultural strategies **1**:124–125
  botanical gardens **6**:519
  C_4 plants **2**:371*T*
  desert ecosystems **2**:469
  ecosystem services **7**:356*T*
  fishing/fisheries **6**:422*T*
  inland deserts **2**:462*F*
  land cover classifications **6**:607*T*
  Mediterranean-type climate ecosystems
    characteristics **5**:212–213
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:215–216
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:216–217
    vertebrate diversity **5**:213*T*, **5**:217
  nitrogen deposition **5**:526*T*
  nonvascular epiphytes **3**:521*T*
  plant diversity **3**:700–701, **6**:82*T*
  plant species **4**:102*T*
  seed dispersal and germination **1**:708–709
  temperate forests
    distribution patterns **7**:116*T*, **7**:117
    diversity patterns **7**:119–121, **7**:119*T*
    dominant floristic features **7**:119
    ecoregion classifications **6**:600–604*T*
    endemism **7**:121
  water funds **7**:356*T*
Chilean jack mackerel **6**:422*T*
Chiliadenus iphionoides **5**:484–485
Chiliadenus montanus **5**:484–485
Chilka Lake **1**:276*T*
Chilodonella **6**:295, **6**:296*F*
Chilomonas **6**:294
Chilopoda
  Amazon ecosystems **1**:246–247, **1**:246*T*, **1**:247
  characteristics **4**:392
  developmental modes **5**:426
  endangered species **3**:226
  geographic distribution
    endemism **5**:428
    environmental characteristics **5**:427–428
    hotspots **5**:429
    islands **5**:429
    regionalization **5**:428–429, **5**:429*T*
    speciation **5**:428
  habitats and adaptations
    burrowing myriapods **5**:430
    evolutionary development **5**:429–430
    open-country myriapods **5**:430
    special habitat myriapods **5**:430–431
  morphology **5**:421–423, **5**:422*F*
  number of species **5**:427*T*
  soil fauna **6**:590*T*
Chilotherium **5**:224*T*
Chilwa (Lake) **4**:461*T*
chimaeras **3**:456–457, **3**:460, **3**:460*F*
Chimaericola **7**:448–449*F*
Chimaericolidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Chimango caracara **1**:624
Chimaphila **1**:271
Chimarra lejea **5**:463*T*
chimpanzees
  conservation biology **2**:243–244
  conservation practices **7**:229*T*
  global diversity **6**:251*T*
  introduced parasites–species decline relationships **2**:528*T*
  oil-palm plantations **5**:603–604
  reciprocal altruism **2**:325
  species diversity **4**:693
  termite fishing **4**:413–414, **4**:415*F*
  tropical Africa **1**:49
China
  animal species **1**:261*T*
  biodiversity richness **1**:501*F*
  botanical gardens **6**:519
  crop plant domestication **2**:658–659
  dam projects **1**:512
  deforestation **1**:502*T*, **1**:677*F*, **2**:448–449
  ecosystems
    coral reefs **1**:500*T*
    laurel forests **1**:270–271
    mangroves **1**:275*T*
    marshes **1**:278–279
    mountain regions **1**:272
    Nature Reserve System (NRS) **3**:372–373, **3**:373*F*, **3**:374*F*
    temperate deciduous forests **1**:271–272
    temperate grasslands **1**:269
    temperate semi-deserts and deserts **1**:269–270
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    tropical moist forest **1**:277*T*
    tropical rain forests **6**:317*T*
    vegetation types and coverage percentage **1**:264–265*T*
  ecosystem services **3**:**372–384**
    background information **3**:372
    Ecosystem Function Zones (EFZs) **3**:382
    Grain-to-Green Program (GTGP)
      background information **3**:379–380
      biodiversity/biophysical effects **3**:379*F*, **3**:380
      future outlook and challenges **3**:381–382
      program goals **3**:379–380
      socioeconomic effects **3**:380–381, **3**:381*F*
      spatial distribution **3**:376*F*, **3**:379*F*, **3**:380
      summary discussion **3**:382
    land-use planning and human development **3**:101
    National Forest Conservation Program (NFCP)
      background information **3**:375–376
      biodiversity/biophysical effects **3**:376–377, **3**:377*F*
      future outlook and challenges **3**:378–379
      program goals **3**:375–376
      socioeconomic effects **3**:377–378, **3**:378*F*
      spatial distribution **3**:376, **3**:376*F*
      summary discussion **3**:382
    Nature Reserve System (NRS)
      background information **3**:372–373
      biodiversity/biophysical effects **3**:373–374
      future outlook and challenges **3**:375
      geographic distribution **3**:373, **3**:374*F*
      number of reserves **3**:373*F*
      program goals **3**:372–373
      socioeconomic effects **3**:374–375
      summary discussion **3**:382
    payment for ecosystem services (PES) **2**:227
    summary discussion **3**:382
  endangered mammals **3**:194*T*
  endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
  ethnobiological studies **3**:321–322
  fish diversity **3**:473
  fishing/fisheries **6**:422*T*
  indigenous peoples **4**:269*T*, **4**:270*T*
  inland deserts **2**:462*F*
  institutional collections and capacity **1**:505*T*
  intercropping systems **2**:383, **2**:384*F*
  maize production **1**:86
  megaherbivores **5**:224*T*
  nitrogen deposition **5**:526*T*
  parasitic infections **5**:671*T*
  plant species **1**:259*T*, **6**:82*T*
  polyculture cropping systems **7**:231
  population regulation **6**:207
  primate endemism **6**:252*T*
  rice production **1**:86, **1**:88
  species richness **1**:501*F*, **4**:271*T*
  temperate forests
    distribution patterns **7**:116–117
    dominant floristic features **7**:118
  threatened birds **7**:181
  wheat production **1**:86
Chinaberry tree **1**:590
Chinaman leatherjackets **2**:8–9
chinampa agriculture **1**:124
chinch bugs **7**:293, **7**:294
Chinese chestnuts **2**:663
Chinese chestnut trees **3**:411
Chinese chives **2**:663
Chinese dove tree **3**:373–374
Chinese giant salamander **3**:373–374
Chinese grass carp **4**:361, **4**:364
Chinese paddy straw mushrooms **3**:637*T*
Chinese pickling melons **2**:663
Chinese river dolphins **3**:192, **5**:65*T*
Chinese sand pears **2**:663
Chinese subtropical forest **1**:262–263*T*
Chinese white pears **2**:663
Chinese yams **2**:663
Chinook salmon **2**:7–8, **2**:10*F*, **3**:260, **4**:518, **6**:522, **6**:523–524, **6**:525*F*

Chionochloa **4**:349
chipmunks
  boreal forests **1**:631–632
  hibernation **7**:36
Chironectes minimus **4**:687
Chironomidae **3**:489–491*T*, **3**:492, **7**:389–391
Chironomoidea **3**:489–491*T*
Chironomus **3**:580–581, **3**:581*F*
Chiropotes albinasus **6**:256–258*T*
Chiropotes satanas **3**:127*T*, **6**:256–258*T*
Chiropotes utahickae **6**:256–258*T*
Chiroptera **2**:587, **2**:587*F*, **4**:690*T*, **4**:691–692, **6**:309*T*
Chirostenotes **2**:519*T*, **2**:520*T*
chitemene system **6**:555–556
chitons
  calcium carbonate ($CaCO_3$) grazing **3**:612
  classifications **4**:381*T*
  physical characteristics **5**:375–376, **5**:375*F*
  spatiotemporal distributions **5**:373–374, **5**:374*F*
Chixulub crater (Mexico) **3**:418–421
Chlamydomonas **6**:29, **6**:33*F*, **6**:291–292
Chlamydomonas acidophila **3**:584
Chlamydophila pneumoniae **2**:636*F*
Chlamydotherium **4**:727*T*
Chlamyphorus truncatus **2**:587*F*, **3**:193
Chloranthaceae **5**:618*F*, **5**:621*F*
Chlorarachniophyta **6**:10*T*
chlordane **5**:693
Chlorella **5**:306, **6**:29, **6**:33*F*, **6**:295
Chlorella vulgaris **2**:726–727, **4**:255
Chloridoid **2**:376*F*
Chloridops kona **1**:29*F*
chlorine (Cl)
  chlorine monoxide (ClO) **7**:297
  chlorofluorocarbons (CFCs)
    atmospheric concentrations **1**:282*T*
    changing concentrations **1**:284–286, **1**:286*F*
    climate impact **1**:281–282, **3**:450
    ecotoxicology **3**:118–119
    enhanced greenhouse effect **4**:19*T*
    environmental ethics **3**:271
    ozone depletion **7**:296–297
  elemental partitioning **1**:554*T*
  methyl chloride (CH_3Cl) **1**:282*T*, **1**:284–286
  organochlorines **3**:581, **5**:693, **5**:699*T*, **5**:700*T*, **6**:155
  ozone depletion **1**:289
  relative abundance **1**:553*F*
  river water composition **3**:562*T*
chloroacetamide **4**:90*T*, **4**:92*T*
Chlorobi **2**:613–614
Chlorobiaceae **6**:10*T*, **6**:11*T*, **6**:14*T*, **6**:25–26
Chloroflexaceae **6**:10*T*, **6**:11*T*, **6**:14*T*
Chloroflexus **4**:119
Chloroflexus aurantiacus **1**:315, **4**:114
chlorofluorocarbons (CFCs)
  atmospheric concentrations **1**:282*T*
  changing concentrations **1**:284–286, **1**:286*F*
  climate impact **1**:281–282, **3**:450
  ecotoxicology **3**:118–119
  enhanced greenhouse effect **4**:19*T*
  environmental ethics **3**:271
  ozone depletion **7**:296–297
Chlorohydra **5**:306
chloromonads **6**:25
Chlorophyceae **6**:478*T*
chlorophyll production
  light-harvesting pigment–protein complexes **6**:9–11, **6**:10*T*
  marine ecosystems **5**:46
  primary photochemical processes **6**:11–12, **6**:11*T*
  total phosphorus/total nitrogen–chlorophyll *a* relationship **3**:348–350, **3**:349*F*
  water-column nitrogen:phosphorus (N:P) ratios **3**:350*F*
Chlorophyta **3**:571*T*, **6**:10*T*
chloropicrin **5**:694
Chloropidae **3**:489–491*T*, **3**:493–494, **3**:493, **5**:463*T*
chloroplast DNA (cpDNA) **3**:660, **3**:666–667
Chlorostoma spp. **5**:33
chlorothalonil **5**:695
chloroxuron **5**:696
chlorpyrifos **5**:693–694, **5**:699*T*, **5**:700*T*
chlorthalonil **5**:695
chlorthiamid **5**:696
Choanata **7**:338
choanoflagellates **5**:47–48, **6**:293–295, **6**:294*F*
Chobe National Park **5**:226–227, **5**:234–235
chocolate **3**:638–639
Choeropsis liberiensis **4**:696–697
Chokoria Sundarbans **1**:276*T*
cholera **4**:503
Chologaster cornuta **7**:55*F*
Chondria armata **6**:132
Chondrichthyes **3**:458, **3**:460, **3**:681*F*, **3**:685*F*
Chondrodendron roseus **5**:712*T*
Chondrodendron tomentosum **2**:432, **5**:712*T*
Chondrohierax wilsonii **7**:185
Chondromyces crocatus **1**:312–313
Chordata/chordates
  characteristics **7**:336
  classification systems **7**:335, **7**:335*F*, **7**:336*F*, **7**:337*F*
  fossil record **3**:684–685, **3**:684*F*, **3**:688
  global diversity **4**:385*T*
  historical perspective **7**:334–335
  known and estimated species **3**:710*T*, **4**:384*T*
  latitudinal range shifts **4**:604*T*
  major groups
    Amniota **7**:339
    Cephalochordata **3**:457, **7**:336
    Choanata **7**:338
    Chordata **7**:336
    Craniata **7**:336–337
    definitions and diagnoses **7**:335–336
    Gnathostomata **7**:338
    Mammalia **4**:681–707, **7**:341
    Osteichthyes **6**:522–523, **7**:338
    Reptilia **7**:339–340
    Sarcopterygii **7**:338
    Sauria **7**:340–341
    Tetrapoda **7**:337*F*, **7**:338–339
    Vertebrata **7**:337–338
  marine chordate diversity **3**:684–685
  marine ecosystems **5**:50*T*
  number of species **5**:669–670*T*
  Phanerozoic diversity **3**:684–685, **3**:688
  phylogenetic relationships **7**:434*F*
  Proterozoic–Cambrian transition **3**:396*F*
  species estimates **3**:710*T*
chord distance measurement **4**:510
Chordeumatida **5**:423–424, **5**:423*F*
Chordopoxvirinae **5**:136
Choreutoidea **5**:387–388*T*
Choromytilus meridionalis **5**:57–58, **5**:60*F*
Chorthippus parallelus **2**:512, **2**:550, **2**:552*F*
chowchilla **2**:75*F*
Christianity
  biodiversity conservation **6**:564
  Christian declaration on nature **3**:272–273, **6**:381, **6**:381
  Judeo–Christian tradition **6**:381
Christmas Island **1**:262–263*T*, **5**:596–597
Chromatiaceae **6**:10*T*, **6**:11*T*, **6**:14*T*, **6**:25–26
chromatin **3**:657–659
chromoblastomycosis **3**:634*T*
Chromolaena odorata **6**:556
Chrysanthemoides monifilera **3**:211
Chrysanthemum **2**:427
Chrysanthemum cinerariaefolium **5**:693
Chrysemys picta **3**:218
Chrysichthys **4**:463
Chrysidoidea **4**:178
Chrysobalanus **5**:623*F*
Chrysobalanus icaco **4**:286–287
Chrysochlorids **4**:690–691
Chrysochromulina **1**:474–475, **1**:474*F*, **6**:29, **6**:293–294
Chrysocyon brachyurus **4**:693
Chrysomela confluens **4**:173
Chrysomelidae **1**:327*T*, **1**:328
Chrysomeloidea **1**:327*T*
chrysomonads **6**:293–295, **6**:294*F*
Chrysopa scelestes **4**:306–307
Chrysophyceae **3**:571*T*, **6**:10*T*, **6**:478*T*
Chrysophyllum perpulchrum **1**:68
Chryspetalidae **7**:417*F*
Chthamalus stellatus **2**:186, **2**:187*F*, **4**:334
Chthamalus tissus **6**:723*T*
chub mackerel **6**:422*T*
chukar partridge **3**:145–146
chum salmon **6**:522, **6**:523
Chundikkulam Lagoon **1**:276*T*
chungungos **5**:68, **5**:70, **5**:77*T*
Chuquiraga oppositifolia **4**:106
churchyards **4**:579, **4**:580*F*
Chusquea **7**:275, **7**:275–276
Chyromyidae **3**:489–491*T*
Chysocampanula **1**:474*F*
chytrid fungus **1**:519–520, **2**:86–87, **4**:664
chytridiomycetes **6**:296–297
chytridiomycosis **1**:519–520, **2**:527–529, **2**:529–530, **2**:530*F*, **3**:166, **3**:696
Chytridiomycota **2**:644, **3**:627

chytrids **5**:47–48
chytriomycosis **2**:86–87, **4**:664
cicadas **4**:392*T*
Cicconiformes **1**:432
Cicer arietinum **2**:390, **2**:662–663
Cichla ocellaris **6**:233–235
cichlids
  adaptive radiation **1**:24–25, **1**:25*F*, **1**:29–30, **1**:432
  biodiversity issues **3**:444
  characteristics **3**:467–471, **3**:470*F*
  East African lake ecosystems **4**:463–464
  interspecies competition **2**:141
  Near East ecosystems **5**:470
  sympatric speciation **6**:625
cichoric acid **6**:137–138, **6**:138*F*
Cichorium intybus **3**:131–132
Cicindelinae **1**:250, **1**:250*T*, **1**:251*F*, **1**:252*F*
Cicindellidae **1**:709–710
Ciconiiformes **1**:606–607, **6**:308*T*
ciguatoxin **6**:131, **6**:132*F*
Cilacap **1**:276*T*
ciliates **5**:568, **5**:569*F*, **6**:289*T*, **6**:290*F*, **6**:295–296, **6**:296*F*, **6**:586, **6**:587*F*
ciliophorans **5**:669–670*T*, **6**:25
Ciliophrys **6**:293
Cimex hemipterus **7**:294
Cimex lectularius **7**:287–288, **7**:294
Cimicidae **7**:294
Cimicomorpha **7**:293
Cinchona calycina **1**:166–167
Cinchona pubescens **5**:712*T*
Cingulata **2**:586, **2**:587*F*, **4**:689–690
cinnabar moth **6**:159*F*
Cinnamomum **1**:270–271
Cinnamomum camphora **5**:712*T*
Circaeasteraceae **1**:260*T*
Circaetus gallicus **5**:474
Circoviridae **5**:134*T*, **5**:136
CIRCUITSCAPE (software program) **4**:512
circuit theory-based models **4**:512
Circus aeruginosus **5**:474
Circus maurus **1**:487–488
Ciridops anna **1**:29*F*
Cirratulidae **7**:417*F*
Cirripedia **2**:401*T*
Cirsium arvense **1**:75
Cirsium hydrophilum **4**:361, **4**:363
Cirsium vulgare **3**:703
ciscoes **3**:464–465
Cis dentatus **7**:220*T*
Cissus gongyloides **4**:289*T*
Cisticola carruthersi **1**:53–54
Cistus **4**:72, **5**:218
Citedolor strepera **3**:365–366
cities
  biodiversity commodification **4**:142
  demographic transition **4**:231–232
  dependence interactions
    biodiversity **4**:237
    ecosystem services **4**:237–238, **4**:238*F*
    general discussion **4**:237
  ecological footprint analysis (EFA) **2**:707
  environmental impact
    ecosystem services **4**:235
    general discussion **4**:234–235
    habitat loss and fragmentation **4**:234–235
    positive impacts **4**:236–237
    protected area systems **4**:235–236, **4**:235*F*
  geographic distribution **4**:235*F*
  human-dominated techno-ecosystem **3**:60–63
  management strategies
    ecosystem services **4**:239*F*, **4**:241–242
    general discussion **4**:238–241
    rural-to-urban migration **4**:240–241*F*, **4**:242
    sustainable development **4**:239–241, **4**:239*F*
  population trends **4**:231, **4**:231*F*
  rural-to-urban migration **4**:232–234, **4**:233*F*
  size distribution **4**:561*F*
  world-at-night map **4**:561*F*
Citisus scoparius **6**:443
Citrullus lanatus **2**:662–663
citrus crops **2**:383, **2**:384*F*, **2**:663, **3**:32*F*
civets **1**:686, **4**:693
clades, definition of **2**:47
cladistics **2:33–45**
  basic concepts
    character types **2**:33*F*, **2**:34–35, **7**:104–106, **7**:105*F*
    cladistic systematization **7**:104–106, **7**:104*T*, **7**:105*F*, **7**:105*T*
    cladograms **2**:33*F*, **2**:36, **2**:36*F*, **2**:52–53, **2**:52*F*, **2**:53*F*, **7**:91–93
    Hennigian terminology **7**:91–93
    parsimony **2**:33*F*, **2**:35–36, **2**:35*F*, **6**:17
    phylogenetic trees **2**:33*F*, **2**:34*F*, **2**:36, **2**:36*F*, **2**:52*F*, **2**:53*F*, **7**:91–93
    relationships **2**:33–35, **2**:33*F*, **2**:34*F*, **2**:52–53, **2**:52*F*, **2**:53*F*
  characters
    character coding **2**:38–39, **2**:39*T*
    character recognition **2**:37
    character types **2**:33*F*, **2**:34–35, **2**:37, **7**:104–106, **7**:105*F*
    homology **2**:37
    phylogenetic evidence **2**:37–38, **2**:38*F*, **2**:38*T*
    research background **2**:36–37
  cladistic analysis
    biogeography **7**:344–345, **7**:344*F*
    character polarization **2**:39–40
    cladogram construction **2**:39, **7**:108, **7**:108*F*, **7**:109*T*
    cladogram rooting **2**:39–40
    ontogenetic criteria **2**:40
    optimization processes **2**:40–41, **2**:41*F*
    outgroup comparisons **2**:40
    progression rule **7**:344*F*, **7**:345–346, **7**:346*F*, **7**:347*F*
  cladistic biogeography
    basic concepts **7**:344–345, **7**:344*F*
    component analysis **7**:348–349, **7**:349, **7**:349*F*, **7**:350*F*, **7**:350*T*
    dispersal biogeography **2**:555
    Heterandria
      cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
      endemic analysis **7**:349–351, **7**:350*F*
      geographical distribution **7**:348*F*
    progression rule **7**:345–346, **7**:346*F*, **7**:347*F*
    research summary **7**:351
    Xiphophorus
      cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
      endemic analysis **7**:349–351, **7**:350*F*
      geographical distribution **7**:348*F*
  cladogram evaluation
    biogeographical analysis **2**:555, **7**:344–345, **7**:344*F*
    character fit **2**:35*F*, **2**:41–42
    character weighting **2**:35*F*, **2**:42
    component analysis **7**:348–349, **7**:349, **7**:349*F*, **7**:350*F*, **7**:350*T*
    consensus cladograms/trees **2**:42–44, **2**:43*F*
    consistency index (CI) **2**:41–42
    group support **2**:44
    lineage hierarchy **7**:108, **7**:108*F*, **7**:109*T*
    rescaled consistency index **2**:42
    Southern Hemisphere ecosystems **6**:614
    systematics **7**:87
  definition **2**:33–34
  historical background **7**:104–106
  partitioned analysis **2**:44–45, **2**:45*F*
  simultaneous analysis **2**:44–45, **2**:45*F*
  systematics **7**:87
Cladium jamaicense **7**:372, **7**:385, **7**:399
Cladocera
  eutrophication impacts **3**:362
  extinction vulnerability **2**:416–417*T*
  lake and pond ecosystems **4**:371–373, **4**:374*F*, **4**:459
  number of species **3**:182*T*, **4**:371*T*
  wetland ecosystems **7**:389–391
cladogenesis **2:46–57**
  basic concepts **2**:47–48, **2**:48*F*, **6**:620, **6**:621*F*
  definition **2**:47
  diversification
    evolutionary theories **2**:46–47
    speciation **2**:47
    synthetic theory **2**:47
  evolutionary rates
    fossil record **2**:54–55
    molecular evolution
      deoxyribonucleic acid (DNA) **2**:56–57, **2**:56*F*
      hemoglobin **2**:55–56, **2**:55*F*, **2**:55*T*
    morphological evolution **2**:54–55, **2**:54*F*
  importance **2**:56–57
  patterns of life **2**:53–54
  systematic classifications
    branching diagrams **2**:49–51, **2**:49*F*, **2**:50*F*, **2**:51*F*, **2**:52*F*, **7**:87, **7**:88*F*
    cladograms **2**:52–53, **2**:52*F*, **2**:53*F*, **7**:91–93
    evolutionary systematics **2**:51–52, **2**:52*F*, **7**:101–102, **7**:102*F*
    grades **2**:51–52, **2**:52*F*

cladogenesis (*continued*)
Hennigian terminology **7**:91–93
monophyly **2**:52–53, **7**:91–93, **7**:92*F*
natural classification systems **2**:49–51, **2**:49*F*, **7**:101–102
paraphyly **2**:52–53, **7**:91–93, **7**:92*F*
research background **2**:48–51
Cladonia pyxidata **1**:236*F*
Cladophora **3**:358
Cladophora fuliginosa **3**:359
Cladophora glomerata **3**:358, **3**:369, **6**:478
Cladosiphon occidentalis **3**:359
Cladosporium **3**:634*T*
clams
chemosynthetic symbioses **5**:120–121
critical pH levels **6**:153*F*
discovered species **2**:5*F*
ecological harvesting practices **7**:228
number of species **3**:182*T*, **4**:371*T*
sandy beaches **2**:131–133
subterranean ecosystems **7**:57
surface water acidification **1**:10*T*
wetland ecosystems **7**:389–391
clam shrimp **3**:182*T*
Clangula hyemalis **1**:236, **2**:549
Claravis godefrida **7**:183
Clark, George Rogers **4**:651–652
Clarkia cocinna **6**:48
Clarkia tembloriensis **4**:248*F*
Clark's nutcracker **2**:571, **6**:51–52
Clarkus **7**:422–424, **7**:426*F*, **7**:429–430
classification and regression tree (CART) analysis **4**:514
Clausocalanus paupulus **4**:386*T*
Claviceps purpurea **3**:633, **3**:638
Clavicipitaceae **2**:146
clay soils **4**:225–226, **4**:225*F*
Clean Air Act (1970) **2**:286
Clean Air Status and Trends Network (CASTNet) **1**:1–5, **5**:533
Clean Water Act (1972) **1**:95–97, **2**:286, **3**:718, **6**:497
Clematis **5**:218
Clements–Gleason debate **6**:429
Clementsian theory
ecological succession paradigm
application
climax community **7**:64
community-level management **7**:64
indicator species **7**:64
progressive nature concept **7**:64
basic principles **7**:63–64
comparison studies **7**:65–66, **7**:65*T*
ethnobiology **3**:324, **3**:325*T*
Clemmys guttata **7**:390
Clemmys insculpta **7**:390
Clemmys marmorata **4**:564
Cleome **2**:376*F*
Cleomeaceae **2**:374*T*
Cleridae **1**:327*T*
Clerodendrum **6**:316*T*
Cleroidea **1**:327*T*
Clethrionomys rutilus **1**:631–632
Clibanarius **5**:16
Cliffortia **5**:219
cliff swallows **6**:571–573
Clifton Beach **1**:276*T*
Climacodium frauenfeldianum **2**:612–613
Climacteridae **4**:33, **4**:34*F*
climate change **2:58–72**
Africa **1**:130*F*
alpine ecosystems **1**:156–157, **1**:156*F*
amphibian species decline
elevated ultraviolet (UV–B) radiation **3**:697
general discussion **3**:696–697
temperature variability–disease immunity relationship **3**:697
weather pattern shift effects **3**:697
Antarctic ecosystems **1**:186–187, **6**:617
Anthropocene epoch **4**:146
Arctic ecosystems **1**:241–242
atmospheric gases **1**:281–282
background information **3**:407–408
biodiversity impacts **7**:203*F*
anthropogenic effects **4**:24
biota types
at-risk organisms **4**:29
organisms likely to benefit **4**:29–30
community impacts **4**:30
current climate change evidence **4**:30–31, **4**:31
fish species **3**:449–450
global biodiversity **3**:689
research methodologies
integrated research **4**:26–27
manipulation experiments **4**:25–26, **4**:26*F*
natural climate variability **4**:25
paleobiological data **4**:25
species distribution models **4**:24–25, **4**:24*T*
research summary **4**:31
response types
adjustment **4**:27
evolution **4**:27
extinction **4**:29
migration **4**:27–29, **4**:28*F*
Southeast Asia **1**:517
biogeochemical feedbacks
carbon cycle **6**:147–149
characteristics **4**:23
ecosystem alterations
climate change effects **6**:149–150
coastal/marine ecosystems **6**:150–151
forest ecosystems **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
marine feedbacks **4**:23
terrestrial feedbacks **4**:23–24
botanical gardens
botanical and ecological research **6**:515–516
ex situ conservation programs **6**:516
germplasm utilization **6**:516
in situ conservation programs **6**:516
public awareness and educational programs **6**:515
research challenges **6**:515
C_4 plants **2**:378
causal factors **2**:60
Central America **2**:30
climate change predictions **3**:408, **3**:408*F*
climate models
dynamical downscaling techniques **2**:89–90
general circulation models (GCMs) **2**:88, **6**:148–149
regional predictions **2**:88–89
statistical downscaling techniques **2**:89
validation methods **2**:88
contemporary conservation theories and practices **2**:260–261
Convention on Biological Diversity (1930) **2**:318
coral reef ecosystems **3**:449–450, **4**:664
deforestation **2**:454–455
desertification **2**:486–487
diversity effects **3**:165
ecological changes and transitions **2**:59–60, **2**:61*F*
ecological models
animal responses **2**:92–94, **2**:93*F*
bottom-up approaches **2**:66–68, **2**:90
combined approaches **2**:68, **2**:90–91
community ecology **2**:94
extinction probability **2**:94–95
future climate projections **2**:90
refined approaches **2**:91
strategic cyclical scaling approach **2**:68–69, **2**:91
three-way linkages **2**:94
top-down approaches **2**:66, **2**:90
vegetation responses **2**:91–92
ecological responses
abundance shifts **2**:85
animal responses **2**:66, **2**:67*F*, **2**:93*F*
animal–vegetation–climate linkage **2**:69
bottom-up research analyses **2**:66–68, **2**:90
coastal/marine ecosystems **6**:150–151
combined analytical approaches **2**:68, **2**:90–91
distribution shifts **2**:85
ecosystem alterations **6**:149–150
extinction rates **2**:86–87, **4**:664
forest ecosystem alterations **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
future research areas **2**:69
general discussion **2**:65–66
phenological shifts **2**:85–86
physical changes **2**:86
strategic cyclical scaling approach **2**:68–69, **2**:91
top-down research analyses **2**:66, **2**:90
vegetation responses **2**:65–66, **2**:91–92
ecological restoration **4**:203
endangered species
freshwater invertebrates **3**:181–182
reptiles **3**:218, **6**:410
environmental impacts **3**:282–283, **7**:203*F*
evolutionary response **3:385–391**
climate complexity **3**:385
environmental versus evolutionary response **3**:386
evolutionary complexity **3**:385
evolutionary genetics

environmental versus evolutionary response **3**:386
evolutionary rates **3**:386–387
general discussion **3**:386
influencing factors **3**:387
natural selection **3**:386
plasticity adaptations **3**:387–388
general discussion **3**:385
population adaptations **3**:385–386
research approaches
climate variables **3**:388
experimental manipulation **3**:389
future predictions **3**:389
geographic variations **3**:388
long-term studies **3**:388
past climate change **3**:388
population dynamics manipulation **3**:389
range limits **3**:388
resurrection ecology **3**:388–389
theoretical perspectives **3**:389–390
research summary **3**:390
extinctions **2:73–78**
background information **3**:407–408
biodiversity loss
communities and ecosystems **2**:76
coral reef ecosystems **3**:449–450, **4**:664
geographic and landscape scale variation **2**:74–75
global warming **4**:664
individuals and populations **2**:75–76, **2**:75*F*
species interaction risks **2**:76–77
temporal match–mismatch (TMM) hypothesis **2**:76
climate change predictions **3**:408, **3**:408*F*
coral reef ecosystems **3**:449–450, **4**:664
current trends **2**:73
global impact **2**:73
habitat loss and fragmentation **3**:408
research summary **2**:77
risk factors
extrinsic risks **2**:74, **2**:75*F*
intrinsic risks **2**:73–74, **2**:74*F*
extreme events **4**:212–213
fire/fire regimes **3**:441
forest canopies **3**:524
forest ecosystems **3**:535, **6**:149–150, **6**:150*T*
future research areas
abrupt climate changes **2**:69–70
adaptability studies **2**:70
ecological applications-driven research **2**:70
interactive/multiscale models **2**:70
transient scenarios **2**:69
global agriculture impacts **3**:427–429, **3**:428*F*
greenhouse effect **2**:58–59
habitat alteration **4**:155
habitat loss and fragmentation **2**:58–59, **2**:80, **3**:221, **3**:408, **4**:56, **6**:62
harmful algal blooms **1**:479–481
historical perspective **2**:82–83
historic land-use patterns **4**:569–570
ice core studies **2**:59–60
indicator species
species distribution changes **4**:257–258, **4**:259*F*
species sensitivity **4**:256–257, **4**:258*F*
infectious diseases **2**:531–532
international organizations **4**:326
marine ecosystems **5:32–44**
abundance impacts **5**:34–35
community impacts **5**:37–38
conservation strategies
assisted colonization **5**:41
ecosystem-based management **5**:40
fisheries management **5**:40–41
goals and priorities **5**:38–39
marine protected areas **5**:39–40
predictive models **5**:39
restoration **5**:41
distribution impacts **5**:35–36
evolutionary impacts
adaptation rates **5**:36
climate tolerance **5**:36–37
general discussion **5**:36
genetic diversity **5**:36–37
local adaptations **5**:36
future research areas **5**:41
global impacts **3**:449, **5**:32
physiological impacts
ocean acidification **5**:33–34
rising seawater temperatures **5**:33
research needs **5**:41
marine mammal extinctions **5**:89–91
marine protected areas **5**:100–101
marine sedimentary habitats **5**:113*F*, **5**:114–115
migration impacts **5**:318
neotropical seasonally dry forests **5**:497–498
paleotemperature assessments **4**:224–226, **4**:225*F*
phenological shifts **5:716–727**
basic concepts **5**:716–717
consequences
microevolution **5**:724–725
population viability **5**:723–724, **5**:724*F*
future outlook **5**:725–726
methodological challenges **5**:717
phenological mismatches
life-history stages **5**:723
trophic levels **5**:722–723, **5**:722*F*
research background **5**:716
research observations
between-population variations **5**:720
between-species variations **5**:718–720
general discussion **5**:717–718
linked observed shifts **5**:720–721
literature reviews **5**:717–718, **5**:719*F*
migration timing **5**:721
reproduction timing **5**:721
research summary **5**:722
spring phenological shifts **5**:718*F*
winter rest onset and termination timing **5**:721–722
plant biodiversity threats **6**:62
plant conservation **6**:83–84
predictive approaches
future research areas **2**:69
global warming forecasts **2**:62
model validation and reliability **2**:62–63
relevance
empirical mapping techniques **2**:64–65
global circulation models (GCMs) **2**:63–64
regional changes **2**:64
regional scale models **2**:65
societal forcing estimations **2**:60–62
protected area systems
challenges **5**:449
Great Limpopo Transfrontier Park **5**:449
Seamounts Marine Management Area **5**:449–450
range shifts **4:599–611**
basic concepts **4**:599
conservation management **4**:607–608
ecological consequences **4**:606–607
ecological drivers **4**:599–600
evidentiary research
elevational range shifts **4**:603–604, **4**:605*T*
global trends **4**:602–603
latitudinal range shifts **4**:603, **4**:604*T*
future challenges **4**:608
geographic dimensions **4**:599–600, **4**:600*F*
range shift types
crash-range shifts **4**:601*F*, **4**:602
lean-range shifts **4**:601*F*, **4**:602
march-range shifts **4**:601*F*, **4**:602
shift mechanisms
climate change-based directional range shifts **4**:600–602
population dynamics **4**:600, **4**:601*F*
range core versus range periphery **4**:600, **4**:601*F*
range shift types **4**:602
time lag response **4**:604–606, **4**:606*F*
riparian zones **6**:467*T*
salmon population **6**:529
savannas **1**:71–72
South America **6**:609
Southeast Asia **1**:517
Southern Hemisphere ecosystems **6**:617
species distribution modeling (SDM) **6**:698–699
species loss **5**:597
Succulent Karoo **1**:489
synergistic effects **2**:58–59
terrestrial impact **2:100–114**
adaptation actions
adaptive management **2**:107
basic concepts **2**:105–106
cooperation and coordination efforts **2**:106
general discussion **2**:105–106
management strategies **2**:106–107
resistance and resilience strategies **2**:106
spatiotemporal scale factors **2**:106

climate change (*continued*)
triage approach **2**:107
vulnerability factors **2**:105–106
adaptation strategies
assisted colonization **2**:109
challenges **2**:109–110
characteristics **2**:107–108
connectivity **2**:108
examples **2**:110*T*
reserve selection and design **2**:107–108
species- and place-specific approaches **2**:108–109
co-benefits **2**:110–111
diseases and parasites **2**:103–104, **2**:110*T*
future outlook **2**:100
interspecific interactions **2**:102–103, **2**:110*T*
multidimensional strategies **2**:110–111, **2**:110*T*
phenological shifts **2**:100–101, **2**:110*T*
population dynamics **2**:102, **2**:110*T*
range shifts **2**:101–102, **2**:110*T*
stressor interactions
adaptation strategies **2**:110*T*
anthropogenic responses **2**:105
ecosystem services **2**:105
fire disturbances **2**:104–105
invasive species **2**:104
land-use changes **2**:104
terrestrial invertebrates **3**:221
threatened birds **7**:185–186
wild species **2:79–99**
climate models
dynamical downscaling techniques **2**:89–90
general circulation models (GCMs) **2**:88
regional predictions **2**:88–89
statistical downscaling techniques **2**:89
validation methods **2**:88
current trends **2**:83–84
ecological models
animal responses **2**:92–94, **2**:93*F*
bottom-up approaches **2**:66–68, **2**:90
combined approaches **2**:68, **2**:90–91
community ecology **2**:94
extinction probability **2**:94–95
future climate projections **2**:90
refined approaches **2**:91
strategic cyclical scaling approach **2**:68–69, **2**:91
three-way linkages **2**:94
top-down approaches **2**:66, **2**:90
vegetation responses **2**:91–92
ecological responses
abundance shifts **2**:85
distribution shifts **2**:85
extinction rates **2**:86–87, **4**:664
fish **3**:449–450
phenological shifts **2**:85–86
physical changes **2**:86
extreme events **2**:84–85
future research areas
applications-driven climatic research **2**:96
ecological impact assessments **2**:95
ecosystem adaptability **2**:95–96
multiscale interactive analysis **2**:96
historical perspective **2**:82–83
importance **2**:80–81
joint atribution theory **2**:87
meta-analyses **2**:87
synergistic effects **2**:81–82
see also global warming; greenhouse effect
Climate Change Science Program (CCSP) **2**:106
climatic regions **1**:581*F*, **1**:582*F*
climax community **7**:64
climbers **6**:315*T*, **6**:316*T*, **6**:318*T*
climbing everlasting daisy **4**:194*T*
climbing ferns **4**:360
Clinostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Clitellata
classifications **7**:408
geographic distribution **7**:418
morphology
anterior ends **7**:410*F*, **7**:412*F*
chaetae **7**:409–410, **7**:414*F*
circulatory system **7**:415
digestive tract **7**:415
general characteristics **7**:408–409
gills **7**:415
gonads **7**:415–416
head **7**:409
musculature **7**:414–415
nephridia **7**:415
nervous system **7**:410–411
posterior ends **7**:410*F*
pygidium **7**:409
segments **7**:409, **7**:413*F*
sense organs **7**:411–414
septa **7**:415
phylogenetic and taxonomic relationships **7**:416–419, **7**:417*F*
reproductive biology **7**:416
Clitocybe dealbata **3**:635*T*
Clitoria triquetrum **1**:165
Cloeon dipterum **5**:471
clomazone **4**:90*T*
clone libraries **6**:349–350, **6**:349*T*
Clonorchis **7**:450–451*F*
closed basins **3**:560
closed forests **1**:292–297, **1**:294*T*, **1**:629–630, **6**:325–327, **6**:326*F*, **6**:326*T*
Closteroviridae **5**:130*F*
Clostridium **6**:32
clothes moths **5**:401–402
cloud ear mushrooms **3**:637*T*
clouded leopard **1**:517
cloud forests
Central America **2**:28–29
South America **3**:111
cloud formation **4**:22–23
clover **1**:114
cloves **2**:663
club fungi **2**:644
club moss **6**:58*T*, **6**:59
Clupea harengus **5**:315
Clupea pallasi **3**:256
clupeomorphs **3**:463–464
Clusia **3**:522*T*, **7**:275–276
Clusiaceae **3**:129*T*, **3**:522*T*, **5**:621–625
Clusiella **3**:522*T*
Clusiidae **3**:489–491*T*
Clusius, Carolus **6**:508
clutch size **5**:676–677
Clytrinae **1**:329
Cnemidophorus **6**:406
Cnemiornis calcitrans **5**:366
Cnidaria
Census of Marine Life project **2**:7*F*
endangered species **3**:202*T*
fossil record **3**:684*F*
global diversity **4**:385*T*
known and estimated species **4**:384*T*
marine ecosystems **5**:50*T*
number of species **4**:371*T*
photosynthetic symbioses **5**:118–120, **5**:119*F*
phylogenetic relationships **7**:434*F*
plankton fraction **6**:25
Proterozoic–Cambrian transition **3**:396*F*
coacervates **5**:630
coach-whip snakes **4**:494, **6**:406, **6**:409
coal
biodiversity impacts **3**:257–259, **3**:264*T*
carbon dioxide (CO_2) emissions **3**:265*T*
land requirements **3**:265*T*
world energy reserves **3**:265*T*
world energy use **3**:251*F*, **3**:252*T*
Coale, Ansley **6**:205–206
coal tits **1**:611–612
coarse particulate organic matter (CPOM) **6**:471, **6**:472*F*, **7**:51–52
coarse woody debris (CWD) **1**:450
coastal beach ecosystems **2:128–136**
sandy beaches
alongshore distribution **2**:131–133
background information **2**:128
beach index (BI) **2**:129
future research areas **2**:135
macrofauna surveys **2**:129–131, **2**:130*F*
macroscale patterns
latitudinal trends **2**:131, **2**:131*F*
morphodynamic effects **2**:131, **2**:132*F*, **2**:133*F*
mesoscale patterns **2**:131–133
microscale patterns **2**:133
microtidal regimes **2**:128–129, **2**:129*F*
morphodynamic states **2**:128–129, **2**:129*F*
research summary **2**:135
species richness hypotheses **2**:133–134
threats and challenges **2**:134–135
tidal regimes **2**:129
two-dimensional model **2**:129, **2**:130*F*
zonation patterns **2**:131–133
Coastal Brazilian Shield **3**:105–107, **3**:106*F*
coastal deserts **2**:458–460, **2**:462*F*, **2**:463*F*
coastal dune vegetation **1**:292–297
coastal ecosystems
Asia **1**:275–279, **1**:275*T*, **1**:276*T*
Australian plant communities **1**:292–297

biodiversity restoration **6**:456–457
Census of Marine Life project **2**:15*F*, **2**:16, **2**:17*F*
climate change effects **6**:150–151
coastal wetland restoration **7**:375–378
conservation status **3**:171–173*T*, **3**:174
ecosystem services **1**:334*T*, **3**:96–98
global extent **7**:370*T*
myriapods **5**:430–431
nature-based coastal protection strategies **5:451–458**
biodiversity impacts **5**:453–456, **5**:455*F*
current status **5**:452–453, **5**:453*T*
historical perspective **5**:452
importance **5**:451–452
research summary **5**:456
nitrogen cycle impacts **5**:543
predation effects **6**:237–239
salt water plants **6**:77
tropical Africa **1**:55
see also wetlands
Coastal Resilience model **5**:460–462*T*
Coastal tailed frog **4**:517
Coastal Zone Management Authority and Institute (CZMAI) **5**:465
coatis **2**:590*T*
Cobb, John B., Jr. **6**:382
cobia **1**:190*T*
cobweb spiders **5**:460*T*
coca **2**:663–664, **4**:287–290
cocaine **2**:426*T*, **2**:427, **5**:712*T*, **6**:136, **6**:136*F*
Coccidioides immitis **3**:634*T*
coccidioidomycosis **3**:634*T*
Coccinella novemnotata **3**:222
Coccinellidae **1**:327*T*
Coccochorella quercifolia **4**:174
coccolithophorids **5**:683–686, **5**:684*F*, **6**:25
Coccoloba **5**:489
Coccus viridis **4**:365
Cocha Cashu Biological Station (Peru) **1**:159*T*, **6**:239–240
Cochliobolus heterostrophus **1**:460
Cochlodinium polykrikoides **1**:477*T*
cockles **1**:190*T*
cockroaches
insecticide resistance **4**:298–299
number of species **4**:392*T*
cocoa **2**:382–383, **2**:383, **3**:32*F*, **5**:605, **6**:554
cocoa beans **3**:638–639
coconut **2**:663, **3**:32*F*, **7**:26
coconut case caterpillar **5**:603–604
coconut crab **3**:220, **3**:222–223, **3**:223*F*, **3**:226
coconut palm borer **5**:603–604
coconut palms **1**:277
Cocos finches **1**:27–28, **1**:28*F*, **1**:618–619, **2**:141
Cocos–Keeling Island **1**:262–263*T*, **5**:449–450, **6**:507*T*
Cocos nucifera **1**:277, **7**:272
Cocus mucifera **2**:663
cod **3**:123*T*, **3**:466–467, **5**:315, **5**:690, **7**:262
codeine **2**:432, **5**:712*T*
coelacanths
characteristics **3**:461, **7**:338
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
speciation **2**:47
Coelastrum **6**:30
Coelenterata **4**:381*T*, **5**:20*T*
Coelodonta antiquitatis **4**:729–730, **4**:729*F*, **5**:225*T*
Coelolepida **5**:385–386, **5**:386*F*, **5**:387–388*T*
Coelopidae **3**:489–491*T*
Coeloptera **3**:182*T*
Coeloscleritophora **5**:374
Coenocorypha pusilla **7**:188
coevolution **2:137–148**
adaptation rates **2**:139–140
character displacement **6**:720
classifications
coincident speciation **2**:137–138
congruent and incongruent phylogenies **2**:137–138, **2**:138*F*
cospeciation **2**:137
escape-and-radiate coevolution **2**:137
reciprocal adaptive responses **2**:137
coevolutionary response **1**:23
consequences
community structure **2**:147–148
phenotypic diversity **2**:146–147
species diversity **2**:147
definition **2**:137
ecological perspectives **3**:7
enemy–victim interactions **2**:142–143
evidentiary sources **2**:138–139
interspecific competition **2**:140–141, **6**:720
mutualism **2**:146, **6**:720–721
natural selection **6**:719–720
parasite–host coevolution
gene-for-gene systems **2**:143–144
origins **5**:663
resistance mechanisms **2**:144–145
sexual selection **2**:144
specialist versus generalist herbivores **2**:145–146
tolerance mechanisms **2**:145
virulence–avirulence evolution **2**:144
plant–animal interactions **6**:42–43
prey–predator interactions
divergent defenses **2**:141
escape space **2**:141
mimicry **2**:141–142
species interactions **6**:721
research summary **6**:724
Coffea arabica **2**:427, **2**:663, **7**:26, **7**:274–275
Coffea canephora **5**:605
coffee
agroforestry **2**:382–383
intercropping patterns **2**:383
low mountain ecosystems **7**:274–275
regional cultivars **3**:32*F*
seed protection **7**:26
species richness **5**:605
Sub-Saharan Africa **2**:663
coffee berry borer **4**:304–305
coffee plantations
multiple cropping systems **1**:120
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
unplanned diversity **3**:12
cognition **1**:43
cognitive maps **3**:324–325
cogon grass **4**:358–359
cohesion species concept **6**:685
coho salmon **6**:522, **6**:523–524
Coimbra-Filho's titi **6**:256–258*T*
coincident speciation **2**:137–138
Coix lachryma-jobi **2**:663
cola drinks **3**:638–639
Cola gigantea **1**:68
colchicine **2**:432, **5**:712*T*
Colchicum autumnale **5**:712*T*, **6**:130–131
cold-desert shrublands
characteristics **6**:333–334
protected area systems **5**:444*T*
cold seeps
Census of Marine Life project **2**:15*F*, **2**:19–20
marine invertebrates **4**:391
cold-winter deserts **5**:444*T*
Cole, Lamont **4**:631
Coleomegilla maculata **5**:316
Coleophoridae **5**:390
Coleoptera
alpine ecosystems **1**:155*F*
Amazon ecosystems **1**:247
Antarctica **1**:184*T*
characteristics **4**:392
freshwater ecosystems **4**:374
insecticide resistance **4**:298–299
major divisions and characteristics **4**:308–312, **4**:309*T*
number of species **4**:371*T*, **4**:392*T*
pollination **6**:140–141, **6**:141*T*
riparian zones **6**:462–463
species and genera **1**:326
Coleorrhyncha **7**:287
Coleps **6**:25–26, **6**:31, **6**:296*F*
Coleridge, Samuel Taylor **4**:649
Coleus forskohlii **3**:209–210, **3**:210*F*
Coliadinae **1**:651
Colias eurytheme **2**:725
Colias meadii **2**:725
Coliiformes **1**:608, **7**:181–182
collared doves **2**:551–552*T*, **5**:476
collared peccary **2**:590*T*
collector curves **6**:349–350, **6**:350*F*
Collembola
Antarctic ecosystems **1**:184*T*
Arctic ecosystems **1**:231*T*
biomonitoring processes **3**:122*T*
characteristics **4**:393*F*, **6**:586
major divisions and characteristics **4**:308–312, **4**:309*T*
number of species **4**:392*T*, **4**:393*F*
soil fauna **6**:586
subterranean ecosystems **7**:55, **7**:55*F*
Collisella **5**:59*F*
Collomia biflora **4**:106
Collomia rawsoniana **3**:208
Collozoum longiforme **6**:292
Colluricincla boweri **2**:75*F*
Colobanthus quitensis **1**:182, **4**:107

Colobinae **4**:692
colobines **6**:253–254
Colobus angolensis prigoginei **6**:256–258*T*
Colobus guereza **6**:254
Colobus guereza percivali **6**:256–258*T*
colobus monkeys **6**:251*T*
Colobus polykomos **7**:229*T*
Colobus satanas satanas **6**:256–258*T*
Colocasia esculenta **2**:663, **6**:553
Colombia
  biodiversity richness **1**:501*F*
  botanical gardens **6**:519
  coral reef ecosystems **1**:500*T*
  deforestation **1**:502*T*
  ecosystem services **3**:100, **7**:356*T*
  education level–fertility relationship **6**:204, **6**:205*F*
  endangered mammals **3**:194*T*
  endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
  indigenous peoples **4**:270*T*
  institutional collections and capacity **1**:505*T*
  Instituto de Investigación de Recursos Biológicos Alexander von Humboldt **1**:507
  land cover classifications **6**:607*T*
  neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
  nonvascular epiphytes **3**:521*T*
  oil-palm plantations **5**:602
  plant species **4**:102*T*
  primate diversity **6**:252*T*
  primate endemism **6**:252*T*
  species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
  threatened birds **7**:181
  tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*, **7**:269–270, **7**:270*T*
  tropical rain forests **6**:314–315, **6**:317*T*
  vascular epiphytes **3**:520*T*
  water funds **7**:356*T*, **7**:362–363
Colombian woolly monkey **6**:256–258*T*
Colophospermum mopane **1**:51, **1**:67
Colorado beetle **4**:307
Colorado Plateau **5**:563
Colorado squawfish **3**:260
colpodid ciliates **6**:295–296, **6**:296*F*
colugos **4**:690*T*, **4**:691
Columba livia **1**:78, **1**:624
Columba livia domestica **5**:476
Columba palumbus **2**:541*F*
Columbian exchange **4**:143
Columbia spotted frogs **4**:516, **4**:518
Columbidae **6**:308*T*
Columbiformes **6**:308*T*
columbine monkeys **4**:692
Colydiidae **1**:327*T*
coma **3**:635*T*
Comatricha typhoides **3**:631*F*
combined zoological/botanical gardens **6**:505
comb jellies **6**:25
Combretaceae **1**:267, **5**:11*T*
Combretum **5**:489
Combretum apiculatum **1**:67
Combretum collinum **1**:68
Combretum leprosum **1**:161
Combretum molle **1**:68
Comicus capensis **3**:730*F*
Comisión Nacional Para el Conocimiento y Uso de la Biodiversidad (CONABIO) **1**:505–506
Commander Islands **4**:723*F*
Commelina benghalensis **6**:557
commensalism **5**:656, **6**:715*T*, **6**:719
commensal pea crab **3**:201–202
commercial farming industry **4**:144
commercial harvesting **6**:444–445
Commidendrum robustum **4**:361
Commiphora **5**:235
Commiphora africana **1**:67
commiphoras **2**:468, **2**:471
Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR) **6**:427
Commission on Genetic Resources for Food and Agriculture (CGRFA) **4**:328
Committee on the Preservation of Natural Conditions within the Ecological Society of America **2**:284
Committee on the Status of Endangered Wildlife in Canada (COSEWIC) **7**:197–199
common ancestry theory **3**:395–396, **3**:397*F*
common beans **1**:132*T*, **2**:663–664
common frog **3**:166
common myna **1**:77–78
common palm civit cat **1**:517
common prickly pear **4**:350–351
common property **6**:274, **6**:274*T*
commons **2:149–154**
  basic concepts **2**:149–150
  biological reserves **2**:154
  definition **2**:149
  externalities **2**:154
  free riders **2**:154
  Gordon model of bionomic equilibrium **2**:150–151, **2**:150*F*, **2**:151*F*
  governance systems **2:155–166**
    advantages/disadvantages **2**:163–164
    centralized regimes **2**:156–157
    challenges **2**:155–156
    complex adaptive systems
      building blocks **2**:161–162
      internal models **2**:161
      parallel adaptive systems **2**:163–164
      polycentric systems **2**:164–165
      properties and mechanisms **2**:161
      rule configurations and situations **2**:162–163, **2**:162*T*
      self-organized resource governance systems **2**:160–161
      tags **2**:161
    local governance systems **2**:163–164
    polycentric systems **2**:164–165
    rules as tools
      aggregation rules **2**:159–160, **2**:162*T*
      boundary rules **2**:157–158, **2**:162*T*
      choice rules **2**:158–159, **2**:162*T*
      information rules **2**:159–160, **2**:162*T*
      payoff rules **2**:159, **2**:162*T*
      position rules **2**:159, **2**:162*T*
      rule change experimentation **2**:160
      rule characteristics **2**:157–158, **2**:162*T*
      rule improvements **2**:160
      scope rules **2**:159–160, **2**:162*T*
    summary discussion **2**:165
  management approaches
    allocated quotas **2**:152
    community control **2**:153
    discounting models **2**:153
    incentives versus resources **2**:152
    international common property resources **2**:152–153
    risk and uncertainty **2**:153–154
  public-based market approaches **5**:150–152
  regulated bionomic equilibrium **2**:151–152
  social cost **2**:154
  spillover effects **2**:154
  tragedy of the commons **2**:149–150, **2**:234–235, **5**:150–152
common shiner **6**:153*F*
common toads **3**:165
common waxbills **1**:83
Communal Areas Management Programme for Indigenous Resources (CAMPFIRE) **3**:197
community-based conservation **4**:316*F*, **4**:320–321, **7**:19–20
Community-Based Natural Resources Management (CBNRM) **2**:297–298
community diversity **2:595–607**
  definition **2**:596
  local species richness controls
    diversity–productivity relationship **2**:597–599, **2**:598*F*
    prediction limitations **2**:599
    research summary **2**:606
    spatial scale **2**:604–605, **2**:605*F*
    species–energy theory **2**:596–597, **2**:597*T*
    theoretical perspectives **2**:605
    Volterra–Gause perspective **2**:595–596
  measurement and analysis **1**:403
  regional species richness controls
    historical biogeographical factors **2**:599–601, **2**:600*F*
    metacommunity processes model **2**:601–602, **2**:602*F*
    phylogenetic patterns **2**:599–601, **2**:600*F*
    proportional sampling **2**:602*F*, **2**:603
    research summary **2**:606
    spatial scale **2**:604–605, **2**:605*F*
    theoretical perspectives **2**:605
  research summary **2**:606
  saturation tests
    caveats **2**:603
    experimental evidence **2**:603–604, **2**:604*F*
    methodologies **2**:602–603
    test results **2**:603
    unsaturated versus saturated patterns **2**:603
  spatial scale **2**:604–605, **2**:605*F*

theoretical perspectives
comparative models **2**:606, **2**:606*F*
design manipulations **2**:605–606
multiscale perspectives **2**:605
Type I/Type II relationships **2**:603
community dominance–diversity relationships **6**:431–432, **6**:432*F*
community ecology **2**:94, **5**:510, **5**:666–667, **5**:681–682
community gardens **6**:505
Community Multiscale Air Quality (CMAQ) model **5**:533
community seed banks **3**:644
Comoros
botanical gardens **6**:507*T*
deforestation **2**:449*T*
energy efficiency measures **3**:254*F*
compensated versus uncompensated demand **7**:321–322, **7**:322*F*
competition
apparent competition **6**:245–246
character displacement **6**:720
competitive exclusion principle
apparent competition (*P*–rule) **6**:671
basic concepts **6**:670–671
exploitative competition (*R*–rule) **6**:671
exploitative competition **6**:716–717
habitat–niche relationship **4**:45–46
interference competition **4**:360, **6**:669–670, **6**:716–717
interspecific competition **2:177–191**
characteristics **6**:716–717
classifications **2**:177–178, **2**:177*T*
coevolution **2**:140–141
competition models
aggregation model **2**:181–182, **2**:182*F*
density-dependent function **2**:178, **2**:178*F*
Lotka–Volterra model **2**:178–180, **2**:178*F*, **2**:179*F*, **6**:716–717, **6**:716*F*, **6**:717*F*
MacArthur model **2**:180–181, **2**:180*F*, **2**:181*F*
resource utilization curves **2**:180–181, **2**:180*F*, **2**:181*F*
field experiments
Anolis lizards **2**:187–188, **2**:187*T*
ants and rodents **2**:186–187, **2**:187*T*
barnacles **2**:186, **2**:187*F*
field observations
African wild dogs and hyenas **2**:190, **2**:190*F*, **2**:190*T*
New England warblers **2**:188–189, **2**:188*F*
Pacific Island birds **2**:189–190, **2**:189*F*
laboratory experiments
Drosophila **2**:185–186, **2**:186*F*
flour beetles (*Tribolium*) **2**:184, **2**:184*T*, **2**:185*F*, **2**:185*T*, **6**:670
protozoa **2**:182–184, **2**:183*F*, **2**:183*T*, **2**:184*F*, **6**:669–670
research results **6**:670
yeasts **2**:182–184, **2**:183*T*
research summary **2**:190–191
two-prey competition **2**:177, **2**:177*F*
intertidal ecosystems **4**:334, **4**:335*F*
introduced species
general discussion **4**:360
interference competition **4**:360
resource competition **4**:360
metapopulation theory **5**:243–244, **5**:248–249, **5**:249*F*, **5**:250*F*
plant biodiversity–climate relationship **2**:123
plant–pollinator interactions **6**:49
population dynamics **6**:180
species coexistence **6**:668–669
species interactions **6**:180, **6**:715*T*, **6**:716–717
complementarity **2:192–200**
area selection
historical perspective **2**:193
hotspots **2**:193, **2**:193*F*, **4**:132
importance **2**:193, **2**:193*F*
biological basis **2**:192–193, **2**:192*F*
research applications
biodiversity indicators **2**:196*F*, **2**:197–200, **2**:197*F*, **2**:198*F*, **2**:199*F*
changing distributions **2**:196–197
pattern variations **2**:200
probability of persistence or occurrence **2**:196
research summary **2**:200
simple species representation
conservation goals **2**:193–194
efficiency factors **2**:194–196, **2**:195*F*
flexibility factors **2**:194, **2**:197*F*
irreplaceability **2**:194
priorities **2**:194
research techniques **2**:194
complementation analysis **1**:323*F*, **1**:324
complex carbohydrates **6**:113*T*
Compositae **3**:129*T*, **6**:316*T*
composts **6**:594
Compsilura cocinnata **5**:402–403
computational general equilibrium (CGE) models **4**:539–540
computer systems and models **2:213–220**
applications
Habitat Suitability Index (HSI) models **2**:218
individual-based models **2**:218–219
metapopulation models **2**:218
characteristics **2**:213
functional role
control **2**:214
description **2**:213
mechanism **2**:213–214
prediction **2**:214
future trends
behavioral dynamics models **2**:219
multimodeling **2**:219
regional assessment **2**:219
limitations
aggregation **2**:216, **2**:216*F*
loss of detail **2**:216
tradeoffs **2**:216, **2**:216*F*
uncertainties **2**:216–217
model types
computational approaches **2**:215
computer models **2**:215
ecological models **2**:215
mathematical models **2**:214–215, **2**:215*F*
physical models **2**:214–216
system simulation models **2**:215–216, **2**:216*F*
tools
dynamic models **2**:217
geographical information systems (GIS) **2**:217–218
statistical models **2**:217
Comstock, John Henry **7**:89
concentrated animal feeding operations (CAFOs) **1**:99
Concholepas concholepas **5**:56–57, **5**:58*F*, **5**:59*F*, **6**:232
Conchostraca
extinction vulnerability **2**:416–417*T*
number of species **3**:182*T*
Condylarthra **2**:583*F*
Condylura cristata **4**:690–691
cone snails **1**:359
Conference of the Parties (COP) **2**:307–308
Congaree National Monument (South Carolina) **7**:385*F*
Conger eels **7**:389–390
Congeria kusceri **7**:57
Congo **4**:432*T*, **6**:519
congruence tests **2**:37
coniferous forests
biodiversity change scenarios **4**:6*F*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
North America **5**:564–565
protected area systems **6**:492*T*, **6**:499*T*
South America **6**:600–604*T*
conifers
characteristics **6**:59–60
number of species **6**:58*T*
threatened species **7**:201*T*
Coniornis **2**:519*T*
Conium maculatum **6**:42
conjunction tests **2**:37
connectivity food webs **3**:498–500, **3**:499*F*, **3**:500*F*
Connochaetes spp. **4**:362
Connochaetes taurinus **1**:52, **3**:191, **5**:313
Conocardium sp. **5**:375*F*
Conocarpus erectus **1**:164
Conocephalinae **3**:727
Conochilis unicornis **1**:10*T*
Conodontophora **3**:685*F*
conodonts **3**:684–685, **7**:337–338
conopeptides **1**:359
Conopholis americana **4**:75
Conopidae **3**:489–491*T*
Conopoidea **3**:489–491*T*
Conospermum **5**:713
Conostylis **5**:221
consequential life cycle analysis (LCA) **4**:627–628
conservancy/megaconservancy networks **2**:301, **2**:302*F*
conservation
agrobiodiversity **1**:133–134, **1**:134*F*

conservation (*continued*)
amphibians **3**:164, **3**:164*T*
as ecosystem components **1**:176
human uses **1**:176
major threats **1**:176–177
population declines **1**:177
Asian ecosystems **1**:279
biodiversity conservation **4:313–323**
basic principles **4**:315
biodiversity-friendly farming **1:418–429**
background information **1**:418–419
biodiversity stewardship **2**:300
chemical use minimization **2**:542
crop production **2**:542–544
disadvantages **2**:544
disturbance events **2**:542–544
influencing factors **2**:544, **2**:554*F*
landscape-scale effects **2**:544
livestock grazing **2**:542–544
local management strategies **1**:419–420
management practices **2**:299–300, **2**:300
native vegetation benefits **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
pest control **2**:544
pollination **2**:544
research summary **2**:545–546
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
structural complex agricultural areas **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
biogeographic research **1**:585–586
botanical gardens **6**:517
conservation biologists **4**:123–124
conservation challenges **4**:122
conservation products
biodiversity utilization **4**:316*F*, 4:321–322
characteristics **4**:321, **4**:321*T*
product dissemination **4**:316*F*, **4**:321
energy sprawl **1**:528–529, **1**:529*F*
ex situ conservation
agrobiodiversity conversation **3**:33
botanical/zoological garden conservation **4**:316*F*, **4**:319, **4**:321*T*
field gene banks **3**:33, **4**:316*F*, **4**:319, **4**:321*T*
in vitro conservation **4**:316*F*, **4**:319
livestock park conservation **4**:316*F*, **4**:319, **4**:321*T*
maintenance strategies **1**:406
plant biodiversity **6**:63
pollen/semen/ovule/DNA conservation **4**:316*F*, **4**:319–320, **4**:321*T*
research programs and methodologies **4**:318–319
seed/embryo collection and storage **4**:316*F*, **4**:319, **4**:319*F*, **4**:321*T*
genetic diversity loss **4**:313–314
habitat–niche relationship **4**:47–48
importance **4**:313
in situ conservation
agrobiodiversity conversation **3**:33
genetic reserve conservation **4**:316*F*, **4**:320, **4**:320*F*, **4**:321*T*
maintenance strategies **1**:405–406
on-farm conservation **3**:33, **4**:316*F*, **4**:320, **4**:321*T*
plant biodiversity **6**:63
research programs and methodologies **4**:316*F*, **4**:320
international organizations **4**:325–326
justice and equity
best management practices **4**:435
ecosystem services **4**:435
natural resources **4**:435
program support **4**:435
protected areas and heritage sites **4**:435–436
sustainability **4**:436
landraces **1**:122–123
major threats **4**:314–315
outside protected areas
biosphere reserves **2**:301
challenges **2**:292–293
conservancy/megaconservancy networks **2**:301, **2**:302*F*
conservation approaches **2**:293–294, **2**:294*T*
land and water protection and management **2**:294, **2**:294*F*, **2**:294*T*
living lands and waters **2**:294*F*, **2**:295–297, **2**:301–303
living landscape framework **2**:294–295, **2**:294*F*
systematic conservation planning **2**:293
working lands and waters **2**:294*F*, **2**:298–299, **2**:301–303
priority management **6:261–272**
carbon sequestration **5**:351–352, **6**:264*F*
carbon storage **5**:351–352, **6**:264*F*
challenges **6**:270–271
cost–benefit analyses **6**:265–266, **6**:266*F*
economic factors **6**:266–269, **6**:267*F*, **6**:268*F*, **6**:269*F*
ecosystem service targets **6**:263–265, **6**:264*F*
land-use patterns **6**:265–266, **6**:266*F*, **6**:268*F*, **6**:269*F*
livestock **6**:264*F*
mapping and modeling approaches **6**:263–265
research summary **6**:269–271
systematic conservation planning **2**:293, **6**:261–263
water provisions **6**:264*F*
protected area systems
characteristics **2**:289–292
classifications **2**:290
geographic distribution **2**:291*F*
global biome coverage **2**:292*F*
global growth estimates **2**:290*F*
research programs and methodologies
community-based conservation **4**:316*F*, 4:320–321
conservation objectives **4**:316*F*, **4**:318
conservation products **4**:316*F*, **4**:321, **4**:321*T*
conservation strategies **4**:316*F*, **4**:318
ecogeographic surveys/preliminary survey missions **4**:316*F*, **4**:317–318
ex situ techniques **4**:316*F*, **4**:318–319
field exploration **4**:316*F*, **4**:318
goals **4**:315–317
in situ techniques **4**:316*F*, **4**:320
modeling approaches **4**:315–317, **4**:316*F*
project commission **4**:316*F*, **4**:317
target taxa selection **4**:315–317, **4**:316*F*
Southeast Asia **1**:520–521
species–area relationships (SARs)
background information **6**:646–648
incidence functions **6**:649–650
loss management strategies **6**:648–649
nestedness measures **6**:649–650
relaxation concept **6**:647
species loss **6**:643*F*, **6**:646–648, **6**:647*F*
summary discussion **6**:650
species concepts
ecological importance **6**:689
genetic differentiation **6**:689
spatiotemporal considerations **6**:689–690
species as real entities **6**:689
sustainable and integrated biodiversity conservation **4**:322, **4**:436
systematic conservation planning **2**:293, **6**:261–263
traditional agriculture **1**:122–123
zoos/zoological parks
ark paradigm **7**:477–478
breeding population research **7**:474
frozen zoos **7**:477
institutional planning **7**:475
management strategies **7**:474–475
population management strategies **7**:475–477
post-ark paradigm **7**:478
record keeping plans **7**:475
research summary **7**:478
birds **1**:623
boreal forests
clear-cutting threats **1**:634
natural reserves and preserves **1**:634
sustainable management **1**:634
climate change-based range shifts **4**:607–608
conservation biology **2:238–248**
basic concepts **2**:238–239
biodiversity-friendly farming **1:418–429**
background information **1**:418–419
chemical use minimization **2**:542
crop production **2**:542–544
disadvantages **2**:544
disturbance events **2**:542–544

influencing factors **2**:544, **2**:554*F*
landscape-scale effects **2**:544
livestock grazing **2**:542–544
local management strategies **1**:419–420
native vegetation benefits **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
pest control **2**:544
pollination **2**:544
research summary **2**:545–546
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
structural complex agricultural areas **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
contemporary conservation theories and practices **2:249–262**
climate change **2**:260–261
direct and indirect anthropogenic impacts **2**:253–254
emerging challenges **2**:261–262
evolutionary development **2**:252–253
future outlook **2**:261–262
human altered landscapes **2**:260–261
management tools and approaches **2**:255–257
philosophical perspectives **2**:249–252
pre-1990 and post-1990 practices **2**:251*T*
species-to-ecosystem shift **2**:252–253
Development by Design (DbD) **2:495–507**
background and characteristics **2**:496–497, **2**:497*F*
framework and components **2**:497–498
research summary **2**:505–506
Wyoming Basins Ecoregion (WBE) study **2**:503–505, **2**:504–505*F*
Diptera **3**:495–496
ecological communities
gap analysis **2**:241, **2**:241*F*, **2**:242*F*
landscape corridors and connectivity **2**:242–243
natural reserves/protected area systems **2**:239–241, **2**:240*F*
reserve selection algorithms **2**:241–242, **4**:262–263
single large or several small (SLOSS) debate **2**:240–241, **4**:495, **6**:649
surrogate data **2**:242*F*
ecological complexity versus biodiversity **2:201–212**
background information **2**:201
basic concepts **2**:204
biotic interactions and relationships **2**:202–204, **2**:203*F*, **2**:204*F*
community complexity **2**:203*F*, **2**:204–205, **2**:205–207, **2**:205*T*, **2**:209–210
complexity measures **2**:205
ecosystem functioning **2**:210
entangled bank metaphor **2**:202–204
functional complexity **2**:203*F*, **2**:204–205, **2**:205*T*, **2**:207–208
historical perspective **2**:201–204
hypothetical relationships **2**:210–211, **2**:211*F*
intricacy measures **2**:205*T*
phylogenetic complexity **2**:203*F*, **2**:204–205, **2**:205*T*, **2**:208, **2**:209*F*, **2**:210
research summary **2**:210–211
schematic diagram **2**:203*F*, **2**:204*F*
significance **2**:209–210
theoretical perspectives **2**:210–211
usefulness **2**:208–209
expanding toolbox **2**:246, **2**:247*F*
extinction-risk assessment **2**:245–248, **2**:246*F*
future challenges **2**:245–248
genetic diversity **2:263–277**
ex-situ conservation **2**:273–274
future outlook **2**:274–275
genetic drift **2**:265–267
genomic research **2**:274–275, **2**:275*F*
heterozygosity and fitness correlations (HFCs) **2**:266
historical perspective **2**:263–265
inbreeding **2**:265–267, **2**:272–273
individual genotypes **2**:272–273
kin relationships **2**:272–273
low diversity–small population relationship **2**:265–267
management units **2**:268–269
measurement methods **2**:244–245
population structure **2**:267–272, **2**:269*F*, **2**:270*F*
research approaches **2**:244–245
skyline plots **2**:267, **2**:267*F*
zoos, aquaria, botanical gardens, and gene banks **2**:245, **4**:316*F*, **4**:319, **4**:321*T*
Global Environment Facility (GEF) **7:136–143**
historical perspective **2**:286–287
marine ecosystems **5:32–44**
abundance impacts **5**:34–35
assisted colonization **5**:41
climate change effects **3**:449, **5**:32
community impacts **5**:37–38
conservation strategies **5**:38–39
distribution impacts **5**:35–36
ecosystem-based management **5**:40
evolutionary impacts **5**:36
fisheries management **5**:40–41
future research areas **5**:41
marine protected areas **5**:39–40
ocean acidification **5**:33–34
predictive models **5**:39
research needs **5**:41
restoration **5**:41
rising seawater temperatures **5**:33
marine protected areas **5:94–104**
benefits **5**:96–97
characteristics and functional role **5**:94–95
climate change threats **5**:100–101
ecosystem-based management **5**:95
effectiveness **5**:100
fish stocks **3**:486
geographic extent **5**:445–446, **5**:446*T*
global commitment **5**:99–100
socioeconomic benefits **5**:98–99, **5**:98*T*
threats and challenges **5**:100–101, **5**:447–448
molecular level diversity **2**:637
species and populations
biodiversity protection **2**:243–244
population viability analysis (PVA) **2**:243–244
small and declining population demographics **2**:244
species–area relationships (SARs)
background information **6**:646–648
incidence functions **6**:649–650
loss management strategies **6**:648–649
nestedness measures **6**:649–650
relaxation concept **6**:647
species loss **6**:643*F*, **6**:646–648, **6**:647*F*
summary discussion **6**:650
tourism industry **7:222–225**
anthropogenic impacts **7**:223–224
biodiversity impacts **7**:222
conservation tourism **7**:224
ecotourism **1**:395–396, **7**:224
geographic distribution **7**:222–223
nature-based tourism **2**:232–233
primary attractions **7**:223
research summary **7**:224–225
translocations **7:236–240**
background information **7**:236
ethical concerns **7**:239
managed relocations **7**:238
motivations **7**:237–238
research summary **7**:239–240
risks and costs **7**:238–239, **7**:239*F*
success assessments **7**:238
traditional conservation practices **7**:236–237
conservation genetics **2:263–277**
ex-situ conservation **2**:273–274
future outlook **2**:274–275
genetic drift **2**:265–267
genomic research **2**:274–275, **2**:275*F*
heterozygosity and fitness correlations (HFCs) **2**:266
historical perspective **2**:263–265
inbreeding **2**:265–267, **2**:272–273
individual genotypes **2**:272–273
kin relationships **2**:272–273
low diversity–small population relationship **2**:265–267
management units **2**:268–269
population structure **2**:267–272, **2**:269*F*, **2**:270*F*
skyline plots **2**:267, **2**:267*F*
contemporary conservation theories and practices **2:249–262**
climate change **2**:260–261
direct and indirect anthropogenic impacts
invasive species **2**:254–255
sustainable development **2**:253–254
emerging challenges **2**:261–262

conservation (*continued*)
evolutionary development
contemporary conservation tools **2**:255–257
direct and indirect anthropogenic impacts **2**:253–254
species-to-ecosystem shift **2**:252–253
future outlook **2**:261–262
human altered landscapes **2**:260–261
management tools and approaches
appropriate conservation landscape scale **2**:257
biodiversity conservation goals and priorities **2**:256–257
civil society-based tools **2**:258, **2**:258*T*
contemporary conservation tools **2**:255–257
economic evaluations **2**:258–260
governance-based tools **2**:258, **2**:258*T*
historical perspective **2**:255–257
landscape ecosystem-based conservation scale **2**:257–258
market-based tools **2**:258, **2**:258*T*
multiple operational tools **2**:258, **2**:258*T*
species-based conservation scale **2**:257
tradeoffs **2**:258–260
philosophical perspectives **2**:249–252
pre-1990 and post-1990 practices **2**:251*T*
ecological restoration **4:199–208**
advantages/disadvantages **4**:199–200
basic concepts **4**:199–200
ecosystem restoration potential **4**:203–204
explicit restoration targets **4**:205
future outlook
biodiversity crisis **4**:205
key challenges **4**:205–206
socioeconomic justifications **4**:206
plant conservation **6**:63, **6**:85–87
restoration by-products **4**:204
restoration justification **4**:204–205
restoration targets
characteristics **4**:200–201
ecosystem drivers **4**:201
ecosystem service restoration **4**:201–202
entire ecosystems **4**:200–201, **4**:202*T*
explicit restoration targets **4**:205
global climate change **4**:203
novel ecosystems **4**:202–203
whole-ecosystem restoration **4**:201
emerging themes **2**:287–288
endemic species **3**:235
environmental ethics **3:309–318**
animal rights **3**:313–314
basic principles **3**:309
conservation biologists
personal choices **3**:317–318
professional considerations **3**:317
role models **3**:316–317, **3**:317*F*
ethical holism **3**:269–270, **3**:270*F*
goals
ecological integrity **3**:310–311, **3**:312*F*
economic factors **3**:309–310
long-term healthy ecosystems **3**:309–310, **3**:310*T*, **3**:311*F*
sustainability **3**:311–312
human rights
challenges **3**:314–315
less-developed countries **3**:315–316, **3**:316*F*
logging industry–owl population controversy **3**:314–315, **3**:315*F*
property rights **3**:315
management versus nonmanagement **3**:312–313, **3**:313*F*
evolutionary significant unit (ESU) **1**:668–669
fish species **3:443–455**
commodity versus wildlife issue **3**:445
ecosystem diversity **3**:452–453
future outlook
marine protected areas **3**:453–454
potential conservation strategies **3**:454
public and political involvement **3**:454–455
industrialized fishing **3**:445
major threats
aquaculture **3**:448–449, **3**:452
general characteristics **3**:447
global climate change **3**:449–450
habitat loss and fragmentation **3**:447–448
human population growth **3**:447
introduced species **3**:448
pollution **3**:449
marine fish biodiversity **3**:450
ocean fishing
bycatch **3**:446, **3**:446, **3**:447, **4**:156
commodity versus wildlife issue **3**:445
economic losses **3**:446–447
government subsidies **3**:446–447
industrialized fishing **3**:445
international fishing conflicts **3**:444–445
management problems **3**:444
overfishing **3**:443–444, **3**:444, **3**:445, **3**:445–446, **3**:446, **4**:156
socioeconomic concerns **3**:444
species diversity
depletion causes and consequences **3**:451–452
extinction stages **3**:452
extinction vulnerability **3**:451–452
within-species genetic diversity **3**:450–451
governance systems **2:155–166**
advantages/disadvantages **2**:163–164
centralized regimes **2**:156–157
challenges **2**:155–156
complex adaptive systems
building blocks **2**:161–162
internal models **2**:161
parallel adaptive systems **2**:163–164
polycentric systems **2**:164–165
properties and mechanisms **2**:161
rule configurations and situations **2**:162–163, **2**:162*T*
self-organized resource governance systems **2**:160–161
tags **2**:161
local governance systems **2**:163–164
polycentric systems **2**:164–165
rules as tools
aggregation rules **2**:159–160, **2**:162*T*
boundary rules **2**:157–158, **2**:162*T*
choice rules **2**:158–159, **2**:162*T*
information rules **2**:159–160, **2**:162*T*
payoff rules **2**:159, **2**:162*T*
position rules **2**:159, **2**:162*T*
rule change experimentation **2**:160
rule characteristics **2**:157–158, **2**:162*T*
rule improvements **2**:160
scope rules **2**:159–160, **2**:162*T*
summary discussion **2**:165
historical perspective **2:278–288**
background information
general discussion **2**:278–279
historic precedents **2**:279–280
prehistoric precedents **2**:278–279
emerging themes **2**:287–288
modern origins
foundations **2**:280–281
Progressive Era **2**:282–283
reform direction and momentum **2**:281–282
twentieth century conservation
challenges **2**:283
conservation program reintegration **2**:286–287
forestry management **2**:283
institutionalization **2**:283
1930s and 1940s **2**:284–285
post-World War II era **2**:285–286
range management **2**:283
recreational activities **2**:283
soil conservation **2**:283–284
wildlife management **2**:283
historic land-use patterns **4:569–582**
anthropogenic impacts **4**:569–570
buildings/built-up areas **4**:579, **4**:580*F*
climate change effects **4**:569–570
conservation implications **4**:581–582
coppice woods **4**:573, **4**:573*F*, **4**:574*F*
crop plants **4**:570, **4**:570*F*
domestic animals **4**:570
fields **4**:570–572, **4**:571*F*
fire cycles **4**:580, **4**:581*F*
forests and woodlands **4**:572–574, **4**:573*F*, **4**:574*F*
grasslands **4**:577
heathland **4**:577, **4**:578*F*
hedges **4**:571, **4**:571*F*
historical perspective **4**:569, **4**:579–581
moorland **4**:577–579, **4**:578*F*
permanent edges and open areas **4**:573–574, **4**:575*F*
terraces **4**:572, **4**:572*F*
wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*, **4**:581*F*
keystone species **4**:456–457

livestock genetic resources
ex situ strategies **1**:645–646
in situ strategies **1**:642–645
mammals **4:708–720**
captive breeding programs **3**:196–197, **4**:710–711
Caribbean hutias and solenodons
background and characteristics **4**:713–714, **4**:713*F*, **4**:714*F*, **4**:715*F*
conservation practices **4**:715
general discussion **4**:713–714
major threats **4**:714–715
case studies
Caribbean hutias and solenodons **4**:713–714
cheetahs and African wild dogs **4**:715–716, **4**:717*F*
general discussion **4**:711
important themes **4**:717–719
saiga antelope **4**:711
charismatic species **3**:197–198
cheetahs and African wild dogs
background and characteristics **4**:715–716, **4**:717*F*
conservation practices **4**:716–717
major threats **4**:716
range areas **4**:718*F*
conservation types
captive breeding programs **3**:196–197, **4**:710–711
government regulation **4**:710
habitat/population management **4**:710
protected area systems **3**:196, **4**:710–711
reintroduction programs **3**:196–197
species protection **4**:710
sustainability **4**:710
economic incentives **3**:197
general discussion **4**:708
government legislation and regulations **3**:195
human impact **4**:709–710
needs assessments **4**:708–709
protected area systems **3**:196, **4**:710–711
reintroduction programs **3**:196–197
research summary **4**:719–720
saiga antelope
background and characteristics **4**:711, **4**:711*F*
conservation practices **4**:712–713
general discussion **4**:711
major threats **4**:711–712, **4**:712*F*
population decline **4**:712*F*
range area **4**:712*F*
Species Survival Commissions (SSC) **3**:195–196
threat assessment **3**:191–192, **4**:709–710
umbrella species **3**:197–198
marine ecosystems **5:32–44**
climate change effects
abundance impacts **5**:34–35
adaptation rates **5**:36
assisted colonization **5**:41
climate tolerance **5**:36–37
community impacts **5**:37–38
conservation strategies **5**:38–39
distribution impacts **5**:35–36
ecosystem-based management **5**:40
evolutionary impacts **5**:36
fisheries management **5**:40–41
genetic diversity **5**:36–37
global impacts **3**:449, **5**:32
local adaptations **5**:36
marine protected areas **5**:39–40
ocean acidification **5**:33–34
physiological impacts **5**:32
predictive models **5**:39
research needs **5**:41
restoration **5**:41
rising seawater temperatures **5**:33
future research areas **5**:41
metapopulation theory **5**:250
nature-based recreation
conservation strategies
ecosystem services **7**:253
health benefits **7**:253
conservation support
backpacking/hiking groups **7**:251–252
declining support **7**:251
demographics **7**:252–253
diversification strategies **7**:252
future outlook **7**:252–253
popular activities **7**:251
public land visits **7**:251–252
two-Americas concept **7**:251–252
neotropical mammals **2**:591–592
people–conservation relationship **2:221–229**, **2:230–237**
benefits
cost–benefits **2**:226
fishing industry **2**:232
forestry sector **2**:232
grasslands **2**:232
nature-based tourism **2**:232–233
payment for ecosystem services (PES) **2**:226–227
poverty reduction **2**:227–228
protected area systems **2**:226
conservation threats
human impact measures and maps **2**:221–222
reduction strategies **2**:222–223
disadvantages **2**:233
general discussion **2**:221
goals and priorities
general discussion **2**:223–224
preferential treatment large-bodied vertebrates **2**:224
shifting baselines **2**:223–224
wilderness ideal **2**:224
historical perspective **2**:230–231
key organizational players **2**:231–232
poverty indicators **2**:233
poverty trends **2**:233–235, **2**:234*F*, **2**:234*T*
research summary **2**:228, **2**:235
victimization factors
conservation refugees **2**:224–225
general discussion **2**:224–225
negative impacts **2**:225–226
protected area systems **2**:224–225
unemployment considerations **2**:225–226
plant communities **6:76–89**
anthropogenic impacts **6**:79–84
Ascension Island **6**:78
biodiversity distribution and loss **6**:79
biodiversity loss responses **6**:63, **6**:85–87
botanical gardens
ex situ conservation programs **6**:513, **6**:513–514, **6**:514
Global Strategy for Plant Conservation (GSPC) **6**:514–515, **6**:515
International Agenda for Botanic Gardens in Conservation **6**:514
Caribbean region **6**:83
crop wild relatives (CWRs) **6**:77
historical perspective **6**:76–79
hotspots **6**:79, **6**:82*T*
importance **6**:76
international agreements **6**:76–79, **6**:80–81*T*, **6**:509
living dead **6**:81
mosses **6**:83
restoration
dominance–diversity curve **6**:440*F*
limitations **6**:440–442, **6**:441*F*
single-species introductions **6**:442–443
species richness improvements **6**:442
succession **6**:443
vegetation management **6**:439–442, **6**:440*F*
salt water plants **6**:77
threatened species **6**:84–85
vascular plant diversity **6**:82*T*
religious traditions **6**:563–566
restoration **6:370–379**, **6:438–452**
active versus passive restoration **6**:439, **6**:439*F*
animal restoration
basic concepts **6**:443–444
habitat restoration **6**:445–446
population sources **6**:446
site assessments **6**:445–446
species reductions or absences **6**:444–445
definition **6**:438
designer ecosystems **6**:451
goals and limitations **6**:438–439
microorganisms
diversity evaluations **6**:446–447
establishment criteria **6**:450–451, **6**:450*F*
free-living saprobes **6**:448
functional groups **6**:448
soil food webs **6**:448–449, **6**:449*F*
spatiotemporal relationships **6**:447–448, **6**:447*F*
symbionts **6**:449–450, **6**:450*F*
plant communities
dominance–diversity curve **6**:440*F*
limitations **6**:440–442, **6**:441*F*

conservation (*continued*)
single-species introductions **6**:442–443
species richness improvements **6**:442
succession **6**:443
vegetation management **6**:439–442, **6**:440*F*
success assessments **6**:451
return-on-investment (ROI) analysis **4:185–198**
analytical examples
action and location allocation **4**:193–194, **4**:195*F*, **4**:196*T*
action and location prioritization **4**:192–193, **4**:193*F*
allocation strategies **4**:193
budget allocations **4**:195*T*
cost–risk–complementarity relationship **4**:195–196
dependent returns **4**:194–195, **4**:197*F*
diminishing returns **4**:190–191, **4**:190*F*, **4**:191*T*
land purchase prioritization **4**:193, **4**:195*T*
restoration strategy prioritization **4**:190, **4**:190*F*
threat considerations **4**:191–192, **4**:192*F*
threatened species recovery projects **4**:193, **4**:194*T*
unknown costs **4**:196–197
basic concepts **4**:185–186
challenges **4**:197
conservation prioritization applications **4**:186, **4**:187*T*
conservation return
basic concepts **4**:188–189
benefit functions **4**:188–189, **4**:189*F*
diminishing returns **4**:188–189
probability of success **4**:189
cost data types **4**:190*F*
framing problems **4**:187–188
investment estimation **4**:189
valuation measures **4**:186–187
soils
agricultural methods
biotic community manipulation **6**:594–595
diversified cropping **6**:594
general characteristics **6**:593–594
organic amendments **6**:594
pollutant bioremediation **6**:594
reduced tillage practices **6**:593–594
twentieth century conservation **2**:283–284
assessment and monitoring programs
basic process **6**:595–596
government policies and regulations **6**:597–598
indicator measures **6**:596
indicator measures
basic concepts **6**:596
biological attributes **6**:597
chemical attributes **6**:596–597
physical attributes **6**:596
research summary **6**:598
twentieth century conservation **2**:283–284
urban soils **6**:595
South American ecosystems **6**:610
stewardship **7:12–22**
contexts
cultural factors **7**:17–18
ecological factors **7**:16–17
economic factors **7**:17
general discussion **7**:16–17
political factors **7**:17–18
etymological background **7**:12–13
rationales
counterarguments **7**:14–16
general discussion **7**:13–14
instrumental arguments **7**:13–14
intrinsic arguments **7**:14
steward types
community-based conservation **4**:320–321, **7**:19–20
cross-boundary stewardship **7**:21
general discussion **7**:18–19
governmental stewards **7**:18–19
nongovernmental organizations (NGOs) **7**:19–20
private companies and individuals **4**:330, **7**:20–21
threatened species lists and assessments **7**:207–208
systematic conservation planning **2**:293, **6**:261–263
temperate forests
conservation status **7**:116*T*, **7**:125–126
future research areas **7**:127
major threats **7**:126–127
protection strategies **7**:126
terrestrial invertebrates
filter approach **3**:224
recovery programs **3**:225
surrogate use **3**:224–225
umbrella/flagship species **3**:225
threatened birds
critically endangered birds
alien and native restriction and control **7**:188
captive breeding programs **7**:188
challenges **7**:189
cross-fostering/cross-breeding **7**:189
habitat restoration **7**:187–188
nest site provision/enhancement **7**:189
reintroduction/translocation programs **7**:188
supplementary feeding programs **7**:188–189
general discussion **7**:186
goals and priorities **7**:186
international cooperation and legislation **7**:187
priority management **7**:186
scale considerations **7**:186
site and habitat conservation **7**:186–187
traditional conservation practices **7:226–235**
background information **7**:226–227
complex ecosystem dynamics
disturbance management **7**:233
ecosystem renewal nurturance **7**:233–234
general discussion **7**:232–233
landscape patchiness management **7**:232–233
multiple-scale ecological processes **7**:233
watershed management **7**:233
cultural belief systems **7**:226–227
ecological monitoring and practices
habitat protection **7**:227*T*, **7**:229
harvest restrictions **7**:227*T*, **7**:229–230
resource monitoring status **7**:227–228, **7**:227*T*
species protection **4**:710, **7**:227*T*, **7**:228, **7**:229*T*
vulnerable life histories protection **7**:227*T*, **7**:228–229
management strategies
agroforestry practices **7**:230–231
multiple-species management **7**:230–231
resource rotation **7**:231
shifting cultivation systems **7**:231–232, **7**:232*F*
succession management **7**:231–232
research summary **7**:234
social–ecological systems **7**:227*F*
translocations **7**:236–237
translocations **7:236–240**
background information **7**:236
ethical concerns **7**:239
managed relocations **7**:238
motivations **7**:237–238
research summary **7**:239–240
risks and costs **7**:238–239, **7**:239*F*
success assessments **7**:238
traditional conservation practices **7**:236–237
tropical rain forests **6**:320
wetlands **7**:371
wildlife management **7**:405–406
see also protected area systems
conservation biogeography **2**:348
conservation botanical gardens **6**:505
conservation breeding programs *see* captive breeding
Conservation International Center for Environmental Leadership in Business **4**:328
Conservation International (CI) **2**:231–232, **4**:325–326
Conservation Reserve Program **3**:719, **7**:371
Conservatoire et Jardin Botaniques (Geneva) **5**:411*F*, **6**:510–511
Consortium for the Barcode of Life (CBOL) **4**:327
conspecific brood parasitism (CBP)
definition **5**:501
prevalence **5**:508–509
species survey **5**:502–503
constant-equivalent consumption measures **3**:27
Constrictotermes **4**:409*T*

Consultative Group on International Agricultural Research (CGIAR) **3**:643, **4**:328
contemporary conservation theories and practices **2:249–262**
  climate change **2**:260–261
  emerging challenges **2**:261–262
  evolutionary development
    contemporary conservation tools **2**:255–257
    direct and indirect anthropogenic impacts
      invasive species **2**:254–255
      sustainable development **2**:253–254
    species-to-ecosystem shift **2**:252–253
  future outlook **2**:261–262
  human altered landscapes **2**:260–261
  management tools and approaches
    contemporary conservation tools
      appropriate conservation landscape scale **2**:257
      biodiversity conservation goals and priorities **2**:256–257
      evolutionary development **2**:255–257
      landscape ecosystem-based conservation scale **2**:257–258
      multiple operational tools **2**:258, **2**:258*T*
      species-based conservation scale **2**:257
    economic evaluations **2**:258–260
    historical perspective **2**:255–257
    tradeoffs **2**:258–260
  philosophical perspectives **2**:249–252
  pre-1990 and post-1990 practices **2**:251*T*
contemporary deforestation **2**:448–449, **2**:448*T*, **2**:449*T*
continental drift **1**:583, **6**:616
continental islands **5**:590–591
Continental Margin Ecosystems on a Worldwide Scale (COMARGE) **2**:17, **2**:18*F*
continental shelves
  Census of Marine Life project
    coastal zone **2**:15*F*, **2**:16, **2**:17*F*
    human edge concept **2**:15–16, **2**:15*F*
    nearshore zone **2**:15*F*, **2**:16, **2**:16*F*
    realms of study **2**:15–16, **2**:15*F*
  ecosystem diversity **3**:453, **5**:52
  marine sedimentary habitats **5**:108*T*
  symbiotic relationships **5**:117*F*
continental slope **2**:17, **5**:108*T*, **5**:117*F*
contingent valuation method **7**:324–325
contraceptive methods **6**:202*F*
Conuropsis carolinensis **2**:281–282, **7**:182
Convention Concerning the Protection of the World Cultural and Natural Heritage **4**:325
Convention on Biological Diversity (1930) **2:306–319**
  Aichi Biodiversity Targets **2**:306–307, **2**:312, **2**:317–318, **7**:204–205
  background information **2**:306–307
  bacterial biodiversity **1**:307–308
  challenges
    aggressive implementation plans **2**:318
    climate change **2**:318
    land-use change **2**:317–318
    marine ecosystem protection **2**:318
    sustainability **2**:318
climate change effects **2**:79–80
community seed banks **3**:645
Conference of the Parties (COP) **2**:307–308
ethical perspectives
  bioprospecting **1**:596
  pharmaceuticals **5**:713–714
future outlook **2**:317–318
Global Biodiversity Outlook
  basic principles **2**:308–311
  biodiversity indicators **2**:309, **2**:310*F*
  tipping points **2**:309, **2**:309
Global Environment Facility (GEF) **7:136–143**
  evolutionary development
    biodiversity strategy **7**:138–139
    operational strategy and programs **7**:138
  functional role **2**:317, **4**:324–325, **7**:136–137
  future outlook **7**:141–142
  GEF benefits index (GBI) **7**:137–138
  GEF performance index (GPI) **7**:137–138
  global policy framework and guidance **7**:137
  gross domestic product index (GDPI) **7**:137–138
  key accomplishments **7**:142
  monitoring and evaluation strategies
    learning objectives **7**:140–141
    portfolio level **7**:140
    results-based management framework **7**:140–141
  origins **4**:324–325, **7**:136–137
  project investment and funding **7**:137, **7**:138*F*
  resource allocations **7**:207
  System forTransparent Allocation of Resources (STAR) **7**:137–138
goals and objectives **2**:307–308
governing bodies **2**:307–308
government-based educational initiatives **3**:133
historical background **4**:326, **6**:78
international awareness efforts **4**:125
international participation
  Clearing House Mechanism (CHM) **2**:317
  cooperation and collaboration efforts **2**:315–317
  ecosystem approach **2**:314, **2**:314–315
  Global Environment Facility (GEF) **2**:317
  implementation resources **2**:317
  novel concepts **2**:314–315
  policy guidance **2**:315, **2**:316
  programs of work **2**:315
mammal conservation **3**:195
national biodiversity strategies and action plans (NBSAPs) **2**:311–314
objectives **3**:713*T*, **3**:714–715, **4**:313, **6**:78
  plant breeding **1**:466–467
  protected area systems **6**:487, **6**:489, **6**:497
  scope **2**:307–308
  Strategic Plan for Diversity (2011-2020) **2**:311, **2**:312, **2**:317–318
  Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA) **2**:307–308
  threatened species **3**:195, **7**:187, **7**:204–205
  Working Group on the Review of Implementation **2**:308
  World Flora online database **6**:509
Convention on International Trade in Endangered Species (CITES)
  background information **4**:325
  conservation status of reptiles **3**:216, **3**:216*T*
  endangered species protection **3**:209, **3**:224
  illegal wildlife trade **1**:518
  mammal conservation **3**:195
  marine invertebrates **3**:199
  marine mammals **5**:88–89
  purpose **3**:713*T*, **3**:714
  seed banks **3**:647
  stewardship **7**:18–19
  threatened species **7**:187, **7**:205
  twentieth century conservation **2**:286
  whaling industry **3**:454
  wildlife management **1**:397, **3**:195
Convention on Migratory Species (CMS) **3**:195, **4**:326, **4**:712–713
Convention on Wetlands of International Importance **2**:286, **3**:195, **3**:713*T*, **4**:325
convex hull volume (CHV) measures **3**:600, **3**:600*F*
Convexitermes **4**:409*T*
Convolvulaceae **1**:24*F*, **1**:277, **6**:316*T*
convulsions **3**:635*T*
Cook Islands **5**:592*F*, **6**:507*T*
Cook, James **4**:651–652
Cook Strait giant weta **4**:194*T*
Coolatai grass **6**:93
coolibah **2**:468
Cooloolidae (Cooloola Monsters) **3**:731
cooperation **2:320–329**
  benefits and challenges **2**:320–321
  evolution and maintenance
    brief exchanges **2**:325–326
    common good theory **2**:326
    entangled fates **2**:324
    group safety **2**:321–322
    group selection **2**:322–323
    host enforcement **2**:324
    intimate mutualism **2**:323
    kin selection theory **2**:322
    mutual enforcement **2**:324–325
    partner choice **2**:323–324
    partner fidelity **2**:324
    sexual reproduction **2**:321–322, **2**:325–326
    symbiotic relationships **2**:322
    vertically transmitted partners **2**:323
  importance

cooperation (*continued*)
ecosystem evolution **2**:328–329
evolutionary transitions **2**:327
general discussion **2**:326–327
modularity **2**:327
natural selection **2**:327–328
Cooperative Holocene Mapping Project **2**:59–60, **2**:61*F*
Cooperative Strategy for the Conservation of Biological Diversity in the Arctic Region **2**:314
Cooper, James Fenimore **2**:280
cooters **7**:390
coots **3**:365–366
Copenhagen Accord **1**:289, **4**:21–22
Copenhagen University Botanic Garden **6**:506–508, **6**:510–511
Copepoda/copepods
American Southeast rivers **3**:185*F*
Census of Marine Life project **2**:18*F*
characteristics **5**:686–688, **5**:687*F*
classifications and sampling methods **4**:381*T*
coral reef ecosystems **2**:333*T*
diversity patterns **4**:385
eutrophication impacts **3**:362
extinction vulnerability **2**:416–417*T*
hydrothermal vent ecosystems **7**:330
lake and pond ecosystems **4**:371–373, **4**:374*F*, **4**:459
marine environments **4**:382*F*
number of species **2**:401*T*, **3**:182*T*, **4**:371*T*, **4**:384*T*
plankton fraction **6**:25
regional distribution **4**:42–45, **4**:42*F*, **5**:688*F*
seasonal cycle **5**:688*F*
spatial distribution **5**:687*F*
species richness **7**:60*T*
wetland ecosystems **7**:389–391
Copernicia tectorum **7**:387
Cope's rule **4**:698–699, **6**:67–68
Copiphora rhinoceros **3**:724–725
copper (Cu)
air pollution impacts **1**:142*F*
biogeochemical cycles **1**:563
inorganic fungicides **5**:694–695
inorganic herbicides **5**:695–696
organic fungicides **5**:695
coppice woods **4**:573, **4**:573*F*, **4**:574*F*
Copra hircus **1**:76–77, **3**:127*T*
coprine **3**:635*T*
Coprinus atramentarius **3**:635*T*
Copsychus saularis **5**:603–604
Copsychus sechellarum **7**:188
Coptotermes **4**:409*T*
Coptotermes acinaciformis **4**:417
Coquerel's sifaka **6**:256–258*T*
Coracias garrulus **5**:474
Coraciiformes **6**:308*T*
Coragyps stratus **5**:603
coral fungi **3**:629–630
Corallina vancouvericnsis **6**:723*T*
corallivores **2**:337–338
Corallomyxa **6**:291
coral reef ecosystems **2**:**330–346**
biodiversity
diversity patterns **2**:333–334
extinction risks **4**:664
interspecific interactions **2**:334
major groups **2**:331
species diversity **2**:331–333, **2**:332*T*, **2**:333*T*
biodiversity-rich countries **1**:500*T*
Census of Marine Life project **2**:15*F*, **2**:16
climate change effects **3**:449–450, **4**:664, **6**:150–151
conservation practices
greenhouse gas emissions **2**:344
local management and restoration **2**:343–344
conservation status **3**:171–173*T*, **3**:174
convergent evolution **3**:613–617, **3**:615*F*, **3**:616*F*
coral bleaching **2**:709–710
corals
characteristics and classifications **2**:334–335
classifications **4**:381*T*
climate change effects **5**:37
conservation practices
greenhouse gas emissions **2**:344
local management and restoration **2**:343–344
coral-dwelling microbes **2**:335
disappearance
anthropogenic impacts **2**:336–337, **2**:336*F*, **2**:337*F*
loss patterns and extent **2**:337
diversity **2**:334–335
evolutionary relationships **2**:334–335
indicator species **7**:206–207*F*
infectious diseases **2**:531–532
loss consequences
ecosystem goods and services **2**:343
species loss **2**:342–343
mortality agents
algal overgrowth **2**:338
anthropogenic causes **2**:340
coral bleaching **2**:338, **2**:338*F*, **2**:339*F*, **2**:340*F*, **2**:709–710, **3**:160–161, **3**:160*F*
corallivore explosions **2**:337–338
direct destruction **2**:340
disease incidence **2**:338–340, **2**:339*F*, **2**:340*F*
greenhouse gas emissions **2**:341–342, **2**:342*F*
invasive species **2**:341
overfishing **2**:341, **5**:61
poor water quality **2**:340–341
nutrition and growth **2**:335–336
photosynthetic symbioses **5**:118–120, **5**:119*F*
reproduction and recruitment **2**:335–336
species loss **5**:597
threatened species **7**:206–207*F*
disappearance
anthropogenic impacts **2**:336–337, **2**:336*F*, **2**:337*F*
loss patterns and extent **2**:337
disease regulation **3**:349
ecosystem diversity **3**:452–453, **5**:52
ecosystem engineers **4**:201
ecosystem services **1**:334*T*, **5**:465–466
El Niño Southern Oscillation (ENSO) events **3**:160–161, **3**:160*F*
endangered species **3**:199, **3**:202*T*
fish stocks **3**:481–482
functional groups
breaking points **3**:621, **3**:622*F*
characteristics **3**:617–619, **3**:619*T*
ecosystem diversity and stability **3**:618–619
ecosystem redundancy **3**:620–621
equivalent species **3**:620–621, **3**:621*F*
keystone species **3**:621–623, **3**:622*F*
species diversity
herbivores **3**:619–620, **3**:620*T*
primary producers **3**:620
trophic levels **3**:619–620
structural characteristics **3**:619
global species richness **3**:710
global wildlife trade **3**:409*T*
grazer-based ecosystem restoration and management **4**:16
grazing effects **4**:8–9, **4**:9*F*, **4**:10*F*, **6**:40
habitat destruction **3**:404, **3**:405*F*, **3**:447–448
indicator species **7**:206–207*F*
Late Devonian extinction event **5**:169–170
Late Permian extinction event **5**:171
loss consequences
ecosystem goods and services **2**:343
species loss **2**:342–343
marine protected areas **5**:100–101
marine sedimentary habitats **5**:108*T*
mortality agents
algal overgrowth **2**:338
anthropogenic causes **2**:340
coral bleaching **2**:338, **2**:338*F*, **2**:339*F*, **2**:340*F*
corallivore explosions **2**:337–338
cyanide fishing **4**:156
direct destruction **2**:340
disease incidence **2**:338–340, **2**:339*F*, **2**:340*F*
greenhouse gas emissions **2**:341–342, **2**:342*F*
invasive species **2**:341
overfishing **2**:341, **5**:61
poor water quality **2**:340–341
mutualistic partnerships **2**:326, **2**:328
nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
nitrogen cycle impacts **5**:543
overfishing **3**:410, **3**:452–453, **5**:54
predation effects **6**:236
protection and restoration models **5**:465–466
spatiotemporal distributions **2**:330–331, **2**:331*F*
species numbers **5**:108–109
threatened species **7**:201*T*, **7**:206–207*F*
ultraviolet radiation effects **7**:300
Coral Triangle **4**:192–193, **4**:193*F*
Corchorus capsularis **2**:663

Cordaixylon **5**:618*F*
cordgrass **4**:349–350, **4**:360, **4**:365–366, **7**:375–378, **7**:388–389
Cordia **5**:489
Cordia interrupta **4**:364
Cordia sinensis **5**:486
Cordia tetrandra **1**:163
Cordyceps subsellis **3**:638
Cordylanthus **4**:75
Cordylanthus maritimus **6**:455
Coregoninae **6**:522–523
Coregonus oxyrhinchus **6**:479
Coregornus lavaratus **3**:123*T*
Coreidae **7**:294
Corethrellidae **3**:489–491*T*
coriander **2**:662–663
Coriandrum sativum **2**:662–663
Corixidae **7**:389–391
Cormohipparion **4**:724–725*T*
corms
  East Asia **2**:663
  North and South America **2**:663–664
  Southeast Asia **2**:663
  Sub-Saharan Africa **2**:663
corn
  biofuel crop production **1**:529*T*, **1**:530, **1**:543–547, **1**:544*F*
  carbon dioxide (CO_2) emissions **4**:295*F*
  see also ethanol production
corn buntings **4**:259
Cornitermes **4**:409*T*
corn rootworm **4**:365
corn smut **3**:633
Cornu aspersum **3**:222–223
Cornus florida **2**:528*T*, **3**:411, **3**:411*F*
Coromandel monsoon forest **1**:262–263*T*
Coronaviridae **5**:130*F*
Corrugatocephalum **7**:456–457*F*
Corsiaceae **6**:315*T*
Corticaria crenicollis **7**:220*T*
Corticaria obsoleta **7**:220*T*
Corticoviridae **5**:134*T*, **5**:135
Cortispongilla barroisi **5**:470*T*
Corvidae **6**:308*T*
Corvus brachyrhynchos **4**:564
Corvus corax **4**:567
Corvus hawaiiensis **5**:367
Corvus splendens **1**:516, **1**:624
Coryceaus spp. **5**:686–687
Corylus **4**:574*F*
Corylus heterophylla **1**:271–272
Coryndon Museum (Kenya) **5**:406
Coryphaenoides rupestris **2**:271
Coryphantha minima **3**:208
Corystospermaceae **5**:617
Corystosperms **5**:618*F*
Corythosaurus **2**:519*T*, **2**:520*T*
Coscinodiscophyceae **1**:476*F*
Cosmetidae **1**:210–211
Cosmibuena **3**:522*T*
cospeciation **2**:137
Cossoidea **5**:386*F*, **5**:387–388*T*
Cossuridae **7**:417*F*
Costa Rica
  bilateral research agreements **5**:714
  biodiversity research **4**:124–125
  botanical gardens **6**:519
  butterflies **1**:657*T*, **1**:658, **1**:658*T*, **1**:659*T*
  Convention on Biological Diversity (1930) **2**:313
  countryside biogeography studies **2**:354, **2**:355*F*
  deforestation **1**:677*F*, **2**:30–31, **2**:31*F*
  ecosystem services **2**:227, **7**:356*T*
  Instituto Nacional de Biodiversidad (INBio) **1**:506
  invasive species **4**:353–354
  Lepidoptera **5**:389–390
  living dead trees **4**:590–591, **4**:591*F*
  neotropical seasonally dry forests **5**:490, **5**:491*T*, **5**:493*T*, **5**:495*F*, **5**:497*F*
  pastureland conversion **4**:665*T*
  Seamounts Marine Management Area **5**:449–450
  species richness **2**:25*T*
  tropical rain forests
    birds **6**:307–309, **6**:308*T*
    butterflies **6**:306–307, **6**:307*T*
    mammals **6**:309*T*
    plant diversity **6**:314–315, **6**:317*T*, **6**:318*T*
  vascular plant diversity **6**:82*T*
  water funds **7**:356*T*
cost–benefit rule **3**:27
Cosumnes River **7**:151*F*, **7**:152–153, **7**:153
Côte d'Ivoire
  botanical gardens **6**:519
  cropland conversion **4**:665*T*
  forest-savanna mosaics **1**:68
  indigenous peoples **4**:270*T*
  termites **4**:430*T*
Cotingidae **6**:308*T*
Coto Brus Valley, Costa Rica **2**:354, **2**:355*F*
Cotoneaster orbicularis **5**:485
cotton **2**:384*F*
cotton bollworms **4**:298–299, **4**:299*F*
cotton grass **1**:136, **1**:275
cotton stainers **7**:294
cotton-top tamarins **6**:256–258*T*
cottonwood **1**:531–532
cottony cushion scale **1**:78, **4**:361
Coturnicops noveboracensis **7**:182
cotylocidium **7**:463, **7**:464–465*F*
Couepia bracteosa **4**:288*T*
Couepia edulis **4**:288*T*
Couepia longipendula **4**:288*T*
Couepia paraensis **1**:163
Couepia uiti **1**:163–164
Couma utilis **4**:288*T*
countryside biogeography **2:347–360**
  basic concepts
    anthropogenic impacts **2**:348–350, **2**:349*F*
    conservation biogeography **2**:348
    island biogeography **2**:348
    marine and land surfaces **2**:349*F*
    research areas **2**:348–350, **2**:349*F*
    vicariance biogeography **2**:348
  case studies
    ancient tropical countrysides **2**:354–356, **2**:356*F*
    Costa Rica **2**:354, **2**:355*F*
    early studies **2**:353–354, **2**:353*F*
    Hawaiian Islands **2**:356, **2**:357*F*
    marine ecosystems **2**:356
  definition **2**:347
  future outlook **2**:358
  human-dominated ecosystems
    basic concepts **2**:350–351
    human-dominated landscapes **2**:349*F*, **2**:350–351, **2**:350*F*
    human-dominated seascapes **2**:351
    human impacts **2**:350
    trait-based studies **2**:356–357
    untouched ecosystems **2**:350
  human influence
    community assemblage and structure **2**:352
    ecosystem function **2**:352–353
    ecosystem services **2**:352–353
    natural capital **2**:352–353
    population diversity **2**:351–352
    sixth extinction wave **2**:351
    species richness/species diversity **2**:351
  research methodologies
    evolutionary bottlenecks **2**:357–358
    future outlook **2**:358
    mathematical models **2**:358
    remote sensing techniques **2**:358
    theoretical foundations **2**:358
    trait-based studies **2**:356–357
countryside landscapes **5**:321–322
Cousin Island **7**:188
Coussapoa **3**:522*T*
covariance effect **7**:6
Cowan–Farquhar model **2**:675
cowbirds
  brood parasitism
    characteristics **5**:502
    cowbird–host systems
      behavioral adaptations **5**:504–505
      egg adaptations **5**:504
      host defenses **5**:505
      nestling and fledgling adaptations **5**:505
    host defenses
      egg rejection **5**:505–506
      nest access **5**:505
    management and control strategies **5**:507
  number of species **5**:502*T*
  population dynamics models **4**:490, **4**:491*F*
  threatened species **7**:184–185
cowfish **3**:467–471
Cowles' succession theory **7**:64–65
cow-parsnip **6**:120*T*
cowpeas **2**:383, **2**:659, **2**:663, **3**:32*F*
cowpea weevil **4**:365
cows **2**:56*F*
cow-wheat **4**:570–572
coyotes
  boreal forests **1**:631–632
  gene flow barriers **4**:517
  human colonization impacts **1**:621–622
  predation effects
    boreal/temperate forests **6**:237

coyotes (*continued*)
midcontinent North American prairies **6**:239
prey–predator interactions **1**:709
coypu **4**:143, **5**:477
crabeater seals **5**:90
crab-eating frogs **5**:13
crabs
classifications **4**:381*T*
endangered species **3**:201–202, **3**:226
mangrove ecosystems
characteristics **5**:15–16
grapsid crabs **5**:15–16, **5**:15*F*
hermit crabs **5**:16
intertidal gradient **4**:337, **4**:339*F*
mud crabs **5**:16
ocypodid crabs **5**:16, **5**:16*F*
number of species **3**:182*T*, **4**:371*T*
sandy beaches **2**:131–133
threatened species **7**:201*T*
crab spiders **1**:207, **5**:460*T*
Cracidae **6**:308*T*
Crambidae **5**:387–388*T*
cramps **3**:635*T*
cranberries **3**:32*F*
craneflies **3**:489–491*T*, **4**:371*T*, **7**:389–391
cranes **6**:479–480, **7**:181–182, **7**:390
Craniata **7**:336–337
Craseonycteridae **4**:691–692
Crassostrea virginica **2**:264, **2**:528*T*, **3**:303–306, **5**:61
Crassulaceae **1**:486*T*, **5**:622*F*
Crassulacean Acid Metabolism (CAM)
desert plants **2**:470–472
photosynthesis **6**:14*T*, **6**:15
tropical rain forests **3**:113
Crataegus aronia **5**:485
Crataegus sinaicus **5**:485
Crataeva benthamii **1**:163
Craterostigmomorpha **1**:246*T*, **5**:421–423
crayfish
critical pH levels **6**:153*F*
exotic species invasions **4**:378
freshwater ecosystems **3**:578–579, **3**:579*T*
interspecific competition **2**:190
number of species **3**:182*T*, **4**:371*T*
subterranean ecosystems **7**:54–55
threatened species **7**:201*T*
wetland ecosystems **7**:389–391
crazy ants **3**:222, **6**:240–241, **7**:260
Cree culture **4**:162–163, **7**:228
Crematogaster **1**:708, **6**:46–47
Crematogaster nigrceps **6**:47
Crenetermes **4**:409*T*
creodonts **1**:693–694
Creontiades dilutus **1**:376
creosote bush **2**:467*F*, **2**:468–469, **4**:485–486
Crepidostomum **7**:450–451*F*
Crepidula fornicata **5**:61
Crescentia alata **4**:15, **4**:592, **4**:592*F*, **7**:274
Crescentia cujete **4**:281–285, **4**:283*F*, **4**:289*T*
Cressa cretica **5**:484–485
crested honeycreepers **1**:29
crested mangabeys **6**:251*T*
crested tits **1**:611–612
Cretaceous period
Age of the Dinosaurs **5**:5–6
Antarctic ecosystems **6**:617
arthropod radiation ecology **2**:399–400
boundary age **5**:1*T*, **5**:2
Cenomanian/Turonian (C/T) mass extinction **5**:176
Cretaceous–Tertiary (K–T) mass extinction **1**:494, **2**:516, **3**:418–421, **3**:688, **5**:6–7
crustaceans **2**:405
early mammals **3**:188–189
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*, **3**:688
great mass extinctions **5**:168*F*, **5**:169*F*, **5**:173–175, **5**:174*F*
mammals **2**:583*F*, **4**:684
marine diversity **5**:168*F*, **5**:169*F*
mass extinctions
event characteristics **1**:494
extinction intensities **1**:494*T*, **5**:162*F*
lesser mass extinctions **5**:176
marine animals **1**:413*F*, **1**:493*F*, **5**:158
notable mass extinctions **5**:161*F*
number of families **3**:416*T*, **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
originations **1**:495, **1**:495*F*, **3**:425*F*
oxygen-18 ($\delta^{18}O$) measurements **1**:557*F*
periodicity **3**:424*F*
temperate forests **7**:122–124, **7**:123*F*
termite species **4**:410*T*
terrestrial plant communities
abundance patterns **6**:74*F*
evolutionary development **6**:65–67, **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
terrestrial plant evolution **6**:70–71
vegetation–climate interactions **2**:682–684
cribellate silk **1**:206–207
Cricetidae **2**:583*F*, **4**:697
crickets
adaptive radiation **1**:32–33
ant-loving crickets **3**:730
camel crickets **3**:732, **3**:732*F*
cave crickets **3**:221, **3**:732, **7**:51–52, **7**:60
characteristics **3**:722–723, **3**:728, **3**:728*F*
Cooloola Monsters **3**:731
food and feeding habits **3**:724–725
Gryllidae
Brachytrupinae **3**:729–730
characteristics **3**:729
Gryllinae **3**:729
Oecanthinae **3**:730
Phalangopsinae **3**:730
hump-winged crickets **3**:727–728
Jerusalem crickets **3**:731–732
leaf-rolling crickets **3**:732
Mogoplistidae **3**:730
mole crickets **3**:728, **3**:728–729
Myrmecophilidae **3**:730
number of species **4**:392*T*
overharvesting/overcollecting **3**:223
paleontology **3**:725
phylogeny and taxonomy **3**:725
raspy crickets **3**:732
recovery programs **3**:225
reproduction and growth **3**:724
Rhaphidophoroidea **3**:732
scaly crickets **3**:730
Schizodactyloidea **3**:730, **3**:730*F*
sound production **3**:723–724
spider crickets **3**:730
splay-footed crickets **3**:730, **3**:730*F*
Stenopelmatoidea
Anostostomatidae (wetas) **3**:731, **3**:731*F*
characteristics **3**:730–731
Cooloolidae (Cooloola Monsters) **3**:731
Gryllacrididae (raspy crickets and leaf-rolling crickets) **3**:732
Stenopelmatidae (Jerusalem Crickets) **3**:731–732
subterranean ecosystems **7**:60
tree crickets **3**:730
true crickets **3**:729
Crick, Francis **5**:630
Criconemoides **7**:424*F*, **7**:429
crimini mushrooms **3**:637–638, **3**:637*T*
Crinoidea **3**:681*F*, **3**:685*F*, **4**:381*T*, **5**:669–670*T*
Cristataria genezarethana **5**:467, **5**:467*F*
critical source area management **1**:103*T*, **1**:107
croakers **3**:467–471
Croatia
botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
Crocidura suaveolens **2**:552*F*
crocodilians
anthropogenic impacts **6**:411
characteristics **6**:401
conservation status **3**:216*T*, **6**:411
hemoglobin evolution **2**:55*F*, **2**:55*T*
Near East ecosystems **5**:472
reproduction and parental care **6**:405–406
river ecosystems **6**:479–480
threatened species **1**:515–516
Crocodylia **6**:401, **7**:340
Crocodylus acutus **6**:401
Crocodylus niloticus **5**:472
Crocodylus porosus **6**:401
Crocosphaera **2**:610–611, **2**:611*F*, **2**:612*T*, **2**:614, **2**:620*F*, **2**:622*F*
Crocosphaera watsonii **2**:612*T*, **2**:613
Crocus **1**:269
Crocuta crocuta **1**:52, **2**:190, **4**:729*F*
crofelemer **1**:594
Croizat, Leon **2**:539–540, **6**:612–613
Cronartium ribicola **3**:633, **3**:633*T*
cropping systems
agricultural production–ecosystem maintenance relationships **1**:116*F*
animal integration **1**:116
bush-fallow rotation systems **1**:114–115
ecosystem services **1**:114
intercropping **2:382–395**
above-ground species interactions
canopy heat capture **2**:387, **2**:387*F*
characteristics **2**:386–387
light interception/light use efficiency **2**:386–387
shading effects **2**:387, **2**:387*F*

spatiotemporal complementarity **2**:387–388
Africa **2**:382–383
Asia **2**:383, **2**:384*F*
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391
crop pest and disease control
biological pest control **2**:391
genetic diversity effects **2**:391–392
push–pull strategies **2**:391
species diversity effects **2**:392
definition **2**:382–383
importance **2**:382–383
intercropping patterns **2**:383
Latin America **2**:382–383
overyielding
above-ground–below-ground species interactions **2**:386
above-ground species interactions **2**:386–387
below-ground species interactions **2**:388–389
classifications **2**:386
crop pest and disease control **2**:391
land equivalent ratio (LER) **2**:386
summary discussion **2**:392
multifunctional agriculture **1**:115–117, **1**:115*F*, **1**:116*F*, **1**:116*T*, **1**:117*F*
overyielding
above-ground–below-ground species interactions **2**:386
above-ground species interactions
canopy heat capture **2**:387, **2**:387*F*
characteristics **2**:386–387
light interception/light use efficiency **2**:386–387
shading effects **2**:387, **2**:387*F*
spatiotemporal complementarity **2**:387–388
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391
classifications **2**:386
crop pest and disease control
biological pest control **2**:391
genetic diversity effects **2**:391–392
push–pull strategies **2**:391
species diversity effects **2**:392
land equivalent ratio (LER) **2**:386
summary discussion **2**:392
resource conservation **1**:113–114
crop production
biodiversity-based agricultural strategies **1**:123–124
biodiversity-friendly farming **2**:542–544
biodiversity–human health relationship **1**:359–360
commercialization
basic concepts **1**:90–91
seed industries **1**:91
crop genetic diversity **1:85–94**
agricultural intensification **1**:86, **1**:122–123, **1**:374–375, **4**:558, **4**:565–566
agrobiodiversity conversation **3**:33
background information **1**:85–86
basic concepts **1**:87–89
commercialization effects
general discussion **1**:91–92
intensified environments **1**:92
non-intensified environments **1**:92–93
crop genetic resources **3**:39–40
genetic narrowing **1**:89–90, **1**:404–405
genetic trends **1**:90
landrace–modern varieties comparisons **1**:89
research summary **1**:93
seed selection evolution **1**:86–87
spatial distribution **1**:87–89, **1**:88*T*
crop homogeneity **1**:460–461, **1**:461*F*
crop plant pathogens
Australia **1**:76*T*, **1**:81
Brazil **1**:76*T*, **1**:83
fungi **3**:633, **3**:633*T*
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81–82
United Kingdom **1**:76*T*, **1**:80
dryland syndrome **2**:483*F*, **2**:484*F*
ecosystem services modeling **5**:349–350
energy crops **1**:542–543, **1**:542*F*
genetic diversity **1**:360
global agriculture impacts **3**:427, **3**:428*T*, **4**:154
global cropland area **4**:558*F*
hemiparasites **4**:75–77
introduced crop plants **1**:75–76
land-use change impacts **1**:542–543, **1**:542*F*
land-use trends **4**:559*F*
monoculture crops **1**:359–360
organism spillover **3**:13–14
parasitic disease vectors **5**:673
plant–human interactions **6**:57–58
pollination **1**:360, **5**:354–355
polycultures **1**:119–120
pre-agriculture land cover **4**:558*F*
ultraviolet radiation effects **7**:298–299
unplanned diversity **3**:12–13
weed management
biodiversity effects **4**:91–92
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
wetland conservation programs **7**:371
crop wild relatives (CWRs) **6**:77
crossbills
beak shape **1**:604*F*
dispersal biogeography **2**:549–550
crossbred animals **1**:639
crossing-over recombination
classifications **6**:353, **6**:354*F*
crossover distribution **6**:356
environmental effects
feedback regulation **6**:357
genetic variation **6**:356–357
ontogenetic memory **6**:357
transgeneration effect **6**:357
genetic interference **6**:355–356
map distance **6**:355
mapping functions **6**:356
recombination rate **6**:355
sex effects **6**:356
Crossota **2**:21*F*
Cross River gorilla **6**:256–258*T*
Crotalus cerastes **2**:473*F*
Croton **5**:489, **6**:316*T*
Croton lechleri **1**:594
CroV virus **5**:136
crowned sifaka **6**:256–258*T*
crown fires **3**:436
crows **1**:516, **1**:624, **4**:564
Crozet archipelago **1**:180–181, **1**:186*T*
Cruciferae **3**:129*T*
Cruciplacolithus neohelis **1**:474*F*
Crudaria **5**:669–670*T*
Crustacea/crustaceans **2:396–418**
acidification effects **3**:584–585, **3**:584*F*
Antarctica **1**:184*T*
architectural diversity **2**:401–403
arthropod radiation ecology **2**:399–400
biodiversity **2**:396–397
characteristics **4**:392
classifications and sampling methods **4**:381*T*
coral reef ecosystems **2**:332*T*
ecological diversity
biological role **2**:408–409
geographic distribution **2**:413
habitats **2**:407–408
life history patterns **2**:409–413
endangered species **3**:202*T*, **3**:226
extinction vulnerability
estimated extinction rates **4**:659*T*
fossil record **2**:413–415
modern crustaceans **2**:415–418, **2**:416–417*T*
fossil record **2**:403–407
freshwater ecosystems **3**:571–572, **3**:572*T*
genetic diversity **3**:665*F*

Crustacea/crustaceans (*continued*)
growth and development **2**:409–413
known and estimated species **4**:384*T*
lake and pond ecosystems **4**:371–373, **4**:461*T*
latitudinal range shifts **4**:604*T*
major global habitats **2**:397–399
mangrove ecosystems
characteristics **5**:15–16
ecosystem engineers **5**:16–17
grapsid crabs **5**:15–16, **5**:15*F*
hermit crabs **5**:16
intertidal gradient **4**:337, **4**:339*F*
mud crabs **5**:16
number of species **5**:20*T*
ocypodid crabs **5**:16, **5**:16*F*
shrimp **5**:16
marine environments **2**:397–399, **5**:106*F*
mating and reproduction **2**:409–413
modern extinctions **5**:368*T*
Near East ecosystems **5**:460*T*, **5**:470*T*
number of species **2**:397–399, **2**:401*T*, **3**:182*T*, **4**:371*T*, **5**:669–670*T*
phylogenetic relationships **1**:203*F*
plankton fraction **6**:25
river ecosystems **6**:478–479, **6**:478*T*
sandy beaches **2**:131–133
taxonomic diversity **2**:400–401, **2**:401*T*
wetland ecosystems **7**:389–391
crustal elements
biological assimilation **1**:554–555
biological transport **1**:554
carbon pools **1**:554, **1**:555*T*
elemental partitioning **1**:553, **1**:554*T*
hydrological cycle **1**:553–554, **1**:555*F*
origins **1**:552–553
orogenesis **1**:554
phase state transitions **1**:553
relative abundances **1**:553*F*
vulcanism **1**:554
weathering processes **1**:553
crusts, desert **2**:462–463
cryopreservation **1**:645–646, **1**:646, **3**:651–652
Cryphonectria parasitica **2**:528*T*, **3**:628–629, **3**:633, **3**:633*T*, **4**:361–362
Cryptantha micrantha **2**:373*T*
Cryptochaetidae **3**:489–491*T*
Cryptochloris wintoni **1**:487–488
Cryptochloris zyli **1**:487–488
cryptococcosis **3**:634*T*
Cryptococcus neoformans **3**:634*T*
Cryptodira **6**:400–401
cryptofauna **2**:332*T*
Cryptomeria **1**:272–273
Cryptomeria japonica **1**:273, **7**:188
cryptomonads **6**:25, **6**:294*F*
Cryptomonas **6**:294
Cryptomonas ovata **6**:33*F*
Cryptophagus confusus **7**:220*T*
Cryptophagus subdepressus **7**:220*T*
Cryptophyceae **6**:478*T*
cryptophycin **6**:131, **6**:131*F*
Cryptophyta **3**:571*T*, **6**:10*T*
Cryptoprocta ferox **1**:55–56, **4**:693
Cryptoprocta spelea **4**:657, **4**:728–729
Cryptops hortensis **5**:425
Cryptorhynchini **1**:328
cryptosporidium **5**:670
Cryptosporidium parvum **6**:297
Cryptotermes **4**:409*T*
crystalline ice plants **4**:360
Ctenizidae **1**:206
Ctenomyidae rodents **7**:432–433, **7**:432*T*
Ctenomys mendocinus **3**:127*T*
Ctenomys opimus **7**:432–433
Ctenomys talarum **2**:587*F*
Ctenopharyngodon idella **4**:361, **4**:364
Ctenophora
fossil record **3**:684*F*
global diversity **4**:385*T*
mangrove ecosystems **5**:20*T*
marine ecosystems **5**:50*T*
number of species **5**:669–670*T*
Ctenostoma aspendum **1**:250*T*
Ctenostoma formicarium **1**:250*T*, **1**:251*F*
Cthamalus stellatus **6**:232
Cuba
botanical gardens **6**:519
endemic plant species **6**:83
lizards **1**:31*F*, **1**:32
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
protected area systems **6**:493–494
solenodons **4**:713–714, **4**:714*F*
Cuban kite **7**:185
Cubitermes **4**:409*T*
cubiu **4**:284–285, **4**:284*F*
cuckoo-dove **2**:189–190, **2**:189*F*
cuckoo-finches **5**:502, **5**:502*T*
cuckoos
brood parasitism
characteristics **5**:501–502
cuckoo–host systems
adaptive behaviors **5**:503
egg adaptations **5**:503
evolutionary development **5**:503
fecundity **5**:503
host defenses **5**:504
nestling and fledgling adaptations **5**:503–504
host defenses
egg rejection **5**:504
mobbing behavior **5**:504
nestling discrimination **5**:504
climate change effects **2**:86
number of species **5**:502*T*
Cucujidae **1**:327*T*
Cucujoidea **1**:327*T*
Cuculidae **5**:501–502, **5**:502*T*, **6**:308*T*
Cuculiformes **6**:308*T*
Cuculus canorus **2**:86
cucumber **2**:663
Cucumeropsis mannii **6**:553
Cucumis anguria **2**:663
Cucumis melo **2**:662–663, **2**:663
Cucumis sativus **2**:663
cucura **4**:287–290
Cucurbita **2**:383
Cucurbitaceae **3**:129*T*, **6**:316*T*
cucurbitacins **2**:427
Cucurbita ficifolia **2**:663–664
Cucurbita foetidissima **4**:248*F*
Cucurbitales **5**:621–625, **5**:623*F*
Cucurbita maxima **2**:663–664
Cucurbita mixta **2**:663–664
Cucurbita moschata **2**:663–664, **3**:129–130
Cucurbita okeechobeensis **3**:208
Cucurbita pepo **2**:663–664, **6**:553
Cucurma longa **2**:663
Cueva de Felipe Reventón (Canary Islands) **7**:59*T*
Cueva del Sobrado (Canary Islands) **7**:59*T*
Cueva del Viento (Canary Islands) **7**:59*T*
Culex mimeticus **5**:463*T*
Culex pipiens **4**:301–302
Culex quinquefasciatus **4**:663–664
Culicidae **3**:489–491*T*, **3**:492, **5**:463*T*
Culicoidea **3**:489–491*T*
Culicomorpha **3**:489–491*T*
cultivated plants/cultivars **5**:551, **5**:553*T*
cultural ecosystem services **1:332–340**
basic concepts **1**:333–335, **1**:333*F*, **2**:490*T*
cultural diversity effects **1**:335–336, **1**:335*F*
cultural services **1**:334*T*
desertification **2**:490*T*
esthetic services **5**:353–354
general discussion **1**:333
habitat services **1**:334*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:353
nature-based recreation models **5**:353
provisioning services **1**:334*T*
regulating services **1**:334*T*
spiritual services **5**:354
sustainable development **1**:337–339
valuation language/value articulating institutions **1**:336–337, **1**:337*T*
Culture Collection, University of Goteborg (CCUG) **3**:652*T*
Cumacea
classifications **4**:381*T*
extinction vulnerability **2**:416–417*T*
Cumberland River system **5**:562–563
cumin **2**:662–663
Cuminum cyminum **2**:662–663
Cumulative Impacts model **5**:460–462*T*
Cunninghamia **1**:272–273
Cunninghamia lanceolata **1**:272
Cunoniaceae **3**:522*T*, **5**:621–625
Cuon alpinus **5**:603
cup fungi **3**:628–629
cuphea **1**:529*T*
Cupressus **1**:268–269
Cupressus chimera **6**:695
cuprobam **5**:695
cupuaçu **4**:290
curacin A **6**:130–131, **6**:130*F*
curare poisons
hunting and fishing **2**:431–432
medicines **2**:432
Curatella americana **1**:165, **7**:274
Curculionidae **1**:327*T*, **1**:328, **4**:347
Curculionoidea **1**:327*T*
Curetinae **1**:652
current global land cover **4**:557*F*, **4**:566*F*
Curtis Prairie (Wisconsin) **6**:441–442

Curtonotidae **3**:489–491*T*
Curvilinear Glasshouse **6**:506–508
curvina **3**:260
Curvitermes **4**:409*T*
cuscus **4**:687*T*
Cuscuta **4**:71
Cuscuta campestris **4**:77
Cuscutaceae **4**:71
cush-cush yam **2**:663–664
cushion bogs **4**:105
cuskeels **3**:466–467
cutaneous melanoma **7**:302
cutaneous mycoses **3**:634*T*
Cutervodesmus adisi **5**:431
cutthroat trout **6**:522, **6**:526
Cuvier, Georges **3**:393
Cuvieronius **4**:724–725*T*, **4**:727*T*, **5**:231–233
Cuvieronius hyodon **5**:225*T*
Cuyabeno, Ecuador **1**:159*T*
cyanide fishing **4**:156
Cyanidium caldarium **4**:114–115
Cyanistes caeruleus **5**:720
cyanobacteria
  Census of Marine Life project **2**:18–19
  characteristics **5**:685*F*, **6**:25–26
  energy flow **3**:246–247
  eutrophication impacts **1**:95–97, **3**:352–354, **3**:353*F*, **3**:356*F*, **3**:357*F*
  evolutionary development **2**:640–641
  freshwater ecosystems **3**:571–572, **3**:571*T*
  nitrogen fixation **2**:611*F*, **2**:614
  nomenclature code **5**:555
  oxygen cycle **1**:557–558
  pharmacologically active compounds **5**:711
  photosynthesis mechanisms **6**:9–11, **6**:10*T*, **6**:14*T*
  photosynthetic symbioses **5**:118–120
  phylogenetic classifications **4**:673–675
  plant-derived drugs and chemicals
    anticancer agents **6**:130–131
    cryptophycin **6**:131, **6**:131*F*
    curacin A **6**:130–131, **6**:130*F*
  river ecosystems **6**:477–478, **6**:478*T*
  temperature tolerance ranges **4**:113–115
Cyanolimnas cerverai **7**:184
Cyanophyta **3**:571*T*
Cyanopsitta spixii **7**:187
Cyanothece **2**:611*F*, **2**:613, **2**:614
Cyathea **6**:59, **6**:316*T*
Cyatheaceae **6**:59
cyathobodonids **6**:294
Cyathocotyle bushiensis **4**:362
cybertaxonomy **7**:94
Cycadaceae **5**:618*F*
Cycadeiodales **5**:613, **5**:615–617, **5**:617*F*, **5**:618*F*
Cycadeoidea **5**:618*F*
cycads **6**:58*T*, **6**:59–60, **7**:201*T*
cycad seed flour **1**:592
Cycas **1**:70
Cycas micronesica **6**:84
Cyclanthaceae **3**:522*T*, **6**:316*T*
Cyclidium **6**:295, **6**:296*F*
Cyclobalanopsis **1**:267
Cyclocoeloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
cyclodienes **5**:693
cyclohexanedione **4**:90*T*
Cycloneuralia **7**:433–435, **7**:433*F*
cyclopeptides **3**:635*T*
Cyclophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
cyclopoids **6**:25
Cyclorrhapha **1**:710–711, **3**:489–491*T*, **3**:491–492
Cyclosquamata **3**:465–466
Cyclotella **6**:30
Cyclotella caspia **6**:30
Cyclotella meneghiniana **3**:356*F*
Cyclothone pseudopallida **4**:386*T*
Cyclura collei **3**:217–218
Cydnidae **7**:294
cyfluthrin **5**:694
Cygnus atratus **3**:365–366, **5**:221
Cygnus cygnus **3**:365–366, **7**:36
Cygnus sumnerensis **5**:366
cyhexatin **5**:694
Cylindera morio **1**:250*T*, **1**:251*F*
Cylindera suturalis **1**:250*T*, **1**:252*F*
Cylindrocarpon lucidum **5**:710–711
Cylindrocopturus **1**:329
Cylindroiulus latestriatus **5**:428
Cylindroiulus madeirae **5**:428
Cylindrospermopsis **6**:33–34
Cylindrospermopsis raciborskii **6**:30
Cymatium parthenopeum **2**:543
Cymbopogon **1**:70
Cymodocea **6**:542*T*
Cymodocea angustata **6**:541*T*
Cymodocea nodosa **6**:541*T*
Cymodocea rotundata **6**:541*T*
Cymodocea serrulata **6**:541*T*
Cymodoceoideae **6**:542*T*
cymoxanil **5**:695
Cynipidae **4**:181
Cynipodea **4**:178
Cynogale bennettii **4**:693
Cynometra alexandri **7**:273
Cynomoriaceae **4**:73
Cynomys spp. **3**:188, **3**:196, **6**:342
Cyperaceae **2**:371*T*, **2**:374*T*, **3**:234–235, **6**:316*T*
Cypereae **2**:376*F*
Cyperus **1**:278–279
Cyperus ferax **1**:165
Cyperus giganteus **1**:164, **1**:165
Cyperus luzulae **1**:165
Cyperus mutisii **1**:164
Cyperus papyrus **1**:53–54, **2**:370–372, **5**:486, **7**:388–389
Cyphoderris strepitans **3**:728*F*
Cyphophthalmi **1**:210, **1**:246*T*
Cyphotilapia **1**:25*F*
cypresses **1**:268–269, **7**:397
Cypria javensis **6**:26
Cypria petenensis **6**:26
Cyprinidae **3**:464
Cyprinodon **4**:115
Cyprus
  animal species **1**:261*T*
  ecosystems
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  wheat production **1**:87–88
Cypselea **2**:376*F*
Cypselosomatidae **3**:489–491*T*
Cyrpinus carpio **3**:478*F*
Cyrpus **6**:507*T*
Cyrtobagous salviniae **6**:669
Cyrtobagus salviniae **4**:347
Cyrtocara **1**:25*F*
Cyrtodactylus cavernicolus **1**:515–516
Cyrtodactylus thirakhupti **1**:515–516
cyrtophorids **6**:295–296, **6**:296*F*
Cystiplana **7**:442–443*F*
Cystophora cristata **5**:90, **5**:314, **7**:38
Cytinaceae **4**:72
Cytinus **4**:72
Cytisus **1**:269
cytochrome P450 **3**:118, **3**:122, **3**:123*T*
Cytophaga **6**:32
Czech Republic
  botanical gardens **6**:519
  vegetation maps **3**:346*T*
  vegetation surveys **3**:341–344*T*

D

dace **1**:10*T*
Dacnitis truttae **5**:670
Dacnonypha **5**:385–386, **5**:387–388*T*
Dacrydium **7**:118–119
Dactyella **2**:644
Dactyladenia barteri **6**:561
Dactylanthus **4**:194*T*
Dactylaria **2**:644
Dactylaria gallopava **4**:114
Dactylis glomerata **1**:269, **6**:459
Dactylogyridea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Dactylopius ceylonicus **4**:351
Dacus oleae **5**:469–470
daddy long legs
  characteristics **1**:209–210, **1**:209*F*
  ecological characteristics **1**:210
  endangered species **3**:225–226
  genera and species **1**:202*T*
  phylogenetic relationships **1**:205*F*
  phylogeny and taxonomy **1**:210
  reproduction and growth **1**:210
dalapon **5**:696
Dalbergia melanoxylon **1**:67
Dalbergia nigra **6**:560
Dalbergia sisoo **5**:486
Dal Lake **1**:276*T*, **1**:279
Dall porpoises **3**:192–193
Dalyellioida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Dama dama mesopotamica **5**:472
Damaliscus lunatus **1**:52
dambo **7**:369*T*
dams
  biodiversity impacts **7**:203*F*

dams (*continued*)
biotic impoverishment **3**:280–281
ecosystem services modeling **5**:351
endangered freshwater invertebrates **3**:180–181
endangered mammals **3**:192
floodplain agriculture **7**:146
freshwater invertebrates **4**:377–378
habitat alteration **3**:573–574, **4**:154
landscape genetics **4**:515
river–floodplain connectivity **7**:150–151, **7**:151*F*, **7**:152–153
river threats **6**:480–481, **6**:481*T*
salmon population **6**:528–529
South American ecosystems **6**:609
Southeast Asia **1**:510–512
damsel bugs **7**:294
damselfish **3**:467–471, **3**:474*F*, **4**:10*F*
damselflies
lake and pond ecosystems **4**:371–373
life cycle **4**:370
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
systematic classifications **7**:91–93, **7**:92*F*
threatened species **7**:201*T*
wetland ecosystems **7**:389–391
Danainae **1**:653–654
Danau Bankau **1**:276*T*
Danaus plexippus **1**:656, **2**:141–142, **3**:222, **5**:316
dance flies **3**:489–491*T*, **3**:492, **3**:495
dandelion **3**:131–132
dangerous anthropogenic interference (DAI) threshold **4**:21–22
Danian stage **5**:174*F*
Daniellia oliveri **1**:50
Danio rerio **6**:361*T*
Dansgaard–Oeschger events **2**:682
Dante's inferno events **2**:516, **4**:650–651
Danube River **7**:147
Daphnia
biomonitoring processes **3**:122*T*
characteristics **6**:26
environmental requirements **6**:31
infectious disease regulation **3**:349
species packing **4**:644
Daphnia dentifera **2**:726–727
Daphnia galeata mendotae **1**:10*T*
Daphnia hyalina **6**:31
Daphnia magna **2**:143, **4**:255
Daphniphyllum **5**:622*F*
dark-eyed juncos **5**:309–310
darkling beetles **1**:327*T*, **1**:328
darkrumped petrel **4**:353
dark-winged fungus gnats **3**:489–491*T*
Darlingtonia **6**:44*T*
Darlington, Philip **2**:539–540
darters **3**:467–471
Dartmoor National Park **5**:445
Darwin, Charles **2**:**419–425**
adaptive radiation **1**:21
Beagle expedition **2**:419–420, **4**:651–652
biodiversity study contributions
competitive natural economy theory **2**:421–423
scientific classification systems **2**:423–424
biogeographic dispersal **1**:578
childhood **2**:419
coevolution theory **2**:138–139
common ancestry theory **3**:395–396
Darwin's five theories
common ancestry theory **3**:395, **3**:397*F*
evolutionary theory **3**:395–396
gradualism **3**:395–396
natural selection **3**:396
populational change **3**:396
dispersal theories **2**:539–540
ecological complexity **2**:201, **2**:202–204
ecosystem processes **3**:588
educational background **2**:419
entangled bank metaphor **2**:202–204
evolutionary theories **2**:46–47, **2**:421, **3**:393–394, **3**:393*F*, **5**:628, **6**:2, **7**:343–344
family life **2**:421
food webs **3**:497–498
general discussion **2**:419
inbreeding depression **4**:245–246
invasive species **2**:590
later publications **2**:421
natural classification systems **2**:49–51, **2**:49*F*, **4**:121, **7**:101–102
natural selection theory **1**:15–16, **1**:695, **2**:263–265, **2**:420–421, **3**:393–394, **4**:651–652
organism-level diversity **2**:638
Origin of the Species **2**:421, **2**:423
post-Darwin research
alternative Darwinism **2**:424
contemporary Darwinian theory **2**:425
early Darwinians **2**:424
genetic studies **2**:424–425
Southern Hemisphere ecosystems **6**:612–613
speciation criteria **6**:631–632, **6**:680–681, **7**:45
taxonomic methods **7**:89
Darwin Core **1**:440–441, **1**:443
Darwinia **5**:221
Darwin's finches *see* Galapagos finches
Daspletosaurus **2**:519*T*, **2**:520*T*
Dasypodidae **4**:727*T*, **6**:309*T*
Dasyprocta fuliginosa **4**:287
Dasyprocta leporina **4**:287, **7**:273
Dasyprocta prymnolopha **4**:287
Dasyprocta punctata **2**:587*F*, **4**:595
Dasyproctidae **6**:309*T*
Dasypus novemcinctus **2**:586
Dasytricha **6**:295–296
Dasyuridae **1**:686, **4**:687–688, **4**:701*T*
Dasyuromorphia **4**:682*F*, **4**:687–688, **4**:687*T*
Dasyurus maculatus **4**:688, **4**:704
date palms **2**:662–663, **3**:32*F*
Datura stramonium **5**:712*T*
Daubentoniidae **6**:250–252, **6**:251*T*
Daucus carota **2**:662–663
Davao Gulf **1**:276*T*
Davidiaceae **1**:260*T*
Davidia involucrata **3**:373–374
Davidsonia **5**:623*F*
Daviesia divaricata **6**:454
dawn redwood **2**:47
Dean's parameter (Ω) **2**:128
death **3**:635*T*
Decapoda
lake and pond ecosystems **4**:371–373
number of species **3**:182*T*, **4**:371*T*
decay index **6**:19
Deccan thorn forest **1**:262–263*T*
Deccan Traps **5**:175
deciduous forests
biodiversity change scenarios **4**:6*F*
climate change studies **2**:61*F*
current global land cover **4**:557*F*
global agriculture impacts **3**:428*T*
litter deposition **6**:120*T*
net primary production (NPP) **1**:560*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
temperate deciduous forests **1**:271
tropical deciduous forests **3**:404
Decticus verrucivorus **3**:225, **3**:225*F*
Deep Creek (Hand) **4**:654
deep ecology **3**:274–276, **3**:275*T*
deep ocean environments
benthic ecosystems **5**:53
diversity regulation **5**:110–111
fish diversity **3**:475–476, **3**:475*F*
global diversity patterns **5**:109
marine sedimentary habitats **5**:108, **5**:108*T*
pelagic ecosystems **5**:49*F*, **5**:52–53
species numbers **5**:108–109
deep-sea fish **2**:12*F*
deep-sea hydrothermal vents **7**:**327–332**
chemosynthetic symbioses **5**:120–121, **5**:120*F*
ecosystem diversity **5**:53, **5**:53
Endeavor Hydrothermal Vents Marine Protected Area **5**:447
energy flow **3**:242
genetic diversity **5**:50
invertebrate diversity
biodiversity variability **7**:331
biogeography **7**:330
gene flow/genetic diversity **7**:330–331
prevalence **7**:329–330
species origins **7**:330
marine sedimentary habitats **5**:108, **5**:108*T*
microbial biodiversity
bacterial–invertebrate symbioses **7**:328–329
free-living microorganisms **7**:328
occurrences and characteristics **7**:327–328
see also hyperthermophiles
deep-soil habitats **7**:51
Deepwater Horizon mobile offshore drilling disaster **2**:14, **3**:255*T*, **3**:256, **4**:654
deer
boreal forests **1**:631–632
characteristics **4**:696
grassland ecosystems **6**:332
Great American Biotic Interchange (GABI) **2**:584*F*
introduced species **1**:76–77
late Quaternary extinction event (LQEE) **4**:724–725*T*
mangrove ecosystems **5**:14

number of species **4**:690*T*
oil-palm plantations **5**:604
risk status classification **1**:644*F*
deer flies **3**:489–491*T*, **3**:492
deer mice **1**:365–368, **1**:366*F*, **4**:363
defensive responses **1**:39–40
deforestation **2:447–456**
Africa **2**:31*F*
biodiversity impacts **4**:144, **4**:663
biodiversity-rich countries **1**:501–502, **1**:502*T*
carbon cycle **1**:676–678, **1**:677*F*, **1**:679*T*
carbon dioxide (CO_2) emissions **4**:294–295
causal factors
anthropogenic impacts **4**:663, **6**:79–84
direct causes
cattle ranching **2**:453
characteristics **2**:452
charcoal production **2**:453
commercial agriculture **2**:452–453, **2**:452*T*
fuelwood collection **2**:453
infrastructural development/industrial projects **2**:453
livestock grazing **2**:453
logging operations **2**:453–454, **2**:454*T*
plantation crops **2**:453
shifting cultivation **2**:452, **2**:452*T*
wildfires **2**:454
general discussion **2**:449–450
underlying causes
characteristics **2**:450
development policies **2**:450
forest product exports **2**:450–451, **2**:451*T*
forests/forest products undervaluation **2**:451
market demands **2**:450–451, **2**:451*T*
population growth/poverty **2**:450
resettlement programs **2**:450
tax incentives **2**:450
tenurial policies **2**:450
weak government policies and regulations **2**:451–452
Central America **2**:30–31, **2**:31*F*, **5**:491–492
commercial farming industry **4**:144
consequences
biodiversity loss **2**:454
economic losses **2**:455
ecosystem services disruption **2**:454
global climate change **2**:454–455
social impacts **2**:455
contemporary deforestation **2**:448–449, **2**:448*T*, **2**:449*T*
definition **2**:447–448
deforesting countries **2**:448*T*, **2**:449*T*
ecological impacts **1**:676–678, **1**:677*F*, **1**:679*T*
future outlook
agricultural expansion **2**:455
deforestation reduction strategies **1**:682–683
economic opportunities **2**:455–456
logging operations **2**:455
protected area systems **2**:455
tree plantations **1**:682, **2**:456
global forest area changes **2**:448*T*, **2**:449*T*
historical perspective **2**:447–448
human impacts **4**:153–154
institutional failure **7**:170–172
landscape legacies **4**:525
land use issues **4**:558
living dead **4**:591–593, **4**:592*F*
neotropical seasonally dry forests **5**:491–492, **5**:494*F*
Oceania **2**:31*F*
oil-palm plantations **5**:607, **6**:320
Reducing Emissions from Deforestation and Forest Degradation (REDD) **1**:682–683, **2**:246–247, **2**:455, **4**:125
research summary **2**:456
South America
Amazon ecosystems **6**:338–339
annual forest change **2**:448*T*
annual forest loss **2**:449*T*
deforestation rates **1**:679*T*
economic factors **1**:677*F*
forest product exports **2**:451*T*
geographic extent **2**:454*T*
historical perspective **2**:447–448
land use change **2**:31*F*
major threats **6**:608–609
neotropical seasonally dry forests **5**:491–492
total forest change **2**:452*T*
Southeast Asia **1**:509–510, **1**:677*F*, **2**:447–448, **2**:448*T*, **2**:449*T*, **4**:562–564, **4**:563*F*
tropical forests
annual forest change **2**:448*T*
C_4 plants **2**:379–381, **2**:380*F*
causal factors **2**:379–381, **2**:380*F*
contemporary deforestation **2**:448–449
deforestation rates **7**:277
ecological impacts **1**:677*F*, **1**:679*T*
habitat destruction **3**:164–165
livestock impacts **6**:338–340, **6**:340*F*
regeneration systems **3:277–286**
challenges **7**:284–285
chronosequence studies **7**:279–280
forest dynamics **7**:278, **7**:279*T*
functional diversity (FD) **7**:282
historical perspective **7**:277–278
prior land use effects **7**:281
recovery processes **7**:281–282
regeneration sources **7**:280–281
seed dispersal–predation relationship **7**:283–284
soil fertility **7**:281
species composition changes **7**:282–283, **7**:283*F*
successional continuum **7**:278
successional trajectories and stages **7**:278–280, **7**:279*T*
tropical rain forests
conservation efforts **6**:320
diversity implications **6**:327–328, **6**:328*T*
model-based assessments **6**:323, **6**:325*T*
1990 remote sensing survey **6**:323
2000 remote sensing survey **6**:323–325
degraded lands **1**:543, **1**:544*F*
dehesa landscapes **4**:574–577
Deinacrida heteracantha **3**:731
Deinacrida rugosa **3**:724
Deinotheriidae **5**:224*T*
Deinotherium **5**:224*T*
Delacour's langur **6**:256–258*T*
Delaware Bay Estuary **5**:28–29, **5**:28*T*, **5**:29*F*
Delaware Bay salt marsh restoration (case study) **7**:376–378, **7**:377*F*
Delhi Sands giant flower-loving flies **3**:496
DeLillo, Don **4**:653, **4**:654
Delonix regia **1**:55–56
Delphinapterus leucas **3**:192–193, **4**:693–694, **5**:314
Delphinium **1**:270
Delphinium confusum **1**:269
delta diversity **6**:534
deltamethrin **5**:694, **5**:699*T*
deltas
conservation status **3**:171–173*T*
definitions and terminology **7**:369*T*
De Materia Medica **5**:705–706
demecolcine **5**:712*T*
demersal fish stocks **3**:482
de Mestral, George **1**:590
demeton-S-methyl **5**:697*T*
Democratic Republic of Congo
botanical gardens **6**:519
deforestation **1**:677*F*, **2**:449*T*, **2**:451*T*
education level–infant mortality relationship **6**:205*F*
endemism **3**:232*T*
forest product exports **2**:451*T*
forest-savanna mosaics **1**:68
primate diversity **6**:252*T*
primate endemism **6**:252*T*
species richness **1**:499*T*
tropical forests **7**:269–270, **7**:270*T*
Demodex **1**:215
Demodex folliculorum **5**:664–665
demographic heterogeneity **6**:178
demographic overfishing **3**:445
demographic transition **4**:231–232, **6**:199–200
Demospongia **3**:685*F*
denaturing/thermal gradient gel electrophoresis (D/TGGE) **5**:264, **5**:264*T*, **5**:265*F*
Dendrobium **6**:316*T*
Dendrobium nobile **3**:208
Dendrobium pauciflorum **3**:208
Dendrobranchiata **2**:416–417*T*
Dendrocalamus strictus **1**:69–70
Dendrocolaptidae **4**:33, **4**:34*F*, **6**:308*T*
Dendrocopos medius **4**:659
Dendroctonus ponderosae **2**:76–77
Dendroctonus rufipennis **4**:218–219
dendrogram-based measures **3**:599*F*
Dendroica adelaidae **1**:618
Dendroica chrysoparia **4**:516–517, **7**:183
Dendroica fusca **5**:309–310
Dendroica kirtlandii **2**:58–59, **5**:370, **7**:184–185

Dendroica striata **5**:311–312
Dendrophthoe falcata **4**:77
Dendrophthoe pentandra **4**:77
Dendropsophus elegans **5**:425
Dendrosicyos socotranus **1**:51–52
denitrification **3**:90*T*, **6**:151–152, **6**:151*F*
Denmark
  botanical gardens **6**:507*T*
  freshwater ecosystems
    anthropogenic impacts **3**:575
    eutrophication **3**:575*F*
    Lake Fure **3**:575*F*, **3**:581–582
    species declines **3**:578–579, **3**:579*T*
    species variability **3**:578*F*
    threatened species **3**:579–580
  nitrogen deposition **5**:529*F*
  nitrogen fertilizer input **3**:576*F*
  Skjern River Restoration Project, Denmark (case study) **7**:373–374, **7**:375*F*
  vegetation surveys **3**:341–344*T*
Dentalium **5**:378–379
Dentispicotermes **4**:409*T*
Deontostoma timmerchioi **7**:421
deoxyribonucleic acid (DNA)
  bioprobes **3**:121–122, **3**:122*F*
  bioprospecting **1**:593
  chloroplast DNA (cpDNA) **3**:660, **3**:666–667
  chromosomal and gene mutations **3**:658*T*
  circular DNA molecules **1**:325
  cloning and sequencing **5**:264*T*, **5**:265–266, **5**:266*F*
  comparative genomics
    basic concepts **3**:664
    protein polymorphism **3**:663
  DNA banks **3**:652, **4**:319–320, **4**:321*T*
  DNA barcoding programs **2**:13–14, **5**:412–413
  DNA diversity
    chloroplast DNA (cpDNA) **3**:666–667
    general discussion **3**:666
    mitochondrial cytochrome b **3**:666
    mitochondrial DNA (mtDNA) **3**:666
    simple sequence repeats (SSRs) **3**:667
    single nucleotide polymorphisms (SNPs) **3**:667
  DNA fingerprinting
    Pfu DNA polymerase **1**:593
    Taq polymerase **1**:593
  DNA polymerase α **5**:569–570, **5**:571*F*
  DNA polymorphism
    amplified fragment length polymorphisms (AFLPs) **3**:663–664
    microsatellites **3**:663, **3**:671*F*
    nucleotide polymorphism **3**:664
    polymerase chain reaction (PCR) analysis **3**:663
    random polymorphic DNA (RAPD) fragments **3**:663, **3**:671*F*
    restriction fragment length polymorphisms (RFLPs) **3**:663
  DNA probe arrays **2**:694*F*, **2**:697
  DNA repair and recombination **6**:358, **6**:367
  DNA replication processes **2**:627–628
  DNA sequencing **2**:627, **5**:254
  double-stranded DNA (dsDNA) analysis **5**:262–264, **5**:263*F*, **5**:264*T*
  eukaryotic genomes/chromosomes **3**:657–659, **3**:658*F*, **6**:360–361, **6**:361*T*
  evolutionary perspective **5**:703–705
  evolutionary rates **2**:56–57, **2**:56*F*
  genetic code **3**:656–657, **3**:657*T*
  homologous recombination **1**:322–324, **1**:324*F*
  hybridization **1**:26–27, **4**:166–167
  information transfer **3**:656–657, **3**:656*F*
  landscape genetics **4**:510
  mammalian evolution **4**:685
  marine viruses
    DNA-containing viruses **5**:142
    dsDNA viruses
      Adenoviridae **5**:134*T*, **5**:135
      Alloherpesviridae **5**:134*T*, **5**:135
      Caudovirales **5**:130*F*, **5**:134–135, **5**:134*T*
      characteristics **5**:134–135, **5**:134*T*
      Corticoviridae **5**:134*T*, **5**:135
      Herpesvirales **5**:130*F*, **5**:134*T*, **5**:135
      Herpesviridae **5**:134*T*, **5**:135
      Iridoviridae **5**:134*T*, **5**:135
      Malacoherpesviridae **5**:134*T*, **5**:135
      Mimiviridae **5**:130*F*, **5**:134*T*, **5**:135
      Myoviridae **5**:130*F*, **5**:134–135, **5**:134*T*
      Nimaviridae **5**:130*F*, **5**:134*T*, **5**:135–136
      Phycodnaviridae **5**:134*T*, **5**:136
      Podoviridae **5**:130*F*, **5**:134*T*, **5**:135
      Poxviridae **5**:134*T*, **5**:136
      Siphoviridae **5**:130*F*, **5**:134*T*, **5**:135
      unassigned viruses **5**:134*T*, **5**:136
    ssDNA viruses
      Anelloviridae **5**:134*T*, **5**:136
      characteristics **5**:134*T*, **5**:136
      Circoviridae **5**:134*T*, **5**:136
      Inoviridae **5**:130*F*, **5**:134*T*, **5**:136
      Microviridae **5**:134*T*, **5**:136
      unassigned viruses **5**:134*T*, **5**:136–137
  mitochondrial cytochrome b **3**:666
  mitochondrial DNA (mtDNA)
    animal genome **3**:659–660
    evolutionary significant unit (ESU) criteria **1**:668–669
    genetic diversity **3**:666
    plant genome **3**:660
  museum research programs **5**:412–413
  mutations **3**:658*T*, **3**:660–661
  noncoding repetitive DNA sequences **3**:659
  nucleic acids **3:568–580**
    gene scrambling
      basic principles **5**:568
      ciliate conjugation **5**:568, **5**:569*F*
      genome rearrangements **5**:569–570, **5**:571*F*
      internal eliminated sequences (IESs) **5**:569, **5**:571*F*
      macronuclear destined sequences (MDSs) **5**:569, **5**:570*F*, **5**:571*F*
      macronuclear formation **5**:568–569, **5**:570*F*
      unscrambling process **5**:569–570, **5**:570*F*, **5**:571*F*
    nucleotide sequences **5**:568, **6**:183
  organelles **3**:659–660
  prokaryotic genomes **3**:657, **3**:658*F*
  protein polymorphism analysis
    comparative genomics **3**:663
    technical innovations **3**:663
  proteorhodopsins **6**:280–282
  ribosomal DNA (rDNA) **5**:254, **5**:259–260, **5**:292, **5**:296
  simple sequence repeats (SSRs) **3**:667
  single nucleotide polymorphisms (SNPs) **3**:667
  small subunit ribosomal DNA (ss rDNA) **6**:345
  structural characteristics **3**:655–656, **3**:655*F*
  terminal restriction fragment length polymorphism (T-RFLP) **5**:254
  ultraviolet radiation effects **7**:298
  see also recombination
Department for International Development (UK) **2**:231–232, **7**:19
De Plantis Libri **7**:89–90
Depressaria pastinacella **2**:139, **2**:143
depth gradients **5**:273–274
Dermaptera **4**:308–312, **4**:309*T*, **4**:392*T*
Dermatophagoides spp. **1**:215
Dermochelys coriacea **2**:9*F*, **3**:217, **3**:218, **3**:256, **6**:400–401
Dermoptera **4**:690*T*, **4**:691, **4**:701*T*
Deschampsia antarctica **1**:182
Deschampsia flexuosa **1**:142, **1**:634
Descurania pinnata **2**:373*T*
Deseado Massif **3**:105–107, **3**:106*F*
deserpidine **5**:712*T*
desert big-horn sheep **3**:193
desert crusts **2**:462–463
desert ironwood **6**:455
desert long-horned grasshoppers **3**:733–734
desert pavement **2**:462
deserts **2:457–478**
  Anatolian–Iranian desert **1**:262–263*T*
  anthropogenic interactions **2**:475–477, **2**:476*T*
  Arabian desert **1**:262–263*T*
  biodiversity change scenarios **4**:6*F*
  biological adaptations
    animal species **2**:471*F*, **2**:472–474, **2**:472*F*, **2**:473*F*, **2**:474*F*
    arid environments **2**:470–472, **2**:470*T*, **2**:471*F*
    plants **2**:470–472, **2**:471*F*, **2**:475*F*
  C_4 plants **2**:371*T*
  characteristics
    Afrotropic deserts **2**:467–468
    Australian deserts **2**:468
    general discussion **2**:467–468
    Indo-Malay region **2**:468
    Nearctic region **2**:468–469
    Neotropic region **2**:469
    Palearctic region **2**:469–470
  climates

rainfall pulses **2**:463–464
winter and summer rainfall patterns **2**:464–466, **2**:465*F*
conservation status **3**:171–173*T*
current global land cover **4**:557*F*
desert biome
aridity index **2**:457, **2**:459*F*, **2**:460*F*, **2**:460*T*, **2**:461*F*
coastal deserts **2**:458–460, **2**:462*F*, **2**:463*F*
definition **2**:457–458, **2**:459*F*, **2**:460*F*
geographic distribution **2**:457–458, **2**:458*F*, **2**:459*F*, **2**:460*F*
inland deserts **2**:458, **2**:461*F*, **2**:462*F*
latitudinal desert belts **2**:458
rainshadow deserts **2**:460
satellite imagery **2**:458*F*
desertification **2:479–494**
anthropogenic impacts **6**:79–84
causal factors
climate change **2**:486–487
proximal causation **2**:486, **2**:487*F*
underlying drivers **2**:486, **2**:487*F*
coupled human–environmental (H–E) systems **2**:491–493, **2**:492*T*, **2**:493*F*
dimensions
ecological dimensions **2**:487–488, **2**:488*F*
ecosystem services **2**:490*T*, **2**:491
general characteristics **2**:487–488
human dimensions **2**:488*F*, **2**:489–491, **2**:490*T*
meteorological dimensions **2**:488*F*, **2**:489
domesticated ruminants **1**:704
Dryland Development Paradigm **2**:491–493, **2**:492*T*, **2**:493*F*
dryland syndrome
environmental/ecological characteristics **2**:481–482, **2**:483*F*, **2**:484*F*
land use **2**:483*F*
socioeconomic characteristics **2**:482–484, **2**:483*F*
ecological dimensions
ecological degradation **2**:488–489
greening trends **2**:489
short- versus long-term ecosystem dynamics **2**:487–488, **2**:488*F*
global distribution **2**:479, **2**:480*F*
habitat destruction **3**:404–405
historical perspective **2**:479–480, **2**:482*F*
land degradation **2**:479, **2**:481*F*
range ecology **6**:337–338
stakeholder perspective **2**:480–481, **2**:483*F*
summary discussion **2**:493
United Nations Convention to Combat Desertification (UNCCD)
formal definitions **2**:484
goals and objectives **2**:484–485, **2**:485*T*
origins and formation **2**:484
10-year strategic plan **2**:485–486, **2**:486*T*
desert puna **4**:104–105
ecosystem processes
primary production **2**:474
soil fertility **2**:474
species interactions **2**:474–475, **2**:475*F*
evolutionary development
modern-day deserts **2**:466
sky-island formation **2**:466–467, **2**:467*F*
geographic distribution **2**:458*F*, **2**:459*F*, **2**:460*F*, **2**:467–468
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
hot deserts **6**:334
Iranian (subalpine) desert **1**:262–263*T*
landforms **2**:460–461, **2**:464*F*
Near East ecosystems **5**:480*F*, **5**:485–486
net primary production (NPP) **1**:560*T*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*, **6**:492, **6**:492*T*, **6**:495*T*, **6**:499*T*
soil characteristics **6**:123*T*
soils **2**:461–464
South America **6**:600–604*T*, **6**:607*T*
subtropical deserts **1**:262–263*T*, **1**:267–268
Succulent Karoo **1:485–490**
biodiversity
biodiversity-friendly farming **2**:300
invertebrates **1**:488
plants **1**:485, **1**:486*T*, **1**:487, **1**:487*T*
vertebrates **1**:487*T*
bioregions **1**:485, **1**:486*F*
characteristics **1**:485, **1**:486*F*
climate change threats **1**:489
land degradation threats **1**:489
nitrogen deposition **5**:526*T*
plant growth forms **1**:488–489
threat management **1**:489
vegetation structure **1**:488–489, **2**:467–468
summary discussion **2**:477
Takla Makan–Gobi desert **1**:262–263*T*
temperate semi-deserts and deserts
characteristics **1**:262–263*T*, **1**:269–270, **6**:333–334
Dzungaria **1**:270
Gobi Desert **1**:270
Middle Asian deserts **1**:270
Tarim Basin **1**:270
Tibetan Plateau **1**:270
Thar desert **1**:262–263*T*
Tibetan desert **1**:262–263*T*
tourism industry **7**:223
tropical Africa **1**:53
Turanian desert **1**:262–263*T*
desert spring communities **3**:186, **3**:187*F*
desert tortoise **2**:528*T*
desert varnish **2**:462
designer ecosystems **6**:451
Desmarestia **4**:12
Desmarest's hutia **4**:713–714
desmids **3**:571–572, **3**:571*T*
Desmodera **7**:424*F*
Desmodus rotundus **2**:352
Desmognathinae **7**:389–390
Desmognathus quadromaculatus **1**:173*F*
Desmoncus spp. **2**:27
Desmopteris gardineri **4**:386*T*
Desmoscolex **7**:422*F*
Desmostachya bipinnata **5**:483
desmostylians
characteristics **5**:68
evolutionary history **5**:69–70
extinction/near extinction over evolutionary time **5**:75
despotism **6**:574–575
Desulfobacter **5**:268*F*
Desulfobacterium **5**:268*F*
Desulfobotulus **5**:268*F*
Desulfobulbus **5**:268*F*
Desulfococcus **5**:268*F*
Desulfomicrobium **5**:268*F*
Desulfomonas **5**:268*F*
Desulfomonile **5**:268*F*
Desulfosarcina **5**:268*F*
Desulfotobacterium **5**:268*F*
Desulfotomaculum **5**:268*F*
Desulfovibrio **5**:268*F*, **6**:32
Desulfurococcus **1**:219–220, **1**:221*T*
Desulfurococcus mobilis **1**:315
Desulfuromonas **5**:268*F*
deterministic stability **7**:7–8, **7**:8*F*
detritus
food chains **3**:505–506, **7**:264
freshwater ecosystems **3**:565–566
deuteromycetes **3**:629*F*, **3**:630
Deuterophlebiidae **3**:489–491*T*
deuterostomes **2**:7*F*, **3**:395, **3**:396*F*
Deutzia amurensis **1**:271–272
Development by Design (DbD) **2:495–507**
background and characteristics **2**:496–497, **2**:497*F*
framework and components
basic concepts **2**:497–498
landscape conservation planning **2**:498
mitigation decision-making approaches **2**:498–500, **2**:499*F*
offset contribution estimation **2**:500–503
research summary **2**:505–506
residual impact/optimal offset determinations **2**:500, **2**:501*F*, **2**:502–503*F*
sustainable landscapes **2**:503–505, **2**:504–505*F*
Wyoming Basins Ecoregion (WBE) study **2**:503–505, **2**:504–505*F*
research summary **2**:505–506
Devonian period
ammonoid differentiation **1**:22*F*
arachnids **1**:204–205
arthropod radiation ecology **2**:399
crustaceans **2**:403–404
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
great mass extinctions **5**:168*F*, **5**:169–170, **5**:169*F*
marine diversity **5**:168*F*, **5**:169*F*
marine invertebrates **4**:390, **4**:390*F*

Devonian period (*continued*)
mass extinctions
event characteristics **1**:493
extinction intensities **1**:494*T*, **5**:162*F*
marine animals **1**:413*F*, **1**:493*F*, **5**:158
notable mass extinctions **5**:161*F*
number of families **3**:416*T*, **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
originations **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
terrestrial plant communities
abundance patterns **6**:74*F*
evolutionary development **6**:65–67, **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
terrestrial plant evolution **6**:70
de Vries, Hugo **6**:2
dew and fog **5**:479–480
dhaincha **2**:389
Dharma **6**:564
dhole **5**:603
Diabrotica virgifera **4**:365
Diaceratherium **5**:224*T*
Diadema antillarum **2**:337, **2**:338–340, **2**:530–531, **3**:620–621, **3**:622*F*, **5**:61
diademed sifaka **6**:256–258*T*
diagnostic species concept **6**:683–684, **7**:91–93, **7**:92*F*
diallate **5**:696
Diamesinae **7**:347*F*
Dianotitan **5**:224*T*
diapause **2:665–669**
adaptability **2**:666
basic concepts **2**:665–666
migration strategies **2**:666–667
research summary **2**:668
storage effect **2**:667–668
variability **2**:666
Diaphanosoma **6**:31
Diaphanosoma brachyurum **6**:31
Diaphanosoma leuchtenbergianum **1**:10*T*
Diaprioidea **4**:178
Diaptomus sanguineus **2**:666
diarrhea **1**:594, **3**:635*T*
Diastatidae **3**:489–491*T*
Diastrammena asynamora **3**:732
Diatomophyceae **5**:684*F*
diatoms
average lifespan **4**:656*T*
Bayesian analysis **1**:476*F*
characteristics **5**:683–686, **5**:684*F*, **6**:25
energy flow **3**:246–247
eutrophication impacts **3**:354–356, **3**:356*F*, **3**:357*F*
evolutionary development **2**:641–642
freshwater ecosystems **3**:571–572, **3**:571*T*
harmful algal blooms **1**:475
marine sedimentary habitats **5**:107
phylogenetic classifications **4**:673–675
river ecosystems **6**:477–478, **6**:478*T*
temperature tolerance ranges **4**:114–115
diazotrophs/diazotrophy **2:608–625**
basic concepts **2**:608–610, **2**:609*F*
biodiversity
dinitrogenase/dinitrogenase reductase **2**:610–611
general discussion **2**:610–611
metabolic diversity **2**:614–615
nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
phylogenetic relationships **2**:611*F*
physiological diversity **2**:614–615
biogeochemical cycles **1**:560–561
emerging patterns
changing environmental conditions **2**:621–623
environmental modifications **2**:619–621
spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
functional role
biological carbon (C) pump **2**:615–618, **2**:616*F*
nutrient cycle linkages **2**:618–619
quantification methods **2**:615–618, **2**:617*F*
Redfield ratio **2**:617–618, **2**:617*F*
habitat characteristics **2**:609*F*
research summary **2**:623
Dicaeum quadricolor **7**:183
dicarboximides **5**:695
Diceratherium **5**:224*T*
Diceratops **2**:519*T*, **2**:520*T*
Dicerorhinus hemitoechus **5**:225*T*
Dicerorhinus kirchbergensis **5**:225*T*
Dicerorhinus sumatrensis **3**:196–197, **4**:695
Diceros bicornis **1**:51, **1**:487–488, **4**:695, **5**:225*T*, **6**:578
Dichapetalaceae **6**:314, **6**:316*T*
Dichapetalum **5**:623*F*
dichlobenil **4**:90*T*, **5**:697
dichlorodiphenyltrichloroethane (DDT)
ecotoxicity **3**:120–121
endangered reptiles **3**:218
environmental threats **3**:166, **5**:693, **6**:155
freshwater ecosystems **3**:581
habitat degradation **3**:406
marine mammals **5**:71–72
peregrine falcons **3**:192–193, **4**:258–259
dichlorvos **5**:693–694, **5**:699*T*
dichopatric speciation **6**:621*F*, **6**:622–624, **6**:622*T*, **6**:623*F*
Dichrosta chyscinerea **1**:708
Dichrostachys cinerea **1**:67
Dicistroviridae **5**:134*T*, **5**:137
Dicliptera dodsonii **3**:205
diclofop methyl **5**:696
Diclybothriidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
dicofol **5**:693
Dicolpomys **4**:727*T*
Dicotyledoneae **2**:374*T*, **3**:522*T*
dicotyledons
Antarctica **1**:186*T*
Arctic ecosystems **1**:231*T*
edible species **3**:129*T*
genetic diversity **3**:665*F*
modern extinctions **5**:368*T*
taxonomic distributions **3**:519*T*
Dicotyles spp. **6**:332–333
dicoumarol **5**:712*T*
Dicrocoelium lanceolatum **5**:658–659
Dicrostonyx spp. **6**:331
Dictydium cancellation **3**:631*F*
Dictyosphaeria cavernosa **3**:359
Dictyosphaerium **6**:30
Dictyosteliomycota (cellular slime molds) **3**:631, **6**:289–293, **6**:289*T*, **6**:290*F*
Dictyostelium discoideum **5**:576
Dicuspiditermes **4**:409*T*
Dicyemida **5**:50*T*, **5**:669–670*T*
Dicynodonts **5**:1*F*, **5**:4–5
Didelimorphia **4**:686–687, **4**:687*T*
Didelphidae **4**:683, **4**:686–687, **6**:309*T*
Didelphimorphia **2**:586, **2**:587*F*, **4**:682*F*, **4**:687*T*, **6**:309*T*
Didelphis albiventris **2**:587*F*
Didelphis virginiana **4**:683, **4**:686–687
Didelphoid **4**:701*T*
Didelphoxyuris **7**:421*F*
Didierea madagascariensis **2**:468
Didinium nasutum **6**:670
Didymeles **5**:621*F*
Didymopanax **3**:522*T*
Didymosphenia geminata **3**:357
dieback **3**:633*T*
Diecros bicornis **5**:449
dieldrin **5**:697*T*, **5**:699*T*
difenzoquat **4**:90*T*
differentiation **2:508–515**
adaptive landscape theory **2**:510, **2**:511*F*, **6**:632
basic concepts **2**:508
divergence **2**:510–511
ecological differentiation **2**:513–514
evolutionary time span **2**:514–515
gene flow **2**:511–513
natural selection **2**:509–511
neutral evolution theory **2**:263–264, **2**:508–509, **2**:509*F*, **6**:191
Difflugia **6**:31, **6**:291
diffusive instability **7**:9–10, **7**:10, **7**:10*F*
diflubenzuron **5**:694
Digenea
Allocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Azygioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Bivesiculoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Brachylaimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Bucephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
characteristics **7**:450–451*F*
Clinostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Cyclocoeloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Diplostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
DNA studies **5**:662–663
Echinostomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452–453
Gorgoderoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453

Gymnophalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Haploporoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Haplosplanchoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Hemiuroidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Heronimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Lepocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
life cycle **7**:463*F*
Microphalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453–454
Microscaphidioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Monorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Opisthorchiodea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Paramphistomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
phylogenetic relationships **7**:441*F*
Plagiorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Pronocephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Schistosomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
species diversity **7**:439–440*T*, **7**:451
Transversotrematoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454–455
digenean trematode **2**:528*T*
Digenea simplex **1**:592, **6**:132
digitalis **5**:706, **5**:711–713
Digitalis lanata **6**:137
Digitalis purpurea **5**:712*T*, **6**:137
Digitalis sp. **5**:706, **6**:136–137, **6**:137*F*
digital media **1**:444–445
Digital Object Identifier (DOI) **1**:441–442
Digitaria **2**:376*F*
Digitaria eriantha **1**:67
Digitaria exilis **2**:663
digitonin **2**:426*T*
digitoxin **5**:706, **5**:712*T*, **6**:136–137, **6**:137*F*
digoxin **5**:712*T*
dik-diks **1**:51, **6**:572–573
Dileptus **6**:296*F*
dill **2**:383
Dilleniaceae **1**:263, **5**:622*F*
dimensionless fall velocity **2**:128
dimethoate **5**:693–694, **5**:697*T*, **5**:699*T*, **5**:700*T*
dimethyl sulfide (CH_3SCH_3) **1**:282*T*, **1**:287
Dina lineata **5**:471
Dinema **6**:294*F*
dingfish **3**:467–471
dingos **1**:709
dinids **3**:467–471
dinitroanilines **4**:90*T*, **4**:92*T*, **5**:696
dinitrogen-fixing (N_2-fixing) microorganisms **2:608–625**
basic concepts **2**:608–610, **2**:609*F*
biodiversity
dinitrogenase/dinitrogenase reductase **2**:610–611
general discussion **2**:610–611
metabolic diversity **2**:614–615
nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
phylogenetic relationships **2**:611*F*
physiological diversity **2**:614–615
biogeochemical cycles **1**:560–561
emerging patterns
changing environmental conditions **2**:621–623
environmental modifications **2**:619–621
spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
functional role
biological carbon (C) pump **2**:615–618, **2**:616*F*
nutrient cycle linkages **2**:618–619
quantification methods **2**:615–618, **2**:617*F*
Redfield ratio **2**:617–618, **2**:617*F*
habitat characteristics **2**:609*F*
research summary **2**:623
dinitrophenols **5**:694, **5**:695, **5**:696–697, **5**:699*T*
Dinizia excelsa **1**:160
Dinobryon **6**:30, **6**:293
Dinobryon divergens **6**:33*F*
dinocap **5**:694, **5**:695, **5**:699*T*
Dinodnavirus **5**:136
Dinofelis **4**:724–725*T*
dinoflagellates
average lifespan **4**:656*T*
characteristics **5**:683–686, **5**:684*F*, **6**:25, **6**:294
coral reef ecosystems **2**:335
eutrophication impacts **1**:95–97, **3**:352–354, **3**:353*F*, **3**:356–358
evolutionary development **2**:641–642
freshwater ecosystems **3**:571–572, **3**:571*T*
harmful algal blooms **1**:470, **1**:471*F*, **1**:479–481
life histories **1**:478*F*
marine ecosystems **5**:48
phylogenetic and taxonomic relationships
group origins **1**:472, **1**:473*F*
toxic species origins **1**:472–473, **1**:473*F*
plant-derived drugs and chemicals
brevetoxin A **6**:131–132, **6**:134*F*
characteristics **6**:131
ciguatoxin **6**:131, **6**:132*F*
Gymnodinium breve **6**:131–132, **6**:134*F*
maitotoxin **6**:131, **6**:133*F*
river ecosystems **6**:477–478
RNA editing **5**:573–574*T*
schematic diagram **6**:290*F*
Dinophyceae **5**:684*F*
Dinophysiales **1**:472, **1**:473*F*
Dinophysis **1**:472–473
Dinophysis norvegica **5**:684*F*
Dinophyta **3**:571*T*, **6**:10*T*
Dinornis giganteus **5**:366
Dinornithidae **1**:606
Dinosaur Park Formation (Canada) **2**:518, **2**:520*T*
dinosaurs
Age of the Dinosaurs **5**:5–6
extinction theories **2:516–522**
Cretaceous–Tertiary (K–T) mass extinction **1**:494, **2**:516, **3**:420, **5**:6–7, **5**:173–175
differential patterns of survival **2**:516, **2**:517*F*
long-term diversity trends **2**:517–518, **2**:519*T*, **2**:520*T*
popular accounts **2**:516
research summary **2**:521–522
sea level changes **2**:517
short-term diversity trends **2**:518–519
taxonomic richness **2**:519–521, **2**:521*F*
thermal radiation scenario **2**:516–517
dinosaur tree **2**:47
Dinotopterus **4**:463–464
Dioclea megacarpa **2**:427
Dioctophyma renale **7**:421
Diodea kuntzei **1**:164
Diomedea amsterdamensis **1**:703
Diomedeidae **7**:181–182
Dionaea **6**:44*T*
Dioncophyllaceae **6**:44*T*
Dionea miscipula **6**:44–45
Dionycha **1**:207
Diopsidae **3**:489–491*T*
Diopsoidea **3**:489–491*T*
Dioscorea **3**:130*T*, **6**:316*T*
Dioscorea alata **2**:663
Dioscorea cayenensis **2**:663
Dioscoreaceae **3**:129*T*, **3**:130*T*
Dioscorea composita **5**:706–707
Dioscorea esculenta **2**:663
Dioscorea mexicana **5**:706–707
Dioscorea spp. **2**:383, **6**:553
Dioscorea trifida **2**:663–664, **4**:281–285, **4**:283*F*, **4**:289*T*
Dioscorides **3**:320–321, **5**:705
Diospyros **1**:69–70
Diospyros kaki **2**:663
Diospyros melanoxylon **1**:267, **6**:226
Diospyros spp. **3**:130*T*, **3**:205
Dipentodontaceae **1**:260*T*
diphenyl ethers **4**:90*T*, **5**:696
Diphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Diphyllobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Diphyllobothrium **7**:456–457*F*
Diphyllobothrium latum **5**:657, **5**:660*T*
Diplacodon **5**:224*T*
Diplasia karataefolia **1**:164
Diplocentridae **1**:209
Diplomeris hirsuta **3**:208
diplomonads **6**:294*F*
Diplopoda
alpine ecosystems **1**:155*F*
Amazon ecosystems **1**:246*T*, **1**:247
characteristics **4**:392
developmental modes **5**:426*F*
endangered species **3**:226
geographic distribution
endemism **5**:428

Diplopoda (*continued*)
environmental characteristics **5**:427–428
Eurasia **5**:427*F*
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429, **5**:429*T*
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430
open-country myriapods **5**:430
special habitat myriapods **5**:430–431
morphology **5**:423–424, **5**:423*F*
number of species **5**:427*T*
soil fauna **6**:590*T*
subterranean ecosystems **7**:53
diplopods **3**:674*F*
Diplostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Diplura **4**:308–312, **4**:309*T*, **6**:590*T*
Dipodomys deserti **4**:43*F*
Dipodomys gravipes **7**:229*T*
Dipodomys merriami **4**:43*F*, **4**:446–447, **4**:446*F*
Dipodomys microps leucotis **7**:229*T*
Dipodomys spectabilis **4**:43*F*, **4**:446–447, **4**:446*F*
Dipor Bheel **1**:276*T*
Diprodontia **4**:687*T*, **4**:688–689
Diprotdontidae **5**:225*T*
Diprotodon **4**:727–728
Diprotodon minor **5**:364*F*
Diprotodon optatum **5**:225*T*, **5**:363–364, **5**:364*F*
Diptera **3**:**488–496**
Antarctica **1**:184*T*
Arctic ecosystems **1**:231*T*
characteristics **4**:392, **5**:674
classifications
Brachycera **3**:489–491*T*, **3**:493
common names **3**:489–491*T*
major families **3**:488–493, **3**:489–491*T*
nematocerous groups **3**:489–491*T*, **3**:493
conservation biology **3**:495–496
feeding behaviors **1**:328–329
freshwater ecosystems **4**:374
grazing effects **1**:709–710
insecticide resistance **4**:298–299
life history diversity
adults **3**:494
larvae
herbivores **3**:493–494
non-free-living larvae **3**:494
parasites **3**:494
parasitoids **3**:494
predators **3**:494
prevalence **3**:493
scavengers **3**:493
major divisions and characteristics **4**:308–312, **4**:309*T*
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:466*T*
number of species **3**:488, **3**:488*T*, **4**:371*T*, **4**:392*T*
plankton fraction **6**:26
pollination **6**:141, **6**:141*T*
prevalence **3**:488, **3**:488*T*
quantity estimates **1**:326–328
soil fauna **6**:590*T*
special associations
aquatic Diptera
freshwater habitats **3**:494–495
phytotelmata **3**:494–495
kleptoparasitism **3**:495
social insects **3**:495
swarms **3**:495
wetland ecosystems **7**:389–391
Dipterocarpaceae **1**:263, **6**:314, **6**:316*T*
dipterocarp savanna forests **1**:69*F*, **1**:70
Dipterocarpus **1**:263, **1**:453–454
Dipterocarpus cinereus **1**:514
Dipterocarpus intricatus **1**:70, **1**:267
Dipterocarpus obtusifolius **1**:70
Dipterocarpus tuberculatus **1**:70, **1**:267
diquat **5**:696
direct allogenous speciation **6**:627
Dirofilaria **7**:429–430
Dirphia avia **4**:595
disaccharides **6**:113*T*
disaster species **5**:164–165
Discula destructiva **2**:528*T*, **3**:411*F*, **3**:633, **3**:633*T*
disease
biodiversity–ecosystem services relationship **1**:348–350, **3**:349
biodiversity loss **1**:519–520, **4**:664
climate change effects **2**:103–104, **2**:110*T*
emerging infectious diseases (EIDs)
human health **1**:368–370, **1**:369*F*
Southeast Asia **1**:521
endangered mammals **3**:191
habitat–niche relationship **1**:519–520, **4**:45–46
human infectious disease regulation **1**:348–350, **3**:349
microbial biodiversity **5**:258
natural history museum research programs **5**:415
parasitic disease vectors
crop production **5**:673
fish **5**:672, **5**:672*F*
humans **5**:670–672, **5**:671*T*
livestock **5**:672
plants **5**:672–673
plant biodiversity threats **6**:62–63
population dynamics **6**:179–180
simple two-species interactions **6**:179–180
ultraviolet radiation effects **7**:302–303
see also infectious diseases
disease triangle **3**:633, **3**:633*F*
disgust **1**:42–43, **1**:42*T*
Dismorphia amphiona **1**:650, **1**:651*F*
Dismorphiinae **1**:650, **1**:651*F*
disodium methanearsenate (DSMA) **5**:696
dispersal biogeography **2**:**539–561**
applied dispersal biogeography **2**:559
basic concepts **7**:343, **7**:343*F*, **7**:347*F*
biogeographic and evolutionary consequences
extinction **2**:558–559
general discussion **2**:558
speciation **2**:558
historical perspective **2**:539–540
long-distance dispersal **2**:539–540, **6**:616–617
neutral theory **5**:511
research methodologies
inferential research
cladistic biogeography **2**:555
dispersal–vicariance analysis (DIVA) **2**:556, **2**:556*F*
nestedness **2**:555–557, **6**:649–650
mathematical models
benefits **2**:557
dispersal models **2**:557
fat-tailed dispersal models **2**:557
reproduction–dispersal models **2**:557–558
natural range expansion
early hominins **2**:550–554, **2**:553*F*
Holocene postglacial spread **2**:550, **2**:551–552*T*, **2**:552*F*
recent range expansion **2**:549–550
observational research
alien species invasion **2**:554–555, **2**:554*F*
barrier dissolution **2**:545*F*, **2**:547–549
direct observations **2**:546–547
natural range expansion **2**:549–550
virgin habitat colonization **2**:547, **2**:548–549*F*
research summary **2**:559–560
Southern Hemisphere ecosystems **6**:616
success determinants
colonization
dispersal barriers/dispersal routes **2**:544–545, **2**:545*F*, **2**:547–549
ecological opportunity **2**:545–546
general characteristics **2**:544
natural range expansion **2**:549–550
physical access **2**:544–545
postdispersal establishment **2**:545–546
preadaptation **2**:544
dispersal mechanisms **2**:543, **2**:544
geographic distribution **2**:546
habitat fragmentation **2**:546
theoretical perspectives
biogeographic perspective
gradual range expansion **2**:542
jump range expansion **2**:542–544, **2**:543
passive dispersal **2**:542–544
spatial spread **2**:541–542
species range **2**:541–542
ecological perspective **2**:540, **2**:541*F*
evolutionary perspective **2**:540–541
population genetics perspective **2**:540–541
Dispholidus typus **6**:409
Disporum **1**:271
dissipative beaches **2**:128–129, **2**:129*F*, **2**:130*F*, **2**:135, **5**:51
dissolved inorganic carbon (DIC) **1**:559–560, **1**:560*F*

dissolved organic matter (DOM) **3**:244–245, **7**:51–52
Disterigma **3**:522*T*
Disterigma empetrifolium **4**:103
Distichilis spicata **3**:115–116
Distichlis humilis **4**:105
Distichlis spicata **6**:454
Distoechurus pennatus **4**:688–689
disturbances **2:562–567**, **2:568–581**
  anthropogenic impacts **2**:566–567
  background and characteristics **2**:568–569, **2**:569*T*
  basic concepts **2**:562–563
  biodiversity-friendly farming **2**:542–544
  carnivores **2**:565–566
  climate change effects **2**:104
  ecological impacts **2**:562–563
  ecological succession **7**:69–70
  ecosystem function measurements **3**:83–84, **3**:83*F*
  El Niño Southern Oscillation (ENSO) **3**:155–156
  extreme events **4**:211–212
  forest ecosystems **3**:529
  freshwater invertebrates **4**:375
  general discussion **2**:568–569
  grazing effects **4**:9–10, **4**:10*F*
  herbivores **2**:565–566
  historical range and variation (HRV)
    applications **2**:573–574
    background information **2**:571–572
    basic concepts **2**:572–573, **2**:572*F*
    edge effect **2**:577*F*
    limitations **2**:574–578, **2**:575*F*
    quantification methods **2**:573–574
    simulation landscapes **2**:574–578, **2**:575*F*, **2**:577*F*
  interaction impacts **2**:566
  intermediate disturbance hypothesis **2**:115–117, **2**:563, **4**:79, **5**:256
  intertidal ecosystems **4**:335–336
  landscape diversity **4**:479–480
  landscape legacies **4**:525
  mechanical disturbances **2**:563–564
  Mediterranean-type climate ecosystems **5**:213–214
  metapopulation theory **5**:243–244
  microbial biodiversity **5**:256
  physico–chemical disturbances **2**:564–565, **2**:565*F*
  plant biodiversity threats **2**:115–117, **2**:123, **3**:118, **6**:62–63
  recovery mechanisms **4**:498–499, **4**:499*F*
  soil biodiversity **6**:588–589
  species distribution modeling (SDM) **6**:698–699
  tropical forest regeneration **7**:277–278
  wetland ecosystems **7**:399–400
  wildland fires
    basic concepts **2**:569–571, **2**:569*T*
    influencing factors **2**:570*F*
    whitebark pine forests **2**:571, **2**:571
ditches **3**:572*T*
dithiocarbamates **5**:695, **5**:699*T*
dithiopyr **4**:90*T*
Ditrysia **5**:386–389, **5**:386*F*, **5**:387–388*T*
Diuraphis noxia **4**:361
diuron **5**:696
divergence **2**:510–511
DIVERSITAS program **4**:406–407
diversity **2:434–446**
  calculation methods **2**:435–437, **2**:435*F*
  community/regional level **2:595–607**
    definition **2**:596
    local species richness controls
      diversity–productivity relationship **2**:597–599, **2**:598*F*
      prediction limitations **2**:599
      research summary **2**:606
      spatial scale **2**:604–605, **2**:605*F*
      species–energy theory **2**:596–597, **2**:597*T*
      theoretical perspectives **2**:605
      Volterra–Gause perspective **2**:595–596
    measurement and analysis **1**:403
    regional species richness controls
      historical biogeographical factors **2**:599–601, **2**:600*F*
      metacommunity processes model **2**:601–602, **2**:602*F*
      phylogenetic patterns **2**:599–601, **2**:600*F*
      proportional sampling **2**:602*F*, **2**:603
      research summary **2**:606
      spatial scale **2**:604–605, **2**:605*F*
      theoretical perspectives **2**:605
    research summary **2**:606
    saturation tests
      caveats **2**:603
      experimental evidence **2**:603–604, **2**:604*F*
      methodologies **2**:602–603
      test results **2**:603
      unsaturated versus saturated patterns **2**:603
    spatial scale **2**:604–605, **2**:605*F*
    theoretical perspectives
      comparative models **2**:606, **2**:606*F*
      design manipulations **2**:605–606
      multiscale perspectives **2**:605
    Type I/Type II relationships **2**:603
  definition **2**:434–435
  diversity indices
    Gini–Simpson index **2**:438–439
    Shannon index **2**:438
    Simpson index **2**:438–439
  diversity principle **4**:148
  molecular level diversity **2:626–637**
    conservation biology **2**:637
    DNA sequencing **2**:627, **5**:254
    future research outlook **2**:635–637
    general discussion **2**:626
    genetic variation
      mutation process **2**:627–628
      phylogenetic divergence **2**:632–633
      process mechanisms **2**:627
    mutation process
      characteristics and classifications **2**:627–628
      cytological considerations **2**:628–629
      empirical research **2**:631–632, **2**:631*T*
      evolutionary models **2**:629–630
      genome size effects **2**:628, **2**:628*F*
      influencing factors **2**:629
      mutation rates **2**:628, **2**:628*T*
      nearly neutral theory **2**:630
      neutral theory **2**:629–630
      nucleotide polymorphism **2**:631–632, **2**:631*T*, **2**:632*F*
      ontogenetic considerations **2**:628–629
      prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
      selection-based models **2**:630–631
      statistical tests **2**:631
    phylogenetic divergence
      amino acid replacement **2**:634, **2**:634*T*
      empirical research **2**:634
      evolutionary models **2**:633
      general discussion **2**:632–633
      nearly neutral theory **2**:633
      neutral theory **2**:633
      prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
      selection-based models **2**:633–634
      statistical tests **2**:634
      substitution process **2**:632–633, **2**:634*T*
    prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
    research background **2**:626–627
    valuation measures **2**:637
  organism level **2:638–647**
    characteristics
      Darwinian theory **2**:638
      energy flow systems **2**:638–639
      functional internal organization **2**:640
      hierarchical organization **2**:640
      information systems **2**:639–640
      Lamarckian theory **2**:638
    evolutionary development
      eukaryotes **2**:641
      general discussion **2**:640–647
      prokaryotes **2**:640–641
    research summary **2**:647
  partitioning methods
    additive partitioning
      absolute species turnover **2**:442–443
      alpha + beta **2**:442–443
      proportional species turnover **2**:444
      Whittaker's species turnover **2**:443–444
    alpha, beta, and gamma diversity
      additive partitioning **2**:442–443
      basic concepts **2**:439–440
      beta richness and diversity **2**:442
      diversity components **2**:440–442, **2**:441*F*
      richness components **2**:440, **2**:441*F*
    multiplicative partitioning
      alpha x beta **2**:439–440
      gamma diversity **2**:439–440
      richness x evenness **2**:439
  phylogenetic diversity **6**:22
  quantification methods **2**:437–438

diversity (*continued*)
research approaches
community approach **2**:648, **2**:649*T*
ecosystem approach **2**:648–649, **2**:649*T*
general discussion **2**:648
individual approach **2**:648, **2**:649*T*
population approach **2**:648, **2**:649*T*
resemblance indices **2**:444–445
resource partitioning **6:429–437**
Clements–Gleason debate **6**:429
community dominance–diversity relationships **6**:431–432, **6**:432*F*
diversity assessment framework **6**:429–430, **6**:430*F*
diversity mechanisms
chance **6**:435
frequency-dependent effects **6**:435
general discussion **6**:435
spatial structure **6**:435–436, **6**:436*F*
niche differentiation
empirical evidence **6**:432–433, **6**:433*F*
relevance **6**:434–435
spatial heterogeneity **6**:432–433, **6**:433*F*
temporal heterogeneity **6**:433–434, **6**:434*F*
theoretical perspectives **6**:430–432, **6**:431*F*, **6**:432*F*
research background **6**:429
research summary **6**:436–437
systems theory
basic concepts **2**:651–652
stability dynamics
basic concepts **2**:652–653
local asymptotic stability **2**:652–653
transient dynamic behavior **2**:653, **2**:653*F*
steady-state dynamics
basic concepts **2**:651–652, **2**:651*F*
colonization rates **2**:651–652, **2**:652*F*
community structure **2**:652
ecosystem processes **2**:652, **2**:652*F*
extinction rates **2**:651–652, **2**:652*F*
number of species **2**:651–652, **2**:652*F*
population densities **2**:652
species composition **2**:652
taxonomic versus functional diversity **2:648–656**
diversity estimations
background information **2**:650
species evenness **2**:650, **2**:651*T*
species richness **2**:650, **2**:651*T*
species richness–species evenness relationship **2**:650–651
ecological coherence **2**:655–656
historical perspective **2**:648
linkages
basic processes **2**:654–655
C3 and C4 grasses **2**:654–655
exotic species invasions **2**:655
island biogeography **2**:652*F*, **2**:655, **2**:655*T*
natural selection **2**:656
research approaches **2**:648, **2**:649*T*
species classifications
characteristics **2**:649–650
ecological niche model **2**:649–650, **2**:649*F*
functional groupings **2**:649*F*, **2**:650
stability dynamics
basic concepts **2**:652–653
local asymptotic stability **2**:652–653
process-based models **2**:654
species interaction models **2**:653–654
transient dynamic behavior **2**:653, **2**:653*F*
steady-state dynamics
basic concepts **2**:651–652, **2**:651*F*
colonization rates **2**:651–652, **2**:652*F*
community structure **2**:652
ecosystem processes **2**:652, **2**:652*F*
extinction rates **2**:651–652, **2**:652*F*
number of species **2**:651–652, **2**:652*F*
population densities **2**:652
species composition **2**:652
systems theory
basic concepts **2**:651–652
stability dynamics **2**:652–653
steady-state dynamics **2**:651–652, **2**:651*F*
diversity of life concepts **3**:134
diversity–productivity relationship **2**:597–599, **2**:598*F*
DIWPA (Diversitas in the West Pacific and Asia) program **4**:406–407
Dixidae **3**:489–491*T*
Djibouti **6**:507*T*
dobsonflies **4**:371*T*, **4**:392*T*
Dobzhansky, Theodosius **2**:424–425, **3**:394, **6**:3–4, **6**:681–682, **7**:46
Dociostaurus maroccanus **3**:221–222
documentary films **4**:654
dodders **4**:70, **4**:77
dodemorph **5**:695
dodine **5**:695
dodo **1**:605–606, **1**:621, **7**:184
Doedicurus **4**:727*T*
dogs
convergent evolution **4**:701*T*
fungi-related diseases **3**:635–636
Great American Biotic Interchange (GABI) **2**:584*F*
hybridization **3**:191
introduced species
Australia **1**:80
Brazil **1**:83
India **1**:82
South Africa **1**:81
United Kingdom **1**:79
United States **1**:76–77
late Quaternary extinction event (LQEE) **4**:724–725*T*
species diversity **1**:685
dog salmon **6**:523
dogwood **2**:528*T*, **3**:411, **3**:411*F*, **6**:120*T*
dogwood anthracnose **3**:411*F*, **3**:633*T*
doka ecosystems **1**:67–68
Dolichogenidea metesae **5**:607
Dolichopodidae **3**:492
Dolichopodide **3**:489–491*T*
Dolichotis patagonum **2**:587*F*
dolphin fish **3**:467–471
dolphins
government legislation and regulations **3**:713*T*, **3**:716
migratory patterns **5**:314
number of species **4**:690*T*
ocean fishing/bycatch **3**:447
river ecosystems **6**:479–480
species diversity **4**:693–694
threat percentage **3**:195
domatia **2**:429
domesticated animals
animal breeding **3:100–107**
challenges **1**:648
conservation planning **1**:639–640
conservation strategies
ex situ strategies **1**:645–646
general discussion **1**:642–645
in situ strategies **1**:642–645
crossbred animals **1**:639
current trends **1**:638–639
diversity challenges **1**:636
future outlook **1**:648
genetic structure **1**:637–638
global breeds **1**:638–639
livestock breeds **1**:637
management strategies
conservation strategies **1**:642–645
Food and Agriculture Organization (FAO) programs **1**:640–642, **1**:641
inventory, description, and characterization assessments **1**:641–642, **1**:642*F*
risk status classification **1**:641–642, **1**:642*F*, **1**:642*T*, **1**:643*F*, **1**:644*F*
new technologies
cloning **1**:647–648
cryopreservation **1**:646
gene mapping **1**:646–647
molecular characterization **1**:646–647
wild relative use **1**:640
desert ecosystems **2**:475–476, **2**:476*T*
domesticated ruminants **1:702–711**
biotic relationships
birds **1**:710
invertebrate fauna **1**:709–710
natural predators **1**:709
parasites and pathogens **1**:710–711
plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
seed dispersal and germination **1**:708–709
small mammals **1**:710
ecological role
desertification **1**:704
erosion **1**:704–705
grazing effects **1**:703–704
land degradation **1**:704–705
nutrient cycling **1**:705–706, **1**:705*F*
over-enrichment factors **1**:705–706
plant diversity–grazing intensity relationship **1**:704*T*
vegetation impacts **1**:703–704
historical perspective **1**:702–703

farming practices **6**:567
forest management impact **1**:453
historic land-use patterns **4**:570
domesticated plants
Amazon ecosystems **4:279–292**
basic concepts **4**:279
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
basic concepts **4**:279
crop plants **2:657–664**
conscious versus unconscious selection **2**:660
crop maintenance **2**:660–661
evolutionary development
African Savanna belt **2**:659
East Asia region **2**:658–659
fertile crescent region **2**:658
founder crops **2**:658
pre-Columbian America **2**:659
growing purposes **2**:661
historic land-use patterns **4**:570, **4**:570*F*
human–crop coevolution **2**:657, **4**:570, **4**:570*F*
native crops **2**:662
reproductive systems
characteristics **2**:659–660
self-pollination **2**:659–660
vegetative propagation **2**:660
sowing and reaping impacts **2**:661–662
traditional agriculture **2**:657–658
domestication mechanisms **4**:279–281
environmental factors **4**:285–287, **4**:286*F*
farming practices **6**:567
genetic diversity opportunities **4**:280–281
genetic drift **4**:279–281, **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:280, **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
domesticated ruminants **1:702–711**
biotic relationships
birds **1**:710
invertebrate fauna **1**:709–710
natural predators **1**:709
parasites and pathogens **1**:710–711
plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
seed dispersal and germination **1**:708–709
small mammals **1**:710
ecological role
desertification **1**:704
erosion **1**:704–705
grazing effects **1**:703–704
land degradation **1**:704–705
nutrient cycling **1**:705–706, **1**:705*F*
over-enrichment factors **1**:705–706
plant diversity–grazing intensity relationship **1**:704*T*
vegetation impacts **1**:703–704
historical perspective **1**:702–703
domestic pigeons **1**:78, **1**:624, **5**:476
Domingo, Esteban **5**:633
Dominica **6**:519
Dominican Republic **5**:493*T*, **5**:495*F*, **5**:497*F*, **6**:493–494, **6**:507*T*
domoic acid **6**:132, **6**:134*F*
donkeys
deoxyribonucleic acid (DNA) evolution **2**:56*F*
desert ecosystems **2**:476*T*
introduced species
Australia **1**:80
Brazil **1**:83
South Africa **1**:81
dorcas gazelle **5**:475–476
Dorcatragus megalotis **1**:51
dories **3**:467–471
dormancy **2:665–669**
adaptability **2**:666
basic concepts **2**:665–666
migration strategies **2**:666–667
research summary **2**:668
storage effect **2**:667–668
variability **2**:666
dormice **4**:573, **7**:27
Dorvilleidae **7**:417*F*
Dorylaimus stagnalis **7**:422–424, **7**:426*F*
Douc langur **3**:192
Douglas fir
climate change effects **6**:150*T*
North America **7**:118
plant tissue chemistry **6**:120*T*
root–soil interface **6**:113*T*
downy mildews **3**:630
Dracaena ombet **6**:78
Draconema **7**:423*F*, **7**:429–430
Dracunculus medinensis **5**:670–672
Draft BioCode **5**:553–554
dragonets **3**:467–471
dragonfish **3**:465, **3**:467–471
dragonflies
freshwater ecosystems
life cycle **4**:370
threatened species **3**:579*T*
lake and pond ecosystems **3**:179, **3**:180*F*, **4**:371–373
Near East ecosystems **5**:460*T*, **5**:464*T*
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
threatened species **7**:201*T*
trophic cascades **7**:261–262, **7**:261*F*
wetland ecosystems **7**:389–391
drainage lakes **3**:560
Dranunculus medinensis **7**:428*T*
Dreiser, Theodore **4**:652
Dreissena bugensis **3**:369
Dreissena polymorpha
environmental ethics **3**:270, **3**:270*F*
eutrophication impacts **3**:369
impacts and distribution **3**:181, **3**:284, **3**:411, **4**:143, **4**:364–365, **6**:93–95
lake and pond ecosystems **3**:181
physical characteristics **5**:374, **5**:381
species diversity theories **6**:26–27
Drepanidinae **1**:29
Drepanidini **1**:29
Drepanis pacifica **1**:29*F*
Drepanocladus uncinatus **1**:232, **1**:239*F*
Drepanoidea **5**:387–388*T*
Drepanotermes **4**:409*T*
driftfish **3**:467–471
drills **1**:49, **6**:251*T*
Driloleirus americanus **3**:225
Drimia maritima **5**:712*T*
Drimys **3**:522*T*, **7**:119
drinking water **1**:95–97, **5**:351
drive-through animal parks **7**:473–474
Dromaeosaurus **2**:519*T*, **2**:520*T*
Dromaius novaehollandia **1**:606
dromedary camels **1**:644*F*, **2**:476*T*, **4**:697
Dromiciops australis **2**:586, **2**:587*F*
Dromiciops gliroides **5**:217
drooping prickly pear **4**:350–351
Drosera **6**:44*T*
Droseraceae **6**:44*T*
Drosera rotundifolia **6**:45
Drosophila
adaptive radiation **1**:28–29
genetic diversity **3**:665*F*
host–parasitoid interactions **5**:675–676
interspecific competition **2**:185–186, **2**:186*F*
Drosophila americana **2**:636*F*
Drosophila ananassae **2**:636*F*
Drosophila heteroneura **1**:29
Drosophila melanogaster
adaptive processes **1**:18
genetic diversity **3**:673–674
genetics research **2**:267–272, **6**:2
inbreeding studies **4**:249*F*
life history measures **4**:636, **4**:636*F*, **4**:637*F*
molecular level diversity **2**:636*F*
mutation rates **2**:628*T*
natural selection **2**:510
nucleotide polymorphism **2**:631–632, **2**:631*T*, **2**:632*F*
random reproduction **2**:509
recombination characteristics **6**:361*T*
Drosophila miranda **2**:636*F*
Drosophila pseudoobscura **2**:185–186, **2**:636*F*
Drosophila serrata **2**:185–186
Drosophila silvestris **1**:29
Drosophila simulans **2**:636*F*
Drosophila virilis **4**:249*F*
Drosophilidae **3**:489–491*T*, **3**:493
Drosophyllum **6**:44*T*
drought conditions

ecological memory **4**:217, **4**:218*F*
ecosystem function measurements **3**:83*F*
El Niño Southern Oscillation (ENSO) **3**:156–158
extreme events **4**:211–212
Mediterranean-type climate ecosystems **5**:213–214
megaherbivores **5**:230
plant biodiversity threats **6**:62–63
wetland ecosystems **7**:395
drug development
economic importance **1**:393–395
fungi-derived drugs **3**:638
plant-derived drugs and chemicals **6:129–139**
algae
domoic acid **6**:132, **6**:134*F*
green algal metabolites **6**:132–135, **6**:134*F*
kainic acid **6**:132, **6**:134*F*
cyanobacteria
anticancer agents **6**:130–131
cryptophycin **6**:131, **6**:131*F*
curacin A **6**:130–131, **6**:130*F*
dinoflagellates
brevetoxin A **6**:131–132, **6**:134*F*
characteristics **6**:131
ciguatoxin **6**:131, **6**:132*F*
Gymnodinium breve **6**:131–132, **6**:134*F*
maitotoxin **6**:131, **6**:133*F*
future outlook **6**:138
herbal remedies
Echinacea **6**:137–138, **6**:138*F*
ginseng **6**:137, **6**:138*F*
higher plants
camptothecin **6**:135–136, **6**:135*F*
paclitaxel **6**:135, **6**:135*F*
marine plants
algae **6**:132
cyanobacteria **6**:130–131
dinoflagellates **6**:131
research background **6**:129–130
terrestrial plants
Digitalis sp. **6**:136–137, **6**:137*F*
herbal remedies **6**:137
higher plants **6**:135
steroidal glycosides **6**:136–137, **6**:137*F*
tropane alkaloids **6**:136, **6**:136*F*
drums **3**:467–471
Drupadia theda **3**:158
Dryandra **5**:221
dry Asian regions
characteristics **1**:267–268
Mediterranean ecosystems **1**:262–263*T*, **1**:268–269
subtropical deserts **1**:262–263*T*, **1**:267–268
temperate grasslands **1**:262–263*T*, **1**:269
temperate semi-deserts and deserts
characteristics **1**:262–263*T*, **1**:269–270
Dzungaria **1**:270
Gobi Desert **1**:270
Middle Asian deserts **1**:270
Tarim Basin **1**:270
Tibetan Plateau **1**:270
Dryas monkey **6**:256–258*T*
Dryas octopetala **1**:232
Dryas punctata **1**:232
Dry Chaco forests **5**:496, **5**:497*F*
dry deciduous forests **1**:69–70, **1**:69*F*
dry forests
conservation status **3**:171–173*T*
global biodiversity trends **4**:563*F*
neotropical seasonally dry forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
precipitation–temperature relationship **1**:59*F*
protected area systems **6**:492*T*, **6**:499*T*
tropical dry forest **1**:266–267, **2**:28, **3**:110–111, **3**:111*F*, **3**:112–113, **6**:600–604*T*
Dryland Development Paradigm **2**:491–493, **2**:492*T*, **2**:493*F*
dryland ecosystems **2**:479, **2**:481*F*, **2**:678–679
see also deserts
dryland syndrome
environmental/ecological characteristics **2**:481–482, **2**:483*F*, **2**:484*F*
land use **2**:483*F*
socioeconomic characteristics **2**:482–484, **2**:483*F*
Drymonia **3**:522*T*
dry mouth **3**:635*T*
Dryobalanops **1**:263
Dryobates syriacus **5**:474
Dryococelus australis **3**:225
Dryomyzidae **3**:489–491*T*
Dryptosaurus **2**:519*T*
dry puna **4**:104
dry semideciduous southern fringe forests **1**:161–162
dry thorn forests **1**:69, **1**:69*F*
dry-zone Toque macaque **6**:256–258*T*
Dubautia **1**:30–32, **1**:30*F*
Dublin Core Metadata Initiative **1**:440–441
Dubos, Rene **6**:382
Dubriar Haor **1**:276*T*
ducks
deoxyribonucleic acid (DNA) evolution **2**:56*F*
hunting declines **7**:246–247, **7**:246*F*
introduced species **4**:362
risk status classification **1**:644*F*
wetland ecosystems **7**:390
Ducktown, Tennessee **6**:154–155
Dudleya traskiae **3**:208–209, **3**:269–270, **3**:270*F*, **3**:313
Dugesia **7**:442–443*F*
Dugesia biblica **5**:470*T*
Dugesia salina **5**:470*T*
Dugong dugon **1**:55, **6**:547–548
Dugongidae **4**:694
dugongs
characteristics **4**:694, **5**:67–68
migratory patterns **5**:314
number of species **4**:690*T*
taxonomy **5**:65*T*
threat percentage **3**:195
Duguetia stenantha **4**:289*T*
duikers **4**:696, **6**:572–573
Du-Le Garden **6**:506
Dulongiaceae **3**:522*T*
Dunaliella **6**:30
Dunaliella salina **1**:592–593
dune ecosystems **5**:452–453, **5**:453*T*
dune grass **6**:121
dune lark **1**:487–488
dune vegetation **1**:292–297, **2**:371*T*
dung beetles **1**:710, **3**:495
dungflies **1**:710, **6**:49*T*
Dupontia fisheri **1**:232, **1**:236, **1**:236*F*
Durban Agreements **2**:318
durian **2**:663
Durio **1**:263
Durio zibethinus **2**:663
Durisols **2**:461–464
durum wheat **2**:662–663
Dusicyon culpaeus **2**:587*F*
dusky seaside sparrow **7**:189
dusky sunbird **1**:487–488
Dust Bowl **2**:283–284
dust mites **1**:215
dust modeling **2**:680
dust plumes **2**:460–461
Dutch elm disease **2**:528*T*, **3**:628–629, **3**:633*T*
Duvernoy's gland **6**:409
dwarf galagos **6**:251*T*
dwarf lemur **6**:251*T*
dwarf mistletoes **4**:77, **5**:673
dwarf sperm whales **3**:192–193
dynamic global vegetation models (DGVMs) **2:670–689**
basic principles and processes
atmospheric carbon dioxide (CO_2) concentrations **2**:675
carbon allocation **2**:676
carbon–nitrogen coupling **2**:675–676
carbon–water coupling **2**:675
dryland ecosystems **2**:678–679
evolutionary development **2**:674–675
fire modeling **2**:678
leaf-to-canopy scaling **2**:676–677
multiple resource limitations **2**:676
phenology **2**:677–678
plant functional types **2**:677
wetlands **2**:678–679

climate change studies **2**:91–92
future research outlook **2**:684
general discussion **2**:672–673
modeling benefits **2**:673
related ecosystem models **2**:673–674
research applications
atmospheric gases and aerosols **2**:680
biophysical land–atmosphere interactions **2**:680–681
carbon cycle **2**:679
hydrological research **2**:679–680
vegetation structure and composition **1**:64*F*, **2**:681–682
vegetation–climate interactions
carbon cycle feedbacks **2**:682
geological record **2**:682–684
Quaternary period **2**:682–684
Dysmicoccus neobrevipes **4**:365
Dysoxylum **1**:263
Dyspnoi **1**:210
dystrophic level **3**:563*T*
Dytiscidae **1**:327*T*
Dzungaria **1**:270

E

eagles **1**:709, **5**:474, **6**:479–480
Earth Day **2**:286, **3**:136, **6**:77–78
earth hoppers **3**:734–735
earthquakes **2**:563
Earth Summit **2**:307–308, **6**:78, **7**:43–44, **7**:136–137
Earth System Science Partnership (ESSP) **4**:326
Earthwatch Institute **3**:138
earthworms
adaptability **4**:393
air pollution impacts **1**:142*F*
biomonitoring processes **3**:122*T*
characteristics **4**:392, **6**:588
classifications **4**:381*T*
endangered species **3**:225
food webs **7**:266*F*, **7**:267*T*
functional role **6**:588
genetic diversity **3**:674*F*
introduced species **1**:78
number of species **4**:371*T*
slash-and-burn systems **6**:557
soils
soil fauna **6**:588, **6**:590*T*
soil quality indicators **6**:597
earwigs **3**:222, **4**:392*T*
East Africa
deforestation **2**:448*T*
nitrogen deposition **5**:526*T*
protected area systems **6**:491*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
East Asia
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
crop plant domestication **2**:658–659
deforestation **1**:677*F*, **2**:448*T*
native crops **2**:663
protected area systems **6**:491*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
temperate forests
distribution patterns **7**:116–117, **7**:116*T*
diversity patterns **7**:119–121, **7**:119*T*, **7**:120*F*
dominant floristic features **7**:118
endemism **7**:121
physiognomic characterization **7**:117
Easter Island **6**:83
Eastern Afromontane region **6**:82*T*
Eastern Atlantic region **3**:474
eastern Bornean orangutan **6**:256–258*T*
eastern chimpanzee **6**:256–258*T*
eastern chipmunk **1**:362
eastern cross-marked surili **6**:256–258*T*
Eastern Europe
deforestation **2**:448*T*
vegetation surveys **3**:341–344*T*
eastern lowland gorilla **6**:256–258*T*
Eastern meadowlark **4**:490*F*
Eastern North America **5**:561*F*, **5**:562–563
Eastern Pacific region **3**:473
Eastern spinner dolphin **5**:87*T*
Eastgate Centre (Harare, Zimbabwe) **1**:590–591
East Indian arrowroot **2**:663
East Kalimantan **3**:156–158, **3**:158*F*
East Kalimantan Hose's surili **6**:256–258*T*
East Melanesian Islands **6**:82*T*
East Pacific Rise **7**:330
East Sabah Hose's surili **6**:256–258*T*
East Sahero-Arabian Desert **2**:469–470
East Timor **1**:513*T*, **4**:560*F*, **6**:507*T*
see also Southeast Asia
Ebenaceae **3**:129*T*, **3**:130*T*, **6**:316*T*
ebola virus **2**:528*T*, **5**:578
ebony **1**:267
Ecballium elaterium **2**:543
Ecdeiocoleaceae **5**:221
Ecdysozoa **3**:396*F*
Echeneidocoelium **7**:450–451*F*
echidna **3**:195, **4**:701*T*
Echimyidae **6**:309*T*
Echinacea **6**:137–138, **6**:138*F*
Echinacea purpurea **6**:138
Echinobothrium **7**:456–457*F*
Echinochasmus **7**:450–451*F*
Echinochla **2**:376*F*
Echinochloa colona **1**:67, **7**:146
Echinochloa pyramidalis **7**:153–154, **7**:387
Echinochloa spectabile **1**:163
Echinochloa stagnina **7**:146
Echinococcus **7**:456–457*F*
Echinodermata
average lifespan **4**:656*T*
Census of Marine Life project **2**:7*F*
fossil record **3**:684*F*
freshwater ecosystems **3**:571–572
global diversity **4**:385*T*
hydrothermal vent ecosystems **7**:329–330
known and estimated species **4**:384*T*
mangrove ecosystems **5**:20*T*
marine ecosystems **5**:50*T*, **5**:106*F*
number of species **5**:669–670*T*
Proterozoic–Cambrian transition **3**:396*F*
sandy beaches **2**:131–133
Echinogammarus syriacus **5**:471
Echinoidea **3**:681*F*, **3**:685*F*, **4**:381*T*, **5**:176, **5**:669–670*T*
Echinostomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452–453
Echiura
classifications **4**:381*T*, **7**:408
geographic distribution **7**:418
global diversity **4**:385*T*
known and estimated species **4**:384*T*
marine ecosystems **5**:50*T*
morphology
anterior ends **7**:410*F*, **7**:412*F*
chaetae **7**:409–410, **7**:414*F*
circulatory system **7**:415
digestive tract **7**:415
general characteristics **7**:408–409
gills **7**:415
gonads **7**:415–416
head **7**:409
musculature **7**:414–415
nephridia **7**:415
nervous system **7**:410–411
posterior ends **7**:410*F*
pygidium **7**:409
segments **7**:409, **7**:413*F*
sense organs **7**:411–414
septa **7**:415
number of species **5**:669–670*T*
phylogenetic and taxonomic relationships **7**:416–419, **7**:417*F*
reproductive biology **7**:416
sampling methods **4**:381*T*
Echiurida **5**:669–670*T*
ecocriticism **4**:653–654
eco-floristic zone concept **6**:329
ecogenomic sensors **2**:690–700
artist's rendition **2**:691*F*
Environmental Sample Processor (ESP)
basic concepts **2**:693
components **2**:692*F*, **2**:693*F*
operational uses
autonomous underwater vehicles (AUVs) **2**:698*F*, **2**:699
open-ocean drifters **2**:696*F*, **2**:697–699
seafloor deployments **2**:697–699, **2**:697*F*
surface moorings **2**:695*F*, **2**:697–699
photograph **2**:692*F*
research background **2**:692–693
sample processing capabilities
DNA probe arrays **2**:694*F*, **2**:697
polymerase chain reaction (PCR) techniques **2**:694–697, **2**:694*F*
ribosomal RNA (rRNA) sequences **2**:693–697, **2**:694*F*
sandwich hybridization assay (SHA) **2**:693–694, **2**:694*F*
future outlook **2**:699
historical perspective **2**:690–691
molecular biology research **2**:690–691
technical challenges **2**:691–693
ecoinformatics **1**:440–445
metadata

ecoinformatics (*continued*)
classifications **1**:441
emerging technologies
digital media **1**:444–445
scientific workflow systems **1**:445
Standard Generalized Markup Language (SGML) **1**:444
World Wide Web (WWW) **1**:444
historical background **1**:440–441
purposes
annotation **1**:442
data access **1**:441
data portals **1**:441
data provenance **1**:442
Globally Unique Identifier (GUID) **1**:441–442
quality control **1**:442
scientific inquiry **1**:441
structured metadata
Access to Biological Collection Data (ABCD) standards **1**:443
Darwin Core **1**:443
Ecological Metadata Language (EML) **1**:443
eXtensible Markup Language (XML) **1**:442–443
geospatial data standards **1**:443–444
machine integration challenges **1**:442
Ecological Asset Inventory and Management (EcoAIM) software **5**:358
ecological character displacement **4**:46–47, **4**:46*F*, **4**:47*F*
ecological conservation **4**:315
ecological differentiation **2**:513–514
ecological footprint analysis (EFA) **2:701–713**
aquaculture **1**:196–198, **1**:196*F*
basic concepts **2**:705
definition **2**:701
environmental impact assessments **3**:291–292
evaluation assessments
conceptual strengths **2**:710
methodological limitations
antitrade/antigrowth considerations **2**:712–713
carbon footprint considerations **2**:712
general discussion **2**:710–711
oversimplification **2**:711–712
scope **2**:710–711
methodological strengths **2**:710
functional role
evolving economic paradigms **2**:702–703
global ecological change **2**:701–702
human ecology **2**:703
patch-disturbance species **2**:703–704
second law of thermodynamics **2**:703
human carrying capacity **2**:704–705
human ecological footprints
biodiversity crisis **2**:708–710
city footprints **2**:707
global development **2**:707–708
global sustainability **2**:706–707
social equity **2**:707–708
methodologies **2**:705–706, **2**:706
research summary **2**:713
ecological genetics **2:714–731**
eco-evolutionary dynamics
case studies **2**:726–727
conceptual developments **2**:727
population–community relationship **2**:726
future outlook **2**:727–728
genetic bases
allelic effects **2**:725
quantitative genetic models **2**:724–725
quantitative trait loci (QTL) mapping **2**:724
research insights **2**:725
genetic correlations
G matrix-based constraints **2**:721
life history tradeoffs **2**:720–721
limited genetic variation **2**:720
pleiotropy **2**:720
genetic variability **2**:715–717
heritability estimation **2**:715, **2**:715–717
historical perspective **2**:714
local adaptation versus gene flow
empirical research **2**:724
gene flow **2**:724
population genetics models **2**:723–724
site-specific adaptation **2**:723
microevolutionary perspective **2**:714–715
natural selection strength
discrete polymorphisms **2**:717
potential biases **2**:717–718
selection gradients **2**:717
phenotypic plasticity
adaptive strategies **2**:721
empirical research **2**:721–722
microevolutionary changes **2**:722–723
norm of reaction **2**:722, **2**:722*F*
theoretical predictions **2**:721–722
phenotypic traits **2**:715–717
quantitative genetics **2**:715
selection patterns
fluctuating selection **2**:718–719
lack of response **2**:719–720
rapid evolution **2**:718
selection strength estimation **2**:716
ecological geobotany **4**:396*T*
ecological integrity **3**:310–311, **3**:312*F*
Ecological Metadata Language (EML) **1**:441, **1**:443
ecological niche model **6:692–705**
basic concepts
geographical models **6**:694–695
niche theory **6**:693–694, **6**:694*F*
spatial autocorrelation (SAC) **6**:693
species response shapes **6**:692*F*, **6**:694
characteristics and functional role **6**:692–693, **6**:692*F*
community studies **6**:702–703
Cupressus chimera case study **6**:695
data collection methodology
biotic interactions and disturbances **6**:698–699
challenges **6**:695–697
environmental data
aquatic systems **6**:700
challenges **6**:699
climate **6**:699
land cover **6**:699–700
multicollinearity **6**:700
soils and substrate **6**:699–700
terrain **6**:699
species data
observational research **6**:696–697
presence-only data **6**:697–698, **6**:701–702
record characteristics **6**:697, **6**:698*F*
sample characteristics **6**:696–697, **6**:696*F*
site-occupancy models **6**:697, **6**:698*F*
modeling approaches
Bayesian methods **6**:702
decision trees **6**:701
general discussion **6**:700–702
generalized additive models (GAMs) **6**:701
generalized linear models (GLMs) **6**:701
genetic algorithms **6**:702
machine learning methods **6**:701, **6**:702*F*
maximum entropy theory (MaxEnt) **6**:702
profile and envelope methods **6**:701–702
new times and places predictions **6**:703
parasites
infection prevalence **5**:665–666, **5**:666*F*
microhabitat segregation **5**:665–666, **5**:667*F*
niche dimensions **5**:665, **5**:665*T*
niche saturation **5**:665, **5**:666*F*
proximate and ultimate causation **5**:665–666, **5**:667*F*
research summary **6**:703–704
species packing **4**:643
species richness **6**:702–703
taxonomic versus functional diversity **2**:649–650, **2**:649*F*
ecological refugees **6**:221*T*
ecological transfer efficiency **7**:265
ecology
basic concepts **3:1–8**
biodiversity theories **3**:4–5
ecosystem theories **3**:6–7
evolutionary theories **3**:7
historical development **3**:1
population dynamics **3**:1–3
research summary **3**:7–8
scaling laws **3**:6
spatial patterns **3**:5–6
species-interaction theories **3**:3–4, **3**:4*F*
definition **3**:1
ecological complexity versus biodiversity **2:201–212**
background information **2**:201
basic concepts **2**:204
biotic interactions and relationships **2**:202–204, **2**:203*F*, **2**:204*F*
community complexity **2**:203*F*, **2**:204–205, **2**:205*T*
entangled bank metaphor **2**:202–204
functional complexity **2**:203*F*, **2**:204–205, **2**:205*T*

historical perspective **2**:201–204
hypothetical relationships **2**:210–211, **2**:211*F*
measures and consequences
community complexity **2**:205–207
complexity measures **2**:205
functional complexity **2**:207–208
intricacy measures **2**:205*T*
phylogenetic complexity **2**:208, **2**:209*F*
usefulness **2**:208–209
phylogenetic complexity **2**:203*F*, **2**:204–205, **2**:205*T*
research summary **2**:210–211
schematic diagram **2**:203*F*, **2**:204*F*
significance
community complexity **2**:209–210
complexity in nature **2**:209–210
ecosystem functioning **2**:210
phylogenetic complexity **2**:210
theoretical perspectives **2**:210–211
ecological restoration **4:199–208**
advantages/disadvantages **4**:199–200
basic concepts **4**:199–200
ecosystem restoration potential **4**:203–204
explicit restoration targets **4**:205
future outlook
biodiversity crisis **4**:205
key challenges **4**:205–206
socioeconomic justifications **4**:206
plant conservation **6**:63, **6**:85–87
restoration by-products **4**:204
restoration justification **4**:204–205
restoration targets
characteristics **4**:200–201
ecosystem drivers **4**:201
ecosystem service restoration **4**:201–202
entire ecosystems **4**:200–201, **4**:202*T*
explicit restoration targets **4**:205
global climate change **4**:203
novel ecosystems **4**:202–203
whole-ecosystem restoration **4**:201
ecological succession **7:63–70**
basic concepts **7**:63
biodiversity
disturbed landscapes **7**:69–70
mature communities **7**:68–70
species gains and losses **7**:69*F*, **7**:70
ecosystem function **3**:93, **3**:93*T*
mechanistic theories
biomass dynamics **7**:67, **7**:67*F*
comparison studies **7**:65–66, **7**:65*T*
descriptive models **7**:65–66, **7**:66*F*
landscape mosaics **7**:66–68, **7**:67*F*
quantitative models **7**:68
vegetation patterns **7**:65–66
organic explanations
background information **7**:63–64
Clementsian theory **7**:63–64
Cowles' theory **7**:64–65
neutral theory **5:510–518**
characteristics
dispersal **5**:511
general discussion **5**:511
speciation **5**:511
stochastic drift **5**:511
definition **5**:510, **5**:511
future outlook
general discussion **5**:516
testing programs **5**:516
theoretical developments **5**:516–517
theory applications **5**:517
general discussion **5**:510
goals **5**:510–511
historical background **5**:511–512
modeling approaches
Hubbell's spatially implicit neutral model **5**:512–514
spatially explicity neutral models **5**:514–515
speciation modes **5**:515
species abundance distributions (SADs) **5**:513–514, **5**:513*F*
species–area relationships (SARs) **5**:514
nonneutral extensions **5**:515–516
research summary **5**:517
schematic diagram **5**:512*F*
species abundance distributions (SADs) **5**:510–511, **5**:513*F*
species–area relationships (SARs) **5**:510–511, **5**:514
EcoMetrix software **5**:358
economically significant C_4 plants **2**:362*T*
economic botany **3**:323
economic growth **3:25–30**
environmental Kuznets curves **3**:25, **3**:28–29, **3**:28*F*
environmental resources **3**:25
gross domestic product (GDP) **3**:25
net national product (NPP)
background information **3**:25–26
social well-being
allocation mechanisms **3**:26
basic concepts **3**:26
constant-equivalent consumption measures **3**:27
cost–benefit rule **3**:27
population changes **3**:26–27
sustainable development **3**:27–28
productivity growth analysis **3**:29–30
Economics of Ecosystems and Biodiversity (TEEB) **4**:330
The Economics of Ecosystems and Biodiversity (TEEB) **1**:333–335, **1**:337–339, **2**:315
econotheism **6**:380–381
EcoPath with EcoSim (EwE) software **5**:459
ecosystem engineers
ecological restoration **4**:201
mangrove ecosystems **5**:16–17
marine ecosystems **5**:60–61
Ecosystem Function Conservation Areas (EFCAs) **3**:101
ecosystem overfishing **3**:445
ecosystems **3:59–63**, **3:169–175**
Africa **1:45–57**
Afro-montane forests **5**:219
alpine life zone **1**:149*T*
climate **1**:46
conservation status **3**:171–173*T*
fire/fire regimes **1**:46–47
future outlook **1**:56–57
geographic characteristics **1**:45
geologic history **1**:45–46
Madagascar **1**:55–56
past climatic fluctuations **1**:47
phytogeographic/ecoclimatic zones
Cape Region (South Africa) **1**:47–48
Mediterranean North Africa **1**:47
tropical Africa **1**:48
plant diversity **6**:314–315
river water composition **3**:562*T*
soils **1**:46
termites **4**:430*T*
tropical Africa
afroalpine ecosystems **1**:53
bushlands **1**:51–52
coastal ecosystems **1**:55
deserts **1**:53
forests **1**:48–49
grasslands **1**:52
lake and pond ecosystems **1**:54–55
montane ecosystems **1**:53
phytogeographic/ecoclimatic zones **1**:48
seasonal tropical vegetation **1**:49–51
shrublands **1**:52–53
thickets **1**:51–52
wetlands **1**:53–54
woodlands **1**:49–51
tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*
vegetation classifications **1**:45
agroecosystems **3:9–15**
background information **3**:9
biodiversity–human health relationship **1**:359–360
future outlook **3**:15
landscape diversity
ecological consequences **3**:13
general discussion **3**:13
organism spillover **3**:13–14
spatial structure **3**:14–15, **3**:14*F*
pest control **1:373–385**
agricultural intensification effects **1**:374–375
agricultural landscapes **1**:380–381
background information **1**:373–374
conservation biological control **1**:382
crop diversity **1**:375–376
habitat heterogeneity **1**:381–382, **1**:381*F*
influencing factors **1**:374*F*
multiple trophic levels **1**:380
natural enemies hypothesis **1**:375–376
natural habitat proximity **1**:381
noncrop habitats **1**:380–381
noncrop plants **1**:376, **1**:376*F*
predator diversity **1**:376–377
research background **1**:373–374
summary discussion **1**:382–383
vegetation diversity **1**:375–376
planned diversity
characteristics **3**:9–11, **3**:10*F*

ecosystems (*continued*)
host–parasite interactions **3**:11–12
pest control **3**:11–12
productivity **3**:9–11, **3**:10*F*
soil processes **3**:12
stability **3**:9–11
planned versus unplanned biodiversity **1**:127*F*, **3**:9, **3**:31–34
termites **4**:432*T*
unplanned diversity
ecological services **3**:12–13
pest control **3**:12–13
soil processes **3**:13
alpine ecosystems **1:148–157**
alpine life zone **1**:148–149, **1**:149*F*, **1**:149*T*
animals and microorganisms
animal diversity **1**:154–155, **1**:155*F*
microbial diversity **1**:155, **1**:155*F*
bioclimatic mountain belts **1**:149*T*
biodiversity functions **1**:155–156
global area **1**:149*T*
global change impacts **1**:156–157, **1**:156*F*
habitat diversity
climatic microhabitats **1**:149–150, **1**:150*F*
substrate diversity **1**:150–151
high-elevation Andean ecosystems **4:96–110**
background information **4**:96
central Andean dwarf *Polylepis* forest **4**:105
cushion bogs **4**:105
endemism **4**:99–100, **4**:99*T*
geographic distribution **4**:96–97, **4**:97*F*
habitat characteristics **4**:96–97
number of species **4**:102*T*
origins and diversification **4**:100–101
páramo ecosystem **4**:101–103, **4**:102*F*, **4**:102*T*
physical characteristics and patterns **4**:97–98
plant diversity **4**:98–99
puna **4**:97–98, **4**:97*F*, **4**:102*T*, **4**:103–104, **4**:104*F*
research summary **4**:107
salar vegetation **4**:105
southern Andean steppes **4**:105–107, **4**:105*F*
species richness **4**:98–99
plant diversity
leaf mass ratio (LMR) **1**:152*F*
morpho-type diversity **1**:151, **1**:151*F*
physio-type diversity **1**:151–153
reproductive diversity **1**:153
specific leaf area (SLA) **1**:152, **1**:152*F*
taxonomic diversity **1**:153–154, **1**:154*F*
Amazon ecosystems **1:158–168**
Amazonian savannas
cerradão **1**:164–165
flooded savannas **1**:165
nonflooded savannas **1**:164–165
background information **1**:158
Biological Dynamics of Forest Fragments Project (BDFFP) **4**:663
deforestation **4**:663
ecosystem function measurements **3**:81*T*
El Niño Southern Oscillation (ENSO) **3**:156–158
floodplain ecosystems
buritizal **1**:164
igapó forest **1**:163–164
lakes **1**:164
mangrove forests **1**:164
permanent swamps **1**:164
pirizal **1**:164
seasonal flooding **1**:162–163
várzea forest **1**:163
indigenous peoples **4**:273*F*, **4**:273*T*
mammals **2**:591*F*
montane fringing ecosystems
Andes foothills **1**:166–167
campos rupestres **1**:167
characteristics **1**:165–166
granite inselbergs **1**:166
Guayana Highland **1**:166
low hills **1**:166
sandstone tepuis **1**:166
myriapods **1**:246*T*, **1**:247, **5**:431
nonflooded terra firme ecosystems
cerradão **1**:161–162
dry semideciduous southern fringe forests **1**:161–162
general discussion **1**:158–160
rain forests **1**:158–160, **1**:159*F*, **1**:159*T*
white sand oligotrophic formations **1**:160–161
plant domestication **4:279–292**
basic concepts **4**:279
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
research summary **1**:167–168
secondary forest **1**:167
threatened plant species **6**:84
transition ecosystems
babassu palm forests **1**:162
bamboo forests **1**:162
characteristics **1**:162
liana forests **1**:162
vascular epiphytes **3**:520*T*
Antarctica **1:179–188**
anthropogenic impacts
alien species invasion **1**:185–186, **1**:186*T*
biological consequences **1**:186–187
historical perspective **1**:185–186
regional and global environmental change effects **1**:186
biodiversity
current studies **1**:183–184
origins and antiquity **1**:184
plants **1**:184*T*
terrestrial invertebrates **1**:184*T*
biological history **1**:179–180
climate change effects **5**:89–91
freshwater ecosystems **1**:183
future trends
anthropogenic impacts **1**:185–186
climate change effects **1**:186–187
geographic characteristics **1**:179, **1**:180*F*
glacial history **1**:179–180
natural colonization processes
colonization mechanisms **1**:185
contemporary colonization **1**:184–185
propagule dispersion **1**:185
tectonic history **1**:179–180
terrestrial biogeographic zones
continental Antarctic zone **1**:181, **1**:181*T*, **1**:182*T*, **1**:184*T*, **1**:186*T*
general characteristics **1**:180–181, **1**:180*F*
maritime Antarctic zone **1**:181, **1**:181*T*, **1**:182*T*, **1**:184*T*, **1**:186*T*
sub-Antarctic zone **1**:180–181, **1**:181*T*, **1**:182*T*, **1**:184*T*, **1**:186*T*
temperature ranges **1**:181*T*
terrestrial biota
fauna **1**:182, **1**:182*T*
flora **1**:182–183, **1**:182*T*
habitats **1**:181–182
microbiota **1**:182*T*, **1**:183
ultraviolet radiation effects **7**:300
Arctic ecosystems
Census of Marine Life project **2**:20–21, **2**:21*F*
climate change effects **5**:89–91
Europe **3**:335*T*
fish diversity **3**:474
psychrophiles **6**:299
terrestrial ecosystems **1:227–244**
biodiversity filters **1**:238
biodiversity patterns **1**:231–235, **1**:231*T*
circumpolar diversity **1**:235–236
current biodiversity status **1**:242
environmental characteristics **1**:227–231, **1**:229*F*
environmental consequences **1**:238–240, **1**:239*F*, **1**:240*F*
general discussion **1**:227
latitudinal biodiversity patterns **1**:232–235, **1**:233–234*T*, **1**:235*F*, **1**:236*F*
natural resources/environmental exploitation **1**:240–241
research organizations **1**:242–243
research summary **1**:243

threats and challenges **1**:241–242
topographic impacts **1**:236–238, **1**:237*F*
vegetation zones **1**:229*F*, **1**:230*T*
Asia **1:254–280**
alpine life zone **1**:149*T*
animal species **1**:258, **1**:261*T*
artificial ecosystems **1**:279
biogeographic provinces **1**:262–263*T*
biomes **1**:258–260, **1**:262–263*T*
boreal and polar Asia
boreal forest **1**:262–263*T*, **1**:273–274
characteristics **1**:273–274
polar and mountain tundra **1**:262–263*T*, **1**:274–275
climatic zonation **1**:257*T*, **1**:258, **1**:262–263*T*
conservation status **1**:279
dry regions
characteristics **1**:267–268
Dzungaria **1**:270
Gobi Desert **1**:270
Mediterranean ecosystems **1**:262–263*T*, **1**:268–269
Middle Asian deserts **1**:270
subtropical deserts **1**:262–263*T*, **1**:267–268
Tarim Basin **1**:270
temperate grasslands **1**:262–263*T*, **1**:269
temperate semi-deserts and deserts **1**:262–263*T*, **1**:269–270
Tibetan Plateau **1**:270
evolutionary history **1**:256*T*
geographic characteristics **1**:255
geologic history **1**:256, **1**:256*T*
humid monsoon region
characteristics **1**:262–263*T*, **1**:270–271
laurel forests **1**:270–271
mountain regions **1**:272–273
pine forests **1**:272
temperate deciduous forests **1**:271
physiographic characteristics **1**:255–258, **1**:255*F*
plant species **1**:258, **1**:259*T*, **1**:260*T*
regionalization **1**:258–260
tropical Asia
characteristics **1**:260–263
humid tropical forest **1**:260–263, **1**:262–263*T*, **1**:266*T*
monsoon forests (moist deciduous forests) **1**:262–263*T*, **1**:263–266
tropical dry forest, scrub, and savanna **1**:262–263*T*, **1**:266–267
tropical moist forest **1**:277*T*
tropical/subtropical humid mountain areas **1**:262–263*T*, **1**:266*T*, **1**:267
vegetation belts **1**:266*T*
vegetation types and coverage percentage **1**:264–265*T*
wetlands and coastal ecosystems **1**:275–279, **1**:275*T*, **1**:276*T*
Australia
alpine life zone **1**:149*T*
biogeographic regions **1**:297–298, **1**:298*F*, **1**:299*F*
characteristics **1**:291–292
floristic groups
biogeographic regions **1**:298–300
diagnostic species **1**:292–297
Mediterranean-type climate ecosystems **5**:221
TWINSPAN floristic groups **1**:292, **1**:294*T*, **1**:300*F*
plant communities **1:291–306**
biogeographic regions **1**:297–298, **1**:298*F*, **1**:299*F*
characteristics **1**:291–292
community diversity **1**:300, **1**:300*F*
evaporative aerodynamics **1**:292, **1**:293*F*
floristic groups **1**:292, **1**:294*T*
Mediterranean-type climate ecosystems **5**:221
research summary **1**:304
species richness **1**:300, **1**:301*T*
structural characteristics **1**:301*T*
Southern Hemisphere **6:612–619**
angiosperms **6**:617–618
background information **6**:612
biotic dispersal **6**:616
climate change effects **6**:617
climatic patterns **6**:614
continental drift **6**:616
dispersal ability **6**:613–614
fossil record **6**:615–616
general area cladograms **6**:614
geographic distribution patterns **6**:613–614
high southern latitude plant growth **6**:618
historical research **6**:612–613
long-distance dispersal **6**:616–617
relaxed clock molecular dating analyses **6**:614–615
research summary **6**:618–619
vicariance biogeography **6**:616
species richness
dryland vegetation continua **1**:302, **1**:303*F*
invasive species **1**:303
invertebrates **1**:300–302, **1**:303*T*
isolated rainforest stands **1**:303–304
overstory and understory strata **1**:300, **1**:301*F*, **1**:302*F*
postfire succession **1**:302–303, **1**:304*F*
vertebrates **1**:300–302, **1**:302*F*
wetland vegetation continua **1**:302
basic concepts **3**:59–60
black box model **3**:60, **3**:60*F*, **3**:61*F*
Central America **2:23–32**
biodiversity
climatic regions/latitudinal range **2**:24–26
endemism **2**:26
floristic elements **2**:25–26
species richness **2**:25, **2**:25*T*
biogeography **2**:23–24, **2**:24*F*
global environmental change
climatic change **2**:30
general discussion **2**:30
land use change **2**:30–31, **2**:31*F*
high-elevation ecosystems **2**:30
life zones **2**:26–27
mangroves **2**:29
seasonally dry tropical forests **2**:28
temperate forests **2**:29–30
tropical cloud forests **2**:28–29
tropical rain forests **2**:27–28
coastal beach ecosystems **2:128–136**
sandy beaches
alongshore distribution **2**:131–133
background information **2**:128
beach index (BI) **2**:129
future research areas **2**:135
macrofauna surveys **2**:129–131, **2**:130*F*
macroscale patterns **2**:131
mesoscale patterns **2**:131–133
microscale patterns **2**:133
microtidal regimes **2**:128–129, **2**:129*F*
morphodynamic states **2**:128–129, **2**:129*F*
research summary **2**:135
species richness hypotheses **2**:133–134
threats and challenges **2**:134–135
tidal regimes **2**:129
two-dimensional model **2**:129, **2**:130*F*
zonation patterns **2**:131–133
coastal ecosystems
Asia **1**:275–279, **1**:275*T*, **1**:276*T*
Australian plant communities **1**:292–297
biodiversity restoration **6**:456–457
Census of Marine Life project **2**:15*F*, **2**:16, **2**:17*F*
climate change effects **6**:150–151
coastal wetland restoration **7**:375–378
conservation status **3**:171–173*T*, **3**:174
ecosystem services **1**:334*T*, **3**:96–98
global extent **7**:370*T*
myriapods **5**:430–431
nature-based coastal protection strategies **5:451–458**
biodiversity impacts **5**:453–456, **5**:455*F*
current status **5**:452–453, **5**:453*T*
historical perspective **5**:452
importance **5**:451–452
research summary **5**:456
nitrogen cycle impacts **5**:543
predation effects **6**:237–239
salt water plants **6**:77
tropical Africa **1**:55
conservation status
assessment methods **3**:169–170
coastal ecosystems **3**:171–173*T*, **3**:174
freshwater ecosystems **3**:171–173*T*, **3**:174
marine ecosystems **3**:171–173*T*, **3**:174
terrestrial ecosystems **3**:170–174, **3**:171–173*T*

ecosystems (*continued*)
cooperation **2**:328–329
coral reef ecosystems **2:330–346**
biodiversity
diversity patterns **2**:333–334
extinction risks **4**:664
interspecific interactions **2**:334
major groups **2**:331
species diversity **2**:331–333, **2**:332*T*, **2**:333*T*
biodiversity-rich countries **1**:500*T*
Census of Marine Life project **2**:15*F*, **2**:16
climate change effects **3**:449–450, **4**:664, **6**:150–151
conservation practices
greenhouse gas emissions **2**:344
local management and restoration **2**:343–344
conservation status **3**:171–173*T*, **3**:174
corals
algal overgrowth **2**:338
anthropogenic causes **2**:340
characteristics and classifications **2**:334–335
classifications **4**:381*T*
climate change effects **5**:37
conservation practices **2**:343–344
coral bleaching **2**:338, **2**:338*F*, **2**:339*F*, **2**:340*F*, **2**:709–710, **3**:160–161, **3**:160*F*
coral-dwelling microbes **2**:335
corallivore explosions **2**:337–338
direct destruction **2**:340
disappearance **2**:336–337, **2**:336*F*, **2**:337*F*
disease incidence **2**:338–340, **2**:339*F*, **2**:340*F*
diversity **2**:334–335
ecosystem goods and services **2**:343
evolutionary relationships **2**:334–335
greenhouse gas emissions **2**:341–342, **2**:342*F*, **2**:344
indicator species **7**:206–207*F*
infectious diseases **2**:531–532
invasive species **2**:341
loss patterns and extent **2**:337
mortality agents **2**:337–338
nutrition and growth **2**:335–336
overfishing **2**:341, **5**:61
photosynthetic symbioses **5**:118–120, **5**:119*F*
poor water quality **2**:340–341
reproduction and recruitment **2**:335–336
species loss **2**:342–343, **5**:597
threatened species **7**:206–207*F*
disappearance
anthropogenic impacts **2**:336–337, **2**:336*F*, **2**:337*F*
loss patterns and extent **2**:337
ecosystem diversity **3**:452–453, **5**:52
ecosystem engineers **4**:201
ecosystem services **1**:334*T*, **5**:465–466
El Niño Southern Oscillation (ENSO) events **3**:160–161, **3**:160*F*
fish stocks **3**:481–482
global species richness **3**:710
grazer-based ecosystem restoration and management **4**:16
grazing effects **4**:8–9, **4**:9*F*, **4**:10*F*, **6**:40
habitat destruction **3**:404, **3**:405*F*, **3**:447–448
indicator species **7**:206–207*F*
Late Devonian extinction event **5**:169–170
Late Permian extinction event **5**:171
loss consequences
ecosystem goods and services **2**:343
species loss **2**:342–343
marine sedimentary habitats **5**:108*T*
mortality agents
algal overgrowth **2**:338
anthropogenic causes **2**:340
coral bleaching **2**:338, **2**:338*F*, **2**:339*F*, **2**:340*F*
corallivore explosions **2**:337–338
cyanide fishing **4**:156
direct destruction **2**:340
disease incidence **2**:338–340, **2**:339*F*, **2**:340*F*
greenhouse gas emissions **2**:341–342, **2**:342*F*
invasive species **2**:341
overfishing **2**:341, **5**:61
poor water quality **2**:340–341
nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
nitrogen cycle impacts **5**:543
overfishing **3**:452–453, **5**:54
protection and restoration models **5**:465–466
spatiotemporal distributions **2**:330–331, **2**:331*F*
species numbers **5**:108–109
threatened species **7**:201*T*, **7**:206–207*F*
cultural ecosystem services **1:332–340**
basic concepts **1**:333–335, **1**:333*F*, **2**:490*T*, **7**:315–316
cultural diversity effects **1**:335–336, **1**:335*F*
cultural services **1**:334*T*
desertification **2**:490*T*
esthetic services **5**:353–354
general discussion **1**:333
habitat services **1**:334*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:353
nature-based recreation models **5**:353
provisioning services **1**:334*T*
regulating services **1**:334*T*
spiritual services **5**:354
sustainable development **1**:337–339
valuation language/value articulating institutions **1**:336–337, **1**:337*T*
definition **3**:59, **3**:169–170
disturbance ecology **2:562–567**
anthropogenic impacts **2**:566–567
basic concepts **2**:562–563
carnivores **2**:565–566
herbivores **2**:565–566
interaction impacts **2**:566
intermediate disturbance hypothesis **2**:115–117, **2**:563, **4**:79
mechanical disturbances **2**:563–564
physico–chemical disturbances **2**:564–565, **2**:565*F*
dune ecosystems **5**:452–453, **5**:453*T*
ecological perspectives **3**:6–7
ecosystem boundaries
analytical studies **3**:241–242
basic concepts **3**:238–239, **3**:238*F*
control mechanisms **3**:241, **3**:241*F*
food web–trophic level relationship **3**:240–241, **3**:241*F*
historical perspective **3**:238–239
trophic dynamic concept **3**:239–240
ecosystem diversity **1**:402, **3**:6–7
ecosystem function **3:90–95**
basic concepts **3**:90
ecological succession **3**:93, **3**:93*T*
ecosystem productivity **3**:91–92, **3**:91*T*, **4**:156–158
ecosystem services **3**:90–91, **3**:90*T*
ecosystem stability **3**:93–94
energy flow **3**:237–238
fire/fire regimes **3**:439
keystone species **3**:92–93, **3**:92*F*, **3**:246–247
Liebig's law of minimum **3**:91–92
nutrient cycling **3**:93
research areas **3**:91, **3**:91*T*
research summary **3**:94–95
ecosystem function measurements **3:64–71**
basic concepts **3**:64
biogeochemical cycles **3**:68–70
comparison studies **3**:64–65
new and regenerated production **3**:69–70
nutrient cycling **3**:68–70
primary production
carbon-14 (^{14}C) measurements **3**:66
carbon-based measurements **3**:65–66
data interpretation and analysis **3**:66–67
fluorescence measurements **3**:66
measurement approaches **3**:65
oxygen-based measurements **3**:65
secondary production
bioenergetic analysis **3**:67
biomass accrual **3**:67
birth and death rate analysis **3**:67–68
egg ratio analysis **3**:68
functional type analysis **3**:68
instar analysis **3**:68
measurement approaches **3**:67
stable isotope studies **3**:70–71
terrestrial ecosystems **3:72–89**
analytical and measurement approaches **3**:72–73, **3**:74–75*T*, **3**:76*F*
biodiversity **3**:84–86, **3**:84*F*
field manipulations **3**:86*F*, **3**:87*F*, **3**:88
functional diversity (FD) **3**:84
monitoring networks **3**:86*F*, **3**:87*F*, **3**:88

natural versus synthesized ecosystems **3**:88
plant functional traits **3**:85*F*, **3**:85*T*, **3**:86
positive-feedback switches **3**:86–87, **3**:86*F*
purpose **3**:72
real ecosystems **3**:88
resistance and resilience strategies **3**:83–84, **3**:83*F*
resource dynamics **3**:72–73, **3**:74–75*T*, **3**:75*F*
synthesized ecosystems **3**:86*F*, **3**:87*F*, **3**:88
taxonomic diversity **3**:84
ecosystem productivity **1**:386–388, **1**:392, **3**:91–92, **3**:91*T*, **4**:156–158
ecosystem services **1:341–356**, **3:96–104**
basic concepts **1**:344, **7**:315–316
biodiversity conservation **4**:435
biodiversity–ecosystem services relationship **1**:341, **3**:342, **3**:342
biodiversity impacts **1**:389–391, **1**:395–396, **1**:396*F*, **7**:316
biophysical measures **5**:349
carrying capacity **7**:75
China **3:372–384**
background information **3**:372
Ecosystem Function Zones (EFZs) **3**:382
Grain-to-Green Program (GTGP) **3**:379–380
land-use planning and human development **3**:101
National Forest Conservation Program (NFCP) **3**:375–376
Nature Reserve System (NRS) **3**:372–373, **3**:373*F*
payment for ecosystem services (PES) **2**:227
summary discussion **3**:382
classifications **3**:96, **3**:97*F*
climate change effects **2**:105
coral reef ecosystems **2**:343
countryside biogeography **2**:352–353
cross-system services **3**:96–98
cultural ecosystem services **1:332–340**
basic concepts **1**:333–335, **1**:333*F*, **2**:490*T*, **7**:315–316
cultural diversity effects **1**:335–336, **1**:335*F*
cultural services **1**:334*T*
desertification **2**:490*T*
esthetic services **5**:353–354
general discussion **1**:333
habitat services **1**:334*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:353
nature-based recreation models **5**:353
provisioning services **1**:334*T*
regulating services **1**:334*T*
spiritual services **5**:354
sustainable development **1**:337–339
valuation language/value articulating institutions **1**:336–337, **1**:337*T*
cultural services
basic concepts **7**:315–316
landscape scale **1**:344–345
regulating services **1**:334*T*
sustainability **7**:76*T*
decision-making processes
coastal and marine spatial planning **3**:100–101
global policy and research efforts **3**:101–102, **3**:102*F*
land-use planning and human development **3**:101
service framework **3**:99–100, **3**:99*F*
watershed management **3**:100
definition **3**:96
desertification **2**:490*T*, **2**:491
ecological restoration **4:199–208**
advantages/disadvantages **4**:199–200
basic concepts **4**:199–200
ecosystem drivers **4**:201
ecosystem restoration potential **4**:203–204
ecosystem service restoration **4**:201–202, **4**:202*T*
entire ecosystems **4**:200–201, **4**:202*T*
explicit restoration targets **4**:205
future outlook **4**:205
global climate change **4**:203
novel ecosystems **4**:202–203
plant conservation **6**:63, **6**:85–87
restoration by-products **4**:204
restoration justification **4**:204–205
restoration targets **4**:200–201
whole-ecosystem restoration **4**:201
ecosystem function **3**:90–91, **3**:90*T*
ecosystem functioning **1**:342–344, **3**:342
ecosystem processes **1**:342–344, **3**:342
energy flow **3**:247
future challenges
distributional effects **3**:102–103
dynamic effects **3**:103
ecosystem services–human health relationships **3**:102
general discussion **3**:102
institutional design **3**:103
monetary and non-monetary valuations **3**:103
research areas **3**:103
shocks and uncertainty **3**:103
tradeoffs and synergies **3**:102
general discussion **3**:96
global status **3**:46*T*
global urbanization impact
dependence interactions **4**:237–238, **4**:238*F*
environmental impact **4**:235
management strategies **4**:239*F*, **4**:241–242
habitat services **1**:334*T*
historical background
initial development **3**:98–99
recent research **3**:99
service disruptions **3**:98–99
service framework **3**:99, **3**:99*F*
human well-being **1**:345, **1**:436, **7**:77–80, **7**:78*F*
Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES) **2**:246–247, **3**:101–102
landowner decisions **5**:347–349
land use/land cover (LULC) patterns **5**:348–349, **5**:351–352
marine ecosystems **5:451–468**
accounting assessment models **5**:455–457
American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries **5**:464–465, **5**:464*F*
bioeconomic models **5**:457–459
challenges **5**:455
coastal zone management (Belize) **5**:465, **5**:465*F*
conceptual framework **5**:452–453
coral reef protection and restoration **5**:465–466
cost–benefit analyses **5**:459–463
coupled social–ecological systems model **5**:453–455, **5**:454*F*
cross-system services **3**:96–98
cultural ecosystem services **5**:451–452, **5**:452*T*
decision-support tools **5**:459–463, **5**:460–462*T*, **5**:463*F*
future outlook **5**:466
integrated ecosystem assessments (IEAs) **5**:454–455
modeling approaches and uses **5**:453
multiple ecosystem services **5**:459–463
production function models **5**:457–459
provisioning services **5**:451–452, **5**:452*T*
regulating services **5**:451–452, **5**:452*T*
revealed references methods **5**:456–457
single ecosystem services **5**:457–459
stated preference methods **5**:457
supporting services **5**:451–452, **5**:452*T*
tradeoffs **5**:459–463
valuation measures **5**:457
wind farms **5**:466
marine sedimentary habitats
filtration processes **5**:112
general discussion **5**:109
nutrient cycling **5**:111
pollutant cycling **5**:111–112
schematic diagram **5**:111*F*
secondary production **5**:112
shoreline stability **5**:112
species diversity connection **5**:112–113
modeling frameworks
animal pollination **5**:354–355
basic principles **5**:349–350
carbon storage and sequestration **5**:351–352
consumable goods **5**:349–350

ecosystems (*continued*)
cost-effectiveness measures **5**:356–357, **5**:357*F*, **5**:459–463
crop production **5**:349–350
cultural ecosystem services **5**:353
decision-making strategies **5**:358
esthetic services **5**:353–354
nature-based recreation models **5**:353
nutrient retention **5**:355–356
provisioning services **5**:350–351
regulatory services **5**:351–352
software **5**:357–358
spiritual services **5**:354
storm peak management **5**:352–353
summary discussion **5**:358–359
supporting services **5**:354–355
uncertainty monitoring **5**:356
nature-based recreation **7**:253
payment for agrobiodiversity conservation services (PACS) **3**:40–41, **3**:41*F*
payment for ecosystem services (PES) **2**:226–227, **2**:298, **3**:40–41, **3**:432–433, **5**:151, **7**:353–355
plant communities **6**:76
population diversity
general characteristics **6**:170–171
global population diversity **6**:171
local population diversity **6**:171
regional population diversity **6**:171
priority management **6:261–272**
carbon sequestration **5**:351–352, **6**:264*F*
carbon storage **5**:351–352, **6**:264*F*
challenges **6**:270–271
cost–benefit analyses **6**:265–266, **6**:266*F*
economic factors **6**:266–269, **6**:267*F*, **6**:268*F*, **6**:269*F*
ecosystem service targets **6**:263–265, **6**:264*F*
land-use patterns **6**:265–266, **6**:266*F*, **6**:268*F*, **6**:269*F*
livestock **6**:264*F*
mapping and modeling approaches **6**:263–265
research summary **6**:269–271
systematic conservation planning **2**:293, **6**:261–263
water provisions **6**:264*F*
provisioning services
basic concepts **2**:490*T*, **7**:315–316
cultural ecosystem services **1**:334*T*
desertification **2**:490*T*
landscape scale **1**:344
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:350–351
sustainability **7**:76*T*
regulating services
basic concepts **2**:490*T*, **7**:315–316
cultural ecosystem services **1**:334*T*
desertification **2**:490*T*
global status **3**:46–47, **3**:46*T*
landscape scale **1**:344
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:351–352
structure–function relationship **3**:46–47, **3**:47*T*
sustainability **7**:76*T*
relationship characteristics and shape **1**:345, **3**:351
relationship examples
agricultural food production **1**:346, **3**:345
general discussion **1**:346
human infectious disease regulation **1**:348–350, **3**:349
pollination **1**:348, **3**:347
soil productive potential regulation **1**:346–348, **3**:347
research areas **3**:91, **3**:91*T*
research summary **1**:355
river–floodplain connectivity
benefits **7**:145–146, **7**:150
fisheries **7**:146–147
flood attenuation **7**:147–150, **7**:148*F*, **7**:149*F*
floodplain agriculture **7**:146
flow levels and events **7**:150
hydrological connectivity **7**:150
overbank flooding **7**:145–146
spatial variables **7**:150
service disruptions **2**:454
service providers **1**:345
social value **7:314–326**
compensated versus uncompensated demand **7**:321–322, **7**:322*F*
contingent valuation method **7**:324–325
estimation methods **7**:322
hedonic valuation estimation methods **7**:322–324
Hicksian demand function **7**:321–322
historical perspective **7**:314–315
Marshallian demand function **7**:321–322
revealed-preference estimation methods **7**:322–324
stated-preference estimation methods **7**:324–325
summary discussion **7**:325
travel cost estimation methods **7**:322–324
valuation measures **7**:317–318, **7**:317*F*
value definition **7**:316–317
value quantification **7**:318–321
spatially variable ecosystem services **3**:53
spatial scale analyses
experimental manipulation effects **3**:349
landscape scale **1**:351–353, **3**:353, **3**:354
local scale **1**:350–351, **3**:349
plant functional diversity **3**:354
regional to global scale **1**:353–355
spatiotemporal scales **1**:345–346
structure–function relationship **3**:46–47, **3**:47*F*
supply–demand–value relationship **5**:348–349
supply–service–value metrics **3**:98, **3**:98*F*
supporting services
animal pollination **5**:354–355
basic concepts **2**:490*T*, **7**:315–316
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:354–355
nutrient retention **5**:355–356
sustainability **7**:76*T*
sustainable agriculture **1**:114
temperate forests **7**:124–125
terrestrial ecosystems **5:347–361**
biophysical measures **5**:349
cost-effectiveness measures **5**:356–357, **5**:357*F*
decision-making strategies **5**:358
landowner decisions **5**:347–349
land use/land cover (LULC) patterns **5**:348–349, **5**:351–352
modeling approaches **5**:349–350
summary discussion **5**:358–359
supply–demand–value relationship **5**:348–349
uncertainty monitoring **5**:356
valuation measures **7:167–179**
accounting price **7**:172
biological resource exploitation **7**:167–168
discontinuous value functions **7**:175–177, **7**:177*F*
ecological functions **7**:168–169
gross domestic product (GDP) **7**:173–174, **7**:174*F*
institutional failure **7**:170–172
local versus global stability **7**:170–172
market failure **7**:169–170
Millennium Ecosystem Assessment (MEA) **7**:168–169
necessity versus luxury perspective **7**:173–174
option value **7**:172
project evaluation **7**:172–173
resource evaluation **7**:172–173
social cost–benefit analysis **7**:172–173
substitution possibilities **7**:174–175
summary discussion **7**:177–178
water funds **7:352–366**
basic concepts **7**:352–353
Brazil **7**:363–364
Colombia **7**:362–363
design and implementation guidelines **7**:360
Ecuador **7**:361–362
financial sources and management **7**:357*F*, **7**:359
governance **7**:359–360
historical perspective **7**:355–357
Latin America and the Caribbean region **7**:354–355*F*, **7**:356*T*

operational and financial model **7**:357–358, **7**:357*F*
public protected area management **7**:358, **7**:358*T*
water fund platform **7**:364–365
watershed management **7**:357–358
endangered ecosystems
assessment methods **3**:170
conservation status **3**:170–174, **3**:171–173*T*
Millennium Ecosystem Assessment (MEA) **3**:170, **3**:284
energy flow **3:237–249**
benefits **3**:237–238
biodiversity effects
experimental research **3**:247
number of species effects **3**:245–246
research background **3**:245–247
top-down versus bottom-up regulation **3**:246–247
ecosystem boundaries
analytical studies **3**:241–242
basic concepts **3**:238–239, **3**:238*F*
control mechanisms **3**:241, **3**:241*F*
food web–trophic level relationship **3**:240–241, **3**:241*F*
historical perspective **3**:238–239
trophic dynamic concept **3**:239–240
future research outlook **3**:247–248
importance **3**:247
multiple energy pathways
background information **3**:242–243
depth–light wavelength relationship **3**:243
lake and pond ecosystems **3**:242–243, **3**:243*F*
spatial heterogeneity **3**:243
nutrient cycling
benefits **3**:237–238
food chains **3**:244–245
internal versus external supplies **3**:243–245
microbial loops **3**:244–245
organic detritus **3**:244–245
species interactions **3**:247
estuarine ecosystems **3:297–308**
biodiversity
characteristics **3**:299–300
ecological functions **3**:300–301
fish **3**:299–300
characteristics and distribution **3**:297–298
Chesapeake Bay case study **3**:303–306, **3**:304*F*, **3**:305*F*, **3**:306*F*
classifications **3**:298–299
climate change effects **6**:150–151
ecological functions
biogeochemical cycles **3**:301
biological properties **3**:301
chemical modifications **3**:301
climate variations **3**:302
ecosystem modifications **3**:301
exchange of materials **3**:301
extreme events **3**:302
freshwater delivery **3**:301
general discussion **3**:300–301
physico–chemical properties **3**:301
sea level changes **3**:302
suspended matter **3**:301
temperature changes **3**:302
transport changes **3**:301
wind **3**:302
eutrophication
aquatic vascular plants **3**:360–361
contaminant–stress relationship **5**:29–30
high nutrient, low growth (HNLG) concept **5**:26–27, **5**:26*T*, **5**:27*F*
macroalgae **3**:358–359
macrophytes **3**:360–361
microfauna **3**:363–364, **3**:363*F*
nutrient ratio changes **5**:28–29, **5**:28*T*, **5**:29*F*
functional role **3**:297–298
future challenges **3**:306–307
habitat loss and degradation **7**:371*T*
marine ecosystems **5**:45
marine sedimentary habitats **5**:108*T*
nitrogen cycle impacts **5**:543
salt water plants **6**:77
Europe **3:333–346**
alpine life zone **1**:149*T*
biological and cultural diversity **3**:333
ecosystem diversity
general discussion **3**:333–334
vegetation classification **3**:333–334
habitat and vegetation diversity
Arctic desert biome **3**:335*T*
biotic interactions **3**:334*T*, **3**:335
boreal forest biome **3**:335*T*, **3**:336–340*T*
continental semidesert biome **3**:335*T*, **3**:336–340*T*
habitat diversity **3**:334–335, **3**:334*T*
human interference **3**:334*T*, **3**:335
Mediterranean biome **3**:335*T*, **3**:336–340*T*
pattern variations **3**:335
prerequisites and sources **3**:334–335, **3**:334*T*
steppe/grassland biome **3**:335*T*, **3**:336–340*T*
taxon diversity **3**:334, **3**:334*T*
temperate deciduous forest biome **3**:335*T*, **3**:336–340*T*
tundra biome **3**:335*T*, **3**:336–340*T*
warm-temperature evergreen forest biome **3**:335*T*, **3**:336–340*T*
warm-temperature semidesert biome **3**:335*T*, **3**:336–340*T*
zonation patterns **3**:335–340, **3**:336–340*T*
pattern variations
Arctic desert biome **3**:335*T*
azonal vegetation **3**:336–340*T*
biomes **3**:335, **3**:335*T*
boreal forest biome **3**:335*T*, **3**:336–340*T*
boreal/nemoral orobiomes **3**:336–340*T*
continental semidesert biome **3**:335*T*, **3**:336–340*T*
habitat diversity **3**:344
Mediterranean biome **3**:335*T*, **3**:336–340*T*
Mediterranean orobiomes **3**:336–340*T*
research literature and surveys **3**:341–344*T*
semidesert orobiomes **3**:336–340*T*
steppe/grassland biome **3**:335*T*, **3**:336–340*T*
synanthropic vegetation **3**:336–340*T*
temperate deciduous forest biome **3**:335*T*, **3**:336–340*T*
tundra biome **3**:335*T*, **3**:336–340*T*
vegetation diversity **3**:335–340
vegetation maps **3**:340–344, **3**:345*F*, **3**:346*T*
warm-temperature evergreen forest biome **3**:335*T*, **3**:336–340*T*
warm-temperature semidesert biome **3**:335*T*, **3**:336–340*T*
zonation patterns **3**:335–340, **3**:336–340*T*
research outlook **3**:344–346
vegetation classification
background information **3**:333–334
dominance approach **3**:334
floristic–sociological approach **3**:334
habitat classification **3**:334
extreme events **4:209–223**
abrupt driver changes **4**:219–220
climate change effects **2**:84–85
climate change studies **2**:66, **2**:93–94
defining criteria
intensity **4**:210, **4**:211*F*
rarity **4**:210
severity/magnitude **4**:210–211, **4**:211*F*
ecological implications **4**:220–221
ecological responses
cascading effects **4**:214–215
general discussion **4**:213
indirect consequences **4**:214
irreversible, slow, delayed, or hysteretic recovery **4**:213–214, **4**:215*F*
key structuring species loss **4**:214
microenvironmental changes **4**:214
response abruptness **4**:213, **4**:214*F*
response characteristics **4**:213–214
response magnitude **4**:214
response rarity **4**:213
emerging patterns and challenges
anthropogenic amplification **4**:217–219
compensatory mechanisms **4**:216–217
compound effects **4**:217
ecological memory **4**:217, **4**:218*F*
event frequency **4**:215–216
resistance mechanisms **4**:216–217
sequences of events **4**:217
species richness role **4**:216–217
estuarine ecosystems **3**:302
event types **4**:211–212
global change effects **4**:219–220
large-scale drivers **4**:212–213

ecosystems (*continued*)
organism- and system-based responses **4**:211, **4**:212*F*
plant biodiversity threats **6**:62–63
research background **4**:209, **4**:210*F*
research methodologies **4**:219, **4**:220*T*
triggers **4**:212
floodplain ecosystems
Amazon ecosystems
buritizal **1**:164
igapó forest **1**:163–164
lakes **1**:164
mangrove forests **1**:164
permanent swamps **1**:164
pirizal **1**:164
seasonal flooding **1**:162–163
várzea forest **1**:163
habitat loss and degradation **7**:371*T*
river–floodplain connectivity **7:144–160**
benefits **7**:144, **7**:145–146, **7**:150
changes and conversion impacts **7**:150–151, **7**:151*F*
ecological processes **7**:145
ecosystem services **7**:145–146
fisheries **7**:146–147
flood attenuation **7**:147–150, **7**:148*F*, **7**:149*F*
floodplain agriculture **7**:146
floodplain reconnection and conservation **7**:152–153
flow levels and events **7**:150
future research outlook **7**:156–158
geomorphic processes **7**:145
hydrological connectivity **7**:150
hydrological processes **7**:145
key floodplain processes **7**:144–145
land and water management challenges **7**:151
research summary **7**:156–158
spatial variables **7**:150
forest ecosystems **3:528–536**
biodiversity
community organization **3**:533–534
disturbance regimes **3**:530–533
ecosystem functions **3**:535
food webs **3**:533–534
forest structure and patterns **3**:530–533, **3**:532*F*
general patterns **3**:530, **3**:531*F*
human activity **3**:534
keystone species **3**:535
latitudinal gradients **1**:498*F*, **3**:530, **3**:530*F*
tree taxa **3**:532*T*
biogeochemical cycles
carbon cycle **6**:149–150, **6**:150*T*
nitrogen cycle **6**:152
classification systems **3**:528–530, **3**:529*F*
disturbance events **3**:529
human activity
biodiversity **3**:534
dramatic climatic shifts **3**:535
forest conversion and fragmentation **3**:534, **4**:154, **4**:156, **5**:496, **5**:497*F*
harvesting effects **3**:534, **4**:156
introduced species **3**:534
pollution **3**:534–535
living dead **4**:590–591, **4**:591*F*
precipitation **3**:529*F*
South America **3**:110
temperature **3**:529*F*
freshwater ecosystems **3:560–569**
acidic and mercury deposition **1:1–14**
acidic deposition–mercury concentration linkage **1**:13
aluminum concentrations **1**:9–10
aquatic biota response **1**:9–10, **1**:10*T*, **1**:11*F*
emission occurrences and sources **1**:1–5
geographic distribution **1**:1–5, **1**:3*F*
mercury contamination **1**:10–13, **1**:11*F*, **1**:12*F*
mercury emissions and deposition **1**:2–5, **1**:4*F*
pH changes **1**:9–10, **1**:10*T*
research background **1**:1
schematic diagram **1**:2*F*
surface water acidification **1**:8–9, **1**:10*T*, **1**:144–145
Antarctica **1**:183
anthropogenic impacts **3:570–586**
acidification and pollution **3**:572–574, **3**:577–578
agricultural and industrial chemicals **3**:575–576, **3**:580–581, **3**:581*F*
areal reductions **3**:573–574
eutrophication **3**:575*F*, **3**:576–577, **3**:576*F*
hydrological cycle changes **3**:574–575
lake eutrophication **3**:581–582
organic pollution **3**:575, **3**:575*F*, **3**:580–581, **3**:581*F*
pollution **1**:512
restoration and conservation programs **3**:574
biodiversity **3**:566–569, **3**:568*F*
biological consequences
agricultural and industrial chemicals **3**:580–581, **3**:581*F*
lake eutrophication **3**:581–582
organic pollution **3**:580–581, **3**:581*F*
biological quality
species declines **3**:578–579, **3**:579*T*
species variability **3**:578–579, **3**:578*F*
threatened species **3**:579–580
climate change effects **6**:150–151
conservation status **3**:171–173*T*, **3**:174
crustaceans **2**:397–399
dissolved oxygen **3**:351, **3**:562
ecosystem function measurements **3:64–71**
basic concepts **3**:64
bioenergetic analysis **3**:67
biogeochemical cycles **3**:68–70
biomass accrual **3**:67
birth and death rate analysis **3**:67–68
carbon-14 (^{14}C) measurements **3**:66
carbon-based measurements **3**:65–66
comparison studies **3**:64–65
data interpretation and analysis **3**:66–67
egg ratio analysis **3**:68
fluorescence measurements **3**:66
functional type analysis **3**:68
instar analysis **3**:68
new and regenerated production **3**:69–70
nutrient cycling **3**:68–70
oxygen-based measurements **3**:65
primary production **3**:65
secondary production **3**:67
stable isotope studies **3**:70–71
ecosystem services **3**:96–98
eutrophication
aquatic vascular plants **3**:359–360
lakes and ponds **3**:583–584
macroalgae **3**:358
macrophytes **3**:359–360
microfauna **3**:361–363
fish
acidic deposition–mercury concentration linkage **1**:13
African/Ethiopian region **3**:472–473
aquaculture **1**:190*T*
Australian region **3**:473
biodiversity issues **3**:444
global biodiversity **3**:471*F*, **3**:472
habitats **3**:472
mercury contamination **1**:10–13, **1**:12*F*
migratory patterns **5**:315
Nearctic region **3**:472
Neotropical region **3**:472
North America **5**:565
Oriental region **3**:473
Palearctic region **3**:472
species richness **3**:571–572, **3**:583–584
Succulent Karoo **1**:487*T*
surface water acidification **1**:9–10, **1**:10*T*, **1**:11*F*
threatened species **3**:578–579, **3**:579*T*
geographic distribution **4**:370*T*
habitat diversity **3**:562–565, **3**:564*F*, **3**:565*F*, **3**:570–571
habitat loss and degradation **1**:510–512, **1**:511*F*
interactive regulatory mechanisms **3**:565–566
invertebrates **4:369–378**
altered water regimes **4**:377–378
background information **4**:369–370
biodiversity loss **4**:376–377
decomposition and recycling functions **4**:375
disturbances **4**:375
ecological functions and processes **4**:375
endemism **4**:374–375
exotic species invasions **4**:378
global patterns **4**:373
habitat depletion and loss **4**:376–377
habitat heterogeneity **4**:375
influencing factors **4**:373–374
isolation factors **4**:374–375

life cycles 4:370, 4:372*F*
major groups 4:370, 4:371*T*
major habitats 4:369–370, 4:369*F*
major threats 4:376–377
material transfer and uptake 4:375–376
number of species 4:370–371, 4:371*T*
origins 4:374
pollution 4:377
primary production 4:376
reduced water quality 4:377
research background 4:373–374
river continuum concept (RCC) 4:373*F*
species interactions 4:375
taxonomic groups 4:371–373
zonal representatives 4:374*F*
lateral and vertical boundaries 3:565*F*
light attenuation 3:560
mean composition 3:562*T*
mean water depth 3:561*F*
Near East 5:470–471, 5:470*F*, 5:470*T*, 5:471*F*
nonextremeophilic Archaea 1:224*T*
number of lakes 3:561*F*
nutrient loading gradient 3:568, 3:568*F*
organic matter 3:562, 3:564*F*, 3:565–566
photosynthesis 3:560, 3:564*F*
physical characteristics 3:560–562
primary productivity 3:563*T*, 3:566–569, 3:568*F*
protected area systems 6:494–498, 6:496*F*, 6:497*F*
river–floodplain connectivity 7:151
Southeast Asia 1:510–512, 1:511*F*, 1:512
species richness
freshwaters 3:571–572, 3:571*T*
freshwater sediments 3:571–572, 3:572*T*
lakes and ponds 3:583–584
site characteristics and ranges 3:572*T*
trophic levels 3:562, 3:563*T*
usefulness 3:570
water storage 3:560–562, 3:561*T*
functional roles 5:582
government legislation and regulations 3:718
grassland ecosystems **4:1–7**
biodiversity
animal species diversity 4:3–4
biodiversity change scenarios 4:5–6, 4:6*F*
characteristics 4:3
ecosystem functionality 4:4–5
future outlook 4:5–6, 4:6*F*
plant species diversity 4:3
Relative Yield Total (RYT) 4:4–5, 4:5*F*
soil microbial diversity 4:4
conservation status 3:171–173*T*
determinant factors 4:1–3, 4:2*F*
ecosystem services 3:96–98
functional diversity (FD) measures 3:604–605, 3:605*F*
geographic extent 4:1, 4:2*F*
neotropical seasonally dry forests 5:496
water availability 4:1–3, 4:2*F*
guilds 4:36
high-elevation Andean ecosystems **4:96–110**
background information 4:96
geographic distribution 4:96–97, 4:97*F*
habitat characteristics 4:96–97
number of species 4:102*T*
páramo ecosystem
characteristics 4:101–103, 7:275–276
geographic distribution 4:96–97, 4:97*F*
northern Andean alpine cushion formations 4:103
northern Andean dwarf *Polylepis* forest 4:103
number of species 4:102*T*
physical characteristics and patterns 4:97–98
plant diversity 4:98–99, 4:99*T*
representative species 4:102*F*
subpáramo zone 4:102–103
superpáramo zone 4:103
true páramo zone 4:103
physical characteristics and patterns 4:97–98
plant diversity
endemism 4:99–100, 4:99*T*
origins and diversification 4:100–101
species richness 4:98–99
puna
central Andean dwarf *Polylepis* forest 4:105
cushion bogs 4:105
desert puna 4:104–105
dry puna 4:104
geographic distribution 4:96–97, 4:97*F*
humid puna 4:104
number of species 4:102*T*
physical characteristics and patterns 4:97–98, 4:103–104
plant diversity 4:98–99
representative species 4:104*F*
salar vegetation 4:105
research summary 4:107
southern Andean steppes
altitudinal zonation 4:106–107, 4:106*F*
bogs and seeps 4:107
cushion plants 4:106–107, 4:106*F*
geographic distribution 4:96–97, 4:97*F*
latitudinal gradients 4:106–107
number of species 4:102*T*
physical characteristics and patterns 4:97–98, 4:105–107
plant diversity 4:98–99
representative species 4:105*F*
historical perspective 3:59–60
historical research 5:581–582
human-dominated ecosystems
basic concepts 2:350–351
human-dominated landscapes 2:349*F*, 2:350–351, 2:350*F*
human-dominated seascapes 2:351
human-dominated techno-ecosystem 3:60–63, 3:62*F*
human impacts 2:350
trait-based studies 2:356–357
untouched ecosystems 2:350
human impacts **4:153–161**
alien species invasion
continental ecosystems 4:155
general discussion 4:155
island biogeography 4:155
onshore ocean waters 4:155–156
pathogens 4:156
current impacts
alien species invasion 4:155
general discussion 4:153–154
habitat alteration 4:153–154
driving forces
environmentally unsound technologies 4:159
faulty economic factors 4:159
faulty political arrangements 4:160
faulty social arrangements 4:159–160
IPAT equation 4:158–159
market failure 4:159
overconsumption 4:159
population growth/overpopulation 4:159
ecosystem deterioration measures
epidemiological environment 4:158
faltering food production 4:157–158
fisheries decline 4:158
net primary production (NPP) 4:156–158
observable symptoms 4:158
water quality decline 4:158
habitat alteration
climate change effects 4:155
conversion to farms 4:154
conversion to grazing 4:154
conversion to infrastructure 4:154
dam projects 3:573–574, 4:154
deforestation 4:153–154
fire suppression 4:140, 4:154
forest fragmentation 4:154
general discussion 4:153–154
mining impacts 4:154
ocean toxification 4:155
onshore ocean siltation 4:155
ozone depletion 4:155
recreational activities 4:155
selective harvesting 4:154
wetland drainage 4:154–155
historical perspective 4:153
overexploitation
aquarium fish 4:156
esthetic resources 4:156
food harvests 4:156
general discussion 4:156
natural products 4:156
scientific and educational use 4:156
solution strategies
proximate steps 4:160
ultimate steps 4:160–161
ultimate solutions
consumption control 4:161
population reduction 4:160–161

ecosystems (*continued*)
socioeconomic and political system revisions **4**:161
technologically benign substitutions **4**:161
intertidal ecosystems **4:332–344**
anthropogenic impacts **5**:56–57, **5**:58*F*, **5**:59*F*
biodiversity measures **4**:339–340
boulder fields **4**:336–337, **4**:336*F*
ecological function **4**:340–342
general discussion **4**:332
habitat–niche relationship **4**:45–46
mangrove forests **4**:337, **4**:338*F*, **4**:339*F*
marine sedimentary habitats **5**:107, **5**:108*T*
pattern variations
competition **4**:334, **4**:335*F*
general discussion **4**:333–334
interaction chains **4**:334
physical disturbances **4**:335–336
predation **4**:334–335
recruitment **4**:333–334
processes and characteristics **4**:332–333
research summary **4**:342–343
sandy beaches and mudflats **4**:337–339
tides and waves **4**:332–333
values and services **4**:342, **4**:342
lake and pond ecosystems
acidic and mercury deposition **1**:11*F*
freshwater invertebrates **4:369–378**
altered water regimes **4**:377–378
background information **4**:369–370
biodiversity loss **4**:376–377
decomposition and recycling functions **4**:375
disturbances **4**:375
ecological functions and processes **4**:375
endemism **4**:374–375
exotic species invasions **4**:378
global patterns **4**:373
habitat depletion and loss **4**:376–377
habitat heterogeneity **4**:375
influencing factors **4**:373–374
isolation factors **4**:374–375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats **4**:376–377
material transfer and uptake **4**:375–376
number of species **4**:370–371, **4**:371*T*
origins **4**:374
pollution **4**:377
primary production **4**:376
reduced water quality **4**:377
research background **4**:373–374
river continuum concept (RCC) **4**:373*F*
species interactions **4**:375
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
geographic distribution **4**:370*T*
habitat loss and degradation **7**:371*T*
landscape legacies **4**:527–528
living dead **4:590–598**
animal populations **4**:594–595
anthropogenic impacts **4**:594–595
deforestation **4**:591–593, **4**:592*F*
extinct megafauna **4**:591–593, **4**:593*F*
food chains **4**:595–596
forest ecosystems **4**:590–591, **4**:591*F*
requirements **4**:593–594
restoration programs **4**:597
small plants **4**:594
wildland patches **4**:596–597, **4**:596*F*
mangrove ecosystems **5:10–22**
Amazon ecosystems **1**:164
aquaculture impacts **1**:193–194
C_4 plants **2**:371*T*
characteristics **7**:271–272
conservation practices **6**:77
definitions and terminology **7**:369*T*
diversity
general discussion **5**:18
genetic diversity **5**:22
global patterns **5**:18
local variations **5**:20–21
regional patterns **5**:20, **5**:21*F*
ecological interconnections **5**:18
economic exploitation and use **5**:22
global extent **7**:370*T*
global patterns
Atlantic-Caribbean-East Pacific (ACEP) region **5**:18–19, **5**:19*F*, **5**:20*T*
faunal diversity **5**:19–20, **5**:20*T*
Indo-West Pacific (IWP) region **5**:18–19, **5**:19*F*, **5**:20*T*
latitudinal range **5**:18, **5**:19*F*
longitudinal differences **5**:18–19, **5**:19*F*
habitat destruction **1**:513, **1**:513*T*, **3**:404
intertidal ecosystems **4**:337, **4**:338*F*, **4**:339*F*
local variations
animal distribution **5**:21–22
tree distribution **5**:20–21
mangrove trees
characteristics and geographical distribution **5**:10–11
families and genera **5**:11*T*
mangrove habitat **5**:10–11, **5**:10*F*
pneumatophores **5**:12*F*
reproduction **5**:11–12, **5**:12*F*
root system **5**:12*F*
salinity/salt tolerance **5**:11, **5**:11*T*
waterlogged soils **5**:11, **5**:12*F*
marine fauna
bivalves **5**:17
characteristics **5**:14–15
crustaceans **5**:15–16, **5**:20*T*
ecosystem engineers **5**:16–17
fish **5**:15, **5**:15*F*, **5**:20*T*
grapsid crabs **5**:15–16, **5**:15*F*
hermit crabs **5**:16
meiofauna **5**:18
molluscs **5**:16–17, **5**:20*T*
mud crabs **5**:16
ocypodid crabs **5**:16, **5**:16*F*
root communities **5**:14–15
shrimp **5**:16
snails **5**:17–18, **5**:17*F*
marine sedimentary habitats **5**:107, **5**:108*T*
nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
protected area systems **6**:492*T*, **6**:499*T*
reproductive adaptations **7**:397
restoration programs **7**:375–378
shrimp farming versus mangrove conservation **3**:51–53, **3**:52*F*
South America **3**:111*F*, **3**:116
Southeast Asia **1**:513, **1**:513*T*
terrestrial fauna
amphibians **5**:13
birds **5**:13–14, **5**:20*T*
characteristics **5**:12–13
insects **5**:13, **5**:13*F*
mammals **5**:14, **5**:20*T*
reptiles **5**:13, **5**:20*T*
threatened species **7**:201*T*
marine ecosystems **5:45–55**
anthropogenic impacts **5:56–63**
artificial reefs **5**:61
ecosystem engineer species **5**:60–61
eutrophication **5**:53–54
invasive species **5**:54, **5**:60–61
major threats **5**:56
mariculture **5**:61
overfishing **3**:410, **5**:54, **5**:61
pollution **5**:61, **6**:150–151
research summary **5**:62
rocky intertidal communities **5**:56–57, **5**:58*F*, **5**:59*F*
rocky subtidal communities **5**:57–58, **5**:60*F*
sea otter populations **5**:58–60
biodiversity
genetic diversity **5**:50
higher taxa **5**:47–49, **5**:50*T*
species **5**:49–50, **5**:50*T*
biodiversity loss **4**:664
biodiversity-rich countries **1**:498
chlorophyll production **5**:46
climate change effects **6**:150–151
conservation **5:32–44**
abundance impacts **5**:34–35
assisted colonization **5**:41
climate change effects **3**:449, **5**:32
community impacts **5**:37–38
conservation strategies **5**:38–39
distribution impacts **5**:35–36
ecosystem-based management **5**:40
evolutionary impacts **5**:36
fisheries management **5**:40–41
future research areas **5**:41
marine protected areas **5**:39–40
ocean acidification **5**:33–34
predictive models **5**:39
research needs **5**:41
restoration **5**:41
rising seawater temperatures **5**:33
conservation status **3**:171–173*T*, **3**:174
continental shelves **5**:52
Convention on Biological Diversity (1930) **2**:318

countryside biogeography studies **2**:356
crustaceans **2**:397–399
dinitrogen-fixing (N_2-fixing)
  microorganisms **2:608–625**
  basic concepts **2**:608–610, **2**:609*F*
  biodiversity **2**:610–611
  biogeochemical cycles **1**:560–561
  changing environmental conditions **2**:621–623
  dinitrogenase/dinitrogenase reductase **2**:610–611
  emerging patterns **2**:619–621
  functional role **2**:615–618, **2**:616*F*
  habitat characteristics **2**:609*F*
  metabolic diversity **2**:614–615
  nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
  nutrient cycle linkages **2**:618–619
  phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
  phylogenetic relationships **2**:611*F*
  physiological diversity **2**:614–615
  Redfield ratio **2**:617–618, **2**:617*F*
  research summary **2**:623
  spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
ecosystem function measurements **3:64–71**
  basic concepts **3**:64
  bioenergetic analysis **3**:67
  biogeochemical cycles **3**:68–70
  biomass accrual **3**:67
  birth and death rate analysis **3**:67–68
  carbon-14 (^{14}C) measurements **3**:66
  carbon-based measurements **3**:65–66
  comparison studies **3**:64–65
  data interpretation and analysis **3**:66–67
  egg ratio analysis **3**:68
  fluorescence measurements **3**:66
  functional type analysis **3**:68
  instar analysis **3**:68
  new and regenerated production **3**:69–70
  nutrient cycling **3**:68–70
  oxygen-based measurements **3**:65
  primary production **3**:65
  secondary production **3**:67
  stable isotope studies **3**:70–71
ecosystem services **5:451–468**
  accounting assessment models **5**:455–457
  American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries **5**:464–465, **5**:464*F*
  bioeconomic models **5**:457–459
  challenges **5**:455
  coastal zone management (Belize) **5**:465, **5**:465*F*
  conceptual framework **5**:452–453
  coral reef protection and restoration **5**:465–466
  cost–benefit analyses **5**:459–463
  coupled social–ecological systems model **5**:453–455, **5**:454*F*
  cross-system services **3**:96–98
  cultural ecosystem services **5**:451–452, **5**:452*T*
  decision-support tools **5**:459–463, **5**:460–462*T*, **5**:463*F*
  future outlook **5**:466
  integrated ecosystem assessments (IEAs) **5**:454–455
  modeling approaches and uses **5**:453
  multiple ecosystem services **5**:459–463
  production function models **5**:457–459
  provisioning services **5**:451–452, **5**:452*T*
  regulating services **5**:451–452, **5**:452*T*
  revealed references methods **5**:456–457
  single ecosystem services **5**:457–459
  stated preference methods **5**:457
  supporting services **5**:451–452, **5**:452*T*
  tradeoffs **5**:459–463
  valuation measures **5**:457
  wind farms **5**:466
ecosystem unit classifications **5**:45–46, **5**:46*F*
eutrophication **5:23–31**
  anthropogenic effects **5**:23–24, **5**:53–54
  aquatic vascular plants **3**:360–361
  contaminant–stress relationship **5**:29–30
  food web complexity **5**:24–25
  harmful algal blooms **5**:25
  high nutrient, low chlorophyll (HNLC) concept **5**:26, **5**:26*T*
  high nutrient, low growth (HNLG) concept **5**:26–27, **5**:26*T*, **5**:27*F*
  lake and pond clean-up effects **5**:24
  macroalgae **3**:358–359
  macrophytes **3**:360–361, **3**:360*F*, **3**:361*F*
  microfauna **3**:363–364, **3**:363*F*
  nearshore ocean nutrient enrichment response **5**:25, **5**:26*F*
  nutrient loading **5**:23–24
  nutrient ratio changes **5**:28–29, **5**:28*T*, **5**:29*F*
  Redfield ratios **5**:27
  silicon additions **5**:27–28
  species shifts **3**:353*F*
extinction risks **4**:664
food chains **5**:50–51
functional roles **5**:50–51
global change effects **4**:664, **5**:257–258
international organizations **4**:329–330
mangrove ecosystems **5**:45–46
  bivalves **5**:17
  characteristics **5**:14–15
  crustaceans **5**:15–16, **5**:20*T*
  ecosystem engineers **5**:16–17
  fish **5**:15, **5**:15*F*, **5**:20*T*
  grapsid crabs **5**:15–16, **5**:15*F*
  hermit crabs **5**:16
  meiofauna **5**:18
  molluscs **5**:16–17, **5**:20*T*
  mud crabs **5**:16
  ocypodid crabs **5**:16, **5**:16*F*
  root communities **5**:14–15
  shrimp **5**:16
  snails **5**:17–18, **5**:17*F*
marine–land comparisons **5**:46–47, **5**:47*F*
marine microbiology **5:292–298**
  Global Ocean Sampling (GOS) program **5**:293
  historical perspective **5**:292–293
marine protected areas **5:94–104**
  benefits **5**:96–97
  Bowie Seamount Marine Protected Area (Sglaan Kinghlas) **5**:446–447
  characteristics and functional role **5**:94–95
  climate change threats **5**:100–101
  conservation strategies **5**:39–40
  ecosystem-based management **5**:95
  effectiveness **5**:100
  Endeavor Hydrothermal Vents Marine Protected Area **5**:447
  fish conservation **3**:453–454
  fish stocks **3**:486
  geographic extent **5**:445–446, **5**:446*T*
  global commitment **5**:99–100
  Great Barrier Reef Marine Park **5**:446
  Papahanaumokuakea Marine National Monument **5**:446
  socioeconomic benefits **5**:98–99, **5**:98*T*
  threats and challenges **5**:447–448
marine sediments **5:105–115**
  benthic organisms and environment **5**:105–107, **5**:106*F*
  diversity regulation **5**:109–110
  ecosystem services **5**:111, **5**:111*F*
  global diversity patterns **5**:109
  habitats **5**:107, **5**:107*F*
  macrofauna **5**:106*F*
  meiofauna **5**:106*F*
  species numbers **5**:108–109
  threats and challenges **5**:113–114
nematodes **7**:429, **7**:431
net primary production (NPP) **1**:559, **1**:560*T*
nonanthropogenic environmental changes and variability **5**:61–62
nonextremeophilic Archaea **1**:224*T*
ocean edge
  coral reefs **5**:52
  intertidal beaches **5**:51
  kelp beds **5**:51
  rocky shores **5**:51–52
open ocean/deep sea
  benthic ecosystems **5**:53
  hydrothermal vents **3**:242, **5**:53
  Mid-Ocean Ridge **5**:53
  pelagic ecosystems **5**:49*F*, **5**:52–53
pelagic ecosystems **5**:46, **5**:46*F*, **5**:47*F*
planktonic organisms **5**:48*F*
primary production **5**:46
protected area systems **6**:494, **6**:495*F*

ecosystems (*continued*)
research summary **5**:115
survival probabilities **5**:47
symbiotic relationships **5:116–126**
bioluminescence **5**:121–123, **5**:122*F*
fish **5**:123
functional roles **5**:117–118
future challenges **5**:125
global distribution **5**:117*F*
historical research **5**:116–117
light-producing symbioses **5**:121–123, **5**:122*F*
nutrient-exchange mutualisms **5**:118–120
protective compound production **5**:124–125
sponges **5**:124–125, **5**:124*F*
squid **5**:123–124
structural characteristics **5**:116–117, **5**:117*F*
transmission modes **5**:116–117, **5**:117*F*
threatened birds **7**:181
threats and challenges
biodiversity loss **5**:113–114
eutrophication **5**:113*F*, **5**:114
exotic species invasions **5**:113*F*, **5**:114
fishing impacts **5**:113–114, **5**:113*F*
global climate change **5**:113*F*, **5**:114–115
habitat destruction and degradation **5**:113*F*, **5**:114
pollution **5**:113*F*, **5**:114
shoreline modifications **5**:113*F*, **5**:114
trophic levels **5**:46–47, **5**:48*F*, **5**:49*F*, **5**:50–51
Mediterranean-type climate ecosystems **5:212–222**
Australia
physical characteristics **5**:220–221
vascular plant diversity **5**:213*T*, **5**:221
vertebrate diversity **5**:213*T*, **5**:221
California
natural disturbances **5**:213–214
physical characteristics **5**:213*T*, **5**:214–215
speciation patterns **5**:214
species richness **5**:213*T*
vascular plant diversity **5**:215
vertebrate diversity **5**:213*T*, **5**:215
Cape Region (South Africa)
physical characteristics **5**:219
vascular plant diversity **5**:213*T*, **5**:219–220
vertebrate diversity **5**:213*T*, **5**:220
characteristics **5**:212–213
Chile
natural disturbances **5**:213–214
physical characteristics **5**:213*T*, **5**:215–216
speciation patterns **5**:214
species richness **5**:213*T*
vascular plant diversity **5**:216–217
vertebrate diversity **5**:213*T*, **5**:217
global distribution **5**:213*F*, **5**:213*T*
Mediterranean Basin
physical characteristics **5**:217–218
vascular plant diversity **5**:218
vertebrate diversity **5**:213*T*, **5**:218–219
natural disturbances **5**:213–214
physical characteristics **5**:213*T*
speciation patterns **5**:214
species richness **5**:213*T*
summary discussion **5**:221–222
Near East **5:478–487**
animal diversity **5:459–477**
agricultural practices **5**:475–477
biodiversity extent **5**:459–460, **5**:460*T*
dry conditions **5**:466–467
endemic versus widespread animals **5**:463–464, **5**:464*F*, **5**:464*T*
freshwater fish diversity **5**:470–471, **5**:470*F*, **5**:470*T*, **5**:471*F*
habitat destruction **5**:472, **5**:473*F*
historic zoogeography **5**:460–463, **5**:461*F*, **5**:462*F*, **5**:463*T*
human impact **5**:471–472
hunting impacts **5**:472
influencing factors **5**:464–467
Oriental representatives **5**:462–463, **5**:463*T*
poisoning **5**:473–475
rainfall factors **5**:465–467, **5**:465*F*, **5**:466*T*
sand adaptability **5**:467–469, **5**:468*F*
substratum **5**:467–469, **5**:467*F*, **5**:468*F*
vegetation **5**:469–470, **5**:469*F*
climate
dew and fog **5**:479–480
rainfall **5**:465–467, **5**:465*F*, **5**:466*T*, **5**:479, **5**:480*F*
temperature regimes **5**:480
edaphic conditions **5**:478–479
flora **5**:480–481, **5**:481*T*
geographical subdivisions **5**:478, **5**:479*F*
geomorphology **5**:478–479
Levant region **5**:459, **5**:459*F*, **5**:460*T*
rock types **5**:478–479
topography **5**:478
vegetation
Ceratonia siliqua open forests **5**:480*F*, **5**:482–483
deserts **5**:480*F*, **5**:485–486
desert savannoid vegetation **5**:480*F*, **5**:486
dryland vegetation patterns **5**:484
general discussion **5**:481–482
mangroves **5**:480*F*, **5**:487
maquis/forests **5**:480*F*, **5**:481–482
Mediterranean savannoid vegetation **5**:480*F*, **5**:483
Mt. Hermon montane forest **5**:480*F*, **5**:482
Mt. Hermon tragacanth vegetation **5**:480*F*, **5**:483–484
Pistacia lentiscus open forests **5**:480*F*, **5**:482–483
Quercus calliprinos woodlands **5**:480*F*, **5**:482
Quercus ithaburensis open forests **5**:480*F*, **5**:482
reed thickets **5**:480*F*, **5**:486
sand vegetation **5**:480*F*, **5**:486
semisteppe bathas **5**:480*F*, **5**:483
shrub–steppes **5**:480*F*, **5**:484–485
shrub–steppes with trees **5**:480*F*, **5**:485
Sudanian tree oases **5**:480*F*, **5**:486
swamps **5**:480*F*, **5**:486
synanthropic vegetation **5**:480*F*, **5**:486–487
wet salinas **5**:480*F*, **5**:486
Ziziphus lotus **5**:480*F*, **5**:483
New Zealand **1**:149*T*
North America **1**:149*T*
ocean ecosystems
characteristics
basic organization **5**:587*F*
distinguishing characteristics **5**:587–588
shared characteristics **5**:586–587
global change impacts **5**:588–589
historical research **5**:582
natural functioning units
biome concept **5**:583
challenges **5**:582–583
coastal upwelling **5**:584*T*, **5**:585*T*, **5**:586, **5**:586*T*
Eastern Boundary Current **5**:584*T*, **5**:585*T*, **5**:586, **5**:586*T*
ecosystem concept **5**:583–586, **5**:584*T*, **5**:585*T*, **5**:586*T*
equatorial upwelling **5**:583–586, **5**:584*T*, **5**:585*T*, **5**:586*T*
low-latitutde gyre **5**:583, **5**:584*T*, **5**:585*T*, **5**:586*T*
Southern Ocean **5**:583, **5**:584*T*, **5**:585*T*, **5**:586*T*
subarctic gyre **5**:584*T*, **5**:585*T*, **5**:586, **5**:586*T*
Oceania **1**:149*T*
paleoecology **5:645–655**
data collection and interpretation **5**:645–646
goals and objectives **5**:645–646
Quaternary data **5**:645
research methodologies
age models **5**:652–653
dating uncertainties **5**:652
mapping intervals **5**:652–653
scaling factors **5**:646–649, **5**:648*F*, **5**:649*F*
space–time variations **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
spatial characteristics **5**:648*F*, **5**:649–650, **5**:649*F*, **5**:651*F*
taxonomic and numerical characteristics **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
temporal characteristics **5**:650–652, **5**:651*F*
temporal resolution **5**:652
zoom lens perspective **5**:647*F*, **5**:648*F*, **5**:649*F*, **5**:651*F*, **5**:653–654
pelagic ecosystems **5:683–691**
benthic biomass distribution **4**:380–381, **4**:380*F*

El Niño Southern Oscillation (ENSO) **5**:687–688, **5**:689*F*, **5**:690*F*
future outlook **5**:690–691
pelagic diversity patterns **4**:385, **4**:386*T*
pelagic fish stocks **5**:688–690, **5**:690*F*
physical characteristics **4**:379, **4**:380*T*
phytoplankton **5**:683–686, **5**:684*F*, **5**:686*F*
primary production **5**:683–686, **5**:685*F*, **5**:686*F*
Russell cycle **5**:690–691
zooplankton **5**:686–688, **5**:687*F*, **5**:688*F*, **5**:689*F*
people–ecosystem relationship
autonomous people **6**:221–222, **6**:221*T*
biosphere people **6**:221*T*
conservation practices **6**:226
ecodevelopment programs **6**:226–227
ecological refugees **6**:221*T*, **6**:225–226
future challenges **6**:227
general characteristics **6**:220–221
land-use changes **6**:223–224, **6**:224*T*
resource ownership, sharing, and utilization **6**:223, **6**:224*T*
subjugated people **6**:221*T*, **6**:222–223
value appropriation **6**:224–225
provisioning services
basic concepts **2**:490*T*
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
regulating services
basic concepts **2**:490*T*
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
research summary **3**:174
schematic diagram **3**:60*F*, **3**:61*F*
South America **3:105–117**
alpine life zone **1**:149*T*
background information **3**:105
biogeography
early history **3**:109
late Tertiary **3**:109
Quaternary period **3**:109–110
geographic distribution **3**:111*F*
high Andean shrublands **3**:111*F*, **3**:114, **3**:115
mangrove ecosystems **3**:111*F*, **3**:116
natural ecosystems **6:599–611**
anthropogenic impacts **6**:608
aquatic ecosystem threats **6**:609
charcoal manufacture **6**:609
climate change **6**:609
conservation priorities **6**:610
converted areas **6**:608
dams **6**:609
deforestation **6**:608–609
extractivism **6**:608
hunting effects **6**:609
land cover classifications **6**:607*T*
logging operations **6**:608, **6**:609
protected area status **6**:610
sedimentation **6**:610
terrestrial ecoregion classifications **6**:599, **6**:600–604*T*, **6**:605*F*
terrestrial ecosystem threats **6**:608–609
threats **6**:608–609
tropical forests **6**:599–608, **6**:606*F*, **6**:606*T*
waterways **6**:609–610
physical characteristics
climate **3**:107–109, **3**:108*F*
geography **3**:105–107
geology and geomorphology **3**:105–107, **3**:106*F*
tectonic history **3**:107
Southern Hemisphere **6:612–619**
angiosperms **6**:617–618
background information **6**:612
biotic dispersal **6**:616
climate change effects **6**:617
climatic patterns **6**:614
continental drift **6**:616
dispersal ability **6**:613–614
fossil record **6**:615–616
general area cladograms **6**:614
geographic distribution patterns **6**:613–614
high southern latitude plant growth **6**:618
historical research **6**:612–613
long-distance dispersal **6**:616–617
relaxed clock molecular dating analyses **6**:614–615
research summary **6**:618–619
vicariance biogeography **6**:616
temperate forests **3**:111*F*, **3**:117
temperate grasslands
distribution and structure **3**:111*F*, **3**:113–114
functional aspects **3**:114–115
tropical forests
distribution and structure **3**:110
functional aspects **3**:112
general patterns **3**:111–112
mountain cloud forests **3**:111
tropical rain forests **3**:110, **3**:111*F*, **3**:112
tropical semideciduous/dry forests **3**:110–111, **3**:112–113
tropical savannas
distribution and structure **3**:111*F*, **3**:113
functional aspects **3**:114
wetlands **3**:116–117
xeric formations
distribution and structure **3**:111*F*, **3**:115–116
functional aspects **3**:116
Southern Hemisphere **6:612–619**
angiosperms **6**:617–618
background information **6**:612
biotic dispersal **6**:616
climate change effects **6**:617
continental drift **6**:616
geographic distribution patterns
climatic patterns **6**:614
dispersal ability **6**:613–614
fossil record **6**:615–616
general area cladograms **6**:614
relaxed clock molecular dating analyses **6**:614–615
research background **6**:613–614
high southern latitude plant growth **6**:618
historical research **6**:612–613
long-distance dispersal **6**:616–617
research summary **6**:618–619
vicariance biogeography **6**:616
subterranean ecosystems **7:49–62**
adaptive behaviors
characteristics **7**:53–55, **7**:54*T*
evolutionary rates **7**:54*T*
genetic expression **7**:53–55, **7**:54*F*
anchialine communities **7**:58, **7**:59*T*
aquatic subterranean habitats **7**:49–50, **7**:50*F*
Ariège (France) **7**:60
bat colonies **7**:60
biogeographical patterns **7**:60–62, **7**:60*T*, **7**:61*F*
Bracken Cave (Texas, USA) **7**:60
characteristics **7**:49–50
chemoautotrophic communities **7**:58–59, **7**:59*T*
colonization and speciation **7**:55–57, **7**:56*F*, **7**:57*F*
conservation and protection strategies **7**:62
Cueva de Felipe Reventón (Canary Islands) **7**:59*T*
Cueva del Sobrado (Canary Islands) **7**:59*T*
Cueva del Viento (Canary Islands) **7**:59*T*
Edwards Aquifer (Texas, USA) **7**:58, **7**:59*T*
energy sources **7**:51–52
epikarst communities **7**:56*F*, **7**:57–58, **7**:59*T*, **7**:60
groundwater calcretes (Western Australia) **7**:59
Gua Salukkan Kallang-Towakkalak (Indonesia) **7**:59–60
hydropetric communities **7**:58, **7**:59*T*
hypotelminorheic communities **7**:59
interstitial communities
Lobau wetlands (Austria) **7**:59
Rhone River (France) **7**:59
Kazumura Cave (Hawaii, USA) **7**:60
Krizna jama (Slovenia) **7**:59*T*
lava tube communities **7**:60
Lobau wetlands (Austria) **7**:59
Logarcek (Slovenia) **7**:59*T*
Mammoth Cave (Kentucky, USA) **7**:59*T*, **7**:60
milieu souterrain superficiel (MSS) **7**:51, **7**:60
Movile Cave (Romania) **7**:58–59, **7**:59*T*
myriapods **5**:431
Organ Cave (West Virginia, USA) **7**:58, **7**:60, **7**:60*T*
Postojna Planina Cave System (cave stream communities) **7**:58, **7**:59*T*
Postojna Planina Cave System (epikarst communities) **7**:56*F*, **7**:57–58, **7**:59*T*
Rhone River (France) **7**:59
Robe River (Australia) **7**:59*T*

ecosystems (*continued*)
Rock Creek (Washington, D. C., USA) **7**:59
Sika-Krka System (Slovenia) **7**:59*T*
species richness **7**:60–62, **7**:60*T*, **7**:61*F*
spring communities **7**:58
taxonomic diversity **7**:52–53, **7**:52*F*
terrestrial communities
Gua Salukkan Kallang-Towakkalak (Indonesia) **7**:59–60
Kazumura Cave (Hawaii, USA) **7**:60
Mammoth Cave (Kentucky, USA) **7**:59*T*, **7**:60
Organ Cave (West Virginia, USA) **7**:60
Vjetrenica (Bosnia and Herzegovina) **7**:60
terrestrial subterranean habitats **7**:50–51
Triadou Aquifer (France) **7**:59*T*
Vjetrenica (Bosnia and Herzegovina) **7**:58, **7**:59*T*, **7**:60
Walsingham Caves (Bermuda) **7**:58, **7**:59*T*
succession **5**:582
supporting services
animal pollination **5**:354–355
basic concepts **2**:490*T*, **7**:315–316
desertification **2**:490*T*
marine ecosystems **5**:451–452, **5**:452*T*
modeling frameworks **5**:354–355
nutrient retention **5**:355–356
sustainability trends **7**:82*F*
termites
functional role **4**:429–431, **4**:430*T*
human disturbances **4**:432–433, **4**:432*T*
soils **4**:431–432, **4**:431*F*
terrestrial ecosystems **4**:429
terrestrial ecosystems
Arctic ecosystems **1:227–244**
biodiversity filters **1**:238
biodiversity patterns **1**:231–235, **1**:231*T*
circumpolar diversity **1**:235–236
current biodiversity status **1**:242
environmental characteristics **1**:227–231, **1**:229*F*
environmental consequences **1**:238–240, **1**:239*F*, **1**:240*F*
general discussion **1**:227
latitudinal biodiversity patterns **1**:232–235, **1**:233–234*T*, **1**:235*F*, **1**:236*F*
natural resources/environmental exploitation **1**:240–241
research organizations **1**:242–243
research summary **1**:243
threats and challenges **1**:241–242
topographic impacts **1**:236–238, **1**:237*F*
vegetation zones **1**:229*F*, **1**:230*T*
biological impoverishment **6**:335–337, **6**:336*F*
conservation status **3**:170–174, **3**:171–173*T*
disturbance ecology **2:562–567**
anthropogenic impacts **2**:566–567
basic concepts **2**:562–563
carnivores **2**:565–566
herbivores **2**:565–566
interaction impacts **2**:566
intermediate disturbance hypothesis **2**:563
mechanical disturbances **2**:563–564
physico–chemical disturbances **2**:564–565, **2**:565*F*
ecosystem function measurements **3:72–89**
analytical and measurement approaches **3**:72–73, **3**:74–75*T*, **3**:76*F*
biodiversity **3**:84–86, **3**:84*F*
field manipulations **3**:86*F*, **3**:87*F*, **3**:88
functional diversity (FD) **3**:84
monitoring networks **3**:86*F*, **3**:87*F*, **3**:88
natural versus synthesized ecosystems **3**:88
plant functional traits **3**:85*F*, **3**:85*T*, **3**:86
positive-feedback switches **3**:86–87, **3**:86*F*
purpose **3**:72
real ecosystems **3**:88
resistance and resilience strategies **3**:83–84, **3**:83*F*
resource dynamics **3**:72–73, **3**:74–75*T*, **3**:75*F*
synthesized ecosystems **3**:86*F*, **3**:87*F*, **3**:88
taxonomic diversity **3**:84
ecosystem services **5:347–361**
biophysical measures **5**:349
cost-effectiveness measures **5**:356–357, **5**:357*F*
decision-making strategies **5**:358
landowner decisions **5**:347–349
land use/land cover (LULC) patterns **5**:348–349, **5**:351–352
modeling approaches **5**:349–350
summary discussion **5**:358–359
supply–demand–value relationship **5**:348–349
uncertainty monitoring **5**:356
global change effects **5**:257–258
net primary production (NPP) **1**:559, **1**:560*T*
nitrogen deposition **5:519–536**
acidification **5**:522, **5**:527*F*
anthropogenic impacts **5**:520–522, **5**:522*F*
biodiversity impacts **5**:522
biological nitrogen fixation **5**:520*F*
direct foliar toxicity **5**:523–525, **5**:527*F*
ecosystem-specific effects **5**:526–528
environmental impact **5**:519–520
eutrophication **5**:522
global distribution **5**:524*F*
global nitrogen budget **5**:523*F*
global nitrogen cycle **5**:520–522, **5**:521*F*
hotspots **5**:526*T*
influencing factors **5**:526–528
nitrogen as nutrient **5**:520, **5**:520*F*
primary drivers **5**:522, **5**:525*F*, **5**:527*F*
research methodologies **5**:525–526, **5**:528*T*, **5**:529*F*
secondary stressors **5**:522–523, **5**:527*F*
taxa-specific responses **5**:528
nonextremeophilic Archaea **1**:224*T*
protected area systems
biogeographic realms **6**:491
biome realms **6**:491–492, **6**:492*T*, **6**:493*F*
data challenges **6**:491
deserts **6**:492, **6**:495*T*
ecoregion protection **2**:591–592, **6**:492, **6**:493*F*
forests **6**:492, **6**:494*F*, **6**:494*T*
islands **6**:493–494
mountains **6**:492–493
World Wildlife Fund (WWF) realms **6**:492*T*
termites **4**:429
trophic cascades **7**:259–260
thermal ecosystems **4:111–120**
biodiversity **4**:118–119
characteristics and definitions **4**:111–112
extreme environments **4**:111
geothermally heated systems **4**:112–113
research summary **4**:119–120
temperature tolerance ranges **4**:113–115
thermophilic evolution **4**:115–118
tropical forests **7:269–276**
biodiversity-rich countries **1**:497–498
climatic characteristics **7**:270–271
general discussion **7**:269
geographic extent **7**:269–270, **7**:270*T*
high mountain ecosystems
characteristics **7**:275
humid forests **7**:275
páramo ecosystem **4**:101–103, **7**:275–276
perhumid forests **7**:275
subhumid forests **7**:275
lowland ecosystems
characteristics **7**:271–272
flooded forests **7**:272
humid forests **7**:273
mangrove ecosystems **7**:271–272
perhumid forests **7**:272–273
savannas **7**:274
subhumid forests **7**:273–274
low mountain ecosystems
characteristics **7**:274–275
humid forests **7**:275
perhumid forests **7**:275
subhumid forests **7**:274–275
major forest ecosystems
classifications **7**:271–272
high mountains **7**:275
lowland ecosystems **7**:271–272
low mountains **7**:274–275
physiographic characteristics **7**:270
soils **7**:271
tundra ecosystems
biodiversity change scenarios **4**:6*F*

boreal forests **1**:262–263*T*, **1**:274–275
Canadian Shield **5**:562
characteristics **6**:330–331, **6**:331*F*
climate change studies **2**:61*F*
current global land cover **4**:557*F*
ecosystem function measurements **3**:77*F*
Europe **3**:335*T*, **3**:336–340*T*
forest–tundra ecotone **1**:630
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
net primary production (NPP) **1**:560*T*
nitrogen deposition **5**:534*F*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
soil characteristics **6**:123*T*
wetlands 7:367–383, 7:384–402
animal species richness **7**:389–391, **7**:389*F*, **7**:390*F*, **7**:393–394, **7**:396*F*
anthropogenic impacts **7**:371*T*
biogeographical variation **7**:387–388
C_4 plants **2**:371*T*
classifications **7**:385–386
continental and regional diversity **7**:387–388
creation and restoration
appropriate hydrologic conditions **7**:378
Bois-des-Bel Peatland, Quebec, Canada (case study 5) **7**:375
coastal wetlands **7**:375–378
conservation programs **7**:371
croplands **7**:371
definitions and terminology **7**:368–371
Delaware Bay salt marsh restoration (case study 6) **7**:376–378, **7**:377*F*
designer wetlands **7**:379
ecological engineering principles **7**:381–382
Florida Everglades (case study 1) **7**:372, **7**:372*F*, **7**:373*F*
forested wetland conservation **7**:371
Gulf of Mexico hypoxia (case study 2) **7**:372–373, **7**:374*F*
habitat loss mitigation **7**:370–371
hydrologic and water quality restoration **7**:371–372
Louisiana delta restoration (case study 7) **7**:378
Mesopotamian marshlands (case study 4) **7**:374
natural successional processes **7**:379
peatlands **7**:374–375
planted versus nonplanted restoration (case study 8) **7**:379, **7**:380*F*
research background **7**:368–371
river delta restoration **7**:378
self-designed wetlands **7**:379
Skjern River Restoration Project, Denmark (case study 3) **7**:373–374, **7**:375*F*
soil properties **7**:378–379
success measures **7**:379–381
dam projects **1**:512
definitions and terminology **7**:367–368, **7**:369*T*
dynamic global vegetation models (DGVMs) **2**:678–679
ecosystem services **1**:334*T*
freshwater invertebrates 4:369–378
altered water regimes **4**:377–378
background information **4**:369–370
biodiversity loss **4**:376–377
decomposition and recycling functions **4**:375
disturbances **4**:375
ecological functions and processes **4**:375
endemism **4**:374–375
exotic species invasions **4**:378
global patterns **4**:373
habitat depletion and loss **4**:376–377
habitat heterogeneity **4**:375
influencing factors **4**:373–374
isolation factors **4**:374–375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats **4**:376–377
material transfer and uptake **4**:375–376
number of species **4**:370–371, **4**:371*T*
origins **4**:374
pollution **4**:377
primary production **4**:376
reduced water quality **4**:377
research background **4**:373–374
river continuum concept (RCC) **4**:373*F*
species interactions **4**:375
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
functional diversity (FD)
acidic environments **7**:395–397, **7**:396*F*
adaptations **7**:394–395
anoxic environments **7**:394–395
carnivorous plants **7**:395–397, **7**:396*F*
drought conditions **7**:395
nutrient availability **7**:395–397
reproductive adaptations **7**:397
salinity/salt tolerance **7**:395
functional role **7**:367, **7**:370*F*
geographic distribution **4**:370*T*
global extent **7**:368, **7**:370*T*
habitat destruction **3**:404, **7**:398
habitat loss and degradation **7**:371*T*
human impacts **4**:154–155
Lobau wetlands (Austria) **7**:59
major regional wetland ecosystems **7**:386–387
major threats **7**:368, **7**:371*T*
anthropogenic impacts **7**:399
global losses **7**:397–398
habitat destruction **3**:404, **7**:398
hydrological alterations **7**:398–399
invasive species **7**:399
landscape fragmentation **7**:399
natural disturbances **7**:399–400
nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
occurrences and characteristics **7**:384–386, **7**:385*F*
plant species richness **7**:388–389, **7**:388*F*, **7**:391–393, **7**:398*F*
restoration and maintenance programs **7**:400–401
schematic diagram **7**:370*F*
South America **3**:116–117
species diversity controls
animal species richness **7**:393–394, **7**:396*F*
influencing factors **7**:391–393, **7**:392*F*
plant species richness **7**:391–393
summary discussion **7**:401
wildlife management **7**:405–406
see also lake and pond ecosystems; rivers
Ecosystem Services Review (ESR) software **5**:358
ecotones 3:142–148
agrobiodiversity **1**:131–132
basic concepts **3**:142, **3**:143*F*, **4**:42–45
biodiversity patterns **3**:144
evolutionary development **3**:144–145
forest–tundra ecotone **1**:630
habitat–niche relationship **4**:42–45
historical research **3**:142–143
measurement approaches **3**:143–144
research summary **3**:147
speciation
conservation implications **3**:146
evolutionary development **3**:144–145
human-generated ecotones **3**:143*F*, **3**:146–147
species range/transitional areas **3**:145–146
ecotourism **1**:395–396, **2**:295–296, **2**:296*F*, **4**:155, **6**:500, **7**:224
ecotoxicology 3:118–126
basic concepts **3**:118–119
biomonitoring processes
bioassays **3**:122*F*
bioindicators **3**:124–125
biological processes **3**:122–123
biomarkers **3**:123*T*
bioprobes **3**:122*F*
buffering capacity determination **3**:123–124
characteristics **3**:121–122
keystone species **3**:125
perturbation measurements **3**:122*F*
receptor/reporter gene assays **3**:121–122, **3**:122*F*
soils **3**:122–123, **3**:122*T*
specificity and resolution **3**:124–125
indicator species **4**:255
pollutant entry, transport and fate **3**:121
research applications
biochemical studies **3**:119

ecotoxicology (*continued*)
biotechnological studies **3**:120
chemical studies **3**:119
ecological studies **3**:120
microbiological studies **3**:119–120
toxicological studies **3**:120–121
research summary **3**:125
Ecotron large-scale controlled environment facility **4**:585
Ecteinascidia turbinata **5**:711
ectomycorrhizae **3**:636, **3**:636*F*
ectoparasites **5**:657, **5**:668, **5**:671*F*
ectopic recombination **6**:354
Ectopistes migratorius **1**:621, **2**:281–282, **7**:182
Ectoprocta **4**:384*T*, **4**:385*T*
ectosymbionts **5**:116–117, **5**:117*F*
Ecuador
biodiversity richness **1**:501*F*
botanical gardens **6**:519
butterflies **1**:657*T*, **1**:658*T*
Convention on Biological Diversity (1930) **2**:313
coral reef ecosystems **1**:500*T*
deforestation **1**:502*T*
ecosystem services **7**:356*T*
endangered mammals **3**:194*T*
endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
indigenous peoples **4**:270
institutional collections and capacity **1**:505*T*
land cover classifications **6**:607*T*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
pastureland conversion **4**:665*T*
plant species **4**:102*T*
rain forests **1**:159*T*
species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
threatened birds **7**:181
tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*
tropical rain forests **6**:314–315, **6**:317*T*, **6**:318*T*
vascular epiphytes **3**:520*T*
water funds **7**:355–357, **7**:356*T*, **7**:361–362
edelweiss **1**:154
Edentata **5**:225*T*
edge effects
habitat loss and fragmentation **3**:406
historical range and variation (HRV) **2**:577*F*
human-generated ecotones **3**:146–147
landscape ecology **4**:496–497
species–area relationships (SARs) **6**:641
Edgertonia floccula **6**:292
edible plants **3:127–132**
food habits **3**:127–128, **3**:127*T*
geographic distribution **3**:130–131, **3**:131*T*
species diversity
abundance rankings **3**:129*T*
edible parts **3**:129–130
food usage **3**:130*T*
prolific species **3**:130*T*
research background **3**:128–129
taxonomic diversity **3**:128–129, **3**:129*T*
standards of consumption **3**:127–128
weeds **3**:131–132
Edith's checkerspot butterfly **4**:257–258, **4**:259*F*
editosomes **5**:572
Edmontonia **2**:519*T*, **2**:520*T*
Edmontosaurus **2**:519*T*, **2**:520*T*
Edriolychnus schmidti **5**:670*F*
educational opportunities **3:133–141**
college and university programs **3**:140
environmental education **3**:136, **3**:136
government-based initiatives **3**:133
informal education experiences
biodiversity experiences and resources
Earthwatch Institute **3**:138
National Biological Information Infrastructure (NBII) **3**:138–139
nonprofit environmental groups **3**:139
online resources **3**:139
resources **3**:138
botanical gardens **3**:138
museums **3**:138
resources **3**:136–137
youth-serving groups **3**:137, **3**:137
zoos **3**:137–138
K–12 education
diversity of life concepts **3**:134
formal science education **3**:135–136
grades 3–5 **3**:134
grades 6–8 **3**:134–135
grades 9–12 **3**:135
grades K–2 **3**:133–134
Project 2061 **3**:133–134
public awareness surveys **3**:139–140
research summary **3**:140–141
systematics research and training **3**:140
Edwards Aquifer (Texas, USA) **7**:58, **7**:59*T*
Edwards Plateau **5**:562–563
eelgrass **3**:360–361, **6**:454
eelgrass limpets **3**:202*T*
eelpouts **3**:467–471, **5**:34–35
eels **5**:315, **5**:473, **6**:479
egalitarianism **6**:574–575
Egeria densa **3**:361
eggplant **2**:663, **3**:32*F*
egg ratio analysis **3**:68
egrets **4**:257–258, **5**:476–477
Egretta garzetta **4**:257–258, **5**:476–477
egusi melon **6**:553
Egypt
botanical gardens **6**:519
flora **5**:481*T*
inland deserts **2**:462*F*
megaherbivores **5**:224*T*
species richness **1**:499*T*
wheat production **1**:87–88
see also Near East ecosystems
Egyptian mongoose **5**:475, **5**:477
Egyptian nightjar **5**:474
Egyptian vulture **5**:474
Ehrlichia ruminantium **2**:636*F*
Ehrlich, Paul **3**:99, **6**:205–206
Eichhornia azurea **1**:165
Eichhornia crassipes **1**:54–55, **1**:164, **1**:279, **3**:359–360, **3**:703, **3**:704*F*, **4**:359–360, **6**:94
Eichhornia paniculata **4**:248*F*
Eigen, Manfred **5**:630
Eimer's organ **4**:690–691
Einiosaurus **2**:519*T*
einkorn wheat **2**:658, **2**:662–663
Eisenia arborea **4**:334, **4**:335*F*
Elachistidae **5**:387–388*T*
Elaeagnaceae **5**:623*F*
Elaeagnus angustifolia **7**:399
Elaeidobius kamerunicus **5**:607
Elaeis **7**:272
Elaeis guineensis **1**:543–547, **2**:663, **5**:602, **6**:320, **6**:508, **6**:553
Elaeis melanococca **5**:600–601
Elaeis oleifera **1**:543–547, **4**:285, **4**:288*T*, **5**:602
Elaeocarpaceae **5**:621–625, **5**:623*F*
elands **1**:50, **2**:472–474, **4**:696, **6**:572–573
Elaphurus davidianus **3**:196, **3**:373
elapids **6**:409
elasmobranchs
Census of Marine Life project **2**:7*F*
characteristics **3**:458–459
chimaeras **3**:460, **3**:460*F*
rays **3**:459–460
sharks
characteristics **3**:456–457, **3**:458–459, **3**:459*F*
extinction theories **2**:517
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
skates **3**:459–460, **3**:460*F*
Elasmotherium **5**:224*T*
Elateridae **1**:327*T*
Elateroidea **1**:327*T*
Elbe River **7**:152–153
electromagnetic radiation
electromagnetic spectrum **6**:390–391, **6**:391*F*, **7**:296, **7**:297*F*
measurement and analysis **6**:390–391
remote sensor systems **6**:391–392
Eleocharis **1**:270, **2**:376*F*
Eleocharis albibracteata **4**:107
Eleocharis parvula **3**:209
elephant bird **1**:601, **1**:606
Elephantidae **5**:224*T*, **5**:225*T*
elephants
characteristics **4**:694–695
conservation biology **2**:243–244
deoxyribonucleic acid (DNA) evolution **2**:56*F*
extant species **5**:225*T*
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
keystone species **3**:93
migration impacts **5**:318
number of species **4**:690*T*
Serengeti ecosystem **2**:487–488, **2**:488*F*
size **1**:601
species diversity **4**:690*T*, **4**:694–695
species interactions **2**:352
Succulent Karoo **1**:487–488
threat percentage **3**:195
trophy hunting **3**:197

tropical Africa **1**:49
see also megaherbivores
elephant seals
adipose tissue **7**:34–35
biologging **2**:19*F*
extinctions **5**:87*T*, **5**:88
genetic diversity **2**:263–264
migratory patterns **5**:314
elephant shrews **4**:690*T*, **4**:697–698
elephant trees **2**:471
Elephas iolensis **5**:225*T*, **5**:231–233
Elephas maximus **3**:188, **4**:694–695, **5**:225*T*, **5**:603
Elephas (Palaeoloxodori)antiquus **4**:729–730, **4**:729*F*
Elettaria cardamomum **2**:663
Eleusine coracana **2**:663
Eleusine indica **1**:120
Eleutherococcus senticosus **6**:137
Eleutherodactylus **1**:170
elevational gradients **3:149–154**
diversity hypotheses
biotic hypotheses
ecotones **3**:152
habitat heterogeneity **3**:152–153
source–sink dynamics **3**:152
climatic variables **3**:150*F*, **3**:151–152
general discussion **3**:151–152
historical hypotheses **3**:153
spatial hypotheses
mid-domain effect (MDE) **3**:152
species–area relationship **3**:152
general discussion **3**:149
historical research **3**:149–150, **3**:150*F*
microbial biogeography **5**:273–274
pattern observation and quantification **3**:150–151, **3**:150*F*
research summary **3**:153
Eligmodontia typus **2**:587*F*
Eliot, T. S. **4**:652
Elizabetha paraensis **1**:162
elk
ecosystem/landscape conservation **7**:405–406
grassland ecosystems **6**:332
habitat selection **4**:65
late Quaternary extinction event (LQEE) **4**:724–725*T*
migratory patterns **5**:313
elkhorn coral **2**:339–340, **2**:528*T*, **2**:529*F*, **3**:201*T*
Elkinsia **5**:618*F*
Ellesmere Island **5**:562
Elliot's chimpanzee **6**:256–258*T*
elm trees **2**:528*T*, **3**:411
El Niño Southern Oscillation (ENSO) **3:155–163**
background information **2**:463–464
biodiversity effects **3**:155, **3**:157*F*
expected disturbance events **3**:155–156
extreme event drivers **4**:213
feedbacks **3**:161–162, **3**:161*F*
global change impacts **5**:588–589
harmful algal blooms **1**:480
marine ecosystems **5**:61–62
Multivariate ENSO Index (MEI) **3**:155, **3**:156*F*
observed effects
drought conditions **3**:156–158
precipitation increases
floods **3**:159
water-limited ecosystems **3**:158–159
sea surface temperatures
coral bleaching **3**:160–161, **3**:160*F*
increased temperatures **3**:159–160
nutrient level shifts **3**:160
ocean current shifts **3**:159–160
wildfires **3**:156–158, **3**:158*F*
pelagic ecosystems **5**:686–688, **5**:689*F*, **5**:690*F*
phenological shifts **5**:720–721
research summary **3**:161–162
elopomorphs **3**:463
El Salvador
botanical gardens **6**:519
deforestation **2**:30–31, **2**:31*F*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
species richness **2**:25*T*
Elton, Charles **3**:239, **3**:588, **4**:121
Eltonian niche **4**:33–34
Elysia chlorotica **5**:119–120, **5**:119*F*
Emballonuridae **6**:309*T*
Emberizidae **6**:308*T*
Embioptera **4**:308–312, **4**:309*T*, **4**:392*T*
Emblica officinalis **1**:70
Emblingiaceae **5**:221
Embolotherium **5**:224*T*
embryophytes **4**:168–169, **6**:10*T*
emerging and re-emerging diseases **3**:285–286
Emerson, Ralph Waldo **2**:280
emetine **5**:712*T*
Emiliania **6**:30, **6**:293–294
Emiliania huxleyi **1**:474*F*, **5**:684*F*
emmer wheat **2**:658, **2**:662–663, **3**:671*F*
Empetrum rubrum **4**:106–107
Empididae **3**:489–491*T*, **3**:492
Empidoidea **3**:489–491*T*
Empidonax traillii **3**:260
Empoasca **1**:120
Emporia **5**:618*F*
empty forest syndrome **6**:253
emu **1**:606
emydids **6**:400–401
encapsulation **5**:678
Encephalartos brevifoliolatus **6**:84
Encephalartos longifolius **3**:211, **3**:213*F*
Encephalartos nubimontanus **6**:84
Encephalartos whitelockii **6**:84
Encephalartos woodii **6**:84
enchytraeid worm **4**:393, **4**:394*F*
Encyclia ionosma **1**:166
Encyclopedia of Life (EOL) **3**:139, **4**:327
endangered species
amphibians **3:164–168**
background and characteristics **3**:164
conservation status **3**:164, **3**:164*T*
environmental threats
chemical pollution **3**:165–166
climate change **3**:165
commercial exploitation **3**:166–167
disease **3**:166
general discussion **3**:164–165
habitat destruction **3**:164–165
introduced species **3**:167
ultraviolet radiaion **3**:165
population declines
developmental deformities **3**:167–168
protected habitats **3**:167–168
birds **7:180–190**
biotic threats
alien species **7**:184
climate change **7**:185–186
general discussion **7**:184
hunting, trapping, poisoning **7**:185
native–native threats **7**:184–185
small ranges **7**:185
causal factors
biotic threats **7**:184
general characteristics **7**:182–183
habitat loss and degradation **7**:183
conservation
critically endangered birds **7**:187–188
general discussion **7**:186
goals and priorities **7**:186
international cooperation and legislation **7**:187
priority management **7**:186
scale considerations **7**:186
site and habitat conservation **7**:186–187
critically endangered birds
alien and native restriction and control **7**:188
captive breeding programs **7**:188
challenges **7**:189
cross-fostering/cross-breeding **7**:189
habitat restoration **7**:187–188
nest site provision/enhancement **7**:189
reintroduction/translocation programs **7**:188
supplementary feeding programs **7**:188–189
current threats **1**:622–623, **1**:622*F*
global and regional trends
criteria **7**:180–181
extinct birds **7**:182
future outlook **7**:181
number of species **7**:181
regions and countries **7**:181
taxonomic groups **7**:181–182
United States/United Kingdom **7**:182
habitat loss and degradation
general discussion **7**:183
habitat degradation **7**:183–184
habitat destruction **7**:183
habitat fragmentation **7**:183–184
causal factors **3**:199–202
definition **3**:199
freshwater invertebrates **3:176–187**
American Southeast rivers **3**:184, **3**:185*F*
anthropogenic impacts
direct harvests **3**:181
global climate change **3**:181–182

endangered species (*continued*)
habitat destruction/degradation **3**:180–181, **3**:180*F*
introduced species **3**:181
pollution **3**:181, **3**:181*F*
causal factors **3**:177, **3**:177*F*
desert spring communities **3**:186, **3**:187*F*
diversity hot spots **3**:176
global conservation status **3**:176–177, **3**:176*T*, **3**:177*F*
global distribution patterns **3**:182–184, **3**:182*T*
global diversity **3**:176
number of species **3**:182–184, **3**:182*T*
protection strategies **3**:184–187
small ranges
causal factors **3**:178
conservation status **3**:177*F*
dispersal ability **3**:178*F*
high endemic hot spots **3**:178–179
high endemic species **3**:179, **3**:180*F*
Lake Baikal **3**:179, **3**:183*F*
narrowly endemic species **3**:177–178, **3**:177*F*, **3**:179
vulnerability factors
small ranges **3**:177–178, **3**:177*F*
sparse populations **3**:179–180
fungi **3**:639–640, **3**:639*F*
IUCN criteria **3**:200–201*T*, **7**:196–197, **7**:197
mammals **3:188–198**
conservation practices
captive breeding programs **3**:196–197, **4**:710–711
charismatic species **3**:197–198
economic incentives **3**:197
government legislation and regulations **3**:195
protected area systems **3**:196, **4**:710–711
reintroduction programs **3**:196–197
Species Survival Commissions (SSC) **3**:195–196
umbrella species **3**:197–198
future outlook **3**:198
geographic locations
endemism **3**:193–194
top ranked countries **3**:193, **3**:194*T*
historical perspective
early/prehistoric era **3**:188–189
Pleistocene extinctions **3**:189
recent extinctions **3**:189–190, **3**:189*F*
major threats
causal factors **3**:190
disease **3**:191
exploitation **3**:191–192
genetic diversity loss **3**:190–191
habitat loss and fragmentation **3**:190
interference **3**:192
introduced species **3**:193
livestock **3**:191
multiple threats **3**:193
persecution **3**:192
physiological factors **3**:190
pollution **3**:192–193
Red List Index (RLI) **3**:194–195, **3**:195*F*, **4**:709
taxonomic percentages **3**:195
marine invertebrates **3:199–204**
causal factors **3**:199–202
definition **3**:199
extinctions **3**:202*T*
IUCN criteria **3**:200–201*T*
recovery strategies **3**:203
research challenges **3**:202–203
species examples **3**:199–202, **3**:201*T*, **3**:202*T*
susceptible habitats and species **3**:202, **3**:202*T*, **3**:203*T*
plants **3:205–215**
causal factors **3**:208–209
conservation implications **3**:212–215
endangerment patterns **3**:206–208, **3**:207*F*
global distribution patterns **3**:207*F*
legal protection mandates
International Plant Protection Convention **3**:209
U.S. Endangered Species Act (1973) **3**:209
occurrences **3**:205–206
rarity
causal factors **3**:208–209
characteristics **3**:206
consequences **3**:209
selected examples
Coleus forskohlii **3**:209–210, **3**:210*F*
Encephalartos longifolius **3**:211, **3**:213*F*
Fitzroya cupressoides **3**:210, **3**:211*F*
Pterostylis truncata **3**:211, **3**:213*F*
Sarracenia rubra ssp. *alabamensis* **3**:210–211, **3**:212*F*
species-environmental matching models **3**:211–212
reptiles **3:216–218**
characteristics and classifications **3**:216
conservation status **3**:216, **3**:216*T*
major threats
climate change effects **3**:218, **6**:410
commercial exploitation **3**:217
direct impacts **3**:216–217
habitat destruction **3**:216–217
introduced species **3**:217–218
pollution **3**:218
saiga antelope **4**:711–712, **4**:712*F*
salmon **6**:523
terrestrial invertebrates **3:219–227**
centipedes **3**:226
conservation practices **3**:224
crabs **3**:226
daddy long legs **3**:225–226
earthworms **3**:225
insects **3**:226
land snails **3**:226
millipedes **3**:226
pill bugs **3**:226
pseudoscorpions **3**:225–226
Red List Index (RLI) **3**:225
spiders **3**:225–226
status assessments **3**:223–224
velvet worms **3**:225, **3**:226*F*
woodlice **3**:226
wildlife management **7**:405
Endangered Species Act (ESA)
basic principles **3**:713*T*, **3**:716–717
genetic diversity protection **2**:264
historical perspective **2**:286, **4**:121–122, **6**:77–78, **7**:314–315
marine invertebrates **3**:199
species protection **2**:108–109
threatened species lists and assessments **7**:199–200
timber industry **7**:314–315
wildlife management **1**:397
Endeavor Hydrothermal Vents Marine Protected Area **5**:447
endemic bird areas (EBAs) **1**:615–616, **7**:186
endemism **3:228–236**
biodiversity-rich countries **1**:497, **1**:499–500
birds **1**:615–616
categorization
affinities **3**:228, **3**:228
evolutionary age **3**:228, **3**:228
local abundance **3**:228–229
spatial distribution **3**:228
Central America **2**:26
conservation practices **3**:235
correlates and causal factors
abiotic environmental factors **3**:232–233
area **3**:232
biological factors
characteristics **3**:233–234
geographic range size–body mass relationship **3**:234*F*
maximum local abundance **3**:233*F*
South Africa **3**:234*F*
biotopes **3**:233
phylogeny **3**:234–235
regional species richness **3**:232
research background **3**:231–232
speciation **3**:235
taxonomy **3**:234–235
endangered mammals **3**:193–194
freshwater invertebrates **4**:374–375
high-elevation Andean ecosystems **4**:99–100, **4**:99*T*
hydrothermal vent ecosystems **4**:391
mammals **4**:702
measurement approaches
biases **3**:230
characteristics **3**:229
percentage versus counts **3**:229–230, **3**:230*F*
units of measurement **3**:229
Mediterranean-type climate ecosystems
Australia **5**:221
California **5**:215
Cape Region (South Africa) **5**:219, **5**:220
Chile **5**:216–217
microbial biogeography **5**:272, **5**:273*F*
myriapods **5**:428
Near East ecosystems **5**:463–464, **5**:464*F*, **5**:464*T*
neotropical mammals **2**:589–592, **2**:591*F*

oceanic islands **5**:592, **5**:592*F*, **5**:592*T*
patterns
centers **3**:231, **3**:231*F*
congruent areas **3**:231, **3**:232*T*
general discussion **3**:230–231
latitudinal gradients **3**:230–231, **3**:230*F*
megadiversity **3**:231, **3**:232*T*
plant biodiversity **6**:61
primates **6**:252*T*
research areas **3**:235
temperate forests **7**:121
theoretical perspectives **3**:229
Endocladia muricata **6**:723*T*
endocommensals **6**:296–297
endocrine disruptors
ecotoxicological studies **3**:118
environmental threats **3**:166
endogamy **2**:322–323
Endomychidae **1**:327*T*
endoparasites **5**:657, **5**:668, **5**:671*F*
endophytic fungi **2**:429, **3**:636
endophytic parasites **4**:70
Endopleura uchi **4**:287–290, **4**:288*T*
Endopterygota **4**:308–312, **4**:309*T*
endosulfan **5**:693, **5**:699*T*
endosymbionts
hydrothermal vent ecosystems **7**:328–329
structural characteristics **5**:116–117, **5**:117*F*
endothal sodium **5**:697
Endothia parasitica **2**:530–531
end-Permian mass extinction **5**:158, **5**:161*F*
endrin **5**:693
enemy-free space **6**:244
energetic food webs **3**:499*F*, **3**:500–501, **3**:507–508
energy flow **3:237–249**
benefits **3**:237–238
biodiversity effects
experimental research **3**:247
number of species effects **3**:245–246
research background **3**:245–247
top-down versus bottom-up regulation **3**:246–247
ecosystem boundaries
analytical studies **3**:241–242
basic concepts **3**:238–239, **3**:238*F*
control mechanisms **3**:241, **3**:241*F*
food web–trophic level relationship **3**:240–241, **3**:241*F*
historical perspective **3**:238–239
trophic dynamic concept **3**:239–240
future research outlook **3**:247–248
importance **3**:247
multiple energy pathways
background information **3**:242–243
depth–light wavelength relationship **3**:243
lake and pond ecosystems **3**:242–243, **3**:243*F*
spatial heterogeneity **3**:243
nutrient cycling
benefits **3**:237–238
food chains **3**:244–245
internal versus external supplies **3**:243–245
microbial loops **3**:244–245
organic detritus **3**:244–245
species interactions **3**:247
Energy Independence and Security Act (2007) **4**:539–540
energy production **7**:203*F*
energy sprawl **1**:528–529, **1**:529*F*
energy use **3:250–266**
biodiversity impacts
carbon dioxide (CO_2) emissions **3**:265*T*
industrial energy
coal **3**:257–259, **3**:264*T*, **3**:265*T*
hydroelectric power **3**:259–260, **3**:264*T*, **3**:265*T*
natural gas **3**:256–257, **3**:264*T*, **3**:265*T*
nuclear energy **3**:259, **3**:264*T*, **3**:265*T*
oil **3**:253–256, **3**:255*T*, **3**:264*T*, **3**:265*T*
land requirements **3**:265*T*
major sources **3**:264*T*
renewable energy **3**:260–262, **3**:264*T*, **3**:265*T*
traditional energy **3**:262–263, **3**:263*T*, **3**:264*T*, **3**:265*T*
world energy reserves **3**:265–266, **3**:265*T*
biofuels **1:528–539**
biodiversity-friendly energy-crop production strategies **1**:535–536
biodiversity restoration **6**:459
biofuel production **1**:528–529, **1**:529*F*, **1**:529*T*
crop selection and management
biodiversity management strategies **1**:533
biodiversity responses **1**:529–530, **1**:530*F*
first-generation biofuel crops **1**:529*T*, **1**:530
second-generation biofuel crops **1**:530*F*, **1**:531–532
energy sprawl concerns **1**:528–529, **1**:529*F*
energy use impacts **3**:261–262
first-generation biofuel crops
economic importance **1**:529*T*, **1**:530
feedstock management **1**:530–531
feedstock selection **1**:530
landscape–scale biodiversity considerations
agroenergy sprawl **1**:534–535
general discussion **1**:533–534
spatial considerations **1**:533–534
life cycle analysis (LCA) **1**:535–536, **4**:628–629
policy development impacts **1**:536
research summary **1**:536–537
second-generation biofuel crops
economic importance **1**:530*F*, **1**:531–532
native polycultures **1**:532–533
perennial monocultures **1**:532
short-rotation woody biomass **1**:531–532
categories
industrial energy **3**:250–253
renewable energy **3**:251
traditional energy **3**:250
energy efficiency measures **3**:251, **3**:254*F*
energy sprawl **1**:528–529, **1**:529*F*
environmental impact **3**:263–266
future trends **3**:263–266
thermodynamic laws **3**:253, **3**:254*T*
world energy reserves **3**:265–266, **3**:265*T*
world energy use **3**:251, **3**:251*F*, **3**:252*T*, **3**:253*T*
England
Convention on Biological Diversity (1930) **2**:314
Dartmoor National Park **5**:445
invasive species
Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346
Fallopia japonica **4**:351–352, **4**:352*F*
Spartina alterniflora **4**:349–350
natural history museums and institutions **5**:408*F*, **5**:409*F*, **5**:411*F*
English sparrows **1**:77–78, **1**:81, **1**:82, **1**:83, **4**:365–366
Engraulis ringens **3**:478
Enhalus **6**:542*T*
Enhalus acoroides **6**:541*T*
enhanced greenhouse effect **4**:18–19, **4**:19*F*, **4**:19*T*, **4**:20*F*
Enhydra lutris
anthropogenic impacts **5**:58–60
biomonitoring processes **3**:125
characteristics **5**:68
fur trade **3**:191–192
Holocene postglacial spread **2**:551–552*T*
keystone species study **4**:443–446, **4**:444*F*, **4**:445*F*
oil spills **3**:256
overharvesting/overcollecting **3**:283
species diversity **4**:693
trophic cascades **7**:259–260
Enhydra lutris kenyoni **5**:81*T*, **5**:84–86
Enhydra lutris nereis **5**:81*T*, **5**:83–84
Enicmus apicalis **7**:220*T*
Enicmus planipennis **7**:220*T*
Enigmatheres canfieldi **3**:201–202
Enlightenment Movement **2**:280–281
Enneapogon cenchroides **1**:67
enoli mushrooms **3**:637*T*
enriched flour **3**:638–639
Ensatina eschscholtzii **2**:512, **6**:622–624, **6**:623*F*
enset **2**:663
Ensete ventricosa **2**:663
Entamoeba coli **5**:656
Entamoeba histolytica **5**:656–657, **5**:660, **5**:660*T*, **5**:670
entangled bank metaphor **2**:202–204
Entelegynae **1**:207
Enterobius vermicularis **5**:660, **5**:660*T*, **5**:671*T*, **7**:425–428, **7**:428*T*
Enterolobium cyclocarpum **4**:592, **6**:52
Enterolobium schomburgkii **1**:161–162
Enteropneusta **5**:669–670*T*
Entisols **6**:123*T*
Entodinium **6**:295–296
Entomophaga maimaiga **5**:395–396

Entomopoxivirinae **5**:136
Entoprocta **3**:684*F*
entoprocts **4**:382*F*, **5**:669–670*T*
Entosiphon **6**:293
Environmental Easement Program **3**:719
environmental ethics **3:267–277**
  biodiversity protection **3:309–318**
    animal rights **3**:313–314
    basic principles **3**:309
    conservation biologists
      personal choices **3**:317–318
      professional considerations **3**:317
      role models **3**:316–317, **3**:317*F*
    environmental philosophies
      deep ecology **3**:274–276, **3**:275*T*
      dominant worldview **3**:275*T*
      humanism **3**:276–277
    goals
      ecological integrity **3**:310–311, **3**:312*F*
      economic factors **3**:309–310
      long-term healthy ecosystems **3**:309–310, **3**:310*T*, **3**:311*F*
      sustainability **3**:311–312
    human flourishing
      artistic expression **3**:273, **3**:274*F*
      general discussion **3**:273–274
      historical understanding **3**:274
      life-support systems **3**:274
      recreational activities **3**:273, **3**:273*F*
      religious inspiration **3**:274
      scientific knowledge **3**:273–274
    human obligations **3**:271–272
    human rights
      challenges **3**:314–315
      less-developed countries **3**:315–316, **3**:316*F*
      logging industry–owl population controversy **3**:314–315, **3**:315*F*
      property rights **3**:315
    intrinsic value
      anthropocentric denials **3**:269
      economic arguments **3**:267
      ethical holism **3**:269–270, **3**:270*F*
      extensionist arguments **3**:267–268
      nonextensionist arguments **3**:268–269, **3**:269*F*
    management versus nonmanagement **3**:312–313, **3**:313*F*
    religious traditions
      Asian traditions **6**:383–384
      Buddhist declaration on nature **3**:272, **6**:383
      Christian declaration on nature **3**:272–273, **6**:381, **6**:381
      Hindu declaration on nature **3**:273, **6**:383
      Jewish declaration on nature **3**:273, **6**:381, **6**:382
      Muslim declaration on nature **3**:273, **6**:381, **6**:382
      pantheistic traditions **6**:384–385
      stewardship **3**:272, **6**:381–383
  environmental philosophies
    deep ecology **3**:274–276, **3**:275*T*
    dominant worldview **3**:275*T*
    humanism **3**:276–277
    ethical holism **3**:269–270, **3**:270*F*
  human flourishing
    artistic expression **3**:273, **3**:274*F*
    general discussion **3**:273–274
    historical understanding **3**:274
    life-support systems **3**:274
    recreational activities **3**:273, **3**:273*F*
    religious inspiration **3**:274
    scientific knowledge **3**:273–274
  human obligations **3**:271–272
  intrinsic value
    anthropocentric denials **3**:269
    economic arguments **3**:267
    ethical holism **3**:269–270, **3**:270*F*
    extensionist arguments **3**:267–268
    nonextensionist arguments **3**:268–269, **3**:269*F*
  religious traditions
    Asian traditions **6**:383–384
    Buddhist declaration on nature **3**:272, **6**:383
    Christian declaration on nature **3**:272–273, **6**:381, **6**:381
    Hindu declaration on nature **3**:273, **6**:383
    Jewish declaration on nature **3**:273, **6**:381, **6**:382
    Muslim declaration on nature **3**:273, **6**:381, **6**:382
    pantheistic traditions
      ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
      indigenous peoples **6**:384, **6**:384–385
      traditional ecological knowledge **6**:385
      traditional social restraints **6**:385–386
    research summary **6**:387–388
    stewardship **3**:272, **6**:381–383
    traditional social protection
      background information **6**:385–386
      biological communities/habitat patches **6**:386
      critical life history stages **6**:386
      long-lived (K-selected) species **6**:386
      stewardship **6**:386
  translocations **7**:239
  virtue ethics **3**:275
environmental impact **3:278–296**
  biotic impoverishment
    characteristics **3**:279–281, **3**:280*T*
    direct biotic depletion
      biotic homogenization **3**:284
      genetic engineering **3**:284–285
      habitat fragmentation and destruction **3**:283–284
      nonhuman life **3**:283
      overharvesting/overcollecting **3**:283
    direct degradation of human life
      cultural diversity loss **3**:286
      cumulative effects **3**:288
      emerging and re-emerging diseases **3**:285–286
      environmental injustice **3**:286–287
      historical perspective **3**:285–286
      political instability **3**:287–288
      quality of life reductions **3**:286
    indirect biotic depletion
      altered biogeochemical cycles **3**:282
      chemical contamination **3**:281–282
      general discussion **3**:280–281
      global climate change **3**:282–283
      soil depletion **3**:281
      water degradation **3**:280–281
  causal factors
    fragmented worldviews **3**:288–289
    gift economies **3**:288–289
    historical perspective **3**:288–289
    overconsumption **3**:289–290, **3**:290*F*
  Development by Design (DbD) **2:495–507**
    background and characteristics **2**:496–497, **2**:497*F*
    framework and components
      basic concepts **2**:497–498
      landscape conservation planning **2**:498
      mitigation decision-making approaches **2**:498–500, **2**:499*F*
      offset contribution estimation **2**:500–503
      research summary **2**:505–506
      residual impact/optimal offset determinations **2**:500, **2**:501*F*, **2**:502–503*F*
      sustainable landscapes **2**:503–505, **2**:504–505*F*
      Wyoming Basins Ecoregion (WBE) study **2**:503–505, **2**:504–505*F*
    research summary **2**:505–506
  government policies and regulations **3**:294–295
  human impact timeline **3**:278–279
  management strategies **3**:294–295
  measurement and analysis
    biological assessments **3**:292–294, **3**:292*T*, **3**:293*F*, **3**:293*T*
    ecological footprint analysis (EFA) **3**:291–292
    environmental quality indicators **3**:290–291, **3**:291*T*
    general sustainability indexes **3**:290–291
    index of biological integrity (IBI) **3**:292–294, **3**:292*T*, **3**:293*F*, **3**:293*T*
  research summary **3**:295–296
  risk assessments **3**:294–295
  single species impacts **3**:278
Environmental Protection Agency (EPA) **1**:95–97, **4**:90–91
Environmental Resources Information Network (ERIN) **1**:506
environmental stress **7:39–44**
  assessment measures
    appraisals **7**:42
    biomonitoring processes **7**:43
    early warning **7**:43
    functional role **7**:42
    predictions **7**:42–43
  biospheric life support system **7**:43
  characteristics **7**:39
  characterization studies **7**:39–40

future trends **7**:43–44
general environmental stress syndrome **7**:41–42, **7**:42*T*
natural versus anthropogenic stress **7**:39
receptors and responses **7**:40–41, **7**:41*T*
spatiotemporal distributions **7**:39–40
ENVISION software **5**:358
Enypniastes **2**:18*F*
enzymes **6**:113*T*
Eoaves **1**:607
Eocene
boundary age **5**:1*T*
C_4 plants **2**:376*F*
crustaceans **2**:406
Eocene–Oligocene boundary **5**:8
global change indicators **4**:225*F*
mammals **2**:583*F*, **4**:685
mass extinctions **5**:176
megaherbivores **5**:224*T*
oxygen-18 ($\delta^{18}O$) measurements **1**:557*F*
Paleocene–Eocene thermal maximum **5**:7–8
terrestrial plant communities **6**:74*F*
vegetation–climate interactions **2**:683
Eopolita protensa **5**:465*F*
Eosalmo driftwoodensis **6**:522–523
Eozygodon **5**:224*T*
Epermenioidea **5**:387–388*T*
Ephedra **1**:268, **5**:618*F*, **6**:59–60
Ephedra foeminea **5**:481–482
Ephedra shunnungiania **2**:432
Ephedra sinica **2**:432, **5**:712*T*
ephedrine **2**:432, **5**:712*T*
Ephelota gemmipara **5**:656
ephemeral plants **2**:464, **2**:464*F*, **2**:471–472
Ephemeroptera
major divisions and characteristics **4**:308–312, **4**:309*T*
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
wetland ecosystems **7**:389–391
Ephippigerinae **3**:727
Ephydridae **3**:489–491*T*
Ephydroidea **3**:489–491*T*
epibatidine **3**:166–167, **5**:711
Epicampes macrura **2**:30
Epidendrum **6**:316*T*
Epidendrum nocturnum **1**:166
Epidendrum rigidum **3**:517
Epidinocarsis lopezi **4**:362
Epifagus virginiana **4**:73, **4**:75
Epigonichthyes **7**:336
epikarst zone **7**:49–50, **7**:50*F*
epilimnion **3**:560, **4**:374*F*
Epilobium angustifolium **1**:634
Epilobium conjugens **4**:106–107
Epimanteoceras **5**:224*T*
Epinephelus lanceolatus **3**:468*F*
Epioblasma lewisii **3**:185*F*
Epiocheirata **1**:212–213
Epipactis veratrifolia **5**:486
Epipedobates **3**:166–167
Epipedobates tricolor **1**:176
epiphytes
ant-associated epiphytes **3**:523–524
atmospheric epiphytes **3**:524
bark epiphytes **3**:524
elevational range shifts **4**:605*T*
hemiepiphytes
characteristics **3**:518
taxonomic distributions **3**:520–521, **3**:522*T*
tropical rain forests **6**:315*T*
humus epiphytes **3**:523
nonvascular epiphytes
characteristics **3**:517–518
taxonomic distributions **3**:520, **3**:521*T*
plant–host interactions **4**:74
tank/trash basket epiphytes **3**:523
tropical forest ecosystems **7**:272
tropical rain forests **6**:315*T*, **6**:316*T*, **6**:318*T*
vascular epiphytes
characteristics **3**:517, **3**:518*T*
taxonomic distributions **3**:519–520, **3**:519*T*, **3**:520*T*
Epipremnum **3**:522*T*
Episernus angulicollis **7**:220*T*
Episoriculus spp. **1**:702
epsilon diversity **6**:534
Epuraea deubeli **7**:220*T*
Epuraea longula **7**:220*T*
Epuraea muehlii **7**:220*T*
equatorial capuchin **6**:256–258*T*
Equatorial Guinea **4**:560*F*, **6**:507*T*
Equidae **2**:583*F*, **4**:690*T*, **4**:695, **4**:727*T*
Equus
late Quaternary extinction event (LQEE) **4**:722, **4**:724–725*T*, **4**:727*T*
species diversity **4**:695
Equus africanus **1**:51, **2**:476*T*
Equus asinus **1**:76–77, **2**:476*T*, **3**:192
Equus burchelli **4**:695, **4**:705, **6**:573
Equus caballus **1**:76–77, **3**:191, **5**:313, **6**:573, **7**:29*F*
Equus capensis **5**:364–365
Equus ferus przewalskii **3**:191
Equus fraternus **6**:52
Equus grevyi **1**:51, **4**:695, **6**:573
Equus hemionus **2**:469–470, **3**:191, **6**:334, **6**:573, **6**:578–579
Equus hemionus hemippus **5**:472
Equus przewalskii **1**:269, **6**:333
Equus quagga **1**:487–488, **4**:705, **5**:235
Equus quagga boehmii **1**:52
Equus zebra **1**:487–488, **4**:695
Eragrostis australasica **2**:468
Eragrostis glomerata **1**:165
Eragrostis hypnoides **1**:165
Eragrostis lehmannian **6**:457
Eragrostis tef **2**:663
Eragrostris lehmanniana **4**:446*F*
Erebidae **5**:387–388*T*, **5**:388*F*, **5**:389–390, **5**:391*F*, **5**:400*F*
Eremina desertorum **5**:468*F*
Eremogryllus hammadae **5**:469
Eremolepidaceae **4**:71
Eremosynaceae **5**:221
Eremotherium **4**:724–725*T*, **4**:727*T*
Eremotherium mirabile **5**:225*T*
Erethizon dorsatum **1**:703
Erethizontidae **6**:309*T*
Ergasilus **6**:26
ergosterol biosynthesis inhibitors (EBIs) **5**:695
ergot **3**:633
Eria **6**:316*T*
Eriachne **2**:376*F*
Eriastrum diffusum **2**:373*T*
Erica **5**:219
Ericaceae **1**:266*T*, **1**:267, **3**:129*T*, **3**:130*T*, **3**:522*T*, **5**:219, **5**:221, **6**:315*T*
Ericanae **5**:622*F*, **5**:625
Erica tetralix **6**:124–125
ericoid mycorrhizae **3**:636
Erie (Lake) **4**:461*T*
Erigeron annuus **6**:433, **6**:434*F*
Erigeron divergens **2**:373*T*
Erigeron leptopetalous **4**:107
Erignathus barbatus **5**:90
Erinaceomorphia **4**:691
Erinaceus spp. **2**:550
Eriobotrya japonica **2**:663
Eriocaulaceae **6**:44*T*
Eriocheir sinensis **4**:465
Eriocraniidae **5**:385–386
Eriocranioidea **5**:386*F*, **5**:387–388*T*
Eriogonum **5**:215
Eriogonum abertianum **2**:373*T*
Eriophorum **1**:270, **1**:277–278, **3**:76*F*
Eriophorum angustifolium **1**:232
Eriophorum vaginatum **1**:136, **1**:232
Eritrea **2**:467–468, **4**:560*F*, **6**:507*T*
Erodium cicutarium **2**:373*T*, **4**:364
erosion
domesticated ruminants **1**:704–705
shield deserts **2**:460–461
soil quality **6**:591, **6**:592*F*
Erotylidae **1**:327*T*
Erpocotyle **7**:448–449*F*
Eruca vesicaria **3**:131–132
erucic acid **7**:26
Eryngium foetidum **4**:288*T*
Erysiphe graminis **3**:633
Erythemis collocatta **4**:119
Erythorxylum **5**:489
Erythrina **2**:427
Erythrina fusca **7**:274
Erythrina ulei **1**:161
Erythroxylon coca **2**:663–664
Erythroxylum coca **4**:287–290, **4**:289*T*, **5**:712*T*
escalation/diversification hypothesis **1**:23
Escherichia coli
molecular level diversity **2**:636*F*
neutral nucleotide polymorphism **2**:631*T*
proteorhodopsins **6**:280–282, **6**:281*F*
random reproduction **2**:509
Eschrichtius robustus **5**:76, **5**:77*T*, **5**:87–88, **5**:87*T*, **6**:236
Eschscholtz's mitred surili **6**:256–258*T*
Eschweilera albiflora **1**:163
Eschweilera tenuifolia **1**:163
Eskimo curlew **7**:182
Esociformes **3**:464–465, **6**:522–523
Esox lucius **4**:463
Espeletia **4**:102*F*, **4**:103, **7**:275–276
Espeletia schultzii **4**:101
Estigmene acrea **5**:400*F*
Estonia

botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
17β-estradiol **3**:118
Estrilda astrild **1**:83
estrogens **3**:118
estuarine ecosystems **3:297–308**
biodiversity
characteristics **3**:299–300
ecological functions
biogeochemical cycles **3**:301
biological properties **3**:301
chemical modifications **3**:301
climate variations **3**:302
ecosystem modifications **3**:301
exchange of materials **3**:301
extreme events **3**:302
freshwater delivery **3**:301
general discussion **3**:300–301
physico–chemical properties **3**:301
sea level changes **3**:302
suspended matter **3**:301
temperature changes **3**:302
transport changes **3**:301
wind **3**:302
fish **3**:299–300
characteristics and distribution **3**:297–298
Chesapeake Bay case study **3**:303–306, **3**:304*F*, **3**:305*F*, **3**:306*F*
classifications **3**:298–299
climate change effects **6**:150–151
eutrophication
aquatic vascular plants **3**:360–361
contaminant–stress relationship **5**:29–30
high nutrient, low growth (HNLG) concept **5**:26–27, **5**:26*T*, **5**:27*F*
macroalgae **3**:358–359
macrophytes **3**:360–361
microfauna **3**:363–364, **3**:363*F*
nutrient ratio changes **5**:28–29, **5**:28*T*, **5**:29*F*
functional role **3**:297–298
future challenges **3**:306–307
habitat loss and degradation **7**:371*T*
marine ecosystems **5**:45
marine sedimentary habitats **5**:108*T*
nitrogen cycle impacts **5**:543
salt water plants **6**:77
ESValue software **5**:358
ethanol production
biofuel crop production **1**:529*T*, **1**:530, **1**:543–547, **1**:544*F*
carbon dioxide (CO_2) emissions **4**:295*F*
croplands and pasturelands **1**:542–543, **1**:542*F*
land-use change carbon dioxide (CO_2) emissions **4:539–554**
Global Trade Analysis Project (GTAP) model **4**:539–540
consumption-side structure **4**:540*F*
economic and environmental impacts **4**:540–544
emission factors **4**:548–549
estimated emissions **4**:549–550, **4**:549*T*
final analysis **4**:551–552, **4**:551*T*
land cover and land-use activities **4**:542*F*
livestock production structure **4**:542*F*, **4**:543*F*, **4**:543*T*
marginal and average estimated emissions (first simulated group) **4**:550, **4**:550*T*
marginal and average estimated emissions (second simulated group) **4**:550–551, **4**:550*T*
marginal and average estimated emissions (third simulated group) **4**:551, **4**:551*T*
modeling process **4**:539–540
production-side structure **4**:541*F*
productivity measures **4**:540–544
research limitations **4**:553
research summary **4**:552–553
simulation models **4**:544–545
2001 simulation model **4**:544–545, **4**:544*T*, **4**:545*T*
2001–2006 simulation model **4**:545–547, **4**:546*T*, **4**:547*T*
2006–2015 simulation model **4**:547–548, **4**:548*T*, **4**:549*T*
well-to-wheel emissions **4**:551–552, **4**:552*T*, **4**:553*T*
Saccharomyces cerevisiae **3**:638
Etheostoma fonticola **7**:58
Ethiopia
botanical gardens **6**:519
desert ecosystems **2**:467–468
fish diversity **3**:472–473
grass and shrub savannas **1**:67
indigenous peoples **4**:270*T*
intercropping systems **2**:382–383
Lepidoptera **5**:389*T*
wheat production **1**:87–88
Ethiopian wolves **3**:191
ethirimol **5**:695
Ethmodiscus **6**:30
ethnobiology **3:319–328**
current trends
change mechanisms **3**:327
cognitive mapping **3**:324–325
ethnobiological classification **3**:326
internationalization **3**:327–328
knowledge variation **3**:326–327
research background **3**:324–325
resource management and valuation **3**:325–326
ritual and spiritual symbolism and practices **3**:327
scientific covalidation **3**:326
definitions **3**:319–320
future research areas **3**:328
historical perspective
ayurvedic scholars **3**:322
Aztec culture **3**:322
botanical research **3**:322
Chinese scholars **3**:321–322
classical Europe **3**:320–321
Middle Ages **3**:321
Moorish era **3**:321
periods and stages **3**:325*T*
prehistoric era **3**:320–321
Renaissance era **3**:322
research areas
Clementsian theory **3**:324, **3**:325*T*
economic botany **3**:323
ethnobotany **3**:323
ethnoecology **3**:323–324
ethnoscience **3**:323
pharmacognosy **3**:323
theoretical science **3**:322–323
ethnobotany **3**:323
ethnoecology **3:319–328**
definitions **3**:319–320, **3**:323–324
historical perspective **3**:323–324
indigenous peoples
traditional agricultural knowledge systems
environmental knowledge **1**:121
experimental approaches **1**:122
farming practices **1**:121
folk taxonomies **1**:121
multiple cropping systems **1**:119–120
research background **1**:120–121
ethnoscience **3**:323
ethnosphere **3**:279
ethylene dibromide **5**:694
Etimonotrysia **5**:387–388*T*
etiological tales of nature **4**:650
eubacteria
common ancestry theory **3**:397*F*
known and estimated species **4**:384*T*
Eubalaena australis **2**:587*F*, **5**:76–78
Eubalaena glacialis **3**:192, **5**:76–78
Eubalaena japonica **5**:76–78
EU Biodiversity Observation Network (E-BONE) **4**:327
Eucalanus subtenuis **4**:385
Eucalyptus **1**:279, **1**:291–292, **4**:360, **5**:221, **6**:615, **7**:125, **7**:223
Eucalyptus amygdalina **4**:173*F*
Eucalyptus baxteri **4**:173*F*
Eucalyptus calophylla **5**:220–221
Eucalyptus diversicolor **5**:220–221
Eucalyptus gomphocephala **5**:220–221
Eucalyptus marginata **5**:220–221, **6**:440*F*, **6**:458
Eucalyptus oblique **4**:173*F*
Eucalyptus patagonica **1**:291–292
Eucalyptus regnans **1**:302–303
Eucalyptus risdonii **4**:173*F*
Eucalyptus spathulata **1**:291–292
Eucarida **2**:401*T*
Euceratherium **4**:724–725*T*
Eucestoda
Bothriocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Caryophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cathetocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cyclophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Diphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Diphyllobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458

Haplobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Lecanicephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Litobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Nippotaeniidea **7**:439–440*T*, **7**:456–457*F*, **7**:458–459
Proteocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Rhinebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Spathebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetrabothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetraphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Trypanorhyncha **7**:439–440*T*, **7**:456–457*F*, **7**:459–460
Euchaetes egle **5**:397
Eucheta hebes **5**:689–690
eucheuma **1**:190*T*
Euchilotermes **4**:409*T*
Euchondrus saulcyi **5**:469–470
Euchondrus septemdentatus **5**:469–470
Eucnemidae **1**:327*T*
Eucommiaceae **1**:260*T*
Eucryphia **7**:119
Eucryphiaceae **5**:621–625, **5**:623*F*
Eucryphia cordifloria **3**:117
Eucyclops serrulatus **5**:471
Eudema hauthalii **4**:107
Eudema nubigena **4**:103
Eudiaptomus **6**:26
Eudicots **5**:618*F*
Eudocinus ruber **5**:13–14
Eudorina **6**:30
Eudorina unicocca **6**:33*F*
Eudorylaimus antarcticus **7**:429
Eudrilus eugeniae **6**:557
Eugenia **3**:130*T*, **5**:489, **6**:316*T*
Eugenia stipitata **4**:289*T*
eugenol **5**:712*T*
Euglandina rosea **3**:222, **4**:359, **4**:361, **5**:382, **5**:434, **6**:455
Euglena **6**:293
Euglena mutabilis **3**:584
euglenids **6**:293–295, **6**:294*F*
euglenoids **6**:25
Euglenophyceae **6**:478*T*
Euglenophyta **3**:571*T*, **6**:10*T*
Euglypha **6**:290–291
Euhamitermes **4**:409*T*
Euhrychiopsis lecontei **4**:376
Eukarya/eukaryotes **3:329–332**
background and characteristics **3**:329
bacterial–eukaryotic symbiosis **5**:642
biodiversity **5**:255
characteristics **5**:305–306
as chimera **3**:329–330
common ancestry theory **3**:397*F*
domain Eukarya **5**:288–289, **5**:289*F*
endosymbiotic theory
challenges **3**:332
evidentiary research **3**:331–332
flagella **3**:332
mitochondria and plastids **3**:330–331
peroxisomes **3**:332
secondary endocytosis **3**:332
evolutionary development **2**:641, **5**:703–705
genes/genetic code **3**:657–659, **3**:658*F*, **6**:360–361, **6**:361*T*
marine ecosystems **5**:47–49
mutation rates **2**:628*F*
mutualistic partnerships **2**:321, **2**:328
number of species **5**:254*T*, **6**:585–586, **6**:586*T*
oceanic food web **5**:49*F*
origin theories **3**:330, **3**:331*F*
phagocytosis **5**:305
phagotrophic protists **5**:306
phototrophs **5**:305–306
phylogenetic relationships **1**:224*F*, **1**:310*F*, **5**:253*F*, **5**:285*F*, **5**:287*F*, **5**:288–289, **5**:289*F*, **6**:584*F*, **7**:162*F*
soil systems
fungal diversity **6**:585–586
macrofauna
earthworms **6**:588
termites **6**:588
mesofauna
collembolans **6**:586
mites **6**:587
nematodes **6**:586
microfauna **6**:586, **6**:587*F*
subterranean ecosystems **7**:53
symbiotic relationships **5**:306
temperature tolerance ranges **4**:113–115
Eukoenenia draco **1**:212
Eukoenenia janetscheki **1**:212
Eukoeneniidae **1**:212
Eulamaops **4**:727*T*
Euleia heracleii **5**:469–470
Eulemur cinereiceps **6**:256–258*T*
Eulemur flavifrons **6**:256–258*T*
Eulemur fulvus **6**:251
Eulemur mongoz **6**:251
Eulemur sanfordi **6**:256–258*T*
Eulepethidae **7**:417*F*
Eulepidoptera **5**:386, **5**:386*F*, **5**:387–388*T*
Eumastacidae (monkey grasshoppers) **3**:733, **3**:733*F*
Eumastacoidea
characteristics **3**:732–733, **3**:733*F*
Eumastacidae (monkey grasshoppers) **3**:733, **3**:733*F*
Proscopiidae (false walking sticks or proscopids) **3**:733
Eumetopias jubatus **5**:447
Eumetopius jubatus **5**:81*T*, **5**:82–83
Eumops auripendulus **2**:587*F*
Eunicida **7**:417*F*
Eunicidae **7**:417*F*
Eunotia exigua **3**:584
Eunotia pectinalis **3**:358
Euonymus **1**:271, **3**:522*T*
Euonymus pauciflora **1**:271–272
Euoplocephalus **2**:519*T*, **2**:520*T*
Euornithopoda **2**:520*T*
Euoticus pallidus pallidus **6**:256–258*T*
Euowenia grata **5**:364*F*
Eupenicillium **5**:555–556
Euphausia **6**:25
Euphausia americana **4**:386*T*
Euphausia brevis **4**:386*T*
Euphausiacea **2**:416–417*T*
Euphausia crystallorophias **4**:386*T*
Euphausia superba **4**:385
Euphausidae **4**:386*T*
Euphorbia **1**:69, **2**:371*T*
Euphorbiaceae
abundance rankings **3**:129*T*
biogeography **2**:371*T*
diversification **5**:623*F*
mangrove ecosystems **5**:11*T*
number of species **2**:374*T*
seasonally dry tropical forests **5**:489
Succulent Karoo **1**:486*T*
taxonomic distributions **3**:522*T*
tropical rain forests **1**:263, **6**:313–318, **6**:316*T*
Euphorbia esula **1**:76
Euphorbia fendleri **2**:373*T*
Euphorbia guerichiana **1**:67
Euphorbia heterophylla **5**:607
Euphrasia **4**:71
Euphrosinidae **7**:417*F*
Euphydryas aurinia **1**:709–710
Euphydryas editha **4**:257–258, **4**:259*F*
Eupithecia **5**:593
Euplectus fauveli **7**:220*T*
Euplotes **6**:295
Eupnoi **1**:210
Eupodotis vigorsii **1**:487–488
Eupomatia **5**:618*F*
Euproctis chrysorrhoea **5**:391*F*, **5**:396
Euprymna scolopes **5**:122*F*, **5**:123–124, **5**:281*F*
Euprymnia scolopoides **2**:323–324
Euptelaceae **1**:260*T*
Eurarchontoglires **4**:685
Eurasia
C_4 plants **2**:371*T*
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*, **4**:729–730, **4**:729*F*
mammalian extinctions **4**:705*T*
megaherbivores **5**:224*T*, **5**:225*T*, **5**:232*F*
myriapods **5**:427*F*, **5**:429*T*
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
Europe
agricultural intensification **4**:565–566
boreal forests **1**:630*T*, **7**:212–213, **7**:212*T*
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
C_4 plants **2**:371*T*
countryside biogeography studies **2**:353–354
deforestation **1**:677*F*, **1**:679*T*, **2**:448*T*
domesticated ruminants **1**:702–703
ecosystems **3:333–346**
alpine life zone **1**:149*T*
biological and cultural diversity **3**:333
conservation status **3**:171–173*T*
ecosystem diversity
general discussion **3**:333–334

Europe (*continued*)
vegetation classification **3**:333–334
forest ecosystems **3**:532*T*
freshwater ecosystems **3**:578–579, **3**:579*T*
habitat and vegetation diversity
Arctic desert biome **3**:335*T*
biotic interactions **3**:334*T*, **3**:335
boreal forest biome **3**:335*T*, **3**:336–340*T*
continental semidesert biome **3**:335*T*, **3**:336–340*T*
habitat diversity **3**:334–335, **3**:334*T*
human interference **3**:334*T*, **3**:335
Mediterranean biome **3**:335*T*, **3**:336–340*T*
pattern variations **3**:335
prerequisites and sources **3**:334–335, **3**:334*T*
steppe/grassland biome **3**:335*T*, **3**:336–340*T*
taxon diversity **3**:334, **3**:334*T*
temperate deciduous forest biome **3**:335*T*, **3**:336–340*T*
tundra biome **3**:335*T*, **3**:336–340*T*
warm-temperature evergreen forest biome **3**:335*T*, **3**:336–340*T*
warm-temperature semidesert biome **3**:335*T*, **3**:336–340*T*
zonation patterns **3**:335–340, **3**:336–340*T*
pattern variations
Arctic desert biome **3**:335*T*
azonal vegetation **3**:336–340*T*
biomes **3**:335, **3**:335*T*
boreal forest biome **3**:335*T*, **3**:336–340*T*
boreal/nemoral orobiomes **3**:336–340*T*
continental semidesert biome **3**:335*T*, **3**:336–340*T*
habitat diversity **3**:344
Mediterranean biome **3**:335*T*, **3**:336–340*T*
Mediterranean orobiomes **3**:336–340*T*
research literature and surveys **3**:341–344*T*
semidesert orobiomes **3**:336–340*T*
steppe/grassland biome **3**:335*T*, **3**:336–340*T*
synanthropic vegetation **3**:336–340*T*
temperate deciduous forest biome **3**:335*T*, **3**:336–340*T*
tundra biome **3**:335*T*, **3**:336–340*T*
vegetation diversity **3**:335–340
vegetation maps **3**:340–344, **3**:345*F*, **3**:346*T*
warm-temperature evergreen forest biome **3**:335*T*, **3**:336–340*T*
warm-temperature semidesert biome **3**:335*T*, **3**:336–340*T*
zonation patterns **3**:335–340, **3**:336–340*T*
research outlook **3**:344–346
river water composition **3**:562*T*
species declines **3**:578–579, **3**:579*T*
threatened species **3**:579–580
vegetation classification
background information **3**:333–334
dominance approach **3**:334
floristic–sociological approach **3**:334
habitat classification **3**:334
energy use **3**:252*T*
freshwater ecosystems **4**:370*T*
habitat–bird species diversity relationship **1**:613–614
historical conservation practices **2**:279–280
invasive species
Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346
Fallopia japonica **4**:351–352
Spartina alterniflora **4**:349–350
lake and pond ecosystems **4**:461*T*
land-use trends **4**:559*F*
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*
maize production **1**:86
megaherbivores **5**:224*T*, **5**:225*T*
native crops **2**:662–663
natural history museums and institutions **5**:405
plant invasions **6**:93, **6**:99*F*
protected area systems **6**:491*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
temperate forests
distribution patterns **7**:116–117, **7**:116*T*
diversity patterns **7**:119–121, **7**:119*T*, **7**:120*F*
dominant floristic features **7**:118
endemism **7**:121
physiognomic characterization **7**:117
threatened plant species **6**:84–85
threatened species **1**:622–623
timber industry **7**:212–213, **7**:212*T*
transboundary protected areas **6**:489*T*
wetland ecosystems **7**:386
wheat production **1**:86, **1**:87–88
European Bank for Reconstruction and Development **7**:207–208
European BIODEPTH project **4**:586–587
European Cooperative Programme for Crop Genetic Resources (ECP/GR) **3**:643–644
European hare **3**:127–128, **3**:127*T*, **7**:28, **7**:28*F*
European Monitoring and Evaluation Program (EMEP) **5**:533
European rabbit **1**:79, **1**:80, **1**:81, **2**:551–552*T*, **4**:361, **7**:262–263
European starlings **1**:78, **1**:81, **2**:551–552*T*, **5**:312
European stoat **5**:370
European tree sparrows **4**:259
European Union **2**:314, **4**:125
European Union Biodiversity Action Plan **2**:314
European Union Habitats and Birds Directive **6**:487
European Union Water Protection Directive (WPD) **1**:95–97, **6**:497
European wild cats **3**:191
Eurosta solidaginis **4**:639–640, **4**:640*F*
Eurotia ceratoides **1**:270
Euryalaceae **1**:260*T*
Eurychoromyiidae **3**:489–491*T*
Eurynorhynchus pygmeus **7**:185
eurypterids **1**:204–205, **2**:400
Eurypygidae **6**:308*T*
Eurytemora **6**:26
Eurytermes **4**:409*T*
Eurytides protesilaus **1**:651*F*
Euscorpiidae **1**:209
Euscorpius flavicaudis **5**:425
Euscorpius germanus **1**:208
eusocial insects
Hymenoptera **4**:181–182
Isoptera **4**:412–413, **4**:412*F*, **4**:414*T*, **4**:421–422
Eutamias amoenus **7**:36
Eutatus **4**:727*T*
Euteleostei
Ateleopodomorpha **3**:465
Cyclosquamata **3**:465–466
Lampriomorpha **3**:466, **3**:466*F*
Paracanthopterygii **3**:466–467
Polymixiomorpha **3**:466
Protacanthopterygii **3**:464–465, **3**:465*F*
Scopelomorpha **3**:466
Stenopterygii **3**:465
Eutermellus **4**:409*T*
Euterpe oleracea **1**:164, **4**:286–287, **4**:288*T*, **4**:290
Euterpe purpurea **1**:166–167
Eutheria **4**:683
see also Placentalia
eutherians **3**:188
Eutreptia **6**:30
eutrophication **3:347–371**
biodiversity impacts **5**:522
biofuel crop production **1**:530
freshwater ecosystems
agricultural and industrial chemicals **3**:580–581, **3**:581*F*
aquatic vascular plants **3**:359–360
biological consequences **3**:581–582
macroalgae **3**:358
macrophytes **3**:359–360
microfauna **3**:361–363
organic pollution **3**:580–581, **3**:581*F*
trophic levels **3**:563*T*
historical perspective
dissolved oxygen **3**:351
nutrient loading **3**:348–351
phosphorus and nitrogen inputs **3**:349*F*
population density–soluble reactive phosphorus (SRP) relationship **3**:348*F*
total phosphorus/total nitrogen–chlorophyll *a* relationship **3**:348–350, **3**:349*F*
water-column nitrogen:phosphorus (N:P) ratios **3**:350–351, **3**:350*F*
lake and pond ecosystems

agricultural and industrial chemicals **3**:580–581, **3**:581*F*
biological consequences **3**:581–582
acidification **3**:584–585, **3**:584*F*
comparison studies **3**:582
ecosystem functions and processes **3**:585
food web manipulation **3**:582–583
Lake Fure, Denmark **3**:575*F*, **3**:581–582
resource and predatory controls **3**:582–583
Secchi transparency **3**:581–582
species richness **3**:583–584
whole-lake studies **3**:582
clean-up results **5**:24
organic pollution **3**:580–581, **3**:581*F*
pollution **4**:465–466
species diversity–lake functioning **4**:462–463
low-light environments
aquatic vascular plants
estuarine ecosystems **3**:360–361
freshwater ecosystems **3**:359–360
marine ecosystems **3**:360–361
inorganic nitrogen toxicity **3**:366–367
invertebrate macrofauna **3**:364
macroalgae
estuarine ecosystems **3**:358–359
freshwater ecosystems **3**:358
marine ecosystems **3**:358–359
macrophytes
estuarine ecosystems **3**:360–361
freshwater ecosystems **3**:359–360
marine ecosystems **3**:360–361, **3**:360*F*, **3**:361*F*
microbial heterotrophs **3**:358
microfauna
estuarine ecosystems **3**:363–364, **3**:363*F*
freshwater ecosystems **3**:361–363
marine ecosystems **3**:363–364, **3**:363*F*
vertebrate macrofauna **3**:364–366, **3**:365*F*
marine ecosystems **5:23–31**
anthropogenic effects **5**:23–24, **5**:53–54
aquatic vascular plants **3**:360–361
contaminant–stress relationship **5**:29–30
food web complexity **5**:24–25
harmful algal blooms **5**:25
lake and pond clean-up effects **5**:24
macroalgae **3**:358–359
macrophytes **3**:360–361, **3**:360*F*, **3**:361*F*
microfauna **3**:363–364, **3**:363*F*
nearshore ocean nutrient enrichment response
harmful algal blooms **5**:25
high nutrient, low chlorophyll (HNLC) concept **5**:26, **5**:26*T*
high nutrient, low growth (HNLG) concept **5**:26–27, **5**:26*T*, **5**:27*F*
nitrogen loading **5**:25, **5**:26*F*
nutrient loading **5**:23–24
species shifts **3**:353*F*
stoichiometric effects
nutrient ratio changes **5**:28–29, **5**:28*T*, **5**:29*F*
Redfield ratios **5**:27
silicon additions **5**:27–28
marine sedimentary habitats **5**:113*F*, **5**:114
microalgae
harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*
long-term human influences **3**:354–356, **3**:355*F*, **3**:356*F*, **3**:357*F*
species shifts **3**:352–354, **3**:353*F*
nitrogen cycle **1**:95–97, **1**:96*F*, **5**:538–540, **5**:539*F*, **5**:542–544
nutrient enrichment impacts **1**:95–97, **1**:96*F*, **3**:351–352, **3**:352*F*, **3**:575*F*, **3**:576–577, **3**:576*F*
oligotrophication **3**:367–369
reversal programs **3**:367–369
structural stability **7**:9
total organic carbon (TOC) content **3**:351, **3**:352*F*
Euwenia **4**:727–728
Euxoa nigricans **6**:159*F*
Evanioidea **4**:178
evergreen broad-leaved forests **1**:270–271, **3**:428*T*, **6**:120*T*
evergreen forests
global agriculture impacts **3**:428*T*
litter deposition **6**:120*T*
net primary production (NPP) **1**:560*T*
protected area systems **5**:444*T*
termites **4**:432*T*
Evolutionarily Distinct and Globally Endangered (EDGE) species **7**:202–203
evolutionarily stable strategies (ESS) **4**:632–633
evolutionary biology
adaptation **1:15–20**
adaptationism
basic concepts **1**:18–19
criticisms **1**:19
post-adaptationism research **1**:19–20
current research areas **1**:19–20
definitions
evolution-based adaptations **1**:16–17
general discussion **1**:16
process-based adaptations **1**:16
evidential perspectives
general discussion **1**:17–18
genomics **1**:18
increased fitness **1**:18
natural selection **1**:17–18
historical perspectives
biblical theories **1**:15
classical era **1**:15
Darwinian theory **1**:15–16
pre-Darwinian theories **1**:15
evolutionary faunas **3**:423–424
evolutionary significant unit (ESU) **1:668–673**
basic concepts **1**:668–669
ecological exchangeability **1**:668–669
evolutionary lineage versus ecological context approach **1**:670–672, **1**:670*F*
hatchery fish **1**:670–672, **1**:670*F*
mammals **4**:704
Pacific salmon **1**:669–670, **1**:670*F*
research summary **1**:672
Evolution Canyon testing microsite **3**:670–671, **3**:673*F*, **3**:674*F*
evolution, theory of **3:392–400**
adaptive processes **1**:16–17
Darwin's five theories
common ancestry theory **3**:395, **3**:397*F*
evolutionary theory **3**:395–396
gradualism **3**:395–396
natural selection **3**:396
populational change **3**:396
evolutionary ecology **3**:7
evolutionary mechanisms
genetic drift **3**:398–399
migration **3**:398
mutations **3**:398
natural selection **3**:399
speciation **3**:399
variation **3**:396–398
historical perspective
ancestor–descendant relationships **3**:393*F*
Darwinian theory **1**:15–16, **3**:393–394
evolutionary developmental biology **3**:395, **3**:396*F*
modern synthesis **3**:394
molecular evolution **3**:394–395
postsynthesis theories **3**:394–395
pre-Darwinian theories **3**:393, **3**:393*F*
Proterozoic–Cambrian transition **3**:396*F*
organic evolution **3**:392–393
scientific theory **3**:392
Exclusive Economic Zones (EEZs) **6**:421, **6**:421–422
Excoecaria **5**:11*T*
Excoecaria agallocha **1**:263–266
Exellodendron barbatum **1**:163
ex situ conservation
agrobiodiversity conversation **3**:33
botanical gardens **6**:513, **6**:513–514, **6**:514, **6**:516
botanical/zoological garden conservation **4**:316*F*, **4**:319, **4**:321*T*
conservation genetics **2**:273–274
field gene banks **3**:33, **4**:316*F*, **4**:319, **4**:321*T*
in vitro conservation **4**:316*F*, **4**:319
livestock park conservation **4**:316*F*, **4**:319, **4**:321*T*
maintenance strategies **1**:406
plant biodiversity **6**:63
pollen/semen/ovule/DNA conservation **4**:316*F*, **4**:319–320, **4**:321*T*
research programs and methodologies **4**:318–319
seed/embryo collection and storage **4**:316*F*, **4**:319, **4**:319*F*, **4**:321*T*
Exoporia **5**:386, **5**:386*F*, **5**:387–388*T*
exotic C_4 plants **2**:379–381, **2**:380*F*
exotic species invasions **1:75–84**
amphibian species decline **3**:695–696
aquaculture **1**:194–196

exotic species invasions (*continued*)
Ascension Island **6**:78
Australia
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:81
crop plant pathogens **1**:76*T*, **1**:81
livestock pathogens **1**:76*T*, **1**:81
nonindigenous plants **1**:76*T*, **1**:80
nonindigenous vertebrates
amphibians and reptiles **1**:80
birds **1**:80
mammals **1**:76*T*, **1**:80
biodiversity impacts **7**:203*F*
Brazil
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:83
crop plant pathogens **1**:76*T*, **1**:83
nonindigenous plants **1**:76*T*, **1**:82–83
nonindigenous vertebrates
birds **1**:83
mammals **1**:76*T*
climate change effects **2**:104
contemporary conservation theories and practices **2**:254–255
dispersal biogeography **2**:554–555, **2**:554*F*
economic impact **1**:75
endangered reptiles **3**:217–218
fire/fire regimes **3**:440
forest ecosystems **3**:534
forest management impact **1**:451
global agriculture impacts **3**:429
globalization effects **3**:701–703, **3**:702*F*
grazer-based ecosystem restoration and management **4**:16–17
impacts and distribution **4:357–368**
basic concepts **4**:357
biodiversity threats **2**:80
competition
general discussion **4**:360
interference competition **4**:360
resource competition **4**:360
continental ecosystems **4**:155
coral reef ecosystems **2**:341
distribution patterns
aquatic habitats **3**:411
biodiversity loss **4**:663
characteristics **3**:410–411
island biogeography **3**:217–218, **3**:410–411, **3**:410*F*, **6**:83
endangered mammals **3**:193
energy flow **3**:246–247
environmental impacts **3**:284
fish species **3**:448
freshwater invertebrates **4**:378
geographic distribution **4**:358
habitat distribution **4**:358–359
habitat modification
general discussion **4**:359
grazing and rooting **4**:359
modified fire regimes **4**:360
modified hydrology **4**:360
modified nutrient regime **4**:360
shading and overgrowth **4**:359–360
herbivory **4**:361
human impacts **4**:143–144
hybridization and introgression **4**:362
impact quantification **4**:366–367
indirect effects
chain reactions **4**:363–364
characteristics **4**:362–363
classical indirect effects **4**:363
pathogen vectors **4**:364
invasional meltdowns
basic concepts **4**:364
introduced animals **4**:364–365
introduced plants **4**:364
mutualism **4**:365
pollination and dispersal **4**:364
island biogeography **4**:155
island vulnerability **4**:359
modern extinctions **5**:370
neotropical mammals **2**:590
onshore ocean waters **4**:155–156
parasitism and disease **4**:361–362
pathogens **4**:156
plant–soil interactions **6**:125
predation **4**:360–361, **6**:240–241
Southeast Asia **1**:518–520, **1**:521, **1**:522*F*, **4**:346–347
terrestrial invertebrates **3**:222
time lag–evolution relationship **4**:365–366
timing factors **4**:358
wetland ecosystems **7**:399
India
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:82
crop plant pathogens **1**:76*T*, **1**:82
livestock pathogens **1**:76*T*, **1**:82
molluscs **1**:82
nonindigenous plants **1**:76*T*, **1**:82
nonindigenous vertebrates
birds **1**:82
mammals **1**:76*T*
invasive species–pathogen relationship **2**:529–530
managed relocations **7**:238–239, **7**:239*F*
marine sedimentary habitats **5**:113*F*, **5**:114
plant invasions **6:90–102**
basic concepts **6**:90–92, **6**:91*F*
current extent **6**:93–95, **6**:94*F*
dispersal pathways **6**:90–92, **6**:91*F*
general discussion **6**:90
historical perspective **6**:92–93
invasibility **6**:97–98, **6**:99*F*
invasion process
dispersal **6**:95–96
establishment and naturalization **6**:96
spread rates **6**:96–97
stages **6**:95–96, **6**:95*F*
invasiveness **6**:97
management strategies
biotic homogenization **6**:100–101*F*, **6**:101
ecological management **6**:101
potential conflicts **6**:101–102
prevention methods **6**:101
removal methods **6**:101
modeling approaches **6**:99–101
plant trait correlations **6**:98–99
population diversity **6**:172–173
range ecology **6**:341
research summary **1**:83
river ecosystems **6**:481*T*, **6**:482
South Africa
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:81
crop plant pathogens **1**:76*T*, **1**:81–82
livestock pathogens **1**:76*T*, **1**:82
nonindigenous plants **1**:76*T*, **1**:81
nonindigenous vertebrates
birds **1**:81
mammals **1**:76*T*, **1**:81
taxonomic versus functional diversity **2**:655
threatened birds **7**:184
United Kingdom
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:79–80
crop plant pathogens **1**:76*T*, **1**:80
nonindigenous plants **1**:76*T*, **1**:79
nonindigenous vertebrates
birds **1**:79
mammals **1**:76*T*, **1**:79
United States
background information **1**:75–76
nonindigenous invertebrates
earthworms **1**:78
general discussion **1**:78
insects and mites **1**:76*T*, **1**:78
microbes **1**:76*T*, **1**:78–79
nonindigenous plants
crop plants **1**:75–76
economic impact **1**:76, **1**:76*T*
nonindigenous vertebrates
amphibians and reptiles **1**:78
exotic birds **1**:76*T*, **1**:77–78
mammals **1**:76–77, **1**:76*T*
expanded perception **3**:635*T*
expansionist worldview **2**:702
exploitation ecosystem hypothesis **3**:503, **3**:508–509
exploitative competition **6**:716–717
eXtensible Markup Language (XML) **1**:442–443
externalities **2**:154
extinctions
anthropogenic impacts **2**:278–279
biodiversity loss **1**:404–405, **2**:74–75, **4**:664
birds
conservation approaches **1**:623
current threats **1**:622–623, **1**:622*F*
global and regional trends **7**:182
historical causes **1**:620–621, **1**:620*T*
human colonization **1**:621–622
natural causes **1**:620
carnivores **1**:693–694
causal factors **3:401–412**
anthropogenic activities and impacts
biodiversity modifications **4**:139–140
biotic environment changes **4**:662–664, **4**:663
deforestation **4**:663
habitat loss and fragmentation **3**:401–402, **3**:401*F*, **4**:662, **4**:663

large-scale development projects 3:402–403
overexploitation 3:401–402, 3:401*F*, 3:402*F*, 4:662
persecution 4:662
unequal natural resource use 3:402
world population growth 3:401–402, 3:402*F*
background extinctions 5:2–3
biotic environment changes 4:662–664, 4:663
deforestation 4:663
disease 1:519–520, 3:411, 3:411*F*, 4:664
distal and proximal causal factors 5:3–4
ecological perspectives 4:664–666
exotic species invasions
aquatic habitats 3:411
biodiversity loss 4:663
characteristics 3:410–411
island biogeography 3:410–411, 3:410*F*
global climate change **2:73–78**
background information 3:407–408
biodiversity loss 2:74–75, 4:664
climate change predictions 3:408, 3:408*F*
communities and ecosystems 2:76
current trends 2:73
extrinsic risks 2:74, 2:75*F*
geographic and landscape scale variation 2:74–75
global impact 2:73
habitat loss and fragmentation 3:408
individuals and populations 2:75–76, 2:75*F*
intrinsic risks 2:73–74, 2:74*F*
research summary 2:77
species interaction risks 2:76–77
temporal match–mismatch (TMM) hypothesis 2:76
habitat conversion 4:664–666, 4:665*T*
habitat destruction
anthropogenic impacts 4:662
coral reef ecosystems 3:404, 3:405*F*
desertification 3:404–405
general discussion 3:403
grasslands 3:404
habitat loss 3:403
mangrove ecosystems 3:404
rain forest loss 3:403–404, 3:403*F*
tropical deciduous forests 3:404
wetlands 3:404
habitat fragmentation
anthropogenic impacts 4:662, 4:663
characteristics 3:405–406
dispersal barriers 3:405–406
edge effects 3:406
interspecific interactions 3:406
tropical forests 3:405–406, 3:405*F*
island biogeography 3:410–411, 3:410*F*
mass extinctions 5:2–3
multiple factors 3:411–412
overexploitation
anthropogenic activities and impacts 3:401–402, 3:401*F*, 3:402*F*
biodiversity loss 4:662
global wildlife trade 3:409–410, 3:409*T*
modern societies 3:409–410
overfishing 3:410
traditional societies 3:409
pollution-based habitat degradation
acid rain 3:407, 3:407*F*
air pollution 3:407
characteristics 3:406
lichen tolerance 3:407
ozone depletion 3:407
pesticides 3:406
smog 3:407
ultraviolet radiation 3:407
water pollution 3:406–407
pseudoextinctions 5:2
temporal and biogeographic scales 5:3
climate change **2:73–78**
background information 3:407–408
biodiversity loss
communities and ecosystems 2:76
geographic and landscape scale variation 2:74–75
global warming 4:664
individuals and populations 2:75–76, 2:75*F*
species interaction risks 2:76–77
temporal match–mismatch (TMM) hypothesis 2:76
climate change predictions 2:94–95, 3:408, 3:408*F*
current trends 2:73
global impact 2:73
habitat loss and fragmentation 3:408
research summary 2:77
risk factors
extrinsic risks 2:74, 2:75*F*
intrinsic risks 2:73–74, 2:74*F*
crustaceans 2:413–415
definition 2:47
dinosaurs **2:516–522**
Cretaceous–Tertiary (K–T) mass extinction 1:494, 2:516, 3:420, 5:6–7, 5:173–175
differential patterns of survival 2:516, 2:517*F*
long-term diversity trends 2:517–518, 2:519*T*, 2:520*T*
popular accounts 2:516
research summary 2:521–522
sea level changes 2:517
short-term diversity trends 2:518–519
taxonomic richness 2:519–521, 2:521*F*
thermal radiation scenario 2:516–517
dispersal biogeography 2:558–559
extinction debt 4:497
extinction mechanisms
general discussion 4:659
isolated populations
Allee effect 4:659
demographic stochasticity 4:659–660, 4:661*F*
environmental stochasticity 4:660, 4:661*F*
genetic bottlenecks 4:660, 4:661*F*
low population densities 4:659
population size changes 4:660, 4:661*F*
metapopulation theory 4:660–661
extinction rates
biodiversity loss 1:404–405, 4:656, 4:656*T*, 4:657*F*
climate change effects 2:86–87, 4:664
coevolution 2:147
comparison studies **2:167–176**
extinction rate versus extinction magnitude 2:173–174, 2:174*F*
functional role 2:167
future extinctions 2:171
metrics 2:167–168
past, recent, and future extinction rates 2:168*F*
present extinctions 2:169–170
research challenges 2:174–175
extinction rate declines 3:423–424
fossil record 3:413–414
future extinctions
estimation methods 4:658–659, 4:659*F*, 4:660
extinction rate versus extinction magnitude 2:173–174, 2:174*F*
global scale studies 2:171–173, 2:172*T*
projected extinctions 2:171, 2:172*T*
uncertainty factors 2:173
inbreeding
captive populations 4:249, 4:249*F*
wild populations 4:249–250
mammals 4:704–705, 4:705*T*
metrics
E/MSY units 2:168–169, 2:168*T*
magnitude measures 2:167–168, 2:168*T*
projected extinction 2:169
realized extinction 2:169
natural extinctions 5:433
present extinctions
estimation methods 4:658, 4:659*T*
extinction patterns 2:170–171, 2:170*F*
human-induced extinctions 2:169–170, 2:170*F*
Red List Index (RLI) 2:169–170, 2:170*F*
recent extinction rates 4:658, 4:659*T*
speciation–extinction rate relationship 3:415, 5:439–440, 5:440*F*
species richness 2:651–652, 2:652*F*
sustainability trends 7:82*F*
extinction risks 4:664, 7:192–194, 7:193
extinct megafauna 4:591–593, 4:593*F*
fish species 3:451–452
forest canopies 3:524
fossil record **3:413–425**
assessment challenges
accurate range estimates 3:415
biased preservation 3:414–415
data collection 3:414
phylogenetic relationships 3:415
preservation unevenness 3:415
pseudoextinction 3:415
speciation–extinction rate relationship 3:415

extinctions (*continued*)
taxon-level bias **3**:414
time measures **3**:413–414
background extinctions
evolutionary faunas **3**:423–424
extinction rate declines **3**:423–424
mass extinctions **5**:2–3
turnover **3**:423
crustaceans **2**:413–415
diversification impacts **3**:424–425, **3**:424*F*, **3**:425*F*
mass extinctions
basic concepts **3**:415–418
biodiversity loss **4**:656
biogeographic impacts **3**:421–422
causal factors **3**:418
climatic changes **3**:422–423
extinction theories **3**:418–421
long-term consequences **3**:421
major mass extinctions **3**:416*T*
periodicity **3**:422–423, **3**:424*F*
quantification **3**:417
research background **3**:413
global biodiversity impacts **3**:689–690
greenhouse effect impacts **4**:29
habitat loss and fragmentation **4**:54, **4**:662, **4**:663
imminent extinction probability
marine mammals
beluga whales **5**:77*T*, **5**:79–80
blue whales **5**:81*T*, **5**:82
California sea otter **5**:81*T*, **5**:83–84
general discussion **5**:76
Mediterranean monk seals **5**:77*T*, **5**:80
Northern sea otter **5**:81*T*, **5**:84–86
right whales **5**:76–78, **5**:77*T*
species and subspecies **5**:77*T*
Steller sea lions **5**:81*T*, **5**:82–83
vaquitas **5**:77*T*, **5**:80
western North Pacific gray whales **5**:77*T*, **5**:78–79
West Indian manatee **5**:81*T*, **5**:83
protected area systems **6**:488*F*
late Quaternary extinction event (LQEE) **4:721–732**
background information **4**:721
causal factors
anthropogenic models **4**:731–732
climatic models **4**:732
global patterns **4**:731–732
comparative studies
Africa **4**:658*F*, **4**:729–730
analytical methods **4**:721–725
Australia **4**:658*F*, **4**:727–728
Eurasia **4**:729–730, **4**:729*F*
Madagascar **4**:658*F*, **4**:728–729
North America **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
Oceania **4**:730–731
South America **4**:726–727, **4**:727*T*
time-transgressive sequence **4**:722*F*, **4**:723*F*
mammals
extinction rates **4**:704–705, **4**:705*T*
future extinction estimations **4**:706–707
geographic patterns **4**:705–706, **4**:705*T*
neotropical mammals **2**:589–592
Pleistocene
major die-offs **3**:189
North America **4**:722–725, **4**:724–725*T*, **4**:727*F*
pre-Quaternary mammals **5:1–9**
Age of the Dinosaurs **5**:5–6
background extinctions **5**:2–3
causal factors **5**:2
Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
distal and proximal causal factors **5**:3–4
early Paleocene faunal recovery **5**:7
geologic timescale **5**:1*T*
late Triassic faunal turnover **5**:4–5
mass extinctions **5**:2–3
Paleocene–Eocene thermal maximum **5**:7–8
Permian–Triassic mass extinction **5**:4
phylogenetic relationships **5**:1*F*
pseudoextinctions **5**:2
summary discussion **5**:8
synapsids **5**:1–2, **5**:4
temporal and biogeographic scales **5**:3
Tertiary faunal evolution **5**:7–8
therapsid radiation **5**:4–5
recent extinctions **3**:189–190, **3**:189*F*
marine invertebrates **3**:202*T*, **4**:664
marine mammals **5:64–93**
cetaceans
characteristics **5**:65*T*, **5**:66–67
evolutionary history **5**:68–69
taxonomy **5**:65*T*
climate change effects **5**:89–91
contributing factors
body size and dimensions **5**:70
fishery–diet/habitat overlap **5**:72
hauling behavior predictability **5**:70–71
rapid recovery potential **5**:71
sea surface-based respiration **5**:70
spatiotemporal distribution predictability **5**:70
subsistence and commercial market demand **5**:71
toxin and contaminant bioaccumulation potential **5**:71–72
desmostylians
characteristics **5**:68
evolutionary history **5**:69–70
extinction/near extinction over evolutionary time **5**:75
geographic regions **5**:92
global change effects **4**:664
identification/monitoring-related hindering factors
availability bias **5**:72
environmental uncertainty effects **5**:73–74
lack of consistent objective criteria **5**:74
observer bias **5**:72–73
population size survey data **5**:73
survey costs **5**:74
vital demographic parameters **5**:73
imminent extinction probability
beluga whales **5**:77*T*, **5**:79–80
blue whales **5**:81*T*, **5**:82
California sea otter **5**:81*T*, **5**:83–84
general discussion **5**:76
Mediterranean monk seals **5**:77*T*, **5**:80
Northern sea otter **5**:81*T*, **5**:84–86
right whales **5**:76–78, **5**:77*T*
species and subspecies **5**:77*T*
Steller sea lions **5**:81*T*, **5**:82–83
vaquitas **5**:77*T*, **5**:80
western North Pacific gray whales **5**:77*T*, **5**:78–79
West Indian manatee **5**:81*T*, **5**:83
major anthropogenic extinctions
North Atlantic gray whale (*Eschrichtius robustus*) **5**:76
species level **5**:75
subspecies level **5**:76
marine otters
characteristics **5**:68
evolutionary history **5**:70
minimization strategies **5**:92–93
natural versus unnatural extinctions **5**:74
patterns and case studies
desmostylians **5**:75
extinction/near extinction over evolutionary time **5**:74–75
general discussion **5**:74–75
major anthropogenic extinctions **5**:75
Odobenidae **5**:74–75
sirenians **5**:75
pinnipeds
characteristics **5**:65*T*, **5**:67
evolutionary history **5**:69
taxonomy **5**:65*T*
polar bears
characteristics **5**:68
evolutionary history **5**:70
recovery evidence
bowhead whales **5**:86–87, **5**:87*T*
elephant seals **5**:87*T*, **5**:88
general discussion **5**:86
gray whales **5**:87–88, **5**:87*T*
humpback whales **5**:87, **5**:87*T*
species and subspecies **5**:87*T*
research summary **5**:91–92
scientific whaling **5**:88–89
significant extinction probability
general discussion **5**:80–82
species and subspecies **5**:81*T*
sirenians
characteristics **5**:65*T*, **5**:67–68
evolutionary history **5**:69
extinction/near extinction over evolutionary time **5**:75
taxonomy **5**:65*T*
species level extinctions
baiji (Yangtze River dolphin/*Lipotes vexillifer*) **5**:75

Caribbean monk seal (*Monachus tropicalis*) **5**:75
Japanese sea lion (*Zalophus japonicus*) **5**:75–76
Steller's sea cow (*Hydrodamalis gigas*) **5**:76
taxonomy
characteristics **5**:65–66, **5**:65*T*
definition **5**:64–66
habitat boundaries **5**:65–66
mass extinctions **5:156–166**
anthropogenic impacts **2**:278–279
basic concepts **3**:415–418
biogeographic impacts **3**:421–422
causal factors **3**:418, **5**:2–3
climatic changes **3**:422–423
crustaceans **2**:413–415
data analysis and interpretation **5**:157–158, **5**:158*F*, **5**:159*F*, **5**:160*F*
extinction theories **3**:418–421
first-contact extinctions
basic concepts **5**:365
Madagascar **5**:365
New Zealand **5**:366
Pacific islands **5**:365–366
fossil diversity curve **5**:157*F*
global biodiversity impacts **3**:689–690
great mass extinctions
biodiversity loss **4**:657*F*
end-Cretaceous **5**:168*F*, **5**:169*F*, **5**:173–175, **5**:174*F*
Late Devonian **5**:168*F*, **5**:169–170, **5**:169*F*
Late Ordovician **5**:167–169, **5**:168*F*, **5**:169*F*
Late Permian **5**:168*F*, **5**:169*F*, **5**:170–172
Late Triassic **5**:168*F*, **5**:169*F*, **5**:172–173
Phanerozoic Eon **1**:493, **1**:493*F*, **1**:494*T*
research challenges **5**:167–169
historical extinctions
extinction paradoxes **5**:367
Hawaii **5**:367
New Zealand **5**:366–367
number of species **5**:368*T*
historical research **5**:156–157, **5**:157*F*
kill curve **5**:160–163, **5**:163*F*
lesser mass extinctions
Cenozoic **5**:176
Mesozoic **5**:169*F*, **5**:176
Neoproterozoic **5**:168*F*, **5**:175–176
Paleozoic **5**:168*F*, **5**:175–176
significance **5**:175–176
long-term consequences **3**:421
magnitude measures
end-of-geologic-period extinctions **5**:158, **5**:161*F*
end-Permian mass extinction **5**:158, **5**:161*F*
extinction events **5**:158–160, **5**:161*F*
extinction intensities **5**:162*F*
general discussion **5**:158
major mass extinctions **3**:416*T*
modern diversity crisis **5**:165–166
modern extinctions **5:362–372**
Africa **5**:364–365
Australia **5**:363–364, **5**:364*F*
current on-going extinctions **5**:371
density-related traits **5**:370–371
extinction cascades **5**:367–368
first-contact extinctions **5**:365
geographical bias **5**:369
historical extinctions **5**:365, **5**:368*T*
inadequate documentation **5**:368–369
inherent biological limitations **5**:369
introduced species **5**:370
methodological limitations **5**:369
North America **5**:363
number of species **5**:368, **5**:368*T*
population dynamics **5**:370
postglacial extinctions **5**:363
research challenges **5**:368
species lifespans **5**:362–363, **5**:362*T*
taxonomic and habitat bias **5**:369
theoretical perspectives **5**:369–370
time factors **5**:371
uneven documentation **5**:369
notable mass extinctions **5:167–177**
background information **5**:167
great mass extinctions **5**:167–169
periodicity **3**:422–423, **3**:424*F*
periodicity hypothesis **5**:160, **5**:162*F*
Permian–Triassic mass extinction **5**:4
Phanerozoic Eon **1**:412–413, **1**:413*F*
quantification **3**:417
recoveries
characteristics **5**:164
delayed diversification **5**:164
disaster species **5**:164–165
large-scale rebounds **5**:164, **5**:165*F*
Lazarus species **5**:164–165
research summary **5**:166
scenarios
abrupt extinction **5**:157
general discussion **5**:157
gradual extinction **5**:157
stepwise extinction **5**:157
self-organized criticality **5**:160–163
temporal and biogeographic scales **5**:3
victim–survivor patterns **5**:163–164
megaherbivores **5**:225*T*, **5**:231–233, **5**:232*F*
modern extinctions **5:362–372**
crustaceans **2**:415–418, **2**:416–417*T*
current on-going extinctions **5**:371
extinction cascades **5**:367–368
first-contact extinctions
basic concepts **5**:365
Madagascar **5**:365
New Zealand **5**:366
Pacific islands **5**:365–366
historical extinctions
extinction paradoxes **5**:367
Hawaii **5**:367
New Zealand **5**:366–367
number of species **5**:368*T*
postglacial extinctions
Africa **5**:364–365
Australia **5**:363–364, **5**:364*F*
North America **5**:363
research challenges **5**:368
species lifespans **5**:362–363, **5**:362*T*
theoretical perspectives **5**:369–370
natural extinctions **5:433–441**
biodiversity trends **5**:439–440
causal factors
environmental changes **5**:433–434
species–species interactions **5**:434
geographic patterns **5**:439
prevalence **5**:433
selectivity factors
body size **5**:438
characteristics **5**:436
dispersal ability **5**:437–438
geographic range **5**:436
local abundance **5**:436
range–abundance relationships **5**:436–437, **5**:437*F*
rarity **5**:436
spatial structure/patchiness **5**:437
specialization **5**:438
speciation–extinction rate relationship **5**:439–440, **5**:440*F*
species duration
age factors **5**:435–436
taxonomic groups **5**:434–435, **5**:434*T*, **5**:435*F*
oceanic islands **5**:594, **5**:595–597
plants **3**:205–206, **6**:62, **6**:73*F*
population diversity **6**:172–173
pre-Quaternary mammal extinctions **5:1–9**
Age of the Dinosaurs **5**:5–6
causal factors
background extinctions **5**:2–3
distal and proximal causal factors **5**:3–4
mass extinctions **5**:2–3
pseudoextinctions **5**:2
temporal and biogeographic scales **5**:3
Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
early Paleocene faunal recovery **5**:7
geologic timescale **5**:1*T*
late Triassic faunal turnover **5**:4–5
Paleocene–Eocene thermal maximum **5**:7–8
Permian–Triassic mass extinction **5**:4
phylogenetic relationships **5**:1*F*
summary discussion **5**:8
synapsids **5**:1–2, **5**:4
Tertiary faunal evolution **5**:7–8
therapsid radiation **5**:4–5
sixth extinction wave **2**:351
extinct megafauna **4**:591–593, **4**:593*F*
extractivism **6**:608
extranuclear recombination **6**:355
extraterrestrial impacts **3**:418–421
extreme environments **4:111–120**
biodiversity **4**:118–119
characteristics **4**:111
climate change studies **2**:66, **2**:93–94
geothermally heated systems **4**:112–113
microbial activities **5**:304
nematodes **7**:429

extreme environments (*continued*)
research summary **4**:119–120
temperature tolerance ranges **4**:113–115
thermal ecosystems **4**:111–112
thermophilic evolution **4**:115–118
extreme events **4:209–223**
abrupt driver changes **4**:219–220
biodiversity impacts **7**:203*F*
climate change effects **2**:84–85
climate change studies **2**:66, **2**:93–94
defining criteria
intensity **4**:210, **4**:211*F*
rarity **4**:210
severity/magnitude **4**:210–211, **4**:211*F*
ecological implications **4**:220–221
ecological responses
cascading effects **4**:214–215
general discussion **4**:213
indirect consequences **4**:214
irreversible, slow, delayed, or hysteretic recovery **4**:213–214, **4**:215*F*
key structuring species loss **4**:214
microenvironmental changes **4**:214
response abruptness **4**:213, **4**:214*F*
response characteristics **4**:213–214
response magnitude **4**:214
response rarity **4**:213
El Niño Southern Oscillation (ENSO) **3**:155–156
emerging patterns and challenges
anthropogenic amplification **4**:217–219
compensatory mechanisms **4**:216–217
compound effects **4**:217
ecological memory **4**:217, **4**:218*F*
event frequency **4**:215–216
resistance mechanisms **4**:216–217
sequences of events **4**:217
species richness role **4**:216–217
environmental impacts **7**:203*F*
estuarine ecosystems **3**:302
event types **4**:211–212
global change effects **4**:219–220
large-scale drivers **4**:212–213
organism- and system-based responses **4**:211, **4**:212*F*
plant biodiversity threats **2**:124–125, **3**:124, **3**:125, **6**:62–63
research background **4**:209, **4**:210*F*
research methodologies **4**:219, **4**:220*T*
triggers **4**:212
extremeophiles **2**:640–641
Exxon Valdez oil spill **3**:256
eyes/photokeratitis **7**:302

F

F1 plant hybrids **1**:460–461
faba beans
intercropping patterns **2**:383, **2**:390
native crops **2**:662–663
regional cultivars **3**:32*F*
Fabaceae **1**:486*T*, **2**:427, **3**:129*T*, **4**:353–354, **5**:221, **5**:489, **5**:621–625, **5**:623*F*
facultative parasites **5**:657
Fagaceae **1**:266*T*, **1**:270–271, **3**:130*T*, **5**:621–625, **5**:623*F*
Fagopyrum esculentum **2**:663
Fagraea **3**:522*T*
Fagus **7**:118
Fagus crenata **1**:271
Fagus japonica **1**:271
Fagus sylvatica **1**:611–612, **2**:552*F*
Faidherbia albida **2**:390
fairy shrimp **3**:182*T*, **4**:371*T*
Falco aesalon **5**:474
Falco biarmicus **5**:474
Falco naumanni **5**:474
Falconidae **6**:308*T*
Falconiformes **6**:308*T*
falcons **2**:56*F*, **5**:474
Falco peregrinus **2**:245, **3**:192–193, **3**:406, **4**:258–259, **7**:188
Falco peregrinus brookei **5**:474
Falco punctatus **6**:455, **7**:188–189
Falco subbuteo **5**:474
Falco tinnunculus **5**:474
Falkland Islands **6**:507*T*
Fallopia japonica **4**:351–352, **4**:352*F*, **6**:107
fallow deer **5**:472
fallowing **6**:567
false elk **4**:724–725*T*
false mole crickets **3**:736, **3**:736*F*
false scorpions **7**:53
false walking sticks **3**:733
Famennian Stage **5**:169–170
Familia Sagrada Church (Barcelona, Spain) **1**:590–591
family planning methods **6**:202, **6**:202*F*
Fanniidae **3**:489–491*T*
Farallon Plate **5**:563
farming practices **6**:567
farmland biodiversity **1:418–429**
background information **1**:418–419
biodiversity stewardship **2**:300
disadvantages **2**:544
influencing factors **2**:544, **2**:554*F*
landscape-scale effects
general discussion **2**:544
landscape connectivity **2**:544–545
landscape heterogeneity **2**:544
local management strategies
chemical use minimization **2**:542
crop production **2**:542–544
disturbance events **2**:542–544
general discussion **1**:419–420
livestock grazing **2**:542–544
native vegetation benefits **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
pest control **2**:544
pollination **2**:544
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
structural complex agricultural areas **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
management practices **2**:299–300, **2**:300
research summary **2**:545–546
farmland birds **4**:259
Farquhar model **2**:674–675
Fascicularia bicolor **3**:517
Fasciola hepatica **5**:664–665
Fastigitermes **4**:409*T*
fatty acids **5**:267, **5**:268*F*, **7**:25–26, **7**:32–33
faucet snail **4**:362
Faulkner, William **4**:652
Fauna and Flora International **2**:231–232
Faurea saligna **1**:68
Faurea speciosa **1**:68
fearful responses **1**:39–40
Feather River **7**:152–153
feathers **1**:600–601
feather stars **4**:381*T*
feathertail glider **4**:687*T*, **4**:688–689
feathertail possum **4**:688–689
Fecampia balanicola **7**:442–443*F*
Fecampiida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
Feddea **6**:83
Federal Agriculture Improvement and Reform Act (1996) **6**:597–598
Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) **4**:90–91
Federal Noxious Weed Act (1974) **3**:715
Felidae
body mass **1**:687*F*
ecological characteristics **1**:688*F*
evolutionary history **2**:583*F*
food categories **1**:690*F*, **1**:691*F*
home range size **1**:687*F*
late Quaternary extinction event (LQEE) **4**:724–725*T*, **4**:727*T*
number of species **4**:690*T*
species diversity **1**:686
tropical rain forests **6**:309*T*
Feliformia **4**:693
Felis **4**:724–725*T*
Felis catus **1**:77
Felis chaus **5**:475
Felis concolor **1**:703, **7**:229*T*
Felis concolor coryi **3**:190–191, **6**:212
Felis lybica **1**:686
Felis silvestris **3**:191, **5**:475
femtoplankton **5**:132, **5**:133*F*
fenitrothion **5**:693–694
fennec fox **4**:693
fennel **2**:383
fens
Asia **1**:277–278
definitions and terminology **7**:369*T*
plant species richness **7**:388–389
fentin **5**:695, **5**:700*T*
fenugreek **2**:662–663
Fenusa pusilla **4**:182–183
feral animals **1**:194–196, **3**:193, **6**:576–577
feral horses **5**:313
feral pigs **1**:76–77, **4**:359, **5**:603–604
Fergusoninidae **3**:489–491*T*
fermentation, microbial **5**:301*T*
Fernow, Bernard **2**:281
ferns
Antarctica **1**:184*T*
characteristics **6**:59
end-Cretaceous mass extinction **5**:174
freshwater ecosystems **3**:571–572
invasive species **4**:346–347, **4**:347*F*
Late Triassic mass extinction **5**:172–173
modern extinctions **5**:368*T*

number of species **6**:58*T*
taxonomic distributions **3**:519*T*
fern wren **2**:75*F*
ferrets **1**:685–686
Ferroglobus **1**:219–220, **1**:221*T*
fertile crescent region **2**:658
fertility and mortality patterns **6**:200, **6**:202–204, **6**:205*F*
fertilizers
biodiversity-friendly farming **2**:542
biofuel crop production **1**:530
historical background **5**:520–522
nitrogen cycle impacts **5**:520–522, **5**:522*F*, **5**:538–540, **5**:540*F*
nutrient management–water quality relationship **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
freshwater invertebrates **4**:377
land-use change impacts **1**:107–108
leaching vulnerability index **1**:102*F*
management criteria **1**:97, **1**:97*T*
nitrogen (N) **1**:96*F*, **1**:98
nutrient input and sources **1**:96*F*, **3**:575*F*, **3**:576–577, **3**:576*F*
phosphorus (P) **1**:96*F*, **1**:98–99
regional nutrient imbalances **1**:98*F*, **1**:99–101, **1**:100*F*, **1**:102*F*
remedial benefical management practices (BMPs)
characteristics **1**:103*T*
cover crops **1**:103*T*, **1**:106
critical source area management **1**:103*T*, **1**:107
erosion potential **1**:103*T*, **1**:105–106, **1**:105*T*, **1**:106*F*
farm input decisions **1**:101, **1**:103*T*
fertilizer application methods **1**:103*T*, **1**:104
fertilizer application rate and frequency **1**:103*T*, **1**:104–105
implementation strategies **1**:101, **1**:103*F*
nutrient-use efficiency improvements **1**:101–104, **1**:103*T*
riparian/buffer zones **1**:103*T*, **1**:106–107
runoff potential **1**:103*T*, **1**:105
soil and plant testing **1**:103*T*, **1**:104
source management **1**:103*T*, **1**:104
transport management **1**:103*T*, **1**:105
research background **1**:95–97
research summary **1**:108
plant community restorations **6**:441*F*, **6**:442
Ferula negevensis **5**:485
Festuca **1**:270, **4**:104*F*
Festuca amplissima **2**:30
Festuca novae-zelandiae **4**:348–349
Festuca orthophylla **4**:104–105
Festuca rubra **2**:390
Festuca sulcata **1**:269
Ficedula hypoleuca **5**:720
Ficus
beetles **1**:329
harvesting restraints **6**:386, **6**:568
hemiepiphytes **3**:522*T*
neotropical seasonally dry forests **5**:489
species numbers **3**:130*T*
tropical Asia **1**:263
tropical rain forests **6**:315*T*, **6**:316*T*
Ficus carica **2**:662–663, **5**:485
Ficus microcarpa **1**:277, **4**:364
Ficus (pseudosycomorus) palmata **5**:485
fiddler crabs **5**:16, **5**:16*F*
field gene banks **3**:33, **3**:652–653, **4**:319, **4**:321*T*
Field Museum (Chicago) **5**:405–406, **5**:408*F*, **5**:409*F*
fields **4**:570–572, **4**:571*F*
field sparrows **4**:490*F*
figs/fig wasps
characteristics **4**:181
figs–fig wasps dependency **2**:326, **4**:364, **5**:438–439, **6**:48–49, **6**:719
introduced species **4**:364, **6**:144–145
native crops **2**:662–663
regional cultivars **3**:32*F*
Fiji
botanical gardens **6**:519
invasive species
aquatic ferns (*Salvinia molesta*) **4**:346–347
fire ants **5**:598*F*
protected area systems **6**:493–494
Fiji petrel **1**:620–621
filamentous bacteria **2**:12*F*
filariasis **5**:668, **5**:671*T*
Filinia **6**:26
Filipendula ulmaria **5**:712*T*
Filistatidae **1**:207
Filoviridae **2**:528*T*
Fimbrystylis **2**:376*F*
fine particulate organic matter (FPOM) **3**:244–245, **6**:471–472, **6**:472*F*, **7**:51–52
finger millet **2**:663, **3**:32*F*
Finland
boreal forests
beetle species **7**:219*T*, **7**:220*F*, **7**:220*T*
forest characteristics **7**:212–213, **7**:212*T*, **7**:216–217, **7**:217*F*
landscape patterns **7**:218*F*, **7**:219*F*
botanical gardens **6**:519
finless porpoises **4**:693–694
fin whales **5**:81*T*, **5**:88–89, **5**:313–314
fire ants **4**:360, **5**:598*F*, **6**:50
fire corals **2**:335, **3**:199, **3**:202*T*
fire/fire regimes
Africa **1**:46–47
biodiversity impacts **1**:517, **7**:203*F*
biodiversity modifications **4**:140
boreal forests **1**:627–628
climate change effects **2**:104
deforestation **2**:454
disturbances
basic concepts **2**:564–565, **2**:565*F*, **2**:569–571, **2**:569*T*
ecological impacts **2**:562–563
influencing factors **2**:570*F*
tropical forest regeneration **7**:277–278
wetland ecosystems **7**:399–400
whitebark pine forests **2**:571, **2**:571
dynamic global vegetation models (DGVMs) **2**:678
ecological impactd **3:435–442**
ecological impacts
disturbance ecology **2**:562–563
ecosystem function **3**:439
ecosystem structure **3**:438–439
fire determinants
fire frequency and severity **3**:436
fire regime characteristics **3**:436
physical conditions **3**:436
phytogeography **3**:436–437
fire suppression policies **3**:440–441
global change effects **3**:441
global extent **3**:435–436
greenhouse gas emissions **3**:441–442
herbivory **3**:440
historical perspective **3**:435
invasive species **3**:440
landscape fragmentation **3**:440
management strategies **3**:440–441
prescription burns **3**:440–441
species response
plants **3**:437–438
species and population recovery **3**:439–440
wildlife **3**:438
ecosystem function measurements **3**:83*F*
El Niño Southern Oscillation (ENSO) **3**:156–158, **3**:158*F*
extreme events **4**:211–212
global extent **3**:435–436
global livestock **6**:342–343
grassland ecosystems **4**:1–2, **4**:2*F*
habitat loss and fragmentation **4**:56
historic land-use patterns **4**:580, **4**:581*F*
human impacts **4**:140, **4**:154
introduced species **4**:360
Mediterranean-type climate ecosystems **5**:213–214, **5**:213*T*
neotropical seasonally dry forests **5**:496–497
plant biodiversity threats **6**:62–63
remote sensing applications **6**:398, **6**:399*F*
savannas **1**:61, **1**:63
Southeast Asia **1**:517
tropical forest regeneration **7**:277–278
wetland ecosystems **7**:399–400
fireflies **5**:13
fire use **7**:233–234
Firmicutes **2**:611*F*, **2**:613–614
first-contact extinctions
basic concepts **5**:365
Madagascar **5**:365
New Zealand **5**:366
Pacific islands **5**:365–366
First Welfare Theorem **5**:145–147
fir trees
boreal forests **1**:628–630
geographic boundaries **1**:629*F*, **1**:630*T*
habitat–bird species diversity relationship **1**:612

fir trees (*continued*)
North America **7**:118
fish **3:456–476**
acidic deposition–mercury concentration linkage **1**:13
adaptations
anatomy and physiology **3**:477–478, **3**:478*F*, **3**:479*F*
reproduction and recruitment **3**:478–479
respiratory constraints **3**:479–480
water level change effects **3**:479*F*
ancient lakes species flocks **4**:463–464, **4**:463*T*
aquaculture **1**:190*T*, **5**:61
aquarium fish **2**:343–344, **4**:156
Asia **1**:278*T*
bioluminescence **5**:123
biomonitoring processes **3**:122*T*, **3**:123*T*
bony fish
bichirs **3**:461–462
bowfin **3**:462–463
characteristics **3**:458*F*, **6**:522–523
classifications **3**:460–461
coelacanths
characteristics **3**:461, **7**:338
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
speciation **2**:47
gars **3**:462, **3**:462*F*
lungfish
characteristics **3**:461, **3**:461*F*, **7**:338
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
major threats **7**:203*F*
paddlefish **3**:462
reedfish **3**:461–462
species diversity **3**:684–685
sturgeon **3**:462, **3**:462*F*
teleosts
Acanthopterygii **3**:467
characteristics **3**:463, **6**:522–523
clupeomorphs **3**:463–464
elopomorphs **3**:463
Euteleostei **3**:464–465
ostariophysans **3**:463–464
osteoglossimorphs **3**:463, **3**:463*F*
tetrapods
characteristics **3**:461, **7**:338–339
classification systems **7**:337*F*
threatened species **7**:201*T*
brood parasitism **5**:503
cartilaginous fishes
characteristics **3**:458–459
chimaeras **3**:460, **3**:460*F*
rays **3**:459–460
sharks
characteristics **3**:456–457, **3**:458–459, **3**:459*F*
extinction theories **2**:517
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
skates **3**:459–460, **3**:460*F*
threatened species **7**:201*T*
Census of Marine Life project **2**:12*F*
Central America **2**:25*T*
climate change effects
abundance impacts **5**:34–35
community impacts **5**:37–38
distribution impacts **5**:35–36
commodity versus wildlife issue **3**:445
conservation practices **3:443–455**
commodity versus wildlife issue **3**:445
ecosystem diversity **3**:452–453
future outlook
marine protected areas **3**:453–454
potential conservation strategies **3**:454
public and political involvement **3**:454–455
industrialized fishing **3**:445
major threats
aquaculture **3**:448–449, **3**:452
general characteristics **3**:447
global climate change **3**:449–450
habitat loss and fragmentation **3**:447–448
human population growth **3**:447
introduced species **3**:448
pollution **3**:449
marine fish biodiversity **3**:450
ocean fishing
bycatch **3**:446, **3**:446, **3**:447
commodity versus wildlife issue **3**:445
economic losses **3**:446–447
government subsidies **3**:446–447
industrialized fishing **3**:445
international fishing conflicts **3**:444–445
management problems **3**:444
overfishing **3**:443–444, **3**:445, **3**:445–446, **3**:446, **4**:156
socioeconomic concerns **3**:444
species diversity
depletion causes and consequences **3**:451–452
extinction stages **3**:452
extinction vulnerability **3**:451–452
within-species genetic diversity **3**:450–451
coral reef ecosystems **2**:332*T*, **2**:333*T*
critical pH levels **6**:153*F*
declining populations **3**:443
elevational range shifts **4**:605*T*
estuarine ecosystems **3**:299–300
eutrophication impacts **3**:364–366, **3**:365*F*
fishing/fisheries **6:420–428**
American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries **5**:464–465, **5**:464*F*
aquarium fish **2**:343–344, **4**:156
biodiversity impacts **4**:144, **7**:203*F*
bioeconomic models **5**:457–459
bycatch **3**:446, **3**:446, **3**:447, **4**:156
catch processing **6**:423
current considerations
discard levels **6**:423
economic factors **6**:422–423
ecosystem impacts **6**:423
world fish catch **6**:422–423, **6**:422*T*
world fleet **6**:422–423
cyanide fishing **4**:156
economic losses **3**:446–447
ecosystem impacts
aquarium fish **4**:156
biodiversity research **3**:483–484, **3**:484*F*
fisheries decline **4**:158
global catch data **3**:483–484, **3**:484*F*
historical trends **3**:483
marine sedimentary habitats **5**:113–114, **5**:113*F*
sustainability **3**:483
Exclusive Economic Zones (EEZs) **6**:421–422
fish-catching methods **3**:482, **6**:423
fisheries science
background information **6**:423–424
economic factors **6**:426
ecosystem management **6**:427
fish growth **6**:425*F*
fish stocks **6**:424–425
government regulation **6**:426–427
international regulations **6**:427
management strategies **5**:40–41, **6**:425
maximum sustainable yield (MSY) **3**:483, **6**:423–424, **6**:424*F*, **6**:425*F*
monitoring strategies **6**:425–426
numbers of fish **6**:425*F*
replenishment **6**:424–425
fishery–diet/habitat overlap **5**:72
fishery management **5**:40–41, **5**:150–152, **6**:421
future outlook **6**:427
government legislation and regulations **3**:713*T*, **3**:717–718
government subsidies **3**:446–447
habitat degradation **4**:566–567
historical perspective **6**:420–421
hunter–gatherer societies **6**:567
individual transferable quota (ITQ) systems **5**:150
international organizations **4**:329
invasive species **5**:61
lake and pond ecosystems **4**:465
marine ecosystem services models **5**:464–465, **5**:464*F*
marine protected areas **5:94–104**
benefits **5**:96–97
characteristics and functional role **5**:94–95
climate change threats **5**:100–101
conservation strategies **5**:39–40
ecosystem-based management **5**:40, **5**:95
effectiveness **5**:100
fish stocks **3**:486
geographic extent **5**:445–446, **5**:446*T*
global commitment **5**:99–100
socioeconomic benefits **5**:98–99, **5**:98*T*
threats and challenges **5**:100–101, **5**:447–448
ocean fishing
bycatch **3**:446, **3**:446, **3**:447, **4**:156

commodity versus wildlife issue **3**:445
economic losses **3**:446–447
government subsidies **3**:446–447
industrialized fishing **3**:445
international fishing conflicts **3**:444–445
management problems **3**:444
overfishing **3**:443–444, **3**:445, **3**:445–446, **3**:446, **4**:156
overfishing
classifications **3**:445
coral reef ecosystems **2**:341, **5**:61
ecological footprint analysis (EFA) **2**:708
human impact **4**:156
marine ecosystems **3**:410, **5**:54
ocean fishing **3**:445–446
prevalence **3**:443–444
salmon **6**:527–528
tragedy of the commons **5**:150
vulnerability assessments **3**:446
parasitic disease vectors **5**:672, **5**:672*F*
payment for ecosystem services (PES) **5**:151
plant defense system uses **2**:431–432
popularity declines **7**:246–247, **7**:246*F*
post-World War II era **6**:421
poverty–conservation relationship **2**:232
recreational fishing **2**:3*F*
river–floodplain connectivity **7**:146–147
socioeconomic concerns **3**:444
trophic dynamic concept **3**:239–240
world fish catch **6**:420*F*, **6**:422–423, **6**:422*T*
fish poisons **1**:589
fish stocks **3:477–487**
adaptations
anatomy and physiology **3**:477–478, **3**:478*F*, **3**:479*F*
reproduction and recruitment **3**:478–479
respiratory constraints **3**:479–480
water level change effects **3**:479*F*
aquaculture impacts **1**:191–193
biodiversity
AquaMaps initiative **3**:484–485, **3**:485*F*, **3**:486*F*
bioquads **3**:484–485
challenges **3**:484
global catch data **3**:483–484, **3**:484*F*
marine protected areas **3**:486
species databases **3**:485
stock management **3**:485–486
commodity versus wildlife issue **3**:445
ecological footprint analysis (EFA) **2**:708
economic factors **6**:426
ecosystem management **6**:427
exploited fish stock distribution
open-ocean habitats **3**:443–444, **3**:481
overall distribution ranges **3**:481
shelf communities **3**:481–482
fish growth **6**:425*F*
general discussion **3**:477
government regulation **6**:426–427
industrialized fishing **3**:445
international fishing conflicts **3**:444–445
international regulations **6**:427
management problems **3**:444
management strategies **6**:425
maximum sustainable yield (MSY) **6**:423–424, **6**:424*F*, **6**:425*F*
monitoring strategies **6**:425–426
numbers of fish **6**:425*F*
replenishment **6**:424–425
respiratory constraints
adaptations **3**:479–480
basic geometrical constraints **3**:479, **3**:480*F*
gill area **3**:479, **3**:480*F*
growth–maturity–survival relationship **3**:480, **3**:481*F*
river–floodplain connectivity **7**:146–147
shelf communities
demersal fish stocks **3**:482
major threats **3**:482–483
neritic stocks **3**:481–482
pelagic fish stocks **3**:482, **5**:688–690, **5**:690*F*
socioeconomic concerns **3**:444
fossil record **3**:688
freshwater ecosystems
acidic deposition–mercury concentration linkage **1**:13
acidification effects **3**:584–585, **3**:584*F*
African/Ethiopian region **3**:472–473
Australian region **3**:473
biodiversity issues **3**:444
energy flow **3**:243*F*
global biodiversity **3**:471*F*, **3**:472
habitats **3**:472
introduced species **4**:378
mercury contamination **1**:10–13, **1**:12*F*
Nearctic region **3**:472
Neotropical region **3**:472
North America **5**:565
Oriental region **3**:473
Palearctic region **3**:472
species richness **3**:571–572, **3**:583–584
Succulent Karoo **1**:487*T*
surface water acidification **1**:9–10, **1**:10*T*, **1**:11*F*
threatened species **3**:578–579, **3**:579*T*
fruit and seed dispersal **6**:50
genetic diversity **3**:665*F*
geographic diversity
characteristics **3**:471–472
freshwater fish
African/Ethiopian region **3**:472–473
Australian region **3**:473
global biodiversity **3**:471*F*, **3**:472
habitats **3**:472
Nearctic region **3**:472
Neotropical region **3**:472
North America **5**:565
Oriental region **3**:473
Palearctic region **3**:472
global biodiversity **3**:471*F*, **3**:472
marine ecosystems
Antarctic region **3**:474
Arctic region **3**:474
characteristics **3**:473
deep sea fish **3**:475–476, **3**:475*F*
Eastern Atlantic region **3**:474
Eastern Pacific region **3**:473
Indo–West Pacific region **3**:473
pelagic regions **3**:475
temperate regions **3**:474–475
Western Atlantic region **3**:473–474
global wildlife trade **3**:409*T*
grazing effects **4:8–17**
direct versus indirect effects **4**:10
disturbance regimes **4**:9–10, **4**:10*F*
grazer-based ecosystem restoration and management
coral reefs **4**:16
lakes and ponds **4**:15–16, **4**:16*F*
herbivory **4**:8–9, **4**:9*F*
host chemistry effects **4**:13
paleo-patterns **4**:13–14
research summary **4**:17
spatiotemporal scale
between-habitat patterns **4**:11–12
functional role **4**:10–11
geographic patterns **4**:11, **4**:11*F*
microhabitat patterns **4**:12–13
within-habitat patterns **4**:12
hagfish
characteristics **3**:456–457, **3**:457, **3**:458*F*, **7**:336–337
classification systems **7**:335, **7**:335*F*, **7**:336*F*
threatened species **7**:201*T*
harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*
ichthyotoxicity **1**:474–475, **1**:474*F*
introduced parasites–species decline relationships **2**:528*T*, **2**:532
lake and pond ecosystems **3**:605–606, **3**:606*F*, **3**:607*F*, **4**:460–461, **4**:461*T*
Late Devonian mass extinction **5**:170
latitudinal range shifts **4**:604*T*
mangrove ecosystems **5**:15, **5**:15*F*, **5**:20*T*
Mediterranean-type climate ecosystems
Australia **5**:213*T*, **5**:221
California **5**:213*T*
Cape Region (South Africa) **5**:213*T*, **5**:220
Chile **5**:213*T*, **5**:217
Mediterranean Basin **5**:213*T*, **5**:218–219
mercury contamination **1**:10–13, **1**:12*F*
Mesoamerican poeciliid fish **7**:348, **7**:348*F*
migratory patterns **5**:314–315
modern extinctions **5**:368*T*
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*, **5**:470–471, **5**:470*F*, **5**:471*F*
North America **5**:565
ocean environment **3**:443
oil spills **3**:192–193, **3**:255–256
overfishing **3**:410, **5**:54, **6**:482, **6**:483*F*, **6**:527–528
parasitic disease vectors **5**:672, **5**:672*F*

fish (*continued*)
pelagic ecosystems **5**:**683–691**
El Niño Southern Oscillation (ENSO) **5**:687–688, **5**:689*F*, **5**:690*F*
future outlook **5**:690–691
pelagic fish stocks **5**:688–690, **5**:690*F*
phytoplankton **5**:683–686, **5**:684*F*, **5**:686*F*
primary production **5**:683–686, **5**:685*F*, **5**:686*F*
Russell cycle **5**:690–691
zooplankton **5**:686–688, **5**:687*F*, **5**:688*F*, **5**:689*F*
Phanerozoic diversity **3**:688
poisonous compounds **1**:589
respiratory constraints
adaptations **3**:479–480
basic geometrical constraints **3**:479, **3**:480*F*
gill area **3**:479, **3**:480*F*
growth–maturity–survival relationship **3**:480, **3**:481*F*
river ecosystems **6**:479, **6**:480*F*
salmon **6**:**522–531**
biodiversity value **6**:529–530
classifications **6**:522–523
climate change effects **6**:529
fossil record **6**:522–523
historical status **6**:526–527
life histories
Atlantic salmon **6**:524
Chinook salmon **6**:524, **6**:525*F*
chum salmon **6**:523
coho salmon **6**:523–524
general characteristics **6**:523
pink salmon **6**:523
sea-run cutthroat trout **6**:526
sockeye salmon **6**:523
steelhead trout **6**:524–526
natural environmental variability **6**:529
population decline **6**:526–527
special adaptations
anadromy **6**:526
homing behavior **6**:526
species **6**:522
threats and challenges
anthropogenic effects **6**:527
habitat destruction **6**:527
harvesting moratorium **6**:527–528, **6**:528*F*
hatchery fish impacts **6**:528
hydropower dam impacts **6**:528–529
overfishing **6**:527–528
sandy beaches **2**:131–133
seagrass ecosystems **6**:547–548
subterranean ecosystems **7**:53–55, **7**:55*F*
Succulent Karoo **1**:487*T*
surface water acidification **1**:9–10, **1**:10*T*, **1**:11*F*
taxonomic diversity
Agnatha
characteristics **3**:457
hagfish **3**:457, **3**:458*F*
lampreys **3**:457–458, **3**:458*F*, **7**:337–338
bony fish
Acanthopterygii **3**:467
bichirs **3**:461–462
bowfin **3**:462–463
characteristics **3**:458*F*, **6**:522–523
classifications **3**:460–461
clupeomorphs **3**:463–464
coelacanths **3**:461
elopomorphs **3**:463
Euteleostei **3**:464–465
gars **3**:462, **3**:462*F*
lungfish **3**:461, **3**:461*F*
ostariophysans **3**:463–464
osteoglossimorphs **3**:463, **3**:463*F*
paddlefish **3**:462
reedfish **3**:461–462
sturgeon **3**:462, **3**:462*F*
teleosts **3**:463, **6**:522–523
tetrapods **3**:461
cartilaginous fishes
characteristics **3**:458–459
chimaeras **3**:460, **3**:460*F*
rays **3**:459–460
sharks **3**:458–459, **3**:459*F*
skates **3**:459–460, **3**:460*F*
Cephalochordata **3**:457
characteristics **3**:456, **3**:456–457
lancelets **3**:457
temperature tolerance ranges **4**:115
threatened species **7**:201*T*
ultraviolet radiation effects **7**:300–301
water-column nitrogen:phosphorus (N:P) ratios **3**:350*F*
wetland ecosystems **7**:389–391, **7**:393–394
see also aquaculture
Fish and Wildlife Service (FWS) **5**:64–66
FishBase **3**:485
Fisher, R. A. **6**:2
Fisher's α **5**:185–187
Fisher's *r* measure **4**:635–637
Fishery Conservation and Management Act (1976) **1**:397, **3**:713*T*, **3**:717–718
Fishery Resources Monitoring System (FIRMS) **4**:329
fishflies **4**:371*T*, **4**:392*T*
Fissiuridon **4**:727–728
Fissurella **5**:56–57, **5**:58*F*, **5**:59*F*
Fitch optimization **2**:40, **2**:41*F*
Fitzroya **7**:119
Fitzroya cupressoides **3**:210, **3**:211*F*
Fitzroya patagonica **3**:117
Flabelligeridae **7**:417*F*
Flacourtiaceae **1**:263, **5**:489
Flacourtia indica **1**:68
flacourts **5**:623*F*
flagellates **6**:25–26, **6**:289*T*, **6**:290*F*, **6**:293–295, **6**:294*F*, **6**:586, **6**:587*F*
flagship species **3**:225
flamingo **1**:604*F*
Flammulina velutipes **3**:637*T*
flashing fireflies **5**:13
flashlight fish **3**:467–471
flatfish **3**:470
flat-footed flies **3**:489–491*T*
flatworms
adaptability **4**:393
American Southeast rivers **3**:185*F*
classifications **4**:381*T*
lake and pond ecosystems **4**:374*F*
number of species **3**:182*T*, **4**:371*T*
predation effects **6**:240
river ecosystems **6**:478*T*
size adaptations **5**:657
see also parasites
Flaveria **2**:376*F*
Flavobacteria **5**:294
Flavobacterium **6**:32
flavorings **1**:42
flax **2**:139, **2**:383, **2**:384*F*, **2**:662–663
fleas
characteristics **4**:393*F*
insecticide resistance **4**:298–299
number of species **4**:392*T*
Flemingia macrophylla **6**:561
flesh flies **3**:489–491*T*, **3**:493
Fleurya aestuans **6**:557
flies **3**:**488–496**
Arctic ecosystems **1**:231*T*
characteristics **4**:392
classifications
Brachycera **3**:489–491*T*, **3**:493
common names **3**:489–491*T*
major families **3**:488–493, **3**:489–491*T*
nematocerous groups **3**:489–491*T*, **3**:493
conservation biology **3**:495–496
insecticide resistance **4**:298–299
life history diversity
adults **3**:494
larvae
herbivores **3**:493–494
non-free-living larvae **3**:494
parasites **3**:494
parasitoids **3**:494
predators **3**:494
prevalence **3**:493
scavengers **3**:493
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*, **5**:466*T*
number of species **3**:488, **3**:488*T*, **4**:392*T*
pollination **6**:141, **6**:141*T*
prevalence **3**:488, **3**:488*T*
soil fauna **6**:590*T*
special associations
aquatic Diptera
freshwater habitats **3**:494–495
phytotelmata **3**:494–495
kleptoparasitism **3**:495
social insects **3**:495
swarms **3**:495
subterranean ecosystems **7**:53
systematic classifications **7**:91–93, **7**:92*F*
flight intercept traps **1**:330
flightless birds
Australia **4**:727–728
island biogeography **1**:605–606
Madagascar **5**:365
Oceania **4**:723*F*, **4**:730–731
Flindersia brayleyana **6**:458–459
floating toilets **1**:512*F*
floodplain ecosystems
Amazon ecosystems
buritizal **1**:164

igapó forest **1**:163–164
lakes **1**:164
mangrove forests **1**:164
permanent swamps **1**:164
pirizal **1**:164
seasonal flooding **1**:162–163
várzea forest **1**:163
habitat loss and degradation **7**:371*T*
river–floodplain connectivity **7:144–160**
benefits **7**:144, **7**:145–146
changes and conversion impacts **7**:150–151, **7**:151*F*
ecological processes **7**:145
ecosystem services
benefits **7**:145–146, **7**:150
fisheries **7**:146–147
flood attenuation **7**:147–150, **7**:148*F*, **7**:149*F*
floodplain agriculture **7**:146
flow levels and events **7**:150
hydrological connectivity **7**:150
overbank flooding **7**:145–146
spatial variables **7**:150
floodplain reconnection and conservation
ecosystem restoration **7**:153
flood-risk reduction programs **7**:154–155, **7**:156*F*, **7**:157*F*
general discussion **7**:152–153
livelihood restoration **7**:153–154
multipurpose reservoir flexibility **7**:155–156, **7**:156*F*, **7**:157*F*
restoration strategies **7**:152–153
future research outlook **7**:156–158
geomorphic processes **7**:145
hydrological processes **7**:145
key floodplain processes **7**:144–145
land and water management challenges
flood control issues **7**:152
flood risks **7**:152
freshwater biological diversity and productivity declines **7**:151
general discussion **7**:151
river-dependent communities impacts **7**:151–152
research summary **7**:156–158
floods
El Niño Southern Oscillation (ENSO) events **3**:159
flooded forests **7**:272
flooded grasslands **6**:600–604*T*
flood management systems **5**:452, **5**:452–453, **5**:453*T*
plant biodiversity threats **6**:62–63
river–floodplain connectivity **7**:147–150, **7**:148*F*, **7**:149*F*
tropical forest regeneration **7**:277–278
wetland ecosystems **7**:399–400
flora *see* plants; vegetation
floral nectar spurs **6**:47–48
Floridaceras **5**:224*T*
Florida Everglades (case study) **7**:372, **7**:372*F*, **7**:373*F*
Florida panther **3**:190–191, **3**:192, **6**:212
Florida scrub jays **7**:186–187
floristic geobotany **4**:396*T*
flounder **3**:467–471
flour beetles (*Tribolium*)
interspecific competition
environmental conditions **2**:184
single-species results **2**:184, **2**:184*T*
two-species results **2**:184–185, **2**:185*F*, **2**:185*T*
Flourensia cernua **1**:708
flower flies **3**:489–491*T*, **3**:492
flowering dogwood **2**:528*T*, **3**:411, **3**:411*F*
flowering plants **5:613–627**
Antarctica **1**:184*T*
background information **5**:613
characteristics **6**:60
diversification
analytical methodologies **5**:619–620
historical patterns **5**:625–626
major events
ASTERID I and ASTERID II Clades diversification **5**:622*F*, **5**:625
ASTERID III and ASTERID IV Clades diversification **5**:622*F*, **5**:625
early diversification events **5**:620–621, **5**:621*F*, **5**:622*F*
general discussion **5**:620–621
Rosid 1 Clade diversification **5**:621–625, **5**:623*F*, **5**:624*F*
molecular clock-based models **5**:619–620
research background **5**:614
emotional responses **1**:41–42
number of species **6**:58*T*
origins
fossil record
challenges **5**:615
early angiosperms **5**:617–619
general discussion **5**:615–617
nineteenth century research **5**:613–614
Phanerozoic **3**:686–687, **3**:687*F*
potential angiosperm sister groups **5**:615–617, **5**:616*F*, **5**:617*F*, **5**:618*F*
phylogenetic and morphological studies **5**:614–615, **5**:618*F*
research background **5**:613–614
research summary **5**:619
parasites **5**:662
tourism industry **7**:223
flower-loving flies **3**:489–491*T*
fluazifop-butyl **5**:697
flukes **2**:645
flumiclorac **4**:90*T*
Fluorencia thurifera **3**:116
fluorescent *in situ* hybridization (FISH) **5**:261, **5**:264*T*, **5**:267
fluorine (F)
chlorofluorocarbons (CFCs)
atmospheric concentrations **1**:282*T*
changing concentrations **1**:284–286, **1**:286*F*
climate impact **1**:281–282, **3**:450
ecotoxicology **3**:118–119
enhanced greenhouse effect **4**:19*T*
environmental ethics **3**:271
ozone depletion **7**:296–297
hydrofluorocarbons (HFCs) **1**:284–286, **4**:19*T*
perfluorocarbons (PFCs) **1**:284–286
fluridone **4**:90*T*
flutriafol **5**:695
Fluvisols **2**:461–464
flycatchers **2**:189–190, **2**:189*F*
flying fish **3**:467
flying foxes **4**:447, **4**:691–692
flying lemurs **4**:691
flying squirrels **4**:701*T*
Foeniculum vulgare **2**:383
fog **5**:479–480
Fokienia **1**:272–273
folk taxonomies **1**:121, **7**:100–101
Fondo de Agua Por La Vida y La Sostenibilidad (FAVS) **7**:362–363
Fondo para la Conservación del Aguaor (FONAG) **7**:355–357, **7**:361–362
fonio **2**:659, **2**:663
Fontaria virginiensis **5**:424–425
Fontigens bottimeri **7**:59
Food, Agriculture, Conservation, and Trade Act (1990) **3**:719, **6**:597–598
Food and Agriculture Organization (FAO) **1**:640–642, **1**:641, **3**:643, **4**:328, **4**:567, **6**:77–78, **6**:322, **6**:422–423, **6**:597–598
foodborne trematodes **5**:671*T*
food caching **1**:688
food chains
ecosystem function measurements **3**:78, **3**:80*F*
energy flow **3**:244–245
living dead **4**:595–596
marine ecosystems **5**:50–51
nutrient cycling **3**:244–245
schematic diagram **3**:497*F*
thermal ecosystems **4**:118–119
trophic interactions **5**:46–47, **7**:264, **7**:265*F*
food crops **1**:542–543, **1**:542*F*
Food, Drug, and Cosmetic Act (FDCA) **4**:90–91
food habits **3**:127–128, **3**:127*T*
Food Security Act (1985) **3**:719, **6**:597–598
foods/food selection **1**:42–43, **1**:42*T*
food spices **2**:432–433
food webs **3:497–510**
allochthonous inputs **3**:505
classifications
characteristics **3**:498–500
connectivity webs **3**:498–500, **3**:499*F*, **3**:500*F*
energetic webs **3**:499*F*, **3**:500–501, **3**:507–508
functional/interaction webs **3**:499*F*, **3**:501–502
community complexity **2**:209–210
dietary shifts **3**:506
ecosystem services **3**:90*T*
El Niño Southern Oscillation (ENSO) **5**:687–688, **5**:689*F*
energy flow **3**:237–238
eutrophication impacts **3**:582–583, **5**:24–25

food webs (*continued*)
exploitation ecosystem hypothesis **3**:503, **3**:508–509
food chains **3**:497*F*
forest ecosystems **3**:533–534
generalized food web theory **6**:231, **6**:231*F*
green-world hypothesis **3**:503–504, **3**:508–509
intermediate levels of complexity **3**:509
lake and pond ecosystems **3**:582–583, **4**:461–462
multispecies interactions **6**:721–722, **6**:721*F*, **6**:722*F*
omnivory **3**:502–503
pelagic ecosystems **5**:687–688, **5**:689*F*
primary production
exploitation ecosystem hypothesis **3**:503, **3**:508–509
general discussion **3**:503–504
multichannel omnivory **3**:504
trophic cascades **3**:504–505
trophic complexity **3**:504
trophic levels **3**:503–504
research background **3**:497–498
research trends
age structure effects **3**:506
carbon:nitrogen:phosphorus (C:N:P) stoichiometry **3**:506–507
detritus **3**:505–506
interaction strengths **3**:507
intermediate levels of complexity **3**:509
modeling approaches **3**:508–509
nutritional balance **3**:506–507
open system communities **3**:505
population and community dynamics **3**:507–508
schematic diagram **3**:498*F*
soil food webs **6**:448–449, **6**:449*F*
species coexistence **6**:674
species diversity **6**:712, **6**:713*F*
species-interaction theories **3**:3–4, **3**:4*F*
trophic levels **3**:239–240, **3**:241*F*, **3**:503–504, **5**:46–47, **5**:48*F*, **5**:49*F*, **7**:266–268, **7**:266*F*, **7**:267*T*
foot and mouth disease **3**:191
foraging behaviors
carnivores **1**:687–690
disturbance ecology **2**:566
foraging lizards **6**:406–408, **6**:407*T*
Foraminifera
average lifespan **4**:656*T*
Cenozoic mass extinctions **5**:176
characteristics **6**:291
end-Cretaceous mass extinction **5**:173–175
marine ecosystems **5**:48, **5**:106*F*
paleotemperature assessments **4**:224–226, **4**:225*F*
schematic diagram **6**:290*F*
species richness **6**:289*T*
zooplankton fraction **6**:25
Foraminitermes **4**:409*T*
Forbes, Edward **1**:578–579
Forbesichthys agassizi **7**:55*F*
Forbes, Stephen Alfred **3**:238–239
Ford, E. B. **6**:3
Fordia splendidissima **3**:156–158
Foreign Assistance Act (1962) **3**:713*T*, **3**:719
forest canopies
animal diversity **3:511–515**
canopy architecture **3**:511–512
invertebrates **3**:512–513, **3**:513*F*, **3**:514*F*
research background **3**:511
research summary **3**:514–515
research techniques **3**:512, **3**:512*F*
vertebrates **3**:513–514
characteristics **1**:449, **3**:511
definition **3**:516
plant diversity **3:516–527**
climate change effects **3**:524
conservation considerations **3**:524
extinction vulnerability **3**:524
fragmentation effects **3**:524
future research areas
databases **3**:526
ethnobotanical studies **3**:526
experimental field work **3**:526
general considerations **3**:524
herbaria **3**:524–526
monographs and inventories **3**:524
research questions and opportunities **3**:525*T*
growth habit diversity
characteristics **3**:523
nutrition acquisition and retention **3**:523
photosynthetic pathways **3**:523
habitat conversion effects **3**:524
habitat diversity
canopy microclimate **3**:521–522
host tree specificity **3**:523
spatial scale **3**:522–523
historical research **3**:517
nutrition acquisition and retention
ant-associated epiphytes **3**:523–524
atmospheric epiphytes **3**:524
bark epiphytes **3**:524
humus epiphytes **3**:523
tank/trash basket epiphytes **3**:523
plant categories
hemiepiphytes **3**:518
nonvascular epiphytes **3**:517–518
parasites **3**:518–519
vascular epiphytes **3**:517, **3**:518*T*
research summary **3**:526
taxonomic distributions
hemiepiphytes **3**:520–521, **3**:522*T*
nonvascular epiphytes **3**:520, **3**:521*T*
parasites **3**:521
vascular epiphytes **3**:519–520, **3**:519*T*, **3**:520*T*
research scope **3**:516–517
tropical forest canopies **3**:709–710
Forest Carbon Partnership Facillity (FCPF) **4**:328–329
forest fires *see* fire/fire regimes; wildfires
Forest Legacy Program **3**:719
Forest Resources Information System (FORIS) **6**:322
forests
acidic and mercury deposition **1:1–14**
emission occurrences and sources **1**:1–5
environmental impacts **1**:5, **1**:5*T*
geographic distribution **1**:1–5, **1**:3*F*
mercury emissions and deposition **1**:2–5, **1**:4*F*
research background **1**:1
schematic diagram **1**:2*F*
soils
alteration consequences **1**:5
aluminum mobilization **1**:5–6
calcium loss **1**:5, **1**:6*F*
geographic distribution **1**:7*F*
nutrient cation loss **1**:5, **1**:6*F*
sulfur and nitrogen accumulations **1**:6
surface water **1**:7*F*
trees
red spruce **1**:6–8, **1**:8*F*
species declines **1**:6–8
species distribution shifts **1**:8
sugar maples **1**:8, **1**:8*F*
Africa
Afro-montane forests **5**:219
tropical Africa **1**:48–49
agroforestry **2**:382–383
Asia
boreal forest **1**:262–263*T*, **1**:273–274
humid tropical forest **1**:260–263, **1**:262–263*T*, **1**:266*T*
laurel forests **1**:270–271
pine forests **1**:273–274
temperate deciduous forests
characteristics **1**:271
China **1**:271–272
Japan **1**:271
Korea **1**:271
Manchuria **1**:271–272
pine forests **1**:272
regional species richness controls **2**:602
tropical moist forest **1**:277*T*
Australia **1**:292–297, **1**:294*T*
biodiversity restoration **6**:456–457
boreal forests **1:626–635**
animal species **1**:631–632, **1**:633*F*
Asia **1**:273–274
biodiversity change scenarios **4**:6*F*
Canadian Shield **5**:562
climate
characteristics **1**:626
precipitation **1**:626–627, **1**:628*F*
seasonal boundaries **1**:626
temperatures **1**:626, **1**:628*F*
climate change studies **2**:61*F*
conservation
clear-cutting threats **1**:634
natural reserves and preserves **1**:634
sustainable management **1**:634
disturbance dynamics **7**:214–215, **7**:214*T*
ecosystem function measurements **3**:77*F*
Europe **3**:335*T*, **3**:336–340*T*
Finland **7**:212–213, **7**:212*T*
geography **1**:626, **1**:627*F*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*

global distribution patterns **7**:117*T*
litter deposition **6**:120*T*
management strategies **7**:213, **7**:213*F*, **7**:214*T*
precipitation–temperature relationship **1**:59*F*
predation effects **6**:236–237, **6**:238*F*
protected area systems **6**:492*T*, **6**:499*T*
soils **1**:627, **6**:123*T*
structure and function **7**:213, **7**:213*F*
succession **1**:632–634
successional dynamics **7**:213–214, **7**:214*F*
timber-oriented management
beetle species **7**:218–219, **7**:219*T*, **7**:220*F*, **7**:220*T*
forest characteristics **7**:216–217, **7**:217*F*
impact mechanisms **7**:215–216, **7**:217*F*, **7**:217*T*
landscape patterns **7**:217–218, **7**:217*F*, **7**:218*F*, **7**:219*F*
production optimization **7**:215, **7**:216*F*
research summary **7**:219–220
vegetation
characteristics **1**:628–630
closed forests **1**:629–630
forest–tundra ecotone **1**:630
geographic boundaries **1**:629*F*, **1**:630*T*
latitudinal classifications **1**:629–630
lichen woodlands **1**:630
longitudinal classifications **1**:630–631, **1**:631*T*, **1**:632*T*, **1**:633*F*
wildfires **1**:627–628
C_4 plants **2**:371*T*
Chinese National Forest Conservation Program (NFCP)
background information **3**:375–376
biodiversity/biophysical effects **3**:376–377, **3**:377*F*
future outlook and challenges **3**:378–379
program goals **3**:375–376
socioeconomic effects **3**:377–378, **3**:378*F*
spatial distribution **3**:376, **3**:376*F*
summary discussion **3**:382
climate change studies **2**:61*F*
closed forests **1**:629–630, **6**:325–327, **6**:326*F*, **6**:326*T*
coniferous forests
biodiversity change scenarios **4**:6*F*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
North America **5**:564–565
protected area systems **6**:492*T*, **6**:499*T*
South America **6**:600–604*T*
conservation status **3**:170–174, **3**:171–173*T*
deciduous forests
biodiversity change scenarios **4**:6*F*
climate change studies **2**:61*F*
current global land cover **4**:557*F*
global agriculture impacts **3**:428*T*
litter deposition **6**:120*T*
net primary production (NPP) **1**:560*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
temperate deciduous forests **1**:271
tropical deciduous forests **3**:404
definition **2**:447–448
dry forests
conservation status **3**:171–173*T*
global biodiversity trends **4**:563*F*
neotropical seasonally dry forests **5**:488–500
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
precipitation–temperature relationship **1**:59*F*
protected area systems **6**:492*T*, **6**:499*T*
tropical dry forest **1**:266–267, **2**:28, **3**:110–111, **3**:111*F*, **3**:112–113, **6**:600–604*T*
ecosystems **3**:528–536
biodiversity
community organization **3**:533–534
disturbance regimes **3**:530–533
ecosystem functions **3**:535
food webs **3**:533–534
forest structure and patterns **3**:530–533, **3**:532*F*
general patterns **3**:530, **3**:531*F*
human activity **3**:534
keystone species **3**:535
latitudinal gradients **1**:498*F*, **3**:530, **3**:530*F*
tree taxa **3**:532*T*
biogeochemical cycles
carbon cycle **6**:149–150, **6**:150*T*
nitrogen cycle **6**:152
classification systems **3**:528–530, **3**:529*F*
climate change effects **6**:149–150, **6**:150*T*
disturbance events **3**:529
human activity
biodiversity **3**:534
dramatic climatic shifts **3**:535
forest conversion and fragmentation **3**:534, **4**:154, **4**:156, **5**:496, **5**:497*F*
harvesting effects **3**:534, **4**:156
introduced species **3**:534
pollution **3**:534–535
living dead **4**:590–591, **4**:591*F*
overexploitation **4**:156
precipitation **3**:529*F*
South America **3**:110
temperature **3**:529*F*
ecosystem services **3**:96–98
evergreen forests **4**:432*T*, **6**:120*T*
flooded forests **7**:272
forested wetland conservation **7**:371
forest fragmentation **4**:154, **5**:496, **5**:497*F*
forestry-related organizations **4**:328–329
forest–tundra ecotone **1**:630
fragmented forests **6**:325–327, **6**:326*F*, **6**:326*T*
fungi-related diseases **3**:633, **3**:633*T*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
Global Forest Resources Assessment **6**:322
grazer-based ecosystem restoration and management **4**:15
habitat–bird species diversity relationship **1**:612–613
habitat fragmentation **4**:564, **4**:565*F*
historical conservation practices **2**:279–280, **2**:281, **2**:283
historic land-use patterns **4**:572–574, **4**:573*F*, **4**:574*F*
humid forests
high mountain ecosystems **7**:275
lowland ecosystems **7**:273
low mountain ecosystems **7**:275
international organizations **4**:328–329
introduced plants **1**:76
landscape legacies **4**:525–527
land-use trends **4**:558, **4**:559*F*
large-scale biodiversity experiments **4**:588–589
latitudinal gradients **1**:498*F*
lichen woodlands **1**:630
litter deposition **6**:119–121, **6**:120*T*
living dead
agroscapes **4**:590–591, **4**:591*F*
animal populations **4**:594–595
anthropogenic impacts **4**:594–595
deforestation **4**:591–593, **4**:592*F*
extinct megafauna **4**:591–593, **4**:593*F*
food chains **4**:595–596
requirements **4**:593–594
restoration programs **4**:597
small plants **4**:594
wildland patches **4**:596–597, **4**:596*F*
logged and managed forests **1**:446–458
abiotic changes
forest floor hydrology **1**:450
insolation **1**:450
temperature **1**:450
wind patterns **1**:450
biotic changes
aquatic fauna **1**:453
behavioral changes **1**:453

forests (*continued*)
characteristics **1**:450–451
domestic animals **1**:453
fungi **1**:451
invertebrates **1**:451–452
mycorrhizae **1**:451
nonforest species **1**:453
plants **1**:451
vertebrates **1**:452–453
ecological interactions **1**:453–454
emerging trends **1**:455–457
environmental impacts **7**:314–315
evolutionary processes **1**:454–455
genetic effects **1**:454–455
management systems
general discussion **1**:447
long-term intensive forest management practices **1**:448
managed versus secondary forests **1**:448
nontimber forest products (NTFPs) **1**:448
shifting agricultural practices **1**:448
timber management **1**:447
plants
early successional phases **1**:451
exotic species invasions **1**:451
forest management impact **1**:451
research background **1**:446–447
structural alterations
coarse woody debris (CWD) **1**:450
environmental impact **1**:448–449
forest canopies **1**:449
fragmentation **1**:449
gaps and edges **1**:449
soil structure **1**:450
spatial patterns **1**:449–450
vertical structure **1**:449
synergistic anthropogenic effects **1**:455
timber management
clearing systems **1**:447
regeneration systems **1**:447
salvage logging **1**:447–448
selective logging systems **1**:447
vertebrates
behavioral changes **1**:453
biotic changes **1**:452–453
birds **1**:453
domestic animals **1**:453
food resource loss or gain **1**:453
habitat loss or gain **1**:452–453
primates **1**:453
long fallow **6**:325–327, **6**:326*F*, **6**:326*T*
mixed forests **2**:61*F*, **3**:428*T*, **4**:563*F*, **5**:526*T*, **6**:492*T*, **6**:499*T*
montane forests **4**:432*T*
Near East ecosystems **5**:480*F*, **5**:481–482
neotropical seasonally dry forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
net primary production (NPP) **1**:560*T*
nitrogen cycle impacts **5**:543
nitrogen deposition **5**:526*T*, **5**:534*F*
Nothofagus forests
characteristics **7**:118–119
habitat–bird species diversity relationship **1**:610–611, **1**:610*F*
phylogenetic analyses **6**:22
old-growth forests **3**:314–315, **3**:315*F*
open forests **6**:325–327, **6**:326*F*, **6**:326*T*
perhumid forests
high mountain ecosystems **7**:275
lowland ecosystems **7**:272–273
low mountain ecosystems **7**:275
pre-agriculture land cover **4**:557*F*, **4**:558*F*
protected area systems **6**:492, **6**:492*T*, **6**:494*F*, **6**:494*T*, **6**:499*T*
reforestation **6:370–379**
Australian reforestation programs
mining impacts **6**:378
North Queensland **6**:377
Brazilian reforestation programs **6**:377–378
challenges **6**:371
ecosystem recovery **6**:374–375
general discussion **6**:370
methodologies
direct seeding **6**:372, **6**:372*F*
natural recovery processes **6**:371–372, **6**:371*F*
planting densities **6**:373
seedling planting **6**:372
species richness **6**:373–374, **6**:374*F*
species selection **6**:372–373, **6**:373*T*
needs **6**:370
Nepal **6**:378–379
New Zealand reforestation programs **6**:377
post-degradation methods **6**:370–371, **6**:371*F*
priority area selection **6**:375, **6**:375*T*
research summary **6**:376
socioeconomic factors **6**:376
specialized conditions/problem areas **6**:375–376
United Kingdom reforestation programs **6**:376–377
regional carbon stocks **1**:678*F*
restoration programs **1**:521–522
riverine forests **4**:432*T*
selective harvesting **4**:154
semideciduous forests **4**:432*T*
short fallow **6**:325–327, **6**:326*F*, **6**:326*T*
South America **6**:600–604*T*
subhumid forests
high mountain ecosystems **7**:275
lowland ecosystems **7**:273–274
low mountain ecosystems **7**:274–275
temperate forests **7:112–127**
biodiversity change scenarios **4**:6*F*
biodiversity trends **7**:119–121, **7**:119*T*, **7**:120*F*
canopy architecture **3**:511–512
Central America **2**:29–30
characterization
dominant floristic features **7**:117–119
physiognomic features **7**:117
conservation
conservation status **7**:116*T*, **7**:125–126
future research areas **7**:127
major threats **7**:126–127
protection strategies **7**:126
conservation status **3**:171–173*T*
current global land cover **4**:557*F*
deforestation **1**:677*F*, **2**:447–448, **2**:448*T*
distribution hypotheses
general discussion **7**:121–124
historical/geological perspectives **7**:122–124, **7**:123*F*
productivity **7**:124
spatiotemporal heterogeneity **7**:124
ecosystem function measurements **3**:77*F*
ecosystem services **1**:334*T*, **7**:124–125
endemism **7**:121
Europe **3**:335*T*, **3**:336–340*T*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
global distribution patterns **7**:112–117, **7**:113*F*, **7**:115*F*, **7**:116*T*, **7**:117*T*
importance **7**:112
litter deposition **6**:120*T*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
predation effects **6**:236–237, **6**:238*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
range size rarity **7**:121
soils **6**:123*T*
South America **3**:111*F*, **3**:117, **6**:600–604*T*
termites **4**:432*T*
termites **4**:432*T*
threatened birds **7**:181
tropical dry forest
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*
seasonally dry tropical forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*

extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
tropical forests
Africa **1**:48–49
biodiversity change scenarios **4**:6*F*
biodiversity-rich countries **1**:497–498
canopy architecture **3**:511–512
current global land cover **4**:557*F*
deforestation
annual forest change **2**:448*T*
C_4 plants **2**:379–381, **2**:380*F*
causal factors **2**:379–381, **2**:380*F*
contemporary deforestation **2**:448–449
ecological impacts **1**:677*F*, **1**:679*T*
ecosystem function measurements **3**:77*F*
ecosystem services **1**:334*T*
El Niño Southern Oscillation (ENSO) events **3**:156–158, **3**:158*F*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
global distribution patterns **7**:117*T*
grazer-based ecosystem restoration and management **4**:15
habitat loss and fragmentation **3**:405–406, **3**:405*F*
livestock impacts **6**:338–340, **6**:340*F*
living dead **4**:596–597
neotropical seasonally dry forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
predation effects **6**:239–240
protected area systems **6**:492*T*, **6**:499*T*
regeneration systems **3:277–286**
challenges **7**:284–285
chronosequence studies **7**:279–280
deforestation rates **7**:277
forest dynamics **7**:278, **7**:279*T*
functional diversity (FD) **7**:282
historical perspective **7**:277–278
prior land use effects **7**:281
recovery processes **7**:281–282
regeneration sources **7**:280–281
seed dispersal–predation relationship **7**:283–284
soil fertility **7**:281
species composition changes **7**:282–283, **7**:283*F*
successional continuum **7**:278
successional trajectories and stages **7**:278–280, **7**:279*T*
seasonally dry tropical forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
South America
distribution and structure **3**:110
functional aspects **3**:112
general patterns **3**:111–112
land cover classifications **6**:607*T*
mountain cloud forests **3**:111
natural ecosystems **6**:599–608, **6**:606*F*, **6**:606*T*
tropical broadleaf forests **6**:600–604*T*
tropical moist forest **6**:600–604*T*
tropical rain forests **3**:110, **3**:111*F*, **3**:112
tropical semideciduous/dry forests **3**:110–111, **3**:112–113
tropical broadleaf forests **3**:428*T*, **6**:600–604*T*
tropical deciduous forests **3**:404
tropical dry forest **1**:266–267, **6**:600–604*T*
tropical forest canopies **3**:709–710
wildfire effects **3**:156–158, **3**:158*F*
tropical moist forest
Asia **1**:277*T*
biodiversity measures **1**:404
conservation status **3**:171–173*T*
nitrogen deposition **5**:526*T*
protected area systems **5**:444*T*
South America **6**:600–604*T*, **6**:607*T*
tropical rain forests
animal diversity **6:304–312**
beetles **6**:306
birds **6**:307–309, **6**:308*T*
butterflies **6**:306–307, **6**:307*T*
frogs **6**:307
future outlook **6**:311–312
general characteristics **6**:306
geographical patterns **6**:310
geographic distribution **6**:304–305
high species richness **6**:310–311
historical perspective **6**:304
mammals **6**:309–310, **6**:309*T*
assessment techniques
Global Forest Resources Assessment **6**:322, **6**:323, **6**:324*T*
model-based assessments **6**:322
multidate remote sensing survey **6**:322–323
biodiversity-rich countries **1**:500–501
Central America **2**:27–28
characteristics **6**:304–305
geographic distribution **6**:304–305
global biodiversity trends **4**:563*F*
habitat destruction **3**:403–404, **3**:403*F*
livestock impacts **6**:338–340, **6**:340*F*
loss and change **6:322–329**
assessment techniques **6**:322
cultural diversity loss **6**:328
current rainforest areas **6**:323, **6**:324*T*
deforestation risks **6**:327–328, **6**:328*T*
diversity implications **6**:327–328
eco-floristic zone concept **6**:329
ecosystem resilience loss **6**:328–329, **6**:329*F*
forest land change and transition **6**:325–327, **6**:326*F*, **6**:326*T*
historic rainforest areas **6**:323, **6**:324*T*
species richness loss **6**:327–328, **6**:328*T*
transition matrices **6**:325–327, **6**:326*F*, **6**:326*T*
net primary production (NPP) **1**:560*T*
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
rain forest types **6**:305–306
seed dispersal and germination **1**:708
soils **6**:123*T*
South America
distribution and structure **3**:110, **3**:111*F*
functional aspects **3**:112
termites **4**:430*T*, **4**:432*T*
weed management
biodiversity effects **4**:91–92

forests (*continued*)
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
wetland ecosystems **7**:371, **7**:385*F*
see also seasonally dry tropical forests
Forest Stewardship Council **4**:330, **7**:20–21
fork-marked lemur **6**:251*T*
formaldehyde (CH_2O) **1**:282*T*
formamidines **5**:694
Formenkreis theory **7**:46
former Soviet Union
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:519, **6**:510*T*
indigenous peoples **4**:269*T*
vegetation surveys **3**:341–344*T*
Formicariidae **6**:308*T*
Formicidae **4**:178, **5**:460*T*, **5**:463*T*, **5**:464*T*
Formicoidea **1**:710
Forster, Johann Reinhold **1**:577–578
fosamine **4**:90*T*
Fossey, Dian **3**:317*F*
fossil diversity curve **5**:157*F*
fossil fuels **1**:283–284, **1**:675*F*, **1**:678, **1**:679*F*, **3**:441–442, **4**:628–629
see also biofuels
fossil nomenclature **5**:555
fossil record **3:537–544**
angiosperms
challenges **5**:615
early angiosperms **5**:617–619
general discussion **5**:615–617
nineteenth century research **5**:613–614
potential angiosperm sister groups **5**:615–617, **5**:616*F*, **5**:617*F*, **5**:618*F*
arachnids **1**:204–205
basic concepts
abundance patterns **3**:538–539
characteristics **3**:537–538
completeness **3**:537–538
correlation and dating assessments **3**:539
preservation quality **3**:538–539
sampling variations **3**:539
stratigraphic ranges **3**:537–538
beetles **1**:326
biodiversity generation **1**:430
biodiversity measures **1**:404
biodiversity origins **1:491–496**
background information **1**:411
community interpretation and analysis **1**:413–414
mass extinctions **1**:412–413, **1**:413*F*
Phanerozoic Eon
general discussion **1**:491
marine diversity timeline **1**:492*F*, **3**:419*F*
mass extinctions **1**:412–413, **1**:413*F*, **1**:493, **1**:493*F*, **1**:494*F*, **1**:494*T*, **3**:419*F*
originations **1**:495, **1**:495*F*, **3**:425*F*
taxa diversity **1**:411*F*, **1**:412, **1**:412*F*
taxa diversity **1**:411–412, **1**:411*F*, **1**:412, **1**:412*F*
usefulness **1**:413
brain size evolution **4**:698
community dynamics **6:414–419**
community age **6**:415*F*, **6**:416
current research areas **6**:417–418
range and composition **6**:414–416, **6**:415*F*
research summary **6**:418
residence times **6**:414, **6**:415*F*
residents and transients **6**:416–417, **6**:417*F*
taxon turnover **6**:415*F*, **6**:416, **6**:416*F*
community interpretation and analysis **1**:413–414
crocodilians **6**:401
crustaceans **2**:403–407, **2**:413–415
differentiation **2**:514–515
early mammals **3**:188–189
evolutionary rates **2**:54–55
extinctions **3:413–425**
assessment challenges
accurate range estimates **3**:415
biased preservation **3**:414–415
data collection **3**:414
phylogenetic relationships **3**:415
preservation unevenness **3**:415
pseudoextinction **3**:415
speciation–extinction rate relationship **3**:415
taxon-level bias **3**:414
time measures **3**:413–414
background extinctions
evolutionary faunas **3**:423–424
extinction rate declines **3**:423–424
mass extinctions **5**:2–3
turnover **3**:423
crustaceans **2**:413–415
diversification impacts **3**:424–425, **3**:424*F*, **3**:425*F*
first-contact extinctions
basic concepts **5**:365
Madagascar **5**:365
New Zealand **5**:366
Pacific islands **5**:365–366
historical extinctions
extinction paradoxes **5**:367
Hawaii **5**:367
New Zealand **5**:366–367
number of species **5**:368*T*
late Quaternary extinction event (LQEE) **4:721–732**
mass extinctions
basic concepts **3**:415–418
biodiversity loss **4**:656
biogeographic impacts **3**:421–422
causal factors **3**:418
climatic changes **3**:422–423
extinction theories **3**:418–421
long-term consequences **3**:421
major mass extinctions **3**:416*T*
modern extinctions **5:362–372**
periodicity **3**:422–423, **3**:424*F*
quantification **3**:417
modern extinctions **5:362–372**
Africa **5**:364–365
Australia **5**:363–364, **5**:364*F*
current on-going extinctions **5**:371
density-related traits **5**:370–371
extinction cascades **5**:367–368
first-contact extinctions **5**:365
geographical bias **5**:369
historical extinctions **5**:366–367, **5**:368*T*
inadequate documentation **5**:368–369
inherent biological limitations **5**:369
introduced species **5**:370
methodological limitations **5**:369
North America **5**:363
number of species **5**:368, **5**:368*T*
population dynamics **5**:370
postglacial extinctions **5**:363
research challenges **5**:368
species lifespans **5**:362–363, **5**:362*T*
taxonomic and habitat bias **5**:369
theoretical perspectives **5**:369–370
time factors **5**:371
uneven documentation **5**:369
research background **3**:413
general discussion **3**:537
Hymenoptera **4**:178
hyperthermophiles **7**:164
indicator species **4**:256–257, **4**:258*F*
insect-pollination angiosperm-radiation hypothesis **6**:48
Isoptera **4**:410–411
late Quaternary extinction event (LQEE) **4:721–732**
background information **4**:721
causal factors
anthropogenic models **4**:731–732
climatic models **4**:732
global patterns **4**:731–732
comparative studies
Africa **4**:658*F*, **4**:729–730
analytical methods **4**:721–725
Australia **4**:658*F*, **4**:727–728
Eurasia **4**:729–730, **4**:729*F*
Madagascar **4**:658*F*, **4**:728–729
North America **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
Oceania **4**:730–731
South America **4**:726–727, **4**:727*T*
time-transgressive sequence **4**:722*F*, **4**:723*F*
Lepidoptera **5**:389
lizards **6**:402
mammals **4**:684
marine invertebrates **4**:389–391, **4**:390*F*
mass extinctions **5:156–166**
basic concepts **3**:415–418
biogeographic impacts **3**:421–422
causal factors **3**:418
climatic changes **3**:422–423
crustaceans **2**:413–415
data analysis and interpretation **5**:158*F*, **5**:159*F*, **5**:160*F*
extinction rates **4**:656, **4**:657*F*
extinction theories **3**:418–421
first-contact extinctions

basic concepts **5**:365
Madagascar **5**:365
New Zealand **5**:366
Pacific islands **5**:365–366
fossil diversity curve **5**:157*F*
historical extinctions
extinction paradoxes **5**:367
Hawaii **5**:367
New Zealand **5**:366–367
number of species **5**:368*T*
historical research **5**:156–157, **5**:157*F*
kill curve **5**:160–163, **5**:163*F*
long-term consequences **3**:421
magnitude measures
end-of-geologic-period extinctions **5**:158, **5**:161*F*
end-Permian mass extinction **5**:158, **5**:161*F*
extinction events **5**:158–160, **5**:161*F*
extinction intensities **5**:162*F*
general discussion **5**:158
major mass extinctions **3**:416*T*
modern diversity crisis **5**:165–166
modern extinctions **5:362–372**
Africa **5**:364–365
Australia **5**:363–364, **5**:364*F*
current on-going extinctions **5**:371
extinction cascades **5**:367–368
first-contact extinctions **5**:365
historical extinctions **5**:365, **5**:368*T*
North America **5**:363
postglacial extinctions **5**:363
research challenges **5**:368
species lifespans **5**:362–363, **5**:362*T*
theoretical perspectives **5**:369–370
periodicity **3**:422–423, **3**:424*F*
periodicity hypothesis **5**:160, **5**:162*F*
Phanerozoic Eon **1**:412–413, **1**:413*F*
quantification **3**:417
recoveries
characteristics **5**:164
delayed diversification **5**:164
disaster species **5**:164–165
large-scale rebounds **5**:164, **5**:165*F*
Lazarus species **5**:164–165
research summary **5**:166
scenarios
abrupt extinction **5**:157
general discussion **5**:157
gradual extinction **5**:157
stepwise extinction **5**:157
self-organized criticality **5**:160–163
victim–survivor patterns **5**:163–164
paleodiversity **3:678–690**
Archean era **3**:682
biodiversity drivers
climate change **3**:689
extinction events **3**:689–690
global heterogeneity **3**:688–689
plate tectonics **3**:688–689
evolutionary development
morphological disparities **3**:679
species richness **3**:678–679
fossil biodiversity
completeness measures **3**:680, **3**:681*F*
rock record completeness **3**:679–680
sampling challenges **3**:679
taxonomic reconstruction measures **3**:680–682
microbial biodiversity **3**:682
Phanerozoic marine biodiversity
Cambrian explosion **3**:683, **3**:684*F*
chordate diversity **3**:684–685
invertebrate diversity **3**:683–684, **3**:685*F*, **3**:686*F*
neoproterozoic animals **3**:682–683
taxonomic turnover **3**:685–686
Phanerozoic terrestrial biodiversity
animal diversity **3**:687–688, **3**:687*F*
chordates **3**:688
invertebrate diversity **3**:687–688, **3**:687*F*
plant diversity **3**:686–687, **3**:687*F*
Proterozoic era **3**:682, **3**:683*F*
research approaches **3**:678
salmon **6**:522–523
snakes **6**:402
Southern Hemisphere ecosystems **6**:615–616
species duration **5**:434–435, **5**:434*T*
species richness gradients **4**:614*F*
statistical estimation models
diversity trends **3**:539–541
nonparametric estimates **3**:541–542
parametric estimates **3**:542–543
rarefaction measurement approach **3**:540–541, **3**:540*F*, **3**:541*F*
research summary **3**:543–544
sampling standardization **3**:540–541, **3**:540*F*, **3**:541*F*
turnover-based estimates **3**:543
taxa diversity **1**:411–412, **1**:411*F*, **1**:412, **1**:412*F*
terrestrial plant communities **6:65–75**
abundance patterns **6**:73, **6**:74*F*
allocation strategies **6**:68–69
body size and shape influences **6**:67–68, **6**:68*F*
climatic influences **6**:69
coevolutionary relationships **6**:69–70
competition **6**:67
equilibrium theories **6**:71–72
evolutionary development **6**:65–67, **6**:66*F*
extinction rates **6**:73*F*
latitudinal gradients **6**:72–73
niche diversification **6**:67
origination rates **6**:71–72, **6**:72*F*
physiological processes **6**:69
reproductive strategies **6**:68–69
research background **6**:65
research summary **6**:73–74
stability and resilience **6**:71
succession **6**:70–71
turtles **6**:400–401
usefulness **1**:413
vertebrate history **7**:334–335
Foudia rubra **7**:188
founder crops **2**:658
founder effect speciation **6**:624, **6**:634–635
founder-flush speciation model **6**:624
fountain grass **3**:208–209
foxes **1**:80, **1**:631–632, **1**:685, **1**:709, **2**:590*T*, **4**:693, **6**:239
foxglove **5**:705–706
Fox, Sidney **5**:631
fox sparrow **1**:605, **5**:312
foxtail millet **2**:659, **2**:663, **3**:32*F*, **3**:131–132
Fragaria xananassa **2**:383
Fragilaria crotonensis **6**:33*F*
fragmented forests **6**:325–327, **6**:326*F*, **6**:326*T*
frailejones **4**:101–102, **4**:102*F*
France
botanical gardens **6**:519
Convention on Biological Diversity (1930) **2**:313
deforestation **1**:677*F*
invasive species **4**:345–346
natural history museums and institutions **5**:408*F*, **5**:409*F*, **5**:411*F*
nitrogen deposition **5**:529*F*
species richness **1**:499*T*
subterranean ecosystems
Ariège **7**:60
Rhone River **7**:59
vegetation surveys **3**:341–344*T*
Francoaceae **5**:216–217
Francois' langur **6**:256–258*T*
Francolinus francolinus **5**:463*T*
Frankel, Otto **2**:263
Frankia **6**:448, **6**:581–582, **6**:582*F*
Franklinia alatamaha **3**:214, **6**:77
Franzen, Jonathan **4**:654
Fraser fir **4**:361
Frasnian-Famennian event **3**:416*T*, **3**:419–420, **5**:161*F*
Frasnian Stage **5**:169–170
Frassasi Caves (Italy) **7**:51
Frattercula spp. **3**:256
Fraxinus **4**:361, **7**:125
Fraxinus americanus **1**:610
Fraxinus lanuginosa **1**:271
Fraxinus mandschurica **1**:271–272
Fraxinus ornus **6**:96–97
Fraxinus pennsylvanica **6**:459
Fraxinus sieboldiana **1**:271
Fraxinus spaethiana **1**:271
free-living algae **3**:517–518
free-living saprobes **6**:448
free riders **2**:154
Fregata aquila **7**:185
French Guiana
botanical gardens **6**:519
land cover classifications **6**:607*T*
nonvascular epiphytes **3**:521*T*
tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*
urbanization **4**:560*F*
French Polynesia **6**:507*T*
freshwater ecosystems **3:560–569**
acidic and mercury deposition **1:1–14**
acidic deposition–mercury concentration linkage **1**:13
aluminum concentrations **1**:9–10
aquatic biota response **1**:9–10, **1**:10*T*, **1**:11*F*

freshwater ecosystems (*continued*)
emission occurrences and sources **1**:1–5
geographic distribution **1**:1–5, **1**:3*F*
mercury contamination **1**:10–13, **1**:11*F*, **1**:12*F*
mercury emissions and deposition **1**:2–5, **1**:4*F*
pH changes **1**:9–10, **1**:10*T*
research background **1**:1
schematic diagram **1**:2*F*
surface water acidification **1**:8–9, **1**:10*T*, **1**:144–145
Antarctica **1**:183
anthropogenic impacts **3:570–586**, **3**:572–574
biodiversity **3**:566–569, **3**:568*F*
climate change effects **6**:150–151
conservation status **3**:171–173*T*, **3**:174
crustaceans **2**:397–399
dissolved oxygen **3**:351, **3**:562
ecosystem function measurements **3:64–71**
basic concepts **3**:64
biogeochemical cycles **3**:68–70
comparison studies **3**:64–65
new and regenerated production **3**:69–70
nutrient cycling **3**:68–70
primary production
carbon-14 (^{14}C) measurements **3**:66
carbon-based measurements **3**:65–66
data interpretation and analysis **3**:66–67
fluorescence measurements **3**:66
measurement approaches **3**:65
oxygen-based measurements **3**:65
secondary production
bioenergetic analysis **3**:67
biomass accrual **3**:67
birth and death rate analysis **3**:67–68
egg ratio analysis **3**:68
functional type analysis **3**:68
instar analysis **3**:68
measurement approaches **3**:67
stable isotope studies **3**:70–71
ecosystem services **3**:96–98
endangered freshwater invertebrates **3:176–187**
American Southeast rivers **3**:184, **3**:185*F*
anthropogenic impacts
direct harvests **3**:181
global climate change **3**:181–182
habitat destruction/degradation **3**:180–181, **3**:180*F*
introduced species **3**:181
pollution **3**:181, **3**:181*F*
causal factors **3**:177, **3**:177*F*
desert spring communities **3**:186, **3**:187*F*
diversity hot spots **3**:176
global conservation status **3**:176–177, **3**:176*T*, **3**:177*F*
global distribution patterns **3**:182–184, **3**:182*T*
global diversity **3**:176
number of species **3**:182–184, **3**:182*T*
protection strategies **3**:184–187
small ranges
causal factors **3**:178
conservation status **3**:177*F*
dispersal ability **3**:178*F*
high endemic hot spots **3**:178–179
high endemic species **3**:179, **3**:180*F*
Lake Baikal **3**:179, **3**:183*F*
narrowly endemic species **3**:177–178, **3**:177*F*, **3**:179
vulnerability factors
small ranges **3**:177–178, **3**:177*F*
sparse populations **3**:179–180
eutrophication
aquatic vascular plants **3**:359–360
macroalgae **3**:358
macrophytes **3**:359–360
microfauna **3**:361–363
fish
acidic deposition–mercury concentration linkage **1**:13
acidification effects **3**:584–585, **3**:584*F*
African/Ethiopian region **3**:472–473
aquaculture **1**:190*T*
Australian region **3**:473
biodiversity issues **3**:444
energy flow **3**:243*F*
global biodiversity **3**:471*F*, **3**:472
habitats **3**:472
introduced species **4**:378
mercury contamination **1**:10–13, **1**:12*F*
migratory patterns **5**:315
Nearctic region **3**:472
Neotropical region **3**:472
North America **5**:565
Oriental region **3**:473
Palearctic region **3**:472
Succulent Karoo **1**:487*T*
surface water acidification **1**:9–10, **1**:10*T*, **1**:11*F*
threatened species **3**:578–579, **3**:579*T*
geographic distribution **4**:370*T*
habitat destruction **3**:164–165
habitat diversity **3**:562–565, **3**:564*F*, **3**:565*F*, **3**:570–571
habitat loss and degradation **1**:510–512, **1**:511*F*
interactive regulatory mechanisms **3**:565–566
invertebrates **4:369–378**
background information **4**:369–370
ecological functions and processes
characteristics **4**:375
decomposition and recycling **4**:375
material transfer and uptake **4**:375–376
primary production **4**:376
global patterns **4**:373
influencing factors
disturbances **4**:375
endemism **4**:374–375
habitat heterogeneity **4**:375
isolation factors **4**:374–375
origins **4**:374
research background **4**:373–374
species interactions **4**:375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats
altered water regimes **4**:377–378
biodiversity loss **4**:376–377
exotic species invasions **4**:378
habitat depletion and loss **4**:376–377
pollution **4**:377
reduced water quality **4**:377
number of species **4**:370–371, **4**:371*T*
river continuum concept (RCC) **4**:373*F*
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
lakes and ponds **4:458–466**
acidification **3**:572–574, **3**:577–578, **3**:584–585, **3**:584*F*
ancient lakes species flocks
characteristics **4**:463–464, **4**:463*T*
fish species flocks **4**:463–464
invertebrates **4**:464, **4**:464*T*
anthropogenic impacts
agricultural and industrial chemicals **3**:580–581, **3**:581*F*
ecosystem functions and processes **3**:585
eutrophication **3**:581–582
organic pollution **3**:580–581, **3**:581*F*
research summary **3**:585–586
biological quality
species declines **3**:578–579, **3**:579*T*
species variability **3**:578–579, **3**:578*F*
threatened species **3**:579–580
characteristics
biota origins and peculiarities **4**:458–459
classifications **4**:458–459
latitudinal gradients **4**:459
species richness–lake size relationship **4**:459–460
species richness–morphometry relationship **4**:460
vertical distribution **4**:459
diversity estimations
fish species **3**:605–606, **3**:606*F*, **3**:607*F*, **4**:460–461, **4**:461*T*
freshwater sediment species **4**:460, **4**:461*T*
general discussion **4**:460
plankton and microorganisms **4**:460
endangered freshwater invertebrates **3:176–187**
causal factors **3**:177, **3**:177*F*
diversity hot spots **3**:176
global conservation status **3**:176–177, **3**:176*T*, **3**:177*F*
global diversity **3**:176
vulnerability factors **3**:177–178
epilimnion **3**:560, **4**:374*F*
eutrophication impacts
acidification **3**:584–585, **3**:584*F*
agricultural and industrial chemicals **3**:580–581, **3**:581*F*
biological consequences **3**:581–582
clean-up results **5**:24
comparison studies **3**:582

ecosystem functions and processes **3**:585
ecosystem structure and function **3**:581–582
food web manipulation **3**:582–583
harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*, **5**:25
Lake Fure, Denmark **3**:575*F*, **3**:581–582
long-term human influences **3**:354–356, **3**:356*F*, **3**:357*F*
organic pollution **3**:575, **3**:575*F*, **3**:580–581, **3**:581*F*
pollution **4**:465–466
resource and predatory controls **3**:582–583
Secchi transparency **3**:581–582
species diversity–lake functioning **4**:462–463
species richness **3**:583–584
species shifts **3**:352–354, **3**:353*F*
whole-lake studies **3**:582
geographic distribution **4**:370*T*
habitat loss and degradation **7**:371*T*
hypolimnion **3**:560, **4**:374*F*
invertebrates **4:369–378**
altered water regimes **4**:377–378
background information **4**:369–370
biodiversity loss **4**:376–377
decomposition and recycling functions **4**:375
disturbances **4**:375
ecological functions and processes **4**:375
endemism **4**:374–375
exotic species invasions **4**:378
global patterns **4**:373
habitat depletion and loss **4**:376–377
habitat heterogeneity **4**:375
influencing factors **4**:373–374
isolation factors **4**:374–375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats **4**:376–377
material transfer and uptake **4**:375–376
number of species **4**:370–371, **4**:371*T*
origins **4**:374
pollution **4**:377
primary production **4**:376
reduced water quality **4**:377
research background **4**:373–374
river continuum concept (RCC) **4**:373*F*
species interactions **4**:375
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
littoral zone **4**:374*F*
major threats
competition for water **4**:464
fisheries practices **4**:465
habitat alteration **4**:464–465
introduced species **4**:378, **4**:465
pollution **4**:465–466
metalimnion **3**:560, **4**:374*F*
nonextremeophilic Archaea **1**:224*T*
nutrient enrichment impacts **1**:95–97, **1**:96*F*, **3**:351, **3**:575*F*, **3**:576–577, **3**:576*F*
protected area systems **5**:444*T*, **6**:494–498, **6**:496*F*, **6**:497*F*
species diversity–lake functioning
biodiversity–ecosystem stability relationship **4**:462
biological productivity **4**:462–463
characteristics **4**:461–462
eutrophication **4**:462–463
food webs **4**:461–462
intra- and interspecies communication systems **4**:462
top-down controls **4**:462
species richness **3**:572*T*, **3**:583–584
lateral and vertical boundaries **3**:565*F*
light attenuation **3**:560
mean composition **3**:562*T*
mean water depth **3**:561*F*
Near East **5**:470–471, **5**:470*F*, **5**:470*T*, **5**:471*F*
number of lakes **3**:561*F*
nutrient loading gradient **3**:568, **3**:568*F*
organic matter **3**:562, **3**:564*F*, **3**:565–566
photosynthesis **3**:560, **3**:564*F*
physical characteristics **3**:560–562
pollution **1**:512
primary productivity **3**:563*T*, **3**:566–569, **3**:568*F*
protected area systems **5**:444*T*, **6**:494–498, **6**:496*F*, **6**:497*F*
river–floodplain connectivity **7**:151
Southeast Asia **1**:510–512, **1**:511*F*, **1**:512
species richness
freshwaters **3**:571–572, **3**:571*T*
freshwater sediments **3**:571–572, **3**:572*T*
lakes and ponds **3**:572*T*, **3**:583–584
site characteristics and ranges **3**:572*T*
trophic levels **3**:562, **3**:563*T*
usefulness **3**:570
water storage **3**:560–562, **3**:561*T*
freshwater eels **5**:315
freshwater marshes **7**:370*T*, **7**:371*T*
freshwater mussels **3**:411–412
fresh water supplies **5**:351
freshwater swamps **7**:370*T*
frilled lizards **7**:223
Fringilla coelebs **1**:618
Fringilla teydea **1**:618
Fringillidae **4**:33, **4**:34*F*, **6**:308*T*
frit flies **3**:489–491*T*, **3**:493
frogfish **3**:466–467
frogs
alpine ecosystems **1**:154–155
Australia **1**:302*F*
biodiversity
conservation programs **1**:176
current biodiversity **1**:169–170
geographic distribution **3**:691–693
historical biogeography **1**:169–170
characteristics **7**:339
critical pH levels **6**:153*F*
ecological genetics **2**:719
environmental threats
chytrid fungus **2**:86–87, **4**:664
disease **3**:166
functional anatomy **1**:170–171
habitat loss and fragmentation **4**:516–517
introduced parasites–species decline relationships **2**:528*T*, **2**:529–530, **2**:530*F*
larval morphology and ecology **1**:174–175
life histories **1**:172
mangrove ecosystems **5**:13
morphology **1**:170–171
phenological shifts **5**:723
postlarval behavior and ecology **1**:175–176
reproductive biology **1**:172, **1**:173*F*
river ecosystems **6**:479–480
South America **3**:115
special features **1**:171
species richness **1**:170*T*
tropical rain forests **1**:263, **6**:307
urbanization threats **4**:564
wetland ecosystems **7**:389–391, **7**:389*F*
frost resistance **3**:83*F*
frozen zoos **7**:477
fructofuranosidase **5**:269*T*
fructose **6**:113*T*
frugivores **6**:50–51
fruit bats **4**:691–692, **5**:14
fruit flies
characteristics **3**:492
classifications **3**:489–491*T*
Near East ecosystems **5**:469–470
fruits
East Asia **2**:663
Europe **2**:662–663
fruit and seed dispersal
plant–animal interactions
ants **6**:50
beetles **6**:50
birds **6**:50
dispersal syndromes **6**:50, **6**:51*T*
disperser types **6**:50
evolutionary dead ends **6**:52
fish **6**:50
frugivores **6**:50–51
importance **6**:49–50
insects **6**:50
mammals **6**:50
plant adaptations **6**:50
reptiles **6**:50
seed predators **6**:51–52
species persistence **6**:61–62
vertebrates **6**:50
geographic distribution **3**:131*T*
global food usage **3**:130*T*
Mediterranean Basin **2**:662–663
Neolithic agriculture **2**:658
North and South America **2**:663–664
Southeast Asia **2**:663
southwest Asia **2**:662–663
F-statistical method **4**:510
Fucus serratus **5**:36–37
Fucus vesiculosus **3**:365, **4**:10
fuelwood collection **2**:453

fugu puffer fish **1**:589
Fugu rubripes **1**:589
Fulica atra **3**:365–366
fumigants **5**:694
functional diversity (FD) **3:587–596**
  basic concepts
    definition **3**:587
    diversity versus composition **3**:588
    ecosystem processes **3**:587–588
    functional groups **3**:588
  ecosystem function measurements
    biodiversity **3**:84–86, **3**:84*F*
    plant functional traits **3**:85*F*, **3**:85*T*, **3**:86
    positive-feedback switches **3**:86–87, **3**:86*F*
    taxonomic diversity **3**:84
  functional diversity (FD) measures **3:597–608**
    basic concepts **3**:587, **3**:597–598, **3**:598*F*
    existing measures
      combined functional richness–function divergence **3**:602–603, **3**:603*F*
      functional divergence **3**:601–602, **3**:602*F*
      functional evenness **3**:601, **3**:601*F*
      functional richness **3**:598–601, **3**:599*F*, **3**:599*T*, **3**:600*F*, **3**:601*F*
      general discussion **3**:598–601
      Rao's quadratic entropy **3**:603, **3**:603*F*, **3**:606–608, **3**:607*F*
      recommended measures **3**:603–604
    key criteria **3**:598, **3**:599*T*
    purpose **3**:597
    research applications
      ecosystem multifunctionality–functional divergence measures **3**:604–605, **3**:605*F*
      general discussion **3**:604–605
      grassland ecosystems **3**:604–605, **3**:605*F*
      lacustrine fish communities **3**:605–606, **3**:606*F*, **3**:607*F*
      phosphorus gradient–hierarchical partitioning measures **3**:606–608, **3**:607*F*
      species richness–functional diversity measures **3**:605–606, **3**:606*F*, **3**:607*F*
  Global Ocean Sampling (GOS) program **5**:295
  historical research **3**:588
  pest control–predator diversity relationship **1**:378
  potential effects
    diversity–productivity relationship
      experimental studies **3**:592–594, **3**:592*F*, **3**:593*F*
      niche differentiation model **3**:589, **3**:591*F*
      sampling effect model **3**:589–592, **3**:590*F*
    diversity–stability relationship
      basic concepts **3**:594
      experimental studies **3**:594–595, **3**:595*F*
      observational studies **3**:594–595, **3**:595*F*
  research summary **3**:595
  taxonomic versus functional diversity **2:648–656**
    diversity estimations
      background information **2**:650
      species evenness **2**:650, **2**:651*T*
      species richness **2**:650, **2**:651*T*
      species richness–species evenness relationship **2**:650–651
    ecological coherence **2**:655–656
    historical perspective **2**:648
    linkages
      basic processes **2**:654–655
      C3 and C4 grasses **2**:654–655
      exotic species invasions **2**:655
      island biogeography **2**:652*F*, **2**:655, **2**:655*T*
    natural selection **2**:656
    research approaches **2**:648, **2**:649*T*
    species classifications
      characteristics **2**:649–650
      ecological niche model **2**:649–650, **2**:649*F*
      functional groupings **2**:649*F*, **2**:650
    stability dynamics
      basic concepts **2**:652–653
      local asymptotic stability **2**:652–653
      process-based models **2**:654
      species interaction models **2**:653–654
      transient dynamic behavior **2**:653, **2**:653*F*
    steady-state dynamics
      basic concepts **2**:651–652, **2**:651*F*
      colonization rates **2**:651–652, **2**:652*F*
      community structure **2**:652
      ecosystem processes **2**:652, **2**:652*F*
      extinction rates **2**:651–652, **2**:652*F*
      number of species **2**:651–652, **2**:652*F*
      population densities **2**:652
      species composition **2**:652
    systems theory
      basic concepts **2**:651–652
      stability dynamics **2**:652–653
      steady-state dynamics **2**:651–652, **2**:651*F*
  wetland ecosystems
    acidic environments **7**:395–397, **7**:396*F*
    adaptations **7**:394–395
    anoxic environments **7**:394–395
    carnivorous plants **7**:395–397, **7**:396*F*
    drought conditions **7**:395
    nutrient availability **7**:395–397
    reproductive adaptations **7**:397
    salinity/salt tolerance **7**:395
functional groups **3:609–623**
  assemblage evolution **3**:611–612, **3**:612*F*
  community structure and function **3**:617, **3**:618*F*
  convergent evolution
    mammals **4**:700, **4**:701*T*
    mobile organisms **3**:613, **3**:614*F*, **3**:615*F*
    research background **3**:612–613
    sessile organisms **3**:613–617, **3**:615*F*, **3**:616*F*
  coral reef ecosystems
    breaking points **3**:621, **3**:622*F*
    characteristics **3**:617–619
    ecosystem diversity and stability **3**:618–619, **3**:619*T*
    ecosystem redundancy **3**:620–621
    equivalent species **3**:620–621, **3**:621*F*
    keystone species **3**:621–623, **3**:622*F*
    species diversity
      herbivores **3**:619–620, **3**:620*T*
      primary producers **3**:620
      trophic levels **3**:619–620
    structural characteristics **3**:619
  functional stasis **3**:611–612
  microorganisms
    free-living saprobes **6**:448
    importance **6**:448
    soil food webs **6**:448–449, **6**:449*F*
    symbionts **6**:449–450, **6**:450*F*
  origins **3**:611–612
  relatedness versus morphological/phylogenetic similarity **3**:609–611, **3**:610*F*
  research background **3**:609
  research summary **3**:623
functional/interaction food webs **3**:499*F*, **3**:501–502
functional landscapes **4**:531–533
fungal dieback **2**:528*T*
fungal opsins **6**:281*F*
fungi **3:624–640**
  air pollution impacts **1**:142*F*
  alpine ecosystems **1**:155
  Antarctica **1**:184*T*
  Arctic ecosystems **1**:231*T*
  background information **3**:624
  biodiversity **3**:625–626, **6**:585–586
  characteristics **2**:524*T*, **3**:624–625, **5**:289–291
  chytrid fungus **3**:166
  domesticated ruminants **1**:710
  endophytic fungi **2**:429
  environmental role
    characteristics **3**:632–633
    human-related parasites
      classifications **3**:633–634
      mycetismus **3**:634, **3**:635*T*
      mycoallergies **3**:635
      mycosis **3**:633–634, **3**:634*T*
      mycotoxicosis **3**:634–635
    mutualism
      characteristics **3**:636
      endophytes **3**:636
      lichens **3**:636–637, **3**:637*F*
      mycorrhizae **2**:391, **3**:636, **3**:636*F*
    nonhuman animals **3**:635–636
    plant parasites
      crop pathology **3**:633, **3**:633*T*
      disease triangle **3**:633, **3**:633*F*
      forest pathology **3**:633, **3**:633*T*
    saprophytes **3**:632–633
  evolutionary development **2**:644
  food webs **7**:266*F*, **7**:267*T*
  forest management impact **1**:451

freshwater sediments **3**:572*T*
functional role **3**:625
global species richness **3**:711
grassland ecosystems **4**:4
growth processes
exoenzymes **3**:626
Hyphae **3**:626, **3**:626*F*
yeasts **3**:626
herbicide use effects **4**:93
industrial applications
antibiotics **3**:638
biopulping/bioremediation **3**:639
bread baking **3**:638
ethanol production **3**:638
fermentation products **3**:638–639
mushroom cultivation **3**:637–638, **3**:637*F*, **3**:637*T*
wine and beer making **3**:638
known and estimated species **4**:384*T*
lake and pond ecosystems **4**:461*T*
life cycles **3**:626, **3**:627*F*
molecular level diversity **2**:635*T*
mutualism **2**:146
mycorrhizal associations **4**:74, **4**:174, **6**:121–122
nematodes **7**:430–431
nomenclature code **5**:550–551, **5**:553*T*, **5**:555–556
nonseed gene banks **3**:652, **3**:652*T*
number of species **5**:254*T*, **6**:585–586, **6**:586*T*
phyla
Ascomycota **3**:628–629, **3**:628*F*
Basidiomycota **3**:628*F*, **3**:629–630, **3**:629*F*
characteristics **3**:626–627
Chytridiomycota **3**:627
deuteromycetes **3**:629*F*, **3**:630
excluded taxa
Dictyosteliomycota (cellular slime molds) **3**:631
general discussion **3**:630
Myxomycota (true slime molds) **3**:630–631, **3**:630*F*, **3**:631*F*
Oomycota **3**:630, **3**:630*F*
Plasmodiophoromycota (endoparasitic slime molds) **3**:631
Glomeromycota **3**:627–628
Zygomycota **3**:627, **3**:628*F*
phylogenetic relationships **5**:289*F*, **5**:290*F*
research areas **3**:625–626
slash-and-burn systems **6**:556
soil food webs **6**:448–449, **6**:449*F*
spatiotemporal relationships **6**:447, **6**:447*F*
species estimates **3**:710*T*
surface area and reproduction **3**:631–632, **3**:632*F*
symbiotic relationships **4**:418–419, **4**:419*F*
taxonomy **3**:625*T*
temperature tolerance ranges **4**:114
termites **4**:418–419, **4**:419*F*
threatened and endangered species **3**:639–640, **3**:639*F*
fungicides
bioprospecting **1**:590
inorganic fungicides **5**:694–695
Near East ecosystems **5**:473–475
organic fungicides
benefits **5**:695
dicarboximides **5**:695
dinitrophenols **5**:695, **5**:699*T*
dithiocarbamates **5**:695, **5**:699*T*
organometallics **5**:695
phthalamides **5**:695
substituted aromatics **5**:695
systemic compounds
acylalanines **5**:695
2-aminopyrimidines **5**:695
benzimidazoles **5**:695
ergosterol biosynthesis inhibitors (EBIs) **5**:695
general discussion **5**:695
organophosphates **5**:695
oxathiins **5**:695
phenylamide **5**:695
pyrimidines **5**:695
quinones **5**:695
thiophantes **5**:695
triazines **5**:695
fungivory **5**:394
fungus gnats **3**:489–491*T*, **3**:493
Funisciurus anerythrus **5**:603
funnel web spiders **1**:206
Furcula **5**:617–619
Fure (Lake) **3**:575*F*, **3**:581–582
Furipteridae **6**:309*T*
Furnariidae **4**:33, **4**:34*F*, **6**:308*T*
fur seals **5**:65*T*, **5**:87*T*, **5**:314
fur trade **3**:191–192
Fuselloviridae **5**:130*F*
futonil **5**:695
fynbos vegetation **2**:300, **5**:213*T*, **5**:219

G

Gabon
botanical gardens **6**:507*T*
tropical rain forests **6**:314–315
Gaboon viper **6**:409
gadoleic acid **7**:32
Gadus morhua **3**:123*T*, **3**:478, **4**:516, **5**:35, **5**:315, **6**:233, **7**:262
gadwells **3**:365–366
Galacticoidea **5**:387–388*T*
galactose **6**:113*T*
Galagidae **4**:692–693, **6**:250–252, **6**:251*T*
Galagoides rondoensis **6**:256–258*T*
Galagoides zanzibaricus zanzibaricus **6**:256–258*T*
galagos **4**:692, **6**:251*T*
Galapagos cormorant **1**:606
Galapagos finches
adaptation rates **2**:140
adaptive radiation **1**:27–28, **1**:28*F*
character displacement **6**:720
ecological character displacement **4**:46, **4**:47*F*
evolutionary development **1**:618–619
fluctuating selection **2**:718–719
hybridization **1**:26–27
inbreeding **2**:266
interspecies competition **2**:141
Galapagos fur seals **5**:87*T*
Galapagos Islands
adaptive radiation **5**:593–594
conservation status **3**:171–173*T*
coral bleaching **3**:160–161
El Niño Southern Oscillation (ENSO) events **3**:160–161
endemism **5**:592*T*
invasive species **4**:352–353
late Quaternary extinction event (LQEE) **4**:723*F*
protected area systems **6**:493–494
species diversification **5**:594
Galapagos seastar **3**:201–202
galaxiids **3**:464–465
Galbulidae **6**:308*T*
Galdikas, Birute **3**:317*F*
Galeodes caspius **1**:213
Galerina autumnalis **3**:635*T*
Galileo **5**:404–405
gall flies **4**:363
Gallicolumba erythroptera **7**:184
Galliformes **6**:308*T*, **7**:181–182
Gallinula chloropus **3**:365–366
Gallinula martinica **2**:549
Gallinula pacifica **7**:184
Gallirallus australis **5**:370, **7**:185
Gallirallus lafresnayus **7**:184
Gallirallus owstoni **7**:188
Gallirallus sylvestris **7**:188
gall midges **3**:489–491*T*, **3**:492
gallotannins **5**:712*T*
galls **3**:633*T*, **6**:40
Gallus gallus **1**:640, **2**:636*F*, **6**:361*T*
gall wasps **4**:181, **4**:349, **6**:40
Gamba grass **6**:93
Gambelia wislizenii **6**:406
Gambel's quail **2**:475*F*
Gambia **6**:507*T*
Gambierdiscus **1**:472–473
Gambierdiscus toxicus **1**:475, **1**:480, **6**:131
Gambusia affinis **3**:167, **4**:361, **5**:668
Gambusia georgei **7**:58
game birds **7**:181–182
game theory **4**:632–633
gamma-aminobutyric acid (GABA) **4**:301
gamma diversity **1**:403, **2**:439–440, **2**:650, **4**:339–340, **6**:534, **6**:534
Gammaproteobacteria **5**:134–135
gamma rays **7**:297*F*
Gammarellis homaris **4**:386*T*
gammarid amphipods **3**:183*F*
Gammarus chevreuxi **6**:7
Gammarus minus **7**:53
Gammarus pulex **4**:255
Ganges river dolphin **3**:192
Gangte society
people–ecosystem relationship
conservation practices **6**:226
ecodevelopment programs **6**:226–227
ecological refugees **6**:221*T*, **6**:225–226
land-use changes **6**:223–224, **6**:224*T*

Gangte society (*continued*)
resource ownership, sharing, and utilization **6**:223, **6**:224*T*
value appropriation **6**:224–225
political and economic subjugation **6**:222
Ganoderma lucidum **3**:637*T*, **3**:638
Garcinia **3**:130*T*
Garcinia mangostana **2**:663
garden rocket **3**:131–132
garibaldi **3**:469, **3**:474*F*
garlic
intercropping patterns **2**:383
native crops **2**:662–663
garlic mustard **6**:106*F*, **6**:107
Garra ghorensis **5**:463*T*
Garra rufa **5**:463*T*
garrigues **1**:268, **4**:577
Garrulus glandarius **5**:474
gars **3**:462, **3**:462*F*
garter snakes **5**:315–316, **6**:411, **7**:390
Garypus titaneus **1**:212–213
gas chromatography–mass spectrometry (GC–MS) **5**:708
Gasterosteiforms **3**:468
Gasterosteus aculeatus **1**:32, **2**:275, **4**:517
gastrointestinal irritants **3**:635*T*
Gastropoda/gastropods
classifications **4**:381*T*
critical pH levels **6**:153*F*
endangered species **3**:226
extinctions **3**:202*T*
fossil record **3**:681*F*, **3**:685*F*
Late Triassic mass extinction **5**:172–173
mangrove ecosystems **5**:17–18, **5**:17*F*
marine sediments **5**:106*F*
number of species **3**:182*T*, **4**:371*T*, **5**:669–670*T*
physical characteristics **4**:392, **5**:379, **5**:379*F*
sandy beaches **2**:131–133
seagrass ecosystems **6**:547–548
spatiotemporal distributions **5**:373–374, **5**:374*F*
species diversity **4**:386–387, **4**:386*F*, **4**:387*F*, **4**:388*F*
Gastrotricha
Antarctica **1**:184*T*
classifications and sampling methods **4**:381*T*
fossil record **3**:684*F*
global diversity **4**:385*T*
lake and pond ecosystems **4**:374*F*
marine ecosystems **5**:50*T*
number of species **4**:371*T*, **5**:669–670*T*
phylogenetic relationships **7**:433*F*
Gaudi, Antoni **1**:590–591
gaudy grasshoppers **3**:735, **3**:735*F*
Gaum rail **7**:188
gaupinol **4**:595–596
gaur **1**:640
Gause, G. F. **2**:182–184, **6**:669–670
Gavialis gangeticus **6**:401
Gaviiformes **7**:181–182
Gavla spp. **3**:256
Gazella atlantica **5**:364–365
Gazella dorcas **5**:475–476
Gazella gazella **5**:475–476, **5**:483
Gazella granti **1**:52
Gazella rufifrons albonotata **1**:52
Gazella rufifrons thomsoni **1**:52
Gazella spp. **6**:333
Gazella subgutturosa **1**:269, **6**:334
Gazella thomsoni **5**:313
gazelles **1**:51, **6**:333, **6**:572–573
Gazza minuta **3**:478*F*
Gecarcoidea natalis **3**:222, **7**:260
geckos **1**:488, **1**:591, **4**:360, **5**:472, **6**:402
geese **1**:644*F*, **1**:710
Geissolomataceae **5**:219–220
Gelechiidae **5**:387–388*T*
Gelechioidea **5**:386*F*, **5**:387–388*T*, **5**:389–390
Geleia **6**:295
gel electrophoresis **5**:708
gelidium **1**:190*T*
Gelisols **4**:225–226
Gemellicystis **6**:30
Gemma gemma **3**:364
gemsbok **1**:487–488
gene banks **2**:245, **3**:33, **4**:319, **4**:321*T*
gene flow **2**:511–513
GENELAND (software program) **4**:511
Gene Ontology (GO) **1**:444
general circulation models (GCMs) **2**:88, **4**:21–22, **4**:22–23, **6**:148–149
general equilibrium (GE) models **4**:294–295
generalized additive models (GAMs) **6**:701
Generalized Inverse Gaussian-Poisson (GIGP) distribution **3**:542–543
generalized linear mixed models (GLMMs) **6**:693–694
generalized linear models (GLMs) **6**:701
genes/genetic code **3:654–661**
chloroplast DNA (cpDNA) **3**:660, **3**:666–667
chromosomal and gene mutations **3**:658*T*
DNA structural characteristics **3**:655–656, **3**:655*F*
eukaryotic genomes/chromosomes **3**:657–659, **3**:658*F*, **6**:360–361, **6**:361*T*
gene banks **3:641–653**
conservation role **3**:641–642
ex situ collections **3**:641–642
nonseed gene banks
animal germplasm samples **1**:645–646, **3**:651–652
botanical gardens **3**:653
cryopreservation **1**:645–646, **3**:651–652
DNA banks **3**:652, **4**:319–320, **4**:321*T*
field gene banks **3**:33, **3**:652–653, **4**:319, **4**:321*T*
microorganisms **3**:652, **3**:652*T*
pollen **3**:650, **4**:319–320, **4**:321*T*
somatic and zygotic plant embryos **3**:650–651
spores **3**:650
vegetative tissues **3**:651
zoos/zoological parks **3**:653
seed banks
collection value **3**:644–645
community seed banks **3**:644
current status **3**:643–644
effectiveness **3**:642
genetic considerations **3**:646
historical perspective **3**:642–643
longevity predictions **3**:645–646
management strategies **3**:646–647
plant conservation **6**:85
scientific principles **3**:645–646
seed storage conditions **3**:645–646
wild species **3**:644
genetic code **3**:656–657, **3**:657*T*
heredity mechanisms **3**:654–655
information transfer **3**:656–657, **3**:656*F*
mitochondrial DNA (mtDNA)
animal genome **3**:659–660
plant genome **3**:660
mutations **3**:658*T*, **3**:660–661
noncoding repetitive DNA sequences **3**:659
organelles **3**:659–660
prokaryotic genomes **3**:657, **3**:658*F*
research background **3**:654–655
RNA structural characteristics **3**:655–656, **3**:655*F*
genetically modified organisms (GMOs) **1**:530, **3**:222
genetic conservation **4**:315
genetic diversity **3:662–677**
analytical methods
comparative genomics
basic concepts **3**:664
protein polymorphism **3**:663
DNA polymorphism
amplified fragment length polymorphisms (AFLPs) **3**:663–664
microsatellites **3**:663, **3**:671*F*
nucleotide polymorphism **3**:664
polymerase chain reaction (PCR) analysis **3**:663
random polymorphic DNA (RAPD) fragments **3**:663, **3**:671*F*
restriction fragment length polymorphisms (RFLPs) **3**:663
expressed sequence tags (ESTs) database **3**:664
microarray analysis **3**:664
protein polymorphism
comparative genomics **3**:663
technical innovations **3**:663
animal breeding **3:100–107**
challenges **1**:648
conservation planning **1**:639–640
conservation strategies
ex situ strategies **1**:645–646
general discussion **1**:642–645
in situ strategies **1**:642–645
crossbred animals **1**:639
current trends **1**:638–639
diversity challenges **1**:636
future outlook **1**:648
genetic structure **1**:637–638
global breeds **1**:638–639
livestock breeds **1**:637

management strategies
conservation strategies **1**:642–645
Food and Agriculture Organization (FAO) programs **1**:640–642, **1**:641
inventory, description, and characterization assessments **1**:641–642, **1**:642*F*
risk status classification **1**:641–642, **1**:642*F*, **1**:642*T*, **1**:643*F*, **1**:644*F*
new technologies
cloning **1**:647–648
cryopreservation **1**:646
gene mapping **1**:646–647
molecular characterization **1**:646–647
wild relative use **1**:640
bacteria **1:317–325**
circular DNA molecules **1**:325
genetic analysis
circular DNA molecules **1**:325
complementation analysis **1**:323*F*, **1**:324
general discussion **1**:321
homologous recombination **1**:322–324, **1**:324*F*
portable regions of homology **1**:324
reversion/suppressor analysis **1**:321–322
transposable elements **1**:324–325
genetic exchange
conjugation **1**:320, **1**:320*T*, **1**:321*F*
importance **1**:319
transduction **1**:320–321, **1**:320*T*, **1**:322*F*
transformation **1**:319–320, **1**:320*T*
mutations
circular DNA molecules **1**:325
conditional mutations **1**:317–318, **1**:318*T*
genetic analysis **1**:321
mutant isolation and characterization **1**:318–319, **1**:318*T*, **1**:319*F*, **1**:320*F*
mutation effects **1**:317–318
occurrences **1**:317
transposable elements **1**:324–325
summary discussion **1**:325
transposable elements **1**:324–325
basic concepts **1**:401
conservation biology
marine ecosystems **5**:36–37
measurement methods **2**:244–245
research approaches **2**:244–245
zoos, aquaria, botanical gardens, and gene banks **2**:245, **4**:316*F*, **4**:319, **4**:321*T*
conservation genetics **2:263–277**
ex-situ conservation **2**:273–274
future outlook **2**:274–275
genetic drift **2**:265–267
genomic research **2**:274–275, **2**:275*F*
heterozygosity and fitness correlations (HFCs) **2**:266
historical perspective **2**:263–265
inbreeding **2**:265–267, **2**:272–273
individual genotypes **2**:272–273
kin relationships **2**:272–273
low diversity–small population relationship **2**:265–267
management units **2**:268–269
population structure **2**:267–272, **2**:269*F*, **2**:270*F*
skyline plots **2**:267, **2**:267*F*
crop production **1**:360
divergence detection **6**:169
endangered mammals **3**:190–191
evidentiary research
allozymic diversity **3**:671–673
cross-phylogeny global allozyme diversity **3**:667–669, **3**:668*F*
cross-phylogeny regional allozyme diversity
Israel **3**:669–670, **3**:669*F*
subterranean mammals **3**:670
DNA diversity
chloroplast DNA (cpDNA) **3**:666–667
general discussion **3**:666
mitochondrial cytochrome b **3**:666
mitochondrial DNA (mtDNA) **3**:666
simple sequence repeats (SSRs) **3**:667
single nucleotide polymorphisms (SNPs) **3**:667
evolutionary significance **3**:673–674
local population diversity
Evolution Canyon testing microsite **3**:670–671, **3**:673*F*, **3**:674*F*
microgeographical studies **3**:670, **3**:672*F*
protein diversity
electrophoretic analysis **3**:664–665, **3**:665*F*
enzyme kinetic diversity **3**:665
genetic partitions **3**:666
genetic population structure **3**:666
loci variation patterns **3**:665–666
marine ecosystems **5**:36–37, **5**:50
molecular evolution
generation and maintenance mechanisms **3**:662
historical perspective **3**:662
theoretical perspectives **3**:662–663
molecular level diversity **2:626–637**
conservation biology **2**:637
DNA sequencing **2**:627, **5**:254
future research outlook **2**:635–637
general discussion **2**:626
genetic variation
mutation process **2**:627–628
phylogenetic divergence **2**:632–633
process mechanisms **2**:627
mutation process
characteristics and classifications **2**:627–628
cytological considerations **2**:628–629
empirical research **2**:631–632, **2**:631*T*
evolutionary models **2**:629–630
genome size effects **2**:628, **2**:628*F*
influencing factors **2**:629
mutation rates **2**:628, **2**:628*T*
nearly neutral theory **2**:630
neutral theory **2**:629–630
nucleotide polymorphism **2**:631–632, **2**:631*T*, **2**:632*F*
ontogenetic considerations **2**:628–629
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
selection-based models **2**:630–631
statistical tests **2**:631
phylogenetic divergence
amino acid replacement **2**:634, **2**:634*T*
empirical research **2**:634
evolutionary models **2**:633
general discussion **2**:632–633
nearly neutral theory **2**:633
neutral theory **2**:633
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
selection-based models **2**:633–634
statistical tests **2**:634
substitution process **2**:632–633, **2**:634*T*
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
research background **2**:626–627
valuation measures **2**:637
nitrogen cycle impacts **5**:540–544
pest control **2**:391–392
theoretical perspectives
ecological diversity **3**:676
environmental stress **3**:676
equilibrium maintenance **3**:674–675
functional value **3**:674
generation mechanisms **3**:674
natural selection **3**:675
selection stabilization **3**:675
selection versus genetic drift **3**:675–676
genetic drift **2**:265–267, **3**:398–399, **4**:250, **4**:279–281, **4**:280*F*, **6**:168, **6**:632
genetic engineering
environmental impacts **3**:284–285
pollinators **6**:145
genetic load theory **6**:188
genetic overfishing **3**:445
genetics research **6**:1–3
genets **4**:693
Genipara americana **4**:289*T*
Genista **1**:269
Genlisea **6**:44*T*
genomics
adaptive processes **1**:18
conservation genetics **2**:274–275, **2**:275*F*
Gentiana **1**:270
Gentianaceae **6**:315*T*
Gentiana sedifolia **4**:103
Genyornis **4**:727–728
Genyornis newtoni **5**:363–364, **5**:364*F*
geobotanical research **4**:395, **4**:396*T*
Geocapromys brownii **4**:713, **4**:714*F*
Geocapromys ingrahami **3**:197, **4**:713, **4**:714*F*
Geochelone **4**:657
Geochelone sulcata **6**:405
geoecology **4**:406
Geoffrey Saint-Hilaire, étienne **2**:424
Geoffroy's cat **2**:590*T*

Geoffroy's spider monkey **6**:256–258*T*
Geoffroy's woolly monkey **6**:256–258*T*
geographical information systems (GIS) **2**:217–218, **4**:509–510, **6**:395
geographic diversity **6**:534
geologic timescale
  paleodiversity **3:678–690**
    Archean era **3**:682
    biodiversity drivers
      climate change **3**:689
      extinction events **3**:689–690
      global heterogeneity **3**:688–689
      plate tectonics **3**:688–689
    evolutionary development
      morphological disparities **3**:679
      species richness **3**:678–679
    fossil biodiversity
      completeness measures **3**:680, **3**:681*F*
      rock record completeness **3**:679–680
      sampling challenges **3**:679
      taxonomic reconstruction measures **3**:680–682
    microbial biodiversity **3**:682
    Phanerozoic marine biodiversity
      Cambrian explosion **3**:683, **3**:684*F*
      chordate diversity **3**:684–685
      invertebrate diversity **3**:683–684, **3**:685*F*, **3**:686*F*
      neoproterozoic animals **3**:682–683
      taxonomic turnover **3**:685–686
    Phanerozoic terrestrial biodiversity
      animal diversity **3**:687–688, **3**:687*F*
      chordates **3**:688
      invertebrate diversity **3**:687–688, **3**:687*F*
      plant diversity **3**:686–687, **3**:687*F*
    Proterozoic era **3**:682, **3**:683*F*
    research approaches **3**:678
  pre-Quaternary mammal extinctions **5**:1*T*
Geometroidea **5**:384–385, **5**:386*F*, **5**:387–388*T*, **5**:389–390
Geomyces destructans **2**:527–529, **2**:529*F*
Geomyidae rodents **7**:432, **7**:432*T*
Geophilomorpha **1**:246*T*, **1**:247, **5**:421–423, **5**:422*F*
Geophilus carpophagus **5**:422*F*
Geophilus persephones **5**:431
Geophilus seurati **5**:425
George (Lake) **4**:461*T*
Georgia **6**:519
Georgia Endangered Plant Stewardship Network **7**:20–21
geospatial data standards **1**:443–444
geosphere **3**:279
Geospiza conirostris **1**:27–28
Geospiza difficilis **1**:27–28
Geospiza fortis **1**:27–28, **2**:140, **2**:718–719, **4**:47*F*
Geospiza fuliginosa **1**:27–28, **4**:47*F*
Geospiza magnirostris **1**:27–28, **4**:47*F*
Geospiza scandens **1**:27–28, **2**:718–719
geothermally heated systems **4**:112–113
geothermal power **3**:253*T*, **3**:262, **3**:265*T*
Geotrichum **3**:638–639
Geotrupidae **1**:710
Geranium multipartitum **4**:103
Gerbillus andersoni allenbyi **5**:472
Gerbillus gerbillus **5**:472
Gerbillus pyramidum **5**:472
gerenuk **5**:235
Germain's langur **6**:256–258*T*
German Saprobic Index **4**:254
Germany
  botanical gardens **6**:519
  natural history museums and institutions **5**:411*F*
  nitrogen deposition **5**:529*F*
  urbanization **4**:560*F*
  vegetation surveys **3**:341–344*T*
Germplasm Resources Information Network (GRIN) **3**:649–650
Geronticus eremita **7**:188
Gerromorpha **7**:293
Gesellschaft für Technische Zusammenarbeiten **7**:19
Gesneriaceae **3**:522*T*, **6**:316*T*
Geum urbanum **4**:517
Ghana
  botanical gardens **6**:519
  deforestation **2**:449*T*
  tropical rain forests **6**:314–315, **6**:317*T*
gharials **6**:401
Gharo Creek **1**:276*T*
gherkins **2**:663
ghost crabs **3**:222, **5**:316
ghostfishing **3**:446
giant African scorpions **3**:223
giant African snail **1**:82, **4**:357, **5**:382
giant anteaters **2**:590*T*, **7**:229*T*
giant armadillo **2**:590*T*, **4**:689–690, **7**:229*T*
giant Cape horse **5**:364–365
giant clams **3**:199, **5**:118–120, **5**:119*F*
giant forest hogs **4**:696
giant ground beetle **3**:222
giant hartebeest **5**:364–365
giant hogweed **3**:705, **6**:93
giant mouse lemur **6**:251*T*
giant North African deer **5**:364–365
giant otter **4**:693
giant pandas **2**:243–244, **3**:190
giant sloths **4**:689
giant waterbugs **7**:293
giardia **5**:670
Giardia intestinalis **6**:293
gibber **2**:462
gibbons **4**:690*T*, **4**:692–693, **6**:251*T*
Gibraltar **6**:519
Gibson Desert **2**:468
gift economies **3**:288–289
Gila achilleifolia **4**:248*F*
Gila elegans **3**:260
Gila monsters **5**:709–710, **5**:711, **7**:229*T*
Gila robusta **3**:260
Gilbertiodendron dewevrei **7**:273
Gilia sinuata **2**:373*T*
gill nets **3**:192, **3**:446, **6**:420
Gilpin–Ayala model **2**:186, **2**:186*F*
ginger
  East Asia **2**:663
  Southeast Asia **2**:663
Gini–Simpson index **2**:438–439, **5**:202–203
Ginkgo **2**:47, **5**:618*F*, **6**:58*T*, **6**:59–60
Ginkgoaceae **1**:260*T*
Ginkgo biloba **4**:225*F*, **4**:226, **4**:227*F*, **6**:59–60, **6**:77
Ginkgoites **5**:555
ginseng **6**:137, **6**:138*F*
ginsenosides **6**:137, **6**:138*F*
Giraffa camelopardalis **1**:487–488, **4**:696, **5**:225–226, **5**:225*T*
Giraffa camelopardalis reticulata **1**:51
giraffes
  characteristics **4**:696
  extant species **5**:225*T*
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  Succulent Karoo **1**:487–488
  tropical Africa **1**:51
Giraffidae **4**:696, **5**:224*T*, **5**:225*T*
Giraffokeryx **4**:696
Gisekia **2**:376*F*
Gisekiaceae **2**:374*T*
Givetian Stage **5**:169–170
glacial cycles **1**:583, **1**:585*F*, **1**:586*F*, **2**:59–60, **2**:466, **5**:171
glacial melting **4**:228–229
glaciers **3**:560–562, **3**:561*T*
Glacken, C.J. **6**:383
Glansdorff, Paul **5**:630
Glanville fritillary butterfly **4**:517
Glaucidiaceae **1**:260*T*
Glen Canyon Dam **3**:260
glendonites **4**:225–226, **4**:225*F*
gliders **4**:687*T*, **4**:688–689, **4**:701*T*
Gliessman, Stephen **1**:112–113
Glis glis **7**:27
global agriculture **3:426–434**
  biodiversity regime shifts and loss thresholds **3**:429–430, **3**:430*T*
  future outlook
    agricultural policies **3**:432–433
    agriculture-related conservation strategies **3**:431
    market-based approaches **3**:432–433
    Millennium Ecosystem Assessment (MEA) **3**:430–431, **3**:431*T*
    payment for ecosystem services (PES) **3**:432–433
    sustainable agriculture **3**:430–431
  impacts
    climate change **3**:427–429, **3**:428*F*
    general discussion **3**:427
    habitat loss **3**:427, **3**:428*T*, **4**:154
    invasive species **3**:429
    water quality **3**:429
    water quantity **3**:429
  prevalence and distribution **3**:426–427, **3**:427*F*
  research summary **3**:433
Global Amphibian Assessment **3**:164*T*
Global Amphibian Assessment (GAA) **1**:176–177
Global Assessment of Soil Degradation **1**:704–705, **6**:597–598
Global Biodiversity (GLOBIO) model **4**:562–564

Global Biodiversity Information Facility (GBIF) **1**:441, **3**:484–485, **4**:327, **5**:407
Global Biodiversity Outlook
basic principles **2**:308–311
biodiversity indicators **2**:309, **2**:310*F*
tipping points **2**:309, **2**:309
global carbon cycle **1**:70–71, **6**:148*F*
global circulation models (GCMs) **2**:63–64
Global Crop Diversity Trust (GCDT) **4**:328
Global Earth Observation System of Systems (GEOSS) **4**:327
Global Environment Facility (GEF) **7:136–143**
evolutionary development
biodiversity strategy
GEF-3 **7**:138–139
GEF-4 and GEF-5 **7**:139
sustainable forest management financing **7**:139–140
operational strategy and programs **7**:138
functional role **2**:317, **4**:324–325, **7**:136–137
future outlook **7**:141–142
GEF benefits index (GBI) **7**:137–138
GEF performance index (GPI) **7**:137–138
global policy framework and guidance **7**:137
gross domestic product index (GDPI) **7**:137–138
key accomplishments **7**:142
monitoring and evaluation strategies
learning objectives **7**:140–141
portfolio level **7**:140
results-based management framework **7**:140–141
origins **4**:324–325, **7**:136–137
project investment and funding **7**:137, **7**:138*F*
resource allocations **7**:207
System forTransparent Allocation of Resources (STAR) **7**:137–138
Global Forest Resources Assessment
background information **6**:322
model-based assessments **6**:323, **6**:324*T*, **6**:325*T*
1990 remote sensing survey **6**:323, **6**:324*T*, **6**:325*T*
2000 remote sensing survey **6**:323–325, **6**:325*T*
globalization
plant species **3:700–706**
biotic homogenization **3**:701
donor and receptor regions **3**:703
general discussion **3**:700
geographic distribution **3**:701–703, **3**:702*F*
historical perspective **3**:700–701, **3**:701*F*
impacts and implications **3**:705
selected examples
alligator weed **3**:703–704, **3**:704*F*
castor oil plant **3**:703, **3**:704*F*
general discussion **3**:703
giant hogweed **3**:705
Spanish Needle **3**:704–705, **3**:704*F*
St. John's wort **3**:704, **3**:704*F*
water hyacinth **3**:703, **3**:704*F*
shipping trade **3**:701*F*
global land cover map **4**:557*F*, **4**:566*F*, **6**:390*F*
Global Learning and Observation to Benefit the Environment (GLOBE) **3**:136, **3**:136
Globally Unique Identifier (GUID) **1**:441–442
global mean surface temperatures **2**:83–84
global nitrogen cycle **3**:282, **5**:111, **5**:307, **5**:307*F*
Global Ocean Biodiversity Initiative (GOBI) **4**:329–330
Global Ocean Observing System (GOOS) **2**:14–15
Global Ocean Sampling (GOS) program
cross-biome sampling studies **5**:294
fragment recruitment **5**:294–295, **5**:295*F*
functional diversity (FD) **5**:295
research background **5**:293
research outlook **5**:297
sampling sites **5**:293*F*
Sargasso Sea **5**:293–294
taxonomic diversity **5**:297
Global Plan of Action for Animal Genetic Resources **1**:640–642, **1**:641
Global Positioning System (GPS) **3**:443, **4**:509–510
global species richness **3:707–711**
current estimates **3**:709, **3**:710*T*
estimation methods
canvassing experts **3**:708
counting **3**:707–708
species–abundance distributions **3**:709
species–body size distributions **3**:709
temporal patterning **3**:708
trophic relations **3**:709
undescribed species **3**:708
well-studied areas extrapolation estimation **3**:708–709
well-studied groups extrapolation estimation **3**:709
future outlook **3**:711
importance **3**:707
sampling considerations **3**:707
uncertainties
environments
coral reef ecosystems **3**:710
deep ocean environments **3**:710
soils **3**:710
tropical forest canopies **3**:709–710
general discussion **3**:709–710
higher taxa
arthropods **3**:711
bacteria **3**:710–711
fungi **3**:711
nematodes **3**:711
Protozoa **3**:711
viruses **3**:710
Global Strategy for Plant Conservation (GSPC) **6**:78, **6**:80–81*T*, **6**:509, **6**:514–515, **6**:515
global sulfur cycle **3**:282, **5**:111
Global Trade Analysis Project (GTAP) model
consumption-side structure **4**:540*F*
corn ethanol production **4**:539–540
economic and environmental impacts **4**:540–544
emission factors **4**:548–549
ethanol production
estimated emissions **4**:549–550, **4**:549*T*
final analysis **4**:551–552, **4**:551*T*
marginal and average estimated emissions
first simulated group **4**:550, **4**:550*T*
second simulated group **4**:550–551, **4**:550*T*
third simulated group **4**:551, **4**:551*T*
research limitations **4**:553
research summary **4**:552–553
well-to-wheel emissions **4**:551–552, **4**:552*T*, **4**:553*T*
land cover and land-use activities **4**:542*F*
livestock production structure **4**:542*F*, **4**:543*F*, **4**:543*T*
modeling process **4**:539–540
production-side structure **4**:541*F*
productivity measures **4**:540–544
simulation models
basic concepts **4**:544–545
2001 simulation model **4**:544–545, **4**:544*T*, **4**:545*T*
2001–2006 simulation **4**:545–547, **4**:546*T*, **4**:547*T*
2006–2015 simulation model **4**:547–548, **4**:548*T*, **4**:549*T*
global urbanization **4:231–244**
demographic transition **4**:231–232
dependence interactions
biodiversity **4**:237
ecosystem services **4**:237–238, **4**:238*F*
general discussion **4**:237
environmental impact
ecosystem services **4**:235
general discussion **4**:234–235
habitat loss and fragmentation **4**:234–235
positive impacts **4**:236–237
protected area systems **4**:235–236, **4**:235*F*
geographic distribution **4**:235*F*
management strategies
ecosystem services **4**:239*F*, **4**:241–242
general discussion **4**:238–241
rural-to-urban migration **4**:240–241*F*, **4**:242
sustainable development **4**:239–241, **4**:239*F*
population trends **4**:231, **4**:231*F*
rural-to-urban migration **4**:232–234, **4**:233*F*, **4**:240–241*F*, **4**:242
global warming **4:224–230**
alpine ecosystems **1**:156–157, **1**:156*F*
amphibian species decline
elevated ultraviolet (UV–B) radiation **3**:697
general discussion **3**:696–697

global warming (*continued*)
temperature variability–disease immunity relationship **3**:697
weather pattern shift effects **3**:697
Anthropocene epoch **4**:146
atmospheric indicators **4**:224
biodiversity–global warming correlation **4**:227–228, **4**:227*F*, **4**:664
biodiversity impacts **7**:203*F*
biotic consequences
helical animal burrows **4**:229
increased leaf herbivory **4**:229
increased plant sclerophylly **4**:229
oceanic eutrophication **4**:229
reduced plant transpiration **4**:229
tropical animal geographic expansion **4**:229
tropical plant geographic expansion **4**:229
wetland eutrophication **4**:229
C_4 plants **2**:378
carbon cycle **6**:147–149
climate models
dynamical downscaling techniques **2**:89–90
general circulation models (GCMs) **2**:88, **6**:148–149
regional predictions **2**:88–89
statistical downscaling techniques **2**:89
validation methods **2**:88
Convention on Biological Diversity (1930) **2**:318
coral reef ecosystems **2**:341–342, **2**:342*F*
desertification **2**:486–487
disease dynamics **2**:86–87, **4**:664
ecological models
animal responses **2**:92–94, **2**:93*F*
bottom-up approaches **2**:66–68, **2**:90
combined approaches **2**:68, **2**:90–91
community ecology **2**:94
extinction probability **2**:94–95
future climate projections **2**:90
refined approaches **2**:91
strategic cyclical scaling approach **2**:68–69, **2**:91
three-way linkages **2**:94
top-down approaches **2**:66, **2**:90
vegetation responses **2**:91–92
ecological responses
abundance shifts **2**:85
distribution shifts **2**:85
extinction rates **2**:86–87, **4**:664
phenological shifts **2**:85–86
physical changes **2**:86
endangered freshwater invertebrates **3**:181–182
environmental consequences
coastal/marine ecosystems **6**:150–151
deforestation **2**:454–455
ecosystem alterations **6**:149–150
enhanced climatic seasonality **4**:228
forest ecosystem alterations **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
humid climates **4**:228
increased stream runoff **4**:228
marine alkalinization **4**:228
past events **4**:227*F*, **4**:228
rising sea level **4**:228
shrinking crysosphere **4**:228–229
soil acidification **4**:225*F*, **4**:228
environmental impacts **3**:282–283, **7**:203*F*
fire/fire regimes **3**:441
forecasts and predictions **2**:62
forest canopies **3**:524
future climate change **4**:21–22, **4**:22*F*
habitat alteration **4**:155
harmful algal blooms **1**:479–481
historical perspective **2**:82–83
infectious diseases **2**:531–532
Late Permian **5**:171
marine ecosystems **5:32–44**
abundance impacts **5**:34–35
community impacts **5**:37–38
conservation strategies
assisted colonization **5**:41
ecosystem-based management **5**:40
fisheries management **5**:40–41
goals and priorities **5**:38–39
marine protected areas **5**:39–40
predictive models **5**:39
restoration **5**:41
distribution impacts **5**:35–36
evolutionary impacts
adaptation rates **5**:36
climate tolerance **5**:36–37
general discussion **5**:36
genetic diversity **5**:36–37
local adaptations **5**:36
future research areas **5**:41
global impacts **3**:449, **5**:32
physiological impacts
ocean acidification **5**:33–34
rising seawater temperatures **5**:33
research needs **5**:41
marine sedimentary habitats **5**:113*F*, **5**:114–115
past climate change **4**:19, **4**:20*F*
plant biodiversity threats **6**:62
plant conservation **6**:83–84
proxy assessments
atmospheric carbon dioxide (CO_2) **4**:225*F*, **4**:226, **4**:227*F*
atmospheric methane (CH_4) **4**:225*F*, **4**:226
biodiversity **4**:226–227
limitations **4**:224–226
paleoprecipitation **4**:225*F*, **4**:226
paleotemperature **4**:224–226
recent climate change **4**:19–21, **4**:21*F*
research summary **4**:229–230
savannas **1**:71–72
Southern Hemisphere ecosystems **6**:617
wild species **2:79–99**
climate models
dynamical downscaling techniques **2**:89–90
general circulation models (GCMs) **2**:88
regional predictions **2**:88–89
statistical downscaling techniques **2**:89
validation methods **2**:88
current trends **2**:83–84
ecological models
animal responses **2**:92–94, **2**:93*F*
bottom-up approaches **2**:66–68, **2**:90
combined approaches **2**:68, **2**:90–91
community ecology **2**:94
extinction probability **2**:94–95
future climate projections **2**:90
refined approaches **2**:91
strategic cyclical scaling approach **2**:68–69, **2**:91
three-way linkages **2**:94
top-down approaches **2**:66, **2**:90
vegetation responses **2**:91–92
ecological responses
abundance shifts **2**:85
distribution shifts **2**:85
extinction rates **2**:86–87, **4**:664
phenological shifts **2**:85–86
physical changes **2**:86
extreme events **2**:84–85
fish **3**:449–450
future research areas
applications-driven climatic research **2**:96
ecological impact assessments **2**:95
ecosystem adaptability **2**:95–96
multiscale interactive analysis **2**:96
historical perspective **2**:82–83
importance **2**:80–81
joint atribution theory **2**:87
meta-analyses **2**:87
synergistic effects **2**:81–82
global water cycle **1**:553–554, **1**:555*F*
Globicephala macrorhynchus **3**:192–193
Globigerina **5**:306, **6**:292
Globigerinoides ruber **6**:291
GLOBIO model **4**:562–564
Globodera rostochiensis **5**:672–673
Globorotalia menardii **6**:291
Globularia arabica **5**:484–485
Glochidion **1**:263
Gloeotrichia **6**:30
Gloethece **2**:614
Gloger, C. L. **1**:579
Gloger's Rule **1**:579
Glomerida **5**:423–424, **5**:423*F*
Glomeromycota **3**:627–628
Glossata **5**:385, **5**:385–386, **5**:386*F*, **5**:387–388*T*
Glossina morsitans **1**:50
Glossina pallidipes **1**:50
Glossinidae **3**:489–491*T*
Glossogobius giurus **4**:465
Glossophaga soricina **2**:587*F*
Glossopteridales **5**:617
Glossopteris **5**:618*F*
Glossotherium **4**:727*T*
Glossotherium harlani **5**:225*T*
glossy buckthorn **7**:399
glossy ibis **5**:476–477
glucose **6**:113*T*
glucosidase **5**:269*T*
glufosinate **4**:90*T*
Gluphisia septentrionis **5**:395

glutamic acid **6**:134*F*
glutathione *S*-transferases (GSTs) **4**:300
Glyceridae **7**:417*F*
Glycine max **1**:543–547, **2**:663
glycines **4**:92*T*
glycogen storage **7**:23
glyphosate **4**:90*T*, **5**:697
Glyptodon **4**:727*T*
Glyptodontidae **4**:689–690, **4**:727*T*
glyptodonts **4**:724–725*T*
Glyptostrobus **1**:272–273
Glyptotermes **4**:409*T*
Glyptotherium **4**:724–725*T*
Gnathabelodon **5**:224*T*
Gnathamitermes **4**:409*T*
Gnathostomata **7**:338
Gnathostomulida
  fossil record **3**:684*F*
  global diversity **4**:385*T*
  marine ecosystems **5**:50*T*
  number of species **5**:669–670*T*
Gnathotitan **5**:224*T*
Gnetales **5**:613, **5**:614*F*
gnetophytes **6**:58*T*, **6**:59–60
Gnetum **5**:618*F*, **6**:59–60
Gnetum gnemon **5**:614*F*
Gnomonia cerastis **4**:174
Gnosenesima **7**:442–443*F*
goatfish **3**:467–471
goats **1:702–711**
  biotic relationships
    birds **1**:710
    invertebrate fauna **1**:709–710
    natural predators **1**:709
    parasites and pathogens **1**:710–711
    plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
    seed dispersal and germination **1**:708–709
    small mammals **1**:710
  desert ecosystems **2**:476*T*
  domestication **1**:702–703
  ecological role
    desertification **1**:704
    grazing effects **1**:703–704
    land degradation **1**:704–705
    nutrient cycling **1**:705–706, **1**:705*F*
    over-enrichment factors **1**:705–706
    plant diversity–grazing intensity relationship **1**:704*T*
    vegetation impacts **1**:703–704
  food habits **3**:127*T*
  introduced species **1**:76–77
    Australia **1**:80
    Brazil **1**:83
    India **1**:82
    South Africa **1**:81
  risk status classification **1**:644*F*
Gobi Desert **1**:270, **2**:469–470
gobies **3**:467–471
Goeldi's monkey **6**:251*T*
gold algae **3**:571–572, **3**:571*T*
golden bamboo lemur **6**:256–258*T*
Golden-bellied Mentawai langur **6**:256–258*T*
golden bowerbird **2**:75*F*
golden-brown mouse lemurs **4**:516, **6**:256–258*T*
golden carpet moth **6**:159*F*
golden-cheeked warbler **7**:183
golden cheeked warblers **4**:516–517
golden dart moth **6**:159*F*
Golden Dragonfish **3**:463, **3**:463*F*
golden eagles **1**:709, **5**:474, **6**:240
golden-headed lion tamarins **1**:666*T*, **6**:256–258*T*
golden langur **6**:256–258*T*
golden lion tamarins **1**:664*T*, **1**:666*T*, **2**:273, **3**:196, **4**:246, **4**:247*F*, **6**:256–258*T*
golden-mantled ground squirrels **7**:36
golden moles **1**:487–488
golden monkey **6**:256–258*T*
golden plover **1**:605
golden-shouldered parrot **7**:183–184
golden toad **2**:76, **2**:86–87, **3**:165, **3**:167–168, **3**:697, **4**:664
Golden V kelp **2**:6*F*
goldfish **2**:55*F*, **2**:55*T*
Gomortega **5**:621*F*
Gomortegaceae **5**:216–217
Gomphotheriidae **2**:583*F*, **4**:727*T*, **5**:224*T*, **5**:225*T*
Gomphotherium **5**:224*T*
Gomphrena **2**:371*T*, **2**:376*F*
Gondwanaland **6**:612–613, **7**:344
  see also Southern Hemisphere ecosystems
Goniadidae **7**:417*F*
Gonibregmatus plurimipes **5**:421–423
Goniomonas **6**:294, **6**:294*F*
Gonocalyx **3**:522*T*
Gonyaulacales **1**:472
Gonyaulux catenella **6**:295
Gonyleptoidea **1**:210–211
Gonyostomum **1**:475
Goodall, Jane **3**:317*F*
Goodeniaceae **1**:277
goosefish **3**:466–467
goosefoot **2**:663–664
goosegrass **1**:120
gophers **7**:432
gopher snakes **6**:408–409
Gopherus agassizii **2**:528*T*
Gorda Ridge **7**:327–328
Gordon model of bionomic equilibrium **2**:150–151, **2**:150*F*, **2**:151*F*
Gorgoderoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
gorgonians **2**:335
Gorgonzola cheese **3**:638–639
Gorgosaurus **2**:519*T*, **2**:520*T*
Gorilla beringei beringei **6**:254, **6**:256–258*T*
Gorilla beringei graueri **6**:256–258*T*
Gorilla gorilla **1**:49, **7**:229*T*
Gorilla gorilla **2**:631*T*
Gorilla gorilla beringei **3**:192
Gorilla gorilla diehli **6**:256–258*T*
Gorilla gorilla gorilla **6**:256–258*T*
gorillas **1**:49, **2**:528*T*, **2**:631*T*, **3**:192, **4**:693, **6**:251*T*, **6**:256–258*T*, **7**:229*T*
Gossipium barbadense **4**:289*T*
Gossypium arboreum **2**:663
Gossypium barbadense **2**:663–664
Gossypium herbaceum **2**:663
Gossypium hirsutum **2**:389, **2**:663–664
Goteborg Botanic Garden **6**:506–508
Götte's larva **7**:462–466, **7**:464–465*F*
gouramies **3**:467–471
governance systems **2:155–166**
  advantages/disadvantages **2**:163–164
  centralized regimes **2**:156–157
  challenges **2**:155–156
  complex adaptive systems
    building blocks **2**:161–162
    internal models **2**:161
    parallel adaptive systems **2**:163–164
    polycentric systems **2**:164–165
    properties and mechanisms **2**:161
    rule configurations and situations **2**:162–163, **2**:162*T*
    self-organized resource governance systems **2**:160–161
    tags **2**:161
  local governance systems **2**:163–164
  polycentric systems **2**:164–165
  rules as tools
    aggregation rules **2**:159–160, **2**:162*T*
    boundary rules **2**:157–158, **2**:162*T*
    choice rules **2**:158–159, **2**:162*T*
    information rules **2**:159–160, **2**:162*T*
    payoff rules **2**:159, **2**:162*T*
    position rules **2**:159, **2**:162*T*
    rule change experimentation **2**:160
    rule characteristics **2**:157–158, **2**:162*T*
    rule improvements **2**:160
    scope rules **2**:159–160, **2**:162*T*
  summary discussion **2**:165
gracilaria **1**:190*T*
Gracillariidae **5**:388*F*, **5**:395–396
Gracillarioidea **5**:386*F*, **5**:387–388*T*, **5**:389–390
grades
  basic concepts **2**:51–52, **2**:52*F*
  branching diagrams **2**:51, **2**:52*F*
  definition **2**:47
gradient speciation model **3**:144–145
gradualism **3**:395–396
graft-chimeras **5**:551
grains
  Europe **2**:662–663
  Mediterranean Basin **2**:662–663
  Neolithic agriculture **2**:658
  southwest Asia **2**:662–663
Grain-to-Green Program (GTGP)
  background information **3**:379–380
  biodiversity/biophysical effects **3**:379*F*, **3**:380
  future outlook and challenges **3**:381–382
  program goals **3**:379–380
  socioeconomic effects **3**:380–381, **3**:381*F*
  spatial distribution **3**:376*F*, **3**:379*F*, **3**:380
  summary discussion **3**:382
Grallatotermes **4**:409*T*
Gramineae **3**:129*T*, **6**:316*T*
graminoids **1**:151, **1**:151*F*
Grammadenia **3**:522*T*
Grammia virguncula **5**:397*F*
grand skink **4**:194*T*
granite inselbergs **1**:166

granivores **2**:470*T*, **2**:474–475, **6**:39–41
grapes **3**:32*F*
Grapes of Wrath, The (Steinbeck) **4**:652
grapevines
  genetic diversity **1**:463
  native crops **2**:662–663
grapsid crabs **5**:15–16, **5**:15*F*
graptolites **5**:167
Graptolithina **3**:685*F*
grasses
  C_3 and C_4 photosynthetic pathways **1**:58–59, **1**:59*F*, **1**:64*F*, **2**:361–362, **2**:375*F*, **2**:654–655
  introduced parasites–species decline relationships **2**:529–530, **2**:529*F*
  plant tissue chemistry **6**:120*T*
  taxonomic versus functional diversity **2**:654–655
grasshoppers **3:722–736**
  Acridoidea
    Acrididae (true grasshoppers and locusts) **3**:735
    characteristics **3**:734
    Pamphagidae (earth hoppers of pamphagids) **3**:734–735
    Romaleidae (lubber grasshoppers) **3**:734*F*, **3**:735
  biotic relationships **1**:710
  bladder grasshoppers **3**:733, **3**:733*F*
  characteristics **3**:722–723
  desert ecosystems **2**:473
  desert long-horned grasshoppers **3**:733–734
  earth hoppers **3**:734–735
  Eumastacoidea
    characteristics **3**:732–733, **3**:733*F*
    Eumastacidae (monkey grasshoppers) **3**:733, **3**:733*F*
    Proscopiidae (false walking sticks or proscopids) **3**:733
  false walking sticks **3**:733
  food and feeding habits **3**:724–725
  gaudy grasshoppers **3**:735, **3**:735*F*
  habitat loss and fragmentation **3**:220–221
  leaf grasshoppers **3**:735–736
  lubber grasshoppers **3**:734*F*, **3**:735
  monkey grasshoppers **3**:733, **3**:733*F*
  Near East ecosystems **5**:460*T*, **5**:469
  New Zealand **4**:194*T*
  number of species **4**:392*T*
  paleontology **3**:725
  pesticides **3**:221–222
  phylogeny and taxonomy **3**:725
  Pneumoroidea **3**:733, **3**:733*F*
  postglacial range expansion **2**:550, **2**:552*F*
  pygmy grasshoppers **3**:736, **3**:736*F*
  Pyrgomorphoidea (gaudy grasshoppers) **3**:735, **3**:735*F*
  reproduction and growth **3**:724
  sound production **3**:723–724
  Tanaoceroidea **3**:733–734
  Tetrigoidea (pygmy grasshoppers and grouse locusts) **3**:736, **3**:736*F*
  Tridactyloidea (false mole crickets and sandgropers) **3**:736, **3**:736*F*
  Trigonopterygoidea (leaf grasshoppers and wingless desert grasshoppers) **3**:735–736
  true grasshoppers **3**:735
  wingless desert grasshoppers **3**:735–736
grasslands
  African savannas **1**:66*F*, **1**:67
  biodiversity restoration **6**:456–457
  C_4 plants **2**:371*T*, **2**:377, **2**:378–379, **2**:379*F*
  characteristics **6**:331–333, **6**:332*F*, **6**:333*F*
  current global land cover **4**:557*F*
  Eastern North America **5**:562–563
  ecosystem function measurements **3**:77*F*
  ecosystems **4:1–7**
    biodiversity
      animal species diversity **4**:3–4
      biodiversity change scenarios **4**:5–6, **4**:6*F*
      characteristics **4**:3
      ecosystem functionality **4**:4–5
      future outlook **4**:5–6, **4**:6*F*
      plant species diversity **4**:3
      Relative Yield Total (RYT) **4**:4–5, **4**:5*F*
      soil microbial diversity **4**:4
    conservation status **3**:171–173*T*
    determinant factors **4**:1–3, **4**:2*F*
    ecosystem services **3**:96–98
    geographic extent **4**:1, **4**:2*F*
    water availability **4**:1–3, **4**:2*F*
  ecosystem services **1**:334*T*
  Europe **3**:335*T*, **3**:336–340*T*
  fire/fire regimes **3**:437–438, **4**:14–15
  flooded grasslands **6**:600–604*T*
  functional diversity (FD) measures **3**:604–605, **3**:605*F*
  global agriculture impacts **3**:428*T*
  global biodiversity trends **4**:563*F*
  grazer-based ecosystem restoration and management **4**:14–15
  habitat–bird species diversity relationship **1**:613
  habitat destruction **3**:404
  healthy/unhealthy grasslands **3**:311*F*
  historic land-use patterns **4**:577
  invasive species
    Bromus tectorum **4**:347–348, **4**:348*F*
    Lantana camara **4**:353
    mouse-ear hawkweed (*Hieracium pilosella*) **4**:348–349, **4**:349*F*
  keystone species study **4**:446–447, **4**:446*F*
  land-use trends **4**:559*F*
  large-scale biodiversity experiments **4**:588, **4**:588*F*
  montane grasslands **6**:600–604*T*
  neotropical seasonally dry forests **5**:496
  net primary production (NPP) **1**:560*T*
  nitrogen cycle impacts **5**:542, **5**:542*F*
  poverty–conservation relationship **2**:232
  pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
  precipitation–temperature relationship **1**:59*F*
  protected area systems **5**:444*T*, **6**:492, **6**:492*T*, **6**:495*T*, **6**:499*T*
  sagebrush–grassland restoration **6**:439–442, **6**:440*F*
  soils **6**:123*T*
  South America **3**:111*F*, **3**:113–114, **3**:114–115, **6**:600–604*T*, **6**:607*T*
  species richness **5**:605
  temperate grasslands **1**:262–263*T*, **1**:269
  threatened birds **7**:181
  tourism industry **7**:223
  tropical Africa **1**:52
  wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*
Grauer's lowland gorilla **6**:256–258*T*
Gravitholus **2**:519*T*, **2**:520*T*
Gray, Asa **1**:578, **2**:281
gray-crowned Central American squirrel monkey **6**:256–258*T*
gray fox **2**:590*T*
Gray-handed crested langur **6**:256–258*T*
gray-headed robin **2**:75*F*
graylings **3**:464–465, **6**:522–523
gray mongoose **4**:365–366
gray mullet **5**:473
gray partridge **3**:123, **4**:259
gray pineapple mealy bug **4**:365
gray-shanked douc **6**:256–258*T*
gray squirrels **1**:79, **2**:551–552*T*, **4**:360, **7**:28, **7**:28*F*
gray whales **5**:65*T*, **5**:76, **5**:87–88, **5**:87*T*, **6**:236
gray wolves *see* wolves
grazing **4:8–17**
  African savannas **1**:65
  antagonistic interactions **6**:39–41
  biodiversity-friendly farming **2**:542–544
  biological impoverishment **6**:334–337, **6**:335*T*
  deforestation **2**:453
  direct versus indirect effects **4**:10
  disturbance regimes **4**:9–10, **4**:10*F*
  domesticated ruminants
    biotic relationships
      birds **1**:710
      invertebrate fauna **1**:709–710
      natural predators **1**:709
      parasites and pathogens **1**:710–711
      plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
      seed dispersal and germination **1**:708–709
      small mammals **1**:710
    ecological role
      desertification **1**:704
      erosion **1**:704–705
      grazing effects **1**:703–704
      land degradation **1**:704–705
      nutrient cycling **1**:705–706, **1**:705*F*
      over-enrichment factors **1**:705–706
      plant diversity–grazing intensity relationship **1**:704*T*
      vegetation impacts **1**:703–704
    historical perspective **1**:702–703
  domestication effects **6**:567
  grazer-based ecosystem restoration and management
    background information **4**:14–15

coral reef ecosystems **4**:16
forests **4**:15
grasslands **4**:14–15
invasive species **4**:16–17
lakes and ponds **4**:15–16, **4**:16*F*
greenhouse effect **6**:340–341
herbivory **1**:65, **4**:8–9, **4**:9*F*
host chemistry effects **4**:13
human impacts **4**:154
introduced species
habitat modification **4**:359
nonindigenous plants **1**:76
nonindigenous vertebrates **1**:76–77, **1**:76*T*
marshes **5**:455*F*
megaherbivores **5**:227–228
nitrogen cycle impacts **5**:542–543
overgrazing **1**:489, **2**:378–379, **3**:311*F*, **4**:560*F*
paleo-patterns **4**:13–14
remote sensing applications **6**:398
research summary **4**:17
riparian zones **6**:337
spatiotemporal scale
between-habitat patterns **4**:11–12
functional role **4**:10–11
geographic patterns **4**:11, **4**:11*F*
microhabitat patterns **4**:12–13
within-habitat patterns **4**:12
Succulent Karoo **1**:489
termites **4**:432*T*
Great American Biotic Interchange (GABI) **2**:582–583, **2**:584*F*
Great Artesian Basin, Australia **3**:186, **3**:187*F*
great auks **1**:617, **1**:621, **7**:182
Great Barrier Reef Marine Park **2**:333*T*, **2**:343–344, **4**:664, **5**:446
Great Basin Desert
biodiversity patterns **5**:563
characteristics **2**:468–469
domesticated ruminants **1**:703
invasive species **4**:347–348
Great Bear (Lake) **4**:461*T*
Great Britain *see* United Kingdom
greater bamboo lemur **6**:256–258*T*
greater sage grouse **2**:503–505, **3**:256–257
greater yams **2**:663
greather short-tailed bat **5**:366–367
great Indian rhinoceros **5**:225*T*, **5**:226
Great Lakes Basin **2**:469–470
Great Lakes (US) **3**:560
Great Limpopo Transfrontier Park **5**:449
great mass extinctions
end-Cretaceous **5**:168*F*, **5**:169*F*, **5**:173–175, **5**:174*F*
Late Devonian **5**:168*F*, **5**:169–170, **5**:169*F*
Late Ordovician **5**:167–169, **5**:168*F*, **5**:169*F*
Late Permian **5**:168*F*, **5**:169*F*, **5**:170–172
Late Triassic **5**:168*F*, **5**:169*F*, **5**:172–173
research challenges **5**:167–169
Great Sandy-Tanami desert **2**:468
Great Slave (Lake) **4**:461*T*
great tits **1**:611–612, **1**:620, **5**:14, **5**:476, **5**:720
Great Victoria Desert **2**:468
grebes **1**:606
Greece
botanical gardens **6**:519
flora **5**:481*T*
green algae
eutrophication impacts **3**:352–354, **3**:353*F*, **3**:356*F*, **3**:357*F*
evolutionary development **2**:641–642
freshwater ecosystems **3**:571–572, **3**:571*T*
impacts and distribution **3**:411
river ecosystems **6**:477–478
surface water acidification **1**:10*T*
temperature tolerance ranges **4**:113–115
green carbon **6**:77
green crabs **4**:366, **5**:54, **5**:61
green fluorescent protein (GFP) **1**:591–592
green frogs **2**:719
greenhouse effect **4**:**18–32**
basic concepts **4**:18
biodiversity impacts
anthropogenic effects **4**:24
biota types
at-risk organisms **4**:29
organisms likely to benefit **4**:29–30
community impacts **4**:30
current climate change evidence **4**:30–31, **4**:31
research methodologies
integrated research **4**:26–27
manipulation experiments **4**:25–26, **4**:26*F*
natural climate variability **4**:25
paleobiological data **4**:25
species distribution models **4**:24–25, **4**:24*T*
research summary **4**:31
response types
adjustment **4**:27
evolution **4**:27
extinction **4**:29
migration **4**:27–29, **4**:28*F*
climate change **2**:58–59
enhanced greenhouse effect **4**:18–19, **4**:19*F*, **4**:19*T*, **4**:20*F*
feedbacks
biogeochemical feedbacks
characteristics **4**:23
marine feedbacks **4**:23
terrestrial feedbacks **4**:23–24
general discussion **4**:22–23
geophysical feedbacks **4**:22–23
global livestock **6**:340–341
global warming
future climate change **4**:21–22, **4**:22*F*
past climate change **4**:19, **4**:20*F*
recent climate change **4**:19–21, **4**:21*F*
radiative forcing effects **4**:20*F*
greenhouse gases
biofuel production **4**:293–294
carbon cycle **1**:679–680, **1**:680*T*, **5**:351–352, **6**:147–149
changing composition
carbon dioxide (CO_2)
biodiversity impacts **6**:126, **6**:127*F*
causal factors **1**:282–284
fossil fuel emissions **1**:283–284
land use changes **1**:284, **4**:**539–554**
monthly average measurements **1**:283*F*
carbon monoxide (CO) **1**:286
chlorofluorocarbons (CFCs) **1**:284–286, **1**:286*F*
future outlook **1**:290
historical perspective **1**:282–284
methane (CH_4)
annual average measurements **1**:285*F*
causal factors **1**:284
methane budget **1**:284
nitrogen oxides (NO_x) **1**:282*T*, **1**:286–287
nitrous oxide (N_2O) **1**:284
nonmethane hydrocarbons **1**:282*T*, **1**:286
sulfur gases **1**:287
climate change **2**:58–59, **2**:82–83
concentrations **1**:281–282
coral reef ecosystems **2**:337, **2**:337*F*, **2**:341–342, **2**:342*F*, **2**:344
direct land-use change (DLUC) **4**:293–294
dynamic global vegetation models (DGVMs) **2**:680
ecosystem alterations
climate change effects **6**:149–150
coastal/marine ecosystems **6**:150–151
forest ecosystems **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:150–151
ecotoxicological studies **3**:118–119
enhanced greenhouse effect **4**:18–19, **4**:19*F*, **4**:19*T*, **4**:20*F*
fire/fire regimes **3**:441–442
forest ecosystems **3**:535, **6**:149–150, **6**:150*T*
future outlook **1**:290
global agriculture impacts **3**:427–429, **3**:428*F*
global trends **7**:83*F*
Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) model **4**:539–540
importance **1**:281–282
indirect land-use change (ILUC) **4**:293–294
life cycle analysis (LCA) **4**:539–540, **4**:628–629
lower atmosphere **1**:289–290
plant biodiversity threats **6**:62
plant conservation **6**:83–84
radiative forcing effects **4**:20*F*
savannas **1**:70–71
stratospheric ozone
concentration deviations **1**:289*F*
current trends **1**:287–289, **1**:289*F*
depletion causes and consequences **1**:289
future outlook **1**:290
occurrences and characteristics **1**:287–289

greenhouse gases (*continued*)
projected trends **1**:289
total ozone column **1**:287–289, **1**:288*F*
Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) model **4**:539–540
green iguana **3**:217
green lacewing **4**:306–307
Greenland
alpine life zone **1**:149*T*
botanical gardens **6**:507*T*
domesticated ruminants **1**:702
protected area systems **6**:493–494
urbanization **4**:560*F*
greenlings **3**:467–471
green manure **6**:594
green peafowl **7**:229*T*
green revolution **1**:87, **1**:123–124, **1**:467–468
green sea turtles **2**:356, **7**:229*T*
green sea urchins **3**:355–356
green sulfur bacteria **6**:25–26
green toads **5**:473
green turtles **2**:279, **3**:217, **3**:218, **5**:472, **6**:410–411
green-world hypothesis **3**:503–504, **3**:508–509
Grenada **6**:519
Grevillea **5**:221
Grevy's zebra **1**:51, **4**:695, **6**:573, **6**:576*F*
Grewia tiliaefolia **1**:266
Grewia villosa **1**:67
grey nurse shark **3**:459*F*
grey seals **2**:56*F*
grey squirrels **3**:191
grey wolf **7**:229*T*
see also wolves
griffon vultures **1**:709, **5**:474
Grifola frondosa **3**:637*T*, **3**:638
grigs **3**:727–728, **3**:728*F*
Grinnellian niche **4**:33–34, **4**:40–42
Griselinia **3**:522*T*
Griseliniaceae **3**:522*T*
grizzly bears **1**:631–632, **2**:242–243, **2**:571, **4**:62, **4**:564
gross domestic product (GDP) **3**:25, **3**:289, **7**:173–174, **7**:174*F*
gross national product (GNP) **3**:289
gross primary production (GPP) **3**:240
ground beetles **1**:326–328
ground finches **1**:27–28, **1**:28*F*, **2**:140
ground fires **3**:436
groundfish **2**:12*F*
ground sloths **4**:689, **4**:724–725*T*
groundwater
freshwater ecosystems
freshwater invertebrates 4:369–378
altered water regimes **4**:377–378
background information **4**:369–370
biodiversity loss **4**:376–377
decomposition and recycling functions **4**:375
disturbances **4**:375
ecological functions and processes **4**:375
endemism **4**:374–375
exotic species invasions **4**:378
global patterns **4**:373
habitat depletion and loss **4**:376–377
habitat heterogeneity **4**:375
influencing factors **4**:373–374
isolation factors **4**:374–375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats **4**:376–377
material transfer and uptake **4**:375–376
number of species **4**:370–371, **4**:371*T*
origins **4**:374
pollution **4**:377
primary production **4**:376
reduced water quality **4**:377
research background **4**:373–374
river continuum concept (RCC) **4**:373*F*
species interactions **4**:375
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
freshwater sediments **3**:572*T*
geographic distribution **4**:370*T*
habitat diversity **3**:570–571
protected area systems **6**:497
water storage **3**:561*T*
groundwater calcretes (Western Australia) **7**:59
groupers **1**:190*T*, **3**:450, **3**:468, **3**:468*F*
Group on Earth Observations Biodiversity Observation Network (GEO BON) **3**:101–102, **4**:326
grouse **4**:578–579
grouse locusts **3**:736
growing seasons **2**:84
growth overfishing **3**:445
Grubbiaceae **5**:219–220
Grubea australis **5**:661*F*
Gruiformes **6**:308*T*, **7**:181–182
grunts **3**:467–471
Grus americana **7**:189
Grus canadensis **7**:189
Grus japonensis **3**:373, **7**:188–189
Gryllacrididae (raspy crickets and leaf-rolling crickets) **3**:732
Gryllidae
adaptive radiation **1**:32–33
Brachytrupinae **3**:729–730
characteristics **3**:729
Gryllinae **3**:729
Oecanthinae **3**:730
Phalangopsinae **3**:730
Gryllinae **3**:729
Grylloblattodea **4**:308–312, **4**:309*T*
Grylloidea
characteristics **3**:728, **3**:728*F*
Gryllidae
Brachytrupinae **3**:729–730
characteristics **3**:729
Gryllinae **3**:729
Oecanthinae **3**:730
Phalangopsinae **3**:730
Gryllotalpidae **3**:728–729
Mogoplistidae **3**:730
Myrmecophilidae **3**:730
Gryllotalpidae **3**:728–729
Gryllus campestris **3**:225
Gryllus veletis **3**:728*F*
Gryposaurus **2**:519*T*, **2**:520*T*
guaco forests **3**:79–80, **3**:81*T*
Guadalupe-Dzhulfian event **3**:415–418, **3**:416*T*
Guadalupe fur seal **5**:81*T*
Guadalupian Stage **5**:169*F*
Guadeloupe **6**:519
Guaiacum **5**:622*F*
guaje **1**:120
Guam **1**:592, **4**:663, **6**:507*T*
Guam rail **1**:621, **1**:664*T*, **1**:666–667, **2**:245
guamuchil **1**:120
guanacaste trees **6**:52
guanacos **2**:476*T*, **2**:590*T*, **3**:115, **3**:127*T*, **6**:334
guancaste **4**:592
guano **7**:51–52
Guarea excelsa **4**:595
Gua Salukkan Kallang-Towakkalak (Indonesia) **7**:59–60
Guatemala
botanical gardens **6**:519
deforestation **2**:30–31, **2**:31*F*
ecosystem services **7**:356*T*
indigenous peoples **4**:270
multiple cropping systems **1**:120
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
oil-palm plantations **5**:601
species richness **2**:25*T*
water funds **7**:356*T*
wheat production **1**:87–88
Guatemalan fir **3**:208
guava **2**:390, **2**:663–664, **4**:287–290
Guayana Highland **1**:166
Guaymas Basin **7**:327–328
guazatine **5**:695
Guembelitria **5**:164–165
guenons **4**:692, **6**:251*T*
guereza **6**:254
guide RNAs (gRNAs) **5**:572, **5**:575*F*
guilds **4:33–38**
characteristics and structure **4**:33–35, **4**:34*F*
ecological boundaries **4**:37
predator guilds **1**:690–692
research benefits
functional community organization **4**:35–36, **4**:36*T*
manageable unit size **4**:35
trophic structure framework **4**:36
species individualism **4**:36–37
vaguely bounded categories **4**:37–38
Guinea
botanical gardens **6**:507*T*
energy efficiency measures **3**:254*F*
forest-savanna mosaics **1**:68
nitrogen deposition **5**:526*T*
plant diversity **6**:82*T*
rice production **1**:88
Guinea-Bissau **6**:507*T*
guinea fowl **1**:644*F*

guinea worm **5**:668, **5**:671*T*
Guizhou snub-nosed monkey **6**:256–258*T*
Guizotia abyssinica **2**:663
Gulf of Alaska beluga whales **5**:77*T*, 5:79–80
Gulf of California **3**:171–173*T*
Gulf of California Desert **2**:468–469
Gulf of Khambhat **1**:276*T*
Gulf of Kutch **1**:276*T*
Gulf of Maine Area Program (GoMA) **2**:16, **2**:17*F*
Gulf of Mexico **2**:14, **2**:14*F*, **2**:333*T*, **3**:171–173*T*, **7**:372–373, **7**:374*F*
Gulf of Oman desert **2**:467–468
Gulf of Thailand **1**:276*T*
gulls **7**:390
Gulo gulo **7**:27
gumwood trees **4**:361
Gund Institute for Ecological Economics **5**:460–462*T*
gunnels **3**:467–471
Gunnera **5**:622*F*
Gunnera spp. **2**:25–26
guppies **2**:140, **2**:718, **3**:467
Gustavia augusta **1**:163
Gustavia hexapetala **1**:163
Gusum, Sweden **1**:140–141, **1**:142*F*
Gutenberg, Johannes **4**:651
Guttiferae **3**:129*T*, **3**:130*T*
Guyana
  botanical gardens **6**:519
  deforestation **1**:677*F*
  land cover classifications **6**:607*T*
  nonvascular epiphytes **3**:521*T*
  tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*
Guyana Shield **3**:105–107, **3**:106*F*
Guyanian Shield **1**:247
Gymnocarpos decander **5**:484–485
Gymnoderma lineare **3**:214–215
Gymnodinium **1**:472, **1**:473*F*, **6**:294
Gymnodinium beii **6**:291
Gymnodinium breve **6**:131–132, **6**:134*F*
Gymnodinium catenatum **1**:472–473, **6**:30
Gymnodinium simplex **6**:294–295
Gymnodiniurri **6**:30
Gymnogyps californianus **1**:621–622, **2**:273, **7**:185
Gymnolaemata **3**:685*F*
Gymnophalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Gymnophiona **1**:169, **1**:170*T*
gymnosperms
  abundance patterns **6**:73, **6**:74*F*
  anthropogenic activities and impacts **3**:401*F*
  Arctic ecosystems **1**:231*T*
  Asia **1**:258, **1**:259*T*
  biome distributions **1**:64*F*
  characteristics **6**:59–60
  edible species **3**:129*T*
  evolutionary development **2**:643, **6**:65–67, **6**:66*F*
  extinction rates **5**:368*T*, **6**:73*F*
  North America **5**:564–565
  origination rates **6**:71–72, **6**:72*F*
  taxonomic distributions **3**:519*T*
gymnostome ciliates **6**:295–296, **6**:296*F*
gymnures **4**:691
Gypaetus barbatus **1**:709, **5**:474
Gypohierax angolensis **5**:603
Gyps fulvus **1**:709, **5**:474
Gypsisols **2**:461–464
gypsy moths **2**:551–552*T*, **4**:183, **4**:361, **5**:395–396, **5**:402–403
Gyrinophilus porphyriticus **4**:516
Gyrocarpus **5**:621*F*
Gyrocotyle **7**:456–457*F*
Gyrocotylidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Gyrodactylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Gyrodactylus **7**:448–449*F*
Gyromitra **3**:635*T*
gyromitrin **3**:635*T*
Gyrotoma alabamensis **3**:185*F*

H

H5N1 influenza virus **1**:519–520
Habenaria **6**:316*T*
Haber reaction **1**:561, **5**:520–522
habitat corridors *see* landscape corridors
Habitat HarshnessHypothesis **2**:134
habitats
  alpine ecosystems
    climatic microhabitats **1**:149–150, **1**:150*F*
    substrate diversity **1**:150–151
  Amazon ecosystems **1**:249–250, **1**:250*T*, **1**:251*F*, **1**:252*F*
  anthropogenic activities and impacts **7**:307*T*, **7**:308*T*
  aquatic subterranean habitats **7**:49–50, **7**:50*F*
  biodiversity dynamics modeling **5:321–328**
    background information **5**:321–322
    phenomenological models
      dose–response models **5**:324, **5**:324*F*
      global species distributions **5**:324
      species–area relationships **5**:322–323
      statistical estimation models **5**:324
    process-based models
      integro-difference equations **5**:324–325
      multispecies metapopulation analysis **5**:325–326, **5**:325*F*, **5**:326*F*
      neutral theory **5**:326–327
      reaction–diffusion models **5**:324–325, **5**:325*F*
    research summary **5**:327
    species–area relationships
      countryside habitats **5**:323–324, **5**:323*F*
      native habitats **5**:322–323, **5**:322*F*
  biodiversity measures **1**:404
  climate change studies **2**:66, **2**:92–94
  coevolution effects **2**:147–148
  conservation and management practices **4**:47–48
  crustaceans **2**:407–408
  deep-soil habitats **7**:51
  European ecosystems
    classifications **3**:334
    diversity
      biotic interactions **3**:334*T*, **3**:335
      habitat diversity **3**:334–335, **3**:334*T*
      human interference **3**:334*T*, **3**:335
      pattern variations **3**:335, **3**:344
      prerequisites and sources **3**:334–335, **3**:334*T*
      taxon diversity **3**:334, **3**:334*T*
    pattern variations
      biomes **3**:335, **3**:335*T*
      habitat diversity **3**:344
      vegetation maps **3**:340–344, **3**:345*F*, **3**:346*T*
  forest canopies **3**:524
  forest management impact **1**:452–453
  global climate change
    background information **3**:407–408
    climate change predictions **3**:408, **3**:408*F*
    habitat loss and fragmentation **2**:58–59, **2**:80, **3**:221, **3**:408, **4**:56
  government legislation and regulations **3**:718
  habitat degradation and pollution **2**:531
  habitat destruction
    biodiversity loss **4**:662
    coral reef ecosystems **3**:404, **3**:405*F*
    desertification **3**:404–405
    endangered reptiles **3**:216–217
    general discussion **3**:403
    grasslands **3**:404
    habitat loss **3**:403
    mangrove ecosystems **3**:404
    marine sedimentary habitats **5**:113*F*, **5**:114
    migration impacts **5**:317–318
    primate threats **6**:253
    rain forest loss **3**:403–404, **3**:403*F*
    tropical deciduous forests **3**:404
    wetlands **3**:404, **7**:398
  habitat heterogeneity **1**:381–382, **1**:381*F*
  habitat loss and fragmentation **4:50–58**
    agricultural growth impacts **4**:50, **4**:51*F*, **4**:56
    agricultural intensification effects **1**:374–375
    amphibian species decline **3**:695
    anthropogenic activities and impacts **3**:401–402, **3**:401*F*, **4**:52–53, **4**:662
    basic concepts
      definition **4**:50–51
      habitat loss versus cover type loss **4**:51–52
      human-caused fragmentation versus natural fragmentation **4**:52–53
      landscape complementation **4**:52
      landscape scale versus patch scale **4**:51, **4**:53*F*
      limitations **4**:50–51
      loss versus fragmentation **4**:51

habitats (*continued*)
precise conceptualization **4**:53–54
spatial extent **4**:52
biodiversity loss **4**:564, **4**:565*F*, **4**:566*F*, **4**:662, **4**:663
biodiversity-rich countries **1**:501–502
characteristics **3**:405–406
climate change effects **2**:58–59, **2**:80, **3**:221, **3**:408, **4**:56, **6**:62
crop wild relatives (CWRs) **6**:77
dispersal barriers **3**:405–406
edge effects **3**:406
endangered freshwater invertebrates **3**:180–181, **3**:180*F*
endangered mammals **3**:190
environmental impacts **3**:283–284
fish species **3**:447–448
forest management impact **1**:452–453
fragmentation effects **4**:50, **4**:52*F*
freshwater invertebrates **4**:376–377
general effects
advantages/disadvantages **4**:54–55
anthropogenic-caused environmental effects **4**:56, **5**:318
extinction probability **4**:54
loss versus fragmentation **4**:54, **4**:54*F*
migration impacts **5**:318
mobility–habitat specialization–matrix quality interactions **4**:55–56
reproductive rates **4**:55
species richness reduction **4**:54
global agriculture impacts **3**:427, **3**:428*T*, **4**:154
global urbanization impact **4**:234–235
infectious disease incidence **1**:361–362, **1**:363*F*
infectious diseases **2**:531
interspecific interactions **3**:406
landscape changes **4**:496–497
landscape genetics **4**:516–517
land use issues **4**:564, **4**:565*F*, **4**:566*F*
loss effects **4**:50, **4**:51*F*
marine sedimentary habitats **5**:113*F*, **5**:114
migration impacts **5**:318
pollution **3**:181, **3**:181*F*
predation effects **6**:237–239
research summary **4**:56
terrestrial invertebrates **3**:220–221
threatened birds **7**:183
tropical forests **3**:405–406, **3**:405*F*
wetlands **7**:371*T*
habitat modification
general discussion **4**:359
grazing and rooting **4**:359
introduced animals **4**:364–365
introduced plants **4**:364
modified fire regimes **4**:360
modified hydrology **4**:360
modified nutrient regime **4**:360
shading and overgrowth **4**:359–360
habitat–niche relationship **4:39–49**
basic concepts **4**:39–40, **4**:40*F*
conservation and management practices **4**:47–48
ecological character displacement **4**:46–47, **4**:46*F*, **4**:47*F*
evolutionary development **4**:46–47
food resources **4**:44*F*
future research outlook **4**:48
historical perspective **4**:42
indirect interactions **4**:45–46
multispecies comparisons **4**:42–45, **4**:42*F*, **4**:43*F*, **4**:44*F*, **4**:45*F*
mutualistic interactions **4**:45
representative species **4**:41*T*
habitat selection **4:59–69**
analytical methods
animal movement models **4**:64
general discussion **4**:63
linear and logistic regression models **4**:63
utilization distributions **4**:63–64
animal location measurements
direct visual observations **4**:60–61, **4**:60*F*
telemetry techniques **4**:60*F*, **4**:62
trapping methods **4**:60*F*, **4**:61–62
visual and auditory surveys **4**:60*F*, **4**:61
basic concepts **4**:59–60
emotional responses **1**:41–42
environmental variable measurements **4**:62–63
inference methods
basic concepts **4**:64
habitat availability **4**:65–66, **4**:65*F*
limitations **4**:64–65
locational imprecision **4**:67
scale and resolution **4**:66, **4**:67*F*
selection decisions **4**:65
spatial and temporal autocorrelation **4**:66–67
location techniques **4**:59–60, **4**:60*F*
research summary **4**:67–68
high-elevation Andean ecosystems **4**:96–97
international conservation organizations **4**:325–326
landscape ecology **4:488–502**
population dynamics
basic concepts **4**:488–491
continuous habitat variability **4**:497–498, **4**:498*F*
core and satellite model **4**:496
disturbance recovery **4**:498–499, **4**:499*F*
edge effects **4**:496–497
environmental gradients **4**:497–498
extinction debt **4**:497
habitat fragmentation **4**:496–497
habitat heterogeneity **1**:381–382, **1**:381*F*, **4**:497–498
habitat preservation **4**:495
immigration **4**:495
interspecific interactions **4**:498–499, **4**:499*F*
landscape changes **4**:496–497
landscape context **4**:493
landscape structure **4**:491–493
metapopulation dynamics **4**:495–496
patch–matrix–corridor model **4**:479, **4**:479*F*, **4**:492–493, **4**:492*F*
population viability analysis (PVA) **4**:493–495
quantification methods **4**:493
research summary **4**:499–500
simulation models **4**:488–491, **4**:490*F*, **4**:491*F*
single large or several small (SLOSS) debate **4**:495
soil moisture prediction **4**:497–498, **4**:498*F*
source–sink dynamics **4**:496, **5**:324–325
spatial patterns **4**:488–491, **4**:489*F*
mangrove habitat **5**:10–11, **5**:10*F*
marine sedimentary habitats
deep-sea habitats **5**:108
intertidal habitats **5**:107, **5**:108*T*
ocean depth ranges **5**:107, **5**:107*F*
subtidal habitats **5**:107–108
natural habitats **1**:381
Near East ecosystems
habitat destruction
coastal sand dunes **5**:472, **5**:473*F*
riverine habitats **5**:473
swamps **5**:472–473, **5**:474*F*, **5**:475*T*
temporary winter rain pools **5**:473, **5**:476*F*
wetland habitats **5**:472–473
noncrop habitats **1**:380–381
planktonic communities **6**:27–28, **6**:28*F*, **6**:29*F*
pollution-based habitat degradation
air pollution
acid rain **3**:407, **3**:407*F*
lichen tolerance **3**:407
ozone depletion **3**:407
smog **3**:407
ultraviolet radiation **3**:407
characteristics **3**:406
pesticides **3**:406
water pollution **3**:406–407
species richness
diversity–productivity relationship **2**:597–599, **2**:598*F*
prediction limitations **2**:599
species–energy theory **2**:596–597, **2**:597*T*
Volterra–Gause perspective **2**:595–596
sustainability trends **7**:82*F*
terrestrial subterranean habitats **7**:50–51
Habitat Suitability Index (HSI) models **2**:218
hachetfish **3**:465
Hadaway, Hunter **2**:690–691
haddock **3**:466–467
Hadenoecus subterraneus **7**:60
Hadesia vasiceki vasiceki **7**:58
Hadley cells **2**:458
Hadogenes troglodytes **1**:207–208
Hadrobregmus pertinax **7**:218–219
Hadropithecus **5**:365
Hadrosauridae **2**:520*T*
Hadrosaurus **2**:519*T*
Haeckel, Ernst **4**:652, **5**:581–582, **7**:87

Haeckel's classification diagram **2**:50*F*, **2**:51, **7**:87, **7**:88*F*
Haematopinus oliveri **3**:226
Haematopus bachmani **5**:57
Haematopus chathamensis **7**:185
Haematopus moquini **6**:232
hagfish
  characteristics **3**:456–457, **3**:457, **3**:458*F*, **7**:336–337
  classification systems **7**:335, **7**:335*F*, **7**:336*F*
  threatened species **7**:201*T*
Hagloidea **3**:727–728, **3**:728*F*
Haikouela lanceolata **7**:334–335
Hainan crested gibbon **6**:256–258*T*
Hainan gibbon **6**:250
hairgrass **1**:182
Hairston-Smith-Slobodkin (HSS) hypothesis **6**:723
Hairy Beggarsticks **3**:704–705
hairy-eared mouse lemur **6**:251*T*
Haiti **5**:493*T*, **5**:495*F*, **5**:497*F*, **6**:493–494, **6**:519
hake **3**:466–467
Hakea **5**:221
Hakea sericea **6**:94*F*
Halacarida **5**:470*T*
Halcyon smyrnensis **5**:463*T*, **5**:603–604
Haldane, J. B. S. **5**:630, **6**:2
Haldane's rule **6**:636–637
halfbeaks **3**:467
Haliaeetus leucocephalus **3**:256, **3**:406, **4**:567
Haliaetus albicilla **5**:474
Halichondria panacea **6**:723*T*
Halidrys dioica **4**:334, **4**:335*F*
Halimeda **6**:132–135, **6**:134*F*
halimedatetraacetate **6**:134–135, **6**:134*F*
halimedatrial **6**:132–135, **6**:134*F*
Halimium **5**:218
Haliotis cracherodii **3**:201*T*, **3**:202*T*, **4**:391
Haliotis kamtschatkana **3**:201–202
Haliotis sorenseni **3**:201*T*
Haliotis spp. **4**:345–346
halkieriids **5**:374*F*
hallucinations **3**:635*T*
Haloarcula **1**:222–223, **1**:223*T*
halobacteria **6**:9–11
Halobacterium **1**:222–223, **1**:223*T*, **2**:640–641, **6**:301
halocarbons **1**:284–286, **1**:286*F*
Halocnemum strobilaceum **1**:268, **1**:270
Halococcus **1**:222–223, **1**:223*T*
Halodule **6**:542*T*
Halodule pinnifolia **6**:541*T*
Halodule uninervis **6**:541*T*
Halodule wrightii **3**:360–361, **6**:541*T*
Haloferax **1**:222–223, **1**:223*T*
Halomonas **6**:301
halon gas **1**:282*T*, **1**:286
Halophila **6**:542*T*
Halophila baillonis **6**:541*T*
Halophila beccarii **6**:541*T*
Halophila capricornii **6**:541*T*
Halophila decipiens **6**:541*T*
Halophila engelmannii **6**:541*T*
Halophila minor **6**:541*T*
Halophila ovalis **6**:541*T*
Halophila spinulosa **6**:541*T*
Halophila stipulacea **6**:541*T*
halophilic Archaea **1**:222–223, **1**:223*T*
Halophiloudeae **6**:542*T*
Haloragis **5**:622*F*
halorhodopsins **6**:281*F*
Halorubrum **1**:223*T*
Halosarcia **2**:376*F*
Halostachys caspica **1**:270, **2**:468
Halothamnus **2**:376*F*
Haloxylon **2**:371*T*, **2**:376*F*
Haloxylon ammodendron **1**:270
Haloxylonetum persici **5**:486
Haloxylon negevensis **5**:485–486
Haloxylon persicum **1**:268, **1**:270, **5**:486
Haloxylon scoparium **5**:484–485
Halteria **6**:31, **6**:296*F*
Hamadryas kingii **4**:106–107
Hamadryas sempervivens **4**:106–107
Hamadryas sempervivoides **4**:107
Hamamelidaceae **5**:622*F*
Hamamelis obtusata **1**:271
Hamamelis virginiana **5**:712*T*
Hamilton's frog **4**:194*T*
hammada **2**:462
hammock spider **3**:222
Hancornia speciosa **1**:165, **4**:286–287, **4**:288*T*
Hand, Dana **4**:654
handfish **3**:466–467
Hangenberg event **5**:169–170
Hanseniella arborea **5**:431
Hanseniella magna **5**:424
Hanseniella submarina **5**:430
hantaviruses **1**:365–368, **1**:366*F*, **3**:349
Haor Basin **1**:276*T*
Hapalemur alaotrensis **6**:256–258*T*
Hapalemur aureus **6**:256–258*T*
Haplobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Haplobothrium **7**:456–457*F*
Haploclastus kayi **3**:225–226
haplodiploidy **5**:676–678
Haplogynae **1**:207
Haplomastodon **4**:727*T*, **5**:231–233
Haplopappus gracilis **2**:373*T*
Haplopharyngida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
Haplopharynx **7**:442–443*F*
Haplophrentis carinatus **5**:375*F*
Haploporoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Haplorhini **4**:692
Haplorrhini **6**:250–252, **6**:251*T*
Haplosplanchoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
haplosporidians **5**:669–670*T*, **6**:296–297
Haplosporidium nelsoni **5**:672, **5**:672*F*
Haptolina **1**:474*F*
haptomonads **6**:293–295
Haptophyta **1**:473–474, **1**:474*F*, **6**:10*T*
haptorids **6**:295–296, **6**:296*F*
harbor porpoise **5**:77*T*
harbor seals
  deoxyribonucleic acid (DNA) evolution **2**:56*F*
  introduced parasites–species decline relationships **2**:528*T*
  oil spills **3**:256
  pollution effects **3**:192–193
Hardy–Weinberg principle **6**:183
hares **1**:80, **1**:81
Harike Lake **1**:276*T*
harmful algal blooms **1**:470–484
  biogeographical range **1**:470
  bloom dynamics **1**:475
  chemical ecology **1**:478–479
  coastal ecosystems **5**:543
  ecogenomic sensors **2**:697*F*
  future outlook **1**:481–482
  global change effects **1**:479–481
  global distribution **1**:470, **1**:471*F*
  isozyme analyses **1**:475
  lake and pond ecosystems **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*, **5**:25
  life histories **1**:478*F*
  microsatellite studies **1**:475, **1**:477*T*
  nutrient enrichment impacts **1**:95–97, **1**:96*F*
  phylogenetic and taxonomic relationships
    dinoflagellates
      group origins **1**:472, **1**:473*F*
      toxic species origins **1**:472–473, **1**:473*F*
    general discussion **1**:470–472
    haptophytes
      group origins **1**:474
      toxic species origins **1**:474–475, **1**:474*F*
    heterokonts
      group origins **1**:475
      toxic species origins **1**:475, **1**:476*F*
  population genetics **1**:475
  species interactions **1**:478–479
  toxin distribution and characteristics **1**:478–479
Harmonia axyridis **1**:379
harpacticoids **4**:381*T*
Harpagornis moorei **5**:366
harp seals **5**:90, **5**:314, **7**:37–38
Harrison, Benjamin **2**:281
hartebeest **3**:191, **4**:696
Hart's tongue fern **7**:52–53
harvester ants **1**:709–710
Harvesters, The (Bruegel) **2**:347, **2**:347*F*
harvestmen
  characteristics **1**:209–210, **1**:209*F*
  ecological characteristics **1**:210
  genera and species **1**:202*T*
  phylogenetic relationships **1**:205*F*
  phylogeny and taxonomy **1**:210
  reproduction and growth **1**:210
  subterranean ecosystems **7**:53
Harvey, William **5**:629–630
Hastigerina pelagica **6**:291
Hatinh langur **6**:256–258*T*
Haury–McGowan–Wiebe diagrams **4**:669
haustorium/haustoria **4**:70–71, **4**:74–75
Havetiopsis **3**:522*T*
Hawaii
  adaptive radiation **5**:593–594

Hawaii (*continued*)
biogeographical analysis **4**:402*F*, **4**:403–405, **4**:404*F*, **4**:405*F*
Bowie Seamount Marine Protected Area (Sglaan Kinghlas) **5**:446–447
coral reef ecosystems **2**:333*T*
countryside biogeography studies **2**:356, **2**:357*F*
ecotones **3**:143*F*
endangered plant species **3**:205–206
endemism **3**:233, **5**:592*F*, **5**:592*T*
extinction rates **5**:595–597
fire ant invasions **5**:598*F*
first-contact extinctions **5**:365–366
future research areas
realistic evolutionary trajectories **5**:597
restoration **5**:599
Hawaii Volcanoes National Park **4**:402*F*, **4**:403–405, **4**:404*F*, **4**:405*F*
historical extinctions **5**:367
innovative behaviors **5**:593
introduced diseases **3**:411
invasive species **4**:353
living dead **6**:81
Papahanaumokuakea Marine National Monument **5**:446
plant invasions **6**:93
species diversification **5**:594, **5**:595*F*
succession/regression theory **4**:397–398, **4**:398*F*
threatened birds **7**:181
Hawaiian bobtail squid **5**:122*F*, **5**:123–124
Hawaiian crows **5**:367
Hawaiian *Drosophila*
adaptive radiation **1**:28–29
hybridization **1**:26–27
Hawaiian duck **7**:184
Hawaiian finches **1**:29
Hawaiian honeycreepers **1**:29, **1**:29*F*, **1**:618–619, **2**:528*T*, **6**:350
Hawaiian monk seal **5**:77*T*
Hawaiian silverswords **1**:30–32, **1**:30*F*
Hawaiian swordtail crickets **1**:32–33
Hawaiian *Tetragnatha* spiders **1**:32, **1**:33*F*
Hawaii Ocean Time-series (HOT) program **2**:616–617
Hawkes Bay **1**:276*T*
hawk moths **6**:47–48, **6**:49*T*, **6**:142
hawks **1**:709
hawkweed **6**:125
hay meadows **4**:577
Hayoceros **4**:724–725*T*
Haytley, Edward **2**:347, **2**:347*F*
hazard **1**:39
hazel **4**:574*F*
headaches **3**:635*T*
Healing Forest Conservancy (HFC) **1**:598
Heard Island **1**:180–181, **1**:186*T*
heat-generating transformations **2**:638–639
heather moors **4**:577–579, **4**:578*F*, **7**:223
heathlands
Australia **1**:292–297, **1**:294*T*
C_4 plants **2**:371*T*
historic land-use patterns **4**:577, **4**:578*F*
heat shock proteins **5**:36–37
heavy metals **1**:137*T*, **1**:140–141, **3**:118–119
Hector's dolphin **3**:192, **5**:81*T*
Hedera helix **5**:485
Hedgehog (*Hh*) signaling gene **7**:53–54
hedgehogs **2**:550, **4**:690*T*, **4**:691
hedgerows **1**:380–381
hedges **4**:571, **4**:571*F*
hedonic valuation estimation methods **7**:322–324
Hedophyllum sessile **6**:723*T*
Hedycarya **5**:621*F*
Hedylidae **5**:384–385, **5**:385*F*, **5**:387–388*T*
Hedyosmum **7**:274–275
Heleomyzidae **3**:489–491*T*
Heliamphora **6**:44*T*
Helianthemum **5**:218
Helianthus annuus **2**:383, **2**:391, **2**:636*F*, **2**:663–664, **4**:167, **6**:626–627, **7**:25
Helianthus anomalus **4**:167, **6**:626–627
Helianthus deserticola **4**:168
Helianthus paradoxus **4**:167
Helianthus petiolaris **2**:636*F*, **4**:168, **6**:626–627
Heliaster solaris **3**:201–202
Helicobacter pylori **2**:636*F*
Heliconia **1**:263
Heliconiaceae **6**:316*T*
Heliconiinae **1**:653
Heliconius **2**:510, **2**:512
Heliconius erato **2**:138
Heliconius melomene **2**:138
Helicoverpa armigera **4**:298–299, **4**:299*F*
Heligmosomoides thomomyos **7**:432, **7**:432*T*
Heliobacteriaceae **6**:10*T*, **6**:11*T*
Heliopora coerulea **3**:199
Heliosphaera **6**:290*F*
heliozoans **6**:293
Helisoma trivolvis **1**:368, **1**:368*F*
Helixanthera mannii **4**:77
Helixanthera parasitica **4**:77
Helix pomatia **3**:222–223
hellbenders **7**:389–390
Hell Creek Formation (US) **2**:518–519, **2**:519–521, **2**:521*F*
hellgrammites **4**:371*T*
Helminthosporium maydis **1**:460
helminths
characteristics and properties **2**:524*T*
disease regulation **1**:348–350, **3**:349
domesticated ruminants **1**:710–711
Heloderma **7**:340
Heloderma suspectum **7**:229*T*
Heloecius cordiformis **4**:337, **4**:339*F*
Helosciomyzidae **3**:489–491*T*
Helwingiaceae **1**:260*T*
Hemeroplanes **5**:392*F*
Hemiauchenia **4**:724–725*T*, **4**:727*T*
Hemiaulus **2**:612*T*
Hemichordata
classifications and sampling methods **4**:381*T*
fossil record **3**:684*F*
global diversity **4**:385*T*
known and estimated species **4**:384*T*
marine ecosystems **5**:50*T*
number of species **5**:669–670*T*
Hemidactylus frenatus **4**:360
Hemiechinus auritus **5**:472
hemiepiphytes
characteristics **3**:518
taxonomic distributions **3**:520–521, **3**:522*T*
tropical rain forests **6**:315*T*
Hemignathini **1**:29
Hemignathus lucidus **1**:29*F*
Hemignathus munroi **1**:29*F*
Hemignathus obscurus **1**:29*F*
Hemignathus parvus **1**:29*F*
Hemignathus virens **1**:29*F*
Hemignathus wilsoni **1**:618–619
Hemigrammocypris rasborella **4**:465
Hemileuca maia **5**:399*F*
Hemimastodon **5**:224*T*
hemimetabolous insects **5**:674–675
Hemingway, Ernest **4**:652
hemiparasitism **4:70–78**
characteristics and classifications **4**:70–71
ecological advantages **4**:75
evolutionary development **4**:73–74
haustorium/haustoria **4**:74–75
host specificity **4**:75
mycorrhizal associations **4**:74
pests
crop production **4**:75–77
dodders **4**:70, **4**:77
mistletoes **4**:70, **4**:77
research summary **4**:77–78
root parasitic weeds **4**:76–77
phylogenetic relationships **4**:71–73, **4**:72*F*
physiological characteristics **4**:75
plant–plant interactions **4**:74
seed dispersal/germination **4**:74
Hemiptera **7:287–295**
Heteroptera
characteristics
coloration and ornamentation **7**:291
feeding behaviors **7**:288
habitat **7**:288
host specificity **7**:290
mimicry **7**:291
morphology **7**:288–292
parasitization **7**:292
physical characteristics **7**:288
predatory behaviors **7**:289
reproduction and growth **7**:289–290
scent glands **7**:288
economic importance **7**:292–293
infraorders
Aradomorpha **7**:294
characteristics **7**:293–294
Cimicomorpha **7**:293
Gerromorpha **7**:293
Leptopodomorpha **7**:293
Nepomorpha **7**:293
Pentatomomorpha **7**:294
number of species **4**:392*T*, **7**:287
research summary **7**:294–295
terminology **7**:287–288
insecticide resistance **4**:298–299
major divisions and characteristics **4**:308–312, **4**:309*T*
number of species **4**:392*T*
suborders

Coleorrhyncha **7**:287
Heteroptera **7**:287
Homoptera **4**:392*T*, **7**:287
research controversies **7**:287
Hemiscorpiidae **1**:209
Hemitragus jemlahicus **2**:551–552*T*
Hemiuroidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
hemlock
boreal forests **1**:628–630
introduced species **4**:361
North America **7**:118
plant secondary chemicals **6**:42
hemoglobin evolution **2**:55–56, **2**:55*F*, **2**:55*T*
hemophilia **6**:3–5
hemp **2**:663
Hennigian argumentation **2**:39, **2**:53, **6**:16, **7**:85, **7**:104–106
Hennig, Willi **7**:89*F*, **7**:90
hen-of-the-woods mushrooms **3**:637*T*
Hensen, Victor **5**:581–582
Henslow, John Stevens **2**:419
Henslow's sparrow **4**:489–490, **4**:490*F*
Hepialidae **5**:386
Hepialoidea **5**:386*F*, **5**:387–388*T*
Heptagyia **7**:345–346, **7**:347*F*
Heracleum mantegazzianum **3**:705, **6**:92, **6**:93
herbaceous plants **5**:528–530, **5**:534*F*
herbaceous vegetation species richness **4:79–86**
control mechanisms
intermediate disturbance hypothesis **4**:79
intraspecific diversity **4**:79–80, **4**:80*F*
dominants and subordinates **4**:82–83, **4**:82*F*, **4**:84
general discussion **4**:79
high species richness effects **4**:81–82
mass ratio hypothesis **4**:80–81, **4**:81*F*
plant traits **4**:80
research summary **4**:84–85
species declines
filter effects **4**:84
founder effects **4**:84
transients **4**:83–84, **4**:83*T*, **4**:84
The Herball, or General Natural History of Plants (Gerard) **1**:588–589, **5**:705
herbals and medicinal plants **1**:359, **5**:705–706, **6**:518–519
herbal teas **3**:130*T*, **3**:131*T*
herbaria **1**:504, **1**:505*T*, **6**:510–511
Herbarium Amboense **1**:589
Herbarium of Apuleius Platonicus **5**:705–706
herbicides **4:87–95**
animal damage control **6**:342
biodiversity effects
increased insect activity **4**:93
weed control practices
community composition and structure changes **4**:91–92
herbicide resistance **4**:92–93, **4**:92*T*
selection pressure **4**:91–92
characteristics and classifications **4**:89, **4**:90*T*
environmental fate **4**:89–90
environmental threats **3**:166
general discussion **5**:695–696
government regulation **4**:90–91
inorganic herbicides **5**:695–696, **5**:699*T*
Near East ecosystems **5**:473–475
organic herbicides
aliphatic acids **5**:696
benzonitriles **5**:697
bipyridiliums **5**:696, **5**:699*T*
carbamates **5**:696
dinitroanilines **5**:696
diphenyl ethers **5**:696
heterocyclic nitrogens **5**:696
organic arsenicals **5**:696
phenol derivatives **5**:696–697
phenoxyaliphatic acids **5**:696
substituted amides **5**:696
substituted ureas **5**:696
thiocarbamates **5**:696
triazines **5**:696, **5**:699*T*
river ecosystems **6**:481*T*, **6**:482
summary discussion **4**:94
toxicity **4**:91*T*
usage **4**:87
herbivores
adaptive radiation **1**:22, **1**:23*F*
African savannas **1**:61, **1**:65
antagonistic interactions **6**:39–41
boreal forests **1**:631–632, **1**:633*F*
convergent evolution **4**:701*T*
coral reef ecosystems **3**:619–620, **3**:620*T*
desert ecosystems **2**:470*T*, **2**:472–474
Diptera **3**:493–494
disturbance ecology **2**:565–566
ecosystem function measurements
food chain controls **3**:78, **3**:80*F*
nutrient cycling **3**:77–78
trophic transfer **3**:76–77
fire/fire regimes **3**:440
food habits **3**:127–128, **3**:127*T*
grasslands **4**:3–4
grazing effects **4:8–17**
direct versus indirect effects **4**:10
disturbance regimes **4**:9–10, **4**:10*F*
grazer-based ecosystem restoration and management
background information **4**:14–15
coral reef ecosystems **4**:16
forests **4**:15
grasslands **4**:14–15
invasive species **4**:16–17
lakes and ponds **4**:15–16, **4**:16*F*
herbivory **4**:8–9, **4**:9*F*
host chemistry effects **4**:13
paleo-patterns **4**:13–14
research summary **4**:17
spatiotemporal scale
between-habitat patterns **4**:11–12
functional role **4**:10–11
geographic patterns **4**:11, **4**:11*F*
microhabitat patterns **4**:12–13
within-habitat patterns **4**:12
hot deserts **6**:334
introduced species **4**:361
keystone species study **4**:446–447, **4**:446*F*
lake ecosystems **6**:233–235
megaherbivores **5:223–239**
conservation challenges **5**:234–236
dietary tolerance **5**:223–226
extant species **5**:225*T*
extinctions **5**:225*T*, **5**:231–233, **5**:232*F*
future outlook **5**:236
grazing impacts **5**:227–228
human impacts **5**:233–234
late Pleistocene extinctions **5**:231–233, **5**:232*F*
life history patterns **5**:228–231, **5**:229*F*, **5**:230*F*
nutritional physiology **5**:223–226
occurrences and characteristics **5**:223, **5**:224*T*
population biomass **5**:227–228, **5**:227*F*
population regulation **5**:228–231, **5**:229*F*, **5**:230*F*
predation vulnerabilities **5**:226–227
vegetation transformations **5**:227–228, **5**:227*F*, **5**:233
oceanic food web **5**:49*F*
parasite–host coevolution
community composition **2**:533
resistance mechanisms **2**:144–145
specialist versus generalist herbivores **2**:145–146
tolerance mechanisms **2**:145
species interactions **6**:718–719
trophic interactions **7**:264, **7**:265*F*
trophic pyramid **4**:670–671
tundra ecosystems **6**:330–331
herbs
food webs **7**:266*F*
tropical rain forests **6**:315*T*, **6**:316*T*, **6**:318*T*
Hericium erinaceus **3**:637*T*
Heritage Rivers Act (Australia) **6**:497
heritage sites **4**:435–436
Heritiera **1**:263, **5**:11*T*
Heritiera littoralis **1**:277
hermaphroditism **5**:659–660
hermatypic corals **2**:332*T*
hermit crabs **5**:16, **5**:316
Hernandia guianensis **1**:161–162
Heronimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Heronimus mollis **7**:450–451*F*
herons **5**:476–477, **6**:479–480, **7**:390
herpes simplex virus **7**:302–303
Herpestes auropunctatus **1**:77, **4**:359
Herpestes edwardsii **4**:365–366
Herpestes ichneumon **5**:475, **5**:477
Herpestidae **1**:686, **1**:687*F*
Herpesvirales **5**:130*F*, **5**:134*T*, **5**:135
Herpesviridae **5**:134*T*, **5**:135
herring **5**:315, **5**:464–465, **5**:464*F*, **6**:422*T*
Hesiod **4**:650
Hesionidae **7**:417*F*
Hesperia comma **5**:370
Hesperiidae **5**:387–388*T*
Hesperornis **2**:519*T*
Hess, Herman **1**:583
Heteralocha acutirostris **1**:604–605
Heterandria
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*

Heterandria (*continued*)
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
Heterangium **5**:618*F*
Heteraxinoides **7**:448–449*F*
Heterelmis comalensis **7**:58
Heterobathmiina **5**:385, **5**:386*F*, **5**:387–388*T*
Heterobathmioidea **5**:386*F*, **5**:387–388*T*
Heterocapsa **6**:294–295
Heterocapsa circularisquama **1**:477*T*, **1**:481–482, **5**:138
Heterocephalus glaber **3**:188
heterocyclic nitrogens **5**:696
Heterohyrax brucei **1**:50, **4**:695
Heterokontophyta **1**:475, **1**:476*F*, **6**:10*T*, **6**:25
Heteromita **6**:294
heteromitids **6**:294
Heteromyias albispecularis **2**:75*F*
Heteromyidae **6**:309*T*
Heteronetta atricapilla **5**:502
Heteroneura **5**:386, **5**:386*F*, **5**:387–388*T*
Heteropogon **1**:69–70
Heteropsis **3**:522*T*
Heteroptera
characteristics
coloration and ornamentation **7**:291
feeding behaviors **7**:288
habitat **7**:288
host specificity **7**:290
mimicry **7**:291
morphology **7**:288–292
parasitization **7**:292
physical characteristics **7**:288
predatory behaviors **7**:289
reproduction and growth **7**:289–290
scent glands **7**:288
economic importance **7**:292–293
freshwater ecosystems **4**:374
infraorders
Aradomorpha **7**:294
characteristics **7**:293–294
Cimicomorpha **7**:293
Gerromorpha **7**:293
Leptopodomorpha **7**:293
Nepomorpha **7**:293
Pentatomomorpha **7**:294
number of species **4**:371*T*, **4**:392*T*, **7**:287
research summary **7**:294–295
terminology **7**:287–288
Heterorhabditis **7**:430–431
Heterorhabditis bacteriophora **5**:673
Heteroscorpionidae **1**:209
Heterosigma **1**:475
Heterosigma akashiwo **1**:477, **1**:477*T*, **1**:480, **5**:137–138
heterosis **1**:639
Heterotermes **4**:409*T*
Heterotheca latifolia **6**:96–97
heterotrophic flagellates **6**:293–295, **6**:294*F*
heterotrophs
marine ecosystems **5**:48*F*
temperature tolerance ranges **4**:113–115
trophic interactions **3**:6–7
trophic pyramid **4**:670–671
Heterozostera **6**:542*T*
Heterozostera tasmanica **6**:541*T*
Hevea brasiliensis **1**:163, **5**:605, **6**:320
Hevea spruceana **1**:163
hexacorals **2**:335
Hexactinellida **3**:685*F*
Hexamita **6**:293, **6**:294*F*
Hexapoda
characteristics **1**:203
phylogenetic relationships **1**:203*F*
Hexathelidae **1**:206
hexazinon **5**:696
hexoses **6**:113*T*
hibernation **7**:36–37
Hibiscadelphus spp. **3**:205, **6**:81
Hibiscadelphus woodii **6**:81
hickory **7**:118
Hicksian demand function **7**:321–322
Hieracium **6**:125
Hieracium lepidulum **4**:349
Hieracium pilosella **4**:348–349, **4**:349*F*
Hieracium praealtum **4**:349
Hieraeetus fasciatus **5**:474
high-elevation ecosystems
Andes **4:96–110**
background information **4**:96
geographic distribution **4**:96–97, **4**:97*F*
habitat characteristics **4**:96–97
number of species **4**:102*T*
páramo ecosystem
characteristics **4**:101–103, **7**:275–276
geographic distribution **4**:96–97, **4**:97*F*
northern Andean alpine cushion formations **4**:103
northern Andean dwarf *Polylepis* forest **4**:103
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98
plant diversity **4**:98–99, **4**:99*T*
representative species **4**:102*F*
subpáramo zone **4**:102–103
superpáramo zone **4**:103
true páramo zone **4**:103
physical characteristics and patterns **4**:97–98
plant diversity
endemism **4**:99–100, **4**:99*T*
origins and diversification **4**:100–101
species richness **4**:98–99
puna
central Andean dwarf *Polylepis* forest **4**:105
cushion bogs **4**:105
desert puna **4**:104–105
dry puna **4**:104
geographic distribution **4**:96–97, **4**:97*F*
humid puna **4**:104
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:103–104
plant diversity **4**:98–99
representative species **4**:104*F*
salar vegetation **4**:105
research summary **4**:107
southern Andean steppes
altitudinal zonation **4**:106–107, **4**:106*F*
bogs and seeps **4**:107
cushion plants **4**:106–107, **4**:106*F*
geographic distribution **4**:96–97, **4**:97*F*
latitudinal gradients **4**:106–107
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:105–107
plant diversity **4**:98–99
representative species **4**:105*F*
Central America **2**:30
characteristics **7**:275
humid forests **7**:275
páramo ecosystem
characteristics **4**:101–103, **7**:275–276
geographic distribution **4**:96–97, **4**:97*F*
physical characteristics and patterns **4**:97–98
plant diversity **4**:98–99, **4**:99*T*
perhumid forests **7**:275
subhumid forests **7**:275
Higher Hamamelididae **5**:621–625
high-input high-diversity (HIHD) bioenergy cropping systems **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
high-input low-diversity (HILD) bioenergy cropping systems **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
highland gray slender loris **6**:256–258*T*
highland purple-faced langur **6**:256–258*T*
highly active antiretroviral therapy (HAART) **1**:594
highly replicated laboratory experiments **4**:587, **4**:587*F*
high-pressure liquid chromatography **5**:708
high-temperature ecosystems **4:111–120**
biodiversity **4**:118–119
characteristics and definitions **4**:111–112
extreme environments **4**:111
geothermally heated systems **4**:112–113
research summary **4**:119–120
temperature tolerance ranges **4**:113–115
thermophilic evolution **4**:115–118
high-throughput sequencing **6**:350–351
hiking
conservation support **7**:251–252
habitat degradation **4**:566–567
per capita contributions **7**:254*F*
popularity **7**:246–247, **7**:246*F*, **7**:251
Hill numbers **5**:203, **5**:204*F*
Hill–Robertson interference **6**:196–197
hill-zone Toque macaque **6**:256–258*T*
Himalayan highlands **1**:262–263*T*
Himalayan tahr **2**:551–552*T*
Himalayas
nitrogen deposition **5**:526*T*
plant diversity **6**:82*T*
Himatione sanguinea **1**:29*F*
Hinduism
biodiversity conservation **6**:564
environmental influences **6**:383–384
Hindu declaration on nature **3**:273, **6**:383

Hindu Kush highlands **1**:262–263*T*
hinterlands **7**:309–310
Hippidion **4**:727*T*
Hippoboscidae **3**:489–491*T*
Hippoboscoidea **3**:489–491*T*
Hippocamelus antisensis **2**:587*F*
Hippocamelus bisculus **3**:193
Hippocratea **5**:489
Hippocrates **5**:705
Hippoglossoides hippoglossoides **4**:631
Hippoglossus hippoglossus **4**:631
Hippopotamidae **4**:696–697, **5**:225*T*
hippopotamus **4**:690*T*, **4**:696, **5**:225*T*, **6**:479–480
Hippopotamus amphibius **1**:53–54, **4**:696–697, **4**:729–730, **4**:729*F*, **5**:225*T*, **6**:466, **6**:479–480
Hippopotamus behemoth **5**:462
Hippopotamus gorgops **5**:225*T*
Hippopotamus lemerlei **4**:657, **5**:365
Hippotragus **4**:696
Hippotragus equinus **1**:50
Hippotragus niger **1**:51
Hirnantia **5**:167
Hirtitermes **4**:409*T*
Hirudinea **4**:371*T*, **5**:669–670*T*, **7**:408
Hirudo medicinalis **1**:359
Hirundo pyrrhonota **6**:571–573
Hirundo rustica **5**:312, **5**:474
Hispaniola
  endemic plant species **6**:83
  lizards **1**:31*F*, **1**:32
  solenodons **4**:713–714, **4**:713*F*, **4**:714*F*
hispid cotton rat **1**:365–368
Histeridae **1**:327*T*
Histeroidea **1**:327*T*
Histoplasma capsulatum **3**:634*T*
histoplasmosis **3**:634*T*
historical biogeography **7**:342–343
historical botanical gardens **6**:505
historical conservation movement **2:278–288**
  background information
    general discussion **2**:278–279
    historic precedents **2**:279–280
    prehistoric precedents **2**:278–279
  emerging themes **2**:287–288
  modern origins
    foundations **2**:280–281
    Progressive Era **2**:282–283
    reform direction and momentum **2**:281–282
  twentieth century conservation
    challenges **2**:283
    conservation program reintegration **2**:286–287
    institutionalization
      forestry management **2**:283
      policies and administration **2**:283
      range management **2**:283
      recreational activities **2**:283
      soil conservation **2**:283–284
      wildlife management **2**:283
    1930s and 1940s **2**:284–285
    post-World War II era **2**:285–286
historical deforestation **2**:447–448
historical extinctions
  extinction paradoxes **5**:367
  Hawaii **5**:367
  New Zealand **5**:366–367
  number of species **5**:368*T*
historical geobotany **4**:396*T*
historic land-use patterns **4:569–582**
  anthropogenic impacts **4**:569–570
  buildings/built-up areas **4**:579, **4**:580*F*
  climate change effects **4**:569–570
  conservation implications **4**:581–582
  coppice woods **4**:573, **4**:573*F*, **4**:574*F*
  crop plants **4**:570, **4**:570*F*
  domestic animals **4**:570
  fields **4**:570–572, **4**:571*F*
  fire cycles **4**:580, **4**:581*F*
  forests and woodlands **4**:572–574, **4**:573*F*, **4**:574*F*
  grasslands **4**:577
  heathland **4**:577, **4**:578*F*
  hedges **4**:571, **4**:571*F*
  historical perspective **4**:569, **4**:579–581
  moorland **4**:577–579, **4**:578*F*
  permanent edges and open areas **4**:573–574, **4**:575*F*
  terraces **4**:572, **4**:572*F*
  wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*, **4**:581*F*
Histriophoca fasciata **5**:90
Hluhluwe-iMfolozi Park **5**:227*F*, **5**:228
Hoagland, Edward **4**:653
hoary bats **5**:313
hoatzin **1**:604, **1**:608
Hochstetter's frog **4**:194*T*
Hodotermes **4**:409*T*
Hodotermitidae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:422*F*
Hodotermopsis **4**:409*T*
Holmesina **4**:724–725*T*, **4**:727*T*
holm oaks **1**:612, **5**:218
Holocene
  climate change studies **2**:59–60, **2**:61*F*, **2**:82–83
  crustaceans **2**:407
  dispersal biogeography **2**:550, **2**:551–552*T*, **2**:552*F*
  vegetation–climate interactions **2**:682
holometabolous insects **5**:674–675
holoparasitism **4**:70–71
Holopedium **6**:26
Holothuroidea **4**:381*T*, **5**:669–670*T*
Holothyrida **1**:216
Holy *see* (Vatican City)
Homalanthus nutans **1**:594, **1**:596–597, **5**:713
Homallotermes **4**:409*T*
Homarus americanus **3**:610*F*
home gardens **1**:120, **4**:320, **6**:567, **6**:595
homing behavior **6**:526
Hominidae **4**:692–693, **4**:693, **6**:250–252, **6**:251*T*
Hominoidea **4**:692–693
Homo erectus **4**:153, **5**:231
Homogammina maculosa **6**:292
homologous recombination **1**:322–324, **1**:324*F*
homology
  cladistic analysis **2**:37
  parsimony analysis **6**:17–18
homoplasy **2**:35*F*, **2**:41–42, **6**:17–18
homoploid hybridization
  general discussion **4**:166–167
  hybrid speciation **4**:168
  hybrid swarms **4**:167–168
  introgression **4**:166–167
  novelty generation **4**:167
  novelty loss **4**:167
Homoptera **4**:392*T*, **7**:287
Homopus signatus signatus **1**:487–488
Homo sapiens
  ecosystem impacts **4**:153
  evolutionary systematics **2**:51–52, **2**:52*F*
  late Pleistocene extinctions **5**:231
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  molecular level diversity **2**:636*F*
  mutation rates **2**:628*T*
  neutral nucleotide polymorphism **2**:631*T*
  recombination characteristics **6**:361*T*
  see also humans
Homotherium **4**:724–725*T*
Honduras
  botanical gardens **6**:519
  deforestation **2**:30–31, **2**:31*F*, **2**:449*T*
  ecosystem services **7**:356*T*
  neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
  species richness **2**:25*T*
  water funds **7**:356*T*
  wheat production **1**:87–88
honeybees
  biodiversity–ecosystem services relationship **1**:348, **3**:347
  honey production **4**:183
  introduced parasites–species decline relationships **2**:532
  introduced species **6**:144–145
  kin selection theory **2**:322
  reproductive behavior **4**:178–179
  see also Hymenoptera
honeyeaters **5**:13–14
honeyguides **5**:502, **5**:502*T*
Hong Kong
  animal species **1**:261*T*
  mangroves **1**:275*T*
  plant species **1**:259*T*
  threatened species **1**:278*T*
  urbanization **4**:560*F*
hooded grebe **7**:186
hooded seals **5**:90, **5**:314, **7**:38
hook-and-line fishing **6**:420
Hooker, Joseph **2**:421, **2**:424, **2**:539–540, **6**:612–613, **7**:343
Hooker's sea lions **5**:81*T*
hookworms **1**:359, **5**:660*T*, **5**:671*T*
Hoolock hoolock **6**:256–258*T*
Hoover, Edgar **6**:205–206
Hopea **1**:263
Hopea ferrea **1**:267
Hoplocarida **2**:401*T*
Hoplognathotermes **4**:409*T*
Hoplophorus **4**:727*T*

Hoplopterus himantopus **5**:476–477
Hoplopterus spinosus **5**:476–477
Hoplostethus atlanticus **3**:468, **3**:482
hoppers **4**:298–299, **4**:392*T*
hops **3**:129–130
Hordeum sativum **2**:383
Hordeum spontaneum **3**:670, **3**:674*F*, **5**:483, **6**:361*T*
Hordeum vulgare **2**:662–663
hormone mimics **2**:426*T*, **5**:694
hormones **6**:113*T*
Hormosira banksii **4**:335
hornbills **1**:617
Horn index **2**:445, **5**:208–209
Horn of Africa **6**:82*T*
horn snails **3**:202*T*
hornworts
  North America **5**:564–565
  number of species **6**:58*T*
horseflies **3**:489–491*T*, **3**:492
horsehair worms **4**:371*T*
horse mackerel **5**:689–690
horses
  adipose tissue **7**:28, **7**:29*F*
  characteristics **4**:695
  deoxyribonucleic acid (DNA) evolution **2**:56*F*
  hybridization **3**:191
  introduced species
    Australia **1**:80
    Brazil **1**:83
    South Africa **1**:81
    United Kingdom **1**:79
    United States **1**:76–77
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  migratory patterns **5**:313
  morphological evolution **2**:54*F*, **2**:55
  risk status classification **1**:644*F*
  threat percentage **3**:195
horseshoe crab **2**:400
horsetails **3**:571–572, **6**:59
Horsfield's tarsier **6**:256–258*T*
horticultural gardens **6**:505, **6**:508
horticultural practices **4**:140–141, **4**:154
Horton Plains red slender loris **6**:256–258*T*
Horvitz–Thompson estimator **5**:204
Hospitalitermes **4**:409*T*
Hostaceae **1**:260*T*
host–disease interactions
  host genetic diversity **2**:533–534
  theoretical perspectives **3**:3
hot deserts **6**:334
hotspots **4:127–136**
  basic concepts **4**:127
  biodiversity indicators **3**:157*F*, **4**:261, **4**:262*F*
  complementarity **2**:193, **2**:193*F*, **4**:132
  conservation planning
    conservation impact **4**:133
    systematic conservation planning **4**:132–133
  criticisms
    biodiversity hotspots *sensu lato*
      inefficiencies **4**:134
      sampling bias **4**:134
    biodiversity hotspots *sensu* Myers/Conservation International (CI)
      competing consensual proposals **4**:134
      duplication of effort **4**:134
      scale **4**:134
      triage approach **4**:134
    lack of congruence **4**:134–135
    lack of overlap **4**:135
  hotspot types
    biodiversity hotspots *sensu lato* **4**:130
    biodiversity hotspots *sensu* Myers/Conservation International (CI) **4**:127–130, **4**:128*F*, **4**:129*F*
    general characteristics **4**:127–130
  mammals **4**:702
  Myriapoda/myriapods **5**:429
  neotropical mammals **2**:591–592
  nitrogen deposition **5**:526*T*
  plant communities **6**:61, **6**:79, **6**:82*T*
  scientific literature reviews
    alternative concepts **4**:132
    ambiguous usage **4**:132
    biodiversity hotspots *sensu lato* **4**:131–132
    biodiversity hotspots *sensu* Myers/Conservation International (CI) **4**:130–131, **4**:131*F*
    complementarity hotspots **4**:132
    usage **4**:130–131
  summary discussion **4**:135
hot springs **7**:327–328
Hough, Franklin B. **2**:281
house crows **1**:624
house finches
  Holocene postglacial spread **2**:551–552*T*
  introduced parasites–species decline relationships **2**:528*T*
houseflies **3**:489–491*T*, **3**:492, **4**:301
Houserock chisel-toothed kangaroo rat **7**:229*T*
house sparrows
  diversity loss **1**:623–624
  Holocene postglacial spread **2**:551–552*T*
  introduced species
    Brazil **1**:83
    India **1**:82
    South Africa **1**:81
    time lag–evolution relationship **4**:365–366
    United States **1**:77–78
  Near East ecosystems **5**:476
  species abundance **1**:601–602
houting **6**:479
hover flies **3**:489–491*T*, **3**:492
howler monkeys **2**:590*T*, **6**:251*T*, **6**:253–254
Hox genes **3**:395, **3**:396*F*
Hoyle, Fred **5**:630
Huarpea **4**:101
Hubbard Brook watershed experimental project **3**:93
Hubbardiidae **1**:212, **1**:212*F*
Hubbell, Stephen **5**:510
Hubbell, Sue **4**:653
Hubbell's unified theory **1**:403–404
Hubei golden snub-nosed monkey **6**:256–258*T*
Hucho taimen **3**:464–465
Hudsonia montana **3**:214
Hudson–Kreitman–Aguadé (HKA) test **2**:634–635
huemel **3**:193
Huia **1**:604–605
hukum **6**:565
Hula Lake **5**:472–473, **5**:474*F*, **5**:475*T*
human carrying capacity **1**:700, **2**:704–705
human-dominated ecosystems
  basic concepts **2**:350–351
  human-dominated landscapes **2**:349*F*, **2**:350–351, **2**:350*F*
  human-dominated seascapes **2**:351
  human-dominated techno-ecosystem **3**:60–63, **3**:62*F*
  human impacts **2**:350
  trait-based studies **2**:356–357
  untouched ecosystems **2**:350
human energy use **3:250–266**
  biodiversity impacts
    carbon dioxide (CO_2) emissions **3**:265*T*
    industrial energy
      coal **3**:257–259, **3**:264*T*, **3**:265*T*
      hydroelectric power **3**:259–260, **3**:264*T*, **3**:265*T*
      natural gas **3**:256–257, **3**:264*T*, **3**:265*T*
      nuclear energy **3**:259, **3**:264*T*, **3**:265*T*
      oil **3**:253–256, **3**:255*T*, **3**:264*T*, **3**:265*T*
    land requirements **3**:265*T*
    major sources **3**:264*T*
    renewable energy **3**:260–262, **3**:264*T*, **3**:265*T*
    traditional energy **3**:262–263, **3**:263*T*, **3**:264*T*, **3**:265*T*
    world energy reserves **3**:265–266, **3**:265*T*
  categories
    industrial energy **3**:250–253
    renewable energy **3**:251
    traditional energy **3**:250
  energy efficiency measures **3**:251, **3**:254*F*
  environmental impact **3**:263–266
  future trends **3**:263–266
  thermodynamic laws **3**:253, **3**:254*T*
  world energy reserves **3**:265–266, **3**:265*T*
  world energy use **3**:251, **3**:251*F*, **3**:252*T*, **3**:253*T*
human flourishing
  artistic expression **3**:273, **3**:274*F*
  general discussion **3**:273–274
  historical understanding **3**:274
  life-support systems **3**:274
  recreational activities **3**:273, **3**:273*F*
  religious inspiration **3**:274
  scientific knowledge **3**:273–274
human footprint measures and maps **2**:221–222
Human Genome Project (HGP) **3**:662
human health–biodiversity relationship **1:357–372**
  biomedical research

bears as models **1**:358–359
cone snails as models **1**:359
nonhuman models **1**:358–359
components
general discussion **1**:357–358
human health impacts **1**:358
species diversity measures **1**:357–358
infectious disease incidence
disease management **1**:360–361, **1**:361*F*
emerging infectious diseases (EIDs) **1**:368–370, **1**:369*F*
global patterns **1**:368–370, **1**:369*F*
hantavirus pulmonary syndrome **1**:365–368, **1**:366*F*
Lyme disease **1**:361–362, **1**:363*F*
schistosomiasis **1**:368, **1**:368*F*
West Nile virus **1**:362–365, **1**:364*F*, **1**:365*F*
zooprophylaxis **1**:360–361
nutritional factors **1**:359–360
preventative/therapeutic agents **1**:359
summary discussion **1**:370
human immunodeficiency virus (HIV) **1**:594, **7**:302–303
human impact
biodiversity **4:137–152**
basic principles
ambivalence **4**:139
environmental diversity **4**:137–138
human diversity **4**:137–138
impact types **4**:138
sociocultural factors **4**:138*T*
species and population levels **4**:138–139
biodiversity commodification
cities **4**:142
colonialism **4**:142–143
commercial farming industry **4**:144
human body **4**:142
international commerce **4**:143–144
modern transportation **4**:143–144
biodiversity conversion
agriculture **4**:141, **4**:154
population growth effects **4**:141–142
raised bed farming **4**:141
rice paddies **4**:141
Vavilov centers **4**:141
biodiversity modifications
extinction hypothesis **4**:139–140
fire/fire regimes **4**:140
horticultural practices **4**:140–141, **4**:154
biodiversity toxification
Anthropocene epoch **4**:146
industrialization **4**:145
militarization **4**:145–146
oil wells and pipelines **4**:145
climate change **4**:24
conservation efforts **4**:148–149
future outlook
conservation versus extermination **4**:147–148
diversity principle **4**:148
governmental limitations **4**:147
indigenous potential **4**:147
protected area systems **4**:147
uncertainty studies **4**:146–147
general discussion **4**:137
humanism **3**:276–277
human rights
challenges **3**:314–315
less-developed countries **3**:315–316, **3**:316*F*
logging industry–owl population controversy **3**:314–315, **3**:315*F*
property rights **3**:315
humans
biodiversity–human health relationship **1:357–372**
biomedical research
bears as models **1**:358–359
cone snails as models **1**:359
nonhuman models **1**:358–359
components
general discussion **1**:357–358
human health impacts **1**:358
species diversity measures **1**:357–358
infectious disease incidence
disease management **1**:360–361, **1**:361*F*
emerging infectious diseases (EIDs) **1**:368–370, **1**:369*F*
global patterns **1**:368–370, **1**:369*F*
hantavirus pulmonary syndrome **1**:365–368, **1**:366*F*
Lyme disease **1**:361–362, **1**:363*F*
schistosomiasis **1**:368, **1**:368*F*
West Nile virus **1**:362–365, **1**:364*F*, **1**:365*F*
zooprophylaxis **1**:360–361
nutritional factors **1**:359–360
preventative/therapeutic agents **1**:359
summary discussion **1**:370
biodiversity impacts **7**:203*F*
desert ecosystems **2**:475–477, **2**:488*F*, **2**:489–491, **2**:490*T*
disturbance ecology **2**:566–567
ecosystem impacts **4:153–161**
alien species invasion
continental ecosystems **4**:155
general discussion **4**:155
island biogeography **4**:155
onshore ocean waters **4**:155–156
pathogens **4**:156
current impacts
alien species invasion **4**:155
general discussion **4**:153–154
habitat alteration **4**:153–154
driving forces
environmentally unsound technologies **4**:159
faulty economic factors **4**:159
faulty political arrangements **4**:160
faulty social arrangements **4**:159–160
IPAT equation **4**:158–159
market failure **4**:159
overconsumption **4**:159
population growth/overpopulation **4**:159
ecosystem deterioration measures
epidemiological environment **4**:158
faltering food production **4**:157–158
fisheries decline **4**:158
net primary production (NPP) **4**:156–158
observable symptoms **4**:158
water quality decline **4**:158
habitat alteration
climate change effects **4**:155
conversion to farms **4**:154
conversion to grazing **4**:154
conversion to infrastructure **4**:154
dam projects **3**:573–574, **4**:154
deforestation **4**:153–154
fire suppression **4**:140, **4**:154
forest fragmentation **4**:154
general discussion **4**:153–154
mining impacts **4**:154
ocean toxification **4**:155
onshore ocean siltation **4**:155
ozone depletion **4**:155
recreational activities **4**:155
selective harvesting **4**:154
wetland drainage **4**:154–155
historical perspective **4**:153
overexploitation
aquarium fish **4**:156
esthetic resources **4**:156
food harvests **4**:156
general discussion **4**:156
natural products **4**:156
scientific and educational use **4**:156
solution strategies
proximate steps **4**:160
ultimate steps **4**:160–161
threatened species **7**:203*F*
ultimate solutions
consumption control **4**:161
population reduction **4**:160–161
socioeconomic and political system revisions **4**:161
technologically benign substitutions **4**:161
environmental impacts **7**:203*F*
fungi-related diseases
classifications **3**:633–634
mycetismus **3**:634, **3**:635*T*
mycosis **3**:633–634, **3**:634*T*
mycotoxicosis **3**:634–635, **3**:635
global diversity **6**:251*T*
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
human–crop plant coevolution **2**:657, **4**:570, **4**:570*F*
human-generated ecotones **3**:143*F*, **3**:146–147
human health issues
botanical gardens **6**:518–519
natural history museum research programs **5**:415
human–microbe relationships **5**:307–308
human well-being **1:435–439**
basic concepts **1**:345
biodiversity loss **1**:435–436
commodity production **1**:436
direct and indirect valuation methods **1**:435–436
ecosystem services **1**:436

humans (*continued*)
esthetic, spiritual, cultural, and recreational benefits **1**:437
human well-being indicators **7**:77–80, **7**:78*F*
insurance benefits **1**:436
knowledge benefits **1**:437
market factors
internalized benefits **1**:438–439
internalized values **1**:438
market efficiency **1**:437
public goods **1**:437–438
option value **1**:436–437
risk factors **1**:436
uncertainty factors **1**:436–437
human–wildlife conflicts **7**:406
infectious disease regulation **1**:348–350, **3**:349
land use issues **4**:558–560
mollusc–human interactions **5**:381–382
nitrogen cycle impacts **5**:520–522, **5**:522*F*, **5**:538–540, **5**:540*F*, **5**:541*F*
parasitic disease vectors **5**:670–672, **5**:671*T*
parasitic nematodes **7**:431–432
plant defense system uses
background information **2**:431–432
food spices **2**:432–433
hunting and fishing **2**:431–432
medicines **2**:432
plant–human interactions **6**:57–58
population stabilization **6:199–209**
cause–effect relationships
economic development **6**:204–206
education levels **6**:203–204, **6**:205*F*
fertility and mortality patterns **6**:202–204
future outlook **6**:208
government regulations and policies **6**:207
historical perspective
pre-1900 population changes **6**:199–200, **6**:200*F*
1900–1950 population change **6**:200
1950–2010 population change **6**:200–201
age profile changes **6**:202, **6**:204*F*
contraceptive methods **6**:202*F*
demographic transition **6**:199–200
family planning methods **6**:202, **6**:202*F*
fertility and mortality patterns **6**:201, **6**:203*F*
growth rates **6**:201–202
International Conference on Population and Development (ICPD) **6**:207–208
population growth responses **6**:206–207
2010–2050 population projections **6**:206
ultraviolet radiation effects
characteristics **7**:301–302
eyes/photokeratitis **7**:302
immune system **7**:302–303
photoaging **7**:302
skin cancer **7**:302
sunburn **7**:302
human waste disposal **1**:512, **1**:512*F*
Humboldt, Alexander von **1**:577–578, **4**:651–652
humid forests
high mountain ecosystems **7**:275
lowland ecosystems **7**:273
low mountain ecosystems **7**:275
humid monsoon region
characteristics **1**:262–263*T*, **1**:270–271
laurel forests **1**:270–271
temperate deciduous forests
characteristics **1**:271
China **1**:271–272
Japan **1**:271
Korea **1**:271
Manchuria **1**:271–272
mountain regions **1**:272–273
pine forests **1**:272
humid puna **4**:104
humid tropical forest **1**:260–263, **1**:262–263*T*, **1**:266*T*
humid tropical mountain areas **1**:262–263*T*, **1**:266*T*, **1**:267
Humiria balsamifera **1**:160
hummingbirds
beak shape **1**:604*F*
evolutionary convergence **1**:617
mangrove ecosystems **5**:13–14
plant–pollinator interactions **6**:47–48
hummock grasslands **1**:292–297, **1**:294*T*
humpbacked flies **3**:489–491*T*, **3**:492
humpback salmon **6**:523
humpback whales **2**:7–8, **5**:77*T*, **5**:81*T*, **5**:87, **5**:87*T*
humped-back model (HBM) **4**:79
hump-winged crickets **3**:727–728
Humulus lupulus **3**:129–130
humus epiphytes **3**:523
Hungary
botanical gardens **6**:507*T*
vegetation surveys **3**:341–344*T*
hunter–gatherer societies
characteristics **4**:162–163
ecological impact **4:162–165**
background information **4**:162
biodiversity **4**:164
community-based conservation **4**:164–165
economic development **4**:164–165
land tenure **4**:164
modernization processes **4**:163–164
natural resource exploitation **4**:162–163
resource management **4**:163
social–ecological connection **4**:164–165
resource management **4**:163, **6**:566–568
tropical forest regeneration **7**:277–278
hunting
biodiversity loss **1**:405, **7**:203*F*
bird extinctions **1**:621
habitat degradation **4**:566–567
hunter–gatherer societies **6**:566–568
mammal conservation **3**:191–192, **3**:197, **4**:709–710
Near East ecosystems **5**:472
neotropical mammals **2**:589–592, **2**:590*T*
plant defense system uses **2**:431–432
popularity declines **7**:246–247, **7**:246*F*
primate threats **6**:253–254
private conservation areas **2**:296–297, **2**:297*T*
South America **6**:609
species reductions or absences **6**:444–445
threatened birds **7**:185
Hura crepitans **1**:163
Hurlburt rarefaction **5**:49
Huron (Lake) **4**:461*T*
hurricanes **6**:62–63, **7**:277–278, **7**:399–400
Huso huso **3**:462, **3**:462*F*
Hutchinson, G. E. **2**:648, **3**:239–240, **3**:609
Hutchinsonian niche **4**:33–34, **4**:42–45
Hutchinson's paradox **6**:26–27
hutias
background and characteristics **4**:713–714, **4**:713*F*, **4**:714*F*, **4**:715*F*
conservation efforts **4**:715
general discussion **4**:713–714
major threats **4**:714–715
reintroduction programs **3**:197
Huxley, Julian **6**:3–4, **7**:101–102
Huxley, Thomas **2**:421, **3**:483, **5**:630
Hwange National Park **5**:234–236
Hyacinthaceae **1**:486*T*
hyacinth macaw **7**:187
Hyaena hyaena **5**:472
Hyaenidae **1**:686, **1**:687*F*
food categories **1**:690*F*, **1**:691*F*
Hyalophora cecropia **5**:395–396
Hyalorrhipis calcarata **5**:469
Hyancinthella **1**:269
Hyancinthoides non-scripta **6**:106*F*, **6**:107
Hyblaeoidea **5**:387–388*T*
hybridization
biodiversity restoration **6**:454
plants **4:166–176**
basic concepts **4**:166
ecological consequences
biological diversity **4**:171–173
community impacts **4**:173–174, **4**:173*F*
resistance abilities **4**:172–173
species richness **4**:173–174, **4**:173*F*
homoploid hybridization
general discussion **4**:166–167
hybrid speciation **4**:168
hybrid swarms **4**:167–168
introgression **4**:166–167
novelty generation **4**:167
novelty loss **4**:167
polyploid hybridization
ancient polyploidy **4**:168–169
angiosperms **4**:169
characteristics **4**:168
classifications **4**:168
diversification **4**:170
polyploid complexes **4**:169–170
recent polyploidy events **4**:170–171, **4**:171*F*, **4**:172*F*
tracheophytes **4**:169
research summary **4**:174
speciation **6**:635
hybrid poplars **1**:544*F*
hybrid vigor **1**:639

Hydnophytum formicarium **5**:13
Hydnoraceae **4**:72*F*, **4**:73
Hydnum repandum **3**:214–215
Hydra **4**:371*T*, **4**:374*F*, **4**:644
Hydrachnidia **4**:371*T*
Hydrangea **3**:522*T*
Hydrangea peruviana **1**:167
hydric soils **7**:378–379
Hydrilla verticillata **3**:361
hydrobiid snails **3**:186, **3**:187*F*
Hydrocharitaceae **2**:374*T*, **6**:542*T*
Hydrocharitoideae **6**:542*T*
Hydrocherus hydrochaeris **4**:697
Hydrochoeridae **4**:727*T*
Hydrochoerus **4**:724–725*T*
Hydrochoerus hydrochaeris **2**:587*F*, **6**:332–333
Hydrodamalis **4**:724–725*T*
Hydrodamalis gigas **3**:283, **4**:694, **5**:76, **5**:365
Hydrodictyon **3**:358
hydroelectric power
  biodiversity impacts **3**:259–260, **3**:264*T*
  carbon dioxide (CO_2) emissions **3**:265*T*
  ecosystem services modeling **5**:351
  land requirements **3**:265*T*
  world energy reserves **3**:265*T*
  world energy use **3**:251*F*, **3**:252*T*, **3**:253*T*
  see also dams
hydrofluorocarbons (HFCs) **1**:284–286, **4**:19*T*
hydrogen (H)
  atmospheric concentrations **1**:282*T*
  elemental partitioning **1**:554*T*
  hydrogen sulfide (H_2S) **1**:561–563
  hydrothermal vent ecosystems **7**:328
  relative abundance **1**:553*F*
hydrogen peroxide (H_2O_2) **1**:281–282, **1**:282*T*
hydrogen sulfide (H_2S) **1**:281–282, **1**:282*T*, **1**:287, **1**:561–563, **6**:154–155, **6**:154*F*
hydroids **4**:384*T*
Hydrolagus colliei **3**:460*F*
hydrolases **5**:269*T*
hydrological cycle **1**:553–554, **1**:555*F*, **4**:360, **7**:371
hydrolysis **1**:555–556
Hydromyles globulosa **4**:386*T*
hydropetric communities **7**:58, **7**:59*T*
Hydrophilidae **1**:327*T*
Hydrophiloidea **1**:327*T*
hydropower **6**:528–529, **6**:609
  see also dams
Hydroptila adana **5**:463*T*
Hydroptila fonsorontina **5**:463*T*
Hydroptila hirra **5**:463*T*
Hydroptila libanica **5**:463*T*
Hydroptila palaestinae **5**:463*T*
Hydroschendyla submarina **5**:425
hydrosphere **3**:279
hydrothermal vents **7:327–332**
  Census of Marine Life project **2**:15*F*, **2**:19–20, **2**:20*F*
  chemosynthetic symbioses **5**:120–121, **5**:120*F*
  discovered species **2**:6*F*
  ecosystem diversity **5**:53, **5**:53
  Endeavor Hydrothermal Vents Marine Protected Area **5**:447
  energy flow **3**:242
  genetic diversity **5**:50
  Lake Baikal **4**:459
  marine invertebrates **4**:391
  marine sedimentary habitats **5**:108, **5**:108*T*
  occurrences and characteristics **7**:327–328
  see also hyperthermophiles
hydrothermal waters **7**:50
hydroxyl radicals (OH) **1**:281–282, **1**:282*T*
Hydrozoa **5**:669–670*T*, **6**:478*T*
Hydrurga leptonyx **5**:90
Hyemoschus aquaticus **1**:49
hyenas **2**:190, **2**:190*F*, **2**:190*T*, **4**:690*T*, **4**:693, **4**:701*T*, **5**:472
hygropetric zone **7**:49–50
Hyla arborea **5**:473
Hylaeogena **1**:329
Hyla pulchella **3**:115
Hyla regilla **3**:165
Hylis foveicollis **7**:220*T*
Hylis procerulus **7**:220*T*
Hylobates **4**:692–693
Hylobates agili agilis **6**:256–258*T*
Hylobates albibarbis **6**:256–258*T*
Hylobates carpenteri **6**:256–258*T*
Hylobates klossii **6**:256–258*T*
Hylobates lar lar **6**:256–258*T*
Hylobates lar vestitus **6**:256–258*T*
Hylobates lar yunnanensis **6**:250
Hylobates moloch **6**:256–258*T*
Hylobates muelleri abbotti **6**:256–258*T*
Hylobates muelleri muelleri **6**:256–258*T*
Hylobates muelleriri funereus **6**:256–258*T*
Hylobates pileatus **6**:256–258*T*
Hylobatidae **4**:692–693, **6**:250–252, **6**:251*T*
Hylochoerus meinertzhageni **1**:52, **4**:696
Hylocoetus dermestoides **7**:218–219
Hylocomium splendens **1**:232
Hymenaea courbaril **4**:286–287, **4**:288*T*, **4**:595
Hymenophysa curta **4**:465
Hymenoptera **4:177–184**
  biological control agents **4**:183
  characteristics **4**:177, **4**:392, **5**:395–396, **5**:674
  classifications **4**:177–178
  fossil record **4**:178
  honey production **4**:183
  major divisions and characteristics **4**:308–312, **4**:309*T*
  number of species **4**:392*T*
  oviposition behavior **4**:179
  parasitoids **4**:179–180
  pest species **4**:182–183
  phylogeny **4**:178
  phytophagy
    characteristics **4**:180–181
    fig wasps **4**:181
    gall makers **4**:181
    leaf-cutting ants **4**:181
    pollen feeders **4**:181
  pollination **4**:183, **6**:141*T*, **6**:142
  predatory behaviors **4**:180
  quantity estimates **1**:326–328
  reproductive behavior **4**:178–179
  social behavior **4**:181–182, **4**:414*T*
Hyolitha
  physical characteristics **5**:374–375, **5**:375*F*
  spatiotemporal distributions **5**:374*F*
Hyophorbe amaricaulis **6**:81
Hyophorbe lagenicaulis **6**:81
hyoscyamine **5**:712*T*
Hyoscyamus niger **5**:712*T*
Hyotitan **5**:224*T*
Hypacrosaurus **2**:519*T*
Hyparrhenia hirta **6**:93
Hyparrhenia rufa **7**:387
Hypera brunneipennis **4**:366
Hyperammina **6**:292
Hypera postica **1**:78
Hyperaspis pantherina **4**:361
hypereutrophic level **3**:563*T*
Hypericum **7**:275–276
Hypericum **5**:623*F*
Hypericum fasciculatum **7**:261–262
Hypericum perforatum **3**:704, **3**:704*F*
Hyperiodrilus africanus **6**:557
hypermetamorphosis **5**:391
Hyperolius horstocki **5**:220
Hyperoodon ampullatus **5**:77*T*
hyperparasitism **5**:657, **5**:675
hyperthermophiles
  characteristics **7**:161
  ecological niches **7**:161–163, **7**:162*F*
  hydrothermal vent ecosystems **7**:328
  metabolic diversity **7**:161–163
  molecular mechanisms **7**:163
  origins **7:161–166**
    early Earth environment **7**:163–164
    fossil record **7**:164
    general discussion **7**:163–164
    molecular fossils
      genetic annealing model **7**:165
      lateral gene transfer **7**:165
      molecular phylogenies **7**:162*F*, **7**:164–165
      small subunit ribosomal RNA (ss rRNA) **7**:162*F*, **7**:164–165
    origin of life theories **7**:165–166
  phylogenetic relationships **7**:161, **7**:162*F*
  research summary **7**:166
  thermophilic Archaea **1**:219–220, **1**:221*F*, **1**:221*T*
Hyphae
  functional role **3**:636
  growth processes **3**:626, **3**:626*F*
hyphochytriomycetes **6**:296–297
Hypochaeris sessiliflora **4**:103
Hypochilidae **1**:206–207
Hypochoeris maculata **3**:206
Hypochrysops **5**:13
Hypogeomys australis **4**:728–729
hypolimnion **3**:560, **4**:374*F*
Hypoprepia fucosa **5**:388*F*
hyporheic zone **7**:50
hypotelminorheic communities **7**:59
hypotension **3**:635*T*
Hypothemus hampei **4**:304–305
hypoxia **3**:351, **5**:543, **7**:372–373, **7**:374*F*
Hypsypops rubicundus **3**:474*F*

Hyptis parkeri **1**:164
Hyracodontidae **5**:224*T*
Hyracoidea **4**:685, **4**:690*T*, **4**:695, **4**:701*T*
Hyracotherium **5**:2
hyrax **4**:690*T*, **4**:695
Hystricomorpha **2**:584*F*, **2**:586
Hystrix indica **5**:463*T*

I

Iberian lynx **7**:262–263
Iberus spp. **3**:222–223
ibex **1**:640, **3**:197
ibis **6**:479–480
Ibn Es-Sûri, Rachid-eddin **3**:321
iboga **1**:594
ibotenic acid **3**:635*T*
ice **6**:607*T*
Ice Age **2**:59–60, **2**:82–83, **2**:466, **4**:228–229
ice crawlers **3**:221
icefish **3**:467–471
Iceland
  botanical gardens **6**:507*T*
  dispersal biogeography **2**:547
  energy efficiency measures **3**:254*F*
  vegetation surveys **3**:341–344*T*
  whaling industry **3**:454
Icelandic ptarmigan **3**:127–128
ice plants **4**:360
Icerya purchasi **1**:78, **4**:361
Ichneumonidae **4**:178, **5**:395–396, **5**:674
Ichneumonoidea **4**:178
Ichthyobodo **6**:293
Ichthyodinium chabelardi **6**:295
Ichthyophthirius **6**:295
ichthyotoxicity **1**:474–475, **1**:474*F*
Ichyrosmilus **4**:724–725*T*
Icosteoidei **3**:469
Icosteus aenigmaticus **3**:469
Ictalurus **3**:365*F*
Icteridae **5**:502, **5**:502*T*, **6**:308*T*
Icterus galbula **4**:174
idiobionts **5**:674–675
Idiomyia **1**:28–29
Idiospermum **5**:621*F*
igapó forest **1**:163–164
Iguana iguana **3**:217
iguanas **3**:217, **5**:217, **6**:402, **6**:411
Iiwis **1**:29, **1**:29*F*
Ilex **1**:270–271
Illacme plenipes **5**:424
illegal wildlife trade **1**:517–518, **1**:518*F*, **1**:519*F*
illegitimate recombination **6**:354–355
Illicium **5**:621*F*
Ilyyocypris harmanni **5**:470*T*
Ilyyocypris nitida **5**:470*T*
imazalil **5**:695
imidacloprid **5**:694
imidazoles **5**:695
imidazolinone **4**:90*T*
imminent extinction probability
  marine mammals
    beluga whales **5**:77*T*, **5**:79–80
    blue whales **5**:81*T*, **5**:82
    California sea otter **5**:81*T*, **5**:83–84
    general discussion **5**:76
    Mediterranean monk seals **5**:77*T*, **5**:80
    Northern sea otter **5**:81*T*, **5**:84–86
    right whales **5**:76–78, **5**:77*T*
    species and subspecies **5**:77*T*
    Steller sea lions **5**:81*T*, **5**:82–83
    vaquitas **5**:77*T*, **5**:80
    western North Pacific gray whales **5**:77*T*, **5**:78–79
    West Indian manatee **5**:81*T*, **5**:83
  protected area systems **6**:488*F*
Immoidea **5**:387–388*T*
immune system **7**:302–303
impala **4**:696, **6**:572–573
Impatiens **1**:451, **6**:316*T*
Impatiens capensis **2**:721–722
Imperata **1**:70
Imperata cylindrica **1**:70, **2**:379–381, **4**:358–359, **5**:605, **6**:556
Important Bird Areas (IBAs) **7**:203–204, **7**:204*F*
Important Plant Areas (IPAs) **7**:203–204
improved short fallow systems **6**:560–561
Inabanga Coast **1**:276*T*
inajá **4**:286–287
inarticulata **3**:685*F*
INBio Agreement **5**:714
inbreeding **4**:**245–252**
  basic concepts **4**:245
  biodiversity restoration **6**:454
  diversity loss **4**:249
  extinction rates
    captive populations **4**:249, **4**:249*F*
    wild populations **4**:249–250
  genetic diversity **2**:265–267, **2**:272–273
  genomic balance **6**:4
  inbreeding depression
    characteristics **4**:245–246
    evidentiary research **4**:245–246, **4**:246*F*
    juvenile mortality **4**:245–246, **4**:246*F*
    lethal equivalent measures **4**:247
    reduction strategies **4**:248–249
    severity measures **4**:246–247, **4**:247*F*
    small populations **4**:248
    susceptibility **4**:247–248, **4**:248*F*
    total reproductive fitness **4**:247
    wild populations **4**:248
  outbreeding/outbreeding depression
    causal factors **4**:250
    characteristics **4**:250
    evidentiary research **4**:250–251
    risk predictions **4**:251–252, **4**:251*F*
incense cedar **6**:152–154
Inceptisols **6**:123*T*
incidental omnivory **3**:503
inclusive fitness **6**:573–574
Incolitermes **4**:409*T*
Inconvenient Truth, An (2006) **4**:654
Incurvariina **5**:387–388*T*
index of biological integrity (IBI) **3**:292–294, **3**:292*T*, **3**:293*F*, **3**:293*T*
India
  animal species **1**:261*T*
  biodiversity richness **1**:501*F*
  botanical gardens **6**:519
  C_4 plants **2**:371*T*
  countryside biogeography studies **2**:354–356, **2**:356*F*
  deforestation **1**:502*T*, **1**:677*F*, **2**:447–448
  ecosystems
    coral reefs **1**:500*T*
    deserts **2**:468
    ecosystem function measurements **3**:82*T*
    mangroves **1**:275*T*, **1**:277
    monsoon forests **1**:263–266
    nutrient cycling **3**:82*T*
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    tropical dry forest **1**:266–267
    tropical moist forest **1**:277*T*
    tropical rain forests **6**:314*T*
    tropical/subtropical humid mountain areas **1**:267
    vegetation types and coverage percentage **1**:264–265*T*
  endangered mammals **3**:194*T*
  endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
  fish diversity **3**:473
  fishing/fisheries **6**:422*T*
  indigenous peoples **4**:270*T*, **4**:275*F*
  institutional collections and capacity **1**:505*T*
  introduced species
    nonindigenous invertebrates
      arthropods **1**:76*T*, **1**:82
      crop plant pathogens **1**:76*T*, **1**:82
      livestock pathogens **1**:76*T*, **1**:82
      molluscs **1**:82
    nonindigenous plants **1**:76*T*, **1**:82
    nonindigenous vertebrates
      birds **1**:82
      mammals **1**:76*T*, **1**:82
  invasive species
    aquatic ferns (*Salvinia molesta*) **4**:346–347
    Lantana camara **4**:352–353
    prickly pear (*Opuntia*) **4**:350–351
  maize production **1**:86
  National Biodiversity Authority (NBA) **1**:507
  native crops **2**:663
  nitrogen deposition **5**:526*T*
  plant species **1**:259*T*
  polyculture cropping systems **7**:231
  population regulation **6**:207
  primate endemism **6**:252*T*
  rice production **1**:86, **1**:88
  shifting cultivation systems **7**:232
  species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
  termites **4**:432*T*
  threatened birds **7**:181
  timber industry **6**:224–225
  tropical forests **7**:269–270, **7**:270*T*
  wheat production **1**:86
Indian black earth **4**:285–287, **4**:286*F*
Indian Botanic Garden **6**:506–508
Indian fig **4**:350–351
Indian meal moth **5**:401–402
Indian mongoose **4**:359

Indian myna **1**:81
Indian Ocean **4**:386*T*
Indian Ocean islands **4**:705*T*, **6**:506*T*, **6**:507*F*, **6**:510*T*
Indian Pipe **2**:643, **4**:74
Indian rhinoceros **4**:695
Indian river dolphins **5**:65*T*, **5**:77*T*
Indicatoridae **5**:502, **5**:502*T*
indicator species **4:253–263**
  biodiversity indicators
    characteristics **4**:259–260
    conservation planning **4**:262
    Global Biodiversity Outlook **2**:309, **2**:310*F*
    indicator reliability
      favorable versus hostile conditions **4**:261
      hot spots **4**:261, **4**:262*F*
      rare species versus endemic species **4**:261
      tropical versus temperate gradients **4**:261
    information gathering concerns **4**:260–261
    reserve selection algorithms **4**:262–263
    suggested taxa **4**:260
  Clementsian ecological succession paradigm **7**:64
  computational approaches **4:264–268**
    ecological applications
      Barro Colorado Island (BCI) **4**:267
      carabid beetles (Belgium) **4**:266–267
      forest ecosystems **4**:267
      oribatid mites **4**:267
    IndVal (indicator value) method
      basic concepts **4**:265
      expanded statistical methods **4**:265–266
      numerical example **4**:265, **4**:266*T*
    research summary **4**:267
    software **4**:267
    TWINSPAN (two-way indicator species analysis) method **4**:264–265
  environmental change indicators
    birds
      farmland birds **4**:259
      general discussion **4**:258–259
      migratory songbirds **4**:259
      peregrine falcons **4**:258–259
    climate change
      species distribution changes **4**:257–258, **4**:259*F*
      species sensitivity **4**:256–257, **4**:258*F*
    general discussion **4**:255
    historical records
      lake acidification **4**:255, **4**:256*F*
      plants–carbon dioxide relationship **4**:255–256, **4**:257*F*
    monitoring strategies **4**:255
  environmental examples
    acid deposition **4**:253–254
    amphibian decline **4**:254–255
    data interpretation and analysis **4**:254
    ecotoxicology **4**:255
    European river management **4**:254
    widespread use **4**:254
    general discussion **4**:253–254
    identification needs **4**:264
    research summary **4**:263
    species assemblages **6**:663–664, **6**:664*F*
    threatened species lists and assessments **7**:205–207, **7**:206–207*F*
    wildlife management **7**:405–406
indigenous knowledge **1**:596
indigenous peoples **4:269–278**
  background information **4**:269
  bio–cultural axiom **4**:277
  biodiversity-rich countries **1**:503
  biological diversity **4**:270–271, **4**:271*T*
  characteristics **4**:162–163, **4**:269–270
  conservation practices **2**:278–279
  cultural diversity **4**:270–271
  ecological impact **4:162–165**
    background information **4**:162
    biodiversity **4**:164
    community-based conservation **4**:164–165
    economic development **4**:164–165
    land tenure **4**:164
    modernization processes **4**:163–164
    natural resource exploitation **4**:162–163
    resource management **4**:163
    social–ecological connection **4**:164–165
  endemic languages **4**:270–271, **4**:270*T*
  estimated populations **4**:269*T*
  ethnoecology
    basic concepts **4**:272–273
    belief system (kosmos) **4**:273
    ecological knowledge (corpus) **4**:273–275, **4**:276*F*
    use strategies and practices (praxis) **4**:275–276
  geographic distribution
    Australia **4**:272*F*
    Brazil **4**:273*F*, **4**:273*T*
    geographic relationships **4**:269–270, **4**:272
    India **4**:275*F*
    Mexico **4**:274*F*
  natural resource appropriation **4**:271–272
  plant domestication **4**:281–285, **4**:282*F*, **4**:284*F*
  priority terrestrial ecoregions **4**:271, **4**:271*T*
  protected area systems **4**:276–277
  resource management **4**:163, **4**:272, **6**:384–385, **6**:384, **6**:566–568
  species richness **4**:271*T*
  traditional agriculture
    biodiversity-based agricultural strategies **1**:123–124
    multiple cropping systems **1**:119–120
    traditional knowledge systems
      environmental knowledge **1**:121
      experimental approaches **1**:122
      farming practices **1**:121
      folk taxonomies **1**:121
      research background **1**:120–121
  tropical forest regeneration **7**:277–278
indigo buntings **6**:702*F*
individual transferable quota (ITQ) systems **5**:150
Indo–Burma region **6**:82*T*
Indochinese gray langur **6**:256–258*T*
Indochinese rain forest **1**:262–263*T*
Indonesia
  animal species **1**:261*T*
  aquatic ferns (*Salvinia molesta*) **4**:346–347
  biodiversity richness **1**:501*F*
  Bogor Botanical Gardens **5**:448–449, **6**:508, **6**:510–511
  deforestation **1**:502*T*, **1**:509–510, **1**:677*F*, **2**:449*T*, **2**:451*T*, **4**:562–564, **4**:563*F*
  ecosystems
    coral reefs **1**:500*T*
    genetic richness **1**:514*T*
    mangroves **1**:275*T*, **1**:277, **1**:513*T*
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    tropical dry forest **1**:266–267
    tropical moist forest **1**:277*T*
    vegetation types and coverage percentage **1**:264–265*T*
  El Niño Southern Oscillation (ENSO) **3**:156–158, **3**:158*F*
  endangered mammals **3**:193–194, **3**:194*T*
  endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
  fish diversity **3**:473
  fishing/fisheries **6**:422*T*
  forest product exports **2**:451*T*
  freshwater habitats **1**:511*F*
  indigenous peoples **4**:270*T*
  institutional collections and capacity **1**:505*T*
  intercropping systems **2**:382–383
  invasive species **4**:346–347
  oil-palm plantations **1**:540–541, **5**:601–602
  plant species **1**:259*T*
  polyculture cropping systems **7**:230–231
  primate diversity **6**:252*T*
  primate endemism **6**:252*T*
  rice production **1**:86, **1**:88
  species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
  threatened birds **7**:181
  tropical forests **7**:269–270, **7**:270*T*
  zoonotic disease **1**:519–520
  see also Southeast Asia
Indotermes **4**:409*T*
Indo–West Pacific region **3**:473
Indri **4**:728–729
Indriidae **6**:250–252, **6**:251*T*
Indri indri **4**:692, **5**:365, **6**:256–258*T*
indris **6**:251*T*, **6**:256–258*T*
Indus–Ganges monsoon forest **1**:262–263*T*
Indus river dolphin **3**:192
industrial energy **3**:250–253
industrialized agriculture **1:85–94**
  agricultural intensification **1**:86, **1**:122–123, **1**:374–375, **4**:558, **4**:565–566
  commercialization
    basic concepts **1**:90–91
    seed industries **1**:91
  crop genetic diversity
    agrobiodiversity conversation **3**:33
    basic concepts **1**:87–89
    commercialization effects
      general discussion **1**:91–92

industrialized agriculture (*continued*)
intensified environments **1**:92
non-intensified environments **1**:92–93
crop genetic resources **3**:39–40
genetic narrowing **1**:89–90
genetic trends **1**:90
landrace–modern varieties comparisons **1**:89
research summary **1**:93
spatial distribution **1**:87–89, **1**:88*T*
seed selection evolution **1**:86–87
Southeast Asia **1**:517, **1**:519*F*, **1**:520*F*
industrial land **4**:579
industrial waterways **6**:609–610
IndVal (indicator value) method
basic concepts **4**:265
expanded statistical methods **4**:265–266
numerical example **4**:265, **4**:266*T*
infectious diseases
amphibian species decline **3**:696
emerging infectious diseases (EIDs)
human health **1**:368–370, **1**:369*F*
Southeast Asia **1**:521
environmental impacts **3**:285–286
human health–biodiversity relationship
disease management **1**:360–361, **1**:361*F*
hantavirus pulmonary syndrome **1**:365–368, **1**:366*F*
Lyme disease **1**:361–362, **1**:363*F*
schistosomiasis **1**:368, **1**:368*F*
West Nile virus **1**:362–365, **1**:364*F*, **1**:365*F*
zooprophylaxis **1**:360–361
infectious disease regulation **1**:348–350, **3**:349
landscape epidemiology **4:503–507**
background information **4**:503
dynamic spatiotemporal mapping approaches **4**:503
future research areas
genetic studies **4**:505–506
geographic profiling **4**:506
mapless approaches **4**:505
research summary **4**:506
static mapping approaches
basic concepts **4**:503–505, **4**:504*F*
host distribution data **4**:505
incidence data **4**:504–505
normalized difference vegetation index (NDVI) **4**:504
vector distribution data **1**:361–362, **1**:363*F*, **4**:505
parasites **2:523–538**
amphibian species decline **3**:696
anthropogenic drivers
general discussion **2**:531
global climate change **2**:531–532
habitat degradation and pollution **2**:531
habitat loss and fragmentation **2**:531
worldwide commerce and travel **2**:532
characteristics and properties **2**:523, **2**:524*T*
disease causes **2**:523
disease management
captive breeding programs **2**:536–537
disease threat assessments **2**:535
intervention methods **2**:535–536
population monitoring **2**:535
threat classifications **2**:534–535
epidemiology
basic principles **2**:523–525
macroparasitic disease **2**:524*T*, **2**:525–527, **2**:526*F*
microparasitic disease **2**:523–525, **2**:525*F*
SIR compartment models **2**:523–525, **2**:525*F*
intervention methods
antiparasitic drug administration **2**:536
culling **2**:535–536
disease management regimes **2**:535–536
habitat management **2**:536
quarantine **2**:536
vaccination **2**:535–536
parasite diversity management **2**:537
positive roles
biodiversity enhancement **2**:533
ecosystem functionality **2**:534
evolutionary diversification **2**:532–533
host conservation efforts **2**:534
host genetic diversity **2**:533–534
interspecific competition **2**:533
threat potential
community-wide effects/ecosystem consequences **2**:530–531
introduced parasites–species decline relationships **2**:527–529, **2**:528*T*, **2**:529*F*
invasive species–pathogen relationship **2**:529–530
ultraviolet radiation effects **7**:302–303
Infectious Hypodermal and Hematopoietic Necrosis Virus (IHHNV) **1**:195–196
Infectious Salmon Anemia (ISA)virus **1**:196
informal education experiences
botanical gardens **3**:138
museums **3**:138
resources **3**:136–137
youth-serving groups **3**:137, **3**:137
zoos **3**:137–138
infrared spectrum **7**:297*F*
Inga **1**:329
Inga cinnamomea **4**:288*T*
Inga edulis **4**:288*T*
Ingolfiella longipes **3**:202*T*
Ingraham's hutia **4**:713–714
in situ conservation
agrobiodiversity conversation **3**:33
botanical gardens **6**:516
genetic reserve conservation **4**:316*F*, **4**:320, **4**:320*F*, **4**:321*T*
maintenance strategies **1**:405–406
on-farm conservation **3**:33, **4**:316*F*, **4**:320, **4**:321*T*
plant biodiversity **6**:63
research programs and methodologies **4**:316*F*, **4**:320
in vitro fertilization (IVF) clinics **1**:645–646, **3**:652
iniquiline communities **6**:45
inland deserts **2**:458, **2**:461*F*, **2**:462*F*
inland waters *see* freshwater ecosystems
inland wetlands **7**:370*T*
see also wetlands
Inocybe **3**:635*T*
Inoviridae **5**:130*F*, **5**:134*T*, **5**:136
Insecta/insects **4:308–312**
Antarctica **1**:184*T*
anthropogenic activities and impacts **7**:307*T*
Arctic ecosystems **1**:231*T*
brood parasitism **5**:503
characteristics **4**:392, **4**:393*F*
critical pH levels **6**:153*F*
domesticated ruminants **1**:709–710
endangered species **3**:226
eusocial insects
Hymenoptera **4**:181–182, **4**:414*T*
Isoptera **4**:412–413, **4**:412*F*, **4**:414*T*, **4**:421–422
extinction rates **4**:659*T*, **5**:439–440
fire/fire regimes **3**:440
food webs **7**:266*F*, **7**:267*T*
forest canopies **3**:512–513, **3**:513*F*, **3**:514*F*
freshwater ecosystems
acidification effects **3**:584–585, **3**:584*F*
freshwaters **3**:571–572
freshwater sediments **3**:571–572, **3**:572*T*
threatened species **3**:578–579, **3**:579*T*
fruit and seed dispersal **6**:50, **6**:51*T*
genetic diversity **3**:665*F*
grassland ecosystems **4**:3
guilds **4**:35–36, **4**:36*T*
herbicide use effects **4**:93
Holocene postglacial spread **2**:550, **2**:551–552*T*
indicator species **7**:206–207*F*
insecticide resistance **4**:298–299
introduced species
Australia **1**:76*T*, **1**:81
Brazil **1**:76*T*, **1**:83
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81
time lag–evolution relationship **4**:365
United Kingdom **1**:76*T*, **1**:79–80
United States **1**:76*T*, **1**:78
known and estimated species **4**:384*T*
lake and pond ecosystems **4**:371–373, **4**:461*T*
Late Permian extinction event **5**:171
major divisions and characteristics **4**:308–312, **4**:309*T*
mangrove ecosystems **5**:13, **5**:13*F*
migratory patterns **5**:316
mites **1**:214–215
modern extinctions **5**:368*T*
molecular level diversity **2**:635*T*
Near East ecosystems **5**:460*T*
nematode predation **7**:430–431
number of species **4**:371*T*, **5**:669–670*T*

plant–insect associations **1**:360, **4**:35–36, **4**:36*T*, **5**:354–355, **6**:48
pollination **6**:141*T*
prevalence **4**:308
quantity estimates **1**:326–328
research summary **4**:312
river ecosystems **6**:478–479, **6**:478*T*
sandy beaches **2**:131–133
social insects
Hymenoptera **4**:181–182
Isoptera **4**:412–413, **4**:412*F*, **4**:414*T*, **4**:421–422
special associations **3**:495
social systems **2**:646
Southeast Asia **1**:514–515
subterranean ecosystems **7**:53
systematic classifications **7**:91–93, **7**:92*F*
threatened species **1**:514–515, **7**:206–207*F*
tropical rain forests **1**:263
urbanization threats **4**:564
wetland ecosystems **7**:389–391
insect growth regulators **5**:694
insecticides
animal damage control **6**:342
fish poisons **1**:589
inorganic insecticides **5**:693
insecticide resistance **4:298–307**
cross-resistance **4**:302
diagnostic assays **4**:302–303, **4**:302*F*
gene homology **4**:301
management strategies **4**:305–306
multiple resistance **4**:302
mutation recurrence **4**:301–302
nonpest species **4**:306–307
origins **4**:299–300
prevalence **4**:298–299, **4**:298*F*, **4**:299*F*
research background **4**:298
research summary **4**:307
resistance gene selection
determinant factors **4**:303
ecological influences **4**:303–304
fitness costs **4**:305
genetic influences **4**:303
operational influences **4**:304–305
resistance mechanisms
characteristics **4**:300
increased detoxification **4**:300, **4**:300*F*
target site alterations **4**:300–301, **4**:300*F*
transgenic plants **4**:307
Near East ecosystems **5**:473–475
neem tree **1**:589–590
organic insecticides
antibiotics **5**:694
botanicals **5**:693
carbamates **5**:694, **5**:699*T*, **5**:700*T*
dinitrophenols **5**:694
formamidines **5**:694
fumigants **5**:694
insect growth regulators **5**:694
microbials **5**:694
new insecticides **5**:694
organochlorines **5**:693, **5**:699*T*, **5**:700*T*
organophosphates **5**:693–694, **5**:699*T*, **5**:700*T*
organotins **5**:694
petroleum oils **5**:694
pyrethroids **5**:694, **5**:699*T*
semiochemicals **5**:694
Insectivora **4**:689, **4**:690*T*, **4**:701*T*, **6**:309*T*
instar analysis **3**:68
Institutions of Sustainability (IoS) framework **1**:338
Instituto de Investigación de Recursos Biológicos Alexander von Humboldt **1**:507
Instituto Nacional de Biodiversidad (INBio) **1**:506
insurance hypothesis **1**:378
integrated aquaculture **1**:198
integrated coastal zone management (ICZM) system **5**:452
Integrated Conservation Development Projects (ICDPs) **2**:231, **2**:253–254
integrated pest management (IPM) systems **4**:306–307
Integrated Valuation of Ecosystem Services and Trade-offs (InVEST) software **5**:357–358, **5**:459, **5**:460–462*T*, **5**:463*F*
integrated weed management (IWM) **4**:88
integrity, ecological **3**:310–311, **3**:312*F*
interaction strengths **3**:507
Inter-American Institute for Global Change Research (IAI) **4**:327
intercropping **2:382–395**
above-ground species interactions
canopy heat capture **2**:387, **2**:387*F*
characteristics **2**:386–387
light interception/light use efficiency **2**:386–387
shading effects **2**:387, **2**:387*F*
spatiotemporal complementarity **2**:387–388
Africa **2**:382–383
Asia **2**:383, **2**:384*F*
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391
crop pest and disease control
biological pest control **2**:391
genetic diversity effects **2**:391–392
push–pull strategies **2**:391
species diversity effects **2**:392
definition **2**:382–383
importance **2**:382–383
intercropping patterns **2**:383
Latin America **2**:382–383
overyielding
above-ground–below-ground species interactions **2**:386
above-ground species interactions **2**:386–387
below-ground species interactions **2**:388–389
classifications **2**:386
crop pest and disease control **2**:391
land equivalent ratio (LER) **2**:386
summary discussion **2**:392
intercrops **1**:375–376, **7**:230–231
interference competition **4**:360, **6**:669–670, **6**:716–717
intergenic transcribed spacer (ITS) analysis **5**:264–265, **5**:264*T*, **5**:265*F*, **5**:266*F*
interglacial cycles **2**:59–60, **2**:82–83
Intergovernmental Oceanographic Commission (IOC) **4**:329
Intergovernmental Panel on Climate Change (IPCC) **2**:73, **2**:246–247, **2**:486–487, **3**:282–283, **4**:324–325, **5**:449
Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES) **2**:246–247, **3**:101–102, **4**:324–325
Interim General Models (IGMs) **1**:568
intermediate beaches **2**:128–129, **2**:129*F*, **2**:130*F*
intermediate disturbance hypothesis **2**:115–117, **2**:563, **4**:79, **5**:256
intermittent lakes **7**:50
internal eliminated sequences (IESs) **5**:569, **5**:571*F*
International Agenda for Botanic Gardens in Conservation **6**:514
International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD) **4**:328
International Association for Plant Taxonomy **4**:327
International Association for Vegetation Science **5**:556
International Association of Botanical and Mycological Societies (IABMS) **6**:511–512
International Association of Botanic Gardens (IABG) **6**:511–512
International Barcode of Life (iBOL) **4**:327
International Biological Program (IBP) **4**:403–405
International Botanical Congress **5**:551
international business strategies **4**:330
International Census of Marine Microbes (ICOMM) **2**:21, **2**:21*F*
International Centre for Agricultural Research in the Dry Areas (ICARDA) **3**:643
International Code of Botanical Nomenclature **5**:551
International Code of Nomenclature **5**:551
International Code of Phylogenetic Nomenclature **5**:554
International Code of Phytosociological Nomenclature **5**:556
International Code of Zoological Nomenclature **5**:552–553

International Commission on Zoological Nomenclature **4**:327
International Committee on Bionomenclature **5**:553–554
International Committee on Prokaryote Systematics (ICPS) **5**:552
International Committee on theTaxonomy of Viruses **5**:552
International Committee on Virus Taxonomy (ICVT) **5**:129
international common property resources **2**:152–153
International Conference on Population and Development (ICPD) **6**:207–208
International Convention for the Conservation of Migratory Species of Wild Animals **2**:286
International Convention for the Regulation of Whaling (1946) **3**:713*T*, **3**:714
International Council for Science (ICSU) **4**:326–327
International Council for the Exploration of the Sea (ICES) **4**:329
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) **3**:643
International Cultivar Registration Authorities (ICRAs) **5**:551
International Dolphin Conservation Act (1992) **3**:716
International Food Policy Research Institute (IFPRI) **4**:328
International Geosphere and Biosphere Program–Global Change in Terrestrial Ecosystems (IGBP–GCTE) **6**:597–598
International Geosphere and Biosphere Program (IGBP) **4**:326–327
International Institute for Applied Systems Analysis (IIASA) **4**:327
International Institute for Environment and Development **4**:330
International Mycological Association **5**:556
international organizations **4:324–331**
  agriculture biodiversity **4**:328
  aquatic biodiversity **4**:329–330
  background and functional role **4**:324
  biodiversity conservation **4**:325–326
  biodiversity data banks **4**:327
  international business strategies **4**:330
  relief organizations **4**:330
  scientific research cooperation **4**:326–327
  United Nations agencies **4**:324–325
International Panel on Climate Change (IPCC) **6**:597–598
International Plant Genetic Resources Institute (IPGRI) **3**:643
International Plant Protection Convention **3**:209
international primate trade **6**:254–255
International Programme of Biodiversity Science (DIVERSITAS) **4**:326–327
International Rice Research Institute (IRRI) **1**:88, **1**:388–389, **3**:643
International Species Information System (ISIS) **1**:662–663, **7**:473
International Trade Center (ITC) **4**:330
International Treaty on Plant Genetic Resources for Food and Agriculture **1**:466–467, **3**:39–40
International Tropical Timber Organization (ITTO) **1**:455, **4**:328–329
International Undertaking on Plant Genetic Resources (2000) **3**:713*T*
International Union for Conservation of Nature and Natural Resources (IUCN)
  background information **4**:324–325
  conservation status of reptiles **3**:216, **3**:216*T*
  endangered species protection **1**:662–663, **2**:108–109, **2**:243–244, **3**:199, **3**:200–201*T*, **3**:209, **3**:223–224, **4**:325–326
  people–conservation relationship **2**:222–223, **2**:231–232
  primate conservation status **6**:255
  projected and realized extinctions **2**:169
  protected area classifications **2**:289–292, **2**:290, **5**:442–443, **5**:443*T*, **6**:485–486, **6**:491, **6**:491*T*
  Red List Index (RLI)
    background information **4**:327
    categories and criteria **7**:196–197, **7**:197
    endangered species protection **3**:223–224, **4**:325–326
    future extinction estimations **4**:660
    geographic distribution **7**:198*F*
    global species extinction risk assessments **7**:194–195
    historical perspective **7**:192
    major biodiversity threats **4**:314–315, **4**:314*F*
    present extinctions **2**:169, **2**:170*F*
    regional and national extinction risk assessments **7**:195–197
    resource allocations **7**:207
    threatened species
      amphibians **1**:176–177, **1**:515, **7**:201*T*
      angiosperms **6**:60
      birds **7**:180–181, **7**:201*T*
      coral reef ecosystems **4**:664
      Denmark **3**:579–580
      Europe **3**:578–579, **3**:579*T*
      fish **3**:451–452, **7**:201*T*
      future extinction estimations **4**:660
      future outlook **7**:208–209
      gymnosperms **6**:60
      indicator species development **7**:205–207, **7**:206–207*F*
      insects **1**:514–515
      invertebrates **7**:201*T*
      mammals **1**:516–517, **3**:194–195, **3**:195*F*, **4**:706–707, **4**:709, **7**:201*T*
      North America **3**:578–579, **3**:579*T*
      plants **6**:84–85, **7**:201*T*
      primates **6**:255, **6**:256–258*T*, **6**:258*T*
      reptiles **1**:515–516, **7**:201*T*
      seedless vascular plants **6**:59
      terrestrial invertebrates **3**:225
      United States **3**:579–580
      vertebrates **7**:201*T*
  species numbers **5**:703–705
  Species Survival Commissions (SSC) **3**:195–196
  Threatened Plants Committee (TPC) **6**:78
International Union for the Protection of New Varieties of Plants (IPOV) **5**:551
International Union of Biological Sciences (IUBS) **4**:326–327, **5**:553–554, **6**:511–512
International Union of Microbiological Sciences (IUMS) **4**:326–327, **5**:552, **5**:553–554
International Union of Soil Science (IUSS) **6**:597–598
International Whaling Commission **3**:454, **6**:421
Internet Assigned Numbers Authority (IANA) **1**:441–442
Internet Engineering Task Force (IETF) **1**:441–442
Internet use impacts **7**:249
interspecific competition **2:177–191**
  characteristics **6**:716–717
  classifications **2**:177–178, **2**:177*T*
  coevolution **2**:140–141
  competition models
    aggregation model **2**:181–182, **2**:182*F*
    density-dependent function **2**:178, **2**:178*F*
    Lotka–Volterra model **2**:178–180, **2**:178*F*, **2**:179*F*, **6**:716–717, **6**:716*F*, **6**:717*F*
    MacArthur model **2**:180–181, **2**:180*F*, **2**:181*F*
    resource utilization curves **2**:180–181, **2**:180*F*, **2**:181*F*
  field experiments
    Anolis lizards **2**:187–188, **2**:187*T*
    ants and rodents **2**:186–187, **2**:187*T*
    barnacles **2**:186, **2**:187*F*
  field observations
    African wild dogs and hyenas **2**:190, **2**:190*F*, **2**:190*T*
    New England warblers **2**:188–189, **2**:188*F*
    Pacific Island birds **2**:189–190, **2**:189*F*
  laboratory experiments
    Drosophila **2**:185–186, **2**:186*F*
    flour beetles (*Tribolium*)
      environmental conditions **2**:184
      single-species results **2**:184, **2**:184*T*
      two-species results **2**:184–185, **2**:185*F*, **2**:185*T*
    protozoa **2**:182–184, **2**:183*F*, **2**:183*T*, **2**:184*F*, **6**:669–670
    yeasts **2**:182–184, **2**:183*T*
  parasite–species interactions **2**:533
  research summary **2**:190–191
  two-prey competition **2**:177, **2**:177*F*
interspecific hybridization
  allopolyploid/amphiploid speciation **6**:627, **6**:627*F*

characteristics **6**:626
direct allogenous speciation **6**:627
introgressive hybrid speciation **6**:626
recombinational hybrid speciation **6**:626
reticulate allogenous speciation **6**:627–628
wild sunflowers **6**:626–627
interstitial communities
Lobau wetlands (Austria) **7**:59
Rhone River (France) **7**:59
intertidal beaches **5**:51
intertidal ecosystems **4:332–344**
anthropogenic impacts **5**:56–57, **5**:58*F*, **5**:59*F*
biodiversity measures **4**:339–340
boulder fields **4**:336–337, **4**:336*F*
ecological function **4**:340–342
general discussion **4**:332
habitat–niche relationship **4**:45–46
mangrove forests **4**:337, **4**:338*F*, **4**:339*F*
marine sedimentary habitats **5**:107, **5**:108*T*
pattern variations
competition **4**:334, **4**:335*F*
general discussion **4**:333–334
interaction chains **4**:334
physical disturbances **4**:335–336
predation **4**:334–335
recruitment **4**:333–334
processes and characteristics **4**:332–333
research summary **4**:342–343
sandy beaches and mudflats **4**:337–339
tides and waves **4**:332–333
values and services **4**:342, **4**:342
intertropical convergence zone (ITCZ) **1**:46, **7**:270–271
intimate mutualism **2**:323
intracellular organelles **2**:327
intraguild interference **1**:378–379, **1**:378*F*
intraspecific parasites **5**:657
introduced species **1:75–84**
amphibians **3**:167
amphibians and reptiles
amphibian species decline **3**:695–696
Australia **1**:80
United States **1**:78
animals
Australia
nonindigenous invertebrates **1**:76*T*, **1**:80–81
nonindigenous vertebrates **1**:76*T*, **1**:80
Brazil
nonindigenous invertebrates **1**:76*T*, **1**:83
nonindigenous vertebrates **1**:76*T*, **1**:83
India
nonindigenous invertebrates **1**:76*T*, **1**:82
nonindigenous vertebrates **1**:76*T*, **1**:82
nonindigenous invertebrates
Australia **1**:76*T*, **1**:80–81
Brazil **1**:76*T*, **1**:83
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81
United Kingdom **1**:76*T*, **1**:79–80
United States **1**:76*T*, **1**:78
South Africa
nonindigenous invertebrates **1**:76*T*, **1**:81
nonindigenous vertebrates **1**:76*T*, **1**:81
United Kingdom
nonindigenous invertebrates **1**:76*T*, **1**:79–80
nonindigenous vertebrates **1**:76*T*, **1**:79
United States
nonindigenous invertebrates **1**:76*T*, **1**:78
nonindigenous vertebrates **1**:76–77, **1**:76*T*
aquaculture **1**:194–196
Ascension Island **6**:78
Australia
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:81
crop plant pathogens **1**:76*T*, **1**:81
livestock pathogens **1**:76*T*, **1**:81
nonindigenous plants **1**:76*T*, **1**:80
nonindigenous vertebrates
amphibians and reptiles **1**:80
birds **1**:80
mammals **1**:76*T*, **1**:80
biodiversity impacts **7**:203*F*
birds
Australia **1**:80
Brazil **1**:83
India **1**:82
South Africa **1**:81
United Kingdom **1**:76*T*, **1**:79
United States **1**:76*T*, **1**:77–78
Brazil
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:83
crop plant pathogens **1**:76*T*, **1**:83
nonindigenous plants **1**:76*T*, **1**:82–83
nonindigenous vertebrates
birds **1**:83
mammals **1**:76*T*, **1**:83
climate change effects **2**:104
contemporary conservation theories and practices **2**:254–255
dispersal biogeography **2**:554–555, **2**:554*F*
economic impact **1**:75
endangered species
freshwater invertebrates **3**:181
mammals **3**:193
reptiles **3**:217–218
exotic birds **1**:76*T*, **1**:77–78
fire/fire regimes **3**:440
forest ecosystems **3**:534
forest management impact **1**:451
global agriculture impacts **3**:429
globalization effects **3**:701–703, **3**:702*F*
grazer-based ecosystem restoration and management **4**:16–17
impacts and distribution **4:357–368**
basic concepts **4**:357
biodiversity threats **2**:80
competition
general discussion **4**:360
interference competition **4**:360
resource competition **4**:360
continental ecosystems **4**:155
coral reef ecosystems **2**:341
distribution patterns
aquatic habitats **3**:411
biodiversity loss **4**:663
characteristics **3**:410–411
island biogeography **3**:217–218, **3**:410–411, **3**:410*F*, **6**:83
endangered mammals **3**:193
energy flow **3**:246–247
environmental impacts **3**:284
fish species **3**:448
freshwater invertebrates **4**:378
geographic distribution **4**:358
habitat distribution **4**:358–359
habitat modification
general discussion **4**:359
grazing and rooting **4**:359
modified fire regimes **4**:360
modified hydrology **4**:360
modified nutrient regime **4**:360
shading and overgrowth **4**:359–360
herbivory **4**:361
human impacts **4**:143–144
hybridization and introgression **4**:362
impact quantification **4**:366–367
indirect effects
chain reactions **4**:363–364
characteristics **4**:362–363
classical indirect effects **4**:363
pathogen vectors **4**:364
invasional meltdowns
basic concepts **4**:364
introduced animals **4**:364–365
introduced plants **4**:364
mutualism **4**:365
pollination and dispersal **4**:364
island biogeography **4**:155
island vulnerability **4**:359
modern extinctions **5**:370
neotropical mammals **2**:590
onshore ocean waters **4**:155–156
parasitism and disease **4**:361–362
pathogens **4**:156
plant–soil interactions **6**:125
predation **4**:360–361, **6**:240–241
Southeast Asia **1**:518–520, **1**:521, **1**:522*F*, **4**:346–347
terrestrial invertebrates **3**:222
time lag–evolution relationship **4**:365–366
timing factors **4**:358
wetland ecosystems **7**:399
India
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:82
crop plant pathogens **1**:76*T*, **1**:82
livestock pathogens **1**:76*T*, **1**:82
molluscs **1**:82
nonindigenous plants **1**:76*T*, **1**:82
nonindigenous vertebrates
birds **1**:82
mammals **1**:76*T*, **1**:82

introduced species (*continued*)
invasive species–pathogen relationship **2**:529–530
lake and pond ecosystems **4**:465
managed relocations **7**:238–239, **7**:239*F*
marine sedimentary habitats **5**:113*F*, **5**:114
plant invasions **6:90–102**
basic concepts **6**:90–92, **6**:91*F*
current extent **6**:93–95, **6**:94*F*
dispersal pathways **6**:90–92, **6**:91*F*
general discussion **6**:90
historical perspective **6**:92–93
invasibility **6**:97–98, **6**:99*F*
invasion process
dispersal **6**:95–96
establishment and naturalization **6**:96
spread rates **6**:96–97
stages **6**:95–96, **6**:95*F*
invasiveness **6**:97
management strategies
biotic homogenization **6**:100–101*F*, **6**:101
ecological management **6**:101
potential conflicts **6**:101–102
prevention methods **6**:101
removal methods **6**:101
modeling approaches **6**:99–101
plant trait correlations **6**:98–99
plants
Australia **1**:76*T*, **1**:80
Brazil **1**:76*T*, **1**:82–83
forest management impact **1**:451
impact predictions and evaluations **4**:355–356
India **1**:76*T*, **1**:82
aquatic ferns (*Salvinia molesta*) **4**:346–347, **4**:347*F*
Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346, **4**:346*F*
background information **4**:345
cheatgrass (*Bromus tectorum*) **4**:347–348, **4**:348*F*
endangered species **3**:208
Fallopia japonica **4**:351–352, **4**:352*F*
impact predictions and evaluations **4**:355–356
Lantana camara **4**:352–353, **4**:353*F*
Mimosa pigra **4**:353–354, **4**:354*F*
mouse-ear hawkweed (*Hieracium pilosella*) **4**:348–349, **4**:349*F*
Pinus pinaster **4**:354–355, **4**:355*F*
prickly pear (*Opuntia*) **4**:350–351, **4**:351*F*
Spartina alterniflora **4**:349–350
pollinators **6**:144–145
South Africa **1**:76*T*, **1**:81
United Kingdom **1**:76*T*, **1**:79
United States **1**:75–76, **1**:76*T*
pollinators **6**:144–145
population diversity **6**:172–173
research summary **1**:83
river ecosystems **6**:481*T*, **6**:482
South Africa
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:81
crop plant pathogens **1**:76*T*, **1**:81–82
livestock pathogens **1**:76*T*, **1**:82
nonindigenous plants **1**:76*T*, **1**:81
nonindigenous vertebrates
birds **1**:81
mammals **1**:76*T*, **1**:81
taxonomic versus functional diversity **2**:655
threatened birds **7**:184
United Kingdom
nonindigenous invertebrates
arthropods **1**:76*T*, **1**:79–80
crop plant pathogens **1**:76*T*, **1**:80
nonindigenous plants **1**:76*T*, **1**:79
nonindigenous vertebrates
birds **1**:79
mammals **1**:76*T*, **1**:79
United States
background information **1**:75–76
nonindigenous invertebrates
earthworms **1**:78
general discussion **1**:78
insects and mites **1**:76*T*, **1**:78
microbes **1**:76*T*, **1**:78–79
nonindigenous plants
crop plants **1**:75–76
economic impact **1**:76, **1**:76*T*
nonindigenous vertebrates
amphibians and reptiles **1**:78
exotic birds **1**:76*T*, **1**:77–78
mammals **1**:76–77, **1**:76*T*
introgressive hybrid speciation **4**:166–167, **6**:626
invasive species
amphibian species decline **3**:695–696
Antarctic ecosystems **1**:185–186, **1**:186*T*
aquaculture **1**:194–196
Australia **1**:303
biodiversity impacts **7**:203*F*
climate change effects **2**:104
contemporary conservation theories and practices **2**:254–255
dispersal biogeography **2**:554–555, **2**:554*F*
disturbance ecology **2**:566
endangered reptiles **3**:217–218
fire/fire regimes **3**:440
forest ecosystems **3**:534
forest management impact **1**:451
global agriculture impacts **3**:429
globalization effects **3**:701–703, **3**:702*F*
grazer-based ecosystem restoration and management **4**:16–17
impacts and distribution
biodiversity threats **2**:80
continental ecosystems **4**:155
coral reef ecosystems **2**:341
distribution patterns
aquatic habitats **3**:411
biodiversity loss **4**:663
characteristics **3**:410–411
island biogeography **3**:217–218, **3**:410–411, **3**:410*F*, **6**:83
endangered mammals **3**:193
energy flow **3**:246–247
environmental impacts **3**:284
fish species **3**:448
freshwater invertebrates **4**:378
human impacts **4**:143–144
island biogeography **4**:155
modern extinctions **5**:370
neotropical mammals **2**:590
onshore ocean waters **4**:155–156
pathogens **4**:156
plant–soil interactions **6**:125
Southeast Asia **1**:518–520, **1**:521, **1**:522*F*, **4**:346–347
terrestrial invertebrates **3**:222
wetland ecosystems **7**:399
invasive species–pathogen relationship **2**:529–530
managed relocations **7**:238–239, **7**:239*F*
marine ecosystems **5**:54, **5**:60–61
marine sedimentary habitats **5**:113*F*, **5**:114
North American biodiversity impacts **5**:566
optimal control theory **3:16–24**
basic concepts **3**:16–19
endogenous risk approach **3**:16
Horan model **3**:22–23
invasion process **3**:17, **3**:17*F*
public policies **3**:16
Reed's model
basic concepts **3**:17
ex ante management (prevention) **3**:18–19, **3**:19*T*
ex post management (control) **3**:19–22, **3**:21*F*
research summary **3**:22
plant invasions **6:90–102**
basic concepts **6**:90–92, **6**:91*F*
current extent **6**:93–95, **6**:94*F*
dispersal pathways **6**:90–92, **6**:91*F*
general discussion **6**:90
historical perspective **6**:92–93
invasibility **6**:97–98, **6**:99*F*
invasion process
dispersal **6**:95–96
establishment and naturalization **6**:96
spread rates **6**:96–97
stages **6**:95–96, **6**:95*F*
invasiveness **6**:97
management strategies
biotic homogenization **6**:100–101*F*, **6**:101
ecological management **6**:101
potential conflicts **6**:101–102
prevention methods **6**:101
removal methods **6**:101
modeling approaches **6**:99–101
plant trait correlations **6**:98–99
population diversity **6**:172–173
predation effects **6**:240–241
range ecology **6**:341
river ecosystems **6**:481*T*, **6**:482
taxonomic versus functional diversity **2**:655
threatened birds **7**:184
inverse texture hypothesis **3**:75
invertebrates
alpine ecosystems **1**:154–155, **1**:155*F*
ancient lakes species flocks **4**:464, **4**:464*T*

Antarctica **1**:186*T*
Australia **1**:300–302, **1**:303*T*
average lifespan **4**:656*T*
coral reef ecosystems **2**:332*T*
deep-sea hydrothermal vents
  biodiversity variability **7**:331
  biogeography **7**:330
  gene flow/genetic diversity **7**:330–331
  prevalence **7**:329–330
  species origins **7**:330
domesticated ruminants **1**:709–710
endangered freshwater invertebrates **3:176–187**
  American Southeast rivers **3**:184, **3**:185*F*
  anthropogenic impacts
    direct harvests **3**:181
    global climate change **3**:181–182
    habitat destruction/degradation **3**:180–181, **3**:180*F*
    introduced species **3**:181
    pollution **3**:181, **3**:181*F*
  causal factors **3**:177, **3**:177*F*
  desert spring communities **3**:186, **3**:187*F*
  diversity hot spots **3**:176
  global conservation status **3**:176–177, **3**:176*T*, **3**:177*F*
  global distribution patterns **3**:182–184, **3**:182*T*
  global diversity **3**:176
  number of species **3**:182–184, **3**:182*T*
  protection strategies **3**:184–187
  small ranges
    causal factors **3**:178
    conservation status **3**:177*F*
    dispersal ability **3**:178*F*
    high endemic hot spots **3**:178–179
    high endemic species **3**:179, **3**:180*F*
    Lake Baikal **3**:179, **3**:183*F*
    narrowly endemic species **3**:177–178, **3**:177*F*, **3**:179
  vulnerability factors
    small ranges **3**:177–178, **3**:177*F*
    sparse populations **3**:179–180
forest canopies **3**:512–513, **3**:513*F*, **3**:514*F*
forest management impact **1**:451–452
fossil record **3**:687–688, **3**:687*F*
freshwater ecosystems **4:369–378**
  background information **4**:369–370
  ecological functions and processes
    characteristics **4**:375
    decomposition and recycling **4**:375
    material transfer and uptake **4**:375–376
    primary production **4**:376
  freshwater sediments **3**:572*T*
  global patterns **4**:373
  influencing factors
    disturbances **4**:375
    endemism **4**:374–375
    habitat heterogeneity **4**:375
    isolation factors **4**:374–375
    origins **4**:374
    research background **4**:373–374
    species interactions **4**:375
  life cycles **4**:370, **4**:372*F*
  major groups **4**:370, **4**:371*T*
  major habitats **4**:369–370, **4**:369*F*
  major threats
    altered water regimes **4**:377–378
    biodiversity loss **4**:376–377
    exotic species invasions **4**:378
    habitat depletion and loss **4**:376–377
    pollution **4**:377
    reduced water quality **4**:377
  number of species **4**:370–371, **4**:371*T*
  river continuum concept (RCC) **4**:373*F*
  taxonomic groups **4**:371–373
  zonal representatives **4**:374*F*
fruit and seed dispersal **6**:50
genetic diversity **3**:665*F*, **3**:668*F*
indicator species **7**:206–207*F*
introduced parasites–species decline relationships **2**:528*T*
introduced species
  Australia **1**:76*T*, **1**:80–81
  Brazil **1**:76*T*, **1**:83
  India **1**:76*T*, **1**:82
  South Africa **1**:76*T*, **1**:81
  United Kingdom **1**:76*T*, **1**:79–80
  United States **1**:76*T*, **1**:78
lake and pond ecosystems **4**:461*T*
latitudinal range shifts **4**:604*T*
marine invertebrates **4:379–391**
  average lifespan **4**:656*T*
  benthic large-scale diversity patterns
    bathymetric gradients **4**:387, **4**:388*F*
    latitudinal gradients **4**:385–387, **4**:386*F*, **4**:387*F*
  benthic small-scale diversity patterns **4**:387–389, **4**:388*T*, **4**:389*F*
  classifications and sampling methods
    benthic species **4**:381–382, **4**:381*T*, **4**:382*F*
    pelagic species **4**:382–383, **4**:383*T*
  cold seeps **4**:391
  diversity control processes **4**:380–381
  endangered species **3:199–204**
    causal factors **3**:199–202
    definition **3**:199
    extinctions **3**:202*T*
    IUCN criteria **3**:200–201*T*
    recovery strategies **3**:203
    research challenges **3**:202–203
    species examples **3**:199–202, **3**:201*T*, **3**:202*T*
    susceptible habitats and species **3**:202, **3**:202*T*, **3**:203*T*
  fossil record **4**:389–391, **4**:390*F*
  global diversity **4**:384–385, **4**:385*T*
  hydrothermal vents **2**:6*F*, **4**:391
  macrofauna **4**:380–381, **4**:381*T*
  megafauna **4**:380–381, **4**:381*T*
  meiofauna **4**:380–381, **4**:381*T*
  ocean characteristics **4**:379–380
  pelagic environment
    benthic biomass distribution **4**:380–381, **4**:380*F*
    pelagic diversity patterns **4**:385, **4**:386*T*
    physical characteristics **4**:379, **4**:380*T*
  Phanerozoic diversity **3**:683–684, **3**:685*F*, **3**:686*F*
  spatio-temporal mosaic theory **4**:388–389, **4**:389*F*
  taxonomy and species diversity **4**:383–384, **4**:384*T*
  ultraviolet radiation effects **7**:300
mites **1**:214–215
modern extinctions **5**:368*T*
mutation rates **2**:628*F*
Phanerozoic diversity **3**:687–688, **3**:687*F*
phenological shifts **5**:719*F*
pollination **6**:49*T*, **6**:140–141, **6**:141*T*
riparian zones **6**:462–463
river ecosystems **6**:478–479, **6**:478*T*
seagrass ecosystems **6**:547–548
slash-and-burn systems **6**:556
storage ecology **7**:24
Succulent Karoo **1**:488
surface water acidification **1**:10*T*
temperature tolerance ranges **4**:115
terrestrial invertebrates **4:392–394**
  Antarctica **1**:184*T*
  conservation practices
    filter approach **3**:224
    recovery programs **3**:225
    surrogate use **3**:224–225
    umbrella/flagship species **3**:225
  endangered species **3:219–227**
    centipedes **3**:226
    conservation practices **3**:224
    crabs **3**:226
    daddy long legs **3**:225–226
    earthworms **3**:225
    insects **3**:226
    land snails **3**:226
    millipedes **3**:226
    pill bugs **3**:226
    pseudoscorpions **3**:225–226
    Red List Index (RLI) **3**:225
    spiders **3**:225–226
    status assessments **3**:223–224
    velvet worms **3**:225, **3**:226*F*
    woodlice **3**:226
  endemism **5**:592*T*
  environmental conditions **4**:393
  life histories **3**:220
  species numbers and diversity **3**:219
  taxonomy
    limited-adapted groups **4**:392–393, **4**:394*F*
    taxonomic impediment **3**:219–220
    well-adapted groups **4**:392, **4**:393*F*, **4**:394*F*
  threats and challenges
    climate change **3**:221
    habitat loss and fragmentation **3**:220–221
    invasive species **3**:222
    overharvesting/overcollecting **3**:222–223
    pest control **3**:221–222
  threatened species **7**:201*T*, **7**:206–207*F*
  wetland ecosystems **7**:389–391
InVEST **3**:101–102
Iocheirata **1**:212–213

Io fluviatilis **3**:185*F*
ioxynil **5**:697
IPAT equation **4**:158–159
ipecac **5**:712*T*
Ipomoea batatas **2**:663–664, **4**:287–290, **4**:289*T*, **6**:553
Ipomoea fistulosa **1**:165
Ipomoea pes-caprae **1**:55, **1**:277, **5**:711–713
ipratropium **5**:712*T*
iprodione **5**:695
Iran
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  ecosystems
    mangroves **1**:275*T*
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  fish diversity **3**:473
  nitrogen deposition **5**:526*T*
  plant species **1**:259*T*
  wheat production **1**:87–88
Iranian (subalpine) desert **1**:262–263*T*
Irano–Anatolian region **6**:82*T*
Iraq
  animal species **1**:261*T*
  botanical gardens **6**:519
  ecosystems
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  energy efficiency measures **3**:254*F*
  plant species **1**:259*T*
Ireland
  botanical gardens **6**:519
  nitrogen deposition **5**:529*F*
  vegetation surveys **3**:341–344*T*
Iridaceae **1**:486*T*, **6**:315*T*
Iridaea boryana **5**:59*F*
iridium concentrations **5**:174–175
Iridomyrmex humilis **3**:208–209
Iridoviridae **5**:134*T*, **5**:135
Iris brevifolia **4**:168
Iris fulva **4**:168
Iris hexagona **4**:168
Iris nelsonii **4**:168
iron (Fe)
  biogeochemical cycles **1**:558–559, **1**:563, **2**:617*F*, **2**:618–619
  carbon cycle management **1**:681
  dinitrogen-fixing (N_2-fixing) microorganisms **2**:617*F*, **2**:618–619
  elemental partitioning **1**:554*T*
  hydrothermal vent ecosystems **7**:328
  inorganic herbicides **5**:695–696
  intercropping systems **2**:390
  relative abundance **1**:553*F*
  river water composition **3**:562*T*
Ironomyiidae **3**:489–491*T*
ironwood **2**:468–469
Irrawaddy Delta **1**:276*T*
Irrawaddy dolphin **5**:81*T*
Isa textula **5**:391*F*
Ischnuridae **1**:209
Ischyropsalidoidea **1**:210
Ischyropsalis **1**:210
Ishmael(Quinn) **4**:653
Islam
  biodiversity conservation **6**:564–565
  Muslim declaration on nature **3**:273
island biogeography **4:395–407**
  basic concepts **2**:348, **4**:395, **5**:511–512
  biodiversity dynamics modeling **5**:326–327
  biodiversity-rich countries **1**:499–500
  birds
    colonization **1**:617–618, **1**:618*F*
    evolutionary development **1**:618–619
    flightless birds **1**:605–606
    fragmentation–extinction rate relationship **1**:619–620
    human colonization impacts **1**:621–622
    introduced species **1**:620
    species–area relationships **1**:618*F*, **1**:619
  deterministic stability **7**:7–8, **7**:8*F*
  domesticated ruminants **1**:702
  ecological perspectives **3**:5
  endangered reptiles **3**:217–218, **3**:217
  endemism **3**:233, **5**:592*T*
  environmental degradation **6**:83
  experimental sampling design
    horizontal approach **4**:406
    transect approach **4**:405–406, **4**:405*F*
    vertical approach **4**:406
  geobotanical research **4**:395, **4**:396*T*
  historical perspective **5**:511–512
  invasive species
    impacts and distribution **3**:217–218, **3**:410–411, **3**:410*F*, **4**:155, **4**:359, **6**:83
    Lantana camara **4**:352–353, **4**:353*F*
    predation effects **6**:240–241
  island–mainland differences model **6**:645–646
  living dead **6**:81
  MacArthur–Wilson equilibrium theory **1**:583–585, **2**:239, **2**:239*F*, **3**:5, **4**:395–396, **4**:396*F*
  mammalian evolution **4**:699
  myriapods **5**:429
  oceanic islands **5:590–599**
    background information and characteristics **5**:590–591, **5**:591*F*
    community characteristics
      disharmony **5**:591–592, **5**:592*F*
      endemism **5**:592, **5**:592*F*, **5**:592*T*
      relictualism **5**:592
    ecological dynamics
      colonization **5**:593
      community formation processes **5**:594–595, **5**:596*F*
      ecological changes **5**:593
      evolutionary changes **5**:593–594
      extinctions **5**:594, **5**:595–597
      species diversification **5**:594, **5**:595*F*
    future research areas
      realistic evolutionary trajectories **5**:597–599
      restoration **5**:599
      taxonomic knowledge **5**:597
      vacant niche space **5**:597
    species characteristics
      innovative behaviors **5**:593
      reduced dispersal capabilities **5**:592–593
      reproductive changes **5**:593
      size changes **5**:593
    species loss
      alien species invasion **5**:596–597, **5**:598*F*
      climate change **5**:597
      extinction rates **5**:595–597
      small population size **5**:597
  PABITRA (Pacific-Asia Biodiversity Transect) initiative **4**:406–407, **4**:406*F*
  protected area systems **6**:493–494
  scaling laws
    spatial scales **4**:395
    timescales **4**:395
  species–area relationships (SARs) **6**:645–646
  taxonomic versus functional diversity **2**:652*F*, **2**:655, **2**:655*T*
  theoretical perspectives
    biome theory **4**:396–397, **4**:397*F*
    MacArthur–Wilson equilibrium theory **1**:583–585, **4**:395–396, **4**:396*F*
    succession/regression theory **4**:397–398, **4**:398*F*
  threatened birds **7**:181
  tropical rain forests **6**:314*T*
  vegetation analysis
    data analysis
      dendograph approach **4**:403*F*, **4**:404*F*, **4**:405
      Hawaii Volcanoes National Park **4**:402*F*, **4**:403–405, **4**:404*F*, **4**:405*F*
      improved display methods **4**:404*F*, **4**:405, **4**:405*F*
      two-way table approach **4**:402*T*, **4**:404*F*, **4**:405
    floristic checklist methods **4**:399
    quantitative sampling methods
      challenges **4**:400
      count-plot method **4**:400, **4**:401–403
      distance methods **4**:400–401, **4**:401–403
      point-centered quarter method **4**:401
      sample plot size **4**:401*F*, **4**:403
      wandering quarter method **4**:401
    relevé sampling method
      basic concepts **4**:399
      cluster analysis **4**:403*F*, **4**:404*F*
      predefinied strata **4**:399
      size determinations **4**:399, **4**:400*F*
      species determinations **4**:399
      species quantity estimations **4**:399–400
    sampling methods
      general discussion **4**:399
      quantitative methods **4**:400
      relevé method **4**:399
island cotton **2**:663–664
island ecosystems
  background information and characteristics **5**:590–591
  oceanic islands **5:590–599**

background information and characteristics **5**:590–591, **5**:591*F*
community characteristics
disharmony **5**:591–592, **5**:592*F*
endemism **5**:592, **5**:592*F*, **5**:592*T*
relictualism **5**:592
ecological dynamics
colonization **5**:593
community formation processes **5**:594–595, **5**:596*F*
ecological changes **5**:593
evolutionary changes **5**:593–594
extinctions **5**:594, **5**:595–597
species diversification **5**:594, **5**:595*F*
future research areas
realistic evolutionary trajectories **5**:597–599
restoration **5**:599
taxonomic knowledge **5**:597
vacant niche space **5**:597
species characteristics
innovative behaviors **5**:593
reduced dispersal capabilities **5**:592–593
reproductive changes **5**:593
size changes **5**:593
species loss
alien species invasion **5**:596–597, **5**:598*F*
climate change **5**:597
extinction rates **5**:595–597
small population size **5**:597
protected area systems **5**:444*T*
Island Rule **4**:699
Islas Malvinas **6**:507*T*
Isle Royal **6**:236–237, **6**:238*F*
Isoberlinia **7**:274–275
Isoberlinia angolensis **1**:68
Isoberlinia doka **1**:68
Isoberlinia tomentosa **1**:50
Isochaeta israelis **5**:470–471, **5**:470*T*
Isoetes braunii **6**:115*T*
Isoetes engelmannii **5**:576
Isoetes lacustris **6**:115*T*
isolation-by-distance (IBD) models **4**:508, **4**:509*F*
Isopoda/isopods
American Southeast rivers **3**:185*F*
characteristics **4**:392
classifications **4**:381*T*
dispersal ability **3**:178*F*
endangered species **3**:226
extinction vulnerability **2**:416–417*T*
lake and pond ecosystems **4**:371–373
marine environments **4**:382*F*
number of species **3**:182*T*, **4**:371*T*
sandy beaches **2**:131–133
seagrass ecosystems **6**:547–548
soil fauna **6**:590*T*
species diversity **4**:386–387, **4**:387*F*
subterranean ecosystems **7**:56, **7**:56*F*
isoproturon **5**:696
Isoptera **4:408–433**
characteristics **6**:588
developmental pathways **4**:412–413, **4**:413*F*
ecosystems
functional role **4**:429–431, **4**:430*T*
human disturbances **4**:432–433, **4**:432*T*
Near East **5**:460*T*
soils **4**:431–432, **4**:431*F*
terrestrial ecosystems **4**:429
families and genera **4**:409*T*
fecundity **4**:421–422
feeding behaviors **4**:413–414, **4**:422*F*
fossil record **4**:410–411
global distribution **4**:408–409, **4**:409*T*, **4**:411–412, **4**:411*F*, **4**:422*F*
global diversification
dispersal ability **4**:426
evolutionary radiation **4**:426–427
higher termite radiation **4**:428–429
major historical events **4**:425–426
one-piece type diversification **4**:427–428, **4**:427*F*
separate type expansion **4**:427–428
life types **4**:414, **4**:422*F*, **4**:423*T*
major divisions and characteristics **4**:308–312, **4**:309*T*
Near East ecosystems **5**:464*T*
nests **4**:414–415, **4**:415*F*, **4**:422*F*
number of species **4**:392*T*, **4**:410*T*, **4**:422*F*
phylogeny **4**:410–411, **4**:410*F*, **4**:417*F*
prey–predator interactions **4**:413, **4**:415*F*
social life
caste differentiation **4**:419–420, **4**:420*F*, **4**:421*F*
characterization **4**:422*F*
eusociality **4**:412–413, **4**:412*F*, **4**:414*T*, **4**:421–422
evolutionary development **4**:424*T*
foraging behaviors **4**:422–424, **4**:425*F*
fungus-growing termites **4**:424–425, **4**:426*F*
life types **4**:423*T*
social diversity **4**:421–422
workers and soldiers **4**:420–421, **4**:421*F*, **4**:423*T*, **4**:424*T*
social systems **2**:646
soil fauna **6**:588, **6**:590*T*
symbiotic relationships
biochemical processes **4**:418*F*
carbon-nitrogen balance symbiosis **4**:416–417, **4**:417–418, **4**:417*F*
cellulose decomposition symbiosis **4**:417
digestive tube **4**:415–416, **4**:416*F*
fungi **4**:418–419, **4**:419*F*
general discussion **4**:408–409
hindgut microenvironment **4**:416*F*
phylogeny **4**:417*F*
symbionts **4**:422*F*
wood feeders versus soil feeders **4**:429, **4**:429*F*
isoxaben **4**:90*T*
isoxaflutole **4**:90*T*
isoxazolidiones **4**:92*T*
Israel
animal species **1**:261*T*
botanical gardens **6**:519
cross-phylogeny regional allozyme diversity **3**:669–670, **3**:669*F*
ecosystems **5:478–487**
climate
dew and fog **5**:479–480
rainfall **5**:465–467, **5**:465*F*, **5**:466*T*, **5**:479, **5**:480*F*
temperature regimes **5**:480
edaphic conditions **5**:478–479
flora **5**:480–481, **5**:481*T*
geographical subdivisions **5**:478, **5**:479*F*
geomorphology **5**:478–479
rock types **5**:478–479
threatened species **1**:278*T*
topography **5**:478
vegetation
Ceratonia siliqua open forests **5**:480*F*, **5**:482–483
deserts **5**:480*F*, **5**:485–486
desert savannoid vegetation **5**:480*F*, **5**:486
dryland vegetation patterns **5**:484
general discussion **5**:481–482
mangroves **5**:480*F*, **5**:487
maquis/forests **5**:480*F*, **5**:481–482
Mediterranean savannoid vegetation **5**:480*F*, **5**:483
Mt. Hermon montane forest **5**:480*F*, **5**:482
Mt. Hermon tragacanth vegetation **5**:480*F*, **5**:483–484
Pistacia lentiscus open forests **5**:480*F*, **5**:482–483
Quercus calliprinos woodlands **5**:480*F*, **5**:482
Quercus ithaburensis open forests **5**:480*F*, **5**:482
reed thickets **5**:480*F*, **5**:486
sand vegetation **5**:480*F*, **5**:486
semisteppe bathas **5**:480*F*, **5**:483
shrub–steppes **5**:480*F*, **5**:484–485
shrub–steppes with trees **5**:480*F*, **5**:485
Sudanian tree oases **5**:480*F*, **5**:486
swamps **5**:480*F*, **5**:486
synanthropic vegetation **5**:480*F*, **5**:486–487
wet salinas **5**:480*F*, **5**:486
Ziziphus lotus **5**:480*F*, **5**:483
vegetation types and coverage percentage **1**:264–265*T*
plant species **1**:259*T*
species reductions or absences **6**:445
Italy
botanical gardens **6**:519
flora **5**:481*T*
invasive species **4**:345–346
vegetation surveys **3**:341–344*T*
wheat production **1**:87–88
Itea **5**:622*F*
Ithomiinae **1**:651*F*, **1**:654
Ituna **2**:138
Iuridae **1**:209
Iva annua **2**:663–664
ivory-billed woodpecker **7**:182
Ivory Coast
botanical gardens **4**:665*T*
cropland conversion **4**:665*T*

Ivory Coast (*continued*)
forest-savanna mosaics **1**:68
indigenous peoples **4**:270*T*
termites **4**:430*T*
ivory gull **7**:186
ivory trade **5**:233–234
Ixodes dammini **5**:415
Ixodes persulcatus **1**:361–362
Ixodes ricinus **1**:361–362
Ixodes scapularis **1**:361–362, **1**:363*F*, **4**:505
Ixodida **1**:216

J

Jaccard index **2**:444–445, **5**:207–208, **6**:659–660
jackals **1**:709, **4**:693, **5**:475, **6**:573
jack beans **2**:663–664
jackfruit **2**:663
jackknife estimator **3**:541–542, **5**:183–184
jackknife support **6**:18–19
jackrabbits **2**:223
jacks **3**:467–471
Jackson, Wes **1**:112
Jacquinia nervosa **2**:28
Jaculus jaculus schlüeri **5**:472
jaguars
adipose tissue **7**:34*F*
conservation biology **2**:243–244
late Quaternary extinction event (LQEE) **4**:724–725*T*
species diversity **4**:693
Jainism **6**:565
Jamaica
botanical gardens **6**:519
endemic plant species **6**:83
lizards **1**:31*F*, **1**:32
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
protected area systems **6**:493–494
Jamaican iguana **3**:217–218
Jamarix **1**:268
Janchen's phylogeny **2**:51, **2**:52*F*
Janzen–Connell effect **6**:318–319
Japan
animal species **1**:261*T*
botanical gardens **6**:519
ecosystems
laurel forests **1**:270–271
mangroves **1**:275*T*
mountain regions **1**:273
temperate deciduous forests **1**:271
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
vegetation types and coverage percentage **1**:264–265*T*
fishing/fisheries **6**:422*T*
indigenous peoples **4**:269*T*
land-use trends **4**:559*F*
nitrogen deposition **5**:526*T*
plant species **1**:259*T*, **6**:82*T*
resource management **6**:386–387, **6**:387*T*
species richness **1**:499*T*
temperate forests
distribution patterns **7**:116–117
dominant floristic features **7**:118
whaling industry **3**:454
Japanese anchovy **6**:422*T*
Japanese barberry **7**:125
Japanese evergreen forest **1**:262–263*T*
Japanese honeysuckle **3**:410–411
Japanese sea lion (*Zalophus japonicus*) **5**:75–76
Japanese stiltgrass **4**:94
Jardim Botânico de São Paulo **6**:507
Jardin des Plantes **5**:405
Jardinella spp. **3**:187*F*
jarrah forests **6**:440*F*
Jasus lalandii **5**:57–58, **5**:60*F*, **6**:233
jatropha **1**:529*T*
Jatropha campestris **1**:67
jaundice **3**:635*T*
Java **1**:262–263*T*, **1**:514*T*, **5**:448–449
Javan rhinoceros **4**:695, **5**:225*T*
Javan slow loris **6**:256–258*T*
Javan surili **6**:256–258*T*
jawless fishes **3**:456–457, **3**:684–685, **7**:335, **7**:337–338
jays **5**:474
Jefferson salamanders **1**:10*T*
Jefferson, Thomas **1**:112
Jeffers, Robinson **4**:652
Jeffrey pine **1**:144
Jehlius cirratus **5**:56–57
Jell-O mold **3**:629*F*
jellyfish **1**:591–592, **2**:21*F*, **4**:371*T*, **4**:381*T*, **6**:25, **7**:24
jelly fungi **3**:629–630
jellynosefish **3**:465
Jena Biodiversity Experiment **4**:586, **4**:586*F*
Jena University Botanic Garden **6**:510–511
Jenny, Hans **6**:111
Jenolan Karst Conservation Reserve **5**:444–445
jerboas **5**:472
Jerusalem artichokes **3**:32*F*
Jerusalem crickets **3**:731–732
jewel beetles **1**:327*T*, **1**:328
Jewish declaration on nature **3**:273, **6**:381, **6**:382
jheels **1**:276*T*
jhum **7**:232
jicama **3**:128
jicaro **4**:15, **4**:592
jirds **5**:472
Jiuzhaigou Nature Reserve **3**:375
Job's tears **2**:663
jock itch **3**:634*T*
Johannsen, Wilhelm **6**:1–2
John Dory **3**:468
Johnson's hutia **4**:705
Joint Group of Experts on the Scientific Aspects of Marine Environment Protection (GESAMP) **4**:329
Joint Nature Conservation Committee (JNCC) **2**:353–354
Joint Photographic Experts Group (JPEG) **1**:445
jojoba bush **2**:468–469, **7**:25–26
Jordan
animal species **1**:261*T*
botanical gardens **6**:507*T*
ecosystems
dryland vegetation patterns **5**:484
threatened species **1**:278*T*
vegetation types and coverage percentage **1**:264–265*T*
plant species **1**:259*T*
see also Near East ecosystems
Jordan Rift Valley **5**:470–471, **5**:470*T*
see also Near East ecosystems
Joshua trees **2**:468–469
Joyce, James **4**:652
Juan de Fuca Plate **5**:563
Juan de Fuca Ridge **7**:327–328
Juan Fernandez fur seal **5**:81*T*
Juan Fernandez Islands **5**:216–217
Jubaea chilensis **5**:217
Judaism
biodiversity conservation **6**:565
Jewish declaration on nature **3**:273, **6**:381, **6**:382
Judeo–Christian tradition **6**:381
Juglandaceae **5**:621–625, **5**:623*F*
Juglans mandschurica **1**:271–272
Juglans regia **2**:662–663
jujubes **2**:383, **2**:384*F*
Julbernardia **7**:274–275
Julbernardia globiflora **1**:68
Julbernardia paniculata **1**:68
Julbernardia seretii **7**:273
Juliania adstringens **2**:26
Julidochromis **1**:25*F*
jumping spiders **1**:207
Junco hyemalis **5**:309–310
Juncus **1**:270
Juncus arabicus **5**:484–485
Juncus effusus **7**:381
Juncus scheuzerioides **4**:107
Junellia **4**:101
Junggar Basin desert **2**:469–470
jungle cat **5**:475
Junin grebe **7**:183–184
juniper **1**:628–630, **6**:342–343
Juniperus **1**:268, **1**:270, **1**:628–630, **6**:342–343
Juniperus bermudiana **6**:78
Juniperus monsperma **4**:212*F*
Juniperus oxycedrus **1**:269
Juniperus phoenicea **5**:480*F*, **5**:485
Juniperus procera **7**:275
Junonia coenia **4**:472
Jurassic period
Age of the Dinosaurs **5**:5–6
amphibians **1**:169
boundary age **5**:1*T*, **5**:2
crustaceans **2**:405
early mammals **3**:188–189
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
marine diversity **5**:168*F*, **5**:169*F*
mass extinctions
extinction intensities **5**:162*F*
marine animals **1**:413*F*, **1**:493*F*
notable mass extinctions **5**:161*F*

number of families **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
originations **1**:495, **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
temperate forests **7**:123*F*
terrestrial plant communities
abundance patterns **6**:74*F*
evolutionary development **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
justice and equity **4:434–441**
affected populations
colonial legacies and restitution **4**:439–440
consequences of conservation **4**:437–438
distribution of action, responsibility, and power **4**:438–439
distribution of rights, title, and access **4**:439
distribution of risks **4**:437–438
distribution of values **4**:439
basic concepts **4**:434–435
biodiversity conservation
best management practices **4**:435
ecosystem services **4**:435
natural resources **4**:435
program support **4**:435
protected areas and heritage sites **4**:435–436
sustainability **4**:436
importance **4**:434
nonhuman organisms **4**:436–437
research summary **4**:440
Justicia **6**:316*T*
Justicia obtusifolia **1**:165
jute **2**:663
juvenile hormone (JH) **2**:427

K

K–12 education
diversity of life concepts **3**:134
formal science education **3**:135–136
grades 3–5 **3**:134
grades 6–8 **3**:134–135
grades 9–12 **3**:135
grades K–2 **3**:133–134
Project 2061 **3**:133–134
Ka'apor capuchin **6**:256–258*T*
Kageneckia angustifolia **4**:96–97, **4**:105–107
kagu **1**:608
kainic acid **1**:592, **6**:132, **6**:134*F*
kaka **7**:184
kakapo **1**:604, **1**:606, **6**:579, **7**:188
kaki **4**:194*T*
Kalametiya Kalapuwa **1**:276*T*
Kalawisan Bay **1**:276*T*
Kalimantan, Indonesia **1**:511*F*, **1**:512*F*
Kaliveli Tank **1**:276*T*
Kalm, Peter **4**:651–652
Kalotermes **4**:409*T*
Kalotermes flavicollis **4**:419, **4**:420*F*
Kalotermitidae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:422*F*, **4**:430*T*
Kalyptorhynchia **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Kamchatkan forest **1**:262–263*T*
Kami **6**:565
Kamptozoa
global diversity **4**:385*T*
marine ecosystems **5**:50*T*
Kandelia **1**:277, **5**:11*T*
kangaroo rats **2**:473, **2**:584*F*, **4**:446–447, **4**:446*F*, **7**:229*T*
kangaroos **4**:687*T*, **4**:688–689, **6**:333
Kant, Immanuel **2**:422
kaolinitic soils **4**:225–226, **4**:225*F*
Kapar Forest Reserve **1**:276*T*
kappaphycus **1**:190*T*
Karagan Lewaya **1**:276*T*
Karenia brevis **1**:470, **1**:477, **1**:477*T*, **3**:357
Karlodinium **1**:472–473
Karnataka, India **2**:354–356, **2**:356*F*
Karoo bustard **1**:487–488
Karoo chat **1**:487–488
karst caves **7**:49–50, **7**:50*F*
karst landscapes **7**:49–50
karyorelictids **6**:295–296, **6**:296*F*
Kashmir langur **6**:256–258*T*
Katagnymeme **2**:611*F*
Katagyneme pelagica **2**:611–612, **2**:612*T*
Katagyneme spiralis **2**:611–612, **2**:612*T*
Kathablepharis **6**:294
Katodinium **6**:294
Katsuwonus pelamis **3**:481
katydids
characteristics **3**:722–723, **3**:725–726, **3**:725*F*
food and feeding habits **3**:724–725
overharvesting/overcollecting **3**:223
paleontology **3**:725
phylogeny and taxonomy **3**:725
recovery programs **3**:225, **3**:225*F*
reproduction and growth **3**:724
sound production **3**:722–723, **3**:723–724
Tettigoniidae
characteristics **3**:726
subfamilies
Acridoxeninae **3**:727
Bradyporinae **3**:726
Conocephalinae **3**:727
Ephippigerinae **3**:727
Phaneropterinae **3**:726–727
Phasmodinae **3**:727
Pseudophyllinae **3**:727
Tettigoniinae **3**:727
Zaprochilinae **3**:727
Kauffman, Stuart **5**:630
kava **5**:704
Kayapó Indians **4**:281–285, **7**:232, **7**:232*F*
kayas **6**:563–564, **7**:229
Kazakhstan
botanical gardens **6**:507*T*
energy efficiency measures **3**:254*F*
megaherbivores **5**:224*T*
Kazumura Cave (Hawaii, USA) **7**:60
Kelletia kelletii **4**:516
Kellet's whelk **4**:516
Kellicottia **6**:26, **6**:31
Kellwasser event **5**:169–170
kelp bass **4**:516
kelp beds **3**:283, **4**:8–9, **4**:443–446, **5**:51, **5**:58–60, **6**:232–233, **6**:233*F*, **6**:234*F*, **6**:247*F*
Kemp's ridley sea turtles **3**:217, **3**:256
Kentish plover **5**:476–477
Kentrophoros **6**:287, **6**:295
Kenya
botanical gardens **6**:519
butterflies **1**:658, **1**:659*T*
floodplain agriculture **7**:146
grass and shrub savannas **1**:67
intercropping systems **2**:382–383
megaherbivores **5**:227–228, **5**:227*F*, **5**:234–236
multifunctional agriculture **1**:117, **1**:117*F*
natural history museums and institutions **5**:406
primate diversity **6**:252*T*
primate endemism **6**:252*T*
termites **4**:430*T*
Keratella **6**:26, **6**:31
Kerguelen archipelago **1**:180–181, **1**:186*T*
kestrels **5**:474
Keteleeria **1**:272–273
Kettlewell, H. B. D. **6**:3
Ketupa zeylonensis **5**:463*T*
Kew Gardens *see* Royal Botanic Gardens (RBG) Kew
Kew Millennium Seed Bank **4**:141
Key Biodiversity Areas (KBAs) **7**:203–204, **7**:204*F*
key deer **3**:192
keyhole limpets **5**:56–57, **5**:58*F*
keystone species **4:442–457**
biomonitoring processes **3**:125
climate change effects **5**:38
community impacts **4**:173
conservation implications **4**:456–457
ecological perspectives **3**:5
ecosystem function **3**:92–93, **3**:92*F*, **3**:246–247
evaluation assessments
concept reevaluation **4**:448, **4**:448*F*, **4**:449*F*
criticisms **4**:447–448
recommended terminology **4**:449
related concepts **4**:448–449
forest ecosystems **3**:535
historical perspective
Enhydra lutris study **4**:443–446, **4**:444*F*, **4**:445*F*
habitat modifiers **4**:447
herbivores **4**:446–447, **4**:446*F*
mutualists **4**:447
Pisaster ochraceus study **4**:442–443, **4**:443*F*
plants **4**:447
pollinators **4**:447
seed dispersers **4**:447
identification processes
challenges **4**:449–451
context dependency **4**:453–454, **4**:454*F*
experimental approaches **4**:451*F*, **4**:452*F*, **4**:453*F*, **4**:453*T*
identifying traits **4**:453*T*

keystone species (*continued*)
nonexperimental approaches **4**:451–452
predictive approaches **4**:452–453
seastar–mussel interaction
strength–seawater temperature relationship **4**:454*F*
species interaction strengths **4**:454–456, 4:454*F*
multispecies interactions **6**:722, **6**:723*T*
wildlife management **7**:406
Khabartal **1**:276*T*
Khabbaki Lake **1**:276*T*
khalifa **6**:564–565
Khao Yai National Park (Thailand) **6**:309
Khaya grandifolia **1**:68
khejri trees **7**:231
kill curve **5**:160–163, **5**:163*F*
killer bees **2**:551–552*T*, **4**:183
killer whales **2**:708, **4**:445–446, **4**:445*F*, **4**:693–694, **5**:58–60, **5**:77*T*, **5**:447
killifish **3**:467, **7**:390
kinetoplastids **5**:572, **5**:573–574*T*
king boletes **3**:629–630
kingfishers **5**:476–477, **7**:390
king salmon **6**:523–524, **6**:525*F*
king snakes **7**:390
kinkajou **2**:590*T*
Kinorhyncha
classifications and sampling methods **4**:381*T*
fossil record **3**:684*F*
global diversity **4**:385*T*
marine ecosystems **5**:50*T*
number of species **4**:384*T*, **5**:669–670*T*
phylogenetic relationships **7**:433*F*
Kinosternon oaxacae **7**:229*T*
kin selection theory **2**:322, **6**:197, **6**:573–574
kipunji **6**:251*T*, **6**:256–258*T*
Kiribati **5**:592*F*, **6**:507*T*
Kirkia acuminata **1**:67
Kirk's red colobus **6**:256–258*T*
Kirtland's warbler **2**:58–59, **5**:370, **7**:184–185
Kissimmee–Okeechobee–Everglades (KOE) region **7**:372, **7**:372*F*, **7**:373*F*
Kitti'shog-nosed bats **1**:601
Kiwa hirsuta **2**:20*F*
kiwis **1**:606, **3**:128
Klang Islands **1**:276*T*
Klebsiella **2**:612*T*, **2**:614–615
Klee diagrams **2**:6, **2**:7*F*
Kleinschmidt theory **7**:46
kleptoparasitism **3**:495, **5**:668
klipspringers **6**:572–573
Kloss' gibbon **6**:256–258*T*
knifefish **3**:464
knock-down resistance (kdr) **4**:301
koalas **4**:687*T*, **4**:688–689
Kobresia **1**:270, **1**:274–275
Kobus ellipsiprymnus **1**:52
Kobus megaceros **1**:53–54
Kochia prostrata **1**:270
kodo millet **3**:32*F*
Koeleria cristata **1**:270
Koeleria gracilis **1**:269
Koeleria macrantha **4**:80*F*
Koenigia islandica **1**:236
Kogia sima **3**:192–193
koinobionts **5**:674–675
kokako **5**:370
kokanee salmon **4**:363–364
kokerbooms **2**:468, **2**:471*F*
Kokia cookei **6**:81
Kokia lanceolata **6**:81
Koliella **6**:29
Kolleru Lake **1**:276*T*
Kolpochoerus **5**:462
Komarov Botanical Institute **5**:411*F*, **6**:508, **6**:510–511
kombu kelp **7**:230
Komodo dragon **7**:223
Komokiacea **5**:48
Koompassia **1**:263
Korangi Creek **1**:276*T*
Korea
animal species **1**:261*T*
botanical gardens **6**:519
ecosystems
laurel forests **1**:270–271
mountain regions **1**:273
temperate deciduous forests **1**:271
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
vegetation types and coverage percentage **1**:264–265*T*
land-use trends **4**:559*F*
plant species **1**:259*T*
Krakatau archipelago **2**:115–117, **2**:547, **2**:548–549*F*, **3**:118
Krameriaceae **4**:72*F*, **4**:73, **5**:622*F*
Kranz anatomy **2**:364, **2**:364*F*
Krapfia **4**:101
krill **4**:385, **6**:25
Krizna jama (Slovenia) **7**:59*T*
Kruger National Park (South Africa) **1**:60*F*, **2**:289–292, **5**:226–227, **5**:227*F*, **5**:234–236, **6**:486
Kryptolebias marmoratus **3**:467
K-strategists **3**:124–125
kudzu **4**:359–360
Kuerhleiko Hu **1**:276*T*
Kuhnia scombercolias **5**:661*F*
Kuhnia scombri **5**:661*F*
Kuhnia sprostonr **5**:661*F*
Kula Ridge **7**:330
Kulczynski index of similarity **1**:297–298, **1**:298*F*, **1**:299*F*
Kullback-Liebler population assessment **7**:407
Kunzea **1**:267
Kurzamra **4**:101
Kuwait
animal species **1**:261*T*
botanical gardens **6**:507*T*
ecosystems
threatened species **1**:278*T*
vegetation types and coverage percentage **1**:264–265*T*
energy efficiency measures **3**:254*F*
plant species **1**:259*T*
Kuznets curves **3**:25, **3**:28–29, **3**:28*F*
Kyoto Protocol **1**:674, **1**:678, **1**:682–683, **4**:21–22
Kyrgyzstan **6**:507*T*

L

Labiatae **3**:129*T*
Labidura herculeana **3**:222
Labiotermes **4**:409*T*
Labrador duck **7**:182
Labroidei **3**:469
Labroides dimidiatus **2**:141–142, **5**:656
labyrinth fish **3**:470
Labyrinthomorpha **5**:669–670*T*
labyrinthulids **6**:296–297
Labyrnaviridae **5**:134*T*, **5**:138
Lac Alaotra bamboo lemur **6**:256–258*T*
Laccadives Islands **1**:262–263*T*
Laccopetalum **4**:101
lace bugs **7**:293
lace corals **3**:199
Lacerta vivipara **2**:76, **6**:410
Lacessititermes **4**:409*T*
Lacey Act (1900) **2**:281, **3**:713*T*, **3**:715
Lacon fasciatus **7**:220*T*
Lacrymaria **6**:295
lactate dehydrogenase **3**:665
lactation **7**:37
Lactoridaceae **5**:216–217
Lactuca sativa **2**:383, **2**:662–663
Lactuca scariola **6**:433, **6**:434*F*
Lacydoniidae **7**:417*F*
lacZ gene **1**:324
Ladoga (Lake) **4**:461*T*
lady beetles **3**:222, **4**:361
ladybird beetles **5**:316
Lady's Smock **6**:106*F*, **6**:107
Lafoensia pacari **1**:161
Lagarocotyle salamandrae **7**:448–449*F*
Lagarocotylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Lagerstroemia lanceolata **1**:266
Lagidium viscacia **2**:587*F*, **3**:127*T*
Lagomorpha **2**:586, **2**:587*F*, **4**:690*T*, **4**:697, **6**:309*T*
lagoons **7**:369*T*
Lagopus muta hyperborea **7**:27–28
Lagopus mutus **3**:127–128
Lagopus scoticus **4**:578–579
Lagostomus maximus **3**:127, **3**:188
Lagothrix cana cana **6**:256–258*T*
Lagothrix flavicauda **2**:587*F*
Lagothrix lugens **6**:256–258*T*
Laguna de Bay **1**:276*T*
Laguncularia **5**:11*T*
Laguncularia racemosa **1**:164, **3**:116
lake and pond ecosystems **4:458–466**
acidic and mercury deposition **1**:11*F*
Amazon ecosystems **1**:164
ancient lakes species flocks
characteristics **4**:463–464, **4**:463*T*
fish species flocks **4**:463–464
invertebrates **4**:464, **4**:464*T*
anthropogenic impacts
acidification **3**:584–585, **3**:584*F*

acidification and pollution **3**:577–578
agricultural and industrial chemicals **3**:575–576, **3**:580–581, **3**:581*F*
comparison studies **3**:582
ecosystem functions and processes **3**:585
eutrophication **3**:575*F*, **3**:576–577, **3**:576*F*, **3**:581–582
food web manipulation **3**:582–583
habitat alteration **3**:573–574
hydrological cycle changes **3**:574–575
Lake Fure, Denmark **3**:575*F*, **3**:581–582
organic pollution **3**:575, **3**:575*F*, **3**:580–581, **3**:581*F*
research summary **3**:585–586
resource and predatory controls **3**:582–583
restoration and conservation programs **3**:574
Secchi transparency **3**:581–582
species richness **3**:583–584
whole-lake studies **3**:582
biological quality
species declines **3**:578–579, **3**:579*T*
species variability **3**:578–579, **3**:578*F*
threatened species **3**:579–580
characteristics
biota origins and peculiarities **4**:458–459
classifications **4**:458–459
latitudinal gradients **4**:459
species richness–lake size relationship **4**:459–460
species richness–morphometry relationship **4**:460
vertical distribution **4**:459
conservation status **3**:171–173*T*
diversity estimations
fish species **3**:605–606, **3**:606*F*, **3**:607*F*, **4**:460–461, **4**:461*T*
freshwater sediment species **4**:460, **4**:461*T*
general discussion **4**:460
plankton and microorganisms **4**:460
ecosystem services **1**:334*T*
endangered freshwater invertebrates **3:176–187**
anthropogenic impacts
direct harvests **3**:181
global climate change **3**:181–182
habitat destruction/degradation **3**:180–181, **3**:180*F*
introduced species **3**:181
pollution **3**:181, **3**:181*F*
causal factors **3**:177, **3**:177*F*
diversity hot spots **3**:176
global conservation status **3**:176–177, **3**:176*T*, **3**:177*F*
global distribution patterns **3**:182–184, **3**:182*T*
global diversity **3**:176
number of species **3**:182–184, **3**:182*T*
protection strategies **3**:184–187
small ranges
causal factors **3**:178
conservation status **3**:177*F*
dispersal ability **3**:178*F*
high endemic hot spots **3**:178–179
high endemic species **3**:179, **3**:180*F*
Lake Baikal **3**:179, **3**:183*F*
narrowly endemic species **3**:177–178, **3**:177*F*, **3**:179
vulnerability factors
small ranges **3**:177–178, **3**:177*F*
sparse populations **3**:179–180
energy flow **3**:242–243, **3**:243*F*
epilimnion **3**:560, **4**:374*F*
eutrophication impacts
biological consequences
acidification **3**:584–585, **3**:584*F*
comparison studies **3**:582
ecosystem functions and processes **3**:585
ecosystem structure and function **3**:581–582
food web manipulation **3**:582–583
Lake Fure, Denmark **3**:575*F*, **3**:581–582
resource and predatory controls **3**:582–583
Secchi transparency **3**:581–582
species richness **3**:583–584
whole-lake studies **3**:582
clean-up results **5**:24
harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*, **5**:25
long-term human influences **3**:354–356, **3**:356*F*, **3**:357*F*
pollution **4**:465–466
species diversity–lake functioning **4**:462–463
species shifts **3**:352–354, **3**:353*F*
fish species **3**:243*F*, **3**:444
freshwater invertebrates **4:369–378**
background information **4**:369–370
ecological functions and processes
characteristics **4**:375
decomposition and recycling **4**:375
material transfer and uptake **4**:375–376
primary production **4**:376
global patterns **4**:373
influencing factors
disturbances **4**:375
endemism **4**:374–375
habitat heterogeneity **4**:375
isolation factors **4**:374–375
origins **4**:374
research background **4**:373–374
species interactions **4**:375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats
altered water regimes **4**:377–378
biodiversity loss **4**:376–377
exotic species invasions **4**:378
habitat depletion and loss **4**:376–377
pollution **4**:377
reduced water quality **4**:377
number of species **4**:370–371, **4**:371*T*
river continuum concept (RCC) **4**:373*F*
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
freshwater sediments **3**:572*T*
geographic distribution **4**:370*T*
grazer-based ecosystem restoration and management **4**:15–16, **4**:16*F*
habitat diversity **3**:570–571, **3**:573–574
habitat loss and degradation **7**:371*T*
hypolimnion **3**:560, **4**:374*F*
indicator species **4**:255, **4**:256*F*
littoral zone **4**:374*F*
major threats
competition for water **4**:464
fisheries practices **4**:465
habitat alteration **4**:464–465
introduced species **4**:378, **4**:465
pollution **4**:465–466
mean water depth **3**:561*F*
metalimnion **3**:560, **4**:374*F*
nonextremeophilic Archaea **1**:224*T*
number of lakes **3**:561*F*
nutrient enrichment impacts **1**:95–97, **1**:96*F*, **3**:351, **3**:575*F*, **3**:576–577, **3**:576*F*
physical characteristics **3**:560–562
predation effects **6**:233–235
protected area systems **5**:444*T*, **6**:494–498, **6**:496*F*, **6**:497*F*
species diversity–lake functioning
biodiversity–ecosystem stability relationship **4**:462
biological productivity **4**:462–463
characteristics **4**:461–462
eutrophication **4**:462–463
food webs **4**:461–462
intra- and interspecies communication systems **4**:462
top-down controls **4**:462
species richness **3**:572*T*, **3**:583–584
trophic cascades **4**:462, **7**:260–261, **7**:261*F*
tropical Africa **1**:54–55
water storage **3**:560–562, **3**:561*T*
Lake Baikal
ancient lakes species flocks
characteristics **4**:463*T*
hydrothermal vents **4**:459
invertebrates **4**:464, **4**:464*T*
ecosystem characteristics **1**:262–263*T*, **1**:279
endemic species **3**:179, **3**:183*F*, **4**:374–375
fish species **3**:444
photosynthesis **3**:560
Lake Fure, Denmark **3**:575*F*, **3**:581–582
Lake Gatun, Panama **6**:233–235
Lake Hula **5**:472–473, **5**:474*F*, **5**:475*T*
Lake Ladoga ringed seal **5**:81*T*
Lake Leonard **1**:276*T*
Lake Nakaumi **1**:276*T*
Lake Okeechobee **7**:372, **7**:372*F*, **7**:373*F*
Lake Saimaa ringed seal **5**:77*T*
Lake Shinji **1**:276*T*
Lake Tanganyika **1**:54–55
lake trout **1**:10*T*, **1**:12*F*, **6**:153*F*
Lake Victoria **1**:54–55, **4**:463–464, **4**:463*T*, **4**:465–466, **4**:663

Lama guanicoe **2**:476*T*, **2**:587*F*, **3**:115, **6**:334
Lama lama **2**:476*T*
Lama paco **2**:476*T*
Lama quanicoe **3**:127*T*
Lamarck, Jean-Baptiste **2**:422, **2**:424, **2**:638, **3**:393
Lambeosaurus **2**:519*T*, **2**:520*T*
Lamellibrachia sp. **7**:328–329
Lamiaceae **3**:129*T*
Lamiales **2**:374*T*
Laminaria augusta **7**:230
Laminariales **4**:345–346, **4**:346*F*
lammergeiers **1**:709
lampreys
  characteristics **3**:456–457, **3**:457–458, **3**:458*F*, **7**:337–338
  migratory patterns **5**:314–315
  river ecosystems **6**:479
Lampriomorpha **3**:466, **3**:466*F*
lamp shells **4**:381*T*
Lampyridae **1**:327*T*
Lamyctes adisi **5**:431
Lamyctes africanus **5**:427–428
Lamyctes coeculus **5**:425–426
Lamyctes emarginatus **5**:427–428
Lanao (Lake) **4**:464
Lance Formation (Canada) **2**:518, **2**:520*T*
lancelets
  characteristics **3**:457, **7**:336
  classification systems **7**:335, **7**:335*F*, **7**:336*F*
lancetfish **3**:465–466
land, arable **6**:592*F*
land-bridge islands **5**:436
land clearing *see* deforestation
land cover
  land-cover pattern mapping **6**:396–397, **6**:396*F*
  species distribution modeling (SDM) **6**:699–700
land degradation
  dryland ecosystems **2**:479, **2**:481*F*
  major degradation processes **6**:592*F*
  Nepal **6**:378–379
  Succulent Karoo **1**:489
land equivalent ratio (LER) **2**:386
Lande, Russ **2**:264–265
LANDFIRE project **2**:575*F*
landlocked salmon **1**:12*F*
land mammals **5**:565
  see also Mammalia/mammals
landraces
  biodiversity conservation **1**:122–123, **3**:34, **3**:35*F*
  crop genetic diversity
    basic concepts **1**:87–89
    commercialization effects
      general discussion **1**:91–92
      intensified environments **1**:92
      non-intensified environments **1**:92–93
    crop genetic resources **3**:39–40
    genetic narrowing **1**:89–90
    genetic trends **1**:90
    landrace–modern varieties comparisons **1**:89
    research summary **1**:93
    spatial distribution **1**:87–89, **1**:88*T*
  crop production **1**:85–86
  genetic evolution **1**:86–87
  planned agrobiodiversity **3**:31–34
  regional cultivars **3**:32*F*
  Vavilov centers **4**:141
Landsat program **4**:567, **6**:393, **6**:393*T*
landscape aesthetics **1**:41
landscape corridors **4:467–475**
  characteristics **4**:467
  connectivity effects **4**:467–468
  conservation biology **2**:242–243
  effectiveness measures **4**:468
  empirical research evidence
    experimental studies
      functional role **4**:469–471
      landscape experiments **4**:471–473
      microcosm/mesocosm studies **4**:470–471, **4**:470*F*
      patch-based studies **4**:471–473, **4**:472*F*
      plot-based studies **4**:470*F*, **4**:471
    observational field studies **4**:469
  future outlook **4**:473–474
  historical perspective **4**:467, **4**:468*F*
  research summary **4**:473
  simulation models **4**:468–469
landscape diversity **4:476–487**
  agrobiodiversity **1**:133
  agroecosystems
    ecological consequences **3**:13
    general discussion **3**:13
    organism spillover **3**:13–14
    spatial structure **3**:14–15, **3**:14*F*
  basic concepts **4**:476, **6**:534
  biodiversity-friendly farming
    general discussion **2**:544
    landscape connectivity **2**:544–545
    landscape heterogeneity **2**:544
  biodiversity planning and management **4**:483
  biofuel production
    landscape–scale biodiversity considerations
      agroenergy sprawl **1**:534–535
      general discussion **1**:533–534
      spatial considerations **1**:533–534
  control mechanisms
    disturbances **4**:479–480
    environmental heterogeneity **4**:480
    fragmentation **4**:480
    patch dynamics **4**:479–480
  environmental impacts **3**:283–284
  functionality **4**:480
  grain and extent **4**:476, **4**:477*F*
  landscape dynamics and change **4**:480–483, **4**:481*F*, **4**:482*F*, **4**:482*T*, **4**:496–497
  multifunctional landscapes **2**:300–301
  nature-based coastal protection strategies **5**:455
  patch dynamics **4**:479–480, **4**:492–493
  planned versus unplanned biodiversity **1**:127*F*, **3**:9, **3**:31–34
  research summary **4**:486
  Sevilleta National Wildlife Refuge (SNWR) case study **4**:483–486, **4**:484*F*, **4**:485*F*
  simulation models **4**:482, **4**:488–491, **4**:490*F*, **4**:491*F*
  structure
    characteristics **4**:476–478
    landscape composition **4**:477*F*, **4**:478–479
    landscape patterns **4**:478–479
    patch identification and description **4**:476–478, **4**:477*F*, **4**:478*F*
    patch–matrix–corridor model **4**:479, **4**:479*F*, **4**:492–493, **4**:492*F*
landscape ecology **4:488–502**
  agricultural landscapes
    characteristics **1**:380–381
    habitat heterogeneity **1**:381–382, **1**:381*F*
    natural habitat proximity **1**:381
    noncrop habitats **1**:380–381
  biodiversity dynamics modeling **5:321–328**
    background information **5**:321–322
    phenomenological models
      dose–response models **5**:324, **5**:324*F*
      global species distributions **5**:324
      species–area relationships **5**:322–323
      statistical estimation models **5**:324
    process-based models
      integro-difference equations **5**:324–325
      multispecies metapopulation analysis **5**:325–326, **5**:325*F*, **5**:326*F*
      neutral theory **5**:326–327
      reaction–diffusion models **5**:324–325, **5**:325*F*
    research summary **5**:327
    species–area relationships
      countryside habitats **5**:323–324, **5**:323*F*
      native habitats **5**:322–323, **5**:322*F*
  Development by Design (DbD) **2**:498
  fire/fire regimes **3**:440
landscape genetics **4:508–523**
  analytical frameworks
    assignment tests **4**:511
    barrier detection **4**:511, **4**:512*F*
    graph theory/network analyses **4**:512–514, **4**:513*F*
    Mantel tests **4**:510–511
    ordination and regression analysis **4**:514
    resistance surfaces **4**:511–512, **4**:513*F*
  isolation-by-distance (IBD) models **4**:508, **4**:509*F*
  research background **4**:508–509
  research challenges
    genetic distance (F_{ST}) measures **4**:518–519
    interdisciplinary communication **4**:517–518
    landscape genomics **4**:519–520
    multispecies comparative studies **4**:518
    spatial analysis method (SAM) **4**:519

spatial replication 4:518
spatial scale effects 4:518
research results
anthropogenic impacts 4:516–517
barrier identification 4:515–516
gene flow barriers 4:515–516
general research trends 4:514–515
landscape variables influences 4:516
sampling strategy and design 4:509–510
spatial patterns and connectivity 4:508–509
landscape legacies **4:524–530**
driving forces
land use 4:524–525
natural disturbances 4:525
ecosystem processes 4:527–528
plant communities 4:525–527
research implications 4:528
site history 4:524
soils 4:527
summary discussion 4:528
nutrient management–water quality relationship **1**:107–108
population dynamics
basic concepts 4:488–491
continuous habitat variability 4:497–498, 4:498*F*
core and satellite model 4:496
disturbance recovery 4:498–499, 4:499*F*
edge effects 4:496–497
environmental gradients 4:497–498
extinction debt 4:497
habitat fragmentation 4:496–497
habitat heterogeneity **1**:381–382, **1**:381*F*, 4:497–498
landscape changes 4:496–497
landscape structure
basic concepts 4:491–493
habitat preservation 4:495
immigration 4:495
interspecific interactions 4:498–499, 4:499*F*
landscape context 4:493
patch–matrix–corridor model 4:479, 4:479*F*, 4:492–493, 4:492*F*
population viability analysis (PVA) 4:493–495
quantification methods 4:493
single large or several small (SLOSS) debate 4:495
metapopulation dynamics 4:495–496
research summary 4:499–500
simulation models 4:488–491, 4:490*F*, 4:491*F*
soil moisture prediction 4:497–498, 4:498*F*
source–sink dynamics 4:496, **5**:324–325
spatial patterns 4:488–491, 4:489*F*
wetland ecosystems **7**:399
wildlife management **7**:405–406
landscape epidemiology **4:503–507**
background information 4:503
dynamic spatiotemporal mapping approaches 4:503
future research areas
genetic studies 4:505–506
geographic profiling 4:506
mapless approaches 4:505
research summary 4:506
static mapping approaches
basic concepts 4:503–505, 4:504*F*
host distribution data 4:505
incidence data 4:504–505
normalized difference vegetation index (NDVI) 4:504
vector distribution data **1**:361–362, **1**:363*F*, 4:505
landscape genetics **4:508–523**
analytical frameworks
assignment tests 4:511
barrier detection 4:511, 4:512*F*
graph theory/network analyses 4:512–514, 4:513*F*
Mantel tests 4:510–511
ordination and regression analysis 4:514
resistance surfaces 4:511–512, 4:513*F*
isolation-by-distance (IBD) models 4:508, 4:509*F*
research background 4:508–509
research challenges
genetic distance (F_{ST}) measures 4:518–519
interdisciplinary communication 4:517–518
landscape genomics 4:519–520
multispecies comparative studies 4:518
spatial analysis method (SAM) 4:519
spatial replication 4:518
spatial scale effects 4:518
research results
anthropogenic impacts 4:516–517
barrier identification 4:515–516
gene flow barriers 4:515–516
general research trends 4:514–515
landscape variables influences 4:516
sampling strategy and design 4:509–510
spatial patterns and connectivity 4:508–509
landscape models **4:531–538**
basic concepts 4:531–533
functional landscapes 4:531–533
modeling approaches
functional role 4:533–535
hybrid/multimodeling models 4:534*T*, 4:535–536
neutral models 4:534*T*, 4:536
process models 4:534*T*, 4:535, 4:535*F*
state-and-transition models 4:533–535, 4:534*F*, 4:534*T*
research summary 4:536–537
landscape mosaics **7**:66–68, **7**:67*F*
landscape patchiness management **7**:232–233
land snails **1**:33, **3**:222, **3**:226, **3**:674*F*, **5**:460*T*, **5**:464*T*, **5**:465*F*
land-use change
agricultural practices 4:556–558
biofuels
direct land-use change (DLUC) 4:293–294
indirect land-use change (ILUC) **4:293–297**
avoidance measures 4:296–297
biomass emission estimation 4:295–296
carbon dioxide (CO_2) emissions 4:294–295
economic models 4:294–295, 4:294*F*, 4:295*F*
estimation uncertainties 4:296
food price increases and rationing 4:296
global market impacts 4:294–295, 4:295*F*
greenhouse gas emissions 4:293–294
land cover change emissions 4:295–296
land cover maps 4:295
marginal land usage 4:296–297
total emission–fuel quantity relationship 4:296
C_4 plants **2**:378–379, **2**:379*F*
carbon dioxide (CO_2) emissions **4:539–554**
Global Trade Analysis Project (GTAP) model 4:539–540
consumption-side structure 4:540*F*
economic and environmental impacts 4:540–544
emission factors 4:548–549
estimated emissions 4:549–550, 4:549*T*
final analysis 4:551–552, 4:551*T*
land cover and land-use activities 4:542*F*
livestock production structure 4:542*F*, 4:543*F*, 4:543*T*
marginal and average estimated emissions (first simulated group) 4:550, 4:550*T*
marginal and average estimated emissions (second simulated group) 4:550–551, 4:550*T*
marginal and average estimated emissions (third simulated group) 4:551, 4:551*T*
modeling process 4:539–540
production-side structure 4:541*F*
productivity measures 4:540–544
research limitations 4:553
research summary 4:552–553
simulation models 4:544–545
2001 simulation model 4:544–545, 4:544*T*, 4:545*T*
2001–2006 simulation model 4:545–547, 4:546*T*, 4:547*T*
2006–2015 simulation model 4:547–548, 4:548*T*, 4:549*T*
well-to-wheel emissions 4:551–552, 4:552*T*, 4:553*T*
soil and vegetation fluxes **1**:284
Convention on Biological Diversity (1930) **2**:317–318
current global land cover 4:557*F*, 4:566*F*
ecosystem boundaries
analytical studies **3**:241–242

land-use change (*continued*)
basic concepts **3**:238–239, **3**:238*F*
control mechanisms **3**:241, **3**:241*F*
food web–trophic level relationship **3**:240–241, **3**:241*F*
historical perspective **3**:238–239
trophic dynamic concept **3**:239–240
environmental impacts
croplands and pasturelands **1**:542–543, **1**:542*F*
global population growth **4**:555–558, **4**:556*F*
marginal, degraded, and abandoned lands **1**:543, **1**:544*F*
native ecosystems **1**:541–542, **1**:545*F*
prime agricultural lands **1**:543
global cropland area **4**:558*F*
global population growth **4**:555–558, **4**:556*F*, **4**:558–560
land-use trends **4**:559*F*
nutrient management–water quality relationship **1**:107–108
overgrazing **4**:560*F*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
land use issues **4:555–568**
agricultural practices **4**:556–558
anthropogenic impacts
historical perspective **4**:560–562
human settlement **4**:558–560, **4**:560*F*, **4**:561*F*
biodiversity impacts
assessment measures **4**:562–564
GLOBIO model **4**:562–564
habitat conversion
Southeast Asia deforestation **4**:562–564, **4**:563*F*
urbanization (Western Washington State) **4**:564
habitat degradation
agricultural intensification effects **4**:565–566
land cover **4**:564–566
nature-based recreation **4**:566–567
timber industry **4**:567
habitat fragmentation **4**:564, **4**:565*F*, **4**:566*F*
historical trends **4**:562*F*
current global land cover **4**:557*F*, **4**:566*F*
ecosystem function measurements **3**:79–80
environmental impacts **4**:555–558
estimation methods
land cover **4**:567
limitations **4**:567–568
general discussion **4**:555
global cropland area **4**:558*F*
global population growth **4**:555–558, **4**:556*F*, **4**:558–560
historical perspective **4**:560–562
human settlement **4**:558–560, **4**:560*F*, **4**:561*F*
land-use trends **4**:559*F*
natural resource extraction **4**:558
overgrazing **4**:560*F*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
Land Use, Land-Use Change and Forestry (LULUCF) **7**:139–140
land-use patterns **4:569–582**
anthropogenic impacts **4**:569–570
biodiversity dynamics modeling **5:321–328**
background information **5**:321–322
phenomenological models
dose–response models **5**:324, **5**:324*F*
global species distributions **5**:324
species–area relationships **5**:322–323
statistical estimation models **5**:324
process-based models
integro-difference equations **5**:324–325
multispecies metapopulation analysis **5**:325–326, **5**:325*F*, **5**:326*F*
neutral theory **5**:326–327
reaction–diffusion models **5**:324–325, **5**:325*F*
research summary **5**:327
species–area relationships
countryside habitats **5**:323–324, **5**:323*F*
native habitats **5**:322–323, **5**:322*F*
buildings/built-up areas **4**:579, **4**:580*F*
climate change effects **2**:104, **4**:569–570
conservation implications **4**:581–582
coppice woods **4**:573, **4**:573*F*, **4**:574*F*
cost–benefit analyses **6**:265–266, **6**:266*F*
crop plants **4**:570, **4**:570*F*
domestic animals **4**:570
economic factors **6**:269*F*
fields **4**:570–572, **4**:571*F*
fire cycles **4**:580, **4**:581*F*
forests and woodlands **4**:572–574, **4**:573*F*, **4**:574*F*
grasslands **4**:577
heathland **4**:577, **4**:578*F*
hedges **4**:571, **4**:571*F*
historical conservation practices **2**:279–280
historical perspective **4**:569, **4**:579–581
landscape legacies **4**:524–525
moorland **4**:577–579, **4**:578*F*
people–ecosystem relationship **6**:223–224, **6**:224*T*
permanent edges and open areas **4**:573–574, **4**:575*F*
systematic conservation planning **6**:268*F*
terraces **4**:572, **4**:572*F*
wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*, **4**:581*F*
language
biodiversity conservation **4**:270–271, **4**:270*T*, **6**:566
extinctions **3**:286
literature–science relationship **4**:649–650
langurs **4**:692, **6**:251*T*
Laniarius mufumbiri **1**:53–54
Laniatores **1**:210, **1**:210–211, **1**:246*T*
La Niña events *see* El Niño Southern Oscillation (ENSO)
Lanius collurio **5**:312
Lanius nubicus **5**:474
lanner falcon **5**:474
Lantana camara
India **1**:82
negative impacts **4**:352–353, **4**:353*F*
pollination and dispersal **4**:364
South Africa **1**:81
United States **1**:77–78
lanternfish **3**:466
Laonasles aenigmamus **4**:704
Laos
animal species **1**:261*T*
botanical gardens **6**:519
ecosystems
genetic richness **1**:514*T*
threatened species **1**:278*T*
tropical moist forest **1**:277*T*
vegetation types and coverage percentage **1**:264–265*T*
endangered mammals **3**:194*T*
indigenous peoples **4**:270
mammals **4**:703–704
plant species **1**:259*T*
rice production **1**:88
urbanization **4**:560*F*
see also Southeast Asia
Laotian black crested gibbon **6**:256–258*T*
Lapham, Increase **2**:281
La Plata dolphins **5**:65*T*, **5**:77*T*
Laponia National Park (Sweden) **6**:486
lapper-faced vulture **5**:474, **7**:185
Laptev Sea walrus **5**:81*T*
Laramide Orogeny **5**:563
larch
boreal forests **1**:628–630
geographic boundaries **1**:629*F*, **1**:630*T*
restoration programs **6**:125
larch sawfly **4**:182–183
Laretia acaulis **4**:105–106
large-headed capuchin **6**:254
largehead hairtail **6**:422*T*
large-mouth bass **1**:10*T*, **1**:12*F*, **3**:123*T*, **3**:365*F*, **7**:261
large-scale biodiversity experiments **4:583–589**
background and characteristics **4**:583–584
community assembly
Biocon experiment **4**:585–586, **4**:585*F*
Ecotron large-scale controlled environment facility **4**:585
highly replicated laboratory experiments **4**:587, **4**:587*F*
historical perspective **4**:584–585, **4**:585*F*
large-scale field experiments **4**:585–586, **4**:585*F*, **4**:586*F*
multisite biodiversity experiments **4**:586–587
research applications
general discussion **4**:587–588
grassland restoration **4**:588, **4**:588*F*
tree diversity experiments **4**:588–589
species removal **4**:587, **4**:588*F*
underlying mechanisms **4**:584
Laris atricilla **3**:256

Larix
boreal forests **1**:628–630
geographic boundaries **1**:629*F*, **1**:630*T*
restoration programs **6**:125
Larix dahurica **1**:269, **1**:273–274, **1**:630*T*
Larix gmelini **1**:272–273
Larix laricina **1**:630*T*, **7**:400
Larix leptolepis **1**:273
Larix potaninii **1**:272–273
Larix sibirica **1**:273–274, **1**:630*T*
Larix sukachzewski **1**:630*T*
larks **1**:710
Larrea tridentata **2**:467*F*, **2**:468–469, **4**:485–486
larvaceans **6**:25
larval parasites **5**:657
larval ticks **1**:361–362, **1**:363*F*
Las Cruces Fenced Marine Coastal Preserve (Chile) **5**:56–57, **5**:58*F*
Las Gemelas Seamounts **5**:449–450
Lasiocampidae **5**:396
Lasiocampoidea **5**:387–388*T*
Lasiognathus amphirhamphus **3**:475*F*
Lasiorhinus krefftii **4**:689
Lasiurus cinereus **5**:313
Lasiurus hirsutus **1**:69
Last Glacial Maximum (LGM) **2**:466, **2**:682
latchets **2**:8–9
late Quaternary extinction event (LQEE) **4:721–732**
background information **4**:721
causal factors
anthropogenic models **4**:731–732
climatic models **4**:732
global patterns **4**:731–732
comparative studies
Africa **4**:658*F*, **4**:729–730
analytical methods **4**:721–725
Australia **4**:658*F*, **4**:727–728
Eurasia **4**:729–730, **4**:729*F*
Madagascar **4**:658*F*, **4**:728–729
North America **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
Oceania **4**:730–731
South America **4**:726–727, **4**:727*T*
time-transgressive sequence **4**:722*F*, **4**:723*F*
Lates niloticus **1**:54–55, **2**:223–224, **3**:468, **4**:360–361, **4**:463, **4**:663
Latimeria **2**:47
Latimeria chalumnae **3**:461, **3**:478–479
Latimeria manadoensis **3**:461
Latin America
biodiversity loss **7**:352–353
C_4 plants **2**:371*T*
crop distribution **1**:88*T*
ecosystems **3:105–117**
background information **3**:105
biogeography
early history **3**:109
late Tertiary **3**:109
Quaternary period **3**:109–110
geographic distribution **3**:111*F*
high Andean shrublands **3**:111*F*, **3**:114, **3**:115
mangrove ecosystems **3**:111*F*, **3**:116
physical characteristics
climate **3**:107–109, **3**:108*F*
geography **3**:105–107
geology and geomorphology **3**:105–107, **3**:106*F*
tectonic history **3**:107
plant diversity **6**:314–315
river water composition **3**:562*T*
temperate forests **3**:111*F*, **3**:117
temperate grasslands
distribution and structure **3**:111*F*, **3**:113–114
functional aspects **3**:114–115
tropical forests
distribution and structure **3**:110
functional aspects **3**:112
general patterns **3**:111–112
mountain cloud forests **3**:111
tropical rain forests **3**:110, **3**:111*F*, **3**:112
tropical semideciduous/dry forests **3**:110–111, **3**:112–113
tropical rain forests **6**:317*T*
tropical savannas
distribution and structure **3**:111*F*, **3**:113
functional aspects **3**:114
wetlands **3**:116–117
xeric formations
distribution and structure **3**:111*F*, **3**:115–116
functional aspects **3**:116
ecosystem services
payment for ecosystem services (PES) **7**:355
water funds **7**:354–355*F*, **7**:356*T*
watershed management
best management practices **7**:358–359, **7**:358*T*
compensating impacts **7**:359
public protected area management **7**:358, **7**:358*T*
endangered mammals **3**:194*T*
endemism **4**:271*T*
energy use **3**:252*T*
floodplain agriculture **7**:146
indigenous peoples **4**:269*T*, **4**:270, **4**:273*F*, **4**:273*T*, **4**:274*F*
intercropping **2**:382–383
land-use trends **4**:559*F*
maize production **1**:86, **1**:88
mammalian extinctions **4**:705*T*
multiple cropping systems **1**:119–120
myriapods **5**:429*T*
natural history museums and institutions **5**:405–406
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
nitrogen deposition **5**:526*T*
oil-palm plantations **5**:602
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
species richness **4**:271*T*
transboundary protected areas **6**:489*T*
wheat production **1**:87–88
see also Central America; South America
latitudinal gradients **4:612–626**
benthic large-scale diversity patterns **4**:385–387, **4**:386*F*, **4**:387*F*
butterflies **1**:657, **1**:657*T*, **1**:658*F*
endemism **3**:230–231, **3**:230*F*
forest ecosystems **3**:530, **3**:530*F*
lake and pond ecosystems **4**:459
microbial biogeography **5**:272–273
parasites
host ranges/host specificity **5**:668–670
reproductive strategies **5**:668
species richness **5**:668, **5**:671*F*
southern Andean steppes **4**:106–107
species richness
climatic regions **4**:614*F*
data assessment and synthesis **4**:624–625, **4**:624*F*
ecological communities **4**:615–618, **4**:615*F*, **4**:616*F*, **4**:617*F*
fossil species **4**:614*F*
general discussion **4**:612
geographic distribution **4**:613*F*
gradient patterns
assemblage characteristics **4**:614–615
assessment considerations **4**:614–615
ecological communities **4**:615–618, **4**:616*F*, **4**:617*F*
hierarchical configuration **4**:618
multiple dimensions **4**:618–620, **4**:619*F*
historical perspective **4**:612–614
influencing factors **4**:617*T*
mammals **4**:613*F*, **4**:615*F*, **4**:700–702
mechanism hypotheses
contributing factors **4**:620–621
evolutionary speed **4**:621–622
geographic area hypothesis **4**:620–621, **4**:621*F*
geometric constraints hypothesis **4**:622–624, **4**:623*F*
Rapoport–rescue hypothesis **4**:622
parasites **5**:668, **5**:671*F*
plant biodiversity–climate relationship **2**:118–119, **3**:119
taxonomic dimensions **4**:616*F*
terrestrial plant communities **6**:72–73
Latvia
botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
laughing gull **3**:256
Laupala **1**:32–33
Lauraceae **1**:266*T*, **1**:267, **1**:270–271, **4**:71, **4**:72*F*, **5**:621*F*, **6**:315*T*, **6**:316*T*
Laurales **5**:618*F*
Laurasia **1**:169–170, **7**:344
Laurasiatheria **4**:685
laurel forests **1**:270–271
Laurelia **7**:119
Laurelia **5**:621*F*
Laurelia philippiana **3**:117
Laurencia poiteaui **3**:359
Laurus **5**:218
Lauxaniidae **3**:489–491*T*
Lauxanioidea **3**:489–491*T*
lava tubes **7**:50–51, **7**:60
Lawas mangroves **1**:276*T*

law of requisite variety **2**:155–156
Laysan finches **1**:29*F*
Lazarus species **5**:164–165, **5**:436
Leaders in Environmental Action for the Future (LEAF) program **7**:253
leaf beetles **1**:327*T*, **1**:328
leaf-cutting ants **4**:181
leaf feeders **7**:266*F*, **7**:267*T*
leaf-footed bugs **7**:294
leaf grasshoppers **3**:735–736
leafhoppers **1**:120
leaf-litter ants **1**:612, **1**:613*F*
leaf-litter skinks **3**:144–145
leafminer flies **3**:489–491*T*, **3**:492
leaf miners **1**:328, **5**:385–386, **5**:392–394, **6**:40–41
leaf monkeys **4**:692, **5**:14
leaf-rolling crickets **3**:732
leaf toughness **2**:428
leafy spurge **1**:76
learning opportunities **3:133–141**
  college and university programs **3**:140
  environmental education **3**:136, **3**:136
  government-based initiatives **3**:133
  informal education experiences
    biodiversity experiences and resources
      Earthwatch Institute **3**:138
      National Biological Information Infrastructure (NBII) **3**:138–139
      nonprofit environmental groups **3**:139
      online resources **3**:139
      resources **3**:138
    botanical gardens **3**:138
    museums **3**:138
    resources **3**:136–137
    youth-serving groups **3**:137, **3**:137
    zoos **3**:137–138
  K–12 education
    diversity of life concepts **3**:134
    formal science education **3**:135–136
    grades 3–5 **3**:134
    grades 6–8 **3**:134–135
    grades 9–12 **3**:135
    grades K–2 **3**:133–134
    Project 2061 **3**:133–134
  public awareness surveys **3**:139–140
  research summary **3**:140–141
  systematics research and training **3**:140
Lear's macaw **7**:187
least-cost path analysis **4**:509*F*, **4**:512
leatherback sea turtles **2**:9*F*, **3**:217, **3**:218, **3**:256
leaves
  geographic distribution **3**:131*T*
  global food usage **3**:130*T*
Lebanon
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  ecosystems
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
Lebellus de re Herbaria Nova **5**:705
Lecanicephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Lecanora tartarea **1**:591
Lechenaultia **5**:221
Lecithoepitheliata **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
Lecythidaceae **6**:316*T*
Ledum **3**:76*F*
Leea **5**:622*F*
leeches **1**:359, **4**:371*T*, **4**:374*F*
leeks
  intercropping patterns **2**:383
  native crops **2**:662–663
legumes
  intercropping systems **2**:389–390
  nitrogen fixation **1**:560–561, **1**:562*F*
Leguminosae **1**:266–267, **3**:129*T*, **3**:130*T*, **6**:313–318, **6**:316*T*
Leidy's comb jelly **3**:284
Leiobunum spp. **1**:210
Leiognathus dussumieri **3**:478*F*
Leiopleura **1**:329
leishmaniasis **5**:671*T*
Leiurus spp. **1**:208
lemmings
  migratory patterns **5**:313–314
  tundra ecosystems **6**:330–331
Lemmus sibiricus **1**:232
Lemmus spp. **6**:331
Lemuridae **6**:250–252, **6**:251*T*
lemurs **4**:690*T*, **4**:692–693, **6**:250–252, **6**:251*T*
lenacil **5**:696
Lennoaceae **4**:72*F*, **4**:73
Lens culinaris **2**:662–663
Lentibulariaceae **6**:44*T*
lentils
  native crops **2**:662–663
  Neolithic agriculture **2**:658
  regional cultivars **3**:32*F*
Lentinula edodes **3**:637*T*, **3**:638
Leonard (Lake) **1**:276*T*
Leontopiihecus rosalia **3**:196
Leontopithecus caissara **6**:256–258*T*
Leontopithecus chrysomelas **6**:256–258*T*
Leontopithecus chrysopygus **6**:256–258*T*
Leontopithecus rosalia **2**:273, **4**:247*F*, **6**:256–258*T*, **6**:560
Leontopodium alpinum **1**:154
leopard cat **1**:517
leopard frog **1**:10*T*, **6**:153*F*
leopard lizards **6**:406
leopards **1**:686, **1**:709, **4**:693, **5**:472, **5**:603
leopard seals **5**:90
Leopold, Aldo **1**:112, **2**:285, **3**:99, **4**:121, **4**:652–653, **6**:77–78, **6**:473–474, **7**:403–405
Leopoldinia piassaba **1**:163
Leopoldinia pulchra **1**:163
Lepanthes **6**:316*T*
Lepeophtheirus salmonis **2**:532
Lepes europaeus **1**:80
Lepidochelys kempii **3**:217, **3**:256
Lepidodactylus **1**:591
Lepidoptera
  attraction to light **5**:400–401, **5**:401*F*
  biogeography **5**:389–390, **5**:389*T*
  celestial navigation **5**:400–401
  characteristics **4**:392, **4**:393*F*, **5**:384–385
  chemical defense **5**:396–397, **5**:397*F*
  classifications
    Ditrysia **5**:386–389, **5**:386*F*
    Glossata **5**:385–386, **5**:386*F*
    Neolepidoptera **5**:386, **5**:386*F*
    phylogenetic relationships **5**:385, **5**:386*F*
    suborders **5**:385, **5**:386*F*, **5**:387–388*T*
  coloration **5**:400–401
  conservation programs **5**:402–403
  defense mechanisms **5**:391*F*, **5**:392*F*, **5**:396–397, **5**:396*F*
  dispersal and migration **5**:397
  diurnality and nocturnality **5**:400–401
  feeding behaviors **1**:328–329
  feeding biology
    adults **5**:394–395, **5**:395*F*
    larvae **5**:392–394
  fossil record **5**:389
  grazing effects **1**:709–710
  importance **5**:401–402
  insecticide resistance **4**:298–299
  key adaptations **5**:397–398, **5**:398*F*
  life history
    cocoons **5**:391–392
    diapause/quiescence **5**:392
    eggs **5**:390
    hypermetamorphosis **5**:391
    larvae **5**:388*F*, **5**:390–391
    pupa stage **5**:391–392, **5**:393*F*
  major divisions and characteristics **4**:308–312, **4**:309*T*
  natural enemies **5**:395–396
  number of species **4**:371*T*, **4**:392*T*
  pheromones **5**:398–400, **5**:399*F*, **5**:400*F*
  pollination **6**:141–142, **6**:141*T*
  quantity estimates **1**:326–328
  Southeast Asia **1**:514–515
  species declines **5**:402–403
  species diversity **5**:387–388*T*, **5**:389–390, **5**:389*T*
Lepidopteris **4**:226, **4**:227*F*
Lepidosauria
  characteristics **6**:401, **7**:340–341
  conservation status **6**:411
  Squamata **6**:401–402, **6**:406–408, **7**:340–341
  tuatara
    characteristics **6**:401, **6**:406
    conservation status **3**:216*T*, **3**:217–218, **6**:411
    historical extinctions **5**:367
Lepidotermes **4**:409*T*
Lepilemur ankaranensis **6**:256–258*T*
Lepilemuridae **6**:250–252, **6**:251*T*
Lepilemur septentrionalis **6**:250, **6**:256–258*T*
Lepocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Lepomis auritus **4**:517
Lepomis macrochirus **4**:465
Leporidae **2**:583*F*, **4**:697, **6**:309*T*
leptin **7**:31
Leptinotarsa decemlineata **4**:307

Leptoceratops **2**:518, **2**:519*T*, **2**:520*T*
Leptodirus hochenwartii **7**:52–53, **7**:52*F*
Leptodora **6**:26
Leptodora kindtii **1**:10*T*
Leptomyxa **6**:291
Leptonychotes weddellii **5**:90
Leptopodomorpha **7**:293
Leptopternis **5**:469
Leptorhychos gatesii **3**:208
Leptosols **2**:461–464
Leptospermum **1**:267
Leptostraca **2**:416–417*T*
Leptura quadrifasciata **7**:218–219
Lepuropet **5**:623*F*
Lepus americanus **1**:631–632, **4**:697
Lepus californicus **2**:473–474
Lepus europeaus **3**:127–128, **3**:127*T*
Lepus timidus **1**:631–632
Lepus timidus scotius **7**:28*F*
Lepus townsendii **2**:223
leshmaniasis **7**:302–303
lesions **3**:633*T*
Leslie matrix model **6**:176
Lesotho
botanical gardens **6**:519
energy efficiency measures **3**:254*F*
Lespedeza bicolor **1**:271–272
lesquerella **1**:529*T*
less-developed countries **3**:315–316, **3**:316*F*
lesser bilby **4**:363–364
lesser dung flies **3**:489–491*T*, **3**:493
lesser galagos **6**:251*T*
lesser kestrels **5**:474
lesser kudu **5**:235
lesser mass extinctions
Cenozoic **5**:176
Mesozoic **5**:169*F*, **5**:176
Neoproterozoic **5**:168*F*, **5**:175–176
Paleozoic **5**:175–176
significance **5**:175–176
Lesser Sunda Islands **1**:262–263*T*
Lessonia nigrescens **4**:517
Lestodon **4**:727*T*
Letharia vulpina **3**:636–637
lettuce
intercropping patterns **2**:383
native crops **2**:662–663
Leucadendron **1**:47–48
Leucaena esculenta **1**:120
Leucania **5**:398, **5**:398*F*
Leucheria leontopodioides **4**:106–107
Leucochloridium macrostomum **5**:658–659
Leucochloridium vogtianum **7**:450–451*F*
Leucopeza semperi **7**:184
Leucopogon **5**:221
Leucopsar rothschildi **7**:187
Levantina caesareana **3**:674*F*
Levant region **5**:459, **5**:459*F*, **5**:460*T*
see also Middle East
Levin's population model **2**:603
Lewis, Meriwether **4**:651–652
Leymus chinensis **5**:531
Liabum megacephalum **1**:167
lianas **1**:162, **3**:521
Liberia **6**:519
Libya **4**:560*F*, **6**:519
Libycosaurus **5**:224*T*
Libytheinae **1**:654
Libytherium **5**:224*T*
Licania **5**:623*F*
Licania heteromorpha **1**:159–160
lice **2**:524*T*, **3**:226, **4**:298–299, **4**:392*T*
Lichenostomus frenatus **2**:75*F*
lichens
air pollution impacts **1**:141–142, **1**:142*F*
air pollution monitors **1**:591
alpine ecosystems **1**:153–154
Antarctica **1**:184*T*
anthropogenic impacts **7**:399
Arctic ecosystems **1**:231*T*
boreal forests **1**:628–630
evolutionary development **2**:644
forest canopies
characteristics **3**:517–518
taxonomic distributions **3**:520, **3**:521*T*
genetic diversity **3**:674*F*
indicator species **7**:206–207*F*
latitudinal range shifts **4**:604*T*
litmus paper **1**:591
modern extinctions **5**:368*T*
mutualism **3**:636–637, **3**:637*F*, **6**:719
nitrogen deposition **5**:534*F*
nomenclature code **5**:555
nonseed gene banks **3**:652*T*
pollution-based habitat degradation **3**:407
threatened species **7**:206–207*F*
lichen woodlands **1**:630
licuri palms **7**:187–188
Liebig, Justice **3**:91–92, **6**:111
Liebig's law of minimum **2**:676, **3**:91–92
Liechtenstein
botanical gardens **6**:507*T*, **6**:519
vegetation surveys **3**:341–344*T*
life cycle analysis (LCA) **4:627–630**
aquaculture **1**:196–198, **1**:197*F*
basic concepts **4**:627
historical perspective **4**:628–629
methodologies **4**:627–628
policy and regulation applications **4**:629
standardization **4**:628
life history evolution **4:631–641**
analytical frameworks
genetic models
characteristics **4**:633
quantitative genetic models **4**:633–634, **4**:634*F*
simple Mendelian models **4**:633, **4**:633*T*
phenotypic models
game theory **4**:632–633
heritability **4**:631–632, **4**:632*T*
optimality analysis **4**:632
basic concepts
Fisher's *r* measure **4**:634–637
local measures **4**:634–637
phenotypic plasticity **4**:639*F*
predictable environments **4**:637–640, **4**:637*F*, **4**:638*F*, **4**:639*F*, **4**:640*F*
static environments **4**:635–637, **4**:636*F*, **4**:637*F*
stochastic and predictable environments **4**:640
stochastic environments **4**:637
Cole's paradox **4**:631
tradeoffs **4**:634, **4**:635*F*
variability considerations **4**:631
life-history omnivores **3**:502–503, **3**:506
Life Science Identifier (LSID) **1**:441–442
lifetime of sun **3**:265*T*
light **7**:296, **7**:297*F*
light-footed clapper rail **6**:454–455
lightning fires **3**:436
light-producing symbioses
bioluminescence **5**:121–123, **5**:122*F*
fish **5**:123
squid **5**:123–124
light trapping **1**:330
lignin **3**:639, **4**:429, **6**:119–121
Liguasan marsh **1**:276*T*
Ligularia **1**:270
Ligularia altaica **1**:269
Ligustrum robustum **4**:364
Liliaceae **3**:129*T*
lima beans **2**:663–664, **4**:287–290
Limacodidae **5**:391*F*, **5**:396
Limaya **7**:347*F*
Limbang mangroves **1**:276*T*
Limenitis archippus **1**:656, **2**:141–142
Limenitis reducta **5**:463*T*
Limicolaria flammea **1**:522*F*
Limnanthes bakeri **3**:209
Limnanthes douglassii **3**:209
Limnocharis flava **1**:164
Limnohalacarus capernaumi **5**:470*T*
Limnothrix redekei **6**:33*F*
Limoniscus violaceus **3**:226
limonoids **2**:427
Limosa lapponica baueri **7**:27–28
Limosa limosa **5**:723
limpets
calcium carbonate ($CaCO_3$) grazing **3**:612, **3**:613*F*
endangered species **3**:201–202
extinctions **3**:202*T*
hydrothermal vent ecosystems **7**:330
number of species **3**:182*T*
predation effects **4**:334–335, **5**:56–57, **5**:58*F*, **6**:232
Linaceae **5**:623*F*
lindane **5**:697*T*, **5**:700*T*
Lindeman, Raymond **3**:239–240
Lindera obtusiloba **1**:271–272
linear regression analysis **4**:514
Linepithema humile **1**:81, **3**:222, **6**:50, **6**:93–95
lingcod **3**:468
linguistic diversity **4**:270–271, **4**:270*T*, **6**:566
Lingulodinium polyedrum **1**:477*T*
linkage disequilibrium (LD) **2**:265–266
Linked Open Data (LOD) **1**:444
Linnaeus, Carolus
biogeographical research **1**:577
dispersal theories **2**:539–540
evolutionary theories **3**:393
literary perspectives **4**:651
portrait **7**:89*F*

Linnaeus, Carolus (*continued*)
systems of nomenclature **5**:547–548
taxonomic methods **4**:121, **7**:89–90, **7**:100
taxonomic versus functional diversity **2**:648
vicariance biogeography **7**:344
Linum usitatissimum **2**:139, **2**:383, **2**:662–663
linuron **5**:696
Linyphia triangularis **3**:222
Linyphiidae **1**:207
Liodidae **1**:327*T*
lionfish **2**:341, **3**:468, **3**:468*F*
lions
adipose tissue **7**:28, **7**:29*F*, **7**:34*F*
disease outbreaks **3**:191, **3**:411
introduced species **4**:362
late Quaternary extinction event (LQEE) **4**:724–725*T*
species diversity **1**:686, **4**:693
lion's mane mushrooms **3**:637–638, **3**:637*T*
lion-tailed macaque **6**:253–254, **6**:256–258*T*
lion tamarins **1**:664*T*, **1**:666*T*, **6**:251*T*, **6**:256–258*T*, **6**:560
Liphistiidae **1**:206
Liphyrinae **1**:652
lipids **5**:267, **5**:268*F*
lipid storage **7:23–38**
adipose tissue
adipocyte size and number **7**:27
anatomical arrangement **7**:27–30, **7**:28*F*, **7**:29*F*
characteristics **7**:26–27, **7**:27*F*
fatness measures **7**:30–31, **7**:30*F*
functional role
adipocyte metabolism **7**:31–32
fatty acid composition **7**:32–33
thermal insulation **7**:33–35, **7**:34*F*
invertebrates **7**:24
life histories
energy storage optimization **7**:35
hibernation **7**:36–37
migration **7**:35–36
reproduction and lactation **7**:37–38
material characteristics **7**:23–24
plants
growth cycles **7**:24–25
lipid composition **7**:25–26
predator protection **7**:26
seed storage **7**:24–25
lipophorins **7**:24
Lipotes vexillifer **3**:192, **5**:75
Lipteninae **1**:652
Liquidambar styraciflua **2**:26
liquified natural gas (LNG) **3**:257
liquified petroleum gas (LPG) **3**:257
Lissachatina fulica **3**:222, **4**:357
Lissamphibia **1**:169
Lissoclinum patellum **1**:310–311
Lissorhoptrus oryzophilus **2**:551–552*T*
listeriosis **7**:302–303
litchi **2**:663
Litchi chinensis **2**:663
literary perspectives **4:649–655**
basic concepts **4**:649
colonial expansion **4**:651
Darwinism **4**:651–652
ecocriticism **4**:653–654
ecology–industrialism relationship **4**:652
environmental and land ethics **4**:652–653
etiological tales of nature **4**:650
language and words **4**:649–650
literature–science relationship **4**:649
mass media **4**:654
natural disasters **4**:654
natural rights and animal fables **4**:653
pastoralism **4**:650–651
Lithaea caustica **3**:116
Lithobates catesbeianus **2**:529–530, **2**:530*F*, **6**:457
Lithobiomorpha **1**:246*T*, **5**:421–423, **5**:422*F*
Lithobius crassipes **5**:425
Lithobius curtipes **5**:428, **5**:430–431
Lithobius emarginatus **5**:428
Lithobius forficatus **5**:428, **5**:430–431, **5**:576
Lithobius microps **5**:430–431
Lithobius mutabilis **5**:430–431
Lithobius validus **5**:422*F*
Lithocarpus
Asia **1**:267, **1**:270–271
North America **7**:118
Lithocranus walleri **1**:51
Lithodes **3**:610*F*
Lithophaga lithophaga **3**:199
lithosphere **3**:279
Lithothamnium sp. **6**:723*T*
Lithuania
botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
litmus paper **1**:591
Litobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Litobothrium daileyi **7**:456–457*F*
Litocranius walleri **5**:235
Litomosoides **7**:428*F*, **7**:432, **7**:432*T*
Litomosoides thomomydis **7**:432*T*
Litomosoides westi **7**:432*T*
Litopterna **2**:583*F*, **4**:727*T*
Litoria nannotis **1**:172
Litoria nyakalensis **1**:173*F*
Litoria prora **1**:173*F*
Litoria rubella **1**:173*F*
little auks **1**:617
little brown bats **5**:313
little egret **5**:476–477
little greenbul **3**:144–145
little-ringed plover **5**:476–477
little spotted kiwi **5**:366–367
littoral zone **4**:374*F*
Littoraria angulifera **5**:17–18
Littoraria flammea **2**:342–343, **3**:202*T*
Littoraria intermedia **5**:17–18
Littoraria pallescens **5**:17–18
Littoraria scabra **5**:17–18
Littorina littorea **4**:10
Littorina obtusata **4**:10
Littorina saxatilis **2**:271, **5**:425
livebearers **3**:467
live fences **1**:380–381
liverworts
alpine ecosystems **1**:153–154, **1**:154*F*
Antarctica **1**:184*T*
Arctic ecosystems **1**:231*T*
boreal forests **1**:628–630
conservation practices **6**:83
forest canopies **3**:520, **3**:521*T*
freshwater ecosystems **3**:571–572
habitat fragmentation **4**:517
North America **5**:564–565
number of species **6**:58*T*
RNA editing **5**:576
livestock
agricultural production evolution **1**:97–98, **1**:98*F*
animal breeding **3:100–107**
challenges **1**:648
conservation planning **1**:639–640
conservation strategies
ex situ strategies **1**:645–646
general discussion **1**:642–645
in situ strategies **1**:642–645
crossbred animals **1**:639
current trends **1**:638–639
diversity challenges **1**:636
future outlook **1**:648
genetic structure **1**:637–638
global breeds **1**:638–639
livestock breeds **1**:637
management strategies
conservation strategies **1**:642–645
Food and Agriculture Organization (FAO) programs **1**:640–642, **1**:641
inventory, description, and characterization assessments **1**:641–642, **1**:642*F*
risk status classification **1**:641–642, **1**:642*F*, **1**:642*T*, **1**:643*F*, **1**:644*F*
new technologies
cloning **1**:647–648
cryopreservation **1**:646
gene mapping **1**:646–647
molecular characterization **1**:646–647
wild relative use **1**:640
biodiversity-friendly farming **2**:542–544
conservation programs **4**:319, **4**:321*T*
deforestation **2**:453
domesticated ruminants **1:702–711**
biotic relationships
birds **1**:710
invertebrate fauna **1**:709–710
natural predators **1**:709
parasites and pathogens **1**:710–711
plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
seed dispersal and germination **1**:708–709
small mammals **1**:710
ecological role
desertification **1**:704
erosion **1**:704–705
grazing effects **1**:703–704
land degradation **1**:704–705
nutrient cycling **1**:705–706, **1**:705*F*

over-enrichment factors **1**:705–706
plant diversity–grazing intensity relationship **1**:704*T*
vegetation impacts **1**:703–704
historical perspective **1**:702–703
domestication **6**:567
global livestock **6:330–344**
anthropogenic impacts **6**:343
biological impoverishment
animal damage control **6**:342
aquatic ecosystems **6**:335–337, **6**:336*F*, **6**:341–342
biogeochemical cycling **6**:342
cropland conversion **6**:341
desertification **6**:337–338
direct/primary influences **6**:335–337, **6**:336*F*, **6**:337
exotic species invasions **6**:341
fire regime alterations **6**:342–343
global impact **6**:334–337, **6**:335*T*
greenhouse effect **6**:340–341
indirect/secondary and tertiary influences **6**:336*F*, **6**:337
pasture conversion **6**:338–340, **6**:340*F*, **6**:341
pathogens **6**:341
riparian zones **6**:337
terrestrial ecosystems **6**:335–337, **6**:336*F*
water diversions **6**:341–342
biological soil crusts **6**:339
livestock-caused degradation **6**:343
Global Trade Analysis Project (GTAP) model **4**:542*F*, **4**:543*F*, **4**:543*T*
grazing effects **2**:542–544
introduced pathogens
Australia **1**:76*T*, **1**:81
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:82
parasitic disease vectors **5**:672
private conservation areas **2**:297*T*
wild mammal threats **3**:191
living dead **4:590–598**
animal populations **4**:594–595
anthropogenic impacts **4**:594–595
conservation practices **6**:81
deforestation **4**:591–593, **4**:592*F*
extinct megafauna **4**:591–593, **4**:593*F*
food chains **4**:595–596
forest ecosystems **4**:590–591, **4**:591*F*
requirements **4**:593–594
restoration programs **4**:597
small plants **4**:594
wildland patches **4**:596–597, **4**:596*F*
living desert museums **7**:473–474
lizardfish **3**:465–466
lizards
adaptive radiation **1**:31*F*, **1**:32
alpine ecosystems **1**:154–155
Australia **1**:302*F*
characteristics **6**:402, **7**:340–341
cladistic analysis **7**:349*F*, **7**:350*F*, **7**:350*T*
conservation status **3**:216*T*, **6**:411
desert ecosystems **2**:470*T*, **2**:472–474
dry Asian regions **1**:268
foraging behaviors **6**:406–408, **6**:407*T*
forest canopies **3**:513–514
interspecific competition **2**:187–188, **2**:187*T*, **2**:533
mangrove ecosystems **5**:13
migratory patterns **5**:315–316
Near East ecosystems **5**:472
past, recent, and future extinction rates **2**:168*F*
South America **3**:115
Succulent Karoo **1**:488
threatened species **1**:515–516
tropical rain forests **1**:263
see also Reptilia/reptiles
llamas **1**:644*F*, **2**:475–476, **2**:476*T*, **4**:697, **4**:724–725*T*
Llanos ecosystems **3**:116–117, **5**:496, **5**:497*F*
Llanos region **3**:105–107, **3**:106*F*
Loa loa **7**:428*T*
Lobatostoma manteri **5**:657, **5**:659*F*, **5**:662*F*
Lobau wetlands (Austria) **7**:59
Lobelia **7**:275–276
loblolly pines **6**:150*T*
Lobochilotes **1**:25*F*
Lobodon carcinophaga **5**:90
lobsters **5**:464–465, **5**:464*F*, **7**:201*T*
Locke, John **1**:112
loco gastropod **6**:232
Locusta migratoria **2**:549–550
locusts **2**:549–550, **3**:221–222, **3**:734, **3**:735
lodgepole pines **6**:125
Loeb, Lawrence **5**:633
Logarcek (Slovenia) **7**:59*T*
loggerhead turtles **2**:356, **5**:472
logging operations
deforestation
direct causes **2**:453–454, **2**:454*T*
sustainable practices **2**:455
underlying causes **2**:451
logged and managed forests **1:446–458**
abiotic changes
forest floor hydrology **1**:450
insolation **1**:450
temperature **1**:450
wind patterns **1**:450
biotic changes
aquatic fauna **1**:453
behavioral changes **1**:453
characteristics **1**:450–451
domestic animals **1**:453
fungi **1**:451
invertebrates **1**:451–452
mycorrhizae **1**:451
nonforest species **1**:453
plants **1**:451
vertebrates **1**:452–453
ecological interactions **1**:453–454
emerging trends **1**:455–457
environmental impacts **7**:314–315
evolutionary processes **1**:454–455
genetic effects **1**:454–455
management systems
general discussion **1**:447
long-term intensive forest management practices **1**:448
managed versus secondary forests **1**:448
nontimber forest products (NTFPs) **1**:448
shifting agricultural practices **1**:448
timber management **1**:447
plants
early successional phases **1**:451
exotic species invasions **1**:451
forest management impact **1**:451
research background **1**:446–447
structural alterations
coarse woody debris (CWD) **1**:450
environmental impact **1**:448–449
forest canopies **1**:449
fragmentation **1**:449
gaps and edges **1**:449
soil structure **1**:450
spatial patterns **1**:449–450
vertical structure **1**:449
synergistic anthropogenic effects **1**:455
timber management
clearing systems **1**:447
regeneration systems **1**:447
salvage logging **1**:447–448
selective logging systems **1**:447
vertebrates
behavioral changes **1**:453
biotic changes **1**:452–453
birds **1**:453
domestic animals **1**:453
food resource loss or gain **1**:453
habitat loss or gain **1**:452–453
primates **1**:453
logging industry–owl population controversy **3**:314–315, **3**:315*F*
South America **6**:608, **6**:609
logistic equation **1**:696, **1**:696*F*, **1**:697, **1**:697*F*, **1**:699–700, **6**:177, **6**:177*F*
Logone River **7**:153–154
logrunner **4**:517
Logtak Lake **1**:276*T*
Lohmanella heptapegoni **5**:470*T*
Loihi Seamount **7**:328
Lolium multiflorum **2**:389–390
Lomaphorus **4**:727*T*
Lomatium peckianum **3**:214
Lonchaeidae **3**:489–491*T*
Lonchocarpus nicou **1**:589
Lonchocarpus utilis **4**:286–287, **4**:288*T*
Lonchopteridae **3**:489–491*T*
London Amendments **1**:289
London, Jack **4**:652
long-distance dispersal **2**:539–540, **6**:616–617
long fallow **6**:325–327, **6**:326*F*, **6**:326*T*
long-horned beetles **1**:327*T*, **1**:328
long-horned buffalo **5**:364–365
Longidorus **7**:422–424
Longipodacrangonyx **7**:56–57, **7**:57*F*
long-leaf pine **6**:455
long-legged buzzards **5**:474
long-legged flies **3**:489–491*T*, **3**:492
Longosomatidae **7**:417*F*
LOng RAnge Navigation (LORAN) system **3**:443
Long's bulrush **3**:209
long-spined sea urchins **2**:530–531

long-tailed bate **4**:194*T*
long-tailed macaques **5**:603–604
long-toed salamanders **4**:516
Lonicera etrusca **5**:481–482
Lonicera japonica **3**:410–411
Lontra felina **5**:68
loons **1**:12*F*, **3**:256, **7**:181–182, **7**:390
Lopez, Barry **4**:653
Lophira alata **1**:50
Lophius americanus **3**:466–467
lophocoronids **5**:385–386
Lophocoronina **5**:386*F*
Lophocoronoidea **5**:386*F*, **5**:387–388*T*
Lophogastrida **2**:416–417*T*
Lophomyrmex quadrispinosus **5**:463*T*
Lophotrochozoa **3**:396*F*
loquats **2**:663
Loranthaceae **4**:71, **6**:315*T*
Loranthus acaciae **5**:476, **5**:483
Lord Howe Island stick insect **3**:225
Lord Howe woodhen **7**:188
Loricifera
  classifications and sampling methods **4**:381*T*
  fossil record **3**:684*F*
  global diversity **4**:385*T*
  marine ecosystems **5**:50*T*
  phylogenetic relationships **7**:433*F*
loriciferans **2**:8*F*
Lorisidae **4**:692–693, **6**:250–252, **6**:251*T*
Loris lydekkerianus grandis **6**:256–258*T*
Loris lydekkerianus nordicus **6**:256–258*T*
Loris tardigradus nycticeboides **6**:256–258*T*
Loris tardigradus tardigradus **6**:256–258*T*
Los Angeles County Natural History Museum **5**:408*F*
Lotka, A. J. **2**:178–180, **2**:638–639, **3**:59
Lotka–Volterra predator–prey system
  basic concepts **2**:178–180, **2**:178*F*, **2**:179*F*, **6**:716–717, **6**:716*F*, **6**:717*F*
  food webs **3**:500–501, **3**:508
  interaction strengths **3**:507
  species coexistence **6**:672
  species packing **4**:643
  structural stability **7**:8–9, **7**:9, **7**:9*F*
Lottia alveus **3**:202*T*
Lottia edmitchelli **3**:202*T*
Lottia gigantea **6**:232
lotus **1**:590, **2**:663
Louisiana delta restoration (case study) **7**:378
louse flies **3**:489–491*T*
lower atmosphere pollution **1**:289–290
low-input high-diversity (LIHD) bioenergy cropping systems **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
low-input low-diversity (LILD) bioenergy cropping systems **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
lowland capped langur **6**:256–258*T*
lowland long jaw galaxid **4**:194*T*
lowland red slender loris **6**:256–258*T*
low mountain ecosystems
  characteristics **7**:274–275
  humid forests **7**:275
  perhumid forests **7**:275
  subhumid forests **7**:274–275
Loxia curvirostra **2**:549–550
Loxigilla noctis **1**:618
Loxioides bailleui **1**:29*F*
Loxoconcha galilea **5**:470–471, **5**:470*T*
Loxodonta **2**:243–244
Loxodonta africana **1**:47, **1**:487–488, **1**:601, **2**:272, **2**:352, **3**:197, **4**:694–695, **5**:225*T*
Loxodonta cyclotis **4**:694–695, **5**:225*T*
Loxomataceae **6**:59
Loxops coccineus **1**:29*F*
lubber grasshoppers **3**:734*F*, **3**:735
Lubbock, John **2**:424
Lucanidae **1**:327*T*
lucerne **2**:391
luciferase **5**:122*F*
Lucilia cuprina **4**:301
Lucretius **1**:15
Ludovia **3**:522*T*
Ludwigia decurrens **1**:164
Ludwigia helminthorriza **1**:164
Ludwig's bustard **1**:487–488
Lueehea seemanni **6**:306
Luehea seemannii **1**:330
Lumbrineridae **7**:417*F*
Lumnitzera **1**:277, **5**:11*T*
lumpfish **3**:467–471
Lunama Kalapuwa **1**:276*T*
lungfish
  characteristics **3**:461, **3**:461*F*, **7**:338
  hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
lung worm parasites **2**:532
Lupinus **5**:215
Lupinus albus **2**:390
Lupinus angustifolius **2**:389–390
Lupinus elegans **6**:458–459
Lupinus plattensis **7**:64
lupus **7**:302–303
Luscinia luscinia **1**:616*F*
Luscinia megarhynchos **1**:616*F*, **4**:573
Luther's larva **7**:462–466, **7**:464–465*F*
Lutra canadensis **7**:390
Lutra felina **5**:77*T*
Lutra lutra **4**:357, **5**:473
Luxembourg
  botanical gardens **6**:507*T*
  energy efficiency measures **3**:254*F*
  urbanization **4**:560*F*
Luzula confusa **1**:236*F*
Luzula racemosa **4**:103
lyases **5**:269*T*
Lycaenidae **5**:387–388*T*
  butterfly caterpillars **1**:652*F*
  geographic diversity
    latitudinal gradients **1**:657, **1**:657*T*, **1**:658*F*
    neotropics **1**:657–658, **1**:658*T*
    regional patterns **1**:658, **1**:659*T*
    species numbers **1**:657*T*
  taxonomy **1**:651, **1**:651*F*
Lycaeninae **1**:652
Lycaon pictus **1**:52, **2**:190, **3**:190, **4**:715–716, **5**:449, **6**:573
Lycidae **1**:327*T*
Lycium arabicum **1**:268
Lycium shawii **5**:485–486
Lycogala epidendrum **3**:631*F*
Lycopersicon esculentum **1**:463, **2**:383, **2**:391, **2**:663–664, **6**:361*T*
Lycopersicon lycopersicum **6**:554
lycophytes **4**:168–169, **6**:58*T*, **6**:59
Lycosoidea **1**:207
Lyell, Charles **1**:578–579, **2**:419–420, **3**:393, **3**:414, **3**:483
Lygaeoidea **7**:294
Lygaeus equestris **5**:316
Lyginopteris **5**:618*F*
Lygodium microphyllum **4**:360
lygus bugs **7**:293
Lymantria dispar **2**:551–552*T*, **4**:183, **4**:361, **5**:395–396, **5**:402–403
Lymantriinae **5**:391*F*, **5**:396
Lyme disease **1**:361–362, **1**:363*F*, **3**:349, **3**:411, **5**:316–317, **5**:415, **7**:302–303
Lymnaea stagnalis **1**:368, **1**:368*F*
Lymnaeidae **4**:374
lymphatic filariasis **5**:671*T*
Lyncodon patagonicus **2**:587*F*
Lyngbya **2**:614
Lyngbya majuscula **6**:130–131, **6**:130*F*
lynx **1**:631–632, **1**:686, **1**:697–698, **1**:698*F*, **3**:196–197
Lynx canadensis **1**:631–632, **3**:196–197
Lynx pardinus **7**:262–263
Lysipomia sphagnophila **4**:103
lyssavirus **2**:551–552*T*
Lystrosaurus **5**:4–5, **5**:164–165
Lythraceae **5**:11*T*
Lythrum salicaria **1**:75–76, **3**:209, **3**:410–411, **6**:93–95, **6**:457, **7**:399
Lytorhynchus diadema **5**:467–469, **5**:472

M

Maackia amurensis **1**:271–272
Maasai culture **2**:224–225
Maastrichtian stage
  dinosaur diversity **2**:517–518, **2**:519*T*, **2**:520*T*
  mass extinctions **3**:415–418, **3**:416*T*, **5**:161*F*, **5**:173–175, **5**:174*F*
  taxonomic richness **2**:521*F*
Macaca fascicularis **5**:603–604
Macaca fuscata **4**:692–693
Macaca maura **6**:256–258*T*
Macaca mulatta **6**:253
Macaca munzala **6**:256–258*T*
Macaca nemestrina **5**:603–604
Macaca nigra **6**:256–258*T*
Macaca pagensis **6**:256–258*T*
Macaca silenus **6**:253–254, **6**:256–258*T*
Macaca sinica aurifrons **6**:256–258*T*
Macaca sinica opisthomelas **6**:256–258*T*
Macaca sinica sinica **6**:256–258*T*
Macaca sylvanus **4**:692–693, **6**:251–252, **6**:256–258*T*
macaques **4**:692–693, **5**:14, **5**:604, **6**:251*T*
maçaranduba **4**:287–290
Macaranga **1**:263

MacArthur–Levins species packing model **4**:645
MacArthur model **1**:402, **2**:180–181, **2**:180*F*, **2**:181*F*
MacArthur, Robert H. **1**:583–585, **2**:539–540, **3**:5, **3**:588
MacArthur–Wilson equilibrium theory **1**:25, **1**:583–585, **2**:239, **2**:239*F*, **3**:5, **3**:508, **4**:395–396, **4**:396*F*
MacCarthy, Cormac **4**:654
Macedonia **6**:519
Macfadyena **5**:489
mackerel **3**:467–471, **6**:422*T*
Macleay's honeyeater **2**:75*F*
Macoma inconspicua **3**:364
Macquarie Island **1**:180–181, **1**:186*T*
Macrauchenia **4**:727*T*
Macraucheniidae **4**:727*T*
macroalgae
  estuarine ecosystems **3**:358–359
  freshwater ecosystems **3**:358
  seagrass ecosystems **6**:547
  ultraviolet radiation effects **7**:299–300
Macrocystis **2**:642
Macrocystis pyrifera **3**:283, **3**:359, **5**:51
Macrodactylus sp. **1**:120
macroecology **6**:651
macrofauna
  marine invertebrates **4**:380–381, **4**:381*T*
  sandy beaches **2**:129–131, **2**:130*F*
  soil fauna **6**:588, **6**:590*T*
macrofungi **1**:142*F*
Macroglossus minimus **5**:14
Macrohectopus **6**:26
Macroheterocera **5**:386*F*, **5**:387–388*T*
macroinvertebrates **3**:571–572, **3**:572*T*
Macrolepidoptera **5**:386–389, **5**:387–388*T*
Macrolobium acaciaefolium **4**:286–287
macronuclear destined sequences (MDSs) **5**:569, **5**:570*F*, **5**:571*F*
macroparasites **2**:524*T*, **2**:525–527, **2**:526*F*, **5**:657
macrophytes **5**:305–306, **6**:477–478
macroplankton **4**:383*T*, **5**:133*F*
Macropodidae **4**:683, **4**:688–689
macropods **4**:689
Macroprotodon cu-cullatus **5**:472
Macropus birdselli **5**:364*F*
Macropus ferragus **5**:363–364, **5**:364*F*
Macropus gouldi **5**:364*F*
Macropus pitonensis **5**:364*F*
Macropus rama **5**:364*F*
Macropus rufogriseus **4**:728
Macropus siva **5**:364*F*
Macropus spp. **6**:333
Macropus stirtoni **5**:364*F*
Macropus thor **5**:364*F*
Macropus titan **5**:363–364, **5**:364*F*
Macropygia mackinlayi **2**:189–190, **2**:189*F*
Macroscelidea **4**:690*T*, **4**:697–698
Macroscelides proboscideusi **4**:697–698
Macrosoma nigramacula **5**:385*F*
Macrostomida **7**:439–440*T*, **7**:440–443, **7**:441*F*, **7**:442–443*F*
Macrotermes **4**:409*T*
Macrotermes bellicosus **4**:419
Macrotermes michaelseni **4**:418, **4**:420*F*
Macrotermes subhyalinus **4**:419
Macrotermitinae **4**:409*T*, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:430*T*
Macrotis leucura **4**:363–364
Macroxenus enghoffi **5**:423*F*
macrozooplankton **3**:362
Mactan Bay **1**:276*T*
Maculina arion **4**:363
Madagascar
  biodiversity richness **1**:501*F*
  botanical gardens **6**:519
  butterflies **1**:658, **1**:659*T*
  coral reef ecosystems **1**:500*T*
  deforestation **1**:502*T*, **1**:677*F*
  desert ecosystems **2**:467–468
  domesticated ruminants **1**:702–703
  ecosystems **1**:55–56
  endangered mammals **3**:193–194, **3**:194*T*
  endemism **1**:501*F*, **3**:232*T*, **3**:233, **4**:271*T*
  first-contact extinctions **5**:365
  fish diversity **3**:472–473
  habitat destruction **3**:216–217
  institutional collections and capacity **1**:505*T*
  invasive species
    aquatic ferns (*Salvinia molesta*) **4**:346–347
    exotic fish **4**:663
    Lantana camara **4**:352–353
  late Quaternary extinction event (LQEE) **4**:657, **4**:658*F*, **4**:723*F*, **4**:728–729
  mammalian evolution **4**:701*T*
  plant species **6**:82*T*
  primates
    endemism **6**:252*T*
    primate diversity **6**:250–252, **6**:251*T*, **6**:252*T*
    threatened species **6**:258*T*
  protected area systems **6**:493–494
  Ranomafana National Park **4**:437–438
  rice production **1**:88
  species richness **1**:501*F*, **4**:271*T*
  termites **4**:409*T*
  threatened species **7**:204*F*
  tourism industry **7**:224
  tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*
  urbanization **4**:560*F*
Madagascar periwinkle **1**:594
Madais fausta **5**:463*T*
Madame Berthe's mouse lemur **6**:256–258*T*
Madhuca indica **6**:226
Madia **1**:30*F*
Madoqua **1**:51
Madoqua kirkii **6**:572–573
Madrean pine–oak woodlands **6**:82*T*
madrone **7**:118
Maerua crassifolia **5**:486
Magelonodae **7**:417*F*
magenta petrel **7**:184
Maggiore (Lake) **4**:461*T*
Magnaporthe grisea **6**:361*T*
magnesium (Mg)
  elemental partitioning **1**:554*T*
  relative abundance **1**:553*F*
  river water composition **3**:562*T*
  root–soil interface **6**:114*F*
Magnolia **7**:118
Magnoliaceae **1**:270–271, **5**:618*F*
Magnoliales **5**:621*F*
Magnolia obovata **1**:271
Magnolidae **4**:72*F*
Magonia glabrescens **1**:161
Maha Lewaya **1**:276*T*
Mahanadian monsoon forest **1**:262–263*T*
Mahasena corbetti **5**:603–604
Mahaweli Ganga floodplain system **1**:276*T*
mahogany **7**:273
mahua **6**:226
Maiasaura **2**:519*T*
maidenhair tree **2**:47, **6**:59–60
Maimonides **3**:321
mainland drill **6**:256–258*T*
maitake mushrooms **3**:637*T*
maitotoxin **6**:131, **6**:133*F*
maize
  crop production **1**:86
  crop rotation systems **1**:116*T*
  genetic diversity
    commercialization effects
      general discussion **1**:91–92
      intensified environments **1**:92
      non-intensified environments **1**:92–93
    genetic narrowing **1**:89–90
    genetic trends **1**:90
    landrace–modern varieties comparisons **1**:89
    remaining gene pool percentage **1**:132*T*
    research summary **1**:93
  intercropping systems **2**:383, **2**:384*F*
  North and South America **2**:663–664
  phenotypic variation **1**:462, **1**:462*F*
  plant conservation **6**:85–86
  regional cultivars **3**:32*F*
  root–soil interface **6**:113*T*
  seed industries **1**:91
  slash-and-burn systems **6**:553
  spatial distribution **1**:87–89, **1**:88, **1**:88*T*
  termites **4**:432*T*
Majorana syriaca **5**:482
major histocompatibility complex (MHC) **2**:534, **2**:629
majority-rule consensus analysis **2**:42–44, **2**:43*F*
Malabar rain forest **1**:262–263*T*
Malacoherpesviridae **5**:134*T*, **5**:135
Malacoihrix fendleri **2**:373*T*
Malacostraca **2**:400–401, **2**:401*T*, **3**:681*F*, **3**:685*F*, **6**:25
Malaise trapping **1**:330
malaria **1**:593–594, **3**:349, **5**:668, **5**:671*T*
malaria parasites **2**:529–530, **5**:660*T*, **5**:671*T*
Malassezia furfur **3**:634*T*
malathion **5**:693–694, **5**:699*T*
Malawi
  botanical gardens **6**:519
  energy efficiency measures **3**:254*F*
  woodland savannas **1**:67–68
Malawi (Lake) **3**:444, **4**:463–464, **4**:463*T*
Malayan rain forest **1**:262–263*T*

Malayan tiger **1**:518*F*, **1**:519*F*
Malayan white-handed gibbon **6**:256–258*T*
Malaysia
  animal species **1**:261*T*
  botanical gardens **6**:519
  butterflies **1**:658, **1**:659*T*
  deforestation **1**:677*F*, **4**:562–564, **4**:563*F*
  ecosystems
    deserts **2**:468
    genetic richness **1**:514*T*
    mangroves **1**:275*T*, **1**:277, **1**:513*T*
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    tropical moist forest **1**:277*T*
    tropical rain forests **6**:314*T*, **6**:317*T*
    vegetation types and coverage percentage **1**:264–265*T*
  endangered mammals **3**:194*T*
  endemism **3**:232*T*
  indigenous peoples **4**:270*T*
  invasive species **4**:353–354
  oil-palm plantations **1**:540–541, **5**:601
  plant species **1**:259*T*
  species richness **1**:499*T*
  termites **4**:432*T*
  see also Southeast Asia
Maldanidae **7**:417*F*
Maldives
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  ecosystems
    mangroves **1**:275*T*
    threatened species **1**:278*T*
  plant species **1**:259*T*
Malesherbiaceae **5**:216–217
Malgas Island **5**:57–58, **5**:60*F*
Mali
  botanical gardens **6**:507*T*
  energy efficiency measures **3**:254*F*
  intercropping systems **2**:382–383
  rice production **1**:88
  seasonal rainfall **2**:465*F*
  urbanization **4**:560*F*
malic acid **2**:363–364, **2**:363*F*
mallards **6**:689–690
Mallee emu-wren **7**:183–184
Mallet, J. **6**:620–621
Mallomonas **6**:293
Mallophaga
  Antarctica **1**:184*T*
  number of species **4**:392*T*
Mallotus **1**:263
Malpighiaceae **5**:489, **5**:621–625, **5**:623*F*, **6**:316*T*
Malpighiales **2**:374*T*
Malta
  botanical gardens **6**:519
  vegetation surveys **3**:341–344*T*
Malthusian overfishing **3**:445
Malthus, Thomas **1**:695, **2**:420–421, **6**:204–206
Maludam swamp forest **1**:276*T*
Malus pumila **2**:662–663
Malvaceae **6**:81
Malva parviflora **4**:364
Mammalia/mammals **4:681–707**
  air pollution impacts **1**:139
  alpine ecosystems **1**:154–155
  anthropogenic activities and impacts **3**:401*F*, **7**:307*T*
  Arctic ecosystems **1**:231*T*
  Asia
    dry Asian regions **1**:268
    endemic species **1**:261*T*
    temperate grasslands **1**:269
    threatened species **1**:278*T*
    tropical rain forests **1**:263
  Australia **1**:302*F*
  average lifespan **4**:656*T*
  biodiversity **3**:188
  biodiversity–population density relationship **7**:74*F*
  biogeographical models **1**:565–566, **1**:566*F*
  boreal forests **1**:631–632, **1**:633*F*
  Cenozoic mass extinctions **5**:176
  Central America **2**:25*T*
  characteristics **4**:681–682, **7**:341
  classifications
    diagnostic characteristics **4**:681–682
    Eutheria **4**:683
    general characteristics **4**:681–682
    Metatheria **4**:683
    Monotremata **4**:682, **4**:682*F*
    Theria **4**:682–683
  climate change studies **2**:66
  climatic tolerance **4**:701–702
  conservation efforts **4:708–720**
    captive breeding programs **3**:196–197
    Caribbean hutias and solenodons
      background and characteristics **4**:713–714, **4**:713*F*, **4**:714*F*, **4**:715*F*
      conservation practices **4**:715
      general discussion **4**:713–714
      major threats **4**:714–715
    case studies
      Caribbean hutias and solenodons **4**:713–714
      cheetahs and African wild dogs **4**:715–716, **4**:717*F*
      general discussion **4**:711
      important themes **4**:717–719
      saiga antelope **4**:711
    cheetahs and African wild dogs
      background and characteristics **4**:715–716, **4**:717*F*
      conservation practices **4**:716–717
      major threats **4**:716
      range areas **4**:718*F*
    conservation types
      captive breeding programs **3**:196–197, **4**:710–711
      charismatic species **3**:197–198
      economic incentives **3**:197
      government regulation **4**:710
      habitat/population management **4**:710
      protected area systems **3**:196, **4**:710–711
      reintroduction programs **3**:196–197
      species protection **4**:710
      sustainability **4**:710
      umbrella species **3**:197–198
    general discussion **4**:708
    government legislation and regulations **3**:195
    human impact **4**:709–710
    needs assessments **4**:708–709
    protected area systems **3**:196
    research summary **4**:719–720
    saiga antelope
      background and characteristics **4**:711, **4**:711*F*
      conservation practices **4**:712–713
      general discussion **4**:711
      major threats **4**:711–712, **4**:712*F*
      population decline **4**:712*F*
      range area **4**:712*F*
    Species Survival Commissions (SSC) **3**:195–196
    threat assessment **3**:191–192, **4**:709–710
  convergent evolution **3**:613, **3**:614*F*
  deoxyribonucleic acid (DNA) evolution **2**:56*F*, **2**:57
  desert ecosystems **2**:470*T*, **2**:472–474, **2**:474*F*
  diagnostic characteristics **4**:681–682
  domesticated ruminants **1**:710
  elevational range shifts **4**:605*T*
  El Niño Southern Oscillation (ENSO) events **3**:156–158
  endangered species **3:188–198**
    conservation practices
      captive breeding programs **3**:196–197, **4**:710–711
      charismatic species **3**:197–198
      economic incentives **3**:197
      government legislation and regulations **3**:195
      protected area systems **3**:196, **4**:710–711
      reintroduction programs **3**:196–197
      Species Survival Commissions (SSC) **3**:195–196
      umbrella species **3**:197–198
    future outlook **3**:198
    geographic locations
      endemism **3**:193–194
      top ranked countries **3**:193, **3**:194*T*
    historical perspective
      early/prehistoric era **3**:188–189
      Pleistocene extinctions **3**:189
      recent extinctions **3**:189–190, **3**:189*F*
    major threats
      causal factors **3**:190
      disease **3**:191
      exploitation **3**:191–192
      genetic diversity loss **3**:190–191
      habitat loss and fragmentation **3**:190
      interference **3**:192
      introduced species **3**:193
      livestock **3**:191
      multiple threats **3**:193
      persecution **3**:192
      physiological factors **3**:190
      pollution **3**:192–193

Red List Index (RLI) **3**:194–195, **3**:195*F*, **4**:709
taxonomic percentages **3**:195
endemism **3**:232*T*, **4**:702
energy requirements **6**:404*F*
evolutionary history **5**:1–2
evolutionary trends
brain size **4**:698
convergent evolution **4**:700, **4**:701*T*
Cope's rule **4**:698–699
Island Rule **4**:699
leg length **4**:699–700
prey–predator interactions **4**:699–700
extinctions
extinction rates **4**:704–705, **4**:705*T*
future extinction estimations **4**:706–707
geographic patterns **4**:705–706, **4**:705*T*
neotropical mammals **2**:589–592
Pleistocene
major die-offs **3**:189
North America **4**:722–725, **4**:724–725*T*, **4**:727*F*
pre-Quaternary mammals **5**:**1–9**
Age of the Dinosaurs **5**:5–6
background extinctions **5**:2–3
causal factors **5**:2
Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
distal and proximal causal factors **5**:3–4
early Paleocene faunal recovery **5**:7
geologic timescale **5**:1*T*
late Triassic faunal turnover **5**:4–5
mass extinctions **5**:2–3
Paleocene–Eocene thermal maximum **5**:7–8
Permian–Triassic mass extinction **5**:4
phylogenetic relationships **5**:1*F*
pseudoextinctions **5**:2
summary discussion **5**:8
synapsids **5**:1–2, **5**:4
temporal and biogeographic scales **5**:3
Tertiary faunal evolution **5**:7–8
therapsid radiation **5**:4–5
recent extinctions **3**:189–190, **3**:189*F*
food habits **3**:127–128, **3**:127*T*
forest canopies **3**:513–514
fossil record **3**:685*F*, **3**:688
freshwater ecosystems **3**:571–572, **3**:578–579, **3**:579*T*
fruit and seed dispersal **6**:50, **6**:51*T*
general discussion **4**:681
genetic diversity **3**:665*F*, **3**:668*F*
geographical variation **1**:614
government legislation and regulations
Endangered Species Act (ESA)
basic principles **3**:713*T*, **3**:716–717
genetic diversity protection **2**:264
historical perspective **2**:286, **4**:121–122, **6**:77–78
marine invertebrates **3**:199
species protection **2**:108–109
wildlife management **1**:397
Marine Mammal Protection Act (1972) **3**:713*T*, **3**:716
grasslands **4**:3–4
historical perspective
early/prehistoric era **3**:188–189
Pleistocene extinctions **3**:189
recent extinctions **3**:189–190, **3**:189*F*
Holocene postglacial spread **2**:550, **2**:551–552*T*
hotspots **4**:702
indicator species **7**:206–207*F*
introduced parasites–species decline relationships **2**:528*T*
introduced species
Australia **1**:76*T*, **1**:80
Brazil **1**:76*T*, **1**:83
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81
United Kingdom **1**:76*T*, **1**:79
United States **1**:76–77, **1**:76*T*
late Quaternary extinction event (LQEE) **4**:**721–732**
background information **4**:721
causal factors
anthropogenic models **4**:731–732
climatic models **4**:732
global patterns **4**:731–732
comparative studies
Africa **4**:658*F*, **4**:729–730
analytical methods **4**:721–725
Australia **4**:658*F*, **4**:727–728
Eurasia **4**:729–730, **4**:729*F*
Madagascar **4**:658*F*, **4**:728–729
North America **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
Oceania **4**:730–731
South America **4**:726–727, **4**:727*T*
time-transgressive sequence **4**:722*F*, **4**:723*F*
latitudinal gradients **1**:498*F*, **1**:499*T*
latitudinal range shifts **4**:604*T*
major threats **7**:203*F*
mangrove ecosystems **5**:14, **5**:20*T*
marine mammals **5**:**64–93**
cetaceans
characteristics **5**:65*T*, **5**:66–67
evolutionary history **5**:68–69
taxonomy **5**:65*T*
desmostylians
characteristics **5**:68
evolutionary history **5**:69–70
extinction/near extinction over evolutionary time **5**:75
extinction-contributing factors
body size and dimensions **5**:70
climate change effects **5**:89–91
fishery–diet/habitat overlap **5**:72
geographic regions **5**:92
hauling behavior predictability **5**:70–71
minimization strategies **5**:92–93
rapid recovery potential **5**:71
research summary **5**:91–92
sea surface-based respiration **5**:70
spatiotemporal distribution predictability **5**:70
subsistence and commercial market demand **5**:71
toxin and contaminant bioaccumulation potential **5**:71–72
extinction patterns and case studies
desmostylians **5**:75
extinction/near extinction over evolutionary time **5**:74–75
general discussion **5**:74–75
major anthropogenic extinctions **5**:75
Odobenidae **5**:74–75
sirenians **5**:75
identification/monitoring-related hindering factors
availability bias **5**:72
environmental uncertainty effects **5**:73–74
lack of consistent objective criteria **5**:74
observer bias **5**:72–73
population size survey data **5**:73
survey costs **5**:74
vital demographic parameters **5**:73
imminent extinction probability
beluga whales **5**:77*T*, **5**:79–80
blue whales **5**:81*T*, **5**:82
California sea otter **5**:81*T*, **5**:83–84
general discussion **5**:76
Mediterranean monk seals **5**:77*T*, **5**:80
Northern sea otter **5**:81*T*, **5**:84–86
right whales **5**:76–78, **5**:77*T*
species and subspecies **5**:77*T*
Steller sea lions **5**:81*T*, **5**:82–83
vaquitas **5**:77*T*, **5**:80
western North Pacific gray whales **5**:77*T*, **5**:78–79
West Indian manatee **5**:81*T*, **5**:83
major anthropogenic extinctions
North Atlantic gray whale (*Eschrichtius robustus*) **5**:76
species level **5**:75
subspecies level **5**:76
marine otters
characteristics **5**:68
evolutionary history **5**:70
natural versus unnatural extinctions **5**:74
pinnipeds
characteristics **5**:65*T*, **5**:67
evolutionary history **5**:69
taxonomy **5**:65*T*
polar bears
characteristics **5**:68
evolutionary history **5**:70
recovery evidence
bowhead whales **5**:86–87, **5**:87*T*
elephant seals **5**:87*T*, **5**:88
general discussion **5**:86
gray whales **5**:87–88, **5**:87*T*
humpback whales **5**:87, **5**:87*T*
species and subspecies **5**:87*T*
scientific whaling effects **5**:88–89
significant extinction probability
general discussion **5**:80–82
species and subspecies **5**:81*T*

Mammalia/mammals (*continued*)
sirenians
characteristics **5**:65*T*, **5**:67–68
evolutionary history **5**:69
extinction/near extinction over evolutionary time **5**:75
taxonomy **5**:65*T*
species level extinctions
baiji (Yangtze River dolphin/*Lipotes vexillifer*) **5**:75
Caribbean monk seal (*Monachus tropicalis*) **5**:75
Japanese sea lion (*Zalophus japonicus*) **5**:75–76
Steller's sea cow (*Hydrodamalis gigas*) **5**:76
taxonomy
characteristics **5**:65–66, **5**:65*T*
definition **5**:64–66
habitat boundaries **5**:65–66
Mediterranean-type climate ecosystems
Australia **5**:213*T*, **5**:221
California **5**:213*T*, **5**:215
Cape Region (South Africa) **5**:213*T*, **5**:220
Chile **5**:213*T*, **5**:217
Mediterranean Basin **5**:213*T*, **5**:218–219
metabolism/energy requirements **6**:404*F*
migratory patterns **5**:313–314
modern extinctions **2**:168*F*, **2**:170*F*, **5**:368*T*
mutation rates **2**:628*F*
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*
neotropical mammals **2:582–594**
background information **2**:582–583
biogeographical–macroecological patterns
areographic patterns **2**:588–589
ecogeographic rules **2**:588–589
latitudinal diversity **2**:588*F*
species richness **2**:588, **2**:589*F*, **2**:591*F*
conservation strategies **2**:591–592
evolutionary history **2**:582–583, **2**:583*F*
future research outlook **2**:592
Great American Biotic Interchange (GABI) **2**:582–583, **2**:584*F*
harvesting threats **2**:590*T*
major threats **2**:589–592, **2**:590*T*
zoogeographic regions
biogeographical limits **2**:583–585, **2**:585*F*
group diversity and classifications **2**:585–586
representative species **2**:586–587, **2**:587*F*
transition zones **2**:583–585, **2**:585*F*
new discoveries
Annamite region **4**:703–704
cryptic species **4**:704
discovery patterns **4**:703, **4**:703*F*
evolutionary significant unit (ESU) **4**:704
general discussion **4**:702–703
number of species **4**:703*F*
North America **5**:565
oil-palm plantations **5**:604–605
oil spills **3**:192–193, **3**:256
parasitic nematodes **7**:431–432
past, recent, and future extinction rates **2**:168*F*
phylogeny
historical perspective **4**:684
ordinal diversification
evolutionary structure **4**:685
genetic patterns **4**:686
marsupials **4**:686–687, **4**:687*T*
monotremes **4**:686
placentals **4**:689, **4**:690*T*
timing considerations **4**:685–686
pre-Quaternary mammals **5**:1*F*
pollination **6**:49*T*, **6**:141*T*, **6**:142–143
pre-Quaternary mammal extinctions **5:1–9**
Age of the Dinosaurs **5**:5–6
causal factors
background extinctions **5**:2–3
distal and proximal causal factors **5**:3–4
mass extinctions **5**:2–3
pseudoextinctions **5**:2
temporal and biogeographic scales **5**:3
Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
early Paleocene faunal recovery **5**:7
geologic timescale **5**:1*T*
late Triassic faunal turnover **5**:4–5
Paleocene–Eocene thermal maximum **5**:7–8
Permian–Triassic mass extinction **5**:4
phylogenetic relationships **5**:1*F*
summary discussion **5**:8
synapsids **5**:1–2, **5**:4
Tertiary faunal evolution **5**:7–8
therapsid radiation **5**:4–5
range size and occupancy distribution **4**:700–702
riparian zones **6**:463
risk status classification **1**:641–642, **1**:643*F*, **1**:644*F*
seagrass ecosystems **6**:547–548
Southeast Asia **1**:516–517
species diversity **4**:682*F*, **6**:712*F*
species richness gradients **1**:499*T*, **4**:613*F*, **4**:615*F*, **4**:700–702
species richness maps **6**:709*F*
Succulent Karoo **1**:487–488, **1**:487*T*
threatened species **1**:516–517, **7**:201*T*, **7**:206–207*F*
tourism industry **7**:223
tropical rain forests **6**:309–310, **6**:309*T*
urbanization threats **4**:564
wetland ecosystems **7**:389–391
wildfire effects **3**:156–158
Mammoth Cave (Kentucky, USA) **7**:59*T*, **7**:60
mammoths
late Pleistocene extinctions **5**:231–233
late Quaternary extinction event (LQEE) **4**:721, **4**:724–725*T*
Mammut **4**:723–725, **4**:724–725*T*
Mammut americanum **5**:225*T*
Mammuthus **4**:722, **4**:724–725*T*
Mammuthus columbi **5**:225*T*
Mammuthus primigenius **4**:729–730, **4**:729*F*, **5**:225*T*
Mammutidae **5**:224*T*, **5**:225*T*
Mamos **1**:29, **1**:29*F*
managed relocations **7**:238
management units **2**:268–269
Man and Nature: Or Physical Geography as Modified by Human Action (Marsh) **2**:280–281
Man and the Biosphere (MAB) Program **2**:301, **4**:325–326, **6**:489
manatee grass **3**:360–361
manatees
characteristics **4**:694, **5**:67–68
extant species **5**:225*T*
extinction probability **5**:81*T*
late Quaternary extinction event (LQEE) **4**:724–725*T*
migratory patterns **5**:314
number of species **4**:690*T*
overharvesting/overcollecting **3**:283
pollution effects **3**:192–193
seagrass ecosystems **6**:547–548
taxonomy **5**:65*T*
threat percentage **3**:195
Manchu–Japanese mixed forest **1**:262–263*T*
Manchuria **1**:271–272
mandrills **1**:49
Mandrillus leucophaeus **1**:49
Mandrillus leucophaeus leucophaeus **6**:256–258*T*
Mandrillus leucophaeus poensis **6**:256–258*T*
Mandrillus sphinx **1**:49
maneb **5**:695, **5**:699*T*
mangaba **4**:286–287
mangal **7**:369*T*
manganese (Mn)
biogeochemical cycles **1**:563
elemental partitioning **1**:554*T*
hydrothermal vent ecosystems **7**:328
relative abundance **1**:553*F*
mange mite **5**:660*T*
Mangifera **1**:263
Mangifera casturi **1**:514
Mangifera rubropetala **1**:514
mango **1**:514, **3**:32*F*
mangosteen **2**:663
mangrove crab **4**:337, **4**:339*F*
mangrove ecosystems
Amazon ecosystems **1**:164
aquaculture impacts **1**:193–194
Asia **1**:275–279, **1**:275*T*
Australia **1**:292–297
C_4 plants **2**:371*T*
Central America **2**:29
conservation practices **6**:77
conservation status **3**:171–173*T*
definitions and terminology **7**:369*T*
ecosystems **5:10–22**
characteristics **7**:271–272
diversity
general discussion **5**:18
genetic diversity **5**:22
global patterns **5**:18

local variations **5**:20–21
regional patterns **5**:20, **5**:21*F*
ecological interconnections **5**:18
economic exploitation and use **5**:22
global patterns
Atlantic-Caribbean-East Pacific (ACEP) region **5**:18–19, **5**:19*F*, **5**:20*T*
faunal diversity **5**:19–20, **5**:20*T*
Indo-West Pacific (IWP) region **5**:18–19, **5**:19*F*, **5**:20*T*
latitudinal range **5**:18, **5**:19*F*
longitudinal differences **5**:18–19, **5**:19*F*
local variations
animal distribution **5**:21–22
tree distribution **5**:20–21
mangrove trees
characteristics and geographical distribution **5**:10–11
families and genera **5**:11*T*
mangrove habitat **5**:10–11, **5**:10*F*
pneumatophores **5**:12*F*
reproduction **5**:11–12, **5**:12*F*
root system **5**:12*F*
salinity/salt tolerance **5**:11, **5**:11*T*
waterlogged soils **5**:11, **5**:12*F*
marine fauna
bivalves **5**:17
characteristics **5**:14–15
crustaceans **5**:15–16, **5**:20*T*
ecosystem engineers **5**:16–17
fish **5**:15, **5**:15*F*, **5**:20*T*
grapsid crabs **5**:15–16, **5**:15*F*
hermit crabs **5**:16
meiofauna **5**:18
molluscs **5**:17, **5**:20*T*
mud crabs **5**:16
ocypodid crabs **5**:16, **5**:16*F*
root communities **5**:14–15
shrimp **5**:16
snails **5**:17–18, **5**:17*F*
marine sedimentary habitats **5**:107, **5**:108*T*
shrimp farming versus mangrove conservation **3**:51–53, **3**:52*F*
South America **3**:111*F*, **3**:116
terrestrial fauna
amphibians **5**:13
birds **5**:13–14, **5**:20*T*
characteristics **5**:12–13
insects **5**:13, **5**:13*F*
mammals **5**:14, **5**:20*T*
reptiles **5**:13, **5**:20*T*
ecosystem services **6**:77
global extent **7**:370*T*
habitat destruction **1**:513, **1**:513*T*, **3**:404
intertidal ecosystems **4**:337, **4**:338*F*, **4**:339*F*
marine ecosystems **5**:45–46
nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
Near East ecosystems **5**:480*F*, **5**:487
protected area systems **6**:492*T*, **6**:499*T*
reproductive adaptations **7**:397
restoration programs **7**:375–378
shrimp farming versus mangrove conservation **3**:51–53, **3**:52*F*
South America **6**:599, **6**:607*T*, **6**:609
Southeast Asia **1**:513, **1**:513*T*
threatened species **7**:201*T*
mangrove finches **1**:27–28, **1**:28*F*
mangrove periwinkle **2**:342–343
Manhattan metric **2**:444–445
Manicaria saccifera **7**:272
Manihot esculenta **2**:382–383, **2**:663–664, **4**:281–285, **4**:283*F*, **4**:289*T*, **6**:553
Manila Bay **1**:276*T*
Manilkara huberi **4**:287–290, **4**:288*T*
Manilkara zapota **2**:663–664, **7**:273
manioc **4**:282–284, **4**:283*F*
Maniola jurtina **6**:7
Manis javanica **5**:604
Mantel tests **4**:510–511
mantids **4**:392*T*
Mantodea **4**:308–312, **4**:309*T*
Manu National Park (Peru) **6**:306
Maoridiamesa **7**:347*F*
maple trees
climate change effects **6**:150*T*
North America **7**:118
map turtles **7**:390
Maputaland **6**:82*T*
Maranhão red-handed howler monkey **6**:256–258*T*
Marantha arundinaceae **4**:287–290, **4**:289*T*
Marathrum **5**:623*F*
Marattia fern **6**:78
marbled murrelet **3**:256, **3**:314–315, **4**:564
marbled teal **7**:185
Marcgravia **3**:522*T*
Marcgraviaceae **3**:522*T*
Marchantia polymorpha **5**:576
march flies **3**:489–491*T*
March of the Penguins (2006) **4**:654
Marco Polo sheep **3**:192
Marcus Island **5**:57–58, **5**:60*F*
MAR-ECO project **2**:19, **2**:20*F*
Margalef, Ramon **3**:588
Margarita Island tufted capuchin **6**:256–258*T*
Margaritifera auricularia **3**:411–412
Margaritifera margaritifera **4**:376
Margarops fuscatus **7**:184–185
marginal lands **1**:543, **1**:544*F*
Marginidae **3**:489–491*T*
mari **4**:286–287
mariculture **5**:61
marigolds **2**:391
marine ecosystems **5:45–55**
anthropogenic impacts **5:56–63**
artificial reefs **5**:61
ecosystem engineer species **5**:60–61
eutrophication **5**:53–54
invasive species **5**:54, **5**:60–61
major threats **5**:56
mariculture **5**:61
overfishing **3**:410, **5**:54, **5**:61
pollution **5**:61, **6**:150–151
research summary **5**:62
rocky intertidal communities **5**:56–57, **5**:58*F*, **5**:59*F*
rocky subtidal communities **5**:57–58, **5**:60*F*
sea otter populations **5**:58–60
biodiversity
genetic diversity **5**:50
higher taxa **5**:47–49, **5**:50*T*
species **5**:49–50, **5**:50*T*
biodiversity loss **4**:664
biodiversity-rich countries **1**:498
chlorophyll production **5**:46
climate change effects **6**:150–151
conservation **5:32–44**
climate change effects
abundance impacts **5**:34–35
adaptation rates **5**:36
assisted colonization **5**:41
climate tolerance **5**:36–37
community impacts **5**:37–38
conservation strategies **5**:38–39
distribution impacts **5**:35–36
ecosystem-based management **5**:40
evolutionary impacts **5**:36
fisheries management **5**:40–41
genetic diversity **5**:36–37
global impacts **3**:449, **5**:32
local adaptations **5**:36
marine protected areas **5**:39–40
ocean acidification **5**:33–34
physiological impacts **5**:32
predictive models **5**:39
research needs **5**:41
restoration **5**:41
rising seawater temperatures **5**:33
future research areas **5**:41
conservation status **3**:171–173*T*, **3**:174
continental shelves **5**:52
Convention on Biological Diversity (1930) **2**:318
countryside biogeography studies **2**:356
crustaceans **2**:397–399
dinitrogen-fixing (N_2-fixing) microorganisms **2:608–625**
basic concepts **2**:608–610, **2**:609*F*
biodiversity
dinitrogenase/dinitrogenase reductase **2**:610–611
general discussion **2**:610–611
metabolic diversity **2**:614–615
nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
phylogenetic relationships **2**:611*F*
physiological diversity **2**:614–615
biogeochemical cycles **1**:560–561
emerging patterns
changing environmental conditions **2**:621–623
environmental modifications **2**:619–621
spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
functional role
biological carbon (C) pump **2**:615–618, **2**:616*F*
nutrient cycle linkages **2**:618–619

marine ecosystems (*continued*)
quantification methods **2**:615–618, **2**:617*F*
Redfield ratio **2**:617–618, **2**:617*F*
habitat characteristics **2**:609*F*
research summary **2**:623
ecosystem function measurements **3:64–71**
basic concepts **3**:64
biogeochemical cycles **3**:68–70
comparison studies **3**:64–65
new and regenerated production **3**:69–70
nutrient cycling **3**:68–70
primary production
carbon-14 (^{14}C) measurements **3**:66
carbon-based measurements **3**:65–66
data interpretation and analysis **3**:66–67
fluorescence measurements **3**:66
measurement approaches **3**:65
oxygen-based measurements **3**:65
secondary production
bioenergetic analysis **3**:67
biomass accrual **3**:67
birth and death rate analysis **3**:67–68
egg ratio analysis **3**:68
functional type analysis **3**:68
instar analysis **3**:68
measurement approaches **3**:67
stable isotope studies **3**:70–71
ecosystem services **5:451–468**
accounting assessment models
basic concepts **5**:455–457
revealed references methods **5**:456–457
stated preference methods **5**:457
American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries **5**:464–465, **5**:464*F*
challenges **5**:455
coastal zone management (Belize) **5**:465, **5**:465*F*
conceptual framework
challenges **5**:455
coupled social–ecological systems model **5**:453–455, **5**:454*F*
integrated ecosystem assessments (IEAs) **5**:454–455
modeling approaches and uses **5**:453
coral reef protection and restoration **5**:465–466
cross-system services **3**:96–98
cultural ecosystem services **5**:451–452, **5**:452*T*
decision-support tools **5**:459–463, **5**:460–462*T*, **5**:463*F*
future outlook **5**:466
multiple ecosystem services **5**:459–463
production function models
basic concepts **5**:457–459
bioeconomic models **5**:457–459
cost–benefit analyses **5**:459–463
tradeoffs **5**:459–463
provisioning services **5**:451–452, **5**:452*T*
regulating services **5**:451–452, **5**:452*T*
single ecosystem services **5**:457–459
supporting services **5**:451–452, **5**:452*T*
valuation measures **5**:457
wind farms **5**:466
ecosystem unit classifications **5**:45–46, **5**:46*F*
eutrophication **5:23–31**
anthropogenic effects **5**:23–24, **5**:53–54
aquatic vascular plants **3**:360–361
contaminant–stress relationship **5**:29–30
food web complexity **5**:24–25
harmful algal blooms **5**:25
lake and pond clean-up effects **5**:24
macroalgae **3**:358–359
macrophytes **3**:360–361, **3**:360*F*, **3**:361*F*
microfauna **3**:363–364, **3**:363*F*
nearshore ocean nutrient enrichment response
harmful algal blooms **5**:25
high nutrient, low chlorophyll (HNLC) concept **5**:26, **5**:26*T*
high nutrient, low growth (HNLG) concept **5**:26–27, **5**:26*T*, **5**:27*F*
nitrogen loading **5**:25, **5**:26*F*
nutrient loading **5**:23–24
species shifts **3**:353*F*
stoichiometric effects
nutrient ratio changes **5**:28–29, **5**:28*T*, **5**:29*F*
Redfield ratios **5**:27
silicon additions **5**:27–28
extinction risks **4**:664
fish diversity
Antarctic region **3**:474
Arctic region **3**:474
characteristics **3**:473
deep sea fish **3**:475–476, **3**:475*F*
Eastern Atlantic region **3**:474
Eastern Pacific region **3**:473
Indo–West Pacific region **3**:473
pelagic regions **3**:475
temperate regions **3**:474–475
Western Atlantic region **3**:473–474
food chains **5**:50–51
functional roles **5**:50–51
global change effects **4**:664, **5**:257–258
international organizations **4**:329–330
mangrove ecosystems **5**:45–46
characteristics **5**:14–15
crustaceans
characteristics **5**:15–16
ecosystem engineers **5**:16–17
grapsid crabs **5**:15–16, **5**:15*F*
hermit crabs **5**:16
mud crabs **5**:16
number of species **5**:20*T*
ocypodid crabs **5**:16, **5**:16*F*
shrimp **5**:16
fish **5**:15, **5**:15*F*, **5**:20*T*
molluscs
bivalves **5**:17
meiofauna **5**:18
number of species **5**:20*T*
snails **5**:17–18, **5**:17*F*
root communities **5**:14–15
marine–land comparisons **5**:46–47, **5**:47*F*
marine microbiology **5:292–298**
Global Ocean Sampling (GOS) program
cross-biome sampling studies **5**:294
fragment recruitment **5**:294–295, **5**:295*F*
functional diversity (FD) **5**:295
research background **5**:293
research outlook **5**:297
sampling sites **5**:293*F*
Sargasso Sea **5**:293–294
taxonomic diversity **5**:297
historical perspective **5**:292–293
Marine Microbial Genome Sequencing Project (MMGSP) **5**:295–297
taxonomic and functional diversity **5**:297
marine protected areas **5:94–104**
benefits
abundance increases **5**:96–97, **5**:97*T*
adjacent fisheries **5**:97
biomass increases **5**:96–97, **5**:97*T*
characteristics **5**:96–97
long-term ecological benefits **5**:97–98
nontarget species **5**:97
reproductive potential **5**:96–97
size increases **5**:96–97, **5**:97*T*
Bowie Seamount Marine Protected Area (Sglaan Kinghlas) **5**:446–447
characteristics and functional role **5**:94–95
climate change threats **5**:100–101
conservation strategies **5**:39–40
ecosystem-based management
characteristics **5**:95
definitions **5**:95
goals and objectives **5**:95
marine spatial planning **5**:96
networks **5**:95–96
effectiveness **5**:100
Endeavor Hydrothermal Vents Marine Protected Area **5**:447
fish conservation **3**:453–454
fish stocks **3**:486
geographic extent **5**:445–446, **5**:446*T*
global commitment **5**:99–100
Great Barrier Reef Marine Park **5**:446
Papahanaumokuakea Marine National Monument **5**:446
socioeconomic benefits **5**:98–99, **5**:98*T*
marine sediments **5:105–115**
benthic organisms and environment **5**:105–107, **5**:106*F*
diversity regulation
deep-sea environments **5**:110–111
shallow water **5**:109–110
ecosystem services
filtration processes **5**:112
general discussion **5**:109
nutrient cycling **5**:111
pollutant cycling **5**:111–112
schematic diagram **5**:111*F*
secondary production **5**:112
shoreline stability **5**:112
species diversity connection **5**:112–113

global diversity patterns **5**:109
habitats
deep-sea habitats **5**:108
intertidal habitats **5**:107, **5**:108*T*
ocean depth ranges **5**:107, **5**:107*F*
subtidal habitats **5**:107–108
macrofauna **5**:106*F*
meiofauna **5**:106*F*
species numbers **5**:108–109
threats and challenges
biodiversity loss **5**:113–114
eutrophication **5**:113*F*, **5**:114
exotic species invasions **5**:113*F*, **5**:114
fishing impacts **5**:113–114, **5**:113*F*
global climate change **5**:113*F*, **5**:114–115
habitat destruction and degradation **5**:113*F*, **5**:114
pollution **5**:113*F*, **5**:114
shoreline modifications **5**:113*F*, **5**:114
nematodes **7**:429, **7**:431
net primary production (NPP) **1**:559, **1**:560*T*
nonanthropogenic environmental changes and variability **5**:61–62
nonextremeophilic Archaea **1**:224*T*
ocean edge
coral reefs **5**:52
intertidal beaches **5**:51
kelp beds **5**:51
rocky shores **5**:51–52
oil spills **3**:255
open ocean/deep sea
benthic ecosystems **5**:53
hydrothermal vents **3**:242, **5**:53
Mid-Ocean Ridge **5**:53
pelagic ecosystems **5**:49*F*, **5**:52–53
pelagic ecosystems **3**:475, **5**:46, **5**:46*F*, **5**:47*F*
planktonic organisms **5**:48*F*
primary production **5**:46
protected area systems **6**:494, **6**:495*F*
research summary **5**:115
survival probabilities **5**:47
symbiotic relationships **5:116–126**
functional roles **5**:117–118
future challenges **5**:125
global distribution **5**:117*F*
historical research **5**:116–117
light-producing symbioses
bioluminescence **5**:121–123, **5**:122*F*
fish **5**:123
squid **5**:123–124
nutrient conversion
nitrogen fixation **5**:121
shipworms **5**:121
worm–whale interaction **5**:121, **5**:121*F*
nutrient-exchange mutualisms
chemosynthetic symbioses **5**:120–121, **5**:120*F*
hydrothermal vents **5**:120–121, **5**:120*F*
nutrient conversion **5**:121
photosynthetic symbioses **5**:118–120
protective compound production
characteristics **5**:124–125
sponges **5**:124–125, **5**:124*F*
structural characteristics **5**:116–117, **5**:117*F*
transmission modes **5**:116–117, **5**:117*F*
threatened birds **7**:181
threats and challenges **5**:447–448
trophic levels **5**:46–47, **5**:48*F*, **5**:49*F*, **5**:50–51
Marine Ecosystems Research Laboratory (MERL) **5**:25
marine invertebrates **4:379–391**
average lifespan **4**:656*T*
benthic large-scale diversity patterns
bathymetric gradients **4**:387, **4**:388*F*
latitudinal gradients **4**:385–387, **4**:386*F*, **4**:387*F*
benthic small-scale diversity patterns **4**:387–389, **4**:388*T*, **4**:389*F*
classifications and sampling methods
benthic species **4**:381–382, **4**:381*T*, **4**:382*F*
pelagic species **4**:382–383, **4**:383*T*
cold seeps **4**:391
diversity control processes **4**:380–381
endangered species **3:199–204**
causal factors **3**:199–202
definition **3**:199
extinctions **3**:202*T*
IUCN criteria **3**:200–201*T*
recovery strategies **3**:203
research challenges **3**:202–203
species examples **3**:199–202, **3**:201*T*, **3**:202*T*
susceptible habitats and species **3**:202, **3**:202*T*, **3**:203*T*
fossil record **4**:389–391, **4**:390*F*
global diversity **4**:384–385, **4**:385*T*
hydrothermal vents **2**:6*F*, **4**:391
macrofauna **4**:380–381, **4**:381*T*
megafauna **4**:380–381, **4**:381*T*
meiofauna **4**:380–381, **4**:381*T*
ocean characteristics **4**:379–380
pelagic environment
benthic biomass distribution **4**:380–381, **4**:380*F*
pelagic diversity patterns **4**:385, **4**:386*T*
physical characteristics **4**:379, **4**:380*T*
Phanerozoic diversity **3**:683–684, **3**:685*F*, **3**:686*F*
spatio-temporal mosaic theory **4**:388–389, **4**:389*F*
taxonomy and species diversity **4**:383–384, **4**:384*T*
ultraviolet radiation effects **7**:300
Marine Mammal Protection Act (1972) **1**:397, **3**:713*T*, **3**:716, **5**:64–66
marine mammals **5:64–93**
cetaceans
characteristics **5**:65*T*, **5**:66–67
evolutionary history **5**:68–69
taxonomy **5**:65*T*
desmostylians
characteristics **5**:68
evolutionary history **5**:69–70
extinction/near extinction over evolutionary time **5**:75
extinctions
climate change effects **5**:89–91
contributing factors
body size and dimensions **5**:70
fishery–diet/habitat overlap **5**:72
hauling behavior predictability **5**:70–71
rapid recovery potential **5**:71
sea surface-based respiration **5**:70
spatiotemporal distribution predictability **5**:70
subsistence and commercial market demand **5**:71
toxin and contaminant bioaccumulation potential **5**:71–72
geographic regions **5**:92
global change effects **4**:664
identification/monitoring-related hindering factors
availability bias **5**:72
environmental uncertainty effects **5**:73–74
lack of consistent objective criteria **5**:74
observer bias **5**:72–73
population size survey data **5**:73
survey costs **5**:74
vital demographic parameters **5**:73
imminent extinction probability
beluga whales **5**:77*T*, **5**:79–80
blue whales **5**:81*T*, **5**:82
California sea otter **5**:81*T*, **5**:83–84
general discussion **5**:76
Mediterranean monk seals **5**:77*T*, **5**:80
Northern sea otter **5**:81*T*, **5**:84–86
right whales **5**:76–78, **5**:77*T*
species and subspecies **5**:77*T*
Steller sea lions **5**:81*T*, **5**:82–83
vaquitas **5**:77*T*, **5**:80
western North Pacific gray whales **5**:77*T*, **5**:78–79
West Indian manatee **5**:81*T*, **5**:83
major anthropogenic extinctions
North Atlantic gray whale (*Eschrichtius robustus*) **5**:76
species level **5**:75
subspecies level **5**:76
minimization strategies **5**:92–93
natural versus unnatural extinctions **5**:74
patterns and case studies
desmostylians **5**:75
extinction/near extinction over evolutionary time **5**:74–75
general discussion **5**:74–75
major anthropogenic extinctions **5**:75
Odobenidae **5**:74–75
sirenians **5**:75
recovery evidence
bowhead whales **5**:86–87, **5**:87*T*
elephant seals **5**:87*T*, **5**:88
general discussion **5**:86

marine mammals (*continued*)
gray whales **5**:87–88, **5**:87*T*
humpback whales **5**:87, **5**:87*T*
species and subspecies **5**:87*T*
research summary **5**:91–92
scientific whaling **5**:88–89
significant extinction probability
general discussion **5**:80–82
species and subspecies **5**:81*T*
species level extinctions
baiji (Yangtze River dolphin/*Lipotes vexillifer*) **5**:75
Caribbean monk seal (*Monachus tropicalis*) **5**:75
Japanese sea lion (*Zalophus japonicus*) **5**:75–76
Steller's sea cow (*Hydrodamalis gigas*) **5**:76
marine otters
characteristics **5**:68
evolutionary history **5**:70
pinnipeds
characteristics **5**:65*T*, **5**:67
evolutionary history **5**:69
taxonomy **5**:65*T*
polar bears
characteristics **5**:68
evolutionary history **5**:70
sirenians
characteristics **5**:65*T*, **5**:67–68
evolutionary history **5**:69
extinction/near extinction over evolutionary time **5**:75
taxonomy **5**:65*T*
taxonomy
characteristics **5**:65–66, **5**:65*T*
definition **5**:64–66
habitat boundaries **5**:65–66
MarineMap program **5**:460–462*T*
Marine Microbial Genome Sequencing Project (MMGSP) **5**:295–297
marine otters
characteristics **5**:68
evolutionary history **5**:70
taxonomy **5**:65*T*
marine plankton
characteristics and functional role **4**:667–668
conservation practices **6**:77
ecosystem services **6**:77
macroscopic patterns **4:667–680**
basic concepts **4**:668
biological oceanography **4**:668–669
comparative ecological analyses **4**:668
macroscopic assembly
functional classifications **4**:673
phylogenetic classifications **4**:673–675, **4**:675*F*
Redfield ratio **4**:675–677, **4**:676*F*
size classifications **4**:671–673
stoichiometric emergence **4**:675–677, **4**:676*F*
temperature dependence **4**:677
trophic pyramid **4**:670–671
phytoplankton distribution **4**:668–669
regional and global change effects **4**:677–678
research summary **4**:678–679
temperature dependence
biochemical reactions **4**:677
biogeochemical cycles **4**:677
whole-organism metabolism **4**:677
variability measures
accumulating method **4**:670, **4**:671*F*, **4**:672*F*
coarse graining method **4**:669–670, **4**:669*F*
multiscale analysis **4**:669
marine toad **3**:167
marine viruses **5:127–144**
biodiversity considerations **5**:127–129
classification systems **5**:129, **5**:129*F*, **5**:134*T*
genomic characteristics **5**:128, **5**:129–132, **5**:129*F*, **5**:130*F*, **5**:131*F*, **5**:134*T*
historical research **5**:127
phenotypic and genotypic diversity
genome size **5**:140–141
host range **5**:140
metagenomics
DNA-containing viruses **5**:142
general discussion **5**:141–142
RNA-containing viruses **5**:142
morphology **5**:140
single-gene surveys **5**:141, **5**:141*F*
taxonomic classifications **5**:140
phenotypic diversity **5**:129–132, **5**:130*F*, **5**:131*F*, **5**:134*T*
replication methods **5**:127–129, **5**:128*F*
research summary **5**:142
seawater concentrations **5**:132–133, **5**:132*F*, **5**:133*F*
taxonomic diversity
dsDNA viruses
Adenoviridae **5**:134*T*, **5**:135
Alloherpesviridae **5**:134*T*, **5**:135
Caudovirales **5**:130*F*, **5**:134–135, **5**:134*T*
characteristics **5**:134–135, **5**:134*T*
Corticoviridae **5**:134*T*, **5**:135
Herpesvirales **5**:130*F*, **5**:134*T*, **5**:135
Herpesviridae **5**:134*T*, **5**:135
Iridoviridae **5**:134*T*, **5**:135
Malacoherpesviridae **5**:134*T*, **5**:135
Mimiviridae **5**:130*F*, **5**:134*T*, **5**:135
Myoviridae **5**:130*F*, **5**:134–135, **5**:134*T*
Nimaviridae **5**:130*F*, **5**:134*T*, **5**:135–136
Phycodnaviridae **5**:134*T*, **5**:136
Podoviridae **5**:130*F*, **5**:134*T*, **5**:135
Poxviridae **5**:134*T*, **5**:136
Siphoviridae **5**:130*F*, **5**:134*T*, **5**:135
unassigned viruses **5**:134*T*, **5**:136
dsRNA viruses
Birnaviridae **5**:134*T*, **5**:137
characteristics **5**:134*T*, **5**:137
Reoviridae **5**:134*T*, **5**:137
general discussion **5**:133–135
negative-sense ssRNA viruses
Bunyaviridae **5**:130*F*, **5**:134*T*, **5**:139–140
characteristics **5**:134*T*, **5**:139
Mononegavirales **5**:134*T*, **5**:139
Orthomyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Paramyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Rhabdoviridae **5**:130*F*, **5**:134*T*, **5**:139
positive-sense ssRNA viruses
Alvernaviridae **5**:134*T*, **5**:138
Astroviridae **5**:134*T*, **5**:138
Bacillarnavirus **5**:134*T*, **5**:138
Caliciviridae **5**:134*T*, **5**:138
characteristics **5**:134*T*, **5**:137
Dicistroviridae **5**:134*T*, **5**:137
Labyrnaviridae **5**:134*T*, **5**:138
Marnaviridae **5**:134*T*, **5**:137–138
Metaviridae **5**:134*T*, **5**:140
Nidovirales **5**:134*T*, **5**:137
Nodaviridae **5**:134*T*, **5**:138–139
Picornavirales **5**:134*T*, **5**:137
Roniviridae **5**:134*T*, **5**:137
Togaviridae **5**:134*T*, **5**:139
ssDNA viruses
Anelloviridae **5**:134*T*, **5**:136
characteristics **5**:134*T*, **5**:136
Circoviridae **5**:134*T*, **5**:136
Inoviridae **5**:130*F*, **5**:134*T*, **5**:136
Microviridae **5**:134*T*, **5**:136
unassigned viruses **5**:134*T*, **5**:136–137
Marintaman Mengalong **1**:276*T*
Marion Island **1**:180–181, **1**:186*T*
Markea **3**:522*T*
market economy **5:145–155**
bioprospecting **5**:148
challenges **5**:152–154
command-and-control policies **5**:150, **5**:152–154
definitions **5**:145
ecotourism **5**:147–150
green products **5**:149
human well-being **1:435–439**
biodiversity loss **1**:435–436
commodity production **1**:436
direct and indirect valuation methods **1**:435–436
ecosystem services **1**:436
esthetic, spiritual, cultural, and recreational benefits **1**:437
insurance benefits **1**:436
internalized benefits **1**:438–439
internalized values **1**:438
knowledge benefits **1**:437
market efficiency **1**:437
option value **1**:436–437
public goods **1**:437–438
risk factors **1**:436
uncertainty factors **1**:436–437
private ownership market economy **5**:145–147
public-based market approaches **5**:150–152
public goods **5**:146

summary discussion **5**:154
sustainable use strategies **5**:147–150
markhor **3**:197
Mark Trail (cartoon) **2**:15
marlin **3**:470
Marmaronetta angustirostris **7**:185
marmosets **4**:690*T*, **6**:251*T*
Marmota flaviventris **5**:722, **7**:36
Marmota marmota **7**:36
marmots **7**:36
Marnaviridae **5**:134*T*, **5**:137–138
Marquesas **5**:592*F*, **5**:595*F*
Marquesia macroura **1**:68
Marseillevirus **6**:585, **6**:585*F*
Marshallian demand function **7**:321–322
Marshall Islands **6**:507*T*
marsh andromeda **1**:136
marsh elder **2**:663–664
marshes
climate change effects **6**:150–151
conservation status **3**:171–173*T*
definitions and terminology **7**:369*T*
global extent **7**:370*T*
grazing effects **5**:455*F*
habitat loss and degradation **7**:371*T*
nature-based coastal protection strategies **5**:453*T*, **5**:454*F*, **5**:455, **5**:455*F*
plant species richness **7**:398*F*
marsh flies **3**:489–491*T*
Marsh, George Perkins **2**:280–281, **3**:98–99, **4**:652
marsh grasses **5**:45–46
marsh harrier **5**:474
marsh tits **1**:611–612
Marsilia polycarpa **1**:164
Marsippospermum reichei **4**:107
marsupials
adaptive radiation **1**:22, **1**:23*F*
characteristics **3**:188, **4**:683, **4**:686–687, **7**:341
cranial and skeletal characteristics **4**:683–684
Dasyuromorphia **4**:687–688, **4**:687*T*
Didelimorphia **4**:686–687, **4**:687*T*
dietary tolerance **5**:225*T*
Diprodontia **4**:687*T*, **4**:688–689
evolutionary history **2**:583*F*
extinction theories **2**:517
Microbiotheria **4**:687, **4**:687*T*
Notoryctemorphia **4**:687*T*, **4**:688
Paucituberculata **4**:687, **4**:687*T*
Peramelemorphia **4**:687*T*, **4**:688
phylogenetic relationships **4**:686–687, **5**:1*F*
reproductive characteristics **4**:684, **4**:684*F*
RNA editing **5**:576–577
species diversity **1**:686, **4**:687*T*
threat percentage **3**:195
transboundary protected areas **6**:489*T*
martens
boreal forests **1**:631–632
skull size comparisons **1**:692*F*
species diversity **1**:685
Martes pennanti **4**:513–514
Martinique **4**:560*F*, **6**:519
Marxan algorithm **2**:501*F*, **5**:460–462*T*
Massachusetts Forest Stewardship Program **7**:20–21
mass effect **3**:145
mass extinctions **5:156–166**
anthropogenic impacts **2**:278–279
basic concepts **3**:415–418
biodiversity generation
fossil record **1**:430
lineages-through-time plot **1**:432*F*
biogeographic impacts **3**:421–422
causal factors **3**:418, **5**:2–3
climatic changes **3**:422–423
crustaceans **2**:413–415
data analysis and interpretation **5**:157–158, **5**:158*F*, **5**:159*F*, **5**:160*F*
extinction theories **3**:418–421
fossil diversity curve **5**:157*F*
global biodiversity impacts **3**:689–690
historical research **5**:156–157, **5**:157*F*
kill curve **5**:160–163, **5**:163*F*
long-term consequences **3**:421
magnitude measures
end-of-geologic-period extinctions **5**:158, **5**:161*F*
end-Permian mass extinction **5**:158, **5**:161*F*
extinction events **5**:158–160, **5**:161*F*
extinction intensities **5**:162*F*
general discussion **5**:158
major mass extinctions **3**:416*T*
modern diversity crisis **5**:165–166
modern extinctions **5:362–372**
current on-going extinctions **5**:371
density-related traits **5**:370–371
extinction cascades **5**:367–368
first-contact extinctions
basic concepts **5**:365
Madagascar **5**:365
New Zealand **5**:366
Pacific islands **5**:365–366
historical extinctions
extinction paradoxes **5**:367
Hawaii **5**:367
New Zealand **5**:366–367
number of species **5**:368*T*
introduced species **5**:370
population dynamics **5**:370
postglacial extinctions
Africa **5**:364–365
Australia **5**:363–364, **5**:364*F*
North America **5**:363
research challenges
general discussion **5**:368
geographical bias **5**:369
inadequate documentation **5**:368–369
inherent biological limitations **5**:369
methodological limitations **5**:369
number of species **5**:368, **5**:368*T*
taxonomic and habitat bias **5**:369
uneven documentation **5**:369
species lifespans **5**:362–363, **5**:362*T*
theoretical perspectives **5**:369–370
time factors **5**:371
notable mass extinctions **5:167–177**
background information **5**:167
great mass extinctions
biodiversity loss **4**:657*F*
end-Cretaceous **5**:168*F*, **5**:169*F*, **5**:173–175, **5**:174*F*
Late Devonian **5**:168*F*, **5**:169–170, **5**:169*F*
Late Ordovician **5**:167–169, **5**:168*F*, **5**:169*F*
Late Permian **5**:168*F*, **5**:169*F*, **5**:170–172
Late Triassic **5**:168*F*, **5**:169*F*, **5**:172–173
research challenges **5**:167–169
lesser mass extinctions
Cenozoic **5**:176
Mesozoic **5**:169*F*, **5**:176
Neoproterozoic **5**:168*F*, **5**:175–176
Paleozoic **5**:168*F*, **5**:175–176
significance **5**:175–176
Phanerozoic Eon **1**:493, **1**:493*F*, **1**:494*T*
periodicity hypothesis **3**:422–423, **3**:424*F*, **5**:160, **5**:162*F*
Permian–Triassic mass extinction **5**:4
Phanerozoic Eon **1**:412–413, **1**:413*F*, **1**:493, **1**:493*F*, **1**:494*F*, **1**:494*T*, **3**:416*T*, **3**:419*F*
quantification **3**:417
rarefaction curves **1**:494*F*
recoveries
characteristics **5**:164
delayed diversification **5**:164
disaster species **5**:164–165
large-scale rebounds **5**:164, **5**:165*F*
Lazarus species **5**:164–165
research summary **5**:166
scenarios
abrupt extinction **5**:157
general discussion **5**:157
gradual extinction **5**:157
stepwise extinction **5**:157
self-organized criticality **5**:160–163
temporal and biogeographic scales **5**:3
victim–survivor patterns **5**:163–164
mass media **4**:654
mass spectrometry **5**:708
Masticophis flagellum **4**:494, **6**:406
Mastigamoeba **6**:291
Mastigella **6**:291
Mastigocladus **4**:119
Mastigophora **5**:669–670*T*
Mastigoproctus giganteus **1**:211–212, **1**:211*F*
mastodons **4**:724–725*T*
Mastotermes **4**:409*T*
Mastotermes darwiniensis **4**:417
Mastotermitidae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:422*F*
masu salmon **6**:522
mathematical models **1:565–575**
characteristics and functional role **1**:565–566, **1**:566*F*
correlative models **1**:566–569, **1**:566*F*
future research areas **1**:572–574
stochastic simulation models **1**:569–572, **1**:572*F*
mating systems **5**:507–508, **6**:572–573, **6**:632–633

matsutake mushrooms **3**:629–630
Matthews, Anne **4**:653
Matthiessen, Peter **4**:653
Maud Island frog **4**:194*T*
Maui Parrotbills **1**:29*F*
Mauritania
  botanical gardens **6**:519
  deforestation **2**:449*T*
  grass and shrub savannas **1**:67
  urbanization **4**:560*F*
Mauritia carana **4**:286–287
Mauritia flexuosa **1**:164, **3**:113, **4**:286–287, **4**:288*T*, **7**:272
Mauritiella aculeata **1**:163
Mauritius
  botanical gardens **6**:519
  living dead **6**:81
  urbanization **4**:560*F*
Mauritius fody **7**:188
Mauritius kestrel **6**:455, **7**:188–189
Mauritius National Plant Protection Legislation **3**:209
Maxillaria **6**:316*T*
Maxillopoda **2**:400–401, **2**:401*T*
maximum likelihood analysis **6**:18
maximum sustainable yield (MSY)
  fisheries science **3**:483, **6**:423–424, **6**:424*F*, **6**:425*F*
  population dynamics **3**:3
  wildlife management **7**:403–405, **7**:404*F*
Maxwell–Boltzmann equation **4**:677
mayflies
  critical pH levels **6**:153*F*
  freshwater ecosystems
    acidification effects **3**:584–585
    organic pollution **3**:580–581, **3**:581*F*
    threatened species **3**:579–580, **3**:579*T*
  number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
  surface water acidification **1**:10*T*
  wetland ecosystems **7**:389–391
Mayorella **6**:291
Mayotte **6**:507*T*
Mayr, Ernst
  adaptive theory **1**:16
  biological species concept **3**:394, **6**:620
  evolutionary systematics **7**:101–102
  founder event principle **6**:624
  genetics research **2**:424–425, **6**:3–4
  species concept theory **6**:681–682
  subspecies taxonomy **7**:46
Mazocraeidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Mbaracayau Forest Biosphere Reserve **6**:267–269, **6**:267*F*
Mbura National Park **2**:225
McDonald Island **1**:180–181, **1**:186*T*
McDonald–Kreitman (MK) test **2**:634–635, **2**:635*T*
McPhee, John **4**:653
meadowlarks **4**:490*F*
meadows **4**:577
Meadows, Dennis **6**:205–206
meadow voles **1**:365–368
mean monthly water temperatures **1**:582*F*
mechanical disturbances **2**:563–564
Mechanitis isthmia **1**:651*F*
Mecoptera **4**:308–312, **4**:309*T*, **4**:392*T*
Medicago sativa **2**:391
medicinal plants **1**:359, **5**:705–706, **6**:518–519
Mediophyceae **1**:476*F*
Mediterranean Basin
  dry regions **1**:262–263*T*, **1**:268–269
  endemism **3**:233
  invasive species **4**:345–346
  Mediterranean-type climate ecosystems
    characteristics **5**:212–213
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:217–218
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:218
    vertebrate diversity **5**:213*T*, **5**:218–219
  native crops **2**:662–663
  nitrogen deposition **5**:526*T*
  plant species **6**:82–83, **6**:82*T*
  regional cultivars **3**:32*F*
  wheat production **1**:86, **1**:87–88
Mediterranean monk seals **5**:77*T*, **5**:80
Mediterranean region
  endangered plant species **3**:206–208, **3**:207*F*
  habitat and vegetation diversity **3**:335*T*, **3**:336–340*T*
  late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*
  temperate forests **7**:116*T*
  vegetation surveys **3**:341–344*T*
Mediterranean sclerophyll **1**:262–263*T*
Mediterranean scrub **6**:600–604*T*
Mediterranean Sea **2**:356, **3**:171–173*T*
Mediterranean-type climate ecosystems **5**:**212–222**
  Australia
    physical characteristics **5**:220–221
    vascular plant diversity **5**:213*T*, **5**:221
    vertebrate diversity **5**:213*T*, **5**:221
  California
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:214–215
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:215
    vertebrate diversity **5**:213*T*, **5**:215
  Cape Region (South Africa)
    physical characteristics **5**:219
    vascular plant diversity **5**:213*T*, **5**:219–220
    vertebrate diversity **5**:213*T*, **5**:220
  characteristics **5**:212–213, **5**:213*T*
  Chile
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:215–216
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:216–217
    vertebrate diversity **5**:213*T*, **5**:217
  global distribution **5**:213*F*, **5**:213*T*
  Mediterranean Basin
    physical characteristics **5**:217–218
    vascular plant diversity **5**:218
    vertebrate diversity **5**:213*T*, **5**:218–219
  natural disturbances **5**:213–214
  physical characteristics **5**:213*T*
  speciation patterns **5**:214
  species richness **5**:213*T*
  summary discussion **5**:221–222
Medullosa **5**:618*F*
meerkats **1**:686
Megacephala aequinoctialis **1**:250*T*, **1**:252*F*
Megacephala bilunata **1**:250*T*, **1**:251*F*
Megacephala klugi **1**:250*T*, **1**:252*F*
Megacephala sobrina punctata **1**:250*T*, **1**:252*F*
Megacephala spinosa **1**:250*T*, **1**:252*F*
Megacerops **5**:224*T*
Megachiroptera **4**:691–692
Megadalapis **5**:365
Megadalapis edwardsi **5**:365
megafauna
  living dead **4**:591–593, **4**:593*F*
  marine invertebrates **4**:380–381, **4**:381*T*
Megagnathotermes **4**:409*T*
megaherbivores **5**:**223–239**
  conservation challenges **5**:234–236
  dietary tolerance **5**:223–226
  extant species **5**:225*T*
  extinctions **5**:225*T*, **5**:231–233, **5**:232*F*
  future outlook **5**:236
  grazing impacts **5**:227–228
  human impacts **5**:233–234
  late Pleistocene extinctions **5**:231–233, **5**:232*F*
  life history patterns **5**:228–231, **5**:229*F*, **5**:230*F*
  nutritional physiology **5**:223–226
  occurrences and characteristics **5**:223, **5**:224*T*
  population biomass **5**:227–228, **5**:227*F*
  population regulation **5**:228–231, **5**:229*F*, **5**:230*F*
  predation vulnerabilities **5**:226–227
  vegetation transformations **5**:227–228, **5**:227*F*, **5**:233
Megaladapis **4**:728
Megalania **4**:727–728
Megalania prisca **5**:363–364, **5**:364*F*
Megalapteryx didinus **5**:366
Megaloceros giganteus **4**:729*F*, **5**:225–226
Megalocerus algericus **5**:364–365
Megalodiscus **7**:450–451*F*
Megalonychidae **4**:727*T*, **6**:309*T*
Megalonyx **4**:724–725*T*
Megaloptera **4**:308–312, **4**:309*T*, **4**:371*T*
Megalopygidae **5**:396
Megalotragus **5**:231–233
Megalotragus priscus **5**:364–365
Megalotylopus **4**:724–725*T*
Megalyridae **4**:178
Megalyroidea **4**:178
Megamerinidae **3**:489–491*T*
Meganteron **4**:724–725*T*
megaplankton **5**:133*F*
Megapodius pritchardii **7**:188
Megaptera novaeangliae **2**:7–8, **5**:87, **5**:87*T*
Megascolides australis **3**:225

Megatheriidae **4**:727*T*, **5**:225*T*
Megatherium **4**:727*T*
Megatherium americanum **5**:225*T*
Megavirus chilensis **5**:135
Mehadiella paveli **7**:60
Mehler Peroxidase reaction **6**:13
Mehuin Preserve (Chile) **5**:57, **5**:59*F*
Meinertellidae **1**:247
meiofauna
  mangrove ecosystems **5**:18
  marine invertebrates **4**:380–381, **4**:381*T*
meiosis **2**:321, **2**:324–325
meiotic drive **1**:16
Mekong Delta **1**:276*T*
Mekong River **1**:512, **7**:147
Melaleuca **5**:221, **7**:125, **7**:399
Melaleuca quinquenervia **4**:360, **6**:96, **6**:442, **7**:399
Melamprosops phaeosoma **6**:350
Melampsora lini **2**:139
Melampyrum arvense **4**:570–572
Melanochromis **1**:25*F*
Melanoplus spretus **3**:221–222
Melanopsis buccinoidea **5**:471
Melanosuchus niger **6**:401, **7**:229*T*
Melanthera marlothiana **1**:67
Melastoma **1**:263
Melastomataceae **3**:522*T*, **6**:316*T*
Meles meles **4**:693
Melia azedarach **1**:590, **5**:476
Meliaceae **1**:263, **5**:11*T*, **6**:316*T*
Melicitus **3**:522*T*
Melilotus officinalis **5**:712*T*
Melinis minutiflora **2**:379–381, **4**:364, **6**:125
Melitaea cinxia **4**:517
Melittomma insulare **5**:603–604
Meloidae **1**:327*T*
melons
  native crops **2**:662–663
  regional cultivars **3**:32*F*
Melospiza melodia **4**:490, **4**:491*F*
Melville, Herman **4**:653
Melyridae **1**:327*T*
membrane hybridization **5**:264*T*, **5**:267
Mendel, Gregor **2**:424–425, **4**:651–652, **6**:1–2
Mendelian populations **6**:168
meniscus midges **3**:489–491*T*
Mentzelia pumila **2**:430
Mephitidae **6**:309*T*
mercury-based fungicides **5**:695
mercury deposition **1:1–14**
  emissions and deposition **1**:2–5, **1**:4*F*
  forest ecosystems **1**:5*T*
  freshwater ecosystems
    acidic deposition–mercury concentration linkage **1**:13
    aluminum concentrations **1**:9–10
    aquatic biota response **1**:9–10, **1**:10*T*, **1**:11*F*
    contamination impacts **1**:10–13, **1**:12*F*
    freshwater fish **1**:10–13, **1**:12*F*, **1**:13
    pH changes **1**:9–10, **1**:10*T*
    schematic diagram **1**:11*F*
    surface water acidification **1**:8–9, **1**:10*T*
  geographic distribution **1**:1–5, **1**:3*F*
  occurrences and sources **1**:1–5
  research background **1**:1
Mercury Deposition Network (MDN) **1**:5
mergansers **1**:604*F*
Merguia **5**:16
Meridiopitymys henseli **1**:702
Meriones sacramenti **5**:472
merlin falcon **5**:474
Merops albicollis **5**:603
Merops apiaster **5**:474
Merops orientalis **5**:463*T*
Merops supeciliosus **5**:473
Merriam, C. Hart **1**:578–579
Merycopotamus **5**:224*T*
Mesembryanthemaceae **1**:486*T*
Mesembryanthemum crystallinum **4**:360
Mesoamerican plant species **6**:82*T*
Mesoamerican poeciliid fish **7**:348, **7**:348*F*
Mesobuthus martensi **1**:208
Mesobuthus spp. **1**:208
Mesocapromys angelcabrarai **4**:714*F*
Mesocapromys auritus **4**:714*F*
Mesocapromys nanus **4**:714*F*
Mesocapromys sanfelipensis **4**:714*F*
Mesocyclops **6**:26
mesofauna **6**:586, **6**:590*T*
Mesonerilla prospera **3**:202*T*
Mesonychoteuthis hamiltoni **5**:406
mesophiles **7**:328
mesoplankton **4**:383*T*, **4**:460, **5**:133*F*
Mesoplodon densirostris **3**:192–193
Mesopotamian marshlands restoration (case study) **7**:374
Mesopropithecus **5**:365
Mesosetum **2**:376*F*
Mesostigmata **1**:216, **6**:587–588
mesotheles **1**:206, **1**:206
Mesotriton alpestris **4**:516–517
mesotrophic level **3**:563*T*
Mesoxylon **5**:618*F*
Mesozoa **3**:684*F*, **5**:669–670*T*
Mesozoic era
  boundary age **5**:1*T*
  fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*
  marine diversity **5**:168*F*, **5**:169*F*
  marine invertebrates **4**:390–391, **4**:390*F*
  mass extinctions **1**:413*F*, **3**:419*F*, **5**:157*F*, **5**:169*F*, **5**:176
  molluscs **5**:374*F*
  temperate forests **7**:123*F*
  terrestrial plant communities
    abundance patterns **6**:74*F*
    evolutionary development **6**:65–67, **6**:66*F*
    extinction rates **6**:73*F*
    origination rates **6**:72*F*
  terrestrial plant evolution **6**:70–71
mesozooplankton **4**:670–671
mesquite
  desert ecosystems **2**:476
  fire regime alterations **6**:342–343
  multiple cropping systems **1**:120
Messerschmidtia **1**:277
Messor messor **5**:483
Mesua **5**:623*F*
Metachirus nudicaudatus **2**:587*F*
metadata **1:440–445**
  classifications **1**:441
  emerging technologies
    digital media **1**:444–445
    scientific workflow systems **1**:445
    Standard Generalized Markup Language (SGML) **1**:444
    World Wide Web (WWW) **1**:444
  historical background **1**:440–441
  purposes
    data discovery
      annotation **1**:442
      data access **1**:441
      data portals **1**:441
      data provenance **1**:442
      Globally Unique Identifier (GUID) **1**:441–442
      quality control **1**:442
    scientific inquiry **1**:441
  structured metadata
    Access to Biological Collection Data (ABCD) standards **1**:443
    Darwin Core **1**:443
    Ecological Metadata Language (EML) **1**:443
    eXtensible Markup Language (XML) **1**:442–443
    geospatial data standards **1**:443–444
    machine integration challenges **1**:442
metalaxyl **5**:695
metalimnion **3**:560, **4**:374*F*
metals **1**:563
Meta menardi **7**:52–53
Metapolybia aztecoides **2**:326
metapopulation theory **5:240–251**
  basic concepts
    local extinctions **5**:241–242
    multispecies dynamics
      competition **5**:243–244
      predators and parasitoids **5**:242–243
      succession **5**:243–244, **5**:243*F*
    patchiness **5**:240–241
    single species dynamics **5**:241–242
  extinction mechanisms **4**:660–661
  general discussion **5**:240
  research results
    conservation **5**:250
    nonequilibrium dynamics **5**:250
    population regulation **5**:250
  spatial patterns **3**:5–6
  spatial variation effects
    basic concepts **5**:244–248
    competition **5**:248–249, **5**:249*F*, **5**:250*F*
    predators and parasitoids **5**:248, **5**:248*F*
    single species dynamics **5**:245–248, **5**:245*F*, **5**:246*F*, **5**:247*F*
Metaryzium anisopliae **5**:694
Metasequoia **2**:47
Metasequoia glyptostroboides **1**:272–273
Metatheria **4**:683
  see also marsupials
Metaviridae **5**:134*T*, **5**:140
Metazoa
  Cambrian explosion **3**:683, **3**:684*F*
  evolutionary development **2**:644–645

Metazoa (*continued*)
group selection **2**:323
photosynthetic symbioses **5**:118–120
reproduction modes **2**:645–646
RNA editing **5**:573–574*T*
social systems **2**:646
temperature tolerance ranges **4**:113–115
metham sodium **5**:695, **5**:696
methane (CH_4)
atmospheric concentrations **1**:282*T*
biodiversity impacts **3**:256–257
biomass emissions **4**:295–296
carbon cycle **1**:679–680, **1**:680*T*
changing concentrations
annual average measurements **1**:285*F*
causal factors **1**:284
methane budget **1**:284
climate change effects **1**:281–282, **3**:407–408
dynamic global vegetation models (DGVMs) **2**:680
ecotoxicological studies **3**:118–119
enhanced greenhouse effect **4**:18–19, **4**:19*F*, **4**:19*T*
fire/fire regimes **3**:441–442
grazing effects **1**:706, **6**:340–341
hydrothermal vent ecosystems **7**:328
mass extinctions **3**:418
past climate change **2**:82–83, **4**:19, **4**:20*F*
proxy assessments **4**:225*F*, **4**:226
radiative forcing effects **4**:20*F*
termites **4**:432–433, **4**:432*T*
Methanobacterium **1**:222*T*
Methanobrevibacter **1**:222*T*
Methanococcoides **1**:222*T*
Methanococcus **1**:222*T*
Methanoculleus **1**:222*T*
methanogenesis, microbial **5**:301*T*
methanogenic Archaea **1**:220–222, **1**:222*T*
Methanogenium **1**:222*T*
Methanogenium frigidum **6**:300
Methanohalobium **1**:222*T*
Methanohalophilus **1**:222*T*
Methanolobus **1**:222*T*
Methanomicrobium **1**:222*T*
Methanoplanus **1**:222*T*
Methanopyrus **1**:222*T*, **2**:640–641
Methanosarcina **1**:222*T*
Methanosphaera **1**:222*T*
Methanospirillum **1**:222*T*
Methanothermus **1**:222*T*
Methanothrix **1**:222*T*
methfuroxam **5**:695
methicillin-resistant *Staphylococcus aureus* **5**:711
methomyl **5**:700*T*
methoprene **5**:694
methoxychlor **5**:693
Methuselah's Beard lichen **1**:591
methyl bromide (CH_3Br) **1**:282*T*, **1**:284–286, **5**:694
methyl chloride (CH_3Cl) **1**:282*T*, **1**:284–286
methylmercury **1**:10–13, **1**:12*F*
methylparathion **5**:693–694
methyl *tert* butyl ether (MTBE) **5**:264, **5**:265*F*
Metopus **6**:31
metribuzin **4**:90*T*
Metridiochoerus **5**:231–233
Metriocnemus knabi **4**:518
Metrioptera roeseli **4**:514
Metrosideros **3**:522*T*, **7**:119
Metrosideros polymorpha **6**:171
Meuse River **7**:152–153
Mexican duck **6**:689–690
Mexican free-tailed bats **5**:313
Mexican mantled howler **6**:256–258*T*
Mexican spider monkey **6**:256–258*T*
Mexico
amphibians **5**:565
biodiversity-based agricultural strategies **1**:123–124
biodiversity richness **1**:501*F*
birds **5**:565
botanical gardens **6**:519
butterflies **1**:657*T*, **1**:658*T*, **5**:565
Chixulub crater **3**:418–421
Comisión Nacional Para el Conocimiento y Uso de la Biodiversidad (CONABIO) **1**:505–506
Convention on Biological Diversity (1930) **2**:313
coral reef ecosystems **1**:500*T*
countryside biogeography studies **2**:353–354
deforestation **1**:502*T*
ecosystem services **7**:356*T*
endangered mammals **3**:194*T*
endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
freshwater fish **5**:565
indigenous peoples **4**:270, **4**:270*T*, **4**:274*F*
institutional collections and capacity **1**:505*T*
land mammals **5**:565
maize production **1**:86
multiple cropping systems **1**:120
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
plants **5**:564–565
protected area systems **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
reptiles **5**:565
seasonal rainfall **2**:465*F*
shifting cultivation systems **7**:232
species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
terrestrial bioregions **5**:563
water funds **7**:356*T*
wheat production **1**:87–88
Miacidae **1**:686
mice
food webs **7**:266*F*, **7**:267*T*
Great American Biotic Interchange (GABI) **2**:584*F*
interspecific competition **2**:186–187, **2**:187*T*
introduced species **1**:80, **1**:82
Michaelis–Menten equation **5**:192
Michelia **1**:270–271
Michener, Charles **7**:102–104
Michigan (Lake) **4**:461*T*
Miconchus **7**:426*F*
Miconia **6**:316*T*
Miconia calvescens **6**:93
Micranthemum micranthemoides **3**:205
microalgae
eutrophication impacts
harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*
long-term human influences **3**:354–356, **3**:356*F*, **3**:357*F*
species shifts **3**:352–354, **3**:353*F*
microarray analysis **3**:664
microbial biodiversity **5:280–291**
background information **5**:280
biological significance **5**:282–283
deep-sea hydrothermal vents
bacterial–invertebrate symbioses **7**:328–329
free-living microorganisms **7**:328
evolutionary development **5**:280–282
future research areas **5**:291
major groups
domain Archaea **5**:286–288, **5**:287*F*, **5**:288*F*
domain bacteria **5**:285*F*, **5**:287*F*
domain Eukarya **5**:288–289, **5**:289*F*
fungi **5**:289–291, **5**:289*F*, **5**:290*F*
protists **5**:288–289, **5**:289*F*
tree of life **5**:285–286, **5**:285*F*
measurement and analysis **5:259–270**
cell counting measurement
characteristics **5**:261
microscopy **5**:261
most probable number (MPN) assays **5**:262
plate counts **5**:261–262
cellular constituent measurements **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
culture-based measurement versus direct measurement **5**:260
environmental heterogeneity **5**:261
future research outlook **5**:269
importance **5**:259–260
measurement challenges **5**:260
measurement methods
cell counting **5**:261
cellular constituents **5**:262–264
characteristics **5**:261
computational challenges **5**:268–269
denaturing/thermal gradient gel electrophoresis (D/TGGE) **5**:264, **5**:264*T*, **5**:265*F*
DNA cloning and sequencing **5**:264*T*, **5**:265–266, **5**:266*F*
enzyme assays **5**:268, **5**:269*T*
fatty acid analysis **5**:267, **5**:268*F*
fluorescent *in situ* hybridization (FISH) **5**:264*T*, **5**:267
intergenic transcribed spacer (ITS) analysis **5**:264–265, **5**:264*T*, **5**:265*F*, **5**:266*F*
membrane hybridization **5**:264*T*, **5**:267
oligonucleotide arrays **5**:264*T*, **5**:267

polymerase chain reaction (PCR) technique **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
process/function measurements **5**:267–268
restriction fragment length polymorphisms (RFLPs) **5**:264*T*, **5**:265
solution hybridization **5**:264*T*, **5**:266–267
substrate utilization patterns **5**:267–268
nucleic acids **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
taxonomic resolution **5**:260–261
units of measurement **5**:259–260
prevalence **5**:280–282, **5**:281*F*
research areas
Delft School tradition **5**:283–284, **5**:283*F*
research background **5**:283
16S ribosomal DNA (rDNA) sequences **5**:284–285
Woesean Reformation tradition **5**:284–285
restoration efforts
diversity evaluations **6**:446–447
establishment criteria **6**:450–451, **6**:450*F*
functional groups
free-living saprobes **6**:448
importance **6**:448
soil food webs **6**:448–449, **6**:449*F*
symbionts **6**:449–450, **6**:450*F*
spatiotemporal relationships **6**:447–448, **6**:447*F*
microbial biogeography 5:271–279
background information
Baas-Becking hypothesis **5**:271
general discussion **5**:271
molecular genetics **5**:271–272
distribution patterns
depth gradients **5**:273–274
distance–decay relationships **5**:274, **5**:274*F*
elevational gradients **5**:273–274
endemism **5**:272, **5**:273*F*
latitudinal gradients **5**:272–273
phylogenetic relationships **5**:273*F*
research background **5**:272
seasonal dynamics **5**:275
taxa–area relationships **5**:274–275, **5**:274*F*
taxa–time relationships **5**:275
future research considerations
habitat type **5**:278
pattern interpretation **5**:277–278
spatial scale **5**:278
taxonomic focus and resolution **5**:277–278
pattern interpretation
challenges **5**:275–276, **5**:276–277
contemporary environmental factors **5**:276, **5**:276*F*
future research considerations **5**:277–278
historical factors **5**:276, **5**:276*F*
hypothetical relationships **5**:275–276, **5**:276*F*
summary discussion **5**:278
microbial insecticides **5**:694
microbial light-driven proton pumps 6:280–285
abundance and diversity **6**:282–283
bacteriorhodopsin relationship tree **6**:281*F*
energy circuit diagram **6**:282*F*
Escherichia coli proton pumps **6**:280–282, **6**:281*F*
future research areas
antenna-assisted proteorhodopsins **6**:283
photochemical reaction cycle **6**:283
historical research **6**:280–282
occurrences **6**:280
phototrophic mechanisms **6**:280–282
spectral tuning **6**:283
Microbiotheria **2**:586, **2**:587*F*, **4**:682*F*, **4**:687, **4**:687*T*
Microcapritermes **4**:409*T*
Microcebus berthae **6**:256–258*T*
Microcebus murinus **4**:692
Microcebus ravelobensis **4**:516, **6**:256–258*T*
Microcebus sambiranensis **6**:256–258*T*
Microcebus tavaratra **6**:256–258*T*
Microcerotermes **4**:409*T*
Microcerotermes parvus **4**:429*F*
Microchiroptera **4**:691–692
Microcosmos (1996) **4**:654
Microcystis **6**:29–30
Microcystis aeruginosa **6**:33*F*
Microdipidops pallidus **4**:43*F*
microfauna **6**:586, **6**:587*F*, **6**:590*T*
microfluidic blocks (MFBs) **2**:694–697, **2**:694*F*
microfungi **1**:142*F*
Microgadus tomcod **3**:123*T*
Microhodotermes **4**:409*T*
Microlepidoptera **5**:386–389, **5**:387–388*T*
Micromeria danaensis **5**:485
Micromesistius poutassou **4**:516
micronekton **4**:383*T*
Micronesia **6**:507*T*
Micronesian atoll **5**:591*F*
Micronesian plant species **6**:82*T*
microorganisms 5:299–308
alpine ecosystems **1**:155, **1**:155*F*
anthropogenic activities and impacts **7**:307*T*
biogeochemical cycles **5**:306–307
biomonitoring processes **3**:122*T*
Census of Marine Life project **2**:15*F*, **2**:21, **2**:21*F*
coral-dwelling microbes **2**:335
crop plant pathogens **1**:80
definition and characteristics **5**:299–300, **5**:299*F*
dinitrogen-fixing (N_2-fixing) microorganisms **2:608–625**
basic concepts **2**:608–610, **2**:609*F*
biodiversity
dinitrogenase/dinitrogenase reductase **2**:610–611
general discussion **2**:610–611
metabolic diversity **2**:614–615
nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
phylogenetic relationships **2**:611*F*
physiological diversity **2**:614–615
biogeochemical cycles **1**:560–561
emerging patterns
changing environmental conditions **2**:621–623
environmental modifications **2**:619–621
spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
functional role
biological carbon (C) pump **2**:615–618, **2**:616*F*
nutrient cycle linkages **2**:618–619
quantification methods **2**:615–618, **2**:617*F*
Redfield ratio **2**:617–618, **2**:617*F*
habitat characteristics **2**:609*F*
research summary **2**:623
ecotoxicity **3**:119–120
eukaryotes/protists
characteristics **5**:305–306
phagocytosis **5**:305
phagotrophic protists **5**:306
phototrophs **5**:305–306
plankton fraction **6**:25–26
symbiotic relationships **5**:306
global nitrogen cycle **5**:307, **5**:307*F*
grassland ecosystems **4**:4
highly replicated laboratory experiments **4**:587, **4**:587*F*
human–microbe relationships **5**:307–308
introduced species **1**:76*T*, **1**:78–79
lake and pond ecosystems **4**:460, **4**:461*T*
marine microbiology 5:292–298
Global Ocean Sampling (GOS) program
cross-biome sampling studies **5**:294
fragment recruitment **5**:294–295, **5**:295*F*
functional diversity (FD) **5**:295
research background **5**:293
research outlook **5**:297
sampling sites **5**:293*F*
Sargasso Sea **5**:293–294
taxonomic diversity **5**:297
historical perspective **5**:292–293
Marine Microbial Genome Sequencing Project (MMGSP) **5**:295–297
taxonomic and functional diversity **5**:297
microbial biodiversity 5:252–258, 5:280–291
background information **5**:252, **5**:280
basic concepts
Archaea **5**:254–255
bacteria **5**:254–255
eukaryotic diversity **5**:255
general discussion **5**:254–255
number of species **5**:254*T*

microorganisms (*continued*)
prokaryotic diversity **5**:255*F*
viruses **5**:255
biological significance **5**:282–283
cell counting measurement
characteristics **5**:261
microscopy **5**:261
most probable number (MPN) assays **5**:262
plate counts **5**:261–262
cellular constituent measurements **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
characterization studies
culture techniques **5**:253–254
genetic methodologies **5**:254
genomic research **5**:252–254
deep-sea hydrothermal vents
bacterial–invertebrate symbioses **7**:328–329
free-living microorganisms **7**:328
ecosystem functions **5**:257
evolutionary development **5**:280–282
future research areas **5**:291
global change effects
microbial diseases **5**:258
terrestrial and marine ecosystems **5**:257–258
maintenance mechanisms
basic principles **5**:255–256
dispersal rates **5**:256–257
disturbances **5**:256
environmental heterogeneity **5**:256
productivity **5**:256
major groups
domain Archaea **5**:286–288, **5**:287*F*, **5**:288*F*
domain bacteria **5**:285*F*, **5**:287*F*
domain Eukarya **5**:288–289, **5**:289*F*
fungi **5**:289–291, **5**:289*F*, **5**:290*F*
protists **5**:288–289, **5**:289*F*
tree of life **5**:253*F*, **5**:285–286, **5**:285*F*
measurement and analysis **5:259–270**
computational challenges **5**:268–269
culture-based measurement versus direct measurement **5**:260
environmental heterogeneity **5**:261
future research outlook **5**:269
importance **5**:259–260
measurement challenges **5**:260
measurement methods **5**:261
taxonomic resolution **5**:260–261
units of measurement **5**:259–260
measurement methods
cell counting **5**:261
cellular constituents **5**:262–264
characteristics **5**:261
computational challenges **5**:268–269
denaturing/thermal gradient gel electrophoresis (D/TGGE) **5**:264, **5**:264*T*, **5**:265*F*
DNA cloning and sequencing **5**:264*T*, **5**:265–266, **5**:266*F*
enzyme assays **5**:268, **5**:269*T*
fatty acid analysis **5**:267, **5**:268*F*
fluorescent *in situ* hybridization (FISH) **5**:264*T*, **5**:267
intergenic transcribed spacer (ITS) analysis **5**:264–265, **5**:264*T*, **5**:265*F*, **5**:266*F*
membrane hybridization **5**:264*T*, **5**:267
oligonucleotide arrays **5**:264*T*, **5**:267
polymerase chain reaction (PCR) technique **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
process/function measurements **5**:267–268
restriction fragment length polymorphisms (RFLPs) **5**:264*T*, **5**:265
solution hybridization **5**:264*T*, **5**:266–267
substrate utilization patterns **5**:267–268
nucleic acids **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
phylogenetic relationships **5**:253*F*
prevalence **5**:280–282, **5**:281*F*
research areas
Delft School tradition **5**:283–284, **5**:283*F*
research background **5**:283
16S ribosomal DNA (rDNA) sequences **5**:284–285
Woesean Reformation tradition **5**:284–285
research background **5**:252
restoration efforts
diversity evaluations **6**:446–447
establishment criteria **6**:450–451, **6**:450*F*
free-living saprobes **6**:448
importance **6**:448
soil food webs **6**:448–449, **6**:449*F*
spatiotemporal relationships **6**:447–448, **6**:447*F*
symbionts **6**:449–450, **6**:450*F*
microbial biogeography **5:271–279**
background information
Baas-Becking hypothesis **5**:271
general discussion **5**:271
molecular genetics **5**:271–272
distribution patterns
depth gradients **5**:273–274
distance–decay relationships **5**:274, **5**:274*F*
elevational gradients **5**:273–274
endemism **5**:272, **5**:273*F*
latitudinal gradients **5**:272–273
phylogenetic relationships **5**:273*F*
research background **5**:272
seasonal dynamics **5**:275
taxa–area relationships **5**:274–275, **5**:274*F*
taxa–time relationships **5**:275
future research considerations
habitat type **5**:278
pattern interpretation **5**:277–278
spatial scale **5**:278
taxonomic focus and resolution **5**:277–278
pattern interpretation
challenges **5**:275–276, **5**:276–277
contemporary environmental factors **5**:276, **5**:276*F*
future research considerations **5**:277–278
historical factors **5**:276, **5**:276*F*
hypothetical relationships **5**:275–276, **5**:276*F*
summary discussion **5**:278
microbial respiration **3**:90*T*
nonseed gene banks **3**:652, **3**:652*T*
parasitoid interactions **5**:678–679
plant–microbiota competitive interactions **6**:118–119
prokaryotes/bacteria
element cycling
aquatic sediments **5**:302*F*, **5**:303
extreme environments **5**:304
mineralization processes **5**:302–303
terrestrial soils **5**:303–304
water column **5**:302–303, **5**:303*F*
metabolic diversity **5**:300–302, **5**:301*T*, **5**:302*F*
principle properties **5**:300
symbiotic relationships **5**:304–305
psychrophiles **6:298–303**
biodiversity **6**:300
definition **6**:298
environmental considerations
Antarctica **6**:299
Arctic region **6**:299
caves **6**:299
oceans **6**:300
permafrost regions **6**:299–300
research background **6**:298–299
sea ice regions **6**:300
upper atmosphere **6**:298–299
evolutionary development **6**:300–301, **6**:302*F*
historical background **6**:298
hydrothermal vent ecosystems **7**:328
phylogenetic relationships **6**:302*F*
physiological/metabolic activities
enzymes **6**:301
growth processes **6**:301
membranes **6**:301–303
rare biosphere **6:345–352**
abundance considerations **6**:351–352
basic concepts **6**:345, **6**:346–347*F*
clone libraries **6**:349–350, **6**:349*T*
collector curves **6**:349–350, **6**:350*F*
global distribution **6**:351
high-throughput sequencing **6**:350–351
loss mechanisms **6**:351
microbial species numbers **6**:345–347
pure cultures versus sequences **6**:348–349
research summary **6**:352
species and phylogenetic concepts **6**:347–348, **6**:348*F*
river ecosystems **6**:472*F*, **6**:477
soils
bioremediation **6**:594
nitrogen deposition effects **5**:530
soil quality indicators **6**:597
microparasellid isopods **3**:178*F*

microparasites **2**:523–525, **2**:524*T*, **2**:525*F*, **5**:657
Microperoryctes murina **4**:688
Micropezidae **3**:489–491*T*
Microphalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453–454
Microphallus **7**:450–451*F*
microphytic soil crusts **6**:339
microplankton **4**:383*T*, **4**:460, **5**:133*F*
Micropogonias undulatus **5**:38
Micropterigidae **5**:385, **5**:386*F*, **5**:388*F*
Micropterigoidea **5**:387–388*T*
Micropterus salmoides **3**:123*T*, **3**:365*F*, **4**:465, **7**:261
microsatellites **3**:659, **3**:663, **3**:671*F*
Microscaphidioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
microsporans **5**:660*T*, **5**:669–670*T*, **5**:670
microsporidians **2**:524*T*, **5**:669–670*T*, **6**:296–297
Microstegium vimineum **4**:94, **6**:125
Microterigoidea **5**:386*F*
Microtermes **4**:409*T*
Microtinae **1**:710
microtubules **2**:327
Microtus pennsylvanicus **1**:365–368
Microviridae **5**:134*T*, **5**:136
microwaves **7**:297*F*
microwhip scorpions **1**:212
microzooplankton
  eutrophication impacts **3**:362
  trophic pyramid **4**:670–671
Mictacea **2**:416–417*T*
Mid-Atlantic Ridge
  Census of Marine Life project **2**:19, **2**:20*F*
  hydrothermal vent ecosystems
    biodiversity variability **7**:331
    biogeographic characterization **7**:330
midcontinent North American prairies **6**:239
Middle Asian deserts **1**:270
Middle East
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
  ecosystems
    conservation status **3**:171–173*T*
    deserts **2**:469–470
  land-use trends **4**:559*F*
  natural history museums and institutions **5**:406
  parasitic infections **5**:671*T*
  temperate forests
    distribution patterns **7**:116–117, **7**:116*T*
    diversity patterns **7**:119–121, **7**:120*F*
    dominant floristic features **7**:118
    endemism **7**:121
    physiognomic characterization **7**:117
  see also Near East ecosystems
midges **3**:489–491*T*, **3**:492, **4**:371–373, **4**:371*T*, **4**:374*F*, **7**:389–391
midocean ridges **5**:53, **7**:327–328, **7**:330
midwife toad **4**:664
Migidae **1**:206
migration **5:309–320**
  background information **5**:309
  birds **1**:605, **1**:605*T*, **1**:614
  definitions **5**:309–310
  distinguishing characteristics **5**:310
  ecological consequences **5**:316–317
  evolutionary mechanisms **3**:398, **5**:310–311
  general discussion **5**:309
  greenhouse effect impacts **4**:27, **4**:28*F*
  human disturbances
    ecological consequences **5**:318
    evolutionary consequences **5**:318–319
    migrant attributes **5**:317–318
  migratory patterns
    amphibians **5**:315–316
    birds **5**:311–313
    Census of Marine Life **2**:7–8, **2**:9*F*, **2**:10*F*
    fish **5**:314–315
    general characteristics **5**:311–313
    insects **5**:316
    mammals **5**:313–314
    reptiles **5**:315–316
Migratory Bird Treaty Act (1918) **3**:713*T*, **3**:715–716
migratory songbirds **4**:259
Mikado pheasant **7**:223
Milankovitch cycles **1**:583, **1**:585*F*, **1**:586*F*, **3**:422–423, **5**:433–434
Milaria calaudra **4**:259
Miletinae **1**:652
milfoil weevil **4**:376
Milichiidae **3**:489–491*T*
milieu souterrain superficiel (MSS) **7**:51, **7**:60
militarization impacts **4**:145–146
milkfish **1**:190*T*, **3**:464
milk thistle **2**:543
milkvetch **2**:391
milkweed leaf beetles **5**:316
Millennium Development Goals (MDGs) **6**:79, **6**:208
Millennium Ecosystem Assessment (MEA)
  agrobiodiversity valuation **3**:35–36
  background information **4**:324–325
  conservation planning **6**:263–265
  desertification **6**:79–84
  ecosystem services
    basic concepts **7**:315–316
    classifications **3**:96, **3**:97*F*
    climate change effects **2**:79–80
    conservation planning **6**:263–265
    desertification **2**:490*T*
    historical background **3**:99
    marine ecosystems **5**:451–452
    plant conservation **6**:79–84
    sustainable agriculture **1**:114
    urban–suburban biodiversity **7**:305–306
    valuation measures **7**:168–169
  endangered ecosystems **3**:170, **3**:284
  environmental ethics **3**:276
  environmental impact assessments **3**:293
  global agriculture **3**:430–431, **3**:431*T*
  urban–suburban biodiversity **7**:305–306
Millennium Seed Bank **4**:328, **6**:63, **6**:86
Millepora boschmai **3**:202*T*
Miller, Stanley **5**:630–631
millet **2**:383, **2**:659, **2**:663, **3**:32*F*
millipedes
  characteristics **4**:392
  developmental modes **5**:426*F*
  endangered species **3**:226
  geographic distribution
    endemism **5**:428
    environmental characteristics **5**:427–428
    Eurasia **5**:427*F*
    hotspots **5**:429
    islands **5**:429
    regionalization **5**:428–429, **5**:429*T*
    speciation **5**:428
  habitats and adaptations
    burrowing myriapods **5**:430
    evolutionary development **5**:429–430
    open-country myriapods **5**:430
    special habitat myriapods **5**:430–431
  invasive species **3**:222
  morphology **5**:423–424
  number of species **5**:427*T*
  soil fauna **6**:590*T*
Milne-Edwards' sifaka **6**:256–258*T*
milpa **7**:232
Milvago chimango **1**:624
Milvus migrans **5**:474
Mimallonoidea **5**:387–388*T*
Mimeutermes **4**:409*T*
mimicry
  beetles **1**:329
  butterflies
    basic concepts **1**:656
    Batesian mimicry **1**:656
    diversity considerations **1**:656–657
    Müllerian mimicry **1**:656
Mimidae **6**:308*T*
Mimiviridae **5**:130*F*, **5**:134*T*, **5**:135
Mimosa **1**:69
Mimosaceae **3**:129*T*
Mimosa pigra **1**:165, **4**:353–354, **4**:354*F*, **4**:364–365, **6**:94
Mimulus cupriphilis **6**:625
Mimulus guttatus **4**:115, **6**:625
mineral leases **2**:497*F*
mineral oils **5**:694
mineral weathering **3**:90*T*
mining impacts
  Australian reforestation programs **6**:378
  biodiversity impacts **3**:257–259, **7**:203*F*
  Brazilian reforestation programs **6**:377–378
  habitat alteration **4**:154
  protected area systems **6**:499
minisatellites **3**:663
minito del monte **4**:687*T*
mink
  adipose tissue **7**:34*F*
  introduced species **1**:79, **3**:192, **4**:357, **4**:362
  skull size comparisons **1**:692*F*
  species diversity **1**:685
  wetland ecosystems **7**:390
minke whales **3**:192–193, **5**:88–89
minnows **1**:10*T*, **6**:479

Miocene
  biotic interchange episodes **2**:545*F*
  boundary age **5**:1*T*
  brain size evolution **4**:698
  C_4 plants **2**:376*F*
  crustaceans **2**:406
  global change indicators **4**:225*F*
  Levantine biota **5**:460–461
  mammals **2**:583*F*
  megaherbivores **5**:224*T*
  membership time calculations **6**:414, **6**:415*F*
  Mid-Miocene Climatic Optimum **5**:8
  oxygen-18 ($\delta^{18}O$) measurements **1**:557*F*
  terrestrial plant communities
    abundance patterns **6**:74*F*
    Australia **1**:292
  vegetation–climate interactions **2**:683
miombo ecosystems **1**:51, **1**:67–68
Miracinonyx **4**:724–725*T*
mire **7**:369*T*
Miridae **7**:293
miriti **4**:286–287
Mirounga angustirostris **2**:263–264, **5**:87*T*, **5**:88
Mirounga leonina **2**:19*F*
miscanthus **1**:529*T*, **1**:532
Misodendraceae **4**:71
Mississippian period *see* Carboniferous period
Mississippi Estuary **5**:28–29, **5**:28*T*
Mississippi River system **5**:562–563, **7**:148*F*, **7**:149*F*, **7**:154, **7**:372–373, **7**:374*F*
Missouri Botanical Garden
  collection size **5**:411*F*
  educational programs **3**:138
  historical perspective **5**:405–406
  importance **6**:508
  monographs and inventories **3**:524
  nonseed gene banks **3**:652
  plant conservation **6**:86–87
  research collections and libraries **6**:510–511
Miss Waldron's red colobus **6**:250, **6**:256–258*T*
mistletoes **3**:518–519, **3**:521, **4**:70, **4**:77, **5**:476, **5**:673, **6**:315*T*
Mitchella **1**:271
mites **1:202–217**
  Arctic ecosystems **1**:231*T*
  characteristics **1**:202–204, **1**:213–215, **1**:214*F*, **2**:524*T*, **4**:392, **4**:394*F*
  classifications **4**:381*T*
  domesticated ruminants **1**:710–711
  ecological characteristics **1**:214–215
  feeding behaviors **1**:204
  food webs **7**:266*F*, **7**:267*T*
  genera and species **1**:202*T*
  habitats **1**:204
  insecticide resistance **4**:298–299
  introduced species
    India **1**:82
    South Africa **1**:81
    United States **1**:78
  lake and pond ecosystems **4**:371–373
  paleontology **1**:204–205
  phylogeny and taxonomy **1**:204, **1**:205*F*, **1**:215–216
  plant defense systems **2**:429
  reproduction and dispersal **1**:215
  reproduction and growth **1**:203–204
  soils
    soil fauna **6**:587, **6**:590*T*
    soil quality indicators **6**:597
  thermal ecosystems **4**:118
  United States **1**:78
Mithrox **3**:610*F*
mitochondria
  eukaryotic endosymbiotic theory **3**:330–331
  mitochondrial cytochrome b **3**:666
  mitochondrial DNA (mtDNA)
    animal genome **3**:659–660
    evolutionary significant unit (ESU) criteria **1**:668–669
    genetic diversity **3**:666
    plant genome **3**:660
    RNA editing **5**:572, **5**:573–574*T*, **5**:575*F*
  mitochondrial tRNA **5**:576–577
  RNA editing
    mitochondrial DNA (mtDNA) **5**:572, **5**:573–574*T*, **5**:575*F*
    mitochondrial tRNA **5**:576–577
mitotic recombination **6**:353, **6**:354*F*, **6**:367
Mitragyna **1**:263
Mitrastemon **4**:72–73
Mitrastemonaceae **4**:72–73, **4**:72*F*
Mitu mitu **7**:185
mixed fire regimes **3**:436
mixed forests **2**:61*F*, **3**:428*T*, **4**:563*F*, **5**:526*T*, **6**:492*T*, **6**:499*T*
Mixotoxodon **4**:727*T*
Mizodendraceae **5**:216–217
Mlao cultivation **7**:146
Mnemiopsis leidyi **3**:284, **4**:358–359
Mnesarchaeoidea **5**:386, **5**:386*F*, **5**:387–388*T*
moas **1**:606, **5**:366
mobile phone use **7**:83*F*
Moby Dick (Melville) **4**:653
moderate-resolution imaging spectroradiometer satellite (MODIS) **2**:463*F*, **4**:295
modern extinctions **5:362–372**
  crustaceans **2**:415–418, **2**:416–417*T*
  current on-going extinctions **5**:371
  density-related traits **5**:370–371
  extinction cascades **5**:367–368
  first-contact extinctions
    basic concepts **5**:365
    Madagascar **5**:365
    New Zealand **5**:366
    Pacific islands **5**:365–366
  historical extinctions
    extinction paradoxes **5**:367
    Hawaii **5**:367
    New Zealand **5**:366–367
    number of species **5**:368*T*
  introduced species **5**:370
  population dynamics **5**:370
  postglacial extinctions
    Africa **5**:364–365
    Australia **5**:363–364, **5**:364*F*
    North America **5**:363
  research challenges
    general discussion **5**:368
    geographical bias **5**:369
    inadequate documentation **5**:368–369
    inherent biological limitations **5**:369
    methodological limitations **5**:369
    number of species **5**:368, **5**:368*T*
    taxonomic and habitat bias **5**:369
    uneven documentation **5**:369
  species lifespans **5**:362–363, **5**:362*T*
  theoretical perspectives **5**:369–370
  time factors **5**:371
modern scientific plant breeding **1**:460–461
Mogoplistidae **3**:730
Mohoua ochrocephala **7**:184
mohua **4**:194*T*
Moina **6**:26
moist deciduous forests **1**:262–263*T*, **1**:263–266
moist sal savannas **1**:70
moist tropical forests **1**:277*T*, **1**:404
Mojave Desert **2**:468–469
Mokelumne River **7**:151*F*
Mola mola **3**:471*F*, **5**:657
molasses grass **4**:364
Moldova **6**:519
mole crickets **3**:728, **3**:728–729
molecular dating analyses **6**:614–615
molecular evolution
  deoxyribonucleic acid (DNA) **2**:56–57, **2**:56*F*
  generation and maintenance mechanisms **3**:662
  hemoglobin **2**:55–56, **2**:55*F*, **2**:55*T*
  historical perspective **3**:662
  random genetic drift
    evolution rates **6**:191–192
    neutral evolution theory **6**:191
    neutral polymorphism **6**:192
  theoretical perspectives **3**:662–663
  theory of evolution **3**:394–395
molecular genetics **3**:655–656, **5**:271–272
molecular level diversity **2:626–637**
  conservation biology **2**:637
  DNA sequencing **2**:627, **5**:254
  future research outlook **2**:635–637
  general discussion **2**:626
  genetic variation
    mutation process
      characteristics and classifications **2**:627–628
      cytological considerations **2**:628–629
      empirical research **2**:631–632, **2**:631*T*
      evolutionary models **2**:629–630
      genome size effects **2**:628, **2**:628*F*
      influencing factors **2**:629
      mutation rates **2**:628, **2**:628*T*
      nearly neutral theory **2**:630
      neutral theory **2**:629–630
      nucleotide polymorphism **2**:631–632, **2**:631*T*, **2**:632*F*
      ontogenetic considerations **2**:628–629

prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
selection-based models **2**:630–631
statistical tests **2**:631
phylogenetic divergence
amino acid replacement **2**:634, **2**:634*T*
empirical research **2**:634
evolutionary models **2**:633
general discussion **2**:632–633
nearly neutral theory **2**:633
neutral theory **2**:633
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
selection-based models **2**:633–634
statistical tests **2**:634
substitution process **2**:632–633, **2**:634*T*
process mechanisms **2**:627
mammalian evolution **4**:685–686
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
research background **2**:626–627
valuation measures **2**:637
molecular toxicology **3**:118
moles
biotic relationships **1**:710
characteristics **4**:690–691
food webs **7**:266*F*, **7**:267*T*
number of species **4**:690*T*
Molinia caerulea **1**:142, **6**:124–125
Mollicy Farms **7**:153
mollies **3**:467
Mollisols **6**:123*T*
Molluginaceae **2**:374*T*
Mollugo **2**:376*F*
Mollugo cerviana **2**:373*T*
Mollugo verticillata **2**:373*T*
Mollusca/molluscs **5:373–383**
ancient lakes species flocks **4**:464, **4**:464*T*
Antarctica **1**:184*T*
aquaculture **1**:190*T*
Cenozoic mass extinctions **5**:176
Census of Marine Life project **2**:7*F*
classifications and sampling methods **4**:381*T*, **5**:382
convergent evolution **3**:613, **3**:615*F*
coral reef ecosystems **2**:332*T*
ecology **5**:380–381
endangered species **3**:202*T*
extinctions **3**:202*T*, **4**:659*T*
feeding behaviors **5**:380–381
fossil record **3**:684*F*
freshwater ecosystems **3**:572*T*
genetic diversity **3**:665*F*
global diversity **4**:385*T*
introduced species **1**:82
known and estimated species **4**:384*T*
lake and pond ecosystems **4**:371–373, **4**:461*T*
major groups
Aplacophora **5**:376, **5**:376*F*
Bivalvia/bivalves
classifications **4**:381*T*
hydrothermal vent ecosystems **4**:391, **7**:328–329
physical characteristics **5**:376–377, **5**:377*F*
sandy beaches **2**:131–133
spatiotemporal distributions **5**:373–374, **5**:374*F*
species diversity **4**:386–387, **4**:387*F*
Cephalopoda
physical characteristics **5**:379–380, **5**:380*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
Coeloscleritophora **5**:374
Gastropoda
classifications **4**:381*T*
physical characteristics **4**:392, **5**:379, **5**:379*F*
sandy beaches **2**:131–133
spatiotemporal distributions **5**:373–374, **5**:374*F*
species diversity **4**:386–387, **4**:386*F*, **4**:387*F*, **4**:388*F*
Hyolitha
physical characteristics **5**:374, **5**:375*F*
spatiotemporal distributions **5**:374*F*
Monoplacophora
physical characteristics **5**:377–378, **5**:377*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
phylogenetic relationships **5**:374
Polyplacophora
classifications **4**:381*T*
physical characteristics **5**:375–376, **5**:375*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
Rostroconchia
physical characteristics **5**:375, **5**:375*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
Scaphopoda
classifications **4**:381*T*
marine environments **4**:382*F*
physical characteristics **5**:378–379, **5**:378*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
mangrove ecosystems
bivalves **5**:17
intertidal gradient **4**:337
meiofauna **5**:18
number of species **5**:20*T*
snails **5**:17–18, **5**:17*F*
marine ecosystems **5**:50*T*, **5**:106*F*
mass extinctions **3**:417–418
modern extinctions **5**:368*T*
mollusc–human interactions **5**:381–382
Near East ecosystems **5**:460*T*
number of species **4**:371*T*, **5**:669–670*T*
phylogenetic relationships **7**:434*F*
physical characteristics **4**:392, **5**:373
Proterozoic–Cambrian transition **3**:396*F*
river ecosystems **6**:478–479, **6**:478*T*
sandy beaches **2**:131–133
spatiotemporal distributions **5**:373–374, **5**:374*F*
species estimates **3**:710*T*
wetland ecosystems **7**:389–391
Moloch horridus **2**:471*F*
Molossidae **6**:309*T*
Molothrus ater **4**:490, **4**:491*F*
Molothrus bonairiensis **1**:618
molting hormones **2**:427
molt migration **5**:312–313
Moluccas **1**:514*T*
Molytini **1**:328
Momotidae **6**:308*T*
Monacha haifaensis **5**:466–467
Monachus monachus **5**:77*T*, **5**:80
Monachus schauinslandi **5**:77*T*
Monachus tropicalis **3**:191–192, **5**:75
Monaco **6**:519
Monarch butterflies **2**:141–142, **3**:222, **5**:316, **5**:396–397
Mongolia
animal species **1**:261*T*
botanical gardens **6**:519
cropland conversion **4**:665*T*
Development by Design (DbD)
landscape-level conservation planning **2**:498–500, **2**:499*F*
mineral leases **2**:497*F*
residual impact/optimal offset determinations **2**:502–503*F*
domesticated ruminants **1**:702–703
ecosystems
marshes **1**:278–279
temperate grasslands **1**:269
temperate semi-deserts and deserts **1**:269–270
threatened species **1**:278*T*
vegetation types and coverage percentage **1**:264–265*T*
megaherbivores **5**:224*T*
plant species **1**:259*T*
urbanization **4**:560*F*
wheat production **1**:87–88
Mongolian gazelles **3**:191
Mongolian–Manchurian steppe **1**:262–263*T*
Mongolian wild horses **3**:191
mongoose lemur **6**:251
mongooses **1**:77, **1**:686, **4**:359, **4**:365–366, **4**:690*T*, **4**:693, **6**:240–241
monilophytes **4**:168–169
monitor lizards
commercial exploitation **3**:217, **6**:411
mangrove ecosystems **5**:13
monkey flower **6**:625
monkey grasshoppers **3**:733, **3**:733*F*
monkeys
food habits **3**:127*T*
harvesting threats **2**:590*T*
introduced species **1**:76–77
mangrove ecosystems **5**:14
number of species **4**:690*T*
species diversity **4**:692–693
tropical Africa **1**:48–49
monk seals **2**:356, **5**:77*T*, **5**:80
Monmonier's maximum difference algorithm **4**:511
Monoclonius **2**:519*T*, **2**:520*T*
Monocotyle **7**:448–449*F*

monocotyledons
  Antarctica **1**:186*T*
  Arctic ecosystems **1**:231*T*
  edible species **3**:129*T*
  modern extinctions **5**:368*T*
  number of species **2**:374*T*
  phylogenetic relationships **5**:618*F*, **5**:621*F*
  taxonomic distributions **3**:519*T*, **3**:522*T*
monoculture crops **1**:359–360
Monodella relicta **5**:470–471, **5**:470*T*
Monodelphis domestica **6**:361*T*
Monodon monoceros **4**:693–694, **5**:90, **5**:314
Monodus **6**:33*F*
monogamy **6**:573
Monogenea **5**:657–658, **5**:659*T*, **5**:661*F*, **5**:669–670*T*, **7**:439–440*T*, **7**:446–447, **7**:462*F*
monogenean trematode **2**:528*T*
Monolistra racovitzae racovitzae **7**:58
monomethyl hydrazine **3**:635*T*
Monomorium mayri **5**:463*T*
Mononchus **7**:421*F*, **7**:422–424, **7**:429–430
Mononegavirales **5**:134*T*, **5**:139
monophenol monooxygenase **5**:269*T*
monophyly **2**:52–53, **6**:683, **7**:91–93, **7**:92*F*
Monopisthocotylea
  Capsalidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
  Dactylogyridea **7**:439–440*T*, **7**:447, **7**:448–449*F*
  Gyrodactylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
  Lagarocotylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Monoplacophora
  fossil record **3**:685*F*
  physical characteristics **5**:377–378, **5**:377*F*
  spatiotemporal distributions **5**:373–374, **5**:374*F*
Monorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Monosiga **6**:293, **6**:294*F*
monoterpenes **2**:427
Monotes glaber **1**:68
Monotremata
  characteristics **3**:188, **4**:682, **4**:686, **7**:341
  convergent evolution **4**:701*T*
  phylogenetic relationships **4**:686, **5**:1*F*
  species diversity **4**:682*F*
  threat percentage **3**:195
monounsaturated fatty acids **7**:32–33
monsoon forests **1**:262–263*T*, **1**:263–266
monsoons **1**:258
Monstera **3**:522*T*
Montana **3**:521*T*
montane ecosystems
  montane forests **4**:432*T*
  montane grasslands **6**:600–604*T*
  tropical Africa **1**:53
Montanoceratops **2**:519*T*
Montastraea annularis **2**:340*F*
Montchadskyella **7**:448–449*F*
Monte Carlo Markov Chain (MCMC) analysis **4**:511
Monteverde cloud forest reserve (Costa Rica) **3**:167–168
Monteverde harlequin frog **2**:86–87, **4**:664
Montreal Protocol (1987) **1**:289, **2**:152–153, **3**:271, **7**:18–19
Montrichardia arborescens **1**:163
Montrichardia linifolia **1**:165
Montserrat **4**:560*F*, **6**:519
Moorea **5**:591*F*
moorhen **3**:365–366
Moorish Idol **3**:467–471
moorland **4**:577–579, **4**:578*F*, **7**:369*T*
Moor macaque **6**:256–258*T*
moose
  boreal forests **1**:631–632
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  riparian zones **6**:466
  wetland ecosystems **7**:390
Moraceae **1**:263, **3**:129*T*, **3**:130*T*, **3**:522*T*, **6**:316*T*
Moraea loubseri **3**:206
Moran effect **4**:213
Mora oleifera **7**:272
Morbillivirus **2**:528*T*
Morchella **3**:637*T*, **3**:639–640
Morchella esculenta **3**:625*T*
Mordellidae **1**:327*T*
Morella faya **4**:360, **6**:121
morel mushrooms **3**:625*T*, **3**:628–629, **3**:628*F*, **3**:637*T*, **3**:639–640
Morenelaphus **4**:727*T*
Morgan, T. H. **6**:3–4
Moringa ovalifolia **1**:67
Moringa peregrina **5**:485, **5**:486
Morisita–Horn index **2**:445, **5**:208–209
Moritella marinus **6**:301
Mormon crickets **3**:727
Mormoopidae **6**:309*T*
Mormotomyiidae **3**:489–491*T*
Morocco
  botanical gardens **6**:519
  desert ecosystems **2**:469–470
  species richness **1**:499*T*
  wheat production **1**:87–88
Morone saxatilis **6**:479
Morone saxitilis **5**:61
Morphinae **1**:652*F*, **1**:653
morphine **1**:359, **2**:426*T*, **2**:427, **2**:432, **5**:712*T*
Morpho achilles **1**:652*F*
morpholines **5**:695
Morula marginalba **4**:334–335
Morus bassanus **3**:256
Morus misozhygia **1**:68
Moschidae **4**:696
Moschopsis rosulata **4**:106–107
mosquitoes
  characteristics **3**:492
  classifications **3**:489–491*T*
  as disease vectors **1**:362–365, **1**:364*F*, **1**:365*F*
  insecticide resistance **4**:301–302
  mangrove ecosystems **5**:13
  number of species **4**:371*T*
  wetland ecosystems **7**:389–391
mosquito fish **3**:167, **4**:361, **7**:390
moss
  air pollution impacts **1**:142*F*
  alpine ecosystems **1**:153–154, **1**:154*F*
  Antarctica **1**:184*T*
  anthropogenic impacts **7**:399
  Arctic ecosystems **1**:231*T*, **1**:239*F*
  boreal forests **1**:628–630
  characteristics **6**:59
  conservation practices **6**:83
  forest canopies **3**:520, **3**:521*T*
  freshwater ecosystems **3**:571–572
  modern extinctions **5**:368*T*
  North America **5**:564–565
  number of species **6**:58*T*
moss animals **4**:381*T*
Motacilla flava **1**:710
moth flies **3**:489–491*T*, **3**:493
moths **5:384–403**
  attraction to light **5**:400–401, **5**:401*F*
  biogeography **5**:389–390, **5**:389*T*
  celestial navigation **5**:400–401
  characteristics **4**:392, **5**:384–385, **5**:385*F*
  chemical defense **5**:396–397, **5**:397*F*
  classifications
    Ditrysia **5**:386–389, **5**:386*F*
    Glossata **5**:385–386, **5**:386*F*
    Neolepidoptera **5**:386, **5**:386*F*
    phylogenetic relationships **5**:385, **5**:386*F*
  suborders **5**:385, **5**:386*F*, **5**:387–388*T*
  coloration **5**:400–401
  conservation programs **5**:402–403
  defense mechanisms **5**:391*F*, **5**:392*F*, **5**:396–397, **5**:396*F*
  dispersal and migration **5**:397
  diurnality and nocturnality **5**:400–401
  elevational range shifts **4**:605*T*
  endangered species **3**:226
  feeding biology
    adults **5**:394–395, **5**:395*F*
    larvae **5**:392–394
  food webs **7**:266*F*, **7**:267*T*
  fossil record **5**:389
  freshwater ecosystems **3**:579–580, **3**:579*T*
  genetics research **6**:3
  importance **5**:401–402
  insecticide resistance **4**:298–299
  key adaptations **5**:397–398, **5**:398*F*
  latitudinal range shifts **4**:604*T*
  life history
    cocoons **5**:391–392
    diapause/quiescence **5**:392
    eggs **5**:390
    hypermetamorphosis **5**:391
    larvae **5**:388*F*, **5**:390–391
    pupa stage **5**:391–392, **5**:393*F*
  mangrove ecosystems **5**:13
  natural enemies **5**:395–396
  number of species **4**:371*T*, **4**:392*T*
  phenological shifts **5**:722
  pheromones **5**:398–400, **5**:399*F*, **5**:400*F*
  pollination **6**:141–142, **6**:141*T*
  population density **6**:159*F*
  Southeast Asia **1**:514–515
  species declines **5**:402–403

species diversity **5**:387–388*T*, **5**:389–390, **5**:389*T*
subterranean ecosystems **7**:60
threatened species **3**:579–580, **3**:579*T*
motor neuron disease **1**:592
mottled woodpecker **4**:659
motts **4**:574–577, **4**:575*F*
Motuweta isolata **5**:366–367
Mougeotia **3**:358
mound springs **3**:186, **3**:187*F*
mountain agile gibbon **6**:256–258*T*
mountain-and-basin deserts **2**:460–461, **2**:464*F*
mountain botanical gardens **6**:505
mountain cloud forests **3**:111
mountain gazelle **5**:475–476
mountain goats **4**:724–725*T*
mountain golden heather **3**:214
mountain gorilla **3**:192, **6**:254, **6**:256–258*T*
mountain hares **1**:631–632
mountain hemlock **2**:571
mountain lions **1**:703
mountain midges **3**:489–491*T*
mountain regions
endemism **3**:233
high-elevation Andean ecosystems **4:96–110**
background information **4**:96
geographic distribution **4**:96–97, **4**:97*F*
habitat characteristics **4**:96–97
number of species **4**:102*T*
páramo ecosystem
characteristics **4**:101–103, **7**:275–276
geographic distribution **4**:96–97, **4**:97*F*
northern Andean alpine cushion formations **4**:103
northern Andean dwarf *Polylepis* forest **4**:103
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98
plant diversity **4**:98–99, **4**:99*T*
representative species **4**:102*F*
subpáramo zone **4**:102–103
superpáramo zone **4**:103
true páramo zone **4**:103
physical characteristics and patterns **4**:97–98
plant diversity
endemism **4**:99–100, **4**:99*T*
origins and diversification **4**:100–101
species richness **4**:98–99
puna
central Andean dwarf *Polylepis* forest **4**:105
cushion bogs **4**:105
desert puna **4**:104–105
dry puna **4**:104
geographic distribution **4**:96–97, **4**:97*F*
humid puna **4**:104
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:103–104
plant diversity **4**:98–99
representative species **4**:104*F*
salar vegetation **4**:105
research summary **4**:107
southern Andean steppes
altitudinal zonation **4**:106–107, **4**:106*F*
bogs and seeps **4**:107
cushion plants **4**:106–107, **4**:106*F*
geographic distribution **4**:96–97, **4**:97*F*
latitudinal gradients **4**:106–107
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:105–107
plant diversity **4**:98–99
representative species **4**:105*F*
high mountain ecosystems
characteristics **7**:275
humid forests **7**:275
páramo ecosystem **4**:101–103, **7**:275–276
perhumid forests **7**:275
subhumid forests **7**:275
low mountain ecosystems
characteristics **7**:274–275
humid forests **7**:275
perhumid forests **7**:275
subhumid forests **7**:274–275
protected area systems **5**:444*T*, **6**:492–493
mountain thornbill **2**:75*F*
mountain whitefish **4**:515
mountain zebras **1**:487–488, **4**:695
Mount Kinabalu (Borneo) **1**:267
Mount Pinatubo volcanic eruption **1**:282–284, **1**:288
Mount St. Helens **6**:446–447, **6**:450–451, **6**:450*F*
Moupin golden snub-nosed monkey **6**:256–258*T*
mousebirds **7**:181–182
mouse deer **4**:696
mouse-ear hawkweed (*Hieracium pilosella*) **4**:348–349, **4**:349*F*
mouse lemur **6**:251*T*
Movile Cave (Romania) **7**:51, **7**:58–59, **7**:59*T*
Mozambique
botanical gardens **6**:519
deforestation **1**:677*F*
woodland savannas **1**:67–68
Mt. Hermon montane forest **5**:480*F*, **5**:482
Mt. Hermon tragacanth vegetation **5**:480*F*, **5**:483–484
Mt. Uaraguess guereza **6**:256–258*T*
Muara Cimanuk **1**:276*T*
mucajá **4**:286–287
Mucor **3**:627
Mucuna pruriens **1**:589, **6**:560–561, **7**:26
mud crabs **5**:16
mudflats **4**:337–339, **5**:107, **5**:108*T*
mud lobsters **5**:15–16
mudminnow **3**:464–465, **6**:522–523
mudpuppies **7**:389–390
mudskippers **5**:15, **5**:15*F*
mud turtles **7**:390
Mudumalai Tiger Reserve (India) **1**:60*F*
Mugil cephalus **5**:473
Mugilomorpha **3**:467
Mugil ramada **5**:473
Muhlenbergia fastigiata **4**:105
Muhlenbergia montana **2**:364*F*
Muir, John **2**:230–231, **2**:281
mulberries **2**:383
mule deer **1**:703, **2**:503–505, **5**:313
mulga **2**:468
Mulinum spinosum **4**:106
mullein **3**:208–209
Müller, Fritz **2**:138, **2**:424
Müller, Hermann **2**:424, **6**:1–3
Müllerian mimicry **1**:656, **2**:424
Mullerornis betsilei **5**:365
Müller's gray gibbon **6**:256–258*T*
Müller's larva **7**:462–466, **7**:464–465*F*
Muller's ratchet **6**:196–197
mullets **1**:190*T*, **3**:467, **5**:15
Mullus barbatus **2**:356
multichannel omnivory **3**:504
multifunctional agriculture **1**:115–117, **1**:115*F*, **1**:116*F*, **1**:116*T*, **1**:117*F*, **2**:300–301
multiparasitism **5**:675
multiple matrix regression analysis **4**:514
Multiple Opening and Closing Net with an Environmental Sensing System (MOCNESS) **2**:13
multiple regression of distance matrices (MRDM) model **4**:514
Multiple-Use Multiple-Use Sustained-Yield Act (1960) **3**:718
multiplicative partitioning
alpha x beta **2**:439–440
gamma diversity **2**:439–440
richness x evenness **2**:439
multipurpose botanical gardens **6**:505
Multipurpose Marine Cadastre (MMC) **5**:460–462*T*
Multiscale Integrated Models of Ecosystem Services (MIMES) **5**:460–462*T*, **5**:463
multisite biodiversity experiments **4**:586–587
multispecies interactions
challenges **6**:721–722
evolutionary change **6**:723
food webs **6**:721–722, **6**:721*F*, **6**:722*F*
indirect interactions **6**:722–723
keystone species **6**:722, **6**:723*T*
metacommunities **6**:723
trophic cascades **6**:722–723
trophic levels **6**:721–722, **6**:721*F*, **6**:722*F*
Multispectral Scanner (MSS) **6**:393, **6**:393*T*
Multituberculata **4**:684
Multituberculates **5**:1*F*
Multivariate ENSO Index (MEI) **3**:155, **3**:156*F*
mummies **3**:127–128
mung beans **2**:663
Munidopsis lentigo **7**:330
Muntiacus muntjak **1**:516–517
Muntiacus putaoensis **4**:703–704
Muntiacus rooseveltorum **4**:703–704
Muntiacus truongsonensis **4**:703–704
Muntiacus vuquangensis **4**:703–704

Muntingia calabura **1**:163
muntjacs **1**:516–517, **4**:703–704
Muraltia **5**:219
Murchison Falls National Park **5**:234–236
Muridae **4**:697, **6**:309*T*
muriquis **6**:251*T*, **6**:256–258*T*
muruci **4**:286–287
Musa accuminata **2**:663
Musa balbisiana **2**:663
Musca domestica **3**:492, **4**:301
Muscardinus avellanarius **4**:573
muscarine **3**:635*T*
Muscicapa striata **5**:474
Muscidae **3**:489–491*T*, **3**:492
muscimol **3**:635*T*
Muscoidea **3**:489–491*T*
Muscomorpha **3**:489–491*T*, **3**:492
Muscovy duck **1**:644*F*
Musculista senhousii **6**:457
Museo Argentino de Ciencias **5**:406
Museo de Historia Natural de la Ciudad de México (MNH) **5**:406
Museo del Mar (Argentina) **5**:406
Museum für Naturkunde (Berlin) **5**:405, **5**:408*F*
Museum National d'Histoire Naturelle (Paris) **5**:405, **5**:408*F*, **5**:409*F*, **5**:411*F*, **6**:510–511
Museum of Comparative Zoology (Harvard) **5**:405–406, **5**:408*F*
Museum of New Zealand Te Papa Tongarewa **5**:406
Museum of Vertebrate Zoology (University of California Berkeley) **5**:405–406
museums **3**:138
  see also natural history museums and institutions
Museum Victoria (Australia) **5**:406
Museu Nacional (Brazil) **5**:406
Museu Paraense Emílio Goeldie Parque Zoobotânico **6**:507
mushrooms
  commercial production **3**:637–638, **3**:637*F*, **3**:637*T*
  modern extinctions **5**:368*T*
  poisonous mushrooms **3**:634, **3**:635*T*
  taxonomy **3**:625*T*
musk deer **4**:696
muskeg **7**:369*T*
musk oxen
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  social behavior **6**:571–573
  tundra ecosystems **6**:330–331
muskrats **2**:551–552*T*, **2**:554–555, **2**:554*F*, **4**:365–366, **4**:701*T*, **7**:390
musk thistles **4**:361, **4**:363
musk turtles **7**:390
Muslim declaration on nature **3**:273, **6**:381, **6**:382
Mus musculus **1**:77, **2**:552*F*, **4**:249*F*, **6**:361*T*
Mus musculus castaneus **2**:631*T*, **2**:636*F*
Mus musculus domesticus **6**:5
Mus musculus musculus **6**:5
mussels
  acidification effects **3**:584–585, **3**:584*F*
  American Southeast rivers **3**:185*F*
  ancient lakes species flocks **4**:464*T*
  aquaculture **1**:190*T*
  biomonitoring processes **3**:125
  chemosynthetic symbioses **5**:120–121
  endangered species **3**:199
  eutrophication impacts **3**:355–356
  exotic species invasions **4**:378
  extinctions **3**:411–412
  freshwater ecosystems **3**:578–579, **3**:579*T*
  intertidal ecosystems **4**:334
  number of species **3**:182*T*, **4**:371*T*
  predation effects **6**:232, **6**:722
  river ecosystems **6**:478*T*
mustard **1**:529*T*
Mustela erminea **5**:370, **7**:28*F*
Mustela lutreola **3**:192, **4**:362
Mustela nigripes **3**:196, **3**:411, **6**:342
Mustela nivalis **4**:693
Mustela vison **1**:79, **3**:192, **4**:357, **4**:362, **7**:390
Mustelidae
  body mass **1**:687*F*
  ecological characteristics **1**:688*F*
  evolutionary history **2**:583*F*
  food categories **1**:690*F*, **1**:691*F*
  home range size **1**:687*F*
  skull size comparisons **1**:692*F*
  species diversity **1**:685, **4**:693
  tropical rain forests **6**:309*T*
mutations **3**:398, **3**:658*T*, **3**:660–661, **5**:629, **6**:633
mute swans **3**:365–366
mutualism
  benefits and challenges **2**:320–321
  coevolution **2**:146, **6**:720–721
  definition **5**:656
  fungi
    characteristics **3**:636
    endophytes **3**:636
    lichens **3**:636–637, **3**:637*F*
    mycorrhizae **2**:391, **3**:636, **3**:636*F*
  habitat–niche relationship **4**:45
  intimate mutualism **2**:323
  introduced species **4**:365
  keystone species study **4**:447
  nutrient-exchange mutualisms
    chemosynthetic symbioses **5**:120–121, **5**:120*F*
    hydrothermal vents **5**:120–121, **5**:120*F*
    nutrient conversion
      nitrogen fixation **5**:121
      shipworms **5**:121
      worm–whale interaction **5**:121, **5**:121*F*
    photosynthetic symbioses **5**:118–120
  plant–animal interactions
    cheating studies **6**:45–46
    fruit and seed dispersal
      ants **6**:50
      beetles **6**:50
      birds **6**:50
      dispersal syndromes **6**:50, **6**:51*T*
      disperser types **6**:50
      evolutionary dead ends **6**:52
      fish **6**:50
      frugivores **6**:50–51
      importance **6**:49–50
      insects **6**:50
      mammals **6**:50
      plant adaptations **6**:50
      reptiles **6**:50
      seed predators **6**:51–52
      species persistence **6**:61–62
      vertebrates **6**:50
    plant–pollinator interactions
      angiosperm radiation hypothesis **6**:48
      basic concepts **6**:47
      cheating studies **6**:47
      competition **6**:49
      desert ecosystems **2**:474–475, **2**:475*F*
      floral nectar spurs **6**:47–48
      insect pollination **6**:48
      pollination syndromes **6**:48–49, **6**:49*T*
      reproductive isolation **6**:47
      species persistence **6**:61–62
      supporting services **1**:360, **5**:354–355
    plant-protecting ants/plant-feeding ants
      benefits to ants/cost to plants **6**:46
      benefits to plants **6**:46–47
      myrmecophytes **6**:46
      myrmecotrophic epiphytes **6**:47
    specialization **6**:45–46
  species interactions **6**:715*T*, **6**:719, **6**:720–721
  see also symbiotic relationships
mutualistic symbionts **6**:449–450
Mya arenaria **3**:364
Myanmar
  animal species **1**:261*T*
  botanical gardens **6**:519
  deforestation **2**:449*T*, **2**:451*T*
  ecosystems
    genetic richness **1**:514*T*
    mangroves **1**:275*T*, **1**:513*T*
    monsoon forests **1**:263–266
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    tropical dry forest **1**:266–267
    tropical moist forest **1**:277*T*
    vegetation types and coverage percentage **1**:264–265*T*
  endangered mammals **3**:194*T*
  forest product exports **2**:451*T*
  indigenous peoples **4**:270
  nitrogen deposition **5**:526*T*
  plant species **1**:259*T*
  rice production **1**:88
  see also Southeast Asia
mycetismus **3**:633–634, **3**:634, **3**:635*T*
Mycetophagus quadripustulatus **7**:220*T*
Mycetophilidae **3**:489–491*T*, **3**:493
myclobutanil **5**:695
mycoallergies **3**:633–634
Mycoplasma agassizii **2**:528*T*
Mycoplasma gallisepticum **2**:528*T*
mycorrhizae
  alpine ecosystems **1**:155, **1**:155*F*
  forest management impact **1**:451
  grassland ecosystems **4**:4

mutualism **2**:391, **3**:636, **3**:636*F*
nitrogen deposition **5**:534*F*
plant hybridization **4**:174
slash-and-burn systems **6**:556
mycosis **3**:633–634, **3**:634*T*
Mycosphaerella fijiensis **6**:557
Mycosphaerella musicola **6**:557
mycotoxicosis **3**:633–634, **3**:634–635, **3**:635
mycotrophs **6**:315*T*
mydas flies **3**:489–491*T*
Mydidae **3**:489–491*T*
mygalomorphs **1**:206, **1**:206
Mylodon **4**:727*T*
Mylodon robustus **5**:225*T*
Mylodontidae **4**:727*T*
Mylohyus **4**:724–725*T*
Mymaridae **5**:674–675
mynas **1**:81
Myocastor coypus **1**:77, **4**:143, **5**:477, **7**:399
Myoglossata **5**:385–386, **5**:386*F*, **5**:387–388*T*
Myomorpha **2**:586, **4**:697
Myotis grisescens **7**:52–53
Myotis lucifugus **2**:527–529, **2**:529*F*, **5**:313
Myoviridae **5**:130*F*, **5**:134–135, **5**:134*T*
Myrciaria dubia **1**:163–164
Myriad Genetics **1**:595
Myriapoda/myriapods **5:421–432**
Amazon ecosystems **1**:246*T*
Antarctica **1**:184*T*
arthropod radiation ecology **2**:399
characteristics and classifications **5**:421–423
Chilopoda
Amazon ecosystems **1**:246–247, **1**:246*T*, **1**:247
characteristics **4**:392
developmental modes **5**:426
endangered species **3**:226
geographic distribution
endemism **5**:428
environmental characteristics **5**:427–428
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429, **5**:429*T*
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430
open-country myriapods **5**:430
special habitat myriapods **5**:430–431
morphology **5**:421–423, **5**:422*F*
number of species **5**:427*T*
soil fauna **6**:590*T*
developmental modes **5**:426, **5**:426*F*
Diplopoda
alpine ecosystems **1**:155*F*
Amazon ecosystems **1**:246*T*, **1**:247
characteristics **4**:392
developmental modes **5**:426*F*
endangered species **3**:226
geographic distribution
endemism **5**:428
environmental characteristics **5**:427–428
Eurasia **5**:427*F*
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429, **5**:429*T*
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430
open-country myriapods **5**:430
special habitat myriapods **5**:430–431
morphology **5**:423–424, **5**:423*F*
number of species **5**:427*T*
soil fauna **6**:590*T*
subterranean ecosystems **7**:53
geographic distribution
endemism **5**:428
environmental characteristics **5**:427–428
Eurasia **5**:427*F*
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429, **5**:429*T*
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430
open-country myriapods **5**:430
special habitat myriapods **5**:430–431
habitats and feeding habits **5**:424–425
morphology
Chilopoda **5**:421–423, **5**:422*F*
Diplopoda **5**:423–424, **5**:423*F*
Pauropoda **5**:424, **5**:424*F*
Symphyla **5**:424, **5**:424*F*
number of species **5**:426–427, **5**:427*T*, **5**:669–670*T*
Pauropoda
Amazon ecosystems **1**:246*T*
characteristics **4**:392
developmental modes **5**:426
geographic distribution
endemism **5**:428
environmental characteristics **5**:427–428
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430
open-country myriapods **5**:430
special habitat myriapods **5**:430–431
morphology **5**:424, **5**:424*F*
number of species **5**:427*T*
reproductive characteristics **5**:425–426
Symphyla
developmental modes **5**:426
geographic distribution
endemism **5**:428
environmental characteristics **5**:427–428
hotspots **5**:429
islands **5**:429
regionalization **5**:428–429
speciation **5**:428
habitats and adaptations
burrowing myriapods **5**:430
evolutionary development **5**:429–430
open-country myriapods **5**:430
special habitat myriapods **5**:430–431
morphology **5**:424, **5**:424*F*
number of species **5**:427*T*
Myricaceae **5**:621–625, **5**:623*F*
Myrica faya **2**:566, **2**:655, **3**:245–247
Myrica gale **7**:395–397
Myriophyllum **5**:622*F*
Myriophyllum **1**:279
Myriophyllum sibiricum **4**:376
Myriophyllum spicatum **4**:364–365, **4**:376
Myriophyllum tenellum **6**:115*T*
Myriophyllum verticillatum **6**:115*T*
Myristicaceae **6**:313–318, **6**:316*T*
Myristica fragrans **2**:663
Myrmecobiidae **4**:687
Myrmecodesmus adisi **5**:430–431
Myrmecophaga tridactyla **2**:587*F*, **7**:229*T*
Myrmecophagidae **6**:309*T*
Myrmecophilidae **3**:730
myrmecophily **1**:655
myrmecophytes **6**:46
Myrmica sabuleti **4**:363
Myrmycodia tuberosa **6**:47
Myrothamnus **5**:622*F*
Myrothamnus flabellifolius **1**:50
myrrh-trees **2**:468
Myrsinaceae **1**:267, **3**:522*T*, **5**:11*T*
Myrtaceae **1**:266*T*, **3**:129*T*, **3**:130*T*, **3**:522*T*, **5**:11*T*, **5**:221, **6**:316*T*
Myrtales **5**:622*F*
Myrtillocactus **2**:26
myrtle warbler **2**:188–189, **2**:188*F*
Myrtus communis **1**:269
Mysatales garridoi **4**:714*F*
Mysatales gundlachi **4**:714*F*
Mysatales melanurus **4**:714*F*
Mysatales prehensilis **4**:714*F*
Mysateles meridionalis **4**:715
Mysida
extinction vulnerability **2**:416–417*T*
invasive species **5**:61
number of species **4**:371*T*
sandy beaches **2**:131–133
Mysidaceans **6**:25
Mysis **6**:26
Mysis relicta **4**:363–364
Mysotis seabrai **1**:487–488
Mystacinobiidae **3**:489–491*T*
Mystacocarida **2**:401*T*, **2**:416–417*T*
mysticeti
characteristics **5**:66–67
taxonomy **5**:65*T*
Mystichthys luzonensis **3**:481
Mysticina tuberculata robusta **5**:366–367
Mythicomyiidae **3**:492
Mytilus **3**:125
Mytilus californianus **4**:442–443, **4**:443*F*, **6**:232, **6**:722, **6**:723*T*
Mytilus edulis **3**:355–356, **6**:547–548
Myxobolus cerebralis **2**:528*T*, **6**:297
myxoma virus **2**:143

myxomycetes **5**:572–575, **5**:573–574*T*, **5**:575*F*
Myxomycota (true slime molds) **3**:630–631, **3**:630*F*, **3**:631*F*, **6**:289–293, **6**:289*T*, **6**:290*F*
Myxophaga **1**:326
Myxosoma cerebralis **4**:362
Myxosporidia **5**:669–670*T*
myxozoans **5**:669–670*T*, **6**:297
Myzostomida **5**:669–670*T*, **7**:408
Myzus persicae **4**:298–299, **4**:299*F*, **4**:302*F*

N

Nabidae **7**:294
Naegleria **5**:573–574*T*, **6**:290–291
Naegleria fowleri **5**:660*T*
Naess, Arne **6**:383–384
Nagoya Protocol **1**:597, **2**:306–307, **2**:308
Nairobi National Museum (Kenya) **5**:406
Naja oxiana **7**:229*T*
Nakaumi (Lake) **1**:276*T*
naked ameboids **6**:290–291
  see also slime molds
naked amoeba **6**:586, **6**:587*F*
naked mole-rats **3**:188
Namaqua dune mole rat **1**:487–488
Nambiquara Indians **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:290*F*
Nam Can mangrove forest **1**:276*T*
Namibia
  botanical gardens **6**:519
  coastal deserts **2**:462*F*
  desert ecosystems **2**:467–468
  private conservation areas **2**:297*T*
  seasonal rainfall **2**:465*F*
  Succulent Karoo **1:485–490**
    biodiversity
      invertebrates **1**:488
      plants **1**:485, **1**:486*T*, **1**:487, **1**:487*T*
      vertebrates **1**:487*T*
    bioregions **1**:485, **1**:486*F*
    characteristics **1**:485, **1**:486*F*, **2**:467–468
    climate change threats **1**:489
    land degradation threats **1**:489
    nitrogen deposition **5**:526*T*
    plant growth forms **1**:488–489
    threat management **1**:489
    vegetation structure **1**:488–489, **2**:467–468
  tree and shrub savannas **1**:67
  urbanization **4**:560*F*
Namyang Bay **1**:276*T*
Nandinaceae **1**:260*T*
Ñandú **3**:115
Nannipus **4**:724–725*T*
Nannochloris **6**:30
Nannopterum harrisi **1**:606
Nannopus palustris tiberiadis **5**:470*T*
nanoplankton **4**:383*T*, **4**:460, **5**:133*F*
napropamide **4**:90*T*
naptalam **4**:90*T*
Narcissus tazetta **5**:484–485
narwhals **4**:693–694, **5**:65*T*, **5**:66–67, **5**:90, **5**:314, **7**:223
Nasalis larvatus **5**:14
Nasalis larvatus larvatus **6**:256–258*T*
Nasalis larvatus orientalis **6**:256–258*T*
Nasella pulchra **6**:455
Nasonia **5**:679
Nasonia vitripennis **5**:677
Nassauvia **4**:101
Nassauvia axillaris **4**:106
Nassauvia lagascae **4**:106, **4**:106–107
Nassauvia pinnigera **4**:106
Nassauvia pygmaea **4**:106–107
Nassauvia pyramidalis **4**:106
Nassula **6**:31
Nastanthus caespitosus **4**:107
Nastanthus spathulatus **4**:105–106
Nasutitermes **4**:409*T*
Nasutitermes exitiosus **4**:417
Nasutitermes latifrons **4**:429*F*
Nasutitermes triodiae **4**:415*F*
Nasutitermitinae **4**:409*T*, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:430*T*
Natalidae **6**:309*T*
Nationaal Herbarium Nederland **5**:411*F*
National Atmospheric Deposition Program (NADP) **1**:1–5, **5**:533
National Biodiversity Authority (NBA) **1**:507
National Biological Information Infrastructure (NBII) **3**:138–139
National Botanic Gardens (Belgium) **6**:510–511
National Botanic Gardens (Ireland) **6**:506–508
National Cancer Institute (NCI) **1**:596–597, **5**:709
National Ecological Observatory Network (NEON) **3**:247–248
National Environmental Education Act (1970) **3**:136
National Environmental Education Act (1990) **3**:136
National Environmental Policy Act (1970) **2**:286, **3**:713*T*, **3**:719
National Forest Conservation Program (NFCP)
  background information **3**:375–376
  biodiversity/biophysical effects **3**:376–377, **3**:377*F*
  future outlook and challenges **3**:378–379
  program goals **3**:375–376
  socioeconomic effects **3**:377–378, **3**:378*F*
  spatial distribution **3**:376, **3**:376*F*
  summary discussion **3**:382
National Forest Management Act (1976) **3**:713*T*, **3**:718, **7**:314–315
National Fruit Collection **3**:652–653
National Marine Fisheries Service (NMFS) **5**:64–66
National Marine Sanctuary System (1972) **3**:718
national monuments **6**:485–486
National Museum of Natural History (NMNH) **5**:405–406, **5**:408*F*, **5**:409*F*
National Museum of Nature and Science (Japan) **5**:406
national parks
  attendance trends
    decline incidence **7**:244
    historical trends **7**:242, **7**:243*F*
    international trends **7**:245, **7**:246*F*
    per capita visitation **7**:244–245, **7**:245*F*, **7**:246*F*
    pervasive declines **7**:247
    popularity declines **7**:245–246, **7**:245*F*, **7**:246*F*
    public surveys **7**:246
    specific-activity trends **7**:246–247, **7**:246*F*
  definition **6**:485–486
  government legislation and regulations **3**:713*T*, **3**:718
  historical background **5**:442
  historical conservation movement **2**:281
National Park Service Act (1916) **3**:712–713, **3**:713*T*, **3**:718
National Regional Nutrient Criteria Program **1**:97, **1**:97*T*
National Science Foundation **7**:19
National Seed Storage Laboratory (NSSL) **3**:643–644, **3**:651
National Wild and Scenic Rivers Act (1968) **6**:496
National Wildlife Refuge System Improvement Act (1997) **3**:713*T*, **3**:718
native crops
  Asia
    East Asia **2**:663
    Southeast Asia **2**:663
    southwest Asia **2**:662–663
  Europe **2**:662–663
  Mediterranean Basin **2**:662–663
  North and South America **2**:663–664
  Sub-Saharan Africa **2**:663
native seed collection **6**:440–442
Natrialba **1**:223*T*
Natrix tessellata **5**:477
Natronobacterium **1**:222–223, **1**:223*T*
Natronococcus **1**:222–223, **1**:223*T*
natterjack toad **3**:166
natuna tarsier **6**:256–258*T*
Natural Assets Information System (NAIS) software **5**:358
natural botanical gardens **6**:505
Natural Capital Project (NatCap) **3**:101–102
natural disasters **4**:654
natural enemies hypothesis **1**:375–376
natural extinctions **5:433–441**
  biodiversity trends **5**:439–440
  causal factors
    environmental changes **5**:433–434
    species–species interactions **5**:434
  geographic patterns **5**:439
  prevalence **5**:433
  selectivity factors
    body size
      causal factors **5**:438
      specialization **5**:438
    characteristics **5**:436
    dispersal ability **5**:437–438
    rarity

geographic range **5**:436
local abundance **5**:436
range–abundance relationships **5**:436–437, **5**:437*F*
spatial structure/patchiness **5**:437
specialization
condition ranges **5**:438
reproductive and life history traits **5**:439
resources **5**:438
species–species interactions **5**:438–439, **5**:439*F*
speciation–extinction rate relationship **5**:439–440, **5**:440*F*
species duration
age factors **5**:435–436
taxonomic groups **5**:434–435, **5**:434*T*, **5**:435*F*
Natural Forest Conservation Program (NFCP) **3**:101
natural gas
biodiversity impacts **3**:256–257, **3**:264*T*
carbon dioxide (CO_2) emissions **3**:265*T*
land requirements **3**:265*T*
world energy reserves **3**:265*T*
world energy use **3**:251*F*, **3**:252*T*
natural gas liquids (NGL) **3**:257
Natural Geography in Shore Areas (NaGISA) project **2**:16
Natural History Museum (London) **5**:408*F*, **5**:409*F*, **5**:411*F*
Natural History Museum (Los Angeles County) **5**:408*F*
natural history museums and institutions **5:404–420**
biodiversity science and conservation programs
advisory roles **5**:414–415
challenges **5**:416
characteristics **5**:414–415
cultural knowledge and traditions **5**:415–416
human health issues **5**:415
future outlook **5**:418–419
historical perspective
Africa **5**:406
Asia **5**:406
early history **5**:404–405
European models **5**:405
museum mission evolution **5**:406–407, **5**:407*F*
North and South America **5**:405–406
Oceania **5**:406
museum collections
collection digitization and databases **5**:409–411
collection management and conservation **5**:411
organizational principles **5**:407–409
size and scope **5**:407–409, **5**:408*F*, **5**:409*F*, **5**:411*F*
purpose **5**:404
research programs
biodiversity research **5**:413–414
biogeographic and environmental studies **5**:413
challenges **5**:413–414
genetic and genomic research **5**:412–413
phenomic research **5**:412–413
taxonomy and systematics **5**:412
training and educational programs
citizen science programs **5**:418
exhibition development and promotion **5**:417–418
graduate and postdoctoral training **5**:416
public awareness **5**:417
public educational programs **5**:418
student and teaching training **5**:416–417
see also botanical gardens; zoos/zoological parks
natural predators **1**:709
see also predators
natural reserves and preserves **5:442–450**
biodiversity stewardship **2**:297
biodiversity sustainability **5**:448–449
biosphere reserves **2**:301, **6**:489
boreal forests **1**:634
classifications **2**:289–292, **2**:290, **5**:442–443, **5**:443*T*
climate change
challenges **5**:449
Great Limpopo Transfrontier Park **5**:449
Seamounts Marine Management Area **5**:449–450
common property resources **2**:154
conservancy/megaconservancy networks **2**:301, **2**:302*F*
conservation biology **2**:239–241
funding considerations **5**:448
future threats **5**:448
general discussion **5**:442
historical background **5**:442
indigenous peoples **4**:276–277
marine protected areas **5:94–104**
benefits
abundance increases **5**:96–97, **5**:97*T*
adjacent fisheries **5**:97
biomass increases **5**:96–97, **5**:97*T*
characteristics **5**:96–97
long-term ecological benefits **5**:97–98
nontarget species **5**:97
reproductive potential **5**:96–97
size increases **5**:96–97, **5**:97*T*
Bowie Seamount Marine Protected Area (Sglaan Kinghlas) **5**:446–447
characteristics and functional role **5**:94–95
climate change threats **5**:100–101
conservation strategies **5**:39–40
ecosystem-based management
characteristics **5**:95
definitions **5**:95
goals and objectives **5**:95
marine spatial planning **5**:96
networks **5**:95–96
effectiveness **5**:100
Endeavor Hydrothermal Vents Marine Protected Area **5**:447
fish conservation **3**:453–454
fish stocks **3**:486
geographic extent **5**:445–446, **5**:446*T*
global commitment **5**:99–100
Great Barrier Reef Marine Park **5**:446
Papahanaumokuakea Marine National Monument **5**:446
socioeconomic benefits **5**:98–99, **5**:98*T*
private conservation areas **2**:296–297, **2**:297*T*
terrestrial extent **5**:443–444, **5**:444*T*
terrestrial protected areas
Canaima National Park **5**:444
Dartmoor National Park **5**:445
Jenolan Karst Conservation Reserve **5**:444–445
North East Svalbard Nature Reserve **5**:444
Selous Game Reserve **5**:445
Snares Islands/Tini Heke Islands **5**:444
Tamshiyacu-Tahuayo Communal Reserve **5**:445
threats and challenges **5**:447–448
tourism industry **7**:224
value **5**:448
natural resources
biodiversity conservation **4**:435
land use issues **4**:558
property rights **6:273–279**
biodiversity **6**:273–274
functional role **6**:273
ownership **6**:274, **6**:274*T*
scope **6**:273
value **6**:273
natural rights **4**:653
natural selection
adaptation **1:15–20**
adaptationism
basic concepts **1**:18–19
criticisms **1**:19
post-adaptationism research **1**:19–20
current research areas **1**:19–20
definitions
evolution-based adaptations **1**:16–17
general discussion **1**:16
process-based adaptations **1**:16
evidential perspectives
general discussion **1**:17–18
genomics **1**:18
increased fitness **1**:18
natural selection **1**:17–18
historical perspectives
biblical theories **1**:15
classical era **1**:15
Darwinian theory **1**:15–16
pre-Darwinian theories **1**:15
adaptive landscape theory **2**:510, **2**:511*F*, **6**:632
basic concepts **3**:386
climate change effects **3**:386
cooperation **2**:327–328
Darwinian theory **1**:15–16, **1**:695, **2**:46–47, **2**:263–265
differentiation **2**:509–511
ecological genetics
discrete polymorphisms **2**:717

natural selection (*continued*)
potential biases **2**:717–718
selection gradients **2**:717
evidential perspectives **1**:17–18
evolutionary mechanisms **3**:399
genetic diversity **3**:675
population diversity **6**:168
speciation **6**:631–632
species interactions **6**:719–720
synthetic theory **2**:47
taxonomic versus functional diversity **2**:656
nature-based coastal protection strategies **5:451–458**
biodiversity impacts **5**:453–456, **5**:455*F*
current status **5**:452–453, **5**:453*T*
historical perspective **5**:452
importance **5**:451–452
research summary **5**:456
nature-based recreation **7:241–257**
backpacking/hiking groups
conservation support **7**:251–252
per capita contributions **7**:254*F*
popularity **7**:251
biophilia versus videophilia
comparison studies **7**:242–244
future outlook **7**:250–251
future research outlook **7**:253–256, **7**:255*F*
historical trends **7**:242, **7**:243*F*
videophilia **7**:244
childhood health and development
children–natural environment interactions **7**:248
early television viewing–attention deficit disorder (ADD) relationship **7**:248–249
future outlook **7**:250–251
future research outlook **7**:253–256, **7**:255*F*
indirect, direct, and vicarious experiences **7**:247
Internet use impacts **7**:249
media depictions–fear relationship **7**:249
recommended television-watching guidelines **7**:248
sedentary lifestyles **7**:247–248
socialization impacts **7**:249
virtual nature experiences **7**:249–250
conservation strategies
ecosystem services **7**:253
health benefits **7**:253
conservation support
backpacking/hiking groups **7**:251–252
declining support **7**:251
demographics **7**:252–253
diversification strategies **7**:252
future outlook **7**:252–253
popular activities **7**:251
public land visits **7**:251–252
two-Americas concept **7**:251–252
declining trends **7**:241–242
future research outlook **7**:253–256, **7**:255*F*
habitat degradation **4**:566–567
importance **7**:241–242
modeling frameworks **5**:353
national and international trends
decline incidence **7**:244
international trends **7**:245, **7**:246*F*
per capita visitation **7**:244–245, **7**:245*F*, **7**:246*F*
pervasive declines **7**:247
popularity declines **7**:245–246, **7**:245*F*, **7**:246*F*
public surveys **7**:246
specific-activity trends **7**:246–247, **7**:246*F*
national park visits **7**:242, **7**:243*F*
Nature Reserve System (NRS)
background information **3**:372–373
biodiversity/biophysical effects **3**:373–374
future outlook and challenges **3**:375
geographic distribution **3**:373, **3**:374*F*
number of reserves **3**:373*F*
program goals **3**:372–373
socioeconomic effects **3**:374–375
summary discussion **3**:382
NatureServe Conservation Status methodology **7**:199
Naturhistoriska riksmuseet (Sweden) **5**:411*F*
Nauru **6**:507*T*
nausea **3**:635*T*
Nauticiella stygivaga **7**:58
Nautilus **2**:47, **5**:380
Navahoceros **4**:724–725*T*
Navajo creation tales **4**:650
Navia pungens **1**:475
Neaphaenops tellkampfi **7**:60
neapolitan triton **2**:543
Nearctic region
desert ecosystems **2**:468–469
Diptera **3**:488*T*
edible plants **3**:130–131, **3**:131*T*
fish diversity **3**:472
indigenous peoples **4**:271*T*
Lepidoptera **5**:389*T*
protected area systems **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
Near East ecosystems **5:478–487**
animal diversity **5:459–477**
biodiversity extent **5**:459–460, **5**:460*T*
endemic versus widespread animals **5**:463–464, **5**:464*F*, **5**:464*T*
freshwater fish diversity **5**:470–471, **5**:470*F*, **5**:470*T*, **5**:471*F*
habitat destruction
coastal sand dunes **5**:472, **5**:473*F*
riverine habitats **5**:473
swamps **5**:472–473, **5**:474*F*, **5**:475*T*
temporary winter rain pools **5**:473, **5**:476*F*
wetland habitats **5**:472–473
historic zoogeography **5**:460–463, **5**:461*F*, **5**:462*F*, **5**:463*T*
human impact
agricultural practices **5**:475–477
habitat destruction **5**:472, **5**:473*F*
historical perspective **5**:471–472
hunting impacts **5**:472
poisoning **5**:473–475
influencing factors
dry conditions **5**:466–467
general discussion **5**:464–467
rainfall **5**:465–467, **5**:465*F*, **5**:466*T*
sand adaptability **5**:467–469, **5**:468*F*
substratum **5**:467–469, **5**:467*F*, **5**:468*F*
vegetation **5**:469–470, **5**:469*F*
Oriental representatives **5**:462–463, **5**:463*T*
climate
dew and fog **5**:479–480
rainfall **5**:465–467, **5**:465*F*, **5**:466*T*, **5**:479, **5**:480*F*
temperature regimes **5**:480
deforestation **1**:679*T*
edaphic conditions **5**:478–479
flora **5**:480–481, **5**:481*T*
geographical subdivisions **5**:478, **5**:479*F*
geomorphology **5**:478–479
Levant region **5**:459, **5**:459*F*, **5**:460*T*
rock types **5**:478–479
topography **5**:478
vegetation
Ceratonia siliqua open forests **5**:480*F*, **5**:482–483
deserts **5**:480*F*, **5**:485–486
desert savannoid vegetation **5**:480*F*, **5**:486
dryland vegetation patterns **5**:484
general discussion **5**:481–482
mangroves **5**:480*F*, **5**:487
maquis/forests **5**:480*F*, **5**:481–482
Mediterranean savannoid vegetation **5**:480*F*, **5**:483
Mt. Hermon montane forest **5**:480*F*, **5**:482
Mt. Hermon tragacanth vegetation **5**:480*F*, **5**:483–484
Pistacia lentiscus open forests **5**:480*F*, **5**:482–483
Quercus calliprinos woodlands **5**:480*F*, **5**:482
Quercus ithaburensis open forests **5**:480*F*, **5**:482
reed thickets **5**:480*F*, **5**:486
sand vegetation **5**:480*F*, **5**:486
semisteppe bathas **5**:480*F*, **5**:483
shrub–steppes **5**:480*F*, **5**:484–485
shrub–steppes with trees **5**:480*F*, **5**:485
Sudanian tree oases **5**:480*F*, **5**:486
swamps **5**:480*F*, **5**:486
synanthropic vegetation **5**:480*F*, **5**:486–487
wet salinas **5**:480*F*, **5**:486
Ziziphus lotus **5**:480*F*, **5**:483
see also Middle East
nearly neutral theory **2**:630, **2**:633
Nebela **6**:290–291
Nebraska Sand Hills **5**:562–563
Necator americanus **7**:428*T*
Nectandra spp. **2**:26
Nectarinia fusca **1**:487–488
Nectarinia olivacea **3**:144–145
Nectarinia osea **5**:476
nectar spurs **6**:47–48
Nectophrynoides occidentalis **1**:172

needle-clawed galagos **6**:251*T*
needlefish **3**:467, **5**:668
neem tree **1**:589–590
neem tree oil **5**:693
negative emotional responses **1**:39–40
negative entropy theory **5**:630
neighbor-joining method **6**:18
Nei's distance measurement **4**:510
Neisseria meningitidis **2**:636*F*
nekton **4**:382–383, **4**:383*T*
Nelder model **6**:160–161
Nelumbium speciosum **2**:663
Nelumbo **5**:621*F*
Nelumbo speciosa **1**:590
Nelusetta ayraudi **2**:8–9
Nemachilus insignis **5**:463*T*
Nemachilus leontine **5**:463*T*
Nemasoma varicorne **5**:425–426
Nemata **7**:433*F*, **7**:434*F*
Nemathelminthes **5**:669–670*T*
Nematoda/nematodes **7**:**420–436**
  adaptability **4**:393
  Antarctica **1**:184*T*
  biodiversity
    biodiversity science **7**:428–429
    epidemiology **7**:425–428
    historical perspective **7**:425–428
    human infection incidence **7**:428*T*
    ubiquitous nature **7**:425–428
  biomonitoring processes **3**:122*T*
  cats **5**:665*T*
  characteristics **2**:524*T*, **4**:392, **4**:394*F*, **6**:586, **7**:420–421
  classifications **4**:381*T*, **7**:434–435
  domesticated ruminants **1**:710–711
  ecological niche model **5**:665*T*
  environmental impact
    aquatic and marine nematodes **7**:431
    feeding behaviors **7**:430–431
    plants **7**:430
    predation **7**:430–431
    research background **7**:430
    soils **7**:430
  fossil record **3**:684*F*
  freshwater ecosystems **3**:571–572
  future research outlook **7**:435
  global diversity **4**:385*T*
  global species richness **3**:711
  habitats
    abundance estimates **7**:430
    characteristics **7**:429
    extreme biotopes **7**:429
    habitat diversity **7**:429–430
  herbicide use effects **4**:93
  known and estimated species **3**:710*T*, **4**:384*T*, **7**:431
  lake and pond ecosystems **4**:371–373, **4**:374*F*
  marine environments **4**:382*F*, **5**:50*T*, **5**:106*F*
  molecular level diversity **7**:433
  morphology
    alimentary canal **7**:422–424
    anterior ends **7**:421*F*, **7**:424*F*, **7**:425*F*, **7**:426*F*
    characteristics **7**:420–421
    cuticle **7**:421–422, **7**:422*F*, **7**:423*F*, **7**:424*F*
    eye spots **7**:423*F*
    head **7**:423*F*
    lateral alae **7**:422*F*
    male nematode **7**:423*F*
    midbody cross-section **7**:423*F*
    mouth and teeth **7**:426*F*
    posterior ends **7**:421*F*, **7**:422*F*, **7**:424*F*, **7**:427*F*, **7**:428*F*
    reproductive system **7**:424–425, **7**:425*F*, **7**:427*F*
    setae **7**:423*F*
    spear **7**:427*F*
    stoma **7**:426*F*
  number of species **3**:710*T*, **4**:371*T*, **4**:384*T*, **5**:669–670*T*, **7**:430, **7**:431
  parasitic disease vectors
    fish **5**:672
    humans **5**:670–672
    livestock **5**:672
    plants **5**:672–673
  pest control **5**:673
  phylogenetic relationships **7**:433–435, **7**:433*F*, **7**:434*F*
  Proterozoic–Cambrian transition **3**:396*F*
  sampling methods **4**:381*T*
  soils
    soil fauna **6**:586, **6**:590*T*
    soil quality indicators **6**:597
  species diversity
    bathymetric gradients **4**:388*F*
    Ctenomyidae rodents **7**:432–433, **7**:432*T*
    Geomyidae rodents **7**:432, **7**:432*T*
    latitudinal gradients **4**:387*F*
    marine environments **4**:387
    number of species **7**:431
    vertebrate parasitic nematodes **7**:431–432
  species estimates **3**:710*T*, **4**:384*T*, **7**:431
  synapomorphies **7**:420–421
Nematomorpha
  fossil record **3**:684*F*
  marine ecosystems **5**:50*T*
  number of species **4**:371*T*, **5**:669–670*T*
  phylogenetic relationships **7**:433*F*, **7**:434*F*
Nemertea
  classifications and sampling methods **4**:381*T*
  global diversity **4**:385*T*
  marine ecosystems **5**:50*T*
  number of species **4**:371*T*
  sandy beaches **2**:131–133
Nemertina **3**:684*F*, **4**:384*T*, **5**:669–670*T*
Nemestrinidae **3**:489–491*T*
Nemestrinoidea **3**:489–491*T*
Nemostoma **1**:210
Neoaplectana carpocapsae **5**:673
Neoaves **1**:607
Neocallimastix **3**:627
Neocapritermes **4**:409*T*
Neoceratodus forsteri **3**:461
Neochoerus **4**:724–725*T*, **4**:727*T*
neoclassical economics **3**:289
Neodermata
  Aspidogastrea **7**:439–440*T*, **7**:441*F*, **7**:450–451*F*, **7**:451
  Cestoda
    Amphilinidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
    Gyrocotylidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
    species diversity **7**:439–440*T*, **7**:455–457
  characteristics **7**:445–447
  Digenea
    Allocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Azygioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Bivesiculoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Brachylaimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Bucephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    characteristics **7**:450–451*F*
    Clinostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Cyclocoeloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Diplostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
    Echinostomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452–453
    Gorgoderoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Gymnophalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Haploporoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Haplosplanchoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Hemiuroidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Heronimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Lepocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
    Microphalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453–454
    Microscaphidioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    Monorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    Opisthorchiodea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    Paramphistomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    Plagiorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    Pronocephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    Schistosomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
    species diversity **7**:439–440*T*, **7**:451
    Transversotrematoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454–455
  DNA studies **5**:662–663
Eucestoda

Neodermata (*continued*)
Bothriocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Caryophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cathetocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cyclophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Diphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Diphyllobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Haplobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Lecanicephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Litobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Nippotaeniidea **7**:439–440*T*, **7**:456–457*F*, **7**:458–459
Proteocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Rhinebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Spathebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetrabothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetraphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Trypanorhyncha **7**:439–440*T*, **7**:456–457*F*, **7**:459–460
Monogenea **7**:439–440*T*, **7**:446–447
Monopisthocotylea
Capsalidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Dactylogyridea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Gyrodactylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Lagarocotylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
neodermis **7**:445–447, **7**:446*F*
number of species and families **7**:439–440*T*
Polyopisthocotylea
Chimaericolidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Diclybothriidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Mazocraeidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Polystomatidea **7**:439–440*T*, **7**:448–449*F*, **7**:449–451
species diversity **7**:439–440*T*
Trematoda **7**:439–440*T*, **7**:450–451*F*, **7**:451
Neodiplorchis **7**:448–449*F*
Neodiplostomum **7**:450–451*F*
Neodiprion lecontei **4**:182–183
neoendemics **3**:228
Neofelis nebulosa **1**:517
Neogene
mammals **2**:583*F*
marine diversity **5**:168*F*, **5**:169*F*
temperate forests **7**:122–124
Neoglyphea neocaledonica **2**:5*F*
Neolepidoptera **5**:386, **5**:386*F*, **5**:387–388*T*
neoliberal economics **2**:702
Neolithic agriculture **2**:658
Neomachilellus adisi **1**:247
Neomeniomorpha **5**:376
Neomorphinae **5**:502*T*
Neomysis mercedis **5**:61
Neophoca cinerea **5**:81*T*
Neophocoena phocoenoides **4**:693–694
Neophocoena phocoenoides asiaorientalis **5**:81*T*
Neophron percnopterus **5**:474
Neopilina galatheae **5**:377–378
Neoplatycephalus richardsoni **2**:8–9
Neoproterozoic era **3**:682–683, **5**:168*F*, **5**:175–176
Neopseusida **5**:385–386
Neopseustoidea **5**:386*F*, **5**:387–388*T*
Neoptera **4**:308–312, **4**:309*T*
Neoraimondia gigantea **3**:115–116
Neosarmatium smithi **5**:15*F*
Neosclerocalyptus **4**:727*T*
Neosittidae **4**:33, **4**:34*F*
Neotermes **4**:409*T*
Neotermes koshunensis **4**:418
Neothoracophorus **4**:727*T*
Neotis ludwigii **1**:487–488
Neotoma albigula **2**:472
Neotragus batesii **1**:49
Neotragus moschatus **6**:572–573
Neotragus pygmaeus **1**:49
neotropics
Amazonian terrestrial arthropods **1**:246–247
butterflies **1**:657–658, **1**:658*T*, **1**:659*T*
desert ecosystems **2**:469
Diptera **3**:488*T*
edible plants **3**:130–131, **3**:131*T*
fish diversity **3**:472
indigenous peoples **4**:271*T*
Lepidoptera **5**:389–390, **5**:389*T*
mammals **2:582–594**
background information **2**:582–583
biogeographical–macroecological patterns
areographic patterns **2**:588–589
ecogeographic rules **2**:588–589
latitudinal diversity **2**:588*F*
species richness **2**:588, **2**:589*F*, **2**:591*F*
conservation strategies **2**:591–592
evolutionary history **2**:582–583, **2**:583*F*
future research outlook **2**:592
Great American Biotic Interchange (GABI) **2**:582–583, **2**:584*F*
harvesting threats **2**:590*T*
major threats **2**:589–592, **2**:590*T*
zoogeographic regions
biogeographical limits **2**:583–585, **2**:585*F*
group diversity and classifications **2**:585–586
representative species **2**:586–587, **2**:587*F*
transition zones **2**:583–585, **2**:585*F*
neotropical seasonally dry forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
oil-palm plantations **5**:602
plant diversity **6**:314*T*
primates
primate diversity **6**:250–252, **6**:251*T*
threatened species **6**:258*T*
protected area systems **2**:591–592, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
species richness **1**:247–249, **1**:248*F*
tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*
Neoxythece elegans **1**:163
Nepal
animal species **1**:261*T*
botanical gardens **6**:519
deforestation **1**:677*F*
ecosystems
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
vegetation types and coverage percentage **1**:264–265*T*
indigenous peoples **4**:270*T*
plant species **1**:259*T*
reforestation programs **6**:378–379
wheat production **1**:87–88
Nepenthaceae **6**:44*T*, **6**:316*T*
Nepenthes **1**:267, **6**:44*T*
Nepenthes alata **6**:45
Nepenthes rajah **2**:643
Nephelium lapaceum **2**:663
nephridia **7**:415
Nephtyidae **7**:417*F*
Nepomorpha **7**:293
Nepticulidae **5**:385*F*
Nepticulina **5**:386, **5**:386*F*, **5**:387–388*T*
Nepticuloidea **5**:386*F*, **5**:387–388*T*
Neptunia oleracea **1**:164
Nereididae **7**:417*F*
Neriidae **3**:489–491*T*
Nerioidea **3**:489–491*T*
Nerium oleander **5**:486
Nesoenas mayeri **7**:188
Nesokia indica **5**:463*T*
Nesolagus netscheri **4**:704
Nesolagus timminsii **4**:704
Nesophontidae **4**:690–691, **4**:706
nestedness **2**:555–557, **6**:649–650
Nestor meridionalis **5**:370, **7**:184

nest parasitism **5**:**501–509**
  adaptive behaviors **5**:501
  background information **5**:501
  conspecific brood parasitism (CBP)
    definition **5**:501
    prevalence **5**:508–509
    species survey **5**:502–503
  cowbird management **5**:507
  fish **5**:503
  host–parasite coevolution
    cowbird–host systems
      behavioral adaptations **5**:504–505
      egg adaptations **5**:504
      host defenses **5**:505
      nestling and fledgling adaptations **5**:505
    cuckoo–host systems
      adaptive behaviors **5**:503
      egg adaptations **5**:503
      evolutionary development **5**:503
      fecundity **5**:503
      host defenses **5**:504
      nestling and fledgling adaptations **5**:503–504
    modeling approaches **5**:506
  host population dynamics **5**:506–507
  insects **5**:503
  major groups **5**:502*T*
  number of species **5**:502*T*
  research outlook **5**:509
  social behavior
    foraging ecology **5**:507
    mating systems **5**:507–508
    spacing behaviors **5**:508
    vocal behaviors **5**:508
  species survey
    black-headed ducks **5**:502, **5**:502*T*
    conspecific brood parasitism (CBP) **5**:502–503
    cuckoo-finches **5**:502, **5**:502*T*
    Cuculidae **5**:501–502, **5**:502*T*
    Icteridae (cowbirds) **5**:502, **5**:502*T*
    Indicatoridae **5**:502, **5**:502*T*
    viduine finches **5**:502, **5**:502*T*
net above-ground primary production (NAPP) **3**:75, **3**:78*F*, **3**:79*F*
Netherlands
  botanical gardens **6**:519
  natural history museums and institutions **5**:411*F*
  nitrogen cycle impacts **5**:543
  nitrogen deposition **5**:529*F*
  urbanization **4**:560*F*
  vegetation surveys **3**:341–344*T*
Netherlands Antilles
  botanical gardens **6**:519
  energy efficiency measures **3**:254*F*
  urbanization **4**:560*F*
net national product (NNP)
  background information **3**:25–26
  social well-being
    allocation mechanisms **3**:26
    basic concepts **3**:26
    constant-equivalent consumption measures **3**:27
    cost–benefit rule **3**:27
    population changes **3**:26–27
    sustainable development **3**:27–28
net primary production (NPP)
  biogeochemical cycles **1**:559, **1**:560*T*
  C_4 plants **2**:373
  climate change effects **5**:37–38
  ecosystem function measurements **3**:73–76, **3**:74–75*T*, **3**:76*F*, **3**:78*F*, **3**:79*F*
  ecosystem productivity **3**:91*T*, **4**:156–158
  plant biodiversity–climate relationship **2**:117–118, **2**:120, **3**:118, **3**:121
  trophic dynamic regulation **3**:240
  valuation measures **3**:58
net-winged midges **3**:489–491*T*
Neurochaetidae **3**:489–491*T*
neurofibromatosis 1 (NF1) tumor suppressor **5**:577
Neuroptera **4**:308–312, **4**:309*T*, **4**:371*T*, **4**:392*T*
neurotoxin shellfish poisoning (NSP) toxins **1**:470
neutral evolution theory **2**:263–264, **2**:508–509, **2**:509*F*, **6**:191
neutral theory **5**:**510–518**
  biodiversity dynamics modeling **5**:326–327
  characteristics
    dispersal **5**:511
    general discussion **5**:511
    speciation **5**:511
    stochastic drift **5**:511
  definition **5**:510, **5**:511
  future outlook
    general discussion **5**:516
    testing programs **5**:516
    theoretical developments **5**:516–517
    theory applications **5**:517
  general discussion **5**:510
  genetic variation processes
    phylogenetic divergence **2**:633
    polymorphism **2**:629–630
  goals **5**:510–511
  historical background **5**:511–512
  modeling approaches
    alternative neutral models
      spatially explicity neutral models **5**:514–515
      speciation modes **5**:515
    Hubbell's spatially implicit neutral model **5**:512–514
    species abundance distributions (SADs) **5**:513–514, **5**:513*F*
    species–area relationships (SARs) **5**:514
  molecular level diversity
    phylogenetic divergence **2**:633
    polymorphism **2**:629–630
  nonneutral extensions **5**:515–516
  research summary **5**:517
  schematic diagram **5**:512*F*
  species abundance distributions (SADs) **5**:510–511, **5**:513*F*
  species–area relationships (SARs) **5**:510–511, **5**:514
Nevadan Orogeny **5**:563
New Biological Nomenclature **5**:556
New Caledonia
  botanical gardens **6**:507*T*
  coral reef ecosystems **2**:333*T*
  disharmony **5**:591–592
  endemism **3**:233
  fire ant invasions **5**:598*F*
  plant species **6**:61, **6**:82*T*
  protected area systems **6**:493–494
New Caledonian rail **7**:184
New England warblers **2**:188–189, **2**:188*F*
Newfane Wildlife Habitat Improvement Group **7**:21
New Guinea
  butterflies **1**:658, **1**:659*T*
  coral reef ecosystems **2**:333*T*
  endangered mammals **3**:194*T*
  fish diversity **3**:473
  termites **4**:409*T*
  vascular plant diversity **6**:82*T*
newts **5**:315–316, **5**:473, **5**:723, **7**:389–390
new world cuckoos **5**:501–502, **5**:502*T*
New World vultures **1**:617
New York Botanical Garden **3**:138, **5**:405–406, **5**:411*F*, **6**:510–511
New Zealand
  birds
    flightless birds **1**:606
    introduced species **1**:620
    New Zealand ground parrot **1**:604, **1**:606, **6**:579, **7**:188
  botanical gardens **6**:519
  C_4 plants **2**:371*T*
  disharmony **5**:591–592
  domesticated ruminants **1**:702–703
  ecosystems
    alpine life zone **1**:149*T*
    conservation status **3**:171–173*T*
    Southern Hemisphere **6**:**612–619**
      angiosperms **6**:617–618
      background information **6**:612
      biotic dispersal **6**:616
      climate change effects **6**:617
      climatic patterns **6**:614
      continental drift **6**:616
      dispersal ability **6**:613–614
      fossil record **6**:615–616
      general area cladograms **6**:614
      geographic distribution patterns **6**:613–614
      high southern latitude plant growth **6**:618
      historical research **6**:612–613
      long-distance dispersal **6**:616–617
      relaxed clock molecular dating analyses **6**:614–615
      research summary **6**:618–619
      vicariance biogeography **6**:616
  endangered plant species **3**:206–208, **3**:207*F*
  endemism **3**:233
  energy use **3**:252*T*
  fish diversity **3**:473, **3**:474
  indigenous peoples **4**:269*T*, **4**:270
  invasive species
    aquatic ferns (*Salvinia molesta*) **4**:346–347

New Zealand (*continued*)
Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346, **4**:346*F*
Fallopia japonica **4**:351–352
mouse-ear hawkweed (*Hieracium pilosella*) **4**:348–349, **4**:349*F*
terrestrial invertebrates **3**:222
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*, **4**:730–731
modern extinctions
first-contact extinctions **5**:366
historical extinctions **5**:366–367
natural history museums and institutions **5**:406
plant species **6**:82*T*
protected area systems **6**:491*T*, **6**:496–497
reforestation programs **6**:377
return-on-investment (ROI) analysis **4**:193, **4**:194*T*
Snares Islands/Tini Heke Islands **5**:444
temperate forests
distribution patterns **7**:116*T*, **7**:117
diversity patterns **7**:119–121, **7**:119*T*, **7**:120*F*
dominant floristic features **7**:119
endemism **7**:121
physiognomic characterization **7**:117
termites **4**:409*T*
threatened birds **7**:181
New Zealand cockle **6**:160*F*
New Zealand flatworm **6**:240
New Zealand Plant Network **6**:86
New Zealand shore plover **4**:194*T*
Nezara viridula **7**:294
Ngami (Lake) **4**:461*T*
Nicaragua
botanical gardens **6**:519
deforestation **2**:30–31, **2**:31*F*
education level–fertility relationship **6**:204, **6**:205*F*
education level–infant mortality relationship **6**:205*F*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
pastureland conversion **4**:665*T*
species richness **2**:25*T*
Nicaragua (Lake) **4**:461*T*
niches
basic concepts **6**:693–694, **6**:694*F*
ecological niche model **6:692–705**
basic concepts
geographical models **6**:694–695
niche theory **6**:693–694, **6**:694*F*
spatial autocorrelation (SAC) **6**:693
species response shapes **6**:692*F*, **6**:694
characteristics and functional role **6**:692–693, **6**:692*F*
community studies **6**:702–703
Cupressus chimera case study **6**:695
data collection methodology
biotic interactions and disturbances **6**:698–699
challenges **6**:695–697
environmental data **6**:699
species data **6**:696–697
environmental data
aquatic systems **6**:700
challenges **6**:699
climate **6**:699
land cover **6**:699–700
multicollinearity **6**:700
soils and substrate **6**:699–700
terrain **6**:699
modeling approaches
Bayesian methods **6**:702
decision trees **6**:701
general discussion **6**:700–702
generalized additive models (GAMs) **6**:701
generalized linear models (GLMs) **6**:701
genetic algorithms **6**:702
machine learning methods **6**:701, **6**:702*F*
maximum entropy theory (MaxEnt) **6**:702
profile and envelope methods **6**:701–702
new times and places predictions **6**:703
research summary **6**:703–704
species data
observational research **6**:696–697
presence-only data **6**:697–698, **6**:701–702
record characteristics **6**:697, **6**:698*F*
sample characteristics **6**:696–697, **6**:696*F*
site-occupancy models **6**:697, **6**:698*F*
species packing **4**:643
species richness **6**:702–703
taxonomic versus functional diversity **2**:649–650, **2**:649*F*
evolutionary development **4**:46–47
habitat–niche relationship **4:39–49**
basic concepts **4**:39–40, **4**:40*F*
conservation and management practices **4**:47–48
ecological character displacement **4**:46–47, **4**:46*F*, **4**:47*F*
evolutionary development **4**:46–47
food resources **4**:44*F*
future research outlook **4**:48
historical perspective **4**:40–42
indirect interactions **4**:45–46
multispecies comparisons **4**:42–45, **4**:42*F*, **4**:43*F*, **4**:44*F*, **4**:45*F*
mutualistic interactions **4**:45
representative species **4**:41*T*
niche complementarity **5**:451–452, **5**:455
niche differentiation
empirical evidence
relevance **6**:434–435
spatial heterogeneity **6**:432–433, **6**:433*F*
temporal heterogeneity **6**:433–434, **6**:434*F*
theoretical perspectives **6**:430–432, **6**:431*F*, **6**:432*F*
parasites
infection prevalence **5**:665–666, **5**:666*F*
microhabitat segregation **5**:665–666, **5**:667*F*
niche dimensions **5**:665, **5**:665*T*
niche saturation **5**:665, **5**:666*F*
proximate and ultimate causation **5**:665–666, **5**:667*F*
plant community evolution **6**:67
species coexistence
classical niche partitioning **6**:671–672
niche construction **6**:676–677
temporal niche partitioning **6**:675
temporal niches **6**:675, **7**:7
Nicholson–Bailey model **5**:679
Nicobar Islands **1**:262–263*T*
Nicotiana tabacum **2**:663–664, **4**:289*T*
nicotinamide adenine dinucleotide phosphate (NADPH) **2**:362*F*
nicotinamide adenine dinucleotide phosphate (NADPH) dehydrogenase **6**:12–13
nicotine **2**:426*T*, **2**:427, **5**:693
Nicrophorus americanus **3**:220–221
Nidovirales **5**:134*T*, **5**:137
Niger
botanical gardens **6**:519
cropland conversion **4**:665*T*
energy efficiency measures **3**:254*F*
intercropping systems **2**:382–383
urbanization **4**:560*F*
Niger Delta red colobus **6**:256–258*T*
Nigeria
botanical gardens **6**:507*T*
deforestation **2**:449*T*, **2**:451*T*
energy efficiency measures **3**:254*F*
forest product exports **2**:451*T*
indigenous peoples **4**:270*T*
rice production **1**:88
termites **4**:430*T*, **4**:432*T*
traditional sacred forests **6**:384
night heron **5**:476–477
nightingales **1**:616*F*, **4**:573
night monkeys **2**:590*T*, **6**:251*T*
Nile crocodiles **5**:472
Nile perch **2**:223–224, **3**:468, **4**:360–361, **4**:465, **4**:663
Nile River **7**:146
Nilgiri langurs **6**:253–254
Nimaviridae **5**:130*F*, **5**:134*T*, **5**:135–136
nine-banded armadillo **2**:590*T*
Ninox novaeseelandiae **7**:189
Niphates digitalis **5**:124*F*
Nippotaenia **7**:456–457*F*
Nippotaeniidea **7**:439–440*T*, **7**:456–457*F*, **7**:458–459
nirvana **6**:564
Nitidulidae **1**:327*T*
Nitocra balnearia **5**:470–471
Nitocra incerta **5**:470*T*
Nitraria retusa **1**:268, **5**:484–485
nitrate fertilizers **3**:166
see also fertilizers
nitrification **3**:90*T*
Nitrobacter **6**:32
nitrogen-fixing prokaryotes **6**:449–450
nitrogen (N)
air pollution impacts **1**:142–144, **1**:143*F*, **1**:146*F*
ammonia (NH_3)

ammonia-oxidizing bacteria (AOB) **5**:257–258
atmospheric concentrations **1**:282*T*
biogeochemical cycles **1**:560–561, **1**:562*F*
emission occurrences and sources **1**:1–5, **1**:137*T*
geographic distribution **1**:3*F*
nitrogen cycle **5**:539*F*
nutrient cycling **1**:706
atmospheric concentrations **1**:282*T*
biodiversity impacts **1**:142–144, **1**:143*F*
biogeochemical cycles
acid rain **6**:151–152
aquatic ecosystems **6**:152, **6**:153*F*
forest ecosystems **6**:152
key reactions **1**:560–561, **1**:562*F*, **6**:581–582
landscape legacies **4**:527
nitrogen oxides (NO_x) **1**:560–561, **1**:562*F*, **6**:151–152
nitrogen saturation **6**:152
soils **6**:581–582, **6**:582*F*
tropospheric ozone precursors **6**:152–154
carbon:nitrogen (C:N) stoichiometry **6**:121–123
carbon:nitrogen:phosphorus (C:N:P) stoichiometry **3**:506–507, **4**:675–677, **4**:676*F*
denitrification **3**:90*T*, **6**:151–152, **6**:151*F*
desert ecosystems **2**:474
dinitrogen-fixing (N_2-fixing) microorganisms **2:608–625**
basic concepts **2**:608–610, **2**:609*F*
biodiversity
dinitrogenase/dinitrogenase reductase **2**:610–611
general discussion **2**:610–611
metabolic diversity **2**:614–615
nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
phylogenetic relationships **2**:611*F*
physiological diversity **2**:614–615
biogeochemical cycles **1**:560–561
emerging patterns
changing environmental conditions **2**:621–623
environmental modifications **2**:619–621
spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
functional role
biological carbon (C) pump **2**:615–618, **2**:616*F*
nutrient cycle linkages **2**:618–619
quantification methods **2**:615–618, **2**:617*F*
Redfield ratio **2**:617–618, **2**:617*F*
habitat characteristics **2**:609*F*
research summary **2**:623
ecosystem function measurements **3**:69–70, **3**:90*T*
elemental partitioning **1**:554*T*
eutrophication **3**:366–367
global nitrogen cycle **3**:282, **5**:111, **5**:307, **5**:307*F*, **5**:520–522, **5**:521*F*
inorganic nitrogen toxicity **3**:366–367
intercropping systems **2**:389–390
nitrate deposition **1**:3*F*, **6**:125–126
nitric acid **1**:1–5, **1**:3*F*
nitrification **1**:561, **1**:562*F*, **3**:90*T*, **6**:151–152, **6**:151*F*, **6**:581–582, **6**:582*F*
nitrogen cycle **5:537–546**
acid rain **6**:151–152
anthropogenic impacts **5**:520–522, **5**:522*F*, **5**:538–540, **5**:540*F*, **5**:541*F*
aquatic ecosystems **6**:152, **6**:153*F*
background information **5**:537
basic concepts **5**:537–538
biodiversity impacts
eutrophication **1**:95–97, **1**:96*F*, **5**:542–544
future outlook **5**:546
genetic diversity loss **5**:540–544
grasslands **5**:542, **5**:542*F*
nitrate deposition **6**:125–126
nitrogen dynamics **5**:544–546, **5**:545*F*
nitrogen oxide (NO_x) acidification **5**:544
dinitrogen-fixing (N_2-fixing) microorganisms **2:608–625**
basic concepts **2**:608–610, **2**:609*F*
biodiversity **2**:610–611
biogeochemical cycles **1**:560–561
changing environmental conditions **2**:621–623
dinitrogenase/dinitrogenase reductase **2**:610–611
emerging patterns **2**:619–621
functional role **2**:615–618, **2**:616*F*
habitat characteristics **2**:609*F*
metabolic diversity **2**:614–615
nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
nutrient cycle linkages **2**:618–619
phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
phylogenetic relationships **2**:611*F*
physiological diversity **2**:614–615
Redfield ratio **2**:617–618, **2**:617*F*
research summary **2**:623
spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
dynamic global vegetation models (DGVMs) **2**:680
ecological importance **5**:537–538
forest ecosystems **6**:152
freshwater ecosystems **3**:572*T*
future outlook **5**:546
global distribution **5**:524*F*
global nitrogen budget **5**:523*F*
historical perspective **5**:520–522
landscape legacies **4**:527–528
nitrogen oxides (NO_x) **1**:560–561, **1**:562*F*, **6**:151–152
nitrogen saturation **6**:152
schematic diagram **5**:521*F*, **5**:523*F*, **5**:539*F*, **6**:151*F*
tropospheric ozone precursors **6**:152–154
nitrogen deposition **5:519–536**
air pollution impacts **1**:142–144, **1**:143*F*, **1**:146*F*
anthropogenic impacts **5**:520–522, **5**:522*F*
biodiversity impacts
acidification **5**:522, **5**:527*F*
direct foliar toxicity **5**:523–525, **5**:527*F*
eutrophication **5**:522
general discussion **5**:522
hotspots **5**:526*T*
management strategies **5**:533
natural recovery potential **5**:532–533
primary drivers **5**:522, **5**:525*F*, **5**:527*F*
secondary stressors **5**:522–523, **5**:527*F*
vulnerability factors **5**:530–531
biological nitrogen fixation **5**:520*F*
environmental impact **5**:519–520
global distribution **5**:524*F*
global nitrogen budget **5**:523*F*
global nitrogen cycle **5**:520–522, **5**:521*F*
influencing factors
ecosystem-specific effects **5**:526–528
taxa-specific responses **5**:528
management strategies
critical load analyses **5**:533, **5**:534*F*
government policies and regulations **5**:533–534
intervention approaches **5**:533–534
modeling approaches **5**:533
monitoring networks **5**:533
natural recovery potential **5**:532–533
nitrogen as nutrient **5**:520, **5**:520*F*
research methodologies **5**:525–526, **5**:528*T*, **5**:529*F*
research summary **5**:534–535
sensitivity factors
abiotic factors **5**:531
biotic factors **5**:531–532
species richness **1**:142–144, **1**:143*F*, **1**:146*F*
taxa-specific responses
herbaceous plants **5**:528–530
higher-taxa responses **5**:530
nonvascular plants **5**:528
soil biota **5**:530
trees **5**:530
vulnerability factors
biodiversity impacts **5**:530–531
exposure characteristics **5**:530–531
sensitivity characteristics **5**:531
site history **5**:532
nitrogen fixation
anthropogenic impacts **5**:520–522, **5**:522*F*
biogeochemical cycles **1**:560–561, **1**:562*F*, **6**:151–152, **6**:151*F*, **6**:581–582, **6**:582*F*
ecosystem services **3**:90*T*
importance **5**:520, **5**:520*F*, **5**:538

nitrogen (N) (*continued*)
nitrogen oxides (NO_x)
atmospheric concentrations **1**:282*T*, **6**:151–152, **6**:151*F*
biogeochemical cycles **1**:560–561, **1**:562*F*, **6**:151–152, **6**:151*F*
changing concentrations **1**:286–287
climate impact **1**:281–282
dynamic global vegetation models (DGVMs) **2**:680
ecotoxicological studies **3**:118–119
emission occurrences and sources **1**:137*T*
eutrophication **3**:366–367
forest ecosystems **1**:5*T*
nitrogen cycle **5**:539*F*, **5**:544
nitrogen:phosphorus (N:P) ratio **2**:617–618, **2**:617*F*
nitrous oxide (N_2O)
atmospheric concentrations **1**:282*T*, **6**:151–152, **6**:151*F*
biogeochemical cycles **1**:560–561, **1**:562*F*, **6**:151–152, **6**:151*F*
biomass emissions **4**:295–296
changing concentrations **1**:284
enhanced greenhouse effect **4**:18–19, **4**:19*F*, **4**:19*T*
nutrient cycling **1**:706
radiative forcing effects **4**:20*F*
source–sink dynamics **1**:284
nitrous oxide (NO_2) **2**:680
nutrient cycling **1**:63–65, **1**:705–706, **1**:705*F*, **3**:348–351, **3**:349*F*
nutrient management–water quality relationship **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
freshwater invertebrates **4**:377
land-use change impacts **1**:107–108
leaching vulnerability index **1**:102*F*
management criteria **1**:97, **1**:97*T*
nutrient input and sources **1**:96*F*, **1**:98, **3**:575*F*, **3**:576–577, **3**:576*F*
regional nutrient imbalances **1**:98*F*, **1**:99–101, **1**:100*F*, **1**:102*F*
remedial benefical management practices (BMPs)
characteristics **1**:103*T*
cover crops **1**:103*T*, **1**:106
critical source area management **1**:103*T*, **1**:107
erosion potential **1**:103*T*, **1**:105–106, **1**:105*T*, **1**:106*F*
farm input decisions **1**:101, **1**:103*T*
fertilizer application methods **1**:103*T*, **1**:104
fertilizer application rate and frequency **1**:103*T*, **1**:104–105
implementation strategies **1**:101, **1**:103*F*
nutrient-use efficiency improvements **1**:101–104, **1**:103*T*
riparian/buffer zones **1**:103*T*, **1**:106–107
runoff potential **1**:103*T*, **1**:105
soil and plant testing **1**:103*T*, **1**:104
source management **1**:103*T*, **1**:104
transport management **1**:103*T*, **1**:105
research background **1**:95–97
research summary **1**:108
peroxyacetyl nitrate (PAN) **1**:286
relative abundance **1**:553*F*
river water composition **3**:562*T*
root–soil interface **6**:114*F*
soil accumulations **1**:6
total phosphorus/total nitrogen–chlorophyll *a* relationship **3**:348–350, **3**:349*F*
water-column nitrogen:phosphorus (N:P) ratios **3**:350–351, **3**:350*F*
Nitrosomonas **6**:32
Nitrosopumilis **5**:294
Nitzschia **1**:475, **6**:291–292
Nitzschia sturionis **2**:528*T*
Niue **6**:507*T*
Noaea **1**:268
Noaea mucronata **5**:483, **5**:484–485
Noah Principle **4**:125
noble fir **3**:639–640, **3**:639*F*
Nocotiana tabacum **5**:693
Noctilio leporinus **2**:587*F*
Noctilionidae **6**:309*T*
Noctiluca **6**:294–295
Noctilucales **1**:472
Noctuidae **5**:387–388*T*, **5**:389–390, **5**:399*F*
noctuid moths **6**:142
Noctuoidea **5**:386*F*, **5**:387–388*T*, **5**:389–390
Nodaviridae **5**:134*T*, **5**:138–139
Noditermes **4**:409*T*
Nodularia spumigena **2**:611*F*, **3**:363–364
noisy scrub-bird **7**:183–184
Nomadacris septemfasciata **2**:549–550
Nomascus **4**:692–693
Nomascus concolor concolor **6**:256–258*T*
Nomascus concolor furvogaster **6**:256–258*T*
Nomascus concolor jingdongensis **6**:256–258*T*
Nomascus concolor lu **6**:256–258*T*
Nomascus gabriellae **6**:256–258*T*
Nomascus hainanus **6**:250, **6**:256–258*T*
Nomascus leucogenys **6**:256–258*T*
Nomascus nasutus **6**:256–258*T*
Nomascus siki **6**:256–258*T*
nomenclature, systems of **5:547–559**
alternative nomenclature systems **5**:556
Code system
algae, fungi, and plants **5**:550–551, **5**:553*T*
ambiregnal organisms **5**:554–555
cultivated plants/cultivars **5**:551, **5**:553*T*
Draft BioCode **5**:553–554
fossils **5**:555
fungi **5**:550–551, **5**:553*T*, **5**:555–556
hybrids **5**:555
lichens **5**:555
PhyloCode **5**:554
prokaryotes **5**:551–552, **5**:553*T*
viruses **5**:552
zoological organisms **5**:552–553, **5**:553*T*
functional role **5**:547
general discussion **5**:547
hierarchical organization **5**:548–549, **5**:548*T*
historical perspective **5**:547–548
name-formation principles **5**:549–550
naming process
name changes **5**:556–557
name determination guidelines **5**:557–558
newly discovered species **5**:557
reference indices **5**:558
nominalism **7**:101
nonadaptive radiations **1**:21
nonallopatric speciation
autopolyploid speciation **6**:626
basic concepts **6**:621*F*, **6**:624–625
cichlids **6**:625
general characteristics **6**:621–624, **6**:622*T*
interspecific hybridization
allopolyploid/amphiploid speciation **6**:627, **6**:627*F*
characteristics **6**:626
direct allogenous speciation **6**:627
introgressive hybrid speciation **6**:626
recombinational hybrid speciation **6**:626
reticulate allogenous speciation **6**:627–628
wild sunflowers **6**:626–627
Mimulus cupriphilis **6**:625
parapatric speciation **6**:621*F*, **6**:622*T*, **6**:625
Rhagoletis fruit flies **6**:625–626
spontaneous thelytokous speciation **6**:626
sympatric speciation **6**:621*F*, **6**:622*T*, **6**:625
noncrop habitats **1**:380–381
noncrop plants **1**:376, **1**:376*F*
nonequilibrium thermodynamics **2**:638–639
nonextremeophilic Archaea **1**:223, **1**:224*T*
nongovernmental organizations (NGOs)
biodiversity conservation **4**:435–436, **4**:438–439, **4**:712–713
bioparasites **1**:597
bioprospecting **1**:596–597
Community-Based Natural Resources Management (CBNRM) **2**:297–298
community seed banks **3**:644
family planning programs **6**:207
illegal wildlife trade **1**:518, **1**:521, **1**:521*F*
livestock genetic resources **1**:645
nature-based recreation support **7**:251, **7**:254*F*
payment for ecosystem services (PES) **5**:151
people–conservation relationship **2**:231–232
United Nations agencies **4**:324–325
nonhomologous chromosome segregation **6**:353
Nonindigenous Aquatic Nuisance Prevention and Control Act (1990) **3**:715

nonmarket regulating services **3**:48–51, **3**:49*T*
non-melanoma skin cancers (NMSCs) **7**:302
nonmethane hydrocarbons **1**:282*T*, **1**:286
non-oxygen (O_2) evolvers *see* bacteriochlorophylls
nonpolychaete worms **5**:20*T*
nonprofit environmental groups **3**:139
nonprotein amino acids **2**:426*T*, **2**:427
nonseed gene banks
  animal germplasm samples **1**:645–646, **3**:651–652
  botanical gardens **3**:653
  cryopreservation **1**:645–646, **3**:651–652
  DNA banks **3**:652, **4**:319–320, **4**:321*T*
  field gene banks **3**:33, **3**:652–653, **4**:319, **4**:321*T*
  microorganisms **3**:652, **3**:652*T*
  pollen **3**:650, **4**:319–320, **4**:321*T*
  somatic and zygotic plant embryos **3**:650–651
  spores **3**:650
  vegetative tissues **3**:651
  zoos/zoological parks **3**:653
nonvascular epiphytes
  characteristics **3**:517–518
  taxonomic distributions **3**:520, **3**:521*T*
nonvascular plants **5**:528, **6**:58–59, **6**:58*T*
nonylphenols **3**:581
noog **2**:663
Norantea **3**:522*T*
Norantea goyasensis **1**:167
Nordic Gene Bank **3**:652–653
norflurazon **4**:90*T*
Norfolk Island **6**:519
Norian event **3**:415–418, **3**:416*T*, **5**:161*F*
normalized difference vegetation index (NDVI) **3**:76, **4**:504
Normapolles **5**:624
Norris, Frank **4**:652
North Africa
  deforestation **2**:448*T*
  desert ecosystems **2**:469–470, **2**:470*T*
  regional carbon stocks **1**:678*F*
North America
  biodiversity patterns **5:560–567**
    anthropogenic impacts **5**:566
    continental characteristics and diversity **5**:560–561
    evolutionary changes **5**:565–566
    invasive species **5**:566
    organism groups
      amphibians **5**:565
      birds **5**:565
      butterflies **5**:565
      freshwater fish **5**:565
      general discussion **5**:564–565
      land mammals **5**:565
      plants **5**:564–565
      reptiles **5**:565
    research summary **5**:566–567
    terrestrial bioregions
      boundaries and characteristics **5**:561–562, **5**:561*F*
      Canadian Shield **5**:561*F*, **5**:562
      Caribbean region **5**:561*F*, **5**:564
      Eastern North America **5**:561*F*, **5**:562–563
      Southernmost North America **5**:561*F*, **5**:563–564
      Southwestern North America **5**:561*F*, **5**:563
      Western North America **5**:561*F*, **5**:563
boreal forests **1**:626, **1**:630*T*, **1**:631*T*, **1**:632*T*, **1**:633*F*
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
butterflies **1**:657*T*
C_4 plants **2**:371*T*
deforestation **1**:679*T*, **2**:448*T*, **2**:452*T*
domesticated ruminants
  historical perspective **1**:702–703
  plant tolerance–grazing relationship **1**:706–708
ecosystems
  alpine life zone **1**:149*T*
  conservation status **3**:171–173*T*
  deserts **2**:470*T*
  forest ecosystems **3**:532*T*
  freshwater ecosystems **3**:578–579, **3**:579*T*
  river water composition **3**:562*T*
  species declines **3**:578–579, **3**:579*T*
freshwater ecosystems **4**:370*T*
grasslands and savannas **6**:331–333
historical conservation practices **2**:279–280
indigenous peoples **4**:269*T*
invasive species
  Bromus tectorum **4**:347–348, **4**:348*F*
  Fallopia japonica **4**:351–352
  Spartina alterniflora **4**:349–350
lake and pond ecosystems **4**:461*T*
land-use trends **4**:559*F*
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
latitudinal gradients **1**:498*F*, **1**:499*T*
mammal extinctions **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
mammalian extinctions **4**:705*T*
marsupials **4**:683, **4**:686–687
megaherbivores **5**:224*T*, **5**:225*T*, **5**:232*F*
myriapods **5**:429*T*
native crops **2**:663–664
natural history museums and institutions **5**:405–406
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
plant invasions **6**:93
postglacial extinctions **5**:363
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional cultivars **3**:32*F*
temperate forests
  distribution patterns **7**:116–117, **7**:116*T*
  diversity patterns **7**:119–121, **7**:119*T*, **7**:120*F*
  dominant floristic features **7**:118
  endemism **7**:121
  physiognomic characterization **7**:117
  termites **4**:409*T*
  transboundary protected areas **6**:489*T*
North American Association for Environmental Education (NAAEE) **3**:136
North American chestnut trees **3**:411
North American Free Trade Agreement (NAFTA) **7**:18–19
North American whipsnakes **6**:409
North Atlantic gray whale (*Eschrichtius robustus*) **5**:76
North Atlantic Oscillation (NAO) **5**:686–688, **5**:686–687, **5**:687*F*, **5**:720–721
North Atlantic right whale **3**:192
North East Svalbard Nature Reserve **5**:444
northern Andean alpine cushion formations **4**:103
northern Andean dwarf *Polylepis* forest **4**:103
northern bald ibis **7**:188
Northern black-and-white ruffed lemurs **6**:256–258*T*
northern bottlenose whale **5**:77*T*
northern brown howler **6**:256–258*T*
Northern elephant seal **5**:87*T*, **5**:88
Northern Europe (vegetation surveys) **3**:341–344*T*
northern gannets **3**:256
northern gray gibbon **6**:256–258*T*
northern gray slender loris **6**:256–258*T*
northern leopard frog **4**:156
northern mitred surili **6**:256–258*T*
northern muriqui **6**:256–258*T*
northern orioles **4**:174
northern pike **1**:12*F*
northern purple-faced langur **6**:256–258*T*
Northern sea otter **5**:81*T*, **5**:84–86
northern spotted owl **2**:225–226, **2**:243–244, **2**:264, **3**:314–315, **3**:315*F*
northern wheatears **7**:185
northern white-cheeked crested gibbon **6**:256–258*T*
North Island brown kiwi **4**:194*T*
North Island Hector's dolphin **5**:77*T*
North Korea
  animal species **1**:261*T*
  botanical gardens **6**:519
  ecosystems
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  see also Korea
North Pacific Subtropical Gyre (NPSG) **5**:45, **5**:48*F*, **5**:49*F*, **5**:52–53
North San Francisco Bay Estuary **5**:28*T*
North Sea **5**:688
North Selangor swamp forest **1**:276*T*
Norway
  boreal forests **1**:626, **7**:212–213, **7**:212*T*
  botanical gardens **6**:519
  Convention on Biological Diversity (1930) **2**:314
  nitrogen deposition **5**:529*F*
  North East Svalbard Nature Reserve **5**:444

Norway (*continued*)
vegetation surveys **3**:341–344*T*
whaling industry **3**:454
Norway rats **1**:365–368, **4**:697, **6**:240–241
Norway spruce **1**:628–629, **7**:212–213
noscapine **5**:712*T*
Nostoc **2**:611*F*
notable mass extinctions **5:167–177**
background information **5**:167
great mass extinctions
end-Cretaceous **5**:168*F*, **5**:169*F*, **5**:173–175, **5**:174*F*
Late Devonian **5**:168*F*, **5**:169–170, **5**:169*F*
Late Ordovician **5**:167–169, **5**:168*F*, **5**:169*F*
Late Permian **5**:168*F*, **5**:169*F*, **5**:170–172
Late Triassic **5**:168*F*, **5**:169*F*, **5**:172–173
research challenges **5**:167–169
lesser mass extinctions
Cenozoic **5**:176
Mesozoic **5**:169*F*, **5**:176
Neoproterozoic **5**:168*F*, **5**:175–176
Paleozoic **5**:168*F*, **5**:175–176
significance **5**:175–176
Nothofagaceae **5**:621–625
Nothofagus dombeyi **4**:217–219, **4**:218*F*
Nothofagus forests
characteristics **1**:291–292, **7**:118–119
habitat–bird species diversity relationship **1**:610–611, **1**:610*F*
phylogenetic analyses **6**:22
South America **3**:117, **4**:105–107
Southern Hemisphere ecosystems **6**:615
species distribution modeling (SDM) **6**:692, **6**:692*F*
Nothofagus pumilio **4**:215*F*
Nothroiherium **4**:727*T*
Nothropus **4**:727*T*
Nothrotheriops **4**:724–725*T*
Nothybidae **3**:489–491*T*
Notiobia **1**:249
Notiomastodon **4**:727*T*, **5**:224*T*
Notiotitanops **5**:224*T*
Notochthamalus scabrosus **5**:56–57
Notodontidae **5**:387–388*T*, **5**:396
Notonectidae **7**:389–391
Notophthalmus viridescens **1**:172
Notoryctemorphia **4**:682*F*, **4**:687*T*, **4**:688
Notoryctes caurinus **4**:688
Notoryctes typhlops **4**:688
Notosolenus **6**:293
Notostraca
extinction vulnerability **2**:416–417*T*
number of species **3**:182*T*
Nototherium **4**:727–728
Notoungulata **2**:583*F*, **4**:727*T*, **5**:225*T*
Notropis lutrensis **4**:364
Noxious Weed Act (1974) **3**:715
nubian shrike **5**:474
Nucella **3**:125
Nucifraga columbiana **2**:571
nuclear energy
biodiversity impacts **3**:259, **3**:264*T*
carbon dioxide (CO_2) emissions **3**:265*T*
land requirements **3**:265*T*
world energy reserves **3**:265*T*
world energy use **3**:251*F*, **3**:252*T*
nuclear fission **3**:259, **3**:264*T*, **3**:265*T*
nuclear magnetic resonance (NMR) spectroscopy **5**:708
nucleic acids **3:568–580**
gene scrambling
basic principles **5**:568
ciliate conjugation **5**:568, **5**:569*F*
genome rearrangements **5**:569–570, **5**:571*F*
internal eliminated sequences (IESs) **5**:569, **5**:571*F*
macronuclear destined sequences (MDSs) **5**:569, **5**:570*F*, **5**:571*F*
macronuclear formation **5**:568–569, **5**:570*F*
unscrambling process **5**:569–570, **5**:570*F*, **5**:571*F*
measurement and analysis **5**:259–260, **5**:262–264, **5**:262*F*, **5**:263*F*, **5**:264*T*
nucleotide sequences **5**:568, **6**:183
origin of life theories **5**:629, **5**:633–636
RNA editing
acceptor stems **5**:576–577
A–>I deamination transitions **5**:577–578
anticodons **5**:576–577
apolipoprotein B **5**:577
basic principles **5**:570–572
C–>U and U–>C editing **5**:575–576, **5**:575*F*
guide RNAs (gRNAs) **5**:572, **5**:575*F*
kinetoplastids **5**:572, **5**:573–574*T*
mitochondrial tRNA **5**:576–577
myxomycetes **5**:572–575, **5**:573–574*T*, **5**:575*F*
paramyxoviruses **5**:573–574*T*, **5**:578
plant mitochondria **5**:573–574*T*, **5**:575–576, **5**:575*F*
types and distribution **5**:570–572, **5**:573–574*T*
U insertion **5**:572–575, **5**:575*F*
structural characteristics **3**:655–656, **3**:655*F*
nucleotide polymorphism **3**:664
nucleotides **3**:655–656, **3**:655*F*
Nuerachne **2**:376*F*
Nuku Pu'u **1**:29*F*
Nullarbor Plains **2**:468
numbat **4**:687, **4**:687*T*, **4**:701*T*
Numenius borealis **7**:182
Nummulites **5**:306
Nuozadu Nature Reserve **3**:374
Nuphar lutea **6**:115–116
nurseryfish **3**:467–471
Nusa Tengarra **1**:514*T*
nutcrackers **2**:571
nuthatches **4**:33–35, **4**:34*F*
nutmeg **2**:663
nutria **2**:590*T*, **7**:399
nutrient cycling
African savannas **1**:63–65
agroecosystems **1**:95–97, **1**:96*F*
dinitrogen-fixing (N_2-fixing) microorganisms **2**:618–619
domesticated ruminants **1**:705–706, **1**:705*F*
ecosystem function measurements
Amazonian forests **3**:81*T*
analytical and measurement approaches **3**:74–75*T*
basic concepts **3**:93
control mechanisms **3**:78
freshwater ecosystems **3**:68–70
herbivores **3**:77–78
India **3**:82*T*
land use impacts **3**:79–80
nutrient capture, retention, and release **3**:74–75*T*, **3**:78
open versus tight nutrient cycles **3**:78–79
rainfall effects **3**:79, **3**:81*F*
soil type effects **3**:79–80
vegetation structure **3**:79–80
energy flow
benefits **3**:237–238
food chains **3**:244–245
internal versus external supplies **3**:243–245
microbial loops **3**:244–245
organic detritus **3**:244–245
environmental impacts **1**:95–97, **1**:96*F*, **3**:282
herbivores **3**:77–78
marine sedimentary habitats **5**:111
riparian zones
high in-stream nutrient levels **6**:464–465
upland runoff **6**:463–464
nutrient-exchange mutualisms
chemosynthetic symbioses **5**:120–121, **5**:120*F*
hydrothermal vents **5**:120–121, **5**:120*F*
nutrient conversion
nitrogen fixation **5**:121
shipworms **5**:121
worm–whale interaction **5**:121, **5**:121*F*
photosynthetic symbioses **5**:118–120
nutrient management–water quality relationship **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
management criteria **1**:97, **1**:97*T*
nitrogen (N) **1**:96*F*, **1**:98
nutrient input and sources **1**:96*F*, **3**:575*F*, **3**:576–577, **3**:576*F*
phosphorus (P) **1**:96*F*, **1**:98–99
research background **1**:95–97
nutrient stress tolerance **3**:83–84, **3**:83*F*
nuts
East Asia **2**:663
Europe **2**:662–663
Mediterranean Basin **2**:662–663
North and South America **2**:663–664
seed protection **7**:26
Southeast Asia **2**:663
southwest Asia **2**:662–663
Nyctaginaceae **2**:374*T*, **5**:489

Nyctea scandiaca **1**:236
Nycteribiidae **3**:489–491*T*
Nyctibiidae **6**:308*T*
Nycticebus javanicus **6**:256–258*T*
Nycticorax nycticorax **5**:476–477
Nymphaeales **5**:618*F*, **5**:621*F*
Nymphalidae **5**:387–388*T*
  butterfly caterpillars **1**:652*F*
  geographic diversity
    latitudinal gradients **1**:657, **1**:657*T*, **1**:658*F*
    neotropics **1**:657–658, **1**:658*T*
    regional patterns **1**:658, **1**:659*T*
    spatiotemporal distributions **1**:660*F*
    species numbers **1**:657*T*
  taxonomy **1**:651*F*, **1**:652–653
  tropical rain forests **6**:307*T*
Nymphalinae **1**:652*F*, **1**:653
Nymphomyiidae **3**:489–491*T*
Nypa **5**:11*T*
Nypa fruticans **1**:277, **7**:272
Nyssa **7**:118
Nyssa aquatica **7**:397
Nyssa sylvatica **2**:26

O

oak trees
  food webs **7**:266*F*
  Holocene postglacial spread **2**:551–552*T*
  Mediterranean North Africa **1**:47
  North America **7**:118
  phenological shifts **5**:723
oak wilt disease **3**:633*T*, **4**:363
oarfish **3**:466, **3**:466*F*
oats **1**:462–463, **2**:383, **2**:662–663, **3**:32*F*, **3**:131–132
Oaxaca mud turtle **7**:229*T*
Oblea **6**:294–295
obligatory parasites **5**:657
Obtectomera **5**:386*F*, **5**:387–388*T*
oca **2**:663–664
Ocean Biogeographic Information System (OBIS) **2**:3, **2**:4*F*, **2**:5*F*, **2**:11*F*, **5**:45
Oceania
  botanical gardens **6**:510*T*
  deforestation **1**:679*T*, **2**:31*F*, **2**:448*T*, **2**:452*T*, **2**:454*T*
  Diptera **3**:488*T*
  ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
  ecosystems
    alpine life zone **1**:149*T*
    conservation status **3**:171–173*T*
  endangered mammals **3**:193–194
  first-contact extinctions **5**:365–366
  fish diversity **3**:473, **3**:474
  indigenous peoples **4**:269*T*, **4**:271*T*
  land-use trends **4**:559*F*
  late Quaternary extinction event (LQEE) **4**:730–731
  logging operations **2**:454*T*
  mammalian extinctions **4**:705*T*
  myriapods **5**:429, **5**:429*T*
  natural history museums and institutions **5**:406
  protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
  regional carbon stocks **1**:678*F*
  regional cultivars **3**:32*F*
  threatened birds **7**:181
oceanic islands **5:590–599**
  background information and characteristics **5**:590–591, **5**:591*F*
  biodiversity
    community characteristics
      disharmony **5**:591–592, **5**:592*F*
      endemism **5**:592, **5**:592*F*, **5**:592*T*
      relictualism **5**:592
    species characteristics
      innovative behaviors **5**:593
      reduced dispersal capabilities **5**:592–593
      reproductive changes **5**:593
      size changes **5**:593
  ecological dynamics
    colonization **5**:593
    community formation processes **5**:594–595, **5**:596*F*
    ecological changes **5**:593
    evolutionary changes **5**:593–594
    extinctions **5**:594, **5**:595–597
    species diversification **5**:594, **5**:595*F*
  future research areas
    realistic evolutionary trajectories **5**:597–599
    restoration **5**:599
    taxonomic knowledge **5**:597
    vacant niche space **5**:597
  species loss
    alien species invasion **5**:596–597, **5**:598*F*
    climate change **5**:597
    extinction rates **5**:595–597
    small population size **5**:597
Oceanites oceanicus **1**:601–602
oceans
  abyssal plains **5**:108*T*
  acidification **5**:33–34
  biodiversity **3**:443
  biological pumps **1**:559–560, **1**:560*F*
  carbon cycle **1**:676
  chemical contamination **3**:282
  conservation status **3**:171–173*T*
  continental shelf **5**:108*T*
  continental slope **5**:108*T*
  crustaceans **2**:397–399
  deep ocean environments **5**:108*T*
  depth ranges **5**:107, **5**:107*F*
  dinitrogen-fixing (N_2-fixing) microorganisms **2:608–625**
    basic concepts **2**:608–610, **2**:609*F*
    biodiversity
      dinitrogenase/dinitrogenase reductase **2**:610–611
      general discussion **2**:610–611
      metabolic diversity **2**:614–615
      nif operon **2**:610, **2**:611*F*, **2**:620*F*, **2**:622*F*
    phenotypic and genotypic diversity **2**:611–614, **2**:612*T*
    phylogenetic relationships **2**:611*F*
    physiological diversity **2**:614–615
  biogeochemical cycles **1**:560–561
  emerging patterns
    changing environmental conditions **2**:621–623
    environmental modifications **2**:619–621
    spatiotemporal variability **2**:619–621, **2**:620*F*, **2**:622*F*
  functional role
    biological carbon (C) pump **2**:615–618, **2**:616*F*
    nutrient cycle linkages **2**:618–619
    quantification methods **2**:615–618, **2**:617*F*
    Redfield ratio **2**:617–618, **2**:617*F*
  habitat characteristics **2**:609*F*
  research summary **2**:623
ecosystems
  alien species invasion **4**:155–156
  characteristics
    basic organization **5**:587*F*
    distinguishing characteristics **5**:587–588
    shared characteristics **5**:586–587
  global change impacts **5**:588–589
  historical research **5**:582
  human impacts
    ocean toxification **4**:155
    onshore ocean siltation **4**:155
  marine sedimentary habitats **5**:108*T*
  natural functioning units
    biome concept **5**:583
    challenges **5**:582–583
    coastal upwelling **5**:584*T*, **5**:585*T*, **5**:586, **5**:586*T*
    Eastern Boundary Current **5**:584*T*, **5**:585*T*, **5**:586, **5**:586*T*
    ecosystem concept **5**:583–586, **5**:584*T*, **5**:585*T*, **5**:586*T*
    equatorial upwelling **5**:583–586, **5**:584*T*, **5**:585*T*, **5**:586*T*
    low-latitutde gyre **5**:583, **5**:584*T*, **5**:585*T*, **5**:586*T*
    Southern Ocean **5**:583, **5**:584*T*, **5**:585*T*, **5**:586*T*
    subarctic gyre **5**:584*T*, **5**:585*T*, **5**:586, **5**:586*T*
  onshore ocean siltation **4**:155
  toxification **4**:155
elemental distribution **1**:561
El Niño Southern Oscillation (ENSO) events
  nutrient level shifts **3**:160
  ocean current shifts **3**:159–160
fish species
  biodiversity issues **3**:443
  ocean fishing
    bycatch **3**:446, **3**:446, **3**:447
    commodity versus wildlife issue **3**:445
    economic losses **3**:446–447
    government subsidies **3**:446–447

oceans (*continued*)
industrialized fishing **3**:445
international fishing conflicts **3**:444–445
management problems **3**:444
overfishing **3**:443–444, **3**:445, **3**:445–446, **3**:446, **4**:156
socioeconomic concerns **3**:444
global species richness **3**:710
global warming effects **4**:229
human impact
alien species invasion **4**:155–156
onshore ocean siltation **4**:155
toxification **4**:155
hydrological cycle **1**:553–554, **1**:555*F*
Late Permian **5**:172
marine ecosystem services **5:451–468**
accounting assessment models
basic concepts **5**:455–457
revealed references methods **5**:456–457
stated preference methods **5**:457
American lobster, Atlantic herring, and Northeast multispecies groundfish fisheries **5**:464–465, **5**:464*F*
challenges **5**:455
classifications **1**:334*T*
coastal zone management (Belize) **5**:465, **5**:465*F*
conceptual framework
challenges **5**:455
coupled social–ecological systems model **5**:453–455, **5**:454*F*
integrated ecosystem assessments (IEAs) **5**:454–455
modeling approaches and uses **5**:453
coral reef protection and restoration **5**:465–466
cross-system services **3**:96–98
cultural ecosystem services **5**:451–452, **5**:452*T*
decision-support tools **5**:459–463, **5**:460–462*T*, **5**:463*F*
future outlook **5**:466
multiple ecosystem services **5**:459–463
production function models
basic concepts **5**:457–459
bioeconomic models **5**:457–459
cost–benefit analyses **5**:459–463
tradeoffs **5**:459–463
provisioning services **5**:451–452, **5**:452*T*
regulating services **5**:451–452, **5**:452*T*
single ecosystem services **5**:457–459
supporting services **5**:451–452, **5**:452*T*
valuation measures **5**:457
wind farms **5**:466
nutrient cycling **3**:244–245
paleotemperature assessments **4**:224–226, **4**:225*F*
pelagic ecosystems **5:683–691**
El Niño Southern Oscillation (ENSO) **5**:687–688, **5**:689*F*, **5**:690*F*
future outlook **5**:690–691
pelagic fish stocks **5**:688–690, **5**:690*F*
phytoplankton **5**:683–686, **5**:684*F*, **5**:686*F*
primary production **5**:683–686, **5**:685*F*, **5**:686*F*
Russell cycle **5**:690–691
zooplankton **5**:686–688, **5**:687*F*, **5**:688*F*, **5**:689*F*
pelagic environment
benthic biomass distribution **4**:380–381, **4**:380*F*
pelagic diversity patterns **4**:385, **4**:386*T*
physical characteristics **4**:379, **4**:380*T*
physical characteristics **4**:379–380
predation effects **6**:236
psychrophiles **6**:300
salt water plants **6**:77
solubility **1**:559–560, **1**:560*F*
sulfur cycle **1**:561–563
symbiotic relationships **5:116–126**
functional roles **5**:117–118
future challenges **5**:125
global distribution **5**:117*F*
historical research **5**:116–117
light-producing symbioses
bioluminescence **5**:121–123, **5**:122*F*
fish **5**:123
squid **5**:123–124
nutrient conversion
nitrogen fixation **5**:121
shipworms **5**:121
worm–whale interaction **5**:121, **5**:121*F*
nutrient-exchange mutualisms
chemosynthetic symbioses **5**:120–121, **5**:120*F*
hydrothermal vents **5**:120–121, **5**:120*F*
nutrient conversion **5**:121
photosynthetic symbioses **5**:118–120
protective compound production
characteristics **5**:124–125
sponges **5**:124–125, **5**:124*F*
structural characteristics **5**:116–117, **5**:117*F*
transmission modes **5**:116–117, **5**:117*F*
water storage **3**:560–562, **3**:561*T*
see also marine ecosystems
Oceans (documentary film) **2**:15
ocean sunfish **3**:467–471
Ochnaceae **5**:623*F*
Ochotona princeps **2**:85
Ochotonidae **4**:697
Ochradenus baccatus **5**:485–486
Ochrobium **6**:32
Ochroma spp. **2**:27–28
Ocnopus **4**:727*T*
Ocodoileus hemionus **1**:703, **2**:503–505
ocotillo cactus **7**:223
octocorals **2**:335
Octodontidae **4**:727*T*
octopus **5**:379–380, **5**:380*F*
octopus tree **2**:468
octylphenols **3**:581
Oculina patagonica **5**:258
ocypodid crabs **5**:16, **5**:16*F*
odd-nosed monkeys **6**:251*T*
Odematopus **3**:522*T*
Odhneriotrema **7**:450–451*F*
Odiniidae **3**:489–491*T*
odobenids **5**:67, **5**:74–75
Odobenus rosmarus **5**:90, **6**:236
Odocoileus **4**:723–725, **4**:724–725*T*
Odocoileus hemionus **5**:313
Odocoileus spp. **6**:332
Odocoileus virginianus **2**:279, **4**:494
Odocoileus virginianus clavium **3**:192
Odonata
lake and pond ecosystems **4**:371–373
latitudinal range shifts **4**:604*T*
major divisions and characteristics **4**:308–312, **4**:309*T*
Near East ecosystems **5**:460*T*, **5**:464*T*
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
wetland ecosystems **7**:389–391
odontoceti
characteristics **4**:693–694, **5**:66–67
taxonomy **5**:65*T*
Odontocheila cayennensis **1**:250*T*, **1**:251*F*
Odontocheila chrysis **1**:250*T*, **1**:251*F*
Odontocheila confusa **1**:250*T*, **1**:252*F*
Odontocheila luridipes **1**:250*T*, **1**:251*F*
Odontocheila margineguttata **1**:250*T*, **1**:251*F*
Odontocheila nigrotarsalis **1**:250*T*, **1**:251*F*
Odontomachus simillimus **3**:222
Odontotermes **4**:409*T*
Odontotermes formosanus **4**:418–419
Odontotermes pauperus **4**:419
Odontotermes smeathmani **4**:419
Odum, E. **2**:707, **3**:93, **3**:588, **5**:581–582
Odum, H. W. **3**:59, **5**:581–582
Oecanthinae **3**:730
Oecophoridae **5**:387–388*T*
Oedimeridae **1**:327*T*
Oedogonium **3**:358
Oenanthe oenanthe **7**:185
Oenocarpus bacaba **4**:286–287, **4**:288*T*
Oenocarpus bataua **4**:288*T*
Oenocarpus disticus **4**:288*T*
Oenocarpus mapora **4**:288*T*
Oenocarpus minor **4**:288*T*
Oenothera speciosa **1**:707
Oestridae **3**:489–491*T*
Oestroidea **3**:489–491*T*
O'Hara, Patrick **3**:274*F*
Ohio River **7**:148*F*, **7**:149*F*
Ohrid (Lake) **4**:463*T*, **4**:464*T*
oil
carbon dioxide (CO_2) emissions **3**:265*T*
energy use impacts **3**:253–256, **3**:255*T*, **3**:264*T*
land requirements **3**:265*T*
oil spills **3**:255, **3**:255*T*
world energy reserves **3**:265*T*
world energy use **3**:251*F*, **3**:252*T*
oilbird **1**:608, **2**:243–244
oil-eating bacteria **1**:592
oilfish **3**:468
oil palm
characteristics **5**:600
historical perspective **5**:600–601
oil-palm plantations **5:600–612**
biodiversity enhancements and benefits **5**:608–609
biodiversity impacts

Asian elephants **5**:606
beneficial wildlife **5**:606–607
charismatic species **5**:605–606
orangutans **5**:605–606
research challenges **5**:607
species diversity **1**:517, **5**:604–605
tigers **5**:606
wildlife habitat value **1**:540–541, **5**:603–604
controversies **5**:600
environmental impacts
deforestation **5**:607, **6**:320
palm oil production effects **5**:607–608
potential benefits **5**:608
wildlife habitats **1**:540–541
historical background **5**:601–602
production and harvests **1**:543–547, **5**:602–603
public perception **5**:603
summary discussion **5**:609
regional cultivars **3**:32*F*
slash-and-burn systems **6**:554
Sub-Saharan Africa **2**:663
tropical rain forest conservation **6**:320
uses **5**:601, **5**:602
oil-palm vulture **5**:603
oils
East Asia **2**:663
Europe **2**:662–663
Mediterranean Basin **2**:662–663
North and South America **2**:663–664
Southeast Asia **2**:663
southwest Asia **2**:662–663
Sub-Saharan Africa **2**:663
oil seed turnips **2**:662–663
oil spills **2**:14, **2**:14*F*
oil wells and pipelines **2**:503–505, **2**:504–505*F*, **4**:145
Okapia johnstoni **1**:49, **4**:696
okapis **4**:696
Okeechobee (Lake) **7**:372, **7**:372*F*, **7**:373*F*
Oksanen–Fretwell theory **3**:78, **3**:80*F*
Olacaceae **4**:71, **6**:315*T*, **6**:316*T*
old-growth forests **3**:314–315, **3**:315*F*, **7**:278, **7**:279*T*, **7**:281–282
old squaw ducks **1**:236
Old World cotton **2**:663
old world cuckoos **5**:501–502, **5**:502*T*
Old World vultures **1**:617
Olea europaea **5**:218, **5**:481–482
Olea europea **1**:47, **2**:662–663
Olentangy River Wetland Research Park **7**:379, **7**:380*F*
Oliarus polyphemus **7**:60
Oligocene
boundary age **5**:1*T*
brain size evolution **4**:698
C_4 plants **2**:376*F*
crustaceans **2**:406
Eocene–Oligocene boundary **5**:8
global change indicators **4**:225*F*
mammals **2**:583*F*
mass extinctions **5**:176
megaherbivores **5**:224*T*
oxygen-18 ($\delta^{18}O$) measurements **1**:557*F*
terrestrial plant communities
abundance patterns **6**:74*F*
Australia **1**:291–292
vegetation–climate interactions **2**:683
Oligochaeta
adaptability **4**:393
Antarctica **1**:184*T*
characteristics **4**:394*F*
classifications **4**:381*T*, **7**:408
endangered species **3**:225
lake and pond ecosystems **4**:371–373, **4**:374*F*
Near East ecosystems **5**:470*T*
number of species **4**:371*T*, **4**:384*T*, **5**:669–670*T*
sampling methods **4**:381*T*
oligomesotrophic level **3**:563*T*
oligonucleotide arrays **5**:264*T*, **5**:267
Oligoryzomys fulvescens **1**:367–368
Oligoryzomys longicaudatus **1**:367
oligosaccharides **6**:113*T*
oligotrich **6**:295–296, **6**:296*F*
oligotrophication **3**:367–369, **3**:563*T*
olive baboons **6**:571–573, **7**:262
olives **2**:662–663, **3**:32*F*
olive sunbird **3**:144–145
olive trees **5**:218, **5**:481–482, **7**:25
Ollala's titi **6**:256–258*T*
Olneya tesota **2**:468–469, **6**:455
Olson, Jerry S. **3**:242–243
Oman
animal species **1**:261*T*
botanical gardens **6**:519
ecosystems
mangroves **1**:275*T*
threatened species **1**:278*T*
vegetation types and coverage percentage **1**:264–265*T*
plant species **1**:259*T*
Ommatoiulus sabulosus **5**:428
Ommatophoca rossii **5**:90
omnivory **3**:498*F*, **3**:500*F*, **3**:502–503
onagers **5**:472
Onchocerca volvulus **5**:664–665, **7**:428*T*
Onchorhynchus rhodurus rhodurus **4**:465
Oncidium **6**:316*T*
oncomiracidium **7**:463
Oncopodoidea **1**:210–211
Oncorhynchus **5**:314–315, **6**:466
Oncorhynchus amago **6**:522
Oncorhynchus clarkii **6**:522
Oncorhynchus gorbuscha **3**:256, **6**:236, **6**:522
Oncorhynchus keta **6**:522
Oncorhynchus kisutch **6**:522
Oncorhynchus mykiss **1**:671, **2**:264, **3**:123*T*, **4**:362, **6**:522
Oncorhynchus nerka **3**:260, **3**:465*F*, **5**:314–315, **6**:522
Oncorhynchus tshawytscha **2**:7–8, **2**:10*F*, **3**:260, **4**:518, **6**:522
Oncorhyncus clarki **4**:363–364
Oncorhyncus nerka **4**:363–364
Ondatra zibethicus **2**:551–552*T*, **2**:554–555, **2**:554*F*, **4**:365–366, **7**:390
on-farm conservation **3**:33, **4**:316*F*, **4**:320, **4**:321*T*
onions
intercropping patterns **2**:383
native crops **2**:662–663
regional cultivars **3**:32*F*
Online Consortium for Library Cataloging (OCLC) **1**:440–441
Onohippidium **4**:727*T*
Onuphidae **7**:417*F*
Onychiurus armatus **7**:55*F*
Onychodiaptomus sanguineous **2**:718–719
Onychophora
arthropod radiation ecology **2**:399
endangered species **3**:225, **3**:226*F*
fossil record **3**:684*F*
global diversity **4**:385*T*
marine ecosystems **5**:50*T*
number of species **5**:669–670*T*
phylogenetic relationships **7**:434*F*
oomycetes **6**:296–297
Oomycota **3**:630, **3**:630*F*
Oophilia amblystomatis **2**:641–642
opahs **3**:466
Opalinata **5**:669–670*T*
opalinids **6**:296–297
Oparin, Alexander **5**:630
open access **6**:274, **6**:274*T*
Open Biomedical Ontology (OBO) **1**:444
open-country myriapods **5**:430
open forests **1**:292–297, **1**:294*T*, **6**:325–327, **6**:326*F*, **6**:326*T*
Open Geospatial Consortium (OGC) **1**:443–444
open ocean/deep sea
benthic ecosystems **5**:53
pelagic ecosystems **5**:49*F*, **5**:52–53
Operation Flood **1**:643
Operophtera brumata **5**:722, **6**:106–107
Opetiidae **3**:489–491*T*
Opheliidae **7**:417*F*
Ophiochla **2**:376*F*
Ophiocordyceps sinensis **3**:638
Ophiodon elongatus **3**:468
Ophioglossaceae **6**:59
Ophiostoma ulmi **3**:628–629, **3**:633, **3**:633*T*
Ophiotermes **4**:409*T*
Ophiuroidea **4**:381*T*, **5**:669–670*T*
Ophryotrocha **7**:330
Ophyiulus pilosus **5**:428
Opiliaceae **4**:71
Opilioacariformes **1**:216
Opiliones
Amazon ecosystems **1**:246*T*
characteristics **1**:209–210, **1**:209*F*
ecological characteristics **1**:210
genera and species **1**:202*T*
major lineages
Cyphophthalmi **1**:210
Dyspnoi **1**:210
Eupnoi **1**:210
Laniatores **1**:210–211
phylogenetic relationships **1**:205*F*
phylogeny and taxonomy **1**:210
reproduction and growth **1**:210
Opisthacanthus cayaporum **1**:208
Opisthocheiron canayerensis **5**:423*F*
Opisthocomus hoazin **1**:604, **1**:608

Opisthopatus cinctipes **3**:224
Opisthorchiodea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Opisthorchis sinensis **5**:660–661, **5**:671*T*
opisthotheles **1**:206
opium **7**:26
opium poppy **1**:359, **6**:57
Opius longicaudatus **6**:668–669
Opius oophilus **6**:668–669
Opius vandenboschi **6**:668–669
Opomyzidae **3**:489–491*T*
Opomyzoidea **3**:489–491*T*
Opomyzoinea **3**:489–491*T*
opossums **2**:584*F*, **2**:590*T*, **4**:686–687, **4**:687*T*, **5**:576–577
opossum shrimp **4**:363–364, **4**:371*T*, **4**:378
Opuntia **3**:130*T*, **4**:350–351, **4**:351*F*, **6**:171
Opuntia aurantiaca **4**:350–351, **4**:351
Opuntia coricolla **4**:361
Opuntia ficus-indica **1**:81, **2**:663–664, **3**:703, **4**:350–351, **4**:351*F*
Opuntia stricta **4**:350–351, **4**:351
Opuntia vulgaris **4**:350–351
orange-fronted parakeet **4**:194*T*
orange roughy **3**:451, **3**:468, **3**:482
oranges **3**:32*F*
orange-tufted sunbirds **5**:476
orangutans **4**:693, **5**:603, **5**:605–606, **6**:251*T*
Orbiculariae **1**:207
Orbignya martiana **4**:287–290
Orbiniidae **7**:417*F*
Orcaella brevirostris **5**:81*T*
orcas *see* killer whales
Orchidaceae **3**:129*T*, **6**:313–318, **6**:316*T*
orchid mycorrhizae **3**:636
orchids **2**:328–329, **3**:409*T*, **5**:551, **6**:47–48
Orcinus orca **4**:445–446, **4**:693–694, **5**:58–60, **5**:447
Orconectes australis australis **7**:54–55
ordinary least-squares (OLS) regression analysis **6**:642–643
Ordovician period
  arthropod radiation ecology **2**:400
  crustaceans **2**:403
  fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
  great mass extinctions **5**:167–169, **5**:168*F*, **5**:169*F*
  marine diversity **5**:168*F*, **5**:169*F*
  marine invertebrates **4**:390, **4**:390*F*
  mass extinctions
    event characteristics **1**:493
    extinction intensities **1**:494*T*, **5**:162*F*
    marine animals **1**:413*F*, **1**:493*F*, **5**:158
    notable mass extinctions **5**:161*F*
    number of families **3**:416*T*, **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
  molluscs **5**:373–374, **5**:374*F*
  originations **1**:495, **1**:495*F*, **3**:425*F*
  periodicity **3**:424*F*
Oreamnos **4**:724–725*T*
Oregon
  nonvascular epiphytes **3**:521*T*
  species richness maps **6**:707–710, **6**:709*F*
  timber industry **7**:314–315
Oregon silverspot butterfly **4**:564
Oregon spotted frog **4**:564
Oreithales **4**:101
Orellia falcata **5**:469–470
Oreobolus obtusangulus **4**:107
oreodonts **4**:701*T*
Oreomyrrhis **4**:194*T*
Oreonax flavicauda **6**:256–258*T*
Oreopanax **3**:522*T*
Oreopolus glacialis **4**:105–106, **4**:106, **4**:106–107
Oreoscopus gutturalis **2**:75*F*
Oreotragus oreotragus **1**:50, **6**:572–573
Orestias **4**:463–464
Organ Cave (West Virginia, USA) **7**:58, **7**:60, **7**:60*T*
organelles **3**:659–660
organic farming **2**:544
organism-level diversity **2:638–647**
  characteristics
    Darwinian theory **2**:638
    energy flow systems **2**:638–639
    functional internal organization **2**:640
    hierarchical organization **2**:640
    information systems **2**:639–640
    Lamarckian theory **2**:638
  evolutionary development
    eukaryotes **2**:641
    general discussion **2**:640–647
    prokaryotes **2**:640–641
  research summary **2**:647
Organization for Economic Cooperation and Development (OECD) **3**:644
organochlorines **3**:581, **5**:693, **5**:699*T*, **5**:700*T*, **6**:155
organometallics **5**:695
organophosphates
  fungicides **5**:695
  insecticides **5**:693–694, **5**:699*T*, **5**:700*T*
organotins **5**:694, **5**:695, **5**:700*T*
organ-pipe corals **2**:335, **3**:199
Orgel, Leslie **5**:631–632
Oribatida **1**:217, **6**:587–588
oribatid mites **4**:267, **6**:587–588
Oriental deciduous forest **1**:262–263*T*
oriental persimmons **2**:663
Origanum isthmicum **5**:485
Origanum jordanicum **5**:485
Origanum petraeum **5**:485
Origanum punonense **5**:485
Origanum ramonense **5**:485
origin of life theories **5:628–636**
  general discussion **5**:628
  historical perspective
    experimental research
      amino acid synthesis **5**:630–631
      prebiotic chemistry simulations **5**:630–631, **5**:631*F*
      self-replicating polynucleotides **5**:631–632, **5**:632*F*
    hypotheses and theories **5**:629–630
    living organisms **5**:628–629
    thermodynamic theories **5**:629
  hyperthermophiles **7**:165–166
  theoretical perspectives
    catalytic hypercycle theory **5**:633–636, **5**:634*F*, **5**:635*F*
    hypercycle–cellular organization transition **5**:635*F*, **5**:636
    informational aspects
      characteristics and origins **5**:633
      maximal information content **5**:633
      nucleic acids **5**:633–636
      stepwise information capacity expansion **5**:634*T*
    quantitative studies
      metabolism **5**:632–633
      mutability **5**:632–633
      quasispecies **5**:633
      self-reproduction **5**:632
    stages of emerging life **5**:635*F*, **5**:636
Origin of the Species (Darwin) **2**:421, **2**:423
orioles **4**:174
ornamental botanical gardens **6**:505
ornamental fish **3**:409*T*
ornamental plants **1**:75–76
ornate spider monkey **6**:256–258*T*
Ornatotholus **2**:519*T*, **2**:520*T*
Ornithigalum **1**:269
Ornithocercus **6**:30
Ornithomimus **2**:519*T*, **2**:520*T*
Ornithorhyncidae **4**:686
Orobanchaceae **4**:71, **4**:72*F*
Orobanche cernua **4**:76
Orobanche crenata **4**:76
Orobanche ramosa **4**:76
Orobothriurus crassimanus **1**:208
Orodromeus **2**:519*T*, **2**:520*T*
orogenesis **1**:554
Oró, Juan **5**:630–631
Orthezia insignis **4**:361
Orthocis linearis **7**:220*T*
Orthognathotermes **4**:409*T*
Orthomyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Orthonecctida **5**:669–670*T*
Orthonyx spaldingii **2**:75*F*
Orthonyx temminckii **4**:517
Orthoporus ornatus **5**:429–430
Orthoptera **3:722–736**
  Acridoidea
    Acrididae (true grasshoppers and locusts) **3**:735
    characteristics **3**:734
    Pamphagidae (earth hoppers of pamphagids) **3**:734–735
    Romaleidae (lubber grasshoppers) **3**:734*F*, **3**:735
  characteristics **3**:722–723
  Eumastacoidea
    characteristics **3**:732–733, **3**:733*F*
    Eumastacidae (monkey grasshoppers) **3**:733, **3**:733*F*
    Proscopiidae (false walking sticks or proscopids) **3**:733
  food and feeding habits **3**:724–725
  Grylloidea
    characteristics **3**:728, **3**:728*F*
    Gryllidae
      Brachytrupinae **3**:729–730
      characteristics **3**:729
      Gryllinae **3**:729
      Oecanthinae **3**:730
      Phalangopsinae **3**:730

Gryllotalpidae **3**:728–729
Mogoplistidae **3**:730
Myrmecophilidae **3**:730
Hagloidea **3**:727–728, **3**:728*F*
major divisions and characteristics **4**:308–312, **4**:309*T*
number of species **4**:392*T*
paleontology **3**:725
phylogeny and taxonomy **3**:725
Pneumoroidea **3**:733, **3**:733*F*
Pyrgomorphoidea (gaudy grasshoppers) **3**:735, **3**:735*F*
reproduction and growth **3**:724
Rhaphidophoroidea **3**:732
Schizodactyloidea **3**:730, **3**:730*F*
sound production **3**:723–724
Stenopelmatoidea
Anostostomatidae (wetas) **3**:731, **3**:731*F*
characteristics **3**:730–731
Cooloolidae (Cooloola Monsters) **3**:731
Gryllacrididae (raspy crickets and leaf-rolling crickets) **3**:732
Stenopelmatidae (Jerusalem Crickets) **3**:731–732
Tanaoceroidea **3**:733–734
Tetrigoidea (pygmy grasshoppers and grouse locusts) **3**:736, **3**:736*F*
Tettigonioidea (katydids)
characteristics **3**:722–723, **3**:725–726, **3**:725*F*
food and feeding habits **3**:724–725
paleontology **3**:725
phylogeny and taxonomy **3**:725
reproduction and growth **3**:724
sound production **3**:723–724
Tettigoniidae
Acridoxeninae **3**:727
Bradyporinae **3**:726
characteristics **3**:726
Conocephalinae **3**:727
Ephippigerinae **3**:727
Phaneropterinae **3**:726–727
Phasmodinae **3**:727
Pseudophyllinae **3**:727
Tettigoniinae **3**:727
Zaprochilinae **3**:727
Tridactyloidea (false mole crickets and sandgropers) **3**:736, **3**:736*F*
Trigonopterygoidea (leaf grasshoppers and wingless desert grasshoppers) **3**:735–736
Orussidae **4**:177–178, **5**:674
Orycteropus afer **4**:695
Orycteropus cafer **1**:608
Oryctes boas **5**:603–604
Oryctolagus cuniculus
adaptation rates **2**:140
biotic relationships **1**:710
diversity loss **1**:621, **3**:193
endangered species **3**:211
grasslands and savannas **6**:333
Holocene postglacial spread **2**:551–552*T*
introduced species **1**:77, **4**:361, **7**:262–263
oryx **2**:472–474, **4**:696, **5**:235, **5**:472
Oryx dammah **1**:53
Oryx gazella **1**:53, **1**:487–488, **5**:235
Oryx leucoryx **2**:243–244, **3**:196, **5**:472
Oryzaephilus surinamensis **6**:670
Oryza glaberrima **1**:88, **2**:663
Oryza longistaminata **1**:53–54, **7**:153–154, **7**:387
Oryza rufipugon **2**:636*F*
Oryza sativa **1**:88, **2**:636*F*, **2**:663, **3**:128
Oryzias latipes **6**:356
Oryzomys longicaudatus **2**:587*F*
Oryzomys palustris **1**:365–368
Osborn, Fairfield **2**:285, **3**:99
Osbornia **5**:11*T*
Osedax frankpressi **5**:121, **5**:121*F*
Osmeriformes **3**:464–465, **6**:522–523
Osmia lignaria **1**:420*F*
Osmoderma eremita **3**:225
osprey **6**:479–480
ospreys **3**:406, **7**:390
ostariophysans **3**:463–464
Osteichthyes
characteristics **6**:522–523, **7**:338
classifications **3**:460–461, **6**:522–523
fossil record **3**:681*F*, **3**:685*F*
osteoglossimorphs **3**:463, **3**:463*F*
Osteolaemus tetraspis **6**:401
Ostertagia ostertagi **5**:672
Ostracoda/ostracods
classifications and sampling methods **4**:381*T*
extinction vulnerability **2**:416–417*T*
fossil record **3**:681*F*, **3**:685*F*
lake and pond ecosystems **4**:371–373, **4**:374*F*
marine environments **4**:382*F*
Near East ecosystems **5**:460*T*
number of species **2**:401*T*, **3**:182*T*, **4**:371*T*, **4**:384*T*
plankton fraction **6**:25
ostrich
deoxyribonucleic acid (DNA) evolution **2**:56*F*
desert ecosystems **2**:472*F*
flight constraints **1**:606
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
introduced species **1**:624
Near East ecosystems **5**:472
risk status classification **1**:644*F*
size **1**:601
Ostrya virginiana **2**:26
Otago skink **4**:194*T*
otariids **5**:67
Otitidae **3**:489–491*T*
Otocephala **3**:463–464
Otocyon megaloitis **4**:693
Otodistomum veliporum **7**:450–451*F*
Otostigmus scabricauda **5**:425
Otostigmus tibialis **5**:425
otters
adipose tissue **7**:33–35
Census of Marine Life project **2**:12*F*
convergent evolution **4**:701*T*
Great American Biotic Interchange (GABI) **2**:584*F*
introduced species **4**:357
mangrove ecosystems **5**:14
number of species **4**:690*T*
river ecosystems **6**:479–480
species diversity **1**:685, **4**:690*T*, **4**:693
Otus scops **5**:474
'O'u **1**:29*F*
Ouachita River **7**:152–153, **7**:153
Oulema melanopus **2**:551–552*T*
Ourisia breviflora **4**:106–107
Our Plundered Planet (Osborn) **2**:285
outbreeding *see* inbreeding
Outer Indus Delta **1**:276*T*
overfishing
classifications **3**:445
coral reef ecosystems **2**:341, **5**:61
ecological footprint analysis (EFA) **2**:708
human impact **4**:156
marine ecosystems **3**:410, **5**:54
ocean fishing **3**:445–446
prevalence **3**:443–444
salmon **6**:527–528
vulnerability assessments **3**:446
overgrazing **1**:489, **2**:378–379, **3**:311*F*, **4**:560*F*
overgrowth **4**:359–360
overhunting **1**:405
Ovibos **4**:724–725*T*
Ovibos moschatus **4**:729–730, **6**:331, **6**:571–573
Ovis **4**:724–725*T*
Ovis ammon cyclocero **6**:334
Ovis ammon polii **3**:192
Ovis aries **2**:476*T*, **3**:127*T*
Ovis canadensis **3**:197, **5**:313, **6**:334
Ovis canadensis nelsoni **3**:193
Ovis orientalis **2**:476*T*
ovule conservation programs **4**:319–320, **4**:321*T*
Oweniidae **7**:417*F*
Owen, Richard **2**:422, **5**:405
owl limpets **6**:232
owls
boreal forests **1**:631–632
food webs **7**:266*F*, **7**:267*T*
introduced species **1**:79
logging industry–owl population controversy **3**:314–315, **3**:315*F*
Near East ecosystems **5**:474
oxadiazole **4**:90*T*
Oxalidaceae **5**:621–625, **5**:623*F*
Oxalis **5**:217, **5**:219
Oxalis compacta **4**:106
Oxalis erythrorhiza **4**:105–106
Oxalis loricata **4**:106–107
Oxalis tuberosa **2**:663–664
oxaloacetate (OAA) **2**:363–364, **2**:363*F*
oxathiins **5**:695
oxbow **7**:369*T*
oxen **4**:724–725*T*
oxidation-reduction (redox) reactions **1**:555–556, **1**:556*T*
Oxisols **4**:228, **6**:123*T*
oxlip **4**:574*F*
oxpecker birds **5**:227
Oxus cobra **7**:229*T*
Oxychlöe **4**:101
Oxycoccus **1**:277–278
oxyfluorofen **5**:696

oxygen (O)
  atmospheric concentrations **1**:282*T*
  biogeochemical cycles
    evolutionary development **1**:558*F*
    key reactions **1**:557–558
    oxygen-18 ($\delta^{18}O$) measurements **1**:556–557, **1**:557*F*
    partial pressure data **1**:559*F*
    sequestration and burial **1**:558–559
  ecosystem function measurements **3**:70–71
  elemental partitioning **1**:554*T*
  oxygen-18 ($\delta^{18}O$) **1**:556–557, **1**:557*F*, **4**:224–226
  relative abundance **1**:553*F*
  sequestration and burial **1**:558–559
oxygen (O_2) evolvers *see* chlorophyll production
oxyreductases **5**:269*T*
Oxyria digyna **1**:232, **1**:236
Oxythyrea noemi **3**:674*F*
Oxytricha **6**:295
Oxytricha nova **5**:569–570, **5**:571*F*
Oxytricha trifallax **5**:569–570, **5**:571*F*
Oxyura jamaicensis **4**:362, **7**:184
Oxyura leucocephala **4**:362, **7**:184
Oxyuroidea **7**:431–432
oyster mushrooms **3**:637–638, **3**:637*T*
oysters
  aquaculture **1**:190*T*
  Chesapeake Bay case study **3**:303–306, **3**:304*F*, **3**:305*F*, **3**:306*F*
  genetic diversity **2**:264
  introduced parasites–species decline relationships **2**:528*T*
  invasive species **5**:61
  oyster reefs **5**:452–453
  parasitic disease vectors **5**:672, **5**:672*F*
ozone (O_3)
  atmospheric concentrations **1**:282*T*
  biodiversity impacts **1**:144
  climate impact **1**:281–282
  dynamic global vegetation models (DGVMs) **2**:680
  emission occurrences and sources **1**:137*T*
  enhanced greenhouse effect **4**:18–19
  habitat alteration **4**:155
  lower atmosphere **1**:289–290
  nitrogen biogeochemical cycle **6**:152–154
  ozone depletion **1**:289, **3**:407, **7**:296–297
  stratospheric ozone
    concentration deviations **1**:289*F*
    current trends **1**:287–289, **1**:289*F*
    depletion causes and consequences **1**:289
    future outlook **1**:290
    occurrences and characteristics **1**:287–289
    projected trends **1**:289
    total ozone column **1**:287–289, **1**:288*F*
  tropospheric ozone precursors **6**:152–154
Ozotoceros bezoarticus **2**:587*F*, **3**:115
Ozotoceros spp. **6**:332–333

P

PABITRA (Pacific-Asia Biodiversity Transect) initiative **4**:406–407, **4**:406*F*
Pablakhali Wildlife Sanctuary **1**:276*T*
Pacala–Socolow concept **1**:683–684, **1**:683*F*
Pachyarmatherium **4**:724–725*T*
Pachycephala melanura dahli **2**:189–190, **2**:189*F*
Pachycephala pectoralis **2**:189–190, **2**:189*F*
Pachycephalosauria **2**:520*T*
Pachycephalosaurus **2**:518, **2**:519*T*, **2**:520*T*
Pachycladon exilis **4**:194*T*
Pachymerium ferrugineum **5**:428
Pachyneuridae **3**:489–491*T*
Pachypodium namaquanum **1**:487
Pachyrachis problematicus **6**:402
Pachyrhizus ahipa **3**:128
Pachyrrhizum tuberosum **4**:288*T*
Pachyschelus **1**:329
Pachytitan **5**:224*T*
Pacific bluefin tuna **2**:7–8, **2**:9*F*
Pacific Decadal Oscillation (PDO) **4**:213
Pacific golden chanterelle **3**:639–640
Pacific golden plover **5**:311–312
Pacific herring **3**:256
Pacific Islands
  birds **2**:189–190, **2**:189*F*
  ecosystem conservation **3**:171–173*T*
  endangered mammals **3**:193–194
  environmental degradation **6**:83
  first-contact extinctions **5**:365–366
  indigenous peoples **4**:269*T*, **4**:271*T*
  interspecific competition **2**:189–190, **2**:189*F*
  invasive species **4**:352–353, **4**:353*F*
  mammalian extinctions **4**:705*T*
  myriapods **5**:429, **5**:429*T*
  protected area systems **6**:491*T*, **6**:492*T*, **6**:493–494, **6**:493*F*, **6**:494*F*, **6**:494*T*
  regional cultivars **3**:32*F*
  tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*
Pacific Ocean **4**:386*T*
Pacific Ocean Shelf Tracking Project (POST) **2**:16, **2**:17*F*
Pacific rats **6**:240–241
Pacific salmon
  aquaculture **3**:448–449
  characteristics **6**:522
  evolutionary significant unit (ESU) criteria
    evolutionary lineage versus ecological context approach **1**:670–672, **1**:670*F*
    hatchery fish **1**:670–672, **1**:670*F*
    research background **1**:668–669
    research criteria **1**:669–670
  life cycle **1**:669–670, **1**:669*F*
  riparian zones **6**:466
Pacific treefrog **3**:165, **3**:696
Pacific yew tree **1**:594, **5**:709, **6**:57
paclitaxel **6**:135, **6**:135*F*
Pacourina edulis **1**:164
pacovi **4**:286–287
paddlefish **3**:462
Padua Botanic Garden **6**:506
Paedocypris progenetica **3**:464
Paeonia **5**:622*F*
Paepalanthus **6**:44*T*
Pagai macaque **6**:256–258*T*
Pagai pig-tailed langur **6**:256–258*T*
Pagophila eburnea **7**:186
Pagophilus groenlandicus **5**:90
Pagurus **3**:610*F*
painted lady butterfly **6**:689–690
painted turtles **3**:218, **7**:390
pain treatment drugs **1**:594
pak-choi **2**:663, **3**:131–132
pakihi **7**:369*T*
Pakistan
  animal species **1**:261*T*
  botanical gardens **6**:519
  deforestation **2**:449*T*
  ecosystems
    deserts **2**:469–470
    mangroves **1**:275*T*, **1**:277
    threatened species **1**:278*T*
    threatened wetlands **1**:276*T*
    vegetation types and coverage percentage **1**:264–265*T*
  education level–infant mortality relationship **6**:205*F*
  plant species **1**:259*T*
Pak Mun Dam **1**:512
Pak Phanang estuary **1**:276*T*
palaeoendemics **3**:228
Palaeolama **4**:724–725*T*, **4**:727*T*
Palaeoloxodon antiquus **5**:225*T*, **5**:231–233
Palaeomastodon **5**:224*T*
Palaeomastodontidae **5**:224*T*
Palaephatoidea **5**:387–388*T*
Palau **6**:519
Palearctic region
  desert ecosystems **2**:469–470
  Diptera **3**:488*T*
  fish diversity **3**:472
  indigenous peoples **4**:271*T*
  Lepidoptera **5**:389*T*
  protected area systems **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
Paleocene
  boundary age **5**:1*T*
  crustaceans **2**:406
  early mammals **3**:188–189
  faunal recovery **5**:7
  global change indicators **4**:225*F*
  mammals **2**:583*F*
  Paleocene–Eocene thermal maximum **5**:7–8
  terrestrial plant communities **6**:74*F*
Paleoclusia chevalieri **5**:624*F*
paleoecology **5:645–655**
  data collection and interpretation **5**:645–646
  environmental change indicators **4**:256–257, **4**:258*F*
  goals and objectives **5**:645–646
  Quaternary data **5**:645
  research methodologies
    age models **5**:652–653
    dating uncertainties **5**:652
    mapping intervals **5**:652–653
    scaling factors **5**:646–649, **5**:648*F*, **5**:649*F*

space–time variations **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
spatial characteristics **5**:648*F*, **5**:649–650, **5**:649*F*, **5**:651*F*
taxonomic and numerical characteristics **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
temporal characteristics **5**:650–652, **5**:651*F*
temporal resolution **5**:652
zoom lens perspective **5**:647*F*, **5**:648*F*, **5**:649*F*, **5**:651*F*, **5**:653–654
Paleogene **2**:583*F*, **5**:168*F*, **5**:169*F*
Paleontological Data Handling Code **5**:556
Paleophatoidea **5**:386, **5**:386*F*
paleoprecipitation **4**:225*F*, **4**:226
Paleopropithecus **5**:365
paleosols **4**:225, **4**:225*F*
Paleosuchus **6**:401
paleotemperature **4**:224–226
Paleozoic era
arthropod radiation ecology **2**:399–400
assemblage evolution **3**:611–612, **3**:612*F*
boundary age **5**:1*T*
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*
marine diversity **5**:168*F*, **5**:169*F*
marine invertebrates **4**:390, **4**:390*F*
mass extinctions **1**:413*F*, **3**:419*F*, **5**:157*F*, **5**:175–176
molluscs **5**:374*F*
Palestine viper **5**:477, **5**:477*F*
Paley, William **3**:394
Palicourea rigida **1**:165
Palila **1**:29*F*
Palintropus **2**:519*T*
Palinura **2**:416–417*T*
Pallopteridae **3**:489–491*T*
Palmae **3**:129*T*, **5**:11*T*, **6**:316*T*
Palmaria **2**:641–642
palmaris **3**:634*T*
palm civit cat **1**:517
palm doves **5**:474
Palmengarten **6**:508
Palmeria dolei **1**:29*F*
palm leafminer **5**:603–604
palms
modern extinctions **5**:368*T*
tropical forest ecosystems **7**:272
Palorchestes **4**:727–728
Palorchestes azcal **5**:364*F*
paloverde **2**:468–469, **2**:475*F*
Palpatores **1**:246*T*
Palpigradi **1**:202*T*, **1**:205*F*, **1**:212, **1**:246*T*
Pameridea roridulae **6**:45
Pamir–Tien Shan highlands **1**:262–263*T*
Pampas **3**:105–107, **3**:106*F*, **3**:111*F*
Pampas cat **2**:590*T*
Pampas deer **3**:115
Pampatherium **4**:727*T*
Pampean Massif **3**:105–107, **3**:106*F*
Pamphagidae (earth hoppers of pamphagids) **3**:734–735
Pamphilioidea **4**:178
Pamplemousses Botanic Gardens **6**:508
Panaeolus **3**:635*T*
Panama
botanical gardens **6**:519
deforestation **2**:30–31, **2**:31*F*
ecosystem services **7**:356*T*
edible plants **3**:130*T*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
species richness **2**:25*T*
tropical forests **6**:600–604*T*
tropical rain forests **6**:314–315, **6**:317*T*, **6**:318*T*
water funds **7**:356*T*
Panandris **5**:363
Panax ginseng **6**:137, **6**:138*F*
Panax quinquefolius **6**:137
Pandanaceae **6**:316*T*
pandans **1**:277
Panda oleosa **1**:49
pandas
conservation biology **2**:243–244
species diversity **1**:686
Pandinus imperator **1**:204, **1**:207–208
Pandion haliaetus **3**:406, **7**:390
Pangaea **5**:5–6
Pangasinin Wetlands **1**:276*T*
pangolin **4**:690*T*, **4**:697, **4**:701*T*, **5**:604
Panicum arizonicum **2**:373*T*
Panicum coloratum **1**:67
Panicum maximum **1**:83, **2**:379–381
Panicum miliaceum **2**:663
Panicum prion **2**:376*F*
Panicum turgidum **1**:69
Panicum virgatum **1**:532, **3**:261
Panochthus **4**:727*T*
Panonychus citri **1**:376
Panoplosaurus **2**:519*T*, **2**:520*T*
Pan paniscus **1**:49, **6**:256–258*T*
panspermia **5**:630
pantheistic traditions
ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
indigenous peoples **6**:384, **6**:384–385
traditional ecological knowledge **6**:385
traditional social protection
background information **6**:385–386
biological communities/habitat patches **6**:386
critical life history stages **6**:386
long-lived (K-selected) species **6**:386
stewardship **6**:386
Panthera **4**:724–725*T*
Panthera leo **1**:47, **3**:411, **4**:362, **4**:693, **6**:573, **7**:29*F*
Panthera onca **2**:243–244, **4**:693
Panthera pardus **1**:50, **1**:709, **3**:373, **4**:693, **5**:603, **6**:573
Panthera pardus nimr **5**:472
Panthera pardus tulliana **5**:472
Panthera tigris **1**:518*F*, **1**:709, **3**:190, **3**:373, **4**:693, **5**:14, **5**:603, **7**:229*T*
Panthera tigris sumatrae **1**:517
Panthera uncia **4**:693
Pantholops hodgsonii **3**:373
pantolestids **4**:701*T*
Pantophthalmidae **3**:489–491*T*
Pan troglodytes **1**:49, **2**:243–244, **2**:631*T*, **5**:603–604, **7**:229*T*
Pan troglodytes ellioti **6**:256–258*T*
Pan troglodytes schweinfurthii **6**:256–258*T*
Pan troglodytes troglodytes **6**:256–258*T*
Pan troglodytes verus **6**:256–258*T*
Panulirus interruptus **4**:516, **6**:233
Papahanaumokuakea Marine National Monument **5**:446
papain **5**:712*T*
papaverine **5**:712*T*
Papaver somniferum **1**:359, **2**:432, **5**:712*T*, **6**:57, **7**:26
papaya **2**:663–664
paperbark tree **4**:360
paper parks **6**:498
Paperveraceae **2**:432
Paphiopedilum druri **3**:208
Pa Phru **1**:276*T*
Papilio dardanus **1**:656, **6**:5
Papilionidae
geographic diversity
latitudinal gradients **1**:657, **1**:657*T*, **1**:658*F*
neotropics **1**:657–658, **1**:658*T*
regional patterns **1**:658, **1**:659*T*
species numbers **1**:657*T*
taxonomy **1**:650, **1**:651*F*
tropical rain forests **6**:307*T*
Papilioninae **1**:650, **1**:651*F*
Papilionoidea **5**:386*F*, **5**:387–388*T*
Papio anubis **6**:571–573, **7**:262
papo-de-peru plant **4**:281–285
Papua New Guinea
birds **2**:189–190, **2**:189*F*
botanical gardens **6**:507*T*
deforestation **1**:677*F*
endangered mammals **3**:194*T*
endemism **3**:232*T*
fish diversity **3**:473
indigenous peoples **4**:270, **4**:270*T*
interspecific competition **2**:189–190, **2**:189*F*
invasive species
aquatic ferns (*Salvinia molesta*) **4**:346–347
Mimosa pigra **4**:353–354
protected area systems **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
tropical rain forests **6**:315–317, **6**:317*T*
urbanization **4**:560*F*
vascular plant diversity **6**:82*T*
papyrus **2**:370–372, **7**:388–389
Parabathynella calmani **5**:470*T*
Parabroteas sarsi **1**:183
Parabuthus spp. **1**:208
Parabuthus transvaalicus **1**:208
Parabuthus villosus **1**:208
Paracanthopterygii **3**:466–467
Paraceratherium **5**:224*T*
Paraceros **4**:727*T*
Paracoccidioides brasiliensis **3**:634*T*
paracoccidioidomycosis **3**:634*T*
Paracornitermes **4**:409*T*
Paradeuterobaris novaguineae **7**:450–451*F*
paradise parrot **7**:183–184

Paradoxurus hermaphroditus **1**:517
Parafontaria laminata **5**:424–425
Paraglaridacris **7**:456–457*F*
Paragonimus **7**:450–451*F*
Paraguay
  botanical gardens **6**:519
  cropland conversion **4**:665*T*
  land cover classifications **6**:607*T*
  pastureland conversion **4**:665*T*
  savannas **6**:600–604*T*
  tropical forests **6**:606*F*, **6**:606*T*
Parahemiurus **7**:450–451*F*
Paraheptgyia **7**:345–346, **7**:347*F*
Parainvenusta englarorum **7**:442–443*F*
Paralabrax clathratus **4**:516
Paralacydoniidae **7**:417*F*
Paralemanea **3**:358
parallel diversification **1**:23–24
paralytic shellfish poisoning (PSP) toxins **1**:471*F*, **1**:472–473
Paramastus episomus **5**:465–467, **5**:469–470
Paramecium **2**:182–184, **2**:183*F*, **2**:183*T*, **2**:184*F*, **6**:295
Paramecium aurelia **2**:182–184, **2**:183*F*, **2**:184*F*, **6**:669–670
Paramecium bursaria **2**:184, **5**:306, **6**:669–670
Paramecium caudatum **1**:698*F*, **2**:182–184, **2**:183*F*, **2**:184*F*, **6**:669–670
Paramegatherium **4**:727*T*
Paramelemorphia **4**:682*F*
páramo ecosystem
  characteristics **4**:101–103, **7**:275–276
  geographic distribution **4**:96–97, **4**:97*F*
  northern Andean alpine cushion formations **4**:103
  northern Andean dwarf *Polylepis* forest **4**:103
  number of species **4**:102*T*
  physical characteristics and patterns **4**:97–98
  plant diversity
    endemism **4**:99–100, **4**:99*T*
    origins and diversification **4**:100–101
    species richness **4**:98–99
  representative species **4**:102*F*
  subpáramo zone **4**:102–103
  superpáramo zone **4**:103
  true páramo zone **4**:103
Paramphistomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Paramylodon **4**:724–725*T*
Paramyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
paramyxoviruses **5**:573–574*T*, **5**:578
Paraná Basin **3**:105–107, **3**:106*F*
Paraonidae **7**:417*F*
Parapanochthus **4**:727*T*
parapatric speciation **3**:144–145, **6**:621*F*, **6**:622*T*, **6**:625, **6**:635
Parapholis incurva **6**:455
paraphyly **2**:52–53, **7**:91–93, **7**:92*F*
Paraphysomonas **6**:293
parapox virus **2**:529–530, **2**:529*F*
Parapristina verticillata **4**:364
paraquat **5**:696, **5**:699*T*
Parasaurolophus **2**:519*T*, **2**:520*T*
Paraserianthes falcataria **5**:605
Parasesarma plicata **5**:15*F*
Parasitaxus **4**:73
parasites **5:656–673**
  adaptations
    aggregation **5**:659
    asexual reproduction **5**:659–660
    complexity reductions and increases **5**:657, **5**:658*F*, **5**:659*F*
    dispersal mechanisms **5**:658, **5**:660*F*
    flowering plants **5**:662
    hermaphroditism **5**:659–660
    host specificity **5**:660
    infection mechanisms **5**:658–659, **5**:660*T*
    life cycles **5**:660–661, **5**:662*F*, **5**:663*F*
    parthenogenesis **5**:659–660
    physiological adaptations **5**:661–662
    reproductive capacity **5**:657–658, **5**:659*T*
    sensory receptors **5**:659*F*
    site specificity **5**:660, **5**:661*F*
    size **5**:657
  agroecosystems **3**:11–12
  amphibian species decline **2**:527–529, **3**:696
  classifications **5**:657
  climate change effects **2**:103–104, **2**:110*T*
  community structure
    community ecology **5**:666–667
    general ecological models **5**:667
    infracommunities **5**:667
  definition **5**:656
  Diptera **3**:494
  disease regulation **1**:348–350
  diversity management **2**:537
  domesticated ruminants **1**:710–711
  ecological niche model
    infection prevalence **5**:665–666, **5**:666*F*
    microhabitat segregation **5**:665–666, **5**:667*F*
    niche dimensions **5**:665, **5**:665*T*
    niche saturation **5**:665, **5**:666*F*
    proximate and ultimate causation **5**:665–666, **5**:667*F*
  economic and hygienic importance
    crop production **5**:673
    fish **5**:672, **5**:672*F*
    human disease vectors **5**:670–672, **5**:671*T*
    livestock **5**:672
    pest control **5**:673
    plants **5**:672–673
  evolutionary development **2**:645
  forest canopies **3**:518–519, **3**:521
  global distribution **5**:668, **5**:669–670*T*
  hemiparasitism **4:70–78**
    characteristics and classifications **4**:70–71
    ecological advantages **4**:75
    evolutionary development **4**:73–74
    haustorium/haustoria **4**:74–75
    host specificity **4**:75
    mycorrhizal associations **4**:74
    pests
      crop production **4**:75–77
      dodders **4**:70, **4**:77
      mistletoes **4**:70, **4**:77
      research summary **4**:77–78
      root parasitic weeds **4**:76–77
    phylogenetic relationships **4**:71–73, **4**:72*F*
    physiological characteristics **4**:75
    plant–plant interactions **4**:74
    seed dispersal/germination **4**:74
  host–parasite interactions **3**:11–12
    cleaning symbiosis **5**:663–664
    host individual effects **5**:664–665
    immune and tissue reactions **5**:664
    population control **5**:664–665
  infectious diseases **2:523–538**
    amphibian species decline **3**:696
    anthropogenic drivers
      general discussion **2**:531
      global climate change **2**:531–532
      habitat degradation and pollution **2**:531
      habitat loss and fragmentation **2**:531
      worldwide commerce and travel **2**:532
    characteristics and properties **2**:523, **2**:524*T*
    disease causes **2**:523
    disease management
      captive breeding programs **2**:536–537
      disease threat assessments **2**:535
      intervention methods **2**:535–536
      population monitoring **2**:535
      threat classifications **2**:534–535
    epidemiology
      basic principles **2**:523–525
      macroparasitic disease **2**:524*T*, **2**:525–527, **2**:526*F*
      microparasitic disease **2**:523–525, **2**:525*F*
      SIR compartment models **2**:523–525, **2**:525*F*
    intervention methods
      antiparasitic drug administration **2**:536
      culling **2**:535–536
      disease management regimes **2**:535–536
      habitat management **2**:536
      quarantine **2**:536
      vaccination **2**:535–536
    parasite diversity management **2**:537
    positive roles
      biodiversity enhancement **2**:533
      ecosystem functionality **2**:534
      evolutionary diversification **2**:532–533
      host conservation efforts **2**:534
      host genetic diversity **2**:533–534
      interspecific competition **2**:533
    threat potential
      community-wide effects/ecosystem consequences **2**:530–531
      introduced parasites–species decline relationships **2**:527–529, **2**:528*T*, **2**:529*F*

invasive species–pathogen relationship **2**:529–530
intraspecific parasites **5**:668, **5**:670*F*
introduced species **4**:361–362
mites **1**:214–215
mollusc–human interactions **5**:381–382
natural history museum research programs **5**:415
nematodes **7**:431–432
number of species **5**:669–670*T*
origins and evolution
complex life cycles **5**:662–663
host–parasite coevolution **5**:663
parasitism **5**:662–663
virulence **5**:663
parasite–host coevolution
gene-for-gene systems **2**:143–144
sexual selection **2**:144
virulence–avirulence evolution **2**:144
population dynamics
ecological strategies **5**:668
population growth **5**:667–668
related phenomena
commensalism **5**:656
mutualism **5**:656
phoresy **5**:656
predation **5**:656–657
symbiosis **5**:656
species interactions **6**:715*T*, **6**:718–719
tropical rain forests **6**:315*T*
zoogeography
as biological markers **5**:670
latitudinal gradients
host ranges/host specificity **5**:668–670
reproductive strategies **5**:668
species richness **5**:668, **5**:671*F*
von Ihering method **5**:670
Parasitica **5**:674
parasitic wasps **4**:306–307, **4**:362
Parasitiformes **1**:216
parasitoids **5:674–682**
characteristics **4**:179–180, **5**:657, **5**:674
Diptera **3**:494
functional role **4**:183, **5**:682
host–parasitoid interactions
community ecology **5**:681–682
host acceptance **5**:676–678
host location strategies **5**:675–676
host resistance **5**:678
metapopulation theory **5**:242–243, **5**:248, **5**:248*F*
Nicholson–Bailey model **5**:679
oviposition strategies **5**:676–678
population dynamics **5**:679–681, **6**:179
importance **5**:682
Lepidoptera **5**:395–396
life histories **5**:674–675
metapopulation theory **5**:242–243, **5**:248, **5**:248*F*
microorganism associations **5**:678–679
plant defense systems **2**:429
toxin injections **5**:678
see also Hymenoptera
Paraspidodera **7**:421*F*, **7**:424*F*, **7**:425*F*, **7**:432–433, **7**:432*T*
Parasthrephia lepydophilla **2**:469
Parathera prostrata **1**:165
parathion **5**:700*T*
Paridae **4**:33–35, **4**:34*F*, **6**:308*T*
Parinari excelsa **1**:159–160, **7**:275
Parinari montana **4**:288*T*
Parkeriaceae **6**:59
Parkia filicoidea **1**:51
Parkinsonia aculeata **7**:274
Parkinsonia africana **1**:67
Parkinsonia microphylla **2**:468–469, **2**:475*F*
Parkinson's disease **1**:592
Parksosaurus **2**:519*T*
Park, Thomas **2**:184
Parnaiba Basin **3**:105–107, **3**:106*F*
Parnassia **5**:623*F*
Parnassinae **1**:650
Parnassius smintheus **4**:516–517
Paromomyidae **4**:701*T*
Parrhinotermes **4**:409*T*
parrotfish **3**:467–471, **6**:40
parrots **7**:181–182
parsimony **2**:33*F*, **2**:35–36, **2**:35*F*, **6**:17
parsley **2**:662–663
parsnip webworm **2**:139, **2**:143
parthenogenesis **5**:659–660
partial equilibrium (PE) models **4**:294–295
partial Mantel tests **4**:510–511
Partnuniella **4**:118
partridge **1**:644*F*
Partula **1**:33
Parulidae **4**:33–35, **4**:34*F*, **6**:308*T*
Paruroctonus boreus **1**:208
Parus **1**:601, **1**:616*T*
Parus ater **1**:611–612
Parus caeruleus **1**:611–612, **1**:620
Parus cristatus **1**:611–612
Parus major **1**:611–612, **1**:620, **5**:14, **5**:476, **5**:603–604, **5**:720
Parus montanus **1**:611–612
Parus palustris **1**:611–612
Parvicorbicula **6**:293
Parvoviridae **5**:130*F*
Paspalum **2**:376*F*, **4**:170
Paspalum fasciculatum **1**:164
Paspalum orbiculatum **1**:165
passenger pigeon **1**:621, **2**:281–282, **4**:156, **7**:182
Passer domesticus **1**:77–78, **1**:601–602, **1**:623–624, **2**:551–552*T*, **4**:365–366, **5**:476
Passerella iliaca **1**:605, **5**:312
Passeriformes
biodiversity generation **1**:432
taxonomic diversity **1**:607–608
tropical rain forests **6**:308*T*
Passerina cyanea **6**:702*F*
passerines
biodiversity generation **1**:432
deoxyribonucleic acid (DNA) evolution **2**:56*F*
as disease vectors **1**:362–365, **1**:364*F*, **1**:365*F*
ecological separations **1**:616*T*
Passer montanus **4**:259
Passiflora **3**:130*T*
Passifloraceae **3**:130*T*, **5**:623*F*, **6**:316*T*
Passiflora edulis **2**:663–664, **4**:289*T*
Passiflora nitida **4**:288*T*
Passiflora quadrangularis **4**:289*T*
passion fruit **2**:663–664
passive sampling hypothesis **6**:641
Pasteurella multocida **2**:531
Pasteuria penetrans **1**:312–313
Pasteur, Louis **5**:628
Pastinaca sativa **2**:139, **2**:143
pastoralism **4**:650–651
pastureland
dryland syndrome **2**:483*F*
energy crops **1**:542–543, **1**:542*F*
introduced species
nonindigenous plants **1**:76
nonindigenous vertebrates **1**:76–77, **1**:76*T*
land-use change impacts **1**:542–543, **1**:542*F*
land-use trends **4**:559*F*
pre-agriculture land cover **4**:558*F*
slash-and-burn systems **6**:560
termites **4**:432*T*
Patagonian Massif **3**:105–107, **3**:106*F*, **3**:111*F*
patch dynamics
landscape diversity **4**:479–480
living dead **4**:596–597, **4**:596*F*
patch–matrix–corridor model **4**:479, **4**:479*F*, **4**:492–493, **4**:492*F*
patchiness **5**:240–241, **7**:232–233
patch use theory **5**:676
Patella ferruginea **3**:201–202
Patelloida latistrigata **4**:334–335
Pathfinder International **2**:231–232
pathogens
biodiversity loss **1**:519–520, **4**:664
crop plants
fungi **3**:633, **3**:633*T*
organism spillover **3**:13–14
crop production
crop plant pathogens
Australia **1**:76*T*, **1**:81
Brazil **1**:76*T*, **1**:83
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81–82
United Kingdom **1**:76*T*, **1**:80
organism spillover **3**:13–14
disease regulation **1**:348–350
domesticated ruminants **1**:710–711
habitat–niche relationship **4**:45–46
invasive species–pathogen relationship **2**:529–530, **4**:156
Lepidoptera **5**:395–396
natural history museum research programs **5**:415
range ecology **6**:341
symbiotic microorganisms **6**:449–450, **6**:450*F*
Patosia **4**:101
patroendemics **3**:228
pattern diversity **6**:534
pau Brasil **6**:560
Paucituberculata **2**:586, **2**:587*F*, **4**:682*F*, **4**:687, **4**:687*T*
Paullinea cupana **4**:289*T*

Paullinia **5**:489
Paulownia elongata **2**:383
Pauropoda
  Amazon ecosystems **1**:246*T*
  characteristics **4**:392
  developmental modes **5**:426
  geographic distribution
    endemism **5**:428
    environmental characteristics **5**:427–428
    hotspots **5**:429
    islands **5**:429
    regionalization **5**:428–429
    speciation **5**:428
  habitats and adaptations
    burrowing myriapods **5**:430
    evolutionary development **5**:429–430
    open-country myriapods **5**:430
    special habitat myriapods **5**:430–431
  morphology **5**:424, **5**:424*F*
  number of species **5**:427*T*
pavement, desert **2**:462
Pavlovsky, Evgenii **4**:503
Pavo cristatus **1**:624
Pavo muticus **7**:229*T*
pax6 genes **7**:53–54, **7**:54*F*
pea aphid **2**:139
Peabody Museum (Yale) **5**:405–406
peaches **2**:383, **2**:663, **3**:32*F*
peach palm **4**:284–285, **4**:284*F*, **4**:287–290
peach-potato aphid **4**:298–299, **4**:299*F*, **4**:302*F*
peacock bass **6**:233–235
pea crab **3**:201–202
peafowl **1**:624
Peale, Charles Wilson **5**:405–406
peanuts **1**:529*T*, **2**:384*F*, **2**:663–664, **4**:287–290, **7**:26
peanut worms **4**:381*T*
pearlfish **3**:466–467
pearl millet **1**:132*T*, **1**:463–465, **2**:659, **2**:663, **3**:32*F*
pearlwort **1**:182
pearly-eyed thrasher **7**:184–185
pearly mussels **3**:177–178, **3**:177*F*, **3**:179, **3**:180*F*, **3**:185*F*
pears
  intercropping patterns **2**:383
  native crops **2**:662–663
Pearson, Karl **6**:2
peas
  intercropping patterns **2**:383
  native crops **2**:662–663
  regional cultivars **3**:32*F*
  root–soil interface **6**:113*T*
peatlands/peat bogs
  acidic environments **7**:395–397, **7**:396*F*
  Bois-des-Bel Peatland, Quebec, Canada (case study) **7**:375
  definitions and terminology **7**:369*T*
  global extent **7**:370*T*
  habitat loss and degradation **7**:371*T*
  restoration programs **7**:374–375
  Southeast Asia **1**:512–513
  tourism industry **7**:223
peat moss **6**:83
peccaries **2**:590*T*, **4**:690*T*, **4**:696, **4**:724–725*T*, **6**:332–333
Pectinariidae **7**:417*F*
Pectis papposa **2**:373*T*
Pedicellarum **3**:522*T*
Pedicularis **4**:71
Pediococcus halophilus **3**:638–639
pekarangan **7**:230–231
Peking cabbage **2**:383
Pelagibacter **5**:293–294
Pelagibacter ubique **5**:294
pelagic ecosystems **5:683–691**
  benthic biomass distribution **4**:380–381, **4**:380*F*
  ecosystem diversity **5**:49*F*, **5**:52–53
  El Niño Southern Oscillation (ENSO) **5**:687–688, **5**:689*F*, **5**:690*F*
  fish diversity **3**:475
  future outlook **5**:690–691
  marine ecosystems **5**:46, **5**:46*F*, **5**:47*F*
  pelagic diversity patterns **4**:385, **4**:386*T*
  pelagic fish stocks **5**:688–690, **5**:690*F*
  physical characteristics **4**:379, **4**:380*T*
  phytoplankton **5**:683–686, **5**:684*F*, **5**:686*F*
  primary production **5**:683–686, **5**:685*F*, **5**:686*F*
  Russell cycle **5**:690–691
  zooplankton **4**:374*F*, **5**:686–688, **5**:687*F*, **5**:688*F*, **5**:689*F*
pelagic fish stocks **3**:482, **5**:688–690, **5**:690*F*
Pelargonium **5**:219
Pelecanoides **1**:617
Pelecanus occidentalis **3**:256
Pelecorhynchidae **3**:489–491*T*
Peleng tarsier **6**:256–258*T*
Pelliciera **5**:11*T*
Pelliecieraceae **5**:11*T*
Pelobates syriacus **5**:473, **5**:476*F*
Pelomyxa **6**:291
Pelorovis **5**:231–233, **5**:462
Pelorovis antiquus **5**:364–365
Peltasperms **5**:618*F*
Peltophorum africanum **1**:67
Pemphigus betae **4**:173
Pemphis **5**:11*T*
Pemphis maritima **1**:55
Penaeaceae **5**:219–220
Pene galilaea **5**:463, **5**:464*F*
Pene sidoniensis **5**:463, **5**:464*F*, **5**:465–467
penguins
  evolutionary convergence **1**:617
  geographical variation **1**:613–614
  threatened species **7**:181–182
penicillin **3**:638
penicilliosis **3**:634*T*
Penicillium **5**:555–556, **5**:710–711
Penicillium brefeldianum **5**:555–556
Penicillium camembertii **2**:644, **3**:638–639
Penicillium chrysogenum **3**:638
Penicillium dodgei **5**:555–556
Penicillium griseofulvum **3**:638
Penicillium marneffei **3**:634*T*
Penicillium notatum **1**:359
Penicillium roqueforti **2**:644, **3**:638–639
Pennant's red colobus **6**:256–258*T*
Pennisetum ciliare **6**:93
Pennisetum glaucum **1**:463–465, **2**:663
Pennisetum purpureum **1**:67
Pennisetum setaceum **3**:208–209
Pennsylvanian period
  crustaceans
    see also Carboniferous period
  marine diversity **5**:168*F*, **5**:169*F*
  terrestrial plant communities
    abundance patterns **6**:74*F*
    evolutionary development **6**:65–67, **6**:66*F*
    extinction rates **6**:73*F*
    origination rates **6**:72*F*
  terrestrial plant evolution **6**:70
Penstemon deustus **4**:518
Penstemon pachyphyllus **4**:518
Penstemon rostriflorus **4**:518
Pentaceratops **2**:518, **2**:519*T*
pentachlorophenol **5**:693, **5**:696–697
pentachlorophenol (PCP) **5**:256, **5**:695
Pentaclethra **7**:272
Pentacomia cribrata **1**:250*T*, **1**:252*F*
Pentacomia egregia **1**:250*T*, **1**:252*F*
Pentacomia lacordairei **1**:250*T*, **1**:251*F*, **1**:252*F*
Pentacomia ventralis **1**:250*T*, **1**:251*F*
Pentapanax **3**:522*T*
Pentapetalum trifasciculandricus **5**:624*F*
pentastomids **2**:400, **5**:669–670*T*
Pentatomidae **7**:294
Pentatomomorpha **7**:294
Penthorum **5**:622*F*
pentoses **6**:113*T*
Pentoxylon **5**:618*F*
people–conservation relationship **2:221–229**, **2:230–237**
  benefits
    cost–benefits **2**:226
    fishing industry **2**:232
    forestry sector **2**:232
    grasslands **2**:232
    nature-based tourism **2**:232–233
    payment for ecosystem services (PES) **2**:226–227
    poverty reduction **2**:227–228
    protected area systems **2**:226
  conservation threats
    human impact measures and maps **2**:221–222
    reduction strategies **2**:222–223
  disadvantages **2**:233
  general discussion **2**:221
  goals and priorities
    general discussion **2**:223–224
    preferential treatment large-bodied vertebrates **2**:224
    shifting baselines **2**:223–224
    wilderness ideal **2**:224
  historical perspective **2**:230–231
  key organizational players **2**:231–232
  poverty indicators **2**:233
  poverty trends **2**:233–235, **2**:234*F*, **2**:234*T*
  research summary **2**:228, **2**:235
  victimization factors
    conservation refugees **2**:224–225
    general discussion **2**:224–225

negative impacts **2**:225–226
protected area systems **2**:224–225
unemployment considerations **2**:225–226
people–ecosystem relationship
autonomous people **6**:221–222, **6**:221*T*
biosphere people **6**:221*T*
conservation practices **6**:226
ecodevelopment programs **6**:226–227
ecological refugees **6**:221*T*, **6**:225–226
future challenges **6**:227
general characteristics **6**:220–221
land-use changes **6**:223–224, **6**:224*T*
resource ownership, sharing, and utilization **6**:223, **6**:224*T*
subjugated people **6**:221*T*, **6**:222–223
value appropriation **6**:224–225
Peperomia **6**:316*T*
peppered moth **2**:717
peppers **2**:383, **2**:663–664, **3**:32*F*, **4**:284–285, **4**:284*F*, **7**:26
peptides **6**:113*T*
Peracarida **2**:401*T*
Peraceras **5**:224*T*
Peramelemorphia **4**:687*T*, **4**:688
Perca fluviatilis **4**:463
perch **1**:10*T*, **1**:12*F*, **3**:467–471, **4**:463, **7**:390
perched aquifers **7**:50
Perciforms **3**:468
Percina tanasi **2**:223
Percomorpha **3**:467–471
Perdix perdix **3**:123, **4**:259
Père David's deer **3**:196
peregrine falcons **2**:245, **3**:192–193, **3**:406, **4**:258–259, **5**:474, **6**:445–446, **7**:188
perennial monocultures **1**:532
Perezia megalantha **4**:106–107
Perezia pilifera **4**:106–107
perfluorocarbons (PFCs) **1**:284–286
perhumid forests
high mountain ecosystems **7**:275
lowland ecosystems **7**:272–273
low mountain ecosystems **7**:275
Pericapritermes **4**:409*T*
peridinians **6**:25
Peridiniopsis berolinensis **6**:294
Peridinium **6**:29–30, **6**:294
Peridinium cinctum **6**:33*F*
periodic parasites **5**:657
peripatric speciation **6**:621*F*, **6**:622*T*, **6**:624, **6**:634–635
Peripatus acacioi **3**:224
Periplaneta americana **7**:431–432
Periploca aphylla **5**:485
Periscelididae **3**:489–491*T*
Perissodactyla
characteristics **4**:695
convergent evolution **4**:701*T*
desert ecosystems **2**:476*T*
dietary tolerance **5**:225*T*
late Quaternary extinction event (LQEE) **4**:727*T*
number of species **4**:690*T*
occurrences and characteristics **5**:224*T*
Paleocene–Eocene thermal maximum **5**:7–8
representative species **2**:587*F*
species diversity **2**:587, **4**:690*T*
tropical rain forests **6**:309*T*
Perissomatidae **3**:489–491*T*
periwinkles **2**:342–343, **3**:202*T*, **4**:10
Periyakarachchi Lagoon **1**:276*T*
Perkinsus marinus **2**:528*T*
permafrost **1**:274–275, **6**:299–300
permanent parasites **5**:657
Permelidae **4**:688
permethrin **5**:694
Permian period
arachnids **1**:204–205
arthropod radiation ecology **2**:400
boundary age **5**:1*T*
crustaceans **2**:404
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
global change indicators **4**:225*F*
great mass extinctions **5**:168*F*, **5**:169*F*, **5**:170–172
marine diversity **5**:168*F*, **5**:169*F*
mass extinctions
event characteristics **1**:493
extinction intensities **1**:494*T*, **5**:162*F*
magnitude measures **5**:158
marine animals **1**:413*F*, **1**:493*F*
notable mass extinctions **5**:161*F*
number of families **3**:416*T*, **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
originations **1**:495, **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
Permian–Triassic mass extinction **5**:4
temperate forests **7**:123*F*
terrestrial plant communities
abundance patterns **6**:74*F*
evolutionary development **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
permutation tail probability (PTP) test **2**:44
Pernettya pumila **4**:106–107
Perognathus alticola **7**:229*T*
Perognathus flavus **4**:43*F*
Perognathus longimembris **4**:43*F*
Perognathus penicillatus **4**:43*F*
Peromyscus leucopus **1**:362, **1**:365–368
Peromyscus maniculatus **1**:365–368, **1**:366*F*, **4**:43*F*, **4**:363
Peromyscus polionotus **6**:5
Peronospora parasitica **3**:630
Peroryctes broadbenti **4**:688
peroxidases **5**:269*T*
peroxyacetyl nitrate (PAN) **1**:286
peroxyl (HO_2) **1**:282*T*
Perrier's sifaka **6**:256–258*T*
Persea **1**:270–271
Persea americana **2**:663–664
Persea theobromifola **3**:208
persistent organic pollutants (POPs) **1**:137*T*
Peru
biodiversity-based agricultural strategies **1**:124
biodiversity richness **1**:501*F*
botanical gardens **6**:507*T*
butterflies **1**:658*T*
coral reef ecosystems **1**:500*T*
deforestation **1**:502*T*
ecosystem services **7**:356*T*
endangered mammals **3**:194*T*
endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
fishing/fisheries **6**:422*T*
indigenous peoples **4**:270, **4**:270*T*
institutional collections and capacity **1**:505*T*
land cover classifications **6**:607*T*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
plant species **4**:102*T*
primate diversity **6**:252*T*
rain forests **1**:159*T*
species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
Tamshiyacu-Tahuayo Communal Reserve **5**:445
threatened birds **7**:181
tropical forests **6**:600–604*T*, **7**:269–270, **7**:270*T*
tropical rain forests **6**:314–315, **6**:317*T*, **6**:318*T*
water funds **7**:356*T*
wheat production **1**:87–88
Perumytilus purpuratus **5**:56–57, **5**:58*F*, **5**:59*F*
pesticides/pest control **5:692–702**
animal damage control **6**:342
biodiversity-friendly farming **2**:542, **2**:544
biodiversity impacts **1:373–385**
agricultural intensification effects **1**:374–375
agricultural landscapes
characteristics **1**:380–381
habitat heterogeneity **1**:381–382, **1**:381*F*
natural habitat proximity **1**:381
noncrop habitats **1**:380–381
background information **1**:373–374
conservation biological control **1**:382
influencing factors **1**:374*F*
multiple trophic levels **1**:380
predator diversity
complementarity **1**:377–378
enhancement factors **1**:379–380
facilitation **1**:377
functional diversity (FD) **1**:378
functional redundancy **1**:378
insurance hypothesis **1**:378
intraguild interference **1**:378–379, **1**:378*F*
risk reduction/risk enhancement **1**:376–377
sampling/selection effects **1**:377
research background **1**:373–374
summary discussion **1**:382–383
vegetation diversity
crop diversity **1**:375–376
natural enemies hypothesis **1**:375–376
noncrop plants **1**:376, **1**:376*F*
biofuel crop production **1**:530
biogeochemical cycles **6**:155
biological pest control **2**:391
classification and use
cost–benefit analyses **5**:692
effect on pest **5**:692*T*

pesticides/pest control (*continued*)
formulations **5**:693
fungicides
inorganic fungicides **5**:694–695
organic fungicides **5**:695
herbicides
aliphatic acids **5**:696
benzonitriles **5**:697
bipyridiliums **5**:696, **5**:699*T*
carbamates **5**:696
dinitroanilines **5**:696
diphenyl ethers **5**:696
general discussion **5**:695–696
heterocyclic nitrogens **5**:696
inorganic herbicides **5**:695–696, **5**:699*T*
organic arsenicals **5**:696
organic herbicides **5**:696
phenol derivatives **5**:696–697
phenoxyaliphatic acids **5**:696
substituted amides **5**:696
substituted ureas **5**:696
thiocarbamates **5**:696
triazines **5**:696, **5**:699*T*
insecticides
inorganic insecticides **5**:693
organic insecticides **5**:693
organic fungicides
acylalanines **5**:695
2-aminopyrimidines **5**:695
benefits **5**:695
benzimidazoles **5**:695
dicarboximides **5**:695
dinitrophenols **5**:695, **5**:699*T*
dithiocarbamates **5**:695, **5**:699*T*
ergosterol biosynthesis inhibitors (EBIs) **5**:695
organometallics **5**:695
organophosphates **5**:695
oxathiins **5**:695
phenylamide **5**:695
phthalamides **5**:695
pyrimidines **5**:695
quinones **5**:695
substituted aromatics **5**:695
systemic compounds **5**:695
thiophantes **5**:695
triazines **5**:695
organic insecticides
antibiotics **5**:694
botanicals **5**:693
carbamates **5**:694, **5**:699*T*, **5**:700*T*
dinitrophenols **5**:694
formamidines **5**:694
fumigants **5**:694
insect growth regulators **5**:694
microbials **5**:694
new insecticides **5**:694
organochlorines **5**:693, **5**:699*T*, **5**:700*T*
organophosphates **5**:693–694, **5**:699*T*, **5**:700*T*
organotins **5**:694
petroleum oils **5**:694
pyrethroids **5**:694, **5**:699*T*
semiochemicals **5**:694
pesticide types **5**:693
type of pest **5**:692, **5**:692*T*
ecotoxicology
aquatic systems **5**:701
avian impacts **5**:701
ecotoxicological studies **3**:118–119
farm-scale observations **5**:700–701, **5**:700*T*
pesticide management **5**:699–701, **5**:701
top-ranked pesticides **5**:700*T*
toxicity classes **5**:699*T*
vertebrates **5**:699*T*
efficiency and applications
bioavailability **5**:697
delivery
formulations **5**:697–698
obstacle deposition **5**:699
soil treatments **5**:697–698
spray applications **5**:698–699, **5**:698*T*
utilization efficiency **5**:697, **5**:697*T*
environmental impacts **3**:166, **3**:282, **3**:575–576, **6**:155
government regulation **4**:90–91
habitat degradation **3**:406
herbicides
biodiversity effects **4**:91–92
characteristics and classifications **4**:89, **4**:90*T*
environmental fate **4**:89–90
government regulation **4**:90–91
toxicity **4**:91*T*
historical perspective **5**:692
intercropping
biological pest control **2**:391
genetic diversity effects **2**:391–392
push–pull strategies **2**:391
species diversity effects **2**:392
moths **5**:402–403
Near East ecosystems **5**:473–475
parasite use **5**:673
river ecosystems **6**:481*T*, **6**:482
terrestrial invertebrates **3**:221–222
Petalomonas **6**:293
Petaurus breviceps **4**:689
petrels **1**:617
Petrochromis **1**:25*F*
Petrogale burbidgei **4**:688–689
Petrogale spp. **6**:333
Petroica macrocephala **7**:189
Petroica traversi **5**:369, **7**:189
petroleum oils **5**:694
petroleum products
biodiversity impacts **3**:253–256, **3**:255*T*, **4**:145
world energy use **3**:251*F*, **3**:252*T*
Petrolisthes spp. **5**:33
Petromyzon marinus **4**:363
Petrosaviaceae **6**:315*T*
Petroselinum sativum **2**:662–663
Petrotilapia **1**:25*F*
pe tsai **2**:663
pet trade **6**:254–255
Peucaea aestivalis **7**:183–184
Peumus **5**:621*F*
Pezophaps **1**:606
Pfiesteria piscicida **1**:96, **2**:531
Phacelia **5**:215
Phacochoerus aethiopicus **4**:696
Phacochoerus africanus **1**:52
Phaeocystis **1**:474–475, **6**:30
Phaeocystis antarctica **6**:349–350
phaeohyphomycosis **3**:634*T*
Phaeophyceae **3**:571*T*, **6**:10*T*
Phaethon lepturus **7**:185
phagocytosis **5**:305
phagotrophic protists **5**:306
phalanger **4**:687*T*, **4**:701*T*
Phalangioidea **1**:209–210
Phalangium opilio **1**:210
Phalangodidae **1**:210–211
Phalangopsinae **3**:730
Phalaris arundinacea **4**:366, **6**:457
phallotoxins **3**:635*T*
Phanerochaete chrysosporium **3**:639
Phaneropterinae **3**:726–727
Phanerozoic Eon
Cambrian explosion **3**:683, **3**:684*F*
fossil record **1**:411*F*, **1**:412*F*, **1**:491, **3**:419*F*
marine diversity timeline **1**:492*F*, **3**:419*F*
mass extinctions **1**:412–413, **1**:413*F*, **1**:493, **1**:493*F*, **1**:494*F*, **1**:494*T*, **3**:416*T*, **3**:419*F*
neoproterozoic animals **3**:682–683
originations **1**:495, **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
taxa diversity **1**:411*F*, **1**:412, **1**:412*F*
phantom crane flies **3**:489–491*T*
phantom midges **3**:489–491*T*, **4**:374*F*, **6**:26
pharmaceuticals
biodiversity **5:703–715**
characteristics
marine organisms **5**:711
microbial products **5**:710–711
terrestrial animals **5**:711
terrestrial plants **5**:711–713, **5**:712*T*
ethical perspectives
bilateral research agreements **5**:714
citizens' rights **5**:715
Convention on Biological Diversity (1930) **5**:713–714
indigenous intellectual property rights **5**:714–715
sick and afflicted populations **5**:715
evolutionary perspective **5**:703–705
general discussion **5**:703
historical perspective
ancient and early research **5**:705
herbals and medicinal plants **5**:705–706
modern drug discovery approaches
bioassays **5**:707–708
isolation techniques **5**:708
research background **5**:706–707
structural determination methods **5**:708
vouchering **5**:707
rational screens
characteristics **5**:708–709
ecological screens **5**:709–710
ethnobotanical screens **5**:710
phylogenetic screens **5**:709
random screens **5**:709

economic importance **1**:393–395
fungi-derived drugs **3**:638
plant-derived drugs and chemicals **6:129–139**
algae
domoic acid **6**:132, **6**:134*F*
green algal metabolites **6**:132–135, **6**:134*F*
kainic acid **6**:132, **6**:134*F*
cyanobacteria
anticancer agents **6**:130–131
cryptophycin **6**:131, **6**:131*F*
curacin A **6**:130–131, **6**:130*F*
dinoflagellates
brevetoxin A **6**:131–132, **6**:134*F*
characteristics **6**:131
ciguatoxin **6**:131, **6**:132*F*
Gymnodinium breve **6**:131–132, **6**:134*F*
maitotoxin **6**:131, **6**:133*F*
future outlook **6**:138
herbal remedies
Echinacea **6**:137–138, **6**:138*F*
ginseng **6**:137, **6**:138*F*
higher plants
camptothecin **6**:135–136, **6**:135*F*
paclitaxel **6**:135, **6**:135*F*
marine plants
algae **6**:132
cyanobacteria **6**:130–131
dinoflagellates **6**:131
research background **6**:129–130
terrestrial plants
Digitalis sp. **6**:136–137, **6**:137*F*
herbal remedies **6**:137
higher plants **6**:135
steroidal glycosides **6**:136–137, **6**:137*F*
tropane alkaloids **6**:136, **6**:136*F*
Pharomachrus mocinno **7**:183
Phascolarctidae **4**:688–689
Phascolarctos cinereus **4**:688–689
Phascolarctos stirtoni **5**:363–364, **5**:364*F*
Phascolomys **4**:727–728
Phascolomys major **5**:364*F*
Phascolomys medius **5**:364*F*
Phascolonus **4**:727–728
Phascolonus gigas **5**:364*F*
Phaseolus acutifolius **2**:663–664
Phaseolus coccineus **2**:663–664
Phaseolus lineatus **1**:165
Phaseolus lunatus **2**:663–664, **4**:281–285, **4**:283*F*, **4**:289*T*
Phaseolus spp. **2**:383, **3**:32*F*, **6**:553
Phaseolus vulgaris **1**:462*F*, **2**:383, **2**:663–664, **6**:2, **6**:361*T*
Phasianidae **1**:623–624, **6**:308*T*
Phasianus colchicus **1**:623–624
Phasmatodea **4**:308–312, **4**:309*T*
Phasmodinae **3**:727
Phayre's langur **6**:253–254
pheasants **1**:623–624, **1**:644*F*
Pheidole megacephala **3**:222, **4**:365
Phellodendron amurense **1**:271–272
Phenacoccus manihoti **4**:183, **4**:361
Phenakospermum guianense **1**:165
phenetic methods **6**:16–17, **7**:85, **7**:102–104, **7**:103*F*, **7**:107*F*
phenmedipham **5**:696
phenolic compounds **2**:426–427, **2**:426*T*, **5**:695, **5**:696–697
phenotype **6:1–8**
basic concepts
artificial selection **6**:4
gene modification **6**:3–5
genomic balance **6**:4
quantitative trait loci (QTL) **6**:4, **6**:5*F*
definition **6**:1
evolutionary models **6**:7–8
genetic load theory **6**:6–7
genotype–environment analysis **6**:5–6
historical research **6**:1–3
phenotypic plasticity **4**:638–639, **4**:639*F*, **6**:7–8
phenotypic variability **6**:6–7
taxonomic diversity **6**:8
phenoxy acid **4**:90*T*
phenoxyaliphatic acids **5**:696
phenylamide **5**:695
phenylcarbamate **4**:90*T*
phenylurea **4**:90*T*
pheromones **5**:675–676
Philadelphus schrenkii **1**:271–272
Philesturnus carunculatus **5**:366–367
Philippine eagles **7**:183
Philippines
animal species **1**:261*T*
biodiversity conservation **1**:520–521
biodiversity richness **1**:501*F*
botanical gardens **6**:519
deforestation **1**:502*T*, **2**:449*T*, **4**:562–564, **4**:563*F*
ecosystems
coral reefs **1**:500*T*, **2**:333*T*
genetic richness **1**:514*T*
mangroves **1**:275*T*, **1**:277, **1**:513*T*
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical dry forest **1**:266–267
tropical moist forest **1**:277*T*
tropical/subtropical humid mountain areas **1**:267
vegetation types and coverage percentage **1**:264–265*T*
education level–fertility relationship **6**:204, **6**:205*F*
education level–infant mortality relationship **6**:205*F*
endangered mammals **3**:194*T*
endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
fish diversity **3**:473
fishing/fisheries **6**:422*T*
indigenous peoples **4**:270*T*
institutional collections and capacity **1**:505*T*
invasive species **4**:346–347
nitrogen deposition **5**:526*T*
pastureland conversion **4**:665*T*
plant species **1**:259*T*, **6**:82*T*
species richness **1**:501*F*, **4**:271*T*
threatened birds **7**:181
see also Southeast Asia
Philodendron **3**:522*T*
Philodendrum **4**:281–285, **4**:282*F*
Philodota **4**:697
Philodromidae **5**:463*T*
Philomachus pugnax **7**:32
Phiomia **5**:224*T*
Phleum alpinum **1**:232
pH levels
aquatic ecosystems **6**:152, **6**:153*F*
freshwater ecosystems **1**:9–10, **1**:10*T*, **3**:584–585, **3**:584*F*
species diversity controls **7**:391–393
wetland ecosystems **7**:393
Phobetron pithecium **5**:391*F*
Phoca groenlandica **5**:314, **7**:37–38
Phoca largha **5**:90
Phocarctos hookeri **5**:81*T*
Phoca vitulina **2**:528*T*, **3**:192–193, **3**:256
Phoca vitulina mellonae **5**:77*T*
phocids **5**:67
phocine distemper virus **2**:528*T*
Phocoena phocoena **5**:77*T*
Phocoena sinus **5**:77*T*, **5**:80
Phocoenoides dalli **3**:192–193
Phoeniconaias minor **1**:54–55
Phoenicopterus andinus **7**:183
Phoenicopterus ruber **1**:54–55
Phoenix **1**:70
Phoenix dactylifera **1**:53, **2**:662–663, **5**:484–485
Pholadomya candida **2**:5*F*
Pholidota **4**:690*T*, **4**:701*T*
Pholoidae **7**:417*F*
Phoradendron piperoides **4**:77
Phoradendron serotinum **4**:77
phorate **5**:693–694
phoresy **3**:495, **5**:656
Phoridae **3**:489–491*T*, **3**:492, **3**:493–494
Phorinda **4**:381*T*
Phormidium **3**:358, **4**:119
Phoronidea
fossil record **3**:684*F*
global diversity **4**:385*T*
hydrothermal vent ecosystems **7**:329–330
marine ecosystems **5**:50*T*
number of species **5**:669–670*T*
phosmet **5**:693–694
phosphoenolpyruvate (PEP) **2**:363–364, **2**:363*F*
phosphoglucose isomerase **3**:665
phosphoglycolate (PG) **2**:361–362, **2**:362*F*
phosphogylcerate (PGA) **2**:361–362, **2**:362*F*
phospholipid fatty acids (PLFAs) **5**:267, **5**:268*F*
phosphorus (P)
biogeochemical cycles **1**:563
carbon:nitrogen:phosphorus (C:N:P) stoichiometry **3**:506–507, **4**:675–677, **4**:676*F*
desert ecosystems **2**:474
elemental partitioning **1**:554*T*
freshwater ecosystems **3**:572*T*
intercropping systems **2**:390
nitrogen:phosphorus (N:P) ratio **2**:617–618, **2**:617*F*

phosphorus (P) (*continued*)
nutrient cycling **1**:63–65, **1**:705–706, **3**:348–351, **3**:349*F*
nutrient management–water quality relationship **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
freshwater invertebrates **4**:377
land-use change impacts **1**:107–108
leaching vulnerability index **1**:102*F*
management criteria **1**:97, **1**:97*T*
nutrient input and sources **1**:96*F*, **1**:98–99, **3**:575*F*, **3**:576–577
regional nutrient imbalances **1**:98*F*, **1**:99–101, **1**:100*F*, **1**:102*F*
remedial benefical management practices (BMPs)
characteristics **1**:103*T*
cover crops **1**:103*T*, **1**:106
critical source area management **1**:103*T*, **1**:107
erosion potential **1**:103*T*, **1**:105–106, **1**:105*T*, **1**:106*F*
farm input decisions **1**:101, **1**:103*T*
fertilizer application methods **1**:103*T*, **1**:104
fertilizer application rate and frequency **1**:103*T*, **1**:104–105
implementation strategies **1**:101, **1**:103*F*
nutrient-use efficiency improvements **1**:101–104, **1**:103*T*
riparian/buffer zones **1**:103*T*, **1**:106–107
runoff potential **1**:103*T*, **1**:105
soil and plant testing **1**:103*T*, **1**:104
source management **1**:103*T*, **1**:104
transport management **1**:103*T*, **1**:105
research background **1**:95–97
research summary **1**:108
organophosphates
fungicides **5**:695
insecticides **5**:693–694, **5**:699*T*, **5**:700*T*
phosphorus and nitrogen inputs **3**:349*F*
phosphorus gradient–hierarchical partitioning measures **3**:606–608, **3**:607*F*
relative abundance **1**:553*F*
root–soil interface **6**:114*F*
total phosphorus/total nitrogen–chlorophyll *a* relationship **3**:348–350, **3**:349*F*
water-column nitrogen:phosphorus (N:P) ratios **3**:350–351, **3**:350*F*
photoaging **7**:302
photoautotrophs
energy flow **3**:242
environmental requirements **6**:28–31
evolutionary development **2**:640–641
temperature tolerance ranges **4**:113–115
Photobacterium leiognathi **5**:121–123, **5**:122*F*
Photobacterium phosphoreum **3**:122*T*, **5**:121–123
photoheterotrophs **2**:640–641
photokeratitis **7**:302
photolithoautotrophs **2**:612*T*, **2**:614
photolyase levels **3**:165
Photorhabdus **7**:430–431
photosynthesis **6:9–15**
adenosine triphosphate (ATP)
carbon dioxide (CO_2) assimilation reactions **6**:13–15, **6**:14*T*
reaction centers **6**:12–13
C_3 plants
anatomical modifications **2**:364*F*
C_3 physiology **6**:14, **6**:14*T*
energy requirements **2**:369*T*
grasses **1**:58–59, **1**:64*F*, **2**:654–655
historical atmospheric conditions **2**:367*T*
photosynthetic carbon reduction cycle **2**:361–362, **2**:362*F*
quantum yield **2**:369*F*
temperature effects **2**:369*F*
C_4 plants **2:361–381**
African savannas **1**:64*F*
animal diversity impacts **2**:377–378
atmospheric carbon dioxide (CO_2) concentrations **2**:361
biogeography
C_4 annuals **2**:373*T*
general characteristics **2**:370–372
global biomes **2**:371*T*
latitudinal gradients **2**:370–372, **2**:370*F*
light intensity **2**:370
moisture availability **2**:372
principle life forms **2**:371*T*
salinity/salt tolerance **2**:371*T*, **2**:372–373
temperature parameters **2**:370–372
C_4 physiology **6**:14*T*, **6**:15
climate change effects **2**:378
ecological dominance **2**:377
economically significant C_4 plants **2**:362*T*
energy requirements **2**:369*T*
evolutionary development
carbon starvation hypothesis **2**:374–377
evolutionary centers **2**:375*F*
evolutionary lineages **2**:369*F*, **2**:373, **2**:374*T*, **2**:375*F*, **2**:376*F*
molecular clock analysis **2**:373–374, **2**:376*F*
polyphyletic origins **2**:373
taxonomic diversity **2**:373
exotic C_4 plants **2**:379–381, **2**:380*F*
grasses **1**:59*F*, **1**:64*F*, **2**:378–379, **2**:379*F*, **2**:654–655, **2**:683
historical atmospheric conditions **2**:367*T*
land-use change impacts **2**:378–379, **2**:379*F*
photosynthetic mechanisms
anatomical modifications **2**:364, **2**:364*F*
basic concepts **2**:362–364, **2**:363*F*
biochemistry **2**:363–364
C_3 photosynthetic carbon reduction cycle **2**:361–362, **2**:362*F*
diffusion barriers **2**:365–366
photorespiration **2**:361–362, **2**:362*F*
requirements **2**:363*T*
schematic diagram **2**:363*F*
single-cell C_4 photosynthesis **2**:364–365, **2**:365*F*
physiological significance
atmospheric carbon dioxide (CO_2) effects **2**:366–368, **2**:366*F*, **2**:368*F*, **2**:369*F*
energy requirements **2**:369*T*
historical atmospheric conditions **2**:367*T*
nitrogen-use efficiency **2**:367*T*, **2**:368–370
photorespiration **2**:366–368, **2**:366*F*
quantum yield **2**:369*F*
temperature effects **2**:367, **2**:369*F*
water-use efficiency **2**:367*T*, **2**:368–370
quantum yield **2**:369*F*
research summary **2**:381
temperature effects **2**:369*F*
tropical deforestation effects **2**:379–381, **2**:380*F*
tropical savannas **1**:58–59, **1**:59*F*
vegetation–climate interactions **2**:683
Crassulacean Acid Metabolism (CAM) **6**:14*T*, **6**:15
ecosystem services **3**:90*T*
excitation energy transfer **6**:9–11
freshwater ecosystems **3**:560, **3**:564*F*
light-harvesting pigment–protein complexes **6**:9–11, **6**:10*T*
nicotinamide adenine dinucleotide phosphate (NADPH) dehydrogenase
carbon dioxide (CO_2) assimilation reactions **6**:13–15, **6**:14*T*
reaction centers **6**:12–13
phosphorylation reactions **6**:9
photosynthetic symbioses **5**:118–120
reaction centers
adenosine triphosphate (ATP) generation **6**:12–13
excitation energy transfer **6**:9–11
nicotinamide adenine dinucleotide phosphate (NADPH) dehydrogenase **6**:12–13
primary photochemical processes **6**:11–12, **6**:11*T*
research summary **6**:15
phototrophs **5**:305–306
phototrophy, microbial **5**:301*T*
photovoltaics **3**:253*T*, **3**:260–262, **3**:265*T*
Phragmites **1**:270, **1**:278–279
Phragmites australis **1**:451, **4**:366, **5**:484–485, **6**:93–95, **6**:115–116, **6**:457, **7**:376–378, **7**:385
Phragmites communis **7**:399
Phragmites frutescens **5**:486
phreatic water **7**:49–50
phrygana **4**:577
phthalamides **5**:695
Phthiraptera **4**:308–312, **4**:309*T*

phycobilinogen **6**:10*T*
Phycodnaviridae **5**:134*T*, **5**:136
Phycomyces **3**:627
phyletic gradualism **2**:47–48, **2**:48*F*
Phylica **5**:219
Phylica nitida **1**:703
Phyllanthus **6**:316*T*
Phyllanthus fluitans **1**:164
Phyllitis scolpendium **7**:52–53
Phyllobates terribilis **1**:176
Phyllocarida **2**:401*T*
Phyllocladus **7**:118–119
Phyllodocida **7**:417*F*
Phyllodocidae **7**:417*F*
Phyllonoma **3**:522*T*
Phyllonorycter apparella **5**:395–396
Phyllonorycter blancardella **5**:388*F*, **5**:395–396
Phyllonorycter salicifoliella **4**:173–174
Phyllospadix **6**:542*T*
Phyllospadix iwatensis **6**:541*T*
Phyllospadix japonicus **6**:541*T*
Phyllospadix scouleri **6**:541*T*
Phyllospadix serrulatus **6**:541*T*
Phyllospadix torreyi **6**:541*T*
Phyllostachys **1**:263
Phyllostomidae **6**:309*T*
PhyloCode **5**:554
phylogenetic biodiversity **1**:430
phylogenetic species concept **6**:683–684, **7**:91–93, **7**:92*F*
phylogeny **6**:**16–23**
  analytical methods
    cladistics **7**:104–106, **7**:104*T*, **7**:105*F*, **7**:105*T*
    operational taxonomic units (OTUs) **5**:252–254, **7**:102–104, **7**:103*F*, **7**:103*T*, **7**:104*T*
    phenetic methods **6**:16–17, **7**:102–104, **7**:103*F*, **7**:107*F*
    phylogenetic methods
      cladistics **6**:17
      computational challenges **6**:17–18
      homology **6**:17–18
      maximum likelihood analysis **6**:18
      neighbor-joining method **6**:18
      parsimony **6**:17
    phylogenetic systematics **7**:104–106
  basic concepts **6**:16
  classifications
    challenges **6**:20–21
    ranks and priority **6**:21
  data sources **6**:19
  endemism **3**:234–235
  historical research **6**:16
  interpretation challenges **6**:19–20, **6**:20*F*
  measures of support
    bootstrap support **2**:44, **6**:18
    Bremer support **2**:44, **6**:19
    jackknife support **6**:18–19
  modern research approaches **6**:16
  phylogenetic trees
    adaptation studies **6**:22
    interpretation challenges **6**:19–20, **6**:20*F*
  research applications
    adaptation studies **6**:22
    evolutionary generalizations **6**:21–22
    phylogenetic diversity **6**:22
    vicariance biogeography **6**:22
  research summary **6**:22
Phymatura brevicollis **7**:220*T*
Physalis angulata **4**:288*T*
Physarum **6**:290*F*
Physarum polycephalum **5**:572–575, **5**:575*F*
Physeter catodon **3**:256
Physeter macrocephalus **2**:272, **5**:88–89, **6**:578–579
physico–chemical disturbances **2**:564–565, **2**:565*F*
Physidae **4**:374
Physocalymma scaberrimum **1**:161
Physostigma venenosum **5**:712*T*
physostigmine **5**:712*T*
Phytolacca dodecandra **1**:592
Phytophaga **1**:328
Phytophthora **4**:193–194, **4**:195*F*, **4**:196*T*
Phytophthora cinnamoni **2**:528*T*, **6**:449
Phytophthora infestans **1**:359–360, **3**:630, **4**:313–314, **6**:297
Phytophthora ramorum **2**:528*T*, **4**:505
phytoplankton
  Census of Marine Life project **2**:18–19, **2**:19*F*
  characteristics **5**:305–306, **5**:683–686, **5**:684*F*, **6**:24–25
  classifications and sampling methods **4**:382–383, **4**:383*T*
  climate change effects **5**:37–38
  color distribution **5**:686*F*
  community structure **6**:25–26
  conservation practices **6**:77
  distribution studies **4**:668–669
  ecosystem services **6**:77
  energy flow **3**:246–247
  environmental requirements **6**:28–31
  eutrophication impacts
    long-term human influences **3**:354–356
    species shifts **3**:352–354, **3**:353*F*
  freshwater ecosystems **3**:563*T*, **3**:565–566
  lake and pond ecosystems **3**:243*F*
  mercury contamination **1**:12*F*
  nitrogen cycle impacts **5**:543
  nutrient cycling **3**:244–245
  oil spills **3**:256
  phenological shifts **5**:719*F*
  predation effects **6**:233–235
  river ecosystems **6**:472*F*, **6**:477, **6**:478*T*
  surface water acidification **1**:10*T*
  trophic pyramid **4**:670–671
  ultraviolet radiation effects **7**:299
  water-column nitrogen:phosphorus (N:P) ratios **3**:350*F*
phytotelmata **3**:494–495
pica **4**:382–383, **4**:383*T*
Pica pica **4**:174
Picea
  Asia **1**:272–273, **1**:273–274
  boreal forests **1**:628–630
  geographic boundaries **1**:629*F*, **1**:630*T*
  North America **7**:118
Picea abies **1**:273–274, **1**:628–629, **7**:212–213
Picea asperata **1**:272–273
Picea brachytyla **1**:272–273
Picea cembra **1**:273–274
Picea engelmannii **1**:612, **1**:631*T*, **2**:571
Picea excelsa **1**:630*T*
Picea glauca **1**:628–629, **1**:630*T*, **1**:632*T*, **3**:535, **6**:120*T*, **6**:123–124, **6**:466
Picea glehnii **1**:273
Picea jezoensis **1**:273, **1**:630*T*
Picea likiangensis **1**:272–273
Picea mariana **1**:628–629, **1**:630*T*, **1**:632*T*, **7**:400
Picea noveitchii **1**:272–273
Picea obovata **1**:273–274, **1**:630*T*
Picea pumila **1**:273–274
Picea rubens **1**:632*T*, **3**:440
Picea sibirica **1**:273–274
Picea sitchensis **6**:125
Picea sylvestris **1**:273–274
Picidae **4**:33–35, **4**:34*F*, **6**:308*T*
Piciformes **6**:308*T*
pickerel **1**:12*F*, **3**:464–465, **7**:390
picloram **5**:696
Picoides albolarvatus **4**:67*F*
Picoides borealis **2**:264, **7**:183–184
picolinic acid **4**:90*T*
picoplankton
  classifications **4**:383*T*, **4**:674
  Global Ocean Sampling (GOS) program
    cross-biome sampling studies **5**:294
    fragment recruitment **5**:294–295, **5**:295*F*
    functional diversity (FD) **5**:295
    research background **5**:293
    research outlook **5**:297
    sampling sites **5**:293*F*
    Sargasso Sea **5**:293–294
    taxonomic diversity **5**:297
  lake and pond ecosystems **4**:460
  Marine Microbial Genome Sequencing Project (MMGSP) **5**:295–297
  marine viruses **5**:132, **5**:133*F*
  taxonomic and functional diversity **5**:297
Picornavirales **5**:134*T*, **5**:137
Picrophilus **1**:220, **1**:221*T*
picrotoxin **5**:712*T*
picture-winged flies **1**:28–29
pie-billed grebe **1**:621
pied bare-faced tamarin **6**:256–258*T*
pied flycatchers **5**:720
pied kingfisher **5**:476–477
pied monarch **2**:75*F*
piedra **3**:634*T*
Piedraia hortae **3**:634*T*
Pieridae
  geographic diversity
    latitudinal gradients **1**:657, **1**:657*T*, **1**:658*F*
    neotropics **1**:657–658, **1**:658*T*
    regional patterns **1**:658, **1**:659*T*
    species numbers **1**:657*T*
  taxonomy **1**:650, **1**:651*F*
  tropical rain forests **6**:307*T*
Pierinae **1**:650–651, **2**:138
Pieris rapae **2**:551–552*T*
pigeon peas **2**:383, **3**:32*F*

pigeons
  introduced species
    Brazil **1**:83
    India **1**:82
    South Africa **1**:81
    United Kingdom **1**:79
    United States **1**:78, **1**:624
  risk status classification **1**:644*F*
pigmy rice rat **1**:367–368
pigs
  animal breeding programs **1**:640
  convergent evolution **4**:701*T*
  introduced species
    Australia **1**:80
    Brazil **1**:83
    India **1**:82
    South Africa **1**:81
    United Kingdom **1**:79
    United States **1**:76–77
  invasive species **3**:222
  number of species **4**:690*T*
  risk status classification **1**:644*F*
pig-tailed macaques **5**:603–604
pika **2**:85
Pikaia gracilans **7**:334–335
pikas **4**:690*T*, **4**:697
pike **1**:12*F*, **3**:365*F*, **3**:464–465, **4**:463, **6**:522–523
Pilargidae **7**:417*F*
Pilbara shrublands **2**:468
Pilgerodendron uviferum **4**:105–107
Piliocolobus badius badius **6**:256–258*T*
Piliocolobus badius temminckii **6**:256–258*T*
Piliocolobus bouvieri **6**:256–258*T*
Piliocolobus epieni **6**:256–258*T*
Piliocolobus gordonorum **6**:256–258*T*
Piliocolobus kirkii **6**:256–258*T*
Piliocolobus pennantii **6**:256–258*T*
Piliocolobus preussi **6**:256–258*T*
Piliocolobus rufomitratus **6**:256–258*T*
Piliocolobus tephrosceles **6**:256–258*T*
Piliocolobus waldronae **6**:250, **6**:256–258*T*
Piliostigma thonningii **1**:51
pill bugs
  characteristics **4**:392
  endangered species **3**:226
pill millipedes **5**:429–430
Pilobolus **3**:627
pilocarpine **5**:711–713, **5**:712*T*
Pilocarpus jaborandi **5**:712*T*
Pilocarpus microphyllus **1**:167
Pilosa **2**:586, **2**:587*F*, **4**:689
Pilostyles **4**:72
pilot whales **5**:314
Pinaceae **4**:354–355
Pinaroloxias inornata **1**:27–28
Pinchot, Gifford **2**:230–231, **2**:282–283
pin-cushion millipedes **5**:429–430
pineapple **2**:663–664, **4**:287–290
pine bark beetles **6**:150*T*
Pine Barrens **5**:562–563
pine forests
  Asia **1**:272, **1**:273–274
  boreal forests **1**:628–630, **7**:218*F*, **7**:219*F*
  geographic boundaries **1**:629*F*, **1**:630*T*
  Mediterranean North Africa **1**:47
  North America **5**:564–565, **7**:118
  soil properties and nutrients **6**:124*F*
pine martens **1**:692*F*
Pinguicula **6**:44*T*
Pinguicula vallisneriifolia **6**:45
Pinguinus impennis **7**:182
pink fairy armadillo **3**:193
pink pigeons **7**:188
pink salmon **3**:256, **6**:236, **6**:522, **6**:523
pinnipeds
  adipose tissue **7**:34
  Census of Marine Life project **2**:12*F*
  characteristics **5**:65*T*, **5**:67
  climate change effects **5**:89–91
  evolutionary history **5**:69
  migratory patterns **5**:314
  species diversity **4**:693
  taxonomy **5**:65*T*
Pinnocaris sp. **5**:375*F*
piñon pine trees **4**:212*F*
Pinus
  Asia **1**:273–274
  boreal forests **1**:628–630
  geographic boundaries **1**:629*F*, **1**:630*T*
  phylogenetic relationships **5**:618*F*
Pinus albicaulis **2**:571, **2**:571
Pinus aristata **7**:118
Pinus armandii **1**:272
Pinus ayacahuite **2**:29–30
Pinus banksiana **1**:630*T*, **1**:631*T*, **1**:632*T*
Pinus brutia **1**:269, **6**:92
Pinus caribaea var. *hondurensis* **7**:274
Pinus caribea **6**:458–459
Pinus cembra **1**:630*T*
Pinus contorta **1**:631*T*, **1**:632*T*, **6**:125, **7**:385
Pinus densiflora **1**:271
Pinus edulis **4**:174, **4**:212*F*
Pinus elliottii **6**:455
Pinus halepensis **1**:47, **1**:269, **2**:541*F*, **4**:355, **5**:481–482, **6**:92
Pinus hartwegii **2**:29–30
Pinus jeffreyi **1**:144
Pinus kesiya **1**:272, **7**:274–275
Pinus koraiensis **1**:271–272
Pinus massonia **4**:77
Pinus massoniana **1**:272
Pinus merkusii **1**:70, **3**:437–438, **7**:274–275
Pinus oocarpa **2**:29–30, **7**:275
Pinus palustris **3**:437–438, **6**:455
Pinus pinaster **4**:354–355, **4**:355*F*, **6**:92, **6**:94*F*
Pinus pinea **6**:92
Pinus ponderosa **1**:144, **3**:209
Pinus pumila **1**:273, **1**:630*T*
Pinus radiata **4**:355
Pinus rigida **6**:456–457
Pinus roxburghiana **1**:272–273
Pinus roxburghii **7**:274–275
Pinus rudis **2**:29–30
Pinus sibirica **1**:272–273, **1**:630*T*
Pinus spp. **2**:25–26
Pinus strobus **2**:29–30, **2**:279, **3**:631–632
Pinus sylvestris **1**:272–273, **1**:628–629, **1**:630*T*, **7**:212–213
Pinus tabulaeformis **1**:269, **1**:272
Pinus taeda **6**:361*T*, **7**:118
Pinus wallichiana **7**:274–275
pinworms **5**:660*T*, **5**:671*T*, **7**:421*F*, **7**:431–432
Piophilidae **3**:489–491*T*
pipefish **3**:467–471
Piper **6**:316*T*
Piperaceae **2**:429–430, **5**:618*F*, **6**:316*T*
Piperales **5**:621*F*
Piper arieianum **2**:429–430
piperazine **5**:695
Piper methysticum **5**:704
Piper nigrum **2**:663, **7**:26
Pipistellus kuhli **5**:474–475
pipistrelle **4**:704
Pipistrellus pipistrellus **4**:704
pipits **2**:556, **2**:556*F*
Pipridae **6**:308*T*
Pipunculidae **3**:489–491*T*
piquiazeiros **4**:287–290
piranhas **3**:464*F*
Piranhea trifoliata **1**:163
pirate perch **3**:466–467
pirizal **1**:164
Pisa Botanic Garden **6**:506
Pisaster **3**:125
Pisaster ochraceus **4**:335, **4**:442–443, **4**:443*F*, **5**:37, **6**:232, **6**:722, **6**:723*T*
piscivores **6**:233–235
Pisionidae **7**:417*F*
pistachio nuts **3**:32*F*
Pistacia atlantica **1**:47, **5**:480*F*, **5**:482
Pistacia khinjuk **5**:480*F*, **5**:485
Pistacia lentiscus **1**:269, **5**:480*F*, **5**:482–483, **6**:443
Pistacia palaestina **5**:481–482
Pistacia palestina **1**:269
Pistia stratiotes **1**:165
Pisum sativum **2**:383, **2**:662–663
Pitcairn **6**:507*T*
pitcher plant **2**:643, **3**:494–495, **6**:45, **7**:396*F*
pitcher plant midge **4**:518
pitcher plant mosquito **5**:724–725
pitfall traps **1**:330
Pithecellobium cochleatum **1**:164
Pithecellobium spp. **1**:120, **5**:489
Pithecelobium multiflorum **1**:163
Pithecia pithecia **3**:127*T*
Pitheciidae **4**:692, **6**:250–252, **6**:251*T*
Pithecophaga jefferyi **7**:183
Pittosporum patulum **4**:194*T*
pit vipers **1**:359
pityriasis versicolor **3**:634*T*
Placentalia
  Afrosoricida **4**:690
  Artiodactyla **4**:695–697
  carnivores **4**:693
  Cetacea **4**:693–694
  characteristics **4**:683, **4**:689, **7**:341
  Chiroptera **4**:690*T*, **4**:691–692
  Cingulata **4**:689–690
  cranial and skeletal characteristics **4**:683–684
  Dermoptera **4**:690*T*, **4**:691
  Erinaceomorphia **4**:691
  Hyracoidea **4**:695

Lagomorpha 4:690*T*, 4:697
Macroscelidea 4:690*T*, 4:697–698
number of species 4:690*T*
Perissodactyla 4:695
Philodota 4:697
phylogenetic relationships 4:689, 5:1*F*
Pilosa 4:689
primates 4:690*T*, 4:692–693
Proboscidea 4:694–695
reproductive characteristics 4:684, 4:684*F*
rodents 4:690*T*, 4:697
Scandentia 4:690*T*, 4:691
Sirenia 4:694
Soricomorpha 4:690–691
species diversity 4:690*T*
Tubulidentata 4:695
placental radiation 1:22
Placentonema gigantissima 7:429
Placidochromis 1:25*F*
Placophora 5:669–670*T*
Placozoa
fossil record 3:684*F*
global diversity 4:385*T*
marine ecosystems 5:50*T*
taxonomic diversity 6:712–713
Plagiodontia aedium 4:713, 4:714*F*, 4:715*F*
Plagiodontia ipaneum 4:705
Plagiodontia velozi 4:705
Plagiopteraceae 1:260*T*
plagiopylids 6:295–296, 6:296*F*
Plagiorchioidea 7:439–440*T*, 7:450–451*F*, 7:454
Plagioscyon squamosissimus 3:260
Plagopterus argentissimus 4:364
plaice 5:315
plain-naped proboscis monkey 6:256–258*T*
plains vizcacha 2:590*T*
plains zebras 4:695
planaria 4:371*T*
planktivores 6:233–235
plankton **6:24–38**
average lifespan 4:656*T*
Census of Marine Life project 2:18–19, 2:19*F*
characteristics and functional role 4:667–668, 6:24–25, 6:291
classifications and sampling methods 4:382–383, 4:383*T*
community patterns and diversity 6:32–33
community structure 6:25–26
conservation practices 6:77
diversity hypotheses and implications 6:35–36
ecosystem services 6:77
energy flow 3:246–247
environmental requirements
phytoplankton 6:28–31
zooplankton 6:31
eutrophication impacts
long-term human influences 3:354–356
species shifts 3:352–354, 3:353*F*
functional groups 6:30
habitat constraints 6:24–25, 6:27–28, 6:28*F*, 6:29*F*
habitat–niche relationship 4:39–40, 4:40*F*
lake and pond ecosystems 4:460
Late Ordovician mass extinction 5:167
latitudinal range shifts 4:604*T*
macroscopic patterns **4:667–680**
basic concepts 4:668
biological oceanography 4:668–669
comparative ecological analyses 4:668
macroscopic assembly
functional classifications 4:673
phylogenetic classifications 4:673–675, 4:675*F*
Redfield ratio 4:675–677, 4:676*F*
size classifications 4:671–673
stoichiometric emergence 4:675–677, 4:676*F*
temperature dependence 4:677
trophic pyramid 4:670–671
phytoplankton distribution 4:668–669
regional and global change effects 4:677–678
research summary 4:678–679
temperature dependence
biochemical reactions 4:677
biogeochemical cycles 4:677
whole-organism metabolism 4:677
variability measures
accumulating method 4:670, 4:671*F*, 4:672*F*
coarse graining method 4:669–670, 4:669*F*
multiscale analysis 4:669
marine ecosystems 5:48*F*
marine viruses 5:132–133, 5:133*F*
mercury contamination 1:12*F*
migratory patterns 5:315
nitrogen cycle impacts 5:543
nutrient cycling 3:244–245
species diversity theories 6:26–27
species richness 6:34–35, 6:289*T*
specific gravity effects 6:24
surface water acidification 1:10*T*
temporal habitat diversity fluctuations 6:33–34, 6:33*F*, 6:34*F*
ultraviolet radiation effects 7:299
Plankton Investigatory Collaborative Autonomous Survey System Operon (PICASSO) ROV system 2:13
Planktothrix agardhii 6:30, 6:33*F*
Planktothrix rubescens 6:30
Plantae 2:641
Plantago barbata 4:107
Plantago purshii 2:373*T*
Plantago sericea 4:103
Plantago uniglumus 4:106–107
plantain 4:286–287, 6:553
plantations
biodiversity restoration 6:458–459
deforestation 1:682, 2:453, 2:456
land use issues 4:558
loss and change 6:325–327, 6:326*F*, 6:326*T*
perennial crops and trees 6:560
shade coffee plantations 1:420–421, 1:421, 1:422*F*, 1:423*F*
tree plantations 1:682, 2:456
plant bugs 7:293
Plant Genetic Resources for Food and Agriculture (PGRFA) 3:645
plant hoppers 7:60
PlantLife 6:86
plants
air pollution impacts 1:142*F*
alpine ecosystems
leaf mass ratio (LMR) 1:152*F*
morpho-type diversity 1:151, 1:151*F*
physio-type diversity 1:151–153
reproductive diversity 1:153
specific leaf area (SLA) 1:152, 1:152*F*
taxonomic diversity 1:153–154, 1:154*F*
angiosperms 6:60
Antarctica 1:184*T*, 1:186*T*
anthropogenic activities and impacts 3:401*F*
anthropogenic impacts 6:79–84
ant–plant interactions
benefits to ants/cost to plants 6:46
benefits to plants 6:46–47
desert ecosystems 2:470*T*, 2:474–475
intraguild interference 1:378*F*
myrmecophytes 6:46
myrmecotrophic epiphytes 6:47
aquatic plants 3:359–360, 4:378
Arctic ecosystems 1:227–231, 1:230*T*, 1:231*T*, 1:236*F*
Ascension Island 6:78
Asia 1:258, 1:259*T*, 1:260*T*, 1:278*T*
Australian plant communities **1:291–306**
biogeographic regions 1:297–298, 1:298*F*, 1:299*F*
characteristics 1:291–292
community diversity 1:300, 1:300*F*
evaporative aerodynamics 1:292, 1:293*F*
floristic groups
biogeographic regions 1:298–300
diagnostic species 1:292–297
TWINSPAN floristic groups 1:292, 1:294*T*, 1:300*F*
species richness
dryland vegetation continua 1:302, 1:303*F*
invasive species 1:303
invertebrates 1:300–302, 1:303*T*
isolated rainforest stands 1:303–304
overstory and understory strata 1:300, 1:301*F*, 1:302*F*
postfire succession 1:302–303, 1:304*F*
vertebrates 1:300–302, 1:302*F*
wetland vegetation continua 1:302
structural characteristics 1:301*T*
beetle–host plant interactions 1:23, 1:24*F*
biodiversity **6:56–64**
angiosperms 6:60
biodiversity distribution and loss 6:79
biodiversity patterns
assessment measures 6:60–61
endemism 6:61
biodiversity restoration 6:456–457
characteristics and classifications 6:56
climate effects **2:115–127**
annual mean temperature 2:119–120, 3:119

plants (*continued*)
area–productivity relationship **2**:117–118, **3**:118
biogeographic history **2**:125–126
competition **2**:123
disturbances **2**:115–117, **2**:123, **3**:118
energy availability **2**:123–124
experimental research **2**:126
extreme events **2**:124–125, **3**:124, **3**:125
general discussion **2**:115
genomic research **2**:126
global-scale distribution **2**:118–119
latitudinal gradients **2**:118–119, **3**:119
net primary production (NPP) **2**:117–118, **2**:120, **3**:118, **3**:121
observational research **2**:121–123, **3**:122
precipitation **2**:120, **3**:120
species–area relationships (SARs) **2**:115, **3**:116, **3**:117
summary discussion **2**:126
theoretical perspectives **2**:120–123
functional roles **6**:56–57
gymnosperms **6**:59–60
major threats
anthropogenic impacts **6**:62
conservation programs **6**:63
habitat destruction **6**:62
intrinsic factors **6**:61–62
natural disturbances **6**:62–63
species interactions **6**:61–62
nonvascular plants **5**:528, **6**:58–59, **6**:58*T*
plant–human interactions **6**:57–58
research outlook **6**:63–64
seedless vascular plants **6**:59
biomonitoring processes **3**:122*T*
biotic homogenization **3**:701, **6**:100–101*F*, **6**:101
C_3 plants
anatomical modifications **2**:364*F*
energy requirements **2**:369*T*
grasses **1**:58–59, **1**:64*F*, **2**:654–655
historical atmospheric conditions **2**:367*T*
photosynthetic carbon reduction cycle **2**:361–362, **2**:362*F*
quantum yield **2**:369*F*
temperature effects **2**:369*F*
C_4 plants **2:361–381**
animal diversity impacts **2**:377–378
atmospheric carbon dioxide (CO_2) concentrations **2**:361
biogeography
C_4 annuals **2**:373*T*
general characteristics **2**:370–372
global biomes **2**:371*T*
latitudinal gradients **2**:370–372, **2**:370*F*
light intensity **2**:370
moisture availability **2**:372
principle life forms **2**:371*T*
salinity/salt tolerance **2**:371*T*, **2**:372–373
temperature parameters **2**:370–372
climate change effects **2**:378
ecological dominance **2**:377
economically significant C_4 plants **2**:362*T*
energy requirements **2**:369*T*
evolutionary development
carbon starvation hypothesis **2**:374–377
evolutionary centers **2**:375*F*
evolutionary lineages **2**:369*F*, **2**:373, **2**:374*T*, **2**:375*F*, **2**:376*F*
molecular clock analysis **2**:373–374, **2**:376*F*
polyphyletic origins **2**:373
taxonomic diversity **2**:373
exotic C_4 plants **2**:379–381, **2**:380*F*
grasses **2**:378–379, **2**:379*F*
historical atmospheric conditions **2**:367*T*
land-use change impacts **2**:378–379, **2**:379*F*
photosynthetic mechanisms
anatomical modifications **2**:364, **2**:364*F*
basic concepts **2**:362–364, **2**:363*F*
biochemistry **2**:363–364
C_3 photosynthetic carbon reduction cycle **2**:361–362, **2**:362*F*
diffusion barriers **2**:365–366
photorespiration **2**:361–362, **2**:362*F*
requirements **2**:363*T*
schematic diagram **2**:363*F*
single-cell C_4 photosynthesis **2**:364–365, **2**:365*F*
physiological significance
atmospheric carbon dioxide (CO_2) effects **2**:366–368, **2**:366*F*, **2**:368*F*, **2**:369*F*
energy requirements **2**:369*T*
historical atmospheric conditions **2**:367*T*
nitrogen-use efficiency **2**:367*T*, **2**:368–370
photorespiration **2**:366–368, **2**:366*F*
quantum yield **2**:369*F*
temperature effects **2**:367, **2**:369*F*
water-use efficiency **2**:367*T*, **2**:368–370
quantum yield **2**:369*F*
research summary **2**:381
temperature effects **2**:369*F*
tropical deforestation effects **2**:379–381, **2**:380*F*
carnivorous plants
benefits **6**:45
fitness costs **6**:45
general discussion **6**:44–45
prey capture mechanisms **6**:44–45, **6**:44*T*
wetland ecosystems **7**:395–397, **7**:396*F*
Central America **2**:25*T*
characteristics and classifications **6**:56
community evolution **6:65–75**
abundance patterns **6**:73, **6**:74*F*
allocation strategies **6**:68–69
body size and shape influences **6**:67–68, **6**:68*F*
climatic influences **6**:69
coevolutionary relationships **6**:69–70
competition **6**:67
equilibrium theories **6**:71–72
evolutionary development **6**:65–67, **6**:66*F*
extinction rates **6**:73*F*
latitudinal gradients **6**:72–73
niche diversification **6**:67
origination rates **6**:71–72, **6**:72*F*
physiological processes **6**:69
plant–animal interactions **6**:43–44
reproductive strategies **6**:68–69
research background **6**:65
research summary **6**:73–74
stability and resilience **6**:71
succession **6**:70–71
conservation practices **6:76–89**
anthropogenic impacts **6**:79–84
Ascension Island **6**:78
biodiversity distribution and loss **6**:79
biodiversity loss responses **6**:63, **6**:85–87
botanical gardens
ex situ conservation programs **6**:513, **6**:513–514, **6**:514
Global Strategy for Plant Conservation (GSPC) **6**:514–515, **6**:515
International Agenda for Botanic Gardens in Conservation **6**:514
Caribbean region **6**:83
crop wild relatives (CWRs) **6**:77
historical perspective **6**:76–79
hotspots **6**:79, **6**:82*T*
importance **6**:76
international agreements **6**:76–79, **6**:80–81*T*, **6**:509
living dead **6**:81
mosses **6**:83
restoration
dominance–diversity curve **6**:440*F*
limitations **6**:440–442, **6**:441*F*
single-species introductions **6**:442–443
species richness improvements **6**:442
succession **6**:443
vegetation management **6**:439–442, **6**:440*F*
salt water plants **6**:77
threatened species **6**:84–85
vascular plant diversity **6**:82*T*
conservation products **4**:321*T*
crop plant domestication **2:657–664**, **4:279–292**
basic concepts **4**:279
conscious versus unconscious selection **2**:660
crop maintenance **2**:660–661
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
evolutionary development
African Savanna belt **2**:659

East Asia region **2**:658–659
fertile crescent region **2**:658
founder crops **2**:658
pre-Columbian America **2**:659
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
growing purposes **2**:661
high care-level plants **4**:289*T*
historic land-use patterns **4**:570, **4**:570*F*
human–crop coevolution **2**:657, **4**:570, **4**:570*F*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
native crops **2**:662
reproductive isolation **4**:281*F*, **4**:285
reproductive systems
characteristics **2**:659–660
self-pollination **2**:659–660
vegetative propagation **2**:660
research summary **4**:290–291
sowing and reaping impacts **2**:661–662
traditional agriculture **2**:657–658
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
crop wild relatives (CWRs) **6**:77
crossbreeding programs **1**:640
cultivated plants/cultivars **5**:551, **5**:553*T*
defense systems **2:426–433**
animal self-medication **2**:432
biotic defenses
ants **2**:428–429
characteristics **2**:428–429
endophytic fungi **2**:429
mites **2**:429
predators/parasitoids **2**:429
chemical defenses
alkaloids **2**:426*T*, **2**:427
background information **2**:426–427
nonprotein amino acids **2**:426*T*, **2**:427
phenolic compounds **2**:426–427, **2**:426*T*
ricin **2**:426*T*, **2**:427–428
saponins **2**:426*T*, **2**:427
terpenes **2**:426*T*, **2**:427
human uses
background information **2**:431–432
food spices **2**:432–433
hunting and fishing **2**:431–432
medicines **2**:432
phenological strategies **2**:429
physical/structural defenses
domatia **2**:429
spines and thorns **2**:428
toughness **2**:428
trichomes **2**:428
plant investments
assumptions **2**:429–430
cost–benefit analyses **2**:430
induced defenses **2**:430
plasticity **2**:431
research summary **2**:433
theoretical perspectives **2**:430–431
density–biomass thinning laws **6**:161
desert ecosystems
biological adaptations **2**:470–472, **2**:471*F*, **2**:475*F*
climates **2**:464, **2**:464*F*
species interactions **2**:474–475, **2**:475*F*
disease outbreaks **3**:411, **3**:411*F*
diversity–productivity relationship **2**:597–599, **2**:598*F*
domesticated plants
Amazon ecosystems **4:279–292**
basic concepts **4**:279
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
basic concepts **4**:279
crop plants **2:657–664**, **4:279–292**, **4**:570, **4**:570*F*
basic concepts **4**:279
domestication mechanisms **4**:280–281
environmental factors **4**:285–287, **4**:286*F*
genetic diversity opportunities **4**:280–281
genetic drift **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
domestication mechanisms **4**:279–281
environmental factors **4**:285–287, **4**:286*F*
farming practices **6**:567
genetic diversity opportunities **4**:280–281
genetic drift **4**:279–281, **4**:280*F*
high care-level plants **4**:289*T*
indigenous peoples **4**:281–285, **4**:282*F*, **4**:283*F*, **4**:284*F*
low care-level plants **4**:288*T*
management systems **4**:287–290, **4**:288*T*, **4**:289*T*, **4**:290*F*
reproductive isolation **4**:280, **4**:281*F*, **4**:285
research summary **4**:290–291
usefulness discovery strategies **4**:281–285, **4**:282*F*, **4**:284*F*
dynamic global vegetation models (DGVMs) **1**:64*F*, **2**:681–682
ecosystem function measurements **3**:85*F*, **3**:85*T*, **3**:86
edible plants **3:127–132**
food habits **3**:127–128, **3**:127*T*
geographic distribution **3**:130–131, **3**:131*T*
species diversity
abundance rankings **3**:129*T*
edible parts **3**:129–130
food usage **3**:130*T*
prolific species **3**:130*T*
research background **3**:128–129
taxonomic diversity **3**:128–129, **3**:129*T*
standards of consumption **3**:127–128
weeds **3**:131–132
elevational range shifts **4**:605*T*
El Niño Southern Oscillation (ENSO) events **3**:158–159
endangered species **3:205–215**
causal factors **3**:208–209
conservation implications **3**:212–215
endangerment patterns **3**:206–208, **3**:207*F*
global distribution patterns **3**:207*F*
legal protection mandates
International Plant Protection Convention **3**:209
U.S. Endangered Species Act (1973) **3**:209
occurrences **3**:205–206
rarity
causal factors **3**:208–209
characteristics **3**:206
consequences **3**:209
selected examples
Coleus forskohlii **3**:209–210, **3**:210*F*
Encephalartos longifolius **3**:211, **3**:213*F*
Fitzroya cupressoides **3**:210, **3**:211*F*
Pterostylis truncata **3**:211, **3**:213*F*
Sarracenia rubra ssp. *alabamensis* **3**:210–211, **3**:212*F*
species-environmental matching models **3**:211–212
end-Cretaceous mass extinction **5**:174
endemism **3**:232*T*, **5**:592*T*
enemy–victim interactions **2**:142–143
ephemeral plants **2**:464, **2**:464*F*, **2**:471–472
evolutionary development **2**:642, **6**:65–67, **6**:66*F*
extinction rates **4**:659*T*, **5**:367, **5**:368*T*
fire/fire regimes **3**:437–438
forest canopies **3:516–527**
characteristics **1**:449
climate change effects **3**:524
conservation considerations **3**:524
extinction vulnerability **3**:524
fragmentation effects **3**:524
future research areas
databases **3**:526

plants (*continued*)
ethnobotanical studies **3**:526
experimental field work **3**:526
general considerations **3**:524
herbaria **3**:524–526
monographs and inventories **3**:524
research questions and opportunities **3**:525*T*
growth habit diversity
characteristics **3**:523
nutrition acquisition and retention **3**:523
photosynthetic pathways **3**:523
habitat conversion effects **3**:524
habitat diversity
canopy microclimate **3**:521–522
host tree specificity **3**:523
spatial scale **3**:522–523
historical research **3**:517
nutrition acquisition and retention
ant-associated epiphytes **3**:523–524
atmospheric epiphytes **3**:524
bark epiphytes **3**:524
humus epiphytes **3**:523
tank/trash basket epiphytes **3**:523
plant categories
hemiepiphytes **3**:518
nonvascular epiphytes **3**:517–518
parasites **3**:518–519
vascular epiphytes **3**:517, **3**:518*T*
research summary **3**:526
taxonomic distributions
hemiepiphytes **3**:520–521, **3**:522*T*
nonvascular epiphytes **3**:520, **3**:521*T*
parasites **3**:521
vascular epiphytes **3**:519–520, **3**:519*T*, **3**:520*T*
forest management impact
early successional phases **1**:451
exotic species invasions **1**:451
general discussion **1**:451
fossil record **3**:686–687, **3**:687*F*
freshwater ecosystems
freshwaters **3**:571–572
freshwater sediments **3**:572*T*
functional diversity (FD) **3**:85*F*, **3**:85*T*, **3**:86
functional roles **6**:56–57
fungi-related diseases
crop pathology **3**:633, **3**:633*T*
disease triangle **3**:633, **3**:633*F*
forest pathology **3**:633, **3**:633*T*
genetic diversity **3**:665*F*
globalization effects **3:700–706**
biotic homogenization **3**:701
donor and receptor regions **3**:703
general discussion **3**:700
geographic distribution **3**:701–703, **3**:702*F*
historical perspective **3**:700–701, **3**:701*F*
impacts and implications **3**:705
selected examples
alligator weed **3**:703–704, **3**:704*F*
castor oil plant **3**:703, **3**:704*F*
general discussion **3**:703
giant hogweed **3**:705
Spanish Needle **3**:704–705, **3**:704*F*
St. John's wort **3**:704, **3**:704*F*
water hyacinth **3**:703, **3**:704*F*
shipping trade **3**:701*F*
global warming effects
increased leaf herbivory **4**:229
increased plant sclerophylly **4**:229
reduced plant transpiration **4**:229
tropical plant geographic expansion **4**:229
grasslands **4**:3, **6**:331–333
grazing effects **4:8–17**
direct versus indirect effects **4**:10
disturbance regimes **4**:9–10, **4**:10*F*
grazer-based ecosystem restoration and management
background information **4**:14–15
coral reef ecosystems **4**:16
forests **4**:15
grasslands **4**:14–15
invasive species **4**:16–17
lakes and ponds **4**:15–16, **4**:16*F*
herbivory **4**:8–9, **4**:9*F*
host chemistry effects **4**:13
paleo-patterns **4**:13–14
research summary **4**:17
spatiotemporal scale
between-habitat patterns **4**:11–12
functional role **4**:10–11
geographic patterns **4**:11, **4**:11*F*
microhabitat patterns **4**:12–13
within-habitat patterns **4**:12
gymnosperms **6**:59–60
hemiparasitism **4:70–78**
characteristics and classifications **4**:70–71
ecological advantages **4**:75
evolutionary development **4**:73–74
haustorium/haustoria **4**:74–75
host specificity **4**:75
mycorrhizal associations **4**:74
pests
crop production **4**:75–77
dodders **4**:70, **4**:77
mistletoes **4**:70, **4**:77
research summary **4**:77–78
root parasitic weeds **4**:76–77
phylogenetic relationships **4**:71–73, **4**:72*F*
physiological characteristics **4**:75
plant–plant interactions **4**:74
seed dispersal/germination **4**:74
herbaceous plants **5**:528–530, **5**:534*F*
herbaceous vegetation species richness **4:79–86**
control mechanisms
intermediate disturbance hypothesis **4**:79
intraspecific diversity **4**:79–80, **4**:80*F*
dominants and subordinates **4**:82–83, **4**:82*F*, **4**:84
general discussion **4**:79
high species richness effects **4**:81–82
mass ratio hypothesis **4**:80–81, **4**:81*F*
plant traits **4**:80
research summary **4**:84–85
species declines
filter effects **4**:84
founder effects **4**:84
transients **4**:83–84, **4**:83*T*, **4**:84
high-elevation Andean ecosystems **4:96–110**
background information **4**:96
geographic distribution **4**:96–97, **4**:97*F*
habitat characteristics **4**:96–97
number of species **4**:102*T*
páramo ecosystem
characteristics **4**:101–103, **7**:275–276
geographic distribution **4**:96–97, **4**:97*F*
northern Andean alpine cushion formations **4**:103
northern Andean dwarf *Polylepis* forest **4**:103
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98
plant diversity **4**:98–99, **4**:99*T*
representative species **4**:102*F*
subpáramo zone **4**:102–103
superpáramo zone **4**:103
true páramo zone **4**:103
physical characteristics and patterns **4**:97–98
plant diversity
endemism **4**:99–100, **4**:99*T*
origins and diversification **4**:100–101
species richness **4**:98–99
puna
central Andean dwarf *Polylepis* forest **4**:105
cushion bogs **4**:105
desert puna **4**:104–105
dry puna **4**:104
geographic distribution **4**:96–97, **4**:97*F*
humid puna **4**:104
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:103–104
plant diversity **4**:98–99
representative species **4**:104*F*
salar vegetation **4**:105
research summary **4**:107
southern Andean steppes
altitudinal zonation **4**:106–107, **4**:106*F*
bogs and seeps **4**:107
cushion plants **4**:106–107, **4**:106*F*
geographic distribution **4**:96–97, **4**:97*F*
latitudinal gradients **4**:106–107
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:105–107
plant diversity **4**:98–99
representative species **4**:105*F*
high southern latitude plant growth **6**:618
Holocene postglacial spread **2**:550, **2**:551–552*T*, **2**:552*F*
hot deserts **6**:334
hotspots **6**:79, **6**:82*T*

hybridization **4**:**166–176**
basic concepts **4**:166
ecological consequences
biological diversity **4**:171–173
community impacts **4**:173–174, **4**:173*F*
resistance abilities **4**:172–173
species richness **4**:173–174, **4**:173*F*
homoploid hybridization
general discussion **4**:166–167
hybrid speciation **4**:168
hybrid swarms **4**:167–168
introgression **4**:166–167
novelty generation **4**:167
novelty loss **4**:167
polyploid hybridization
ancient polyploidy **4**:168–169
angiosperms **4**:169
characteristics **4**:168
classifications **4**:168
diversification **4**:170
polyploid complexes **4**:169–170
recent polyploidy events **4**:170–171, **4**:171*F*, **4**:172*F*
tracheophytes **4**:169
research summary **4**:174
importance **6**:76
Important Plant Areas (IPAs) **7**:203–204
indicator species **4**:255–256, **4**:257*F*
insecticide resistance **4**:**298–307**
cross-resistance **4**:302
diagnostic assays **4**:302–303, **4**:302*F*
gene homology **4**:301
management strategies **4**:305–306
multiple resistance **4**:302
mutation recurrence **4**:301–302
nonpest species **4**:306–307
origins **4**:299–300
prevalence **4**:298–299, **4**:298*F*, **4**:299*F*
research background **4**:298
research summary **4**:307
resistance gene selection
determinant factors **4**:303
ecological influences **4**:303–304
fitness costs **4**:305
genetic influences **4**:303
operational influences **4**:304–305
resistance mechanisms
characteristics **4**:300
increased detoxification **4**:300, **4**:300*F*
target site alterations **4**:300–301, **4**:300*F*
transgenic plants **4**:307
introduced parasites–species decline relationships **2**:528*T*
introduced species
Australia **1**:76*T*, **1**:80
Brazil **1**:76*T*, **1**:82–83
extinctions **3**:410–411
forest management impact **1**:451
grazer-based ecosystem restoration and management **4**:16–17
impact predictions and evaluations **4**:355–356
impacts and distribution **4**:378
India **1**:76*T*, **1**:82
aquatic ferns (*Salvinia molesta*) **4**:346–347, **4**:347*F*
Asian laminarian kelp (*Undaria pinnatifida*) **4**:345–346, **4**:346*F*
background information **4**:345
cheatgrass (*Bromus tectorum*) **4**:347–348, **4**:348*F*
endangered species **3**:208
Fallopia japonica **4**:351–352, **4**:352*F*
impact predictions and evaluations **4**:355–356
Lantana camara **4**:352–353, **4**:353*F*
Mimosa pigra **4**:353–354, **4**:354*F*
mouse-ear hawkweed (*Hieracium pilosella*) **4**:348–349, **4**:349*F*
Pinus pinaster **4**:354–355, **4**:355*F*
prickly pear (*Opuntia*) **4**:350–351, **4**:351*F*
Spartina alterniflora **4**:349–350
pollinators **6**:144–145
South Africa **1**:76*T*, **1**:81
United Kingdom **1**:76*T*, **1**:79
United States **1**:75–76, **1**:76*T*
island biogeography
analytical methods
data analysis **4**:403–405
floristic checklist methods **4**:399
quantitative methods **4**:400
relevé method **4**:399
sampling methods **4**:399
data analysis
dendograph approach **4**:403*F*, **4**:404*F*, **4**:405
Hawaii Volcanoes National Park **4**:402*F*, **4**:403–405, **4**:404*F*, **4**:405*F*
improved display methods **4**:404*F*, **4**:405, **4**:405*F*
two-way table approach **4**:402*T*, **4**:404*F*, **4**:405
experimental sampling design
horizontal approach **4**:406
transect approach **4**:405–406, **4**:405*F*
vertical approach **4**:406
invasive species **3**:410–411
quantitative sampling methods
challenges **4**:400
count-plot method **4**:400, **4**:401–403
distance methods **4**:400–401, **4**:401–403
point-centered quarter method **4**:401
sample plot size **4**:401*F*, **4**:403
wandering quarter method **4**:401
relevé sampling method
basic concepts **4**:399
cluster analysis **4**:403*F*, **4**:404*F*
predefinied strata **4**:399
size determinations **4**:399, **4**:400*F*
species determinations **4**:399
species quantity estimations **4**:399–400
keystone species study **4**:447
known and estimated species **4**:384*T*
lake and pond ecosystems **4**:461*T*
landscape legacies **4**:525–527
Late Permian extinction event **5**:171
Late Triassic mass extinction **5**:172–173
latitudinal range shifts **4**:604*T*
living dead **4**:**590–598**
animal populations **4**:594–595
anthropogenic impacts **4**:594–595
conservation practices **6**:81
deforestation **4**:591–593, **4**:592*F*
extinct megafauna **4**:591–593, **4**:593*F*
food chains **4**:595–596
forest ecosystems **4**:590–591, **4**:591*F*
requirements **4**:593–594
restoration programs **4**:597
small plants **4**:594
wildland patches **4**:596–597, **4**:596*F*
mangrove ecosystems **4**:337
medicinal plants **1**:359, **5**:705–706, **6**:518–519
Mediterranean-type climate ecosystems
Australia **5**:213*T*
California **5**:213*T*, **5**:215
Cape Region (South Africa) **5**:213*T*
Chile **5**:213*T*, **5**:216–217
Mediterranean Basin **5**:213*T*
mites **1**:214–215
modern extinctions **2**:168*F*, **5**:368*T*
molecular level diversity **2**:635*T*
mutation rates **2**:628*F*
native crops **2**:662
Near East ecosystems
animal diversity **5**:469–470, **5**:469*F*
Ceratonia siliqua open forests **5**:480*F*, **5**:482–483
deserts **5**:480*F*, **5**:485–486
desert savannoid vegetation **5**:480*F*, **5**:486
dryland vegetation patterns **5**:484
flora **5**:480–481, **5**:481*T*
general discussion **5**:481–482
mangroves **5**:480*F*, **5**:487
maquis/forests **5**:480*F*, **5**:481–482
Mediterranean savannoid vegetation **5**:480*F*, **5**:483
Mt. Hermon montane forest **5**:480*F*, **5**:482
Mt. Hermon tragacanth vegetation **5**:480*F*, **5**:483–484
Pistacia lentiscus open forests **5**:480*F*, **5**:482–483
Quercus calliprinos woodlands **5**:480*F*, **5**:482
Quercus ithaburensis open forests **5**:480*F*, **5**:482
reed thickets **5**:480*F*, **5**:486
sand vegetation **5**:480*F*, **5**:486
semisteppe bathas **5**:480*F*, **5**:483
shrub–steppes **5**:480*F*, **5**:484–485
shrub–steppes with trees **5**:480*F*, **5**:485
Sudanian tree oases **5**:480*F*, **5**:486
swamps **5**:480*F*, **5**:486
synanthropic vegetation **5**:480*F*, **5**:486–487
wet salinas **5**:480*F*, **5**:486
Ziziphus lotus **5**:480*F*, **5**:483
nematodes **7**:430

plants (*continued*)
nitrogen deposition effects **5**:523–525, **5**:527*F*
nitrogen fixation **1**:560–561, **1**:562*F*
nomenclature code **5**:550–551, **5**:553*T*
nonseed gene banks
pollen **3**:650
somatic and zygotic plant embryos **3**:650–651
spores **3**:650
vegetative tissues **3**:651
nonvascular plants **5**:528, **6**:58–59, **6**:58*T*
North America **5**:564–565
nutrient uptake **3**:74–75*T*, **3**:78, **3**:90*T*
ornamental plants **1**:75–76
paleoecology **5:645–655**
data collection and interpretation **5**:645–646
goals and objectives **5**:645–646
Quaternary data **5**:645
research methodologies
age models **5**:652–653
dating uncertainties **5**:652
mapping intervals **5**:652–653
scaling factors **5**:646–649, **5**:648*F*, **5**:649*F*
space–time variations **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
spatial characteristics **5**:648*F*, **5**:649–650, **5**:649*F*, **5**:651*F*
taxonomic and numerical characteristics **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
temporal characteristics **5**:650–652, **5**:651*F*
temporal resolution **5**:652
zoom lens perspective **5**:647*F*, **5**:648*F*, **5**:649*F*, **5**:651*F*, **5**:653–654
parasites
hemiparasitism **4:70–78**
characteristics and classifications **4**:70–71
dodders **4**:70, **4**:77
ecological advantages **4**:75
evolutionary development **4**:73–74
haustorium/haustoria **4**:74–75
host specificity **4**:75
mistletoes **4**:70, **4**:77
mycorrhizal associations **4**:74
pests **4**:75–77
phylogenetic relationships **4**:71–73, **4**:72*F*
physiological characteristics **4**:75
plant–plant interactions **4**:74
research summary **4**:77–78
root parasitic weeds **4**:76–77
seed dispersal/germination **4**:74
parasite–host coevolution
resistance mechanisms **2**:144–145
specialist versus generalist herbivores **2**:145–146
tolerance mechanisms **2**:145
parasitic disease vectors **5**:672–673
past, recent, and future extinction rates **2**:168*F*
pathogens **6**:449–450
Phanerozoic diversity **3**:686–687, **3**:687*F*
pharmacologically active compounds **5**:711–713, **5**:712*T*
phenological shifts **5**:719*F*
phenology **6:103–108**
change consequences **6**:106–107
data characteristics **6**:103–104, **6**:103*F*
research challenges **6**:107
research evidence **6**:105–106
seasonal change effects **6**:103–104
temperature drivers **6**:104–105, **6**:105*F*, **6**:106*F*
phytotelmata **3**:494–495
pioneer guilds **1**:451
plant–animal interactions **6:39–55**
antagonistic interactions
carnivorous plants **6**:44–45
coevolution **6**:42–43
community and ecosystem effects **6**:43–44
evolutionary responses **6**:41–42
plant consumers **6**:39–41
carnivorous plants
benefits **6**:45
fitness costs **6**:45
general discussion **6**:44–45
nonprey guests **6**:45
prey capture mechanisms **6**:44–45, **6**:44*T*
chemical defenses
digestibility reducers **6**:41
plant secondary chemicals **6**:42
repellents **6**:42
toxins **6**:41
classifications **6**:39, **6**:39*T*, **6**:40*F*
evolutionary responses
chemical defenses **6**:41–42
compensatory growth **6**:42
escape defenses **6**:42
mechanical defenses **6**:42
tolerance mechanisms **6**:42
fruit and seed dispersal
ants **6**:50
beetles **6**:50
birds **6**:50
dispersal syndromes **6**:50, **6**:51*T*
disperser types **6**:50
evolutionary dead ends **6**:52
fish **6**:50
frugivores **6**:50–51
importance **6**:49–50
insects **6**:50
mammals **6**:50
plant adaptations **6**:50
reptiles **6**:50
seed predators **6**:51–52
species persistence **6**:61–62
vertebrates **6**:50
mutualism
cheating studies **6**:45–46
fruit and seed dispersal **6**:49–50, **6**:61–62
plant–pollinator interactions **2**:474–475, **2**:475*F*, **6**:47
plant-protecting ants/plant-feeding ants **6**:46
specialization **6**:45–46
research summary **6**:52
trophic levels **6**:40*F*
plant breeding **1:459–469**
commodity agriculture **1**:467
controversies **1**:459
crop homogeneity **1**:461*F*
exploration and trade **1**:465–466
gene pool concepts **1**:463–465, **1**:465*F*
genetic bottlenecks **1**:463
green revolution genetics **1**:467–468
legal frameworks **1**:466–467
modern scientific plant breeding **1**:460–461, **1**:461*F*
phenotypic variation **1**:461–463, **1**:462*F*
quality guidelines **1**:461*F*
reduced genetic diversity **1**:461–463
research summary **1**:468
plant community evolution **6**:67
plant-derived drugs and chemicals **6:129–139**
algae
domoic acid **6**:132, **6**:134*F*
green algal metabolites **6**:132–135, **6**:134*F*
kainic acid **6**:132, **6**:134*F*
cyanobacteria
anticancer agents **6**:130–131
cryptophycin **6**:131, **6**:131*F*
curacin A **6**:130–131, **6**:130*F*
dinoflagellates
brevetoxin A **6**:131–132, **6**:134*F*
characteristics **6**:131
ciguatoxin **6**:131, **6**:132*F*
Gymnodinium breve **6**:131–132, **6**:134*F*
maitotoxin **6**:131, **6**:133*F*
future outlook **6**:138
herbal remedies
Echinacea **6**:137–138, **6**:138*F*
ginseng **6**:137, **6**:138*F*
higher plants
camptothecin **6**:135–136, **6**:135*F*
paclitaxel **6**:135, **6**:135*F*
marine plants
algae **6**:132
cyanobacteria **6**:130–131
dinoflagellates **6**:131
research background **6**:129–130
terrestrial plants
Digitalis sp. **6**:136–137, **6**:137*F*
herbal remedies **6**:137
higher plants **6**:135
steroidal glycosides **6**:136–137, **6**:137*F*
tropane alkaloids **6**:136, **6**:136*F*
plant diversity–grazing intensity relationship **1**:703–704, **1**:704*T*
plant–human interactions **6**:57–58
plant–insect associations **1**:360, **4**:35–36, **4**:36*T*, **5**:354–355, **6**:48
plant invasions **6:90–102**
basic concepts **6**:90–92, **6**:91*F*
current extent **6**:93–95, **6**:94*F*
dispersal pathways **6**:90–92, **6**:91*F*
general discussion **6**:90

historical perspective **6**:92–93
invasibility **6**:97–98, **6**:99*F*
invasion process
dispersal **6**:95–96
establishment and naturalization **6**:96
spread rates **6**:96–97
stages **6**:95–96, **6**:95*F*
invasiveness **6**:97
management strategies
biotic homogenization **6**:100–101*F*, **6**:101
ecological management **6**:101
potential conflicts **6**:101–102
prevention methods **6**:101
removal methods **6**:101
modeling approaches **6**:99–101
plant trait correlations **6**:98–99
plant–soil interactions **6:109–128**
basic concepts **6**:109–111, **6**:110*F*, **6**:111*F*
conservation and management practices
anthropogenic impacts **6**:125
atmospheric carbon dioxide (CO_2) concentrations **6**:126, **6**:127*F*
ecosystem functions **6**:126–127
exotic species invasions **6**:125
general discussion **6**:125
nutrient pollution **6**:125–126
restoration programs **6**:125
feedback processes **6**:111*F*
geographic patterns **6**:123, **6**:123*T*
litter deposition **6**:119–121, **6**:120*T*
nutrient acquisition
individual plants **6**:116–117, **6**:117*F*, **6**:118*F*, **6**:119*F*
root growth patterns **6**:119*F*
spatial patterns **6**:119*F*
stand scale **6**:117–118
plant–microbiota competitive interactions **6**:118–119
plant tissue chemistry **6**:119–121, **6**:120*T*
regional characteristics **6**:123, **6**:123*T*
root–soil interface
large-scale processes **6**:123
mesoscale processes **6**:116–117
nutrient and water uptake **6**:113–115, **6**:114*F*
reduction–oxidation potential **6**:115*T*
rhizosphere **6**:111–113, **6**:112*F*, **6**:113*T*
small-scale processes **6**:111–113
soil structure **6**:115–116
schematic diagram **6**:110*F*
soil biota **6**:121–123, **6**:122*F*
within-biome variations **6**:123–125, **6**:124*F*
plant tolerance–grazing relationship **1**:708, **2**:428–429
restoration efforts
dominance–diversity curve **6**:440*F*
limitations **6**:440–442, **6**:441*F*
single-species introductions **6**:442–443
species richness improvements **6**:442
succession **6**:443
vegetation management **6**:439–442, **6**:440*F*
return-on-investment (ROI) analysis **4**:187*T*
river ecosystems **6**:477–478
RNA editing **5**:573–574*T*, **5**:575–576, **5**:575*F*
root activity **3**:90*T*
salt water plants **6**:77
savannas **6**:331–333
seasonal rainfall **2**:464–466
seedless vascular plants **6**:59
slash-and-burn systems **6**:556
Southeast Asia **1**:514, **1**:514*T*
speciation–extinction rate relationship **5**:439–440, **5**:440*F*
species–area relationships (SARs) **2**:115, **3**:116, **3**:117
species estimates **3**:710*T*
species richness gradients **1**:499*T*
spines and thorns **1**:708
storage ecology
growth cycles **7**:24–25
lipid composition **7**:25–26
predator protection **7**:26
seed storage **7**:24–25
Succulent Karoo **1**:485, **1**:486*T*, **1**:487, **1**:487*T*, **1**:488–489
temperate/cold-desert shrublands **6**:333–334
temperature tolerance ranges **4**:115
termite effects **4**:431–432, **4**:431*F*
thermal ecosystems **4**:112
threatened species **1**:278*T*, **7**:201*T*
Tilman biomass study **1**:386–388, **1**:387*F*
tourism industry **7**:223
transgenic plants **1**:467–468, **4**:307
transpiration **3**:90*T*
tropical rain forests **6:313–321**
conservation efforts **6**:320
economic value and benefits **6**:319–320
environmental conditions **6**:313
explanatory theories **6**:318–319
geographic regions **6**:314*T*, **6**:316*T*
life-form varieties **6**:313–318, **6**:315*T*, **6**:316*T*
lowland tropical rain forests **6**:317*T*
species assessments **6**:313, **6**:314*T*
species richness trends **6**:317*T*
tundra ecosystems **6**:330–331, **6**:331*F*
ultraviolet radiation effects **7**:298–299
wetland ecosystems **7**:388–389, **7**:388*F*, **7**:391–393, **7**:398*F*
see also endemism; flowering plants; vascular plant life-forms
plasmids **3**:652*T*
Plasmodiophora brassicae **3**:631, **6**:297
plasmodiophorids **6**:296–297
Plasmodiophoromycota (endoparasitic slime molds) **3**:631
Plasmodium **6**:297
Plasmodium relictum **2**:528*T*, **2**:529–530
Plasmodium relictum capristranoae **4**:362
Plasmopara viticola **3**:630
plastids **3**:330–331
Platanista gangetica **3**:192, **5**:77*T*
Platanista minor **3**:192
Platanus **5**:621*F*
plate tectonics **1**:580–583, **1**:584*F*, **3**:688–689
Plathymenia foliosa **1**:161–162
Platichthys flessus **4**:465
Platnick biodiversity model **7**:91, **7**:91*F*
Plato **3**:98–99, **7**:100
Platonia insignis **4**:286–287, **4**:288*T*
Platybelodon **5**:224*T*
Platycerabos **4**:724–725*T*
Platycladus **1**:267
Platydesmida **5**:423*F*
Platygastroidea **4**:178
Platygonus **4**:723–725, **4**:724–725*T*, **4**:727*T*
Platyhelminthes **7:437–469**
adaptability **4**:393
Antarctica **1**:184*T*
defining characteristics **7**:438
DNA studies **5**:662–663
evolutionary development **2**:645
excretory system/osmoregulation **7**:460–461
feeding and digestion control **7**:460
fossil record **3**:684*F*
general characteristics **7**:437
global diversity **4**:385*T*
host associations **7**:441*F*, **7**:467–468
larval stages **7**:464–465*F*
life cycle **7**:462–466, **7**:462*F*, **7**:463*F*, **7**:464–465*F*
major groups
Neodermata
Aspidogastrea **7**:439–440*T*, **7**:441*F*, **7**:450–451*F*, **7**:451
characteristics **7**:445–447
Digenea **7**:439–440*T*, **7**:450–451*F*, **7**:451
Monogenea **7**:439–440*T*, **7**:446–447
Monopisthocotylea **7**:439–440*T*, **7**:441*F*, **7**:447, **7**:448–449*F*
neodermis **7**:445–447, **7**:446*F*
number of species and families **7**:439–440*T*
Polyopisthocotylea **7**:439–440*T*, **7**:441*F*, **7**:448–449*F*, **7**:449
species diversity **7**:439–440*T*
Trematoda **7**:439–440*T*, **7**:450–451*F*, **7**:451
number of species and families **7**:439–440*T*
species diversity **7**:438–440, **7**:439–440*T*
Turbellaria
Bothrioplanida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443–444
Catenulida **7**:439–440*T*, **7**:440, **7**:441*F*, **7**:442–443*F*
Dalyellioida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Fecampiida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
Haplopharyngida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
Kalyptorhynchia **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Lecithoepitheliata **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443

Platyhelminthes (*continued*)
Macrostomida **7**:439–440*T*, **7**:440–443, **7**:441*F*, **7**:442–443*F*
number of species and families **7**:439–440*T*
Polycladida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
Prolecithophora **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
Proseriata **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Temnocephalida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Tricladida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
Typhloplanoida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444–445
marine ecosystems **5**:50*T*
medical importance **7**:468
Neodermata
Aspidogastrea **7**:439–440*T*, **7**:441*F*, **7**:450–451*F*, **7**:451
Cestoda
Amphilinidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Gyrocotylidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
species diversity **7**:439–440*T*, **7**:455–457
characteristics **7**:445–447
Digenea
Allocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Azygioidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Bivesiculoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Brachylaimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Bucephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
characteristics **7**:450–451*F*
Clinostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Cyclocoeloidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Diplostomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452
Echinostomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:452–453
Gorgoderoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Gymnophalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Haploporoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Haplosplanchoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Hemiuroidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Heronimoidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Lepocreadioidea **7**:439–440*T*, **7**:450–451*F*, **7**:453
Microphalloidea **7**:439–440*T*, **7**:450–451*F*, **7**:453–454
Microscaphidioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Monorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Opisthorchiodea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Paramphistomoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Plagiorchioidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Pronocephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
Schistosomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
species diversity **7**:439–440*T*, **7**:451
Transversotrematoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454–455
DNA studies **5**:662–663
Eucestoda
Bothriocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Caryophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cathetocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:457
Cyclophyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:457–458
Diphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Diphyllobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Haplobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Lecanicephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Litobothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:458
Nippotaeniidea **7**:439–440*T*, **7**:456–457*F*, **7**:458–459
Proteocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Rhinebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Spathebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetrabothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetraphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Trypanorhyncha **7**:439–440*T*, **7**:456–457*F*, **7**:459–460
Monogenea **7**:439–440*T*, **7**:446–447
Monopisthocotylea
Capsalidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Dactylogyridea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Gyrodactylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
Lagarocotylidea **7**:439–440*T*, **7**:447, **7**:448–449*F*
neodermis **7**:445–447, **7**:446*F*
number of species and families **7**:439–440*T*
Polyopisthocotylea
Chimaericolidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Diclybothriidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Mazocraeidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Polystomatidea **7**:439–440*T*, **7**:448–449*F*, **7**:449–451
species diversity **7**:439–440*T*
Trematoda **7**:439–440*T*, **7**:450–451*F*, **7**:451
nervous system **7**:460
number of species **4**:371*T*, **5**:669–670*T*
ontogeny **7**:462–466, **7**:464–465*F*
organ systems **7**:442–443*F*
phylogenetic relationships **7**:434*F*, **7**:441*F*, **7**:466–467
Proterozoic–Cambrian transition **3**:396*F*
reproductive system **7**:461–462
Platypezidae **3**:489–491*T*
Platypodinae **1**:328
Platyprepia virginalis **5**:396*F*
platypus
convergent evolution **4**:701*T*
hemoglobin evolution **2**:55*F*, **2**:55*T*
threat percentage **3**:195
platys **3**:467
Platyspiza crassirostris **1**:27–28
Platystomatidae **3**:489–491*T*
Plautus alle **1**:617
Plaxhaplous **4**:727*T*
playas **2**:461, **2**:464*F*, **7**:369*T*
Plecoptera
major divisions and characteristics **4**:308–312, **4**:309*T*
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
wetland ecosystems **7**:389–391
Plectus antarcticus **7**:429
Plegadis falcinellus **5**:476–477
pleiotropy **2**:720
Pleistocene
Antarctic glacial history **1**:179–180
Australian vertebrate species **5**:363–364, **5**:364*F*
biotic interchange episodes **2**:545*F*, **2**:552*F*
boundary age **5**:1*T*
C_4 plants **2**:376*F*
crustaceans **2**:407
desert development **2**:466
glacial cycles **1**:583, **1**:585*F*, **1**:586*F*, **2**:466, **4**:721–725
Levantine biota **5**:461–462
mammal ages **2**:583*F*
mammal extinctions
major die-offs **3**:189
North America **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:727*F*
mass extinctions **4**:656–657, **4**:659*T*
megaherbivores **5:223–239**
extinctions **5**:225*T*, **5**:231–233, **5**:232*F*
grazing impacts **5**:228
occurrences and characteristics **5:223–239**
oxygen-18 ($\delta^{18}O$) measurements **1**:557*F*
termite species **4**:410*T*
terrestrial plant communities **6**:74*F*

plesiomorphy **2**:33*F*, **2**:34–35, **2**:35*F*
Plesiorycteropus **4**:728–729
Plesiorycteropus madagascariensis **4**:657, **5**:365
Plesippus **4**:724–725*T*
Plethodon
  extinction rate–speciation rate relationship **1**:431*F*
  maximum likelihood estimation **1**:431*F*
  phylogenetic reconstruction **1**:431–432, **1**:431*F*
Plethodon larselli **4**:564
Plethodontidae **7**:389–390
Pleurochrysis carterae **1**:474*F*
Pleurodira **6**:400–401
Pleuronectes platessa **5**:315
Pleuronema **6**:295
Pleuropogon sabinei **1**:236
Pleurothallis **6**:316*T*
Pleurotus ostreatus **3**:637*T*
Plicosepalus acaciae **5**:483
Pliny, the elder **3**:320–321
Pliocene
  biotic interchange episodes **2**:545*F*
  boundary age **5**:1*T*
  C_4 plants **2**:376*F*
  crustaceans **2**:406
  Levantine biota **5**:461
  mammals **2**:583*F*, **4**:722–725, **4**:724–725*T*
  megaherbivores **5**:224*T*
  oxygen-18 ($\delta^{18}O$) measurements **1**:557*F*
  South America **3**:107, **3**:109
  temperate forests **7**:122
  terrestrial plant communities **6**:74*F*
  vegetation–climate interactions **2**:683
Ploceidae **5**:502, **5**:502*T*
Plodia interpunctella **5**:401–402
Plotosus lineatus **3**:464
plovers **1**:710, **5**:476–477
Pluchea indica **4**:365
Plumariidae **4**:178
Plumbaginaceae **5**:11*T*
plums **2**:662–663
Plusioporus setiger **5**:424–425
Plutella xylostella **4**:301
Pluvialis **1**:601
Pluvialis dominica **1**:605
Pluvialis fulva **5**:311–312
Pneumocystis carinii **3**:634*T*, **5**:660*T*, **5**:670, **6**:297
Pneumocystis jiroveci **2**:644, **6**:297
Pneumocystis pneumonia **3**:634*T*
Pneumoroidea **3**:733, **3**:733*F*
Poa alopecurus **4**:106–107
Poa arctica **1**:239*F*
Poa botryoides **1**:269
Poa bulbosa **1**:269, **1**:270
Poaceae **1**:486*T*, **2**:371*T*, **2**:374*T*, **3**:129*T*, **3**:234–235, **4**:347–348
poaching **1**:516–517, **1**:519*F*, **1**:521, **1**:521*F*, **6**:444–445, **6**:499
Poales **2**:374*T*
Poa pratensis **6**:433, **6**:434*F*
Poa spania **4**:194*T*
pocket gophers **2**:584*F*, **7**:229*T*, **7**:432
pocosin **7**:369*T*
Podiceps gallardoi **7**:186
Podiceps taczanowskii **7**:183–184
Podilymbus gigas **1**:621
Podilymbus podiceps **1**:621
Podoaceae **1**:260*T*
Podocarpus **2**:25–26, **5**:618*F*, **7**:119, **7**:275
Podophrya **6**:295
podophyllotoxin **5**:712*T*
Podophyllum peltatum **5**:712*T*
Podoviridae **5**:130*F*, **5**:134*T*, **5**:135
Poecilia reticulata **2**:140, **2**:718
poeciliid fish **7**:348, **7**:348*F*
Poeciliopsis lucida **2**:513
Poeciliopsis monacha **2**:513
Poecilochaetidae **7**:417*F*
poetry **4**:650
Pogonathria squarrosa **1**:67
Pogonatum alpinum **1**:236*F*
Pogonomyrmex occidentalis **1**:709
Pogonophora
  classifications **4**:381*T*, **7**:408
  marine ecosystems **5**:50*T*
  number of species **5**:669–670*T*
  sampling methods **4**:381*T*
point diversity **2**:650, **6**:534
poison arrows **2**:431–432
poison dart frogs **3**:166–167, **5**:711
poisoning **7**:185
poisonous mushrooms **3**:634, **3**:635*T*
Poisson distribution **6**:158–160, **6**:160*F*
Poland
  botanical gardens **6**:519
  vegetation surveys **3**:341–344*T*
Polanyi, Karl **6**:383
polar Asia
  characteristics **1**:273–274
  tundra **1**:262–263*T*, **1**:274–275
polar bears
  adipose tissue **7**:33–35, **7**:34*F*
  characteristics **5**:68
  climate change effects **5**:89–91
  evolutionary history **5**:70
  taxonomy **5**:65*T*
polar ice **3**:560–562, **3**:561*T*
polar regions
  C_4 plants **2**:371*T*
  Census of Marine Life project **2**:20–21
  conservation status **3**:171–173*T*
  current global land cover **4**:557*F*
  global agriculture impacts **3**:428*T*
  global biodiversity trends **4**:563*F*
  pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
  protected area systems **6**:492*T*, **6**:499*T*
polecats
  skull size comparisons **1**:692*F*
  species diversity **1**:685
political subjugation **6**:222–223
pollack **6**:422*T*
pollen feeders **4**:181
pollen gene banks **3**:650, **4**:319–320, **4**:321*T*
pollen studies *see* paleoecology
Pollicipes polymerus **6**:723*T*
pollination **6:140–146**
  animal pollination **5**:354–355
  basic concepts **6**:140
  biodiversity–ecosystem services relationship **1**:348, **3**:347
  biodiversity-friendly farming **2**:544
  biodiversity threats **2**:80
  competition **6**:49
  conservation biology
    conservation programs **4**:319–320, **4**:321*T*
    management strategies **6**:145–146
    plant–pollinator mutualism threats **6**:145
  mutualistic partnerships **2**:328–329, **4**:447
  pollination syndromes **6**:48–49, **6**:49*T*
  pollinators
    agriculture
      economic importance **6**:144–145
      genetic engineering effects **6**:145
      introduced species **6**:144–145
      plant–pollinator mutualism threats **6**:145
      supporting services **1**:360, **5**:354–355
    coevolutionary interactions **6**:143
    invertebrates
      beetles **6**:49*T*, **6**:140–141, **6**:141*T*
      butterflies and moths **6**:49*T*, **6**:141–142, **6**:141*T*
      flies **6**:49*T*, **6**:141, **6**:141*T*
      flower visitation **6**:49*T*, **6**:140–141
      Hymenoptera **6**:141*T*, **6**:142
    keystone species study **4**:447
    native vegetation **1**:419, **1**:419–420, **1**:419*F*, **1**:420*F*
    natural systems
      cheaters **6**:143–144
      deception **6**:144
      obligate and facultative mutualisms **6**:143
      pollen and nectar robbers **6**:143–144
    organism spillover **3**:13–14
    plant–pollinator interactions
      angiosperm radiation hypothesis **6**:48
      basic concepts **6**:47
      cheating studies **6**:47
      competition **6**:49
      desert ecosystems **2**:474–475, **2**:475*F*
      floral nectar spurs **6**:47–48
      insect pollination **6**:48
      pollination syndromes **6**:48–49, **6**:49*T*
      reproductive isolation **6**:47
      species persistence **6**:61–62
      supporting services **1**:360, **5**:354–355
    pollinator classes **6**:140–141, **6**:141*T*
    vertebrates
      birds **6**:49*T*, **6**:141*T*, **6**:143
      mammals **6**:49*T*, **6**:141*T*, **6**:142–143
pollock **3**:466–467
pollution
  air pollution **1:136–147**
    biodiversity impacts
      air quality guidelines **1**:145–146
      birds **1**:139
      evidentiary research **1**:139–141
      evolutionary responses **1**:138–139
      lichen biodiversity **1**:141–142, **1**:142*F*

pollution (*continued*)
long-term and short-term effects **1**:137–138
mammals **1**:139
nitrogen deposition effects **1**:142–144, **1**:143*F*, **1**:146*F*
ozone effects **1**:144
smelting industry impacts **1**:140–141, **1**:141*F*, **1**:142*F*
species interaction effects **1**:139, **1**:140*F*
surface water acidification **1**:144–145
synergistic interactions **1**:138
depositional and environmental pathways **1**:137
environmental impact **1**:136–137, **1**:137*T*
global impacts **1**:145–146, **1**:146*F*
historical perspective **1**:136
lower atmosphere **1**:289–290
major sources **1**:136–137, **1**:137*T*
summary discussion **1**:146
biodiversity impacts **7**:203*F*
biogeochemical cycles **6:147–156**
carbon biogeochemical cycle perturbations
atmospheric carbon dioxide (CO_2) **6**:147–149, **6**:148*F*
climate change effects **6**:147–149
coastal/marine ecosystems **6**:150–151
ecosystem alterations **6**:149–150
forest ecosystem alterations **6**:149–150, **6**:150*T*
freshwater ecosystems **6**:151
historical perspective **6**:147
nitrogen biogeochemical cycle perturbations
acid rain **6**:151–152
aquatic ecosystems **6**:152, **6**:153*F*
forest ecosystems **6**:152
nitrogen oxides (NO_x) **6**:151–152
nitrogen saturation **6**:152
troposheric ozone precursors **6**:152–154
sulfur biogeochemical cycle perturbations **6**:154–155, **6**:154*F*
summary discussion **6**:155
trace pollutants **6**:155
bioremediation **6**:594
chemical pollution **3**:165–166
ecosystem services modeling **5**:351
endangered mammals **3**:192–193
endangered reptiles **3**:218
environmental impacts **3**:281–282, **7**:203*F*
forest ecosystems **3**:534–535
freshwater ecosystems
anthropogenic impacts
acidification **3**:572–574, **3**:577–578
agricultural and industrial chemicals **3**:575–576, **3**:580–581, **3**:581*F*
areal reductions **3**:573–574
eutrophication **3**:575*F*, **3**:576–577, **3**:576*F*
hydrological cycle changes **3**:574–575
lake eutrophication **3**:581–582
organic pollution **3**:575, **3**:575*F*, **3**:580–581, **3**:581*F*
pollution **1**:512
restoration and conservation programs **3**:574
invertebrates **4**:377
Southeast Asia **1**:512
habitat degradation
air pollution
acid rain **3**:407, **3**:407*F*
lichen tolerance **3**:407
ozone depletion **3**:407
smog **3**:407
ultraviolet radiation **3**:407
characteristics **3**:406
pesticides **3**:406
water pollution **3**:406–407
habitat degradation and pollution **2**:531
industrialization impacts **4**:145
infectious diseases **2**:531
lake and pond ecosystems **3**:181, **3**:181*F*, **4**:465–466
lower atmosphere **1**:289–290
marine ecosystems **5**:61, **6**:150–151
marine sedimentary habitats **5**:113*F*, **5**:114
soil conservation **6**:594
trace pollutants **6**:155
water quality **4**:158, **5**:351
polyandry **6**:573
polybrominated diphenyl ethers (PBDEs) **5**:71–72
polycentric governance systems **2**:164–165
Polychaeta/polychaete worms
classifications **4**:381*T*, **7**:408
coral reef ecosystems **2**:332*T*, **2**:333*T*
endangered species **3**:202*T*
fossil record **3**:681*F*, **3**:685*F*
geographic distribution **7**:418
hydrothermal vent ecosystems **7**:329, **7**:330
mangrove ecosystems **5**:20*T*
marine environments **4**:382*F*, **5**:106*F*
morphology
anterior ends **7**:410*F*, **7**:412*F*
chaetae **7**:409–410, **7**:414*F*
circulatory system **7**:415
digestive tract **7**:415
general characteristics **7**:408–409
gills **7**:415
gonads **7**:415–416
head **7**:409
musculature **7**:414–415
nephridia **7**:415
nervous system **7**:410–411
posterior ends **7**:411*F*
pygidium **7**:409
segments **7**:409, **7**:413*F*
sense organs **7**:411–414
septa **7**:415
number of species **4**:371*T*, **4**:384*T*, **5**:669–670*T*, **7**:408
phylogenetic and taxonomic relationships **7**:416–419, **7**:417*F*
reproductive biology **7**:416
sampling methods **4**:381*T*
sandy beaches **2**:131–133
polychaete worms **6**:25
polychlorinated biphenyls (PCBs) **3**:118–119, **3**:218, **3**:581, **5**:71–72, **6**:155
polychloroterpenes **5**:693
Polycladida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
polyculture cropping systems **1**:375–376, **7**:230–231
polycyclic aromatic hydrocarbons (PAHs) **3**:118–119
Polydesmida **1**:247, **5**:423–424, **5**:423*F*, **5**:430
Polydesmus **5**:423*F*
Polydesmus collaris **5**:424–425
polydnaviruses **5**:678
Polydora uncata **3**:364
Polygalaceae **5**:623*F*, **6**:315*T*
Polygala negevensis **5**:484–485
Polygonaceae **2**:371*T*, **3**:129*T*, **4**:351–352, **5**:489
Polygonales **2**:374*T*, **5**:622*F*
Polygonanthus amazonicus **1**:163
Polygonum **1**:270
Polygonum viviparum **1**:236*F*
Polykrikos **6**:294–295
Polylepis **4**:103, **4**:105, **7**:275
polymerase chain reaction (PCR) technique
dinitrogen-fixing (N_2-fixing) microorganisms **2**:611
DNA polymorphism **3**:663
double-stranded DNA (dsDNA) analysis **5**:262–264, **5**:263*F*, **5**:264*T*
drug development **1**:394
ecogenomic sensors **2**:694–697, **2**:694*F*
genetic analysis **1**:389, **1**:593
microbial biodiversity measurement **5**:254, **5**:260, **6**:345
Polymixiomorpha **3**:466
Polynesian ground dove **7**:184
Polynesian plant species **6**:82*T*
Polynoidae **7**:417*F*
Polyommatinae **1**:652
Polyopisthocotylea
Chimaericolidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Diclybothriidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Mazocraeidea **7**:439–440*T*, **7**:448–449*F*, **7**:449
Polystomatidea **7**:439–440*T*, **7**:448–449*F*, **7**:449–451
Polyphaga **1**:326, **1**:327*T*
polyphagy **5**:394
polyphyly **7**:91–93, **7**:92*F*
Polyplacophora
classifications **4**:381*T*
physical characteristics **5**:375–376, **5**:375*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
Polyplaston **6**:295–296
polyploid hybridization
ancient polyploidy
angiosperms **4**:169
prevalence **4**:168–169
tracheophytes **4**:169

characteristics 4:168
classifications 4:168
diversification 4:170
polyploid complexes 4:169–170
recent polyploidy events 4:170–171, 4:171*F*, 4:172*F*
Polypodium polypodioides 3:517
Polypompholyx 6:44*T*
Polyrachis simplex 5:463*T*
Polyrhachis sokolova 5:13
Polyscias 3:522*T*
Polysiphonia 3:359
Polysphondylium 6:290*F*
Polystomatidea 7:439–440*T*, 7:448–449*F*, 7:449–451
Polytoca 1:70
polyunsaturated fatty acids 7:25–26, 7:32–33
Polyxenida 5:423–424, 5:423*F*
Polyxenus lagurus 5:425–426, 5:428
Polyzonium germanicum 5:426
Pomacanthus imperator 3:471*F*
pomace flies 1:28–29, 3:489–491*T*, 3:493
Pomarea dimidiata 7:184
Pomatias olivieri 3:674*F*
Pomatomus saltatrix 3:469
Pomatoschistus minutus 3:478*F*
pompano 3:467–471
pompeii worm 7:329
pom-pon mushrooms 3:637*T*
ponderosa pine trees 1:144, 6:150*T*, 6:152–154, 6:712*F*
pond snails 4:374
Pongamia pinnata 1:277
Pongidae 4:693
Pongo abelii 6:256–258*T*
Pongo pygmaeus morio 6:256–258*T*
Pongo pygmaeus pygmaeus 6:256–258*T*
Pongo pygmaeus wurmbii 6:256–258*T*
Pontania gracilis 4:173–174
Pontania pomum 4:173–174
Pontic steppe 1:262–263*T*
Pontonia pinnophylax 6:547–548
Pontoporia blainvillei 5:77*T*
ponyfish 5:122*F*, 5:123
poplars
biofuel crop production 1:529*T*, 1:531–532, 1:544*F*
boreal forests 1:628–630, 1:630*T*
energy use impacts 3:261
North America 7:118
plant tissue chemistry 6:120*T*
root–soil interface 6:113*T*
Popophyra 6:296*F*
Popper, Karl 1:16
poppy seeds 7:26
population density 6:157–167
measurement methods
as a causal agent 6:157–158
general discussion 6:157
as a response variable 6:157
spatial density variation
consequences
genetic/evolutionary consequences 6:163
population dynamics 6:162–163
determinant factors
external forcing 6:162
generation mechanisms 6:161–162
lifestyle characteristics 6:162
population dynamics 6:162
spatial patterns
density–biomass thinning laws
–3/2 thinning rule (plants) 6:161
–4/3 thinning rule (animals) 6:161
general discussion 6:161
statistical patterns 6:158–160, 6:160*F*
Taylor's power law 6:160–161
variability 6:158–160, 6:159*F*
within-species occupancy-abundance patterns 6:161
summary discussion 6:167
temporal density patterns
consequences
cascade effects 6:164–165
evolutionary consequences 6:165
increased extinction probability 6:164
determinant factors
external forcing 6:164
life history characteristics 6:164
population dynamics 6:164
measurement methods
distribution shape 6:163
fluctuation magnitude 6:163
red-shifted variance 6:163
Taylor's power law 6:163–164
trans-specific patterns
–3/2 thinning rule 6:165, 6:166*F*
cross-species comparisons 6:165
over-species density distributions 6:165
trait-based relationships
breeding systems 6:165
dispersal ability 6:166
geographic range 6:166–167, 6:166*F*
homozygosity 6:166
latitudinal gradients 6:166
reproductive investment 6:165–166
resource usage 6:166
Taylor's power law 6:167
trophic status 6:166
population diversity
agrobiodiversity 1:132, 1:132*T*
countryside biogeography 2:351–352
ecosystem services
general characteristics 6:170–171
global population diversity 6:171
local population diversity 6:171
regional population diversity 6:171
introduced species 6:172–173
natural selection 6:168
population dynamics 6:175–181
basic concepts 3:1–3, 6:175
density-dependent dynamics
continuous time 6:177, 6:177*F*
discrete time 6:177–178
density-independent dynamics
single species with age structure 6:176–177
single species without age structure 6:175–176
evolutionary response studies 3:389
historical perspectives 2:420–421
landscape ecology 4:488–502
basic concepts 4:488–491
continuous habitat variability 4:497–498, 4:498*F*
core and satellite model 4:496
disturbance recovery 4:498–499, 4:499*F*
edge effects 4:496–497
environmental gradients 4:497–498
extinction debt 4:497
habitat fragmentation 4:496–497
habitat heterogeneity 1:381–382, 1:381*F*, 4:497–498
landscape changes 4:496–497
landscape structure
basic concepts 4:491–493
habitat preservation 4:495
immigration 4:495
interspecific interactions 4:498–499, 4:499*F*
landscape context 4:493
patch–matrix–corridor model 4:479, 4:479*F*, 4:492–493, 4:492*F*
population viability analysis (PVA) 4:493–495
quantification methods 4:493
single large or several small (SLOSS) debate 4:495
metapopulation dynamics 4:495–496
research summary 4:499–500
simulation models 4:488–491, 4:490*F*, 4:491*F*
soil moisture prediction 4:497–498, 4:498*F*
source–sink dynamics 4:496, 5:324–325
spatial patterns 4:488–491, 4:489*F*
modern extinctions 5:370
parasites
ecological strategies 5:668
population growth 5:667–668
parasitoids 5:679–681, 6:179
simple two-species interactions
basic concepts 6:178–179
competition 6:180
disease dynamics 6:179–180
host–parasitoid dynamics 6:179
prey–predator interactions 6:179
spatial density variation
consequences 6:162–163
determinant factors 6:162
stochastic population processes 6:178
summary discussion 6:180
temporal density patterns 6:164
population genetics 6:182–198
conservation genetics 2:263–277
ex-situ conservation 2:273–274
future outlook 2:274–275
genetic drift 2:265–267
genomic research 2:274–275, 2:275*F*
heterozygosity and fitness correlations (HFCs) 2:266
historical perspective 2:263–265
inbreeding 2:265–267, 2:272–273
individual genotypes 2:272–273
kin relationships 2:272–273

population genetics (*continued*)
low diversity–small population relationship **2**:265–267
management units **2**:268–269
population structure **2**:267–272, **2**:269*F*, **2**:270*F*
skyline plots **2**:267, **2**:267*F*
deterministic population genetics
allele and genotype frequencies **6**:183
genetic load
basic concepts **6**:188
mutational load **6**:188
segregational load **6**:188
substitutional load **6**:188–189
multiple loci
close linkage evolution **6**:189–190
epistatic selection **6**:189
selection absence **6**:189
mutations
evolutionary time span **6**:184
trait selection **6**:184
types and rates **6**:184
mutation–selection balance **6**:187–188
quantitative trait-based selection
change rates **6**:186
multiple character selection **6**:186
response predictions **6**:185–186, **6**:186*F*
single locus selection
basic model **6**:184–185
directional selection **6**:185
favorable mutations **6**:185
gene frequency change time course **6**:185
survival rates **6**:185
variation maintenance
definition **6**:186–187
frequency-dependent selection **6**:187
heterozygote advantage **6**:186–187
meiotic drive **6**:187
relative fitness variations **6**:187
dispersal biogeography **2**:540–541
drift–deterministic force interactions
balanced polymorphisms **6**:195
diffusion equations
basic concepts **6**:195
fixation probabilities **6**:196
stationary distributions **6**:195–196
directional selection effects **6**:194–195
group selection **6**:197
Hill–Robertson interference **6**:196–197
kin selection **6**:197
population subdivision
basic concepts **6**:193–194
island model **6**:194
partitioning variability **6**:194
structural models **6**:194
selection targets **6**:195
shifting balance theory **6**:197
harmful algal blooms **1**:475
landscape genetics **4:508–523**
analytical frameworks
assignment tests **4**:511
barrier detection **4**:511, **4**:512*F*
graph theory/network analyses **4**:512–514, **4**:513*F*
Mantel tests **4**:510–511
ordination and regression analysis **4**:514
resistance surfaces **4**:511–512, **4**:513*F*
isolation-by-distance (IBD) models **4**:508, **4**:509*F*
research background **4**:508–509
research challenges
genetic distance (F_{ST}) measures **4**:518–519
interdisciplinary communication **4**:517–518
landscape genomics **4**:519–520
multispecies comparative studies **4**:518
spatial analysis method (SAM) **4**:519
spatial replication **4**:518
spatial scale effects **4**:518
research results
anthropogenic impacts **4**:516–517
barrier identification **4**:515–516
gene flow barriers **4**:515–516
general research trends **4**:514–515
landscape variables influences **4**:516
sampling strategy and design **4**:509–510
spatial patterns and connectivity **4**:508–509
molecular level diversity **2:626–637**
conservation biology **2**:637
DNA sequencing **2**:627
future research outlook **2**:635–637
general discussion **2**:626
genetic variation
mutation process **2**:627–628
phylogenetic divergence **2**:632–633
process mechanisms **2**:627
mutation process
characteristics and classifications **2**:627–628
cytological considerations **2**:628–629
empirical research **2**:631–632, **2**:631*T*
evolutionary models **2**:629–630
genome size effects **2**:628, **2**:628*F*
influencing factors **2**:629
mutation rates **2**:628, **2**:628*T*
nearly neutral theory **2**:630
neutral theory **2**:629–630
nucleotide polymorphism **2**:631–632, **2**:631*T*, **2**:632*F*
ontogenetic considerations **2**:628–629
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
selection-based models **2**:630–631
statistical tests **2**:631
phylogenetic divergence
amino acid replacement **2**:634, **2**:634*T*
empirical research **2**:634
evolutionary models **2**:633
general discussion **2**:632–633
nearly neutral theory **2**:633
neutral theory **2**:633
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
selection-based models **2**:633–634
statistical tests **2**:634
substitution process **2**:632–633, **2**:634*T*
prediction models **2**:634–635, **2**:635*T*, **2**:636*F*
research background **2**:626–627
valuation measures **2**:637
molecular variation
nucleotide sequences
interpretation process **6**:183
measurement **6**:183
protein electrophoresis **6**:183
phenotypic variation
concealed variability **6**:182
discrete variation **6**:182
general discussion **6**:182
interpretation process **6**:182–183
quantitative variation **6**:182
random genetic drift
basic concepts **6**:190, **6**:190*F*
coalescent process
basic concepts **6**:192–193, **6**:193*F*
neutrality tests **6**:193
genetic differentiation **6**:190–191
homozygosity increases **6**:190
molecular evolution
evolution rates **6**:191–192
neutral evolution theory **6**:191
neutral polymorphism **6**:192
research summary **6**:197–198
Population Management Plan (PMP) **7**:476
populations **6:168–174**
definition **6**:168
divergence detection
DNA variation **6**:169
genetic diversity **6**:169
migration rates **6**:169
morphological variation **6**:169
population size **6**:169
protein variation **6**:169
divergence indices **6**:169–170
divergence processes
characteristics **6**:168
genetic drift **6**:168
influencing factors **6**:168–169
natural selection **6**:168
geographic extent
average range size **6**:172
estimation methods **6**:171–172
populations per area **6**:172
species diversity **6**:172
human-induced changes
California **6**:172–173
extinctions **6**:172–173
genetic diversity loss **6**:173
habitat impact **6**:172–173
habitat loss–extinction rate relationship
estimation methods **6**:173, **6**:173*F*
extinction rates **6**:173, **6**:173*F*
population–area relationship **6**:173, **6**:173*F*
introduced species **6**:172–173
management strategies
invasion control **6**:173–174

preservation **6**:174
restoration **6**:174
Western Australia **6**:172–173
importance
direct economic value **6**:170
ecosystem service value
general characteristics **6**:170–171
global population diversity **6**:171
local population diversity **6**:171
regional population diversity **6**:171
esthetic value **6**:170
general characteristics **6**:170
genetic value **6**:170
species conservation value **6**:170
Mendelian populations **6**:168
species variability **6**:168
population stabilization **6:199–209**
cause–effect relationships
economic development **6**:204–206
education levels **6**:203–204, **6**:205*F*
fertility and mortality patterns **6**:202–204
future outlook **6**:208
government regulations and policies **6**:207
historical perspective
pre-1900 population changes **6**:199–200, **6**:200*F*
1900–1950 population change **6**:200
1950–2010 population change **6**:200–201
age profile changes **6**:202, **6**:204*F*
contraceptive methods **6**:202*F*
demographic transition **6**:199–200
family planning methods **6**:202, **6**:202*F*
fertility and mortality patterns **6**:201, **6**:203*F*
growth rates **6**:201–202
International Conference on Population and Development (ICPD) **6**:207–208
population growth responses **6**:206–207
2010–2050 population projections **6**:206
population viability analysis (PVA) **6:210–219**
basic concepts **3**:2, **6**:210
conservation biology **2**:243–244
criticisms **6**:218–219
functional role
extinction probability–area relationship **6**:213–214, **6**:214*F*
management decision-making **6**:214
research applications **6**:213–214
risk measures **6**:214, **6**:215*F*
landscape structure **4**:493–495
methods and tools
demographic models
analytical methods and approximations **6**:215
general discussion **6**:215
matrix population models **6**:215
Monte Carlo simulation models **6**:215–216, **6**:216*F*
numerical methods **6**:215
time series analysis methods **6**:215
problematic issues
density dependence **6**:217
general discussion **6**:217
genetic issues **6**:217
habitat dynamics **6**:217
historical data **6**:217–218
metapopulation models **6**:217
risk spreading **6**:217
similar species extrapolations **6**:218
space **6**:217
sensitivity analyses **6**:216–217
stochastic population processes and extinctions
demographic population fluctuation
catastrophes **6**:212, **6**:212*F*
demographic stochasticity **4**:659–660, **4**:661*F*, **6**:210–211, **6**:211*F*
environmental stochasticity **4**:660, **4**:661*F*, **6**:211–212, **6**:211*F*
deterministic extinction processes **6**:213
extinction probability–carrying capacity relationship **6**:212–213, **6**:213*F*
extinction vortices **6**:213
genetic stochastic processes **6**:212
testing methods **6**:218
Populus
boreal forests **1**:628–630, **1**:630*T*
energy use impacts **3**:261
Holocene postglacial spread **2**:551–552*T*
North America **7**:118
Populus angustifolia **4**:173
Populus balsamifera **1**:630*T*, **2**:636*F*, **6**:120*T*, **6**:123–124
Populus diversifolia **1**:270
Populus euphratica **2**:470
Populus fremontii **4**:173
Populus koreana **1**:271–272
Populus suaveolens **1**:630*T*
Populus tremula **1**:269, **1**:630*T*, **2**:636*F*
Populus tremuloides **1**:612, **1**:628–629, **1**:630*T*, **1**:632*T*, **6**:342–343, **7**:259
Populus ussuriensis **1**:271–272
Poraqueiba paraensis **4**:286–287, **4**:288*T*
Poraqueiba sericea **4**:287–290, **4**:288*T*
Poratia obliterata **5**:425–426
porcelain crabs **5**:33
Porcula salvania **3**:226
porcupine fish **3**:467–471
porcupines
boreal forests **1**:631–632
Great American Biotic Interchange (GABI) **2**:584*F*
range ecology **1**:703
Porifera
Census of Marine Life project **2**:7*F*
classifications and sampling methods **4**:381*T*
convergent evolution **3**:615*F*
fossil record **3**:681*F*, **3**:684*F*
global diversity **4**:385*T*
hydrothermal vent ecosystems **7**:329–330
marine ecosystems **5**:50*T*
marine environments **4**:382*F*
Near East ecosystems **5**:470*T*
number of species **4**:371*T*, **5**:669–670*T*
phylogenetic relationships **7**:434*F*
Proterozoic–Cambrian transition **3**:396*F*
river ecosystems **6**:478*T*
Porites corals **4**:664
Porites pukoensis **3**:201, **3**:202*T*
Poritiinae **1**:652
Porotermes **4**:409*T*
Porphyra **2**:641–642
Porphyrospatha **3**:522*T*
porpoises **4**:690*T*, **4**:693–694, **5**:65*T*
portabella mushrooms **3**:637–638, **3**:637*T*
Portieria hornemannii **5**:711
Portugal
botanical gardens **6**:507*T*
intercropping systems **2**:382–383
wheat production **1**:87–88
Portulacaceae **2**:374*T*, **2**:376*F*
Portulaca oleracea **3**:131–132
Posidonia **6**:542*T*
Posidonia angustifolia **6**:541*T*
Posidonia australis **6**:541*T*
Posidonia coriacea **6**:541*T*
Posidonia denhartogii **6**:541*T*
Posidonia kirkmanii **6**:541*T*
Posidonia oceanica **4**:358–359, **6**:541*T*
Posidonia ostenfeldii **6**:541*T*
Posidonia robertsoniae **6**:541*T*
Posidonia sinuosa **6**:541*T*
Posidonoideae **6**:542*T*
positive emotional responses
foods/food selection **1**:42–43, **1**:42*T*
habitat selection **1**:41–42
research background **1**:40–41
restorative responses **1**:43
symmetry **1**:40–41
Posoqueria **3**:522*T*
post-ark paradigm **7**:478
postglacial range expansion **2**:550
Postmonorchis orthopristis **7**:450–451*F*
Postojna Planina Cave System (cave stream communities) **7**:58, **7**:59*T*
Postojna Planina Cave System (epikarst communities) **7**:56*F*, **7**:57–58, **7**:59*T*
postpartum infecundity **6**:202–204
postzygotic reproductive isolation **6**:636–637
Potaliaceae **3**:522*T*
Potamochoerus larvatus **1**:55–56
Potamochoerus porcus **1**:640
Potamocorbula amurensis **5**:61
Potamogeton **3**:360–361, **3**:578*F*
Potamogetonaceae **6**:542*T*
Potamopyrgus antipodarum **4**:465
Potamornis **2**:519*T*, **2**:520*T*
potassium (K)
desert ecosystems **2**:474
elemental partitioning **1**:554*T*
relative abundance **1**:553*F*
river water composition **3**:562*T*
root–soil interface **6**:114*F*
potato blight **4**:313–314
potatoes **2**:383, **2**:384*F*, **2**:475–476, **3**:32*F*, **3**:128, **5**:672–673, **6**:85–86
pothole **7**:369*T*
Pothos **3**:522*T*
Pottingeriaceae **1**:260*T*
pottos **6**:251*T*
pot worms **6**:590*T*
Pouchet, Félix-Archimède **5**:629–630

Poulton, E. B. **6**:681–682
Pound, Ezra **4**:649–650
Pourouma cecropiaefolia **4**:287–290, **4**:289*T*
Pouteria caimito **4**:284*F*, **4**:285, **4**:287–290, **4**:289*T*
Pouteria macrophylla **4**:288*T*
Pouteria ucuqui **4**:288*T*
poverty **6:220–228**
  environmental impacts **3**:286–287, **7**:171
  people–conservation relationship **2:230–237**
    benefits
      cost–benefits **2**:226
      fishing industry **2**:232
      forestry sector **2**:232
      grasslands **2**:232
      nature-based tourism **2**:232–233
      payment for ecosystem services (PES) **2**:226–227
      poverty reduction **2**:227–228
      protected area systems **2**:226
    disadvantages **2**:233
    historical perspective **2**:230–231
    key organizational players **2**:231–232
    poverty indicators **2**:233
    poverty reduction **2**:227–228
    poverty trends **2**:233–235, **2**:234*F*, **2**:234*T*
    research summary **2**:235
  people–ecosystem relationship
    autonomous people **6**:221–222, **6**:221*T*
    biosphere people **6**:221*T*
    conservation practices **6**:226
    ecodevelopment programs **6**:226–227
    ecological refugees **6**:221*T*, **6**:225–226
    future challenges **6**:227
    general characteristics **6**:220–221
    land-use changes **6**:223–224, **6**:224*T*
    resource ownership, sharing, and utilization **6**:223, **6**:224*T*
    subjugated people **6**:221*T*, **6**:222–223
    value appropriation **6**:224–225
  sustainability–biodiversity relationship **7**:77–80
  wealth versus poverty **6**:220
powdery mildew **3**:633
Powell, John Wesley **2**:281
Poxviridae **5**:130*F*, **5**:134*T*, **5**:136
Pradosia schomburgkiana **1**:160
Praeovibos **4**:724–725*T*
prairie dogs **3**:188, **3**:196, **4**:10, **6**:342, **7**:405–406
prairie ecosystems **2**:61*F*, **6**:239, **6**:458–459, **7**:236–237
Prairie Pothole Region **6**:239
Praomys minor **5**:462
Praon palitans **4**:362
Prasium majus **5**:485
Pratylenchus **7**:421*F*, **7**:422–424, **7**:427*F*
prawns
  aquaculture **1**:190*T*
  classifications **4**:381*T*
  number of species **3**:182*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
Precambrian era
  arthropod radiation ecology **2**:399–400
  fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*
  mass extinctions **1**:413*F*, **3**:419*F*
  molluscs **5**:373–374, **5**:374*F*
precipitation
  boreal forests **1**:626–627, **1**:628*F*
  climate change effects **3**:165, **4**:225*F*, **4**:226
  desert climates
    rainfall pulses **2**:463–464
    winter and summer rainfall patterns **2**:464–466, **2**:465*F*
  dryland syndrome **2**:481–482, **2**:483*F*, **2**:484*F*
  ecosystem function measurements **3**:79, **3**:81*F*
  El Niño Southern Oscillation (ENSO) events
    floods **3**:159
    water-limited ecosystems **3**:158–159
  forest ecosystems **3**:529*F*
  Near East ecosystems **5**:465–467, **5**:465*F*, **5**:466*T*
  plant biodiversity–climate relationship **2**:120, **3**:120
  riparian zones **6**:467*T*
  seasonally dry tropical forests **5**:488–490
pre-Darwinian evolutionary theories **3**:393, **3**:393*F*
predators
  characteristics **6**:230
  definition **5**:656–657
  Diptera **3**:494
  ecological role **6:229–249**
    background information
      disturbance models **6**:230–231, **6**:231*F*
      generalized food web theory **6**:231, **6**:231*F*
      indirect effects **6**:230
      keystone species **6**:230–231
      predator characteristics **6**:230
      species interactions **6**:230
      top-down forces **6**:230, **6**:231*F*
      trophic cascades **6**:231, **6**:233*F*, **6**:234*F*
      trophic groups **6**:230
    case studies
      boreal forests **6**:236–237, **6**:238*F*
      fragmented coastal scrub habitats **6**:237–239
      kelp forests **6**:232–233, **6**:233*F*, **6**:234*F*
      lake ecosystems **6**:233–235
      midcontinent North American prairies **6**:239
      nonnative island predators **6**:240–241
      oceanic systems **6**:236
      research summary **6**:241–242
      rivers and streams **6**:235–236
      rocky shores **6**:232
      temperate forests **6**:236–237, **6**:238*F*
      tropical forests **6**:239–240
      tropical reef systems **6**:236
    future outlook **6**:247–248, **6**:248*F*
    generalist predators
      apparent competition **6**:245–246
      community stability **6**:246
      enemy-free space **6**:244
      numerical responses **6**:245–246
      saturating functional responses **6**:244–245
      switching **6**:244
    generality considerations **6**:246–247, **6**:247*F*
    interaction pathways **6**:229*F*
    research background **6**:229–230
    seed dispersal–predation relationship **7**:283–284
    theoretical studies
      density-independent mortality factors **6**:242–243
      generalist predators **6**:244
      numerical and functional responses **6**:243
      predator–biodiversity relationship **6**:242–243
      specialist predators **6**:243–244
habitat–niche relationship **4**:45–46
intertidal ecosystems **4**:334–335
livestock management effects **6**:342
nematodes **7**:430–431
parasitism **5**:656–657
pest control–predator diversity relationship
  complementarity **1**:377–378
  enhancement factors **1**:379–380
  facilitation **1**:377
  functional diversity (FD) **1**:378
  functional redundancy **1**:378
  insurance hypothesis **1**:378
  intraguild interference **1**:378–379, **1**:378*F*
  risk reduction/risk enhancement **1**:376–377
  sampling/selection effects **1**:377
plant defense systems **2**:429
prey–predator interactions
  agricultural intensification effects **1**:374–375
  carnivores **1**:687–690
  climate change effects **2**:75–76, **2**:75*F*
  coevolution **6**:721
  domesticated ruminants **1**:709
  ecosystem services **3**:90*T*
  Lotka–Volterra predator–prey system **6**:717–718, **6**:718*F*, **7**:8–9, **7**:9, **7**:9*F*
  mammalian evolution **4**:699–700
  megaherbivores **5**:226–227
  metapopulation theory **5**:242–243, **5**:248, **5**:248*F*
  population dynamics **6**:179
  predator guilds **1**:690–692
  simple two-species interactions **6**:179
  surplus killing **1**:688
  theoretical perspectives **3**:3–4, **3**:4*F*
seed dispersal–predation relationship **7**:283–284
trophic pyramid **4**:670–671

pre-Quaternary mammal extinctions **5:1–9**
Age of the Dinosaurs **5**:5–6
causal factors
background extinctions **5**:2–3
distal and proximal causal factors **5**:3–4
mass extinctions **5**:2–3
pseudoextinctions **5**:2
temporal and biogeographic scales **5**:3
Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
early Paleocene faunal recovery **5**:7
geologic timescale **5**:1*T*
late Triassic faunal turnover **5**:4–5
Paleocene–Eocene thermal maximum **5**:7–8
Permian–Triassic mass extinction **5**:4
phylogenetic relationships **5**:1*F*
summary discussion **5**:8
synapsids **5**:1–2, **5**:4
Tertiary faunal evolution **5**:7–8
therapsid radiation **5**:4–5
Presbytis chrysomelas chrysomelas **6**:256–258*T*
Presbytis chrysomelas cruciger **6**:256–258*T*
Presbytis comata **6**:256–258*T*
Presbytis hosei canicrus **6**:256–258*T*
Presbytis hosei sabana **6**:256–258*T*
Presbytis melalophos mitrata **6**:256–258*T*
Presbytis melalophos sumatrana **6**:256–258*T*
Presbytis potenziani potenziani **6**:256–258*T*
Presbytis potenziani siberu **6**:256–258*T*
President's Committee of Advisors on Science and Technology (PCAST) **3**:133
Preston's canonical hypothesis **6**:645
Pretoria National Botanical Garden **6**:506–508
Preuss' red colobus **6**:256–258*T*
prey–predator interactions
agricultural intensification effects **1**:374–375
carnivores **1**:687–690
climate change effects **2**:75–76, **2**:75*F*
coevolution **6**:721
ecosystem services **3**:90*T*
introduced species **4**:360–361
Lotka–Volterra predator–prey system
competition models **6**:717–718, **6**:718*F*
food webs **3**:500–501, **3**:508
interaction strengths **3**:507
species coexistence **6**:672
structural stability **7**:8–9, **7**:9, **7**:9*F*
mammalian evolution **4**:699–700
megaherbivores **5**:226–227
metapopulation theory **5**:242–243, **5**:248, **5**:248*F*
population dynamics **6**:179
predator guilds **1**:690–692
simple two-species interactions **6**:179
surplus killing **1**:688
termites **4**:413, **4**:415*F*
theoretical perspectives **3**:3–4, **3**:4*F*
Priapula
global diversity **4**:385*T*
marine ecosystems **5**:50*T*
phylogenetic relationships **7**:433*F*
Proterozoic–Cambrian transition **3**:396*F*
Priapulida
classifications and sampling methods **4**:381*T*
fossil record **3**:684*F*
known and estimated species **4**:384*T*
number of species **5**:669–670*T*
phylogenetic relationships **7**:434*F*
prickly pear (*Opuntia*) **2**:467*F*, **2**:476, **2**:663–664, **4**:350–351, **4**:351*F*, **5**:401–402
Pridontes maximus **7**:229*T*
Prigogine–Glansdorff principle **5**:630, **5**:633
Prigogine, Ilja **5**:630
Prigogine's Angolan colobus **6**:256–258*T*
primary production
desert ecosystems **2**:474
ecosystem function measurements **3**:77–78
carbon-14 (^{14}C) measurements **3**:66
carbon-based measurements **3**:65–66
data interpretation and analysis **3**:66–67
fluorescence measurements **3**:66
measurement approaches **3**:65
oxygen-based measurements **3**:65
ecosystem services **1**:342–344
eutrophication impacts **5**:23–24, **5**:25, **5**:26*F*, **5**:27*F*
exploitation ecosystem hypothesis **3**:503
food webs
eutrophication impacts **5**:24–25
general discussion **3**:503–504
trophic levels **3**:503–504
freshwater invertebrates **4**:376
herbivores **3**:77–78
marine ecosystems **5**:46
multichannel omnivory **3**:504
number of species effects **3**:245–246
pelagic ecosystems **5**:683–686, **5**:685*F*, **5**:686*F*
trophic cascades **3**:504–505, **7**:258*F*, **7**:259–260
trophic complexity **3**:504
trophic pyramid **4**:670–671
primates
characteristics and occurrences **4**:692–693
conservation strategies **6:250–260**
conservation status **6**:255
future outlook **6**:258–259
global diversity **6**:250–252, **6**:251*T*, **6**:252*T*
Global Strategy for Primate Conservation **6**:255–258
historical perspective **6**:250
major threats
general characteristics **6**:252–253
habitat destruction **6**:253
hunting **6**:253–254
live capture **6**:254–255
primate endemism **6**:252*T*
evolutionary history **2**:583*F*
forest management impact **1**:453
fruit and seed dispersal **6**:50
global diversity **6**:250–252, **6**:251*T*, **6**:252*T*
global wildlife trade **3**:409*T*
major threats
general characteristics **6**:252–253
habitat destruction **6**:253
hunting **6**:253–254
live capture **6**:254–255
number of species **4**:690*T*
Paleocene–Eocene thermal maximum **5**:7–8
primate endemism **6**:252*T*
representative species **2**:587*F*
social systems **2**:646
species diversity **2**:586, **4**:690*T*
threat percentage **3**:195
tropical rain forests **6**:309*T*
prime agricultural lands **1**:543
Primelephas **5**:224*T*
Primula **1**:270
Primula cunefolia **4**:516
Primula elatior **4**:516–517, **4**:574*F*
Primula magellanica **4**:107
Primula scotica **3**:206
Primula sieboldii **4**:515
Primula veris **6**:458
Primula vulgaris **6**:5
Prince Edward Island **1**:186*T*
principal component analysis (PCA) **3**:600, **3**:600*F*, **4**:514
Principle of Organic Selection **6**:7–8
Prinia spp. **5**:603–604
Priodontes maximus **2**:587*F*, **4**:689–690
Prionailurus bengalensis **1**:517
Prionodiaptomus incarum **3**:188
Prionodura newtoniana **2**:75*F*
Prionotropis hystrix rhodanica **3**:220–221
Prioria copaifera **7**:272
Pristiphora erichsonii **4**:182–183
private conservation areas **2**:296–297, **2**:297*T*
private ownership market economy **5**:145–147
private property **6**:274–275, **6**:274*T*
private reserves **7**:224
Proasellus coaxalis **5**:471
Proboscidae **4**:685, **4**:690*T*, **4**:694–695, **4**:701*T*, **4**:727*T*, **5**:224*T*, **5**:225*T*
proboscis monkey **5**:14
Procambarus clarkii **3**:696
Procapra gutturosa **3**:191
Procapritermes **4**:409*T*
Procaridiidae **2**:416–417*T*
Procarpa spp. **6**:333
Procastoroides **4**:724–725*T*
Procavia capensis **4**:695
Procavia johnstoni **1**:50
Procellaria parkinsoni **7**:188
Procellariiformes **7**:181–182
Prochilodus mariae **5**:315
prochloraz **5**:695
Prochlorococcus **4**:671–673, **4**:675*F*, **5**:134–135, **5**:293–294, **5**:302–303
Prochloron didemni **1**:310–311
Procoptodon **4**:727–728
Procornitermes **4**:409*T*
Proctoptodon goliah **5**:364*F*
Proctoptodon pusio **5**:364*F*
Proctoptodon rapha **5**:364*F*
Proctotrupoidea **4**:178, **5**:395–396

Procubitermes **4**:409*T*
Procyon cancrivorus **2**:587*F*
Procyonidae
  body mass **1**:687*F*
  evolutionary history **2**:583*F*
  food categories **1**:690*F*, **1**:691*F*
  home range size **1**:687*F*
  species diversity **1**:686, **4**:693
  tropical rain forests **6**:309*T*
Procyon lotor **7**:390
Prodeinotherium **5**:224*T*
production efficiency **7**:264–265
productivity growth analysis **3**:29–30
Proganochelys **7**:339–340
Progoneata **1**:203*F*
Program on Ecosystem Change and Society **4**:326–327
Program on International Dimensions of Global Environmental Change **4**:326–327
progression rule **7**:344*F*, **7**:345–346, **7**:346*F*, **7**:347*F*
Progressive Era **2**:282–283
Progura gallinacea **5**:364*F*
Progura naracoortensis **5**:364*F*
Project 2061 **3**:133–134
Project Tiger **3**:315–316
prokaryotes
  biodiversity **5**:255*F*
  evolutionary development **2**:640–641
  functional role
    element cycling
      aquatic sediments **5**:302*F*, **5**:303
      extreme environments **5**:304
      mineralization processes **5**:302–303
      terrestrial soils **5**:303–304
      water column **5**:302–303, **5**:303*F*
    metabolic diversity **5**:300–302, **5**:301*T*, **5**:302*F*
    principle properties **5**:300
    symbiotic relationships **5**:304–305
  genes/genetic code **3**:657, **3**:658*F*
  mutation rates **2**:628*F*
  nomenclature code **5**:551–552, **5**:553*T*
  number of species **6**:583
  phylogenetic relationships **6**:584*F*
  soil food webs **6**:448–449, **6**:449*F*
  symbiotic relationships **5**:304–305
  temperature tolerance ranges **4**:113–115
Prolagidae **4**:697
Prolagus sardus **1**:702
Prolecithophora **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
Prolecithoplana **7**:442–443*F*
Prolemur simus **6**:256–258*T*
Promecotheca cumingii **5**:603–604
Promeostoma **7**:442–443*F*
Promirotermes **4**:409*T*
pronamide **4**:90*T*
pronghorn
  energy use impacts **3**:256–257
  grassland ecosystems **6**:332
  Great American Biotic Interchange (GABI) **2**:584*F*
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  sustainable landscape development study **2**:503–505
Pronocephaloidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
propanil **4**:90*T*, **5**:696
property rights **6:273–279**
  basic concepts
    biodiversity **6**:273–274
    functional role **6**:273, **6**:275–276
    ownership **6**:274, **6**:274*T*
    scope **6**:273
    value **6**:273
  biodiversity protection
    challenges **6**:275
    cost–benefit distributions **6**:278
    exclusivity **6**:278
    general discussion **6**:277–278
    human rights **3**:315
    private and social goals alignment **6**:278
    uncertainty **6**:277–278
  classifications
    biodiversity protection **6**:275
    common property **6**:274, **6**:274*T*
    open access **6**:274, **6**:274*T*
    ownership **6**:274, **6**:274*T*
    private property **6**:274–275, **6**:274*T*
    property rights over goods **6**:275
    property rights over services **6**:275
    state property **6**:274, **6**:274*T*
  evolutionary development
    causal factors **6**:277
    ecosystem use **6**:277
    single-species use **6**:277
    use expectations **6**:277
  functional role
    basic concepts **6**:273, **6**:275–276
    externalities **6**:276
    geographic scale **6**:276–277
    transactions costs **6**:276
    uncertainty **6**:276
  future challenges **6**:278–279
  research summary **6**:278–279
Prophalangopsis obscura **3**:727–728
Propithecus candidus **6**:256–258*T*
Propithecus coquereli **6**:256–258*T*
Propithecus coronatus **6**:256–258*T*
Propithecus diadema **6**:256–258*T*
Propithecus edwardsi **6**:256–258*T*
Propithecus perrieri **6**:256–258*T*
Propithecus tattersalli **6**:256–258*T*
Propithecus verreauxi **1**:55–56
Propleopus **4**:727–728
Propleopus oscillans **5**:364*F*
propoconazol **5**:695
proportional sampling **2**:602*F*, **2**:603
Propraopus **4**:727*T*
propyzamide **5**:696
Prorhinotermes **4**:409*T*
Prorocentrales **1**:472, **1**:473*F*
Prorocentrum minimum **1**:480, **3**:368–369
Prorodon **6**:295
Prosaurolophus **2**:519*T*, **2**:520*T*
proscillaridin **5**:712*T*
Proscopiidae (false walking sticks or proscopids) **3**:733
Proseriata **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Prosobonia cancellata **7**:184
prosobranchs **4**:464*T*
proso millet **3**:32*F*
Prosopis africana **1**:67
Prosopis caldenia Burkart **1**:708–709
Prosopis cineraria **7**:231
Prosopis juliflora **3**:115–116, **6**:457
Prosopis juliiflora **6**:78
Prosopis spp. **1**:69, **1**:120, **6**:342–343
Prosopium williamsoni **4**:515
prospect **1**:39
Prostigmata **1**:216–217, **6**:587–588
Prostratin **1**:594, **1**:596–597
Protacanthopterygii **3**:464–465, **3**:465*F*, **6**:522–523
Protaphorura **4**:393*F*
Proteaceae **5**:219, **5**:221, **5**:621*F*
protected area systems **6:485–503**
  basic concepts **2**:289–292, **2**:290, **6**:485–486
  biodiversity conservation **4**:435–436
  biodiversity stewardship **2**:297
  biodiversity sustainability **5**:448–449
  biosphere reserves **2**:301, **6**:489
  characteristics **2**:289–292
  classifications **2**:290, **5**:442–443, **5**:443*T*, **6**:485–486
  climate change
    challenges **5**:449
    Great Limpopo Transfrontier Park **5**:449
    Seamounts Marine Management Area **5**:449–450
  conservancy/megaconservancy networks **2**:301, **2**:302*F*
  conservation biology **2**:239–241, **2**:240*F*
  creation process **6**:488–489
  current status
    effectiveness measures
      challenges **6**:498
      land conversions **6**:498, **6**:499*T*
    freshwater ecosystems **6**:494–498, **6**:496*F*, **6**:497*F*
    global, regional, and national coverage **6**:489–491, **6**:490*F*, **6**:491*T*
    International Union for Conservation of Nature and Natural Resources (IUCN) categories **6**:491, **6**:491*T*
    major threats
      encroachment **6**:498–499
      financial resources **6**:500
      general threats **6**:498
      isolation factors **6**:499–500
      mining impacts **6**:499
      poaching/bush meat **6**:499
      tourism industry **6**:500
    marine ecosystems **6**:494, **6**:495*F*
    terrestrial habitats
      biogeographic realms **6**:491
      biome realms **6**:491–492, **6**:492*T*, **6**:493*F*
      data challenges **6**:491
      deserts **6**:492, **6**:495*T*
      ecoregion protection **2**:591–592, **6**:492, **6**:493*F*

forests **6**:492, **6**:494*F*, **6**:494*T*
islands **6**:493–494
mountains **6**:492–493
World Wildlife Fund (WWF) realms **6**:492*T*
World Database on Protected Areas (WDPAs) **6**:489
current trends
loss of existing protected areas **6**:500–501
multiple-use protected areas **6**:501
number and size accretion **6**:500
Reducing Emissions from Deforestation and Forest Degradation (REDD) **6**:500
sustainable management **6**:501
deforestation **2**:455
funding considerations **5**:448
future threats **5**:448
general discussion **5**:442
geographic distribution **2**:291*F*
global biome coverage **2**:292*F*
global growth estimates **2**:290*F*
global urbanization impact **4**:235–236, **4**:235*F*
government legislation and regulations **6**:485–486
historical background **5**:442, **6**:77–78, **6**:486
indigenous peoples **4**:276–277
international approaches **6**:489
island ecosystems **5**:444*T*
mammal conservation **3**:196, **4**:710–711
marine protected areas **5:94–104**
benefits
abundance increases **5**:96–97, **5**:97*T*
adjacent fisheries **5**:97
biomass increases **5**:96–97, **5**:97*T*
characteristics **5**:96–97
long-term ecological benefits **5**:97–98
nontarget species **5**:97
reproductive potential **5**:96–97
size increases **5**:96–97, **5**:97*T*
Bowie Seamount Marine Protected Area (Sglaan Kinghlas) **5**:446–447
characteristics and functional role **5**:94–95
climate change threats **5**:100–101
conservation strategies **5**:39–40
ecosystem-based management
characteristics **5**:95
definitions **5**:95
goals and objectives **5**:95
marine spatial planning **5**:96
networks **5**:95–96
effectiveness **5**:100
Endeavor Hydrothermal Vents Marine Protected Area **5**:447
fish conservation **3**:453–454
fish stocks **3**:486
geographic extent **5**:445–446, **5**:446*T*
global commitment **5**:99–100
Great Barrier Reef Marine Park **5**:446
Papahanaumokuakea Marine National Monument **5**:446
socioeconomic benefits **5**:98–99, **5**:98*T*
neotropical seasonally dry forests **5**:492–496, **5**:493*T*, **5**:495*F*
paper parks **6**:498
people–conservation relationship
benefits **2**:226
victimization factors **2**:224–225
selection criteria **6**:486–487
selection methods
basic knowledge and processes **6**:487
imminent extinction probability **6**:488*F*
networks and connectivity **6**:487–488
species protection **6**:488, **6**:488*F*
Southeast Asia **1**:520
terrestrial extent **5**:443–444, **5**:444*T*
terrestrial protected areas
Canaima National Park **5**:444
Dartmoor National Park **5**:445
Jenolan Karst Conservation Reserve **5**:444–445
North East Svalbard Nature Reserve **5**:444
Selous Game Reserve **5**:445
Snares Islands/Tini Heke Islands **5**:444
Tamshiyacu-Tahuayo Communal Reserve **5**:445
threatened species lists and assessments **7**:203–204
threats and challenges **5**:447–448
transboundary protected areas **6**:489, **6**:489*T*
value **5**:448
proteinases **5**:269*T*
protein electrophoresis **6**:183
Protemnodon **4**:727–728
Protemnodon anak **5**:364*F*
Protemnodon brehus **5**:364*F*
Protemnodon roechus **5**:364*F*
Proteobacteria **2**:611*F*, **5**:294
Proteocephalidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Proteocephalus **7**:456–457*F*
proteorhodopsins **6:280–285**
abundance and diversity **6**:282–283
bacteriorhodopsin relationship tree **6**:281*F*
energy circuit diagram **6**:282*F*
Escherichia coli proton pumps **6**:280–282, **6**:281*F*
future research areas
antenna-assisted proteorhodopsins **6**:283
photochemical reaction cycle **6**:283
historical research **6**:280–282
occurrences **6**:280
phototrophic mechanisms **6**:280–282
spectral tuning **6**:283
Protermes **4**:409*T*
Proteroiulus fuscus **5**:428, **5**:429–430
Proterozoic biodiversity **3**:682, **3**:683*F*
Proteus anguinus anguinus **7**:52
prothiocarb **5**:695
protists
biodiversity **5**:288–289, **5**:289*F*
characteristics **5**:305–306
evolutionary development **2**:641
nomenclature code **5**:554–555
plankton fraction **6**:25–26
RNA editing **5**:576
soil fauna **6**:586, **6**:587*F*
protoceratids **4**:701*T*
Protoctista **3**:710*T*
Protocyon **4**:724–725*T*
Protogonyaulax **6**:294
Protohamitermes **4**:409*T*
protohorofauna **2**:582, **2**:583*F*
Protolepis cuprealata **5**:389
Protoperidinium **6**:294–295, **6**:294*F*
Protopterus **3**:461*F*
Protopterus aethiopicus **1**:53–54
Protoschizomidae **1**:212
protosomes **3**:395
protostelids **6**:289–293
Protostomia **2**:7*F*
Prototheoridae **5**:386
protoveratrine **5**:712*T*
Protozoa **6:286–297**
ameboid protozoa **6**:289–293, **6**:289*T*, **6**:290*F*
Antarctica **1**:184*T*
characteristics and properties **2**:524*T*, **4**:392–393, **6**:286–288
ciliated protozoa **6**:289*T*, **6**:295–296, **6**:296*F*
definition **6**:286
domesticated ruminants **1**:710–711
ecosystem functions **6**:288–289
endocommensals **6**:296–297
eutrophication impacts **3**:361–363
flagellated protozoa **6**:289*T*, **6**:293–295, **6**:294*F*
freshwater ecosystems **3**:572*T*
functional roles **6**:286–288
global species richness **3**:711
interspecific competition **2**:182–184, **2**:183*F*, **2**:183*T*, **2**:184*F*, **6**:669–670
known and estimated species **4**:384*T*, **6**:288–289, **6**:289*T*
lake and pond ecosystems **4**:371–373, **4**:374*F*, **4**:460, **4**:461*T*
morphological characteristics **6**:288
nomenclature code **5**:554–555
nonseed gene banks **3**:652*T*
oceanic food web **5**:49*F*
slime molds **6**:289–293, **6**:289*T*, **6**:290*F*
soils
soil fauna **6**:586, **6**:587*F*, **6**:590*T*
soil quality indicators **6**:597
species coexistence **6**:669–670
species-interaction theories **3**:4–5
species richness and diversity
ameboid protozoa **6**:289–293, **6**:289*T*, **6**:290*F*
ciliated protozoa **6**:289*T*, **6**:295–296, **6**:296*F*
endocommensals **6**:296–297
flagellated protozoa **6**:289*T*, **6**:293–295, **6**:294*F*
slime molds **6**:289–293, **6**:289*T*, **6**:290*F*
Protura **4**:308–312, **4**:309*T*, **4**:392*T*
Provitellotrema crenimugilis **7**:450–451*F*
Prunus armeniaca **2**:663
Prunus avium **2**:662–663
Prunus domestica **2**:662–663

Prunus persica **2**:663
Prunus serotina **6**:121
Prymnesium **1**:474–475, **1**:474*F*, **6**:293–294
Prymnesium parvum **1**:474*F*, **3**:357–358
Prymnesium polylepsis **1**:474–475, **1**:474*F*
Przewalski's horse **1**:269
Psalteriomonas **6**:294
Psammobates geometricus **5**:220
Psammobates tentorius trimeni **1**:487–488
Psammotermes **4**:409*T*
Pseduomyrmex concolor **6**:46–47
Pselaphidae **1**:327*T*
Psephotus chrysopterygius **7**:183–184
Psephotus pulcherrimus **7**:183–184
Pseudacris regilla **3**:696
Pseudeuglenes pentatomus **7**:220*T*
Pseudobodo **6**:293
Pseudobradya barroisi **5**:470–471, **5**:470*T*
Pseudocanthotermes **4**:409*T*
Pseudocerastes fieldi **5**:467–469
Pseudochactidae **1**:209
pseudoephedrine **5**:712*T*
pseudoextinction **3**:415, **5**:2
Pseudogymnophallum alcae **7**:450–451*F*
Pseudokuhnia minor **5**:661*F*
Pseudomonas **2**:612*T*, **2**:614–615, **6**:32
Pseudomonas fluorescens **1**:27, **1**:33, **5**:256
Pseudomonas syringae **2**:636*F*, **3**:349
Pseudomyrmex **2**:146
Pseudomyrmex ferruginea **2**:428
Pseudonester xanthophyrs **1**:29*F*
Pseudoneureclipsis palmonii **5**:463*T*
Pseudo-nitzschia **1**:475, **1**:476*F*, **6**:347–348, **6**:348*F*
Pseudonitzschia australis **6**:132
Pseudo-nitzschia pungens **1**:481–482
Pseudophyllinae **3**:727
Pseudopontia paradoxa **1**:650
Pseudopontinae **1**:650
Pseudoryx nghetinhensis **4**:704
Pseudoscorpiones **1**:202*T*, **1**:205*F*, **1**:212–213, **1**:213*F*, **1**:246*T*, **3**:225–226
Pseudosinella christianseni, **7**:55*F*
Pseudotrichomonas **6**:297
Pseudotropheus **1**:25*F*
Pseudotsuga **7**:118
Psidium acutangulum **4**:288*T*
Psidium catleianum **2**:655
Psidium guajava **2**:390, **2**:663–664, **4**:287–290, **4**:288*T*
Psidium guianensis **4**:286–287, **4**:288*T*
Psilidae **3**:489–491*T*, **3**:493–494
psilocin **3**:635*T*
Psilocybe **3**:635*T*
psilocybin **3**:635*T*
Psilotum nudum **3**:206
Psittacidae **6**:308*T*
Psittaciformes **6**:308*T*, **7**:181–182
Psittacus erithacus **1**:49
Psittirostra psittacea **1**:29*F*
Psittirostrini **1**:29
psocids **4**:392*T*
Psocoptera **4**:308–312, **4**:309*T*, **4**:392*T*
Psoralea corylifolia **5**:712*T*
psoralen **5**:712*T*
psychedelic mushrooms **3**:635*T*
Psychodidae **3**:489–491*T*, **3**:493
Psychodomorpha **3**:489–491*T*
Psychotria **6**:316*T*
psychrophiles **6**:298–303
  biodiversity **6**:300
  definition **6**:298
  environmental considerations
    Antarctica **6**:299
    Arctic region **6**:299
    caves **6**:299
    oceans **6**:300
    permafrost regions **6**:299–300
    research background **6**:298–299
    sea ice regions **6**:300
    upper atmosphere **6**:298–299
  evolutionary development **6**:300–301, **6**:302*F*
  historical background **6**:298
  hydrothermal vent ecosystems **7**:328
  phylogenetic relationships **6**:302*F*
  physiological/metabolic activities
    enzymes **6**:301
    growth processes **6**:301
    membranes **6**:301–303
Pteridaceae **5**:11*T*
Pteridomonas **6**:293, **6**:294*F*
Pteridophyllaceae **1**:260*T*
pteridophytes
  abundance patterns **6**:73, **6**:74*F*
  Antarctica **1**:186*T*
  Arctic ecosystems **1**:231*T*
  evolutionary development **6**:65–67, **6**:66*F*
  extinction rates **6**:73*F*
  nonseed gene banks **3**:651
  North America **5**:564–565
  origination rates **6**:71–72, **6**:72*F*
  tropical rain forests **6**:313–318, **6**:316*T*
Pterobothrium **7**:456–457*F*
Pterobranchia **5**:669–670*T*
Pterocarpus **5**:489
Pterocarpus macrocarpus **1**:70
Pterocarpus marsupium **1**:266
Pterocarpus officinalis **1**:164, **7**:272
Pterocarya rhoifolia **1**:271
Pterocladia capillacea **4**:334, **4**:335*F*
Pteroclididae **1**:604
Pterodroma cahow **2**:279, **7**:185
Pterodroma macgillivrayi **1**:620–621
Pterodroma magentae **7**:184
Pterodroma phaeopyia **4**:353
Pterolepis nitidus **3**:478*F*
Pteronura brasiliensis **2**:587*F*, **4**:693
Pterophoroidea **5**:387–388*T*
Pteropoda **4**:386*T*
Pterostemon **5**:622*F*
Pterostylis truncata **3**:211, **3**:213*F*
Pterydophyta **3**:129*T*
Pterygota **4**:308–312, **4**:309*T*
Pterygotrigla polyommata **2**:8–9
Ptilagrostis **1**:270
Ptilidium ciliare **1**:232
Ptilinopus solomonensis **1**:617–618
Ptilocercidae **4**:691
Ptiloris victoriae **2**:75*F*
Ptychocheilus lucius **3**:464
Ptychocheitus lucius **3**:260
Ptychopteridae **3**:489–491*T*
Ptychopteromorpha **3**:489–491*T*
puava **7**:233
public goods **1**:437–438, **5**:145–147
Puccinia graminis **1**:359–360, **3**:633
Puccinia myrsiphylli **4**:362
Puccinia piloselloidarum var. *hieracii* **4**:349
puccinia rust infections **3**:349
Pudica **7**:432–433, **7**:432*T*
Pueraria montana **4**:359–360
Pueraria phaseoloides **6**:560–561
Puerto Rican parrot **6**:212, **7**:184–185
Puerto Rico
  botanical gardens **6**:507*T*
  endemic plant species **6**:83
  lizards **1**:31*F*, **1**:32
  protected area systems **6**:493–494
  termites **4**:432*T*
  urbanization **4**:560*F*
puff adder **6**:409
puffballs **3**:629–630
pufferfish **3**:467–471
puffins
  beak shape **1**:604*F*
  oil spills **3**:256
Puffinus griseus **2**:7–8
Puffinus mauretanicus **7**:182
Pulcherricium caeruleum **3**:629*F*
Pulicaria incisa **5**:485–486
Pull of the Recent **1**:491
pulmonates **4**:464*T*, **5**:460*T*, **5**:464*T*
pulse management **7**:233
pulses
  Europe **2**:662–663
  Mediterranean Basin **2**:662–663
  Neolithic agriculture **2**:658
  North and South America **2**:663–664
  Southeast Asia **2**:663
  southwest Asia **2**:662–663
  Sub-Saharan Africa **2**:663
Puma concolor **1**:709, **2**:587*F*
Puma concolor stanleyana **3**:197
pumas
  conservation practices **7**:229*T*
  Great American Biotic Interchange (GABI) **2**:584*F*
  harvesting threats **2**:590*T*
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  prey–predator interactions **1**:709
pumpkins **2**:383, **2**:663–664, **3**:32*F*, **3**:129–130
puna
  central Andean dwarf *Polylepis* forest **4**:105
  cushion bogs **4**:105
  desert puna **4**:104–105
  dry puna **4**:104
  geographic distribution **4**:96–97, **4**:97*F*
  humid puna **4**:104
  number of species **4**:102*T*
  physical characteristics and patterns **4**:97–98, **4**:103–104
  plant diversity
    endemism **4**:99–100, **4**:99*T*

origins and diversification **4**:100–101
species richness **4**:98–99
representative species **4**:104*F*
salar vegetation **4**:105
punctuated equilibrium **2**:48, **2**:48*F*
pupfish **3**:467
purple bacteria **6**:10*T*, **6**:25–26
purple loosestrife **1**:75–76, **3**:209, **3**:410–411, **6**:457, **7**:399
purple-winged ground-dove **7**:183
purse web spiders **1**:206
purslane **3**:131–132
Pusa caspica **5**:81*T*
Pusa hispida **5**:90
Pycnocycla saxatilis **5**:485
Pycnogonida
arthropod radiation ecology **2**:400
classifications and sampling methods **4**:381*T*
number of species **5**:669–670*T*
phylogenetic relationships **1**:203*F*
Pycnonotus barbatus **5**:474
Pycnonotus barbatus capensis **5**:476
Pycnonotus goiavier **5**:603–604
Pycnonotus jocosus **4**:364
Pycnonotus zeylanicus **7**:187
Pyelorhamphus **5**:363
Pygathrix cinerea **6**:256–258*T*
Pygathrix nemaeus **3**:192, **6**:256–258*T*
Pygathrix nigripes **6**:256–258*T*
pygmy button daisy **4**:194*T*
pygmy grasshoppers **3**:736, **3**:736*F*
pygmy hippopotamus **4**:657, **4**:696–697
pygmy hogs **4**:696
pygmy pine **6**:456–457
pygmy right whales **5**:65*T*
pygmy sperm whales **5**:65*T*
Pynchon, Thomas **4**:653
Pyralidae **5**:387–388*T*
Pyraloidea **5**:386*F*, **5**:387–388*T*, **5**:389–390
pyramid of production **3**:239–240, **7**:265, **7**:265*F*
pyrazon **4**:90*T*
pyrazophos **5**:693–694, **5**:695
Pyrenestes ostrinus **1**:604–605, **3**:144–145
pyrethroids **2**:426*T*, **4**:301, **5**:694, **5**:699*T*
pyrethrum **5**:693
Pyrgomorphoidea (gaudy grasshoppers) **3**:735, **3**:735*F*
Pyrgotidae **3**:489–491*T*
pyridate **4**:90*T*
pyridine
fungicides **5**:695
herbicides **5**:696
pyrifenox **5**:695
pyrimidines **5**:695
pyrimidinyl oxybenzoate **4**:90*T*
Pyrobaculum aerophilum **1**:219–220, **1**:221*T*
Pyrobaculum islandicum **1**:315
Pyrococcus **1**:219–220, **1**:221*F*, **1**:221*T*
Pyrococcus furiosus **1**:593, **5**:703–704
Pyrocystis **6**:30
Pyrodictium **1**:219–220, **1**:221*T*
Pyrodictium occultum **7**:163
Pyrodinium **1**:472–473
Pyrodinium bahamense **1**:472–473
Pyrolobus **1**:219–220, **1**:221*T*
Pyrotag sequencing **2**:13–14
Pyrotheria **2**:583*F*
Pyrrhocoridae **7**:294
Pyrrhophyta **6**:25
Pyrrhus furiosus **5**:709
pyrrolizidine alkaloids **5**:397
Pyrus bretschneideri **2**:663
Pyrus communis **2**:662–663
Pyrus pyrifolia **2**:663
pyruvates **2**:363–364, **2**:363*F*
Pythia **5**:17*F*
Python brongersmai **5**:604
Python curtus **5**:604
Python regius **2**:357
Pyura praeputialis **5**:60–61
Pyura stolonifera **5**:60–61

Q

Qaidam Basin **2**:469–470
Qatar
animal species **1**:261*T*
botanical gardens **6**:507*T*
ecosystems
mangroves **1**:275*T*
threatened species **1**:278*T*
vegetation types and coverage percentage **1**:264–265*T*
energy efficiency measures **3**:254*F*
plant species **1**:259*T*
Qiangtang Nature Reserve **3**:373
qinghaosu **1**:593–594
Qinling golden snub-nosed monkey **6**:256–258*T*
Quabbin Preserve, Massachusetts **3**:313
Quaestora **5**:618*F*
quagga **1**:487–488, **4**:705
quagga mussels **4**:378
quail **1**:644*F*
Qualea grandiflora **1**:165
quality of life **3**:286
quantitative genetics **2**:715
quantitative trait loci (QTL)
basic concepts **6**:4, **6**:5*F*
livestock genetic resources **1**:647
mapping and modeling approaches **2**:724
Quapoya **3**:522*T*
Quararibea cordata **4**:289*T*
Quaternary period
boundary age **5**:1*T*
crustaceans **2**:407
late Quaternary extinction event (LQEE) **4:721–732**
background information **4**:721
causal factors
anthropogenic models **4**:731–732
climatic models **4**:732
global patterns **4**:731–732
comparative studies
Africa **4**:658*F*, **4**:729–730
analytical methods **4**:721–725
Australia **4**:658*F*, **4**:727–728
Eurasia **4**:729–730, **4**:729*F*
Madagascar **4**:658*F*, **4**:728–729
North America **4**:658*F*, **4**:722–725, **4**:724–725*T*, **4**:726*F*
Oceania **4**:730–731
South America **4**:726–727, **4**:727*T*
time-transgressive sequence **4**:722*F*, **4**:723*F*
Levantine biota **5**:460–463
mammals **2**:583*F*
mass extinctions **3**:425*F*, **4**:657*F*
paleoecology **5:645–655**
data collection and interpretation **5**:645–646
goals and objectives **5**:645–646
Quaternary data **5**:645
research methodologies
age models **5**:652–653
dating uncertainties **5**:652
mapping intervals **5**:652–653
scaling factors **5**:646–649, **5**:648*F*, **5**:649*F*
space–time variations **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
spatial characteristics **5**:648*F*, **5**:649–650, **5**:649*F*, **5**:651*F*
taxonomic and numerical characteristics **5**:646–649, **5**:647*F*, **5**:648*F*, **5**:649*F*
temporal characteristics **5**:650–652, **5**:651*F*
temporal resolution **5**:652
zoom lens perspective **5**:647*F*, **5**:648*F*, **5**:649*F*, **5**:651*F*, **5**:653–654
prehistoric human dispersal **4**:721, **4**:722*F*, **4**:723*F*
South America **3**:109–110
vegetation–climate interactions **2**:682–684
Queen Anne's lace **1**:462*F*
Queen Charlotte Islands **5**:563
queen conch **3**:199
Queensland (Australia) **3**:167–168
Queensland Museum (Australia) **5**:406
Quelea quelea **1**:601–602
Quercetea ilicis **3**:340
Quercus **1**:270–271, **2**:551–552*T*, **2**:552*F*, **3**:130*T*, **7**:118
Quercus acatenanguensis **2**:29–30
Quercus alba **5**:390
Quercus boissieri **5**:482
Quercus calliprinos **1**:269, **5**:480*F*, **5**:481–482, **5**:482
Quercus candicans **2**:28–29
Quercus cerris **1**:269
Quercus coccifera **1**:47
Quercus dumosa **4**:173–174
Quercus gambelii **4**:174
Quercus grisea **4**:174
Quercus humboldtii **4**:516–517
Quercus ilex **1**:47, **1**:268, **1**:612, **5**:218
Quercus ithaburensis **1**:269, **5**:480*F*, **5**:482
Quercus liaotungensis **1**:271–272
Quercus libani **5**:482
Quercus mongolica **1**:269, **1**:271
Quercus oleoides **4**:595, **7**:274
Quercus oocarpa **2**:29–30
Quercus petraea **1**:612–613
Quercus robur **1**:269, **5**:723, **6**:106*F*

Quercus rubra **4**:363
Quercus semicarpifolia **1**:272–273
Quercus serrata **1**:271
Quercus skinnerii **2**:28–29
Quercus spp.
  Central America **2**:25–26
  phenological studies **6**:106–107
Quercus suber **1**:47, **1**:709
Quercus variabilis **1**:271–272
Questidae **7**:417*F*
quilworts **6**:59
quinclorac **4**:90*T*
quinidine **5**:712*T*
quinine **2**:426*T*, **2**:427, **5**:711–713, **5**:712*T*
Quinn, David **4**:653
quinoa **2**:475–476, **2**:484*F*, **2**:663–664, **3**:36, **3**:40, **3**:41*F*
quinones **5**:695
Quinqueserialis **7**:450–451*F*
quokka **6**:455

R

rabbitfish **3**:460, **3**:467–471
rabbits
  adaptation rates **2**:140
  biotic relationships **1**:710
  characteristics **4**:697
  deoxyribonucleic acid (DNA) evolution **2**:56*F*
  diversity loss **1**:621
  Great American Biotic Interchange (GABI) **2**:584*F*
  Holocene postglacial spread **2**:551–552*T*
  introduced species
    Australia **1**:80
    diversity loss **3**:193
    South Africa **1**:81
    United Kingdom **1**:79
  number of species **4**:690*T*
  risk status classification **1**:644*F*
  species diversity **4**:697
  wetland ecosystems **7**:390
Rabdophaga rigidae **4**:173–174
Rabdophaga salicisbrassicoides **4**:173–174
Rabdophaga strobiloides **4**:173–174
rabies outbreaks **2**:527–529
rabies virus **2**:551–552*T*
raccoons **1**:686, **2**:584*F*, **7**:390
rachypithecus vetulus nestor **6**:256–258*T*
racoons **4**:690*T*
RADAMBRASIL project **1**:160
radiations **1**:430
radiolarians **6**:25, **6**:289*T*, **6**:290*F*, **6**:292
radionucleotides **3**:118–119
radio waves **7**:297*F*
radishes **2**:383
Radopholus similis **6**:557
Radula flaccid **4**:517
Rafflesia **1**:267
Rafflesia arnoldii **4**:71–72
Rafflesiaceae **4**:71, **4**:72*F*, **6**:315*T*
rafflesia flower **7**:223
Raffles' silvered langur **6**:256–258*T*
ragfish **3**:467–471
ragged tooth shark **3**:459*F*
Raillardiopsis **1**:30*F*
rails **1**:605–606, **1**:621, **7**:181–182, **7**:390
rainbowfish **3**:467
rainbow trout **2**:264, **2**:528*T*, **3**:123*T*, **4**:362, **4**:465, **6**:153*F*
rainfall
  climate change effects **3**:165
  desert climates
    rainfall pulses **2**:463–464
    winter and summer rainfall patterns **2**:464–466, **2**:465*F*
  dryland syndrome **2**:481–482, **2**:483*F*, **2**:484*F*
  ecosystem function measurements **3**:79, **3**:81*F*
  megaherbivores **5**:230
  Near East ecosystems **5**:465–467, **5**:465*F*, **5**:466*T*, **5**:479, **5**:480*F*
rain forests
  Amazon ecosystems **1**:158–160, **1**:159*F*, **1**:159*T*
  characteristics **6**:305–306
  habitat destruction **3**:403–404, **3**:403*F*
  see also tropical rain forests
rainshadow deserts **2**:460
raised bed farming **1**:123–124, **4**:141
Raja binoculata **3**:460*F*
rakkuyu **2**:663
Rallidae **1**:605–606, **6**:308*T*
Rallus longirostrus levipes **6**:454–455, **6**:456–457
Rallus owstoni **1**:621, **1**:664*T*, **1**:666–667, **2**:245
Rallus wakensis **1**:621
rambutan **2**:663
Ramosmania rodriguesiana **6**:81
Ramphastidae **6**:308*T*
Ramsar Convention
  endangered species protection **4**:325, **7**:187
  government legislation and regulations **3**:195, **3**:713*T*
  international agreements **3**:713*T*
  protected area systems **6**:487, **6**:489, **6**:496*F*, **6**:497*F*, **6**:498
  twentieth century conservation **2**:286
  wetland ecosystems **7**:367–368, **7**:386
Ramsaya **4**:727–728
Ramsayia curvirostris **5**:364*F*
Rana boylii **3**:695–696
Rana cancrivora **5**:13
Rana cascadae **2**:76, **3**:165
Rana catesbeiana **1**:78, **3**:167, **3**:695–696, **7**:389–390
Rana clamitans **2**:719
Rana draytonii **3**:695–696
Rana luteiventris **4**:514, **4**:516, **4**:518, **4**:564
Rana pipiens **4**:156, **7**:389–390
Rana temporaria
  environmental threats
    disease **3**:166
    weather pattern shift effects **3**:697
  habitat loss and fragmentation **4**:516–517
  landscape genetics **4**:519
  phenological shifts **5**:723
  phenotypic plasticity **2**:721–722
ranaviruses **3**:166, **3**:696
Rancholabrean mammalian age stage **4**:721–725
Randia **5**:489
random polymorphic DNA (RAPD) fragments **3**:663, **3**:671*F*
random reproduction **2**:508–509, **2**:509*F*
range ecology
  domesticated ruminants **1**:702–703
  global livestock **6:330–344**
    anthropogenic impacts **6**:343
    biological impoverishment
      animal damage control **6**:342
      aquatic ecosystems **6**:335–337, **6**:336*F*, **6**:341–342
      biogeochemical cycling **6**:342
      cropland conversion **6**:341
      desertification **6**:337–338
      direct/primary influences **6**:335–337, **6**:336*F*, **6**:337
      exotic species invasions **6**:341
      fire regime alterations **6**:342–343
      global impact **6**:334–337, **6**:335*T*
      greenhouse effect **6**:340–341
      indirect/secondary and tertiary influences **6**:336*F*, **6**:337
      pasture conversion **6**:338–340, **6**:340*F*, **6**:341
      pathogens **6**:341
      riparian zones **6**:337
      terrestrial ecosystems **6**:335–337, **6**:336*F*
      water diversions **6**:341–342
    biological soil crusts **6**:339
    livestock-caused degradation **6**:343
  historical conservation practices **2**:283
  rangelands
    dryland syndrome **2**:483*F*
    general characteristics **6**:331–333
    grasslands **6**:331–333, **6**:332*F*, **6**:333*F*
    healthy/unhealthy rangelands **3**:311*F*
    historical conservation practices **2**:283
    hot deserts **6**:334
    introduced species
      mammals **1**:76–77, **1**:76*T*
      nonindigenous plants **1**:76
    savannas **6**:331–332, **6**:332*F*, **6**:333*F*
    South America **6**:608
    temperate/cold-desert shrublands **6**:333–334
    tundra **6**:330–331, **6**:331*F*
    weed management
      biodiversity effects **4**:91–92
      chemical weed control **4**:88–89
      general discussion **4**:88
      integrated weed management (IWM) **4**:88
      tools and methodologies **4**:88
      wildland weed management **4**:93–94
Rangifer **4**:724–725*T*
Rangifer tarandus **3**:255, **4**:359, **5**:313, **5**:723, **6**:331
Rangifer tarandus platyrhynchus **7**:27
Ranomafana National Park **4**:437–438
Ransomus rodentorum **7**:432, **7**:432*T*

Ranunculids **5**:621*F*
Ranunculus **4**:101
Ranunculus circinatus **6**:115*T*
Ranunculus flammula **2**:721–722
Ranunculus insignis **3**:209
Ranunculus nivicola **3**:209
Ranunculus verticillata **3**:209
Rao's quadratic entropy **3**:603, **3**:603*F*, **3**:606–608, **3**:607*F*, **5**:205
Rapa Nui **6**:83
rapeseed oil **1**:529*T*, **1**:530, **1**:543–547, **7**:26
Raphanus raphanistrum **2**:720
Raphanus sativus **2**:383
Raphia taedigera **7**:272
Raphidioptera **4**:308–312, **4**:309*T*
Raphus cucullatus **1**:606, **1**:621, **7**:184
rapid heartbeat **3**:635*T*
Rapoport's Rule **2**:588–589, **3**:230–231, **4**:622
rare biosphere **6:345–352**
  abundance considerations **6**:351–352
  basic concepts **6**:345, **6**:346–347*F*
  clone libraries **6**:349–350, **6**:349*T*
  collector curves **6**:349–350, **6**:350*F*
  global distribution **6**:351
  high-throughput sequencing **6**:350–351
  loss mechanisms **6**:351
  microbial species numbers **6**:345–347
  pure cultures versus sequences **6**:348–349
  research summary **6**:352
  species and phylogenetic concepts **6**:347–348, **6**:348*F*
Rare Breeds Survival Trust **1**:645
rarefaction measurement approach **3**:414
rarity
  causal factors **3**:208–209
  characteristics **3**:206
  consequences **3**:209
Rarotonga monarch **7**:184
raspy crickets **3**:732
Rassenkreise theory **7**:46
ratfish **3**:460, **3**:460*F*
rat opossum **4**:687*T*
rats
  introduced species
    Australia **1**:76*T*, **1**:80
    Brazil **1**:83
    India **1**:76*T*, **1**:82
    predation effects **6**:240–241
    South Africa **1**:76*T*, **1**:81
    United Kingdom **1**:76*T*, **1**:79
    United States **1**:76*T*, **1**:77
  invasive species **3**:222
  oil-palm plantations **5**:603–604
rat snakes **6**:408–409, **7**:390
rattlesnakes **6**:409
Rattus **5**:603–604
Rattus exulans **5**:366–367, **6**:240–241
Rattus norvegicus **1**:77, **1**:365–368, **4**:697, **6**:240–241
Rattus rattus **1**:77, **1**:365–368, **1**:621, **5**:366–367, **6**:240–241
Rattus tiomanicus **5**:604
Raukahua **3**:522*T*
Raupo swamp **7**:369*T*
Raup's kill curve **5**:160–163, **5**:163*F*
Rauvolfia canescens **5**:712*T*
Rauvolfia serpentina **5**:704, **5**:706, **5**:712*T*
Rauvolfia spp. **5**:712*T*
Ravenala madagascariensis **1**:55–56
ravens **4**:567
Rawa Biru **1**:276*T*
ray-finned fish **6**:522–523
Ray, John **7**:89–90
rays **3**:456–457, **3**:459–460, **7**:203*F*
rctocephalus philippii **5**:81*T*
reaction–diffusion models **5**:324–325, **5**:325*F*
reactive oxygen species (ROS) **3**:119
Reaumuria **1**:270
Reaumuria hirtella **5**:484–485
Reaumuria negevensis **5**:484–485
Reaumuria palestina **1**:268
reciprocal altruism **2**:324–325
reciprocal evolution **3**:7
recombination **6:353–369**
  biodiversity effects
    domestication evolution **6**:361–362
    eukaryotic genomes **6**:360–361
    genome-based recombination rate **6**:360–361, **6**:361*T*
    intraspecific variations **6**:361
  chiasmata **6**:353
  classifications
    crossing-over recombination **6**:353, **6**:354*F*
    ectopic recombination **6**:354
    extranuclear recombination **6**:355
    general discussion **6**:353
    horizontal transfer **6**:354–355
    illegitimate recombination **6**:354–355
    mitotic recombination **6**:353, **6**:354*F*
    nonhomologous chromosome segregation **6**:353
    nonreciprocal exchange **6**:353–354
  crossing-over recombination
    classifications **6**:353, **6**:354*F*
    crossover distribution **6**:356
    environmental effects
      feedback regulation **6**:357
      genetic variation **6**:356–357
      ontogenetic memory **6**:357
      transgeneration effect **6**:357
    genetic interference **6**:355–356
    map distance **6**:355
    mapping functions **6**:356
    recombination rate **6**:355
    sex effects **6**:356
  eukaryotic genomes **6**:361*T*
  evolutionary development
    experimental evolution studies
      adaptive trait selection **6**:365
      increased recombination **6**:366, **6**:366*F*
      recombination rate reduction **6**:366
      stress adaptations **6**:365–366
    life history traits **6**:365
    theoretical assumption tests
      background information **6**:364–365
      final effects **6**:365
      initial assumptions **6**:364–365
    theoretical models **6**:363–364
  evolutionary effects
    evolutionary novelties **6**:362–363
    genetic adaption **6**:362
    horizontal gene transfer **6**:363
    sequence polymorphism **6**:363
  genetic mechanisms and control
    DNA repair and recombination **6**:358
    genetic modifiers **6**:357–358
    hierarchical controls **6**:360*T*
    meiotic mutations **6**:358
    molecular mechanisms **6**:359
    recombination hotspots **6**:359–360
    recombination nodules (RNs) **6**:358–359
    selection processes **6**:357–358
    synaptonemal complex (SC) ultrastructure **6**:358–359
    transcription promoters **6**:360
  homologous recombination **1**:322–324, **1**:324*F*
  recombination rate
    genome-based recombination rate **6**:360–361
    intraspecific variations **6**:361
  research applications
    artificial sequence optimization **6**:367
    cancer treatment
      chromosome translocations **6**:367
      DNA repair and recombination **6**:367
      mitotic recombination **6**:367
    genetic hybridization/breeding **6**:366–367
    genetic mapping **6**:366
    targeted gene replacement **6**:367–368
  research summary **6**:368
recombinational hybrid speciation **6**:626
recreational activities
  habitat alteration **4**:155
  human flourishing **3**:273, **3**:273*F*
  nature-based recreation **7:241–257**
    backpacking/hiking groups
      conservation support **7**:251–252
      per capita contributions **7**:254*F*
      popularity **7**:251
    biophilia versus videophilia
      comparison studies **7**:242–244
      future outlook **7**:250–251
      future research outlook **7**:253–256, **7**:255*F*
      historical trends **7**:242, **7**:243*F*
      videophilia **7**:244
    childhood health and development
      children–natural environment interactions **7**:248
      early television viewing–attention deficit disorder (ADD) relationship **7**:248–249
      future outlook **7**:250–251
      future research outlook **7**:253–256, **7**:255*F*
      indirect, direct, and vicarious experiences **7**:247
      Internet use impacts **7**:249
      media depictions–fear relationship **7**:249

recreational activities (*continued*)
recommended television-watching guidelines **7**:248
sedentary lifestyles **7**:247–248
socialization impacts **7**:249
virtual nature experiences **7**:249–250
conservation strategies
ecosystem services **7**:253
health benefits **7**:253
conservation support
backpacking/hiking groups **7**:251–252
declining support **7**:251
demographics **7**:252–253
diversification strategies **7**:252
future outlook **7**:252–253
popular activities **7**:251
public land visits **7**:251–252
two-Americas concept **7**:251–252
declining trends **7**:241–242
future research outlook **7**:253–256, **7**:255*F*
habitat degradation **4**:566–567
importance **7**:241–242
modeling frameworks **5**:353
national and international trends
decline incidence **7**:244
international trends **7**:245, **7**:246*F*
per capita visitation **7**:244–245, **7**:245*F*, **7**:246*F*
pervasive declines **7**:247
popularity declines **7**:245–246, **7**:245*F*, **7**:246*F*
public surveys **7**:246
specific-activity trends **7**:246–247, **7**:246*F*
national park visits **7**:242, **7**:243*F*
twentieth century conservation **2**:283
recreational fishing **2**:3*F*
recruitment overfishing **3**:445
Recurvirostra avocetta **5**:476–477
Recurvirostra spp. **1**:710
red algae
evolutionary development **2**:641–642
freshwater ecosystems **3**:571–572, **3**:571*T*
marine ecosystems **5**:48
red-backed shrike **5**:312
red-backed voles **1**:631–632
red-bellied monkey **6**:256–258*T*
red-bellied newts **5**:315–316
red-billed dioch **1**:601–602
red-breast sunfish **4**:517
red clover **2**:139
red-cockaded woodpecker **1**:397, **2**:264, **7**:183–184
Red Cross **4**:330
red-crowned cranes **3**:373, **7**:188–189
red deer **2**:551–552*T*, **2**:719–720, **3**:127*T*, **4**:516, **5**:723
red-eared slider turtles **3**:218
red fescue **2**:390
Redfield ratio **2**:617–618, **2**:617*F*, **3**:506–507, **4**:675–677, **4**:676*F*, **5**:27, **6**:471
red fire ants **1**:78
red fox **1**:631–632, **2**:528*T*, **2**:590*T*, **5**:475
red-headed pine sayfly **4**:182–183
Redi, Francesco **5**:629–630
red jungle fowl **1**:640
red knots **7**:185
red land crab **3**:222, **7**:260
red-legged frogs **3**:695–696
Red List Index (RLI)
background information **4**:327
categories and criteria **7**:196–197, **7**:197
endangered species protection **3**:223–224, **4**:325–326
future extinction estimations **4**:660
geographic distribution **7**:198*F*
global species extinction risk assessments **7**:194–195
historical perspective **7**:192
major biodiversity threats **4**:314–315, **4**:314*F*
present extinctions **2**:169, **2**:170*F*
regional and national extinction risk assessments **7**:195–197
resource allocations **7**:207
threatened species
amphibians **1**:176–177, **1**:515
angiosperms **6**:60
birds **7**:180–181
coral reef ecosystems **4**:664
Denmark **3**:579–580
Europe **3**:578–579, **3**:579*T*
fish **3**:451–452
future extinction estimations **4**:660
future outlook **7**:208–209
gymnosperms **6**:60
indicator species development **7**:205–207, **7**:206–207*F*
insects **1**:514–515
mammals **1**:516–517, **3**:194–195, **3**:195*F*, **4**:706–707, **4**:709
North America **3**:578–579, **3**:579*T*
plants **6**:84–85
primates **6**:255, **6**:256–258*T*, **6**:258*T*
reptiles **1**:515–516
seedless vascular plants **6**:59
terrestrial invertebrates **3**:225
United States **3**:579–580
red mangrove **7**:397
red maple **6**:433*F*
red moki **5**:33
red mullet **2**:356
red-nosed bearded saki **6**:256–258*T*
red oak **4**:363
redox reactions **1**:555–556, **1**:556*T*
red palm oil **5**:602
Red Queen hypothesis **2**:56, **2**:143, **5**:435–436, **5**:704, **6**:70
Red River Delta **1**:276*T*
Red River estuary **1**:276*T*
red ruffed lemur **6**:256–258*T*
red-rumped swallow **5**:474
red salmon **6**:523
Red Sea coastal desert **2**:467–468, **2**:469–470
red-shanked douc **6**:256–258*T*
red shiner **4**:364
red-sided garter snakes **5**:315–316
red-spotted newts **1**:172
red spruce **1**:6–8, **1**:8*F*, **3**:440
red squirrels
boreal forests **1**:631–632
hybridization **3**:191
interspecific competition **2**:190
introduced parasites–species decline relationships **2**:529–530, **2**:529*F*
resource competition **4**:360
whitebark pine forests **2**:571
red swamp crayfish **3**:696
red-tailed hawks **4**:516
red tides **3**:357, **3**:449
Reducing Emissions from Deforestation and Forest Degradation (REDD)
background information **4**:328–329
Brazil **7**:170–172
deforestation **1**:682–683, **2**:246–247, **2**:455, **4**:125
international awareness efforts **4**:125
market-based approaches **5**:152
plant conservation **6**:86
protected area systems **1**:521, **6**:500
Southeast Asia **1**:513, **1**:521
sustainable forest management financing **7**:139–140
threatened species **7**:187
Reduncinae **4**:696
Reduviidae **7**:292–293, **7**:293–294
red wolf **2**:268–269
redwoods
North America **7**:118
tourism industry **7**:223
reedbucks **4**:696, **6**:572–573
reed canary grass **1**:529*T*
reedfish **3**:461–462
reed grass **7**:385
reedmace swamp **7**:369*T*
reeds **1**:278–279, **2**:543, **6**:115–116
reedswamp **7**:369*T*
reef fish **2**:12*F*
reflective beaches **2**:128–129, **2**:129*F*, **2**:130*F*, **2**:135, **5**:51
reforestation **6**:**370–379**
Australian reforestation programs
mining impacts **6**:378
North Queensland **6**:377
Brazilian reforestation programs **6**:377–378
challenges **6**:371
ecosystem recovery **6**:374–375
general discussion **6**:370
methodologies
direct seeding **6**:372, **6**:372*F*
natural recovery processes **6**:371–372, **6**:371*F*
planting densities **6**:373
seedling planting **6**:372
species richness **6**:373–374, **6**:374*F*
species selection **6**:372–373, **6**:373*T*
needs **6**:370
Nepal **6**:378–379
New Zealand reforestation programs **6**:377
post-degradation methods **6**:370–371, **6**:371*F*
priority area selection **6**:375, **6**:375*T*

research summary **6**:376
socioeconomic factors **6**:376
specialized conditions/problem areas **6**:375–376
United Kingdom reforestation programs **6**:376–377
refuge **1**:39
reg **2**:462
Regalecus glesne **3**:466*F*
Regional Action Plan for Amazonian Biodiversity **2**:314
regional diversity **6**:534
Regosols **2**:461–464
regression analysis **1**:698–699
regulated bionomic equilibrium **2**:151–152
regulating services **3:45–54**
characteristics **3**:47*F*
definition **3**:45–46
global status **3**:45–46, **3**:46*T*
private benefits versus public benefits **3**:51–53, **3**:52*F*
research summary **3**:53–54
spatially variable ecosystem services **3**:53
structure–function relationship **3**:46–47, **3**:47*F*, **3**:47*T*
Thailand (case study) **3**:51–53, **3**:52*F*
valuation measures
challenges **3**:47–48, **3**:48*F*
nonmarket regulating services **3**:48–51, **3**:49*T*
Regulidae **6**:308*T*
Reid's paradox **2**:550
reindeer
introduced species **4**:359
migratory patterns **5**:313
tundra ecosystems **6**:330–331
reintroduction programs
mammal conservation **3**:196–197
translocations **7**:236–237
reishi mushrooms **3**:637*T*
Reissia **7**:347*F*
Reithrodontomys megalotis **4**:43*F*
relative tidal range (RTR) **2**:129
relevé sampling method
basic concepts **4**:399
cluster analysis **4**:403*F*, **4**:404*F*
predefinied strata **4**:399
size determinations **4**:399, **4**:400*F*
species determinations **4**:399
species quantity estimations **4**:399–400
relief organizations **4**:330
religious traditions
biodiversity conservation **6**:563–566
biodiversity protection **6:380–388**
Asian traditions **6**:383–384
background information **6**:380
basic concepts **6**:380–381
Buddhist declaration on nature **3**:272, **6**:383
Christian declaration on nature **3**:272–273, **6**:381, **6**:381
Hindu declaration on nature **3**:273, **6**:383
Jewish declaration on nature **3**:273, **6**:381, **6**:382
monotheistic traditions
Christian declaration on nature **3**:272–273, **6**:381, **6**:381
Jewish declaration on nature **3**:273, **6**:381, **6**:382
Judeo–Christian tradition **6**:381
Muslim declaration on nature **3**:273, **6**:381, **6**:382
stewardship **6**:381–383
Muslim declaration on nature **3**:273, **6**:381, **6**:382
pantheistic traditions
ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
indigenous peoples **6**:384, **6**:384–385
traditional ecological knowledge **6**:385
traditional social restraints **6**:385–386
research summary **6**:387–388
stewardship **3**:272, **6**:381–383
traditional social protection
background information **6**:385–386
biological communities/habitat patches **6**:386
critical life history stages **6**:386
long-lived (K-selected) species **6**:386
stewardship **6**:386
Remanella **6**:296*F*
Remipedia **2**:400–401, **2**:401*T*, **2**:416–417*T*
Remirea maritima **1**:55
remoras **3**:467–471
remote sensing **6:389–399**
advantages **6**:389
applications
canopy gap dynamics **6**:398
fire/fire regimes **6**:398, **6**:399*F*
general discussion **6**:395–397
grazing **6**:398
species richness and distribution **6**:398, **6**:398*F*, **6**:398*T*
vegetation and land-cover patterns **6**:396–397, **6**:396*F*
vegetation change measurement **6**:397–398, **6**:397*F*
basic concepts **6**:389–390, **6**:390*F*
data sources
analytical methods
general discussion **6**:394–395
geographical information systems (GIS) **6**:395
image classification **6**:395
interpretation process **6**:394–395
satellite imagery
characteristics **6**:392–393
coarse-resolution sensors **6**:394
emerging technologies **6**:394
Landsat program **6**:393, **6**:393*T*
SPOT program **6**:393–394
very high-resolution images **6**:392, **6**:392*F*
electromagnetic radiation
electromagnetic spectrum **6**:390–391, **6**:391*F*
measurement and analysis **6**:390–391
sensor systems **6**:391–392
global land cover map **6**:390*F*
research summary **6**:398–399
tropical rain forests
current rainforest areas
1990 remote sensing survey **6**:323, **6**:324*T*, **6**:325*T*
2000 remote sensing survey **6**:323–325, **6**:325*T*
multidate remote sensing survey **6**:322–323
remote sensing techniques **2**:358, **4**:509–510
renewable energy
biodiversity impacts **3**:260–262, **3**:264*T*
carbon dioxide (CO_2) emissions **3**:265*T*
land requirements **3**:265*T*
wildlife habitat restoration **1:540–551**
cropping system comparisons **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
government policies and regulations **1**:540–541
land allocations
global-scale assessments **1**:547–549
landscape-scale assessments **1**:547–549, **1**:548*F*
local-scale assessments **1**:549
regional-scale assessments **1**:547–549
land-use change impacts
croplands and pasturelands **1**:542–543, **1**:542*F*
marginal, degraded, and abandoned lands **1**:543, **1**:544*F*
native ecosystems **1**:541–542, **1**:545*F*
prime agricultural lands **1**:543
oil-palm plantations **1**:540–541
summary discussion **1**:549
world energy reserves **3**:265*T*
world energy use **3**:251, **3**:252*T*, **3**:253*T*
Renewable Fuel Standard (RFS2) **4**:539–540
Renggeria **3**:522*T*
renosterveld **2**:300, **5**:219
Rensch, Bernhard **3**:394, **7**:46
Reoviridae **5**:134*T*, **5**:137
reproductive revolution **6**:202
reproductive success **6**:632–633
Reptilia/reptiles **6:400–413**
active reptiles **6**:404–405
adaptive radiation **1**:22, **1**:23*F*
alpine ecosystems **1**:154–155
ancestral and derived characters
characteristics **6**:402–403
metabolism
aerobic/anaerobic capacity **6**:404
body size **6**:404, **6**:404*F*
conversion efficiency **6**:403–404
energy requirements **6**:403–404
resting metabolic rate **6**:403–404, **6**:404*F*
sex determination **6**:402–403, **6**:410
temperature requirements **6**:403
water balance **6**:403
anthropogenic activities and impacts **7**:307*T*
Asia
dry Asian regions **1**:268
endemic species **1**:261*T*
threatened species **1**:278*T*
tropical rain forests **1**:263

Reptilia/reptiles (*continued*)
Central America **2**:25*T*
characteristics **7**:339–340
climate change studies **2**:66, **2**:92–93, **6**:410
conservation issues
climate change effects **6**:410
crocodilians **6**:411
declining populations **6**:409–410
Lepidosauria **6**:411
turtles **6**:410–411
desert ecosystems **2**:470*T*, **2**:472–474, **2**:473*F*
elevational range shifts **4**:605*T*
El Niño Southern Oscillation (ENSO) events **3**:156–158
endangered species **3:216–218**
characteristics and classifications **3**:216
conservation status **3**:216, **3**:216*T*
major threats
climate change effects **3**:218, **6**:410
commercial exploitation **3**:217
direct impacts **3**:216–217
habitat destruction **3**:216–217
introduced species **3**:217–218
pollution **3**:218
endemism **3**:232*T*
endothermy evolution **6**:404–405
extant evolutionary lineages
crocodilians **6**:401
Lepidosauria **6**:401
rhynchocephalians **6**:401
species diversity **6**:400–401
Squamata **6**:401–402
turtles **6**:400–401
forest canopies **3**:513–514
fossil record **3**:685*F*
freshwater ecosystems **3**:571–572, **3**:578–579, **3**:579*T*
fruit and seed dispersal **6**:50
genetic diversity **3**:665*F*
geographical variation **1**:614
global wildlife trade **3**:409*T*
grassland ecosystems **4**:3
group characteristics
crocodilians **6**:405–406
general discussion **6**:405
Lepidosauria
Squamata **6**:406–408
tuatara **6**:406
lizards **6**:406–408, **6**:407*T*
snakes **6**:408–409
turtles
gigantism **6**:405
hatchling characteristics **6**:405
nest environment **6**:405
introduced parasites–species decline relationships **2**:528*T*
introduced species
Australia **1**:80
United States **1**:78
late Quaternary extinction event (LQEE) **4**:721
major threats **7**:203*F*
mangrove ecosystems **5**:13, **5**:20*T*
Mediterranean-type climate ecosystems
Australia **5**:213*T*, **5**:221
California **5**:213*T*
Cape Region (South Africa) **5**:213*T*, **5**:220
Chile **5**:213*T*, **5**:217
Mediterranean Basin **5**:213*T*, **5**:218–219
migratory patterns **5**:315–316
modern extinctions **5**:368*T*
Near East ecosystems **5**:460*T*, **5**:464*T*
North America **5**:565
Permian–Triassic mass extinction **5**:4
phylogenetic relationships **5**:1*F*, **6**:400
river ecosystems **6**:479–480
seagrass ecosystems **6**:547–548
Southeast Asia **1**:515–516
species richness **6**:708*F*
species richness gradients **4**:613*F*
species richness maps **6**:708*F*, **6**:709*F*
Succulent Karoo **1**:487–488, **1**:487*T*
threatened species **1**:515–516, **7**:201*T*
tourism industry **7**:223
United States **1**:78
urbanization threats **4**:564
wetland ecosystems **7**:389–391
wildfire effects **3**:156–158
rescinnamine **5**:712*T*
res communes **6**:274
reserpine **2**:432, **5**:712*T*
resistance surface analysis **4**:511–512, **4**:513*F*
res nullius **6**:274
resource competition **4**:360
resource concentration hypothesis **6**:641
Resource Description Framework (RDF) **1**:444
resource management
ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
indigenous peoples **4**:272, **6**:384–385, **6**:384, **6**:566–568
traditional belief systems **6**:386, **6**:567–568
resource partitioning **6:429–437**
Clements–Gleason debate **6**:429
community dominance–diversity relationships **6**:431–432, **6**:432*F*
diversity assessment framework **6**:429–430, **6**:430*F*
diversity mechanisms
chance **6**:435
frequency-dependent effects **6**:435
general discussion **6**:435
spatial structure **6**:435–436, **6**:436*F*
niche differentiation
empirical evidence
relevance **6**:434–435
spatial heterogeneity **6**:432–433, **6**:433*F*
temporal heterogeneity **6**:433–434, **6**:434*F*
theoretical perspectives **6**:430–432, **6**:431*F*, **6**:432*F*
research background **6**:429
research summary **6**:436–437
respiration, microbial **5**:301*T*, **5**:302*F*
res privatae **6**:274–275
res publicae **6**:274
restingas **1**:161, **6**:600–604*T*
Restionaceae **5**:219
restoration **6:438–452**
active versus passive restoration **6**:439, **6**:439*F*
animal restoration
basic concepts **6**:443–444
habitat restoration **6**:445–446
population sources **6**:446
site assessments **6**:445–446
species reductions or absences **6**:444–445
biodiversity restoration **6:453–460**
biologically diverse landscapes
landscape objectives **6**:458
novel/experimental opportunities **6**:458–459
regional efforts **6**:458
community restoration
abiotic constraints **6**:456–457
basic concepts **6**:456–457
biotic constraints **6**:457
management solutions **6**:457–458
genetic considerations
captive breeding programs **6**:454
challenges **6**:454
gene pool manipulation **6**:454
greenhouse propagation **6**:454
guidance **6**:454
research background **6**:453–454
goals and priorities **6**:453
single species reintroduction
abiotic constraints **6**:455
biotic constraints **6**:455
management solutions **6**:455–456
purposes **6**:454–455
definition **6**:438
designer ecosystems **6**:451
ecological restoration **4:199–208**
advantages/disadvantages **4**:199–200
basic concepts **4**:199–200
ecosystem restoration potential **4**:203–204
explicit restoration targets **4**:205
future outlook
biodiversity crisis **4**:205
key challenges **4**:205–206
socioeconomic justifications **4**:206
plant conservation **6**:63, **6**:85–87
restoration by-products **4**:204
restoration justification **4**:204–205
restoration targets
characteristics **4**:200–201
ecosystem drivers **4**:201
ecosystem service restoration **4**:201–202
entire ecosystems **4**:200–201, **4**:202*T*
explicit restoration targets **4**:205
global climate change **4**:203
novel ecosystems **4**:202–203
whole-ecosystem restoration **4**:201
goals and limitations **6**:438–439
microorganisms
diversity evaluations **6**:446–447
establishment criteria **6**:450–451, **6**:450*F*

functional groups
free-living saprobes **6**:448
importance **6**:448
soil food webs **6**:448–449, **6**:449*F*
symbionts **6**:449–450, **6**:450*F*
spatiotemporal relationships **6**:447–448, **6**:447*F*
plant communities
dominance–diversity curve **6**:440*F*
limitations **6**:440–442, **6**:441*F*
single-species introductions **6**:442–443
species richness improvements **6**:442
succession **6**:443
vegetation management **6**:439–442, **6**:440*F*
success assessments **6**:451
restorative responses **1**:43
restriction fragment length polymorphisms (RFLPs) **3**:663, **5**:264*T*, **5**:265
Retama **1**:268
Retama raetam **5**:484–485
reticulate allogenous speciation **6**:627–628
Reticulitermes **4**:409*T*
Reticulitermes flavipes **4**:417, **5**:283*F*
Reticulitermes speratus **4**:417
Reticulomyxa **6**:291
Retortamonas **6**:293
Retronects **7**:442–443*F*
return-on-investment (ROI) analysis **4:185–198**
analytical examples
action and location allocation **4**:193–194, **4**:195*F*, **4**:196*T*
action and location prioritization **4**:192–193, **4**:193*F*
allocation strategies **4**:193
budget allocations **4**:195*T*
cost–risk–complementarity relationship **4**:195–196
dependent returns **4**:194–195, **4**:197*F*
diminishing returns **4**:190–191, **4**:190*F*, **4**:191*T*
land purchase prioritization **4**:193, **4**:195*T*
restoration strategy prioritization **4**:190, **4**:190*F*
threat considerations **4**:191–192, **4**:192*F*
threatened species recovery projects **4**:193, **4**:194*T*
unknown costs **4**:196–197
basic concepts **4**:185–186
challenges **4**:197
conservation prioritization applications **4**:186, **4**:187*T*
conservation return
basic concepts **4**:188–189
benefit functions **4**:188–189, **4**:189*F*
diminishing returns **4**:188–189
probability of success **4**:189
cost data types **4**:190*F*
framing problems **4**:187–188
investment estimation **4**:189
valuation measures **4**:186–187
Reunion **4**:560*F*, **6**:519
revealed-preference estimation methods **7**:322–324
Rhabdadenia macrostoma **1**:165
Rhabditida **4**:394*F*
rhabditid nematodes **7**:431
Rhabdodendron **5**:622*F*
Rhabdoviridae **5**:130*F*, **5**:134*T*, **5**:139
Rhagamys orthodon **1**:702
Rhagionidae **3**:489–491*T*
Rhagoletis pomonella **2**:140, **2**:725, **6**:625–626
Rhamnaceae **5**:623*F*
rhamnose **6**:113*T*
Rhamnus disperma **5**:485
Rhamnus frangula **7**:399
Rhamnus lycioides **5**:481–482
Rhamnus punctata **5**:485
Rhamnus purshiana **5**:712*T*
Rhamphochromis **1**:25*F*
Rhaphidophora **3**:522*T*
Rhaphidophoroidea **3**:732
Rhaphiomidas terminatus abdominalis **3**:496
Rhea americana **3**:115
rheas
deoxyribonucleic acid (DNA) evolution **2**:56*F*
flight constraints **1**:606
Rheedia brasiliensis **1**:163, **4**:288*T*
Rheedia macrophylla **4**:288*T*
Rheidae **1**:606
Rheocyclops carolinianus **3**:185*F*
rhesus macaques **6**:253
Rheum **1**:270
Rhigosum brevispinosum **1**:67
Rhigosum virgatum **1**:67
Rhincodon typus **7**:338
Rhinebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Rhinebothrium megacanthophallus **7**:456–457*F*
Rhine River **7**:152–153
rhinoceros
characteristics **4**:695
convergent evolution **4**:701*T*
deoxyribonucleic acid (DNA) evolution **2**:56*F*
extant species **5**:225*T*
mangrove ecosystems **5**:14
number of species **4**:690*T*
species diversity **4**:695
Succulent Karoo **1**:487–488
threat percentage **3**:195
see also megaherbivores
Rhinoceros sondaicus **4**:695, **5**:225*T*
Rhinoceros unicornis **4**:695, **5**:225*T*
Rhinocerotidae **5**:224*T*, **5**:225*T*
Rhinochenus stigma **4**:595
Rhinocyllus conicus **4**:361, **4**:363
Rhinophoridae **3**:489–491*T*
Rhinopithecus avunculus **6**:250, **6**:256–258*T*
Rhinopithecus bieti **6**:256–258*T*
Rhinopithecus brelichi **6**:256–258*T*
Rhinopithecus roxellana hubeiensis **6**:256–258*T*
Rhinopithecus roxellana qinlingensis **6**:256–258*T*
Rhinopithecus roxellana roxellana **6**:256–258*T*
Rhinopithecus strykeri **6**:259
Rhino Protection Units (RPUs) **1**:521*F*
Rhinotermes **4**:409*T*
Rhinotermitidae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:422*F*, **4**:430*T*
Rhinotitan **5**:224*T*
Rhipicephalus appendiculatus **2**:103–104
Rhizobium **3**:122*T*, **6**:448, **6**:581–582, **6**:582*F*
Rhizocephala **2**:416–417*T*
Rhizophora **1**:277, **5**:11*T*
Rhizophoraceae **1**:277, **5**:11*T*
Rhizophora harrisonii **1**:164
Rhizophora mangle **1**:164, **3**:116, **7**:397
Rhizophora racemosa **1**:164
rhizopod amebae **6**:289*T*, **6**:290–291
Rhizopsammia wellingtoni **3**:202*T*
Rhizopus **3**:627, **3**:628*F*, **3**:634*T*
Rhizopus nigricans **3**:638
Rhizosolenia **2**:611*F*, **2**:612*T*
Rhizosolenia alata **4**:386*T*
rhizosphere **6**:111–113, **6**:112*F*, **6**:113*T*
Rhododendron **1**:270, **1**:272–273, **6**:316*T*, **7**:118
Rhododendron ponticum **6**:107
Rhodomonas **6**:29
Rhodomonas pusilla **6**:33*F*
Rhodophiala rhodolirion **4**:106
Rhodophyta **3**:571*T*, **6**:10*T*
Rhodospata **3**:522*T*
Rhodospirillaceae **6**:10*T*, **6**:11*T*, **6**:14*T*
Rhodospirillum **6**:32
Rhoiptelaceae **1**:260*T*
Rhone River (France) **7**:59
Rhopalocera **5**:460*T*, **5**:463*T*
Rhopilema nomadica **3**:284
Rhus tripartita **5**:485
rhynchocephalians **6**:401, **7**:340–341
Rhynchocyon chrysopygus **4**:697–698
Rhyncholestes raphanurus **2**:587*F*
Rhynchomonas **6**:293, **6**:294*F*
Rhynchomonas nasuta **6**:288–289
Rhynchophorus palmarum **5**:603–604
Rhynchophorus phoenicis **5**:603–604
Rhynchopsitta pachyrhyncha **7**:183
Rhynchotermes **4**:409*T*
Rhynchotherium **5**:224*T*
Rhyncotherium **4**:724–725*T*
Rhynochetos jubatus **1**:608
Rhytisma acerinum **1**:139
Rhyzopertha dominica **6**:670
Ribautiella amazonica **5**:431
ribbonfish **3**:466
ribbon seal **5**:90
ribbon worms
classifications **4**:381*T*
number of species **4**:371*T*
Riberoia ondatrae **2**:528*T*
Ribes **5**:622*F*
ribonucleic acid (RNA)
chromosomal and gene mutations **3**:658*T*
evolutionary perspective **5**:703–705
genetic code **3**:656–657, **3**:657*T*
information transfer **3**:656–657, **3**:656*F*
marine viruses
dsRNA viruses

ribonucleic acid (RNA) (*continued*)
Birnaviridae **5**:134*T*, **5**:137
characteristics **5**:134*T*, **5**:137
Reoviridae **5**:134*T*, **5**:137
negative-sense ssRNA viruses
Bunyaviridae **5**:130*F*, **5**:134*T*, **5**:139–140
characteristics **5**:134*T*, **5**:139
Mononegavirales **5**:134*T*, **5**:139
Orthomyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Paramyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Rhabdoviridae **5**:130*F*, **5**:134*T*, **5**:139
positive-sense ssRNA viruses
Alvernaviridae **5**:134*T*, **5**:138
Astroviridae **5**:134*T*, **5**:138
Bacillarnavirus **5**:134*T*, **5**:138
Caliciviridae **5**:134*T*, **5**:138
characteristics **5**:134*T*, **5**:137
Dicistroviridae **5**:134*T*, **5**:137
Labyrnaviridae **5**:134*T*, **5**:138
Marnaviridae **5**:134*T*, **5**:137–138
Metaviridae **5**:134*T*, **5**:140
Nidovirales **5**:134*T*, **5**:137
Nodaviridae **5**:134*T*, **5**:138–139
Picornavirales **5**:134*T*, **5**:137
Roniviridae **5**:134*T*, **5**:137
Togaviridae **5**:134*T*, **5**:139
RNA-containing viruses **5**:142
nucleic acids **3:568–580**
nucleotide sequences **5**:568, **6**:183
RNA editing
acceptor stems **5**:576–577
A–>I deamination transitions **5**:577–578
anticodons **5**:576–577
apolipoprotein B **5**:577
basic principles **5**:570–572
C–>U and U–>C editing **5**:575–576, **5**:575*F*
guide RNAs (gRNAs) **5**:572, **5**:575*F*
kinetoplastids **5**:572, **5**:573–574*T*
mitochondrial tRNA **5**:576–577
myxomycetes **5**:572–575, **5**:573–574*T*, **5**:575*F*
paramyxoviruses **5**:573–574*T*, **5**:578
plant mitochondria **5**:573–574*T*, **5**:575–576, **5**:575*F*
types and distribution **5**:570–572, **5**:573–574*T*
U insertion **5**:572–575, **5**:575*F*
origin of life theories **5**:633–636, **5**:635*F*
ribosomal RNA (rRNA) sequences **2**:693–697, **2**:694*F*
ribosomal subunits **5**:259–260
RNA editing
acceptor stems **5**:576–577
A–>I deamination transitions **5**:577–578
anticodons **5**:576–577
apolipoprotein B **5**:577
basic principles **5**:570–572
C–>U and U–>C editing **5**:575–576, **5**:575*F*
guide RNAs (gRNAs) **5**:572, **5**:575*F*
kinetoplastids **5**:572, **5**:573–574*T*
mitochondrial tRNA **5**:576–577
myxomycetes **5**:572–575, **5**:573–574*T*, **5**:575*F*
paramyxoviruses **5**:573–574*T*, **5**:578
plant mitochondria **5**:573–574*T*, **5**:575–576, **5**:575*F*
types and distribution **5**:570–572, **5**:573–574*T*
U insertion **5**:572–575, **5**:575*F*
structural characteristics **3**:655–656, **3**:655*F*
ribosomal RNA (rRNA) sequences **2**:693–697, **2**:694*F*
ribulose-1,5-bisphosphatecarboxylase–oxygenase (RUBISCO) **6**:13–15, **6**:14*T*
ribulose-1,5, bisphosphate (RuBP) **2**:361–362, **2**:362*F*
Ricardoestesia **2**:519*T*, **2**:520–521, **2**:520*T*
rice
crop production **1**:86, **1**:388–389
genetic diversity
commercialization effects
general discussion **1**:91–92
intensified environments **1**:92
non-intensified environments **1**:92–93
genetic evolution **1**:86–87
genetic narrowing **1**:89–90
genetic trends **1**:90
landrace–modern varieties comparisons **1**:89
remaining gene pool percentage **1**:132*T*
research summary **1**:93
intercropping patterns **2**:384*F*
native crops
East Asia **2**:663
Southeast Asia **2**:663
plant conservation **6**:85–86
regional cultivars **3**:32*F*
rice paddies **1**:279, **4**:141
seed industries **1**:91
spatial distribution **1**:87–89, **1**:88, **1**:88*T*
standards of consumption **3**:128
rice beans **2**:663
rice rat **1**:365–368
rice water weevil **2**:551–552*T*
Richardiidae **3**:489–491*T*
Richelia **2**:611*F*, **2**:612*T*, **2**:621, **2**:622*F*
Richelia intracellularis **2**:612–613, **2**:612*T*
ricin **2**:426*T*, **2**:427–428
Ricinoides afzelli **1**:212
Ricinoididae **1**:212, **1**:212*F*
Ricinulei **1**:202*T*, **1**:205*F*, **1**:212, **1**:212*F*, **1**:246*T*
Ricinus communis **2**:427–428, **3**:703, **3**:704*F*, **7**:26
Ricker model **6**:177–178
Ridgeia piscesae **7**:328–329
Riftia pachyptila **4**:391, **7**:328–329
Riftia tubeworms **5**:120–121
Rift Valley **3**:444, **5**:470–471, **5**:470*T*
rights of nature concept **3**:268–269, **3**:269*F*
right whales **3**:192, **5**:65*T*, **5**:76–78, **5**:77*T*, **5**:81*T*
Rimicaris exoculata **7**:329
rinderpest **4**:156, **6**:332
ringed seal **5**:90
ring-tailed lemurs **6**:251*T*
ringworm **3**:634*T*
Rio Declaration **2**:79–80, **2**:307–308, **2**:484, **4**:326–327, **5**:713–714, **6**:78, **7**:43–44, **7**:136–137
Rio de Janeiro Botanic Gardens **6**:507
Riodinidae **6**:307*T*
Riodininae **1**:651, **1**:652*F*
riparian zones **6:461–468**
biodiversity **6**:337
characteristics **6**:461
definitions and terminology **7**:369*T*
flow regulation **6**:466–467, **6**:467*T*
functional role
biodiversity focal points **6**:461–463, **6**:463*F*
catchments and floodplains **6**:461–463, **6**:462*F*, **6**:463*F*
nutrient buffering effectiveness
high in-stream nutrient levels **6**:464–465
upland runoff **6**:463–464
reciprocal energy source function **6**:465–466
suspended sediment removal **6**:465
global extent **7**:370*T*
large animal influences **6**:466
Pacific salmon **6**:466
research summary **6**:467
threat management **6**:466–467, **6**:467*T*
water quality–nutrient management relationship **1**:103*T*, **1**:106–107
river dolphins **5**:65*T*, **5**:66–67
riverine forests **4**:432*T*
river otters **5**:473, **6**:479–480, **7**:390
rivers **6:469–484**
American Southeast rivers **3**:184, **3**:185*F*
assessments and management strategies **6**:482–484
biogeochemical cycles **6**:150–151, **6**:470–471
channelization **3**:573–574
climate change effects **6**:150–151
conceptual models
background information **6**:471–473
flood pulse concept **6**:474–475, **6**:475*F*
geomorphic–trophic hypothesis **6**:477
resource/nutrient spiraling **6**:473–474, **6**:473*F*
river continuum concept (RCC) **4**:373*F*, **6**:471–473, **6**:472*F*
riverine ecosystem synthesis (RES) **6**:477
serial discontinuity concept **6**:474, **6**:474*F*
telescoping ecosystem model (TEM) **6**:475–477, **6**:476*F*
conservation status **3**:171–173*T*
dam projects **1**:510–512, **3**:573–574
ecosystem services **1**:334*T*
endangered freshwater invertebrates **3:176–187**

American Southeast rivers **3**:184, **3**:185*F*
anthropogenic impacts
direct harvests **3**:181
global climate change **3**:181–182
habitat destruction/degradation **3**:180–181, **3**:180*F*
introduced species **3**:181
pollution **3**:181, **3**:181*F*
causal factors **3**:177, **3**:177*F*
diversity hot spots **3**:176
global conservation status **3**:176–177, **3**:176*T*, **3**:177*F*
global distribution patterns **3**:182–184, **3**:182*T*
global diversity **3**:176
number of species **3**:182–184, **3**:182*T*
protection strategies **3**:184–187
small ranges
causal factors **3**:178
conservation status **3**:177*F*
dispersal ability **3**:178*F*
high endemic hot spots **3**:178–179
high endemic species **3**:179, **3**:180*F*
Lake Baikal **3**:179, **3**:183*F*
narrowly endemic species **3**:177–178, **3**:177*F*, **3**:179
vulnerability factors
small ranges **3**:177–178, **3**:177*F*
sparse populations **3**:179–180
fish species **3**:444
freshwater invertebrates **4:369–378**
background information **4**:369–370
ecological functions and processes
characteristics **4**:375
decomposition and recycling **4**:375
material transfer and uptake **4**:375–376
primary production **4**:376
global patterns **4**:373
influencing factors
disturbances **4**:375
endemism **4**:374–375
habitat heterogeneity **4**:375
isolation factors **4**:374–375
origins **4**:374
research background **4**:373–374
species interactions **4**:375
life cycles **4**:370, **4**:372*F*
major groups **4**:370, **4**:371*T*
major habitats **4**:369–370, **4**:369*F*
major threats
altered water regimes **4**:377–378
biodiversity loss **4**:376–377
exotic species invasions **4**:378
habitat depletion and loss **4**:376–377
pollution **4**:377
reduced water quality **4**:377
number of species **4**:370–371, **4**:371*T*
river continuum concept (RCC) **4**:373*F*
taxonomic groups **4**:371–373
zonal representatives **4**:374*F*
geographic distribution **4**:370*T*
habitat diversity **3**:570–571
importance **6**:469
Louisiana delta restoration (case study) **7**:378
mean composition **3**:562*T*
nature-based coastal protection strategies **5**:452–453, **5**:453*T*
Near East ecosystems **5**:473
organic pollution **3**:575
physical characteristics **6**:469–470, **6**:470*F*
predation effects **6**:235–236
protected area systems **6**:494–498, **6**:496*F*, **6**:497*F*
river delta restoration **7**:378
river–floodplain connectivity **7:144–160**
benefits **7**:144, **7**:145–146
changes and conversion impacts **7**:150–151, **7**:151*F*
ecological processes **7**:145
ecosystem services
benefits **7**:145–146, **7**:150
fisheries **7**:146–147
flood attenuation **7**:147–150, **7**:148*F*, **7**:149*F*
floodplain agriculture **7**:146
flow levels and events **7**:150
hydrological connectivity **7**:150
overbank flooding **7**:145–146
spatial variables **7**:150
floodplain reconnection and conservation
ecosystem restoration **7**:153
flood-risk reduction programs **7**:154–155, **7**:156*F*, **7**:157*F*
general discussion **7**:152–153
livelihood restoration **7**:153–154
multipurpose reservoir flexibility **7**:155–156, **7**:156*F*, **7**:157*F*
restoration strategies **7**:152–153
future research outlook **7**:156–158
geomorphic processes **7**:145
hydrological processes **7**:145
key floodplain processes **7**:144–145
land and water management challenges
flood control issues **7**:152
flood risks **7**:152
freshwater biological diversity and productivity declines **7**:151
general discussion **7**:151
river-dependent communities impacts **7**:151–152
research summary **7**:156–158
riverine biota
amphibians **6**:479–480
birds **6**:479–480
fish **6**:479, **6**:480*F*
invertebrates **6**:478–479, **6**:478*T*
large animals **6**:479–480
microbiota **6**:477
phytoplankton **6**:472*F*, **6**:477, **6**:478*T*
plants **6**:477–478
reptiles **6**:479–480
Southeast Asia **1**:510–512, **1**:511*F*, **1**:512*F*
species richness **3**:572*T*
threats and challenges
dams **6**:480–481, **6**:481*T*
human activity **6**:480–481
interbasin connections **6**:481–482
introduced species **6**:481*T*, **6**:482
nutrient loading **6**:481*T*, **6**:482
overharvesting **6**:481*T*, **6**:482, **6**:483*F*
stream channelization **6**:481–482, **6**:481*T*
toxic substance loading **6**:481*T*, **6**:482
water storage **3**:561*T*
RIVPACS (UK) **4**:254
roach **4**:463
roadkill **3**:192, **3**:217
Road to Survival, The (Vogt) **2**:285
robber flies **3**:489–491*T*, **3**:492
Robe River (Australia) **7**:59*T*
robins **4**:174, **5**:723
robust grasshopper **4**:194*T*
robust tufted capuchin **6**:256–258*T*
Roccella tinctoria **1**:591
rock **6**:607*T*
Rock Creek (Washington, D. C., USA) **7**:59
rocket **3**:131–132
rockfalls **3**:436
rockfish **3**:467–471
rock hyrax **4**:695
rock lobsters **5**:57–58, **5**:60*F*, **6**:233
rock ptarmigans **7**:27–28
Rocky Mountains **5**:563
Rocky Mountain tailed frogs **4**:517
rocky shores
anthropogenic impacts **5**:56–57, **5**:58*F*, **5**:59*F*
ecosystem diversity **5**:51–52
predation effects **6**:232
rocky subtidal communities **5**:57–58, **5**:60*F*
Rodale, JI **1**:112
Rodentia/rodents
boreal forests **1**:631–632
Cape Region (South Africa) **5**:220
characteristics **4**:697
Chile **5**:217
convergent evolution **4**:701*T*
desert ecosystems **2**:470*T*, **2**:472–474
as disease vectors **1**:365–368, **1**:366*F*
El Niño Southern Oscillation (ENSO) events **3**:158–159
evolutionary history **2**:583*F*
food webs **7**:266*F*, **7**:267*T*
genetic diversity **3**:674*F*
habitat–niche relationship **4**:43–44, **4**:43*F*
hantaviruses **1**:365–368, **1**:366*F*
interspecific competition **2**:186–187, **2**:187*T*
introduced species **1**:76–77
late Quaternary extinction event (LQEE) **4**:727*T*
mangrove ecosystems **5**:14
migratory patterns **5**:313–314
Near East ecosystems **5**:472
nematode parasites **7**:432, **7**:432*T*
number of species **4**:690*T*
representative species **2**:587*F*
species diversity **2**:586, **4**:690*T*, **4**:697
threat percentage **3**:195
tropical rain forests **6**:309*T*
tundra ecosystems **6**:330–331
Rodolia cardinalis **4**:361
roe deer **4**:515–516, **5**:472
Rollinia mucosa **4**:284*F*, **4**:285, **4**:287–290, **4**:289*T*

Roloway monkey **6**:256–258*T*
Romaleidae (lubber grasshoppers) **3**:734*F*, **3**:735
Romanes, George **2**:424
Romania
  botanical gardens **6**:519
  vegetation surveys **3**:341–344*T*
Romantic Movement **2**:280–281
Rondo dwarf galago **6**:256–258*T*
Roniviridae **5**:134*T*, **5**:137
Roosevelt, Theodore **2**:282–283
rooting **4**:359
root parasites **3**:633*T*, **4**:70, **4**:76–77
roots **3**:90*T*, **3**:130*T*, **3**:131*T*
root–soil interface **6**:111–113, **6**:112*F*, **6**:113*T*
ropaeolum polyphyllum **4**:105*F*
Ropalomeridae **3**:489–491*T*
Roquefort cheese **3**:638–639
Roridula **6**:44*T*
Roridulaceae **5**:219–220, **6**:44*T*
Roridula gorgonias **6**:45
rorquals **5**:65*T*
Rosaceae **3**:129*T*, **3**:130*T*, **5**:621–625, **5**:623*F*
Rosalina **6**:291–292
Rosa palustris **6**:115*T*
Rosen, Walter G. **4**:122
roses **6**:120*T*
rosewood **6**:560
Rosid 1 Clade diversification **5**:621–625, **5**:623*F*, **5**:624*F*
Rosidae **4**:72*F*
Rosid II Clade **5**:622*F*
Rosmarinetea officinalis **3**:340
Ross seal **5**:90
Rostrhamus sociabilis **7**:186
Rostroconchia
  physical characteristics **5**:375, **5**:375*F*
  spatiotemporal distributions **5**:373–374, **5**:374*F*
rosy periwinkle **1**:359, **6**:57
rosy wolf snail **4**:359, **4**:361
rot **3**:633*T*
Rotatoria **5**:669–670*T*, **6**:26
rotenone **1**:589
Rothschild, Walter **7**:46
Rotifera
  acidification effects **3**:584–585, **3**:584*F*
  Antarctica **1**:184*T*
  eutrophication impacts **3**:362
  fossil record **3**:684*F*
  global diversity **4**:385*T*
  lake and pond ecosystems **4**:371–373, **4**:374*F*
  marine ecosystems **5**:50*T*
  number of species **4**:371*T*
roughies **3**:467–471
Roundtable for Sustainable Palm Oil (RSPO) **5**:608–609
roundtail chub **3**:260
roundworms **4**:371*T*, **5**:660*T*, **5**:671*T*
Rousettus aegyptiacus **5**:474–475
rove beetles **1**:328, **5**:674
Royal Botanic Garden Edinburgh **6**:510–511
Royal Botanic Garden (Madrid) **6**:506–508, **6**:510–511
Royal Botanic Gardens Melbourne **6**:506–508
Royal Botanic Gardens (RBG) Kew
  collection size **5**:411*F*
  conservation biology **2**:245
  historical perspective **5**:405, **6**:78
  importance **6**:506
  nonseed gene banks **3**:652
  plant conservation **4**:328, **6**:86
  research collections and libraries **6**:510–511
  seed banks **4**:141, **4**:328, **6**:86
  Vavilov centers **4**:141
Royal Botanic Gardens Sydney **6**:506–508
Royal Medicinal Plant Garden **5**:405
Royal National Park (Australia) **2**:289–292, **6**:486
Royal Ontario Museum **5**:406
royal terns **3**:256
r-strategists **3**:124–125
rubber **5**:605, **6**:320
rubber tree **2**:383
Rubiaceae **1**:267, **2**:427, **3**:129*T*, **3**:522*T*, **5**:11*T*, **5**:489, **6**:313–318, **6**:316*T*
Rubia danaensis **5**:485
Rubia tenuifolia **5**:481–482
Rubisco **2**:361–362, **2**:362*F*, **2**:366–368, **2**:366*F*, **2**:654–655, **2**:675
Rubus **4**:170
Rubus alceifolius **4**:364
Rubus proceus **1**:80
ruddy duck **7**:184
ruffed grouse **4**:62
ruffed lemurs **6**:251*T*, **6**:256–258*T*
Rufiji River **7**:146
rufous bushchat **5**:474
Rugitermes **4**:409*T*
Ruibus **3**:130*T*
ruminants
  adaptive radiation **1**:22, **1**:23*F*
  convergent evolution **4**:701*T*
Rumphius **3**:322
Rungwecebus kipunji **6**:256–258*T*
running buffalo clover **3**:214
Rupicapra rupicapra **5**:313
Ruppia maritima **3**:360–361
Ruschia **5**:219
Russell cycle **5**:690–691
Russia
  boreal forests
    beetle species **7**:219*T*, **7**:220*F*, **7**:220*T*
    forest characteristics **7**:216–217, **7**:217*F*
    geographic distribution **1**:626, **1**:630*T*
    landscape patterns **7**:218*F*, **7**:219*F*
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:507*T*, **6**:510*T*
  C_4 plants **2**:371*T*
  ecosystem conservation **3**:171–173*T*
  energy use **3**:252*T*
  fishing/fisheries **6**:422*T*
  indigenous peoples **4**:269*T*, **4**:270*T*
  natural history museums and institutions **5**:411*F*
  regional carbon stocks **1**:678*F*
  salt marshes **1**:277
  temperate forests **7**:116*T*
  temperate grasslands **1**:269
  wheat production **1**:86, **1**:87–88
Russian olive **7**:399
Russian sea otter **5**:87*T*
Russian wheat aphid **4**:361
rust **2**:139
rust flies **3**:489–491*T*
rust fungus **3**:629–630, **3**:633*T*, **4**:349, **4**:362
Rutaceae **3**:129*T*, **3**:522*T*
Rutilus rutilus **4**:463
Ruvettus pretiosus **3**:470
Ruyschia **3**:522*T*
Rwanda
  botanical gardens **6**:519
  Convention on Biological Diversity (1930) **2**:313
  energy efficiency measures **3**:254*F*
rye **2**:662–663
ryegrass **2**:389–390
ryegrass infections **3**:349
Rynchospora **2**:376*F*
Ryukyu Islands **1**:262–263*T*

S

Sabah **1**:277*T*
Sabatia angularis **4**:248*F*
Sabatinca **5**:388*F*
Sabellariidae **7**:417*F*
Sabellida **7**:417*F*
Sabellidae **7**:417*F*
saber-toothed blennies **2**:141–142
saber-toothed cats **4**:724–725*T*
saber-toothed species **1**:693
Sabiaceae **5**:621*F*
Sabinia **1**:270
sablefish **3**:467–471
sables **1**:631–632
Saccharomyces **2**:182–184
Saccharomyces cerevisiae
  bread baking **3**:638
  characteristics **3**:628–629
  ethanol production **3**:638
  evolutionary development **2**:644
  molecular level diversity **2**:636*F*
  mutation rates **2**:628*T*
  wine and beer making **3**:638
Saccharomyces paradoxus **2**:636*F*
Saccharum **1**:69–70, **1**:543–547
Saccharum officinarum **2**:663, **6**:553
Saccocoelioides octavus **7**:450–451*F*
Sacculina carcini **5**:658*F*
sac fungi **2**:644
Sacoglottis **7**:272
Sacramento River **7**:147, **7**:155, **7**:156*F*, **7**:157*F*
sacred sites **6**:384, **6**:384–385, **6**:563–564, **6**:567, **6**:568, **7**:229
Saculamonas ecuadoriensis **5**:576
saddle fungi **3**:628–629
Safeguard for Agricultural Varieties **1**:645
Sagan, Carl **5**:630–631
Sagara Anakan **1**:276*T*
sagebrush **2**:468–469

sagebrush–grassland restoration **6**:439–442, **6**:440*F*
Sagenopteris **5**:616*F*
Sageretia thea **5**:485
Sagita plantonis **4**:386*T*
Sagitta pseudoserratodentata **4**:386*T*
Sagittarius serpentarius **1**:608
saguaro cactus **2**:468–469, **2**:475*F*, **7**:223
sagui **6**:560
Saguinus bicolor **6**:256–258*T*
Saguinus fuscicollis **7**:284
Saguinus leucopus **6**:256–258*T*
Saguinus mystax **7**:284
Saguinus oedipus **6**:256–258*T*
Sahafary sportive lemur **6**:250, **6**:256–258*T*
Sahara Desert **2**:469–470, **2**:480, **2**:482*F*
Sahel region **3**:143*F*, **7**:146, **7**:153–154
Saiga **4**:724–725*T*
saiga antelope
  case study
    background and characteristics **4**:711, **4**:711*F*
    conservation efforts **4**:712–713
    general discussion **4**:711
    major threats **4**:711–712, **4**:712*F*
    population decline **4**:712*F*
    range area **4**:712*F*
  ecological role **1**:705
  illegal hunting **3**:191–192
  migration **5**:316–317
  temperate/cold-desert shrublands **1**:269, **6**:334
Saiga tatarica **1**:269, **1**:705, **3**:191–192, **4**:711, **6**:334
sailfish **3**:470
Saimiri oerstedii citrinellus **6**:256–258*T*
Saimiri oerstedii oerstedii **6**:256–258*T*
Saimiri sciureus **2**:587*F*
Saint Helena **6**:519
Saint-Hilaire, Geoffrey *see* Geoffrey Saint-Hilaire, étienne
Saint Kitts and Nevis **6**:519
Saint Lucia **6**:519
Saint Pierre and Miquelon **6**:507*T*
Saint Vincent and the Grenadines **6**:519
sakis **6**:251*T*, **6**:256–258*T*
sal **1**:266, **6**:225, **7**:118, **7**:273–274
salamander fish **3**:464–465
salamanders
  biodiversity
    geographic distribution **3**:691–693
    historical biogeography **1**:169–170
    urbanization threats **4**:564
  biodiversity generation
    extinction rate–speciation rate relationship **1**:431*F*
    maximum likelihood estimation **1**:431*F*
    phylogenetic reconstruction **1**:431–432, **1**:431*F*
  characteristics **3**:691–693, **7**:339
  critical pH levels **6**:153*F*
  dichopatric speciation **6**:622–624, **6**:623*F*
  functional anatomy **1**:170–171
  introduced parasites–species decline relationships **2**:528*T*
  larval morphology and ecology **1**:173–174
  life histories **1**:171–172
  migratory patterns **5**:315–316
  morphology **1**:170–171
  postlarval behavior and ecology **1**:175–176
  reproductive biology **1**:171–172, **1**:173*F*
  special features **1**:171
  species richness **1**:170*T*
  subterranean ecosystems **7**:58
  surface water acidification **1**:10*T*
  wetland ecosystems **7**:389–390
salar vegetation **4**:105
Saldidae **7**:293
Salicornia pulvinata **4**:105
saline lakes **3**:560, **3**:561*T*
Salix **1**:359, **1**:613–614, **1**:628–630, **1**:630*T*, **3**:258
Salix arctica **1**:236*F*
Salix eriocephala **4**:173–174
Salix martiana var. *humboltiana* **1**:163
Salix sericea **4**:173–174
Salix viminalis **6**:115*T*
Salmo **2**:528*T*
Salmo gairdneri **4**:255, **4**:465
salmon **6:522–531**
  anthropogenic impacts **3**:579, **5**:318–319
  aquaculture **1**:190*T*, **1**:196, **3**:448–449
  biodiversity value **6**:529–530
  Census of Marine Life project **2**:16, **2**:17*F*
  characteristics **3**:464–465, **3**:465*F*
  classifications **6**:522–523
  climate change effects **6**:529
  evolutionary significant unit (ESU) criteria **1**:668–669
  fossil record **6**:522–523
  historical status **6**:526–527
  introduced parasites–species decline relationships **2**:528*T*, **2**:532
  introduced species **4**:363–364
  life histories
    Atlantic salmon **6**:524
    Chinook salmon **6**:524, **6**:525*F*
    chum salmon **6**:523
    coho salmon **6**:523–524
    general characteristics **6**:523
    pink salmon **6**:523
    sea-run cutthroat trout **6**:526
    sockeye salmon **6**:523
    steelhead trout **6**:524–526
  mercury contamination **1**:12*F*
  migratory patterns **5**:314–315
  natural environmental variability **6**:529
  Pacific salmon
    characteristics **6**:522
    evolutionary significant unit (ESU) criteria
      evolutionary lineage versus ecological context approach **1**:670–672, **1**:670*F*
      hatchery fish **1**:670–672, **1**:670*F*
      research background **1**:668–669
      research criteria **1**:669–670
    life cycle **1**:669–670, **1**:669*F*
    riparian zones **6**:466
  population decline **6**:526–527
  special adaptations
    anadromy **6**:526
    homing behavior **6**:526
  species **6**:522
  surface water acidification **1**:10*T*
  threats and challenges
    anthropogenic effects **6**:527
    habitat destruction **6**:527
    harvesting moratorium **6**:527–528, **6**:528*F*
    hatchery fish impacts **6**:528
    hydropower dam impacts **6**:528–529
    overfishing **6**:527–528
  urbanization threats **4**:564
Salmonella enterica **2**:636*F*
Salmonidae **6**:522–523
Salmoniformes **3**:464–465, **6**:522–523
Salmoninae **6**:522–523
Salmo salar **3**:260, **4**:516, **5**:314–315, **5**:318–319, **6**:522
Salmo trutta **3**:181
Salpinctes obsoletus exsul **1**:620
Salpingoeca **6**:293
salps **6**:25
Salsola **2**:371*T*
Salsola kali **2**:376*F*
Salsola kamarovii **2**:364*F*
Salsola tetrandra **5**:485–486
Salsola vermiculata **5**:484–485
Salsola villosa **1**:268
saltbush **2**:469
salt cedars **3**:284, **4**:360, **4**:378, **7**:399
salt marshes
  Asia **1**:277
  Australian plant communities **1**:292–297
  biodiversity restoration **6**:456–457
  C_4 plants **2**:371*T*
  definitions and terminology **7**:369*T*
  Delaware Bay salt marsh restoration (case study) **7**:376–378, **7**:377*F*
  ecosystem services **6**:77
  marine sedimentary habitats **5**:107, **5**:108*T*
  nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*, **5**:455*F*
  plant species richness **7**:388–389, **7**:388*F*
  restoration programs **7**:375–378, **7**:376–378, **7**:377*F*
  South America **6**:607*T*
salt marsh grasses **6**:627, **6**:627*F*, **7**:375–378
salt water plants **6**:77
salvage logging **1**:447–448
Salvertia convalarriodora **1**:165
Salvertia convallariodora **1**:167
Salvia azurea **1**:707
Salvia nutans **1**:269
Salvia superba **1**:269
Salviniaceae **4**:346–347
Salvinia molesta **4**:346–347, **4**:347*F*, **6**:669
Salvinia radula **1**:165

Sambirano mouse lemurs **6**:256–258*T*
Sambucus mexicana **1**:120
Samburu National Park (Kenya) **1**:60*F*
Samoa **1**:596–597, **2**:313, **4**:560*F*, **5**:592*F*, **6**:507*T*
Samoan moorhen **7**:184
San Benedicto rock wren **1**:620
San Carlos de Rio Negro **1**:159*T*, **1**:161*F*
sandalwood tree **1**:267, **4**:71
Sand County Almanac, A (Leopold) **2**:285
Sand County Almanac, The (Leopold) **4**:652–653
sand dunes **2**:464*F*
sanderlings **5**:312
sandflats **5**:107, **5**:108*T*
sand flies
  classifications **3**:489–491*T*
  mangrove ecosystems **5**:13
sandgropers **3**:736
sandgrouse **1**:604, **2**:549–550
sandhill cranes **7**:189
Sand Hills **5**:562–563
sand lances **3**:467–471
sandpipers **7**:32
sand tiger shark **3**:459*F*
sandwich hybridization assay (SHA) **2**:693–694, **2**:694*F*
sandy beaches
  alongshore distribution **2**:131–133
  background information **2**:128
  beach index (BI) **2**:129
  future research areas **2**:135
  intertidal ecosystems **4**:337–339
  macrofauna surveys **2**:129–131, **2**:130*F*
  macroscale patterns
    latitudinal trends **2**:131, **2**:131*F*
    morphodynamic effects **2**:131, **2**:132*F*, **2**:133*F*
  mesoscale patterns **2**:131–133
  microscale patterns **2**:133
  microtidal regimes **2**:128–129, **2**:129*F*
  morphodynamic states **2**:128–129, **2**:129*F*
  nature-based coastal protection strategies **5**:452–453, **5**:453*T*
  research summary **2**:135
  species richness hypotheses **2**:133–134
  threats and challenges **2**:134–135
  tidal regimes **2**:129
  two-dimensional model **2**:129, **2**:130*F*
  zonation patterns **2**:131–133
Sanford's lemur **6**:256–258*T*
San Francisco garter snake **6**:411
Sangihe tarsier **6**:256–258*T*
Sanguilevator yearsleyi **7**:456–457*F*
Sanguinarea canadensis **6**:50
Sanje River mangabey **6**:256–258*T*
San Marino
  botanical gardens **6**:507*T*, **6**:519
  vegetation surveys **3**:341–344*T*
Sanmiguelia **5**:617–619
San Quintin kangaroo rat **7**:229*T*
Santa Barbara live-forever **3**:269–270, **3**:270*F*
Santa Catalina Island, California **4**:663
Santalaceae **4**:71, **4**:72*F*, **6**:315*T*
Santalales **5**:622*F*
Santalum album **1**:267
Santa Marta white-fronted capuchin **6**:256–258*T*
Santa Rosa National Park, Costa Rica **5**:490, **5**:491*T*
Sao Francisco Basin **3**:105–107, **3**:106*F*
saola **4**:704
Sao Tome and Principe **6**:507*T*
Sapajus apella margaritae **6**:256–258*T*
Sapajus flavius **6**:256–258*T*
Sapajus robustus **6**:256–258*T*
Sapajus xanthosternos **6**:256–258*T*
sap feeders **6**:39–41
Sapindaceae **5**:489, **6**:314, **6**:316*T*
Sapium marmieri **1**:162
sapodilla **2**:663–664
saponins **2**:426*T*, **2**:427
Sapotaceae **3**:129*T*, **6**:316*T*
Saprobic Index **4**:254
Saprolegnia **3**:696
Saprolegnia oogonium **3**:630*F*
saprolegniasis **2**:528*T*
Saprolegnia sporangium **3**:630*F*
saprophytes **3**:632–633
saps **3**:130*T*, **3**:131*T*
SAR11 group **5**:292
SAR86 group
  abundance and diversity **6**:282–283
  gene sequence studies **5**:292, **5**:296
  phototrophic mechanisms **6**:280–282
  spectral tuning **6**:283
Sarawak
  Sarawak Mangrove Forest Reserve **1**:276*T*
  tropical moist forest **1**:277*T*
Sarcodina **5**:669–670*T*
Sarcomastigophora **5**:669–670*T*
Sarcophagidae **3**:489–491*T*, **3**:493
Sarcophilus laniarius **4**:688, **5**:364*F*
Sarcopoterium spinosum **5**:482
Sarcopterygii **2**:7*F*, **7**:338
Sarcoptes scabiei **1**:217, **5**:660*T*
Sarcosuchus imperator **6**:401
Sardina pilchardus **2**:356
sardines **2**:356, **3**:450–451, **5**:688–690, **5**:690*F*
Sardinops sagax **5**:688–690
Sargasso Sea **5**:293–294
Sargasso Sea Gyre **5**:45
Sargassum **2**:642
Sargent, Charles S. **2**:281
Sargentodoxaceae **1**:260*T*
Sarotherodon melanotheron **3**:481
Sarracenia **6**:44*T*
Sarracenia alata **6**:45
Sarraceniaceae **6**:44*T*
Sarracenia rubra ssp. *alabamensis* **3**:210–211, **3**:212*F*
Sasa **1**:271
Saskatchewan Wetland Conservation Corporation **7**:20–21
Satan eurystomus **7**:58
satellite imagery
  characteristics **6**:392–393
  coarse-resolution sensors **6**:394
  emerging technologies **6**:394
  land cover estimation **4**:567
  Landsat program **6**:393, **6**:393*T*
  SPOT program **6**:393–394
saturated fatty acids **7**:25–26, **7**:32–33
saturation tests
  caveats **2**:603
  experimental evidence **2**:603–604, **2**:604*F*
  methodologies **2**:602–603
  test results **2**:603
  unsaturated versus saturated patterns **2**:603
Saturniidae **5**:387–388*T*, **5**:391*F*, **5**:396, **5**:399*F*
Satyrinae **1**:653
sauco **1**:120
Saudi Arabia
  animal species **1**:261*T*
  botanical gardens **6**:519
  ecosystems
    flora **5**:481*T*
    mangroves **1**:275*T*
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  see also Near East ecosystems
Sauria **7**:340–341
Saurolophus **2**:519*T*
Saurornitholestes **2**:519*T*, **2**:520*T*
Saururus cernuus **6**:115*T*
Saussurea **1**:270
Savannah River Site corridor experiments **4**:471–473, **4**:472*F*
savannas
  adaptive radiation **1**:22, **1**:23*F*
  Africa **1:58–74**
    biodiversity **1**:70, **1**:71*F*
    characteristics **1**:58–59
    classifications
      arid savannas **1**:65–67, **1**:66*F*
      dystrophic savannas **1**:65–67, **1**:66*F*
      eutrophic versus mesic savannas **1**:65–67, **1**:66*F*
      forest-savanna mosaics **1**:66*F*, **1**:68
      general discussion **1**:65–67
      grass and shrub savannas **1**:66*F*, **1**:67
      life-form associations **1**:66*F*, **1**:67
      shrublands **1**:66*F*, **1**:67
      tree and shrub savannas **1**:66*F*, **1**:67
      woodland savannas **1**:66*F*, **1**:67–68
    climate change effects **1**:71–72
    emotional responses **1**:41
    future outlook **1**:71–72
    geographic distribution **1**:59–60, **1**:59*F*
    global carbon cycle **1**:70–71
    precipitation–temperature relationship **1**:59*F*
    species richness **1**:71*F*
    structure–function drivers
      bottom-up controls **1**:60–61
      characteristics **1**:60–61
      fire/fire regimes **1**:61, **1**:63

herbivory **1**:61, **1**:65
nutrient intake regulation mechanisms **1**:63–65
regulating factors **1**:63
top-down controls **1**:61
water and nutrient intake **1**:60–61
woody vegetation dynamics **1**:61–63, **1**:62*F*, **1**:64*F*
termites **4**:430*T*, **4**:432*T*
tree cover variations **1**:60*F*, **1**:61–63, **1**:62*F*, **1**:64*F*
tropical biomes **1**:49–51, **1**:58–59
Amazonian savannas
cerradão **1**:164–165
flooded savannas **1**:165
nonflooded savannas **1**:164–165
Asia
biodiversity **1**:70
biome distributions **1**:64*F*
characteristics **6**:333
classifications
dipterocarp savanna forests **1**:69*F*, **1**:70
dry deciduous forests **1**:69–70, **1**:69*F*
dry thorn forests **1**:69, **1**:69*F*
general characteristics **1**:68–69
moist sal savannas **1**:70
climate change effects **1**:71–72
future outlook **1**:71–72
geographic distribution **1**:59–60, **1**:59*F*
global carbon cycle **1**:70–71
tree cover variations **1**:60*F*
biodiversity change scenarios **4**:6*F*
C_4 plants **2**:371*T*
characteristics **6**:331–332, **6**:332*F*, **6**:333*F*
conservation status **3**:171–173*T*
current global land cover **4**:557*F*
ecosystem function measurements **3**:77*F*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
neotropical seasonally dry forests **5**:496
net primary production (NPP) **1**:560*T*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:495*T*, **6**:499*T*
seed dispersal and germination **1**:708
soils **6**:123*T*
South America **3**:111*F*, **3**:113, **3**:114, **6**:600–604*T*, **6**:607*T*
tourism industry **7**:223
tropical Asia **1**:262–263*T*, **1**:266–267
tropical forest ecosystems **7**:274
wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*, **4**:581*F*
sawback turtles **7**:390
sawflies **4**:177–178, **4**:392*T*, **5**:674, **6**:142
sawgrass **7**:372, **7**:385, **7**:399
Sawit Watch **5**:608–609
Saxifraga **1**:270
Saxifraga caespitosa **1**:236*F*
Saxifragaceae **3**:522*T*, **5**:622*F*
Saxifraga magellanica **4**:106–107
Saxifraga nivalis **1**:236*F*
Saxifraga oppositifolia **1**:232
Saxifragella **4**:101
Saxifragodes albowiana **4**:106–107
Sayornis phoebe **7**:52–53
Scabiosa alpestris **1**:269
Scaevola **1**:277
Scalibregmatidae **7**:417*F*
scaling laws **6:532–539**
basic concepts
complete and incomplete similarity **6**:535*F*, **6**:536*F*, **6**:537–538
multiscale analysis **6**:534–536, **6**:535*F*, **6**:536*F*
scaling relations
general discussion **6**:538
number scales completely with area ($N \cong A^1$) **6**:538
species diversity distribution ($S(n_i \leq n_k) \cong S(A_i \leq A_k) \cong H(A_i \leq A_k)$) **6**:538
species number scales incompletely with area ($S \cong A^{\beta}$) **6**:538
species number scales incompletely with numbers ($S \cong N^{1/4}$) **6**:538
species number scales with total numbers ($S \cong N^{1/4} \cong A^{1/4}$) **6**:538
steeper scaling exponent ($S \cong A^{\beta}$, where $\beta_{island} > \beta_{block}$) **6**:538
steeper scaling exponent ($S \cong A^{\beta}$, where $\beta_{province} > \beta_{block}$) **6**:538
scope measures **6**:536–537
biodiversity
alpha diversity **1**:402–403, **6**:534, **6**:534
beta diversity **1**:402–403, **6**:534, **6**:534
evolutionary biology **6**:535
gamma diversity **1**:402–403, **6**:534, **6**:534
genetic variability **6**:533–534
marine ecology **6**:535
scalable quantities **6**:532
scaling relations
general discussion **6**:538
number scales completely with area ($N \cong A^1$) **6**:538
species diversity distribution ($S(n_i \leq n_k) \cong S(A_i \leq A_k) \cong H(A_i \leq A_k)$) **6**:538
species number scales incompletely with area ($S \cong A^{\beta}$) **6**:538
species number scales incompletely with numbers ($S \cong N^{1/4}$) **6**:538
species number scales with total numbers ($S \cong N^{1/4} \cong A^{1/4}$) **6**:538
steeper scaling exponent ($S \cong A^{\beta}$, where $\beta_{island} > \beta_{block}$) **6**:538
steeper scaling exponent ($S \cong A^{\beta}$, where $\beta_{province} > \beta_{block}$) **6**:538
taxonomic diversity distributions **6**:532–533, **6**:533*F*
terrestrial ecology **6**:535
ecological perspectives **3**:6
island biogeography
spatial scales **4**:395
timescales **4**:395
scale effects
conservation planning **6**:539
diversity estimates **6**:538–539
scalloped hammerhead sharks **5**:449–450
scaly crickets **3**:730
Scandentia **4**:690*T*, **4**:691
Scaphopoda
classifications **4**:381*T*
marine environments **4**:382*F*
number of species **5**:669–670*T*
physical characteristics **5**:378–379, **5**:378*F*
spatiotemporal distributions **5**:373–374, **5**:374*F*
Scaptomyza **1**:28–29
Scarabaeidae **1**:327*T*, **1**:328, **1**:710
Scarabaeoidea **1**:327*T*
scarab beetles **1**:120, **1**:327*T*, **1**:328
scarlet ibis **5**:13–14, **7**:390*F*
scarlet runner beans **2**:663–664
Scathophagidae **1**:710, **3**:489–491*T*
Scatopsidae **3**:489–491*T*
scats **3**:467–471
scattered trees **1**:420–421, **1**:423*F*
Scelidodon **4**:727*T*
Scelidoiherium **4**:727*T*
Scenedesmus **6**:30
Scenedesmus quadricauda **6**:33*F*
Scenopinidae **3**:489–491*T*
Scenopoeetes dentirostris **2**:75*F*
Schausiella santarosensis **4**:595
Schedorhinotermes **4**:409*T*
Schefflera **3**:522*T*, **6**:316*T*
Scheldt Estuary **5**:28*T*
Schendylops oligopus **5**:421–423
Schiedea globosa **2**:636*F*
Schima **1**:270–271
Schinus terebinthifolium **4**:357, **6**:442, **7**:399
Schisandra **5**:621*F*
Schistocerca gregaria **2**:549–550, **6**:357–358
Schistosoma **7**:450–451*F*
Schistosoma mansoni **1**:368, **1**:368*F*
Schistosomatoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454
schistosomiasis
anti-schistosomiasis soap **1**:592
disease regulation **3**:349
as disease vectors **1**:368, **1**:368*F*
mollusc–human interactions **5**:381–382
Schizachyrium condensatum **4**:364
Schizachyrium scoparium **1**:707
Schizodactyloidea **3**:730, **3**:730*F*
schizoendemics **3**:228
Schizolobium amazonicum **1**:161–162
Schizomida **1**:202*T*, **1**:205*F*, **1**:212, **1**:212*F*, **1**:246*T*
Schizopera taricheana **5**:470*T*
Schizophora **3**:489–491*T*, **3**:491–492
Schizosaccharomyces **2**:182–184
Schlegelia **3**:522*T*
Schoenefeldia gracilis **1**:67
Schoenoplectus tabernaemontani **7**:379
Schouteden's blue monkey **6**:256–258*T*

Schradera **3**:522*T*
Schreckensteinioidea **5**:387–388*T*
Schrödinger, Erwin **5**:630
Schumacher, E. F. **6**:384
Schurz, Carl **2**:281
Schwalbea americana **3**:206
Sciadoceridae **3**:489–491*T*
Sciadophyllum **3**:522*T*
Sciadopityaceae **1**:260*T*
Sciadopitys **1**:272–273
Sciaridae **3**:489–491*T*, **3**:493–494
Scientific Committee on Ocean Research **4**:326–327
Scientific Committee on Problems of the Environment **4**:326–327, **7**:18
scientific whaling **5**:88–89
scientific workflow systems **1**:445
Scierospongia **3**:685*F*
Scincus scincus **5**:467–469, **5**:468*F*, **5**:472
Sciomyzidae **3**:489–491*T*
Sciomyzoidea **3**:489–491*T*
Scirpus **1**:270, **1**:278–279
Scirpus longii **3**:209
Sciuridae **2**:583*F*, **6**:309*T*
Sciurocheirus alleni alleni **6**:256–258*T*
Sciuromorpha **2**:586, **4**:697
Sciurus carolinensis **1**:79, **2**:190, **2**:551–552*T*, **3**:191, **4**:360, **7**:28*F*
Sciurus griseus **4**:564
Sciurus ignitus **2**:587*F*
Sciurus vulgaris **2**:190, **2**:529*F*, **3**:191, **4**:360
Sclater, Philip **1**:578–579
Scleria geniculata **1**:165
Sclerocalyptus **4**:727*T*
Scleropages formosus **3**:463, **3**:463*F*
Scleropauropus tarumamirimi **5**:431
sclerophyll oaks **1**:268
sclerophyll shrubs **1**:277
Sclerotinia sclerotorium **2**:644
Scolopendra abnormis **3**:226
Scolopendra cingulata **5**:422*F*, **5**:425
Scolopendra gigantea **5**:421–423, **5**:425
Scolopendra morsitans **3**:222
Scolopendromorpha **1**:246–247, **1**:246*T*, **5**:421–423, **5**:422*F*
Scolopendropsis bahiensis **5**:421–423
Scolopendropsis duplicata **5**:421–423
Scolytinae **1**:328
Scomber australasicus **5**:661*F*
Scombroidei **3**:470
Scombrolabracid **3**:467–471
Scopelomorpha **3**:466
scopolamine **5**:712*T*
scops owl **5**:474
Scorpaeniforms **3**:468
Scorpio maurus **1**:208
Scorpiones
  Amazon ecosystems **1**:246*T*, **1**:247
  characteristics **1**:207–208, **1**:207*F*
  ecological characteristics **1**:208
  genera and species **1**:202*T*
  major lineages **1**:209
  phylogenetic relationships **1**:205*F*
  phylogeny and taxonomy **1**:208–209
  reproduction and growth **1**:208
scorpionfish **3**:467–471
scorpionflies **4**:392*T*
Scorpionidae **1**:209
Scorpionoidea **1**:209
scorpions **1:202–217**
  characteristics **1**:207–208, **1**:207*F*, **4**:392
  ecological characteristics **1**:208
  endangered species **3**:225–226
  genera and species **1**:202*T*
  major lineages **1**:209
  Near East ecosystems **5**:460*T*, **5**:468–469
  overharvesting/overcollecting **3**:223
  paleontology **1**:204–205
  phylogenetic relationships **1**:205*F*
  phylogeny and taxonomy **1**:204, **1**:208–209
  reproduction and growth **1**:208
Scorpiopidae **1**:209
Scotland **1**:677*F*
Scots pine **1**:628–629, **7**:212–213
Scottnema lindsayae **3**:94, **7**:429
Scrippsiella trochoidea **6**:30
scripture **4**:650–651
Scrophulariaceae **1**:486*T*, **2**:374*T*, **3**:129*T*, **4**:71, **6**:315*T*
scrub
  Australia **1**:292–297, **1**:294*T*
  predation effects **6**:237–239
  protected area systems **6**:495*T*
  South America **6**:600–604*T*, **6**:607*T*
  termites **4**:432*T*
  tropical Asia **1**:262–263*T*, **1**:266–267
scuds **3**:182*T*, **4**:371*T*
sculpin fish **3**:467–471
Scutelleridae **7**:294
scuticociliates **6**:295–296, **6**:296*F*
Scutigera coleoptrata **5**:422*F*
Scutigerella immaculata **5**:424–425
Scutigeromorpha **1**:246*T*, **5**:421–423, **5**:422*F*
scuttle flies **3**:489–491*T*, **3**:492
Scydmaenidae **1**:327*T*
Scylla serrata **5**:16
Scyphiphora **5**:11*T*
Scyphozoa **5**:669–670*T*
sea anemones **4**:381*T*, **5**:118–120, **5**:119*F*
sea bass **3**:468
sea cows
  characteristics **4**:694
  late Quaternary extinction event (LQEE) **4**:724–725*T*
  number of species **4**:690*T*
  taxonomy **5**:65*T*
sea cucumbers **2**:18*F*, **4**:381*T*
seadragons **3**:467–471
sea fans **2**:339–340, **4**:381*T*
sea firs **4**:381*T*
seafloor spreading **1**:583, **1**:584*F*
sea-gooseberries **6**:25
seagrass **6:540–550**
  biological diversity
    invertebrates **6**:547–548
    macroalgae **6**:547
    microbial communities **6**:546–547
    species richness **6**:546–547
    vertebrates **6**:548
  bioprospecting **1**:593
  characteristics **6**:540, **6**:541*F*
  distribution and controls
    global scale **6**:543–544, **6**:544*F*, **6**:545*F*
    local species diversity **6**:545–546
    regional scale **6**:544–545, **6**:545*F*
  ecosystem services **6**:77
  evolutionary development **6**:540–541
  functional role **6**:548–549
  genera and families **6**:542*T*
  genera and species **6**:541*T*
  genetic diversity **6**:546
  geographic distribution **6**:543*F*
  low-light environments **3**:360–361, **3**:360*F*, **3**:361*F*
  management strategies **6**:549
  nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
  nitrogen cycle impacts **5**:543
  origins **6**:540–541
  seagrass insulation **1**:593
  seagrass meadow **6**:540, **6**:541*F*
  species richness
    biogeography **6**:542–543, **6**:543*F*
    biological diversity **6**:546–547
    taxonomic status and uncertainty **6**:541–542, **6**:542*T*
  threatened species **7**:201*T*
  threats and challenges **6**:549
  ultraviolet radiation effects **7**:299–300
seahorses **3**:410, **3**:467–471
sea ice regions **6**:300
sea lamprey **4**:363
sea lettuce **3**:358–359
sea level changes
  current trends **2**:84
  estuarine ecosystems **3**:302
  extinction theories **2**:517
sea lice **2**:532
sea lilies **4**:381*T*
sea lions
  convergent evolution **4**:701*T*
  ecological footprint analysis (EFA) **2**:708
  migratory patterns **5**:314
  overharvesting/overcollecting **3**:283
  pollution effects **3**:192–193
  species diversity **4**:693
  taxonomy **5**:65*T*
seals
  adipose tissue **7**:34–35
  convergent evolution **4**:701*T*
  deoxyribonucleic acid (DNA) evolution **2**:56*F*
  ecological footprint analysis (EFA) **2**:708
  migratory patterns **5**:314
  number of species **4**:690*T*
  overharvesting/overcollecting **3**:283
  species diversity **4**:693
  taxonomy **5**:65*T*
seamounts **2**:15*F*, **2**:19–20, **2**:20*F*, **5**:53
Seamounts Marine Management Area **5**:449–450
Sea of Cortez, The (Steinbeck) **4**:652
sea otters
  anthropogenic impacts **5**:58–60
  biomonitoring processes **3**:125
  characteristics **5**:68

ecological footprint analysis (EFA) **2**:708
extinctions **5**:87*T*
fur trade **3**:191–192
Holocene postglacial spread **2**:551–552*T*
introduced parasites–species decline relationships **2**:528*T*
kelp beds **3**:283, **4**:8–9, **4**:443–446, **5**:51, **5**:58–60, **6**:232–233, **6**:233*F*, **6**:234*F*, **6**:247*F*
keystone species study **4**:443–446, **4**:444*F*, **4**:445*F*
oil spills **3**:192–193, **3**:256
overharvesting/overcollecting **3**:283
species diversity **4**:693
trophic cascades **7**:259–260
sea pens **4**:381*T*
sea robins **3**:467–471
sea-run cutthroat trout **6**:522, **6**:526
seas *see* marine ecosystems; oceans
sea scorpions **2**:400
sea slugs **2**:333*T*, **5**:118–120, **5**:119*F*
seasonally dry tropical forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
seasonal rainfall **2**:464–466
sea spiders **2**:400, **4**:381*T*
sea squirts **4**:381*T*
sea stars **2**:17*F*, **4**:381*T*, **4**:442–443, **4**:443*F*, **5**:37, **6**:232, **6**:722, **6**:723*T*
sea surface temperatures
climate change effects **2**:60
coastal deserts **2**:463*F*
coral reef ecosystems **2**:341–342, **2**:342*F*
El Niño Southern Oscillation (ENSO) events
coral bleaching **3**:160–161, **3**:160*F*
increased temperatures **3**:159–160
nutrient level shifts **3**:160
ocean current shifts **3**:159–160
pelagic ecosystems **5**:683–686
sea swallows **2**:141–142
sea turtles
anthropogenic impacts **6**:410–411
bycatch **3**:447
Census of Marine Life project **2**:12*F*
characteristics **6**:400–401
countryside biogeography studies **2**:356
habitat destruction **5**:317
migratory patterns **5**:314
sea urchins
calcium carbonate ($CaCO_3$) grazing **3**:612
classifications **4**:381*T*
coral reef ecosystems **2**:337, **2**:338–340, **5**:61
disease **5**:61
ecological footprint analysis (EFA) **2**:708–709
eutrophication impacts **3**:355–356
kelp beds **3**:283, **4**:8–9, **4**:443–446, **5**:51, **5**:58–60, **6**:232–233, **6**:233*F*, **6**:234*F*, **6**:247*F*
keystone species study **4**:443–446, **4**:444*F*, **4**:445*F*
landscape genetics **4**:516
mass extinctions **3**:417–418
sea otter predation **4**:8–9, **4**:443–446, **5**:58–60, **6**:232–233, **6**:233*F*, **6**:234*F*, **6**:247*F*, **7**:259–260
sea-viewing wide field-of-view sensor (SeaWiFS) satelliite **2**:620*F*
seawater **1**:561
seaweed flies **3**:489–491*T*
seaweeds
disease prevention **1**:592
grazing effects **4**:10
marine ecosystems **5**:45–46
Sebastes alutus **3**:450
Secale cereale **2**:662–663
Secchi transparency **3**:581–582
Sechelleptus seychellarum **5**:423–424
Sechura Desert **2**:469
secondary production measurement
bioenergetic analysis **3**:67
biomass accrual **3**:67
birth and death rate analysis **3**:67–68
egg ratio analysis **3**:68
functional type analysis **3**:68
instar analysis **3**:68
measurement approaches **3**:67
second law of thermodynamics **2**:703
Second Welfare Theorem **5**:145–147
secretary bird **1**:608
Securinega virosa **1**:67
Secutor ruconius **3**:478*F*
sedges **1**:275, **1**:278–279, **2**:375*F*, **7**:369*T*
Sedgwick, Adam **2**:419
Sedili Kecil swamp forest **1**:276*T*
sediment-hosted hydrothermal systems **7**:327–328
seed banks
collection value **3**:644–645
community seed banks **3**:644
current status **3**:643–644
effectiveness **3**:642
historical perspective **3**:642–643
management strategies
base versus active collections **3**:647–648
basic procedures **3**:646–647
characterization **3**:649
design principles **3**:650
distribution to users **3**:649–650
duplication **3**:649
evaluation assessments **3**:649
packaging **3**:648
regeneration systems **3**:650
seed cleaning **3**:647
seed collecting **3**:647
seed drying **3**:648
storage standards **3**:647–648
storage temperature **3**:648–649
viability monitoring **3**:649
plant conservation **6**:63, **6**:85
scientific principles
genetic considerations **3**:646
longevity predictions **3**:645–646
seed storage conditions **3**:645–646
wetland ecosystems **7**:397
wild species **3**:644
seed bugs **7**:294
seedless vascular plants **6**:59
seeds
fire-stimulated germination **3**:437–438
geographic distribution **3**:131*T*
global food usage **3**:130*T*
seed dispersal
domesticated ruminants **1**:708–709
forest management impact **1**:453–454
grazer-based ecosystem restoration and management **4**:15
hemiparasites **4**:74
keystone species study **4**:447
mutualistic partnerships **2**:328–329, **4**:447
plant–animal interactions
ants **6**:50
beetles **6**:50
birds **6**:50
dispersal syndromes **6**:50, **6**:51*T*
disperser types **6**:50
evolutionary dead ends **6**:52
fish **6**:50
frugivores **6**:50–51
importance **6**:49–50
insects **6**:50
mammals **6**:50
plant adaptations **6**:50
reptiles **6**:50
seed predators **6**:51–52
species persistence **6**:61–62
vertebrates **6**:50
seed dispersal–predation relationship **7**:283–284
tropical forest regeneration **7**:283–284
seed industries **1**:91
seed production **1**:85–86
Seed Savers Exchange **1**:460–461
seed shrimp **3**:182*T*
seed storage **7**:24–25
seepage lakes **3**:560
segmented sea worms **4**:381*T*
Sehima nervosum **1**:69
Seidlitzia rosmarinus **5**:486
seines **6**:421
sei whales **5**:81*T*, **5**:88–89
Selaginella **6**:316*T*
self-cleaning lotus paint **1**:590
self-designed wetlands **7**:379
self-pollination **2**:659–660
Selous Game Reserve **5**:445
semaphore cactus **4**:361

semen conservation programs **4**:319–320, **4**:321*T*
Semicossyphus pulcher **6**:233
semideciduous forests **4**:432*T*
semi-deserts
  Asia **1**:262–263*T*, **1**:269–270
  protected area systems **5**:444*T*
semiochemicals
  butterfly–ant symbiotic relationship **1**:655–656
  organic insecticides **5**:694
semipalmated sandpipers **7**:185
semispecies taxonomy **7**:46–47
Semnopithecus ajax **6**:256–258*T*
Semnopithecus priam thersites **6**:256–258*T*
Semper's warbler **7**:184
Senecio **5**:217, **5**:219, **7**:275–276
Senecio adglacialis **4**:103
Senecio cleefi **4**:103
Senecio cocuyanus **4**:103
Senecio guicanensis **4**:103
Senecio heritieri **3**:206–208
Senecio magellanicus **4**:106–107
Senecio pasqui-andinus **4**:103
Senecio supremus **4**:103
Senecio vulgaris **3**:703
Senegal
  botanical gardens **6**:519
  floodplain agriculture **7**:146
  grass and shrub savannas **1**:67
  termites **4**:430*T*
Senegalia mellifera **4**:515
Senegal River **7**:146, **7**:151–152, **7**:153–154
Senna spectabilis **6**:561
sennoside **5**:712*T*
sensory rhodopsins **6**:281*F*
Sentinelese Island **6**:221–222
Sepsidae **3**:489–491*T*
sequential gel electrophoresis **3**:663
Sequoia **7**:118
Sequoiadendron **7**:118
Serangoon estuary **1**:276*T*
Serbia and Montenegro
  botanical gardens **6**:519
  vegetation surveys **3**:341–344*T*
Serengeti ecosystem
  colonial legacies and restitution **4**:439–440
  disease outbreaks **3**:191, **3**:411
  interspecific competition **2**:190, **2**:190*F*, **2**:190*T*
  megaherbivores **5**:226–227
  short- versus long-term ecosystem dynamics **2**:487–488, **2**:488*F*
serial overfishing **3**:445
Sericornis keri **2**:75*F*
Serjania **5**:489
Serpulidae **7**:417*F*
Serradium hirsutipes **5**:430–431
Serradium semiaquaticum **5**:430–431
Serratia marcescens **2**:339–340, **2**:528*T*, **2**:529*F*
Serritermes **4**:409*T*
Serritermitidae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*
service corridors **7**:203*F*
sesame oil **2**:663
Sesamum indicum **2**:663
Sesarma **5**:15–16, **5**:15*F*
Sesarma meinerti **5**:13
Sesbania aculeata **2**:389
Sesbania exasperata **1**:165
sesquiterpenes **2**:427
sessile oaks **1**:612–613
Sesuvium **2**:376*F*
Setaria italica **2**:663, **3**:131–132
Setodes alala **5**:463*T*
Setonix brachyurus **6**:455
17β-estradiol **3**:118
several-little-all-too-small (SLATS) strategy **6**:650
Sevier Orogeny **5**:563
Sevilleta National Wildlife Refuge (SNWR)
  case study **4**:483–486, **4**:484*F*, **4**:485*F*
sewage **1**:512, **1**:512*F*
Sewell, Anna **4**:653
sex change patterns **6**:574
sex ratios **5**:677
sexual selection
  parasite–host coevolution **2**:144
  speciation **6**:632–633
  species concept theory **6**:682
Seychelles **6**:493–494, **6**:519
Seychelles magpie-robin **7**:188
Seychelles paradise flycatcher **7**:184
Seychelles warbler **1**:601–602, **7**:188
shad **3**:260, **5**:314–315
shade coffee plantations **1**:420–421, **1**:421, **1**:422*F*, **1**:423*F*
shading and overgrowth **4**:359–360
Shaman Pharmaceuticals **1**:598
Shannon index **1**:403, **2**:205–207, **2**:438, **2**:650, **2**:651*T*, **4**:616*F*, **5**:202–203, **6**:652–653
Shannon–Weaver index **1**:141*F*, **2**:209–210, **2**:438, **4**:10*F*, **4**:339–340, **5**:49, **6**:32–33
Shan States Phayre's langur **6**:256–258*T*
sharks
  bioprospecting **1**:591
  bycatch **3**:447
  Census of Marine Life project **2**:12*F*
  characteristics **3**:456–457, **3**:458–459, **3**:459*F*
  endangered species **5**:449–450
  extinction theories **2**:517
  hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
  major threats **7**:203*F*
  overfishing **3**:410
  predation effects **6**:236
  river ecosystems **6**:479–480
  shark swimsuits **1**:591
sheep **1:702–711**
  biotic relationships
    birds **1**:710
    invertebrate fauna **1**:709–710
    natural predators **1**:709
    parasites and pathogens **1**:710–711
    plant tolerance–grazing relationship **1**:706–708, **1**:707*F*
    seed dispersal and germination **1**:708–709
    small mammals **1**:710
  desert ecosystems **2**:476*T*
  domestication **1**:702–703
  ecological role
    desertification **1**:704
    grazing effects **1**:703–704
    land degradation **1**:704–705
    nutrient cycling **1**:705–706, **1**:705*F*
    over-enrichment factors **1**:705–706
    plant diversity–grazing intensity relationship **1**:704*T*
    vegetation impacts **1**:703–704
  introduced species
    Australia **1**:80
    Brazil **1**:83
    India **1**:82
    South Africa **1**:81
    United Kingdom **1**:79
    United States **1**:76–77
  risk status classification **1**:644*F*
sheep blowfly **4**:301
sheephead fish **6**:233
sheepshead mushrooms **3**:637*T*
shelf fungi **3**:629–630
Shelford, Victor **3**:239
shellfish reefs **5**:452–453, **5**:453*T*
Shengjin Hu **1**:276*T*
shepherd's purse **3**:131–132
Sherman's Lagoon (cartoon) **2**:15
Sherwin–Williams biodiversity model **7**:91
Shewanella **5**:293–294
shield deserts **2**:460–461
shifting balance theory **6**:197
shifting cultivation **2**:452, **2**:452*T*, **7**:231–232
Shigella flexneri **2**:636*F*
shiitake mushrooms **3**:637–638, **3**:637*F*, **3**:637*T*
Shijiu Hu **1**:276*T*
shiners **7**:390
shingle-backed lizards **7**:32
Shinji (Lake) **1**:276*T*
Shinto **6**:565
shipping trade **3**:700–701, **3**:701*F*
shipworms **5**:121
Shizosaccharomyces pombe **6**:355
shoal grass **3**:360–361
Shorea **1**:263
Shorea albida **7**:272
Shorea cuspidata **1**:514
Shorea henryana **1**:267
Shorea kuantanensis **1**:514
Shorea robusta **1**:70, **1**:266, **6**:225, **7**:118, **7**:273–274
Shorea roxburghi **1**:70
Shorea siamensis **1**:70, **1**:267, **1**:453–454
shore bugs **7**:293
shore flies **3**:489–491*T*
shore limpets **3**:202*T*
shoreline protection strategies **5:451–458**
  biodiversity impacts **5**:453–456, **5**:455*F*
  current status **5**:452–453, **5**:453*T*
  historical perspective **5**:452

importance **5**:451–452
research summary **5**:456
shore pine **7**:385
short fallow **6**:325–327, **6**:326*F*, **6**:326*T*
short-finned pilot whales **3**:192–193
short-horned grasshoppers **4**:194*T*
Shortridge's langur **6**:256–258*T*
short-rotation woody biomass **1**:531–532
short-tailed pythons **5**:604
short-toed eagles **5**:474
short tussock **4**:348–349, **4**:349*F*
shrew opossums **4**:687, **4**:687*T*
shrews
boreal forests **1**:631–632
characteristics **4**:690–691
food webs **7**:266*F*, **7**:267*T*
Great American Biotic Interchange (GABI) **2**:584*F*
number of species **4**:690*T*
wetland ecosystems **7**:390
shrikes **1**:604*F*
shrimp
aquaculture **1**:190*T*, **1**:192*F*, **1**:194, **3**:448–449
classifications **4**:381*T*
discovered species **2**:3–8, **2**:5*F*
fishing bycatch **3**:446
hydrothermal vent ecosystems **7**:329
mangrove ecosystems **3**:51–53, **3**:52*F*, **5**:15–16, **5**:16
number of species **4**:371*T*
shrimpfish **3**:467–471
shrubby tororaro **4**:194*T*
shrublands
African savannas **1**:66*F*, **1**:67
Australia **1**:292–297, **1**:294*T*
biodiversity change scenarios **4**:6*F*
boreal forests **1**:628–630
C_4 plants **2**:371*T*
characteristics **6**:333–334
conservation status **3**:171–173*T*
current global land cover **4**:557*F*
energy use impacts **3**:262–263
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
neotropical seasonally dry forests **5**:496
nitrogen biogeochemical cycle **6**:152–154
nitrogen deposition **5**:526*T*, **5**:534*F*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
protected area systems **6**:492, **6**:492*T*, **6**:495*T*, **6**:499*T*
South America **3**:111*F*, **3**:114, **3**:115, **6**:600–604*T*, **6**:607*T*
threatened birds **7**:181
transition matrices **6**:325–327, **6**:326*F*, **6**:326*T*
tree and shrub savannas **1**:66*F*, **1**:67
tropical Africa **1**:52–53
shy albatross **2**:543
Sialis **6**:26
siamangs **6**:251*T*, **6**:256–258*T*
Siberian boreal forests **1**:630*T*
Siberian taimen **3**:464–465
Siberian tigers **3**:373
Siberut pig-tailed langur **6**:256–258*T*
Siboglinidae **7**:417*F*
Sichuan–Yunnan highlands **1**:262–263*T*
sick building syndrome (SBS) **3**:635
side-blotched lizards **2**:720–721
Siderastrea glynni **3**:202*T*
sidewinders **2**:473*F*
siduron **4**:90*T*
Sierra Club **3**:139
Sierra Leone **6**:519
Sierra Madre de Chiapas **5**:563–564
Sierra Madre del Sur **5**:563–564
Sierra Madre Occidental **5**:563
Sierra Madre Oriental **5**:563
Sierra Nevadas **5**:563
sifakas **6**:251*T*, **6**:256–258*T*
Sigalionidae **7**:417*F*
Sigmodon hispidus **1**:365–368, **4**:471
Signor-Lipps effect **5**:157–158, **5**:158*F*, **5**:159*F*
Signor–Lipps effect **3**:414–415, **3**:538–539
Sika-Krka System (Slovenia) **7**:59*T*
Sikhism **6**:565
Silene danaensis **5**:485
Silent Spring (Carson) **2**:285–286, **3**:281–282, **4**:652–653
silicon (Si)
elemental partitioning **1**:554*T*
relative abundance **1**:553*F*
river water composition **3**:562*T*
silica supplies **3**:354–355
silk
moths **5**:398
spiders **1**:205–206
silky sifaka **6**:256–258*T*
Silurian period
arachnids **1**:204–205
arthropod radiation ecology **2**:399
crustaceans **2**:403
fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
marine diversity **5**:168*F*, **5**:169*F*
mass extinctions
extinction intensities **5**:162*F*
marine animals **1**:413*F*, **1**:493*F*
notable mass extinctions **5**:161*F*
number of families **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
originations **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
silverback jackal **6**:573
silverfish **4**:392*T*, **7**:91–93, **7**:92*F*
silver flies **3**:489–491*T*
silver salmon **6**:523–524
silversides **3**:467
silverswords **1**:30–32, **1**:30*F*
silvery gibbon **6**:256–258*T*
Silybum marianum **2**:543
Simaethistoidea **5**:387–388*T*
Simaruba amara **1**:161–162
simazine **5**:696, **5**:699*T*
Simias concolor concolor **6**:256–258*T*
Simias concolor siberu **6**:256–258*T*
similarity tests **2**:37
Simmondsia chinensis **2**:468–469, **7**:25–26
Simocephalus **6**:26
Simon, Julian **6**:206
simple sequence repeats (SSRs) **3**:667
simple sugars **6**:113*T*
simponas **6**:251*T*
Simpson Desert **2**:468
Simpson, G. G. **2**:201, **3**:394, **6**:3–4, **7**:101–102
Simpson index **2**:438–439, **2**:650, **2**:651*T*, **6**:652–653
Simuliidae **3**:489–491*T*
Sinai *see* Near East ecosystems
Singapore
animal species **1**:261*T*
botanical gardens **6**:519
deforestation **1**:510, **4**:663
ecosystems
mangroves **1**:275*T*, **1**:513*T*
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical moist forest **1**:277*T*
tropical rain forests **6**:318*T*
plant species **1**:259*T*
urbanization **4**:560*F*
see also Southeast Asia
single-large-or-plentifully-patchy (SLOPP) debate **6**:649
single large or several small (SLOSS) debate **2**:240–241, **4**:495, **6**:649
single nucleotide polymorphisms (SNPs) **2**:274–275, **3**:664, **3**:667
Single Population Analysis and Record Keeping System (SPARKS) **1**:662–663, **7**:475–476
single-species metapopulation equation **2**:603
Sinnakarachchi Lagoon **1**:276*T*
Sinomastodon **5**:224*T*
Siphonaptera **4**:308–312, **4**:309*T*, **4**:392*T*, **4**:393*F*
Siphonaria compressa **3**:202*T*
Siphonodon **5**:623*F*
Siphoviridae **5**:130*F*, **5**:134*T*, **5**:135
Siponaria lessoni **5**:59*F*
Sipuncula
classifications **4**:381*T*, **7**:408
fossil record **3**:684*F*
geographic distribution **7**:418
global diversity **4**:385*T*
known and estimated species **4**:384*T*
marine ecosystems **5**:50*T*
morphology
anterior ends **7**:410*F*, **7**:412*F*
chaetae **7**:409–410, **7**:414*F*
circulatory system **7**:415
digestive tract **7**:415
general characteristics **7**:408–409
gills **7**:415
gonads **7**:415–416
head **7**:409
musculature **7**:414–415
nephridia **7**:415
nervous system **7**:410–411
posterior ends **7**:410*F*
pygidium **7**:409
segments **7**:409, **7**:413*F*

Sipuncula (*continued*)
sense organs **7**:411–414
septa **7**:415
number of species **5**:669–670*T*
phylogenetic and taxonomic relationships **7**:416–419, **7**:417*F*
reproductive biology **7**:416
sampling methods **4**:381*T*
Siranda Lake **1**:276*T*
SIR compartment models **2**:523–525, **2**:525*F*
sirenians
adipose tissue **7**:34
Census of Marine Life project **2**:12*F*
characteristics **2**:586, **4**:694, **5**:65*T*, **5**:67–68
evolutionary history **4**:685, **5**:69
extinction/near extinction over evolutionary time **5**:75
number of species **4**:690*T*
representative species **2**:587*F*
species diversity **4**:690*T*
taxonomy **5**:65*T*
sirens **7**:389–390
Siricoidea **4**:178
Siskiyou Mountains **5**:563
Sistema Nacional de Datos Biológicos (SNDB) **1**:506–507
sit-and-wait lizards **6**:406–408, **6**:407*T*
Siteroptes graminum **1**:215
Sitka spruce **6**:125
Sittidae **4**:33–35, **4**:34*F*, **6**:308*T*
Sivatherium **5**:231–233
Sivatherium maurusium **5**:225*T*
sixth extinction wave **2**:351
skates **3**:456–457, **3**:459–460, **3**:460*F*
Skeletonema costatum **3**:355–356, **4**:674, **5**:684*F*, **6**:30
Skellam's equation **5**:324–325, **5**:325*F*
skin cancer **7**:302
skinks **3**:144–145, **4**:194*T*, **5**:472, **6**:402
skipjack tuna **3**:481, **6**:422*T*
skipper butterflies **5**:370
skipper flies **3**:489–491*T*
Skjern River Restoration Project, Denmark (case study) **7**:373–374, **7**:375*F*
skunks **1**:685, **2**:584*F*, **2**:590*T*
sky-island formation **2**:466–467, **2**:467*F*
Sky Islands Wilderness Network (SIWN) **2**:258
skylarks **4**:259
Sladentiaceae **1**:260*T*
slash-and-burn systems **6**:**551–562**
alternative systems
improved short fallow systems **6**:560–561
pastures **6**:560
plantation crops **6**:560
basic concepts **6**:551
carbon sequestration **6**:561–562
clearing effects **6**:554
consequences
bacteria and fungi reductions **6**:556
heat evolution **6**:554–555
invertebrate loss **6**:556
plant species loss **6**:556
soil properties and nutrients **6**:555–556
vertebrate loss **6**:556
weeds **6**:556
cultivation phase effects
characteristics **6**:556–557
nutrient export **6**:558
pest and disease suppression **6**:557–558
prolonged cropping phase effects **6**:558
tillage **6**:556–557
weed and crop residue management **6**:558
weeding operations **6**:557
fallow/crop sequences **6**:553–554
fallow length **6**:552, **6**:552*F*
fallow phase effects **6**:558–559
geographic distribution **6**:553
historical perspective **6**:553
hunter–gatherer societies **6**:567
labor requirements **6**:553
purpose
biomass removal **6**:551
declining crop yields **6**:551–552
economic factors **6**:551
nutrient release and addition **6**:551
pest and disease avoidance **6**:551
weed avoidance and suppression **6**:551
scale dependence
spatiotemporal scales **6**:559–560, **6**:560*F*
topographic preferences and fragmentation **6**:559
traditional conservation practices **7**:231–232
tropical forests **6**:339
slash–mulch systems **7**:232
slash pine **6**:455
sleepers **3**:467–471
sleeping sickness **5**:660*T*, **5**:668, **5**:671*T*
slender loris **6**:251*T*
sliders **7**:390
slime eels *see* hagfish
slime molds
Dictyosteliomycota (cellular slime molds) **3**:631, **6**:289–293, **6**:289*T*, **6**:290*F*
Myxomycota (true slime molds) **3**:630–631, **3**:630*F*, **3**:631*F*, **6**:289–293, **6**:289*T*, **6**:290*F*
nomenclature code **5**:555
Plasmodiophoromycota (endoparasitic slime molds) **3**:631
schematic diagram **6**:290*F*
species richness **6**:289–293, **6**:289*T*
Slockoceros **4**:724–725*T*
Sloping Land Conversion Program (SLCP) **3**:101
SLOSS debate **2**:240–241, **4**:495, **6**:649
sloths **2**:584*F*, **2**:590*T*, **4**:689, **4**:690*T*, **4**:724–725*T*
slough **7**:369*T*
Slovakia
botanical gardens **6**:519
vegetation surveys **3**:341–344*T*
Slovenia
botanical gardens **6**:519
vegetation maps **3**:346*T*
slow lorises **6**:251*T*
small-headed flies **3**:489–491*T*
smallmouth bass **1**:12*F*, **6**:153*F*
small spikerush **3**:209
small subunit ribosomal DNA (ss rDNA) **6**:345
small vesper mouse **1**:365–368
smelt **3**:464–465, **6**:522–523
smelting industry **1**:140–141, **1**:141*F*, **1**:142*F*
Smilax aspera **2**:543
Smilodon **4**:724–725*T*, **4**:727*T*
Smith, Adam **3**:288–289
Smithsonian Institution **5**:405–406, **5**:408*F*, **5**:409*F*
smog **3**:407
smoke aerosols **3**:442
smokers **2**:6*F*, **4**:391
smut fungus **3**:629–630, **3**:633*T*
snail darter **2**:223
snail kite **7**:186
snails
acidification effects **3**:584–585, **3**:584*F*
adaptive radiation **1**:33
American Southeast rivers **3**:185*F*
characteristics **4**:392
coral reef ecosystems **2**:333*T*
critical pH levels **6**:153*F*
as disease vectors **1**:368, **1**:368*F*
endangered species **3**:226
freshwater ecosystems **4**:374
genetic diversity **3**:674*F*
Great Artesian Basin, Australia **3**:186, **3**:187*F*
invasive species **1**:522*F*, **3**:222, **4**:359
mangrove ecosystems **4**:337, **5**:17–18, **5**:17*F*
Near East ecosystems **5**:460*T*, **5**:464*T*
number of species **3**:182*T*, **4**:371*T*
river ecosystems **6**:478*T*
surface water acidification **1**:10*T*
snakeflies **4**:392*T*
snakeheads **3**:467–471
snakes
alpine ecosystems **1**:154–155
Australia **1**:302*F*
brown tree snakes
dispersal biogeography **2**:543, **2**:558–559
diversity loss **1**:621
invasive species **3**:218, **3**:410–411, **3**:410*F*, **4**:143, **4**:357, **7**:184
characteristics **6**:402, **6**:408–409, **7**:340–341
commercial exploitation **6**:411
conservation status **3**:216*T*, **6**:411
desert ecosystems **2**:470*T*, **2**:472–474, **2**:473*F*
foraging behaviors **6**:409
forest canopies **3**:513–514
mangrove ecosystems **5**:13
Near East ecosystems **5**:472
oil-palm plantations **5**:606–607
prey types **6**:408
river ecosystems **6**:479–480
South America **3**:115
threatened species **1**:515–516

venom evolution **6**:409
wetland ecosystems **7**:390
see also Reptilia/reptiles
snapper **3**:467–471
snapping turtles **7**:390
Snares Islands **5**:444
Sneath, Peter **7**:102–104
snipe flies **3**:489–491*T*
snooks **3**:467–471
snow
snow cover **2**:680–681
South America **6**:607*T*
snowblindness **7**:302
Snow, John **4**:503
snow leopards **4**:693
Snow Leopard, The (Matthiessen) **4**:653
snowshoe hares **1**:631–632, **4**:697
snow valleys **1**:274–275
snowy owls **1**:236
Soay sheep **2**:726–727, **3**:127*T*
social behavior **6:571–579**
characteristics **6**:571
ecological consequences
anthropogenic activities and impacts **6**:577–578
demographic change **6**:577–578
environmental determinism **6**:575–577
inappropriate interventions **6**:578–579
social change **6**:576–577
formation patterns and mechanisms
benefits **6**:571–573
demographic factors **6**:573–574
ecological challenges **6**:572–573
egalitarianism/despotism **6**:574–575
influencing factors **6**:571–573, **6**:572*F*
kinship **6**:573–574
mating systems **6**:572–573
organizations and networks **6**:575, **6**:576*F*
phenotypic constraints **6**:572–573
social insects
Hymenoptera **4**:181–182
Isoptera **4**:412–413, **4**:412*F*, **4**:414*T*, **4**:421–422
special associations **3**:495
social parasitism **5:501–509**
adaptive behaviors **5**:501
background information **5**:501
conspecific brood parasitism (CBP)
definition **5**:501
prevalence **5**:508–509
species survey **5**:502–503
cowbird management **5**:507
fish **5**:503
host–parasite coevolution
cowbird–host systems
behavioral adaptations **5**:504–505
egg adaptations **5**:504
host defenses **5**:505
nestling and fledgling adaptations **5**:505
cuckoo–host systems
adaptive behaviors **5**:503
egg adaptations **5**:503
evolutionary development **5**:503
fecundity **5**:503
host defenses **5**:504
nestling and fledgling adaptations **5**:503–504
modeling approaches **5**:506
host population dynamics **5**:506–507
insects **5**:503
major groups **5**:502*T*
number of species **5**:502*T*
research outlook **5**:509
social behavior
foraging ecology **5**:507
mating systems **5**:507–508
spacing behaviors **5**:508
vocal behaviors **5**:508
species survey
black-headed ducks **5**:502, **5**:502*T*
conspecific brood parasitism (CBP) **5**:502–503
cuckoo-finches **5**:502, **5**:502*T*
Cuculidae **5**:501–502, **5**:502*T*
Icteridae (cowbirds) **5**:502, **5**:502*T*
Indicatoridae **5**:502, **5**:502*T*
viduine finches **5**:502, **5**:502*T*
social well-being
allocation mechanisms **3**:26
basic concepts **3**:26
constant-equivalent consumption measures **3**:27
cost–benefit rule **3**:27
population changes **3**:26–27
sustainable development **3**:27–28
Society Islands **5**:591*F*, **5**:592*F*, **5**:595*F*, **6**:93
sociocultural factors **6:563–570**
language **4**:270–271, **4**:270*T*, **6**:566
religious traditions **6**:563–566
research summary **6**:569
resource management
ecosystem-like concepts and applications **6**:386–387, **6**:387*T*
indigenous peoples **4**:272, **6**:384–385, **6**:384, **6**:566–568
traditional belief systems **6**:386, **6**:567–568
urban versus rural perceptions **6**:568–569, **6**:569*T*
sociological geobotany **4**:396*T*
sockeye salmon **3**:260, **3**:465*F*, **6**:522, **6**:523
Socorro dove **7**:188
sodium (Na)
elemental partitioning **1**:554*T*
inorganic herbicides **5**:695–696
relative abundance **1**:553*F*
river water composition **3**:562*T*
sodium arsenite **5**:696, **5**:699*T*
Soergelia **4**:724–725*T*
soft corals **2**:335
soft-shelled turtles **5**:473, **7**:390
Soil and Water Assessment Tool (SWAT) **5**:355–356
Soil Conservation Service (SCS) **2**:283–284
soils **6:580–589**, **6:590–598**
abiotic–biotic interactions **6**:580, **6**:582*F*
acidic deposition
alteration consequences **1**:5
aluminum mobilization **1**:5–6
calcium loss **1**:5, **1**:6*F*
geographic distribution **1**:7*F*
nutrient cation loss **1**:5, **1**:6*F*
sulfur and nitrogen accumulations **1**:6
acidification **1**:143*F*, **4**:225*F*, **4**:228
Africa **1**:46
agroecosystems
planned diversity **3**:12
unplanned diversity **3**:13
air pollution impacts **1**:143*F*
anthropogenic soils **4**:285–287, **4**:286*F*
arable land **6**:592*F*
biodiversity
biotic diversity **6**:582*F*, **6**:583
disturbance effects **6**:588–589
domains
bacteria **6**:583–584
characteristics **6**:583
eukaryotes **6**:585–586, **6**:586*T*
phylogenetic relationships **6**:583
prokaryotes **6**:583
Protozoa **6**:586, **6**:587*F*
tree of life **6**:584*F*
evolutionary history **6**:582–583
research summary **6**:589
spheres of influence **6**:582*F*, **6**:583
viruses **6**:584–585
biodiversity–ecosystem services relationship **1**:346–348, **3**:347
biological soil crusts **6**:339
biomonitoring processes **3**:122–123, **3**:122*T*
bioturbation **3**:90*T*
boreal forests **1**:627
carbon cycle management **1**:681–682
carbon depletion **4**:527
coarse woody debris (CWD) **1**:450
composition **6**:590
conservation and restoration
agricultural methods
biotic community manipulation **6**:594–595
diversified cropping **6**:594
general characteristics **6**:593–594
organic amendments **6**:594
pollutant bioremediation **6**:594
reduced tillage practices **6**:593–594
twentieth century conservation **2**:283–284
assessment and monitoring programs
basic process **6**:595–596
government policies and regulations **6**:597–598
indicator measures **6**:596
indicator measures
basic concepts **6**:596
biological attributes **6**:597
chemical attributes **6**:596–597
physical attributes **6**:596
research summary **6**:598
twentieth century conservation **2**:283–284
urban soils **6**:595
deep-soil habitats **7**:51
degraded soils **1**:543, **1**:544*F*
deserts **2**:461–464, **2**:474
ecosystem components

soils (*continued*)
abiotic–biotic interactions **6**:580, **6**:582*F*
organizing functions **6**:580
soil-forming factors **6**:580, **6**:581*F*
ecosystem function measurements **3**:77*F*, **3**:79–80
environmental impacts **3**:281
enzyme assays **5**:268, **5**:269*T*
food webs **3**:90*T*, **7**:266*F*, **7**:267*T*
forest alteration effects **1**:450
formation mechanisms **6**:580, **6**:581*F*
functional role **6**:580, **6**:590–591
global change indicators **4**:225–226, **4**:225*F*, **4**:228
global species richness **3**:710
grassland ecosystems **4**:2, **4**:2*F*, **4**:4
grazing effects **1**:705–706, **6**:337–338
hydric soils **7**:378–379
landscape legacies **4**:527
major degradation processes **6**:592*F*
major processes
decomposition **6**:580–581
nitrogen cycle **6**:581–582, **6**:582*F*
microbial activities **5**:303–304
microorganisms
bioremediation **6**:594
nitrogen deposition effects **5**:530
soil quality indicators **6**:597
nematodes **7**:430
nitrogen depletion **4**:527
nitrogen deposition effects **5**:530
nitrogen fixation **1**:560–561, **1**:562*F*, **6**:581–582, **6**:582*F*
nonextremeophilic Archaea **1**:224*T*
pesticide treatments **5**:697–698
phosphorus gradient–hierarchical partitioning measures **3**:606–608, **3**:607*F*
planned diversity **3**:12
plant–soil interactions **6**:**109–128**
basic concepts **6**:109–111, **6**:110*F*, **6**:111*F*
conservation and management practices
anthropogenic impacts **6**:125
atmospheric carbon dioxide (CO_2) concentrations **6**:126, **6**:127*F*
ecosystem functions **6**:126–127
exotic species invasions **6**:125
general discussion **6**:125
nutrient pollution **6**:125–126
restoration programs **6**:125
feedback processes **6**:111*F*
geographic patterns **6**:123, **6**:123*T*
litter deposition **6**:119–121, **6**:120*T*
nutrient acquisition
individual plants **6**:116–117, **6**:117*F*, **6**:118*F*, **6**:119*F*
root growth patterns **6**:119*F*
spatial patterns **6**:119*F*
stand scale **6**:117–118
plant–microbiota competitive interactions **6**:118–119
plant tissue chemistry **6**:119–121, **6**:120*T*
regional characteristics **6**:123, **6**:123*T*
root–soil interface
large-scale processes **6**:123
mesoscale processes **6**:116–117
nutrient and water uptake **6**:113–115, **6**:114*F*
reduction–oxidation potential **6**:115*T*
rhizosphere **6**:111–113, **6**:112*F*, **6**:113*T*
small-scale processes **6**:111–113
soil structure **6**:115–116
schematic diagram **6**:110*F*
soil biota **6**:121–123, **6**:122*F*
within-biome variations **6**:123–125, **6**:124*F*
polyphasic characteristics **6**:580
prior land use effects **7**:281
productive potential regulation **1**:346–348, **3**:347
savannas **1**:63–65
slash-and-burn systems
bacteria and fungi reductions **6**:556
invertebrate loss **6**:556
soil properties and nutrients **6**:555–556
soil fauna **5**:530, **6**:590*T*
soil fertility **2**:474, **7**:281
soil food webs **6**:448–449, **6**:449*F*
soil moisture prediction **4**:497–498, **4**:498*F*
soil organic matter **1**:113–114
soil quality
anthropogenic impacts **6**:591–593
assessment and monitoring programs
basic process **6**:595–596
government policies and regulations **6**:597–598
indicator measures **6**:596
erosion effects **6**:591, **6**:592*F*
high-quality soils **6**:591
indicator measures
basic concepts **6**:596
biological attributes **6**:597
chemical attributes **6**:596–597
physical attributes **6**:596
soil types **1**:581*F*
species distribution modeling (SDM) **6**:699–700
termites **4**:431–432, **4**:431*F*
thermal ecosystems **4**:112
tropical forests **7**:271
twentieth century conservation **2**:283–284
uncultured free-living organisms **1**:315
unplanned diversity **3**:13
water storage **3**:561*T*
wetland restoration programs **7**:378–379
Sokal, Robert **7**:102–104
Solanaceae **3**:129*T*, **3**:130*T*, **3**:522*T*
Solanum **6**:316*T*
Solanum melongena **2**:663
Solanum mozinianum **1**:120
Solanum pennelli **1**:463
Solanum sessiliflorum **4**:284–285, **4**:284*F*, **4**:289*T*
Solanum tuberosum **2**:383, **2**:663–664, **3**:128
solar energy **3**:253*T*, **3**:260–262, **3**:265*T*
solar heating **4**:112
Solarium **3**:130*T*
Solaster endeca **2**:17*F*
soldier flies **3**:489–491*T*, **3**:493
sole **3**:467–471
Solenodon cubanus **2**:587*F*, **4**:714*F*
Solenodon paradoxus **4**:713–714, **4**:713*F*, **4**:714*F*
solenodons
background and characteristics **4**:713–714, **4**:713*F*, **4**:714*F*
conservation efforts **4**:715
general discussion **4**:713–714
major threats **4**:714–715
Solenodontidae **4**:690–691
Solenogastres **5**:669–670*T*
Solenopsis geminata **4**:360
Solenopsis invicta **1**:78, **4**:360, **6**:50
Solenopsis xyloni **4**:360
solfataras **1**:219–220
Solidago altissima **4**:639–640
Solidago missouriensis **1**:707
Solifugae **1**:202*T*, **1**:205*F*, **1**:213, **1**:213*F*, **1**:246*T*
solitaires **1**:605–606
Solomon Islands **4**:270*T*, **4**:560*F*, **5**:592*F*, **5**:598*F*, **6**:507*T*
Solonchaks **2**:461–464
Solonetz **2**:461–464
solution caves **7**:50–51
solution hybridization **5**:264*T*, **5**:266–267
Somalia
botanical gardens **6**:507*T*
desert ecosystems **2**:467–468
domesticated ruminants **1**:702–703
urbanization **4**:560*F*
Somalian white-collared guenon **6**:256–258*T*
Somatiidae **3**:489–491*T*
somber-bellied Mentawai surili **6**:256–258*T*
songbirds **1**:607–608
song sparrows **4**:490, **4**:491*F*
song thrushes **4**:259
sonic hedgehog (*shh*) signaling gene **7**:54*F*
Sonneratia **1**:277, **5**:11*T*, **5**:14
Sonneratiaceae **5**:11*T*
Sonoran Desert **2**:468–469
Sonoran Orogeny **5**:563
sooty shearwaters **2**:7–8
sooty tern **6**:78
Sophora chrysophylla **2**:655
Sophora inhambanensis **1**:55
Sophora toromiro **6**:83
Sorbus alnifolia **1**:271
Sorbus aucuparia **4**:517
Sordaria fimicola **3**:673–674
Sørensen index **2**:444–445, **5**:207–208, **6**:659–660, **6**:660*F*
Sorex araneus **2**:552*F*
sorghum **1**:132*T*, **2**:383, **2**:659, **2**:663, **3**:32*F*
Sorghum bicolor **2**:636*F*, **2**:663
Soricidae **2**:583*F*, **6**:309*T*
Soricomorpha **2**:586–587, **2**:587*F*, **4**:690–691
Sorites **6**:291–292
Soulé, Michael **2**:255–257
source–sink dynamics **3**:152, **4**:496, **5**:324–325, **5**:325*F*

Souroubea **3**:522*T*
sourthern purple-faced langur **6**:256–258*T*
Sousa chinensis **5**:77*T*
Sousa teuszi **5**:81*T*
South Africa
  biodiversity and wine initiative (BWI) **2**:300
  biodiversity-friendly farming **2**:300
  botanical gardens **6**:519
  butterflies **1**:658, **1**:659*T*
  Cape Region (South Africa) **5**:213*T*
  Community-Based Natural Resources Management (CBNRM) **2**:297–298
  Convention on Biological Diversity (1930) **2**:313
  deforestation **2**:448*T*
  desert ecosystems **2**:470*T*
  endangered plant species **3**:206–208, **3**:207*F*
  endemism **3**:232*T*, **3**:234*F*
  grasslands and savannas **6**:331–333, **6**:333*F*
  Great Limpopo Transfrontier Park **5**:449
  introduced species
    nonindigenous invertebrates
      arthropods **1**:76*T*, **1**:81
      crop plant pathogens **1**:76*T*, **1**:81–82
      livestock pathogens **1**:76*T*, **1**:82
    nonindigenous plants **1**:76*T*, **1**:81
    nonindigenous vertebrates
      birds **1**:81
      mammals **1**:76*T*, **1**:81
  invasive species
    Lantana camara **4**:352–353
    Pinus pinaster **4**:354–355, **4**:355*F*
    prickly pear (*Opuntia*) **4**:350–351
  maize production **1**:86
  Mediterranean-type climate ecosystems
    characteristics **5**:212–213
    natural disturbances **5**:213–214
    physical characteristics **5**:213*T*, **5**:219
    speciation patterns **5**:214
    species richness **5**:213*T*
    vascular plant diversity **5**:213*T*, **5**:219–220
    vertebrate diversity **5**:213*T*, **5**:220
  megaherbivores **5**:227–228, **5**:227*F*, **5**:234–236
  natural history museums and institutions **5**:406
  nitrogen deposition **5**:526*T*
  phytogeographic/ecoclimatic zones **1**:47–48
  plant invasions **6**:93–95, **6**:94*F*
  seasonal rainfall **2**:465*F*
  South African National Biodiversity Institute (SANBI) **1**:506, **6**:84
  species richness **1**:499*T*
  Succulent Karoo **1:485–490**
    biodiversity
      biodiversity-friendly farming **2**:300
      invertebrates **1**:488
      plants **1**:485, **1**:486*T*, **1**:487, **1**:487*T*
      vertebrates **1**:487*T*
    bioregions **1**:485, **1**:486*F*
    characteristics **1**:485, **1**:486*F*
    climate change threats **1**:489
    land degradation threats **1**:489
    nitrogen deposition **5**:526*T*
    plant growth forms **1**:488–489
    threat management **1**:489
    vegetation structure **1**:488–489, **2**:467–468
  temperate forests
    distribution patterns **7**:116*T*, **7**:117
    diversity patterns **7**:119–121, **7**:119*T*, **7**:120*F*
    dominant floristic features **7**:119
    endemism **7**:121
  termites **4**:432*T*
  threatened plant species **6**:84
South African National Biodiversity Institute (SANBI) **1**:506, **6**:84
South America
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
  butterflies **1**:658*T*
  C_4 plants **2**:371*T*
  Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
  deforestation
    Amazon ecosystems **4**:663, **6**:338–339
    annual forest change **2**:448*T*
    annual forest loss **2**:449*T*
    deforestation rates **1**:679*T*
    economic factors **1**:677*F*
    forest product exports **2**:451*T*
    geographic extent **2**:454*T*
    historical perspective **2**:447–448
    land use change **2**:31*F*
    major threats **6**:608–609
    total forest change **2**:452*T*
  desert ecosystems **2**:469, **2**:591*F*
  ecosystems **3:105–117**
    alpine life zone **1**:149*T*
    background information **3**:105
    biogeography
      early history **3**:109
      late Tertiary **3**:109
      Quaternary period **3**:109–110
    conservation status **3**:171–173*T*
    geographic distribution **3**:111*F*
    high Andean shrublands **3**:111*F*, **3**:114, **3**:115
    mangrove ecosystems **3**:111*F*, **3**:116
    natural ecosystems **6:599–611**
      anthropogenic impacts **6**:608
      aquatic ecosystem threats **6**:609
      charcoal manufacture **6**:609
      climate change **6**:609
      conservation priorities **6**:610
      converted areas **6**:608
      dams **6**:609
      deforestation **6**:608–609
      extractivism **6**:608
      hunting effects **6**:609
      land cover classifications **6**:607*T*
      logging operations **6**:608, **6**:609
      protected area status **6**:610
      sedimentation **6**:610
      terrestrial ecoregion classifications **6**:599, **6**:600–604*T*, **6**:605*F*
      terrestrial ecosystem threats **6**:608–609
      threats **6**:608–609
      tropical forests **6**:599–608, **6**:606*F*, **6**:606*T*
      waterways **6**:609–610
  physical characteristics
    climate **3**:107–109, **3**:108*F*
    geography **3**:105–107
    geology and geomorphology **3**:105–107, **3**:106*F*
    tectonic history **3**:107
  river water composition **3**:562*T*
  Southern Hemisphere **6:612–619**
    angiosperms **6**:617–618
    background information **6**:612
    biotic dispersal **6**:616
    climate change effects **6**:617
    climatic patterns **6**:614
    continental drift **6**:616
    dispersal ability **6**:613–614
    fossil record **6**:615–616
    general area cladograms **6**:614
    geographic distribution patterns **6**:613–614
    high southern latitude plant growth **6**:618
    historical research **6**:612–613
    long-distance dispersal **6**:616–617
    relaxed clock molecular dating analyses **6**:614–615
    research summary **6**:618–619
    vicariance biogeography **6**:616
  temperate forests **3**:111*F*, **3**:117
  temperate grasslands
    distribution and structure **3**:111*F*, **3**:113–114
    functional aspects **3**:114–115
  tropical forests
    distribution and structure **3**:110
    functional aspects **3**:112
    general patterns **3**:111–112
    mountain cloud forests **3**:111
    tropical rain forests **3**:110, **3**:111*F*, **3**:112
    tropical semideciduous/dry forests **3**:110–111, **3**:112–113
  tropical savannas
    distribution and structure **3**:111*F*, **3**:113
    functional aspects **3**:114
  wetlands **3**:116–117
  xeric formations
    distribution and structure **3**:111*F*, **3**:115–116
    functional aspects **3**:116
endangered mammals **3**:194*T*
energy use **3**:252*T*
fire/fire regimes **3**:435–436
fish diversity **3**:472
floodplain agriculture **7**:146
forest product exports **2**:451*T*
freshwater ecosystems **4**:370*T*
grasslands and savannas **6**:331–333, **6**:332*F*

South America (*continued*)
high-elevation Andean ecosystems **4:96–110**
background information **4**:96
geographic distribution **4**:96–97, **4**:97*F*
habitat characteristics **4**:96–97
number of species **4**:102*T*
páramo ecosystem
characteristics **4**:101–103, **7**:275–276
geographic distribution **4**:96–97, **4**:97*F*
northern Andean alpine cushion formations **4**:103
northern Andean dwarf *Polylepis* forest **4**:103
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98
plant diversity **4**:98–99, **4**:99*T*
representative species **4**:102*F*
subpáramo zone **4**:102–103
superpáramo zone **4**:103
true páramo zone **4**:103
physical characteristics and patterns **4**:97–98
plant diversity
endemism **4**:99–100, **4**:99*T*
origins and diversification **4**:100–101
species richness **4**:98–99
puna
central Andean dwarf *Polylepis* forest **4**:105
cushion bogs **4**:105
desert puna **4**:104–105
dry puna **4**:104
geographic distribution **4**:96–97, **4**:97*F*
humid puna **4**:104
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:103–104
plant diversity **4**:98–99
representative species **4**:104*F*
salar vegetation **4**:105
research summary **4**:107
southern Andean steppes
altitudinal zonation **4**:106–107, **4**:106*F*
bogs and seeps **4**:107
cushion plants **4**:106–107, **4**:106*F*
geographic distribution **4**:96–97, **4**:97*F*
latitudinal gradients **4**:106–107
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:105–107
plant diversity **4**:98–99
representative species **4**:105*F*
intercropping **2**:382–383
invasive species **4**:350–351
land-use trends **4**:559*F*
late Quaternary extinction event (LQEE) **4**:658*F*, **4**:723*F*, **4**:726–727, **4**:727*T*
latitudinal gradients **1**:498*F*, **1**:499*T*
logging operations **2**:454*T*
mammalian extinctions **4**:705*T*
mammals **2**:582–583, **2**:583*F*, **4**:701*T*
marsupials **4**:683
mass extinctions **5**:6–7
megaherbivores **5**:224*T*, **5**:225*T*, **5**:232*F*
myriapods **5**:429*T*
native crops **2**:663–664
natural history museums and institutions **5**:405–406
Neartic–Neotropical transition zone **2**:585*F*
neotropical mammals **2:582–594**
background information **2**:582–583
biogeographical–macroecological patterns
areographic patterns **2**:588–589
ecogeographic rules **2**:588–589
latitudinal diversity **2**:588*F*
species richness **2**:588, **2**:589*F*, **2**:591*F*
conservation strategies **2**:591–592
evolutionary history **2**:582–583, **2**:583*F*
future research outlook **2**:592
Great American Biotic Interchange (GABI) **2**:582–583, **2**:584*F*
harvesting threats **2**:590*T*
major threats **2**:589–592, **2**:590*T*
zoogeographic regions
biogeographical limits **2**:583–585, **2**:585*F*
group diversity and classifications **2**:585–586
representative species **2**:586–587, **2**:587*F*
transition zones **2**:583–585, **2**:585*F*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
nonvascular epiphytes **3**:521*T*
oil-palm plantations **5**:602
parasitic infections **5**:671*T*
plant invasions **6**:93
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
temperate forests
distribution patterns **7**:116*T*, **7**:117
diversity patterns **7**:119–121, **7**:119*T*
dominant floristic features **7**:119
endemism **7**:121
threatened birds **7**:181
transboundary protected areas **6**:489*T*
tropical forests **7**:269–270, **7**:270*T*
tropical rain forests
distribution **6**:324*T*
loss and change **6**:324*T*
mammals **2**:591*F*
model-based assessments **6**:324*T*
species richness loss **6**:328*T*
vascular epiphytes **3**:520*T*
South American otter **4**:143, **5**:477
South American palm weevil **5**:603–604
South Asia
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
deforestation **1**:677*F*, **2**:448*T*, **2**:449*T*, **2**:451*T*
forest product exports **2**:451*T*
indigenous peoples **4**:269*T*
nitrogen deposition **5**:526*T*
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
temperate forests
distribution patterns **7**:116*T*, **7**:117
dominant floristic features **7**:118
South Atlantic Gyre **5**:45
South Chinese rain forest **1**:262–263*T*
Southeast Asia **1:509–527**
biodiversity threats
climate change **1**:517
emerging infectious diseases (EIDs) **1**:521
illegal wildlife trade **1**:517–518, **1**:518*F*, **1**:519*F*
industrialized agriculture **1**:517, **1**:519*F*, **1**:520*F*
invasive species **1**:518–520, **1**:521, **1**:522*F*, **4**:346–347
overexploitation **1**:517–518
wildfires **1**:517
botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
C_4 plants **2**:371*T*
challenges and opportunities
biodiversity conservation **1**:520–521
forest restoration **1**:521–522
protected area systems **1**:520
definition **1**:509
deforestation **1**:509–510, **1**:677*F*, **2**:447–448, **2**:448*T*, **2**:449*T*, **4**:562–564, **4**:563*F*
habitat loss and degradation
blackwater streams **1**:512–513
conservation strategies **1**:513
freshwater habitats **1**:510–512, **1**:511*F*
lowland rainforests **1**:509–510
mangrove ecosystems **1**:513, **1**:513*T*
montane forests **1**:510
peat swamp forests **1**:512–513
recovery prospects **1**:513
rivers and streams **1**:510–512, **1**:511*F*, **1**:512*F*
indigenous peoples **4**:269*T*
invasive species **1**:518–520, **1**:521, **1**:522*F*, **4**:346–347
native crops **2**:663
nitrogen deposition **5**:526*T*
oil-palm plantations **1**:540–541, **5**:601–602
polyculture cropping systems **7**:231
protected area systems **6**:491*T*, **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
regional cultivars **3**:32*F*
species status and trends
amphibians **1**:515
birds **1**:516
insects **1**:514–515
mammals **1**:516–517
plants **1**:514, **1**:514*T*
reptiles **1**:515–516
tropical rain forests **6**:314*T*, **6**:316*T*, **6**:317*T*

Southeastern Europe (vegetation surveys) **3**:341–344*T*
Southern Africa
  protected area systems **6**:491*T*
  regional carbon stocks **1**:678*F*
  regional cultivars **3**:32*F*
Southern Africa Biodiversity Strategy **2**:314
Southern Africa Development Corporation (SADC) **3**:643–644
Southern black-and-white ruffed lemurs **6**:256–258*T*
southern boobook **7**:189
southern Bornean orangutan **6**:256–258*T*
Southern California Coastal Wetland Recovery Project **6**:458
Southern Hemisphere ecosystems **6:612–619**
  angiosperms **6**:617–618
  background information **6**:612
  biotic dispersal **6**:616
  climate change effects **6**:617
  continental drift **6**:616
  geographic distribution patterns
    climatic patterns **6**:614
    dispersal ability **6**:613–614
    fossil record **6**:615–616
    general area cladograms **6**:614
    relaxed clock molecular dating analyses **6**:614–615
    research background **6**:613–614
  high southern latitude plant growth **6**:618
  historical research **6**:612–613
  long-distance dispersal **6**:616–617
  research summary **6**:618–619
  vicariance biogeography **6**:616
Southernmost North America **5**:561*F*, **5**:563–564
southern muriqui **6**:256–258*T*
Southern Oscillation (SO) **5**:720–721
southern pocket gopher **7**:229*T*
Southern right whale **5**:81*T*
southern springbok **5**:364–365
Southern Subtropical Gyres **4**:674, **4**:675*F*
southern white-cheeked crested gibbon **6**:256–258*T*
South Georgia Island **1**:180–181, **1**:186*T*
South Island saddleback **5**:366–367
South Korea
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  ecosystems
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  see also Korea
South Pacific Subtropical Gyre (SPSG) **5**:45
southwest Asia
  botanical gardens **6**:505–506, **6**:506*T*, **6**:507*F*, **6**:510*T*
  native crops **2**:662–663
Southwestern North America **5**:561*F*, **5**:563
southwestern willow flycatchers **3**:260
sow bugs **3**:182*T*, **4**:371*T*
soybeans **1**:116*T*, **1**:132*T*, **1**:530, **1**:543–547, **2**:383, **2**:384*F*, **2**:663, **3**:32*F*
soy sauce **3**:638–639
spadefish **3**:467–471
spadefoot toad **2**:473, **5**:473
Spain
  botanical gardens **6**:519
  flora **5**:481*T*
  intercropping systems **2**:382–383
  invasive species **4**:345–346
  vegetation maps **3**:346*T*
  vegetation surveys **3**:341–344*T*
  wheat production **1**:87–88
Spalax ehrenhergi **3**:675–676
Spallanzani, Lazzaro **5**:629–630
Spanish imperial eagle **7**:185, **7**:262–263
Spanish Needle **3**:704–705, **3**:704*F*
sparrowhawks **5**:474
sparrows **1**:77–78, **1**:81, **1**:82, **1**:83, **4**:365–366, **4**:490*F*
Spartina alterniflora
  introduced species **4**:349–350, **4**:360, **4**:365–366, **6**:93–95
  restoration programs **6**:454, **6**:458, **7**:375–378
  root–soil interface **6**:115*T*
  speciation **6**:627, **6**:627*F*
  wetland ecosystems **7**:375–378, **7**:388–389
Spartina anglica **4**:350, **4**:360, **4**:365–366, **6**:627, **6**:627*F*, **7**:375–378
Spartina brasiliensis **1**:164
Spartina foliosa **4**:350, **6**:454–455, **6**:456–457
Spartina maritima **4**:350, **4**:360, **4**:365–366, **6**:627, **6**:627*F*
Spartina patens **7**:388–389
Spartina townsendii **4**:350, **6**:627, **6**:627*F*, **7**:375–378
Spathebothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Spathebothrium **7**:456–457*F*
spatial autocorrelation (SAC) **6**:693
Spatially Explicit Species Assemblage Models (SESAMs) **1**:567
spatial patterns
  ecological perspectives **3**:5–6
  landscape genetics **4:508–523**
    analytical frameworks
      assignment tests **4**:511
      barrier detection **4**:511, **4**:512*F*
      graph theory/network analyses **4**:512–514, **4**:513*F*
      Mantel tests **4**:510–511
      ordination and regression analysis **4**:514
      resistance surfaces **4**:511–512, **4**:513*F*
    isolation-by-distance (IBD) models **4**:508, **4**:509*F*
    research background **4**:508–509
    research challenges
      genetic distance (F_{ST}) measures **4**:518–519
      interdisciplinary communication **4**:517–518
      landscape genomics **4**:519–520
      multispecies comparative studies **4**:518
      spatial analysis method (SAM) **4**:519
      spatial replication **4**:518
      spatial scale effects **4**:518
    research results
      anthropogenic impacts **4**:516–517
      barrier identification **4**:515–516
      gene flow barriers **4**:515–516
      general research trends **4**:514–515
      landscape variables influences **4**:516
    sampling strategy and design **4**:509–510
spatio-temporal mosaic theory **4**:388–389, **4**:389*F*
spearfish **3**:470
spear-winged flies **3**:489–491*T*
speciation **6:620–628**, **6:629–639**
  basic concepts
    accepted definitions **6**:629–630
    commonly used concepts
      biological species concept **6**:630–631
      cohesion species concept **6**:631
      evolutionary species concept **6**:631
      genealogical species concept **6**:631
      general discussion **6**:630
      genotypic species clusters **6**:630
      morphological species concept **6**:630
      phenetic species concept **6**:630
      phylogenetic species concept **6**:631
      recognition species concept **6**:630–631
    processes and characteristics **6**:620, **6**:621*F*
  biological species concept **6**:620
  definition **6**:629
  dispersal biogeography **2**:558
  diversification **2**:47
  ecotones
    conservation implications **3**:146
    evolutionary development **3**:144–145
    human-generated ecotones **3**:143*F*, **3**:146–147
    species range/transitional areas **3**:145–146
  endemism **3**:235
  evolutionary mechanisms **3**:399
  evolutionary processes
    general characteristics **6**:631–632
    genetic drift **6**:632
    mutations **6**:633
    natural selection **6**:631–632
    sexual selection **6**:632–633
  future outlook **6**:638
  future research areas
    developmental biology studies **6**:638
    genealogical models **6**:637–638
  genetic-based species concept **6**:620–621
  genetic patterns and processes
    postzygotic reproductive isolation **6**:636–637

speciation (*continued*)
research benefits **6**:636–637
species differences **6**:636–637
within-species versus between-species variation **6**:637
geographical modes
characteristics **6**:633–634
classical allopatric speciation **6**:634
founder effect speciation **6**:634–635
hybridization **6**:635
parapatric speciation **6**:635
peripatric speciation **6**:634–635
sympatric speciation **6**:635–636
keystone species
biomonitoring processes **3**:125
ecological perspectives **3**:5
multispecies interactions **6**:722, **6**:723*T*
myriapods **5**:428
neutral theory **5**:511
reproductive barriers **6**:620–621, **6**:622*T*
research summary **6**:627–628
speciation–extinction rate relationship **3**:415, **5**:439–440, **5**:440*F*
speciation modes
allopatric speciation
basic concepts **6**:621–624, **6**:621*F*
dichopatric speciation **6**:621*F*, **6**:622–624, **6**:622*T*, **6**:623*F*
general characteristics **6**:621–624, **6**:622*T*
peripatric speciation **6**:621*F*, **6**:622*T*, **6**:624
salamanders **6**:622–624, **6**:623*F*
general characteristics **6**:621–624
neutral theory **5**:515
nonallopatric speciation
allopolyploid/amphiploid speciation **6**:627, **6**:627*F*
autopolyploid speciation **6**:626
basic concepts **6**:621*F*, **6**:624–625
cichlids **6**:625
direct allogenous speciation **6**:627
general characteristics **6**:621–624, **6**:622*T*
interspecific hybridization **6**:626
introgressive hybrid speciation **6**:626
Mimulus cupriphilis **6**:625
parapatric speciation **6**:621*F*, **6**:622*T*, **6**:625
recombinational hybrid speciation **6**:626
reticulate allogenous speciation **6**:627–628
Rhagoletis fruit flies **6**:625–626
spontaneous thelytokous speciation **6**:626
sympatric speciation **6**:621*F*, **6**:622*T*, **6**:625
subspecies taxonomy **7:45–48**
historical background
binominal versus trinominal nomenclature **7**:45–46
divergence theories **7**:46
levels of variation **7**:45
theoretical challenges **6**:620
underlying assumptions **6**:629
species abundance distributions (SADs) **5**:510–511, **5**:513–514, **5**:513*F*, **6**:655, **6**:656*F*
species–area relationships (SARs) **6:640–650**
biodiversity conservation
background information **6**:646–648
incidence functions **6**:649–650
loss management strategies **6**:648–649
nestedness measures **6**:649–650
relaxation concept **6**:647
species loss **6**:643*F*, **6**:646–648, **6**:647*F*
summary discussion **6**:650
biodiversity dynamics modeling **5:321–328**
background information **5**:321–322
countryside habitats **5**:323–324, **5**:323*F*
native habitats **5**:322–323, **5**:322*F*
biogeographic research **1**:583–585
birds **6**:659–660, **6**:660*F*, **6**:662*F*
diversity patterns **6**:710
functional form
curve shapes **6**:643–644, **6**:643*F*
extreme value model **6**:644–645
power-function model **6**:644
random placement model **6**:644–645
self-similarity model **6**:644
neutral theory **5**:510–511, **5**:514
parameter interpretation
alternative interpretations **6**:646
canonical slope values **6**:645
island–mainland differences model **6**:645–646
isolation effects **6**:646
slope parameter **6**:645
summary discussion **6**:646
plant biodiversity–climate relationship **2**:115, **3**:116, **3**:117
research background **6**:640
sampling practice
basic concepts **6**:641–642
independent sample area **6**:642, **6**:642*F*
nested sample area **6**:642, **6**:642*F*
species assemblages **6**:659–660, **6**:660*F*, **6**:662*F*
statistical practice **6**:642–643
underlying mechanisms
area per se hypothesis **6**:640–641
edge effects **6**:641
habitat diversity hypothesis **6**:640
multiple causes **6**:641
passive sampling hypothesis **6**:641
research background **6**:640
resource concentration hypothesis **6**:641
species richness **6**:640, **6**:640*F*
species assemblages **6:651–666**
assemblage properties
general discussion **6**:663–664
historical variability range **6**:665
indicator species **6**:663–664, **6**:664*F*
knowledge gaps **6**:664–665
definition **6**:651
descriptive processes
abundance lists **6**:652*T*, **6**:653*T*
beta diversity
basic concepts **6**:659–660
biogeographic scales **6**:660*F*, **6**:661, **6**:662*F*
characterizations **6**:659–660, **6**:660*F*
habitat and dispersal limitation scales **6**:660*F*, **6**:661
sampling scale **6**:660–661, **6**:660*F*
temporal beta diversity **6**:662, **6**:662*F*
characteristics **6**:652
correlation/covariation measures
abundance versus mass **6**:657*F*, **6**:658
abundance versus occupancy/range size **6**:657*F*, **6**:658
abundance versus traits **6**:658
energy versus mass **6**:656–658, **6**:657*F*
general discussion **6**:655–658
traits versus traits **6**:658–659
mean and univariate descriptions **6**:652–653, **6**:652*T*, **6**:653*T*
spatiotemporal variations
climatic forcing scale **6**:663
evolutionary scale **6**:663
indicator species **6**:664*F*
population dynamics scale **6**:663
sampling and detection errors **6**:662–663
species–time relationships **6**:661–662, **6**:662*F*
temporal beta diversity **6**:662, **6**:662*F*
variation measures
basic principles **6**:653
body mass distribution **6**:654–655, **6**:654*F*
lognormal distributions **6**:653–654, **6**:654*F*, **6**:656*F*, **6**:657*F*
range size and occupancy distribution **6**:655, **6**:656*F*, **6**:657*F*
species abundance distributions (SADs) **6**:655, **6**:656*F*
species- versus individual-level distributions **6**:652*T*, **6**:653, **6**:653*T*, **6**:654*F*
traits and trait distribution **6**:655, **6**:656*F*
extinction risks **7**:192–194, **7**:193
process drivers **6**:651
research summary **6**:665
spatial scale **6**:651–652
Species at Risk Act **2**:225–226, **7**:197–199
species coexistence **6:667–678**
background information **6**:667
coexistence mechanisms
basic concepts **6**:671
landscape heterogeneity
exclusion mechanisms **6**:676
spillover effects **6**:676
local coexistence models
food web effects **6**:674
local habitat heterogeneity **6**:674–675
mechanistic models **6**:673–674, **6**:673*F*
nonlinear consumption **6**:675

storage effect **6**:675
temporal niche partitioning **6**:675
open communities
autecology **6**:675
colonization–extinction dynamics **6**:675–676
landscape heterogeneity **6**:676
population size effects **6**:675
constraints
community contexts **6**:669
competition **6**:668–669
competitive exclusion principle
apparent competition (*P*– rule) **6**:671
basic concepts **6**:670–671
exploitative competition (*R*– rule) **6**:671
general discussion **6**:668–669
natural enemies **6**:669
spatial contexts **6**:669
experimental methods
competition
beetles **6**:670
interference competition **6**:669–670
Protozoa **6**:669–670
research results **6**:670
general discussion **6**:669–670
predation **6**:670
future outlook
Allee effect **6**:676–677
niche construction **6**:676–677
speciation modes **6**:677
transient coexistence **6**:676
general considerations
permanence **6**:668
scale **6**:667
species interactions **6**:667–668
research summary **6**:677
traditional models
classical niche partitioning **6**:671–672
limiting similarity model **6**:672
Lotka–Volterra model **6**:672
manipulative field experiments **6**:672–673
species concepts **6:679–691**
alternative species concepts
autapomorphic species concept **7**:93*T*
biological species concept **7**:93*T*
ecological species concept **6**:682, **7**:93*T*
Hennigian species concept **7**:93*T*
phylogenetic species concept **7**:93*T*
recognition species concept **6**:682–683
sexual selection **6**:682
systematic classifications **7**:93, **7**:93*T*
basic concepts **6**:680
biases **6**:680
biodiversity conservation
ecological importance **6**:689
genetic differentiation **6**:689
spatiotemporal considerations **6**:689–690
species as real entities **6**:689
combined species concepts
cohesion species concept **6**:685
evolutionary and lineage concepts **6**:685
Darwinian theory
morphological species criterion **6**:681
polytypic species criterion **6**:681
dissenting perspectives
evolutionary unit approach **6**:686
general discussion **6**:685–686
genotypic cluster criterion and assignment tests **6**:686–688, **6**:687*F*
phenetic species concept **6**:686
race versus species theory **6**:688
taxonomic practice **6**:685–686
historical perspective
Bayesian species delimitation **6**:684
diagnostic/phylogenetic species concept **6**:683–684, **7**:91–93, **7**:92*F*
genealogy **6**:684
monophyly **6**:683, **7**:91–93, **7**:92*F*
interbreeding theory **6**:681–682
species distribution modeling (SDM) **6:692–705**
basic concepts
geographical models **6**:694–695
niche theory **6**:693–694, **6**:694*F*
spatial autocorrelation (SAC) **6**:693
species response shapes **6**:692*F*, **6**:694
characteristics and functional role **6**:692–693, **6**:692*F*
community studies **6**:702–703
Cupressus chimera case study **6**:695
data collection methodology
biotic interactions and disturbances **6**:698–699
challenges **6**:695–697
environmental data
aquatic systems **6**:700
challenges **6**:699
climate **6**:699
land cover **6**:699–700
multicollinearity **6**:700
soils and substrate **6**:699–700
terrain **6**:699
species data
observational research **6**:696–697
presence-only data **6**:697–698, **6**:701–702
record characteristics **6**:697, **6**:698*F*
sample characteristics **6**:696–697, **6**:696*F*
site-occupancy models **6**:697, **6**:698*F*
greenhouse effect impacts **4**:24–25, **4**:24*T*
modeling approaches
Bayesian methods **6**:702
decision trees **6**:701
general discussion **6**:700–702
generalized additive models (GAMs) **6**:701
generalized linear models (GLMs) **6**:701
genetic algorithms **6**:702
machine learning methods **6**:701, **6**:702*F*
maximum entropy theory (MaxEnt) **6**:702
profile and envelope methods **6**:701–702
new times and places predictions **6**:703
research summary **6**:703–704
species richness **6**:702–703
species diversity **2:434–446**, **6:706–714**
air pollution impacts
nitrogen deposition **1**:142–144, **1**:143*F*, **1**:146*F*
ozone (O_3) **1**:144
surface water acidification **1**:144–145
background information **6**:706
basic concepts **1**:401–402
biodiversity policies **6**:713
calculation methods **2**:435–437, **2**:435*F*
classification systems **6**:713
coevolution effects **2**:147
countryside biogeography **2**:351
definition **2**:434–435
diversity indices
body size diversity **6**:712, **6**:712*F*
characteristics **6**:711–713
food web position **6**:712, **6**:713*F*
Gini–Simpson index **2**:438–439
internal variability **6**:713
Shannon index **2**:438
Simpson index **2**:438–439
species abundance curves **6**:711–713, **6**:711*F*
taxonomic diversity **6**:712–713
edible plants
abundance rankings **3**:129*T*
edible parts **3**:129–130
food usage **3**:130*T*
prolific species **3**:130*T*
research background **3**:128–129
taxonomic diversity **3**:128–129, **3**:129*T*
fish species
depletion causes and consequences **3**:451–452
extinction stages **3**:452
extinction vulnerability **3**:451–452
human health–biodiversity relationship **1**:357–358
landscape composition **4**:477*F*, **4**:478–479
Lepidoptera **5**:387–388*T*, **5**:389–390, **5**:389*T*
mammals **4**:682*F*, **6**:712*F*
marine ecosystems **5**:49–50, **5**:50*T*
marsupials **4**:687*T*
measurement and analysis **1**:402–403, **2**:650, **6**:534, **6**:534
measurement and estimation methods **5:195–211**
biases **5**:197
biotic similarity
abundance-based similarity indices **5**:208–209, **5**:209*F*
basic concepts **5**:197
incidence-based similarity indices **5**:207–208
phylogenetic similarity indices **5**:210
total abundance–shared species-based similarity indices **5**:209–210
characteristics **5**:196, **5**:196*F*
functional diversity **5**:196–197
phylogenetic diversity **5**:196–197, **5**:197*F*
sampling data
characteristics **5**:197–198

species diversity (*continued*)
individual-based (abundance) data **5**:198
sample-based (incidence) data **5**:198
species diversity estimation methods
cladistic diversity **5**:205
Gini–Simpson index **5**:202–203
Hill numbers **5**:203, **5**:204*F*
metrics **5**:202–203
phylogenetic entropy **5**:205
Rao's quadratic entropy **3**:603, **3**:603*F*, **5**:205
Shannon entropy **5**:202–203
taxonomic/phylogenetic diversity **5**:204–207, **5**:205*F*, **5**:206*F*
species richness estimation methods
abundance-based coverage estimator (ACE) **5**:201
Chao estimator **5**:200–201
extrapolation estimators **5**:201–202, **5**:202*F*
incidence-based coverage estimator (ICE) **5**:201
nonparametric asymptotic richness estimators **5**:200–201
rarefaction measurement approach **5**:199–200, **5**:199*F*
species accumulation curve **5**:198–200, **5**:198*F*
taxonomic diversity **5**:196–197
nematodes
Ctenomyidae rodents **7**:432–433, **7**:432*T*
Geomyidae rodents **7**:432, **7**:432*T*
number of species **7**:431
vertebrate parasitic nematodes **7**:431–432
nitrogen dynamics **5**:544–546, **5**:545*F*
oil-palm plantations **1**:517, **5**:604–605
partitioning methods
additive partitioning
absolute species turnover **2**:442–443
alpha + beta **2**:442–443
proportional species turnover **2**:444
Whittaker's species turnover **2**:443–444
alpha, beta, and gamma diversity
additive partitioning **2**:442–443
basic concepts **2**:439–440
beta richness and diversity **2**:442
diversity components **2**:440–442, **2**:441*F*
richness components **2**:440, **2**:441*F*
multiplicative partitioning
alpha x beta **2**:439–440
gamma diversity **2**:439–440
richness x evenness **2**:439
pest control **2**:392
placentals **4**:690*T*
plant biodiversity–climate relationship **2:115–127**
annual mean temperature **2**:119–120, **3**:119
area–productivity relationship **2**:117–118, **3**:118
disturbances **2**:115–117, **2**:123, **3**:118
experimental research **2**:126
general discussion **2**:115
genomic research **2**:126
global-scale distribution **2**:118–119
latitudinal gradients **2**:118–119, **3**:119
net primary production (NPP) **2**:117–118, **2**:120, **3**:118, **3**:121
observational research **2**:121–123, **3**:122
precipitation **2**:120, **3**:120
species–area relationships (SARs) **2**:115, **3**:116, **3**:117
summary discussion **2**:126
theoretical perspectives
biogeographic history **2**:125–126
competition **2**:123
disturbances **2**:123
energy availability **2**:123–124
extreme events **2**:124–125, **3**:124, **3**:125
general discussion **2**:120–123
observational research **2**:121–123, **3**:122
Protozoa **6**:289–293, **6**:289*T*
quantification methods **2**:437–438
reptiles **6**:400–401
research scope **6**:706–707
resemblance indices **2**:444–445
species packing **4**:646
species richness
data collection and analysis **6**:707–710, **6**:708*F*, **6**:709*F*
determinant factors **6**:710–711
diversity indices
body size diversity **6**:712, **6**:712*F*
characteristics **6**:711–713
food web position **6**:712, **6**:713*F*
internal variability **6**:713
species abundance curves **6**:711–713, **6**:711*F*
taxonomic diversity **6**:712–713
diversity patterns **6**:708*F*, **6**:710
species–area relationships (SARs) **6**:710
species–energy theory **2**:596–597, **2**:597*T*
species interactions **6:715–725**
above-ground–below-ground species interactions **2**:386
above-ground species interactions
canopy heat capture **2**:387, **2**:387*F*
characteristics **2**:386–387
light interception/light use efficiency **2**:386–387
shading effects **2**:387, **2**:387*F*
spatiotemporal complementarity **2**:387–388
air pollution impacts **1**:139, **1**:140*F*
background information **6**:715
basic concepts **3**:3–4, **3**:4*F*
below-ground species interactions
hydraulic lift effect **2**:391
iron/zinc (Fe/Zn) nutrition **2**:390
legume-based crop mixtures **2**:389–390
nitrogen complementary utilization **2**:389–390
phosphorus facilitation **2**:390
root spatial distribution **2**:388–389
spatiotemporal complementarity **2**:388–389
vesicular arbuscular mycorrhizae (VAM) **2**:391
water capture/water use **2**:390–391
water use efficiency **2**:391
climate change effects **2**:76–77
desert ecosystems **2**:474–475, **2**:475*F*
energy flow **3**:247
evolutionary change
character displacement **6**:720
competition **6**:720
multispecies interactions **6**:723
mutualism **6**:720–721
natural selection **6**:719–720
predation **6**:721
freshwater invertebrates **4**:375
harmful algal blooms **1**:478–479
human impact **6**:723–724
interaction strengths **4**:454–456, **4**:454*F*
multispecies interactions
challenges **6**:721–722
evolutionary change **6**:723
food webs **6**:721–722, **6**:721*F*, **6**:722*F*
indirect interactions **6**:722–723
keystone species **6**:722, **6**:723*T*
metacommunities **6**:723
trophic cascades **6**:722–723
trophic levels **6**:721–722, **6**:721*F*, **6**:722*F*
pairwise interactions
asymmetrical interactions **6**:715*T*, **6**:719
characteristics **6**:715–717
commensalism **6**:715*T*, **6**:719
competition **6**:715*T*, **6**:716–717
herbivory **6**:718–719
mutualism **6**:715*T*, **6**:719
parasitism **6**:715*T*, **6**:718–719
predation **6**:715*T*, **6**:717–718
symbiotic relationships **6**:719
plant biodiversity **6**:61–62
research summary **6**:724
simple two-species interactions
basic concepts **6**:178–179
competition **6**:180
disease dynamics **6**:179–180
host–parasitoid dynamics **6**:179
prey–predator interactions **6**:179
stability dynamics **2**:653–654
species packing **4:642–648**
basic concepts
ecological niche model **4**:643
general discussion **4**:644–645
laboratory experiments **4**:643–644
Lotka–Volterra model **4**:643
MacArthur–Levins model **4**:645
multiple species model **4**:645
niche separation–anatomical difference relationship **4**:645–646
biological metaphor **4**:642
research summary **4**:646–647
species diversity **4**:646
species richness
African savannas **1**:70, **1**:71*F*
agricultural intensification effects **1**:374–375
air pollution impacts

nitrogen deposition **1**:142–144, **1**:143*F*, **1**:146*F*
ozone (O_3) **1**:144
surface water acidification **1**:144–145
Amazon ecosystems **1**:247–249, **1**:248*F*
amphibians **1**:170*T*, **6**:709*F*
animal biodiversity **7**:389–391, **7**:389*F*, **7**:390*F*, **7**:393–394, **7**:396*F*
assessment measures **6**:60–61
Australia
dryland vegetation continua **1**:302, **1**:303*F*
invasive species **1**:303
invertebrates **1**:300–302, **1**:303*T*
isolated rainforest stands **1**:303–304
overstory and understory strata **1**:300, **1**:301*F*, **1**:302*F*
postfire succession **1**:302–303, **1**:304*F*
structural characteristics **1**:301*T*
vertebrates **1**:300–302, **1**:302*F*
wetland vegetation continua **1**:302
biodiversity-rich countries **1**:497
biogeographical–macroecological patterns **2**:588, **2**:589*F*, **2**:591*F*
biogeographical models **1:565–575**
characteristics and functional role **1**:565–566, **1**:566*F*
correlative models **1**:566–569, **1**:566*F*
future research areas **1**:572–574
stochastic simulation models **1**:569–572, **1**:572*F*
birds **6**:709*F*
Central America **2**:25, **2**:25*T*
conservation and restoration **6**:442
countryside biogeography **2**:351
data collection and analysis **6**:707–710, **6**:708*F*, **6**:709*F*
determinant factors **6**:710–711
diversity estimations **2**:650, **2**:651*T*
diversity indices
body size diversity **6**:712, **6**:712*F*
characteristics **6**:711–713
food web position **6**:712, **6**:713*F*
internal variability **6**:713
species abundance curves **6**:711–713, **6**:711*F*
taxonomic diversity **6**:712–713
diversity patterns **6**:708*F*, **6**:710
endemism **3**:232
fossil record **3**:678–679
freshwater ecosystems
freshwaters **3**:571–572, **3**:571*T*
freshwater sediments **3**:571–572, **3**:572*T*
site characteristics and ranges **3**:572*T*
global species richness **3:707–711**
current estimates **3**:709, **3**:710*T*
estimation methods
canvassing experts **3**:708
counting **3**:707–708
species–abundance distributions **3**:709
species–body size distributions **3**:709
temporal patterning **3**:708
trophic relations **3**:709
undescribed species **3**:708
well-studied areas extrapolation estimation **3**:708–709
well-studied groups extrapolation estimation **3**:709
future outlook **3**:711
importance **3**:707
sampling considerations **3**:707
uncertainties
arthropods **3**:711
bacteria **3**:710–711
coral reef ecosystems **3**:710
deep ocean environments **3**:710
fungi **3**:711
general discussion **3**:709–710
nematodes **3**:711
Protozoa **3**:711
soils **3**:710
tropical forest canopies **3**:709–710
viruses **3**:710
habitat loss and fragmentation **4**:54
herbaceous vegetation **4:79–86**
control mechanisms
intermediate disturbance hypothesis **4**:79
intraspecific diversity **4**:79–80, **4**:80*F*
dominants and subordinates **4**:82–83, **4**:82*F*, **4**:84
general discussion **4**:79
high species richness effects **4**:81–82
mass ratio hypothesis **4**:80–81, **4**:81*F*
plant traits **4**:80
research summary **4**:84–85
species declines
filter effects **4**:84
founder effects **4**:84
transients **4**:83–84, **4**:83*T*, **4**:84
high-elevation Andean ecosystems **4**:98–99
lake and pond ecosystems
species richness–lake size relationship **4**:459–460
species richness–morphometry relationship **4**:460
land-use change impacts **1**:541–542
latitudinal gradients
biodiversity-rich countries **1**:499*T*
climatic regions **4**:614*F*
data assessment and synthesis **4**:624–625, **4**:624*F*
ecological communities **4**:615–618, **4**:615*F*, **4**:616*F*, **4**:617*F*
fossil species **4**:614*F*
general discussion **4**:612
geographic distribution **4**:613*F*
gradient patterns
assemblage characteristics **4**:614–615
assessment considerations **4**:614–615
ecological communities **4**:615–618, **4**:616*F*, **4**:617*F*
hierarchical configuration **4**:618
multiple dimensions **4**:618–620, **4**:619*F*
historical perspective **4**:612–614
influencing factors **4**:617*T*
mammals **4**:613*F*, **4**:615*F*, **4**:700–702
mechanism hypotheses
contributing factors **4**:620–621
evolutionary speed **4**:621–622
geographic area hypothesis **4**:620–621, **4**:621*F*
geometric constraints hypothesis **4**:622–624, **4**:623*F*
Rapoport–rescue hypothesis **4**:622
parasites **5**:668, **5**:671*F*
plant biodiversity–climate relationship **2**:118–119, **3**:119
taxonomic dimensions **4**:616*F*
local controls
diversity–productivity relationship **2**:597–599, **2**:598*F*
prediction limitations **2**:599
research summary **2**:606
spatial scale **2**:604–605, **2**:605*F*
species–energy theory **2**:596–597, **2**:597*T*
theoretical perspectives **2**:605
Volterra–Gause perspective **2**:595–596
mammals **6**:709*F*
marine ecosystems **5**:49–50
marine protected areas **5**:97*T*
measurement and estimation methods **5:195–211**
biases **5**:197
biotic similarity
abundance-based similarity indices **5**:208–209, **5**:209*F*
basic concepts **5**:197
incidence-based similarity indices **5**:207–208
phylogenetic similarity indices **5**:210
total abundance–shared species-based similarity indices **5**:209–210
characteristics **5**:196, **5**:196*F*
functional diversity **5**:196–197
phylogenetic diversity **5**:196–197, **5**:197*F*
sampling data
characteristics **5**:197–198
individual-based (abundance) data **5**:198
sample-based (incidence) data **5**:198
species diversity estimation methods
cladistic diversity **5**:205
Gini–Simpson index **5**:202–203
Hill numbers **5**:203, **5**:204*F*
metrics **5**:202–203
phylogenetic entropy **5**:205
Rao's quadratic entropy **3**:603, **3**:603*F*, **5**:205
Shannon entropy **5**:202–203
taxonomic/phylogenetic diversity **5**:204–207, **5**:205*F*, **5**:206*F*
species richness estimation methods
abundance-based coverage estimator (ACE) **5**:201
Chao estimator **5**:200–201
extrapolation estimators **5**:201–202, **5**:202*F*
incidence-based coverage estimator (ICE) **5**:201

species richness (*continued*)
nonparametric asymptotic richness estimators **5**:200–201
rarefaction measurement approach **5**:199–200, **5**:199*F*
species accumulation curve **5**:198–200, **5**:198*F*
taxonomic diversity **5**:196–197
Mediterranean-type climate ecosystems **5**:213*T*
Near East ecosystems **5**:480–481, **5**:481*T*
nitrogen dynamics **5**:544–546, **5**:545*F*
planktonic communities **6**:34–35
plant biodiversity **6**:60–61, **7**:388–389, **7**:388*F*, **7**:391–393, **7**:398*F*
plant biodiversity–climate relationship **2:115–127**
annual mean temperature **2**:119–120, **3**:119
area–productivity relationship **2**:117–118, **3**:118
disturbances **2**:115–117, **2**:123, **3**:118
experimental research **2**:126
general discussion **2**:115
genomic research **2**:126
global-scale distribution **2**:118–119
latitudinal gradients **2**:118–119, **3**:119
net primary production (NPP) **2**:117–118, **2**:120, **3**:118, **3**:121
observational research **2**:121–123, **3**:122
precipitation **2**:120, **3**:120
species–area relationships (SARs) **2**:115, **3**:116, **3**:117
summary discussion **2**:126
theoretical perspectives
biogeographic history **2**:125–126
competition **2**:123
disturbances **2**:123
energy availability **2**:123–124
extreme events **2**:124–125, **3**:124, **3**:125
general discussion **2**:120–123
observational research **2**:121–123, **3**:122
plant hybridization **4**:173–174, **4**:173*F*
productivity **3**:9–11, **3**:10*F*
Protozoa **6**:289–293, **6**:289*T*
reforestation **6**:373–374, **6**:374*F*
regional controls
historical biogeographical factors **2**:599–601, **2**:600*F*
metacommunity processes model **2**:601–602, **2**:602*F*
phylogenetic patterns **2**:599–601, **2**:600*F*
proportional sampling **2**:602*F*, **2**:603
research summary **2**:606
spatial scale **2**:604–605, **2**:605*F*
theoretical perspectives **2**:605
remote sensing applications **6**:398, **6**:398*F*, **6**:398*T*
reptiles **6**:708*F*, **6**:709*F*
return-on-investment (ROI) analysis **4**:187*T*
riparian zones **6**:461–463, **6**:463*F*
sandy beaches
alongshore distribution **2**:131–133
causal factors **2**:133–134
future research areas **2**:135
macroscale patterns
latitudinal trends **2**:131, **2**:131*F*
morphodynamic effects **2**:131, **2**:132*F*, **2**:133*F*
mesoscale patterns **2**:131–133
microscale patterns **2**:133
research summary **2**:135
threats and challenges **2**:134–135
zonation patterns **2**:131–133
saturation tests
caveats **2**:603
experimental evidence **2**:603–604, **2**:604*F*
methodologies **2**:602–603
test results **2**:603
unsaturated versus saturated patterns **2**:603
seasonally dry tropical forests **5**:488–490
spatial scale **2**:604–605, **2**:605*F*
species–area relationships (SARs) **6:640–650**, **6**:710
biodiversity conservation
background information **6**:646–648
incidence functions **6**:649–650
loss management strategies **6**:648–649
nestedness measures **6**:649–650
relaxation concept **6**:647
species loss **6**:643*F*, **6**:646–648, **6**:647*F*
summary discussion **6**:650
diversity patterns **6**:710
functional form
curve shapes **6**:643–644, **6**:643*F*
extreme value model **6**:644–645
power-function model **6**:644
random placement model **6**:644–645
self-similarity model **6**:644
parameter interpretation
alternative interpretations **6**:646
canonical slope values **6**:645
island–mainland differences model **6**:645–646
isolation effects **6**:646
slope parameter **6**:645
summary discussion **6**:646
plant biodiversity–climate relationship **2**:115, **3**:116, **3**:117
research background **6**:640
sampling practice
basic concepts **6**:641–642
independent sample area **6**:642, **6**:642*F*
nested sample area **6**:642, **6**:642*F*
statistical practice **6**:642–643
underlying mechanisms
area per se hypothesis **6**:640–641
edge effects **6**:641
habitat diversity hypothesis **6**:640
multiple causes **6**:641
passive sampling hypothesis **6**:641
resource concentration hypothesis **6**:641
vascular plants **6**:640*F*
species distribution modeling (SDM) **6**:702–703
species richness–functional diversity measures **3**:605–606, **3**:606*F*, **3**:607*F*
theoretical perspectives
comparative models **2**:606, **2**:606*F*
design manipulations **2**:605–606
multiscale perspectives **2**:605
tropical rain forests **6**:317*T*, **6**:327–328, **6**:328*T*
Type I/Type II relationships **2**:603
wetland ecosystems
animal species **7**:389–391, **7**:389*F*, **7**:390*F*, **7**:393–394, **7**:396*F*
plant communities **7**:388–389, **7**:388*F*, **7**:391–393, **7**:398*F*
Species Survival Commissions (SSC) **3**:195–196
Species Survival Plan (SSP) **7**:475–477
speckled alder **7**:395–397
spectacled bear **7**:229*T*
Speculitermes **4**:409*T*
Spelaeogriphacea **2**:416–417*T*
Speleiacris tabulae **3**:221
Spencer, Herbert **2**:422, **4**:652
Speonomus colluvii **7**:60
Speoplatyrhinus poulsoni **7**:55*F*
Spermolepis echinata **2**:373*T*
Spermophilus lateralis **7**:36
Spermophylus parryi **1**:631–632
sperm whales **3**:192–193, **3**:256, **5**:65*T*, **5**:66–67, **5**:88–89, **5**:314, **6**:578–579
Speyeria zerene **4**:564
Sphaeradenia **3**:522*T*
Sphaeridae **4**:464*T*
Sphaeroceridae **3**:489–491*T*, **3**:493
Sphaeroceroidea **3**:489–491*T*
Sphaerocystis **6**:30
Sphaerodoridae **7**:417*F*
Sphaeroeca **6**:293
Sphaeroma **5**:14–15
Sphaerotermes **4**:409*T*
Sphaerotholus **2**:519*T*
Sphagnum **1**:136, **1**:277–278, **4**:577–579, **7**:396–397, **7**:396*F*
Sphalloplana holsingeri **3**:185*F*
Spheniscidae **1**:613–614
Sphenisciformes **7**:181–182
Sphenoclea zeylanica **1**:164
Sphenodon **2**:47, **7**:340–341
Sphenodon guentheri **5**:367
Sphenodon punctatus **5**:367, **6**:8, **6**:401
Sphenops sepsoides **5**:467–469, **5**:472
Sphigurus mexicanus **2**:587*F*
Sphincterochila cariosa **5**:465–467, **5**:469–470
Sphincterochila fimbriata **5**:465–467, **5**:466–467
Sphincterochila prophetarum **5**:466
Sphincterochila zonata **5**:466–467
Sphingidae **5**:387–388*T*, **5**:392*F*, **5**:401–402
Sphingobacterium comitans **1**:312–313
Sphyraena barracuda **3**:260
Sphyrospermum **3**:522*T*

spices
Europe **2**:662–663
food selection **1**:42
Mediterranean Basin **2**:662–663
North and South America **2**:663–664
plant defense system uses **2**:432–433
Southeast Asia **2**:663
southwest Asia **2**:662–663
Sub-Saharan Africa **2**:663
Spicotermes **4**:409*T*
spider crickets **3**:730
spider mites **6**:587–588
spider monkeys **2**:590*T*, **6**:251*T*
spiders **1:202–217**
adaptive radiation **1**:32, **1**:33*F*
alpine ecosystems **1**:155*F*
Amazon ecosystems **1**:246*T*
Arctic ecosystems **1**:231*T*
characteristics **1**:202–204, **1**:205–206, **1**:205*F*, **4**:392
ecological characteristics **1**:205–206
endangered species **3**:225–226
feeding behaviors **1**:204
food webs **7**:266*F*, **7**:267*T*
forest canopies **3**:512–513
habitats **1**:204
insecticide resistance **4**:298–299
major lineages
araneomorphs **1**:206–207
Entelegynae **1**:207
Haplogynae **1**:207
Liphistiidae **1**:206
mygalomorphs **1**:206
Near East ecosystems **5**:460*T*, **5**:463*T*
overharvesting/overcollecting **3**:223
paleontology **1**:204–205
phylogeny and taxonomy **1**:204, **1**:206
reproduction and growth **1**:203–204, **1**:206
riparian zones **6**:462–463
subterranean ecosystems **7**:53
Spiegelman, Solomon **5**:632
spike-nosed frog **1**:173*F*
Spilanthes acmella **4**:289*T*
spilled oil **3**:255, **3**:255*T*
spillover effects **2**:154
Spilosoma lutea **6**:159*F*
spinach beet **3**:131–132
spines and thorns **2**:428
spinifex **2**:468
Spinitermes **4**:409*T*
spinner dolphins **3**:192
spiny lobsters **4**:516, **5**:315, **6**:233
spiny octopus **7**:223
Spionida **7**:417*F*
Spionidae **7**:417*F*
Spiraea **1**:269
Spiraea trilobata **1**:271–272
Spiratrella coccinea **5**:124*F*
spirit sanctuaries **7**:229
spiritual services
human well-being **1**:437
valuation measures **5**:354
Spirochaetes **2**:611*F*, **2**:613–614
Spirogyra **3**:358
Spiroloculina **6**:291–292
Spirostachys africana **1**:67
Spix's macaw **7**:187
Spizella pusilla **4**:490*F*
Spizellomyces **5**:573–574*T*
splay-footed crickets **3**:730, **3**:730*F*
Spodoptera littoralis **5**:474–475
Spodosols **6**:123*T*
Spondias lutea **1**:161–162
Spondias mombin **4**:286–287, **4**:288*T*
sponges
classifications **4**:381*T*
convergent evolution **3**:615*F*
global diversity **4**:385*T*
hydrothermal vent ecosystems **7**:329–330
mangrove ecosystems **5**:20*T*
marine environments **4**:382*F*
number of species **4**:371*T*
protective compound production **5**:124–125, **5**:124*F*
river ecosystems **6**:478*T*
symbiotic relationships **5**:124–125, **5**:124*F*
spongillaflies **4**:371*T*
Spongospora subterranea **3**:631
spontaneous thelytokous speciation **6**:626
spoon-billed sandpipers **7**:185
spores **3**:650
Sporobolus marginatus **1**:69
Sporobolus pyramidalis **1**:52
Sporobolus robustus **7**:153–154
Sporobolus virginicus **3**:115–116
Sporothrix schenckii **3**:634*T*
sporotrichosis **3**:634*T*
sporozoans **6**:296–297
sportive lemur **6**:251*T*
SPOT program **6**:393–394
spotted eagles **5**:474
spotted flycatcher **5**:474
spotted hyenas **2**:190, **2**:190*F*, **2**:190*T*
spotted knapweed **4**:94, **4**:363
spotted mantis squid **2**:356
spotted owls **2**:264, **4**:564, **7**:186
spotted salamander **1**:10*T*, **6**:153*F*
spotted seal **5**:90
spotted turtles **7**:390
springbok **1**:487–488
spring communities **7**:58
springs **7**:50
springtails
Antarctic ecosystems **1**:184*T*
Arctic ecosystems **1**:231*T*
characteristics **4**:393*F*, **6**:586
lake and pond ecosystems **4**:374*F*
number of species **4**:392*T*
soil fauna **6**:590*T*
subterranean ecosystems **7**:53, **7**:55*F*
sprossers **1**:616*F*
spruce
boreal forests **1**:628–630, **1**:632*T*, **1**:633*F*, **7**:218*F*, **7**:219*F*
climate change effects **6**:150*T*
geographic boundaries **1**:629*F*, **1**:630*T*
greenhouse effect impacts **4**:28*F*
habitat–bird species diversity relationship **1**:612
North America **7**:118
soil properties and nutrients **6**:124*F*
whitebark pine forests **2**:571
Spumella **6**:293, **6**:294*F*
spurge **3**:262–263
spur-winged plover **5**:476–477
squacco herons **5**:476–477
Squamata **6**:401–402, **6**:406–408, **7**:340–341
squamous cell carcinoma (SCC) **7**:302
squash **2**:383, **2**:663–664
squash bugs **7**:294
squat lobster **7**:330
squids
bioluminescence **5**:123–124
microbial symbionts **5**:281*F*
physical characteristics **5**:379–380, **5**:380*F*
squidworms **2**:19*F*
Squilla mantis **2**:356
squirrelfish **3**:467–471
squirrel galagos **6**:251*T*
squirrel monkeys **6**:251*T*
squirrels **1**:631–632, **2**:190, **2**:584*F*, **2**:590*T*, **4**:564
squirting cucumber **2**:543
Sri Lanka
animal species **1**:261*T*
botanical gardens **6**:519
ecosystems
mangroves **1**:275*T*
monsoon forests **1**:263–266
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical dry forest **1**:266–267
tropical moist forest **1**:277*T*
tropical rain forests **6**:314*T*
tropical/subtropical humid mountain areas **1**:267
vegetation types and coverage percentage **1**:264–265*T*
fish diversity **3**:473
invasive species **4**:346–347
nitrogen deposition **5**:526*T*
plant species **1**:259*T*, **6**:82*T*
primate endemism **6**:252*T*
stability **7:1–11**
basic concepts **7**:1–2
changing environmental conditions
chaos–stability relationship **7**:8, **7**:8, **7**:9*F*
deterministic stability **7**:7–8, **7**:8*F*
temporal niches **7**:7
temporal stability **7**:5–7, **7**:6*F*, **7**:7*F*
deterministic stability **7**:7–8, **7**:8*F*
disturbance considerations **7**:2–3
diversity–stability hypothesis **7**:2
evolutionary stability **7**:10
spatial stability **7**:9–10, **7**:10, **7**:10*F*
structural stability
characteristics **7**:8–9
eutrophication effects **7**:9
Lotka–Volterra predator–prey system **7**:8–9, **7**:9, **7**:9*F*
temporal stability **7**:5–7, **7**:6*F*, **7**:7*F*
theoretical perspectives
algebraic interpretations
discrete-time models **7**:5, **7**:8

stability (*continued*)
eigenvalues **7**:5
general discussion **7**:4–5
multiple species **7**:5
resilience/strength of return **7**:5
singles species **7**:3*F*, **7**:4–5
ecological versus mathematical perspectives **7**:5*T*
equilibrium **7**:3–4, **7**:3*F*
geometric interpretations
carrying capacity **7**:3*F*, **7**:4
local versus global stability **7**:4
multiple species **7**:4
singles species **7**:4
stable isotopes
ecosystem function measurements **3**:70–71
oxygen-18 ($\delta^{18}O$) **1**:556–557, **1**:557*F*
Stachybotrys chartarum **3**:635
Stachys aegyptiaca **5**:484–485
Stachystemon **5**:623*F*
Stachyuraceae **1**:260*T*
Stackhousia **5**:623*F*
Stactobia pacatoria **5**:463*T*
stag beetles **3**:223
Stagetes borealis **7**:220*T*
staghorn coral **3**:201*T*
stalk-eyed flies **3**:489–491*T*
Standard Generalized Markup Language (SGML) **1**:444
Staphylinidae **1**:327*T*, **1**:328, **1**:709–710, **5**:674
Staphylinoidea **1**:327*T*
Staphylococcus **5**:710–711
Staphylococcus aureus **2**:636*F*
Staphylothermus **1**:219–220, **1**:221*T*
starfish **3**:125, **4**:45–46, **4**:335
stargazers **3**:467–471
starlings
diversity loss **1**:623–624
hemoglobin evolution **2**:55*F*, **2**:55*T*, **2**:56
introduced species **1**:78, **1**:81
species abundance **1**:601–602
star-nosed moles **4**:690–691
star phylogeny **6**:20, **6**:20*F*
stasis **2**:47–48
stated-preference estimation methods **7**:324–325
state property **6**:274, **6**:274*T*
Stator limbatus **2**:266
steady-state worldview **2**:702–703
Steatornis caripensis **1**:608, **2**:243–244, **7**:52–53
steelhead trout **3**:464–465, **6**:522, **6**:524–526
Stegoceras **2**:519*T*, **2**:520*T*
Stegodibelodon **5**:224*T*
Stegodon spp. **5**:225*T*, **5**:231–233
Stegodontidae **5**:224*T*, **5**:225*T*
Stegolophodon **5**:224*T*
Stegomastodon **4**:724–725*T*, **4**:727*T*
Stegomastodon platensis **5**:225*T*
Stegomastodon waringi **5**:225*T*
Stegotetrabelodon **5**:224*T*
Steinbeck, John **4**:652
Steinernema **7**:430–431
Stein, Gertrude **4**:652
Stelis **6**:316*T*
Stellarietea mediae **3**:340
Stelleroidea **3**:685*F*
Steller sea lions **5**:81*T*, **5**:82–83, **5**:447
Steller's sea cow (*Hydrodamalis gigas*) **3**:191–192, **3**:283, **4**:694, **5**:67–68, **5**:76, **5**:365
Stemmiulida **5**:423*F*
Stemmiulus adisi **5**:423*F*
stem parasites **4**:70
stem rust **1**:359–360
Stenella frontalis **3**:192–193
Stenella longirostris **3**:192
Stenella longirostris orientalis **5**:87*T*
Stenodactylus **5**:467–469
Stenodactylus petrii **5**:472
Stenodactylus sthenodacylus **5**:472
Stenolaemata **3**:685*F*
Stenopelmatidae (Jerusalem Crickets) **3**:731–732
Stenopelmatoidea
Anostostomatidae (wetas) **3**:731, **3**:731*F*
characteristics **3**:730–731
Cooloolidae (Cooloola Monsters) **3**:731
Gryllacrididae (raspy crickets and leaf-rolling crickets) **3**:732
Stenopelmatidae (Jerusalem Crickets) **3**:731–732
Stenopodidea **2**:416–417*T*
Stenopterygii **3**:465
Stenorhinus hemitoechus **4**:729–730, **4**:729*F*
Stenorhinus kirchbergensis **4**:729–730, **4**:729*F*
Stentor **6**:295
Stephanodiscus hantzschii **6**:33*F*
Stephanoeca **6**:293
Stephanoidea **4**:178
Stephanomeria malheurensis **3**:209
Stephen Island wrens **1**:621, **5**:366–367
steppes
adaptive radiation **1**:22, **1**:23*F*
biodiversity change scenarios **4**:6*F*
C_4 plants **2**:371*T*
current global land cover **4**:557*F*
dry Asian regions **1**:269
Europe **3**:335*T*, **3**:336–340*T*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
protected area systems **6**:492*T*, **6**:499*T*
southern Andean steppes
altitudinal zonation **4**:106–107, **4**:106*F*
bogs and seeps **4**:107
cushion plants **4**:106–107, **4**:106*F*
geographic distribution **4**:96–97, **4**:97*F*
latitudinal gradients **4**:106–107
number of species **4**:102*T*
physical characteristics and patterns **4**:97–98, **4**:105–107
plant diversity
endemism **4**:99–100, **4**:99*T*
origins and diversification **4**:100–101
species richness **4**:98–99
representative species **4**:105*F*
Sterculaceae **5**:11*T*
Sterculea **1**:69–70
Sterculiaceae **1**:263, **1**:277
Stereospermum kunthianum **1**:67
Sterna fusca **6**:78
Sterna maxima **3**:256
Sterna paradisaea **1**:605, **5**:311–312
Sternbergia clusiana **5**:484–485
Sternidae **1**:710
steroidal glycosides **6**:136–137, **6**:137*F*
stewardship **7:12–22**
biodiversity stewardship **2**:297, **2**:300
contexts
cultural factors **7**:17–18
ecological factors **7**:16–17
economic factors **7**:17
general discussion **7**:16–17
political factors **7**:17–18
etymological background **7**:12–13
outside protected areas **2**:297, **2**:300
rationales
counterarguments **7**:14–16
general discussion **7**:13–14
instrumental arguments **7**:13–14
intrinsic arguments **7**:14
research summary **1**:167–168
steward types
community-based conservation **4**:320–321, **7**:19–20
cross-boundary stewardship **7**:21
general discussion **7**:18–19
governmental stewards **7**:18–19
nongovernmental organizations (NGOs) **7**:19–20
private companies and individuals **4**:330, **7**:20–21
threatened species lists and assessments **7**:207–208
Sthanturus andersoni **5**:364*F*
Sthanturus atlas **5**:364*F*
Sthanturus brownei **5**:364*F*
Sthanturus gilli **5**:364*F*
Sthanturus maddocki **5**:364*F*
Sthanturus occidentalis **5**:364*F*
Sthanturus oreas **5**:364*F*
Sthanturus orientalis **5**:364*F*
Sthanturus pales **5**:364*F*
Sthanturus tindalei **5**:364*F*
St. Helena giant earwig **3**:222
St. Helena redwood **6**:77
Sthenopis auratus **5**:393*F*
Sthenurus **4**:727–728
sticklebacks **3**:467–471, **5**:314–315
Stigeoclonium **3**:358
stigmasterol **5**:712*T*
Stigmatopelia senegalensis **5**:476
Stigmella aceris **5**:385*F*
Stigonema tomentosum **1**:160
Stilbaceae **5**:219–220
stiletto flies **3**:489–491*T*
stilt-legged flies **3**:489–491*T*
Stilton cheese **3**:638–639
stilts **5**:476–477
stimulants
East Asia **2**:663
North and South America **2**:663–664
Sub-Saharan Africa **2**:663
stinging corals **6**:295

stingless bees **4**:183
stinkbugs **7**:293, **7**:294
Stipa baicalensis **1**:269
Stipa capillata **1**:269, **1**:270
Stipagrostis **2**:376*F*
Stipagrostis uniplumis **1**:67
Stipa ichu **2**:30
Stipa krylovii **1**:269
Stipa lessingiana **1**:269
Stipa stenophylla **1**:269
Stipiturus mallee **7**:183–184
St. John's wort **3**:704, **3**:704*F*, **7**:261–262
St. Maarten **2**:187–188
stoats **1**:685–686, **1**:692*F*, **5**:370, **7**:28, **7**:28*F*
stochastic simulation models **1**:569–572, **1**:572*F*
Stockholm Convention on Persistent Organic Pollutants (2001) **7**:137
Stockoceros **4**:722
Stolotermes **4**:409*T*
stomatal index **4**:225*F*, **4**:226, **4**:227*F*
Stomatopoda **2**:416–417*T*
Stomoxys calcitrans **3**:492
stonefish **3**:468
stoneflies
  American Southeast rivers **3**:185*F*
  freshwater ecosystems
    organic pollution **3**:580–581, **3**:581*F*
    threatened species **3**:579*T*
  number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
  surface water acidification **1**:10*T*
  wetland ecosystems **7**:389–391
stoneworts **3**:571–572, **3**:571*T*
stony corals **3**:199, **3**:202*T*, **6**:295
storage ecology **7:23–38**
  adipose tissue
    adipocyte size and number **7**:27
    anatomical arrangement **7**:27–30, **7**:28*F*, **7**:29*F*
    characteristics **7**:26–27, **7**:27*F*
    fatness measures **7**:30–31, **7**:30*F*
    functional role
      adipocyte metabolism **7**:31–32
      fatty acid composition **7**:32–33
      thermal insulation **7**:33–35, **7**:34*F*
  invertebrates **7**:24
  life histories
    energy storage optimization **7**:35
    hibernation **7**:36–37
    migration **7**:35–36
    reproduction and lactation **7**:37–38
  material characteristics **7**:23–24
  plants
    growth cycles **7**:24–25
    lipid composition **7**:25–26
    predator protection **7**:26
    seed storage **7**:24–25
storm peak management **5**:352–353
stormwater management **6**:459
strangler figs **3**:520–521, **4**:74, **7**:272–273
Strategic Plan for Diversity (2011-2020) **2**:311, **2**:312, **2**:317–318, **6**:78, **6**:80–81*T*
stratigraphic records **1**:430
Stratiomyidae **3**:489–491*T*, **3**:493
Stratiomyomorpha **3**:489–491*T*
stratospheric ozone
  concentration deviations **1**:289*F*
  current trends **1**:287–289, **1**:289*F*
  depletion causes and consequences **1**:289
  future outlook **1**:290
  occurrences and characteristics **1**:287–289
  projected trends **1**:289
  total ozone column **1**:287–289, **1**:288*F*
strawberries **2**:383
strawberry poison dart-frog **4**:518
straw-headed bulbul **7**:187
streams
  climate change effects **6**:150–151
  conservation status **3**:171–173*T*
  dam projects **1**:510–512
  freshwater invertebrates **4:369–378**
    background information **4**:369–370
    ecological functions and processes
      characteristics **4**:375
      decomposition and recycling **4**:375
      material transfer and uptake **4**:375–376
      primary production **4**:376
    global patterns **4**:373
    influencing factors
      disturbances **4**:375
      endemism **4**:374–375
      habitat heterogeneity **4**:375
      isolation factors **4**:374–375
      origins **4**:374
      research background **4**:373–374
      species interactions **4**:375
    life cycles **4**:370, **4**:372*F*
    major groups **4**:370, **4**:371*T*
    major habitats **4**:369–370, **4**:369*F*
    major threats
      altered water regimes **4**:377–378
      biodiversity loss **4**:376–377
      exotic species invasions **4**:378
      habitat depletion and loss **4**:376–377
      pollution **4**:377
      reduced water quality **4**:377
    number of species **4**:370–371, **4**:371*T*
    river continuum concept (RCC) **4**:373*F*
    taxonomic groups **4**:371–373
    zonal representatives **4**:374*F*
  geographic distribution **4**:370*T*
  habitat diversity **3**:570–571
  organic pollution **3**:575, **3**:580–581, **3**:581*F*
  predation effects **6**:235–236
  protected area systems **6**:494–498, **6**:496*F*, **6**:497*F*
  research summary **3**:585–586
  Southeast Asia **1**:510–512, **1**:511*F*, **1**:512*F*
  species richness **3**:572*T*, **3**:578–579, **3**:578*F*
  stream channelization **3**:573–574, **6**:481–482, **6**:481*T*
  threatened species **3**:579–580
stream salamander **4**:516
Streblidae **3**:489–491*T*
Streblospio benedecti **4**:384
Strepsiptera **4**:308–312, **4**:309*T*, **4**:392*T*
Strepsirrhini **6**:250–252, **6**:251*T*
Streptococcus pneumoniae **2**:636*F*
Streptococcus pyogenes **2**:636*F*
Streptomyces griseus **1**:359
Streptopelia decaocto **2**:551–552*T*, **5**:476
Streptopelia (Stigmatopelia) senegalensis **5**:474
Streptostachys **2**:376*F*
stress
  environmental stress **7:39–44**
    assessment measures
      appraisals **7**:42
      biomonitoring processes **7**:43
      early warning **7**:43
      functional role **7**:42
      predictions **7**:42–43
    biospheric life support system **7**:43
    characteristics **7**:39
    characterization studies **7**:39–40
    future trends **7**:43–44
    general environmental stress syndrome **7**:41–42, **7**:42*T*
    natural versus anthropogenic stress **7**:39
    receptors and responses **7**:40–41, **7**:41*T*
    spatiotemporal distributions **7**:39–40
strict consensus analysis **2**:42–44, **2**:43*F*
Striga **4**:71, **5**:673
Striga asiatica **4**:76–77, **4**:362
Striga gesneroides **4**:76–77
Striga hermonthica **4**:76–77
Strigamia maritima **5**:425
Strigea falconispalumbi **5**:660–661, **5**:663*F*
Strigidae **6**:308*T*
Strigiformes **6**:308*T*
Strigops habroptilus **1**:604, **1**:606, **6**:579, **7**:188
striped bass **6**:479
striped maple **6**:433*F*
stripe-naped proboscis monkey **6**:256–258*T*
Strix occidentalis **4**:564, **7**:186
Strix occidentalis caurina **2**:225–226, **2**:243–244, **2**:264, **3**:314–315, **3**:315*F*
Strombidium **6**:31, **6**:295
Strombidium purpureum **6**:295
Strombus gigas **3**:199, **6**:547–548
Stronglyocentrotus purpuratus **3**:283
Strongylocentrotus droebachiensis **3**:355–356
Strongyloides stercoralis **5**:664, **5**:671*T*, **7**:428*T*
Strongylophthalmyiidae **3**:489–491*T*
strophanthin **5**:711–713, **5**:712*T*
Strophanthus gratus **5**:712*T*
Stropharia rugoso-annulata **3**:637*T*
structured metadata
  Access to Biological Collection Data (ABCD) standards **1**:443
  Darwin Core **1**:443
  Ecological Metadata Language (EML) **1**:443
  eXtensible Markup Language (XML) **1**:442–443
  geospatial data standards **1**:443–444
  machine integration challenges **1**:442
Struthio camelus **1**:52, **1**:601, **1**:606, **1**:624, **2**:472*F*
Struthio camelus syriacus **5**:472

Struthiomimus **2**:519*T*, **2**:520*T*
Struthioniformes **1**:607
Strychnos cocculoides **1**:68
Strychnos guianensis **5**:712*T*
Strychnos spinosa **1**:68
Stryker's snub-nosed monkey **6**:259
Stuart Little(White) **4**:653
Stuart's marsupial mouse **4**:687*T*, **4**:688
studbooks **7**:474–475
Stufenreihe **2**:51, **2**:51*F*
stunting **3**:633*T*
sturgeon
  characteristics **3**:462, **3**:462*F*
  introduced parasites–species decline relationships **2**:528*T*
  migratory patterns **5**:314–315
Sturnella magna **4**:490*F*
Sturnus vulgaris **1**:78, **1**:601–602, **1**:623–624, **2**:551–552*T*, **5**:312
Stygimoloch **2**:519*T*, **2**:520*T*
Styginae **1**:651–652
Stygiolobus **1**:221*T*
stygobionts **7**:52, **7**:55–57, **7**:59*T*
Stygoparnus comalensis **7**:58
stygophiles **7**:52, **7**:55–57
stylasterine corals **2**:335
Stylidium **5**:221
Stylocheiron maximum **4**:386*T*
Stylocheiron suhmi **4**:386*T*
Stylonychia lemnae **5**:569–570
Styphlotrema **7**:450–451*F*
Stypopodium zonale **4**:12
Styracosaurus **2**:519*T*, **2**:520*T*
Styrax obassia **1**:271
Styrax officinalis **5**:482
Styrax tonkinensis **5**:712*T*
Styx infernalis **1**:651–652
Suada **2**:371*T*
Suaeda aralocaspica **2**:364–365, **2**:365*F*
Suaeda asphaltica **5**:485–486
Suaeda divaricata **2**:469
Suaeda foliosa **4**:105
Suaeda monoica **5**:486
Suaeda palestina **1**:268
Suaeda salsa **2**:376*F*
subalpine fir **2**:571
Subantarctic fur seals **5**:87*T*
subcutaneous mycoses **3**:634*T*
subhumid forests
  high mountain ecosystems **7**:275
  lowland ecosystems **7**:273–274
  low mountain ecosystems **7**:274–275
subhumid lands **6**:492
subjugated people **6**:221*T*, **6**:222–223
submarine canyons **5**:53
subpáramo zone **4**:102–103
Sub-Saharan Africa **2**:663
Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA) **2**:307–308
subspecies taxonomy **7:45–48**
  historical background
    binominal versus trinominal nomenclature **7**:45–46
    divergence theories **7**:46
    levels of variation **7**:45
  modern perspectives
    phylogenetic species concept **7**:46–47
    semispecies taxonomy **7**:46–47
substituted amides **5**:696
substituted aromatics **5**:695
substituted nitrites **5**:697
substituted ureas **5**:696
subterranean ecosystems **7:49–62**
  adaptive behaviors
    characteristics **7**:53–55, **7**:54*T*
    evolutionary rates **7**:54*T*
    genetic expression **7**:53–55, **7**:54*F*
  anchialine communities **7**:58, **7**:59*T*
  aquatic subterranean habitats **7**:49–50, **7**:50*F*
  Ariège (France) **7**:60
  bat colonies **7**:60
  biogeographical patterns **7**:60–62, **7**:60*T*, **7**:61*F*
  Bracken Cave (Texas, USA) **7**:60
  characteristics **7**:49–50
  chemoautotrophic communities **7**:58–59, **7**:59*T*
  colonization and speciation **7**:55–57, **7**:56*F*, **7**:57*F*
  conservation and protection strategies **7**:62
  Cueva de Felipe Reventón (Canary Islands) **7**:59*T*
  Cueva del Sobrado (Canary Islands) **7**:59*T*
  Cueva del Viento (Canary Islands) **7**:59*T*
  Edwards Aquifer (Texas, USA) **7**:58, **7**:59*T*
  energy sources **7**:51–52
  epikarst communities **7**:56*F*, **7**:57–58, **7**:59*T*, **7**:60
  groundwater calcretes (Western Australia) **7**:59
  Gua Salukkan Kallang-Towakkalak (Indonesia) **7**:59–60
  hydropetric communities **7**:58, **7**:59*T*
  hypotelminorheic communities **7**:59
  interstitial communities
    Lobau wetlands (Austria) **7**:59
    Rhone River (France) **7**:59
  Kazumura Cave (Hawaii, USA) **7**:60
  Krizna jama (Slovenia) **7**:59*T*
  lava tube communities **7**:60
  Lobau wetlands (Austria) **7**:59
  Logarcek (Slovenia) **7**:59*T*
  Mammoth Cave (Kentucky, USA) **7**:59*T*, **7**:60
  milieu souterrain superficiel (MSS) **7**:51, **7**:60
  Movile Cave (Romania) **7**:58–59, **7**:59*T*
  myriapods **5**:431
  Organ Cave (West Virginia, USA) **7**:58, **7**:60, **7**:60*T*
  Postojna Planina Cave System (cave stream communities) **7**:58, **7**:59*T*
  Postojna Planina Cave System (epikarst communities) **7**:56*F*, **7**:57–58, **7**:59*T*
  Rhone River (France) **7**:59
  Robe River (Australia) **7**:59*T*
  Rock Creek (Washington, D. C., USA) **7**:59
  Sika-Krka System (Slovenia) **7**:59*T*
  species richness **7**:60–62, **7**:60*T*, **7**:61*F*
  spring communities **7**:58
  taxonomic diversity **7**:52–53, **7**:52*F*
  terrestrial communities
    Gua Salukkan Kallang-Towakkalak (Indonesia) **7**:59–60
    Kazumura Cave (Hawaii, USA) **7**:60
    Mammoth Cave (Kentucky, USA) **7**:59*T*, **7**:60
    Organ Cave (West Virginia, USA) **7**:60
    Vjetrenica (Bosnia and Herzegovina) **7**:60
  terrestrial subterranean habitats **7**:50–51
  Triadou Aquifer (France) **7**:59*T*
  Vjetrenica (Bosnia and Herzegovina) **7**:58, **7**:59*T*, **7**:60
  Walsingham Caves (Bermuda) **7**:58, **7**:59*T*
subtidal habitats **5**:57–58, **5**:60*F*, **5**:107–108
subtropical deserts **1**:262–263*T*, **1**:267–268
Subulitermes **4**:409*T*
succession **7:63–70**
  basic concepts **7**:63
  biodiversity
    disturbed landscapes **7**:69–70
    mature communities **7**:68–70
    species gains and losses **7**:69*F*, **7**:70
  boreal forests **1**:632–634
  ecosystem function **3**:93, **3**:93*T*
  ecosystem research **5**:582
  habitat–bird species diversity relationship **1**:612–613, **1**:612*F*, **1**:612*T*, **1**:613*F*
  mechanistic theories
    biomass dynamics **7**:67, **7**:67*F*
    comparison studies **7**:65–66, **7**:65*T*
    descriptive models **7**:65–66, **7**:66*F*
    landscape mosaics **7**:66–68, **7**:67*F*
    quantitative models **7**:68
    vegetation patterns **7**:65–66
  metapopulation theory **5**:243–244, **5**:243*F*
  organic explanations
    background information **7**:63–64
    Clementsian theory
      basic principles **7**:63–64
      climax community **7**:64
      community-level management **7**:64
      comparison studies **7**:65–66, **7**:65*T*
      indicator species **7**:64
      progressive nature concept **7**:64
    Cowles' theory **7**:64–65
  restoration efforts **6**:443
  terrestrial plant communities **6**:70–71
  vegetation succession **3**:90*T*
Succulent Karoo **1:485–490**
  biodiversity
    biodiversity-friendly farming **2**:300
    invertebrates **1**:488
    plants **1**:485, **1**:486*T*, **1**:487, **1**:487*T*
    vertebrates **1**:487*T*
  bioregions **1**:485, **1**:486*F*
  characteristics **1**:485, **1**:486*F*
  climate change threats **1**:489
  land degradation threats **1**:489
  nitrogen deposition **5**:526*T*
  plant growth forms **1**:488–489
  plant species **6**:82*T*

threat management **1**:489
vegetation structure **1**:488–489, **2**:467–468
sucking lice **3**:226, **4**:392*T*
sucking millipedes **5**:429–430
suctorians **6**:295–296, **6**:296*F*
Sudan
botanical gardens **6**:519
deforestation **2**:449*T*, **2**:451*T*
domesticated ruminants **1**:702–703
forest product exports **2**:451*T*
forest-savanna mosaics **1**:68
indigenous peoples **4**:270*T*
tropical forests **7**:269–270, **7**:270*T*
Sudanian tree oases **5**:480*F*, **5**:486
Sudbury, Ontario **1**:140–141, **1**:141*F*, **6**:155
Sudbury smelter complex **1**:140–141, **1**:141*F*
sudden oak death **2**:528*T*, **4**:505
Suessiales **1**:472
sugar **2**:663
sugar beets **1**:529*T*, **7**:24–25
sugar cane **1**:529*T*, **1**:543–547, **2**:663, **3**:32*F*, **4**:432*T*, **6**:553
sugaring **5**:394
sugar maples **1**:8, **1**:8*F*, **6**:433*F*, **7**:125
sugar pine **6**:152–154
Suidae **4**:696
Suillus americanus **3**:631–632
Suisun thistles **4**:361, **4**:363
Sulawesi **1**:514*T*
Sulawesi black macaque **6**:256–258*T*
Sulfolobus **1**:219–220, **1**:221*T*, **2**:640–641
sulfonylurea **4**:90*T*
sulfur (S)
acidic deposition **3**:577–578
biogeochemical cycles **1**:558–559, **1**:561–563, **6**:154–155, **6**:154*F*
carbon disulfide (CS_2) **1**:282*T*
carbonyl sulfide (COS) **1**:282*T*, **1**:287
dimethyl sulfide (CH_3SCH_3) **1**:282*T*, **1**:287
elemental partitioning **1**:554*T*
global sulfur cycle **3**:282, **5**:111, **6**:154–155, **6**:154*F*
hydrogen sulfide (H_2S) **1**:281–282, **1**:282*T*, **1**:287, **1**:561–563, **6**:154–155, **6**:154*F*
hydrothermal vent ecosystems **7**:328, **7**:329–330
inorganic herbicides **5**:695–696
relative abundance **1**:553*F*
river water composition **3**:562*T*
root–soil interface **6**:114*F*
soil accumulations **1**:6
sulfate deposition **1**:3*F*, **3**:577–578, **6**:125–126
sulfur dioxide (SO_2)
atmospheric concentrations **1**:282*T*, **1**:287
biodiversity impacts **1**:140–141
biogeochemical cycles **6**:154–155, **6**:154*F*
climate change effects **1**:281–282
ecotoxicological studies **3**:118–119
emission occurrences and sources **1**:1–5, **1**:137*T*
forest ecosystems **1**:5*T*
smelting industry impacts **1**:140–141
sulfuric acid **1**:1–5, **1**:3*F*
Sumatra **1**:262–263*T*, **1**:514*T*, **5**:606
Sumatran orangutan **6**:256–258*T*
Sumatran rhinoceros **3**:196–197, **4**:695
Sumatran tigers **1**:517, **5**:606
Sumatran white-handed gibbon **6**:256–258*T*
summer squash **6**:553
Sumner, William Graham **4**:652
sunbirds **1**:617
sunburn **7**:302
Suncus etruscus **1**:601
Sundaland **5**:526*T*, **6**:82*T*
Sundarbans **1**:276*T*
Sundew **2**:643
sundews **1**:136
sunfish **3**:467–471, **3**:471*F*, **5**:657, **7**:390
sunflower oil **1**:529*T*, **2**:663–664, **7**:25
sunflowers **2**:383, **2**:391, **3**:32*F*, **4**:167
sunis **6**:572–573
Sunn pests **7**:293
sun spiders **1**:213
superficial mycoses **3**:634*T*
Superior (Lake) **4**:461*T*
superpáramo zone **4**:103
superparasitism **5**:675
Superstitionidae **1**:209
superthermophiles **7**:328
supertramp dove **1**:617–618
surface albedo **2**:680–681
surface energy partitioning **2**:680–681
surface fires **3**:436
surface roughness **2**:680–681
surface temperatures **2**:83–84
surface water
acidic deposition **1**:7*F*, **1**:8–9, **1**:10*T*, **1**:144–145
biodiversity impacts **1**:144–145
habitat diversity **3**:570–571
mercury contamination **1**:12
water storage **3**:560–562, **3**:561*T*
surfperch **3**:467–471
surgeonfish **3**:467–471
surilis **6**:251*T*
Suriname
botanical gardens **6**:519
land cover classifications **6**:607*T*
tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*
surplus killing **1**:688
Sus barbatus **1**:517, **5**:603–604
Sus bucculentus **4**:703–704
Sus salvanicus **4**:696
Sus scrofa **1**:77, **1**:517, **1**:640, **4**:359, **5**:603–604
Sus scrofa vittatus **5**:603
sustainability
biodiversity **7:71–84**
biodiversity–population density relationship **7**:74*F*
carrying capacity **7**:75
demographic trends **7**:72*F*
ecological resilience **7**:80–81, **7**:81*F*
ecosystem services **7**:76*T*
global trends **7**:81–83, **7**:82*F*, **7**:83*F*
human well-being indicators **7**:77–80, **7**:78*F*
as natural capital component **7**:76–77
poverty reduction **7**:77–80
sustainability–biodiversity relationship **7**:71–75
sustainable development **7**:75
theory of inclusive wealth **7**:73, **7**:75–76
biodiversity conservation **4**:322, **4**:436
environmental ethics **3**:311–312
fishing/fisheries **3**:483
global urbanization **4**:239–241, **4**:239*F*
social well-being **3**:27–28
sustainable agriculture **1:111–118**
agrobiodiversity **1**:129–130
basic concepts **1**:111
cropping systems
agricultural production–ecosystem maintenance relationships **1**:116*F*
animal integration **1**:116
bush-fallow rotation systems **1**:114–115
ecosystem services **1**:114
multifunctional agriculture **1**:115–117, **1**:115*F*, **1**:116*F*, **1**:116*T*, **1**:117*F*, **2**:300–301
resource conservation **1**:113–114
developmental background **1**:111–113
global agriculture **3**:430–431
global marketplace **1**:117
indicators
criteria **1**:113–114
ecosystem services **1**:114
resource conservation **1**:113–114
Svalbard Global Seed Vault **4**:328, **6**:86
Svalbard reindeer **7**:27
Svalbard rock ptarmigans **7**:27–28
Swainson's thrush **7**:187
swamp eels **3**:467–471
swamps
Amazon ecosystems **1**:164
Asia **1**:277–278
conservation status **3**:171–173*T*
definitions and terminology **7**:369*T*
global extent **7**:370*T*
habitat loss and degradation **7**:371*T*
Near East ecosystems **5**:472–473, **5**:474*F*, **5**:475*T*, **5**:480*F*, **5**:486
United States **7**:385*F*
swans **3**:365–366
swarms **3**:495
swash activity **2**:134
Swash Exclusion Hypothesis **2**:134
Swaziland **6**:507*T*
sweat bees **1**:420*F*
Sweden
boreal forests **7**:212–213, **7**:212*T*
botanical gardens **6**:519
natural history museums and institutions **5**:411*F*
nitrogen deposition **5**:529*F*
nonvascular epiphytes **3**:521*T*
Swedish Museum **5**:411*F*
sweep netting **1**:330

sweet alyssum **1**:376*F*
sweet-blue lupine **2**:389–390
sweet cherries **2**:662–663
sweet gale **7**:395–397
sweetgum **1**:529*T*
sweet peppers **3**:32*F*
sweet potatoes **1**:132*T*, **3**:32*F*, **4**:287–290, **6**:553
sweet prickly pear **4**:350–351
swiddens **4**:140–141, **6**:551, **6**:567, **7**:231–232
  see also slash-and-burn systems
Swietenia macrophylla **1**:161–162, **4**:595, **6**:560, **7**:273
Swietenia spp. **2**:25–26
swift fox **3**:192
swifts **1**:604*F*
switchgrass **1**:529*T*, **1**:532, **2**:391, **3**:261
switching **6**:244
Switzerland
  botanical gardens **6**:519
  natural history museums and institutions **5**:411*F*
  vegetation surveys **3**:341–344*T*
sword beans **2**:663
swordfish **2**:8–13, **2**:12*F*, **3**:410, **3**:467–471
swordtail crickets **1**:32–33
swordtail fish **3**:467
Syagrus coronata **7**:187–188
sycamore trees **1**:139
Sydney Opera House **1**:590–591
Syllidae **7**:417*F*
Sylvia communis **5**:474
Sylviidae **6**:308*T*
Sylvilagus aquaticus **7**:390
Sylvilagus brasiliensis **2**:587*F*, **6**:332–333
Sylvilagus palustris **7**:390
Symbiodinium **5**:118–119, **5**:119*F*, **5**:306, **6**:292–293, **6**:295
Symbiodinium microadriaticum **2**:335, **5**:118–119, **6**:291–292
symbiotic relationships
  algae **5**:306, **7**:300
  bacteria **5**:304–305
  cooperation **2**:322
  definition **5**:656
  fungi **4**:418–419, **4**:419*F*
  Isoptera
    biochemical processes **4**:418*F*
    carbon-nitrogen balance symbiosis **4**:416–417, **4**:417–418, **4**:417*F*
    cellulose decomposition symbiosis **4**:417
    digestive tube **4**:415–416, **4**:416*F*
    fungi **4**:418–419, **4**:419*F*
    general discussion **4**:408–409
    hindgut microenvironment **4**:416*F*
    phylogeny **4**:417*F*
    symbionts **4**:422*F*
  marine ecosystems **5:116–126**
    functional roles **5**:117–118
    future challenges **5**:125
    global distribution **5**:117*F*
    historical research **5**:116–117
    light-producing symbioses
      bioluminescence **5**:121–123, **5**:122*F*
      fish **5**:123
      squid **5**:123–124
    nutrient conversion
      nitrogen fixation **5**:121
      shipworms **5**:121
      worm–whale interaction **5**:121, **5**:121*F*
    nutrient-exchange mutualisms
      chemosynthetic symbioses **5**:120–121, **5**:120*F*
      hydrothermal vents **5**:120–121, **5**:120*F*
      nutrient conversion **5**:121
      photosynthetic symbioses **5**:118–120
    protective compound production characteristics **5**:124–125
      sponges **5**:124–125, **5**:124*F*
    structural characteristics **5**:116–117, **5**:117*F*
    transmission modes **5**:116–117, **5**:117*F*
  microorganisms **6**:449–450, **6**:450*F*
  plant–animal interactions **6**:61–62
  species interactions **6**:719
Symetrodonta **4**:684
symian immunodeficiency virus (SIV) **4**:516
symmetry **1**:40–41
sympatric speciation **3**:144–145, **6**:621*F*, **6**:622*T*, **6**:625, **6**:635–636
Symphalangus **4**:692–693
Symphalangus syndactylus **6**:256–258*T*
Symphonia globulifera **4**:286–287
Symphyla
  Amazon ecosystems **1**:246*T*
  characteristics **4**:392
  developmental modes **5**:426
  geographic distribution
    endemism **5**:428
    environmental characteristics **5**:427–428
    hotspots **5**:429
    islands **5**:429
    regionalization **5**:428–429
    speciation **5**:428
  habitats and adaptations
    burrowing myriapods **5**:430
    evolutionary development **5**:429–430
    open-country myriapods **5**:430
    special habitat myriapods **5**:430–431
  morphology **5**:424, **5**:424*F*
  number of species **5**:427*T*
Symphylella vulgaris **5**:428
Symphysodon discus **3**:481
Symphyta **4**:177–178, **5**:674
symplesiomorphy **7**:104–106
Symplocaceae **1**:267
Symplocos **7**:274–275
Symplocos pichindensis **1**:167
Synacanthotermes **4**:409*T*
synapomorphy **2**:34–35, **2**:34*F*, **2**:53, **7**:104–106, **7**:105*F*
synapsids **5:1–9**
  early Paleocene faunal recovery **5**:7
  evolutionary history **5**:1–2
  extinctions
    causal factors
      background extinctions **5**:2–3
      distal and proximal causal factors **5**:3–4
      mass extinctions **5**:2–3
      pseudoextinctions **5**:2
      temporal and biogeographic scales **5**:3
    Cretaceous–Tertiary (K–T) mass extinction **5**:6–7
    late Triassic faunal turnover **5**:4–5
    Permian–Triassic mass extinction **5**:4
  geologic timescale **5**:1*T*
  origins **5**:4
  Paleocene–Eocene thermal maximum **5**:7–8
  phylogenetic relationships **5**:1*F*
  summary discussion **5**:8
  Tertiary faunal evolution **5**:7–8
  therapsid radiation **5**:4–5
Synarundinaria alpina **1**:53
Synbranchiforms **3**:468
Syncarida **2**:401*T*, **3**:182*T*
Synceros caffer **6**:572–573
Syncerus caffer **1**:51, **1**:487–488, **4**:362
Synchaeta **6**:26
Synchytrium endobioticum **3**:627
Syndesmis **7**:442–443*F*
Synechoccous lividus **1**:315
Synechococcus lividus **4**:114
Synechococcus sp. **3**:357, **4**:671–673, **4**:675*F*, **5**:134–135, **5**:293–294, **5**:302–303, **5**:685*F*
Synecococcus **6**:33*F*
Syngonium **3**:522*T*
Synneuridae **3**:489–491*T*
Synodontis **4**:463
Syntermes **4**:409*T*
synthetic nitrogen fertilizers **5**:520–522, **5**:522*F*, **5**:538–540, **5**:540*F*
Synthliborhamphus antiquus **5**:447
Synura **6**:293
Syria
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  ecosystems
    flora **5**:481*T*
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
  wheat production **1**:87–88
  see also Near East ecosystems
Syrian woodpeckers **5**:474
Syringa amurensis **1**:271–272
Syringodium **6**:542*T*
Syringodium filiforme **3**:360–361, **6**:541*T*
Syringodium isoetifolium **6**:541*T*
Syringogastridae **3**:489–491*T*
Syrphidae **3**:489–491*T*, **3**:492
Syrrhaptes paradoxus **2**:549–550
systematic conservation planning **2**:293, **6**:261–263
systematic rearing **1**:330
systematics **7:85–96**
  alternative species concepts **7**:93, **7**:93*T*
  attributes **7**:87–89
  background information **7**:85

basic concepts **7**:85
biodiversity perspectives **7**:90–91, **7**:91*F*
cladistics **7**:87
comparative methods **7**:86
evidentiary research **7**:87–89
Hennigian terminology **7**:91–93
historical perspective **7**:86, **7**:89–90
Linnean nomenclature **7**:90
mission and vision
data dissemination **7**:94
future outlook **7**:93–94
phylogenetic classifications **7**:94
species discovery and inventory **7**:93–94
museum research programs **5**:412
philosophical perspectives **7**:86
predictive classifications **7**:91–93, **7**:92*F*
research collections **7**:86
taxon inventories **7**:86
taxonomic importance
biomimicry and adaptations **7**:95
communication tools **7**:94–95
conceptual framework **7**:94
environmental baseline **7**:95
evolutionary record **7**:95
general discussion **7**:94
information storage and retrieval **7**:95
taxon scholars **7**:86
tokogenetic relationships **7**:91, **7**:91*F*
see also taxonomy
System forTransparent Allocation of Resources (STAR) **7**:137–138
systemic mycoses **3**:634*T*
systems of nomenclature **5**:547–559
alternative nomenclature systems **5**:556
Code system
algae, fungi, and plants **5**:550–551, **5**:553*T*
ambiregnal organisms **5**:554–555
cultivated plants/cultivars **5**:551, **5**:553*T*
Draft BioCode **5**:553–554
fossils **5**:555
fungi **5**:550–551, **5**:553*T*, **5**:555–556
hybrids **5**:555
lichens **5**:555
PhyloCode **5**:554
prokaryotes **5**:551–552, **5**:553*T*
viruses **5**:552
zoological organisms **5**:552–553, **5**:553*T*
functional role **5**:547
general discussion **5**:547
hierarchical organization **5**:548–549, **5**:548*T*
historical perspective **5**:547–548
name-formation principles **5**:549–550
naming process
name changes **5**:556–557
name determination guidelines **5**:557–558
newly discovered species **5**:557
reference indices **5**:558
Syzygium **1**:267, **6**:316*T*
Syzygium aromaticum **2**:663, **5**:712*T*

T

Tabanidae **3**:489–491*T*, **3**:492
Tabanomorpha **3**:489–491*T*
Tabebuia **5**:489
Tabebuia caraiba **1**:165
Tabellaria flocculosa **6**:33*F*
Tabernanthe iboga **1**:594
tabinau **7**:233
taboos
ecological monitoring and practices
habitat protection **7**:227*T*, **7**:229
harvest restrictions **7**:227*T*, **7**:229–230
species protection **7**:227*T*, **7**:228, **7**:229*T*
vulnerable life histories protection **7**:227*T*, **7**:228–229
religious traditions **6**:563–564
Tacca leontopetaloides **2**:663
Tachigali myrmecophila **6**:46–47
Tachinidae **3**:489–491*T*, **3**:492, **5**:395–396, **5**:674
Tachiniscidae **3**:489–491*T*
Tachycineta bicolor **5**:720
Tachyglossidae **4**:686
Tachyglossus aculeatus **4**:686
Tadarida brasiliensis **5**:313, **7**:53, **7**:262
tadpole shrimp
number of species **3**:182*T*, **4**:371*T*
species duration **5**:434–435
Taenia solium **5**:664–665
Taeniopoda reticulate **3**:734*F*
Tagetes erecta **2**:391
Tagging of Pacific Predators (TOPP) **2**:18–19, **2**:19*F*
Taghanic event **5**:169–170
taiga **1**:262–263*T*, **2**:371*T*, **5**:534*F*, **6**:492*T*, **6**:499*T*
taimen **3**:464–465
Taiwan
animal species **1**:261*T*
botanical gardens **6**:519
ecosystems
mangroves **1**:275*T*
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical moist forest **1**:277*T*
vegetation types and coverage percentage **1**:264–265*T*
fish diversity **3**:473
plant species **1**:259*T*
urbanization **4**:560*F*
Taiwania **1**:272–273
Tajikistan **6**:519
Takla Makan **1**:270
Takla Makan–Gobi desert **1**:262–263*T*
Taklimakan Desert **2**:469–470
Talauma mexicana **2**:28–29
Talisia esculenta **4**:288*T*
tallgrass prairie **1**:706–708
tall tussock **4**:349
Tallula watsoni **5**:394
Talpa spp. **1**:710
Talpidae **4**:690–691, **6**:309*T*
Talun-kebun **7**:230–231
tamarack **7**:400
tamarins **1**:664*T*, **1**:666*T*, **6**:251*T*, **6**:256–258*T*, **7**:284
Tamarix aphylla **5**:484–485
Tamarix nilotica **5**:484–485
Tamarix ramosissima **1**:270, **2**:476–477, **6**:93–95
Tamarix spp. **2**:470, **3**:284, **4**:360
Tamiasciurus hudsonicus **2**:571
Tamiasciurus hudsonius **1**:631–632
Tamias striatus **1**:362
Tamshiyacu-Tahuayo Communal Reserve **5**:445
Tanacetum **1**:270
Tanacetum sibiricum **1**:269
Tanacetum sinaicum **5**:484–485
Tanaidacea
classifications **4**:381*T*
extinction vulnerability **2**:416–417*T*
marine environments **4**:382*F*
Tanaoceroidea **3**:733–734
Tana River **7**:146, **7**:151–152
Tana River mangabey **6**:256–258*T*
Tana River red colobus **6**:256–258*T*
tandem mass spectroscopy **5**:708
Tanganyika
ancient lakes species flocks **4**:463–464, **4**:463*T*
fish diversity **3**:472–473, **4**:463*T*
Lake Tanganyika **1**:54–55, **3**:560, **4**:463–464, **4**:463*T*, **4**:464*T*
tangle-veined flies **3**:489–491*T*
tank epiphytes **3**:523
tannia **6**:553
tannins **2**:426–427, **2**:426*T*
tanoaks **7**:118
Tansley, A. G. **3**:59–60, **3**:239
Tantulocarida **2**:401*T*, **2**:416–417*T*
Tanyderidae **3**:489–491*T*
Tanypezidae **3**:489–491*T*
Tanzania
botanical gardens **6**:519
deforestation **2**:449*T*, **2**:451*T*
energy efficiency measures **3**:254*F*
floodplain agriculture **7**:146
forest product exports **2**:451*T*
grass and shrub savannas **1**:67
indigenous peoples **4**:270*T*
primate diversity **6**:252*T*
primate endemism **6**:252*T*
Selous Game Reserve **5**:445
Serengeti ecosystem **3**:411
woodland savannas **1**:67–68
Taoism **6**:384, **6**:565
Tapera naevia **5**:501–502
taperebá **4**:286–287
tapeworms **4**:364, **5**:657–658, **5**:659*T*, **7**:439–440*T*, **7**:455–457
Taphrocerus **1**:329
Tapinanthus bangwensis **4**:77
Tapinanthus globiferus **4**:77
Tapiridae **2**:583*F*, **6**:309*T*
tapirs **1**:517, **2**:590*T*, **3**:195, **4**:690*T*, **4**:695, **6**:309*T*, **6**:332–333, **7**:229*T*
Tapirus bairidi **7**:229*T*
Tapirus indicus **1**:517
Tapirus spp. **4**:724–725*T*, **6**:332–333

Tapirus terrestris **2**:587*F*
tarantulas **1**:205–206, **1**:206, **3**:223, **3**:225–226
Taraxacum officinale **3**:131–132
tarbush **1**:708
Tardigrada
  Antarctica **1**:184*T*
  arthropod radiation ecology **2**:399–400
  classifications and sampling methods **4**:381*T*
  fossil record **3**:684*F*
  global diversity **4**:385*T*
  marine ecosystems **5**:50*T*
  number of species **4**:371*T*, **5**:669–670*T*
  phylogenetic relationships **7**:434*F*
Taricha rivularis **5**:315–316
Tarim Basin **1**:270
taro **2**:663, **3**:32*F*, **6**:553
tarsiers **1**:516, **4**:692, **6**:251*T*, **6**:256–258*T*
Tarsiidae **6**:250–252, **6**:251*T*
Tarsius bancanus **1**:516
Tarsius bancanus natunensis **6**:256–258*T*
Tarsius pelengensis **6**:256–258*T*
Tarsius sangirensis **6**:256–258*T*
Tartarocreagis infernalis **1**:213*F*
Tasmania **2**:371*T*, **4**:377
Tasmanian devil **4**:687*T*, **4**:688
Tasmanian wolf/tiger **4**:687–688
Tattersall's sifaka **6**:256–258*T*
Tatu estuary **1**:276*T*
Taurotragus oryx **1**:50, **6**:572–573
Tavarata mouse lemurs **6**:256–258*T*
tawheed **6**:564–565
Taxodium **7**:397
taxol **1**:594, **2**:432, **5**:709
taxonomic diversity
  ecosystem function measurements **3**:84
  functional diversity (FD) **3**:84
  measurement and analysis **1**:403, **3**:84, **6**:534, **6**:534
  taxonomic versus functional diversity **2:648–656**
    diversity estimations
      background information **2**:650
      species evenness **2**:650, **2**:651*T*
      species richness **2**:650, **2**:651*T*
      species richness–species evenness relationship **2**:650–651
    ecological coherence **2**:655–656
    historical perspective **2**:648
    linkages
      basic processes **2**:654–655
      C3 and C4 grasses **2**:654–655
      exotic species invasions **2**:655
      island biogeography **2**:652*F*, **2**:655, **2**:655*T*
    natural selection **2**:656
    research approaches **2**:648, **2**:649*T*
    species classifications
      characteristics **2**:649–650
      ecological niche model **2**:649–650, **2**:649*F*
      functional groupings **2**:649*F*, **2**:650
    stability dynamics
      basic concepts **2**:652–653
      local asymptotic stability **2**:652–653
      process-based models **2**:654
      species interaction models **2**:653–654
      transient dynamic behavior **2**:653, **2**:653*F*
    steady-state dynamics
      basic concepts **2**:651–652, **2**:651*F*
      colonization rates **2**:651–652, **2**:652*F*
      community structure **2**:652
      ecosystem processes **2**:652, **2**:652*F*
      extinction rates **2**:651–652, **2**:652*F*
      number of species **2**:651–652, **2**:652*F*
      population densities **2**:652
      species composition **2**:652
    systems theory
      basic concepts **2**:651–652
      stability dynamics **2**:652–653
      steady-state dynamics **2**:651–652, **2**:651*F*
taxonomy **7:97–111**
  alternative species concepts **7**:93, **7**:93*T*
  attributes **7**:87–89
  basic concepts **7**:85
  biodiversity perspectives **7**:90–91, **7**:91*F*
  characters and attributes **7**:98–99, **7**:105*T*
  cladistics **7**:87
  comparative methods **7**:86
  current research approaches **7**:109
  endemism **3**:234–235
  evidentiary research **7**:87–89
  evolutionary classification
    basic concepts **7**:106–109, **7**:107*F*
    cladistic classification **7**:108, **7**:108*F*, **7**:109*T*
    phenetic classification **7**:107, **7**:107*F*
  evolutionary theory **7**:99–100
  functional role **7**:98
  genetic relatedness **7**:99–100
  Hennigian terminology **7**:91–93
  historical perspective **7**:86, **7**:89–90
  importance
    biomimicry and adaptations **7**:95
    communication tools **7**:94–95
    conceptual framework **7**:94
    environmental baseline **7**:95
    evolutionary record **7**:95
    general discussion **7**:94
    information storage and retrieval **7**:95
  individual organisms and life cycles **7**:98
  Linnean nomenclature **7**:90
  museum research programs **5**:412
  operational procedures **7**:106
  philosophical perspectives
    cladistics **7**:104–106, **7**:104*T*, **7**:105*F*, **7**:105*T*
    Darwinism **7**:101–102
    empirical methods **7**:100–101
    essentialism **7**:100
    evolutionary systematics **7**:101–102, **7**:102*F*
    folk taxonomies **1**:121, **7**:100–101
    general discussion **7**:100
    idealism **7**:100
    inductive methods **7**:100–101
    nominalism **7**:101
    operational taxonomic units (OTUs) **5**:252–254, **7**:102–104, **7**:103*F*, **7**:103*T*, **7**:104*T*
    phenetic methods **6**:16–17, **7**:102–104, **7**:103*F*, **7**:107*F*
    phylogenetic systematics **7**:86, **7**:104–106
  phylogenetic relationships **7**:91, **7**:91*F*, **7**:99–100
  predictive classifications **7**:91–93, **7**:92*F*
  research collections **7**:86
  research summary **7**:109–110
  special versus general classifications **7**:99–100
  species concepts **6:679–691**
    alternative species concepts
      autapomorphic species concept **7**:93*T*
      biological species concept **7**:93*T*
      ecological species concept **6**:682, **7**:93*T*
      Hennigian species concept **7**:93*T*
      phylogenetic species concept **7**:93*T*
      recognition species concept **6**:682–683
      sexual selection **6**:682
      systematic classifications **7**:93, **7**:93*T*
    basic concepts **6**:680
    biases **6**:680
    biodiversity conservation
      ecological importance **6**:689
      genetic differentiation **6**:689
      spatiotemporal considerations **6**:689–690
      species as real entities **6**:689
    combined species concepts
      cohesion species concept **6**:685
      evolutionary and lineage concepts **6**:685
    Darwinian theory
      morphological species criterion **6**:681
      polytypic species criterion **6**:681
    dissenting perspectives
      evolutionary unit approach **6**:686
      general discussion **6**:685–686
      genotypic cluster criterion and assignment tests **6**:686–688, **6**:687*F*
      phenetic species concept **6**:686
      race versus species theory **6**:688
      taxonomic practice **6**:685–686
    historical perspective
      Bayesian species delimitation **6**:684
      diagnostic/phylogenetic species concept **6**:683–684, **7**:91–93, **7**:92*F*
      genealogy **6**:684
      monophyly **6**:683, **7**:91–93, **7**:92*F*
    interbreeding theory **6**:681–682
  subspecies taxonomy **7:45–48**
    historical background

binominal versus trinominal nomenclature **7**:45–46
divergence theories **7**:46
levels of variation **7**:45
modern perspectives
phylogenetic species concept **7**:46–47
semispecies taxonomy **7**:46–47
systematization **7**:97, **7**:101–102, **7**:102*F*
taxon inventories **7**:86
taxon scholars **7**:86
tokogenetic relationships **7**:91, **7**:91*F*, **7**:99
see also systematics
Taxus **1**:272–273, **5**:618*F*
Taxus baccata **1**:612, **6**:135
Taxus brevifolia **1**:594, **2**:432, **5**:709, **6**:57, **6**:135, **6**:135*F*
Tayabas Bay **1**:276*T*
Tayassuidae **2**:583*F*, **4**:696, **4**:727*T*, **6**:309*T*
Tayassu spp. **6**:332–333
Tayassu tajacu **2**:587*F*
Taylor's power law **6**:160–161, **6**:163–164, **6**:167
Tbilisi Declaration **3**:136
tea **2**:663, **3**:130*T*, **3**:131*T*, **7**:274–275
teak **1**:266, **6**:226, **7**:273–274
teal **3**:365–366
tea tree **2**:383
Tectona grandis **1**:69–70, **6**:226, **7**:273–274
tectonic activity **7**:328
Tectura testudinalis **3**:613*F*
teff **2**:663, **3**:32*F*
Tegeticula yuccasella **1**:709–710, **2**:146
tegu lizards **6**:411
Teius teyou **3**:115
Teleoceras **5**:224*T*
Teleogryllus commodus **3**:724
teleological theory **2**:422
Teleostei/teleosts
Acanthopterygii
Atherinomorpha **3**:467
Mugilomorpha **3**:467
Percomorpha **3**:467–471
characteristics **3**:463, **6**:522–523
clupeomorphs **3**:463–464
elopomorphs **3**:463
Euteleostei
Ateleopodomorpha **3**:465
Cyclosquamata **3**:465–466
Lampriomorpha **3**:466, **3**:466*F*
Paracanthopterygii **3**:466–467
Polymixiomorpha **3**:466
Protacanthopterygii **3**:464–465, **3**:465*F*
Scopelomorpha **3**:466
Stenopterygii **3**:465
ostariophysans **3**:463–464
osteoglossimorphs **3**:463, **3**:463*F*
telescopefish **3**:465–466
Telescopium **5**:17–18, **5**:17*F*
Telespyza cantans **1**:29*F*
television-watching guidelines **7**:248
Tellervinae **1**:654
telomere end-binding protein α (TEBPα) **5**:569–570, **5**:570*F*
telsontails **4**:392*T*
temephos **5**:693–694
Temminck's red colobus **6**:256–258*T*
Temnocephala **7**:442–443*F*
Temnocephalida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
Tempasuk Plain **1**:276*T*
tempeh **3**:638–639
temperate America **2**:448*T*
temperate Asia **2**:448*T*
temperate/cold-desert shrublands **6**:333–334
temperate deciduous forests
characteristics **1**:271
China **1**:271–272
Europe **3**:335*T*, **3**:336–340*T*
Japan **1**:271
Korea **1**:271
Manchuria **1**:271–272
mountain regions **1**:272–273
pine forests **1**:272
regional species richness controls **2**:602
temperate forests **7:112–127**
biodiversity change scenarios **4**:6*F*
biodiversity trends **7**:119–121, **7**:119*T*, **7**:120*F*
canopy architecture **3**:511–512
Central America **2**:29–30
characterization
dominant floristic features **7**:117–119
physiognomic features **7**:117
conservation
conservation status **7**:116*T*, **7**:125–126
future research areas **7**:127
major threats **7**:126–127
protection strategies **7**:126
conservation status **3**:171–173*T*
current global land cover **4**:557*F*
deforestation **1**:677*F*, **2**:447–448, **2**:448*T*
distribution hypotheses
general discussion **7**:121–124
historical/geological perspectives **7**:122–124, **7**:123*F*
productivity **7**:124
spatiotemporal heterogeneity **7**:124
ecosystem function measurements **3**:77*F*
ecosystem services **1**:334*T*, **7**:124–125
endemism **7**:121
Europe **3**:335*T*, **3**:336–340*T*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
global distribution patterns **7**:112–117, **7**:113*F*, **7**:115*F*, **7**:116*T*, **7**:117*T*
importance **7**:112
litter deposition **6**:120*T*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
predation effects **6**:236–237, **6**:238*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
range size rarity **7**:121
soils **6**:123*T*
South America **3**:111*F*, **3**:117, **6**:600–604*T*
termites **4**:432*T*
temperate grasslands **1**:262–263*T*, **1**:269
temporal match–mismatch (TMM) hypothesis **2**:76
temporal niches **7**:7
temporal stability **7**:5–7, **7**:6*F*, **7**:7*F*
temporary parasites **5**:657
temporary winter rain pools **5**:473, **5**:476*F*
tendu **6**:226
Tenebrionidae **1**:327*T*, **1**:328, **3**:513
teniposide **5**:712*T*
Tennessee River system **5**:562–563
Tenrecidae **4**:690
tenrecs **4**:690, **4**:690*T*
Tentaculata **5**:669–670*T*
Tenthredinidae **4**:182–183
Tenthredinoidea **4**:178
Tenuiros tritermes **4**:409*T*
tepary beans **2**:663–664
Tephritidae **3**:489–491*T*, **3**:492, **3**:493–494, **5**:460*T*
Tephritoidea **3**:489–491*T*
Tephrosia patulaial **1**:589
Teratomyzidae **3**:489–491*T*
Terebellida **7**:417*F*
Terebellidae **7**:417*F*
Terebralia **5**:17–18, **5**:17*F*
Terebralia palustris **5**:17–18
Teredo **5**:14–15
Teritiary period
brain size evolution **4**:698
Cretaceous–Tertiary (K–T) mass extinction **1**:494, **2**:516, **3**:418–421, **5**:6–7
faunal evolution **5**:7–8
fossil record **1**:411*F*, **1**:412*F*
mass extinctions **1**:413*F*, **4**:657*F*
megaherbivores **5**:224*T*
South America **3**:107, **3**:109
Termes **4**:409*T*
Termes hospes **4**:429*F*
Terminalia **1**:69–70
Terminalia chebula **6**:225
Terminalia tomentosa **1**:266
terminal restriction fragment length polymorphism (T-RFLP) **5**:254
termites **4:408–433**
characteristics **6**:588
developmental pathways **4**:412–413, **4**:413*F*
ecosystems
functional role **4**:429–431, **4**:430*T*
human disturbances **4**:432–433, **4**:432*T*
Near East **5**:460*T*
soils **4**:431–432, **4**:431*F*
terrestrial ecosystems **4**:429
families and genera **4**:409*T*
fecundity **4**:421–422
feeding behaviors **4**:413–414, **4**:422*F*
fossil record **4**:410–411
global distribution **4**:408–409, **4**:409*T*, **4**:411–412, **4**:411*F*, **4**:422*F*
global diversification

termites (*continued*)
dispersal ability **4**:426
evolutionary radiation **4**:426–427
higher termite radiation **4**:428–429
major historical events **4**:425–426
one-piece type diversification **4**:427–428, **4**:427*F*
separate type expansion **4**:427–428
life types **4**:414, **4**:422*F*, **4**:423*T*
mangrove ecosystems **5**:13
microbial assemblages **5**:283*F*
mound architecture **1**:590–591
Near East ecosystems **5**:464*T*
nests **4**:414–415, **4**:415*F*, **4**:422*F*
number of species **4**:392*T*, **4**:410*T*, **4**:422*F*
phylogeny **4**:410–411, **4**:410*F*, **4**:417*F*
prey–predator interactions **4**:413, **4**:415*F*
social life
caste differentiation **4**:419–420, **4**:420*F*, **4**:421*F*
characterization **4**:422*F*
eusociality **4**:412–413, **4**:412*F*, **4**:414*T*, **4**:421–422
evolutionary development **4**:424*T*
foraging behaviors **4**:422–424, **4**:425*F*
fungus-growing termites **4**:424–425, **4**:426*F*
life types **4**:423*T*
social diversity **4**:421–422
workers and soldiers **4**:420–421, **4**:421*F*, **4**:423*T*, **4**:424*T*
social systems **2**:646
soil fauna **6**:588, **6**:590*T*
symbiotic relationships
biochemical processes **4**:418*F*
carbon-nitrogen balance symbiosis **4**:416–417, **4**:417–418, **4**:417*F*
cellulose decomposition symbiosis **4**:417
digestive tube **4**:415–416, **4**:416*F*
fungi **4**:418–419, **4**:419*F*
general discussion **4**:408–409
hindgut microenvironment **4**:416*F*
phylogeny **4**:417*F*
symbionts **4**:422*F*
wood feeders versus soil feeders **4**:429, **4**:429*F*
Termitidae **4**:409*T*, **4**:410*T*
Termitimyces **4**:429
Termitinae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:417*F*, **4**:422*F*, **4**:430*T*
Termitogeton **4**:409*T*
Termitomyces **4**:418–419, **4**:419*F*
Termopsidae **4**:409*T*, **4**:410–411, **4**:410*F*, **4**:410*T*, **4**:422*F*, **4**:427–428, **4**:427*F*
terns **1**:710, **7**:390
Teropidae **4**:691–692
terpenes **2**:426*T*, **2**:427
Terpsiphone corvina **7**:184
terraces **4**:572, **4**:572*F*
Terranova **7**:429–430
terrapins **7**:390
terra-preta-do-índio **4**:285–287, **4**:286*F*
terrestrial arthropods
Amazon ecosystems **1:245–253**
biotopes **1**:249–250
future research areas **1**:252
genetic characteristics **1**:251–252
habitat specificity **1**:249–250, **1**:250*T*, **1**:251*F*, **1**:252*F*
morphological characteristics **1**:250–251
number of species **1**:245, **1**:246*T*
origins and geographical distribution
Andes **1**:247
Carabidae **1**:247, **1**:250*T*
Geophilomorpha **1**:247
Meinertellidae **1**:247
neotropics **1**:246–247
nonflooded and/or floodplain forests **1**:247
Polydesmida **1**:247
Scolopendromorpha **1**:246–247
Scorpiones **1**:247
treeless floodplains **1**:247
species richness **1**:247–249, **1**:248*F*
taxonomy **1**:245, **1**:246*T*
terrestrial ecosystems
Arctic ecosystems **1:227–244**
biodiversity filters **1**:238
biodiversity patterns **1**:231–235, **1**:231*T*
circumpolar diversity **1**:235–236
current biodiversity status **1**:242
environmental characteristics **1**:227–231, **1**:229*F*
environmental consequences **1**:238–240, **1**:239*F*, **1**:240*F*
general discussion **1**:227
latitudinal biodiversity patterns **1**:232–235, **1**:233–234*T*, **1**:235*F*, **1**:236*F*
natural resources/environmental exploitation **1**:240–241
research organizations **1**:242–243
research summary **1**:243
threats and challenges **1**:241–242
topographic impacts **1**:236–238, **1**:237*F*
vegetation zones **1**:229*F*, **1**:230*T*
biological impoverishment **6**:335–337, **6**:336*F*
conservation status **3**:170–174, **3**:171–173*T*
disturbance ecology **2:562–567**
anthropogenic impacts **2**:566–567
basic concepts **2**:562–563
carnivores **2**:565–566
herbivores **2**:565–566
interaction impacts **2**:566
intermediate disturbance hypothesis **2**:563
mechanical disturbances **2**:563–564
physico–chemical disturbances **2**:564–565, **2**:565*F*
ecosystem function measurements **3:72–89**
analytical and measurement approaches
biomass production **3**:73–76, **3**:74–75*T*, **3**:76*F*, **3**:77*F*, **3**:78*F*, **3**:79*F*
carbon assimilation **3**:75*F*, **3**:77*F*
community structure approaches **3**:72–73
decomposition processes **3**:80–81
long-term, large-scale ecosystem processes **3**:73, **3**:74–75*T*, **3**:76*F*
net primary production (NPP) **3**:73–76, **3**:74–75*T*, **3**:76*F*, **3**:78*F*, **3**:79*F*
nutrient cycling **3**:74–75*T*, **3**:78
short-term resouce dynamics **3**:73, **3**:74–75*T*, **3**:76*F*
trophic transfer **3**:74–75*T*, **3**:76–77
water dynamics **3**:74–75*T*, **3**:82–83
whole-ecosystem approaches **3**:72–73
decomposition processes
detrital food web composition **3**:82, **3**:82*T*
importance **3**:80–81, **3**:82*T*
litter deposition **3**:82, **3**:82*T*
mineralization processes **3**:80–81
nutrient immobilization **3**:80–81
field manipulations **3**:86*F*, **3**:87*F*, **3**:88
functional diversity (FD)
biodiversity **3**:84–86, **3**:84*F*
plant functional traits **3**:85*F*, **3**:85*T*, **3**:86
positive-feedback switches **3**:86–87, **3**:86*F*
taxonomic diversity **3**:84
herbivores
food chain controls **3**:78, **3**:80*F*
nutrient cycling **3**:77–78
trophic transfer **3**:76–77
monitoring networks **3**:86*F*, **3**:87*F*, **3**:88
natural versus synthesized ecosystems **3**:88
nutrient cycling
Amazonian forests **3**:81*T*
control mechanisms **3**:78
herbivores **3**:77–78
India **3**:82*T*
land use impacts **3**:79–80
nutrient capture, retention, and release **3**:74–75*T*, **3**:78
open versus tight nutrient cycles **3**:78–79
rainfall effects **3**:79, **3**:81*F*
soil type effects **3**:79–80
vegetation structure **3**:79–80
purpose **3**:72
real ecosystems **3**:88
resistance and resilience strategies **3**:83–84, **3**:83*F*
resource dynamics **3**:72–73, **3**:74–75*T*, **3**:75*F*
synthesized ecosystems **3**:86*F*, **3**:87*F*, **3**:88
trophic transfer
basic concepts **3**:74–75*T*, **3**:76–77
bottom-up and top-down food chain controls **3**:78, **3**:80*F*
herbivore performance **3**:76–77
nutrient availability and quality **3**:76–77
plant production **3**:76–77

ecosystem services **5**:**347–361**
biophysical measures **5**:349
landowner decisions **5**:347–349
land use/land cover (LULC) patterns **5**:348–349, **5**:351–352
modeling frameworks
animal pollination **5**:354–355
basic principles **5**:349–350
carbon storage and sequestration **5**:351–352
consumable goods **5**:349–350
cost-effectiveness measures **5**:356–357, **5**:357*F*
crop production **5**:349–350
cultural ecosystem services **5**:353
decision-making strategies **5**:358
esthetic services **5**:353–354
nature-based recreation models **5**:353
nutrient retention **5**:355–356
provisioning services **5**:350–351
regulatory services **5**:351–352
software **5**:357–358
spiritual services **5**:354
storm peak management **5**:352–353
summary discussion **5**:358–359
supporting services **5**:354–355
uncertainty monitoring **5**:356
supply–demand–value relationship **5**:348–349
global change effects **5**:257–258
net primary production (NPP) **1**:559, **1**:560*T*
nitrogen deposition **5**:**519–536**
anthropogenic impacts **5**:520–522, **5**:522*F*
biodiversity impacts
acidification **5**:522, **5**:527*F*
direct foliar toxicity **5**:523–525, **5**:527*F*
eutrophication **5**:522
general discussion **5**:522
hotspots **5**:526*T*
primary drivers **5**:522, **5**:525*F*, **5**:527*F*
secondary stressors **5**:522–523, **5**:527*F*
biological nitrogen fixation **5**:520*F*
environmental impact **5**:519–520
global distribution **5**:524*F*
global nitrogen budget **5**:523*F*
global nitrogen cycle **5**:520–522, **5**:521*F*
influencing factors
ecosystem-specific effects **5**:526–528
taxa-specific responses **5**:528
nitrogen as nutrient **5**:520, **5**:520*F*
research methodologies **5**:525–526, **5**:528*T*, **5**:529*F*
taxa-specific responses
herbaceous plants **5**:528–530
higher-taxa responses **5**:530
nonvascular plants **5**:528
soil biota **5**:530
trees **5**:530
nonextremeophilic Archaea **1**:224*T*
protected area systems
biogeographic realms **6**:491
biome realms **6**:491–492, **6**:492*T*, **6**:493*F*
data challenges **6**:491
deserts **6**:492, **6**:495*T*
ecoregion protection **2**:591–592, **6**:492, **6**:493*F*
forests **6**:492, **6**:494*F*, **6**:494*T*
islands **6**:493–494
mountains **6**:492–493
World Wildlife Fund (WWF) realms **6**:492*T*
subterranean ecosystems
Gua Salukkan Kallang-Towakkalak (Indonesia) **7**:59–60
Kazumura Cave (Hawaii, USA) **7**:60
Mammoth Cave (Kentucky, USA) **7**:59*T*, **7**:60
Organ Cave (West Virginia, USA) **7**:60
Vjetrenica (Bosnia and Herzegovina) **7**:60
termites **4**:429
trophic cascades **7**:259–260
terrestrial invertebrates **4**:**392–394**
Antarctica **1**:184*T*
conservation practices
filter approach **3**:224
recovery programs **3**:225
surrogate use **3**:224–225
umbrella/flagship species **3**:225
endangered species **3**:**219–227**
centipedes **3**:226
conservation practices **3**:224
crabs **3**:226
daddy long legs **3**:225–226
earthworms **3**:225
insects **3**:226
land snails **3**:226
millipedes **3**:226
pill bugs **3**:226
pseudoscorpions **3**:225–226
Red List Index (RLI) **3**:225
spiders **3**:225–226
status assessments **3**:223–224
velvet worms **3**:225, **3**:226*F*
woodlice **3**:226
endemism **5**:592*T*
environmental conditions **4**:393
life histories **3**:220
species numbers and diversity **3**:219
taxonomy
limited-adapted groups **4**:392–393, **4**:394*F*
taxonomic impediment **3**:219–220
well-adapted groups **4**:392, **4**:393*F*, **4**:394*F*
threats and challenges
climate change **3**:221
habitat loss and fragmentation **3**:220–221
invasive species **3**:222
overharvesting/overcollecting **3**:222–223
pest control **3**:221–222
terrestrial subterranean habitats **7**:50–51
terrestrial vertebrates **6**:709*F*
Tertiary period
boundary age **5**:1*T*, **5**:2
Cretaceous–Tertiary (K–T) mass extinction **1**:494, **2**:516, **3**:418–421, **3**:688
crustaceans **2**:406
fossil record **3**:685*F*
mass extinctions
extinction intensities **5**:162*F*
marine animals **1**:493*F*
notable mass extinctions **5**:161*F*
number of families **3**:419*F*, **3**:425*F*, **5**:157*F*
originations **1**:495*F*, **3**:425*F*
temperate forests **7**:122–124, **7**:123*F*
termite species **4**:410*T*
terrestrial plant communities
abundance patterns **6**:74*F*
Australia **1**:291–292
evolutionary development **6**:65–67, **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
Tessaria absinthoides **4**:105
tesselation analysis **4**:511
testacea **6**:586, **6**:587*F*
testate amebae **6**:290–291, **6**:290*F*
see also Protozoa
Testudines **6**:400–401
Testudo kleinmanni **5**:472
Tethinidae **3**:489–491*T*
Tetrabothriidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetrabothrius **7**:456–457*F*
Tetracarpaea **5**:622*F*
Tetracentraceae **1**:260*T*
Tetrachondra patagonica **4**:106–107
Tetragnatha spiders **1**:32, **1**:33*F*
tetrahydrocannabinol **5**:712*T*
Tetrahymena **6**:295, **6**:296*F*
Tetrahymena thermophila **5**:568–569, **6**:355
tetrahymenids **6**:295–296, **6**:296*F*
Tetralophodon **5**:224*T*
Tetrameles **5**:624–625
Tetrameryx **4**:724–725*T*
Tetranychus urticae **4**:298–299, **4**:299*F*
Tetraodontiformes **3**:470–471
Tetraopes **1**:24*F*
Tetraopini **1**:24*F*
Tetrao tetrix **1**:631–632
Tetrao urogallus **1**:631–632
Tetraphyllidea **7**:439–440*T*, **7**:456–457*F*, **7**:459
Tetrapoda
characteristics **3**:461, **7**:338–339
classification systems **7**:337*F*
Phanerozoic diversity **3**:688
Tetrarthrosoma syriaca **3**:674*F*
Tetraspora **3**:358
Tetrasporopsis **3**:358
Tetrastigma **4**:71–72
Tettigoniidae
characteristics **3**:726
Near East ecosystems **5**:460*T*, **5**:464*T*
subfamilies
Acridoxeninae **3**:727
Bradyporinae **3**:726

Tettigoniidae (*continued*)
Conocephalinae **3**:727
Ephippigerinae **3**:727
Phaneropterinae **3**:726–727
Phasmodinae **3**:727
Pseudophyllinae **3**:727
Tettigoniinae **3**:727
Zaprochilinae **3**:727
Tettigoniinae **3**:727
Tettigonioidea (katydids)
characteristics **3**:722–723, **3**:725–726, **3**:725*F*
food and feeding habits **3**:724–725
paleontology **3**:725
phylogeny and taxonomy **3**:725
reproduction and growth **3**:724
sound production **3**:723–724
Tettigoniidae
characteristics **3**:726
subfamilies
Acridoxeninae **3**:727
Bradyporinae **3**:726
Conocephalinae **3**:727
Ephippigerinae **3**:727
Phaneropterinae **3**:726–727
Phasmodinae **3**:727
Pseudophyllinae **3**:727
Tettigoniinae **3**:727
Zaprochilinae **3**:727
Tevnia jerichonana **7**:328–329
Texas cougars **3**:197
Thailand
animal species **1**:261*T*
botanical gardens **6**:507*T*
cropland conversion **4**:665*T*
deforestation **4**:562–564, **4**:563*F*
ecosystems
genetic richness **1**:514*T*
mangroves **1**:275*T*, **1**:277, **1**:513*T*
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical dry forest **1**:266–267
tropical moist forest **1**:277*T*
vegetation types and coverage percentage **1**:264–265*T*
endangered mammals **3**:194*T*
fishing/fisheries **6**:422*T*
invasive species **4**:353–354
pastureland conversion **4**:665*T*
plant species **1**:259*T*
regulating services case study **3**:51–53, **3**:52*F*
rice production **1**:88
shrimp farming versus mangrove conservation **3**:51–53, **3**:52*F*
see also Southeast Asia
Thailand monsoon forest **1**:262–263*T*
Thais kiosquiformis **5**:17–18
Thais lapillus **6**:232
Thalassarche cauta **2**:543
Thalassia **6**:542*T*
Thalassia hemprichii **6**:541*T*
Thalassia testudinum **3**:360–361, **6**:541*T*
Thalassina anomala **5**:15–16
Thalassinidea **2**:416–417*T*
Thalassiosira leptoporous **6**:30
Thalassiosira nordenskioldi **6**:30
Thalassiosira weissflogii **5**:687*F*, **6**:30
Thalassisobates litoralis **5**:430–431
Thalassodendron **6**:542*T*
Thalassodendron ciliatum **6**:541*T*
Thalassodendron pachyrhizum **6**:541*T*
Thalassoma bifasciatum **6**:574
Thalia geniculata **1**:164
Thambetochen **4**:730–731
Thamnophilidae **6**:308*T*
Thamnophis sirtalis parietalus **5**:315–316
Thamnophis sirtalis tetrataenia **6**:411
Thanatus fornicatus **5**:463*T*
Thandamannar Lagoon **1**:276*T*
Thar desert **1**:262–263*T*
Thaumaleidae **3**:489–491*T*
Thaumetopoea pityocampa **4**:607
Theaceae **1**:267, **1**:270–271
Theba pisana **5**:466–467, **5**:468*F*
Theclinae **1**:651*F*, **1**:652
The Economics of Ecosystems and Biodiversity (TEEB) **1**:333–335, **1**:337–339, **2**:315
Thelepogon elegans **1**:69
Thelyphonidae **1**:211–212, **1**:246*T*
thelytoky **4**:179
Themada triandra **1**:67
thematic gardens **6**:505
ThematicMapper (TM) **6**:393, **6**:393*T*
Themeda quadrivalis **1**:69
Themeda triandra **1**:52, **1**:69, **3**:437–438
The Nature Conservancy (TNC) **2**:231–232, **4**:325–326, **7**:152–153, **7**:253, **7**:355
Theobroma bicolor **4**:289*T*
Theobroma cacao **2**:383, **2**:663–664, **4**:287–290, **4**:288*T*, **5**:605, **6**:554
Theobroma grandiflorum **4**:279–280, **4**:288*T*
Theobroma speciosum **4**:287–290, **4**:288*T*
theobromine **2**:432
Theocritus **4**:650
Theodoxus jordani **5**:470–471
Theophrastus **3**:320–321, **5**:705
theophylline **5**:712*T*
theory of evolution **3:392–400**
adaptive processes **1**:16–17
Darwin's five theories
common ancestry theory **3**:395, **3**:397*F*
evolutionary theory **3**:395–396
gradualism **3**:395–396
natural selection **3**:396
populational change **3**:396
evolutionary ecology **3**:7
evolutionary mechanisms
genetic drift **3**:398–399
migration **3**:398
mutations **3**:398
natural selection **3**:399
speciation **3**:399
variation **3**:396–398
historical perspective
ancestor–descendant relationships **3**:393*F*
Darwinian theory **1**:15–16, **3**:393–394
evolutionary developmental biology **3**:395, **3**:396*F*
modern synthesis **3**:394
molecular evolution **3**:394–395
postsynthesis theories **3**:394–395
pre-Darwinian theories **3**:393, **3**:393*F*
Proterozoic–Cambrian transition **3**:396*F*
organic evolution **3**:392–393
scientific theory **3**:392
theory of inclusive wealth **7**:73, **7**:75–76
Theragra chalcogramma **5**:58–60
Theraphosa leblondi **1**:205–206
Theraphosidae **1**:206
Therapsids **4**:684, **5**:4–5
Therevidae **3**:489–491*T*
Theria **4**:682–683, **7**:341
Theridiidae **1**:207
Therioaphis trifolii **4**:362
Theriodictis **4**:727*T*
Theritas nr. *hemon* **1**:651*F*
Thermacarus **4**:118
thermal ecosystems **4:111–120**
biodiversity **4**:118–119
characteristics and definitions **4**:111–112
extreme environments **4**:111
geothermally heated systems **4**:112–113
research summary **4**:119–120
temperature tolerance ranges **4**:113–115
thermophilic evolution **4**:115–118
thermally opportunistic organisms **4**:111–112
thermally tolerant organisms **4**:111–112
Thermococcus **1**:219–220, **1**:221*F*, **1**:221*T*
thermodynamics **5**:629
Thermofilum **1**:219–220, **1**:221*T*
Thermofilum pendens **1**:315
thermophiles
biodiversity **4**:118–119
characteristics **7**:161
characteristics and definitions **4**:111–112
ecological niches **7**:161–163, **7**:162*F*
evolutionary development **4**:115–118
hydrothermal vent ecosystems **7**:328
metabolic diversity **7**:161–163
molecular mechanisms **7**:163
origins **7:161–166**
early Earth environment **7**:163–164
fossil record **7**:164
general discussion **7**:163–164
molecular fossils
genetic annealing model **7**:165
lateral gene transfer **7**:165
molecular phylogenies **7**:162*F*, **7**:164–165
small subunit ribosomal RNA (ss rRNA) **7**:162*F*, **7**:164–165
origin of life theories **7**:165–166
phylogenetic relationships **7**:161, **7**:162*F*
research summary **4**:119–120, **7**:166
temperature tolerance ranges **4**:113–115
thermophilic Archaea **1**:219–220, **1**:221*F*, **1**:221*T*
Thermoplasma **1**:220, **1**:221*T*, **4**:112
Thermoplasma acidophilum **1**:314
Thermoproteus **1**:219–220, **1**:221*T*
Thermosbaenacea **2**:416–417*T*
Thermotoga maritima **7**:163

Thermus aquaticus **1**:315, **1**:593, **2**:640–641, **4**:116–117, **5**:709–710
Theropoda **2**:520*T*
Thescelosaurus **2**:519*T*, **2**:520*T*
thiabendazole **5**:695
thick-billed parrot **7**:183
thick-headed flies **3**:489–491*T*
thick-tailed galagos **6**:251*T*
Thiobacillus **6**:32
Thiobacillus ferrooxidans **5**:308, **6**:448–449
thiobencarb **5**:696
thiocarbamates **4**:90*T*, **4**:92*T*, **5**:696
Thiomicrospira **4**:391
thiophantes **5**:695
Thioploca **2**:12*F*
Thirakhupt's bent-toed gecko **1**:515–516
thiram **5**:695, **5**:699*T*
Thisbe irenea **1**:652*F*
thistles **4**:361, **4**:363
Thlaspic aerulescens **6**:96
Thlaspietea rotundifolii **3**:340
Thomas' camel **5**:364–365
Thomas' rope squirrels **5**:603
Thomomys umbrinusemotus **7**:229*T*
Thomson's gazelles **5**:313
Thoracica **2**:416–417*T*
Thoracocarpus **3**:522*T*
Thoracostoma **7**:423*F*
Thoracotermes **4**:409*T*
Thoreau, Henry David **2**:280, **3**:271, **4**:649–650, **4**:652
thorny smilax **2**:543
thrashers **4**:64
Thraupidae **6**:308*T*
thread worms **4**:381*T*
Threatened Plants Committee (TPC) **6**:78
threatened species lists and assessments **7:191–211**
  Alliance for Zero Extinction (AZE) **6**:488, **7**:204, **7**:204*F*
  challenges **7**:208–209
  Evolutionarily Distinct and Globally Endangered (EDGE) species **7**:202–203
  future outlook **7**:208–209
  historical perspective **7**:192
  importance **7**:191–192
  Important Bird Areas (IBAs) **7**:203–204, **7**:204*F*
  Important Plant Areas (IPAs) **7**:203–204
  Key Biodiversity Areas (KBAs) **7**:203–204, **7**:204*F*
  research applications
    biodiversity status determination **7**:200, **7**:201*T*, **7**:202*F*
    conservation action support **7**:200–203, **7**:203*F*
    environmental stewardship **7**:207–208
    general discussion **7**:200
    indicator species development **7**:205–207, **7**:206–207*F*
    major threats **7**:203*F*
    policy-based actions **7**:204–205
    protected area systems **7**:203–204
    resource allocations **7**:207
    scientific research guidance **7**:200–203, **7**:203*F*
  risk assessments
    classification systems **7**:194–195
    Endangered Species Act (ESA) **7**:199–200
    geographic distribution **7**:198*F*, **7**:202*F*
    global species extinction risk assessments **7**:194–195
    IUCN Red List categories and criteria **7**:196–197, **7**:197
    NatureServe Conservation Status methodology **7**:199
    regional and national extinction risk assessments **7**:195–197
    vulnerable species **7**:192–194, **7**:193
Three Gorges Dam **3**:260
three-spine sticklebacks
  adaptive radiation **1**:32, **1**:32*F*, **2**:725
  gene flow barriers **4**:517
  genomic research **2**:275
three-toed sloth **2**:590*T*
Threskiornithidae **6**:308*T*
thrips **4**:298–299, **4**:392*T*, **6**:141*T*
Thuja
  Asia **1**:272–273
  boreal forests **1**:628–630
Thuja standishii **1**:273
Thujopsis **1**:272–273
Thujopsis dolobrata **1**:273
Thunnus obesus **3**:478*F*, **5**:35
Thunnus orientalis **2**:7–8, **2**:9*F*
Thunnus thynnus **2**:7–8, **2**:10*F*
Thylacinidae **1**:686, **4**:687–688, **4**:687*T*
Thylacinus cynocephalus **4**:687–688
Thylacoleo **4**:727–728
Thylacoleo carnifex **5**:363–364, **5**:364*F*
Thylacomyidae **4**:688
Thymallinae **6**:522–523
Thymeliaceae **1**:263
Thyridia **2**:138
Thyridoidea **5**:387–388*T*
Thyropteridae **6**:309*T*
Thyrsopteridaceae **5**:216–217
Thysanoessa macrura **4**:386*T*
Thysanoptera **4**:308–312, **4**:309*T*, **4**:392*T*
Thysanura **4**:308–312, **4**:309*T*, **4**:392*T*
Tibet
  C_4 plants **2**:371*T*
  domesticated ruminants **1**:702–703
  marshes **1**:278–279
  temperate semi-deserts and deserts **1**:269–270
  Tibetan Plateau **1**:270, **1**:272
  wheat production **1**:87–88
Tibetan antelope **3**:373
Tibetan desert **1**:262–263*T*
ticks **1:202–217**
  characteristics **1**:213–215, **1**:214*F*, **2**:524*T*, **4**:392
  climate change effects **2**:103–104
  as disease vectors **1**:361–362, **1**:363*F*, **4**:505
  domesticated ruminants **1**:710–711
  ecological characteristics **1**:214–215
  genera and species **1**:202*T*
  insecticide resistance **4**:298–299
  natural history museum research programs **5**:415
  phylogeny and taxonomy **1**:205*F*, **1**:215–216
  reproduction and dispersal **1**:215
tidal freshwater marsh **7**:369*T*, **7**:370*T*
tidal salt marshes **7**:370*T*
Tidestromia **2**:376*F*
Tidestromia lanuginosa **2**:373*T*
tiger beetles **1**:250, **1**:250*T*, **1**:251*F*, **1**:252*F*, **1**:326–328
tiger flatheads **2**:8–9
tiger moths **5**:396–397, **5**:396*F*, **5**:397*F*
tiger pear **4**:350–351
tiger quoll **4**:687*T*, **4**:688, **4**:704
tigers
  adipose tissue **7**:34*F*
  conservation practices **7**:229*T*
  mangrove ecosystems **5**:14
  oil-palm plantations **5**:603, **5**:606
  prey–predator interactions **1**:709
  species diversity **1**:686, **4**:693
  suitable habitats **3**:190
tiger salamanders
  acid rain **3**:166
  disease **3**:166
Tigridia acesta **1**:652*F*
Tiktaalik roseae **3**:461
tilapia **1**:190*T*, **1**:195
Tilapia zilli **5**:470, **5**:471*F*
Tilia amurensis **1**:271–272
Tilia japonica **1**:271
Tilia mandschurica **1**:271–272
Tilia mongolica **1**:269
Tiliqua rugosa **7**:32
tillage **6**:556–557, **6**:593–594
Tillandsia usneoides **3**:517
Tilman biomass study **1**:386–388, **1**:387*F*
timber industry **7:212–221**
  boreal forests
    disturbance dynamics **7**:214–215, **7**:214*T*
    Europe **7**:212–213
    managed versus natural forests
      beetle species **7**:218–219, **7**:219*T*, **7**:220*F*, **7**:220*T*
      forest characteristics **7**:216–217, **7**:217*F*
      impact mechanisms **7**:215–216, **7**:217*F*, **7**:217*T*
      landscape patterns **7**:217–218, **7**:217*F*, **7**:218*F*, **7**:219*F*
    management strategies **7**:213, **7**:213*F*, **7**:214*T*
    production optimization **7**:215, **7**:216*F*
    research summary **7**:219–220
    structure and function **7**:213, **7**:213*F*
    successional dynamics **7**:213–214, **7**:214*F*
  deforestation **2**:451
  ecosystem services modeling **5**:350–351
  environmental impacts **7**:314–315
  Europe **7**:212*T*
  habitat degradation **4**:567
  India **6**:224–225

timber industry (*continued*)
landscape legacies **4**:525
land use issues **4**:558, **4**:567
logged and managed forests **1**:446–458
abiotic changes
forest floor hydrology **1**:450
insolation **1**:450
temperature **1**:450
wind patterns **1**:450
biodiversity impacts **7**:203*F*
biotic changes
aquatic fauna **1**:453
behavioral changes **1**:453
characteristics **1**:450–451
domestic animals **1**:453
fungi **1**:451
invertebrates **1**:451–452
mycorrhizae **1**:451
nonforest species **1**:453
plants **1**:451
vertebrates **1**:452–453
ecological interactions **1**:453–454
emerging trends **1**:455–457
environmental impacts **7**:314–315
evolutionary processes **1**:454–455
genetic effects **1**:454–455
management systems
general discussion **1**:447
long-term intensive forest management practices **1**:448
managed versus secondary forests **1**:448
nontimber forest products (NTFPs) **1**:448
shifting agricultural practices **1**:448
timber management **1**:447
plants
early successional phases **1**:451
exotic species invasions **1**:451
forest management impact **1**:451
research background **1**:446–447
structural alterations
coarse woody debris (CWD) **1**:450
environmental impact **1**:448–449
forest canopies **1**:449
fragmentation **1**:449
gaps and edges **1**:449
soil structure **1**:450
spatial patterns **1**:449–450
vertical structure **1**:449
synergistic anthropogenic effects **1**:455
timber management
clearing systems **1**:447
regeneration systems **1**:447
salvage logging **1**:447–448
selective logging systems **1**:447
tropical forests **2**:379–381, **2**:380*F*
vertebrates
behavioral changes **1**:453
biotic changes **1**:452–453
birds **1**:453
domestic animals **1**:453
food resource loss or gain **1**:453
habitat loss or gain **1**:452–453
primates **1**:453
overharvesting/overcollecting **3**:283
poverty–conservation relationship **2**:232
time–series analysis **1**:698–699
Timor-Leste **1**:513*T*, **4**:560*F*, **6**:507*T*
see also Southeast Asia
Tinamidae **6**:308*T*
Tinamiformes **1**:607, **6**:308*T*
tin-based insecticides **5**:694
tinea capitis **3**:634*T*
tinea corporus **3**:634*T*
tinea nigra **3**:634*T*
Tineidae **5**:401–402
Tineoidea **5**:386*F*, **5**:387–388*T*, **5**:389–390
Tini Heke Islands **5**:444
Tintinnidium **6**:25–26, **6**:31
tintinnids **6**:295–296, **6**:296*F*
Tintinnopsis **6**:295, **6**:296*F*
tipping points **2**:309, **2**:309, **4**:213, **4**:214*F*
tippler's bane **3**:635*T*
Tipula carinitrons **1**:232
Tipulidae
characteristics **3**:492
classifications **3**:489–491*T*
life history diversity **3**:493–494
wetland ecosystems **7**:389–391
Tipulomorpha **3**:489–491*T*
Tirari-Stuart stony desert **2**:468
tiretrack eels **3**:467–471
Tischerioidea **5**:386, **5**:386*F*, **5**:387–388*T*
titanium (Ti)
elemental partitioning **1**:554*T*
relative abundance **1**:553*F*
Titanochaeta **1**:28–29
Titanotylopus **4**:724–725*T*
Titicaca (Lake) **4**:463–464, **4**:463*T*, **4**:464*T*
titi monkeys **2**:590*T*, **6**:251*T*, **6**:256–258*T*
titmice
food webs **7**:266*F*, **7**:267*T*
guilds **4**:33–35
tits **1**:611–612, **5**:720
Tityus spp. **1**:208
toadfish **3**:466–467
toads **3**:165
characteristics **7**:339
critical pH levels **6**:153*F*
geographic distribution **3**:691–693
introduced parasites–species decline relationships **2**:528*T*
South America **3**:115
wetland ecosystems **7**:389–391
toadstools **3**:634
tobacco **2**:384*F*, **2**:663–664
tobacco whitefly **4**:298–299, **4**:299*F*
Tockus leucomelas **5**:603
Togaviridae **5**:134*T*, **5**:139
Togo
botanical gardens **6**:519
deforestation **2**:449*T*
Tokelau **7**:228
Tolypocladium inflatum **3**:625–626, **3**:638
tomatoes **2**:383, **2**:391, **2**:663–664, **3**:32*F*, **6**:554
Tomistoma schlegeli **6**:401
tomtits **7**:189
Tonga **5**:592*F*, **6**:507*T*
Tongan megapode **7**:188
Tongariro National Park (New Zealand) **6**:486
Tonkin black-crested gibbon **6**:256–258*T*
Tonkin snub-nosed monkey **6**:250, **6**:256–258*T*
tooth-billed bowerbird **2**:75*F*
toothed whales
characteristics **4**:693–694, **5**:66–67
migratory patterns **5**:314
taxonomy **5**:65*T*
topi **4**:696
topminnows **3**:467
Topobea **3**:522*T*
Torgos tracheliotos **7**:185
Torgos tracheliotus **5**:474
Toronto, Canada **2**:707
Torontoceros **4**:724–725*T*
Torosaurus **2**:519*T*, **2**:520*T*
Torreya taxifolia **3**:215
Torricelliaceae **1**:260*T*
tortoises
characteristics **6**:400–401
conservation status **3**:216*T*
gigantism **6**:405
Tortricoidea **5**:386*F*, **5**:387–388*T*, **5**:390
Tortula andicola **4**:103
total organic carbon (TOC) **3**:351, **3**:352*F*
totemism **6**:563–564
toucans
beak shape **1**:604*F*
evolutionary convergence **1**:617
tourism industry **7**:**222–225**
anthropogenic impacts **7**:223–224
biodiversity impacts **7**:222
Chinese nature reserves **3**:375
conservation tourism **7**:224
ecotourism **1**:395–396, **7**:224
geographic distribution **7**:222–223
nature-based tourism **2**:232–233
poverty–conservation relationship **2**:232–233
primary attractions **7**:223
protected area systems **6**:500
research summary **7**:224–225
toxiferine **5**:712*T*
Toxodon **4**:727*T*
Toxodon platensis **5**:225*T*
Toxodontidae **4**:727*T*, **5**:225*T*
toxoplasma **5**:660*T*
Toxoplasma gondii **2**:528*T*, **5**:660, **6**:297
Toxostoma redivivum **4**:41*T*, **4**:64
Toynbee, Arnold **6**:382–383
trace gases **2**:680
trace metals **1**:563
trace pollutants **6**:155
Tracheloraphis **6**:295
Trachemys scripta **3**:218
tracheophytes **3**:686–687, **3**:687*F*, **4**:169
Trachinoids **3**:469
Trachurus trachurus **5**:689–690
Trachypithecus cristatus vigilans **6**:256–258*T*
Trachypithecus delacouri **6**:256–258*T*
Trachypithecus francoisi **6**:256–258*T*
Trachypithecus geei **6**:256–258*T*
Trachypithecus germaini **6**:256–258*T*
Trachypithecus hatinhensis **6**:256–258*T*

Trachypithecus johnii **6**:253–254
Trachypithecus phayrei **6**:253–254
Trachypithecus phayrei crepusculus **6**:256–258*T*
Trachypithecus phayrei phayrei **6**:256–258*T*
Trachypithecus phayreis hanicus **6**:256–258*T*
Trachypithecus pileatus durga **6**:256–258*T*
Trachypithecus pileatus pileatus **6**:256–258*T*
Trachypithecus pileatus tenebricus **6**:256–258*T*
Trachypithecus poliocephalus leucocephalus **6**:256–258*T*
Trachypithecus poliocephalus poliocephalus **6**:256–258*T*
Trachypithecus shortridgei **6**:256–258*T*
Trachypithecus vetulus monticola **6**:256–258*T*
Trachypithecus vetulus philbricki **6**:256–258*T*
Trachypithecus vetulus vetulus **6**:256–258*T*
Trachysphaera lobata **5**:423*F*
trade **3**:700–701, **3**:701*F*
tradeoffs **3**:102, **4**:634, **4**:635*F*
traditional agriculture **1:119–125**
  background information **1**:119
  benefits **1**:124–125
  biodiversity-based agricultural strategies **1**:123–124
  biodiversity conservation **1**:122–123
  biodiversity features **1**:119–120
  definition **1**:119
  ecological interaction benefits **1**:122
  global patterns **2**:657–658
  polyculture cropping systems **1**:375–376, **7**:230–231
  research summary **1**:124–125
  traditional knowledge systems
    environmental knowledge **1**:121
    experimental approaches **1**:122
    farming practices **1**:121
    folk taxonomies **1**:121
    research background **1**:120–121
traditional conservation practices **7:226–235**
  background information **7**:226–227
  complex ecosystem dynamics
    disturbance management **7**:233
    ecosystem renewal nurturance **7**:233–234
    general discussion **7**:232–233
    landscape patchiness management **7**:232–233
    multiple-scale ecological processes **7**:233
    watershed management **7**:233
  cultural belief systems **7**:226–227
  ecological monitoring and practices
    habitat protection **7**:227*T*, **7**:229
    harvest restrictions **7**:227*T*, **7**:229–230
    resource monitoring status **7**:227–228, **7**:227*T*
    species protection **4**:710, **7**:227*T*, **7**:228, **7**:229*T*
    vulnerable life histories protection **7**:227*T*, **7**:228–229
  management strategies
    agroforestry practices **7**:230–231
    multiple-species management **7**:230–231
    resource rotation **7**:231
    shifting cultivation systems **7**:231–232, **7**:232*F*
    succession management **7**:231–232
  research summary **7**:234
  social–ecological systems **7**:227*F*
  translocations **7**:236–237
traditional ecological knowledge **6**:385
traditional energy **3**:250
traditional medicine **3**:192
TRAFFIC **1**:518
tragacanth vegetation **5**:480*F*, **5**:483–484
tragedy of the commons **2**:149–150, **2**:234–235, **5**:150–152
Tragelaphus euryceros **1**:49
Tragelaphus imberbis **5**:235
Tragelaphus scriptus **1**:50
Tragelaphus spekei **1**:53–54
Tragelaphus strepsiceros **1**:51
Tragopogon dubius **4**:171*F*, **4**:172*F*
Tragopogon mirus **4**:171, **4**:171*F*
Tragopogon miscellus **4**:171, **4**:171*F*
Tragopogon porrifolius **4**:171*F*, **4**:172*F*
Tragulidae **4**:696
Tragulus kanchil **3**:188
Tragus berteronianus **1**:708
Trametes versicolor **3**:629*F*
trancelike states **3**:635*T*
transboundary protected areas **6**:489, **6**:489*T*
transfer RNA (tRNA) **5**:634–635, **5**:635*F*
transgenic plants **1**:467–468, **4**:307
translocations **7:236–240**
  background information **7**:236
  ethical concerns **7**:239
  managed relocations **7**:238
  motivations **7**:237–238
  research summary **7**:239–240
  risks and costs **7**:238–239, **7**:239*F*
  success assessments **7**:238
  traditional conservation practices **7**:236–237
Transversotrema **7**:450–451*F*
Transversotrematoidea **7**:439–440*T*, **7**:450–451*F*, **7**:454–455
trap crops **1**:376
trap-door spiders **1**:206, **3**:223
Trapellaceae **1**:260*T*
trap fishing **6**:421
trapping **7**:185, **7**:203*F*
trash basket epiphytes **3**:523
trash fish **1**:192–193
travel cost estimation methods **7**:322–324
Traviunoidea **1**:210–211
trawl fishing **3**:453, **3**:482, **6**:420–421
tree cotton **2**:663
treecreepers **1**:618–619, **4**:33–35, **4**:34*F*
tree crickets **3**:730
tree finches **1**:27–28, **1**:28*F*
tree of life **5**:253*F*
trees
  acidic deposition
    red spruce **1**:6–8, **1**:8*F*
    species declines **1**:6–8
    species distribution shifts **1**:8
    sugar maples **1**:8, **1**:8*F*
  elevational range shifts **4**:605*T*
  food webs **7**:266*F*
  indicator species **4**:258*F*
  large-scale biodiversity experiments **4**:588–589
  latitudinal gradients **1**:498*F*
  living dead **4:590–598**
    deforestation **4**:591–593, **4**:592*F*
    extinct megafauna **4**:591–593, **4**:593*F*
    food chains **4**:595–596
    forest ecosystems **4**:590–591, **4**:591*F*
    requirements **4**:593–594
    restoration programs **4**:597
    wildland patches **4**:596–597, **4**:596*F*
  nitrogen deposition effects **5**:530
  North America **5**:564–565
  species richness gradients **4**:613*F*
  tree plantations **1**:682, **2**:456, **4**:432*T*
  tropical rain forests **6**:315*T*, **6**:316*T*, **6**:318*T*
tree shrews **4**:690*T*, **4**:691
tree sparrows **4**:259
tree swallows **5**:720
Trema **1**:263
Tremarctos **4**:724–725*T*
Tremarctos ornatus **2**:587*F*, **7**:229*T*
trematode parasites
  amphibian species decline **3**:696
  characteristics **7**:450–451*F*
  digenean trematode **2**:528*T*
  disease regulation **3**:349
  as disease vectors **1**:368, **1**:368*F*
  domesticated ruminants **1**:710–711
  evolutionary development **2**:645
  foodborne trematodes **5**:671*T*
  life cycles **5**:663*F*
  mollusc–human interactions **5**:381–382
  monogenean trematode **2**:528*T*, **5**:657–658
  number of species **5**:669–670*T*
  phylogenetic relationships **7**:439–440*T*
  reproductive capacity **5**:657–658, **5**:659*T*
  species diversity **7**:451
  see also parasites
Trematomus bernacchii **5**:33
Tremella reticulata **3**:629*F*
Trends in International Mathematics and Science Study (TIMSS) **3**:135
Trepomonas **6**:293
Treptodemoides fukenensis **7**:450–451*F*
triacylglycerols
  adipose tissue **7**:26–27
  fatty acid composition **7**:32–33
  invertebrates **7**:24
  plant lipids **7**:25–26
  reproduction and lactation **7**:37–38
  seeds **7**:25
Triadobatrachus massinoti **1**:169
Triadou Aquifer (France) **7**:59*T*
Trianthema **2**:376*F*
Triassic period
  amphibians **1**:169
  boundary age **5**:1*T*
  crustaceans **2**:404–405
  early mammals **3**:188–189
  fossil record **1**:411*F*, **1**:412*F*, **3**:684*F*, **3**:685*F*
  global change indicators **4**:225*F*

Triassic period (*continued*)
great mass extinctions **5**:168*F*, **5**:169*F*, **5**:172–173
late Triassic faunal turnover **5**:4–5
mammals **4**:684
marine diversity **5**:168*F*, **5**:169*F*
mass extinctions
event characteristics **1**:493–494
extinction intensities **1**:494*T*, **5**:162*F*
marine animals **1**:413*F*, **1**:493*F*, **5**:158
notable mass extinctions **5**:161*F*
number of families **3**:416*T*, **3**:419*F*, **3**:425*F*, **4**:657*F*, **5**:157*F*
originations **1**:495, **1**:495*F*, **3**:425*F*
periodicity **3**:424*F*
Permian–Triassic mass extinction **5**:4
temperate forests **7**:123*F*
terrestrial plant communities
abundance patterns **6**:74*F*
evolutionary development **6**:66*F*
extinction rates **6**:73*F*
origination rates **6**:72*F*
triazines
fungicides **5**:695
herbicides **4**:90*T*, **5**:696, **5**:700*T*
toxicity **5**:699*T*
triazinones **5**:696
triazoles
fungicides **5**:695
herbicides **4**:92*T*, **5**:696
triazolinone **4**:90*T*
triazolopyrimidine sulfonanilide **4**:90*T*
Tribolium castaneum
insecticide resistance **4**:301–302
interspecific competition **2**:184, **2**:184*T*, **2**:185*F*, **2**:185*T*, **6**:670
recombination characteristics **6**:357–358, **6**:361*T*
Tribolium confusum **2**:184, **2**:184*T*, **2**:185*F*, **2**:185*T*, **6**:670
Tribophyceae **6**:10*T*
Triceratops **2**:519*T*, **2**:520*T*
Trichechidae **4**:694
Trichechus **4**:724–725*T*
Trichechus inunguis **3**:192–193, **5**:81*T*
Trichechus manatus **2**:587*F*, **5**:81*T*, **5**:83, **5**:225*T*, **5**:314
Trichechus senegalensis **5**:81*T*
Trichilia **5**:489
trichinella **7**:302–303
Trichinella spiralis **5**:656–657
Trichinotus blochi **5**:662*F*
Trichius fasciatus **7**:218–219
trichlopyr **5**:696
trichloroacetic acid (TCA) **5**:696
Trichobranchidae **7**:417*F*
Trichoceridae **3**:489–491*T*
Trichoderma **6**:447
Trichodesmium **2**:611*F*, **2**:612*T*, **2**:614, **2**:620*F*, **2**:622*F*, **4**:671–673, **5**:302–303
Trichodesmium contortium **2**:611–612, **2**:611*F*
Trichodesmium erythraeum **2**:611–612, **2**:611*F*
Trichodesmium hildebrandt **2**:611–612, **2**:611*F*
Trichodesmium tenue **2**:611–612, **2**:611*F*
Trichodesmium thiebautii **2**:611–612, **2**:611*F*
Trichogramma **5**:679
Trichogrammatidae **5**:674–675
trichomes **2**:428
Trichomitopsis **6**:293
Trichomonas tenax **5**:660*T*
Trichomonas vaginalis **5**:660*T*, **6**:293
Trichomycterus **4**:463–464
Trichoplax adhaerens **6**:712–713
Trichoplusia ni **5**:399*F*
Trichoptera
major divisions and characteristics **4**:308–312, **4**:309*T*
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*
number of species **3**:182*T*, **4**:371*T*, **4**:392*T*
phylogenetic relationships **5**:385, **5**:386*F*
wetland ecosystems **7**:389–391
Trichosporon beigelii **3**:634*T*
Trichosurus vulpecula **5**:370
Trichotermes **4**:409*T*
Trichuris **7**:422–424, **7**:427*F*, **7**:432–433, **7**:432*T*
Trichuris fossor **7**:432*T*
Trichuris trichura **6**:132, **7**:425–428, **7**:428*T*
Tricladida **5**:470*T*, **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
Triconadonta **4**:684
Tridacna gigas **5**:380–381
Tridacna squamosa **5**:119*F*
Tridactyloidea (false mole crickets and sandgropers) **3**:736, **3**:736*F*
tridemorph **5**:695
trifluralin **5**:696
Trifolium alexandrinum **2**:383
Trifolium microcephalum **3**:214
Trifolium repens **2**:386
Trifolium stoloniferum **3**:214
triforine **5**:695
trigger fish **3**:467–471
Triglochin maritima **4**:105
Trigonalyoidea **4**:178
Trigonella foenum-graecum **2**:662–663
Trigonia **5**:623*F*
Trigonopterygoidea (leaf grasshoppers and wingless desert grasshoppers) **3**:735–736, **3**:736, **3**:736*F*
Trigonostylopoidea **2**:583*F*
trilobites
fossil record **3**:685*F*
Late Devonian mass extinction **5**:169–170
Late Ordovician mass extinction **5**:167
radiation ecology **2**:400
Trinervitermes **4**:409*T*
Trinidad and Tobago
botanical gardens **6**:519
energy efficiency measures **3**:254*F*
Trinidadian white-fronted capuchin **6**:256–258*T*
Trionyx nigricans **6**:564–565
Trionyx triunguis **5**:473
Triops cancriformis **5**:434–435
Trioxys utilis **4**:362
triphenyltin (TPT) **5**:695
Triphyophyllum **6**:44*T*
Triplax rufipes **7**:220*T*
Triplostegiaceae **1**:260*T*
tripodfish **3**:465–466
Triptilion **4**:101
Tristerix aphyllus **4**:73
Tristoma **7**:448–449*F*
Tristramella **5**:470
Tristramella sacra **5**:471*F*
triterpenes **2**:427
Triticum aestivum **1**:462–463, **2**:662–663
Triticum dicoccoides **1**:707–708, **3**:670, **3**:671*F*, **3**:672*F*, **5**:483
Triticum monococcum **2**:662–663
Triticum turgidum **2**:662–663
Triturus cristatus **2**:552*F*
Triturus vittatus **5**:473
Triuridaceae **6**:315*T*
Trochetia melanoxylon **3**:208
Trochetiopsis erythroxylon **6**:77
Trochilidae **6**:308*T*
Trochochaetidae **7**:417*F*
Trochodendraceae **1**:260*T*
Trochodendrales **5**:621*F*, **5**:622*F*
Trochoidea davidiana **5**:463
Trochoidea picardi **5**:463, **5**:472, **5**:473*F*
Trochoidea simulata **5**:466–467
troglobionts **7**:52, **7**:52*F*, **7**:55–57, **7**:59*T*, **7**:61*F*
Troglodytidae **6**:308*T*
Trogloglanis pattersoni **7**:58
troglophiles **7**:52
Troglotayoscidae **1**:209
Trogonidae **6**:308*T*
Trogoniformes **6**:308*T*
Troguloidea **1**:210
Trogulus torosus **1**:209–210
Troodon **2**:518, **2**:519*T*, **2**:520*T*
Tropaeolaceae **2**:138
Tropaeolum tuberosum **2**:663–664
tropane alkaloids **6**:136, **6**:136*F*
Tropheus **1**:25*F*
trophic cascades **7:258–263**
apex predators **7**:262–263
aquatic ecosystems **5**:61, **7**:259–260
basic concepts **7**:258–259, **7**:258*F*
consequences **7**:261–262, **7**:261*F*
cross-habitat effects **7**:261–262, **7**:261*F*
food webs **3**:504–505, **7**:258*F*
functional role **7**:260–261
Hairston-Smith-Slobodkin (HSS) hypothesis **6**:723
lake and pond ecosystems **4**:462, **7**:260–261, **7**:261*F*
multispecies interactions **6**:722–723
prey–predator interactions **6**:231, **6**:233*F*, **6**:234*F*, **7**:258–259
research summary **7**:263
terrestrial ecosystems **7**:259–260
trophic levels **7:264–268**
bioenergetics **7**:264–266
El Niño Southern Oscillation (ENSO) events **3**:158–159
food chains **5**:46–47, **7**:264, **7**:265*F*
food webs **3**:239–240, **3**:241*F*, **3**:503–504, **5**:46–47, **5**:48*F*, **5**:49*F*, **7**:266–268, **7**:266*F*, **7**:267*T*
freshwater ecosystems **3**:562, **3**:563*T*
guilds **4**:36
historical research **3**:239–240

lake and pond ecosystems **3**:562, **3**:563*T*, **4**:462
marine ecosystems **5**:46–47, **5**:48*F*, **5**:49*F*, **5**:50–51
multispecies interactions **6**:721–722, **6**:721*F*, **6**:722*F*
pest control **1**:380
phenological mismatches **5**:722–723, **5**:722*F*
plant–animal interactions **6**:39, **6**:40*F*
pyramid of production **3**:239–240, **7**:265, **7**:265*F*
schematic diagram **7**:265*F*
trophic interactions **3**:6–7
universality **7**:264–266
tropical America **2**:448*T*
tropical Asia
characteristics **1**:260–263
deforestation **2**:448*T*
humid tropical forest **1**:260–263, **1**:262–263*T*, **1**:266*T*
monsoon forests (moist deciduous forests) **1**:262–263*T*, **1**:263–266
tropical dry forest, scrub, and savanna **1**:262–263*T*, **1**:266–267
tropical moist forest **1**:277*T*
tropical rain forests **6**:314*T*
tropical/subtropical humid mountain areas **1**:262–263*T*, **1**:266*T*, **1**:267
vegetation belts **1**:266*T*
tropical dry forest
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*
seasonally dry tropical forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
tropical forests **7:269–276**
Africa **1**:48–49
Asia
tropical dry forest, scrub, and savanna **1**:262–263*T*, **1**:266–267
tropical moist forest **1**:277*T*
tropical/subtropical humid mountain areas **1**:262–263*T*, **1**:266*T*, **1**:267
biodiversity change scenarios **4**:6*F*
canopy architecture **3**:511–512
climatic characteristics **7**:270–271
cloud forests **2**:28–29
current global land cover **4**:557*F*
deforestation
annual forest change **2**:448*T*
C_4 plants **2**:379–381, **2**:380*F*
causal factors **2**:379–381, **2**:380*F*
contemporary deforestation **2**:448–449
deforestation rates **7**:277
ecological impacts **1**:677*F*, **1**:679*T*
habitat destruction **3**:164–165
livestock impacts **6**:338–340, **6**:340*F*
regeneration systems **3:277–286**
challenges **7**:284–285
chronosequence studies **7**:279–280
forest dynamics **7**:278, **7**:279*T*
functional diversity (FD) **7**:282
historical perspective **7**:277–278
prior land use effects **7**:281
recovery processes **7**:281–282
regeneration sources **7**:280–281
seed dispersal–predation relationship **7**:283–284
soil fertility **7**:281
species composition changes **7**:282–283, **7**:283*F*
successional continuum **7**:278
successional trajectories and stages **7**:278–280, **7**:279*T*
ecosystem function measurements **3**:77*F*
ecosystem services **1**:334*T*
El Niño Southern Oscillation (ENSO) events **3**:156–158, **3**:158*F*
general discussion **7**:269
geographic extent **7**:269–270, **7**:270*T*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
global distribution patterns **7**:117*T*
grazer-based ecosystem restoration and management **4**:15
habitat loss and fragmentation **3**:405–406, **3**:405*F*
livestock impacts **6**:338–340, **6**:340*F*
living dead **4**:596–597
major forest ecosystems
biodiversity-rich countries **1**:497–498
classifications **7**:271–272
high mountain ecosystems
characteristics **7**:275
humid forests **7**:275
páramo ecosystem **4**:101–103, **7**:275–276
perhumid forests **7**:275
subhumid forests **7**:275
lowland ecosystems
characteristics **7**:271–272
flooded forests **7**:272
humid forests **7**:273
mangrove ecosystems **7**:271–272
perhumid forests **7**:272–273
savannas **7**:274
subhumid forests **7**:273–274
low mountain ecosystems
characteristics **7**:274–275
humid forests **7**:275
perhumid forests **7**:275
subhumid forests **7**:274–275
mountain cloud forests **3**:111
neotropical seasonally dry forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
nitrogen deposition **5**:526*T*
physiographic characteristics **7**:270
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
precipitation–temperature relationship **1**:59*F*
predation effects **6**:239–240
protected area systems **6**:492*T*, **6**:499*T*
regeneration systems **3:277–286**
challenges **7**:284–285
deforestation rates **7**:277
forest dynamics **7**:278, **7**:279*T*
functional diversity (FD) **7**:282
historical perspective **7**:277–278
seed dispersal–predation relationship **7**:283–284
species composition changes **7**:282–283, **7**:283*F*
successional continuum **7**:278
successional trajectories and stages
chronosequence studies **7**:279–280
forest dynamics **7**:279*T*
influencing factors **7**:278–280
prior land use effects **7**:281
recovery processes **7**:281–282
regeneration sources **7**:280–281
soil fertility **7**:281
seasonally dry tropical forests **5:488–500**
animal diversity **5**:490
Central America **2**:28
characteristics **5**:488–490
climate change effects **5**:497–498
deforestation **5**:491–492, **5**:494*F*
extent and geographical distribution **5**:491, **5**:492*F*, **5**:493*T*
fires **5**:496–497
forest fragmentation **5**:496, **5**:497*F*
grassland ecosystems **5**:496
land change drivers **5**:494*F*
management practices **5**:498
protected area systems **5**:492–496, **5**:493*T*, **5**:495*F*
savannas **5**:496
secondary succession **5**:490
shrublands **5**:496

tropical forests (*continued*)
structural and functional characteristics **5**:489–490, **5**:489*T*, **5**:491*T*
soils **7**:271
South America
distribution and structure
general patterns **3**:111–112
mountain cloud forests **3**:111
tropical rain forests **3**:110
tropical semideciduous/dry forests **3**:110–111
functional aspects
tropical rain forests **3**:112
tropical semideciduous/dry forests **3**:112–113
general patterns **3**:111–112
land cover classifications **6**:607*T*
mountain cloud forests **3**:111
natural ecosystems **6**:599–608, **6**:606*F*, **6**:606*T*
tropical broadleaf forests **6**:600–604*T*
tropical moist forest **6**:600–604*T*
tropical rain forests
distribution and structure **3**:110, **3**:111*F*
functional aspects **3**:112
tropical semideciduous/dry forests
distribution and structure **3**:110–111
functional aspects **3**:112–113
tropical broadleaf forests **3**:428*T*, **6**:600–604*T*
tropical deciduous forests **3**:404
tropical dry forest **1**:266–267, **2**:28, **6**:600–604*T*
tropical forest canopies **3**:709–710
tropical moist forest
biodiversity measures **1**:404
conservation status **3**:171–173*T*
nitrogen deposition **5**:526*T*
South America **6**:600–604*T*, **6**:607*T*
tropical Asia **1**:277*T*
tropical semideciduous forests **3**:110–111, **3**:112–113
wildfire effects **3**:156–158, **3**:158*F*
tropical moist forest
biodiversity measures **1**:404
conservation status **3**:171–173*T*
nitrogen deposition **5**:526*T*
protected area systems **5**:444*T*
South America **6**:600–604*T*
tropical Asia **1**:277*T*
tropical racers **6**:409
tropical rain forests
animal diversity **6:304–312**
Asia **1**:263
beetles **6**:306
birds **6**:307–309, **6**:308*T*
butterflies **6**:306–307, **6**:307*T*
frogs **6**:307
future outlook **6**:311–312
general characteristics **6**:306
geographical patterns **6**:309–310
geographic distribution **6**:304–305
high species richness **6**:310–311
historical perspective **6**:304
mammals **6**:309–310, **6**:309*T*
Asia **1**:263
Australia **1**:303–304
biodiversity-rich countries **1**:500–501
Central America **2**:27–28
characteristics **6**:304–305
geographic distribution **6**:304–305
global biodiversity trends **4**:563*F*
habitat destruction **3**:403–404, **3**:403*F*
livestock impacts **6**:338–340, **6**:340*F*
loss and change **6:322–329**
assessment techniques
Global Forest Resources Assessment **6**:322, **6**:323, **6**:324*T*
model-based assessments **6**:322
multidate remote sensing survey **6**:322–323
current rainforest areas
assessment evaluations **6**:325
global distribution **6**:324*T*
model-based assessments **6**:323, **6**:324*T*, **6**:325*T*
1990 remote sensing survey **6**:323, **6**:324*T*, **6**:325*T*
2000 remote sensing survey **6**:323–325, **6**:325*T*
diversity implications
challenges **6**:327–328
cultural diversity loss **6**:328
deforestation risks **6**:327–328, **6**:328*T*
ecosystem resilience loss **6**:328–329, **6**:329*F*
species richness loss **6**:327–328, **6**:328*T*
eco-floristic zone concept **6**:329
forest land change and transition **6**:325–327, **6**:326*F*, **6**:326*T*
historic rainforest areas **6**:323, **6**:324*T*
transition matrices **6**:325–327, **6**:326*F*, **6**:326*T*
net primary production (NPP) **1**:560*T*
plant diversity **6:313–321**
conservation efforts **6**:320
economic value and benefits **6**:319–320
environmental conditions **6**:313
explanatory theories **6**:318–319
geographic regions **6**:314*T*, **6**:316*T*
life-form varieties **6**:313–318, **6**:315*T*, **6**:316*T*
lowland tropical rain forests **6**:317*T*
species assessments **6**:313, **6**:314*T*
species richness trends **6**:317*T*
precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
rain forest types **6**:305–306
seed dispersal and germination **1**:708
soils **6**:123*T*
South America
distribution and structure **3**:110, **3**:111*F*
functional aspects **3**:112
termites **4**:430*T*, **4**:432*T*
tropical reef systems **6**:236
tropical semideciduous forests **3**:110–111, **3**:112–113
Tropic of Cancer **2**:458
Tropic of Capricorn **2**:458
TROPICOS database **3**:526
Tropocyclops prasinus mexicanus **1**:10*T*
troposheric ozone precursors **6**:152–154
Troposodon **4**:727–728
Troposodon minor **5**:364*F*
trout
anthropogenic impacts **3**:579
aquaculture **1**:190*T*
biomonitoring processes **3**:123*T*
characteristics **3**:464–465, **6**:522
critical pH levels **6**:153*F*
eutrophication impacts **3**:365*F*
lake and pond ecosystems **3**:181
life histories **6**:524–526
mercury contamination **1**:12*F*
surface water acidification **1**:10*T*
troutperch **3**:466–467
true bugs **7:287–295**
Heteroptera
characteristics
coloration and ornamentation **7**:291
feeding behaviors **7**:288
habitat **7**:288
host specificity **7**:290
mimicry **7**:291
morphology **7**:288–292
parasitization **7**:292
physical characteristics **7**:288
predatory behaviors **7**:289
reproduction and growth **7**:289–290
scent glands **7**:288
economic importance **7**:292–293
infraorders
Aradomorpha **7**:294
characteristics **7**:293–294
Cimicomorpha **7**:293
Gerromorpha **7**:293
Leptopodomorpha **7**:293
Nepomorpha **7**:293
Pentatomomorpha **7**:294
number of species **4**:392*T*, **7**:287
research summary **7**:294–295
terminology **7**:287–288
insecticide resistance **4**:298–299
migratory patterns **5**:316
number of species **4**:371*T*, **4**:392*T*
suborders
Coleorrhyncha **7**:287
Heteroptera **7**:287
Homoptera **4**:392*T*, **7**:287
research controversies **7**:287
true crickets **3**:729
true flies *see* Diptera
true grasshoppers **3**:735
true lemurs **6**:251*T*
true midges **4**:374*F*
true páramo zone **4**:103
trumpetfish **3**:467–471
trunkfish **3**:467–471
Trusan-Sundar mangroves **1**:276*T*
Truxillo coca **6**:136
Trypanoplasma borreli **5**:572

Trypanorhyncha **7**:439–440*T*, **7**:456–457*F*, **7**:459–460
Trypanosoma brucei **6**:293
Trypanosoma brucei brucei **5**:672
Trypanosoma cruzi **7**:292–293
trypanosomiasis **3**:93
Tsavo National Park **5**:234–236
tse tse flies **3**:93, **3**:489–491*T*
Tsuga
  Asia **1**:272–273
  boreal forests **1**:628–630
  introduced species **4**:361
  North America **7**:118
Tsuga diversifolia **1**:273
Tsuga dumosa **1**:272–273
Tsuga mertensiana **2**:571
Tsuga sieboldii **1**:271
Tsuga yunnanensis **1**:272–273
Tuamotu sandpiper **7**:184
Tuan, Yi-Fu **6**:384
tuatara
  characteristics **6**:401, **6**:406
  conservation status **3**:216*T*, **3**:217–218, **6**:411
  historical extinctions **5**:367
Tubastraea floreana **3**:202*T*
tube-eye fish **3**:466
tubenoses **7**:181–182
tuberculosis **7**:302–303
tubers
  East Asia **2**:663
  North and South America **2**:663–664
  Southeast Asia **2**:663
  Sub-Saharan Africa **2**:663
tube worms
  chemosynthetic symbioses **5**:120–121, **5**:120*F*
  classifications **4**:381*T*
Tubifex tubifex **3**:364
tubocurarine **5**:711–713, **5**:712*T*
Tubulidentata **4**:690*T*, **4**:695
tucuman **4**:286–287
tufted beardgrass **4**:364
Tulipa **1**:269
Tumbes **6**:82*T*
tumeric **2**:663
tuna **3**:447, **3**:467–471, **6**:422*T*
tundra ecosystems
  biodiversity change scenarios **4**:6*F*
  boreal forests **1**:262–263*T*, **1**:274–275
  C_4 plants **2**:371*T*
  Canadian Shield **5**:562
  characteristics **6**:330–331, **6**:331*F*
  climate change studies **2**:61*F*
  current global land cover **4**:557*F*
  ecosystem function measurements **3**:77*F*
  Europe **3**:335*T*, **3**:336–340*T*
  forest–tundra ecotone **1**:630
  global agriculture impacts **3**:428*T*
  global biodiversity trends **4**:563*F*
  net primary production (NPP) **1**:560*T*
  nitrogen deposition **5**:534*F*
  pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*
  precipitation–temperature relationship **1**:59*F*
  protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
  soil characteristics **6**:123*T*
Tungshih mangroves **1**:276*T*
tunicates
  characteristics **7**:336
  classification systems **7**:335, **7**:335*F*, **7**:336*F*
  microhabitat creation **5**:60–61
Tunisia **5**:481*T*
  botanical gardens **6**:507*T*
  wheat production **1**:87–88
  see also Near East ecosystems
Tuoba poseidonis **5**:425
Tuosu Hu **1**:276*T*
Tupaiidae **4**:691
Tupidanthus **3**:522*T*
Turanian desert **1**:262–263*T*
Turbellaria
  adaptability **4**:393, **5**:657–658
  Bothrioplanida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443–444
  Catenulida **7**:439–440*T*, **7**:440, **7**:441*F*, **7**:442–443*F*
  classifications and sampling methods **4**:381*T*
  Dalyellioida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
  Fecampiida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
  Haplopharyngida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
  Kalyptorhynchia **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
  Lecithoepitheliata **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
  Macrostomida **7**:439–440*T*, **7**:440–443, **7**:441*F*, **7**:442–443*F*
  number of species **3**:182*T*, **4**:371*T*, **5**:669–670*T*, **7**:439–440*T*
  plankton fraction **6**:25
  Polycladida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:443
  Prolecithophora **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
  Proseriata **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
  reproductive capacity **5**:657–658, **5**:659*T*
  river ecosystems **6**:478*T*
  Temnocephalida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444
  Tricladida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:445
  Typhloplanoida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444–445
Turdidae **6**:308*T*
Turdus merula **5**:312, **5**:474
Turdus migratorius **4**:174, **5**:723
Turdus philomelos **4**:259
Turing, Alan **7**:9–10
Turing instability **7**:9–10, **7**:10, **7**:10*F*
Turkana (Lake) **4**:461*T*
Turkey
  animal species **1**:261*T*
  botanical gardens **6**:519
  ecosystems
    dry regions **1**:268–269
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  nitrogen deposition **5**:526*T*
  plant species **1**:259*T*
  wheat production **1**:87–88
turkeys **1**:644*F*
Turkmenistan **6**:519
Turks and Caicos Islands **4**:560*F*, **6**:519
turlough **7**:369*T*
Turner, Frederick Jackson **4**:652
turnips
  native crops
    East Asia **2**:663
    Europe **2**:662–663
    Mediterranean Basin **2**:662–663
    southwest Asia **2**:662–663
Tursiops truncatus **3**:256, **5**:81*T*
turtle grass **3**:360–361
turtles
  anthropogenic impacts **6**:410–411
  Ascension Island **6**:78
  characteristics **6**:400–401, **7**:339–340
  conservation status **3**:216*T*, **6**:410–411
  countryside biogeography studies **2**:356
  endangered species **3**:217
  geographic distribution **6**:400–401
  gigantism **6**:405
  hatchling characteristics **6**:405
  migratory patterns **2**:9*F*
  Near East ecosystems **5**:472
  nest environment **6**:405
  river ecosystems **6**:479–480
  threatened species **1**:515–516
  tourism industry **7**:223
  urbanization threats **4**:564
  wetland ecosystems **7**:390
  see also Reptilia/reptiles
tusked weta **5**:366–367
tussock grasslands **1**:292–297, **1**:294*T*
Tuvalu **6**:507*T*
TWINSPAN (two-way indicator species analysis) method **4**:264–265
twisted-wing parasites **4**:392*T*
two-prey competition **2**:177, **2**:177*F*
two-spotted spider mite **4**:298–299, **4**:299*F*
two-toed sloth **2**:590*T*
Tylecodon paniculatus **1**:487
Typha **1**:270, **1**:278–279
Typha angustifolia **4**:173–174
Typha domingensis **5**:486, **7**:372
Typha latifolia **4**:173–174
Typhlocaris galilea **5**:470–471, **5**:470*F*, **5**:470*T*
Typhlocirolana reichi **5**:470–471, **5**:470*T*
Typhlocirolana steinitzi **5**:470*T*
Typhlocoelum **7**:450–451*F*
Typhloglomeris caeca **7**:60
Typhlomolge rathbuni **7**:58
Typhlonectes **7**:339
Typhloplanoida **7**:439–440*T*, **7**:441*F*, **7**:442–443*F*, **7**:444–445
Typhula sp. **6**:449

Tyrannidae **6**:308*T*
Tyrannosaurus **2**:519*T*, **2**:520–521, **2**:520*T*
Tyrannus dominicensis **1**:618
Tyto alba **5**:606–607, **6**:689–690

U

uakaris **6**:251*T*
Uapaca bojeri **1**:55–56
Uapaca kirkiana **1**:68
Ucides **5**:16, **5**:16*F*
Udzungwa red colobus **6**:256–258*T*
Uganda
  botanical gardens **6**:519
  cropland conversion **4**:665*T*
  deforestation **2**:449*T*
  energy efficiency measures **3**:254*F*
  megaherbivores **5**:227–228, **5**:227*F*, **5**:234–236
  primate diversity **6**:252*T*
  tropical rain forests **6**:314–315, **6**:317*T*
U insertion **5**:572–575, **5**:575*F*
Ukraine
  botanical gardens **6**:519
  vegetation surveys **3**:341–344*T*
  wheat production **1**:86
'Ula'aihawane **1**:29*F*
ulluco **2**:663–664
Ullucus tuberosus **2**:663–664
Ulmaceae **1**:263
Ulmus **7**:118
Ulmus americana **2**:528*T*, **3**:411
Ulmus davidiana **1**:271
Ulmus macrocarpa **1**:271–272
Ulmus mexicana **7**:275
Ulmus pumila **1**:270
Ultisols **4**:225–226, **4**:228
ultraoligotrophic level **3**:563*T*
ultraviolet radiation **7:296–303**
  characteristics **7**:296–297
  environmental threats
    amphibians **3**:165
    climate change **3**:450
    elevated ultraviolet (UV–B) radiation **3**:450, **3**:697, **4**:155
  habitat degradation **3**:407
  measurement and analysis **7**:297–298
  ozone depletion **1**:289
  radiation effects
    aquatic systems
      amphibians **7**:300–301
      fish **7**:300–301
      macroalgae **7**:299–300
      marine invertebrates **7**:300
      plankton **7**:299
      research background **7**:299
      seagrass **7**:299–300
    biogeochemical cycles **7**:301
    biological systems
      aquatic systems **7**:299
      DNA repair **7**:298
      terrestrial plants **7**:298–299
    humans
      characteristics **7**:301–302
      eyes/photokeratitis **7**:302
      immune system **7**:302–303
      photoaging **7**:302
      skin cancer **7**:302
      sunburn **7**:302
Ulva **3**:358–359
Umbelliferae **3**:129*T*
Umbonium **5**:381
umbrella species **3**:197–198, **3**:225, **7**:406
umbrella-thorn acacia **2**:469–470
uncertainty
  common property resources **2**:153–154
  ecosystem services **3**:103, **5**:356
  global species richness
    environments
      coral reef ecosystems **3**:710
      deep ocean environments **3**:710
      soils **3**:710
      tropical forest canopies **3**:709–710
    general discussion **3**:709–710
    higher taxa
      arthropods **3**:711
      bacteria **3**:710–711
      fungi **3**:711
      nematodes **3**:711
      Protozoa **3**:711
      viruses **3**:710
  modeling frameworks **5**:356
  paleoecological data **5**:652
  property rights
    biodiversity protection **6**:277–278
    decision-making concerns **6**:276
UN Commission on Sustainable Development **4**:326–327
UN Conference on Environment and Development **7**:43–44
UN Conference on the Law of the Sea **6**:421
UN Convention on the Law of the Sea **6**:421
UN Convention to Combat Desertification (UNCCD) **7**:137
Undaria pinnatifida **4**:345–346, **4**:346*F*
UNESCO
  background information **4**:324–325
  UNESCO Biosphere Reserve Program **2**:301, **6**:227, **6**:489, **7**:18–19
  UNESCO World Heritage Sites **6**:506
UN Framework Convention on Climate Change (UNFCCC) **2**:318, **4**:326
ungava seal **5**:77*T*
Unguitermes **4**:409*T*
ungulates
  adaptive radiation **1**:22, **1**:23*F*
  biodiversity effects **3**:188
  convergent evolution **4**:701*T*
  desert ecosystems **2**:475–476
  extinction theories **2**:517
  food habits **3**:127–128
  fruit and seed dispersal **6**:50
  grassland ecosystems **6**:332
  hot deserts **6**:334
  migratory patterns **5**:313
  prey–predator interactions **4**:699–700
  temperate/cold-desert shrublands **6**:334
  threat percentage **3**:195
  tundra ecosystems **6**:330–331
  see also megaherbivores
UN High Commissioner for Refugees (UNHCR) **4**:330
Unified Medical Language System (UMLS) **1**:444
Uniform Resource Identifier (URI) **1**:441–442
Unikonts **2**:641
Uniramia **6**:712–713
United Arab Emirates
  animal species **1**:261*T*
  botanical gardens **6**:519
  ecosystems
    mangroves **1**:275*T*
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  energy efficiency measures **3**:254*F*
  plant species **1**:259*T*
United Fruit Company **5**:601
United Kingdom
  botanical gardens **6**:519
  Dartmoor National Park **5**:445
  edible plants **3**:130*T*
  Fallopia japonica **4**:351–352, **4**:352*F*
  flora **5**:481*T*
  introduced species
    nonindigenous invertebrates
      arthropods **1**:76*T*, **1**:79–80
      crop plant pathogens **1**:76*T*, **1**:80
    nonindigenous plants **1**:76*T*, **1**:79
    nonindigenous vertebrates
      birds **1**:79
      mammals **1**:76*T*, **1**:79
  invasive species **4**:351–352, **4**:352*F*
  natural history museums and institutions **5**:408*F*, **5**:409*F*, **5**:411*F*
  nitrogen deposition **5**:529*F*
  reforestation programs **6**:376–377
  species richness **1**:499*T*
  threatened birds **7**:182
  vegetation surveys **3**:341–344*T*
United Nation Framework Convention on Climate Change (UNFCCC) **1**:521
United Nations **4**:324–325, **7**:18–19
United Nations Conference on Environment and Development (UNCED) **4**:313, **6**:597–598, **7**:43–44
United Nations Conference on the Human Environment **4**:325, **6**:78
United Nations Conference on the Law of the Sea **6**:421
United Nations Conference on Trade and Development (UNCTAD) **1**:598–599
United Nations Convention on Biological Diversity **2**:286
United Nations Convention on the Law of the Sea **2**:286, **4**:329–330
United Nations Convention to Combat Desertification (UNCCD)
  formal definitions **2**:484
  goals and objectives **2**:484–485, **2**:485*T*

origins and formation **2**:484
10-year strategic plan **2**:485–486, **2**:486*T*
United Nations Development Program (UNDP) **2**:231–232
United Nations Educational, Scientific and Cultural Organization (UNESCO) *see* UNESCO
United Nations Environment Programme (UNEP) **1**:704–705, **4**:324–325, **4**:664–666
United Nations University/Institute for Natural Resources in Africa (UNU/INRA) **3**:140
United States
amphibians **5**:565
birds **5**:565
botanical gardens **6**:519
butterflies **1**:657*T*
deforestation **1**:677*F*
endangered mammals **3**:194*T*
endangered plant species **3**:205–206, **3**:207*F*, **3**:209
endemism **3**:232*T*
energy use **3**:252*T*
fishing/fisheries **6**:422*T*
freshwater ecosystems **3**:579–580
freshwater fish **5**:565
government legislation and regulations **3:712–721**
biodiversity protection **3**:712–713, **3**:713*T*
federal lands
management strategies **3**:718
National Forest Management Act (1976) **3**:713*T*, **3**:718
National Marine Sanctuary System (1972) **3**:718
National Park Service Act (1916) **3**:713*T*, **3**:718
National Wildlife Refuge System Improvement Act (1997) **3**:713*T*, **3**:718
habitats and ecosystems
federal lands **3**:718
management strategies **3**:718
private land **3**:719, **7**:371
international agreements
biodiversity protection **3**:713*T*
Convention on Biological Diversity (1930) **3**:713*T*, **3**:714–715
Convention on International Trade in Endangered Species (CITES) **3**:713*T*, **3**:714
importance **3**:713–714
International Convention for the Regulation of Whaling (1946) **3**:713*T*, **3**:714
local ordinances **3**:720–721
national legislation
background information **3**:715
Bald Eagle Protection Act (1940) **3**:713*T*, **3**:716
conservation of genetic resources **3**:715
Endangered Species Act (ESA) **3**:713*T*, **3**:716–717
Fishery Conservation and Management Act (1976) **3**:713*T*, **3**:717–718
Foreign Assistance Act (1962) **3**:713*T*, **3**:719
habitats and ecosystems **3**:718
Lacey Act (1900) **3**:713*T*, **3**:715
Marine Mammal Protection Act (1972) **3**:713*T*, **3**:716
Migratory Bird Treaty Act (1918) **3**:713*T*, **3**:715–716
National Environmental Policy Act (1970) **3**:713*T*, **3**:719
nonnative species control **3**:715
Wild Bird Conservation Act (1992) **3**:713*T*, **3**:716
research summary **3**:721
state legislation **3**:719–720
habitat–bird species diversity relationship **1**:614, **1**:614*F*
historical conservation practices **2**:279–280
indigenous peoples **4**:270*T*
introduced species
background information **1**:75–76
nonindigenous invertebrates
earthworms **1**:78
general discussion **1**:78
insects and mites **1**:76*T*, **1**:78
microbes **1**:76*T*, **1**:78–79
nonindigenous plants
crop plants **1**:75–76
economic impact **1**:76, **1**:76*T*
nonindigenous vertebrates
amphibians and reptiles **1**:78
exotic birds **1**:76*T*, **1**:77–78
mammals **1**:76–77, **1**:76*T*
land mammals **5**:565
maize production **1**:86
natural history museums and institutions **5**:405–406, **5**:408*F*, **5**:409*F*, **5**:411*F*
nonvascular epiphytes **3**:521*T*
parasitic infections **5**:671*T*
protected area systems **6**:492*T*, **6**:493*F*, **6**:494*F*, **6**:494*T*
regional carbon stocks **1**:678*F*
reptiles **5**:565
seasonal rainfall **2**:465*F*
species richness **1**:499*T*
threatened birds **7**:181, **7**:182
threatened species **3**:579–580
wheat production **1**:86
United States Agency for International Development (USAID) **2**:231–232
university botanical gardens **6**:505
University of Copenhagen Botanic Garden **6**:506–508, **6**:510–511
University of Helsinki BotanicGarden **6**:510–511
University of Oslo Botanic Garden **6**:506–508
University of Tokyo Botanic Garden **6**:506–508
University of Vienna Botanic Garden **6**:506–508
UN Program for Reducing Emissions from Deforestation and Forest Degradation (UN-REDD) **4**:328–329
unsaturated fatty acids **7**:25–26, **7**:32–33
Upanishads **7**:226
upland cotton **2**:663–664
upland sandpiper **5**:312
Upper Guinea red colobus **6**:256–258*T*
upper Kellwasser Event **5**:157
upper respiratory tract syndrome **2**:528*T*
Uppsala University Botanic Garden **6**:510–511
upside-down flies **3**:489–491*T*
uracils **4**:90*T*, **5**:696
Urania lelius **5**:395*F*
Uraniidae **5**:395*F*
uranium (U) **3**:265*T*
Urbania **4**:101
urbanization, global **4:231–244**
biodiversity impacts **4**:564, **7**:203*F*
demographic transition **4**:231–232
dependence interactions
biodiversity **4**:237
ecosystem services **4**:237–238, **4**:238*F*
general discussion **4**:237
environmental impact
ecosystem services **4**:235
general discussion **4**:234–235
habitat loss and fragmentation **4**:234–235
positive impacts **4**:236–237
protected area systems **4**:235–236, **4**:235*F*
geographic distribution **4**:235*F*
management strategies
ecosystem services **4**:239*F*, **4**:241–242
general discussion **4**:238–241
rural-to-urban migration **4**:240–241*F*, **4**:242
sustainable development **4**:239–241, **4**:239*F*
population trends **4**:231, **4**:231*F*
rural-to-urban migration **4**:232–234, **4**:233*F*, **4**:240–241*F*, **4**:242
urban sewage **1**:512, **1**:512*F*
urban–suburban biodiversity **7:304–313**
anthropogenic activities and impacts **7**:307*T*, **7**:308*T*
city-scale decisions **7**:309–310
dryland syndrome **2**:483*F*
evaluation criteria **7**:305–306
global scale **7**:310
governmental policies **7**:308*T*
hinterlands **7**:309–310
historical perspective **7**:304
individual and household scales **7**:306–308
land use issues **4**:560*F*, **4**:561*F*
Millennium Ecosystem Assessment (MEA) **7**:305–306
neighborhood scale **7**:308–309
regional scale **7**:310
settlement–biome interactions **7**:304–305
socioecological theories **7**:310–312
urban soil conservation **6**:595
ureas **4**:92*T*

urease **5**:269*T*
Urey, Harold Clayton **5**:630–631
Urochloa **2**:376*F*
Urodacidae **1**:209
Urodacus manicatus **1**:208
Urodacus yaschenkoi **1**:208
Urodoidea **5**:387–388*T*
Uroglena **6**:29–30, **6**:293
uronic acids **6**:113*T*
Urophora affinis **4**:363
Urophora quadrifasciata **4**:363
Uropygi **1**:202*T*, **1**:205*F*, **1**:211–212, **1**:211*F*, **1**:246*T*
Urosalpinx cinerea **5**:61
Ursidae
  body mass **1**:687*F*
  evolutionary history **2**:583*F*
  food categories **1**:690*F*, **1**:691*F*
  home range size **1**:687*F*
  late Quaternary extinction event (LQEE) **4**:724–725*T*, **4**:727*T*
  species diversity **1**:686
  tropical rain forests **6**:309*T*
Ursus **1**:709
Ursus arctos **2**:242–243, **2**:550, **2**:552*F*, **3**:190, **4**:564
Ursus arctos horribilis **2**:571
Ursus arctos syriacus **5**:472
Ursus horriblis **4**:62
Ursus maritimus see polar bears
Ursus spelaeus **4**:729*F*
Ursus thibetanus **3**:192
Urticales **5**:621–625, **5**:623*F*
Urtinotherium **5**:224*T*
Uruguay
  botanical gardens **6**:519
  land cover classifications **6**:607*T*
  savannas **6**:600–604*T*
U.S. Agency for International Development (USAID) **3**:719, **7**:19
U.S. Agricultural Research and Marketing Act (1946) **3**:713*T*
U.S. Army Corps of Engineers **7**:152–153
U.S. Climate Change Science Program (CCSP) **2**:106
U.S. Endangered Species Act (1973) **2**:224, **2**:243–244, **3**:209
U.S. Environmental Protection Agency (EPA) **1**:95–97, **4**:90–91
U.S. Fish and Wildlife Service **7**:19
U.S. Forest Service **2**:282–283, **7**:19
U.S. National Park Service **7**:19
Usnea longissima **1**:591
Ustilago maydis **3**:633, **6**:355
U.S. Wilderness Act **2**:224
Uta Hick's bearded saki **6**:256–258*T*
Uta stansburiana **2**:720–721
Utetheisa bella **5**:397, **5**:397*F*
utilitarianism **4**:434–435
Utisols **6**:123*T*
Utricularia **6**:44*T*
Uvigerina **6**:291–292
uxi **4**:287–290
Uzbekistan
  botanical gardens **6**:519
  energy efficiency measures **3**:254*F*

V

vacant niche space **5**:597
Vaccinium **3**:76*F*, **3**:130*T*, **7**:275–276
Vaccinium myrtillus **1**:155*F*
Vaccinium uliginosum **1**:273–274
Vaccinium vitis-idea **1**:155*F*
Vaejovidae **1**:209
Vaejovis littoralis **1**:208
Valeriana alpifolia **4**:103
Valeriana (Aretiastrum) sedifolia **4**:106–107
Valgipes **4**:727*T*
Vallisneria americana **6**:477–478
vampire bats **2**:352
Vampyrella **6**:290–291
Vanessa cardui **6**:689–690
Vangidae **4**:33
Vangueria infausta **1**:68
Vanilla fragrans **2**:663–664
van Leeuwenhoek, Antonie **5**:116–117
Vannella **6**:290–291
Van't Hoff–Arrhenius equation **3**:677
vanua **7**:233
Vanuatu
  botanical gardens **6**:519
  indigenous peoples **4**:270*T*
  island ecosystems **5**:592*F*
vaquitas **5**:77*T*, **5**:80
Varanus griseus **5**:472
Varanus indicus **5**:13
Varecia rubra **6**:256–258*T*
Varecia variegata editorum **6**:256–258*T*
Varecia variegata subcincta **6**:256–258*T*
Varecia variegata variegata **6**:256–258*T*
varied white-fronted capuchin **6**:256–258*T*
variegated black-and-white ruffed lemurs **6**:256–258*T*
variegated spider monkey **6**:256–258*T*
varnish, desert **2**:462
várzea forest **1**:163
vascular plant life-forms
  abundance patterns **6**:73, **6**:74*F*
  air pollution impacts **1**:142*F*
  aquatic vascular plants **3**:359–360
  Australia **1**:302*F*
  biodiversity–population density relationship **7**:74*F*
  coevolutionary relationships **6**:69–70
  endangered species **3:205–215**
    causal factors **3**:208–209
    conservation implications **3**:212–215
    endangerment patterns **3**:206–208, **3**:207*F*
    global distribution patterns **3**:207*F*
    legal protection mandates
      International Plant Protection Convention **3**:209
      U.S. Endangered Species Act (1973) **3**:209
    occurrences **3**:205–206
    rarity
      causal factors **3**:208–209
      characteristics **3**:206
      consequences **3**:209
    species-environmental matching models **3**:211–212
  evolutionary development **6**:65–67, **6**:66*F*
  extinction rates **6**:73*F*
  global centers of diversity **6**:82*T*
  Holocene postglacial spread **2**:551–552*T*
  indicator species **7**:206–207*F*
  Mediterranean-type climate ecosystems
    Australia **5**:213*T*, **5**:221
    California **5**:213*T*, **5**:215
    Cape Region (South Africa) **5**:213*T*, **5**:219–220
    Chile **5**:213*T*, **5**:216–217
    Mediterranean Basin **5**:213*T*, **5**:218
  number of species **6**:586*T*
  origination rates **6**:71–72, **6**:72*F*
  plant invasions **6**:93–95
  river ecosystems **6**:477–478
  temperature tolerance ranges **4**:115
  thermal ecosystems **4**:112
  threatened species **7**:206–207*F*
  tropical rain forests **6**:315*T*
  vascular epiphytes
    characteristics **3**:517, **3**:518*T*
    taxonomic distributions **3**:519–520, **3**:519*T*, **3**:520*T*
  see also plants
vascular wilt **3**:633*T*
Vatairea guianensis **1**:163
Vatica **1**:263
Vatican City **6**:507*T*
Vavilov centers **4**:141
Vedas **7**:226
vegetables
  Europe **2**:662–663
  intercropping patterns **2**:383
  Mediterranean Basin **2**:662–663
  Neolithic agriculture **2**:658
  North and South America **2**:663–664
  Southeast Asia **2**:663
  southwest Asia **2**:662–663
  Sub-Saharan Africa **2**:663
vegetarian finches **1**:27–28, **1**:28*F*
vegetation
  Africa
    Cape Region (South Africa) **1**:47–48
    classifications **1**:45
    Mediterranean North Africa **1**:47
    tropical Africa
      afroalpine ecosystems **1**:53
      bushlands **1**:51–52
      coastal ecosystems **1**:55
      deserts **1**:53
      forests **1**:48–49
      grasslands **1**:52
      lake and pond ecosystems **1**:54–55
      montane ecosystems **1**:53
      phytogeographic/ecoclimatic zones **1**:48
      seasonal tropical vegetation **1**:49–51
      shrublands **1**:52–53
      thickets **1**:51–52
      wetlands **1**:53–54
      woodlands **1**:49–51
  anthropogenic activities and impacts **3**:401*F*, **7**:307*T*
  Arctic ecosystems **1**:227–231, **1**:230*T*, **1**:231*T*, **1**:236*F*

biodiversity restoration **6**:456–457
boreal forests
  characteristics **1**:628–630
  geographic boundaries **1**:629*F*, **1**:630*T*
  latitudinal classifications
    closed forests **1**:629–630
    forest–tundra ecotone **1**:630
    general discussion **1**:629–630
    lichen woodlands **1**:630
  longitudinal classifications **1**:630–631, **1**:631*T*, **1**:632*T*, **1**:633*F*
climate change studies **2**:59–60, **2**:61*F*, **2**:65–66, **2**:82–83, **2**:83*F*, **2**:91–92
disease outbreaks **3**:411, **3**:411*F*
dynamic global vegetation models (DGVMs) **1**:64*F*, **2**:681–682
dynamic models **2**:65–66, **2**:91–92
ecosystem function measurements **3**:79–80
El Niño Southern Oscillation (ENSO) events **3**:158–159
European ecosystems
  classifications
    background information **3**:333–334
    dominance approach **3**:334
    floristic–sociological approach **3**:334
    habitat classification **3**:334
  diversity
    Arctic desert biome **3**:335*T*
    azonal vegetation **3**:336–340*T*
    biotic interactions **3**:334*T*, **3**:335
    boreal forest biome **3**:335*T*, **3**:336–340*T*
    boreal/nemoral orobiomes **3**:336–340*T*
    continental semidesert biome **3**:335*T*, **3**:336–340*T*
    habitat diversity **3**:334–335, **3**:334*T*
    human interference **3**:334*T*, **3**:335
    Mediterranean biome **3**:335*T*, **3**:336–340*T*
    Mediterranean orobiomes **3**:336–340*T*
    pattern variations **3**:335
    prerequisites and sources **3**:334–335, **3**:334*T*
    research literature and surveys **3**:341–344*T*
    semidesert orobiomes **3**:336–340*T*
    steppe/grassland biome **3**:335*T*, **3**:336–340*T*
    synanthropic vegetation **3**:336–340*T*
    taxon diversity **3**:334, **3**:334*T*
    temperate deciduous forest biome **3**:335*T*, **3**:336–340*T*
    tundra biome **3**:335*T*, **3**:336–340*T*
    warm-temperature evergreen forest biome **3**:335*T*, **3**:336–340*T*
    warm-temperature semidesert biome **3**:335*T*, **3**:336–340*T*
    zonation patterns **3**:335–340, **3**:336–340*T*
  pattern variations
    Arctic desert biome **3**:335*T*
    azonal vegetation **3**:336–340*T*
    biomes **3**:335, **3**:335*T*
    boreal forest biome **3**:335*T*, **3**:336–340*T*
    boreal/nemoral orobiomes **3**:336–340*T*
    continental semidesert biome **3**:335*T*, **3**:336–340*T*
    Mediterranean biome **3**:335*T*, **3**:336–340*T*
    Mediterranean orobiomes **3**:336–340*T*
    research literature and surveys **3**:341–344*T*
    semidesert orobiomes **3**:336–340*T*
    steppe/grassland biome **3**:335*T*, **3**:336–340*T*
    synanthropic vegetation **3**:336–340*T*
    temperate deciduous forest biome **3**:335*T*, **3**:336–340*T*
    tundra biome **3**:335*T*, **3**:336–340*T*
    vegetation diversity **3**:335–340
    vegetation maps **3**:340–344, **3**:345*F*, **3**:346*T*
    warm-temperature evergreen forest biome **3**:335*T*, **3**:336–340*T*
    warm-temperature semidesert biome **3**:335*T*, **3**:336–340*T*
    zonation patterns **3**:335–340, **3**:336–340*T*
fire/fire regimes **3**:437–438
herbaceous vegetation species richness **4:79–86**
  control mechanisms
    intermediate disturbance hypothesis **4**:79
    intraspecific diversity **4**:79–80, **4**:80*F*
  dominants and subordinates **4**:82–83, **4**:82*F*, **4**:84
  general discussion **4**:79
  high species richness effects **4**:81–82
  mass ratio hypothesis **4**:80–81, **4**:81*F*
  plant traits **4**:80
  research summary **4**:84–85
  species declines
    filter effects **4**:84
    founder effects **4**:84
  transients **4**:83–84, **4**:83*T*, **4**:84
high-elevation Andean ecosystems **4:96–110**
  background information **4**:96
  geographic distribution **4**:96–97, **4**:97*F*
  habitat characteristics **4**:96–97
  number of species **4**:102*T*
  páramo ecosystem
    characteristics **4**:101–103
    geographic distribution **4**:96–97, **4**:97*F*
    northern Andean alpine cushion formations **4**:103
    northern Andean dwarf *Polylepis* forest **4**:103
    number of species **4**:102*T*
    physical characteristics and patterns **4**:97–98
    plant diversity **4**:98–99, **4**:99*T*
    representative species **4**:102*F*
    subpáramo zone **4**:102–103
    superpáramo zone **4**:103
    true páramo zone **4**:103
  physical characteristics and patterns **4**:97–98
  plant diversity
    endemism **4**:99–100, **4**:99*T*
    origins and diversification **4**:100–101
    species richness **4**:98–99
  puna
    central Andean dwarf *Polylepis* forest **4**:105
    cushion bogs **4**:105
    desert puna **4**:104–105
    dry puna **4**:104
    geographic distribution **4**:96–97, **4**:97*F*
    humid puna **4**:104
    number of species **4**:102*T*
    physical characteristics and patterns **4**:97–98, **4**:103–104
    plant diversity **4**:98–99
    representative species **4**:104*F*
    salar vegetation **4**:105
  research summary **4**:107
  southern Andean steppes
    altitudinal zonation **4**:106–107, **4**:106*F*
    bogs and seeps **4**:107
    cushion plants **4**:106–107, **4**:106*F*
    geographic distribution **4**:96–97, **4**:97*F*
    latitudinal gradients **4**:106–107
    number of species **4**:102*T*
    physical characteristics and patterns **4**:97–98, **4**:105–107
    plant diversity **4**:98–99
    representative species **4**:105*F*
historic land-use patterns **4:569–582**
  anthropogenic impacts **4**:569–570
  buildings/built-up areas **4**:579, **4**:580*F*
  climate change effects **4**:569–570
  conservation implications **4**:581–582
  coppice woods **4**:573, **4**:573*F*, **4**:574*F*
  crop plants **4**:570, **4**:570*F*
  domestic animals **4**:570
  fields **4**:570–572, **4**:571*F*
  fire cycles **4**:580, **4**:581*F*
  forests and woodlands **4**:572–574, **4**:573*F*, **4**:574*F*
  grasslands **4**:577
  heathland **4**:577, **4**:578*F*
  hedges **4**:571, **4**:571*F*
  historical perspective **4**:569, **4**:579–581
  moorland **4**:577–579, **4**:578*F*
  permanent edges and open areas **4**:573–574, **4**:575*F*
  terraces **4**:572, **4**:572*F*
  wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*, **4**:581*F*
island biogeography
  analytical methods
    data analysis **4**:403–405
    floristic checklist methods **4**:399
    quantitative methods **4**:400
    relevé method **4**:399
    sampling methods **4**:399

vegetation (*continued*)
data analysis
dendograph approach **4**:403*F*, **4**:404*F*, **4**:405
Hawaii Volcanoes National Park **4**:402*F*, **4**:403–405, **4**:404*F*, **4**:405*F*
improved display methods **4**:404*F*, **4**:405, **4**:405*F*
two-way table approach **4**:402*T*, **4**:404*F*, **4**:405
experimental sampling design
horizontal approach **4**:406
transect approach **4**:405–406, **4**:405*F*
vertical approach **4**:406
invasive species **3**:410–411
quantitative sampling methods
challenges **4**:400
count-plot method **4**:400, **4**:401–403
distance methods **4**:400–401, **4**:401–403
point-centered quarter method **4**:401
sample plot size **4**:401*F*, **4**:403
wandering quarter method **4**:401
relevé sampling method
basic concepts **4**:399
cluster analysis **4**:403*F*, **4**:404*F*
predefinied strata **4**:399
size determinations **4**:399, **4**:400*F*
species determinations **4**:399
species quantity estimations **4**:399–400
late Pleistocene extinctions **5**:233
living dead **4:590–598**
animal populations **4**:594–595
anthropogenic impacts **4**:594–595
deforestation **4**:591–593, **4**:592*F*
extinct megafauna **4**:591–593, **4**:593*F*
food chains **4**:595–596
forest ecosystems **4**:590–591, **4**:591*F*
requirements **4**:593–594
restoration programs **4**:597
small plants **4**:594
wildland patches **4**:596–597, **4**:596*F*
Madagascar **1**:55–56
megaherbivores **5**:227–228, **5**:227*F*, **5**:233
Near East ecosystems
animal diversity **5**:469–470, **5**:469*F*
Ceratonia siliqua open forests **5**:480*F*, **5**:482–483
deserts **5**:480*F*, **5**:485–486
desert savannoid vegetation **5**:480*F*, **5**:486
dryland vegetation patterns **5**:484
flora **5**:480–481, **5**:481*T*
general discussion **5**:481–482
mangroves **5**:480*F*, **5**:487
maquis/forests **5**:480*F*, **5**:481–482
Mediterranean savannoid vegetation **5**:480*F*, **5**:483
Mt. Hermon montane forest **5**:480*F*, **5**:482
Mt. Hermon tragacanth vegetation **5**:480*F*, **5**:483–484
Pistacia lentiscus open forests **5**:480*F*, **5**:482–483
Quercus calliprinos woodlands **5**:480*F*, **5**:482
Quercus ithaburensis open forests **5**:480*F*, **5**:482
reed thickets **5**:480*F*, **5**:486
sand vegetation **5**:480*F*, **5**:486
semisteppe bathas **5**:480*F*, **5**:483
shrub–steppes **5**:480*F*, **5**:484–485
shrub–steppes with trees **5**:480*F*, **5**:485
Sudanian tree oases **5**:480*F*, **5**:486
swamps **5**:480*F*, **5**:486
synanthropic vegetation **5**:480*F*, **5**:486–487
wet salinas **5**:480*F*, **5**:486
Ziziphus lotus **5**:480*F*, **5**:483
nitrogen deposition effects **5**:523–525, **5**:527*F*
pest control
crop diversity **1**:375–376
natural enemies hypothesis **1**:375–376
noncrop plants **1**:376, **1**:376*F*
remote sensing applications **6**:396–397, **6**:396*F*
riparian zones **6**:467*T*
seasonally dry tropical forests **5**:488–490
seasonal rainfall **2**:464–466
species distribution modeling (SDM) **6**:699–700
Succulent Karoo **1**:485, **1**:486*T*, **1**:488–489
ultraviolet radiation effects **7**:298–299
vegetation succession **3**:90*T*
vegetative propagation **2**:660
Velcro **1**:590
Velkovrhia enigmatica **7**:58
Velocitermes **4**:409*T*
velvet beans **1**:589, **7**:26
velvet stem mushrooms **3**:637*T*
velvet worm **3**:224, **3**:225, **3**:226*F*
Vendian period **3**:684*F*, **3**:685*F*, **5**:157*F*, **5**:161*F*
Venezuela
biodiversity richness **1**:501*F*
botanical gardens **6**:519
Canaima National Park **5**:444
coral reef ecosystems **1**:500*T*
deforestation **1**:502*T*, **2**:449*T*, **2**:451*T*
ecosystem services **7**:356*T*
endemism **1**:501*F*, **3**:232*T*, **4**:271*T*
forest product exports **2**:451*T*
institutional collections and capacity **1**:505*T*
land cover classifications **6**:607*T*
neotropical seasonally dry forests **5**:493*T*, **5**:495*F*, **5**:497*F*
plant species **4**:102*T*
species richness **1**:499*T*, **1**:501*F*, **4**:271*T*
tropical forests **6**:600–604*T*, **6**:606*F*, **6**:606*T*, **7**:269–270, **7**:270*T*
tropical rain forests **6**:314–315, **6**:317*T*
vascular epiphytes **3**:520*T*
water funds **7**:356*T*
Venice system **3**:298–299
vents *see* hydrothermal vents
Venus flytrap **2**:643, **6**:44–45
Veratrum album **5**:712*T*
Verbanaceae **4**:352–353
Verbascum thapsus **3**:208–209
Verhulst, Pierre François **1**:696, **6**:177
vermiform animals **4**:381–382, **4**:382*F*
Vermileonidae **3**:489–491*T*
Vermivora bachmanii **5**:317–318, **7**:182
Vernadsky, V. I. **3**:239–240
vernal pools **7**:369*T*
Vernonia baldwinii **1**:707
Verrucositermes **4**:409*T*
vertebrates **7:333–341**
alpine ecosystems **1**:154–155
Antarctica **1**:186*T*
Australia **1**:300–302, **1**:302*F*
Cenozoic mass extinctions **5**:176
characteristics **7**:337–338
chordates
classification systems **7**:335, **7**:335*F*, **7**:336*F*, **7**:337*F*
global diversity **4**:385*T*
historical perspective **7**:334–335
known and estimated species **4**:384*T*
major groups
Amniota **7**:339
Cephalochordata **3**:457, **7**:336
Choanata **7**:338
Chordata **7**:336
Craniata **7**:336–337
definitions and diagnoses **7**:335–336
Gnathostomata **7**:338
Mammalia **4:681–707**, **7**:341
Osteichthyes **6**:522–523, **7**:338
Reptilia **7**:339–340
Sarcopterygii **7**:338
Sauria **7**:340–341
Tetrapoda **7**:337*F*, **7**:338–339
Vertebrata **7**:337–338
classification systems **7**:335, **7**:335*F*, **7**:336*F*, **7**:337*F*
end-Cretaceous mass extinction **5**:173–175
estimated extinction rates **4**:658*F*, **4**:659*T*
forest canopies **3**:513–514
forest management impact
behavioral changes **1**:453
biotic changes **1**:452–453
birds **1**:453
domestic animals **1**:453
food resource loss or gain **1**:453
habitat loss or gain **1**:452–453
primates **1**:453
freshwater ecosystems **3**:571–572
fruit and seed dispersal **6**:50
genetic diversity **3**:665*F*, **3**:668*F*
historical perspective **7**:334–335
indicator species **7**:206–207*F*
introduced species
Australia **1**:76*T*, **1**:80
Brazil **1**:76*T*, **1**:83
India **1**:76*T*, **1**:82
South Africa **1**:76*T*, **1**:81
United Kingdom **1**:76*T*, **1**:79
United States **1**:76–77
Mediterranean-type climate ecosystems
Australia **5**:213*T*, **5**:221
California **5**:213*T*, **5**:215

Cape Region (South Africa) **5**:213*T*, **5**:220
Chile **5**:213*T*, **5**:217
Mediterranean Basin **5**:213*T*, **5**:218–219
modern extinctions **5**:368*T*
molecular level diversity **2**:635*T*
Near East ecosystems **5**:460*T*
parasitic nematodes **7**:431–432
people–conservation relationship **2**:224
pesticide toxicity **5**:699*T*
physiological and structural characteristics **7**:333–334, **7**:334*T*
pollination **6**:49*T*, **6**:141*T*, **6**:142–143
prevalence **7**:333
return-on-investment (ROI) analysis **4**:187*T*
riparian zones **6**:463
seagrass ecosystems **6**:547–548
slash-and-burn systems **6**:556
social systems **2**:646
species richness maps **6**:709*F*
Succulent Karoo **1**:487*T*
temperature tolerance ranges **4**:115
threatened species **7**:201*T*, **7**:206–207*F*
true bug–vertebrate interactions **7**:289
wetland ecosystems **7**:389–391
see also fish
vertically transmitted partners **2**:323
Verticillium lecanii **5**:694
vertigo **3**:635*T*
Vertisols **6**:123*T*
vesicomyid clams **7**:329
vesicular arbuscular mycorrhizae (VAM) **2**:391, **3**:636
Vespa orientalis **5**:463*T*
vesper mouse **1**:365–368
Vespertilionidae **6**:309*T*
Vespidae **5**:460*T*, **5**:463*T*, **5**:464*T*
Vespoidea **4**:178
Vestiaria coccinea **1**:29*F*
Vestimentifera **4**:391, **7**:328–329, **7**:408
veterinary chemicals **1**:710–711, **2**:544
Vexillata **7**:432, **7**:432*T*
Vexillata armandae **7**:423*F*
Vexillata convoluta **7**:432*T*
Vexillata vexillata **7**:432*T*
Vibrio **2**:614–615, **6**:32
Vibrio cholerae **4**:503
Vibrio diazotrophicus **2**:611*F*, **2**:612*T*
Vibrio fischeri **2**:323–324, **5**:121–123, **5**:122*F*, **5**:123–124, **5**:281*F*
Vibrio marinus **6**:300
Vibrio psychroerythrus **6**:300
Vibrio shiloi **5**:258
Vibrio splendidus **4**:674
Viburnum **1**:271
vicariance biogeography **7:342–351**
basic concepts **2**:348, **7**:343, **7**:343*F*, **7**:346–348
component analysis **7**:348–349, **7**:349, **7**:349*F*, **7**:350*F*, **7**:350*T*
dispersal–vicariance analysis (DIVA) **2**:556, **2**:556*F*
Heterandria
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
historical biogeography **7**:342–343
Hooker's hypothesis **7**:343
Mesoamerican poeciliid fish **7**:348, **7**:348*F*
phylogenetic analyses **6**:22
plate tectonics **7**:343–344
research summary **7**:351
Southern Hemisphere ecosystems **6**:616
theoretical perspectives **2**:539–540, **7**:343–344
Xiphophorus
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
viceroy butterflies **2**:141–142
Vicia faba **1**:590, **2**:383, **2**:390, **2**:662–663
Victoria amazonica **1**:164
Victoria (Lake) **3**:444, **4**:463–464, **4**:463*T*, **4**:465–466, **4**:663
Victoria's riflebird **2**:75*F*
Vicugna vicugna **2**:243–244
vicuña **2**:243–244
videophilia **7**:244
viduine finches **5**:502, **5**:502*T*
Vietnam
animal species **1**:261*T*
botanical gardens **6**:519
deforestation **4**:562–564, **4**:563*F*
ecosystems
genetic richness **1**:514*T*
mangroves **1**:275*T*, **1**:277, **1**:513*T*
threatened species **1**:278*T*
threatened wetlands **1**:276*T*
tropical moist forest **1**:277*T*
vegetation types and coverage percentage **1**:264–265*T*
endangered mammals **3**:194*T*
mammals **4**:703–704
pastureland conversion **4**:665*T*
plant species **1**:259*T*
see also Southeast Asia
Vigna calcarata **2**:663
Vigna mungo **2**:663
Vigna radiata **2**:663
Vigna unguiculata **2**:383, **2**:663
Villa Luz Cave (Mexico) **7**:51
vinblastine **1**:359, **5**:712*T*
vinchlozin **5**:695
vincristine **1**:359, **5**:711–713, **5**:712*T*
vinegar eels **7**:427
vinegar flies **3**:489–491*T*, **3**:493
vinegaroons **1**:211–212
Viola **1**:271
Viola atropurpurea **4**:106
Violaceae **3**:522*T*
Viola pygmaea **4**:103
violet click beetle **3**:226
Vipera palaestinae **5**:477, **5**:477*F*
viperids **6**:409
Virchow, Rudolf **5**:628
Vireo atricapillus **7**:184–185
Vireo gilvus **4**:174
Vireonidae **6**:308*T*
vireos **4**:174, **7**:184–185
Virgil **4**:650
virgin forest concept **4**:569
Virgin Islands (British) **6**:519
Virgin Islands (US) **4**:560*F*, **6**:519
viriobenthos **5**:132–133
virioplankton **5**:132–133, **5**:133*F*
Virola elongata **1**:163
Virola surinamensis **1**:163
virtue ethics **3**:275
viruses
aquatic viruses **6**:584–585, **6**:584*F*
biodiversity **5**:127–129, **5**:255
characteristics and properties **2**:524*T*, **6**:584–585
classification systems **5**:129, **5**:129*F*
food webs **6**:584–585, **6**:584*F*
genomic characteristics **5**:128, **5**:129–132, **5**:129*F*, **5**:130*F*, **5**:131*F*
global species richness **3**:710
habitat–niche relationship **4**:45–46
historical research **5**:127
Holocene postglacial spread **2**:551–552*T*
known and estimated species **4**:384*T*
marine viruses **5:127–144**
biodiversity considerations **5**:127–129
classification systems **5**:129, **5**:129*F*, **5**:134*T*
dsDNA viruses
Adenoviridae **5**:134*T*, **5**:135
Alloherpesviridae **5**:134*T*, **5**:135
Caudovirales **5**:130*F*, **5**:134–135, **5**:134*T*
characteristics **5**:134–135, **5**:134*T*
Corticoviridae **5**:134*T*, **5**:135
Herpesvirales **5**:130*F*, **5**:134*T*, **5**:135
Herpesviridae **5**:134*T*, **5**:135
Iridoviridae **5**:134*T*, **5**:135
Malacoherpesviridae **5**:134*T*, **5**:135
Mimiviridae **5**:130*F*, **5**:134*T*, **5**:135
Myoviridae **5**:130*F*, **5**:134–135, **5**:134*T*
Nimaviridae **5**:130*F*, **5**:134*T*, **5**:135–136
Phycodnaviridae **5**:134*T*, **5**:136
Podoviridae **5**:130*F*, **5**:134*T*, **5**:135
Poxviridae **5**:134*T*, **5**:136
Siphoviridae **5**:130*F*, **5**:134*T*, **5**:135
unassigned viruses **5**:134*T*, **5**:136
dsRNA viruses
Birnaviridae **5**:134*T*, **5**:137
characteristics **5**:134*T*, **5**:137
Reoviridae **5**:134*T*, **5**:137
genomic characteristics **5**:128, **5**:129–132, **5**:129*F*, **5**:130*F*, **5**:131*F*, **5**:134*T*
historical research **5**:127
negative-sense ssRNA viruses
Bunyaviridae **5**:130*F*, **5**:134*T*, **5**:139–140
characteristics **5**:134*T*, **5**:139
Mononegavirales **5**:134*T*, **5**:139
Orthomyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Paramyxoviridae **5**:130*F*, **5**:134*T*, **5**:139
Rhabdoviridae **5**:130*F*, **5**:134*T*, **5**:139

viruses (*continued*)
phenotypic and genotypic diversity
DNA-containing viruses **5**:142
genome size **5**:140–141
host range **5**:140
metagenomics **5**:141–142
morphology **5**:140
RNA-containing viruses **5**:142
single-gene surveys **5**:141, **5**:141*F*
taxonomic classifications **5**:140
phenotypic diversity **5**:129–132, **5**:130*F*, **5**:131*F*, **5**:134*T*
positive-sense ssRNA viruses
Alvernaviridae **5**:134*T*, **5**:138
Astroviridae **5**:134*T*, **5**:138
Bacillarnavirus **5**:134*T*, **5**:138
Caliciviridae **5**:134*T*, **5**:138
characteristics **5**:134*T*, **5**:137
Dicistroviridae **5**:134*T*, **5**:137
Labyrnaviridae **5**:134*T*, **5**:138
Marnaviridae **5**:134*T*, **5**:137–138
Metaviridae **5**:134*T*, **5**:140
Nidovirales **5**:134*T*, **5**:137
Nodaviridae **5**:134*T*, **5**:138–139
Picornavirales **5**:134*T*, **5**:137
Roniviridae **5**:134*T*, **5**:137
Togaviridae **5**:134*T*, **5**:139
replication methods **5**:127–129, **5**:128*F*
research summary **5**:142
seawater concentrations **5**:132–133, **5**:132*F*, **5**:133*F*
ssDNA viruses
Anelloviridae **5**:134*T*, **5**:136
characteristics **5**:134*T*, **5**:136
Circoviridae **5**:134*T*, **5**:136
Inoviridae **5**:130*F*, **5**:134*T*, **5**:136
Microviridae **5**:134*T*, **5**:136
unassigned viruses **5**:134*T*, **5**:136–137
taxonomic diversity
dsDNA viruses **5**:134–135, **5**:134*T*
dsRNA viruses **5**:134*T*, **5**:137
general discussion **5**:133–135
negative-sense ssRNA viruses **5**:134*T*, **5**:139
positive-sense ssRNA viruses **5**:134*T*, **5**:137
ssDNA viruses **5**:134*T*, **5**:136
Marseillevirus **6**:585, **6**:585*F*
natural history museum research programs **5**:415
nomenclature code **5**:552
nonseed gene banks **3**:652, **3**:652*T*
number of species **5**:254*T*, **6**:586*T*
oceanic food web **5**:49*F*
parasitoids **5**:678
phenotypic diversity **5**:129–132, **5**:130*F*, **5**:131*F*
replication methods **5**:127–129, **5**:128*F*
RNA editing **5**:573–574*T*, **5**:578
zoonotic disease **1**:519–520
Visayan hornbills **7**:185
Viscaceae **4**:71, **6**:315*T*
viscachas **3**:188
Viscum album **4**:77, **5**:673
Viscum cruciatum **4**:77
visible light **7**:297*F*
Vitaceae **3**:129*T*
vitamins **6**:113*T*
Vitellaria paradoxa **1**:50
viticulture **2**:300
Vitis **5**:622*F*
Vitis vinifera **1**:463, **2**:662–663
Viverridae
body mass **1**:687*F*
food categories **1**:690*F*, **1**:691*F*
home range size **1**:687*F*
species diversity **1**:686
viverrids **4**:693, **4**:701*T*
Vjetrenica (Bosnia and Herzegovina) **7**:58, **7**:59*T*, **7**:60
vleis **7**:369*T*
Voandzeia subterranea **2**:663
Vochysiaceae **6**:316*T*
Vochysia haenkiana **1**:161
Vogt, William **2**:285, **3**:99
volatile organic compounds (VOCs) **1**:281–282, **1**:679–680, **1**:680*T*, **2**:680
volcanism
dispersal biogeography **2**:547, **2**:548–549*F*
mass extinctions **3**:418
plant biodiversity–climate relationship **2**:115–117, **3**:118
tropical forest regeneration **7**:277–278
tropical soils **7**:271
voles
alpine ecosystems **1**:154–155
biotic relationships **1**:710
boreal forests **1**:631–632
food webs **7**:266*F*, **7**:267*T*
Volterra–Gause perspective **2**:595–596, **4**:40–42
Volterra, Vito **2**:178–180
Volvariella volvacea **3**:637*T*
Volvox **6**:30
Volvox aureus **6**:33*F*
Vombatus hacketti **5**:364*F*
Vombat usursinus **4**:363–364
vomiting **3**:635*T*
von Baer, Karl-Ernst **3**:393–394, **5**:629–630
von Ihering method **5**:670
Vorticella **6**:295
vulcanism **1**:554
Vulpes **1**:709
Vulpes lagopus **1**:236
Vulpes velox **3**:192
Vulpes vulpes **1**:80, **1**:631–632, **5**:475
Vulpes zerda **4**:693
vultures
evolutionary convergence **1**:617
oil-palm plantations **5**:603
veterinary chemicals **2**:544

W

Wächtershäuser, Günther **5**:631
wadis **2**:461
wad/wadden **7**:369*T*
Wagner optimization **2**:40, **2**:41*F*
Wahlund effect **6**:190–191
Wake Island rail **1**:621
Wallabaia indra **5**:364*F*
wallabies **4**:687*T*, **4**:688–689, **6**:333
Wallacea **6**:82*T*
Wallace, Alfred Russell
biogeographic regions **1**:578, **1**:579*F*
dispersal theories **2**:539–540
evolutionary theories **2**:421, **2**:424, **4**:651–652, **5**:628
Southern Hemisphere ecosystems **6**:612–613
vicariance biogeography **7**:343
Wallace, David Foster **4**:654
walleye pollock **5**:58–60
walleye trout **1**:10*T*, **1**:12*F*
Wallis & Futuna Islands
deforestation **2**:449*T*
island ecosystems **5**:592*F*
walnuts
native crops **2**:662–663
seed protection **7**:26
walruses
adipose tissue **7**:34–35
characteristics **5**:67
extinctions **5**:90
predation effects **6**:236
species diversity **4**:693
taxonomy **5**:65*T*
Walsingham Caves (Bermuda) **7**:58, **7**:59*T*
warble flies **3**:489–491*T*
warbler finches **1**:27–28, **1**:28*F*
warblers
beak shape **1**:604*F*
guilds **4**:33–35, **4**:34*F*
habitat–niche relationship **4**:44, **4**:45*F*
interspecific competition **2**:188–189, **2**:188*F*
warbling vireos **4**:174
warfare effects **3**:192
wart biter katydids **3**:225, **3**:225*F*
warthogs **4**:696
waru-warus **1**:123–124
Washington
nonvascular epiphytes **3**:521*T*
urbanization–biodiversity impacts **4**:564
Wasmannia auropunctata **5**:598*F*
wasps
characteristics **4**:392
Near East ecosystems **5**:460*T*, **5**:463*T*, **5**:464*T*
number of species **4**:392*T*
pollination **6**:141*T*, **6**:142
waste disposal **1**:512, **1**:512*F*
Wasur Biru **1**:276*T*
water
anthropogenic impacts
acidification and pollution **3**:572–574, **3**:577–578
agricultural and industrial chemicals **3**:575–576, **3**:580–581, **3**:581*F*
areal reductions **3**:573–574
eutrophication **3**:575*F*, **3**:576–577, **3**:576*F*
hydrological cycle changes **3**:574–575
lake eutrophication **3**:581–582

organic pollution **3**:575, **3**:575*F*, **3**:580–581, **3**:581*F*
pollution **1**:512
restoration and conservation programs **3**:574
drinking water **1**:95–97, **5**:351
ecosystem services modeling **5**:351
environmental impacts **3**:280–281
geothermally heated systems **4**:112–113
hydrological cycle **1**:553–554, **1**:555*F*
intercropping systems **2**:390–391
pollination **6**:141*T*
pollution
anthropogenic impacts
freshwater ecosystems **3**:572–574, **3**:577–578
water quality decline **4**:158
soil erosion **6**:591
South America **6**:607*T*
Southeast Asia **1**:512
water bears **4**:371*T*, **4**:381*T*
water beetles **4**:371–373, **4**:371*T*, **7**:389–391
waterbirds **1**:606
water boatmen **4**:371–373, **4**:374*F*, **7**:293
waterbucks **4**:696
water buffalo **1**:80, **1**:81, **4**:354, **4**:364–365, **6**:333
waterbugs **4**:371–373
Water Conservation Orders (New Zealand) **6**:496–497
water fleas
characteristics **6**:26
impacts and distribution **4**:378
number of species **3**:182*T*, **4**:371*T*
water funds **7:352–366**
basic concepts **7**:352–353
Brazil **7**:363–364
Colombia **7**:362–363
design and implementation guidelines
administrative management **7**:360
evaluation and monitoring strategies **7**:361
feasibility studies **7**:360
governance **7**:360–361
implementation **7**:361
service assessments **7**:360
Ecuador **7**:361–362
financial sources and management **7**:357*F*, **7**:359
governance **7**:359–360
historical perspective **7**:355–357
Latin America and the Caribbean region **7**:354–355*F*, **7**:356*T*
operational and financial model **7**:357–358, **7**:357*F*
water fund platform **7**:364–365
watershed management
best management practices **7**:358–359, **7**:358*T*
compensating impacts **7**:359
general discussion **7**:357–358
public protected area management **7**:358, **7**:358*T*
water hyacinth
Asian ecosystems **1**:279
eutrophication **3**:359–360
freshwater ecosystems **3**:359–360, **4**:378
globalization effects **3**:703, **3**:704*F*
impacts and distribution **4**:359–360, **4**:378, **6**:94
lake and pond ecosystems
Amazon ecosystems **1**:164
introduced species **4**:465
watermelon **2**:662–663
watermilfoil **4**:364–365
water mites **4**:371*T*, **4**:374, **4**:374*F*
water mold **2**:644, **3**:630, **3**:696
water opossum **4**:701*T*
Water Protection Directive (WPD) **1**:95–97
water quality–nutrient management relationship **1:95–110**
agricultural production evolution **1**:97–98, **1**:98*F*
concentrated animal feeding operations (CAFOs) **1**:99
farm input–nutrient balance **1**:99, **1**:99*T*
freshwater invertebrates **4**:377
land-use change impacts **1**:107–108
leaching vulnerability index **1**:102*F*
management criteria **1**:97, **1**:97*T*
nitrogen (N) **1**:96*F*, **1**:98
nutrient input and sources **1**:96*F*, **3**:575*F*, **3**:576–577, **3**:576*F*
phosphorus (P) **1**:96*F*, **1**:98–99
regional nutrient imbalances **1**:98*F*, **1**:99–101, **1**:100*F*, **1**:102*F*
remedial benefical management practices (BMPs)
characteristics **1**:103*T*
cover crops **1**:103*T*, **1**:106
critical source area management **1**:103*T*, **1**:107
erosion potential **1**:103*T*, **1**:105–106, **1**:105*T*, **1**:106*F*
farm input decisions **1**:101, **1**:103*T*
fertilizer application methods **1**:103*T*, **1**:104
fertilizer application rate and frequency **1**:103*T*, **1**:104–105
implementation strategies **1**:101, **1**:103*F*
nutrient-use efficiency improvements **1**:101–104, **1**:103*T*
riparian/buffer zones **1**:103*T*, **1**:106–107
runoff potential **1**:103*T*, **1**:105
soil and plant testing **1**:103*T*, **1**:104
source management **1**:103*T*, **1**:104
transport management **1**:103*T*, **1**:105
research background **1**:95–97
research summary **1**:108
water scorpions **7**:293
watershed management
ecosystem services **3**:100
mercury contamination **1**:12
nitrogen cycle impacts **5**:540, **5**:541*F*
storm peak management **5**:352–353
traditional conservation practices **7**:233
water funds
best management practices **7**:358–359, **7**:358*T*
compensating impacts **7**:359
general discussion **7**:357–358
public protected area management **7**:358, **7**:358*T*
watershed characteristics **6**:469–470, **6**:470*F*
water shrews **7**:390
water snake **5**:477, **7**:390
water spiders **4**:374*F*
water storage **3**:560–562, **3**:561*T*
water striders **4**:371–373, **4**:374*F*, **7**:293
water tupelo **7**:397, **7**:397*F*
water vapor **1**:281, **4**:18–19
water voles **4**:701*T*
waterways **6**:609–610
wave attenuation **5**:452–453, **5**:453*T*, **5**:454*F*
Wawa, Ontario **6**:155
waxbills **1**:83
wax palms **4**:516
wealth versus poverty **6**:220
weasels
food webs **7**:266*F*, **7**:267*T*
Great American Biotic Interchange (GABI) **2**:584*F*
number of species **4**:690*T*
skull size comparisons **1**:692*F*
species diversity **1**:685, **4**:693
Weatershiemia **5**:618*F*
weather events **2**:84–85
weatherfish **3**:464
weathering processes **1**:6*F*, **1**:678–679
Weber, Max **6**:383
Web Ontology Language (OWL) **1**:444
webspinners **4**:392*T*
Weddell seal **5**:90
Wedelia trilobata **6**:96–97
wedge-type millipedes **5**:429–430
weeds
benefits **1**:120
characteristics **4**:87–88, **4**:87*T*
economic impact **1**:76*T*
edible plants **3**:131–132
herbicide use
community composition and structure changes **4**:91–92
herbicide resistance **4**:92–93, **4**:92*T*
increased insect activity **4**:93
selection pressure **4**:91–92
historic land-use patterns **4**:570–572
impacts and distribution **4**:87–88
pest control **1**:376
root parasitic weeds **4**:76–77
slash-and-burn systems **6**:556, **6**:557
United States **1**:75–76, **1**:76*T*
weed management
biodiversity effects **4**:91–92
chemical weed control **4**:88–89
general discussion **4**:88
integrated weed management (IWM) **4**:88
tools and methodologies **4**:88
wildland weed management **4**:93–94
weed science **4**:88
weeper capuchin **6**:254
weeverfish **3**:467–471
weevils **1**:327*T*, **1**:328, **3**:512–513, **4**:361, **4**:363, **4**:376

Wegener, Alfred L. **1**:583
Weiner, Alan **5**:634–635
Weinmannia **3**:522*T*, **7**:119
Weinmannia balbisiana **1**:167
Weinmannia racemosa **6**:692*F*
Weismann, August **2**:424, **6**:1–2
Weissmann, Charles **5**:633
weka **5**:370, **7**:185
Weldon, W. F. R. **6**:2
Welwitschia **5**:618*F*, **6**:59–60
Welwitschia mirabilis **1**:53, **2**:468
welwitschia plant **7**:223
West Africa
  deforestation **2**:448*T*
  nitrogen deposition **5**:526*T*
  oil-palm plantations **5**:600–601
  protected area systems **6**:491*T*
  regional carbon stocks **1**:678*F*
  regional cultivars **3**:32*F*
West African manatee **5**:81*T*
West Anatolian temperate forest **1**:262–263*T*
West Asia
  deforestation **2**:448*T*
  regional carbon stocks **1**:678*F*
  regional cultivars **3**:32*F*
West Coast Aquatic (WCA) Management Board **3**:100–101
Western Atlantic region **3**:473–474
Western Australia Mulga shrublands **2**:468
western Bornean orangutan **6**:256–258*T*
western chimpanzee **6**:256–258*T*
western corn rootworm **4**:365
western cross-marked surili **6**:256–258*T*
Western Europe
  deforestation **2**:448*T*
  vegetation surveys **3**:341–344*T*
Western Ghats **5**:526*T*, **6**:82*T*
western grey squirrel **4**:564
western hoolock gibbon **6**:256–258*T*
western lowland gorilla **6**:256–258*T*
Western North America **5**:561*F*, **5**:563
western North Pacific gray whales **5**:77*T*, **5**:78–79
Western North Pacific harbor seal **5**:81*T*
western Phayre's langur **6**:256–258*T*
western pond turtle **4**:564
Western purple-faced langur **6**:256–258*T*
Western Sahara **6**:507*T*
western toad **3**:165
western woolly lemur **6**:256–258*T*
West Indian manatee **5**:81*T*, **5**:83, **5**:225*T*
West Nile virus **1**:362–365, **1**:364*F*, **1**:365*F*, **3**:349, **3**:411, **4**:504, **5**:316–317, **7**:237–238
West Virginia
  birds **6**:307–309, **6**:308*T*
  butterflies **6**:306–307, **6**:307*T*
  mammals **6**:309*T*
West Yunnan black crested gibbon **6**:256–258*T*
wetas **3**:222, **3**:731, **3**:731*F*
wetlands **7**:**367–383**, **7**:**384–402**
  animal species richness **7**:389–391, **7**:389*F*, **7**:390*F*, **7**:393–394, **7**:396*F*
  anthropogenic impacts **7**:371*T*
  Asia **1**:275–279, **1**:275*T*, **1**:276*T*
  Australian plant communities **1**:292–297
  biodiversity restoration **6**:456–457
  biogeographical variation **7**:387–388
  C_4 plants **2**:371*T*
  classifications **7**:385–386
  climate change effects **6**:150–151
  conservation status **3**:171–173*T*
  continental and regional diversity **7**:387–388
  creation and restoration
    appropriate hydrologic conditions **7**:378
    Bois-des-Bel Peatland, Quebec, Canada (case study 5) **7**:375
    coastal wetlands **7**:375–378
    conservation programs
      croplands **7**:371
      forested wetland conservation **7**:371
    definitions and terminology **7**:368–371
    Delaware Bay salt marsh restoration (case study 6) **7**:376–378, **7**:377*F*
    designer wetlands **7**:379
    ecological engineering principles **7**:381–382
    Florida Everglades (case study 1) **7**:372, **7**:372*F*, **7**:373*F*
    Gulf of Mexico hypoxia (case study 2) **7**:372–373, **7**:374*F*
    habitat loss mitigation **7**:370–371
    hydrologic and water quality restoration **7**:371
    Louisiana delta restoration (case study 7) **7**:378
    Mesopotamian marshlands (case study 4) **7**:374
    natural successional processes **7**:379
    peatlands **7**:374–375
    planted versus nonplanted restoration (case study 8) **7**:379, **7**:380*F*
    research background **7**:368–371
    river delta restoration **7**:378
    self-designed wetlands **7**:379
    Skjern River Restoration Project, Denmark (case study 3) **7**:373–374, **7**:375*F*
    soil properties **7**:378–379
    success measures **7**:379–381
  dam projects **1**:512
  definitions and terminology **7**:367–368, **7**:369*T*
  dynamic global vegetation models (DGVMs) **2**:678–679
  ecosystem services **1**:334*T*
  freshwater invertebrates **4**:**369–378**
    background information **4**:369–370
    ecological functions and processes
      characteristics **4**:375
      decomposition and recycling **4**:375
      material transfer and uptake **4**:375–376
      primary production **4**:376
    global patterns **4**:373
    influencing factors
      disturbances **4**:375
      endemism **4**:374–375
      habitat heterogeneity **4**:375
      isolation factors **4**:374–375
      origins **4**:374
      research background **4**:373–374
      species interactions **4**:375
    life cycles **4**:370, **4**:372*F*
    major groups **4**:370, **4**:371*T*
    major habitats **4**:369–370, **4**:369*F*
    major threats
      altered water regimes **4**:377–378
      biodiversity loss **4**:376–377
      exotic species invasions **4**:378
      habitat depletion and loss **4**:376–377
      pollution **4**:377
      reduced water quality **4**:377
    number of species **4**:370–371, **4**:371*T*
    river continuum concept (RCC) **4**:373*F*
    taxonomic groups **4**:371–373
    zonal representatives **4**:374*F*
  functional diversity (FD)
    acidic environments **7**:395–397, **7**:396*F*
    adaptations **7**:394–395
    anoxic environments **7**:394–395
    carnivorous plants **7**:395–397, **7**:396*F*
    drought conditions **7**:395
    nutrient availability **7**:395–397
    reproductive adaptations **7**:397
    salinity/salt tolerance **7**:395
  functional role **7**:367, **7**:370*F*
  geographic distribution **4**:370*T*
  global extent **7**:368, **7**:370*T*
  global warming effects **4**:229
  habitat destruction **3**:404, **7**:398
  habitat loss and degradation **7**:371*T*
  human impacts **4**:154–155
  Lobau wetlands (Austria) **7**:59
  major regional wetland ecosystems **7**:386–387
  major threats **7**:368, **7**:371*T*
    anthropogenic impacts **7**:399
    global losses **7**:397–398
    habitat destruction **3**:404, **7**:398
    hydrological alterations **7**:398–399
    invasive species **7**:399
    landscape fragmentation **7**:399
    natural disturbances **7**:399–400
  nature-based coastal protection strategies **5**:452–453, **5**:453*T*, **5**:454*F*
  Near East ecosystems
    general discussion **5**:472–473
    riverine habitats **5**:473
    swamps **5**:472–473, **5**:474*F*, **5**:475*T*
    temporary winter rain pools **5**:473, **5**:476*F*
  occurrences and characteristics **7**:384–386, **7**:385*F*
  plant species richness **7**:388–389, **7**:388*F*, **7**:391–393, **7**:398*F*
  protected area systems **6**:494–498, **6**:496*F*, **6**:497*F*
  restoration and maintenance programs **7**:400–401
  schematic diagram **7**:370*F*
  South America **3**:116–117, **6**:600–604*T*, **6**:607*T*

species diversity controls
animal species richness **7**:393–394, **7**:396*F*
influencing factors **7**:391–393, **7**:392*F*
plant species richness **7**:391–393
summary discussion **7**:401
threatened birds **7**:181
threatened wetlands **1**:276*T*
tropical Africa **1**:53–54
Wetlands International **2**:231–232
Wetlands Reserve Program **7**:371
wet meadows **7**:369*T*
wet prairies **7**:369*T*
Wettinia cladospadyx **1**:166–167
wet-zone Toque macaque **6**:256–258*T*
whalefish **3**:467–471
whales
adipose tissue **7**:34
Census of Marine Life project **2**:12*F*
convergent evolution **4**:701*T*
deoxyribonucleic acid (DNA) evolution **2**:56*F*
government legislation and regulations **3**:454, **3**:713*T*, **3**:716
migratory patterns **5**:313–314
number of species **4**:690*T*
overharvesting/overcollecting **3**:283
shipworms **5**:121
species diversity **4**:693–694
subspecies level extinctions **5**:76
taxonomy **5**:65*T*
threat percentage **3**:195
tourism industry **7**:223
whaling industry **3**:454
worm–whale nutrient conversion **5**:121, **5**:121*F*
Whalleyanoidea **5**:387–388*T*
wheat
crop production **1**:86
crop rotation systems **1**:116*T*
genetic diversity
commercialization effects
general discussion **1**:91–92
intensified environments **1**:92
non-intensified environments **1**:92–93
genetic evolution **1**:86–87
genetic narrowing **1**:89–90
genetic trends **1**:90
landrace–modern varieties comparisons **1**:89
plant breeding **1**:462–463
remaining gene pool percentage **1**:132*T*
research summary **1**:93
intercropping systems **2**:383, **2**:384*F*
native crops **2**:662–663
Neolithic agriculture **2**:658
plant conservation **6**:85–86
regional cultivars **3**:32*F*
root–soil interface **6**:113*T*
seed industries **1**:91
spatial distribution **1**:87–89, **1**:88*T*
wheat stem sawfly **4**:182–183
whelks **3**:125, **4**:334–335, **6**:232
whip scorpions **1**:211–212, **1**:211*F*
whipsnakes **6**:408–409
whip spiders **1**:211
whiptail lizards **6**:406
whirling disease **2**:528*T*
whisk ferns **6**:59
white abalone **3**:201*T*
white adipose tissue **7**:26–27, **7**:27*F*, **7**:33–35
white ash **1**:610
whitebark pine forests **2**:571, **2**:571
white-bellied spider monkey **6**:256–258*T*
white bur-sage **2**:468–469
white button pizza mushrooms **3**:625*T*, **3**:637–638, **3**:637*T*
white-collared lemur **6**:256–258*T*
white-eared pocket mouse **7**:229*T*
White, E. B. **4**:653
white-eyelid mangabeys **6**:251*T*
white-faced capuchin monkeys **5**:603
white fir **6**:152–154
whitefish **3**:123*T*, **3**:464–465, **6**:522–523
white flies **4**:298–299, **4**:299*F*
white-footed mouse **1**:362, **1**:365–368
white-footed tamarin **6**:256–258*T*
white-fronted geese **1**:710
White, Gilbert **4**:651–652, **6**:111
white goby **4**:465
white-headed duck **7**:184
white-headed langur **6**:256–258*T*
white-headed woodpecker **4**:67*F*
white-lipped peccary **2**:590*T*
white lupine **2**:390
White, Lynn, Jr. **6**:381–383
white-naped mangabey **6**:256–258*T*
white nose syndrome **2**:527–529, **2**:529*F*, **7**:52–53
white-nose syndrome **4**:156
white oak **5**:390
white perch **1**:12*F*
white pine **2**:279
white pine blister rust **3**:633*T*
white pox disease **2**:528*T*, **2**:529*F*
white rhinos **3**:198, **4**:695, **5**:225*T*, **5**:226
see also megaherbivores
white rot **4**:429
white-rumped sandpiper **5**:312
white sand oligotrophic formations **1**:160–161
white spot syndrome virus **5**:135–136
white spruce ecosystems **1**:632*T*, **1**:633*F*, **3**:535, **6**:120*T*, **6**:123–124, **6**:466
white sucker **1**:12*F*
white-tailed deer **2**:279, **4**:494
white-tailed eagles **5**:474
white-throated bee-eaters **5**:603
white-throated packrats **2**:472
white-toothed shrew **1**:601
white whales **4**:693–694
white-whiskered spider monkey **6**:256–258*T*
whiting **3**:466–467, **6**:422*T*
Whittaker's species turnover **2**:443–444
whooper swans **3**:365–366, **7**:36
whooping cranes **7**:189
Wielandiella **5**:616
Wilberforce, Samuel **2**:421
wild ass **1**:644*F*, **6**:334, **6**:573, **6**:576*F*, **6**:578–579
Wild Bird Conservation Act (1992) **3**:713*T*, **3**:716
wild boar **1**:640, **4**:359, **5**:603–604
wild botanical gardens **6**:505
wild carrot **1**:462*F*
wild dogs **5**:449, **5**:603
wildebeest **2**:533, **3**:191, **4**:362, **4**:490, **4**:491*F*, **4**:696, **5**:313, **6**:332
wild emmer wheat **3**:670, **3**:671*F*, **3**:672*F*
Wildenow, Karl **1**:577–578
Wilderness Act (1964) **2**:286
wilderness areas **2**:284
wildfires
basic concepts **2**:564–565, **2**:565*F*
boreal forests **1**:627–628
deforestation **2**:454
ecological impacts **2**:562–563
El Niño Southern Oscillation (ENSO) **3**:156–158, **3**:158*F*
global livestock **6**:342–343
Southeast Asia **1**:517
wildflowers **6**:458–459, **7**:223
wild horses **6**:333
wild ibex **1**:640
wildland weed management **4**:93–94
Wildlife Conservation Society **2**:231–232
wildlife corridors *see* landscape corridors
Wildlife Habitat Council **7**:20–21
wildlife management **7:403–407**
biofuels/renewable energy industry **1:540–551**
cropping system comparisons **1**:543–547, **1**:544*F*, **1**:545*F*, **1**:548*F*
government policies and regulations **1**:540–541
land allocations
global-scale assessments **1**:547–549
landscape-scale assessments **1**:547–549, **1**:548*F*
local-scale assessments **1**:549
regional-scale assessments **1**:547–549
land-use change impacts
croplands and pasturelands **1**:542–543, **1**:542*F*
marginal, degraded, and abandoned lands **1**:543, **1**:544*F*
native ecosystems **1**:541–542, **1**:545*F*
prime agricultural lands **1**:543
oil-palm plantations **1**:540–541
summary discussion **1**:549
conservation biology **2**:239–241
damage control **7**:406
ecosystem/landscape conservation **7**:405–406
endangered species **7**:405
fire/fire regimes **3**:438
goals **7**:403
harvesting considerations **7**:403–405
historical conservation practices **2**:282, **2**:283
human–wildlife conflicts **7**:406
indicator species **7**:405–406
international organizations **4**:325–326
keystone species **7**:406

wildlife management (*continued*)
management strategies
applied techniques
behavioral techniques **7**:406–407
biological controls **7**:407
direct population control **7**:407
habitat manipulation **7**:406–407
immunization **7**:407
mechanical techniques **7**:406
characteristics **7**:406–407
population assessments **7**:407
maximum sustainable yield (MSY) **7**:403–405, **7**:404*F*
private conservation areas **2**:297*T*
protected area systems **2**:239–241
tourism industry **7**:223
translocations **7:236–240**
background information **7**:236
ethical concerns **7**:239
managed relocations **7**:238
motivations **7**:237–238
research summary **7**:239–240
risks and costs **7**:238–239, **7**:239*F*
success assessments **7**:238
traditional conservation practices **7**:236–237
umbrella species **3**:197–198, **7**:406
wild parsnip **2**:139, **2**:143
wild pigs
mangrove ecosystems **5**:14
oil-palm plantations **1**:517
wild radish **2**:720
Wild Rivers Act (Australia) **6**:497
wild species
climate change effects **2:79–99**
climate models
dynamical downscaling techniques **2**:89–90
general circulation models (GCMs) **2**:88
regional predictions **2**:88–89
statistical downscaling techniques **2**:89
validation methods **2**:88
current trends **2**:83–84
ecological models
animal responses **2**:92–94, **2**:93*F*
bottom-up approaches **2**:66–68, **2**:90
combined approaches **2**:68, **2**:90–91
community ecology **2**:94
extinction probability **2**:94–95
future climate projections **2**:90
refined approaches **2**:91
strategic cyclical scaling approach **2**:68–69, **2**:91
three-way linkages **2**:94
top-down approaches **2**:66, **2**:90
vegetation responses **2**:91–92
ecological responses
abundance shifts **2**:85
distribution shifts **2**:85
extinction rates **2**:86–87, **4**:664
fish **3**:449–450
phenological shifts **2**:85–86
physical changes **2**:86
extreme events **2**:84–85
future research areas
applications-driven climatic research **2**:96
ecological impact assessments **2**:95
ecosystem adaptability **2**:95–96
multiscale interactive analysis **2**:96
historical perspective **2**:82–83
importance **2**:80–81
joint atribution theory **2**:87
meta-analyses **2**:87
synergistic effects **2**:81–82
importance **2**:79–80
major threats
exotic species invasions **2**:80
habitat change **2**:80
nutrient loading **2**:80
overharvesting/overexploitation **2**:80
pollution **2**:80
seed banks **3**:644
wild sunflowers **6**:626–627
wildwood vegetation **4**:569, **4**:572–574
wild yaks **3**:191, **3**:373
Wilkesia **1**:30–32, **1**:30*F*
Willamette Basin **6**:265–266, **6**:266*F*, **6**:268*F*
Williamonia **5**:618*F*
Williamsoniella **5**:618*F*
Williamsoniella **5**:616
willows
biofuel crop production **1**:529*T*, **1**:531–532
boreal forests **1**:628–630, **1**:630*T*
energy use impacts **3**:258
geographical variation **1**:613–614
medicinal properties **1**:359
willow tits **1**:611–612
Wilson, Edward O. **1**:583–585, **2**:539–540, **3**:5, **4**:123, **6**:563–564
Wilson's petrel **1**:601–602
wilts **3**:633*T*
wind
soil erosion **6**:591
wind farms **5**:466
Windermere (Lake) **4**:461*T*
Windhausenia **4**:727*T*
window flies **3**:489–491*T*
wind pollination **6**:141*T*
wind power **3**:253*T*, **3**:262, **3**:265*T*
wind spiders **1**:213, **1**:213*F*
windstorms **2**:562–563
wine cap mushrooms **3**:637*T*
wine making **2**:300, **3**:638
winged beans **3**:32*F*
Winged Migration (2001) **4**:654
wingless desert grasshoppers **3**:735–736
Winteraceae **3**:522*T*, **5**:618*F*
Winterales **5**:621*F*
winter crane flies **3**:489–491*T*
winter mushrooms **3**:637*T*
winter wheat **1**:462–463
Wisconsin **3**:521*T*
wisent **3**:196
witchweeds **4**:76–77, **4**:362, **5**:673
within-habitat diversity **6**:534
within-phenotype niche width (WPNW) **2**:140
Wiwaxia **5**:382
Wobern Abbey experimental garden **4**:585*F*
Woese, Carl **5**:631–632
Wolbachia **4**:179, **5**:395–396, **5**:678–679
wolffish **3**:467–471
wolftrap anglerfish **3**:475*F*
Wollemia **2**:47
Wolong Nature Reserve **3**:374
wolverines
adipose tissue **7**:27, **7**:34*F*
boreal forests **1**:631–632
wolves
conservation practices **7**:229*T*
dispersal biogeography **2**:546–547
ecosystem engineers **4**:201
ecosystem/landscape conservation **7**:405–406
genetic diversity **2**:268, **2**:269*F*
habitat selection **4**:64
late Quaternary extinction event (LQEE) **4**:724–725*T*
Near East ecosystems **5**:472, **5**:475
predation effects **6**:236–237, **6**:238*F*, **7**:259–260, **7**:260*F*
prey–predator interactions **1**:709
reintroduction programs **3**:196
social behavior **6**:573
species diversity **1**:685, **4**:693
tourism industry **7**:223
trophic cascades **7**:259, **7**:260*F*
urbanization threats **4**:564
wombats **4**:363–364, **4**:687*T*, **4**:689
Wombling analysis **4**:511
WOMBSOFT (software program) **4**:511
Wonambi **4**:727–728
Wonambi naracoortensis **5**:364*F*
wood
carbon dioxide (CO_2) emissions **3**:265*T*
energy use impacts **3**:262–263, **3**:263*T*, **3**:264*T*
land requirements **3**:265*T*
world energy reserves **3**:265*T*
world energy use **3**:251*F*, **3**:252*T*, **3**:253*T*
wood ash **6**:555
woodchucks **7**:36
wood ear mushrooms **3**:637*T*
wood frog **6**:153*F*
wood gnats **3**:489–491*T*
woodlands
African savannas **1**:66*F*, **1**:67–68
Australia **1**:292–297, **1**:294*T*
biodiversity change scenarios **4**:6*F*
conservation status **3**:171–173*T*
current global land cover **4**:557*F*
ecosystem services **1**:334*T*
global agriculture impacts **3**:428*T*
global biodiversity trends **4**:563*F*
habitat–bird species diversity relationship **1**:613
historic land-use patterns **4**:572–574, **4**:573*F*, **4**:574*F*
land-use trends **4**:559*F*
nitrogen deposition **5**:526*T*
pre-agriculture land cover **4**:557*F*, **4**:558*F*, **4**:566*F*

precipitation–temperature relationship **1**:59*F*
protected area systems **5**:444*T*, **6**:492*T*, **6**:499*T*
Serengeti ecosystem **2**:487–488, **2**:488*F*
South America **6**:607*T*
tropical Africa **1**:49–51
woodlice **3**:226
wood mouse **1**:367
wood-pasture ecosystems **4**:574–577, **4**:575*F*, **4**:576*F*, **4**:581*F*
woodpecker finches **1**:27–28, **1**:28*F*, **4**:33, **4**:34*F*
woodpeckers
beak shape **1**:604*F*
guilds **4**:33–35, **4**:34*F*
woodpigeons **2**:541*F*
wood turtles **7**:390
wood wasps **4**:177–178
woolly adelgids **4**:361
woolly lemurs **6**:251*T*
woolly mammoths **5**:231–233
woolly monkeys **6**:251*T*
wooly spider monkey **4**:692
workflow systems **1**:445
World Association of Zoos and Aquariums (WAZA) **1**:662–663, **4**:325–326
world-at-night map **4**:561*F*
World Bank **2**:231–232, **4**:324–325, **7**:18–19, **7**:207–208
World Climate Research Program **4**:326–327
World Conference on Education for All **3**:136
World Conservation Monitoring Centre (WCMC) **4**:327
World Conservation Union **3**:209, **4**:664–666
World Database on Protected Areas (WDPAs) **6**:489
World Education **2**:231–232
world energy use **3**:251, **3**:251*F*, **3**:252*T*, **3**:253*T*
World Flora online database **6**:509
World Health Organization (WHO) **5**:670–672
World Heritage Convention **4**:325, **6**:489
World Heritage Sites program **4**:325, **6**:489, **6**:506, **7**:18–19
World Network of Biosphere Reserves (WNBR) **4**:325–326
World Resources Institute (WRI) **4**:324–325, **4**:664–666
World Trade Organization (WTO) **7**:18–19
worldwide commerce and travel **2**:532
World Wide Web Consortium (W3C) **1**:444
World Wide Web (WWW) **1**:444
World Wildlife Fund (WWF) **2**:231–232, **3**:139, **4**:271*T*, **4**:325–326
Wormaldia kisoensis **7**:345, **7**:346*F*
Wormaldia mohri **7**:346*F*
worm-lizards **3**:216*T*
worms
classifications **4**:381*T*
mangrove ecosystems **4**:337
river ecosystems **6**:478*T*
worm–whale nutrient conversion **5**:121, **5**:121*F*
see also Platyhelminthes
woundfin **4**:364
Wrangel Island **4**:723*F*
wrasse **3**:467–471
Wrightea tomentosa **1**:70
Wright, Sewall **3**:675, **4**:245–246, **6**:2
Wright's F-statistical method **4**:510
Wular Lake **1**:276*T*
Wulff, Evgenii **2**:539–540
Wuliang black crested gibbon **6**:256–258*T*
Wyeomyia smithii **5**:724–725
wyerone **1**:590
Wyoming Basins Ecoregion (WBE) study **2**:503–505, **2**:504–505*F*
Wyoming toad **3**:696
Wytham Wood (England) **7**:266–268, **7**:266*F*, **7**:267*T*

X

Xanthoma sagittifolium **4**:289*T*
Xanthopan morganii preadicta **5**:398
Xanthopan morgani praedicta **2**:138
Xanthophyceae **3**:571*T*, **6**:478*T*
xanthorhodopsins **6**:281*F*, **6**:283
Xanthorhoë fluctuata **6**:159*F*
Xanthorrhoea caespitosa **1**:304*F*
Xanthosoma sagittifolium **6**:553
Xanthotis macleayana **2**:75*F*
xanthotoxin **5**:712*T*
Xenarthra
characteristics **4**:689
evolutionary history **2**:583*F*
late Quaternary extinction event (LQEE) **4**:727*T*
mammalian evolution **4**:685
number of species **4**:690*T*
tropical rain forests **6**:309*T*
Xenasteiidae **3**:489–491*T*
Xenicus longipes **1**:621
Xenicus lyalli **1**:621
Xenicus lyallii **5**:366–367
xenoestrogens (XEs) **3**:118–119
Xenophyophora **5**:48, **6**:292
Xenopus laevis **3**:695–696
Xenoturbellida **7**:434*F*
Xenungulata **2**:583*F*
xeric formations **3**:111*F*, **3**:115–116, **6**:600–604*T*
xeric shrublands **5**:526*T*, **6**:492*T*, **6**:495*T*, **6**:499*T*
xerophytes **2**:471
Xeropicta vestalis **5**:466–467, **5**:469–470
Xerus inauris **2**:473–474, **2**:474*F*
Xi Jiang Delta **1**:276*T*
Xingu River, Brazil **1**:159*T*, **1**:161*F*
Xiphias gladius **2**:8–13, **2**:12*F*, **3**:410
Xiphinema **7**:422–424, **7**:430
Xiphocaris elongata **3**:247
Xiphophorus
cladistic analysis **7**:348, **7**:348*F*, **7**:349–351, **7**:350*F*
endemic analysis **7**:349–351, **7**:350*F*
geographical distribution **7**:348*F*
Xiphosura
arthropod radiation ecology **2**:400
number of species **5**:669–670*T*
phylogenetic relationships **1**:203*F*
Xiphydrioidea **4**:178
X-ray crystallography **5**:708
X-rays **7**:297*F*
Xyeloidea **4**:178
Xylia kerrii **1**:70
Xylocarpus **1**:277, **5**:11*T*
Xylomyidae **3**:489–491*T*
Xylophagidae **3**:489–491*T*
Xylophagomorpha **3**:489–491*T*
xylose **6**:113*T*

Y

yaks **1**:640, **1**:644*F*, **3**:191
yams
intercropping patterns **2**:383
native crops
Southeast Asia **2**:663
Sub-Saharan Africa **2**:663
regional cultivars **3**:32*F*
slash-and-burn systems **6**:553
Yancheng Biosphere Reserve **3**:374
Yancheng Marshes **1**:276*T*
yang **6**:565
Yangtze River **7**:150
Yangtze River dolphin **5**:75
Yangtze River finless porpoise **5**:81*T*
Yanomono, Peru **1**:159*T*
yeast infections **3**:634, **3**:634*T*
yeasts
bread baking **3**:638
growth processes **3**:626
interspecific competition **2**:182–184, **2**:183*T*
nonseed gene banks **3**:652*T*
wine and beer making **3**:638
yellow-bellied marmots **5**:722
yellow-billed hornbills **5**:603
yellow-billed magpie **5**:215
yellow birch **1**:610
yellow-breasted capuchin **6**:256–258*T*
yellow-cheeked crested gibbon **6**:256–258*T*
yellow clover aphid **4**:362
yellowfin tuna **6**:422*T*
yellow–green algae **3**:571–572, **3**:571*T*
yellowhead **7**:184
yellow-legged frogs **3**:695–696
yellow perch **1**:10*T*, **1**:12*F*, **4**:517, **6**:153*F*
yellow pitcher plants **7**:396*F*
yellow rail **7**:182
Yellow River **7**:150–151
yellow rocket **1**:75
Yellow Sea **3**:171–173*T*
Yellowstone National Park Act (1872) **2**:289–292, **5**:442, **6**:77–78, **6**:486
Yellowstone National Park ecosystem **4**:113, **5**:449
Yellowstone to Yukon initiative **2**:258
yellowtail **1**:190*T*

yellow-tailed woolly monkey **6**:256–258*T*
yellow wagtails **1**:710
yellow water lily **6**:115–116
Yemen
  animal species **1**:261*T*
  botanical gardens **6**:507*T*
  ecosystems
    mangroves **1**:275*T*
    threatened species **1**:278*T*
    vegetation types and coverage percentage **1**:264–265*T*
  plant species **1**:259*T*
Yersinia pestis **2**:551–552*T*, **2**:636*F*
yeti crab **2**:20*F*
yew **1**:612
yin **6**:565
Yoda's law **6**:161
Yolo Bypass **7**:155, **7**:156*F*, **7**:157*F*
Yoruba culture **6**:384
Yosemite National Park (United States) **3**:167–168
Younger Dryas glacial period **2**:69–70
yphlogammarus mrazeki **7**:58
Ypomeutoidea **5**:386*F*
Yponomeutoidea **5**:387–388*T*
ysanu **2**:663–664
Yucatán black howler monkey **6**:256–258*T*
Yucatán spider monkey **6**:256–258*T*
yucca **2**:468–469
Yucca brevifolia **2**:468–469
Yucca elata **1**:707–708, **1**:709–710
yucca moth
  domesticated ruminants **1**:709–710
  mutualism **2**:146
Yugoslavia (vegetation surveys) **3**:341–344*T*
Yunnanozoon lividum **7**:334–335
Yunnan snub-nosed monkey **6**:256–258*T*
Yunnan white-handed gibbon **6**:250

Z

Zacryptocerus **1**:329
Zaglossus **4**:686
Zaglossus attenboroughi **4**:686
Zaglossus bartoni **4**:686
Zaglossus bruijni **4**:686
Zaglossus hacketti **5**:363–364, **5**:364*F*
Zaglossus ramsayi **5**:364*F*
Zaire
  butterflies **1**:658, **1**:659*T*
  indigenous peoples **4**:270*T*
  woodland savannas **1**:67–68
Zaleya **2**:376*F*
Zalophus californianus **3**:192–193
Zalophus japonicus **5**:75–76
Zambia
  botanical gardens **6**:519
  deforestation **1**:677*F*, **2**:449*T*, **2**:451*T*
  energy efficiency measures **3**:254*F*
  forest product exports **2**:451*T*
  woodland savannas **1**:67–68
Zamia **4**:281–285, **4**:282*F*
Zamiaceae **5**:618*F*
Zannichellia palustris **5**:484–485
Zantedeschia aethiopica **5**:220
Zanthoxylum **3**:522*T*, **5**:489
Zanthoxylum bungeanum **2**:383
Zanzibar dwarf galago **6**:256–258*T*
Zapata rail **7**:184
Zapodidae **6**:309*T*
Zaprochilinae **3**:727
Zea mays **1**:543–547, **1**:544*F*, **2**:364*F*, **2**:636*F*, **2**:663–664, **4**:281–285, **4**:283*F*, **4**:289*T*, **6**:553
zebra mussels
  environmental ethics **3**:270, **3**:270*F*
  eutrophication impacts **3**:369
  government legislation and regulations **3**:715
  impacts and distribution **3**:181, **3**:284, **3**:411, **4**:143, **4**:364–365, **4**:378, **6**:93–95
  lake and pond ecosystems **3**:181
  life cycle **4**:370
  physical characteristics **5**:374, **5**:381
  river ecosystems **6**:482
zebras **1**:487–488, **4**:695, **5**:235, **5**:313
Zenaida graysoni **7**:188
Zen Buddhism **6**:383–384
zero-inflated Poisson distribution **6**:160
Zeugloptera **5**:385, **5**:386*F*, **5**:387–388*T*
Zeus faber **3**:468
Zhangjiajie Nature Reserve **3**:375
Zheiran gazelle **1**:269, **6**:334
Zimbabwe
  botanical gardens **6**:519
  cattle ranching **1**:395, **1**:396*F*
  deforestation **2**:449*T*, **2**:451*T*
  ecotourism **1**:395
  forest product exports **2**:451*T*
  invasive species **4**:347
  megaherbivores **5**:227–228, **5**:227*F*, **5**:234–236
  wildlife management **1**:395, **1**:396*F*
  woodland savannas **1**:67–68
zinc nutrition **2**:390
zineb **5**:695
Zingiberaceae **3**:129*T*, **6**:313–318, **6**:316*T*
Zingiberidae **1**:263
Zingiber officinale **2**:663
Zizeeria karsandra **5**:463*T*
Ziziphus **1**:69
Ziziphus jujuba **2**:383
Ziziphus lotus **5**:480*F*, **5**:483
Ziziphus mucronata **1**:67
Ziziphus spinachristi **5**:480*F*, **5**:483
Zoarces viviparus **5**:34–35
Zoarcoidei **3**:469
Zoellneriallum andinum **4**:106
Zonocerus elegans **3**:735*F*
zoobenthos **6**:478–479, **6**:478*T*
Zoological Information Management System (ZIMS) **1**:662–663, **7**:473
Zoological Museum (Russia) **5**:405
zoological organisms **5**:552–553, **5**:553*T*
zooplankton
  Census of Marine Life project **2**:18–19, **2**:19*F*
  characteristics **5**:686–688, **5**:687*F*, **6**:24–25
  classifications and sampling methods **4**:382–383, **4**:383*T*
  climate change effects **5**:37
  community structure **6**:25–26
  El Niño Southern Oscillation (ENSO) **5**:689*F*
  energy flow **3**:246–247
  environmental requirements **6**:31
  eutrophication impacts **3**:361–363, **3**:363*F*
  freshwater ecosystems **3**:565–566
  harmful algal blooms **3**:356–358, **3**:449, **3**:575*F*, **3**:576–577, **3**:576*F*
  lake and pond ecosystems **3**:243*F*, **4**:374*F*, **4**:459
  marine viruses **5**:133*F*
  mercury contamination **1**:12*F*
  pelagic ecosystems **4**:374*F*, **5**:683–686
  regional distribution **5**:688*F*
  river ecosystems **6**:472*F*, **6**:478–479, **6**:478*T*
  seasonal cycle **5**:688*F*
  spatial distribution **5**:687*F*
  surface water acidification **1**:10*T*
  trophic pyramid **4**:670–671
  ultraviolet radiation effects **7**:300
  water-column nitrogen:phosphorus (N:P) ratios **3**:350*F*
zooprophylaxis **1**:360–361
zoos/zoological parks **7:470–479**
  animal welfare **7**:472–473
  ark paradigm **7**:477–478
  biodiversity conservation
    advantages/disadvantages **4**:319
    ark paradigm **7**:477–478
    breeding population research **7**:474
    conservation products **4**:321*T*
    frozen zoos **7**:477
    institutional planning **7**:475
    management strategies **7**:474–475
    population management strategies **7**:475–477
    post-ark paradigm **7**:478
    record keeping plans **7**:475
    research summary **7**:478
  biodiversity-rich countries **1**:504, **1**:505*T*
  conservation biology **2**:245, **4**:316*F*, **4**:319, **4**:321*T*, **7**:471–472
  definition **7**:470
  emerging trends **7**:473–474
  gene banks **3**:653
  historical perspective **7**:470–471
  inbreeding depression **4**:246, **4**:246*F*
  learning opportunities **3**:137–138
  post-ark paradigm **7**:478
  primate trade **6**:254–255
  public support **7**:471–472
  zoo husbandry **7**:472–473
  see also natural history museums and institutions
Zootermopsis **4**:409*T*
zooxanthellae
  coral reef ecosystems **2**:335
  marine ecosystems **5**:45–46
Zoraptera **4**:308–312, **4**:309*T*, **4**:392*T*
zorapterans **4**:392*T*
Zostera **6**:542*T*
Zostera asiatica **6**:541*T*
Zostera capensis **6**:541*T*

Zostera capricorni **6**:541*T*
Zostera caulescens **6**:541*T*
Zostera japonica **6**:541*T*
Zostera marina **1**:593, **3**:360–361, **3**:360*F*, **3**:361*F*, **5**:689–690, **6**:454, **6**:541*T*
Zostera mucronata **6**:541*T*
Zostera mulleri **6**:541*T*
Zostera noltii **6**:541*T*
Zostera novazelandica **6**:541*T*
Zosteroideae **6**:542*T*
Zygaenidae **5**:396
Zygaenoidea **5**:386*F*, **5**:387–388*T*
Zygnematophyceae **3**:571*T*
Zygodon pichinchensis **4**:103
Zygodontomys brevicauda **1**:367–368
Zygolophodon **5**:224*T*
Zygomaturus **4**:727–728
Zygomaturus trilobus **5**:363–364, **5**:364*F*
zygomycosis **3**:634*T*
Zygomycota **3**:627, **3**:628*F*
Zygopachylus albomarginis **1**:210
Zygophyceae **6**:478*T*
Zygophyllaceae **2**:374*T*
Zygophyllales **2**:374*T*
zygophylls **2**:469
Zygophyllum **2**:376*F*
Zygophyllum dumosum **1**:268, **5**:485–486
Zygophyllum simplex **2**:364*F*
Zygopini **1**:328
Zygosaccharomyces rouxii **3**:638–639